

Current Protocols in Molecular Biology

Frederick M. Ausubel, Roger Brent, Robert E. Kingston, David D. Moore, J.G. Seidman, John A. Smith,
Kevin Struhl (eds.)
Current Protocols in Molecular Biology
Copyright © 2003 John Wiley & Sons, Inc. All rights reserved.

John Wiley & Sons Inc; ringbou edition (December 4, 2003)

ISBN: 047150338X

Foreword

Phillip A. Sharp¹

The breadth of knowledge required for modern research in molecular biology is truly staggering. A single series of experiments can encompass genetic manipulation of specific bacterial strains and their phage, the appropriate use of several highly specific enzymes, and the culturing and genetic engineering of mammalian cells. For individuals who have followed the evolution of research in molecular biology, it was surprising to watch the mammalian λ immunoglobulin gene being cloned in the prokaryotic λ phage. Neither immunologists nor bacterial geneticists expected the ultimate union when the phage and gene were named. Now a single individual must master both fields to be successful. Research that depends upon such breadth of knowledge needs protocol books that explain why certain methods are used and provide references for further reading. This manual fulfills that need.

All scientists realize that they figuratively stand on the shoulders of their predecessors. Nothing more concretely reflects this dependence than a laboratory protocol book. It is a monument to the hours of labor that students, technicians and other professionals have expended in developing methods that work every time. Good methods have another characteristic: they have been honed to the minimum of time and effort required to achieve the desired results. Most investigators collect personal protocol books of “tried and true” recipes. This new manual with its continuously growing and evolving set of protocols is an excellent basis from which to begin such a collection.

There are two unique features of this manual that are in tune with the rapid advances in molecular biology. First, the manual evolves continuously with the addition of new protocols in emerging areas of research. These new protocols are added as quarterly supplements to the core manual. Second, the manual evolves through a “network” of the users. Readers are encouraged to contribute corrections and improvements on techniques in each supplement. Thus, the book is designed to grow and change as biology grows and changes.

Molecular biology has always been an intensely experimental science. Young people, more commonly, have the energy and intensity to push the field forward. The excellent group of young scientists who have contributed to this book are representatives of this tradition.

¹Cambridge, Massachusetts

Preface

Frederick M. Ausubel, Roger Brent, Robert E. Kingston, David D. Moore, J. G. Seidman, John A. Smith, and Kevin Struhl

Although mastery of the techniques in this manual will enable the reader to pursue research in molecular genetics, the manual is not intended to be a substitute for graduate-level courses in molecular biology or a comprehensive textbook in the field. Introductory texts that we recommend include: *Molecular Biology of the Gene* (4th ed.), by J.D. Watson, N.H. Hopkins, J.W. Roberts, J.A. Steitz, and A.M. Weiner; *Molecular Genetics: An Introductory Narrative*, by G.S. Stent and R. Calendar; *From Genes to Clones*, by E.-L. Winnacker; and *Genetics and Molecular Biology*, by R. Schleif. In addition, *An Introduction to Genetic Analysis*, by D.T. Suzuki, A.J.F. Griffiths, J.H. Miller, and R.C. Lewontin, is a good place to learn classical genetics, and *The Molecular Biology of the Cell*, by B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J.D. Watson is a compendium of useful information about many aspects of cellular and molecular biology. Finally, we strongly recommend that readers gain first-hand experience in basic techniques and safety procedures by working in a molecular biology laboratory.

An inevitable hazard of manual writing is that protocols become obsolete or that new techniques are developed—usually just as the pages are going to press. This problem is particularly severe in a fast-moving field such as molecular genetics. To safeguard our manual from inexorable obsolescence, and as a means of correcting errors in a timely manner, we provide a quarterly updating service. The looseleaf format facilitates the replacement of pages that contain errors or have become obsolete and will accommodate the addition of new protocols to the appropriate section of the manual. The publisher can provide further details about subscribing to the quarterly supplements.

How to Use This Manual

Format and Organization

This publication is available in looseleaf, CD-ROM and online formats. The material covered in each version is identical. For looseleaf purchasers, four binders are provided to accommodate the growth of the manual via the quarterly update service. The looseleaf format of the binders allows easy insertion of new pages, units, and chapters that are added. The index and table of contents are updated with each supplement.

CD-ROM purchasers receive a completely new disc containing the entire manual every quarter. The CD-ROM User's Guide describes in detail the many features designed for accessing information presented in the CD. Topics of interest can be located primarily by using either the search index (by typing in the desired word or string of words) or by scanning through the expandable table of contents and selecting the desired units. In addition, hyperlinks jump between related units when cross-references are selected and from solutions and reagents listed in the materials list to their corresponding recipes.

Subjects in this manual are organized by chapters and sections, and protocols are contained in units. Units generally describe a method and include one or more protocols with listings of materials, the protocol steps and annotations, recipes for unique reagents and solutions, and commentaries on the “hows” and “whys” of the method. Overview units contain theoretical discussions that lay the foundation for subsequent protocols. Page numbering in the looseleaf version reflects this modular arrangement; for example, page 15.1.3 refers to Chapter 15 (PCR), Unit 1 (standard procedures and optimization), page 3 of that particular unit.

Many reagents and procedures are employed repeatedly throughout the manual. Rather than duplicate this information, cross-references among units are used extensively. Early chapters (and appendices) describe these commonly used techniques (basic microbiology and basic manipulations of enzymes, DNA, and RNA), while later chapters describe their application (constructing libraries, DNA sequencing, mutagenesis, transfection, and protein analysis). Thus, whenever a particular enzyme is used in a protocol, the appropriate unit in Chapter 3—describing reaction conditions for that enzyme—is cross-referenced (e.g., unit 3.7 for reverse transcriptase). Similarly, throughout the book readers are referred to unit 1.3 for spreading or streaking a plate, to unit 2.1A for phenol extraction/alcohol precipitation, to unit 2.5A for agarose gel electrophoresis, and so on. By turning to these units, the reader will find instructions for the techniques, recipes for relevant reagents, and commentary. As a result, protocols in the later chapters of the book—which can be lengthy and complex—are not overburdened with steps describing auxiliary procedures required to prepare, purify, and analyze the sample or molecule of interest.

Introductory and Explanatory Information

Because this publication is first and foremost a compilation of molecular biology techniques we have not provided extensive instructive material. We have, however, included explanatory information where required to help readers gain an intuitive grasp of the procedures. First, the chapter, section, and unit introductions describe how the protocols that follow connect to each other. Second, annotations to the actual protocol steps describe what is happening. Third, the commentary sections that close each unit give background information and review the relevant literature so that readers are aware of the historical and theoretical development of a technique, as well as alternative approaches. Fourth, key references are sometimes included to inform users of particularly useful background reading or original descriptions or applications of the technique. Finally, in some cases special “overview” units have been included early in the chapter.

Protocols

Many units in the manual contain groups of protocols. The basic protocol is presented first in each unit and is generally the recommended approach. Alternate protocols are provided where different equipment or reagents can be employed to achieve similar ends, where the starting material requires a variation in approach, or where requirements for the end product differ from those in the basic protocol. Support protocols describe additional steps that are required to perform the basic or alternate protocols; these steps are separated from the core protocol because they might be applicable to other uses in the manual, or because they are performed in a time frame separately from the basic protocol steps.

Reagents and Solutions

Reagents required for a protocol are listed in the “materials” list before the procedure begins. Many are common stock solutions, others are commonly used buffers or media, while others are solutions unique to a particular protocol. Recipes for the latter solutions are provided in each unit, following the protocols (and before the commentary) under the heading Reagents and Solutions. It is important to note that the names of some of these special solutions might be similar from unit to unit (e.g., hybridization solution, high-salt solution, etc.) while the recipes differ; thus, make certain that reagents are prepared from the proper recipes. On the other hand, recipes for commonly used buffers and media are provided once in an appendix or in early units in the manual. These universal recipes are cross-referenced parenthetically in the materials lists rather than repeated with every usage.

NOTE: Deionized, distilled water should be used in all protocols in this manual, and in the preparation of all reagents and solutions.

Commercial Suppliers

In some instances throughout the manual, we have recommended commercial suppliers of chemicals, biological materials, kits, or equipment. This has been avoided wherever possible, because preference for a specific brand is subjective and is generally not based on extensive comparison testing. Our guidelines for recommending a supplier are that (1) the particular brand has actually been found to be of superior quality, or (2) the item is difficult to find in the marketplace. An appendix lists the full names and addresses of all recommended suppliers, but these are by no means the only vendors of biological supplies. Readers may experiment with substituting their own favorite brands.

Safety Considerations

Anyone carrying out these protocols will encounter the following hazardous or potentially hazardous materials: (1) radioactive substances, (2) toxic chemicals and carcinogenic or teratogenic reagents, (3) pathogenic and infectious biological agents, and (4) recombinant DNA. Most governments regulate the use of these materials; it is essential that they be used in strict accordance with local and national regulations. Cautionary notes are included in many instances throughout the manual, but we emphasize that users must proceed with the prudence and precaution associated with good laboratory practice, and that all materials be used in strict accordance with local and national regulations.

Reader Response

Most of the protocols included in this manual are used routinely in our own laboratories. These protocols work for us; to make them work for you we have annotated critical steps and included critical parameters and troubleshooting guides in the commentaries to most units. However, the successful evolution of this manual depends upon readers' observations and suggestions. Consequently, a self-mailing, reader-response survey is included at the back of the manual (and is included with each supplement); we encourage readers to send in their comments. Electronic contact information is available at <http://www.currentprotocols.com>.

Acknowledgments

Continuing to keep this manual up to date would be impossible without massive assistance from the Current Protocols staff at John Wiley & Sons. Among those who help us, we are extremely grateful to Joe Ingram, Virginia Chanda, Scott Holmes, Tom Cannon, Tom Downey, Susan Lieberman, Allen Ranz, and Joseph White. We are particularly indebted to Sarah Greene and her staff at Greene Publishing Associates who initially conceived of this project and helped to shape it with skill and unbounded patience.

We are especially grateful to our co-workers who have helped with the manual by contributing material to it, by commenting on the chapters, or by field-testing the procedures. To those people—in our own labs, and in academic and industrial labs all over the world—we offer our deepest thanks.

Editors & Contributors

EDITORIAL BOARD

Frederick M. Ausubel
Massachusetts General Hospital & Harvard
Medical School

Roger Brent
The Molecular Sciences Institute, Berkeley,
California

Robert E. Kingston
Massachusetts General Hospital & Harvard
Medical School

David D. Moore
Baylor College of Medicine

J.G. Seidman
Harvard Medical School

John A. Smith
University of Alabama at Birmingham

Kevin Struhl
Harvard Medical School

GUEST EDITORS

Lisa M. Albright
Austin, Texas

Donald M. Coen
Harvard Medical School

Carolyn L. Smith
Baylor College of Medicine

Ajit Varki
University of California San Diego

Mary C. Wildermuth
University of California Berkeley

CONTRIBUTORS

Susan M. Abmayr
Pennsylvania State University
University Park, Pennsylvania

Lonnie D. Adams
The Upjohn Company
Kalamazoo, Michigan

Lisa M. Albright
Austin, Texas

R. Donald Allison
University of Florida
Gainesville, Florida

Roland Annan
SmithKline Beecham
King of Prussia, Pennsylvania

Anthony T. Annunziato
Boston College
Chestnut Hill, Massachusetts

Oscar M. Aparicio
University of Southern California
Los Angeles, California

Carmen A. Argmann
Institut de Génétique et de Biologie
Moléculaire et Cellulaire
Illkirch, France

Alejandro Aruffo
Bristol-Myers Squibb
Princeton, New Jersey

Frederick M. Ausubel
Massachusetts General Hospital and
Harvard Medical School
Boston, Massachusetts

Johan Auwerx
Institut de Génétique et de Biologie
Moléculaire et Cellulaire and Institut
Clinique de la Souris
Illkirch, France

Omar Bagasra
Thomas Jefferson University
Philadelphia, Pennsylvania

Jim Bahre
Fuji Medical Systems, USA
Stamford, Connecticut

Albert S. Baldwin, Jr.
University of North Carolina
Chapel Hill, North Carolina

Cynthia Bamdad
Clinical Micro Sensors
Pasadena, California

David Baulcombe
John Innes Centre
Norwich, United Kingdom

Andreas D. Baxevasis
National Human Genome Research
Institute
Bethesda, Maryland

E.L. Bearer
Brown University
Providence, Rhode Island

C.R. Bebbington
Celltech
Slough, United Kingdom

Daniel M. Becker
Pennie & Edwards
Menlo Park, California

Sabine D. Bell
University of Texas
Austin, Texas

Timothy P. Bender
University of Virginia
Charlottesville, Virginia

Claude Besmond
Hôpital Robert Debré
Paris, France

Stephen M. Beverley
Harvard Medical School
Boston, Massachusetts

Michael Bittner
National Human Genome Research
Institute, NIH
Bethesda, Maryland

Seth Blackshaw
Harvard Medical School
Boston, Massachusetts

Justin Blethrow
University of California at San Francisco
San Francisco, California

Kenneth D. Bloch
Massachusetts General Hospital
Boston, Massachusetts

Mark S. Boguski
National Center for Biotechnology
Information
Bethesda, Maryland

Juan S. Bonifacino
National Institute of Child Health and
Human Development
Bethesda, Maryland

Angie L. Bookout
Howard Hughes Medical Institute
University of Texas Southwestern Medical
Center
Dallas, Texas

Ann Boyle
Current Protocols
Madison, Connecticut

Reinhard I. Boysen
Monash University
Victoria, Australia

Allan R. Brasier
University of Texas
Galveston, Texas

Michael Brenowitz
Albert Einstein College
Bronx, New York

Roger Brent
The Molecular Sciences Institute
Berkeley, California

Terry Brown
University of Manchester Institute of
Science and Technology
Manchester, United Kingdom

Bernard H. Brownstein
Howard Hughes Medical Institute and
Washington University School of Medicine
St. Louis, Missouri

Mark Brunswick
Uniformed Services University of the
Health Sciences
Bethesda, Maryland

Mikail Bubunenko
National Cancer Institute at Frederick
Frederick, Maryland

Kim Budelier
Qiagen, Inc.
Chatsworth, California

William Buikema
University of Chicago
Chicago, Illinois

Linda Buonocore
Yale University School of Medicine
New Haven, Connecticut

Stephen Buratowski
Harvard Medical School
Boston, Massachusetts

Richard R. Burgess
University of Wisconsin
Madison, Wisconsin

Dean Burgi
Genomix
Foster City, California

Michael C. Byrne
Genetics Institute
Cambridge, Massachusetts

Steven A. Carr
SmithKline Beecham
King of Prussia, Pennsylvania

Andrés E. Carrasco
University of Buenos Aires-CONICET
School of Medicine
Buenos Aires, Argentina

Miles W. Carroll
Oxford BioMedica
Oxford, United Kingdom

A. Nigel Carter
The Salk Institute
La Jolla, California

Anthony Celeste
Genetics Institute
Cambridge, Massachusetts

Constance Cepko
Harvard Medical School
Boston, Massachusetts

Bulbul Chakravarti
Keck Graduate Institute of Applied Life
Sciences
Claremont, California

Deb N. Chakravarti
Keck Graduate Institute of Applied Life
Sciences
Claremont, California

Marie-France Champy
Institut Clinique de la Souris
Illkirch, France

Lawrence Chan
Baylor College of Medicine
Houston, Texas

Yuan Chang
Hillman Cancer Center
University of Pittsburgh
Pittsburgh, Pennsylvania

David D. Chaplin
Howard Hughes Medical Institute and
Washington University School of Medicine
St. Louis, Missouri

Claudia A. Chen
National Institute of Mental Health
Bethesda, Maryland

Jin-Long Chen
Tularik, Inc.
South San Francisco, California

J. Michael Cherry
Stanford University School of Medicine
Stanford, California

Jonathan D. Chesnut
Invitrogen Corporation
Carlsbad, California

Cheng-Ming Chiang
Case Western Reserve University School of
Medicine
Cleveland, Ohio

Lewis A. Chodosh
University of Pennsylvania
Philadelphia, Pennsylvania

Piotr Chomczynski
University of Cincinnati College of
Medicine
Cincinnati, Ohio

Joanne Chory
The Salk Institute for Biological Studies
La Jolla, California

Valentina Ciccarone
Life Technologies, Inc.
Rockville, Maryland

Donald M. Coen
Harvard Medical School
Boston, Massachusetts

Nelson B. Cole
University of Pennsylvania
Philadelphia, Pennsylvania

Robert N. Cole
The Johns Hopkins University Medical
School
Baltimore, Maryland

Thomas Coleman
Fox Chase Cancer Center
Philadelphia, Pennsylvania

John E. Coligan
National Institute of Allergy and Infectious
Diseases
Bethesda, Maryland

Donald Coling
University of California, Berkeley
Berkeley, California

Martine A. Collart
Harvard Medical School
Boston, Massachusetts

James F. Collawn
The Salk Institute
La Jolla, California

David A. Conner
Harvard Medical School
Boston, Massachusetts

H. Edward Conrad
University of Illinois
Urbana, Illinois

Nina Costantino
National Cancer Institute at Frederick
Frederick, Maryland

Darryl Conte Jr.
University of Massachusetts Medical School
Worcester, Massachusetts

Helen M. Cooper
Ludwig Institute for Cancer Research
Melbourne, Australia

Norman Cooper
National Institute of Allergy and Infectious
Diseases
Bethesda, Maryland

Lynn M. Corcoran
Walter and Eliza Hall Institute
Victoria, Australia

Brendan Cormack
Harvard Medical School
Boston, Massachusetts

Nina Costantino
National Cancer Institute at Frederick
Frederick, Maryland

Michele Cottler-Fox
University of Maryland
Greenebaum Cancer Center
Baltimore, Maryland

Donald L. Court
National Cancer Institute at Frederick
Frederick, Maryland

Francis Cruz
Brigham and Women's Hospital
Boston, Massachusetts

Carolyn L. Cummins
Howard Hughes Medical Institute

University of Texas Southwestern Medical
Center
Dallas, Texas

Rita Das
Harvard Medical School
Boston, Massachusetts

Simanti Datta
National Cancer Institute at Frederick
Frederick, Maryland

Pieter J. de Jong
Children's Hospital Oakland Research
Institute
Oakland, California

Job Dekker
University of Massachusetts Medical School
Worcester, Massachusetts

Esteban C. Dell'Angelica
National Institute of Child Health and
Human Development
Bethesda, Maryland

Joseph DeRisi
University of California, San Francisco
San Francisco, California

Sandra Diaz
University of California, San Diego
La Jolla, California

Andree Dierich
Institut de Génétique et de Biologie
Moléculaire et Cellulaire and Institut
Clinique de la Souris
Illkirch, France

Michael P. DiGiovanna
Yale University School of Medicine
New Haven, Connecticut

Jack E. Dixon
University of California at San Diego
La Jolla, California

Tamara L. Doering
Cornell University Medical College
New York, New York

Robert L. Dorit
Yale University
New Haven, Connecticut

David Dorris
Harvard Medical School
Boston, Massachusetts

Josée Dostie
University of Massachusetts Medical School
Worcester, Massachusetts

Dennis Dowhan
Baylor College of Medicine
Houston, Texas

Tingting Du
University of Massachusetts Medical School
Worcester, Massachusetts

Allan Duby
Medical City
Dallas, Texas

Barbara Dunn
GenPharm International
Mountain View, California

Ben Dunn
University of Florida
Gainesville, Florida

Christopher D. Earl
Plant Resources Venture Fund
Cambridge, Massachusetts

Patricia L. Earl
National Institute of Allergy and Infectious
Diseases
Bethesda, Maryland

Richard L. Eckert
Case Western Reserve School of Medicine
Cleveland, Ohio

Diane G. Edmondson
M.D. Anderson Cancer Center
Houston, Texas

Karen L. Elbing
Clark and Elbing LLP
Boston, Massachusetts

Elaine A. Elion
Harvard Medical School
Boston, Massachusetts

Andrew Ellington
University of Texas
Austin, Texas

Orna Elroy-Stein
Tel Aviv University
Tel Aviv, Israel

Isam-Eldin Eltoum
University of Alabama at Birmingham
Birmingham, Alabama

JoAnne Engebrecht
State University of New York
Stony Brook, New York

Paul T. Englund
Johns Hopkins Medical School
Baltimore, Maryland

Douglas J. Epstein
University of Pennsylvania School of
Medicine
Philadelphia, Pennsylvania

Jeffrey D. Esko
University of Alabama
Birmingham, Alabama

Michael J. Eveleigh
ADI Diagnostics
Rexdale, Ontario

Hua-Ying Fan
Massachusetts General Hospital
Boston, Massachusetts

Rhonda Feinbaum
Massachusetts General Hospital
Boston, Massachusetts

Gregg B. Fields
Florida Atlantic University
Boca Raton, Florida

Russell L. Finley, Jr.
Wayne State University School of Medicine
Detroit, Michigan

Michael Finney
MJ Research
Watertown, Massachusetts

Achim Fischer
F. Hoffmann-La Roche AG
Basel, Switzerland

Thomas A. Fleisher
Warren Grant Magnuson Clinical Center
Bethesda, Maryland

Maximillian T. Follettie
Genetics Institute
Cambridge, Massachusetts

Susan L. Forsburg
Molecular and Cell Biology Laboratory
The Salk Institute for Biological Studies
La Jolla, California

John J. Fortin
Tropix, Inc.
Bedford, Massachusetts

Cecil H. Fox
Molecular Histology Laboratory
Gaithersburg, Maryland

Nicole Francis
Massachusetts General Hospital
Boston, Massachusetts

Paula Franco
University of Buenos Aires-CONICET

School of Medicine
Buenos Aires, Argentina

Verna Frasca
Amersham Pharmacia Biotech
Piscataway, New Jersey

Hudson H. Freeze
The Burnham Institute
La Jolla, California

Eirik Frengen
The Biotechnology Centre of Oslo
University of Oslo
Oslo, Norway

Andra R. Frost
University of Alabama at Birmingham
Birmingham, Alabama

Michiko N. Fukuda
La Jolla Cancer Research Foundation
La Jolla, California

Minoru Fukuda
La Jolla Cancer Research Foundation
La Jolla, California

Steven A. Fuller
Nabi
Rockville, Maryland

Dmitry V. Fyodorov
University of California, San Diego
La Jolla, California

Sean R. Gallagher
UVP, Inc.
Upland, California

Subinay Ganguly
SmithKline Beecham
King of Prussia, Pennsylvania

Yuan Gao
Deakin University
Victoria, Australia

Henrik Garoff
Karolinska Institute, Novum Research
Center
Huddinge, Sweden

Paul A. Garrity
University of California, Los Angeles
Los Angeles, California

Anke Geick
Ribopharma AG
Kulmbach, Germany

Joseph V. Geisberg
Harvard Medical School
Boston, Massachusetts

David H. Gelfand
Roche Molecular Systems
Alameda, California

C. Ronald Geyer
University of Florida
Gainesville, Florida

Nele Gheldof
University of Massachusetts Medical School
Worcester, Massachusetts

Chris Gilbert
Cancer Research UK
London Research Institute
Clare Hall Laboratory
Hertfordshire, U.K.

Michael Gilman
Cold Spring Harbor Laboratory
Cold Spring Harbor, New York

Warren Gish
Washington University School of Medicine
St. Louis, Missouri

Arthur Glatfelter
National Human Genome Research
Institute, NIH
Bethesda, Maryland

Jane Glazebrook
University of Minnesota
St. Paul, Minnesota

Erica A. Golemis
Fox Chase Cancer Center
Philadelphia, Pennsylvania

Chris Gooden
National Human Genome Research
Institute, NIH
Bethesda, Maryland

Gregory A. Grant
Washington University School of Medicine
St. Louis, Missouri

Michael E. Greenberg
Harvard Medical School
Boston, Massachusetts

John M. Greene
Massachusetts General Hospital
Boston, Massachusetts

David Greenstein
Massachusetts General Hospital
Boston, Massachusetts

Mitchell S. Gross
SmithKline Beecham
King of Prussia, Pennsylvania

Barbara Grossmann
Amersham Life Science
Cleveland, Ohio

Stefan Grünwald
Pharmingen
San Diego, California

James Gulick
Cincinnati Children's Hospital
University of Cincinnati
Cincinnati, Ohio

Tod Gulick
Massachusetts General Hospital
Charleston, Massachusetts

Jeno Gyuris
Mitotix, Inc.
Cambridge, Massachusetts

Philipp Hadwiger
Ribopharma AG
Kulmbach, Germany

Lars Hagel
Amersham Pharmacia Biotech
Uppsala, Sweden

John Hanson
MJ Research
Watertown, Massachusetts

Anna Haramis
EMBL
Heidelberg, Germany

Gerald W. Hart
The Johns Hopkins University Medical
School
Baltimore, Maryland

Grant Hartzog
Harvard Medical School
Boston, Massachusetts

Robert G. Hawley
American Red Cross
Rockville, Maryland and
The George Washington University
Washington, D.C.

Pamela Hawley-Nelson
Life Technologies, Inc.
Rockville, Maryland

Bradley K. Hayes
University of Alabama
Birmingham, Alabama

Jeffrey J. Hayes
University of Rochester Medical Center
Rochester, New York

Milton T. W. Hearn
Monash University
Victoria, Australia

Olaf Heidenreich
University of Tübingen
Tübingen, Germany

Joseph S. Heilig
University of Colorado
Boulder, Colorado

Peter Heinrich
Consortium für Elektrochemische Industrie
Munich, Germany

William J. Henzel
Genentech, Inc.
South San Francisco, California

David E. Hill
Applied Biotechnology
Cambridge, Massachusetts

James P. Hoeffler
Invitrogen Corporation
Carlsbad, California

Timothy Hoey
University of California, Berkeley
Berkeley, California

Charles S. Hoffman
Boston College
Chestnut Hill, Massachusetts

Diane Hollenbaugh
Bristol-Myers Squibb
Seattle, Washington

Shelley B. Hoover
Molecular Histology Laboratory
Gaithersburg, Maryland

Peter Hornbeck
University of Maryland
Baltimore, Maryland

Hayden Huang
Brigham and Women's Hospital
Boston, Massachusetts

Helen M. Huls
Baylor College of Medicine
Houston, Texas

Tony Hunter
The Salk Institute for Biological Studies
La Jolla, California

John G.R. Hurrell
FluorRx
Carmel, Indiana

Charles B-C. Hwang
Harvard Medical School
Boston, Massachusetts

Panayiotis A. Ioannou
The Murdoch Institute for Research into
Birth Defects
Royal Children's Hospital
Melbourne, Australia

Nina Irwin
Cambridge, Massachusetts

Vishwanath R. Iyer
University of Texas at Austin
Austin, Texas

Kenneth A. Jacobs
Genetics Institute
Cambridge, Massachusetts

Yongsu Jeong
University of Pennsylvania School of
Medicine
Philadelphia, Pennsylvania

Timm Jessen
Hoechst AG
Frankfurt am Main, Germany

Sulay D. Jhaveri
University of Texas
Austin, Texas

Yuan Jiang
National Human Genome Research
Institute, NIH
Bethesda, Maryland

Matthias John
Ribopharma AG
Kulmbach, Germany

Stephen Albert Johnston
University of Texas Southwestern Medical
Center
Dallas, Texas

Bryan Jones
Bristol-Myers Squibb
Seattle, Washington

Jingyue Ju
Incyte Pharmaceuticals
Palo Alto, California

Bechara Kachar
National Institute on Deafness and Other
Communication Disorders
Bethesda, Maryland

Mustak A. Kaderbhai
University College of Wales
Penglais, United Kingdom

James T. Kadonaga
University of California, San Diego
La Jolla, California

Jason A. Kahana
Dana-Farber Cancer Institute
Boston, Massachusetts

Steven R. Kain
Clontech Laboratories
Palo Alto, California

Fumiaki Katagiri
University of Minnesota
St. Paul, Minnesota

Randal J. Kaufman
University of Michigan
Ann Arbor, Michigan

Ning Ke
Iowa State University
Ames, Iowa

Marie Keaveney
Harvard Medical School
Boston, Massachusetts

Anthony D. Keefe
Massachusetts General Hospital
Boston, Massachusetts

Leslie A. Kerrigan
University of California, San Diego
La Jolla, California

Jae Bum Kim
Harvard Medical School
Boston, Massachusetts

Jonghwan Kim
University of Texas at Austin
Austin, Texas

Jung-Sik Kim
Georgetown University School of Medicine
Washington, D.C.

Robert E. Kingston
Massachusetts General Hospital and
Harvard Medical School
Boston, Massachusetts

Carol M. Kissinger
Millipore
Burlington, Massachusetts

Christoph A. Klein
Ludwig-Maximilians-University of Munich
Munich, Germany

Lloyd B. Klickstein
Brigham and Women's Hospital
Boston, Massachusetts

Joan H.M. Knoll
Harvard Medical School
Boston, Massachusetts

Mikhail G. Kolonin
Wayne State University School of Medicine
Detroit, Michigan

Elena Kotova
Fox Chase Cancer Center
Philadelphia, Pennsylvania

Martha F. Kramer
Harvard Medical School
Boston, Massachusetts

Nandini Krishnamurthy
University of California
Berkeley, California

Anuj Kumar
Yale University
New Haven, Connecticut

Thomas A. Kunkel
National Institute of Environmental Health
Sciences
Research Triangle Park, North Carolina

David M. Kurnit
University of Michigan Medical Center
Ann Arbor, Michigan

Joshua LaBaer
Harvard Medical School
Cambridge, Massachusetts

Mariana Lagos-Quintana
The Rockefeller University
New York, New York

Edward R. LaVallie
Genetics Institute
Cambridge, Massachusetts

Gretchen Lawler
Purdue University
West Lafayette, Indiana

Carolyn Lee
Georgetown University School of Medicine
Washington, D.C.

Chung-Sheng Brian Lee
Harvard Medical School
Boston, Massachusetts

Kyu-Min Lee
Massachusetts General Hospital
Boston, Massachusetts

Richard T. Lee
Brigham and Women's Hospital
Boston, Massachusetts

Michael Lenardo
National Institute of Allergy and Infectious
Diseases
Bethesda, Maryland

Mark E. Levenstein
University of California, San Diego
La Jolla, California

Peng Liang
Vanderbilt-Ingram Cancer Center
Nashville, Tennessee

Peter Lichter
Deutsches Krebsforschungs-zentrum
Heidelberg, Germany

Peter Liljeström
Karolinska Institute, Novum Research
Center
Huddinge, Sweden

Paul D. Ling
Baylor College of Medicine
Houston, Texas

Robert J. Linhardt
University of Iowa
Iowa City, Iowa

Peter Linsley
Bristol-Myers Squibb
Seattle, Washington

Silvia López
University of Buenos Aires-CONICET
School of Medicine
Buenos Aires, Argentina

John Lueders
National Human Genome Research
Institute, NIH
Bethesda, Maryland

Victoria Lundblad
Baylor College of Medicine
Houston, Texas

George Lunn
Baltimore, Maryland

Ming-Juan Luo
Harvard Medical School
Boston, Massachusetts

Karol Mackey
Molecular Research
Cincinnati, Ohio

Thomas L. Madden
National Center for Biotechnology
Information
Bethesda, Maryland

David J. Mangelsdorf
Howard Hughes Medical Institute
University of Texas Southwestern Medical
Center
Dallas, Texas

Adriana E. Manzi
University of California, San Diego
La Jolla, California

Chris D. Martin
Tropix Inc.
Bedford, Massachusetts

David Martin
National Institute of Allergy and Infectious
Diseases
Bethesda, Maryland

John M. McCoy
Genetics Institute
Cambridge, Massachusetts

Michael McPhaul
University of Texas Southwestern Medical
Center
Dallas, Texas

Scott Medberry
Amersham Pharmacia Biotech
San Francisco, California

Annemarie H. Meijer
Leiden University
Leiden, The Netherlands

Jill Meisenhelder
The Salk Institute
La Jolla, California

Craig C. Mello
University of Massachusetts Medical School
Worcester, Massachusetts

Andrew R. Mendelsohn
The Molecular Sciences Institute
Berkeley, California

Adriana Miele
University of Massachusetts Medical School
Worcester, Massachusetts

James J. Mond
Uniformed Services University of the
Health Sciences
Bethesda, Maryland

Rebecca I. Montgomery
Northwestern University
Chicago, Illinois

Butch Moomaw
Hamamatsu Photonic Systems
Spring Branch, Texas

David D. Moore
Baylor College of Medicine
Houston, Texas

Malcolm Moos, Jr.
Center for Biologics Evaluation and
Research
Food and Drug Administration
Bethesda, Maryland

Zarmik Moqtaderi
Harvard Medical School
Boston, Massachusetts

Richard Mortensen
University of Michigan Medical Center
Ann Arbor, Michigan

Bernard Moss
National Institute of Allergy and Infectious
Diseases
Bethesda, Maryland

Paul R. Mueller
California Institute of Technology
Pasadena, California

Cheryl Isaac Murphy
Aquila Biopharmaceuticals
Worcester, Massachusetts

Richard M. Myers
Stanford University School of Medicine
Stanford, California

Geeta Narlikar
Massachusetts General Hospital
Boston, Massachusetts

Robert J. Nelson
Dakota Scientific
Sioux Falls, South Dakota

Rachael L. Neve
McLean Hospital
Belmont, Massachusetts

Tagvor G. Nishanian
Georgetown University School of Medicine
Washington, D.C.

Paul E. Nisson
Life Technologies, Inc.
Gaithersburg, Maryland

B. Tracy Nixon
Pennsylvania State University
University Park, Pennsylvania

Marjorie Oettinger
Massachusetts General Hospital
Boston, Massachusetts

Osamu Ohara
Shionogi Research Laboratories
Osaka, Japan

Kazuhiro Oka
Baylor College of Medicine
Houston, Texas

Hiroto Okayama
Osaka University
Osaka, Japan

Salvatore Oliviero
Harvard Medical School
Boston, Massachusetts

Amos Oppenheim
National Cancer Institute at Frederick
Frederick, Maryland

Kazutoyo Osoegawa
Children's Hospital Oakland Research
Institute
Oakland, California

B.F. Francis Ouellette
University of British Columbia
Vancouver, British Columbia

Pieter B.F. Ouwkerk
Leiden University
Leiden, The Netherlands

Alejandra Paganelli
University of Buenos Aires-CONICET
School of Medicine
Buenos Aires, Argentina

Nicholas M. Papadopoulos
Warren Grant Magnuson Clinical Center
Bethesda, Maryland

Ophelia Papoulas
Harvard Medical School
Boston, Massachusetts

Arthur B. Pardee
Dana Farber Cancer Institute
Boston, Massachusetts

Jaehong Park
Harvard Medical School
Cambridge, Massachusetts

Mukesh Patel
Whitehead Institute for Biomedical
Research
Cambridge, Massachusetts

Yvonne Paterson
University of Pennsylvania
Philadelphia, Pennsylvania

Scott D. Patterson
Amgen, Inc.
Thousand Oaks, California

Warren Pear
University of Pennsylvania
Philadelphia, Pennsylvania

Fred A. Pereira
Baylor College of Medicine
Houston, Texas

Heather Perry-O'Keefe
Millipore
Burlington, Massachusetts

Jean M. Pesola
Harvard Medical School
Boston, Massachusetts

Karina Petat-Dutter
Ludwig-Maximilians-University of Munich
Munich, Germany

Kevin J. Petty
University of Texas Southwestern Medical
Center
Dallas, Texas

Sébastien Pfeffer
The Rockefeller University
New York, New York

Mary C. Phelan
Molecular Pathology Laboratory Network
Maryville, Tennessee

Helen Piwnica-Worms
Washington University School of Medicine
St. Louis, Missouri

Anne Pizard
Harvard Medical School
Boston, Massachusetts

Jack D. Pollard, Jr.
Mintz Levin Cohn Glosky Ferris and
Popeo
Boston, Massachusetts

Roy M. Pollock
Ariad Pharmaceuticals, Inc.
Cambridge, Massachusetts

Kornelia Polyak
Harvard Medical School
Boston, Massachusetts

Roger Pomerantz
Thomas Jefferson University
Philadelphia, Pennsylvania

Huntington Potter
Harvard Medical School
Boston, Massachusetts

Leland D. Powell
University of California, San Diego School
of Medicine
La Jolla, California

Pedro A. Prieto
Ross Laboratories
Columbus, Ohio

Thomas Quertermous
Massachusetts General Hospital
Boston, Massachusetts

Elisabeth Raleigh
New England Biolabs
Beverly, Massachusetts

Ali Ramezani
American Red Cross
Rockville, Maryland

Ayoub Rashtchian
Life Technologies, Inc.
Gaithersburg, Maryland

Hervé E. Recipon
National Center for Biotechnology
Information
Bethesda, Maryland

K.J. Reddy
State University of New York
Binghamton, New York

Robin Reed
Harvard Medical School
Boston, Massachusetts

Mark Reichardt
Lakeside Biotechnology
Chicago, Illinois

Ann Reynolds
University of Washington
Seattle, Washington

Randall K. Ribaldo
National Institute of Allergy and Infectious
Diseases
Bethesda, Maryland

Eric J. Richards
Washington University
St. Louis, Missouri

Paul Riggs
New England Biolabs
Beverly, Massachusetts

Phil Robakiewicz
Worcester Polytechnic Institute
Worcester, Massachusetts

Jeffrey Robbins
Cincinnati Children's Hospital
University of Cincinnati
Cincinnati, Ohio

Melissa Rogers
University of South Florida
Tampa, Florida

Sharon Rogers
Lakeside Biotechnology
Chicago, Illinois

M.R. Rolfe
Celltech Ltd.
Slough, United Kingdom

William G. Romanow
Corning Costar
Portsmouth, New Hampshire

John K. Rose
Yale University School of Medicine
New Haven, Connecticut

Martin Rosenberg
SmithKline Beecham
King of Prussia, Pennsylvania

Sharon Y. Roth
M.D. Anderson Cancer Center
Houston, Texas

Robert R. Roussel
Dartmouth College
Hanover, New Hampshire

Hadara Rubinfeld
The Weizmann Institute of Science
Rehovot, Israel

Colleen A. Ryan
Boston College
Chestnut Hill, Massachusetts

Elizabeth F. Ryder
Worcester Polytechnic Institute
Worcester, Massachusetts

Nicoletta Sacchi
National Cancer Institute
Frederick, Maryland

Brad St. Croix
Harvard Medical School
Boston, Massachusetts

Thomas P. St. John
ICOS Corporation
Bothwell, Washington

Joachim Sasse
Shriners Hospital for Crippled Children
Tampa, Florida

N.M. Sawtell
Children's Hospital Medical Center
Cincinnati, Ohio

Stephen J. Scharf
Cetus Corporation
Emeryville, California

David G. Schatz
Howard Hughes Medical Institute and Yale
University School of Medicine
New Haven, Connecticut

Paul Schendel
Genetics Institute
Cambridge, Massachusetts

Timothy D. Schlabach
Applied Biosystems
Foster City, California

Gavin R. Schnitzler
Tufts University School of Medicine
Boston, Massachusetts

Joachim Schorr
Diagen GmbH
Hilden, Germany

Gregory D. Schuler
National Center for Biotechnology
Information
Bethesda, Maryland

R.K. Scopes
La Trobe University
Bundoora, Australia

Brian Seed
Massachusetts General Hospital and
Harvard Medical School
Boston, Massachusetts

Bartholomew M. Sefton
The Salk Institute
San Diego, California

Rony Seger
Department of Biological Regulation
The Weizmann Institute of Science
Rehovot, Israel

Christine E. Seidman
Harvard Medical School
Boston, Massachusetts

Edward Sekinger
Harvard Medical School
Boston, Massachusetts

Richard F. Selden
TKT, Inc.
Cambridge, Massachusetts

Kentaro Semba
The Salk Institute
La Jolla, California

Donald Senear
University of California, Irvine
Irvine, California

Ilya Serebriiskii
Fox Chase Cancer Center
Philadelphia, Pennsylvania

Thikkavarapu Seshamma
Thomas Jefferson University
Philadelphia, Pennsylvania

Raj Shankarappa
University of Pittsburgh School of Medicine
Pittsburgh, Pennsylvania

Allan Shatzman
SmithKline Beecham
King of Prussia, Pennsylvania

Yoav Shaul
Department of Biological Regulation
The Weizmann Institute of Science
Rehovot, Israel

Jen Sheen
Massachusetts General Hospital and
Harvard Medical School
Boston, Massachusetts

Shirish Shenolikar
Duke University Medical Center
Durham, North Carolina

L.A. Sherman
Purdue University
West Lafayette, Indiana

Penny Shockett
Yale University School of Medicine
New Haven, Connecticut

Kevan M. Shokat
University of California at San Francisco
San Francisco, California

Gene P. Siegal
University of Alabama at Birmingham
Birmingham, Alabama

Pam A. Silver
Dana-Farber Cancer Institute
Boston, Massachusetts

Michael H. Simonian
Beckman Instruments, Inc.
Fullerton, California

Harinder Singh
University of Chicago
Chicago, Illinois

Hazel Sive
Whitehead Institute for Biomedical
Research
Cambridge, Massachusetts

Kimmen V. Sjölander
University of California
Berkeley, California

Barton E. Slatko
New England Biolabs
Beverly, Massachusetts

Alan J. Smith
Stanford University Medical Center
Stanford, California

Carolyn L. Smith
National Institute of Neurological Disorders
and Stroke
Bethesda, Maryland

David F. Smith
University of Georgia
Atlanta, Georgia

Donald B. Smith
University of Edinburgh
Edinburgh, Scotland

John A. Smith
University of Alabama at Birmingham
Birmingham, Alabama

Michael Snyder
Yale University
New Haven, Connecticut

Timothy A. Springer
Center for Blood Research
Harvard Medical School
Boston, Massachusetts

Patrick Stanssens
Keygene N.V.
The Netherlands

David F. Stern
Yale University School of Medicine
New Haven, Connecticut

Julie M. Stone
University of Missouri
Columbia, Missouri

William M. Strauss
Whitehead Institute
Cambridge, Massachusetts

Kevin Struhl
Harvard Medical School
Boston, Massachusetts

John T. Stults
Genentech, Inc.
South San Francisco, California

Cynthia Sung
Bioengineering and Physical Science
Program/NIH
Bethesda, Maryland

Jesper Q. Svejstrup
Cancer Research U.K.
London Research Institute
Clare Hall Laboratory
Hertfordshire, U.K.

Jonathan C. Swaffield
University of Texas Southwestern Medical
Center
Dallas, Texas

Stanley Tabor
Harvard Medical School
Boston, Massachusetts

Tomoko M. Tabuchi
University of Massachusetts Medical School
Worcester, Massachusetts

Miyoko Takahashi
Spectral Diagnostics, Inc.
Toronto, Ontario

Christophe Thiriet
University of Rochester Medical Center
Rochester, New York

Lynn Thomason
National Cancer Institute at Frederick
Frederick, Maryland

Nancy E. Thompson
University of Wisconsin
Madison, Wisconsin

Douglas A. Treco
TKT, Inc.
Cambridge, Massachusetts

Steven J. Triezenberg
Michigan State University
East Lansing, Michigan

Thomas Tuschl
The Rockefeller University
New York, New York

Peter van der Geer
University of California, San Diego
La Jolla, California

Ajit P. Varki
University of California, San Diego
La Jolla, California

Baruch Velan
Israel Institute of Biological Research
Ness Ziona, Israel

Jacob Villegas
University of Texas Southwestern Medical
Center
Dallas, Texas

Joseph M. Vitolo
University of Rochester Medical Center
Rochester, New York

Hans-Peter Vornlocher
Ribopharma AG
Kulmbach, Germany

Pieter Vos
Keygene N.V.
The Netherlands

Daniel Voytas
Iowa State University
Ames, Iowa

Todd Waldman
Georgetown University School of Medicine
Washington, D.C.

Simon Watkins
University of Pittsburgh Medical School
Pittsburgh, Pennsylvania

John H. Weis
Harvard Medical School
Boston, Massachusetts

Jane M. Weisemann
National Center for Biotechnology
Information
Bethesda, Maryland

Douglas C. Weiser
Duke University Medical Center
Durham, North Carolina

Eric L. Weiss
Northwestern University
Evanston, Illinois

Nicole Wendler
Ludwig-Maximilians-University of Munich
Munich, Germany

Maryann Z. Whitley
Genetics Institute
Cambridge, Massachusetts

Michael Whitt
Yale University School of Medicine
New Haven, Connecticut

Alan Williams
Amersham Pharmacia Biotech
Piscataway, New Jersey

David S. Wilson
Zyomyx
Hayward, California

Helen Wilson
National Cancer Institute at Frederick
Frederick, Maryland

Kate Wilson
Australian Institute of Marine Science
Townsville, Australia

Kenneth J. Wilson
Applied Biosystems
Foster City, California

Fred Winston
Harvard Medical School
Boston, Massachusetts

Scott E. Winston
Nabi
Rockville, Maryland

C. Richard Wobbe
Harvard Medical School
Boston, Massachusetts

Barbara Wold
California Institute of Technology
Pasadena, California

Paula Wolf
Massachusetts Institute of Technology
Cambridge, Massachusetts

Tyra G. Wolfsberg
National Center for Biotechnology
Information
Bethesda, Maryland

Carolyn A. Worby
University of California at San Diego
La Jolla, California

Jerry L. Workman
Pennsylvania State University
University Park, Pennsylvania

Shwu-Yuan Wu
Case Western Reserve University School of

Medicine
Cleveland, Ohio

Linda S. Wyatt
National Institute of Allergy and Infectious
Diseases
Bethesda, Maryland

Dong Xu
Digital Biology Laboratory
University of Missouri-Columbia
Columbia, Missouri

Jianming Xu
Baylor College of Medicine
Houston, Texas

Ying Xu
Department of Biochemistry and Molecular
Biology
University of Georgia
Athens, Georgia

Annie Yang
Harvard Medical School
Boston, Massachusetts

Yanhai Yin
Iowa State University
Ames, Iowa

Wayne M. Yokoyama
University of California School of Medicine
San Francisco, California

Natasha E. Zachara
The Johns Hopkins University Medical
School
Baltimore, Maryland

Ken Zaret
Fox Chase Cancer Center
Philadelphia, Pennsylvania

Rolf Zeller
University of Utrecht
The Netherlands

Chao Zhang
University of California at San Francisco
San Francisco, California

Ning Zhang
Tularik, Inc.
South San Francisco, California

Heng Zhou
Harvard Medical School
Boston, Massachusetts

Zhaolan Zhou
Children's Hospital and Harvard Medical
School
Boston, Massachusetts

Dietlind Zohlh fer
Ludwig-Maximilians-University of Munich
Munich, Germany

Lou Zumstein
Baylor College of Medicine
Houston, Texas

Contents

CHAPTER 1 ESCHERICHIA COLI, PLASMIDS, AND BACTERIOPHAGES

Introduction

Section I Escherichia coli

- Unit 1.1 Media Preparation and Bacteriological Tools
- Unit 1.2 Growth in Liquid Media
- Unit 1.3 Growth on Solid Media
- Unit 1.4 Selected Topics from Classical Bacterial Genetics

Section II Vectors Derived from Plasmids

- Unit 1.5 Introduction to Plasmid Biology
- Unit 1.6 Minipreps of Plasmid DNA
- Unit 1.7 Large-Scale Preparation of Plasmid DNA
- Unit 1.8 Introduction of Plasmid DNA into Cells

Section III Vectors Derived from Lambda and Related Bacteriophages

- Unit 1.9 Introduction to Lambda Phages
- Unit 1.10 Lambda as a Cloning Vector
- Unit 1.11 Plating Lambda Phage to Generate Plaques
- Unit 1.12 Growing Lambda-Derived Vectors
- Unit 1.13 Preparing Lambda DNA from Phage Lysates

Section IV Vectors Derived from Filamentous Phages

- Unit 1.14 Introduction to Vectors Derived from Filamentous Phages
- Unit 1.15 Preparing and Using M13-Derived Vectors

Section V Specialized Techniques

- Unit 1.16 Recombineering: Genetic Engineering in Bacteria Using Homologous Recombination

CHAPTER 2 PREPARATION AND ANALYSIS OF DNA

Introduction

Section I Manipulation of DNA

- Unit 2.1A Purification and Concentration of DNA from Aqueous Solutions
- Unit 2.1B Purification of DNA by Anion-Exchange Chromatography
- Unit 2.2 Preparation of Genomic DNA from Mammalian Tissue
- Unit 2.3 Preparation of Genomic DNA from Plant Tissue
- Unit 2.4 Preparation of Genomic DNA from Bacteria

Section II Resolution and Recovery of Large DNA Fragments

- Unit 2.5A Agarose Gel Electrophoresis
- Unit 2.5B Pulsed-Field Gel Electrophoresis
- Unit 2.6 Isolation and Purification of Large DNA Restriction Fragments from Agarose Gels

Section III Resolution and Recovery of Small DNA Fragments

Unit 2.7 Separation of Small DNA Fragments by Conventional Gel Electrophoresis

Unit 2.8 Capillary Electrophoresis of DNA

Section IV Analysis of DNA Sequences by Blotting and Hybridization

Unit 2.9A Southern Blotting

Unit 2.9B Dot and Slot Blotting of DNA

Unit 2.10 Hybridization Analysis of DNA Blots

Section V Synthesis and Purification of Oligonucleotides

Unit 2.11 Synthesis and Purification of Oligonucleotides

Unit 2.12 Purification of Oligonucleotides Using Denaturing Polyacrylamide Gel Electrophoresis

CHAPTER 3 ENZYMATIC MANIPULATION OF DNA AND RNA

Introduction

Section I Restriction Endonucleases

Unit 3.1 Digestion of DNA with Restriction Endonucleases

Section II Restriction Mapping

Unit 3.2 Mapping by Multiple Endonuclease Digestions

Unit 3.3 Mapping by Partial Endonuclease Digestions

Section III Enzymes for Modifying and Radioactively Labeling Nucleic Acids

Unit 3.4 Reagents and Radioisotopes Used to Manipulate Nucleic Acids

Unit 3.5 DNA-Dependent DNA Polymerases

Unit 3.6 Template-Independent DNA Polymerases

Unit 3.7 RNA-Dependent DNA Polymerases

Unit 3.8 DNA-Dependent RNA Polymerases

Unit 3.9 DNA-Independent RNA Polymerases

Unit 3.10 Phosphatases and Kinases

Unit 3.11 Exonucleases

Unit 3.12 Endonucleases

Unit 3.13 Ribonucleases

Unit 3.14 DNA Ligases

Unit 3.15 RNA Ligases

Section IV Construction of Hybrid DNA Molecules

Unit 3.16 Subcloning of DNA Fragments

Unit 3.17 Constructing Recombinant DNA Molecules by the Polymerase Chain Reaction

Section V Specialized Applications

Unit 3.18 Labeling and Colorimetric Detection of Nonisotopic Probes

Unit 3.19 Chemiluminescent Detection of Nonisotopic Probes

Unit 3.20 Recombinational Cloning

CHAPTER 4 PREPARATION AND ANALYSIS OF RNA

Introduction

Section I Preparation of RNA from Eukaryotic and Prokaryotic Cells

Unit 4.1 Preparation of Cytoplasmic RNA from Tissue Culture Cells

Unit 4.2 Guanidine Methods for Total RNA Preparation

Unit 4.3 Phenol/SDS Method for Plant RNA Preparation

Unit 4.4 Preparation of Bacterial RNA

Unit 4.5 Preparation of Poly(A)⁺ RNA

Section II Analysis of RNA Structure and Synthesis

Unit 4.6 S1 Analysis of Messenger RNA Using Single-Stranded DNA Probes

Unit 4.7 Ribonuclease Protection Assay

Unit 4.8 Primer Extension

Unit 4.9 Analysis of RNA by Northern and Slot Blot Hybridization

Unit 4.10 Identification of Newly Transcribed RNA

CHAPTER 5 CONSTRUCTION OF RECOMBINANT DNA LIBRARIES

Introduction

Section I Overview of Recombinant DNA Libraries

Unit 5.1 Genomic DNA Libraries

Unit 5.2 cDNA Libraries

Section II Preparation of Insert DNA from Genomic DNA

Unit 5.3 Size Fractionation Using Sucrose Gradients

Unit 5.4 Size Fractionation Using Agarose Gels

Section III Preparation of Insert DNA from Messenger RNA

Unit 5.5 Conversion of mRNA into Double-Stranded cDNA

Unit 5.6 Ligation of Linkers or Adapters to Double-Stranded cDNA

Section IV Production of Genomic DNA and cDNA Libraries

Unit 5.7 Production of a Genomic DNA Library

Unit 5.8A Production of a Complete cDNA Library

Unit 5.9 Construction of Bacterial Artificial Chromosome (BAC/PAC) Libraries

Section V Amplification of Transformed or Packaged Libraries

Unit 5.10 Amplification of a Bacteriophage Library

Unit 5.11 Amplification of Cosmid and Plasmid Libraries

CHAPTER 6 SCREENING OF RECOMBINANT DNA LIBRARIES

Introduction

Section I Plating Libraries and Transfer to Filter Membranes

Unit 6.1 Plating and Transferring Bacteriophage Libraries

Unit 6.2 Plating and Transferring Cosmid and Plasmid Libraries

Section II Hybridization with Radioactive Probes

Unit 6.3 Using DNA Fragments as Probes

Unit 6.4 Using Synthetic Oligonucleotides as Probes

Section III Purification of Bacteriophage, Cosmid, and Plasmid Clones

Unit 6.5 Purification of Bacteriophage Clones

Unit 6.6 Purification of Cosmid and Plasmid Clones

Section IV Screening with Antibodies

Unit 6.7 Immunoscreening of Fusion Proteins Produced in Lambda Plaques

Unit 6.8 Immunoscreening after Hybrid Selection and Translation

Section V Yeast Artificial Chromosome Libraries

Unit 6.9 Overview of Strategies for Screening YAC Libraries and Analyzing YAC Clones

Unit 6.10 Analysis of Isolated YAC Clones

Section VI Specialized Strategies for Screening Libraries

Unit 6.11 Use of Monoclonal Antibodies for Expression Cloning

Unit 6.12 Recombination-Based Assay (RBA) for Screening Bacteriophage Lambda Libraries

CHAPTER 7 DNA SEQUENCING

Introduction

Unit 7.1 DNA Sequencing Strategies

Unit 7.2 Constructing Nested Deletions for Use in DNA Sequencing

Unit 7.3 Preparation of Templates for DNA Sequencing

Unit 7.4A DNA Sequencing by the Dideoxy Method

Unit 7.4B Dideoxy DNA Sequencing with Chemiluminescent Detection

Unit 7.5 DNA Sequencing by the Chemical Method

Unit 7.6 Denaturing Gel Electrophoresis for Sequencing

Unit 7.7 Computer Manipulation of DNA and Protein Sequences

Unit 7.8 Polony DNA Sequencing

CHAPTER 8 MUTAGENESIS OF CLONED DNA

Introduction

Unit 8.1 Oligonucleotide-Directed Mutagenesis without Phenotypic Selection

Unit 8.2A Mutagenesis with Degenerate Oligonucleotides: Creating Numerous Mutations in a Small DNA Sequence

Unit 8.2B Gene Synthesis: Assembly of Target Sequences Using Mutually Priming Long Oligonucleotides

Unit 8.3 Random Mutagenesis by PCR

Unit 8.4 Linker-Scanning Mutagenesis of DNA

Unit 8.5 Directed Mutagenesis Using the Polymerase Chain Reaction

CHAPTER 9 INTRODUCTION OF DNA INTO MAMMALIAN CELLS

Introduction

Section I Transfection of DNA into Eukaryotic Cells

Unit 9.1 Calcium Phosphate Transfection

Unit 9.2 Transfection Using DEAE-Dextran

Unit 9.3 Transfection by Electroporation

Unit 9.4 Transfection of Cultured Eukaryotic Cells Using Cationic Lipid Reagents

Unit 9.5 Selection of Transfected Mammalian Cells

Section II Uses of Fusion Genes in Mammalian Transfection

Unit 9.6 Overview of Genetic Reporter Systems

Unit 9.7A Isotopic Assays for Reporter Gene Activity

Unit 9.7B Nonisotopic Assays for Reporter Gene Activity

Unit 9.7C Use of the *A. Victoria* Green Fluorescent Protein to Study Protein Dynamics in Vivo

Unit 9.8 Direct Analysis of RNA after Transfection

Section III Transduction of Genes Using Retrovirus Vectors

Unit 9.9 Overview of the Retrovirus Transduction System

Unit 9.10 Preparation of a Specific Retrovirus Producer Cell Line

Unit 9.11 Transient Transfection Methods for Preparation of High-Titer Retroviral Supernatants

Unit 9.12 Large-Scale Preparation and Concentration of Retrovirus Stocks

Unit 9.13 Detection of Helper Virus in Retrovirus Stocks

Unit 9.14 Retrovirus Infection of Cells In Vitro and In Vivo

Section IV Inactivation of Genes in Mammalian Cells

Unit 9.15 Human Somatic Cell Gene Targeting

CHAPTER 10 ANALYSIS OF PROTEINS

Introduction

Section I Quantitation of Proteins

Unit 10.1A Spectrophotometric and Colorimetric Determination of Protein Concentration

Unit 10.1B Quantitative Amino Acid Analysis

Section II Electrophoretic Separation of Proteins

Unit 10.2A One-Dimensional SDS Gel Electrophoresis of Proteins

Unit 10.2B One-Dimensional Electrophoresis Using Nondenaturing Conditions

Unit 10.3 Two-Dimensional Gel Electrophoresis Using the ISO-DALT System

Unit 10.4 Two-Dimensional Gel Electrophoresis

Unit 10.5 Overview of Digital Electrophoresis Analysis

Section III Detection of Proteins

- Unit 10.6 Staining Proteins in Gels
- Unit 10.7 Detection of Proteins on Blot Transfer Membranes
- Unit 10.8 Immunoblotting and Immunodetection

Section IV Purification of Proteins by Conventional Chromatography

- Unit 10.9 Gel-Filtration Chromatography
- Unit 10.10 Ion-Exchange Chromatography
- Unit 10.11A Immunoaffinity Chromatography
- Unit 10.11B Metal-Chelate Affinity Chromatography

Section V Purification of Proteins by High-Performance Liquid Chromatography

- Unit 10.12 HPLC of Peptides and Proteins: Preparation and System Set-Up
- Unit 10.13 HPLC of Peptides and Proteins: Standard Operating Conditions
- Unit 10.14 Reversed-Phase Isolation of Peptides

Section VI Specialized Applications

- Unit 10.15 Purification of Recombinant Proteins and Study of Protein Interaction by Epitope Tagging
- Unit 10.16 Immunoprecipitation
- Unit 10.17 Synthesizing Proteins In Vitro by Transcription and Translation of Cloned Genes
- Unit 10.18 Metabolic Labeling with Amino Acids
- Unit 10.19 Isolation of Proteins for Microsequence Analysis
- Unit 10.20 Capillary Electrophoresis of Proteins and Peptides
- Unit 10.21 Overview of Peptide and Protein Analysis by Mass Spectrometry
- Unit 10.22 Protein Identification and Characterization by Mass Spectrometry
- Unit 10.23 Difference Gel Electrophoresis (DIGE) Using CyDye DIGE Fluor Minimal Dyes
- Unit 10.24 Solution Radioimmunoassay of Proteins and Peptides

CHAPTER 11 IMMUNOLOGY

Introduction

Section I Immunoassays

- Unit 11.1 Conjugation of Enzymes to Antibodies
- Unit 11.2 Enzyme-Linked Immunosorbent Assays (ELISA)
- Unit 11.3 Isotype Determination of Antibodies

Section II Preparation of Monoclonal Antibodies

- Unit 11.4 Immunization of Mice
- Unit 11.5 Preparation of Myeloma Cells
- Unit 11.6 Preparation of Mouse Feeder Cells for Fusion and Cloning
- Unit 11.7 Fusion of Myeloma Cells with Immune Spleen Cells
- Unit 11.8 Cloning of Hybridoma Cell Lines by Limiting Dilution
- Unit 11.9 Freezing and Recovery of Hybridoma Cell Lines
- Unit 11.10 Production of Monoclonal Antibody Supernatant and Ascites Fluid
- Unit 11.11 Purification of Monoclonal Antibodies

Section III Preparation of Polyclonal Antisera

Unit 11.12 Production of Polyclonal Antisera

Unit 11.13 In Vitro Antibody Production

Unit 11.14 Purification of Immunoglobulin G Fraction from Antiserum, Ascites Fluid, or Hybridoma Supernatant

Section IV Preparation of Antipeptide Antibodies

Unit 11.15 Introduction to Peptide Synthesis

Unit 11.16 Synthetic Peptides for Production of Antibodies that Recognize Intact Proteins

Section V Determination of Specific Antibody Titer and Isotype

Unit 11.17 Determination of the Specific Antibody Titer

Section VI Preparation and Use of Specialized Antibodies

Unit 11.18 Identification of Polyol-Responsive Monoclonal Antibodies for Use in Immunoaffinity Chromatography

CHAPTER 12 DNA-PROTEIN INTERACTIONS

Introduction

Unit 12.1 Preparation of Nuclear and Cytoplasmic Extracts from Mammalian Cells

Unit 12.2 Mobility Shift DNA-Binding Assay Using Gel Electrophoresis

Unit 12.3 Methylation and Uracil Interference Assays for Analysis of Protein-DNA Interactions

Unit 12.4 DNase I Footprint Analysis of Protein-DNA Binding

Unit 12.5 UV Crosslinking of Proteins to Nucleic Acids

Unit 12.6 Purification of DNA-Binding Proteins Using Biotin/Streptavidin Affinity Systems

Unit 12.7 Detection, Purification, and Characterization of cDNA Clones Encoding DNA-Binding Proteins

Unit 12.8 Rapid Separation of Protein-Bound DNA from Free DNA Using Nitrocellulose Filters

Unit 12.9 Analysis of DNA-Protein Interactions Using Proteins Synthesized In Vitro from Cloned Genes

Unit 12.10 Purification of Sequence-Specific DNA-Binding Proteins by Affinity Chromatography

Unit 12.11 Determination of Protein-DNA Sequence Specificity by PCR-Assisted Binding-Site Selection

Unit 12.12 Yeast One-Hybrid Screening for DNA-Protein Interactions

CHAPTER 13 YEAST

Introduction

Section I Basic Techniques of Yeast Genetics

Unit 13.1 Preparation of Yeast Media

Unit 13.2 Growth and Manipulation of Yeast

Unit 13.3 Genome-Wide Transposon Mutagenesis in Yeast

Section II Yeast Vectors

Unit 13.4 Yeast Cloning Vectors and Genes

Unit 13.6 Yeast Vectors and Assays for Expression of Cloned Genes

Section III Manipulation of Yeast Genes

Unit 13.7 Introduction of DNA into Yeast Cells

- Unit 13.8 Cloning Yeast Genes by Complementation
- Unit 13.9 Manipulation of Plasmids from Yeast Cells
- Unit 13.10 Manipulation of Cloned Yeast DNA

Section IV Preparation of Yeast DNA, RNA, and Proteins

- Unit 13.11 Preparation of Yeast DNA
- Unit 13.12 Preparation of Yeast RNA
- Unit 13.13 Preparation of Protein Extracts from Yeast

Section V *Schizosaccharomyces pombe*

- Unit 13.14 Overview of *Schizosaccharomyces pombe*
- Unit 13.15 *S. pombe* Strain Maintenance and Media
- Unit 13.16 Growth and Manipulation of *S. pombe*
- Unit 13.17 Introduction of DNA into *S. pombe* Cells

CHAPTER 14 IN SITU HYBRIDIZATION AND IMMUNOHISTOCHEMISTRY

Introduction

- Unit 14.1 Fixation, Embedding, and Sectioning of Tissues, Embryos, and Single Cells
- Unit 14.2 Cryosectioning
- Unit 14.3 In Situ Hybridization to Cellular RNA
- Unit 14.4 Detection of Hybridized Probe
- Unit 14.5 Counterstaining and Mounting of Autoradiographed In Situ Hybridization Slides
- Unit 14.6 Immunohistochemistry
- Unit 14.7 In situ Hybridization and Detection Using Nonisotopic Probes
- Unit 14.8 In situ Polymerase Chain Reaction and Hybridization to Detect Low-Abundance Nucleic Acid Targets
- Unit 14.9 Whole-Mount In Situ Hybridization and Detection of RNAs in Vertebrate Embryos and Isolated Organs
- Unit 14.10 Principles and Application of Fluorescence Microscopy
- Unit 14.11 Basic Confocal Microscopy
- Unit 14.12 Measurement of In Situ Hybridization
- Unit 14.13 Morphological, Biochemical, and Flow Cytometric Assays of Apoptosis
- Unit 14.14 Whole-Mount Histochemical Detection of β -Galactosidase Activity
- Unit 14.15 Overview of Image Analysis, Image Importing, and Image Processing using Freeware
- Unit 14.16 Three-Dimensional Reconstruction of Tissues

CHAPTER 15 THE POLYMERASE CHAIN REACTION

Introduction

- Unit 15.1 Enzymatic Amplification of DNA by PCR: Standard Procedures and Optimization
- Unit 15.2 Direct DNA Sequencing of PCR Products
- Unit 15.3 Ligation-Mediated PCR for Genomic Sequencing and Footprinting
- Unit 15.4 Molecular Cloning of PCR Products
- Unit 15.5 Enzymatic Amplification of RNA by PCR (RT-PCR)
- Unit 15.6 cDNA Amplification Using One-Sided (Anchored) PCR
- Unit 15.7 Quantitation of Rare DNAs by PCR
- Unit 15.8 High-Throughput Real-Time Quantitative Reverse Transcription PCR

CHAPTER 16 PROTEIN EXPRESSION

Introduction

Section I Expression of Proteins in *Escherichia coli*

- Unit 16.1 Overview of Protein Expression in *E. coli*
- Unit 16.2 Expression Using the T7 RNA Polymerase/Promoter System
- Unit 16.3 Expression Using Vectors with Phage λ Regulatory Sequences
- Unit 16.4A Introduction to Expression by Fusion Protein Vectors
- Unit 16.4B Enzymatic and Chemical Cleavage of Fusion Proteins
- Unit 16.5 Expression and Purification of lacZ and trpE Fusion Proteins
- Unit 16.6 Expression and Purification of Maltose-Binding Protein Fusions
- Unit 16.7 Expression and Purification of Glutathione-S-Transferase Fusion Proteins
- Unit 16.8 Expression and Purification of Thioredoxin Fusion Proteins

Section II Expression of Proteins in Insect Cells Using Baculovirus Vectors

- Unit 16.9 Overview of the Baculovirus Expression System
- Unit 16.10 Maintenance of Insect Cell Cultures and Generation of Recombinant Baculoviruses
- Unit 16.11 Expression and Purification of Recombinant Proteins Using the Baculovirus System

Section III Expression of Proteins in Mammalian Cells

- Unit 16.12 Transient Expression of Proteins Using COS Cells
- Unit 16.13 Expression and Purification of Epitope-Tagged Multisubunit Protein Complexes from Mammalian Cells
- Unit 16.14 Inducible Gene Expression Using an Autoregulatory, Tetracycline-Controlled System
- Unit 16.15 Overview of the Vaccinia Virus Expression System
- Unit 16.16 Preparation of Cell Cultures and Vaccinia Virus Stocks
- Unit 16.17 Generation of Recombinant Vaccinia Viruses
- Unit 16.18 Characterization of Recombinant Vaccinia Viruses and Their Products
- Unit 16.19 Gene Expression Using the Vaccinia Virus/T7 RNA Polymerase Hybrid System
- Unit 16.20 Expression of Proteins Using Semliki Forest Virus Vectors
- Unit 16.21 Overview of the HIV-1 Lentiviral Vector System
- Unit 16.22 Generation of HIV-1-Based Lentiviral Vector Particles
- Unit 16.23 Amplification Using CHO Cell Expression Vectors
- Unit 16.24 Helper-Dependent Adenoviral Vectors

CHAPTER 17 PREPARATION AND ANALYSIS OF GLYCOCONJUGATES

Introduction

Section I Special Considerations for Glycoproteins and Their Purification

- Unit 17.1 Special Considerations for Glycoproteins and Their Purification
- Unit 17.2 Special Considerations for Proteoglycans and Glycosaminoglycans and Their Purification
- Unit 17.3 Special Considerations for Glycolipids and Their Purification

Section II Detection of Saccharides on Glycoconjugates

- Unit 17.4 Metabolic Radiolabeling of Animal Cell Glycoconjugates
- Unit 17.5 Chemical Labeling of Carbohydrates by Oxidation and Sodium Borohydride Reduction
- Unit 17.6 Detection and Analysis of Proteins Modified by O-Linked N-Acetylglucosamine

- Unit 17.7 Lectin Analysis of Proteins Blotted onto Filters
- Unit 17.8 Detection of Glycophospholipid Anchors on Proteins
- Unit 17.9 Direct Chemical Analysis of Glycoconjugates for Carbohydrates
- Unit 17.10A Inhibition of N-Linked Glycosylation
- Unit 17.10B Inhibition of Glycolipid Biosynthesis
- Unit 17.11 Synthetic Glycosides as Primers of Oligosaccharide Biosynthesis and Inhibitors of Glycoprotein and Proteoglycan Assembly

Section III Release of Saccharides from Glycoconjugates

- Unit 17.12 Sialidases
- Unit 17.13A Endoglycosidase and Glycoamidase Release of N-Linked Oligosaccharides
- Unit 17.13B Analysis of Glycosaminoglycans with Polysaccharide Lyases
- Unit 17.14A Preparation of Glycopeptides
- Unit 17.14B Detection of Individual Glycosylation Sites on Glycoproteins
- Unit 17.15A β -Elimination for Release of O-Linked Glycosaminoglycans from Proteoglycans
- Unit 17.15B β -Elimination for Release of O-GalNAc-Linked Oligosaccharides from Glycoproteins and Glycopeptides
- Unit 17.16 Acid Hydrolysis for Release of Monosaccharides
- Unit 17.17A Enzymatic Release of Oligosaccharides from Glycolipids
- Unit 17.17B Endo- β -Galactosidases and Keratanase

Section IV Analysis of Saccharides Released from Glycoconjugates

- Unit 17.18 Analysis of Monosaccharides
- Unit 17.19A Total Compositional Analysis by High-Performance Liquid Chromatography or Gas-Liquid Chromatography
- Unit 17.19B Composition of Labeled Monosaccharides from Glycosaminoglycans
- Unit 17.20 Analysis of Oligosaccharide Negative Charge by Anion-Exchange Chromatography
- Unit 17.21A HPLC Methods for the Fractionation and Analysis of Negatively Charged Oligosaccharides and Gangliosides
- Unit 17.21B Fractionation and Analysis of Neutral Oligosaccharides by HPLC
- Unit 17.22A Nitrous Acid Degradation of Glycosaminoglycans
- Unit 17.22B Analysis of Disaccharides and Tetrasaccharides Released from Glycosaminoglycans
- Unit 17.23 Analysis of Sulfate Esters by Solvolysis or Hydrolysis

CHAPTER 18 ANALYSIS OF PROTEIN PHOSPHORYLATION

Introduction

- Unit 18.1 Overview of Protein Phosphorylation
- Unit 18.2 Labeling Cultured Cells with $^{32}\text{P}_i$ and Preparing Cell Lysates for Immunoprecipitation
- Unit 18.3 Phosphoamino Acid Analysis
- Unit 18.4 Analysis of Phosphorylation of Unlabeled Proteins
- Unit 18.5 Detection of Phosphorylation by Enzymatic Techniques
- Unit 18.6 Production of Antibodies That Recognize Specific Tyrosine-Phosphorylated Peptides
- Unit 18.7 Assays of Protein Kinases Using Exogenous Substrates
- Unit 18.8 Permeabilization Strategies to Study Protein Phosphorylation
- Unit 18.9 Phosphopeptide Mapping and Identification of Phosphorylation Sites
- Unit 18.10 Use of Protein Phosphatase Inhibitors
- Unit 18.11 Design and Use of Analog-Sensitive Protein Kinases
- Unit 18.12 The Detection of MAPK Signaling
- Unit 18.13 Isolation of Phosphopeptides by Immobilized Metal Ion Affinity Chromatography

CHAPTER 19 INFORMATICS FOR MOLECULAR BIOLOGISTS

Introduction

- Unit 19.1 Internet Basics for Biologists
- Unit 19.2 Sequence Databases: Integrated Information Retrieval and Data Submission
- Unit 19.3 Sequence Similarity Searching Using the BLAST Family of Programs
- Unit 19.4 Protein Databases on the Internet
- Unit 19.5 Basic Protein Sequence Analysis
- Unit 19.6 Analysis and Management of Microarray Gene Expression Data

CHAPTER 20 ANALYSIS OF PROTEIN INTERACTIONS

Introduction

- Unit 20.1 Interaction Trap/Two-Hybrid System to Identify Interacting Proteins
- Unit 20.2 Affinity Purification of Proteins Binding to GST Fusion Proteins
- Unit 20.3 Phage-Based Expression Cloning to Identify Interacting Proteins
- Unit 20.4 Surface Plasmon Resonance for Measurements of Biological Interest
- Unit 20.5 Detection of Protein-Protein Interactions by Coprecipitation
- Unit 20.6 Identification of Protein Interactions by Far Western Analysis
- Unit 20.7 Two-Hybrid Dual Bait System
- Unit 20.8 Interaction Trap/Two-Hybrid System to Identify Loss-of-Interaction Mutant Proteins

CHAPTER 21 CHROMATIN ASSEMBLY AND ANALYSIS

Introduction

- Unit 21.1 Micrococcal Nuclease Analysis of Chromatin Structure
- Unit 21.2 Separation of Histone Variants and Post-Translationally Modified Isoforms by Triton/Acetic Acid/Urea Polyacrylamide Gel Electrophoresis
- Unit 21.3 Chromatin Immunoprecipitation for Determining the Association of Proteins with Specific Genomic Sequences In Vivo
- Unit 21.4 DNase I and Hydroxyl Radical Characterization of Chromatin Complexes
- Unit 21.5 Isolation of Histones and Nucleosome Cores from Mammalian Cells
- Unit 21.6 Assembly of Nucleosomal Templates by Salt Dialysis
- Unit 21.7 Chromatin Assembly Using *Drosophila* Systems
- Unit 21.8 Analysis of Protein Co-Occupancy by Quantitative Sequential Chromatin Immunoprecipitation
- Unit 21.9 Defining In Vivo Targets of Nuclear Proteins by Chromatin Immunoprecipitation and Microarray Analysis
- Unit 21.10 Identifying Chromosomal Targets of DNA-Binding Proteins by Sequence Tag Analysis of Genomic Enrichment (STAGE)
- Unit 21.11 Mapping Chromatin Interactions by Chromosome Conformation Capture
- Unit 21.12 Paired-End diTagging for Transcriptome and Genome Analysis

CHAPTER 22 NUCLEIC ACID ARRAYS

Introduction

- Unit 22.1 Overview of Nucleic Acid Arrays
- Unit 22.2 Preparation of mRNA for Expression Monitoring
- Unit 22.3 Profiling Human Gene Expression with cDNA Microarrays
- Unit 22.4 Overview of mRNA Expression Profiling Using Microarrays
- Unit 22.5 Pattern Discovery in Expression Profiling Data

CHAPTER 23 MANIPULATING THE MOUSE GENOME

Introduction

- Unit 23.1 Overview of Gene Targeting by Homologous Recombination
- Unit 23.2 Mouse Embryo Fibroblast (MEF) Feeder Cell Preparation
- Unit 23.3 Mouse Embryonic Stem (ES) Cell Culture
- Unit 23.4 Mouse Embryonic Stem (ES) Cell Isolation
- Unit 23.5 Production of a Heterozygous Mutant Cell Line by Homologous Recombination (Single Knockout)
- Unit 23.6 Production of a Homozygous Mutant Embryonic Stem Cell Line (Double Knockout)
- Unit 23.7 Chimeric Mouse Production by Microinjection
- Unit 23.8 Mouse Colony Management
- Unit 23.9 Transgenic Mouse Production By Zygote Injection
- Unit 23.10 Transgenic Mouse Colony Management
- Unit 23.11 Modification and Production of BAC Transgenes
- Unit 23.12 Regulation of Transgene Expression Using Tetracycline

CHAPTER 24 GENERATION AND USE OF COMBINATORIAL LIBRARIES

- Unit 24.1 Overview of Agents from Combinatorial Nucleic Acid and Protein Libraries
- Unit 24.2 Design, Synthesis, and Amplification of DNA Pools for Construction of Combinatorial Pools and Libraries
- Unit 24.3 In Vitro Selection of RNA Aptamers to a Protein Target by Filter Immobilization
- Unit 24.4 Peptide Aptamers: Dominant “Genetic” Agents for Forward and Reverse Analysis of Cellular Processes
- Unit 24.5 Protein Selection Using mRNA Display

CHAPTER 25 DISCOVERY AND ANALYSIS OF DIFFERENTIALLY EXPRESSED GENES IN SINGLE CELLS AND CELL POPULATIONS

Introduction

Section A Nucleic Acid Amplification from Individual Cells

- Unit 25A.1 Laser Capture Microdissection
- Unit 25A.2 Preparation of Single Cells from Solid Tissues for Analysis by PCR

Section B Molecular Methods for Discovery of Differentially Expressed Genes

- Unit 25B.1 Production of a Subtracted cDNA Library
- Unit 25B.2 PCR-Based Subtractive cDNA Cloning
- Unit 25B.3 Differential Display of mRNA by PCR
- Unit 25B.4 Restriction-Mediated Differential Display (RMDD)
- Unit 25B.5 AFLP-Based Transcript Profiling
- Unit 25B.6 Serial Analysis of Gene Expression (SAGE)
- Unit 25B.7 Representational Difference Analysis
- Unit 25B.8 Gene Expression Analysis of a Single or Few Cells

CHAPTER 26 GENE SILENCING

Introduction

- Unit 26.1 Overview of RNA Interference and Related Processes
- Unit 26.2 Gene Silencing by RNAi in Mammalian Cells
- Unit 26.3 RNA Interference in *Caenorhabditis Elegans*

- Unit 26.4 Cloning of Small RNA Molecules
- Unit 26.5 RNA Interference in Cultured Drosophila Cells
- Unit 26.6 RNAi in Transgenic Plants

CHAPTER 27 RNA-PROTEIN INTERACTIONS

Introduction

- Unit 27.1 Agarose Gel Separation/Isolation of RNA-Protein Complexes
- Unit 27.2 Identification of RNA Binding Proteins by UV Cross-Linking
- Unit 27.3 Purification of Functional RNA-Protein Complexes using MS2-MBP
- Unit 27.4 RNA Immunoprecipitation for Determining RNA-Protein Associations In Vivo

CHAPTER 28 MAMMALIAN CELL CULTURE

Introduction

- Unit 28.1 Preparation, Culture, and Immortalization of Mouse Embryonic Fibroblasts
- Unit 28.2 Isolation and Immortalization of Lymphocytes
- Unit 28.3 Establishment and Culture of Human Skin Fibroblasts

CHAPTER 29 MOUSE PHENOTYPING

Introduction

Section A General Considerations in Mouse Phenotyping

- Unit 29A.1 Uses of Forward and Reverse Genetics in Mice to Study Gene Function
- Unit 29A.2 Minimizing Variation Due to Genotype and Environment
- Unit 29A.3 Collection of Blood and Plasma from the Mouse

Section B Metabolic Exploration of the Mouse

- Unit 29B.1 Evaluation of Energy Homeostasis
- Unit 29B.2 Lipid and Bile Acid Analysis
- Unit 29B.3 Evaluation of Glucose Homeostasis

Appendix 1 Standard Measurements, Data, and Abbreviations

- 1A Common Abbreviations
- 1B Useful Measurements and Data
- 1C Characteristics of Amino Acids
- 1D Characteristics of Nucleic Acids
- 1E Radioactivity
- 1F Safe Use of Radioisotopes
- 1G Centrifuges and Rotors
- 1H Safe Use of Hazardous Chemicals
- 1I Commonly Used Detergents
- 1J Common Conversion Factors
- 1K Compendium of Drugs Commonly Used in Molecular Biology Research

Appendix 2 Commonly Used Reagents and Equipment

2 Commonly Used Reagents and Equipment

Appendix 3 Commonly Used Techniques in Biochemistry and Molecular Biology

3A Detection and Quantitation of Radiolabeled Proteins and DNA in Gels and Blots

3B Silanizing Glassware

3C Dialysis and Ultrafiltration

3D Quantitation of DNA and RNA with Absorption and Fluorescence Spectroscopy

3E Introduction of Restriction Enzyme Recognition Sequences by Silent Mutation

3F Techniques for Mammalian Cell Tissue Culture

3G Importing Biological Materials

3H Kinetic Assay Methods

3I Statistics for the Molecular Biologist: Group Comparisons

Appendix 4 Suppliers

4 Selected Suppliers of Reagents and Equipment

Appendix 5 Vectors

5 Vectors

CHAPTER 1

Escherichia coli, Plasmids, and Bacteriophages

INTRODUCTION

Mastery of current DNA technology requires familiarity with a small number of basic concepts and techniques. The goal of this chapter is to present this information concisely, yet in enough detail to be useful when a procedure goes wrong. Section I is devoted to *Escherichia coli*. Recipes are provided for media that support *E. coli* growth, as well as instructions for making the simple tools needed to work with bacterial cells. Growth of *E. coli* in liquid and solid media is then detailed. The final unit in Section I describes a few detailed aspects of *E. coli* biology learned from classical bacterial genetic studies, the understanding of which is especially relevant to the techniques used in modern DNA work.

The remainder of the chapter discusses vectors used to introduce foreign DNA into *E. coli*. For the purposes of this chapter, vectors are said to be derived from plasmids, from bacteriophage lambda and related phages, or from filamentous phages. (Many modern vectors incorporate elements from more than one of these classes, and it is likely that this classification scheme will be hopelessly outdated by the time this chapter is revised.) Section II is concerned with plasmid vectors. Following a brief introduction to plasmid biology, procedures are described for purifying small and large amounts of plasmid DNA (“minipreps” and large preps). Finally, procedures for reintroducing plasmid DNA into bacterial cells are described. Section III covers vectors derived from bacteriophages. The biology of bacteriophage lambda is first introduced, followed by detailed aspects of biology that are especially significant when lambda derivatives are used as cloning vectors. Protocols in this section describe techniques for manipulating lambda-derived vectors, making single plaques, making and titering phage stocks, and isolating phage DNA. Finally, Section IV covers the biology and manipulation of vectors derived from filamentous phages.

This chapter will be meaningful primarily to readers with some knowledge of the principles of molecular biology. Several books on molecular biology are recommended in the preface. For further advanced reading in the topics of this chapter, we recommend five books, all from Cold Spring Harbor Laboratory: *Methods in Molecular Genetics* (Miller, 1972), *Advanced Bacterial Genetics* (Davis et al., 1980), *The Bacteriophage Lambda* (Hershey, 1971), *Lambda II* (Hendrix et al., 1983), and *Experiments with Gene Fusions* (Silhavy et al., 1984).

Many terms and jargon used by molecular biologists are introduced in this chapter. These terms are italicized at their first mention, and are defined below.

alpha fragment peptide containing the amino terminus of β -galactosidase, the *lacZ* gene product. Alpha fragments lack enzymatic activity, but can associate with omega fragments (see below) to form proteins whose β -galactosidase activity has been restored.

alpha-complementation restoration of β -galactosidase activity to omega fragments by association with alpha fragments.

amplification increase in copy number of some plasmids that occurs when host protein synthesis is inhibited.

cloning site site on a vector into which foreign DNA is inserted.

competent state in which bacterial or yeast cells are able to take up foreign DNA (for example, as the result of calcium treatment).

cos site of action of phage lambda *ter* function. Cos site is cut by *ter* to yield two cohesive ends (cos ends).

dilution, 10^x-fold a solution or suspension that contains 1/10^x as much (10^{-x} as much) of the dissolved or suspended species as does the starting liquid. For example, to “do a 10²-fold dilution” is to dilute a solution 100×.

early-log phase period during the growth of a culture after the lag period. During early log phase growth, cells have begun exponential growth.

efficiency of plating (EOP) titer of bacterial colonies or phage plaques under some experimental conditions divided by the titer of bacteria or phage obtained by growth on some reference medium.

exponential growth period during which the number of cells in the culture increases as an exponential function of time, that is, during which cell number = ke^t .

F factor genetic element found in some strains of *E. coli* and related species. F encodes proteins used in formation of sex pili which allow its transfer from bacterium to bacterium.

female strain strain that does not contain the F factor and that receives genetic information when crossed with a strain containing F.

helper phages bacteriophages that encode essential proteins and that allow other phages which do not encode these essential proteins to grow.

incompatible phenomenon in which two plasmids cannot replicate in the same cell without continual selection for both of them.

incompatibility group consists of plasmids that cannot be maintained together in the same cell. Compatible plasmids belong to different incompatibility groups.

induction the onset of transcription of a new gene or operon, usually in response to some environmental stimulus. Phage induction or lysogenic induction describe the process in which prophage excise from the chromosome of bacteria that harbor them and begin to grow lytically.

inoculation introduction of cells into a container of sterile growth medium.

lag period period just after inoculation of a culture when cells have not yet begun to grow exponentially.

late-log phase last period of exponential growth of a culture, after which growth slows and then stops altogether due to nutrient exhaustion or accumulation of waste products.

lawn uniform layer of bacteria that covers the surface of a plate.

log phase period during growth of a culture in which cells are growing exponentially.

low-copy-number plasmids plasmids found in less than about 20 copies per cell when cells containing them are grown in rich medium.

lysogen *E. coli* cell or strain that harbors a dormant bacteriophage.

male strain strain of bacteria that contains the F factor.

male-specific phages bacteriophages that only grow on male strains because they adsorb to sex pili.

marker detectable genetic difference between one organism and another (usually wild-type) organism of the same species.

minimal medium growth medium for cells that contains only salts, vitamins, trace elements, and simple compounds which serve as carbon, nitrogen, and phosphorous sources.

miniprep small-scale preparation or purification of some desired species, usually of plasmid or phage DNA.

mobilization transfer of DNA from one cell to another caused by a mobile genetic element such as the F factor.

multiplicity of infection (MOI) ratio of infecting bacteriophage to host cells.

nonsense suppression the insertion of amino acids into proteins at positions where translation would normally not occur because the mRNA contains a UAG (amber), UAA (ochre) or UGA nonsense codon.

nonsense suppressor tRNA that inserts amino acids at nonsense codons. The term is sometimes used for the genes encoding these tRNAs.

omega fragment protein containing the carboxy-terminal fragment of β-galactosidase. This protein lacks enzymatic activity, but β-galactosidase activity can be restored when the peptide is complexed with an alpha fragment.

ori (origin) site on genome at which DNA replication begins.

outgrowth the growth of freshly transformed cells under nonselective conditions for enough time to allow proteins encoded by the foreign DNA to be expressed.

overnight a small, freshly saturated liquid culture of bacteria.

packaging extract extract from special strains of *E. coli* that contains bacteriophage lambda head proteins, tail proteins, and packaging proteins. Phage DNA added to such an extract is assembled into phage particles.

par site on some plasmids which ensures that each daughter cell receives a plasmid copy.

pilot protein protein in the coat of filamentous phages that helps phage DNA enter the cell.

plates petri dishes filled with solid medium, used to grow separated bacterial colonies or plaques. The term is sometimes used to refer to 96-well microtiter dishes.

plating out the placement of bacteria or phage on plates so that colonies or plaques are formed.

polylinker stretch of DNA that contains contiguous restriction sites.

prophage dormant bacteriophage, usually integrated into the host chromosome, that replicates with the host bacterium.

relaxed control applies to plasmids whose replication does not depend on the bacterial cell cycle.

replicative form double-stranded circular filamentous phage DNA found inside infected cells.

replicator stretch of DNA on a phage or plasmid that enables the phage or plasmid to replicate.

rich medium growth medium that contains complex organic molecules (peptides, nucleotides, etc.). Typical components of rich media include tryptone (made from beef) and yeast extract (made from yeast).

rolling-circle replication mechanism of replication sometimes used by circular molecules in which DNA polymerase continually circumnavigates the template, and thus synthesizes a long tail.

satellite colonies small colonies that grow around a large colony on a plate containing selective medium. These are usually composed of cells unable to grow on selective medium, but which are able to grow near the large colony because the cells in the large colony neutralize the selective agent.

saturated culture culture of cells in liquid medium that has stopped growing because nutri-

ents are exhausted or because waste products have accumulated.

SOS response response of *E. coli* to DNA damage or other treatments that inhibit DNA replication. Lambda-derived phages are induced during this response.

stringent control applies to plasmids whose replication is synchronized with the *E. coli* cell cycle.

temperate describes bacteriophages capable of lysogenic growth.

transfection introduction of bacteriophage DNA into competent *E. coli* cells. Also describes the introduction of any DNA (including plasmid DNA) into cells of higher eukaryotes.

transformation introduction of plasmid DNA into *E. coli* or yeast. Also used to denote any of a number of changes in cultured higher eukaryotic cells to characteristics more typical of cancer cells (immortal growth, loss of contact inhibition, etc.).

LITERATURE CITED

- Davis, R., Botstein, D., and Roth, J.R. 1980. Advanced Bacterial Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Hendrix, R., Roberts, J., Stahl, F., and Weisberg, R. 1983. Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Hershey, A.D. 1971. The Bacteriophage Lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Miller, J. 1972. Methods in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Silhavy, T., Berman, M.L., and Enquist, L.W. 1984. Experiments with Gene Fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

ESCHERICHIA COLI

SECTION I

Escherichia coli is a rod-shaped bacterium with a circular chromosome about 3 million base pairs (bp) long. It can grow rapidly on *minimal medium* that contains a carbon compound such as glucose (which serves both as a carbon source and an energy source) and salts which supply nitrogen, phosphorus, and trace metals. *E. coli* grows more rapidly, however, on a *rich medium* that provides the cells with amino acids, nucleotide precursors, vitamins, and other metabolites that the cell would otherwise have to synthesize. The purpose of this first section is to provide basic information necessary to grow *E. coli*. A more detailed introduction to certain aspects of *E. coli* biology may be found in UNIT 1.4.

When *E. coli* is grown in liquid culture, a small number of cells are first *inoculated* into a container of sterile medium. After a period of time, called the *lag period*, the bacteria begin to divide. In rich medium a culture of a typical strain will double in number every 20 or 30 min. This phase of *exponential growth* of the cells in the culture is called *log phase* (sometimes subdivided into *early-log*, *middle-log*, and *late-log phases*). Eventually the cell density increases to a point at which nutrients or oxygen become depleted from the medium, or at which waste products (such as acids) from the cells have built up to a concentration that inhibits rapid growth. At this point, which, under normal laboratory conditions, occurs when the culture reaches a density of $1-2 \times 10^9$ cells/ml, the cells stop dividing rapidly. This phase is called *saturation* and a culture that has just reached this density is said to be freshly saturated.

With very few exceptions, bacterial strains used in recombinant DNA work are derivatives of *E. coli* strain K-12. Most advances in molecular biology until the end of the 1960s came from studies of this organism and of bacteriophages and plasmids that use it as a host. Much of the cloning technology in current use exploits facts learned during this period.

Media Preparation and Bacteriological Tools

UNIT 1.1

Recipes are provided below for minimal liquid media, rich liquid media, solid media, top agar, and stab agar. Tryptone, yeast extract, agar (Bacto-agar), nutrient broth, and Casamino Acids are from Difco. NZ Amine A is from Hunko Sheffield (Kraft).

MINIMAL MEDIA

Ingredients for these media should be added to water in a 2-liter flask and heated with stirring until dissolved. The medium should then be poured into separate bottles with loosened caps and autoclaved at 15 lb/in² for 15 min. Do not add nutritional supplements or antibiotics to any medium until it has cooled to <50°C. After the bottles cool to below 40°C, the caps can be tightened and the concentrated medium stored indefinitely at room temperature. All recipes are on a per liter basis.

M9 medium, 5×

- 30 g Na₂HPO₄
- 15 g KH₂PO₄
- 5 g NH₄Cl
- 2.5 g NaCl
- 15 mg CaCl₂ (optional)

Escherichia coli, Plasmids, and Bacteriophages

1.1.1

M63 medium, 5×

10 g (NH₄)₂SO₄
68 g KH₂PO₄
2.5 mg FeSO₄·7H₂O
Adjust to pH 7 with KOH

A medium, 5×

5 g (NH₄)₂SO₄
22.5 g KH₂PO₄
52.5 g K₂HPO₄
2.5 g sodium citrate·2H₂O

Before they are used, concentrated media should be diluted to 1× with sterile water and the following sterile solutions, per liter:

1 ml 1 M MgSO₄·7H₂O
10 ml 20% carbon source (sugar or glycerol)
and, if required:
0.1 ml 0.5% vitamin B1 (thiamine)
5 ml 20% Casamino Acids *or*
L amino acids to 40 μg/ml *or*
DL amino acids to 80 μg/ml
Antibiotic (see Table 1.4.1)

RICH MEDIA

Unless otherwise specified, rich media should be autoclaved for 25 min. Antibiotics and nutritional supplements should be added only after the solution has cooled to 50°C or below. A flask containing liquid at 50°C feels hot but can be held continuously in one's bare hands. All recipes are on a per liter basis.

H medium

10 g tryptone
8 g NaCl

Lambda broth

10 g tryptone
2.5 g NaCl

LB medium

10 g tryptone
5 g yeast extract
5 g NaCl
1 ml 1 N NaOH

The original recipe for LB medium (sometimes referred to as Luria or Lenox broth), does not contain NaOH. There are many different recipes for LB that differ only in the amount of NaOH added. We use this formula in our own work. Even though the pH is adjusted to near 7 with NaOH, the medium is not very highly buffered, and the pH of a culture growing in it drops as it nears saturation.

NZC broth

10 g NZ Amine A
5 g NaCl
2 g MgCl₂·6H₂O
Autoclave 30 min
5 ml 20% Casamino Acids

Superbroth

32 g tryptone
20 g yeast extract
5 g NaCl
5 ml 1 N NaOH

TB (terrific broth)

12 g Bacto tryptone
24 g Bacto yeast extract
4 ml glycerol

Add H₂O to 900 ml and autoclave, then add to above sterile solution 100 ml of a sterile solution of 0.17 M KH₂PO₄ and 0.72 M K₂HPO₄.

Tryptone broth

10 g tryptone
5 g NaCl

2× TY medium

16 g tryptone
10 g yeast extract
5 g NaCl

TYGPN medium

20 g tryptone
10 g yeast extract
10 ml 80% glycerol
5 g Na₂HPO₄
10 g KNO₃

SOLID MEDIA

Liquid media can be solidified with agar. For minimal plates, dissolve the agar in water and autoclave separately from the minimal medium; autoclaving the two together will give rise to an insoluble precipitate. For rich plates, autoclave the agar together with the other ingredients of the medium. Cool the agar to about 50°C and add other ingredients if necessary. At this temperature, the medium will stay liquid indefinitely, but it will rapidly solidify if its temperature falls much below 45°C. Finally, pour the medium into sterile disposable petri dishes (*plates*) and allow to solidify.

Freshly poured plates are wet and unable to absorb liquid spread onto them. Moreover, plates that are even slightly wet tend to exude moisture underneath bacteria streaked on them, which can cause the freshly streaked bacteria to float away. So for most applications, dry the plates by leaving them out at room temperature for 2 or 3 days, or by leaving them with the lids off for 30 min in a 37°C incubator or in a laminar flow hood. Store dry plates at 4°C, wrapped in the original bags used to package the empty plates. Plates should be inverted when incubated or stored.

Minimal Plates

Autoclave 15 g agar in 800 ml water for 15 min. Add sterile concentrated minimal medium and carbon source. After medium has cooled to about 50°C, add supplements and antibiotics. Pouring 32 to 40 ml medium into each plate, expect about 25 to 30 plates per liter.

Rich Plates

To ingredients listed below, add water to 1 liter and autoclave 25 min. Pour LB and H plates with 32 to 40 ml medium, in order to get 25 to 30 plates per liter. Pour lambda plates with about 45 ml medium for about 20 plates per liter.

H plates

10 g tryptone
8 g NaCl
15 g agar

Lambda plates

10 g tryptone
2.5 g NaCl
10 g agar

LB plates

10 g tryptone
5 g yeast extract
5 g NaCl
1 ml 1 N NaOH
15 g agar or agarose

Additives

Antibiotics (if required):

Ampicillin to 50 µg/ml
Tetracycline to 12 µg/ml
Other antibiotics, see Table 1.4.1

Galactosides (if required):

Xgal to 20 µg/ml
IPTG to 0.1 mM
Other galactosides, see Table 1.4.2

TOP AGAR

Top agar is used to distribute phage or cells evenly in a thin layer over the surface of a plate. In a typical application, molten top agar is mixed with bacteria and the mixture poured onto a plate to make a thin layer that is allowed to solidify. This layer of cells then grows denser, forming the opaque *lawn* of cells. Top agar contains less agar than plates, and so stays molten for days when it is kept at 45° to 50°C. Top agarose is sometimes used when DNA is to be prepared directly from phage, and is also used when libraries are plated out to be screened by plaque lifting.

Prepare top agar in 1-liter batches, autoclave for 15 min to melt, cool to 50°C, swirl to mix, pour into separate 100-ml bottles, reautoclave, cool, and store at room temperature. Before use, melt the agar by heating in a water bath or microwave oven (see *UNIT 1.11*) then cool to and hold at 45° to 50°C.

H top agar

10 g tryptone
8 g NaCl
7 g agar

LB top agar

10 g tryptone
5 g yeast extract
5 g NaCl
7 g agar

Lambda top agar

10 g tryptone
2.5 g NaCl
7 g agar

Top agarose

10 g tryptone
8 g NaCl
6 g agarose

STAB AGAR

Stab agar is used for storing bacterial strains (see *UNIT 1.3*). The recipe below is for 1 liter.

Stab agar

10 g nutrient broth
5 g NaCl
6 g agar
10 mg cysteine·Cl
10 mg thymine

Thymine is included so that thy⁻ bacteria can grow. Cysteine is thought to increase the amount of time bacteria can survive in stabs.

TOOLS

Inoculating Loops

Inoculating loops are used to move small numbers of bacteria or phage to a plate or to a new container of liquid medium. Inoculating loops may be purchased from any general scientific supply company. However, most researchers prefer to use loops made in the laboratory. These are made by inserting both ends of a 10-in. piece of 28-G platinum wire into an inoculating loop holder (also widely available) and twirling the holder while tugging on the middle of the wire with the point of a pencil (see Fig. 1.1.1).

Sterilize the loop by holding it in a bunsen burner flame until it is red hot. Cool the loop by touching it to a sterile portion of the surface of an agar plate until it stops sizzling.

Sterile Toothpicks

The broad side of flat wooden toothpicks may also be used for streaking out bacteria. Round wooden toothpicks, or the pointed end of flat toothpicks, are sometimes used to pick individual colonies or phage plaques. To sterilize, place toothpicks in a small beaker, cover the beaker with foil, and autoclave. Alternatively, simply autoclave the whole box of toothpicks and hold them in the middle when picking them up out of the opened box. It is convenient to put used toothpicks into another smaller beaker which, when full, is covered with foil and autoclaved. Used toothpicks can be saved, reautoclaved, and used again (see Fig. 1.1.2).

Spreaders

Spreaders are used to distribute liquid containing bacterial cells evenly over a plate. They are made by heating and bending a piece of 4-mm glass tubing (see Fig. 1.1.3). Less durable spreaders can be made from a Pasteur pipet. Before each use, sterilize the spreader by dipping the triangular part into a container of ethanol, passing the spreader through a gas flame to ignite the ethanol, and letting the flame burn out. Be careful not to ignite the

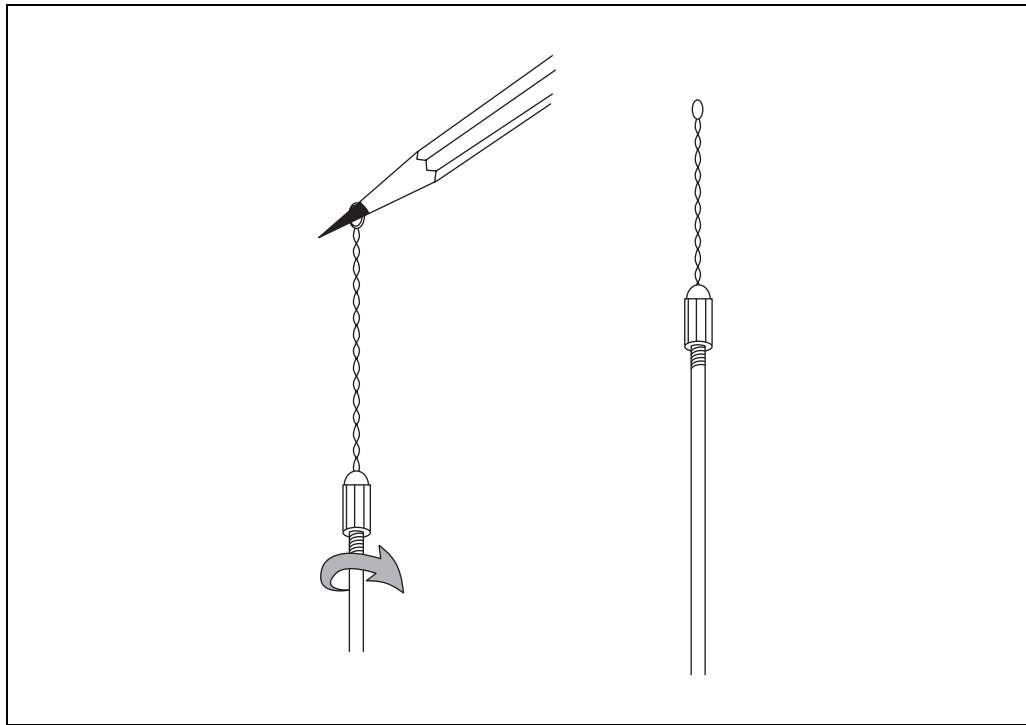


Figure 1.1.1 Making a loop.

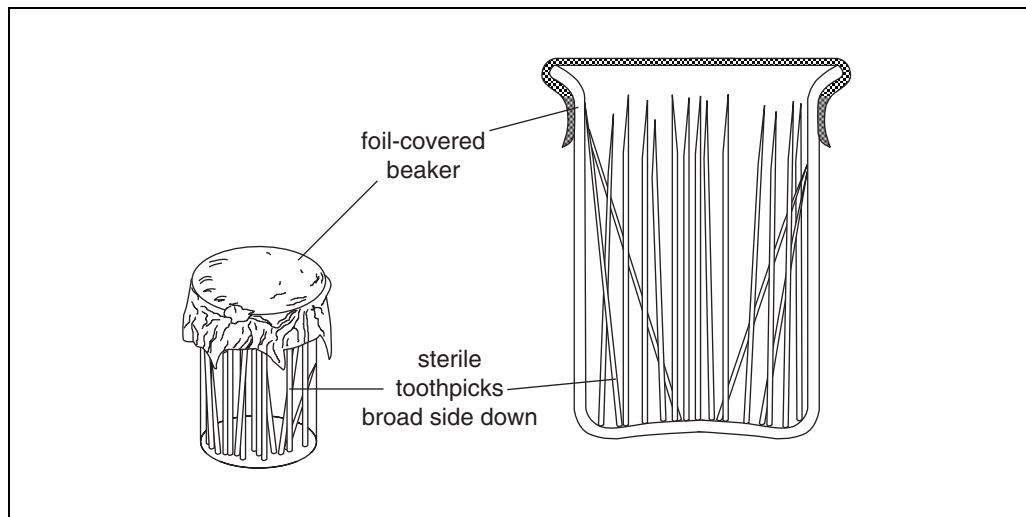


Figure 1.1.2 Preparing sterile toothpicks.

ethanol in the container. Cool the spreader by touching it to the surface of an agar plate that has not yet been spread with cells.

Glass Beads

Although spreaders are useful for many applications, when processing large numbers of plates it can become a time-consuming process. A popular practice is to use 4-mm glass beads that have been sterilized. A half dozen or more beads distribute the liquid culture on surface of the agar when the plate is shaken horizontally in all directions. The beads are then discarded and the plate inverted and placed in the incubator. Stacks of plates can be handled together when plating many culture samples.

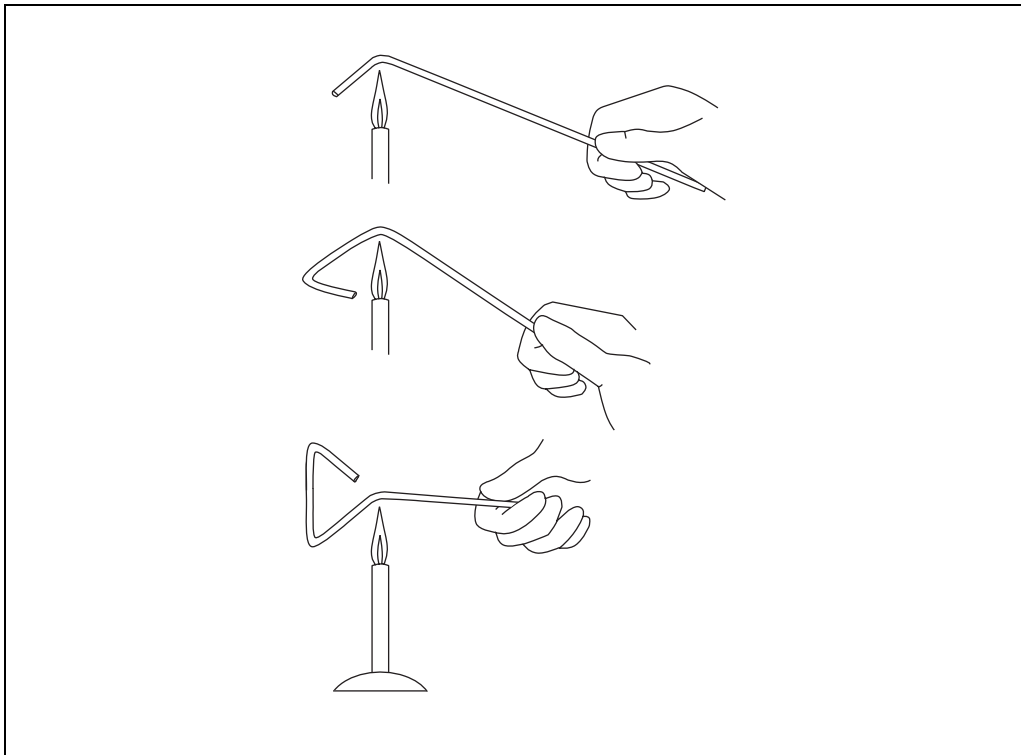


Figure 1.1.3 Making a spreader.

Contributed by Karen Elbing
Clark & Elbing LLP
Boston, Massachusetts

Roger Brent
The Molecular Sciences Institute
Berkeley, California

Growth in Liquid Media

UNIT 1.2

GROWING AN OVERNIGHT CULTURE

BASIC
PROTOCOL 1

Small freshly saturated cultures of *E. coli* are called *overnights*. To make an overnight, remove the cap from a sterile 16- or 18-mm culture tube. Working quickly to minimize contact of the tube with the possibly contaminated air, use a sterile pipet to transfer 5 ml of liquid medium into the tube. Inoculate the liquid with a single bacterial colony by touching a sterile inoculating loop to the colony, making certain that some of the cells have been transferred to the loop, and then dipping the loop into the liquid and shaking it a bit. Replace the tube's cap, and place the tube on a roller drum at 60 rpm, 37°C. Grow until the culture is freshly saturated (at a density of $1\text{--}2 \times 10^9$ cells/ml, which typically takes at least 6 hr).

GROWING LARGER CULTURES

BASIC
PROTOCOL 2

Larger cultures are generally inoculated with overnight cultures diluted 1:100. Use an Erlenmeyer or baffle flask whose volume is at least 5 times the volume of the culture. Grow the culture at 37°C with vigorous agitation (~300 rpm) to ensure proper aeration. If it is necessary to grow a culture without shaking (for example, if the strain is temperature-sensitive for growth and no low-temperature shaker is available), then, to ensure that the cells get adequate aeration, grow the culture in an Erlenmeyer flask whose volume is at least 20 times that of the culture.

MONITORING GROWTH

BASIC
PROTOCOL 3

With a Count Slide

Take a clean count slide (or hemacytometer) and cover it with a clean cover slip. Dip a 0.1- or 1-ml pipet into the culture medium, allow a small drop of liquid to form on the end of the pipet, and touch it lightly to the surface of the slide at the periphery of the cover slip. The liquid will quickly spread under the cover slip. Put the slide on the stage of a phase-contrast microscope set to 400 \times , and focus on the cells. Each cell in a small square is equivalent to 2×10^7 cells/ml (see Fig. 1.2.1).

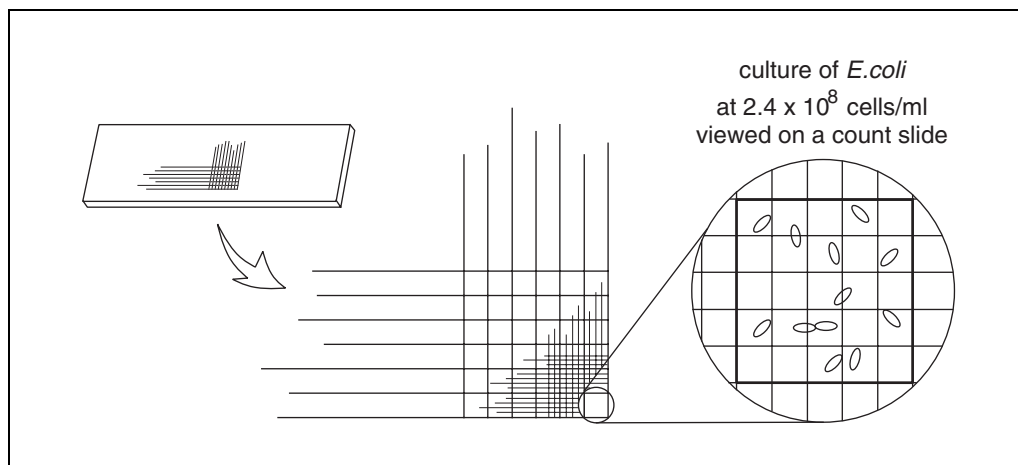


Figure 1.2.1 Monitoring growth via a hemacytometer or count slide.

Escherichia coli,
Plasmids, and
Bacteriophages

With a Spectrophotometer

The concentration of cells in a culture can also be determined with a spectrophotometer by measuring the amount of 600-nm light scattered by the culture. The level of absorbance (A) at 600 nm will depend on the distance between the cuvette and the detector and will vary among spectrophotometers, often by a factor of 2. It is thus wise to calibrate each instrument by recording the OD_{600} (sometimes expressed as A_{600}) of a culture that contains a known number of cells determined by some other method, such as observation on a count slide or titering for viable colonies (*UNIT 1.3*).

If the culture is visibly turbid, also measure a *10-fold dilution* of it. For a culture grown in rich medium, a good rule of thumb is that each 0.1 OD unit is roughly equivalent to 1×10^8 cells/ml.

Calculate the number of cells/ml from whichever suspension (the undiluted or the diluted) has an $OD_{600} < 1$.

Contributed by Karen Elbing
Clark & Elbing LLP
Boston, Massachusetts

Roger Brent
The Molecular Sciences Institute
Berkeley, California

Growth on Solid Media

UNIT 1.3

TITERING AND ISOLATING BACTERIAL COLONIES BY SERIAL DILUTIONS

BASIC
PROTOCOL 1

Bacteria are grown from single colonies to ensure that each cell in a population is descended from a single founder cell, and thus to help ensure that each cell in the culture has the same genetic makeup. One way to generate single colonies is to titer a culture with serial dilutions and to pick colonies from one of the dilution plates. In this procedure, a small, measured amount of a bacterial culture is diluted into fresh liquid in another tube. A small amount of liquid is taken from this tube and diluted into another fresh tube. This process is repeated several times. Equal volumes of liquid are then taken from each of the dilution tubes and plated on petri plates. The plates are incubated overnight at 37°C; well-separated single colonies will arise on some of the dilution plates. The number of living bacteria in the culture is calculated from the number of colonies formed on the dilution plates.

A typical saturated culture contains 1×10^9 cells/ml. Phage suspensions can also be titered; a concentrated phage stock might typically contain 1×10^{11} phage/ml.

Materials

LB medium (UNIT 1.1)

LB plates (UNIT 1.1)

Sterile 16- or 18-mm-diameter culture tubes

1. Use pipets to introduce 5 ml LB medium into three sterile culture tubes. Line the tubes up, or label them so that they can be distinguished.
2. Using a pipettor, transfer 5 μ l from the suspension of cells into the first tube of LB medium. Set the vortexer to a mild setting and agitate the tube 5 sec.
3. Put a new tip on the pipettor and transfer 5 μ l from the first tube of LB medium into the second tube, and vortex the second tube. Take 5 μ l from the second tube and repeat step 3 until you have serial dilutions in all three tubes.

The first dilution tube now contains a 1×10^3 -fold dilution, generated by diluting the culture by a factor of one thousand (i.e., it contains 1×10^{-3} as many cells/ml as were present in the original culture). The second tube contains a 1×10^6 -fold dilution, generated by diluting the original culture by a factor of one million (i.e., it contains 1×10^{-6} as many cells/ml as the original culture), etc.

Many investigators prefer to perform serial dilutions with different volumes and different factors of dilution. These parameters can be modified in steps 1 to 3.

4. Spread 100 μ l of liquid from the culture and from each dilution tube onto separate, labeled, dry LB plates (as described on p. 1.3.2). Incubate overnight at 37°C.

During this incubation, each living bacterial cell will grow into a separate colony on the plate.

5. Count the colonies from these plates. Since only 100 μ l was plated from the undiluted culture and from each dilution tube, each plate has $1/10$ as many colonies on it as were present in each milliliter of liquid in the corresponding tube. Therefore, one can determine the number of cells that were present per milliliter of the culture by counting the number of colonies on a plate, and then multiplying that number by 10 times the factor of dilution.

Escherichia coli,
Plasmids, and
Bacteriophages

1.3.1

For example, if 22 colonies were observed on the plate corresponding to the 1×10^6 -fold dilution, then the number of living cells in each milliliter of the original culture was $22 \times 10 \times 10^6$, or 2.2×10^8 cells/ml.

- Any of the single colonies may be saved for further use. Store plates at 4°C wrapped in parafilm or in the plastic sleeve in which the plates were supplied.

Commentary

Titering by serial dilutions is a good way to determine the number of any kind of living organism present in a suspension. The organisms do not even need to be able to grow into colonies—i.e., the concentration of living bacteriophage in a tube can be determined by titering with serial dilutions and counting the number of plaques made when an aliquot of each dilution is plated on a lawn of phage-sensitive bacteria (see *UNIT 1.11*).

It is sometimes useful to use smaller factors of dilution. Mixing $50 \mu\text{l}$ of the culture into 5 ml medium will give dilutions of $100\times$. Mixing $100 \mu\text{l}$ into $900 \mu\text{l}$ will give dilutions of $10\times$.

BASIC PROTOCOL 2

ISOLATING SINGLE COLONIES BY STREAKING A PLATE

Another way to isolate single colonies is called streaking or streaking for single colonies. This method is easier and faster than serial dilutions for isolating single colonies, but it cannot be used to count the number of cells in a culture. An inoculum of bacteria is streaked across one side of an agar plate with an inoculating loop or sterile toothpick. The resterilized loop or a fresh toothpick is then passed once through the first streak and streaked across a fresh part of the plate (see Fig. 1.3.1). This process is repeated at least once more, and the plate is incubated at 37°C until colonies become visible. If single colonies must be isolated from many bacteria, it is convenient to divide a plate into 4, 6, or 8 sectors and to streak for single colonies in each sector.

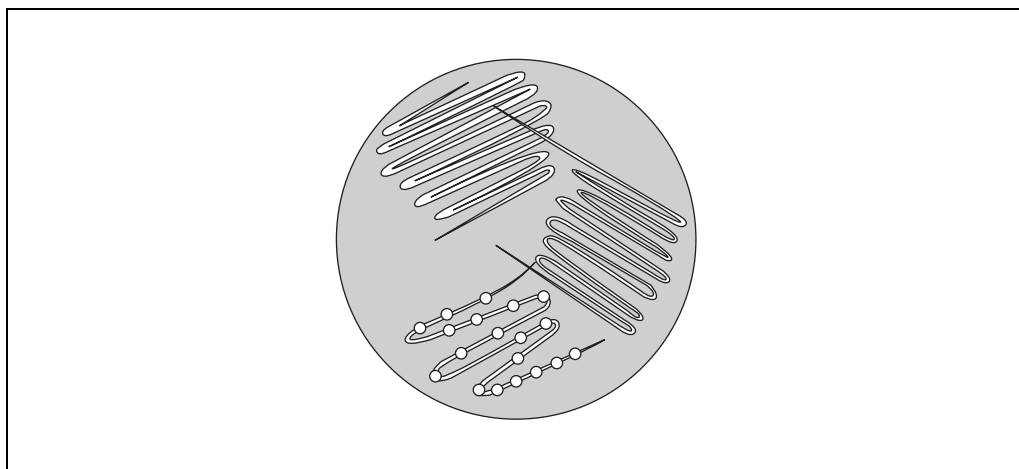


Figure 1.3.1 Streaking a plate.

ISOLATING SINGLE COLONIES BY SPREADING A PLATE

It is sometimes necessary to distribute a liquid culture of bacteria evenly over the surface of a plate (for example, when plasmid-containing colonies are to be isolated after treatment of cells with plasmid DNA and calcium chloride, *UNIT 1.8*). This is usually done with a glass spreader. From 0.05 to 1 ml of liquid is pipetted onto a dry plate (see *UNIT 1.1*) and spread using a circular motion as shown in Figure 1.3.2. Alternatively, the edge of the spreader can be used to make a raster pattern on the plate's surface. The plate can be turned at right angles and the process repeated. Evenly spread plates should be placed in the incubator with the lids ajar until they are completely dry.

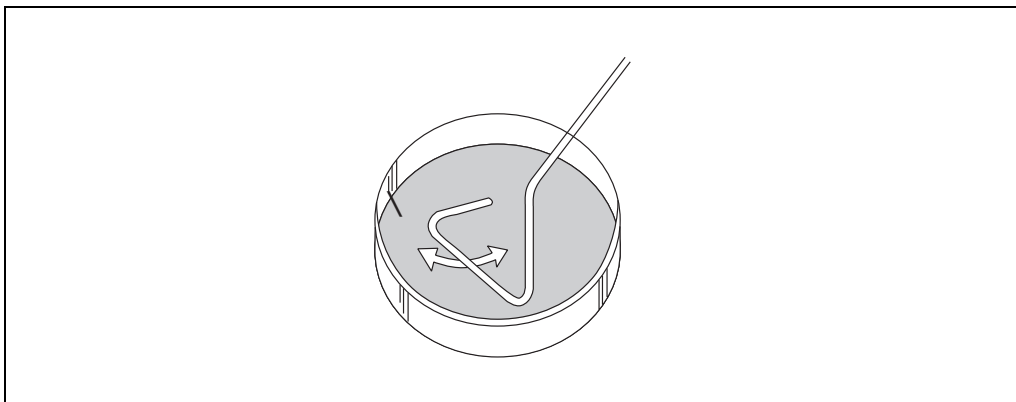


Figure 1.3.2 Spreading a plate.

REPLICA PLATING

Replica plating is a convenient way to test many colonies for their ability to grow under different conditions. In this technique, bacterial colonies are transferred from one plate to another in a way that maintains the original pattern of colonies. This technique has many applications to recombinant DNA work. As an example, consider the plasmid pBR322, which contains two antibiotic resistance genes, encoding resistance to ampicillin and tetracycline (see Fig. 1.5.9). A piece of foreign DNA cloned into the tetracycline resistance gene inactivates it; cells that carry such a plasmid are ampicillin resistant but tetracycline sensitive. These cells can be identified by replica plating colonies from ampicillin-containing master plates onto plates containing tetracycline. Tetracycline-sensitive colonies can be identified by their inability to grow on the tetracycline plates, rescued from the master plate, and analyzed further.

This procedure requires two specialty items: a replica block and sterile velvets. The replica block is a wooden or metal cylinder that fits snugly inside a petri plate (see Fig. 1.3.3). One method for constructing these has been described by Adams (1965). A metal ring is used to secure the velvets to the block. Squares of velvet should be cut so as to cover the base (a diameter of 14 cm is suggested). These velvets can be washed, autoclaved, and reused. If velvets are not available, pieces of sterile filter paper or disposable replica plates can be used ("Repli-Plate" Colony Transfer Pads, American Laboratories #59901). Replica plating also requires a master plate composed of well-separated colonies. The master plate can be a fresh plate onto which 50 to 100 colonies have been gridded (using toothpicks and the grid in Fig. 1.3.4), or it can be a plate on which bacteria were spread that have now grown up into well-separated colonies.

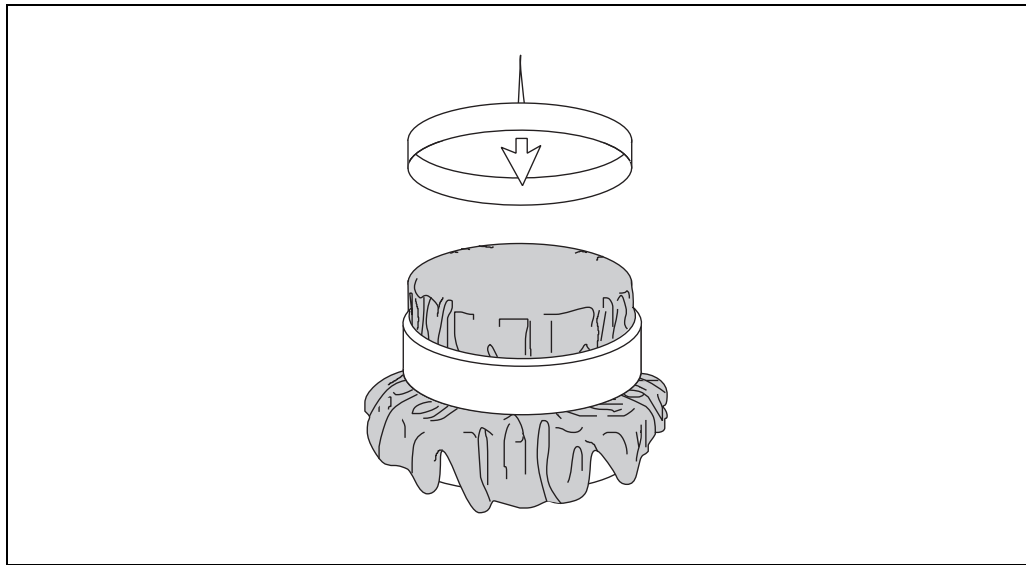


Figure 1.3.3 Replica block with velvet.

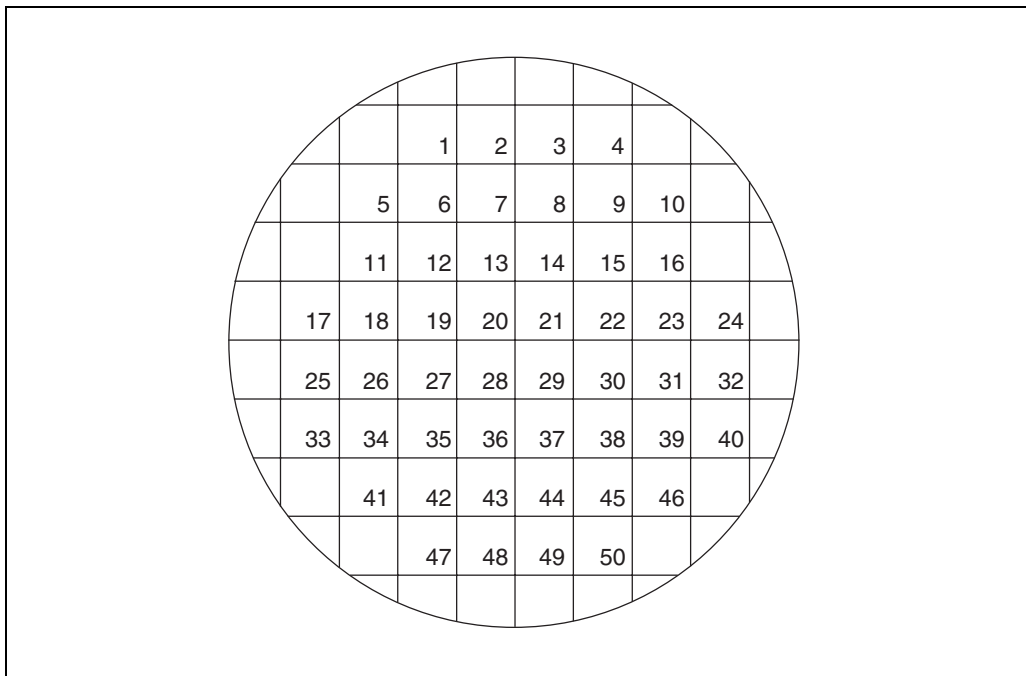


Figure 1.3.4 Grid for replica plating.

Mark the top of the master plate to enable alignment with the grid. Press the plate down *lightly* onto the velvet. Do not bear down hard on the plate; pressing too hard will cause the colonies to run together on the velvet or may even cause the plate to collapse. Press new plates, oriented like the master plate, *lightly* onto the imprinted velvet to transfer the colonies. As many as 10 plates per velvet can often be replica plated.

**SUPPORT
PROTOCOL 2**

STRAIN STORAGE AND REVIVAL

Most strains of *E. coli* can be stored for years in stab vials, or indefinitely if frozen at -70°C . It is prudent to check the genetic markers of a strain revived from storage. Ways to verify the presence of other selective markers are described in Table 1.4.4.

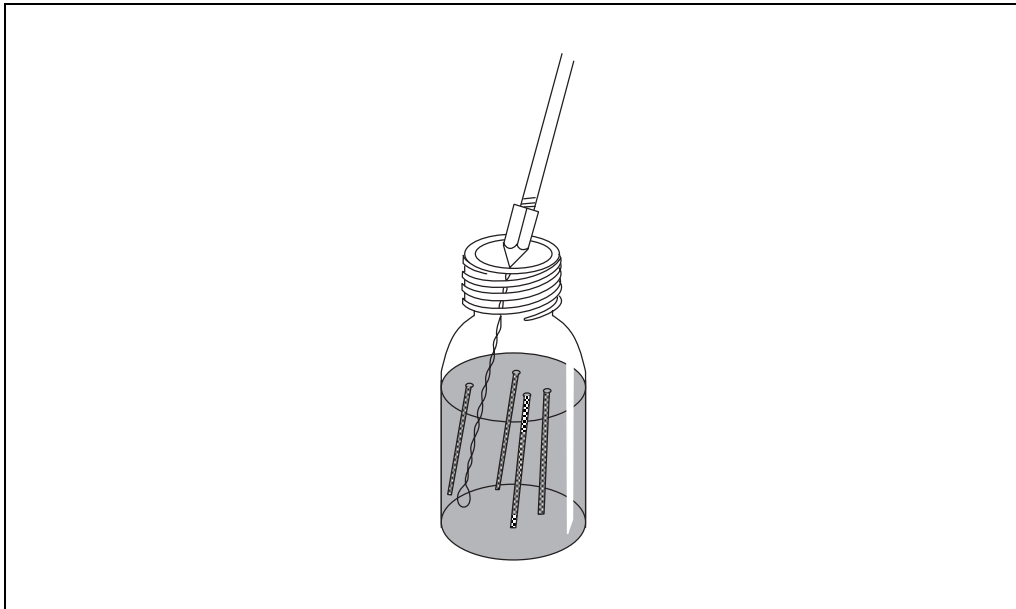


Figure 1.3.5 Inoculating a stab vial.

Stabs

Use airtight, autoclavable vials with rubber or Teflon caps (not cardboard). These are available from Wheaton Glassware and John's Scientific ($\frac{1}{4}$ -oz. Bijoux bottles, #15690-001). Fill the vials $\frac{2}{3}$ full with stab agar (*UNIT 1.1*). Inoculate them with a single colony (see Fig. 1.3.5) by collecting most of the cells in the colony with an inoculating loop, then repeatedly poking the loop deeply into the agar. Leave the cap of the stab vial slightly loose and incubate 8 to 12 hr at 37°C, or until cloudy tracks of bacterial growth are evident. Seal the vials tightly and store them in a cool (15° to 22°C), dark place. To revive a stored strain, flame sterilize an inoculating loop (*UNIT 1.1*), allow it to cool, insert it into the stab agar, and move the loop around until a gobbet of bacteria-laden agar is stuck onto the loop. Smear the gobbet onto one section of an LB plate and streak for single colonies (Fig. 1.3.1).

Frozen Stocks

Add 2 ml of a mid-log culture or 1 ml of a freshly saturated culture to a stab vial or a Nunc vial (Nunc #1087) containing 1 ml glycerol solution or DMSO solution (see recipes). Vials can be stored at -20° to -70°C , but most strains remain viable longer if stored at -70°C . Revive stored cells by scraping off splinters of solid ice with a toothpick or sterile pipet and streaking these splinters onto an LB plate (*UNIT 1.1*). Do not allow the contents of the vial to thaw.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

Glycerol solution

65% glycerol (vol/vol)
0.1 M MgSO_4
0.025 M Tris-Cl, pH 8

DMSO solution

7% dimethylsulfoxide (vol/vol)

The only advantage DMSO seems to have over glycerol for frozen stocks is that it is easier to pipet because it is less viscous. Use a bottle of reagent- or spectrophotometric-grade DMSO that has been kept tightly sealed.

LITERATURE CITED

Adams, J.N. 1965. Automotive pistons for use as bases in velvetreen replication. *J. Bacteriol.* 89:1627.

KEY REFERENCE

Lederberg, J. and Tatum, E.L. 1953. Novel genotypes in mixed cultures of biochemical mutants of bacteria. *Cold Spring Harbor Symp. Quant. Biol.* 18:75.

A primary reference in this field of research.

Contributed by Karen Elbing
Clark & Elbing LLP
Boston, Massachusetts

Roger Brent
The Molecular Sciences Institute
Berkeley, California

Selected Topics from Classical Bacterial Genetics

Current cloning technology exploits many facts learned from classical bacterial genetics. This unit covers those that are critical to understanding the techniques described in this book.

ANTIBIOTICS

Antibiotics are chemicals that kill microorganisms but are relatively nontoxic to eukaryotic organisms. Antibiotics are very important for the techniques described in this book; genes encoding resistance to them are carried on plasmid and phage vectors and cells that contain the vector are identified by their ability to grow and form colonies in the presence of the antibiotic. Table 1.4.1 gives stock and working concentrations, and mechanisms of action of most of the antibiotics that are used in recombinant DNA work.

Antibiotics are usually added to freshly autoclaved solid medium after it has cooled to below 50°C. In an emergency, antibiotics can be added directly to existing plates using the same final concentration as above (assume that a plate 100 mm in diameter contains a total medium volume of 30 ml). Allow the antibiotic time to diffuse away from the very surface of the plate; an hour is usually sufficient. Since many antibiotics (especially ampicillin) lose potency at room temperature, plates are usually stored at 4°C. In addition, rifampicin and tetracycline should be stored in the dark (see Table 1.4.1).

THE LAC OPERON

Many of the techniques described in this book were made possible by early studies of the *E. coli* *lac* operon. The *lac* operon consists of three genes—*lacZ*, *lacY*, and *lacA* (see Fig. 1.4.1). When the cell grows on rich medium or glucose minimal medium, transcription is blocked by *lac* repressor (product of the neighboring *lacI* gene) which binds to a single site (operator) upstream of *lacZ* and prevents RNA polymerase from binding to the promoter. When the cell grows on medium that contains lactose or certain related compounds (see Table 1.4.2), *lac* repressor no longer binds the operator, and RNA polymerase synthesizes a single mRNA which encodes *lacZ*, *lacY*, and *lacA*. (In the wild-type *lac* operon, transcription initiation also requires the presence of a cAMP-CRP activator complex; all *lac* promoters used in

cloning experiments are independent of this control.) Two of these genes are necessary for growth on lactose. *lacY* encodes a permease which is necessary for the uptake of lactose and certain related sugars. *lacZ* encodes a β -galactosidase, which cleaves lactose into glucose and galactose, which the cell then utilizes. The third structural gene, *lacA*, encodes an enzyme called galactoside acetyltransferase (GAT). This enzyme is not required for lactose metabolism, but is thought to play a role in cellular detoxification of nonmetabolizable lactopyranosides (Lewendon et al., 1995; Wang et al. 2002).

Alpha-Complementation

Vectors such as the pUC series and the M13mp series (see UNITS 1.5, 1.14, & 1.15) contain a piece of DNA that encodes an alpha fragment of β -galactosidase. These vectors exploit a phenomenon called *alpha-complementation* (see Fig. 1.4.2), which was discovered by Ullman, Jacob, and Monod in 1967. They showed that a cell that bears any of a number of deletions of the 5' end of the *lacZ* gene synthesizes an inactive C-terminal fragment of β -galactosidase, called an *omega* (ω) fragment. Similarly, a cell that bears a deletion of the 3' end of *lacZ* encodes an inactive N-terminal fragment of β -galactosidase called an *alpha* (α) fragment. However, if a cell contains two genes, one directing the synthesis of an alpha fragment, the other directing synthesis of an omega fragment, the β -galactosidase activity is observed. Many vectors incorporate a *lac* α -fragment gene, which is small and easily manipulated. Exploitation of these vectors requires use of a strain carrying the complementing ω -fragment gene to allow assembly of an active complex. This gene is often carried on an F' plasmid (see below). When these vectors are used, cells containing them are grown on medium containing IPTG, which inactivates *lac* repressor and thus derepresses ω peptide synthesis, and Xgal, which is turned blue by the enzymatic activity of β -galactosidase (see Table 1.4.2). On this medium, these vector-containing cells possess β -galactosidase activity and turn blue.

Lactose Analogs

There are many compounds related to lactose that were first used for the biochemical and genetic analysis of *lac* operon activity and are

Escherichia coli,
Plasmids, and
Bacteriophages

1.4.1

Table 1.4.1 Antibiotics, Their Modes of Action, and Modes of Bacterial Resistance^a

Antibiotic ^b	Stock conc. (mg/ml)	Final conc. (μg/ml)	Mode of action	Mode of resistance
Ampicillin ^c	4	50	Bacteriocidal; only kills growing <i>E. coli</i> ; inhibits cell wall synthesis by inhibiting formation of the peptidoglycan cross-link	β-lactamase hydrolyzes ampicillin before it enters the cell
Chloramphenicol, in methanol	10	20	Bacteriostatic; inhibits protein synthesis by interacting with the 50S ribosomal subunit and inhibiting the peptidyltransferase reaction	Chloramphenicol acetyltransferase inactivates chloramphenicol
D-Cycloserine, ^d in 0.1 M sodium phosphate buffer, pH 8	10	200	Bacteriocidal; only kills growing <i>E. coli</i> ; inhibits cell wall synthesis by preventing formation of D-alanine from L-alanine and formation of peptide bonds involving D-alanine	Mutations destroy the D-alanine transport system
Gentamycin	10	15	Bacteriocidal; inhibits protein synthesis by binding to the L6 protein of the 50S ribosomal subunit	Aminoglycoside acetyltransferase and aminoglycosidenucleotidyltransferase inactivate gentamycin; mutations in <i>rplF</i> (encodes the L6 protein) prevent the gentamycin from binding
Kanamycin	10	30	Bacteriocidal; inhibits protein synthesis; inhibits translocation and elicits miscoding	Aminoglycoside phosphotransferase, also known as neomycin phosphotransferase, aminoglycoside acetyltransferase, and aminoglycoside nucleotidyltransferase; inactivates kanamycin
Kasugamycin	10	1000	Bacteriocidal; inhibits protein synthesis by altering the methylation of the 16S RNA and thus an altered 30S ribosomal subunit	Mutations prevent kasugamycin from binding to the ribosome; mutations decrease uptake of kasugamycin
Nalidixic acid, pH to 11 with NaOH	5	15	Bacteriostatic; inhibits DNA synthesis by inhibiting DNA gyrase	Mutations in the host DNA gyrase prevent nalidixic acid from binding
Rifampicin, ^e in methanol	34	150	Bacteriostatic; inhibits RNA synthesis by binding to and inhibiting the β subunit of RNA polymerase; rifampicin sensitivity is dominant.	Mutation in the β subunit of RNA polymerase prevents rifampicin from complexing; rifampicin resistance is recessive
Spectinomycin	10	100	Bacteriostatic; inhibits translocation of peptidyl tRNA from the A site to the P site	Mutations in <i>rpsE</i> (encodes the S5 protein) prevent spectinomycin from binding; spectinomycin sensitivity is dominant and resistance is recessive
Streptomycin	50	30	Bacteriocidal; inhibits protein synthesis by binding to the S12 protein of the 30S ribosomal subunit and inhibiting proper translation; streptomycin sensitivity is dominant	Aminoglycoside phosphotransferase inactivates streptomycin; mutations in <i>rpsL</i> (encodes the S12 protein) prevent streptomycin from binding; streptomycin resistance is recessive
Tetracycline, ^e in 70% ethanol	12	12	Bacteriostatic; inhibits protein synthesis by preventing binding of aminoacyl tRNA to the ribosome A site	Active efflux of drug from cell

^aData assembled from Foster (1983), Gottlieb and Shaw (1967), and Moazed and Noller (1987).

^bAll antibiotics should be stored at 4°C, except tetracycline, which should be stored at -20°C. All antibiotics should be dissolved in sterile distilled H₂O unless otherwise indicated. Antibiotics dissolved in methanol often can be dissolved in the less hazardous ethanol.

^cCarbenicillin, at the same concentration, can be used in place of ampicillin. Carbenicillin can be stored in 50% ethanol/50% water at -20°C.

^dD-cycloserine solutions are unstable. They should be made immediately before use.

^eLight-sensitive; store stock solutions and plates in the dark.

1.4.2

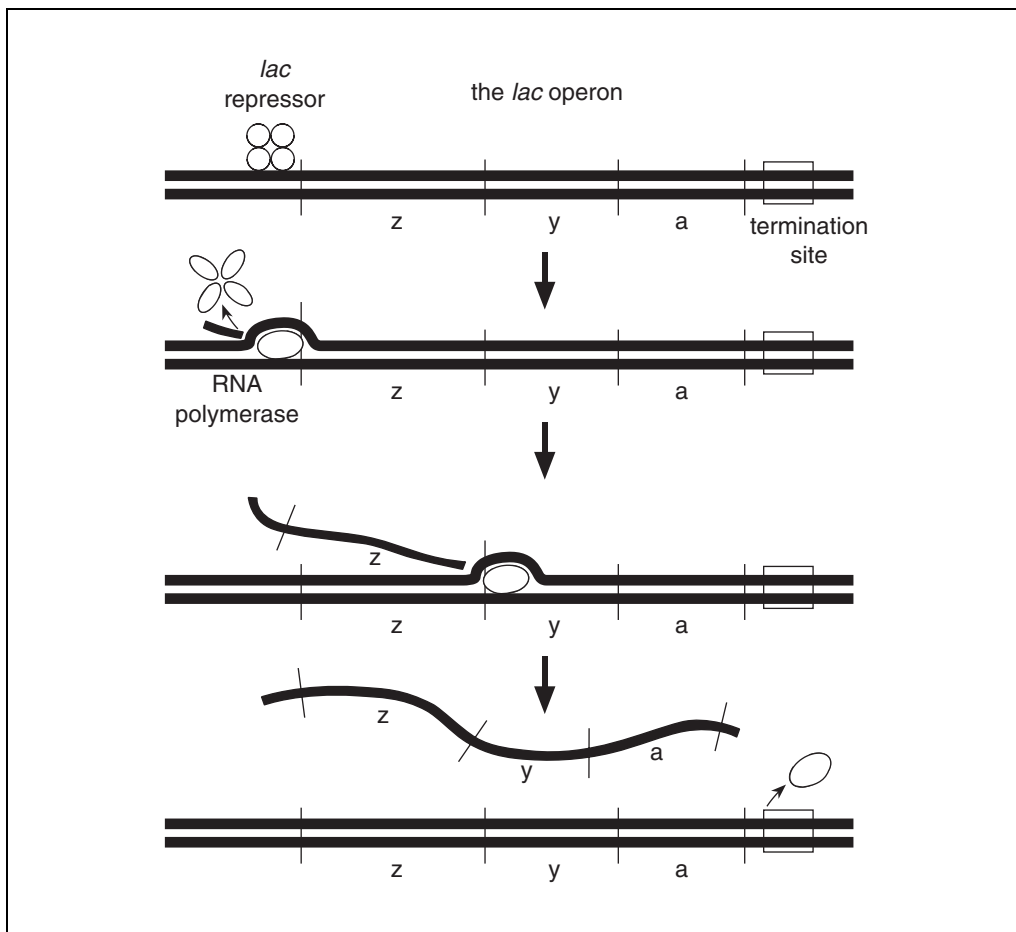


Figure 1.4.1 The first line shows the genes of the *lac* operon. *lacZ*, *lacY*, and *lacA* transcription is repressed by *lac* repressor. In the next line, inactive repressor is no longer bound to the operator, and the genes are being transcribed by RNA polymerase. In the third line, RNA polymerase transcribes the operon. Finally, RNA polymerase encounters a transcription terminator and, in a reaction requiring rho protein, RNA polymerase is detached from the DNA and the nascent transcript.

used in the cloning technology described in this book. These are described in Table 1.4.2.

THE F FACTOR

The F (fertility) factor is a genetic unit found in some strains of *E. coli*. Bacteria that contain the F factor are used in many techniques described in this book, mainly for two reasons. First (as described later in this chapter), possession of F allows a cell to be infected by vectors based on filamentous phages, which bind to cell surface structures called pili elaborated by F-containing cells. Second, defective *lacZ* genes that encode the ω fragment of β -galactosidase (described above) are commonly carried on F' factors.

The F factor is found in three alternative forms: as double-stranded, single-copy, circular extrachromosomal plasmid DNA (F⁺); as plasmid DNA like F⁺ but also including other bacterial genes (F'); and as a stretch of linear

DNA integrated into various sites on the bacterial chromosome (Hfr). Possession of the F factor confers on *E. coli* the ability to donate DNA in bacterial crosses (or matings). For this reason cells that carry F are sometimes called *male*. F or F' plasmids can transfer themselves to other cells, and may occasionally cause transfer of other plasmids. This latter process is called *mobilization*. Mutations called *tra* prevent F from transferring itself or mobilizing other plasmids. Integrated Tra⁺ F factors (Hfr) can cause transfer of chromosomal DNA to other cells, but the recipient usually does not receive the F sequence.

NONSENSE SUPPRESSORS

Some vectors used in recombinant DNA research (e.g., plasmid π VX and phage Charon 4a) contain nonsense mutations in essential genes. These vectors must be propagated in special strains of *E. coli*. In these strains, trans-

Table 1.4.2 Lactose Analogs Used in DNA Cloning Technology

Galactoside	Stock concentration ^a	Use	Characteristics	Reference
Isopropyl-1-thio- β -D-galactoside (IPTG)	100 mM	Very effective inducer	Nonmetabolizable inducer	Barkley and Bourgeois, 1978 (pp. 177-220)
5-Bromo-4-chloro-3-indolyl- β -D-galactoside (Xgal)	20 mg/ml (dissolved in <i>N,N</i> dimethyl formamide)	Identification of <i>lacZ</i> ⁺ bacteria, especially useful for detecting β -galactosidase made by recombinant vectors	Noninducing chromogenic substrate of β -galactosidase (cleavage of Xgal results in blue color); production of blue color independent of <i>lacY</i> gene product	Miller, 1972
Orthonitrophenyl- β -D-galactoside (ONPG)	10 mM	β -galactosidase assays	Chromogenic substrate of β -galactosidase (cleavage of ONPG results in yellow color)	Miller, 1972 (pp. 352-355)
6-O- β -D-Galactopyranosyl D-glucose (allolactose)			Inducer of the lactose operon in vivo; lactose is converted into allolactose by β -galactosidase	Zabin and Fowler, 1978 (pp. 89-121)
Phenyl- β -D-galactoside (Pgal)	2 mg/ml	Selection for <i>lac</i> constitutive mutants	Noninducing substrate of β -galactosidase; uptake partly dependent on <i>lacY</i> gene product	Miller, 1978 (pp. 31-88)
Orthonitrophenyl- β -D-thiogalactoside (TONPG)	10 mM	Selection for <i>lac</i> ⁻ mutants	Transported into cells by <i>lac</i> permease (the <i>lacY</i> gene product); inhibits cell growth at high concentration	Miller, 1978 (pp. 31-88)

^aStock solutions should be dissolved in sterile water unless otherwise noted.

lation of messages does not always stop when the ribosome encounters a chain termination codon (amber or ochre), but sometimes continues, with a new amino acid installed at the end of the growing polypeptide. This process is called *nonsense suppression* and strains of *E. coli* in which it occurs are said to contain *nonsense suppressors*. In a strain that contains an efficient or a “strong” suppressor, suppression might occur 50% of the time an amber codon is encountered.

The mechanism of nonsense suppression is a simple one: the cell contains a mutant species of tRNA in which the anticodon loop has mutated so that it base pairs with the UAG amber codon or the UAA ochre codon. Nonsense suppressors commonly used in cloning technology are given in Table 1.4.3. UGA (opal) suppressors also exist but are rarely used.

GENETIC MARKERS

Genetic markers in *E. coli* are named according to the convention proposed by Demerec et al. (1966). All genes of a given strain are presumed to be in the wild-type state unless otherwise noted in the genotype (see Table 1.4.4). Gene names have three italicized lowercase letters, sometimes followed by an italic uppercase letter, and sometimes also followed by an italic arabic number that specifies the precise mutation (allele) in question (e.g., *lacY1*, *trp-31*). The three-letter combination is usually a mnemonic intended to suggest the function of the gene. Proper notation omits superscript + and – in a genotype, but these are sometimes used redundantly for clarity. Deletion mutations are described by Δ , followed by the names of deleted genes in parentheses, followed by the allele number [e.g., $\Delta(lac-pro)X111$]. The delta may be replaced by “del” or “d.” Sometimes a

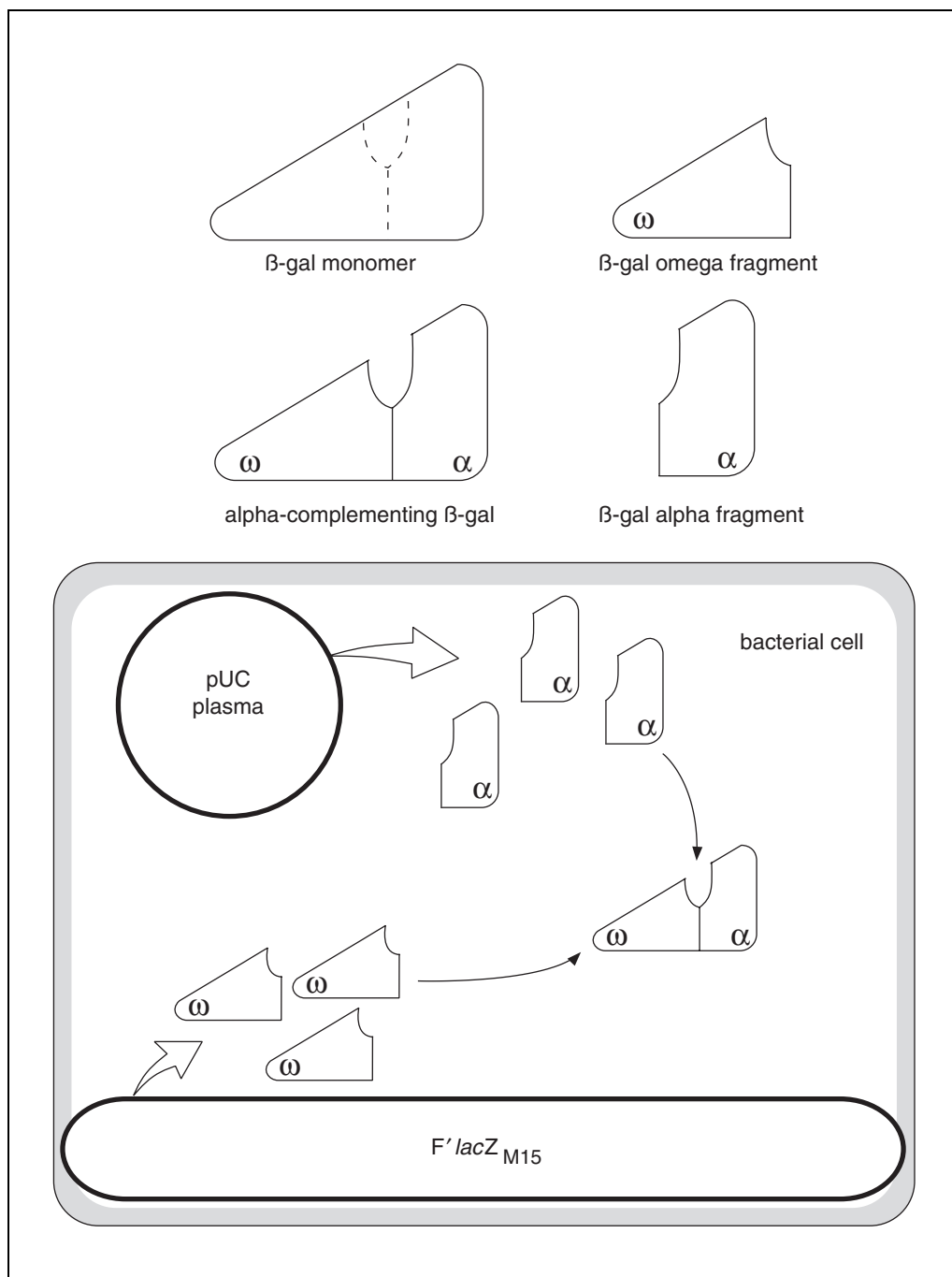


Figure 1.4.2 Alpha-complementation.

Table 1.4.3 Commonly Used Nonsense Suppressors^a

Suppressor	Map position ^b	Type of suppressor	Amino acid inserted	tRNA gene
<i>supD</i> (<i>su1</i>)	43	Amber	Serine	<i>serU</i>
<i>supE</i> (<i>su2</i>)	16	Amber	Glutamine	<i>glnU</i>
<i>supF</i> (<i>su3</i>)	27	Amber	Tyrosine	<i>tyrT</i>
<i>supB</i> (<i>suB</i>)	16	Ochre/amber	Glutamine	<i>glnU</i>
<i>supC</i> (<i>suC</i>)	27	Ochre/amber	Tyrosine	<i>tyrT</i>

^aData compiled from Bachmann (1983) and Celis and Smith (1979).

^bGiven in minutes; see Bachmann (1983) for description.

Table 1.4.4 Examples of Genotype and Phenotype Nomenclature

Genotype	Phenotype	Description of phenotype
<i>Some straightforward examples</i>		
<i>trp-31</i>	Trp ⁻	Requires tryptophan for growth on minimal media
<i>uvrA</i>	UV ^s	Sensitive to UV light
<i>recA</i>	Rec ⁻	Recombination defective
<i>Some common examples that are not so straightforward</i>		
<i>supE44</i>	Sup ⁺	Carries a tRNA suppressor gene. The <i>mutant</i> gene product, not the wild type, suppresses nonsense mutations; wild type is indicated as <i>sup⁰</i> , Sup ⁻ , and does not suppress
<i>rpsL104</i>	Str ^r	Resistant to streptomycin (this makes a mutant ribosomal protein, small subunit, the target of the drug)
<i>rpsE</i>	Spc ^r	Resistant to spectinomycin (also codes for a ribosomal protein, a different one)
<i>gyrA</i>	Nal ^r	Resistant to nalidixic acid (the mutation affects DNA gyrase)
<i>One mutation may create several phenotypes</i>		
<i>dam-3</i>	Dam ⁻ , 2-AP ^s UV ^s	DNA not methylated at adenines in GATC Sensitive to 2-aminopurine Sensitive to UV light
<i>hsdS</i>	<i>EcoK</i> R ⁻ , <i>EcoK</i> M ⁻	Neither restricts nor modifies DNA that enters the cell
<i>Some mutations lead to counterintuitive phenotypes</i>		
<i>recD</i>	ExoV ⁻ but Rec ⁺	Exonucleolytic activity of the RecBCD protein is defective, but the recombinational activity is intact

phenotype designation (see box) in parentheses follows the genotype designation, if the former is not obvious from the latter [e.g., *rpsL104* (Str^r)]. However, this usage is by no means universal.

Table 1.4.5 lists commonly used genetic markers, with methods for verifying their presence or absence in bacterial cells. Genotypes of several strains used for different applications are listed in Table 1.4.6.

GENOTYPE AND PHENOTYPE

Genotype indicates what genes are mutated in a strain. A genotype is a theoretical construct describing a genetic constitution that would explain the phenotype of the strain. It is derived from considerations of the strain's behavior and ancestry.

Phenotype describes the observable behavior of the strain—e.g., Lac⁻ fails to grow on lactose as a sole carbon source. Phenotypes are in Roman type, the first letter is capitalized, and the letters are always followed by superscript + or - (sometimes r, resistant, or s, sensitive). A phenotype is a datum to be explained.

Genotype and phenotype names are usually related, but the relationship is not always obvious. Examples are provided in Table 1.4.4.

DNA RESTRICTION, MODIFICATION, AND METHYLATION

This section and the next two describe *E. coli* functions that can prevent cloning the sequence of interest. *E. coli* has at least four restriction systems that identify foreign DNA and destroy it. These systems, encoded by *hsdRMS*, *mcrA*, *mcrB*, and *mrr*, can be avoided by using host strains in which they are disabled by mutation. Restriction of DNA and the content of methylated bases in the DNA are interrelated as described below. To select the appropriate strain, it is necessary to know the content of methylated bases in the DNA to be cloned.

The *EcoK* restriction system, encoded by the *hsdRMS* genes, is the best understood of the *E. coli* restriction mechanisms (Bickle, 1982). It attacks DNA that carries the site:



Table 1.4.5 Commonly Used Genetic Markers and How to Test Them^a

Nutritional markers	Streak or replica plate colonies of the strain onto plates with and without the nutrient to be tested, but which contain all other necessary nutrients.
Antibiotic resistance markers	Streak or replica plate colonies of the strain onto plates with and without the antibiotic.
<i>Other markers</i>	
<i>lacZ</i> ⁺	Streak strain on an LB plate with Xgal and IPTG (UNIT 1.4). Colonies should turn blue. Colonies of control <i>lacZ</i> ⁻ strain should not turn blue.
<i>lacZΔM15</i> ^b	Transform strain with pUC plasmid and with control plasmid such as pBR322. Streak transformants onto LB/ampicillin plate with Xgal and IPTG. Colonies bearing pUC plasmid should turn blue, while colonies bearing pBR322 should not.
F ⁺ or F'	Spot M13 phage onto a lawn of the cells. Small plaques should appear (see UNIT 1.15).
<i>recA</i>	Using a toothpick, make a horizontal stripe of cells across an LB plate. Also make a stripe of <i>recA</i> ⁺ control cells. Cover half of the plate with a piece of cardboard, and irradiate the plate with 300 ergs/cm ² of 254 nm UV light from a hand-held UV source (typically 20 sec exposure from a lamp held 50 cm over the plate). <i>recA</i> ⁻ cells are very sensitive to killing by UV light, and <i>recA</i> ⁻ cells in the unshielded part of the plate should be killed by this level of irradiation.
<i>recBCD</i>	Spot dilutions of λ <i>gam</i> ⁻ (UNIT 1.9) on a lawn of cells side by side with dilutions of λ <i>gam</i> ⁺ . The <i>gam</i> ⁻ plaques should be almost as big as the <i>gam</i> ⁺ plaques.
<i>hsdS</i> ⁻	(1) Use the strain and a wild-type strain to plate out serial dilutions of a λ-like phage stock grown on an <i>hsdS</i> ⁻ or <i>hsdR</i> ⁻ host. If the phage stock came from an <i>hsdS</i> ⁻ host, then it should make plaques with 1 × 10 ⁴ to 1 × 10 ⁶ higher efficiency on the putative <i>hsdS</i> ⁻ host than on a wild-type host. If the phage stock came from an <i>hsdR</i> ⁻ (<i>hsdS</i> ⁺ <i>hsdM</i> ⁺) host, it should make plaques with the same efficiency on both strains. (2) Suspend one of the fresh plaques from the putative <i>hsdS</i> ⁻ host in 1 ml lambda dilution buffer. Titer this suspension on the putative <i>hsdS</i> ⁻ strain and on a wild-type strain. The suspension should make plaques at 1 × 10 ⁴ to 1 × 10 ⁶ higher efficiency on the <i>hsdS</i> ⁻ strain than on the wild-type strain. One plaque contains ~1 × 10 ⁷ phage.
<i>hsdR</i> ⁻	(1) Perform step 1 described above, using a plate stock made from an <i>hsdS</i> ⁻ host.
(<i>hsdS</i> ⁺ <i>hsdM</i> ⁺)	(2) Suspend one of the fresh plaques in 1 ml lambda dilution buffer. Titer this suspension on the putative <i>hsdR</i> ⁻ strain and on a wild-type strain. This suspension should make plaques with the same efficiency on the <i>hsdR</i> ⁻ as on a wild-type strain.
<i>dam</i>	Transform the strain and a wild-type strain with a plasmid that contains recognition sites for the enzymes <i>Mbo</i> I or <i>Bcl</i> I. Prepare plasmid DNA from both strains and verify that plasmid DNA isolated from the <i>dam</i> ⁻ strain is sensitive to digestion by the enzyme.
<i>dcm</i>	Transform the strain and a wild-type strain with a plasmid that contains recognition sites for <i>Scr</i> FI. Prepare plasmid DNA from both strains to verify that only plasmid DNA from the <i>dcm</i> strain is fully sensitive to digestion by the enzyme. Half of the <i>Scr</i> FI sites will be cut even when the DNA is <i>dcm</i> -methylated.
<i>lon</i>	Streak LB plate for single colonies. Also streak a control plate of a wild-type strain. Incubate at 37°C. Colonies of the <i>lon</i> ⁻ strain should be larger, glistening, and mucoidal.

^aCommonly used protocols in this table are media preparation (UNIT 1.1), streaking and replicating a plate (UNIT 1.3), and growing lambda-derived vectors (UNIT 1.12).

^bEncodes omega fragment of β-galactosidase.

Table 1.4.6 Commonly Used *Escherichia coli* Strains

Strain ^a	Genotype	Reference ^b
AR58	<i>sup</i> ⁰ <i>galK2 galE::Tn10</i> (λ cI857 Δ H1 <i>bio</i> ⁻ <i>uvrB kil</i> ⁻ <i>cIII</i> ⁻) Str ^r	A. Shatzman, pers comm. ^d
AR120	<i>sup</i> ⁰ <i>galK2 nad::Tn10</i> (Tet ^r) (λ cI ⁺ <i>ind</i> ⁺ <i>p_L-lacZ fusion</i>) Str ^r	A. Shatzman, pers comm. ^d
AS1 ^f	<i>endA1 thi-1 hsdR17(r_K⁻m_K⁺) supE44</i> (λ cI ⁺)	A. Shatzman, pers comm. ^d
BNN102 ^f	C600 <i>hflA150 chr::Tn10 mcrA1 mcrB</i>	Young and Davis, 1983 ^c
BW313 ^s	<i>Hfr lysA</i> ⁻ <i>dut ung thi-1 recA spoT1</i>	Kunkel et al., 1987 ^{c,d}
C600	<i>thi-1 thr-1 leuB6 lacY1 tonA21 supE44 mcrA</i>	Appleyard, 1954 ^{c,e}
CJ236 ^s	<i>dut1 ung1 thi-1 relA1/pCJ105</i> (Cm ^r)	Kunkel et al., 1987 ^c ; Joyce and Grindley, 1984
DH1	<i>recA1 endA1 thi-1 hsdR17 supE44 gyrA96</i> (Nal ^r) <i>relA1</i>	Hanahan, 1983 ^c ; D. Hanahan, pers. comm. ^{d,e}
DH5 α F ^h	F ['] / <i>endA1 hsdR17(r_K⁻m_K⁺) supE44 thi-1 recA1 gyrA</i> (Nal ^r) <i>relA1</i> Δ (<i>lacZYA-argF</i>) _{U169} (<i>m80lacZ</i> Δ M15)	See DH1 references
DK1	<i>hsdR2 hsdM</i> ⁺ <i>hsdS</i> ⁺ <i>araD139</i> Δ (<i>ara-leu</i>) ₇₆₉₇ Δ (<i>lac</i>) _{X74} <i>galU galK rpsL</i> (Str ^r) <i>mcrA mcrB1</i> Δ (<i>sr1-recA</i>) ₃₀₆	D. Kurnit and B. Seed, pers. comm. ^{d,e}
ER1451	F ['] <i>traD36 proAB lacI^q</i> Δ (<i>lacZ</i>)M15/ <i>endA gyrA96 thi-1 hsdR2</i> (or <i>hsdR17</i>) <i>supE44</i> Δ (<i>lac-proAB</i>) <i>mcrB1 mcrA</i>	Raleigh et al., 1988 ^{d,e}
HB101 ⁱ	Δ (<i>gpt-proA</i>) ₆₂ <i>leuB6 thi-1 lacY1 hsdS_B20 recA rpsL20</i> (Str ^r) <i>ara-14 galK2 xyl-5 mtl-1 supE44 mcrB_B</i>	Boyer and Roulland-Dussoix, 1969 ^{c,d,e}
JM101 ^j	F ['] <i>traD36 proA</i> ⁺ <i>proB</i> ⁺ <i>lacI^q lacZ</i> Δ M15/ <i>supE thi</i> Δ (<i>lac-proAB</i>)	Yanisch-Perron et al., 1985 ^{c,d,e}
JM105 ^j	F ['] <i>traD36 proA</i> ⁺ <i>proB</i> ⁺ <i>lacI^q lacZ</i> Δ M15/ Δ (<i>lac-pro</i>) _{X111} <i>thi rpsL</i> (Str ^r) <i>endA sbcB supE hsdR</i>	See JM101 references
JM107 ^j	F ['] <i>traD36 proA</i> ⁺ <i>proB</i> ⁺ <i>lacI^q lacZ</i> Δ M15/ <i>endA1 gyrA96</i> (Nal ^r) <i>thi hsdR17 supE44 relA1</i> Δ (<i>lac-proAB</i>) <i>mcrA</i>	See JM101 references
JM109 ^j	F ['] <i>traD36 proA</i> ⁺ <i>proB</i> ⁺ <i>lacI^q lacZ</i> Δ M15/ <i>recA1 endA1 gyrA96</i> (Nal ^r) <i>thi hsdR17 supE44 relA1</i> Δ (<i>lac-proAB</i>) <i>mcrA</i>	See JM101 references
K38	HfrC (λ)	Russel and Model, 1984
KM392	<i>hsdR514(r_K⁻m_K⁺) supE44 supF58 lacY galK2 galT22 metB1 trp55 mcrA</i> Δ <i>lac</i> _{U169} <i>proC::Tn5</i>	T. St. John, pers. comm. ^d ; K. Moore ^e
LE392	<i>hsdR514(r_K⁻m_K⁺) supE44 supF58 lacY galK2 galT22 metB1 trp55 mcrA</i>	Borck, et al., 1976 ^c ; N. Murray, pers. comm. ^d ; L. Enquist ^e
MC1061	<i>hsdR2 hsdM</i> ⁺ <i>hsdS</i> ⁺ <i>araD139</i> Δ (<i>ara-leu</i>) ₇₆₉₇ Δ (<i>lac</i>) _{X74} <i>galE15 galK16 rpsL</i> (Str ^r) <i>mcrA mcrB1</i>	Casadaban and Cohen, 1980 ^c ; M. Casadaban ^{d,e}
MM294	<i>endA thiA hsdR17 supE44</i>	Backman et al., 1976 ^c ; M. Meselson ^{d,e}

continued

and results in double-strand cleavage at a variable distance from the site, leading eventually to degradation of the resulting fragments. DNA is not attacked if it lacks the site, or if the site is present but methylated at the adenines shown (^mA).

The *EcoK* enzyme is both genetically and enzymatically complex. The HsdR, HsdM, and HsdS subunits are required for restriction of an unmethylated substrate. The same complex will methylate the same substrate, but at a very slow rate, so that an unmethylated target rarely

survives. A substrate methylated on only one strand (hemimethylated) will be methylated on the other strand by the three-protein complex, but will not be cut. HsdM and HsdS together can methylate either an unmethylated or a hemimethylated substrate. The three-protein complex is inactive for restriction if any of the three subunits is defective, but can still methylate if HsdR is defective.

In summary, a strain defective in the *hsdR* gene is described as having the phenotype HsdR⁻M⁺ (or, equivalently, *EcoK* R⁻M⁺ or

Table 1.4.6 Commonly Used *Escherichia coli* Strains, continued

Strain ^a	Genotype	Reference ^b
NM539 ^k	<i>supF hsdR</i> (P2 <i>cox3</i>)	Frischauf et al., 1983 ^c ; Lindahl and Sunshine, 1972 ^d ; N. Murray ^e
P2392	<i>hsdR514</i> (<i>r_K⁻m_K⁺</i>) <i>supE44 supF58 lacY galK2 galT22 metB1 trp55 mcrA</i> (P2)	L. Klickstein, pers. comm. ^d
PR722	F' Δ (<i>lacIZ</i>) _{E65} <i>pro⁺/proC::Tn5</i> Δ (<i>lacIZYA</i>) _{U169} <i>hsdS20 ara-14 galK2 rpsL20</i> (Str ^r) <i>xyl-5 mtl-1 supE44 leu</i>	P. Riggs, pers.comm. ^d
Q359	<i>hsdR⁻ hsdM⁺ supE tonA</i> (ϕ 80 ^r) (P2)	Karn et al., 1980 ^{c,d,e}
RR1	Δ (<i>gpt-proA</i>) ₆₂ <i>leuB6 thi-1 lacY1 hsdS_B20 rpsL20</i> (Str ^r) <i>ara-14 galK2 xyl-5 mtl-1 supE44 mcrB_B</i>	Bolivar et al., 1977
Y1088 ^l	<i>supE supF metB trpR hsdR⁻ hsdM⁺ tonA21 strA</i> Δ <i>lac</i> _{U169} <i>mcrA proC::Tn5/pMC9</i>	Huynh et al., 1985 ^c ; Miller et al., 1984 ^d ; R. Young ^e ; M. Calos (pMC9) ^e
Y1089 ^l	Δ <i>lac</i> _{U169} <i>proA⁺ Δ(lon) araD139 strA hflA150 chr::Tn10/pMC9</i>	See Y1088 references
Y1090 ^l	Δ <i>lac</i> _{U169} <i>proA⁺ Δ(lon) araD139 strA supF trpC22::Tn10 mcrA/pMC9</i>	See Y1088 references

^aThe original *E. coli* K-12 strain was an F⁺ λ lysogen, but most K-12 derivatives in common use have been cured of the F factor and prophage and these are indicated only when present. All other genes in these strains are presumed to be wild-type except for the genotype markers noted in the second column.

^bReference for all *mcr* and *mrr* genotypes is Raleigh et al., 1988.

^cReference for genotype of strain.

^dSource of additional genotype information.

^eThought to be responsible for original strain construction.

^fAS1 is also known as MM294c1⁺. BNN102 is also known as C600 *hflA*.

^gBoth CJ236 and BW313 are commonly used in oligonucleotide-directed mutagenesis. pCJ105, the plasmid CJ236 carries, is not relevant for this application.

^hThree strains are in circulation. DH5 is a derivative of DH1 that transforms at slightly higher efficiency. DH5 α and DH5 α F' are derivatives that carry a deletion of the *lac* operon and a Φ 80 prophage that directs synthesis of the omega fragment of β -galactosidase. DH5 α F' carries an F' factor as well. DH5 α and DH5 α F' are proprietary strains and the cells are prepared in some way that allows them to be transformed with slightly higher efficiency than DH5.

ⁱIn this strain, the area of the chromosome that contains the *hsd* genes was derived from the related B strain of *E. coli*.

^jThe continued presence of the F' factor in JM strains can be insured by starting cultures only from single colonies grown on minimal plates that do not contain proline. These strains encode the omega fragment of *lacZ* and are frequently used with vectors that direct the synthesis of the *lacZ* alpha fragment. These strains are frequently used with M13 vectors for DNA sequencing (*UNITS 1.14, 1.15, & 7.4*).

^kIt is not known whether this strain has markers other than those listed.

^lpMC9, the plasmid in the Y strains listed here, directs the synthesis of large amounts of *lac* repressor. It also confers resistance to tetracycline and ampicillin (Lebrowski et al., 1984, *EMBO. J.* 3:3117-3121).

R_K⁻M_K⁺; see box): it will methylate newly introduced DNA but will not restrict it. However, a strain defective in either *hsdM* or *hsdS* will neither restrict nor methylate, and has the phenotype HsdR⁻M⁻ (or *EcoK* R⁻M⁻ or R_K⁻M_K⁻).

In contrast with *EcoK*, the other three restriction systems of *E. coli* K-12—*mcrA*, *mcrB*, and *mrr*—specifically attack DNA that is methylated at particular sequences, rather than DNA that is not. Either methylated cytosine residues or methylated adenine residues can create problems (see below).

The action of either *mcrA* or *mcrB* reduces the number of clones recovered from libraries made with genomic DNA from other organ-

isms, and leads to bias against recovery of specific fragments from those libraries (Raleigh et al., 1988; Whittaker et al., 1988; Woodcock et al., 1988, 1989; *mrr* has not been tested). For *McrB* there is evidence that a nuclease is responsible for these effects (E. Sutherland and E.A. Raleigh, unpub. observ.), but no such evidence is available for the other two systems.

Even without biochemical characterization, something can be said of the recognition sites for these systems. *McrA* restricts DNA modified by the *HpaII* (5' C^mCGG) methylase and possibly other methylases. *McrB* restricts DNA modified by any one of 14 other modification

methylases, which led to the suggestion that the McrB recognition site is 5' G^mC (Raleigh and Wilson, 1986). Mrr restricts DNA modified by the *Hha*II (5' G^mANTC) or *Pst*I (5' CTGC^mAG) methylases, but *not* that modified by the *Eco*RI methylase, among others (Heitman and Model, 1987).

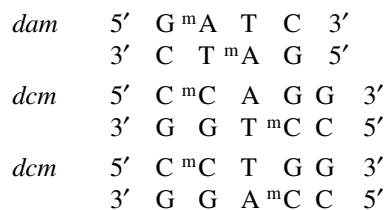
Many commonly used *E. coli* strains are McrA⁻; including (from Table 1.4.5) BNN102 (also known as C600*hflA*), C600, JM107, JM109, LE392, Y1088, and Y1090. Of the strains listed in Table 1.4.5, only BNN102, HB101, and MC1061 are McrB⁻; and only HB101 is Mrr⁻ (see also Raleigh et al., 1988).

A strain should be used which lacks the appropriate methyl-specific restriction system(s) when cloning genomic DNA from an organism containing methylated bases. All mammals and higher plants, and many prokaryotes, contain methylcytosine (Ehrlich and Wang, 1981), so McrA⁻B⁻ strains should be used for libraries of DNA from these organisms. Bacteria and lower eukaryotes may contain methyladenine, so Mrr sensitivity should be considered. However, the important experimental organisms *Drosophila melanogaster* and *Saccharomyces cerevisiae* contain no detectable methylated bases.

In addition, any time DNA is methylated in vitro during a manipulation, an appropriate restriction-deficient host should be used as DNA recipient. Methylases are used to generate novel restriction enzyme specificities or to protect cDNA from subsequent digestion (see *UNITS 3.1 & 5.6*). For example, the *Alu*I methylase (*M.Alu*I) is sometimes used to protect *Hind*III sites. McrB will restrict DNA modified by *M.Alu*I.

Once the DNA introduced into *E. coli* has been replicated, the foreign methylation pattern will be lost (and the *E. coli* methylation pattern will be acquired) unless the clone carries a methylase activity. Once successfully introduced, clones can be freely transferred among Mcr⁺ Mrr⁺ *E. coli* strains, since the methylation pattern will no longer be foreign. It is important that the clone be passed through an HsdM⁺ strain before trying to introduce it into an HsdR⁺ strain.

The normal methylation pattern of *E. coli* DNA is the product of three methylases. The *Eco*K methylase modifies the sequences indicated above. The *dam* and *dcm* gene products are also methylases (Marinus, 1987). The recognition sites for these are:



These modifications will render DNA resistant or partially resistant to some restriction endonucleases used for in vitro work (see Table 3.1.1), such as *Mbo*I and *Bcl*II (for Dam-modified DNA) or *Eco*RII (for Dcm-modified DNA). The Dam and Dcm methylases are *not* associated with any *E. coli* restriction function. Loss of Dam and/or Dcm methylation will not make the DNA sensitive to *Eco*K restriction, although loss of K modification will. However, Dam and Dcm modification confer sensitivity to Mrr and Mcr analogues in *Streptomyces* species (MacNeil, 1988).

RECOMBINATION AND ITS EFFECTS ON CLONED DNA INSERTS

During propagation in *E. coli*, DNA inserted into vectors is sometimes rearranged by the proteins involved in DNA recombination. Fortunately, although the genetics and enzymology of recombination in *E. coli* are still not well understood, there are mutant strains available that can provide solutions to two common cloning problems.

Problem 1. The DNA contains dispersed repeated sequences. Recombination occurs between these repeated sequences, causing loss of pieces of the DNA (see Fig. 1.4.3).

For plasmid libraries, this problem can be solved by propagating the DNA in a *recA*⁻ host, where homologous recombination does not occur. For libraries made using λ-derived vectors, the vector must also be recombination-defective (*red*). However, only about 30% to 50% of the cells are viable in such a strain, and libraries, particularly phage libraries, may be hard to propagate. Phage λ vectors that are *red* will not make high-titer lysates in *recA* strains, and *red gam* phage will not grow at all, unless the *recBCD* enzyme is also inactivated (see below). Many λ vectors are *red gam* to make use of the Spi⁻ selection or to make room for larger insert pieces (see *UNIT 1.10*).

Problem 2. The inserted DNA contains closely spaced inverted repeat sequences (palindromes or interrupted palindromes). Such stretches of DNA are not stably propagated in either phage or plasmid vectors. Available

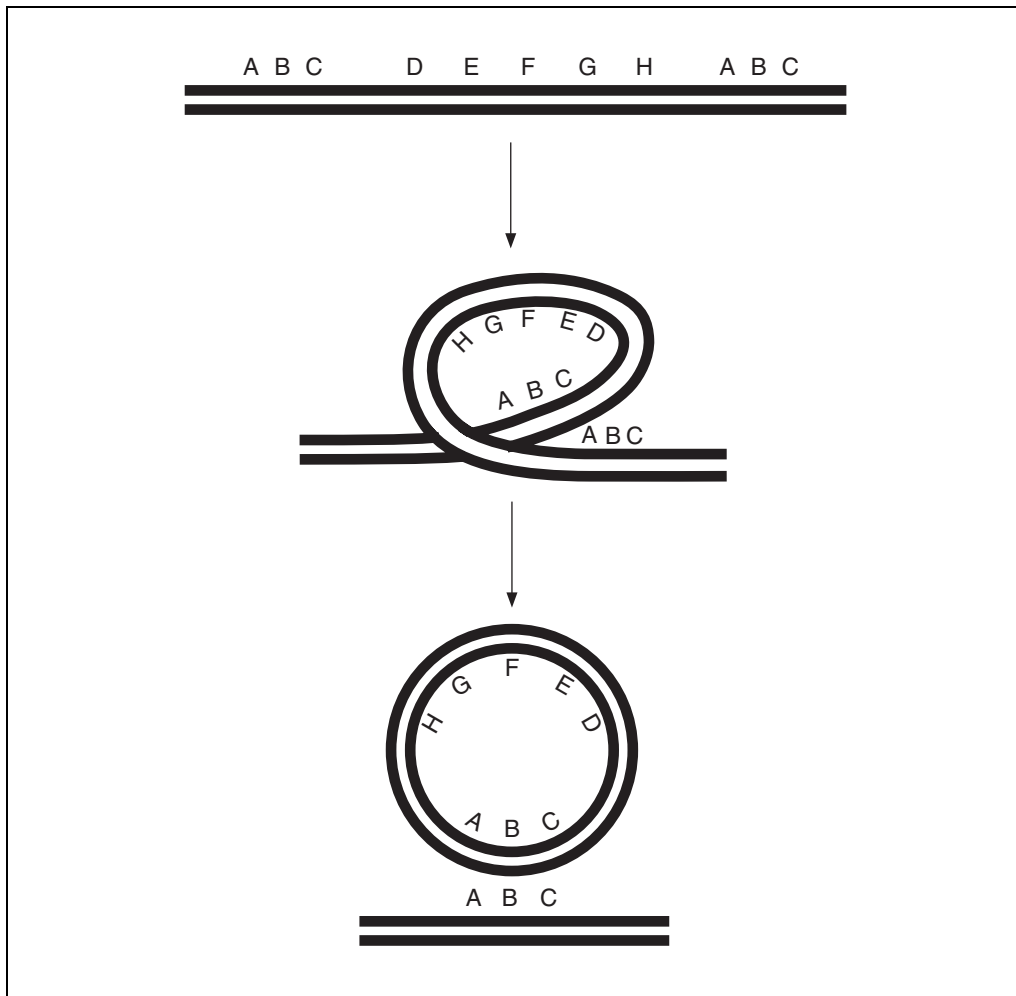


Figure 1.4.3 Recombination between homologous direct repeats.

knowledge is consistent with the idea that large (>300-bp) palindromes can sometimes form an alternative, hairpin structure that resembles an intermediate found in normal recombination called a Holliday junction, and are then acted upon by the host recombination system in such a way that the hairpin is eliminated or made smaller.

There are strains of bacteria from which phage and plasmid clones containing palindromes are recovered at higher frequency. These bacteria have inactivated exonuclease V (ExoV; encoded by the *recB*, *recC*, and *recD* genes) or the SbcC product (encoded by the *sbcC* gene). Many strains permissive for palindromes have defects in *recB recC* combined with a defect in *sbcA* (which probably encodes the RecE protein) or *sbcB* (which encodes exonuclease I).

Involvement of ExoV (RecBCD enzyme) in palindrome stabilization was first noticed by Leach and Stahl (1983) using artificially constructed palindromes in λ phages, and hosts

mutant in *recB*, *recC*, and *sbcB*. Because *recB recC* strains are sick, they tend to accumulate two additional mutations, one in *sbcB* (the gene for exonuclease I; suppressor of *recBC*), and one in *sbcC*, the biochemical nature of which is unknown (Lloyd and Buckman, 1985). Together these mutations restore recombination and increase cell viability. Wertman et al. (1986) and Wyman et al. (1986) found that both the *recB recC* defect and the *sbcB* defect independently contributed to stabilization of cloned palindromes in λ libraries.

Involvement of *recD* was first investigated by Wertman et al. (1986) and Wyman et al. (1986). *recD* codes for the nuclease activity of ExoV, as distinct from the recombination activity of that enzyme. Strains mutant in *recD* alone are Rec⁺ and healthy, and are not known to accumulate secondary mutations. Such strains were the best hosts for palindrome stabilization using λ -derived phages in the studies cited above.

The effect of *sbcC* was examined by Chalker et al. (1988), who found that *sbcC* alone was better at maintaining palindromes in λ than *recD* alone and better than *recB recC sbcB*, but that *recD sbcC* strains attained the highest palindrome stability (*recA recD sbcC* strains also maintained the palindrome). A further advantage of *sbcC* mutant strains is that palindromes were maintained in plasmids as well as phage (Chalker et al., 1988), whereas ExoV⁻ deficient mutants are poor hosts for plasmids regardless of palindrome content (see below).

EFFECTS OF RECOMBINATION-DEFECTIVE STRAINS ON VECTORS

Lambda-derived vectors or clones that are *red gam* (see *UNIT 1.10*, especially the *spi*⁻ selection) must be propagated on ExoV⁻ hosts, because the long linear multimers that are the normal substrate for lambda packaging are exonucleolytically degraded by ExoV in the absence of the Gam protein (Stahl, 1986). These phage will grow well on ExoV⁻ RecA⁺ hosts, reasonably well on ExoV⁻ RecA⁻ hosts, very poorly on ExoV⁺ RecA⁺ (the phage are packaged by an alternative packaging mechanism using circular dimers produced by recombination that depends on *E. coli* proteins), and not at all on ExoV⁺ RecA⁻. With ExoV⁻ RecA⁺ hosts, clones carrying a recombination hot spot called a Chi site (5' GCTGGTGG 3') will outgrow those that don't, resulting in a biased library. The ExoV⁻ RecA⁺ host can be *recBC sbcB* or *recD*, and the ExoV⁻ RecA⁻ host can be *recB recC sbcB recA* or *recD recA*.

Most cloning plasmids, including ColE1 derivatives like pBR322, and p15A derivatives like the pACYC vectors, are very unstable in ExoV⁻ deficient strains unless selection is maintained. Moreover, they are often unstable or difficult to maintain even with selection. Both *recB recC sbcA* strains (Basset and Kushner, 1984) and *recD* strains (Biek and Cohen, 1986, and references therein) behave this way. However, RecA⁻ suppresses this effect. Instability is probably due to recombination-initiated rolling-circle replication of the plasmids that leads to synthesis of long linear multimers, which fail to segregate properly at cell division (Silberstein and Cohen, 1987). The problem is particularly severe with very high-copy-number ColE1 derivatives such as pUC vectors, which may be impossible to establish at all in *recD* strains (E.A. Raleigh, unpublished observation).

From the above considerations, a universal host strain for phage and plasmid cloning vec-

tors would have the markers *recA recD sbcC hsdR mcrA mcrB mrr*. Although no such strain has been made, a *recD sbcC hsdR mcrA mcrB* strain, DL491, has been reported recently (Whittaker et al., 1988).

LITERATURE CITED

- Appleyard, R.K. 1954. Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from *Escherichia coli* K12. *Genetics* 39:440-452.
- Bachmann, B.J. 1983. Linkage map of *Escherichia coli* K-12, edition 7. *Microbiol. Rev.* 47:180-230.
- Backman, K., Ptashne, M., and Gilbert, W. 1976. *Proc. Natl. Acad. Sci. U.S.A.* 73:4174-4178.
- Barkley, M.D. and Bourgeois, S. 1978. Repressor recognition of operator and effectors. In *The Operon* (J. Miller, ed.) pp. 177-220. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Bassett, C.L. and Kushner, S.R. 1984. Exonucleases I, III and V are required for stability of ColE1-related plasmids in *Escherichia coli*. *J. Bacteriol.* 157:661-664.
- Bickle, T. 1982. ATP-dependent restriction endonucleases. In *Nucleases* (S.M. Linn and R.G. Roberts, eds.) p. 85. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Biek, D.P. and Cohen, S.N. 1986. Identification and characterization of *recD*, a gene affecting plasmid maintenance and recombination in *Escherichia coli*. *J. Bacteriol.* 167:594-603.
- Bolivar, F. Rodriguez, R.L., Greene, T.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W. 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* 2:95-113.
- Borck, J., Beggs, J.D., Brammer, W.J., Hopkins, A.S., and Murray, N.E. 1976. The construction in vitro of transducing derivatives of phage lambda. *Mol. Gen. Genet.* 146:199-207.
- Boyer, H.W. and Roulland-Dussoix, D. 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J. Mol. Biol.* 41:459-472.
- Casadaban, M.J. and Cohen, S.N. 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.* 138:179-207.
- Celis, J.E. and Smith, J.D. 1979. Nonsense Mutations and tRNA Suppressors. Academic Press, London.
- Chalker, A.F., Leach, D.R.F., and Lloyd, R.G. 1988. *Escherichia coli sbcC* mutants permit stable propagation of DNA replicons containing a long palindrome. *Gene* 71:201-205.
- Demerec, M., Adelberg, E.A., Clark, A.J., and Hartman, P.E. 1966. A proposal for uniform nomenclature in bacterial genetics. *Genetics* 54:61-66.
- Ehrlich, M. and Wang, R.Y. 1981. 5-methylcytosine in eukaryotic DNA. *Science* 212:1350-1357.

- Foster, T.J. 1983. Plasmid-determined resistance to antimicrobial drugs and toxic metal ions in bacteria. *Microbiol. Rev.* 47:361-409.
- Frischauf, A.-M., Lehrach, H., Poustka, A., and Murray, N. 1983. Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* 170:827-842.
- Gottlieb, D. and Shaw, P.D. 1967. Antibiotics. I. Mechanism of Action. Springer-Verlag, New York.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166:557-580.
- Heitman, J. and Model, P. 1987. Site-specific methylases induce the SOS DNA repair response in *Escherichia coli*. *J. Bacteriol.* 169:3243-3250.
- Huynh, T.V., Young, R.A., and Davis, R.W. 1985. Construction and screening of cDNA libraries in λ gt10 and λ gt11. In *DNA Cloning, Vol 1: A practical approach* (D.M. Glover, ed.) pp 49-78. IRL Press, Oxford.
- Joyce, C.M. and Grindley, N.D.F. 1984. Method for determining whether a gene of *Escherichia coli* is essential: Application to the *polA* gene. *J. Bacteriol.* 158:636.
- Karn, J., Brenner, S., Barnett, L., and Cesareni, G. 1980. Novel bacteriophage λ cloning vector. *Proc. Natl. Acad. Sci. U.S.A.* 77:5172-5176.
- Kunkel, T.A. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* 154:367-383.
- Leach, D.R.F. and Stahl, F.W. 1983. Viability of λ phages carrying a perfect palindrome in the absence of recombination nucleases. *Nature* 305:448.
- Lewendon, A., Ellis, J., Shaw, W.V. 1995. Structural and mechanistic studies of galactoside acetyltransferase, the *Escherichia coli* *LacA* gene product. *J. Biol. Chem.* 270:26326-31.
- Lindahl, G. and Sunshine, M. 1972. Excision-deficient mutants of bacteriophage P2. *Virology* 49:180-187.
- Lloyd, R.G. and Buckman, C. 1985. Identification and genetic analysis of *sbcC* mutations in commonly used *recBC sbcB* strains of *Escherichia coli* K-12. *J. Bacteriol.* 164:836-844.
- MacNeil, D. 1988. Characterization of a unique methyl-specific restriction system in *Streptomyces avermitilis*. *J. Bacteriol.* 170:5607-5612.
- Marinus, M.G. 1987. DNA methylation in *Escherichia coli*. *Ann. Rev. Genet.* 21:113-132.
- Miller, J. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Miller, J.H. 1978. The *lacI* gene: Its role in *lac* operon control and its uses as a genetic system. In *The Operon* (J. Miller, ed.) pp. 31-88. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Miller, J.H., Lebkowski, J.S., Greisen, K.S., and Calos, M.P. 1984. Specificity of mutations induced in transfected DNA by mammalian cells. *EMBO J.* 3:3117-3121.
- Moazed, D. and Noller, H.F. 1987. Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature* 327:389-394.
- Raleigh, E.A. and Wilson, G. 1986. *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. *Proc. Natl. Acad. Sci. U.S.A.* 83:9070-9074.
- Raleigh, E.A., Murray, N.E., Revel, H., Blumenthal, R.M., Westaway, D., Reith, A.D., Rigby, P.W.J., Elhai, J., and Hanahan, D. 1988. *McrA* and *McrB* restriction phenotypes of some *E. coli* strains and implications for gene cloning. *Nucl. Acids Res.* 16:1563-1575.
- Russel, M. and Model, P. 1984. Replacement of the *flp* gene of *Escherichia coli* by an inactive gene cloned on a plasmid. *J. Bacteriol.* 159:1034-1039.
- Silberstein, Z. and Cohen, A. 1987. Synthesis of linear multimers of *oriC* and pBR322 derivatives in *Escherichia coli* K-12: Role of recombination and replication functions. *J. Bacteriol.* 169:3131-3137.
- Stahl, F.W. 1986. Roles of double-strand breaks in generalized genetic recombination. In *Progress in Nucleic Acid Research and Molecular Biology, Vol. 33* (W.E. Cohn and K. Moldave, eds.) pp. 169-194. Academic Press, San Diego.
- Ullman, A., Jacob, F., and Monod, J. 1967. Characterization by in vitro complementation of a peptide corresponding to an operator-proximal segment of the β -galactosidase structural gene of *Escherichia coli*. *J. Mol. Biol.* 24:339-343.
- Wang, X.G., Olsen, L.R., Roderick, S.L. 2002. Structure of the *lac* operon galactoside acetyltransferase. *Structure* 10:581-588.
- Wertman, K.F., Wyman, A.R., and Botstein, D. 1986. Host/vector interactions which affect the viability of recombinant phage lambda clones. *Gene* 49:253-262.
- Whittaker, P.A., Campbell, A.J.B., Southern, E.M., and Murray, N.E. 1988. Enhanced recovery and restriction mapping of DNA fragments cloned in a new λ vector. *Nucl. Acids Res.* 16:6725-6736.
- Woodcock, D.M., Crowther, P.J., Diver, W.P., Graham, M.W., Bateman, C., Baker, D.J., and Smith, S.S. 1988. *RglB* facilitated cloning of highly methylated eukaryotic DNA: The human L1 transposon, plant DNA and DNA methylation in vitro with human DNA methyltransferase. *Nucl. Acids Res.* 16:4465-4482.
- Woodcock, D.M., Crowther, P.J., Doherty, J., Jefferson, S., De Cruz, E., Noyer-Weidner, M., Smith, S.S., Michael, M.Z., and Graham, M.W. 1989. Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucl. Acids Res.* 17:3469-3478.

- Wyman, A.R., Wertman, K.F., Barker, D., Helms, C., and Petri, W.H. 1986. Factors which equalize the representation of genome segments in recombinant libraries. *Gene* 49:263-271.
- Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103-119.
- Young, R.A. and Davis, R.W. 1983. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. U.S.A.* 80:1194-1198.
- Zabin, I. and Fowler, A.V. 1978. β -galactosidase, the lactose permease protein, and thiogalactoside transacetylase. *In The Operon* (J. Miller, ed.) pp. 89-122. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

KEY REFERENCES

- Beckwith, J. and Zipser, D., eds. 1970. *The Lactose Operon*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- An introduction to early work with the lactose operon.*
- Miller, 1972. See above.
- Contains excellent introductions to experimental techniques for working with E. coli and λ -derived phages.*
- Miller, J., ed. 1981. *The Operon*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

An updated and thorough introduction to regulatory mechanisms, starting with the lactose operon.

- Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M., and Umberger, H.E., eds. 1996. *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington, D.C.

Encyclopedic coverage of the biology of these useful bacteria.

- Ullman, A. and Perrin, D. 1970. Complementation in β -galactosidase. *In The Lactose Operon* (J. Beckwith and D. Zipser, eds.) pp. 143-172. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Definitive review of alpha-complementation.

Contributed by Elisabeth A. Raleigh
New England Biolabs
Beverly, Massachusetts

Karen Elbing
Clark & Elbing LLP
Boston, Massachusetts

Roger Brent
The Molecular Sciences Institute
Berkeley, California

Introduction to Plasmid Biology

UNIT 1.5

Plasmids are self-replicating, extrachromosomal DNA molecules found in virtually all bacterial species. In nature, plasmids occur in exuberant profusion, varying in structure, size, mode of replication, number of copies per bacterial cell, ability to propagate in different bacteria, transferability between bacterial species, and perhaps most important, in the traits they carry. Most prokaryotic plasmids are double-stranded circular DNA molecules; however, linear plasmids have been identified in both gram-positive and gram-negative bacteria. The size of plasmids varies widely, from several kilobases to hundreds of kilobases. Replication of plasmids depends on host-cell proteins but also may require plasmid-encoded functions. Plasmid replication may be synchronized with the bacterial cell cycle, resulting in a low number of plasmid molecules per bacterial cell, or independent of the host cell cycle, allowing for the proliferation of hundreds of plasmid copies per cell. Some plasmids freely transfer their DNA across bacterial species, some only transfer their DNA into bacteria of the same species, and some do not transfer their DNA at all. Plasmids carry genes that specify a wide variety of functions including: resistance to antibiotics, resistance to heavy metals, sensitivity to mutagens, sensitivity or resistance to bacteriophages, production of restriction enzymes, production of rare amino acids, production of toxins, determination of virulence, catabolism of complicated organic molecules, ability to form symbiotic relationships, and ability to transfer DNA across kingdoms.

Starting in the 1970s, vectors for propagation, manipulation, and delivery of specific DNA sequences were constructed with fragments from naturally occurring plasmids, primarily *Escherichia coli* plasmids. All plasmid vectors contain three common features: a **replicator**, a **selectable marker**, and a **cloning site**. The replicator is a stretch of DNA that contains the site at which DNA replication begins (the origin of replication or *ori*), and that also includes genes encoding whatever plasmid-encoded RNAs and proteins are necessary for replication. The selectable marker, necessary for following and maintaining the presence of the plasmid in cells, is usually dominant and is usually a gene encoding resistance to some

antibiotic. The cloning site is a restriction endonuclease cleavage site into which foreign DNA can be inserted without interfering with the plasmid's ability to replicate or to confer the selectable phenotype on its host. Over the years, plasmid vectors have increased in sophistication; in addition to these three basic elements many vectors now include features that make them particularly suitable for specific types of experiments.

UNITS 1.6 & 1.7 describe procedures for making plasmid DNA. The process by which plasmids are introduced into *E. coli* is called **transformation**. Transformation protocols are given in UNIT 1.8. For a list of vectors and their salient features see APPENDIX 5.

REPLICATORS

One of the ways that replicators are classified is based on the number of plasmid molecules maintained per bacterial cell under some set of standard growth conditions, the so-called **copy number** of the plasmid. This book defines **high-copy-number plasmids** as those which exist in ≥ 20 copies per bacterial cell grown in liquid LB medium, and **low-copy-number plasmids** as those which exist in < 20 copies per cell. High-copy-number-plasmids tend to be smaller than low-copy-number plasmids. They are more commonly used in molecular biological techniques since it is easier to prepare large quantities of pure plasmid DNA from cells that bear them. Low-copy-number plasmids are utilized when it is important to control the gene dosage of a cloned sequence, for example when the DNA sequence or the protein it encodes makes the host organism sick (see Table 1.5.1).

High-copy-number plasmids tend to be under **relaxed control** of replication. These relaxed plasmids initiate DNA replication in a process controlled by plasmid-encoded functions (see Mechanism of Replication and Copy-Number Control), and replication does not depend on the unstable host replication initiation proteins synthesized at the start of the bacterial cell cycle. Because their replication depends only on the stable host enzymatic replication machinery, the copy number of these plasmids can be greatly increased, or amplified, by treatment of the plasmid-containing cells with protein synthesis inhibitors such as chlorampheni-

Table 1.5.1 Characteristics of Commonly Used Plasmid Replicators

Replicator	Prototype plasmid	Size (bp)	Markers on prototype	Copy ^a number	References
pMB1	pBR322	4,362	Ap ^r , Tet ^r	High; 100-300	Bolivar et al., 1977
ColE1	pMK16	~4,500	Kan ^r , Tet ^r , ColE1 ^{imm}	High; >15	Kahn et al., 1979
p15A	pACYC184	~4,000	Eml ^r , Tet ^r	High; ~15	Chang et al., 1978
pSC101	pLG338	~7,300	Kan ^r , Tet ^r	Low; ~6	Stoker et al., 1982
F	pDF41	~12,800	TrpE	Low; 1 to 2	Kahn et al., 1979
R6K	pRK353	~11,100	TrpE	Low; <15	Kahn et al., 1979
R1 (R1 <i>drd-17</i>)	pBEU50	~10,000	Ap ^r , Tet ^r	Low at 30°C; high above 35°C ^b	Uhlin et al., 1983
RK2	pRK2501	~11,100	Kan ^r , Tet ^r	Low; 2 to 4	Kahn et al., 1979
λ <i>dv</i>	λ <i>dvgal</i>	— ^c	Gal	—	Jackson et al., 1972

^aCopy numbers are for the prototype plasmid. Plasmid vectors that contain replicators derived from these plasmids may have different copy numbers due to introduction of mutations into the replicator. For example, pUC series (pMB1-derived) has copy numbers of 1000-3000.

^bTemperature sensitive.

^cNot known.

col or spectinomycin. These protein synthesis inhibitors prevent synthesis of the replication-initiation proteins required for chromosomal replication, allowing the replication enzymes to be commandeered for plasmid replication. High-copy-number plasmids usually do not have any mechanism to ensure correct segregation of the plasmids to the daughter cells; they rely on random assortment to partition at least one copy of the plasmid to each daughter.

Low-copy-number plasmids are usually under *stringent control*. Initiation of replication of these plasmids depends on unstable proteins synthesized at the start of the bacterial cell cycle and thus is synchronized with the replication of the bacterial chromosome. As copy number decreases, random segregation of plasmid copies is not sufficient to ensure that each daughter cell acquires a copy of the plasmid. However, most low-copy-number plasmids carry genes that guarantee their maintenance in the bacterial population. Stabilizing loci that encode a mechanism for postsegregational killing of plasmid-free daughter cells and a system for efficiently resolving plasmid multimers so that they can be appropriately partitioned have been identified. In addition, *cis*- and *trans*-acting partitioning loci (*par*) that are thought to constitute an active mechanism for distribution of plasmid copies to daughter cells have also been identified.

MECHANISM OF REPLICATION AND COPY-NUMBER CONTROL

While there are many different replicators, the majority of plasmid vectors used in routine recombinant DNA work contain one of the functionally similar replicators derived from plasmids ColE1 or pMB1 (for example, the popular pBluescript series contain a ColE1 origin, and the pUC plasmids are derived from pMB1). The ColE1 replicator is a 600-nucleotide DNA fragment that contains the origin of replication (*ori*), encodes an RNA primer, and encodes two negative regulators of replication initiation. All enzymatic functions for replication of the plasmid are provided by the bacterial host.

Plasmid replication begins with transcription of an RNA primer upstream of the *ori* (RNAII; see Fig. 1.5.1) by the host RNA polymerase. RNAII is elongated through and terminated downstream of the *ori*. Interaction of a specific secondary structure in the nascent RNAII transcript with the DNA template results in formation of a persistent hybrid between RNAII and the DNA template such that RNAII remains paired with the DNA template at the *ori*. The RNAII transcript is cleaved by RNaseH at the *ori* sequence. This processed RNAII primer is extended by DNA polymerase I to initiate plasmid replication. Regulation of the proper RNAII secondary structure controls initiation of DNA replication and is responsible

for determining the number of plasmid molecules per cell.

Both ColE1 and pMB1 plasmids are high-copy-number plasmids, maintained at between 15 and 25 copies per bacterial cell, respectively. The copy number of these plasmids is regulated by an antisense RNA transcript, RNAI, and the protein ROP, the product of the *rop* gene. RNAI and the ROP protein act in concert to intercept formation of the proper RNAII secondary

structure. RNAI is exactly complementary to the 5' end of the RNAII transcript. RNAI base pairs with the 5' end of RNAII, preventing the formation of the specific secondary structure necessary for establishment of a persistent hybrid between RNAII and the DNA template that is a prerequisite for maturation of the RNAII primer. ROP protein stabilizes the interaction between the antisense RNAI and the primer RNAII. Mutations that disrupt or destabilize

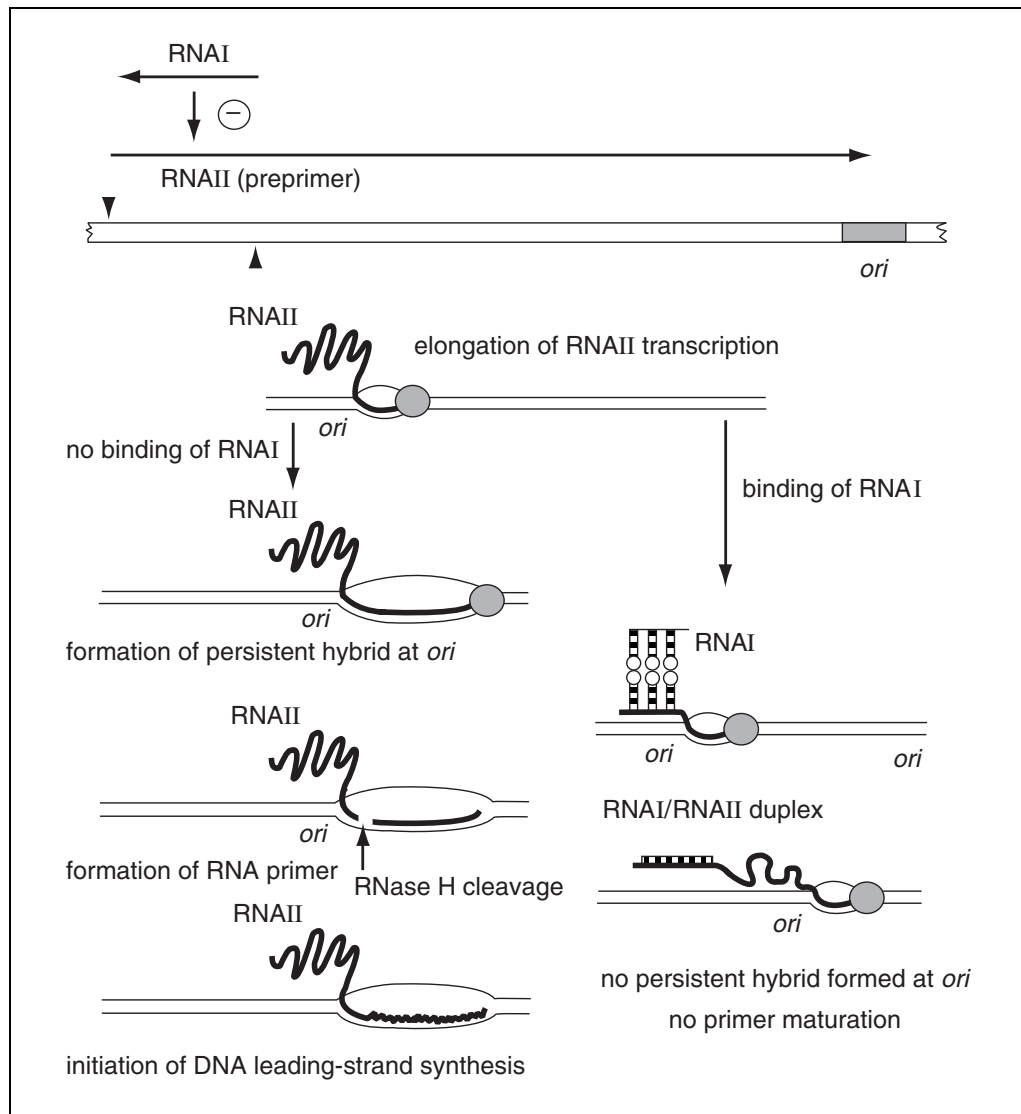


Figure 1.5.1 ColE1 replication initiation and copy number control. ColE1 replication requires the formation of a RNA primer to initiate DNA synthesis. The RNA primer is derived from processing of a transcript, RNAII (bold), that starts upstream of the *ori*. Appropriate processing of the RNAII transcript is dependent on formation of a persistent hybrid between RNAII and the *ori*/DNA template. If the proper RNAII secondary structure forms, then the RNA/DNA duplex is maintained at the *ori* and RNase H can cleave the RNAII transcript to generate the primer for DNA synthesis. Processing of RNAII to form the primer is regulated by a second transcript, RNAI. RNAI is complementary to the 5' end of RNAII. If the RNAI transcript forms a duplex with RNAII, RNAII cannot take on the secondary structure necessary to create the persistent hybrid at the *ori* and maturation of the RNAII primer is inhibited. The copy number of ColE1 plasmids is determined by the balance between successful RNAII processing events and those inhibited by RNAI. Adapted from Gerhart et al. (1994) with permission from the *Annual Review of Microbiology*.

the RNAI/RNAII interaction result in a higher plasmid copy number. For example, the pUC plasmid series have pMB1 replicators that do not contain an intact *rop* gene, which accounts for a 2-fold increase in copy number over plasmids with intact pMB1 replicons. Interestingly, additional mutation(s) in the origin are the main factor in the ultra-high-copy number (1000 to 3000) of this plasmid series (K. Struhl, unpub. observ.).

PLASMID INCOMPATIBILITY

For experiments that require that more than one plasmid vector be maintained in a bacterial cell at the same time, another critical feature of plasmid replicators is whether or not two plasmid replicators are *compatible*. Two plasmids are said to be *incompatible* with one another, and hence belong to the same *incompatibility group*, if they cannot stably co-exist. Plasmids are generally incompatible if they share any function required for the regulation of plasmid replication. For example, ColE1- and pMB1-derived plasmids are incompatible with one another but are compatible with p15A plasmids. The incompatibility of ColE1 and pMB1 plasmids is a consequence of two facts: first, that plasmid DNA replication of these plasmids is negatively controlled by RNAI which acts in *trans* on other plasmids with the same primer RNA, and second, that these plasmids lack a mechanism to ensure that each plasmid in a cell replicates once per cell cycle. Therefore, if a cell that contains a pMBI-derived plasmid is subsequently transformed with a ColE1-derived plasmid, cells selected to contain the ColE1 plasmid will usually have lost the pMBI plasmid.

SELECTABLE MARKERS

In order to guarantee that a plasmid vector is taken up by and maintained in bacterial cells, there must be a way to select for plasmid-containing cells. Genes encoding resistance to antibiotics such as ampicillin, tetracycline, kanamycin and chloramphenicol are the most common bacterial selectable markers for plasmid vectors. Typically, cells are transformed with plasmid DNA using the technique described in *UNIT 1.8* and then plated out on LB plates that contain the proper antibiotic (see recipes in *UNIT 1.1*). Only the bacterial cells containing the plasmid will grow on the selective medium; the antibiotic-resistance phenotype conferred is dominant to the antibiotic-sensitive phenotype of cells that do not possess the plasmid vector. Another dominant select-

able marker that is occasionally used is the immunity to infection by phage lambda (lambda repressor). In addition, recessive markers are sometimes used in plasmid selections; for example, *leuB⁻ E. coli* cannot grow in the absence of leucine, and selection for growth of these strains in the absence of leucine allows isolation of colonies transformed with a plasmid that contains a gene that complements *leuB*.

For experiments that involve introduction of plasmid vectors into systems other than bacteria, such as yeast or mammalian cells, it is usually necessary to select for the presence of plasmid DNA in these hosts. Selectable markers for yeast and mammalian systems are generally distinct from those used in *E. coli* (see Plasmid Vectors for Yeast and Plasmid Vectors for Expression in Cultured Mammalian Cells). In cases where a vector is needed that will be used in both *E. coli* and another host, it is necessary for the plasmid vector to carry selectable markers for both hosts.

CLONING SITE

Today's plasmid vectors contain a multiple cloning site (MCS) or *polylinker* cloning region that can include >20 tandemly arranged restriction endonuclease sites. The sites in the polylinker are almost always designed to be unique within the vector sequence so that cutting the vector with a restriction endonuclease in the polylinker and cloning foreign DNA into this site does not disrupt other critical features of the vector. The plethora of sites in the polylinker ensures that the appropriate enzyme sites will be available for cloning most DNA fragments, provides unique reference restriction sites for rapid restriction mapping of the insert, and generally allows for a great deal of flexibility when manipulating the cloned DNA.

The sequences that directly flank the polylinker site are often useful for manipulation or analysis of insert DNA. Many polylinker sites are flanked by sequences for which there are commercially available complementary oligonucleotides, for example the M13 reverse, -20, and -40 primers, that can be used for priming polymerase chain reactions (PCR) or DNA sequencing reactions. Such primers are useful tools for amplification or sequencing of any DNA fragment inserted into the polylinker. Some polylinkers are bordered by 8-bp-cutter restriction sites, like *Not1*. These sites occur infrequently in DNA and thus allow for the easy excision of an intact insert fragment from the plasmid vector.

Many plasmid vectors have been developed with bacteriophage—SP6, T7, or T3—promoters flanking the polylinker cloning sites. These promoters can be used *in vitro* or *in vivo*, if the bacteriophage RNA polymerase is also present, to produce large quantities of RNA transcripts from DNA inserted into the polylinker.

In order to make it easier to identify plasmids that contain insert DNA, the polylinkers of some vectors have been engineered so that introduction of DNA into the polylinker results in a scorable phenotype. The most common example of this is insertion of the polylinker of many basic cloning vectors, the pUC series for example, into the *lacZ* gene α fragment. In the appropriate genetic background, production of the *lacZ* α fragment allows for formation of an active β -galactosidase enzyme which results in the formation of blue colonies on Xgal/IPTG indicator plates. Cloning into these polylinkers prevents production of a functional *lacZ* α fragment, allowing for rapid identification of plasmids containing inserts as white colonies on Xgal/IPTG indicator plates. Vectors have also been developed that allow for direct selection of plasmids containing inserts by location of the polylinker in the *ccdB* (control of cell death) gene which causes cell death in *E. coli*. Disruption of the *ccdB* gene by introduction of an insert into the polylinker allows the cells to survive, and thus only recombinants will grow under conditions where the *ccdB* gene is expressed.

CHOOSING A PLASMID VECTOR

Some of the basic features to consider when selecting a plasmid vector include the size of the vector, its copy number, the polylinker, and the ability to select and/or screen for inserts. Large plasmids, >15 kb, do not transform well and frequently give lower DNA yields. Consider the final size of the vector plus insert when planning an experiment and wherever possible use smaller vectors. The higher the copy number the more vector DNA is produced, but high-copy-number vectors may not be applicable to all situations (see Mechanism of Replication and Copy-Number Control). When choosing a plasmid vector, consider both what sites are present in the polylinker and the order of the sites. If it will be necessary to manipulate the cloned sequence subsequent to insertion into the polylinker, plan ahead to ensure that the necessary sites will remain available in the polylinker. Selections or screens for identification of recombinant clones are useful for experiments where cloning efficiency is expected

to be low or when generation of a large number of clones is necessary. However, for routine subcloning experiments the advent of PCR (UNIT 15.1) has made it possible to rapidly screen through a large number of transformants to identify potential recombinant molecules, obviating the need for histochemical and genetic screening methods.

The primary factor in choosing a plasmid vector is to understand and anticipate the experiments for which the recombinant clone will be used. The specialized functions of plasmid vectors are generally the key to selecting the correct vector for an experiment. For example, completely different types of vectors would be used for generating large quantities of DNA, expressing a fusion protein in bacteria, or for undertaking a two-hybrid screen in yeast. Once the type of vector required is determined then deciding upon a particular vector is dependent on both the details of the ancillary vector features, for example the type of promoter used to express a recombinant protein, and the properties of the replicator, polylinker, and selectable marker.

PLASMID VECTORS FOR PRODUCTION OF SINGLE-STRANDED DNA

Plasmids have been developed that contain a filamentous phage origin of replication in addition to a plasmid *ori*. These “phagemid” vectors (UNIT 1.14) can be grown and propagated as plasmids. However, upon super-infection of a plasmid-containing cell with a wild-type helper phage, the phage *ori* becomes active, and single-stranded DNA (ssDNA) is produced and secreted. There are usually (+) and (–) versions of these vectors where the phage *ori* is in opposite orientation so that it is possible to produce ssDNA from either DNA strand. For many years, ssDNA was the substrate of choice for DNA sequencing (UNIT 7.4) and oligonucleotide-directed mutagenesis (UNIT 8.1). Development of sequencing and mutagenic protocols that use double-stranded templates has made the production of ssDNA a less frequently utilized feature of plasmid vectors. The pBluescriptI, pBluescriptII, and pBS phagemid vectors derived from the general cloning vector pUC19 (Fig. 1.5.2) are examples of phagemids that incorporate the f1 filamentous phage *ori*.

PLASMID VECTORS FOR CLONING LARGE INSERTS

Cosmid vectors, plasmids carrying a lambda phage *cos* site (e.g., pWE15, Fig. 1.5.3), were

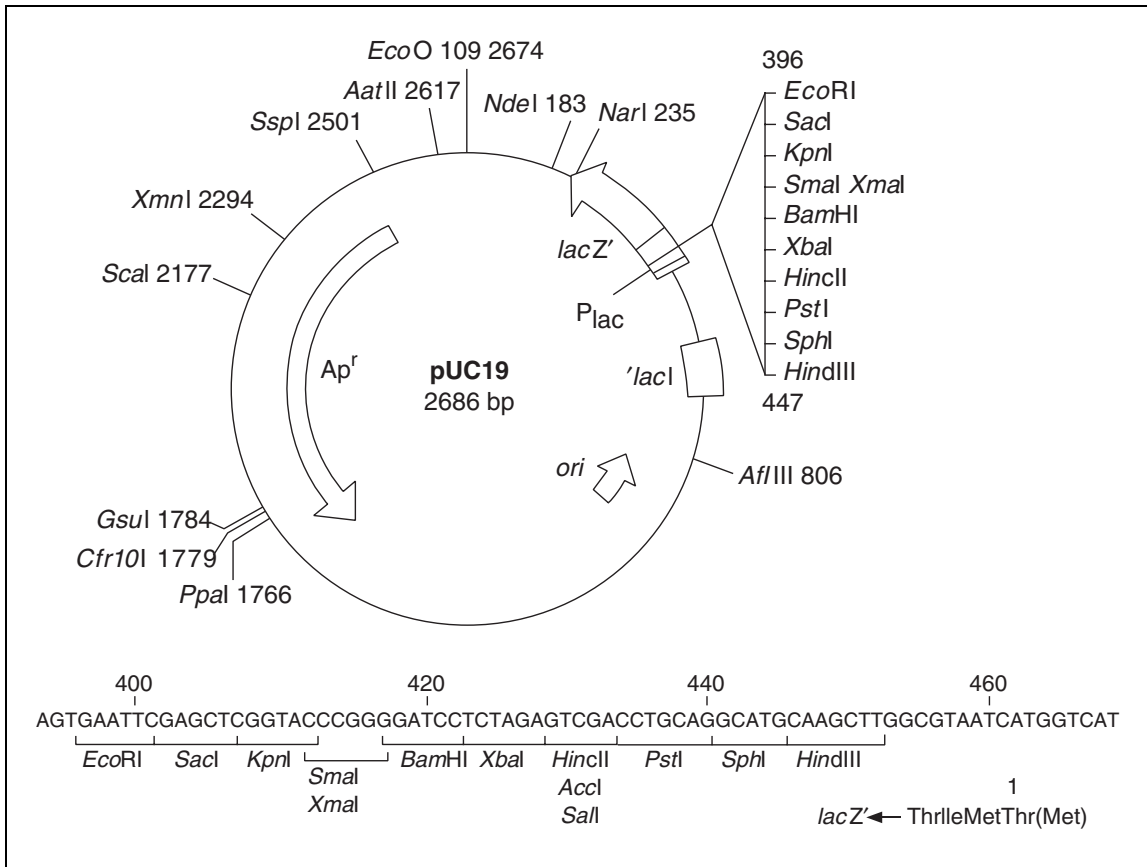


Figure 1.5.2 Map of pUC19.

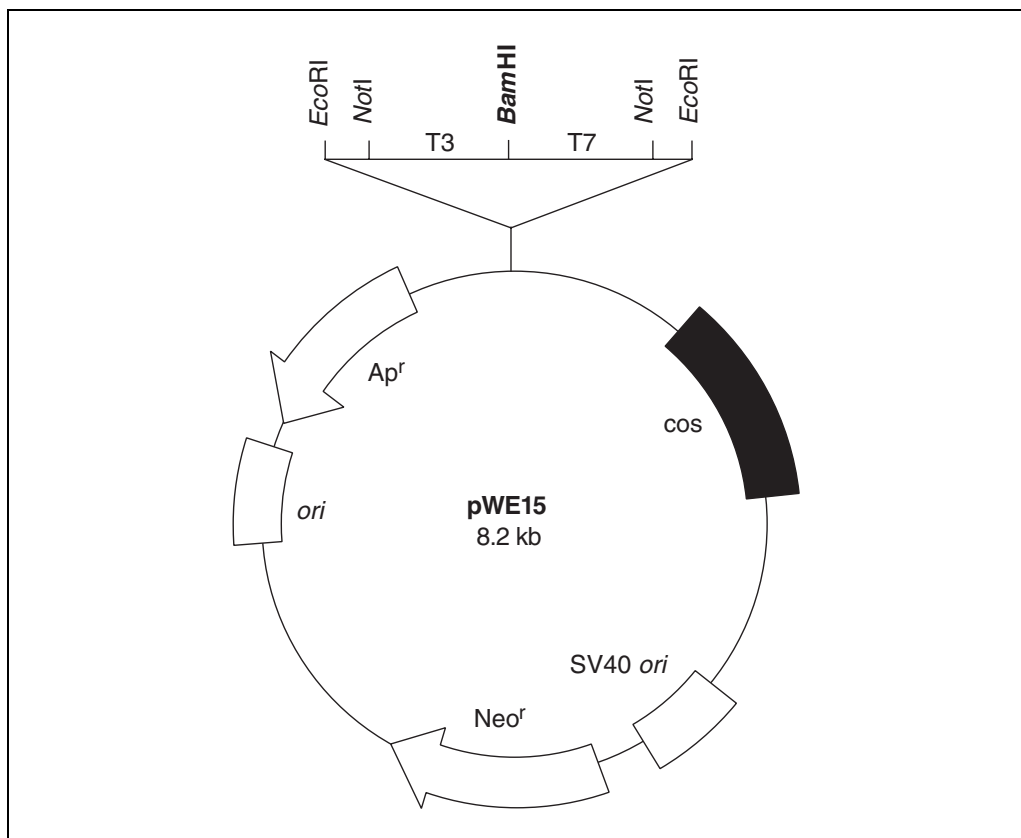


Figure 1.5.3 Map of pWE15 (adapted from Wahl et al., 1987, with permission).

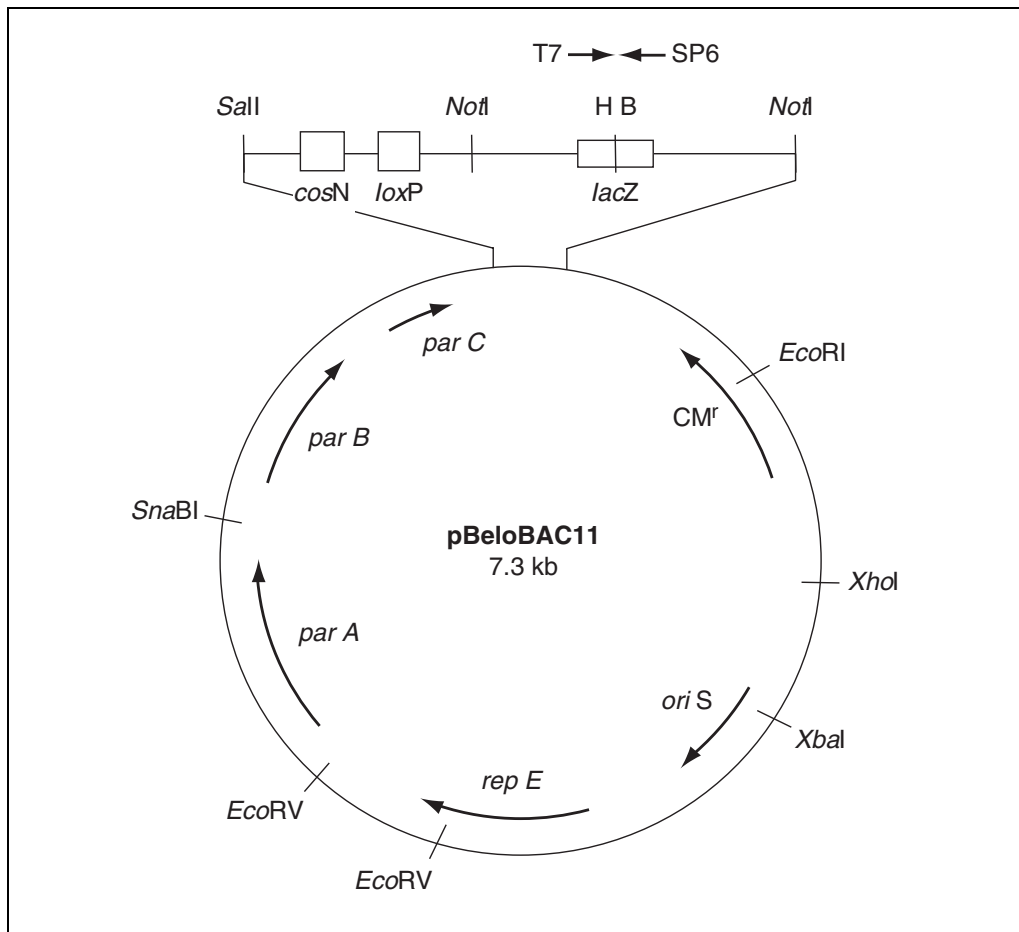


Figure 1.5.4 Map of pBeloBAC11 (adapted from Shizuya et al., 1992, with permission). Abbreviation: CM, chloramphenicol.

developed to facilitate cloning of large DNA fragments (UNIT 1.10). Cosmids can be transformed into cells like plasmids and once in the cells, replicate using their plasmid *ori*. The ColE1 type replicators are the most commonly used in these vectors, and cosmids can generally be maintained at high-copy-number in *E. coli*. Cosmids can also be packaged into lambda phage heads. In order for packaging to occur the *cos* sites must be separated by 40 to 50 kb, the approximate length of the wild-type lambda genome. Many cosmid vectors are between 5 to 10 kb in size and therefore can accept inserts of 30 to 45 kb. Because of the size of inserts they can accept, the high efficiency of packaging recombinant molecules into phage, and the efficiency of infection of cosmid-containing phage heads, cosmid vectors are frequently used for making genomic libraries. Unfortunately, propagation of insert-containing cosmid vectors in *E. coli* sometimes results in the deletion of all or a portion of the insert. To address this problem, a new set of cosmid vectors have been developed that replace the ColE1 replica-

tor with the F factor replicator. These “fosmid” vectors are maintained at low-copy-number, 1 to 2 copies per cell, and are more stable than higher-copy-number cosmid vectors when grown in *E. coli* (Kim et al., 1992).

Another type of plasmid cloning vector, called bacterial artificial chromosome (BAC; Fig. 1.5.4), has been developed using the F factor replicator for propagation of very large pieces of DNA (100 to 500 kb). The vectors are used in a similar manner to yeast artificial chromosome (YAC) vectors but have the advantage of being manipulated solely in *E. coli*.

PLASMID VECTORS FOR EXPRESSION OF LARGE QUANTITIES OF RECOMBINANT PROTEINS

There are a wide variety of vectors for expressing high levels of recombinant proteins in *E. coli*, insect, and mammalian cells (Chapter 16). The general goal of expressing proteins in any of these systems is to produce large quantities of a particular protein upon demand.

While the features of the protein expression systems vary considerably, the basic properties outlined for expressing proteins in *E. coli* are common to all of them. Expression vectors are usually designed such that production of the foreign protein is tightly regulated. This is necessary because the host cellular machinery is co-opted to produce large quantities of the foreign protein, the sheer amount of which may be toxic to the cell, and/or the foreign protein may encode a function that will inhibit cell growth or kill the host cell. Generally, expression vectors are configured so the polylinker cloning site is downstream of an inducible promoter. One of the promoters commonly used in *E. coli* expression vectors is the hybrid *trp/lac* promoter (*trc*) which contains the *lacO* operator site. This promoter is turned off in the presence of the *lacI^q* repressor; the repressor gene is either carried by the bacterial host or is encoded on the expression vector itself. Expression of the foreign protein from the *trc* promoter is induced by the addition of IPTG.

The quantity of foreign protein produced will be determined by both the rate of transcription of the gene and the efficiency of translation of the mRNA. Therefore, in addition to regulated highly inducible promoters, many *E. coli* protein expression vectors are designed to optimize translation of the foreign protein in bacterial cells. These expression vectors include an efficient ribosome binding site and an ATG start codon upstream of the polylinker cloning site. Usually the cloning sites in the polylinker are designed so that it is possible to make an in-frame fusion to the protein of interest in all three reading frames.

For many experiments, high levels of pure recombinant protein are required, and some expression vectors are designed to create tagged or fusion proteins that facilitate purification of the recombinant protein. Tag sequences may be located 5' or 3' to the polylinker, creating either amino- or carboxy-terminal-tagged fusion proteins. Six polyhistidine residues (*UNIT 10.11B*), the FLAG epitope, and glutathione-S-transferase protein (*UNIT 16.7*) are some of the sequences that are appended to proteins to assist in purification. Tagged or fusion proteins can be rapidly and efficiently purified using an appropriate affinity column designed to tightly bind the tag or fusion region. Many protein expression vectors are created with specific protein cleavage sites adjacent to the tag or fusion sequences to allow for removal of these sequences from the purified protein. This feature may prove to be essential if the tag

or fusion sequence impairs the function of the protein in the relevant assays.

PLASMID VECTORS FOR REPORTER GENE FUSIONS

Plasmid vectors have been designed to simplify the construction and manipulation of reporter gene fusions, where a promoter of interest is used to drive an easily scored marker gene (*UNIT 9.6*). Gene fusions provide a rapid and simple method for following the expression pattern conferred by a particular promoter. There are generally two types of reporter fusions, transcriptional and translational fusions. For transcriptional fusions, the ATG start codon is provided by the marker gene. In translational fusions, the 5' untranslated region and ATG are provided by the gene of interest; in fact these constructs may fuse a large portion, or even entire coding region, to the amino terminus of a marker. In the reporter vector, the polylinker cloning site is located directly upstream of the reporter gene for insertion of the promoter fragment. In vectors that are designed for expression in eukaryotic cells, a polyadenylation signal is located downstream of the reporter gene. There are a variety of reporter genes used including chloramphenicol acetyltransferase, luciferase, β -galactosidase, secreted alkaline phosphatase, human growth hormone, β -glucuronidase, and green fluorescent protein (also see *UNITS 9.6-9.7C*). An example of a reporter vector pEGFP, is shown in Figure 1.5.5. The choice of reporter gene will depend on the cell type or organism, whether the assay will be done in vivo or in vitro, and whether quantitative or qualitative data is desired. The replicator, selectable marker, and other elements of the reporter vector will also have to be compatible with the system.

PLASMID VECTORS FOR YEAST

Yeast plasmid vectors contain the same basic features as *E. coli* vectors—replicator, selectable marker, and cloning site (*UNIT 13.4*). There are two primary types of replicators used in yeast plasmid vectors, autonomously replicating sequences (ARS) derived from the yeast chromosomes and the natural yeast 2 μ m plasmid replicator. ARS-containing plasmids frequently contain yeast centromeric sequences to ensure their stable maintenance in the population of cells. Both ARS and 2 μ m vectors average 10 to 30 copies per cell. Many yeast vectors are “shuttle” vectors that can be maintained in both *S. cerevisiae* and *E. coli* (e.g., pRS303; see Fig. 1.5.6). These vectors have an *E. coli* plas-

mid replicator, frequently pMB1 derived, and a yeast replicator. Conversely, the yeast integrating plasmid vectors, used for introduction of genes into the yeast chromosome, have a bacterial replicator but no yeast replicator.

The selectable markers in yeast are for the most part recessive markers, usually cloned yeast genes that are used to complement muta-

tions in a biosynthetic pathway. For example, the *URA3* marker carried on a plasmid is used to restore the ability to grow in the absence of uracil to a *ura3⁻* yeast. Therefore, yeast selectable markers, unlike most bacterial selectable markers, are strain dependent.

There are a wide variety of yeast vectors available that have a range of specialized func-

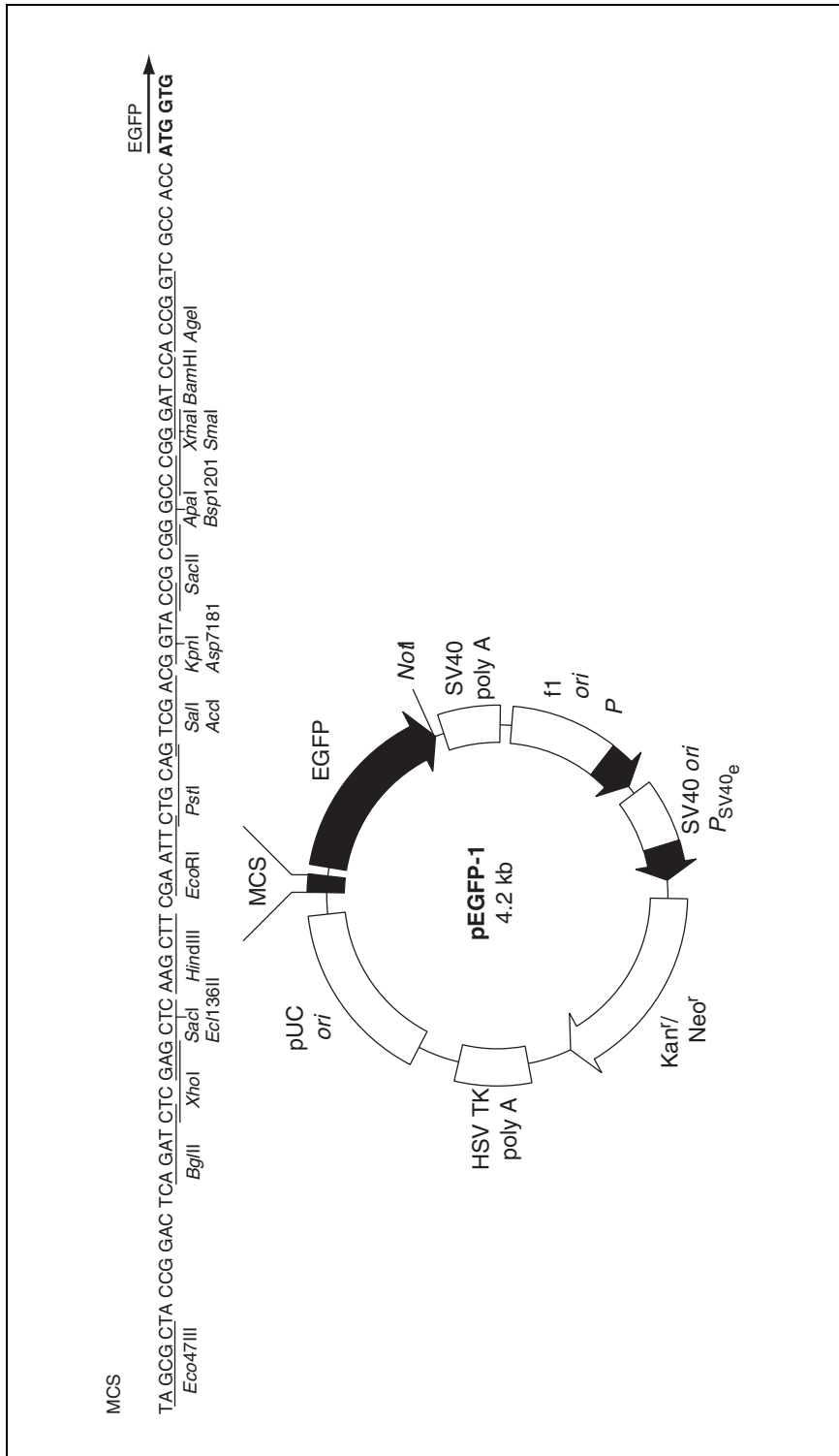


Figure 1.5.5 Map of pEGFP-1.

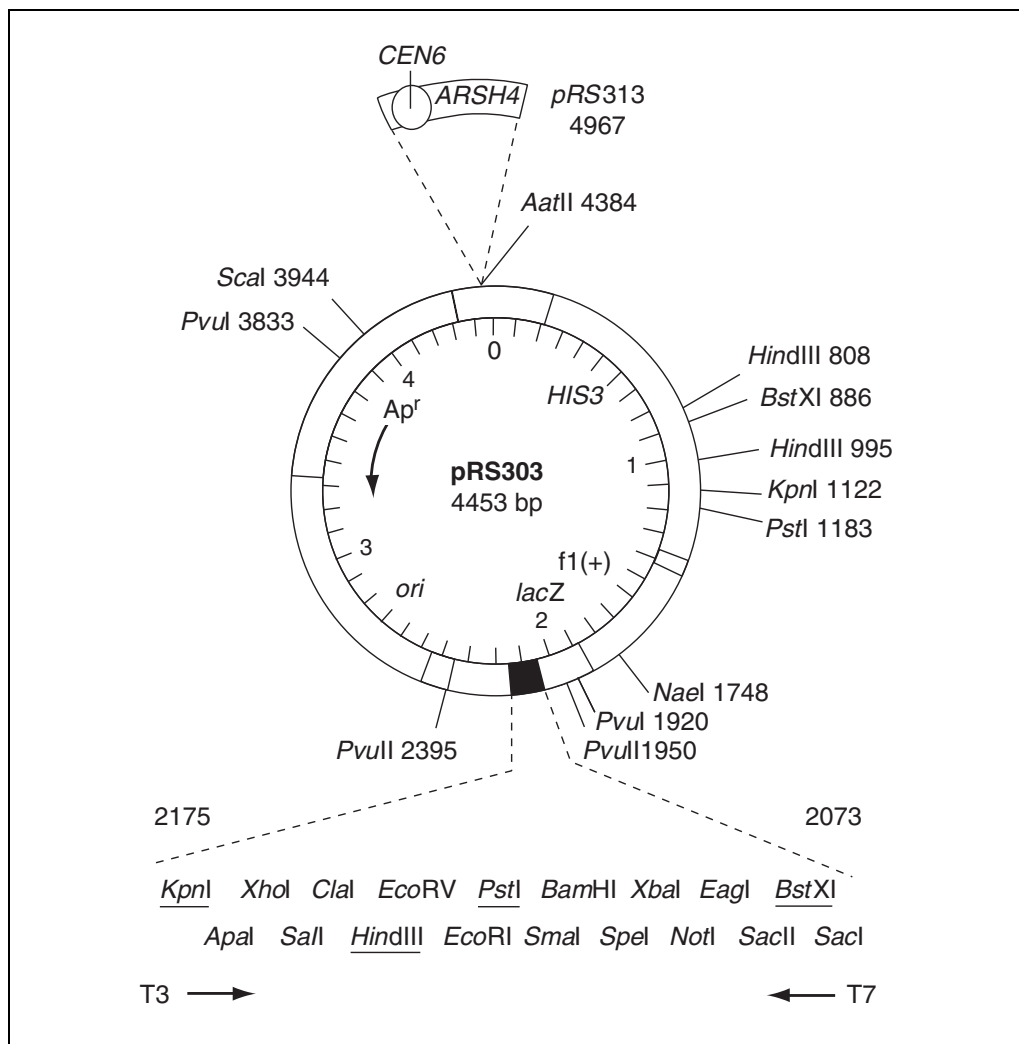


Figure 1.5.6 Map of pRS303 (adapted from Sikorski and Heiter, 1989, with permission).

tions—e.g., expression of recombinant proteins in yeast, integration of sequences into the yeast genome, and cloning of very large fragments (hundreds of kilobases) of genomic DNA.

PLASMID VECTORS FOR EXPRESSION IN CULTURED MAMMALIAN CELLS

The type of vector that is used in mammalian cells depends on whether the experiment involves transient transfection into mammalian cells or the generation of stable mammalian cell lines carrying the construct of interest (UNITS 9.1-9.5). Virtually any plasmid vector that contains an appropriate construct for expression in mammalian cells can be used in transient assays. In transient assays, the plasmid vector carrying the DNA of interest is transfected into mammalian cells, the cells are harvested some time later (24 to 96 hr), and the pertinent assay

is performed. It is not necessary for plasmid vectors used in these assays to carry a mammalian selectable marker or to replicate in mammalian cells; therefore easily manipulatable bacterial plasmids, like pUC, are usually the vectors of choice.

Plasmid vectors carrying a selectable marker that functions in mammalian cells are necessary for the generation of stable transgenic lines (UNITS 9.5; e.g., pcDNA3.1, Fig. 1.5.7). In order to generate a stable mammalian cell line, the plasmid DNA is transfected into the mammalian cells, and over a period of several weeks the DNA of interest is selected for based on expression of the vector-borne marker. Many mammalian vectors cannot replicate in mammalian cells, and the only way to maintain the DNA of interest and the selectable marker is for the vector to randomly integrate into the mammalian genome. However, there are some plasmid vectors that carry the simian

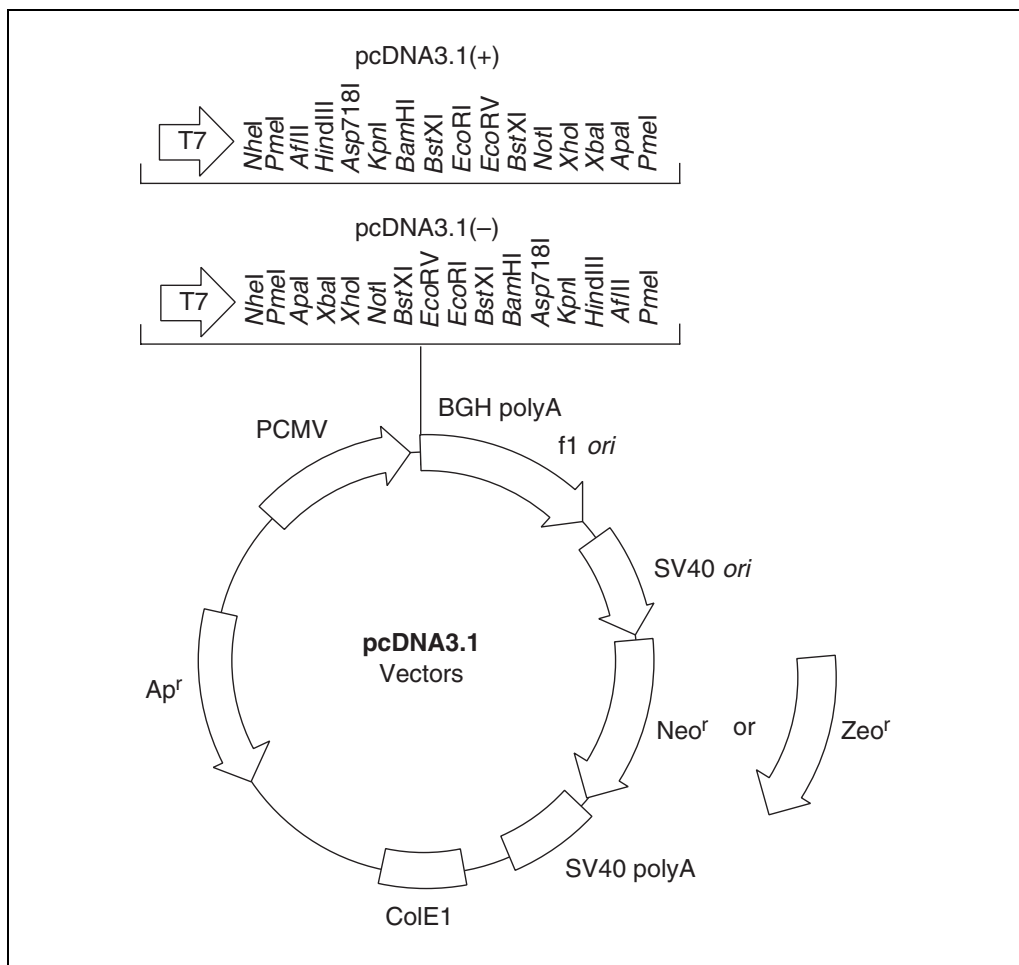


Figure 1.5.7 Map of pcDNA3.1.

virus (SV40) or bovine papilloma virus (BPV) *ori* that can replicate in mammalian cells if the necessary viral replication proteins are provided; in most cases the viral replication proteins must be provided by the host cell line, limiting the range of cell types in which these vectors are useful.

The most critical general feature of plasmid vectors (and also viral vectors) for creation of stable mammalian cell lines is the selectable marker (see *UNIT 9.5*). Unlike transient assays, formation and maintenance of stable cell lines requires selection. The number of selectable markers available for mammalian cells is limited; resistance to hygromycin, puromycin, G418, and neomycin are the predominant markers used. Since expression of foreign proteins or assaying the expression of a mammalian gene may require multiple plasmids and/or integration of constructs into the mammalian chromosome, careful planning must take place to ensure that all constructs can be selected for with the limited number markers. Further, since selection may be necessary over a long period

of time—weeks for generation of the lines and years for their maintenance—it is important to recognize that some of the antibiotics used in selection are very expensive and thus cost may be a factor in experimental design.

Specialized plasmid vectors are also used for production of viruses that can infect mammalian cells. Plasmid vectors have been designed to produce infectious retroviral particles when transfected into the appropriate packaging cell line (*UNIT 9.9*). These retroviral vectors are then used to create stable transgenic lines in mammalian cell types not amenable to transfection. In addition, plasmid vectors have been designed to allow easy insertion of DNA sequences into vaccinia virus for the purpose of creating recombinant viruses that overexpress recombinant proteins (*UNIT 16.15*).

PLASMID VECTORS FOR NON-*E. COLI* BACTERIA

Three features required of all bacterial plasmid vectors are that they replicate (unless they are suicide vectors), carry a selectable marker,

Escherichia coli,
Plasmids, and
Bacteriophages

1.5.11

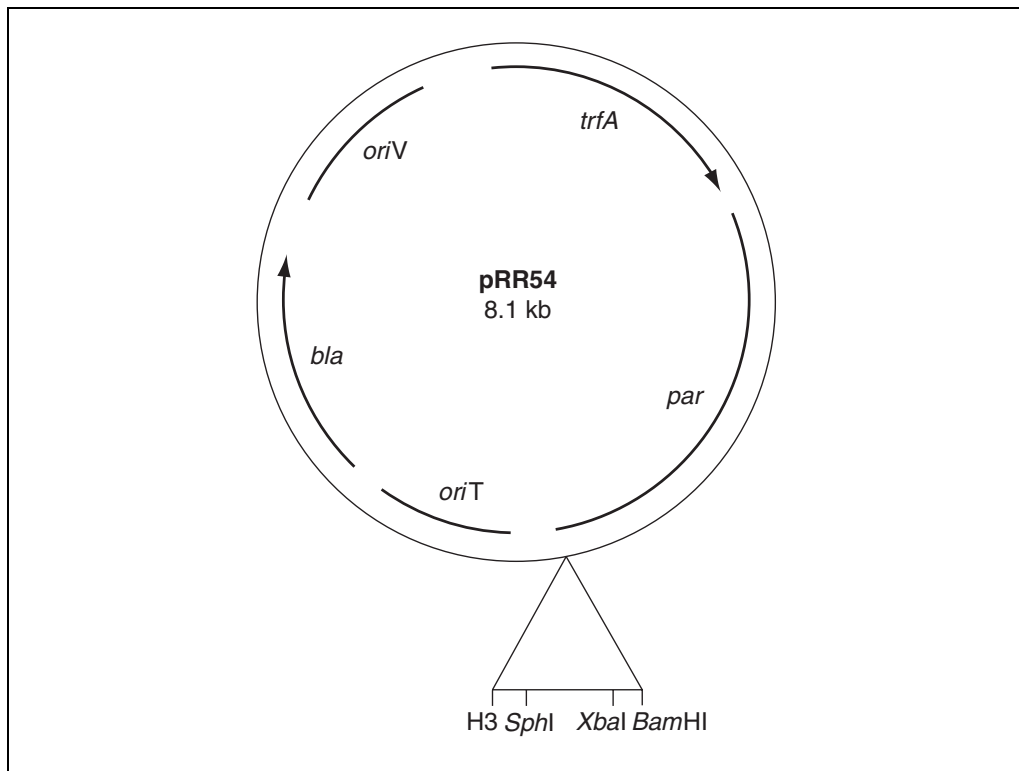


Figure 1.5.8 Map of pRR54 (adapted from Roberts et al., 1990, with permission).

and can be easily introduced into host cells. The first thing to consider when selecting a plasmid vector for use in a non-*E. coli* host is whether or not it can replicate and be stably maintained in the particular strain. Different plasmid replicators have different host ranges, some have a narrow host range and can only replicate in a specific strain while others are promiscuous and can replicate in a wide variety of host (e.g., pRR54, Fig. 1.5.8). Unfortunately, the ColE1-type replicators have a narrow host range, thus many standard *E. coli* vectors cannot be maintained in other bacteria. However, a number of broad-host-range replicators, such as RK2 and RSF1010, have been well characterized and used to construct vectors that can replicate in many gram-negative bacterial species (and in the case of RSF1010, some gram-positive species as well).

Antibiotic resistance genes are used as selectable markers in non-*E. coli* bacterial hosts; however, the quantity of an antibiotic used to select against non-plasmid-containing cells is usually higher than the quantity used for *E. coli* selection. Furthermore, some bacterial strains are inherently resistant to particular antibiotics, thus it is important to determine whether the selectable marker carried on a plasmid vector is functional in a particular strain.

The favored method for introduction of plasmid DNA into bacterial host cells varies widely with the bacterial strain. Unlike *E. coli*, many bacteria cannot be efficiently transformed by chemical procedures or electroporation. In these cases bacterial mating is used to introduce plasmid DNA into the desired bacterial host. Mating to transfer a plasmid vector from an *E. coli* host where the vector is maintained and manipulated to a recipient bacterium requires both *cis*- and *trans*-acting functions. The *tra* (or *mob*) genes encode the *trans*-acting proteins necessary to transfer the plasmid DNA from one bacterium to another, and they are usually located on a helper plasmid that is distinct from the plasmid vector. Any plasmid vector that is to be mobilized must contain the *cis*-acting site called *oriT* where the DNA is cleaved and transfer is initiated.

MAPS OF PLASMIDS

Figures 1.5.2-1.5.11 present maps of plasmids that are in widespread use or are examples of plasmids whose special functions make them useful for particular techniques described in this manual. Note that the trend in development of vectors is to include multiple features on a single vector, and many of these examples span the vector categories described.

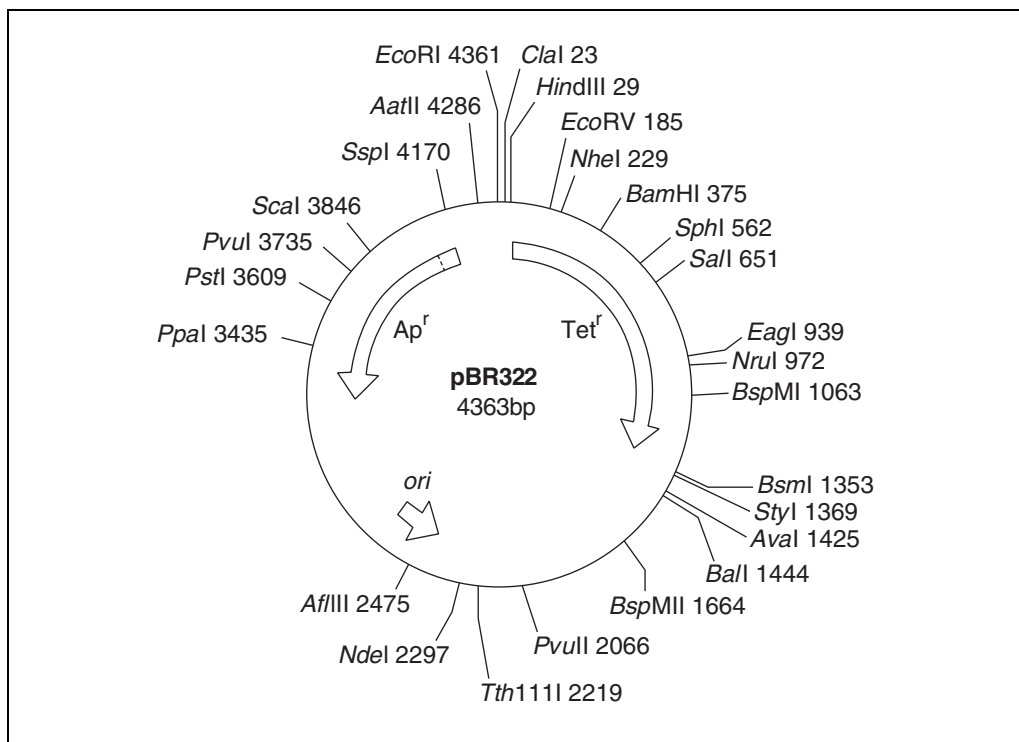


Figure 1.5.9 Map of pBR322 (Bolivar et al. 1977; sequence in Sutcliffe, 1978).

pBR322 is one of the classic cloning vectors from which many other vectors are derived. It contains an amplifiable pMB1 replicator and genes encoding resistance to ampicillin and tetracycline. Insertion of DNA into a restriction site in either drug-resistance gene usually inactivates it and allows colonies bearing plasmids with such insertions to be identified by their inability to grow on medium with that antibiotic (see Fig. 1.5.9; Bolivar et al., 1977; sequence in Sutcliffe, 1978).

pUC19 belongs to a family of plasmid vectors that contains a polylinker inserted within the alpha region of the *lacZ* gene. The polylinkers are the same as those used in the m13mp series (Fig. 1.14.2). pUC19 and pUC18 have the same polylinker but in opposite orientations. Under appropriate conditions (see UNIT 1.4 for a description), colonies that bear plasmids containing a fragment inserted into the polylinker form white colonies instead of blue ones. These pMB1-derived plasmids (see Fig. 1.5.2) maintain a very high-copy-number (1000 to 3000 per genome). Wild-type and recombinant plasmids confer ampicillin resistance and can be amplified with chloramphenicol (Norrandar et al., 1983). In addition wild-type plasmids confer a LacZ⁺ phenotype to appropriate cells (e.g., JM101 cells, UNIT 1.4).

pBluescript is a commonly used phagemid cloning vector that contains a polylinker in-

serted into the alpha region of the *lacZ* gene and T3 and T7 promoter sequences flanking the cloning sites. The f1 (+) filamentous phage origin of replication in pBluescript SK⁺ allows for the recovery of the sense strand of the *lacZ* gene as ssDNA; the pBluescript SK(-) vector (Fig. 1.5.10) with the f1 origin in the opposite orientation, f1(-), facilitates recovery of the other strand. The position of the polylinker in the alpha region of the *lacZ* gene allows for identification of inserts based on a blue/white color screen under the appropriate conditions. The T3 and T7 promoters are recognized by bacteriophage RNA polymerases. Transcription from these promoters reads into the polylinker from either side. RNA transcripts of any DNA cloned into the polylinker can thus be produced by run-off transcription in vitro.

pWE15 is an example of a cosmid vector used for cloning DNA fragments ~35 to 45 kb (see Fig. 1.5.3; Wahl et al., 1987). The *cos* sites allow the DNA to be cut and packaged into phage heads by the appropriate lambda proteins. There is a single unique *Bam*HI cloning site flanked by T3 and T7 promoter sequences. These promoters are particularly useful for production of labeled RNA probes corresponding to the ends of the insert DNA, and these can be used to identify overlapping cosmids for chromosome walking and construction of cosmid contigs. *Not*I sites flanking the cloning site can

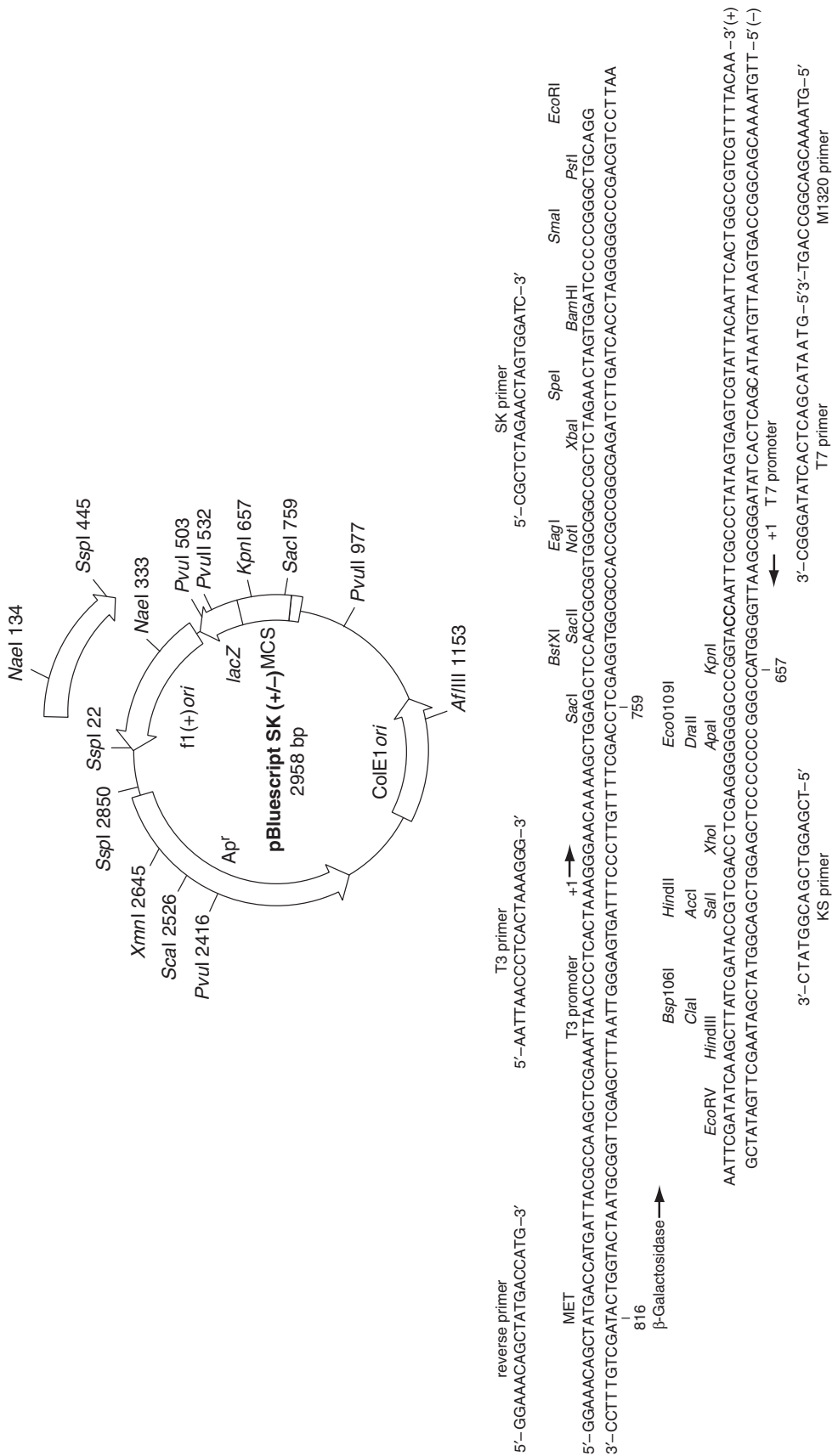


Figure 1.5.10 Map of pBluescript SK (+/-).

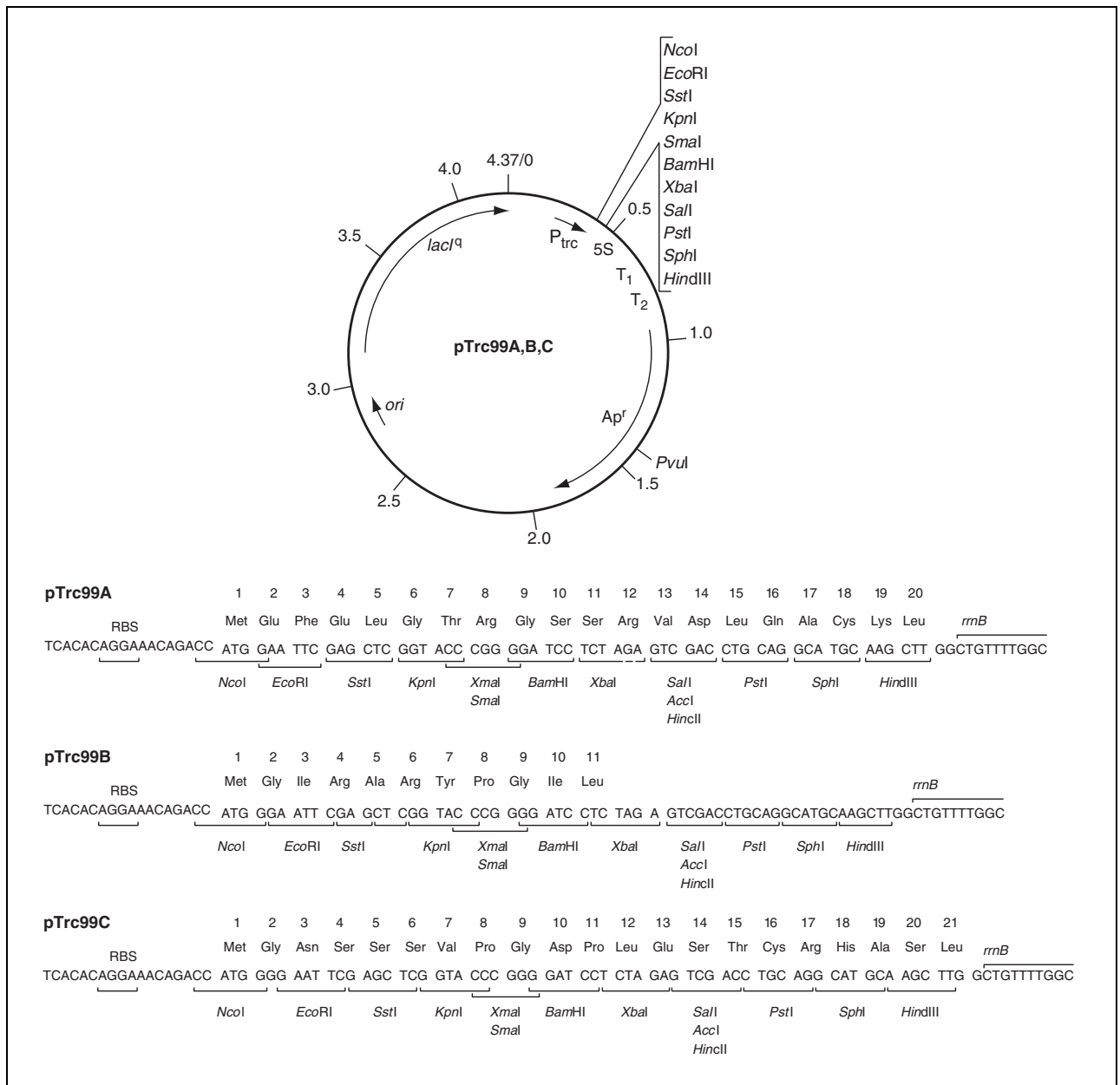


Figure 1.5.11 Map of pTrc99A,B,C.

potentially be used to excise an intact insert fragment from the vector. The ColE1-derived *ori* and ampicillin resistance gene allow for replication and selection in bacteria. The SV40 promoter (included in SV40 *ori*) which drives the neomycin phosphotransferase gene enables selection in eukaryotic cells.

pBeloBAC11 is an example of the family of bacterial-artificial-chromosome vectors based on the low-copy-number F factor replicator (see Fig. 1.5.4; Shizuya et al., 1992). BAC vectors are used for cloning large DNA fragments (100 to 500 kb) in *E. coli* and are used commonly in genome-mapping strategies. *oriS*, *repE*, *parA*, *parB*, and *parC* genes are the

essential genes that compose the F factor replicator. *oriS* and *repE* genes are required for unidirectional replication of the plasmid, and *parABC* loci stably maintain the copy number at one to two per *E. coli* genome. There are two unique cloning sites (*HindIII* and *BamHI*) inserted into the *lacZ* alpha region. Other useful features of the cloning region are (1) T7 and SP6 promoter sequences flanking the cloning sites, (2) *NotI* restriction sites flanking the cloning sites for potential excision of the insert, and (3) presence of the *loxP* and *cosN* sites that can be cleaved by specific enzymes. The ends generated by cleavage at *loxP* or *cosN* can be used as fixed reference points in building an

ordered restriction map by end labeling and partial restriction digestion.

pEGFP-1 is a selectable vector for monitoring promoter activity in mammalian cells via fluorescence of a green fluorescent protein (GFP) derivative (Clontech; see Fig. 1.5.5; Yang et al., 1996). The vector contains a neomycin resistance gene downstream of the SV40 early promoter for selection of stably transformed mammalian cells. It has a polylinker located upstream of the EGFP gene, so that the function of promoter sequences introduced into the polylinker can be assessed based on EGFP activity. The EGFP gene is modified from wild-type GFP to ensure expression in mammalian cells, it has silent base mutations that correspond to human codon-usage preferences, and sequences flanking the coding region have been converted to a Kozak consensus translation initiation signal. The vector backbone contains an *f1 ori* for production of ssDNA, a pUC-derived *ori* for propagation in *E. coli* and a kanamycin resistance gene for selection in bacteria.

A series of yeast shuttle vectors (pRS304, 305, and 306) has been created to facilitate manipulation of DNA in *Saccharomyces cerevisiae* (see Fig. 1.5.6; Sikorski and Hieter, 1989). These vectors have a backbone derived from pBLUESCRIPT into which the features necessary for replication and maintenance in yeast have been introduced. The members of this series of plasmids differ only in the yeast selectable marker incorporated; pRS303 carries the *HIS3* marker that complements a non-reverting *his3* chromosomal mutation in specific yeast strains. These plasmids contain an autonomously replicating sequence as well as a centromere sequence, *CEN6*, that ensures stable maintenance in yeast cells.

pCDNA3.1 is a selectable cloning and expression vector for use in mammalian cells. The features of this vector include a neomycin resistance gene driven by the SV40 early promoter (contained within the SV40 *ori*) and terminated by an SV40 polyadenylation signal for selection in mammalian cells (see Fig. 1.5.7). In addition, due to the inclusion of the SV40 *ori*, the vector can replicate as an episome in cells expressing the SV40 large T antigen. The polylinker cloning site is located downstream of strong cytomegalovirus enhancer-promoter sequences and upstream of the bovine growth hormone gene termination signals for high-level expression of protein-coding sequences cloned into this vector. This vector also contains some of the more standard features of

other plasmid vectors, including a ColE1 replicator for propagation in *E. coli*, the ampicillin resistance gene for selection in *E. coli*, the *f1 ori* for production of ssDNA, and the T7 promoter sequence for in vitro transcription of DNA inserted into the polylinker.

pRR54 is an example of a broad-host-range mobilizable plasmid vector. This vector contains replicator and stabilization sequences derived from the natural RK2 broad-host-range plasmid (see Fig. 1.5.8; Roberts et al., 1990). *oriV* is the vegetative origin of replication, *trfA* encodes *trans*-acting functions necessary for replication, and *par* encodes a locus that enhances stability of the plasmid. This plasmid can be mated into diverse gram-negative species as long as the appropriate mobilization machinery is provided in *trans* because it contains the origin of conjugal transfer, *oriT*. The plasmid carries the β -lactamase gene, allowing for ampicillin/carbenecillin selection of plasmid containing bacteria.

The pTrc series of plasmid expression vectors facilitates regulated expression of genes in *E. coli*. These vectors carry the strong hybrid *trp/lac* promoter, the *lacZ* ribosome-binding site (RBS), the MCS of pUC18 that allows insertion in three reading frames, and the *rrnB* transcription terminators (see polylinker sequences given below the vector diagram in Fig. 1.5.11). These vectors are equally useful for expression of unfused proteins (resulting from insertion in the *NcoI* site) or for expression of fusion proteins (using one of the cloning sites in the correct translational frame). The presence of the *lacI^q* allele on the plasmid ensures complete repression of the hybrid *trp/lac* promoter during cloning and growth in any host strain (see Amann et al., 1988, for further details).

LITERATURE CITED

- Amann, E., Ochs, B., and Abel, K.-J. 1988. Tightly regulated *tac* promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* 69:301-315.
- Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heynecker, H.L., and Boyer, H.W. 1977. Construction of useful cloning vectors. *Gene* 2:95-113.
- Chang, A.C.Y. and Cohen, S.N. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.* 134:1141-1156.
- Gerhart, E., Wagner, H., and Simmons, R.W. 1994. Antisense RNA control in bacteria, phages, and plasmids. *Annu. Rev. Microbiol.* 48:713-742.

- Jackson, D.A., Symons, R.M., and Berg, P. 1972. Biochemical method for inserting new genetic information into DNA of Simian Virus 40 circular DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 69:2904-2909.
- Kahn, M., Kolter, R., Thomas, C., Figurski, D., Meyer, R., Remaut, E., and Helinski, D.R. 1979. Plasmid cloning vehicles derived from plasmids ColE1, F, R6K, RK2. *Methods Enzymol.* 68:268-280.
- Kim, U.J., Shizuya, H., de Jong, P.J., Birren, B., and Simon, M.I. 1992. Stable propagation of cosmid-sized human DNA inserts in an F-factor based vector. *NAR* 20:1083-1085.
- Norrandner, J., Kempe, T., and Messing, J. 1983. Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. *Gene* 26:101-106.
- Roberts, R.C., Burioni, R., and Helinski, D.R. 1990. Genetic characterization of the stabilizing functions of a region of broad-host-range plasmid RK2. *J. Bacteriol.* 172:6204-6216.
- Shizuya, H., Birren, B., Kim U.J., Mancin, V., Slepak, T., Tachiiri, Y., and Simon, M.I. 1992. A bacterial system for cloning large human DNA fragments. *Proc. Natl. Acad. Sci. U.S.A.* 89:8794-8797.
- Sikorski, R.S. and Hieter, P.I. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* 122:19-27.
- Stoker, N.G., Fairweather, N., and Spratt, B.G. 1982. Versatile low-copy-number plasmid vectors for cloning in *Escherichia coli*. *Gene* 18:335-341.
- Sutcliffe, J.G. 1978. Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. *Cold Spring Harbor Symp. Quant. Biol.* 43:77-90.
- Uhlin, B.E., Schweickart, V., and Clark, A.J. 1983. New runaway-replication-plasmid cloning vectors and suppression of runaway replication by novobiocin. *Gene* 22:255-265.
- Wahl, G.M., Lewis, K.A., Ruiz, J.C., Rothenberg, B., Zhao, J., and Evans, G.A. 1987. Cosmid vectors for rapid genomic walking, restriction mapping, and gene transfer. *Proc. Natl. Acad. Sci. U.S.A.* 84:2160-2164.
- Yang, T., Cheng, L., and Kain, S.R. 1996. Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. *Nucl. Acids Res.* 24:4592-4593.

Contributed by Rhonda Feinbaum
Massachusetts General Hospital
Boston, Massachusetts

Minipreps of Plasmid DNA

Although there are a large number of protocols for the isolation of small quantities of plasmid DNA from bacterial cells (minipreps), this unit presents four procedures based on their speed and success: the alkaline lysis prep, a modification of the alkaline lysis prep that is performed in 1.5-ml tubes or 96-well microtiter dishes, the boiling method, and a lithium-based procedure. A support protocol provides information on storing plasmid DNA.

ALKALINE LYSIS MINIPREP

The alkaline lysis procedure (Birnboim and Doly, 1979, and Birnboim, 1983) is the most commonly used miniprep. Plasmid DNA is prepared from small amounts of many different cultures (1 to 24) of plasmid-containing bacteria. Bacteria are lysed by treatment with a solution containing sodium dodecyl sulfate (SDS) and NaOH (SDS denatures bacterial proteins, and NaOH denatures chromosomal and plasmid DNA). The mixture is neutralized with potassium acetate, causing the covalently closed plasmid DNA to reanneal rapidly. Most of the chromosomal DNA and bacterial proteins precipitate—as does the SDS, which forms a complex with potassium—and are removed by centrifugation. The reannealed plasmid DNA from the supernatant is then concentrated by ethanol precipitation.

Materials

LB medium (UNIT 1.1) containing appropriate antibiotic (Table 1.4.1)

Glucose/Tris/EDTA (GTE) solution

TE buffer (APPENDIX 2)

NaOH/SDS solution

Potassium acetate solution

95% and 70% ethanol

10 mg/ml DNase-free RNase (optional; UNIT 3.13)

1.5-ml disposable microcentrifuge tubes

1. Inoculate 5 ml sterile LB medium with a single bacterial colony. Grow to saturation (overnight).
2. Spin 1.5 ml of cells 20 sec in a microcentrifuge at maximum speed to pellet. Remove the supernatant with a Pasteur pipet.

The spins in steps 2 and 6 can be performed at 4°C or at room temperature. Longer spins make it difficult to resuspend cells.

3. Resuspend pellet in 100 µl GTE solution and let sit 5 min at room temperature.

Be sure cells are completely resuspended.

4. Add 200 µl NaOH/SDS solution, mix by tapping tube with finger, and place on ice for 5 min.

5. Add 150 µl potassium acetate solution and vortex at maximum speed for 2 sec to mix. Place on ice for 5 min.

Be sure mixing is complete.

6. Spin 3 min as in step 2 to pellet cell debris and chromosomal DNA.

7. Transfer supernatant to a fresh tube, mix it with 0.8 ml of 95% ethanol, and let sit 2 min at room temperature to precipitate nucleic acids.

BASIC PROTOCOL

**Escherichia coli,
Plasmids, and
Bacteriophages**

1.6.1

Contributed by JoAnne Engebrecht, Roger Brent, and Mustak A. Kaderbhai

Current Protocols in Molecular Biology (1991) 1.6.1-1.6.10

Copyright © 2000 by John Wiley & Sons, Inc.

8. Spin 1 min at room temperature to pellet plasmid DNA and RNA.
9. Remove supernatant, wash the pellet with 1 ml of 70% ethanol, and dry pellet under vacuum.
10. Resuspend the pellet in 30 μ l TE buffer and store as in support protocol. Use 2.5 to 5 μ l of the resuspended DNA for a restriction digest.

Contaminating RNA may interfere with detection of DNA fragments on the agarose gel; it can be destroyed by adding 1 μ l of a 10 mg/ml RNase solution (DNase-free) to the digestion mixture.

ALTERNATE PROTOCOL

ALKALINE LYSIS IN 96-WELL MICROTITER DISHES

Escherichia coli cells that contain plasmids are grown and lysed, and the plasmid DNA is precipitated—all in the wells of 96-well microtiter dishes. This procedure makes it possible to perform hundreds of rapid plasmid preps in a day. It is based on an unpublished procedure by Brian Seed of Massachusetts General Hospital.

Additional Materials

TYGPN medium (UNIT 1.1)

70% ethanol, ice-cold

Isopropanol

96-well microtiter plates (Dynatech PS plates or equivalent)

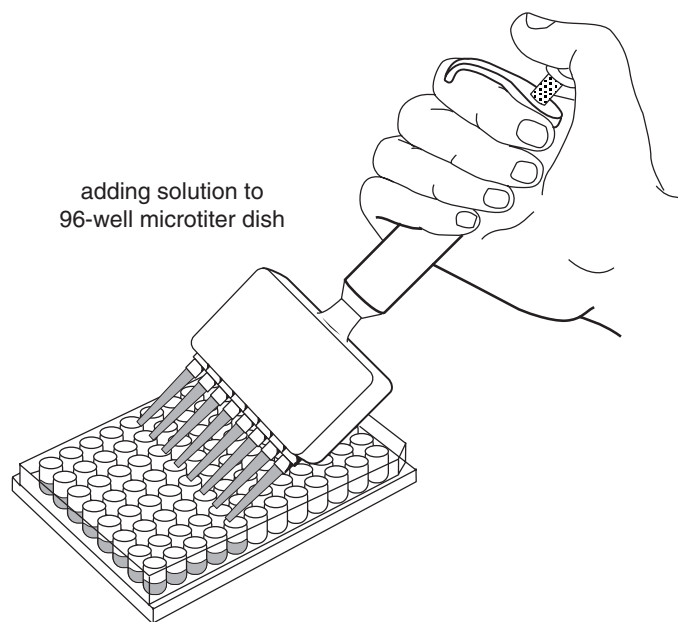
Multichannel pipetting device (8-prong Costar; 12-prong Titer Tek)

Multitube vortexer

Sorvall RT-6000 low-speed centrifuge, or equivalent, with microplate carrier in H-1000B rotor

1. Add 0.3 ml sterile TYGPN medium to each well of a 96-well microtiter plate (see sketch 1.6A). Inoculate each well with a single plasmid-containing colony.

The 96-well microtiter plates used must have U-shaped bottoms. To take full advantage of this protocol, one should perform all pipetting steps with a multichannel pipetting device.



Sketch 1.6A

2. Grow bacteria to saturation at 37°C (~48 hr).

All subsequent steps are performed at room temperature unless otherwise noted.

Potassium nitrate in the TYGPN medium presumably acts as a terminal electron acceptor when the bacteria in the wells are growing anaerobically, resulting in high cell densities.

3. Spin saturated cultures in H-1000B rotor with microplate carrier for 10 min at 2000 rpm (600 × g), 4°C. Decant supernatant with brief flick.

A microplate carrier is available for the Beckman JS-4.2 rotor. The same rpm values can be used for the JS-4.2 as those given here for the Sorvall H-1000B (see APPENDIX 1 for rotor conversion values).

4. Resuspend cells in the well bottoms by clamping the plate in a multitube vortexer and running it 20 sec at setting 4.
5. Add 50 µl GTE solution to each well.
6. Add 100 µl NaOH/SDS solution to each well. Wait 2 min.
7. Add 50 µl potassium acetate solution to each well.
8. Cover with plate tape or parafilm. Agitate vigorously in vortexer 20 sec at setting 4. Spin 5 min at 2000 rpm (600 × g), 4°C.
9. Insert a pipet tip just at the edge of the U in the bottom of the well. Remove 200 µl from each well and transfer to a new plate.

Do not try to recover all the fluid in each well.

10. Add 150 µl isopropanol to each well of the new plate. Cover with plate tape, agitate, and chill 30 min at -20°C.
11. Spin 25 min at 2000 rpm, 4°C, and decant supernatant. Wash pellets with cold 70% ethanol, gently decant supernatant, wash with 95% ethanol, and again gently decant supernatant.

Pellets often shrink visibly during the 70% ethanol wash, as impurities in them are dissolved.

Restriction enzymes will not cut well if the DNA is contaminated with even very small amounts of NaOH/SDS or potassium acetate solutions. It is therefore very important to decant the supernatants from the isopropanol precipitation and ethanol washes thoroughly, but not so vigorously that the pellets are flung out of the well bottoms.

If pellets become detached during either washing step, the plate should be respun at 2500 rpm for 5 min to bring the pellets to the bottom of the wells again.

12. Air dry pellets for 30 min, then resuspend in 50 µl TE buffer. Store as in support protocol and use 10-µl aliquots for digestion.

BOILING MINIPREP

Bacteria that contain plasmid DNA are broken open by treatment with lysozyme, Triton (a nonionic detergent), and heat. The chromosomal DNA remains attached to the bacterial membrane and is pelleted to the bottom of a centrifuge tube during a brief spin. Plasmid DNA is precipitated from the supernatant with isopropanol (Holmes and Quigley, 1981). This procedure is recommended for preparing small amounts of plasmid DNA from 1 to 24 cultures. It is extremely quick, but the quality of DNA produced is lower than that from the alkaline lysis miniprep.

Materials

LB medium (*UNIT 1.1*) containing appropriate antibiotic (Table 1.4.1)

STET solution

Hen egg white lysozyme

Isopropanol, ice-cold

TE buffer (*APPENDIX 2*)

10 mg/ml DNase-free RNase (optional; *UNIT 3.13*)

1.5-ml disposable microcentrifuge tubes

Boiling water bath (100°C)

1. Inoculate 5 ml sterile LB medium with a single bacterial colony. Grow at 37°C at least until mid-log phase ~6 hr, (a freshly saturated overnight culture works even better; see *UNIT 1.2*).
2. Transfer 1.5 ml of the saturated culture to a 1.5-ml microcentrifuge tube and pellet the cells by spinning 20 sec in microcentrifuge at maximum speed. Discard supernatant with a Pasteur pipet.

The spins in steps 2, 6, and 7 can be performed at 4°C or room temperature. Longer spins make it difficult to resuspend cells.

3. Resuspend the bacteria in 300 µl of STET solution containing 200 µg lysozyme. Vortex to achieve complete suspension.

Be sure cells are completely resuspended in order to maximize the number of cells exposed to the lysozyme and consequently the yield of plasmid DNA.

4. Place tube on ice for 30 sec to 10 min.

The time required for this step can vary between the limits indicated without affecting the yield or quality of the plasmid DNA.

5. Place tube in a boiling water bath (100°C) 1 to 2 min.

Heat and detergents cause the weakened cell walls to break, releasing plasmid DNA and RNA, but not the larger bacterial chromosome which remains attached to or trapped inside the lysed cells.

6. Spin in microcentrifuge 15 to 30 min at maximum speed.

The pellet, which should be fairly gummy, contains bacterial debris as well as chromosomal DNA. The supernatant contains plasmid DNA and RNA.

7. Pipet off supernatant into a new tube, carefully, without dislodging pellet. Mix with 200 µl (an equal volume) of cold isopropanol. Place at -20°C for 15 to 30 min. Spin 5 min in microcentrifuge at maximum speed.

The cold isopropanol precipitates the plasmid DNA and cellular RNA. Considerably shorter incubation periods (e.g., 2 to 5 min) may be sufficient for precipitation.

8. Remove the supernatant by inverting the tube and flicking it several times. Dry the

pellet by placing under a vacuum until it looks flaky.

If a vacuum source is unavailable, the pellet can be air dried.

9. Resuspend the pellet in 50 μ l TE buffer and store as in the support protocol. Use 5 μ l of the resuspended DNA for a restriction digest.

Contaminating RNA may interfere with detection of DNA fragments on the agarose gel; it can be destroyed by adding 1 μ l of a 10 mg/ml RNase solution (DNase-free) to the digestion mixture.

LITHIUM MINIPREP

Plasmid DNA is obtained from *E. coli* grown on plates as colonies or in liquid cultures. Bacterial cells harboring plasmid DNA are sequentially treated with Triton X-100/LiCl and phenol/chloroform. These steps solubilize plasmid DNA while precipitating chromosomal DNA with cellular debris. The debris is removed by centrifugation. This isolation procedure yields preparations of plasmid DNA that are virtually devoid of chromosomal DNA.

The procedure described here (originally presented by He et al., 1990) for small-scale isolation of plasmid DNA can also be readily extended for large-scale preparations as described in the annotation to the final step. The merit of the approach is that it is extremely reliable and rapid—requiring no more than 20 min of simple and economical operations for a preparation. The final plasmid DNA preparations are of a purity and quality usable for most biological applications.

Materials

TELT solution

LB medium (*UNIT 1.1*) containing appropriate antibiotic (Table 1.4.1)

1:1 (w/v) phenol/chloroform (*UNIT 2.1*)

100% ethanol, prechilled to -20°C

TE buffer (*APPENDIX 2*)

10 mg/ml DNase-free RNase A (optional; *UNIT 3.13*)

1.5-ml disposable microcentrifuge tubes

NOTE: All steps are performed at room temperature.

1. To isolate plasmid DNA from transformant colonies grown on agar plates, prepare the cells as follows:
 - a. Using a microspatula, scoop out an entire bacterial colony grown to 2- to 5-mm diameter on an LB agar plate. Transfer the colony to a 1.5-ml microcentrifuge tube containing 100 μ l TELT solution.
 - b. Vortex thoroughly to suspend the cells. Proceed to step 3.
2. To isolate plasmid DNA from liquid cultures, prepare the cells as follows:
 - a. Inoculate a colony of bacteria into 1.8 ml of sterile LB medium supplemented with appropriate antibiotic. Grow to saturation with shaking for 18 to 24 hr at 30°C (see *UNIT 1.2*).

Glass test tubes with plastic caps are suitable. Place the tubes at a suitably inclined angle to achieve good agitation.

- b. Carefully transfer the entire culture volume into a 1.5-ml microcentrifuge tube.

At this stage the volume of the liquid culture will have been reduced to ~ 1.5 ml.

BASIC PROTOCOL

**Escherichia coli,
Plasmids, and
Bacteriophages**

1.6.5

- c. Pellet the cells by spinning in a microcentrifuge ($10,000 \times g$) for 20 sec.
The spins in steps 2, 4, 7, and 9 can be performed at 4°C or room temperature. Centrifugation for longer periods or at higher speeds makes it difficult to resuspend the cells in the following step.
- d. With the tube held in a vertical position, aspirate the supernatant using a long-tipped Pasteur pipet connected to a vacuum line.
- e. Add 100 μl of TELT solution to the pellet and resuspend by vortexing.
Ensure that the cells are thoroughly suspended. See annotation to step 11 for scaled-up DNA preparations.
3. Add 100 μl of 1:1 phenol/chloroform and thoroughly vortex for 5 sec.
This mixture may be left at room temperature for ≤ 15 min. Plasmid yield will elevate with increasing duration of incubation; however, incubation periods > 15 min may result in phenol-mediated modification of DNA.
4. Microcentrifuge 1 min at $15,000 \times g$ (maximum speed).
5. Using a pipettor or similar device, carefully withdraw 75 μl of the upper aqueous phase and transfer the contents into a clean microcentrifuge tube.
Do not agitate the resolved phases. If mixing occurs, recentrifuge. When collecting the top layer avoid picking the debris at the interface.
6. To the supernatant, add 150 μl of chilled 100% ethanol. Mix the contents well to precipitate the plasmid DNA.
7. Pellet the nucleic acids by microcentrifuging 5 min at maximum speed.
8. Discard the supernatant by inverting the tube. When all the supernatant has drained, hold the tube in the same position for a few seconds and wipe off the last droplet from the rim of the tube by touching the edge of a Kleenex paper tissue.
9. Wash the pellet with 1 ml of cold 100% ethanol and harvest the nucleic acid pellet as in steps 6 and 7.
After centrifugation, decant the supernatant carefully as the pellet may be loose.
10. Cap the tube. Stab a small hole in the cap with a thumbtack or a syringe needle. Place the tube in a vacuum desiccator (without desiccant). Apply vacuum until the nucleic acid pellet appears completely dry.
A water-pumped vacuum line suffices for the purpose and usually takes ≤ 15 min.
11. Dissolve the pellet in 30 μl TE buffer. Vortex the contents well, capturing most of the DNA around the inner surface of the microcentrifuge tube. Store as in the support protocol and use 2 to 5 μl of DNA solution in a final 20- μl reaction volume for restriction digestion.

Contaminating RNA may interfere with detection of DNA fragments on the agarose gel; it can be destroyed by adding 1 μl of a 10 mg/ml RNase solution (DNase-free) to the digestion mixture.

For scaled-up plasmid DNA preparations (He et al., 1991), increase the amounts of TELT solution and 1:1 phenol/chloroform in direct proportion to the culture volume used. For cultures ≤ 5 ml, transfer the cells after suspension in TELT buffer into a microcentrifuge tube. Wash the final nucleic acids pellet twice with 1 ml of 100% ethanol. For cultures between 5 and 100 ml, use Corex glass tubes for treatment with TELT and phenol/chloroform and for centrifugations (Sorvall RC-5C centrifuge at $6000 \times g$).

STORAGE OF PLASMID DNA

Plasmids can be maintained for a short period (up to 1 month) in bacterial strains simply by growing on selective plates and storing at 4°C. For permanent storage, bacteria harboring the plasmid should be grown to saturation in the presence of the appropriate selective agent. An equal volume (~1 to 2 ml) of bacteria should be added to sterile 100% glycerol or a DMSO-based solution (recipe in *UNIT 1.3*) and frozen at -70°C in sterile vials. Cells taken from storage should again be grown on a selective plate (*UNIT 1.1*), and the plasmid DNA should be checked by restriction analysis (*UNIT 3.1*).

Plasmid DNA can be stored in TE buffer at 4°C for several weeks or preserved for several years by storing at -20° or -70°C. Most investigators prefer to store plasmids as frozen DNA, due to the widely held belief that plasmids stored in bacteria are sometimes lost, are rearranged, or accumulate insertion sequences and transposons during storage or on revival. Although such rearrangements certainly occur during storage of plasmids in bacteria in stab vials, we are unaware of any report of rearrangements affecting plasmids stored in frozen cells.

REAGENTS AND SOLUTIONS***Glucose/Tris/EDTA (GTE) solution***

50 mM glucose
25 mM Tris·Cl, pH 8.0
10 mM EDTA
Autoclave and store at 4°C

NaOH/SDS solution

0.2 N NaOH
1% (wt/vol) sodium dodecyl sulfate (SDS)
Prepare immediately before use

5 M potassium acetate solution, pH 4.8

29.5 ml glacial acetic acid
KOH pellets to pH 4.8 (several)
H₂O to 100 ml
Store at room temperature (do not autoclave)

STET solution

8% (wt/vol) sucrose
5% (wt/vol) Triton X-100
50 mM EDTA
50 mM Tris·Cl, pH 8.0
Filter sterilize and store at 4°C

TELT solution

2.5 M LiCl
50 mM Tris·Cl, pH 8.0
62.5 mM Na₂ EDTA
4% (wt/vol) Triton X-100
Store as 1- to 5-ml aliquots frozen at -20°C (do not filter sterilize or autoclave)

COMMENTARY

Background Information

Isolation of small quantities of plasmid DNA from bacterial cells is essential for the analysis of recombinant clones (UNITS 3.2 & 6.3). A myriad of plasmid DNA miniprep methods now exist and investigators are generally remarkably loyal to his or her own particular protocol. This unit presents three of the most widely used and reliable methods—alkaline lysis, the boiling method, and a lithium-based miniprep. The success of any of these procedures is largely a function of the expertise of the investigator; choice of method is therefore determined by personal preference as well as the size and type of the plasmid and the host strain of *E. coli*. With practice, all three protocols yield plasmid DNA of sufficient quantity and quality for use in most enzymatic manipulations (Chapters 3 & 7), and in most bacterial (UNIT 1.8) and yeast (UNIT 13.7) transformation procedures.

The procedures presented in this unit allow for the preferential recovery of circular plasmid DNA over linear chromosomal DNA. Treatments with either base or detergent (used in all three procedures) disrupt base pairing and cause the linear chromosomal DNA to denature and separate. In contrast, because of its supercoiled configuration, covalently closed circular plasmid DNA is unable to separate and readily reforms a correctly paired superhelical structure under renaturing conditions.

In the alkaline lysis miniprep, treatment with SDS and NaOH breaks open bacterial cells. Subsequent addition of potassium acetate preferentially reanneals covalently closed plasmid DNA, while chromosomal DNA and proteins are trapped in a complex formed between the potassium and SDS. The lysis treatment in the boiling miniprep causes chromosomal DNA to remain attached to the bacterial membrane, while plasmid DNA remains in the supernatant. In the lithium method, treatment of bacterial cells with Triton X-100/LiCl results in dissolution of the inner bacterial plasma membrane. However, this treatment has no effect on the overall morphology of the cells as observed microscopically, nor does it lead to release of plasmid DNA from the cells. Subsequent addition of phenol/chloroform leads to denaturation and precipitation of intracellular proteins. The concomitant rapid shrinkage of the cells preferentially expels soluble, super-

coiled plasmid DNA into the medium while retaining chromosomal DNA and denatured cellular protein with the bulk of the cell mass. Bacterial morphology, particularly of the cell wall envelope, is preserved under these conditions for at least 30 min, at which point cell lysis ensues. In all preparations, chromosomal DNA is removed with cellular debris by centrifugation, and soluble, supercoiled plasmid DNA is concentrated by ethanol precipitation.

In general, all three methods provide plasmid DNA of comparable yield and quality suitable for most biological applications. Yield of DNA is determined more by the type of plasmid than by method of isolation. Plasmids derived from pBR322 generally give lower (but, for most applications, sufficient) yields compared with the more recently derived pUC-like vectors, which contain a lesion in the plasmid-encoded *rop* gene, causing the plasmid to be maintained in high copy number in the cell (see UNIT 1.5).

All three methods are successful for the isolation of small (<10-kb) plasmids. Larger plasmids are generally poorly isolated using the lithium method, with yields <10% of those in the pUC-size range. These larger-sized plasmids may be retained along with chromosomal DNA in the mesh-like structure of the cell wall. In rare instances this problem may extend to plasmids that exist in low copy number and to those contained in strains with unusual cell wall compositions (see discussion of strain considerations below). The alkaline prep is better suited for efficient isolation of large-sized or low-copy-number plasmids (see critical parameters).

Strain background is also a consideration when selecting a miniprep procedure. When lysed with detergent or heat (as in the boiling or lithium procedures), strains such as HB101 and its derivatives may release a large amount of carbohydrate that contaminates plasmid DNA and inhibits many restriction endonucleases. In addition, HB101-related strains express endonuclease A which, if not inactivated, may degrade plasmid DNA in the presence of magnesium (during restriction enzyme digestion). Plasmids harbored in these strains may be better isolated using the alkaline lysis method.

In general, the alkaline lysis miniprep can be used for a variety of plasmid types and

sizes carried in most host strains. The boiling method is perhaps the most forgiving of the three, and the lithium miniprep, although restricted for use with small plasmids in certain hosts, is the most rapid method.

Critical Parameters and Troubleshooting

The successful isolation of small quantities of DNA from bacterial cells, regardless of the protocol, is largely dependent on the strain of *E. coli* used. For example, DNA isolated from strains C600, DH1 and LE392 is of good quality, while DNA isolated from strains HB101 and the JM100 series is of lesser quality (see Table 1.4.5). The latter strains have high nuclease activity (endonuclease A, which is not completely inactivated by boiling), necessitating further purification with a phenol extraction (UNIT 2.1) or an additional precipitation with ammonium acetate (UNIT 2.1). If one protocol fails to yield DNA with a particular strain of *E. coli*, it would be best to try an alternative method. The alkaline lysis miniprep seems to be the most consistent regardless of strain used, while the lithium method incorporates a phenol extraction and may be well suited for isolation of plasmid from strains with high nuclease activity.

Failure to isolate DNA using any of the protocols could be due to DNase contamination of the RNase. This problem can be circumvented by leaving out or using considerably less RNase. Other procedure-specific guidelines are described below.

Alkaline lysis miniprep. The DNA yield can generally be increased by adding 500 mg/ml of lysozyme to the 100- μ l suspension of cells in the glucose/Tris/EDTA (GTE) solution (step 3).

With a few variations, this protocol can also be used to isolate low-copy-number plasmids (UNIT 1.5), as follows: Increase the number of starting cells to 3 ml and include lysozyme in the GTE solution as described above. Because yields are lower with strains harboring low-copy-number plasmids, it will be necessary to use at least 5 μ l of DNA for a restriction digest.

If the isolated DNA fails to cut with restriction endonucleases, the most common cause is inadequate washing of the pellets after the ethanol precipitation step. Precipitating the DNA a second time with ethanol, or washing the pellets from the first precipitation with 70% ethanol, will usually clean up the DNA enough for restriction enzyme cutting.

Boiling miniprep. Failure to isolate DNA is

sometimes caused by incomplete cell lysis. If this appears to be the case, try using a new bottle of lysozyme powder. Apparent failure to recover DNA is also sometimes caused by DNase contamination of the freshly prepared plasmid DNA, or of the RNase. If miniprep DNA is contaminated with DNase, it can usually be cleaned up by an additional precipitation with ammonium acetate (UNIT 2.1) or by phenol extraction (UNIT 2.1). This method should not be used with endonuclease A-containing strains (such as HB101).

Lithium miniprep. This procedure eliminates many of the lengthy fractionation steps commonly employed in other minipreps of plasmid DNA and is quite effective for isolating plasmids ≤ 10 kbp. The high quality of the plasmid DNA obtained by this method has been verified by the fact that no problems are encountered with regard to inhibition of some sensitive restriction enzyme activities—for instance *NdeI*, which is sensitive to trace quantities of impurities (see Fig. 1.6.1). Any DNA cleavage inhibition is likely to be due to contamination of the DNA with cellular debris from the phenol/chloroform interface. However, this can be readily overcome by reducing the volume of DNA in the reaction mixture. The plasmid DNA is suitable for direct transformation of bacteria. If speed is required, DNA recovered from the aqueous phase following phenol/chloroform extraction (after step 4) can be used for transformation. However, such DNA should be used immediately to minimize the likelihood of phenol-mediated modification of DNA.

This technique has been used to purify shuttle vectors from yeast for high-frequency back-transformation in *E. coli* (Ward, 1990).

Anticipated Results

Two to five micrograms of DNA are obtained from 1.5 ml of a culture of cells containing a pBR322-derived plasmid by using either the alkaline lysis miniprep, the boiling protocol, or the lithium method. Three- to five-fold higher yields can be expected from pUC-derived plasmids. DNA yield from alkaline lysis in 96-well microtiter dishes (0.3 ml of culture/well) is ~ 2 μ g for high-copy-number plasmids (e.g., pUC plasmids). All three methods provide plasmid DNA of comparable quality, suitable for use in many applications.

Time Considerations

Using the alkaline lysis procedure, it is possible, with practice, to produce twelve samples

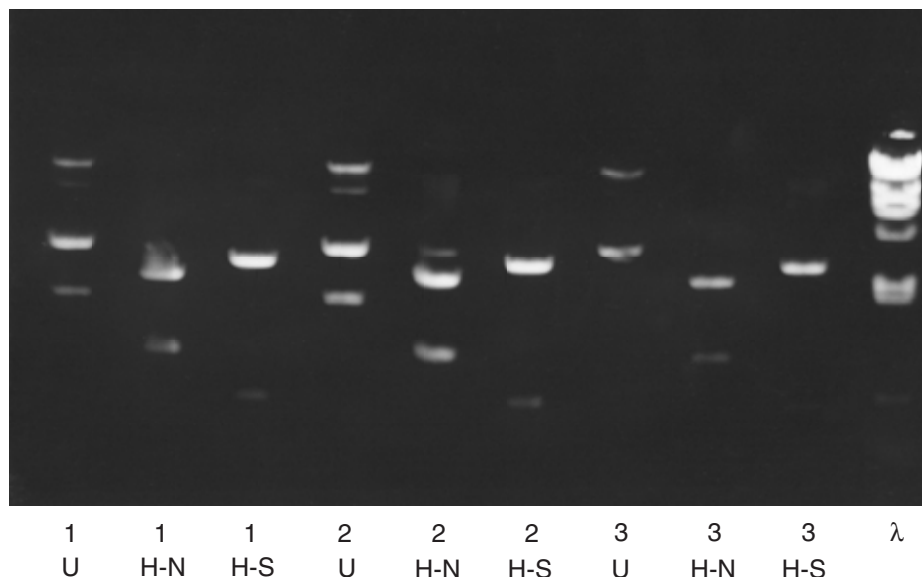


Figure 1.6.1 Restriction mapping of three preparations of plasmid DNA (pVS-WH) from transformant of *E. coli* TB strain. The three clones are numbered 1, 2, and 3 and were treated as follows: **U**, uncleaved DNA; **H-N**, plasmid DNA doubly cut with *Hind*III and *Nde*I; **H-S**, plasmid DNA doubly cut with *Hind*III and *Sph*I; λ , *Hind*III-digested λ DNA. A volume of 2.5 μ l of plasmid DNA (about 1 μ g) was used for digestion. The samples were separated on a 0.7% agarose gel containing 80 mM Tris-phosphate (pH 8.0) and 1 mM EDTA using Pharmacia Submarine electrophoresis tank with a current of 70 mA for 1 hr. All of the three clones are positive.

of DNA from saturated bacterial cultures in <1 hr. Although it is possible to do a large number of samples in a single day, we recommend using the alkaline lysis in 96-well microtiter dishes for such mass screenings. With the latter protocol, it is easy to process two plates in 4 hr. However, an incubation period of 48 hr is often required for cells to grow to saturation in the wells of the dish.

Starting with saturated cultures, the boiling method allows for the isolation of twelve DNA samples in 1 hr.

The lithium method is extremely quick—DNA samples can be processed in 20 min. With proper organization and availability of the materials, twelve plasmid preparations can be accomplished in substantially less than 1 hr. The entire operation involving plasmid preparation, restriction digestion, and agarose gel electrophoresis can be easily completed within half a day.

Literature Cited

- Birboim, H.C. 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. *Meth. Enzymol.* 100:243-255.
- Birboim, H.C. and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* 7:1513-1523.

He, M., Kaderbhai, M.A. Adcock, I., and Austen, B.M. 1991. An improved and rapid procedure for isolating RNA-free *Escherichia coli* plasmid DNA. *Gene Anal. Tech.* 8:107-110.

He, M., Wilde, A., and Kaderbhai. 1990. A simple single-step procedure for small-scale preparation of *Escherichia coli* plasmids. *Nucl. Acids Res.* 18:1660.

Holmes, D.S. and Quigley, M. 1981. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* 114:193-197.

Ward, A.C. 1990. Single-step purification of shuttle vectors from yeast for high frequency back-transformation into *E. coli*. *Nucl. Acids Res.* 18:5319.

Contributed by JoAnne Engebrecht
State University of New York
Stony Brook, New York

Roger Brent
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

Mustak A. Kaderbhai (lithium miniprep)
University College of Wales
Penglais, Aberystwyth, United Kingdom

Large-Scale Preparation of Plasmid DNA

UNIT 1.7

Although the need for large quantities of plasmid DNA has diminished as techniques for manipulating small quantities of DNA have improved, occasionally large amounts of high-quality plasmid DNA are desired. This unit describes the preparation of milligram quantities of highly purified plasmid DNA. The first part of the unit describes three methods for preparing crude lysates enriched in plasmid DNA from bacterial cells grown in liquid culture. The second part describes three methods for purifying plasmid DNA in such lysates away from contaminating RNA and protein.

Methods for crude lysate preparation—alkaline lysis (see Basic Protocol 1), boiling (see Alternate Protocol 1), and Triton lysis (see Alternate Protocol 2)—separate chromosomal and plasmid DNA by exploiting the structural differences between these molecules. Plasmids are covalently closed and smaller than chromosomal DNA. When the cell lysate is centrifuged to pellet chromosomal DNA and cellular debris, these differences permit plasmid DNA to remain in the supernatant (see Key References for the theory behind each protocol). All three protocols yield a solution greatly enriched for plasmid DNA but that still contains significant amounts of RNA and protein. These contaminants must be removed if certain procedures, including 5' end labeling with T4 polynucleotide kinase and transfection of higher eukaryotic cells, are to be performed with the DNA. Accordingly, three procedures are described for purifying plasmid DNA from crude lysates. CsCl/ethidium bromide density gradient centrifugation (see Basic Protocol 2) permits separation by the different capacities of covalently closed plasmid DNA and chromosomal DNA to bind the intercalating agent ethidium bromide. Because binding of ethidium bromide lowers the density of DNA and plasmid DNA can bind less ethidium bromide than chromosomal DNA, plasmid DNA forms bands in a region of greater density (lower in the tube) than chromosomal DNA (see Fig. 1.7.1). Polyethylene glycol (PEG) precipitation (see Alternate Protocol 3) takes advantage of the inverse relationship between macromolecular size and concentration of PEG required for precipitation. There are two chromatographic methods for purifying plasmid DNA (see Alternate Protocol 4): anion-exchange chromatography, which exploits the strong negative charge of nucleic acids, and size-exclusion chromatography, which takes advantage of the large size of plasmid DNA molecules relative to other molecules present in the crude lysate.

PREPARATION OF CRUDE LYSATE BY ALKALINE LYSIS

Alkaline lysis is probably the most generally useful plasmid preparation procedure. It is fairly rapid, very reliable, and yields reasonably clean crude DNA that can be further purified by any of the purification methods described in this unit. Plasmid-bearing *E. coli* cells are lysed with lysozyme. The lysate is treated with NaOH/SDS solution and potassium acetate and centrifuged to separate plasmid DNA from proteins and chromosomal DNA. The supernatant is treated with isopropanol to precipitate plasmid DNA.

Materials

- LB medium or enriched medium (e.g., superbroth or terrific broth; UNIT 1.1)
containing ampicillin or other appropriate selective agent (Table 1.4.1)
- Plasmid-bearing *E. coli* strain
- Glucose/Tris/EDTA solution (UNIT 1.6)
- 25 mg/ml hen egg white lysozyme in glucose/Tris/EDTA solution (prepare fresh)
- 0.2 M NaOH/1% (w/v) SDS [prepare fresh from 10 M NaOH and 10% (w/v)
SDS stocks]
- 3 M potassium acetate solution, pH ~5.5 (see recipe)

**BASIC
PROTOCOL 1**

***Escherichia coli*,
Plasmids, and
Bacteriophages**

Isopropanol
70% ethanol

Sorvall GSA, GS-3, or Beckman JA-10 rotor (or equivalent)
High-speed centrifuge tubes with ≥ 20 -ml capacity (e.g., Oak Ridge centrifuge tubes)
Sorvall SS-34 or Beckman JA-17 rotor (or equivalent)

Grow and concentrate cells

1. Inoculate 5 ml LB medium or enriched medium containing selective agent (most commonly ampicillin) with a single colony of *E. coli* containing the desired plasmid. Grow at 37°C with vigorous shaking overnight.
2. Inoculate 500 ml LB medium or enriched medium containing selective agent in a 2-liter flask with ~1 ml of overnight culture. Grow at 37°C until culture is saturated ($OD_{600} \cong 4$).

To increase yields, maximize aeration using a flask with high surface area (whose volume exceeds the culture volume—i.e., is >2 liters), add baffles, and shake at >400 rpm. Alternatively, treat cultures of cells growing logarithmically with chloramphenicol to amplify the plasmids (see UNIT 1.5). Growing the bacteria in medium that supports higher cell densities also increases the yield. These media include M9, terrific broth, and LB medium containing 0.1% (w/v) glucose (UNIT 1.1). These media can increase plasmid yields 2- to 10-fold; different plasmids respond to the media differently. Most plasmids commonly used today, particularly derivatives of the pUC series (Fig. 1.5.2), grow at a copy number high enough to routinely yield 1 to 5 mg plasmid DNA from a 500-ml culture grown in LB medium.

An important consideration when using enriched medium is the method to be used for final purification of plasmid DNA. Increased yield poses no problems when using CsCl/ethidium bromide or PEG purification. In addition, commercially available chromatography columns have been designed for large-scale plasmid purification. For example, the larger Qiagen-tip columns can purify 500 μ g to 10 mg of plasmid DNA.

3. Collect cells by centrifuging 10 min at $6000 \times g$ (~6000 rpm in Sorvall GSA/GS-3 or Beckman JA-10 rotors), 4°C.

If necessary the pellets can be stored frozen indefinitely at -20° or -70° C.

4. Resuspend pellet from 500-ml culture in 4 ml glucose/Tris/EDTA solution and transfer to high-speed centrifuge tube with ≥ 20 -ml capacity.

Lyse the cells

5. Add 1 ml of 25 mg/ml hen egg white lysozyme in glucose/Tris/EDTA solution. Resuspend the pellet completely in this solution and allow it to stand 10 min at room temperature.

Neither glucose nor lysozyme is absolutely necessary for the success of the procedure. Glucose serves as a buffer in step 6 when the pH of the solution is greatly increased by addition of NaOH. Glucose provides buffering in the range of pH 12 and, by preventing the pH from rising too drastically in step 6, increases the efficiency of precipitation in step 7 (when the pH is lowered by addition of potassium acetate).

Lysozyme assists in the destruction of bacterial cell walls and subsequent release of plasmid DNA. Bacterial debris and soluble proteins are precipitated in step 7. One problem that can reduce recovery of plasmid DNA is inefficient separation of plasmid DNA from cellular debris. Lysozyme helps increase yield by reducing the amount of plasmid DNA trapped in partially degraded cell material and subsequently lost by precipitation at step 7. Omitting lysozyme reduces plasmid recovery by 5% to 10%.

The effort and expense required to include glucose and lysozyme in step 5 is negligible. The efficiency gained in streamlining the procedure by omitting them is also negligible. However, the potential for loss of plasmid DNA when these components are not included is measurable and worth avoiding. It should be noted that some commercially available chromatographic systems (e.g., Qiagen) do not use lysozyme for lysate preparation. The Qiagen modified alkaline lysis procedure relies on inefficient bacterial lysis (facilitated by sodium hydroxide lysis and acidic potassium acetate neutralization) to reduce contamination of plasmid DNA with chromosomal DNA. Although omitting lysozyme reduces the recovery of plasmid DNA, the Qiagen protocol yields high-quality DNA. When using Qiagen-tips for large-scale plasmid preparation, the manufacturer's recommendations should be followed.

When chromatographic methods are used for final purification of plasmid DNA, it is essential to degrade RNA that contaminates the lysate and will copurify with plasmid DNA. Treating the lysate with RNase A is the most efficient and economical method for degrading RNA. This can be accomplished at any step in the preparation of crude lysate, but it is most convenient to do it at step 5, by adding RNase A to the glucose/Tris/EDTA solution to a final concentration of 50 µg/ml.

6. Add 10 ml freshly prepared 0.2 M NaOH/1% SDS and mix by stirring gently with a pipet until solution becomes homogeneous and clears. Let stand 10 min on ice.

The solution should become very viscous.

7. Add 7.5 ml of 3 M potassium acetate solution and again stir gently with a pipet until viscosity is reduced and a large precipitate forms. Let stand 10 min on ice.
8. Centrifuge 10 min at $20,000 \times g$ (13,000 rpm in Sorvall SS-34; 12,500 rpm in Beckman JA-17), 4°C.

A large, fairly compact pellet will form; this contains most of the chromosomal DNA, SDS-protein complexes, and other cellular debris. Plasmid DNA remains in the translucent supernatant.

Addition of ~0.5 ml chloroform before the centrifugation can help reduce floating material.

Precipitate plasmid DNA

9. Decant the supernatant into a clean centrifuge tube. Pour it through several layers of cheesecloth if any floating material is visible. Add 0.6 vol isopropanol, mix by inversion, and let stand 5 to 10 min at room temperature.

If the supernatant is cloudy or contains floating material, repeat centrifugation (step 8) before adding isopropanol.

10. Recover nucleic acids by centrifuging 10 min at $15,000 \times g$ (11,500 rpm in SS-34 rotor; 10,500 rpm in JA-17 rotor), room temperature.
11. Wash the pellet with 2 ml of 70% ethanol; centrifuge briefly at $15,000 \times g$, room temperature, to collect pellet. Aspirate ethanol and dry pellet under vacuum.

The pellet can be stored indefinitely at 4°C.

PREPARATION OF CRUDE LYSATE BY THE BOILING METHOD

The boiling method is extremely simple and fast but typically yields crude DNA containing more contaminating bacterial DNA and proteins than other methods. In this protocol a bacterial cell lysate is boiled to denature chromosomal DNA and protein. These denatured macromolecules are precipitated by centrifugation, whereas plasmid DNA remains in the supernatant. The plasmid DNA is then precipitated with isopropanol.

Additional Materials (also see *Basic Protocol 1*)

STET solution (*UNIT 1.6*)

10 mg/ml hen egg white lysozyme in 25 mM Tris·Cl, pH 8.0 (prepare fresh)

Boiling and ice-water baths

Sorvall HB-4 rotor (or equivalent) and appropriate centrifuge tube

1. Grow and concentrate cells (see *Basic Protocol 1*, steps 1 to 3).
2. Resuspend pellet from a 500-ml culture in 20 ml STET solution and transfer to a glass tube or flask.
3. Add 2 ml hen egg white lysozyme and mix solution by inverting several times. Heat to near boiling over an open flame, then incubate 1 min in a boiling water bath.

Be certain that enough room remains between the top of the solution and the top of the tube to permit safe handling and to prevent boiling over. Submerge the tube deep enough into the boiling water to allow the entire solution to heat rapidly.

4. Place tube into an ice-water bath to cool.
5. Pour solution into centrifuge tube and centrifuge 20 min at $\geq 25,000 \times g$, preferably in a swinging-bucket rotor (e.g., 12,000 rpm in HB-4), room temperature.

After boiling, the solution will be extremely viscous due to denatured chromosomal DNA. It will tend to behave as a goeey, semisolid mass. Therefore, be careful not to allow the solution to overflow when pouring it into centrifuge tubes.

Centrifugation in a swinging-bucket rotor permits concentration of chromosomal DNA and denatured proteins at the bottom of the tube in a more compact pellet than is possible in a fixed-angle rotor. However, fixed-angle rotors can be used—e.g., Sorvall SS-34 at $47,000 \times g$ (20,000 rpm), Beckman JA-17 at $40,000 \times g$ (17,000 rpm), Beckman 70Ti at $200,000 \times g$ (44,000 rpm), or SW-41 at $100,000 \times g$ (25,000 rpm).

6. Decant supernatant to a clean centrifuge tube.

*The supernatant can be used without further treatment for purifying plasmid DNA by CsCl/ethidium bromide equilibrium gradient centrifugation (see *Basic Protocol 2*). If the volume is greater than necessary and inconvenient to handle, plasmid DNA can be precipitated with isopropanol. Precipitation is required for purification by PEG precipitation (see *Alternate Protocol 3*) or column chromatography (see *Alternate Protocol 4*).*

7. Add 0.6 vol isopropanol, mix by inversion, and let stand 5 to 10 min at room temperature.
8. Pellet nucleic acids by centrifuging 10 min at $15,000 \times g$ (11,500 rpm in SS-34 rotor; 10,500 rpm in JA-17 rotor), room temperature.
9. Wash the pellet with 2 ml of 70% ethanol. Centrifuge briefly to collect pellet. Aspirate ethanol and dry pellet under vacuum.

The pellet can be stored indefinitely at 4°C.

PREPARATION OF CRUDE LYSATE BY TRITON LYSIS

ALTERNATE PROTOCOL 2

The method described below is a modification of that described by Clewell and Helinski (1970, 1972) in which Brij-58 and sodium deoxycholate were used. In this protocol plasmid DNA is extracted from a bacterial cell lysate that has been treated with Triton X-100 (TX-100). This is a very gentle procedure and is therefore useful for isolating very large plasmids such as cosmids. Plasmid DNA can be further purified by any of the methods described in the second part of this unit.

Additional Materials (also see Basic Protocol 1)

Sucrose/Tris/EDTA solution (see recipe)

10 mg/ml hen egg white lysozyme in 25 mM Tris·Cl, pH 8.0 (prepare fresh)

0.5 M EDTA (APPENDIX 2)

10 mg/ml DNase-free RNase (UNIT 3.13)

Triton lysis solution (see recipe)

1:1 (v/v) buffered phenol/chloroform (UNIT 2.1)

24:1 (v/v) chloroform/isoamyl alcohol (UNIT 2.1)

Additional reagents and equipment for phenol/chloroform extraction (UNIT 2.1)

1. Grow and concentrate cells (see Basic Protocol 1, steps 1 to 3).
2. Resuspend pellet from a 500-ml culture in 5 ml sucrose/Tris/EDTA solution. Transfer to appropriate centrifuge tube (tube should be less than $\sim\frac{1}{3}$ full).
3. Add to tube:
 - 1.5 ml 10 mg/ml hen egg white lysozyme in 25 mM Tris·Cl
 - 2 ml 0.5 M EDTA
 - 25 μ l 10 mg/ml DNase-free RNase.

Let stand 15 min on ice.

4. Overlay with 2.5 ml Triton lysis solution. Mix gently but thoroughly by inversion and let stand 20 min at 4°C.

This solution should not be vortexed or shaken hard because that will shear chromosomal DNA and prevent it from precipitating in the next step. The solution will become extremely viscous as the cells lyse. Streaks of opaque material will be visible and may remain throughout the incubation.

5. Centrifuge 70 min at 40,000 \times g (18,000 rpm in Sorvall SS-34; 17,000 rpm in Beckman JA-17), 4°C.
6. Decant supernatant carefully to a clean centrifuge tube.

Avoid contaminating plasmid DNA with the gelatinous pellet. The pellet contains chromosomal DNA and cellular debris. The pellet may detach from the bottom of the tube, so it is helpful to hold a Pasteur pipet against the mouth of the tube while pouring to prevent the pellet from sliding into the clean tube. The integrity of the pellet varies greatly between preparations; occasionally it is necessary to leave behind some of the viscous supernatant to avoid contaminating it with pellet material.

This supernatant can be used without further treatment for purifying plasmid DNA by CsCl/ethidium bromide equilibrium gradient centrifugation (see Basic Protocol 2). If the volume is greater than necessary and inconvenient to handle, plasmid DNA can be extracted and precipitated with isopropanol. Extraction and precipitation are required for purification by PEG precipitation (see Alternate Protocol 3) or column chromatography (see Alternate Protocol 4).

***Escherichia coli,
Plasmids, and
Bacteriophages***

1.7.5

7. Extract the supernatant with 1:1 buffered phenol/chloroform, then with 24:1 chloroform/isoamyl alcohol. Add 0.6 vol isopropanol to the final aqueous phase and let stand 5 to 10 min at room temperature.

Phenol extraction and isopropanol precipitation are done as described in UNIT 2.1.

8. Pellet nucleic acids by centrifuging 10 min at $15,000 \times g$ (11,500 rpm in SS-34 rotor; 10,500 rpm in JA-17 rotor), room temperature.
9. Wash the pellet with 2 ml of 70% ethanol and centrifuge briefly to collect pellet. Aspirate ethanol and dry pellet under vacuum.

The pellet can be stored indefinitely at 4°C.

BASIC PROTOCOL 2

PURIFICATION OF PLASMID DNA BY CsCl/ETHIDIUM BROMIDE EQUILIBRIUM CENTRIFUGATION

This purification procedure yields high-quality plasmid DNA free of most contaminants, but it requires the use of ethidium bromide (a mutagen) and often requires long ultracentrifuge runs to establish the density gradient. A crude bacterial cell lysate is mixed with cesium chloride (CsCl) and ethidium bromide and centrifuged to equilibrium. Ethidium bromide is removed by passing plasmid DNA over a cation exchange column, and CsCl is removed by ethanol precipitation.

Materials

Pellet from crude lysate of plasmid-bearing bacterial cell culture (see Basic Protocol 1 or Alternate Protocol 1 or 2)

TE buffer, pH 7.5 (APPENDIX 2)

Cesium chloride

10 mg/ml ethidium bromide (APPENDIX 2)

CsCl/TE solution (see recipe)

Dowex AG50W-X8 cation-exchange resin (see recipe)

TE buffer (pH 7.5)/0.2 M NaCl

100% and 70% ethanol

Beckman VTi65 or VTi80 rotor (or equivalent)

5-ml quick-seal ultracentrifuge tubes

3-ml syringes with 20-G needles

Additional reagents and equipment for ethanol precipitation (UNIT 2.1)

CAUTION: Ethidium bromide is a mutagen and environmental hazard. It should be handled carefully with gloves and disposed of properly. Methods for disposal vary between different institutions. Consult the institution's environmental safety office for the preferred means of storage and disposal of ethidium bromide-containing waste.

1. Resuspend pellet from a crude lysate preparation in 4 ml TE buffer. Add 4.4 g CsCl, dissolve, and add 0.4 ml of 10 mg/ml ethidium bromide. If using supernatants resulting from boiling or Triton lysis preparations, add 1.1 g CsCl/ml supernatant and 0.1 ml of 10 mg/ml ethidium bromide/ml supernatant.

Ethidium bromide will form a complex with protein remaining in the solution to form a deep red flocculent precipitate. This can be removed by centrifuging the lysate-CsCl/ethidium bromide solution 5 min at $\sim 2000 \times g$, room temperature. After this procedure, the protein-ethidium bromide complex will form a disc at the top of the solution. The solution can be pipetted out from beneath the disc or poured carefully, allowing the floating disc to adhere to the side of the tube.

2. Transfer the solution to a 5-ml ultracentrifuge tube. Top up the tube, if necessary, with CsCl/TE solution and seal tube. Band plasmid by centrifuging 3.5 hr at $500,000 \times g$ (77,000 rpm in VTi80 rotor) or ≥ 14 hr at $350,000 \times g$ (65,000 rpm in VTi80 rotor; 58,000 rpm in VTi65 rotor), 20°C .

This centrifugation must be done at a temperature no lower than 15°C . Because of the high concentration of CsCl and the high centrifugal force necessary to establish the gradient, lower temperatures will cause the CsCl at the bottom of the tube (where the density is highest) to precipitate during the run. CsCl precipitate moves the center of mass towards the bottom of the tube. This can unbalance the rotor and cause breakage of the rotor and destruction of the centrifuge at least, and serious personal injury at worst. Also, equilibrium is achieved more quickly by warmer gradients.

Other rotors can be used, including fixed-angle and swinging-bucket rotors (e.g., centrifuge ≥ 24 hr at 56,000 rpm in Ti70 rotor). These rotors require longer centrifugation times than vertical rotors, but allow larger volumes or more samples to be included in the run.

3. Carefully remove the tube from the centrifuge. Insert a 20-G needle gently into the top of the tube. Recover the plasmid band (the lower of the two bands) by inserting a 3-ml syringe with a 20-G needle attached into the side of the tube ~ 1 cm below the plasmid band as shown in Figure 1.7.1. Insert the needle with the beveled side up.

Do not allow the gradient to be mixed by rough handling or turbulence. Be certain not to cover the top of the first needle with gloved finger. The needle serves to provide an inlet for air to displace the volume of solution being withdrawn.

If chromosomal DNA has been thoroughly removed in previous steps, only the plasmid band may be visible. Large amounts of plasmid DNA will be visible in the gradient in ordinary light. Smaller amounts can be visualized more easily by side illumination with low-intensity shortwave UV light. Prolonged exposure of the DNA–ethidium bromide complex to UV light will cause damage to the DNA and should be avoided. RNA may be detected as a diffuse region of fluorescence at the bottom of the tube.

CAUTION: *To avoid potentially serious eye injury by UV light, wear UV-blocking glasses or face shield. Wear gloves when handling ethidium bromide.*

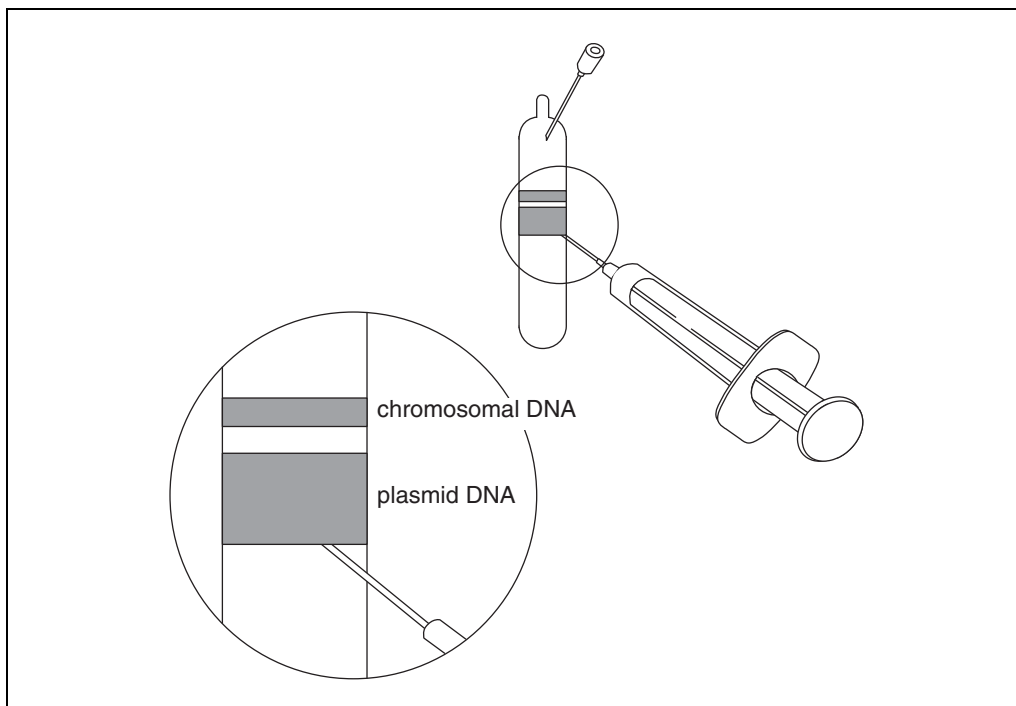


Figure 1.7.1 Collecting plasmid DNA from a CsCl gradient.

Protein–ethidium bromide complexes will pellet on the outside edge of the tube if not removed earlier. To prevent contamination of plasmid DNA, avoid this area when inserting the collection needle.

There should be no resistance in the syringe when drawing off the plasmid DNA. If there is resistance, check that needles are clear. Occasionally the needle will become clogged if a piece of tube enters it. Do not try to draw harder on the syringe, as this may create turbulence in the tube when the obstruction is sucked in and cause mixing of the gradient. Instead, insert another needle and use it to draw off the band. Leave the clogged needle in place in the tube. (If the clogged needle is removed, the tube will empty through the hole that remains.) The air inlet needle can also become clogged; if it does, remove it and allow air to enter through the remaining hole. If plasmid DNA is drawn through a very small opening in a clogged needle, it may be sheared.

4. If higher-purity plasmid DNA is required, perform a second ultracentrifugation to eliminate any contaminating RNA or chromosomal DNA. Add plasmid DNA band to another ultracentrifuge tube, top up with CsCl/TE solution containing 1.0 mg/ml ethidium bromide, and repeat steps 2 and 3.
5. Pour a Dowex AG50W-X8 column, 1.5 to 2 times the volume of the plasmid DNA/ethidium bromide solution, in a glass or plastic column. Pass several volumes of TE buffer/0.2 M NaCl through the column to wash and equilibrate it.

The column can be set up in a Pasteur pipet plugged with a little glass wool or in a commercially purchased plastic column.

6. Load the plasmid DNA/ethidium bromide solution directly from syringe to top of resin bed without disturbing the resin.
7. Begin collecting the solution flowing through the column immediately after loading the plasmid solution. Wash the column with a volume of TE buffer/0.2 M NaCl equal to twice that of the volume of plasmid solution loaded.

The final volume collected from the column should be three times the volume of plasmid DNA solution removed from the gradient.

As the plasmid DNA flows through the column the ethidium bromide will be retained in the resin and form a red band at the top. All of the DNA will flow through in the volume recommended. This procedure will dilute the CsCl sufficiently to allow the DNA to be precipitated.

Ethidium bromide can also be removed by extracting the plasmid DNA–ethidium bromide solution with an equal volume of TE-saturated n-butanol (UNIT 2.1). Shake the tube or vortex it vigorously to maximize the efficiency of extraction. Remove the organic upper phase and extract the aqueous phase repeatedly until no red color remains. Dilute the solution 3-fold with TE buffer to dilute the CsCl. This procedure generates contaminated organic solvent waste. Follow correct procedure for its disposal.

8. Ethanol precipitate plasmid DNA, using 2 vol of 100% ethanol at room temperature or -20°C , and centrifuge 10 min at $10,000 \times g$, 4°C .

Do not cool this solution below -20°C , as this may cause the CsCl to precipitate.

An alternative to ethanol precipitation at this step is to dialyze the plasmid DNA, from which the ethidium bromide has been removed, against 500 to 1000 vol TE buffer. Dialysis buffer should be changed at least twice with ≥ 2 hr between changes at room temperature, or 4 hr at 4°C (see APPENDIX 3C).

9. Wash pellet with 70% ethanol and dry under vacuum. Resuspend pellet in TE buffer and store at 4°C .

PLASMID DNA PURIFICATION BY PEG PRECIPITATION

ALTERNATE
PROTOCOL 3

Polyethylene glycol (PEG) precipitation is a rapid, reliable, and convenient method for plasmid DNA purification. It can be stopped at any step without affecting the ultimate recovery of plasmid DNA. No ultracentrifugation is required and the use of ethidium bromide (a mutagen) is avoided. RNA and chromosomal DNA contaminants are removed from a crude lysate pellet by treating it with RNase, NaOH/SDS, and potassium acetate. Plasmid DNA is extracted from the supernatant and precipitated with PEG. This method is suitable for preparing plasmid DNA for procedures that require plasmid DNA free of any contaminants.

Additional Materials (also see Basic Protocol 2)

- Glucose/Tris/EDTA solution (UNIT 1.6)
- 0.2 M NaOH/1% (w/v) SDS [prepare fresh from 10 M NaOH and 10% (w/v) SDS stocks]
- 10 mg/ml DNase-free RNase (UNIT 3.13)
- 3 M potassium acetate solution, pH ~5.5 (see recipe)
- Buffered phenol (UNIT 2.1)
- 24:1 (v/v) chloroform/isoamyl alcohol (UNIT 2.1)
- 10 M ammonium acetate (APPENDIX 2)
- PEG solution (see recipe)
- 3 M sodium acetate, pH 5.5 (APPENDIX 2)
- Sorvall SS-34 or Beckman JA-17 rotor (or equivalent)
- Sorvall HB-4 rotor

Remove contaminants from crude lysate pellet

1. Resuspend the pellet obtained in the final step of crude lysate preparation in 1 ml glucose/Tris/EDTA solution.
2. Add RNase to a final concentration of 20 µg/ml and incubate 20 min at 37°C.
3. Add 2 ml freshly prepared 0.2 M NaOH/1% SDS, mix by inverting the tube, and let stand 5 to 10 min at room temperature.
4. Add 1.5 ml of 3 M potassium acetate solution, mix by inverting the tube, and let stand 5 to 10 min at room temperature.
5. Centrifuge 10 min at 20,000 × g (11,000 rpm in SS-34 rotor; 12,500 rpm in HB-4 or JA-17 rotor), room temperature.

If restriction fragments of the plasmid are to be prepared for use as probes in filter hybridizations, gel purification will rid the desired fragments of any contaminating traces of chromosomal DNA. In that case steps 1 to 5 can be omitted. Resuspend the pellet from the crude lysate in 4 ml TE buffer and add RNase to 20 µg/ml final. Incubate 20 min at 37°C and treat as described in steps 7 to 12 of this protocol.

6. Transfer the supernatant to a clean tube.

The white precipitate is primarily SDS-potassium complex. It includes any chromosomal DNA that remained in the crude lysate. The amount remaining after crude lysate preparation varies. It may not be necessary to remove the chromosomal DNA completely, depending on the procedures to which the plasmid DNA is to be subjected. However, some procedures, such as hybrid selection (UNIT 6.8), require that plasmid DNA be free of any contaminants.

7. Extract plasmid DNA with buffered phenol, then with 24:1 chloroform/isoamyl alcohol.

***Escherichia coli,
Plasmids, and
Bacteriophages***

1.7.9

8. Add $\frac{1}{4}$ vol of 10 M ammonium acetate (2 M final concentration) to the aqueous phase and mix. Add 2 vol of 100% ethanol and place tube in dry ice for 10 min.
9. Recover plasmid DNA by centrifuging 10 min at $10,000 \times g$ (8000 rpm in HB-4 or SS-34 rotor; 8500 rpm in JA-17 rotor), 4°C.
10. Wash the pellet with 70% ethanol and dry briefly under vacuum.

PEG precipitate plasmid DNA

11. Resuspend the pellet in 2 ml TE buffer and add 0.8 ml PEG solution. Incubate 1 to 15 hr at 0°C.

The percentage of plasmid DNA recovered with PEG precipitation increases with time of incubation at 0° to 4°C. Usually $\geq 50\%$ (~0.5 to 3 mg) can be recovered by centrifuging the PEG-precipitated DNA solution after 1 hr at 0°C. Incubating the supernatant ≥ 12 hr at 4°C will permit complete recovery of remaining plasmid DNA while plasmid DNA collected after the first centrifugation can be used for other procedures.

12. Recover plasmid DNA by centrifuging 20 min at $10,000 \times g$ (8000 rpm in HB-4 or SS-34 rotor; 8500 rpm in JA-17 rotor; 10,000 rpm in most microcentrifuges), 4°C.
13. Resuspend pelleted plasmid DNA in 1 ml TE buffer. Ethanol precipitate plasmid DNA using 3 M sodium acetate, pH 5.5.

ALTERNATE PROTOCOL 4

PLASMID DNA PURIFICATION BY ANION-EXCHANGE OR SIZE-EXCLUSION CHROMATOGRAPHY

Chromatographic methods for purifying plasmid DNA take advantage of distinctions between the physical properties of plasmid DNA and those of molecules that copurify with it in the crude lysate. Nucleic acids are negatively charged and can therefore be purified away from contaminants using anion-exchange chromatography (see *UNIT 2.1B* for a protocol and discussion of one anion-exchange method). Similarly, the large size of plasmid DNA allows it to be purified away from smaller contaminants by gel-filtration chromatography. The specific properties of most matrices provided by commercial suppliers are unknown, due to the reluctance of the suppliers to divulge proprietary information, but they typically exploit one or both of these methods.

This protocol describes modifications for preparing a crude lysate for chromatographic purification of plasmid DNA and discusses features of three commercially available columns. Because no single protocol is appropriate for all of the chromatography methods, it is important to adhere to the manufacturer's suggested methodology. Most columns are supplied as kits that include reagents for preparing crude lysate using alkaline lysis (see Basic Protocol 1; e.g., Qiagen Plasmid Kits). Most manufacturers will also provide the columns separately.

Preparation of Crude Lysate

It is important to match the bacterial culture volume and cell density to the quantity of DNA expected and the plasmid DNA binding capacity of the column to be used. For large-scale plasmid preparation, most commercial suppliers provide a range of column sizes, with plasmid DNA yields of up to 10 mg. If too much culture volume is used, alkaline lysis will be inefficient, the column will be overloaded, and the performance of the system reduced. Furthermore, the excessive viscosity of the lysate will require vigorous mixing, resulting in shearing of bacterial genomic DNA and subsequent contamination of the plasmid DNA.

Successful chromatographic purification of DNA using the Qiagen-tips requires that the bacterial cells be incompletely lysed, so lysozyme should be omitted from the preparation of crude lysate.

The most frequent contaminant of plasmid DNA prepared by chromatographic methods is high-molecular-weight RNA. This contamination is reduced by adding 50 µg/ml RNase A (from frozen 1 mg/ml stock, *UNIT 3.13*) to the resuspended cell pellet. The reagents provided with Qiagen-tips contain RNase A. However, RNase A is stable for no more than several months at 4°C, the storage temperature of the working solutions. The Qiagen protocol uses RNase at 100 µg/ml; following addition of the RNase to the bacterial pellet resuspension buffer (P1), it is stable for 6 months at 4°C. The reagents provided by 5 Prime→3 Prime include a mixture of RNase A and RNase T1. The manufacturer suggests adding the RNases to a resuspended crude lysate pellet (see Basic Protocol 1, step 11) because the pellet is resuspended in a smaller volume at this step and less RNase is required.

Column flow is greatly impeded or completely prevented by the presence of solid material in the lysate when it is loaded onto the column. The most common source of solids is floating material from the precipitation (see Basic Protocol 1, step 7). To be certain that no floating material is loaded onto the column, decant the supernatant through cheesecloth, add chloroform before centrifugation, or recentrifuge the supernatant. The Qiagen protocol allows the lysate to be loaded on the Qiagen-tip column directly following removal of the protein and cellular debris by precipitation and centrifugation (see Basic Protocol 1, steps 7 to 9). Alternatively, Qiagen supplies a Qiafilter which rapidly filters bacterial lysates without centrifugation. The Qiafilter is available in syringe form for filtration of lysates from ≤250 ml of culture, or in cartridge format for filtering lysates from ≤2.5 liters of culture. These filtration units efficiently remove protein, cell debris, and even small SDS precipitates. Following lysate clearing using either centrifugation or filtration through a Qiafilter, the plasmid DNA can be purified directly on a Qiagen-tip column. The pZ523 protocol requires the crude lysate plasmid DNA to be precipitated with isopropanol. This reduces the volume of material to be loaded on the column and allows the buffer in which plasmid DNA is loaded to be optimized for the column. Isopropanol precipitation can be used with either of the alternate protocols for preparation of crude lysate, and the final pellet can be resuspended in the buffer appropriate for the chromatographic matrix to be used.

Column Capacity

Overloading columns with large lysate volumes can affect separation characteristics, leading to reduced yield and purity. Suppliers offer a range of column sizes to allow maximum plasmid DNA binding and recovery from large bacterial lysate volumes. The maximum plasmid-binding capacities of Qiagen tips 2500 and 10000 are 2.5 mg and 10 mg, respectively. The pZ523 column has a capacity of 4 to 5 mg and does not require that DNA bind to the column, so overloading is less likely to be a problem. The standard protocols for most other commercial columns are adjusted to provide a “good” yield for a plasmid that is maintained at a moderate copy number when cells that bear it are grown in ~500 ml LB medium.

To optimize plasmid DNA recovery, culture volume, culture medium, plasmid copy number, size of the insert, and the host strain should be considered. Very-low-copy-number plasmids and cosmids often require large culture volumes to yield significant amounts of DNA. When preparing DNA from high-copy-number plasmids such as pUC and its derivatives, culture volumes one-fourth to one-half the standard volume may be appropriate. When using media that support growth to high cell density such as terrific broth (*UNIT 1.1*; Tartoff and Hobbs, 1987), culture volumes one-fourth to one-tenth the standard

may be appropriate. The protocol provided by Qiagen includes a table of recommended culture volumes for plasmid isolation from a variety of sources, including high- to very-low-copy-number plasmids.

Plasmid DNA exceeding the capacity of the column will in no way prevent recovery of DNA. The excess DNA will simply run through the column and be discarded. One way to increase the yield is to recover and save the material that flows through the column when it is initially loaded. Some columns can be regenerated following elution of the plasmid DNA and the initial flowthrough reloaded. Qiagen resin is stable for up to 6 hr after equilibration and Qiagen-tips may be reused within 6 hr for the same sample. Beyond this time the separation characteristics of the resin will begin to change. The manufacturer's recommendations should be followed. Columns such as pZ523, which require vacuum and centrifugation respectively, collapse during use and cannot be reused. In this case, excess DNA in the flowthrough can be purified on a second column of the same type.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

CsCl/TE solution

100 ml TE buffer, pH 7.5 (APPENDIX 2)

100 g CsCl

Store indefinitely at room temperature

Dowex AG50W-X8 cation exchange resin

Prepare large batches (200 to 400 ml packed resin) of Dowex AG50W-X8 resin (100- to 200-mesh; Bio-Rad) by performing the following series of washing steps. Use a large Buchner funnel and filter paper to collect the resin between changes of wash solution.

1. Wash resin in ≥ 10 vol of 0.5 N NaOH until no color is observed in wash solution (resin will retain its buff color).
2. Wash with 5 to 10 vol of 0.5 N HCl.
3. Wash with 5 to 10 vol of 0.5 M NaCl.
4. Wash with 5 to 10 vol of distilled H₂O.
5. Wash with 5 to 10 vol of 0.5 N NaOH.
6. Wash with distilled H₂O until pH = 9.
7. Store prepared resin indefinitely in 0.5 M NaCl/0.1 M Tris (pH 7.5) at 4°C.

Polyethylene glycol (PEG) solution

30% (w/v) PEG 8000

1.6 M NaCl

Store indefinitely at 4°C

Potassium acetate solution (3 M), pH ~5.5

294 g potassium acetate (3 M final)

50 ml 90% formic acid (1.18 M final)

H₂O to 1 liter

Store indefinitely at room temperature

Sucrose/Tris/EDTA solution

25% (w/v) sucrose

50 mM Tris-Cl, pH 8.0 (APPENDIX 2)

100 mM EDTA, pH 8.0 (APPENDIX 2)

Store indefinitely at 4°C

Triton lysis solution

3% (v/v) Triton X-100
200 mM EDTA, pH 8.0 (APPENDIX 2)
150 mM Tris·Cl, pH 8.0 (APPENDIX 2)
Store indefinitely at 4°C

COMMENTARY

Critical Parameters and Troubleshooting

Preparation of crude lysate by alkaline lysis. This is a reliable procedure and one in which few irretrievable disasters can occur. One potential problem is failure of chromosomal DNA and proteins to precipitate after addition of potassium acetate solution. The cause of this is probably improper pH of the potassium acetate solution. The NaOH/SDS solution denatures linear (chromosomal) DNA. When the solution is neutralized in the presence of a high salt concentration (as when potassium acetate solution is added), the linear DNA precipitates. This precipitation presumably is due to inter-strand reassociation of denatured, linear DNA molecules at multiple sites. It results in the formation of a large, insoluble DNA network. Protein–SDS complexes also precipitate under these conditions. If the pH of the potassium acetate solution is not ~5.5, these precipitates will not form. If precipitation fails to occur because the pH is incorrect, the preparation can be saved by adding concentrated formic or acetic acid dropwise to the solution. Mix after each addition until the viscosity decreases, which will happen suddenly. A precipitate will then appear.

Preparation of crude lysate by the boiling method. The drawback to this procedure is that if it fails there are few chances for recovery. The only easily assayed step is the final recovery of plasmid DNA. Failure is often caused by inactive lysozyme or incorrect boiling time. An indication of failure will be that the solution is not extremely viscous following step 4. The remedy for inactive lysozyme is to try a new preparation of lysozyme. The optimal boiling time may vary slightly between bacterial strains. Using a strain for which the correct time is already known is the simplest remedy to this problem. Alternatively, the correct time may be determined empirically by performing the boiling miniprep procedure in UNIT 1.6.

Preparation of crude lysate by Triton lysis. This method is much gentler than the others described in this unit because of the relatively mild conditions used to disrupt the cells (i.e.,

use of Triton X-100 and lysozyme rather than boiling or severe changes in ionic strength and pH). Because the conditions used are near the lower limits of their effectiveness, solutions must be prepared correctly and the lysozyme must be active. Although powdered lysozyme is stable for long periods when stored properly, it is occasionally necessary to purchase and use a new bottle. The solution should become extremely viscous when Triton lysis solution is added, indicating that lysis has occurred. Centrifugation is also critical for separating plasmid DNA from the bulk of contaminants. If the pellet is not reasonably firm at this stage it will pour out of the tube when the plasmid DNA-containing supernatant is decanted. If this happens, repeat the centrifugation. If it happens routinely, increase the reagent volumes proportionately throughout the procedure or centrifuge for longer times or at higher *g* forces, if necessary using an ultracentrifuge.

Plasmid DNA purification by CsCl/ethidium bromide centrifugation. It is important that the density of the DNA–CsCl/ethidium bromide solution be correct for this procedure to work. Therefore, be precise (using graduated plastic tubes) when measuring the solution volume as this determines the amount of CsCl to be added. If the amount of added CsCl is incorrect, the position of the bands will be high (if too much was added) or low (if too little was added). Ideally the bands should appear slightly above the center of the tube. It is also important to allow sufficient time for the establishment of the gradient during centrifugation. If the bands appear diffuse at the end of the run, resolution is not adequate and centrifugation should be continued. In a gradient that has achieved equilibrium, bands are well defined.

Occasionally an ultracentrifuge tube breaks and its contents leak into the rotor compartment. This can occur if the tube is defective or if it is not filled or sealed properly. If most of the solution remains in the rotor compartment it can be pipetted into a clean tube, topped up with CsCl/TE solution, and centrifuged again. The quality of the DNA should not be affected,

***Escherichia coli,
Plasmids, and
Bacteriophages***

1.7.13

although the yield will be lower due to loss of material. The rotor should be rinsed with warm water if it comes in contact with CsCl solution. Cesium chloride is very corrosive and can cause pitting and weakening of the rotor. The centrifuge chamber should be inspected after every run, and cleaned with warm water and dried thoroughly if red stains are evident.

If the CsCl precipitates, as evidenced by a large, white, crystalline pellet observed when precipitating plasmid DNA in the final steps, warm the solution to room temperature and centrifuge it at room temperature. If this problem persists, be certain the solution is diluted 3-fold with TE buffer before adding ethanol.

Plasmid DNA purification by PEG precipitation. The purity of plasmid DNA obtained by polyethylene glycol (PEG) precipitation of crude bacterial lysate depends on the amount of chromosomal DNA remaining in the solution when the PEG solution is added. Steps 1 to 5 of Alternate Protocol 3 remove any remaining traces of chromosomal DNA. These steps are not necessary for PEG precipitation to succeed. For size fractionation by PEG to be effective, the concentration of DNA in the crude lysate must be $>10 \mu\text{g/ml}$. This is not a concern when applied to plasmid DNA purification, where the concentrations should be orders of magnitude greater than that figure.

Plasmid DNA purification by anion-exchange or size-exclusion chromatography. The columns available in kit form from a number of manufacturers are quite reliable. Almost all necessary reagents are provided, including common buffers such as TE buffer. Because the composition of the matrix is undisclosed, it is impossible to evaluate the procedure carefully and attempt to optimize it. Therefore, users of kits are strongly encouraged to follow the manufacturer's recommended procedures. Additional discussion of one type of anion-exchange procedure (Qiagen) can be found in the Commentary of *UNIT 2.1B*.

A major drawback to some prepared columns is the limited capacity of the matrix. A 500-ml culture of plasmid-containing bacteria can often yield 2 to 8 mg of plasmid DNA in the crude lysate. Some columns routinely yield only 500 to 1000 μg purified plasmid DNA, which is adequate for most purposes. Recovering the column flowthrough and rechromatographing it is the most practical method of increasing recovery of plasmid DNA. Not all columns can be reused, however, and recovery may require use of additional columns. Alter-

natively, smaller culture volumes can be used as suggested. If optimal recovery of DNA is desired, CsCl/ethidium bromide centrifugation or PEG precipitation should be used.

DNA obtained from chromatographic purification of plasmid DNA is comparable in quality to that prepared by the other methods. It is of sufficient purity for virtually any procedure for which it can be used. For example, plasmid DNA prepared using Qiagen-tips is at least equivalent in purity to that obtained following two rounds of CsCl-gradient centrifugation. It is suitable for all applications from cloning to transfection, radioactive and fluorescent sequencing, and gene therapy research. Qiagen also offers an endotoxin-free plasmid DNA kit which yields plasmid DNA containing lower levels of endotoxins than DNA purified by two rounds of CsCl centrifugation, which is essential in gene therapy studies.

The two most frequent contaminants are chromosomal DNA and high-molecular-weight RNA. These contaminants may be detected by the presence of large, diffuse ethidium bromide-binding material in agarose gel electrophoresis of purified plasmid DNA. To prevent such contamination, follow the manufacturer's suggestions for the use of RNase.

Only common laboratory equipment is required for chromatographic purification of plasmid DNA. Another consideration with commercial kits is the large amount of packaging material and waste. In addition to the plastic columns and excess packaging, kits contain standard reagents supplied in plastic bottles. These reagents are solutions of buffers, salts, ethanol, and detergent—all of which can be, and usually are, prepared in the lab. Most suppliers will provide the column without the reagents. However, the major advantage of using prepared anion-exchange column kits includes the absence of organic extractions and exposure to toxic chemicals such as phenol, chloroform, ethidium bromide, and CsCl.

The toxicity of DNA prepared by several methods has been assessed by performing a biological assay. Crude lysate from a 1-liter culture of plasmid-containing bacteria was prepared by the alkaline lysis procedure and divided into four equal aliquots. The aliquots were then subjected to purification by CsCl/ethidium bromide centrifugation, PEG precipitation, or chromatography on Qiagen and pZ523 columns. Purified plasmid DNA was injected into *Drosophila* embryos, and the frequency of germline transformation and killing

of injected embryos was determined. No significant differences attributable to the method of purification were observed.

Anticipated Results

Most plasmids currently used are derivatives of the pUC series (Fig. 1.5.2). These plasmids contain an origin of replication significantly more efficient than that of the previous generation of pBR322-derived plasmids. This allows recovery of 1 to 5 mg of plasmid DNA (free of contaminating bacterial products) from a 500-ml culture following any of the crude lysate preparation methods or PEG precipitation. Purification by CsCl/ethidium bromide density gradient centrifugation yields 75% to 90% of the amount of plasmid DNA obtained using PEG precipitation. Yields obtained from column chromatography are limited by the capacity of the column and are generally <1 mg.

Time Considerations

Cell growth and concentration require one overnight growth period to collect the starting 5-ml culture and most of the next day for its outgrowth and concentration. Crude lysate preparation can be completed in 90 min using alkaline lysis, ~35 min using the boiling method, and 2.5 to 3 hr using Triton lysis. Plasmid DNA purification by CsCl/ethidium bromide density gradient centrifugation takes 4 hr to 3 days depending upon the quality of plasmid DNA desired and the type of centrifuge and rotor used. PEG precipitation of the crude lysate can yield pure plasmid DNA in as little as 2 hr; however, complete recovery may require 13 to 16 hr. Column chromatography takes 30 to 90 minutes. High-quality plasmid DNA recovery from an *E. coli* strain containing the desired plasmid using any combination of these procedures takes 1 day to 1 week.

In addition to the total time necessary to obtain pure plasmid DNA, it is important to consider the hands-on time required for individual steps of the different procedures. PEG precipitation can yield pure plasmid DNA 4 or 5 hr after harvesting cells but requires direct attention and manipulations every 10 to 30 min. Purification using CsCl/ethidium bromide centrifugation takes 6 to 18 hr after harvesting cells, but the final 4-5 hr are taken up by a centrifugation step that requires no direct attention. Column chromatography can yield purified DNA within 3 hr of harvesting cells. With Qiagen protocols, plasmid purification times can also be reduced by using a Qiafilter to clear the bacterial lysate. In addition, most of these

columns operate by gravity flow for maximum handling convenience; the columns never run dry. Therefore, if pure plasmid DNA is needed the same day the large culture is harvested, it is perhaps best to prepare crude lysates by alkaline lysis and purify plasmid DNA by PEG precipitation, collecting the PEG precipitate after ~1 hr at 0°C, or by chromatography. If it has been a long day and the priority is to go home, crude lysate can be prepared by the boiling method and plasmid DNA purified by the CsCl/ethidium bromide centrifugation. The variety of methods and the opportunity to interrupt them at different steps (such as at the precipitation steps) facilitate selection of methods most convenient for specific work situations.

Literature Cited

- Clewell, D.B. and Helinski, D.R. 1970. Properties of a deoxyribonucleic acid-protein relaxation complex and strand specificity of the relaxation event. *Biochemistry* 9:4428-4440.
- Clewell, D.B. and Helinski, D.R. 1972. Nature of ColE1 plasmid replication in *Escherichia coli* in the presence of chloramphenicol. *J. Bacteriol.* 110:667-676.
- Tartoff, K.D. and Hobbs, C.A. 1987. Improved media for growing plasmid and cosmid clones. *BRL FOCUS* 9:2-12.

Key References

- Birnboim, H.C. 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods Enzymol.* 100:243-255.
- Birnboim, H.C. 1988. Citation classic. *Current Contents (Life Sciences)* 31(45):12.
- Birnboim, H.C. and Doly, J. 1979. A rapid alkaline extraction method for screening recombinant plasmid DNA. *Nucl. Acids Res.* 7:1513-1523.

These three references describe preparation of crude lysate by alkaline lysis.

- Holmes, D.S. and Quigley, M. 1981. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* 114:193-197.

Describes preparation of crude lysate by boiling.

- Clewell and Helinski, 1970, 1972. See above.

These two references describe preparation of crude lysate by Triton X-100 lysis.

- Radloff, R., Bauer, W., and Vinograd, J. 1967. A dye-buoyant-density method for the detection and isolation of closed circular duplex DNA: The closed circular DNA in HeLa cells. *Proc. Natl. Acad. Sci. U.S.A.* 57:1514-1521.

Describes plasmid DNA purification by CsCl/ethidium bromide centrifugation.

Lis, J.T. 1980. Fractionation of DNA fragments by polyethylene glycol induced precipitation. *Methods Enzymol.* 65:347-353.

Lis, J.T. and Schleif, R. 1975. Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. *Nucl. Acids Res.* 2:383-389. (Correction: the photograph of Fig. 1 should be interchanged with the photograph of Fig. 2. Lis and Schleif, 1975, *Nucl. Acids Res.* 2:757.)

These two references describe plasmid DNA purification by PEG precipitation.

Contributed by J.S. Heilig
University of Colorado
Boulder, Colorado

Karen L. Elbing
Clark and Elbing LLP
Boston, Massachusetts

Roger Brent
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

Introduction of Plasmid DNA into Cells

UNIT 1.8

Transformation of *E. coli* can be achieved using any of the four protocols in this unit. The first method (see Basic Protocol 1) using calcium chloride gives good transformation efficiencies, is simple to complete, requires no special equipment, and allows storage of competent cells. The one-step method (see Alternate Protocol 1) is considerably faster and also gives good transformation efficiencies (although they are somewhat lower). However, because it was developed relatively recently, its reproducibility and reliability are not as well established.

If considerably higher transformation efficiencies are needed, the third method (see Basic Protocol 2) using electroporation should be followed. Although this procedure is simple, fast, and reliable, it requires an electroporation apparatus. As in the calcium chloride protocol, prepared cells can be stored. The final method described (see Alternate Protocol 2) describes an adaptation based on electroporation that allows direct transfer of vector DNA from yeast into *E. coli*.

TRANSFORMATION USING CALCIUM CHLORIDE

**BASIC
PROTOCOL 1**

Escherichia coli cells are grown to log phase. Cells are concentrated by centrifugation and resuspended in a solution containing calcium chloride. Exposure to calcium ions renders the cells able to take up DNA, or *competent*. Plasmid DNA is mixed with the cells and presumably adheres to them. The mixture of DNA and cells is then heat shocked, which allows the DNA to efficiently enter the cells. The cells are grown in nonselective medium to allow synthesis of plasmid-encoded antibiotic resistance proteins, then plated on antibiotic-containing medium to allow identification of plasmid-containing colonies.

Materials

Single colony of *E. coli* cells
LB medium (UNIT 1.1)
CaCl₂ solution (see recipe), ice cold
LB plates (UNIT 1.1) containing ampicillin (Table 1.4.1)
Plasmid DNA (UNITS 1.6 & 1.7)

Chilled 50-ml polypropylene tubes
Beckman JS-5.2 rotor or equivalent
42°C water bath

Additional reagents and equipment for growth of bacteria in liquid media (UNIT 1.2)

NOTE: All materials and reagents coming into contact with bacteria must be sterile.

Prepare competent cells

1. Inoculate a single colony of *E. coli* cells into 50 ml LB medium. Grow overnight at 37°C with moderate shaking (250 rpm; see UNIT 1.2).

Alternatively, grow a 5-ml culture overnight in a test tube on a roller drum.

2. Inoculate 4 ml of the culture into 400 ml LB medium in a sterile 2-liter flask. Grow at 37°C, shaking at 250 rpm, to an OD₅₉₀ of 0.375.

This procedure requires that cells be growing rapidly (early- or mid-log phase). Accordingly, it is very important that the growing cells have sufficient air. A 1-liter baffled flask can be used instead of the 2-liter flask. Overgrowth of culture (beyond OD₅₉₀ of 0.4) decreases the efficiency of transformation.

***Escherichia coli*,
Plasmids, and
Bacteriophages**

3. Aliquot culture into eight 50-ml prechilled, sterile polypropylene tubes and leave the tubes on ice 5 to 10 min.

Cells should be kept cold for all subsequent steps.

Larger tubes or bottles can be used to centrifuge cells if volumes of all subsequent solutions are increased in direct proportion.

4. Centrifuge cells 7 min at $1600 \times g$ (3000 rpm in JS-5.2), 4°C. Allow centrifuge to decelerate without brake.

We have not attempted to determine whether deceleration without braking is critical to this procedure. However, we do not use the brake for this step or for the subsequent centrifugation steps.

5. Pour off supernatant and resuspend each pellet in 10 ml ice-cold CaCl_2 solution.

Resuspension should be performed very gently and all cells kept on ice.

6. Centrifuge cells 5 min at $1100 \times g$ (2500 rpm), 4°C. Discard supernatant and resuspend each pellet in 10 ml ice-cold CaCl_2 solution. Keep resuspended cells on ice for 30 min.

7. Centrifuge cells 5 min at $1100 \times g$, 4°C. Discard supernatant and resuspend each pellet in 2 ml ice-cold CaCl_2 solution.

It is important to resuspend this final pellet well. The suspension may be left on ice for several days. For many strains (e.g., DH1) competency increases with increasing time on ice, and reaches a maximum at 12 to 24 hr. This is not true for MC1061 cells, which should be frozen immediately.

8. Dispense cells into prechilled, sterile polypropylene tubes (250- μl aliquots are convenient). Freeze immediately at -70°C .

Assess competency of cells

9. Use 10 ng of pBR322 to transform 100 μl of competent cells according to the steps provided below. Plate appropriate aliquots (1, 10, and 25 μl) of the transformation culture on LB/ampicillin plates and incubate at 37°C overnight.
10. Calculate the number of transformant colonies per aliquot volume (μl) $\times 10^5$: this is equal to the number of transformants per microgram of DNA.

*Transformation efficiencies of 10^7 to 10^8 and 10^6 to 10^7 are obtained for *E. coli* MC1061 and DH1, respectively. Competency of strains decreases very slowly over months of storage time.*

Transform competent cells

11. Aliquot 10 ng of DNA in a volume of 10 to 25 μl into a sterile 15-ml round-bottom test tube and place on ice.

Plasmid DNA can be used directly from ligation reactions. When this is done, more DNA is usually used. However, if there is $>1 \mu\text{g}$ of DNA in the ligation reaction, or if the ligation reaction is from low gelling/melting temperature agarose, it is wise to dilute the ligation mix (see UNIT 3.16).

12. Rapidly thaw competent cells by warming between hands and dispense 100 μl immediately into test tubes containing DNA. Gently swirl tubes to mix, then place on ice for 10 min.

Competent cells should be used immediately after thawing. Remaining cells should be discarded rather than refrozen.

- Heat shock cells by placing tubes into a 42°C water bath for 2 min.

Alternatively, incubate at 37°C for 5 min.

- Add 1 ml LB medium to each tube. Place each tube on a roller drum at 250 rpm for 1 hr at 37°C.

- Plate aliquots of transformation culture on LB/ampicillin or other appropriate antibiotic-containing plates.

It is convenient to plate several different dilutions of each transformation culture. The remainder of the mixture can be stored at 4°C for subsequent platings.

- When plates are dry, incubate 12 to 16 hr at 37°C.

ONE-STEP PREPARATION AND TRANSFORMATION OF COMPETENT CELLS

ALTERNATE PROTOCOL 1

This procedure is considerably easier than Basic Protocol 1 because it eliminates the need for centrifugation, washing, heat shock, and long incubation periods (Chung et al., 1989). Moreover, competent cells made by this simple procedure can be directly frozen at -70°C for long-term storage. A variety of strains can be made competent by this procedure, and the transformation frequency can be as high as that achieved by Basic Protocol 1. However, frequency is considerably lower than can be obtained by electroporation.

Additional Materials (also see *Basic Protocol 1*)

2× transformation and storage solution (TSS; see recipe), ice cold
LB medium (UNIT 1.1) containing 20 mM glucose

- Dilute a fresh overnight culture of bacteria 1:100 into LB medium and incubate at 37°C until the cells reach an OD₆₀₀ of 0.3 to 0.4.

The procedure will work if cells are harvested at other stages of the growth cycle (including stationary phase), but with reduced efficiency.

- Add a volume of ice-cold 2× TSS equal to that of the cell suspension, and gently mix on ice.

For long-term storage, freeze small aliquots of the suspension in a dry ice/ethanol bath and store at -70°C. To use frozen cells for transformation, thaw slowly and then use immediately.

Cells can also be used if pelleted by centrifugation 10 min at 1000 × g, 4°C, and this may increase the frequency of transformation (according to Chung et al., 1989). Discard supernatant and resuspend cell pellet at one-tenth of original volume in 1× TSS (prepared by diluting 2× TSS). Proceed with transformation as in step 3.

- Add 100 μl competent cells and 1 to 5 μl DNA (0.1 to 100 ng) to an ice-cold polypropylene or glass tube. Incubate 5 to 60 min at 4°C.

As is the case for related procedures, the transformation frequency as measured by transformants/μg DNA is relatively constant at amounts of DNA <10 ng. However, the frequency decreases at higher concentrations. The time of incubation at 4°C is relatively unimportant.

- Add 0.9 ml LB medium containing 20 mM glucose and incubate 30 to 60 min at 37°C with mild shaking to allow expression of the antibiotic resistance gene. Select transformants on appropriate plates.

It is unnecessary to heat shock the transformation mixture. The expected transformation frequency should range between 10⁶ and 10⁷ colonies/μg DNA.

Escherichia coli,
Plasmids, and
Bacteriophages

1.8.3

HIGH-EFFICIENCY TRANSFORMATION BY ELECTROPORATION

Electroporation with high voltage is currently the most efficient method for transforming *E. coli* with plasmid DNA. The procedure described may be used to transform freshly prepared cells or to transform cells that have been previously grown and frozen. With freshly grown cells, it routinely gives more than 10^9 bacterial transformants per microgram of input plasmid DNA.

Materials

- Single colony of *E. coli* cells
- LB medium (UNIT 1.1)
- H₂O, ice cold
- 10% glycerol, ice cold
- SOC medium (see recipe)
- LB plates (UNIT 1.1) containing antibiotics (Table 1.4.1)
- 1-liter centrifuge bottle, 50-ml narrow-bottom polypropylene tube, and microcentrifuge tubes, chilled ice cold
- Beckman J-6M centrifuge (or equivalent)
- Beckman JS-4.2 rotor (or equivalent) and adaptors for 50-ml narrow-bottom tubes
- Electroporation apparatus with a pulse controller or 200- or 400-ohm resistor
- Chilled electroporation cuvettes, 0.2-cm electrode gap
- Additional reagents and equipment for growth of bacteria in liquid media (UNIT 1.2)

NOTE: All materials and reagents coming into contact with bacteria must be sterile.

Prepare the cells

1. Inoculate a single colony of *E. coli* cells into 5 ml LB medium. Grow 5 hr to overnight at 37°C with moderate shaking (see UNIT 1.2).
2. Inoculate 2.5 ml of the culture into 500 ml LB medium in a sterile 2-liter flask. Grow at 37°C, shaking at 300 rpm, to an OD₆₀₀ of ~0.5 to 0.7.

Best results are obtained by harvesting cells at an OD₆₀₀ of ~0.5 to 0.6.

3. Chill cells in an ice-water bath 10 to 15 min and transfer to a prechilled 1-liter centrifuge bottle.

Cells should be kept at 2°C for all subsequent steps.

4. Centrifuge cells 20 min at 4200 rpm in Beckman J-6M, 2°C.
5. Pour off supernatant and resuspend the pellet in 5 ml ice-cold water. Add 500 ml ice-cold water and mix well. Centrifuge cells as in step 4.
6. Pour off supernatant immediately and resuspend the pellet by swirling in remaining liquid.

Because the pellet is very loose, the supernatant must be poured off immediately. The pellet can be made tighter by substituting ice-cold sterile HEPES (1 mM, pH 7.0) for the ice-cold water in step 5.

7. Add another 500 ml ice-cold water, mix well, and centrifuge again as in step 4.
8. Pour off supernatant immediately and resuspend the pellet by swirling in remaining liquid.

- 9a. If fresh cells are to be used for electroporation, place suspension in a prechilled, narrow-bottom, 50-ml polypropylene tube, and centrifuge 10 min at 4200 rpm in Beckman J-6M centrifuge with JS-4.2 rotor and adaptors, 2°C.

Fresh cells work better than frozen cells.

Estimate the pellet volume (usually ~500 μ l from a 500-ml culture) and add an equal volume of ice-cold water to resuspend cells (on ice). Aliquot 50- to 300- μ l cells into prechilled microcentrifuge tubes. The cell density is $\sim 2 \times 10^{11}$ /ml.

- 9b. If frozen cells are to be used for electroporation, add 40 ml ice-cold 10% glycerol to the cells and mix well. Centrifuge cells as described in step 9a.

Estimate the pellet volume and add an equal volume of ice-cold 10% glycerol to resuspend cells (on ice). Place 50- to 300- μ l aliquots of cells into prechilled microcentrifuge tubes and freeze on dry ice (not in liquid nitrogen). Store at -80°C .

Prolonged incubation of cells in ice-water at all stages can increase transformation efficiency of some strains, such as BW313/P3 and MC1061/P3, >3-fold.

Transform the cells

10. Set the electroporation apparatus to 2.5 kV, 25 μ F. Set the pulse controller to 200 or 400 ohms.

The pulse controller is necessary when high-voltage pulses are applied over short gaps in high-resistance samples (see Background Information).

11. Add 5 pg to 0.5 μ g plasmid DNA in 1 μ l to tubes containing fresh or thawed cells (on ice). Mix by tapping the tube or by swirling the cells with the pipettor.

12. Transfer the DNA and cells into a cuvette that has been chilled 5 min on ice, shake slightly to settle the cells to the bottom, and wipe the ice and water from the cuvette with a Kimwipe.

The volume of DNA added to the cells should be kept small. Adding DNA up to one-tenth of the cell volume will decrease the transformation efficiency 2- to 3-fold. Also, since the resistance of the sample should be high, make sure that addition of the DNA to the cells does not increase the total salt concentration in the cuvette by >1 mM.

13. Place the cuvette into the sample chamber.

If using a homemade apparatus, connect the electrodes to the cuvette.

14. Apply the pulse by pushing the button or flipping the switch.

15. Remove the cuvette. Immediately add 1 ml SOC medium and transfer to a sterile culture tube with a Pasteur pipet. Incubate 30 to 60 min with moderate shaking at 37°C.

If the actual voltage and time constant of the pulse are displayed on the electroporation apparatus, check this information. Verify that the set voltage was actually delivered, and record the time constant of the pulse so that you may vary it later if necessary (see Critical Parameters).

16. Plate aliquots of the transformation culture on LB plates containing antibiotics.

DIRECT ELECTROPORETIC TRANSFER OF PLASMID DNA FROM YEAST INTO *E. COLI*

The use of “shuttle vectors”—plasmids that can be grown successfully in at least two different organisms—facilitates the transfer of DNA between, for example, yeast and *E. coli*. In this adaptation of the electroporation protocol, plasmid DNA from a shuttle vector is transformed directly from yeast into *E. coli*. Components of the interaction trap/two-hybrid system (UNIT 20.1) are used as an example in this protocol. The transfer and selection of a “prey” plasmid from the yeast strain EGY48 into the *E. coli* strain KC8 is described here, but the approach can be adapted for use with other yeast and *E. coli* strains.

Additional Materials (also see Basic Protocol 2)

Single colony of *E. coli* KC8 cells (UNIT 20.1)

Streak colony of Trp⁻ plasmid-harboring EGY48 yeast cells on Gal/Raff/Xgal/CM plates (UNIT 20.1), no older than 2 weeks

M9 minimal medium and plates (UNIT 1.1) containing 100 µg/ml ampicillin (Table 1.4.1) and standard concentrations of leucine, histidine, and uracil

Additional reagents and equipment for growth and manipulation of yeast (UNIT 13.2) and for plasmid DNA miniprep (UNIT 1.6) or PCR (UNIT 15.1)

1. Prepare electrocompetent KC8 cells (see Basic Protocol 2, steps 1 to 9a), resuspending the final cell pellet in ice-cold water to obtain an OD₆₀₀ of 100.

Fresh KC8 cells work better in this electroporation method than frozen ones.

To measure OD, dilute 5 µl of the cell suspension with water to 1 ml and measure the OD₆₀₀. If necessary add more water to the suspension to get an OD₆₀₀ of 100.

2. Distribute 65-µl aliquots of the electrocompetent *E. coli* KC8 cells into ice-cold microcentrifuge tubes.
3. With a sterile wooden or plastic stick, scrape off ~10 µl of yeast from a streak colony of EGY48 harboring the respective “prey” plasmid derivative of pJG4-5 and grown on Gal/Raff/Xgal/CM plates. Resuspend the yeast cells in the KC8 suspension by swirling the stick used for scraping off the cells.

Avoid scraping off plate medium when collecting the yeast streak cells.

Keep the microcentrifuge tube on ice as much as possible. Try to get an even distribution of the two cell types but do not vortex. Yeast grown on plates other than Gal/Raff/Xgal/CM will probably work as well; do not worry if the yeast colony used is blue.

4. Set the electroporation apparatus to 1.5 kV, 25 µF, and the pulse controller to 100 ohms. Transfer the cell suspension into a 0.2-cm cuvette that has been chilled 5 min on ice, shake slightly to settle the cells to the bottom, and wipe the ice and water from the cuvette with a Kimwipe.

The use of Pasteur pipettes will facilitate placing the cell suspension at the bottom of the cuvette. Avoid any air bubbles.

5. Place the cuvette in the sample chamber of the apparatus and pulse. Take the cuvette out and place it on ice for ≥45 sec. Meanwhile, change the settings in preparation for the second pulse.

The expected time constant for the first pulse is 2.2 to 2.4 msec.

6. Set the electroporation apparatus to 2.5 kV, 25 µF, and the pulse controller to 200 ohms. Wipe the cuvette again, place it in the sample chamber, and pulse.

The expected time constant for the second pulse is 4.2 to 4.8 msec.

7. Remove the cuvette, immediately add 1 ml LB medium, and transfer the suspension into a microcentrifuge tube. Incubate 45 min at room temperature.

Incubation of the suspension after electroporation for 1 hr at 37°C decreases the yield of transformants, probably due to prolonged adhesion of the E. coli cells to the yeast cell debris.

8. Spread 150 μ l of the suspension evenly onto M9 minimal medium plates containing 100 μ g/ml ampicillin and leucine, histidine, and uracil. Incubate \geq 24 hr at 37°C.

A slight yeast background might appear on the plates but single E. coli colonies are easy to pick. Between 50 and 200 KC8 colonies have been obtained per plate employing 150 μ l out of the 1 ml LB suspension.

9. Pick a single KC8 colony, inoculate it into 1.5 to 5 ml M9 minimal medium (Leu⁺, His⁺, Ura⁺, 100 μ g/ml Amp) or LB (100 μ g/ml Amp) and grow at 37°C. Harvest at an appropriate OD to prepare miniprep DNA (UNIT 1.6) or perform PCR analysis (UNIT 15.1).

Using M9 minimal medium to grow KC8 in liquid culture is an additional safety measure but not absolutely necessary. It ensures isolation of the plasmid whose marker complements the auxotrophic defect in KC8; in addition, slightly increased plasmid copy number and improved DNA quality have been reported.

REAGENTS AND SOLUTIONS

CaCl₂ solution

60 mM CaCl₂

15% glycerol

10 mM PIPES [piperazine-*N,N'*-bis(2-hydroxypropanesulfonic acid)], pH 7

Filter sterilize using a disposable filter unit, or autoclave

Store at room temperature (stable for years)

SOC medium

0.5% yeast extract

2% tryptone

10 mM NaCl

2.5 mM KCl

10 mM MgCl₂

10 mM MgSO₄

20 mM glucose

Store at room temperature (stable for years)

Transformation and storage solution (TSS), 2×

Dilute sterile (autoclaved) 40% (w/v) polyethylene glycol (PEG) 3350 to 20% PEG in sterile LB medium containing 100 mM MgCl₂. Add dimethyl sulfoxide (DMSO) to 10% (v/v) and adjust to pH 6.5.

COMMENTARY

Background Information

Calcium and one-step transformation

Transformation of *E. coli* was first described by Mandel and Higa (1970). Subsequent modifications to improve transformation efficiencies have included prolonged exposure of cells to CaCl₂ (Dagert and Ehrlich, 1974), substitution of calcium with other cations such as Rb⁺

(Kushner, 1978), Mn²⁺, and K⁺, and addition of other compounds such as dimethyl sulfoxide, dithiothreitol, and cobalt hexamine chloride (Hanahan, 1983). Basic Protocol 1 given here provides good transformation efficiencies, permits long-term storage of competent cells, and is relatively uncomplicated to perform. Variations on this protocol can be obtained from the references provided. Alternate Proto-

***Escherichia coli,*
Plasmids, and
Bacteriophages**

1.8.7

col 1, a one-step preparation and transformation procedure, is considerably faster.

Transformation by electroporation

Electroporation has become a valuable technique for transfer of nucleic acids into eukaryotic and prokaryotic cells. The method can be applied to many different *E. coli* strains and to other gram-negative and gram-positive bacteria.

In this technique, a high-voltage electric field is applied briefly to cells, apparently producing transient holes in the cell membrane through which plasmid DNA enters (Shigekawa and Dower, 1988). The field strength used for mammalian cell and plant protoplast electroporation is usually 0.5 to 1 kV/cm. The field strength needed for high-efficiency transformation of *E. coli* is much greater, usually ~12.5 kV/cm. Under these conditions up to 10^{10} transformants/ μg plasmid DNA have been reported (Calvin and Hanawalt, 1988; Miller et al., 1988; Dower et al., 1988). Recently, field strength up to 8 kV/cm was also used successfully to electroporate both mammalian cells and plant protoplasts.

The capacitor discharge circuit of the electroporation apparatus typically generates an electrical pulse with an exponential decay waveform. The voltage across the electrodes rises rapidly to a peak voltage, which then declines over time as follows:

$$V_t = V_0 [e^{-t/T}]$$

where V_t = voltage at a given time t after the time of V_0 , V_0 = initial voltage, t = time (sec), T = pulse time constant = RC , R = resistance of circuit (ohms); and C = capacitance of circuit (Farads).

The pulse time constant is ~5 to 10 msec for electroporating *E. coli* cells and ranges from 5 μsec to 50 msec for higher eukaryotic cells.

The pulse controller contains a number of different-sized resistors, any one of which is placed in parallel with the sample, and one resistor of fixed (20-ohm) resistance, which is placed in series with the sample. The resistor placed in parallel with the sample (usually 200 or 400 ohms) swamps out the effect of changes in the resistance of the sample on the total resistance of the circuit, thus determining the total resistance across the capacitor, and controlling the time constant (T) of the capacitor discharge. The 20-ohm resistor in series with the sample protects the circuitry by limiting the current should a short circuit (arc) occur and the capacitor discharge instantly.

The pulse controller is required when high-voltage electroporation pulses are delivered to high-resistance samples across narrow electrode gaps. In this procedure, the resistance of washed *E. coli* in the 0.2-cm cuvette is ~5000 ohms. A pulse controller is not necessary when samples of low (<1000-ohm) resistance are used, for example for electroporation of mammalian cells suspended in PBS.

Direct transfer of plasmid DNA from yeast into *E. coli*

Alternate Protocol 2 presents an application of electrophoretic transformation whereby a "shuttle vector" (a vector designed to be used in at least two different organisms to facilitate interspecies transfer of DNA) can be directly transferred between two species. Shuttle vectors have become increasingly popular in recent years, with those designed to facilitate the transfer of plasmid DNA between yeast and *E. coli* proving to be particularly useful. With the widespread use of the two-hybrid system (or interaction trap; see UNIT 20.1), transfer of plasmid DNA from yeast into *E. coli* using shuttle vectors has become a common task. As an alternative to rescuing a shuttle plasmid from a yeast clone and subsequently transforming an appropriate *E. coli* strain, the procedure described in Alternate Protocol 2 bypasses the need for plasmid isolation and offers a one-step method to obtain the same result. The direct transfer method was first mentioned by Marcil and Higgins (1992) and further modified by Karen Clemens (NIH, Bethesda, Md., pers. comm.). Alternate Protocol 2 comprises an optimization of the procedure and an adaptation to the interaction trap; however, although outlined for that specific case, the procedure is generally applicable to other yeast and *E. coli* strains. As presented, this protocol accomplishes the transfer of a "prey" plasmid from the yeast strain EGY48, designed to be used in interactor hunts, into the *E. coli* strain KC8. EGY48 contains three different plasmids (bait, prey, and lacZ reporter), all of which confer ampicillin resistance if transferred into *E. coli*. Therefore, the TRP1 selectable marker of the new interactor (prey plasmid), which complements the Trp^- phenotype of EGY48 as well as KC8, is used to rescue the plasmid. During the electroporation procedure the yeast cells are destroyed in the first pulse and the liberated plasmids transformed into *E. coli* KC8 cells in the second pulse. Selection of the prey plasmid is achieved by choosing the correct medium on which to plate the transformed KC8.

Critical Parameters

Calcium transformation

In Basic Protocol 1, preparation of competent cells with a high transformation efficiency is thought to depend on (1) harvesting bacterial cultures in logarithmic phase of growth, (2) keeping cells on ice throughout the procedure, and (3) prolonged CaCl_2 exposure.

At least 30 min of growth in nonselective medium (*outgrowth*) after heat shock is necessary for plasmids containing the pBR322 tetracycline resistance promoter and gene to express enough of the protein to allow the cells to form colonies with an efficiency of 1 on tetracycline plates. Cells expressing the common plasmid-encoded ampicillin resistance (β -lactamase) gene may not require such prolonged outgrowth to form colonies on ampicillin plates. When an ampicillin-resistant plasmid is used, transformation mixtures should be diluted so that transformed colonies arise at a relatively low density (≤ 500 cells/plate). Otherwise, the β -lactamases present in the colonies may lower the ampicillin level in the plate near them, and permit growth of weakly ampicillin-resistant *satellite colonies*. This problem can be ameliorated if carbenicillin (a related antibiotic slightly less sensitive to destruction by the pBR322 β -lactamase) is substituted for ampicillin in the medium. Carbenicillin should be used at a concentration of 50 to 100 $\mu\text{g}/\text{ml}$.

Usually only 3% to 10% of cells are competent to incorporate plasmid DNA. Transformation frequencies decrease with increasing plasmid size (Hanahan, 1983). The number of transformants obtained usually increases linearly with increasing numbers of plasmid molecules up to a point, reached at ~ 10 ng DNA/100 μl competent cells in the procedure given here. After this point the number of transformants does not increase linearly with increasing numbers of plasmid DNA molecules.

Transformation by electroporation

Although the procedure works with cells grown to many different densities, best results are obtained when cells are harvested at an OD_{600} of 0.5 to 0.6. After the cells are centrifuged in water, the cell pellet is very loose and the supernatant should be poured off as quickly as possible to prevent a big loss in yield; to achieve this, it is best to handle no more than two centrifuge bottles simultaneously. Fresh

cells in either water or 10% to 20% glycerol usually work better than frozen cells.

As described above, the relevant parameters for exponential electroporation pulses are the time constant of the exponential curve (how long the pulse lasts) and the initial voltage or field strength (how strong the pulse is). Generally speaking, successful electroporation of *E. coli* requires long, strong pulses. In this procedure the capacitance of the circuit is relatively large (25 μF), ensuring a relatively long pulse; fine tuning of the pulse time constant is achieved by varying the size of the resistor placed in parallel with the sample.

The SOC medium must be added immediately after electroporation. Do not let the electroporated cells sit in the cuvette.

In the procedure given here, the number of transformants increases linearly with input DNA over a very wide range (from 5 pg to 500 ng). The amount of plasmid DNA added can be as little as 4 to 5 pg in 50 μl of cells, and as much as 0.5 μg in 300 μl of cells, without affecting the transformation efficiency significantly. However, transformation efficiency falls off with < 5 pg of DNA; in one experiment, the number of transformants obtained with 1 pg of plasmid DNA was 30 times lower than with 5 pg.

The size of the DNA does not seem to be important for this procedure; plasmid DNA can be as large as 14 kb without significant effects on transformation efficiency. Religated DNA (still in the ligation mix) can be transformed almost as efficiently as intact supercoiled DNA.

The general recommendations given for electroporation apply to the direct transfer method as well. Although outlined for a specific case, the procedure described in Alternate Protocol 2 should be generally applicable to other yeast and *E. coli* strains. The yeast streak colony should not be older than 2 weeks. Scraping medium off the plate when collecting the yeast cells should be avoided. Fresh KC8 cells work better than frozen ones. Incubation of the 1 ml LB suspension after electroporation at 37°C for 1 hr decreases the yield of transformants, probably due to prolonged adhesion of the *E. coli* cells to the yeast cell debris. Using M9 minimal medium to grow KC8 in liquid culture is an additional safety measure, but is not absolutely necessary. It ensures isolation of the plasmid whose marker complements the auxotrophic defect in KC8; in addition slightly increased plasmid copy number and improved DNA quality have been reported.

Anticipated Results

Calcium and one-step transformation. In Basic Protocol 1, transformation efficiencies of 10^7 to 10^8 and 10^6 to $10^7/\mu\text{g}$ plasmid DNA should be obtained for *E. coli* MC1061 and DH1, respectively. In Alternate Protocol 1, transformation frequencies should range from 10^6 to 10^7 colonies/ μg DNA.

Transformation by electroporation. Using Basic Protocol 2, efficiencies of $2.5\text{--}14 \times 10^{10}$ transformants/ μg have been obtained with superpure pUC19 DNA (from BRL) and $6.2\text{--}12 \times 10^9$ transformants/ μg with home-made pUC18 DNA and cDNA libraries in MC1061/P3. In addition, 5×10^9 transformants/ μg have been obtained with the expression plasmid CDM8 in MC1061/P3. Similar results may be anticipated with MC1061 (see Table 1.4.5) and with many other commonly used lab strains. Using direct plasmid transfer from yeast to *E. coli* (Alternate Protocol 2), between 50 and 200 KC8 colonies have been obtained per plate when 150 μl out of the 1 ml LB suspension was employed.

Time Considerations

Calcium and one-step transformation. Growth of competent cells from an aliquot of an overnight culture to logarithmic phase requires ~3 hr. In Basic Protocol 1, cells are then exposed to calcium as long as overnight. Once competent cells are available, transformation requires ~90 min for either strain. In Alternate Protocol 1, preparation and transformation of competent cells requires 1 to 2 hr.

Transformation by electroporation. In Basic Protocol 2, once the culture of bacterial cells is ready to be harvested, the cells can be washed and concentrated within an hour. Electroporation takes only a few minutes, and growth and plating of the transformed cells should not take more than 90 min. The time frame for direct plasmid transfer (Alternate Protocol 2) is essentially the same as for the electroporation basic protocol. It should be noted that using M9 minimal medium slows down the growth of *E. coli*.

Literature Cited

Calvin, N.M. and Hanawalt, P.C. 1988. High-efficiency transformation of bacterial cells by electroporation. *J. Bacteriol.* 170:2796-2801.

Chung, C.T., Niemela, S.L., and Miller, R.H. 1989. One-step preparation of competent *Escherichia coli*: Transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. U.S.A.* 86:2172-2175.

Dagert, M. and Ehrlich, S.D. 1974. Prolonged incubation in calcium chloride improves competence of *Escherichia coli* cells. *Gene* 6:23-28.

Dower, W.J., Miller, J.F., and Ragsdale, C.W. 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucl. Acids Res.* 16:6127-6145.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166:557-580.

Kushner, S.R. 1978. An improved method for transformation of *Escherichia coli* with *ColEI* derived plasmids. In *Genetic Engineering* (H.W. Boyer and S. Nicosia, eds.) pp. 17-23. Elsevier/North Holland, Amsterdam.

Mandel, M. and Higa, A. 1970. Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.* 53:159-162.

Marcil, R. and Higgins, D.R. 1992. Direct transfer of plasmid DNA from yeast to *E. coli* by electroporation. *Nucl. Acids Res.* 20:917.

Miller, J.F., Dower, W.J., and Tompkins, L.S. 1988. High-voltage electroporation of bacteria: Genetic transformation of *Campylobacter jejuni* with plasmid DNA. *Proc. Natl. Acad. Sci. U.S.A.* 85:856-860.

Shigekawa, K. and Dower, W.J. 1988. Electroporation of eukaryotes and prokaryotes: A general approach to the introduction of macromolecules into cells. *BioTechniques* 6:742-751.

Key References

Dower et al., 1988. See above.

The paper from which the second basic protocol was derived, and the highest-efficiency E. coli transformation by electroporation published to date.

Hanahan, 1983. See above.

An extremely thorough explanation of the parameters affecting transformation efficiency.

Contributed by Christine E. Seidman
(calcium) and Kevin Struhl (one-step)
Harvard Medical School
Boston, Massachusetts

Jen Sheen (electroporation)
Massachusetts General Hospital
Boston, Massachusetts

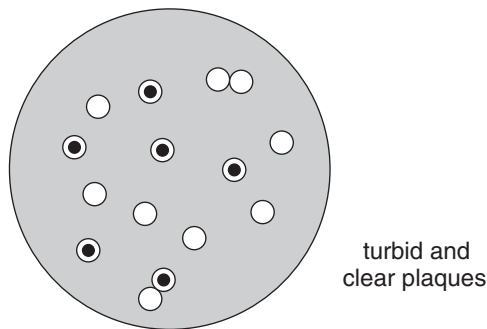
Timm Jessen (direct transfer)
Hoechst AG
Frankfurt am Main, Germany

VECTORS DERIVED FROM LAMBDA AND RELATED BACTERIOPHAGES

Introduction to Lambda Phages

The biology of lambda-derived vectors is extremely well understood. This introduction to the biology of λ and related phages is included to help readers use the λ vectors that are employed in the manual (e.g., UNIT 1.11 and Chapters 5 and 6), and to help readers understand new λ vectors that are being developed.

Lambda is a *temperate* phage, which means



Sketch 1.9A

that it can grow lytically or lysogenically. When λ infects a cell, it injects its DNA into the host bacterium. When it grows lytically, it makes many copies of its genome, packages them into new phage particles, and lyses the host cell to release the progeny phage. When λ grows lysogenically, it inserts its DNA into the chromosome of the host cell. The integrated phage DNA is replicated along with the host chromosome (but sometimes, e.g., when the cell's DNA is damaged, phage DNA is excised from the chromosome and the phage begins to grow

lytically). Lysogens are immune to superinfection by additional λ phages carrying a homologous immunity region. Plaques of temperate phages are turbid because they contain both cells lysed by phage that have grown lytically and cells spared from further phage infection because they have become lysogens (see sketch 1.9A).

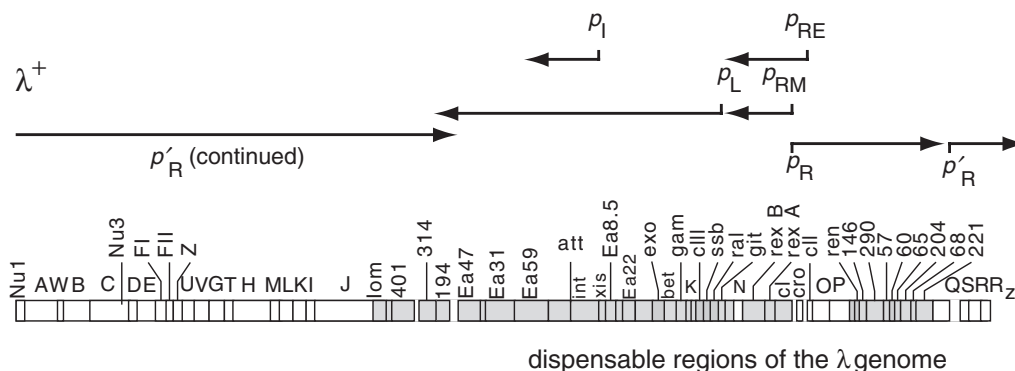
The λ genome is grouped into discrete blocks of related genes. This fact is quite convenient to the molecular biologist, as it has allowed construction of many phage-based cloning vectors which have deletions in large stretches of DNA nonessential to lytic growth (see gray areas in sketch 1.9B). In addition to these large areas, much of the immunity region, including the *cII*, *cro*, *cI* and *rex* genes, is inessential for lytic growth of otherwise wild-type phage; even the *N* gene is not essential if *t_{R2}* is deleted.

LYTIC GROWTH

Early Gene Expression

Lytic growth results in the production of progeny phage and the eventual lysis of the host cell. Lytic growth begins either on infection, if a phage does not grow lysogenically, or after the inactivation of existing *cI* repressor in a lysogen (see box E).

Transcription initiates from the early promoters, *P_R* and *P_L*. The *P_R* transcript terminates at *t_{R1}* and encodes the *cro* gene. The *P_L* tran-



Sketch 1.9B

Contributed by Karen Lech and Roger Brent
Current Protocols in Molecular Biology (1987) 1.9.1-1.9.4
 Copyright © 2000 by John Wiley & Sons, Inc.

script continues to t_{L1} and encodes the N gene. The product of the N gene is an antitermination factor; that is, it allows the transcripts which initiate at P_L or P_R to proceed through t_{L1} and t_{R1} respectively. The N -antiterminated P_R transcript, which terminates at t_{R2} , encodes the O , P , and Q genes. The O and P proteins are required for phage DNA replication; Q protein is another antitermination protein, in this case, specific for transcripts initiating at a promoter located to the right of Q called P_R' . Q -antiterminated transcription from P_R' proceeds through another terminator called t_{R65} and the late region, then through the head and tail genes which have been joined to the same transcription unit when the incoming phage DNA circularized, and finally terminates in b . The P_R' transcript encodes proteins necessary for head and tail assembly, DNA packaging, and host cell lysis.

This sequential expression of phage functions allows for the replication of the λ genome, its subsequent packaging into phage heads, and lysis of the host in the correct temporal order.

A. REQUIREMENTS FOR LYTIC GROWTH

N antitermination factor. Causes RNA polymerase to read through early terminators and expresses Q protein as well as the replication proteins.

Q antitermination factor. Causes RNA polymerase to read through t_{R65} and expresses the proteins required for head and tail assembly and host cell lysis.

Replication proteins. Phage O and P proteins work together with host proteins to replicate DNA.

Packaging and lysis proteins. Other proteins essential for phage growth.

DNA Replication

Replication during lytic growth requires both host proteins and phage-encoded proteins. Lytic DNA replication can be divided into an early and late phase. The early phase begins with the injection of linear phage DNA into the bacterial cell. This DNA has, at either end, complementary, single-stranded cohesive ends generated by cleavage of the cos site. These sticky ends base pair, and are ligated by host enzymes to yield a covalently closed circular molecule. After supercoiling by host topoisom-

erases, this molecule is able to initiate DNA replication. Replication begins at a unique site (called ori), and proceeds bidirectionally. This type of replication results first in the formation of a theta-shaped replication intermediate and later in the production of two daughter circles.

Approximately 15 min after phage infection, replication switches to the late phase. This is characterized by *rolling-circle replication*, which produces long polymers (called *concatemers*) of repeated, full-length phage genomes. The gam gene product protects the concatemers from degradation by an exonuclease encoded by the host $recB$ and $recC$ genes. Concatemers are substrates for packaging (see below and UNITS 5.8 & 5.9).

B. REQUIREMENTS FOR DNA REPLICATION

cos ends. 12-bp cohesive overhangs that pair and cause circularization in vivo. These are generated by cutting at cos sites, which occurs during packaging.

ori . Site at which DNA replication starts.

O and P proteins. O protein binds DNA; P protein interacts with host-encoded $dnaB$ protein.

gam protein. Inhibits *E. coli* exonuclease V ($recBC$ nuclease) and thus protects the end of the rolling circle concatemer from degradation by this enzyme.

Packaging and Lysis

Late in infection, DNA replication and packaging of λ occur concurrently. Once concatemeric DNA structures are formed, they are condensed into λ proheads (incomplete head particles). Cos sites are recognized by λ proteins, and the DNA between two cos sites is cleaved from the concatemer coincident with its packaging into a single prohead (linearizing the DNA and regenerating the cohesive ends). The remaining head proteins then assemble, and the tail, which has assembled independently, attaches to the head to form the intact phage. DNA located between two cos sites will only be packaged in a form that can be injected if it is between 38 and 53 kb long (a fact important in choosing the proper vector to construct a library; see UNITS 5.1 and 5.2). The last event during lytic growth is the lysis of the host. Phage-encoded proteins disrupt the bacterial inner membrane and degrade the cell wall.

C. REQUIREMENTS FOR PACKAGING AND LYSIS

cos sites. DNA is cleaved here, and DNA between the sites is packaged into phage heads.

NuI and A proteins. Components of the “terminase” that cleaves at *cos* sites.

B, C, D, E, and Nu3 proteins. Involved in the structure or assembly of the phage head.

G, H, I, J, K, L, M, T, U, V, and Z proteins. Involved in the structure or assembly of the tail and tail fibers.

R, S, and Rz proteins. *S* disrupts the inner membrane and the *R* and *Rz* proteins conspire to degrade the cell wall to achieve cellular lysis.

growth) with the joining of the cohesive ends and circularization of the genome. *cII* protein then binds to the P_{int} promoter and activates its transcription. *Int* protein is made from this transcript. *Int* protein catalyses a recombination event between the phage sequence, *attP*, and a site on the bacterial chromosome, *attB*, resulting in integration of the phage.

Immunity Regions

D. REQUIREMENTS FOR LYSOGENIZATION

imm region. Contains the promoters and genes (such as *cI*) that are essential for establishment and maintenance of lysogeny (see also section on immunity region).

cII and cIII proteins. *cII* is required for *int* synthesis; *cIII* protects *cII* from degradation by host proteases.

int protein. Along with host proteins, catalyzes the integration of the phage into the chromosome.

attP site. Required for integration of the phage into the host chromosome.

cos ends. Necessary for circularization of the molecule upon infection.

LYSOGENIC GROWTH

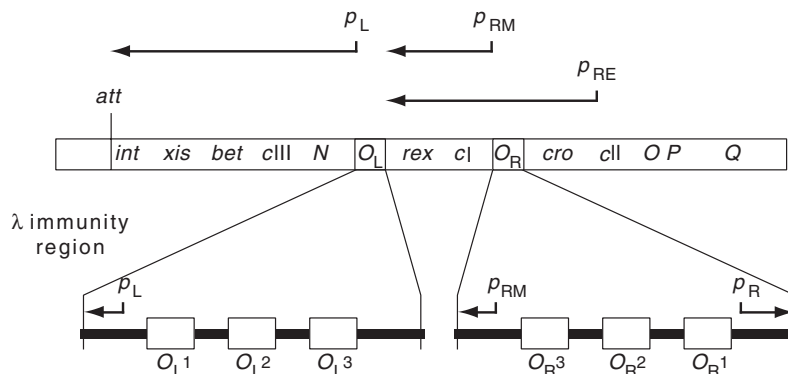
Gene Expression

Following infection, wild-type λ phage sometimes shuts down the majority of its genome and integrates into the bacterial chromosome; such a phage is called a *prophage* and the cell that contains it is called a *lysogen*. The product of the *cI* gene, λ repressor, is essential for lysogenic growth. This protein binds to the operators, O_L and O_R , blocking transcription of the early genes (from P_L and P_R) and preventing lytic growth. Binding to O_R also stimulates transcription from P_{RM} , a promoter that transcribes the *cI* gene. Lambda repressor thus maintains the lysogenic state by repressing the transcription of genes necessary for lytic growth as well as stimulating transcription of itself. There are mutant strains of *E. coli* called *hfl⁻*, used for the construction of λ gt10 libraries (UNIT 1.11 and Chapter 5), in which a wild-type phage almost always becomes a lysogen.

The actual sequence of events leading to integration begins (as it does during lytic

A number of bacteriophages—including 434, 21, 82, and 80—are related to phage λ , as evidenced by significant stretches of DNA homology. Unique to each phage, however, is its immunity (*imm*) region (see sketch 1.9C). This region includes a number of important regulatory sites and genes such as P_L and P_R , *cI* and *cro*, as well as O_R and O_L , and P_{RM} . A host cell stably lysogenized with a particular lambdoid phage is immune to infection by a second phage carrying the same *imm* region as the lysogen, because transcription from P_L and P_R of the incoming phage is repressed. A lyso-

Sketch 1.9C



gen can, however, be infected by a phage carrying a different *imm* region. This phenomenon occurs because the repressor encoded by a particular immunity region specifically recognizes and represses only its own promoters.

Induction

Lambda lysogens are *induced* when they excise from the host chromosome and undergo lytic growth. Excision requires the phage *int* and *xis* (excise) gene products. These are synthesized from the *N*-antiterminated P_L transcript. Initiation of P_L transcription occurs when the *cI* product (λ repressor) is inactivated. Induction occurs spontaneously at a low frequency in λ lysogens. However, it is more often due to (1) cleavage of λ repressor during the SOS response (see below), or (2) destabilization of a thermosensitive mutant repressor at nonpermissive temperatures.

In the laboratory, the bacterial *SOS response* is often induced by DNA damage caused by exposure of the cells to agents such as UV light or mitomycin C. During this response, LexA, a bacterial repressor protein, is inactivated and various bacterial genes are induced. Repressor proteins of λ and related bacteriophages resemble LexA and are also inactivated. Some clon-

ing vectors contain a temperature-sensitive mutation in the *cI* gene (*cI^{ts}*). As a result, the repressor is stable and behaves like the wild-type repressor at 30°C, but is unstable at 42°C. *cI^{ts}* mutations allow lytic growth to be induced simply by increasing the temperature at which the lysogen is grown.

E. REQUIREMENTS FOR INDUCTION OF A LYSOGEN

***int* and *xis* proteins.** Necessary for the excision of the phage DNA from the host chromosome.

Lytic growth requirements. Necessary for a productive burst of phage (see box A).

KEY REFERENCE

Hendrix, R., Roberts, J., Stahl, F., and Weisberg, R. 1983. Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Contributed by Karen Lech and Roger Brent
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

Lambda as a Cloning Vector

ADVANTAGES OF USING LAMBDA

About the middle third of the λ genome is dispensable for lytic growth. Derivatives of phage λ that are used as cloning vectors typically contain restriction sites that flank some or all of these dispensable genes. The major advantage to using lambda-derived cloning vectors is that DNA can be inserted and packaged into phages in vitro. Although the efficiency of packaging only approaches 10%, phages, once packaged, form plaques on *E. coli* with an efficiency of 1. While techniques for transformation of bacterial plasmids have improved greatly in recent years (see UNIT 1.8), the best routinely attainable frequencies are about 10^8 transformants/ μg pBR322, which means that less than 1 in 1000 plasmids become transformed into cells.

SELECTIONS FOR INSERTED DNA

Some phage vectors exploit the fact that λ has a minimum genome size requirement for packaging, so that vector phages that do not contain inserts above a minimum size are never detected. Other phage vectors utilize some genetic means to distinguish between phages that still retain the block of nonessential genes and recombinant phages in which the block of genes has been replaced with foreign DNA. Since the selection for the insert does not depend on the size of the inserted fragment, many of these vectors allow small fragments to be cloned.

Size selection. Lambda cannot be packaged into phage heads if its genome is less than 78% or more than 105% of the length of wild-type λ DNA. When a vector is chosen on the basis of size selection, the region dispensable for growth is cut out; the left and right arms of the phage are then purified and ligated to cut foreign DNA under conditions that favor concatemer formation. Those phages whose left and right arms have not been joined to an insert will have *cos* sites too close together to be packaged into viable phage. The existence of the size requirement for efficient phage packaging makes it impossible to use a single phage vector to clone fragments of all sizes: phages that can accommodate large inserts cannot be packaged if they contain small inserts, and vice versa. There are two special size selection tricks that have sometimes found use in phage cloning. In the first, packaged phage are treated with EDTA or other chelating agents. The population of surviving phage is enriched for phages with shorter than wild-type genomes. In the second, phage are plated on a mutant strain of *E. coli* called *pel⁻* which allows plaque formation by phages of

up to 110% of wild-type length and which severely inhibits plaque formation by phages of less than wild-type length.

***spi⁻* selection.** *red⁺ gam⁺* phages do not form plaques on a host lysogenic for the unrelated bacteriophage P2. These phages are said to be *spi⁺* (sensitive to P2 interference). *red⁻ gam⁻* phages do form plaques on P2 lysogens, and so are said to be *spi⁻* (Zissler et al., 1971). P2 inhibition of *red⁺ gam⁺* phage growth only occurs if the P2 is wild-type for a gene called *old*. Commonly used vectors like λ 2001 (Karn et al., 1984) and λ EMBL3 (Frischauf et al., 1983) contain a fragment with the *red⁺* and *gam⁺* genes. These vectors will not form plaques on a P2 lysogen unless the *red⁺ gam⁺* fragment is deleted and replaced with a piece of foreign DNA.

***hfl⁻* selection.** The product of the λ *cII* gene is necessary for infecting phages to synthesize repressor efficiently. The *E. coli hflA* and *hflB* genes encode products whose effect is to decrease the stability of the *cII* gene product (Banuett et al., 1986; Hoyt et al., 1982). When wild-type λ infects an *hfl⁻* strain, so much repressor is made that the phage almost always lysogenizes the infected cell, causing it to form either no plaque or an extremely turbid plaque. Vectors like λ gt10 contain a restriction site in the *cI* gene. Insertion of foreign DNA into this site inactivates the *cI* gene. The vector phage does not form plaques on the *hfl⁻* strain, but phages containing inserts instead of the *cI* gene form clear, normal-sized plaques.

MAPS OF LAMBDA-DERIVED CLONING VECTORS

Maps are presented for phage cloning vectors that are currently used to construct libraries, as well as maps for not-so-modern vectors which were employed to construct libraries that are still important.

The top of each page shows a simplified version of the map of wild-type lambda. The next lines show the changes that were made in the wild-type λ genome to generate that derivative. Deletions of lambda DNA are indicated by parentheses, and insertions of new DNA are indicated by bars. The bottom lines show transcripts, a genetic map, and a physical map of the resulting λ derivatives. The genetic nomenclature used, especially for deletions and changes in restriction sites, is extremely complicated. More explanation can be found in UNIT 1.9, in Lambda II by Hendrix et al. (1983), and in the articles cited in the figure legends.

Escherichia coli,
Plasmids, and
Bacteriophages

Contributed by Karen Lech, Roger Brent, and Nina Irwin

Current Protocols in Molecular Biology (1987) 1.10.1-1.10.11

Copyright © 2000 by John Wiley & Sons, Inc.

1.10.1

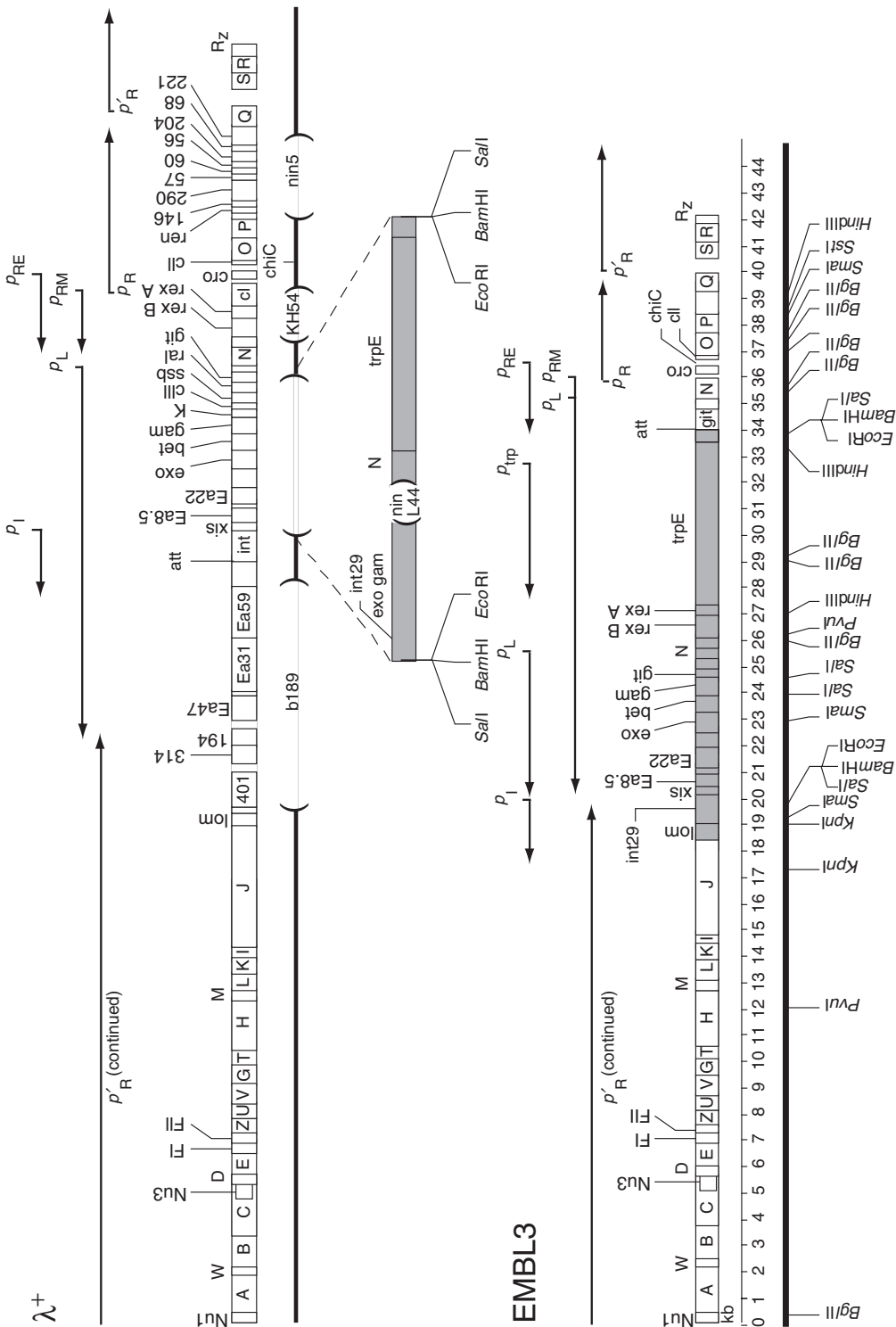


Figure 1.10.1 Wild-type lambda. Role of the major phage transcripts (arrows) and phage genes (boxes) in the lifecycle of bacteriophage λ is described in *UVITS 1.9 & 1.10*.

Figure 1.10.2 λ EMBL3. This phage vector is used for cloning large (10.4 to 20 kb) fragments. It contains a polylinker with cloning sites for *Bam*HI, *Sal*I, and *Eco*RI; in EMBL4, the polylinker is reversed. Recombinant phages containing DNA inserted into these sites become *cl'*, *gam'*, *red'*, and *int'*, and thus have the *Spi* phenotype. Although the phage is said to carry *cl857*, we believe that it does not carry a *cl* gene, and it has been drawn accordingly. The polylinker sequence is GGATCTGGTCCGACGGATCCGGGAATCCAGATCC. EMBL4's full genotype is λ sbh1 λ 1° b189 < polylinker (SalI-EcoRI) *int*29 *nin*L44 *cl857* *trpE* polylinker (EcoRI-SalI) > KH54 *chiC* srl λ 4° *nin*5 srl λ 5° (Frischauff et al., 1983).

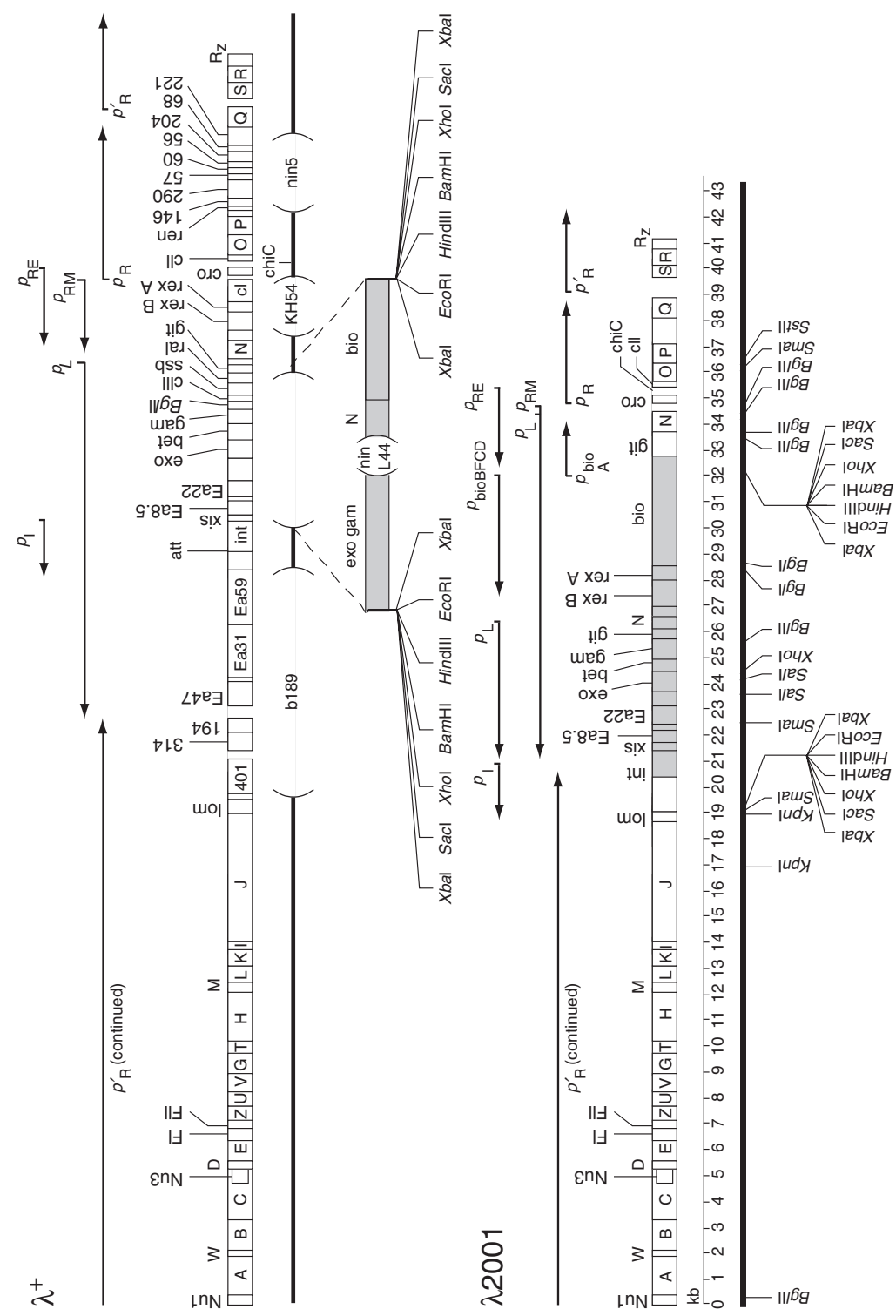


Figure 1.10.3 $\lambda 2001$. This phage is used for cloning large (10.4 to 20 kb) DNA fragments. Phages containing foreign DNA become cl^- , gam^- , int^- , and red^- . Since recombinants are red^- and gam^- , they have the Spi^- phenotype, and form plaques on a strain lysogenic for bacteriophage P2 (see *UNITS 1.9* and *1.10*). The vector contains cloning sites for *Bam*HI, *Eco*RI, *Hind*III, *Sac*I, *Xba*I, and *Xho*I. The polylinker sequence is **TCTAGAGCTCGAGGATCCCAAGCTTCAATTC**TAGA. $\lambda 2001$'s full genotype is λ sbhl $\lambda 1^\circ$ b189 Eint(linker) srl $\lambda 3^\circ$ ninL44 EshndIII $\lambda 4$ (bio) (linker) δ (sbhl) $\lambda 3$ -sbhl $\lambda 4$ KH54 srl $\lambda 4^\circ$ chiC nin5 srl $\lambda 5^\circ$ shndIII $\lambda 6^\circ$ (Kam et al., 1984).

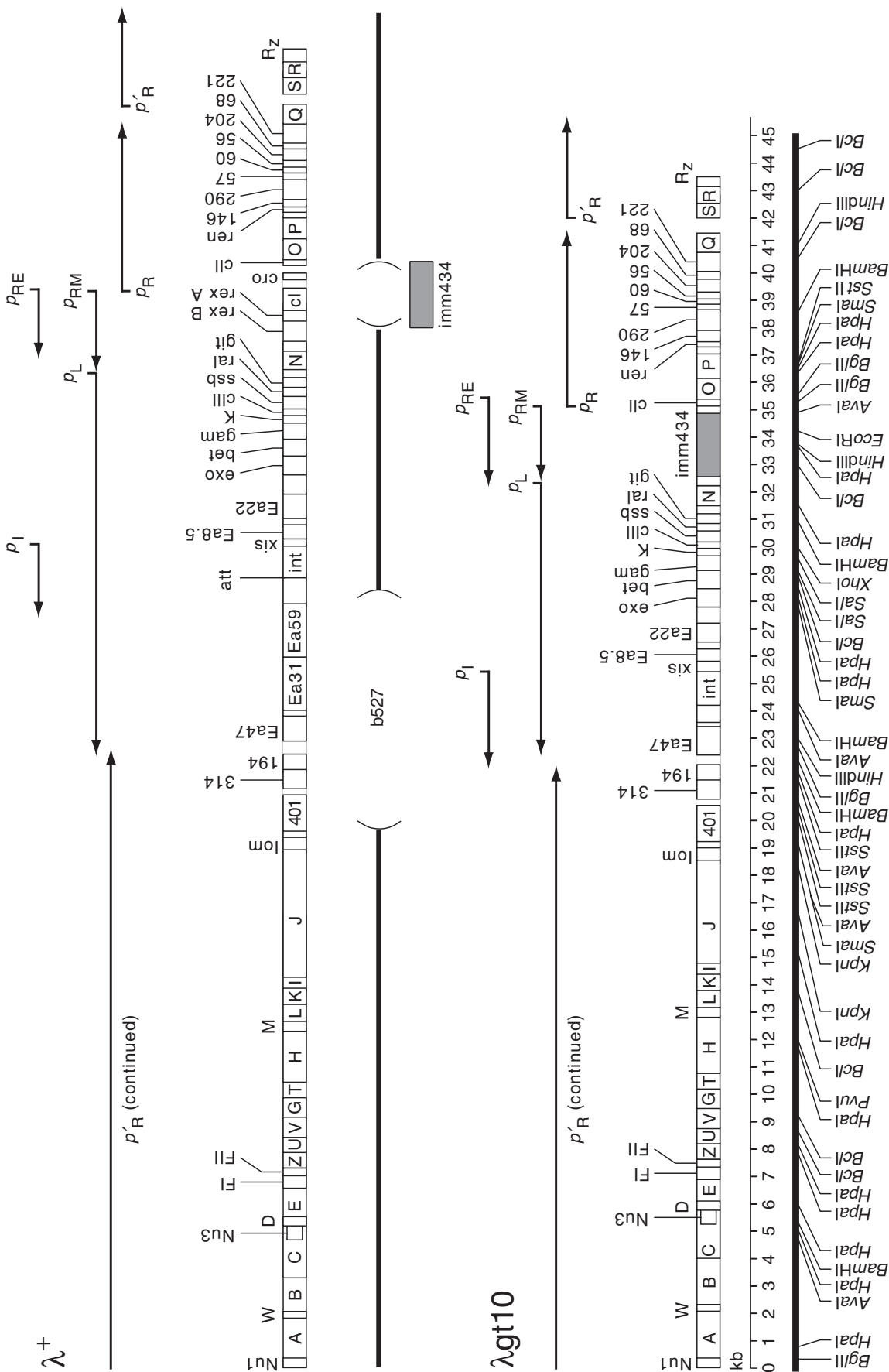


Figure 1.10.4 λ_{gt10} . This phage vector accepts small (0 to 5 kb) DNA fragments which are inserted into an RI site in its *ci* gene. Insertion of DNA into this site inactivates the *ci* gene and enables insert-containing phages to form plaques on a *hrfI* host. Since insert-containing phage form plaques very efficiently on this host, and since the plaques formed by insert-containing phage are usually very healthy, λ_{gt10} is very frequently used to construct libraries for which only very small amounts of DNA are available, for example in the construction of cDNA libraries. Insert-containing λ_{gt10} phage are *ci⁻* and *int⁺* but *red⁺*. Its full genotype is λ_{b527} *srI* λ_3^+ *imm⁺* *srI* λ_4^+ *srI* λ_5^+ (Huynh et al., 1984).

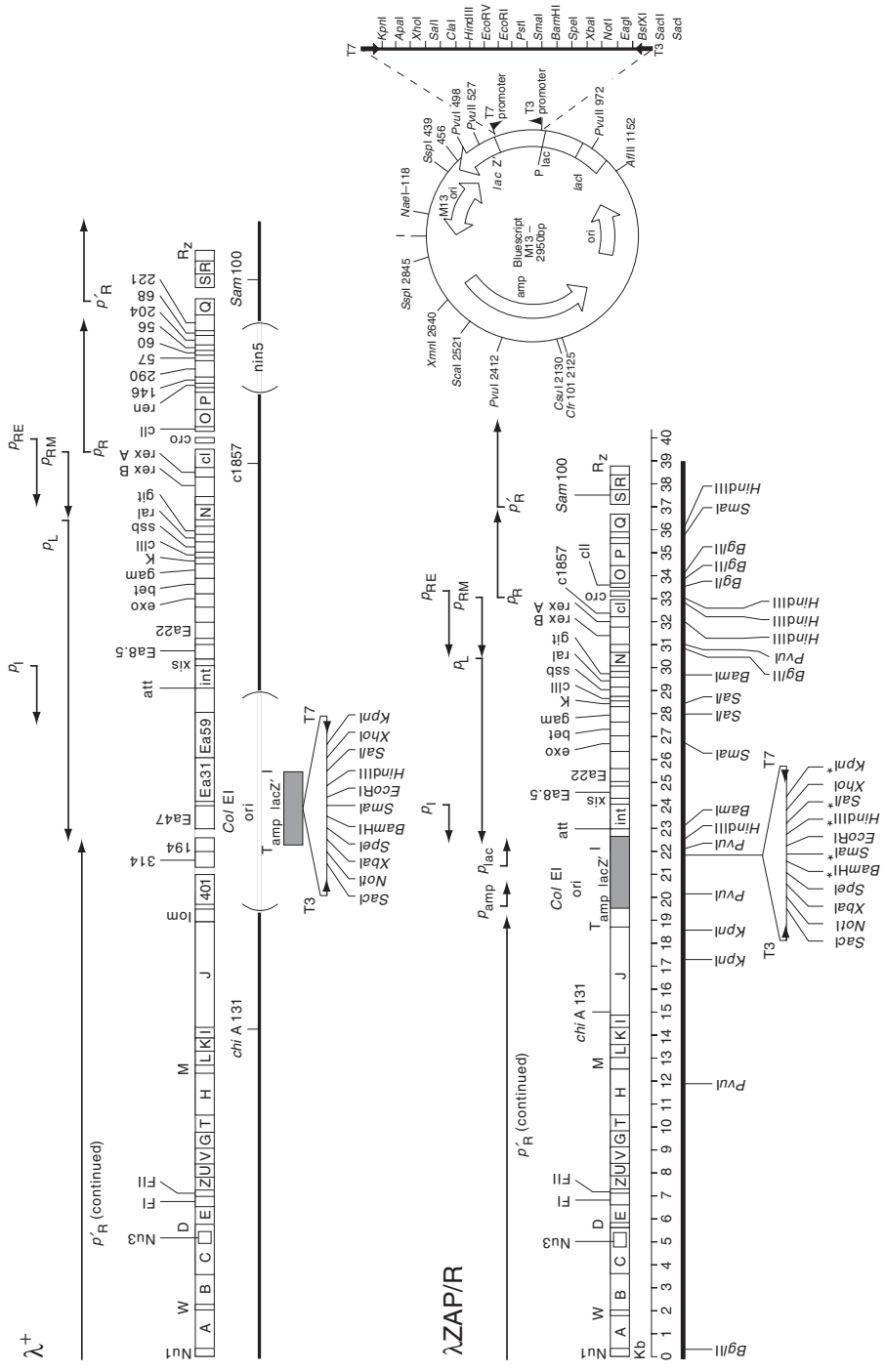


Figure 1.10.8 λ ZAP. λ ZAP carries pBluescript SK (-), which is excised in vivo upon infection with f1 or M13 helper phages. Inserts are cloned into λ ZAP within Bluescript sequences. Specifically, inserts are cloned within a polylinker located within *lacZ*. As with λ gt11, a fusion protein may be expressed if the insert DNA is in frame with the *lacZ* sequence; thus, libraries made in this vector can be screened with antibodies. In λ ZAP, T7 and T3 promoters flank the inserts, which allows RNA probes to be easily obtained. pBluescript M13 (-), the excised plasmid (shown on the right), is normally propagated as a double-stranded circular DNA, but infection with a helper phage enables the plasmid to be propagated as single-stranded DNA. These properties facilitate sequencing of the insert, site-directed mutagenesis, and the construction of unidirectional cloning sites are available. λ ZAP/R is shown in the figure. Within the polylinker, unique *XhoI*, *EcoRI*, *SpeI*, *XbaI*, *NofI*, and *SacI* λ ZAP vectors are available from Stratagene.

THE COSMID, A USEFUL LAMBDA-DERIVED PLASMID VECTOR

Cosmids were developed to allow cloning of large pieces of DNA. How they work is diagrammed in Figure 1.10.9. Cosmids contain a selectable marker, a plasmid origin of replication, a site into which DNA can be inserted, and a *cos* site from phage λ . The vector is cut with a restriction enzyme and mixed with pieces of DNA to be cloned. DNA ligase joins cut vector and insert fragments into concatemeric molecules. The ligation mixture is then mixed with a *packaging extract* (see UNIT 1.11) containing the proteins necessary to package

naked phage DNA into phage heads. Whenever two *cos* sites are present on a concatemer and separated by 40,000 to 50,000 nucleotides, they will be cut and packaged into phage heads. The cosmid-containing phages are infectious, they inject DNA into cells, but the DNA is plasmid DNA: the phage adsorbs to the host and the cosmid DNA is injected into the cell, which circularizes due to its sticky ends. The annealed ends are then covalently joined by the host's ligase, and the resulting large circular molecule replicates as a plasmid (see map, Fig. 1.10.10).

Cosmid vectors within cells replicate using

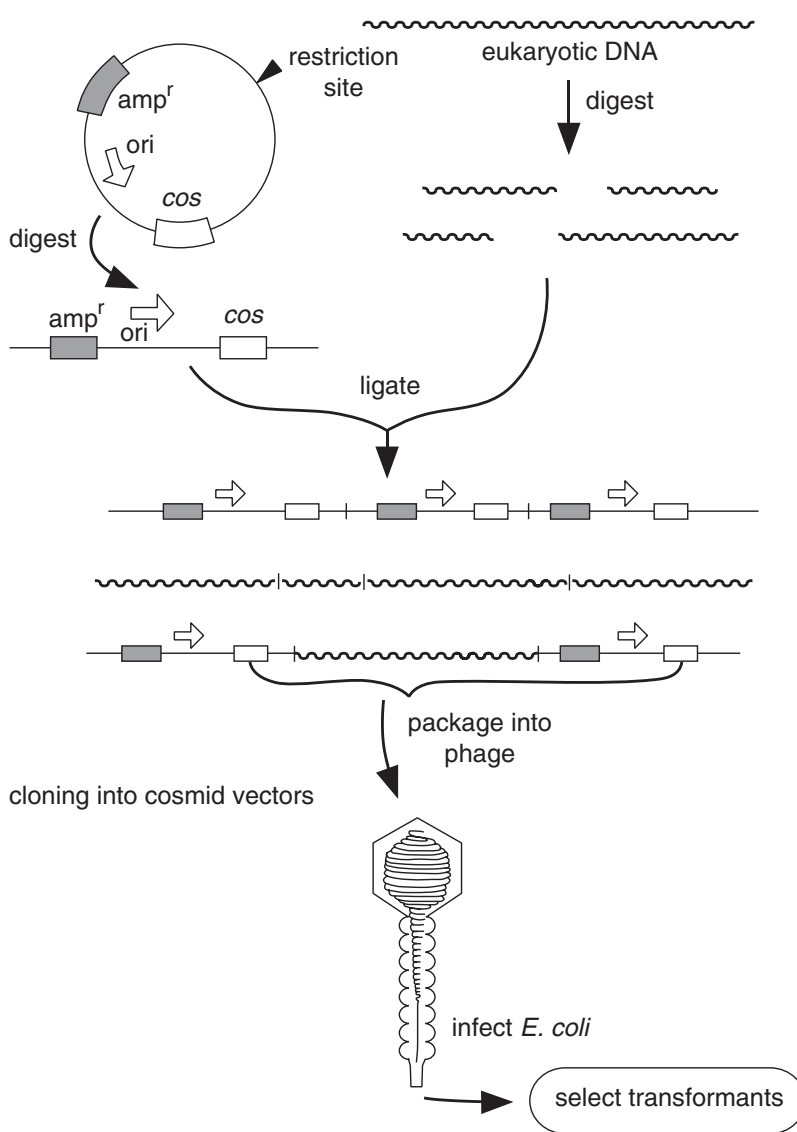


Figure 1.10.9 Cloning into cosmids.

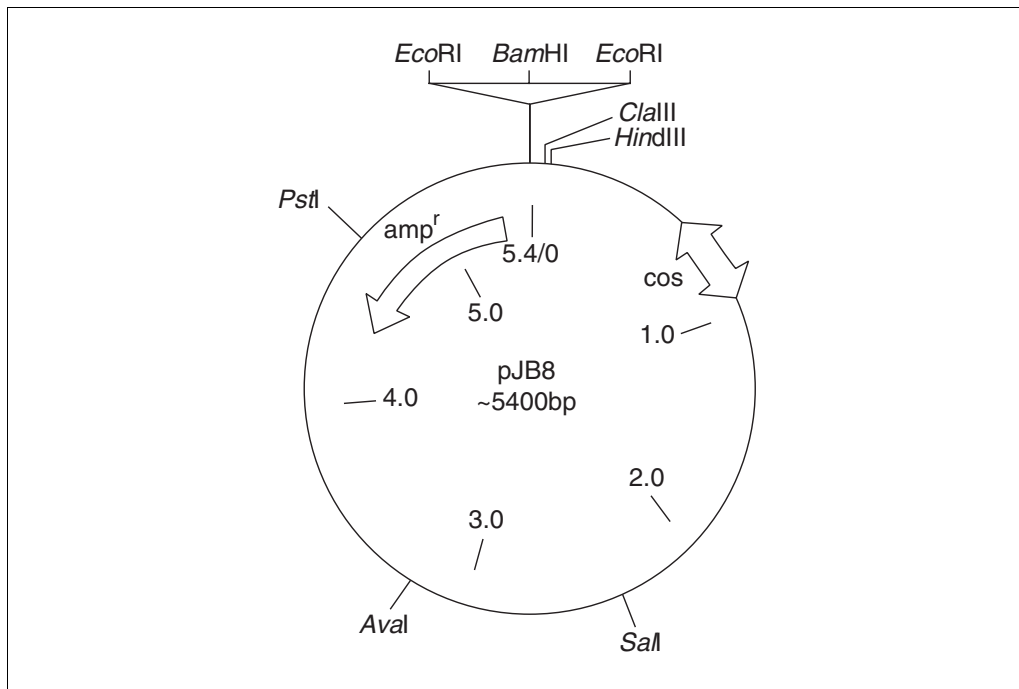


Figure 1.10.10 pJB8 is a commonly used cosmid vector. It contains a *cos* site, an ampicillin resistance gene, and a pMB1 replicator, so it can be amplified with chloramphenicol. In a typical application, libraries are constructed by insertion of random *Sau3a* partial-digestion fragments of DNA into its *Bam*HI site. The inserts can be excised by cleaving the plasmid at its flanking *Eco*RI sites (Ish-Horowitz and Burke, 1981).

their pBR322 origins. Intracellular cosmids sometimes rearrange DNA inserted into them, perhaps because the time it takes pBR322-dependent replication to replicate the 50,000-bp cosmid is almost as long as the generation time of *E. coli*, so that cosmids which have deleted sections of DNA have a growth advantage on antibiotic-containing medium. Although this problem can be ameliorated by propagating the cosmids in cells that are less prone to rearrange vector DNA (UNIT 1.4), another approach may well become more popular. Cosmid vectors (called *lorist vectors*) have recently been developed that are said to circumvent this problem by using the λ *ori* and *O* and *P* proteins to replicate (see, for example, Gibson et al., 1987). It takes only a few minutes for these vectors to replicate inside cells, and insertions in them are said to be more stable.

LITERATURE CITED

- Banuett, F., Hoyt, A.M., McFarlane, L., Echols, H., And Herskowitz, I. 1986. *hflB*, a new *Escherichia coli* locus regulating lysogeny and the level of bacteriophage lambda cII protein. *J. Mol. Biol.* 187:213-224.
- Blattner, F.R., Williams, B.G., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Furlong, L.-A., Grunwald, D.J., Kiefer, D.O., Moore, D.D., Schumm, J.W., Sheldon, E.O., and Smithies, O. 1977. Charon phages: Safer derivatives of bacteriophage lambda for DNA cloning. *Science* 196:161-169.
- deWet, J.R., Daniels, D.L., Schroeder, J.L., Williams, B.G., Denniston-Thompson, K., Moore, D.D., and Blattner, F.R. 1980. Restriction maps for twenty-one Charon vector phages. *J. Virol.* 33:401-410.
- Dunn, I.S. and Blattner, F.R. 1987. Charon 36 and 40: Multi-enzyme, high capacity, recombination deficient replacement vectors with polylinkers and polystuffers. *Nucl. Acids Res.* 15:2677-2698.
- Frischauf, A.-M., Lehrach, H., Polstka, A., and Murray, N.M. 1983. Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* 170:827-842.
- Gibson, T.J., Coulson, A.R., Sulston, J.E., and Little, P.F.R. 1987. Lorist 2, a cosmid with transcriptional terminators insulating vector genes from interference by promoters within the insert: Effect on DNA yield and cloned insert frequency. *Gene* 53:275-281.
- Hendrix, R., Roberts, J., Stahl, F., and Weisberg, R. 1983. Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

- Hoyt, M.A., Knight, D.M., Das, A., Miller, H.I., and Echols, H. 1982. Control of phage lambda development by stability and synthesis of *cII* protein: Role of the viral *cIII* and host *hflA*, *himA*, and *himD* genes. *Cell* 31:565-573.
- Huynh, T.V., Young, R.A., and Davis, R.W. 1984. Constructing and screening cDNA libraries in λ gt10 and λ gt11. *In* DNA Cloning Techniques: A Practical Approach (D. Glover, ed.) pp. 49-78. IRL Press, Oxford.
- Ish-Horowitz, D. and Burke, J.F. 1981. Rapid and efficient cosmid cloning. *Nucl. Acids Res.* 9:2989-2998.
- Karn, J., Matthes, H.W.P., Gait, M.T., and Brenner, S. 1984. A new selection cloning vector, λ 2001, with sites for *Xba*I, *Bam*HI, *Hind*III, *Eco*RI, *Sst*I, and *Xho*I. *Gene* 32:217-224.
- Williams, B.G. and Blattner, F.R. 1979. Construction and characterization of the hybrid bacteriophage lambda Charon vectors for DNA cloning. *J. Virol.* 29:555-575.
- Young, R.A. and Davis, R.W. 1983. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. U.S.A.* 80:1194-1198.

- Zissler, J., Signer, E., and Schaefer, F. 1971. The role of recombination in growth of bacteriophage lambda. *In* The Bacteriophage Lambda (A.D. Hershey, ed.) pp. 455-475. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

KEY REFERENCE

- Murray, N.E. 1983. Phage Lambda and Molecular Cloning. *In* Lambda II, (R.W. Hendrix, J. Roberts, F. Stahl, and R. Weisberg, eds.) pp. 395-432. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Provides a thorough introduction to the biology of cloning vectors in common use before 1983.

Contributed by Karen Lech and Roger Brent
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

Nina Irwin
Harvard University
Cambridge, Massachusetts

Plating Lambda Phage to Generate Plaques

ISOLATING A SINGLE PLAQUE BY TITERING SERIAL DILUTIONS

This procedure is used to isolate pure populations of phage from a single plaque and provide the titer of the phage stock. Serial dilutions are made of a phage lysate. In separate tubes, aliquots of each dilution are mixed with *E. coli*. Phage are allowed to adsorb to the cells and the cell/phage mixture is then heated to 37°C, causing the phage to inject their DNA into the cells. Top agar is added to each tube, and the mixture is poured onto rich plates, which are incubated at 37°C until plaques appear. Each plaque contains phages derived from a single infecting phage.

Materials

Lambda broth (UNIT 1.1)
 0.2% maltose
 10 mM MgSO₄
 Lambda top agar (UNIT 1.1)
 Suspension medium (SM)
 Fresh lambda plates (UNIT 1.1), prewarmed to 37°C

Microwave oven or boiling water bath
 45° to 50°C water bath
 8 × 80–mm tubes
 37°C water bath or heat block
 Capillary tubes or toothpicks

NOTE: All materials coming into contact with *E. coli* must be sterile.

1. Grow a culture of *E. coli* to saturation in lambda broth + 0.2% maltose + 10 mM MgSO₄.

Growth in maltose induces production in E. coli of the λ receptor (lamB protein), which is necessary for maltose transport. Mg⁺⁺ ions also aid phage adsorption.

2. Melt top agar in microwave oven set to defrost setting, or in boiling water bath for 15 min. After agar is melted, let bottle cool at room temperature for 5 min, then place bottle of melted agar in 45° to 50°C water bath.

The cap to the bottle of top agar must be loose before putting it in microwave! Watch the microwave oven to make sure that the contents of the bottle of top agar do not boil over. If agar boils over, carefully take the bottle out and swirl it around to see if there are any unmelted flecks of agar. If there are, reinsert into oven and microwave longer, inspecting occasionally, until the agar is completely melted.

Be sure top agar is left in the water bath enough time to cool to 45° to 50°C. Cells will be killed by even brief exposure to agar that is hotter than 65°C.

3. Add 0.3 ml of the *E. coli* culture to five 8 × 80–mm tubes.
4. Make serial dilutions of the phage lysate in SM (see UNIT 1.3 for serial dilutions).

Dilution factors of 100-fold are usually used. Label the dilution tubes so that they do not get mixed up.

5. Add 0.1 ml of the first dilution to one tube of *E. coli*, 0.1 ml of the next dilution to the next tube, etc. Label the tubes of *E. coli*/phage mixture, so that they do not get mixed up. Incubate tubes at room temperature for 20 min.

The phage adsorb to the E. coli during this step.

6. Move tubes to a 37°C water bath or heat block for 10 min. While tubes are in

water bath, label 5 fresh, prewarmed lambda plates to correspond with the labels on the dilution tubes.

During this step, the phage inject their DNA into the cells.

Plates should be fresh, but not so fresh that they are wet on their surfaces, nor so wet that they will exude moisture and cause the lawn of cells in top agar to slide away.

7. Remove tubes from the water bath. Add 2.5 ml top agar to one tube, vortex lightly to mix, and pour the contents of tube onto a plate. Spread agar over the entire surface of the plate by tilting it gently.
8. Place the plates in a 37°C incubator. Plaques of lambda-derived phages will appear after 6 to 8 hr, but will be easier to score, count, and pick if left for 12 hr.
9. From one of the dilution plates that is not too crowded with plaques, pick a single plaque with a sterile capillary tube or toothpick. To save the plaque for future use (for example, to make a plate stock), cut out a plug of agar containing the plaque with a capillary tube, and blow the plug into a tube containing 1 ml of SM (or place the tip of the toothpick in the liquid and agitate gently). If desired, count the number of plaques on one of the dilution plates and use this number to compute the number of viable phage in the starting stock.

Background Information

Titration of bacteriophage, and isolation of phage from single plaques, was first described by d'Herelle in 1920. The best general background to bacteriophage growth protocols is probably found in Stent (1971).

Genes encoding the tail proteins of most lambda-derived cloning vectors come from phage λ . Vectors with these tail proteins, said to be “h λ ” (for host range of λ), adsorb to the cell *lamB* protein, which is involved in maltose uptake. These vectors make plaques with sharp boundaries. Some vectors have tail proteins derived from phage 80 and are said to be “h80.” These vectors adsorb to the host *tonA* protein, which is involved in ion transport. They make slightly larger plaques with fuzzier borders than those made by h λ phages. Stocks of h80 phages usually have a higher titer than stocks of corresponding h λ phages.

Troubleshooting

It is sometimes helpful to include two other control plates in the procedure. One is a plate that contains a lawn made from a separate tube of the *E. coli* culture that has not been infected with phage. The other is a plate that contains top agar that did not contain any cells. These two controls provide benchmarks for growth of the lawn, as it becomes denser during the time in the incubator. If the lawn appears crinkled, then the top agar/*E. coli*/phage mixture was probably too cold by the time it was poured onto the plate. If this occurs, try warming the plates up to 37°C before pouring the lawn onto them, or pouring the lawns more quickly after the top agar is added to the *E. coli*/phage mixture. If the top agar layer floats off the plate, then the plates were too wet. If this occurs, use dryer plates.

Time Considerations

It takes from 6 to 8 hr of incubation at 37°C before plaques appear, and is often 12 to 14 hr before differences in morphology can be reliably distinguished.

ISOLATING SINGLE PLAQUES BY STREAKING ON A LAWN OF CELLS

Phages are streaked for single plaques on a plate containing a pre-poured lawn of *E. coli*. This procedure is easier than titering by serial dilutions, and is recommended if only a few isolated plaques are needed.

Materials

- LB medium (UNIT 1.1)
- Lambda top agar (UNIT 1.1)
- Rich plate (UNIT 1.1), prewarmed to 37°C
- 32-G platinum wire loop or sterile 1½ × ⅛ in. strips of paper

1. Grow a lambda-sensitive strain of *E. coli* to saturation in 5 ml LB medium.
2. Add 0.2 ml of the saturated culture to 2 ml of melted top agar (cooled to about 45°C; see step 2, p. 1.11.1), and pour evenly over the top of a prewarmed, rich plate.

In the recipe for top agar (UNIT 1.1), agar can be replaced with 6 g agarose.

3. After top agar has hardened, cool plates by placing in refrigerator ≥15 min.
4. Spot 100 µl of λ stock culture (usually around 10⁸ phage/ml) on the corner of the plate.
5. Using the techniques described in UNIT 1.3 to streak out single bacterial colonies, *lightly* streak out the phage using a thin wire loop or the edge of a sterile 1½ × ⅛ in. piece of paper.

Paper should be cut into strips and sterilized by autoclaving (dry) in screw-cap vials.

PHAGE TRANSFECTION AND IN VITRO PACKAGING

Construction of a library with lambda-derived cloning vectors results in a population of phage DNA molecules. In order for these molecules to be replicated, they must be introduced into cells so that they can grow as phage. Phage DNA is typically introduced into cells either by infection after packaging into phage particles *in vitro* or, much less frequently, by *transfection*, that is, transformation of phage DNA that has been circularized by treatment with DNA ligase into competent cells.

In vitro packaging uses lysates of phage-infected *E. coli* called *packaging extracts*. These lysates contain empty phage heads, unattached phage tails, and the phage-encoded proteins required for DNA packaging (see UNIT 1.9). If ATP is present, then concatemeric phage DNA mixed with the extract is cut at one *cos* site by the terminase (probably a complex of the *A* and *Nul* proteins) and loaded into the phage head by an unknown mechanism. DNA continues to be loaded into the phage head until the terminase encounters and cuts the next *cos* site on the molecule. The phage tails then attach themselves to the filled heads.

In the early 1980s, production of high-efficiency packaging extracts was considered one of the biggest inconveniences in library construction. To avoid the work involved, most researchers now use frozen packaging extracts purchased from commercial suppliers (see Enquist and Sternberg, 1979). Commercial extracts typically yield 2 × 10⁸ to 2 × 10⁹ plaque-forming particles per µg of concatemeric phage DNA. Phage DNA can also be introduced into *E. coli* by transformation of competent cells. In this procedure, linear or circular phage DNA is mixed with competent *E. coli*, and the mixture of cells and DNA is treated as in UNIT 1.8. After the heat shock step, the calcium-treated cells are mixed with *E. coli* that have been grown in lambda broth and maltose as described in first basic protocol. Molten top agar is then added, the mixture is poured

onto a lambda plate, allowed to solidify, and incubated at 37°C until plaques appear. This procedure typically yields 10⁴ plaques/μg of phage DNA.

REAGENTS AND SOLUTIONS

Suspension medium (SM), per liter

5.8 g NaCl

2 g MgSO₄·7H₂O

50 ml 1 M Tris·Cl, pH 7.5

0.01% gelatin (Difco)

LITERATURE CITED

Enquist, L. and Sternberg, N. 1979. Packaging of bacteriophage λ in vitro. *Meth. Enzymol.* 68:281-298.

Stent, G.S. 1971. *Molecular Genetics: An Introductory Narrative*, W.H. Freeman, NY.

Contributed by Karen Lech and Roger Brent
Massachusetts General Hospital and Harvard Medical School
Boston, Massachusetts

Growing Lambda-Derived Vectors

It is often necessary to grow large quantities of lambda-derived vectors, so that DNA can be made from them (see UNIT 1.14). The following basic protocol tells how to make a phage stock by plate lysis, while the alternate protocol tells how to make a liquid lysate. Storage of phage lysates is described in the support protocol.

BASIC PROTOCOL

MAKING A STOCK OF PHAGE BY PLATE LYSIS

Phage from a single plaque are mixed with cells and top agar and poured onto a plate. Although there are initially more cells than phage in the lawn, the phage increase their number more rapidly than the cells, and eventually lyse most cells on the plate. The top agar is then scraped off the plate and the phage in it are extracted and saved.

Materials

Fresh lambda top agar (UNIT 1.1)
 Fresh lambda plates (UNIT 1.1), prewarmed to 37°C
 Suspension medium (SM; UNIT 1.11)
 Chloroform
 45° to 50°C water bath
 Capillary tube or toothpick
 Beckman JA-21 rotor or equivalent

NOTE: All materials coming into contact with *E. coli* must be sterile.

1. Dilute a fresh overnight culture of a lambda-sensitive strain of *E. coli* 50-fold and grow in culture tube on a roller drum at 37°C.
2. Melt 100-ml bottle of fresh lambda top agar by heating in microwave oven set to defrost. Place in 45° to 50°C water bath.
3. When cell density reaches 2 to 3 × 10⁸/ml (OD₆₀₀ = 0.4), place 0.75 to 1 ml of cells into a test tube. Pick a single fresh plaque with capillary tube or toothpick, blow the plug into cells, and vortex lightly for 10 sec.
4. Add 7.5 ml top agar, vortex gently, and pour equal amounts of the mixture onto three prewarmed fresh lambda plates.

If agar in the top agar and plate recipes is replaced by agarose, the extractable DNA will be an acceptable substrate for restriction enzymes and ligases.

5. Incubate plates at 37°C, typically for 4 to 6 hr, until plaques are clearly visible and when 90% to 100% of the lawn is lysed.
6. Pipet 3 ml SM onto plate. Using clean microscope slide, scrape top agar from all 3 plates into a small centrifuge tube (e.g., a 15-ml screw-top tube for the Beckman JA-21 rotor).
7. Add 3 drops chloroform. Vortex vigorously for 10 sec.
8. Leave at room temperature 10 min.

As an alternative to steps 6 through 8, 3 ml SM and 3 drops chloroform can be dribbled onto the plate, the plate left overnight at 4°C, and the liquid decanted and centrifuged as below.

9. Centrifuge 10 min at 10,000 rpm in JA-21 rotor (11,400 × g), 4°C.
10. Gently decant and save supernatant.

MAKING A LIQUID LYSATE

Host bacteria grown to saturation are infected with 10^5 to 10^8 phage/ml. Following phage adsorption, the infection mixture is diluted into a rich medium and shaken vigorously until cell lysis. Any remaining viable cells are lysed with chloroform, and cell debris is removed with a low-speed spin.

Additional Materials

LB medium (UNIT 1.1)
Lambda dilution buffer
10 mM $MgCl_2$ /10 mM $CaCl_2$
NZC medium (UNIT 1.1)
Beckman JA-20 rotor or equivalent

1. Grow an overnight culture of a lambda-sensitive strain of *E. coli* in LB medium at 37°C.

Lambda-sensitive strains will support lytic growth. This can be tested by spotting 10 μ l of a lambda lysate onto a lawn of bacteria. If the strain is lambda-sensitive, a plaque will form where the phage were spotted.

2. Using a sterile toothpick or capillary tube, pick a single plaque (see step 3 of basic protocol), blow it into a tube that contains 0.4 ml lambda dilution buffer, and place tube at 4°C for 2 hr to allow phage to elute.

Alternatively, 10^5 to 10^8 phage from a liquid lysate or plate stock can be used.

3. Combine 0.1 ml eluted phage with 0.1 ml of saturated culture and 0.1 ml of 10 mM $MgCl_2$ /10 mM $CaCl_2$ solution and incubate 15 min in a 37°C water bath.

Incubating with Mg^{++} and Ca^{++} allows the phage to adsorb to the bacteria.

4. Transfer this solution to 50 ml of NZC medium and shake *vigorously* at 37°C until lysis occurs (usually between 6 and 8 hr).

Good aeration is important for high yields.

5. The culture should be checked frequently after 6 hr, and harvested immediately upon clearing.

6. Add a few drops of chloroform to lyse any remaining cells, transfer the solution to Co-rex or Nalgene tubes (being careful to leave the chloroform behind), and spin 10 min at 10,000 rpm ($12,100 \times g$), 4°C, to pellet the cell debris.

7. Save as much of the lysate as desired. Transfer to a screw-cap tube, add a few drops of chloroform, vortex briefly, and store at 4°C.

The titer of the phage should be determined as described in UNIT 1.11.

STORING PHAGE LYSATES

Phage stocks should be stored in SM (UNIT 1.11) plus a few drops of chloroform at 4°C. Screw-cap glass tubes with rubber-lined or teflon caps, capable of holding at least 1 ml, are generally used, although disposable plastic tubes work well. Lambda titers drop over a period of several years. Stocks should be dated and checked occasionally to determine phage titer. Addition of 0.1% gelatin to the SM slows down the rate at which phage titers drop when stored under these conditions. Lysates can also be stored frozen in 15% glycerol.

REAGENTS AND SOLUTIONS

Lambda dilution buffer

20 mM Tris·Cl, pH 8
20 mM $MgCl_2$

ALTERNATE PROTOCOL

SUPPORT PROTOCOL

*Escherichia coli,
Plasmids, and
Bacteriophages*

1.12.2

COMMENTARY

Background Information

Plate lysis is typically used to make small phage stocks, which are then used as intermediates for making larger stocks for preparation of phage DNA. However, if agar is replaced by agarose in the plate recipe and in the top agar (see *UNIT 1.1*), then DNA extracted from phage in the lysate is an acceptable substrate for restriction enzymes and ligases.

Liquid lysates generally give slightly lower yields than plate lysates. Less work is required, however, especially when large quantities of phage are needed. This fact, taken together with the fact that there is no contaminating agar in the lysate, makes this method preferable for isolating phage for DNA preps. In the liquid lysis procedure, cells are diluted into NZC medium after phage adsorption. This is a rich medium that contains amino acids but no glucose or other sugars. In cells grown in this medium, sugar receptors like LamB protein are fractionally induced (so phage can adsorb to the cell), but cells are not covered with so many receptors that adsorption to their debris depletes the yield of phage from the lysate. In addition, cells grown on this medium seem to cause infecting phage to grow lytically, which helps to obtain high titers of turbid phages.

Critical Parameters

LB medium and top agar can replace lambda plates and top agar for the plate lysis method, but the phage yield is usually lower. It is important that the plates be freshly poured, and that the top agar be freshly melted. The plates do not even have to be dry; a loose lawn is no problem as long as the plates are not tilted. If the yield is less than expected, it may be because the plates were not fresh, or because the particular phage used does not give a burst as large as that given by wild-type lambda. The size of single plaques made by the phage sometimes gives a clue that the second problem obtains. If plaques made by the phage in question are smaller than those made by closely related phages, it is often worthwhile to reduce the number of starting cells, or to increase the starting number of phage by combining several plaques, and to incubate the plates longer until

the lawn is completely lysed.

In the liquid lysis method, good aeration is essential for high yields. Flasks should be no more than $\frac{1}{5}$ full. Cultures should be watched carefully as they approach lysis and harvested soon after lysis is complete to prevent released phage from reinfecting cell debris. Lysis usually takes between 6 and 8 hr. If there is no sign of lysis after 8 hr, there probably will not be any, and one can either go on to add the chloroform and complete the procedure, or abandon hope and start over.

Anticipated Results

The plate lysis procedure yields ~10 ml of a phage stock containing 10^{10} to 10^{11} phage/ml. The liquid lysis procedure usually yields a stock with 5×10^9 to 3×10^{10} phage/ml.

Time Considerations

The plate lysate procedure usually takes about 8 hr. The liquid lysate procedure takes 2 hr to elute the phage from the plaque, and 6 to 9 hr between the time the cells are infected and the time the lysate is harvested. The time required to elute the phage is very flexible. Two hours will ensure that more than 90% of the phages are eluted from the plaque; shorter elution times allow less phage to be eluted. Elution times as short as 30 min generally work but do not always give good lysis. Phage may be eluted for longer periods of time; overnight is often convenient. Once the host cells are infected, it takes ~6 to 9 hr for lysis and preparation of the lysate.

Key Reference

Blattner, F.R., Williams, B.G., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Furlong, L.-A., Grunwald, D.J., Kiefer, D.O., Moore, D.D., Schumm, J.W., Sheldon, E.O., and Smithies, O. 1977. Charon phages: Safer derivatives of bacteriophage lambda for DNA cloning. *Science* 196:161-169.

Origin of this liquid lysate procedure.

Contributed by Karen Lech
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

Preparing Lambda DNA from Phage Lysates

UNIT 1.13

DNA extracted from lambda-derived vectors is typically subcloned into plasmid or filamentous phage vectors. The first two protocols describe methods for isolating phage DNA from large- and medium-scale liquid lysates. These two methods use either density-gradient centrifugation or ion-exchange chromatography to purify the phage particles. The third protocol describes a rapid procedure for isolating phage DNA, suitable for small-scale liquid lysates.

PREPARING DNA BY STEP- AND EQUILIBRIUM-GRADIENT CENTRIFUGATION

BASIC
PROTOCOL

A scaled-up liquid lysate is used to prepare a large quantity of highly purified phage DNA. Phage is separated from cellular debris by a CsCl step gradient followed by equilibrium-gradient centrifugation. Two alternate sets of steps are provided for extracting λ DNA from CsCl-purified phage particles. In the first method, CsCl is removed by dialysis and the DNA is extracted by phenol and chloroform. In the second approach, phage DNA is extracted directly from CsCl-purified phage particles with high-grade formamide.

Materials

- 5 \times polyethylene glycol (PEG) solution
- Suspension medium (SM; UNIT 1.11)
- Potassium chloride
- CsCl solutions
- Low-salt buffer
- Buffered phenol (UNIT 2.1)
- Chloroform
- 2 M Tris·Cl (pH 8.5)/0.2 M EDTA (optional, for formamide extraction)
- Formamide (*very* high grade, preferably recrystallized; optional)
- TE buffer, pH 8.0 (APPENDIX 2)
- Beckman JA-10, JA-20, SW-28, and VTi50 rotors and bottles/tubes (or equivalents)
- 3-ml syringe with 25-G needle
- Beckman VTi50 quick-seal tubes
- Additional materials for preparing liquid phage lysate (UNIT 1.12), titering lambda phage (UNIT 1.11), and quantitation of DNA (APPENDIX 3)

Prepare and concentrate the phage

1. Use 25 ml liquid lysate to make a 1000-ml lysate (see UNIT 1.12).
2. Split the lysate into two JA-10 centrifuge bottles and spin 10 min at 10,000 rpm (17,700 \times g), 4°C, to remove cell debris.
3. Transfer supernatant to a 1000-ml graduated cylinder and add 5 \times PEG solution to a final concentration of 1 \times . Invert gently to mix. Let sit overnight at 4°C.

The PEG solution causes the phage to precipitate.

4. Remove ~50 ml of supernatant and save for step 5. Pour off remaining supernatant, being careful not to lose any of the white precipitate.
5. Transfer precipitate to Nalgene centrifuge tube. Rinse cylinder with saved supernatant and transfer to centrifuge tube. Spin 10 min in JA-20 rotor at 5000 rpm (3000 \times g), 4°C.

Escherichia coli,
Plasmids, and
Bacteriophages

Contributed by Karen Lech

Current Protocols in Molecular Biology (1990) 1.13.1-1.13.10

Copyright © 2000 by John Wiley & Sons, Inc.

1.13.1

Supplement 10

- Place centrifuge tubes on ice. Remove the top layer, being careful not to remove any of the thick white phase, which contains the PEG solution and the phage.
- Resuspend the white phase in a minimum volume of SM. Transfer to a 125-ml flask.
The volume of the suspension medium added should not be more than three times the volume of the white phase.

- Determine the amount of solid KCl needed to make a 1 M solution. Add this amount of KCl in four aliquots of approximately equal size, mixing well after each addition. Let sit on ice for 15 to 30 min.
Adding KCl precipitates the PEG solution slowly while leaving the phage behind.

- Transfer to Nalgene centrifuge tube and spin 10 min in a JA-20 rotor at 10,000 rpm ($12,100 \times g$), 4°C .
The PEG solution will be pelleted and the phage will remain in the supernatant.

- Measure the phage titer and keep the supernatant in a 16- or 18-mm glass test tube.
The phage titer should be $\sim 1 \times 10^{12}$ to 5×10^{13} pfu/ml.

Isolate the phage particles

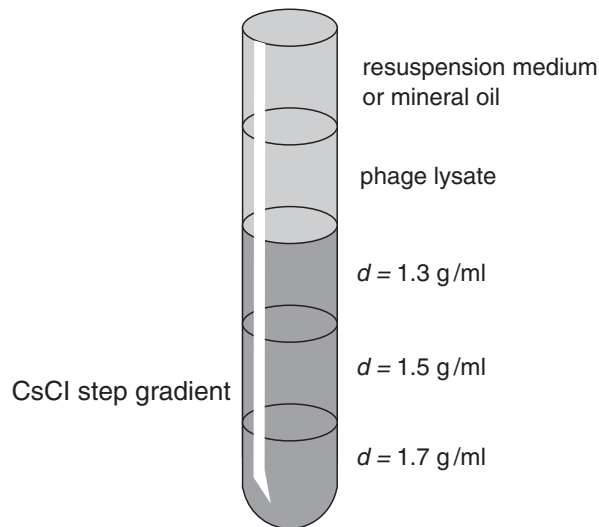
- Pour a CsCl step gradient in an SW-28 centrifuge tube as follows:

First layer: 3.5 ml CsCl solution, $d = 1.7 \text{ g/ml}$
 Second layer: 2.5 ml CsCl solution, $d = 1.5 \text{ g/ml}$
 Third layer: 2.5 ml CsCl solution, $d = 1.3 \text{ g/ml}$

The layers must be added very slowly to avoid mixing.

- Carefully layer the lambda lysate (supernatant from step 10) on top of this gradient (see sketch 1.13A).
- Fill tube to just below the top with SM. Spin 2 hr in an SW-28 rotor at 24,000 rpm ($104,000 \times g$), 4°C .
- Recover the phage band by inserting a 3-ml syringe with a 25-G needle into the side of the tube just below the band and drawing it into the syringe (see Fig. 1.7.1 in CsCl/EtBr plasmid prep, UNIT 1.7).

Usually, three bands are visible: one blue phage band at the lowest gradient, one blue



Sketch 1.13A

band containing empty phage heads, and one white cell debris band. Occasionally one or two of these bands are missing. This is not a problem as long as the band that is visible is in the $d = 1.5$ g/ml layer. If no band is visible, there are probably not enough phage to make it worthwhile continuing the procedure. If the layers are not clearly defined, but a bluish band appears at about the right place, recover the band and determine its density by weighing 100 μ l of the lysate. If its density is ~ 1.5 , it probably is phage.

15. Transfer the phage to Beckman VTi50 quick-seal tubes. Fill tubes with CsCl solution ($d = 1.5$ g/ml).

About half the weight of the phage particle is protein; the other half is DNA. Phage DNA is denser than phage proteins. Thus, phages with genomes larger than wild-type are denser than phages with wild-type-length genomes, and phages with smaller genomes are less dense.

16. Spin 24 hr in VTi50 rotor at 30,000 rpm ($81,500 \times g$), 4°C.
17. Remove the band as shown in Figure 1.7.1 of CsCl/EtBr plasmid prep (UNIT 1.7), using a 3-ml syringe and a 25-G needle. Only one band should be visible.

Extract the phage DNA

Two methods are presented below for extracting DNA from purified phage. In steps 18a to 21a, cesium chloride is removed by dialysis prior to extraction of DNA with phenol and chloroform. In steps 18b to 21b, DNA is extracted directly from the isolated phage with high-grade formamide. The latter approach is quicker and possibly gentler.

Dialysis and phenol/chloroform extraction:

- 18a. Dialyze, with stirring, in 500 ml low-salt buffer for ≥ 4 hr at 4°C. Repeat twice.
- 19a. Extract three times by agitating gently for 20 min with an equal volume of buffered phenol.
- 20a. Extract twice using an equal volume of chloroform.
- 21a. Dialyze, with stirring, in 500 ml TE buffer, pH 8.0, for 8 hr at 4°C. Change buffer once. Proceed to step 22.

Residual phenol and chloroform are removed by dialysis rather than ethanol precipitation because purified phage DNA is very difficult to resuspend.

Formamide extraction:

- 18b. Measure the volume of phage band from step 17. Add 0.1 vol of 2 M Tris-Cl (pH 8.5)/0.2 M EDTA and invert to mix.
- 19b. Add 1 vol formamide, mix, and let stand 30 min at room temperature.
- 20b. Add 2 vol (each equal to the original volume of phage band in step 18b) of 100% ethanol at room temperature. Mix gently and microcentrifuge 1 to 2 min.
- 21b. Discard supernatant and rinse pellet with 70% ethanol. Remove all droplets of ethanol with a drawn-out pipet and dissolve the moist pellet in TE buffer, pH 8.0.
22. Measure DNA concentration as described in APPENDIX 3.

PREPARING DNA USING DEAE-CELLULOSE COLUMN CHROMATOGRAPHY

This protocol employs an ion-exchange resin to preferentially bind contaminants in crude phage lysate (*E. coli* DNA, RNA, and protein) while phage particles pass through the column. The result is a highly purified phage preparation. The DNA is then extracted from the purified phage with organic solvents and precipitated with ethanol. The method is simple, rapid, and does not require DNase, RNase, or CsCl density gradient centrifugation of either phage particles or DNA. The high-quality λ DNA obtained is suitable for cloning, sequencing, restriction enzyme digestion, ligation, and in vitro packaging.

Materials

TM buffer
2% sodium dodecyl sulfate (SDS; optional)
0.1 M EDTA (optional)
Sodium chloride
Polyethylene glycol (PEG) 8000
DEAE-cellulose (microgranular anion exchanger; Whatman DE52 #4057-050)
0.05 N HCl
10 M NaOH
Sodium azide
5 M NaCl
Ice-cold 100% isopropanol
TE buffer, pH 8.0 (APPENDIX 2)
25:24:1 phenol/chloroform/isoamyl alcohol (UNIT 2.1)
3 M sodium acetate, pH 6.0
70% ethanol and ice-cold 100% ethanol

Beckman JA-14 and JA-20 rotors (or equivalents)
15- and 30-ml Corex centrifuge tubes
10-ml disposable syringe (1.4 cm-i.d., optional; Becton Dickinson) *or*
1.5 \times 10-cm standard glass or disposable column (Bio-Rad)
Glass-fiber filter or glass wool (optional)

Additional reagents and equipment for preparing liquid or plate lysate (UNIT 1.12),
titering lambda phage (UNIT 1.11), agarose gel electrophoresis (UNIT 2.5), and
phenol extraction/ethanol precipitation (UNIT 2.1)

Prepare concentrated crude phage lysate

1. Prepare liquid or plate lysate and determine the phage titer. The lysate should contain $1\text{--}2 \times 10^{10}$ pfu/ml in a volume of 200 ml.

When using plate lysates, collect phage in TM buffer (instead of SM) from six big (150 mm) or fifteen small (90 mm) petri dishes and adjust volume to 200 ml.

Alternatively, a 0.7% agarose minigel can be used to detect λ DNA in phage lysates. Treat 20 μ l lysate with 2 μ l of 2% SDS and 2 μ l of 0.1 M EDTA for 5 min at room temperature. Load onto gel. Include size standard and control (nontreated) lanes. In this gel, the SDS/EDTA-treated lysate sample should show a distinct λ DNA band compared to the nontreated sample (Fig. 1.13.1).

2. Add 5.8 g NaCl (0.5 M final) and 20 g PEG 8000 (10% wt/vol final) to 200 ml phage lysate.

If the phage titer is low, it may be necessary to add NaCl first to release phage adhering to the debris. Remove debris by centrifuging 10 min in JA-14 at 6000 rpm (5500 \times g), 4°C, and then add PEG. Dissolve the PEG flakes by gentle stirring and place on wet ice 1 hr. This is enough time to precipitate most of the phage.

3. Separate precipitated phage by centrifuging 10 min in a JA-14 rotor at 6000 rpm, 4°C.

Drain liquid from the bottles by placing in an inverted position; and shake by hand to remove any remaining liquid.

4. Resuspend phage pellet in 3 ml TM buffer and transfer to a 15-ml Corex centrifuge tube. Add 3 ml chloroform, mix gently, and centrifuge 10 min in a JA-20 rotor at 5000 rpm ($3000 \times g$), 4°C. Collect the phage-containing upper aqueous phase without disturbing the PEG interface.

Corex tubes make it easy to see the PEG interface and the well-separated supernatant.

5. Add 3 ml TM buffer to the tube, mix, and centrifuge as in step 4. Save the aqueous fraction and combine with aqueous fraction from step 4. Adjust the volume to 6 ml with TM buffer.

This step removes phage particles trapped at the PEG interface.

Prepare the DEAE-cellulose column

6. Prepare a slurry of DEAE-cellulose by adding several volumes of 0.05 N HCl. Make sure the pH of the solution is below 4.5.

It is convenient to prepare enough DEAE-cellulose for 30 to 40 columns. Each column requires ~9 ml resin (for 200 ml phage lysate).

7. While stirring, add 10 M NaOH until the pH approaches 7.5.
8. Let the resin settle, decant or aspirate, and equilibrate the resin with TM buffer (see manufacturer's instructions).

Make sure the DEAE-cellulose is completely equilibrated with TM buffer. This may require repeating step 8 four or five times. This process also removes fine particles from the resin.

9. Adjust the slurry to 75% resin and 25% TM buffer.

For long-term storage (several months), add NaN_3 to a final concentration of 0.02% and place at 4°C.

10. Pour 9 ml resin in a 10-ml disposable syringe (1.4-cm i.d.) or a 1.5×10 -cm column. This will result in a bed height of 5 to 6 cm, appropriate for 200 ml of phage lysate.

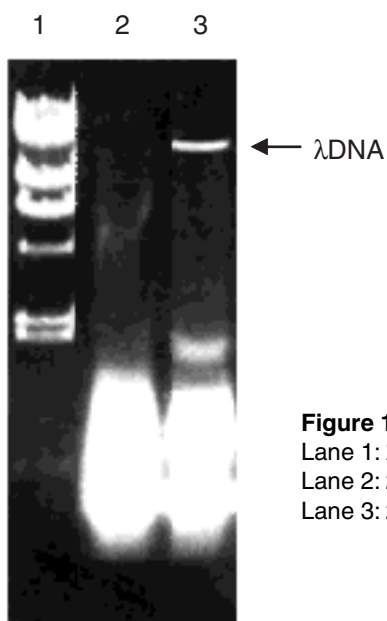


Figure 1.13.1 Agarose minigel of bacteriophage lambda lysates. Lane 1: λ DNA digested with *Hind*III. Lane 2: 20 μ l of lysate without SDS and EDTA. Lane 3: 20 μ l of lysate treated with 2% SDS and 0.1 M EDTA.

When using a disposable syringe, place a glass-fiber filter or glass wool at the bottom of the syringe to support the resin. Attach a small piece of tubing with a stopcock to the syringe tip to control the flow. Column preparation is easier with a standard glass or disposable column. See UNIT 10.10 for complete discussion of ion-exchange chromatography and Figure 3.4.1 for a sketch of homemade column setup.

Isolate the DNA

11. Load 6 ml crude phage solution from step 5 onto DEAE-cellulose column and elute with 10 ml TM buffer. Discard the first 3 ml of the eluate (void volume); collect the next 13 ml in a 30-ml Corex tube.

The 13-ml eluate may be collected in 1-ml fractions, if desired. Generally, phage elute immediately after the void volume. Fractions containing high-titer phage appear bluish due to light scattering. This indicates that the column is working well. Phage purification can also be confirmed by analyzing 5 μ l of each fraction on an agarose minigel (Fig. 1.13.1).

12. Add 2 ml of 5 M NaCl (400 mM final) and 10 ml of ice-cold 100% isopropanol (40% final) to the 13-ml eluate. Place at -20°C for 15 min.
13. Centrifuge 10 min in JA-14 rotor at 5000 rpm ($3800 \times g$), 4°C , and discard supernatant.

Invert centrifuge tubes until liquid is completely drained. Using tissue paper, wipe the walls of the centrifuge tube without touching the phage pellet. The pellet will appear as a feathery design adhering to wall of the tube. If there is not a sufficient pellet, abandon the experiment and start again from new lysate.

14. Resuspend the phage pellet in 0.8 ml TE buffer, pH 8.0, and divide equally into two 1.5-ml microcentrifuge tubes.
15. Extract once with phenol, twice with phenol/chloroform/isoamylalcohol, and finally with chloroform. Repeat extractions until no visible protein precipitate is seen at the interface.
16. Precipitate λ DNA by adding $\frac{1}{10}$ vol of 3 M sodium acetate, pH 6.0, and 2 vol ice-cold 100% ethanol.

In general, λ DNA forms a fibrous precipitate immediately after the addition of ice-cold ethanol at room temperature. A white cloudy precipitate indicates leftover protein. The λ DNA may be spooled out with a glass rod or condensed into a globule by gentle mixing.

17. Rinse the DNA pellet with 70% ethanol, air dry, and resuspend in TE buffer, pH 8.0. After dissolving completely, measure DNA concentration (UNIT 1.7) and adjust to 200 $\mu\text{g}/\text{ml}$ in TE buffer. Store at 4°C .

It may take several hours for the λ DNA to go into solution.

PREPARING DNA FROM SMALL-SCALE LIQUID LYSATES

ALTERNATE PROTOCOL

This protocol is useful for making small quantities of DNA to be used for restriction analysis. Phage are concentrated by centrifugation and their capsids are destroyed with phenol. The DNA is then ethanol precipitated.

Additional Materials

- 5 mg/ml DNase (*UNIT 3.12*)
- 10 mg/ml DNase-free RNase (*UNIT 3.13*)
- 0.05 M Tris·Cl, pH 8.0
- 3 M sodium acetate, to pH 4.8 with acetic acid

1. To approximately 50 ml liquid phage lysate (*UNIT 1.12*), add 10 μ l of 5 mg/ml DNase and 25 μ l of 10 mg/ml DNase-free RNase. Incubate 1 hr at 37°C.

This treatment will degrade the bacterial DNA and RNA released during lysis. The viscosity of the mixture should decrease.

2. Centrifuge 1½ hr at 27,000 rpm in an SW-28 rotor (132,000 \times g), 4°C.

Alternatively, pellet the phage by spinning 2¼ hr in JA-20 rotor at 20,000 rpm (48,000 \times g), 4°C. The pellet obtained in this manner will resuspend somewhat more easily.

3. Discard supernatant. Invert the tubes on an absorbent surface, e.g., paper towels, to remove any remaining liquid. Resuspend the phage pellet in 200 μ l of 0.05 M Tris·Cl, pH 8.0.

A small translucent pellet should be visible after the tubes are inverted.

4. Transfer the solution to a microcentrifuge tube and add 200 μ l buffered phenol. Vortex 20 min or shake 20 min in microcentrifuge tube shaker. Spin 2 min in microcentrifuge and save the aqueous (top) layer. Repeat phenol extraction.

Phenol denatures the phage capsids and releases the DNA. This denatured capsid protein appears as a thick white precipitate at the phenol/water interface. Vigorous agitation is necessary because the pellet is difficult to resuspend.

There should be less white precipitate after the second phenol extraction. If there is still a large amount at the interface, do a third extraction.

5. Add 200 μ l chloroform, shake well, and spin in microcentrifuge briefly. Save the aqueous (top) layer. Repeat.
6. Add 20 μ l of 3 M sodium acetate, pH 4.8, and precipitate DNA with 2 vol of 100% ethanol at room temperature. Spin in microcentrifuge for 10 min.
7. Remove supernatant. Wash pellet by adding 1 ml of 70% ethanol and spinning 5 min.
8. Remove supernatant. Dry pellet under a vacuum and resuspend the DNA in 100 μ l TE buffer, pH 8.0.

The DNA will resuspend more quickly if the pellet is still slightly wet.

3 μ l of the DNA suspension should be used for a restriction digest.

REAGENTS AND SOLUTIONS

CsCl solutions

$d = 1.3$ g/ml: 31.24 g CsCl + 68.76 ml H₂O

$d = 1.5$ g/ml: 45.41 g CsCl + 54.59 ml H₂O

$d = 1.7$ g/ml: 56.24 g CsCl + 43.76 ml H₂O

The equation for preparing the CsCl solutions is: % w/w = 137.48 - 138.11/d. This calculation assumes the CsCl has no water in it. Typically, CsCl from the shelf will have adsorbed water from the air; and the densities of these stock solutions will be lower than claimed. However, in our experience, the phage band is always found in the middle layer of the step gradient.

Low-salt buffer

0.05 M NaCl

0.05 M Tris·Cl, pH 7.5

0.01 M MgSO₄

5× polyethylene glycol (PEG) solution, 600 ml

207 g Carbowax (PEG 6000)

6 g dextran sulfate

49.5 g NaCl

350 ml H₂O

TM buffer

50 mM Tris·Cl, pH 7.5

10 mM MgSO₄

COMMENTARY

Background Information

This unit provides three methods for extracting and purifying DNA from intact phage particles. The three methods differ in the amount of starting material or phage lysate, and in the manner in which phage are concentrated and purified before lysing and DNA release.

The large-scale preparation (basic protocol) is useful when large quantities of highly purified DNA (over 200 μg) are required—for example, in the construction of libraries. It is based on traditional methods of bacteriophage lambda isolation, involving precipitation of phage from the lysate by polyethylene glycol (PEG) and subsequent purification by CsCl step and equilibrium density-gradient centrifugation (Yamamoto et al., 1970; Davis et al., 1980; Maniatis et al., 1982). It is necessary to use two different gradients to purify phage away from the considerable amount of cellular debris present in a large lysate. Following phage purification, DNA is extracted from capsids by one of two approaches. In one method, phage proteins are removed by a series of phenol extractions and resulting phage DNA is dialyzed in TE buffer to remove any remaining phenol or chloroform. In the other method, DNA is extracted directly from the capsids with high-grade formamide.

The first alternate protocol describes a moderate-scale procedure, capable of yielding enough pure lambda DNA (about 200 μg) for most standard DNA manipulations, such as cloning, sequencing, in vitro packaging, and Southern blotting. Phage particles are separated from cellular components in the lysate by ion-exchange chromatography, resulting in a high degree of purification in a single step. The chromatographic purification works as follows. DEAE (diethyl aminoethyl)-cellulose is an anion-exchange resin—i.e., it has positively charged groups that adsorb negatively charged molecules in buffer of near neutral pH and medium ionic strength. The major cellular contaminants in bacteriophage lysates are negatively charged (polyanionic) molecules (DNA, RNA, and proteins) which are preferentially adsorbed onto the positively charged groups of DEAE-cellulose (Creaser and Taussig, 1957; Benson and Taylor, 1984; Shuang-Young, 1986).

Two different ways to use column chromatography with Whatman DE52 anion exchange resin for phage purification have been reported. In the procedure described here, cellular contaminants are bound to the column, leaving phage particles free (Reddy et al., 1988;

White and Rosenzweig, 1989). Because SDS and proteinase K are not used in the nucleic acid extraction, the resulting DNA can be used with restriction endonucleases and DNA-modifying enzymes without further purification, and has been used successfully to make genomic expression libraries (Webb et al., 1989). Alternatively, binding phage particles to the resin with subsequent elution using a high-salt buffer (Helms et al., 1985) is used for processing several clones simultaneously to obtain a few micrograms of DNA for subcloning.

The second alternate protocol describes a rapid, small-scale method for phage purification which yields small amounts of moderately pure DNA, suitable for restriction analysis. In this procedure, phage are pelleted out of solution with a high-speed spin, and host nucleic acids are destroyed using RNase and DNase. (These enzymes have no effect on phage DNA as it is still packaged in phage heads.) Phage DNA is subsequently extracted from the capsids with phenol and is concentrated with ethanol precipitation.

Critical Parameters

Regardless of procedure, most failures to isolate lambda DNA are due to low phage titer in the initial lysate. Thus, hours of wasted time can be spared by first titrating the phage lysate. Phage can be titered by dilution plating (UNIT 1.11); if the titer is less than 10^9 pfu/ml, a new lysate should be made. An alternative and usually superior method to determine the amount of DNA is to treat 20 μ l of lysate with SDS/EDTA followed by electrophoresis on an agarose minigel (UNIT 2.5). The gel assay takes less time and provides a more accurate estimation of DNA present in both viable and nonviable phage particles. Figure 1.13.1 (lane 3) shows a typical gel resolving relative amounts of cellular nucleic acids and λ DNA (~50 ng) present in a lysate. One should be able to visualize ≥ 10 ng of λ DNA in the band. If there is no visible band, it is advisable to start from fresh lysate.

In chromatographic purification, the ratio of lysate to column size is an important variable. A 9-ml bed volume is sufficient for 200 ml of liquid lysate (at 10^{10} pfu/ml); if more lysate is used (or, if lysate has more highly concentrated cellular debris), the bed volume should be increased proportionately. If the column is overloaded, cellular DNA and RNA may coelute with the phage (to assess this, monitor column fractions with sample of crude lysate by running on an agarose gel). If DNA is contami-

nated, column fractions can be pooled and treated with DNase and RNase (UNITS 3.12 & 3.13) before organic solvent extraction of phage DNA. One should carefully prepare, equilibrate, and pack DEAE-cellulose according to instructions in the Whatman information leaflet (see also UNIT 10.10). Note that the type of DEAE-cellulose (Whatman DE52), as well as the pH and ionic strength of the buffer (use TM buffer recipe provided), are important. Chloroform-lysed cultures often contain very viscous, high-molecular-weight chromosomal DNA which can block the DEAE column. Such a lysate can be treated with DNase before loading onto the column.

After organic extraction, it may be preferable to dialyze rather than precipitate phage DNA, because large DNA can be difficult to resuspend. Lambda DNA is large and easily sheared after it is isolated. When intact DNA is desired (e.g., when phage arms are to be prepared for making a library), the DNA should be treated with special care: shaken, swirled, or vortexed very gently to mix.

Anticipated Results

Yield of λ DNA from each of these procedures depends upon the phage titer and the original lysate volume. If the initial lysate titer is around $1-2 \times 10^{10}$ pfu/ml, then 1 ml of phage lysate ought to yield ~1 μ g of DNA; that is, 2×10^{10} phage particles gives ~1 μ g purified phage DNA. Thus, the large-scale procedure, starting with 1000 ml of lysate, gives ~1 mg of phage DNA; the DEAE-column procedure (200 ml lysate) results in ~200 μ g DNA; and the small-scale procedure (50 ml lysate), provides ~50 μ g DNA.

Time Considerations

The large-scale prep employing gradient purification typically takes at least 3 days: ~30 min the first day to set up the PEG precipitation; ~5 hr the next day to spin down and collect the precipitate, run the step gradient, and load the equilibrium gradient; and ~30 min the third day to collect and phenol extract the phage band, followed by two dialysis steps lasting 8 hr. The phage lysates made as intermediates can be stored for months at 4°C without decreasing in titer. The small-scale prep takes ~4 to 6 hr and it is fairly easy to process a number of samples at once.

Column-purified phage and DNA isolation can be completed in 4 hr, starting from liquid or plate lysate. Preparation of the crude phage solution for the column takes ~2 hr. Another 2

hr are needed for column setup, column loading, elution, and DNA extraction. It is usually most convenient to process four to five different columns simultaneously. The DEAE-cellulose slurry can be prepared in advance and stored at 4°C for several months.

Literature Cited

- Benson, S. and Taylor, R.K. 1984. A rapid small-scale procedure for isolation of phage lambda DNA. *Biotechniques* 2:126-127.
- Creaser, E.H. and Taussig, A. 1957. The purification and chromatography of bacteriophages on anion-exchange cellulose. *Virology* 4:200-208.
- Helms, C., Graham, M.Y., Dutchik, J.E., and Olson, M.V. 1985. A new method for purifying lambda DNA from phage lysates. *DNA* 4:39-49.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. 1982. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Reddy, K.J., Kuwabara, T., and Sherman, L.A. 1988. A simple and efficient procedure for the isolation of high-quality phage lambda DNA using a DEAE-cellulose column. *Anal. Biochem.* 168:324-331.
- Shuang-Young X. 1986. A rapid method for preparing phage lambda DNA from agar plate lysates. *Gene Anal. Techn.* 3:90-91.
- Webb, R., Reddy, K.J., and Sherman, L.A. 1989. Lambda ZAP: Improved strategies for expression library construction and use. *DNA* 8:69-73.
- White, B.A. and Rosenzweig, S. 1989. A reliable method for the purification of bacteriophage lambda DNA. *Biotechniques* 7:694-695.

Key References

PEG precipitation

Yamamoto, K.R., Alberts, B.M., Benzinger, R., Lawhorne, L., and Treiber, G. 1970. Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification. *Virology* 40:734-744.

Step gradient and equilibrium gradient

Davis, R.W., Botstein, D., and Roth, J.R. 1980. *A Manual for Genetic Engineering: Advanced Bacterial Genetics*, pp. 70-113. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Includes a number of other protocols for preparation of DNA from lambda.

Ion-exchange purification

Reddy et al. 1988. See above.

Describes the details of the DEAE-cellulose column chromatography of lambda DNA upon which this protocol is based. Also describes a large-scale isolation procedure of lambda DNA from a 1-liter liquid lysate using a 45-ml bed volume DEAE column.

Contributed by Karen Lech
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

K.J. Reddy and L.A. Sherman
(column purification)
Purdue University
West Lafayette, Indiana

VECTORS DERIVED FROM FILAMENTOUS PHAGES

Introduction to Vectors Derived from Filamentous Phages

Many vectors in current use are derived from filamentous phages. These vectors are used because DNA inserted into them can be recovered in two forms—double-stranded circles and single-stranded circles. Foreign DNA is inserted into double-stranded vector DNA and, then, reintroduced into cells by transformation. Once inside the cells, double-stranded DNA replicates, giving rise both to new double-stranded circles and to single-stranded circles derived from one of the two strands of the vector. Single-stranded circles are packaged into phage particles and secreted from cells (which do not lyse). Centrifugation of a culture of infected cells yields a supernatant that is full of particles containing only a single strand of the phage DNA. This ready availability of single-stranded DNA has made possible new pro-

cedures for sequencing DNA (Chapter 7), mutagenesis (Chapter 8), and other techniques described in this book.

Techniques to isolate double-stranded DNA using these vectors are described in *UNIT 1.15*.

DEVELOPMENT AND USE OF FILAMENTOUS PHAGE VECTORS

Prototypes of the filamentous phage vectors are the M13mp derivatives (see Figs. 1.14.1 and 1.14.2). These vectors were developed by Joachim Messing and his co-workers, who also developed and disseminated simple and powerful techniques for working with them. The M13mp vectors are viable phages. Foreign DNA is inserted into a polylinker (a stretch of DNA that contains contiguous restriction sites) located in an inessential region

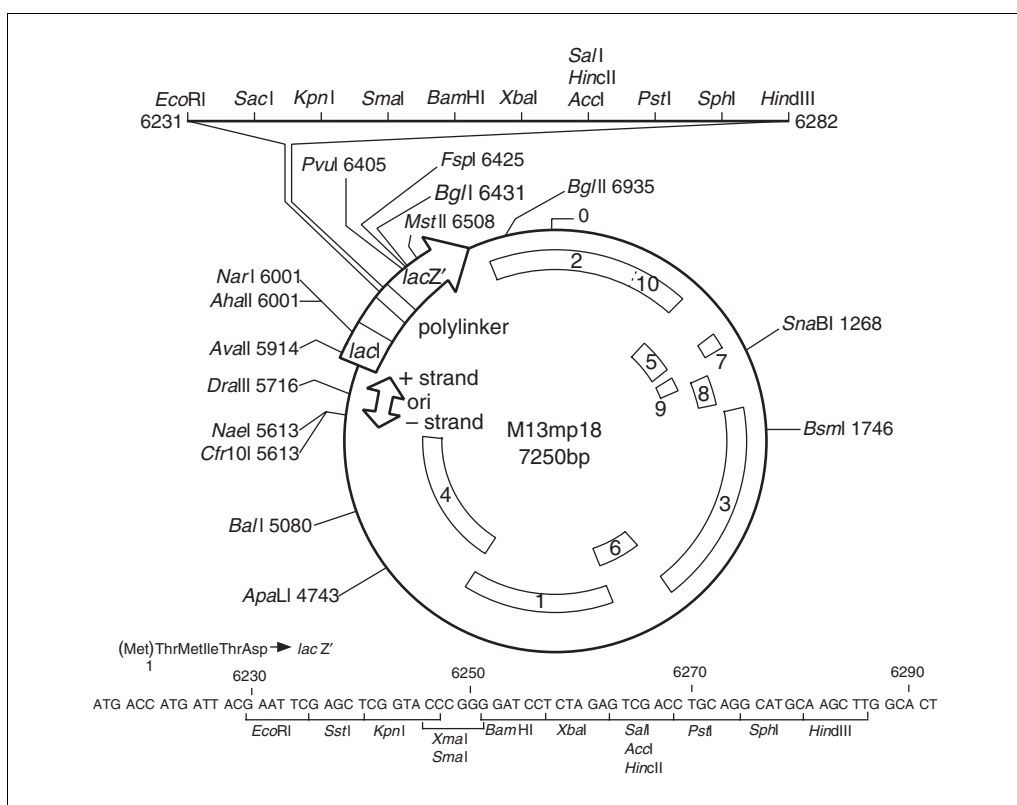


Figure 1.14.1 M13mp18. M13mp18 is one of the M13mp vectors made by Messing and colleagues. Insertion of DNA into the polylinker inactivates the *lacZ* alpha fragment. When insert-containing phages are plated under appropriate conditions (*UNIT 1.15*), they form colorless plaques; vectors that do not contain inserts form blue plaques (Yanisch-Perron et al., 1985, and references therein).

Contributed by David Greenstein and Roger Brent
Current Protocols in Molecular Biology (1990) 1.14.1-1.14.5
Copyright © 2000 by John Wiley & Sons, Inc.

Escherichia coli,
Plasmids, and
Bacteriophages

1.14.1

Supplement 13

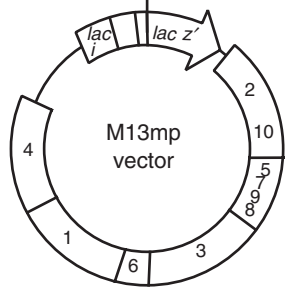
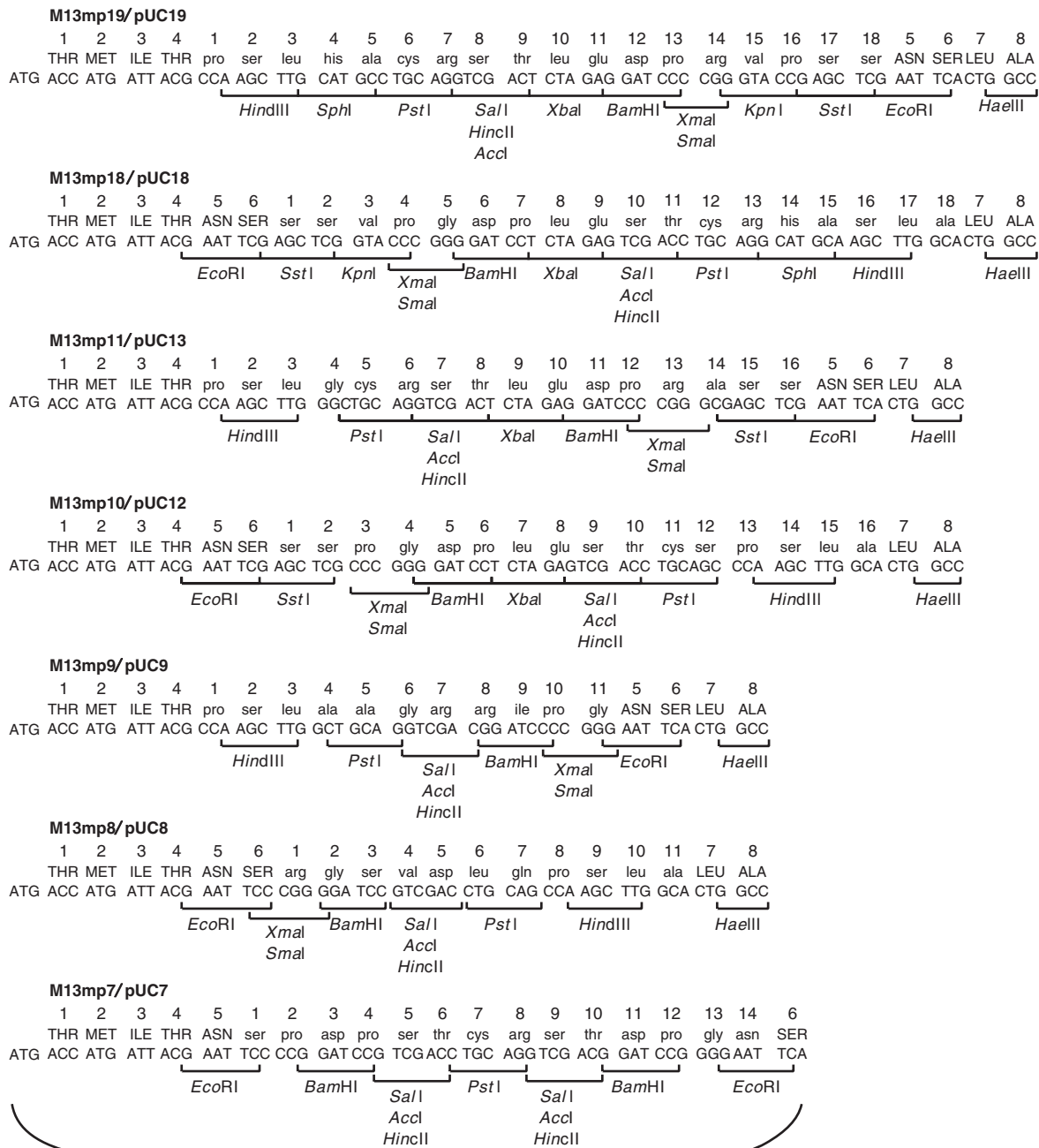


Figure 1.14.2 M13mp/pUC polylinkers. Sequence of polylinkers in the commonly used members of these two series of vectors (Yanisch-Perron et al., 1985, and references therein). Amino acids that have been added to the *lacZ* gene product by insertion of the polylinker are shown in lower case letters. The bracket shows location of polylinkers on vector.

1.14.2

of the phage genome. The polylinker is embedded in-frame within an alpha fragment of the *lacZ* gene. M13mp derivatives form blue plaques on lawns of cells that contain the *lacZ* omega fragment on plates with Xgal and IPTG. Double-stranded DNA is purified from phage-infected cells, cut within the polylinker with restriction enzymes, and foreign DNA is ligated to the cut vector. Insert-bearing phages form white plaques. Infected cells from the center of white plaques are grown in liquid culture then centrifuged to yield supernatants full of phage particles containing a single strand of DNA. Phage particles are concentrated and pure single-stranded DNA is extracted from them and used in other procedures.

Two inadequacies of the M13mp phages sparked development of subsequent generations of vectors. First, it is not always easy to obtain large amounts of double-stranded DNA. This is because DNA is obtained by lysing infected cells, but infected cells grow more slowly than uninfected cells, and there is no easy selection for cells that retain the phage. Second, phages that contain insertions of more than a few hundred nucleotides sometimes give rise to progeny phage in which some of the inserted DNA was deleted. This may be due to the fact that cells infected with large phages grow even more slowly than cells infected with small phages, so that cells containing deletion derivatives of phages with large inserts have a growth advantage.

In order to overcome these disadvantages, many vectors were developed (e.g., pEMBL) that contain a short stretch of DNA that includes the phage origin of replication. These vectors also contain a pMB1-derived origin of replication, a β -lactamase (ampicillin-resistance) gene, and a polylinker embedded within an alpha-complementing fragment of *lacZ*. The vectors, which are much smaller than the M13mp phages, are introduced into the cell by transformation; their continued presence can be ensured by growth of the cells in ampicillin-containing medium. Since these vectors do not encode phage proteins required for DNA replication, vector DNA replicates inside the cells using the pMB1 origin of replication. However, when cells containing these vectors are superinfected with wild-type *helper phage*, the phage origin of replication becomes active and single-stranded progeny phages are secreted into the medium together with progeny helper phage. Contamination with helper phage usually does not interfere with most applications because the helper does not contain genes en-

coding *lacZ* or ampicillin resistance, nor is it likely to contain any piece of nucleic acid used as a hybridization probe. Vectors developed even more recently (e.g., pUC118, pBluescribe) are designed to be used with a slightly improved helper phage and also often include SP6 or T7 phage polymerase promoters reading into the polylinker (see Fig. 1.5.3 and *UNIT 1.15* for details).

LIFECYCLE OF FILAMENTOUS PHAGES

The filamentous phages (f1, M13, and fd) are single-stranded DNA phages that are male-specific—that is, they infect male *E. coli* strains that contain an F factor (F⁺, F', or Hfr; see *UNIT 1.4*). These three phages are essentially independent isolates of the same phage, and differ only by a few nucleotide substitutions. The phage particle is shaped like a long thin tube. Its coat is composed primarily of thousands of monomers of the gene 8 product. The genome consists of a single-stranded, circular DNA molecule, 6407 nucleotides long, which runs down the length of the tube. In addition to coding for ten proteins, the genome contains an intergenic space (IG) between gene 4 and gene 2. The IG contains origins for (+)- and (–)-strand DNA synthesis, a signal for packaging the (+) strands into phage particles, and a transcription terminator. (A genetic map of the phage is shown in Figure 1.14.3.) Filamentous phages are useful cloning vectors because there appears to be no size limit for packaging; if longer genomes are generated, they are packaged into longer phage particles. The phage particle contains ~2700 monomers of the major coat protein, the product of gene 8, and several minor coat proteins at the ends. One of the minor coat proteins, the product of gene 3, attaches to the receptor at the tip of the F pilus of the host *E. coli*. Upon binding, the pilus is thought to retract, bringing the phage in contact with the bacterial cell surface. The coat proteins are removed from the phage particle and inserted into the bacterial cell membrane. After uncoating, the infecting circular single-stranded DNA is brought into the cytoplasm in a process that remains unclear.

Phage DNA replication takes place in three stages within the infected cell. In stage one (complementary strand synthesis), a complementary (–) strand of phage DNA is synthesized. This synthesis converts the infecting single-stranded circular DNA, the (+) strand, into double-stranded replicative-form (RF) DNA. Synthesis of the (–) strand begins at the (–)

strand origin (within the IG), and is carried out by the host *E. coli* replicative enzymes.

In the second stage of replication, the intracellular pool of RF DNA is increased by the sequential action of both replication origins. Rolling-circle type replication from the (+)-strand origin is followed by conversion of the progeny single-stranded circles to double strands by replication starting at the (–)-strand origin. The resulting RF DNA molecules are intermediates in DNA synthesis and transcription templates for the synthesis of the phage-encoded proteins. Transcription proceeds in the same direction as (+)-strand rolling-circle replication. It is suspected that transcription into the replication fork, in the direction of the (–) strand, would inhibit replication, as proposed for the replication of the *E. coli* chromosome (Brewer, 1988). Although this idea has not been rigorously tested, it is attractive because it provides an explanation for two common cloning problems: why certain pieces of DNA are not clonable, and why, occasionally, only one of two possible orientations of an insert are sometimes recovered in a subcloning experiment.

The (+)-strand replication origin that directs rolling-circle replication has a bipartite structure. It consists of an essential core origin region (~50 bp) and an adjacent A+T-rich “enhancer” sequence (~100 bp) which increases replication ~100-fold. The core origin binds the

initiator protein (gene 2 protein) and the enhancer binds the *E. coli* integration host factor (IHF). The polylinker cloning site in cloning vectors of the M13mp series disrupts the replication enhancer sequence, but these vectors have acquired compensatory mutations in gene 2 that restore efficient replication. Consequently, virtually any sequence can be introduced into the polylinker without disrupting DNA replication.

The gene 2 protein is a multifunctional protein that plays several roles in phage DNA replication. It binds cooperatively to the (+)-strand origin in two steps, bends the origin DNA, and introduces a specific nick in the (+) strand of RF DNA. The 3'-hydroxyl end of the nick serves as the primer for (+)-strand rolling-circle replication. (This activity makes the gene 2 protein a useful enzyme for producing uniquely nicked DNA molecules in vitro.) After nicking, the DNA molecules still need the gene 2 protein for unwinding and replication. Finally, upon completion of a round of synthesis, gene 2 protein cleaves and circularizes the displaced single strand (see UNIT 1.15).

The third stage of replication (single-strand production) occurs late in infection. This stage of DNA replication is asymmetric because the (–)-strand origin functions at a reduced level late in the infection and therefore mainly (+) strands are produced. These single-stranded circles, (+) strands, are packaged into phage

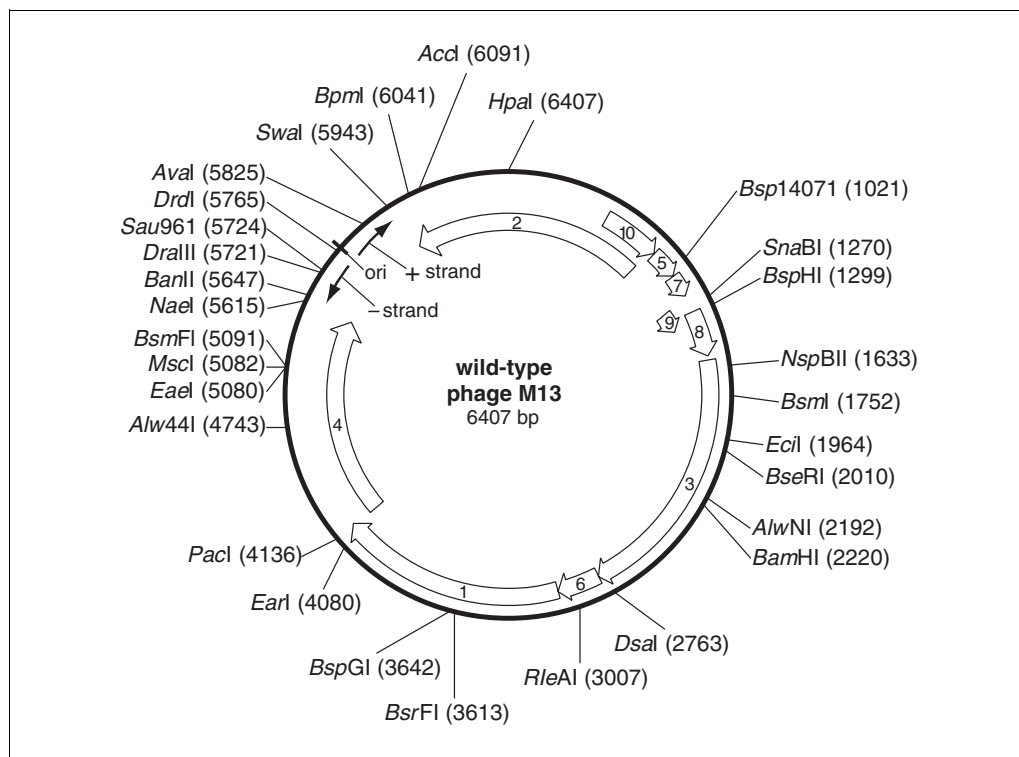


Figure 1.14.3 Lifecycle of wild-type M13, described in this unit and in Rasched and Oberer (1986).

particles for export instead of being converted into RF DNA.

Ordinarily, infection with filamentous phages is not lethal and the host cells do not lyse, although their growth rate slows ~2-fold. The reduced growth rate of the host cells accounts for the turbid plaques that the filamentous phage form on a lawn of sensitive cells. Surprisingly, several of the phage proteins have been found to be lethal to *E. coli* when expressed from a plasmid with enhanced replication (high copy number). Evidently, DNA replication in the infected cell is tightly regulated to prevent this lethality. The mechanism whereby this regulation is achieved is not completely understood, but the phage-encoded single-stranded DNA binding protein (gene 5 protein) appears to be required for it. In the steady state of infection, the RF DNA is maintained as a plasmid with a copy number of ~20 to 40, and phage are continuously exported at a rate of 100 to 200/hr. The infected state is quite stable with ~1 cell in every 1000 cells becoming uninfected each generation.

LITERATURE CITED

- Beck, E. and Zink, B. 1981. Nucleotide sequence and genome organization of filamentous bacteriophage f1 and bacteriophage fd. *Gene* 16:35-38.
- Brewer, B.J. 1988. When polymerases collide: Replication and the transcriptional organization of the *E. coli* chromosome. *Cell* 53:679-686.

- Dente, L., Cesareni, G., and Cortese, R. 1983. pEMBL: A new family of single-stranded plasmids. *Nucl. Acids Res.* 11:1645-1655.
- Greenstein, D. and Horiuchi, K. 1988. Integration host factor interacts with the DNA replication enhancer of filamentous phage f1. *Proc. Natl. Acad. Sci. U.S.A.* 85:6262-6266.
- Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* 101:20-78.
- Messing, J., Gronenborn, B., Muller-Hill, B., and Hofschneider, P.H. 1977. Single-strand filamentous DNA phage as a carrier for in vitro recombinant DNA. *Proc. Natl. Acad. Sci. U.S.A.* 74:3642-3646.
- Rasched, I. and Oberer, E. 1986. Ff coliphages: Structural and functional relationships. *Microbiol. Rev.* 50:401-427.
- Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103-119.
- Zinder, N.D. and Boeke, J.D. 1982. The filamentous phage as vector for recombinant DNA: A review. *Gene* 19:1-10.
- Zinder, N.D. and Horiuchi, K. 1985. Multiregulatory element of filamentous bacteriophages. *Microbiol. Rev.* 9:101-106.

Contributed by David Greenstein
and Roger Brent
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

Preparing and Using M13-Derived Vectors

Cloning vectors derived from filamentous phage are extremely useful because they allow cloned DNA to be isolated as either single- or double-stranded DNA. This unit contains protocols for preparing both forms of DNA and for characterizing inserts in M13-derived vectors. A protocol is also presented for preparing single-stranded DNA from plasmids using superinfection with helper phage. This method is advantageous because it allows cloned DNA to be maintained in the form of a plasmid while permitting single-stranded DNA to be isolated for DNA sequencing.

BASIC PROTOCOL

ISOLATING SINGLE M13-DERIVED VECTORS

To ensure homogeneity of stocks of native or insert-containing vectors, it is necessary to start from a single infected cell, or from a single plaque. This is achieved by plating out serial dilutions of an original culture stock.

Materials

- E. coli* strain infected with M13-derived vector (e.g., JM101 infected with M13mp18)
- 2× TY medium (UNIT 1.1)
- 20 mg/ml IPTG in H₂O (stored in aliquots at −20°C)
- 20 mg/ml Xgal in dimethylformamide (stored in aliquots at −20°C)
- 45°C top agarose (UNIT 1.1)
- H plates, prewarmed to 37°C (UNIT 1.1)
- 5-ml Falcon tubes with caps or equivalent plastic tubes
- 37°C incubator

NOTE: All materials coming into contact with *E. coli* must be sterile.

1. Make a series of 1:10 dilutions of the infected strain or a phage stock in 2× TY medium. Place 100 μl of each dilution into separate 5-ml Falcon tubes with caps. Label the tubes in order to keep better track of them.

The vector is usually stored in an infected host at −80°C in 20% glycerol.

The vector can also be stored as isolated DNA or as phage. DNA is introduced into cells by transfection (UNIT 1.11); phage, by infection (described in steps 2 to 5 below). If the vector contains a plasmid origin and a drug-resistance marker, vector-containing cells can be isolated as single colonies on antibiotic-containing medium. If the vector forms viable plaques, vector-containing cells can be isolated by streaking single colonies from the tiny patch of cells in the center of the plaques.

The following steps assume the vector is an M13mp vector, which contains a stretch of DNA that encodes the alpha fragment of β-galactosidase. These vectors make blue plaques on cells containing the β-galactosidase omega fragment (UNIT 1.4). Insertion of DNA into the polylinker inactivates the gene encoding the alpha fragment and gives rise to vectors that have colorless plaques.

2. To each tube add the following:

- 200 μl noninfected bacteria grown to saturation
- 10 μl IPTG
- 40 μl Xgal
- 3 ml 45°C top agarose

It is convenient to prepare a larger batch of the first three ingredients beforehand.

3. Mix by rapidly inverting the tubes twice and pour on individual H plates prewarmed to 37°C.

4. Let the top agarose harden 10 min at room temperature and transfer the plates to a 37°C incubator.
5. After overnight growth, save only the plates that contain less than ~100 plaques.

PREPARING SINGLE-STRANDED PHAGE DNA FROM M13-DERIVED VECTORS

BASIC PROTOCOL

Cells containing the vector and the insert are grown up in liquid medium. The cells are collected by centrifugation, and single-stranded DNA is prepared from the phage particles in the supernatant.

Materials

E. coli strain (male-type, e.g., JM105)

M13-derived vector

2× TY medium (UNIT 1.1)

PEG solution (UNIT 1.7)

TE buffer (APPENDIX 2)

Buffered phenol (UNIT 2.1)

3 M sodium acetate

100% ethanol and cold 70% ethanol

Sterile toothpicks

37°C shaking water bath

Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5)

1. Inoculate 50 ml of 2× TY medium with 0.5 ml of an overnight culture of *E. coli*. The overnight should be grown from a single colony on a minimal plate that does not contain proline. Dispense 2-ml aliquots in 10-ml culture tubes. Inoculate each tube with a colorless phage plaque picked with a sterile toothpick.

Many male strains of E. coli used for propagation of single-stranded phage (e.g., JM101 or JM105; Table 1.4.5) carry a deletion that removes the proA and proB genes, but also contain an F' plasmid that contains the proA⁺ and proB⁺ genes. Picking a colony from a plate that lacks proline thus ensures that the cells contain the F' plasmid, and that the phage will be able to infect them.

2. Shake at 37°C for 5 to 8 hr.
3. Transfer 1.5 ml of the culture to a microcentrifuge tube. Centrifuge 5 min at room temperature.
4. Pour 1.25 ml of the supernatant into a fresh microcentrifuge tube.

To store the strain, resuspend the bacterial pellet by vortexing, adjust to 30% glycerol, and freeze at -80°C.

5. Add 250 µl PEG solution. Mix and leave for 15 min at room temperature.
6. Centrifuge 5 min and discard the supernatant.

The phage pellet should be visible at this stage.

7. Centrifuge again for 5 min. Carefully remove all remaining traces of PEG using a drawn-out Pasteur pipet.
8. Add to the viral pellet 200 µl TE buffer. Resuspend the pellet completely, then add 100 µl buffered phenol.

Escherichia coli,
Plasmids, and
Bacteriophages

1.15.2

9. Vortex 1 min, let stand for 5 min, vortex again, and centrifuge 5 min.
10. Remove 175 μ l from the upper aqueous phase; transfer to a new microcentrifuge tube.
11. Add 20 μ l of 3 M sodium acetate and 400 μ l of 100% ethanol.
12. Leave \geq 1 hr at -20°C to precipitate the DNA, or 15 min in a dry ice/ethanol bath.
13. Centrifuge 5 min, pour off the supernatant, add 1 ml cold 70% ethanol, centrifuge again, discard the supernatant, and dry briefly in a vacuum desiccator.

For some applications of the single-stranded DNA, such as oligonucleotide-directed mutagenesis (UNIT 8.1), it is advisable to use very clean DNA. For these applications, aspirate off the supernatants from this step using a drawn-out Pasteur pipet.

14. Dissolve the pellet in 25 μ l TE buffer.

Yield can be as high as 10 μ g DNA/ml culture.

15. Analyze 1 to 2 μ l of the DNA solution by electrophoresis on a 1% agarose gel and store the remainder at -20°C .

BASIC PROTOCOL

PREPARING DOUBLE-STRANDED REPLICATIVE-FORM DNA

The disadvantage of many protocols for isolating double-stranded replicative-form (RF) DNA is that the DNA is isolated late in the infectious cycle, at a time when it is least abundant. The following protocol involves isolation of double-stranded DNA from chloramphenicol-treated cells. Addition of chloramphenicol at a low concentration (15 μ g/ml) soon after infection prevents the accumulation of the phage gene 5 protein, which inhibits (–)-strand DNA synthesis. Enough gene 2 protein is synthesized in the first 15 min to allow accumulation of RF DNA.

Materials

F⁺ or Hfr *E. coli* strain
 2 \times TY medium (UNIT 1.1)
 Recombinant phage
 20% glucose
 1 mg/ml chloramphenicol in ethanol, freshly prepared
 Sorvall SS-34 rotor or equivalent

1. Inoculate 20 ml of 2 \times TY medium plus 0.1 ml of 20% glucose with uninfected bacteria. Incubate in a 37 $^{\circ}\text{C}$ shaking water bath until OD₆₀₀ = 0.8 to 1.0.
2. Infect the cells with the recombinant phage at an MOI (see glossary in chapter introduction) of 20 to 50. Incubate 15 min at 37 $^{\circ}\text{C}$.
3. Add 0.3 ml of 1 mg/ml chloramphenicol in ethanol (final concentration, 15 μ g/ml) to the culture. Incubate the culture an additional 2 hr.
4. Centrifuge 10 min at 4000 \times g (6000 rpm) to harvest the cells. Prepare double-stranded DNA using the usual procedures for isolation of plasmid DNA (see UNITS 1.6 & 1.7).

PREPARING SINGLE-STRANDED DNA FROM PLASMIDS USING HELPER PHAGE

BASIC PROTOCOL

Cells containing plasmids with filamentous phage origins (usually the f1 origin) are infected with helper phage. The helper phage provides the gene 2 protein that drives the plasmid into the f1 mode of replication. The gene 2 protein nicks the (+)-strand origin of the plasmid and initiates rolling-circle replication, resulting in the production of single-stranded circles of the plasmid DNA. The helper phage also provides the DNA packaging and export functions. Single strands of the plasmid are packaged into phage coats and secreted into the supernatant. It is important to remember that only the (+) strand is packaged efficiently. Therefore, only the DNA strand of the insert that is in the same 5'→3' orientation as the phage (+)-strand origin will be packaged.

Materials

F⁺ or Hfr *E. coli* strain (Table 1.4.5) containing a plasmid (pUC118, pBS, or equivalent; see commentary)

2× TY medium (UNIT 1.1)

37°C shaking water bath and 65°C water bath

Sorvall SS-34 rotor or equivalent

1. Grow cells in 2× TY (containing an appropriate antibiotic) at 37°C to an OD₆₀₀ of 0.1.

A fresh overnight culture started from a single colony can be diluted 1:50 and grown to an OD₆₀₀ of 0.1. Generally, 1- to 5-ml cultures will yield enough DNA for sequencing (typical yields are 0.2 to 1 μg plasmid single strands/ml culture).

2. Infect the cells at an MOI of 20. Some plasmids and bacterial strains seem to require addition of more phage (MOI of 50), while some give good yields with addition of less phage (MOI 5 to 10).

When performing this protocol for the first time, it is helpful to try several different MOIs (1, 10, 20, and 50). Use the least amount of phage necessary to give a good yield of plasmid single strands because a portion of the input phage will always be recovered.

3. Grow the cells 4.5 hr at 37°C with vigorous shaking.

For convenience, the infected cells can be grown overnight and the supernatant collected the following day. In many cases, this longer growth period will result in contamination of the supernatant with chromosomal DNA from lysed bacteria.

4. Centrifuge 10 min at 4000 × g (6000 rpm). Collect the supernatant and heat at 65°C for 15 min to kill any residual bacteria. Prepare single-stranded DNA from the supernatant (steps 4 to 13 of the second basic protocol, preparing single-stranded phage DNA from M13-derived vectors). Analyze the single-stranded DNA on an agarose gel (UNIT 2.5) with the helper phage serving as a control.

INTRODUCTION OF PHAGE DNA INTO CELLS

Both double-stranded and single-stranded vector DNA can be introduced into CaCl₂-treated competent bacteria by transfection, just as if the vector DNA molecules were plasmids. Single-stranded DNA usually yields about ten times fewer transformants than the same amount of double-stranded molecules.

Special applications may require vector DNA grown on an F- strain. The fact that phage DNA can be introduced into cells by transformation also makes it possible to produce phages from F-hosts. However, since phages produced by infection of these cells cannot infect neighboring cells, phages introduced into F- cells should always contain a drug-resistance marker so that transformed cells can be selected on antibiotic-containing plates.

Methods of transformation described in *UNIT 1.8* are used to introduce vector DNA into cells, with the following adaptations where appropriate.

1. If the vector contains a plasmid replicator and a drug-resistance gene, then simply select transformed cells on antibiotic-containing plates.
2. If the vector is able to form plaques, add 200 μ l of noninfected cells at late log phase ($OD_{600} = 0.6$ to 0.8)—grown as in *UNIT 1.2*—to the plating-out mixture. Add 2.5 ml H top agar (*UNIT 1.1*), plate the mixture as a lawn on an H plate (*UNIT 1.1*), and incubate at 37°C until plaques appear (as in first basic protocol).

DETERMINING SIZE OF INSERTS IN SINGLE-STRANDED VECTORS

This method allows a quick comparison of a large number of viral DNAs without purification of the single-stranded DNA (Messing, 1983).

The size of inserts is estimated by comparing the mobility of DNA purified from insert-containing phage to the mobility of DNA from phages lacking inserts or containing inserts of known size. This procedure allows one to compare the sizes of single-stranded circular molecules that are several kilobases long; however, the resolution on agarose gels usually does not allow detection of inserts that are <300 bp long.

Materials

2% sodium dodecyl sulfate (SDS)
Loading buffer

1. Pipet 1.5 ml of an infected culture into a microcentrifuge tube.
2. Spin for 1 min. Remove 20 μ l of the supernatant and mix with 1 μ l of 2% SDS and 3 μ l loading buffer. Repeat this process with supernatants of cultures of vector phage without inserts, and of other phage which contain inserts of known size.
3. Run these samples on a 0.7% agarose gel and analyze as described in *UNIT 2.5*. If desired, DNA can be transferred from gel to membrane and assayed with a radioactive probe, for example, to determine that the fragment has the right sequence as well as the right size.

DETERMINING INSERT ORIENTATION

Only one strand of the inserted DNA is contained in the phage's (+) strand and made into phage particles. This means that phage DNAs containing identical inserts in opposite orientations will hybridize with each other along the stretch of inserted DNA (see sketch 1.15A). The structures thus formed have different electrophoretic mobility than two unassociated single-stranded viral DNAs. This property of single-stranded phage vectors can be used to determine if two phages have an insert in opposite orientations (Howarth et al., 1981), which can be useful for sequencing both ends of a fragment of DNA.

Additional Materials

5 M NaCl

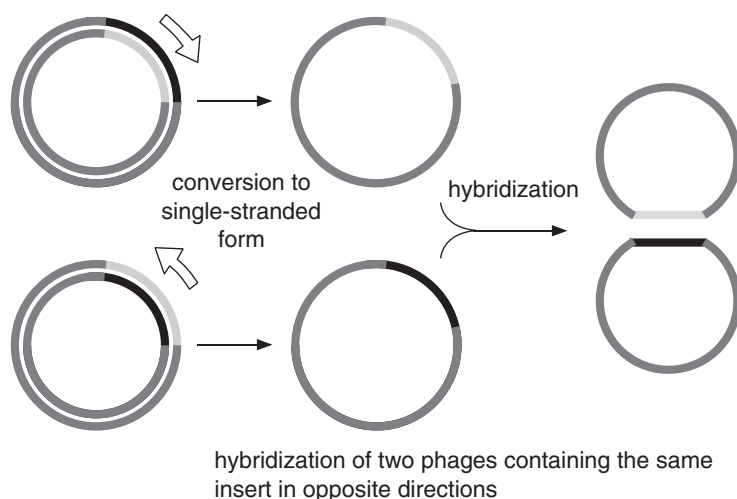
1. Mix 20 μ l of two supernatants with 1 μ l of 2% SDS in a microcentrifuge tube. In separate tubes mix 40 μ l of each individual supernatant with 2 μ l of 2% SDS.
2. Add 2 μ l of 5 M NaCl to each of the three tubes and incubate 30 min at 60°C.
3. Run all three samples on a 0.7% agarose gel and visualize bands by staining with ethidium bromide as described in *UNIT 2.5*.

If many samples must be screened, it is possible to grow, lyse, and process recombinant infected cells in wells of microtiter plates.

It is also possible to determine the relative orientation of inserts in purified preparations of single-stranded DNA. To do this, mix 5 μ l DNA from each of two independent plaques with 0.5 μ l of 5 M NaCl. Incubate 30 min at 60°C, then analyze on an agarose gel as described above.

REAGENTS AND SOLUTIONS**Loading buffer**

50% glycerol
0.2 M EDTA, pH 8.3
0.05% bromphenol blue



Sketch 1.15A

COMMENTARY

Background Information

In M13 cloning vectors (Messing, 1983), a portion of the *E. coli lac* operon which bears a polylinker (a stretch of DNA which contains contiguous recognition sites for many different restriction enzymes) is inserted in the intergenic region (Messing et al., 1977). On an omega-fragment producing host, these vectors form blue plaques, and vectors with inserts form white plaques.

A plaque formed by M13 vectors is a zone of infected cells within a lawn of noninfected cells. Infected cells grow more slowly than uninfected cells. Plaques made by phage that contain inserts are usually smaller than those made by wild-type phage. This is because phage DNA which contains inserts is replicated more slowly, and it takes longer for cells infected with those phages to produce progeny plaques. Many people prefer to store the recombinant phages not as plaques but as purified single-stranded DNA.

Other single-stranded cloning vectors contain the origins of replication of phages like ϕ 1. These vectors usually also contain the pBR322 plasmid origin of replication, a drug resistance coding gene, and a polylinker inserted in frame into the portion of the *lacZ* gene coding for the alpha peptide. Most of these vectors lack gene 2, or the other phage genes necessary to form single-strand phage particles.

In cells these vectors are double-stranded and replicate using the pBR322 origin of replication. However, when cells are infected with wild-type helper phages, the plasmids replicate using the phage origin of replication, and single-stranded copies of the plasmid are encapsidated into phage filaments, which are then secreted into the culture medium along with copies of the helper. Contamination with helper phage does not interfere with most applications because the helper does not contain genes encoding *lacZ* or ampicillin resistance, nor is it likely to contain any piece of nucleic acid used as a hybridization probe.

Perhaps the most common application of these vectors is in sequencing using the dideoxy method (UNIT 7.4). In this procedure, small fragments of DNA from a plasmid or lambda-derived vector are subcloned into M13mp vectors. Vectors containing the inserts are identified by the fact that they form colorless plaques on a lawn of omega-fragment producing cells growing on Xgal + IPTG plates (see UNIT 1.4).

For further background information see

lifecycle of filamentous phages, UNIT 1.14, and key references listed there.

Critical Parameters

Isolating single M13-derived vectors. Plates should be used immediately after plaques have appeared. Phages should be freshly plated every time a single plaque is to be picked for DNA amplification. Phages diffuse, and use of an old plate can result in cross contamination. Cells to be infected and lawns of infected cells must be grown at 37°C. If cells are grown at temperatures below 34°C, sex pili do not form and phages cannot infect.

Preparing single-stranded phage DNA from M13-derived vectors. It is critical to remove PEG from the preparations, since traces of PEG inhibit the activity of many DNA polymerases. Contaminated templates can be extracted with chloroform prior to ethanol precipitation. However if step 7 (of the first basic protocol) has been followed carefully, chloroform extraction is not necessary.

Preparing replicative-form (RF) DNA. Typical yields of double-stranded DNA from chloramphenicol-treated cells are 50 to 200 μ g per 20-ml culture depending on how well the particular recombinant phage grows. Too much chloramphenicol will result in reduced yields, while too little will result in contamination with single-stranded DNA.

Sometimes after the cesium chloride gradient is equilibrated, a third band appears between the plasmid DNA band (lowest) and the genomic DNA band (highest). This is made of single-stranded DNA and should be left behind.

Preparing single-stranded DNA from plasmids using helper phage. Start with a single colony of male *E. coli* strain that is harboring a recombinant plasmid containing a filamentous phage replication origin such as pUC118 (Viera and Messing, 1987) or pBS (Stratagene).

The strain of helper phage used must be appropriate for the experiment. Several helper phages such as IR-1 (Enea and Zinder, 1982), R408 (Russell et al., 1986), M13K07 (Viera and Messing, 1987), and VCSM13 (derived from M13K07; Stratagene) are available. The phages IR-1 and R408 are more stable than VCSM13 and M13K07, but the latter phages contain a kanamycin-resistance gene that aids in selection. The helper phages VCSM13 and R408 were designed to favor the production of plasmid single strands. VCSM13 has a phage (+)-strand replication origin with

a defective replication “enhancer” sequence. Therefore, it doesn’t compete as well for the gene 2 protein and hence favors the production of plasmid single strands. In contrast, the helper phage R408 has a defective packaging signal (see *UNIT 1.14*) causing the plasmid single strands to be preferentially packaged into phage particles. Regardless of the helper phage used, the single-stranded DNA prepared from the supernatant will contain some DNA from the helper phage. This is usually not a problem for DNA sequencing because a primer can be chosen that is specific for the plasmid single strands.

Introduction of phage DNA into cells.

Transformation of CaCl₂-treated competent bacteria is carried out as described in *UNIT 1.8*. If the vector forms plaques, it is not necessary to phenotypically express, and the cells can be plated out immediately following the heat shock step. Using F⁻ competent cells is a common mistake, since the recombinant phage will not form plaques on a lawn of F⁻ cells. However, F⁻ competent cells can be used if they are diluted with uninfected F⁺ cells following transformation and plated together. In this case, the F⁻ cells will produce phage that will make plaques on the mixed lawn. As long as the F⁺ cells are in abundance, plaques will be seen.

Determining size of inserts in single-stranded vectors. Single-stranded DNA is analyzed from a phage supernatant after extraction from the phage particle. It is important to vortex the phage with the SDS to remove the coat proteins from the phage DNA. Failure to do so will result in incomplete extraction and smearing of the DNA bands. If the culture supernatant contains bacteria or chromosomal DNA from lysed bacteria, the phage DNA will be contaminated and other DNA bands will be visible. If this is a recurrent problem, the position of these extra bands can be determined by running an uninfected culture supernatant as a control.

Time Considerations

Plaques can be seen after ~4 hr. The color of plaques can be determined after a few more hours of incubation.

When preparing single-stranded phage DNA, incubation and processing samples are performed the same day. The only limiting factor is the number of samples to process at the same time. If there are too many, it is possible to store the supernatants (step 5, first basic protocol) at 4°C for 24 hr and then recentrifuge before proceeding to the next step.

Preparing single-stranded DNA with helper

phage requires a shorter incubation period (4.5 hr) after inoculation with the phage. Preparing RF DNA from chloramphenicol-treated cells can be accomplished within one day after the incubation. An additional day may be required if the sample is to be purified on a CsCl gradient.

Introducing phage DNA into cells, determining insert size, and determining insert orientation each takes ~2 hr.

Literature Cited

- Enea, V. and Zinder, N.D. 1982. Interference resistant mutants of phage f1. *Virology* 122:222-226.
- Howarth, A.J., Gardner, R.C., Messing, J., and Sheperd, R.J. 1981. Nucleotide sequence of a naturally occurring deletion of a mutant of a cauliflower mosaic virus. *Virology* 112:678.
- Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* 101:20-78.
- Messing, J., Gronenborn, B., Muller-Hill, B., and Hofschneider, P.H. 1977. Single-strand filamentous DNA phage as a carrier for in vitro recombined DNA. *Proc. Natl. Acad. Sci. U.S.A.* 74:3642-3646.
- Russell, M., Kidd, S., and Kelley, M.R. 1986. An improved filamentous helper phage for sequencing single-stranded plasmid DNA. *Gene* 45:333-338.
- Viera, J. and Messing, J. 1987. Production of single-stranded plasmid DNA. *Methods Enzymol.* 153:3-11.

Key References

- Dente, L., Sollazzo, M., Cesareni, G., and Cortese, R. 1985. The pEMBL family of single-stranded vectors. In *DNA cloning: A Practical Approach* (D.M. Glover, ed.) pp. 101-108. IRL Press, Oxford.
- Reviews biology of non-M13-derived single-stranded vectors.*
- Howarth et al. 1981. See above.
- Describes techniques for determining insert orientation.*
- Messing, J. 1983. See above.
- Maintenance, propagation, and titration of filamentous phage vectors are extensively described.*

Contributed by David Greenstein
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

Claude Besmond
Centre de Recherche Inserm
Paris, France

Escherichia coli,
Plasmids, and
Bacteriophages

1.15.8

Recombineering: Genetic Engineering in Bacteria Using Homologous Recombination

UNIT 1.16

Over the past few years, *in vivo* technologies have emerged which, due to their efficiency and simplicity, already complement and may one day replace standard genetic engineering techniques. The bacterial chromosome and episomes can be engineered *in vivo* by homologous recombination using PCR products and synthetic oligonucleotides (oligos) as substrates. This is possible because bacteriophage-encoded recombination functions efficiently recombine sequences with homologies as short as 35 to 40 bases. This technology, termed recombineering, allows DNA sequences to be inserted or deleted without regard to presence or location of restriction sites.

To perform recombineering, a bacterial strain expressing a bacteriophage recombination system is required. The phage enzymes can be expressed from either their own promoter or from a heterologous regulated promoter. Expressing the genes from their endogenous phage promoter confers the advantage of tight regulation and coordinate expression, which results in higher recombination frequencies. This is an important advantage, since in many cases high recombination frequencies will be essential to obtaining a desired recombinant. The authors of this unit routinely use a defective λ prophage located on the *E. coli* chromosome, and have recently transferred the critical elements of this prophage to a number of different plasmids (Thomason et al., 2005; also see Commentary). In this prophage system, the phage recombination functions are under control of the bacteriophage λ temperature-sensitive *ci857* repressor. At low temperatures (30° to 34°C), the recombination genes are tightly repressed, but when the temperature of the bacterial culture is shifted to 42°C, they are expressed at high levels from the λ *p_L* promoter. In the plasmid construct of Datsenko and Wanner (2000), the recombination genes are located on a plasmid and expressed from the arabinose promoter. The Datsenko and Wanner plasmid and some of the authors' plasmid constructs have temperature-sensitive origins of DNA replication. The plasmid-based systems have the advantage of mobility—they can be transferred among different *E. coli* strains or to *Salmonella typhimurium* and possibly other gram-negative bacteria. Using the prophage system located on the bacterial chromosome is more facile if the recombineering is targeted to a plasmid. After induction of the recombination functions, the modifying DNA, either a double-stranded (ds) PCR product or a synthetic single-stranded (ss) oligonucleotide (oligo), is introduced into the prophage-containing strain by electroporation. Recombinants are obtained either by selection or screening of the population of cells surviving electroporation. Once the desired construct is obtained, the prophage can be removed by another recombination. Alternatively, engineered alleles on the chromosome can be moved into a different host by P1 transduction. Plasmids with temperature-sensitive replication origins can be lost from the recombinant strain by growth at the appropriate temperature.

Preparation of electrocompetent cells that have expressed the recombineering functions from the λ promoter and their transformation with double-stranded DNA (dsDNA) or single-stranded DNA (ssDNA) is the first procedure described in this unit (see Basic Protocol 1). Support Protocol 1 describes a two-step method of making genetic

Escherichia coli,
Plasmids, and
Bacteriophages

1.16.1

Contributed by Lynn Thomason, Donald L. Court, Mikail Bubunenko, Nina Costantino, Helen Wilson, Simanti Datta, and Amos Oppenheim

Current Protocols in Molecular Biology (2005) 1.16.1-1.16.21
Copyright © 2005 by John Wiley & Sons, Inc.

Supplement 70

alterations without leaving any unwanted changes. In the latter protocol, the first of the two series of steps is a recombineering reaction that replaces the sequence to be modified with an antibiotic-resistance cassette and a counter-selectable marker (e.g., *sacB*, which is toxic when cells are grown on medium containing sucrose; Gay et al., 1985). The second series of steps is a subsequent recombination that replaces the antibiotic cassette and the counter-selectable marker with the desired genetic alteration. Moreover, the same selection and counter-selection can be reused to make further modifications. Support Protocol 2 describes retrieving a genetic marker (cloning) from the *E. coli* chromosome or a coelectroporated DNA fragment and moving it onto a plasmid. Whereas the above protocols generally use a selection to find the recombinants, Support Protocols 3 and 4 describe methods to screen for unselected mutations. A protocol is also included describing removal of the defective prophage (see Basic Protocol 2). Alternate Protocols 1 and 2 are methods for recombineering with an intact prophage and introducing mutations onto bacteriophage λ , respectively.

STRATEGIC PLANNING

Before attempting to modify the *E. coli* chromosome or a plasmid, the DNA sequence of the desired final construct should be determined. A DNA-analysis computer program such as Gene Construction Kit (GCK; Textco Software; <http://www.textco.com/>) or Vector NTI (Invitrogen) is invaluable for this task. Having the sequence of both the original genome arrangement and the designed final construct as electronic files facilitates the design of oligonucleotides to be used as primers for PCR or as ssDNA recombination substrates themselves. The computer-determined sequences also allow rapid design of primers to analyze and verify the potential recombinants. One must be aware of gene-regulation issues when designing the constructs. Bringing in a promoter with an antibiotic cassette can help in establishing drug resistance; however, transcription from this promoter can extend beyond the drug marker and affect distal genes. The authors of this unit have designed several drug cassettes with their promoter, open reading frame, and transcription terminator region, as described in Yu et al. (2000); primers for amplification are listed in Table 1.16.2. One must also be careful of possible polarity effects and avoid creating unwanted fusion proteins when generating recombinants.

The DNA substrate used for recombineering depends on the desired change. For a sizeable insertion, such as a drug cassette, a PCR product is generated that contains 40 to 50 base pairs of flanking homology to the chromosomal or plasmid target at each end. This homology is provided at the 5' end of each synthetic primer. Following this region of chromosomal homology, ~20 bases of homology to the drug cassette provides the primers to amplify the cassette sequence. Thus, two primers are designed that will each be 60 to 70 nucleotides (nt) long: the 5' ends provide homology to the targeted region and the 3' ends provide homology to the cassette (see Fig. 1.16.1). Careful primer design is crucial (see above). The efficiency of recombineering with dsDNA can approach 0.1% (Yu et al., 2000). If deletions, small substitutions, or base changes are desired, a synthetic single-stranded oligo of ~70 to 100 nt can be used. The oligo should have 35 to 40 nt of complete homology flanking the alteration. Recombineering with ssDNA in wild-type *E. coli* containing the defective λ prophage gives efficiencies approaching 1% (Ellis et al., 2001), and if host mismatch repair is inactivated, either by mutation or by using an oligo that creates a C-C mismatch, a 20% to 25% recombination frequency is achievable (Costantino and Court, 2003). This extremely high frequency means that, for oligo recombination, it is possible to create recombinants without a selection and find them by screening. Since an oligonucleotide corresponding to the lagging strand of DNA replication is some 20-fold more efficient than its complement, it is worthwhile to determine the direction of replication through one's region of interest and use the oligo

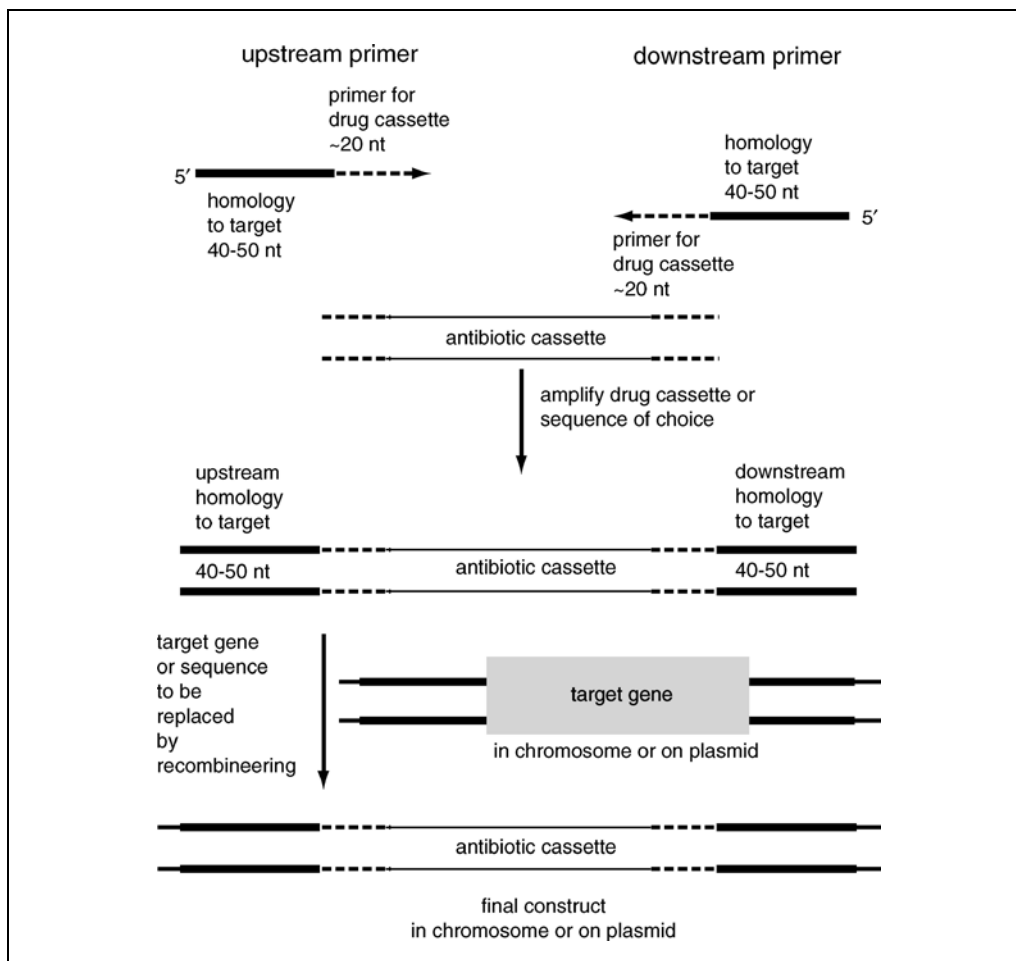


Figure 1.16.1 Targeting of an antibiotic cassette. Two primers with 5' homology to the target are used to PCR amplify the antibiotic cassette. The PCR product is introduced by electroporation into cells induced for the Red recombineering functions. The Red functions catalyze the insertion of the cassette at the target site, which may be on the bacterial chromosome or on a plasmid.

that corresponds to the lagging strand (Costantino and Court, 2003). It is also possible to rescue or clone gene(s) from the bacterial chromosome onto a plasmid. To do this, create a linear PCR product of the plasmid using primers with homology flanking the target sequence, designed so that the sequence will be incorporated onto the circular plasmid in the appropriate orientation. The PCR-amplified linear plasmid will require an origin of DNA replication and a selective marker.

MAKING ELECTROCOMPETENT CELLS AND TRANSFORMING WITH LINEAR DNA

This basic protocol describes making electrocompetent cells that are preinduced for the recombination functions and transforming them with the appropriate DNA to create the desired genetic change. As noted in Strategic Planning, the phage recombination functions are repressed by the phage λ temperature-sensitive cI857 repressor, so that they are not expressed when the cells are grown at low temperature (30°C to 34°C) but are highly expressed when the culture is shifted to 42°C. See Commentary for additional considerations before executing the procedure.

BASIC PROTOCOL 1

Escherichia coli,
Plasmids, and
Bacteriophages

1.16.3

Table 1.16.1 Bacterial Strains Commonly Used for Recombineering

Strain	Genotype
DY329	W3110 $\Delta lacU169 nadA::Tn10 gal490 pgl\Delta8 \lambda cI857\Delta(cro bioA)$ (Tet ^R)
DY330	W3110 $\Delta lacU169 gal490 pgl\Delta8 \lambda cI857 \Delta(cro-bioA)$
DY331	W3110 $\Delta lacU169 srlA::Tn10 \Delta recA gal490 pgl\Delta8 \lambda cI857 \Delta(cro-bioA)$ (Tet ^R)
DY378 ^a	W3110 $\lambda cI857 \Delta(cro-bioA)$
DY380	<i>mcrA</i> $\Delta(mrr-hsdRMS-mcrBC) \phi 80dlacZ\Delta M15 lacX74 deoR recA1 endA1 araD139 \Delta(ara, leu)7697 galU gal490 pgl\Delta8 rpsL nupG \lambda(cI857ind1) \Delta\{(cro-bioA)\} <> tetRA\}$ (Tet ^R)
DY441	DY329 with <i>cat-sacB</i> inserted between <i>cI857</i> and <i>rexA</i> (Tet ^R , Cm ^R)
HME5	W3110 $\Delta lacU169 \lambda cI857 \Delta(cro-bioA)$
HME45 ^b	W3110 $gal490 pgl\Delta8 \lambda cI857 \Delta(cro-bioA)$
HME63	W3110 $\Delta lacU169 \lambda cI857 \Delta(cro-bioA) galKam mutS <> amp$
HME64	W3110 $\Delta lacU169 \lambda cI857 \Delta(cro-bioA) galKam uvrD <> kan$

^aDH10B derivative (Invitrogen).

^bGives less background on low concentrations of chloramphenicol than DY378.

Materials

- Purified PCR product or oligonucleotide primers with ~40 to 50 bases of flanking homology on either side of desired change (also see *UNIT 15.1*)
- Bacterial strain expressing the defective lambdoid prophage recombination system λ Red (Table 1.16.1; strains are available from the Court Laboratory; court@ncifcrf.gov)
- LB medium and plates (*UNIT 1.1*), without antibiotic
- Medium lacking carbon source: M9 medium (*UNIT 1.1*) or 1 × TM buffer (*APPENDIX 2*)
- Selective plates—minimal plates (*UNIT 1.1*) if selecting for prototrophy *or* rich plates (also *UNIT 1.1*) containing the following concentrations of antibiotic (depending on drug cassette used):
 - 30 μ g/ml ampicillin
 - 30 μ g/ml kanamycin
 - 10 μ g/ml chloramphenicol
 - 12.5 μ g/ml tetracycline
 - 50 μ g/ml spectinomycin
- 30° to 32°C incubator
- 32° and 42°C shaking water baths
- 125- and 250-ml Erlenmeyer flasks, preferably baffled
- Refrigerated low-speed centrifuge with Sorvall SA-600 rotor (or equivalent)
- 35- to 50-ml plastic centrifuge tubes
- 0.1-cm electroporation cuvettes (Bio-Rad), chilled
- Electroporator (Bio-Rad *E. coli* Pulser is used in authors' laboratory)
- Additional reagents and equipment for PCR (*UNIT 15.1*), gel electrophoresis of DNA (*UNIT 10.2*), purification of DNA by ethanol precipitation (*UNIT 2.1*; optional; commercially available PCR cleanup kit may be substituted), electroporation (*UNIT 1.8*), isolation of bacterial colonies by streaking (*UNIT 1.3*), restriction enzyme digestion (*UNIT 3.1*), and DNA sequencing (Chapter 7).

Table 1.16.2 PCR Primers and Suggested Source of Template for Amplifying Drug Cassettes (from Yu et al., 2000)

Gene	Source	Primer sequence
Ampicillin	pBluescript SK(+) (Stratagene)	5' CATTCAAATATGTATCCGCTC 5' AGAGTTGGTAGCTCTTGATC
Tetracycline	<i>Tn10</i>	5' CAAGAGGGTCATTATATTTTCG 5' ACTCGACATCTTGGTTACCG
Chloramphenicol	pPCR-Script Cam (Stratagene)	5' TGTGACGGAAGATCACTTCCG 5' ACCAGCAATAGACATAAGCG
Kanamycin	<i>Tn5</i>	5' TATGGACAGCAAGCGAACCG 5' TCAGAAGAACTCGTCAAGAAG
Spectinomycin	DH5 α PRO (Clontech)	5' ACCGTGGAAACGGATGAAGGC 5' AGGGCTTATTATGCACGCTTAA

Prepare DNA for transformation

1. Design and procure the oligos to use for PCR-mediated generation of a dsDNA product, or for use in single-stranded oligo engineering.

The sequences of the primers used to amplify the common drug cassettes are listed in Table 1.16.2. Remember to add the homologous targeting sequence to the 5' ends of the oligos.

UNIT 15.1 describes general considerations for primer design.

2. Make the PCR product (UNIT 15.1), examine it by gel electrophoresis (UNIT 10.2), and gel purify by isolating the desired band if unwanted products are obtained.

If the DNA is gel purified, avoid exposing it to ultraviolet light, which will damage it and result in lower recombination frequencies.

3. Clean up the PCR product by ethanol precipitation (UNIT 2.1) or using a commercially available kit to remove salt.

If a plasmid template is used to construct a PCR-amplified drug cassette, any intact circular plasmid remaining will transform the cells efficiently and give unwanted background. This background can be minimized by using a linear plasmid template for the PCR and by digesting the completed PCR reaction with DpnI before using it for electroporation. Always include a control reaction of uninduced cells transformed with the PCR product, to give a measure of any unwanted intact plasmid background.

Prepare bacterial cultures

4. Inoculate the suitable bacterial strain (Table 1.16.1) from frozen glycerol stock or a single colony into 3 to 5 ml LB medium. Shake at 30° to 32°C overnight.

Most of the authors' strains containing the defective lambdaoid prophage are W3110 derivatives; however the prophage can be moved into other backgrounds. See Commentary for details. Plasmids expressing the recombination functions can be put into any strain of choice.

Either 30° or 32°C is acceptable for the low temperature throughout the procedure, since either temperature allows good repression by the cI857 repressor. The cultures will grow more rapidly at 32°C.

5. Equilibrate two shaking water baths to 32°C and 42°C, respectively. Add ~0.5 ml of the overnight culture to 35 ml of LB medium in a 250-ml (baffled) Erlenmeyer flask.

This is a 70-fold dilution. One must make sure that one's dilution is at least 50-fold. Higher dilutions will also work, but the cells will take longer to grow to the appropriate density.

***Escherichia coli,
Plasmids, and
Bacteriophages***

1.16.5

If targeting the recombineering to a plasmid, or expressing from a plasmid, add antibiotic as appropriate to maintain selection during growth. If an alternate method of inducing the recombination functions is used (i.e., addition of arabinose), the inducer should be added to the medium. In this case remember to include an additional flask containing an uninduced culture as a negative control. The authors have found that expression of the λ Red functions from the plasmids of Datsenko and Wanner (2000) is enhanced by use of 10 mM arabinose (for Ara⁺ strains).

6. Place the flask in the 32°C shaking water bath and grow cells at 32°C with shaking for ~2 hr (the time will vary with different strains and dilutions).

The cells are ready when the A₆₀₀ is between 0.4 and 0.6. It is important not to over-grow the cells, since stationary phase cells do not express the recombination functions well.

Induce recombination functions

7. Transfer half the culture to a 125-ml (baffled) Erlenmeyer flask and place that flask in the 42°C water bath. Shake 15 min at 220 rpm to induce. Leave the remainder of the culture at 32°C; this will be used as the uninduced control that lacks recombination activity. While the cells are inducing, fill an ice bucket with an ice-water slurry.
8. Immediately after inducing for 15 min at 42°C, rapidly cool the flask in the ice-water slurry with gentle swirling. Leave on ice for ≥ 5 min. Follow the same cooling protocol with the uninduced 32°C culture. While the cells are on ice, precool the centrifuge to 4°C and chill the necessary number of 35- to 50-ml plastic centrifuge tubes, labeled for induced and uninduced cells.

The temperature shift is unnecessary when a chemical inducer like arabinose is used.

Make electrocompetent cells

9. Transfer both the induced and uninduced cultures to the appropriately labeled chilled 35- to 50-ml centrifuge tubes. Centrifuge 7 min at $4600 \times g$ (6700 rpm in a Sorvall SA-600 rotor), 4°C. Aspirate or pour off supernatant.
10. Add 1 ml ice-cold distilled water to the cell pellet in the bottom of each tube and gently resuspend cells with a large pipet tip (do not vortex). Add another 30 ml ice-cold distilled water to each tube, seal, and gently invert to mix, again without vortexing. Centrifuge tubes again as in step 9.

All subsequent resuspensions of cells through step 16 should be done gently and without vortexing. Preparation of the cells for electroporation washes out any added chemical inducing agent.

11. Decant the 30-ml supernatant very carefully from the soft pellet in each tube and resuspend each cell pellet in 1 ml ice-cold distilled water.

Remove tubes from the centrifuge promptly. The pellet is very soft and care should be taken not to dislodge it, especially when processing multiple tubes.

12. Transfer resuspended cells to microcentrifuge tubes. Microcentrifuge 30 to 60 sec at maximum speed, 4°C. Carefully aspirate supernatant. In each of the tubes, resuspend the cell pellet in 200 μ l cold distilled water, which will provide enough material for four or five electroporations.

For routine procedures when optimal recombination frequency is not necessary, e.g., when selection is used to find recombinants, electrocompetent cells can be stored at -80°C after resuspending the cell pellet in 15% (v/v) glycerol. For highest efficiency, use freshly processed cells.

Introduce DNA by electroporation

13. Chill the desired number of 0.1-cm electroporation cuvettes on ice. Turn on the electroporator and set to 1.80 kV.

Brands of cuvettes and electroporator other than Bio-Rad may work, but have not been tested in the authors' laboratory. 0.2-cm cuvettes may require different electroporation conditions (consult electroporator instruction manual) and standardization to obtain optimal recombination frequencies.

14. In microcentrifuge tubes on ice, mix 100 to 150 ng of salt-free PCR fragment (from step 3) or 10 to 100 ng of single-stranded oligonucleotide with 50 to 100 μ l of the suspension of induced or uninduced cells (from step 12). Do the mixing and subsequent electroporation rapidly; do not leave the DNA-cell mixes on ice for extended periods. Be sure to include the following electroporation reactions and controls:

- a. Induced cells plus DNA.

This is the culture that should yield the designed recombinants.

- b. Induced cells without DNA.

This is a control to identify contamination, determine the reversion frequency, and obtain some idea of the efficiency of the selection.

- c. Uninduced cells plus DNA.

This control tells whether there is some contaminating factor in the DNA that is contributing to the selected colonies (for example, intact plasmid template from the PCR reaction will give rise to drug-resistant colonies here).

15. Introduce the DNA into the cells by electroporation (UNIT 1.8).

The time constant should be greater than 5 msec for optimal results. Low time constants indicate problems with the cells, the DNA, or even the equipment.

16. Immediately after electroporation, add 1 ml LB medium to the cuvette using a micropipettor with a 1000- μ l pipet tip. If transforming with a drug cassette, transfer the electroporation mix to sterile culture tubes and incubate the tubes with shaking at 30° to 34°C for 1 to 2 hr to allow expression of the antibiotic resistance gene.

Even if not selecting for drug resistance, it is still recommended that the cells outgrow to recover from the shock of electroporation. It has been observed in the authors' laboratory that omitting the outgrowth reduces the cell viability ~10-fold.

An alternative, and in the author's experience more reliable, method for outgrowth is to spread appropriate dilutions of cells (see step 18 below) on a sterile 82-mm diameter nitrocellulose filter atop a rich (LB) plate. Incubate this plate for 3 hr at 30° to 34°C. After the incubation, transfer the filter to the appropriate selective drug plate using sterile forceps. This method is preferable because the cells are less dense and more efficient outgrowth is achieved.

Determine cell titers

17. Make serial 1:10 dilutions of the electroporation mix through 10^{-6} using M9 medium or 1 \times TM (Tris Magnesium) buffer, dispensing 0.9 ml M9 or TMG and 0.1 ml of the cell suspension per tube.

The dilutions can be made in rich medium if a selection for antibiotic resistance is applied.

18. To determine total viable cell count, spread 100 μ l of 10^{-5} and 10^{-6} dilutions on LB plates (rich plates without drug). Incubate the plates at 30° to 34°C for 1 to 2 days, depending on the growth requirements of the recipient strain.

19. To determine recombinant cell count, plate cells on selective plates as follows depending on the anticipated recombinant frequency.
 - a. If efficient recombination is expected, spread both 10 and 100 μl of the 10^{-1} and 10^{-2} dilutions.
 - b. If low numbers of recombinants are expected, spread 100 μl each of a 1:5 and 1:10 dilution.

The authors routinely obtain ten-fold or higher recombinant yields with the prophage system than with the Datsenko and Wanner plasmids.

For the no-DNA and uninduced controls, plate 200 μl directly on selective plates.

Since targeting to the chromosome results in a lower copy number of the drug cassette than is present with a multicopy plasmid, antibiotic concentrations must be adjusted accordingly. The authors routinely use the drug concentrations recommended above for chromosomal constructs. Minimal plates are used for selection based on prototrophy.

20. Incubate plates at the appropriate temperature (30° to 34°C).

At 30°C, colonies may take two days to come up on LB plates and 3 to 4 days on minimal plates. Candidates should be purified by streaking for single colonies and retested for the appropriate phenotype.

Analyze recombinants

21. Once recombinant clones are identified, confirm the presence of the desired mutation(s) by PCR analysis (UNIT 15.1) followed by DNA sequencing (Chapter 7) or restriction digestion analysis, if appropriate.

The design of the PCR primers depends on the changes made. For an antibiotic cassette or other insertion, the primer pair used to amplify the insertion can also be used to confirm its presence. These primers will not determine whether the cassette has integrated at the desired location, however. The recombinant junctions can be confirmed with the help of two additional primers (all four should have compatible annealing temperatures) pointing outwards from the cassette; use one primer flanking the cassette and one internal cassette primer to amplify the unique junctions created by the recombination reaction. A primer pair that hybridizes to the external flanking sequences on each side rather than the insertion itself can also be used to demonstrate loss of the target sequences and presence of the insertion. Unwanted mutations can be introduced by heterologies (variations) in the synthetic primer population (Oppenheim et al. 2004); therefore, it is important to confirm the final construct by sequence analysis, especially the regions derived from the original primers.

SUPPORT PROTOCOL 1

MANIPULATING *cat-sacB* FOR COUNTER-SELECTION AND GENE REPLACEMENT

This protocol describes a two-step method to create precise genetic changes without otherwise altering the DNA sequence. First, *cat-sacB* (or another counter-selectable cassette) is placed on the DNA; this is then replaced with the desired alteration in a second recombineering. The final construct will not have a drug marker.

Additional Materials (also see Basic Protocol 1)

Template for amplification of *cat-sacB*: bacterial strain DY441 (DY329 with a *cat-sacB* insertion on the *E. coli* chromosome) or the plasmid pEL04 (Lee et al., 2001). pEL04 has previously been called both pK04 and *pcat-sacB*. Both DY441 and pEL04 are available from the Court Laboratory; court@ncifcrf.gov

SpeI and *DpnI* restriction endonucleases (UNIT 3.1)

Primer L *sacB*: 5'-homology sequence- ATC AAA GGG AAA ACT GTC CAT
AT -3'

Primer R *cat*: 5'- homology sequence- TGT GAC GGA AGA TCA CTT CG -3'

Invitrogen Platinum High Fidelity enzyme
 LB-Cm plates: LB plates (*UNIT 1.1*) containing 10 µg/ml chloramphenicol
 M63 minimal glycerol plates with sucrose (see recipe) or LB plates (*UNIT 1.1*)
 lacking NaCl but containing 6% (w/v) sucrose
 Medium lacking carbon source: e.g., M9 medium (*UNIT 1.1*)
 Thermal cycler (MJ Research PTC-100)
 Additional reagents and equipment for PCR (*UNIT 15.1*) and purification of DNA
 (*UNIT 2.1*)

Amplify the *cat-sacB* element for recombineering

1. If using pEL04, completely cleave *pcat-sacB* with *SpeI*.

This step is unnecessary if the chromosomal insertion of cat-sacB is used as template for PCR. Failure to digest when the plasmid is used as template will give a high background of intact plasmid transformants.

2. Amplify *cat-sacB*, either from the cleaved pEL04 or from DY441 with the Invitrogen Platinum High Fidelity enzyme and an MJ Research thermal cycler, using 50 pmol of each primer and the following cycling program:

1 cycle:	2 min	94°C	(denaturation)
9 cycles:	15 sec	94°C	(denaturation)
	30 sec	55°C	(annealing)
	3.5 min	68°C	(extension)
19 cycles:	15 sec	94°C	(denaturation)
	30 sec	55°C	(annealing)
	3.5 min		
	(adding 5 sec/cycle)	68°C	(extension)
1 cycle:	7 min	68°C	(extension)
1 cycle:	indefinite	4°C	(hold)

The PCR primers used to amplify the cat-sacB element are at the 3' end of chimeric primers that include 5' segments of bacterial target homology ~40 to 50 nt in length. Hence each primer is 60 to 70 nt long.

Since the PCR product is greater than 3 kb, it can be difficult to amplify. The above conditions have been used successfully in the authors' laboratory.

3. If pEL04 was used as template, digest the completed PCR reaction with *DpnI* to specifically eliminate the methylated plasmid template. Purify the PCR product to remove salt (*UNIT 2.1*).

Perform electroporation and recombination to insert *cat-sacB* at the desired location

4. Insert the *cat-sacB* cassette into the chromosome as described above (see Basic Protocol 1, steps 4 to 21; also see Background Information for important tips) using the following techniques specific for the *cat-sacB* cassette.
 - a. Select chloramphenicol-resistant (Cm^R) colonies and purify on LB-Cm plates to isolate single colonies.
 - b. Test several Cm^R isolates for sensitivity to sucrose, either on minimal glycerol plates containing 5% sucrose or on LB plates lacking NaCl but containing 6% sucrose (Blomfield et al., 1991). Use the parental transformation strain as a sucrose-resistant control. Determine that the insertion is correct before proceeding with the next step of the procedure.

*The sucrose-resistant parent serves as a control for growth in the presence of sucrose. Sucrose sensitivity needs to be tested because the PCR process generates mutations that inactivate *sacB* in a fraction of the clones.*

Perform electroporation and recombination to replace *cat-sacB* with chosen allele

5. Use a confirmed Cm^R/sucrose-sensitive candidate from step 4 as the starting bacterial strain for a second round of recombineering by carrying out Basic Protocol 1, steps 1 to 21; however, suspend the electroporated cells at step 16 of Basic Protocol 1 in a final volume of 10 ml instead of 1 ml of LB medium, and incubate with aeration (applied via shaking in a shaking water bath) at 30° to 32°C for 4 hr to overnight for outgrowth following electroporation.

*The higher dilution promotes better cell recovery and allows complete segregation of recombinant chromosomes that no longer carry the *cat-sacB* cassette from nonrecombinant sister chromosomes that still contain it. If outgrowth is inadequate and sister chromosomes are not fully segregated, the presence of the cassette on one chromosome will confer sucrose sensitivity to the entire cell, thus preventing recovery of recombinants. This is generally true in counter-selection experiments and illustrates a common problem encountered with them.*

6. Centrifuge cells 7 min at 4600 × g, 4°C. Remove supernatant, then wash the cells twice, each time by resuspending in 1 ml minimal medium lacking a carbon source, such as M9, centrifuging again as before, and removing the supernatant. Resuspend and dilute the cells for plating, and spread appropriate dilutions of the cells on either LB sucrose plates or minimal glycerol plates (if the defective prophage was moved to another strain, the metabolic requirements of that strain must be considered). In parallel, be sure to include a control of electroporated cells to which no PCR product was added, to determine the frequency of spontaneous sucrose-resistant mutants; these spontaneous mutants will retain the Cm^R cassette.

*The frequency of spontaneous sucrose-resistant cells is normally ~1 in 10⁴. Thus, in the recombination experiment, the sucrose-resistant colonies that arise are of two types, spontaneous mutants like those found in the control, and deletions caused by replacing *cat-sacB* by recombination with the PCR product. The frequency of the latter is optimally 10- to 100-fold greater than that of spontaneous mutation.*

7. Purify ~10 sucrose-resistant colonies by streaking to isolate single colonies (UNIT 1.3) and then test for chloramphenicol sensitivity (Cm^S)—spontaneous mutants will be Cm^R while the recombinants will be Cm^S. Screen the Cm^S/sucrose-resistant colonies by PCR (see Basic Protocol 1, step 21), then sequence to confirm the presence of the desired change.

Screen a minimum of ten colonies. If high-efficiency recombination is not achieved, more colonies will need to be screened. Note that the frequency of spontaneous mutants remains relatively constant and provides an internal control for determining the efficiency of the recombination.

RETRIEVAL OF ALLELES ONTO A PLASMID BY GAP REPAIR

Often it is desirable to retrieve a DNA sequence from the bacterial chromosome, either to clone and amplify a gene, to express a gene under a given promoter, or to create a gene or operon fusion. To do this with recombineering, a PCR product with homology to the target at the ends is made from a linearized plasmid DNA and introduced into cells expressing the Red system. This homology will allow recombination with the sequence to be retrieved, yielding a circular plasmid containing the sequence. It is important to linearize the plasmid DNA used as template. This retrieval method works with ColEI and p15A (pACYC) replicons, but not with pSC101 (Lee et al., 2001).

**SUPPORT
PROTOCOL 2**

**Genetic
Engineering in
Bacteria Using
Homologous
Recombination**

1.16.10

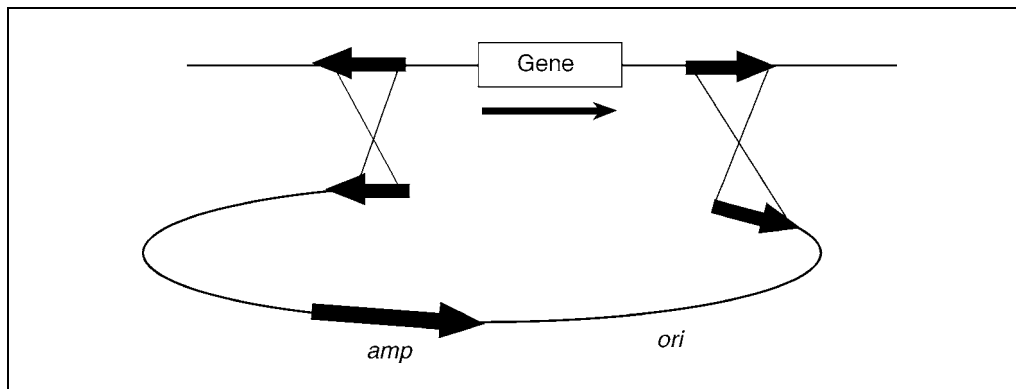


Figure 1.16.2 Cloning genes by gap repair of a plasmid. A linear plasmid with flanking homology to the target at the ends (indicated by dark arrows) is generated by PCR. The plasmid is introduced by electroporation into cells expressing the Red functions, which catalyze recombination of the vector with the target site, resulting in incorporation of the gene onto the plasmid.

If the desired gene is not present on the chromosome, it can be provided as PCR product and introduced into the cells along with the PCR-amplified vector DNA by coelectroporation. Always remember to provide flanking homology so that the plasmid can recombine with the additional PCR product (it is easier to provide the homology on the short product to be cloned rather than on the vector).

Additional Materials (also see *Basic Protocol 1*)

- Plasmid onto which sequence of choice is to be rescued
- Restriction enzyme(s) (*UNIT 3.1*) that do not cut within plasmid region to be amplified
- Synthetic chimeric primers providing homology to sequence flanking gene of choice and to the plasmid sequence to be amplified
- Additional reagents and equipment for PCR (*UNIT 15.1*)

Amplify linear plasmid PCR product with homology to the target

1. Design primers with homology that flanks the desired target.

Drawing a sketch of the plasmid as a gapped circle interacting with the target sequence will help one visualize the recombination reaction (see Fig. 1.16.2), since the plasmid template has the linear ends pointing toward each other. Both of the primers will have ~50-nt of bacterial sequence homology at the 5' ends linked to 3' plasmid sequence. It is not necessary to amplify the entire plasmid. Any portion can be amplified as long as the minimal requirements of a selectable marker and an origin of DNA replication are met. If the DNA to be retrieved is adjacent to an antibiotic resistance gene, only a plasmid replication origin need be amplified; the origin can be used to retrieve both the desired sequence and the nearby drug marker. Avoid having other regions of the plasmid that are homologous to the bacterial chromosome (e.g., lac); these can lead to unwanted rearrangements.

2. To minimize background, digest the plasmid with one or more restriction enzyme(s) that do not cut within the region to be amplified. Amplify the linear plasmid by PCR using the primers (reaction conditions will need to be established empirically).

The amplified product will be a linear gapped plasmid with flanking homology to either side of the allele to be rescued from the chromosome.

Use the least amount of plasmid DNA possible for the PCR template, to minimize the background of false positives. Digestion of the completed PCR reaction mix with DpnI will help remove the template plasmid. Purify the PCR product to remove salt before proceeding.

Transform induced cells with the linear plasmid and select recombinants

3. Introduce the linear plasmid PCR product into the strain (see Basic Protocol 1, steps 4 to 21). If necessary, also add the PCR product to be co-electroporated. Select for the marker on the plasmid, and transform the uninduced cells with the linear plasmid PCR product, to determine the background of intact plasmid present. Purify candidate colonies and screen them with PCR. Isolate the recombinant plasmids and use them to retransform a standard cloning strain such as DH5 α or XL2 Blue, to generate pure clones.

Transformation into a recA mutant host ensures that the newly engineered plasmid does not undergo additional rearrangement.

If the DNA of interest is linked to a drug marker and it is retrieved with a replication origin lacking a drug marker, dilute the electroporation mix into 10 ml LB medium and grow the culture overnight nonselectively. The next day, isolate plasmid DNA and transform into a high-efficiency cloning strain, selecting for the drug resistance of the rescued marker. The transformation should be done with a low concentration of DNA, to minimize uptake of multiple plasmids into the same cell. The advantage of using only the plasmid origin for retrieval is that any possible background of religated vector is eliminated.

Gap repair is less efficient than targeting to the chromosome. The maximal yield achieved in the authors' laboratory is ~ 1000 recombinants/ 10^8 cells. A possible side reaction is joining of the plasmid ends without incorporation of the chromosomal marker. This is caused by small (>5 base) repeats near the linear ends (Zhang et al., 2000); it is likely that the Red functions facilitate the short repeat recombination that generates this background. The short repeat recombination can be reduced or eliminated by designing primers that are free of such repeated sequences.

SCREENING FOR UNSELECTED RECOMBINANTS

When recombinant frequencies approach 1/1000, direct screening can often be used to find recombinant colonies from total viable cells plated out nonselectively on LB. The authors have successfully used the nonradioactive Roche DIG (digoxigenin) system for colony hybridization to detect the recombinant bacterial colonies (L. Thomason, unpub. observ.). For this method to be feasible, the sequence inserted by recombineering must be unique to the recombinant and absent in the starting strain. The authors have used a 21-nt long DIG-labeled oligonucleotide probe to detect insertion of the same sequence. For larger insertions or fusion proteins such as GFP derivatives, a labeled PCR product or gel-purified fragment could also be used as a probe. Both oligo probes and larger DNA fragments could be radioactively labeled with ^{32}P .

The authors have also successfully isolated unselected recombinants when the genetic change confers a slow growth phenotype by simply looking for colonies that grow more slowly than the majority class. Another useful method is the mismatch amplification mutation assay-PCR (MAMA-PCR; Cha et al., 1992; Swaminathan et al., 2001). MAMA-PCR is capable of identifying single base changes by screening colonies.

Additional Materials (also see Basic Protocol 1)

Flanking primers for PCR analysis of mutation of interest (also see UNIT 15.1)
Additional reagents and equipment for colony hybridization (e.g., UNIT 6.6)

1. Introduce the genetic change of interest (see Basic Protocol 1, steps 1 to 18).

Possible changes include small in-frame deletions or a protein tag. For subtle changes, it is helpful to engineer a restriction-site change that can be detected in a PCR product. For detection by hybridization, confirm beforehand that the sequence does not exist in the parental bacterial strain—e.g., using a BLAST search (UNIT 19.3; Altschul et al., 1990).

2. Assuming an expected viable cell count in the transformation mix of $\sim 10^8$, plate the cells nonselectively on rich plates so that the expected number of colonies per plate is ~ 500 .

Six plates should be adequate if the recombination is efficient. More crowded plates will mean that fewer plates must be screened, but if the plates are too crowded it will be more difficult to locate positive clones.

3. Screen for recombinants by performing colony hybridization using established procedures (e.g., UNIT 6.6). If using a nonradioactive labeling method, follow the conditions suggested by the manufacturer.

The required length of oligonucleotide probes will depend on the sensitivity of the system.

The authors have detected positive colonies at frequencies of as low as 5×10^{-2} to 1×10^{-3} .

4. Streak positive candidates to obtain pure clones and retest.

Since the recombination only alters one of several copies of the chromosome existing in a single cell, the colony from that cell will be heterozygous for the allele. Thus, on re-streaking, some fraction of the colonies will not give a positive signal. The signal can again be detected by colony hybridization, or if a new restriction site has been designed into the construct, perform PCR and cut the product with the appropriate restriction enzyme to detect the recombinant (also see UNIT 3.1).

SCREENING FOR UNSELECTED PLASMID RECOMBINANTS

With the extremely high oligonucleotide recombination frequencies obtainable in the absence of mismatch repair (20% to 25% of total viable cells), direct sequencing of unselected plasmid clones can be used to find recombinants. If the oligonucleotide carrying the mutation to be introduced creates a C-C mismatch when paired at the target site, the recombination can be done in wild-type cells containing the defective prophage, since a C-C mismatch is not repaired. Otherwise, a bacterial strain with prophage and mutant for the mismatch repair system can be used. It is helpful if the mutation to be inserted creates a restriction site change that can be monitored by digestion of a PCR amplified fragment covering the region of interest.

Additional Materials (also see Basic Protocol 1)

Cells expressing Red but mutant for host mismatch repair system *or* oligonucleotide that creates C-C mismatch when annealed to target DNA strand: e.g., HME63 and/or HME64 (see Table 1.16.1)
High-efficiency cloning strain (lacking the Red system)

1. Perform Basic Protocol 1, introducing both the plasmid and the oligonucleotide by co-electroporation into cells mutant for the host mismatch repair system, or using an oligo that creates a C-C mismatch when annealed to the target DNA strand.
2. After the outgrowth, dilute the electroporation mix into 10 ml broth containing the appropriate antibiotic for plasmid selection and grow the culture overnight.
3. Isolate plasmid DNA from this culture.
4. Using a low DNA concentration, transform the engineered plasmid into a high efficiency cloning strain (lacking the Red system), selecting for drug resistance. Purify colonies.

Use a low DNA concentration to minimize uptake of more than one molecule/cell.

SUPPORT PROTOCOL 4

Escherichia coli,
Plasmids, and
Bacteriophages

1.16.13

5. Screen to find the mutation, either by direct sequencing or, if there is a restriction site change, by amplifying a fragment with PCR and digesting it with the appropriate enzyme(s).

The authors recommend looking at 25 to 50 colonies. This procedure has been successfully used for plasmid mutagenesis (L.Thomason, unpub. observ.).

REMOVAL OF THE PROPHAGE BY RECOMBINEERING

Once the mutational changes are introduced, the defective λ prophage can be removed if necessary. This is done by another Red-mediated recombination reaction. Alternatively, the prophage can be removed by P1 transduction. In both cases, the desired recombinant can be selected on minimal plates since it will grow in the absence of biotin.

Materials

Oligonucleotide primers for amplifying the bacterial *attB* site:

5' GAGGTACCAGGCGCGGTTTGATC 3'

5' CTCCGGTCTTAATCGACAGCAAC 3'

E. coli K12 strain lacking the prophage (e.g., W3110), but containing the *attB* and biotin (*bio*) genes

M63 minimal glucose plates with and without biotin (see recipe)

Additional reagents and equipment for recombineering (see Basic Protocol 1)

1. Amplify the bacterial *attB* site by PCR using 50 pmol of each of the primers listed above, with an *E. coli* K12 strain lacking the prophage (e.g., W3110) as template (the PCR product is \sim 2.5 kb), and the following program:

1 cycle:	2 min	94°C	(denaturation)
9 cycles:	15 sec	94°C	(denaturation)
	30 sec	65°C	(annealing)
	3 min	68°C	(extension)
19 cycles:	15 sec	94°C	(denaturation)
	30 sec	65°C	(annealing)
	3 min		
	(adding 5 sec/cycle)	68°C	(extension)
1 cycle:	7 min	68°C	(extension)
1 cycle:	indefinite	4°C	(hold)

*The technique used here is "colony PCR" in which the PCR reaction is prepared without allowing any volume for the template DNA; a fresh colony of the *E. coli* K12 strain is then touched with a sterile inoculating loop and mixed into the PCR reaction.*

2. Delete the prophage by recombineering (see Basic Protocol 1, steps 1 to 18). using the *attB* PCR product for recombination.
3. Wash the cells in minimal salts twice and resuspend them in the same medium for plating. Select for the desired recombinant on minimal glucose plates lacking biotin but containing vitamin B1 by incubating 2 hr at 32°C then shifting to 42°C until colonies appear.

*LB has trace amounts of biotin, enough to complement a *bio* mutant for growth; therefore the cells must be washed free of LB for this selection to be successful. Those cells in which the prophage has been deleted will be *bio*⁺, since the PCR product brings in a wild-type (wt) biotin gene, and will grow at 42°C, since the removal of the prophage makes the cells temperature resistant. As a control, confirm that the prophage-containing strain, subjected to the same washes, does not plate on minimal glucose plates lacking biotin, but does plate when biotin is present.*

The 32°C incubation allows time for the recombinant chromosomes to segregate away from those still containing the prophage, which expresses a killing function at high temperature (see Commentary). Streak the recombinant colonies to purify them. As an additional confirmation that the prophage has been deleted, the same PCR primers can be used to confirm the presence of the 2.5-kb band in the purified strain.

RECOMBINEERING WITH AN INTACT λ PROPHAGE

Basic Protocol 1 describes a method for recombineering using a defective lambdoid prophage. It is also possible to recombineer using an intact λ prophage (Court et al., 2003). If the phage of interest has a temperature-sensitive *cI857* repressor, intact *exo*, *beta*, and *gam* genes, and is able to lysogenize the host, it can exist in the prophage state and can be induced by high temperature to express the recombination functions. Such a phage will provide the necessary functions for recombineering; it can also itself serve as a target for such engineering. Since the recombineering efficiency is lower with this method than with the defective prophage, a selection for the recombinants should be applied. To recombineer using an intact phage, construct and confirm a bacterial lysogen (Arber et al., 1983) with the phage of choice. Once the phage exists in the prophage state, it can be induced by a temperature shift, as described below, and PCR product or oligo containing the desired genetic change can be introduced by electroporation.

Additional Materials (also see Basic Protocol 1)

Bacterial lysogen carrying the λ *cI857* bacteriophage of choice as a prophage

Chloroform

1 × TM buffer (APPENDIX 2)

82-mm nitrocellulose filters, sterile

39°C water bath

Additional reagents and equipment for plating λ phage to generate plaques (plaque purification; UNIT 1.11)

1. Grow the host bacterial strain (lysogen) to mid-log phase at 32°C (see Basic Protocol 1, step 6).
2. Induce recombination functions, electroporate cells, and plate to determine viable cell counts (see Basic Protocol 1, steps 7 to 18), except at step 7 of Basic Protocol 1, transfer half the culture to 42°C and shake vigorously for 4 to 5 min at 220 to 225 rpm, instead of 15 min, before chilling both flasks rapidly in an ice-water slurry for ≥ 5 min.

The induction time must be shorter to prevent phage lytic replication and resultant cell killing. The shorter induction time means that lower levels of the recombination functions are produced. The only situation in which the 15-min induction time should be used is when changes are being targeted to the bacteriophage chromosome itself (steps 3b and 4b).

If the mutation of interest is targeted to the bacterial chromosome

- 3a. Select for recombinants by plating an entire electroporation mix on one selective plate (because recombinant levels are reduced by the lower induction time). Use a sterile 82-mm diameter nitrocellulose filter atop a rich (LB) plate and incubate > 3 hr at 30° to 32°C, then transfer the filter to the appropriate drug plate using sterile forceps.

The number of recombinants is generally less than 500 per electroporation mix. Usually, approximately half of the surviving cells will have spontaneously lost the prophage. It is possible to screen for non-lysogens by testing candidate colonies for their ability to plate λ , since the prophage renders the cells immune to phage infection; the cured cells will also be viable at 42°C while those containing the prophage will not (see Commentary).

ALTERNATE PROTOCOL 1

Escherichia coli,
Plasmids, and
Bacteriophages

1.16.15

- 4a. Use PCR (*UNIT 15.1*) and subsequent DNA sequencing (Chapter 7) to confirm the mutation.

If mutations have been targeted to the bacteriophage itself

In this case, the 15-min induction time can be used.

- 3b. Dilute the electroporation mix into 5 ml LB medium and aerate by shaking in a shaking water bath 90 min at 39°C. Add 0.25 ml chloroform to complete cell lysis and release the phage particles. Dilute the lysate and plate for single plaques (see *UNIT 1.11*).
- 4b. Plaque-purify the positive candidates (*UNIT 1.11*). Resuspend a plaque in 50 µl of sterile water and use 20 µl of this suspension as template for a PCR reaction (*UNIT 15.1*; reduce the amount of water accordingly). Resuspend another plaque from the same plate in 1 mM of 1× TM buffer to grow a stock. Reconfirm the mutation after growing the stock.

ALTERNATE PROTOCOL 2

TARGETING AN INFECTING PHAGE λ WITH THE DEFECTIVE PROPHAGE STRAINS

Sometimes it may be useful to target genetic changes to bacteriophage λ derivatives. A strain carrying the Red system on the prophage or a plasmid can be infected with a phage and recombineering can then be targeted to the incoming phage chromosome. A procedure for this method follows. Ideally the construction should be designed so that the plaque morphology of the recombinant phage will differ from that of the parent. For example, a PCR product able to both introduce the mutation of interest and correct a known mutation (such as an amber or temperature-sensitive allele) can be recombined onto a phage containing the known mutation. Select for correction of the known mutation and screen among these recombinants for the mutation of interest. If no selection exists, plaque hybridization can sometimes be used to identify recombinant phages. While the authors have only tested this method for phage λ , theoretically it may be possible to introduce genetic changes onto the chromosome of any phage able to propagate in the defective prophage host (Oppenheim et al., 2004).

Additional Materials (also see Basic Protocol 1)

- 10% (w/v) maltose stock solution, filter sterilized
- Tris-magnesium (TM) buffer: 10 mM Tris·Cl, pH 7.4 (*APPENDIX 2*)/10 mM MgSO₄
- High-titer lysate of the bacteriophage to be engineered
- PCR product with desired sequence changes and flanking homology to the target on the phage chromosome
- Lambda plates and lambda top agar (see *UNIT 1.1*, adjust NaCl to 5 g per liter in both the plates and the top agar)
- Chloroform
- Appropriate bacterial indicator strain
- Additional reagents and equipment for working with λ bacteriophages (*UNITS 1.9-1.13*)

Grow host strain

1. Grow strain with defective prophage to mid-log phase at 32°C (see Basic Protocol 1, step 6), except supplement the LB medium with 0.4% maltose (added from 10% w/v maltose stock solution).

Maltose induces the phage receptor on the bacterial cell surface, ensuring efficient adsorption of the phage to the host cells.

Adsorb bacteriophage

2. Harvest the cells by centrifuging 7 min at $4600 \times g$, 4°C . Resuspend pellet in 1 ml Tris-magnesium (TM) buffer. Infect the cells with the phage to be engineered at a multiplicity of 1 to 3 phages per cell.

The cells will be at $\sim 1 \times 10^8/\text{ml}$ before concentration. UNITS 1.9-1.13 contain protocols for working with λ bacteriophages.

3. Let the phage adsorb to the cells for 15 min at room temperature.

Phages other than λ may require different adsorption conditions.

Induce recombination functions and phage DNA entry; electroporate cells

4. Transfer the infected cells to 5 ml of 42°C LB medium and shake vigorously for 15 min. At end of incubation, chill rapidly on ice.

This serves to both induce the Red functions and allow phage infection to proceed.

5. Make electrocompetent cells and introduce transforming DNA (i.e., introduce PCR product with mutation of interest; see Basic Protocol 1, steps 8 to 14).

Outgrow infected cells to allow phage production

6. Dilute electroporation mix into 5 ml warm LB medium and shake vigorously for 90 min at 42°C to allow the phage to complete a lytic cycle. Add 0.25 ml chloroform to completely lyse infected cells. Plate phage on lambda plates with lambda top agar (see UNIT 1.11) using the appropriate bacterial indicator strain as a host; apply a selection if possible; or plate non-selectively and screen for the desired mutation with plaque hybridization.

Amber mutations can be specifically selected (Oppenheim et al., 2004).

REAGENTS AND SOLUTIONS

Use Milli-Q purified water or equivalent in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

M63 minimal glucose plates with biotin

Per liter:

3 g KH_2PO_4

7 g K_2HPO_4

2 g $(\text{NH}_4)_2\text{SO}_4$

0.5 ml of 1 mg/ml FeSO_4

0.2% (v/v) glucose

0.001% (w/v) biotin (omit for control)

1 ml 1% (w/v) vitamin B1 (thiamine)

1 ml 1 M MgSO_4

15 g agar

Pour 40 ml of the agar-containing medium per plate.

Store up to several months at 4°C .

M63 minimal glycerol plates with sucrose

Per liter:

3 g KH_2PO_4

7 g K_2HPO_4

2 g $(\text{NH}_4)_2\text{SO}_4$

0.5 ml of 1 mg/ml FeSO_4

0.2% (v/v) glycerol

continued

***Escherichia coli,
Plasmids, and
Bacteriophages***

1.16.17

5% (w/v) sucrose
 0.001% (w/v) biotin
 1 ml 1% (w/v) vitamin B1 (thiamine)
 1 ml 1 M MgSO₄
 15 g agar
 Pour 40 ml of the agar-containing medium per plate.
 Store up to several months at 4°C.

COMMENTARY

Background Information

Bacteriophage λ encodes three genes important for recombineering. The *exo* and *bet* genes, respectively, encode a 5' to 3' double-strand exonuclease, Exo, and a single strand annealing protein, Beta, which together can recombine a double-stranded PCR product with short flanking homologies into the desired genetic target. The λ *gam* gene encodes a protein, Gam, that inhibits the RecBCD enzyme, which will otherwise degrade linear DNA introduced into the bacterial cell. Only the Beta single-strand annealing function is required to recombine single-stranded oligos containing the desired alterations. A cryptic lambdoid prophage, *rac*, is present in some strains of *E. coli*, and encodes RecE and RecT functions that are analogous to Exo and Beta, respectively. Unlike some other in vivo genetic engineering methods (Russell et al., 1989), recombineering does not require the host *recA* function. A strain mutant for *recA* provides more controllable recombination, since the strain is recombination proficient only when the phage functions are induced. A more detailed discussion of the molecular mechanism of recombineering can be found in Court et al. (2002).

Recently the Red genes have been moved to several plasmids (Table 1.16.3; S. Datta, manuscript in preparation) having different DNA replication origins. Here, the essential control elements of the prophage system are retained and recombination functions are

induced by a temperature shift as in Basic Protocol 1. The plasmids are especially useful when one wants to create a mutation in a particular bacterial strain, rather than create the mutations in the prophage-containing strains and subsequently move them into a different background.

Critical Parameters and Troubleshooting

Induction times

When inducing the recombination functions from the defective prophage, the proper induction time is essential. Longer induction times will cause decreased cell viability, since the prophage Kil function is also induced by the temperature shift. The Gam function, necessary for efficient transformation of ds-DNA, is also toxic to the cells (Sergueev et al., 2001).

Storage of induced cells

It is also important to use the induced cells promptly, since the induced phage functions will decay over time, especially at 32°C. Although one pioneering laboratory member reports that induced cells can be successfully stored on ice for several hours before electroporation, the authors do not recommend this procedure. As detailed in the Basic Protocol 1, competent cells may be frozen in 15% glycerol, although their recombineering efficiency has been less than that of the freshly prepared cells.

Table 1.16.3 Plasmids Containing the Red System Under *cl857* Control

Plasmid designation	Plasmid origin	Approximate copy number/cell
pSIM5(Cm ^R) pSIM6(Amp ^R)	pSC101 <i>ts</i>	16
pSIM7(Cm ^R) pSIM8(Amp ^R)	pBBR1	30-40
pSIM9(Cm ^R)	pRK2 <i>ts</i>	20-40

Amplification of drug cassettes

The authors strongly recommend that each drug cassette be amplified with only one standard set of primers. Confusion has arisen in the authors' laboratory when individuals have designed their own primers for the drug cassettes. Remember that it is best to have a promoter, the open reading frame, and a transcriptional terminator on each cassette. Standard priming sequences for the commonly used cassettes are listed in Table 1.16.2.

The *cat-sacB* element

The *cat-sacB* DNA segment described in Support Protocol 1 is amplified from plasmid pEL04, previously known as both pKO4 and *pcat-sacB*, or from the strain DY441, using colony PCR. If a different antibiotic resistance gene is required, the *cat* gene on *pcat-sacB* can be replaced with another drug marker. Use 50 bases of homology upstream of the 5' *cat* primer and downstream of the 3' *cat* primer to cross in the other marker. The sequence of pEL04 is available from the Court Laboratory (court@ncifcrf.gov). Since the *cat-sacB* PCR product is 3264 bp, it is not always easy to amplify. The authors have used the PCR conditions detailed above, but other conditions and any other high-fidelity enzyme should also work. This is really an empirical problem, and, in the opinion of the authors, the biggest factors, if sufficient extension times and appropriate temperatures are used, are the PCR machine itself and the polymerase. An MJ Research PTC-100 machine is used in the authors' laboratory. Adjusting the Mg²⁺ concentration may also improve yield. There has been some confusion regarding the primers for amplification of the *cat-sacB* cassette. The authors have now settled on an optimal primer pair for amplification of this DNA, and strongly recommend that only this set of primers be used. Please note that both primer sequences are different than those previously given in this protocol.

Detecting recombinants

If one has engineered a drug marker into the DNA, drug resistance will serve as the selection. The two-step *cat-sacB* replacement allows precise modification without leaving an antibiotic-resistance cassette or other marker behind. The authors have also had success using colony hybridization with a labeled oligo probe, although a unique sequence hybridizing only to the probe must be inserted (or a unique sequence must be created by a deletion) for this method to be an option.

Mismatch repair minus conditions for oligonucleotide recombination

The authors have recently found (Costantino and Court, 2003) that extremely high levels of recombinants can be obtained when recombineering with oligos if the bacterial strain is mutant for the host methyl-directed mismatch repair system (the *mut HLS* system). Under these conditions ~20% to 25% of the total viable cells are recombinant. Strains with the prophage and mutations in the mismatch repair system are available from the Court laboratory: these are HME63 and HME64 and contain a *mutS* <> *amp* and an *uvrD* <> *kan* mutation, respectively. Either mutation will eliminate mismatch repair; the latter strain can be used for engineering onto plasmids expressing ampicillin resistance. The same elevated level of recombination is obtained in mismatch repair proficient strains if the incoming oligo creates a C-C mismatch when annealed to the target DNA, since C-C mispairs are not subject to mismatch repair.

Counter-selections

The *cat-sacB* cassette is a counter-selection; it can be selected either for (by chloramphenicol resistance) or against (sucrose sensitivity). Several such counter-selections are available and may prove useful for one's particular situation; tetracycline resistance is another example (Bochner et al., 1980; Maloy and Nunn, 1981). If removing a bacteriostatic drug marker (such as tetracycline), enrichment for recombinants is possible using ampicillin. Nonrecombinant bacteria will still express resistance to the *tet* marker encoded on the chromosome; when the electroporated culture is propagated in the presence of both tetracycline and ampicillin, only nonrecombinant cells will grow; these will be killed by the ampicillin (Murphy et al., 2000). Recombinant cells will not grow but will not be killed. 2-deoxy D-galactose is a toxic analog of galactose on which only mutants of galactokinase, the product of the *galK* gene, will grow (Alper and Ames, 1975). Other counter-selections may be devised.

In vivo assembly with overlapping oligos

Complementary oligos will anneal in vivo when introduced by electroporation into cells expressing Exo and Beta protein (Yu et al., 2003). Multiple overlapping oligos can be used to build moderately sized DNA products (100 to 150 bp) in an in vivo reaction similar to PCR assembly (Stemmer et al., 1995). By adding

flanking homologies to chromosomal targets, these overlapping oligos are restored to linear dsDNA *in vivo* and recombined into the target.

Anticipated Results

As discussed above, ideally, transformation with dsDNA and ssDNA can approach frequencies of 0.1%, and 25%, respectively. Because of this high efficiency, there should rarely if ever be experiments where no recombinants are found. However, in practice, recombineering frequency may be partially context dependent; certain areas of the chromosome appear “hotter” for recombination than others (Ellis et al., 2001). If recombinants are not obtained, check the construct design and redesign primers if necessary, as this is one reason for failure. If the recombineering reaction does not work, the authors strongly recommend doing a control experiment using known strains and oligos or PCR products, as described in Yu et al. (2000), Ellis et al. (2001), and Costantino and Court (2003) to verify that the predicted number of recombinants are obtained in these control reactions. It has been the authors’ experience that most recombineering failures occur because the protocol is not executed carefully or properly.

Recombineering onto plasmids

Genetic alterations can also be targeted to intact circular plasmids. If the desired change can be engineered with an oligonucleotide, the mismatch repair-deficient strains can be used and unselected clones screened for the mutation. Otherwise, it is helpful to have a selection for the desired recombinant. If one has a selection, the plasmid may either be already resident in the bacterial cell or may be introduced by co-electroporation. The former method is less optimal, because expression of the λ Gam function can lead to formation of linear multimers of some plasmid replicons (Cohen and Clark, 1986). Hence, co-electroporation is preferable. If introducing an unselected mutation, co-electroporation is essential, since it is unlikely that more than one resident plasmid in a cell will become modified. In either case, the recombinant plasmid clones should be purified and the DNA isolated, characterized, and used to transform a new host.

Engineering onto BACs

Red recombineering has also been optimized for engineering bacterial artificial chromosomes (Copeland et al., 2001; Lee et al., 2001; Swaminathan and Sharan, 2004), which are able to accommodate hundreds of kilobases

of foreign DNA. DY380 is the bacterial host of choice for use with BACs; it is a DH10B derivative containing the defective prophage. The plasmids listed in Table 1.16.3 may also be useful in manipulating BACs.

Mutating essential genes

Mutation of an essential gene usually results in cell death. Such a mutation can be recovered in a diploid state; however, the cell will maintain a wild-type copy of the gene as well as a mutant version. This diploid state can usually be detected by PCR analysis.

Moving the prophage to a different background

If a different host is required for recombineering, the defective prophage can be moved by P1 transduction (Miller, 1972; Yu et al., 2000). Grow a P1 lysate on DY329 and transduce the strain of choice. Select for tetracycline resistance and screen for temperature sensitivity or an inability to grow in the absence of biotin. Check the genotype of the strain first and supplement the minimal plates to complement any additional auxotrophies when performing the biotin screen.

Time Considerations

The basic recombineering protocol can be executed in one day, with bacterial cultures started the evening before. The recombinants may take several days to grow on the selective plates, and it may take another day or so to confirm the recombinants.

Literature Cited

- Alper, M.D. and Ames, B.N. 1975. Positive selection of mutants with deletions of the *gal-chl* region of the *Salmonella* chromosome as a screening procedure for mutagens that cause deletions. *J. Bacteriol.* 121:259-266.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Arber, W., Enquist, L., Hohn, B., Murray, N.E., and Murray, K. 1983. Experimental methods for use with lambda. In *Lambda II* (R. Hendrix, J. Roberts, F. Stahl, and R. Weisberg, eds.) pp. 433-471. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Blomfield, I.C., Vaughn, V., Rest, R.F., and Eisenstein, B.I. 1991. Allelic exchange in *Escherichia coli* using the *Bacillus subtilis sacB* gene and a temperature-sensitive pSC101 replicon. *Mol. Microbiol.* 5:1447-1457.
- Bochner, B.R., Huang, H.C., Schieven, G.L., and Ames, B.N. 1980. Positive selection for loss of tetracycline resistance. *J. Bacteriol.* 143:926-933.

- Cha, R.S., Zarbl, H., Keohavong, P., and Thilly, W.G. 1992. Mismatch amplification mutation assay (MAMA): Application to the c-H-ras gene. *PCR Methods Appl.* 2:14-20.
- Cohen, A. and Clark, A. J. 1986. Synthesis of linear plasmid multimers in *Escherichia coli* K-12. *J. Bacteriol.* 167:327-35.
- Copeland, N.G., Jenkins, N.A., and Court, D.L. 2001. Recombineering: A powerful new tool for mouse functional genomics. *Nat. Rev. Genet.* 2:769-779.
- Costantino, N. and Court, D.L. 2003. Enhanced levels of lambda Red-mediated recombinants in mismatch repair mutants. *Proc. Natl. Acad. Sci. U.S.A.* 100:15748-15753.
- Court, D.L., Sawitzke, J.A., and Thomason, L.C. 2002. Genetic engineering using homologous recombination. *Annu. Rev. Genet.* 36:361-388.
- Court, D.L., Swaminathan, S., Yu, D., Wilson, H., Baker, T., Bubunenko, M., Sawitzke, J., and Sharan, S.K. 2003. Mini-lambda: A tractable system for chromosome and BAC engineering. *Gene* 315:63-69.
- Datsenko, K.A. and Wanner, B.L. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97:6640-6645.
- Ellis, H.M., Yu, D., DiTizio, T., and Court, D.L. 2001. High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 98:6742-6746.
- Gay, P., Le Coq, D., Steinmetz, M., Berkelman, T., and Kado, C.I. 1985. Positive selection procedure for entrapment of insertion sequence elements in Gram-negative bacteria. *J. Bacteriol.* 164:918-921.
- Lee, E.C., Yu, D., Martinez de Velasco, J., Tesarollo, L., Swing, D.A., Court, D.L., Jenkins, N.A., and Copeland, N.G. 2001. A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* 73:56-65.
- Maloy, S.R. and Nunn, W.D. 1981. Selection for loss of tetracycline resistance by *Escherichia coli*. *J. Bacteriol.* 145:1110-1111.
- Miller, J.M. 1972. Generalized transduction: Use of P1 in strain construction. In *Experiments in Molecular Genetics* (J.M. Miller, ed.) pp. 201-205. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Murphy, K.C., Campellone, K.G., and Poteete, A.R. 2000. PCR-mediated gene replacement in *Escherichia coli*. *Gene* 246:321-330.
- Oppenheim A.B., Rattray, A.J., Bubunenko, M., Thomason, L.C., and Court, D.L. 2004. *In vivo* recombineering of bacteriophage λ by PCR fragments and single-strand oligonucleotides. *Virology* 319:185-189.
- Russell, C.B., Thaler, D.S., and Dahlquist, F.W. 1989. Chromosomal transformation of *Escherichia coli* recD strains with linearized plasmids. *J. Bacteriol.* 171:2609-2613.
- Sergueev, K., Yu, D., Austin, S., and Court, D. 2001. Cell toxicity caused by products of the P_L operon of bacteriophage lambda. *Gene* 272:227-235.
- Stemmer, W.P., Cramer, A., Ha, K.D., Brennan, T.M., and Heyneker, H.L. 1995. Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* 164:49-53.
- Swaminathan, S. and Sharan, S.K. 2004. Bacterial artificial chromosome engineering. *Methods Mol. Biol.* 256:89-106.
- Swaminathan, S., Ellis, H.M., Waters, L.S., Yu, D., Lee, E.C., Court, D.L., and Sharan, S.K. 2001. Rapid engineering of bacterial artificial chromosomes using oligonucleotides. *Genesis* 29:14-21.
- Thomason, L.C., Myers, R.S., Oppenheim, A., Costantino, N., Sawitzke, J.A., Datta, S., Bubunenko, M., and Court, D.L. 2005. Recombineering in prokaryotes. In *Phages: Their Role in Bacterial Pathogenesis and Biotechnology*. pp. 383-399. ASM Press, Herndon, Virginia.
- Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G., and Court, D.L. 2000. An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 97:5978-5983.
- Yu, D., Sawitzke, J.A., Ellis, H., and Court D.L. 2003. Recombineering with overlapping single-stranded DNA oligonucleotides: Testing a recombination intermediate. *Proc. Natl. Acad. Sci. U.S.A.* 100:7207-7212.
- Zhang, Y., Muylers, J.P., Testa, G., and Stewart, A.F. 2000. DNA cloning by homologous recombination in *Escherichia coli*. *Nature Biotechnol.* 12:1314-1317.

Contributed by Lynn Thomason,
Donald L. Court, Mikail Bubunenko,
Nina Costantino, Helen Wilson,
Simanti Datta, and Amos Oppenheim
National Cancer Institute at Frederick
Frederick, Maryland

CHAPTER 2

Preparation and Analysis of DNA

INTRODUCTION

The ability to prepare and isolate pure DNA from a variety of sources is an important step in many molecular biology protocols. Indeed, the isolation of genomic, plasmid, or DNA fragments from restriction digests and polymerase chain reaction (PCR) products has become a common everyday practice in almost every laboratory. This chapter therefore begins with protocols for purification of genomic DNA from bacteria, plant cells, and mammalian cells (*UNITS 2.1-2.4*). These protocols consist of two parts: a technique to lyse the cells gently and solubilize the DNA, followed by one of several basic enzymatic or chemical methods to remove contaminating proteins, RNA, and other macromolecules. The basic approaches described here are generally applicable to a wide variety of starting materials. A brief collection of general protocols for further purifying and concentrating nucleic acids is also included.

The last decade has shown a dramatic departure from the use of traditional DNA purification methods outlined in *UNITS 2.2-2.4*, with a concomitant increase in the use of purpose-specific kits for the isolation and purification of DNA. For example, kits for purification of DNA using pre-made anion-exchange columns packaged with all necessary solutions to lyse the cells and solubilize the DNA are available from many molecular biology companies. A variety of kits based on binding of DNA to glass beads are also available. The uses of both types of kits are discussed in *UNIT 2.1B*.

The use of kits has two main advantages: it saves time and makes the process of DNA purification a relatively easy and straightforward process. The purification of DNA by anion-exchange chromatography (*UNIT 2.1B*) is readily becoming the accepted standard for quick and efficient large-scale (more than 100 μg of DNA) production of DNA from bacteria, mammalian tissue, and plant tissue. In most cases, the cell lysis and solubilization of DNA is relatively unchanged compared to traditional methods, with anion-exchange chromatography columns having replaced labor and time-intensive techniques such as cesium chloride centrifugation for the isolation of relatively pure DNA. Purification kits are usually available in several sizes and configurations, allowing the researcher to have variability concerning the processing and purification of their DNA.

A variety of techniques exist for the isolation of small amounts of plasmid DNA from minipreps and for DNA fragments from restriction digests/PCR products from agarose gels (with removal of unincorporated nucleoside triphosphates, reaction products, and small oligonucleotides from PCR reactions). These are detailed in *UNITS 2.1A, 2.1B, 2.6 & 2.7*. Likewise, kits are available from several molecular biology companies, usually based on silica-gel technology, for each of these applications (*UNIT 2.1B*). As with large-scale DNA isolation and purification, these kits provide a quick and efficient means to recover purified DNA that can be used for subsequent cloning or other modifications.

Virtually all protocols in molecular biology require, at some point, fractionation of nucleic acids. Chromatographic techniques are appropriate for some applications and may be used for separation of plasmid from genomic DNA as well as separation of genomic DNA from debris in a cell lysate (*UNIT 2.1B*). Gel electrophoresis, however, has much greater

resolution than alternative methods and is generally the fractionation method of choice. Gel electrophoretic separations can be either analytical or preparative, and can involve fragments with molecular weights ranging from less than 1000 Daltons to more than 10^8 . A variety of electrophoretic systems have been developed to accommodate such a large range of applications.

In general, the use of electrophoresis to separate nucleic acids is simpler than its application to resolve proteins. Nucleic acids are uniformly negatively charged and, for double-stranded DNA, reasonably free of complicating structural effects that affect mobility. A variety of important variables affect migration of nucleic acids on gels. These include the conformation of the nucleic acid, the pore size of the gel, the voltage gradient applied, and the salt concentration of the buffer. The most basic of these variables is the pore size of the gel, which dictates the size of the fragments that can be resolved. In practice, this means that larger-pore agarose gels are used to resolve fragments >500 to 1000 bp (*UNITS 2.5A & 2.6*) and smaller-pore acrylamide or sieving agarose gels (*UNIT 2.7*) are used for fragments <1000 bp. A protocol for resolution of very large pieces of DNA may also be resolved on agarose gels using pulsed-field gel electrophoresis (*UNIT 2.5B*). Finally, the powerful analytical technique of capillary electrophoresis of DNA (*UNIT 2.8*) may be used to assess the purity of synthetic oligonucleotides, analyze quantitative PCR results, and compare DNA fragment lengths from restriction fragment length polymorphism (RFLP) and variable number of tandem repeat (VNTR) analyses.

Frequently it is desirable to identify an individual fragment in a complex mixture that has been resolved by gel electrophoresis. This is accomplished by a technique termed Southern blotting, in which the fragments are transferred from the gel to a nylon or nitrocellulose membrane and the fragment of interest is identified by hybridization with a labeled nucleic acid probe. Section IV of this chapter gives a complete review of methods and materials required for immobilization of fractionated DNA (*UNIT 2.9*) and associated hybridization techniques (*UNIT 2.10*). These methods have greatly contributed to the mapping and identification of single and multicopy sequences in complex genomes, and facilitated the initial eukaryotic cloning experiments.

Other commonly encountered applications of gel electrophoresis include resolution of single-stranded RNA or DNA. Polyacrylamide gels containing high concentrations of urea as a denaturant provide a very powerful system for resolution of short (<500-nucleotide) fragments of single-stranded DNA or RNA. Such gels can resolve fragments differing by only a single nucleotide in length, and are central to all protocols for DNA sequencing (see *UNIT 7.6*). Such gels are used for other applications requiring resolution of single-stranded fragments, particularly including the techniques for analyzing mRNA structure by S1 analysis (*UNIT 4.6*), ribonuclease protection (*UNIT 4.7*), or primer extension (*UNIT 4.8*). Denaturing polyacrylamide gels are also useful for preparative applications, such as small-scale purification of radioactive single-stranded probes and large-scale purification of synthetic oligonucleotides (*UNIT 2.12*).

Resolution of relatively large single-stranded fragments (>500 nucleotides) can be accomplished using denaturing agarose gels. This is of particular importance to the analysis of mRNA populations by northern blotting and hybridization. A protocol for use of agarose gels containing formaldehyde in resolution of single-stranded RNA is presented in *UNIT 4.9*. The use of denaturing alkaline agarose gels for purification of labeled single-stranded DNA probes is described in *UNIT 4.6*.

Gels and Electric Circuits

Gel electrophoresis units are almost always simple electric circuits and can be understood using two simple equations. Ohm's law, $V = IR$, states that the electric field, V (measured in volts), is proportional to current, I (measured in milliamps), times resistance, R (measured in ohms). When a given amount of voltage is applied to a simple circuit, a constant amount of current flows through all the elements and the decrease in the total applied voltage that occurs across any element is a direct consequence of its resistance. For a segment of a gel apparatus, resistance is inversely proportional to both the cross-sectional area and the ionic strength of the buffer. Usually the gel itself provides nearly all of the resistance in the circuit, and the voltage applied to the gel will be essentially the same as the total voltage applied to the circuit. For a given current, decreasing either the thickness of the gel (and any overlying buffer) or the ionic strength of the buffer will increase resistance and, consequently, increase the voltage gradient across the gel and the electrophoretic mobility of the sample.

A practical upper limit to the voltage is usually set by the ability of the gel apparatus to dissipate heat. A second useful equation, $P = I^2R$, states that the power produced by the system, P (measured in watts), is proportional to the resistance times the square of the current. The power produced is manifested as heat, and any gel apparatus can dissipate only a particular amount of power without increasing the temperature of the gel. Above this point small increases in voltage can cause significant and potentially disastrous increases in temperature of the gel. It is very important to know how much power a particular gel apparatus can easily dissipate and to carefully monitor the temperature of gels run above that level.

Two practical examples illustrate applications of the two equations. The first involves the fact that the resistance of acrylamide gels increases somewhat during a run as ions related to polymerization are electrophoresed out of the gel. If such a gel is run at constant current, the voltage will increase with time and significant increases in power can occur. If an acrylamide gel is being run at high voltage, the power supply should be set to deliver constant power. The second situation is the case where there is a limitation in number of power supplies, but not gel apparatus. A direct application of the first equation shows that the fraction of total voltage applied to each of two gels hooked up in series (one after another) will be proportional to the fraction of total resistance the gel contributes to the circuit. Two identical gels will each get 50% of the total voltage and power indicated on the power supply.

Finally, it should be noted that some electrophoretic systems employ lethally high voltages, and almost all are potentially hazardous. It is very important to use an adequately shielded apparatus, an appropriately grounded and regulated power supply, and most importantly, common sense when carrying out electrophoresis experiments.

David D. Moore and Dennis Dowhan

MANIPULATION OF DNA

SECTION I

This section begins with a protocol describing basic techniques for purifying and concentrating DNA samples by extraction and precipitation (*UNIT 2.1A*) and anion-exchange chromatography (*UNIT 2.1B*), followed by similar protocols for purifying DNA from mammalian tissue (*UNIT 2.2*), plant tissue (*UNIT 2.3*), and bacteria (*UNIT 2.4*). Various properties and useful measurements of DNA are presented in *APPENDIX 1* (see Tables A.1B.1 and A.1B.2 and Fig. A.1B.1).

IMPORTANT NOTE: The smallest amount of contamination of DNA preparations by recombinant phages or plasmids can be disastrous. Many person-years have been wasted reisolating previously cloned sequences that contaminated preparations of DNA used to create recombinant DNA libraries (see Hall, 1987, for an account of such a mistake) and many researchers have been embarrassed to find that the “extra” genes they found on their Southern blots were actually contaminating plasmid DNA. All materials used for preparation of plasmid or phage DNA should be kept separate from those used for preparation of genomic DNA, and disposable items should be used wherever possible. Particular care should be taken to avoid contamination of commonly used rotors.

Purification and Concentration of DNA from Aqueous Solutions

UNIT 2.1A

This unit presents basic procedures for manipulating solutions of single- or double-stranded DNA through purification and concentration steps. These techniques are useful when proteins or solute molecules need to be removed from aqueous solutions, or when DNA solutions need to be concentrated for reasons of convenience. The Basic Protocol, using phenol extraction and ethanol precipitation, is appropriate for the purification of DNA from small volumes (<0.4 ml) at concentrations ≤ 1 mg/ml. Isopropanol may also be used to precipitate DNA, as described in Alternate Protocol 1. Three support protocols outline methods to buffer the phenol used in extractions (see Support Protocol 1), concentrate DNA using butanol (see Support Protocol 2), and extract residual organic solvents with ether (see Support Protocol 3). An alternative to these methods is nucleic acid purification using commercially available silica membrane spin columns, presented in Alternate Protocol 2. These protocols may also be used for purifying RNA.

The final two alternate protocols provide modifications to the basic protocol that are used for concentrating RNA and extracting and precipitating DNA from larger volumes and from dilute solutions (see Alternate Protocol 3), and for removing low-molecular-weight oligonucleotides and triphosphates (see Alternate Protocol 4).

PHENOL EXTRACTION AND ETHANOL PRECIPITATION OF DNA

BASIC PROTOCOL

This protocol describes the most commonly used method of purifying and concentrating DNA preparations. The DNA solution is first extracted with a phenol/chloroform/isoamyl alcohol mixture to remove protein contaminants, and then precipitated with 100% ethanol. The DNA is pelleted after the precipitation step, washed with 70% ethanol to remove salts and small organic molecules, and resuspended in buffer at a concentration suitable for further experimentation.

Preparation and Analysis of DNA

2.1.1

Materials

≤1 mg/ml DNA to be purified
25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol (made with buffered phenol; see Support Protocol 1)
3 M sodium acetate, pH 5.2 (APPENDIX 2)
100% ethanol, ice cold
70% ethanol, room temperature
TE buffer, pH 8.0 (APPENDIX 2)
Speedvac evaporator (Savant)

1. Add an equal volume of phenol/chloroform/isoamyl alcohol to the DNA solution to be purified in a 1.5-ml microcentrifuge tube.

DNA solutions containing ≤0.5 M monovalent cations can be used. Extracting volumes ≤100 μl is difficult; small volumes should be diluted to obtain a volume that is easy to work with.

High salt concentrations can cause the inversion of the aqueous and organic phases. If this happens, the organic phase can be identified by its yellow color.

2. Vortex vigorously for 10 sec and microcentrifuge 15 sec at maximum speed, room temperature.

Phases should be well separated. If DNA solution is viscous or contains a large amount of protein, it should be microcentrifuged longer (1 to 2 min).

3. Carefully remove the top (aqueous) phase containing the DNA using a 200-μl pipettor and transfer to a new tube. If a white precipitate is present at the aqueous/organic interface, repeat steps 1 to 3.

If starting with a small amount of DNA (<1 μg), recovery can be improved by reextracting the organic phase with 100 μl TE buffer, pH 8.0. This aqueous phase can be pooled with that from the first extraction.

4. Add $\frac{1}{10}$ vol of 3 M sodium acetate, pH 5.2, to the solution of DNA. Mix by vortexing briefly or by flicking the tube several times with a finger.

If the solution contained a high concentration of NaCl or sodium acetate (0.3 to 0.5 M) prior to the phenol extraction step, then no additional salt should be added. It is advisable to make appropriate dilutions to keep NaCl and sodium acetate concentrations below 0.5 M.

For high concentrations of DNA (>50 to 100 μg/ml), precipitation is essentially instantaneous at room temperature. If ethanol precipitation is not desirable, residual organic solvents can be removed by ether extraction (see Support Protocol 3). In this case, no salt should be added.

To prevent carryover of residual phenol, the aqueous phase can be reextracted with 24:1 (v/v) chloroform/isoamyl alcohol. However, this should not be necessary if the final pellet is washed well with 70% ethanol, or if an additional ethanol precipitation step is included.

5. Add 2 to 2.5 vol (calculated after salt addition) of ice-cold 100% ethanol. Mix by vortexing and place in crushed dry ice for 5 min or longer.

This precipitation step can also be done in a -70°C freezer for 15 min or longer, or in a -20°C freezer for at least 30 min. A slurry of dry ice and ethanol may also be used, but tube labels are less often lost when crushed dry ice is used.

6. Microcentrifuge 5 min at maximum speed and remove the supernatant.

For large pellets the supernatant can simply be poured off. For small pellets (<1 μg), aspirate off the ethanol supernatant with a pipetting device such as a Pasteur pipet or mechanical pipettor. This is best accomplished by drawing off liquid from the side of the

tube opposite that against which the DNA precipitate was pelleted. Start at the top and move downward as the liquid level drops.

7. Add 1 ml room-temperature 70% ethanol. Invert the tube several times and microcentrifuge as in step 6.

If the DNA molecules being precipitated are very small (<200 bases), use 95% ethanol at this step.

8. Remove the supernatant as in step 6. Dry the pellet in a desiccator under vacuum or in a Speedvac evaporator.

The DNA pellet will not stick well to the walls of the tube after the 70% ethanol wash and care must be taken to avoid aspirating the pellet out of the tube.

9. Dissolve the dry pellet in an appropriate volume of water if it is going to be used for further enzymatic manipulations requiring specific buffers. Dissolve in TE buffer, pH 8.0, if it is going to be stored indefinitely.

DNA pellets will not dissolve well in high-salt buffers. To facilitate resuspension, the DNA concentration of the final solution should be kept at <1 mg/ml.

If DNA is resuspended in a volume of TE buffer or water to yield a DNA concentration of <1 mg/ml, small quantities (<25 µg) of precipitated plasmids or restriction fragments should dissolve quickly upon gentle vortexing or flicking of the tube. However, larger quantities of DNA may require vortexing and brief heating (5 min at 65°C) to resuspend. High-molecular-weight genomic DNA may require one to several days to dissolve and should be shaken gently (not vortexed) to avoid shearing, particularly if it is to be used for cosmid cloning or other applications requiring high-molecular-weight DNA. Gentle shaking on a rotating platform or a rocking apparatus is recommended.

PRECIPITATION OF DNA USING ISOPROPANOL

Equal volumes of isopropanol and DNA solution are used in precipitation. Note that the isopropanol volume is half that of the given volume of ethanol in precipitations. This allows precipitation from a large starting volume (e.g., 0.7 ml) in a single microcentrifuge tube. Isopropanol is less volatile than ethanol and takes longer to remove by evaporation. Some salts are less soluble in isopropanol (compared to ethanol) and will be precipitated along with nucleic acids. Extra washings may be necessary to eliminate these contaminating salts.

PREPARATION OF BUFFERED PHENOL AND PHENOL/CHLOROFORM/ISOAMYL ALCOHOL

For some purposes, fresh liquefied phenol (88% phenol) can be used without further purification. However, for purification of DNA prior to cloning and other sensitive applications, phenol must be redistilled before use, because oxidation products of phenol can damage and introduce breaks into nucleic acid chains. Redistilled phenol for use in nucleic acid purification is commercially available, but must be buffered before use. Appropriately buffered phenol is also commercially available, but is somewhat more expensive and should not be stored for long periods of time (e.g., >6 months).

CAUTION: Phenol can cause severe burns to skin and damage clothing. Gloves, safety glasses, and a lab coat should be worn whenever working with phenol, and all manipulations should be carried out in a fume hood. A glass receptacle should be available exclusively for disposing of used phenol and chloroform.

**ALTERNATE
PROTOCOL 1**

**SUPPORT
PROTOCOL 1**

**Preparation and
Analysis of DNA**

2.1.3

Materials

8-hydroxyquinoline
Liquefied phenol
50 mM Tris base (unadjusted pH ~10.5)
50 mM Tris·Cl, pH 8.0 (APPENDIX 2)
Chloroform
Isoamyl alcohol

1. Add 0.5 g of 8-hydroxyquinoline to a 2-liter glass beaker containing a stir bar.
2. Gently pour in 500 ml liquefied phenol or melted crystals of redistilled phenol (melted in a water bath at 65°C).

The phenol will turn yellow due to the 8-hydroxyquinoline, which is added as an antioxidant.

3. Add 500 ml of 50 mM Tris base.
4. Cover the beaker with aluminum foil. Stir 10 min at low speed with magnetic stirrer at room temperature.
5. Let phases separate at room temperature. Gently decant the top (aqueous) phase into a suitable waste receptacle. Remove what cannot be decanted with a 25-ml glass pipet and a suction bulb.
6. Add 500 ml of 50 mM Tris·Cl, pH 8.0. Repeat steps 4 to 6 so that two successive equilibrations with Tris·Cl are performed, ending with removal of the second Tris·Cl phase.

The pH of the phenol phase can be checked with indicator paper and should be 8.0. If it is not, the Tris·Cl equilibration should be repeated until this pH is obtained.

7. Add 250 ml of 50 mM Tris·Cl, pH 8.0, or TE buffer, pH 8.0, and store at 4°C in brown glass bottles or clear glass bottles wrapped in aluminum foil.

Phenol prepared with 8-hydroxyquinoline as an antioxidant can be stored ≤ 2 months at 4°C.

8. For use in DNA purification procedure (see Basic Protocol), mix 25 vol phenol (bottom yellow phase of stored solution) with 24 vol chloroform and 1 vol isoamyl alcohol. Store up to 2 months at 4°C wrapped in foil or in a dark glass bottle.

SUPPORT PROTOCOL 2

CONCENTRATION OF DNA USING BUTANOL

It is generally inconvenient to handle large volumes or dilute solutions of DNA. Water molecules (but not DNA or solute molecules) can be removed from aqueous solutions by extraction with *sec*-butanol (2-butanol). This procedure is useful for reducing volumes or concentrating dilute solutions before proceeding with the Basic Protocol.

Additional Materials (also see Basic Protocol)

sec-Butanol
25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol (made with buffered phenol;
see Support Protocol 1)
Polypropylene tube

1. Add an equal volume of *sec*-butanol to the sample and mix well by vortexing or by gentle inversion (if the DNA is of high molecular weight). Perform extraction in a polypropylene tube, as butanol will damage polystyrene.

2. Centrifuge 5 min at $1200 \times g$ (2500 rpm), room temperature, or in a microcentrifuge for 10 sec at maximum speed.
3. Remove and discard the upper (*sec*-butanol) phase.
4. Repeat steps 1 to 3 until the desired volume of aqueous solution is obtained.
5. Extract the lower, aqueous phase with 25:24:1 phenol/chloroform/isoamyl alcohol and ethanol precipitate (see Basic Protocol, steps 1 to 9) or remove *sec*-butanol by two ether extractions (see Support Protocol 3).

Addition of too much sec-butanol can result in complete loss of the water phase into the sec-butanol layer. If this happens, add 1/2 vol water back to the sec-butanol, mix well, and spin. The DNA can be recovered in this new aqueous phase and can be further concentrated with smaller amounts of sec-butanol.

The salt concentration will increase in direct proportion to the volume decrease. The DNA can be precipitated with ethanol to readjust the buffer conditions.

REMOVAL OF RESIDUAL PHENOL, CHLOROFORM, OR BUTANOL BY ETHER EXTRACTION

SUPPORT PROTOCOL 3

DNA solutions that have been purified by extraction with phenol and chloroform (see Basic Protocol) or concentrated with *sec*-butanol (see Support Protocol 2) can often be used without ethanol precipitation for enzymatic manipulations or in gel electrophoresis experiments if the organic solvents are removed by extraction with ether. Traces of ether are subsequently removed by evaporation. This procedure is useful only if the solute concentrations in the starting solution are compatible with what is needed in later steps. It is quite useful in purifying high-molecular-weight DNA, as mechanical shearing of large nucleic acid molecules can occur during precipitation with ethanol.

CAUTION: Ether is highly flammable and its vapors can cause drowsiness. All manipulations with ether should be carried out in a well-ventilated fume hood.

Materials

Diethyl ether
TE buffer, pH 8.0 (APPENDIX 2)
Polypropylene tube

1. Mix diethyl ether with an equal volume of water or TE buffer, pH 8.0, in a polypropylene tube. Vortex vigorously for 10 sec and let the phases separate.

Ether is the top phase.

2. Add an equal volume of hydrated ether to the DNA sample. Mix well by vortexing or by gentle inversion (if the DNA is of high molecular weight).
3. Microcentrifuge 5 sec at maximum speed or let the phases separate by setting the tube upright in a test tube rack.
4. Remove and discard the top (ether) layer. Repeat steps 2 and 3.
5. Remove ether by leaving the sample open under a fume hood for 15 min (small volumes, <100 μ l), or under vacuum for 15 min (larger volumes).

The DNA solution will be free of organic solvents and will have salt concentrations that are roughly three-fourths of those that were in the aqueous solution before phenol extraction (solute concentrations are lowered in the two phenol/chloroform/isoamyl alcohol extractions steps).

Preparation and Analysis of DNA

2.1.5

DNA PURIFICATION USING SILICA MEMBRANE SPIN COLUMNS

The use of glass beads or silica gel particles has become a popular method for isolating DNA. The evolution of this principle has resulted in the introduction of silica membrane spin columns. The basic principle of silica gel solid support spin columns is fairly simple. DNA is bound to the silica membrane spin column in the presence of a high concentration of chaotropic salt, contaminants are washed away, and the DNA is then eluted from the silica membrane in water or a low-salt buffer. The procedure below includes steps for when a bacterial lysate is used as the starting material.

The major advantage of silica membrane spin columns is the fact that the silica is bound to a solid support, which eliminates the problem of glass-bead contamination of the DNA sample. This method of DNA purification is quick and convenient, and can produce a high yield of pure DNA. Silica membrane spin columns are available from many companies, including Qiagen, Promega, Invitrogen, and Novagen, as kits including the columns and all appropriate buffers necessary for DNA purification.

Materials

- 1 × 10⁹ cell/ml bacterial culture containing DNA of interest *or* 0.1 to 1 mg/ml DNA to be purified (5 to 10 µg DNA total)
- 6 M sodium iodide (NaI) solution (filter through filter paper, store up to 3 months in the dark at 4°C)
- Resuspension buffer (see recipe)
- Lysis solution: 0.2 M NaOH/1.0% (w/v) SDS (store indefinitely at room temperature)
- Neutralization/binding solution (see recipe)
- Wash buffer (see recipe)
- TE buffer, pH 8.5 (*APPENDIX 2*) *or* nuclease-free H₂O
- Silica membrane spin columns (e.g., Qiagen, Promega, Invitrogen, Novagen)
- 1.5-ml microcentrifuge tubes

Prepare DNA solution

1. Harvest 1 to 5 ml of bacterial culture by centrifuging 1 min at 10,000 × *g*, room temperature, then discard the supernatant.

To clean a DNA solution not generated from bacterial lysates, add 3 vol of 6 M NaI solution to DNA in a 1.5-ml microcentrifuge tube, mix, and proceed to step 6.
2. Add 250 µl cell resuspension buffer, resuspend the cell pellet by vortexing or pipetting, and transfer to 1.5-ml microcentrifuge tube.
3. Add 250 µl lysis solution and mix by inversion (do not vortex). Allow lysis to proceed 3 to 5 min; do not allow lysis reaction to proceed for more than 5 min.
4. Add 350 µl neutralization/binding solution and mix by inversion.
5. Microcentrifuge 10 min at maximum speed (a white pellet will form in the bottom of the tube).

Purify DNA on silica membrane spin column

6. Apply the supernatant to the silica membrane spin column.
7. Microcentrifuge spin column inside its collection tube for 1 min at maximum speed. Remove the spin column from the collection tube and discard the flowthrough. Reinsert the spin column into the collection tube.

8. Wash the spin column by adding 750 μ l of wash buffer and microcentrifuging 1 min at maximum speed. Remove the spin column from the collection tube and discard the flowthrough.
9. Reinsert the spin column in the collection tube and microcentrifuge for an additional 1 min to remove any residual wash buffer (ethanol) from the column membrane.
10. Transfer the spin column to a 1.5-ml microcentrifuge tube and add 75 to 100 μ l nuclease-free water or TE buffer, pH 8.5, to the center of the membrane. Let stand for 2 to 10 min, then microcentrifuge 1 min at maximum speed.
11. Collect DNA and store at 4°C until use.

PURIFICATION AND CONCENTRATION OF RNA AND DILUTE SOLUTIONS OF DNA

*ALTERNATE
PROTOCOL 3*

The following adaptations to the purification procedure (see Basic Protocol) are used if RNA or dilute solutions of DNA are to be purified.

Purification and Concentration of RNA

The procedure outlined in the Basic Protocol is identical for purification of RNA, except that 2.5 vol ethanol should be used routinely for the precipitation (step 5). It is essential that all water used directly or in buffers be treated with diethylpyrocarbonate (DEPC) to inactivate RNase (see *UNIT 4.1*, reagents and solutions, for instructions).

Dilute Solutions of DNA

When DNA solutions are dilute (<10 μ g/ml) or when <1 μ g of DNA is present, the ratio of ethanol to aqueous volume should be increased to 3:1 and the time on dry ice (step 5) extended to 30 min. Microcentrifugation should be carried out for 15 min in a cold room to ensure the recovery of DNA from these solutions.

Nanogram quantities of labeled or unlabeled DNA can be efficiently precipitated by the use of carrier nucleic acid. A convenient method is to add 10 μ g of commercially available tRNA from *E. coli*, yeast, or bovine liver to the desired DNA sample. The DNA will be co-precipitated with the tRNA. The carrier tRNA will not interfere with most enzymatic reactions, but will be phosphorylated efficiently by polynucleotide kinase and should not be added if this enzyme will be used in subsequent radiolabeling reactions.

Recovery of small quantities of short DNA fragments and oligonucleotides can be enhanced by adding magnesium chloride to a concentration of <10 mM before adding ethanol (step 4). However, DNA precipitated from solutions containing >10 mM magnesium or phosphate ions is often difficult to redissolve and such solutions should be diluted prior to ethanol precipitation.

DNA in Large Aqueous Volumes (>0.4 to 10 ml)

Larger volumes can be accommodated by simply scaling up the amounts used in the basic protocol or by using butanol concentration (see Support Protocol 2). For the phenol extraction (see Basic Protocol, steps 1 through 3), tightly capped 15- or 50-ml polypropylene tubes should be used as polystyrene tubes cannot withstand the phenol/chloroform mixture. Centrifugation steps should be performed for 5 min at speeds not exceeding $1200 \times g$ (2500 rpm), room temperature. The ethanol precipitate (step 6) should be centrifuged in thick-walled Corning glass test tubes (15- or 30-ml capacity) for 15 min in fixed-angle rotors at $8000 \times g$ (10,000 rpm), 4°C. Glass tubes should be silanized (*APPENDIX 3B*) to facilitate recovery of small amounts of DNA (<10 μ g).

**Preparation and
Analysis of DNA**

2.1.7

REMOVAL OF LOW-MOLECULAR-WEIGHT OLIGONUCLEOTIDES AND TRIPHOSPHATES BY ETHANOL PRECIPITATION

The use of ammonium acetate in place of sodium acetate allows the preferential precipitation of longer DNA molecules. Thus, small single- or double-stranded oligonucleotides (less than ~30 bp) and unincorporated nucleotides used in radiolabeling or other DNA modification reactions can be effectively removed from DNA solutions by two rounds of ethanol precipitation in the presence of ammonium acetate. This approach is not sufficient to completely remove large quantities of linkers as used in cloning procedures (UNIT 3.16). If the nucleic acid is to be phosphorylated, this protocol should not be used because T4 polynucleotide kinase is inhibited by ammonium ions. Although the removal of unincorporated nucleoside triphosphates, reaction products, and small oligonucleotides is effective, it is not absolute and the procedure should not be used to purify DNA from these small molecules prior to detailed biochemical or analytical studies.

Additional Materials (also see Basic Protocol)

4 M ammonium acetate, pH 4.8

1. Add an equal volume of 4 M ammonium acetate, pH 4.8, to the DNA solution. Mix well.
2. Add 2 vol (calculated *after* salt addition) of ice-cold 100% ethanol. Vortex and set tube in crushed dry ice for 5 min.

This precipitation step can also be done in a -70°C freezer for 15 min or longer, or in a -20°C freezer for at least 30 min. A slurry of dry ice and ethanol may also be used, but tube labels are less often lost when crushed dry ice is used.

3. Microcentrifuge 5 min at high speed, room temperature. Carefully remove supernatant and redissolve pellet in 100 µl TE buffer, pH 8.0.
4. Repeat steps 1 to 3.

Reprecipitation is required, particularly if the DNA solution from step 1 contained Mg²⁺ or other divalent or polyvalent cations that will facilitate the precipitation of the oligonucleotides.

5. Proceed with steps 7 to 9 in Basic Protocol.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Neutralization/binding solution

Add 477.65 g guanidine HCl and 49.09 g potassium acetate to 500 ml water and stir to dissolve. Adjust pH to ~4.2 with acetic acid, dilute solution to 1 liter with water, and filter sterilize. Store indefinitely at 4°C.

Resuspension buffer

50 mM Tris·Cl, pH 8.0 (APPENDIX 2)
10 mM EDTA (APPENDIX 2)
100 mg RNase A
Store indefinitely at 4°C.

Wash buffer

1 part 10 mM Tris·Cl, pH 7.5 (APPENDIX 2)
1 part 100 mM NaCl (APPENDIX 2)
4 parts 100% ethanol (final 80%)
Store indefinitely at room temperature

COMMENTARY

Background Information

It is often necessary to purify or concentrate a solution of DNA prior to further enzymatic manipulations or analytical studies. The most commonly used method for deproteinizing DNA is extraction with phenol, which efficiently denatures proteins and probably dissolves denatured protein (Kirby, 1957). Chloroform is also a useful protein denaturant with somewhat different properties—it stabilizes the rather unstable boundary between an aqueous phase and a pure phenol layer. The phenol/chloroform mixture reduces the amount of aqueous solution retained in the organic phase (compared to a pure phenol phase), maximizing the yield (Penman, 1966; Palmiter, 1974). Isoamyl alcohol prevents foaming of the mixture upon vortexing and aids in the separation of the organic and aqueous phases (Marmur, 1961). Denatured protein forms a layer at the interface between the aqueous and organic phases and is thus isolated from the bulk of the DNA in the aqueous layer. This procedure is rapid, inexpensive, and easy to perform.

Ethanol precipitation is useful for concentrating DNA solutions and for removing residual phenol and chloroform from the deproteinized aqueous solution. It is also useful for providing DNA that is relatively free of solute molecules when buffer conditions need to be changed. In the presence of relatively high (0.1 to 0.5 M) concentrations of monovalent cations, ethanol induces a structural transition in nucleic acid molecules which causes them to aggregate and precipitate from solution (Eickbush and Moudrianakis, 1978). However, because most salts and small organic molecules are soluble in 70% ethanol, ethanol precipitation and washing of the pellet will effectively desalt DNA. Although sodium chloride, sodium acetate, and ammonium acetate are each capable of inducing precipitation, it is more difficult to remove sodium chloride due to its lower solubility in 70% ethanol.

The silica membrane spin column protocol, modified from principles originally described in Vogelstein and Gillespie (1979), provides a simple, nontoxic method for removing DNA from contaminating impurities (UNIT 2.6). In the presence of high chaotropic salt concentrations, DNA binds to a silica membrane inside a spin column. The resulting precipitate is washed to remove NaI (or other chaotropic salts like guanidine HCl or sodium perchlorate) along with impurities from the original sample, and

subsequent suspension in water or TE buffer causes dissociation (elution) of the DNA from the silica membrane. Because fewer manipulations are required, this method is faster and easier to perform than organic-based extraction methods. However, the yields may somewhat lower, generally ranging from 50% to 75% of the starting material. The procedure seems to work best with DNA fragments larger than 500 bp, as some shorter fragments may bind tightly and irreversibly to the silica membrane. An advantage of silica membrane spin columns relative to glass beads (Vogelstein and Gillespie, 1979) is that they help to decrease shearing of DNA fragments that are larger than 3 to 10 kb.

One alternative for purifying DNA from residual protein is the use of StrataClean Resin (Stratagene), a nontoxic slurry of hydroxylated silica particles. Acidic hydroxy groups on the resin appear to bind proteins in a manner similar to phenolic hydroxyls, and at or near neutral pH display a high affinity for protein and low affinity for DNA. Protein bound to the resin is separated by centrifugation from nucleic acids remaining in solution; two or three extractions with the resin may be required to completely remove protein from a nucleic acid sample. Another product, Phase Lock Gel (an inert silica-based blend of intermediate density available from 5 Prime → 3 Prime; APPENDIX 4) improves recoveries in standard organic extractions by reducing loss of sample at the interface. During centrifugation, the normally fuzzy interface is compacted tightly below or within the gel. The gel/interface complex migrates discretely between the organic and aqueous phases, thus creating a tight partition which allows recovery of virtually all of the aqueous phase.

Critical Parameters

The oxidation products of phenol can damage nucleic acids and only redistilled phenol should be used. For complete deproteinization, extractions should be repeated until no protein precipitate remains at the aqueous/organic interface.

In general, alcohol precipitation of nucleic acids requires the presence of at least 0.1 M monovalent cation in the starting aqueous solution. Precipitation of nucleic acids at low concentrations requires cooling to low temperatures to give good recovery. Precipitation of nucleic acids at high concentrations (>0.25

mg/ml after addition of ethanol) is very rapid at room temperature. Formation of a visible precipitate after adding alcohol and mixing well indicates complete precipitation, and no chilling or further incubation is needed.

In organic extraction, loss of nucleic acid at the interface and into the organic phase is minimized by back-extracting the organic phase. The most critical parameter in the purification of DNA using silica membrane spin columns is the pH of the chaotropic salt solution. For efficient binding to the silica membrane, the salt solution should be at pH 6.5. Many different companies sell silica membrane spin columns as kits supplied with all necessary reagents and buffers. One variable is the type of chaotropic salt solution supplied with the kit (NaI, guanidine HCl, or guanidine isothiocyanate). For the most part, these salt solutions are only a variation on the same theme and may be interchangeable between kits or columns from different vendors.

Anticipated Results

These procedures should result in virtually complete removal of proteins and quantitative recovery of nucleic acids. However, sequential extractions or precipitations require care and attention to detail to prevent accumulation of small losses at each step. It is particularly important to carefully recover the aqueous phase and reextract the organic phase to ensure full recovery of small amounts of DNA from phenol/chloroform extractions.

The yield of nucleic acids resulting from the silica membrane spin column procedure can be similarly improved (to $\leq 80\%$ recovery) by subjecting supernatants to an additional binding step and increasing the amount of elution buffer to 100 μ l or more.

Time Considerations

Approximately 90 min should be allowed for carrying out steps 1 through 9 of the Basic Protocol on twelve DNA samples in microcentrifuge tubes. Phenol buffering should be started ≥ 1 hr before equilibrated phenol is needed. Nucleic acids should not be left in the presence of phenol, but can be indefinitely precipitated in alcohol or dried after precipitation. The silica membrane spin column protocol can be performed on twelve samples in 15 to 20 min.

Literature Cited

- Eickbush, T.H. and Moudrianakis, E.N. 1978. The compaction of DNA helices into either continuous supercoils or folded-fiber rods and toroids. *Cell* 13:295-306.
- Hall, S.S. 1987. Going for the gene. *The Boston Globe Magazine*, Aug. 2, p. 20.
- Kirby, K.S. 1957. A new method for the isolation of deoxyribonucleic acids: Evidence on the nature of bonds between deoxyribonucleic acid and protein. *Biochem. J.* 66:495-504.
- Marmur, J. 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. *J. Mol. Biol.* 3:208-218.
- Palmiter, R.D. 1974. Magnesium precipitation of ribonucleoprotein complexes. Expedient techniques for the isolation of undegraded polysomes and messenger ribonucleic acid. *Biochemistry* 13:3606-3615.
- Penman, S. 1966. RNA metabolism in the HeLa cell nucleus. *J. Mol. Biol.* 17:117-130.
- Vogelstein, B. and Gillespie, D. 1979. Preparative and analytical purification of DNA from agarose. *Proc. Nat. Acad. Sci. U.S.A.* 76:615-619.

Contributed by David Moore and
Dennis Dowhan
Baylor College of Medicine
Houston, Texas

Purification of DNA by Anion-Exchange Chromatography

Column chromatography has recently evolved to provide a rapid and effective alternative to more laborious methods for preparing high-quality DNA, such as CsCl-gradient centrifugation. This protocol describes the use of a column made of a unique anion-exchange resin that selectively binds nucleic acids, allowing rapid separation of DNA from contaminating RNA, proteins, carbohydrates, and metabolites. The procedure below employs columns supplied by QIAGEN; other preparation methods are available from other suppliers. A crude nucleic acid sample (usually a cleared cell lysate) is applied to the QIAGEN-tip under conditions that favor binding. Contaminants in the sample are washed from the column with a moderate-salt buffer, and DNA is eluted using a high-salt buffer.

Materials

Plasmid- or phage-bearing bacterial culture or mammalian, plant, or bacterial cell culture

Buffer QBT (equilibration buffer; see recipe)

QIAGEN-tip of appropriate size (Table 2.1B.1)

Buffer QC (washing buffer; see recipe)

Buffer QF (eluting buffer; see recipe)

TE buffer, pH 8.0 (APPENDIX 2)

Isopropanol, room temperature

70% (v/v) ethanol, ice cold

Beckman JS-13, Sorvall HB-4 or HB-6, or equivalent rotor

Additional reagents and equipment for preparation of mammalian, plant, or bacterial genomic DNA (cell lysate; UNITS 2.2-2.4), alkaline lysis preparation of plasmid DNA (UNIT 1.7), or preparation of phage lysates (UNIT 1.13)

1. To isolate plasmid, cosmid, or phage DNA, prepare a cell lysate from a strain harboring the appropriate vector. To isolate genomic DNA, prepare cell lysate from mammalian, plant, or bacterial culture. Remove a sample from the cleared lysate and

Table 2.1B.1 Recommended Loading Volumes, DNA Capacities, and Culture Volumes for QIAGEN-tips

	QIAGEN-tip 100	QIAGEN-tip 500	QIAGEN-tip 2500	QIAGEN-tip 10000
<i>Column DNA capacity</i>				
Plasmid or ds phage	100 µg	500 µg	2500 µg	10000 µg
Cosmid or λ phage	60 µg	300 µg	1500 µg	6000 µg
Mammalian or <i>B. subtilis</i> genomic	80 µg	400 µg	2000 µg	8000 µg
<i>Culture volume for plasmid DNA</i>				
High-copy-number plasmid	25 ml	100 ml	500 ml	2.5 liters
Low-copy-number plasmid	100 ml	500 ml	2.5 liters	5 liters
<i>Culture volume for genomic DNA</i>				
<i>B. subtilis</i>	8 ml	40 ml	200 ml	800 ml
Mammalian (no. cells)	2×10^7	1×10^8	5×10^8	2×10^9

save for an analytical gel to determine whether growth and lysis conditions were optimal.

For plasmid preps, lysozyme treatment is not necessary. Plasmid and phage preps should include RNase A treatment, because the long RNA molecules can compete with DNA for binding sites on the resin. When the sample is loaded on the QIAGEN-tip, it must have a salt concentration of 750 mM and a pH of 7.0, and should be free of particulate matter, which will clog the column.

Bacterial lysates can be cleared by a conventional centrifugation or by filtration through a special filter unit available from QIAGEN. The QIAfilter is available in syringe or cartridge format and is suitable for small to large bacterial culture volumes. It has been designed to rapidly clear bacterial lysates and completely remove SDS precipitates without centrifugation.

For endotoxin-free DNA, QIAGEN offers EndoFree Plasmid Kits containing an endotoxin-removal buffer that is briefly incubated with the cleared lysate prior to loading on the QIAGEN-tip. Plasmid DNA purified using EndoFree Plasmid Kits contains only negligible amounts of endotoxin.

Anionic detergents such as SDS will inhibit DNA binding to the resin; thus, their use in preparing these cell lysates should be avoided if possible (see Tables 10.15.1 and 10.15.2 for physical and chemical properties of commonly used detergents). If anionic detergents must be included in any of the extraction or lysis buffers, they may be removed by precipitation with acidic potassium acetate (UNIT 1.7) followed by centrifugation. Small SDS precipitates that cannot be separated by conventional centrifugation are completely removed by the QIAfilter process.

2. Equilibrate an appropriately sized QIAGEN-tip with Buffer QBT by applying 2 bed volumes of buffer and allowing it to empty by gravity flow.

The flow of buffer will stop when the meniscus reaches the upper column frit. Do not force out the remaining buffer.

3. Apply cleared lysate from step 1 to the equilibrated column and allow it to enter the resin by gravity flow. Remove a small sample of flowthrough (50 to 500 μ l, depending on the column size) and save for analytical gel electrophoresis.
4. Wash column twice with 3 to 6 bed volumes of Buffer QC (depending on column size) and allow the buffer to move through the QIAGEN-tip by gravity flow. Remove a small sample of flowthrough and save for analytical gel electrophoresis.

This step removes cellular contaminants and any remaining RNA from the DNA.

5. Elute pure DNA by applying 1 to 2 bed volumes of Buffer QF (depending on column size) and allowing it to flow through by gravity flow. Remove a small sample of eluate and save for analytical gel electrophoresis.
6. Add 0.7 vol isopropanol to eluate and immediately centrifuge 30 min at 15,000 \times g (\geq 9500 rpm in Beckman JS-13), 4°C, to precipitate DNA. Remove and discard supernatant.
7. Wash pellet with ice-cold 70% ethanol, air dry 10 min, and redissolve in a suitable volume of TE buffer, pH 8.0.

Overdrying will make the pellet difficult to dissolve. Do not pipet high-molecular-weight DNA up and down when dissolving, as this may cause shearing.

8. *Optional:* To analyze the purification procedure, precipitate the samples taken in steps 1, 3, 4, and 5 with 0.7 vol isopropanol. Rinse pellets with 70% ethanol, drain well, and resuspend in 10 μ l TE buffer, pH 8.0. Use 2 μ l of each for analysis on a 1% agarose gel (UNIT 2.5A; see Critical Parameters and Troubleshooting).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Buffer QBT

750 mM NaCl
50 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS)
Adjust pH to 7.0 with HCl
15% (v/v) isopropanol
0.15% (v/v) Triton X-100
Store indefinitely at room temperature

Buffer QC

1.0 M NaCl
50 mM MOPS
Adjust pH to 7.0 with HCl
15% (v/v) isopropanol
Store indefinitely at room temperature

Buffer QF

1.25 M NaCl
50 mM Tris·Cl
15% isopropanol
Adjust pH to 8.5 with HCl
Store indefinitely at room temperature

COMMENTARY

Background Information

For many routine applications in molecular biology, DNA used as starting material must be very pure. Until recently, traditional CsCl/ethidium bromide gradients (for plasmid and phage DNA; *UNITS 1.7 & 1.13*) or organic extraction (for genomic DNA; *UNIT 2.1A*) have been widely used. Although they provide a highly purified product, these methods are generally time-consuming, may involve the use of hazardous amounts of toxic reagents, and do not yield endotoxin-free plasmid DNA. A simple and time-saving alternative to these methods, without sacrifice in quality, is provided by ion-exchange chromatography. This technique was originally used for protein purification (see *UNIT 10.10* for a complete discussion), but is now also widely used for nucleic acid separation.

The general principles of ion-exchange purification hold true for nucleic acid separations: negatively charged nucleic acids are applied to an oppositely charged chromatographic matrix. The best choice for these separations is an anion-exchange resin carrying positively charged groups that adsorb negatively charged molecules in buffers near neutral pH and of medium ionic strength. In particular, the high surface density of anion-exchange (DEAE)

groups—together with the large pore size of the commercially available QIAGEN Anion-Exchange Resin—permits attachment and retention of the highly charged nucleic acid molecules on the column. Subsequent washing of the column to remove impurities contained in the lysate results in a very high degree of purification of the DNA.

The separation range of conventional anion exchangers for nucleic acids extends only up to ~0.5 M salt. Because binding and elution of the components of a lysate are limited to a narrow range of salt concentrations, the elution peaks of proteins, RNA, and DNA overlap with one another, preventing effective separation of DNA from contaminants. Because of the extremely high charge density (QIAGEN Resin contains ~5-fold more DEAE groups per unit surface area of resin than do conventional anion exchangers), the resin described here separates various nucleic acid forms over a broader range of salt concentration (up to 1.6 M, the elution point for double-stranded DNA; Fig. 2.1B.1). Because the charge density of the resin is inversely proportional to the pH of the buffer, the binding, washing, and elution profiles are strongly affected by the pH of the relevant buffers (Fig. 2.1B.2).

Anion-exchange methods have been used in conjunction with a variety of applications whose success depends critically upon the purity of the nucleic acid sample, including transfection of mammalian cells (Chen et al., 1991), DNA sequencing (Hall et al., 1989; Voss et al., 1990), PCR (Jones and Winistorfer, 1992; Jung

et al., 1992), and runoff transcription (Kayne et al., 1988; Luytjes et al., 1989). In addition, the method has been used to recover plasmid and viral DNA from mammalian cells (Lutze and Winegar, 1990).

Further alternatives to traditional purification procedures include gel filtration as well as

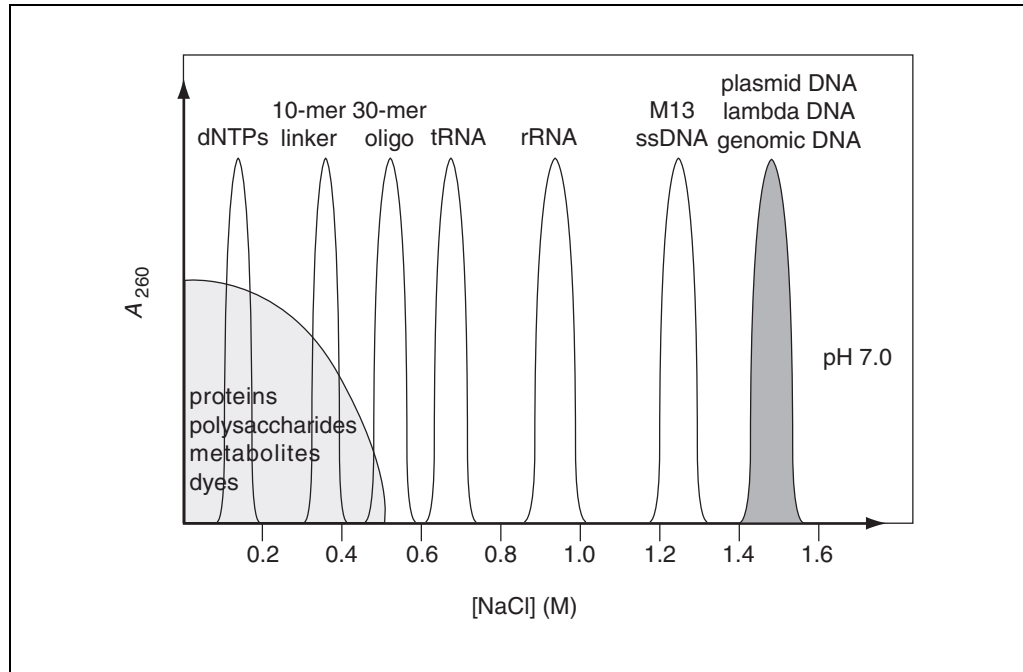


Figure 2.1B.1 Separation of nucleic acids at neutral pH on QIAGEN Anion-Exchange Resin (reproduced with permission from QIAGEN, 1997).

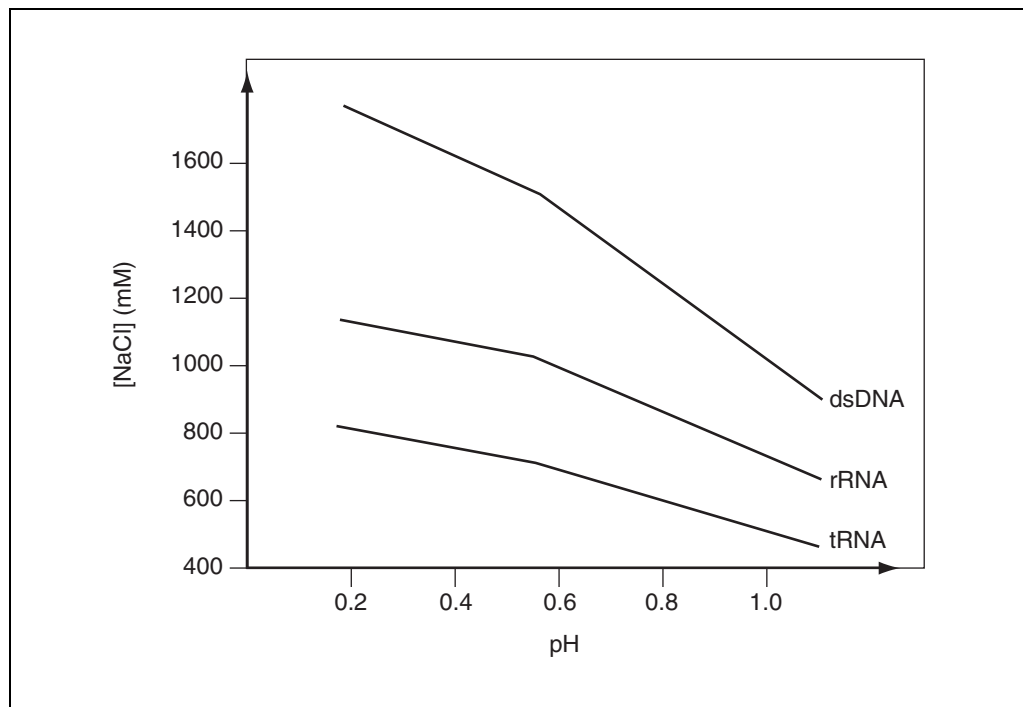


Figure 2.1B.2 Elution points of different nucleic acids from QIAGEN Anion-Exchange Resin as a function of pH (reproduced with permission from QIAGEN, 1997).

Table 2.1B.2 Troubleshooting Guide for DNA Purification by Anion-Exchange Chromatography (adapted from QIAGEN, 1997)

Problem	Possible cause(s)	Solution
No DNA in sample lysate before loading	Plasmid failed to propagate	Check that conditions for optimum plasmid growth were met
	Lysate was incorrectly prepared	Check age of buffers and prepare fresh if necessary
DNA in flowthrough fraction	Column overloaded with DNA	Check yield against capacity of column (see Table 2.1B.1). Purify excess DNA by passing through new tip.
	SDS or other ionic detergent in the DNA sample	Ensure that SDS is removed from lysate before column loading by dialyzing or precipitating with cold potassium acetate
High-molecular-weight RNA in eluate	RNase A digestion insufficient	Check culture volume against recommended volumes in Table 2.1B.1 and reduce if necessary. Check effectiveness of RNase A solution.
		Recover DNA by precipitating the eluate, digesting with RNase A, and purifying on a new tip
DNA found in wash fraction	Column overloaded with DNA	Check the culture volume and yield against the capacity of the respective QIAGEN-tip. Reduce culture volume accordingly.
	Wash buffer incorrect	Check pH and salt concentration of Buffer QC. Recover DNA by precipitation and purify on another column.
No DNA in eluate	Elution buffer incorrect	Check pH and salt concentration of Buffer QF. Recover DNA by eluting with fresh buffer.
Little or no DNA upon precipitation	DNA failed to precipitate	Check isopropanol batch. Make sure that precipitate is centrifuged at $>15,000 \times g$ for 30 min; if necessary, centrifuge longer at higher speeds.
	DNA pellet lost (isopropanol pellets are glassy and may be difficult to see)	Mark expected location of pellet before centrifugation
	DNA poorly resuspended	Check that DNA was completely resuspended. Wash any DNA off walls, particularly if glass tubes are used, as up to half of total DNA may be smeared on walls. Use a swinging-bucket rotor so that the pellet will be localized at very bottom of tube.
DNA difficult to resuspend	Pellet overdried	Air dry pellet; do not use a vacuum. Warm DNA solution slightly to dissolve.
	Residual isopropanol in pellet	Wash pellet in 70% ethanol
	Too much salt in pellet	Use isopropanol at room temperature. Increase volume of TE buffer.
RNA in eluate	RNase A digestion insufficient	Check culture volume against recommended volumes in Table 2.1B.1 and reduce if necessary. Increase volume of wash buffer, recover DNA by precipitating eluate, digest with RNase A, and purify.
Genomic DNA contamination in eluate (for plasmid preps)	Mixing too vigorous	Lysate must be handled gently to prevent shearing. Reduce culture volume if lysate is too viscous for gentle mixing.
	Lysis time too prolonged	Make sure that lysis step does not last longer than 5 min

continued

Preparation and Analysis of DNA

2.1.15

Table 2.1B.2 Troubleshooting Guide for DNA Purification by Anion-Exchange Chromatography (adapted from QIAGEN, 1997), continued

Problem	Possible cause(s)	Solution
DNA is nicked, sheared, or degraded	Endonuclease-containing host	Consider changing <i>E. coli</i> host strain
	Nuclease contamination present	Check buffers for nuclease contamination and replace if necessary. Use autoclaved glass and plastic and wear gloves.
	DNA poorly buffered	Resuspend DNA in TE, pH 8.0, to inhibit nuclease activity and maintain stable pH during storage
	Shearing during resuspension	Resuspend DNA gently without vortexing or vigorous pipetting
DNA does not perform well in later experiments	DNA nicked	Check percentage of nicked DNA on agarose gel
	Too much salt	Make sure isopropanol is at room temperature for precipitation and pellet is washed with 70% ethanol
	Residual protein	Check culture volume against recommended volumes in Table 2.1B.1 and reduce if necessary. Increase volume of wash buffer.

additional separation methods (e.g., G-25/G-50, Select, and pZ spin columns, available from 5 Prime → 3 Prime, and Wizard DNA purification resin, available from Promega; also see *UNIT 1.13*). Gel-filtration columns purify DNA by allowing large nucleic acid molecules to pass through while retarding the migration of contaminants (small RNAs and oligonucleotides) through the gel. In general, these alternatives are comparable in terms of the purity of the sample obtained, although some may permit carryover of residual impurities from the starting sample. Opinions vary, however, regarding the relative values of the different methods. For example, recent reports assert that plasmid DNA purified by QIAGEN Anion-Exchange Resin is at least comparable (and may be superior) to preparations purified using CsCl. Determinations were made by measuring transfection efficiencies (Ehlert et al., 1993) and by observations of plasmid preparations using electron microscopy (Schleef and Heilmann, 1993). In a series of biological assays, various purified preparations of plasmid DNA injected into *Drosophila* embryos were subsequently analyzed for the ability to generate germline-transformed flies. No significant differences were observed among CsCl-, PEG-, QIAGEN-, or pZ-purified plasmids, although PEG-derived preparations appeared to be slightly less toxic to the cells (Joseph Heilig, pers. comm.).

Critical Parameters and Troubleshooting

QIAGEN Anion-Exchange Resin has different binding capacities for different classes of nucleic acids: the binding capacity for RNA, for example, is roughly twice that for plasmid DNA. Larger DNA molecules (e.g., phage, cosmid, BACs, P1s, and genomic DNAs) are bound at an even lower capacity than plasmid DNA. Therefore, when selecting a column size, the type of nucleic acid to be purified should be taken into account. It is also important to choose a column with the appropriate binding capacity for the culture volume used, to avoid overloading (see Table 2.1B.1). For plasmid, phage, and cosmid preparations, this is achieved by careful control of culture conditions and inclusion of an RNase treatment in the lysate preparation. While the use of super-rich medium such as Superbroth (*UNIT 1.1*) or Terrific broth (Tartoff and Hobbs, 1987; *UNIT 1.7*) can be advantageous for culturing cosmids and plasmids maintained at low copy number, high-copy-number plasmids do not require growth in super-rich medium to give a reasonable yield. In these cases, use of super-rich medium can result in extraordinarily high cell densities with a corresponding increase in lysate viscosity due to the presence of large amounts of nucleic acid or protein. Before loading the column, ensure that the lysate is clear and free of particulate matter and adjust the loading volume as indicated in Table 2.1B.1.

Another critical factor is the pH of the column buffers used for binding, washing, and eluting (Fig. 2.1B.2). For optimal recovery, the final pH should be adjusted as recommended in the recipes (see Reagents and Solutions); changes in pH may lead to loss of DNA.

Many factors may cause low or no yield or incomplete purification of DNA. Checking an aliquot of each fraction on an analytical gel (Fig. 2.1B.3) can help locate the source of the problem. Table 2.1B.2 describes various problems that may be encountered and provides suggestions for how to deal with them.

Anticipated Results

The expected yield for a preparation of genomic DNA from mammalian cell culture is ~30 to 40 μg per 10^7 cells. Low-copy-number plasmid and cosmid preparations from cells grown overnight in LB medium yield 0.2 to 1 μg DNA per milliliter of overnight culture, while high-copy-number plasmids prepared under these conditions yield 3 to 5 $\mu\text{g}/\text{ml}$. When grown in super-rich medium, the yield of cosmids and low-copy-number plasmids is in-

creased to 2.5 to 5 μg DNA/ml, and that of high-copy-number plasmids to 5 to 25 $\mu\text{g}/\text{ml}$.

Time Considerations

With most types of cells, preparing the original DNA sample (lysate) and running the column (including equilibration time, sample loading, washing, and elution) should take 1.5 to 4 hr, depending upon the volume of lysate and the corresponding column size.

Literature Cited

- Chen, W., Andres, D., Goldstein, J., Russell, D., and Brown, M. 1991. cDNA cloning and expression of the peptide-binding beta subunit of Rat p21^{ras} farnesyltransferase, the counterpart of yeast DPR1/RAM1. *Cell* 66:327-334.
- Ehlert, F., Bierbaum, P., and Schorr, J. 1993. Importance of DNA quality for transfection efficiency. *BioTechniques* 14:546.
- Hall, B., Betts, P., and Wootton, J. 1989. DNA sequence analysis of artificially evolved *ebg* enzyme and *ebg* repressor gene. *Genetics* 123:635-648.
- Jones, D. and Winistorfer, S. 1992. Recombinant circle PCR and recombinant PCR for site-specific mutagenesis without PCR production purification. *BioTechniques* 12:528-535.

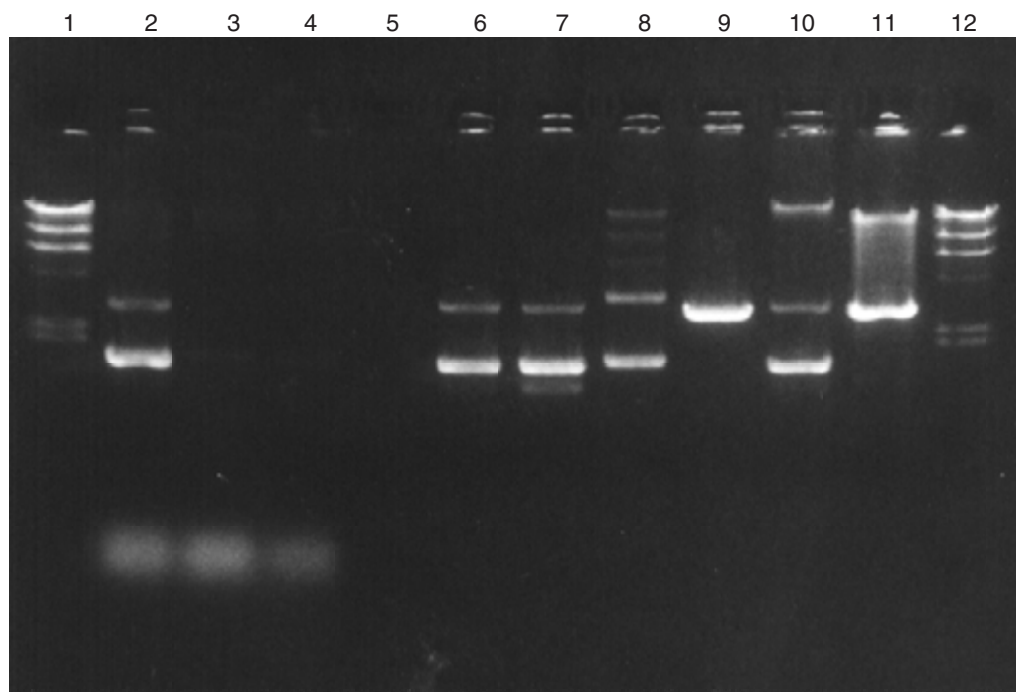


Figure 2.1B.3 Agarose analytical gel (1%) comparing DNA composition of QIAGEN-tip elution fractions at different stages of plasmid purification (lanes 2 to 6) or containing different types of plasmid DNA (lanes 7 to 11). Lanes: 1, lambda *Hind*III marker; 2, cleared lysate before column purification; 3, flowthrough fraction; 4,5, first and second Buffer QC washes; 6, eluted plasmid DNA; 7, fraction containing denatured supercoiled DNA; 8, fraction containing multimeric forms of supercoiled plasmid DNA; 9, fraction containing linear plasmid DNA (pTZ19/*Eco*RI); 10, fraction contaminated with bacterial chromosomal DNA; 11, fraction 10 digested with *Eco*RI; 12, lambda *Hind*III marker. (reproduced with permission from QIAGEN, 1997).

- Jung, M., Dritschilo, A., and Kasid, U. 1992. Reliable and efficient direct sequencing of PCR double-stranded genomic DNA template. *PCR Methods Appl.* 1:171-174.
- Kayne, P., Kim, U., Han, M., Mullen, J., Yoshizaki, F., and Grunstein, M. 1988. Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. *Cell* 55:27-39.
- Lutze, L. and Winegar, R.A. 1990. A quick and efficient method for the recovery of plasmid or viral DNA from mammalian cells. *Nucl. Acids Res.* 18:6150.
- Luytjes, W., Krystal, M., Enami, M., Parvin, J., and Palese, P. 1989. Amplification, expression, and packaging of a foreign gene by influenza virus. *Cell* 59:1108-1113.
- QIAGEN. 1997. QIAGEN Plasmid Purification Handbook. QIAGEN, Valencia, Calif.
- Schleef, M. and Heimann, P. 1993. Cesium chloride or column preparation? An electron microscopical view of plasmid preparations. *BioTechniques* 14:544.
- Tartoff, K.D. and Hobbs, C.A. 1987. Improved media for growing plasmid and cosmid clones. *Focus (BRL)* 9:12.
- Voss, H., Zimmerman, J., Schwager, C., Stegemann, J., Erfle, H., Stuckey, K., and Ansorge, W. 1990. Automated fluorescent sequencing of cosmic DNA. *Nucl. Acids Res.* 18:1066.

Contributed by Kim Budelier
QIAGEN, Inc.
Valencia, California

Joachim Schorr
QIAGEN GmbH
Hilden, Germany

Preparation of Genomic DNA from Mammalian Tissue

UNIT 2.2

**BASIC
PROTOCOL**

Tissue is rapidly frozen and crushed to produce readily digestible pieces. The processed tissue is placed in a solution of proteinase K and SDS and incubated until most of the cellular protein is degraded. The digest is deproteinized by successive phenol/chloroform/isoamyl alcohol extractions, recovered by ethanol precipitation, and dried and resuspended in buffer.

Materials

Tissues, whole or cultured cells
Liquid nitrogen
Digestion buffer (see recipe)
PBS (*APPENDIX 2*), ice cold
7.5 M ammonium acetate
70% and 100% ethanol
TE buffer, pH 8 (*APPENDIX 2*)
Incubator or water bath at 50°C, with shaker

Additional reagents and equipment for trypsinizing adherent cells (*APPENDIX 3F*) and phenol/chloroform/isoamyl alcohol extraction (*UNIT 2.1A*)

Prepare cells

Beginning with whole tissue:

1a. As soon as possible after excision, quickly mince tissue and freeze in liquid nitrogen.

If working with liver, remove the gallbladder, which contains high levels of degradative enzymes.

2a. Starting with between 200 mg and 1 g, grind tissue with a prechilled mortar and pestle, or crush with a hammer to a fine powder (keep the tissue fragments, if crushing is incomplete).

3a. Suspend the powdered tissue in 1.2 ml digestion buffer per 100 mg tissue. There should be no clumps.

Beginning with tissue culture cells:

1b. Pellet suspension culture out of its serum-containing medium. Trypsinize adherent cells and collect cells from the flask. Centrifuge 5 min at $500 \times g$, 4°C, and discard supernatant.

2b. Resuspend cells with 1 to 10 ml ice-cold PBS. Centrifuge 5 min at $500 \times g$ and discard supernatant. Repeat this resuspension and centrifugation step.

3b. Resuspend cells in 1 vol digestion buffer. For $<3 \times 10^7$ cells, use 0.3 ml digestion buffer. For larger numbers of cells use 1 ml digestion buffer/ 10^8 cells.

Lyse and digest cells

4. Incubate the samples with shaking at 50°C for 12 to 18 hr in tightly capped tubes.

The samples will be viscous. After 12 hr incubation the tissue should be almost indiscernible, a sludge should be apparent from the organ samples, and tissue culture cells should be relatively clear.

**Preparation and
Analysis of DNA**

2.2.1

Extract nucleic acids

5. Thoroughly extract the samples with an equal volume of phenol/chloroform/isoamyl alcohol.

CAUTION: *Phenol is extremely caustic.*

6. Centrifuge 10 min at $1700 \times g$ in a swinging bucket rotor.

If the phases do not resolve well, add another volume of digestion buffer, omitting proteinase K, and repeat the centrifugation.

If there is a thick layer of white material at the interface between the phases, repeat the organic extraction.

Purify DNA

7. Transfer the aqueous (top) layer to a new tube and add $\frac{1}{2}$ vol of 7.5 M ammonium acetate and 2 vol (of original amount of top layer) of 100% ethanol. The DNA should immediately form a stringy precipitate. Recover DNA by centrifugation at $1700 \times g$ for 2 min.

This brief precipitation in the presence of high salt reduces the amount of RNA in the DNA. For long-term storage it is convenient to leave the DNA in the presence of ethanol.

Alternatively, to prevent shearing of high-molecular-weight DNA, omit steps 7 to 9 and remove organic solvents and salt from the DNA by at least two dialysis steps against at least 100 vol TE buffer. Because of the high viscosity of the DNA, it is necessary to dialyze for a total of at least 24 hr.

8. Rinse the pellet with 70% ethanol. Decant ethanol and air dry the pellet.

It is important to rinse well to remove residual salt and phenol.

9. Resuspend DNA at ~ 1 mg/ml in TE buffer until dissolved. Shake gently at room temperature or at 65°C for several hours to facilitate solubilization. Store indefinitely at 4°C .

From 1 g mammalian cells, ~ 2 mg DNA can be expected.

If necessary, residual RNA can be removed at this step by adding 0.1% sodium dodecyl sulfate (SDS) and 1 $\mu\text{g/ml}$ DNase-free RNase (UNIT 3.13) and incubating 1 hr at 37°C , followed by organic extraction and ethanol precipitation, as above.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Digestion buffer

100 mM NaCl
10 mM Tris·Cl, pH 8 (APPENDIX 2)
25 mM EDTA, pH 8 (APPENDIX 2)
0.5% SDS
0.1 mg/ml proteinase K
Store at room temperature

The proteinase K is labile and must be added fresh with each use.

COMMENTARY

Background Information

There are a number of different procedures for the preparation of genomic DNA. They all start with some form of cell lysis, followed by deproteination and recovery of DNA. The main differences between various approaches lie in the extent of deproteination and in molecular weight of the DNA produced. The isolation procedure described here is relatively brief and relies on the powerful proteolytic activity of proteinase K combined with the denaturing ability of the ionic detergent SDS. Use of proteinase K for DNA purification was described by Gross-Bellard et al. (1973) and Enrietto et al. (1983). EDTA is included in the digestion buffer to inhibit DNases.

Critical Parameters

To minimize the activity of endogenous nucleases, it is essential to rapidly isolate, mince, and freeze tissue. Tissue culture cells should be cooled and washed quickly. As soon as the tissue is frozen or the tissue culture cells are added to the lysis buffer, DNA is protected from action of nucleases throughout this protocol. It is important that the tissue be well dispersed, and not left in large lumps, to permit rapid and efficient access to proteinase K and SDS.

It is crucial to generate high-molecular-weight DNA for construction of phage (>60 kb) or cosmid (>120 kb) genomic libraries. Two main precautions should be taken to maximize molecular weight: (1) minimize shearing forces by gentle (but thorough) mixing during extraction steps, and (2) after the extraction, remove organic solvents and salt from the DNA by dialysis, rather than by ethanol precipitation. Additional precautions must be taken to prepare very high-molecular-weight DNA for the construction of P1 or BAC libraries.

The absence of both cellular proteins and proteinase K in the final DNA solution is important for susceptibility of the genomic DNA to restriction enzyme action; therefore, care should be exercised in deproteination. To remove protein completely it may be necessary to repeat the proteinase K digestion. In general, highly pure DNA has an A_{260}/A_{280} ratio >1.8, while 50% protein/50% DNA mixtures have A_{260}/A_{280} ratios of ~1.5.

Troubleshooting

Failure of the organic phase to separate cleanly from the aqueous phase is generally due to a very high concentration of DNA and/or cellular debris in the aqueous phase. Dilution with more digestion buffer and reextraction can remedy this problem.

Upon addition of the room-temperature ethanol to the extracted DNA solution, the DNA should precipitate in long, stringy fibers. If there is no precipitate or if the precipitate is flocculent, the DNA is either degraded or not purified away from cellular debris. Improper handling of the tissue before digestion or too much tissue in the digestion reaction are possible causes of such problems.

Anticipated Results

Approximately 2 mg DNA should be obtained from 1 g tissue or 10^9 cells. The DNA should be at least 100 kb long and should be digestible with restriction enzymes.

Time Considerations

This protocol involves effort on 2 days: tissue preparation on the first day followed by overnight lysis, and extraction/precipitation on the second day. Actual time spent on the procedure, however, will be less than 1 hr each day. The DNA can be stored indefinitely in the presence of ethanol at 4° or in TE buffer at -20°C.

Literature Cited

- Enrietto, P.J., Payne, L.N., and Hayman, M.J. 1983. A recovered avian myelocytomatosis virus that induces lymphomas in chickens. Pathogenic properties and their molecular basis. *Cell* 35:369-379.
- Gross-Bellard, M., Oudet, P., and Chambon, P. 1973. Isolation of high-molecular-weight DNA from mammalian cells. *Eur. J. Biochem.* 36:32-38.

Contributed by William M. Strauss
Harvard Medical School and Beth
Israel Deaconess Medical Center
Boston, Massachusetts

Preparation of Genomic DNA from Plant Tissue

UNIT 2.3

This unit describes two methods for preparing genomic DNA from plant tissue. In the first method, plant cells are lysed with ionic detergent, treated with protease, and subsequently purified by cesium chloride (CsCl) density gradient centrifugation. The second method is based upon a series of treatments with the nonionic detergent cetyltrimethylammonium bromide (CTAB) to lyse cells and purify nucleic acid. Nucleic acid is recovered from the final CTAB solution by isopropanol or ethanol precipitation. The first method, although somewhat more lengthy, results in highly purified nucleic acid. The second method requires fewer manipulations, results in very high yields (~10-fold higher per gram fresh tissue depending on species and condition of starting material), and produces DNA that is less pure but nonetheless suitable in quality for use in many molecular biology manipulations.

PREPARATION OF PLANT DNA USING CSCL CENTRIFUGATION

Plant cells are lysed by the detergent *N*-lauroylsarcosine (Sarkosyl), and the lysate is digested with proteinase K. After clearing insoluble debris from the lysate, nucleic acids are precipitated and DNA is purified on a cesium chloride (CsCl) density gradient.

**BASIC
PROTOCOL**

Materials

For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

Plant tissue, fresh
Liquid nitrogen
Extraction buffer (see recipe)
10% (w/v) *N*-lauroylsarcosine (Sarkosyl)
Isopropanol
TE buffer, pH 8.0
Cesium chloride
10 mg/ml ethidium bromide
CsCl-saturated isopropanol (equilibrate over a CsCl-saturated aqueous phase)
Ethanol
3 M sodium acetate, pH 5.2
250-ml centrifuge bottle
55°C water bath
Beckman JA-14, JA-20 or JA-21, and VTi80 rotors (or equivalents)
5-ml quick-seal ultracentrifuge tubes
15-G needle and 1-ml syringe

Prepare plant tissue

1. Harvest 10 to 50 g fresh plant tissue.

Plants may be placed in the dark for 1 to 2 days prior to harvest to reduce the starch content in the tissues.

Younger plants are the preferred source of tissue because they have a lower polysaccharide content.

2. Rinse tissue with deionized water to remove adhering debris and blot dry.
3. Freeze tissue with liquid nitrogen and grind to a fine powder in a mortar and pestle.

Keep the tissue frozen throughout this procedure by occasionally adding liquid nitrogen.

**Preparation and
Analysis of DNA**

2.3.1

Contributed by Eric Richards, Mark Reichardt, and Sharon Rogers

Current Protocols in Molecular Biology (1994) 2.3.1-2.3.7

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 27

Lyse and digest cells

4. Transfer frozen powder to a 250-ml centrifuge bottle and immediately add 5 to 10 ml extraction buffer per gram fresh plant tissue. Stir gently to disperse tissue.
5. Add 10% *N*-lauroylsarcosine to a final concentration of 1%. Incubate 1 to 2 hr at 55°C.

*It is important to add N-lauroylsarcosine **after** the tissue is resuspended in extraction buffer. If N-lauroylsarcosine is included in extraction buffer, premature lysis of the plant cells will interfere with tissue dispersal and lead to unwanted shearing of DNA.*

The lysate should be clear, green, and slightly viscous. From this point on solutions should be handled gently to reduce shearing of the DNA—use a wide-bore pipet and do not vortex or mix vigorously.

6. Centrifuge lysate 10 min at $5500 \times g$ (6000 rpm in a Beckman JA-14 rotor), 4°C, to pellet debris. Save the supernatant and centrifuge again if necessary to remove undigested debris.

Precipitate the DNA

7. Add 0.6 vol isopropanol to the supernatant and gently mix. A nucleic acid precipitate should be visible; if not, incubate 30 min at -20°C .
8. Centrifuge 15 min at $7500 \times g$ (8000 rpm in a Beckman JA-14 rotor), 4°C. Discard supernatant.

Do not let the nucleic acid pellet dry or it will become extremely difficult to dissolve.

Carry out CsCl centrifugation

9. Resuspend pellet in 9 ml TE buffer. If necessary, incubate at 55°C to aid resuspension. Add 9.7 g of solid CsCl and mix gently until dissolved.

To minimize depurination, limit 55°C incubation to ≤ 2 hr.

10. Incubate 30 min on ice. Centrifuge 10 min at $7500 \times g$ (8000 rpm in a JA-20 rotor), 4°C, and save supernatant.

This clearing spin removes some of the insoluble debris remaining in the lysate. In addition, a small separate phase may form on the top of the solution after centrifugation; this is due to residual Sarkosyl in the lysate. The Sarkosyl phase can be removed by filtering the supernatant through two layers of cheesecloth. Collect the supernatant but discard the Sarkosyl phase.

11. Add 0.5 ml of 10 mg/ml ethidium bromide and incubate 30 min on ice.

CAUTION: Ethidium bromide is a mutagen. Be careful and wear gloves.

12. Centrifuge 10 min at $7500 \times g$, 4°C.

A large RNA pellet should form. At this point much of the unwanted constituents in the lysate—RNA, protein, and carbohydrates—have been removed.

13. Transfer the supernatant to two 5-ml quick-seal ultracentrifuge tubes and seal tubes.

Make sure tubes are full, balanced, and well-sealed.

14. Centrifuge 4 hr at $525,000 \times g$ (80,000 rpm in a Beckman VTi80 rotor), 20°C, or overnight at $300,000 \times g$ (60,000 rpm in VTi80 rotor), 20°C.

Collect and purify DNA

15. Gently remove the tube. Punch a hole in the top (to provide an air inlet) with a large-bore (15-G) collecting needle. Recover the DNA band by inserting needle, attached to a 1-ml syringe, through tube wall directly below the band (see Fig. 1.7.1).

This operation is identical to that used during plasmid purification, except that only one band should be visible.

CAUTION: If UV illumination is used to visualize the DNA, wear UV protective glasses or a face shield. Minimize exposure of gradient to visible light to reducing nicking of DNA caused by ethidium bromide.

16. Remove the ethidium bromide by repeatedly extracting the collected DNA with CsCl-saturated isopropanol.
17. Add 2 vol water and 6 vol ethanol to the DNA solution and mix. Incubate 1 hr at -20°C .

DNA may precipitate immediately as a single white mass; it can be collected using a Pasteur pipet with a hook introduced at the tip or by brief centrifugation.

18. Centrifuge 10 min at $7500 \times g$, 4°C .
19. Resuspend pellet in TE buffer and reprecipitate DNA by adding $\frac{1}{10}$ vol of 3 M sodium acetate and 2 vol ethanol. Incubate at -20°C if precipitate is not visible and collect DNA by centrifugation.
20. Briefly air dry the final pellet and resuspend in 0.5 to 2 ml TE buffer.

A DNA concentration of 100 ng/ μl is generally convenient for most purposes.

PREPARATION OF PLANT DNA USING CTAB

Alternatively, the nonionic detergent cetyltrimethylammonium bromide (CTAB) is used to liberate and complex with total cellular nucleic acids. This general procedure has been used on a wide array of plant genera and tissue types. Many modifications have been published to optimize yields from particular species. The protocol is relatively simple, fast, and easily scaled from milligram to grams of tissue; it requires no cesium chloride density gradient centrifugation.

Additional Materials

*For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.*

CTAB extraction solution (see recipe)

CTAB/NaCl solution (*UNIT 2.4*)

CTAB precipitation solution (see recipe)

2% (v/v) 2-mercaptoethanol (2-ME)

High-salt TE buffer (see recipe)

24:1 (v/v) chloroform/octanol or chloroform/isoamyl alcohol

80% ethanol

Pulverizer/homogenizer: mortar and pestle, blender, Polytron (Brinkmann),
or coffee grinder

Organic solvent-resistant test tube or beaker

65°C water bath

Beckman JA-20 rotor or equivalent *or* microcentrifuge

ALTERNATE PROTOCOL

Preparation and Analysis of DNA

2.3.3

Extract nucleic acids

1. Add 2-ME to the required amount of CTAB extraction solution to give a final concentration of 2% (v/v). Heat this solution and CTAB/NaCl solution to 65°C.

Approximately 4 ml of 2-ME/CTAB extraction solution and 0.4 to 0.5 ml CTAB/NaCl solution are required for each gram of fresh leaf tissue. With lyophilized, dehydrated, or dry tissues such as seeds, 2-ME/CTAB extraction solution should be diluted 1:1 with sterile water. 2-ME should be used in a fume hood.

2. Chill a pulverizer/homogenizer with liquid nitrogen (−196°C) or dry ice (−78°C). Pulverize plant tissue to a fine powder and transfer the frozen tissue to an organic solvent-resistant test tube or beaker.

Use young tissue and avoid larger stems and veins to achieve the highest DNA yield with minimal polysaccharide contamination.

3. Add warm 2-ME/CTAB extraction solution to the pulverized tissue and mix to wet thoroughly. Incubate 10 to 60 min at 65°C with occasional mixing.

A 60-min incubation results in larger DNA yields. If maximum yield is not important, 10 min should be adequate. If the tissue contains large amounts of phenolic compounds, 1% (v/v) polyvinylpyrrolidone (mol. wt. = 40,000) may be added to absorb them.

4. Extract the homogenate with an equal volume of 24:1 chloroform/octanol or chloroform/isoamyl alcohol. Mix well by inversion. Centrifuge 5 min at 7500 × g (8000 rpm in JA-20; ~10,000 rpm in a microcentrifuge, for smaller samples), 4°C. Recover the top (aqueous) phase.

Octanol, rather than isoamyl alcohol, is used because it may enhance isolation of nuclei. Slower centrifugation speeds are possible if centrifugation time is increased accordingly; a microcentrifuge may be used for small-scale preparations (≤150 mg starting tissue). After centrifugation, two phases should be evident with tissue debris at the interface.

5. Add 1/10 vol 65°C CTAB/NaCl solution to the recovered aqueous phase and mix well by inversion.
6. Extract with an equal volume of chloroform/octanol. Mix, centrifuge, and recover as in step 4 above.

The aqueous phase may still be light yellow-brown in color.

Precipitate nucleic acids

7. Add exactly 1 vol CTAB precipitation solution. Mix well by inversion. If precipitate is visible, proceed to step 8. If not, incubate mixture 30 min at 65°C.
8. Centrifuge 5 min at 500 × g (2000 rpm in JA-20; ~2700 rpm in microcentrifuge), 4°C.

Do not increase the speed or time of centrifugation as the pellet may become very difficult to resuspend. If there is no pellet, add more CTAB precipitation solution (up to 1/10 the total volume). Incubate 1 hr to overnight at 37°C. Centrifuge 5 min at 500 × g, 4°C.

9. Remove but do not discard the supernatant and resuspend pellet in high-salt TE buffer (0.5 to 1 ml per gram of starting material). If the pellet is difficult to resuspend, incubate 30 min at 65°C. Repeat until all or most of pellet is dissolved.

Polysaccharide contamination may make it excessively difficult to resuspend the pellet. Read the A₂₆₀ of the supernatant and discard pellet if nucleic acids are present in the supernatant.

10. Precipitate the nucleic acids by adding 0.6 vol isopropanol. Mix well and centrifuge 15 min at $7500 \times g$, 4°C .

Ethanol can be used for the precipitation, but isopropanol may yield cleaner pellets.

11. Wash the pellet with 80% ethanol, dry, and resuspend in a minimal volume of TE (0.1 to 0.5 ml per gram of starting material).

Residual CTAB is soluble and is removed by the 80% ethanol wash. Further purification of the DNA with RNase A and proteinase K may be done using standard methods (UNITS 2.4 & 3.13).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

CTAB extraction solution

2% (w/v) CTAB
100 mM Tris·Cl, pH 8.0
20 mM EDTA, pH 8.0
1.4 M NaCl
Store at room temperature (stable several years)

CTAB precipitation solution

1% (w/v) CTAB
50 mM Tris·Cl, pH 8.0
10 mM EDTA, pH 8.0
Store at room temperature (stable several years)

Extraction buffer

100 mM Tris·Cl, pH 8.0
100 mM EDTA, pH 8.0
250 mM NaCl
100 $\mu\text{g}/\text{ml}$ proteinase K (add fresh before use)
Store indefinitely at room temperature without proteinase K

High-salt TE buffer

10 mM Tris·Cl, pH 8.0
0.1 mM EDTA, pH 8.0
1 M NaCl
Store at room temperature (stable for several years)

COMMENTARY

Background Information

CsCl gradient purification

This protocol is an adaptation of common DNA isolation procedures—cell lysis by detergent, protease treatment, and CsCl gradient purification. Because whole cells are lysed, DNA purified using this protocol will correspond to both the nuclear genome and cytoplasmic (mitochondrial and chloroplast) genomes. Methods for purifying nuclear DNA—free of plastid and mitochondrial DNA contamination—have been described by Watson and Thompson (1986). In addition, a miniprep pro-

tol for isolation of total plant DNA has been described by Dellaporta et al. (1983); the miniprep protocol is similar to the protocol described here, except that it omits the CsCl gradient centrifugation.

CTAB purification

The alternate protocol, cetyltrimethylammonium bromide (CTAB) DNA isolation, was initially used in bacteria (Jones, 1953; UNIT 2.4) and later modified to obtain DNA from plants (Murray and Thompson, 1980). CTAB forms an insoluble complex with nucleic acids when the initial NaCl concentration is lowered to

~0.5 M (Rogers and Bendich, 1985). Polysaccharides, phenolic compounds, and other enzyme-inhibiting contaminants found in plant cells are efficiently removed in the supernatant because most do not precipitate under the conditions described. The nucleic acid-CTAB complex is only soluble in high salt; detergent is removed by raising the NaCl concentration and precipitating the nucleic acids. The residual CTAB is removed by washing the nucleic acid pellet with 80% ethanol; CTAB is more soluble in ethanol and is discarded with the wash solution.

The CTAB method of DNA isolation is widely used on plants because of its versatility. Total genomic DNA has been isolated from many genera of monocotyledons and dicotyledons (Murray and Thompson, 1980; Rogers and Bendich, 1985). Various types of tissue can be used, including whole seedlings, leaves, cotyledons, seeds/grains, endosperm, embryos, tissue culture callus, and pollen. In addition, milligram amounts of tissue can be used when sample size is limiting. The starting material can be lyophilized, dehydrated (even mummified), frozen, or fresh.

The authors' (M.R. and S.R.) laboratory has successfully used this protocol to isolate DNA from *Arabidopsis thaliana*, *Zea mays*, *Gossypium hirsutum*, *Flaveria* sp., *Linum usitatissimum*, *Petunia hybrida*, *Glycine* sp., *Nicotiana tabacum*, and *Lycopersicon esculentum*. In addition, each year *The Plant Molecular Biology Reporter* publishes several modifications to this general protocol optimized for species from which nucleic acid extraction is difficult.

Critical Parameters

The aim of any genomic DNA preparation technique is to isolate high-molecular-weight DNA of sufficient purity. Two factors affect the size of the DNA isolated: shear forces and nuclease activity. As noted in the protocols, lysates should be treated gently to minimize shear forces. Plant cells are rich in nucleases. To reduce nuclease activity, the tissue should be frozen quickly and thawed only in the presence of an extraction buffer that contains detergent and a high concentration of EDTA.

Plant DNA isolated using the basic protocol should be in the range of 50 kb in length, which is quite acceptable for most applications. *Arabidopsis* DNA isolated using this protocol can be digested with restriction enzymes and ligated efficiently into cloning vectors. However, in some cases it may be necessary to modify the steps in order to reduce contamination by poly-

saccharides, phenolics, and other compounds that interfere with DNA isolation. Polysaccharides pose the most common problem affecting plant DNA purity. These carbohydrates are difficult to separate from the DNA itself, and they inhibit many enzymes commonly used in cloning procedures. The recommended procedure for polysaccharide removal is chloroform extraction of lysates in the presence of 1% CTAB and 0.7 M NaCl, as described by Murray and Thompson (1980) and in the alternate protocol. If the DNA pellet obtained using the alternate protocol is excessively difficult to resuspend, it may be due to the presence of polysaccharides that were not removed during CTAB precipitation. The DNA should be soluble in TE buffer; passage of the solution over an anion-exchange column should remove much of the contamination (Fang et al., 1992; UNIT 2.1B). If phenolic compounds are a problem, 1% polyvinylpyrrolidone (mol. wt. = 40,000; Sigma) can be included during tissue homogenization to absorb them. Chloroform/octanol is preferred over chloroform/isoamyl alcohol for organic extractions because it has been reported to isolate nuclei more efficiently (Watson and Thompson, 1986). Finally, RNase A and proteinase K digestion followed by phenol extraction can be performed if further purification is required.

Anticipated Results

***CsCl* gradient purification.** Yields should be in the range of 10 to 40 μ g DNA (50-kb length) per gram of fresh plant tissue. Isolated DNA should digest well with restriction enzymes and be ligated efficiently into cloning vectors.

CTAB purification. Yields should be in the range of 100 to 500 μ g DNA per gram of fresh plant tissue. The greatest yields will always be obtained using the youngest, freshest tissue available. DNA \geq 50 kb can be obtained if care is taken not to shear it (by using wide bore pipettes and gentle mixing) and if nucleases are avoided (by keeping tissue frozen or lyophilized and thawing or rehydrating only in the presence of CTAB extraction solution). In addition to DNA, RNA is also efficiently liberated and purified by this method and can be separated from the DNA if desired (e.g., see UNIT 4.3 and Taylor and Powell, 1982).

Time Considerations

***CsCl* gradient purification.** Approximately 4 to 6 hr are required to work through the protocol to the point where the lysate is loaded

onto the CsCl gradient. The gradient can be centrifuged overnight at $300,000 \times g$ or 4 hr at $525,000 \times g$. Approximately 3 to 4 hr are required to process the banded DNA.

CTAB purification. This procedure should take between 2 and 6 hr depending on the quantity of starting material, desired purity, and yield.

Literature Cited

- Dellaporta, S.L., Wood, J., and Hicks, J.B. 1983. A plant DNA miniprep: Version II. *Plant Mol. Biol. Rep.* 1:19-21.
- Fang, G., Hammar, S., and Grumet, R. 1992. A quick and inexpensive method for removing polysaccharides from plant genomic DNA. *BioTechniques* 13:52-56.
- Jones, A.S. 1953. The isolation of bacterial nucleic acids using cetyltrimethylammonium bromide (cetavlon). *Biochim. Biophys. Acta.* 10:607-612.
- Murray, M.G. and Thompson, W.F. 1980. Rapid isolation of high-molecular-weight plant DNA. *Nucl. Acids Res.* 8:4321-4325.

Rogers, S.O. and Bendich, A.J. 1985. Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol. Biol.* 5:69-76.

Taylor, B. and Powell, A. 1982. Isolation of plant DNA and RNA. *BRL Focus* 4(3):4-6.

Watson, J.C. and Thompson, W.F. 1986. Purification and restriction endonuclease analysis of plant nuclear DNA. *Methods Enzymol.* 118:57-75.

Contributed by Eric Richards
(CsCl preparation)
Washington University
St. Louis, Missouri

Mark Reichardt and Sharon Rogers
(CTAB preparation)
Lakeside Biotechnology
Chicago, Illinois

Preparation of Genomic DNA from Bacteria

MINIPREP OF BACTERIAL GENOMIC DNA

Bacteria from a saturated liquid culture are lysed and proteins removed by digestion with proteinase K. Cell wall debris, polysaccharides, and remaining proteins are removed by selective precipitation with CTAB, and high-molecular-weight DNA is recovered from the resulting supernatant by isopropanol precipitation.

Materials

- TE buffer (APPENDIX 2)
- 10% sodium dodecyl sulfate (SDS)
- 20 mg/ml proteinase K (stored in small single-use aliquots at -20°C)
- 5 M NaCl
- CTAB/NaCl solution
- 24:1 chloroform/isoamyl alcohol
- 25:24:1 phenol/chloroform/isoamyl alcohol (UNIT 2.1)
- Isopropanol
- 70% ethanol

1. Inoculate a 5-ml liquid culture with the bacterial strain of interest. Grow in conditions appropriate for that strain (i.e., appropriate medium, drug selection, temperature) until the culture is saturated. This may take several hours to several days, depending on the growth rate.
2. Spin 1.5 ml of the culture in a microcentrifuge for 2 min, or until a compact pellet forms. Discard the supernatant.
3. Resuspend pellet in 567 μl TE buffer by repeated pipetting. Add 30 μl of 10% SDS and 3 μl of 20 mg/ml proteinase K to give a final concentration of 100 $\mu\text{g/ml}$ proteinase K in 0.5% SDS. Mix thoroughly and incubate 1 hr at 37°C .

The solution should become viscous as the detergent lyses the bacterial cell walls. There should be no need to predigest the bacterial cell wall with lysozyme.

4. Add 100 μl of 5 M NaCl and mix thoroughly.

This step is very important since a CTAB–nucleic acid precipitate will form if salt concentration drops below about 0.5 M at room temperature (Murray and Thompson, 1980). The aim here is to remove cell wall debris, denatured protein, and polysaccharides complexed to CTAB, while retaining the nucleic acids in solution.

5. Add 80 μl of CTAB/NaCl solution. Mix thoroughly and incubate 10 min at 65°C .
6. Add an approximately equal volume (0.7 to 0.8 ml) of chloroform/isoamyl alcohol, mix thoroughly, and spin 4 to 5 min in a microcentrifuge.

This extraction removes CTAB–protein/polysaccharide complexes. A white interface should be visible after centrifugation.

7. Remove aqueous, viscous supernatant to a fresh microcentrifuge tube, leaving the interface behind. Add an equal volume of phenol/chloroform/isoamyl alcohol, extract thoroughly, and spin in a microcentrifuge for 5 min.

With some bacterial strains the interface formed after chloroform extraction is not compact enough to allow easy removal of the supernatant. In such cases, most of the interface can be fished out with a sterile toothpick before removal of any supernatant. Remaining CTAB precipitate is then removed in the phenol/chloroform extraction.

8. Transfer the supernatant to a fresh tube. Add 0.6 vol isopropanol to precipitate the nucleic acids (there is no need to add salt since the NaCl concentration is already

high). Shake the tube back and forth until a stringy white DNA precipitate becomes clearly visible. At this point it is possible to transfer the pellet to a fresh tube containing 70% ethanol by hooking it onto the end of a micropipet that has been heat-sealed and bent in a Bunsen flame. Alternatively, the precipitate can be pelleted by spinning briefly at room temperature.

If no stringy DNA precipitate forms in the above step, this implies that the DNA has sheared into relatively low-molecular-weight pieces. If this is acceptable, i.e., if DNA is to be digested to completion with restriction endonucleases for Southern blot analysis, chromosomal DNA can often still be recovered simply by pelleting the precipitate in a microcentrifuge.

9. Wash the DNA with 70% ethanol to remove residual CTAB and respin 5 min at room temperature to repellet it. Carefully remove the supernatant and briefly dry the pellet in a lyophilizer.
10. Redissolve the pellet in 100 μ l TE buffer.

This may take some time (up to 1 hr) since the DNA is of high molecular weight. 15 μ l of this DNA will typically digest to completion with 10 U EcoRI in 1 hr, which is sufficient to be clearly visible on an agarose gel, or to give a good signal during Southern hybridization.

REMOVAL OF POLYSACCHARIDES FROM EXISTING GENOMIC DNA PREPS

Steps 4 through 10 of the basic protocol can be adapted for removing polysaccharides and other contaminating macromolecules from existing bacterial chromosomal DNA preparations. Simply adjust the NaCl concentration of the DNA solution to 0.7 M and add 0.1 vol CTAB/NaCl solution. A white interface after the chloroform/isoamyl extraction indicates that contaminating macromolecules have been removed. The CTAB extraction step (steps 5 and 6) can be repeated several times until no interface is visible.

**SUPPORT
PROTOCOL**

MINIPREP OF BACTERIAL GENOMIC DNA

1. Grow bacterial strain to saturation.
2. Spin 1.5 ml for 2 min in microcentrifuge.
3. Resuspend in 567 μ l TE buffer, 3 μ l of 10% SDS, 3 μ l of 20 mg/ml proteinase K. Mix and incubate 1 hr at 37°C.
4. Add 100 μ l of 5 M NaCl. Mix thoroughly.
5. Add 80 μ l of CTAB/NaCl solution. Mix. Incubate 10 min at 65°C.
6. Extract with an equal volume of chloroform/isoamyl alcohol. Spin 5 min in microcentrifuge.
7. Transfer aqueous phase to a fresh tube. Extract DNA with phenol/chloroform/isoamyl alcohol. Spin 5 min in microcentrifuge.
8. Transfer aqueous phase to a fresh tube. Extract DNA with 0.6 vol isopropanol. Wash precipitate with 70% ethanol. Remove supernatant and briefly dry pellet in lyophilizer.
9. Resuspend pellet in 100 μ l TE buffer.

**SHORT
PROTOCOL**

**Preparation and
Analysis of DNA**

2.4.2

LARGE-SCALE CsCl PREP OF BACTERIAL GENOMIC DNA

This procedure is essentially a scale-up of the chromosomal miniprep described in the basic protocol, followed by additional purification on a cesium chloride gradient. This procedure may be used if large amounts of exceptionally clean genomic DNA are required, e.g., for the construction of genomic libraries.

Additional Materials

Cesium chloride
10 mg/ml ethidium bromide
CsCl-saturated isopropanol or H₂O-saturated butanol
3 M sodium acetate, pH 5.2

Beckman JA-20 rotor or equivalent
50-ml Oak Ridge centrifuge tubes
Wide-bored pipet
4-ml sealable centrifuge tubes
Beckman VTi80 rotor
3-ml plastic syringe with 15-G needle

Preparation and lysis of cells

1. Grow 100 ml culture of bacterial strain to saturation.
2. Pellet cells for 10 min at 4000 × g (e.g., in a Beckman JA-20 rotor at 6000 rpm). Discard supernatant.

This, and the following steps, can be conveniently carried out using 50-ml Oak Ridge centrifuge tubes.

3. Resuspend cells gently in 9.5 ml TE buffer. Add 0.5 ml of 10% SDS and 50 μl of 20 mg/ml proteinase K. Mix thoroughly and incubate 1 hr at 37°C.

Precipitation and purification of DNA

4. Add 1.8 ml of 5 M NaCl and mix thoroughly.
5. Add 1.5 ml CTAB/NaCl solution. Mix thoroughly and incubate 20 min at 65°C.
6. Add an equal volume of chloroform/isoamyl alcohol. Extract thoroughly. Spin 10 min at 6000 × g (JA-20 rotor at 7000 rpm), room temperature, to separate phases.
7. Transfer aqueous supernatant to a fresh tube using a wide-bored pipet.

The supernatant will probably be very viscous if the yield is high. An additional chloroform/isoamyl alcohol extraction, or a phenol/chloroform/isoamyl alcohol extraction, is optional but should not be necessary if the material is to be purified on a cesium chloride gradient.

8. Add 0.6 vol isopropanol and mix until a stringy white DNA pellet precipitates out of solution and condenses into a tight mass. Transfer the precipitate to 1 ml of 70% ethanol in a fresh tube, by hooking it on the end of a Pasteur pipet that has been bent and sealed in a Bunsen flame.
9. Spin the pellet 5 min at 10,000 × g (JA-20 rotor at 9900 rpm). Remove supernatant and redissolve the pellet in 4 ml TE buffer. This may take several hours to overnight—the DNA can be placed at 60°C to hasten the process.
10. Measure the DNA concentration on a spectrophotometer. Adjust concentration to 50 to 100 μg/ml. This will give 200 to 400 μg chromosomal DNA per 4 ml gradient.

It is not advisable to spin larger quantities of chromosomal DNA on such a small gradient.

11. Add 4.3 g CsCl per 4 ml TE buffer. Dissolve. Add 200 μ l of 10 mg/ml ethidium bromide. Transfer to 4-ml sealable centrifuge tubes. Adjust volume and balance tubes with CsCl in TE buffer (1.05 g/ml). Seal tubes. Spin 4 hr in a Beckman VTi80 rotor at 70,000 rpm, 15°C, or overnight at 55,000 rpm, 15°C.
12. Visualize gradient under longwave UV lamp. A single band should be visible. Remove band using a 15-g needle and a 3-ml plastic syringe.

If the DNA is intact high-molecular-weight chromosomal DNA it will appear very viscous as the band is withdrawn from the gradient; hence, it is important to use a wide-bore needle to avoid mechanical shearing of the DNA. If the band appears right at the top of the gradient, then the gradient is too dense. Reduce the amount of CsCl added in step 11.
13. Remove the ethidium bromide by sequential extractions with CsCl-saturated isopropanol or water-saturated butanol, as described in *UNIT 1.7*.
14. Dialyze overnight against 2 liters TE buffer to remove CsCl.
15. Transfer DNA solution to a fresh tube. If required, precipitate chromosomal DNA as described above (steps 8 and 9) by adding $\frac{1}{10}$ vol of 3 M sodium acetate and 0.6 vol isopropanol, and resuspend at desired concentration.

LARGE-SCALE CsCl PREP OF BACTERIAL GENOMIC DNA

1. Grow 100 ml culture of bacterial strain to saturation.
2. Spin 10 min at 4000 \times g.
3. Resuspend pellet in 9.5 ml TE buffer, 0.5 ml of 10% SDS, and 50 μ l of 20 mg/ml proteinase K. Mix and incubate 1 hr at 37°C.
4. Add 1.8 ml of 5 M NaCl. Mix thoroughly.
5. Add 1.5 ml CTAB/NaCl solution and mix. Incubate 20 min at 65°C.
6. Extract with an equal volume of chloroform/isoamyl alcohol. Spin 10 min at 6000 \times g, room temperature.
7. Transfer aqueous phase to a fresh tube. Extract with phenol/chloroform/isoamyl alcohol if necessary. Spin as in step 6.
8. Transfer aqueous phase to a fresh tube. Precipitate DNA with 0.6 vol isopropanol. Wash precipitate with 70% ethanol. Remove supernatant and resuspend pellet in 4 ml TE buffer.
9. Measure DNA concentration. Adjust concentration to give 50 to 100 μ g/ml. Add 4.3 g CsCl per 4 ml TE buffer. Add 200 μ l of 10 mg/ml ethidium bromide. Transfer to sealable centrifuge tubes. Spin 4 hr at 70,000 rpm, 15°C.
10. Visualize gradient with UV light. Remove band.
11. Extract ethidium bromide with CsCl-saturated isopropanol.
12. Dialyze overnight against 2 liters TE buffer.

SHORT PROTOCOL

Preparation and Analysis of DNA

2.4.4

REAGENTS AND SOLUTIONS

CTAB/NaCl solution (10% CTAB in 0.7 M NaCl)

Dissolve 4.1 g NaCl in 80 ml water and slowly add 10 g CTAB (hexadecyltrimethyl ammonium bromide) while heating and stirring. If necessary, heat to 65°C to dissolve. Adjust final volume to 100 ml.

COMMENTARY

Background Information

Most commonly used protocols for the preparation of bacterial genomic DNA consist of lysozyme/detergent lysis, followed by incubation with a nonspecific protease and a series of phenol/chloroform/isoamyl alcohol extractions prior to alcohol precipitation of the nucleic acids (Meade et al., 1984; Silhavy et al., 1982). Such procedures effectively remove contaminating proteins, but are not effective in removing the copious amounts of exopolysaccharides that are produced by many bacterial genera, and which can interfere with the activity of molecular biological enzymes such as restriction endonucleases and ligases. In this procedure, however, the protease incubation is followed by a CTAB extraction whereby CTAB complexes both with polysaccharides and with residual protein; both groups of contaminating molecules are effectively removed in the subsequent emulsification and extraction with chloroform/isoamyl alcohol. This procedure is effective in producing digestible chromosomal DNA from a variety of gram-negative bacteria, including those of the genera *Pseudomonas*, *Agrobacterium*, *Rhizobium*, and *Bradyrhizobium*, all of which normally produce large amounts of polysaccharides. If large amounts of exceptionally clean DNA are required, the procedure can be scaled up and the DNA purified on a cesium chloride gradient, as described in the alternate protocol. The method can also be used to extract high-molecular-weight DNA from plant tissue (Murray and Thompson, 1980).

Critical Parameters

The most critical parameter is the salt (NaCl) concentration of the solution containing the

lysed bacteria prior to adding CTAB. If the NaCl concentration is <0.5 M then the nucleic acid may also precipitate; indeed, CTAB is frequently used for just this process (Murray and Thompson).

It is also important to maintain all solutions above 15°C, as the CTAB will precipitate below this temperature.

Anticipated Results

The typical yield from both the miniprep and the large-scale prep is 0.5 to 2 mg DNA per 100 ml starting culture (10^8 to 10^9 cells/ml).

Time Considerations

The miniprep takes ~2 hr, including the 1-hr incubation. The large-scale prep takes slightly longer to reach the point where the CsCl gradients loaded. Subsequent steps will spread over 1 to 2 days, depending on the time of the CsCl gradient spin.

Literature Cited

- Meade, H.M., Long, S.R., Ruvkun, C.B., Brown, S.E., and Ausubel, F.M. 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. *J. Bacteriol.* 149:114-122.
- Murray, M.G. and Thompson, W.F. 1980. Rapid isolation of high-molecular-weight plant DNA. *Nucl. Acids Res.* 8:4321-4325.
- Silhavy, T.J., Berman, M.L., and Enquist, L.W. 1984. Experiments with Gene Fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Contributed by Kate Wilson
Australian Institute of Marine Science
Townsville, Australia

RESOLUTION AND RECOVERY OF LARGE DNA FRAGMENTS

This section describes the application of agarose gel electrophoresis to both analytical and preparative separation of DNA fragments. Standard agarose gels separate DNA fragments from ~0.5 to 25 kb, whereas pulsed-field agarose gels resolve molecules from ~10 to >2000 kb. Descriptions of standard and pulsed-field agarose gel electrophoresis as well as parameters affecting resolution of large DNA fragments are presented in *UNITS 2.5A* and *2.5B*. Three different protocols that employ agarose gels to prepare DNA fragments are described in *UNIT 2.6*. A general description of gels as electric circuits can be found in the introduction to this chapter.

Agarose Gel Electrophoresis

UNIT 2.5A

Agarose gel electrophoresis is a simple and highly effective method for separating, identifying, and purifying 0.5- to 25-kb DNA fragments (see *UNIT 2.5B* for larger fragments). The protocol can be divided into three stages: (1) a gel is prepared with an agarose concentration appropriate for the size of DNA fragments to be separated; (2) the DNA samples are loaded into the sample wells and the gel is run at a voltage and for a time period that will achieve optimal separation; and (3) the gel is stained or, if ethidium bromide has been incorporated into the gel and electrophoresis buffer, visualized directly upon illumination with UV light.

RESOLUTION OF DNA FRAGMENTS ON STANDARD AGAROSE GELS

BASIC PROTOCOL

Materials

- Electrophoresis buffer (TAE or TBE; *APPENDIX 2*)
- Ethidium bromide solution (see recipe)
- Electrophoresis-grade agarose
- 10× loading buffer (see recipe)
- DNA molecular weight markers (see Fig. 2.5A.3)
- 55°C water bath
- Horizontal gel electrophoresis apparatus
- Gel casting platform
- Gel combs (slot formers)
- DC power supply

Preparing the gel

1. Prepare an adequate volume of electrophoresis buffer (TAE or TBE; see Critical Parameters, Electrophoresis buffers) to fill the electrophoresis tank and prepare the gel.

To facilitate visualization of DNA fragments during the run, ethidium bromide solution can be added to the electrophoresis buffer to a final concentration of 0.5 µg/ml. If buffer is prepared for the electrophoresis tank and the gel separately, be sure to bring both to an identical concentration of ethidium bromide.

CAUTION: *Ethidium bromide is a mutagen and potential carcinogen. Gloves should be worn and care should be taken when handling ethidium bromide solutions.*

2. Add the desired amount of electrophoresis-grade agarose to a volume of electrophoresis buffer sufficient for constructing the gel (see Table 2.5A.1). Melt the agarose in

Preparation and Analysis of DNA

2.5A.1

Contributed by Daniel Voytas

Current Protocols in Molecular Biology (2000) 2.5A.1-2.5A.9

Copyright © 2000 by John Wiley & Sons, Inc.

Table 2.5A.1 Appropriate Agarose Concentrations for Separating DNA Fragments of Various Sizes

Agarose (%)	Effective range of resolution of linear DNA fragments (kb)
0.5	30 to 1
0.7	12 to 0.8
1.0	10 to 0.5
1.2	7 to 0.4
1.5	3 to 0.2

a microwave oven or autoclave and swirl to ensure even mixing. Gels typically contain 0.8 to 1.5% agarose (see Critical Parameters, Agarose Concentration).

Melted agarose should be cooled to 55°C in a water bath before pouring onto the gel platform. This prevents warping of the gel apparatus. Gels are typically poured between 0.5 and 1 cm thick. Remember to keep in mind that the volume of the sample wells will be determined by both the thickness of the gel and the size of the gel comb.

3. Seal the gel casting platform if it is open at the ends. Pour in the melted agarose and insert the gel comb, making sure that no bubbles are trapped underneath the combs and all bubbles on the surface of the agarose are removed before the gel sets.

Most gel platforms are sealed by taping the open ends with adhesive tape. As an added measure to prevent leakage, hot agarose can be applied with a Pasteur pipet to the joints and edges of the gel platform and allowed to harden.

Loading and running the gel

4. After the gel has hardened, remove the tape from the open ends of the gel platform and withdraw the gel comb, taking care not to tear the sample wells.

Most gel platforms are designed so that 0.5 to 1 mm of agarose remains between the bottom of the comb and the base of the gel platform. This is usually sufficient to ensure that the sample wells are completely sealed and to prevent tearing of the agarose upon removal of the comb. Low percentage gels and gels made from low gelling/melting temperature agarose should be cooled at 4°C to gain extra rigidity and prevent tearing.

5. Place the gel casting platform containing the set gel in the electrophoresis tank. Add sufficient electrophoresis buffer to cover the gel to a depth of about 1 mm (or just until the tops of the wells are submerged). Make sure no air pockets are trapped within the wells.
6. DNA samples should be prepared in a volume that will not overflow the gel wells by addition of the appropriate amount of 10× loading buffer. Samples are typically loaded into the wells with a pipettor or micropipet. Care should be taken to prevent mixing of the samples between wells.

Be sure to include appropriate DNA molecular weight markers (see Fig. 2.5.3).

7. Be sure that the leads are attached so that the DNA will migrate into the gel toward the anode or positive lead. Set the voltage to the desired level, typically 1 to 10 V/cm of gel, to begin electrophoresis. The progress of the separation can be monitored by the migration of the dyes in the loading buffer.

CAUTION: To prevent electrical shocks, the gel apparatus should always be covered and kept away from heavily used work spaces.

- Turn off the power supply when the bromphenol blue dye from the loading buffer has migrated a distance judged sufficient for separation of the DNA fragments. If ethidium bromide has been incorporated into the gel, the DNA can be visualized by placing on a UV light source and can be photographed directly (see Support Protocol 2).

Gels that have been run in the absence of ethidium bromide can be stained by covering the gel in a dilute solution of ethidium bromide (0.5 $\mu\text{g}/\text{ml}$ in water) and gently agitating for 10 to 30 min. If necessary, gels can be destained by shaking in water for an additional 30 min. This serves to remove excess ethidium bromide which causes background fluorescence and makes visualization of small quantities of DNA difficult.

MINIGELS AND MIDIGELS

Small gels—minigels and midigels—can generally be run faster than larger gels and are often employed to expedite analytical applications. Because they use narrower wells and thinner gels, minigels and midigels also require smaller amounts of DNA for visualization of the separated fragments. Aside from a scaling down of buffer and gel volumes, the protocol for running minigels or midigels is similar to that described above for larger gels. Similarly, the parameters affecting the mobility of DNA fragments, discussed below, are the same for both large and small gels.

A number of manufacturers offer smaller versions of their large electrophoresis models. An important feature to consider when selecting a mini- or midigel apparatus is the volume of buffer held by the gel tank. As smaller gels are typically run at high voltages (>10 V/cm), electrophoresis buffers are quickly exhausted, and it is therefore advantageous to choose a gel apparatus with a relatively large buffer reservoir. Minigel boxes can also be easily constructed from a few simple materials (Maniatis et al., 1982). A small (e.g., 15 cm long \times 8 cm wide \times 4 cm high) plastic box can be equipped with male connectors and platinum wire electrodes at both ends to serve as a minimal gel tank (see Fig. 2.5A.1).

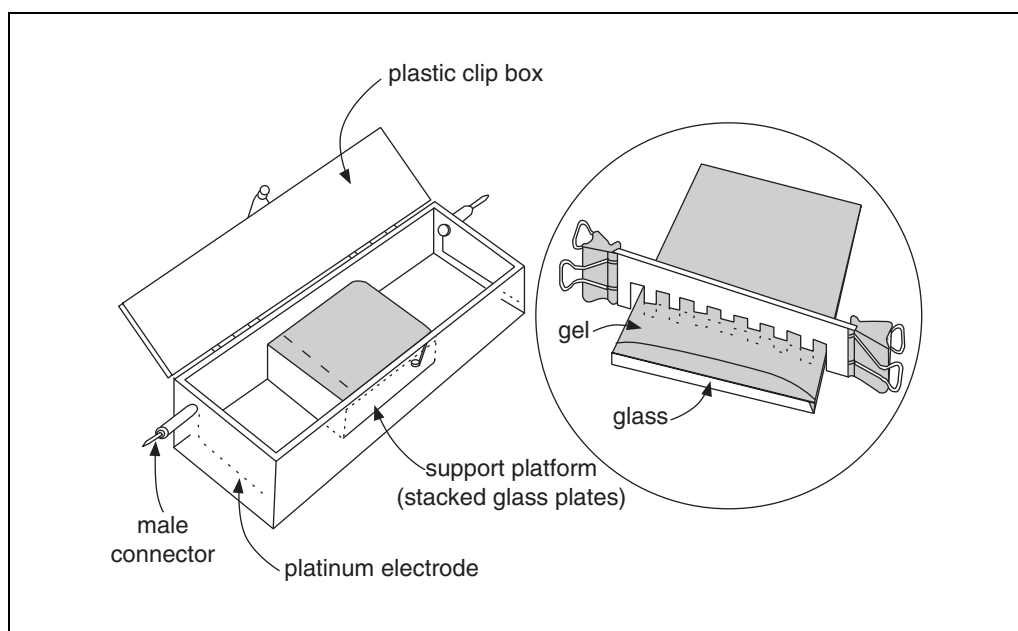


Figure 2.5A.1 Minigel apparatus.

SUPPORT PROTOCOL 1

Preparation and Analysis of DNA

2.5A.3

Although trays for casting small gels are commercially available, gels can also be poured onto glass lantern slides or other small supports without side walls. Such gels are held on the support simply by surface tension. After pouring the gel (e.g., 10 ml for a 5 cm × 8 cm gel), the comb is placed directly onto the support and held up by metal paper-binding clamps placed to the side.

Since there is no agarose at the bottom of such wells, extra care must be taken to prevent separation of the gel from the support when removing the comb.

PHOTOGRAPHY OF DNA IN AGAROSE GELS

DNA can be photographed in agarose gels stained with ethidium bromide by illumination with UV light ($>2500 \mu\text{W}/\text{cm}^2$). A UV transilluminator is typically used for this purpose, and commercial models are available designed specifically for DNA visualization and photography (e.g., UVP, Bio-Rad, Stratagene).

CAUTION: UV light is damaging to eyes and exposed skin. Protective eyewear should be worn at all times while using a UV light source.

The traditional means for photographing agarose gels uses a bellows-type camera equipped with a Polaroid film holder. An orange filter (Kodak Wratten #23A) is required to achieve a desirable image of light transmitted by fluorescing DNA. Polaroid type 667 film (ASA 3000) offers ideal sensitivity, allowing as little as several nanograms of DNA to be detected on film after making adjustments of exposure time. Gel photography systems can be set up by purchasing and assembling the individual components (e.g., camera, film holder, filters, and UV transilluminator). Alternatively, complete photography systems are commercially available (Bio-Rad, UVP).

A major disadvantage of traditional photography is the high cost of film. This can have a significant impact on a laboratory budget, particularly if large numbers of gels are processed on a routine basis. Because of this, traditional photography systems have been largely replaced with video imaging systems (e.g., Alpha Innotech Corp., Bio-Rad, Stratagene, and UVP; also see *UNIT 10.5*). Such systems use a CCD camera to capture the gel image and a video monitor to adjust focus and exposure time. A thermal printer generates inexpensive prints suitable for data documentation. A second advantage of imaging systems over traditional photography is that many systems come with video-processing computers and analysis software. This enables the degree of DNA fluorescence to be quantified (and thereby the relative amount of nucleic acid), and images can be stored for documentation and publication. Although the commercially available imaging systems are convenient, they can also be assembled on a less costly basis by purchasing the individual components and using public domain or commercially available software (Griess et al., 1995; Harriman and Wabl, 1995; Sutherland et al., 1987).

REAGENTS AND SOLUTIONS

Agarose gel

Gels typically contain ~1% agarose in 1× TAE or TBE (see *APPENDIX 2* for recipes; see Critical Parameters for choice of buffer). Electrophoresis-grade agarose powder is added to 1× gel buffer and melted by boiling for several minutes. Be sure all agarose particles are completely melted. To facilitate visualization of DNA fragments during the run, ethidium bromide can be added to 0.5 $\mu\text{g}/\text{ml}$ in the gel.

Ethidium bromide solution

1000× stock solution, 0.5 mg/ml:

50 mg ethidium bromide

100 ml H₂O

Working solution, 0.5 µg/ml:

Dilute stock 1:1000 for gels or stain solution

Protect from light.

10× loading buffer

20% Ficoll 400

0.1 M disodium EDTA, pH 8 (*APPENDIX 2*)

1.0% sodium dodecyl sulfate

0.25% bromphenol blue

0.25% xylene cyanol (optional; runs ~50% as fast as bromphenol blue and can interfere with visualization of bands of moderate molecular weight, but can be helpful for monitoring very long runs)

COMMENTARY

Background Information

Virtually all scientific investigations involving nucleic acids use agarose gel electrophoresis as a fundamental tool. Two papers that offer detailed descriptions of this technique and its practical use for DNA analysis are McDonnell et al. (1977) and Southern (1979).

Voltage applied at the ends of an agarose gel generates an electric field with a strength defined by the length of the gel and the potential difference at the ends (V/cm). DNA molecules exposed to this electric field migrate toward the anode due to the negatively charged phosphates along the DNA backbone. The migration velocity is limited by the frictional force imposed by the gel matrix. While charge and/or size can affect the rate at which macromolecules will pass through a gel, the charge to mass ratio is the same for DNA molecules of different lengths. It is the size of the DNA, therefore, that determines the rate at which it passes through the gel, thereby allowing an effective separation of DNA fragment-length mixtures by electrophoresis.

Critical Parameters

While agarose gel electrophoresis is straightforward and easy to perform, several factors should be considered before undertaking an electrophoretic separation. Parameters affecting the mobility of the DNA (i.e., applied voltage and agarose concentration) should be chosen for optimal resolution of the desired DNA fragment(s). The protocol should be designed such that the progress of the electrophoretic separation can be monitored and the

results accurately interpreted. This includes using such tools as tracking dyes and molecular weight markers.

Parameters affecting the migration of DNA through agarose gels

Agarose concentration. Molecules of linear, duplex DNA (form III) travel through gel matrices at a rate that is inversely proportional to the log₁₀ of their molecular weight (Helling et al., 1974). The molecular weight of a fragment of interest can therefore be determined by comparing its mobility to the mobility of DNA standards of known molecular weight. This is the most valuable feature of agarose gel electrophoresis, as it provides a reproducible and accurate means of characterizing DNA fragments by size.

Agarose concentration plays an important role in electrophoretic separations, as it determines the size range of DNA molecules that can be adequately resolved. Figure 2.5A.2A shows the effects of agarose concentration on mobility of fragments of different molecular weights. For most analyses, concentrations of 0.5% to 1.0% agarose are used to separate 0.5- to 30-kb fragments. However, low agarose concentrations (0.3 to 0.5%) are used to separate large DNA fragments (20 to 60 kb), and high agarose concentrations (1 to 1.5%) can resolve small DNA fragments (0.2 to 0.5 kb). See also Table 2.5A.1.

Applied voltage. In general, DNA fragments travel through agarose at a rate that is proportional to the applied voltage. With increasing voltages, however, large DNA molecules mi-

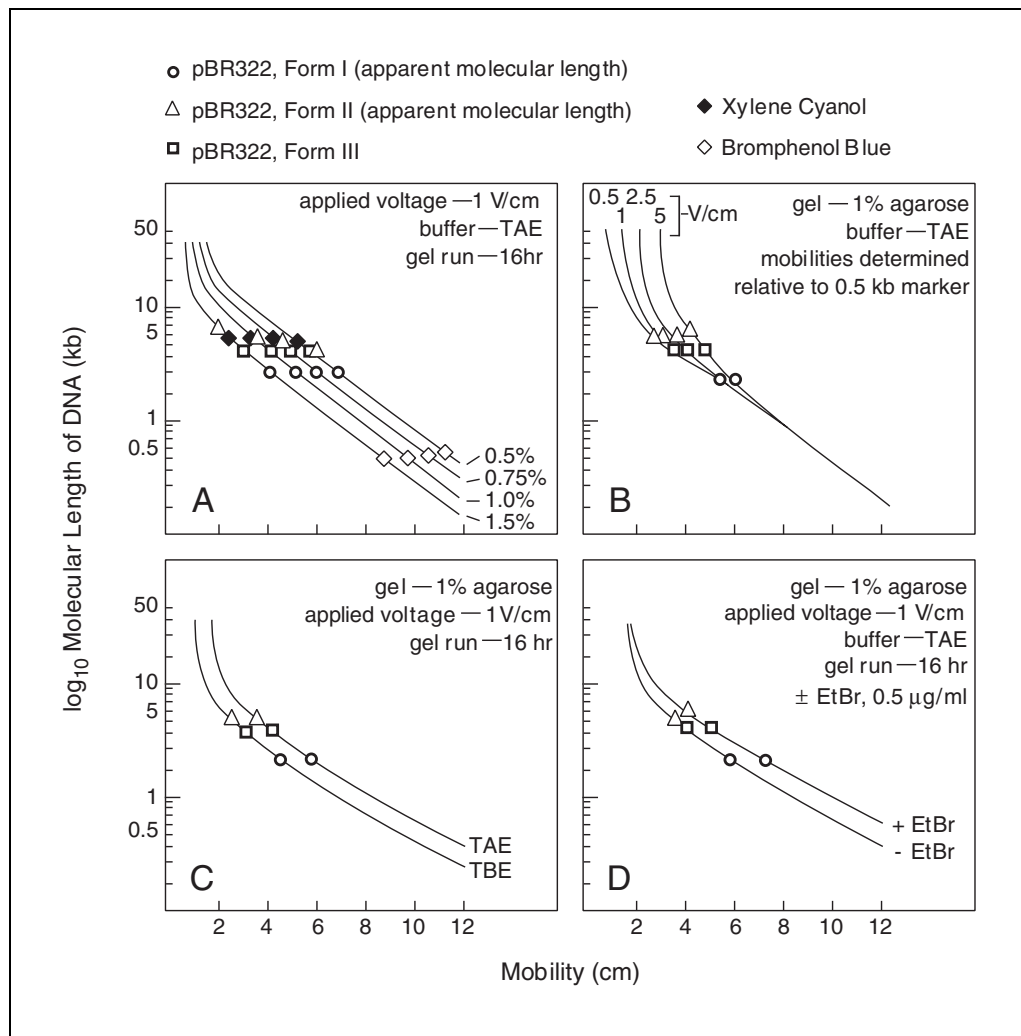


Figure 2.5A.2 Effect of (A) agarose concentration, (B) applied voltage, (C) electrophoresis buffer, and (D) ethidium bromide on migration of DNA through agarose gels.

grate at a rate proportionately faster than small DNA molecules. Consequently, higher voltages are significantly less effective in resolving large DNA fragments, as shown in Figure 2.5A.2B. For separating large DNA molecules, it is best to run gels at both low agarose concentrations and low applied voltages (~1 V/cm, 0.5% agarose).

Electrophoresis buffers. The two most widely used electrophoresis buffers are Tris/acetate (TAE) and Tris/borate (TBE). While these buffers have slightly different effects on DNA mobility (Figure 2.5A.2C), the predominant factor that should be considered in choosing between the two is their relative buffering capacity. Tris/acetate is the most commonly used buffer despite the fact that it is more easily exhausted during extended or high-voltage electrophoresis. Tris/borate has a significantly greater buffering capacity, but should be

avoided for purification of DNA from gels (see gel purification protocols).

DNA conformation. Closed circular (form I), nicked circular (form II), and linear duplex (form III) DNA of the same molecular weight migrate at different rates through agarose gels (Thorne, 1967). In the absence of ethidium bromide, closed circular supercoiled DNAs such as plasmids migrate faster than their linear counterpart DNAs (Fig. 2.5A.2D). Supercoiling essentially winds the molecules up, giving them a smaller hydrodynamic radius and allowing them to pass more readily through the gel matrix. Nicked or relaxed circular molecules that have lost all of their superhelicity migrate appreciably slower than either supercoils or linear molecules (Fig. 2.5A.2D; Johnson and Grossman, 1977).

The intercalating dye (ethidium bromide) is commonly incorporated into the gel and run-

ning buffer. The dye reduces the mobility of linear duplexes (Fig. 2.5A.2D) and has a particularly pronounced effect on the mobility of closed circular DNA. Ethidium bromide changes the superhelical density of closed circular molecules by inducing positive supercoils. With increasing concentrations of ethidium bromide, negative supercoils are gradually removed, causing a concomitant decrease in the mobility of the DNA molecule. This occurs until a critical free-dye concentration is reached where no more superhelical turns remain (usually between 0.1 to 0.5 $\mu\text{g}/\text{ml}$). As still more ethidium bromide is bound, positive superhelical turns are generated which, like negative supercoils, cause an increase in the electrophoretic mobility of the molecules (i.e., Fig. 2.5A.2D). By running gels at different concentrations of ethidium bromide, therefore, form I DNAs can easily be distinguished from other topoisomers.

Monitoring and interpreting separations of DNA through agarose gels

The DNA sample. In choosing the amount of DNA to be loaded, the width of the well plus the depth of the gel and the number and size of DNA fragments should be considered. Between 5 and 200 ng of a single DNA fragment can be loaded into a 0.5-cm-wide \times 0.2-cm-deep sample well; 5 ng approaches the minimal amount of an individual DNA fragment that can be detected by ethidium bromide staining, and 200 ng approximates the most that can be resolved before overloading occurs (the trailing and smearing characteristic of overloading become more pronounced with DNA fragments above 10 kb). For samples of DNA containing several fragments, between 0.1 and 0.5 μg of DNA is typically loaded per 0.5 cm sample well. Up to 10 μg of DNA can be adequately resolved for samples containing numerous fragments of different sizes (e.g., restriction digests of genomic DNA).

Tracking dyes. The most common means of monitoring the progress of an electrophoretic separation is by following the migration of tracking dyes that are incorporated into the loading buffer. Two widely used dyes displaying different electrophoretic mobilities are bromphenol blue and xylene cyanol. Xylene cyanol typically migrates with DNA fragments around 5 kb (Fig. 2.5A.2A) and bromphenol blue usually comigrates with DNA molecules around 0.5 kb (Fig. 2.5A.2A). Bromphenol Blue therefore provides an index of the mobility of the fastest fragments and is particularly

valuable in determining the length of the gel over which the separation of DNA has occurred. Xylene cyanol is useful for monitoring the progress of longer runs. Both dyes can interfere with the visualization of the fragments that comigrate with them.

Ethidium bromide. Ethidium bromide is commonly used for direct visualization of DNA in gels. The dye intercalates between the stacked bases of nucleic acids and fluoresces red-orange (560 nm) when illuminated with UV light (260 to 360 nm). This allows very small quantities of DNA to be detected (<5 ng) (Sharp et al., 1973).

Ethidium bromide is frequently added to the gel and running buffer prior to electrophoresis. While this has a slight effect on the mobility of the DNA (Fig. 2.5A.2D), it eliminates the need to stain the gel upon completion of the separation. An added advantage to running gels with ethidium bromide is that the mobility of the DNA can be monitored throughout the run until the desired separation is achieved.

Molecular weight markers. Among the samples loaded onto a gel, at least one lane should contain a series of DNA fragments of known sizes so that a standard curve can be constructed to allow the calculation of the sizes of unknown DNA fragments. The most commonly used molecular weight markers are restriction digests of phage λ DNA or, for smaller fragments, the plasmid pBR322. Figure 2.5A.3 shows the migration pattern and fragment sizes for restriction digests of λ DNA and pBR322 that are frequently used as molecular weight markers.

Aside from λ restriction fragments, many commercial preparations of molecular weight markers are also available. These products usually cover a wide range of DNA sizes, and some manufacturers also offer supercoiled (form I) markers for calculating plasmid sizes.

Troubleshooting

Common problems encountered in agarose gel electrophoresis are described below, along with several possible causes.

Poor resolution of DNA fragments. The most frequent cause of poor DNA resolution is improper choice of agarose concentration. Low percentage agarose gels should be used to resolve high-molecular-weight DNA fragments and high percentage gels for low-molecular-weight DNAs (see Table 2.5A.1). Fuzzy bands, encountered particularly with small DNA fragments, result from diffusion of the DNA through

the gel. This is especially true when gels are run for long periods of time at low voltages.

Band smearing. Trailing and smearing of DNA bands is most frequently observed with high-molecular-weight DNA fragments. This is often caused by overloading the DNA sample or running gels at high voltages. DNA samples loaded into torn sample wells will also cause extensive smearing, as the DNA will tend to run in the interface between the agarose and the gel support.

Melting of the gel. Melting of an agarose gel during an electrophoretic separation is a sign that either the electrophoresis buffer has been omitted in the preparation of the gel or has become exhausted during the course of the run. For high-voltage electrophoresis over long time periods, TBE should be used instead of TAE as it has a greater buffering capacity. Also, minigel and midigel boxes, which typically have small buffer reservoirs, tend to exhaust buffers more readily than larger gel boxes.

Anticipated Results

DNA fragments of ~0.5 to 25 kb are well resolved using the Basic Protocol. See following description of pulsed-field gel electrophoresis (UNIT 2.5B) for resolving molecules from ~10 to >2000 kb.

Time Considerations

Among the parameters that influence the length of time required to complete an electrophoretic separation, the one that has the greatest effect is the applied voltage. Most large agarose gels are run overnight (~16 hr) at voltages between 1 and 1.5 V/cm. While gels can be run much faster, particularly if the gel apparatus is cooled, resolution of larger DNA fragments is lost when higher voltages are used. Because the resolution required depends on the relative molecular weights of the fragments of interest, the time required for adequate separation is best determined empirically. As described above, this is most easily done by including ethidium bromide in the gel and then monitoring the progress of the run directly by visualization with UV light.

Mini- and midigels are typically run at high voltages relative to gel size (>10 V/cm) and are often completed in less than 1 hr. Two consequences of high-voltage runs should be kept in mind: First, as mentioned above, buffers become quickly exhausted and therefore TBE should be used, or gel tanks with large buffer capacities. Second, high voltages provide poor resolution of high-molecular-weight DNA fragments (see previous discussion). If large DNA fragments are to be separated, it may be advisable to use larger gels and/or lower applied voltages.

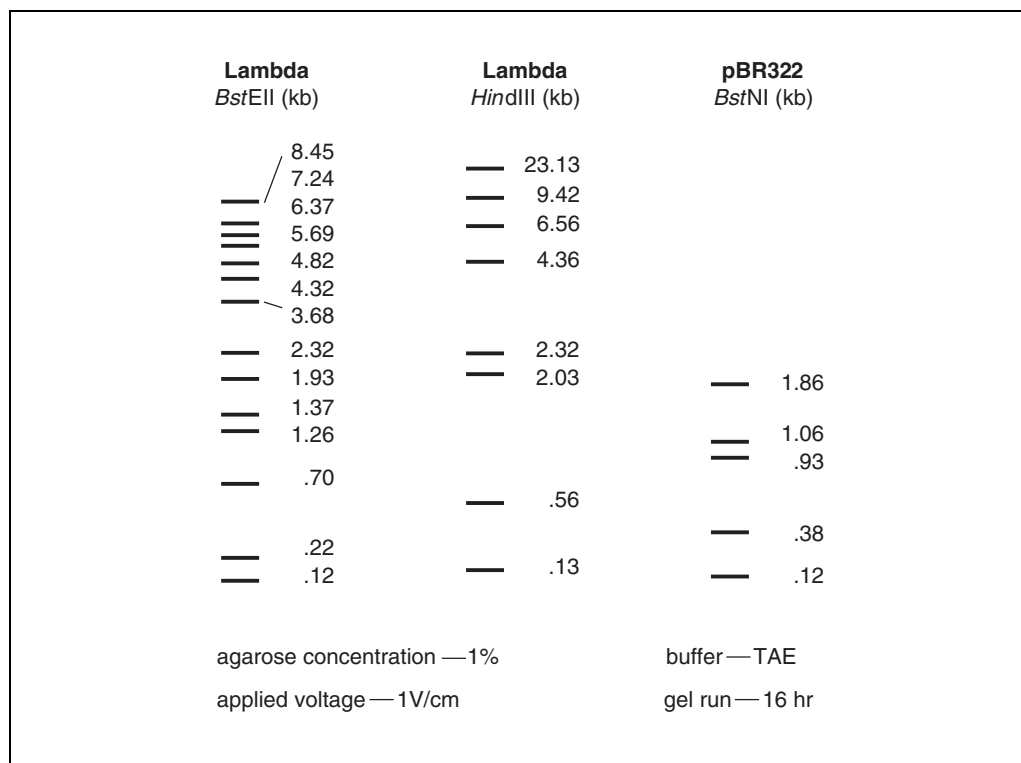


Figure 2.5A.3 Migration pattern and fragment sizes of common DNA molecular weight markers.

Literature Cited

- Griess, G.A., Louie, D., and Serwer, P. 1995. A desktop, low-cost video fluorometer for quantitation of macromolecules after gel electrophoresis. *Appl. Theor. Electrophor.* 4:175-177.
- Harriman, W.D., and Wabl, M. 1995. A video technique for the quantification of DNA in gels stained with ethidium bromide. *Anal. Biochem.* 228:336-342.
- Helling, R.B., Goodman, H.M., and Boyer, H.W. 1974. Analysis of R. *EcoRI* fragments of DNA from lambdoid bacteriophages and other viruses by agarose gel electrophoresis. *J. Virol.* 14:1235-1244.
- Johnson, P.H. and Grossman, L.I. 1977. Electrophoresis of DNA in agarose gels. Optimizing separations of conformational isomers of double- and single-stranded DNAs. *Biochemistry* 16:4217-4224.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. 1982. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- McDonnell, M.W., Simon, M.N., and Studier, F.W. 1977. Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. *J. Mol. Biol.* 110:119-146.
- Sharp, P.A., Sugden, B., and Sambrook, J. 1973. Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose. *Biochemistry* 12:3055-3063.
- Southern, E. 1979. Gel electrophoresis of restriction fragments. *Meth. Enzymol.* 68:152-176.
- Sutherland, J.C., Lin, B., Monteleone, D.C., Mugavero, J., Sutherland, B.M., and Trunk, J. 1987. Electronic imaging system for direct and rapid quantitation of fluorescence from electrophoretic gels: Application to ethidium bromide-stained DNA. *Anal. Biochem.* 163:446-457.
- Thorne, H.V. 1967. Electrophoretic characterization and fractionation of polyoma virus DNA. *J. Mol. Biol.* 24:203.

Contributed by Daniel Voytas
Iowa State University
Ames, Iowa

Pulsed-Field Gel Electrophoresis

UNIT 2.5B

DNA molecules longer than 25 kb are poorly resolved by standard agarose gel electrophoresis (UNIT 2.5A). These longer molecules can be resolved using several techniques that periodically change the direction of the electric field in the gel. The simplest and most generally useful of the pulsed-field techniques is field inversion electrophoresis (see Basic Protocol), which can be tuned to resolve molecules from ~10 to 2000 kb (or more with specialized equipment). To resolve molecules beyond the range of field inversion, it is necessary to use some sort of field-angle alternation electrophoresis such as CHEF (contour-clamped homogeneous electric field; see Alternate Protocol). High-molecular-weight DNA samples and size markers to be resolved by these techniques can be prepared while embedded in agarose blocks (see Support Protocol).

FIELD-INVERSION ELECTROPHORESIS

**BASIC
PROTOCOL**

Materials

1% agarose gel, standard or pulsed-field grade (e.g., SeaKem FastLane; FMC Bioproducts)

GTBE buffer (see recipe) *or* 0.5× TBE buffer (APPENDIX 2)

Samples embedded in agarose (Support Protocol), or liquid samples

Peristaltic pump (Cole-Parmer Masterflex or equivalent)

Programmable switching device (MJ Research PPI-200 or equivalent)

Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5A)

NOTE: Some power supplies have pulsed-DC rather than constant-voltage output and are unacceptable for pulsed-field gels. These can usually be recognized because their output is fixed, or adjustable in steps, rather than continuously variable.

1. Prepare a 1% agarose gel for a horizontal gel apparatus using GTBE or 0.5× TBE buffer.

Make the gel only as thick as necessary for the samples so that it will consume little power and heat up as little as possible.

Ethidium bromide can be incorporated into the gel, but is recommended only for gels resolving fragments <100 kb (see Critical Parameters).

2. Allow gel to set, then carefully remove comb. Insert into wells any samples that have been prepared in agarose blocks (see Support Protocol).

If blocks fit tightly into wells, it may be easiest to draw them down into the wells by inserting a pipettor with a 0.4-mm (or thinner) gel-loading tip into the back of the well to remove the air from under the blocks. If the blocks do not fit tightly, add melted agarose (55°C in gel buffer) to the well to hold the block in place.

3. Place gel into gel box, cover with buffer to a depth of 2 to 3 mm, and load any samples that are in liquid. To avoid shearing DNA >100 kb, cut ~5 mm off ends of pipet tips with a razor blade and pipet gently. At least one lane should contain bromphenol blue.
4. Adjust peristaltic pump for an appropriate flow (5 to 10 ml/min for a minigel and 20 to 50 ml/min for a large gel). Connect tubing ends to recirculation ports of gel box or place directly in buffer tanks.
5. Paying careful attention to polarity, connect programmable switching device to a constant-voltage DC power supply and connect gel apparatus to switching device.

**Preparation and
Analysis of DNA**

2.5B.1

Contributed by Michael Finney

Current Protocols in Molecular Biology (2000) 2.5B.1-2.5B.9

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 51

Set switching device for an appropriate switching regime but don't begin switching yet. Start gel running.

See Critical Parameters, including Table 2.5B.2, for a guide to time intervals, voltage, and other parameters. The most commonly used ratio of forward to reverse time is 3:1.

6. Allow bromphenol blue to migrate 1 cm, then start switching device and peristaltic pump.
7. Complete run and stain gel with ethidium bromide (UNIT 2.5A). Photograph as for a standard agarose gel (Fig. 2.5B.1).

The gel may be Southern blotted (UNIT 2.9A); note that the acid depurination step is essential for transfer.

ALTERNATE PROTOCOL

CHEF ELECTROPHORESIS

It is possible to resolve DNA molecules several million bases in length by periodically changing the angle of the electric field in the gel. There are a number of variations to the basic setup but all require specially constructed gel boxes, which can be quite expensive. In addition, some types of alternating-angle electrophoresis setups dissipate so much power in the gel that special cooling equipment must be used, adding to the expense.

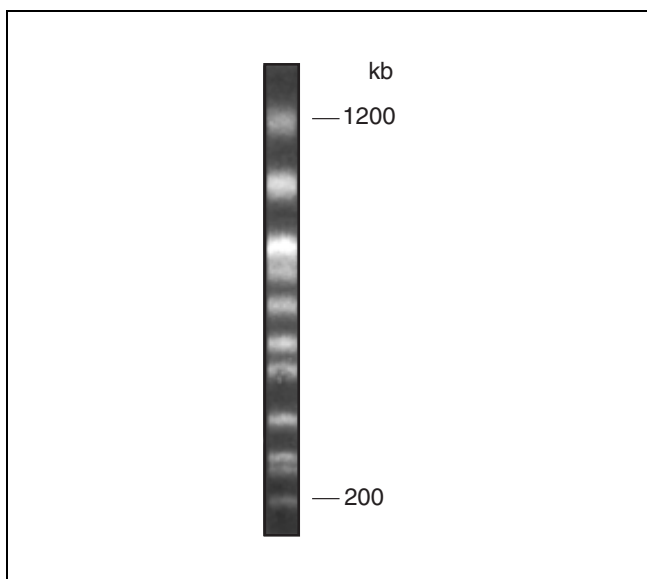
Running CHEF or other alternating-angle gels, described here, is very similar to running field-inversion gels, so the factors to be considered in that protocol apply to this one, as well. See Critical Parameters for a guide to estimating optimal conditions for running alternating-angle gels.

Additional Materials (also see Basic Protocol)

CHEF electrophoresis voltage-divider circuitry and gel box (see Background Information).

1. Prepare a 1% agarose gel for a CHEF gel apparatus using GTBE or 0.5× TBE buffer.
2. Allow gel to set, then carefully remove comb. Insert into wells any samples that have been prepared in agarose blocks (see Support Protocol).
3. Place gel into gel box, cover with buffer to a depth of 2 to 3 mm, adjust recirculation to ≥ 100 ml/min, and monitor buffer temperature. Wait 15 min after buffer has reached

Figure 2.5B.1 Chromosomes of *Saccharomyces cerevisiae* separated by field inversion.



desired running temperature to ensure that the gel has equilibrated at the correct temperature.

4. Load any samples that are in liquid.
5. Paying careful attention to polarity, connect programmable switching device to constant-voltage DC power supply, voltage divider circuitry, and gel apparatus. Set the switching device for an appropriate switching regime and start gel.
6. Complete run and stain gel with ethidium bromide (*UNIT 2.5A*). Photograph as for a standard agarose gel.

PREPARATION OF HIGH-MOLECULAR-WEIGHT DNA SAMPLES AND SIZE MARKERS

SUPPORT PROTOCOL

Very long DNA molecules are extremely fragile and cannot survive the standard manipulations of molecular biology. These molecules can, however, be prepared and manipulated easily while embedded in agarose blocks. This protocol describes the preparation of very high-molecular-weight DNA.

Materials

1% agarose

Sample to be prepared (e.g., tissue culture cells, nematode worms, nuclei, yeast, bacteria, or phage; Table 2.5B.1)

Lysis buffer (see recipe)

Storage buffer (see recipe)

400 mM phenylmethylsulfonyl fluoride (PMSF) in ethanol

10 mM Tris-Cl, pH 8.0 (*APPENDIX 2*)

Appropriate restriction enzyme and buffer (*UNIT 3.1*)

Block molds or petri plates

1. Prepare block molds (see Fig. 2.5B.2) by sealing one end with tape. If block molds are unavailable, samples can be prepared in agarose poured as a puddle on the bottom of a petri dish and blocks cut to size using a razor blade.
2. Suspend sample at room temperature in water or an appropriate buffer or medium at twice the desired final concentration.

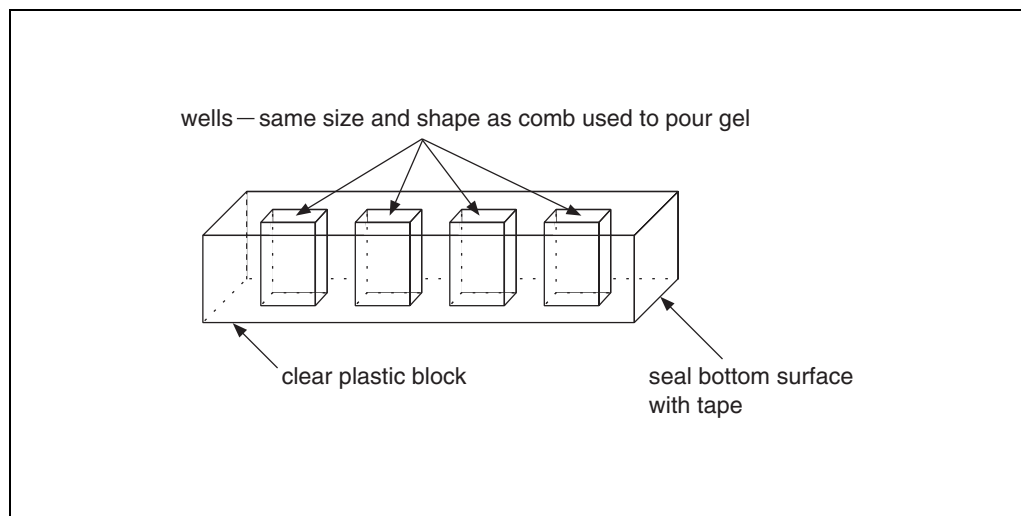


Figure 2.5B.2 Block molds for high-molecular-weight DNA samples. These can be made in the laboratory, or may be purchased from pulsed-field gel box manufacturers.

Preparation and Analysis of DNA

2.5B.3

Table 2.5B.1 Preparation of High-Molecular-Weight DNA Samples and Size Markers

Starting material	Preparation
Bacteria and phage	Resuspend bacteria (<i>UNIT 1.2</i>) or phage particles (<i>UNIT 1.13</i>) at a concentration calculated to yield the desired amount of DNA per lane; e.g., 5×10^8 <i>E. coli</i> per ml will yield ~100 ng DNA in an average lane.
Lambda ladders for size markers	Start with a concentrated stock of phage λ particles (<i>UNIT 1.13</i>). This procedure does not work well with some lots of commercial λ DNA, possibly because of damaged cohesive ends. Try several dilutions of phage stock to see which works best. The second incubation in lysis buffer should be done at 25° rather than 37°C; during this incubation, the cohesive ends of the molecules will anneal, giving multimers of varying lengths. These concatemers will stay together during electrophoresis provided that the gel is run at <25°C.
Nematodes	Anesthetize worms by resuspending in 10 mM NaN ₃ , then place in agarose.
Nuclei	Isolate as in <i>UNIT 4.10</i> . Approximately 1 μ g DNA is contained in 10 ⁵ mammalian nuclei.
Tissue culture cells	Cells should be washed several times in a medium containing no serum, as serum may inhibit proteinase K.
Yeast	<i>Saccharomyces cerevisiae</i> cells must have their cell walls removed before being embedded in agarose as described in <i>UNIT 13.13</i> , Basic Protocol.

The samples may also be suspended in the buffer in which they were originally prepared, or in a buffer such as TE buffer (APPENDIX 2). See Table 2.5B.1 for guidelines to preparing specific starting materials.

3. Add equal volume of 1% agarose, melted and cooled to 50°C. Quickly mix and aliquot solution into block mold. Let agarose solidify on ice.
4. Remove tape from mold and carefully push hardened blocks into a 50-ml conical tube containing at least 20 vol lysis buffer. Incubate overnight at 50°C, preferably with gentle agitation.
5. Pour off lysis buffer, add fresh lysis buffer, and incubate overnight at 37°C.
6. Pour off lysis buffer and replace with 20 vol storage buffer. Store at 4°C or proceed with steps 7 to 9 for sample to be digested with restriction enzymes.
7. Wash sample three times, 1 hr at room temperature each time, with at least 10 vol storage buffer supplemented with 1 mM PMSF.

CAUTION: PMSF is a powerful covalent inhibitor of proteases. It is both toxic and volatile and should always be handled in a hood. It is unstable in aqueous solution, so solutions should always be freshly prepared from a stock of 400 mM in ethanol stored at -20°C (PMSF will precipitate at -20°C and redissolve when warmed).

8. Wash sample three times, 30 min at room temperature each time, in at least 10 vol of 10 mM Tris·Cl, pH 8.0.

- Place sample in 1.5-ml microcentrifuge tube and remove excess liquid. Add an amount of 3× restriction buffer (containing the restriction enzyme) equal to half the volume of the block. Incubate at appropriate temperature.

For particular batches of samples and lots of enzyme, it may be necessary to titrate the amount of enzyme and digestion time to minimize DNA degradation.

REAGENTS AND SOLUTIONS

GTBE buffer

50 ml 10× TBE buffer (*APPENDIX 2*; 0.5× final)

50 ml 2 M glycine (0.1 M final)

900 ml H₂O

Store at room temperature

Lysis buffer

100 mM EDTA, pH 8.0 (*APPENDIX 2*)

10 mM Tris·Cl, pH 8.0 (*APPENDIX 2*)

1% (w/v) N-lauroylsarcosine sodium salt (Sarkosyl)

100 μg/ml proteinase K (add just before use from a 20 mg/ml stock)

Store at room temperature, *without proteinase K*

Storage buffer

10 mM Tris·Cl, pH 8.0 (*APPENDIX 2*)

10 mM EDTA, pH 8.0 (*APPENDIX 2*)

Store at room temperature

COMMENTARY

Background Information

In DNA gel electrophoresis, negatively charged DNA molecules are pulled through a gel by an electric field. An agarose gel presents a DNA molecule with a set of pores of varying sizes. Small molecules pass easily through most pores and move rapidly through the gel. Larger molecules have to “squeeze,” or change conformation, to get through the smaller pores and are slowed more than smaller molecules. The mobility of a molecule is related to the fraction of pores that it can easily move through, a process known as “sieving.”

DNA molecules larger than a certain size must squeeze to get through even the largest pores and cannot be sieved by the gel. Molecules in this size range all migrate at about the same rate, called “limiting mobility.” For ordinary agarose gel electrophoresis (*UNIT 2.5A*), limiting mobility generally occurs between 20 and 40 kb, depending on the exact conditions of the gel run. Schwartz and Cantor (1984) showed that it was possible to resolve molecules that would otherwise run at limiting mobility by cyclically varying the orientation of the electric field in the gel during the run.

The principle involved in alternating-angle electrophoresis is relatively simple. As a large molecule squeezes through a pore, it is forced into an extended conformation. It takes some time (from milliseconds to minutes, depending on the molecule’s size) to squeeze through the first pore but once the extended conformation is attained, the molecule can continue to squeeze rapidly through successive pores, moving at limiting mobility. If a large molecule moving through a gel is suddenly forced to change direction by a change in the orientation of the electric field, it must first change to a new conformation that will allow it to move in the new direction. The larger a molecule is, the longer it will take to change to the new conformation, and this time difference can be used to separate molecules. For example, although a 100-kb and a 200-kb molecule move through the gel matrix at the same rate in the steady state of limiting mobility, every time the direction of the field changes the 100-kb molecule will get a “head start” on the 200-kb molecule.

There are a number of implementations of alternating-angle electrophoresis. Schwartz and Cantor (1984) originally used an apparatus

that resulted in nonuniform electric fields. Nonuniform fields can focus bands, making them very sharp, but inevitably cause the lanes to run in various crooked shapes.

More practical systems use uniform electric fields, which result in gel lanes that run straight. Several methods of achieving uniform fields have been invented, including CHEF (contour-clamped homogeneous electric fields; Chu et al., 1986), and simply running the gel in a rectangular box and periodically rotating it. The angle between the two fields must be somewhat greater than 90° , and may be much greater.

In CHEF electrophoresis, the gel is surrounded by a set of electrodes whose voltages can be fixed in such a way that a uniform electric field is set up in the gel. Thus CHEF gels require two pieces of equipment not necessary for field inversion electrophoresis: a voltage-divider circuit to set the electrode voltages, and a multi-electrode gel box. Both are available commercially and plans have been published (Chu, 1989). Figure 2.5B.3 shows a popular version of CHEF voltage divider circuitry.

Alternating-angle electrophoresis using a fixed switching time gives optimum resolution in a single size range, but does separate DNA outside this size range, albeit with reduced resolution (see critical parameters). It is there-

fore advantageous to vary the pulse durations during a run.

Field-inversion electrophoresis employs the limiting value for the angle between two uniform fields, 180° (Carle et al., 1986). Analysis of field-inversion electrophoresis is somewhat more complicated than that of alternating-angle electrophoresis because different pulse times are used in the forward and reverse directions. Like alternating-angle electrophoresis, field inversion slows molecules that would otherwise run at limiting mobility by taking advantage of the time required for molecules to change conformation.

One difference between the two techniques is that in field inversion a single pulse duration separates only a relatively narrow range of sizes. To get resolution over a broad size range, it is necessary to use a range of reverse times; this is accomplished with a time ramp—i.e., progressively increasing the forward and reverse intervals from a lower limit to an upper limit. It is preferable to cycle repeatedly through the range of times rather than to use one long time ramp, as there is some hysteresis in the behavior of large molecules (the largest molecules will move more slowly during short intervals if they immediately follow longer intervals).

Varying the reverse interval also prevents a possible artifact of field inversion. If the reverse

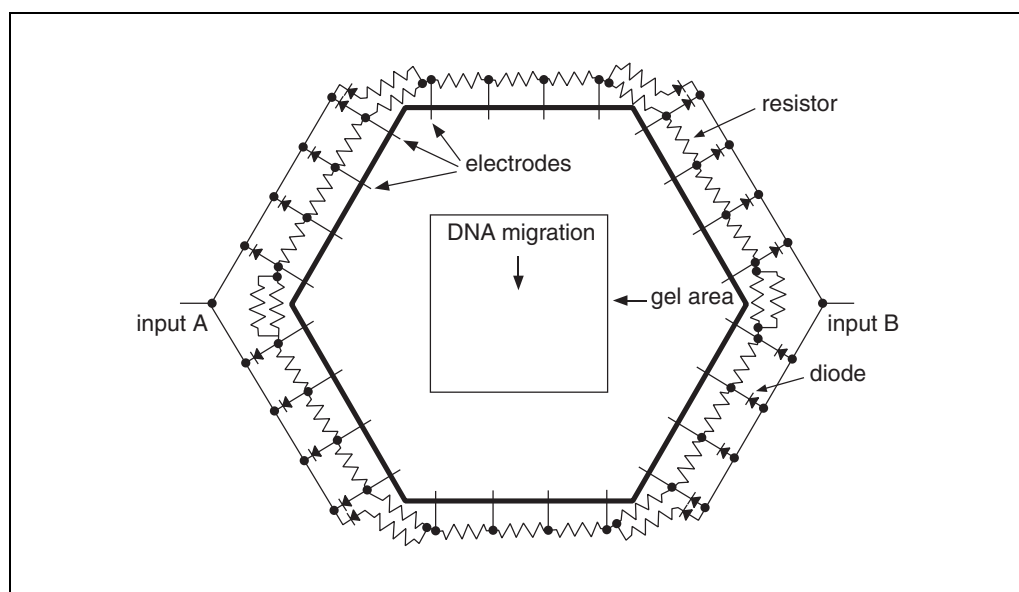


Figure 2.5B.3 Circuitry for clamping electrode voltages of a 24-electrode hexagonal CHEF gel (redrawn from Chu, 1989). Positions of the 24 electrodes are shown schematically, as well as the direction of the resulting DNA migration. This circuitry can be driven directly by an inverting gel controller connected to input A and input B. Power dissipation in the resistors limit this circuit to ~ 250 V input. All resistors are 470Ω , 1% tolerance, 3 W and diodes are type 1N4004 (1 A, 400 V).

time is not long enough, large molecules may not be sufficiently disrupted by the reversal, and may actually migrate faster than shorter molecules. This effect can be completely overcome by using a time ramp whose longest reverse time is long enough to disrupt all molecules of interest.

A phenomenon that affects all forms of pulsed-field electrophoresis is “trapping” of the largest molecules, preventing them from entering the gel. Trapping is dependent on both voltage and temperature. At high temperatures and field strengths, molecules as small as 1 Mb may be trapped; at low temperatures and field strengths, molecules >5 Mb will enter the gel.

Critical Parameters

All forms of pulsed-field electrophoresis depend on differences in the time it takes for molecules of various sizes to change directions in a gel; these times are affected by a number of factors. The effects of some of these factors have been investigated in detail (Birren et al., 1988) and are discussed below.

Voltage. Voltage is measured in V/cm of gel length. It is best to measure this directly by dipping the two probes of a voltmeter into the buffer at the ends of a running gel, and dividing the reading by the length of the gel in centimeters. Most horizontal submarine gel boxes apply ~80% of the power-supply voltage to the gel (the rest is lost in the buffer tanks), so voltage can be estimated by multiplying the

power supply readout by 0.8 and dividing the result by the length of the gel. Sizes resolved by pulsed-field gels can be changed over a wide range merely by changing the voltage. Higher voltage increases the sizes separated, but also may result in trapping the largest molecules. Pulsed-field gels are generally run at 5 to 10 V/cm.

Temperature. Pulsed-field gels can be run over a broad range of temperatures. A given time interval will separate larger molecules at 30°C than at 10°C. Choose a temperature that is convenient and adjust times for best resolution. The upper limit of resolution field-inversion gels is affected by trapping, and at 8 V/cm varies from 2000 kb at 4°C to ~1000 kb at 30°C.

Although different temperatures may be used, it is critical that the entire gel be at the same temperature to prevent the “smile” effect. This is best accomplished by recirculating the buffer using a peristaltic pump. In many cases, pulsed-field gels can be run on the benchtop with no cooling. Gels that generate a lot of heat can often be run in a cold room.

Buffer. Because pulsed-field gels are usually run at higher voltages than standard gels, they have the potential to generate quite a bit of heat. For this reason, they are often run in 0.5× TBE buffer, which carries little current. This works well, especially for molecules less than a few hundred kilobases.

DNA has higher mobility in TAE buffer, but TAE carries a lot of current, causing gels to heat

Table 2.5B.2 Empirical Equations for Fragment Resolution and Velocity Using Pulsed-Field Gels^{a,b}

Field-inversion gels

Maximum resolved size ^c (kb)	$0.13 \times (T + 40) \times V^{1.1} \times (3 - A)^{0.6} \times t^{0.875}$
Minimum resolved size (kb)	0.75 × maximum size
Velocity of 10-kb fragment (cm/hr)	$\frac{0.0016 \times (T + 25) \times V^{1.6} \times (R - 1)}{A \times (R + 1)}$

CHEF or other alternating-angle gels

Maximum well-resolved size ^c (kb)	$0.034 \times (T + 40) \times V^{1.1} \times (3 - A)^{0.6} \times t^{0.875}$
Minimum well-resolved size (kb)	0.75 × maximum size
Velocity of 10-kb fragment (cm/hr)	$\frac{0.0012 \times (T + 25) \times V^{1.6} \times \cos(\theta/2)}{A}$

^aEquations assume use of 0.5× TBE buffer; for GTBE and TAE buffers, sizes separated will be slightly larger and gels will run ~20% and 30% faster, respectively.

^bVariables: *T*, temperature in °C; *V*, field strength, volts/cm; *A*, % agarose (multiply by 0.8 for pulsed-field grade agarose); *t*, pulse time (reverse time for field-inversion gels) in sec; *R*, forward-to-reverse time ratio; *θ*, reorientation angle.

^cField inversion does not resolve fragments outside this range; alternating-angle gels will resolve fragments outside this range, but not as well.

up. A good compromise is GTBE buffer, which is 0.5× TBE with 100 mM glycine added. GTBE increases the mobility of the DNA without significantly increasing the current.

Ethidium bromide. Ethidium slows the re-orientation of molecules, possibly by making DNA stiffer. Addition of ethidium bromide to the gel can help resolve molecules <100 kb, but is not recommended for larger molecules.

Agarose. Low agarose concentration shifts the sizes of molecules resolved toward the larger range, and also speeds the migration of all molecules. Special pulsed-field grade agarose is available from several companies. These agaroses make gels with large pore sizes, so that larger molecules can be resolved, and with high gel strength, so that low-percentage gels can be used. A gel poured with 1% pulsed-field grade agarose will give similar results to a gel poured with 0.8% regular agarose, but will be much stronger.

Pulse times. For alternating-angle electrophoresis, best resolution is obtained by using pulse times just long enough to resolve the largest molecules of interest (or better yet, using a time ramp with a maximum that is just long enough).

Pulse times are much more critical for field-inversion electrophoresis, because a given reverse pulse time resolves a narrower range of sizes. Use a range of reverse times broad enough to separate all molecules of interest. The ratio between forward and reverse times is generally in the range of 2.5:1 to 3.5:1. Lower ratios give better separation but make gels run more slowly. The most commonly used ratio is 3:1.

Table 2.5B.2 lists equations predicting the sizes resolved and the speed of gel runs for field-inversion and alternating-angle gels. For example, for a field inversion gel run at 12° at 8 V/cm in 0.8% agarose with reverse pulses of 10 sec and forward pulses of 30 sec:

$$\begin{aligned} \text{Size} &= 0.13 \times (12 + 40) \times 8^{1.1} \times \\ &\quad (3 - 0.8)^{0.6} \times 10^{0.875} \\ &= 0.13 \times 52 \times 9.85 \times 1.6 \times 7.5 = 800 \text{ kb} \end{aligned}$$

$$\text{Minimum size} = 0.75 \times 800 = 600 \text{ kb}$$

$$\begin{aligned} \text{Velocity} &= \frac{0.0016 \times (12 + 25) \times 8^{1.6} \times (3 - 1)}{0.8 \times (3 + 1)} \\ &= \frac{0.0016 \times 37 \times 27.9 \times 2}{0.8 \times 4} = 1.0 \text{ cm/hr} \end{aligned}$$

Thus, the gel will resolve fragments from 600 to 800 kb, and a 10-kb fragment will move at 1 cm/hr.

Troubleshooting

Size range resolved too high or too low. Refer to Table 2.5B.2 and adjust whatever parameter is convenient (typically voltage or pulse times) to change the sizes resolved.

Bands broadening or disappearing at top of gel. Decrease trapping by decreasing temperature or voltage.

Bands smeared. Run samples into gel for longer time, at lower voltage, or both.

Excessive smile. More heat is being produced than can be dissipated effectively. Either reduce the heating by decreasing voltage, gel thickness, or buffer depth, or increase buffer recirculation.

Anticipated Results

Molecules of interest should be well resolved in straight lanes (see Fig. 2.5B.2).

Time Considerations

The time required for running a gel can be estimated from Table 2.5B.2. The support protocol on sample preparation requires two overnight incubations, and restriction enzyme digestion (if desired) takes most of another day.

Literature Cited

Birren, B.W., Lai, E., Clark, S.M., Hood, L. and Simon, M.I. 1988. Optimized conditions for pulsed field electrophoretic separations of DNA. *Nucl. Acids Res.* 16:7563-7581.

Carle, G.F., Frank, M., and Olson, M.V. 1986. Electrophoretic separations of large DNA molecules by periodic inversion of the electric field. *Science* 232:65-68.

Chu, G. 1989. Pulsed field electrophoresis in contour-clamped homogeneous electric fields for the resolution of DNA by size or topology. *Electrophoresis* 10:290-295.

Chu, G., Vollrath, D., and Davis, R.W. 1986. Separation of large DNA molecules by contour clamped homogeneous electric fields. *Science* 234:1582-1585.

Schwartz, D.C. and Cantor, C.R. 1984. Separation of yeast chromosome-sized DNAs by pulsed-field gradient electrophoresis. *Cell* 37:67-75.

Key Reference

Schwartz and Cantor, 1984. See above.

Describes the original use of pulsed fields to separate large molecules.

Contributed by Michael Finney
MJ Research
Watertown, Massachusetts

**Preparation and
Analysis of DNA**

2.5B.9

Isolation and Purification of Large DNA Restriction Fragments from Agarose Gels

UNIT 2.6

This unit describes methods for recovering and purifying DNA restriction fragments from agarose gels. The first Basic Protocol involves electroelution of the fragment of interest from standard agarose gels using buffer-filled dialysis bags, followed by concentration and purification using an Elutip column (see Basic Protocol 1). This approach can be used effectively for fragments of all sizes from 50 to 20,000 bp. Electrophoresis directly onto NA-45 paper (see Basic Protocol 2) provides relatively high yields for fragments ≤ 2000 bp. Fragments ≥ 1000 bp can also be separated on low gelling/melting agarose gels and purified by phenol extraction (see Basic Protocol 3), β -agarase digestion of the gel (Alternate Protocol 1), or via silica membrane spin columns (Alternate Protocol 2). Removing linkers from a fragment using a column rather than a gel (Alternate Protocol 3) is included, followed by a method for estimating DNA concentrations in solution (Support Protocol).

ELECTROELUTION FROM AGAROSE GELS

**BASIC
PROTOCOL**

Following digestion with appropriate restriction enzymes, the DNA of interest is electrophoresed on a preparative agarose gel. The portion of the gel containing the restriction fragment to be purified is then physically removed from the remainder of the gel. This agarose slice is placed into a buffer-filled piece of dialysis tubing and again subjected to electrophoresis. The restriction fragment migrates out of the gel slice into the buffer, and the DNA is further purified and concentrated using an Elutip-d column. This procedure is effective with fragment sizes ranging from 50 to 20,000 bp, although for eluting fragments ≤ 2000 bp, the use of one of the other protocols (e.g., elution from NP-45 paper; see Basic Protocol 2) is recommended.

Materials

- DNA encoding sequence of interest
- Appropriate restriction enzymes and buffers (*UNIT 3.1*)
- Ethidium bromide solution (*UNIT 2.5A*)
- TAE buffer (*APPENDIX 2*)
- Elutip high-salt solution (see recipe)
- 2.5 M NaCl
- Elutip low-salt solution (see recipe)
- 100% and 70% ethanol
- TE buffer, pH 8.0 (*APPENDIX 2*)
- Spectrapor 3 dialysis membrane tubing (11.5-mm diameter with MWCO of 3500; Baxter)
- Elutip-d columns (Schleicher & Schuell)
- Small syringe (e.g., 5-ml)
- Additional reagents and equipment for restriction enzyme digestion (*UNIT 3.1*) and agarose gel electrophoresis and photography (*UNIT 2.5A*)

Choose and electrophorese a preparative agarose gel

1. Digest 0.1 to 25 μ g DNA to completion with appropriate restriction enzyme(s).

The relative abundance of the DNA fragment to be isolated in the sample to be electrophoresed and the use to which this fragment will be put are the most important factors in determining the amount of DNA to be digested. For frequently used cloning

**Preparation and
Analysis of DNA**

2.6.1

vectors and fragments for nick translation, enough sample should be digested to give ~10 μg of DNA in the target band. For preparation of fragments to be used in a single cloning experiment, 1 μg of DNA in the target band is more than sufficient. It is worthwhile to electrophorese a small aliquot of the preparative digest on a minigel before loading the preparative gel—this will allow the efficiency of the digestion to be ascertained as well as verify the existence of the relevant restriction sites.

2. Load sample onto the appropriate agarose gel and electrophorese.

The molecular weight of the DNA fragment to be isolated, as well as the weights of other fragments in the sample, should be considered when choosing the dimensions and percentage of the gel to be poured and the rate at which the gel will be electrophoresed. Using a wide slot (e.g., 4 to 5 cm) on a large gel (e.g., 20 \times 20 cm) is best when large amounts of DNA are to be separated or when the DNA fragment of interest is somewhat difficult to resolve from other fragments. Using a minigel is best when a small amount of a well-separated DNA fragment is to be isolated. It is best to load only a single sample on a preparative gel and to minimize or eliminate markers—this will avoid contamination of the fragment of interest.

3. After electrophoresis, stain the gel with ethidium bromide solution and photograph the gel.

Do not be surprised to see a few extra bands on the preparative gel. Because so much more DNA is loaded onto the gel, minor products of the restriction digest are sometimes visible. If the gel was electrophoresed too rapidly or if too much DNA was loaded into the preparative slot, a “flame” of DNA running from the bands back up to the slot will be visible. When this flame is present, it is likely that trapping has occurred (i.e., that small amounts of DNA of several different molecular weights migrated with the target band).

Electroelute the DNA

4. Using UV light for visualizing bands, carefully cut out the target band with a scalpel.

Cut the narrowest gel slice possible that allows the removal of the bulk of the target band. Occasionally, it is worthwhile to cut out only the heart of the target band if it has migrated close to another band on the gel. Be sure to minimize exposure of eyes, skin, and the DNA to UV light.

CAUTION: *Wear safety goggles or full face mask—UV burns of the cornea (which will occur with even brief exposure of unprotected eyes) are extremely painful and can result in infections and loss of vision. A strong UV source can also “sunburn” sensitive skin in exposed areas such as the face and hands in <1 min.*

5. Rinse dialysis tubing with TAE buffer.

It is not necessary to store the tubing in buffer for long periods of time—5 min in TAE buffer is sufficient.

6. Tightly tie off one end of the tubing with two knots. Slide the gel slice into the tubing. Fill the tubing with TAE buffer until it is almost completely expanded. Knot or clip the top of the tubing closed.

It is essential that no leakage from the sealed tubing occur. Spectra/Por closures (e.g., Thomas Scientific) are often used to clip the ends of the bags.

7. Place the sealed dialysis bag in a horizontal gel electrophoresis apparatus (any box with two wires connected to electrodes). Fill the apparatus with 1 \times TAE buffer until the dialysis bag is just covered with solution.

If too much buffer is added, the dialysis bag will float freely and not always be in the proper orientation for efficient electroelution.

8. Electroelute at a constant voltage of ~2 V/cm between the two wires. For a 50- to 500-bp target fragment, electroelute 30 to 45 min. For a 500- to 2000-bp fragment, electroelute 2 hr. For a 2000- to 4000-bp fragment, electroelute 4 hr. For larger fragments, electroelute at 1 V/cm overnight.

Under no circumstances should the running buffer heat up—the yield will be severely reduced if this occurs.

9. After electroelution is complete, reverse the polarity of the electrodes and turn on voltage at 100 V for 30 sec.

This will free any DNA that has stuck to the dialysis tubing.

10. Carefully open the top of the dialysis bag and collect TAE buffer with a Pasteur pipet. Massage gel slice out of the bag and wash the bag with a pipet of TAE buffer.

It is worthwhile to stain the gel slice with ethidium bromide to ensure that all the DNA has been eluted from the gel. If some DNA remains in the gel, repeat electroelution (step 4). Depending on size of gel slice, the final volume of TAE buffer will range from 2 and 25 ml.

Concentrate and purify electroeluted DNA with Elutip-d

11. Wet the Elutip-d column with Elutip high-salt solution. Place 2 ml high-salt solution in a small syringe. Attach the syringe to an Elutip-d column and push the high-salt solution through the column.

Elutip-d miniature affinity columns contain a reverse-phase resin. Do not cut off the tip of the column (as recommended in some protocols), as this will cause solutions to pass through the column too rapidly. Also, do not allow any solutions to flow in the reverse direction (i.e., toward the wide portion of the column), and do not allow the column to dry out during the procedure.

12. Equilibrate the column with Elutip low-salt solution. Place 5 ml low-salt solution in the syringe and push the solution through the column.

13. Adjust salinity of the DNA sample to be purified to 0.2 M NaCl with 2.5 M NaCl. Take up the DNA solution in the syringe and load onto the column.

The DNA is loaded onto the column in low salt. Loading is conveniently done by placing a weight on the syringe such that ~20 drops per minute of the DNA-containing solution pass through the column. If the flow rate is too high, the DNA will flow through too quickly and severe reduction in yield will occur. If the DNA solution being applied to the column appears to contain particulates (which it should not), a disposable 0.45- μ m cellulose acetate filter can be placed between the syringe and the column.

14. Wash the column with 5 ml low-salt solution.

The flow rate through the column during this step should be approximately twice as fast as during the DNA loading step.

15. Elute the DNA with 400 μ l high-salt solution. Collect the solution in a 1.5-ml microcentrifuge tube.

The flow rate should be the same as during the DNA loading step. It is worthwhile to elute a second time with 400 μ l high-salt solution.

16. Add 1 ml of 100% ethanol to 400 μ l eluted DNA. Precipitate at -20°C . Pellet DNA, wash in 70% ethanol, and dry. Resuspend in desired amount of TE buffer, pH 8.0, and use for further manipulations and quantitation (support protocol).

For fragments >1 kb, the yield after electroelution and purification should be 80% to 90%. For smaller fragments, the yield is generally 50% to 60%.

ELECTROPHORESIS ONTO NA-45 PAPER

This procedure is simpler than Basic Protocol 1, as it does not require the manipulation of gel slices and dialysis bags. It is particularly suitable for small DNA fragments (i.e., <2000 bp). After preparative gel electrophoresis in a gel containing ethidium bromide, a piece of NA-45 paper is inserted into the gel ahead of the fragment of interest. Further electrophoresis results in migration of the fragment onto the paper, which then is removed from the gel. The fragment is eluted from the paper with high salt.

Materials

Ultrapure agarose (e.g., SeaKem GTG agarose, FMC Bioproducts)

NA-45 paper (Schleicher & Schuell)

TE buffer, pH 8.0 (*APPENDIX 2*)

NA-45 elution buffer (see recipe)

Buffered phenol (*UNIT 2.1*)

25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol (*UNIT 2.1*)

95% and 70% ethanol, ice cold

Flat forceps, two pairs

Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5A*) and DNA extraction and precipitation (*UNIT 2.1*)

1. Digest DNA to completion with appropriate restriction enzymes and load onto gel prepared with ultrapure agarose (see Basic Protocol 1, steps 1 and 2).

Ethidium bromide is added to the gel at a final concentration of 0.5 µg/ml.

2. Electrophorese at appropriate voltage until DNA fragment of interest is well resolved from contaminants, as assessed by visualization with hand-held UV lamp.
3. Stop electrophoresis and, with a clean scalpel or razor blade, cut slits just above and below the fragment.
4. Insert a small piece of NA-45 paper into each of the slits by carefully separating the opening with a flat forceps while placing the paper in the slit with another set of forceps.

The NA-45 paper is cut to the size of a rectangle as wide as the gel band of interest and as long as the depth of the gel. Do not touch the paper with your hands!

5. Gently push on the top and bottom of the gel to ensure that the slits are in contact with the paper. Electrophorese 10 min at the same voltage and current as the prior run, or until the DNA has migrated onto the paper.

The paper below the band will bind the DNA fragment of interest, while the paper above will capture any higher-molecular-weight contaminants.

6. Carefully remove the paper containing the DNA fragment of interest with forceps. Wash paper gently three times in TE buffer, pH 8.0.

A 12-well microtiter plate is convenient for this; simply move the paper from one well to the next.

7. Remove the washed paper to a 1.5-ml microcentrifuge tube containing 400 µl NA-45 elution buffer. Heat to 70°C—15 min for short fragments (<500 bp) or 1 hr for larger fragments (>1500 bp).

Intermediate times are appropriate for intermediate-length fragments.

8. Extract once with 400 µl buffered phenol. Reextract aqueous phase twice with phenol/chloroform/isoamyl alcohol, then extract twice with chloroform.

9. Add 1 ml of 95% ethanol and precipitate overnight at -20°C .
10. Microcentrifuge at high speed and wash pellet with 70% ethanol chilled to -20°C . Resuspend DNA in TE buffer, pH 8.0.

ISOLATION OF DNA FRAGMENTS USING LOW GELLING/MELTING TEMPERATURE AGAROSE GELS

BASIC PROTOCOL 3

Following preparative gel electrophoresis using low gelling/melting temperature agarose, the gel slice containing the band of interest is removed from the gel. This agarose slice is then melted and subjected to phenol extraction or to an Elutip-d column.

For some applications (including restriction enzyme digestions, *UNIT 3.1*, and ligations, *UNIT 3.16*) the melted gel slice containing the fragment can be used directly, without further purification by extraction or Elutip-d purification.

Materials

DNA to be isolated

Appropriate restriction enzymes and buffers (*UNIT 3.1*)

Low gelling/melting temperature agarose (SeaPlaque; FMC Bioproducts)

Ethidium bromide solution (*UNIT 2.5A*)

TE buffer, pH 8.0 (*APPENDIX 2*)

Buffered phenol (*UNIT 2.1*)

Elutip high-salt and low-salt solutions (see recipe)

Scalpel

Elutip-d column (Schleicher & Schuell)

Beckman JS-13 swinging-bucket rotor (or equivalent)

Additional reagents and equipment for restriction enzyme digestion (*UNIT 3.1*), agarose gel electrophoresis (*UNIT 2.5A*), and ethanol precipitation (*UNIT 2.1*)

1. Digest DNA sample to completion with appropriate restriction enzymes. Pour, load, and electrophorese on a 1% low gelling/melting temperature agarose gel.

Both the quantity of DNA that can be loaded and the quality of resolution are less on a low gelling/melting temperature gel than on a standard agarose gel.

2. Stain the gel using ethidium bromide solution and cut out the target band with a scalpel. Melt the gel slice at 65°C and add enough TE buffer, pH 8.0, to decrease the agarose percentage to $\leq 0.4\%$.

3. If DNA is to be used for ligation (*UNIT 3.16*), transformation (*UNIT 1.12*), or restriction endonuclease digestion (*UNIT 3.1*), use the molten solution directly. To isolate a fragment directly using an Elutip-d column, proceed to steps 8 and 9. To remove agarose by phenol extraction, proceed to step 4.

Fragments separated by low gelling/melting temperature agarose gels can also be isolated by β -agarase digestion of the agarose (see Alternate Protocol 1) or using a silica membrane spin column (see Alternate Protocol 2).

4. Add an equal volume of buffered phenol (*not* phenol/chloroform) and mix vigorously for 5 to 10 min. Centrifuge 10 min at $15,800 \times g$ (10,000 rpm in a Beckman JS-13 rotor), room temperature.

After centrifugation, a thick white interface will be visible.

5. Collect the aqueous phase and set aside. Reextract the phenol phase and interface with an equal volume of TE buffer, pH 8.0. Centrifuge as in step 4.

Preparation and Analysis of DNA

2.6.5

6. Collect second aqueous phase. If a large interface still appears, reextract a third time.
7. Ethanol precipitate the combined aqueous phase. Further purify the DNA solution via an Elutip-d column (steps 8 and 9) or use directly after resuspending in the appropriate buffer.

Generally, the resulting pellet will be rather large, as it contains a fair amount of agarose. If the DNA is not further purified, it is helpful to add BSA to 500 µg/ml final concentration. This helps reduce the inhibition of enzymatic reactions that is characteristic of agarose-containing DNA solutions.
8. Add 10 to 20 vol Elutip low-salt solution.
9. Purify the DNA fragment using an Elutip-d column (see Basic Protocol, steps 11 to 16), maintaining all solutions at 37°C. Ethanol precipitate and quantitate as described in the support protocol.

**ALTERNATE
PROTOCOL 1**

**RECOVERY OF DNA FROM LOW GELLING/MELTING TEMPERATURE
AGAROSE GELS USING β-AGARASE DIGESTION**

The following protocol is an alternative method for recovery of DNA from low gelling/melting temperature agarose. A gel slice containing the DNA of interest is treated with β-agarase, which digests the long-chain polysaccharides in the agarose to monosaccharides and short oligosaccharides. The resulting DNA in solution can be directly subjected to many further manipulations such as ligation (UNIT 3.16), transformation (UNIT 1.12) and restriction endonuclease digestion (UNIT 3.1). Alternatively, ethanol precipitation may be used to remove oligosaccharide digestion products. This protocol may also be used in conjunction with purification via Elutip-d columns (second basic protocol) and for purifying small DNA fragments from sieving agarose gels (UNIT 2.8).

Additional Materials (also see Basic Protocol 3)

β-agarase I (New England Biolabs or Calbiochem)
β-agarase buffer (see recipe)

Additional reagents and materials for dialysis (APPENDIX 3D) and isopropanol precipitation of DNA (UNIT 2.1)

1. Prepare a low gelling/melting temperature agarose gel (see Basic Protocol 3, steps 1 and 2).
2. Transfer gel slice to a clean tube. Wash twice for 30 min each on ice with 2 vol of 1× β-agarase buffer. Add an approximately equal volume of β-agarase buffer to the washed gel slice.

Alternatively, if trying to keep volumes low, add 1/10 vol β-agarase buffer to the gel slice and proceed directly to step 3.
3. Melt the gel completely by heating 10 min at 65°C.
4. Equilibrate molten agarose to 40°C (allow ~10 min). Add 1 U β-agarase for every 200 µl of 1% agarose and continue incubation for 1 hr.

If equilibration of the gel slice in β-agarase buffer was bypassed in step 2, double the amount of β-agarase added here. β-agarase exhibits 50% activity in TBE buffer and 80% activity in TAE buffer.

5. If DNA is to be used for ligation (*UNIT 3.16*), transformation (*UNIT 1.12*), or restriction endonuclease digestion (*UNIT 3.1*), use the molten β -agarase/DNA solution directly. To further purify large (>50 kb) DNA fragments, dialyze the DNA solution to remove carbohydrates and β -agarase. To purify smaller fragments, follow steps 6 to 9.

This gentle treatment will avoid mechanical shearing of large DNA molecules.

6. To purify small (<50 kb) fragments, adjust the salt concentration of the β -agarase/DNA solution to 0.5 M NaCl and add an equal volume of isopropanol. Chill 15 min on ice.

Sodium acetate (0.3 M final) or LiCl (0.8 M final) may also be used for this precipitation step.

7. Centrifuge 15 min at $15,000 \times g$ to pellet any undigested carbohydrates and transfer supernatant to a clean tube.
8. Add 2 to 3 vol isopropanol, mix thoroughly, and chill 30 min at 0°C.

Carrier RNA (1 μ g) may be added to ensure recovery of small quantities of DNA. Large volumes of isopropanol are needed to dilute salt concentrations and ensure DNA precipitation.

9. Centrifuge 15 min at $15,000 \times g$, discard supernatant, and dry pellet. Resuspend in TE buffer, pH 8.0, or appropriate buffer required for subsequent manipulation.

RECOVERY OF DNA FROM AGAROSE USING SILICA MEMBRANE SPIN COLUMNS

ALTERNATE PROTOCOL 2

The following rapid approach is based on the silica membrane column purification method described in *UNIT 2.1*. A gel slice containing DNA fractionated through an agarose gel is melted and passed through a silica membrane column in the presence of high salt. Under these conditions, DNA is adsorbed onto the silica membrane. The gel contaminants are subsequently washed away and DNA is eluted in low-salt buffer. This procedure is suitable for fragments >100 bp. It is actually intended for recovering DNA from gels made from standard agarose, but can also be used with low gelling/melting temperature agarose. Although it is more rapid than standard electroelution or organic extraction methods, this approach may result in somewhat lower yields. Silica membrane spin columns for purification from agarose gels are available from many companies (including Qiagen, Promega, Invitrogen, and Novagen) as kits which include silica membrane spin columns and all appropriate buffers necessary for DNA purification.

Additional Materials (also see Basic Protocol 3)

6.0 M NaI solution (filter through filter paper, store up to 3 months in the dark at 4°C)

Binding buffer (see recipe)

Wash buffer (*UNIT 2.1A*)

TE buffer, pH 8.0 (*APPENDIX 2*) or nuclease-free H₂O

1.5-ml microcentrifuge tubes

45° to 50°C water bath

Silica membrane spin columns (e.g., Qiagen, Promega, Invitrogen, Novagen)

1. Digest DNA sample to completion, pour, load, and electrophorese through a 1% agarose gel, and stain gel using ethidium bromide (see Basic Protocol 3, steps 1 and 2). Cut out the target band with a clean scalpel and transfer to a 1.5-ml microcentrifuge tube.

Preparation and Analysis of DNA

2.6.7

2. Estimate the volume of the agarose (usually ~100 μ l) and add 2.5 to 3.0 vol of 6.0 M sodium iodide (250 to 300 μ l).

When using either this procedure or kits with NaI as the chaotropic salt, recovery may be reduced if the agarose gel was run with TBE rather than TAE buffer; TBE may inhibit the binding of DNA to silica membranes. To resolve this problem, reduce the pH of the buffer by the addition of 1/10 vol of sodium phosphate, pH 6 (APPENDIX 2). Sodium perchlorate or guanidine isothiocyanate can be used as the chaotropic salt to dissolve the agarose gel. It is important keep the pH of the gel/salt solution at pH 6.5 for the DNA to bind the silica membrane efficiently.

3. Incubate 5 min at 45° to 50°C to dissolve the agarose.

Mix the contents once or twice during this incubation. If all agarose has not completely dissolved after 5 min, return the tube to the water bath for 1 to 2 min, but do not leave the mixture heating longer than required to dissolve all agarose.

4. Place the tube at room temperature, add 2 vol binding buffer (700 to 800 μ l), and mix well.
5. Apply the supernatant to the silica membrane spin column.
6. Microcentrifuge the spin column inside its collection tube for 1 min at maximum speed. Remove the spin column from the collection tube and discard the flowthrough. Reinsert the spin column into the collection tube.
7. Wash the spin column by adding 750 μ l of wash buffer and microcentrifuging 1 min at maximum speed. Remove the spin column from the collection tube and discard the flow-through.
8. Reinsert the spin column in the collection tube and microcentrifuge for an additional 1 min to remove any residual wash buffer (ethanol) from the column membrane.
9. Transfer the spin column to a 1.5-ml microcentrifuge tube and add 75 to 100 μ l nuclease-free water or TE buffer, pH 8.5, to the center of the membrane. Let stand 2 to 10 min, then microcentrifuge 1 min at maximum speed.
10. Collect DNA and store at 4°C until use.

Typical yields are 60% to 80%, but may be lower for smaller fragments (e.g., ~50% for a fragment <500 bp).

ALTERNATE PROTOCOL 3

REMOVAL OF OLIGONUCLEOTIDE FRAGMENTS USING A SEPHACRYL S-300 COLUMN

In this protocol, DNA samples are passed over a Sephacryl S-300 column and the eluted fractions that contain the DNA of interest are recovered. This provides a rapid, nonelectrophoretic method for isolating a DNA fragment from linkers or other synthetic oligonucleotides.

Additional Materials (also see Basic Protocol 3)

- 50% (v/v) Sephacryl S-300 (Pharmacia Biotech) in TE buffer, pH 8.0 (Sephacryl/TE), stored covered at 4°C
- 0.3 M NaCl in TE buffer, pH 8.0 (NaCl/TE solution)
- DNA sample in NaCl/TE solution
- Glass wool plug and 6-in. Pasteur pipet, silanized (APPENDIX 3B)

1. Pour ~10 ml of a resuspended 50% Sephacryl S-300/TE buffer mixture into a tube. Allow the Sephacryl to settle.

2. Remove the TE buffer and add ~10 ml NaCl/TE solution. Mix, allow the Sephacryl to settle, and repeat.
3. Place a silanized glass wool plug inside a 6-in. silanized Pasteur pipet. Pour the equilibrated Sephacryl into the column. Continue adding mixture until the level of Sephacryl is up to the indentation on the pipet (~1 in. from the top). After the column is packed, wash with 2 ml NaCl/TE solution.
4. Load 200 μ l DNA sample onto the column and collect the first 200 μ l that elutes—this is fraction 1.
5. Add 100- μ l aliquots of NaCl/TE solution to the column and collect twelve to fourteen 100- μ l fractions.
6. Combine fractions 8 to 11. Add 2.5 vol ethanol and precipitate. Centrifuge and recover the pellet, which contains the oligonucleotide-free DNA fragment of interest.

It is worthwhile to calibrate the batch of Sephacryl using a plasmid on hand. Collect and ethanol precipitate each 100- μ l fraction individually so that the plasmid-containing fractions can be identified by electrophoresis.

RAPID ESTIMATION OF DNA CONCENTRATION BY ETHIDIUM BROMIDE DOT QUANTITATION

SUPPORT PROTOCOL

In this protocol, the sample DNA is mixed with ethidium bromide and spotted on a piece of plastic wrap next to DNA standards. Comparison of the fluorescence of the sample to these standards provides an estimate of the DNA concentration in the sample. This is a simple method to estimate the concentration of DNA in dilute solution, and is useful for determining recovery of isolated fragments.

Materials

DNA: purified DNA for standards and isolated DNA to be quantitated
(previous protocols)

TE buffer, pH 8.0 (*APPENDIX 2*)

1 μ g/ml ethidium bromide (store wrapped in foil at 4°C)

Plastic wrap

UV transilluminator

Additional reagents and equipment for gel photography (*UNIT 2.5A*)

CAUTION: Ethidium bromide is a mutagen and should be handled with care.

1. Prepare the following DNA standards in TE buffer, pH 8.0, using any solution of pure DNA: 0, 1, 2.5, 5, 7.5, 10, and 20 μ g/ml.

These standards can be stored for several months at 4°C.

2. Add 4 μ l of each standard DNA solution to 4 μ l of 1 μ g/ml ethidium bromide. Prepare 4 μ l of the DNA solution of unknown concentration in the same way. Mix well.
3. Place a piece of plastic wrap on a UV transilluminator. Spot the standard and unknown DNA/ethidium bromide solutions side-by-side on the plastic wrap. Photograph under UV light. Estimate DNA concentration of the unknown sample by comparison to the fluorescence of the standards.

The linear range of the assay is from 0 to 20 μ g/ml. If the fluorescence of the DNA sample exceeds that of the standards, dilute the unknown DNA solution until it can be estimated.

Preparation and Analysis of DNA

2.6.9

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

***β*-agarase buffer, 10×**

100 mM bis-Tris, pH 6.5

10 mM EDTA

Filter sterilize and store indefinitely at room temperature

Binding buffer (6 M guanidine HCl)

Add 0.75 g Na₂SO₃ and 57.3 g guanidine HCl (Sigma) to 35 ml H₂O and stir until dissolved. Adjust volume to 100 ml with water, sterilize, and store 3 to 4 months in the dark (in aluminum foil) at 4°C. Discard if precipitate is observed.

Elutip high-salt solution

1 M NaCl

20 mM Tris-Cl, pH 7.5 (*APPENDIX 2*)

1 mM EDTA

Filter sterilize and store indefinitely at room temperature

Elutip low-salt solution

0.2 M NaCl

20 mM Tris-Cl, pH 7.5 (*APPENDIX 2*)

1 mM EDTA

Filter sterilize and store indefinitely at room temperature

NA-45 elution buffer

1 M NaCl

0.05 M arginine (free base)

Filter sterilize and store indefinitely at room temperature

COMMENTARY

Background Information

There are probably as many approaches to the isolation and purification of DNA fragments as there are molecular biologists. Not surprisingly, many of these approaches actually work. In the 1970s, the first successful protocols for isolating DNA fragments from agarose gels were developed—these included the freeze-squeeze method (Thuring et al., 1975), gel dissolution using potassium iodide (Blin et al., 1975), and electroelution (Wienand et al., 1978). One of the most intriguing purification techniques involved specific binding of DNA fragments to powdered glass. Because only certain types of glass worked for this procedure and because the reasons why they worked were uncharacterized, it was not uncommon to see scientists crushing every piece of glassware they could get their hands on. The glass bead dilemma was resolved by Vogelstein and Gillespie (1979), and has been improved upon by the introduction of silica membrane spin columns to replace glass beads.

The most important considerations in choosing an isolation/purification system (whether it is based on physical or chemical properties of DNA) are speed of the technique, yield of purified DNA, and ability of the DNA to be further manipulated. Many protocols, including those presented here, satisfy all three criteria. The basic electroelution/Elutip protocol described here (Weinand et al., 1978) is particularly reproducible and reliable, and the resulting DNA is suitable for a wide range of enzymatic reactions as well as several biological applications (including the production of transgenic mice by single-cell embryo microinjection).

The major objection to the electroelution procedure is that it requires manipulating a small gel fragment with forceps, and introducing this into a dialysis bag. Occasionally the geometry of the gel slice and the orientation of the electric field work against each other and it is difficult to know when all of the fragment has eluted; this uncertainty can result in poor quan-

titative recovery. Two of the other methods described—electrophoresis onto NA-45 paper and purification with low gelling/melting agarose—eliminate the difficulties of getting the gel slice into the dialysis bag and often give better yields.

Electrophoresis onto NA-45 paper, using paper at both the leading and trailing edges of the band, is most effective for eliminating contamination with higher-molecular-weight fragments. It does, however, give lower yields for fragments >2 kb. In contrast, low gelling/melting agarose is effective for DNA fragments of virtually all sizes. A major advantage of using low gelling/melting agarose is the ability to perform subsequent ligations (*UNIT 3.16*) in the gel slice without removing the DNA fragment. If desired, DNA fragments in low gelling/melting temperature agarose gels may be treated with β -agarase I, an agarose-digesting enzyme isolated from *Pseudomonas atlantica*. β -agarase I cleaves carbohydrate bonds, reducing agarose to a mixture of monosaccharides and short oligosaccharides which can no longer gel. Many further enzymatic reactions can be subsequently carried out directly without interference from contaminating carbohydrate residues.

A related compound, the agarose-digesting enzyme mixture GELase (Epicentre Technologies), has somewhat different properties and can be used with denaturing gels or regular agarose as well as with low gelling/melting agarose; see manufacturer's specifications for using these gels.

An alternative for isolating relatively large DNA fragments (>500 bp) is the silica membrane spin column method (see Alternate Protocol 2). This is a very rapid method that has the advantage of concentrating purified DNA from agarose gels without the need for ethanol precipitation. Commercial kits for this procedure are now widely available (e.g., Qiagen, Promega, Invitrogen, Novagen).

Critical Parameters and Troubleshooting

The quality of the preparative DNA digest and the resolution of the preparative gel are critical for obtaining optimal yields of fragments from all protocols. If the preparative gel is overloaded, cross-contamination will occur as a small amount of each fragment becomes "trapped" in the other bands. This low level of contamination can be significant if the fragment is to be used for some cloning procedures, or as a probe in some hybridizations. If it is

necessary to obtain a large amount of highly pure fragment, the first preparation should be re-electrophoresed and repurified.

For the Elutip purification used in the electroelution or low gelling/melting temperature agarose protocols, it is essential that the DNA solution be passed through the column slowly to adsorb it completely. For the phenol extraction purification from low gelling/melting temperature agarose, it is essential to use straight buffered phenol and not phenol/chloroform/isomyl alcohol mixtures.

Transfer to NA-45 paper works well as long as the paper is carefully introduced in front of the band of interest, so that air bubbles are not trapped between the paper and the gel. It is convenient to examine the gel under UV illumination with a hand-held lamp prior to removing the NA-45 paper, to ensure that the fragment has migrated into the paper. If the fragment of interest is <2000 bp, another method should be employed, because the yields can be poor (because of inefficient elution) with small fragments.

Low gelling/melting temperature agarose does not have the integrity of regular agarose. Extreme care should be taken when pulling the combs out of the gel, as wells tear easily. Because low gelling/melting temperature agarose remelts at 65°C, a low voltage (6 to 7 V/cm of gel) should be maintained to keep the gel from overheating. In 4% gels using a TAE buffer system, the bromphenol blue marker migrates at ~50 bp.

The β -agarase treatment is straightforward and reliable. Incomplete digestion can be remedied in most cases by adding more enzyme (up to double the amount recommended). If difficulties are encountered when using the Sephacryl column to purify DNA fragments away from small oligonucleotides, the column must be calibrated using plasmid DNA as described in the protocol.

The most critical parameters in purification of DNA using silica membrane spin columns are the molar concentration and pH of the chaotropic salt solution. For efficient binding to the silica membrane, the salt solution should be at pH 6.5 with the chaotropic salt (NaI, NaClO₄, guanidine HCl, guanidine isothiocyanate) at 2 M. Some companies recommend that the agarose gel be made with TAE and not TBE. This problem may be overcome by reducing the pH of the gel/salt solution with the addition of 1/10 volume of 1.0 M sodium phosphate, pH 6.5.

Anticipated Results

Electroelution, although more tedious and time-consuming to perform than the other methods, provides the most efficient recoveries of highly purified DNA: 80% to 90% for fragments ≥ 1000 bp and 50% to 60% for smaller fragments. Electrophoresis onto NA-45 paper is appropriate for fragments ≤ 2000 bp, giving yields of 50% to 90%. Recovery from conventional or sieving low gelling/melting temperature agaroses is generally 70% for fragments ≤ 1000 bp.

Typical yields for silica membrane spin column purification are 60% to 80% for large DNA fragments. The procedure seems to work best with DNA fragments > 500 bp; smaller-length fragments apparently bind tightly and irreversibly to the silica membrane.

Time Considerations

The gel slice containing the target band should be cut out as soon as possible after the preparative gel has been electrophoresed—this will minimize diffusion of the DNA in the band. The gel slice can be stored for hours to days by wrapping the damp gel slice in a piece of plastic wrap and refrigerating. Depending on the size of the fragment to be isolated, electroelution should take 2 to 4 hr. After electroelution, the DNA solution can be stored for days at 4°C before purifying. Once Elutip purification has been started, however, it should be carried to the stage of ethanol precipitation. The Elutip-d column itself is completed in ~ 15 min.

NA-45 elution takes ~ 1 hr for fragments < 500 bp and longer (≤ 2 to 3 hr) for those > 1500

bp. Purification of fragments separated on low gelling/melting temperature agarose, using either phenol extraction or β -agarase digestion, will take ~ 1 hr. Purification via silica membrane spin column requires 20 to 40 min. Linker removal using Sephacryl S-300 takes ~ 30 min.

Literature Cited

- Blin, N., Gabain, A.V., and Bujard, H. 1975. Isolation of large molecular DNA from agarose gels for further digestion by restriction enzymes. *FEBS Lett.* 53:84-86.
- Thuring, R.W., Sanders, J.B., and Borst, P.A. 1975. Freeze-squeeze method for recovering long DNA from agarose gels. *Anal. Biochem.* 66:213-220.
- Vogelstein, B. and Gillespie, D. 1979. Preparative and analytical purification of DNA from agarose. *Proc. Nat. Acad. Sci. U.S.A.* 76:615-619.
- Wienand, U., Schwarz, Z., and Felix, G. 1978. Electrophoretic elution of nucleic acids from gels adapted for subsequent biological tests: Application for analysis of mRNAs from maize endosperm. *FEBS Lett.* 98:319-323.

Contributed by David Moore and
Dennis Dowhan
Baylor College of Medicine
Houston, Texas

Joanne Chory
The Salk Institute
La Jolla, California

Randall K. Ribaldo (NA-45 paper)
National Institute of Allergy and
Infectious Diseases
Bethesda, Maryland

RESOLUTION AND RECOVERY OF SMALL DNA FRAGMENTS

This section describes the use of polyacrylamide gels and sieving agarose gels for analytical or preparative separation of small double-stranded DNA fragments. Polyacrylamide gels provide somewhat better resolution as well as significantly higher capacity. Sieving agarose gels are much easier to pour and run, however, and are particularly useful for simple analytical applications.

The use of denaturing polyacrylamide gels to separate single-stranded polynucleotides is described in *UNITS 2.12 & 7.6*.

Separation of Small DNA Fragments by Conventional Gel Electrophoresis

Large amounts of small (<1000-bp) DNA fragments can be separated by conventional electrophoretic means. The purified fragments can then be used for cloning, sequencing, and labeling. In this unit, the techniques of DNA separation via both nondenaturing polyacrylamide and sieving agarose electrophoresis are discussed.

Basic Protocol 1 outlines the pouring and electrophoresis of nondenaturing polyacrylamide gels. Elution of the labeled or unlabeled separated DNA fragments from the gels by either passive diffusion (see Basic Protocol 1) or electroelution (see Alternate Protocol) is described.

Sieving agarose (see Basic Protocol 2) is a specially treated type of agarose designed to be used at high concentrations. Poured and run like conventional agarose gels, this matrix can resolve small DNA fragments much like a nondenaturing polyacrylamide gel.

NONDENATURING POLYACRYLAMIDE GEL ELECTROPHORESIS

This protocol describes the preparation of polyacrylamide gels for separation of small double-stranded DNA fragments. After gel setup, DNA samples are loaded, electrophoresed through the gel, and finally purified away from the gel slices.

Materials

- 10× and 1× TBE electrophoresis buffer, pH 8.0 (*APPENDIX 2*)
- 29:1 (w/w) acrylamide/bisacrylamide (see recipe; solutions are also available commercially from National Diagnostics)
- TEMED (*N,N,N',N'*-tetramethylethylenediamine; store at 4°C)
- 10% (w/v) ammonium persulfate (APS) in water (store ≤1 month at 4°C)
- 5× loading buffer (see recipe)
- DNA samples
- DNA-molecular-weight markers: e.g., pBR322 restriction digested with *Hin*FI or M13 digested with *Hpa*II
- 0.5 µg/ml ethidium bromide
- Elution buffer, pH 7.5 (see recipe)
- 100% and 70% ethanol
- TE buffer, pH 7.5 (*APPENDIX 2*)
- 3 M sodium acetate, pH 5.2 (*APPENDIX 2*)

Glass plates, spacers, and combs for casting gels
 Acrylamide gel electrophoresis apparatus
 DC power supply
 Thin-layer chromatography (TLC) plate with fluorescent indicator (e.g., Silica Gel F-254 or IB-F; for UV shadowing)
 Longwave UV transilluminator
 3-ml small-bore disposable syringe
 Syringe equipped with silanized glass wool plug (UNIT 5.6) or 2- μ m filter
 Centrifuge with Beckman JA-20 rotor or equivalent
 Additional reagents and equipment for ethanol precipitation (UNIT 2.1A)

Prepare the gel

1. Assemble the gel casting apparatus.

Gel spacer and casting systems have been developed to avoid leakage. Those that avoid sealing the gel with tape are best, and recently gel casting boots that lack bottom spacers have become available (Life Technologies). Greasing the side and/or bottom spacers or pouring an agarose plug for the gel is not necessary if some care is taken to ensure that the bottom of the plate assembly is completely sealed. The gel plates should be thoroughly cleaned by washing them with warm soapy water and then rinsing with 70% ethanol. However, if the plates are particularly dirty or if complete removal of any residual nucleic acids is required, the plates may be soaked in 0.1 M NaOH for 30 min prior to washing. When the gel is particularly thin (<1 mm), silanizing one or both plates (APPENDIX 3B) facilitates post-electrophoretic separation of the gel from the plate.

2. Prepare the gel solution in a flask that has a wide mouth and a spout for pouring (see Table 2.7.1 for appropriate acrylamide concentrations for resolving DNA fragments of different sizes). For a nondenaturing 5% polyacrylamide gel of 20 cm \times 16 cm \times 1.6 mm, 60 ml of gel solution is sufficient; this can be made by mixing the following:

6 ml 10 \times TBE buffer
 10 ml 29:1 acrylamide/bisacrylamide
 44 ml H₂O.

CAUTION: Acrylamide is a neurotoxin; always wear gloves, safety glasses, and a surgical mask when working with acrylamide powder.

The migration distance (D) of double-stranded DNA through a nondenaturing gel is inversely proportional to the log of its molecular weight: i.e., $D \approx -\log(MW)$. Pick a concentration of acrylamide that will allow the desired DNA fragments to have migrated approximately one-half to three-fourths of the way through the gel when the loading dye has reached the bottom of the gel. Also, note that the base composition of a sequence affects its electrophoretic mobility and may cause aberrant migration.

Commercially prepared polyacrylamide solutions (National Diagnostics) are available and highly recommended, since they have long shelf lives and avoid manipulation of the neurotoxic acrylamide powder.

Table 2.7.1 Concentrations of Acrylamide Giving Maximum Resolution of DNA Fragments^a

Acrylamide (%)	Size of fragments separated (bp)	Migration of bromphenol blue marker (bp)	Migration of xylene cyanol marker (bp)
3.5	100 to 1000	100	460
5.0	100 to 500	65	260
8.0	60 to 400	45	160
12.0	50 to 200	20	70
20.0	5 to 100	12	45

^aData are compiled from articles by Maniatis and Ptashne (1973a,b) and Maniatis et al. (1975).

3. Vigorously agitate the solution ~1 min with magnetic stirring to ensure complete mixing.
4. Add 34 μl TEMED and swirl the flask to ensure thorough mixing. Immediately add 250 μl of 10% (w/v) APS and mix thoroughly. Pour the acrylamide between the gel plates and insert the comb. Clamp the comb in place at the top of the gel to avoid separation of the gel from the plates as the acrylamide polymerizes. Allow the gel to polymerize for ~30 min.

IMPORTANT NOTE: Polymerization begins as soon as APS is added to the mixture, so all succeeding actions must be performed promptly.

CAUTION: Be sure to wear safety glasses while pouring the gel since splashing of the neurotoxic, unpolymerized acrylamide is common.

For thick gels, the acrylamide should be poured directly from the mixing flask, but for thinner ones, a syringe fitted with a needle is useful. By pouring the gel slowly with a tilt 45° relative to the bench top and starting from one corner, bubbles may be largely avoided. Also, the gel should be polymerized lying flat to avoid undesirable hydrostatic pressure on the gel bottom.

TEMED may be stored indefinitely at 4°C, but the ability of APS solution to efficiently initiate the free radical-induced acrylamide polymerization diminishes greatly over time. A new APS stock should be prepared each month and stored at 4°C.

Run the gel

5. After polymerization is complete, remove the comb and any bottom spacers from the gel. Wash the gel plates free of spilled acrylamide and be sure that the spacers are properly seated and clean.
6. Fill the lower reservoir of the electrophoresis tank with 1× TBE buffer. Initially, place the gel into the lower tank at an angle to avoid the formation of air bubbles between the plates and the gel bottom. Clamp the gel plates to the top of the electrophoresis tank and fill the upper reservoir with 1× TBE so that the wells are covered.

A syringe with a bent needle may be used to remove air bubbles trapped under the gel, which will disrupt the current flow.

7. Use a DC power supply to prerun and warm the gel for at least 30 min at 5 V/cm (constant voltage).
8. Add 5× loading buffer to DNA samples and molecular-weight markers (to 1× final) and load on gel.

Load an amount of DNA that correlates with the visualization technique to be used. If the sample is to be UV shadowed (UNIT 2.12), then 2 μg of DNA will be required per band in a 2 cm \times 2 cm \times 1.6 mm well. With ethidium bromide staining, the detection limit is only 15 ng DNA per band; for good resolution using this method, only 25 μg of material should be loaded per 2 cm \times 2 cm \times 1.6 mm well.

Plastic disposable pipet tips are available in a variety of styles and sizes. Choose one that fits the application for use in loading the gel. Alternatively, particularly for larger volumes, use a micropipet or pulled plastic capillary, made as follows: Break a 1-ml disposable plastic (polystyrene) pipet in half and heat the center portion over a Bunsen burner until it is just soft (barely melted). Remove from heat for 1 to 2 sec and, holding vertically, quickly pull to generate a thin capillary. After the plastic has cooled, cut as desired. To calibrate, draw up a premeasured volume of water and mark capillary.

9. Run the gel at ~5 V/cm, taking care to avoid excessive heating. Run the gel until the desired resolution has been obtained as determined empirically or from Table 2.7.1.

From 2 to 10 V/cm is acceptable. If the gel is noticeably warm to the touch, the samples in the middle will run faster or may even be denatured.

Shorter electrophoresis times may be achieved by running the gel at higher voltage in a cold room so long as the temperature of the gel remains below the denaturation temperature of the sample.

10. Turn off the power supply and detach the gel plates from electrophoresis apparatus. Carefully pry apart the plates so that the gel is still attached to one plate.
- 11a. If desired (but only for samples $\geq 2 \mu\text{g}$), visualize the DNA with UV shadowing, using a TLC plate with a fluorescent indicator (see *UNIT 2.12*).
- 11b. Alternatively, stain the gel while still attached to the plate by submerging it for 5 to 10 min in 0.5 $\mu\text{g/ml}$ ethidium bromide. If necessary, soak the gel and plate in water for 10 to 30 min to remove nonintercalated ethidium bromide to lower the background absorption.
12. Carefully wrap the gel and plate with plastic wrap. Invert, place the gel onto a UV transilluminator, and photograph.

Longwave UV light transmits through plastic wrap. Alternatively, the gel can be put directly on the transilluminator. If a photograph is not required, a longwave UV light may be shined onto the stained preparative gel to locate the DNA fragment of interest. It is important to avoid unnecessarily long UV exposure, which will damage the nucleic acids. Unpolymerized acrylamide absorbs strongly at 211 nm and may also cause shadowing that is confined to the edges and wells of the gel.

Recover the DNA

13. Cut out the desired DNA band with a scalpel or razor blade.
14. Crush the gel into many fine pieces by pushing it through a 3-ml small-bore disposable syringe to aid the diffusion of the DNA from the matrix.

If planning to use electroelution, omit this step and proceed to the Alternate Protocol.

15. Collect the pieces in an appropriately sized tube.
16. Add 2 vol elution buffer for every volume of gel. Incubate the tube with rotation or in a shaking air incubator at room temperature until desired level of elution is obtained.

Since elution is a diffusion-controlled process, increasing the amount of buffer will improve elution efficiency. Also, note that larger DNAs will take longer to diffuse from the gel. If speed is essential and high yields are dispensable, enough sample can be obtained for most experiments in only a few hours of extraction. Increasing the temperature to 37°C will also speed the process. Yield may be increased upon repeated elutions. Small fragments (<300 bp) should be mostly eluted in 4 hr, but large fragments (>750 bp) will need to be eluted overnight.

17. Depending on the volume of the elution, pellet the gel fragments at room temperature for 10 min in a tabletop centrifuge or 1 min in a microcentrifuge. Pipet off the supernatant solution, taking care to avoid the polyacrylamide pieces.
18. Recover any residual DNA by rinsing the gel with a small volume of elution buffer. Recentrifuge if necessary and combine the two supernatant solutions.

If necessary, remove any remaining acrylamide pieces by filtering the supernatant through a syringe equipped with a disposable 0.2- μm filter.

Also, if the volume of elution buffer is too large to allow for convenient precipitation, it may be reduced by successive extractions against equal volumes of butanol to concentrate the sample. About $\frac{1}{5}$ vol of the aqueous layer is extracted into the organic butanol layer for every volume of butanol used. If too much butanol is added and the water is completely extracted into the butanol, simply add more water and concentrate again.

- Precipitate the DNA by adding 2 vol of 100% ethanol and chilling 30 min at -20°C or 10 min at -70°C . Pellet DNA by centrifuging 10 min at $12,000 \times g$.

It is generally not necessary to add carrier to aid precipitation since the small acrylamide polymers released from the gel slice will suffice. If carrier is necessary, then use 10 μg of a carrier such as tRNA or glycogen, depending on the application.

- Redissolve the DNA pellet in 100 μl TE buffer, pH 7.5. If necessary, transfer to a microcentrifuge tube. Add 10 μl of 3 M sodium acetate, pH 5.2, reprecipitate the DNA with 2 vol of 100% ethanol, and chill 30 min at -20°C or 10 min at -70°C . Recover the DNA by centrifugation as in step 19.
- Rinse the pellet twice with 70% ethanol. Allow pellet to dry, and resuspend in TE buffer, pH 7.5, if appropriate.

ELECTROELUTION OF SMALL DNA FRAGMENTS FROM POLYACRYLAMIDE GELS

**ALTERNATE
PROTOCOL**

When working with small DNAs (<300 bp), if time is at a premium, the passive elution by diffusion described in Basic Protocol 1 (steps 15 to 19) may be replaced with the following electroelution procedure. Recovery of DNA fragments should be similar to that obtained with passive elution.

Additional Materials (also see Basic Protocol 1)

0.5 \times TBE electrophoresis buffer, pH 8.0 (APPENDIX 2)

Small dialysis bag

Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5A)

1. Prepare and run the gel as described in the previous protocol (see Basic Protocol 1, steps 1 to 12). Cut out the DNA band of interest but do not crush.

The gel slabs are not crushed to allow easy monitoring of the electroelution process and to speed the elution of the DNA.

2. Place the gel slab in a small dialysis bag with an appropriate molecular weight cutoff, and add enough 0.5 \times TBE buffer to surround and immerse the slab.

Many commercial electroelution apparatuses with less cumbersome handling procedures are also available (e.g., from Schleicher & Schuell and Sialomed).

3. Place bag in a small horizontal electrophoresis apparatus containing 0.5 \times TBE buffer.
4. Electrophorese the DNA out of the polyacrylamide gel at ~ 4 V/cm across the apparatus for 2 hr for small DNAs (<300 bp) or 6 hr for longer DNAs.

Because elution times are variable and especially if near-complete recovery is required, the gel should be UV shadowed and stained again after elution to ensure that the DNA has been quantitatively removed. Should some DNA remain in the gel, elution should be continued.

5. Recover the 0.5 \times TBE buffer, now containing the eluted DNA, from the electrophoresis apparatus. Reverse the polarity of the apparatus for ~ 1 min to free any bound DNA, and rinse the gel slab and inner surface of the dialysis bag to recover residual DNA.
6. Add 0.1 vol of 3 M sodium acetate and ethanol precipitate (see Basic Protocol 1, steps 20 and 21).

A second ethanol precipitation may be done if desired.

**Preparation and
Analysis of DNA**

2.7.5

SIEVING AGAROSE GEL ELECTROPHORESIS

Sieving agarose is agarose chemically modified to give low gelling ($\leq 35^{\circ}\text{C}$) and melting ($\leq 65^{\circ}\text{C}$) temperatures. Sieving agarose gels are poured and run like conventional agarose gels, but resolve small DNA fragments like nondenaturing polyacrylamide gels. They finely resolve nucleic acid fragments <1000 bp and can distinguish fragments as small as 8 bp.

Materials

Sieving agarose (NuSieve GTG agarose)
TAE or TBE electrophoresis buffer, pH 8.0 (*APPENDIX 2*)

Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5A*) and isolating DNA using low gelling/melting temperature agarose (*UNIT 2.6*)

1. Melt 2% to 4% sieving agarose in the appropriate buffer as outlined by the manufacturer. Pour gel of desired size for ordinary agarose gel apparatus.

Sieving gels generally resolve DNA more slowly and with sharper resolution in TBE than in TAE. Smaller fragments (<300 bp) are best separated in TBE and larger ones in TAE.

2. Load sample and run gel as for an ordinary agarose gel (*UNIT 2.5A*).

In a 4% NuSieve GTG agarose gel in TBE buffer, bromphenol blue will migrate at a rate equal to DNA fragments of <20 bp and xylene cyanol at a rate equal to 150 bp.

The gel can be run with ethidium bromide (0.5 mg/ml) included; this will cause only a slight decrease in DNA mobility, due to unwinding of the helix as a result of intercalation of the dye.

CAUTION: Ethidium bromide is a mutagen.

3. Isolate the fragment (see *UNIT 2.6*, Alternate Protocol for low gelling/melting temperature agarose, steps 2 to 9).

Various manufactures have developed methods for the removal of ethidium bromide from agarose gel slices. Spin columns (Supelco) made of a negatively charged matrix that attracts ethidium bromide and repels DNA offer quick, easy handling along with high recovery.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Acrylamide/bisacrylamide 29:1 (w/w)

29 g acrylamide
1 g bisacrylamide
H₂O to 100 ml
Store ≤ 1 month at 4°C

CAUTION: Acrylamide is a neurotoxin. Always wear gloves when working with the unpolymerized monomer.

Elution buffer, pH 7.5

10 mM Tris-Cl, pH 7.5
50 mM NaCl
1 mM EDTA, pH 8.0

Loading buffer for nondenaturing PAGE, 5 \times

50 mM EDTA, pH 8.0
50 mM Tris-Cl, pH 8.0
50% (v/v) glycerol

COMMENTARY

Background Information

Polyacrylamide gel electrophoresis (PAGE) offers high resolution of low-molecular-weight nucleic acids. In particular, small DNA fragments (<500 bp) that are poorly resolved by ordinary agarose gels are easily separated on polyacrylamide gels. Depending on the pore size of the gel (3.5% to 20% polyacrylamide), a separation from 10 to 1000 bp can be achieved. The concentrations of acrylamide that give the maximum resolution of DNA fragments of different sizes have been empirically determined, as shown in Table 2.7.1.

Polyacrylamide gels have a much higher capacity for DNA than agarose gels. Up to 15 μg of material can be loaded per 2 cm \times 2 cm \times 1.6 mm well. This is particularly important for preparation of significant amounts of small fragments. Elution of fragments from polyacrylamide gels yields DNA that is generally devoid of contaminating material that could interfere with enzymes used in cloning, sequencing, or labeling DNA. These two qualities make polyacrylamide gels the preferred means for purifying significant quantities of small fragments.

A polyacrylamide gel is formed by the polymerization of acrylamide monomers into long chains, which are further covalently attached by a cross-linking agent, most commonly *N,N'*-methylenebisacrylamide. Polymerization of a polyacrylamide gel is initiated by free radicals provided by ammonium persulfate and stabilized by TEMED. The reaction takes ~10 to 20 min to go to completion, but the reaction rate can be varied by adjusting the TEMED and ammonium persulfate concentrations.

The pore size of a polyacrylamide gel is determined by the total percentage of acrylamide (the sum of the weights of the acrylamide monomer and cross-linker). Historically, this has been expressed as %T. For example, the 5%T gel described above would contain 5% (w/v) of acrylamide plus bisacrylamide. As the %T increases, the pore size decreases. An appropriate %T for various ranges of fragment sizes can be determined from Table 2.7.1. The migration distance (*D*) of double-stranded DNA through a nondenaturing gel is inversely proportional to the log of its molecular weight—that is, $D \approx -\log(\text{MW})$. The base composition of a sequence also affects its electrophoretic mobility and may cause aberrant migration with respect to what would be expected

based on size alone; hence, size markers of similar composition should be used to confirm the size of the fragment of interest. To separate fragments over a wide range of molecular weights, a pore-gradient gel can be used. In such a gel the pore size is larger at the top than at the bottom, and becomes more restrictive as the fragment runs down the gel. Such gradient gels are difficult to pour, however, and are not commonly used. A general description of gels as electric circuits can be found in the introduction to this chapter.

The unique pore size of sieving agarose allows separation of much smaller fragments than ordinary agarose. Sieving agarose gels in the range of 2% to 4% separate in the range of 8 to 1000 bp and provide an alternative to polyacrylamide gels. Although the bands on sieving gels are somewhat more diffuse and the resolution is slightly poorer, sieving gels are easier and faster to pour and run, making them useful for a variety of applications, such as checking ligation of linker monomers into ladders (UNIT 3.16). Like low gelling/melting temperature agarose, sieving agarose is used to purify DNA fragments because of its unique properties of melting at 65°C and remaining liquid at $\leq 35^\circ\text{C}$. However, DNA fragments prepared from polyacrylamide gels are generally cleaner than those prepared from sieving agarose and can be more reproducibly used as substrates for a variety of enzymes.

Critical Parameters and Troubleshooting

The most important parameter for the successful separation of small DNA fragments by polyacrylamide gels is the polymerization reaction itself. It is important to use only high-quality electrophoresis-grade reagents when running the gels. Acrylamide and bisacrylamide both break down in solution to acrylic acid, which affects the mobility of molecules through the gel matrix. Acrylamide solutions should be protected from light and should not be stored for more than a few months. Commercially prepared polyacrylamide solutions with long shelf lives (due to the incorporation of a gaseous inhibitor that prevents the initiation of polymerization) are available (National Diagnostics) and highly recommended. Ammonium persulfate is stable for ~1 month at 4°C. Clean plates are also essential in order to avoid the introduction of bubbles into the gel when pouring.

One of the most common problems encountered in polyacrylamide gels is "smiling," in which the lanes in the center of an overheated gel run faster than the lanes at the sides. This is caused by uneven dissipation of heat by the gel: the sides are cooler than the center and samples run faster at higher temperatures. There are several ways to avoid smiling, the simplest of which is to run the gel at lower voltage. An alternative is to use an apparatus that incorporates a mechanism such as a metal plate to disperse heat evenly throughout the gel, or an active cooling mechanism.

Sieving agarose does not have the integrity of regular agarose and is even more fragile than the common low gelling/melting temperature agaroses. The manufacturers recommend against using sieving agarose concentrations of <2%. Extreme care should be taken when pulling the combs out of the gel, as wells tear easily. It is usually best to remove the comb after the gel has been submerged in the electrophoresis tank. Because this type of agarose melts at 65°C, a low voltage (6 to 7 V/cm of gel) should be maintained to keep the gel from heating too much. In a 4% NuSieve GTG agarose gel in TBE buffer, bromphenol blue will migrate at a rate equal to DNA fragments of <20 bp and xylene cyanol at a rate equal to 150 bp.

Anticipated Results

Because of the high capacity of acrylamide relative to agarose, up to 2 µg of a fragment >250 bp and up to 5 µg of a smaller fragment can be purified on a 2 cm × 2 cm × 1 mm lane by the method presented in Basic Protocol 1. Up to 25 µg of material can be purified on the larger 2 cm × 2 cm × 1.6 mm preparative gels. After several hours of shaking at 37°C, the eluted yield should be 60% to 75% for larger fragment and >85% for smaller fragments. Essentially quantitative recovery will be obtainable by overnight elution. Similar recoveries can be obtained using the faster but slightly more cumbersome electroelution alternate method (Alternate Protocol).

Yields from sieving agarose gels (Basic Protocol 2) are similar to those from conventional low gelling/melting temperature agarose (≥70%) as long as fragments are <1000 bp. These gels resolve best below 500 bp.

Time Considerations

Although a polyacrylamide gel is less convenient than an agarose gel to set up and polymerize, this process should take <1 hr in

either case. After polymerization, gels can be stored overnight or even for several weeks, provided precautions are taken to prevent the slots from drying out: for instance, using a paper towel soaked with buffer and encased in plastic wrap to keep the gel hydrated. In general, the comb should be left in the gel and the top should be covered with plastic wrap. The limiting factor in polyacrylamide gel electrophoresis is heating of the gel, which results in smiling. Inclusion of a mechanism to disperse the heat and/or actively cool the gel can minimize this problem and greatly shorten electrophoresis times. If the Alternate Protocol for electroelution of fragments is followed, purified fragments can be obtained in <8 hr after pouring the gel.

The sieving agarose procedure is short, especially for minigels, and can be completed in 2 hr. After the addition of ethanol to the extracted fragment, the protocol can be interrupted for as long as desired.

Literature Cited

- Maniatis, T. and Ptashne, M. 1973a. Multiple repressor binding at the operators in bacteriophage. *Proc. Natl. Acad. Sci. U.S.A.* 70:1531-1535.
- Maniatis, T. and Ptashne, M. 1973b. Structure of the operators. *Nature* 246:133-136.
- Maniatis, T., Jeffrey, A., and van deSande, H. 1975. Chain length determination of small double- and single-stranded DNA molecules by polyacrylamide gel electrophoresis. *Biochemistry* 14:3787-3794.

Key Reference

- Chrambach, A. and Rodbard, D. 1971. Polyacrylamide gel electrophoresis. *Science* 172:440-451.

Provides background chemistry of polymerization of acrylamide and reviews the size separation characteristics of gels with different acrylamide percentages.

FMC Marine Colloids product information.

Sieving agarose was developed by FMC Marine Colloids. Literature describing its properties and use is available from the company.

Contributed by Joanne Chory
The Salk Institute for Biological Studies
La Jolla, California

Jack D. Pollard, Jr.
Harvard Medical School and Massachusetts
General Hospital
Boston, Massachusetts

Capillary Electrophoresis of DNA

DNA fragments are traditionally separated and analyzed by slab gel electrophoresis. The gel matrices are usually either polyacrylamide (UNIT 10.2A) or agarose (UNIT 2.5A), and separations are achieved in the presence (for ssDNA) or the absence (for dsDNA) of dissociating agents such as urea or formamide. The slab gel systems have the advantage of analyzing multiple samples in the same separation at low cost, but normally take several hours to complete. The DNA is typically visualized with stains, UV shadowing, intercalating dyes such as ethidium bromide, and on occasion by radioactivity. Capillary electrophoresis (CE), an alternative to conventional slab gel electrophoresis, has developed over the past few years into a very powerful tool for the separation of DNA fragments. CE offers a number of advantages over slab gel separations in terms of speed, resolution, sensitivity, and data handling. This is partly because the CE separation occurs inside a small-diameter (50- to 100- μm) quartz capillary in the presence of high (kilovolt-level) separating voltages. Separation times are generally only a few minutes. The DNA is detected either by UV absorption or by fluorescent labeling, both of which eliminate the need to use mutagenic substances (e.g., ethidium bromide) or dispose of radioactive waste. The quantity of DNA required for the separation is in the nanogram range. Single-base resolution can be readily obtained on fragments up to several hundred base pairs in size. In the presence of appropriate standards, fragments can be accurately sized, based on relative electrophoretic mobility.

The separation of DNA fragments by CE occurs within the walls of a fused-silica capillary. Since the negatively charged nature of this surface has a dramatic impact on the resolution achieved during the separations, the vast majority of CE separations are done in “coated” capillaries whose surface has been modified to be chemically inert to the DNA. The capillaries are filled with a sieving matrix, and the DNA fragments are separated on the basis of size, analogously to standard slab gel separations. The matrix is either a chemically cross-linked gel, such as polyacrylamide, or a flowable polymer, such as modified cellulose or non-cross-linked polyacrylamide. Single-stranded DNA (ssDNA) fragments as small as 5 bases are readily separated with single-base resolution. The analysis of synthetic oligonucleotides in a flowable matrix is described in this unit (see Basic Protocol 1) as an example of this type of application. Fragments of double-stranded DNA (dsDNA) as large as 20 kb are also separated, although not with single-base-pair resolution. The only difference between these separations is the separation matrix.

CE has found increasing use in a number of analytical applications where DNA separations are required. These include assessment of the purity of synthetic oligonucleotides and their modifications, analysis of PCR products, sequencing of fluorescent DNA, analysis of restriction maps, accurate sizing of restriction fragments for genetic analysis, forensic analysis of biological samples, genotyping (see Alternate Protocol), and analysis of conformational polymorphisms. Additional applications continue to be developed. An area of growing interest is the ability to analyze low levels of PCR products in biological fluids, as presented below (see Basic Protocol 2). Rapid progress is also being made in the development of multicapillary automated DNA sequencing instruments using laser fluorescence detection.

CE is an analytical technique rarely used in preparative mode. This is largely because only small quantities of DNA can be loaded onto a capillary. Amplifying DNA by PCR after separation can circumvent this problem. In general, however, preparative separation of DNA fragments is best achieved by slab gel electrophoresis (UNITS 2.5A, 2.7, & 2.12) or high-performance liquid chromatography (HPLC) methods (UNITS 10.12–10.14).

INSTRUMENTATION

CE separation in its simplest form can be achieved by passing a high voltage between two buffer reservoirs that are joined by a fused silica capillary filled with liquid or gel. This results in an electric field that drives the molecules of interest from one end of the capillary to the other. The capillaries are generally 20 to 80 cm long and 50 to 100 μm in internal diameter, with total volumes in the 1- to 2- μl range. For comparison, the volume of a slab gel lane is ~ 1000 μl . The capillaries are thin walled, which allows for dissipation of the Joule heating resulting from the high voltages (10 to 30 kV) that are necessary for high-performance electrophoretic separations. This minimizes convective effects that could result in band broadening during electrophoresis. The fused-silica capillary is coated on the outside with a polyimide layer that eliminates oxidation of the fused-silica glass and confers excellent tensile strength to the otherwise fragile capillary. The polyimide sheathing is carefully burned from a small portion of the capillary to expose a section of the silica. This clear section of the capillary is inserted into the light path of a UV or fluorescence detector and becomes the on-column flow cell. As the DNA molecules migrate through the capillary as a result of the electric field, they pass through the detector light path and are measured by UV or fluorescence detection. In effect, the separation column itself becomes a very-low-volume flow cell.

IMPORTANT NOTE: Removal of the polyimide coating makes the capillaries susceptible to breakage. Capillaries that are not provided in cartridges by the manufacturer should be handled with care to avoid breakage.

The combination of high field strength and large surface-area-to-volume ratio of the capillaries results in rapid and very efficient separations of both ssDNA and dsDNA. Sample loading can be accomplished from as little as 1 μl , with starting sample concentrations of ~ 1 $\mu\text{g}/\text{ml}$ for UV detection and ~ 1 pg/ml or less for laser-induced fluorescence detection. Clearly, with respect to sensitivity, speed, and versatility, CE offers significant advantages over gel electrophoresis for the separation of nucleic acids.

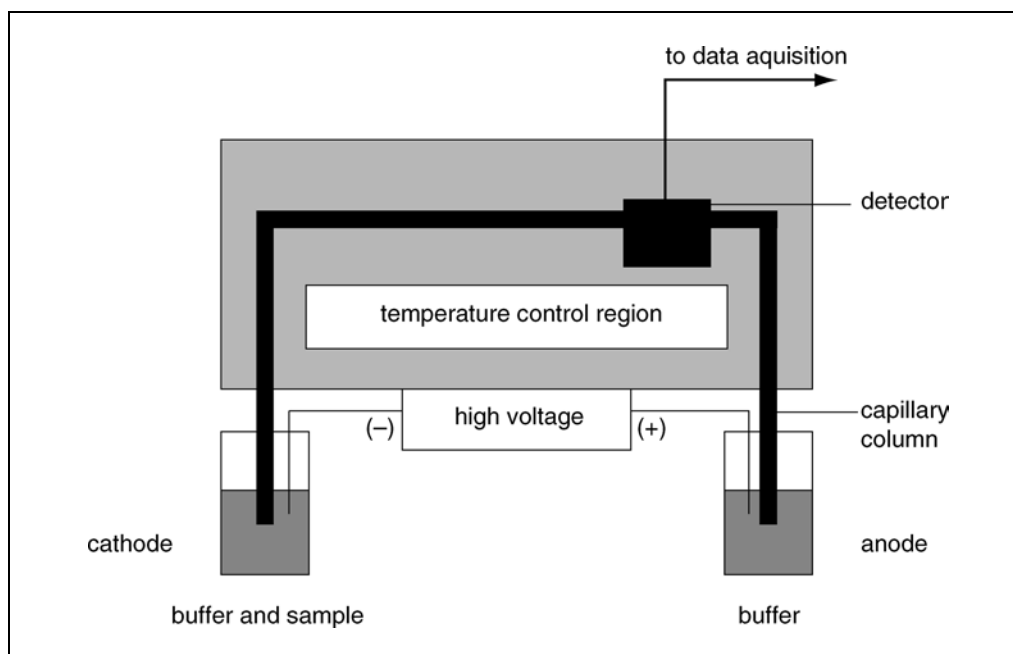


Figure 2.8.1 Schematic of a CE instrument configured for DNA separations.

Table 2.8.1 Capillary Electrophoresis Systems

Manufacturer	Model number ^a	Autosampler	Detection method ^b	Column temperature
Agilent	CE System	48-position, 10° to 40°C	Diode array UV-vis	15° to 60°C
Amersham	MegaBACE 500 16, 32 or 48-capillary DNA sequencer	96-well plate, no cooling	4-color scanning LIF detection	27° to 44°C
	MegaBACE 1000 96-capillary DNA sequencer	Single 96-well plate, no cooling	4-color scanning LIF detection	27° to 44°C
	MegaBACE 4000 384-capillary DNA sequencer	Single 384-well plate, no cooling	4-color scanning LIF detection	27° to 44°C
Applied Biosystems	Prism 310	482- or 96-sample position, cooled 96-well	4-color LIF with CCD	30° to 60°C
	Prism 3100 16-capillary DNA sequencer	96-well plate with sample cooling	CCD	26° to 65°C
	Prism 3730 96-capillary DNA sequencer	96-well plates with autoloader	CCD	18° to 70°C
Beckman Coulter	P/ACE MDQ DNA System	Up to 96-well plate	UV and diode array detection	15° to 60°C
	CEQ 8000 8-capillary DNA sequencer	96-well plate with cooling and heating to 90°C	4-color CCD	30° to 60°C

^aAll units listed have a single capillary unless otherwise noted.

^bLIF, laser-induced fluorescence detector, used with either a PMT (photomultiplier tube) or a CCD (charge-coupled device) camera for visualization; UV, ultraviolet detector.

As stated previously, in its simplest form capillary electrophoretic separation can be achieved by passing a current between cathodic and anodic buffer reservoirs via a liquid-filled glass capillary. In practice, the basic CE instrument also requires a suitable sample injection module, a detector, adequate temperature control, and isolation for the user from the high voltages used for the separations. A schematic of the basic instrument is shown in Figure 2.8.1. There have been a number of changes recently in the features of commercially available instruments; these instruments and their capabilities are summarized in Table 2.8.1.

SEPARATION THEORY

CE is part of the family of electrophoretic techniques that separate species based upon their size and ionic properties. An ion (*i*) placed in an electric field will move in the direction parallel to the field towards the oppositely charged electrode with a velocity (v_i) defined as follows:

$$v_i = \mu_i E = \mu_i V / L$$

where μ_i is the mobility of the ion, E is the electric field in volts per centimeter, V is the voltage across the column, and L is the total column length. The electrophoretic mobility of a given ion is equal to:

$$m_i = q_i / 6\pi\eta a_i$$

where q_i is the charge on the ion, η is the viscosity of the buffer or gel matrix, and a_i is the radius of the ion.

Since DNA has a constant size-to-charge ratio, a sieving matrix must be added to the capillary in order to discriminate based only on size, rather than charge and size. In CE there are two types of gels employed in DNA separations: cross-linked gels (static gels) and non-cross-linked gels (flowable polymers or polymer networks). Cross-linked gels are fixed gels that are polymerized inside the capillary, usually covalently bound to the capillary surface, and are not removed from the capillary between runs. Flowable polymers are viscous hydrophilic polymer solutions that can be pumped into the capillary. The same flowable polymer matrix can be used repeatedly when small molecules such as synthetic oligonucleotides are being analyzed. The time between injections is sufficient for the preceding sample to clear the detector. Alternatively, the polymer can be used once, discarded, and replaced with fresh matrix prior to the next injection. This approach is preferred where larger DNA molecules are present in the samples—e.g., for fragment analysis and DNA sequencing analysis. Usually, a coated capillary is utilized to eliminate the charge effects that are contributed by the native silica surface. With cellulose-derived polymers or some specially modified acrylamides, however, uncoated capillaries may be used, because of the strong interaction of the polymer with the inner surface of the bare fused-silica capillary, in essence forming its own coating.

With either a cross-linked or non-cross-linked gel in the capillary, the matrix offers a frictional resistance to the movement of the DNA through the gel medium that is proportional to the size of the species. The frictional resistance can vary with the molecular weight, concentration, and chemical composition of the flowable gel polymer or the pore size in the cross-linked gel, and must be optimized for the particular size of the DNA to be separated. A detailed description of the theory of DNA motility in entangled polymer solutions can be found in Grossman (1991).

STRATEGIC PLANNING

The most common approach to the separation of both ssDNA and dsDNA by CE uses a coated capillary and an uncharged sieving matrix. This is very similar to slab gel electrophoresis, but in a silica capillary. The separation matrix, as mentioned, can take the form of a cross-linked polyacrylamide gel or flowable polymer such as hydroxypropyl methyl cellulose (HPMC), hydroxyethylcellulose (HEC), polyethylene oxide (PEO), or non-cross-linked linear polyacrylamide. The cross-linked gel is polymerized directly inside the capillary and can be reused for 30 to 100 separations before losing resolution. The capillary is then discarded, since the polyacrylamide gel cannot be regenerated. The flowable polymer has the advantage that it can be expelled from the capillary by pressure at the end of each electrophoretic separation; fresh matrix is then reloaded into the capillary prior to the next separation. These capillaries have lifetimes of several hundred injections. The eventual loss of the surface coating is the major reason for replacement; another common reason is mechanical breakage.

The selection of the appropriate matrix can significantly affect the quality of the separation. Cross-linked polyacrylamide is best used for the separation of synthetic oligonucleotides—both native and modified versions. However, flowable polymers can also be used for oligonucleotide analysis and for the separations of automated sequencing ladders. Where dsDNA fragment analysis is required, only flowable polymers are routinely used. The general rule for matrix selection is that the larger the DNA fragment, the weaker the sieving capabilities of the matrix.

Separation buffers frequently are variants of Tris/borate/EDTA (TBE) mixtures and are buffered at alkaline pH. Urea is often included in the buffer, as a denaturant, when analyzing ssDNA (e.g., synthetic oligonucleotides). Samples are loaded onto the capillary by electrokinetic, or pressure, injection. Separation times range from 10 to 45 min, at

voltages between 1 and 10 kV. DNA fragments are detected in the UV spectrum at 260 nm, either in the presence or absence of ethidium bromide. Sensitivity can be increased by at least two orders of magnitude through the use of fluorescence detection. Since DNA possesses no native fluorescence, intercalating dyes such as cyanine derivatives (Zhu et al., 1994) or rhodamine derivatives must be added to the electrophoresis buffer, or covalently attached to the DNA prior to the electrophoretic separation. In addition to increasing the sensitivity of detection, these intercalating dyes can improve resolution and sharpening of the bands by physically disrupting the DNA structure. The selection of specific dyes is dictated by their excitation and emission spectra and the compatibility with the detection systems of individual instruments. Specific examples of intercalating dyes are: thiazole orange (Aldrich), YO-PRO-1, YOYO-1, and SYBR Green (Molecular Probes). The added sensitivity is particularly useful when analyzing PCR products that have been amplified from biological fluids. Targets of interest are frequently present in small amounts, and the presence of salts and proteins make their direct analysis by CE impractical. However, after completion of the PCR reaction, the sample can be diluted with water and an aliquot analyzed using fluorescent detection.

The mobility of a given DNA fragment may not be constant over a series of injections. This variability can have a variety of causes: aging of the polymer (polyacrylamide), loss of capillary coating, or depletion of the conductivity of the running buffers. The absolute mobility of DNA in a given sample will be dependent upon the salt content (and hence the conductance) of that sample. The presence of high salt will significantly reduce the electrophoretic mobility of the DNA. One solution is to dilute the sample in water and load for longer times; alternatively, the sample can be desalted (*UNIT 10.9*) prior to injection. Where accurate sizing is important, it is essential to incorporate sizing standards into the sample prior to electrophoresis.

The CE analysis of synthetic oligonucleotides requires the selection of a matrix that optimizes resolution of low-molecular-weight oligonucleotides. The separation of fluorescently labeled fragments from an automated sequencing ladder represents a specialized CE application and requires the selection of a matrix with a greater resolution range. These ladders range in size from 20 to more than 1000 bases and can be separated with single-base resolution to a high degree of accuracy. Currently, three automated CE instruments are commercially available as DNA sequencers (see Table 2.8.1). A detailed discussion of the general principles associated with dideoxy sequencing can be found in *UNIT 7.4*.

SEPARATION OF OLIGONUCLEOTIDES

In this protocol synthetic oligonucleotides are analyzed for purity by CE using a replaceable, flowable polymer as the separation matrix. The running buffer and separation matrix contain 7 M urea to keep the DNA in its single-stranded configuration. The sample is loaded onto the capillary at the cathode by electrochemical injection. After loading, the sample vial is replaced by the cathode buffer reservoir, and the electrophoresis is continued. The matrix does not need to be replaced between each separation, but should be replaced at the beginning of each series of separations, e.g., at the beginning of each day. Each time fresh matrix is loaded, the capillary must be equilibrated before samples are run. No further equilibration is required between samples. The electrophoretic separation should provide single-base resolution for DNAs of at least 100 bases. A poly(A)₄₀₋₆₀ size ladder (see Fig. 2.8.2) should be analyzed initially in order to confirm that resolution is optimal.

NOTE: The filled capillary can be stored on the instrument overnight, but if >1 day elapses between runs, the capillary should be stored at 4°C with both ends capped.

BASIC PROTOCOL 1

Preparation and Analysis of DNA

2.8.5

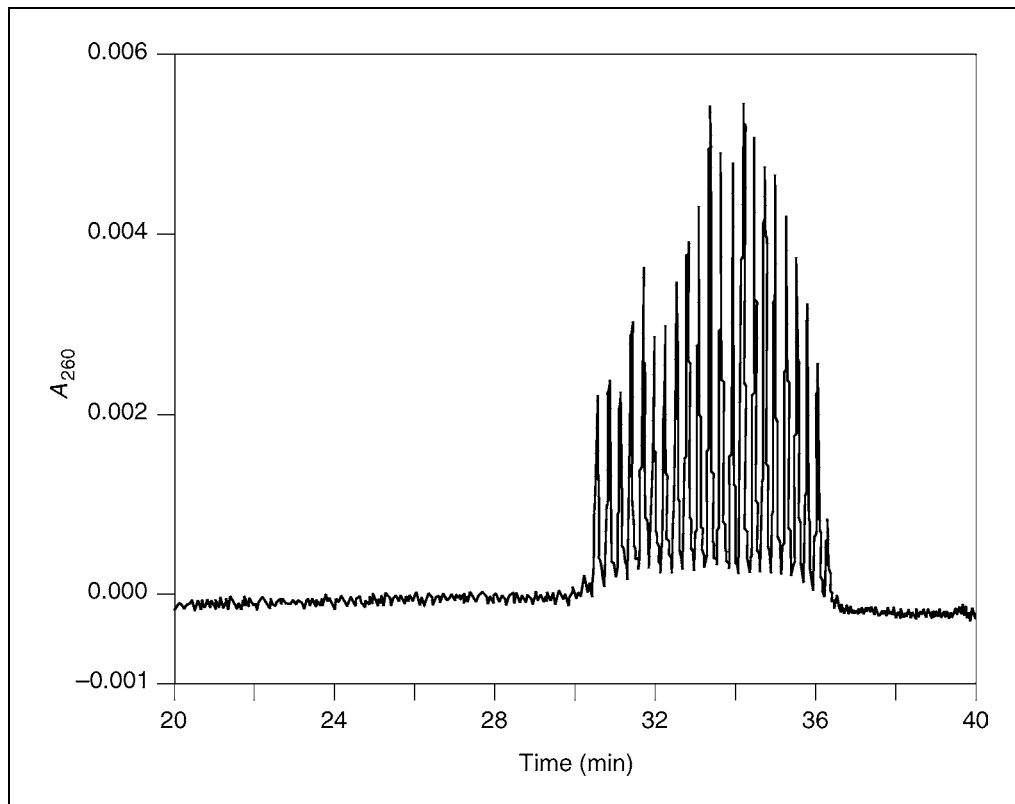


Figure 2.8.2 CE separation of a standard poly(A)₄₀₋₆₀ mixture of synthetic oligonucleotides.

NOTE: The following protocol demonstrates the use of a P/ACE 5510 CE instrument from Beckman Coulter. However, other instruments (Table 2.8.1) are capable of comparable separations when operated in accordance with manufacturers' instructions.

NOTE: The selection of a matrix is often instrument-dependent. It is recommended that a kit be used initially for a reference separation. Linear polyacrylamide is frequently used for oligonucleotide separations.

Materials

ssDNA 100-R separation kit (Beckman Coulter) including:

60-cm, 100- μ m-i.d. coated capillary

ssDNA 100-R separation gel solution

Running buffer: Tris-borate electrophoresis buffer (reconstitute and store up to 30 days at 4°C)

Poly(A)₄₀₋₆₀ sizing standard (dissolve at 100 μ g/ml [3 OD₂₆₀ units/ml] in water and store indefinitely at -20°C)

Dried ssDNA oligonucleotide sample

CE instrument (e.g., Beckman Coulter P/ACE 5510 or equivalent; see Table 2.8.1)

1. Reverse standard polarity of the CE instrument electrodes (see manufacturer's instructions).
2. Rinse capillary on-instrument with deionized water for 5 min.
3. Fill the capillary with ssDNA 100-R gel solution using a 20-min pressure rinse from the matrix vial (based on a 20-psi rinse pressure).

This solution can be stored on-instrument for 5 days, but should then be discarded.

4. Equilibrate the capillary in running buffer by running a voltage ramp from 0 to 8.1 kV over 20 min and holding at 8.1 kV for 10 min.

5. Replace the inlet reservoir with a container of water and inject for 1 sec at 7.5 kV.
6. Position the sizing standard vial at the inlet and inject for 10 sec at 7.5 kV.
7. Replace with running buffer reservoir and carry out the electrophoresis at 8.1 kV for 40 min at 30°C.
8. At completion of run, confirm that the separation of the standards is satisfactory by comparison with the example provided in the kit.
9. Prepare samples to be analyzed by dissolving in water to ~10 µg/ml.
10. Load samples onto autosampler and inject for 10 sec at 7.5 kV.
11. Carry out the electrophoresis at 8.1 kV for 40 min at 30°C.
12. Repeat steps 10 and 11 until resolution begins to deteriorate (at least 15 runs); then replace separation gel by returning to step 2.

QUANTITATIVE PCR ANALYSIS

Quantitative PCR can be used in conjunction with CE separation to amplify and quantitate any DNA target sequence—by the use of either an intercalating dye (coinjected with the samples) or a covalently modified, fluorescently labeled oligonucleotide primer. The size of the expected product is determined by coinjection of sizing standards. Quantitation is achieved by the coamplification of a second target sequence of known concentration, or by the addition of a known quantity of DNA to each sample.

Another application of this method is the direct measurement of viral load by reverse transcription (RT)-PCR of the viral RNA. This is achieved by the procedure known as competitive PCR analysis (Piatak, 1993). A known amount of a standard DNA template is included in the reaction mixture to compete for amplification with the target DNA. The sequence of the competing DNA is designed such that the PCR product is similar, but not identical, in size to the target DNA. The small quantities of DNA that are produced by this process require coinjection of an intercalating dye (e.g., YO-PRO-1 or SYBR Green I) as well as fluorescence detection.

Since dsDNA is being analyzed in these applications, it is not necessary to include denaturant in the electrophoresis buffer.

Materials

LIFluor dsDNA 1000 kit (Beckman Coulter) containing:

- Gel buffer mixture (containing separating gel and Tris/borate/EDTA buffer)
- Enhance intercalating dye
- 65-cm, 100-µm-i.d. coated capillary
- Standard sizing ladder: *Hae*III restriction digest of ϕ X-174 DNA (dissolve at 10 µg/ml in deionized water and store at –20°C)
- PCR reaction mixes containing amplicon
- CE instrument with fluorescent detection (e.g., Beckman Coulter P/ACE 5510 or equivalent)

1. Prepare gel buffer mixture according to manufacturer's instructions and add 0.4 µg/ml intercalating dye.

Store up to 30 days at 4°C.

2. Reverse standard polarity of CE instrument electrodes (see manufacturer's instructions).

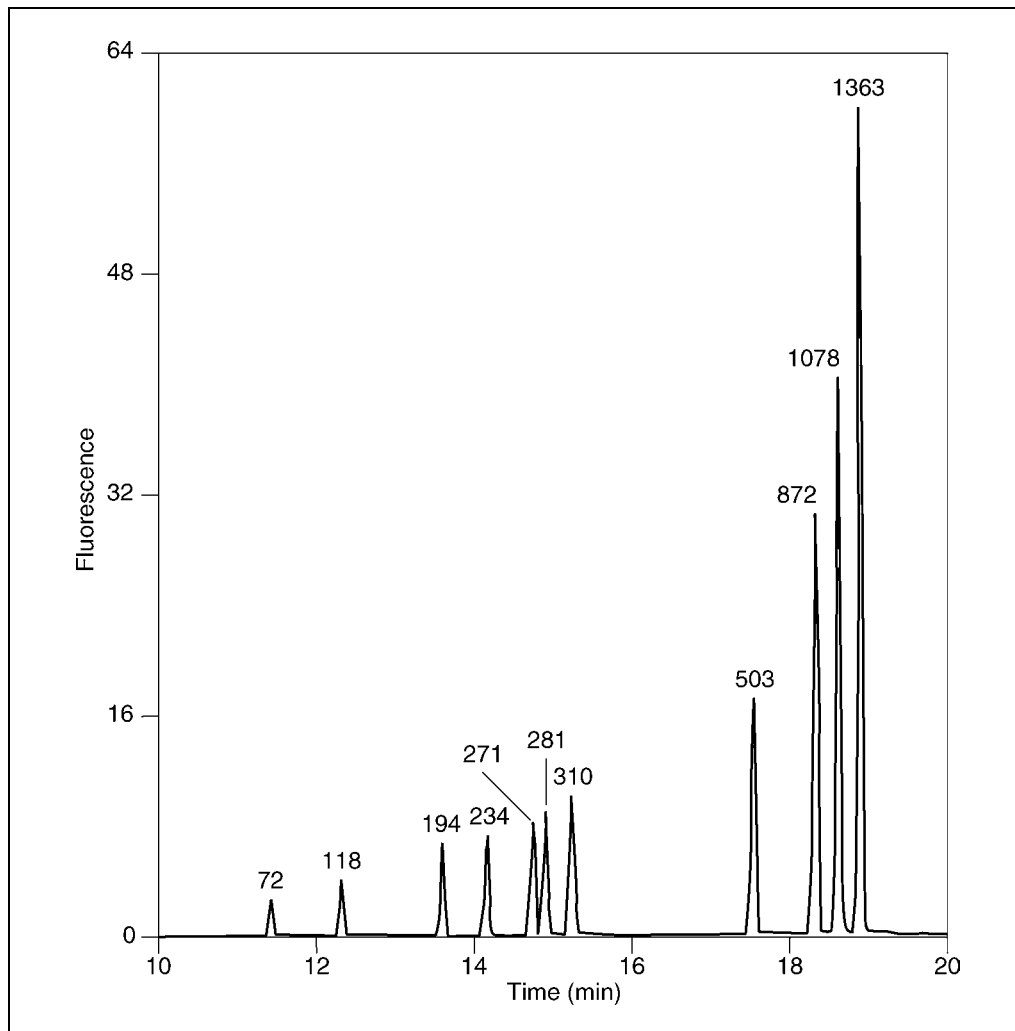


Figure 2.8.3 A sizing ladder from a *Hae*III digest of ϕ X-174 DNA. Fragment size in bp is indicated.

3. If the capillary is new, rinse with gel buffer for 10 min at high (20 psi) pressure.
This step is not necessary between runs.
4. Place standard sizing ladder in inlet position and load for 10 sec at low (0.5 psi) pressure.
See Figure 2.8.3 for a chromatogram of this standard.
The buffer ions become depleted over a series of injections. Consequently the inlet gel reservoir should be replaced after 30 injections.
5. Perform electrophoresis at 9.4 kV for 30 min at 25°C.
6. At the completion of the separation, replace the gel matrix using a 3-min high-pressure wash.
7. Assess the resolution and the linearity of area quantitation of the electrophoretic separation, and compare with the expected profile.
8. Dilute the PCR reaction mixes 10-fold with water and analyze by repeating steps 4 to 6, as with the sizing ladder.

Some applications may require the coinjection of the sizing ladder with the sample. In these instances the sizing ladder is loaded for 10 sec, followed by the sample for 10 sec.

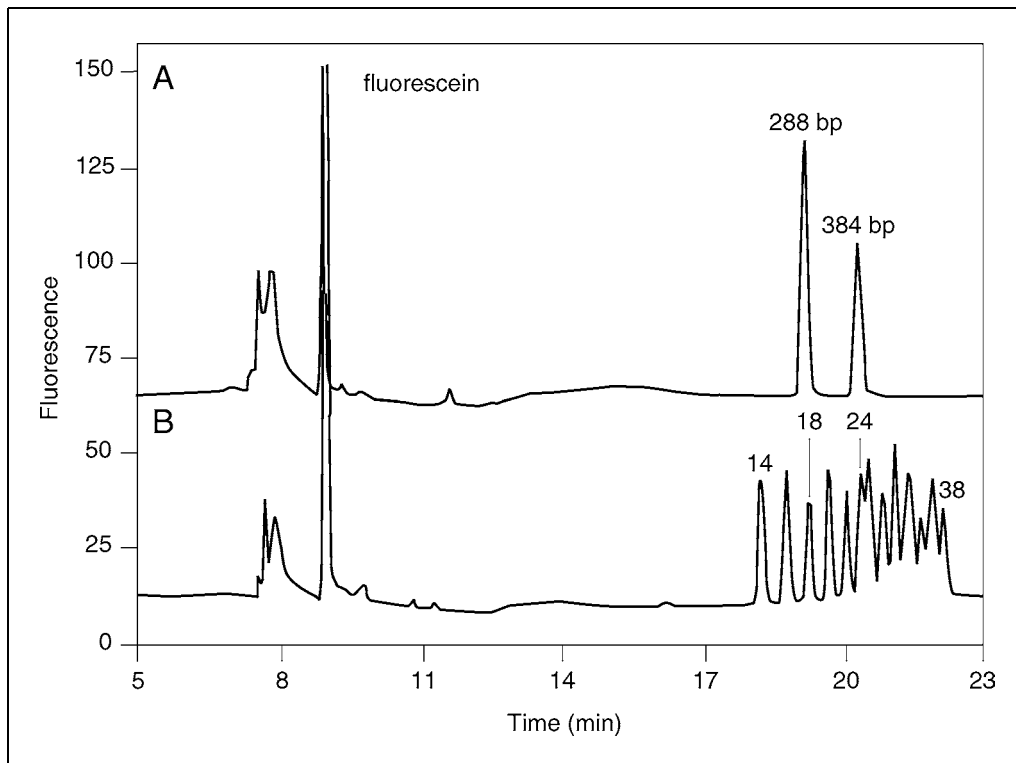


Figure 2.8.4 An example of VNTR analysis with (A) the D1S80 locus PCR products and (B) the D1S80 allelic ladder with a repeat unit of 16 bp and a range of 14 to 38 copies (as labeled on peaks).

GENOTYPING

The use of capillary gel electrophoresis in genotyping can be viewed as an extension of the fragment-sizing application of RT-PCR analysis (Basic Protocol 2). Two major differences are the source of the target DNA and the analysis method used. Genotyping procedures begin with genomic DNA, which is either assessed directly by restriction fragment length polymorphism (RFLP; Jarcho, 1994) analysis or amplified and then subjected to variable number of tandem repeat (VNTR; Sozer et al., 1998) analysis. The fragments can be detected by using either UV or fluorescence detectors. The same Beckman Coulter dsDNA 1000 analysis kit that is used for quantitative PCR can be used for genotyping, since the fragment size ranges are similar.

In the case of VNTR analysis, where the number of repeats is being measured, it is important to select standards whose size corresponds approximately to the expected number of repeats. Standards can be obtained from a number of suppliers. For example, a D1S80 allelic ladder, with a range of 300 to 700 bp (see Fig. 2.8.4) and a 16-bp repeating unit, is available from Lifecodes. It is advisable to coinject the sizing standards with the sample for the greatest accuracy. RFLP analysis simply requires pattern matching. However, size standards are required for the alignment of multiple analyses.

The CE setup, sample loading, and running conditions for this method are identical to those for quantitative PCR (see Basic Protocol 2).

COMMENTARY

Background Information

The application of capillary electrophoresis to the separation of DNA fragments is a relatively recent development. The inherent

advantages of this technology—high resolution, excellent sensitivity, and rapid separation times—provide significant improvements over the conventional slab gel electrophoresis

ALTERNATE PROTOCOL

Preparation and Analysis of DNA

2.8.9

technology. This has led to rapid acceptance of CE as an essential tool for the analysis of DNA fragments of all sizes.

When applied to the analysis of synthetic oligonucleotides (see Basic Protocol 1), the technique's ability to obtain single-base resolution has proven extremely useful in diagnosing synthesis problems and in obtaining overall purity estimates for the final products (Cohen et al., 1988). The chemistry of DNA synthesis involves sequential addition of bases to the growing oligonucleotide chain in a prescribed order (UNIT 2.11). After each addition, the growing chain is "capped" in order to terminate the portion of the oligonucleotide that has not completed the coupling reaction. This yields a truncation series. If present in sufficient quantity, these species will compete with the full-length oligonucleotide in some applications. Since they are shorter than the full-length oligonucleotide they can readily be resolved and quantitated by CE.

After synthesis has been completed, the bases have to be deprotected. If deprotection is incomplete it can interfere with the base-pairing properties of the oligonucleotide. These species, which appear larger than the

full-length oligonucleotide, can also be readily resolved by CE. Figure 2.8.5 illustrates a crude oligonucleotide containing coupling failures (shorter migration times) and incompletely deprotected species (longer migration times). The resolution capacity of CE using a matrix optimized for oligonucleotides can extend to the analysis of oligonucleotides in excess of 100 bases. An example of the separation of a standard oligonucleotide mixture ranging from 40 to 60 bases is presented in Figure 2.8.2. Single-base resolution is readily observed.

An increasing number of oligonucleotide applications require modification of the basic oligonucleotide probe—e.g., biotinylation, fluorescent dye modification, phosphorylation, base modification, addition of phosphorothioate (antisense) backbones, DNA-RNA hybridization, and 5'-amino- or 5'-thiol-modification—prior to conjugation to other molecular species. In all instances, the extent of modification to produce these unique oligonucleotides can be readily assessed by CE due to its high resolution. Recent developments in the area of quantitative gene expression measurements (Freeman et al., 1999) and single nucleotide polymorphism

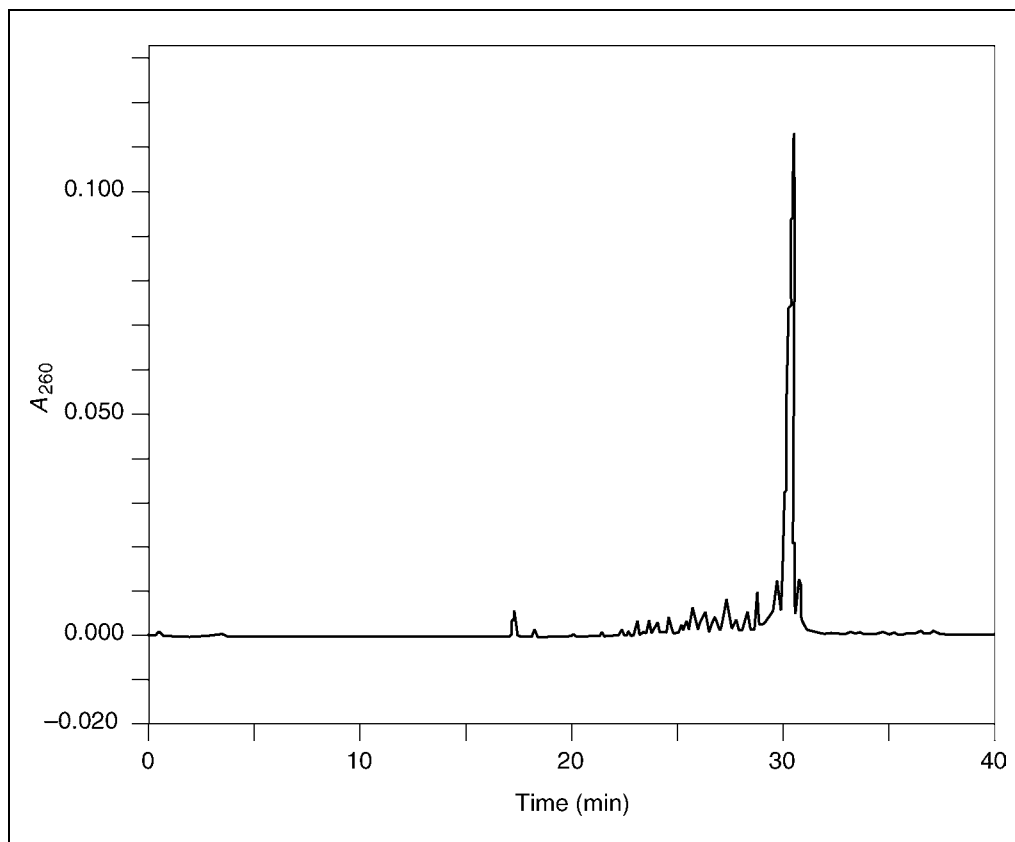


Figure 2.8.5 CE separation of a synthetic 20-mer oligonucleotide. The coupling failure products (shorter migration times) and incomplete deprotection products (longer migration times) are clearly visible.

(SNP) detection have utilized the principle of fluorescence energy transfer via oligonucleotide probes. Assays based on quantitative RT-PCR have been developed—e.g., TaqMan (Applied Biosystems), Invader (Third Wave), and Black Hole Quenchers (BioSearch Technologies)—that require either paired single-labeled or dual-labeled oligonucleotides. These reagents are readily analyzed for purity by CE using Basic Protocol 1. CE is particularly useful in this instance since MALDI mass spectrometry (*UNIT 10.21*), a commonly used alternative, fragments the fluorescent label, resulting in multiple molecular species.

Small interfering RNAs (siRNA) can be used for the down-regulation of individual genes (Chapter 26; Elbashir et al., 2001). These molecules all share the same features: 2 thymidine deoxynucleotides at the 3' end followed by 19 ribonucleotides, terminating in a 5'-hydroxyl. These siRNA probes are readily characterized for purity by CE using Basic Protocol 1.

The development of matrices to extend this single-base resolution to ssDNA that ranges in size from 20 to 1000 bases has allowed for development of important applications in the field of automated fluorescence-based DNA sequencing. The most commonly performed sequencing chemistry is the “dye terminator” chemistry, in which the sequence-terminating dideoxy nucleotide also contains the fluorescent reporter group. Consequently, the sequence ladder is labeled in the 3'-hydroxyl position. This means that sequencing reactions can be performed using a primer of any

sequence. The alternate “dye primer” chemistry employs a primer that is labeled at the 5'-hydroxyl and is restricted to a small number of commonly used vector sequences—e.g., M13 forward and reverse, T3, T7, and SP6 sequences. An example of this type of application is shown in Figure 2.8.6. Significant effort is being directed towards the development of matrix formulations that will further extend the length of the current sequence reads. Commercial CE sequencing instruments all have shorter run times than slab gel-based automated sequencers and retain comparable sensitivity and accuracy. Each of these applications requires the analysis of ssDNA and is performed under dissociating conditions in the presence of a flowable matrix.

The sizing of larger fragments that have been amplified from genomic DNA has proven to be a very effective method for studying genetic variability in populations. These fragments can be analyzed by capillary gel electrophoresis either under dissociating conditions (short fragments, high size accuracy) or as dsDNA (large fragments, lower resolution). Genomic DNA from eukaryotes contains a large number of tandem-repeating sequences that vary in size from 2 to several dozen bp in size. This polymorphism can be used to advantage when studying human identity or individual heredity. The smaller repeats, commonly referred to as microsatellite or short tandem repeats (STR), are best analyzed as ssDNA (Butler et al., 1994). To determine identity or heredity it is necessary to accurately determine the number of copies of (2- to 5-bp-long) identical sequence repeats in the

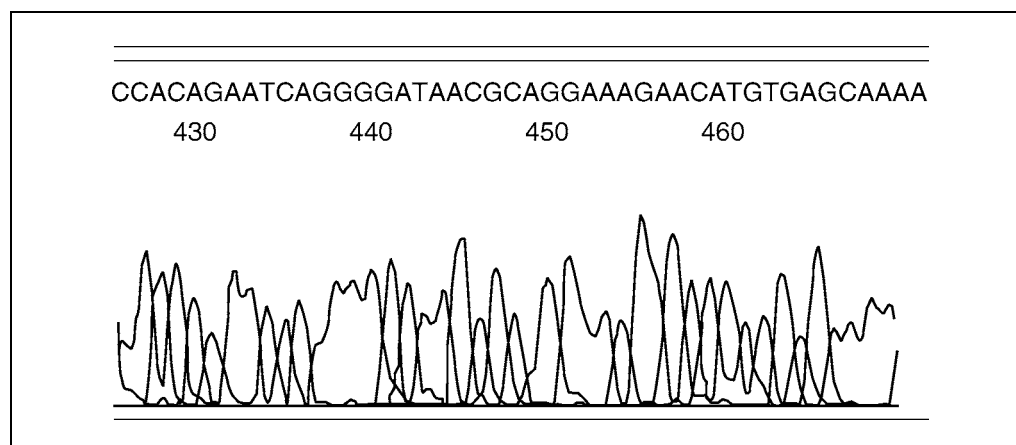


Figure 2.8.6 A portion of the chromatographic output from a PE Biosystems model 373A DNA sequencer using dye terminator sequencing chemistry. Although shown in black and white here, the direct output is color-coded to more clearly illustrate the peaks corresponding to each base in the sequence.

selected fragment. Fragment lengths are usually <300 bp and require single-nucleotide resolution.

Larger repeats are commonly referred to as variable number of tandem repeats (VNTR) and produce larger-sized fragments that are analyzed as dsDNA. The sizes of the repeats are 10 to 12 nucleotides and up, and fragment sizes can exceed 4000 bp. An example of this type of separation was presented in Figure 2.8.4 and explained in the Alternate Protocol. The required resolution for this type of separation is on the order of 4%.

In contrast to the accurate sizing of DNA, fragment pattern matching can also be a very useful analytical tool. The loss or gain of a restriction site between individuals is sufficiently common that restriction fragment length polymorphism (RFLP) analysis can demonstrate individual identity to a high degree of certainty (Ulfelder et al., 1992). These patterns are readily analyzed by CE. RFLP has found application in the fields of forensic medicine and pedigree testing. Where a polymorphic mutation produces a detectable phenotype, it can also be used as a diagnostic for inherited diseases. It requires the separation of dsDNA in the presence of an intercalating dye for visualization. The matrix is generally a flowable polymer, and fragment lengths are from hundreds to thousands of bp in size.

Mutation detection can also be performed using ssDNA. A denatured DNA fragment can adopt a sequence-specific conformation upon refolding, which will affect its electrophoretic mobility. This property is taken advantage of in single-stranded conformational polymorphism (SSCP) separations (Ren et al., 1997). The DNA region of interest is PCR-amplified from genomic or cDNA to give sufficient copies of a small (<300-bp) fragment. The strands are separated and allowed to refold. The wild-type and mutant fragments adopt different conformations and are resolved with single-base-pair resolution by CE using a nondissociating medium and a flowable polymer matrix.

CE can accurately quantitate and size DNA fragments (Rossomando et al., 1991; Fasco et al., 1995). This can be of considerable value when used to quantitate levels of viruses and pathogenic bacteria. The difficulty arises when quantitation is attempted in the presence of high levels of other DNA, RNA, and proteins. Fluorescence detection coupled with PCR—which can amplify very low levels of DNA in a highly specific manner—can be used to surmount this problem. This amplification can be

combined with the inherent sensitivity of CE through the incorporation of fluorescently labeled primers into the amplified DNA. In this fashion samples such as blood can be analyzed at high dilution, thereby reducing the levels of interfering substances to manageable levels. RNA can be amplified by RT-PCR to give the dsDNA fragment that is subsequently quantitated. This is ideal for measuring low levels of viral messenger RNA, i.e., viral load. The fragments are <1000 bp and separations are usually performed under nondenaturing conditions. The use of appropriate quantitative calibration standards is essential. One such standard, a sizing ladder, is shown in Figure 2.8.3, and an example of this technique is presented in Basic Protocol 2.

A majority of CE separations are performed at the standard alkaline pH in the presence of borate, which is an effective buffer in this pH range. The buffer may also contain 6 to 8 M urea, a denaturant that keeps the DNA in its simple single-stranded conformation when required. The urea is omitted from the buffer for analyses where secondary structure plays an important role in the separation, e.g., single-nucleotide polymorphisms or conformational polymorphisms.

The type of matrix that is selected for the actual separation can dramatically influence the quality of the separation that is achieved for a given application. Instrument manufacturers frequently supply kits that have been optimized for a particular application (see Table 2.8.2). These can be very helpful as a first, and maybe the only step required for optimizing individual applications. In general, cross-linked polyacrylamide that is polymerized inside, and covalently attached to, the capillary is best suited for smaller-fragment separations. However, since the column is reused multiple times, many things can reduce resolution—for example, capillary plugging, bubble formation, or drying of the capillary end—and require that the capillary be discarded. In the absence of such external parameters, the lifetime of the capillary is ultimately dependent on the breakdown of the polyacrylamide (or other hydrophilic polymer) matrix.

Flowable polymers have the advantage of wide fragment-separation ranges. These polymers can have a variety of chemical origins (see Table 2.8.2) and have the advantage of being replaced after each separation. This is achieved by applying pressure to the inlet end of the capillary. However, some of these polymers are quite viscous and require considerable pressure within the instrument

Table 2.8.2 Supplies and Kits for CE Systems

Manufacturer	Part number	Supplies and kits	Application information
Agilent	192-1311	μPAGE-10 (10% T, 0% C) 100-μm i.d.	
	192-5211	μPAGE-5 (5% T, 5% C) 75-μm i.d.	
	192-3211	μPAGE-3 (3% T, 3% C) 75-μm i.d.	
Amersham	25-6001-0	MegaBACE SNUPe genotyping application kit	Single-base primer extension for SNP analysis
Applied Biosystems	402838	POP-4	Gel for microsatellite, SNP, differential display, AFLP, and other genotyping applications using capillary #402839 and buffer #402824 on the 310 platform
	4316355	POP-4	Gel for microsatellite, SNP, differential display, AFLP, and other genotyping applications using capillary array #4315930 and buffer #402824 on the 3100 platform
	4313087	POP-5	Gel for high-throughput DNA sequencing on the 3730 platform
	402844	POP-6	Sequencing applications on the 310 platform, including template suppression reagent (TSR) for 67-cm capillary #402840 and buffer #402824
	4316357	POP-6	Sequencing applications on the 3100 platform with 50-cm capillary #4315930 and buffer #402824
Beckman Coulter	477480	eCAP ssDNA 100-R kit, including gel, caps, and standard	Oligonucleotides, RNA, and antisense DNA from 10 to 100 bases
	477407	eCAP dsDNA 1000 kit	Analysis of dsDNA fragments from 72 to 1000 bp
	477486	eCAP dsDNA 20,000 kit	Analysis of dsDNA fragments from 1,000 to 20,000 bp
	477409	EnhanceCE Dye	Intercalating dye for LIF applications

(up to 1600 psi) to load the capillary. Linear polyacrylamide, which contains no crosslinker, is a flowable polymer that can produce a very effective sieving matrix. The sieving properties are dependent upon both the polymer concentration and the average chain length (molecular weight). In general, lower polymer concentrations of higher average molecular weight are preferred for the separation of high molecular weight DNA, the reverse being true for lower molecular weight DNA. The selection of an optimal combination of these conditions must be balanced by viscosities that are compatible with the pumping capabilities of the CE instrumentation. It is essential to avoid purchasing incompatible components by matching the instrument capabilities to the matrix that has been selected for the given application.

With bare fused silica, either cellulose-based gels (e.g., hydroxyethylcellulose; Aldrich) or acrylamide-based gels (PE Biosystems) can be used. Replacing coated or gel-filled capillaries can be quite expensive; as premature failure of the capillary can normally be attributed to excessive voltage during separation or to inappropriate storage, it is worthwhile to take precautions for avoiding these problems. All manufacturers of CE instrumentation offer UV absorbance detection systems that will work for most general applications. Beckman Coulter and Bio-Rad offer laser-based fluorescence detectors for high-sensitivity CE applications. PE Biosystems offers a single-capillary laser fluorescence instrument for genetic analysis and DNA sequencing, while Beckman Coulter offers similar instrumentation in an eight-capillary format.

Preparation and Analysis of DNA

2.8.13

Critical Parameters

Oligonucleotide purity

The most common size of synthetic oligonucleotide probes is in the range of 20 to 30 bases. However, some applications require synthetic probes that are 100 bases or more in length. Single-base resolution over this range is essential in order to obtain an accurate assessment of purity. The salt content of the oligonucleotide should be kept to a minimum (<50 mM) in order to obtain optimal resolution. The oligonucleotide should be dissolved in water, or serially diluted in water from the stock TBE solution, prior to electrophoresis. Concentrations should be in the 1 µg/20 ml range for UV absorbance detection. The presence of alkaline-pH buffers and urea in the separation matrix and electrophoresis buffer are essential for single-stranded separations. It is important to run a standard—poly(A)₄₀₋₆₀—at the beginning of each set of analyses in order to confirm that the electrophoretic resolution is optimal (see Fig. 2.8.2).

DNA sequencing

In addition to the items above, this application requires single-base resolution over the range of 20 to 1000 bases. Automated, DNA sequencing instruments rigidly control the separation conditions to minimize temperature and power fluctuations and to eliminate variability in the matrix loading protocols. Consequently, most of the critical parameters are associated with template and primer quality and quantity issues. High-quality, salt-free plasmid, cosmid, or PCR-derived DNA is essential. The optimal molar ratio of template to primer is 1:1 and an imbalance of either component outside a ratio of 4:1 will give unusable sequences. Template concentrations should be in the 1 µg/20 µl range. The presence of proteins, salts, detergents, etc. in the template can inhibit the DNA polymerase and kill the reaction. It is essential to remove the excess fluorescent primers or dideoxynucleotides before separating the sequencing ladder on the CE. Further details on the essential parameters of the sequencing reaction itself can be found in Ausubel et al. (2003).

Fragment sizing

Selection of an appropriate separation matrix is probably the most important issue when analyzing microsatellite repeats. Resolution of the smaller fragments at the one to two nucleotide level is required. The DNA should be single-stranded and separated with denaturants. The use of internal size standards is es-

sential since the absolute mobilities can change from run to run and from sample to sample. When analyzing the VNTR fragments, which have larger repeating units (see Alternate Protocol), base pair resolution in the 3% to 6% range is required. These fragments are analyzed as double-stranded DNA without denaturants. Since the fragments are longer, lower matrix concentrations are preferred. Again, internal standards are essential.

The fragments are either obtained directly from genomic DNA, or PCR-amplified DNA. Considerable care has to be taken to remove particulate matter and salts from these samples prior to CE separation. The internal diameter of most capillaries is only 50 to 100 µm, and they are sensitive to plugging. Sometimes plugged capillaries can be salvaged by high-pressure back-flushing, but replacement is usually necessary. Salt concentrations should be <50 mM to minimize sample loading problems.

RFLP mutation screening

The same criteria for VNTR analysis are applied to mutation screening—the major exception being that the application requires pattern matching between control and mutant DNA, rather than absolute sizing. The double-stranded fragments must assume their sequence-dependent conformations. The addition of 10% glycerol to the flowable polymer significantly improves the discrimination of the electrophoretic separation.

Quantitation (RT-PCR)

The readout from this application is the absolute level of viral RNA in a biological sample. The PCR amplification is specific, but the target is present in a mixture of other nucleic acids, proteins and salts. The use of a fluorescent label or the addition of an intercalating dye, such as YO-PRO-1 or SYBR Green I, as well as the ability to dilute out the interfering substances, are essential features of this application. The fragments are separated as double-stranded PCR products and are quantitated by peak height on the electropherogram. The presence of salts in the starting material and in the PCR reaction can cause variable sample loading if electrokinetic sample injection is used. Sample loading by pressure injection bypasses these problems.

Absolute quantitation of the PCR product has to take into account both the length of the DNA and the degree of incorporation of the intercalating dye. This is achieved by coinjecting a standard of known concentration with the sample. If this standard contains a ladder

of fragments of known length then the quantity of the amplicon can be accurately determined. This procedure carries the added advantage of confirming the correct fragment size.

The difficulties encountered with this type of quantitation are not associated with the CE quantitation but with the PCR amplification step. It is essential that a dilution series is performed on the PCR mix to ensure that the reaction is linear with respect to cycle number. In the case of competitive PCR, it is essential that the competing template concentration approximates that of the target template. Alternatively, expression levels can be quantitated relative to a “housekeeping” gene. This is a gene, such as β -actin, that is constitutively expressed in the cells of interest. In this procedure the β -actin gene is coamplified with the gene of interest prior to quantitation by CE. As indicated previously, this field has recently undergone major expansion. The introduction of automated PCR-based instrumentation and the availability of high-quality fluorescently labeled oligonucleotides has produced a number of robust RT-PCR-based kits. They have superseded the intercalating dye approach for some applications.

Troubleshooting

Mechanical and electrical problems that might be encountered during instrument use are addressed in the troubleshooting sections of the manufacturers’ manuals. One difficulty is the loss of electrical continuity during a run. This loss of current can be caused if a small bubble forms in the column during injection or if the power generated by the run is so great that the solution boils or outgasses. Purging the column after a failed run removes any bubbles in the column. The vial containing the sample must have sufficient liquid to cover the end of the column during injection to prevent introduction of air into the column. On the autosampler, the vials can dry out during a long run, so sample temperature control or the use of the correct cap on the vial is required to slow evaporation. Reducing the buffer concentration or the run voltage can eliminate bubble formation during a run. Degassing the buffer is also useful if the outgassing problems continue.

On occasion capillaries can become plugged. This is normally due to the capillary matrix being allowed to dry out. When not in use, the capillary can remain installed on the CE instrument, but both ends must remain submerged in buffer. If it is removed from the instrument the capillary should be stored refrigerated, with both ends capped to prevent evapo-

ration. The capillary can also become plugged from insoluble material that is present in the sample. Sometimes it is possible to apply pressure to the anode outlet and blow out the plug. More likely, however, the capillary will need to be replaced.

Loss of resolution on a separation of standards can have several causes: e.g., the buffer, the separation matrix, or the capillary. A process of elimination should follow the course of first trying fresh buffer, then new matrix, and finally changing out the capillary. Where separation of standards is normal, but resolution or signal strength is poor, the cause is likely excess salts and/or buffers in the samples. Examining the behavior of the electrical current during the run will probably assist in troubleshooting. Most CE instruments simultaneously monitor both detector output and current flow throughout the electrophoresis run.

A drop in sensitivity or signal strength can be observed when multiple injections are made from a single standard vial (especially if the standard is dissolved in deionized water). This phenomena can be due to salt contamination of the sample by residual matrix buffer on the outside of the capillary (Guttman and Schwartz, 1995). Keep in mind that any source of additional ions to the sample (e.g., salt, buffer, or other ion species) will affect an electrokinetic injection. The ions of the sample will be effectively diluted by the ions from the contaminant. Sample preparation to remove variable salt contamination will improve the consistency of injections (Ruiz-Martinez et al., 1998).

Anticipated Results

The separation of oligonucleotides should show single-base resolution from 5 bases to > 100 bases. The synthetic efficiency of the assembly process is in excess of 99% (Pon et al., 1996), such that the yield of full-length product for a 100-mer should still exceed 30% of the total DNA present. For a 25-mer the yield of full-length product should exceed 80%. The addition of various functional moieties to the oligonucleotide will normally be seen as an increase in size (or decrease in mobility). The addition of a group as small as a phosphate group at the 5'-hydroxyl will make the oligonucleotide appear longer by a single base. The size increase, however, is insufficient proof of successful modification. This must be verified by obtaining a molecular weight by mass spectrometry (e.g., MALDI-TOF; UNIT 10.21).

The DNA-sequencing application (as stated previously) can have single-base resolution up

to 1000 bases or more, though performance today on commercial multicapillary instruments is ~600 bases or more in 2 hr. The smallest sequencing fragments will be detected as sharp, tight bands usually ~1 to 3 sec in width. The larger (slower-moving) fragments will show considerable band broadening due to diffusion as well as reduced separation efficiency of the matrix.

The microsatellite and VNTR analyses will show well-resolved peaks for the various multimers that are reflective of the number of repeats in the sample. These should be well resolved from the sizing control that has been added to, or coinjected with, the sample of interest.

The variability of these repeats at a given locus should give sufficient information to clearly identify relatedness between individuals. The expression levels of various isotypes should be sufficiently well discriminated to use this technique as a potential diagnostic test for disease states.

Quantitative PCR data should be sufficiently accurate to detect a 3- to 5-fold change in viral load level. The linearity of response of the standard and target DNAs should extend to the low to sub-nanogram level.

Time Considerations

These will vary dependent upon the type of separations that are required. Preparation of the stock solutions can take up to 1 hr. Filling the capillaries with fresh gel can take 1 hr. The analysis of synthetic oligonucleotides is normally completed within 30 min and sample preparation time is minimal. The DNA sequencing application probably requires the longest separation times. A 2-hr cycle time from injection to injection is normally sufficient to obtain 500 to 600 bases of high-quality sequence. Regeneration and loading times are <15 min. Preparing the stock solutions from the kit takes 24 hr. Filling the capillary with gel takes 30 min. Analysis of microsatellite fragments can take as little as 15 min, whereas the analysis of VNTR fragments takes up to 30 min. Regeneration times are <10 min. The analysis component of quantitative PCR is similar. However, this application requires multiple runs in order to obtain linearity curves for the PCR products and the calibration standards.

Literature Cited

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl,

K. (eds.) 2003. *Current Protocols in Molecular Biology*. John Wiley & Sons, Hoboken, N.J.

Butler, J., McCord, B., Jung, J., and Allen, R. 1994. Rapid analysis of short tandem repeat HUMTH01 by capillary electrophoresis. *BioTechniques* 10:1062-1068.

Cohen, A., Najarian, D., Paulus, A., Guttman, A., Smith, J., and Karger, B. 1988. Rapid separation and purification of oligonucleotides by high-performance, capillary, gel electrophoresis. *Proc. Natl. Acad. Sci. U.S.A.* 85:9660-9663.

Elbashir, E.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. 2001. Duplexes of 21-nucleotide RNAs mediated RNA interference in cultured mammalian cells. *Nature* 411:494-498.

Fasco, M., Treanor, C., Spivack, S., Figge, H., and Kaminsky, L. 1995. Quantitative RNA-polymerase chain reaction-DNA analysis by capillary electrophoresis and laser induced fluorescence. *Anal. Biochem.* 224:140-147.

Freeman, W.M., Walker, S.J., and Vrana, K.E. 1999. Quantitative RT-PCR: Pitfalls and potential. *BioTechniques* 26:112-125.

Grossman, P.D. 1991. Effect of molecular orientation and entangled polymer additives on the electrophoresis of biopolymers in free solution. Ph.D. Thesis, University of California, Berkeley.

Guttman, A. and Schwartz, H. 1995. Artifacts related to sample introduction in capillary gel electrophoresis affecting separation performance and quantitation. *Anal. Chem.* 67:2279-2283.

Jarcho, J. 1994. Restriction fragment length polymorphism analysis. In *Current Protocols in Human Genetics* (N.C. Dracopoli, J.L. Haines, B.R. Korf, D.T. Moir, C.C. Morton, C.E. Seidman, J.G. Seidman, and D.R. Smith, eds.) pp. 2.7.1-2.7.15. John Wiley & Sons, New York.

Piatak, M., Luk, K.-C., Williams, B., and Lifson, J.D. 1993. Quantitative competitive PCR for accurate quantitation of HIV DNA and RNA species. *BioTechniques* 14:70-81.

Pon, R., Buck, G., Hager, K., Naeve, C., Niece, R., Robertson, M., and Smith, A. 1996. Multifacility survey of oligonucleotide synthesis and an examination of the performance of unpurified primers in automated DNA sequencing. *BioTechniques* 21:680-685.

Ren, J., Ulvik, A., Ueland, P.U., and Refsum, H. 1997. Analysis of single-strand conformation polymorphism by capillary electrophoresis with laser-induced fluorescence detection using short-chain polyacrylamide as sieving medium. *Anal. Biochem.* 245:79-84.

Rossomando, E., White, L., and Ulfelder, K. 1991. Capillary electrophoresis: Separation and quantitation of reverse transcriptase polymerase chain reaction products from polio virus. *J. Chromatogr. B.* 656:159-168.

Ruiz-Martinez, M., Salas-Solano, O., Carilho, E., Kotler, L., and Karger, B. 1998. A sample purification method for ruffed and high-performance DNA sequencing by capillary electrophoresis

- using replaceable polymer solutions. A. Development of the cleanup protocol. *Anal. Chem.* 70:1516-1527.
- Sozer, A.C., Kelly, C.M., and Demers, D.B. 1998. Molecular analysis of paternity. *In Current Protocols in Human Genetics* (N.C. Dracopoli, J.L. Haines, B.R. Korf, D.T. Moir, C.C. Morton, C.E. Seidman, J.G. Seidman, and D.R. Smith, eds.) pp. 14.4.1-14.4.26. John Wiley & Sons, New York.
- Ulfelder, K., Schwartz, H., Hall, J., and Sunzeri, F. 1992. Restriction fragment length polymorphism analysis of *ERBB2* oncogene by capillary electrophoresis. *Anal. Biochem.* 200:260-267.
- Zhu, H., Clark, S., Benson, S., Rye, H., and Glazer, A. 1994. High sensitivity capillary electrophoresis of double stranded DNA fragments using monomeric and dimeric fluorescent intercalating dyes. *Anal. Chem.* 66:1941-1949.

Key Reference

- Ulfelder, K.J. and McCord, B. 1997. Chapter 11. Separation of DNA by capillary electrophoresis. *In Handbook of Capillary Electrophoresis*, 2nd ed. (J.P., Landers, ed.) pp. 347-378. CRC Press, Boca Raton, Fla.

An excellent reference on all aspects of capillary electrophoresis separations. Chapter 11, on DNA, goes into much greater depth than is possible here on the theory of separation, selection of buffers, and selection of gel matrices.

Contributed by Alan Smith
Stanford University
Stanford, California

Robert J. Nelson
Dakota Scientific
Sioux Falls, South Dakota

ANALYSIS OF DNA SEQUENCES BY BLOTTING AND HYBRIDIZATION

UNIT 2.9A

Southern Blotting

Southern blotting is the transfer of DNA fragments from an electrophoresis gel to a membrane support. The transfer or a subsequent treatment results in immobilization of the DNA fragments, so the membrane carries a semipermanent reproduction of the banding pattern of the gel. After immobilization, the DNA can be subjected to hybridization analysis (UNIT 2.10), enabling bands with sequence similarity to a labeled probe to be identified.

When setting up a Southern transfer, choices must be made between different types of membrane, transfer buffer, and transfer method. The most popular membranes are made of nitrocellulose, uncharged nylon, or positively charged nylon. Although these materials have different properties, the three types of membrane are virtually interchangeable for many applications. The main advantage of nylon membranes is that they are relatively robust and so can be reprobbed ten or more times before falling apart. Nitrocellulose membranes are fragile and can rarely be reprobbed more than three times; however, these are still extensively used, as they give lower backgrounds with some types of hybridization probe. The properties and advantages of the different membranes are discussed more fully in the commentary.

The basic protocol describes Southern blotting via upward capillary transfer of DNA from an agarose gel onto a nylon or nitrocellulose membrane, using a high-salt transfer buffer to promote binding of DNA to the membrane. With the high-salt buffer, the DNA becomes bound to the membrane during transfer but not permanently immobilized. Immobilization is achieved by UV irradiation (for nylon) or baking (for nitrocellulose). A support protocol describes how to calibrate a UV transilluminator for optimal UV irradiation of a nylon membrane.

The first alternate protocol details transfer using nylon membranes and an alkaline buffer, and is primarily used with positively charged nylon membranes. The advantage of this combination is that no post-transfer immobilization step is required, as the positively charged membrane binds DNA irreversibly under alkaline transfer conditions. The method can also be used with neutral nylon membranes but less DNA will be retained.

The second alternate protocol describes a transfer method based on a different transfer-stack setup. The traditional method of upward capillary transfer of DNA from gel to membrane described in the first basic and alternate protocols has certain disadvantages, notably the fact that the gel can become crushed by the weighted filter papers and paper towels that are laid on top of it. This slows down the blotting process and may reduce the amount of DNA that can be transferred. The downward capillary method described in the second alternate protocol is therefore more rapid than the basic protocol and can result in more complete transfer.

Although the ease and reliability of capillary transfer methods makes this far and away the most popular system for Southern blotting with agarose gels, it unfortunately does not work with polyacrylamide gels, whose smaller pore size impedes the transverse movement of the DNA molecules. The third alternate protocol describes an electroblotting procedure that is currently the most reliable method for transfer of DNA from a polyacrylamide gel.

Southern Blotting**2.9.1**

Contributed by Terry Brown

Current Protocols in Molecular Biology (1999) 2.9.1-2.9.20

Copyright © 2004 by John Wiley & Sons, Inc.

SOUTHERN BLOTTING ONTO A NYLON OR NITROCELLULOSE MEMBRANE WITH HIGH-SALT BUFFER

BASIC PROTOCOL

The procedure is specifically designed for blotting an agarose gel onto an uncharged or positively charged nylon membrane. With the minor modifications detailed in the annotations, the same protocol can also be used with nitrocellulose membranes.

The protocol is divided into three stages. First, the agarose gel is pretreated by soaking in a series of solutions that depurinate, denature, and neutralize the DNA and gel matrix. The second stage is the transfer itself, which occurs by upward capillary action. Finally, the DNA is immobilized on the membrane by UV irradiation (for nylon) or baking (for nitrocellulose).

Materials

- DNA samples to be analyzed
- 0.25 M HCl
- Denaturation solution: 1.5 M NaCl/0.5 M NaOH (store at room temperature)
- Neutralization solution: 1.5 M NaCl/0.5 M Tris·Cl, pH 7.0 (store at room temperature)
- 20× and 2× SSC (APPENDIX 2)
- Oblong sponge slightly larger than the gel being blotted
- Whatman 3MM filter paper sheets
- Nylon *or* nitrocellulose membrane (see Table 2.9.1 for suppliers)
- UV-transparent plastic wrap (e.g., Saran Wrap) for nylon membranes
- UV transilluminator (UNIT 2.5) or UV light box (e.g., Stratagene Stratalink) for nylon membranes
- Additional reagents and equipment for restriction endonuclease digestion (UNIT 3.1) and agarose gel electrophoresis (UNIT 2.5)

Table 2.9.1 Properties of Materials used for Immobilization of Nucleic Acids^a

	Nitrocellulose	Supported nitrocellulose	Uncharged nylon	Positively charged nylon	Activated papers
Application	ssDNA, RNA, protein	ssDNA, RNA, protein	ssDNA, dsDNA, DNA, protein	ssDNA, dsDNA, RNA, protein	ssDNA, RNA
Binding capacity (μg nucleic acid/cm ²)	80-100	80-100	400-600	400-600	2-40
Tensile strength	Poor	Good	Good	Good	Good
Mode of nucleic acid attachment ^b	Noncovalent	Noncovalent	Covalent	Covalent	Covalent
Lower size limit for efficient nucleic acid retention	500 nt	500 nt	50 nt or bp	50 nt or bp	5 nt
Suitability for reprobing	Poor (fragile)	Poor (loss of signal)	Good	Good	Good
Commercial examples	Schleicher & Schuell BA83, BA85; Amersham Hybond-C; PALL Biodyne A	Schleicher & Schuell BA-S; Amersham Hybond-C extra	Amersham Hybond-N; Stratagene Duralon-UV; Du Pont NEN GeneScreen	Schleicher & Schuell Nytran; Amersham Hybond-N ⁺ ; Bio-Rad ZetaProbe; PALL Biodyne B; Du Pont NEN GeneScreen Plus	Schleicher & Schuell APT papers

^aThis table is based on Brown (1991), with permission from BIOS Scientific Publishers Ltd.

^bAfter suitable immobilization procedure (see text).

Preparation and Analysis of DNA

2.9.2

CAUTION: Wear gloves from step 2 of the protocol onward to protect your hands from the acid and alkali solutions and to protect the membrane from contamination.

Prepare the gel

1. Digest the DNA samples with appropriate restriction enzyme(s), run in an agarose gel with appropriate DNA size markers, stain with ethidium bromide, and photograph with a ruler laid alongside the gel so that band positions can later be identified on the membrane.

The gel should contain the minimum agarose concentration needed to resolve bands in the area of interest (see commentary to UNIT 2.5A) and should be ≤ 7 mm thick, preferably less. The amount of DNA that must be loaded depends on the relative abundance of the target sequence that will subsequently be sought by hybridization probing (see recommendations in commentary to this unit and UNIT 2.10).

2. Rinse the gel in distilled water and place in a clean glass dish containing ~10 gel volumes of 0.25 M HCl. Shake slowly on a platform shaker for 30 min at room temperature.

This step results in partial depurination of the DNA fragments, which in turn leads to strand cleavage. The length reduction improves the transfer of longer molecules. The step is not necessary with PCR products < 4 kb in length, or if efficient transfer of molecules > 5 kb is not required. The time needed for the acid treatment depends on the concentration and thickness of the gel. To check that the treatment has been sufficient, watch the xylene cyanol and bromophenol blue dyes. When these change color to green and yellow respectively, the gel has equilibrated with the acid. Adequate depurination takes a further 10 min.

3. Pour off the HCl and rinse the gel with distilled water. Add ~10 vol denaturation solution and shake as before for 20 min. Replace with fresh denaturation solution and shake for a further 20 min.

Denaturation unzips the DNA to give single-stranded molecules that have unpaired bases and are suitable for subsequent hybridization analysis.

4. Pour off the denaturation solution and rinse the gel with distilled water. Add ~10 vol neutralization solution, shake as before for 20 min, then replace with fresh neutralization solution and shake for a further 20 min.

The aim of the neutralization step is to bring the gel pH down to < 9.0 . At higher pH, the transferred DNA will not bind to nitrocellulose. After neutralization, the gel pH can be checked using a pH meter. If the pH is > 9.0 , carry out a third washing in neutralization solution. Neutralization is less critical with a nylon membrane but should still be carried out.

Set up the transfer

5. Using Figure 2.9.1 as a guide, place an oblong sponge, slightly larger than the gel, in a glass or plastic dish (if necessary, use two or more sponges placed side by side). Fill the dish with enough 20 \times SSC to leave the soaked sponge about half-submerged in buffer.

The sponge forms the support for the gel. Any commercial sponge will do; before a sponge is used for the first time, it should be washed thoroughly with distilled water to remove any detergents that may be present. As an alternative, use a solid support with wicks made out of Whatman 3MM filter paper (see Fig. 2.9.1). An electrophoresis tank should not be used, as the high-salt transfer buffer will corrode the electrodes.

6. Cut three pieces of Whatman 3MM paper to the same size as the sponge. Place these on the sponge and wet them with 20 \times SSC.

7. Place the gel on the filter paper and squeeze out air bubbles by rolling a glass pipet over the surface.
8. Cut four strips of plastic wrap and place over the edges of the gel.

This is to prevent the buffer from “short-circuiting”—i.e., so that it flows through rather than around the gel.
9. Cut a piece of nylon membrane just large enough to cover the exposed surface of the gel. Pour distilled water ~0.5 cm deep in a glass dish and wet the membrane by placing it on the surface of the water. Allow the membrane to submerge, then leave for 5 min.

If a nitrocellulose membrane is being used, submerge in distilled water; replace the water with 20× SSC and leave for 10 min. Avoid handling nylon and nitrocellulose membranes even with gloved hands—use clean blunt-ended forceps instead.
10. Place the wetted membrane on the surface of the gel. Try to avoid getting air bubbles under the membrane; remove any that appear by carefully rolling a glass pipet over the surface.

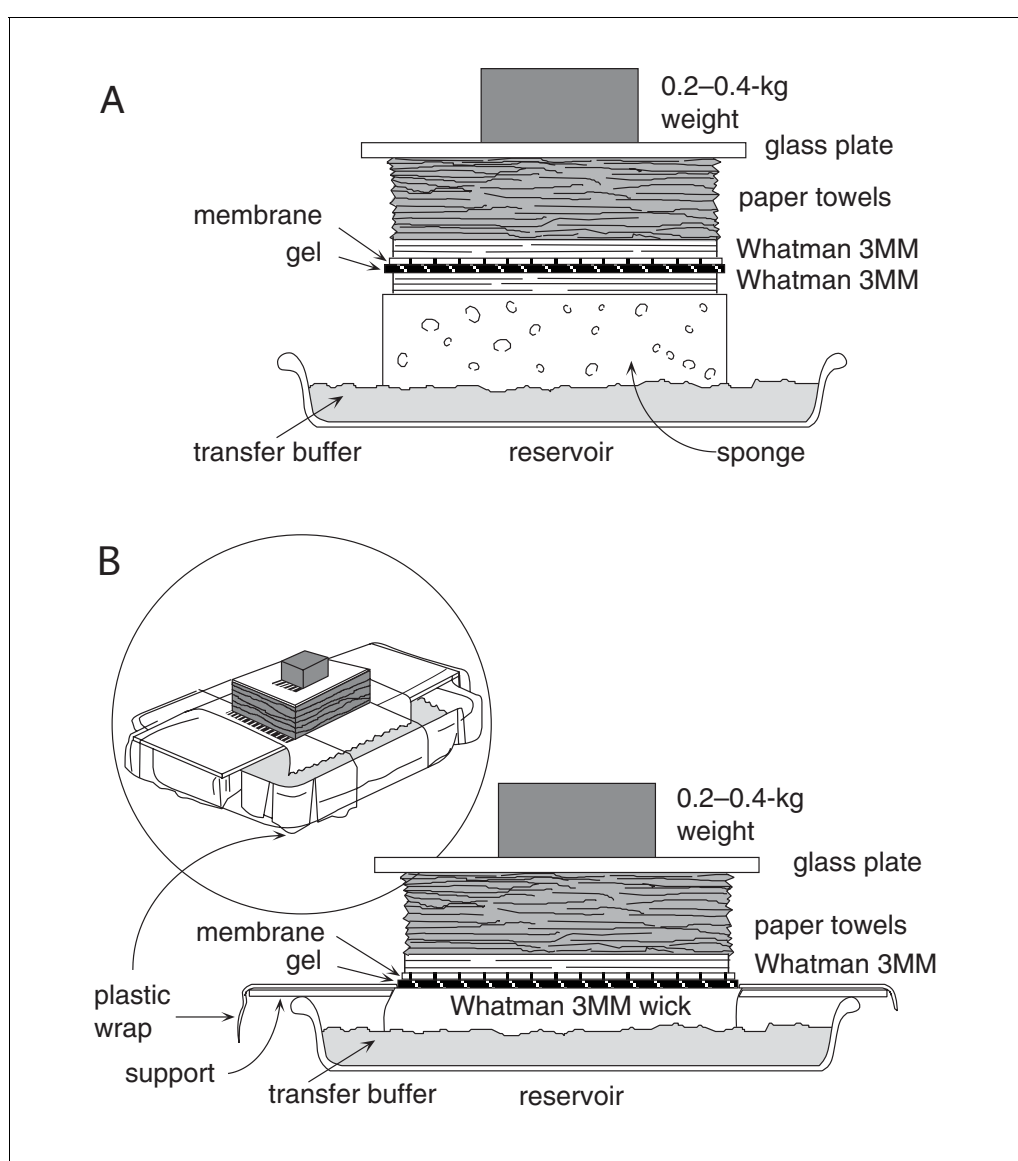


Figure 2.9.1 Two alternative transfer pyramid setups for Southern blotting via upward capillary transfer. (A) Sponge method. (B) Whatman 3MM filter paper wick method.

For blots of substantial amounts of plasmid or other very-low-complexity DNA, it is important to lay the membrane down precisely the first time, as detectable transfer can take place almost immediately, and moving the membrane may cause “mobile” bands as streaks on the autoradiograph.

11. Flood the surface of the membrane with 20× SSC. Cut five sheets of Whatman 3MM paper to the same size as the membrane and place these on top of the membrane.
12. Cut paper towels to the same size as the membrane and stack these on top of the Whatman 3MM papers to a height of ~4 cm.
13. Lay a glass plate on top of the structure and place a weight on top to hold everything in place. Leave overnight.

The weight should be sufficient to compress the paper towels to ensure good contact throughout the stack. Excessive weight, however, will crush the gel and retard transfer.

An overnight transfer is sufficient for most purposes. Extend the transfer time if the gel concentration is >1%, or transfer of fragments >20 kb is desired. Make sure that the reservoir of 20× SSC does not run dry during the transfer. If the gel contains large amounts of DNA (see annotation to step 10), high transfer efficiencies may not be required and shorter transfer times (~1 hr) may be used.

Disassemble the transfer pyramid

14. Remove the paper towels and filter papers and recover the membrane. Mark in pencil the position of the wells on the membrane and ensure that the up-down and back-front orientations are recognizable.

Pencil is preferable to pen, as ink marks may wash off the membrane during hybridization. An alternative for a nylon membrane only is to cut slits with a razor blade to mark the positions of the wells (do not do this before transfer or the buffer will short-circuit). The best way to record the orientation of the membrane is by making an asymmetric cut at one corner.

15. Rinse the membrane in 2× SSC, then place it on a sheet of Whatman 3MM paper and allow to dry.

The rinse has two purposes: to remove agarose fragments that may adhere to the membrane and to leach out excess salt. The gel can be retained with ethidium bromide to assess the efficiency of transfer.

Immobilize the DNA

16. Wrap the membrane UV-transparent plastic wrap, place DNA-side-down on a UV transilluminator (254-nm wavelength) and irradiate for the time determined from the support protocol.

CAUTION: Exposure to UV irradiation is harmful to the eyes and skin. Wear suitable eye protection and avoid exposure of bare skin.

Nitrocellulose membranes should not be UV irradiated. Instead, place between two sheets of Whatman 3MM paper and bake under vacuum for 2 hr at 80°C. This results in noncovalent attachment of the DNA to the membrane; the vacuum is needed to prevent the membrane from igniting.

UV crosslinking is recommended for a nylon membrane as this leads to covalent attachment and enables the membrane to be reprobed several times. The membrane must be completely dry before UV crosslinking; check the manufacturer’s recommendations. A common procedure is to bake for 30 min at 80°C prior to irradiation. Plastic wrap is used to protect the membrane during irradiation, but it must be UV-transparent. A UV light box can be used instead of a transilluminator (follow manufacturer’s instructions).

17. Store membranes dry between sheets of Whatman 3MM paper for several months at room temperature. For long-term storage, place membranes in a desiccator at room temperature or 4°C.

CALIBRATION OF A UV TRANSILLUMINATOR

When immobilizing DNA on a nylon membrane by UV crosslinking as described in the basic protocol, the intensity of irradiation is critical. Too little results in submaximal immobilization, and too much causes DNA degradation. UV transilluminators must therefore be calibrated to determine their output. The calibration should be repeated from time to time as the output of the lights changes with age. If a UV exposure meter (e.g., Baxter Scientific Blak-Ray) is available, it can be used to calculate the period of irradiation needed to produce the optimal exposure described in the manufacturer's support literature for the membrane (see also *UNIT 3.19*). Alternatively, the optimal exposure time can be determined experimentally as described in this protocol.

UV light boxes designed specifically for crosslinking are now commercially available. Most have automatic devices that vary output according to the age of the lights, but it is still a good idea to calibrate them regularly.

Additional Materials

DNA probe labeled to a specific activity of 10^8 dpm/ μ g

Additional reagents and equipment for DNA dot blotting (*UNIT 2.9B*) and hybridization analysis (*UNIT 2.10*)

CAUTION: Exposure to UV irradiation is harmful to the eyes and skin. Wear suitable eye protection and avoid exposure of bare skin.

1. Prepare five identical series of DNA dot blots (*UNIT 2.9B*) on nylon membrane strips, each with a range of DNA quantities from 1 to 100 pg.

The dot blots should be prepared on nylon membrane from the batch that is being used for Southern transfers.

2. Dry the nylon strips, wrap each one in UV-transparent plastic wrap, and place on the UV transilluminator, DNA-side-down.
3. Put on eye protection and cover exposed skin, then switch on the transilluminator. Remove individual strips after 30 sec, 45 sec, 1 min, 2 min, and 5 min.
4. Hybridize the strips with a suitable DNA probe labeled to a specific activity of 10^8 dpm/ μ g and prepare autoradiographs (*UNIT 2.10*).
5. Determine which nylon strip gives the most intense hybridization signals; the exposure time used for that strip is the optimal exposure time.

Note that the optimal exposure time depends on the dryness of the membrane. The test membrane strips should therefore be dried in exactly the same way as the Southern blots.

At no or low UV dose, the signal will be weak; with increasing dose, it will reach a plateau. At long exposures the signal will again decline, as the radiation will damage the DNA so that it hybridizes poorly.

**ALTERNATE
PROTOCOL**

**SOUTHERN BLOTTING ONTO A NYLON MEMBRANE WITH AN
ALKALINE BUFFER**

With a positively charged nylon membrane, the transferred DNA becomes covalently linked to the membrane if an alkaline transfer buffer is used. The protocol can also be carried out with uncharged nylon membranes using the minor modifications detailed in the annotations to the protocol steps below, but the transfer will be less complete. However, this transfer technique is not suitable for nitrocellulose membranes, as these do not retain DNA at pH >9.0 and fall apart after long exposure to alkali. The resistance of the nylon membrane to alkali should be checked before use, as some types of nylon membrane are less resistant than others.

Additional Materials

- 0.4 M (for charged membrane) *or* 0.25 M (for uncharged membrane) NaOH
- 0.25 M NaOH/1.5 M NaCl for uncharged membrane
- Positively charged *or* uncharged nylon membrane (see Table 2.9.1 for suppliers)

CAUTION: Wear gloves to protect your hands from the acid and alkali solutions and to protect the membrane from contamination.

1. Prepare a gel and treat with 0.25 M HCl as described in steps 1 and 2 of the basic protocol.
2. Rinse the gel with distilled water. Pour 10 gel volumes of 0.4 M NaOH into the dish and shake slowly on a platform shaker for 20 min.

This is the denaturation step. If an uncharged nylon membrane is being used, use 0.25 M NaOH in place of 0.4 M NaOH.

3. Follow steps 5 to 14 of the basic protocol to carry out the transfer, using 0.4 M NaOH as the transfer solution in place of 20× SSC. Leave to transfer ≥2 hr.

Alkaline transfer is quicker than high-salt transfer, so the blot can be taken apart any time after 2 hr.

A positively charged nylon membrane does not have to be prewetted, but can be placed directly onto the gel. If uncharged nylon is being used, use 0.25 M NaOH/1.5 M NaCl as the transfer solution. Check that the paper towels are resistant to the alkali solution; some types go brown and should not be used.

4. Remove the paper towels and filter paper and recover the membrane. Rinse the membrane in 2× SSC, place on a sheet of Whatman 3MM filter paper, and allow to air dry. Store as described in step 17 of the basic protocol.

The rinse has two purposes: to remove agarose fragments that may adhere to the membrane and to neutralize the membrane. Baking or UV crosslinking is not needed with a positively charge membrane; in fact, UV crosslinking is detrimental. With an uncharged membrane, immobilize as described in step 16, basic protocol.

**ALTERNATE
PROTOCOL**

SOUTHERN BLOTTING BY DOWNWARD CAPILLARY TRANSFER

One disadvantage with the upward capillary method (basic protocol) is that the gel can become crushed by the weighted filter papers and paper towels that are laid on top of it, reducing capillary action. The problem worsens as the transfer proceeds, as the paper towels become soaked in transfer buffer, increasing the pressure on the gel. This retards the blotting process and as a result the transfer must be carried out for 16 to 24 hr.

This alternate protocol describes a simple downward capillary transfer system that does not cause excessive pressure to be placed on the gel. Transfer from a 4-mm-thick 1% agarose gel is complete in just 1 hr. The protocol is suitable for all types of membrane and

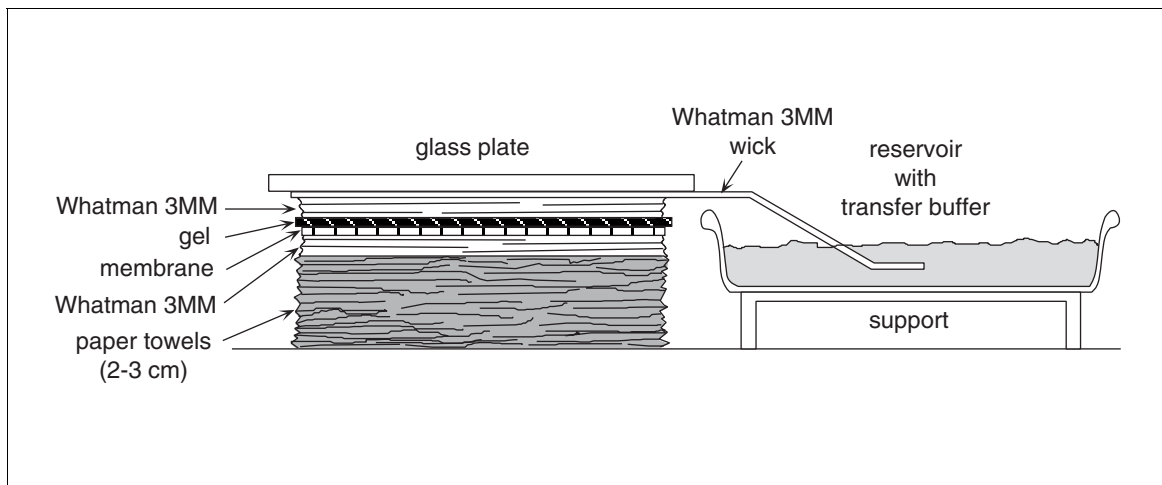


Figure 2.9.2 Transfer pyramid setup for Southern blotting via downward capillary transfer. Adapted with permission from Academic Press.

can be used with either high-salt or alkaline transfer buffers.

CAUTION: Wear gloves to protect your hands from the acid and alkali solutions used in gel preparation and transfer and to protect the membrane from contamination.

1. Prepare a gel as described in steps 1 to 4 of the basic protocol (high-salt transfer) or steps 1 to 2 of the first alternate protocol (alkaline transfer).
2. Make a stack of paper towels 2 to 3 cm high in a glass dish. The towels should be slightly wider than the gel.
3. Using Figure 2.9.2 as a guide, place four pieces of Whatman 3MM filter paper on top of the paper towels. Wet a fifth filter paper with transfer buffer and place on top.
4. Wet the membrane as described in step 9 of the basic protocol.

The membrane can be slightly larger than the gel.
5. Place the membrane on the top filter paper. Remove bubbles by rolling a glass pipet over the surface of the membrane.

Plastic wrap should be placed around the gel to prevent short-circuiting, as in step 8 of the basic protocol.
6. Place the gel on the membrane and remove bubbles.

No part of the gel should extend over the edge of the membrane.
7. Soak three pieces of Whatman 3MM paper cut to the same size as the gel with transfer buffer and place on top of the gel.
8. Place two larger pieces of Whatman 3MM paper together and soak in transfer buffer. As shown in Figure 2.9.2, place the glass dish containing transfer buffer on a support and use the two pieces of soaked Whatman 3MM paper to form a bridge between the gel and the reservoir.
9. Place a light plastic cover (e.g., a gel plate) over the top of the stack to reduce evaporation. Leave for 1 hr.
10. Remove the paper towels and filter papers and recover the membrane. Treat the membrane as described in steps 16 and 17 of the basic protocol (high-salt transfer) or step 4 of the first alternate protocol (alkaline transfer).

**ELECTROBLOTTING FROM A POLYACRYLAMIDE GEL
TO A NYLON MEMBRANE**

Capillary transfer, although reliable and needing no special equipment, has the disadvantage that it does not work with polyacrylamide gels, whose pore sizes are too small for effective transverse diffusion of DNA. Polyacrylamide gels must therefore be blotted by electrophoretic transfer. This requires a transfer buffer of low ionic strength, so nylon membranes are generally used, with uncharged nylon being slightly preferable to positively charged.

In conjunction with UV crosslinking (support protocol), this method transfers small DNA fragments and retains them quantitatively on the membrane; complete transfer and retention is crucial to the success of many procedures, including genomic sequencing and polymerase chain reaction (PCR) methods for quantitation of rare DNAs (Chapter 15). The polyacrylamide gel containing the samples of interest is placed in contact with an appropriately prepared membrane. These are then sandwiched into an electroblotting apparatus and the DNA is transferred out of the gel onto the membrane using electric current. After a buffer rinse, the membrane is ready to be UV-crosslinked.

There are several types of electroblotting apparatus on the market and the precise details of the procedure depend on the equipment being used. This protocol refers specifically to the Trans-Blot electroblotter (Bio-Rad). With a different piece of equipment, modify the procedure in accordance with the manufacturer's instructions.

Additional Materials

- 0.5× TBE electrophoresis buffer (*APPENDIX 2*)
- Scotch-Brite pads (supplied with Trans-Blot apparatus)
- Trans-Blot electroblotting cell (Bio-Rad) with cooling coil, or other electroblotting apparatus (*UNIT 10.8*)
- Additional reagents and equipment for nondenaturing (*UNIT 2.7*) or denaturing (*UNIT 2.12*) polyacrylamide gel electrophoresis and for electroblotting (*UNIT 10.8*)

CAUTION: Wear gloves to protect the membrane from contamination.

1. Run DNA samples in a nondenaturing or denaturing polyacrylamide gel.
2. When electrophoresis is almost complete, cut a piece of nylon membrane sufficient in size to cover the relevant parts of the gel. Pour distilled water ~0.5 cm deep in a glass dish. Wet the membrane by floating it on the surface of the water, then submerge it. Leave for 5 min.

The membrane need not be exactly the same size as the gel. Ensure that the membrane is large enough to cover all the DNA bands, but do not worry if not all parts of the gel are in contact with the membrane. Avoid handling the membrane even with gloved hands—use clean blunt-ended forceps instead.

3. When electrophoresis is complete, remove one glass plate from the electrophoresis apparatus and stain and photograph the gel (if nondenaturing). Cut a piece of Whatman 3MM filter paper slightly larger than the gel. Lay the filter paper on the surface of the gel and remove trapped air bubbles by rolling a glass pipet over the surface. The gel should adhere to the filter paper. Lift the gel off the glass plate by peeling the filter paper away.

This manipulation is designed to prevent breakage of a thin polyacrylamide gel. If the gel is sufficiently sturdy, it can instead be lifted directly onto a piece of Whatman 3MM paper that has been soaked in 0.5× TBE electrophoresis buffer.

4. Soak two Scotch-Brite pads in 0.5× TBE and remove air pockets by repeated

squeezing and agitation. Cut seven pieces of Whatman 3MM paper to the same size as the gel and soak these for 15 to 30 min in 0.5× TBE.

5. Place the opened gel holder of the Trans-Blot cell in a shallow tray, with the grey panel resting flat on the bottom, and place one of the saturated Scotch-Brite pads on the inner surface of the grey panel. Place three soaked filter papers on the pad. To ensure that there are no trapped air bubbles, build up the stack of filter papers one by one, carefully searching for trapped bubbles and removing them by rolling a glass pipet over the surface of the top paper.

Removing air bubbles is crucial to the success of electroblotting. It is also important that the filter papers be totally saturated with buffer. If heat build-up and bubble formation during transfer are a problem, the filters were probably not completely saturated.

6. Flood the filter paper carrying the gel with 0.5× TBE and place on top of the filter-paper stack. Flood the surface of the gel with 0.5× TBE and place the prewetted membrane onto the gel.

Again, it is important to remove air bubbles. Placing the stack on a light box makes trapped bubbles stand out more clearly.

7. Flood the surface of the membrane with 0.5× TBE and place the remaining four sheets of saturated Whatman 3MM paper on top, followed by the second saturated Scotch-Brite pad. Close the gel holder.

8. Half-fill the Trans-Blot cell with 0.5× TBE and place the gel holder in the cell with the grey panel facing towards the cathode. Fill the cell with 0.5× TBE and electroblot at 30 V (~125 mA) for 4 hr under the conditions recommended by the manufacturer.

Higher voltages can be used to reduce the transfer time (e.g., 40 V for 2 hr).

Some protocols suggest using precooled 0.5× TBE when filling the cell as well as performing the electroblotting in a cold room. This may not provide sufficient cooling to prevent heat build up and bubble formation, especially if a voltage >30 V is used. For more efficient transfer, the temperature must be kept at 4°C using the cooling coil. With other makes of equipment, it may be possible to control the temperature of the blot by circulating water from a cooling water bath.

9. Switch off the power and remove the gel holder. Take the assembly apart, marking the orientation of the membrane in pencil or by cutting a corner as in step 14 of the basic protocol.
10. If the gel was nondenaturing, denature the membrane by placing it for 10 min, DNA-side-up, on three pieces of Whatman 3MM paper soaked in 0.4 M NaOH.

This denaturation step is required for subsequent rehybridization. If the gel was run under denaturing conditions, this step is not required.

11. Rinse the membrane in 2× SSC, place on a sheet of Whatman 3MM paper, and allow to dry. Immobilize the DNA and store membranes as in steps 16 and 17 of the basic protocol.

COMMENTARY

Background Information

Southern blotting has been one of the cornerstones of DNA analysis since its first description by E.M. Southern (1975). Immobilization of DNA by binding to nitrocellulose, either powdered or in sheet form, had been utilized in biochemistry and molecular biology for several years (Hall and Spiegelman, 1961; Nygaard and Hall, 1963), but Southern was the first to show how immobilization of size-fractionated DNA fragments could be carried out in a reliable and efficient manner. The advent of Southern transfer and the associated hybridization techniques (UNIT 2.10) made it possible for the first time to obtain information about the physical organization of single and multicopy sequences in complex genomes. This expedited the first successful cloning experiments with eukaryotic genes and was directly responsible for breakthroughs such as the discovery of introns (Doel et al., 1977). The later application of Southern transfer and hybridization to the study of restriction fragment length polymorphisms (RFLPs) opened up new possibilities such as genetic fingerprinting (Jeffreys et al., 1985) and prenatal diagnosis of genetic disease (Davies, 1986).

The term Southern blotting, now used to describe any type of DNA transfer from gel to membrane, originally referred solely to capillary transfer onto nitrocellulose. The fragility of nitrocellulose membranes prompted the search for alternative types of support matrix, resulting in the introduction of nylon membranes in the early 1980s. In comparison with nitrocellulose, nylon membranes have less stringent requirements regarding the composition of the transfer buffer, and in recent years several novel buffer formulations have been devised. The most important innovation has been the use of alkaline transfer buffers with positively charged nylon membranes, resulting in immediate covalent attachment of the DNA to the membrane and thereby eliminating the need for the post-transfer immobilization steps required with nitrocellulose and uncharged nylon. Attempts have also been made to reduce the time needed for transfer by either changing the architecture of the transfer assembly (Lichtenstein et al., 1990; Chomczynski, 1992) or using noncapillary methods such as electroblotting, first described for protein transfers (Towbin et al., 1979; UNIT 10.8).

Critical Parameters

The amount of DNA that must be loaded depends on the relative abundance of the target sequence to which hybridization will take place (UNIT 2.10). The detection limit for a radioactive probe with a specific activity of 10^8 to 10^9 dpm/ μ g is about 0.5 pg DNA. Thus, for human genomic DNA, 10 μ g—equivalent to 1.5 pg of a single-copy gene 500 bp in length—is a reasonable minimum quantity to load. For more information about appropriate DNA quantities, see the commentary to UNIT 2.10.

Optimization of the parameters that influence Southern blotting must be carried out in conjunction with hybridization analysis, as the efficiency of transfer can be assessed only from the appearance of the autoradiograph obtained after the membrane is probed. It can therefore take a considerable amount of time to identify the precise conditions needed for optimal transfer, and there is often a temptation to make do with a suboptimal setup, so long as it works sufficiently well. This is understandable, but quality control of Southern transfers should not be neglected, as quite dramatic improvements can sometimes be achieved simply by using a more appropriate membrane or transfer system. The following paragraphs describe the most important factors that should be considered.

Choice of membrane

Nylon and nitrocellulose membranes have very different properties (Table 2.9.1). It is generally accepted that the main advantage of a nylon membrane is its greater tensile strength and the fact that the DNA can be bound covalently by UV crosslinking (Li et al., 1987) or, in the case of positively charged nylon, by transfer with an alkaline buffer (Reed and Mann, 1985). Nylon membranes can therefore be reprobbed up to about 12 times without becoming broken or losing their bound DNA. In contrast, nitrocellulose membranes are fragile and do not bind DNA covalently; baking to immobilize the DNA leads to a relatively weak hydrophobic attachment through exclusion of water. After about three rounds of hybridization, a nitrocellulose membrane is usually in pieces, with most of the DNA leached off (Haas et al., 1972). If reprobbing is important, then a nylon membrane (charged or uncharged) should be used.

The most clearly recognized disadvantage of nylon membranes (both charged and uncharged) is the amount of background signal

seen after hybridization. This can be a problem with any type of hybridization probe but is especially bad with some nonradioactive DNA probes (*UNITS 3.18 & 3.19*). There are several ways of reducing background hybridization signals (see *UNIT 2.10*, troubleshooting), but these may not be fully effective with a nylon membrane; if background is a problem, the only answer may be to change to nitrocellulose.

Nylon membranes are able to bind about five times more DNA per cm² than nitrocellulose (Table 2.9.1), but this is not a factor in Southern blotting, as the maximum binding capacity of the membrane is never approached. Size limits for efficient DNA retention are more important. Nylon retains DNA fragments down to 50 nucleotides in length, but nitrocellulose is inefficient with molecules <500 nucleotides. The only reason that PCR products shorter than 500 bp give reasonable hybridization signals after transfer to nitrocellulose is because there is normally a lot of DNA in the band to start with. Nitrocellulose is not a good choice for restriction-digested genomic DNA if the target band for hybridization probing is likely to be <500 bp.

In addition to nitrocellulose, uncharged nylon, and positively charged nylon, there are several other transfer matrices that are used less frequently (Table 2.9.1). Supported nitrocellulose membranes, an interesting innovation of the late 1980s, attempt to improve the tensile strength of nitrocellulose by supporting the matrix on a more rigid platform. In practice these membranes tend to combine the worst feature of nitrocellulose (loss of bound DNA) with elevated background caused by the support material. A second alternative to the “traditional” types of membrane are the activated cellulose papers, ABM- (Noyes and Stark, 1975) and APT-papers (Seed, 1982). The cellulose matrix of these papers carries aromatic groups that, after chemical activation, bind DNA covalently. These papers have relatively low binding capacities and for routine work offer no advantages over nylon (which they predated). They can, however, bind oligonucleotides down to just 2 nucleotides in length, considerably below the lower limit for nylon or nitrocellulose.

Many commercial suppliers market transfer membranes of one type or another. Quality is variable, especially with charged nylon membranes. The products listed in Table 2.9.1 reflect the author’s experiences and do not necessarily include all the reliable brands.

Transfer buffer

With nitrocellulose, the transfer buffer must provide a high ionic strength to promote binding of the DNA to the membrane (Southern, 1975; Nagamine et al., 1980). Several formulations exist, but 20× SSC is recommended because it is easy to make up and can be stored for several months at room temperature. The solution does not need to be filtered prior to use as a transfer buffer, but will need filtering if the same stock is to be used to prepare hybridization solutions (*UNIT 2.10*). Lower SSC concentrations (e.g., 10×) should not be used with nitrocellulose membranes, as the lower ionic strength may result in loss of smaller DNA fragments during transfer. In addition, alkaline transfer is not suitable for nitrocellulose membranes, as they do not retain DNA at pH 9.0 and fall apart after long exposure to alkali (Dyson, 1991).

Nylon membranes are able to bind DNA under a variety of conditions (acid, neutral, alkaline, high or low ionic strength), but a high-salt buffer such as 20× or 10× SSC appears to be beneficial (Khandjian, 1987). The addition of 2 mM *N*-laurylsarcosine (Sarkosyl) to the buffer may aid transfer to certain brands of nylon membrane (Chomczynski, 1992). Positively charged nylon membranes can be used with SSC buffers as described in the basic protocol, but this does not exploit their full potential. The ability of these membranes to bind DNA covalently after transfer in an alkaline buffer (0.4 M NaOH) is a major advantage, and the technique is well worth assessing if low efficiency is suspected with high-salt transfers. The only real problem with alkaline blotting is that high backgrounds may result if a chemiluminescent detection system (*UNIT 3.19*) is used. Alkaline transfer can also be carried out with uncharged nylon, but less DNA is retained than with charged nylon (Dyson, 1991).

Duration of transfer

The most difficult parameter to evaluate is duration of transfer. In a capillary system, rate of transfer depends on the size of the DNA, thickness of the gel, and agarose concentration. Upward capillary transfer is slow, as the architecture of the blot crushes the gel and retards diffusion of the DNA. With a high-salt buffer, it takes ~18 hr to obtain acceptable transfer of a 15-kb molecule from a 5-mm-thick 0.7% agarose gel; with the same gel 90% of the 1-kb molecules will be transferred in 2 hr. This problem is partially alleviated by the depurination step in the basic protocol, which breaks

larger molecules into fragments 1 to 2 kb in length, thereby reducing the time needed for their transfer. Even so, the blot should be left for at least 12 hr. If the gel is thicker than 5 mm or has an agarose concentration >1.0%, it cannot be assumed that the larger fragments will have transferred to a sufficient extent after 12 hr and a longer blotting period (up to 24 hr) may be necessary. The required blotting period can be assessed only by trial and error: if it is possible that the target band for the hybridization probe is >10 kb and a clear signal is not seen, the blot should be repeated using a longer transfer period.

Alkaline blots are more rapid, with most of the DNA being transferred during the first 2 hr. Therefore, if transfer of large bands is essential, alkaline transfer onto positively charged nylon should be considered as an alternative to a high-salt blot. Prolonging the alkaline transfer beyond 2 hr apparently does no harm (Chomczynski, 1992), and often these blots are left overnight.

More rapid transfers can be achieved with the downward capillary blot procedure developed by Chomczynski, described in the second alternate protocol. Variables in this system have been comprehensively evaluated (Chomczynski, 1992) and a downward blot is recommended if maximal transfer efficiency is required. If speed is more important than efficiency (e.g., if the gel being blotted contains a restriction-digested plasmid with few bands, each containing a relatively large amount of DNA) then any one of several quick-blot techniques can be used (e.g., Smith and Summers, 1980). These are not described in the protocols because in practice an overnight blot is usually just as convenient—everyone has to sleep sometime.

Transfer method

Capillary transfer is still the most popular method of Southern blotting. Its advantages are simplicity, reliability, and the lack of special equipment requirements. Alternative transfer methods have not yet achieved widespread use but are gaining prominence for certain applications.

Electroblotting, originally developed for protein transfers (see *UNIT 10.8*), does not work well with high-salt buffers and so is less appropriate for nitrocellulose membranes; it is worth attempting with nitrocellulose only if there is a really good reason for using this type of membrane (Smith et al., 1984). The technique has been applied to both nylon membranes and

activated papers (Bittner et al., 1980; Stellwag and Dahlberg, 1980), but problems often arise with overheating, which leads to bubble formation between the layers of the sandwich enclosing the gel, in turn leading to uneven transfer. Electroblotting is therefore only recommended for polyacrylamide gels, which cannot be blotted by capillary transfer.

The second alternative to capillary transfer is vacuum blotting, which was likewise first used with protein gels (Peferoen et al., 1982). The transfer buffer is drawn through the gel by vacuum pressure, enabling a 5-mm-thick 1% agarose gel to be blotted onto nitrocellulose (in 20× SSC) or nylon (in water) in as little as 30 min. The vacuum pressure must be controlled carefully to avoid compressing the gel and retarding transfer, but once mastered, the technique is very efficient. In a side-by-side comparison, a vacuum-blotted gel usually gives a significantly higher hybridization signal than a capillary blot. Reliable vacuum blotters are the StrataVac (Stratagene), VacuGene (Pharmacia Biotech), Model 785 (Bio-Rad), and TransVac (Hoefer). A protocol for their use is not provided here, as each device has its own special features; the individual manufacturer's instructions should be followed. Vacuum blotting should be seriously considered in any laboratory where Southern transfers are the rate-limiting step in the research program.

Immobilization technique

The aim of immobilization is to attach as much of the transferred DNA as possible to the membrane as tightly as possible. The standard procedures for the three common types of membranes, in increasing order of desirability, are as follows: for nitrocellulose, baking in a vacuum (leading to relatively weak noncovalent interactions), for uncharged nylon, UV irradiation (leading to covalent crosslinking), and for positively charged nylon, immobilization during alkaline transfer (giving covalent linkages).

Baking and alkaline transfer are straightforward procedures that cannot be improved on if carried out according to the instructions. UV crosslinking is more variable. Calibration of the UV source (see support protocol) is essential even if a light box rather than a transilluminator is used. Calibration should be repeated from time to time to check that the performance of the lights has not declined, and should also be carried out with each new batch of membrane, especially if the supplier has been

changed. Before irradiation, a nylon membrane must be thoroughly dried. Some manufacturers recommend air drying; others suggest a 30-min bake at 80°C. The manufacturer's recommendations should also be consulted if UV crosslinking is to be carried out after a nonalkaline transfer to a positively charged nylon membrane; with some brands, UV treatment can result in a loss of DNA.

Troubleshooting

Troubleshooting Southern blotting can be difficult and laborious, as transfer efficiency cannot be assessed accurately until the membrane has been subjected to hybridization analysis. The gel can be restained with ethidium bromide immediately after blotting in order to determine if a large quantity of DNA remains in the gel, which indicates a major problem with the transfer; however, the absence of DNA on the gel does not mean that all of it has become bound to the membrane. Bound DNA can sometimes be visualized by staining the membrane with methylene blue, but this is unreliable. Often a problem is only recognized when a less-than-perfect autoradiograph is obtained after probing, and even then it may be difficult to decide whether the transfer or hybridization technique is to blame. Discussion of how to troubleshoot Southern transfer is therefore provided in more detail in the commentary of *UNIT 2.10*.

Early warning signs of problems may be provided by the appearance of the membrane before and during transfer. Membranes should be pure white when removed from their wrapping and should remain white during the transfer and immobilization procedures. Membranes that are discolored, dusty, or dirty should not be used, and a membrane that is difficult to wet should be treated with suspicion. Yellow patches on a nitrocellulose membrane after transfer indicate inadequate neutralization. The only permissible color changes are a slight pinkness after transfer, from ethidium bromide staining of the DNA, or blue marks from transfer of the gel dye markers.

Anticipated Results

The Southern transfer procedures described in this unit should all yield a clear white membrane, possibly with pink and blue patches as described above. No data is generated until the membrane is subjected to autoradiography; anticipated results of autoradiography are discussed in the commentary to *UNIT 2.10*.

Time Considerations

The time required for Southern blotting is variable, as discussed in critical parameters. The basic protocol can be carried out in 24 hr as follows: restricting the DNA and running the agarose gel, 4 hr; preparing the gel, 2 hr; setting up the transfer, 30 min; transfer, overnight; taking the blot apart and drying the membrane, 30 min; immobilization, 2 hr for nitrocellulose, 5 min for nylon. More rapid transfer systems allow the procedure to be completed in a working day, mainly by reducing transfer time.

Literature Cited

- Bittner, M., Kupferer, P., and Morris, C.F. 1980. Electrophoretic transfer of proteins and nucleic acids from slab gels to diazobenzyloxymethyl cellulose or nitrocellulose sheets. *Anal. Biochem.* 102:459-471.
- Brown, T.A. (ed.) 1991. *Molecular Biology Labfax*. BIOS Scientific Publishers, Oxford.
- Chomczynski, P. 1992. One-hour downward alkaline capillary transfer for blotting of DNA and RNA. *Anal. Biochem.* 201:134-139.
- Davies, K.E. (ed.) 1986. *Human Genetic Diseases: A Practical Approach*. IRL Press at Oxford University Press, Oxford.
- Doel, M.T., Houghton, M., Cook, E.A., and Carey, N.H. 1977. The presence of ovalbumin mRNA coding sequences in multiple restriction fragments of chicken DNA. *Nucl. Acids Res.* 4:3701-3713.
- Dyson, N.J. 1991. Immobilization of nucleic acids and hybridization analysis. In *Essential Molecular Biology: A Practical Approach*, Vol. 2 (T.A. Brown, ed.) pp. 111-156. IRL Press at Oxford University Press, Oxford.
- Haas, M., Vogt, M., and Dulbecco, R. 1972. Loss of simian virus 40 DNA-RNA hybrids from nitrocellulose membrane: Implications for the study of virus-host interactions. *Proc. Natl. Acad. Sci. U.S.A.* 69:2160-2164.
- Hall, B.D. and Spiegelman, S. 1961. Sequence complementarity of T2-DNA and T2-specific RNA. *Proc. Natl. Acad. Sci. U.S.A.* 47:137-146.
- Jeffreys, A.J., Wilson, V., and Thein, S.L. 1985. Hypervariable 'minisatellite' regions in human DNA. *Nature* 314:67-73.
- Khandjian, E.W. 1987. Optimized hybridization of DNA blotted and fixed to nitrocellulose and nylon membranes. *Bio/Technology* 5:165-167.
- Li, J.K., Parker, B., and Kowalic, T. 1987. Rapid alkaline blot-transfer of viral dsRNAs. *Anal. Biochem.* 163:210-218.
- Lichtenstein, A.V., Moiseev, V., and Zaboikin, M.M. 1990. A procedure for DNA and RNA transfer to membrane filters avoiding weight-induced gel flattening. *Anal. Biochem.* 191:187-191.

- Nagamine, Y., Sentenac, A., and Fromageot, P. 1980. Selective blotting of restriction DNA fragments on nitrocellulose membranes at low salt concentrations. *Nucl. Acids Res.* 8:2453-2460.
- Noyes, B.E. and Stark, G.R. 1975. Nucleic acid hybridization using DNA covalently coupled to cellulose. *Cell* 5:301-310.
- Nygaard, A.P. and Hall, B.D. 1963. A method for the detection of RNA-DNA complexes. *Biochem. Biophys. Res. Commun.* 12:98-104.
- Peferoen, M., Huybrechts, R., and De Loof, A. 1982. Vacuum-blotting: A new simple and efficient transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to nitrocellulose. *FEBS Letts.* 145:369-372.
- Reed, K.C. and Mann, D.A. 1985. Rapid transfer of DNA from agarose gels to nylon membranes. *Nucl. Acids Res.* 13:7207-7221.
- Seed, B. 1982. Diazotizable arylamine cellulose papers for the coupling and hybridization of nucleic acids. *Nucl. Acids Res.* 10:1799-1810.
- Smith, G.E. and Summers, M.D. 1980. The bidirectional transfer of DNA and RNA to nitrocellulose or diazobenzoyloxymethyl paper. *Anal. Biochem.* 109:123-129.
- Smith, M.R., Devine, C.S., Cohn, S.M., and Lieberman, M.W. 1984. Quantitative electrophoretic transfer of DNA from polyacrylamide or agarose gels to nitrocellulose. *Anal. Biochem.* 137:120-124.
- Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
- Stellwag, E.J. and Dahlberg, A.E. 1980. Electrophoretic transfer of DNA, RNA and protein onto diazobenzoyloxymethyl (DBM)-paper. *Nucl. Acids Res.* 8:299-317.
- Towbin, H., Staehelin, T., and Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* 76:4350-4354.

Key Reference

Southern, 1975. See above.

First description of capillary transfer from gel to membrane.

Contributed by Terry Brown
University of Manchester Institute of
Science and Technology
Manchester, United Kingdom

UNIT 2.9B

Dot and Slot Blotting of DNA

Dot and slot blotting are simple techniques for immobilizing bulk unfractionated DNA on a nitrocellulose or nylon membrane. Hybridization analysis (UNIT 2.10) can then be carried out to determine the relative abundance of target sequences in the blotted DNA preparations. Dot and slot blots differ only in the geometry of the blot, a series of spots giving a hybridization pattern that is amenable to analysis by densitometric scanning.

Samples are usually applied to the membrane using a manifold attached to a suction device. The basic protocol describes such a procedure for dot or slot blotting on an uncharged nylon membrane; annotations to the steps detail the minor modifications that are needed if blotting onto nitrocellulose. The first alternate protocol describes the more major changes required for blotting with a positively charged nylon membrane. A second alternate protocol describes preparation of dot blots by spotting the samples onto the membrane by hand.

CAUTION: In all of the protocols, wear gloves to protect your hands from the alkali solution and to protect the membrane from contamination. Avoid handling nitrocellulose and nylon membranes even with gloved hands—use clean blunt-ended forceps instead.

- Nagamine, Y., Sentenac, A., and Fromageot, P. 1980. Selective blotting of restriction DNA fragments on nitrocellulose membranes at low salt concentrations. *Nucl. Acids Res.* 8:2453-2460.
- Noyes, B.E. and Stark, G.R. 1975. Nucleic acid hybridization using DNA covalently coupled to cellulose. *Cell* 5:301-310.
- Nygaard, A.P. and Hall, B.D. 1963. A method for the detection of RNA-DNA complexes. *Biochem. Biophys. Res. Commun.* 12:98-104.
- Peferoen, M., Huybrechts, R., and De Loof, A. 1982. Vacuum-blotting: A new simple and efficient transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to nitrocellulose. *FEBS Letts.* 145:369-372.
- Reed, K.C. and Mann, D.A. 1985. Rapid transfer of DNA from agarose gels to nylon membranes. *Nucl. Acids Res.* 13:7207-7221.
- Seed, B. 1982. Diazotizable arylamine cellulose papers for the coupling and hybridization of nucleic acids. *Nucl. Acids Res.* 10:1799-1810.
- Smith, G.E. and Summers, M.D. 1980. The bidirectional transfer of DNA and RNA to nitrocellulose or diazobenzoyloxymethyl paper. *Anal. Biochem.* 109:123-129.
- Smith, M.R., Devine, C.S., Cohn, S.M., and Lieberman, M.W. 1984. Quantitative electrophoretic transfer of DNA from polyacrylamide or agarose gels to nitrocellulose. *Anal. Biochem.* 137:120-124.
- Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
- Stellwag, E.J. and Dahlberg, A.E. 1980. Electrophoretic transfer of DNA, RNA and protein onto diazobenzoyloxymethyl (DBM)-paper. *Nucl. Acids Res.* 8:299-317.
- Towbin, H., Staehelin, T., and Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* 76:4350-4354.

Key Reference

Southern, 1975. See above.

First description of capillary transfer from gel to membrane.

Contributed by Terry Brown
University of Manchester Institute of
Science and Technology
Manchester, United Kingdom

UNIT 2.9B

Dot and Slot Blotting of DNA

Dot and slot blotting are simple techniques for immobilizing bulk unfractionated DNA on a nitrocellulose or nylon membrane. Hybridization analysis (UNIT 2.10) can then be carried out to determine the relative abundance of target sequences in the blotted DNA preparations. Dot and slot blots differ only in the geometry of the blot, a series of spots giving a hybridization pattern that is amenable to analysis by densitometric scanning.

Samples are usually applied to the membrane using a manifold attached to a suction device. The basic protocol describes such a procedure for dot or slot blotting on an uncharged nylon membrane; annotations to the steps detail the minor modifications that are needed if blotting onto nitrocellulose. The first alternate protocol describes the more major changes required for blotting with a positively charged nylon membrane. A second alternate protocol describes preparation of dot blots by spotting the samples onto the membrane by hand.

CAUTION: In all of the protocols, wear gloves to protect your hands from the alkali solution and to protect the membrane from contamination. Avoid handling nitrocellulose and nylon membranes even with gloved hands—use clean blunt-ended forceps instead.

DOT AND SLOT BLOTTING OF DNA ONTO UNCHARGED NYLON AND NITROCELLULOSE MEMBRANES USING A MANIFOLD

Dot and slot blots are usually prepared with the aid of a manifold and suction device. This is quicker and more reproducible than manual blotting and is the method of choice if a number of blots are to be prepared at any one time. Many commercial manifolds are available, most of them with interchangeable units that provide a choice of dot- and slot-blot geometries (Fig. 2.9.3).

In this protocol, the DNA to be transferred is heat-denatured and applied to the membrane in a salt buffer. After blotting, the membrane is treated with denaturation and neutralization solutions and the DNA immobilized by UV irradiation (for nylon) or baking (for nitrocellulose).

Materials

6× and 20× SSC (APPENDIX 2)

DNA samples to be analyzed

Denaturation solution: 1.5 M NaCl/0.5 M NaOH (store at room temperature)

Neutralization solution: 1 M NaCl/0.5 M Tris·Cl, pH 7.0 (store at room temperature)

Uncharged nylon *or* nitrocellulose membrane (see Table 2.9.1, UNIT 2.9A, for suppliers)

Whatman 3MM filter paper sheets

Dot/slot blotting manifold (e.g., Bio-Rad Bio-Dot SF *or* Schleicher and Schuell Minifold II)

UV-transparent plastic wrap (e.g., Saran Wrap)

UV transilluminator (UNIT 2.5A) for nylon membranes

1. Cut a piece of nylon membrane to the size of the manifold. Pour 6× SSC to a depth of ~0.5 cm in a glass dish; place the membrane on the surface and allow to submerge. Leave for 10 min.

A nitrocellulose membrane should be wetted in 20× rather than 6× SSC.

2. Cut a piece of Whatman 3MM filter paper to the size of the manifold. Wet in 6× SSC.

Use 20× SSC if transferring onto nitrocellulose.

3. Place the Whatman 3MM paper in the manifold and lay the membrane on top of it. Assemble the manifold according to the manufacturer's instructions, ensuring that there are no air leaks in the assembly.

4. To each DNA sample, add 20× SSC and water to give a final concentration of 6× SSC in a volume of 200 to 400 μl. Denature the DNA by placing in a water bath or oven for 10 min at 100°C, then place in ice.

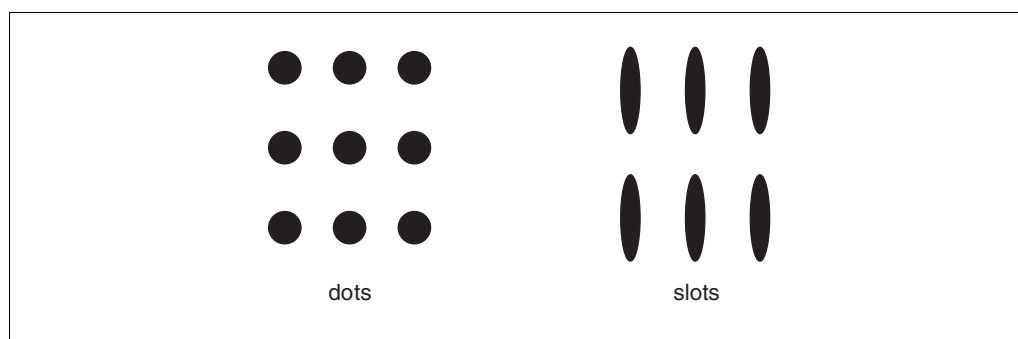


Figure 2.9.3 Dot (left) and slot (right) blot manifold architectures.

The amount of DNA that should be blotted will depend on the relative abundance of the target sequence that will subsequently be sought by hybridization probing (see commentaries to UNITS 2.9A & 2.10).

If using a nitrocellulose membrane, add an equal volume of 20× SSC to each sample after placing in ice.

5. Switch on the suction to the manifold device, apply 500 µl of 6× SSC to each well, and allow the SSC to filter through, leaving the suction on.

For a nitrocellulose membrane, use 20× SSC. The suction should be adjusted so that 500 µl of buffer takes ~5 min to pass through the membrane, as higher suction may damage the membrane.

Wells that are not being used can be blocked off by placing masking tape over them or by applying 500 µl of 3% (w/v) gelatin to each one (the former method is preferable, as gelatin may lead to a background signal after hybridization). Alternatively, keep all wells open and apply 6× or 20× SSC instead of sample to the extra wells.

6. Spin the DNA samples in a microcentrifuge for 5 sec. Apply to the wells being careful to avoid touching the membrane with the pipet. Allow the samples to filter through.

If any of the samples contain particulate material, blockage of the wells can be a major problem. If this occurs, add additional 6× SSC (20× SSC for nitrocellulose) and try to remove the blockage by resuspending the particles. The extra SSC must filter through the membrane, so often the blockage recurs. Increasing suction is not advisable, as it risks damaging the membrane. The best solution is to troubleshoot the DNA preparative method to avoid carryover of particles.

7. Dismantle the apparatus and place the membrane on a piece of Whatman 3MM paper soaked in denaturation solution. Leave for 10 min.
8. Transfer the membrane to a piece of Whatman 3MM paper soaked in neutralization solution. Leave for 5 min.
9. Place the membrane on a piece of dry Whatman 3MM paper and allow to dry.
10. Wrap the dry membrane in UV-transparent plastic wrap, place DNA-side-down on a UV transilluminator, and immobilize the DNA by irradiating for the appropriate time (determined as described in UNIT 2.9A support protocol).

CAUTION: Exposure to UV irradiation is harmful to the eyes and skin. Wear suitable eye protection and avoid exposure of bare skin.

UV irradiation causes DNA to become covalently bound to the nylon membrane. The membrane must be completely dry before UV crosslinking; check the manufacturer's recommendations. A common procedure is to bake for 30 min at 80°C prior to irradiation. Plastic wrap is used to protect the membrane during irradiation, but it must be UV transparent. A UV light box (e.g., Stratagene Stratalinker) can be used instead of a transilluminator (follow manufacturer's instructions).

Nitrocellulose membranes should not be UV irradiated. Instead, place between two sheets of Whatman 3 MM paper and bake under vacuum for 2 hr at 80°C.

11. Store the membrane dry between sheets of Whatman 3MM filter paper for several months at room temperature. For long-term storage, place the membrane in a desiccator at room temperature or 4°C.

DOT AND SLOT BLOTTING OF DNA ONTO A POSITIVELY CHARGED NYLON MEMBRANE USING A MANIFOLD

ALTERNATE
PROTOCOL

Positively charged nylon membranes bind DNA covalently at high pH (see UNIT 2.9A). Samples for dot or slot blotting can therefore be applied in an alkaline buffer, which promotes both denaturation of the DNA and binding to the membrane. The procedure is therefore quicker than blotting in salt buffer, as the post-blotting denaturation, neutralization, and immobilization steps are omitted.

Additional Materials

Positively charged nylon membrane (see Table 2.9.1, UNIT 2.9A, for suppliers)
0.4 M and 1 M NaOH (APPENDIX 2)
200 mM EDTA, pH 8.2 (APPENDIX 2)
2× SSC (APPENDIX 2)

1. Cut a piece of positively charged nylon membrane to the appropriate size. Pour distilled water to a depth of ~0.5 cm in a glass dish; place the membrane on the surface and allow to submerge. Leave for 10 min.
2. Prepare the blotting manifold as described in steps 2 and 3 of the basic protocol, using distilled water instead of SSC.
3. Add 1 M NaOH and 200 mM EDTA, pH 8.2, to each sample to give a final concentration of 0.4 M NaOH/10 mM EDTA. Heat for 10 min in a water bath or oven at 100°C. Microcentrifuge each tube for 5 sec.

The alkali/heat treatment denatures the DNA. The amount of DNA that should be blotted will depend on the relative abundance of the target sequence that will subsequently be sought by hybridization probing (see commentaries to UNITS 2.9A & 2.10).

4. Apply the samples to the membrane as described in steps 5 and 6 of the basic protocol, but prewash the membrane with 500 µl distilled water per well.
5. After applying the samples, rinse each well with 500 µl of 0.4 M NaOH and dismantle the manifold.
6. Rinse the membrane briefly in 2× SSC and air dry.
7. Store membrane as described in step 11 of the basic protocol.

MANUAL PREPARATION OF A DNA DOT BLOT

ALTERNATE
PROTOCOL

Dot blots can also be set up by hand simply by spotting small aliquots of each sample on to the membrane and waiting for the blot to dry. Repeated applications enable a sufficiently large volume of a dilute DNA sample to be blotted, but applying more than 30 µl is tedious and leads to untidy dots. In fact, manual application rarely produces results of publishable quality. It should be used only if no manifold is available. The protocol below details the procedure used for uncharged nylon membranes; annotations describe the changes needed to blot nitrocellulose and positively charged nylon membranes.

CAUTION: Wear gloves to protect your hands from the alkali solution and to protect the membrane from contamination. Avoid handling nitrocellulose and nylon membranes even with gloved hands—use clean blunt-ended forceps instead.

1. Cut a strip of uncharged nylon membrane to the desired size and mark out a grid of 0.5-cm × 0.5-cm squares with a blunt pencil. Pour 6× SSC to a depth of ~0.5 cm in a glass dish; place membrane on the surface and allow to submerge. Leave 10 min.

A nitrocellulose membrane should be wetted in 20× instead of 6× SSC, and a positively

Preparation and
Analysis of DNA

2.9.18

charged nylon membrane should be wetted in distilled water.

2. To each DNA sample, add $\frac{1}{2}$ vol of 20× SSC to give a final concentration of 6× SSC in the minimum possible volume. Denature the DNA by placing in a water bath or oven for 10 min at 100°C, then place in ice.

The amount of DNA that should be blotted will depend on the relative abundance of the target sequence that will subsequently be sought by hybridization probing (see commentaries to UNITS 2.9A & 2.10). The sample volume should be no more than 30 μ l and if possible much less. If necessary, reduce volume by ethanol precipitation (UNIT 2.1) before adding SSC.

If using positively charged nylon, add 1 M NaOH and 200 mM EDTA, pH 8.2, to each sample to give a final concentration of 0.4 M NaOH/10 mM EDTA, then heat as described.

If using a nitrocellulose membrane, add an equal volume of 20× SSC to each sample after placing on ice.

3. Place the wetted membrane over the top of an open plastic box so that the bulk of the membrane is freely suspended.
4. Spin each sample in a microcentrifuge for 5 sec, spot onto the membrane using a pipet, and allow to dry.

Do not touch the membrane with the pipet when applying the samples. Up to 2 μ l can be spotted in one application. If the sample volume is >2 μ l, it should be applied in successive 2- μ l aliquots, with each spot being allowed to dry before the next aliquot is applied on top. Drying can be aided with a hair dryer, but be careful that the blower does not spread the sample over the surface of the membrane. Try to keep the diameter of each dot to <4 mm.

5. For an uncharged nylon or nitrocellulose membrane, denature, neutralize, and immobilize as described in steps 7 to 10 of the basic protocol.

For a positively charged nylon membrane, rinse and dry as described in step 6 of the first alternate protocol.

6. Store membrane as described in step 11 of the basic protocol.

COMMENTARY

Background Information

Dot or slot blotting followed by hybridization analysis (UNIT 2.10) was first developed by Kafatos et al. (1979). The procedure is used to determine the relative abundance of a target sequence in a series of DNA samples. If a manifold is used, a large number of samples can be applied at once, enabling many different DNAs to be screened in a single hybridization experiment. The technique has found many applications over the years. For instance, in genome analysis, information on the genetic significance of a DNA sequence can often be obtained by using the sequence as a hybridization probe to dot blots of DNA prepared from related species. The rationale is that most genes have homologs in related organisms; for example, a coding sequence from the human genome will probably hybridize to related sequences in

dot blots prepared from DNA of various mammals. An intergenic or intronic region, which is less likely to have homologs in the other species, will probably not show widespread hybridization. This is the so-called “zoo blot” approach; blots containing DNA from a variety of related species are available ready-made from a number of suppliers.

Critical Parameters

The key requirement with dot blotting is that the DNA be fully denatured after transfer, or at least that all the samples be denatured to the same extent. The underlying assumption of dot blot analysis—that it can be used for meaningful comparisons of sequence abundance in different DNA samples—holds only if denaturation is precisely controlled. Variations in denaturation result in different samples having

different amounts of hybridizable DNA; if this occurs, the relative intensities displayed by two dots after hybridization will not be representative of the amount of target DNA that each contains.

The protocols for blotting uncharged nylon and nitrocellulose membranes attempt to ensure complete denaturation through the use of two denaturation steps—a heat denaturation before application to the membrane and an alkaline denaturation after application. Heat denaturation on its own is rarely adequate, as the DNA can renature fairly extensively before application to the membrane, even if plunged into ice on removal from the incubator. Blotting, whether manual or with a manifold, takes time, with some samples being blotted more quickly than others, so differential renaturation is a possibility. The second denaturation step, when the membrane is placed on a filter paper soaked in alkali, is intended to bring all the DNA back to an equal standing. Note that these problems do not arise with alkaline blotting onto positively charged nylon, as the high pH of the blotting solution maintains the DNA in a denatured state. Alkaline blotting is therefore the method of choice for DNA dot and slot blots where comparisons between different samples are to be made.

A second variable results from the purity of the DNA samples. With a Southern transfer, the gel electrophoresis step helps to fractionate away impurities, so the DNA that is transferred is relatively clean. Dot/slot blotting with bulk DNA lacks the benefit of a gel fractionation step, and the resulting co-blotted impurities can have unpredictable effects on hybridization, possibly reducing signal by blocking access to the hybridization sites, or increasing signal by trapping the probe. This must be taken into account if the signal intensity is to be used to estimate the absolute amount of target DNA, through comparison with a control dilution series. Copy number reconstruction by dot blot analysis is particularly suspect, as comparison between blots of cellular and plasmid DNA are reliable only if both types of DNA are scrupulously purified.

Troubleshooting

As with Southern blotting (*UNIT 2.9A*), most

problems with dot and slot blots become apparent only after hybridization analysis. The warning signs detailed in the commentary to *UNIT 2.9A* also hold for dot/slot blotting; other problems are described in *UNIT 2.10* (see Table 2.10.4 for troubleshooting).

Anticipated Results

The procedures yield a clear white membrane carrying applied DNA in amounts up to the carrying capacity of the matrix (Table 2.9.1, *UNIT 2.9A*). No data is generated until the membrane is subjected to autoradiography; anticipated results of autoradiography are discussed in *UNIT 2.10*.

Time Considerations

A manifold or manual blot can be set up and ready for sample application in as little as 15 min. After sample application, it takes about 3 hr to complete the protocol with a nitrocellulose membrane (most of this being the baking step), 60 min with an uncharged nylon membrane, and 30 min with a positively charged nylon membrane. The rate-determining step is sample application. Manifold application of clean samples (where no blockages occur) takes 5 min, but application by hand can take several hours if the sample volume is large and multiple additions have to be made.

Literature Cited

Kafatos, F.C., Jones, C.W., and Efstratiadis, A. 1979. Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. *Nucl. Acids Res.* 7:1541-1552.

Key Reference

Dyson, N.J. 1991. Immobilization of nucleic acids and hybridization analysis. *In* Essential Molecular Biology: A Practical Approach, Vol. 2 (T.A. Brown, ed.) pp. 111-156. IRL Press at Oxford University Press, Oxford.

Describes dot and slot blotting in some detail.

Contributed by Terry Brown
University of Manchester Institute of
Science and Technology
Manchester, United Kingdom

Hybridization Analysis of DNA Blots

The principle of hybridization analysis is that a single-stranded DNA or RNA molecule of defined sequence (the “probe”) can base-pair to a second DNA or RNA molecule that contains a complementary sequence (the “target”), with the stability of the hybrid depending on the extent of base pairing that occurs. Experimentally, the analysis is usually carried out with a probe that has been labeled and target DNA that has been immobilized on a membrane support. Hybridization analysis is sensitive and permits detection of single-copy genes in complex genomes. The technique has widespread applications in molecular biology.

The first stage in a hybridization experiment is to immobilize the denatured nucleic acids on a suitable solid support. Methods for achieving this with gel-fractionated and bulk DNA are described in *UNITS 2.9A & 2.9B*. The labeled probe is then applied in a solution that promotes hybridization. After a suitable incubation, the membrane is washed so that any nonspecifically bound probe is removed, leaving only probe that is base-paired to the target DNA. By controlling the stringency of the washing conditions, decisions can be made about whether to target sequences that are 100% complementary to the probe, or allow some mismatching so that sequences with lower degrees of similarity are also detected. The latter approach (heterologous probing) is used to study related sequences in a single or more than one genome.

Hybridization analysis was originally carried out with long (100 to 1000 bp), radioactively labeled DNA probes. Other types of probe (RNA, oligonucleotide) have more recently been introduced, as have nonradioactive labeling and detection strategies. In addition, improvements in understanding of the factors that influence hybrid stability and hybridization rate have led to a proliferation of reagents and protocols for hybridization analysis. Finding one’s way through the maze can be a daunting task, especially as protocols that work well with one probe-target combination may not work so well if either member of the partnership is changed. The approach taken here is to present as the basic protocol an unsophisticated procedure for hybridization analysis with a radiolabeled DNA probe. Despite its lack of embellishments, the protocol gives acceptable results with Southern and dot blots on nitrocellulose and nylon (uncharged and charged) membranes. The alternate protocol describes a similar method for probing DNA blots with a radiolabeled RNA probe. A support protocol for stripping blots to ready them for reprobing is also provided.

Relevant units elsewhere in the manual include the following: *UNITS 3.18 & 3.19* describe the preparation of nonradioactive probes and their use in hybridization analysis; *UNIT 4.9* covers hybridization analysis of immobilized RNA; *UNIT 6.3* describes hybridization analysis of recombinant clone libraries; and *UNIT 6.4* explains how to use labeled oligonucleotides as hybridization probes.

These hybridization protocols should not be read in isolation. The commentary describes various modifications that can be introduced, including changes to prehybridization, hybridization, and wash solution formulations, and alterations to incubation times and conditions, the latter including a discussion of the wash conditions compatible with different degrees of stringency. The intention is provide the reader with sufficient data to make well-informed decisions about how to modify the basic and alternate protocols for specific applications.

CAUTION: Investigators should wear gloves for all procedures involving radioactivity and should be careful not to contaminate themselves and their clothing. When working with ^{32}P , investigators should frequently check themselves and the working area for

radioactivity using a hand-held monitor. Any radioactive contamination should be cleaned up using appropriate procedures. Radioactive waste should be placed in appropriately designated areas for disposal. Follow the guidelines provided by your local radiation safety adviser.

HYBRIDIZATION ANALYSIS OF A DNA BLOT WITH A RADIOLABELED DNA PROBE

BASIC PROTOCOL

This protocol is suitable for hybridization analysis of Southern transfers (*UNIT 2.9A*) and dot and slot blots (*UNIT 2.9B*) with a radioactively labeled DNA probe 100 to 1000 bp in length. The steps employ nylon membranes (uncharged or positively charged) but are suitable for nitrocellulose if modified as described in the annotations. The commentary describes how to tailor the protocol for individual requirements.

A hybridization experiment can be divided into three stages. First, the membrane is incubated in a prehybridization solution containing reagents that block nonspecific DNA binding sites on its surface, thereby reducing background hybridization. In this protocol, the blocking agents are Denhardt solution and denatured salmon sperm DNA; alternatives are described in the commentary. In the second stage, the prehybridization solution is replaced by fresh buffer containing the labeled probe, and an overnight incubation is carried out to allow the probe to bind to target sequences in the immobilized DNA. During this hybridization step, the probe pairs not only with target sites that have 100% complementarity with the probe, but also with related sequences. In the final stage of the experiment the membrane is washed with a series of solutions that gradually remove bound probe molecules until only highly matched hybrids remain.

Materials

DNA to be used as probe

Aqueous prehybridization/hybridization (APH) solution, room temperature and 68°C

2× SSC/0.1% (w/v) SDS

0.2× SSC/0.1% (w/v) SDS, room temperature and 42°C

0.1× SSC/0.1% (w/v) SDS, 68°C

2× and 6× SSC (*APPENDIX 2*)

Hybridization oven (e.g., Hybridiser HB-1, Techne) *or* 68°C water bath *or* incubator

Hybridization tube *or* sealable bag and heat sealer

Additional reagents and equipment for DNA labeling by nick translation or random oligonucleotide priming (*UNIT 3.5*), measuring the specific activity of labeled DNA and separating unincorporated nucleotides from labeled DNA (*UNIT 3.4*), and autoradiography (*APPENDIX 3*)

1. Label the probe DNA to a specific activity of $>1 \times 10^8$ dpm/μg by nick translation or random oligonucleotide priming. Measure the specific activity and remove unincorporated nucleotides.

The probe should be a double-stranded DNA fragment, ideally 100 to 1000 bp in length. Usually the probe DNA is obtained as a cloned fragment (Chapter 1) which is purified from the vector by restriction digestion (UNIT 3.1) followed by recovery from an agarose gel (UNIT 2.6).

2. Wet the membrane carrying the immobilized DNA in 6× SSC.

The membrane is blotted as described in UNIT 2.9A. Do not handle the membrane: use clean blunt-ended forceps.

Preparation and Analysis of DNA

2.10.2

3. Place the membrane, DNA-side-up, in a hybridization tube and add ~1 ml APH solution per 10 cm² of membrane.

Prehybridization and hybridization are usually carried out in glass tubes in a commercial hybridization oven. Alternatively, a heat-sealable polyethylene bag can be used. The membrane should be placed in the bag, all edges sealed using a heat sealer, and a corner cut off. The APH solution is then pipetted into the bag through the cut corner and resealed.

4. Place the tube in the hybridization oven and incubate 3 hr with rotation at 68°C.

If using a bag, shake slowly in a suitable incubator or water bath.

If using a nylon membrane, reduce the prehybridization period to 15 min, but warm the prehybridization/hybridization solution to 68°C before adding to the membrane.

5. Denature the probe DNA by heating for 10 min in a water bath or incubator at 100°C. Place in ice.

Step 5 should be done immediately before step 6, with a minimum delay between them.

6. Pour the APH solution from the hybridization tube and replace with an equal volume of prewarmed (68°C) APH solution. Add denatured probe and incubate with rotation overnight at 68°C.

The probe concentration in the hybridization solution should be 10 ng/ml if the specific activity is 10⁸ dpm/μg, or 2 ng/ml if the specific activity is 1 × 10⁹ dpm/μg. If using a bag, cut off a corner, pour out the prehybridization solution, add the hybridization solution plus probe, and reseal. It is very difficult to avoid contaminating the bag sealer with radioactivity; furthermore, the sealing element (the part that gets contaminated) is often difficult to clean. Hybridization bags are therefore not recommended.

7. Pour out the APH solution, using the appropriate disposal method for radioactive waste, and add an equal volume of 2× SSC/0.1% SDS. Incubate with rotation for 10 min at room temperature, changing the wash solution after 5 min.

CAUTION: *All wash solutions must be treated as radioactive waste and disposed of appropriately.*

To reduce background, it may be beneficial to increase the volume of the wash solutions by 100%. If using a bag, transfer the membrane to a plastic box for the washes.

8. Replace the wash solution with an equal volume of 0.2× SSC/0.1% SDS and incubate with rotation 10 min at room temperature, changing the wash solution after 5 min (this is a low-stringency wash; see commentary).

9. If desired, carry out two further washes as described in step 8 using prewarmed (42°C) 0.2× SSC/0.1% SDS for 15 min each at 42°C (moderate-stringency wash).

10. If desired, carry out two further washes using prewarmed (68°C) 0.1× SSC/0.1% SDS for 15 min each at 68°C (high-stringency wash).

11. Pour off the final wash solution, rinse the membrane in 2× SSC at room temperature, and blot excess liquid. Wrap in plastic wrap.

Do not allow the membrane to dry out if it is to be reprobbed.

12. Set up an autoradiograph (APPENDIX 3).

HYBRIDIZATION ANALYSIS OF A DNA BLOT WITH A RADIOLABELED RNA PROBE

ALTERNATE PROTOCOL

Purified RNA polymerases from bacteriophages such as SP6, T3, and T7 (*UNIT 3.8*) are very efficient at synthesizing RNA *in vitro* from DNA sequences cloned downstream of the appropriate phage promoter (Little and Jackson, 1987). Several micrograms of RNA can be obtained from 1 μg of DNA template in a 10-min reaction. If a radiolabeled ribonucleotide is added to the reaction mixture, the polymerase synthesizes uniformly labeled RNA with specific activities up to and beyond 10^9 dpm/ μg . The fact that RNA probes are single-stranded gives them certain advantages over DNA probes. RNA probes do not need to be denatured before being added to the hybridization solution, and remain fully available for hybridization during the course of the incubation. In contrast, the “potency” of a double-stranded DNA probe gradually declines as the complementary strands of the probe reanneal during the hybridization reaction, reducing the amount of probe available to hybridize to the target.

The hybridization protocol for an RNA probe is not greatly different from that with labeled DNA. Formamide is usually included in the prehybridization/hybridization solutions, because in the presence of formamide an RNA-DNA hybrid is more stable than the equivalent DNA-DNA hybrid. Carrying out the hybridization with formamide also permits the incubation to be at a lower temperature without loss of stringency. Single-strand-specific RNases are added to one or more of the wash solutions to remove nonhybridized probe molecules, considerably reducing background.

The protocol includes details of probe preparation (Mundy et al., 1991). The hybridization procedure is suitable for both nitrocellulose and nylon membranes, though backgrounds may be higher with nylon.

Additional Materials

- TE buffer, pH 8.0 (*APPENDIX 2*)
 - Labeling buffer
 - Nucleotide mix
 - 200 mM dithiothreitol (DTT), freshly prepared
 - 20 U/ μl human placental ribonuclease inhibitor
 - [α - ^{32}P]UTP: 20 mCi/ml (800 Ci/mmol) or 10 mCi/ml (400 Ci/mmol)
 - SP6 or T7 RNA polymerase (*UNIT 3.8*)
 - RNase-free DNase I (*UNIT 3.12*)
 - 0.25 M EDTA, pH 8.0 (*APPENDIX 2*)
 - Formamide prehybridization/hybridization (FPH) solution
 - 2 \times SSC (*APPENDIX 2*) containing 25 $\mu\text{g}/\text{ml}$ RNase A + 10 U/ml RNase T1 (*UNIT 3.13*)
- Additional reagents and equipment for cloning and purifying plasmid DNA (Chapter 1), phenol extraction and ethanol precipitation (*UNIT 2.1*), restriction digestion of DNA (*UNIT 3.1*), measuring the specific activity of and separating unincorporated nucleotides from labeled RNA (*UNIT 3.4*), and autoradiography (*APPENDIX 3*)

Prepare the RNA probe

1. Clone into a suitable vector (Table 2.10.1) the DNA fragment that will be transcribed into the RNA probe.

DNA must be of high purity, so use a method that includes a CsCl/ethidium bromide equilibrium centrifugation step (UNIT 1.7).

2. Linearize the DNA by restriction digestion immediately downstream of the cloned fragment.

Preparation and Analysis of DNA

2.10.4

Linearization introduces an endpoint for RNA synthesis. This ensures that enzymes and substrates are not wasted by transcribing downstream vector DNA, and also increases the specificity of the probe by excluding unwanted sequences.

3. Purify the DNA from the restriction enzyme by phenol extraction and ethanol precipitation. Resuspend in TE buffer, pH 8.0, at a concentration of 1 mg/ml.

4. Mix the following at room temperature:

- 4 μ l labeling buffer
- 1.5 μ l nucleotide mix
- 1 μ l 200 mM DTT
- 1 μ l (20 U) human placental ribonuclease inhibitor
- 2 μ g purified plasmid DNA from step 3
- 100 to 200 μ Ci [α -³²P]UTP
- H₂O to a final volume of 20 μ l.

These are mixed at room temperature, as the spermidine in the labeling buffer may precipitate on ice.

5. Add 5 U SP6 or T7 RNA polymerase. Incubate for 1 hr at 40°C for SP6 or at 37°C for T7.

6. Add 2 U RNase-free DNase I and incubate at 10 min 37°C. Stop the reaction by adding 2 μ l of 0.25 M EDTA, pH 8.0.

DNase treatment degrades the template. This step may not be necessary for probes to be used in hybridization analysis, but is worth doing to be on the safe side.

7. Measure the specific activity of the RNA by acid precipitation and remove unincorporated nucleotides by the spin-column procedure.

The specific activity should be at least 7×10^8 dpm/ μ g, preferably $>10^9$ dpm/ μ g.

The labeled probe can be stored at -20°C for 2 days before use.

Carry out hybridization analysis

8. Carry out the prehybridization incubation as described in steps 2 to 4 of the basic protocol, but use FPH solution and incubate at 42°C.

9. Replace the FPH solution with an equal volume of fresh prewarmed solution. Add the labeled probe and incubate overnight with rotation at 42°C.

The probe concentration in the hybridization solution should be 1 to 5 ng/ml. Hybridizations in formamide solutions are carried out at lower temperatures than aqueous hybridizations. However, if background hybridization is a problem, raise the incubation temperature to 65°C.

10. Wash the membrane as described in steps 7 to 8 of the basic protocol.

11. Replace the wash solution with an equal volume of 2 \times SSC containing 25 μ g/ml RNase A + 10 U/ml RNase T1; incubate with rotation for 30 min at room temperature.

The RNase wash decreases background hybridization.

12. Carry out moderate- and high-stringency washes as desired, rinse the membrane in 2 \times SSC, and set up autoradiograph as in steps 9 to 12 of the basic protocol.

Table 2.10.1 Selection of Cloning Vectors Incorporating Promoters for Bacteriophage RNA Polymerases

Vector	Size (bp)	Markers ^a	Promoters
pBluescript	2950	amp, <i>lacZ'</i>	T3, T7
pGEM series	2746-3223	amp, <i>lacZ'</i>	SP6, T7
pGEMEX-1	4200	amp	SP6, T3, T7
pSELECT-1	3422	tet, <i>lacZ'</i>	SP6, T7
pSP18, 19, 64, 65	2999-3010	amp	SP6
pSP70, 71, 72, 73	2417-2464	amp	SP6, T7
pSPORT1	4109	amp, <i>lacZ'</i>	SP6, T7
pT3/T7 series	2700, 2950	amp, <i>lacZ'</i>	T3, T7
pWE15	8800	amp, neo	T3, T7
pWE16	8800	amp, dhfr	T3, T7

^aAbbreviations: amp, ampicillin resistance; dhfr, dihydrofolate reductase; *lacZ'*, β -galactosidase α -peptide; neo, neomycin phosphotransferase (kanamycin resistance); tet, tetracycline resistance.

REMOVAL OF PROBES FROM HYBRIDIZED MEMBRANES

If the DNA has been immobilized on the membrane by UV crosslinking (for uncharged nylon membranes) or by alkaline transfer (for positively charged nylon), the matrix-target DNA interaction (which is covalent in nature) is much stronger than the target-probe interaction (which results from hydrogen bonding). It is therefore possible to remove (or “strip off”) the hybridized probe DNA without removing the membrane-bound target DNA. Nylon membranes can therefore be reused several times—a dozen reprobings are routinely possible. Hybridized probe DNA can also be stripped from nitrocellulose membranes, but the weakness of the hydrophobic interactions that bind the target DNA to the matrix, plus the fragility of nitrocellulose, limits the lifetime of these membranes to three reprobings at most.

This protocol describes three methods for probe stripping, in order of increasing harshness. The treatment that is needed depends on how tightly the probe anneals to the target, which in turn depends on the GC content of the probe and the number of base pairs that are formed. To start with, use the least harsh treatment, monitoring the results by autoradiography of the stripped membrane. If a hybridization signal can still be seen, move up to a harsher method. To strip probe from a northern membrane (with immobilized RNA), use the mild treatment only. Do not add NaOH to RNA; it will be destroyed.

Additional Materials

Mild stripping solution
Moderate stripping solution
0.4 M NaOH
0.1% (w/v) SDS, 100°C

CAUTION: Although the stripping solutions may not become highly radioactive, they should still be disposed of as radioactive waste.

- 1a. *Mild treatment:* Wash the membrane in several hundred milliliters of mild stripping solution for 2 hr at 65°C.
- 1b. *Moderate treatment:* Wash the membrane in 0.4 M NaOH for 30 min at 45°C. Then rinse twice in several hundred milliliters of moderate stripping solution for 10 min at room temperature.

SUPPORT PROTOCOL

Preparation and Analysis of DNA

2.10.6

- 1c. *Harsh treatment:* Pour several hundred milliliters of boiling 0.1% SDS onto the membrane. Cool to room temperature.

If a membrane is to be reprobbed, it must not be allowed to dry out between hybridization and stripping. If it becomes dry, the probe may bind to the matrix.

2. Place membrane on a sheet of dry Whatman 3MM filter paper and blot excess liquid with a second sheet. Wrap the membrane in plastic wrap and set up an autoradiograph.

If signal is still seen after autoradiography, rewash using harsher conditions.

3. The membrane can now be rehybridized. Alternatively, it can be dried and stored for later use.

Membranes can be stored dry between Whatman 3MM paper for several months at room temperature. For long-term storage, place the membranes in a desiccator at room temperature or 4°C.

REAGENTS AND SOLUTIONS

Aqueous prehybridization/hybridization (APH) solution

- 5× SSC (APPENDIX 2)
- 5× Denhardt solution (APPENDIX 2)
- 1% (w/v) SDS

Add 100 µg/ml denatured salmon sperm DNA (see below) just before use

Alternatives to Denhardt solution and denatured salmon sperm DNA as blocking agents are listed in Table 2.10.5 (see discussion in critical parameters).

Denatured salmon sperm DNA

Dissolve 10 mg Sigma type III salmon sperm DNA (sodium salt) in 1 ml water. Pass vigorously through a 17-G needle 20 times to shear the DNA. Place in a boiling water bath for 10 min, then chill. Use immediately or store at -20°C in small aliquots. If stored, reheat to 100°C for 5 min and chill on ice immediately before using.

Formamide prehybridization/hybridization (FPH) solution

- 5× SSC (APPENDIX 2)
- 5× Denhardt solution (APPENDIX 2)
- 50% (w/v) formamide
- 1% (w/v) SDS

Add 100 µg/ml denatured salmon sperm DNA (see above) just before use

Alternatives to Denhardt solution and denatured salmon sperm DNA as blocking agents are listed in Table 2.10.5 (see discussion in critical parameters).

Commercial formamide is usually satisfactory for use. If the liquid has a yellow color, deionize as follows: add 5 g of mixed-bed ion-exchange resin [e.g., Bio-Rad AG 501-X8 or 501-X8(D) resins] per 100 ml formamide, stir at room temperature for 1 hr, and filter through Whatman no. 1 paper.

CAUTION: Formamide is a teratogen. Handle with care.

Labeling buffer

- 200 mM Tris·Cl, pH 7.5
- 30 mM MgCl₂
- 10 mM spermidine

Mild stripping solution

- 5 mM Tris·Cl, pH 8.0
- 2 mM EDTA
- 0.1× Denhardt solution (APPENDIX 2)

Moderate stripping solution

200 mM Tris·Cl, pH 7.0
0.1× SSC (APPENDIX 2)
0.1% (w/v) SDS

Nucleotide mix

2.5 mM ATP
2.5 mM CTP
2.5 mM GTP
20 mM Tris·Cl, pH 7.5
Store at −20°C

COMMENTARY

Background Information

Hybridization between complementary polynucleotides was implicit in the Watson-Crick model for DNA structure and was initially exploited, via renaturation kinetics, as a means of studying genome complexity. In these early applications, the two hybridizing molecules were both in solution—an approach that is still utilized in “modern” techniques such as nuclease protection transcript mapping (UNITS 4.6 & 4.7) and oligonucleotide-directed mutagenesis (Chapter 8). The innovative idea of immobilizing one hybridizing molecule on a solid support was first proposed by Denhardt (1966) and led to methods for identification of specific sequences in genomic DNA (dot blotting; UNIT 2.9B) and recombinant clones (UNITS 6.3 & 6.4). A second dimension was subsequently introduced by Southern (1975), who showed how DNA molecules contained in an electrophoresis gel could be transferred to a membrane (UNIT 2.9A), enabling genetic information relating to individual restriction fragments to be obtained by hybridization analysis.

Since the pioneering work of Denhardt and Southern, advances in membrane hybridization have been technical rather than conceptual. As reviewed by Dyson (1991), the detailed protocols have become more sophisticated, largely because of advances in understanding of the factors that influence hybrid stability and hybridization rate.

Hybrid stability is expressed as the melting temperature or T_m , which is the temperature at which the probe dissociates from the target DNA. For DNA-DNA hybrids, the T_m can be approximated from the equation of Meinkoth and Wahl (1984):

$$T_m = 81.5^\circ\text{C} + 16.6 (\log M) + 0.41(\%GC) - 0.61 (\%form) - 500/L$$

and for RNA-DNA hybrids from the equation of Casey and Davidson (1977):

$$T_m = 79.8^\circ\text{C} + 18.5(\log M) + 0.58(\%GC) - 11.8(\%GC)^2 - 0.56(\%form) - 820/L$$

where M is the molarity of monovalent cations, %GC is the percentage of guanosine and cyto-

sine nucleotides in the DNA, %form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The practical considerations that arise from these two equations are summarized Table 2.10.2A.

The second important consideration in hybridization analysis is the rate at which the hybrid is formed. Hybrid formation cannot occur until complementary regions of the two molecules become aligned, which occurs purely by chance; however, once a short nucleating region of the duplex has formed, the remaining sequences base-pair relatively rapidly. The rate at which the probe “finds” the target, which is influenced by a number of factors (Table 2.10.2B), is therefore the limiting step in hybrid formation (Britten and Davidson, 1985). However, in practical terms hybridization rate is less important than hybrid stability, as in most protocols hybridization is allowed to proceed for so long that factors influencing rate become immaterial.

Critical Parameters

To be successful, a hybridization experiment must meet two criteria:

(1) *Sensitivity*. Sufficient probe DNA must anneal to the target to produce a detectable signal after autoradiography.

(2) *Specificity*. After the last wash, the probe must be attached only to the desired target sequence (or, with heterologous probing, family of sequences).

Parameters influencing these two criteria are considered in turn, followed by other miscellaneous factors that affect hybridization.

Factors influencing sensitivity

The sensitivity of hybridization analysis is determined by how many labeled probe molecules attach to the target DNA. The greater the number of labeled probe molecules that anneal, the greater the intensity of the hybridization signal seen after autoradiography.

Probe specific activity. Of the various factors that influence sensitivity, the one that most

frequently causes problems is the specific activity of the probe. Modern labeling procedures, whether nick translation, random oligonucleotide priming (UNIT 3.5), or in vitro RNA synthesis (alternate protocol), routinely provide probes with a specific activity of $>10^8$ dpm/ μ g. This is the minimum specific activity that should be used in hybridization analysis of genomic DNA, even if the target sequences are multicopy. If the specific activity is $<10^8$ dpm/ μ g, hybridization signals will be weak or possibly undetectable, and no amount of adjusting the hybridization conditions will compensate. If there is a problem in obtaining a specific activity of $>10^8$ dpm/ μ g, it is important to troubleshoot the labeling protocol before attempting to use the probe in hybridization analysis.

If the probe is labeled to 10^8 to 10^9 dpm/ μ g, it will be able to detect as little as 0.5 pg of target DNA. Exactly what this means depends on the size of the genome being studied and the copy number of the target sequence. For human genomic DNA, 0.5 pg of a single-copy sequence 500 bp in length corresponds to 3.3 μ g

of total DNA. This is therefore the minimum amount of human DNA that should be used in a dot blot or Southern transfer if a single-copy gene is being sought.

Amount of target DNA. There is, however, a second argument that dictates that rather more than 3.3 μ g of DNA should be loaded with each dot or Southern blot. During hybridization, genuine target sequences (100% homologous to the probe) and heterologous target sequences (related but not identical to the probe) compete with one another, with the homologous reactions always predominant. Ideally this competition should be maintained until the end of the incubation period so that maximum discrimination is seen between homologous and heterologous signals. This occurs only if the membrane-bound DNA is in excess, so that target sequences are continually competing for the available probe (Anderson and Young, 1985). If the probe is in excess then the homologous reaction may reach completion (i.e., all genuine target sites become filled) before the end of the incubation, leaving a period when only

Table 2.10.2 Factors Influencing Hybrid Stability and Hybridization Rate^a

Factor	Influence
A. Hybrid stability^b	
Ionic strength	T_m increases 16.6°C for each 10-fold increase in monovalent cations between 0.01 and 0.40 M NaCl
Base composition	AT base pairs are less stable than GC base pairs in aqueous solutions containing NaCl
Destabilizing agents	Each 1% of formamide reduces the T_m by about 0.6°C for a DNA-DNA hybrid. 6 M urea reduces the T_m by about 30°C
Mismatched base pairs	T_m is reduced by 1°C for each 1% of mismatching
Duplex length	Negligible effect with probes >500 bp
B. Hybridization rate^b	
Temperature	Maximum rate occurs at 20-25°C below T_m for DNA-DNA hybrids, 10-15°C below T_m for DNA-RNA hybrids
Ionic strength	Optimal hybridization rate at 1.5 M Na^+
Destabilizing agents	50% formamide has no effect, but higher or lower concentrations reduce the hybridization rate
Mismatched base pairs	Each 10% of mismatching reduces the hybridization rate by a factor of two
Duplex length	Hybridization rate is directly proportional to duplex length
Viscosity	Increased viscosity increases the rate of membrane hybridization; 10% dextran sulfate increases rate by factor of ten
Probe complexity	Repetitive sequences increase the hybridization rate
Base composition	Little effect
pH	Little effect between pH 5.0 and pH 9.0

^aThis table is based on Brown (1991) with permission from BIOS Scientific Publishers.

^bThere have been relatively few studies of the factors influencing membrane hybridization. In several instances extrapolations are made from what is known about solution hybridization. This is probably reliable for hybrid stability, less so for hybridization rate.

heterologous hybridization is occurring and during which discrimination between the homologous and heterologous signals becomes reduced. The problem is more significant with a double-stranded rather than a single-stranded probe, as with double-stranded probe reannealing between the two probe polynucleotides gradually reduces the effective probe concentration to such an extent that it always becomes limiting towards the end of the incubation.

In practical terms it is difficult to ensure that the membrane-bound DNA is in excess. The important factor is not just the absolute amount of DNA (which is dependent on the efficiency of immobilization and how many times the membrane has been reprobed) but also the proportion of the DNA that is composed of sequences (homologous and heterologous) able to hybridize to the probe. Rather than attempting complex calculations whose results may have factor-of-ten errors, it is advisable simply to blot as much DNA as possible: 10 μ g is sufficient with most genomes. Assuming that the probe is labeled adequately and used at the correct concentration in the hybridization solution, a clear result will be obtained after autoradiography for a few hours with a simple genome (e.g., yeast DNA) or a few days with a more complex one (e.g., human DNA).

Labels other than ^{32}P . The discussion so far has assumed that the probe is labeled with ^{32}P . The lower emission energy of ^{35}S results in reduced sensitivity, and this isotope is in general unsuitable for hybridization analysis of genomic DNA. ^{35}S can be used only if the blotted DNA is exceptionally noncomplex (e.g., restricted plasmid DNA), or if the DNA is highly concentrated (e.g., colony and plaque blots; *UNIT 6.3*). Note that a membrane hybridized to a ^{35}S -labeled probe has to be dried before autoradiography, so probe stripping is not possible.

Nonradioactive probes are a more realistic option for hybridization analysis of genomic DNA and are becoming increasingly popular as the problems involved in their use are gradually ironed out. Their advantages include greater safety, the fact that large amounts of probe can be prepared in one batch and stored for years, and the rapidity of the detection protocols. Their main disadvantage is that the sensitivity of most nonradioactive detection systems is lower than that of ^{32}P autoradiography, which means that the blot and hybridization have to be carried out at maximum efficiency if a satisfactory signal is to be seen. For details on hybridization analysis with nonra-

dioactive probes, see *UNITS 3.18 & 3.19* and Mundy et al. (1991).

Using an inert polymer to increase sensitivity. In addition to adjusting the parameters discussed above, an improvement in sensitivity can also be achieved by adding an inert polymer such as 10% (w/v) dextran sulfate (molecular weight 500,000) or 8% (w/v) PEG 6000 to the hybridization solution. Both induce an increase in hybridization signal, 10-fold with a single-stranded probe and as much as 100-fold if the probe is double-stranded (Wahl et al., 1979; Amasino, 1986). The improvement is thought to arise from formation of interlocked meshes of probe molecules, which anneal en masse at target sites. Increased hybridization signals are a major bonus in detecting single-copy sequences in complex genomes, but this must be balanced with the fact that the polymers make the hybridization solutions very viscous and difficult to handle.

Factors influencing specificity

Ensuring specificity in homologous hybridization experiments. The hybridization incubation is carried out in a high-salt solution that promotes base-pairing between probe and target sequences. In $5\times$ SSC, the T_m for genomic DNA with a GC content of 50% is about 96°C . Hybridization is normally carried out at 68°C , so the specificity of the experiment is not determined at this stage. Specificity is the function of post-hybridization washes, the critical parameters being the ionic strength of the final wash solution and the temperature at which this wash is carried out.

The highly stringent wash conditions described in the basic and alternate protocols should destabilize all mismatched heteroduplexes, so that hybridization signals are obtained only from sequences that are perfectly homologous to the probe. For DNA and RNA probes (as opposed to oligonucleotides), problems with lack of specificity after the highly stringent wash occur only if the hybridizing sequences are very GC-rich, resulting in a relatively high T_m . If the high-stringency wash does not remove all nonspecific hybridization, temperature can be increased by a few degrees. The equations above for calculating T_m can be used as a guide for selecting the correct temperature for the final wash, but trial and error is more reliable. Note that a membrane that has been autoradiographed can be rewashed at a higher stringency and put back to expose again, the only limitation being decay of the label and the need for a longer exposure the second time.

Table 2.10.3 High-Salt Solutions Used in Hybridization Analysis

Stock solution	Composition
20× SSC	3.0 M NaCl/0.3 M trisodium citrate
20× SSPE ^a	3.6 M NaCl/0.2 M NaH ₂ PO ₄ /0.02 M EDTA, pH 7.7
Phosphate solution ^b	1 M NaHPO ₄ , pH 7.2 ^c

^aSSC may be replaced with the same concentration of SSPE in all protocols.

^bPrehybridize and hybridize with 0.5 M NaHPO₄ (pH 7.2)/1 mM EDTA/7% SDS [or 50% formamide/0.25 M NaHPO₄ (pH 7.2)/0.25 M NaCl/1 mM EDTA/7% SDS]; perform moderate-stringency wash in 40 mM NaHPO₄ (pH 7.2)/1 mM EDTA/5% SDS; perform high-stringency wash in 40 mM NaHPO₄ (pH 7.2)/1 mM EDTA/1% SDS.

^cDissolve 134 g Na₂HPO₄·7H₂O in 1 liter water, then add 4 ml 85% H₃PO₄. The resulting solution is 1 M Na⁺, pH 7.2.

Designing washes for heterologous hybridization. Calculations of T_m become more critical if heterologous probing is being attempted. If the aim is to identify sequences that are merely related, not identical, to the probe (e.g., members of a multigene family, or a similar gene in a second organism), then it is useful to have an idea of the degree of mismatching that will be tolerated by a “moderate-” or “low-” stringency wash. The best way to approach this is to first establish the lowest temperature at which only homologous hybridization occurs with a particular SSC concentration. Then assume that 1% mismatching results in a 1°C decrease in the T_m (Bonner et al., 1973) and reduce the temperature of the final wash accordingly (for example, if sequences with ≥90% similarity with the probe are being sought, decrease the final wash temperature by 10°C). If the desired degree of mismatching results in a wash temperature of <45°C, then it is best to increase the SSC concentration so that a higher temperature can be used. Doubling the SSC concentration results in a ~17°C increase in T_m , so washes at 45°C in 0.1× SSC and 62°C in 0.2× SSC are roughly equivalent. Note that in these extreme cases it may also be necessary to reduce the hybridization temperature to as low as 45°C (aqueous solution) or 32°C (formamide solution).

This approach sometimes works extremely well (as shown when the heterologous targets are eventually sequenced), but the assumption that a 1% degree of mismatching reduces the T_m of a heteroduplex by 1°C is very approximate. Base composition and mismatch distribution influence the actual change in T_m , which can be anything between 0.5° and 1.5°C per 1% mismatch (Hyman et al., 1973). Unfortunately trial and error is the only alternative to the “rational” approach described here.

Other parameters relevant to hybridization analysis

Length of prehybridization and hybridization incubations. The protocols recommend prehybridization for 3 hr with nitrocellulose and 15 min for nylon membranes. Inadequate prehybridization can lead to high backgrounds, so these times should not be reduced. They can, however, be extended without problem.

Hybridizations are usually carried out overnight. This is a rather sloppy aspect of the procedure, because time can have an important influence on the result, especially if, as described above, an excess amount of a single-stranded probe is being used. The difficulties in assigning values to the parameters needed to calculate optimum hybridization time has led to the standard “overnight” incubation, which in fact is suitable for most purposes. The exception is when hybridization is being taken to its limits, for instance in detection of single-copy sequences in human DNA, when longer hybridization times (up to 24 hr) may improve sensitivity if a single-stranded probe is being used. Note that this does not apply to double-stranded probes, as gradual reannealing results in only minimal amounts of a double-stranded probe being free to hybridize after ~8 hr of incubation.

Formamide hybridization buffers. Formamide destabilizes nucleic acid duplexes, reducing the T_m by an average of 0.6°C per 1% formamide for a DNA-DNA hybrid (Meinkoth and Wahl, 1984) and rather less for a DNA-RNA hybrid (Casey and Davidson, 1977; Kafatos et al., 1979). It can be used at 50% concentration in the hybridization solution, reducing the T_m so that the incubation can be carried out at a lower temperature than needed with an aqueous solution. Originally formamide was used with nitrocellulose membranes as a means of prolonging their lifetime, as the

lower hybridization temperature results in less removal of target DNA from the matrix. More recently formamide has found a second use in reduction of heterologous background hybridization with RNA probes. RNA-DNA hybrids are relatively strong, and heterologous duplexes remain stable even at high temperatures. The destabilizing effect of formamide is therefore utilized to maximize the discrimination between homologous and heterologous hybridization with RNA probes.

Formamide probably confers no major advantage on DNA-DNA hybridization with a nylon membrane. In fact it introduces two problems, the hazardous nature of the chemical itself, and an apparent reduction in hybridization rate. The latter point is controversial (Hutton, 1977), but for equivalent sensitivity a formamide hybridization reaction usually has to incubate for longer than an aqueous one.

Alternatives to SSC. Although SSC has been used in hybridization solutions for many years, there is nothing sacrosanct about the formulation, and other salt solutions can be employed (Table 2.10.3). There is little to choose between these alternatives. SSPE and phosphate solutions have a greater buffering power and may confer an advantage in formamide hybridization solutions. Alternatively, the buffering power of SSC can be increased by adding 0.3% (w/v) tetrasodium pyrophosphate.

Probe length. Probe length has a major influence on the rate of duplex formation in solution hybridization (Wetmur and Davidson, 1986), but the effect is less marked when the target DNA is immobilized. In membrane hybridization a more important factor is the specificity of the probe. The probe should never be too long (>1000 bp), as this increases the chance of heterologous duplexes remaining stable during a high-stringency wash. Neither should the probe contain extensive vector sequences, as these can hybridize to their own target sites, wrecking the specificity of the experiment.

Mechanics of hybridization. Traditionally hybridization has been carried out in plastic bags. This technique is messy, radiochemical spills being almost unavoidable, and can lead to detrimental contact effects if too many membranes are hybridized in a single bag. Hybridization incubators are now available from a number of companies and are recommended as a distinct advance over the plastic bag technology. Rotation of the hybridization tube results in excellent mixing, reducing hot spots caused by bubbles and dust and leading to very evenly

hybridized membranes. Good-quality results are possible even when ten or more minigel Southern are hybridized in a single 8.5 × 3.0-inch tube.

If bags are used, they should be of stiff plastic to prevent the sides collapsing on to the membrane, which will lead to high background. The volume of hybridization solution should be sufficient to fill the bag, and no more than two membranes should be hybridized in each bag.

Troubleshooting

Problems in blotting and hybridization reveal themselves when the autoradiograph is developed. A guide to the commonest problems and how to solve them is given in Table 2.10.4 (based on Dyson, 1991).

A particularly troublesome problem is high background signal across the entire membrane. This is due to the probe attaching to nucleic acid binding sites on the membrane surface, the same sites that bind DNA during the blotting procedure. Prehybridization/hybridization solutions contain reagents that block these sites and hence reduce background hybridization. The most popular blocking agent is Denhardt solution, which contains three polymeric compounds (Ficoll, polyvinylpyrrolidone, and BSA) that compete with nucleic acids for the membrane-binding sites. The formulations used in the basic and alternate protocols also include denatured salmon sperm DNA (any complex DNA that is nonhomologous with the target is acceptable) which also competes with the probe for the membrane sites. Blocking agents are included in the prehybridization solution to give them a head start over the probe. With a nylon membrane, the blocking agents may have to be left out of the hybridization solution, as they can interfere with the probe-target interaction. When the membranes are washed, the Denhardt solution and salmon sperm DNA are replaced with SDS, which acts as a blocking agent at concentrations ≥1%.

Other blocking agents can also be used (Table 2.10.5). With DNA blots, the main alternatives to Denhardt are heparin (Singh and Jones, 1984) and milk powder (BLOTTO; Johnson et al., 1984), although Denhardt is generally more effective, at least with nylon membranes. Note that BLOTTO contains RNases and so can be used only in DNA-DNA hybridizations. With an RNA probe, denatured salmon sperm DNA is sometimes replaced by 100 µg/ml yeast tRNA, which has the advantage that it does not need to be sheared before

Table 2.10.4 Troubleshooting Guide for DNA Blotting and Hybridization Analysis^a

Problem	Possible cause ^b	Solution
Poor signal	Probe specific activity too low	Check labeling protocol if specific activity is $<10^8$ dpm/ μ g.
	Inadequate depurination	Check depurination if transfer of DNA >5 kb is poor.
	Inadequate transfer buffer	1. Check that 20 \times SSC has been used as the transfer solution if small DNA fragments are retained inefficiently when transferring to nitrocellulose. 2. With some brands of nylon membrane, add 2 mM Sarkosyl to the transfer buffer. 3. Try alkaline blotting to a positively charged nylon membrane.
	Not enough target DNA	Refer to text for recommendations regarding amount of target DNA to load per blot.
	Poor immobilization of DNA	See recommendations in <i>UNIT 2.9A</i> commentary.
	Transfer time too short	See recommendations in <i>UNIT 2.9A</i> commentary.
	Inefficient transfer system	Consider vacuum blotting as an alternative to capillary transfer.
	Probe concentration too low	1. Check that the correct amount of DNA has been used in the labeling reaction. 2. Check recovery of the probe after removal of unincorporated nucleotides. 3. Use 10% dextran sulfate in the hybridization solution. 4. Change to a single-stranded probe, as reannealing of a double-stranded probe reduces its effective concentration to zero after hybridization for 8 hr.
	Incomplete denaturation of probe	Denature as described in the protocols.
	Incomplete denaturation of target DNA	When dot or slot blotting, use the double denaturation methods described in <i>UNIT 2.9B</i> , or blot onto positively charged nylon.
	Blocking agents interfering with the target-probe interaction	If using a nylon membrane, leave the blocking agents out of the hybridization solution.
	Final wash was too stringent	Use a lower temperature or higher salt concentration. If necessary, estimate T_m as described in <i>UNIT 6.4</i> .
	Hybridization temperature too low with an RNA probe	Increase hybridization temperature to 65°C in the presence of formamide (see alternate protocol).
	Hybridization time too short	If using formamide with a DNA probe, increase the hybridization time to 24 hr.
Inappropriate membrane	Check the target molecules are not too short to be retained efficiently by the membrane type (see Table 2.9.1).	

continued

Table 2.10.4 Troubleshooting Guide for DNA Blotting and Hybridization Analysis^a, continued

Problem	Possible cause ^b	Solution
	Problems with electroblotting	Make sure no bubbles are trapped in the filter-paper stack. Soak the filter papers thoroughly in TBE before assembling the blot. Used uncharged rather than charged nylon.
Spotty background	Unincorporated nucleotides not removed from labeled probe	Follow protocols described in <i>UNIT 3.4</i> .
	Particles in the hybridization buffer	Filter the relevant solution(s).
	Agarose dried on the membrane	Rinse membrane in 2× SSC after blotting.
Patchy or generally high background	Baking or UV crosslinking when membrane contains high salt	Rinse membrane in 2× SSC after blotting.
	Insufficient blocking agents	See text for of discussion of extra/alternative blocking agents.
	Part of the membrane was allowed to dry out during hybridization or washing	Avoid by increasing the volume of solutions if necessary.
	Membranes adhered during hybridization or washing	Do not hybridize too many membranes at once (ten minigel blots for a hybridization tube, two for a bag is maximum).
	Bubbles in a hybridization bag	If using a bag, fill completely so there are no bubbles.
	Walls of hybridization bag collapsed on to membrane	Use a stiff plastic bag; increase volume of hybridization solution.
	Not enough wash solution	Increase volume of wash solution to 2 ml/10 cm ² of membrane.
	Hybridization temperature too low with an RNA probe	Increase hybridization temperature to 65°C in the presence of formamide (see alternate protocol).
	Formamide needs to be deionized	Although commercial formamide is usually satisfactory, background may be reduced by deionizing immediately before use.
	Labeled probe molecules are too short	1. Use a ³² P-labeled probe as soon as possible after labeling, as radiolysis can result in fragmentation. 2. Reduce amount of DNase I used in nick translation (<i>UNIT 3.5</i>).
Probe concentration too high	Check that the correct amount of DNA has been used in the labeling reaction.	
Inadequate prehybridization	Prehybridize for at least 3 hr with nitrocellulose or 15 min for nylon.	
Probe not denatured	Denature as described in the protocols.	
Inappropriate membrane type	If using a nonradiocative label, check that the membrane is compatible with the detection system.	
Hybridization with dextran sulfate	Dextran sulfate sometimes causes background hybridization. Place the membrane between Schleicher and Schuell no. 589 WH paper during hybridization, and increase volume of hybridization solution (including dextran sulfate) by 2.5%.	

continued

Table 2.10.4 Troubleshooting Guide for DNA Blotting and Hybridization Analysis^a, continued

Problem	Possible cause ^b	Solution
Extra bands	Not enough SDS in wash solutions	Check the solutions are made up correctly.
	Final wash was not stringent enough	Use a higher temperature or lower salt concentration. If necessary, estimate T_m as described in <i>UNIT 6.4</i> .
	Probe contains nonspecific sequences (e.g., vector DNA)	Purify shortest fragment that contains the desired sequence.
	Target DNA is not completely restriction digested	Check the restriction digestion (<i>UNIT 3.1</i>).
Nonspecific background in one or more tracks	Formamide not used with an RNA probe	RNA-DNA hybrids are relatively strong but are destabilized if formamide is used in the hybridization solution.
	Probe is contaminated with genomic DNA	Check purification of probe DNA. The problem is more severe when probes are labeled by random printing. Change to nick translation.
	Insufficient blocking agents	See text for of discussion of extra/alternative blocking agents.
	Final wash did not approach the desired stringency	Use a higher temperature or lower salt concentration. If necessary, estimate T_m as described in <i>UNIT 6.4</i> .
Cannot remove probe after hybridization	Probe too short	Sometimes a problem with probes labeled by random priming. Change to nick translation.
	Membrane dried out after hybridization	Make sure the membrane is stored moist between hybridization and stripping.
	Poor retention of target DNA during probe stripping	1. Check calibration of UV source if cross-linking on nylon. 2. Use a less harsh stripping method (support protocol).
Decrease in signal intensity when reprobbed		

^aBased on Dyson (1991).

^bWithin each category, possible causes are listed in decreasing order of likelihood.

use. If a cDNA clone is used as the probe, or for the *in vitro* synthesis of an RNA probe, then blockage of sites with high affinity for poly(A)⁺ sequences often reduces background. This is achieved by using 10 µg/ml of poly(A) DNA as the blocking agent.

Anticipated Results

Using either a nitrocellulose or nylon membrane and a probe labeled to $\geq 10^8$ dpm/µg, there should be no difficulty in detecting 10 pg of a single copy sequence in human DNA after 24 hr autoradiography.

Time Considerations

The hybridization experiment can be completed in 24 hr, the bulk of this being taken up by the overnight incubation. Prehybridization takes 3 hr for a nitrocellulose membrane or 15 min for nylon. Post-hybridization washing to high stringency can usually be completed in 1.5 hr. If a single-copy sequence in human DNA is being probed, the hybridization step may be extended to 24 hr, with a concomitant increase in the length of the experiment as a whole.

The length of time needed for autoradiography depends on the abundance of the target sequences in the blotted DNA. Adequate exposure can take anything from overnight to several days.

Table 2.10.5 Alternatives to Denhardt/Denatured Salmon Sperm DNA as Blocking Agents in DNA Hybridization^a

Blocking agent	Composition	Storage and use
BLOTTO	5% (w/v) nonfat dried milk/0.02% (w/v) NaN ₃ in H ₂ O	Store at 4°C; use at 4% final concentration
Heparin (porcine grade II)	50 mg/ml in 4× SSC	Store at 4°C. Use at 500 µg/ml with dextran sulfate or 50 µg/ml without
Yeast tRNA	10 mg/ml in H ₂ O	Store at 4°C; use at 100 µg/ml
Homopolymer DNA	1 mg/ml poly(A) or poly(C) in H ₂ O	Store at 4°C; use at 10 µg/ml in water; appropriate targets: poly(A) for AT-rich DNA, poly(C) for GC-rich DNA

^aThis table is based on Brown (1991) with permission from BIOS Scientific Publishers.

Literature Cited

- Amasino, R.M. 1986. Acceleration of nucleic acid hybridization rate by polyethylene glycol. *Anal. Biochem.* 152:304-307.
- Anderson, M.L.M. and Young, B.D. 1985. Quantitative filter hybridisation. In *Nucleic Acid Hybridisation: A Practical Approach* (B.D. Hames and S.J. Higgins, eds.) pp. 73-111. IRL Press at Oxford University Press, Oxford.
- Bonner, T.I., Brenner, D.J., Neufeld, B.R., and Britten, R.J. 1973. Reduction in the rate of DNA reassociation by sequence divergence. *J. Mol. Biol.* 81:123-135.
- Britten, R.J. and Davidson, E.H. 1985. Hybridisation strategy. In *Nucleic Acid Hybridisation: A Practical Approach* (B.D. Hames and S.J. Higgins, ed.) pp. 3-46. IRL Press at Oxford University Press, Oxford.
- Brown, T.A. (ed.) 1991. *Molecular Biology Labfax*. BIOS Scientific Publishers, Oxford.
- Casey, J. and Davidson, N. 1977. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes as high concentrations of formamide. *Nucl. Acids Res.* 4:1539-1552.
- Denhardt, D.T. 1966. A membrane-filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Commun.* 23:641-646.
- Dyson, N.J. 1991. Immobilization of nucleic acids and hybridization analysis. In *Essential Molecular Biology: A Practical Approach*, Vol. 2 (T.A. Brown, ed.) pp. 111-156. IRL Press at Oxford University Press, Oxford.
- Hutton, J.R. 1977. Renaturation kinetics and thermal stability of DNA in aqueous solutions of formamide and urea. *Nucl. Acids Res.* 4:3537-3555.
- Hyman, R.W., Brunovskis, I., and Summers, W.C. 1973. DNA base sequence homology between coliphages T7 and ϕ 11 and between T3 and ϕ 11 as determined by heteroduplex mapping in the electron microscope. *J. Mol. Biol.* 77:189-196.
- Johnson, D.A., Gautsch, J.W., Sportsman, J.R., and Elder, J.H. 1984. Improved technique utilizing non-fat dry milk for analysis of proteins and nucleic acids transferred to nitrocellulose. *Gene Anal. Tech.* 1:3.
- Kafatos, F.C., Jones, C.W., and Efstratiadis, A. 1979. Determination of nucleic acid sequence homologies and relative concentrations by a dot blot hybridization procedure. *Nucl. Acids Res.* 7:1541-1552.
- Little, P.F.R. and Jackson, I.J. 1987. Application of plasmids containing promoters specific for phage-encoded RNA polymerases. In *DNA Cloning: A Practical Approach*, Vol. 3 (D.M. Glover, ed.) pp. 1-18. IRL Press at Oxford University Press, Oxford.
- Meinkoth, J. and Wahl, G. 1984. Hybridization of nucleic acids immobilized on solid supports. *Anal. Biochem.* 138:267-284.
- Mundy, C.R., Cunningham, M.W., and Read, C.A. 1991. Nucleic acid labelling and detection. In *Essential Molecular Biology: A Practical Approach*, Vol. 2 (T.A. Brown, ed.) pp. 57-109. IRL Press at Oxford University Press, Oxford.
- Singh, L. and Jones, K.W. 1984. The use of heparin as a simple cost-effective means of controlling background in nucleic acid hybridization procedures. *Nucl. Acids Res.* 12:5627-5638.
- Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
- Wahl, G.M., Stern, M., and Stark, G.R. 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization by using dextran sulfate. *Proc. Natl. Acad. Sci. U.S.A.* 76:3683-3687.
- Wetmur, J.G. and Davidson, N. 1986. Kinetics of renaturation of DNA. *J. Mol. Biol.* 31:349-370.

Key Reference

Dyson, N.J. 1991. See above.

Provides a detailed account of factors influencing hybridization.

Contributed by Terry Brown
University of Manchester Institute of
Science and Technology
Manchester, United Kingdom

**Preparation and
Analysis of DNA**

2.10.16

SYNTHESIS AND PURIFICATION OF OLIGONUCLEOTIDES

Synthesis and Purification of Oligonucleotides

Avery's realization that DNA carries the genetic information lead chemists on a 40-year search that has culminated in efficient, automated oligonucleotide synthesis on solid-phase supports. Modern nucleic acid synthesizers utilize phosphite triester chemistries that employ stable phosphoramidite monomers to build the growing polymer. These robust reactions allow both chemists and molecular biologists to easily generate specific ribo- and deoxyribo oligonucleotides with a variety of labels, modified linkages, and nonstandard bases attached throughout the chain. The synthesis of short (less than 40-nucleotide) probes and primers requires no more special expertise than the ability to read the synthesizer operator's manual, and longer oligonucleotides (up to 150 nucleotides) can be synthesized with a little more care.

The introductory section of this unit provides strategies on the maximization of synthetic yield, the generation of sequences containing site specific modifications, and the isolation of synthetic oligonucleotides. Protocols describe monitoring the progress of synthesis via the trityl assay (see Basic Protocol 1 and Support Protocol) and methods for deprotection of DNA (see Basic Protocol 2) and RNA (see Basic Protocol 3) oligonucleotides.

This unit augments the detailed instructions provided by the manufacturers of oligonucleotide synthesizers. A functional understanding of the synthesis chemistries coupled to insights on the mechanical operation of the synthesizer will allow the user to minimize input time and maximize oligonucleotide output.

INTRODUCTION TO CHEMICAL NUCLEIC ACID SYNTHESIS

Many hydroxyl and amine moieties make nucleic acids very animated molecules with rich chemistries of their own that can interfere with the phosphite triester reactions used to couple the nucleotide monomers; therefore, protection strategies are necessary in chemical synthesis to mask the functional groups on the monomers so that the only significant reaction is the desired 3' to 5' sequential condensation of monomers to the growing oligonucleotide. (Note that enzymatic polymerizations occur with the opposite directionality and require no masking of monomer functionality.) These synthetic protecting groups must be chosen so that they can be removed easily to expose the natural nucleotides.

The fully protected monomers for nucleic acid synthesis are generally called phosphoramidites (Fig. 2.11.1). In traditional protection schemes the nucleophilic amino moieties on the bases are protected with either isobutyryl (N-2 of guanine) or benzoyl (N-6 of adenine and N-4 of cytidine) groups, both of which can be removed at the completion of synthesis by ammoniolysis. However, recent advances have led to the wide spread use of phenoxyacetyl (PAC) protection of adenosine, dimethylformamide (DmF) protection of guanosine, and acyl protection of cytosine to yield oligonucleotides that can be deprotected rapidly under very mild conditions (Reddy et al., 1994). The 5' primary hydroxyl of the ribose sugar is protected with a dimethoxytrityl (DMT) ether moiety which is removed by mild protic acids at the beginning of each coupling cycle. The efficiency of synthesis at each coupling cycle can be monitored by detecting the release of the chromophore trityl cation. To synthesize nucleic acids with the natural phosphodiester backbone, the 3' secondary hydroxyl function of the ribose sugar is

derivatized with a highly reactive phosphitylating agent. The phosphate oxygen on this moiety is usually masked by β -cyanoethoxy and diisopropylamine protecting groups. By insulating the phosphate oxygen with alternative groups, modified phosphate backbones may be accessed. Finally, for ribonucleic acids the secondary 2' hydroxyl of the ribose is shielded throughout the chemical synthesis by the *tert*-butyldimethylsilyl group.

The genius of using these protecting groups for automated nucleic acid synthesis is that they yield nearly lesion-free natural nucleic acids with high efficiency through simple hydrolysis, nucleophilic displacement, and redox chemistries. In a standard synthesis cycle, the nucleotide chain grows from an initial protected nucleoside bound via its terminal 3' hydroxyl to a solid support. Reagents and solvents are pumped through the support to induce the consecutive removal and addition of sugar protecting groups in order

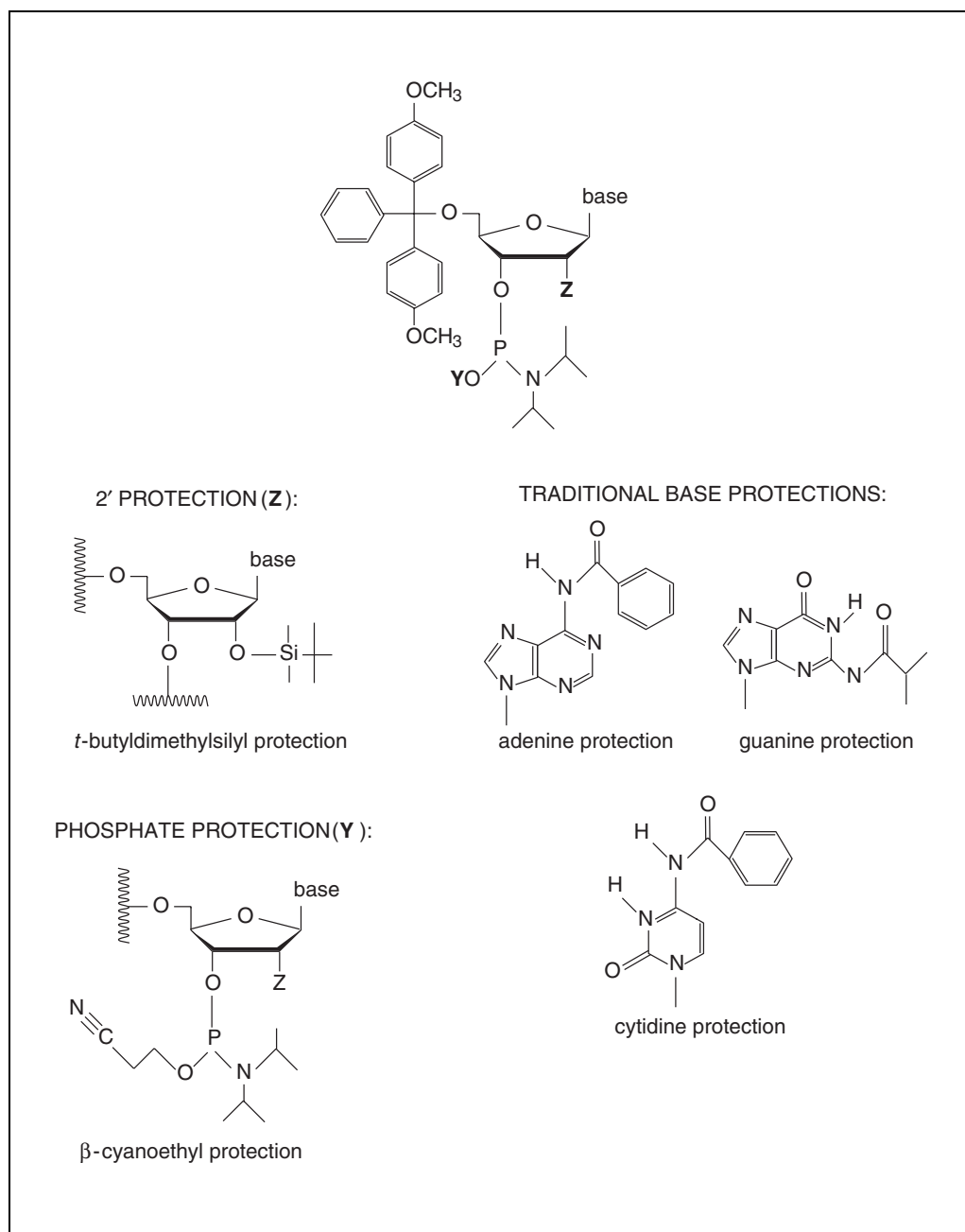


Figure 2.11.1 Structure of nucleoside phosphoramidite units, showing traditional protection groups for 2' hydroxyl, phosphate, and base moieties. For DNA, Z = H (no protection required); for RNA, Z = O-protection group.

to isolate the reactivity of a specific chemical moiety on the monomer and affect its stepwise addition to the growing oligonucleotide chain. This design eliminates the need to purify synthetic intermediates or unreacted reagents, because they are simply rinsed off the column at the end of each chemical step. Assembly of the protected oligonucleotide chain is carried out in four chemical steps: deblocking, activation/coupling, oxidation, and capping (Fig. 2.11.2). Cleavage and deprotection then reveal the single-stranded nucleic acid.

Deblocking

The synthesis cycle begins with the removal of the acid-labile DMT ether from the 5' hydroxyl of the 3' terminal nucleoside. This is usually accomplished by using dichloroacetic acid (DCA) in dichloromethane. The resulting trityl cation chromophore can be quantitated to determine coupling efficiency (see Basic Protocol 1). After deblocking, the 5' hydroxyl is the only reactive nucleophile capable of participating in the subsequent coupling step. Since the nitrogenous bases of the growing DNA chain are susceptible to acid-catalyzed depurination, the deblocking step is short, and an acetonitrile rinse thoroughly removes the deblocking agent from the support. Also, coupling efficiency and accuracy are increased by this wash, since premature detritylation of the incoming phosphoramidite monomer is prevented.

Activation/Coupling

Following deblocking of the 5' hydroxyl group, the next protected phosphoramidite is delivered to the reaction column along with the weakly acidic activator tetrazole ($pK_a = 4.8$). Nucleophilic attack of the previously freed 5' hydroxyl upon the incoming monomer's diisopropylamine-protected phosphorus, which was activated via protonation by tetrazole, elongates the nucleic acid chain. Because this protonated phosphoramidite is so reactive, the coupling reaction is usually complete within 30 sec. A molar excess of tetrazole over the phosphoramidite ensures complete activation, and a molar excess of phosphoramidite over free 5' hydroxyls of the growing chain promotes efficient coupling. To optimize the coupling efficiency, the amounts of reagents injected and the coupling time can be varied (see Synthesizing Long Oligonucleotides).

Capping

In spite of these efficiency measures, a small percentage of the support-bound nucleoside's 5' hydroxyls do not couple to the incoming activated monomer. They must be rendered inactive to minimize deletion products and simplify the purification process. Usually, acetic anhydride and *N*-methylimidazole dissolved in pyridine and tetrahydrofuran (THF) act to create an acylating agent that "caps" the unextended 5' hydroxyls. The 5' acetyl ester cap is unreactive in all subsequent cycles and is removed during the final ammonia deprotection step. Additional acetonitrile washing subsequent to capping can increase synthetic yield. After coupling and capping, the internucleotide linkage is a trivalent phosphite triester that is extremely unstable and must be oxidized to a phosphotriester which will ultimately yield natural DNA.

Oxidation

In the last step of the cycle, the unstable phosphite triester linkages are oxidized to a more stable phosphotriester by 0.02 M iodine dissolved in water/pyridine/THF. An iodine-pyridine adduct forms to the phosphite triester and is subsequently displaced by water to yield phosphorus oxidized to the pentavalent state. Pyridine also neutralizes the hydrogen iodide byproduct. Because the oxidizer contains water, the support is rinsed several times with acetonitrile following this reaction. One cycle of monomer addition is then complete, and another cycle begins with the removal of the 5' DMT from the previously added monomer.

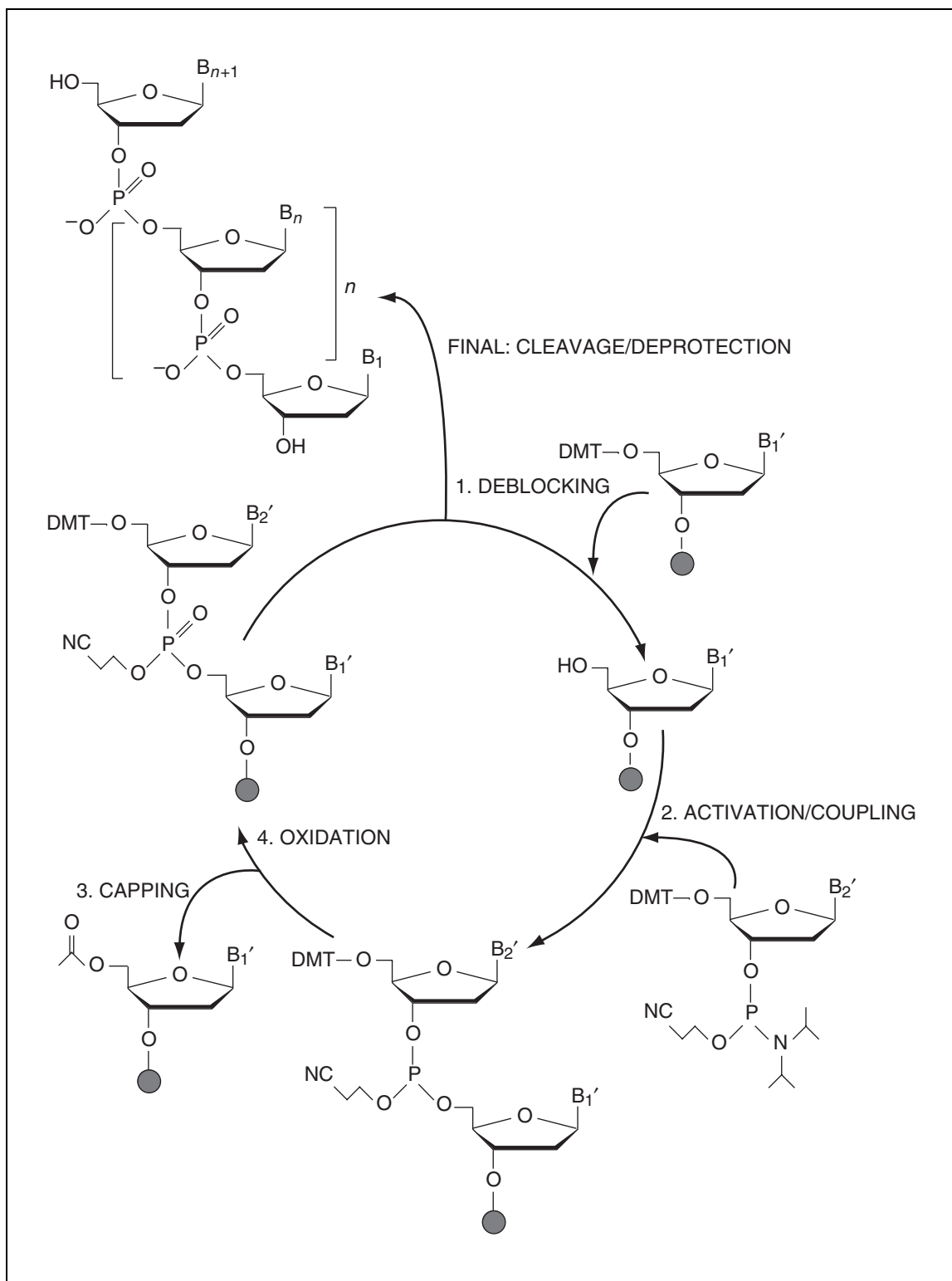


Figure 2.11.2 Steps in the assembly of the protected oligonucleotide chain.

Cleavage/Deprotection

At the end of the synthesis, the final trityl can either be removed with a final acid wash (“trityl-off”), or be left on for purification purposes (“trityl-on”). The oligonucleotide itself is removed from the support with concentrated ammonium hydroxide. Additionally, this treatment deprotects the phosphorus by β -elimination of the cyanoethyl group and removes the protecting groups from the heterocyclic bases to yield a single-stranded nucleic acid.

RNA Synthesis

RNA chemical synthesis is identical to that used for DNA except for the need for an additional protecting group at the 2' hydroxyl of ribose. This position is usually protected with *tert*-butyldimethyl silyl groups, which are stable throughout the synthesis (Fig. 2.11.1). They are removed at the final deprotection step by the basic fluoride ion. The remaining positions on both the sugar and the bases are protected in the same fashion as for DNA. By adjusting several parameters in the DNA synthesis protocol—including the coupling times, monomer delivery rate, frequency of washing steps, and types of capping reagents—stepwise coupling efficiencies of up to 99% can be obtained (Wincott et al., 1995; G. Glick, pers. comm.). However, for the casual user this yield represents the exception rather than the norm, and only shorter oligoribonucleotides (<20 bases) should initially be attempted.

STRATEGIES FOR NUCLEIC ACID SYNTHESIS

A Checklist for Nucleic Acid Synthesis

Consistency and planning are the keys to reliable nucleic acid synthesis. Organizing these repetitive tasks into a standard operating protocol will streamline efficiency and produce better-quality oligonucleotides.

1. Plan syntheses to optimize machine use. Oligonucleotides of similar size should be combined in parallel runs, since synthesizing many short oligonucleotides followed by a longer one is faster than mixing the sets on dual-column synthesizers.
2. Determine the total number of bases to be incorporated. Be sure there are enough of the required reagents and phosphoramidites available for the entire synthesis. Consult the synthesizer manual for the amount of reagent needed for each coupling (also see Troubleshooting). Remember that different synthetic scales (e.g., 0.25 μ mol versus 1 μ mol) require different amounts of reagents. Syntheses should be planned so that a phosphoramidite is almost completely exhausted. Also, phosphoramidites that have been dissolved for >2 weeks should be replaced.
3. Consider special programming requirements. Many synthesis options pertaining to the scale of the synthesis, backbone composition, and presence of protecting groups can be modified. Create a log for users to fill out detailing exact synthesis requirements, and check the log prior to synthesis. A computerized log book is especially useful and allows for an organized oligonucleotide nomenclature (e.g., R20.17 may refer to Rebecca's 20-mer, the 17th 20-mer made on the system).
4. If the synthesizer lacks a trityl monitor, set up a fraction collector to monitor the chromophoric trityl cation release (see Basic Protocol 1).
5. Initialize the system. If previous oligonucleotides were cleaved from their supports automatically, rinse the columns for 30 sec with acetonitrile to remove any remaining traces of ammonium hydroxide. If the machine has been at rest for >6 hr, immediately

before starting the synthesis (and following the addition of fresh reagents) remove any stale reagents or moisture from the lines by priming them. This will maximize the first coupling step's efficiency. Check the reagent and phosphoramidite flow rates to ensure that reagents are being properly delivered.

6. Start the synthesis. Confirm that the flow through the lines and columns is not obstructed and monitor the first few trityl releases. An abortive synthesis of a 60-mer sequence wastes much more material than a failed 3-mer run.

Synthesizing Long Oligonucleotides (≥ 100 Bases)

Modern synthesizers can routinely produce sequences of 150 or more nucleotides in usable amounts ($\geq 10 \mu\text{g}$). Several strategies can be employed to enhance the yield from syntheses longer than 100 bases.

1. Exclude water from the system (the importance of this cannot be overstated; see Critical Parameters, Excluding water from solvents). Replace reagents on the machine with fresh ones before all long syntheses. This is particularly important for the phosphoramidites and especially for guanosine phosphoramidite, since it decomposes more quickly than the other two protected bases (Zon et al., 1985).
2. Monitor trityl releases for shorter runs prior to attempting the synthesis of long oligonucleotides to limit wastage of expensive reagents. In general, if the stepwise efficiency of synthesis is $<99\%$, alter parameters to increase the efficiency on shorter sequences before attempting to synthesize a long oligonucleotide (see Basic Protocol 1).
3. Use dichloroacetic acid (DCA) rather than trichloroacetic acid (TCA) for deblocking if the synthesizer is compatible with this reagent. Depurination (cleavage of the glycosidic bond) under acidic conditions is a prominent side reaction that ultimately limits DNA synthesis. DCA tends to show much better synthetic yields than trichloroacetic acid, especially for longer oligonucleotides (R. Pon, pers. comm.). Use a 2% (v/v) DCA/1,2-dichloroethane mixture.
4. Modify the synthesis protocol to increase the coupling time of the phosphoramidite. Also, additional methylene chloride wash steps included prior and subsequent to deblocking, along with increased acetonitrile washing subsequent to capping, lead to increased yields (G. Glick, pers. comm.).
5. Increase the phosphoramidite concentration to enhance the coupling efficiency—e.g., use a concentration of 50 mg/ml (double the normal concentration; 20-fold molar excess over the synthetic polynucleotide chain) for longer sequences.
6. Using a support matrix such as control pore glass (CPG) with a loading capacity of $<40 \mu\text{mol/g}$ can greatly increase yields of long oligonucleotide. Furthermore, the pore size of the support should be 1000 Å for >100 -mers and 2000 Å for 200-mers (Gait, 1984) to alleviate molecular crowding and steric effects. For a typical 1- μmol scale synthesis of a 150-mer, 20 mg of a support with a loading capacity of 5 mmol/g is used.

If it is too difficult to synthesize the desired sequence in a reasonable yield, or if oligomers >150 bases are desired, the nucleic acids may be made in segments and ligated (Bartel and Szostak, 1993) following PCR with a proofreading polymerase such as *Pfu* or *Tth*. Note, however, that since PCR is an inherently mutagenic procedure, any products generated should be checked by sequencing. Also, very long synthetic oligonucleotides (300 to 600 bases) have also been synthesized directly, and in spite of incredibly low

yields, rare full-length products have been successfully amplified by PCR (Ciccarelli et al., 1991). Finally, mutually primed synthesis (UNIT 8.2) can also be a suitable option for oligomers >150 bases.

Synthesizing RNA

Since many structural and mechanistic studies are underway concerning catalytic RNAs, the catalog of commercially available modified RNA monomers has recently bloomed (Table 2.11.1). RNA chemical synthesis has become as routine as that of DNA and typically uses identical 5'-dimethoxytrityl- β -cyanoethyl-protected phosphoramidites except for an additional protecting group on the 2' hydroxyl. *tert*-Butyldimethylsilyl protection of the 2'-hydroxyl group is the basis of most commercially available RNA phosphoramidites, since the silyl group is stable to both acid and base and can be removed with fluoride ion. Recently, however, RNA monomers with 2'-acetyl groups (FPMP, CEE) have appeared; these have the advantage of being conveniently removed at pH 2 just prior to use of the particular sample of RNA. Strategies for ribophosphoramidite protection are an active area of research, and recent work with 5'-silyl ethers in conjunction with 2'-orthoester protection has proven particularly interesting (S. Scaaringe, pers. comm.). For the casual user of RNA, it is often easier just to purchase small quantities of the required sequence from a ribo-oligonucleotide synthesis company such as Baron Consulting, Dharmacon Research, Genosys, or Peninsula Labs (see APPENDIX 4).

Isomeric purity of the phosphoramidites is often variable because of the difficulties inherent in distinguishing between the vicinal hydroxyls of ribose. If a homogeneous population of 5'-to-3'-linked oligoribonucleotides is required (as in most cases), then a thin-layer chromatography (TLC) or ^{31}P nuclear magnetic resonance (NMR) analysis of the starting phosphoramidites should be performed to establish their isomeric composition. While phosphorus NMR facilities are not generally available to molecular biologists, TLC is both inexpensive and straightforward. Recommended solvent systems to separate the 2' and 3' ribonucleotide phosphoramidites are 1:1 ether/chloroform, or 40:58:2 or 50:46:4 dichloromethane/hexane/triethylamine (Usman et al., 1987).

With a few slight modifications, the procedures and precautions described for DNA synthesis chemistry apply to RNA as well. Since stepwise coupling efficiency is lower than that of DNA, even greater care should be taken to exclude water completely from the closed system. Because the 2' hydroxyl is often protected with sterically hindering protecting groups, reaction times for RNA reagents tend to be longer, and adjustments should be made to phosphoramidite concentrations and coupling times, as detailed below. As is true for all RNA work, equipment and reagents that will contact unprotected oligonucleotides should be RNase free (UNITS 3.12 & 4.1) to avoid degradation of the synthesized material.

Depending on the synthesizer and coupling program used, RNA phosphoramidites are suspended in dry acetonitrile at a concentration of 0.1 to 1.0 M with a 6- to 10-fold excess of reagents delivered per 300 sec coupling. *S*-ethyltetrazole and DCI have also been found to be a more effective activators than the traditional tetrazole (Sproat et al., 1995; Vargeese et al, 1998). Also, additional methylene chloride washing steps included prior to and subsequent to deblocking, along with increased acetonitrile washing subsequent to capping, lead to increased yields (G. Glick, pers. comm.).

Table 2.11.1 Suppliers of Unnatural and Modified Phosphoramidites

Phosphoramidite ^a	Suppliers ^b
<i>Fast/mild deprotecting DNA monomers</i>	
Ac-C-3'-CED	Cruachem, Glen
DmF-G-3'-CED	Glen, SPS, PE
Ibu-C-3'-CED	Glen, PE
N-PAC-A-3'-CED	BioGenex, ChemGenes, Cruachem, PE, PerSeptive, Sigma, SPS
N-PAC-C-3'-CED	ChemGenes, PerSeptive
N-PAC-G-3'-CED	BioGenex, ChemGenes, Cruachem, PE, PerSeptive, Sigma, SPS
<i>Fast/mild deprotecting RNA monomers</i>	
Ac-rC-2'-tBDSilyl-3'-CED	Glen
DmF-rG-2'-tBDSilyl-3'-CED	Glen, PE, SPS
Ibu-rC-2'-tBDSilyl-3'-CED	PE
N-PAC-rA-2'-tBDSilyl-3'-CED	BioGenex, ChemGenes, Cruachem, PE, Sigma, SPS
N-PAC-rC-2'-tBDSilyl-3'-CED	ChemGenes, Sigma
N-PAC-rG-2'-tBDSilyl-3'-CED	BioGenex, ChemGenes, Cruachem, Sigma, SPS
2'-FPMP-rA-3'-CED	Sigma
2'-FPMP-rC-3'-CED	Sigma
2'-FPMP-rG-3'-CED	Sigma
2'-FPMP-rU-3'-CED	Sigma
2'-FPMP-rI-3'-CED	Sigma
<i>Modified DNA monomers</i>	
α -Anomers of deoxyribose	Appligene, Interactiva
O6-Methyl-G-3'-CED	Glen
2'-Deoxy-nebularine-3'-CED (degenerate base)	Glen
2-Amino-A-3'-CED	Glen
2-Aminopurine-3'-CED	Glen
2-Thio-T-3'-CED	Glen
3'-Deoxy-A-3'-CED	Glen
3-Nitropyrrole-2'-deoxy-3'-CED (M) (universal base)	Glen
4-Methyl-indole-3'-CED	Glen
4-Thio-T-3'-CED	Glen
5-Halogenated-C-3'-CED	ChemGenes, Cruachem, Glen, Sigma
5-Halogenated-dU-3'-CED	ChemGenes, Cruachem, Glen, Sigma
5-Methyl-C-3'-CED	ChemGenes, Glen, Sigma
5-Methyl-iso-C-3'-CED (alternative base pair)	Glen
5-Nitroindole-3'-CED (universal base)	Glen
7-Deaza-A-3'-CED	ChemGenes, Glen
7-Deaza-G-3'-CED	ChemGenes, Glen
8-Halogenated-A-3'-CED	Glen
8-Halogenated-G-3'-CED	ChemGenes, Glen
8-Oxo-A-3'-CED	Glen
8-Oxo-G-3'-CED	Glen
8-Oxo-A-3'-CED	Glen
8-Oxo-G-3'-CED	Glen

continued

Table 2.11.1 Suppliers of Unnatural and Modified Phosphoramidites, continued

Phosphoramidite ^a	Suppliers ^b
ARA-C-3'-CED	ChemGenes, Glen
C-5 Propyne C-3'-CED (helix stability)	Glen
C-5 Propyne U-3'-CED (helix stability)	Glen
dU-3'-CED	ChemGenes, Glen, Sigma
Etheno-A-3'-CED	ChemGenes, Glen
Etheno-C-3'-CED	ChemGenes
Inosine-3'-CED (universal base)	ChemGenes, PE, Sigma, SPS
Iso-G-3'-CED (alternative base pair)	Glen
K-3'-CED (degenerate base)	Glen
N3-Methyl-T-3'-CED	ChemGenes
N6-Methyl-A-3'-CED	Glen
O4-Methyl-T-3'-CED	Glen
O4-Methyl-T-3'-CED	Glen
O6-Methyl-G-3'-CED	Cruachem, Glen
P-3'-CED (degenerate base)	Glen
Purine-deoxyribose phosphoramidite	ChemGenes
<i>Convertible deoxynucleosides</i>	
O4-Triaz-dU-3'-CED	Glen
O4-Triaz-T-3'-CED	Glen
O6-Phenyl-I-3'-CED	Glen
S6-DNP-G-3'-CED	Glen
TMP-F-dU-3'-CED	Glen
<i>Radiolabeled deoxyphosphoramidites</i>	
	Cambridge Isotopes
<i>Modified RNA monomers</i>	
2',3'-Diacetyl-rG-3'-CED	ChemGenes
2',3'-Diacetyl-rU-3'-CED	ChemGenes
2'-CED-rC-3'-tBDSilyl	BioGenex
2'-CED-rU-3'-tBDSilyl	BioGenex
2'-O-Allyl-(2-amino)-rA-3'-CED	BM
2'-O-Allyl-rA-3'-CED	BM
2'-O-Allyl-rC-3'-CED	BM
2'-O-Allyl-rG-3'-CED	BM
2'-O-Allyl-rI-3'-CED	BM
2'-O-Allyl-rU-3'-CED	BM
2'-OMe-2-amino-rA-3'-CED	Glen
2'-OMe-2-aminopurine-3'-CED	Glen
2'-OMe-3-deaza-5-aza-rC-3'-CED	Glen
2'-OMe-Ac-rC-3'-CED	Glen
2'-OMe-DMF-rG-CE-3'-CED	Glen
2'-OMe-propynyl-rC-3'-CED	Glen
2'-OMe-propynyl-rU-3'-CED	Glen
2'-OMe-rA-3'-CED	Glen
2'-OMe-rC-3'-CED	Glen
2'-OMe-rG-3'-CED	Glen
2'-OMe-rI-3'-CED	Glen
2'-OMe-rU-3'-CED	Glen
2'-OMe-TMP-5-F-rU-CE-3'-CED	Glen

*continued***Preparation and
Analysis of DNA****2.11.9**

Table 2.11.1 Suppliers of Unnatural and Modified Phosphoramidites, continued

Phosphoramidite ^a	Suppliers ^b
2'-tBDSilyl-rI-3'-CED	Dalton, PE, Sigma
2'-tBDSilyl-rT-3'-CED	ChemGenes
5-Fluoro-3'-tBDSilyl-rU-3'-CED	ChemGenes
5-Halogenated-2'-tBDSilyl-rC-3'-CED	ChemGenes
5-Halogenated-2'-tBDSilyl-rU-3'-CED	ChemGenes, Glen
5-Methyl-2'-tBDSilyl-rC-3'-CED	ChemGenes
7-Deaza-2'-tBDSilyl-rA-3'-CED	ChemGenes
7-Deaza-2'-tBDSilyl-rG-3'-CED	ChemGenes
8-Bromo-2'-tBDSilyl-rA-3'-CED	ChemGenes
8-Bromo-3'-tBDSilyl-rA-3'-CED	ChemGenes
Etheno-2'-tBDSilyl-rA-3'-CED	ChemGenes
Etheno-2'-tBDSilyl-rC-3'-CED	ChemGenes
N-PAC-2'-CED-rA-3'-tBDSilyl	BioGenex
N-PAC-2'-CED-rG-3'-tBDSilyl	BioGenex
N-PAC-2'-OMe-rA-3'-CED	BioGenex, Cruachem
N-PAC-2'-OMe-rG-3'-CED	BioGenex, Cruachem
N-PAC-8-Bromo-2'-tBDSilyl-rA-3'-CED	ChemGenes
N3-(thiobenzoyl)ethyl-2'-tBDSilyl-rU-3'-CED	ChemGenes
N3-Methyl-2'-tBDSilyl-rU-3'-CED	ChemGenes
Labeling monomers	
5'-Acridine-3'-CED	Clontech, Cruachem, Glen
5'-Amino-modifiers-3'-CED	Clontech, Glen
5'-Amino-modifiers-3'-CED (terminal)	Glen, SPS, PE
5'-Biotin-modifiers-3'-CED	BioGenex (highly rigid cyclohexane spacer), Clontech, Glen
5'-Biotin-modifiers-3'-CED (terminal)	Cruachem, Glen, PE, Sigma, SPS, ChemGenes, BM
5'-BODIPY-3'-CED (terminal)	ABI
5'-Carboxy-modifier-3'-CED	Glen
5'-Cholesteryl-3'-CED	Glen
5'-Cholesteryl-3'-CED (terminal)	Clontech
5'-Dansyl-3'-CED (terminal)	Clontech
5'-Digoxigenin-3'-CED	PE, PerSeptive
5'-DNP-3'-CED	Clontech
5'-DNP-3'-CED (terminal)	Clontech, Cruachem, Glen
5'-FAM-3'-CED (terminal)	ABI, BM, Cruachem, Glen, PE
5'-Fluorescein-3'-CED	BioGenex (highly rigid cyclohexane spacer), ChemGenes, Clontech, Cruachem, Glen
5'-HEX-3'-CED (terminal)	BM, Cruachem, Glen, PE
5'-JOE-3'-CED (terminal)	ABI
5'-Phosphate-3'-CED	ChemGenes, Clontech, Glen, SPS
5'-Psoralen-3'-CED (terminal)	ChemGenes, Clontech, Glen
5'-Pyrene-3'-CED (terminal)	Clontech
5'-ROX-3'-CED (terminal)	ABI
5'-TAMARA-3'-CED (terminal)	ABI
5'-TET-3'-CED (terminal)	BM, Cruachem, Glen, PE
5'-Texas Red-3'-CED (terminal)	Genosys

continued

Table 2.11.1 Suppliers of Unnatural and Modified Phosphoramidites, continued

Phosphoramidite ^a	Suppliers ^b
5'-Thio-modifiers-3'-CED (terminal)	ChemGenes, Cruachem, Glen, SPS
3'-Acridine-CPG	Clontech, Cruachem, Glen
3'-Amine-CPG	ChemGenes, Clontech, Glen (photolabile), SPS
3'-Biotin-CPG	ChemGenes, Clontech, Glen
3'-Carboxyl-CPG (photolabile)	Glen
3'-Cholesteryl-CPG	Clontech, Glen
3'-DabcyI-CPG	Glen
3'-Digoxigenin-CPG	Clontech, Cruachem
3'-DNP-CPG	Clontech
3'-Fluorescein-CPG	SPS
3'-Glycerol-CPG	Glen
3'-Phosphate-CPG	ChemGenes, Glen, PE, SPS
3'-TAMRA-CPG	SPS
3'-Thio-CPG	Glen, SPS
<i>Uniquely structured oligonucleotides</i>	
Branched oligonucleotide synthesis	Clontech
Cyclic oligonucleotides (up to 30)	Glen
Deoxyribose spacer-3'-CED	Glen
C9 spacer-3'-CED	Glen
C18 spacer-3'-CED	Glen
<i>Non-enzymatically-extendable 3' ends</i>	
2',3'-Dideoxy-A-CPG	Glen
2',3'-Dideoxy-C-CPG	Glen
3'-C3 spacer-CPG	Glen
3'-Deoxy-A-CPG	Glen
3'-Deoxy-C-CPG	Glen
3'-Deoxy-G-CPG	Glen
3'-Deoxy-T-CPG	Glen
5'-Amino-T-3'-CED	Glen
5'-OMe-T-3'-CED	Glen
<i>Alternative backbones</i>	
H-phosphate chemistries lead to phosphothioates, phosphoramidates, or phosphotriesters	Glen
Methyl phosphonate linkages for DNA and RNA	ChemGenes, Glen
Sulfurizing reagents to convert to phosphothioates	Glen
Various 5' supports (3'-3' or 5'-5' linkages or opposite-sense synthesis)	Glen

^aAll phosphoramidites are 5' DMT and 3' cyanoethyl protected (CED) unless otherwise noted.

^bABI, ABI Biotechnology; BM, Boehringer Mannheim; Glen, Glen Research; Interactiva, Interactiva Biotechnologie; PerSeptive, PerSeptive BioSystems; PE, Perkin-Elmer; SPS, Solid Phase Science. See APPENDIX 4 for supplier contact information.

Incorporation of Modified Nucleosides

Chemical nucleic acid synthesis allows for the incorporation of unnatural or modified bases, as well as a variety of labeling moieties, into an oligonucleotide. This can be extremely useful for testing models of structural interactions between enzymes and nucleic acids, selecting labeled molecules from a population of unlabeled ones, or gaining insights into the parameters which govern nucleic acid structure and chemistry. Modified backbone chemistries such as phosphorothioates, phosphoramidates, and phosphotriesters are also readily available. In general, the bases themselves can be obtained commercially and are handled like any other phosphoramidite; however, consult the company which supplies the analog about necessary modifications to programs or reagents (see Table 2.11.1). Typically, the only adjustment needed is to dissolve the modified base at a somewhat higher concentration than normal to overcome problems associated with reactivity. Most of the methods used to increase the yield of long and ribo-oligonucleotides may be applied to the synthesis of modified nucleic acids.

When synthesizing modified oligonucleotides, compatibilities of the chemistries, placement of modifications relative to other chemical groups, and 5' to 3' directionality are all factors to consider. Generally when an oligonucleotide is end-labeled/modified, a long flexible tether is added to allow greater accessibility. Stretches of four deoxythymines are often used for this purpose. Also, adding deoxythymines 5' to the label (5'-TTTT-label-3') can aid in separating labeled molecules from unlabeled ones by size. Note that some tagging phosphoramidites allow for the enzymatic extension or kinasing of the modified oligonucleotide, while others do not. Finally, oligonucleotides may also be synthesized directly on solid glass supports (Cohen et al., 1997).

Terminal transferase can be used as an alternative means of incorporating modified bases at the 3' end of an oligonucleotide (Ratliff, 1982; *UNIT 3.6*). This enzyme is tolerant of a variety of substrates, and has been used to add deoxynucleotide triphosphates derivatized at virtually every position (C-8 on adenine, any of the amino groups, C-5 on cytosine, O-6 on guanosine) to DNA. It also functions, though less well, with RNA bases. It can use any DNA oligonucleotide that is at least 2 bases long [d(pXpX)] and contains a free 3' hydroxyl as a primer. A potential problem in preparing homogeneous polynucleotides using terminal transferase is that a statistically random number of bases is added to the 3' end of the template (with the exception of molecules such as cordycepin, which act as chain terminators due to the absence of 3' hydroxyl). If a single species is desired, it can be gel purified (see *UNIT 2.12*). Polynucleotide phosphorylase may also be used to incorporate modified bases at the 3' end (Gillam and Smith, 1980).

A more controlled means of introducing modified nucleotides relies on T4 RNA ligase and substrates of the form A(5')ppX (where X can be virtually any molecule, including for example ribose or amino acids, in a pyrophosphate linkage with adenosine; Uhlenbeck and Gumpert, 1982). The minimal template for reactions of this form is a trinucleoside containing a free 3' hydroxyl. RNA reacts much better than DNA, and single-strand molecules act as better templates than double-stranded ones. Since 3' hydroxyl groups are required, substrates of the form A(5')ppXp will undergo only a single round of addition, unlike the similar reaction with terminal transferase. In some cases, the compound A(5')ppXp can be generated directly by RNA ligase from pXp and ATP, although the substrate requirements for the X moiety are much more strict than in the ligation reaction. Thus, while virtually any dinucleotide of the form A(5')ppX can be added to an oligonucleotide, only a few compounds (primarily sterically "small" derivatives of natural bases) can be used by the enzyme to form A(5')ppXp from pXp.

T4 RNA ligase can catalyze the ligation of single-stranded oligonucleotides in the presence of ATP and various analogs (Kinoshita et al., 1997). Templates prepared by terminal transferase or by T4 RNA ligase that contain modified nucleotides (or other adducts) at their 3' termini may be able to act as substrates in this reaction. This would allow modified nucleotides to be introduced into the middle of a longer chain. However, the substrate specificity of the enzyme for the 3' hydroxyl donor is highly substrate dependent and will have to be determined empirically.

Synthesizing Degenerate Oligonucleotides

Current combinatorial and “irrational” nucleic acid design methodologies focus on the ability to create large pools of random sequences from which useful sequences may be culled (Szostak, 1992). Also, random mutagenesis using degenerate oligonucleotides allows for the exploration of “sequence space” surrounding a given protein or RNA structure. Sequences can be produced that give a completely random distribution of nucleotides at a given position or, alternatively, the sequence can be biased or “doped” toward a particular base with only a low level of randomization.

Most synthesizers can be programmed for in-line degenerate mixing of bases, which is useful if only a few positions are to be randomized. A potential problem with this method is that if mixing is incomplete, the sequence will be skewed toward whichever phosphoramidite enters the column first, since the reaction of the activated phosphoramidite with the free 5' hydroxyl is extremely fast. Therefore, while in-line mixing will generate all base substitutions at a given position, the distribution of these substitutions may not be uniform. If a statistically random distribution of nucleotides is required or if long stretches of random sequence are to be made, it is better to manually mix the phosphoramidites together and use this mixture for the degenerate position. A true random distribution may be obtained by mixing A, C, G, and T phosphoramidites in a 3:3:2:2 molar ratio to compensate for the faster coupling times of G and T phosphoramidites (D. Bartel, pers. comm.). On synthesizers where phosphoramidites are loaded without detaching the bottle, the mixing generated by sequential loading of each phosphoramidite into the extra bottle is sufficient to generate randomized sequences.

Oliphant and Struhl (Oliphant et al., 1986; Oliphant, 1989) have constructed degenerate oligonucleotides using mixed phosphoramidites, but have modified the synthesis protocol by deleting the capping step during the random additions. This increases the overall yield of long oligonucleotides, since sequences that fail to elongate are not terminated, and the size of the final product is more heterogeneous. This method is particularly useful if deletions, as well as randomized bases, in a given sequence are required.

Hermes et al. (unpub. observ.) have developed a detailed protocol for producing statistically mutated oligonucleotides. This method employs in-line mixing between pure phosphoramidite contained in separate bottles and equimolar mixtures of the four bases contained in an additional bottle. The obvious advantage of such a method is that doped and clean sequences can be synthesized on the same oligonucleotide. Whether or not this method is employed, it is an example of how to dissect the chemistry of mixed-site oligonucleotide synthesis. Hermes et al. (1989) have shown that mutations introduced by this method are indeed statistically random.

Regardless of the strategy employed, the level of misincorporation of an oligonucleotide should be decided in advance by the mutagenesis frequency desired. Quantitatively, this level is given by the probability distribution:

$$P(n, m, x) = [m! / (m - n)! n!] [x]^n [1 - x]^{m \times n}$$

where P is the probability of finding n errors in an oligonucleotide m in length with x level of misincorporation (fraction “wrong” nucleotides delivered). This equation describes a Poisson distribution. If primarily single mutations are desired, then x should be maximized for $n = 1$; if multiple mutations (e.g., doubles, or triples in a single oligonucleotide) are necessary, x should be correspondingly higher. If the mix is optimized for n mutations, then $n - 1$ and $n + 1$ mutations will occur in roughly equal amounts and n mutations will be the most frequent.

Cloning randomized oligonucleotides can be difficult, since a complementary wild-type sequence will generate mismatches that may result in biased correction *in vivo*. To avoid this problem, second-strand synthesis can be primed from a nonrandom portion of the sequence, or mutually primed synthesis (UNIT 8.2) can be utilized. Alternatively, Derbyshire et al. (1986) describe the direct cloning of doped single-strand material with “sticky ends” into a double-stranded cloning vector. Finally, Reidhaar-Olson and Sauer (1988) describe the synthesis of complementary oligonucleotides containing inosine (which can pair with any of the four natural bases) directly across from randomized codons. This method resulted in the successful introduction of a wide variety of mutations into the gene for lambda repressor, although there was a slight compositional bias in cloned sequences.

STRATEGIES FOR OLIGONUCLEOTIDE PURIFICATION

Deprotected nucleic acids may be purified and isolated by a variety of methods. The method of choice will depend on the availability of resources, the purity required—some methods cannot separate ($n-1$)-mers from n -mers—and time considerations. Any of the methods described below can be used to clean up crude material.

Isolation Methods

Precipitation

Direct precipitation of the nucleic acids constitutes a fast and efficient purification from contaminating small molecules such as urea and phenol, but does not allow for purification of abortive synthesis products from the full-length one. If oligonucleotide size separation is required, this method should be used in conjunction with some of the other methods described. After deprotection, resuspend the whitish pellet obtained in water. Add $MgCl_2$ to a final concentration of 10 mM and mix along with 5 vol ethanol. Precipitation should be immediate. Freeze the sample briefly at -20° or $-70^\circ C$. Centrifuge the precipitated material, wash with 95% ethanol, dry, and resuspend in water.

Precipitated deoxyoligonucleotides can be used for sequencing or cloning. They can also be used in PCR reactions, although the efficiency of amplification may be reduced as compared to gel-purified oligonucleotides. If the DNA is to be phosphorylated, a more thorough purification procedure is necessary, since T4 polynucleotide kinase is inhibited by lingering ammonium ions.

Sizing columns

Size-exclusion chromatography is extremely useful as a final purification step, especially when small-molecule contamination occurs with otherwise pure oligonucleotides, but (like precipitation) does not effect purification of abortive synthesis products from the full-length one. If oligonucleotide size separation is required, this method should be used in conjunction with some of the other methods described. Oligonucleotides purified by PAGE (UNIT 2.12) might contain small amounts of low-molecular-weight contaminants such as urea or phenol that might inhibit enzymatic reactions; sizing columns are a simple way to decontaminate these samples. Spin columns (UNIT 3.4) are simple to prepare

and use, but gravity-flow columns give better, more reproducible separation. The type of resin used should be adjusted based on the size of the oligonucleotide being purified (e.g., Sephadex G-25 for 25-mers or less, G-50 for longer sequences).

Reversed-phase cartridges

A hydrophobic matrix may be used to separate full-length from abortive synthesis products if the final trityl group is left on following the final monomer coupling reaction. The resulting hydrophobically tagged full-length “trityl-on” oligonucleotide can be separated easily from failure sequences, which lack trityl groups and do not efficiently bind the hydrophobic matrix. Several companies supply columns designed specifically for the purification of “trityl-on” oligonucleotides (e.g., Applied Biosystems oligonucleotide purification cartridges). The procedure is simple and can be performed on a number of samples in parallel within only a few minutes. The yield from such columns is excellent, often >80% of the applied sample. However, although a majority of failure sequences are removed using this method, many shorter sequences are co-purified with the desired full-length material. Some of these fragments are due to cleavage of depurinated DNA. These apurinic molecules can be eliminated prior to cleavage from the column by treatment with lysine (Horn and Urdea, 1988). In addition, if care is not taken to wash and elute samples from these columns slowly, some inhibitors (particularly of ligation reactions) may co-elute.

Denaturing PAGE

Denaturing polyacrylamide gel electrophoresis separates oligonucleotides with single-residue resolution and is the method of choice for purifying full-length oligonucleotides (see *UNIT 2.12*). However, the compatibility of the chemistries of modified nucleotides incorporated into the nucleic acids and acrylamide matrix should be checked (thiolated oligonucleotides seem to undergo Michael addition to the acrylamide, which renders them irreversibly capped).

HPLC

Liquid chromatography can also separate oligonucleotides with single-residue resolution, but its chief advantage is speed. Total run time can be as short as 30 min. The use of alkali perchlorate salts has made ion exchange the HPLC method of choice given that long oligonucleotide (>40 residues) may be easily purified in large scales ($\geq 25 \mu\text{mol}$; Sproat et al., 1995; Warren and Vella, 1994). However, secondary structural migration anomalies are generally more severe than those found with PAGE. Depending on the system employed, the amount of oligonucleotide that can be purified in a single chromatographic run can be comparable to PAGE, and sample recovery is typically >70%. For labs with an HPLC system and a need to routinely purify short oligonucleotides with no secondary structure, this method is ideal.

Oligonucleotides can be purified with HPLC by charge differences through ion exchange or hydrophobicity if the final trityl group is left on following the ultimate monomer coupling reaction. Only the “trityl-on” systems use buffers that can be lyophilized. A more complete treatment of the complexities of oligonucleotide purification by HPLC can be found in Applied Biosystems User Bulletin no. 13 on oligonucleotide purification (Applied Biosystems, 1988).

Confirming the Oligonucleotide Sequence

Most oligonucleotides that are used for cloning need not be checked immediately after synthesis, since the clones themselves will be checked after biological or enzymatic amplification. However, in cases where a sequence will be used in a structural application such as mobility-shift assays (*UNIT 12.1*), filter-binding assays (*UNIT 12.8*), or crystallography,

it is desirable to confirm the sequence. For almost all oligonucleotides, this usually requires chemical sequencing (*UNIT 7.5*). Although sequencing will confirm that the correct product was made, it cannot determine the homogeneity. How unnatural or protected bases will react during chemical sequencing, or how they will affect mobility on a sequencing gel, is not predictable. In order to determine what fraction of molecules contain only the natural bases (A, T, C, and G), it is necessary to digest the DNA enzymatically to completion and to examine its composition by a comparison with standard bases separated by HPLC. At least one HPLC buffer system has been developed specifically for examining modified nucleosides in chemically synthesized oligonucleotides (Eadie et al., 1987). The insertion and deletion rates for chemical nucleic acid synthesis are non-negligible. The rate of insertions (presumably due to DMT cleavage via tertrazole) has been measured to be as high as 0.4% per position, and the rate of deletions (presumably due to incomplete capping) has been found to be as high as 0.5% per position (A. Keefe and D. Wilson, pers. comm.). Therefore, for an oligonucleotide 100 bases in length and assuming a random, noncorrelated mechanism of action, only about 40% of the sequences will be the intended one. Finally, matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS; *UNITS 10.20 & 10.21*), which has recently emerged as a new biotechnology tool, may be used to determine the sequence of deoxy and ribo-oligonucleotides of up to 60 bases (Zhu et al., 1997).

BASIC PROTOCOL 1

MONITORING DNA SYNTHESIS USING THE TRITYL ASSAY

A trityl cation is released from the 5' end of the growing oligonucleotide during each synthesis cycle, and the yield of each step of the synthesis can be determined by spectrophotometrically measuring the amount of trityl cation liberated. This procedure is the simplest and most rapid means available for the identification of problems with synthesized oligonucleotides prior to deprotection and purification. Also, in-line quantitative trityl monitors can be interfaced to most synthesizers (Ana-Gen Technologies).

Materials

0.1 M *para*-toluene sulfonic acid (TSA; monohydrate) in acetonitrile
15-ml glass tubes (graduated, if possible)

1. Collect the trityl cation solution in 15-ml glass tubes after each step.

It is helpful to use graduated tubes so that a uniform final volume of acetonitrile can be attained before assay.

2. Dilute the first three and last three fractions to 10 ml with 0.1 M TSA in acetonitrile. Mix thoroughly.

CAUTION: Handle the solution consisting of dichloroacetic acid (or trichloroacetic acid) in dichloromethane and acetonitrile with gloves, because it is corrosive as well as toxic. Avoid prolonged contact with toxic acetonitrile fumes. Do not pipet by mouth.

Although 2.5 ml of deblocking solution is initially released, because acetonitrile is a volatile liquid the volume may change during the course of a synthesis. Fractions may sit for several days before being assayed without affecting the results. Samples that have evaporated to dryness must be thoroughly redissolved. The acid ensures protonation of the trityl groups, making them more strongly colored. It is misleading merely to visualize the yellow/orange color of the fractions, however, since variable volumes of differing acidity are often released.

3. Dilute each sample 20- to 50-fold in the same solution. Measure the absorbance at 498 nm versus acetonitrile/TSA.

These dilutions are necessary because most spectrophotometers cannot accurately measure absorbances >2.

4. Calculate the stepwise coupling efficiency and absolute yield for the synthesis as a whole. The stepwise efficiency is given by:

$$(\text{stepwise efficiency})^n = \frac{\text{average absorbance of last three fractions}}{\text{average absorbance of first three fractions}}$$

where n is the number of trityl nucleotides in the oligonucleotide (equal to the length of the oligonucleotide for trityl-off syntheses).

The absolute yield of product is given by:

$$\mu\text{mol DMT} = \frac{(\text{absorbance of last fraction}) \times (\text{dilution factor}) \times (10 \text{ ml})}{70 \text{ ml}/\mu\text{mol}}$$

where 70 μmol is the extinction coefficient of DMT.

The average stepwise efficiency is useful in determining the relative efficiency of each cycle of the synthesis. The absolute yield is useful for determining how many milligrams of product are present for subsequent purifications, although not all of the product will be of the desired length.

The average absorbance of the first and last few steps is used to avoid discrepancies in individual trityl assays. Clearly, if the synthesis was performed on a 1- μmol column, the absolute yield in the first few fractions should be close to 1 μmol . A low absolute yield (absorbance) in the first trityl release, followed by higher absolute yields (absorbances) in the next few fractions, is sometimes observed. This is because some spontaneous detritylation occurs during storage of the columns and the trityls are subsequently rinsed off during the prime lines program. In this case, the average stepwise efficiency should be calculated with fractions 2 through 4. It should be noted that visual assessment of the trityl fractions cannot begin to detect subtle differences (<5%) which may be critical in terms of overall yield.

5. Perform troubleshooting assay (see Support Protocol) if average stepwise efficiency is low.

USING THE TRITYL ASSAY FOR TROUBLESHOOTING

If the average stepwise efficiency for the oligonucleotide synthesis is low, each fraction should be assayed in order to diagnose the problem (what counts as a “low” yield depends on the length and quantity of oligonucleotide desired). In general, synthetic efficiency should be >99% per step, although lower stepwise efficiencies can be tolerated for short oligonucleotides (<40 bases) or where yield is not critical.

Three classes of failures can be detected by trityl assays. A low absolute yield at the first step followed by similarly low absorbances that results in a low overall yield is commonly due to inefficient purging of activator or phosphoramidite lines prior to the synthesis. Such a problem frequently occurs when a synthesizer has not been used for several days. Purge the lines with dry reagents prior to starting the run to avoid inefficient synthesis. Most machines have a priming program precisely for this purpose.

If the stepwise efficiency of each step is low, there is a systematic problem with one of the common reagents, such as the acetonitrile. Often this is due to moisture in one or more of the reagents, and it is more likely if reagents have not been recently replaced. The phosphoramidites should be used until they are almost completely exhausted during a series of syntheses, so that fresh chemicals will be diluted as little as possible by older, potentially wet material (see Critical Parameters, Excluding water from solvents).

SUPPORT PROTOCOL

Preparation and Analysis of DNA

2.11.17

Finally, individual trityl assays are most useful in determining when phosphoramidites have become defective. In this case, drops in stepwise efficiency will only be seen at the coupling steps involving the phosphoramidites in question.

Many problems, such as inefficient oxidation or product depurination, cannot be detected by the trityl cation assay. Therefore, the trityl assay procedure should be used in conjunction with HPLC or gel electrophoresis for product analysis, especially if a homogenous population of oligonucleotides is essential.

DEPROTECTION OF DNA OLIGONUCLEOTIDES

After synthesis is complete, the DNA may be cleaved from the column and protecting groups removed by treatment with ammonia. Although very extended treatment in base can harm DNA, hours of ammoniolysis are still preferred to ensure complete deprotection, since a homogeneous population of “natural” oligonucleotides at slightly lower yield is better than a mixed population of partially deprotected, “unnatural” molecules.

To cleave the DNA from the support matrix and remove protecting groups completely, the support bound product must be treated with concentrated ammonia at 55° to 60°C overnight (≥ 12 hr). Even with such extended treatment, deoxyguanosine may not be completely deprotected (Schulhof et al., 1987). Raising the temperature at which oligonucleotides are deprotected has been reported to speed up the process (≥ 5 hr at 70°C).

Phosphoramidites with more labile protecting groups such as phenoxyacetyl or dimethylformamide masking adenosine and guanosine have recently become commercially available (Table 2.11.1). These allow essentially complete deprotection within 30 to 60 min at 70°C (Reddy et al., 1994). Also, by replacing the traditional benzyl protection of cytosine with acetyl and using a 1:1 mixture of aqueous ammonium hydroxide and aqueous methylamine, oligonucleotides synthesized with traditional purine protections may be completely deprotected in 5 min at 65°C (Reddy et al., 1995). Finally, anhydrous ammonia gas-phase deprotection of oligonucleotides has recently been described; this provides a convenient method for parallel deprotection of as many columns will fit in a reactor. Since no water is present, the fully deprotected oligonucleotides remain adsorbed to the column matrix, thereby guaranteeing that no cross-contamination will occur. The oligonucleotides can then be eluted with water and desalted or further purified. Using PAC-protected monomers, the cleavage and deprotection processes can be completed in ~30 min (Boal et al., 1996).

Materials

Concentrated (14.8 N) ammonium hydroxide (see recipe)
Triethylamine
3:1 (v/v) concentrated ammonium hydroxide/ethanol
n-Butanol

Screw-cap plastic vial (preferably fitted with rubber O ring)
Heat block or oven, 55° to 60°C
0.2- μ m filter

1. In a screw-cap plastic vial, suspend the synthesis support matrix or the already support-cleaved oligonucleotide in ~1.0 ml concentrated ammonium hydroxide for a 1- μ mol synthesis.

Depending on the synthesizer or nucleic acid provider, the oligonucleotide may come attached to a support matrix or free in an ammonium hydroxide solution. The volume of ammonium hydroxide in which product is eluted from the synthesizer is variable. The ammonium hydroxide used should not have been diluted by excessive vapor loss.

2. Place the sample in heat block or oven for ≥ 12 hr at 55° to 60°C .
Seal vial tightly with Parafilm (if not fitted with a rubber O ring).
3. After cleavage from the support and deprotection are complete, spin the sample briefly in a tabletop centrifuge to pool the ammonia and support that may have collected in the cap. Let the vial cool to room temperature before opening it to avoid sample boil-over.
4. Filter off the support by passing the liquid through an $0.2\text{-}\mu\text{m}$ filter and wash the filter with 0.3 ml of $3:1$ (v/v) ammonium hydroxide/ethanol.
5. Precipitate the oligonucleotide from the resulting supernatant by adding 10 vol n -butanol and vortexing for 15 sec. Centrifuge 10 min at $16,000 \times g$, 4°C .
6. Remove and discard the single aqueous n -butanol phase to reveal the white oligonucleotide pellet.

Oligonucleotides 20 residues or shorter with good trityl responses are typically suitable for use directly in DNA sequencing, PCR amplification, and gel-shift analysis. However, if more assuredly homogenous material is required, methods such as denaturing PAGE and HPLC may be employed to further purify the full length product.

7. If a yellowish liquid or crusty pellet remains, rather than a white powder, resuspend the pellet in 0.1 ml distilled water and precipitate again with n -butanol as described above.

Generally, it is not necessary to add additional salt for precipitation.

Further extraction will aid in removing any residual ammonia or volatile organics. If the yellow color does not disappear, it will ultimately be removed by almost any of the standard purification methods.

8. Lyophilize the sample to dryness in a Speedvac evaporator.

This deprotection is primarily intended for oligonucleotides with the trityl group off. When the trityl group is left on, care must be taken that it is not prematurely hydrolyzed from the DNA by acid conditions. During lyophilization, a drop of triethylamine must be added regularly to the sample in order to maintain its basicity (Applied Biosystems, 1988). Heating of the samples should be avoided.

DEPROTECTION OF RNA OLIGONUCLEOTIDES

Techniques and recommendations for deprotecting RNA are similar to those for DNA. Ammoniolysis cleaves the RNA from the support and frees the bases of their protecting groups. Additionally, the $2'$ hydroxyl protecting group must be removed to reveal a functional RNA sequence. Reagents, water, and plasticware to which RNA is exposed must be sterile. The protocol detailed below is for use with nucleosides bearing the standard isobutyryl and benzyl protection on the bases and $2'$ silyl protection. Protection options are becoming available that both increase yield and streamline deprotection time (Table 2.11.1).

Materials

100% ethanol
 $3:1$ (v/v) concentrated (14.8 N) ammonium hydroxide/ethanol
 3 M sodium acetate, pH 5.2 (APPENDIX 2)
 Triethylamine trihydrofluoride
 Screw-cap plastic vial (preferably fitted with rubber O ring)
 Heat block or oven, 55° to 60°C
 $0.2\text{-}\mu\text{m}$ filter
 Sephadex G-25 column (Amersham Pharmacia Biotech)

BASIC PROTOCOL 3

Preparation and Analysis of DNA

2.11.19

1. In a screw-cap plastic vial, suspend the synthesis support matrix or the already support-cleaved oligonucleotide in ~1.2 ml of 3:1 (v/v) ammonium hydroxide/ethanol for up to a 1- μ mol scale.

Depending on the synthesizer or nucleic acid provider, the oligo may come attached to a support matrix or free in an ammonium hydroxide solution. The volume of ammonium hydroxide in which product is eluted from the synthesizer is variable. The ammonium hydroxide used should not have been diluted by excessive vapor loss.

2. Place the sample in a heat block or oven for 12 to 16 hr at 55° to 60°C.

Seal vial tightly with Parafilm (if not fitted with a rubber O ring). If fast-cleaving phosphoramidites such as PAC-protected purines and acyl-protected cytosine are used, the deprotection time of the bases can be a little as 10 min in methylamine at 65°C (Wincott et al., 1995).

3. After cleavage from the support and base deprotection are complete, spin the sample briefly in a tabletop centrifuge to pool the ammonia and support that may have collected in the cap. Let the vial cool to room temperature before opening it to avoid sample boil-over.

4. Filter off the support by passing the liquid through an 0.2- μ m filter and wash the filter with 0.3 ml of 3:1 (v/v) ammonium hydroxide/ethanol.

5. Evaporate the combined solutions to dryness in the Speedvac evaporator without heating. Resuspend the pellet in 0.2 ml of 100% ethanol and evaporate to dryness without heating.

6. Treat the dried residue with triethylamine trihydrofluoride (0.3 ml for a 0.2- μ mol or 0.5 ml for a 1- μ mol synthesis). Allow to rotate in a foil-covered screw-cap vial in the dark for at least 24 but no more than 48 hr.

Alternative methods exist to remove the 2' silyl protecting groups either under dilute acidic conditions (Kawahara et al., 1996) or with anhydrous triethylamine/hydrogen fluoride in N-methylpyrrolidinone (Wincott et al., 1995).

7. To desalt via ethanol precipitation, add an equal volume of water to the triethylamine trihydrofluoride solution and immediately dilute with $\frac{1}{10}$ vol of 3.0 M sodium acetate, pH 5.2. Precipitate by adding 3 vol of 100% ethanol and chilling for ~20 min at -80°C. Centrifuge 10 min at 16,000 \times g, 4°C.

8. Remove and discard the single ethanol layer to reveal the white oligonucleotide pellet.

9. If a yellowish liquid or crusty pellet remains, rather than a white powder, resuspend the pellet in 0.1 ml distilled water and repeat the sodium acetate/ethanol precipitation as described in step 7. If desired, desalt the RNA on a desalting matrix such as Sephadex G-25 (UNIT 3.4).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Acetonitrile, dry

Since large volumes are used for rinsing lines and dissolving other reagents, acetonitrile is one of the most costly reagents in nucleic acid synthesis. If the DNA synthesizer is used infrequently, it may be useful to produce small amounts of dry acetonitrile immediately prior to synthesis since the acetonitrile will accumulate water after the bottle has been opened. Moisture greatly diminishes synthetic yield. In this case acetonitrile can be prepared in any laboratory equipped for routine

continued

distillations. However, the time and effort involved in setting up and maintaining a still should be balanced against the cost of obtaining dry acetonitrile. The use of molecular sieving pouches that do not release metal ions is recommended; they are often available from the synthesizer manufacturer and should keep the acetonitrile to <20 ppm water. Some commercial suppliers now market bulk solvents specifically for nucleic acid synthesis (e.g., Baker “low-water” acetonitrile, 0.002% or 20 ppm water; Burdick and Jackson acetonitrile, 0.001% or 10 ppm water). The cost of these special dry reagents is about the same as that of HPLC-grade acetonitrile (which typically contains 0.01% or 100 ppm water), which is the starting material for a laboratory distillation.

CAUTION: *Acetonitrile vapor is poisonous.*

Ammonium hydroxide

Purchase concentrated (14.8 N) ammonium hydroxide in 0.5- to 1.0-liter bottles, and store it at 4°C to reduce the ammonia gas found in the vapor phase above the liquid.

CAUTION: *Concentrated ammonium hydroxide is extremely caustic. Breathing the vapors is harmful. Always use this compound in the fume hood, as it is possible to be quickly overcome by ammonia fumes and blinded.*

Ammonia is volatile and its concentration will decrease on repeated opening, until it is no longer completely functional as a deprotecting reagent. When this occurs, obtain a fresh supply.

Oxidizer

Prepare 0.02 M iodine in 7:2:1 (v/v/v) THF/pyridine/water. Store ≤6 months at room temperature.

CAUTION: *Pyridine is toxic in both liquid and vapor forms.*

CAUTION: *Iodine is harmful if inhaled. Beware of spills containing both iodine and ammonia since explosive compounds can be formed.*

Most commercial sources of THF contain BHT, a free-radical scavenger that prevents the buildup of explosive peroxides. This compound has no effect on synthesis chemistry. Any commercial grade of pyridine is acceptable; use resublimed iodine.

Phosphoramidites

Prepare phosphoramidites in extremely dry, commercially sealed acetonitrile according to the procedure recommended by the synthesizer manufacturer. Store ≤6 months at 4°C and discard after 2 weeks.

Some companies recommend dissolving phosphoramidites to give equal molarities of the four bases, while others recommend a standard weight/volume ratio that will give slightly different final molar concentrations. The actual phosphoramidite concentration may matter for some applications—e.g., when making mixed-site oligonucleotides.

COMMENTARY

Critical Parameters

By optimizing the reagents and protocols used in oligonucleotide synthesis, it is reliably possible to make products of greater length and yield, while minimizing the costs associated with unproductive runs.

Excluding water from solvents

The most critical factor in any synthesis is how the reagents are handled to excluded water

from the system. From the moment a bottle is opened, it is in contact with water in the air, and all the solvents used are hygroscopic and will absorb water vapor, which reduces yields. This problem is so severe that it is advisable to avoid large-scale, lengthy, or important runs on rainy or highly humid days.

Special anhydrous reagents can be purchased from most manufacturers of nucleic acid synthesizers. Additionally, some chemical

suppliers are now beginning to market solvents specifically for DNA synthesis. Adding molecular sieving pouches to the activator and acetonitrile used before oxidation is extremely useful in preserving the anhydrous environment of the phosphoramidite coupling reaction.

Phosphoramidites are the most sensitive to water contamination because they are easily hydrolyzed, which renders them unreactive. Precautions must be taken to avoid exposing them to even small amounts of water. Very dry acetonitrile (<0.003% water), kept as a separate stock and sealed under argon, should be used to dissolve the phosphoramidites. The acetonitrile should be introduced through the septum on the amidite bottles via a syringe. Glass syringes can be dried in a 300°C drying oven, then cooled in a desiccator prior to use. Plastic syringes can be dried in a 45°C vacuum oven. Disposable plastic syringes from air-tight sterile casings are dry enough to use with acetonitrile to be dispensed into phosphoramidites with no additional precautions. The syringe should be filled with argon or helium prior to drawing up the acetonitrile, so that wet air is not introduced into the system. On some synthesizers, this can be done directly via the phosphoramidite ports. Otherwise, an argon line should be used.

An argon line is generally a helpful tool for DNA synthesis chemistry. It consists of an argon tank with a regulator connected to a piece of plastic (Tygon) tubing. The tubing is then fitted with either a Pasteur pipet or a syringe and needle. Gas flow can be roughly determined via an in-line “bubbler” (e.g., Aldrich). Argon from a tank is dry enough so that an in-line desiccator is unnecessary. It is necessary to flush the line for several seconds prior to use. Empty bottles can be dried under an argon stream, which will help to exclude condensation from the air. Partially used reagents should be sealed under an argon layer to prevent equilibration with water vapor in the air (the heavier argon will exclude air from the containers).

Choosing synthesis columns

Automated DNA synthesis generally takes place on a solid support made of controlled-pore glass (CPG). This is a porous, nonswelling particle, 125 to 177 μm in diameter, that is derivatized with a terminal nucleotide attached to a long spacer arm. The accessibility of the growing end of the oligonucleotide chain is influenced by the pore size of the particle. It has been recommended that oligonucleotides up to 50 bases in length should be synthesized

on CPG with $\sim 500\text{-\AA}$ pores, while longer oligonucleotides should be synthesized on CPG with $\sim 1000\text{-\AA}$ or larger pores. Be sure that the column geometry is compatible with the trityl monitor. Also, when packing columns be sure to purge the column first with dry acetonitrile and then with dry argon to ensure that all CPG particles are caught between the filters, since loose particles can damage the synthesizer.

CPG columns are available that contain ~ 0.2 to $10.0\ \mu\text{mol}$ of linked nucleotides. In general larger loadings are used for larger oligonucleotides since as the size of the oligonucleotide increases, the overall yield decreases. Also, subsequent purification steps invariably involve losses, and the amount of product of correct length must be kept high to ensure that there is enough material. Oligonucleotides >35 bases should be synthesized on $1.0\text{-}\mu\text{mol}$ CPG columns. See Synthesizing Long Oligonucleotides for discussion of choosing a column when synthesizing ≥ 100 -base oligonucleotides.

Alternatively, columns that are more heavily derivatized with the 3' terminal nucleotide are available as solid supports (e.g., Fractosil, from Merck), and a higher yield of product per column (and per amount of reagent delivered) can be obtained using these columns. However, Fractosil is not recommended for synthesizing oligonucleotides >20 bases long.

Synthesis chemistry

Regardless of reagent and final product purity, there are inherent limitations in the chemistry used for the synthesis of oligonucleotides. Therefore, it is necessary to understand the differences between chemically synthesized and biologically derived DNA.

Failure to produce “natural” DNA can be due to the synthesis chemistry used. Methylation of some bases can occur during deprotection with thiophenol in methyl phosphoramidite-based syntheses. Under standard synthesis conditions, these methylated bases can account for up to 7% of the nucleotides present, with *N*-3-methyl-dT being the primary modified base (Farrance et al., 1989). However, this problem does not always occur—some such syntheses have been reported to contain 99.9% dA, dG, dT, and dC. In general, this problem can be avoided by increasing the time of the thiophenol deprotection step to between 60 and 90 min.

With both methyl and β -cyanoethyl chemistries, the glycosidic bond of *N*-protected purine phosphoramidites is subject to hydroly-

sis during DMT removal with acid. Such depurination eventually leads to strand cleavage during the ammonia deprotection step. N-protected adenosine is more sensitive to depurination than guanosine, and it is most sensitive when located at the 5' end of an oligonucleotide chain (Tanaka and Letsinger, 1982).

In order to maintain low levels of depurination, the deblocking step should not be longer than ~1 min. Additionally, the mildest effective acid (DCA rather than TCA) should be used. Occasionally, however, protected nucleotides are supplied as monomethoxytrityl (MMT) esters, which are 10-fold more resistant to acidic detritylation. Oligonucleotides synthesized with MMT as a protecting group are more susceptible to depurination, since the deblocking step must be correspondingly lengthened to achieve complete detritylation and high yields. In cases where only one base is being added as the MMT compound, it is recommended that this step be performed manually, so that the nucleotides are subjected to only one lengthy acid treatment.

Advances in oligonucleotide synthesis chemistry may mitigate the problems associated with depurination. Recently, nucleotides derivatized with protecting groups that render synthesized material less sensitive to depurination have become commercially available (Schulhof et al., 1987; Pharmacia, 1989).

Reagents

Most reagents may be purchased from the synthesizer manufacturer or from companies that specialize in reagents for nucleic acid synthesis; however, some labs choose to make their own reagents to reduce costs. Notes on the preparation and storage of certain reagents are provided above. Other reagents should probably be purchased from commercial sources, as the preparation of anhydrous materials is more difficult and expensive than most molecular biology labs can support. An appropriate text should be consulted on the medical dangers of all solutions and reagents used in DNA synthesis (see Key References). Most suppliers of materials have a fairly extensive list of products available on-line.

Troubleshooting

Rescue after reagent depletion

Most synthesizer problems require a visit from the service engineer. However, syntheses during which reagent or solvent bottles run empty can sometimes be rescued. Certain in-

struments respond to depleted reagents during a synthesis by continuing the synthesis or stopping. For instruments that stop, it may be possible to recover oligonucleotides, depending on which reservoir was depleted. It is important that the lines be rinsed after an empty reservoir has been detected (some machines perform this step automatically). Otherwise, the oligonucleotide should be resynthesized. Many machines have the capacity to restart in the middle of a cycle, and if not, the cycle can be finished manually. The exact procedure will vary depending on which reagent has been exhausted.

Acetonitrile. This is quite serious. The lines cannot be rinsed of whatever reagent they last contained; thus, determine whether to continue with the synthesis based on the reactivity of the last reagent in the lines.

Deblocking reagent. Refill the reservoir. The machine can continue at the beginning of the cycle that was interrupted, but full-length yield will greatly suffer.

Phosphoramidites. Refill the reservoir and perform manual coupling. In this way, the chains that did elongate will not form $n + 1$ products, and those that did not will have a chance to elongate. Programmed synthesis can be resumed with the next cycle.

Capping reagent. Perform a manual capping and complete the cycle manually, at which point programmed synthesis can resume. This ensures that capping is efficient for this cycle and avoids accumulation of $n - 1$ sequences.

Oxidizer. Running out of oxidizer is particularly dangerous, because the unstable phosphite linkages remain on the column for long periods. It is best to discard the material; otherwise, perform a manual oxidation and continue with the synthesis.

Ammonium hydroxide. Some material may be left on the column if ammonium hydroxide runs out during the course of synthesis. Refill the bottle and pump more ammonium hydroxide through the column, or remove the column and treat with ammonium hydroxide.

Thiophenol. Refill the thiophenol reservoir. Restart from the deprotection step. At worst, a small amount of product may have cleaved because of inefficient methyl deprotection.

Routine maintenance

To avoid other purely mechanical problems with a DNA synthesizer, create a regular maintenance schedule to perform a few minor maintenance tasks. Change the various frits and filters that remove debris before they enter the system; furthermore, remove the deposits of

salt and debris as these may affect flow rates and decrease volumes delivered. Change the filters on the acetonitrile bottle more often, because the flow of this reagent is much greater than any other. Also, change any rubber O rings every few months. Rinse all the instrument's lines with base and organic solvents thoroughly every 500 hr of machine use or approximately every 3 months. Finally, rinse the lines extensively with dry acetonitrile. Finish the general maintenance by checking the flow rates of the instrument after cleaning and then synthesize a small oligonucleotide. Monitor the trityls carefully in order to confirm that the cleaning did not affect any aspect of the synthesis cycle. If this regimen is followed, many minor delays encountered in the synthesis schedule can be avoided.

Anticipated Results

When the synthesizer is working properly (>98% coupling efficiency), typical yields of deoxyoligonucleotides 20 bases in length at the 1- μ mol scale, as measured from the trityl response, are nearly 60%. Also, these shorter deoxyoligonucleotides are typically suitable for DNA sequencing, PCR amplification, and gel-shift analysis without extensive purification. However, if very homogenous material is required, further purification must be done, because many sequences will be truncated (*UNIT 2.12*). Longer DNA sequences are generally obtained in lesser yield and quality and should be purified. Sequences of >120 bases are generally obtained in \leq 10% yield as measured by the trityl response. Yields for RNA sequences are typically much lower than for their DNA counterparts of a given length, and the products must be extensively purified.

Time Considerations

On a typical day when most of the bottles on the synthesizer need refilling, it may take 30 to 60 min to fill the bottles, empty waste, install columns, and rinse and prime the lines prior to starting a synthesis. The most time-consuming syntheses are those involving doped oligonucleotides, which require a high level of precision in distributing the various phosphoramidites to the appropriate bottles. A computer interfaced with the instrument can increase synthesis accuracy and throughput. Instruments take varying amounts of time to synthesize oligonucleotides depending on the number of columns in use, the length of the oligonucleotide, and the synthesis program. Deprotection and isolation of a standard DNA oligonu-

cleotide will probably take one working day, while RNA oligonucleotides take longer because of the extra deprotection and desalting steps.

Literature Cited

- Applied Biosystems. 1988. User Bulletin no. 13. Applied Biosystems, Foster City, Calif.
- Bartel, D. and Szostak, J.W. 1993. Isolation of new ribozymes from a pool of random sequences. *Science* 261:1411-1418.
- Boal, J.H., Wilk, A., Harindranath, N., Max, E.E., Kempe, T., and Beaucage, S.L. 1996. Cleavage of oligodeoxyribonucleotides from controlled-pore glass supports and their rapid deprotection by gaseous amines. *Nucl. Acids Res.* 24:3115-3117.
- Ciccarelli, R.B., Gunyuzlu, P., Huang, J., Scott, C., and Oakes, F.T. 1991. Construction of synthetic genes using PCR after automated DNA synthesis of their entire top and bottom strands. *Nucl. Acids Res.* 19:6007-6013.
- Cohen, G., Deutsch, J., Fineberg, J., and Levine, A. 1997. Covalent attachment of DNA oligonucleotides to glass. *Nucl. Acids Res.* 25:911-912.
- Derbyshire, K.M., Salvo, J.J., and Grindley, N. 1986. A simple and efficient procedure for saturation mutagenesis using mixed oligodeoxynucleotides. *Gene* 46:145-152.
- Eadie, J.S., McBride, L.J., Efcavitch, J.W., Hoff, L.B., and Cathcart, R. 1987. High-performance liquid chromatographic analysis of oligodeoxyribonucleotide base composition. *Anal. Biochem.* 165:442-447.
- Farrance, I.K., Eadie, J.S., and Ivarie, R. 1989. Improved chemistry for oligodeoxyribonucleotide synthesis substantially improves restriction enzyme cleavage of a synthetic 35mer. *Nucl. Acids Res.* 17:1231-1245.
- Gait, M.J. (ed.). 1984. *Oligonucleotide Synthesis: A Practical Approach*. IRL Press, Washington, D.C.
- Gillam, S. and Smith, M. 1980. Use of *E. coli* polynucleotide phosphorylase for the synthesis of oligodeoxyribonucleotides of defined sequence. *Methods Enzymol.* 65:687-701.
- Hermes, J.D., Parekh, S.M., Blacklow, S.C., Kuster, H., and Knowles, J.R. 1989. A reliable method for random mutagenesis: The generation of mutant libraries using spiked deoxyribonucleotide primers. *Gene* 184:143-151.
- Horn, T. and Urdea, M.S. 1988. Solid supported hydrolysis of apurinic sites in synthetic oligonucleotides for rapid and efficient purification on reverse-phase cartridges. *Nucl. Acids Res.* 16:11559-11571.
- Kawahara, S., Wada, T., and Sekine, M. 1996. Unprecedented mild acid-catalyzed desilylation of the 2-*O*-*tert*-butyldimethylsilyl group from chemically synthesized oligoribonucleotides intermediates via neighboring group participation

- of the internucleotide phosphate residue. *J. Am. Chem. Soc.* 118:9461-9468.
- Kinoshita, Y., Nishigaki K., and Husimi Y. 1997. Fluorescence-, isotope- or biotin-labeling of the 5'-end of single-stranded DNA/RNA using T4 RNA ligase. *Nucl. Acids Res.* 25:3747-3748.
- Oliphant, R. 1989. Functional Sequences from Random DNA. Harvard University Thesis, Boston, Mass.
- Oliphant, R., Nussbaum, A.L., and Struhl, K. 1986. Cloning of random-sequence oligodeoxynucleotides. *Gene* 44:177-183.
- Pharmacia. 1989. *Analects* Vol. 17, no. 2. Pharmacia Biotech, Piscataway, NJ.
- Ratliff, R.L. 1982. Terminal deoxynucleotidyltransferase. In *The Enzymes*, Vol. XV (P.D. Boyer, ed.) pp. 105-118. Academic Press, San Diego.
- Reddy, M.P., Hanna, N.B., and Farooqui, F. 1994. Fast cleavage and deprotection of oligonucleotides. *Tetrahedron Lett.* 35:4311-4314.
- Reddy, M.P., Farooqui, F., Hanna, N.B. 1995. Methylamine deprotection provides increased yield of oligoribonucleotides. *Tetrahedron Lett.* 36:8929-8932.
- Reidhaar-Olson, J.F. and Sauer, R.T. 1988. Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences. *Science* 241:53-57.
- Schulhof, J.C., Molko, D., and Teoule, R. 1987. The final deprotection step in oligonucleotide synthesis is reduced to a mild and rapid ammonia treatment by using labile base-protecting groups. *Nucl. Acids Res.* 15:397.
- Sproat, B., Colonna, F., Mullah, B., Tsou, D., Andrus, A., Hampel, A., and Vinayak, R. 1995. An efficient method for the isolation and purification of oligoribonucleotides. *Nucleosides Nucleotides* 14:255-273.
- Szostak, J. 1992. In vitro genetics. *Trends Biochem. Sci.* 17:89-93.
- Tanaka, T. and Letsinger, R.L. 1982. Syringe method for stepwise chemical synthesis of oligonucleotides. *Nucl. Acids Res.* 10:3249.
- Uhlenbeck, O.C. and Gumpert, R.I. 1982. T4 RNA ligase. In *The Enzymes*, Vol. XV (P.D. Boyer, ed.) pp. 31-58. Academic Press, San Diego.
- Usman, N., Ogilvie, K.K., Jiang, M.Y., and Cederagren, R.J. 1987. Automated chemical synthesis of long oligoribonucleotides using 2'-O-silylated ribonucleotide 3'-O-phosphoramidites on a controlled-pore glass support: Synthesis of a 43-nucleotide sequence similar to the 3' half molecule of an *E. coli* formylmethionine tRNA. *J. Am. Chem. Soc.* 109:7845-7854.
- Vargeese, C., Carter, J., Yegge, J., Karivjansky, S., Settle, A., Kropp, E., Peterson, K., and Peiken, W. 1998. Efficient activation of nucleoside phosphoramidites with 4,5-dicyanoimidazole during oligonucleotide synthesis. *Nucl. Acids Res.* 26:1046-1050.
- Warren, W.J. and Vella, G. 1994. Analysis and purification of synthetic oligonucleotides by high-performance liquid chromatography. *Methods Mol. Biol.* 233-264.
- Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S., and Usman, N. 1995. Synthesis, deprotection, analysis and purification of RNA and ribozymes. *Nucl. Acids Res.* 23:2677-2684.
- Zhu, Y., He, L., Srinivasan, R., and Lubman, D. 1997. Improved resolution in the detection of oligonucleotides up to 60-mers in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry using pulsed-delayed extraction with a simple high voltage transistor switch. *Rapid Commun. Mass Spectrom.* 11:987-992.
- Zon, G., Gallo, K.A., Samson, C.J., Shao, K., Summers, M.F., and Byrd, R.A. 1985. Analytical studies of "mixed sequence" oligodeoxyribonucleotides synthesized by competitive coupling of either methyl or β -cyanoethyl-*N,N*-diisopropylamino phosphoramidite reagents, including 2'-deoxyinosine. *Nucl. Acids Res.* 13:8181-8196.

Key References

Applied Biosystems, 1988. See above.

A well-organized overview of synthetic oligonucleotide synthesis, purification, and quantitation.

Bretherick, L., 1986. Hazards in the Chemical Laboratory, 4th ed. Alden Press, Oxford.

A guide to hazardous chemical handling.

Gait, 1986. See above.

The seminal text on synthetic oligonucleotide synthesis that provides critical insight.

Internet Resources

http://www.interactiva.de/oligoman/intro_c1.html#b1

Web site detailing synthesis chemistries, procedures, and reagents for solid-phase oligonucleotide chemistry.

<http://www.medstv.unimelb.edu.au/ABRFNews/1994/December1994/dec94ponoligo.html>

Web site providing general advice on oligonucleotide synthesis.

Contributed by Andrew Ellington
University of Texas
Austin, Texas

Jack D. Pollard, Jr.
Harvard Medical School and Massachusetts
General Hospital
Boston, Massachusetts

Purification of Oligonucleotides Using Denaturing Polyacrylamide Gel Electrophoresis

UNIT 2.12

**BASIC
PROTOCOL**

Several methods exist for the purification of oligonucleotides following chemical synthesis (UNIT 2.11). The advantages of purification on denaturing polyacrylamide gels are speed, simplicity, and high resolution. These gels can resolve oligonucleotides from 2 to 300 bases long, depending on the percentage of polyacrylamide used (see Table 2.12.1). This method is thus useful for isolating not only chemically synthesized deoxyribonucleotides (UNIT 2.11) but also small RNAs or other single-stranded oligonucleotides. After gel setup, samples are loaded onto a urea-based denaturing gel, separated by electrophoresis, and finally recovered from the crushed gel slice by freeze/thaw and elution.

Materials

Nucleic acid samples

Urea

10× and 1× TBE buffer, pH 8 (APPENDIX 2)

40% acrylamide/2% bisacrylamide (UNIT 7.6)

TEMED (*N,N,N',N'*-tetramethylethylenediamine)

10% ammonium persulfate (APS) in water (store ≤1 month at 4°C)

Urea loading buffer (see recipe)

3 M sodium acetate, pH 5.2 (APPENDIX 2)

TE buffer, pH 7.5 (APPENDIX 2)

Acrylamide gel electrophoresis apparatus

Glass plates, spacers, and combs for pouring gels

50°, 60° and 90°C water baths

DC power supply

Thin-layer chromatography (TLC) plate with fluorescent indicator (e.g., Silica Gel F-254 or IB-F, Merck)

Hand-held short-wave (254-nm) UV lamp

Small-bore (5-ml) syringe (e.g., Becton Dickinson)

15-ml centrifuge tube capable of withstanding high temperatures

Rotary shaker

0.2-μm filter (Gelman Sciences)

Additional reagents and equipment for oligonucleotide deprotection (UNIT 2.11), gel electrophoresis (UNIT 2.7), phenol extraction, and ethanol and butanol precipitation (UNIT 2.1)

Prepare the sample

1. Follow the appropriate deprotection protocol (UNIT 2.11) to prepare the nucleic acid sample for electrophoresis.

Be sure to lyophilize the sample to dryness. The samples will generally appear as an off-white powder following deprotection and lyophilization. If a yellowish liquid or crusty pellet remains, rather than an off-white powder, resuspend the pellet in 0.5 ml distilled water and add $\frac{1}{10}$ vol of 3.0 M sodium acetate, pH 5.2. Add 3 vol of 100% ethanol and chill ~20 min at -80°C to precipitate. Centrifuge 10 min at 16,000 × g, 4°C. Decant and save the supernatant. Wash the pellet in 70% ethanol and lyophilize to dryness.

**Preparation and
Analysis of DNA**

2.12.1

Contributed by Andrew Ellington and Jack D. Pollard, Jr.

Current Protocols in Molecular Biology (1998) 2.12.1-2.12.7

Copyright © 1998 by John Wiley & Sons, Inc.

Supplement 42

Table 2.12.1 Concentrations of Acrylamide Giving Optimum Resolution of DNA Fragments Using Denaturing PAGE^a

Acrylamide (%)	Fragment sizes separated (bases)	Migration of bromphenol blue (bases)	Migration of xylene cyanol (bases)
30	2-8	6	20
20	8-25	8	28
10	25-35	12	55
8	35-45	19	75
6	45-70	26	105
5	70-300	35	130
4	100-500	~50	~230

^aData, from Maniatis et al. (1975), are for single-stranded DNA; RNA will migrate slightly more slowly than DNA of the same sequence and length.

Prepare the gel

2. Assemble the gel-casting apparatus.

Gel spacer and casting systems have been developed to avoid leakage. Those that avoid the need to seal the gel with tape are best, and recently gel casting boots that lack bottom spacers have become available (GIBCO/BRL). Greasing the side/bottom spacers or pouring an agarose plug for the gel is not necessary if some care is taken to ensure that the bottom of the plate assembly is completely sealed. The gel plates should be thoroughly cleaned by washing them with warm soapy water and then rinsing them in 70% ethanol/water. However, if the plates are particularly dirty or if the complete removal of any residual nucleic acids is required, the plates may be soaked in 0.1 M NaOH for 30 min prior to washing. If the gel is particularly thin (<1 mm), silanizing one or both plates (APPENDIX 3B) facilitates post-electrophoretic separation of the gel from the plate.

3. Prepare the gel solution (see Table 2.12.1 for appropriate acrylamide concentrations for resolving single-stranded DNAs). For a denaturing acrylamide gel of 20 cm × 16 cm × 1.6 mm, 60 ml of gel solution is sufficient, made by mixing the following:

25.2 g urea (7 M final)
6 ml 10× TBE buffer
Desired amount of 40% acrylamide/2% bisacrylamide needed for resolution
H₂O to 60 ml final.

CAUTION: Acrylamide is a neurotoxin; always wear gloves, safety glasses, and a surgical mask when working with acrylamide powder.

Commercially prepared polyacrylamide solutions (National Diagnostics) are available and highly recommended since they have long shelf lives and avoid manipulation of the neurotoxic acrylamide powder.

Pick a concentration of acrylamide that will allow the single-stranded nucleic acid to migrate approximately one-half to three-fourths of the way through the gel when the loading dye has reached the bottom of the gel. This allows for good separation of non-full-length and full-length products.

Use a flask that has a wide mouth and a spout for pouring.

4. Heat the mixture to speed its dissolution by immersing the flask in a 60°C water bath or under hot running tap water. Once most of the urea and acrylamide have dissolved, vigorously agitate the solution for ~20 min with magnetic stirring to ensure complete mixing.

5. Add 40 μl TEMED and swirl the flask to ensure thorough mixing. Immediately add 300 μl of 10% APS and mix thoroughly. Pour the acrylamide between the gel plates and insert the comb. Clamp the comb in place at the top of the gel to avoid separation of the gel from the plates as the acrylamide polymerizes. Allow the gel to polymerize for ~ 30 min.

IMPORTANT NOTE: *Polymerization begins as soon as the TEMED and APS are mixed, so all succeeding steps must be performed promptly.*

CAUTION: *Be sure to wear safety glasses while pouring the gel, since it is easy to splash the neurotoxic unpolymerized acrylamide.*

For thick gels, the acrylamide can be poured directly from the mixing flask, but for thinner ones, a large syringe fitted with a needle is useful. By pouring the gel slowly with a tilt of 45° relative to the bench top and starting from one corner, bubbles may largely be avoided. Also, letting the gel polymerize while it is lying flat helps minimize undesirable hydrostatic pressure on the gel bottom.

TEMED may be stored indefinitely at 4°C , but the ability of APS to efficiently initiate the free radical-induced acrylamide polymerization diminishes greatly over time. Make a new APS stock every month and store at 4°C .

Run the gel

6. After polymerization is complete, remove the comb and any bottom spacers from the gel. Wash the gel plates free of spilled acrylamide, and be sure that the spacers are properly seated and clean.
7. Fill the lower reservoir of the electrophoresis tank with $1\times$ TBE. Initially, place the gel into the lower tank at an angle to avoid formation of air bubbles between the plates and the gel bottom. Clamp the gel plates to the top of the electrophoresis tank and fill the upper reservoir with $1\times$ TBE so that the wells are covered.

A syringe with a bent needle may be used to remove air bubbles trapped under the gel, which will disrupt the current flow.

8. Using a DC power supply, prerun and warm the gel for at least 30 min at 20 to 40 V/cm (constant voltage).
9. Add $1\times$ urea loading buffer to the oligonucleotide pellet from step 1 and resuspend it by heating it 5 min at 90°C .

The amount of sample that can be loaded depends on the efficiency of the synthesis reaction. At least 10 μg of material in a single band 2 cm wide is required to cast a clear UV shadow. The longer the oligonucleotide, the less full-length product will be obtained.

Use an amount of loading buffer that is consistent with loading $\sim 25\%$ of a 0.2- μmol synthesis of a 20-mer oligonucleotide per $2\text{ cm} \times 2\text{ cm} \times 1.6\text{ mm}$ well. This will give sharp bands with good resolution. Up to 4-fold more may be added, but the resolution will suffer.

10. Rinse the wells thoroughly with $1\times$ TBE solution immediately prior to gel loading.

The 7 M urea dissolved in the gel will start to diffuse from the wells, creating a dense layer at the bottom of the wells that prevents sample loading and decreases resolution. Rinsing eliminates this problem.

11. Load the samples.

Tracking dyes such as bromphenol blue and xylene cyanol may be added to the samples or in empty lanes to monitor migration (see Table 2.12.1 for migration data).

12. Electrophorese the gel at 20 to 40 V/cm (constant voltage) until the positions of the tracking dye(s) from the loading buffer indicate that the oligonucleotide has migrated one-half to three-fourths of the way through the gel.

The speed of electrophoresis is directly proportional to the voltage gradient across the gel. The current in the circuit and the heat generated for higher-percentage gels (>15% acrylamide) are corresponding smaller; since the increased acrylamide concentration leads to greater resistance. While some heating of the gel during electrophoresis is desirable since it helps to denature the sample, temperatures >65°C should be avoided. All gels should be monitored to make sure that they do not generate so much heat that the plates crack. For example, while a 20% gel can be electrophoresed at 800 V with few problems, an 8% gel run under the same conditions would likely generate too much heat for the apparatus to dissipate.

13. When the oligonucleotide is sufficiently resolved, turn off the power supply and detach the plates from the electrophoresis tank. Pry off the top plate. Cover the gel with plastic wrap (taking care to avoid bubbles and folds) and invert the plate onto a TLC plate with a fluorescent indicator. Using a spatula, peel a corner of the gel away from the plate and onto the plastic wrap. Pry off the remaining plate and place another sheet of plastic wrap on top of the gel.

Recover the oligonucleotide

14. Visualize the bands on the gel by briefly exposing them to short-wave (254 nm) radiation from a handheld lamp. The bands will appear as black shadows on a green background. Outline the bands using a marking pen.

Avoid unnecessarily long UV exposure, which will damage the nucleic acids.

The desired band is generally the darkest one on the gel (excluding material that runs at the dye front); it should also be the slowest-migrating band unless deprotection was incomplete. Lighter bands containing partially protected oligonucleotides, if present, will migrate considerably above the major fully deprotected band. If the stepwise efficiency of the synthesis is low, a smear may be seen instead of a clear band. The top of the smear should be cut out of the gel.

Unpolymerized acrylamide absorbs strongly at 211 nm and may also cause shadowing that is confined to the edges and wells of the gel.

15. Cut out the bands directly with a clean scalpel or razor blade.
16. Chop the gel slabs into fine particles by forcing the gel through a small-bore syringe to aid the diffusion of the oligonucleotide from the matrix. Place the crushed gel slab in a 15-ml centrifuge tube capable of withstanding high temperatures.
17. Add 3 ml TE buffer for every 0.5 ml of gel slab. Freeze the sample for 30 min at -80°C or until frozen solid. Quickly thaw it in a hot water bath (-50°C) and let soak 5 min at 90°C. Elute on a rotary shaker overnight at room temperature.

This freeze/rapid thaw approach (Chen and Ruffner, 1996) greatly decreases elution time and increases yield by allowing ice crystals to break apart the acrylamide matrix. Recovery of a 20-mer oligonucleotide is typically 80% after 3 hr of rotary shaking, making this technique comparable in yield to electroelution (UNIT 2.7).

Since elution is a diffusion-controlled process, using more buffer will aid in elution efficiency. Also, note that longer oligonucleotides will take longer to diffuse from the gel. If speed is essential and high yields are dispensable, enough sample can be obtained for most experiments in only a few hours of extraction. Increasing the temperature to 37°C will also speed the process. Yield may be increased by repeated elutions.

18. Centrifuge the tube 2 min at 1000 × g, room temperature, to pellet the gel fragments. Use a syringe to remove the supernatant, then filter off any remaining acrylamide

fragments by passing the suspension through an 0.2- μ m filter and into a fresh 15-ml centrifuge tube.

19. Concentrate the sample by extracting against 1 vol *n*-butanol. Remove the upper butanol layer and repeat until the volume of the lower, aqueous layer is convenient for precipitation.

About 1/5 vol of the aqueous layer is extracted into the organic butanol layer for every volume of butanol used. If too much butanol is added and the water is completely extracted into the butanol, simply add more water and concentrate again.

20. Add 3.0 M sodium acetate, pH 5.2 (to a final concentration of 0.3 M). Add 2 vol of 100% ethanol to precipitate DNA or 3 vol for RNA. Chill 20 min at -20°C . Pellet the oligonucleotide by centrifuging 10 min at $12,000 \times g$, 4°C .

Do not attempt to precipitate small oligonucleotides (20 bases) in the presence of ammonium ions. If the samples prove refractory to precipitation, use a 1:1 (v/v) mix of ethanol/acetone or 6 vol of acetone for precipitation. A rinse with 95% ethanol will remove undesired salts.

21. Redissolve the oligonucleotide in TE buffer if appropriate.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Urea loading buffer

- 8 M urea
- 20 mM EDTA
- 5 mM Tris·Cl, pH 7.5 (APPENDIX 2)
- 0.5% (w/v) xylene cyanol, bromphenol blue, or both

Add 1 vol loading buffer to sample if a solution, or enough to dissolve a dry sample.

COMMENTARY

Background Information

The traditional alternative to gel purification of oligonucleotides has been high-performance liquid chromatography (HPLC). Although alkali perchlorate salts HPLC systems can achieve very high resolution of small and medium-sized oligonucleotides (<60 bases), electrophoresis provides superior capacity and resolution over a greater range of sizes and is simpler to set up and operate. Separation times using HPLC may be faster (<30 min) than for gels, but the time required for initialization of the system and product workup tend to negate this advantage.

Purification of oligonucleotides on low-pressure reverse-phase cartridges is technically simpler than gel electrophoresis, and faster (<2 hr). However, these cartridges offer no separation of desired product from failed sequences and, if not used properly, allow contamination of the final product with low-molecular-weight compounds that often inhibit subsequent enzy-

matic manipulation of oligonucleotides. For short oligonucleotides synthesized in high yield, very simple purification methods (e.g., gel filtration or ethanol precipitation) are adequate for some applications, such as sequencing or PCR primers, that do not require absolutely homogeneous material.

The high resolution and capacity of polyacrylamide gels makes them the method of choice for the purification of oligonucleotides. Urea disrupts hydrogen bonding between bases and thus allows oligonucleotides to be resolved almost exclusively on the basis of molecular weight as opposed to secondary structure. However, it should be noted that oligonucleotides of equivalent length but different sequence will still migrate slightly differently. Thus, mixed sequences will appear as broader bands than homogeneous sequences (Applied Biosystems, 1984). Also, RNA electrophoreses through the gel more slowly than does DNA of comparable size. Finally, when modified nu-

cleotides have been incorporated into the nucleic acids, the compatibility of their chemistries with that of the acrylamide matrix should be checked before PAGE purification (in particular, oligonucleotides bearing thio groups seem to undergo Michael addition to the acrylamide, which renders them irreversibly capped).

Critical Parameters

For most applications, the separation of oligonucleotides from mononucleotides and protecting groups provides adequate purification. In those cases where separation of oligonucleotides from nearby failure sequences is essential, however, the most critical parameters to be considered are the percentage of acrylamide and the amount of sample loaded. If maximum resolution is desired, then only 50 to 100 μg of material should be loaded per 2 cm \times 2 cm \times 1.6 mm well. The percentages of acrylamide that will give optimal resolution for different-sized oligonucleotides are given in Table 2.12.1, and can also be determined empirically by running a small portion of the starting material on trial gels and staining with ethidium bromide. By running long (20- to 30-cm) gels, oligonucleotides of lengths (n) up to 100 bases can be cleanly separated from the $n - 1$ and $n + 1$ products. If an oligonucleotide contains extensive self-complementary sequences or polyguanosine tracts, it may not be completely denatured in 7 M urea, and thus it may not be possible to separate it cleanly from failed synthesis products. To overcome this difficulty, samples can be electrophoresed on gels containing 20 M formamide instead of urea (Frank et al., 1981).

Troubleshooting

All of the problems that apply to non-denaturing PAGE are relevant here (see *UNIT 2.7*, Commentary). However, most failures in purification will occur because the initial synthesis reaction has been inefficient. In almost all cases, it is better to resynthesize a poor-yielding oligonucleotide than to attempt to isolate a small amount of full-length product from a starting material seriously contaminated with failure sequences. If the oligonucleotide cannot be resynthesized, relatively small amounts of product can be visualized by autoradiography providing the starting material is end-labeled with polynucleotide kinase and [γ - ^{32}P]ATP (the starting material should not contain residual ammonium, which inhibits the enzyme; see

UNIT 3.10). It is convenient to end label a small amount of starting material with radioactive ATP; the remainder should be phosphorylated using nonradioactive ATP so that it will not migrate differently from the labeled tracer. Smaller amounts of starting material should be loaded on thinner (~ 0.75 - to 1.0-mm) gels in narrower lanes (~ 1.0 cm).

Anticipated Results

In general, the yield of purified oligonucleotides from denaturing PAGE decreases as the percentage of acrylamide increases. With crushed gel slices, an average yield of 50% may be expected.

Greater recoveries can be obtained by increasing the volume of elution solution added to the gel slice or by performing serial elutions from the same gel slice. Methods employing more active transfers (e.g., electroelution; Smith, 1980; Vorndam and Kerschner, 1986) may give more efficient recoveries. Also, samples (especially large synthetic RNAs) that prove particularly refractory to elution with aqueous buffers may be eluted easily with 6 vol formamide (for >5 hr at room temperature) followed by a brief (~ 1 hr) elution with an aqueous buffer. Isoamyl alcohol may be used to concentrate the formamide/aqueous buffer extracts to a convenient precipitation volume (J. Urbach, pers. comm.).

Time Considerations

It is usually most convenient to set up and run the gel on one day, elute the oligonucleotide overnight, then phenol extract and ethanol precipitate the sample the following day. However, a deprotected oligonucleotide can be ready for molecular biology applications in as little as 6 hr: setup and polymerization of gel, 1 hr; running of gel, 2 hr; fragment elution, 2 hr; product recovery, 1 hr.

Literature Cited

- Applied Biosystems. 1984. Evaluation and Purification of Synthetic Oligonucleotides. User Bulletin no. 13. Applied Biosystems, Foster City, Calif.
- Chen, Z. and Ruffner, D.E. 1996. Modified crush-and-soak method for recovering oligodeoxynucleotides from polyacrylamide gel. *Biotechniques* 21:820-822.
- Frank, R., Muller, D., and Wolff, C. 1981. Identification and suppression of secondary structures formed from deoxyoligonucleotides during electrophoresis in denaturing polyacrylamide gels. *Nucl. Acids Res.* 9:4967-4979.

- Maniatis, T., Jeffrey, A., and deSande, H.V. 1975. Chain length determination of small double- and single-stranded DNA molecules by polyacrylamide gel electrophoresis. *Biochemistry* 14:3787-3794.
- Smith, H.O. 1980. Recovery of DNA from gels. *Methods Enzymol.* 65:371-379.
- Vorndam, A.V. and Kerschner, J. 1986. Purification of small oligonucleotides by polyacrylamide gel electrophoresis and transfer to diethylaminoethyl paper. *Anal. Biochem.* 152:221-225.

Contributed by Andrew Ellington
University of Texas
Austin, Texas

Jack D. Pollard, Jr.
Harvard Medical School and
Massachusetts General Hospital
Boston, Massachusetts

CHAPTER 3

Enzymatic Manipulation of DNA and RNA

INTRODUCTION

Many of the revolutionary changes that have occurred in the biological sciences over the past few years can be directly attributed to the ability to manipulate DNA in defined ways. The major tools for this genetic engineering are the enzymes that catalyze specific reactions on DNA molecules. This chapter reviews the properties of the principal enzymes that are critical for carrying out most of the important reactions involved in recombinant DNA technology. In addition, it describes protocols for many of the basic techniques such as restriction mapping, radioactive labeling of nucleic acids, and construction of hybrid DNA molecules (cloning). The protocols are described in a manner that includes all the necessary individual components. For convenience, commercial suppliers offer self-contained kits in which many of the components are combined.

From a historical perspective, the discovery of restriction enzymes that cleave DNA at discrete nucleotide sequences was probably the breakthrough that ushered in the rest of the technology. First, the cleavage sites provide specific landmarks for obtaining a physical map of the DNA. Second, the ability to produce specific DNA fragments by cleavage with restriction enzymes makes it possible to purify these fragments by molecular cloning. Third, DNA fragments generated by restriction endonuclease treatment are basic substrates for the wide variety of enzymatic manipulations of DNA that are now possible.

This chapter begins with a protocol for cleaving DNA with restriction endonucleases. In addition to the basic reaction, it includes methods for cleavage with multiple enzymes, for partial digestion of DNA, and for analysis of multiple samples. Structural maps of DNA are generated by restriction mapping that involves digestion with multiple enzymes. This section also contains basic information about all the enzymes that are commercially available. For each enzyme the recognition sequence, type of termini that are produced upon cleavage, buffer conditions, and conditions for thermal inactivation are described.

Another major section of this chapter describes the properties and reaction conditions for other enzymes that are used to manipulate DNA molecules. These include the following: DNA polymerases which synthesize DNA from double-stranded or primed, single-stranded templates; exonucleases which degrade DNA stepwise from the 5' and/or 3' ends; ligases which join double-stranded segments of DNA; RNA polymerases which synthesize RNA from double-stranded DNA templates; phosphatases which remove 5' terminal phosphate residues from nucleic acids; kinases which phosphorylate 5' terminal hydroxyl residues; and several other enzymes that have more specialized applications. The major uses of these enzymes are described, as are some specific protocols for radioactively labeling RNA or DNA.

General techniques for constructing hybrid DNA molecules are presented in the fourth section. Because there are numerous tricks of the trade, this section deals with the general principles for designing the best strategy for any particular cloning experiment. Some specific examples are discussed in more detail, e.g., subcloning restriction fragments via cohesive or blunt ends, directional cloning, joining DNA fragments with incompatible ends, and ligations involving oligonucleotide linkers (*UNIT 3.16*). The polymerase chain

Digestion of DNA with Restriction Endonucleases

Restriction endonucleases recognize short DNA sequences and cleave double-stranded DNA at specific sites within or adjacent to the recognition sequences. Restriction endonuclease cleavage of DNA into discrete fragments is one of the most basic procedures in molecular biology. The first method presented in this unit is the cleavage of a single DNA sample with a single restriction endonuclease (see Basic Protocol). A number of common applications of this technique are also described. These include digesting a given DNA sample with more than one endonuclease (see Alternate Protocol 1), digesting multiple DNA samples with the same endonuclease (see Alternate Protocol 2), and partially digesting DNA such that cleavage only occurs at a subset of the restriction sites (see Alternate Protocol 3). A protocol for methylating specific DNA sequences and protecting them from restriction endonuclease cleavage is also presented (see Support Protocol). A collection of tables describing restriction endonucleases and their properties (including information about recognition sequences, types of termini produced, buffer conditions, and conditions for thermal inactivation) is given at the end of this unit (see Table 3.1.1, Table 3.1.2, Table 3.1.3, and Table 3.1.4).

DIGESTING A SINGLE DNA SAMPLE WITH A SINGLE RESTRICTION ENDONUCLEASE

Restriction endonuclease cleavage is accomplished simply by incubating the enzyme(s) with the DNA in appropriate reaction conditions. The amounts of enzyme and DNA, the buffer and ionic concentrations, and the temperature and duration of the reaction will vary depending upon the specific application.

Materials

- DNA sample in H₂O or TE buffer (*APPENDIX 2*)
- 10× restriction endonuclease buffers (see recipe)
- Restriction endonucleases (Table 3.1.1 and Table 3.1.3)
- 10× loading buffer (*UNIT 2.5*)
- 0.5 M EDTA, pH 8.0 (optional; *APPENDIX 2*)

Additional reagents and equipment for agarose or polyacrylamide gel electrophoresis (*UNIT 2.5* or *UNIT 2.7*), DNA extraction (optional; *UNIT 2.1*), and ethanol precipitation (optional; *UNIT 2.1*)

1. Pipet the following into a clean microcentrifuge tube:

- x μ l DNA (0.1 to 4 μ g DNA in H₂O or TE buffer)
- 2 μ l 10× restriction buffer (Table 3.1.1)
- 18 – x μ l H₂O.

A 20- μ l reaction is convenient for analysis by electrophoresis in polyacrylamide or agarose gels. The amount of DNA to be cleaved and/or the reaction volume can be increased or decreased provided the proportions of the components remain constant.

2. Add restriction endonuclease (1 to 5 U/ μ g DNA) and incubate the reaction mixture 1 hr at the recommended temperature (in general, 37°C).

In principle, 1 U restriction endonuclease completely digests 1 µg of purified DNA in 60 min using the recommended assay conditions. However, crude DNA preparations, such as those made by rapid procedures (UNIT 1.6), often require more enzyme and/or more time for complete digestion (see Critical Parameters). The volume of restriction endonuclease added should be less than 1/10 the volume of the final reaction mixture, because glycerol in the enzyme storage buffer may interfere with the reaction.

3. Stop the reaction and prepare it for agarose or acrylamide gel electrophoresis (UNIT 2.5 or UNIT 2.7) by adding 5 µl (20% of reaction vol) 10× loading buffer.

The reaction can also be stopped by chelating Mg²⁺ with 0.5 µl of 0.5 M EDTA (12.5 mM final concentration). If the digested DNA is to be used in subsequent enzymatic reactions (e.g., ligation or “filling-in” reactions), addition of EDTA should be avoided. Alternatively, many enzymes can be irreversibly inactivated by incubating 10 min at 65°C (see Table 3.1.1). Some enzymes that are partially or completely resistant to heat inactivation at 65°C may be inactivated by incubating 15 min at 75°C. When the enzyme(s) is completely resistant to heat inactivation, DNA may be purified from the reaction mixture by extraction with phenol and precipitation in ethanol (UNIT 2.1).

Alternatively, DNA may be purified conveniently using a silica matrix suspension as described in UNIT 2.1 (also commercially available as GeneClean from Bio101.)

ALTERNATE PROTOCOL 1

DIGESTING DNA WITH MULTIPLE RESTRICTION ENDONUCLEASES

It is often desirable to cleave a given DNA sample with more than one endonuclease. Two or more enzymes may be added to the same reaction mixture if all are relatively active in the same buffer and at the same temperature. Many enzymes are active in a wide variety of buffer solutions. It is frequently possible to choose a standard buffer solution in which two or more enzymes will retain activity (see Commentary and Table 3.1.2). Alternatively, most restriction endonucleases and some DNA-modifying enzymes are active to some extent in potassium glutamate- and potassium acetate-based buffers (see recipe for 10× restriction endonuclease buffers). Hence, these buffers may be useful for digesting DNA with multiple enzymes. However, if the reaction conditions needed are too dissimilar, follow the procedure below.

1. Digest the DNA with the enzyme(s) that is active at the lower NaCl concentration (see Basic Protocol, steps 1 and 2).

If optimal digestion conditions differ only in incubation temperature, cleave the DNA with one enzyme, then shift the temperature and add the second enzyme (the order of cleavage does not matter; Table 3.1.1). However, many enzymes with optimal activity at high temperatures are also active at 37°C. In these cases, the enzymes can be added simultaneously and the reaction mixture incubated at 37°C (it may be necessary to add more of the “high-temperature” enzyme than usual).

2. For enzymes active at higher salt concentrations, add 1 M NaCl (1 to 3 µl for a 20-µl reaction) so that the final concentration is suitable for digestion by the next enzyme(s). Add enzyme(s) for the second reaction and incubate appropriately.

Purification of the DNA fragments between digestions is the most reliable method to ensure complete, multiple digestions. However, it is much more laborious and is rarely necessary.

3. Stop the reaction for electrophoretic analysis or further enzymatic treatment (see Basic Protocol, step 3).

Digestion of DNA with Restriction Endonucleases

3.1.2

DIGESTING MULTIPLE SAMPLES OF DNA

This procedure minimizes the number of pipetting steps when multiple samples are to be digested with the same enzyme(s) and, hence, saves time. More importantly, by minimizing the number of transfers from the tube containing the restriction enzyme, the potential for contamination of the enzyme is reduced.

1. For each sample to be tested, add a constant *volume* of DNA to a separate microcentrifuge tube.

It is critical to use a different pipet tip for each DNA sample in order to prevent cross-contamination.

2. Prepare a “premix solution” containing sufficient 10× restriction endonuclease buffer and water for digesting all the samples. Place solution on ice.

For example, if ten 3- μ l samples of DNA are each to be digested in a 20- μ l reaction mixture, the premix will contain 20 μ l of 10× restriction buffer and 150 μ l water. It is prudent to make up enough solution for at least one more sample than is to be tested.

3. Add sufficient restriction endonuclease(s) for digesting all the samples. Mix quickly by flicking the tube and replace on ice.

The solution to which the enzyme is added should not be more concentrated than 3× buffer.

4. Add the appropriate amount of solution containing the restriction endonuclease (17 μ l for above example) to each tube of DNA and incubate the reactions 1 hr at the appropriate temperature.

For most analytical purposes, the same pipet tip can be used to dispense the restriction endonuclease solution provided that care is taken to avoid direct contact with the DNA at the bottom of the tubes. For preparative purposes, it is advisable to use a different pipet tip for each sample.

5. Stop the reactions for electrophoretic analysis or further enzyme digestion (see Basic Protocol, step 3).

PARTIAL DIGESTION OF DNA WITH RESTRICTION ENDONUCLEASES

For some purposes, it is useful to produce DNA that has been cleaved at only a subset of the restriction sites. This is particularly important for cloning segments of DNA in which the site(s) used for cloning is also present internally within the segment, and is also useful for restriction mapping (UNIT 3.3). Partial cleavage is accomplished using reduced (and therefore limiting) concentrations of restriction enzyme. Here a set of serial dilutions of enzyme is set up so that one or more of the conditions used is likely to produce the appropriate partial digestion product. If necessary, the procedure can be used for analytical purposes, and the reaction can be repeated and/or scaled up using digestion conditions that have been empirically determined. Partial cleavage can also be achieved by varying the time of digestion, but this is more tedious and less reliable.

1. Make up a 100- μ l reaction mixture containing DNA in 1× restriction enzyme buffer.
2. Divide up reaction mixture such that tube 1 contains 30 μ l, tubes 2 to 4 contain 20 μ l, and tube 5 contains 10 μ l. Place tubes on ice.
3. Add the selected restriction endonuclease (3 to 10 U/ μ g DNA) to tube 1, mix quickly by flicking the tube, and place the tube back on ice.

**ALTERNATE
PROTOCOL 2**

**ALTERNATE
PROTOCOL 3**

**Enzymatic
Manipulation of
DNA and RNA**

3.1.3

- Using a different pipet tip, add 10 μl from tube 1 into tube 2, mix quickly, and place back on ice. Continue the serial dilution process by successively pipetting 10 μl from tube 2 to 3, 3 to 4, and 4 to 5. When finished, all five tubes should contain 20 μl and be on ice.

Figure 3.1.1 illustrates the serial dilution process.

It is critical to use a new pipet tip for each dilution step.

- Incubate all five tubes for 15 min at the appropriate temperature for the restriction endonuclease and stop the reactions for electrophoretic analysis or further enzymatic treatment (see Basic Protocol, step 3).

By virtue of the serial dilution process, the amount of enzyme per microgram of DNA has been varied over a 54-fold range. The extent of digestion is determined by gel electrophoresis of the samples. For most purposes, the desired partial digestion products can be obtained directly.

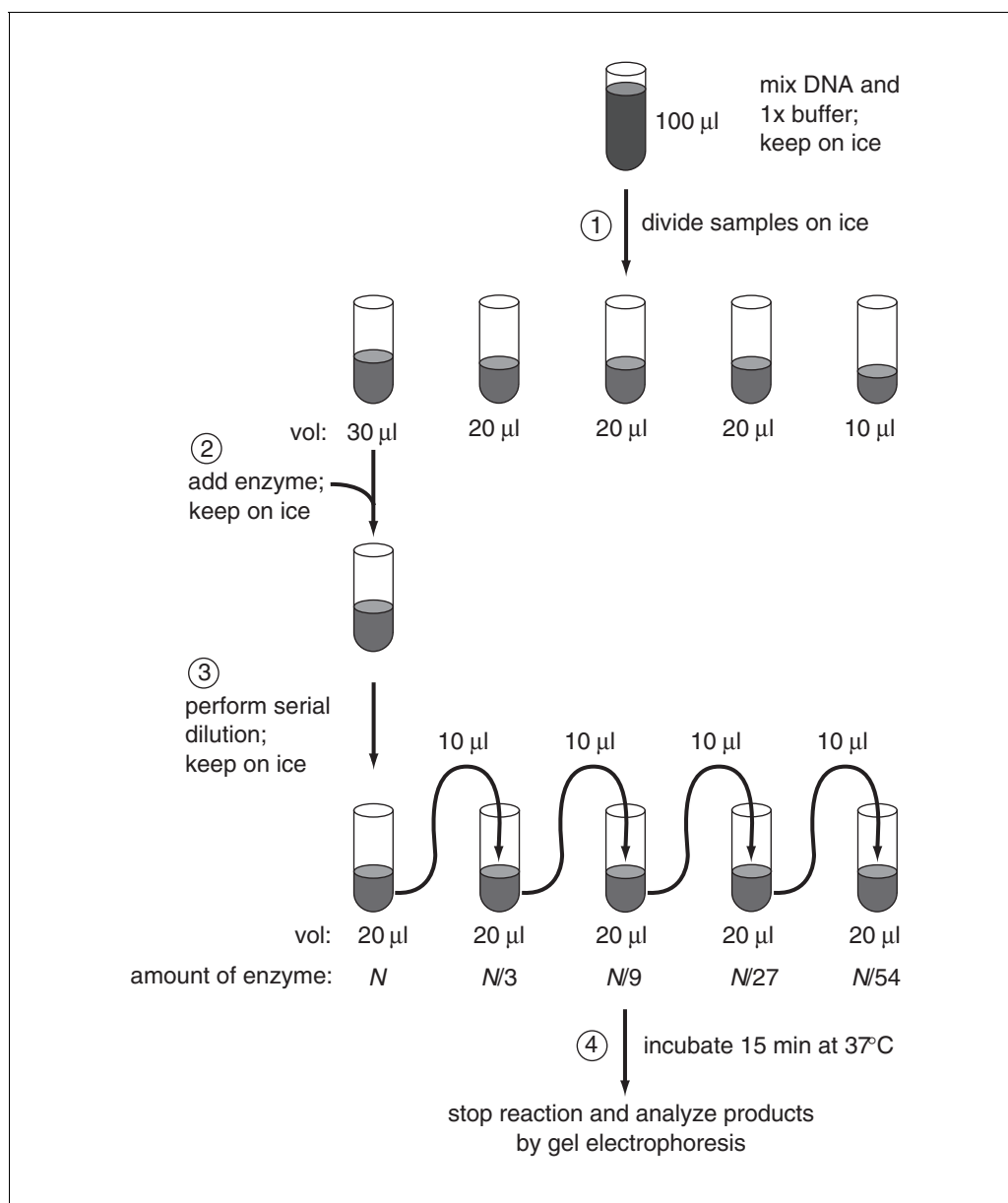


Figure 3.1.1 Partial digestion by serial dilution.

METHYLATION OF DNA

A number of commercially available methylases covalently join methyl groups to adenine or cytosine residues within specific target sequences (i.e., the *EcoRI* methylase methylates an A residue within the *EcoRI* recognition sequence). Methylation of these sites renders them resistant to cleavage by the corresponding restriction endonuclease.

Additional Materials (also see Basic Protocol)

10× methylase buffer (see recipe)
S-adenosylmethionine (SAM)

1. Set up a reaction (typically 20 μ l) containing DNA (at a final concentration of ~20 to 200 μ g/ml) in 1× methylase buffer (Table 3.1.3).
2. Add SAM to achieve a final concentration of 80 μ M.

SAM is the methyl group donor.

3. Add a sufficient amount of methylase to completely protect the DNA from cleavage by the corresponding restriction endonuclease.

A unit of methylase protects 1 μ g bacteriophage λ DNA under the recommended conditions from cleavage by the corresponding restriction endonuclease.

4. Incubate the reaction mixture 1 hr at the appropriate temperature (usually 37°C).

Methylase reactions frequently contain EDTA, which may inhibit subsequent endonuclease reactions. Methylated DNA may be purified from the reaction mixture by extraction with phenol/chloroform and precipitation in ethanol (UNIT 2.1).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Methylase buffer, 10×

NaCl
Tris·Cl, pH 7.5 (APPENDIX 2)
EDTA
2-mercaptoethanol (2-ME) or dithiothreitol (DTT)
S-adenosylmethionine (SAM)

The concentrations of buffer components depend upon the methylase (Table 3.1.3)

Restriction endonuclease buffers

10× sodium chloride–based buffers
100 mM Tris·Cl, pH 7.5 (APPENDIX 2)
100 mM MgCl₂
10 mM dithiothreitol (DTT)
1 mg/ml bovine serum albumin (BSA)
0, 0.5, 1.0, or 1.5 M NaCl

The concentration of NaCl depends upon the restriction endonuclease (Table 3.1.1). The four different NaCl concentrations listed above are sufficient to cover the range for essentially all commercially available enzymes except, those requiring a buffer containing KCl instead of NaCl (see recipe below). Autoclaved gelatin (at 1 mg/ml) can be used instead of BSA. See Critical Parameters.

continued

Note that most restriction enzymes are provided by the suppliers with the appropriate buffers.

10× potassium acetate–based buffer

660 mM potassium acetate
330 mM Tris·acetate, pH 7.9
100 mM magnesium acetate
5 mM DTT
1 mg/ml BSA (optional)

10× potassium chloride–based buffer

60 mM Tris·Cl, pH 8.0 (APPENDIX 2)
60 mM MgCl₂
200 mM KCl
60 mM 2-mercaptoethanol (2-ME)
1 mg/ml BSA

2× potassium glutamate–based buffer

200 mM potassium glutamate
50 mM Tris·acetate, pH 7.5
100 µg/ml BSA
1 mM 2-ME

COMMENTARY

Background Information

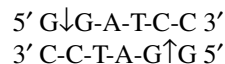
Restriction endonucleases are enzymes that cleave DNA in a sequence-dependent manner. Such cleavage is used for a wide number of applications in molecular biology, including (a) establishment of an endonuclease map of a plasmid or bacteriophage clone; (b) fragmentation of genomic DNA prior to electrophoretic separation and Southern blotting; (c) generation of fragments that can be subcloned in appropriate vectors; and (d) generation of fragments to be used as labeled probes in both Southern (UNIT 2.9) and northern (UNIT 4.9) blotting, as well as in nuclease protection analysis (UNIT 4.7).

Members of a large subgroup of these enzymes (type II restriction endonucleases) recognize short nucleotide sequences and cleave double-stranded DNA at specific sites within or adjacent to the sequences. The recognition sequences are generally, but not always, 4 to 6 nucleotides in length and are usually characterized by dyad symmetry (the 5' to 3' nucleotide sequence of one DNA strand is identical to that of the complementary strand). Two enzymes, *NotI* and *SfiI*, recognize an 8-bp sequence found infrequently in genomic DNA and hence are extremely useful for cleaving DNA into large fragments. More than 600 type II restriction endonucleases have been isolated.

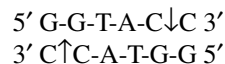
Some type II restriction endonucleases cleave at the axis of symmetry, yielding “flush” or “blunt” ends. Others make staggered cleavages, yielding overhanging single-stranded 3'

or 5' ends known as cohesive termini or “sticky ends.”

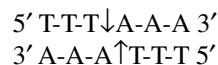
*Bam*HI cleavage generates cohesive 5' overhanging ends:



*Kpn*I cleavage generates cohesive 3' overhanging ends:

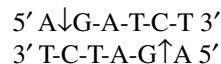


*Dra*I cleavage generates blunt ends:



Endonucleases with the same recognition sequences are called isoschizomers. Isoschizomers may or may not cleave DNA identically to produce the same ends. Conversely, two or more restriction endonucleases recognizing identical or different sequences may generate identical DNA fragment termini (Table 3.1.4). These endonucleases are said to produce compatible ends. DNA fragments with compatible termini may be ligated with DNA ligases to produce hybrid DNA molecules (UNIT 3.16).

For example, *Bgl*II recognizes a different six-nucleotide sequence from *Bam*HI, but generates cohesive termini compatible with those of *Bam*HI. *Bgl*II cleaves:



The hybrid sites generated by joining *Bam*HI and *Bgl*II cohesive ends cannot be cleaved by either enzyme.

Restriction endonucleases, generally found in prokaryotic organisms, are probably important for degrading foreign DNA (particularly bacteriophage DNA). Organisms that produce restriction endonucleases protect their own genomes by methylating nucleotides within the endonuclease recognition sequences. A specific methylase covalently links methyl groups to adenine or cytosine nucleotides within target sequences, thus rendering them resistant to cleavage by the restriction enzyme.

Several sources provide comprehensive and up-to-date listings of restriction endonucleases, including their restriction sites, isoschizomers, and reaction conditions. One source is the catalogs of commercial suppliers of enzymes. The *Biotech Buyers' Guide* (ACS, 1995) provides free of charge an annually updated list of restriction enzymes and their suppliers. Another comprehensive source listing all known restriction endonucleases is a special annual supplement to *Nucleic Acids Research* (e.g., Roberts, 1994). This provides an additional listing of all commercial suppliers of each enzyme. The most complete and up-to-date version of Roberts' information is maintained as a text file database called REBASE, available online from a variety of electronic mail and network servers (see UNIT 19.1 for a complete listing of sources).

Critical Parameters

Purity of DNA

The efficiency of the restriction endonuclease reaction is very dependent upon the purity of the DNA. Contaminants found in some DNA preparations (e.g., protein, phenol, chloroform, ethanol, EDTA, SDS, high salt concentration) may inhibit restriction endonuclease activity. Such impurities are often present in DNA samples prepared by miniprep procedures (UNIT 1.6). The decreased reaction efficiency associated with impure DNA preparations may be overcome by increasing the number of enzyme units added to the reaction mixture (up to 10 to 20 U per microgram DNA), increasing the reaction volume to dilute potential inhibitors, or increasing the duration of incubation. Some preparations of DNA (particularly minipreps) are contaminated by DNases. Because DNases require Mg²⁺ for enzyme activity, DNA in such preparations is stable in its storage buffer (which contains EDTA), but is rapidly degraded upon addition

of restriction endonuclease buffer. This problem can be overcome only by repurifying the DNA.

Digestion of genomic DNA (prepared as in UNITS 2.2-2.4) can be facilitated by the addition of the polycation spermidine (final concentration 1 to 2.5 mM), which acts by binding negatively charged contaminants. However, as spermidine will precipitate DNA at 4°C, it should be added after the other components of the reaction mixture have been incubated at the appropriate temperature for a few minutes. Finally, some preparations of DNA require repurification (phenol and chloroform extractions and ethanol precipitation) prior to digestion with restriction enzymes.

Degree of methylation

Some restriction endonucleases are inhibited by methylation of nucleotides within their recognition sequences (Table 3.1.1). In general, *Escherichia coli* host strains from which plasmids are harvested contain two nucleotide-sequence-specific methylases: *dam*, which methylates adenine in the sequence GATC, and *dcm*, which methylates the internal cytosine residue in the sequences CC(A/T)GG (Table 3.1.3). Thus, plasmid DNA from normal strains may be cleaved partially or not at all by restriction endonucleases that are sensitive to methylation. This can be avoided by preparing plasmid DNA from strains that lack these methylases (UNIT 1.4).

Mammalian DNA contains occasional 5-methylcytosine residues, usually at the 5' side of guanosine residues. The degree of methylation varies from site to site, and is strongly influenced by the cell type from which the DNA is isolated. Methylation patterns in eukaryotic genomic DNA can be investigated by using the different methylation sensitivities of isoschizomers. For example, *Msp*I cleaves CCGG even when the internal cytosine is methylated, whereas *Hpa*II, which also cleaves CCGG, is very sensitive to such methylation.

In some situations, it is useful to take advantage of a restriction enzyme's inability to cleave methylated nucleotide sequences. Methylases recognizing sequences close to a restriction enzyme's recognition sequences can inhibit cleavage at those sites, thereby altering the enzyme's apparent sequence specificity. Alternatively, when using synthetic linkers to modify the termini of a DNA fragment (UNIT 3.16), it may be important to protect internal restriction enzyme sites by methylation prior to enzyme cleavage of linkers.

Other factors influencing DNA cleavage

Larger amounts (up to 20-fold more) of some enzymes are necessary to cleave supercoiled plasmid or viral DNA as compared to the amount needed to cleave linear DNA (Fuchs and Blakesley, 1983). In addition, some enzymes cleave their defined sites with different efficiency, presumably due to differences in flanking nucleotides. In general, cleavage rates for different sites recognized by a given enzyme differ by less than a factor of 10. Although such variability is usually irrelevant, it can be significant in experiments involving partial digestion. It may be difficult to cleave DNA at a particular site without extensive cleavage at other sites. A few restriction endonucleases such as *NarI*, *NaeI*, *SacII*, and *XmaIII* show extreme variability such that some sites are very difficult to cleave.

Buffer conditions

The typical restriction endonuclease buffer contains magnesium chloride, sodium or potassium chloride, Tris-Cl, 2-mercaptoethanol (2-ME) or dithiothreitol (DTT), and bovine serum albumin (BSA). A divalent cation, usually Mg^{++} , is an absolute requirement for enzyme activity. Buffer, typically Tris-Cl, is necessary to maintain the optimal pH for enzyme function. Sulfhydryl reagents may be useful for stabilization of some restriction enzymes, but may also stabilize potential contaminants. Some restriction endonucleases are very sensitive to the concentration of sodium or potassium ion, while others are active over a wide range of ionic strengths.

For each restriction endonuclease, optimal reaction conditions are recommended by the manufacturer. However, strict adherence to these recommendations would require the investigator to stock a large number of buffers. Because many enzymes retain most of their activity over a wide range of reaction conditions (Table 3.1.2), most manufacturers have begun to recommend a panel of four buffers for use with their restriction endonucleases. These buffers may be provided with the purchase of enzymes or may be purchased separately, or they may be prepared in the laboratory and stored at $-20^{\circ}C$, as 10 \times concentrates, for more than a year. The buffer systems typically include a core buffer with varying NaCl concentrations, pH, and specific ion requirements. Manufacturers' catalogs include descriptions of these buffers and indicate enzyme activities in each buffer. Unfortunately, these buffers

have not been standardized among the various companies; different manufacturers may recommend different buffers for the same endonuclease. In general, we recommend purchasing restriction endonucleases from a single company based upon price and using the buffers that are provided with the enzymes.

As an alternative to the four-buffer system described above, many laboratories use potassium glutamate-based (McClelland et al., 1988) and potassium acetate-based (O'Farrell et al., 1980) buffers (known as "universal" buffers) for digestion of DNA with restriction endonucleases. The key features of these buffers is that they include potassium glutamate or potassium acetate instead of sodium chloride and Tris-acetate instead of Tris-Cl. Most restriction endonucleases are active in these buffers, although some enzymes are less active (sometimes only 20%) than under optimal conditions specified by the manufacturer. Several DNA-modifying enzymes are also active in these buffers, including T4 DNA polymerase and T4 DNA ligase. "Universal" buffers are useful for digestion of DNA with multiple restriction endonucleases, particularly when the endonucleases are incompatible in any one of the standard buffers.

Some restriction endonucleases "relax" their recognition sequence specificity in "nonoptimal" reaction conditions (including high endonuclease concentrations, high glycerol concentrations, low ionic strength, Mn^{2+} instead of Mg^{2+} , and high pH). This "star" activity cleaves DNA at other sites besides those containing the "correct" sequence. For example, *EcoRI* star activity cleaves some but not all sequences of the form AATT (usually sites with a 5 out of 6 match to GAATTC are cleaved better than sites with a 4 out of 6 match). The cleavage products all contain cohesive ends identical to those generated by the true *EcoRI* activity.

Additional comments

Enzymes should be stored at $-20^{\circ}C$. While in use, enzymes should be carefully maintained on ice.

Be extremely careful to avoid contaminating enzyme solutions, particularly with plasmid DNA, other restriction endonucleases, or DNase I.

Some restriction enzymes are expensive, others relatively inexpensive. Their use is dictated in part by their cost. Ironically, expensive enzymes are often of lower quality.

Anticipated Results

Complete cleavage of the DNA by a restriction endonuclease should generate a set of discrete DNA fragments that are bounded by the restriction sites. Upon analysis by gel electrophoresis (UNITS 2.5 & 2.7), the cleavage products should be visualized as sharp bands.

Time Considerations

In principle, if 1 U of restriction endonuclease digests 1 μ g of DNA in 1 hr, then incubations for longer durations might be expected to permit conservation of expensive enzymes. In practice, however, some restriction endonucleases have only limited stability in the reaction mixture. Thus, enzyme reactions are usually carried out for 30 min to 2 hr unless very large amounts of DNA or expensive enzymes are involved. In addition, be aware that extended incubations often reveal low levels of contaminating nuclease activities, which may confound experimental results.

Literature Cited

- ACS (American Chemical Society). 1995. *Biotech Buyers' Guide 1995*. ACS, Washington, D.C.
- Fuchs, R. and Blakesley, R. 1983. Guide to the use of type II restriction endonucleases. *Methods Enzymol.* 100:3-38.
- McClelland, M., Hanish, J., Nelson, M., and Patel, Y. 1988. KGB: A single buffer for all restriction endonucleases. *Nucl. Acids Res.* 16:364.
- O'Farrell, P.H., Kutter, E., and Nakanishe, M. 1980. *Mol. Gen. Genet.* 179:411-435.
- Roberts, R.J. 1994. Restriction enzymes and their isoschizomers. *Nucl. Acids Res.* 20 (Supp.#1):2167-2180.

Contributed by Kenneth D. Bloch
Massachusetts General Hospital
Boston, Massachusetts

Barbara Grossmann (Table 3.1.1)
Amersham Life Science, Inc.
Cleveland, OH

Tables appear on following pages.

Table 3.1.1 Recognition Sequences and Reaction Conditions of Commercially Available Restriction Endonucleases^a

Name	Site ^b	Salt ^c	Rxn. temp. ^d	Inact. temp. ^d	Isoschizomers	Comments ^e
Restriction endonucleases						
<i>AatI</i>	AGG↓CCT	M	37	75	<i>Eco147I, StuI</i>	<i>dcm</i>
<i>AatII</i>	GACGT↓C	M(K)	37	60		
<i>AccI</i>	GT↓(A/C)(G/T)AC	— ^f	37	90		↑ activity at 55°C
<i>AccII</i>	CG↓CG	L	37		<i>Bsh1236I, BstUI, MvnI, ThaI</i>	
<i>AccIII</i>	T↓CCGGA	H	60		<i>BseAI, BsiMI, BspEI, Kpn2I, MroI</i>	
<i>Acc65I</i>	G↓GTACC	H(K)	37	65	<i>Asp718I, KpnI</i>	
<i>AccB7I</i>	CCAN ₄ ↓NTGG	M	37		<i>PflMI, Van91I</i>	
<i>AciI</i>	C↓CGC GGC↑G	H	37	65		
<i>AcsI</i>	Pu↓AATTPy	— ^f	50		<i>ApoI</i>	
<i>AcyI</i>	GPu↓CGPyC	H	50	65	<i>BbiII, BsaHI, Hin1I, Hsp92I</i>	
<i>AfaI</i>	GT↓AC	M	37	70	<i>Csp6I, RsaI</i>	
<i>AflII</i>	C↓TTAAG	M	37	65	<i>BfrI, Bst98I</i>	
<i>AflIII</i>	A↓CPuPyGT	H	37	100		
<i>AgeI</i>	A↓CCGGT	L	25		<i>PinAI</i>	
<i>AhdI</i>	GACN ₃ ↓N ₂ GTC	— ^f	37	65	<i>AspEI, Eam1105I, EclHKI</i>	
<i>AluI</i>	AG↓CT	M	37	70		
<i>AlwI</i>	GGATCN ₄ ↓ CCTAGN ₅ ↑	— ^f	37	65		<i>dam</i>
<i>Alw21I</i>	G(A/T)GC(A/T)↓C	L	37	65	<i>AspHI, BsiHKAI, HgiAI</i>	
<i>Alw26I</i>	GTCTCN↓ CAGAGN ₅ ↑	L	37	65	<i>BsmAI</i>	
<i>Alw44I</i>	G↓TGCAC	M	37	65	<i>ApaLI</i>	
<i>AlwNI</i>	CAGN ₃ ↓CTG	— ^f	37	65		
<i>AocI</i>	CC↓TNAGG	L	37		<i>Bsu36I, CvnI, Eco81I, MstII, SauI</i>	
<i>Aor51HI</i>	AGC↓GCT	M	37		<i>Eco47III</i>	
<i>ApaI</i>	GGGCC↓C	L	37	65	<i>Bsp120I</i>	<i>dcm</i>
<i>ApaLI</i>	G↓TGCAC	L	37		<i>Alw44I</i>	
<i>ApoI</i>	Pu↓AATTPy	H	50		<i>AcsI</i>	
<i>AscI</i>	GG↓CGCGCC	— ^f	37	65		
<i>AseI</i>	AT↓TAAT	H(K)	37	65	<i>AsnI, VspI</i>	Star activity
<i>AsnI</i>	AT↓TAAT	H	37		<i>AseI, VspI</i>	
<i>AspI</i>	GACN↓N ₂ GTC	H	37		<i>Tth111I</i>	
<i>Asp700I</i>	GAAN ₂ ↓N ₂ TTC	H	37		<i>XmnI</i>	
<i>Asp718I</i>	G↓GTACC	H	37		<i>Acc65I, KpnI</i>	
<i>AspEI</i>	GACN ₃ ↓N ₂ GTC	L	37	65	<i>AhdI, Eam1105I, EclHKI</i>	
<i>AspHI</i>	G(A/T)GC(A/T)↓C	H	37		<i>Alw21I, BsiHKAI, HgiAI</i>	
<i>AsuI</i>	G↓GNCC	M	37		<i>BsiZI, Bsu54I, Cfr13I, Sau96I</i>	
<i>AvaI</i>	C↓PyCGPuG	M	37	100	<i>BcoI, NspIII</i>	↑ activity at 45°C
<i>AvaII</i>	G↓G(A/T)CC	M	37	65	<i>Eco47I, SinI</i>	<i>dcm</i>
<i>AviII</i>	TGC↓GCA	H	37		<i>FdiII, FspI, MstI</i>	
<i>AvrII</i>	C↓CTAGG	M	37		<i>BlnI</i>	
<i>BalI</i>	TGG↓CCA	L	25	65	<i>MluNI, MscI</i>	<i>dcm</i> ; active >18 hr
<i>BamHI</i>	G↓GATCC	H	37	60	<i>BstI</i>	Star activity
<i>BanI</i>	G↓GPuPuCC	L	50	70		

continued

3.1.10

Table 3.1.1 Recognition Sequences and Reaction Conditions of Commercially Available Restriction Endonucleases^a continued

Name	Site ^b	Salt ^c	Rxn. temp. ^d	Inact. temp. ^d	Isoschizomers	Comments ^e
<i>Ban</i> II	GPuGCPy↓C	M	37	60		
<i>Ban</i> III	AT↓CGAT	M	37	70	<i>Bsi</i> XI, <i>Bsp</i> 106I, <i>Bsp</i> DI, <i>Cla</i> I	
<i>Bbe</i> I	GGCGC↓C	L	37	65	<i>Ehe</i> I, <i>Kas</i> I, <i>Nar</i> I	
<i>Bbi</i> II	GPu↓CGPyC	L	37		<i>Acy</i> I, <i>Bsa</i> HI, <i>Hin</i> 1I, <i>Hsp</i> 92I	
<i>Bbr</i> PI	CAC↓GTG	H	37		<i>Eco</i> 72I, <i>Pma</i> CI, <i>Pml</i> II	
<i>Bbs</i> I	GAAGACN ₂ ↓ CTTCTGN ₆ ↑	L	37	65	<i>Bpu</i> AI, <i>Bsc</i> 91I	
<i>Bbu</i> I	GCATG↓C	L	37	65	<i>Sph</i> I	
<i>Bbv</i> I	GCAGCN ₈ ↓ CGTCGN ₁₂ ↑	L	37	65	<i>Bst</i> 71I	
<i>Bca</i> 77I	(A/T)↓CCGG(A/T)	— ^f	— ^g		<i>Bsa</i> WI	
<i>Bcg</i> I	↓N ₁₀ GCAN ₆ TCGN ₁₂ ↓ ↑N ₁₂ CGTN ₆ AGCN ₁₀ ↑	H	37	65		Requires SAM
<i>Bcl</i> II	T↓GATCA	H(K)	50	100	<i>Bsi</i> QI	<i>dam</i>
<i>Bcn</i> I	CC↓(G/C)GG	H	37		<i>Nci</i> I	
<i>Bco</i> I	C↓PyCGPuG	H	65		<i>Ava</i> I, <i>Nsp</i> III	
<i>Bfa</i> I	C↓TAG	— ^f	37		<i>Mae</i> I	
<i>Bfr</i> I	C↓TTAAG	M	37		<i>Afl</i> III, <i>Bst</i> 98I	
<i>Bgl</i> II	GCCN ₄ ↓NGGC	H	37	65		
<i>Bgl</i> III	A↓GATCT	H	37	100		
<i>Bln</i> I	C↓CTAGG	H	37		<i>Avr</i> II	
<i>Blp</i> I	GC↓TNAGC	— ^f	37		<i>Bpu</i> 1102I, <i>Cel</i> III, <i>Esp</i> I	
<i>Bmy</i> I	G(A/G/T)GC(A/C/T)↓C	— ^f	37		<i>Bsp</i> 1286I	
<i>Bpm</i> I	CTGGAGN ₁₆ ↓ GACCTCN ₁₄ ↑	H	37	65	<i>Gsu</i> I	
<i>Bpu</i> 1102I	GC↓TNAGC	— ^f	37		<i>Blp</i> I, <i>Cel</i> III, <i>Esp</i> I	
<i>Bpu</i> AI	GAAGACN ₂ ↓ CTTCTGN ₆ ↑	H	37		<i>Bbs</i> I, <i>Bsc</i> 91I	
<i>Bsa</i> I	GGTCTCN↓ CCAGAGN ₅ ↑	— ^f	60	65		10% activity at 37°C
<i>Bsa</i> AI	PyAC↓GTPu	H	37			
<i>Bsa</i> BI	GATN ₂ ↓N ₂ ATC	M	60		<i>Bsi</i> BI, <i>Bsr</i> BrI, <i>Mam</i> I	<i>dam</i> ; 20% activity at 37°C
<i>Bsa</i> HI	GPu↓CGPyC	— ^f	37	65	<i>Acy</i> I, <i>Bbi</i> III, <i>Hin</i> 1I, <i>Hsp</i> 92I	↑ activity at 37°C
<i>Bsa</i> JI	C↓CN ₂ GG	M	60			20% activity at 37°C
<i>Bsa</i> MI	GAATGCN↓ CTTAC↑GN	H	65		<i>Bsm</i> I	
<i>Bsa</i> OI	CGPuPy↓CG	M	50		<i>Bsi</i> EI, <i>Mcr</i> I	
<i>Bsa</i> WI	(A/T)↓CCGG(A/T)	M	60		<i>Bca</i> 77I	
<i>Bsc</i> 91I	GAAGACN ₂ ↓ CTTCTGN ₆ ↑	H	37		<i>Bbs</i> I, <i>Bpu</i> AI	
<i>Bsc</i> BI	GGN↓NCC	— ^f	55		<i>Nla</i> IV	
<i>Bse</i> AI	T↓CCGGA	— ^f	55		<i>Acc</i> III, <i>Bsi</i> MI, <i>Bsp</i> EI, <i>Kpn</i> 2I, <i>Mro</i> I	
<i>Bse</i> RI	GAGGAGN ₁₀ ↓ CTCCTCN ₈ ↑	L	37	65		
<i>Bsg</i> I	GTGCAGN ₁₆ ↓ CACGTCN ₁₄ ↑	— ^f	37	65		
<i>Bsh</i> I	GG↓CC		37		<i>Bsp</i> KI, <i>Hae</i> III, <i>Pal</i> I	
<i>Bsh</i> 1236I	CG↓CG	H	37	65	<i>Acc</i> II, <i>Bst</i> UI, <i>Mvn</i> I, <i>Tha</i> I	

continued

3.1.11

Table 3.1.1 Recognition Sequences and Reaction Conditions of Commercially Available Restriction Endonucleases^a continued

Name	Site ^b	Salt ^c	Rxn. temp. ^d	Inact. temp. ^d	Isoschizomers	Comments ^e
<i>Bsi</i> BI	GATN ₂ ↓N ₂ ATC		55		<i>Bsa</i> BI, <i>Bsr</i> BrI, <i>Mam</i> I	
<i>Bsi</i> CI	TT↓CGAA	L	65		<i>Bst</i> BI, <i>Csp</i> 45I, <i>Nsp</i> V, <i>Nsp</i> 7524V, <i>Sfu</i> I	
<i>Bsi</i> EI	CGPuPy↓CG	M	60		<i>Bsa</i> OI, <i>Mcr</i> I	
<i>Bsi</i> HKAI	G(A/T)GC(A/T)↓C	H	60		<i>Alw</i> 2II, <i>Asp</i> HI, <i>Hgi</i> AI	
<i>Bsi</i> LI	CC↓(A/T)GG	— ^f	60		<i>Bst</i> GI, <i>Bst</i> NI, <i>Bst</i> OI, <i>Eco</i> RII, <i>Mva</i> I	
<i>Bsi</i> MI	T↓CCGGA	— ^f	60		<i>Acc</i> III, <i>Bse</i> AI, <i>Bsp</i> EI, <i>Kpn</i> 2I, <i>Mro</i> I	
<i>Bsi</i> QI	T↓GATCA	— ^f	55		<i>Bcl</i> I	
<i>Bsi</i> WI	C↓GTACG	H	60		<i>Sp</i> II, <i>Sun</i> I	50% activity at 37°C
<i>Bsi</i> XI	AT↓CGAT	— ^f	65		<i>Ban</i> III, <i>Bsp</i> 106I, <i>Bsp</i> DI, <i>Cla</i> I	
<i>Bsi</i> YI	CCN ₅ ↓N ₂ GG	M	55		<i>Bsi</i> I	
<i>Bsi</i> ZI	G↓GNCC	— ^f	65		<i>Asu</i> I, <i>Bsu</i> 54I, <i>Cfr</i> 13I, <i>Sau</i> 96I	
<i>Bsl</i> I	CCN ₅ ↓N ₂ GG	H	55		<i>Bsi</i> YI	30% activity at 37°C
<i>Bsm</i> I	GAATGCN↓ CTTAC↑GN	M	65	90	<i>Bsa</i> MI	20% activity at 37°C
<i>Bsm</i> AI	GTCTCN↓ CAGAGN ₅ ↑	H	55		<i>Alw</i> 26I	10% activity at 37°C; 30% activity at 65°C
<i>Bsm</i> BI	CGTCTCN↓ GCAGAGN ₅ ↑	H	55			
<i>Bsm</i> FI	GGGACN ₁₀ ↓ CCCTGN ₁₄ ↑	L	37			
<i>Bso</i> FI	GC↓NGC	L	55	65	<i>Fnu</i> 4HI	
<i>Bsp</i> 106I	AT↓CGAT	— ^f	37		<i>Ban</i> III, <i>Bsi</i> XI, <i>Bsp</i> DI, <i>Cla</i> I	
<i>Bsp</i> 120I	G↓GGCCC	L	37		<i>Apa</i> I	
<i>Bsp</i> 1286I	G(A/G/T)GC(A/C/T)↓C	L	30		<i>Bmy</i> I	<i>dam</i>
<i>Bsp</i> 1407I	T↓GTACA	— ^f	50	70	<i>Ssp</i> BI	
<i>Bsp</i> 90I	GTA↓TAC	— ^f	45		<i>Bst</i> 1107I	
<i>Bsp</i> CI	CGAT↓CG	— ^f	37		<i>Pvu</i> I, <i>Xor</i> II	
<i>Bsp</i> DI	AT↓CGAT	— ^f	37	65	<i>Ban</i> III, <i>Bsi</i> XI, <i>Bsp</i> 106I, <i>Cla</i> I	<i>dam</i>
<i>Bsp</i> EI	T↓CCGGA	H	37		<i>Acc</i> III, <i>Bse</i> AI, <i>Bsi</i> MI, <i>Kpn</i> 2I, <i>Mro</i> I	
<i>Bsp</i> HI	T↓CATGA	H(K)	37	65	<i>Rca</i> I	<i>dam</i>
<i>Bsp</i> KI	GG↓CC	— ^f	40		<i>Bsh</i> I, <i>Hae</i> III, <i>Pal</i> I	
<i>Bsp</i> MI	ACCTGCN ₄ ↓ TGGACGN ₈ ↑	M	37			
<i>Bsp</i> WI	GCN ₅ ↓N ₂ GC	— ^f	42		<i>Mwo</i> I	
<i>Bsr</i> I	ACTGGN↓ TGAC↑CN	H(K)	65			
<i>Bsr</i> BI	GAG↓CGG CTC↑GCC	M	37		<i>Bst</i> D102 I	
<i>Bsr</i> BrI	GATN ₂ ↓N ₂ ATC	L	65		<i>Bsa</i> BI, <i>Bsi</i> BI, <i>Mam</i> I	
<i>Bsr</i> DI	GCAATGN ₂ ↓ CGTTAC↑N ₂	M	60			
<i>Bsr</i> FI	Pu↓CCGGPy	M	37		<i>Cfr</i> 10I	
<i>Bsr</i> GI	T↓GTACA	M	37		<i>Ssp</i> BI	
<i>Bss</i> HII	G↓CGCGC	L ^f	50	90		

continued

3.1.12

Table 3.1.1 Recognition Sequences and Reaction Conditions of Commercially Available Restriction Endonucleases^a continued

Name	Site ^b	Salt ^c	Rxn. temp. ^d	Inact. temp. ^d	Isoschizomers	Comments ^e
<i>Bss</i> SI	C↓TCGTG	H	37			
<i>Bst</i> I	G↓GATCC	H(K)	55	85	<i>Bam</i> HI	Star activity; 10% activity at 37°C
<i>Bst</i> 1107I	GTA↓TAC	H(K)	37		<i>Bsp</i> 90I	
<i>Bst</i> 71I	GCAGCN ₈ ↓ CGTCGN ₁₂ ↑	H	50		<i>Bbv</i> I	
<i>Bst</i> 98I	C↓TTAAG	M	37		<i>Af</i> III, <i>Bfr</i> I	
<i>Bst</i> BI	TT↓CGAA	M	65		<i>Bsi</i> CI, <i>Csp</i> 45I, <i>Nsp</i> 7524V, <i>Nsp</i> V, <i>Sfu</i> I	10% activity at 37°C
<i>Bst</i> D102I	GAG↓CGG	M	37		<i>Bsr</i> BI	
<i>Bst</i> EII	G↓GTNACC	H	60	85	<i>Eco</i> O651	Star activity; 10% activity at 37°C
<i>Bst</i> GII	CC↓(A/T)GG	— ^f	— ^g		<i>Bsi</i> LI, <i>Bst</i> NI, <i>Bst</i> OI, <i>Eco</i> RII, <i>Mva</i> I	
<i>Bst</i> NI	CC↓(A/T)GG	H	60		<i>Bsi</i> LI, <i>Bst</i> GII, <i>Bst</i> OI, <i>Eco</i> RII, <i>Mva</i> I	Difficult ligation; 30% activity at 37°C
<i>Bst</i> OI	CC↓(A/T)GG	M	60		<i>Bsi</i> LI, <i>Bst</i> GII, <i>Bst</i> NI, <i>Eco</i> RII, <i>Mva</i> I	
<i>Bst</i> UI	CG↓CG	H	60		<i>Acc</i> II, <i>Bsh</i> 236I, <i>Mvn</i> I, <i>Tha</i> I	20% activity at 37°C
<i>Bst</i> XI	CCAN ₅ ↓NTGG	H	50	65		
<i>Bst</i> YI	Pu↓GATCPy	L	60		<i>Mfl</i> I, <i>Xho</i> II	30% activity at 37°C
<i>Bst</i> ZI	C↓GGCCG	H	50		<i>Eag</i> I, <i>Ecl</i> XI, <i>Eco</i> 52I, <i>Xma</i> III	
<i>Bsu</i> 23I	T↓CCGGA	— ^f	37	65	<i>Acc</i> III, <i>Bse</i> AI, <i>Bsi</i> MI, <i>Bsp</i> EI, <i>Kpn</i> 2I, <i>Mro</i> I	
<i>Bsu</i> 36I	CC↓TNAGG	H	37		<i>Aoc</i> I, <i>Cvn</i> I, <i>Eco</i> 81I, <i>Mst</i> II, <i>Sau</i> I	
<i>Bsu</i> 54I	G↓GNCC	— ^f	— ^g		<i>Asu</i> I, <i>Bsi</i> ZI, <i>Cfr</i> 13I, <i>Sau</i> 96I	
<i>Cac</i> 8I	GCN↓NGC	H	37	65		
<i>Ccr</i> I	C↓TCGAG	H	37		<i>Pae</i> R7I, <i>Xho</i> I	
<i>Cel</i> III	GC↓TNAGC	H	37		<i>Blp</i> I, <i>Bpu</i> 1102I, <i>Esp</i> I	
<i>Cfo</i> I	GCG↓C	L	37		<i>Hha</i> I, <i>Hin</i> P1I	
<i>Cfr</i> 9I	C↓CCGGG	— ^f	37		<i>Psp</i> AI, <i>Sma</i> I, <i>Xma</i> I	
<i>Cfr</i> 10I	Pu↓CCGGPy	L ^f	37	100	<i>Bsr</i> FI	
<i>Cfr</i> 13I	G↓GNCC	M	37	100	<i>Asu</i> I, <i>Bsi</i> Z I, <i>Bsu</i> 54 I, <i>Sau</i> 96 I	
<i>Cla</i> I	AT↓CGAT	M	37	65	<i>Ban</i> III, <i>Bsi</i> XI, <i>Bsp</i> 106I, <i>Bsp</i> DI	<i>dam</i>
<i>Cpo</i> I	CG↓G(A/T)CCG	H	30	60	<i>Csp</i> I, <i>Rsr</i> II	
<i>Csp</i> 6I	G↓TAC	L	37		<i>Afa</i> I, <i>Rsa</i> I	
<i>Csp</i> I	CG↓G(A/T)CCG	H(K)	30		<i>Cpo</i> I, <i>Rsr</i> II	
<i>Csp</i> 45I	TT↓CGAA	M	37	65	<i>Bsi</i> CI, <i>Bst</i> BI, <i>Nsp</i> V, <i>Nsp</i> 7524V, <i>Sfu</i> I	
<i>Cvi</i> JI	PuG↓CPy	— ^f	— ^g			
<i>Cvn</i> I	CC↓TNAGG	M	37	65	<i>Aoc</i> I, <i>Bsu</i> 36I, <i>Eco</i> 81I, <i>Mst</i> II, <i>Sau</i> I	
<i>Dde</i> I	C↓TNAG	H	37	70		
<i>Dpn</i> I	GA↓TC	H	37	65		Both strands methylated A ^m
<i>Dpn</i> II	↓GATC	H	37	65	<i>Mbo</i> I, <i>Nde</i> II, <i>Sau</i> 3A I	<i>dam</i>
<i>Dra</i> I	TTT↓AAA	M	37	65		
<i>Dra</i> II	PuG↓GNCCPy	H	37		<i>Eco</i> O109I	
<i>Dra</i> III	CACN ₃ ↓GTG	H	37	65		Star activity
<i>Drd</i> I	GACN ₄ ↓N ₂ GTC	M(K)	37			

continued

Table 3.1.1 Recognition Sequences and Reaction Conditions of Commercially Available Restriction Endonucleases^a continued

Name	Site ^b	Salt ^c	Rxn. temp. ^d	Inact. temp. ^d	Isoschizomers	Comments ^e
<i>DsaI</i>	C↓CPuPyGG	H	55			
<i>DsaV</i>	↓CCNGG	H	60		<i>ScrFI</i>	
<i>EaeI</i>	Py↓GGCCPu	M(K)	37	65		<i>dcm</i>
<i>EagI</i>	C↓GGCCG	H	37	65	<i>BstZI, EclXI, Eco52I, XmaIII</i>	
<i>Eam1104I</i>	CTCTTCN↓ GAGAAGN ₄ ↑	H(K)	37	65	<i>EarI, Ksp632I</i>	
<i>Eam1105I</i>	GACN ₃ ↓N ₂ GTC	— ^f	37	65	<i>AhdI, AspEI, EclHKI</i>	
<i>EarI</i>	CTCTTCN↓ GAGAAGN ₄ ↑	L	37	65	<i>Eam1104I, Ksp632I</i>	
<i>EclHKI</i>	GACN ₃ ↓N ₂ GTC	M	37	65	<i>AhdI, AspEI, Eam1105I</i>	
<i>EclXI</i>	C↓GGCCG	H	37		<i>BstZI, EagI, Eco52I, XmaIII</i>	
<i>Ecl136II</i>	GAG↓CTC	— ^f	37	65	<i>EcoICRI, SacI, SstI</i>	
<i>Eco105I</i>	TAC↓GTA	L	37		<i>SnaBI</i>	
<i>Eco130I</i>	C↓C(A/T)(A/T)GG	H	37		<i>EcoT14I, StyI</i>	
<i>Eco47I</i>	G↓G(A/T)CC	H	37	100	<i>AvaII, SniI</i>	
<i>Eco47III</i>	AGC↓GCT	H	37	100	<i>Aor51HI</i>	
<i>Eco52I</i>	C↓GGCCG	H	37	80	<i>BstZI, EagI, EclXI, XmaIII</i>	
<i>Eco57I</i>	CTGAAGN ₁₆ ↓ GACTTCN ₁₄ ↑	L	37	65		
<i>EcoO65I</i>	G↓GTNACC	H	37	70	<i>BstEII</i>	
<i>Eco72I</i>	CAC↓GTG	— ^f	37	65	<i>BbrPI, PmaCI, PmlI</i>	
<i>Eco81I</i>	CC↓TNAGG	L	37	90	<i>AocI, Bsu36I, CvnI, MstII, SauI</i>	
<i>EcoICRI</i>	GAG↓CTC	M	37	65	<i>Ecl136II, SacI, SstI</i>	
<i>EcoNI</i>	CCTN ₂ ↓N ₃ AGG	— ^f	37			
<i>EcoO109I</i>	(A/G)G↓GNCC(C/T)	L	37	65	<i>DraII</i>	<i>dcm</i>
<i>EcoRI</i>	G↓AATTC	H	37	65		Star activity
<i>EcoRII</i>	↓CC(A/T)GG	M	37	60	<i>BstGII, BsiLI, BstNI, BstOI, MvaI</i>	<i>dcm</i>
<i>EcoRV</i>	GAT↓ATC	H	37	65		Star activity
<i>EcoT14I</i>	C↓C(A/T)(A/T)GG	H	37		<i>Eco130I, StyI</i>	
<i>EcoT22I</i>	ATGCA↓T	H	37	100	<i>NsiI, Ppu10I</i>	
<i>EheI</i>	GGC↓GCC	L	37	70	<i>BbeI, KasI, NarI</i>	
<i>EspI</i>	GC↓TNAGC	H	37	100	<i>BlpI, Bpu1102I, CclII</i>	
<i>Esp3I</i>	CGTCTCN↓ GCAGAGN ₅ ↑	— ^f	37	65		
<i>FdiII</i>	TGC↓GCA	L	50	100	<i>AviII, FspI, MstI</i>	
<i>Fnu4HI</i>	GC↓NGC	L	37	65	<i>BsoFI, ItaI</i>	Difficult ligation
<i>FokI</i>	GGATGN ₉ ↓ CCTACN ₁₃ ↑	L	37	65		
<i>FseI</i>	GGCCGG↓CC	M	30			
<i>FspI</i>	TGC↓GCA	M	37	65	<i>AviII, FdiII, MstI</i>	
<i>GsuI</i>	CTGGAGN ₁₆ ↓ GACCTCN ₁₄ ↑	— ^f	37	65	<i>BpmI</i>	
<i>HaeII</i>	PuGCGC↓Py	M	37	65		
<i>HaeIII</i>	GG↓CC	M	37	90	<i>BshI, BspKI, PstII</i>	Very stable; active at 70°C
<i>HapII</i>	C↓CGG	— ^f	37		<i>HpaII, MspI</i>	
<i>HgaI</i>	GACGCN ₅ ↓ CTGCGN ₁₀ ↑	M	37	65		

continued

Table 3.1.1 Recognition Sequences and Reaction Conditions of Commercially Available Restriction Endonucleases^a continued

Name	Site ^b	Salt ^c	Rxn. temp. ^d	Inact. temp. ^d	Isoschizomers	Comments ^e
<i>Hgi</i> AI	G(A/T)GC(A/T)↓C	H	37	65	<i>Alw</i> 2II, <i>Asp</i> HI, <i>Bsi</i> HKAI	
<i>Hgi</i> DI	GPu↓CGPyC				<i>Acy</i> I, <i>Bbi</i> II, <i>Bsa</i> HI, <i>Hin</i> II	
<i>Hha</i> I	GCG↓C	M	37	90	<i>Cfo</i> I, <i>Hin</i> P1I	
<i>Hin</i> I	GPu↓CGPyC	L	37	80	<i>Acy</i> I, <i>Bbi</i> II, <i>Bsa</i> HI, <i>Hsp</i> 92I	
<i>Hinc</i> II	GTPy↓PuAC	H	37	70	<i>Hind</i> II	
<i>Hind</i> II	GTPy↓PuAC	M	37	65	<i>Hinc</i> II	
<i>Hind</i> III	A↓AGCTT	M	37	90		Star activity
<i>Hinf</i> I	G↓ANTC	H	37	80		Star activity
<i>Hin</i> P1I	G↓CGC	L	37		<i>Cfo</i> I, <i>Hha</i> I	
<i>Hpa</i> I	GTT↓AAC	M(K)	37	90		
<i>Hpa</i> II	C↓CGG	L(K)	37	90	<i>Hap</i> II, <i>Msp</i> I	No cleavage when internal C is methylated
<i>Hph</i> I	GGTGAN ₈ ↓ CCACTN ₇ ↑	L(K) ^f	37	65		<i>dam</i>
<i>Hsp</i> 92I	GPu↓CGPyC	L	37		<i>Acy</i> I, <i>Bbi</i> II, <i>Bsa</i> HI, <i>Hin</i> II	
<i>Hsp</i> 92II	CATG↓	H	37		<i>Nla</i> III	
<i>Ita</i> I	GC↓NGC	H	37	65	<i>Bso</i> FI, <i>Fnu</i> 4HI	
<i>Kas</i> I	G↓GCGCC	M	37	65	<i>Bbe</i> I, <i>Ehe</i> I, <i>Nar</i> I	Site-dependent activity
<i>Kpn</i> I	GGTAC↓C	L	37	60	<i>Acc</i> 65I, <i>Asp</i> 718I	Star activity
<i>Kpn</i> 2I	T↓CCGGA	— ^f	55		<i>Acc</i> III, <i>Bse</i> AI, <i>Bsi</i> MI, <i>Bsp</i> EI, <i>Mro</i> I	
<i>Ksp</i> I	CCGC↓GG	H	37		<i>Sac</i> II, <i>Sst</i> II	
<i>Ksp</i> 632I	CTCTTCN↓ GAGAAGN ₄ ↑	H(K)	37		<i>Eam</i> 1104I, <i>Ear</i> I	
<i>Mae</i> I	C↓TAG	H	45		<i>Bfa</i> I	
<i>Mae</i> II	A↓CGT	H	50			
<i>Mae</i> III	↓GTNAC	H	45			
<i>Mam</i> I	GATN ₂ ↓N ₂ ATC	H	37		<i>Bsa</i> BI, <i>Bsi</i> BI, <i>Bsr</i> BrI	
<i>Mbo</i> I	↓GATC	H	37	65	<i>Dpn</i> II, <i>Nde</i> II, <i>Sau</i> 3AI	<i>dam</i>
<i>Mbo</i> II	GAAGAN ₈ ↓ CTTCTN ₇ ↑	L(K)	37	65		<i>dam</i>
<i>Mcr</i> I	CGPuPy↓CG	H	37		<i>Bsa</i> OI, <i>Bsi</i> EI	
<i>Mfe</i> I	C↓AATTG	— ^f	37	65	<i>Mun</i> I	
<i>Mfl</i> II	Pu↓GATCPy	L	37		<i>Bst</i> YI, <i>Xho</i> II	
<i>Mlu</i> I	A↓CGCGT	H	37	100		
<i>Mlu</i> NI	TGG↓CCA	— ^f	37	65	<i>Bal</i> I, <i>Msc</i> I	
<i>Mnl</i> II	CCTCN ₇ GGAGN ₆	M	37	65		
<i>Mro</i> I	T↓CCGGA	L	37	100	<i>Acc</i> III, <i>Bse</i> AI, <i>Bsi</i> MI, <i>Bsp</i> EI, <i>Kpn</i> 2I	
<i>Msc</i> I	TGG↓CCA	M	37		<i>Bal</i> I, <i>Mlu</i> NI	Star activity
<i>Mse</i> I	T↓TAA	M	37	65	<i>Tru</i> 9I	
<i>Msl</i> II	CAPyN ₂ ↓N ₂ PuTG	M	37	65		
<i>Msp</i> I	C↓CGG	M	37	90	<i>Hap</i> II, <i>Hpa</i> II	No cleavage when 5' C is methylated
<i>Msp</i> A1I	C(A/C)G↓C(G/T)G	H	37	65	<i>Nsp</i> BII	

continued

Table 3.1.1 Recognition Sequences and Reaction Conditions of Commercially Available Restriction Endonucleases^a continued

Name	Site ^b	Salt ^c	Rxn. temp. ^d	Inact. temp. ^d	Isoschizomers	Comments ^e
<i>Mst</i> I	TGC↓GCA	H	37		<i>Avi</i> II, <i>Fdi</i> II, <i>Fsp</i> I	
<i>Mst</i> II	CC↓TNAGG	H	37	65	<i>Aoc</i> I, <i>Bsu</i> 36I, <i>Cvn</i> I, <i>Eco</i> 81I, <i>Sau</i> I	
<i>Mun</i> I	C↓AATTG	M	37	65	<i>Mfe</i> I	
<i>Mva</i> I	CC↓(A/T)GG	H	37	100	<i>Bsi</i> LI, <i>Bst</i> GII, <i>Bst</i> NI, <i>Bst</i> OI, <i>Eco</i> RII	
<i>Mvn</i> I	CG↓CG	M	37		<i>Acc</i> II, <i>Bsh</i> 1236I, <i>Bst</i> UI, <i>Tha</i> I	
<i>Mwo</i> I	GCN ₅ ↓N ₂ GC	— ^f	60		<i>Bsp</i> WI	
<i>Nae</i> I	GCC↓GGC	L	37	100	<i>Ngo</i> AIV, <i>Ngo</i> MI	Site-dependent activity
<i>Nar</i> I	GG↓CGCC	L	37	65	<i>Ehe</i> I, <i>Kas</i> I	Stable at 37°C for 24 hr; site-dependent activity
<i>Nci</i> II	CC↓(G/C)GG	L	37	80	<i>Bcn</i> I	Difficult ligation
<i>Nco</i> I	C↓CATGG	H	37	80		
<i>Nde</i> I	CA↓TATG	— ^f	37	65		<i>t</i> _{1/2} ~15 min at 37°C
<i>Nde</i> II	↓GATC	H	37	65	<i>Dpn</i> II, <i>Mbo</i> I, <i>Sau</i> 3AI	
<i>Ngo</i> AIV	G↓CCGGC	— ^f	37		<i>Nae</i> I, <i>Ngo</i> MI	
<i>Ngo</i> MI	G↓CCGGC	— ^f	37		<i>Nae</i> I, <i>Ngo</i> AIV	
<i>Nhe</i> I	G↓CTAGC	M	37	65		
<i>Nla</i> III	CATG↓	L ^f	37	65	<i>Hsp</i> 92II	
<i>Nla</i> IV	GGN↓NCC	L ^f	37	65	<i>Bsc</i> BI	
<i>Not</i> I	GC↓GGCCGC	H	37	100		
<i>Nru</i> I	TCG↓CGA	H	37	80	<i>Spo</i> I	<i>dam</i>
<i>Nsi</i> II	ATGCA↓T	H	37		<i>Eco</i> T22I, <i>Ppu</i> 10I	
<i>Nsp</i> I	PuCATG↓Py	L	37	65	<i>Nsp</i> 7524I	
<i>Nsp</i> III	C↓PyCGPuG	— ^f	37	85	<i>Ava</i> I, <i>Bco</i> I	
<i>Nsp</i> V	TT↓CGAA	L	50		<i>Bsi</i> CI, <i>Bst</i> BI, <i>Csp</i> 45I, <i>Nsp</i> 7524V, <i>Sfu</i> I	
<i>Nsp</i> 7524I	PuCATG↓Py	— ^f	37		<i>Nsp</i> I	
<i>Nsp</i> 7524V	TT↓CGAA	L	37	70	<i>Bsi</i> CI, <i>Bst</i> BI, <i>Csp</i> 45I, <i>Nsp</i> V, <i>Sfu</i> I	
<i>Nsp</i> BII	C(A/C)G↓C(G/T)G	L	37		<i>Msp</i> A1I	
<i>Pac</i> I	TTAAT↓TAA	L	37			
<i>Pae</i> R7I	C↓TCGAG	— ^f	37		<i>Ccr</i> I, <i>Xho</i> I	Site-dependent activity
<i>Pai</i> I	GG↓CC	M	37	65	<i>Bsh</i> I, <i>Bsp</i> KI, <i>Hae</i> III	
<i>Pfl</i> MI	CCAN ₄ ↓NTGG	H	37	65	<i>Acc</i> B7I, <i>Van</i> 91I	
<i>Pin</i> AI	A↓CCGGT	M	37	65	<i>Age</i> I	
<i>Ple</i> I	GAGTCN ₄ ↓ CTCAGN ₅ ↑	— ^f	37	65		
<i>Pma</i> CI	CAC↓GTG	L	37	60	<i>Bbr</i> PI, <i>Eco</i> 72I, <i>Pm</i> II	
<i>Pme</i> I	GTTT↓AAAC	— ^f	37	65		
<i>Pm</i> II	CAC↓GTG	L	37		<i>Bbr</i> PI, <i>Eco</i> 72I, <i>Pma</i> CI	
<i>Ppu</i> 10I	A↓TGCAT	— ^f	37	65	<i>Eco</i> T22I, <i>Nsi</i> I	
<i>Ppu</i> MI	PuG↓G(A/T)CCPy	— ^f	37		<i>Psp</i> 5II	<i>dcm</i>
<i>Psh</i> AI	GACN ₂ ↓N ₂ GTC	K	37			
<i>Psp</i> 1406I	AA↓CGTT	— ^f	37	65		
<i>Psp</i> 5II	PuG↓G(A/T)CCPu	H	37		<i>Ppu</i> MI	

continued

Table 3.1.1 Recognition Sequences and Reaction Conditions of Commercially Available Restriction Endonucleases^a continued

Name	Site ^b	Salt ^c	Rxn. temp. ^d	Inact. temp. ^d	Isoschizomers	Comments ^e
<i>Psp</i> AI	C↓CCGGG	— ^f	37	65	<i>Cfr</i> 9I, <i>Sma</i> I, <i>Xma</i> I	
<i>Pst</i> I	CTGCA↓G	H	37	70		Star activity
<i>Pvu</i> I	CGAT↓CG	H	37	100	<i>Bsp</i> CI, <i>Xor</i> II	
<i>Pvu</i> II	CAG↓CTG	M	37	95		Star activity
<i>Rca</i> I	T↓CATGA	M	37	65	<i>Bsp</i> HI	
<i>Rsa</i> I	GT↓AC	M	37	65	<i>Afa</i> I, <i>Csp</i> 6I	
<i>Rsr</i> II	CG↓G(A/T)CCG	L	37	65	<i>Cpo</i> I, <i>Csp</i> I	
<i>Sac</i> I	GAGCT↓C	L	37	60	<i>Ecl</i> 136II, <i>Eco</i> ICRI, <i>Sst</i> I	
<i>Sac</i> II	CCGC↓GG	L	37	80	<i>Ksp</i> I, <i>Sst</i> II	Site-dependent activity
<i>Sal</i> I	G↓TCGAC	H	37	80		Star activity
<i>Sap</i> I	GCTCTTCN↓ CGAGAAGN ₄ ↑	— ^f	37	65		
<i>Sau</i> I	CC↓TNAGG	H	37		<i>Aoc</i> I, <i>Bsu</i> 36I, <i>Cvn</i> I, <i>Eco</i> 81I, <i>Mst</i> II	
<i>Sau</i> 3AI	↓GATC	H	37	70	<i>Dpn</i> II, <i>Mbo</i> I, <i>Nde</i> II	
<i>Sau</i> 96I	G↓GNCC	L	37		<i>Asu</i> I, <i>Bsi</i> ZI, <i>Bsu</i> 54I, <i>Cfr</i> 13I	<i>dcm</i>
<i>Sca</i> I	AGT↓ACT	H	37	100		Star activity
<i>Scr</i> FI	CC↓NGG	H	37	65	<i>Dsa</i> V	
<i>Sdy</i> I	GGNCC	— ^f	— ^g		<i>Asu</i> I, <i>Bsi</i> ZI, <i>Bsu</i> 54I, <i>Cfr</i> 13I, <i>Sau</i> 96I	
<i>Sex</i> AI	A↓CC(A/T)GGT	M	37			
<i>Sfa</i> NI	GCATCN ₅ ↓ CGTAGN ₉ ↑	H	37	65		
<i>Sfc</i> I	C↓TPuPyAG	— ^f	37	65		
<i>Sfi</i> I	GGCCN ₄ ↓NGGCC	M	50			
<i>Sfu</i> I	TT↓CGAA	H	37		<i>Bsi</i> CI, <i>Bst</i> BI, <i>Csp</i> 45I, <i>Nsp</i> 7524V, <i>Nsp</i> V	
<i>Sgf</i> I	GCGAT↓CGC	L	37			
<i>Sgr</i> AI	CPu↓CCGGPyG	— ^f	37			
<i>Sin</i> I	G↓G(A/T)CC	L	37	65	<i>Ava</i> II, <i>Eco</i> 47I	
<i>Sma</i> I	CCC↓GGG	L(K) ^f	30	65	<i>Cfr</i> 9I, <i>Psp</i> AI, <i>Xma</i> I	
<i>Sna</i> BI	TAC↓GTA	M	37		<i>Eco</i> 105I	
<i>Spe</i> I	A↓CTAGT	M	37	65		
<i>Sph</i> I	GCATG↓C	H	37	100	<i>Bbu</i> I	
<i>Sp</i> II	C↓GTACG	H	55		<i>Bsi</i> WI, <i>Sun</i> I	
<i>Spo</i> I	TCG↓CGA	M(K)	37	65	<i>Nru</i> I	
<i>Srf</i> I	GCCC↓GGGC	M	37	65		
<i>Sse</i> 8387I	CCTGCA↓GG	— ^f	37	60		
<i>Ssp</i> I	AAT↓ATT	— ^f	37	65		
<i>Ssp</i> BI	T↓GTACA	— ^f	50	70	<i>Bsr</i> GI	
<i>Sst</i> I	GAGCT↓C	M	37	65	<i>Ecl</i> 136II, <i>Eco</i> ICRI, <i>Sac</i> I	
<i>Sst</i> II	CCGC↓GG	M	37		<i>Ksp</i> I, <i>Sac</i> II	
<i>Stu</i> I	AGG↓CCT	M	37	65	<i>Aat</i> I	<i>dcm</i>
<i>Sty</i> I	C↓C(A/T)(A/T)GG	H	37	65	<i>Eco</i> O130I, <i>Eco</i> T14I	
<i>Sun</i> I	C↓GTACG	M	55		<i>Bsi</i> WI, <i>Sp</i> II	
<i>Swa</i> I	ATTT↓AAAT	H	25			

continued

Table 3.1.1 Recognition Sequences and Reaction Conditions of Commercially Available Restriction Endonucleases^a continued

Name	Site ^b	Salt ^c	Rxn. temp. ^d	Inact. temp. ^d	Isoschizomers	Comments ^e
<i>Taq</i> I	T↓CGA	H	65	90	<i>Tth</i> HB8I	<i>dam</i>
<i>Tfi</i> I	G↓A(A/T)TC	L	65			Star activity; 10% activity at 37°C
<i>Tha</i> I	CG↓CG	L	60		<i>Acc</i> II, <i>Bsh</i> 1236I, <i>Bst</i> UI, <i>Mvn</i> I	
<i>Tru</i> 9I	T↓TAA	M	65		<i>Mse</i> I	
<i>Tsp</i> 45I	↓GT(G/C)AC	L	65			
<i>Tsp</i> 509I	↓AATT	— ^f	65			
<i>Tsp</i> RI	N ₂ CAGTGN ₂ ↓	— ^f	65			
<i>Tth</i> 111I	GACN↓N ₂ GTC	— ^f	65		<i>Asp</i> I	Difficult ligation; 10% activity at 37°C
<i>Tth</i> HB8I	T↓CGA	H	65		<i>Taq</i> I	
<i>Van</i> 91I	CCAN ₄ ↓NTGG	H(K)	37	65	<i>Acc</i> B7I, <i>Pfi</i> MI	
<i>Vsp</i> I	AT↓TAAT	H	37		<i>Ase</i> I, <i>Asn</i> I	
<i>Xba</i> I	T↓CTAGA	M	37	70		<i>dam</i>
<i>Xcm</i> I	CCAN ₅ ↓N ₄ TGG	M(K)	37	65		
<i>Xho</i> I	C↓TCGAG	H	37	80	<i>Ccr</i> I, <i>Pae</i> R7I	
<i>Xho</i> II	Pu↓GATCPy	L	37	65	<i>Bst</i> YI, <i>Mfi</i> I	
<i>Xma</i> I	C↓CCGGG	L	37	65	<i>Cfr</i> 9I, <i>Psp</i> AI, <i>Sma</i> I	
<i>Xma</i> III	C↓GGCCG	L	25		<i>Bst</i> ZI, <i>Eag</i> I, <i>Ecl</i> XI, <i>Eco</i> 52I	
<i>Xmn</i> I	GAAN ₂ ↓N ₂ TTC	L	37	65	<i>Asp</i> 700I	Star activity
<i>Xor</i> II	CGAT↓CG	L	37		<i>Bsp</i> CI, <i>Pvu</i> I	
Intron-encoded endonucleases						
I- <i>Ceu</i> I	TAAGTATAACGGTCCCTAA↓GGTAGCGA ^f ATTGATATTGCCAG↑GATTCCATCGCT					Rxn. temp. 37°C ^f
I- <i>Ppo</i> I	ATGACTCTCTTAA↓GGTAGCCAA ^f TACTGAGAG↑AATTCCATCGGTTT					Rxn. temp. 37°C ^f
I- <i>Sce</i> I	TAGGGATAA↓CAGGGTAAT ^f ATCCC↑TATTGTCCCATTA					Rxn. temp. 37°C ^f
PI- <i>Psp</i> I	TGGCAAACAGCTATTAT↓GGGTATTATGGG ^f ACCGTTTGTTCGAT↑AATACCCATAATACCCA					Rxn. temp. 65°C ^f
PI- <i>Sce</i> I	ATCTATGTCGGGTGC↓GGAGAAAGAGGTAATGAAATGGCA ^f TAGATACAGCC↑CACGCCTTTTCTCCATTACTTTACCGT					Rxn. temp. 37°C ^f
PI- <i>Tli</i> I	GGTCTTTTATGCGGACAC↓TGACGGCTTTATG ^f CCAAGAAATACGCC↑TGTGACTGCCGAAATAC					Rxn. temp. 50°C ^f

^aCompiled from Amersham (1994-95), Boehringer Mannheim (1994), GIBCO/BRL (1994), New England Biolabs (1995), Pharmacia Biotech (1994), Promega (1994), Stratagene (1995). See APPENDIX 4 for addresses and phone numbers of suppliers.

^bAbbreviations: N, any nucleotide (G, A, T, C); Pu, either purine (G or A); Py, either pyrimidine (C or T).

^cRecommended concentrations of NaCl (or KCl, indicated by parenthetical K) where L is <50 mM, M is 50-100 mM, and H is >100 mM (see recipe for restriction endonuclease buffers).

^dReaction temperature is the temperature (°C) at which the reaction should be performed; inactivation temperature indicates the temperature (°C) at which the enzyme is inactivated after 15 min of incubation.

^eAbbreviations and other terminology: *dam*, activity blocked by *dam* or overlapping *dam* methylation; *dcm*, activity blocked by *dcm* or overlapping *dcm* methylation; difficult ligation, the enzyme has single-bp 5' overhanging ends that are difficult to ligate with T4 DNA ligase; SAM, S-adenosylmethionine; site-dependent activity refers to marked differences in rates of cleavage at various sites by a particular enzyme (probably determined by the surrounding sequence); star activity refers to altered specificity of a restriction enzyme that causes it to cleave sequences that are similar but not identical to its defined recognition sequence—conditions that may provoke this altered specificity include elevated pH, high glycerol concentration, low ionic strength, and high enzyme to DNA ratio; *t*_{1/2}, half-life.

^fRefer to manufacturer's recommendations and Table 3.1.2 for information on salt requirements.

^gRefer to manufacturer's recommendations for reaction temperature.

Table 3.1.2 Effect of NaCl Concentration on Restriction Endonuclease Activity^a

Enzyme	0 mM NaCl	50 mM NaCl	100 mM NaCl	150 mM NaCl	Enzyme	0 mM NaCl	50 mM NaCl	100 mM NaCl	150 mM NaCl
<i>AatII</i>	+	++	++	+	<i>EcoRI</i>	^b	+++	+++	+++
<i>AccI</i>	+++	+++	+	+	<i>EcoRV</i>	+	+	+	+++
<i>AccI</i>	+	+++	+++	+++	<i>Fnu4HI</i>	+++	+++	++	+
<i>AflIII</i>	+	+++	++	++	<i>FokI</i>	+++	+++	+++	+++
<i>AhaII</i>	+	++	+++	+++	<i>FspI</i>	+	+++	++	++
<i>AluI</i>	+	+++	+++	++	<i>HaeII</i>	+++	+++	+++	++
<i>AlwI</i>	+++	+++	++	+	<i>HaeIII</i>	+++	+++	+++	+++
<i>AlwNI</i>	++	+++	+++	+	<i>HgaI</i>	+++	+++	+	+
<i>ApaI</i>	+++	+++	++	+	<i>HgiAI</i>	+	+	++	+++
<i>ApaLI</i>	+++	+++	++	+	<i>HhaI</i>	+	+	+	++
<i>AscI</i>	+	++	++	+	<i>HincII</i>	++	+++	+++	+++
<i>AseI</i>	+	+++	+++	+++	<i>HindIII</i>	++	+++	+++	++
<i>AvaI</i>	+++	+++	+++	+++	<i>Hinfl</i>	++	+++	+++	+++
<i>AvaII</i>	+++	+++	++	+	<i>HinPI</i>	+++	+++	+++	+++
<i>AvrII</i>	+++	+++	+++	++	<i>HpaI</i>	+	+++	+++	+
<i>BalI</i>	+++	++	++	+	<i>HpaII</i>	+++	+++	++	+
<i>BamHI</i>	+	++	+++	+++	<i>HphI</i>	+++	+++	+++	+
<i>BanI</i>	+++	+++	++	++	<i>KasI</i>	+	+++	+	+
<i>BanII</i>	+++	+++	+++	+++	<i>KpnI</i>	+++	+	+	+
<i>BbvI</i>	+++	+++	+++	+++	<i>MboI</i>	++	+++	+++	+++
<i>BcgI</i>	+++	+++	+++	+++	<i>MboII</i>	+++	+++	+++	+++
<i>BclI</i>	+	+++	+++	+	<i>MluI</i>	++	+++	+++	++
<i>BglI</i>	+	+++	+++	+++	<i>MnII</i>	+++	+++	+++	++
<i>BglIII</i>	++	+++	+++	+++	<i>MseI</i>	+++	+++	++	+
<i>BsaHI</i>	++	+++	+++	+	<i>MspI</i>	+++	+++	+++	+++
<i>BsmI</i>	+++	+++	+++	+++	<i>NaeI</i>	+++	+++	+++	+
<i>Bsp1286</i>	+++	+++	++	+	<i>NarI</i>	+++	+++	+	+
<i>BspHI</i>	++	+++	+++	+++	<i>NciI</i>	+++	+++	++	+
<i>BspMI</i>	++	++	+++	+++	<i>NcoI</i>	+	++	+++	+++
<i>BspMII</i>	++	++	+++	+++	<i>NdeI</i>	+	+	++	+++
<i>BssHIII</i>	+++	+++	+++	+++	<i>NheI</i>	+++	+++	+++	++
<i>BstBI</i>	+++	+++	++	+	<i>NlaIII</i>	+	+	+	+
<i>BstEII</i>	+	++	+++	+++	<i>NlaIV</i>	+	+	+	+
<i>BstNI</i>	++	++	+++	+++	<i>NotI</i>	+	+++	+++	+++
<i>BstUI</i>	+++	+++	++	+	<i>NruI</i>	+	+++	+++	+++
<i>BstXI</i>	++	+++	+++	+++	<i>NsiI</i>	++	++	++	++
<i>BstYI</i>	+++	+++	++	++	<i>PaeR7I</i>	+++	+++	+++	+
<i>Bsu36I</i>	+	++	+++	++	<i>PflMI</i>	++	+++	+++	++
<i>ClaI</i>	+++	+++	+++	++	<i>PleI</i>	+++	+	+	+
<i>DdeI</i>	++	+++	+++	+++	<i>PpuMI</i>	+++	+++	++	+
<i>DpnI</i>	+	++	+++	+++	<i>PstI</i>	+++	+++	+++	+++
<i>DraI</i>	++	+++	+	+	<i>PvuI</i>	+	++	+++	+++
<i>DraIII</i>	++	+++	+++	++	<i>PvuII</i>	+++	+++	+++	+++
<i>EaeI</i>	++	+++	++	+	<i>RsaI</i>	+++	+++	+++	+++
<i>EagI</i>	++	+++	+++	+++	<i>RsrII</i>	+++	++	+	+

continued

Table 3.1.2 Effect of NaCl Concentration on Restriction Endonuclease Activity^a, continued

Enzyme	0 mM	50 mM	100 mM	150 mM	Enzyme	0 mM	50 mM	100 mM	150 mM
	NaCl	NaCl	NaCl	NaCl		NaCl	NaCl	NaCl	NaCl
<i>Eam</i> 1105I	+	++	+	+	<i>Sac</i> I	+++	++	++	+
<i>Eco</i> NI	+++	+++	+++	+++	<i>Sac</i> II	+++	+	+	+
<i>Eco</i> O109	+++	+++	+++	+++	<i>Sal</i> I	+	+	++	+++
<i>Eco</i> 57I	+++	+++	+++	++	<i>Sau</i> 3AI	+++	+++	+++	+++
<i>Sau</i> 96I	+++	+++	+++	+++	<i>Stu</i> I	+++	+++	+++	+++
<i>Sca</i> I	+	+++	+++	++	<i>Sty</i> I	+	++	+++	+++
<i>Scr</i> FI	++	+++	+++	+++	<i>Taq</i> I	+++	+++	+++	++
<i>Sfa</i> NI	+	+	+++	+++	<i>Th</i> 111I	+++	+++	+++	+
<i>Sma</i> I	+	+	+	+	<i>Xba</i> I	+	+++	+++	+++
<i>Sna</i> BI	+++	+++	++	+	<i>Xca</i> I	+++	+	+	+
<i>Spe</i> I	++	+++	+++	++	<i>Xho</i> I	++	+++	+++	+++
<i>Sph</i> I	+	+	+++	+++	<i>Xma</i> I	+++	+++	++	+
<i>Ssp</i> I	++	+++	+++	++	<i>Xmn</i> I	+++	+++	+	+

^aReprinted by permission of New England Biolabs. The activity of each enzyme listed is compared at specified NaCl concentrations to its activity in recommended assay buffer. Recommended assay buffers in some cases differ widely in terms of pH and specific ion requirements. The conditions here varied only the NaCl concentration. All buffers contained 10 mM Tris·Cl (pH 7.5) and 100 µg/ml bovine serum albumin. All incubations were done for 60 min at the optimum temperature for each enzyme. Scoring is as follows:

+ <10% of the activity can be obtained using these conditions compared to the recommended conditions.

++ between 100% and 20% of the activity can be obtained using these conditions compared to the recommended conditions.

+++ between 30% and 100% of the activity can be obtained using these conditions compared to the recommended conditions.

^bNot recommended because of star activity, which refers to cleavage at sites other than the usual recognition sequence. Star activity occurs under nonoptimal reaction conditions, such as low ionic strength, high endonuclease concentrations, high glycerol concentrations, high pH, and when Mn²⁺ is used in place of Mg²⁺.

Table 3.1.3 Recognition Sequences and Reaction Conditions of Commercially Available Methylases^a

Methylase	Recognition sequence ^b	NaCl	Tris·Cl ^c	EDTA	ME ^d
		(mM)			
<i>Alu</i> I	AGC ^m T0	50	10	5	—
<i>Bam</i> HI	GGATC ^m C	50	10	10	5
<i>Cla</i> I	ATCGA ^m T	0	50	10	5
<i>Cp</i> G	C ^m G	50	10	0.1	1 ^e
<i>dam</i>	GA ^m TC	0	50	10	5
<i>Eco</i> RI	GAA ^m TTC	100	100	1	0
<i>Fnu</i> DII	C ^m GCG	0	50	10	5
<i>Hae</i> III	GGC ^m C	50	50	10	1D
<i>Hha</i> I	GC ^m GC	0	50	10	5
<i>Hpa</i> II	CC ^m GG	0	50	10	5
<i>Msp</i> I	C ^m CGG	100	50	10	5
<i>Pst</i> I	CTGCA ^m G	0	50	10	5
<i>Taq</i> I	TCGA ^m	100	10	0	6 ^f

^aAll reaction mixtures contain 80 µM S-adenosylmethionine and are incubated at 37°C.

^bSuperscript m signifies methylated nucleotide.

^cpH 7.5.

^d2-mercaptoethanol or dithiothreitol (D).

^eReaction buffer includes 160 µM S-adenosylmethionine.

^fReaction buffer includes 6 mM MgCl₂.

Table 3.1.4 Cross Index of Recognition Sequences and Restriction Endonucleases^{a,b}

	AATT	ACGT	AGCT	ATAT	CATG	CCGG	CGCG	CTAG	GATC	GCGC	GGCC	GTAC	TATA	TCGA	TGCA	TTAA
↓****	Tsp 509I								DpnII MboI ^c Sau3AI							
*↓***		MaeII				MspI HpaII ^c		BfaI		HinPI		Csp6I		TaqI		MseI
↓			AluI				BstUI		DpnI		HaeIII	RsaI				
***↓*										HhaI						
****↓					NlaIII											
A↓****T	ApoI		HindIII		AflIII	AgeI BsrFI BsaWI	MluI AflIII	SpeI	BglII BstYI						Ppu10I	
A*↓***T		Psp1406I												ClaI BspDI		AseI
A**↓**T				SspI						Eco47III	StuI	ScaI				
A***↓*T																
A****↓T					NspHI					HaeII					NsiI	
C↓****G	MunI				NcoI StyI DsaI	XmaI AvaI	DsaI	AvrII StyI			EagI EaeI	BsiWI	SfcI	XhoI AvaI	SfcI	AflII
C*↓***G				NdeI												
C**↓**G		PmlI BsaAI	PvuII MspA1I			SmaI	MspA1I									
C***↓*G							SacII		PvuI BsiEI		BsiEI					
C****↓G															PstI	
G↓****C	EcoRI ApoI					NgoMI BsrFI	BssHII	NheI	BamHI BstYI	KasI BanI	Bsp120I	Acc65I BanI		SalI	ApaLI	
G*↓***C		BsaHI								NarI BsaHI			AccI	AccI		
G**↓**C			Ecl136II	EcoRV		NaeI				EheI			Bst1107I	HincII		HpaI HincII
G***↓*C																
G****↓C		AatII	SacI BanII BsiHKAI Bsp1286I		SphI NspHI					BbeI HaeII	ApaI BanII Bsp1286I	KpnI				Bsp1286I BsiHKAI
T↓****A					BspHI	BspEI BsaWI		XbaI	BclI		EaeI	BsrGI				
T*↓***A														BstBI		
T**↓**A		SnaBI BsaAI					NruI			FspI	MscI					DraI
T***↓*A																
T****↓A																

^aReprinted by permission of New England Biolabs.

^bSequences at the top of each column are written 5' to 3'. Asterisks at the left of each row are place holders for nucleotides within a recognition sequence, and arrows indicate the point of cleavage. Sequences of complementary strands and their cleavage sites are implied. Enzymes written in bold type recognize only one sequence, while those in light type have multiple recognition sequences.

^cSequence cleaved identically by two or more enzymes that are affected differently by DNA modification at that site (see Table 3.1.1).

reaction (PCR) provides additional versatility for constructing recombinant DNA molecules. *UNIT 3.17* gives a general approach and discusses particular scenarios for incorporating specific sequences onto the ends of DNA fragments, for creating in-frame fusion proteins, and for creating deletions and insertions using inverse PCR.

A final section on specialized applications describes the labeling and various means of detecting nonisotopically labeled probes. Biotinylated and digoxigenin-labeled probes are becoming more widely used in place of their radioactive counterparts for the hybridization applications described in *UNITS 2.9, 4.9 & 14.3*. Aside from eliminating the concerns of working with hazardous materials, the use of these probes offers increased stability and reasonable sensitivity to colorimetric (*UNIT 3.18*) or chemiluminescent (*UNIT 3.19*) detection techniques.

Kevin Struhl

RESTRICTION MAPPING

Construction of an accurate map of the sites where restriction endonucleases cleave DNA—restriction mapping—is critical for almost all subsequent manipulations of genetic material. Restriction mapping is based upon the cleavage of DNA at or near specific nucleotide sequences with restriction endonucleases and the determination of the sizes of the resulting DNA fragments by gel electrophoresis (see *UNITS 2.5* and *2.7*).

Restriction maps are often determined by digesting recombinant DNA molecules in which the DNA of interest is cloned into plasmid or phage vectors. The most common approach is to deduce the map after digesting the DNA of interest with a variety of restriction enzymes, either individually or in combination. This method requires the fewest experimental manipulations, and it is easy to use for cloned DNA segments up to 20 kb in length. Restriction mapping of larger unknown regions gets progressively more difficult. This method is best used for generating an initial restriction map with enzymes that cleave the DNA relatively infrequently.

In an alternative method, DNA fragments are end labeled (*UNITS 3.5* and *3.10*), purified by electrophoresis (*UNITS 2.6* and *2.7*), and subjected to partial endonucleolytic cleavage. This method defines the position of restriction sites with respect to a single defined position, the site of labeling. It is valuable in situations where a restriction endonuclease cleaves the DNA in many positions, or in mapping very large DNA molecules.

It is important to note that careful restriction mapping is critical for subcloning (*UNIT 3.16*) and is useful for nucleotide sequence analysis (*UNIT 7.4*). Conversely, nucleotide sequence data can predict restriction endonuclease cleavage sites and thereby generate a complete restriction map.

Mapping by Multiple Endonuclease Digestions

UNIT 3.2

The DNA of interest is cleaved with a variety of restriction endonucleases, either individually or in combination, and the resulting products are separated by agarose or acrylamide gel electrophoresis. By determining the sizes of DNA fragments produced by endonuclease cleavage, the restriction map is deduced progressively from simple situations where enzymes cleave the DNA once or twice to more complex situations where cleavage occurs more frequently.

BASIC
PROTOCOL

Materials

Restriction endonucleases (Table 3.1.1)

10× restriction endonuclease buffers (*UNIT 3.1*)

Additional reagents and equipment for agarose or polyacrylamide gel electrophoresis (*UNITS 2.5* or *2.7*) and DNA molecular weight markers (*UNITS 2.5A*)

1. In each of several separate tubes, cleave the DNA of interest (up to 20 kb in length) with a restriction endonuclease that can be expected to cleave infrequently.

In general, inexpensive restriction endonucleases with six base-pair recognition sequences are used first (e.g., EcoRI, HindIII, BamHI, PstI, KpnI, XbaI, Sall, and XhoI). DNA fragments produced by cleavage with these enzymes are also easily subcloned into standard plasmid or phage vectors.

2. After cleavage, remove a small portion of each reaction mixture and separate the products by agarose or polyacrylamide gel electrophoresis. Store the remainders of the reaction mixtures on ice.

It is essential that the DNA is completely cleaved by the restriction endonuclease. See critical parameters.

Enzymatic
Manipulation
of DNA and RNA

Contributed by Kenneth D. Bloch

Current Protocols in Molecular Biology (1995) 3.2.1-3.2.5

Copyright © 2000 by John Wiley & Sons, Inc.

3.2.1

- Using DNA fragments of known lengths for comparison (see *UNIT 2.5* for the preparation of size standards), calculate the lengths of all the fragments. Determine the number of cleavage sites for each restriction endonuclease.

A restriction endonuclease cleaving a DNA molecule N times will generate N restriction fragments if the molecule is circular and $N + 1$ fragments if the molecule is linear.

- For each sample from step 1, transfer an aliquot into separate tubes and add a different restriction endonuclease to each tube. In this way, the DNA in each tube will be digested by two enzymes. One aliquot from each sample should not be cleaved with a second enzyme.

Obviously, it is wasteful and unnecessary to perform secondary digestions with enzymes that do not cleave the DNA (as determined from step 3).

- Using agarose gel electrophoresis, compare DNA fragments resulting from digestion with the two restriction enzymes with the fragments resulting from digestion with the first enzyme alone and with the second enzyme alone. Calculate the lengths of all the restriction fragments.
- As more restriction enzymes are used to digest the DNA and the resulting restriction fragments are analyzed, an unambiguous, internally consistent map of restriction endonuclease cleavage sites can be determined. See commentary and examples.

COMMENTARY

Critical Parameters

Completeness of digestion

When digestion is complete, the cleavage products should be present in equimolar quantities. If DNA is visualized by staining with ethidium bromide (*UNIT 2.5*), band intensity is proportional to fragment length. For example, an 8-kb fragment stains twice as intensely as a 4-kb fragment. Bands that appear fainter than expected for their molecular weight are probably the result of incomplete cleavage. Bands that appear more intense than expected are probably due to the presence of two (or more) DNA fragments of similar length. If DNA is end labeled at the 5' (*UNIT 3.10*) or 3' (*UNIT 3.5*) termini and visualized by autoradiography, all radiolabeled restriction fragments should produce bands of equal intensity. If DNA is uniformly labeled (such as by nick translation as described in *UNIT 3.5*), band intensities are proportional to fragment length.

Internal consistency of the results

Sum of fragment lengths. It is obvious that the sum of the lengths of the restriction fragments should equal the total length of the DNA molecule. In other words, for each enzyme that is tested, the sum of all the fragments should be equal.

If, for a given digestion, the sum is lower than expected, check carefully for fragments of similar length. Such “comigrating” fragments can be identified by their increased band intensity (two comigrating fragments should produce a band that is twice as intense as expected from the molecular weight) or by band “broadening” or “fuzziness” (due to similar but distinct electrophoretic mobilities of two fragments). Alternatively, a low sum of fragment lengths could be due to the existence of several small fragments, which may be difficult to visualize in the agarose gel or which may have “run off” the bottom of the gel.

If the sum is larger than expected, the most likely explanation is that some of the bands arise from incomplete digestion. Another possibility, less likely in general, is that the DNA being tested is actually a mixture of DNAs. Finally, some DNA fragments have anomalous electrophoretic mobilities (especially in acrylamide gels), usually because of unusual bending properties dictated by specific DNA sequences.

Number of bands. As mentioned, the number of DNA fragments produced in a given reaction is directly related to the number of times that the restriction enzyme cleaves the DNA. Thus, it follows that when DNA is cleaved

by two (or more) enzymes, the number of fragments should be the sum of the fragments generated by the individual enzymes. For example, if cleavage of a circular molecule by enzyme A generates three fragments (indicative of three sites) and cleavage by enzyme B generates five fragments (indicative of five sites), cleavage by enzymes A + B should generate eight fragments. If fewer fragments are seen, it means either that some fragments are comigrating or that very small DNA fragments have been generated due to the proximity of A and B sites.

Additional comments

If the DNA to be studied is subcloned into a plasmid or phage vector for which the restriction map is known, restriction mapping of the DNA insert is facilitated by digestion with a series of restriction endonucleases that cleave the vector at least once. In this way, restriction sites in the DNA are mapped with respect to known sites in the vector (see Example 3.2.2).

Utilization of nucleic acid hybridization probes (UNIT 3.5) and of the Southern transfer technique (UNIT 2.9) permits restriction mapping of a small segment of DNA within a larger fragment or even within total cellular DNA.

Anticipated Results

When sufficient enzyme digestions are performed, it should be possible to obtain an unambiguous and internally consistent map.

Time Considerations

Depending on the resolution that is desired, it will take 1 to 10 days to generate a restriction map. Additional mapping can be done at any time.

EXAMPLES

Example 3.2.1: Restriction Mapping a Plasmid of Unknown Structure

Consider a 9-kb circular plasmid DNA, pPROTO, with an unknown restriction map. Cleavage with restriction endonuclease A or with enzyme B generates a 9-kb linear DNA molecule, thus indicating that each enzyme recognizes a single site in pPROTO. To determine the location of enzyme A's cleavage site with respect to enzyme B's cleavage site, pPROTO is digested with both enzymes. Two restriction fragments, 4 kb and 5 kb in size, are generated. Therefore, these endonucleases cleave sequences that are located 4 kb apart.

Restriction endonuclease C cleaves

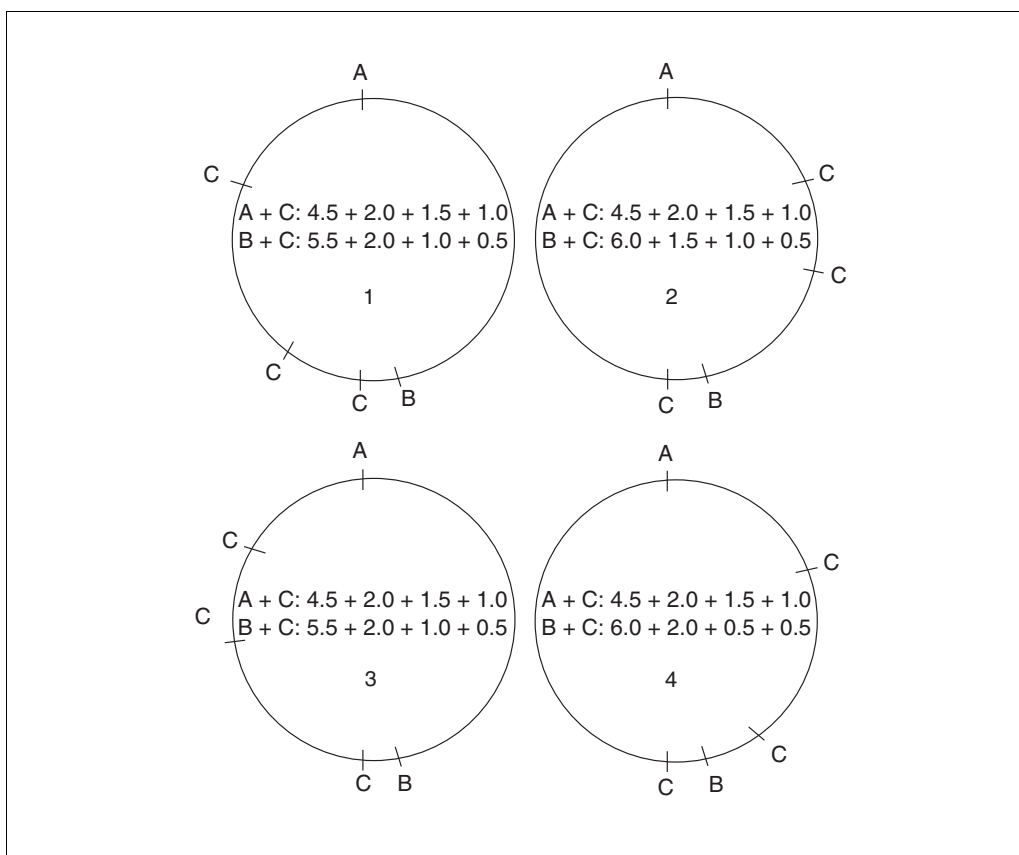


Figure 3.2.1 Restriction mapping of pPROTO (see Example 3.2.1).

pPROTO DNA at three sites. Gel electrophoresis reveals three fragments that are 1 kb, 2 kb, and 6 kb in length. When pPROTO is digested with enzymes A and C, four bands are seen, as expected. The 6-kb fragment generated by enzyme C alone “disappears” and is replaced by two more rapidly migrating fragments 1.5 and 4.5 kb in size, whereas the 1- and 2-kb fragments produced by endonuclease C are unchanged. Thus, the A site is localized within the 6-kb enzyme C fragment.

At this stage, there are four possible restriction maps (Fig. 3.2.1). Within the 6 kb fragment generated by enzyme C, the A site could be located in two possible positions (1.5 kb from either of the two possible ends), and the 1- and 2-kb C fragments could be located in two possible orders with respect to the 6-kb fragment. The correct map can be determined from the results obtained upon cleavage with enzymes B + C. The predictions for each of the possible maps are shown in Figure 3.2.1.

To map restriction sites for endonuclease D, the same logic is followed. The process is greatly simplified once the restriction sites for endonucleases A, B, and C have been mapped as described above. In some cases, ambiguities

in positions of D sites can be resolved by appropriate triple digests.

The general rule is to build up the restriction map from simple situations where enzymes cleave once or twice in the DNA to more complex situations where cleavage occurs more frequently.

Example 3.2.2: Restriction Mapping a Cloned DNA Insert within a Plasmid of Known Structure

Frequently, restriction mapping is performed on a region of DNA within the context of surrounding DNA for which the nucleotide sequence and/or restriction map is known. Consider the following approach to mapping pCDNA, which contains an insert of cDNA that has been ligated into the plasmid pBR322 at the *Pst*I site. pBR322, a 4363-bp *E. coli* plasmid vector, is cleaved once by *Pst*I at a site that is 754 bp from the unique *Eco*RI cleavage site (UNIT 1.5). Utilization of the known restriction map for pBR322 permits rapid restriction mapping of the cDNA insert.

Cleavage of pBR322 with *Pst*I generates a single 4363-bp DNA fragment. Cleavage of single pCDNA with *Pst*I generates two fragments,

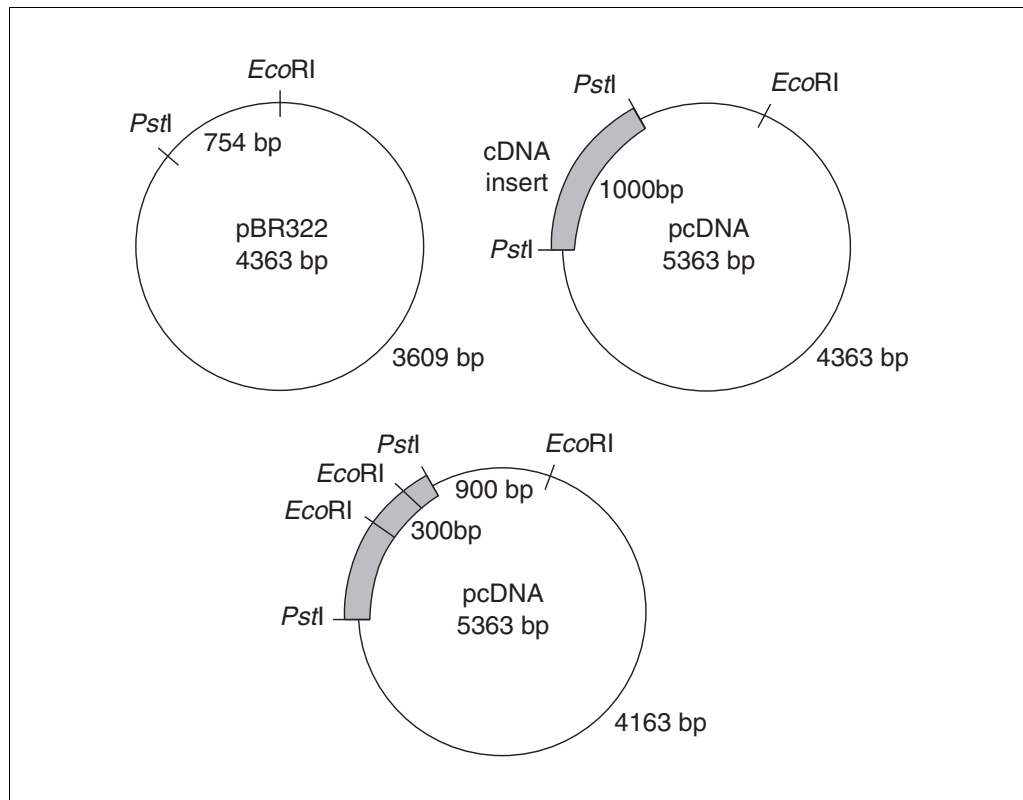


Figure 3.2.2 Restriction mapping pCDNA (see Example 3.2.2).

4363 and 1000 bp in size, demonstrating that the cDNA insert is 1 kb in size (Fig. 3.2.2).

Cleavage of pBR322 with *EcoRI* generates a single 4363-bp DNA molecule. Cleavage of pCDNA generates three fragments, 300, 900, and 4163 bp in size. Knowledge of the location of the cDNA insert (cloned into the *PstI* site) with respect to the vector *EcoRI* site allows one to conclude that the cDNA insert has two *EcoRI* cleavage sites that are located 146 and 446 bp from one end of the cDNA.

Similarly, parallel digestions of the recombinant DNA molecule and the cloning vector, using a panel of restriction endonucleases, permit efficient restriction mapping of the DNA of interest.

Contributed by Kenneth D. Bloch
Harvard Medical School
Boston, Massachusetts

Mapping by Partial Endonuclease Digestions

BASIC PROTOCOL

A DNA fragment radiolabeled at one of its two ends is purified by gel electrophoresis and subjected to partial cleavage by a restriction endonuclease. Analysis of the resulting products by polyacrylamide (or agarose) gel electrophoresis enables one to define the distance of restriction sites from the labeled end.

Materials

Reagents and equipment for digesting DNA with restriction enzymes (UNIT 3.1), labeling DNA (UNIT 3.5 or 3.10), agarose or polyacrylamide gel electrophoresis (UNIT 2.5 or 2.7), and autoradiography (APPENDIX 3)

1. Cleave the DNA fragment of interest with a restriction endonuclease that can be expected to cleave infrequently.
2. "End label" the products of digestion with ^{32}P . 5' termini can be labeled by successive treatment with calf intestine phosphatase and T4 polynucleotide kinase (UNIT 3.10), and 3' termini can be labeled with the Klenow fragment of *E. coli* DNA polymerase I (UNIT 3.5) or with T4 DNA polymerase (UNIT 3.5).

Each restriction fragment is labeled at both ends.

3. Cleave fragments with a second restriction endonuclease and fractionate the products by gel electrophoresis.
4. Purify the DNA fragments that are now radiolabeled at one end. See UNIT 2.7 and 2.6, respectively, for methods of isolating DNA from acrylamide or agarose gels.

Only the fragments generated by cleavage with both restriction endonucleases will be labeled at one end. Fragments from step 2 that are not cleaved by the second enzyme will retain the label at both ends and hence be useless for the subsequent analysis.

5. Partially digest one of the isolated DNA fragments with a restriction enzyme which cleaves relatively frequently (usually with a 4-bp recognition site).

Because some sites are cleaved more readily than others, it is useful to generate a series of enzyme reactions in which the DNA is cleaved to various extents. This is easily accomplished by serially diluting the enzyme, as described in UNIT 3.1.

6. Separate the products of the partial digestion by gel electrophoresis (usually polyacrylamide) and visualize the bands by autoradiography.
7. Determine the size of the restriction fragments by comparison with radiolabeled DNA fragments of known size.

MspI-cleaved pBR322 DNA (UNIT 1.5) that has been end labeled is an excellent size standard for fragments less than 650 bp in length.

COMMENTARY

Background Information

This technique for partial digestion of DNA fragments radiolabeled at one end is rapid and relatively simple for fine structure restriction mapping. Unlike restriction mapping by multiple endonuclease cleavage (UNIT 3.2), the analysis is not complicated by the presence of many cleavage sites for a given enzyme. However, it can be complicated if

some sites are cleaved much better than others.

If the restriction fragment to be mapped has been generated by two restriction endonucleases with different cohesive termini, it may be possible to radiolabel one end of the fragment by "filling in" with the appropriate [α - ^{32}P]NTP. For example, a fragment that has been generated by *EcoRI* (5' overhang, AATT) and *BamHI* (5' overhang, GATC) can

be radiolabeled, exclusively, at the *Bam*HI site by using [α - 32 P]GTP in the reaction.

Critical Parameters

High salt concentrations can alter the electrophoretic mobility of DNA fragments (especially small fragments). When performing fine structure restriction mapping, it may be necessary to “desalt” the sample by ethanol precipitation and by washing with 70% ethanol.

The major artifact associated with this method is that recognition sites for a given restriction endonuclease are not always cleaved with equal efficiency. This problem becomes more severe with increased distance of the “inefficient” site to the labeled site because cleavage at more “efficient” sites becomes more likely to eliminate the diagnostic fragment. Thus, restriction maps generated by this method are occasionally “missing” a site. A standard restriction digest using unlabeled DNA is very useful for confirming the positions of cleavage.

It is relatively difficult to map cleavage sites that are located close to either of the two ends. Sites near the unlabeled end will be very similar in size to the undigested fragment, whereas sites near the labeled end will be extremely small. These problems can be resolved by choosing different fragments and/or by using appropriate gels to visualize small DNA fragments.

Anticipated Results

By varying the amount of enzyme to produce different degrees of partial digestion, a series of bands should be generated. Each band represents the location of a restriction site with respect to the labeled site.

Time Considerations

Preparation of the labeled DNA fragment should take about 1 to 2 days, and the remainder of the procedure (partial cleavage, gel electrophoresis, and autoradiography) should take an additional 1 to 5 days. The locations of restric-

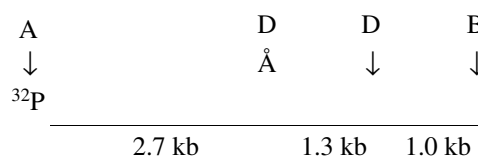
tion sites for 5 to 20 different enzymes can easily be determined from the same preparation of the labeled DNA fragment.

EXAMPLES

Example 3.3.1: Mapping Restriction Sites by Partial Cleavage of End-Labeled DNA.

Plasmid pPROTO (see sketch below) is digested with enzyme A, and the ends of the linearized molecule are radiolabeled. The DNA is then digested with enzyme B, and the 5-kb fragment is isolated.

The uniquely end-labeled fragment is partially digested with enzyme D, and the products are fractionated by agarose gel electrophoresis. A “ladder” of radiolabeled DNA fragments, 2.7, 4.0, and 5.0 kb in size, is evident by autoradiography. This demonstrates conclusively that enzyme D cleaves pPROTO DNA at sites that are 2.7 and 4 kb from enzyme A’s cleavage site as shown below.



This result can be confirmed by reversing the order of restriction enzyme cleavage. When the 32 P label is at the B cleavage site, partial digestion should produce bands of 1.0, 2.3, and 5.0 kb.

Key References

- Boseley, P.G., Moss, T., and Birnstiel, M.L. 1980. 5' labeling and poly(dA) tailing. *Meth. Enzymol.* 65:478-494.
- Danna, A.J. 1980. Determination of fragment order through partial digests and multiple enzyme digests. *Meth. Enzymol.* 65:449-467.

Contributed by Kenneth D. Bloch
Harvard Medical School
Boston, Massachusetts

ENZYMES FOR MODIFYING AND RADIOACTIVELY LABELING NUCLEIC ACIDS

This section describes enzymes other than restriction endonucleases that are used to manipulate nucleic acids. An increasing number of DNA polymerases, RNA polymerases, and exonucleases, each with unique properties, are becoming available to the molecular biologist. A basic knowledge of these differences helps one to choose the optimal enzyme for a given procedure.

Increasingly, commercially produced enzymes are purified from strains containing cloned genes that have been designed to overproduce the enzyme. These enzymes offer several advantages. First, they are significantly less expensive. Second, enzyme purification procedures are less extensive and more rapid, resulting in more homogenous enzymes with higher specific activities. Finally, for enzymes normally encoded by bacteriophages, the use of cloned genes eliminates contamination by other nucleic acid-metabolizing enzymes that are encoded by the phage genome.

IMPORTANT NOTE: Enzymes are expensive and sensitive reagents that are easily inactivated or destroyed by careless use. Enzymes should always be stored at -20°C in freezers that maintain a constant temperature. For use, the tube containing the enzyme should be placed quickly on ice, and the enzyme should be removed with a clean pipet tip (sterility is not essential, but can be helpful). After the enzyme is removed, the tube should be immediately returned to -20°C . Do not let enzymes stay for extended periods on ice and do not expose them to temperatures greater than 0°C .

UNIT 3.4

Reagents and Radioisotopes Used to Manipulate Nucleic Acids

STOCK SOLUTIONS

It is convenient to have preparations of the following stock solutions on hand for the enzyme reactions and applications described in *UNITS 3.5 through 3.15* and in subsequent chapters of the manual. See *APPENDIX 2* for recipes.

Solutions should be prepared in deionized water and stored in aliquots. Most of the solutions are stable for years at -20°C or at room temperature if kept sterile. Dithiothreitol should be stored at -20°C and is stable for months. DTT should not be autoclaved, and it should be kept at 0°C during use.

Acid precipitation solution	1 M NaCl
10 mg/ml autoclaved gelatin	5 M NaCl
10 mg/ml bovine serum albumin (BSA), Pentax Fraction V	10 mg/ml sonicated salmon or herring sperm DNA
Buffered phenol (<i>UNIT 2.1</i>)	500 $\mu\text{g}/\text{ml}$ sonicated salmon or herring sperm DNA
0.1 M dithiothreitol (DTT)	Standard enzyme diluent (SED; <i>APPENDIX 2</i>)
10 mg/ml <i>E. coli</i> tRNA	TE buffer (<i>APPENDIX 2</i>)
0.5 M EDTA, pH 8.0	1 M Tris-Cl, pH 7.5
Ethanol	1 M Tris-Cl, pH 8.0
Glycerol	
1 M KCl	
0.2 M MgCl_2	

Reagents and Radioisotopes Used to Manipulate Nucleic Acids

3.4.1

Contributed by Kevin Struhl

Current Protocols in Molecular Biology (1993) 3.4.1-3.4.11

Copyright © 2000 by John Wiley & Sons, Inc.

10× ENZYME BUFFERS

Listed below are recipes for 10× buffers (or 5× buffer for *Bal* 31 nuclease). Nucleoside triphosphates should not be included in the 10× or 5× buffers because the high concentration of Mg⁺⁺ ions will lead to the formation of insoluble complexes. Bovine serum albumin (BSA) or gelatin are not essential for activity, but they help stabilize the enzymes. See also Table 3.4.1 which follows, listing 1× reaction conditions.

***Bal* 31 nuclease (5×)**

0.25 M Tris·Cl, pH 8.0
50 mM MgCl₂
50 mM CaCl₂
3 M NaCl
0.25 mg/ml BSA or gelatin

***BAP* (bacterial alkaline phosphatase)**

0.5 M Tris·Cl, pH 8.0
10 mM ZnCl₂

***CIP* (calf intestine phosphatase)**

0.2 M Tris·Cl, pH 8.0
10 mM MgCl₂
10 mM ZnCl₂
0.5 mg/ml BSA or gelatin

***DNase I* (deoxyribonuclease I)**

0.5 M Tris·Cl, pH 7.5
0.1 M MgCl₂ (single-strand breaks; or 0.1 M MnCl₂, double-strand breaks)
0.5 mg/ml BSA

***E. coli* DNA ligase**

400 mM Tris·Cl, pH 8
0.1 M MgCl₂
50 mM DTT
0.5 mg/ml BSA

***E. coli* DNA polymerase I or**

***Klenow* fragment**

0.5 M Tris·Cl, pH 7.5
0.1 M MgCl₂
10 mM DTT
0.5 mg/ml BSA or gelatin

***E. coli* RNA polymerase**

0.4 M Tris·Cl, pH 8.0
0.1 M MgCl₂
50 mM DTT
0.5 M KCl
0.5 mg/ml BSA or gelatin

***Exonuclease III* (*exo III*)**

0.5 M Tris·Cl, pH 7.5
50 mM MgCl₂
50 mM DTT
0.5 mg/ml BSA or gelatin

***Exonuclease VII* (*exo VII*)**

0.7 M Tris·Cl, pH 8.0
80 mM EDTA
0.1 M β-mercaptoethanol
0.5 mg/ml BSA or gelatin

***Klenow* fragment (see *E. coli* DNA polymerase I)**

λ *exonuclease*

0.7 M glycine·KOH, pH 9.4
25 mM MgCl₂
0.5 mg/ml BSA or gelatin

***Mung bean* nuclease**

0.3 M sodium acetate, pH 5.0
0.5 M NaCl
10 mM zinc acetate
0.5 mg/ml BSA or gelatin

***Poly(A)* polymerase**

400 mM Tris·Cl, pH 8.0
0.1 M MgCl₂
25 mM MnCl₂
2.5 M NaCl
0.5 mg/ml BSA

Reverse transcriptase

0.5 M Tris·Cl, pH 8.0
50 mM MgCl₂
50 mM DTT
0.5 M KCl
0.5 mg/ml BSA or gelatin

***RNase H* (*ribonuclease H*)**

0.2 M HEPES·KOH, pH 8.0
0.5 M KCl
40 mM MgCl₂
10 mM DTT
0.5 mg/ml BSA or gelatin

***S1* nuclease**

0.5 M sodium acetate, pH 4.5
10 mM zinc acetate
2.5 M NaCl
0.5 mg/ml BSA or gelatin

Sequenase (see T7 DNA polymerase)

SP6 RNA polymerase

0.4 M Tris·Cl, pH 7.5
0.1 M MgCl₂
50 mM DTT
0.5 mg/ml BSA or gelatin

T3 RNA polymerase (see SP6 polymerase)

T4 DNA ligase

0.5 M Tris·Cl, pH 7.5
50 mM MgCl₂
50 mM DTT
0.5 mg/ml BSA or gelatin

T4 DNA polymerase

0.5 M Tris·Cl, pH 8.0
50 mM MgCl₂
50 mM DTT
0.5 mg/ml BSA or gelatin

T4 polynucleotide kinase

0.5 M Tris·Cl, pH 7.5 (forward reaction) or 500 mM imidazole·Cl, pH 6.6 (exchange reaction)
0.1 M MgCl₂
50 mM DTT
0.5 mg/ml BSA or gelatin

T4 RNA ligase

0.5 M HEPES, pH 8.3
100 mM MgCl₂
50 mM DTT
0.5 mg/ml BSA

T7 DNA polymerase (native and modified)

0.4 M Tris·Cl, pH 8.0
0.1 M MgCl₂ (native or chemically modified; or 50 mM MgCl₂, genetically modified)
50 mM DTT
0.5 M NaCl
0.5 mg/ml BSA or gelatin

T7 gene 6 exonuclease (see exonuclease III)

T7 RNA polymerase (see SP6 polymerase)

Taq DNA polymerase

0.1 M Tris·Cl, pH 8.4
x MgCl₂
500 mM KCl
1 mg/ml gelatin

The concentration (x) of MgCl₂ depends upon the sequence and primer set of interest. The optimal concentration is first determined empirically using a 10× amplification buffer minus MgCl₂. Based upon these results, a 10× amplification buffer containing MgCl₂ at the optimal concentration is prepared (UNIT 15.1).

Terminal transferase

1 M sodium cacodylate, pH 7.0
10 mM CoCl₂
1 mM DTT
0.5 mg/ml BSA or gelatin

ENZYME REACTION CONDITIONS AND APPLICATIONS

General guidelines for enzyme reaction conditions are provided in Table 3.4.1. Because many parameters vary depending on the application, *UNITS 3.5-3.15*, which describe specific classes of enzymes, should be consulted for further details.

Some applications of nucleic acid-modifying enzymes are indicated in Table 3.4.2. For some applications, more than one enzyme can be used to carry out the reaction. More detailed information is provided in *UNITS 3.5-3.15*.

Table 3.4.1 General Guidelines for Enzyme Reaction Conditions

Many parameters, including reaction volume, dNTP concentration, time, and temperature, will vary depending on application. Dilute appropriate 10× or 5× enzyme buffer (see recipes preceding this section) to 1× and set up reactions as described in table. Stop reactions at 75°C for 10 min, or by adding 2 μl of 0.5 M EDTA (except where indicated otherwise under “Comments”). Enzymes are described in further detail in the units indicated in parentheses.

Enzyme	Rxn. vol. (μl)	Amt. DNA (μg)	Amt. enz. (U)	dNTPs (mM each)	Rxn. temp. (°C)	Time (min)	Comments
DNA Ligases							
<i>E. coli</i> DNA ligase (3.14)	50	1	10	100 NAD	10-25	2-16 hr	
T4 DNA ligase (3.14)	50	1	1	500 ATP	12-30	1-16	
DNA Polymerases							
<i>E. coli</i> DNA polymerase I (3.5)	25	1	3	20	20-37	15-30	<i>a</i>
T4 DNA polymerase (3.5)	50	2	5	100	11	20	
T7 DNA polymerase (3.5)	50	2	5	300	37	20	
Modified T7 DNA polymerase (3.5)	50	2	10	300	37	20	
<i>Taq</i> DNA polymerase (3.5 & 3.17)	100	0.1-1	2.5	200	94, 55, 72	1-2	<i>a, b</i>
Exonucleases							
Exonuclease III (3.11)	50	2	10	—	37	1-30	
Exonuclease VII (3.11)	50	1	0.2	—	37	30	<i>a</i>
Lambda exonuclease (3.11)	50	2	10	—	37	1-30	
T7 gene 6 exonuclease (3.11)	50	2	5	—	37	1-30	
Kinases							
T4 polynucleotide kinase (3.10)							
5′ labeling (forward)	30	1-50 pmol ^c	20	50 ^d	37	60	<i>a</i>
oligonucleotides	30	1-10 ^c	20	1000 ATP	37	60	<i>a</i>
5′ labeling (exchange)	30	1-50 pmol ^c	20	60 ^d	37	60	<i>a, e</i>
Nucleases							
Bal 31 nuclease (3.12)	50	2	10	—	30	1-30	<i>a</i>
Deoxyribonuclease I (DNase I) (3.12)	100	2	1	—	37	1-30	<i>a</i>
Mung bean nuclease (3.12)	100	1	15	—	37	30	<i>a</i>
S1 nuclease (3.12)	100	2	10	—	37	30	<i>a</i>
Phosphatases							
BAP (3.10)	50	1-20 pmol ^f	0.1	—	60	30	<i>a</i>
CIP (3.10)	50	1-20 pmol ^f	0.1	—	37	30	<i>a</i>
RNA Modifying Enzymes							
<i>E. coli</i> RNA polymerase (3.8)	50	2 ^g	10	300 NTPs	37	30	
SP6, T7, T3 RNA polymerases (3.8)	50	2 ^g	10	400 NTPs	37	30	<i>h</i>
Poly(A) polymerase (3.9)	50	12.5 RNA	5	250 ATP	37	30	
Reverse transcriptase (3.7)	50	1 mRNA	40	40	37	30	<i>i</i>
RNase H (3.13)	100	2 RNA:DNA	1	—	37	20	<i>a</i>
T4 RNA ligase (3.15)	50	2 ^g	1	2000 ATP	17	10 hr	<i>a</i>
Terminal transferase (3.6)	50	4 ^f	10	20	37	30	

^aStop reactions as follows: DNase I—5 μl 0.5 M EDTA; *E. coli* DNA polymerase—75°C for 10 min or 1 μl 0.5 M EDTA; *Taq* DNA polymerase—store at -20°C; Exonuclease VII—phenol extract and ethanol precipitate; T4 polynucleotide kinase—1 μl 0.5 M EDTA, then phenol extract and ethanol precipitate; Bal 31 nuclease—75°C for 10 min or 5 μl 0.5 M EDTA; S1 nuclease, mung bean nuclease, and RNase H—1 μl 0.5 M EDTA; BAP—add SDS to 0.1% and proteinase K to 100 μg/ml, incubate at 37°C for 30 min, then phenol extract twice and ethanol precipitate; CIP—75°C for 10 min or phenol extract and ethanol precipitate; T4 RNA ligase—2 μl 0.5 M EDTA.

^bDNA is template (genomic); add 0.2-1 μM each oligonucleotide primer.

^cForward—dephosphorylated DNA (5′ ends); oligonucleotide—linkers; exchange—phosphorylated DNA (5′ ends).

^dpmol[γ-³²P]ATP, specific activity >3000 Ci/mmol (forward—150 μCi; exchange—180 μCi).

^eAdd 5 mM ADP and use imidazole in the buffer instead of Tris.

^fBAP and CIP—pmol DNA termini; terminal transferase—pmol DNA as 3′ termini.

^gRNA polymerases—use DNA with a promoter; T4 RNA ligase—use single-stranded DNA or RNA.

^hAdd 1 mM spermidine for SP6.

ⁱAdd 100 μCi [α-³²P]dNTP, specific activity >400 Ci/mmol, and 1 μg oligo(dT)₁₂₋₁₈.

3.4.4

Table 3.4.2 Applications of Nucleic Acid–Modifying Enzymes
Enzymes are described in further detail in the units indicated in parentheses.

Application	Enzyme
Blunt-end generation	
by removal of 3' protruding ends	T4 and T7 DNA polymerase (3.5)
by filling in 3' recessed ends	T4 and T7 DNA polymerase (3.5); Klenow fragment (3.5)
cDNA synthesis	
from RNA	Reverse transcriptase (3.7)
of second strand	Klenow fragment (3.5)
Cloning of DNA fragments	CIP (3.10); T4 DNA ligase (3.14); Klenow fragment (3.5); restriction endonucleases (3.1)
Degradation of DNA	
Exonucleases—double strand	
5'→3'	T7 gene 6 and λ exonuclease (3.11)
3'→5'	Exonuclease III (3.11)
5'→3' and 3'→5'	<i>Bal</i> 31 nuclease (3.12)
Exonuclease—single strand (3' & 5')	Exonuclease VII (3.11)
Endonucleases—double strand	
nonspecific	DNase I and micrococcal nucleases (3.12)
specific	Restriction endonucleases (3.1)
Endonucleases—single strand	S1, <i>Bal</i> 31, and mung bean nucleases (3.12)
Degradation of RNA	Ribonuclease A and/or T1 (3.13); micrococcal nuclease (3.12)
Degradation of RNA in RNA:DNA duplexes	Ribonuclease H (3.13) BAP and CPI (3.10)
DNA sequencing	Modified T7 DNA polymerase (3.5); <i>Taq</i> polymerase (3.5); Klenow fragment (3.5); T4 polynucleotide kinase (3.10); reverse transcriptase (3.7)
Intron mapping	Exonuclease VII (3.11); S1 nuclease (3.12 & 4.6)
Labeling DNA at 5' ends	T4 polynucleotide kinase (3.10)
Labeling DNA at 3' ends	Modified T7 DNA polymerase (3.5); Klenow fragment (3.5); reverse transcriptase (3.7); terminal transferase (3.6)
Labeling RNA at 3' ends	Poly(A) polymerase (3.9); T4 RNA ligase (3.15)
Ligations of DNA	<i>E. coli</i> and T4 DNA ligase (3.14)
Ligations of RNA	T4 RNA ligase (3.15)
Mapping mRNA 5' ends	S1 or mung bean nuclease (3.12 & 4.6); reverse transcriptase (3.7 & 4.8); ribonuclease A (4.8)
Methylation of DNA	Methylases (3.1)
Nested deletion construction	<i>Bal</i> 31 (3.12 & 7.2); exonuclease III (3.11 & 7.2) + S1 or mung bean nuclease (3.12 & 7.2)
Nick translation	DNase I (3.12); <i>E. coli</i> DNA polymerase I (3.5)
Oligonucleotide extension	T7 DNA polymerase (3.5); Klenow fragment (3.5)
Phosphorylation	T4 polynucleotide kinase (3.10)
Poly(A) tailing of RNA	Poly(A) polymerase (3.9)
Polymerase chain reaction (PCR)	<i>Taq</i> DNA polymerase (3.5 & 3.17)
Primed synthesis	Klenow fragment (3.5)
Replacement synthesis	T4 and T7 DNA polymerase (3.5)
Restriction mapping	Restriction endonucleases (3.1)
Tailing DNA	Terminal transferase (3.6)
Tailing RNA	Poly(A) polymerase (3.9)
Transcription, promoter-specific	<i>E. coli</i> , SP6, T3, and T7 RNA polymerase (3.8)

3.4.5

NUCLEOSIDE TRIPHOSPHATES

Nucleoside triphosphates (also referred to as *ribonucleoside* triphosphates) have a limited shelf life in solution and hence should be stored as aliquots at -20°C , where they are stable for up to 1 year. They should be HPLC-(or comparably) purified. Pharmacia is the recommended supplier. Deoxyribonucleoside triphosphates (dNTPs) or ribonucleoside triphosphates (NTPs) can be purchased as ready-made 100 mM solutions, which is the preferred method of shipping and storage. Alternatively, they can be purchased in lyophilized form and prepared in deionized water as follows:

1. Dissolve in water to an expected concentration of 30 mM; adjust to pH 7 with 1 M NaOH.

dNTPs and NTPs will undergo acid catalyzed hydrolysis unless they are neutralized.

2. Determine the actual concentration of each dNTP and NTP spectrophotometrically, using the extinction coefficients given in Table 3.4.3 (see also APPENDIX 3).
3. Prepare 5 mM working solutions for each NTP and dNTP from concentrated stocks of nucleoside triphosphates.
4. Prepare NTP and dNTP solutions (mixes) containing equimolar amounts of all four RNA or DNA precursors as follows:

5 mM 4NTP mix: 5 mM each of ATP, UTP, CTP, GTP

5 mM 4dNTP mix: 5 mM each of dATP, dTTP, dCTP, dGTP

0.5 mM 4dNTP mix: 0.5 mM each of dATP, dTTP, dCTP, dGTP

5. For radioactive labeling purposes, prepare stocks lacking one particular NTP or dNTP but containing equimolar amounts of the remaining three precursors:

5 mM 3NTP mix (minus UTP): 5 mM of ATP, GTP, CTP for radiolabeling RNA with radioactive UTP

5 mM and 0.5 mM 3dNTP mixes (minus dATP): 5 mM and 0.5 mM each of dTTP, dCTP, dGTP for radiolabeling DNA with radioactive ATP.

Other 3NTP or 3dNTP mixes are made by omitting the precursor corresponding to the selected radiolabel. Various alternate nomenclatures exist for identifying the precursors that are present (or absent) in these solutions. For example, in the instances above, 5 mM 3NTP mix (minus UTP) might be referred to as 5 mM (A,C,G)TP mix. Similarly, 5 mM 3dNTP mix (minus dATP) might be named 5 mM d(C,G,T)TP mix. Whatever the nomenclature, remember that the molarity of the mix refers to the concentration of each precursor present in the solution.

Table 3.4.3 Properties of the Nucleoside Triphosphates

Nucleoside triphosphate	λ max (pH 7.0)	ϵ max $\times 10^{-3}$ (pH 7.0)	Absorbance ratio (280/260) (pH 7.0)	pK _a of base
ATP	259	15.4	0.15	4.0
CTP	271	12.8 ^a	0.97	4.8
GTP	252	13.7	0.66	3.3
				9.3
UTP	262	10.0	0.38	9.5
dATP	259	15.4	0.15	3.6
dCTP	280	13.1 ^a	0.98	4.3
dGTP	253	13.7	0.66	3.5
				9.7
dTTP	267	9.6	0.73	9.3

^aPerform the spectral analysis at pH 2.0.

RADIOISOTOPES FOR LABELING NUCLEIC ACIDS

^{32}P is the most common isotope used to radioactively label nucleic acids. It emits energetic β particles, it has the highest specific activity (9200 Ci/mmol in pure form), and it has a fairly short half-life (14 days). ^{32}P -labeled NTP or dNTP precursors can be purchased at various specific activities. In general, α -labeled precursors are purchased at 400 to 800 Ci/mmol, whereas γ -labeled precursors are purchased at 3000 to 7000 Ci/mmol. ^{32}P is preferred for preparation of highly radioactive probes and for most autoradiographic procedures. Its strong energy gives optimal sensitivity, especially when autoradiography is carried out with intensifying screens. In addition, enzymes are not inhibited by ^{32}P -labeled nucleoside triphosphates because their structures are virtually identical to nonradioactive nucleoside triphosphates.

^{35}S emissions are less energetic, the specific activity is several-fold lower (1500 Ci/mmol in pure form), and it has a longer half-life (87 days; see Table A.1E.2). ^{35}S -labeled nucleoside triphosphates have a thio moiety replacing one of the oxygens covalently bound to the phosphate. This is a significant perturbation that inhibits many enzymes. Because the lower energy of ^{35}S does not cause extensive damage to nucleic acids, it is used for the preparation of more stable, but lower specific activity, probes. Recently, ^{35}S has replaced ^{32}P as the preferred isotope for the dideoxy sequencing procedure because its lower energy results in sharper images in autoradiography. Exposure to ^{35}S is less dangerous to laboratory workers than exposure to ^{32}P , but ^{35}S is more difficult to detect with a standard radioactivity monitor.

^3H has the lowest specific activity (29 Ci/mmol in pure form) and the longest half-life (12 years). Its emission is too weak for most autoradiographic procedures, although it is used for in situ hybridizations. Its primary use is for quantitative analysis of nucleic acid synthesis and degradation. For some specific purposes, ^{14}C and ^{125}I are used for radiolabeling nucleic acids.

CAUTION: Investigators should wear gloves for all procedures involving radioactivity and should be careful about contaminating themselves or their clothing. All experiments involving ^{32}P should be performed behind lucite screens to minimize exposure. When working with ^{32}P , investigators should frequently check themselves and the working area for radioactivity with a hand-held minimonitor. When finished, radioactive waste should be placed in appropriately designated areas (not public garbage cans!) for eventual disposal. The working area should be checked for radioactivity and, if any is found, cleaned up to protect other laboratory workers and other unsuspecting people.

BASIC PROTOCOL

Measuring Radioactivity in DNA and RNA by Acid Precipitation

For many of the radioactive labeling procedures discussed in this chapter, it is useful to quantitate the amount of radioactive precursors that have been incorporated into nucleic acid. The amount of radioactivity for a known amount of DNA is called the specific activity (usually given in units of cpm/ μg of nucleic acid). The most common method for the measurement of radioactivity in nucleic acids is based on the fact that DNA or RNA molecules greater than 20 nucleotides in length are quantitatively precipitated in strong acids, whereas dNTP or NTP precursors remain in solution. The precipitated nucleic acids are easily separated from the radioactive precursors by filtration on discs of glass microfiber paper.

Materials

500 $\mu\text{g}/\text{ml}$ sonicated salmon sperm DNA in TE buffer
Ice-cold acid precipitation solution (APPENDIX 2)
100% ethanol

Reagents and
Radioisotopes Used
to Manipulate
Nucleic Acids

3.4.7

Glass microfiber filters (2.4-cm diameter, Whatman GF/A)
Filtration device
Heat lamp (optional)
Scintillation fluid and vials

1. Add a known volume (typically 1 μ l) of a reaction mixture containing radioactive precursors to a disposable glass tube containing 100 μ l of salmon sperm DNA (500 μ g/ml in TE buffer).
2. Spot 10 μ l of the mixture onto a glass microfiber filter.
3. To the remaining 90 μ l, add 1 ml of ice-cold acid precipitation solution and incubate on ice 5 to 10 min. The solution should appear cloudy due to the precipitated nucleic acid.
4. Collect the precipitate by filtering the solution through a second glass microfiber filter. Rinse the tube with 3 ml acid precipitation solution and pour through the filter. Wash the filter four more times with 3 ml acid precipitation solution, followed by 3 ml ethanol.
5. Dry both filters under a heat lamp, and place them in separate vials containing 3 ml of a toluene-based scintillation fluid. Measure the radioactivity in a liquid scintillation counter.

Drying the filters is unnecessary with scintillation fluids that can accommodate aqueous samples.

6. Determine the incorporation of radioactivity into nucleic acid from the ratio of cpm on the second filter (which measures radioactivity in nucleic acid) to cpm on the first filter (which measures total radioactivity in the sample).

For example, if both filters contain an equal number of cpm, the incorporation is 11% because only 10% of the original mixture (step 1) was placed on filter 1 whereas 90% of the original mixture was subjected to acid precipitation.

Separating Radioactively Labeled DNA from Unincorporated dNTP Precursors by Column Chromatography

After most radioactive labeling procedures, it is desirable to separate the DNA from unincorporated dNTPs. Removal of the radioactive precursors helps to reduce the background in hybridization experiments and protects investigators from unnecessary exposure to radioactivity. Gel filtration, which separates DNA and dNTPs by virtue of the large difference in molecular weight, is the most efficient method. It is also possible to perform the separation by ethanol precipitation, especially if ammonium acetate is used instead of sodium chloride (*UNIT 2.1*). However, ethanol precipitation is less efficient at removing the dNTPs and is more likely to result in radioactive contamination of the microcentrifuge and the investigator.

For DNA that is greater than 100 nucleotides, the two most common materials for column chromatography are Sephadex G-50 and Bio-Gel P-60. Other resins may be more useful for specialized applications (e.g. Bio-Gel P-2 is useful for separating oligonucleotides from dNTPs). Before carrying out the separation procedure described below, it is important to “swell the resin.” This is accomplished by adding 30 g of either resin to 300 ml TE buffer in a 500-ml screw-capped bottle and then heating to 65°C for a few hours (overnight at room temperature is also sufficient). After cooling to room temperature, excess TE buffer is removed and replaced by fresh TE buffer ($\frac{1}{2}$ volume of the resin). Store at room temperature or 4°C.

**BASIC
PROTOCOL**

**Enzymatic
Manipulation
of DNA and RNA**

3.4.8

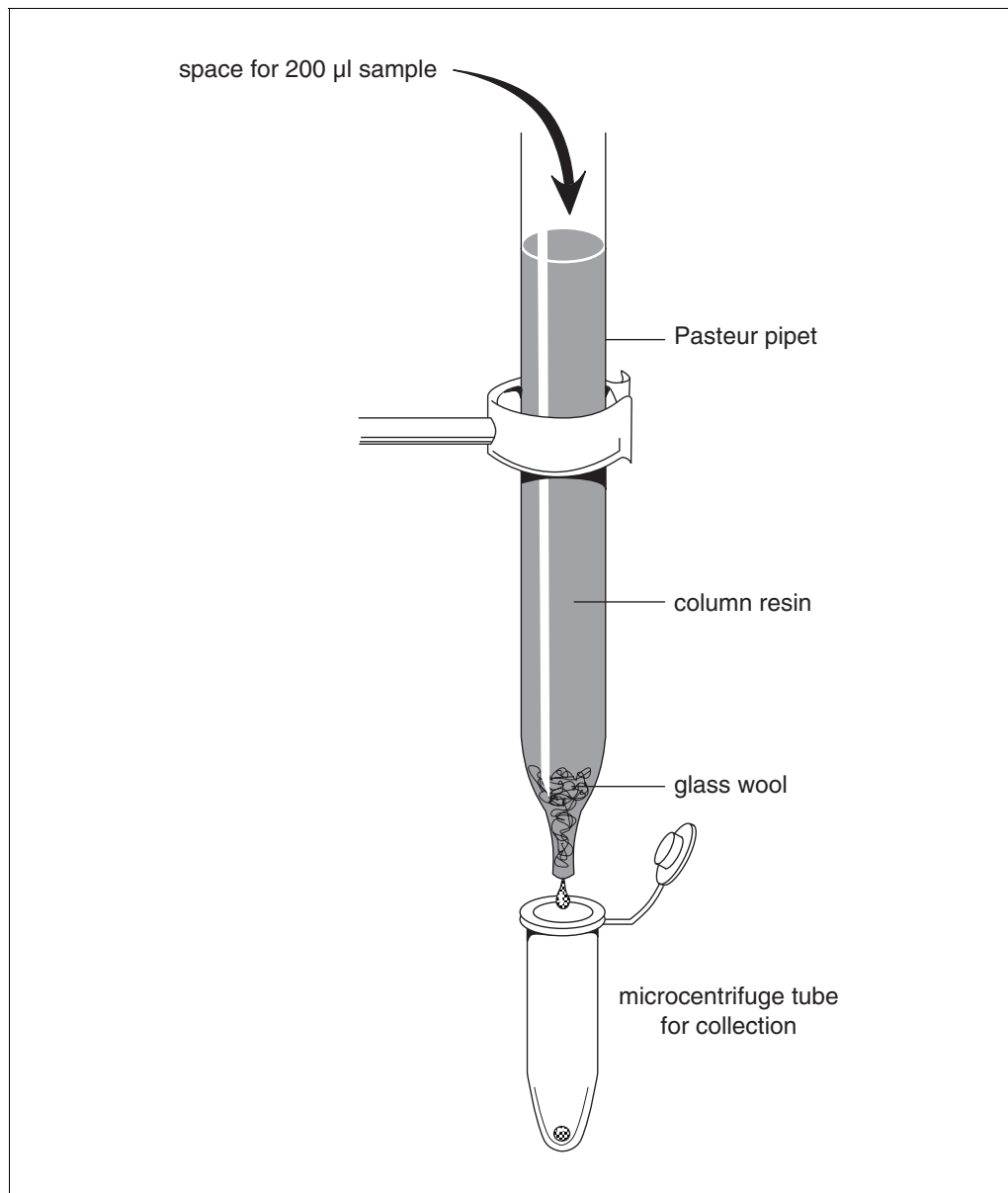


Figure 3.4.1 Separating DNA from dNTPs by column chromatography.

Materials

TE buffer (*APPENDIX 2*)

6 1/2-in. Pasteur pipet

Silanized glass wool (*APPENDIX 3*)

Swelled column resin (e.g., Sephadex G-50 or Bio-Gel P-60 from Bio-Rad)

1. Plug the bottom of a 6 1/2-in. Pasteur pipet with clean glass wool.

The easiest method is to push the glass wool down to the bottom with a 9-in. Pasteur pipet. A small amount of glass wool is sufficient to plug the bottom.

2. Swirl the bottle containing the column resin to make an even suspension, and fill the pipet with the suspension. Repeat until the pipet is almost completely filled with packed column resin, leaving enough space for the addition of 200 µl (see Fig. 3.4.1). This should take about 5 min.

Several columns can be prepared and processed simultaneously.

3. After placing the column behind a lucite shield to protect against exposure to radioactivity, apply the DNA sample (in a volume no greater than 200 μ l) to the column, while collecting the drops in a microcentrifuge tube. After the sample has run in, wash the column with two 200- μ l additions of TE buffer.

It takes 1 to 2 min for each 200- μ l addition to run into the column. At this time, the DNA should be nearing the bottom of the column, while the dNTPs should be near the top of the column. This can be ascertained with a hand-held "minimonitor."

4. Place a new microcentrifuge tube beneath the column, and add 200 μ l TE buffer to the top of the column. Repeat four times, thus generating five fractions each containing 200 μ l.
5. Determine the amount of radioactivity in each tube with a hand-held minimonitor. Almost all of the DNA should be in the first two fractions that contain high amounts of radioactivity. The amount of radioactivity in subsequent fractions, which represents a variety of products including degraded DNA (most often seen in nick translation reactions containing too much DNase I), should be less than in the "peak" fractions. The unincorporated dNTPs should still be on the column.

With experience, the investigator will know when the radioactive DNA will come through the column. By following the progress of the chromatography with a minimonitor, fewer fractions need be collected.

6. For many applications, the fractions containing the DNA can be used directly. If necessary, the DNA can be ethanol precipitated (UNIT 2.1). The column and the unused radioactive fractions should be disposed of properly as radioactive waste.

Spin-Column Procedure for Separating Radioactively Labeled DNA from Unincorporated dNTP Precursors

**ALTERNATE
PROTOCOL**

This method involves the same principle as described in the basic protocol above for conventional column chromatography. It differs in that the packing and running of the column is accomplished by centrifugation rather than by gravity. This method is more rapid than conventional chromatography and is useful when many samples are involved; however, removal of the dNTPs may be less quantitative. Spin columns can be purchased commercially at moderate expense.

CAUTION: Use extreme care that centrifuge and work area do not become contaminated with radioactivity. Whenever possible, place samples behind a lucite shield.

Additional Materials

5-ml disposable syringe

1. Plug the bottom of a 5-ml disposable syringe with clean, silanized glass wool.
2. Swirl the column resin to make an even suspension, and fill the syringe with the suspension.

Several columns can be prepared and processed simultaneously.

3. Place the syringe containing the resin into a polypropylene tube that is suitable for centrifugation in a desktop centrifuge (see Fig. 3.4.2). Spin 2 to 3 min at a setting of 4 in order to pack the column.

Centrifugation should be carried out in a swinging-bucket rotor, not fixed-angle rotors. This ensures that the sample passes uniformly through the column. In addition, do not use Sorvall or Beckman floor-model centrifuges that are meant for higher speed centrifugation.

**Enzymatic
Manipulation
of DNA and RNA**

3.4.10

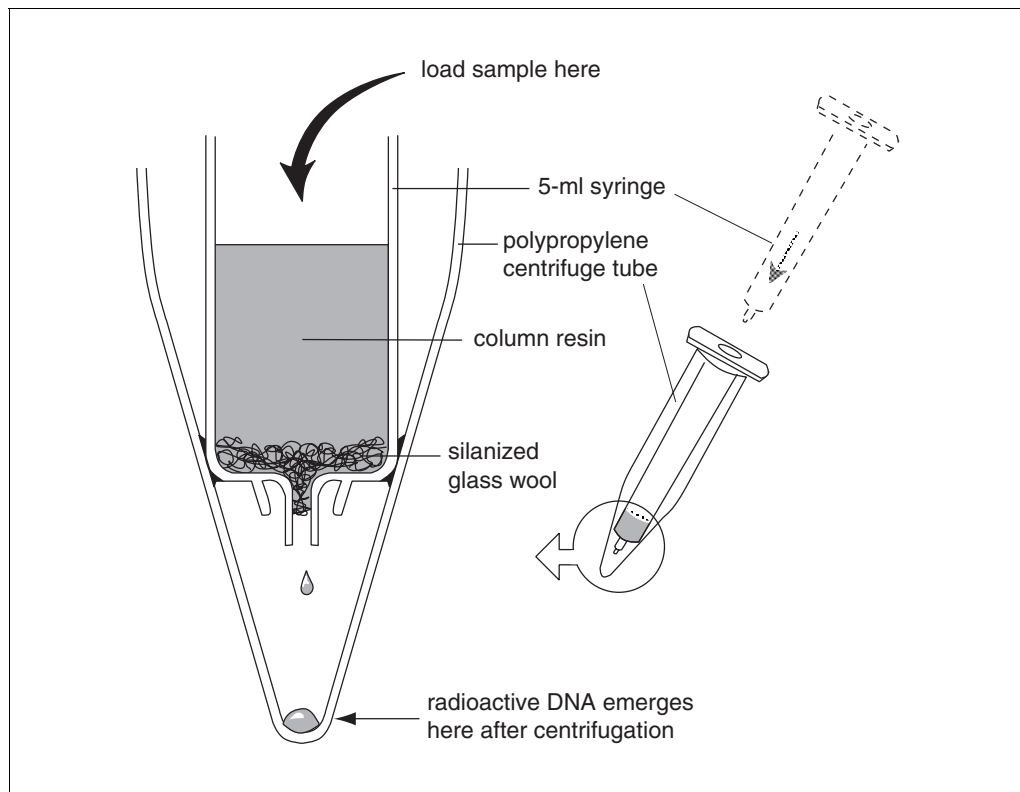


Figure 3.4.2 Preparation of a spin column. A 5-ml syringe containing the column resin is placed upside down in a polypropylene centrifuge tube.

The most convenient equipment for this purpose is a desktop centrifuge.

The resin should not be packed too tightly or too loosely. Therefore, adjust the time and speed in the tabletop centrifuge according to individual laboratory conditions. A general guideline for speed is ~1200 rpm.

4. Dilute the radioactive sample with TE buffer to a final volume of 100 μ l and load the sample in the center of the column.

If sample is loaded on the side of the column resin, some dNTPs may pass around the gel and elute with the DNA.

5. Place the syringe containing the loaded column into a new polypropylene tube, and spin 5 min at a setting of 5 to 6.
6. Save the liquid at the bottom of the tube containing the labeled DNA. Discard the syringe, which retains the unincorporated dNTPs, into a radioactive waste unit.

Contributed by Kevin Struhl
Harvard Medical School
Boston, Massachusetts

DNA-Dependent DNA Polymerases

All DNA polymerases add deoxyribonucleotides to the 3'-hydroxyl terminus of a primed double-stranded DNA molecule (Kornberg, 1980), as shown in Figure 3.5.1.

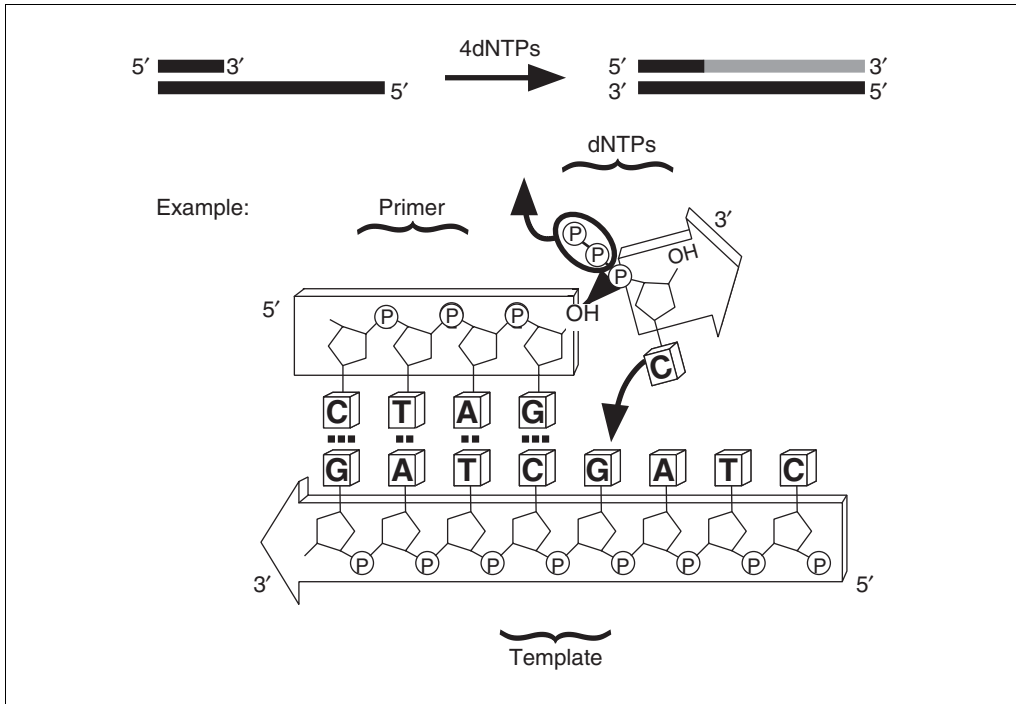


Figure 3.5.1 DNA polymerase 5'→3' polymerase activity.

Synthesis is exclusively in a 5'→3' direction with respect to the synthesized strand. Each nucleotide that is incorporated during polymerization is complementary to the one opposite to it in the template (dA pairs with dT, dC with dG). The reaction requires the four deoxyribonucleoside triphosphates (dNTPs) and magnesium ions.

Many DNA polymerases have a 3'→5' exonuclease inherently associated with the polymerase activity, as shown in Figure 3.5.2.

The 3'→5' exonuclease activity removes a single nucleotide at a time, releasing a nucleoside 5' monophosphate. In the absence of dNTPs, this activity will catalyze

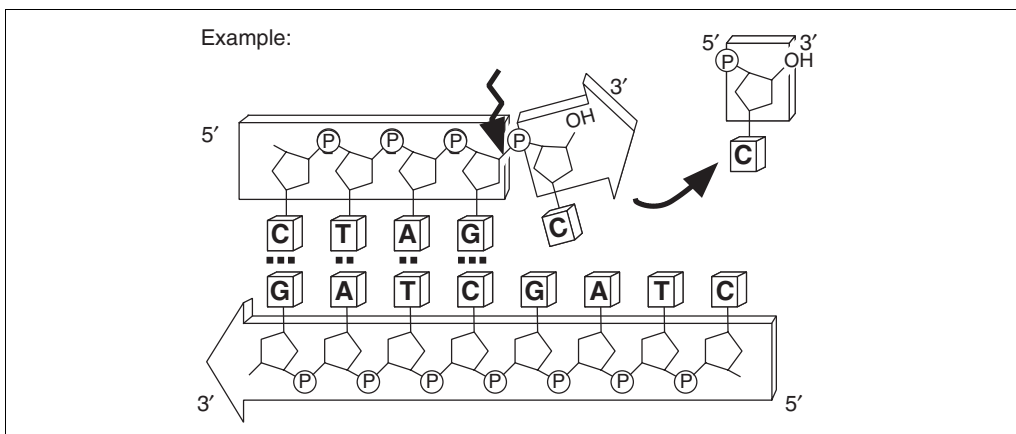


Figure 3.5.2 DNA polymerase 3'→5' exonuclease activity.

stepwise degradation from a free 3'-hydroxyl end of both single- and double-stranded DNA. In the presence of dNTPs, the exonuclease activity on double-stranded DNA is inhibited by the polymerase activity. During DNA synthesis, the exonuclease activity performs a proofreading function by removing misincorporated nucleotides.

In addition to the 3'→5' exonuclease activity, some DNA polymerases (e.g., *E. coli* DNA polymerase I) also have an associated 5'→3' exonuclease activity. This activity degrades double-stranded DNA from a free 5'-hydroxyl end (Fig. 3.5.3).

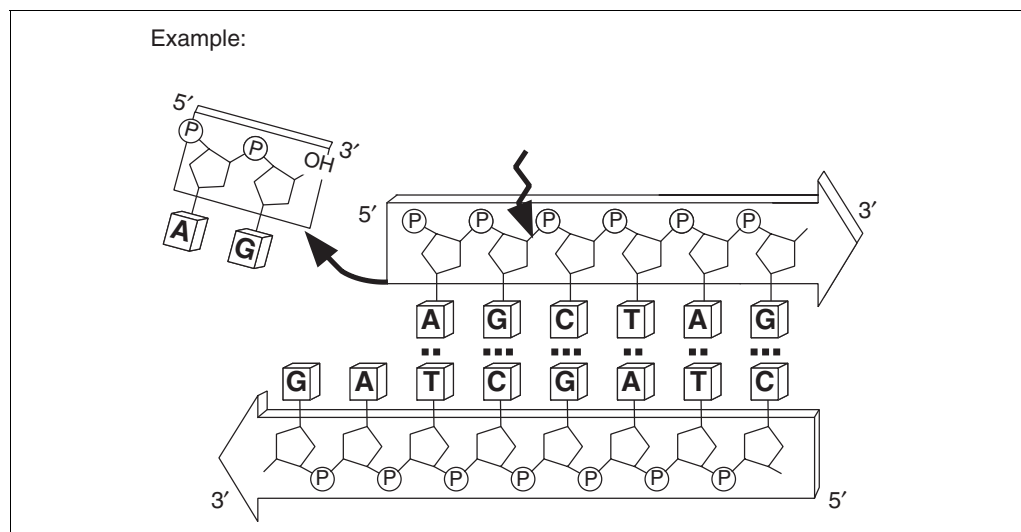


Figure 3.5.3 DNA polymerase 5'→3' exonuclease activity.

The 5'→3' exonuclease activity removes from one to several nucleotides at a time, releasing predominantly nucleoside 5' phosphates, but also some larger oligonucleotides up to 10 nucleotides in length. The 5'→3' exonuclease activity of *E. coli* DNA polymerase I enables it to initiate synthesis from nicks in duplex DNA. The 5'→3' exonuclease degrades the DNA ahead of the synthesizing polymerase, resulting in translocation of the nick (Kelley et al., 1970). This reaction, known as nick translation, is important for uniformly labeling duplex DNA molecules (Rigby et al., 1977; see Fig. 3.5.4). The 5'→3' exonuclease activity of *E. coli* DNA polymerase I, which is located at the N-terminus of the molecule, can be removed either by protease treatment or by deletion of the relevant part of the gene. The resulting DNA polymerase, which retains the 3' to 5' exonuclease, is referred to as the *E. coli* DNA polymerase I large fragment, or the Klenow fragment.

An important property of DNA polymerases is their processivity. Processivity is the ability of a polymerase molecule to incorporate nucleotides continuously on a given primer without dissociating from the primer template. Most DNA polymerases (e.g., *E. coli* DNA polymerase I, Klenow fragment, T4 DNA polymerase) have low processivity; they dissociate from a primer template after incorporating fewer than 10 nucleotides. In contrast, T7 DNA polymerase is highly processive; it can incorporate thousands of nucleotides from a given primer without dissociating from the primer template. This is a useful property when long stretches of DNA are being synthesized. See Table 3.5.1 for a summary of properties of DNA polymerases.

Table 3.5.1 Properties of DNA Polymerases

Enzyme	3'→5' exonuclease	5'→3' exonuclease	Polymerase rate	Processivity
<i>E. coli</i> DNA polymerase	Low	Present	Intermediate	Low
Klenow fragment	Low	None	Intermediate	Low
Reverse transcriptase	None	None	Slow	Intermediate
T4 DNA polymerase	High	None	Intermediate	Low
Native T7 DNA polymerase	High	None	Fast	High
Chemically modified T7 DNA polymerase	Low	None	Fast	High
Genetically modified T7 DNA polymerase ($\Delta 28$)	None	None	Fast	High
<i>Taq</i> DNA polymerase	None	Present	Fast	High

***ESCHERICHIA COLI* DNA POLYMERASE I** (Lehman, 1981)**ENZYME**

The enzyme is the product of the *E. coli polA* gene. The *polA* gene has been cloned in bacteriophage λ , and the enzyme is overproduced after temperature induction of a lysogenic strain. DNA polymerase I is a single polypeptide of molecular weight 109,000. In addition to its DNA-dependent DNA polymerase activity, it also has both a 3'→5' exonuclease activity and a 5'→3' exonuclease activity. The 3'→5' exonuclease activity is much less active than that of T4 or T7 DNA polymerase.

Reaction Conditions

For 25- μ l reaction:

50 mM Tris·Cl, pH 7.5

10 mM MgCl₂

1 mM DTT

50 μ g/ml BSA

20 μ M 4dNTP mix (one or more [α -³²P]dNTPs; UNIT 3.4)

1 μ g DNA

3 U *E. coli* DNA polymerase I

The volume of reaction, concentration of the 4 dNTPs, and amount of DNA will vary, depending upon the specific application. Stop the reaction by adding 1 μ l of 0.5 M EDTA or by heating to 75°C for 10 min.

Nick translation reactions should be carried out at 15°C (typically for 1 to 3 hr) to prevent the synthesis of “snapback” DNA that occurs at higher temperatures. Other reactions involving DNA polymerase I are performed at 20° to 37°C for 15 to 30 min.

For preparing radioactively labeled DNA of high specific activity, the concentration of one (or more) of the dNTP precursors is reduced to 1 to 2 μ M. For example, in a 25- μ l reaction, 25 μ Ci of ³²P-labeled dATP at 400 to 800 Ci/mmol, or 100 μ Ci at 3000 Ci/mmol, can be used instead of 20 μ M dATP. For efficient synthesis of DNA, the concentration of any of the dNTPs should not go below 1 μ M, the approximate K_m . If dNTP precursors of higher specific activity (3000 to 5000 Ci/mmol) are used, the reaction volume should be decreased or the amount of radioactive dNTP should be increased to maintain the 1 μ M concentration. Reactions with two radioactive precursors (both at 1 μ M) are frequently inefficient because the reaction rate is slowed by the low concentrations of the labeled precursors.

Uniform Labeling of DNA by Nick Translation

The nick translation reaction is a rapid, easy, and relatively inexpensive method for producing uniformly radioactive DNA of high specific activity. It is frequently used for the preparation of sequence-specific probes for screening libraries (UNIT 6.3), for genomic DNA blots (UNIT 2.9), and for RNA blots (UNIT 4.9). A typical nick translation reaction is carried out with a DNA fragment that has been purified by gel electrophoresis; for some applications, it is unnecessary to purify the fragment away from the plasmid or phage vector DNA sequences.

The mechanism of nick translation involves the combined activities of the 5'→3' polymerase and 5'→3' exonuclease activities of *E. coli* DNA polymerase I. Given a nicked duplex DNA molecule, the polymerase will translocate the nick, removing nucleotides ahead of it using its 5'→3' exonuclease activity, while simultaneously synthesizing DNA at the 3' end. The nicks in the DNA template are generally produced by adding trace amounts of DNase I to the reaction mixture (UNIT 3.12) (Fig. 3.5.4).

1. Prepare reaction mix for 0.25 µg of DNA as follows:

2.5 µl 0.5 mM 3dNTP mix (minus dATP; UNIT 3.4)
 2.5 µl 10× *E. coli* DNA polymerase I buffer (UNIT 3.4)
 10 µl 3000 Ci/mmol [α -³²P]dATP (100 µCi)
 1 µl DNase I [diluted 10,000-fold into standard enzyme diluent (APPENDIX 2)
 from a 1 mg/ml stock solution; refer to critical parameters]
 1 µl *E. coli* DNA polymerase I (5 to 15 U)

Keep ice-cold until all additions are made.

For generating probes with slightly lower specific activity, 2.5 µl (25 µCi) of 800 Ci/mmol [α -³²P]dATP can be used (an additional 7.5 µl water must be added to reaction mixture to maintain a constant volume). Lower-specific-activity probes are useful for most purposes, and are less expensive to prepare.

The reaction can be scaled up or down depending on the amount of probe that is desired. However, it is critical to maintain the above concentrations of DNA, [α -³²P]dATP, and enzymes for an efficient reaction.

The amount of DNase I is critical for optimal nick translation (refer to critical parameters). A molten gel slice of DNA in low gelling/melting temperature agarose can be added directly to the reaction mixture. For reproducible results, it is best to prepare and store the 1 mg/ml DNase I solution as described in UNIT 3.12.

2. Add 0.25 µg DNA in a total volume of 8 µl (add water if necessary). The total reaction volume should be 25 µl. Immediately incubate the reaction mixture at 12° to 14°C for 15 to 45 min, depending on the DNase calibration curve (refer to critical parameters).

If the DNA of interest is purified by electrophoresis in a low gelling/melting temperature agarose gel (UNIT 2.6), 8 µl of a molten gel slice (at 37°C) can be added directly to the reaction mixture without purification of the DNA fragment from the agarose. Although the resulting mixture will solidify, the nick translation reaction will take place (sometimes at a reduced efficiency).

3. After the time determined for the particular batch of DNase I, stop the reaction by adding 1 µl of 0.5 M EDTA, 3 µl of 10 mg/ml tRNA, and 100 µl TE buffer.

If DNA was added as a molten gel slice, the stopped reaction mixture must be remelted by incubating 10 min at 70°C.

4. Phenol extract the reaction mixture as described in UNIT 2.1. Transfer the aqueous (top) phase into a fresh tube.

If the volume of the mixture exceeds 400 μl , a small amount may seep around the cap and contaminate the microcentrifuge with ^{32}P .

5. Separate the labeled DNA from the unincorporated radioactive precursors by chromatography on a small Sephadex G-50 or Bio-Gel P-60 column as described in UNIT 3.4.

Alternatively, this separation can be achieved by ethanol precipitation; however, the separation is not as good and the procedure is more likely to lead to radioactive contamination of the investigator and equipment.

6. If desired, remove a 1- μl aliquot and determine the amount of ^{32}P incorporation (UNIT 3.4). A specific activity of 10^8 cpm/ μg should be easily obtained.

Commercial kits are available which can be used to produce nick translated DNA fragments. However, they do not produce very high-specific-activity probes because they do not optimize the DNase to DNA polymerase I activities as described below in critical parameters. Furthermore, they are expensive.

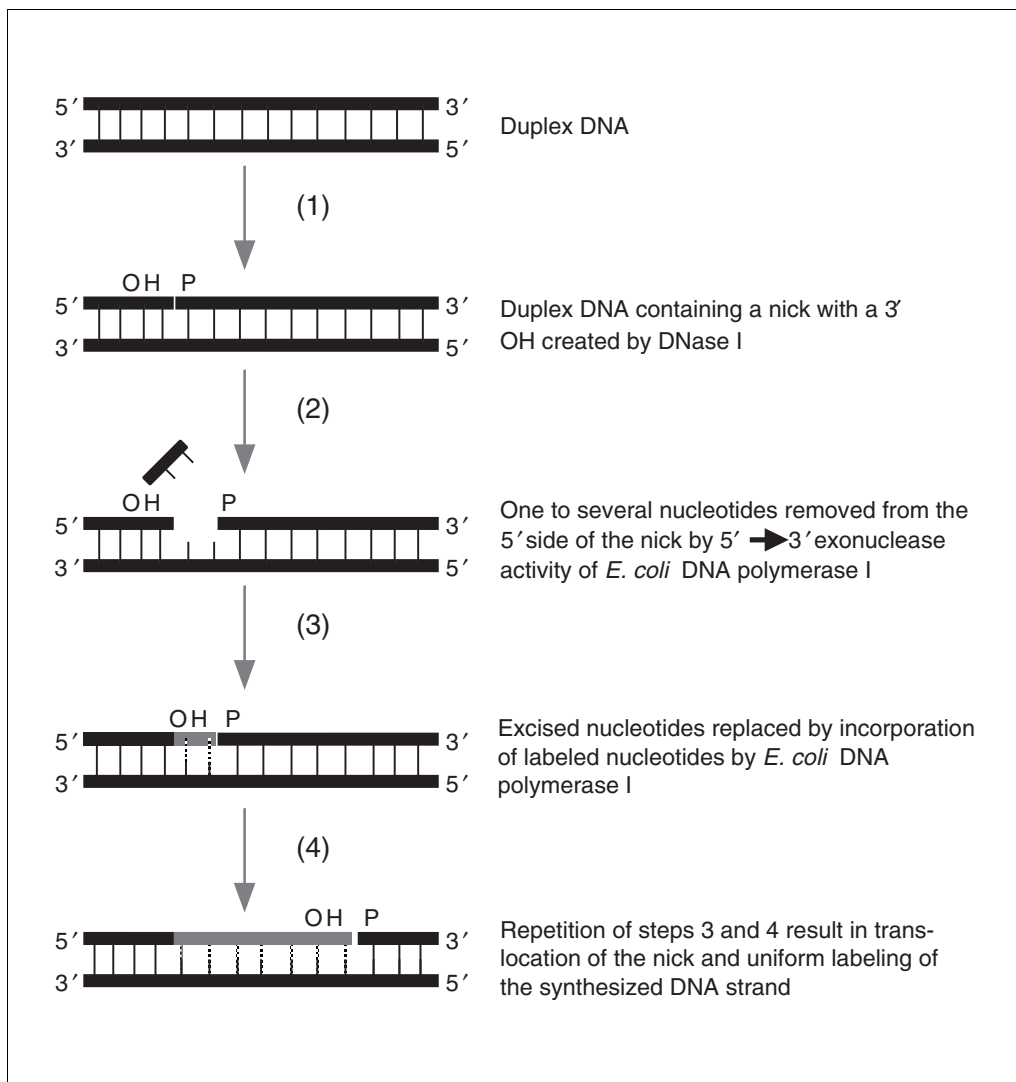


Figure 3.5.4 Nick translation activity of *Escherichia coli* DNA polymerase I.

Critical Parameters of Nick Translation

1. For production of high-specific-activity probes by nick translation, it is crucial to optimize the amount of DNase I and the time of reaction. If the reaction time is too short or the amount of DNase I is too low, insufficient label will be incorporated. In contrast, if the reaction time is too long or the amount of DNase I is too high, the DNA will be degraded. To optimize the nick translation reaction for a given preparation of the 1 mg/ml DNase I stock solution, parallel reactions should be set up in which the amount of DNase I and the time of reaction are varied. By measuring the incorporation of ^{32}P dATP into DNA by acid precipitation (*UNIT 3.4*), optimal conditions can be selected. At the appropriate concentration of DNase I (typically 1 μl of a 10^4 dilution of a 1 mg/ml solution), a maximum incorporation of 30 to 40% should occur between 15 and 45 min. Once the DNase I stock solution is “calibrated,” nick translations are carried out routinely by using the optimal DNase I dilution and reaction time.
2. Nick-translation can be carried out on DNA fragments purified by gel electrophoresis or on intact DNA molecules. Gel-purified DNA fragments are usually essential for screening genomic or cDNA libraries because vector or *E. coli* DNA sequences in the probe will hybridize to every clone in the library. The purification of DNA fragments also leads to the production of more specific probes, which is helpful for increasing the signal and reducing the background in hybridization experiments. For example, consider a 1-kb DNA fragment cloned into a 4-kb plasmid vector or a 39-kb bacteriophage λ vector. If intact DNA molecules are nick translated, only 20% or 2.5% of the labeled DNA will represent the DNA region of interest; in contrast, if the purified fragment is used, 100% of the labeled DNA will represent the region of interest.
3. For nick translation of very short fragments (<500 bp), it may be necessary to increase the amount of DNase I in order to ensure that most of the molecules are nicked at least once.
4. Nick translations of electrophoretically separated DNA fragments can be carried out without purification of the DNA away from the gel matrix provided that low gelling/melting temperature agarose is used. The gel slice containing the DNA fragment is melted, and the reaction occurs in the resolidified gel. This procedure saves time, but radioactive labeling of the DNA may not be as efficient.

Other Applications of *E. Coli* DNA Polymerase I

1. Labeling 3' ends of DNA molecules. This procedure is generally performed with the Klenow fragment of *E. coli* DNA polymerase I (see following section).
2. Repairing overhanging 3' or 5' ends to generate blunt ends. This procedure is generally performed with the Klenow fragment of *E. coli* DNA polymerase I (see following section). For repairing overhanging 3' ends, T4 DNA polymerase is usually preferred because its 3'→5' exonuclease is more active.

KLENOW FRAGMENT OF *ESCHERICHIA COLI* DNA POLYMERASE I

The Klenow fragment, molecular weight 76,000, consists of the C-terminal, 70% of *E. coli* DNA polymerase I. It retains the DNA polymerase and 3'→5' exonuclease activity of *E. coli* DNA polymerase I, but lacks the 5'→3' exonuclease activity.

Originally, the Klenow fragment was generated by proteolytic digestion of intact *E. coli* DNA polymerase I (Jacobsen et al., 1974). However, now it is overproduced directly from the cloned gene for DNA polymerase I modified to delete the N-terminal 323 amino acid residues of the polymerase (Joyce and Grindley, 1983).

The reaction conditions for the Klenow fragment are the same as those described earlier in this unit for the intact DNA polymerase I. The protocols that follow below describe procedures for labeling 3' ends of DNA, for repairing 3' or 5' overhanging ends to blunt ends, and for uniformly labeling DNA by random-sequence oligonucleotide-primed synthesis.

Labeling the 3' Ends of DNA

1. In a 20- μ l reaction mixture, digest 0.1 to 4 μ g DNA with a restriction endonuclease that generates 5' overhanging ends (UNIT 3.1).

DNA fragments with blunt ends can be labeled inefficiently by replacement of the nucleotide at the 3'-hydroxyl terminus. For endonucleases that produce 3' overhanging ends, labeling of 3' termini must be carried out by replacement synthesis using T4 DNA polymerase (see Fig. 3.5.5 and accompanying text).

2. Add 20 μ Ci of the desired [α -³²P]dNTP (400 to 800 Ci/mmol) and 1 μ l of appropriate 5 mM 3dNTP mix (UNIT 3.4). If higher specific activities are required, add 80 μ Ci of the radioactive dNTP at 3000 Ci/mmol.

Since the Klenow fragment incorporates nucleotides that are complementary to the single-stranded, 5' extensions, the choice of ³²P-labeled dNTP depends on the restriction endonuclease used to cleave the DNA. For example, labeling of BamHI ends (GATC) can be accomplished with any of the radioactive precursors, whereas labeling of EcoRI ends (AATT) requires either radioactive dATP or dTTP.

3. Add 1 U of the Klenow fragment and incubate 15 min at 30°C.

It is unnecessary to inactivate the restriction endonuclease, to change buffers, or to repurify the DNA prior to adding the Klenow fragment.

4. Stop the reaction with 1 μ l of 0.5 M EDTA or by heating to 75°C for 10 min. If desired, remove unincorporated dNTP precursors from labeled DNA (UNIT 3.4).

Critical Parameters of Labeling 3' Ends

For DNA fragments that are 500 bp in length, the above procedure will generate DNA with a specific activity of 10⁷ cpm/ μ g. For situations where higher specific activities are desired (e.g., for hybridization probes), use radioactive dNTPs of higher specific activity (3000 to 5000 Ci/mmol). When the specific activity is relatively unimportant (e.g., restriction mapping or preparation of size standards), add 1 μ l of a mixture containing all four unlabeled dNTPs (0.5 mM each dNTP) to the ³²P-labeled dNTP.

This procedure is extremely useful for visualizing small DNA molecules after gel electrophoresis. Unlike staining with ethidium bromide where band intensity is proportional to fragment length, all end-labeled fragments will have the same band intensity.

This is the method of choice for generating ³²P-labeled size standards for autoradiography. Besides the fact that band intensities are independent of fragment length, end-labeled

DNAs are stable for several months. Although radioactive decay reduces the specific activity and increases exposure time, DNA is not broken internally, as is the case for uniformly labeled molecules.

DNA can be labeled selectively at one end by cleaving with two different restriction endonucleases and labeling with a ^{32}P -labeled dNTP that is complementary to only one of the two 5' extensions. For example, if DNA is cleaved with *Eco*RI (AATT) and *Bam*HI (GATC), the *Bam*HI ends can be selectively labeled by using ^{32}P -labeled dGTP. Alternatively, DNA can be cleaved with one enzyme, radiolabeled, and then cleaved with the second enzyme. In this case, it is important to inactivate the Klenow fragment by heating to 75°C for 15 min prior to addition of the second restriction enzyme.

For some applications, e.g., DNA sequencing by the Maxam-Gilbert method (UNIT 7.5) or preparation of hybridization probes for S1 analysis (UNIT 4.6), it is important that all molecules are of identical length. To accomplish this, carry out steps 1 to 3 as above,

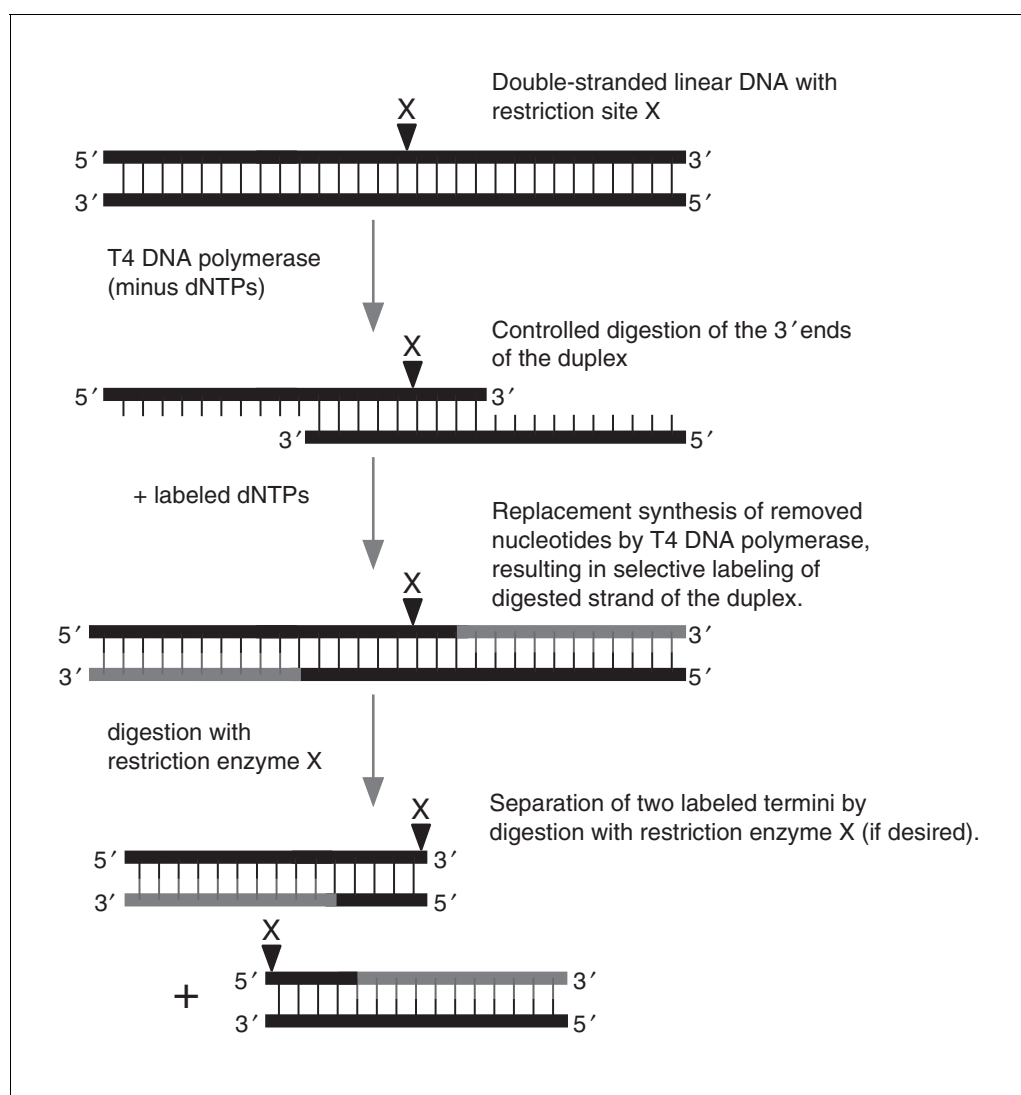


Figure 3.5.5 T4 DNA polymerase replacement synthesis.

but after the initial incubation, add 1 μ l of the 5 mM solution containing all four dNTP and incubate for an additional 15 min. This “cold chase” ensures that the Klenow fragment will extend to the end of every molecule.

Repairing 3' or 5' Overhanging Ends to Generate Blunt Ends

For many cloning experiments, it is necessary to convert the ends generated by restriction endonucleases into blunt ends (UNIT 3.16).

1. In a 20- μ l reaction, digest 0.1 to 4 μ g DNA with a restriction endonuclease.

DNA can be treated with exonucleases such as Bal 31, λ exonuclease, exonuclease III, or endonucleases such as S1 or mung bean nuclease if desired.

2. Add 1 μ l of 0.5 mM each dNTP.

It is unnecessary to inactivate the restriction endonuclease, to change buffers, or to repurify the DNA prior to adding the Klenow fragment.

3. Add 1 to 5 U of the Klenow fragment and incubate at 30°C for 15 min.

Repair of 5' extensions is carried out by polymerase activity, whereas repair of 3' extensions is carried out by 3' to 5' exonuclease activity. Due to the relative inactivity of exonuclease, this method is not desirable in cases where extensive repair of overhanging 3' ends is required. In such situations, T4 DNA polymerase (a much more expensive enzyme) or native T7 DNA polymerase are better choices.

4. Stop the reaction by heating to 75°C for 10 min or by adding 1 μ l of 0.5 M EDTA.

For restriction fragments produced by cleavage with different endonucleases, it is possible to repair one end selectively. This is done by cleaving with enzyme 1, repairing the ends, inactivating the Klenow fragment by heat (75°C for 10 min), and cleaving with enzyme 2.

Labeling of DNA by Random Oligonucleotide-Primed Synthesis

Random oligonucleotide-primed synthesis is an alternative to nick translation for producing uniformly radioactive DNA of high specific activity (Feinberg and Vogelstein, 1983). To carry out the labeling procedure, cleave the DNA with a restriction endonuclease, and, if desired, purify the DNA fragment containing the sequence of interest by gel electrophoresis. Denature the resulting linear DNA molecule(s) by boiling; anneal to random-sequence oligodeoxynucleotides (typically six bases in length); and then incubate with the Klenow fragment in the presence of dNTPs. In this way, the hexanucleotides prime the DNA of interest at various positions along the template, and are extended to generate double-stranded DNA that is uniformly labeled on both strands.

1. Cleave the DNA of interest with an appropriate restriction endonuclease (UNIT 3.1).

It is very useful, though not absolutely essential, to linearize the DNA for the subsequent denaturation step.

2. If desired, purify the DNA fragment of interest by gel electrophoresis using any of the procedures described in UNIT 2.6. If gel purification is unnecessary, ethanol precipitate the DNA (UNIT 2.1) and resuspend in TE buffer.

It is essential to remove the restriction endonuclease buffer from the DNA because the Mg^{++} ions will make it difficult to denature the DNA.

If low gelling/melting temperature agarose is used for the gel, it is usually unnecessary to purify the DNA away from the gel matrix.

3. Prepare the following reaction mix on ice:

- 2.5 μ l 0.5 mM 3dNTP mix (minus dATP; *UNIT 3.4*)
- 2.5 μ l 10 \times Klenow fragment buffer (*UNIT 3.4*)
- 5 μ l 3000 Ci/mmol [α -³²P]dATP (50 μ Ci)
- 1 μ l Klenow fragment (3 to 8 U)

For generating probes with slightly lower specific activity, 2.5 μ l (25 μ Ci) of 800 Ci/mmol [α -³²P]dATP can be used (an additional 2.5 μ l of water must be added to reaction mixture to maintain a constant volume). Lower-specific-activity probes are useful for most purposes, and are less expensive to prepare.

4. Combine the DNA of interest (30 to 100 ng) with the random hexanucleotides (1 to 5 μ g) in a total volume of 14 μ l. Boil the DNA mixture for 2 to 3 min and then place on ice.

If the DNA has not been purified from the gel matrix, place the boiled DNA at 37°C. Random hexanucleotides can be purchased from Pharmacia.

5. Add 11 μ l of the reaction mix from step 3 to the denatured DNA (final volume is 25 μ l), and immediately incubate the mixture 2 to 4 hr at room temperature.

6. Stop the reaction by adding 1 μ l of 0.5 M EDTA, 3 μ l of 10 mg/ml tRNA, and 100 μ l TE buffer.

If DNA was added as a molten gel slice, the stopped reaction mixture must be remelted by incubation at 70°C for 10 min.

7. Phenol extract the reaction mixture as described in *UNIT 2.1*. Transfer the aqueous (top) phase into a fresh tube.

8. Separate the labeled DNA from the unincorporated radioactive precursors by chromatography on a small Sephadex G-50 or Bio-Gel P-60 column as described in *UNIT 3.4*.

Alternatively, this separation can be achieved by ethanol precipitation; however, the separation is not as good and the procedure is more likely to lead to radioactive contamination of the investigator and equipment.

9. If desired, remove a 1- μ l aliquot and determine the amount of ³²P incorporation by acid precipitation (*UNIT 3.4*). A specific activity of 10⁸ cpm/ μ g should be easily obtained.

Other Applications of Klenow Fragment

1. DNA sequencing by the dideoxy method (*UNIT 7.4*).
2. Synthesis of the second strand for the cloning of cDNAs (*UNIT 5.5*).
3. Extension of oligonucleotide primers on single-stranded templates. Some specific applications include dideoxy DNA sequencing (mentioned above), synthesis of hybridization probes, and in vitro mutagenesis (*UNITS 8.1 & 8.3*).
4. Converting single-stranded oligonucleotides to double-stranded DNA by mutually primed synthesis (*UNIT 8.2*).

T4 DNA POLYMERASE

T4 DNA polymerase, the product of gene 43 of bacteriophage T4, is either prepared from cells of *E. coli* that have been infected with the phage or has been overproduced from its cloned gene. T4 DNA polymerase is a single polypeptide of molecular weight 112,000. In addition to its DNA-dependent DNA polymerase activity, it has a very active single-stranded and double-stranded 3'→5' exonuclease. It lacks a 5'→3' exonuclease activity. T4 DNA polymerase by itself has low processivity; however, in the presence of several T4 accessory proteins, it becomes very processive (Nossal, 1984). This polymerase has been produced from a clone that overexpresses its gene (gene 43) (Lin et al., 1987).

Reaction Conditions

For 50- μ l reaction:

50 mM Tris·Cl, pH 8.0

5 mM MgCl₂

5 mM DTT

2 μ g DNA

100 μ M 4dNTP mix (*UNIT 3.4*)

50 μ g/ml BSA

10 U T4 DNA polymerase

Incubate 20 min at 11°C. Stop reaction by adding 2 μ l of 0.5 M EDTA or by heating to 75°C for 10 min. The volume of reaction, concentration of 4 dNTPs, and the temperature of the reaction will vary, depending upon the individual application.

Effect of Triphosphate Concentration

For reaction conditions that do not require radioactive labeling of the DNA, high concentrations (100 μ M) of dNTPs are used to maximize the ratio of polymerase to exonuclease activity. In labeling experiments, the concentration of the labeled dNTP is reduced to 1 to 2 μ M; this maximizes the specific activity of the DNA. Levels lower than 1 μ M labeled dNTP should not be used because once the dNTPs are exhausted the exonuclease activity will degrade the DNA.

Effect of Temperature

For labeling 3' termini with T4 DNA polymerase, the temperature of the reaction should be maintained at 11°C. At higher temperatures, if only a subset of the four deoxyribonucleoside triphosphates are added, the exonuclease will degrade the template beyond the nucleotide which the polymerase could replace.

Buffer Compatibility

For most applications, T4 DNA polymerase is used following digestion with restriction endonucleases. For many restriction enzymes, cleavage can be carried out in T4 DNA polymerase buffer, and T4 DNA polymerase can be used directly. If buffers for the restriction enzyme and T4 DNA polymerase are incompatible, the DNA should be digested with the restriction enzyme and then repurified by phenol extraction, ethanol precipitation, and resuspension in TE buffer (*UNIT 2.1*) prior to treatment with T4 DNA polymerase.

Applications

1. Radioactive labeling of the 3' termini of DNA fragments (Goulian et al., 1968). DNA fragments containing 5' protruding ends are incubated with the appropriate [α -³²P]dNTPs at 1 to 2 μ M at 11°C for 20 min. The Klenow fragment of *E. coli* DNA polymerase I carries out the same reaction and is considerably less expensive.

2. Selective and extensive labeling of the 3' termini of a linear duplex DNA molecule (Challberg and Englund, 1980). This technique, known as "replacement synthesis" (Fig. 3.5.5), involves first incubating the duplex DNA fragment with T4 DNA polymerase in the absence of dNTPs. Under these conditions, the strong 3'→5' exonuclease activity will degrade selectively from the 3' ends. After an empirically determined period of time, the four dNTPs are added (including one radioactively labeled [α -³²P]dNTP). The polymerase activity of T4 DNA polymerase then extends the 3' ends the length of the template. Optimal labeling is achieved when the 3'→5' exonuclease activity removes 30 to 40% of the nucleotides from each end. If exonuclease action proceeds further, the strands will separate and hence not serve as substrates for the polymerase activity.

The resulting molecules are selectively labeled at their 3' ends to an extent determined by the length of time of the exonuclease reaction. To obtain a DNA fragment labeled from only one of its 3' ends, the molecules are subsequently cut with an appropriate restriction enzyme, and the resulting fragments are resolved by electrophoresis.

3. Converting the ends of any duplex DNA fragment to blunt-ended structures suitable for blunt-end ligation for cloning. Although the Klenow fragment of *E. coli* DNA polymerase I can also catalyze this reaction, T4 DNA polymerase is the enzyme of choice for removing 3' protruding ends because of its strong 3'→5' exonuclease activity. In the presence of high concentrations of all four dNTPs, the degradation will stop when the enzyme reaches the duplex region. Similarly, if the termini have a 5' protruding region, the enzyme will simply extend the recessed 3' termini until the end is blunt. For these applications, the DNA is incubated with T4 polymerase and 100 μ M each of the four dNTPs at 11°C for 20 min.

ENZYME

NATIVE T7 DNA POLYMERASE (Tabor et al., 1987)

This enzyme was originally isolated from bacteriophage T7-infected *E. coli*. The purified T7 DNA polymerase is a complex of two proteins in a one-to-one stoichiometry: the T7-encoded gene 5 protein (molecular weight 80,000), and the *E. coli*-encoded thioredoxin (molecular weight 12,000). Both proteins have been overproduced in *E. coli* from genes cloned into plasmid vectors.

Purified T7 gene 5 protein has a 3'→5' exonuclease activity and a nonprocessive DNA polymerase activity. Thioredoxin acts as an accessory protein to increase the affinity of T7 gene 5 protein for the primer template, rendering DNA synthesis processive for thousands of nucleotides. The high processivity of the T7 gene 5/thioredoxin complex (T7 DNA polymerase) is in marked contrast to the low processivity of *E. coli* DNA polymerase I, Klenow fragment, or T4 DNA polymerase.

Native T7 DNA polymerase (T7 gene 5 protein/thioredoxin complex) has a very active single-stranded and double-stranded DNA 3'→5' exonuclease activity in addition to its polymerase activity. This activity is a detriment to the use of this enzyme for DNA sequence analysis, and must be reduced or eliminated (modified T7 DNA polymerase) to be used for this application.

Reaction Conditions

For 50- μ l reaction:

40 mM Tris·Cl, pH 7.5
10 mM MgCl₂
5 mM DTT
50 mM NaCl
2 μ g DNA
300 μ M 4dNTP mix (*UNIT 3.4*)
50 μ g/ml BSA
5 U T7 DNA polymerase

Incubate at 37°C for 20 min. Stop reaction by adding 2 μ l of 0.5 M EDTA or by heating to 75°C for 10 min. The volume of reaction, concentration of 4 dNTPs, and temperature of the reaction will vary, depending upon the individual application.

Applications

1. Since T7 DNA polymerase is highly processive, it can be used for extensive synthesis of DNA on long templates (e.g., M13). It will extend thousands of nucleotides from the same primer template without dissociating and is largely unaffected by secondary structures that impede *E. coli* DNA polymerase I, T4 DNA polymerase, or reverse transcriptase.
2. Native T7 DNA polymerase is the enzyme of choice for the synthesis of the complementary strand during site-directed mutagenesis (Tabor and Richardson, 1989; Bebenek and Kunkel, 1989; see *UNIT 8.1*).
3. Native T7 DNA polymerase can be used analogously to T4 DNA polymerase for labeling 3' termini either by simple extension or by replacement synthesis. T7 and T4 DNA polymerases have comparably strong 3'→5' exonuclease activities.
4. Native T7 DNA polymerase can be used analogously to T4 DNA polymerase to convert the ends of any duplex DNA fragment (either 5' or 3' protruding) to blunt-ended structures.

MODIFIED T7 DNA POLYMERASE (Tabor and Richardson, 1987a, 1989)

Native T7 DNA polymerase has a very active 3'→5' exonuclease activity in addition to its polymerase activity. In modified T7 DNA polymerase, this exonuclease activity has either been reduced selectively by a chemical reaction or inactivated completely by genetic modification that deletes 28 amino acid residues from the exonuclease domain of the enzyme. Inactivation of the 3'→5' exonuclease activity increases the T7 DNA polymerase activity by 3- to 9-fold, by increasing the processivity of the enzyme and its ability to synthesize through regions of secondary structure.

Reaction Conditions

For 50- μ l reaction:

40 mM Tris·Cl, pH 7.5
5 or 10 mM MgCl₂ (see below)
5 mM DTT
50 mM NaCl
2 μ g DNA
300 μ M 4dNTP mix (*UNIT 3.4*)
50 μ g/ml BSA
10 U modified T7 DNA polymerase

ENZYME

**Enzymatic
Manipulation
of DNA and RNA**

3.5.13

Incubate at 37°C for 20 min. Stop reaction by adding 2 µl of 0.5 M EDTA or by heating to 75°C for 10 min. The volume of reaction, concentration of 4 dNTPs, amount of DNA and enzyme, and temperature of the reaction will vary, depending upon the individual application. Genetically modified T7 DNA polymerase has a lower MgCl₂ optimum (5 mM) than the chemically modified polymerase (10 mM).

Applications

1. Modified T7 DNA polymerase has the ideal characteristics for a DNA sequencing enzyme: high processivity, lack of 3'→5' exonuclease, and the lack of discrimination against deoxynucleotide analogs (dideoxynucleotides, [α -³⁵S]deoxynucleotides, and deoxyinosine nucleotides; see *UNIT 7.4* and Tabor and Richardson, 1987b).
2. Modified T7 DNA polymerase efficiently incorporates dNTPs that are present at very low levels (<0.1 µM) (Tabor and Richardson, 1987b). This is useful for the preparation of labeled substrates.
3. Because of its high specific activity and lack of associated 3'→5' exonuclease activity, modified T7 DNA polymerase is useful for labeling the 3' termini of DNA fragments with 5' protruding ends. It should not be used to make blunt-end fragments because it leaves a one-base overhang at the 5' end.

ENZYME

TAQ DNA POLYMERASE

Native *Taq* DNA polymerase is isolated from the extreme thermophile *Thermus aquaticus* (Lawyer et al., 1989). The enzyme is a double-stranded DNA polymerase with a molecular mass of 94 kDa. It has a temperature optimum for polymerization of 75° to 80°C and does not appear to have any 3'→5' exonuclease activity (D.H. Gelfand and S. Stoffel, unpublished results; Tindall and Kunkel, 1988). The enzyme is a single polypeptide chain with high turnover number and high processivity.

For Conditions for Polymerase Chain Reaction (PCR) see *UNIT 15.1*

Applications

1. *Taq* polymerase is active over a broad temperature range and has a high optimum temperature of polymerization, which makes it ideal for use in the polymerase chain reaction (see Chapter 15). The high temperature allows more specific pairing of oligonucleotide primers, which increases the efficiency and specificity and hence the sensitivity of PCR. The fact that the enzyme withstands repeated cyclings from the annealing temperatures used in PCR to the temperature required to separate the strands affords the advantage of adding the enzyme only once to a sample prior to running the amplification cycles.
2. *Taq* polymerase has a high turnover number, high processivity, and the capability to use deoxynucleoside triphosphate analogs, which makes it an excellent enzyme for DNA sequencing (Innis et al., 1988). These qualities—combined with its high temperature optimum—allow sequencing on templates that may exhibit strong secondary structure at lower temperatures, obviating use of conventional sequencing enzymes (see *UNIT 15.2*).

LITERATURE CITED

- Bebenek, K. and Kunkel, T.A. 1989. The use of native T7 DNA polymerase for site-directed mutagenesis. *Nucl. Acids Res.* 17:5408.
- Challberg, M.D. and Englund, P.T. 1980. Specific labeling of 3' termini with T4 DNA polymerase. *Methods Enzymol.* 65:39-43.
- Feinberg, A.P. and Vogelstein, B. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6-13.
- Goulian, M., Lucas, Z.J., and Kornberg, A. 1968. Enzymatic synthesis of deoxyribonucleic acid. XXV. Purification and properties of deoxyribonucleic acid polymerase induced by infection with phage T4. *J. Biol. Chem.* 243:627-638.
- Innis, M.A., Myambo, K.B., Gelfand, D.H., and Brow, M.D. 1988. DNA sequencing with *Thermus aquaticus* DNA polymerase, and direct sequencing of PCR-amplified DNA. *Proc. Natl. Acad. Sci. U.S.A.* 85:9436-9440.
- Jacobsen, H., Klenow, H., and Overgaard-Hansen, K. 1974. The N-terminal amino-acid sequences of DNA polymerase I from *Escherichia coli* and of the large and the small fragments obtained by a limited proteolysis. *Eur. J. Biochem.* 45:623-627.
- Joyce, C.M. and Grindley, N.D.F. 1983. Construction of a plasmid that overproduces the large proteolytic fragment (Klenow fragment) of DNA polymerase I of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 80:1830-1834.
- Kornberg, A. 1980. DNA Replication. W.H. Freeman, New York.
- Lawyer, F.C., Stoffel, S., Saiki, R.K., Myambo, K., Drummond, R., and Gelfand, D.H. 1989. Isolation, characterization, and expression of *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *J. Biol. Chem.* 264:6427-6437.
- Lehman, I.R. 1981. DNA polymerase I of *Escherichia coli*. In *The Enzymes*, Vol. 14A (P.D. Boyer, ed.) pp. 16-38. Academic Press, San Diego.
- Lin, T.C., Rush, J., Spicer, E.K., and Konigsberg, W.H. 1987. Cloning and expression of T4 DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 84:7000-7004.
- Nossal, N.G. 1984. Prokaryotic DNA replication systems. *Annu. Rev. Biochem.* 53:581-615.
- Tabor, S. and Richardson, C.C. 1987a. Selective oxidation of the exonuclease domain of bacteriophage T7 DNA polymerase. *J. Biol. Chem.* 262:15330-15333.
- Tabor, S. and Richardson, C.C. 1987b. DNA sequencing analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 84:4767-4771.
- Tabor, S. and Richardson, C.C. 1989. Selective inactivation of the exonuclease activity of bacteriophage T7 DNA polymerase by in vitro mutagenesis. *J. Biol. Chem.* 264:6447-6458.
- Tabor, S., Huber, H.E., and Richardson, C.C. 1987. *Escherichia coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J. Biol. Chem.* 262:16212-16223.
- Tindall, K.R. and Kunkel, T.A. 1988. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* 27:6008-6013.

Contributed by Stanley Tabor and Kevin Struhl
Harvard Medical School
Boston, Massachusetts

Stephen J. Scharf (*Taq* polymerase)
Cetus Corporation
Emeryville, California

David H. Gelfand (*Taq* polymerase)
Roche Molecular Systems
Alameda, California

Template-Independent DNA Polymerases

UNIT 3.6

TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE (Terminal Transferase)

ENZYME

This enzyme, purified from calf thymus, catalyzes the incorporation of deoxynucleotides to the 3'-hydroxyl termini of DNA accompanied by the release of inorganic phosphate (Fig. 3.6.1).

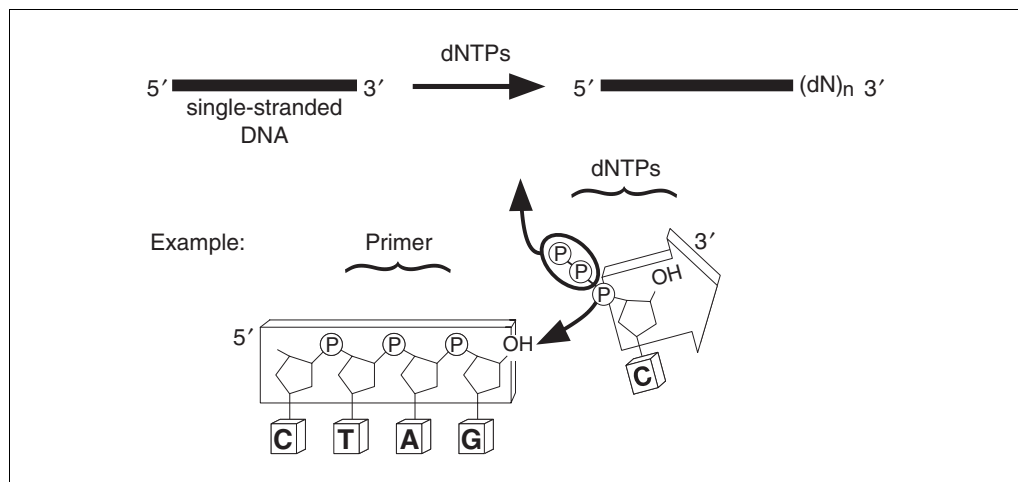


Figure 3.6.1 Terminal transferase activity.

A template is not required and will not be copied. Divalent cations are required, and the preference for the nucleotide to be incorporated will vary depending upon the divalent cation used. Single-stranded DNA is preferred as a primer; for double-stranded DNA, the extensions are most efficient when the ends have 3' protruding termini. However, in the presence of Co^{2+} , terminal transferase will prime any 3' terminus (protruding, blunt, or recessed) although not with uniform efficiency. The presence of Co^{2+} will also allow terminal transferase to catalyze the limited polymerization of ribonucleotides (Deng and Wu, 1983).

Reaction Conditions

For 50- μl reaction:

100 mM sodium cacodylate, pH 7.0

1 mM CoCl_2

0.1 mM DTT

4 pmol DNA (as 3' termini)

20 μM dNTPs (UNIT 3.4)

50 $\mu\text{g/ml}$ BSA

10 U terminal transferase

Incubate at 37°C for 30 min. Stop reaction by adding 2 μl of 0.5 M EDTA or by heating to 75°C for 10 min. The selection of which dNTPs to use and its concentration will vary, depending upon the individual application. Under the conditions outlined above, using dCTP, terminal transferase will add approximately 10 deoxycytidines to each 3' end in 30 min.

Enzymatic
Manipulation
of DNA and RNA

3.6.1

Contributed by Stanley Tabor

Current Protocols in Molecular Biology (1987) 3.6.1-3.6.2

Copyright © 2000 by John Wiley & Sons, Inc.

Applications

1. Cloning DNA fragments. It is used to synthesize a homopolymer “tail” at each end of the DNA to be cloned and a complementary homopolymer tail at each end of a linearized vector. The vector and insert DNAs are subsequently annealed by virtue of their complementary tails.
2. Labeling the 3' termini of DNA with ^{32}P . For DNA sequence analyses, incorporation can be limited to a single nucleotide by using [α - ^{32}P]cordycepin triphosphate (3'-deoxyribonucleoside triphosphate). This analog is a chain terminator; without a free 3' hydroxyl group, no additional nucleotides can be incorporated (Tu and Cohen, 1980).
3. Incorporating nonradioactive tags onto the 3' termini of DNA fragments. With use of biotin-11-dUTP [8-(2,4-dinitrophenyl-2,6-aminoethyl)aminoadenosine-5'-triphosphate or 2'-deoxyuridine-5'-triphosphate-5'-allylaminebiotin], these nucleotides serve as receptor sites for either fluorescent dyes or avidin conjugates (Vincent et al., 1982).
4. Synthesizing model polydeoxynucleotide homopolymers.

LITERATURE CITED

- Deng, G.R. and Wu, R. 1983. Terminal transferase: Use in the tailing of DNA and for in vitro mutagenesis. *Meth. Enzymol.* 100:96-116.
- Tu, C.-P.D. and Cohen. S.N. 1980. 3'-end labeling of DNA with [α - ^{32}P]cordycepin-5'-triphosphate. *Gene* 10:177-183.
- Vincent, C., Tchen, P., Cohen-Solal, M., and Kourilsky, P. 1982. Synthesis of 8-(2,4-dinitrophenyl-2,6-aminoethyl)aminoadenosine-5'-triphosphate: Biological properties and potential uses. *Nucl. Acids Res.* 10:6787-6796.

KEY REFERENCE

- Ratliff, R.L. 1981. Terminal deoxynucleotidyltransferase. *In* The Enzymes, Vol. 14A (P.D. Boyer, ed.) pp. 105-118. Academic Press, NY.

Contributed by Stanley Tabor
Harvard Medical School
Boston, Massachusetts

RNA-Dependent DNA Polymerases

UNIT 3.7

ENZYME

REVERSE TRANSCRIPTASE

Reverse transcriptases are derived from retroviruses such as avian myeloblastosis virus (AMV) or Moloney murine leukemia virus (MMLV), which use them to make DNA copies of their RNA genomes. The AMV enzyme is purified from isolated AMV virus (Verma, 1977), whereas the MMLV enzyme is purified from overproducing *E. coli* cells containing the cloned gene (Roth et al., 1985). AMV and MMLV reverse transcriptases are multifunctional enzymes, but they are mainly used as RNA-directed DNA polymerases. Specifically, deoxyoligonucleotides (either oligo(dT) polymers, a collection of random DNA sequences, or a specific sequence) are used as primers for extension on RNA (usually messenger RNA) templates (UNIT 5.5). The DNA synthesized from the RNA template is referred to as complementary DNA (cDNA).

Reverse transcriptase also has a DNA-directed DNA polymerase activity (Fig. 3.7.1). The incorporation of dNTPs is very slow (~5 nucleotides/sec), nearly 100 times slower than the rate of T7 DNA polymerase (Tabor and Richardson, 1987). The DNA polymerase activity of reverse transcriptase lacks a 3'→5' exonuclease activity.

A third activity of reverse transcriptase will degrade the RNA in an RNA:DNA hybrid (Fig. 3.7.2). This so-called RNase H activity is an exonuclease acting processively from either the 5' or 3' terminus. This activity is useful for selectively destroying parts of an RNA molecule to which DNA molecules (typically sequence-specific oligonucleotides) have been hybridized.

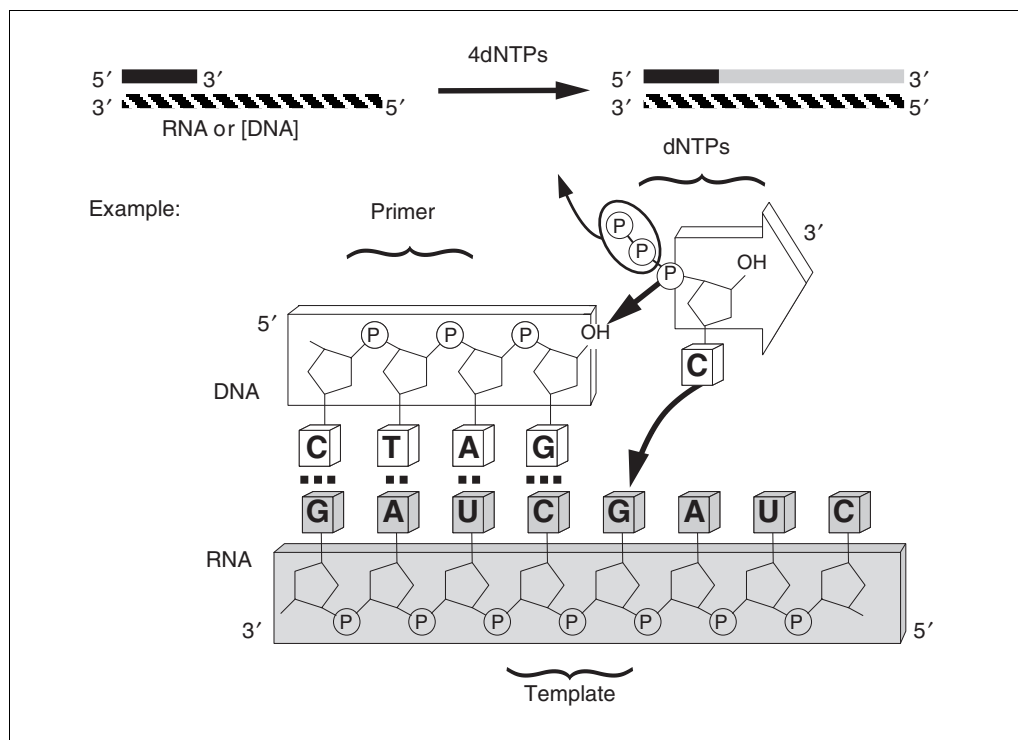


Figure 3.7.1 Reverse transcriptase 5'→3' DNA polymerase activity.

Enzymatic
Manipulation
of DNA and RNA

3.7.1

Contributed by Stanley Tabor

Current Protocols in Molecular Biology (1987) 3.7.1-3.7.3

Copyright © 2000 by John Wiley & Sons, Inc.

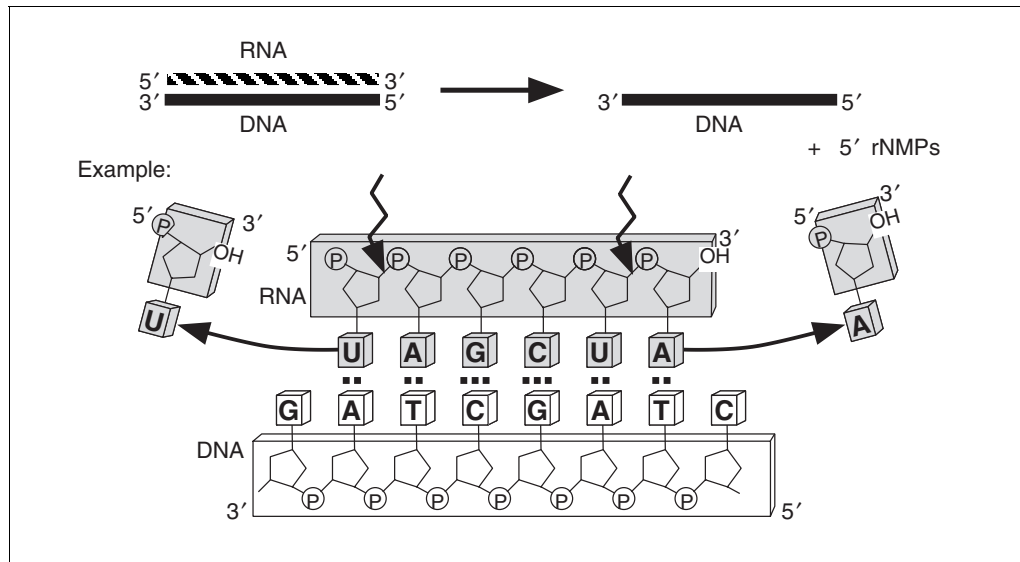


Figure 3.7.2 Reverse transcriptase 5'→3' and 3'→5' exonuclease activity (RNase H activity).

Reaction Conditions

For 50- μ l reaction:

50 mM Tris-Cl, pH 8.2

5 mM MgCl₂

5 mM DTT

50 mM KCl

1 μ g mRNA

1 μ g oligo(dT)₁₂₋₁₈

40 μ M 4 dNTPs (UNIT 3.4)

100 μ Ci [α -³²P]dNTP (specific activity >400 Ci/mmol)

50 μ g/ml BSA

40 U reverse transcriptase

Incubate at 37°C for 30 min. Stop reaction by adding 2 μ l of 0.5 M EDTA or by heating to 75°C for 10 min. If necessary, the RNA template can be destroyed by adding 10 μ l of 5 M NaOH and incubating at 37°C overnight. The volume of reaction, concentration of 4 dNTPs, and temperature of the reaction will vary, depending upon the individual application.

Applications

1. Synthesizing cDNA for insertion into bacterial cloning vectors (see UNIT 5.5). Two different types of primers are used for the synthesis of cDNA: oligo(dT) primers and a population of randomly generated oligodeoxyribonucleotides. Oligo(dT) primers bind exclusively to the poly(A) tracts of mRNA; as a result, the synthesis of cDNA is biased towards the region corresponding to the 3' end of an mRNA template. In contrast, a mixed population of oligonucleotides will anneal at random positions on the RNA; consequently, the DNA synthesized will be complementary to all parts of the RNA template at a similar frequency (Taylor et al., 1976).
2. Filling in and labeling the 3' termini of DNA fragments with 5' protruding ends. Reverse transcriptase is useful for this reaction because, unlike *E. coli* DNA polymerase I or the Klenow fragment, it lacks a 3'→5' exonuclease activity.
3. For DNA sequencing, in place of Klenow fragment (UNIT 7.4). Due to its lack of a 3'→5' exonuclease, reverse transcriptase will often synthesize through regions that impede the progress of the Klenow fragment (Zagursky et al., 1985).

LITERATURE CITED

- Roth, M.J., Tanese, N., and Goff, S.P. 1985. Purification and characterization of murine retroviral reverse transcriptase expressed in *Escherichia coli*. *J. Biol. Chem.* 260:9326-9335.
- Tabor, S. and Richardson, C.C. 1987. DNA sequencing with a modified T7 DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* In Press.
- Taylor, J.M., Illmensee, R., and Summers, J. 1976. Efficient transcription of RNA into DNA by avian sarcoma virus polymerase. *Biochem. Biophys. Acta* 442:324-330.
- Verma, I.M. 1977. Reverse transcriptase. In *The Enzymes*, Vol. 14A (P.D. Boyer, ed.) pp. 87-104. Academic Press, NY.
- Zagurski, R.J., Baumeister, K., Lomax, N., and Berman, M.L. 1985. Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. *Gene Anal. Technol.* 2:89-94.

Contributed by Stanley Tabor
Harvard Medical School
Boston, Massachusetts

DNA-Dependent RNA Polymerases

There are two types of DNA-dependent RNA polymerases commonly used for nucleic acid research. One class is represented by *E. coli* RNA polymerase. This is a large, multisubunit enzyme which recognizes two DNA sequences, the -10 and -35 elements, within a large (~40 bp) promoter region. *E. coli* RNA polymerase was the DNA-dependent RNA polymerase used originally for the synthesis of transcripts *in vitro*. However, its use has several problems. It terminates prematurely *in vitro*; therefore, transcripts over 500 bases are difficult to obtain. Also, it has a relatively low specificity for promoter sequences; transcripts frequently initiate from other sequences, such as the 5' ends of the templates.

More recently, a different class of DNA-dependent RNA polymerases has become available. They are the bacteriophage T7, T3, and SP6 RNA polymerases, which are encoded by members of a related family of bacteriophages. Each is a single subunit enzyme that recognizes specific, 20 bp promoter sequences. Each initiates transcription specifically and exclusively from its own promoter sequence.

The phage RNA polymerases have a number of advantages over *E. coli* RNA polymerase for the synthesis of high levels of specific transcripts. First, they are extremely processive enzymes; transcripts thousands of bases in length are readily obtained without the enzyme dissociating from the template. Second, transcription is both extensive and rapid; from 2 µg of template DNA, 60 µg of specific transcripts can be generated in 30 min. Third, initiation is extremely specific for the individual promoter sequence. This allows the generation of radioactive RNA probes that are strand specific. Finally, the T7 and T3 RNA polymerases have been purified from cells overexpressing their genes. This results in homogeneous enzyme preparations that are inexpensive and have extremely high specific activities.

ENZYME

ESCHERICHIA COLI RNA POLYMERASE

The *E. coli* holoenzyme has five subunits (structure $\alpha_2\beta\beta'\omega$) and a total molecular weight of ~450,000. It will transcribe a given DNA template, native or denatured, into an RNA copy utilizing ribonucleoside triphosphates as precursors. Transcription is initiated preferentially from promoters that contain conserved sequences located 10 and 35 bp upstream from the start of transcription, and usually ceases *in vivo* at sequences called terminators. The specificity and extent of transcription will depend strongly on the quality of the DNA preparation, the strength of the promoter and terminator sequences, and the kind and concentration of mono- and divalent cations in the reaction mixture. The core enzyme, which lacks the ω subunit, does not recognize promoter sequences and hence initiates transcription randomly on the template.

Reaction Conditions for Promoter-Specific Transcription

For 50-µl reaction:

40 mM Tris·Cl, pH 8.0

10 mM MgCl₂

5 mM DTT

50 mM KCl

2 µg double-stranded DNA containing a promoter

300 µM 4 NTPs (UNIT 3.4)

50 µg/ml BSA

10 U *E. coli* RNA polymerase

Incubate 30 min at 37°C. Stop reaction by adding 2 µl of 0.5 M EDTA or by heating to 75°C for 10 min. The volume of reaction and concentration of DNA and 4 NTPs will vary, depending upon the individual application.

Applications

For most purposes, *E. coli* RNA polymerase is used for synthesizing transcripts only when it is not possible to clone the DNA into a vector containing a phage RNA polymerase promoter. *E. coli* RNA polymerase holoenzyme is useful for determining which segments of cloned DNA contain promoter sequences for expression in *E. coli*. The core enzyme, which lacks the σ subunit, synthesizes short transcripts relatively uniformly on the DNA template in the presence of high concentrations of random primers and low ribonucleoside triphosphate concentrations. The transcripts using native DNA (<500 nucleotides) are longer than those on denatured DNA (<100 nucleotides).

PHAGE RNA POLYMERASES: SP6, T7, T3

T7 and T3 RNA polymerases, the gene 1 products of the respective bacteriophage, were originally purified from infected *E. coli* cells. Now, both polymerases are purified from *E. coli* cells that are overexpressing the cloned genes (Davanloo et al., 1984; Tabor and Richardson, 1985; Morris et al., 1986). SP6 RNA polymerase is purified from bacteriophage SP6-infected *Salmonella typhimurium* (strain LT2) (Butler and Chamberlin, 1982).

All three of these homologous bacteriophage RNA polymerases are single polypeptides of molecular weight 90,000 to 100,000. Each is highly specific for its own promoter, which consists of a specific 20-bp sequence. Transcription is both very rapid (10 times the rate of *E. coli* RNA polymerase in vitro) and extremely processive.

Reaction Conditions

For 50-µl reaction:

40 mM Tris-Cl, pH 7.5

10 mM MgCl₂

5 mM DTT

2 µg DNA template containing appropriate phage promoter

400 µM 4 NTPs (UNIT 3.4)

50 µg/ml BSA

1 mM spermidine (for SP6 reaction mixture only)

10 U RNA polymerase (T7, T3, or SP6)

Incubate at 37°C for 30 min. Stop reaction by adding 2 µl of 0.5 M EDTA or by heating to 75°C for 10 min. The volume of reaction and concentration of DNA and 4 dNTPs will vary, depending upon the individual application. The reaction conditions given above are used to synthesize large amounts of unlabeled transcripts; over 60 µg of transcripts can be generated under these conditions. For producing labeled transcripts of high specific activity, the concentration of labeled ribonucleoside triphosphate should be reduced to 10 µM. Spermidine, which is helpful for SP6 RNA polymerase, should be added to the reaction mixture at room temperature in order to prevent precipitation of DNA.

Applications

Phage T7, T3, and SP6 RNA polymerases are used for extensive, highly specific transcription of DNA sequences inserted downstream from the appropriate T7, T3, or SP6 promoter (Fig. 3.8.1). Plasmids have been constructed with polylinker cloning sites

ENZYME

Enzymatic
Manipulation
of DNA and RNA

3.8.2

adjacent to these promoters (see *UNIT 1.5*). Often the plasmid is cleaved with a restriction enzyme prior to transcription, resulting in the synthesis of “runoff” transcripts that terminate at the end of the DNA template.

1. Phage polymerases are used to generate homogeneous single-stranded RNA probes that are uniformly labeled to a high specific activity. These are useful for detection of homologous DNA or RNA sequences by standard hybridization techniques. With use of [α - 32 P]UTP, over 90% of the radioactivity can be incorporated into the transcripts. For hybridization experiments, RNA probes, because they are single stranded, result in a greater sensitivity as compared to nick-translated DNA probes, which are double-stranded and hence can self-anneal.
2. Uniformly labeled transcripts are used to map the ends of RNA or DNA (*UNIT 4.7*). After hybridization is allowed to proceed to completion, the products are cleaved with ribonucleases under conditions where double-stranded RNA is not cleaved. This procedure is similar to mapping the ends of RNA and DNA using end-labeled DNA fragments and S1 nuclease (*UNIT 4.6*).
3. Uniformly labeled transcripts are used for genomic DNA sequencing.
4. Transcripts synthesized by phage polymerases are used as precursor RNAs for studies of RNA splicing and processing.
5. The RNA transcripts can be translated *in vitro* in the presence of radioactively labeled amino acids to generate radiopure proteins. Proteins generated in this manner can be tested for their functional activity and are useful for substrates for a variety of *in vitro* systems (Hope and Struhl, 1985).
6. T7 RNA polymerase can be used *in vivo* to express cloned genes at very high levels exclusively under the control of a T7 RNA polymerase promoter (Tabor and Richardson, 1985; Studier and Moffatt, 1986).

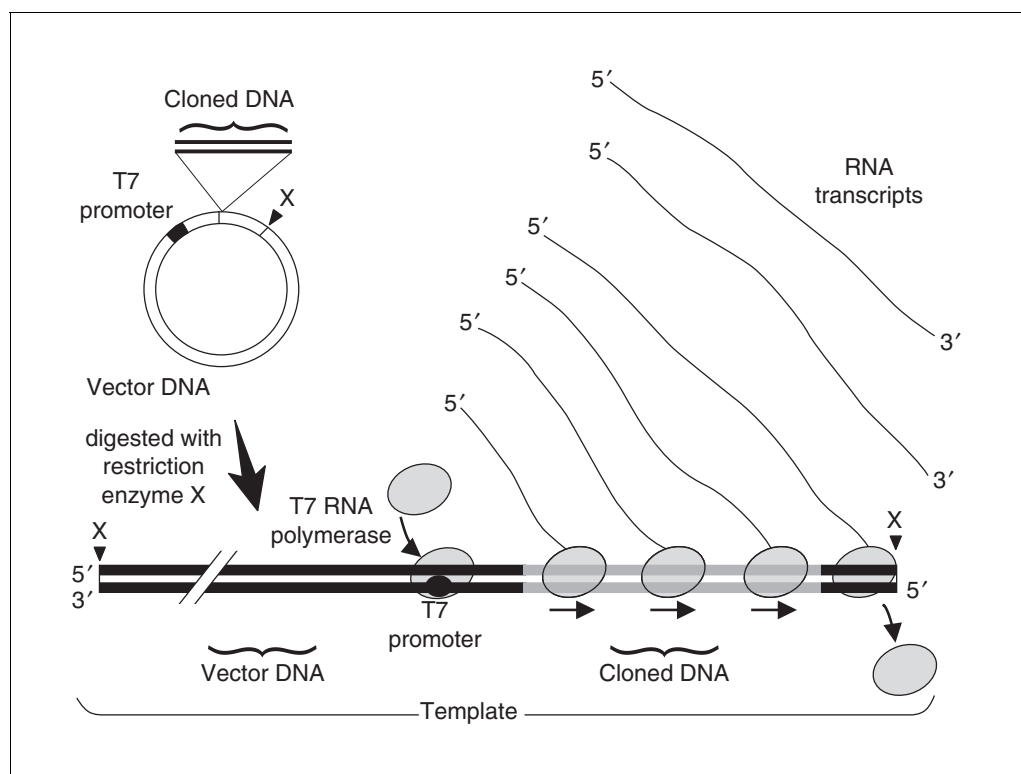


Figure 3.8.1 Runoff transcripts.

LITERATURE CITED

- Butler, E.T. and Chamberlin, M. 1982. Bacteriophage SP6-specific RNA polymerase. I. Isolation and characterization of the enzyme. *J. Biol. Chem.* 257:5772-5778.
- Chamberlin, M.J. and Ryan, T. 1982. Bacteriophage DNA-dependent RNA polymerases. *In The Enzymes*, Vol. 15B (P.D. Boyer, ed.) pp. 87-109. Academic Press, NY.
- Davanloo, P., Rosenberg, A.H., Dunn, J.J., and Studier, F.W. 1984. Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 81:2035-2039.
- Green, M.R., Maniatis, T., and Melton, D.A. 1983. Human β -globin pre-mRNA synthesized in vitro accurately spliced in *Xenopus* oocyte nuclei. *Cell* 32:681-694.
- Hope, I.A. and Struhl, K. 1985. GCN4 protein, synthesized in vitro, binds HIS3 regulatory sequences: Implications for general control of amino acid biosynthesis in yeast. *Cell* 43:177-188.
- Krieg, P.A. and Melton, D.A. 1984. Functional messenger RNAs are produced by SP6 in vitro transcription of cloned cDNAs. *Nucl. Acids Res.* 12:7057-7070.
- Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K., and Green, M.R. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucl. Acids Res.* 12:7035-7056.
- Morris, C.E., Klement, J.F., and McAllister, W.T. 1986. Cloning and expression of the bacteriophage T3 RNA polymerase gene. *Gene* 41:193-200.
- Studier, F.W. and Moffatt, B.A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189:113-130.
- Tabor, S. and Richardson, C.C. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. U.S.A.* 82:1074-1078.

KEY REFERENCE

- Chamberlin, M.J. 1982. Bacterial DNA-dependent RNA polymerases. *In The Enzymes*, Vol. 15B (P.D. Boyer, ed.) pp. 61-86. Academic Press, NY.

Contributed by Stanley Tabor
Harvard Medical School
Boston, Massachusetts

DNA-Independent RNA Polymerases

ENZYME

POLY(A) POLYMERASE

The enzyme, purified from *E. coli*, catalyzes the incorporation of AMP residues onto the free 3'-hydroxyl terminus of RNA, utilizing ATP as a precursor (Fig. 3.9.1; Edmonds, 1982).

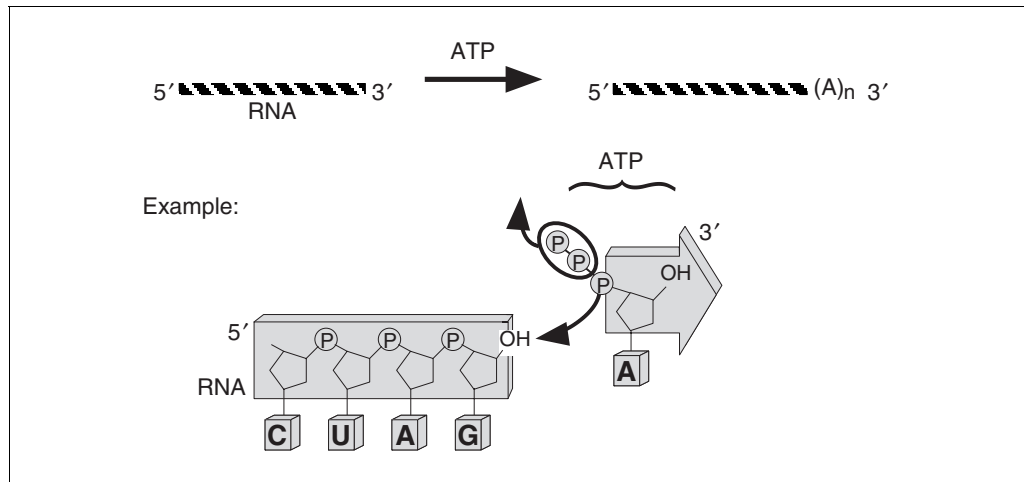


Figure 3.9.1 Poly(A) polymerase activity.

Reaction Conditions

For 50- μ l reaction:
 40 mM Tris-Cl, pH 8.0
 10 mM MgCl₂
 2.5 mM MnCl₂
 250 mM NaCl
 250 μ g/ml RNA
 250 μ M ATP (UNIT 3.4)
 50 μ g/ml BSA
 5 U poly(A) polymerase

Incubate at 37°C for 30 min. Stop reaction by adding 2 μ l of 0.5 M EDTA or by heating to 75°C for 10 min. The volume of reaction and concentration of ATP will vary, depending upon the individual application.

Applications

1. Labeling the 3' ends of RNA with [α^{32} P]ATP. Labeled RNA prepared by this procedure has been used for hybridization probes. For cloned genes, however, the synthesis of RNA probes using phage RNA polymerases is much more efficient and produces RNA with a much higher specific activity. For labeling of cellular RNA, cDNA synthesis using reverse transcriptase is the preferred method.
2. Cloning RNA that lacks a poly(A) tail (Gething et al., 1980). A poly(A) tail is synthesized using a poly(A) polymerase, and subsequently a cDNA copy is made by using oligo(dT) as a primer.

LITERATURE CITED

- Edmonds, M. 1982. Poly(A) adding enzymes. *In* The Enzymes, Vol. 15B (P.D. Boyer, ed.) pp. 218-245. Academic Press, NY.
- Gething, M.J., Bye, J., Skehel, J. and Waterfield, M. 1980. Cloning and DNA sequence of double-stranded haemagglutinin genes from H2 and H3 strains elucidates antigenic shift and drift in human influenza virus. *Nature (London)* 287:301-306.

Contributed by Stanley Tabor
Harvard Medical School
Boston, Massachusetts

Phosphatases and Kinases

ENZYME

ALKALINE PHOSPHATASES: BACTERIAL ALKALINE PHOSPHATASE AND CALF INTESTINE PHOSPHATASE

Bacterial alkaline phosphatase (BAP) from *E. coli* and calf intestine phosphatase (CIP) from veal are commonly used in nucleic acid research. Both enzymes catalyze the hydrolysis of 5'-phosphate residues from DNA, RNA, and ribo- and deoxyribonucleoside triphosphates (Fig. 3.10.1). The dephosphorylated products possess 5'-hydroxyl termini which can subsequently be radioactively labeled using [γ - 32 P]ATP and T4 polynucleotide kinase (p. 3.10.3).

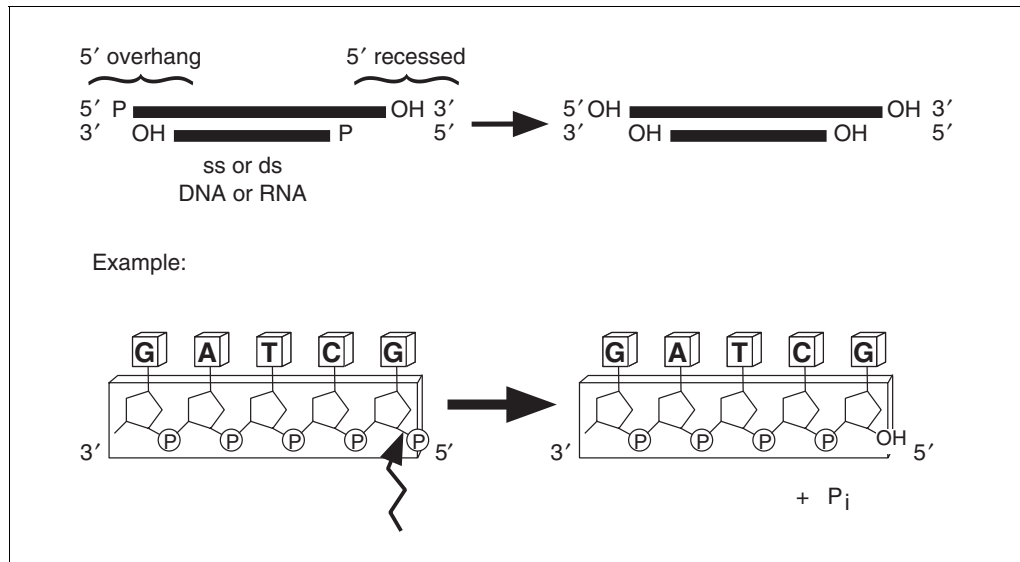


Figure 3.10.1 Phosphatase activity.

Both phosphatases require Zn^{2+} for activity. The primary difference between them is the stability of the two enzymes. CIP is readily inactivated by heating to 70°C for 10 min and/or extraction with phenol. On the other hand, BAP is much more resistant to these treatments. Thus, for most purposes, CIP is the enzyme of choice. Furthermore, CIP has a 10- to 20-fold higher specific activity than BAP.

BAP Reaction Conditions for Dephosphorylation of DNA

For 50- μ l reaction:

- 50 mM Tris-Cl, pH 8.0
- 1 mM $ZnCl_2$
- 1 to 20 pmol DNA termini
- 0.1 U BAP

Incubate at 60°C for 30 min. Stop reaction by adding 0.1% SDS and 100 μ g/ml proteinase K and incubating at 37°C for 30 min. Then, extract twice with phenol and precipitate the DNA with ethanol (UNIT 2.1).

CIP Reaction Conditions for Dephosphorylation of DNA

For 50- μ l reaction:

- 20 mM Tris-Cl, pH 8.0
- 1 mM $MgCl_2$
- 1 mM $ZnCl_2$
- 1 to 20 pmol DNA termini
- 0.1 U CIP

Incubate at 37°C for 30 min. Stop reaction by heating to 75°C for 10 min or extracting with phenol, and then precipitate with ethanol. *NOTE: CIP is heat labile; 10 min at 75°C effectively inactivates CIP.*

The volume of reaction and concentration of DNA will vary, depending upon the individual application. In general, phosphatase treatment can be done directly following cleavage by a restriction endonuclease, thus minimizing the number of manipulations. A useful conversion factor is that 1 µg of a 3-kb linear DNA contains 1 pmol of 5' termini.

Applications

1. Dephosphorylation of 5' termini of nucleic acids prior to labeling with [γ - 32 P]ATP and T4 polynucleotide kinase. 5'- 32 P end-labeled DNA is used for sequencing by the Maxam-Gilbert procedure (UNIT 7.5), RNA sequencing by specific RNase digestions, and in mapping studies using specific DNA or RNA fragments.
2. Dephosphorylation of 5' termini of vector DNA in order to prevent self-ligation of vector termini (UNIT 3.16).

T4 POLYNUCLEOTIDE KINASE

T4 polynucleotide kinase, the product of the phage T4 *pseT* gene, was originally purified from T4-infected *E. coli* cells (Richardson, 1981). Recently, the *pseT* gene has been cloned into *E. coli*, and the enzyme has been overproduced from this strain (Midgley and Murray, 1985).

The forward reaction of T4 polynucleotide kinase catalyzes the transfer of the terminal (γ) phosphate of ATP to the 5'-hydroxyl termini of DNA and RNA (Fig. 3.10.2). This reaction is very efficient and hence is the general method for labeling 5' ends or for phosphorylating oligonucleotides.

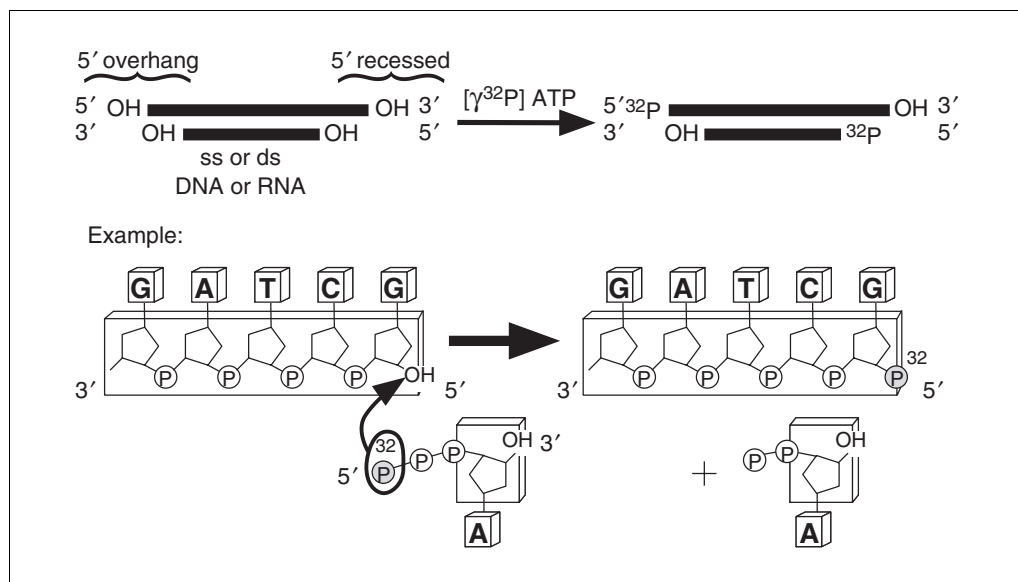


Figure 3.10.2 Kinase activity.

The exchange reaction of T4 polynucleotide kinase catalyzes the exchange of 5'-terminal phosphates. In this reaction, which requires an excess of ADP, the 5'-terminal phosphate is transferred to ADP and subsequently rephosphorylated by the transfer from the γ phosphate of [γ - 32 P]ATP (Fig. 3.10.3). The exchange reaction is less efficient than the forward reaction; thus, it is rarely used.

ENZYME

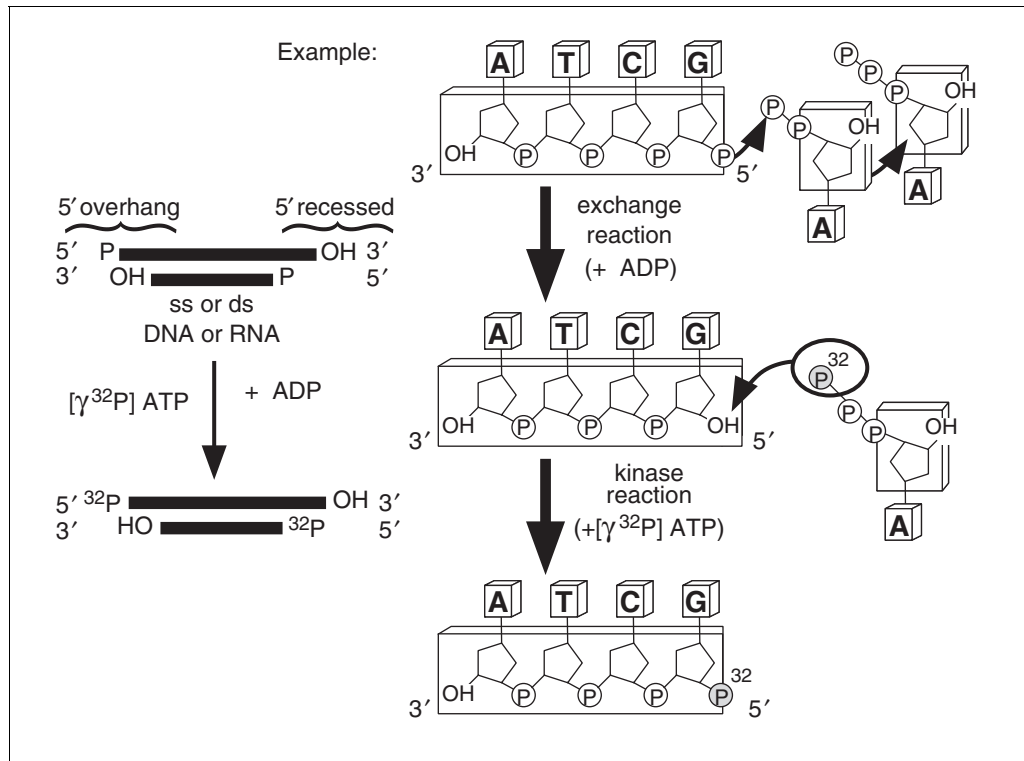


Figure 3.10.3 Kinase exchange activity.

Finally, polynucleotide kinase is a 3' phosphatase. Some commercial preparations of polynucleotide kinase are prepared from the phage T4 strain am N81 *pseT1*, which has a mutated *pseT* gene. This derivative lacks the 3' phosphatase activity.

**BASIC
PROTOCOL**

Labeling 5' Ends by the Forward Reaction

For 30- μl reaction:

50 mM Tris-Cl pH 7.5

10 mM MgCl_2

5 mM DTT

1 to 50 pmol dephosphorylated DNA, 5' ends

50 pmol (150 μCi) $[\gamma\text{-}^{32}\text{P}] \text{ATP}$ (specific activity $>3000 \text{ Ci/mmol}$)

50 $\mu\text{g/ml}$ BSA

20 U T4 polynucleotide kinase

Incubate at 37°C for 60 min. Stop reaction by adding 1 μl of 0.5 M EDTA or by heating to 75°C for 10 min. Extract with phenol/chloroform. Separate the labeled DNA from the unincorporated labeled nucleotides by filtration on Sephadex G-100 (UNIT 3.4) or by centrifugation through a spin column containing Sephadex G-50 (UNIT 3.4).

The volume of reaction and the concentration of DNA and $[\gamma\text{-}^{32}\text{P}] \text{ATP}$ will vary, depending upon the application.

Phosphorylating Synthetic Oligonucleotides by the Forward Reaction

For 30- μ l reaction:

50 mM Tris·Cl, pH 7.5
10 mM MgCl₂
5 mM DTT
1 to 10 μ g oligonucleotide linker
1 mM ATP
50 μ g/ml BSA
20 U T4 polynucleotide kinase

Incubate at 37°C for 60 min. Stop reaction by adding 1 μ l of 0.5 M EDTA. If desired, [γ -³²P]ATP can be added to trace the reaction.

Labeling 5' Termini by the Exchange Reaction

For 30- μ l reaction:

50 mM imidazole·Cl, pH 6.6
10 mM MgCl₂
5 mM DTT
1 to 50 pmol phosphorylated DNA, 5' ends
5 mM ADP
60 pmol (180 μ Ci) [γ -³²P]ATP (specific activity >3000 Ci/mmol)
50 μ g/ml BSA
20 U T4 polynucleotide kinase

Incubate at 37°C for 60 min. Stop reaction by adding 1 μ l of 0.5 M EDTA. Extract with buffered phenol. Separate the labeled DNA from the unincorporated labeled nucleotides by either Sephadex G-100 gel filtration (UNIT 3.4), or centrifugation through a spin column containing Sephadex G-50 (UNIT 3.4).

The volume of reaction and the concentration of DNA and [γ -³²P]ATP will vary, depending upon the application.

Critical Parameters

1. The ATP concentration should be at least 1 μ M for the forward reaction and 2 μ M for the exchange reaction.
2. T4 polynucleotide kinase prepared from an *E. coli* strain overproducing the enzyme should be used. This enzyme has a higher purity, specific activity, concentration, and is less expensive than that obtained from T4 infected cells.
3. The DNA should be well purified. Contamination by small oligonucleotides or tRNA (often used for the purification of small DNA fragments) can significantly lower the efficiency with which the desired DNA is labeled. This is because oligonucleotides can contribute a large number of 5' ends even if they constitute a small percentage of the total weight of the preparation.
4. The enzyme is inhibited by low levels of phosphate buffer or ammonium salts. In addition, the enzyme works extremely poorly on DNAs that have been purified from agarose gels (even the more purified low gelling/melting temperature agarose). To radioactively label such DNA, it is necessary to perform further purification procedures such as chromatography on DEAE cellulose (UNIT 2.6) or hydroxylapatite. DNAs purified by electrophoresis in polyacrylamide gels also sometimes require further purification for effective labeling.
5. For linear pBR322, 1 pmol of 5' ends = 1.6 μ g DNA.

**BASIC
PROTOCOL**

**BASIC
PROTOCOL**

**Enzymatic
Manipulation
of DNA and RNA**

3.10.4

6. 5' protruding ends are phosphorylated more efficiently than blunt ends, and labeling of 5' recessed ends is very inefficient. Labeling of blunt or recessed ends can be improved by denaturing the DNA (by boiling or by NaOH treatment) and then immediately carrying out the reaction. However, if alternative strategies are possible, it is preferable to avoid labeling of 5' recessed ends by T4 polynucleotide kinase.
7. Recently, it has been shown that macromolecular exclusion molecules (e.g., PEG 8000) greatly improve the rate and overall efficiency of all the kinase reactions. For example, PEG 8000 caused a 1000-fold increase in the rate of labeling of a *Pst*I restriction fragment with recessed 5' termini, and the overall efficiency was increased from 20 to 70% using much lower enzyme levels (Harrison and Zimmerman, 1986).

Applications

1. DNA sequencing by the chemical degradation (Maxam-Gilbert) technique.
2. Defining specific protein–DNA interactions by DNase I footprinting or protection from DNA-damaging chemicals such as dimethyl sulfate or methidiumpropyl-EDTA (MPE).
3. Mapping of restriction sites by partial digestion of 5' end-labeled DNA fragments (*UNIT 3.3*).
4. Mapping the termini of RNA transcripts (*UNIT 4.6*).
5. Mapping the position of introns in DNA.
6. Synthesis of substrates for the assay of DNA and RNA ligases.
7. Labeling oligonucleotides for purification by gel electrophoresis (*UNIT 2.7*).
8. Ligation of oligonucleotides into DNA vectors; when oligonucleotides are chemically synthesized, they have 5'-hydroxyl termini; for cloning purposes, it is often desirable to phosphorylate them (*UNIT 3.16*).

LITERATURE CITED

- Chaconas, G. and van de Sande, J.H. 1980. 5'-³²P labeling of RNA and DNA restriction fragments. *Meth. Enzymol.* 65:75-88.
- Harrison, B. and Zimmerman, S.B. 1986. T4 polynucleotide kinase: Macromolecular crowding increases the efficiency of reaction at DNA termini. *Anal. Biochem.* 158:307-315.
- Midgley, C.A. and Murray, N.E. 1985. T4 polynucleotide kinase; cloning of the gene (*pseT*) and amplification of its product. *EMBO J.* 4:2695-2703.
- Reid, T.W. and Wilson, I.B. 1971. *Escherichia coli* alkaline phosphatase. In *The Enzymes*, Vol. 4 (P.D. Boyer, ed.) pp. 373-415. Academic Press, NY.
- Richardson, C.C. 1981. Bacteriophage T4 polynucleotide kinase. In *The Enzymes*, Vol. 14A (P.D. Boyer, ed.) pp. 299-314, Academic Press, NY.

Contributed by Stanley Tabor
Harvard Medical School
Boston, Massachusetts

Exonucleases

SINGLE-STRANDED 5'→3' and 3'→5' EXONUCLEASES

Exonuclease VII (exo VII) (Vales et al., 1982)

ENZYME

Exonuclease VII from *E. coli* consists of two subunits, the products of the *xseA* and *xseB* genes. It is a processive single-stranded exonuclease that acts from both the 3' and 5' ends of single-stranded DNA (Fig. 3.11.1). The products of exo VII are small oligonucleotides. Exo VII is unique among the nucleases discussed here in that it does not require Mg^{2+} ; it retains full activity in the presence of 10 mM EDTA.

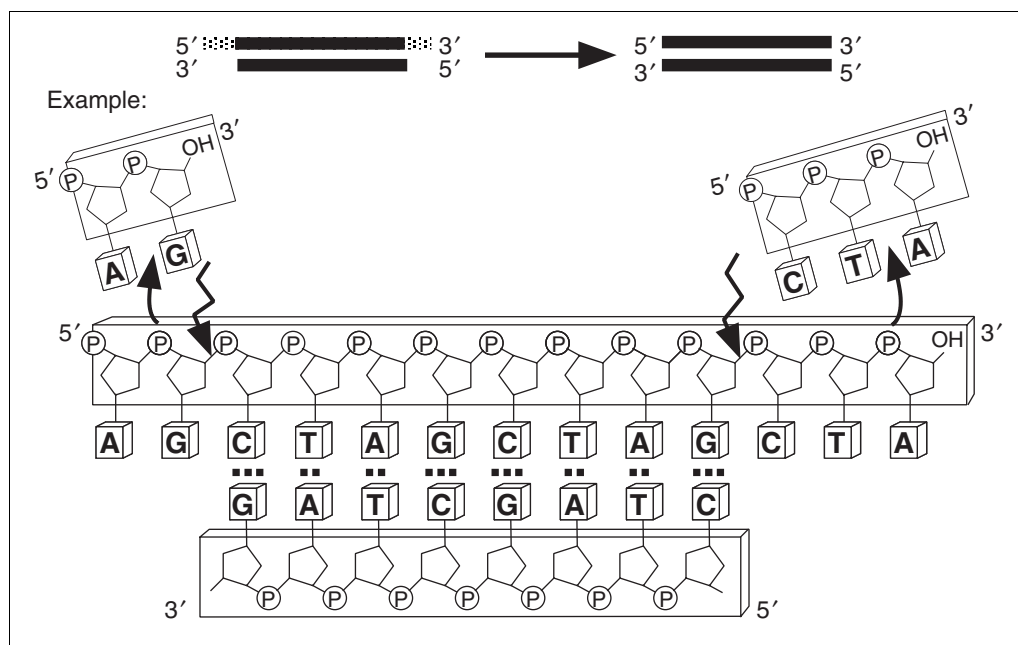


Figure 3.11.1 Exonuclease VII activity.

Reaction conditions

- For 50- μ l reaction:
- 70 mM Tris·Cl, pH 8
 - 8 mM EDTA
 - 10 mM β -mercaptoethanol
 - 1 μ g DNA
 - 50 μ g/ml BSA
 - 0.2 U exo VII

Incubate at 37°C for 30 min. Stop reaction by extraction with phenol and precipitation with ethanol. Note that exo VII is not inhibited by EDTA.

Applications

1. For mapping the positions of introns in genomic DNA (Berk and Sharp, 1977) (UNIT 4.6).
2. To excise segments of DNA that have been inserted into plasmid vectors by the poly(dA-dT) tailing method (Goff and Berg, 1978).

DOUBLE-STRANDED 5'→3' EXONUCLEASES

ENZYME

Lambda Exonuclease (λ exo) (Little, 1981)

This exonuclease, which is purified from *E. coli* cells that have been infected with bacteriophage λ , catalyzes the stepwise and processive hydrolysis of duplex DNA from 5'-phosphoryl termini liberating 5' mononucleotides (Fig. 3.11.2). λ exo will not degrade 5'-hydroxyl termini.

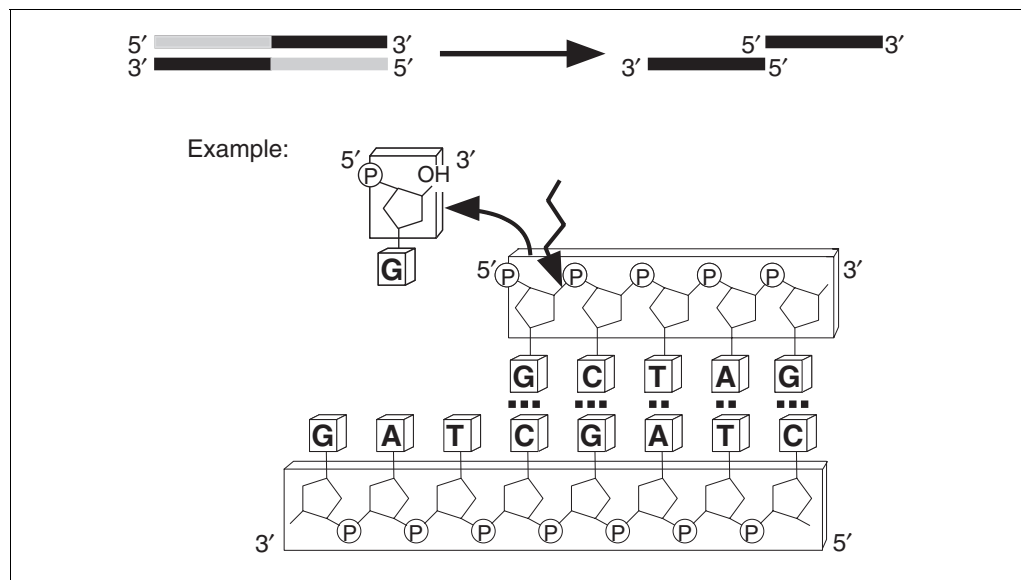


Figure 3.11.2 5'→3' exonuclease activity: T7 gene 6 exonuclease and λ exonuclease.

Reaction conditions

For 50- μ l reaction:

67 mM glycine·KOH, pH 9.4

2.5 mM MgCl₂

2 μ g DNA

50 μ g/ml BSA

10 U λ exo

Incubate at 37°C for 1 to 30 min, depending upon extent of digestion required. Stop reaction by adding 2 μ l of 0.5 M EDTA or by heating to 75°C for 10 min.

Applications

1. Converting double-stranded DNA into single-stranded DNA that can be subjected to the dideoxy sequencing method (UNIT 7.4).
2. Removing 5' protruding ends from duplex DNA for tailing with terminal transferase.

ENZYME

T7 Gene 6 Exonuclease (Kerr and Sadowski, 1972)

This enzyme is the product of gene 6 of bacteriophage T7. Gene 6 has been cloned onto a plasmid, and the enzyme is purified from overexpressing *E. coli* cells. The gene 6 exonuclease is similar to λ exo in that it catalyzes the stepwise hydrolysis of duplex DNA from the 5' termini liberating 5' mononucleotides. However, unlike λ exo, the enzyme has low processivity and it will remove both 5'-hydroxyl and 5'-phosphoryl termini (Thomas and Olivera, 1978).

Reaction conditions

For 50- μ l reaction:

50 mM Tris·Cl, pH 7.5

5 mM MgCl₂

5 mM DTT

20 mM KCl

2 μ g DNA

50 μ g/ml BSA

5 U T7 gene 6 exo

Incubate at 37°C for 1 to 30 min, depending upon extent of digestion required. Stop reaction by adding 2 μ l of 0.5 M EDTA or by heating to 75°C for 10 min.

Applications

T7 gene 6 exonuclease is used for the same purposes as described for λ exonuclease. However, the T7 enzyme, because it is less processive than λ exonuclease, is preferred for the controlled and uniform digestion from 5' ends. In addition, the T7 enzyme can degrade from 5'-phosphoryl termini. Finally, the T7 enzyme is available from an over-producing strain; thus, it is more homogeneous and less expensive than enzyme from T7 infected cells. It is currently available from U.S. Biochemical.

DOUBLE-STRANDED 3'→5' EXONUCLEASES

Exonuclease III (exo III) (Rogers and Weiss, 1980)

Exonuclease III, the product of the *E. coli xthA* gene, is made from *E. coli* cells overproducing the protein. Exo III is a multifunctional enzyme that catalyzes hydrolysis of several types of phosphodiester bonds in double-stranded DNA. The main application of exo III is as a 3'→5' double-strand specific exonuclease that catalyzes release of 5' nucleotides from the 3'-hydroxy end of double-stranded DNA (Fig. 3.11.3).

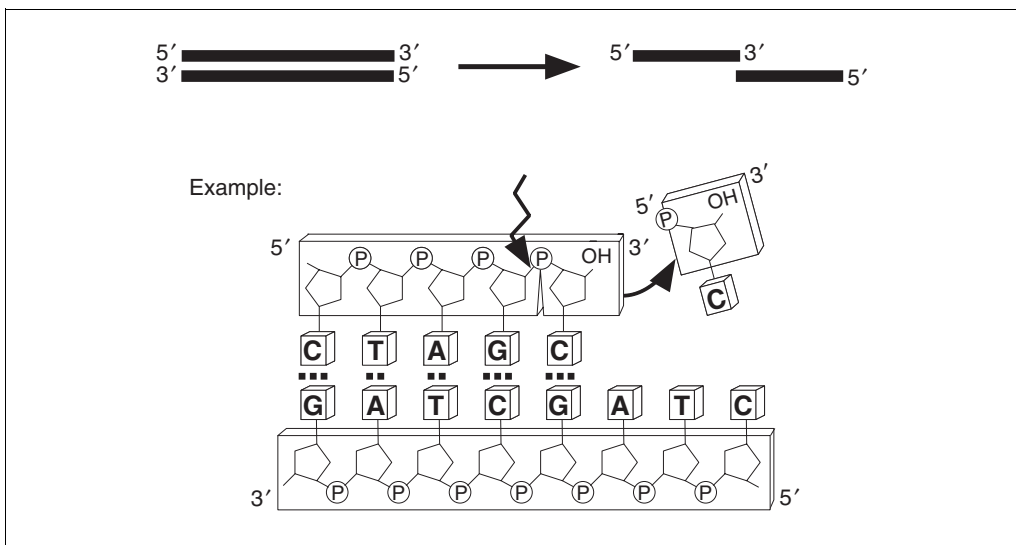


Figure 3.11.3 Exonuclease III activity.

The exonuclease activity is nonprocessive, which is ideal for generating uniform single-stranded regions in double-stranded DNA. Its rate of degradation is dependent on the base composition of the DNA: C>>A~T>>G. As a result, different termini will be degraded at different rates.

In addition to the 3'→5' exonuclease, exo III has three other activities: a 3' phosphatase, an RNase H, and an endonuclease specific for apurinic or apyrimidinic sites.

ENZYME

Enzymatic
Manipulation
of DNA and RNA

3.11.3

Reaction conditions

For 50- μ l reaction:

50 mM Tris-Cl, pH 7.5

5 mM MgCl₂

5 mM DTT

2 μ g DNA

50 μ g/ml BSA

10 U exo III

Incubate at 37°C for 1 to 30 min, depending upon the extent of digestion required. Stop the reaction by adding 2 μ l of 0.5 M EDTA or by heating to 75°C for 10 min. One unit of exo III will remove ~ 200 nucleotides from each 3' recessed end of 1 μ g of a 5000-bp linear double-stranded DNA template in 10 min at 37°C.

Applications

All of the applications described below utilize the nonprocessive 3'→5' double-stranded exonuclease activity of exo III to generate uniform single-stranded regions in double-stranded DNA.

1. Preparing strand-specific radioactive probes, in conjunction with the Klenow fragment; this procedure is analogous to replacement synthesis using the exonuclease and polymerase functions of T4 DNA polymerase (Fig. 3.5.5).
2. Preparing single-stranded DNA templates for sequencing by the dideoxy technique (Sanger et al., 1977) (UNIT 7.4).
3. Constructing unidirectional deletions; since exo III is double-strand specific, it will degrade preferentially from duplex DNA having a 3'-recessed end as opposed to one with a 5'-recessed end (Henikoff, 1984). This property is useful for constructing a set of unidirectional deletions from a given position in cloned DNA, that are used for DNA sequencing without prior restriction site mapping.

LITERATURE CITED

- Berk, A.J. and Sharp, P.A. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. *Cell* 12:721-732.
- Goff, S. and Berg, P. 1978. Excision of DNA segments introduced into cloning vectors by the poly(dA-dT) joining method. *Proc. Natl. Acad. Sci. U.S.A.* 75:1763-1767.
- Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28:351-359.
- Kerr, C. and Sadowski, P.D. 1972. Gene 6 exonuclease of bacteriophage T7. I. Purification and properties of the enzyme. *J. Biol. Chem.* 247:305-310.
- Little, J.W. 1981. Lambda exonuclease. In *Gene Amplification and Analysis, Vol. 2: Structural analysis of nucleic acids* (J.G. Chirikjian and T.S. Papas, eds.) pp. 135-145. Elsevier/North Holland, NY, Amsterdam, Oxford.
- Rogers, S.G. and Weiss, B. 1980. Exonuclease III of *Escherichia coli* K-12, an AP exonuclease. *Meth. Enzymol.* 65:201-211.
- Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74:5463-5467.
- Thomas, K.R. and Olivera, B.M. 1978. Processivity of DNA exonucleases. *J. Biol. Chem.* 253:424-429.
- Vales, L.D., Rabin, B.A., and Chase, J.W. 1982. Subunit structure of *Escherichia coli* exonuclease VII. *J. Biol. Chem.* 257:8799-8805.

Contributed by Stanley Tabor
Harvard Medical School
Boston, Massachusetts

Endonucleases

BAL 31 NUCLEASE

ENZYME

Bal 31 nuclease from *Alteromonas espejiana* is a single strand-specific endodeoxyribonuclease (Gray et al., 1981; Wei et al., 1983) (Fig. 3.12.1). On duplex circular DNA, it degrades at nicks or at transient single-stranded regions created by supercoiling (Lau and Gray, 1979). On duplex linear DNA, it degrades from both the 5' and 3' termini at both ends, resulting in a controlled shortening of the DNA.

Bal 31 also acts as a ribonuclease, catalyzing the hydrolysis of ribosomal and tRNA (Gray et al., 1975). *Bal 31* nuclease requires both Ca^{++} and Mg^{++} . Digestions can be terminated with EGTA, a specific chelator of Ca^{++} , without affecting the Mg^{++} concentration. *Bal 31* nuclease is active in SDS and urea (Gray et al., 1975).

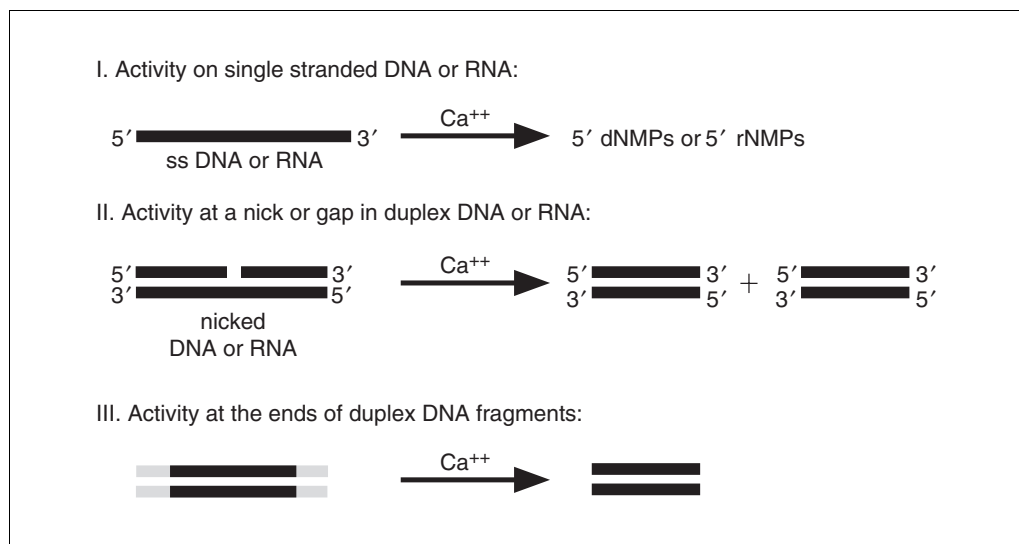


Figure 3.12.1 *Bal 31* activities.

Reaction Conditions

For 50- μl reaction:

50 mM Tris-Cl, pH 7.5

10 mM CaCl_2

10 mM MgCl_2

600 mM NaCl

2 μg DNA

50 $\mu\text{g/ml}$ BSA

10 U *Bal 31* nuclease

Incubate at 30°C for 1 to 30 min, depending upon extent of digestion required. Stop the reaction by adding 5 μl of 0.5 M EDTA or by heating to 75°C for 10 min, and analyze products by agarose gel electrophoresis (UNIT 2.5). Due to the high NaCl concentration in the *Bal 31* buffer, small volumes of the reaction mixture should be loaded. For subsequent enzymatic treatments, it is usually necessary to remove the NaCl by precipitating the DNA with ethanol (UNIT 2.1).

One unit is defined as the amount of enzyme that catalyzes the removal of 200 bp from each end of linearized pBR322 in 10 min at 30°C using a DNA concentration of 50 $\mu\text{g/ml}$ in a 50- μl reaction.

Contributed by Stanley Tabor and Kevin Struhl

Current Protocols in Molecular Biology (1989) 3.12.1-3.12.6

Copyright © 2000 by John Wiley & Sons, Inc.

Critical Parameters

1. *Bal* 31 activity is inhibited by RNA that contaminates crude DNA preparations. Thus, DNA should be purified using CsCl gradients or PEG precipitation (UNIT 1.7), or gel-purified (UNIT 2.6).
2. The unit definition for *Bal* 31 nuclease should be used only as a rough estimate. For each preparation of DNA, the extent of digestion should be monitored by agarose gel electrophoresis (UNIT 2.5).
3. To reduce the rate of digestion, lower the temperature of the reaction. A temperature of 20°C reduces the rate of *Bal* 31 nuclease 3-fold from that observed at 30°C.
4. *Bal* 31 activity will vary with different DNA templates. GC-rich regions are degraded more slowly than AT-rich regions. Consequently, the two ends of a given DNA fragment can be degraded at different rates. Also, a series of deletions of a given DNA fragment will be nonuniform at the nucleotide level; some sequences will have a very low frequency of appearance at the junction of the deletion.
5. *Bal* 31-digested DNA fragments can be ligated directly at low frequency. To increase the efficiency of ligation, digested DNA should be extracted with phenol and precipitated with ethanol (UNIT 2.1). Then the ends of the DNA should be repaired with Klenow fragment or T4 DNA polymerase (UNIT 3.5), prior to the ligation reaction with T4 DNA ligase (UNIT 3.16).

Applications

1. Cloning for creating deletions of different sizes in a controlled manner. The cloned DNA is resected at specific sites with a restriction enzyme, and the resulting linear DNA is digested with *Bal* 31 nuclease for various intervals of time (UNIT 7.3).
2. Mapping restriction sites in a DNA fragment (Legerski et al., 1978).
3. Investigating secondary structure of supercoiled DNA (Lau and Gray, 1979) and alterations in the helix structure of duplex DNA caused by treatment with mutagenic agents (Legerski et al., 1977).
4. Creating successive deletions needed for linker-scanning mutagenesis (UNIT 8.4).

ENZYME

S1 NUCLEASE

S1 nuclease from *Aspergillus oryzae* is a highly specific single-stranded endonuclease (Vogt, 1980; Fig. 3.12.2). The rate of digestion of single-stranded DNA is 75,000 times faster than that of double-stranded DNA.

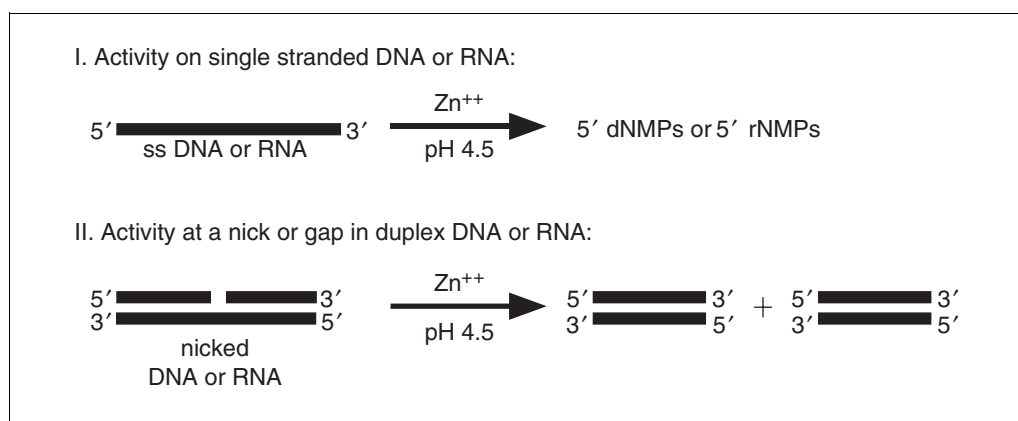


Figure 3.12.2 S1 nuclease activities.

The pH optimum is 4.0 to 4.3, and the rate drops 50% at pH 4.9. Reactions are normally run at pH 4.6 to avoid depurination of the DNA. The enzyme requires low levels of Zn⁺⁺, and it is largely unaffected by NaCl concentrations in the range of 10 to 300 mM. It is stable to urea, SDS, and formamide.

Reaction Conditions

For 100- μ l reaction:

50 mM sodium acetate, pH 4.6

1 mM zinc acetate

250 mM NaCl

2 μ g DNA

50 μ g/ml BSA

10 U S1 nuclease

Incubate at 37°C for 30 min. Stop reaction by adding 1 μ l of 0.5 M EDTA. The volume of reaction, amount of DNA, units of S1 nuclease, temperature, and time of reaction will vary with different applications of S1 nuclease.

Applications

Most applications of S1 nuclease make use of its ability to trim protruding single-stranded ends of DNA and RNA without significant nibbling of blunt duplex ends.

1. Mapping the 5' and 3' ends of RNA transcripts by the analysis of S1-resistant RNA:DNA hybrids (*UNIT 4.6*).
2. Mapping the location of introns by digesting a hybrid of mature mRNA with ³²P-labeled genomic DNA. S1 will cleave at the single-stranded loops created by introns within these hybrid molecules.
3. Digesting the hairpin structures formed during the synthesis of cDNA by reverse transcriptase (*UNIT 5.5*).
4. Removing single-stranded termini of DNA fragments to produce blunt ends for ligation (*UNIT 3.16*).
5. Creating small deletions at restriction sites. It will nibble at the ends of a linear duplex fragment at a very slow rate.
6. Rendering the ends flush after successive deletions by exo III for linker-scanning mutagenesis (*UNIT 8.4*).

MUNG BEAN NUCLEASE

Mung bean nuclease (from sprouts of mung bean) is a highly specific single-stranded endonuclease with properties similar to those of S1 nuclease (Kroeker et al., 1976). Reactions are normally performed at pH 5.0, which is optimal for degrading single-stranded DNA. Activity against double-stranded DNA is essentially negligible at pH 5.0, but it increases as the pH is reduced below this point. Mung bean nuclease activity is significantly decreased at NaCl concentrations >200 mM.

One unit is defined as the amount of enzyme that produces 1 μ g of acid-soluble material in 1 min at 37°C using single-stranded salmon sperm DNA as the substrate.

ENZYME

Reaction Conditions

For 100- μ l reaction:

30 mM sodium acetate, pH 5.0

50 mM NaCl

1 mM zinc acetate

5% (vol/vol) glycerol

1 μ g DNA

50 μ g/ml BSA

15 U mung bean nuclease

Incubate 30 min at 37°C. Stop reaction by adding 1 μ l of 0.5 M EDTA. The volume of reaction, amount of DNA, units of enzyme, temperature, and time of reaction will vary depending on the application.

Applications

Mung bean nuclease is used for most of the same purposes as described for S1 nuclease. However, mung bean nuclease is more precise than S1 nuclease for cleaving immediately adjacent to the last hybridized base pair without removing any of the base-paired nucleotides. For this reason, mung bean nuclease will typically generate single bands in transcript mapping experiments (UNIT 4.6), whereas S1 often results in multiple bands. In addition, mung bean nuclease is preferred over S1 for precisely deleting overhanging bases that result from restriction endonuclease cleavage. Unlike S1 nuclease, mung bean nuclease will not cleave the strand opposite a nick in duplex DNA. The disadvantages of mung bean nuclease are that it is more expensive than S1, and that its activity is more sensitive to reaction conditions.

ENZYME

MICROCOCCAL NUCLEASE

Micrococcal nuclease from *Staphylococcus aureus* is a relatively nonspecific nuclease that cleaves single- and double-stranded DNA and RNA to oligo- and mononucleotides with 3' phosphates (Alexander et al., 1961). The enzyme is more active on single-stranded nucleic acids. Cleavage of DNA or RNA occurs preferentially at AT- or AU-rich regions although all sequences are ultimately cleavable. The enzyme is strictly dependent on calcium for activity and hence can be inactivated by Ca⁺⁺-specific chelating agents such as EGTA.

One unit is defined as the amount of enzyme that hydrolyzes 1 μ mol of acid-soluble oligonucleotides from native DNA per minute at 37°C, pH 8.8.

Reaction Conditions

Typical micrococcal nuclease digestions are performed in solutions containing 10 mM Tris-Cl, pH 8.0, and 1 mM CaCl₂. However, the enzyme is somewhat more active at higher pH. The volume of reaction, amount and source of nucleic acid, units of enzyme, temperature, ionic strength, and time of reaction will vary greatly depending on the application. The reactions can be stopped by EDTA or the Ca⁺⁺-specific chelator EGTA.

Applications

1. Studies of chromatin structure (Kornberg, 1977).
2. Removing nucleic acid from crude cell-free extracts without destroying enzyme activities. Digestion of nucleic acid is performed under mild buffer and ionic conditions in the presence of CaCl₂. When digestion is complete, micrococcal nuclease is inactivated with EGTA (Pelham and Jackson, 1976). Such treated extracts are generally active for all processes except those requiring calcium.

Endonucleases

3.12.4

DEOXYRIBONUCLEASE I (DNase I)

DNase I from bovine pancreas is an endonuclease that degrades double-stranded DNA to produce 3'-hydroxyl oligonucleotides (Moore, 1981; Fig. 3.12.3). The enzyme requires divalent cations; the specificity of the reaction differs depending upon which divalent cation is present (Campbell and Jackson, 1980). In the presence of Mg^{++} , DNase I produces nicks in duplex DNA, while in the presence of Mn^{++} the enzyme produces double-stranded breaks in the DNA.

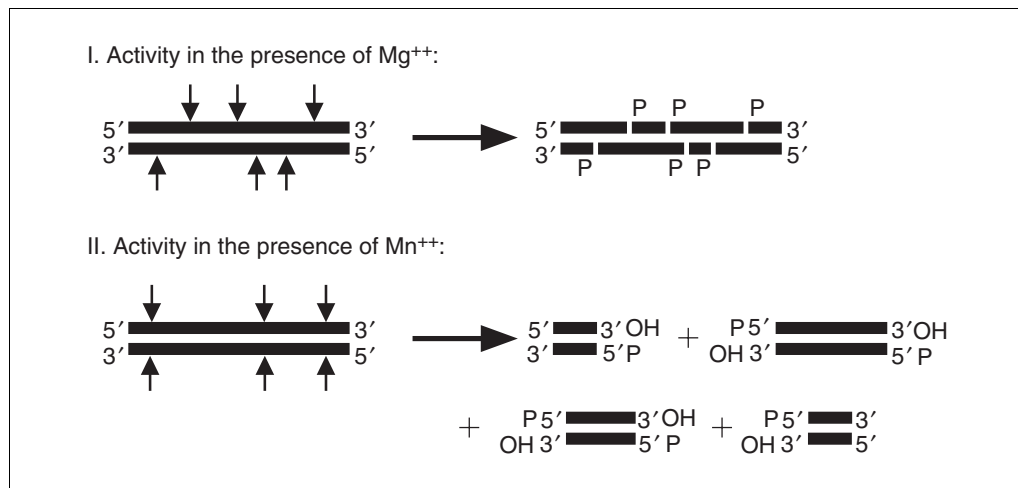


Figure 3.12.3 DNase I activities.

Reaction Conditions

For 100- μ l reaction:

50 mM Tris-Cl, pH 7.5

10 mM $MgCl_2$ (for single-strand nicks; to create double-strand breaks, replace with 10 mM $MnCl_2$)

2 μ g DNA

50 μ g/ml BSA

1 μ l DNase I (the concentration depends on the application)

Incubate at $37^\circ C$ for 1 to 30 min, depending upon the degree of digestion desired. Stop the reaction by adding 5 μ l of 0.5 M EDTA. For nick translations, the DNase I reaction is carried out simultaneously with the DNA polymerase I reaction (UNIT 3.5).

Preparation and Storage of DNase I Solutions

DNase I is usually purchased as a lyophilized powder at a concentration of 2000 to 3000 U/mg protein. For almost all applications, it is useful to make up a solution containing DNase I that can be stored for long periods of time without loss of enzyme activity. One method is to dissolve 1 mg DNase I in 1 ml of a 50% (wt/vol) solution of glycerol containing 20 mM Tris-Cl, pH 7.5, plus 1 mM $MgCl_2$; this solution can be stored in liquid form at $-20^\circ C$. Alternatively, 1 mg DNase I can be dissolved in 1 ml of a solution containing 20 mM Tris-Cl, pH 7.5, plus 1 mM $MgCl_2$, then aliquotted into small microcentrifuge tubes (100 10- μ l aliquots are convenient). The aliquots should be quick-frozen on dry ice, then stored at $-80^\circ C$. For use, an aliquot of DNase I is thawed on ice and an appropriate amount is added to the reaction mixture (for nick translations, the enzyme must first be diluted; see UNIT 3.5). Aliquots are not refrozen; they are simply discarded after use. DNase I solutions prepared by either method are stable for at least 1 year. It is important that DNase I be dissolved without vortexing (to minimize denaturation) and at a concentration of at least 1 mg/ml (to improve long-term stability).

RNase-Free DNase I

For many purposes, DNase I should be free of RNase. High-grade commercial preparations such as Worthington grade DPRF can be satisfactory. Alternatively, DNase I can be dissolved at 1 mg/ml in 0.1 M iodoacetic acid plus 0.15 M sodium acetate at a final pH of 5.3. The solution is then heated 40 min at 55°C and cooled. Finally, 1 M CaCl₂ is added to the solution to 5 mM. This homemade preparation of RNase-free DNase I should be stored frozen in small aliquots.

Applications

1. Nick translation (*UNIT 3.5*). Introduces nicks in duplex DNA which then serve as primer sites to initiate DNA synthesis by *E. coli* DNA polymerase I.
2. Cloning random DNA fragments by catalyzing double-stranded cleavage of DNA in the presence of Mn⁺⁺ (Anderson, 1981).

LITERATURE CITED

- Alexander, M., Heppel, L.A., and Hurwitz, J. 1961. The purification and properties of micrococcal nuclease. *J. Biol. Chem.* 236:3014-3019.
- Anderson, S. 1981. Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucl. Acids Res.* 9:3015-3027.
- Campbell, V.W. and Jackson, D.A. 1980. The effect of divalent cations on the mode of action of DNase I. The initial reaction products produced from covalently closed circular DNA. *J. Biol. Chem.* 255:3726-3735.
- Gray, H.B., Ostrander, D.A., Hodnett, J.L., Legerski, R.J., and Robberson, D.L. 1975. Extracellular nucleases of *Pseudomonas Bal 31*. I. Characterization of single strand-specific deoxyriboendonuclease and double-strand deoxyriboexonuclease activities. *Nucl. Acids Res.* 2:1459-1492.
- Gray, H.B., Winston, T.P., Hodnett, J.L., Legerski, R.J., Nees, D.W., Wei, C.-F., and Robberson, D.L. 1981. The extracellular nuclease from *Alteromonas espejiana*: An enzyme highly specific for nonduplex structure in nominally duplex DNAs. In *Gene Amplification and Analysis, Vol. 2: Structural analysis of nucleic acids* (J.G. Chirikjian and T.S. Papas, eds.) pp. 169-203. Elsevier/North Holland, New York, Amsterdam, Oxford.
- Kornberg, R.D. 1977. Structure of chromatin. *Ann. Rev. Biochem.* 46:931-954.
- Kroeker, W.D., Kowalski, D., and Laskowski, M. 1976. Mung bean nuclease I. Terminally directed hydrolysis of native DNA. *Biochemistry* 15:4463-4467.
- Lau, P.P. and Gray, H.B. 1979. Extracellular nucleases of *Alteromonas espejiana Bal 31*. IV. The single strand-specific deoxyriboendonuclease activity as a probe for regions of altered secondary structure in negatively and positively supercoiled closed circular DNA. *Nucl. Acids Res.* 6:331-357.
- Legerski, R.J., Gray, H.B. and Robberson, D.L. 1977. A sensitive endonuclease probe for lesions in DNA helix structure produced by carcinogenic or mutagenic agents *J. Biol. Chem.* 252:8740-8746.
- Legerski, R.J., Hodnett, J.L., and Gray, H.B. 1978. Extracellular nucleases of *Pseudomonas Bal 31*. III. Use of the double-strand deoxyriboexonuclease activity as a basis of a convenient method for the mapping of fragments of DNA produced by cleavage with restriction enzymes. *Nucl. Acids Res.* 5:1445-1464.
- Moore, S. 1981. Pancreatic DNase. In *The Enzymes, Vol. 14A* (P.D. Boyer, ed.) pp. 281-298. Academic Press, San Diego.
- Pelham, H.R.B. and Jackson, R.J. 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur. J. Biochem.* 67:247-256.
- Vogt, V.M. 1980. Purification and properties of S1 nuclease from *Aspergillus*. *Methods Enzymol.* 65:248-254.
- Wei, C.-F., Alianell, G.A., Bencen, G.H., and Gray, H.B. 1983. Isolation and comparison of two molecular species of the *Bal 31* nuclease from *Alteromonas espejiana* with distinct kinetic properties. *J. Biol. Chem.* 258:13506-13512.

Contributed by Stanley Tabor and Kevin Struhl
Harvard Medical School
Boston, Massachusetts

Ribonucleases (RNases) with different sequence specificities are used for a variety of analytical purposes, including RNA sequencing, mapping, and quantitation. One very common application for RNase A (see below) is hydrolyzing RNA that contaminates DNA preparations. Two other commonly used RNases, RNase H and RNase T1, are also described below.

Many commercially available RNases are sequence-specific endoribonucleases. This property has been used for enzymatic sequencing of RNA. For example, a combination of three different RNases and *Staphylococcus aureus* nuclease can be used in RNA sequence determination (see Table 3.13.1).

Table 3.13.1 RNases Used in RNA Sequencing^a

RNase	Sequence specificity
T1	Gp↓N
U2	Ap↓N
CL3	C(A/G)p↓N
<i>S. aureus</i> nuclease ^b	Np↓A/U

^aAdapted from Boehringer Mannheim.

^b*Staphylococcus aureus* nuclease cleaves both RNA and DNA.

RIBONUCLEASE A

Ribonuclease A (RNase A) from bovine pancreas is an endoribonuclease that specifically hydrolyzes RNA after C and U residues (Richards and Wyckoff, 1971). Cleavage occurs between the 3'-phosphate group of a pyrimidine ribonucleotide and the 5'-hydroxyl of the adjacent nucleotide. The reaction generates a 2':3' cyclic phosphate which then is hydrolyzed to the corresponding 3'-nucleoside phosphates. RNase A activity can be inhibited specifically by an RNase inhibitor (e.g., RNasin from Promega), a protein isolated from human placenta.

Reaction Conditions

RNase A is active under an extraordinarily wide range of reaction conditions, and it is extremely difficult to inactivate. At low salt concentrations (0 to 100 mM NaCl), RNase A cleaves single-stranded and double-stranded RNA as well the RNA strand in RNA:DNA duplexes. However, at NaCl concentrations of 0.3 M or above, RNase A becomes specific for cleavage of single-stranded RNA. Removal of RNase A from a reaction solution generally requires treatment with proteinase K followed by multiple phenol extractions and ethanol precipitation.

DNase-Free RNase A

To prepare RNase A free of DNase, dissolve RNase A in TE buffer at 1 mg/ml, and boil 10 to 30 min. Store aliquots at -20°C to prevent microbial growth.

Applications

1. Mapping and quantitating RNA species using the ribonuclease protection assay (UNIT 4.7). It is used in conjunction with RNase T1.
2. Hydrolyzing RNA that contaminates DNA preparations (UNITS 1.6 & 1.7).
3. RNA sequencing.

ENZYME

4. Blunt-ending double-stranded cDNA (*UNIT 5.5*). It is used in conjunction with RNase H.

ENZYME

RIBONUCLEASE H

Ribonuclease H (RNase H) from *E. coli* is an endoribonuclease that specifically hydrolyzes the phosphodiester bonds of RNA in RNA:DNA duplexes to generate products with 3' hydroxyl and 5' phosphate ends (Berkower et al., 1973). It will not degrade single-stranded or double-stranded DNA or RNA. RNase H cleavage can be directed to specific sites by hybridizing short deoxyoligonucleotides to the RNA (Donis-Keller, 1979).

One unit is defined as the amount of enzyme that produces 1 nmol of acid-soluble ribonucleotides from poly(A)-poly(dT) in 20 min at 37°C.

Reaction Conditions

For 100- μ l reaction:

- 20 mM HEPES·KOH, pH 8.0
- 50 mM KCl
- 4 mM MgCl₂
- 1 mM DTT
- 2 μ g RNA:DNA duplex
- 50 μ g/ml BSA
- 1 U ribonuclease H

Incubate 20 min at 37°C. Stop reaction by adding 1 μ l of 0.5 M EDTA. The volume of reaction, amount of DNA, units of enzyme, temperature, time, and method of stopping the reaction will vary depending on the application.

Applications

1. Facilitating the synthesis of double-stranded cDNA by removing the mRNA strand of the RNA:DNA duplex produced during first strand synthesis of cDNA (*UNIT 5.5*). It is also used later in the protocol to degrade residual RNA in conjunction with RNase A.
2. Creating specific cleavages in RNA molecules by using synthetic deoxyoligonucleotides to create local regions of RNA:DNA duplexes (Donis-Keller, 1979).

ENZYME

RIBONUCLEASE T1

Ribonuclease T1 (RNase T1) from *Aspergillus oryzae* is an endoribonuclease that specifically hydrolyzes RNA after G residues (Uchida and Egami, 1971). Cleavage occurs between the 3'-phosphate group of a guanine ribonucleotide and the 5'-hydroxyl of the adjacent nucleotide. The reaction generates a 2':3' cyclic phosphate which then is hydrolyzed to the corresponding 3'-nucleoside phosphates.

Reaction Conditions

The enzyme is active under a wide range of reaction conditions, and it is difficult to inactivate. At low salt concentrations (0 to 100 mM NaCl), it cleaves single-stranded and double-stranded RNA as well as the RNA strand in RNA:DNA duplexes. However, at NaCl concentrations of 0.3 M or above, it becomes specific for cleavage of single-stranded RNA. Removal of RNase T1 from a reaction solution generally requires treatment with proteinase K followed by multiple phenol extractions and ethanol precipitation.

Ribonucleases

3.13.2

Applications

1. Mapping and quantitating RNA species using the ribonuclease protection assay (*UNITS 4.7 & 9.8*). It is used in conjunction with RNase A.
2. RNA sequencing.
3. Determining the level of RNA transcripts synthesized in vitro from DNA templates containing a “G-less cassette.” Because accurately initiated transcripts contain no G residues, treatment of the reaction products with RNase T1 preferentially destroys all nonspecific transcripts (Sawadogo and Roeder, 1985).

LITERATURE CITED

- Berkower, I., Leis, J., and Hurwitz, J. 1973. Isolation and characterization of an endonuclease from *Escherichia coli* specific for ribonucleic acid in ribonucleic acid-deoxyribonucleic acid hybrid structures. *J. Biol. Chem.* 248:5914-5921.
- Boehringer Mannheim Biochemicals. Biochemicals for Molecular Biology (catalog). Indianapolis.
- Donis-Keller, H. 1979. Site specific enzymatic cleavage of RNA. *Nucl. Acids Res.* 7:179-192.
- Richards, F.M. and Wyckoff, H.W. 1971. Bovine pancreatic ribonuclease. *In The Enzymes*, Vol. IV (P.D. Boyer, ed.) pp. 647-806. Academic Press, San Diego.
- Sawadogo, M. and Roeder, R.G. 1985. Factors involved in specific transcription by human RNA polymerase II: Analysis by a rapid and quantitative in vitro assay. *Proc. Natl. Acad. Sci. U.S.A.* 82:4394-4398.
- Uchida, T. and Egami, F. 1971. Microbial ribonucleases with special reference to RNases T1, T2, N1, and U2. *In The Enzymes*, Vol. IV (P.D. Boyer, ed.) pp. 205-250. Academic Press, San Diego.

Contributed by Kevin Struhl
Harvard Medical School
Boston, Massachusetts

DNA Ligases

DNA ligases catalyze the formation of phosphodiester bonds between juxtaposed 5' phosphate and a 3'-hydroxyl terminus in duplex DNA. This activity can repair single-stranded nicks in duplex DNA (Fig. 3.14.1) and join duplex DNA restriction fragments having either blunt ends (Fig. 3.14.2) or homologous cohesive ends (Fig. 3.14.3). Two ligases are used for nucleic acid research—*E. coli* ligase and T4 ligase. These enzymes differ in two important properties. One is the source of energy: T4 ligase uses ATP, while *E. coli* ligase uses NAD. Another important difference is their ability to ligate blunt ends; under normal reaction conditions, only T4 DNA ligase will ligate blunt ends.

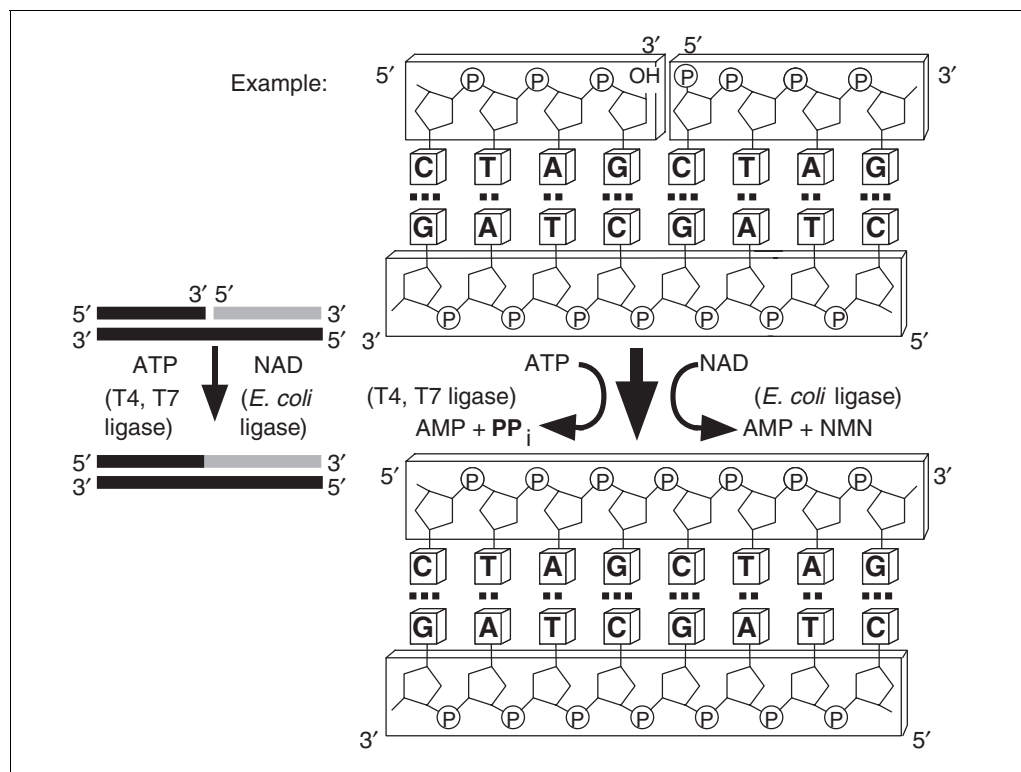


Figure 3.14.1 DNA ligase activity at a nick.

ENZYME

T4 DNA LIGASE

T4 DNA ligase, the product of gene 30 of phage T4, was originally purified from phage-infected cells of *E. coli*. The phage T4 gene 30 has been cloned, and the enzyme is now prepared from overproducing strains. Using ATP as a cofactor, T4 DNA ligase catalyzes the repair of single-stranded nicks in duplex DNA and joins duplex DNA restriction fragments having either blunt or cohesive ends. It is the only ligase that efficiently joins blunt-end termini under normal reaction conditions. It appears that T4 DNA ligase activity may be stimulated by T4 RNA ligase (UNIT 3.15). See UNIT 3.16 for a detailed ligation protocol.

Reaction Conditions

For 50- μ l reaction:

40 mM Tris·Cl, pH 7.5

10 mM MgCl₂

10 mM DTT

1 μ g DNA

0.5 mM ATP

50 μ g/ml BSA

1 “Weiss” U T4 DNA ligase

Incubate at 12° to 30°C for 1 to 16 hr. Stop reaction by adding 2 μ l of 0.5 M EDTA or by heating to 75°C for 10 min. The volume of reaction, concentration of DNA, and the temperature and time of the reaction will vary, depending upon the individual application. One Weiss unit is equivalent to 60 cohesive-end units.

Ligation of cohesive ends is usually carried out at 12° to 15°C to maintain a good balance between annealing of the ends and activity of the enzyme. Higher temperatures make it difficult for the ends to anneal, whereas lower temperatures diminish ligase activity. Blunt-end ligations are typically performed at room temperature since annealing is not a factor (the enzyme is not particularly stable above 30°C). Blunt-end ligations require about 10 to 100 times more enzyme than cohesive-end ligations to achieve an equal efficiency. T4 DNA ligase is not inhibited by tRNA, but it is strongly inhibited by NaCl concentrations >150 mM. Macromolecular exclusion molecules (e.g., PEG 8000) have been shown to greatly increase the rate of both cohesive-end and blunt-end joining by T4 DNA ligase (Pfeiffer and Zimmerman, 1983). An inherent consequence of macromolecular crowding is that all ligations are intermolecular; thus, this technique is not suitable for the ligation and circularization of inserts and vectors that are required for most cloning experiments.

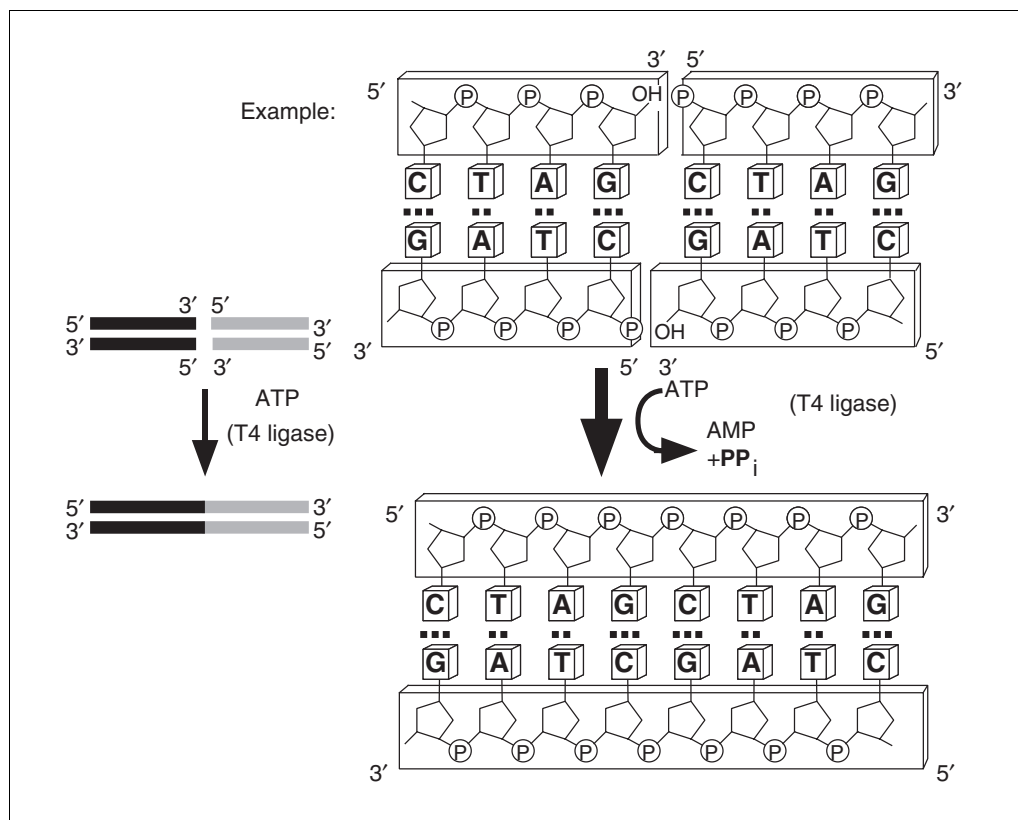


Figure 3.14.2 DNA ligase activity at blunt ends.

(e.g., DTT) in the reaction. PEG 8000 greatly increases the rate of cohesive end joining by *E. coli* DNA ligase (Harrison and Zimmerman, 1983). Interestingly, the presence of macromolecular exclusion molecules also enables *E. coli* DNA ligase to efficiently join blunt-end termini, a reaction it is unable to carry out in their absence.

Modrich-Lehman units measure the ability to form poly d(A-T) circles. One Modrich-Lehman unit is equivalent to 6 Weiss units (Modrich and Lehman, 1975).

Applications

E. coli DNA ligase can be used as an alternative to T4 DNA ligase when blunt-end ligations are not required. Transformation using DNA ligated with *E. coli* DNA ligase has a lower background that results from aberrant ligations compared with T4 DNA ligase, since T4 DNA ligase has a much lower specificity for the structure of the termini.

LITERATURE CITED

- Harrison, B. and Zimmerman, S.B. 1983. Macromolecular crowding allows blunt-end ligation by DNA ligases from rat liver or *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 80:5852-5856.
- Modrich, T. and Lehman, I.R. 1975. Enzymatic joining of polynucleotides. *J. Biol. Chem.* 245:3626-3631.
- Pfeiffer, B.H. and Zimmerman, S.B. 1983. Polymer-stimulated ligation: Enhanced blunt- or cohesive-end ligation of DNA or deoxyribonucleotides by T4 DNA ligase in polymer solutions. *Nucl. Acids Res.* 11: 7853-7871.

KEY REFERENCE

- Engler, M.J. and Richardson, D.C. 1982. DNA ligases. *In* The Enzymes, Vol. 15B (P.D. Boyer, ed.) pp. 3-30. Academic Press, San Diego.

Contributed by Stanley Tabor
Harvard Medical School
Boston, Massachusetts

RNA Ligases

ENZYME

T4 RNA LIGASE (Uhlenbeck and Gumport, 1982)

T4 RNA ligase, the product of the phage gene 63, is purified from phage-infected cells. It catalyzes the ATP-dependent covalent joining of single-stranded 5'-phosphoryl termini of DNA or RNA to single-stranded 3'-hydroxyl termini of DNA or RNA (Fig. 3.15.1).

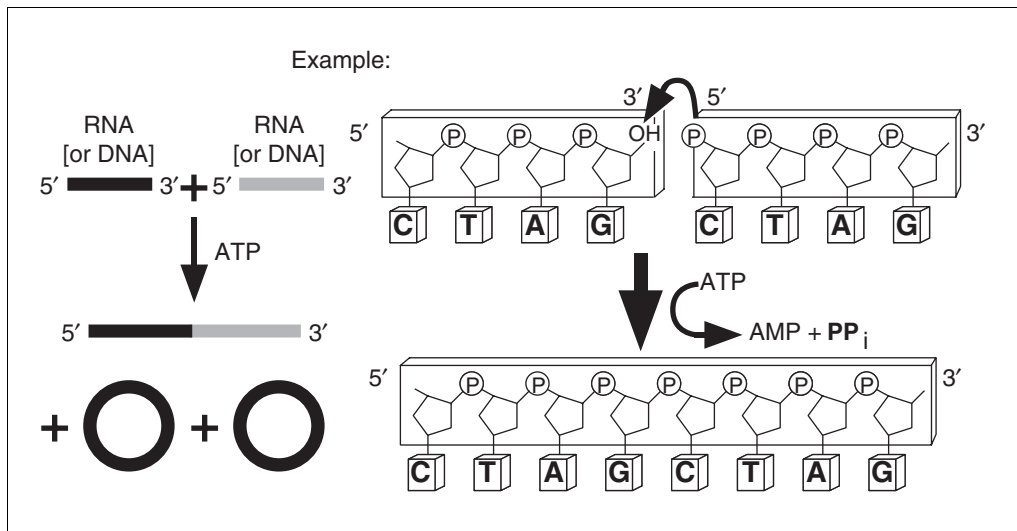


Figure 3.15.1 RNA ligase activity.

Reaction Conditions

For 50- μ l reaction:

- 50 mM HEPES, pH 8.3
- 10 mM MgCl₂
- 5 mM DTT
- 2 μ g single-stranded DNA or RNA
- 2 mM ATP
- 50 μ g/ml BSA
- 1 U T4 RNA ligase

Incubate at 17°C for 10 hr. Stop the reaction by adding 2 μ l of 0.5 M EDTA.

Applications

1. Radioactive labeling of 3' termini of RNA (Uhlenbeck and Gumport, 1982). The reaction contains the RNA molecule as the acceptor, a [5'-³²P]nucleoside 3', 5'-bis(phosphate) (e.g., 5'-³²P-Cp) as the donor, and ATP as the energy source. The product formed by this reaction is the RNA molecule containing one additional nucleotide at the 3' terminus and a ³²P-labeled phosphate in the last internucleotide linkage.
2. Circularizing deoxy- and ribo-oligonucleotides (Brennan et al., 1983).
3. Ligating oligomers for oligonucleotide synthesis (Romaniuk and Uhlenbeck, 1983). One such use is to synthesize oligomers that contain internally labeled oligomers at specific residues.
4. Stimulating the blunt-end ligation activity of T4 DNA ligase (Sugino et al., 1977).

LITERATURE CITED

- Brennan, C.A., Manthey, A.E., and Gumport, R.I. 1983. Using T4 RNA ligase with DNA substrates. *Meth. Enzymol.* 100:38-52.
- Romaniuk, P.J. and Uhlenbeck, O.C. 1983. Joining of RNA molecules with RNA ligase. *Meth. Enzymol.* 100:52-59.
- Sugino, A., Goodman, H.M., Heyneker, H.L., Shine, J., Boyer, H.W., and Cozzarelli, N.R. 1977. Interaction of bacteriophage T4 RNA and DNA ligases in joining of duplex DNA at base-paired ends. *J. Biol. Chem.* 252:3987-3994.

KEY REFERENCE

- Uhlenbeck, O.C. and Gumport, R.I. 1982. T4 RNA ligase. *In* The Enzymes, Vol. 15B (P.D. Boyer, ed.) pp. 31-60. Academic Press, NY.

Contributed by Stanley Tabor
Harvard Medical School
Boston, Massachusetts

CONSTRUCTION OF HYBRID DNA MOLECULES

Given the large number of restriction endonucleases and other enzymes for manipulating DNA, it is possible to create any DNA molecule of interest. As it is impossible to cover all possible ways of manipulating DNA molecules, this section will discuss the general principles involved along with several specific techniques that are frequently used. Chapter 8 describes methods for creating point mutations in DNA.

UNIT 3.16

BASIC PROTOCOL

Subcloning of DNA Fragments

In order to construct new DNA molecules, the starting DNAs are treated with appropriate restriction endonucleases and other enzymes if necessary. The individual components of the desired DNA molecule are purified by agarose or polyacrylamide gel electrophoresis, combined, and treated with DNA ligase. The products of the ligation mixture (along with control mixtures) are introduced into competent *E. coli* cells, and transformants are identified by an appropriate genetic selection. DNA is prepared from the colonies or plaques and subjected to restriction endonuclease mapping in order to determine if the desired DNA molecule was created. All cloning experiments follow the steps outlined below.

Materials

Calf intestine phosphatase (CIP) and buffer (optional; *UNIT 3.10*)
 dNTP mix (0.5 mM each; *UNIT 3.4*)
 Klenow fragment of *E. coli* DNA polymerase I or T4 DNA polymerase (optional; *UNIT 3.5*)
 Oligonucleotide linkers (optional)
 10 mM ATP
 0.2 mM dithiothreitol (DTT)
 T4 DNA ligase (measured in cohesive-end units; *UNIT 3.14*)
 2× T4 DNA ligase buffer

Additional reagents and equipment for restriction endonuclease digestions (*UNIT 3.1*), transformation of *E. coli* cells (*UNIT 1.8*), DNA minipreps (*UNIT 1.6*), and agarose or polyacrylamide gel electrophoresis (*UNIT 2.5* or *2.7*).

1. In a 20- μ l reaction mixture, cleave the individual DNA components with appropriate restriction endonuclease. After the reaction is complete, inactivate the enzymes by heating 15 min to 75°C. If no further enzymatic treatments are necessary, proceed to step 6.

Reaction mixtures can be done in any volume; 20 μ l is convenient for gel electrophoresis (step 6). Many of the subsequent enzymatic manipulations (steps 2 to 5) can be carried out sequentially without further buffer changes.

2. If the 5' phosphates of one of the DNAs are to be removed (see Example 3.16.1), add 2 μ l of 10× CIP buffer and 1 U CIP; incubate 30 to 60 min at 37°C as described in *UNIT 3.10*. After the reaction is complete, inactivate CIP by heating 15 min to 75°C. If no further enzymatic treatments are necessary, proceed to step 6.
3. If one or both ends generated by a restriction endonuclease must be converted to blunt ends (see Example 3.16.4), add 1 μ l of a solution containing all 4 dNTPs (0.5 mM each) and an appropriate amount of the Klenow fragment of *E. coli* DNA polymerase I or T4 DNA polymerase; carry out the filling-in or trimming reaction as described in *UNIT 3.5*. After the reaction is complete, inactivate the enzymes by

heating 15 min to 75°C. If oligonucleotide linkers are to be added (see Example 3.16.6), proceed to step 4. If a DNA fragment containing only one blunt end is desired, cleave the reaction products with an appropriate restriction endonuclease. If no further enzymatic treatments are necessary, proceed to step 6.

4. Add 0.1 to 1.0 µg of an appropriate oligonucleotide linker, 1 µl of 10 mM ATP, 1 µl of 0.2 M DTT, and 20 to 100 cohesive-end units of T4 DNA ligase; incubate overnight at 15°C. Inactivate the ligase by heating 15 min to 75°C.

In general, the oligonucleotide linkers are phosphorylated prior to ligation (UNIT 3.10). However, the linkers do not have to be phosphorylated, in which case the products will contain only one linker on each end.

5. Cleave the products from step 4 with the restriction endonuclease recognizing the oligonucleotide linker, adjusting the buffer conditions if necessary. If only one of the two ends is to contain a linker, cleave the products with an additional restriction enzyme (see Examples 3.16.2 and 3.16.6).
6. Isolate the desired DNA segments by gel electrophoresis, or by other methods if appropriate.

Electrophoresis in agarose gels is the most common method. Purification is not essential for many cloning experiments, but it is usually very helpful. See critical parameters.

7. Using longwave UV light for visualization of the DNA, cut out the desired band(s) and purify the DNA away from the gel material using the procedures described in UNIT 2.6.

It is critical to use longwave UV light sources to prevent damage to the DNA. If low gelling/melting temperature agarose is used, ligation reactions can usually be performed directly in the gel slice (see alternate protocol).

8. Set up the following ligation reaction:

9 µl component DNAs (0.1 to 5 µg)
10 µl 2× ligase buffer
1 µl 10 mM ATP
20 to 500 U (cohesive end) T4 DNA ligase

Incubate 1 to 24 hr at 15°C.

Simple ligations with two fragments having 4-bp 3' or 5' overhanging ends require much less ligase than more complex ligations or blunt-end ligations. The quality of the DNA will also affect the amount of ligase needed.

Controlled cloning experiments usually involve several different ligation reactions. For simplicity, appropriate DNAs should be added to each tube (adding water if necessary) so that the volume is 9 µl. Immediately prior to use, pre-mix on ice the remaining ingredients (2× buffer, ATP, enzyme) in sufficient amounts for all the reactions. To start the ligation reactions, add 11-µl aliquots of the pre-mix to each tube containing DNA. The amounts and ratios of the component DNAs and the controls for these ligation reactions are discussed below.

9. Introduce 1 to 10 µl of the ligated products into competent *E. coli* cells and select for transformants by virtue of the genetic marker present on the vector.
10. From individual *E. coli* transformants, purify plasmid or phage DNAs by miniprep procedures and determine their structures by restriction mapping.

If the frequency of transformants containing the desired hybrid DNA molecule is too low, initial screening by filter hybridization to immobilized colonies (UNIT 6.2) or plaques (UNIT 6.1) or by genetic techniques is necessary.

LIGATION OF DNA FRAGMENTS IN GEL SLICES

This alternate protocol saves considerable time in comparison to the basic protocol because it eliminates purification of the DNA fragments away from the gel matrix (Struhl, 1985). It is suitable for most simple cloning experiments and is particularly valuable for carrying out a set of hybrid constructions involving a variety of different DNA fragments. However, as the cloning efficiency may be reduced, this method should be employed only when it is desired to make one (or relatively few) specific molecules.

Additional Materials

Low gelling/melting temperature agarose (SeaPlaque, FMC Marine Colloids)
TAE buffer (*APPENDIX 2*)

1. Treat the starting DNAs with appropriate restriction endonucleases and other enzymes, as described in steps 1 to 6 of the basic protocol.
2. Subject the treated DNAs to electrophoresis in low gelling/melting temperature agarose using TAE buffer.

It is critical that the agarose be of high quality. SeaPlaque is a good choice. The agarose concentration should be kept as low as possible (0.7% is suitable for most applications).

3. Cut out the desired band(s) in the smallest possible volume (20 to 50 μ l) using a clean razor blade, and place the gel slice in a microcentrifuge tube.
4. Melt the gel slices containing DNA at 70°C for at least 10 min.

This temperature is hot enough to melt the agarose without denaturing the DNA.

5. In separate tubes for each ligation reaction, combine the gel slices containing appropriate DNAs (and water if necessary) for a total volume of 9 μ l. Place the tubes at 37°C for a few minutes.

The gel slices should remain molten at 37°C.

6. To each tube containing DNA, add 11 μ l of an ice-cold mixture containing 2 \times buffer, ATP, and T4 DNA ligase. Mix immediately by flicking the tube and place on ice. Then incubate the reaction mixtures 1 to 48 hr at 15°C.

DNA fragments can still be ligated even though the reaction mixture has resolidified into a gel.

7. After the ligation reaction is complete, remelt the gel slices 5 to 10 min at 73°C and add 5 μ l of the ligated products to 200 μ l of competent *E. coli* cells (*UNIT 1.8*).

The remelted gel must be diluted at least 30-fold so that it does not resolidify when the cells are placed on ice.

8. Carry out steps 9 and 10 of the basic protocol.

*Ligation in gel slices is suitable for most cloning applications. If the method does not work, DNA fragments can be purified from the gel by various techniques discussed in *UNIT 2.6*. Ligations are then carried out by the basic protocol.*

REAGENTS AND SOLUTIONS

2 \times T4 DNA ligase buffer

100 mM Tris·Cl, pH 7.5
20 mM MgCl₂
20 mM DTT

COMMENTARY

Background Information

The essence of recombinant DNA technology is the joining of two or more separate segments of DNA to generate a single DNA molecule that is capable of autonomous replication in a given host. The simplest constructions of hybrid DNA molecules involve the cloning of insert sequences into plasmid or bacteriophage cloning vectors. The insert sequences can derive from essentially any organism, and they may be isolated directly from the genome (UNITS 5.3 and 5.4), from mRNA (UNIT 5.5), or from previously cloned DNA segments (in which case, the procedure is termed subcloning). Alternatively, insert DNAs can be created directly by DNA synthesis.

Critical Parameters

Many different factors must be considered before embarking on any specific DNA manipulation. These include the number, type, and concentration of DNA fragments, the preparation and purification of DNA fragments, and the selection or screening systems that might be necessary for identifying the DNA molecule of interest. Another practical consideration is the number of hybrid DNA molecules that are desired for any particular experiment. For example, when generating genomic or cDNA libraries, it is critical to obtain the maximum number of recombinants. Strategies for producing large libraries are discussed elsewhere (UNITS 5.1 and 5.2).

On the other hand, when it is desired to make one (or relatively few) specific DNA molecule, it is important to maximize the frequency of generating the “correct” molecule rather than the absolute cloning efficiency. In other words, it is better to obtain a low number of colonies, mostly containing the desired DNA rather than a large number of colonies few of which contain the desired DNA. In this way, screening for the correct colonies or plaques is unnecessary, and preparation and analysis of DNA from *E. coli* transformants is minimized; these represent the major time-consuming steps in the overall cloning process.

Although it is obvious that the experimental conditions should be designed to favor the desired ligation events, the crucial factor for optimizing the frequency of correct molecules is to reduce the “background” of undesired events that will result in transformable DNA molecules. One major cause of background involves

“errors” in the preparation of the starting DNA components (e.g., incomplete digestion by restriction endonucleases). Such errors can be significantly reduced, but never entirely eliminated. The other major source of background is undesired ligation between molecules in the reaction mixture. This problem is addressed by considering all possible ways in which the starting components (and probable contaminants) can be joined to produce DNA molecules that can transform *E. coli* via the genetic selection being imposed. By appropriate “tricks,” the major “side reactions” can often be minimized; these are discussed below.

DNA preparation and the purification of specific fragments. The efficiency of obtaining the desired recombinant DNA molecule is greatly improved by purifying the individual components by electrophoresis in agarose gels. The extra effort involved is more than compensated by the elimination (in most cases) of the need to screen the recombinants and the resulting preparation and analysis of fewer DNAs. Equally importantly, gel purification makes it possible to use crude miniprep DNAs (UNIT 1.6) as the initial starting material instead of CsCl-purified DNAs, thus eliminating time, expense, and tedium.

Purified DNA fragments are relatively free of minor contaminants that can cause major background problems. For example, if a circular plasmid vector is cleaved at 99% efficiency with a single restriction enzyme, the uncleaved 1% of the molecules are fully infectious even though they cannot be visualized by ethidium bromide staining. This background, which is unacceptably high for many experiments, can be greatly reduced by gel purification of the linearized DNA. Gel purification is extremely useful in situations where only one of several fragments is needed or where fragments are generated by partial digestion. It is also useful for removal of small oligonucleotide linkers or other low-molecular-weight contaminants.

Electrophoretically separated DNA fragments should be isolated in the smallest possible volume of gel slice. Purity is more important than yield. DNA fragments can be purified away from the agarose (or acrylamide) by any of the procedures described in UNIT 2.6. It is important to use high-quality agarose such as SeaKem or SeaPlaque, since other preparations may contain inhibitors of DNA ligase that are hard to remove. DNA purified from acrylamide

gels does not usually contain such inhibitors. Finally, for most applications, ligation reactions can be performed directly in slices of low gelling/melting temperature agarose (Struhl, 1985; see alternate protocol on p. 3.16.3). By eliminating the need to purify the DNA away from the agarose, considerable time and effort is saved.

Intramolecular versus intermolecular ligation of DNA fragments. The absolute and relative concentrations of the input DNA fragments influence the frequencies of specific ligation events. The first consideration is whether the end of a particular DNA molecule will be joined to the other end of the same molecule (intramolecular ligation) or to a separate DNA molecule (intermolecular ligation). The second consideration is whether intermolecular events will be between molecules of the same type (vector–vector or insert–insert) or between molecules of a different type (vector–insert). Most hybrid constructions involve one or more intermolecular events followed by a final intramolecular event to generate a circular plasmid molecule.

Intramolecular events. Intramolecular events are a major source of background in many cloning experiments primarily because self-ligation of vector molecules produces undesired *E. coli* transformants. However, since intramolecular ligation can occur only if the ends of the molecule are cohesive or blunt, it can be almost completely eliminated by using molecules with heterologous ends. This is achieved by cleaving the vector with two restriction endonucleases that produce incompatible ends. For situations where the vector must be cleaved with a single endonuclease, intramolecular joining can be prevented by removing both 5' phosphates of each molecule with calf intestinal phosphatase. These dephosphorylated vector molecules can be ligated to phosphorylated insert DNAs even though at each end only 1 of the 2 strands will be covalently joined in vitro. *Escherichia coli* can repair the single-stranded nicks in vivo to generate the desired DNA molecules.

Intramolecular ligation of insert fragments does not generally affect the background of unwanted colonies. However, it strongly influences the frequency of obtaining the correct colonies because self-ligated insert DNA cannot be joined to the vector. Thus, in situations where the insert DNA fragments can cyclize, relatively high concentrations of DNA (20 to 100 µg/ml) should be used in order to favor

intermolecular ligation events. This is particularly true for short DNA fragments which are much more likely to cyclize than large fragments (for a theoretical discussion, see Dugaiczky et al., 1975).

Intermolecular events. Almost all cloning experiments using plasmid vectors require the joining of two or more separate DNA molecules followed by circularization of the product. Obviously, intermolecular ligation events become increasingly favored over self-joining as the total concentration of the relevant ends becomes higher. Thus, optimal cloning efficiency occurs at DNA concentrations that are high enough to permit sufficient intermolecular joining, but not so high as to reduce intramolecular ligation. For situations when the background is low, DNA concentrations between 1 and 50 µg/ml are acceptable.

If we exclude self-ligation events, the frequency of joining different DNA segments is directly related to the molarity of each component in terms of specific ends to be joined. For example, if equimolar amounts of vector (V) and insert (I) are present, 50% of the bimolecular products will be V–I, 25% will be V–V, and 25% will be I–I. However, because in most cases only molecules containing vector sequences will generate *E. coli* transformants, it is preferable to use higher molar concentrations of insert fragment(s) because this will favor the desired V–I products over the V–V background. This consideration becomes less important if V–V events are precluded by phosphatase treatment or if V–V events cannot generate viable molecules.

Amount of DNA. In general, it is not necessary to use very much DNA. For simple hybrid constructions, roughly 1 to 10% of the vector DNA molecules are converted to the recombinant DNA of interest. With normal transformation efficiencies, 10⁶ colonies per microgram of pure vector DNA, ligated products containing as little as 1 ng of treated vector DNA should generate 10 to 100 transformants. This is more than enough if the background is not a problem. In practical terms, if starting DNA fragments are easily visualized by conventional ethidium bromide staining, there is enough for most experiments. This is true even when performing ligation in gel slices (see alternate protocol) where only a fraction of the electrophoretically separated DNA is used.

More difficult constructions such as those involving blunt ends, three fragments, or *Bal* 31 nuclease usually require more DNA. How-

ever, even in these cases, there is sufficient DNA in a fraction of a gel slice such that the alternate protocol can be used. In general, 25 to 100 ng of each DNA component (in a standard 20- μ l reaction) is usually enough to obtain the desired recombinants. Of course, the cloning efficiency will be increased by using higher concentrations of the components. If this proves to be necessary, the DNA can be concentrated easily after purifying it away from the agarose (UNIT 2.6).

Selection, screening, and enrichment procedures. For some hybrid constructions, colonies containing the desired DNA molecule can be distinguished from background colonies by genetic methods. In such cases, techniques for background reduction become less important, which often means that time can be saved. As an extreme but relatively infrequent example, if one fragment contains the gene for ampicillin resistance and another fragment contains the gene for tetracycline resistance, the recombinant containing both fragments can be selected directly on plates containing both drugs; there is no need to purify the component DNAs prior to ligation.

Cloning vectors have been designed to select or screen for recombinants. Perhaps those most commonly used are the pUC or M13mp series of vectors which contain the α -complementation region of the *E. coli lacZ* gene and produce blue plaques or colonies on appropriate indicator plates (UNIT 1.4). Insertion of a DNA fragment into the *lacZ* region generally causes a white colony or plaque, which is easily seen among a background of thousands of blue transformants. Another common vector, pBR322, contains two selectable markers, ampicillin and tetracycline resistance. Insertion into one of these marker genes renders the colony sensitive to that drug; such colonies can be screened by replica plating (UNIT 1.5).

In many cases, the desired recombinant contains (or lacks) specific regions of DNA that are present in either or both of the starting molecules. Such recombinants can be identified by colony or plaque filter hybridization (UNITS 6.1 and 6.2) using an appropriate 32 P-labeled probe (UNIT 3.5). These screening procedures can identify the correct recombinant from a background of hundreds or thousands of transformants. However, they take several days to perform and are unnecessary if the background is acceptably low.

Finally, the proportion of correct colonies can be increased by appropriate enzymatic

treatment after the ligation reaction. For example, if the desired product lacks an *EcoRI* site, but undesired products contain such a site, the background can be reduced simply by *EcoRI* cleavage after the ligation reaction. Most restriction endonucleases are active in ligation buffer.

Controls. It is critical to perform appropriate control experiments simultaneously with the reaction of interest. In this way, it is often possible to determine the relative success of the hybrid construction. Such knowledge is particularly important because preparation and analysis of prospective recombinant DNAs is the most time-consuming part of the overall process. Obviously, it is worth minimizing the number of colonies to be examined from successful experiments as well as avoiding further analysis from cloning attempts that were clearly unsuccessful.

NOTE: *The basic controls for any experiment are to set up and analyze parallel ligation reactions, each lacking a single DNA component.* In order to keep the volume of the ligation reactions constant, the omitted DNA should be replaced by an equal volume of water. In a successful hybrid construction, the number of colonies obtained when all DNA components are present should be higher than any of the control reactions. If this is the case, many or most of the colonies should contain the molecule of interest.

Two additional controls should be carried out as part of the transformation procedure. First, an aliquot of competent cells should be incubated in the absence of DNA. The appearance of transformants is indicative of contamination in the cells or transformation buffers. Second, another aliquot of competent cells should be incubated with 1 ng of a control DNA (either plasmid or phage depending on the experiment) to measure the transformation efficiency; 1000 colonies or plaques should be achieved routinely.

Anticipated Results

For simple hybrid constructions, approximately 100 to 10,000 colonies or plaques should be generated, most of which should be the desired molecule. For more complex situations, the expected number of colonies should be between 1 and 1000; depending on the individual situation 5 to 80% of the molecules should contain the desired structure. For many cloning experiments, one correct colony is enough.

Time Considerations

The enzyme reactions can usually be performed in 1 to 4 hr (except in experiments involving ligation of oligonucleotide linkers). Gel electrophoresis and DNA fragment purification should take an additional 1 to 4 hr; shorter times are obtained by using minigels (UNIT 2.5) and by using the alternate protocol. Thus, in one day, it should be possible to go from uncleaved starting DNAs to the ligation reaction. On day 2, the ligated products can be introduced into *E. coli*, and the resulting transformants will be obtained on day 3. After colony or plaque purification, DNA from the transformants can be prepared and analyzed on day 4.

EXAMPLES

Example 3.16.1: Subcloning DNA Fragments with Homologous and Cohesive Ends

Consider the cloning of a 2-kb *EcoRI* fragment into a vector containing a single *EcoRI* site. The major source of background is uncleaved or self-ligated vector DNA. In most cases, self-ligation should be reduced by treating the *EcoRI* cleaved vector DNA with calf intestinal alkaline phosphatase (see Fig. 3.16.1 below and UNIT 3.10). Although phosphatase treatment does lower the absolute efficiency of obtaining the desired DNA molecule, the great

reduction in background is necessary to obtain a decent relative frequency of the desired molecule. The concentration of insert DNA(s) should be relatively high in order to facilitate ligation to the vector. It should also be noted that the insert DNA can contribute to background transformants by self-ligation, especially when it is difficult to purify away from “vector” fragments derived from the original plasmid.

Because all *EcoRI* ends are equivalent, the 2-kb insert can be cloned in either of the two possible orientations with respect to the vector sequences. These two outcomes can be distinguished only by restriction mapping using an enzyme that cleaves asymmetrically within the insert DNA. In addition, it is possible to insert multiple copies of the fragment into the vector. Generally, multiple insertion is infrequent unless the DNA concentrations are very high. When multiple insertions occur, the inserts are essentially always oriented in the same direction. Molecules containing tandem copies oriented in the opposite direction are incapable of replication in *E. coli* unless the insert fragment is very short (less than 30 bp).

For all the reasons mentioned above, cloning of fragments with homologous and cohesive ends is best avoided when possible. Directional cloning, discussed below, is preferable.

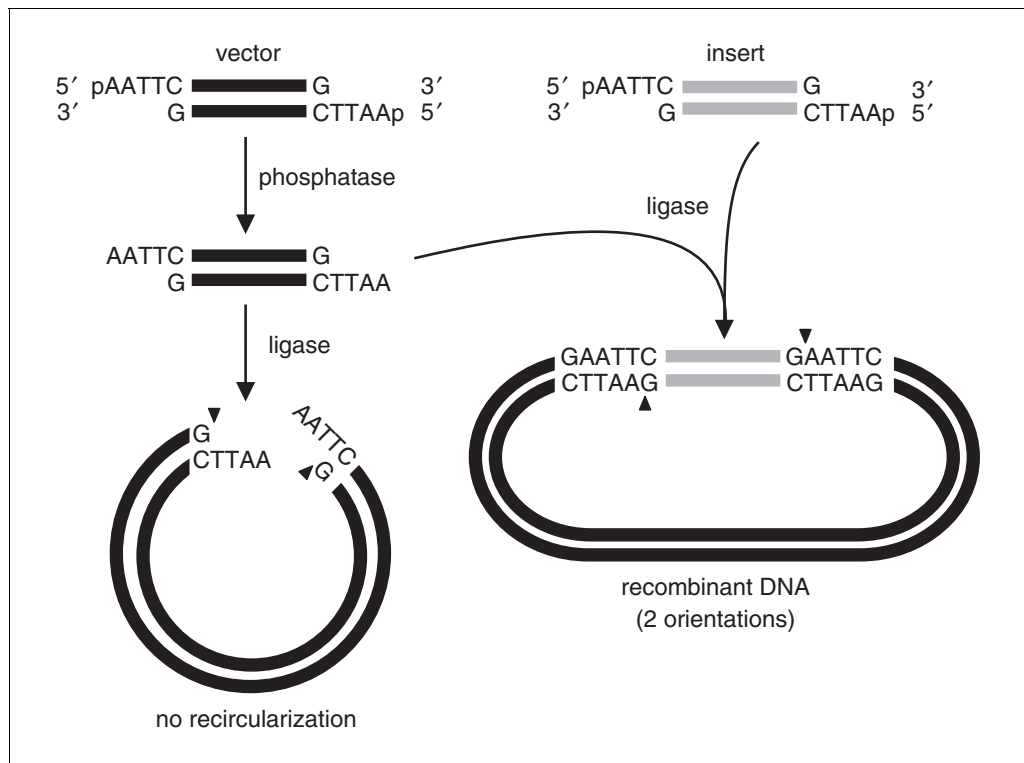


Figure 3.16.1 Treatment of cleaved vector with CIP in order to reduce background.

Example 3.16.2: Directional Cloning Using Fragments with Heterologous Ends

This is by far the most efficient method for cloning, and it should be used wherever possible. In the simplest example, consider the joining of two fragments, each produced by cleavage with enzymes A and B. Self-ligation of either fragment is theoretically eliminated, and ligation of two vector fragments produces a “perfect inverted repeat” molecule that cannot be stably maintained in *E. coli*. Thus, the background is very low, and if enzymes A and B produce 5′ or 3′ overhangs, the efficiency is very high. Moreover, almost all the transformants will contain recombinant molecules in which one copy of the insert DNA is oriented in a defined direction with respect to the vector. Molecules containing two copies of the insert cannot be formed, and those with three copies will be very rare.

The major problems associated with directional cloning come from errors in generating the DNA segments, particularly incomplete cleavage of the vector by one of the restriction endonucleases. Vector molecules that are singly cleaved can self-ligate and generate background. Ideally, molecules cleaved by both enzymes can be purified away from those that are incompletely cleaved by gel electrophoresis. However, this is not possible when the sites for cleavage by enzymes A and B are close together. Moreover, if the two sites are extremely close together, as is the case for the polylinkers in the pUC and M13mp vectors, it can be difficult to cleave the molecules with both enzymes.

The high efficiency of directional cloning makes it possible to create molecules composed of three or more segments. A typical three-piece ligation will consist of fragments generated by enzymes A + B, B + C, and A + C. The desired recombinant will occur less frequently than in the case for two-fragment ligation. However, the low background makes it possible to obtain the desired molecule at high enough frequency to permit individual analysis of colonies.

Example 3.16.3: Blunt-End Ligation

Although T4 DNA ligase can join fragments with blunt ends, ligation efficiency is greatly reduced as compared to fragments with cohesive ends. Blunt-end ligation is facilitated by using higher concentrations of DNA and ~10 times more ligase. Unlike the ligation of cohesive ends which requires compatible termini, blunt fragments are equally ligatable no matter which endonuclease is used to produce the

ends. The joining of blunt ends produced by different enzymes almost always results in a product that cannot be cleaved by either enzyme.

The most difficult ligations are those in which all the relevant ends are blunt (e.g., cloning an *Sma*I fragment into an *Sma*I cleaved vector). The problems are similar to those described in Example 3.16.1 and are compounded by the inefficiency of blunt-end ligation.

Cloning involving blunt ends is considerably easier if one of the ligation events involves cohesive ends. For example, two fragments generated by *Sma*I and *Eco*RI are much more easily joined than two *Sma*I-generated segments. In addition, such heterologous ends permit directional cloning and the background is very low (see Example 3.16.2).

Example 3.16.4: Joining DNA Fragments with Incompatible Ends

Consider the case of joining an *Eco*RI end (overhang 5′ AATT) to an *Sac*I end (overhang 3′ AGCT). To join such incompatible termini, it is necessary to convert them to blunt ends. For 5′ overhangs, this is accomplished by “filling in” the ends with the Klenow fragment of *E. coli* DNA polymerase I in the presence of all 4 dNTPs (see Fig. 3.16.2 below and UNIT 3.5). This reaction should be carried out in the presence of all 4 dNTPs in order to prevent more extensive exonuclease action. For 3′ overhangs, the 3′ to 5′ exonuclease activity of T4 DNA polymerase or Klenow enzyme is used to remove the protruding nucleotides (UNIT 3.5). The T4 enzyme is preferred because it has a much more active exonuclease. However, the Klenow enzyme will do the job, and it is considerably less expensive than T4 DNA polymerase. Both polymerases are active in all restriction endonuclease buffers and hence can be added (along with the 4 dNTPs) directly to the reaction mixture after cleavage.

Once the relevant ends have been blunted, the ligation procedure proceeds as in Example 3.16.3. However, the additional step for generating the blunt ends usually results in a lowered efficiency of ligation; the more ends to be blunted, the worse the problem. The fill-in reaction for 5′ overhangs is usually more efficient than the trimming reaction for 3′ extensions. Thus, when incompatible ends must be joined, it is preferable if some of the required ligation reactions involve joining of cohesive ends; e.g., the joining of a *Bam*HI–*Eco*RI (filled) fragment to a *Bam*HI–*Sac*I (trimmed) fragment.

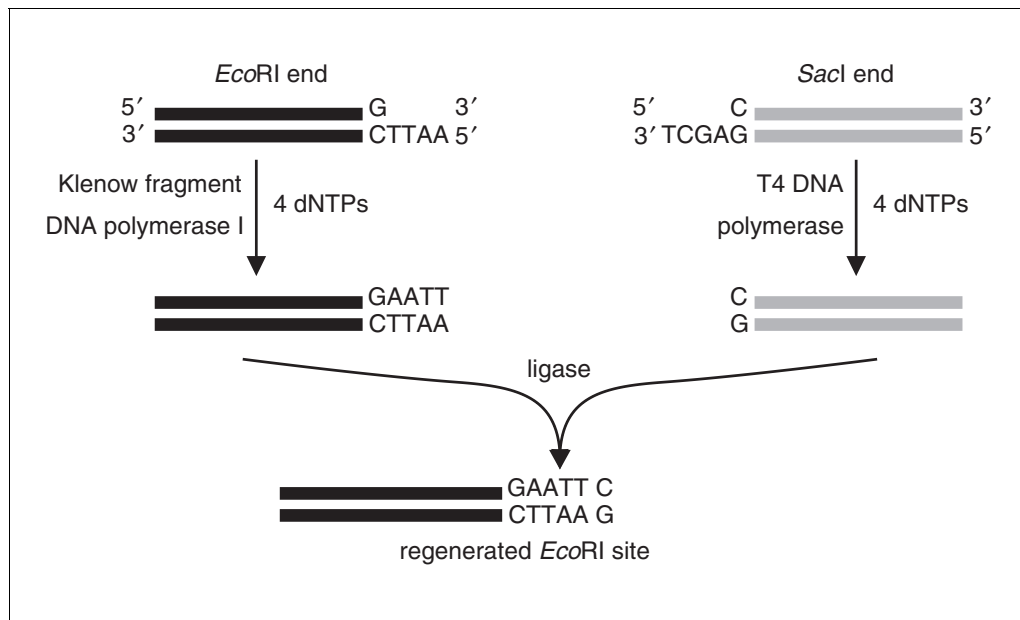


Figure 3.16.2 Joining DNA fragments with incompatible ends.

In many cases, the joining of incompatible ends generates a product that cannot be cleaved by either of the restriction enzymes used to produce the ends. In such cases, the structure of the molecule at the fusion point can be determined only by DNA sequencing.

Sometimes, however, one of enzyme sites should be regenerated, thus making it possible to analyze the prospective molecules by restriction mapping. For example, a filled in *EcoRI* site has the terminal sequence GAATT. If this is joined to a fragment whose 5' terminal nucleotide is C, the *EcoRI* site will be restored. In principle, a *SacI* site (GAGCTC) that has been trimmed will regenerate an *EcoRI* site that has been filled in. Products where the *EcoRI* and *SacI* ends have been joined aberrantly will not result in the regeneration of the *EcoRI* site.

Example 3.16.5: Ligation of Fragments Generated by Partial Cleavage

Sometimes a given DNA fragment will contain internal restriction sites that are the same as those on one (or both) of the ends. In this case, the fragment can be obtained only by partial restriction endonuclease digestion. Contrary to popular opinion, it is not difficult to obtain hybrid DNA molecules when one or more of the components is generated by partial cleavage by a restriction endonuclease(s). By gel purifying the desired partial cleavage product(s), the ligation reactions can be set up in the normal way, and they should be equally efficient. The main difficulty with partial cleavage,

obtaining enough of the specific fragment, can usually be overcome by starting with more DNA (1 to 5 μg should be enough) prior to gel electrophoresis. To ensure the proper degree of partial digestion, the method of serial enzyme dilution should be performed (UNIT 3.1).

In some cases, it will be difficult to purify the desired partial cleavage product from undesired products. This is especially true if the starting molecule contains many cleavage sites or if cleavage sites are close together. However, if the ligation reaction has a low background, such impurities are only a minor problem. For example, if the “band” that is cut from the gel contains four different clonable fragments, roughly 25% of the transformants should be the ones of interest. Moreover, if the fragment is generated by partial cleavage at one end and complete cleavage by another enzyme at the other end, many of the contaminating fragments cannot be cloned into the vector and hence will not contribute to the background.

Example 3.16.6: Ligations Involving Oligonucleotide Linkers

Synthetic linkers are self-complementary oligonucleotides, typically 8 to 12 bases in length, that anneal to form blunt-ended, double-stranded DNA which contains a site for restriction endonuclease cleavage. For example, GGAATTCC is an 8-bp *EcoRI* linker (the underlined nucleotides represent the recognition site). Oligonucleotide linkers facilitate

the cloning of blunt-ended DNA fragments, and they are valuable for “introducing” new restriction sites at desired positions. A wide variety of linkers, each containing a specific recognition sequence, is available from commercial suppliers. Linkers are usually obtained in the nonphosphorylated form. For most applications, they are phosphorylated using ATP and T4 polynucleotide kinase (UNIT 3.10). After the kinase reaction, the phosphorylated linkers can be added directly to the ligation reaction.

Ligations involving linkers are performed in three stages. First, the DNA of interest, which must contain blunt ends, is joined to the linker, as illustrated in Figure 3.16.3. This reaction is performed with 0.1 to 1 µg of linker, which represents a 100- to 1000-fold molar excess of linker over fragment. The high concentration of ends makes this blunt-end ligation reaction more efficient than normal. Generally, the linker is phosphorylated prior to ligation. In this case, the ligation products will typically contain several linkers on each end of the fragment. Some experiments involve ligation of nonphosphorylated linkers, in which case the products will contain only one linker on each end.

Second, after heat treatment to inactivate the ligase, the products of this reaction are cleaved with the restriction enzyme that recognizes the linker, as illustrated in Figure 3.16.3. Because of the high concentration of restriction sites due to the linker, this step usually requires incubation for several hours with a large amount of restriction enzyme (20 to 50 U). It is important that the fragment does not contain an internal site that is recognized by the restriction enzyme (unless the fragment has been methylated previously by the methylase corresponding to the

restriction enzyme). If possible, the ligation products should be cleaved with a second restriction enzyme to produce a fragment with heterologous ends suitable for cloning (see Example 3.16.2).

Third, the fragment is purified by gel electrophoresis and then ligated by conventional methods. The gel purification step also removes unligated linkers, which would interfere with this second ligation step. Linkers can also be removed by other methods (chromatography on Sepharose CL-4B or Sephacryl S-300; see UNITS 2.6 and 5.6), but electrophoresis is more efficient and allows for purification of the desired fragment away from undesired fragments.

Example 3.16.7: Cloning Synthetic Oligonucleotides

Double-stranded oligonucleotides suitable for cloning can be obtained by annealing two single-stranded oligonucleotides that have been synthesized chemically. Generally, such oligonucleotides are phosphorylated with T4 kinase (UNIT 3.10) prior to annealing. Double-stranded and phosphorylated oligonucleotides can also be generated by “mutually primed synthesis” of appropriate single-stranded oligonucleotides (UNIT 8.2). Oligonucleotides obtained by either method can be cloned by standard ligation procedures using cohesive or blunt ends.

The main problem associated with cloning oligonucleotides is that of multiple insertion. This is because small amounts of oligonucleotide represent a large molar excess. This problem is best dealt with by making a series of 10-fold serial dilutions of oligonucleotide and setting up parallel ligation reactions using

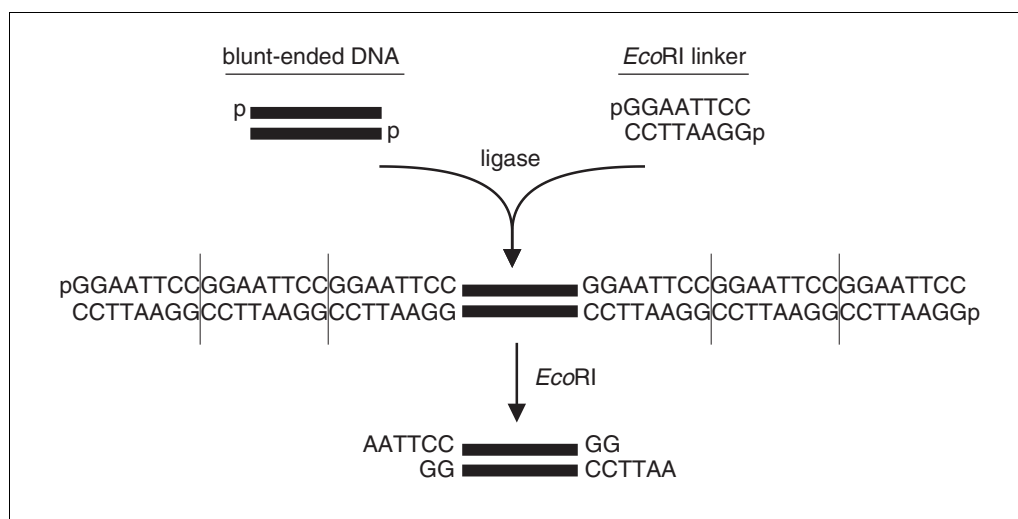


Figure 3.16.3 Joining blunt-ended DNA to *EcoRI* linker.

these various concentrations. If single insertions are desired, transformants should be picked off plates that represent the lowest amount of oligonucleotide needed to increase the number of colonies significantly above the background (i.e., the absence of oligonucleotide). In general, optimal conditions for single insertion occur when the oligonucleotide is present in a 5- to 20-fold molar excess over the other fragments in the ligation mixture.

Literature Cited

- Dugaiczyk, A., Boyer, H.W., and Goodman, H.M. 1975. Ligation of *Eco*RI endonuclease-generated DNA fragments into linear and circular structures. *J. Mol. Biol.* 96:174-184.
- Struhl, K. 1985. A rapid method for creating recombinant DNA molecules. *Biotechniques* 3:452-453.

Contributed by Kevin Struhl
Harvard Medical School
Boston, Massachusetts

Constructing Recombinant DNA Molecules by the Polymerase Chain Reaction

UNIT 3.17

Any two segments of DNA can be ligated together into a new recombinant molecule using the polymerase chain reaction (PCR). The DNA can be joined in any configuration, with any desired junction-point reading frame or restriction site, by incorporating extra nonhomologous nucleotides within the PCR primers. Cloning by PCR is often more rapid and versatile than cloning with standard techniques that rely on the availability of naturally occurring restriction sites and require microgram quantities of DNA. It is not necessary to know the nucleotide sequence of the DNA being subcloned by this technique, other than the two short flanking regions (~20 bp) that serve as anchors for the two oligonucleotide primers used in the amplification process. Moreover, PCR can be performed on low-abundance or even degraded DNA (or RNA) sources.

This unit describes using PCR to construct hybrid DNA molecules. The main objective is to give an overview of how PCR can be exploited to accomplish numerous cloning strategies; it is assumed that the reader is already familiar with basic molecular biology techniques including PCR amplification (UNIT 15.1) and subcloning (UNIT 3.16). The basic protocol outlines the PCR amplification and cloning strategies. A troubleshooting guide for problems most frequently encountered in PCR cloning, and three specific examples of this technique—for creating (1) in-frame fusion proteins, (2) recombinant DNA products, and (3) deletions and inversions by inverse PCR—are presented in the Commentary.

SUBCLONING DNA FRAGMENTS

In this protocol, synthetic oligonucleotides incorporating new unique restriction sites are used to amplify a region of DNA to be subcloned into a vector containing compatible restriction sites. The amplified DNA fragment is purified, subjected to enzymatic digestion at the new restriction sites, and then ligated into the vector. Individual subclones are analyzed by restriction endonuclease digestion and either sequenced or tested in a functional assay. The procedure is summarized in Figure 3.17.1.

Materials

Template DNA (1 to 10 ng of plasmid or phage DNA; 20 to 300 ng of genomic or cDNA)

Oligonucleotide primers (0.6 to 1.0 mM; UNIT 8.5)

Mineral oil

TE-buffered phenol (UNIT 2.1) and chloroform

100% ethanol

TE buffer, pH 8.0 (APPENDIX 2)

Klenow fragment of *E. coli* DNA polymerase I (UNIT 3.5)

Vector DNA

Calf intestinal phosphatase (UNIT 3.10)

Additional reagents and equipment for phosphorylating synthetic oligonucleotides (UNIT 3.10), enzymatic amplification of DNA by PCR (UNIT 15.1), agarose and polyacrylamide gel electrophoresis (UNITS 2.5 & 2.7), DNA extraction and precipitation (UNIT 2.1), purification of DNA by glass beads, electroelution from agarose gels, or from low-gelling/melting temperature agarose gels (UNIT 2.6), restriction endonuclease digestion (UNIT 3.1), ligation of DNA fragments (UNIT 3.16), transformation of *E. coli* (UNIT 1.8), plasmid DNA minipreps (UNIT 1.6), and DNA sequence analysis (UNIT 7.4)

**BASIC
PROTOCOL**

**Enzymatic
Manipulation
of DNA and RNA**

Contributed by Elaine A. Elion

Current Protocols in Molecular Biology (1993) 3.17.1-3.17.10

Copyright © 2000 by John Wiley & Sons, Inc.

3.17.1

Supplement 24

Amplify the target DNA

1. Prepare the template DNA. If using an impure DNA preparation (i.e., not purified by CsCl gradients), heat sample 10 min at 100°C to inactivate nucleases.

Plasmid, phage, genomic, or cDNA, obtained from either rapid preparations or purified on CsCl gradients, can be used as the source of target DNA.

2. Prepare the oligonucleotide primers. If the PCR product is to be cloned by blunt-end ligation, phosphorylate the 5' hydroxyl of the oligonucleotide primers.

A 5' phosphate on the ends of the PCR products will be needed to form the phosphoester linkage to the 3'OH of the vector during ligation. This step is essential if the vector has been treated with a phosphatase.

Because the purity of the oligonucleotides does not seem to affect the PCR reaction, primer purification (as detailed in UNIT 8.5) may not be necessary.

3. Set up a standard amplification reaction and overlay with mineral oil as described in UNIT 15.1. Carry out PCR in an automated thermal cycler for 20 to 25 cycles under the following conditions: denature 60 sec at 94°C, hybridize 1 min at 50°C, and extend 3 min at 72°C. Extend an additional 10 min at 72°C in the last cycle to make products as complete as possible.

Include negative controls of no template DNA and each oligonucleotide alone, as well as several different oligonucleotide:template ratios.

For a discussion of optimization of amplification conditions see UNIT 15.1. A thermostable DNA polymerase with 3'→5' exonuclease proofreading activity can be used instead of Taq DNA polymerase to reduce the amount of nucleotide misincorporation during amplification. Pfu DNA polymerase (Stratagene) and Vent DNA polymerase (New England Biolabs) have this activity (follow manufacturers' instructions).

Recover the amplified fragment

4. Analyze an aliquot (e.g., 4 to 8 µl) of each reaction mix by agarose or polyacrylamide gel electrophoresis to verify that the amplification has yielded the expected product.
5. Recover amplified DNA from PCR reaction mix. Remove mineral oil overlay from each sample, then extract sample once with buffered chloroform to remove residual mineral oil. Extract once with buffered phenol and then precipitate DNA with 100% ethanol.

Carrier tRNA may be added during precipitation if desired.

6. Microcentrifuge DNA 10 min at high speed, 4°C. Dissolve pellet in 20 µl TE buffer. Purify desired PCR product from unincorporated nucleotides, oligonucleotide primers, unwanted PCR products, and template DNA using glass beads, electroelution, or phenol extraction of low gelling/melting temperature agarose.

Unused oligonucleotide primers can inhibit the ability of the restriction enzymes to digest the amplified PCR product. Amplified DNA that is greater than 100,000 Da (>150 bp) can be rapidly separated from the primers using a Centricon 100 microconcentrator unit from Amicon (follow manufacturers' instructions). If the gel analysis in step 4 shows that amplification yielded only the desired PCR product, this microconcentrated DNA can be used directly for cloning.

Prepare amplified fragment and vector for ligation

- 7a. If the PCR fragment is to be cloned by blunt-end ligation, repair the 3' ends with DNA polymerase I (Klenow fragment).

This step is necessary because Taq DNA polymerase adds a nontemplated nucleotide (usually dA) to the 3' ends of PCR fragments.

- 7b. If primers contain unique restriction sites, digest half the amplified DNA in 20 μ l with the appropriate restriction enzyme(s). Use an excess of enzyme, and digest for several hours.

Reserve the undigested half for future use, if necessary.

8. Prepare the recipient vector for cloning by digesting 0.2 to 2 μ g in 20 μ l with compatible restriction enzymes. If necessary, treat vector DNA with calf intestinal phosphatase (*UNIT 3.1*) to prevent recircularization during ligation.
9. Separate the linearized vector from uncut vector by agarose or low-gelling/melting temperature gel electrophoresis. Recover linearized vector from the gel by adsorption to glass beads, electroelution, or phenol extraction of low-gelling/melting temperature agarose.

Ligate amplified fragment and vector

10. Ligate the PCR fragment into the digested vector following the procedure outlined in *UNIT 3.16*.
11. Transform an aliquot of each ligation into *E. coli*. Prepare plasmid miniprep DNA from a subset of transformants.

Analyze recombinant plasmids

12. Digest the plasmid DNA of the selected transformants with the appropriate restriction endonuclease. Analyze the digestions by agarose gel electrophoresis to confirm fragment incorporation.
13. Sequence the amplified fragment portion of the plasmid DNA to check for mutations. Alternatively, screen the subset of transformants using a biochemical or genetic functional assay if available.

This analysis is critical because the Taq DNA polymerase can introduce mutations into the amplified fragment.

COMMENTARY

Background Information

The main benefit of cloning by PCR is that unique restriction sites can be introduced on either side of any segment region of amplified DNA to allow its ligation into a recipient vector (Mullis and Faloona, 1987; Chapter 15) in any configuration. The incorporation of additional nucleotides at the 5' ends of the oligonucleotide primers permits the creation of novel restriction sites or changes in reading frame and coding sequence. The oligonucleotide primers can also be designed to contain mismatches, deletions, or insertions in the region of homology (*UNIT 8.5*). However, it is not necessary to always incorporate a new restriction site in the primer. Amplified PCR fragments can also be subcloned by blunt-end or sticky-end ligation using preexisting restriction sites within the amplified DNA. For example, PCR might be used to amplify a target DNA that already contains appropriate restriction sites, but is available in limited quantities. Finally, sequential polym-

erase chain reactions can be used to generate more complex recombinant PCR products which can subsequently be subcloned into a recipient vector.

The most obvious disadvantage of PCR cloning is the need to verify that the subcloned PCR product does not contain mutations generated during the polymerase chain reaction. In cases where longer DNA segments (i.e., >1 kb) are being amplified, it may be more advantageous to use a polymerase which has a 3'→5' exonuclease activity (e.g., Pfu polymerase, Stratagene) to reduce the chances of generating mutations. After subcloning, several independent PCR products should be analyzed by DNA sequencing to be sure that the recombinant DNA molecule is not mutated. Sequencing can be laborious when a large fragment of DNA is subcloned. However, subcloned PCR products can be prescreened by either biological or biochemical functional assays if they are available. In some instances, it may be more desir-

able to break down the cloning into several steps that might involve the introduction of a needed restriction site within a short piece of DNA first.

Critical Parameters

In general, the DNA preparation, purification, and ligation guidelines outlined in *UNIT 3.16* should be applied to PCR cloning to ensure recovery of the desired ligation products. However, the following points deserve special consideration.

Design of oligonucleotide primers. Primers should only hybridize to the sequence of interest. This can be predicted in instances where sequence information is available. In general, primers with homology of 16 to 20 nucleotides to the target DNA and a GC content of ~50% should be chosen. A longer oligonucleotide of ~25 nucleotides should be used for AT-rich regions. In instances where genomic DNA is the source of the target DNA, the oligonucleotide primers should contain at least 20 nucleotides of homology to the target DNA to ensure that they anneal specifically (Arnheim and Erlich, 1992).

When using primers to introduce a specific restriction site, a sequence within the target DNA should be selected that requires the addition of the fewest noncomplementary nucleotides to create the new site, if possible. Special consideration should be given to the choice of site itself, as restriction endonucleases vary in their ability to cleave recognition sequences within ten nucleotides of the end of a DNA duplex (consult Table 8.5.1 for the efficacies of different restriction enzymes in cleaving terminal recognition sequences). It is also recommended that four to five additional nucleotides be added on the 5' side of the restriction site in the primer. Because DNA duplexes "breathe" at termini, potentially interfering with the ability of a restriction enzyme to cleave (Innis et al., 1990), it is useful to use the GCGC "clamp" sequence that is most thermostable (Sheffield et al., 1989).

Finally, the sequence of the primer should be checked for internal complementarity to avoid secondary structure formation that will interfere with hybridization of the primer to the target DNA. The 3' ends of the two primers being used must not be complementary, so that the formation of primer-dimers that will compete with the synthesis of the desired PCR product will be avoided.

Additional details on primer design are discussed in *UNIT 15.1*.

DNA polymerase. Commercially available *Taq* DNA polymerase (Perkin-Elmer Cetus) lacks the 3'→5' proofreading exonuclease activity used by DNA polymerase I Klenow fragment and T4 DNA polymerase to reduce error frequency (Kornberg, 1992). This absence of proofreading activity in *Taq* DNA polymerase is thought to result in a heightened error frequency. Old estimates indicate that the average rate of misincorporation is 8.5×10^{-6} nucleotides per cycle (Goodenow et al., 1989; Fucharoen et al., 1989). Two other thermostable DNA polymerases possessing proofreading 3'→5' exonuclease activity have recently become commercially available: Pfu DNA polymerase, purified from *Pyrococcus furiosus* (Stratagene) and Vent DNA polymerase, purified from *Thermococcus litoralis* (New England Biolabs and Promega). Both are more thermostable than *Taq* DNA polymerase. Pfu DNA polymerase is 12-fold more accurate than *Taq* DNA polymerase, as assayed by the method of Kohler et al. (1991). Vent DNA polymerase is 4-fold more accurate than *Taq* DNA polymerase (Cariello et al., 1991). Although it is difficult to compare the relative error frequencies of three enzymes because they were assayed by different methods, the use of either Vent or Pfu DNA polymerases may reduce the amount of misincorporation.

Removal of unincorporated nucleotide triphosphates. It is recommended that the amplified PCR fragment be purified from unincorporated nucleotides and primers. Any method of purification that involves electrophoresis can also separate the desired PCR product from any undesired DNA species produced during amplification. Typical methods of DNA purification include electrophoresis through low-gelling/melting temperature agarose or electrophoresis through agarose followed by DNA purification by electroelution or adsorption to glass beads (*UNIT 2.6*). However, amplified DNA can be more rapidly purified from unincorporated nucleotide triphosphates and primers using a Centricon microconcentration unit (Amicon). The disadvantage of using the microconcentrator is that undesirable PCR products and the starting template DNA will copurify with the amplified PCR fragment.

Troubleshooting

PCR amplification. The use of appropriately designed primers should allow the amplification of the DNA segment of interest. Occasionally, however, primers may not be specific, leading to the amplification of undesired DNA

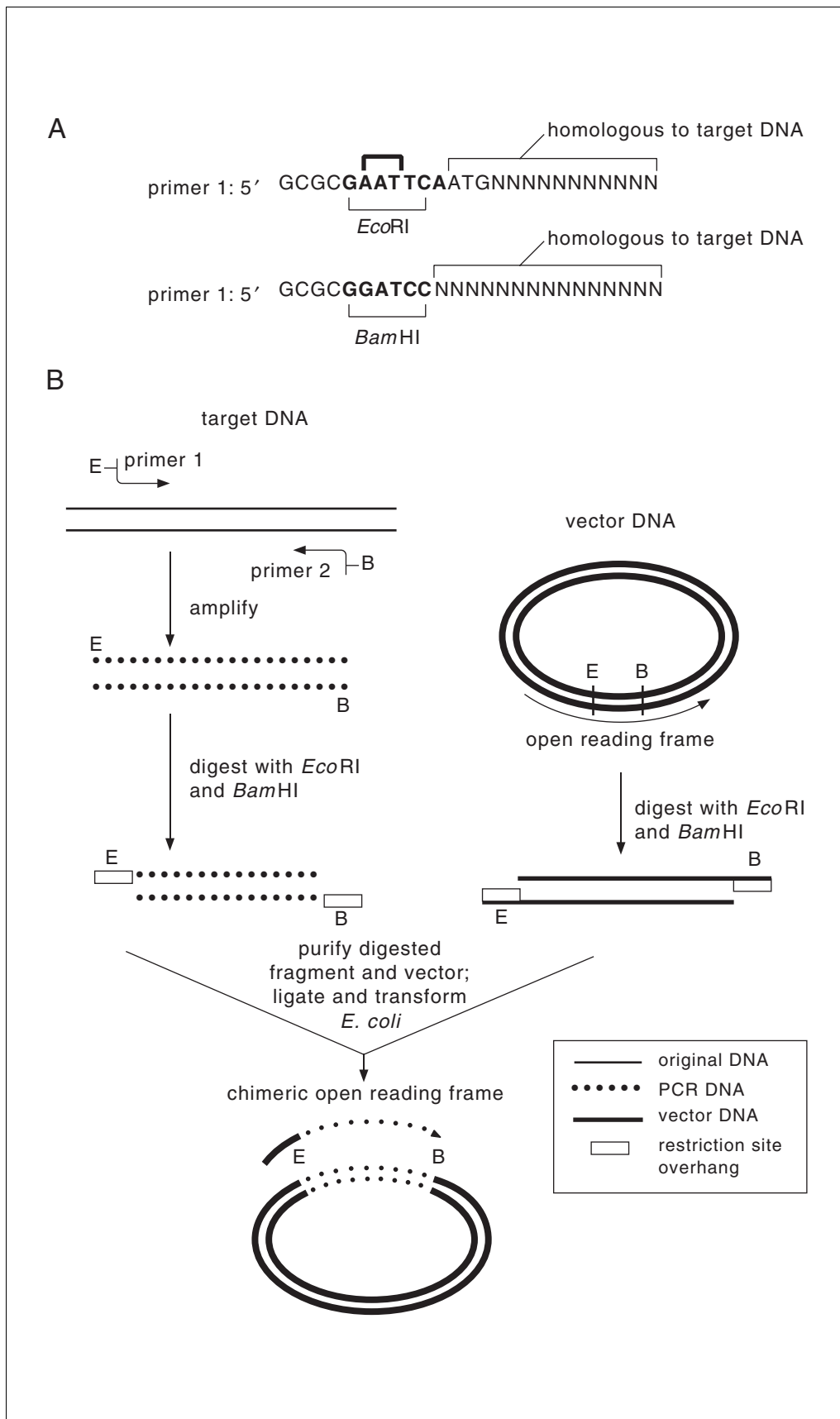


Figure 3.17.1 Introducing unique restriction sites and creating an in-frame fusion protein by PCR. Abbreviations: E, *EcoRI*; B, *BamHI*. For a full description, see Example 3.17.1 in Commentary.

segments. The specificity of primer to template hybridization will depend upon temperature and salt (see *UNITS 15.1 & 6.4* for a thorough discussion). The highest annealing temperature possible should be used to reduce nonspecific associations. Some nonspecific amplifications can be avoided by employing a “hot-start” technique—i.e., adding the DNA polymerase to a prewarmed sample (D’Aquila et al., 1991). In addition, purifying the PCR product by gel electrophoresis will help ensure that the proper DNA fragment is subcloned. If the synthesized primer does not bind with specificity, it may be simplest to have another one synthesized.

In setting up the amplification cycle, keep in mind that for cloning, fidelity is more important than yield, so it is better to keep a low cycle number and not to raise the MgCl₂ concentration too much. 1.5 mM MgCl₂ in the amplification buffer should be sufficient for most primers.

Cloning of the amplified fragment. “Sticky-end” ligation of the amplified DNA can sometimes be difficult, due to poor cutting of the terminal restriction site by the desired restriction endonuclease. Careful choice of restriction sites and the addition of extra nucleotides to the 5′ end of the primer (see critical parameters) will facilitate digestion.

Other common explanations for poor cutting by restriction endonucleases include:

(1) Blockage of the duplex terminus by bound *Taq* DNA polymerase. If this is the case, the amplified DNA can be treated with proteinase K to remove associated protein. This is done by adding 50 μg/ml proteinase K in 10 mM Tris-Cl (pH 7.8)/5 mM EDTA/0.5% (v/v) SDS to the sample and incubating 30 min at 37°C. The sample must be extracted with phenol/chloroform (*UNIT 2.1*) to remove the proteinase K.

(2) Inefficient extension by *Taq* DNA polymerase. The nonduplex ends generated in this fashion can be repaired by filling in with Klenow fragment (*UNIT 3.5*).

(3) An insufficient number of extra nucleotides 5′ to the restriction site in the primer. In this case, the primer should be resynthesized. Prolonged (typically overnight) DNA digestion should also be tried.

(4) 5′ terminal breathing of duplex DNA. Duplex formation can be stabilized by the inclusion of 0.1 mM spermidine in the digestion reaction.

An alternative approach is to internalize the restriction site by concatemering the PCR products prior to restriction endonuclease di-

gestion (Jung et al., 1990). To do this, 5′ phosphorylated oligonucleotide primers are used in the PCR amplification, or following amplification, the DNA fragment is phosphorylated using T4 polynucleotide kinase (*UNIT 3.10*). The amplified fragments are then concatemered by ligation using T4 DNA ligase (*UNIT 3.16*) prior to restriction endonuclease treatment. Aliquots of the amplified DNA before and after concatemering, and after digestion, can be compared on an agarose gel to confirm that the procedure worked.

Blunt-end ligations are often inhibited by nonflushed ends in the PCR fragment, due to the presence of a nontemplate-directed nucleotide (usually dATP) added by *Taq* DNA polymerase (Clark, 1988). Treatment of the PCR fragment with Klenow polymerase in the presence of dNTPs (*UNIT 3.5*) to make the ends flush should circumvent this problem. In the event that this simple approach does not work, the PCR fragment can be subcloned into a *MstII* (or *Bsu36 I*) site (CC↓TNAGG) which leaves a 5′ dT overhang. This approach requires a recipient vector with a unique *MstII* site.

Anticipated Results

All of the cloning approaches outlined are reliable and should result in efficient recovery of the desired recombinant molecules.

Time Considerations

Once the oligonucleotides have been synthesized, the PCR amplification, purification, ligation, and transformation steps can all be done within 2 days. The appropriate subclones can then be sequenced or tested in a functional assay immediately thereafter.

EXAMPLES

Example 3.17.1: Creating In-Frame Fusion Proteins by PCR

PCR cloning is particularly useful for creating in-frame fusions between two open reading frames, as is often done for synthesizing fusion proteins with *E. coli* expression vectors (Chapter 16). The essence of this type of subcloning involves incorporating additional noncomplementary nucleotides within the oligonucleotide primer that will encode the junction sequences of the amplified PCR fragment. Consider the introduction of a unique *EcoRI* site into a piece of target DNA that is to be fused with an open reading frame in the recipient vector, as depicted in Figure 3.17.1.

For this experiment, the primers should be

designed as indicated in panel A. Each primer is designed to contain a unique restriction site not present within the target DNA. The primer carrying the *Eco*RI site contains an additional nucleotide (shown in bold) to allow the ATG of the amplified target DNA to be in-frame with the open reading frame in the vector (bold bracket). The second primer contains a unique *Bam*HI site. Both oligonucleotide primers are designed to be homologous to and anneal with ~20 nucleotides of DNA flanking the target DNA. They are oriented such that their 3' hydroxyl ends point toward the target DNA. These unique restriction sites are 5' to the region of the primer that is homologous to the target DNA. Each new restriction site is separated from the 5' end of the oligonucleotide by four additional nucleotides to facilitate enzymatic digestion of the amplified DNA. Shown in this example is a GC clamp (Myers et al., 1985) which favors duplex formation at the ends of the amplified fragment.

Panel B depicts the sequence of events in this experiment. Following primer annealing and PCR amplification, the amplified DNA is first digested with restriction endonucleases that cleave at the new restriction sites, then purified by gel electrophoresis. The recipient vector DNA is digested with either the same or compatible restriction endonucleases and is purified before being ligated to the recipient vector.

In this example, one oligonucleotide contains additional bases to create an in-frame fusion with the plasmid-borne open reading frame. More elaborate primers can be designed to include additional restriction sites in different reading frames to allow subcloning of the same PCR fragment into multiple recipient fusion vectors that may have nonidentical cloning sites. This is efficient both in terms of labor and the cost of having to synthesize a new oligonucleotide primer. Note that an ATG can also be incorporated into the 5' oligonucleotide primer to create an open reading frame with a new translational start (e.g., in the construction of a promoter-exon fusion).

Example 3.17.2: Creating a Recombinant DNA Molecule by Sequential PCR Amplifications

Consider creating a chimeric DNA molecule by sequential polymerase chain reactions rather than by ligation. This technique is useful for complex cloning schemes that involve fusing together more than two pieces of DNA, as depicted in the creation of the gene fusion

shown in Figure 3.17.2. In this example, two PCR products are made from noncontiguous regions of DNA (that are also nonhomologous) in separate reactions.

Two of the first-round amplification primers are designed to contain 5' extensions that are homologous to a portion of the other target gene (see primers 1b and 1c). In this example, the primers for target gene 1 are labeled as 1a and 1b. Primer 1a contains a unique *Eco*RI site; primer 1b contains a 5' extension that is homologous to a region in target gene 2 that will be amplified (thin line of arrow). Primers 1c and 1d are for amplifying target gene 2: primer 1c contains a 5' extension that is homologous to a portion of target gene 1 that will be amplified (bold line of arrow), and primer 1d contains a unique *Bam*HI restriction site.

Because primers 1b and 1c contain complementary 5' extensions, two PCR products containing a region of overlapping homology are generated. The two PCR fragments are purified away from the primers, then mixed together and annealed by denaturation and renaturation. Four DNA species are generated in this reaction: two heteroduplexes associating at the region of overlapping homology and two parental homoduplexes. The recessed 3' ends of the heteroduplexes are extended by *Taq* DNA polymerase to produce a single fragment that is equal in length to the sum of the two overlapping fragments.

In a second round of amplification, the combined heteroduplex DNA species is amplified by adding the outside set of primers (1b and 1d) to the PCR assay. These primers will now have complete homology to the amplified heteroduplex DNA species. (Note that the parental homoduplexes will not be amplified because only one of each outside primers will anneal to each parental homoduplex.)

The complementary primers used in the first polymerase chain reaction step can be designed to either insert a restriction site at the junction between the joined PCR products, or alter a reading frame.

Example 3.17.3: Inserting a Restriction Site by Inverse PCR

Consider deleting a segment of DNA from a plasmid by inserting a unique restriction site by inverse PCR as depicted in Figure 3.17.3. Divergent primers containing the same novel restriction site (R) are annealed to two portions of the plasmid. (Note that R can be any restriction site not found in the plasmid.) The primers are oriented with their 3' ends facing away from

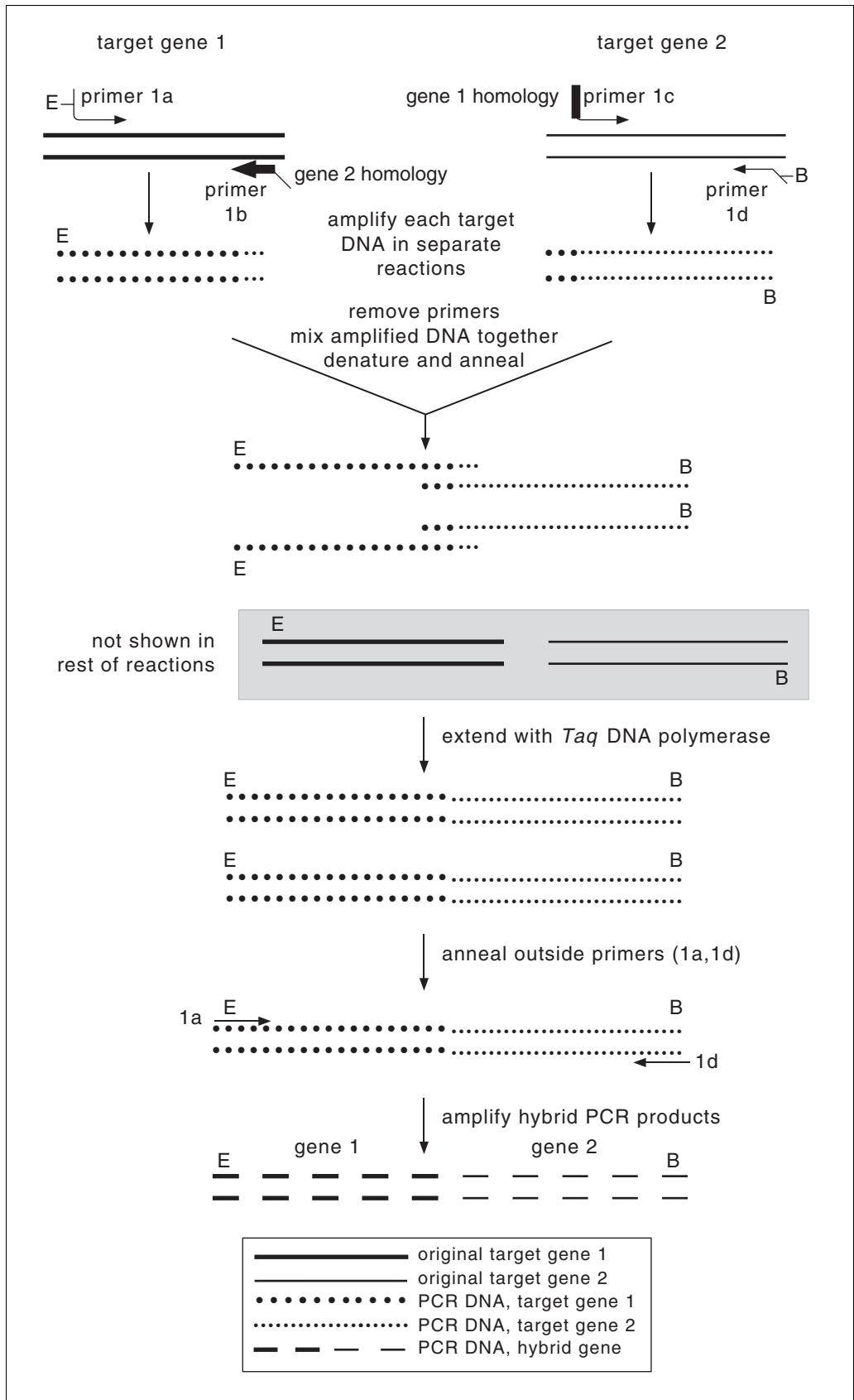


Figure 3.17.2 Creating a recombinant DNA molecule by sequential PCR amplifications. Primer 1b has a region of homology to target gene 2 (open box); primer 1c has a region of homology to target gene 1 (closed box). Abbreviations: E, *Eco*RI; B, *Bam*HI. For a complete description, see Example 3.17.2 in Commentary.

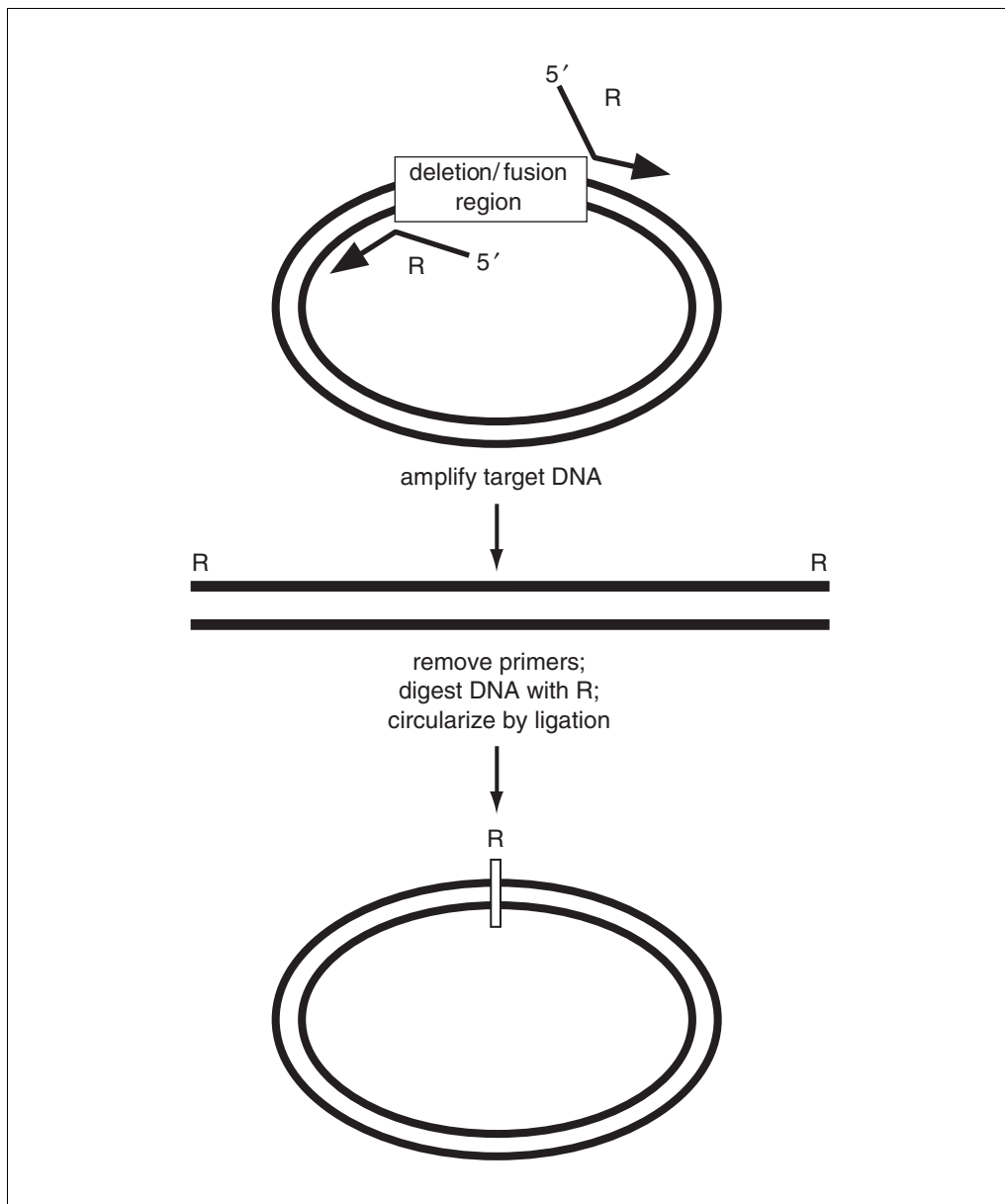


Figure 3.17.3 Inserting a restriction site by inverse PCR. Abbreviation: R, restriction enzyme not found in target plasmid. For a complete description, see Example 3.17.3 in Commentary.

each other so that sequences flanking the region to be deleted will be amplified. The full-length PCR product is purified and digested at the new restriction site and the ends are ligated in a unimolecular reaction. The amount of spacing between the two primers will determine whether the final product contains all of the original sequence or a deletion. Insertions can also be made at specific sites by including 5' extensions in the primers. Mismatches within the body of the primer can also be used to introduce mutations.

This method is useful for rapid introduction of desired restriction sites, and is limited only by the size of the plasmid and the ability of

DNA polymerase to synthesize complete products. This methodology allows the amplification of DNA flanking a region of known sequence. It is also useful for cloning DNA that has not yet been sequenced and for making hybridization probes (Ochman et al., 1990).

Literature Cited

- Arnheim, N. and Erlich, H. 1992. Polymerase chain reaction strategy. *Annu. Rev. Biochem.* 61:131-156.
- Cariello, N.F., Swenberg, J.A., and Skopek, T.R. 1991. Fidelity of *Thermococcus litoralis* DNA polymerase (Vent) in PCR determined by denaturing gradient gel electrophoresis. *Nucl. Acids Res.* 19:4193-4198.

- Clark, J.M. 1988. Novel nontemplated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucl. Acids Res.* 16:9677-9689.
- D'Aquila, R.T., Bechtel, L.J., Videler, J.A., Eron, J.J., Gorczyca, P., and Kaplan, J.C. 1991. Maximizing sensitivity and specificity of PCR by pre-amplification heating. *Nucl. Acids Res.* 19:3749.
- Fucharoen, S., Fucharoen, G., Fucharoen, P., and Fukamaki, Y. 1989. A novel *ochre* mutation in the beta-thalassemia gene of a Thai identified by direct cloning of the entire beta-globin gene amplified using polymerase chain reactions. *J. Biol. Chem.* 264:7780-7783.
- Goodenow, M., Huet, T., Saurin, W., Kwok, S., Sninsky, J., and Wain-Hobson, S. 1989. HIV-1 isolates are rapidly evolving quasi-species: Evidence for viral mixtures and preferred nucleotide substitutions. *J. Acquired Immunol. Defic. Syndr.* 2:344-352.
- Innis, M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. (eds.) 1990. PCR Protocols. Academic Press, San Diego.
- Jung, V., Pestka, S.B., and Pestka, S. 1990. Efficient cloning of PCR-generated DNA containing terminal restriction endonuclease sites. *Nucl. Acids Res.* 18:6156.
- Kornberg, A. and Baker, T.A. 1992. DNA Replication, 2nd ed. W.H. Freeman, New York.
- Mullis, K.B. and Faloona, F.A. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* 155:335-350.
- Ochman, H., Medhora, M.M., Garza, D., and Hartl, D.L. 1990. Amplification of flanking sequences by inverse PCR. In PCR Protocols. (M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White, eds.) pp. 219-227. Academic Press, San Diego, Calif.
- Sheffield, V.C., Cox, D.R., Lerman, L.S., and Myers, R.M. 1989. Attachment of a 40 base pair G+C-rich sequence (GC clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proc. Natl. Acad. Sci. U.S.A.* 86:232-236.

Key Reference

Innis et al., 1990. See above.

Provides an in-depth analysis of PCR methods and techniques.

Contributed by Elaine A. Elion
Harvard Medical School
Boston, Massachusetts

Labeling and Colorimetric Detection of Nonisotopic Probes

UNIT 3.18

Although a number of different nonisotopic labels have been described in the literature, biotin and digoxigenin are used most frequently and are commercially available. Either label can be easily incorporated into DNA probes and be detected colorimetrically; a number of fluorochromes, as well as alkaline phosphatase and horseradish peroxidase (which produce colored precipitates) are available directly conjugated to anti-digoxigenin antibodies and to avidin. Chemiluminescent detection methods (UNIT 3.19) and indirect immunofluorescent techniques (UNIT 14.6) also provide sensitive alternatives for many molecular biology applications. Because one of the advantages of nonisotopically labeled probes is their long shelf life (≥ 2 years), many micrograms of DNA can be labeled in one reaction to provide probes of constant quality for multiple experiments.

The two basic protocols describe incorporating biotinylated nucleotides into DNA probes by nick translation and random-primed synthesis. The support protocol describes colorimetric detection of the probes, which also serves to check the extent of nucleotide incorporation. An alternate protocol describes adaptations of the basic protocols for incorporation of digoxigenin-modified nucleotides.

PREPARATION OF BIOTINYLATED PROBES BY NICK TRANSLATION

BASIC
PROTOCOL

Nick translation is frequently used to label DNA probes for nonisotopic detection procedures. The DNA to be labeled can be an isolated fragment or—as is often the case for in situ hybridization probes—the intact phage, cosmid, or plasmid clone. The biotin-avidin system is the most widely used nonisotopic labeling method and the biotinylated deoxynucleotides and detection reagents required are commercially available from several sources.

The protocol is quite similar to the standard nick translation procedure (UNIT 3.5) except for some modifications—the biotin-11-dUTP is substituted for dTTP in a standard nick translation reaction mixture and the DNase I concentration is adjusted to ensure a size range of 100 to 500 nucleotides. A minimum of 2 hr incubation time is necessary for optimal incorporation of the modified deoxynucleotide. The reaction provides sufficient probe for 15 to 50 in situ hybridizations or 2 to 5 Southern blots.

Materials

- E. coli* DNA polymerase I (UNIT 3.5) and 10 \times buffer (UNIT 3.4)
 - 0.5 mM 3dNTP mix (minus dTTP; UNIT 3.4)
 - 0.5 mM biotin-11-dUTP stock
 - 100 mM 2-mercaptoethanol (2-ME)
 - Test DNA
 - 1 mg/ml DNase I stock (UNIT 3.12) prepared in 0.15 M NaCl/50% glycerol
 - DNA molecular weight markers (UNIT 2.5A)
 - 0.5 M EDTA, pH 8.0 (APPENDIX 2)
 - 10% (w/v) SDS
 - 100% ethanol
 - SDS column buffer
 - 1-ml syringe
- Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5) and removal of unincorporated nucleotides by Sephadex G-50 spin columns (UNIT 3.4)

Enzymatic
Manipulation
of DNA and RNA

1. Prepare a 100- μ l reaction mix as follows:

10 μ l 10 \times *E. coli* DNA polymerase I buffer
10 μ l 0.5 mM 3dNTP mix
10 μ l 0.5 mM biotin-11-dUTP stock
10 μ l 100 mM 2-ME
2 μ g DNA
20 U *E. coli* DNA polymerase I
DNase I stock diluted 1:1000 in cold H₂O immediately before use
H₂O to 100 μ l.

Incubate reaction 2 to 2.5 hr at 15°C.

Other biotinylated nucleotides such as biotin-14-dATP and biotin-16-dUTP can be used. Adjust the mixture of unsubstituted deoxynucleotides accordingly. The quantity of DNase I depends on the individual lot used. The enzyme can be diluted in water as described, or in a buffer that is compatible with the nick translation reaction. See notes on DNase I titration in critical parameters.

2. Place reaction on ice. Remove a 6- μ l aliquot, boil it 3 min and place on ice 2 min.
3. Load aliquot on an agarose minigel, along with suitable size markers (0.1- to 10-kb range). Run gel quickly (15 V/cm) in case additional incubation is necessary.
4. If the digested DNA is between 100 and 500 nucleotides, proceed to step 5. If the probe size is between 500 and 1000 nucleotides (or larger) add a second aliquot of DNase I and incubate further.

Additional DNase I is added in a more concentrated form than in the initial reaction to avoid significant volume changes. For example, if 10 μ l of a 1:1000 dilution of DNase I is added initially, the second addition might be 1 μ l of a 1:100 dilution incubated for ~1 hr. Monitor the additional incubation carefully to avoid complete digestion of the probe. If little or no digestion occurs, purify a new DNA sample by phenol extraction and ethanol precipitation (UNIT 2.1) and repeat the nick translation.

5. Add 2 μ l of 0.5 M EDTA, pH 8.0 (10 mM final) and 1 μ l of 10% SDS (0.1% final) to the reaction. Heat 10 min at 68°C to stop the reaction and inactivate the DNase I.
6. Prepare a Sephadex G-50 spin column in a 1-ml syringe. Wash syringe and silanized glass wool plug with 2 ml of 100% ethanol, then 4 ml water. Pack G-50 resin to the 1-ml mark. Wash column 3 to 4 times with 100 μ l SDS column buffer before loading the sample. An additional wash after loading the sample is not necessary.

Alternatively, the G-50 resin can be stored in SDS column buffer, reducing the number of times the column should be washed to 2 to 3 times. SDS prevents the biotinylated DNA from sticking nonspecifically to the resin and the glass wool plug.

7. Separate the biotinylated probe from unincorporated nucleotides. The eluted probe concentration should be ~20 ng/ μ l (for 2 μ g nick-translated DNA) and is ready to use without further treatment. Probe can be stored at -20°C for years without loss of activity.

Biotinylated DNA should not be subjected to phenol extraction, as biotin causes the probe to partition to the phenol/water interface or completely into the phenol if heavily biotinylated.

8. Assess the extent of the biotinylation reaction and the probe quality by colorimetric (support protocol) or chemiluminescent (UNIT 3.19) detection.

Chemiluminescent detection is 5×10^3 times more sensitive than colorimetric detection (Beck and Koster, 1990).

PREPARATION OF BIOTINYLATED PROBES BY RANDOM OLIGONUCLEOTIDE-PRIMED SYNTHESIS

BASIC PROTOCOL

This protocol outlines a priming reaction using random octamers that have been biotinylated at the 5' end. Octamers are the optimal length for efficient hybridization of 5'-labeled oligonucleotides because the additional two bases act to minimize the steric hindrance from the 5' end label. This ensures that every probe molecule generated contains at least one biotin, but biotin-16-dUTP is also included in the reaction so that additional biotin is incorporated into some of the probe molecules. Other biotinylated nucleotides (e.g., biotin-14-dATP and biotin-11-dUTP) can be used in place of the biotin-16-dUTP.

The starting DNA template should be linear. Best results are obtained with isolated insert DNA that is ≥ 200 bp. The use of shorter templates will lead to ineffective probe synthesis because of the inability for many random primers to bind on a short target. The amount of template that can be labeled by this procedure ranges from 25 ng to 2 μ g.

Materials

Linear template DNA
Biotinylated random octamers
dNTP/biotin mix
5 U/ μ l Klenow fragment (UNIT 3.5)
TE buffer, pH 7.5 (APPENDIX 2)
0.5 M EDTA, pH 8.0 (APPENDIX 2)
4 M LiCl
100% and 70% ethanol, ice-cold

1. Place 500 ng to 1 μ g template DNA in a 1.5-ml microcentrifuge tube. Add nuclease-free water to a total volume of 34 μ l.

Up to 2 μ g template DNA can be biotinylated without scaling up the reaction volumes.

2. Denature the DNA for 5 min in boiling water. Place on ice 5 min and microcentrifuge briefly.
3. Add the following to the sample in the order listed:
 - 10 μ l biotinylated random octamers
 - 5 μ l dNTP/biotin mix
 - 1 μ l (5 U) Klenow fragment.

Incubate 1 hr at 37°C.

If desired, an additional 2.5 U of Klenow fragment can be added after 30 min of incubation to "boost" the reaction. Continuing the incubation for up to 6 hr can increase the yield.

4. Terminate reaction by adding 3 μ l of 0.5 M EDTA, pH 8.0. Precipitate the probe by adding 5 μ l of 4 M LiCl and 150 μ l ice-cold 100% ethanol. Place 30 min on dry ice.
5. Microcentrifuge 10 min at top speed, room temperature, and wash DNA pellet with ice-cold 70% ethanol.
6. Resuspend DNA pellet in 20 μ l TE buffer, pH 7.5. Assess the quality of the biotinylation reaction by colorimetric (support protocol) or chemiluminescent (UNIT 3.19) detection.

Chemiluminescent detection is 5×10^3 times more sensitive than colorimetric detection (Beck and Koster, 1990).

COLORIMETRIC DETECTION OF BIOTINYLATED PROBES

This support protocol provides a method for checking the extent of the biotinylated incorporation into the labeled probe and serves as the basis for detecting biotinylated DNA probes hybridized to Southern blots. Because the detectability of biotinylated probes is a function of the number of biotin molecules per kilobase, rather than the specific activity of a corresponding radiolabeled probe, the extent of biotinylated deoxynucleotide incorporation is checked by a colorimetric assay (similar assessment can be made by chemiluminescent detection; *UNIT 3.19*). Under the standard nick translation conditions provided, ~50 biotin molecules are incorporated per kilobase of DNA.

In this procedure, biotinylated DNA is spotted on a nitrocellulose filter. The streptavidin–alkaline phosphatase conjugate binds strongly to the biotin and is visualized when the enzyme is provided with substrates that produce a colored precipitate. As little as 2 pg biotinylated DNA should be visible. If the protocol is being utilized as a method for detecting DNA probes by Southern analysis, follow standard Southern blot protocols (*UNIT 2.9*) substituting the biotinylated probe (~100 ng/ml) for the radiolabeled probe and proceed to steps 3 to 8 below.

Additional Materials

- Biotinylated standard DNA (basic protocols or GIBCO/BRL) and test DNA
- DNA dilution buffer
- Alkaline phosphatase pH 7.5 (AP 7.5) buffer
- Blocking buffer: 3% (w/v) BSA fraction V in AP 7.5 buffer
- 1 mg/ml streptavidin–alkaline phosphatase (AP) conjugate (GIBCO/BRL)
- Alkaline phosphatase pH 9.5 (AP 9.5) buffer
- 75 mg/ml nitroblue tetrazolium (NBT)
- 50 mg/ml 5-bromo-4-chloro-3-indoyl phosphate (BCIP)
- TE buffer, pH 8.0 (*APPENDIX 2*)

- Small piece of membrane (e.g., 5 × 3 cm²): nitrocellulose *or* uncharged nylon
- Sealable bags

NOTE: To avoid nonspecific background, wear powder-free gloves when handling the membranes.

1. Prepare biotinylated standard DNA in concentrations of 0, 1, 2, 5, 10 and 20 pg/μl in dilution buffer. Dilute the biotinylated test DNA in a similar fashion.

6× SSC in the dilution buffer prevents the DNA spot from spreading on the nitrocellulose, which diminishes the color intensity of the spot.

- 2a. *For nitrocellulose membrane:* Spot 1 μl of each dilution on a small piece of nitrocellulose. Bake the membrane ~1 hr at 80°C. Proceed directly to step 3.

*Nitrocellulose membranes cannot be used in chemiluminescent detection procedures (*UNIT 3.19*).*

- 2b. *For nylon membrane:* Spot 1 μl of each dilution on the nylon membrane. Air-dry and cross-link DNA to the membrane by UV illumination (*UNIT 3.19*). Proceed to step 3 or develop membrane as described for chemiluminescent detection (*UNIT 3.19*).

This cross-linking step is critical. If done incorrectly, ≤50% of the DNA can be washed off the membrane during the detection steps, resulting in inaccurate probe assessment.

If chemiluminescent detection is used, the test DNA should be diluted in a series from 10⁻¹ to 10⁻⁶, and must be visible at a 10⁻³ dilution. If it is not, the probe is not sufficiently biotinylated.

3. Float membrane in a small volume of AP 7.5 buffer for 1 min to rehydrate. Place membrane and 10 ml blocking buffer in a sealable bag (cut to size). Avoid trapping air bubbles when sealing. Block for 1 hr at 37°C.
4. Dilute 10 µl streptavidin–AP conjugate with 10 ml AP 7.5 buffer (1 µg/ml final). Cut corner of bag and squeeze out blocking buffer. Replace with the streptavidin–AP solution and reseal. Incubate 10 min at room temperature with agitation on a platform shaker.

Do not allow the membrane to dry out at any stage.

5. Remove membrane from bag and transfer to a shallow dish. Wash in 200 ml AP 7.5 buffer (twice, 15 min each time) and in 200 ml AP 9.5 buffer (once, 10 min) with gentle agitation.
6. Add 33 µl of 75 mg/ml NBT to 7.5 ml AP 9.5 buffer and invert to mix. Add 25 µl of 50 mg/ml BCIP and mix gently.

Be sure to add reagents in this order and mix gently to prevent precipitation of the reagents.

7. Incubate membrane with NBT/BCIP solution in a shallow dish in low light, checking periodically until color development is satisfactory (usually 15 to 60 min).
8. Stop reaction by washing with TE buffer, pH 8.0. Check incorporation of biotinylated dUTP by comparing the intensities of standard and test DNA. If the probe is at least half as intense as the standard DNA at the corresponding dilution, it should be suitable as an in situ hybridization probe.

PREPARATION AND DETECTION OF DIGOXIGENIN-LABELED DNA PROBES

ALTERNATE PROTOCOL

The digoxigenin-based detection system is an alternative nonisotopic labeling method offered by Boehringer Mannheim. Detection is achieved by incubation with antidigoxigenin antibodies coupled directly to one of several fluorochromes or enzymes, or by indirect immunofluorescence (*UNIT 14.6*). The availability of uncoupled antibodies also permits signal-amplification protocols to be employed. Biotin- and digoxigenin-labeled probes can be visualized simultaneously using a different fluorochrome for each probe.

Digoxigenin-11-dUTP can be incorporated into DNA by either of the nick translation or random oligonucleotide–primed synthesis protocols. Boehringer Mannheim advises that nick translation incorporation is not as efficient as random-priming; however, nick translation affords greater control over the final probe size. Nick translated probes perform well with in situ hybridizations, with no apparent sensitivity problems.

The protocol is virtually identical to that for biotin-11-dUTP incorporation by nick translation. The standard 100-µl reaction is set up as described in the basic protocol, but 10 µl of a 10× digoxigenin-11-dUTP/dTTP stock solution (see reagents and solutions) is substituted for 10 µl of a 10× biotin-11-dUTP stock solution. The amount of DNase I added to the sample and the time of incubation remain the same and do not depend on the modified deoxynucleotide in the reaction. After incubation at 15°C for 2 hr, an aliquot is run on a minigel to check the probe size. When the correct size has been obtained, the reaction is stopped and the probe is separated from the unincorporated nucleotides using a G-50 spin column. It is advisable to avoid phenol extraction of digoxigenin-labeled probes.

The support protocol for colorimetric detection of biotinylated probes can be modified for use with digoxigenin-labeled probes. If incorporation is to be checked, one can use as a control either digoxigenin-labeled standard DNA (obtained from Boehringer Mannheim as part of the Genius kit) or digoxigenin-labeled DNA that has already been used

Enzymatic Manipulation of DNA and RNA

3.18.5

successfully. Anti-digoxigenin antibody conjugated to alkaline phosphatase is substituted for streptavidin–alkaline phosphatase. Boehringer Mannheim recommends a 1:5000 dilution. The other steps in the support protocol are unchanged and can be followed exactly. A protocol for chemiluminescent detection of digoxigenin-labeled probes is discussed in *UNIT 3.19*.

REAGENTS AND SOLUTIONS

Alkaline phosphatase pH 7.5 (AP 7.5) buffer

0.1 M Tris·Cl, pH 7.5
0.1 M NaCl
2 mM MgCl₂
Autoclave or filter sterilize and store at room temperature

Alkaline phosphatase pH 9.5 (AP 9.5) buffer

0.1 M Tris·Cl, pH 9.5
0.1 M NaCl
50 mM MgCl₂
Autoclave or filter sterilize and store at room temperature ≤1 yr

Biotin-11-dUTP, 0.5 mM

Prepare a 0.5 mM stock of lyophilized biotin-11-dUTP (e.g., Sigma) in 20 mM Tris·Cl, pH 7.5 (at 289 nm, the molar extinction coefficient $\epsilon = 7100$; see Table A.3D.1). Check the pH of the stock; if <7.5, add a few microliters of 1 M Tris·Cl, pH 7.5, to adjust the pH to 7.5. Store at -20°C .

Biotinylated random octamers

Obtain commercially (e.g., Millipore or New England Biolabs) or prepare on a DNA synthesizer (*UNIT 2.11*). Prepare a 5 \times stock (60 O.D./ml) in 250 mM Tris·Cl, pH 8.0 (*APPENDIX 2*), 1 M HEPES, pH 6.6, 25 mM MgCl₂, and 10 mM DTT. Store at -20°C .

Digoxigenin-11-dUTP/dTTP, 10 \times stock

0.375 mM dTTP
0.125 mM digoxigenin-11-dUTP
20 mM Tris·Cl, pH 7.8
Store at -20°C

DNA dilution buffer

0.1 $\mu\text{g}/\mu\text{l}$ sheared salmon sperm DNA prepared in 6 \times SSC (*APPENDIX 2*)
Store at 4°C (for long-term storage, aliquot and store at -20°C)

dNTP/biotin mix

Prepare a mixture of 1 mM dATP, 1 mM dCTP, 1 mM dGTP, and 0.65 mM dTTP. Add biotin-16-dUTP (Enzo Biochem) to 0.35 mM final. Store at -20°C .

SDS column buffer

10 mM Tris·Cl, pH 8.0
1 mM EDTA
0.1% (w/v) SDS
Autoclave or filter sterilize

COMMENTARY

Background Information

Nonisotopically labeled probes have become increasingly popular in recent years, with researchers using them to complement or even replace radioactive probes in techniques such as in situ hybridization (Chapter 14), or Southern (UNIT 2.9) and northern (UNIT 4.9) blotting. These probes eliminate the concerns of working with radioisotopes and accumulating hazardous radioactive waste. A further advantage of this method of generating probes is their stability and long shelf life (~2 yr).

The synthesis of biotinylated nucleotides was first described by Langer et al. (1981). Biotin, a water-soluble vitamin, is covalently attached to the C5 position of the pyrimidine ring via an allylamine linker arm. Langer and colleagues demonstrated that biotin-11-dUTP can replace dTTP as a substrate for *E. coli* DNA polymerase I, although the initial incorporation rate is lower. Furthermore, denaturation and hybridization characteristics are basically unchanged (provided that incorporation does not exceed 50 biotin molecules/kb), making biotinylated probes suitable for nonisotopic in situ hybridization. The biotinylated dNTPs used in the protocols (biotin-11 and biotin-16-dUTP; biotin-14-dATP) vary in the length of the allylamine linker arm (i.e., biotin-16-dUTP has five more CH₂ groups than biotin-11-dUTP). The linker provides sufficient distance between the nucleotide and biotin to minimize steric hindrance between the synthesized probe and the streptavidin bridge needed for detection (see below).

For many, nick translation is the method of choice for biotinylating DNA and the technique outlined in UNIT 3.5 has been adapted successfully to incorporate biotinylated nucleotides (first basic protocol; Cherif et al., 1989). The DNA need not be linearized first, biotin incorporation is routinely efficient, and the probe length can be adjusted according to experimental needs. Although DNA fragments <500 nucleotides can be nick translated, oligomers are not labeled efficiently this way. Random-priming using biotinylated octamers (second basic protocol; Feinberg and Vogelstein, 1983) is probably a more efficient method for preparing biotinylated probes, as every probe molecule labeled will have at least one biotin reporter molecule incorporated into it. In addition, Moyzis et al. (1988) describe labeling oligomers with biotinylated nucleotides using terminal deoxynucleotidyl transferase (UNIT 3.6) and sub-

sequent use of these probes for in situ hybridization.

Biotin is detected directly by avidin or streptavidin conjugated to a fluorochrome or enzyme such as alkaline phosphatase or horseradish peroxidase (Lichter et al., 1988). Avidin is a 68,000-Da glycoprotein derived from egg white and has a very high affinity for biotin ($K_d = 10^{-15}$). Streptavidin is a 60,000-Da protein from *Streptomyces avidinii* that also has a high affinity for biotin ($K_d = 10^{-15}$). Either protein may be used for biotin detection. Reaction of the alkaline phosphatase conjugated to the avidin- or streptavidin-biotin complex with enzyme substrates results in formation of colored precipitates (support protocol). Detection of biotinylated and digoxigenin-labeled probes can also be accomplished via anti-biotin (Langer-Safer et al., 1982) or anti-digoxigenin antibodies, followed by a secondary antibody conjugated to a fluorochrome or enzyme.

Although there are many references in the literature for the use of biotinylated probes in Southern blotting, most researchers still prefer autoradiographic methods, citing superior sensitivity as the primary reason. With the advent of chemiluminescent techniques (UNIT 3.19; Carlson et al., 1990; Wilchek and Bayer, 1988) the use of biotinylated probes for Southern and northern blotting is becoming more common.

For in situ hybridizations, biotinylated probes are rapidly approaching the level of sensitivity afforded by using radioactively labeled probes. Detection of probes as small as 1 kb can be accomplished within a few hours post-hybridization. This is in contrast to using probes generated by isotopic techniques, where developing the emulsion film may take anywhere from several days to several weeks.

However, there are disadvantages to using biotinylated probes. First is the relative difficulty of determining the success of a probe synthesis. Unlike radioactive probes, one cannot get a quick gauge of the success of a biotinylation reaction by using a mini-monitor or by putting a sample in a scintillation counter. Biotinylated probes require a comparative dilution series with a known biotinylated standard (support protocol) and even then, quantitation is difficult because assessing the intensities cannot be done via densitometry. Second, biotinylated probes are inefficient in detecting extremely rare mRNA transcripts, because background noise of the membrane occurs at a rate constant to that of the signal.

Critical Parameters and Troubleshooting

In the nick translation protocol, optimal DNase I concentration must be determined by titration, as the size distribution of the probe depends on the amount of DNase I added. For nonisotopic *in situ* hybridizations, the optimal probe size is between 100 and 500 nucleotides, since larger probes sometimes have trouble accessing the target DNA and tend to cause increased background. Each new stock of DNase I should be tested when it is prepared to determine the appropriate working concentration. This is best accomplished by setting up several reactions and varying the amount of DNase I added. The time of incubation should not be shortened; at least 2 hr are necessary for optimal incorporation, as the initial incorporation rate of biotinylated nucleotides is less than that of unsubstituted dNTP. Even once a DNase I stock has been titrated, different amounts of the enzyme may be needed for different samples. In general, phage and miniprep DNA require more DNase I than CsCl-purified plasmid or cosmid DNA.

For optimal incorporation, the test DNA should be as free of impurities as possible. CsCl-purified DNA is nick translated more efficiently than miniprep DNA, although the latter can certainly be labeled successfully. Miniprep DNA should be treated with RNase A (UNIT 3.13), then reextracted and ethanol precipitated. Biotinylated DNA should never be phenol extracted, as biotin causes the DNA to partition to the phenol-water interface. It is advisable to avoid phenol when working with digoxigenin-labeled probes as well.

In the random-priming protocol, it is easy to forget the DNA denaturation that precedes the reaction. Biotinylation will not occur if the DNA is not denatured. As above, the optimal probe size for *in situ* hybridizations is 100–500 nucleotides and the parameters of this protocol account for this. If the probe yield is low (see below), increase the incubation (step 3) up to 6 hr and add an additional 2.5 U Klenow enzyme to boost the reaction.

For the colorimetric detection protocol, nylon membranes must be cross-linked correctly (UNIT 3.19) to properly assess probe quality. If crosslinking is omitted, the signal will be much weaker and can account for the loss of $\leq 50\%$ of the DNA.

Anticipated Results

Nick translation of 2 μg DNA yields a biotinylated probe digested to a size range of 100 to

500 nucleotides at a final concentration of 20 ng/ μl . Approximately 50 biotin molecules/kb will be incorporated.

A random priming reaction that begins with 1 μg of starting template yields ~ 260 ng biotinylated probe. Because the random priming reaction uses biotinylated random octamers, every probe molecule that is generated contains at least 1 biotin molecule. Additional incorporation of biotinylated dUTP provides more highly biotinylated probes for increased sensitivity.

Colorimetric detection of the probe spotted on nitrocellulose using streptavidin-conjugated alkaline phosphatase should allow visualization of as little as 2 pg DNA. Note that nylon membranes can also be used in chemiluminescent detection procedures (UNIT 3.19); ≥ 100 fg of DNA can be detected in a standard dot blot assay.

Time Considerations

Setting up and incubating the nick translation reactions will take ~ 3 hr. During this time, the agarose minigel and spin columns can be prepared. Running, staining, and photographing the gel takes ~ 1 hr. If the probe is the proper size, the reaction can be terminated and loaded on the spin column and be ready to use in 15 min. Probes that are not sufficiently digested will require more time for additional incubation and gel steps. As long as EDTA and SDS are not added to the tubes, the reactions can be stored at -20°C for additional incubation at a later time. If dot blots are desired, an additional 3.5 to 4.5 hr will be required.

The random priming reaction requires ~ 2 hr to denature the DNA, incubate with the primers, and precipitate and recover the labeled product.

The detection procedure requires ~ 1.5 hr to dilute the samples and prepare the membrane, ~ 2 hr to link the streptavidin-AP conjugate to the samples on the membrane, and 15 to 60 min to develop the membrane.

Literature Cited

- Beck, S. and Koster, H. 1990. Applications of di-oxetane chemiluminescent probes to molecular biology. *Anal. Chem.* 62:2258-2270.
- Carlson, D.D., Superko, C., Mackey, J., Gaskill, M.E., and Hansen, P. 1990. Chemiluminescent detection of nucleic acid hybridization. *Focus* 12:9-12.
- Cherif, D., Bernard, O., and Berger, R. 1989. Detection of single copy genes by nonisotopic *in situ* hybridization on human chromosomes. *Hum. Genet.* 81:358-362.

- Feinberg, A.P. and Vogelstein, B. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6-13.
- Langer, P.R., Waldrop, A.A., and Ward, D.C. 1981. Enzymatic synthesis of biotinylated polynucleotides: Novel nucleic acid affinity probes. *Proc. Natl. Acad. Sci. U.S.A.* 78:6633-6637.
- Langer-Safer, P.R., Levine, M., and Ward, D.C., 1982. Immunological method for mapping genes on *Drosophila* polytene chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* 79:4381-4385.
- Lichter, P., Cremer, T., Borden, J., Manuelidis, L., and Ward, D.C. 1988. Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. *Hum. Genet.* 8:224-234.
- Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., and Wu, J.R. 1988. A highly conserved repetitive DNA sequence, (TTAGGG)_n, present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* 85:6622-6626.
- Wilchek, M. and Bayer, E.A. 1988. The avidin-biotin complex in bioanalytical applications. *Anal. Biochem.* 171:1-32.

Contributed by Ann Boyle
Miles Research Center
West Haven, Connecticut

Heather Perry-O'Keefe (random priming)
Millipore Corporation
Burlington, Massachusetts

Chemiluminescent Detection of Nonisotopic Probes

UNIT 3.19

With recent advances in techniques for detecting chemiluminescent substrates, hybridization with nonisotopic rather than radiolabeled probes is becoming more common. In the basic protocol, nylon membranes carrying transferred nucleic acids are prepared for hybridization with biotinylated probes (UNIT 3.18) by UV cross-linking. This is a critical step in the procedure and the support protocol provides a detailed description of light-source calibration. After hybridization, the target nucleic acid is detected through a series of steps that lead to an enzyme-catalyzed light reaction.

The alternate protocol describes chemiluminescent detection based upon antibody recognition of digoxigenin-labeled probes. For both biotinylated and digoxigenin-labeled probes, chemiluminescent detection is more sensitive than colorimetric detection (UNIT 3.18) and has the added advantage that the membrane can be used for multiple film exposures, then stripped and redetected with different probes.

CAUTION: Exposure to UV radiation poses a significant health hazard. Wear UV-protected goggles and shield exposed skin.

CHEMILUMINESCENT DETECTION OF BIOTINYLATED PROBES

This protocol describes the series of steps leading to the reaction of alkaline phosphatase with a chemiluminescent substrate that results in the highly sensitive detection of hybridized target DNA. Nylon membranes carrying immobilized nucleic acid are first subjected to UV cross-linking, then hybridized with biotinylated probes. Subsequent detection reactions link the biotinylated probes to alkaline phosphatase through a streptavidin bridge. The alkaline phosphatase then cleaves a chemiluminescent dioxetane substrate to give a localized light emission (Fig. 3.19.1). Many of the reagents used in the protocol are available commercially; Table 3.19.1 lists suppliers of these detection kits.

**BASIC
PROTOCOL**

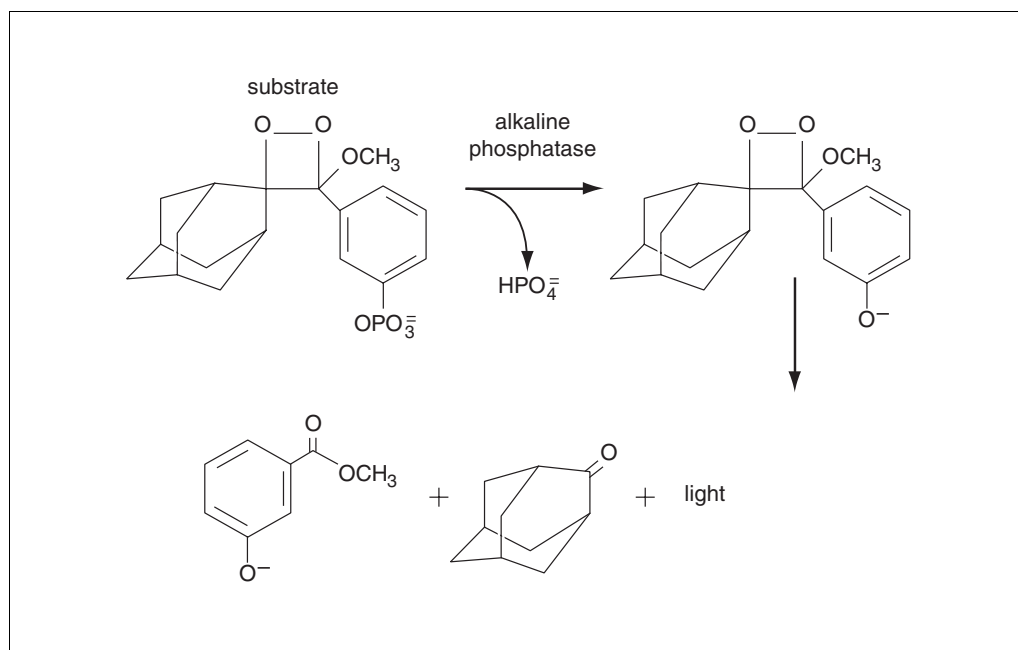


Figure 3.19.1 Chemiluminescence reaction. The enzyme alkaline phosphatase cleaves a phosphate group off the chemiluminescent substrate. The intermediate is unstable and quickly decomposes, emitting light in the process. Reprinted with permission from Millipore.

**Enzymatic
Manipulation
of DNA and RNA**

Contributed by Heather Perry-O'Keefe and Carol M. Kissinger

Current Protocols in Molecular Biology (1994) 3.19.1-3.19.8

Copyright © 2000 by John Wiley & Sons, Inc.

3.19.1

Supplement 26

Table 3.19.1 Chemiluminescent Detection Kits

Kit	Supplier ^a
AmpliProbe	ImClone Systems
ECL	Amersham
Flash	Stratagene
Gene Images	U.S. Biochemical
Genius	Boehringer Mannheim
LightSmith	Promega
PhotoGene	GIBCO/BRL
Phototope	NEB
PolarPlex	Millipore
SouthernLight	Tropix

^aSee suppliers' addresses in APPENDIX 4.

Materials

Uncharged nylon membrane blotted via neutral transfer with DNA (UNIT 2.9) or RNA (UNIT 4.9)

Biotinylated probes (UNIT 3.18)

Blocking solution

Wash buffers I and II

1 mg/ml streptavidin (see reagents and solutions)

0.38 mg/ml biotinylated alkaline phosphatase

Chemiluminescent dioxetane substrate (Table 3.19.2)

Substrate buffer, pH 9.6

Blotting paper (Whatman 3MM or equivalent)

Calibrated UV source (support protocol)

Heat-sealable hybridization bags

Additional reagents and equipment for Southern and northern hybridizations (UNITS 2.9 & 4.9)

Cross-link and hybridize the membrane

1. Fasten the corners of the blotted membrane to a dry piece of blotting paper, nucleic acid-side up, with paper clips. Place in an incubator and dry. The drying time and temperature can range from 15 to 30 min at 42° to 80°C. Alternatively, dry overnight at room temperature.

When dry, the membrane will have curled-up edges. It is very important that the membrane be completely dry to ensure even, reproducible cross-linking and successful detection.

2. Expose membrane, nucleic acid-side up, to the UV source. Cross-link for the calculated optimal time period.

Determine optimal cross-linking time as described in the support protocol. The UV source should have sufficient power at 254 nm to generate a minimum of 100 $\mu\text{W}/\text{cm}^2$. Most transilluminators have power outputs in the 500 to 1000 $\mu\text{W}/\text{cm}^2$ range. Hand-held lamps will have slightly lower outputs.

3. Hybridize membrane with biotinylated probes.

Both Southern (UNIT 2.9) and northern (UNIT 4.9) blots may be hybridized to biotinylated probes. Nitrocellulose membranes will quench the chemiluminescence reactions and must not be used.

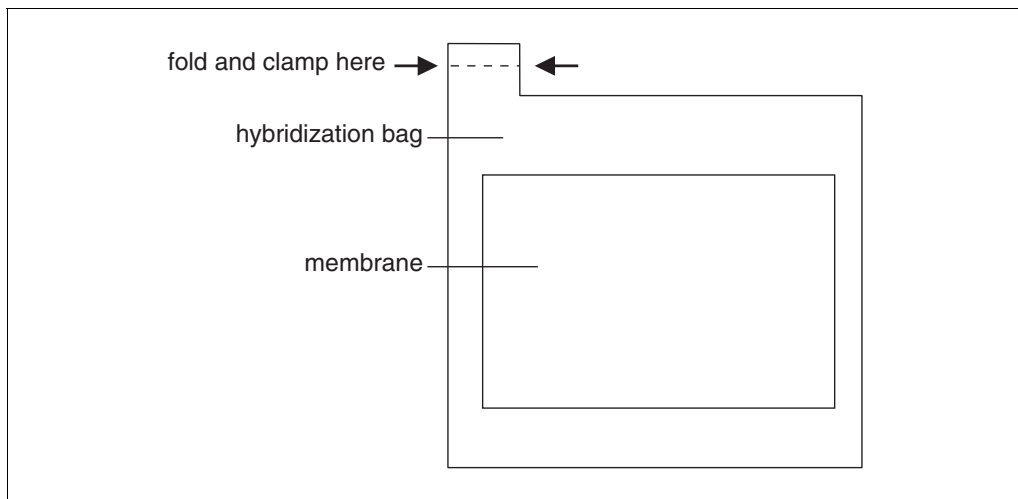


Figure 3.19.2 Hybridization bag spout. After a buffer has been added to the bag, the spout is folded over and held closed with a bulldog clamp or several paper clips. It is important to leave sufficient room for the membrane to float free in solution and the bag should be free of trapped air to ensure maximum contact between the reagents and the membrane.

Detect the hybridized nucleic acids

The amount of wash and detection solutions needed during the procedure depend on the size of the membrane being used. Calculate the volume of the membrane as 1 vol (V) = membrane area ($\text{cm}^2 \times 0.05 \text{ ml/cm}^2$). For example, a 10×10 -cm membrane would have $V = 100 \text{ cm}^2 \times 0.05 \text{ ml/cm}^2 = 5 \text{ ml}$. All subsequent reaction volumes are based on this calculation.

4. Following hybridization, place membrane in a hybridization bag.

Be careful not to touch the membrane with bare hands or gloves that contain powder, as both oil and powder residues will contribute to nonspecific background on the membrane.

5. Seal bag, leaving a small area in one corner that can be made into a spout through which buffers can be easily added and removed (Fig. 3.19.2).
6. Carefully add 1 vol blocking solution to hybridization bag. Incubate 1 min at room temperature with moderate shaking. Drain and discard solution.
7. Add 1 mg/ml streptavidin to 1 vol blocking solution such that its final concentration is $1 \mu\text{g/ml}$ and add to hybridization bag. Incubate 4 min at room temperature with moderate shaking. Drain and discard solution.
8. Add 10 vol of wash buffer I to hybridization bag. Wash membrane 4 min at room temperature with moderate shaking. Drain and discard solution. Repeat once and discard solution.
9. Add biotinylated alkaline phosphatase to 1 vol blocking solution such that its final concentration is $0.5 \mu\text{g/ml}$ and add to hybridization bag. Incubate 4 min at room temperature with moderate shaking. Drain and discard solution.
10. Wash membrane twice as in step 9, using 10 vol of wash buffer II.
11. Dilute chemiluminescent dioxetane substrate to $1 \times$ final concentration with substrate buffer. Add 0.5 vol of substrate to hybridization bag. Incubate 4 min at room temperature with moderate shaking. Open bag and drain as thoroughly as possible.

The chemiluminescent substrate should be brought to room temperature before use.

Table 3.19.2 Chemiluminescent Substrates for Detection of Nonisotopic Probes^a

Dioxetane substrate	Buffer	Source ^c
Lumigen-PPD ^b (0.33 mM)	2-amino-2-methyl-1-propanol (pH 9.6)/0.88 mM MgCl ₂ /750 mM CTAB/1.13 mM fluorescein surfactant	BM, GB, LU, MI, NEB
Lumi-Phos 530 ^b (0.33 mM)	2-amino-2-methyl-1-propanol (pH 9.6)/0.88 mM MgCl ₂ /750 mM CTAB/1.13 mM fluorescein surfactant	BM, GB, LU, MI, NEB
AMPPD (0.25 mM)	1 mM DEA/1 mM MgCl ₂ , pH 10	TR
CSPD (0.25 mM)	1 mM DEA/1 mM MgCl ₂ , pH 10	TR

^aAbbreviations: AMPPD, disodium 3-(4-methoxyspiro{1,2-dioxetane-3,2-tricyclo[3.3.1.1^{3,7}] decan}-4-yl)phenyl phosphate; CSPD, AMPPD with substituted chlorine group on adamantane chain; CTAB, cetyltrimethylammonium bromide; DEA, diethanolamine; Lumigen-PPD and Lumi-Phos 530: 4-methoxy-4-(3-phosphate phenyl)-spiro-(1,2,-dioxetane-3,2-adamantane), disodium salt.

^bLumi-Phos 530 has a fluorescence enhancer; Lumigen-PPD does not.

^cAbbreviations: BM, Boehringer Mannheim; GB, GIBCO/BRL; LU, Lumigen; MI, Millipore; NEB, New England Biolabs; TR, Tropic. Addresses and phone numbers of suppliers are provided in APPENDIX 4.

Observe chemiluminescent reaction

- Smooth out any wrinkles in bag and reseal. Place in a film cassette with the nucleic acid side of the membrane facing up and put X-ray film on top.

It is important that there be uniform contact between the bag and X-ray film to insure sharp images.

- Close cassette and expose 10 to 20 min.

The chemiluminescent reaction starts immediately and reaches a plateau after several hours, remaining constant for days. If necessary, adjust exposure times to optimize the signal.

ALTERNATE PROTOCOL

CHEMILUMINESCENT DETECTION OF DIGOXIGENIN-LABELED PROBES

An alternative to the biotinylated streptavidin–alkaline phosphatase method of detecting blotted nucleic acids is a system based on antibody detection—the digoxigenin system. It utilizes an anti-digoxigenin antibody conjugated to alkaline phosphatase which recognizes digoxigenin-dUTP that has been incorporated into the probe (UNIT 3.18). The detection procedure is similar to that for the biotinylated probes (basic protocol) except that longer blocking steps are required for the antibody detection.

A positively charged nylon membrane blotted with DNA (UNIT 2.9) or RNA (UNIT 4.9) is hybridized to a digoxigenin-labeled probe (UNIT 3.18). Chemiluminescent detection is based on the use of the Genius kit (Boehringer-Mannheim; see Table 3.19.1), the only kit that works with digoxigenin-labeled probes. Follow manufacturer's instructions for blocking, antibody-binding, and developing steps. The membrane is exposed to X-ray film for ~20 min. The enzymatic reaction reaches a steady state after 3 to 5 hr at room temperature. If necessary, adjust exposure times to optimize the signal. The membrane can be used for multiple exposures, and can be stripped and redetected with different probes.

CALIBRATING AN ULTRAVIOLET LIGHT SOURCE

SUPPORT PROTOCOL

Additional Materials

UV source: transilluminator, hand-held UV lamp, homemade box with germicidal bulbs, or cross-linker (e.g., Stratalinker, Stratagene #400071)
Radiometer and radiometer sensor

UV cross-linking of nucleic acids to nylon membrane is a critical step that significantly affects the quality of the data. Inadequate cross-linking will lead to insufficient binding and the nucleic acids will wash off the membrane. Excessive cross-linking renders the nucleic acids inaccessible to the probe. Therefore, calibration of the UV source is important.

UV sources are calibrated with a radiometer, which determines the power output in $\mu\text{W}/\text{cm}^2$. Radiation safety offices at academic research institutions can often provide radiometers for calibration. The UV light must be turned on and the bulbs warmed up for 1 to 2 min before calibration. The calibration procedure itself takes only a few minutes and should be repeated as the bulbs age. Transilluminator power output is calibrated at the glass surface and hand-held lamp output is calibrated at the distance at which the entire membrane will be evenly exposed to the UV light. The two Stratalinker instruments—models 1800 and 2400—have power outputs of 3000 and 4000 $\mu\text{W}/\text{cm}^2$, respectively.

Once the power output is measured, the optimal exposure time for the light source can be calculated as follows:

$$T = \frac{(33,000 \mu\text{W}/\text{cm}^2) (1 \text{ sec})}{P \mu\text{W}/\text{cm}^2}$$

where T is the exposure time in seconds and P is the power output of the UV source in $\mu\text{W}/\text{cm}^2$. For example, if the radiometer indicates a power output of 752 $\mu\text{W}/\text{cm}^2$, the time required for cross-linking with that transilluminator is:

$$T = \frac{(33,000 \mu\text{W}/\text{cm}^2) (1 \text{ sec})}{752 \mu\text{W}/\text{cm}^2} = 44 \text{ sec}$$

This determination can also be done empirically. Transfer duplicate lanes of nucleic acid to a membrane and expose different portions to UV light for increasing amounts of time varying by 1-min increments. Remember to cover previously exposed portions of the membrane with aluminum foil. There should be a marked difference in the amount of detectable and hybridizable nucleic acids.

REAGENTS AND SOLUTIONS

Biotinylated Alkaline Phosphatase

0.38 mg/ml biotin conjugate
3M NaCl
1 mM MgCl_2
0.1 mM ZnCl_2
30 mM triethanolamine acetate (TEA), pH 7.5
Store at 4°C

Blocking Solution

5% w/v SDS

17 mM Na₂HPO₄

8 mM NaH₂PO₄

Store at room temperature

Filter sterilize with a 0.45- μ m filter if necessary

Streptavidin, 1 mg/ml

Prepare a 1 mg/ml solution by reconstituting lyophilized streptavidin powder in 0.01 M sodium phosphate, pH 7.2/0.15 M NaCl/0.05% (v/v) NaN₃.

CAUTION: NaN₃ is hazardous. Handle very carefully.

To make a 0.01 M sodium phosphate (pH 7.4) solution, combine 68.4 ml of 0.1 M Na₂HPO₄ and 31.6 ml of 0.1 M NaH₂PO₄ and dilute 1:10 with H₂O.

Wash buffer I

Dilute blocking solution 1:10 with water just prior to use

Wash buffer II, 10 \times stock

100 mM Tris·Cl, pH 9.5

100 mM NaCl

10 mM MgCl₂

Prepare stock and store at 4°C

Dilute 1:10 with water just prior to use

COMMENTARY

Background Information

Chemiluminescent detection systems have proven to be particularly adaptable to current molecular biology techniques. The reaction is based on the use of enzyme-cleavable dioxetanes, first described by Schaap et al. (1987), and can be carried out at ambient temperature, in aqueous buffer, and at a moderate alkaline pH (for review, see Beck and Koster, 1990). Chemiluminescent detection can be used in Southern and northern hybridizations (UNITS 2.9 & 4.9); the UV cross-linking and membrane detection described here are simple variations on the traditional method of visualizing nucleic acids with radioactive probes. This technique has also been applied to DNA sequencing (Creasey et al., 1991) with the development of sequencing kits for this purpose (Uniplex and Multiplex sequencing kits, Millipore). Phage and colony hybridizations (UNITS 6.3 & 6.4) and immunoblotting and immunodetection (UNIT 10.8) also lend themselves to this technology and kits are available for these procedures as well (Amersham and U.S. Biochemicals).

The chemiluminescent substrate used in these protocols is typically a 1,2-dioxetane substituted with adamantyl and phosphate-protected aryl moieties (Fig. 3.19.1). The large adamantyl moiety imparts stability because of steric hindrance. To activate chemilumines-

cence, the following reactions are carried out. Streptavidin, which has several binding sites for biotin, is applied to the membrane and binds to the biotin incorporated into the probe. After washing, biotinylated alkaline phosphatase is added to the membrane and in turn binds the streptavidin. The streptavidin thus serves as a bridge linking alkaline phosphatase to the hybridized probe. Addition of chemiluminescent dioxetane substrates to alkaline phosphatase initiates cleavage of the phosphate group. The resulting 1,2-dioxetane anion decomposes to produce light (Fig. 3.19.1).

In addition to the obvious safety and disposal advantages, chemiluminescence affords several other benefits. The probes are stable and large amounts of probe can be synthesized at one time and stored, because only small amounts are used in an experiment. In addition, hybridization buffer containing the probe can be reused several times before the signal becomes weak. Another advantage is the sensitivity of the technique—chemiluminescent detection is 5×10^3 times more sensitive than colorimetric detection (Beck and Koster, 1990). Short exposure times are also advantageous—in many cases a 10- to 20-min exposure is all that is required. Moreover, the chemiluminescent reaction is sustained for many days; thus, one blot can easily be exposed many times. Furthermore, because the reaction

Table 3.19.3 Guide to Troubleshooting Chemiluminescent Detection Techniques

Problem	Possible cause	Solution
Low or insufficient signal	Insufficient biotin labeling—check by comparing sample and control dilution series	If probe is only visible at 10^{-1} and 10^{-2} dilutions, repeat the biotinylation reaction. If visible at dilutions of 10^{-3} or greater, consider other possible causes.
	Inadequate cross-linking	Make sure the membrane is completely dry before cross-linking. Calibrate UV source with a radiometer.
	Incorrect pH of the chemiluminescent substrate	Check pH of substrate buffer and adjust to pH 9.6.
Uniform or uneven high background	Inadequate washing	Increase number of washes, ensuring the membrane floats freely in bag during detection steps.
	Contaminated wash solutions	Mold can grow in the Wash I solution, leading to spotting on the membrane. Filter through a 0.45- μ m filter.
	Unbuffered blocking and Wash I solutions	Make sure that SDS solutions are made up in phosphate buffer as described.

produces light, photographic paper can be placed directly on the blot, eliminating the need to photograph data captured on X-ray film. However, photographic paper is significantly slower than X-ray film, and the exposure time must be increased. Finally, as for blots probed with radioactivity, those probed with chemiluminescence can be stripped and reprobed.

Critical Parameters

Several parameters are critical to the detection of chemiluminescent probes. Most important of these is the cross-linking of nucleic acids to the membrane. The membrane should be completely dry in order for cross-linking to be uniform and reproducible. The UV source should be calibrated either quantitatively or qualitatively (support protocol). If the nucleic acids are inadequately cross-linked, they will wash off the membrane. If the cross-linking is excessive, the nucleic acids will become inaccessible to the probe.

For detection, it is important that the pH of the substrate buffer be near pH 9.6 (a range of pH 9.5 to 9.7 is acceptable) to ensure optimal steady light emission. If the pH is too low, the reaction will be quenched and no light will be given off. If the pH is too high, the reaction will occur instantly and a flash of light will be given off. Nonfat dry milk, a popular component of

blocking solutions, cannot be used in this procedure because it contains biotin and will cause intense background problems.

It is important to adhere to the length of time detailed, as well as concentration of reagents needed, for both the streptavidin and the biotinylated alkaline phosphatase incubation steps. If either of these parameters is significantly altered, the signal:noise ratio will drastically change.

In the detection procedure, the minimum number of washes needed are detailed. If necessary, the wash times may be lengthened and the number of washes can be increased to eliminate background signal. The wash with buffer II directly precedes addition of the chemiluminescent substrate and is perhaps the most critical because it removes residual SDS and establishes the correct pH of the membrane. An attempt should be made to thoroughly drain the hybridization bag prior to adding the chemiluminescent reagent.

Troubleshooting

Low or insufficient signal and high or uneven background are two commonly encountered detection problems. Possible causes and solutions for these problems are described in Table 3.19.3.

Anticipated Results

Chemiluminescent visualization of Southern blots has proven to be sensitive enough to detect a single-copy gene from as little as 1 μ g of human genomic DNA. When applied to northern blots, this approach can detect abundant to mildly rare transcripts. At this time chemiluminescent detection is not recommended for the detection of very rare mRNA transcripts.

Time Considerations

Successful detection reactions can be done in ~30 min. Increasing in the duration of any of the wash steps will have little or no ill effect on the final results, but will obviously increase the total length of the protocol.

Literature Cited

Beck, S. and Koster, H. 1990. Applications of dioxetane chemiluminescent probes to molecular biology. *Anal. Chem.* 62:2258-2270.

Creasey, A., D'Angio, L., Dunne, T.S., Kissinger, C., O'Keefe, T., Perry-O'Keefe, H., Moran, L.S., Roskey, M., Schildkraut, I., Sears, L.E., and Slatko, B. 1991. Application of a novel chemiluminescence-based DNA detection method to single vector and multiplex DNA sequencing. *BioTechniques* 11:102-109.

Schaap, A.P., Handley, R.S., and Giri, B.P. 1987. Chemical and enzymatic triggering of 1,2-dioxetanes. 1:Aryl esterase-catalyzed chemiluminescence from a naphthyl-acetate substituted dioxetane. *Tetrahedron Lett.* 28:935-938.

Key Reference

Beck, S. and Koster, H. 1990. See above.

Reviews the chemistry behind the techniques described and discusses the applications for this methodology.

Heather Perry-O'Keefe and
Carol M. Kissinger
Millipore Corporation
Burlington, Massachusetts

Recombinational Cloning

Large-scale experiments requiring protein expression from thousands of genes require an efficient method for cloning the genes into protein expression vectors. Traditional cloning methods based on restriction enzyme digestion and ligation are not practical. Recombinational cloning, by contrast, allows high-throughput cloning of genes into protein expression vectors. Recombinational cloning is based on site-specific recombination for transferring DNA sequences into multiple expression systems. It can be used to transfer a single insert into multiple different expression systems (Fig. 3.20.1), or to simultaneously transfer many different DNA sequences into a single expression vector.

The methods described in this unit are designed for use with two commercially available recombinational cloning systems, Gateway (a registered trademark of Invitrogen) and Creator (a registered trademark of BD Clontech). Basic Protocol 1 describes the amplification of target genes and addition of the required recombination sites by PCR, Basic Protocols 2 and 3 describe generation of master clones, and Basic Protocols 4 and 5 address generation of expression clones.

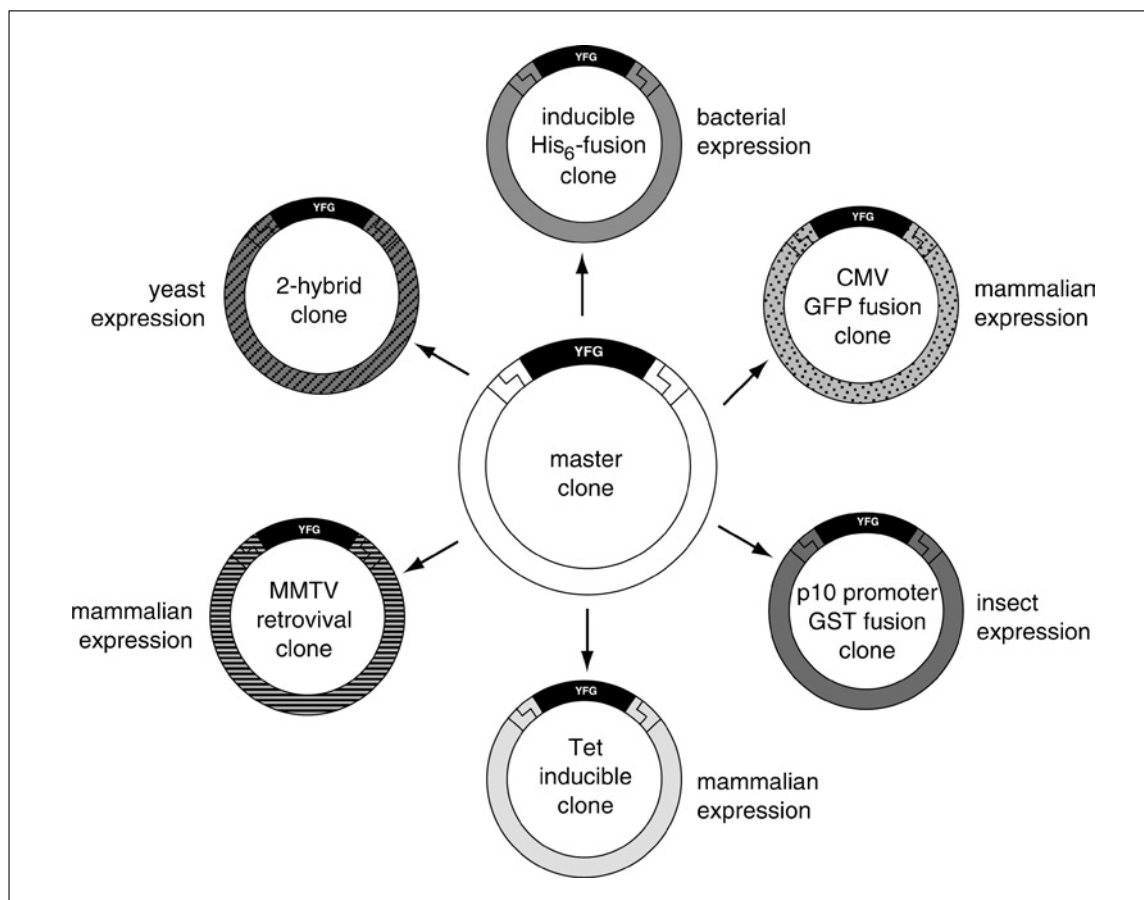


Figure 3.20.1 Advantages of universal cloning technology. Recombination-based universal cloning technology enables efficient parallel transfer of your favorite gene (YFG) from a master clone into various different expression systems for protein production and functional analysis.

AMPLIFICATION OF TARGET GENES BY PCR

The sequence of PCR primer used to amplify a gene sequence is composed of a gene-specific portion along with a common tail sequence for adding flanking recombination sites. If the required flanking recombination sequences are long, as is often the case in the Gateway scheme, the primers may acquire errors during their synthesis. In this case, the gene-specific primers can be designed to contain a tail sequence including only part of the recombination site. The rest of the recombination site can then be added efficiently by a secondary PCR employing a universal primer set (Gateway scheme; Fig. 3.20.2). It is best to use just one PCR step to build a recombination-competent fragment, which is usually possible with the Creator cloning scheme (Fig. 3.20.2).

All primers must have a similar melting temperature and must be anchored in start or stop positions of the open reading frame. The gene-specific 5' primer must start with ATG, and the 3' primer must start at the end of the common stop codon. On a practical level, such a primer is designed using bases from the relevant end (5' or 3') of the target sequence and adding additional bases until the desired melting temperature (often 60°C) is reached. The recombination sequence can be appended once the design of the gene-specific portion is complete (Fig. 3.20.3). The primers place the coding sequences in-frame with the recombination sites. The Kozak sequence is added in front of ATG in all 5' oligonucleotides for better protein expression; the common stop codon, in the case of the 3'-open oligonucleotide, can be changed to encode leucine to generate C-terminal fusion proteins.

Materials

- 10× *Pfx* amplification buffer (Invitrogen)
- 10× *PCRx* enhancer (Invitrogen)
- 10 mM 4dNTPs (Invitrogen)
- 50 mM MgSO₄

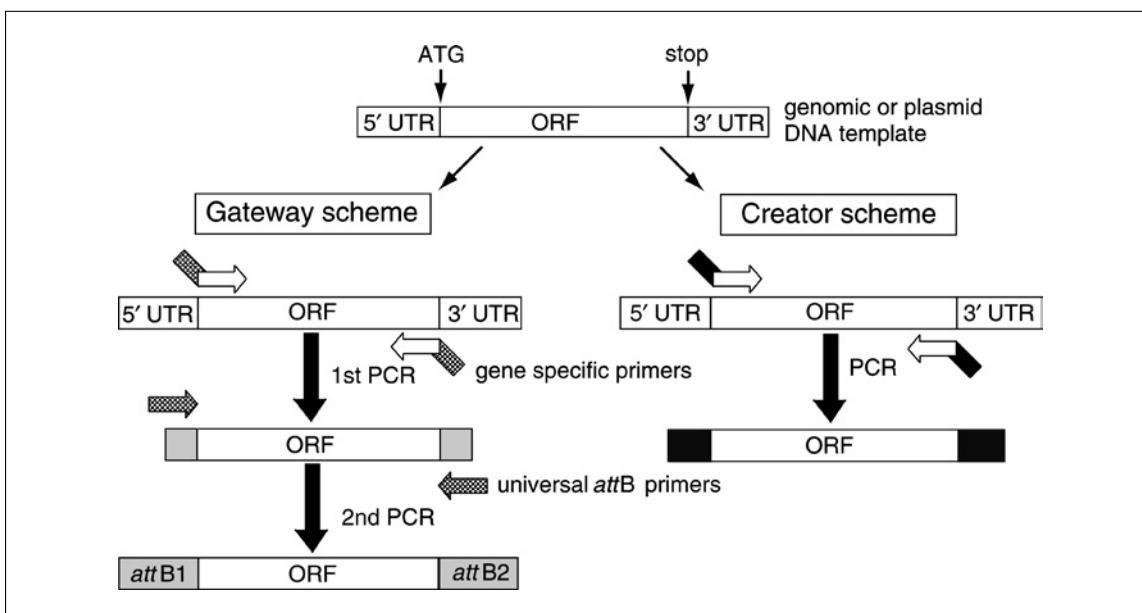


Figure 3.20.2 Schematic view of primer design and PCR amplification. Only the open reading frame (ORF) will be amplified from ATG to the stop codon, using gene-specific primers. Note that in the Gateway scheme the gene of interest will be amplified first with a gene-specific primer set (containing partial *attB* sequences) and then reamplified with a universal primer set. In the Creator scheme only one primer set is used.

Gateway scheme		
gene specific	5' oligo	5'- <u>AAAGCAGGCTCCACC</u> ^{Kozak} <i>atg[18-25 gene-specific nucleotides]</i> -3'
	3' closed oligo	5'- <u>ACAAGAAAGCTGGGTC</u> <i>cta[18-25 gene-specific nucleotides]</i> -3' stop anticodon
	3' open oligo	5'- <u>ACAAGAAAGCTGGGTC</u> <i>caa[18-25 gene-specific nucleotides]</i> -3' Leu anticodon
universal	5' oligo	5'-GGGGACAAGTTTGTACAAA <u>AAAGCAGGCTCCACC</u> <i>atg</i> -3'
	3' oligo	5'-GGGGACCACTTTGTACA <u>AAAGCTGGGT</u> -3'
Creator scheme		
	5' oligo	5'- <u>GAAGTTATCAGTCGACACC</u> ^{Kozak} <i>atg[18-25 gene-specific nucleotides]</i> -3'
	3' closed oligo	5'- <u>ATGGTCTAGAAAGCTTCC</u> <i>cta[18-25 gene-specific nucleotides]</i> -3' stop anticodon
	3' open oligo	5'- <u>ATGGTCTAGAAAGCTTCC</u> <i>caa[18-25 gene-specific nucleotides]</i> -3' Leu anticodon

Figure 3.20.3 Sequence information for primers. Kozak sequences (CACC) were placed just in front of ATG sequences in both schemes. In the Gateway scheme, the partial *att* sequences are underlined in the case of gene-specific primers. Universal primers contain complete *att* sequences and the matching sequences to gene-specific primers are underlined. In the Creator scheme, the homologous sequences to pDNR-Dual vector are underlined. To obtain C-terminal fusion proteins in the case of open primers in both schemes, the common stop anticodon (CTA) is changed to the leucine anticodon (CAA).

2.5 U/μl *Pfx* DNA polymerase (Invitrogen)

DNA template: 10 ng/μl first-strand cDNA mix (1:4 brain/placenta) or 20 ng/μl plasmid DNA

830 nM gene specific 5'-oligonucleotide primer in sterile H₂O (for primer design see Basic Protocol 1 introduction)

830 nM gene specific 3'-oligonucleotide primer in sterile H₂O (for primer design see Basic Protocol 1 introduction)

200 μM universal 5'-oligonucleotide primer (BD Clontech; dilute from stock with sterile H₂O and store up to 6 months at -20°C)

200 μM universal 3'-oligonucleotide primer (BD Clontech; dilute from stock with sterile H₂O and store up to 6 months at -20°C)

PCR tubes (thin-walled) or 96-well plates

Multichannel pipettor (optional)

Thermal cycler

Perform first-round PCR

1. Prepare the master mix for the first-round 50-μl PCR (multiply the volume for each component by the number of reactions plus two or three to allow for losses during transfer) as follows (38 μl per tube or well):

5 μl 10× *Pfx* amplification buffer (1× final)

5 μl 10× PCR_x enhancer (1× final)

1.5 μl 10 mM 4dNTPs (0.3 mM final)

1 μl 50 mM MgSO₄ (1.0 mM final)

0.4 μl 2.5 U/ml *Pfx* DNA polymerase (1 U final)

10 μl DNA template (100 or 200 ng final)

15.1 μl PCR-quality H₂O.

For cloning human genes, a first-strand cDNA mix of brain plus placenta (1:4) works well as template. First-strand cDNA, 100 ng per reaction, can be added to the master mix. To clone genes from specific tissue sources or other organisms, genomic DNA or an appropriate first-strand cDNA sample can be used (see UNIT 5.5 & 15.5 for making first-strand cDNA).

Prepare plasmid DNA using any commercially available miniprep kit or standard miniprep method (e.g., UNIT 1.6), but note that the DNA quantitation (APPENDIX 3D) is important. A total of 200 ng of plasmid DNA should be sufficient.

2. Dispense 38 μ l master mix into PCR tubes for individual reactions or into each well of a 96-well plate for a large-scale format.

For 96-well plate format a multichannel pipettor should be used.

3. Add 6 μ l 830 nM gene-specific 5'-oligonucleotide primer and 6 μ l 830 nM gene-specific 3'-oligonucleotide primer to each reaction (final concentration 100 nM each primer).

When primer stocks are delivered at 200 μ M, they may be diluted to 830 nM in two steps. (1) Add 5 μ l primer stock to 115 μ l PCR quality water and mix well. (2) Add 20 μ l diluted primer to 180 μ l PCR-quality water and mix well.

- 4a. For amplification with first-strand cDNA as a template: Carry out PCR using the following amplification cycles:

Initial step:	2 min	94°C	(denaturation)
20 cycles:	15 sec	94°C	(denaturation)
	1 min	52°C	(annealing)
	4 to 6 min	68°C	(extension)
Final step:	12 min	68°C	(polishing).

- 4b. For amplification with plasmid DNA as a template: Carry out PCR using the following amplification cycles:

Initial step:	2 min	94°C	(denaturation)
6 cycles:	15 sec	94°C	(denaturation)
	1 min	52°C	(annealing)
	4 to 6 min	68°C	(extension)
Final step:	12 min	68°C	(polishing).

For target ORF sequences <2 kb, an extension time of 4 to 6 min is enough. If the target is larger (2 to 4 kb), extension times of up to 8 to min are recommended. The extension time for larger genes (>4 kb) should be determined experimentally.

For the Creator cloning scheme, as little as a total of 15 cycles for first-strand cDNA or 6 cycles for plasmid DNA have been shown to amplify enough PCR fragments for use with In-Fusion enzyme reactions in creating master clones (unpub. observ). Amplification with a second primer set is not needed. Proceed to step 9.

Perform second-round PCR (Gateway system only)

5. Prepare the master mix for the second-round 50- μ l PCR (multiply the volume for each component by the number of reactions plus two or three to allow for losses during transfer) as follows (30 μ l per tube or well):

0.125 μ l 200 μ M universal 5' primer (0.5 μ M final)

0.125 μ l 200 μ M universal 3' primer (0.5 μ M final)

3 μ l 10 \times Pfx amplification buffer (1 \times final)

3 μl 10 \times PCR \times enhancer (1 \times final)
1.5 μl 10 mM 4dNTPs (0.3 mM final)
0.6 μl 50 mM MgSO₄ (1.0 mM final)
0.4 μl 2.5 U/ μl Pfx DNA polymerase (1 U final)
21.25 μl PCR-quality H₂O.

6. Aliquot 30 μl of the master mix into fresh PCR tubes or 96-well plates.
7. Add 20 μl amplified DNA from step 4a or 4b.
8. Carry out PCR using the following amplification cycles:

Initial step:	2 min	94°C	(denaturation)
6 cycles:	15 sec	94°C	(denaturation)
	1 min	52°C	(annealing)
	4 to 6 min	68°C	(extension)
Final step:	12 min	68°C	(polishing).

9. Store amplified DNA up to 6 months at -20°C . Proceed to purification (see Support Protocol) or preparation of the master clone (see Basic Protocol 2 or 3).

RAPID GEL PURIFICATION OF PCR PRODUCTS

Amplification of a gene of interest from known plasmid DNA template using PCR is almost 100% efficient in generating the expected PCR fragment; however, it has been observed that several additional minor products can be amplified from the first-strand cDNA mix. Small products, especially in the Gateway scheme, will recombine with higher efficiency than the larger expected product, resulting in vectors containing aberrant sequences instead of the desired product. Moreover, PCR reaction mixtures may contain unused recombination site primers that could later interfere with the site-specific recombination reaction. When smaller, nonspecific PCR products are present in the reaction, it is advisable to purify the desired PCR products using agarose gel electrophoresis to avoid cloning contaminants or undesirable PCR products.

Materials

PCR samples (Basic Protocol 1)
GelStar nucleic acid stain (Cambrex; <http://www.cambrex.com>)
Razor blades *or* 15-ml syringe with attached intramedic polyethylene tubing (Becton Dickinson PE330) or comparable
Low-frequency UV light table
Eye protection for UV light
Filter tubes (Ultrafree-DA) and collection tubes (Millipore) *or* 96-well filter plates (Unifilter 350 GF/D plate) and collection plates (Millipore or equivalent)
Benchtop centrifuge (adaptor for 96-well plates, optional)
Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5A) and determination of DNA concentration (APPENDIX 3D)

1. Prepare a 1% agarose gel containing 0.001% (v/v) GelStar nucleic acid stain, load PCR samples, and run the gel (see UNIT 2.5A).

In large cloning projects, a saw-tooth well arrangement will reduce the chance of contamination from neighboring bands (see Critical Parameters and Troubleshooting).

SUPPORT PROTOCOL

Enzymatic Manipulation of DNA and RNA

3.20.5

- After separation, excise bands using a razor blade or a 15-ml syringe with intramedic polyethylene tubing, while visualizing the bands with GelStar nucleic acid stain on a low-frequency UV light box. Collect the gel fragments into filter tubes for individual clones or into wells of a 96-well filter plate for larger-scale experiments.

Stiff tubing mounted on a 15-ml syringe allows quick excision of the bands and rapid ejection (under air pressure) into the filter plate. After ejecting the plugged band, cut off the tip of the tubing to create a clean end for the next band extraction.

- Place the filter tube into a new collection tube. For large-scale experiments, place the filter plate on top of a matching empty 96-well collection plate.
- Freeze the plate at -20°C for 1 hr, then thaw quickly at 37°C .

This step is optional because most PCR fragments <2 kb can be eluted easily without it. This step helps for fragments >2 kb.

- Centrifuge 15 min at $3000 \times g$, room temperature.

The eluate contains the DNA fragment, which can be verified by examining over a UV box.

- Discard the top filter tube or plate. Store the bottom tube or plate up to 6 months at -20°C . Determine DNA concentration (see APPENDIX 3D)

BASIC PROTOCOL 2

CAPTURE OF ORFS TO MAKE MASTER CLONES FOR THE GATEWAY SYSTEM: BP REACTION

The amplified PCR fragment of a gene of interest must first be transferred into a master vector to generate a master clone. The transfer process requires recombination between the PCR fragment and the master vector. This protocol describes the BP recombination reaction, so called because it mediates transfer of a gene of interest from an *attB*-PCR fragment or *attB*-expression clone to an *attP*-containing master vector to create a master clone (Fig. 3.20.4). For an efficient BP reaction, the *attB* substrate should be in linear form (PCR fragment or restriction-digested linearized plasmid), and the *attP*-containing master vector should be supercoiled. Because the master vector has the *ccdB* gene, which is toxic to common *E. coli* strains, the *ccdB*-resistant *E. coli* strain DB3.1 is used for propagation of any *ccdB*-containing vectors. M13 forward and reverse primers can be used for sequence verification of master clone DNA.

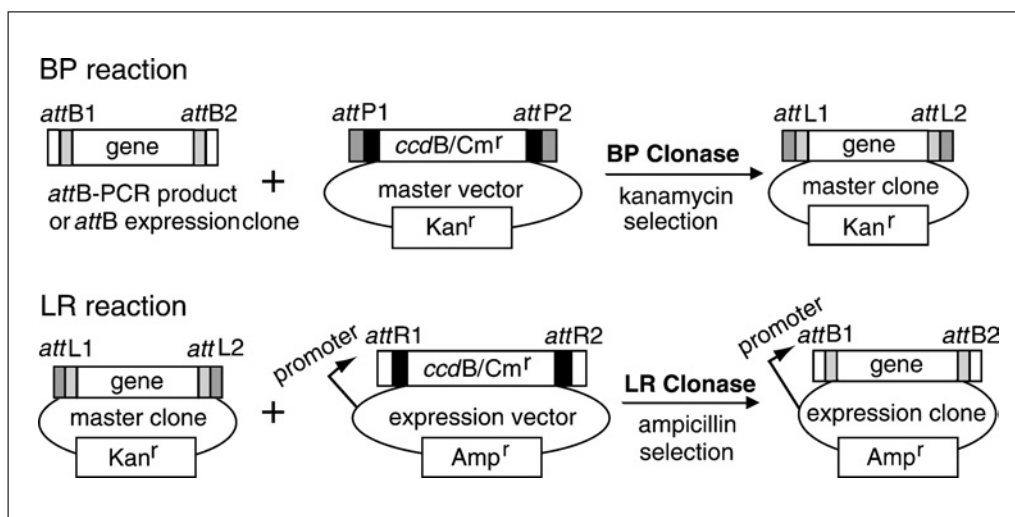


Figure 3.20.4 Gateway BP and LR reactions. BP and LR Clonase (registered trademarks of Invitrogen) facilitate the recombination between *attB* and *attP* and between *attL* and *attR*, generating *attL* and *attB* sequences, respectively.

Materials

- 150 ng/μl master vector (e.g., pDONR221; Invitrogen)
- 5× BP Clonase buffer (Invitrogen)
- BP Clonase enzyme mix (Invitrogen)
- 10 to 20 ng/μl *attB*-PCR product: prepared according to the Gateway scheme (Basic Protocol 1) and purified (Support Protocol)
- 2 μg/μl proteinase K (Invitrogen)
- LB plates supplemented with 50 μg/ml kanamycin (or other appropriate antibiotic; *UNIT 1.1*)
- LB medium supplemented with 50 μg/ml kanamycin (or other appropriate antibiotic; *UNIT 1.1*)
- 50% (v/v) glycerol, sterile
- Plasmid DNA isolation kit (e.g., Qiagen; optional)
- 96-well plate or 1.5-ml microcentrifuge tubes
- Multichannel pipettor
- 25°C incubator (optional)
- Additional reagents and equipment for growth of *E. coli* competent cells (*UNIT 1.8*), DNA miniprep (*UNIT 1.6*), and DNA sequencing (Chapter 7)

1. Prepare the master reaction mix on ice for the 10-μl BP reaction (multiply the volume for each component by the number of reactions plus two or three to allow for losses during transfer) as follows (5 μl per tube or well):

- 1 μl 150 ng/μl pDONR221 (15 ng/μl final)
- 2 μl 5× BP Clonase buffer (1× final)
- 2 μl BP Clonase enzyme mix.

2. Dispense 5 μl master reaction mix into 1.5-ml microcentrifuge tubes or each well of a 96-well plate.

For the 96-well plate format, a multichannel pipettor should be used.

3. Add 5 μl 10 to 20 ng/μl PCR product (50 to 100 ng).

3. Incubate 1 hr at 25°C.

Although, the manufacturer's instructions call for incubation at room temperature, reactions incubated at a controlled temperature of 25°C are more robust.

4. Add 2 μl 2 μg/μl proteinase K solution and incubate 10 min at 37°C.

5. Add 1 to 2 μl BP reaction mixture to 50 μl *E. coli* competent cell culture, and perform transformation.

The BP reaction mixture may be stored up to 6 months at -20°C.

Transformation efficiency of $>1 \times 10^8$ cfu/μg DNA is recommended. Successful transformation with larger genes can sometimes be increased by using cells with higher transformation efficiencies.

Refer to UNIT 1.8 for more information about bacterial transformation.

6. Plate the transformation mix on an LB plate supplemented with 50 μg/ml kanamycin.

Antibiotics should be carefully chosen to be appropriate to the master vector (in this case, pDONR221).

7. Incubate the plate overnight at 37°C.
8. Pick several (one to four) colonies and inoculate each into 1 ml LB medium containing 50 µg/ml kanamycin. Incubate overnight at 37°C.
9. Combine 0.7 ml overnight culture with 0.3 ml 50% (v/v) glycerol and store up to 6 months at –80°C.
10. Isolate plasmid DNA from the remainder of the culture using standard miniprep methods (e.g., *UNIT 1.6*) or a commercial kit (e.g., Qiagen).
11. Sequence the master clone plasmid DNA covering the ORF and recombinational sequence region (see Chapter 7). Verify the sequence against the targeted DNA sequence information using bioinformatics tools (e.g., BLAST at <http://www.ncbi.nlm.nih.gov/blast> or Sequencher at <http://www.genecodes.com/sequencher>).

The BP reactions are not perfect. Smaller fragments can be cloned more efficiently than the desired fragment and occasionally aberrant recombination products can be observed.

It is highly recommended to fully sequence the insert and recombination sites to confirm that the correct, high-quality clone is captured in the master vector.

The authors usually pick four isolates per gene, fully sequence them, and choose only one isolate to represent the high-quality, fully sequenced clone.

Once validated, it is not necessary to verify the sequence again in the expression clone. Refer to Chapter 7 for more information.

12. Pick the clone with the highest quality gene sequence.

**BASIC
PROTOCOL 3**

CAPTURE OF ORFS TO MAKE MASTER CLONES FOR THE CREATOR SYSTEM: IN-FUSION CLONING

Capturing the coding sequences into the Creator master vector is mediated by homologous recombination between the end sequences of PCR fragments and the end sequences of linearized master vector (Fig. 3.20.5). Thus, there are no sequence constraints on this flanking sequence as long as the ends of the PCR product match the ends of the linearized master vector; therefore, this reaction can work well with other vector and sequence combinations. Moreover, because this reaction is based on homologous recombination, many thermostable polymerases in addition to the In-Fusion enzyme listed here, could be adapted to this process. In theory, any proofreading thermostable DNA polymerase can be adapted with enhancer agents (i.e., DMSO, polyethyleneglycol (PEG), *E. coli* single-stranded DNA-binding protein (SSB), Gene 32 protein) instead of commercial In-Fusion enzyme kit (BD Clontech). M13 or T7 forward primers can be used for DNA sequencing.

Materials

- 20 U/µl In-Fusion polymerase (BD Clontech; diluted from 10× stock before use with provided buffer)
- 500 µg/ml bovine serum albumin (BSA)
- 10× In-Fusion reaction buffer (BD Clontech)
- 100 ng/µl linearized pDNR master vector (BD Clontech)
- 10 to 20 ng/µl PCR product: prepared according to the Creator scheme (Basic Protocol 1) and purified (Support Protocol)
- TE buffer, pH 8.0 (*APPENDIX 2*)
- E. coli* high-efficiency competent cell culture (e.g., DH5α; Invitrogen)
- LB plates supplemented with 100 µg/ml ampicillin (or other appropriate antibiotic) and 40 µg/ml Xgal (*UNIT 1.1*)

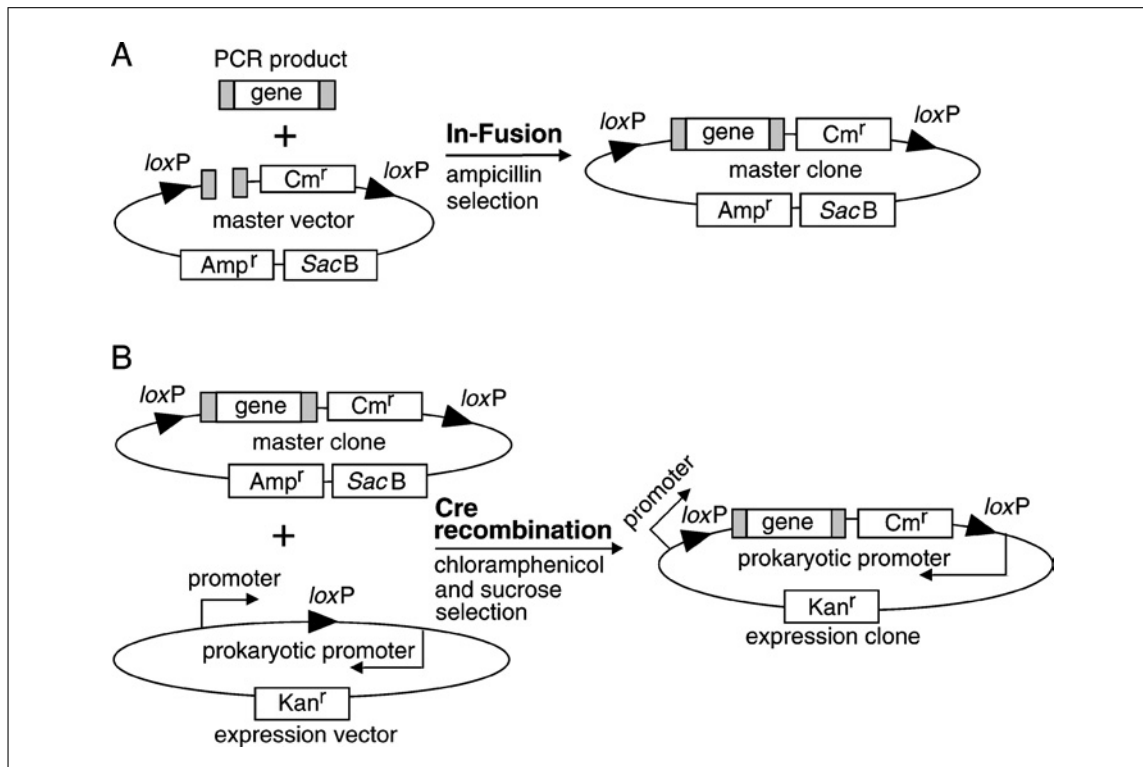


Figure 3.20.5 Schematic diagram of Creator In-Fusion and Cre transfer reactions. In-Fusion facilitates the homologous sequence recombination between PCR product and linearized pDNR-Dual vector. Cre facilitates the transfer of the insert from master clone to expression vector generating expression clones. The chloramphenicol resistance (Cm^r) gene is then under the control of the prokaryotic promoter and is actively expressed.

LB medium supplemented with 100 µg/ml ampicillin (or other appropriate antibiotic; *UNIT 1.1*)

50% (v/v) glycerol, sterile

Plasmid DNA isolation kit (e.g., Qiagen; optional)

96-well plate or 1.5-ml-microcentrifuge tubes

Multichannel pipettor

25°C incubator (optional)

Additional reagents and equipment for growth of *E. coli* competent cells (*UNIT 1.8*), DNA miniprep methods (*UNIT 1.6*) and DNA sequencing (see Chapter 7)

1. Prepare the master reaction mix on ice for the 10-µl In-Fusion reaction (multiply the volume for each component for the number of reactions plus two or three to allow for losses during transfer) as follows (3 µl per tube or well):

0.5 µl 20 U/µl diluted In-Fusion polymerase (10 U final)

1 µl 500 µg/ml BSA (500 ng final)

1 µl 10× In-Fusion reaction buffer (1× final)

0.5 µl 100 ng/µl linearized pDNR master vector (50 ng final).

2. Dispense 3 µl master reaction mix into 1.5-ml microcentrifuge tubes or each well of a 96-well plate.

For 96-well plate format, a multichannel pipettor should be used.

3. Add 7 μ l 10 to 15 ng/ μ l PCR product (50 to 100 ng).
4. Incubate 30 min at 25°C.

Although, the manual calls for room temperature, reactions incubated at a controlled temperature of 25°C are more robust. A longer incubation time is not beneficial.
5. Dilute the reaction mix with 20 μ l TE buffer, pH 8.0.
6. Add 2.5 μ l of In-Fusion reaction mixture to 50 μ l *E. coli* high-efficiency competent cell culture and perform transformation.

Transformation efficiency of $>1 \times 10^8$ cfu/ μ g DNA is recommended. Low efficiencies may result in failed reactions. Do not use more than 5 μ l of diluted reaction mix for 50 μ l competent cells.

Refer to UNIT 1.8 for more information about bacterial transformation.
7. Plate the transformation mix on an LB plate supplemented with 100 μ g/ml ampicillin and 40 μ g/ml Xgal.

Antibiotics should be carefully chosen to be appropriate to the master vector (in this case, pDNR).
8. Incubate the plate overnight at 37°C.
9. Pick several (one to four) white colonies and inoculate each into 1 ml LB medium containing 100 μ g/ml ampicillin. Incubate overnight at 37°C.
10. Proceed as in Basic Protocol 2, steps 9 through 12, to create glycerol stock cultures and sequence plasmid DNA covering the ORF and recombinational sequence region.

Even though the blue-white selections work well, it is necessary to verify the identity of the master clone insert by sequencing.

**BASIC
PROTOCOL 4**

GENERATING GATEWAY EXPRESSION CLONES

Expression clones are generated by the LR reaction between the *attL*-containing master clone and *attR*-containing Gateway-compatible expression vectors (see Fig. 3.20.4). Either the commercial Gateway expression vectors or Gateway-compatible expression vectors generated in-house can be used. Conversion to a Gateway-compatible vector involves a simple insertion of a blunt-ended conversion cassette by ligation into the blunt-ended expression vector that is to be adapted to the Gateway system. The conversion cassette, commercially available for the Gateway system, includes a *ccdB* death gene and a chloramphenicol resistance (*Cm^r*) gene flanked by *attR* sites at the ends (Fig. 3.20.6).

Materials

- Expression vector (e.g., pDual; Stratagene; or any expression vector with a different selectable marker than the master clone)
- Appropriate restriction enzymes (e.g., NEB)
- T4 DNA polymerase *or* Klenow fragment (e.g., NEB)
- Calf intestine alkaline phosphatase (CIAP) and buffer (e.g., NEB; optional)
- TE buffer, pH 8.0 (APPENDIX 2)
- T4 DNA ligase and buffer (e.g., NEB)
- Gateway recombination cassette (rfA, rfB, or rfC; Invitrogen), for making in-house Gateway-compatible expression vectors
- Sterile H₂O
- E. coli* competent cell cultures (e.g., DB3.1 and DH5 α ; Invitrogen)
- LB plates supplemented with 30 μ g/ml chloramphenicol (or other antibiotic appropriate for the expression vector; UNIT 1.1)

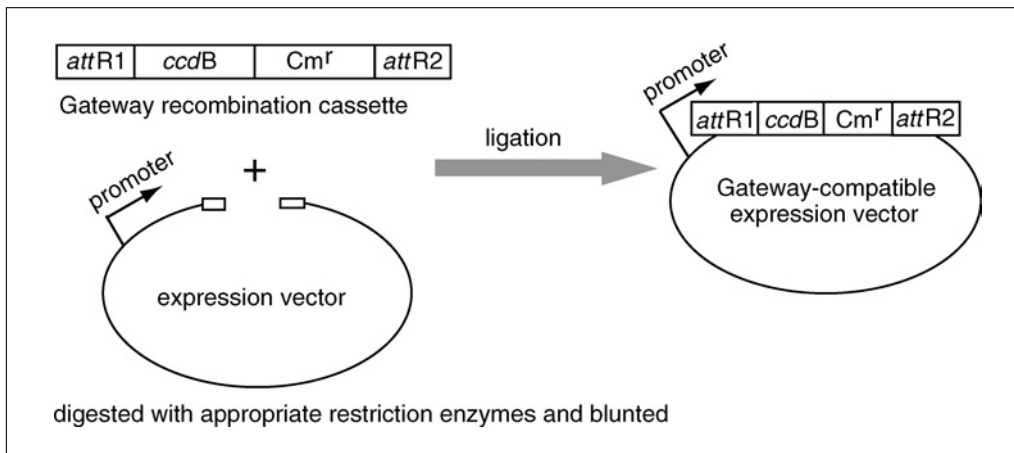


Figure 3.20.6 Conversion of an expression vector into a Gateway-compatible vector. An expression vector of choice is linearized and blunted at favorable restriction sites. It is then ligated with an available blunt-ended cassette containing *ccdB* and *Cm^r* genes. The Gateway-compatible expression vector sequence should be analyzed to confirm the correct orientation.

- LB medium supplemented with 30 $\mu\text{g}/\text{ml}$ chloramphenicol (or other antibiotic appropriate for the expression vector; *UNIT 1.1*)
- 50% (v/v) glycerol, sterile
- 150 ng/ μl Gateway expression vector (e.g., pDEST27; Invitrogen), for commercial expression vector
- LR Clonase (Invitrogen)
- 5 \times LR Clonase buffer (Invitrogen)
- 150 ng/ μl master clone plasmid DNA (Basic Protocol 2)
- 2 $\mu\text{g}/\mu\text{l}$ proteinase K (Invitrogen)
- LB plates supplemented with 100 $\mu\text{g}/\text{ml}$ ampicillin (or other appropriate antibiotic; *UNIT 1.1*)
- LB medium supplemented with 100 $\mu\text{g}/\text{ml}$ ampicillin (or other appropriate antibiotic; *UNIT 1.1*)
- 96-well plates or 1.5-ml microcentrifuge tubes
- Multichannel pipettor
- 25°C incubator
- Additional reagents and equipment for growth of *E. coli* competent cells (*UNIT 1.8*), linearizing the expression vector at the multiple cloning sites using restriction enzymes (*UNIT 3.1*), generating blunt ends using T4 DNA polymerase or Klenow fragment (*UNITS 3.5 & 3.16*), and analyzing plasmid DNA using restriction enzymes (*UNIT 3.1*)

Prepare a Gateway-compatible expression vector (optional)

1. Linearize the expression vector at the multiple cloning sites using restriction enzymes (*UNIT 3.1*), and generate blunt ends using T4 DNA polymerase or Klenow fragment (*UNITS 3.5 & 3.16*).
If necessary, perform dephosphorylation using CIAP (see UNIT 3.10) to decrease background from self-ligation of the vector.
2. Adjust the linearized expression vector to a final concentration of 20 to 50 ng/ μl in TE buffer, pH 8.0.
3. Prepare the ligation reaction mix in a microcentrifuge tube as follows:
 - 1 to 5 μl of 20 to 50 ng/ μl dephosphorylated linearized vector
 - 1 μl 10 \times T4 DNA ligase buffer
 - 2 μl 5 ng/ μl Gateway recombination cassette

1 μl 1 U/ μl T4 DNA ligase
Sterile H_2O to a final volume of 10 μl .

If a polypeptide fusion is to be added at either end of the gene, ensure that the Gateway recombination cassette is in the appropriate reading frame.

4. Incubate 1 hr at room temperature.
5. Add 1 to 2 μl reaction mixture to 50 μl *E. coli* DB3.1 competent cell culture and perform transformation.

DB3.1 is a strain modified to allow growth of plasmids containing ccdB. Cells other than DB3.1 will be killed by the new expression vector that will carry a copy of the ccdB death gene.

Transformation efficiency of $>1 \times 10^8$ cfu/ μg DNA is recommended.

Refer to UNIT 1.8 for more information about bacterial transformation.

6. Plate the transformation mix on an LB plate supplemented with 30 $\mu\text{g}/\text{ml}$ chloramphenicol. Incubate overnight at 37°C.

Antibiotics should be carefully chosen to be appropriate to the recombination cassette (in this case chloramphenicol).

7. Pick several (one to four) colonies and inoculate each into 1 ml LB supplemented with 30 mg/ml chloramphenicol. Incubate overnight at 37°C.

8. Combine 0.7 ml overnight culture with 0.3 ml 50% glycerol and store up to 6 months at -80°C .

9. Isolate plasmid DNA from the remainder of the culture using standard miniprep methods (e.g., UNIT 1.6) or a commercial kit (e.g., Qiagen). Analyze plasmid DNA using appropriate restriction enzymes (UNIT 3.1) or by sequencing to identify clones with the proper orientation. Quantify plasmid DNA from a selected clone (APPENDIX 3D).

Transformants need to be screened for orientation because both orientations will be generated after blunt-end ligation. Appropriate restriction enzymes will cut the vector and cassette in an asymmetric fashion, yielding fragment lengths that depend on orientation.

Transfer an ORF into the expression vector using the Gateway LR reaction

10. Prepare the master reaction mix on ice for the 20- μl LR reaction (multiply the volume for each component for the number of reactions plus two or three to allow for losses during transfer) as follows (18 μl per tube or well):

2 μl 150 ng/ μl commercial Gateway (e.g., pDEST27) or Gateway-compatible (from step 9) expression vector
4 μl LR Clonase
4 μl 5 \times LR Clonase buffer
8 μl TE buffer, pH 8.0.

11. Mix well and dispense 18 μl master reaction mix into each well of a 96-well plate or 1.5-ml microcentrifuge tube.

For a 96-well plate format, a multichannel pipettor should be used.

12. Add 2 μl master clone plasmid DNA to the 18 μl master reaction mix.

13. Incubate the reaction 1 hr at 25°C.

Although, the manual calls for room temperature, reactions incubated at a controlled temperature of 25°C are recommended.

14. Add 2 μl proteinase K solution and incubate at 37°C for 10 min.
The reaction mixture can be stored up to 6 months at –20°C after this step.
15. Add 2 to 3 μl of reaction mixture to 50 μl *E. coli* competent cell culture (e.g., DH5 α) and perform transformation.
Any standard cloning strain is fine.
Transformation efficiency of $>1 \times 10^8$ cfu/ μg DNA is recommended.
Refer to UNIT 1.8 for more information about bacterial transformation.
16. Plate the transformation mix on an LB plate supplemented with 100 $\mu\text{g}/\text{ml}$ ampicillin. Incubate overnight at 37°C.
Antibiotics should be carefully chosen to be appropriate to the expression vector, in this case, ampicillin for the commercial vector pDEST27.
17. Pick colonies and analyze the plasmid DNA using restriction enzymes.
Typically, the transfer reaction is very efficient (90% to 95%). For large-scale projects, in many cases, it is not necessary to pick several colonies and analyze them.
18. Choose the best clone and prepare for the desired use or store as a 15% glycerol stock up to 6 months at –80°C.

GENERATING CREATOR EXPRESSION CLONES

Cre recombinase facilitates the transfer reaction of the ORF from the master clone to a Creator-compatible expression vector in generating expression clones (Fig. 3.20.5). The process can be performed using the commercially available vectors or Creator-compatible vectors generated in-house. The generation of Creator-compatible expression vectors involves (1) the amplification of the Creator conversion cassette with primers that contain restriction sites, (2) restriction of expression vector, and (3) ligation of the prepared cassette into the target expression vector (Fig. 3.20.7).

Materials

- Creator expression vector (e.g., pLP-CMVneo; BD Clontech) or other expression vector to make Creator-compatible vector
- LoxP* primers (BD Clontech)
- LoxP* template (e.g., Creator acceptor vector; BD Clontech)
- Restriction enzymes and buffers (NEB)
- T4 DNA ligase and buffer (NEB)
- E. coli* competent cell culture (e.g., DH5 α ; Invitrogen)
- LB plates supplemented with 100 $\mu\text{g}/\text{ml}$ kanamycin (or other appropriate antibiotic; *UNIT 1.1*)
- LB plates supplemented with 30 $\mu\text{g}/\text{ml}$ chloramphenicol (or other appropriate antibiotic; *UNIT 1.1*) and 7% (w/v) sucrose
- Cre recombinase (New England Biolabs)
- 10 \times Cre reaction buffer (New England Biolabs)
- 20 ng/ μl master clone plasmid DNA (Basic Protocol 3)
- 96-well plate or 1.5-ml microcentrifuge tubes
- Multichannel pipettor
- Additional reagents and equipment for PCR (Basic Protocol 1), agarose gel electrophoresis (Support Protocol; *UNIT 2.5A*), growth of *E. coli* competent cells (*UNIT 1.8*), and analyzing the plasmid DNA using restriction enzymes (*UNIT 3.1*) or recombination-specific PCR methods (BD Clontech)

BASIC PROTOCOL 5

Enzymatic Manipulation of DNA and RNA

3.20.13

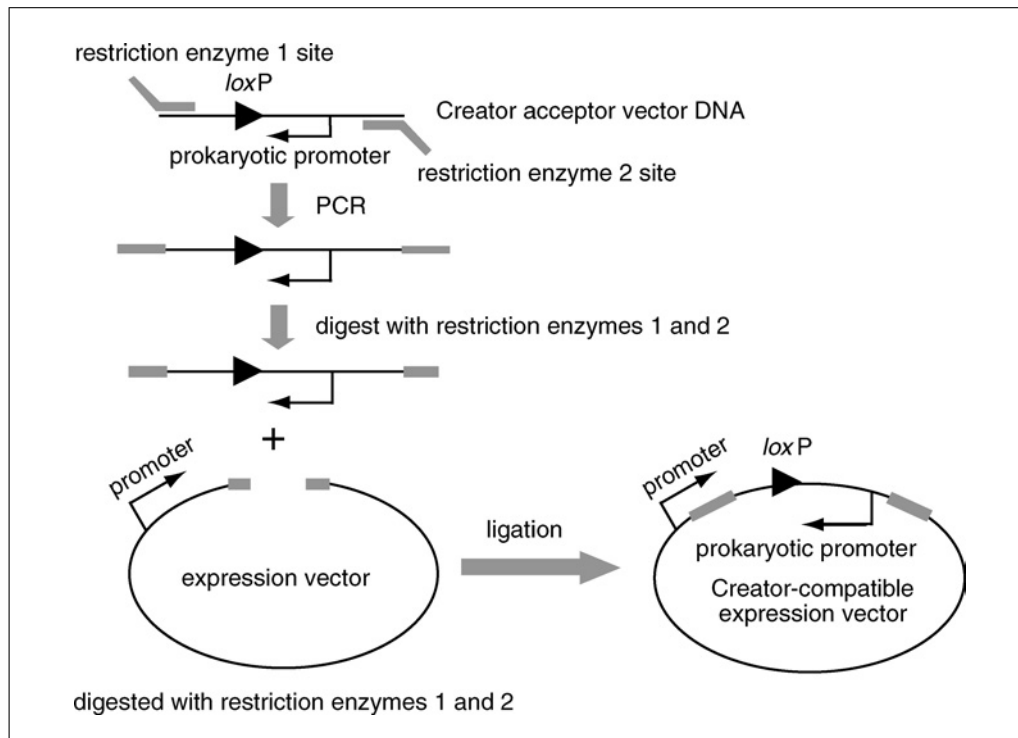


Figure 3.20.7 Conversion of an expression vector into a Creator-compatible vector. The region flanking the *loxP* site and prokaryotic promoter is amplified with specific PCR primers containing specific restriction enzyme sites. Then, the PCR products and an expression vector are digested with appropriate restriction enzyme. A ligation reaction generates a Creator-compatible expression vector.

Prepare Creator-compatible expression vector (optional)

1. Design PCR primers to amplify the *loxP* template as follows:

5' primer: 5'-RS-ATA ACT TCG TAT AGC ATA CAT TAT-3'
 3' primer: 5'-RS-CAC GTC AGG TGG CAC TTT TCG-3'.

*Restriction sites (RS) should be added, allowing the ligation of the cassette into the target vector. The target vector should have restriction enzyme sites that are not present in the *loxP* cassette.*

Available restriction sites can be found in the Creator Acceptor vector construction manual.

2. Amplify the *loxP* cassette by PCR (see Basic Protocol 1, steps 1 to 3 and 4b).
3. Digest the PCR *loxP* fragment and the expression vector with the selected restriction enzymes.
4. Purify the digested PCR fragment (161 bp) by agarose gel electrophoresis (UNIT 2.5A).
5. Ligate (see Basic Protocol 4, step 3) the PCR-amplified *loxP* cassette and the linearized expression vector.

*Use of the appropriate matched restriction enzymes for preparing the *loxP* cassette and the linearized expression vector are necessary for the ligation reaction to occur.*

6. Add 2.5 μ l of ligation mixture to 50 μ l *E. coli* competent cell culture, e.g., DH5 α , and perform transformation.

Transformation efficiency of $>1 \times 10^8$ cfu/ μ g DNA is recommended.

Refer to UNIT 1.8 for more information about bacterial transformation.

7. Plate the transformation mix on an LB plate supplemented with 100 µg/ml kanamycin. Incubate overnight at 37°C.

Antibiotics should be carefully chosen to be appropriate to the expression vector.

8. Pick colonies and prepare overnight cultures and glycerol stocks (see Basic Protocol 4, steps 7 and 8). Prepare plasmid DNA (UNIT 1.6 or Qiagen) and quantify (APPENDIX 3D).

Transfer an ORF into the expression vector using Creator Cre reaction

9. Prepare the master reaction mix on ice for the 20-µl LR reaction (multiply the volume for each component for the number of reactions plus two or three to allow for losses during transfer) as follows (10 µl per tube or well):

Reaction mix per well includes:

200 ng commercial Creator (e.g., pLP-CMVneo) or Creator-compatible (step 8) expression vector

1 µl Cre recombinase

2 µl 10× Cre reaction buffer

Bring volume to 10 µl with H₂O.

10. Dispense 10 µl master reaction mix into each well of a 96-well plate.

For a 96-well plate format, a multichannel pipettor should be used.

11. Add 10 µl 20 ng/µl master clone plasmid DNA to the 10 µl master reaction mix.

Master clone plasmid DNA (150 to 200 ng) per reaction works well.

12. Incubate 30 min at 37°C.

13. Heat inactivate the reaction 15 min at 70°C.

14. Add 2 to 3 µl reaction mixture to 50 µl *E. coli* competent cell culture, e.g., DH5α, and perform transformation.

Transformation efficiency of $>2 \times 10^8$ cfu/µg DNA is recommended.

More than 5 µl of reaction mix can be toxic to bacterial competent cells.

Refer to UNIT 1.8 for more information about bacterial transformation.

15. Plate the transformation mix on an LB plate supplemented with 30 µg/ml chloramphenicol and 7% (w/v) sucrose. Incubate the plates overnight at 37°C.

Antibiotics should be carefully chosen to be appropriate to the expression vector (in this case, pLP-CMVneo).

16. Pick good-sized colonies.

It is important to pick good-sized colonies rather than small ones. It has been reported that the small colonies could be background and may harbor undesirable recombinants (Creator DNA cloning kit manual).

17. Analyze the plasmid DNA of picked colonies using restriction enzymes (UNIT 3.1) or recombination-specific PCR methods (BD Clontech).

Typically, the transfer reaction is very efficient (90% to 95%). For large-scale projects, in many cases, it is not necessary to pick several colonies and analyze them.

18. Choose the best clone and and prepare for the desired use or store as a 15% glycerol stock up to 6 months at –80°C.

COMMENTARY

Background Information

Overview of universal cloning strategy

As molecular biology has matured, the number of experimental approaches that can be conceived and executed with a cloned cDNA copy of a gene has grown exponentially. Proteins can be expressed *in vitro* or *in vivo* in a wide variety of organisms, ranging from bacteria to human cells; their expression can drive the production of purified protein for structural studies, or their regulated expression can be used for complementing genetic phenotypes; epitope tags (e.g., green fluorescent protein, GST, or HA) can be added as markers for localization, for use in purification, or as part of two-hybrid assay systems. Each of the myriad applications that utilize protein expression demands the subcloning of the relevant cDNA into an appropriate plasmid vector that adds the relevant selectable markers, transcriptional and/or translational signals, any needed polypeptide tags, and other relevant sequences. Using traditional subcloning methods based on restriction enzymes and ligation, the effort required to transfer cDNAs from vector to vector can be substantial and time consuming. If one aspires to do large-scale experiments involving the examination of thousands of genes simultaneously using automated and parallel approaches, it would be impossible using standard methods. To move the focus of the community's efforts away from the mechanics of subcloning and put it back on biology, recombinational cloning has been developed. This approach facilitates the transfer of DNA molecules from one vector to another in a single concerted step that is highly efficient, *in-frame*, and does not introduce mutations.

Recombinational cloning is based on site-specific recombination to transfer DNA sequences into multiple expression systems (Walhout et al., 2000; Reboul et al., 2001, 2003; Brizuela et al., 2002; Yamada et al., 2003; Marsischky and LaBaer, 2004). It is independent of the insert DNA sequence to be cloned, separating it from the limitations of classical restriction enzyme-based cloning. It enables rapid and efficient transfer of single inserts into multiple different expression systems for protein expression and functional analysis. It also allows the simultaneous transfer of a many different DNA sequences into an expression vector in parallel, enabling high throughput applications (Fig. 3.20.1). Presently, the Gateway system (Invitrogen)

and the Creator system (BD Clontech) are commercially available.

The strategy employs the use of a set of master clones from which numerous different expression clones can be easily produced. The master clones contain the coding sequences (also called open reading frames, ORFs) situated in inert (non-protein-expressing) vectors, which are endowed with recombinational sequences. The use of inert vectors for the master clones avoids inadvertent protein expression and consequent negative selection of clones that contain inserts toxic to the propagating bacteria. Using recombination, the coding sequences can then be transferred to any protein expression vector that has the recombination sequences situated in the appropriate position and reading frame. Virtually any existing expression vector can be easily adapted to contain recombinational sequences for either of the existing commercial systems. Based upon the design of the original master clones, the expression clones can either produce native protein (when master clones contain a naturally positioned stop codon) or carboxyl-terminal tagged proteins (when the stop codon is omitted from the master clone). The transfer process is a conservative molecular reaction. Once the sequence of the master clones has been verified, there is no need to verify the daughter clones. Thus, one library of clones in master vectors becomes many libraries of clones in expression vectors.

The steps involved in making expression-ready master clones include (1) identifying the genes that will be cloned, (2) obtaining sequence information regarding the coding sequences, (3) determining whether to include or exclude the stop codon, (4) designing gene-targeted oligos for PCR, (5) constructing master clones using high-fidelity PCR and capture into the master vector, (6) bacterial transformation and picking colonies, (7) DNA sequencing of the insert and important flanking sequences, and (8) sequence analysis and selection of master clones that meet the acceptance criteria. Once the master clones are constructed and their sequences verified, they can be transferred to appropriately constructed expression vectors to generate expression-ready clones.

As a matter of a common lexicon, the term vector is used to describe plasmids without an insert, whereas the term clone refers to a vector containing an insert. For example, one

might state that the coding sequences from a set of master clones were transferred to an expression vector to produce expression clones. Each commercial system also uses its own terminologies. The Gateway system refers to master vectors as entry vectors and expression vectors as destination vectors. The Creator system refers to master vectors as donor vectors and expression vectors as acceptor vectors. Each system supports the insertion of a simple cassette that adds the appropriate recombinational signals and relevant selectable markers into virtually any existing expression vector adapting it for use in that system.

The recombination reactions themselves are equilibrium reactions that result in a mixture of end products, starting substrates, and intermediates. The desired product may only be present as a small fraction of this mix, but by using the appropriate positive and negative selections, it is the only viable product that is recovered. This results in an overall efficiency that exceeds 96% in most cases, which can enable high throughput automation. Experiments can be easily expanded to genome scale once a complete high-quality collection has been constructed in a recombinational cloning system.

The Gateway system

Gateway technology is based on the bacteriophage lambda (λ) site-specific recombination system, which facilitates the integration of λ genome into the *Escherichia coli* (*E. coli*) chromosome and the switch between the lytic and lysogenic pathways (Ptashne, 1992; Hartley et al., 2000; Walhout et al., 2000). Lambda-based recombination involves the DNA recombination sequences and the proteins that mediate the recombination reaction (*i.e.*, Clonase enzyme mix). Recombination occurs between specific attachment (*att*) sites on the interacting DNA molecules. The DNA segments flanking the recombination sites are changed such that after recombination the *att* sites become hybrid sequences from the *att* sequences of each parental vector. For example, *attL* sites comprise sequences from *attB* and *attP* sites. Strand exchange occurs within a core region that is common to all *att* sites. For more detailed information about λ recombination, see published references and reviews (Landy, 1989; Ptashne, 1992; <http://www.invitrogen.com/gateway>).

The *att* sites serve as the binding site for recombination proteins and have been well characterized (Weisberg and Landy, 1983). Upon λ integration, recombination occurs between

attB (bacteria) and *attP* (phage) sites to give rise to *attL* (left) and *attR* (right) sites. The *attB* and *attP* sites are different sizes, and because the *attL* and *attR* sites are hybrids, their sizes fall in between. The actual crossover occurs between homologous core regions (15 bp) on the two sites, but surrounding sequences are required as they contain the binding sites for the recombination proteins (Landy, 1989). In developing the Gateway system, mutations were introduced into the core of the *att* sites to create different sets that are incompatible with each other, generating *attB1* and *attP1* sites and *attB2* and *attP2* sites. The *attB1* recombine only with *attP1*, not with *attP2*. This allows a gene to be flanked by *attL1* and *attL2* sites on either side that can recombine with the corresponding *attR* sites and maintain the correct orientation. The different species of recombination sites (*attB*, *attP*, *attL*, and *attR*) are found in different vector constructs. The system is designed so that the *attB* site, the smallest of the four, appears in the final expression clone introducing an additional 9 amino acids into any fusion proteins.

Lambda recombination is catalyzed by a mixture of enzymes that bind to specific *att* sequences. The recombination proteins differ depending upon which direction the reaction is run. When the reaction is run to mix *attB* and *attP* sites, such as when making master clones, it is catalyzed by the bacteriophage λ integrase (Int) and *E. coli* integration host factor (IHF) proteins (BP Clonase enzyme mix), whereas, if the reaction is between *attL* and *attR* sites it is catalyzed by the bacteriophage λ Int and excisionase (Xis) proteins, and the *E. coli* IHF protein (LR Clonase enzyme mix). For more information about the recombination enzymes, see published references and reviews (Nash, 1977; Nash and Robertson, 1981; Landy, 1989; Ptashne, 1992). Two recombination reactions constitute the basis of the Gateway technology. BP reaction, catalyzed by BP Clonase enzyme mix, facilitates recombination of an *attB* substrate (*attB*-PCR product or a linearized *attB*-containing expression clone) with an *attP* substrate (master vector) to create an *attL*-containing master clone (Fig. 3.20.4). The BP cloning efficiency tends to drop rapidly with increasing insert DNA size larger than 2 kb (Marsischky and LaBaer, 2004). On the other hand, LR Reaction, catalyzed by LR Clonase enzyme mix, facilitates recombination of an *attL* substrate (master clone) with an *attR* substrate (expression vector) to create an *attB*-containing expression clone (Fig. 3.20.4).

Recombination reaction is an equilibrium reaction; therefore tight selection is required to ensure that only the desired product is obtained. In the Gateway system, positive selection for the correct master clone is mediated by an antibiotic resistance marker. To avoid contamination from starting vectors, there is counter-selection by the death gene (*ccdB*), which is toxic to general cloning bacteria (Bernard and Couturier, 1992). Thus, only bacteria that contain master clones where the death gene has been successfully replaced by the desired insert will survive selection. Vectors carrying the death gene can be propagated using a special *E. coli* strain (DB3.1), which is resistant to the *ccdB* gene. The combined positive and negative selection provides an overall BP or LR reaction efficiency of 95% to 100%. For more detail refer to the Gateway manual from the manufacturer.

The Creator system

The BD Clontech Creator system relies on homologous recombination and Cre-*loxP* recombination generating master clone and expression clone, respectively. PCR amplified ORF containing homologous sequences in each ends will be transferred into a master vector in the In-Fusion reaction. Because it is simply a homologous recombination event, many thermostable DNA polymerase could be adapted in place of the In-Fusion enzyme. The Cre-*loxP* site-specific recombination reaction catalyzes the transfer of ORF from a master clone to an expression vector creating expression clone (Fig. 3.20.5). Cre, a 38-kDa recombinase protein from bacteriophage P1, mediates recombination between or within DNA sequences at specific locations called *loxP* sites (Sternberg et al., 1981; Abremski and Hoess, 1984; Sauer, 1994). These sites consist of two 13-bp inverted repeats separated by an 8-bp spacer region that provides directionality to the recombination reaction (<http://www.clontech.com/clontech/products/families/creator>). The 8-bp spacer region in the *loxP* site has a defined orientation that forces a gene to be transferred in a fixed orientation and reading frame.

Master vectors contain two *loxP* sites that flank a multicloning site (MCS) and the open reading frame for the chloramphenicol resistance gene (*Cm^r*; Fig. 3.20.5). Master vectors also contain the ampicillin gene for propagation and selection in *E. coli*, and the sucrose gene (*sacB*) from *B. subtilis* for negative selec-

tion of correct recombinants. Expression vectors contain a single *loxP* site, followed by a bacterial promoter, which drives expression of the chloramphenicol marker after Cre-*loxP*-mediated recombination. After transfer, the gene of interest will become linked to the specific expression elements for which the expression vector was designed. Furthermore, if the coding sequence for the gene of interest is in frame with the upstream *loxP* site in the master vector, it will automatically be in frame in the expression vector. The expression of C-terminal fusion proteins is complicated because the chloramphenicol resistance gene blocks access to C-terminal fusion tags. To enable C-terminal fusions, the gene of interest should be collected into a master vector that has a splice donor site (pDNR-Dual), and any C-terminal tag will be spliced in-frame to the gene of interest.

During the recombination reaction, the DNA sequences between the *loxP* sites on the master clone are inserted at the *loxP* site of the expression vector. This reaction will put the chloramphenicol resistance gene under control of the constitutively active bacterial promoter and the gene of interest under control of the heterologous promoter (Fig. 3.20.5). Because the recombination is site-specific, the reading frame is always maintained. Correct recombinants can be selected using positive selection on chloramphenicol (selects for the recombinant) and counter selection on sucrose (selects against the master vector). Sucrose in the presence of endogenous sucrose is toxic to general *E. coli* strains (Pelicic et al., 1996). For more details, refer to the Creator manual from the manufacturer.

Modification for use with other vectors

The Gateway and Creator recombinational cloning technology platforms provide various forms of expression systems (e.g., bacterial, mammalian, insect, and in vitro transcription/translation). In many cases, however, users may want to modify their own expression vectors into recombination-compatible systems (Figs. 3.20.6 and 3.20.7). For that purpose, those commercial systems also provide a way to convert other expression vectors into recombination-compatible ones by simply inserting the recombination cassette, using PCR and restriction-mediated cloning steps. For this purpose, refer to standard cloning manuals or to the Gateway and Clontech vector conversion manuals for more details.

Critical Parameters and Troubleshooting

Amplification of target DNA

Primer mapping. When performing agarose gel electrophoresis in handling multiple genes or large cloning projects using a 96-well plate format, it is helpful if the resulting gel bands form a saw tooth pattern in which the expected fragments progressively increase in size across the gel lanes, but alternate in size to avoid lane-to-lane contamination (Fig. 3.20.8). This effect is most easily planned at the stage of primer design, so that the genes are amplified in an order that will result in this pattern.

DNA polymerase. Fidelity and PCR success rate (% of genes with positive PCR product) are crucial parameters in high-throughput PCR cloning. The choice of polymerase has a dramatic effect on the success rate. In the authors' experience, *Pfu* Turbo (Stratagene) and *Pfx* (Invitrogen) were comparable and gave the highest success rates in cloning up to 4 kbp from human brain and placenta first-strand cDNA mix. In addition, Herculase (Stratagene), GC-rich polymerase (Roche), and Elongase (Invitrogen) were successfully used in other systems (Walhout et al., 2000; Reboul et al., 2001). Running pilot experiments on small subsets of clones is recommended before executing a large-scale experiment in order to

adjust reaction parameters, e.g., the amount of polymerase, melting temperature, cycle number, Mg^{2+} concentration, and GC content of the primers.

Template. PCR reactions using plasmid template are highly efficient. As little as 200 ng of template plasmid DNA is enough to reduce the number of PCR cycles and the rate of mutations. When the plasmid form for the target gene is not available, the cDNAs must be amplified from cDNA libraries, first-strand cDNA, or genomic DNA. Genomic DNA is an excellent source for a simple organism because it is essentially a normalized library where each gene is represented equally. For eukaryotes, the authors have found increased success rates from first-strand cDNA with highest rates in the brain and placenta first-strand cDNA mixture (1:4) for general human gene cloning purposes; however, for cloning certain cell- or tissue-type-specific genes, use of template DNA generated from those cells or tissues is recommended.

Generation of master clones

Pilot experiments. The concentration of purified PCR product can vary from gene to gene, especially in large-scale projects, and can influence the success rate of generating master clones, even though all were successfully amplified. Consequently, running a pilot

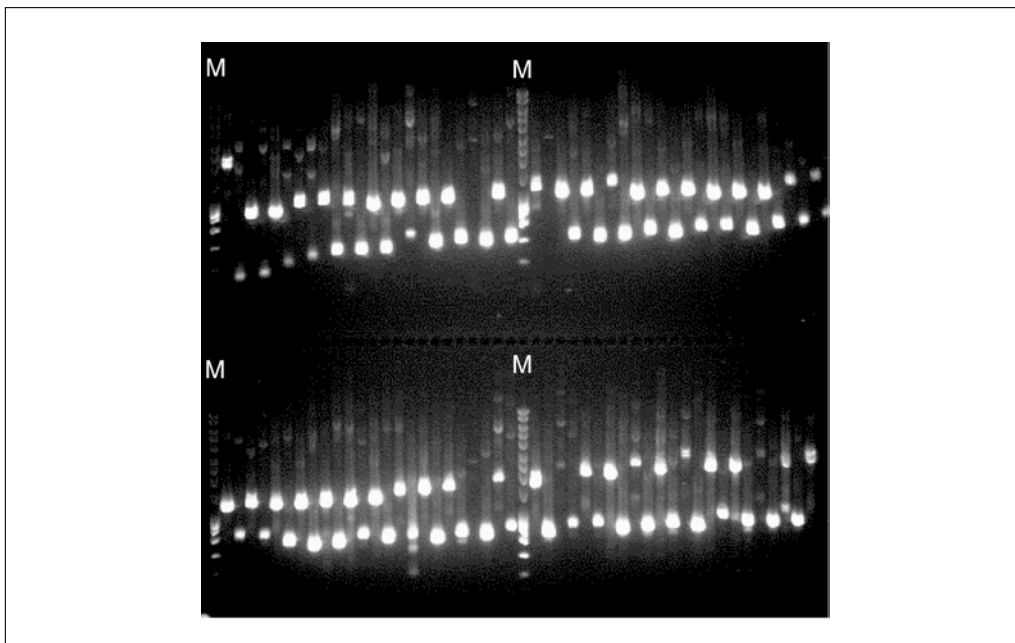


Figure 3.20.8 Examples of a saw-tooth pattern in agarose gel analysis of PCR product. PCR mixtures (96 samples) were run on a 1% agarose gel. A primer-ordering algorithm generates a map of oligonucleotide primers in 96-well format with alternating expected product size, so that the expected PCR bands would appear in a saw-tooth fashion. The letter M indicates DNA size markers.

experiment with a small number of genes is strongly recommended, using positive and negative controls for estimating the range of PCR product concentration against the fixed master vector concentration.

Transformation. The efficiency of transformation into competent *E. coli* cells is critical, especially in dealing with large-scale projects in a 96-well format. Transformation efficiencies of $>1 \times 10^8$ cfu/ μ g DNA are recommended. For high-throughput projects LB agar can be prepared in 48-well grid square bioassay trays (Genetix), and the bacterial colonies can be automatically picked using robotics.

Selection. Because the recombination reaction does not go to completion, a mixture containing the starting plasmids, one or more intermediates, the desired product, and one or more byproducts is obtained. Thus, stringent antibiotic selection conditions are required to select only the desired recombinant containing the target gene in its correct orientation.

Verification. Enzymatic amplification using PCR can introduce unwanted random mutations in the DNA sequence. In addition, primer mutations can be introduced during synthesis of the oligonucleotide primers. The enzymatic contribution to this mutation rate can be decreased by reducing the number of amplification cycles. Nevertheless, each clone should be sequenced in its full length to verify the correct sequence of the target gene and the recombination site has been obtained. Once a master clone is verified, it is not necessary to resequence after it is transferred into an expression vector.

Generation of expression clones

Transformation. As in generation of the master clone, the efficiency of transformation into competent *E. coli* cells is critical in dealing with large-scale projects in a 96-well format. Bacterial competent cells with transformation efficiencies of $>2 \times 10^8$ cfu/ μ g DNA are recommended for the Creator system pipeline.

Pilot experiments. Unusual recombination events occasionally occur, especially in retroviral expression vectors. It is advisable to confirm that the expression clones harbor correct inserts in proper orientation by analyzing plasmid DNA of some expression clones using specific restriction enzymes in a pilot experiment.

Anticipated Results

Amplification of target DNA

From both plasmid DNA or genomic DNA (especially prokaryotic cells) as a template,

PCR should amplify only the gene of interest, and the overall success rate is $>95\%$. In the case of first-strand cDNA, the PCR success rate is $\sim 70\%$ because of the complexity and relatively low representation of each gene. Upon electrophoresis of first-strand PCR product the authors frequently observe a few bands in addition to those for the desired products, most likely because of the complexity of the mixture. Purification of the desired products alleviates that problem except in cases where the product bands are smears. In those cases PCR conditions should be optimized again.

Generation of master clones

In the authors' experience, a $>95\%$ success rate for generating master clones from plasmid DNA can be expected, but with first-strand cDNA the situation is a bit more complicated. Generally, upon electrophoresis, the PCR product band is faint compared to those from plasmid DNA template because of a lower DNA concentration. This is a result of the lower success rate in PCR steps ($\sim 70\%$) for first-strand DNA templates. Moreover, after DNA sequencing and analysis, more clones may be eliminated for the first-strand cDNA template because of more random mutations introduced during higher numbers of PCR cycles required as a result of the more complex and heterogeneous DNA template. Overall, based on the authors' experience, a $>50\%$ success rate is expected from a first-strand cDNA cloning scheme, compared to $>95\%$ overall success rate using plasmid DNA as template.

Generation of expression clones

The transfer reaction is robust, and a $\sim 95\%$ success rate is expected. The few ineffective transfers should be easily rescuable after an additional round of transfer reaction. Their failure is most often due to random losses during plasmid DNA preparation in the 96-well format. Using highly competent cells (2×10^8 cfu/ μ g DNA competency), at least fifty colonies in a 10-cm petri dish are normally obtained from a standard bacterial transformation protocol.

Because the entire process of transferring inserts into expression vectors is very robust, it is easily automated. In addition, because the reaction does not involve any amplification or rearrangements, if the insert DNA sequence has been fully verified and is in-frame with the recombination sites, it is not necessary to resequence the insert DNA sequence again after constructing expression clones.

Any functioning expression vector can be converted to accept inserts in the Gateway or Creator systems. It should be noted that the final expression vector is in the same reading frame, which will ensure that all fusion proteins will be in-frame. The expression vector should not use the same antibiotic marker as those of the master vector.

Time Considerations

The basic procedures can be completed in ~1 week with activities broken down as follows: 1 day for PCR and gel plugging, 1 day for capture reactions and transformation, 1 day for colony picking and DNA minipreps, 1 to 2 days for sequencing, 1 day for transfer of ORF into expression vector, and 1 to 2 days for DNA preparation and analysis.

Literature Cited

- Abremski, K. and Hoess, R. 1984. Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein. *J. Biol. Chem.* 259:1509-1514.
- Bernard, P. and Couturier, M. 1992. Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J. Mol. Biol.* 226:735-745.
- Brizuela, L., Richardson, A., Marsischky, G., and LaBaer, J. 2002. The FLEXGene repository: Exploiting the fruits of the genome projects by creating a needed resource to face the challenges of the post-genomic era. *Arch. Med. Res.* 33:318-324.
- Hartley, J.L., Temple, G.F., and Brasch, M.A. 2000. DNA cloning using in vitro site-specific recombination. *Genome Res.* 10:1788-1795.
- Landy, A. 1989. Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annu. Rev. Biochem.* 58:913-949.
- Marsischky, G. and LaBaer, J. 2004. Many paths to many clones: A comparative look at high-throughput cloning methods. *Genome Res.* 14:2020-2028.
- Nash, H.A. 1977. Integration and excision of bacteriophage lambda. *Curr. Top. Microbiol. Immunol.* 78:171-199.
- Nash, H.A. and Robertson, C.A. 1981. Purification and properties of the *Escherichia coli* protein factor required for lambda integrative recombination. *J. Biol. Chem.* 256:9246-9253.
- Pelicic, V., Reyat, J.M., and Gicquel, B. 1996. Expression of the *Bacillus subtilis* *sacB* gene confers sucrose sensitivity on mycobacteria. *J. Bacteriol.* 178:1197-1199.
- Ptashne, M. 1992. A Genetic Switch: Phage (Lambda) and Higher Organisms. Cell Press, Cambridge, MA.
- Reboul, J., Vaglio, P., Tzellas, N., Thierry-Mieg, N., Moore, T., Jackson, C., Shin-I, T., Kohara, Y., Thierry-Mieg, D., Thierry-Mieg, J., Lee, H., Hitti, J., Doucette-Stamm, L., Hartley, J.L., Temple, G.F., Brasch, M.A., Vandenhoute, J., Lamesch, P.E., Hill, D.E., and Vidal, M. 2001. Open-reading-frame sequence tags (OSTs) support the existence of at least 17,300 genes in *C. elegans*. *Nat. Genet.* 27:332-336.
- Reboul, J., Vaglio, P., Rual, J.F., Lamesch, P., Martinez, M., Armstrong, C.M., Li, S., Jacotot, L., Bertin, N., Janky, R., Moore, T., Hudson, J.R. Jr, Hartley, J.L., Brasch, M.A., Vandenhoute, J., Boulton, S., Endress, G.A., Jenna, S., Chevet, E., Papisotiropoulos, V., Tolia, P.P., Ptacek, J., Snyder, M., Huang, R., Chance, M.R., Lee, H., Doucette-Stamm, L., Hill, D.E., and Vidal, M. 2003. *C. elegans* ORFeome version 1.1: Experimental verification of the genome annotation and resource for proteome-scale protein expression. *Nat. Genet.* 34:35-41.
- Sauer, B. 1994. Site-specific recombination: Developments and applications. *Curr. Opin. Biotechnol.* 5:521-527.
- Sternberg, N., Hamilton, D., Austin, S., Yarmolinsky, M., and Hoess, R. 1981. Site-specific recombination and its role in the life cycle of bacteriophage P1. *Cold Spring Harb. Symp. Quant. Biol.* 45:297-309.
- Walhout, A.J., Temple, G.F., Brasch, M.A., Hartley, J.L., Lorson, M.A., van den Heuvel, S., and Vidal, M. 2000. GATEWAY recombinational cloning: Application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol.* 328:575-592.
- Weisberg, R.A. and Landy, A. 1983. Site-Specific Recombination in Phage Lambda. In *Lambda II* (R.A. Weisberg, ed.) pp. 211-250. Cold Spring Harbor Press, Cold Spring Harbor, NY.
- Yamada, K., Lim, J., Dale, J.M., Chen, H., Shinn, P., Palm, C.J., Southwick, A.M., Wu, H.C., Kim, C., Nguyen, M., Pham, P., Cheuk, R., Karlin-Newmann, G., Liu, S.X., Lam, B., Sakano, H., Wu, T., Yu, G., Miranda, M., Quach, H.L., Tripp, M., Chang, C.H., Lee, J.M., Toriumi, M., Chan, M.M., Tang, C.C., Onodera, C.S., Deng, J.M., Akiyama, K., Ansari, Y., Arakawa, T., Banh, J., Banno, F., Bowser, L., Brooks, S., Carninci, P., Chao, Q., Choy, N., Enju, A., Goldsmith, A.D., Gurjal, M., Hansen, N.F., Hayashizaki, Y., Johnson-Hopson, C., Hsuan, V.W., Iida, K., Karnes, M., Khan, S., Koesema, E., Ishida, J., Jiang, P.X., Jones, T., Kawai, J., Kamiya, A., Meyers, C., Nakajima, M., Narusaka, M., Seki, M., Sakurai, T., Satou, M., Tamse, R., Vaysberg, M., Wallender, E.K., Wong, C., Yamamura, Y., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R. 2003. Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* 302:842-846.

Internet Resources

<http://www.invitrogen.com/gateway>

Highly recommended website for more information and troubleshooting related to Invitrogen Gateway Technology.

<http://www.clontech.com/clontech/products/families/creator>

Highly recommended website for more information and troubleshooting related to BD Clontech Creator Technology.

Contributed by Jaehong Park and
Joshua LaBaer
Harvard Medical School
Cambridge, Massachusetts

CHAPTER 4

Preparation and Analysis of RNA

INTRODUCTION

The ability to isolate clean intact RNA from cells is essential for experiments that measure transcript levels, for cloning of intact cDNAs, and for functional analysis of RNA metabolism. RNA isolation procedures frequently must be performed on numerous different cell samples, and therefore are designed to allow processing of multiple samples simultaneously. This chapter begins by describing several methods commonly used to isolate RNA, and concludes with methods used to analyze RNA expression levels and synthesis rates.

The difficulty in RNA isolation is that most ribonucleases are very stable and active enzymes that require no cofactors to function. The first step in all RNA isolation protocols therefore involves lysing the cell in a chemical environment that results in denaturation of ribonuclease. The RNA is then fractionated from the other cellular macromolecules under conditions that limit or eliminate any residual RNase activity. The cell type from which RNA is to be isolated and the eventual use of that RNA will determine which procedure is appropriate. No matter which procedure is used, it is important that the worker use care (e.g., wearing gloves) not to introduce any contamination that might include RNase during work up of the samples, and particularly when the samples are prepared for storage at the final step.

While the RNA isolation protocols describe methods that can be performed using common laboratory reagents, several kits for RNA isolation are commercially available. These kits offer the dual advantage of ease of use and (at least in theory) of reagents that have been tested for effectiveness. These kits frequently work well and are widely used. The disadvantages of using kits are that they are more expensive per sample than isolations that are done using “home made” solutions, and that the kits do not offer flexibility for cell types that require special conditions. The cost disadvantage is frequently outweighed in situations where only a few RNA isolations are performed; however, preparing reagents from scratch can take time, and in the event that any of the reagents are not working properly, troubleshooting will require further time. In situations where numerous samples are routinely processed, significant cost savings can be realized by avoiding the use of kits.

One of the primary uses of RNA isolation procedures is the analysis of gene expression. In order to elucidate the regulatory properties of a gene, it is necessary to know the structure and amount of the RNA produced from that gene. The second part of this chapter is devoted to techniques that are used to analyze RNA. Procedures such as S1 nuclease analysis and ribonuclease protection can be used to do fine-structure mapping of any RNA. These techniques allow characterization of 5' and 3' splice junctions as well as the 5' and 3' ends of RNA. Both of these procedures, as well as northern analysis, can also be used to accurately determine the steady-state level of any particular message.

After determining the steady-state level of a message, many investigators wish to examine whether that level is set by the rate of transcription of the gene. Alterations

in steady-state level might also reflect changes in processing or stability of the RNA. The final section of the chapter describes the “nuclear run-off” technique, which determines the number of active RNA polymerase molecules that are traversing any particular segment of DNA. This procedure is used to analyze directly how the rate of transcription of a gene varies when the growth state of a cell is changed.

Robert E. Kingston

PREPARATION OF RNA FROM EUKARYOTIC AND PROKARYOTIC CELLS

Three methods are presented for preparing RNA from eukaryotic cells. The first two are rapid and can be used to prepare several RNA samples at the same time. As written, they are intended for production of RNA to be used for analysis by S1 nuclease or ribonuclease protection. Modifications of each protocol are given that should be used if intact full-length RNA is a priority.

Both of these protocols require limited hands-on time. The first (*UNIT 4.1*) utilizes a gentle detergent to lyse the cell. Its main advantage is that it requires no high-speed centrifuge spins, thus allowing preparation of numerous samples without having to find several high-speed rotors. In the second protocol (*UNIT 4.2*) cells are lysed using 4 M guanidinium isothiocyanate. This protocol requires very few manipulations, gives clean RNA from many sources, and is the method of choice when working with tissues that have high levels of endogenous RNase. It does require a high-speed centrifuge run, which limits the number of samples that can be prepared at the same time.

In the third protocol (*UNIT 4.3*) the cell is lysed with phenol and SDS. It produces clean, full-length RNA from large quantities of plant cells. This protocol also works well with several mammalian cells and tissues. All three protocols can be used with cells from any higher eukaryote. In particular, many laboratories use the guanidinium procedure when preparing RNA from plant tissue.

These protocols produce total RNA, which contains primarily ribosomal RNA and transfer RNA. Many techniques require messenger RNA that is largely free of contaminating rRNA and tRNA. The isolation of poly(A)⁺ RNA, which is highly enriched for mRNA, is described in *UNIT 4.5*. A protocol is also presented for extracting RNA from gram-negative and gram-positive bacteria (*UNIT 4.4*).

The major source of failure in any attempt to produce RNA is contamination by ribonuclease. RNases are very stable enzymes and generally require no cofactors to function. Therefore, a small amount of RNase in an RNA preparation will create a real problem. To avoid contamination problems, the following precautions can be taken:

1. *Solutions.* Any water or salt solutions used in RNA preparation should be treated with the chemical diethylpyrocarbonate (DEPC). This chemical inactivates ribonucleases by covalent modification. Solutions containing Tris cannot be effectively treated with DEPC because Tris reacts with DEPC to inactivate it. See Reagents and Solutions for instructions regarding DEPC treatment.
2. *Glassware and plastic.* Labware used in the preparation of RNA should be treated to remove residual RNase activity. Autoclaving will not fully inactivate many RNases. Glassware can be baked at 300°C for 4 hr. Certain kinds of plasticware (e.g., some conical centrifuge tubes and pipets) can be rinsed with chloroform to inactivate RNase. When done carefully, this rinse is an effective treatment. Keep in mind, however, that many plastics (e.g., gel boxes) will melt when treated with chloroform. Plasticware straight out of the package is generally free from contamination and can be used as is.
3. Hands are a major source of contaminating RNase. Wear gloves.

Preparation of Cytoplasmic RNA from Tissue Culture Cells

BASIC PROTOCOL

Cells are washed with ice-cold phosphate-buffered saline and kept on ice for all subsequent manipulations. The pellet of harvested cells is resuspended in a lysis buffer containing the nonionic detergent Nonidet P-40. Lysis of the plasma membranes occurs almost immediately. The intact nuclei are removed by a brief microcentrifuge spin, and sodium dodecyl sulfate (SDS) is added to the cytoplasmic supernatant to denature protein. Protein is digested with protease and removed by extractions with phenol/chloroform and chloroform. Cytoplasmic RNA is recovered by ethanol precipitation and quantitated by measuring its absorbance at 260 and 280 nm.

Materials

Cell monolayer or suspension
 Phosphate-buffered saline, ice cold (PBS; *APPENDIX 2*)
 Lysis buffer, ice cold (see recipe)
 20% (w/v) sodium dodecyl sulfate (SDS)
 20 mg/ml proteinase K
 25:24:1 phenol/chloroform/isoamyl alcohol (*UNIT 2.1A*)
 24:1 chloroform/isoamyl alcohol
 3 M DEPC-treated (see recipe) sodium acetate, pH 5.2 (*APPENDIX 2*)
 100% ethanol
 75% ethanol/25% 0.1 M DEPC-treated (see recipe) sodium acetate, pH 5.2
 Beckman JS-4.2 rotor or equivalent
 Rubber policeman

Additional reagents and equipment for removing contaminating DNA (see Support Protocol)

CAUTION: Diethylpyrocarbonate (DEPC) is a suspected carcinogen and should be handled carefully (*APPENDIX 1H*).

IMPORTANT NOTE: Water and sodium acetate should be treated with DEPC (see recipe).

Wash cells

1a. *For monolayer cells:* Rinse cell monolayer three times with ice-cold PBS. Scrape cells into a small volume of cold PBS with a rubber policeman. Transfer to a centrifuge tube on ice. Collect cells by centrifuging in a Beckman JS-4.2 rotor 5 min at $300 \times g$ (1000 rpm), 4°C, or in a microcentrifuge 15 sec at maximum speed, 4°C.

For a 10-cm dish, collect cells in 1 ml PBS. For a 15-cm dish collect in 3 to 5 ml PBS.

1b. *For suspension cultures:* Pellet by centrifuging in a Beckman JS-4.2 rotor 5 min at $300 \times g$ (1000 rpm), 4°C. Resuspend pellet in one-half original culture volume ice-cold PBS. Repeat once.

This procedure, as written, is used for up to 2×10^7 cells (two 10-cm dishes or ~20 ml suspension culture). The procedure can be scaled up for larger cell quantities by increasing volumes appropriately and using larger, conical tubes.

Lyse cells

2. Resuspend cells in 375 μ l ice-cold lysis buffer. Incubate 5 min on ice.

The suspension should clear rapidly, indicating cell lysis.

Cells are best suspended by careful but vigorous vortexing. Avoid foaming.

3. If the cells are not already in a microcentrifuge tube, transfer them into one. Microcentrifuge 2 min at maximum speed, 4°C.
4. Transfer supernatant to a clean tube containing 4 µl of 20% SDS. Mix immediately by vortexing.

The supernatant is the cytoplasmic extract. It is usually slightly cloudy and yellow-white, depending on the cells. The pellet, which contains nuclei and some cell debris, should be considerably smaller than the whole cell pellet obtained in step 1a or 1b and white in color.

5. Add 2.5 µl of 20 mg/ml proteinase K. Incubate 15 min at 37°C.

Extract with phenol/chloroform/isoamyl and chloroform/alcohol

6. Add 400 µl of 25:24:1 phenol/chloroform/isoamyl alcohol, room temperature. Vortex thoroughly—i.e., at least 1 min. Microcentrifuge ≥5 min at maximum speed, room temperature.

With protease treatment, there should be only a small amount of precipitate at the interface between the two phases, although this can vary depending on the cell type. For some cells, the protease step can be safely omitted. In this case, the white precipitate at the interface can be considerable. If a very large precipitate forms after the first organic extraction and little or no aqueous phase can be recovered, first try spinning for a few minutes more. If the precipitate fails to collapse to the interface, remove and discard the organic phase from the bottom of the tube. Add 400 µl chloroform/isoamyl alcohol. Vortex well and spin ~2 min. The precipitate should have largely disappeared. Recover the upper aqueous phase and proceed.

7. Transfer the aqueous (upper) phase to a clean tube, avoiding precipitated material from the interface. Add 400 µl phenol/chloroform/isoamyl alcohol and repeat the extraction.
8. Transfer the aqueous phase to a clean tube. Add 400 µl of 24:1 chloroform/isoamyl alcohol. Vortex 15 to 30 sec and microcentrifuge 1 min at maximum speed, room temperature.
9. Again, transfer the aqueous (upper) phase to a clean tube.

Precipitate RNA

10. Add 40 µl of 3 M DEPC-treated sodium acetate, pH 5.2, and 1 ml of 100% ethanol. Mix by inversion. Incubate 15 to 30 min on ice or store at –20°C overnight.
11. Recover the RNA by microcentrifuging for 15 min at maximum speed, 4°C.
12. If necessary, remove contaminating DNA (see Support Protocol).
13. Rinse the pellet with 1 ml of 75% ethanol/25% 0.1 M sodium acetate, pH 5.2 solution.

Analyze purity

14. Dry and resuspend in 100 µl DEPC-treated water. Dilute 10 µl into 1 ml water to determine the A_{260} and A_{280} . Store the remaining RNA at –70°C.

REMOVAL OF CONTAMINATING DNA

If RNA is isolated from cells transiently transfected with cloned DNA, substantial amounts of this DNA will copurify with the RNA in this procedure. This contaminating DNA will interfere with analysis of the RNA by nuclease protection assays, especially if uniformly labeled probes are used. To remove this DNA, perform the following steps after step 11 of the preparation (see Basic Protocol).

Additional Materials (also see *Basic Protocol*)

TE buffer, pH 7.4 (*APPENDIX 2*)
100 mM MgCl₂/10 mM dithiothreitol (DTT; see *APPENDIX 2* for both components)
2.5 mg/ml RNase-free DNase I (see recipe)
25 to 50 U/μl placental ribonuclease inhibitor (e.g., RNAsin from Promega Biotec)
or vanadyl-ribonucleoside complex
DNase stop mix (see recipe)

1. Redissolve the RNA in 50 μl TE buffer, pH 7.4.
2. Prepare on ice a cocktail containing (per sample) 10 μl of 100 mM MgCl₂/10 mM DTT, 0.2 μl of 2.5 mg/ml RNase-free DNase, 0.1 μl of 25 to 50 U/μl placental ribonuclease inhibitor or vanadyl-ribonucleoside complex, and 39.7 μl TE buffer. Add 50 μl of this cocktail to each RNA sample. Mix and incubate 15 min at 37°C.
3. Stop the DNase reaction by adding 25 μl DNase stop mix.
4. Extract once with phenol/chloroform/isoamyl alcohol and once with chloroform/isoamyl alcohol.
5. Add 325 μl of 100% ethanol and precipitate 15 to 30 min on ice or overnight at -20°C.

Resume Basic Protocol at step 13.

REAGENTS AND SOLUTIONS

Diethylpyrocarbonate (DEPC) treatment of solutions

Add 0.2 ml DEPC to 100 ml of the solution to be treated. Shake vigorously to get the DEPC into solution. Autoclave the solution to inactivate the remaining DEPC. Many investigators keep the solutions they use for RNA work separate to ensure that “dirty” pipets do not go into them.

CAUTION: *Wear gloves and use a fume hood when using DEPC, as it is a suspected carcinogen (also see APPENDIX 1H).*

DNase stop mix

50 mM EDTA
1.5 M sodium acetate
1% (w/v) sodium dodecyl sulfate (SDS)

The SDS may come out of solution at room temperature. Heat briefly to redissolve.

Lysis buffer

50 mM Tris·Cl, pH 8.0 (*APPENDIX 2*)
100 mM NaCl
5 mM MgCl₂
0.5% (v/v) Nonidet P-40
Prepare with DEPC-treated H₂O (see recipe)
Filter sterilize

If the RNA is to be used for northern blot analysis or the cells are particularly rich in ribonuclease, add ribonuclease inhibitors to the lysis buffer: 1000 U/ml placental ribonuclease inhibitor (e.g., RNAsin) plus 1 mM DTT or 10 mM vanadyl-ribonucleoside complex.

RNase-free DNase I

Commercially prepared enzymes such as Worthington grade DPRF are satisfactory. If supplied as a powder, redissolve in TE buffer containing 50% (v/v) glycerol and store at -20°C . See also *UNIT 4.10*, reagents and solutions, for homemade preparation of RNase-free DNase.

COMMENTARY

Background Information

Most procedures for isolating RNA from eukaryotic cells involve lysing and denaturing cells to liberate total nucleic acids. Additional steps are then required to remove DNA. This procedure allows rapid preparation of total cytoplasmic RNA by using a nonionic detergent to lyse the plasma membrane, leaving the nuclei intact. The nuclei and hence the bulk of the cellular DNA are then removed with a simple brief centrifugation.

Variations of this technique are in wide use, and its precise origins are obscure. Versions of this procedure were used in some of the early S1 nuclease mapping papers (Berk and Sharp, 1977; Favoloro et al., 1980). The protocol described here is a considerable simplification of the earlier methods, omitting, for example, removal of nuclei by centrifugation through sucrose. It is fast and streamlined, designed for preparing total cytoplasmic RNA from many cultures simultaneously for nuclease protection analysis. It is scaled for small cultures—i.e., 1 to 2 dishes of adherent cells or 10 to 20 ml of suspension culture.

The Basic Protocol works well for many cell types. The protocol takes no special precautions for ribonucleases and may not yield northern blot-quality RNA from some cells. If full-length RNA is required, ribonuclease inhibitors should be added to the lysis buffer (as described; see Reagents and Solutions) or the guanidinium isothiocyanate method should be used (*UNIT 4.2*). Finally, if RNA is isolated from transiently transfected cells, the RNA should be further treated with deoxyribonuclease to remove transfected DNA (see Support Protocol). This modification is especially critical if the RNA is to be assayed by nuclease protection using uniformly labeled probes.

Troubleshooting

Degradation of RNA by ribonuclease is best avoided by working quickly and keeping everything cold until SDS is added to the cytoplasmic

extract. For cells with which ribonuclease is a problem, inhibitors can be added to the lysis buffer (see Reagents and Solutions), but in most cases this is unnecessary.

Note that for some cell lines, it may be possible to omit the proteinase K step and to proceed directly to organic extraction after removal of the nuclei and addition of SDS.

DNA contamination is only a problem when preparing RNA from cells transfected with cloned DNA in a transient expression assay. In this case, add the DNase digestion steps outlined in the Support Protocol.

Anticipated Results

Yields vary widely, depending on the cell line. Expect 30 to 100 μg from a confluent 10-cm dish of most fibroblast lines or 1×10^7 lymphoid cells. Ratios of A_{260} to A_{280} should fall in the range 1.7 to 2.0. RNA at 1 mg/ml has an A_{260} of 25.

Time Considerations

Depending on the number of samples being processed, it is possible to proceed from harvesting the cells to the ethanol precipitation step in 1 to 2 hr. This is the best interim stopping point. The RNA may be recovered, redissolved, and quantitated later the same day or the following day.

Literature Cited

- Berk, A. J. and Sharp, P. A. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* 12: 721-732.
- Favoloro, J., Treisman, R., and Kamen, R. 1980. Transcription maps of polyoma virus-specific RNA: Analysis by two-dimensional nuclease S1 gel mapping. *Meth. Enzymol.* 65:718-749.

Contributed by Michael Gilman
Cold Spring Harbor Laboratory
Cold Spring Harbor, New York

Guanidine Methods for Total RNA Preparation

Three different methods for RNA preparation using guanidine are presented in this unit — a single-step isolation method employing liquid-phase separation to selectively extract total RNA from tissues and cultured cells (see Basic Protocol) and two methods that rely on a CsCl step gradient to isolate total RNA (see Alternate Protocols 1 and 2).

SINGLE-STEP RNA ISOLATION FROM CULTURED CELLS OR TISSUES

BASIC PROTOCOL

Cultured cells or tissues are homogenized in a denaturing solution containing 4 M guanidine thiocyanate. The homogenate is mixed sequentially with 2 M sodium acetate (pH 4), phenol, and finally chloroform/isoamyl alcohol or bromochloropropane. The resulting mixture is centrifuged, yielding an upper aqueous phase containing total RNA. In this single-step extraction the total RNA is separated from proteins and DNA that remain in the interphase and in the organic phase. Following isopropanol precipitation, the RNA pellet is redissolved in denaturing solution (containing 4 M guanidine thiocyanate), reprecipitated with isopropanol, and washed with 75% ethanol.

Materials

- Denaturing solution (see recipe)
- 2 M sodium acetate, pH 4 (see recipe)
- Water-saturated phenol (see recipe)
- 49:1 (v/v) chloroform/isoamyl alcohol or bromochloropropane
- 100% isopropanol
- 75% ethanol (prepared with DEPC-treated water; UNIT 4.1)
- DEPC-treated water (UNIT 4.1) or freshly deionized formamide (see recipe)
- Glass Teflon homogenizer
- 5-ml polypropylene centrifuge tube
- Sorvall SS-34 rotor (or equivalent)

CAUTION: Phenol is a poison and causes burns. When handling phenol, use gloves and eye protection.

NOTE: Carry out all steps at room temperature unless otherwise stated.

Homogenize cells

- 1a. *For tissue:* Add 1 ml denaturing solution per 100 mg tissue and homogenize with a few strokes in a glass Teflon homogenizer.
- 1b. *For cultured cells:* Either centrifuge suspension cells and discard supernatant, or remove the culture medium from cells grown in monolayer cultures. Add 1 ml denaturing solution per 10^7 cells and pass the lysate through a pipet seven to ten times.

Do not wash cells with saline. Cells grown in monolayer cultures can be lysed directly in the culture dish or flask.

The procedure can be carried out in sterile, disposable, round-bottom polypropylene tubes with caps; no additional treatment of the tubes is necessary. Before using, test if the tubes can withstand centrifugation at $10,000 \times g$ with the mixture of denaturing solution and phenol/chloroform.

2. Transfer the homogenate into a 5-ml polypropylene tube. Add 0.1 ml of 2 M sodium acetate, pH 4, and mix thoroughly by inversion. Add 1 ml water-saturated phenol,

Preparation and Analysis of RNA

4.2.1

mix thoroughly, and add 0.2 ml of 49:1 chloroform/isoamyl alcohol or bromochloropropane. Mix thoroughly and incubate the suspension 15 min at 0° to 4°C.

Make sure that caps are tightly closed when mixing. The volumes used are per 1 ml denaturing solution.

Bromochloropropane is less toxic than chloroform and its use for phase separation decreases possibility of contaminating RNA with DNA (Chomczynski and Mackey, 1995).

3. Centrifuge 20 min at $10,000 \times g$ (9000 rpm in SS-34 rotor), 4°C. Transfer the upper aqueous phase to a clean tube.

The upper aqueous phase contains the RNA, whereas the DNA and proteins are in the interphase and lower organic phase. The volume of the aqueous phase is ~1 ml, equal to the initial volume of denaturing solution.

Isolate RNA

4. Precipitate the RNA by adding 1 ml (1 vol) of 100% isopropanol. Incubate the samples 30 min at -20°C . Centrifuge 10 min at $10,000 \times g$, 4°C, and discard supernatant.

For isolation of RNA from tissues with a high glycogen content (e.g., liver), a modification of the single-step method is recommended to diminish glycogen contamination (Puissant and Houdebine, 1990). Following this isopropanol precipitation, wash out glycogen from the RNA pellet by vortexing in 4 M LiCl. Sediment the insoluble RNA 10 min at $5000 \times g$. Dissolve the pellet in denaturing solution and follow the remainder of the protocol.

5. Dissolve the RNA pellet in 0.3 ml denaturing solution and transfer into a 1.5-ml microcentrifuge tube.
6. Precipitate the RNA with 0.3 ml (1 vol) of 100% isopropanol for 30 min at -20°C . Centrifuge 10 min at $10,000 \times g$, 4°C, and discard supernatant.
7. Resuspend the RNA pellet in 75% ethanol, vortex, and incubate 10 to 15 min at room temperature to dissolve residual amounts of guanidine contaminating the pellet.
8. Centrifuge 5 min at $10,000 \times g$, 4°C, and discard supernatant. Dry the RNA pellet in a vacuum for 5 min.

Do not let the RNA pellet dry completely, as this greatly decreases its solubility. Avoid drying the pellet by centrifugation under vacuum. Drying is not necessary for solubilization of RNA in formamide.

9. Dissolve the RNA pellet in 100 to 200 μl DEPC-treated water or freshly deionized formamide by passing the solution a few times through a pipe tip. Incubate 10 to 15 min at 55° to 60°C . Store RNA dissolved in water at -70°C and RNA dissolved in formamide at either -20° or -70°C .

RNA dissolved in formamide is protected from degradation by RNase and can be used directly for formaldehyde-agarose gel electrophoresis in northern blotting (Chomczynski, 1992). However, before use in RT-PCR, RNA should be precipitated from formamide by adding 4 vol ethanol and centrifuging 5 min at $10,000 \times g$.

Quantitate RNA

10. Quantitate RNA by diluting 5 μl in 1 ml alkaline water and reading the A_{260} and A_{280} (APPENDIX 3D).

Water used for spectrophotometric measurement of RNA should have $\text{pH} > 7.5$. Acidic pH affects the UV absorption spectrum of RNA and significantly decreases its A_{260}/A_{280} ratio (Willfinger et al., 1996). Typically, distilled water has $\text{pH} < 6$. Adjust water to a slightly alkaline pH by adding concentrated Na_2HPO_4 solution to a final concentration of 1 mM.

CsCl PURIFICATION OF RNA FROM CULTURED CELLS

ALTERNATE PROTOCOL 1

Cells are washed free of medium and are then lysed by placing them in a 4 M guanidine solution. The viscosity of the solution is reduced by drawing the lysate through a 20-G needle and the RNA is pelleted through a CsCl step gradient. The supernatant fluid from this gradient is then carefully removed to allow complete separation of RNA, found in the pellet, from contaminating DNA and protein. Finally, the RNA in the pellet is dissolved, ethanol precipitated, and quantitated spectrophotometrically at A_{260} .

Additional Materials (also see Basic Protocol)

Phosphate-buffered saline (PBS; *APPENDIX 2*)

Guanidine solution (see recipe)

5.7 M cesium chloride (CsCl), DEPC-treated (see recipe)

TES solution (see recipe)

3 M sodium acetate, pH 5.2 (*APPENDIX 2*)

100% ethanol

Rubber policeman

6-ml syringe with 20-G needle

Beckman JS-4.2 and SW 55 rotors (or equivalents)

13 × 51-mm silanized (*APPENDIX 3B*) and autoclaved polyallomer ultracentrifuge tube

Additional reagents and equipment for quantitating RNA (*APPENDIX 3D*)

CAUTION: DEPC is a suspected carcinogen and should be handled carefully.

NOTE: The following solutions should be treated with DEPC to inhibit RNase activity: sodium acetate, water, and 5.7 M CsCl (*UNIT 4.1*).

NOTE: Carry out steps 1 to 4 at room temperature.

Lyse the cells

For monolayer culture

- 1a. Wash cells at room temperature by adding 5 ml PBS per dish, swirling dishes, and pouring off. Repeat wash.
- 2a. Add 3.5 ml guanidine solution for $\leq 10^8$ cells, dividing the solution equally between the dishes. The cells should immediately lyse in place. Recover the viscous lysate by scraping the dishes with a rubber policeman. Remove lysate from dishes using a 6-ml syringe with 20-G needle. Combine lysates.

For suspension culture

- 1b. Pellet $\leq 10^8$ cells by centrifuging 5 min at $300 \times g$ (1000 rpm in JS-4.2 rotor), room temperature, and discarding supernatant. Wash cells once at room temperature by resuspending the pellet in an amount of PBS equal to half the original volume and centrifuging. Discard supernatant.
- 2b. Add 3.5 ml guanidine solution to the centrifuge tube.
3. Draw the resultant extremely viscous solution up and down four times through a 6-ml syringe with 20-G needle. Transfer the solution to a clean tube.

It is critical that chromosomal DNA be sheared in this step in order to reduce viscosity. This allows complete removal of the DNA in the centrifugation step.

Isolate the RNA

4. Place 1.5 ml of 5.7 M CsCl in a 13 × 51–mm silanized and autoclaved polyallomer ultracentrifuge tube. Layer 3.5 ml of cell lysate on top of CsCl cushion to create a step gradient. The interface should be visible.

Silanizing the tube decreases the amount of material that sticks to the sides of the tube and thus decreases the level of contamination of the final RNA.

5. Centrifuge 12 to 20 hr at $150,000 \times g$ (35,000 rpm in SW 55 rotor), 18°C. Set centrifuge for slow acceleration and deceleration in order to avoid disturbing the gradient.
6. Remove the supernatant very carefully (see Fig. 4.2.1). Place the end of the Pasteur pipet at the top of the solution and lower it as the level of the solution lowers. Leave ~100 μ l in the bottom, invert the tube carefully, and pour off the remaining liquid.

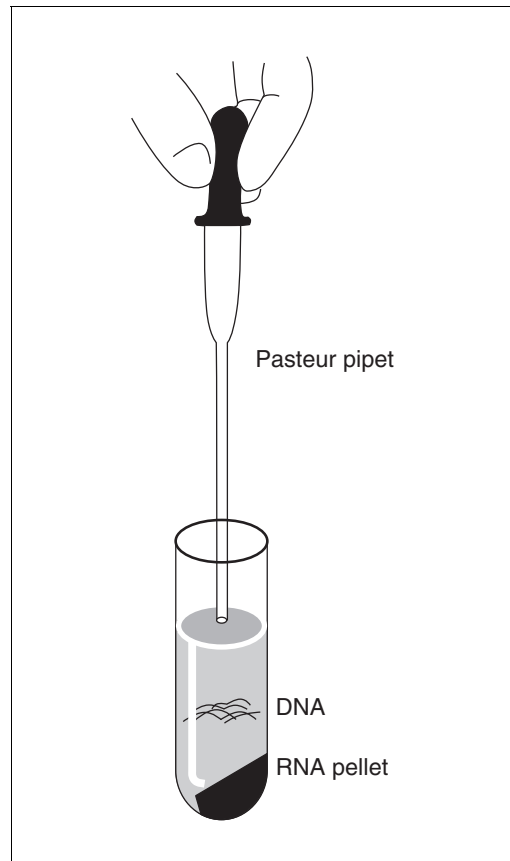
There should be a white band of DNA at the interface—care must be taken to remove this band completely, as it contains cellular DNA.

7. Allow the pellet to drain 5 to 10 min, then resuspend it in 360 μ l TES solution by repeatedly drawing the solution up and down in a pipet. Allow the pellet to resuspend 5 to 10 min at room temperature. Transfer to a clean microcentrifuge tube.

It is critical to allow ample time for resuspension of this pellet or the yield of RNA will be significantly decreased.

8. Add 40 μ l of 3 M sodium acetate, pH 5.2, and 1 ml of 100% ethanol. Precipitate the RNA 30 min on dry ice/ethanol. Microcentrifuge 10 to 15 min at 4°C and discard supernatant.

Figure 4.2.1 Technique for removing supernatant from a CsCl step gradient.



9. Resuspend the pellet in 360 μ l water and repeat step 8.

RNA dissolves more readily in water than in a salt solution.

Quantitate the RNA

10. Drain the pellet 10 min and dissolve in \sim 200 μ l water. Quantitate by diluting 10 μ l to 1 ml in alkaline water and reading the A_{260} and A_{280} (see Basic Protocol, step 10, and APPENDIX 3D). Store RNA at -70°C either as an aqueous solution or as an ethanol precipitate.

This protocol produces RNA that is clean enough for northern, S1, or SP6 analysis. If cleaner RNA is desired, step 7 can be modified with the following: After resuspending the pellet in TES solution, extract with 360 μ l of 4:1 (v/v) chloroform/1-butanol and save the supernatant. Extract the chloroform by adding 360 μ l TES solution. Combine the supernatants, add 0.1 vol of 3 M sodium acetate, pH 5.2, and ethanol precipitate as in step 8.

CsCl PURIFICATION OF RNA FROM TISSUE

Additional precautions must be taken when purifying RNA from tissue, as certain organs such as pancreas and spleen have very high endogenous levels of RNase. (Liver and intestine, however, have relatively low levels.) This protocol was originally described in Chirgwin et al. (1979) and modified by Richard Selden.

Additional Materials (also see Alternate Protocol 1)

Liquid nitrogen
Tissue guanidine solution (see recipe)
20% (w/v) *N*-lauroylsarcosine (Sarkosyl)
Cesium chloride (CsCl)
Tissue resuspension solution (see recipe)
25:24:1 phenol/chloroform/isoamyl alcohol (UNIT 2.1)
24:1 chloroform/isoamyl alcohol
Tissuemizer
Sorvall SS-34 and Beckman SW 28 rotors (or equivalents)
SW 28 polyallomer tube silanized (APPENDIX 3B) and autoclaved

1. Rapidly remove tissue from the animal and quick-freeze it in liquid nitrogen.

The sample should be removed from the animal in pieces \leq 2 g or it will be difficult to do the further workup.

RNA is very unstable in tissue once removed from the body so it is critical to quick-freeze the tissue. Placing the tissue in guanidine and then waiting to grind it will result in degraded RNA.

2. Add 20 ml tissue guanidine solution for \sim 2 g of tissue. Immediately grind the tissue in a tissuemizer with two or three 10-sec bursts for complete grinding.

Tissue guanidine solution, unlike the guanidine solution used in the Basic Protocol, lacks Sarkosyl. It is important that Sarkosyl not be present at this stage or a frothy mess will result.

3. Centrifuge 10 min at $12,000 \times g$ (10,000 rpm in SS-34 rotor), 12°C .
4. Collect the supernatant and add 0.1 vol of 20% Sarkosyl. Heat 2 min at 65°C .
5. Add 0.1 g CsCl/ml of solution, dissolve the CsCl, then layer the sample over 9 ml of 5.7 M CsCl in an SW 28 silanized, autoclaved polyallomer tube. Centrifuge overnight at $113,000 \times g$ (25,000 rpm in SW 28 rotor), 22°C .

ALTERNATE PROTOCOL 2

Preparation and Analysis of RNA

4.2.5

6. Carefully remove the supernatant (see Alternate Protocol 1, step 6, and Fig. 4.2.1). Invert the tube to drain. Cut off bottom of tube (containing RNA pellet) and place it in a 50-ml plastic tube.
7. Add 3 ml tissue resuspension buffer and allow pellet to resuspend overnight at 4°C.
It is difficult to resuspend this pellet. Occasionally the sample may have to be left longer than overnight. The high concentrations of 2-mercaptoethanol and Sarkosyl prevent RNA degradation during this resuspension.
8. Extract the solution sequentially with 25:24:1 phenol/chloroform/isoamyl alcohol, then with 24:1 chloroform/isoamyl alcohol (see UNIT 2.1).
9. Add 0.1 vol of 3 M sodium acetate, pH 5.2, and 2.5 vol of 100% ethanol. Precipitate RNA 30 min on dry ice/ethanol, microcentrifuge 10 to 15 min at 4°C, discard supernatant, and resuspend in water. Quantitate the RNA and store (see Alternate Protocol 1, step 10).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

5.7 M CsCl, DEPC-treated

Dissolve CsCl in 0.1 M EDTA, pH 8.0. Add 0.002 vol DEPC, shake 20 to 30 min, and autoclave. Weigh the bottle of solution before and after autoclaving and make up the weight lost to evaporation during autoclaving with DEPC-treated water (UNIT 4.1) to ensure that the solution is actually 5.7 M when used.

Denaturing solution

Stock solution: Mix 293 ml water, 17.6 ml of 0.75 M sodium citrate, pH 7.0, and 26.4 ml of 10% (w/v) *N*-lauroylsarcosine (Sarkosyl). Add 250 g guanidine thiocyanate and stir at 60° to 65°C to dissolve. Store up to 3 months at room temperature.

Working solution: Add 0.35 ml 2-mercaptoethanol (2-ME) to 50 ml of stock solution. Store up to 1 month at room temperature.

Final concentrations are 4 M guanidine thiocyanate, 25 mM sodium citrate, 0.5% Sarkosyl, and 0.1 M 2-ME.

Formamide

Prepare freshly deionized formamide by stirring with 1 g AG 501-X8 ion-exchange resin (Bio-Rad) per 10 ml formamide for 30 min and filter at room temperature. Alternatively, use a commercially available stabilized, ultrapure formamide (Formazol, Molecular Research Center).

Guanidine solution

Mix 550 ml water with 1.64 g sodium acetate (anhydrous) and 472.8 g guanidine thiocyanate, and stir to dissolve, heating slightly (to 65°C) if necessary to get the guanidine into solution. Add 15.4 mg dithiothreitol (DTT) and 50 ml of 10% (w/v) *N*-lauroylsarcosine (Sarkosyl). Adjust pH to ~5.5 with acetic acid, dilute solution to 1 liter with water, and filter through a Nalgene filter. Store up to one month at room temperature.

Final concentrations are 4 M guanidine isothiocyanate, 20 mM sodium acetate, 0.5% Sarkosyl, and 0.1 mM DTT.

Sodium acetate, 2 M

Add 16.42 g sodium acetate (anhydrous) to 40 ml water and 35 ml glacial acetic acid. Adjust solution to pH 4 with glacial acetic acid and dilute to 100 ml final with

continued

water (solution is 2 M with respect to sodium ions). Store up to 1 year at room temperature.

TES solution

10 mM Tris-Cl, pH 7.4
5 mM EDTA
1% (w/v) SDS
Store up to 1 year at room temperature

Tissue guanidine solution

Dissolve 590.8 g guanidine thiocyanate in ~400 ml DEPC-treated water (*UNIT 4.1*). Add 25 ml of 2 M Tris-Cl, pH 7.5, and 20 ml of 0.5 M Na₂EDTA, pH 8.0 (*APPENDIX 2*). Stir overnight. Adjust the volume to 950 ml and filter. Finally, add 50 ml of 2-mercaptoethanol. Store up to three months at room temperature.

Tissue resuspension solution

5 mM EDTA
0.5% (w/v) *N*-lauroylsarcosine (Sarkosyl)
5% (v/v) 2-mercaptoethanol
Store up to 1 month at room temperature

Water-saturated phenol

Dissolve 100 g phenol crystals in water at 60° to 65°C. Aspirate the upper water phase and store up to 1 month at 4°C.

Do not use buffered phenol in place of water-saturated phenol.

COMMENTARY

Background Information

Guanidine thiocyanate is one of the most effective protein denaturants known. The use of guanidine to lyse cells was originally developed to allow purification of RNA from cells high in endogenous ribonucleases (Cox, 1968; Ullrich et al., 1977; Chirgwin et al., 1979).

The single-step method of RNA isolation described in the Basic Protocol is based on the ability of RNA to remain water soluble in a solution containing 4 M guanidine thiocyanate, pH 4, in the presence of a phenol/chloroform organic phase. Under such acidic conditions, most proteins and small fragments of DNA (50 bases to 10 kb) will be found in the organic phase while larger fragments of DNA and some proteins remain in the interphase. The fragmentation of DNA during homogenization helps to remove DNA from the water phase.

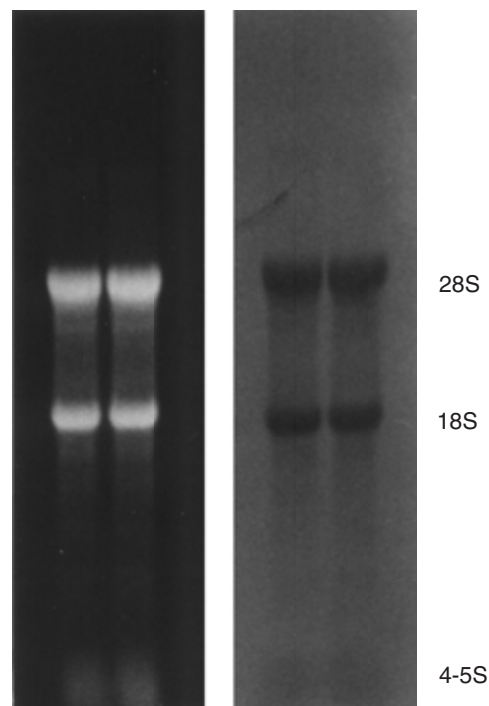
Since its introduction (Chomczynski and Sacchi, 1987), the single-step method has become widely used for isolating RNA from a large number of samples. In addition, the procedure permits recovery of total RNA from small quantities of tissue or cells, making it suitable for gene expression studies whenever the quantity of tissue or cells available is limited. The protocol presented here is an updated version of the original method that further

shortens the time for RNA isolation. All commercial application of the method is restricted by a U.S. patent (Chomczynski, 1989).

The two alternate protocols present methods, based on the observed fact that RNA is denser than DNA or protein, for separating RNA from other cellular macromolecules in the guanidine lysate on a CsCl step gradient (Glisin et al., 1974). A method using hot phenol and guanidine thiocyanate has also been described (Ferimisco et al., 1982).

In Alternate Protocol 1, cultured cells are lysed with a solution that contains 4 M guanidine as well as a mild detergent. This lysis is virtually instantaneous and the cells are thus rapidly taken from an intact state to a completely denaturing environment. In Alternate Protocol 2, tissues are homogenized in a guanidine solution without detergent. The protocols then take advantage of the fact that RNA can be separated from DNA and protein by virtue of its greater density. These protocols have received widespread use because they require very few manipulations. This increases the chance of producing intact RNA and reduces hands-on time for the experimenter. The disadvantage is that they require an ultracentrifuge and rotor, which generally limits the number of samples that can easily be processed

Figure 4.2.2 Rat liver RNA (5 μ g) isolated using the Basic Protocol was electrophoresed in a formaldehyde 1% agarose gel containing ethidium bromide (left), transferred to a hybridization membrane and stained with methylene blue stain (Molecular Research Center; Herrin and Schmidt, 1988, right). Shown are 28S (4.7 kb) and 18S (1.9 kb) ribosomal RNAs, and 4S to 5S (0.10 to 0.15 kb) RNA containing mix of tRNA and 5S ribosomal RNA.



simultaneously. These protocols should be used when very high quality RNA from a limited number of samples is required.

There are several commercial kits for total RNA isolation utilizing guanidine-based methods, the majority based on the single-step method. They can be divided into two groups. The first group, exemplified by the RNA Isolation Kit from Stratagene, includes kits containing denaturing solution, water-saturated phenol, and sodium acetate buffer prepared according to the single-step protocol described here (see Basic Protocol). The use of these kits saves the time needed to make components of the single-step method, but at a substantially higher price. The second group of kits is based on a commercial version of the single-step method combining denaturing solution, phenol, and buffer in a single monophasic solution. These kits offer an improved yield and shorter RNA isolation time (Chomczynski and Mackey, 1995). In this second group, the authors have tested and can recommend the following kits: Isogen (Nippon Gene), RNA-Stat 60 (Tel-Test), RNAzol B (Cinna Scientific), Tri-Pure Isolation Reagent (Boehringer Mannheim), TRI Reagent (Molecular Research Center), and TRIzol Reagent (Life Technologies). All the kits in the second group, except RNAzol B, allow simultaneous isolation of DNA and proteins from a sample used for RNA isolation.

Critical Parameters

As with any RNA preparative procedure, care must be taken to ensure that solutions are free of ribonuclease. Solutions that come into contact with the RNA after adding the guanidine solution are all treated with DEPC, with the exception of the TES solution (Tris inactivates DEPC). Most investigators wear gloves at all times when working with RNA solutions, as hands are a likely source of ribonuclease contamination (see introduction to Chapter 4).

The two Alternate Protocols rely on a thorough separation of DNA and protein from RNA in the step gradient. The use of silanized tubes, as well as careful technique when removing the supernatant, are important. Finally, low yields may result from failing to allow sufficient time for resuspension of the RNA pellet after centrifugation. This pellet is not readily soluble, and sufficient time and vortexing should be allowed to dissolve it.

There are two important points to consider when using the single-step protocol. First, fresh tissue is preferable for RNA isolation. Alternatively, tissue should be frozen immediately in liquid nitrogen and stored at -70°C . In the latter case, tissue should be pulverized in liquid nitrogen and homogenized, using a Polytron or Waring blender, in denaturing solution without thawing. Second, it is important not to let the final RNA pellet dry completely, as that will greatly decrease its solubility. This is critical in

all RNA isolation methods. Partially solubilized RNA has an A_{260}/A_{280} ratio <1.6 . Solubility of RNA can be improved by heating at 55° to 60°C with intermittent vortexing or by passing the RNA solution through a pipet tip.

Anticipated Results

The single-step method yields the whole spectrum of RNA molecules, including small (4S to 5S) RNAs. The amount of isolated RNA depends on the tissue used for isolation. Typically, 100 to 150 μg of total RNA is isolated from 100 mg of muscle tissue and up to 800 μg is isolated from 100 mg of liver. The yield of total RNA from 10^7 cultured cells ranges from 50 to 80 μg for fibroblasts and lymphocytes and 100 to 120 μg for epithelial cells. The A_{260}/A_{280} ratio of the isolated RNA is >1.8 .

The electrophoretic pattern of RNA isolated by the single-step method is exemplified in Figure 4.2.2 which shows the results of formaldehyde-agarose gel electrophoresis of rat liver RNA.

Time Considerations

The isolation of total RNA by the single-step method can be completed in <4 hr. The procedure can be interrupted at one of the isopropanol precipitations or at the ethanol wash steps. Store samples at -20°C if the procedure is interrupted at these steps. Avoid keeping samples in denaturing solution for >30 min.

In the alternate protocols, harvesting the RNA and setting up the gradient takes very little time (~ 1 hr for six samples) and is conveniently done in the evening, allowing the high-speed centrifuge run to go overnight. In a pinch, the guanidine cell lysate can be quick frozen in dry ice/ethanol and stored at -70°C . When the RNA is dissolved after the gradient, it can be stored as an ethanol precipitate indefinitely at any of the precipitation steps. The entire protocol requires 2 to 3 hr of hands-on time for 6 to 12 samples.

Literature Cited

- Chirgwin, J.J., Przybyla, A.E., MacDonald, R.J., and Rutter, W.J. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18:5294.
- Chomczynski, P. 1989. Product and process for isolating RNA. U.S. Patent #4,843,155.
- Chomczynski, P. 1992. Solubilization in formamide protects RNA from degradation. *Nucl. Acids Res.* 20:3791-3792.

- Chomczynski, P. and Mackey, K. 1995. Substitution of chloroform by bromochloropropane in the single-step method of RNA isolation. *Anal. Biochem.* 225:163-164.
- Chomczynski, P. and Sacchi, N. 1987. Single-step method of RNA isolation by acid guanidine thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156-159.
- Cox, R.A. 1968. The use of guanidine chloride in the isolation of nucleic acids. *Methods Enzymol.* 12:120-129.
- Ferimisco, J.R., Smart, J.E., Burrige, K., Helfman, D.M., and Thomas, G.P. 1982. Co-existence of vinculin and a vinculin-like protein of higher molecular weight in smooth muscle. *J. Biol. Chem.* 257:11024-11031.
- Glisin, V., Crkvenjakov, R., and Byus, C. 1974. Ribonucleic acid isolated by cesium chloride centrifugation. *Biochemistry* 13:2633.
- Herrin, D.L. and Schmidt, G.W. 1988. Rapid, reversible staining of northern blots prior to hybridization. *BioTechniques* 6:196-200.
- Puissant, C. and Houdebine, L.M. 1990. An improvement of the single-step method of RNA isolation by acid guanidine thiocyanate-phenol-chloroform extraction. *BioTechniques* 8:148-149.
- Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W.J., and Goodman, H.M. 1977. Rat insulin genes: Construction of plasmids containing the coding sequences. *Science* 196:1313.
- Wilfinger, W.W., Mackey, K. and Chomczynski, P. 1997. Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. *BioTechniques* 22:474-476.

Key References

- Chirgwin et al., 1979. See above.
Describes the use of guanidine to lyse cells.
- Chomczynski and Sacchi, 1978. See above.
Original description of the single-step method.

Contributed by Robert E. Kingston
(CsCl isolation)
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

Piotr Chomczynski (single-step isolation)
University of Cincinnati College of Medicine
Cincinnati, Ohio

Nicoletta Sacchi (single-step isolation)
Laboratory of Molecular Oncology
National Cancer Institute
Frederick, Maryland

Phenol/SDS Method for Plant RNA Preparation

UNIT 4.3

**BASIC
PROTOCOL**

This protocol is divided into two stages: (1) lysis of the cells and removal of proteins by phenol/SDS extraction, and (2) separation of RNA from DNA and other impurities by selective precipitation using LiCl.

Materials

Diethylpyrocarbonate (DEPC)
Liquid nitrogen
Grinding buffer
Phenol equilibrated with TLE solution
Chloroform
8 M and 2 M LiCl
3 M sodium acetate
100% ethanol

Polytron (Brinkmann PT 10/35)
Beckman JA-10, JA-20, and JA-14 rotors (or equivalent)
50-ml Oak Ridge tube
Sarstedt tube

Sodium acetate, water, and LiCl solutions should be treated with DEPC to inhibit RNase activity. See *UNIT 4.1*, reagents and solutions, for instructions. **CAUTION:** *DEPC is a suspected carcinogen and should be handled carefully.*

Homogenize tissue and extract protein

1. Cool a mortar and pestle by pouring a little liquid nitrogen over it.
2. Weigh 15 g frozen plant tissue. If using freshly harvested tissue, quick-freeze in liquid nitrogen.

The time between harvesting of the tissue and freezing should be minimized; once tissue is frozen, work quickly so it does not have a chance to thaw.

3. Grind plant tissue in the mortar and pestle until tissue becomes a fine powder.

Add liquid nitrogen as needed to keep tissue frozen.

4. Immediately transfer to a 500-ml beaker containing 150 ml grinding buffer plus 50 ml TLE-equilibrated phenol.
5. Homogenize the mixture with Polytron for ~2 min at moderate speed (setting 5-6).
6. Add 50 ml chloroform. Use Polytron at low speed to mix in the added chloroform.

It is not necessary to add isoamyl alcohol to the chloroform for the following extractions.

7. Pour the slurry into a 500-ml Nalgene centrifuge bottle and heat 20 min at 50°C.

All extractions involving TLE-equilibrated phenol and chloroform should be done in screw-cap tubes or bottles resistant to those chemicals.

8. Centrifuge mixture 20 min at 10,000 rpm (17,700 × g), 4°C, in JA-10 rotor.
9. Take off as much aqueous layer as possible without disturbing the interface and transfer it to a clean 500-ml Nalgene bottle. Add 50 ml TLE-equilibrated phenol to this aqueous layer, shake bottle to mix phenol and aqueous phase, then add 50 ml chloroform.

TLE-equilibrated phenol and chloroform are added to the freshly removed aqueous layer

**Preparation and
Analysis of RNA**

4.3.1

to reduce possibility of degradation of RNA while steps 10 and 11 are performed.

10. Remove remaining aqueous layer together with interface from initial phenol extraction and transfer it to a 50-ml Oak Ridge tube. Centrifuge this material 20 min at 10,000 rpm ($17,700 \times g$), 4°C, in JA-20 rotor.

11. Remove aqueous layer and combine with aqueous phase already separated in step 9.

Steps 10 and 11 are done to recover the large volume of aqueous phase that is difficult to separate from the interface in a 500-ml bottle.

12. Vigorously shake the 500-ml bottle containing the combined aqueous layers to mix TLE-equilibrated phenol and chloroform with aqueous phase. Centrifuge mixture 15 min at 10,000 rpm ($17,700 \times g$), 4°C, in JA-10 rotor and remove aqueous layer to fresh 500-ml bottle.

13. Reextract the aqueous phase with TLE-equilibrated phenol and chloroform until no interface is obtained (usually a total of three extractions).

The interface should be small on these steps, so recentrifuging as in steps 10 and 11 is not necessary.

14. Extract aqueous phase one last time with chloroform.

This removes traces of TLE-equilibrated phenol in the aqueous layer which can cause problems with the lithium chloride precipitation.

Selectively precipitate RNA

These steps result in removal of contaminating DNA. If there is no need to remove DNA—for example, when poly(A)⁺ selection is the next step—simply ethanol precipitate the nucleic acid.

15. Transfer the aqueous phase to a clean 250-ml Nalgene bottle and add 8 M LiCl ($\frac{1}{3}$ vol) to bring solution to a final concentration of 2 M LiCl. Precipitate overnight at 4°C.

16. Collect precipitate by centrifugation for 20 min at 10,000 rpm ($15,300 \times g$), 4°C, in JA-14 rotor. Rinse pellet with a few milliliters of 2 M LiCl.

17. Resuspend pellet in 5 ml water and transfer to a 15-ml Sarstedt tube. Add 8 M LiCl to bring concentration of LiCl to 2 M and precipitate the RNA at 4°C for at least 2 hr.

18. Recover the RNA by centrifugation for 20 min at 10,000 rpm ($12,100 \times g$), 4°C, in JA-20 rotor. Rinse pellet with 2 M LiCl.

19. Resuspend RNA pellet in 2 ml water. Add 200 μ l of 3 M sodium acetate and 5.5 ml 100% ethanol. Precipitate at -20° overnight or in dry ice/ethanol for 30 min.

RNA can be stored in ethanol at -20° or -70° C indefinitely.

20. Recover RNA by centrifugation for 15 min at 10,000 rpm ($17,700 \times g$), 4°C, in JA-20 rotor. Resuspend RNA in 1 ml water. Dilute 10 μ l to 1 ml and measure the A_{260} and A_{280} . $1 \text{ OD}_{260} = 40 \mu\text{g/ml RNA}$.

REAGENTS AND SOLUTIONS

Grinding buffer

0.18 M Tris
0.09 M LiCl
4.5 mM EDTA
1% sodium dodecyl sulfate (SDS)
pH to 8.2 with HCl

This buffer is equivalent to TLE solution with $1/10$ vol 10% SDS added.

Phenol

Equilibrate freshly liquefied phenol (250 ml for a 15-g prep) with TLE solution (see below) on the day of preparation. First, extract with an equal volume of TLE solution plus 0.5 ml of 15 M NaOH (this should bring the pH close to 8.0), then extract two more times with TLE solution.

TLE solution

0.2 M Tris
0.1 M LiCl
5 mM EDTA
pH to 8.2 with HCl

COMMENTARY

Background Information

The method described here can be used to prepare RNA from a variety of eukaryotic tissues. The critical factor in isolating RNA from eukaryotic tissues is inactivating the endogenous RNase and preventing introduction of RNase from external sources. In general, protocols for making RNA from eukaryotic organisms involve lysing the cells in the presence of a strong denaturant and deproteinizing agent which inhibits RNase as well as strips the protein away from the RNA. In this protocol, the RNA is then separated from DNA and other impurities by selective precipitation in high salt.

Literature Review

The use of phenol in RNA preparations originated with a method described by Kirby (1968). The phenol/SDS procedure for RNA extraction described here is taken most directly from a protocol developed by Palmiter (1974). However, this procedure has evolved from Palmiter's through contributions from many different laboratories. The most widely used alternative method, developed by Chirgwin et al. (1979), involves the use of guanidinium isothiocyanate instead of the phenol/SDS mixture to disrupt cells and inactivate nucleases.

Critical Parameters

The most essential factor in making high-quality RNA from eukaryotic tissues is to eliminate RNase activity. The endogenous RNase is quickly inactivated by the phenol/SDS extraction. After the extractions, it is very important not to introduce RNase from external sources. In making RNA from plant tissues, special attention must be paid to effectively grinding up the tissue. Plant tissues are frequently fibrous and contain organic compounds that can make fresh tissue difficult to break up. Therefore, it is recommended that the tissue is frozen before grinding in a mortar and pestle.

Anticipated Results

The yield of RNA varies widely, depending on the plant tissue from which the RNA is extracted. A tissue that is good for making RNA, such as pea seedlings, should yield about 7 mg of total RNA from 15 g of starting material; however, mature *Arabidopsis* plants yield only about 3 mg from the same amount of tissue. The quality of RNA will also vary due to differences in levels of carbohydrates and secondary metabolites in the tissues used.

This protocol can easily be scaled up or down by appropriate changes in the volumes used. It can be successfully used with less than 1 g of tissue.

Time Considerations

Two samples will take 2 to 3 hr to process to the first LiCl precipitation. After the addition of LiCl to the aqueous phase of the phenol extractions, the RNA can be stored at 4°C for days or even weeks.

Literature Cited

- Chirgwin, J.M., Przbyla, A.E., MacDonald, R.J., and Rutter, W.J. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18:5294.
- Kirby, K.S. 1968. Isolation of nucleic acids with phenolic solvents. *Meth. Enzymol.* 12B:87.
- Palmiter, R.D. 1974. Magnesium precipitation of ribonucleoprotein complexes: Expedient techniques for the isolation of undegraded polysomes and messenger ribonucleic acid. *Biochemistry* 13:3606.

Preparation of Bacterial RNA

UNIT 4.4

Procedures for isolating RNA from bacteria involve disruption of the cells, followed by steps to separate the RNA from contaminating DNA and protein. Lysis strategies differ in the protocols presented below, including chemical degradation of gram-negative cell walls using sucrose/detergent or lysozyme, and sonication to break open gram-positive cell walls. Combinations of enzymatic degradation, organic extraction, and alcohol or salt precipitation are employed in the procedures to isolate the RNA from other cellular components, and various inhibitors of ribonuclease activity (diethylpyrocarbonate, vanadyl-ribonucleoside complex, and aurintricarboxylic acid) are described. If extremely high-quality RNA is required (e.g., for gene expression studies), the first basic protocol employs CsCl step-gradient centrifugation to remove all traces of contaminating DNA.

NOTE: Water and all other solutions should be treated with DEPC to inhibit RNase activity. See UNIT 4.1, reagents and solutions, for instructions.

CAUTION: DEPC is a suspected carcinogen and should be handled carefully with gloves. ATA causes irritation on contact with skin, eyes, and respiratory system. VRC is harmful if inhaled or swallowed. Use only in a well-ventilated area. Avoid contact with skin.

ISOLATION OF HIGH-QUALITY RNA FROM GRAM-NEGATIVE BACTERIA

BASIC
PROTOCOL

This protocol produces high-quality RNA suitable for northern blotting, S1 mapping, and primer extension from *E. coli* or cyanobacteria. Bacterial cells are lysed in a sucrose/Triton X-100 solution; subsequent organic extraction and ethanol precipitation yield total nucleic acids. The high quality of the resulting RNA is due to removal of contaminating DNA and proteins in a CsCl step-gradient centrifugation. A Beckman TL-100 ultracentrifuge with a TLA-100.3 rotor reduces the time required for RNA pelleting, but because many laboratories may not have access to this rotor, conditions for using an SW-41 rotor are also provided. In addition, two effective RNase inhibitors—vanadyl-ribonucleoside complex (VRC) and aurintricarboxylic acid (ATA; Hallick et al., 1977)—are included in this procedure. Lysozyme and proteinase digestions are not required.

Materials

- 100-ml *E. coli* culture or 500-ml cyanobacteria culture
- Stop buffer
- STET lysing solution
- Buffered phenol (UNIT 2.1)
- Chloroform (UNIT 2.1)
- 3 M sodium acetate, pH 6.0
- 200 mM and 10 mM vanadyl-ribonucleoside complex (VRC; GIBCO/BRL)
- 1:1 buffered phenol/chloroform
- DEPC-treated water (UNIT 4.1)
- Cesium chloride, solid
- CsCl cushion: 5.7 M CsCl in 100 mM EDTA, pH 7.0
- 100% and 70% ethanol, ice cold

Preparation and
Analysis of RNA

4.4.1

Beckman JA-14 and JA-17 rotors
15-ml polypropylene tube (Sarstedt)
Beckman TL-100 ultracentrifuge with TLA-100.3 rotor and 13 × 51-mm polycarbonate centrifuge tubes, *or* Beckman L5-65 ultracentrifuge with SW-41 rotor and 14 × 89-mm ultraclear centrifuge tubes

Lyse the bacteria

1. Grow a 100-ml culture of *E. coli* or 500-ml culture of cyanobacteria to log phase and stop growth by adding $\frac{1}{20}$ vol stop buffer. Place the culture on ice.

Ice cubes can be added directly to the culture to reduce the temperature.

Stop buffer contains the nuclease inhibitor ATA. This inhibitor can affect certain enzymes and should not be used if RNA will be needed for primer extension or S1 nuclease analysis (see critical parameters).

2. Harvest cells by centrifuging 5 min in a Beckman JA-14 rotor at 6000 rpm (5500 × g), 4°C. Resuspend pellet in 2 ml STET lysing solution and add 100 μl of 200 mM VRC. Transfer to 15-ml polypropylene tube.
3. Add 1 ml buffered phenol and vortex 1 min. Add 1 ml chloroform and vortex 1 min. Centrifuge 10 min in a JA-17 rotor at 8500 rpm (10,000 × g), 4°C, and collect the top aqueous phase.

After centrifugation, the cellular debris will form a thick crust at the interphase. Avoid disturbing the interphase while collecting upper aqueous phase.

4. Precipitate nucleic acids by adding $\frac{1}{10}$ vol of 3 M sodium acetate and 2 vol ice-cold 100% ethanol. Centrifuge 10 min in a JA-17 rotor at 8500 rpm (10,000 × g), 4°C. Resuspend pellet in 2 ml of 10 mM VRC.

The nucleic acid will form a visible precipitate immediately after the addition of cold ethanol. There is no need to place tubes at low temperatures. Formation of nucleic acid precipitate is a good indication that the cells lysed properly.

5. Extract twice with 1:1 phenol/chloroform and reprecipitate as in step 4.

This step further reduces protein contamination. The size of the nucleic acid pellet should be smaller than the pellet seen in step 4.

Purify RNA on CsCl gradients

- 6a. *If using TLA-100.3 rotor:* resuspend the nucleic acid pellet in 2 ml DEPC-treated water. Add 1 g solid CsCl and dissolve it completely. Layer 2.25 ml of this solution onto a 0.75-ml CsCl cushion in a 13 × 51-mm TLA-100.3 polycarbonate tube.

- 6b. *If using SW-41 rotor:* resuspend the nucleic acid pellet in 6 ml DEPC-treated water. Add 4.5 g solid CsCl and adjust volume to 9 ml with DEPC-treated water. Layer this solution onto a 3-ml CsCl cushion in a 14 × 89-mm ultraclear SW-41 tube.

Whichever rotor is employed, it is very important that the two layers remain well-separated. Gently overlay the cushion by releasing the liquid in a controlled manner.

Other swinging-bucket ultracentrifuge rotors may be used, but the sample volume, rotor speed, and run time must be adjusted accordingly (MacDonald et al., 1987).

- 7a. *For TLA-100.3 rotor:* centrifuge 1 hr at 80,000 rpm (280,000 × g), 20°C.

- 7b. *For SW-41 rotor:* centrifuge 24 hr at 30,000 rpm (150,000 × g), 20°C.

Recover RNA

8. Immediately after centrifugation, carefully remove the DNA at the interface and then remove the upper CsCl layer with a sterile Pasteur pipet. Pour off remaining supernatant and mark the position of the RNA pellet. Wipe walls of centrifuge tube with tissue.

Do not let the centrifuge tubes sit after the completion of run. The RNA pellet may become loosened which will cause difficulty in removing all liquid from the pellet.

9. Resuspend RNA pellet in 0.36 ml DEPC-treated water and transfer to a 1.5-ml microcentrifuge tube using a 1000- μ l pipettor.
10. Add $\frac{1}{10}$ vol of 3 M sodium acetate and 2.5 vol ice-cold 100% ethanol. Precipitate 20 min at -70°C . Microcentrifuge 5 min at high speed, 4°C , to pellet RNA.
11. Add 1 ml ice-cold 70% ethanol and microcentrifuge 5 min at high speed, 4°C .
12. Air dry the RNA pellet and dissolve in 200 μ l DEPC-treated water. Quantify by measuring the A_{260} and A_{280} (APPENDIX 3). Adjust to a final concentration of 4 $\mu\text{g}/\mu\text{l}$. Place at -70°C for long-term storage or store as an ethanol precipitate.

ISOLATION OF RNA FROM GRAM-POSITIVE BACTERIA

This protocol—designed primarily for use with gram-positive cells—uses sonication to break open the cell wall, detergent to lyse the membranes, and protease digestion to degrade cellular protein. Organic extraction and ethanol precipitation yield total nucleic acids. DNA is removed enzymatically and the RNA is repurified.

Materials

- 10-ml bacteria culture
- Lysis buffer
- 25:24:1 phenol/chloroform/isoamyl alcohol (UNIT 2.1)
- 24:1 chloroform/isoamyl alcohol
- 5 M NaCl
- 100% and 70% ethanol, ice-cold
- DNase digestion buffer
- 2.5 mg/ml RNase-free DNase I (UNIT 4.1)
- TE buffer, pH 8.0 (APPENDIX 2)
- Sorvall SS-34 rotor (or equivalent)
- Microtip sonicator

Lyse the bacteria

1. Harvest the cells from a 10-ml bacteria culture by centrifuging in a Sorvall SS-34 rotor 10 min at 10,000 rpm ($12,000 \times g$), 4°C .
2. Resuspend cells in 0.5 ml lysis buffer. Transfer to microcentrifuge tube and freeze on dry ice.
3. Thaw and sonicate three times for 10 sec with a microtip sonicator. Use a power setting of about 30 W. The cell suspension should clear, indicating lysis.

Avoid foaming the lysate.

4. Incubate 60 min at 37°C .

This incubation allows digestion of bacterial protein.

BASIC PROTOCOL

Preparation and Analysis of RNA

4.4.3

Isolate and recover RNA

5. Add an equal volume of 25:24:1 phenol/chloroform/isoamyl alcohol and microcentrifuge 5 min at high speed, room temperature. Remove aqueous (top) layer to a clean microcentrifuge tube.
6. Reextract once with an equal volume of 25:24:1 phenol/chloroform/isoamyl alcohol, then extract once with an equal volume of 24:1 chloroform/isoamyl alcohol.
7. To 400 μ l aqueous phase, add 15 μ l of 5 M NaCl and fill microcentrifuge tube with ice-cold 100% ethanol. Mix and incubate 15 to 30 min on ice or overnight at -20°C .
8. Spin down precipitated RNA in microcentrifuge tube 15 min at 4°C . Rinse pellet with 500 μ l ice-cold 70% ethanol and air dry.
9. Redissolve pellet in 95 μ l DNase digestion buffer. Add 4 μ l of 2.5 mg/ml RNase-free DNase I. Incubate 60 min at 37°C .
10. Extract once with 25:24:1 phenol/chloroform/isoamyl alcohol. Add 100 μ l TE buffer to remaining organic phase, mix thoroughly, and microcentrifuge 5 min at high speed, room temperature. Pool the two aqueous phases.
11. Extract once with chloroform/isoamyl alcohol.
12. Add 10 μ l of 5 M NaCl to the aqueous phase and mix. Add 600 μ l of 100% ethanol. Precipitate overnight at -20°C or 15 min on dry ice/ethanol. Collect the precipitate by microcentrifuging 15 to 30 min at high speed, 4°C .
13. Rinse pellet with 500 μ l ice-cold 70% ethanol and air dry. Redissolve in 100 μ l DEPC-treated water. Dilute 10 μ l into 1 ml water and quantify RNA by measuring the A_{260} and A_{280} (APPENDIX 3). Store the remaining RNA at -70°C or as an ethanol precipitate.

ALTERNATE PROTOCOL

RAPID ISOLATION OF RNA FROM GRAM-NEGATIVE BACTERIA

The following rather simple procedure works well for *Escherichia coli* and other gram-negative bacteria. Lysozyme is used to strip off the cell walls and the resulting protoplasts are lysed with detergent. Diethylpyrocarbonate (DEPC), a potent inactivator of ribonuclease, is added to the lysate. Salt is then added to coprecipitate the detergent, protein, and chromosomal DNA, which are removed by centrifugation. RNA is the predominant component of the supernatant fluid and is recovered by ethanol precipitation. RNA prepared by this method contains small amounts of DNA and protein but should be adequate for most kinds of analyses.

Additional Materials

- 10-ml gram-negative bacteria culture
- Protoplasting buffer
- 50 mg/ml lysozyme
- Gram-negative lysing buffer
- Diethylpyrocarbonate (DEPC)
- Saturated NaCl: 40 g NaCl in 100 ml DEPC-treated H_2O (stir until solution reaches saturation)

Lyse the bacteria

1. Collect the cells from a 10-ml gram-negative bacteria culture by centrifuging 10 min in a Sorvall SS-34 rotor at 10,000 rpm ($12,000 \times g$), 4°C .

2. Resuspend in 10 ml protoplasting buffer. Add 80 μ l of 50 mg/ml lysozyme. Incubate 15 min on ice.

Lysozyme digests the cell walls, leaving behind protoplasts (in effect, naked cells).

3. Collect the protoplasts by centrifuging 5 min in an SS-34 rotor at 7000 rpm (5900 \times g), 4°C.

The gentler spin is used because protoplasts are fragile.

4. Resuspend in 0.5 ml gram-negative lysing buffer. Add 15 μ l DEPC. Mix gently and transfer to a microcentrifuge tube.

The lysate should become clear and viscous. Avoid excessive agitation which shears DNA.

5. Incubate 5 min at 37°C.

Isolate and recover RNA

6. Chill on ice. Add 250 μ l saturated NaCl. Mix by inversion. A substantial precipitate should form.

The precipitate contains SDS, protein, and DNA.

7. Incubate 10 min on ice. Microcentrifuge 10 min at high speed, at room temperature or 4°C.

8. Remove the supernatant to two clean microcentrifuge tubes. Add to each tube 1 ml ice-cold 100% ethanol and precipitate 30 min on dry ice or overnight at -20°C.

9. Microcentrifuge 15 min at high speed, 4°C.

10. Rinse pellet in 500 μ l ice-cold 70% ethanol and air dry. Redissolve in 100 μ l DEPC-treated water. Dilute 10 μ l into 1 ml water and determine the A_{260} and A_{280} (APPENDIX 3). Store the remaining RNA at -70°C.

REAGENTS AND SOLUTIONS

Stop buffer

200 mM Tris·Cl, pH 8.0
20 mM EDTA
20 mM sodium azide
20 mM aurintricarboxylic acid (ATA; Sigma)

Do not include ATA if RNA is needed for primer extension or S1 nuclease mapping.
Store in a brown bottle at room temperature.

DNase digestion buffer

20 mM Tris·Cl, pH 8.0
10 mM MgCl₂
Store at room temperature

Gram-negative lysing buffer

10 mM Tris·Cl, pH 8.0
10 mM NaCl
1 mM sodium citrate
1.5% (w/v) sodium dodecyl sulfate (SDS)
Store at room temperature

Lysis buffer

30 mM Tris, pH 7.4
100 mM NaCl
5 mM EDTA
1% (w/v) SDS
Add proteinase K to 100 µg/ml just before use
Store at room temperature

Protoplasting buffer

15 mM Tris-Cl, pH 8.0
0.45 M sucrose
8 mM EDTA
Store at 4°C

STET lysing solution

8% (w/v) sucrose
5% (v/v) Triton X-100
50 mM EDTA
50 mM Tris-Cl, pH, 7.0
Prepare from DEPC-treated stock solutions and store at 4°C.

COMMENTARY

Background Information

The RNA isolation procedures presented here involve disruption of the bacteria, removal of DNA and protein, and precipitation of the remaining RNA. The purity of the RNA preparation varies, depending on the protocol followed.

For many experiments, including gene expression studies, obtaining high-quality RNA is crucial. The high-quality RNA procedure (first basic protocol) employs a lysing regime that does not involve lysozyme or protease digestions to separate total nucleic acids from cell debris. The purity of the RNA preparation is the result of pelleting through a CsCl step gradient.

Traditionally, RNA isolated from DNA and protein on CsCl step gradients by high-speed centrifugations (UNIT 4.2; Glisin et al., 1974) required 4.5 to 26 hr, depending upon the type of ultracentrifugation conditions used (MacDonald et al., 1987). The use of a TL-100 ultracentrifuge and TLA-100.3 rotor significantly reduces the time needed to pellet the RNA, making this a quick procedure as well.

The second basic protocol is an adaptation of published methods for preparing RNA from eukaryotic cells (McKnight, 1978; Thomas, 1980) for use with the gram-positive bacterium *Bacillus subtilis* (Gilman and Chamberlin, 1983). It is a relatively straightforward tech-

nique that uses protease digestion and organic extraction to remove protein and nuclease digestion to remove DNA. The main complication is that sonication is often required to facilitate lysis of *B. subtilis*. This procedure may also be used for *E. coli*.

A rapid procedure for obtaining bacterial RNA is detailed in the alternate protocol. This protocol provides a relatively simple method for rapidly isolating RNA from *E. coli*—without organic extractions, protease, or nuclease treatment (Summers, 1970). As a result, the RNA preparation may contain small amounts of DNA and protein, but these contaminants will not usually interfere with subsequent analysis.

Critical Parameters

As with any RNA preparative procedure, care must be taken to ensure that solutions are free of ribonuclease. Solutions that come into contact with the RNA should be treated with DEPC to inhibit RNase activity. Most investigators wear gloves at all times when working with RNA solutions, as hands are a likely source of ribonuclease contamination (see introduction to Chapter 4).

In all three protocols, lysis is an important step. Cell lysis must be complete in the first basic RNA protocol, which can be easily confirmed during the first ethanol precipitation (step 4). In a well-lysed culture, the nucleic acid present in the aqueous phase will form a

visible precipitate immediately after the addition of ice-cold ethanol. In the second basic protocol, care should be taken to avoid foaming the lysate (which contains SDS) during sonication. In the alternate protocol, care should be exercised to avoid manipulations that might shear the chromosomal DNA. Sheared fragments of DNA will not be efficiently removed from the lysate by the salt precipitation.

Also critical when preparing RNA by the first basic protocol is the volume limitation of the centrifuge tubes. Although the TLA-100.3 rotor will complete the RNA pelleting in 1 hr, it will only accommodate a 3-ml volume of sample/CsCl cushion. The SW-41 rotor takes longer but can accommodate a 9-ml sample. When preparing the gradient, care must be taken to maintain the layers—the interface between the CsCl cushion and the solution of RNA layered on top should be clearly visible.

The final RNA pellet may appear pink due to bound ATA. This will not affect northern blot hybridizations. ATA affects certain enzymes and should be omitted from the stop buffer if the RNA is to be used for primer extension or S1 mapping. RNA obtained without ATA in the stop buffer can still be used for primer extension and S1 mapping because other RNase inhibitors (VRC and DEPC) are used in the procedure.

Anticipated Results

For the first basic protocol, it is possible to get ~5 mg total RNA from 100 ml *E. coli*, of which ~95% is ribosomal. With cyanobacteria cultures, 500 ml yields 0.2 to 1 mg total RNA.

For the second basic and alternate protocols, from 10 ml cells at $OD_{600} = 1$ in rich medium, expect yields of 0.5 to 1 mg RNA, more than 95% of which is stable rRNA and tRNA. Yields from cells growing more slowly in poorer media are lower, but this difference is predominantly due to the stable RNAs whose rate of synthesis is tightly coupled to growth rate.

Time Considerations

The first basic protocol requires 90 min for cell lysis, deproteinization, and setup of CsCl gradients. Centrifugation times are 1 and 24 hr for the TLA-100.3 and SW-41 rotors, respectively. Recovery and precipitation of the RNA pellet requires 40 min.

For the second basic protocol, allow ~3 hr

to get to the first ethanol precipitation. The RNA can be stored indefinitely at this stage. The alternate protocol is more rapid. It can be completed (i.e., to the ethanol precipitation) in <2 hr.

Literature Cited

- Gilman, M.Z. and Chamberlin, M.J. 1983. Developmental and genetic regulation of *Bacillus subtilis* genes transcribed by sigma-28-RNA polymerase. *Cell* 35:285-293.
- Glisin, V., Crkvenjakov, R., and Byus, C. 1974. Ribonucleic acid isolated by cesium chloride centrifugation. *Biochemistry* 13:2633-2637.
- Hallick, R.B., Chelm, B.K., Gray, P.W., and Orozco, E.M. 1977. Use of aurintricarboxylic acid as an inhibitor of nucleases during nucleic acid isolation. *Nucleic Acids Res.* 4:3055-3064.
- MacDonald, R.J., Swift, G.H., Przybyla, A.E., and Chirgwin, J.M. 1987. Isolation of RNA using guanidinium salts. *Methods Enzymol.* 152:219-227.
- McKnight, G.S. 1978. The induction of ovalbumin and conalbumin mRNA by estrogen and progesterone in chick oviduct explant cultures. *Cell* 14:403-413.
- Summers, W.C. 1970. A simple method for extraction of RNA from *E. coli* utilizing diethylpyrocarbonate. *Anal. Biochem.* 33:459-463.
- Thomas, P.S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. U.S.A.* 77:5201-5205.

Key Reference

- Reddy, K.J., Webb, R., and Sherman, L.A. 1990. Bacterial RNA isolation with one hour centrifugation in a table-top ultracentrifuge. *BioTechniques* 8:250-251.

Original description of the method for isolating high-quality RNA from E. coli and cyanobacteria.

Summers, 1970. See above.

A brief but illuminating comparison of basic methods for isolating RNA from E. coli.

Contributed by K.J. Reddy
(high-quality RNA)
State University of New York
Binghamton, New York

Michael Gilman
Cold Spring Harbor Laboratory
Cold Spring Harbor, New York

Preparation of Poly(A)⁺ RNA

UNIT 4.5

BASIC
PROTOCOL

This protocol separates poly(A)⁺ RNA from the remainder of total RNA, which is largely rRNA and tRNA. Total RNA is denatured to expose the poly(A) (polyadenylated) tails. Poly(A)-containing RNA is then bound to oligo(dT) cellulose, with the remainder of the RNA washing through. The poly(A)⁺ RNA is eluted by removing salt from the solution, thus destabilizing the dT:rA hybrid. The column can then be repeated to remove contaminating poly(A)⁻ RNA.

Materials

Diethylpyrocarbonate (DEPC)
5 M NaOH
Oligo(dT) cellulose
0.1 M NaOH
Poly(A) loading buffer
10 M LiCl
Middle wash buffer
2 mM EDTA/0.1% sodium dodecyl sulfate (SDS)
3 M sodium acetate
RNase-free TE buffer

Silanized column (APPENDIX 3)
Silanized SW-55 centrifuge tubes (APPENDIX 3)
Beckman SW-55 rotor or equivalent

The following solutions should be treated with DEPC to inhibit RNase activity: water, 10 M LiCl, 3 M sodium acetate. See UNIT 4.1, reagents and solutions, for instructions. CAUTION: DEPC is a suspected carcinogen and should be handled carefully.

Pour oligo(dT) column

1. Wash a silanized column with 10 ml of 5 M NaOH, then rinse it with water.
A silanized glass Pasteur pipet plugged with silanized glass wool or a small disposable column with a 2-ml capacity can be used. It is important to silanize the column to prevent RNA from sticking to the glass or plastic.
2. Add 0.5 g dry oligo(dT) cellulose powder to 1 ml of 0.1 M NaOH. Pour the slurry into the column and rinse the column with ~10 ml water.
3. Equilibrate the column with 10 to 20 ml of loading buffer. The pH of the output should be near 7.5 at the end of the wash.

Fractionate of poly(A)⁺ RNA

4. Heat ~2 mg total RNA in water to 70°C for 10 min. Add LiCl to 0.5 M final concentration from a 10 M LiCl stock solution.

Heating the RNA is necessary to disrupt any secondary structure that might form. It is important not to have too large a column for the amount of RNA selected. This is because the final poly(A)⁺ RNA will be so dilute that precipitation and workup of the sample will be very inefficient. Therefore, use a much smaller column when poly(A)⁺-selecting 500 µg or less of RNA, and scale down all of the steps accordingly. Generally, 1 ml of oligo(dT) cellulose is sufficient for 5 to 10 mg input RNA.

5. Pass the RNA solution through the oligo(dT) column. Wash the column with 1 ml poly(A) loading buffer. Make certain to save the eluant from this loading step.
6. Pass the eluant through the column twice more.

The starting RNA is passed through the column three times to ensure that all of the poly(A)⁺ RNA has stuck to the oligo(dT).

Preparation and
Analysis of RNA

4.5.1

Contributed by Robert E. Kingston

Current Protocols in Molecular Biology (1993) 4.5.1-4.5.3

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 21

7. Rinse the column with 2 ml middle wash buffer.
8. Elute the RNA into a fresh tube with 2 ml of 2 mM EDTA/0.1% SDS.
9. Reequilibrate the oligo(dT) column, as in step 3. Take the eluted RNA and repeat the poly(A)⁺ selection, as described in steps 4 to 8.

This second oligo(dT) column removes small levels of contaminating poly(A)⁻ RNA. It can be omitted if such contaminants will not create a problem, e.g., when RNA is to be used for S1 analysis.

10. Precipitate the eluted RNA by adjusting the salt concentration to 0.3 M sodium acetate using a 3 M sodium acetate stock solution. Add 2.5 vol ethanol and transfer the solution to two silanized SW-55 tubes.
11. Incubate RNA overnight at -20°C or on dry ice/ethanol for 30 min. Collect the precipitate by centrifuging 30 min at 50,000 rpm (304,000 × g), 4°C, in a Beckman SW-55 rotor.

This high speed centrifugation is required to pellet the very dilute RNA.

12. Pour off ethanol and allow pellets to air dry. Resuspend RNA in 150 μl of RNase-free TE buffer and pool the samples. Quality of RNA can be checked by heating 5 μl at 70°C for 5 min and analyzing on a 1% agarose gel (UNIT 2.5).

REAGENTS AND SOLUTIONS

Middle wash buffer

0.15 M LiCl
10 mM Tris·Cl, pH 7.5
1 mM EDTA
0.1% sodium dodecyl sulfate

Poly(A) loading buffer

0.5 M LiCl
10 mM Tris·Cl, pH 7.5
1 mM EDTA
0.1% sodium dodecyl sulfate

COMMENTARY

Background Information

Most messenger RNAs contain a poly(A) tail, while structural RNAs do not. Poly(A) selection therefore enriches for messenger RNA. The technique has proved essential for construction of cDNA libraries. It is also useful when analyzing the structure of low-abundance mRNAs. Removing the ribosomal and tRNAs from a preparation increases the amount of RNA that can be clearly analyzed by S1 analysis, for example, thus allowing detection of a low level message.

Literature Review

Aviv and Leder (1972) first used oligo(dT)

cellulose to bind poly(A)⁺ message and thus achieved fractionation of mRNA. The basic technique has undergone slight modification since then. Some protocols substitute poly(U) Sephadex for oligo(dT) (e.g., Moore and Sharp, 1984). Poly(U) Sephadex has somewhat longer stretches of nucleotides and a better flow rate than does oligo(dT) cellulose.

Critical Parameters

It is critical—even more so than in most RNA techniques—to have RNase-free solutions when doing poly(A) selection. This is because in most instances the 5' end of the message is needed. Therefore, no breaks be-

tween the 5' and 3' end of the message can be tolerated, as the broken message is separated from its poly(A)⁺ tail.

A second critical aspect is that the size of the column be matched to the amount of RNA being selected. The size of the column determines the volume in which the poly(A)⁺ RNA is eluted. If this volume is very large, then the poly(A)⁺ RNA will be extremely dilute. The more dilute the RNA, the more difficult it is to quantitatively precipitate. Also, a greater fraction of a dilute RNA solution will be lost due to nonspecific sticking of the RNA to the sides of the column and tubes used during the preparation. The capacity of oligo(dT) cellulose is generally supplied by the manufacturer and tends to be quite high. Thus, a very small column should be used when doing poly(A) selection on small quantities of RNA, and the elution volumes should be scaled down as well.

Anticipated Results

Approximately 1% of the input RNA should be retrieved as poly(A)⁺ RNA. The RNA should

appear as a smear from 20 kb down (with greatest intensity in the 5- to 10-kb range) on an agarose gel, with no evidence of rRNA bands.

Time Considerations

It will take approximately 1 hr to prepare and equilibrate the column. Running the column will take half an hour. The RNA is stable once it is in ethanol.

Literature Cited

- Aviv, H. and Leder, P. 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. Natl. Acad. Sci. U.S.A.* 69:1408-1412.
- Moore, C.L. and Sharp, P.A. 1984. Site-specific polyadenylation in a cell-free reaction. *Cell* 36:581-591.

Contributed by Robert E. Kingston
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

ANALYSIS OF RNA STRUCTURE AND SYNTHESIS

This section presents methods used to determine the level, structure, size, and synthesis rate of RNA. The first four protocols describe commonly used methods for analyzing in detail RNA structure and amount: S1 analysis, ribonuclease protection, primer extension, and northern blots. Both S1 analysis (UNIT 4.6) and ribonuclease protection (UNIT 4.7) use a single-stranded probe that is complementary to the sequence of the measured RNA. These protocols can be used to determine both the endpoint and the amount of a specific RNA. The S1 technique uses an end-labeled single-stranded DNA probe. This allows unambiguous determination of the 5' end of a message and results in low background on the final gel; this means, however, that the technique is not as sensitive as ribonuclease protection. The latter technique increases sensitivity by utilizing a body-labeled RNA probe, although high background problems may occur.

Primer extension (UNIT 4.8) employs a labeled oligomer of defined sequence that is extended to the end of any homologous RNA by the enzyme reverse transcriptase. The major strengths of this technique are that no extensive probe preparation is needed and that it allows mapping of RNA across discontinuities, such as splice sites. High background may also occur with this technique, resulting in lowered sensitivity. This background is caused by random priming as well as possible termination and pause sites in the RNA for reverse transcriptase.

The fourth technique presented for analysis of RNA structure is northern blot hybridization (UNIT 4.9). In this protocol, RNA is separated on an agarose gel and transferred to nitrocellulose. The size and amount of any specific RNA is determined by hybridizing a labeled specific probe to the nitrocellulose filter. This allows determination of the size of the entire message and also is very sensitive to message level. One cannot determine precise endpoints of a message using this protocol.

The final section of the chapter describes the nuclear runoff technique (UNIT 4.10), which allows determination of the number of active RNA polymerase molecules on a given eukaryotic gene. It is commonly used to determine how the transcription rates of genes vary in response to the growth state of a cell.

UNIT 4.6

S1 Analysis of Messenger RNA Using Single-Stranded DNA Probes

This method takes advantage of the ability of oligonucleotides to be efficiently labeled to a high specific activity at the 5' end through the use of kinase. The oligonucleotide is hybridized to a specific single-stranded template containing the complementary sequence to the oligonucleotide, and this hybrid is extended through the use of the Klenow fragment of *E. coli* DNA polymerase I. The mixture is cut with a restriction enzyme to give the probe a defined 3' end, and the probe is isolated on an alkaline agarose gel. Before using this protocol it is first helpful to have an M13 clone. If this is unavailable, a double-stranded plasmid clone of the region to be studied may be used, as described in the alternate protocol. A second alternate protocol is presented that describes the use of long oligonucleotides as probes for S1 analysis. This alternate protocol is useful for rapid and easy quantitation of the level of mRNA produced from a characterized promoter.

For the mapping of the 5' end of an RNA species, hybridization of the probe to RNA is then carried out. S1 nuclease is added to digest all of the unhybridized portion of the probe (see Fig. 4.6.1). Electrophoresis of the hybrid on a denaturing polyacrylamide

S1 Analysis of
mRNA Using
Single-Stranded
DNA Probes

4.6.1

Contributed by John M. Greene and Kevin Struhl
Current Protocols in Molecular Biology (1988) 4.6.1-4.6.13
Copyright © 2000 by John Wiley & Sons, Inc.

gel allows a determination of the length of the remaining DNA fragment. This length equals the distance between the 5' end of the probe to the 5' end of the RNA, defining the transcriptional start site to the nucleotide. By performing the hybridization reaction in vast probe excess, quantitation of the relative amounts of RNA can be estimated between samples.

S1 ANALYSIS OF mRNA USING M13 TEMPLATE

Materials

- Diethylpyrocarbonate (DEPC)
- Low gelling/melting temperature agarose (UNIT 2.6)
- Alkaline pour buffer
- Alkaline running buffer
- [γ - 32 P]ATP (10 mCi/ml, 6000 Ci/mmol)
- 100 μ g/ml oligonucleotide primer (UNIT 2.11)
- 10 \times polynucleotide kinase buffer
- T4 polynucleotide kinase (UNIT 3.10)
- 18 μ g M13mp template DNA containing sequence of interest
- 10 \times TM buffer (APPENDIX 2)
- 4 mM dNTP mix (UNIT 3.4)
- Klenow fragment of *E. coli* DNA polymerase I (UNIT 3.5)
- 10 \times restriction buffer (UNIT 3.1)

BASIC PROTOCOL

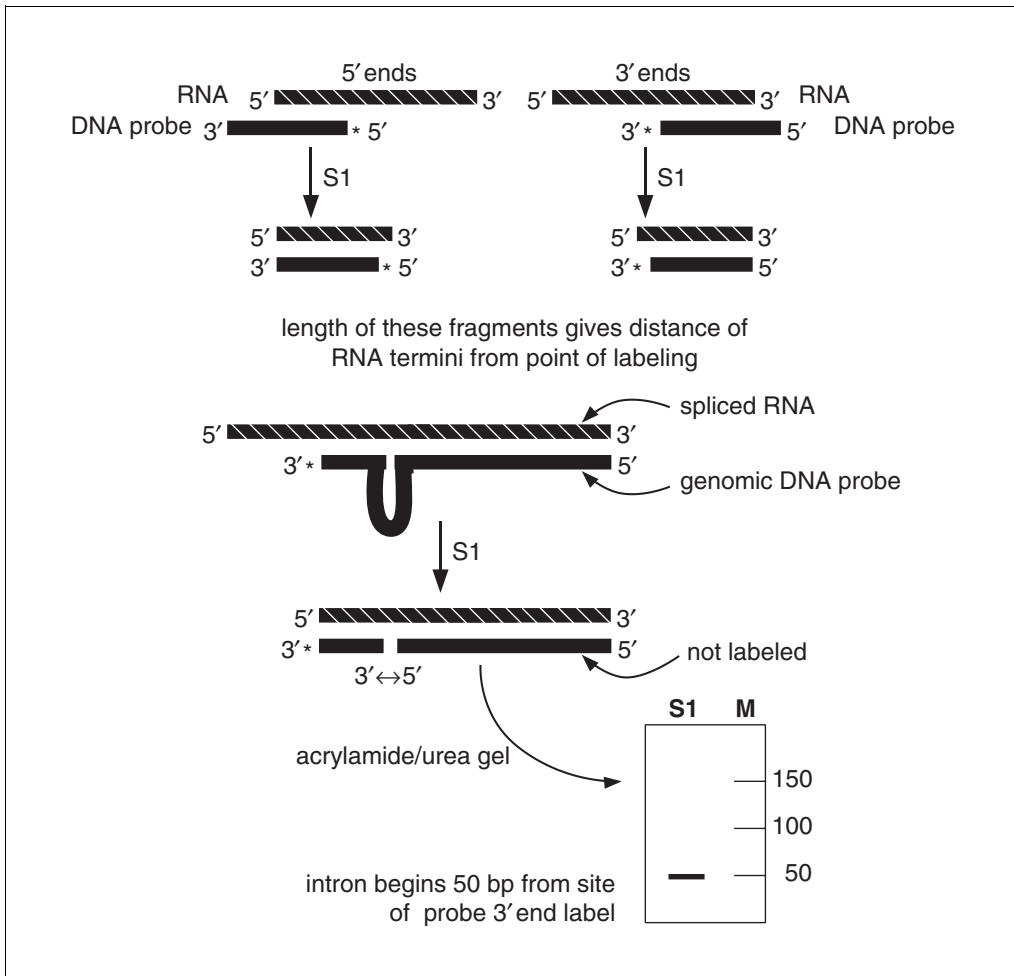


Figure 4.6.1 S1 mapping of RNA 5' ends and intron boundaries.

Preparation and Analysis of RNA

4.6.2

40 U restriction endonuclease (Table 3.1.1)
5 M ammonium acetate
100% ethanol
Alkaline loading buffer
TE buffer (APPENDIX 2)
10 mg/ml tRNA
Buffered phenol (UNIT 2.1A)
3 M and 0.3 M sodium acetate, pH 5.2
70% ethanol/30% DEPC-treated H₂O
S1 hybridization solution
2× S1 nuclease buffer (UNIT 3.12)
2 mg/ml single-stranded calf thymus DNA
S1 nuclease (UNIT 3.12)
S1 stop buffer
Formamide loading buffer (UNIT 2.12)

Additional reagents and equipment for ethanol precipitation (UNIT 2.1A) and agarose and denaturing polyacrylamide gel electrophoresis (UNITS 2.5 and 7.6, respectively)

Water and sodium acetate should be treated with DEPC to inhibit RNase activity. See UNIT 4.1, reagents and solutions, for instructions. CAUTION: DEPC is a suspected carcinogen and should be handled carefully.

Synthesize single-stranded probe

1. Prepare a 1.2% low gelling/melting temperature agarose gel in 1× alkaline pour buffer (see UNIT 2.5 for a description of how to pour horizontal agarose gels). Use a comb with fairly wide teeth (8 mm works well). When the gel has solidified, soak it overnight in 1× alkaline running buffer.

If the agarose is boiled in alkaline buffer, the gel will not form properly, hence the overnight soaking step. An alternative that eliminates soaking, but is somewhat more work, is to boil the agarose in water. When the agarose has cooled to 50° to 60°C, add alkaline running buffer to 1× and cast the gel.

2. To prepare kinased oligonucleotide, mix the following and incubate 30 min at 37°C:

20 µl [γ -³²P]ATP (10 mCi/ml, 200 µCi total)
1 µl 100 µg/ml oligonucleotide primer (100 ng; ideally a 20- to 30-mer)
2.5 µl 10× polynucleotide kinase buffer
4 U T4 polynucleotide kinase

The oligonucleotide used must hybridize to the RNA of interest and should be selected to produce a probe that results in a 50- to 250-base-long protected fragment after S1 analysis.

3. Heat 5 min at 65°C to inactivate kinase.
4. Anneal oligonucleotide to probe template. If a single-stranded template (e.g., an M13mp clone) is used, add the following to the kinased oligonucleotide: 18 µg template DNA in 55 µl water plus 9 µl of 10× TM buffer. Hybridize 15 min at 40°C.

To make probe using a double-stranded template, see alternate protocol.

5. Extend oligonucleotide primer to synthesize a full-length probe. To hybridized probe template, add the following: 9 µl of 4 mM dNTP mix (final concentration 400 µM) plus 2 µl Klenow fragment of *E. coli* DNA polymerase I (10 U). Incubate 30 min at 37°C. Heat 5 min at 65°C to inactivate Klenow fragment, then place on ice.
6. Cut probe to an appropriate length with 10 µl of 10× appropriate restriction buffer plus 40 U appropriate restriction enzyme.

Incubate 45 min at 37°C. Inactivate restriction enzyme by heating 5 min at 65°C.

This step creates the 3' unlabeled end of the probe. When analyzing a specific transcription start site, for example, this end of the probe should be 20 to 300 bases upstream of the start site. This allows easy separation of undigested probe from the band that forms as a result of hybridization to appropriately initiated RNA.

7. Add 100 µl of 5 M ammonium acetate and 500 µl ethanol. Precipitate at –20°C overnight or in dry ice/ethanol for 15 min. Spin 15 min in microcentrifuge, then dry pellet.

Isolate single-stranded probe

8. Load the alkaline agarose gel (see UNIT 2.5). Resuspend the probe in 25 µl of alkaline loading buffer and load carefully onto gel. Note which lane you load as the dye may not be visible by the end of the run. Run 4 hr at 1.8 V/cm maximum.

Gel can heat up and melt readily in alkaline running buffer if voltage is too high.

CAUTION: Follow all rules for safe electrophoresis as outlined in UNIT 2.5—cover gel box to provide radioactive shielding.

9. Isolate the probe from the gel. Shut off power and carefully slide the gel onto a glass plate behind a radiation shield; cover the gel with plastic wrap. In a darkroom, place a piece of X-ray film on the gel such that a corner of the film and the gel are in alignment. Expose 3 min, then develop the film.
10. Check alignments of the bands shown on the film with the gel by using a minimonitor. The upper, darker band is the probe, the lower band is unhybridized oligonucleotide (see Fig. 4.6.2). Trim away excess agarose with a razor, using the minimonitor to be sure you are not accidentally discarding probe. When down to a small block of agarose, cut into thin slices starting at the bottom of the rectangle (see Fig. 4.6.2). Check each slice with the minimonitor and remove the slice(s) that contains the upper

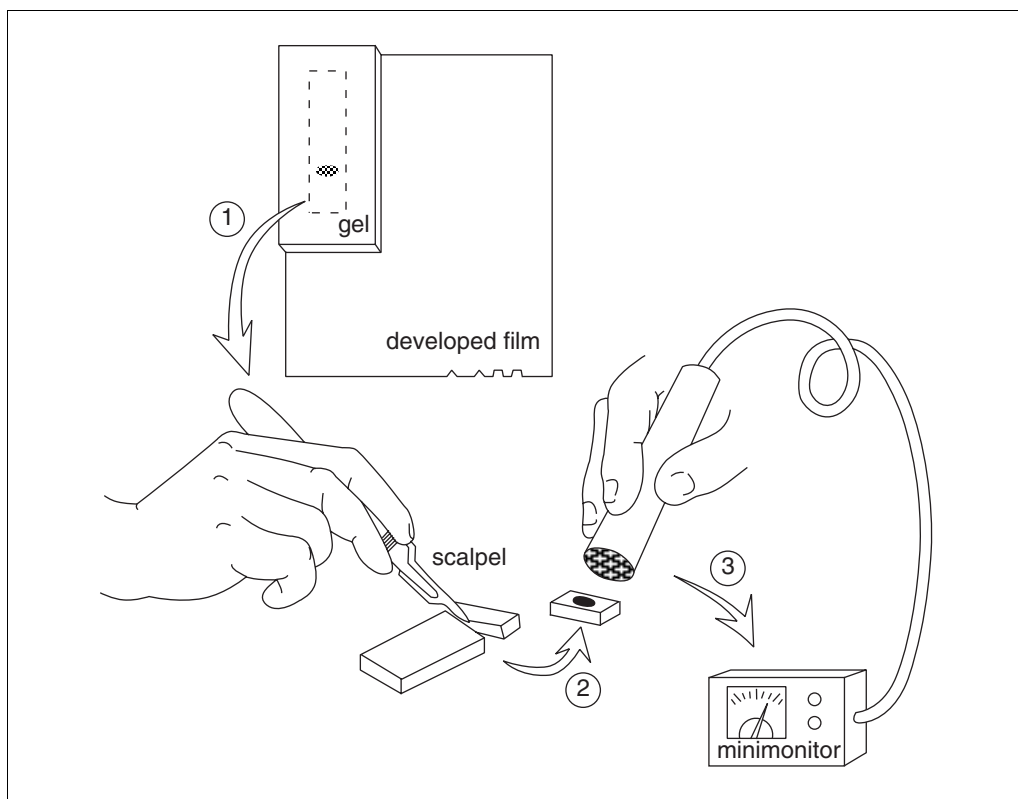


Figure 4.6.2 Excision of single-stranded end-labeled probe from alkaline low gelling/melting temperature agarose gel.

radioactive band to one or two microcentrifuge tubes.

CAUTION: *Minimize exposure—use radioactive shielding.*

11. Melt the slice(s) at 65°C and determine the approximate volume with a pipettor. Add an equal volume of TE buffer. Heat at 65°C for 10 min.
12. Add 1 µl of 10 mg/ml tRNA; phenol extract (*NOT* phenol/chloroform) twice, avoiding the white interface both times.
Chloroform can dramatically reduce the yield from these extractions.
13. Add $\frac{1}{10}$ vol of 3 M sodium acetate (pH 5.2), 2 vol ethanol, and precipitate. Resuspend pellet(s) in 100 µl of 0.3 M sodium acetate.
14. Count 1 µl for Cerenkov counts in a scintillation counter to determine cpm/µl.

Hybridize single-stranded probe and digest with S1 nuclease

This hybridization protocol uses an 80% formamide solution, and thus is very stringent. In our hands, it produces very low background. Less stringent hybridization conditions may be used; in particular, aqueous hybridization (steps 8 to 14 of the second alternate protocol).

15. Add amount of probe equal to 5×10^4 Cerenkov counts to up to 50 µg RNA on ice. Adjust final volume to 100 µl and the salt to 0.3 M sodium acetate. Add 250 µl ethanol and precipitate. Wash with 70% ethanol/30% DEPC-treated water and dry pellet inverted on Kimwipes for 30 min.

Do not use a Speedvac evaporator, as pellet will become virtually impossible to resuspend.

If the purpose of performing the S1 analysis is to quantitate RNA levels, it is important have probe excess. For most uses, 5×10^4 Cerenkov counts will be a large excess of probe. This can be verified empirically either by using 2- to 3-fold different amounts of the same RNA sample and affirming that there is an appropriate change in signal (see support protocol).

16. Resuspend the pellet in 20 µl S1 hybridization solution. Draw liquid through pipettor tip at least 50 times and vortex vigorously.

It is critical to resuspend well.

17. Denature the samples 10 min at 65°C. Hybridize overnight at 30°C.

18. The next morning, prepare following mix (volumes are per reaction):

150 µl 2× S1 nuclease buffer
3 µl 2 mg/ml single-stranded calf thymus DNA
147 µl H₂O
300 U S1 nuclease

Add 300 µl mix to each hybridization reaction. Incubate 60 min at 30°C.

The optimal amount of S1 nuclease may vary somewhat according to nuclease lot and the precise probe and RNA structure. Generally, between 100 and 1000 U/ml in the reaction works best.

19. Add 80 µl S1 stop buffer to each reaction, then add 1 ml 100% ethanol and precipitate. Wash pellet with 70% ethanol; dry 5 min in a Speedvac evaporator.
20. Resuspend pellet in 3 µl TE buffer and add 4 µl formamide loading dye. Boil tubes 3 min and place on ice.
21. Analyze 3 to 5 µl on a denaturing polyacrylamide/urea gel (*UNIT 7.4*) of the appropriate percentage for the expected size of the protected band (see commentary).

**PREPARATION OF SINGLE-STRANDED END-LABELED PROBE
AND S1 ANALYSIS OF mRNA**

1. Pour a 1.2% low gelling/melting temperature alkaline agarose gel.
2. To prepare kinased oligonucleotide mix the following and incubate 30 min at 37°C:
 - 20 μ l [γ -³²P]ATP (10 mCi/ml, 6000 Ci/mmol)
 - 1 μ l 100 μ g/ml oligonucleotide primer
 - 2.5 μ l 10 \times polynucleotide kinase buffer
 - 4 U polynucleotide kinaseHeat 5 min at 65°C to inactivate kinase.
3. Anneal oligonucleotide to probe template: Add 9 μ l of 10 \times TM buffer to 18 μ g template DNA in 55 μ l water and incubate 15 min at 40°C.
4. Extend oligonucleotide primer to make full-length probe: Add 9 μ l of 4 mM dNTP mix, then add 2 μ l Klenow fragment, and incubate 30 min at 37°C. Heat inactivate Klenow fragment 5 min at 65°C.
5. Cut probe to defined length: Add 10 μ l of 10 \times restriction buffer and 40 U appropriate restriction enzyme; incubate 45 min at 37°C. Treat 5 min at 65°C to inactivate restriction enzyme.
6. Ethanol precipitate (100 μ l of 3 M ammonium acetate, 500 μ l ethanol) and load onto the alkaline gel in 25 μ l alkaline loading buffer. Run gel at 1.8 V/cm for 4 hr.
7. Isolate probe. Add equal volume of TE buffer and incubate 10 min at 65°C.
8. Phenol extract twice, add 1 μ l of 10 mg/ml tRNA, ethanol precipitate, and quantitate.
9. Make up a 100- μ l solution containing 5×10^4 counts probe and up to 50 μ g RNA on ice.
10. Ethanol precipitate (250 μ l ethanol) and dry inverted 30 min.
11. Resuspend pellet thoroughly in 20 μ l S1 hybridization solution. Denature 10 min at 65°C. Hybridize overnight at 30°C.
12. The next morning, add the following mix to each reaction and incubate 60 min at 30°C:
 - 150 μ l 2 \times S1 nuclease buffer
 - 3 μ l 2 mg/ml single-stranded calf thymus DNA
 - 147 μ l H₂O
 - 300 U S1 nuclease
13. Add 80 μ l S1 stop buffer and ethanol precipitate. Wash pellet with 70% ethanol and dry 5 min in Speedvac evaporator. Resuspend pellet in 3 μ l TE buffer and add 4 μ l formamide loading dye. Boil 3 min and place on ice.
14. Load 3 to 5 μ l on denaturing polyacrylamide/urea gel.

**ALTERNATE
PROTOCOL**

**SYNTHESIS OF SINGLE-STRANDED PROBE FROM DOUBLE-STRANDED
PLASMID TEMPLATE**

A double-stranded plasmid can also be used as template for making the single-stranded S1 probe. It is first necessary, however, to denature the template as described here.

Additional Materials

- 10× NaOH/EDTA solution
- 1.5 M ammonium acetate, pH 4.5

1. To 18 µg DNA, add enough 10× NaOH/EDTA to achieve a 1× final concentration. Incubate 5 min at room temperature.
2. Add 1.5 vol of 1.5 M ammonium acetate, pH 4.5, to neutralize the solution. Add 2.5 vol ethanol and precipitate at -70°C for 15 min.
3. Collect the pellet by centrifugation, rinse with 70% ethanol, and dry 5 min in a Speedvac evaporator.
4. Resuspend the pellet in 55 µl water and add the template to the reaction of step 4 of basic protocol in place of the single-stranded template.

**ALTERNATE
PROTOCOL**

**QUANTITATIVE S1 ANALYSIS OF mRNA USING OLIGONUCLEOTIDE
PROBES**

This protocol is ideal for measuring the amount of RNA in situations where the structure of the RNA is already known. It is similar to the basic protocol except that the hybridization probes are synthetic oligonucleotides (40 to 80 nucleotides in length) that are ^{32}P -labeled at the 5' end. Necessary controls for this procedure are described in the support protocol.

Additional Materials

- 2 pmol each oligonucleotide probe
- 4 M ammonium acetate
- BioGel P-2 (or equivalent resin; optional)
- 3× aqueous hybridization solution (optional)
- 0.5 M EDTA
- 0.1 M NaOH

Additional reagents and equipment for acid precipitation (*UNIT 3.4*) and denaturing polyacrylamide gel electrophoresis (*UNIT 7.6*)

Design of the oligonucleotides

1. For each RNA to be analyzed, the oligonucleotide should contain at least 40 residues that are complementary to the RNA coding strand. It is essential that the 5' end of the oligonucleotide be complementary to the RNA, and it is useful if the 5' terminal nucleotides contain dG or dC residues.

The rate and optimal conditions of forming RNA:DNA duplexes as well as their stability are strongly influenced by the length of the duplex region. Oligonucleotides with 40 to 80 complementary residues are preferred over shorter oligonucleotides. The use of dG or dC residues at the 5' terminus minimizes fraying at the ends of the RNA:DNA duplex.

2. For experiments in which it is also desired to determine the 5' termini of the RNA(s), the 3' end of the oligonucleotide should extend at least 4 nucleotides beyond the RNA coding sequence (i.e., upstream of the upstream-most RNA initiation site). If the levels of RNA species with different 5' ends are to be quantitated, the oligonucleotide should be designed so that each RNA species will contain at least 40 complementary residues; this minimizes variability in physical properties of RNA:DNA duplexes.

By including the additional nucleotides, bands resulting from RNA:DNA duplexes are easily distinguished from the band representing the probe. A single probe can be used to quantitate RNA species provided that the 5' end of these RNAs map relatively close together (less than 20 to 30 nucleotides). If the 5' ends map further apart, individual (preferably non-overlapping) oligonucleotides must be used.

3. For the probe representing the control RNA, 5' end determination is unimportant, so the oligonucleotide can be complementary to internal RNA sequences. However, the 3' end of the oligonucleotide should contain at least 4 additional oligonucleotides that are not complementary to the RNA. The additional nucleotides should be chosen such that in the RNA:DNA hybrids purines are opposite purines and pyrimidines are opposite pyrimidines.

A major advantage of using the "internal probe" for the control RNA is that the resulting RNA:DNA duplexes generate a single band on the autoradiogram. Maximizing the mismatch at the noncomplementary nucleotides facilitates S1 cleavage and hence the distinction between bands resulting from RNA:DNA duplexes and from the band representing the probe.

4. It is crucial to design the oligonucleotides such that the RNA:DNA duplexes from the expected transcripts are sufficiently different in size to be separated by gel electrophoresis. A minimal spacing of 4 nucleotides among different species is desirable.

Prepare the hybridization probe

5. Set up the following T4 polynucleotide kinase reaction in a final volume of 25 μ l as described in *UNIT 3.10*:

2 pmol each oligonucleotide
150 μ Ci [γ -³²P]ATP (3000 to 7000 Ci/mmol)
2.5 μ l 10 \times T4 polynucleotide kinase buffer
10 U T4 polynucleotide kinase

For optimal sensitivity, use [γ -³²P]ATP at the highest possible specific activity. Crude preparations of "carrier-free" [γ -³²P]ATP work equally as well as more purified preparations which have a somewhat lower specific activity. If supplies of an oligonucleotide are limiting or if few samples are to be analyzed, the amount of oligonucleotide can be reduced. Conversion factor: 1 pmol of an oligonucleotide of length 45 = 15 ng.

6. Incubate at 37°C for 30 to 60 min. Stop the reaction by heating at 75°C for 10 min.

*It is useful to monitor the extent of phosphorylation by acid precipitation of the oligonucleotide (*UNIT 3.4*). Assuming that 2 pmol of 2 oligonucleotides are incubated in the above reaction conditions, the theoretical incorporation of labeled ATP into DNA should be 5% to 15%.*

7. Add 1 μ l of 10 mg/ml tRNA, 26 μ l of 4 M ammonium acetate, and 110 μ l ethanol and carry out an ethanol precipitation as described in *UNIT 2.1*. Resuspend the sample in 26 μ l H₂O, add 26 μ l of 4 M ammonium acetate and 110 μ l ethanol, and repeat the ethanol precipitation.

Ethanol precipitation from a 2 M ammonium acetate solution precipitates the oligonucleotide, but not the unincorporated ATP. Alternatively, the unincorporated precursors can be separated from the labeled oligonucleotide by column chromatography (see Fig. 3.4.1). Due to the small size of the oligonucleotide, it is necessary to use a resin such as BioGel P-2. This procedure generates sufficient probe for at least 50 to 100 hybridization reactions. The probe can be used for at least 6 weeks (store at -20°C) with minor effect on the quality of the results (of course, the exposure time must be increased to compensate for radioactive decay).

Hybridize, treat with S1 nuclease, and analyze product

Hybridization, S1 digestion, and product analysis using the oligonucleotide probe can be done exactly as described in the basic protocol (start at step 15). Care should be taken to ensure that the temperature and time of hybridization are such that hybridization goes to completion (see support protocol and critical parameters). Alternatively, an aqueous hybridization can be performed, as described below:

8. For each RNA preparation to be analyzed, set up the following 30 μ l hybridization reaction.

20 μ l RNA (containing up to 50 μ g RNA)
9 μ l 3 \times aqueous hybridization solution
1 μ l probe mixture (0.3 ng each oligonucleotide or approximately 10⁵ cpm)

It is useful to make up a premix containing enough hybridization solution and probe mixture for all the hybridization reactions. If supplies of the oligonucleotide are limiting, the amount of oligonucleotides can be reduced further (probably at least as low as 0.1 ng for each). In this case, control reactions should be carried out to determine if hybridization was complete.

9. Heat the reaction mixture to 75°C for 10 min, and then incubate overnight at 55°C (or other optimal hybridization temperature—see support protocol and critical parameters).

It is best to carry out the reactions in an incubator or covered waterbath to minimize evaporation to the top of the tube.

10. Briefly spin each tube in a microcentrifuge to collect the condensate from the top of the tube. Place tubes at 37°C.

11. To each tube, add 270 μ l of S1 nuclease mix containing 100 to 300 U of S1 nuclease (see basic protocol, step 18, for preparation of S1 nuclease mix). Incubate 30 to 60 min at 37°C. Stop the reaction by adding 3 μ l of 0.5 M EDTA, 1 μ l of 10 mg/ml tRNA, and 0.7 ml ethanol.

For any specific application, the amount of S1 nuclease should be determined empirically. The optimal amount of S1 nuclease should completely degrade the unhybridized probe but should not degrade the RNA:DNA hybrids.

12. Place the mixture on dry ice for 10 to 15 min, ethanol precipitate, and wash with ethanol as described in UNIT 2.1. Resuspend in 10 μ l of 0.1 M NaOH.

13. Combine 3 μ l of the resuspended products with 3 μ l of formamide loading dye. Heat to 90°C for 2 min and analyze on a denaturing polyacrylamide gel (UNIT 7.4).

14. Scan the resultant autoradiogram with a densitometer and measure the band intensities corresponding to the RNA(s) of interest and the control RNA(s). In this way, the amount of the RNA of interest is normalized to the amount of control RNA.

The absolute intensities of the bands will depend on the relative amount of the RNA species and the total amount of RNA loaded. When comparing RNA levels for a number of samples, it is often useful to run a second gel in which the volumes of the samples remaining from step 13 are adjusted to equalize the band intensities for the control RNA.

CONTROLS FOR QUANTITATIVE S1 ANALYSIS OF mRNA

For accurate quantitation, it is essential that (1) the probe is in excess, (2) the hybridization reaction goes to completion, (3) the RNA:DNA duplexes are equally stable, and (4) the S1 nuclease reaction proceeds properly. Although these conditions should be satisfied with the basic and alternate protocols described above, it is important to verify them with appropriate controls. First, varying amounts of a given RNA sample should be hybridized to a constant amount of probe to ensure that the band intensity is directly proportional to the amount of RNA added; i.e., the assay should be linear. Second, samples from a given hybridization reaction should be taken at various times after combining the RNA and DNA, and the resultant aliquots should be treated with S1 nuclease. Hybridization is complete when the band intensity does not increase after further incubation. This control is particularly important when using low concentrations of the labeled oligonucleotides. Third, the hybridization temperature should be varied to determine the optimal temperature. The temperature should be high enough to promote efficient hybridization, but low enough such that there is no preferential loss of RNA:DNA duplexes (shorter duplexes are less stable than longer ones). Fourth, for a standard determination, the level of S1 nuclease should be varied to ensure complete degradation of the probe without loss of the desired signal. Once the parameters for a given assay are established, it is unnecessary to repeat them for subsequent RNA determinations.

REAGENTS AND SOLUTIONS***Alkaline loading buffer***

30 mM NaOH
1 mM EDTA, pH 8
10% Ficoll
0.025% bromocresol green

Alkaline pour buffer, 50×

2.5 M NaCl
50 mM EDTA, pH 8
Dilute to 1× for working solution

Alkaline running buffer, 50×

1.5 M NaOH
50 mM EDTA, pH 8
Dilute to 1× for working solution

3× aqueous hybridization solution

3 M NaCl
0.5 M HEPES, pH 7.5
1 mM EDTA, pH 8

S1 hybridization solution

80% deionized formamide
40 mM PIPES, pH 6.4
400 mM NaCl
1 mM EDTA, pH 8
Store in 1-ml aliquots at -70°C

10× NaOH/EDTA solution

2 N NaOH
2 mM EDTA, pH 8

10× polynucleotide kinase buffer

700 mM Tris·Cl, pH 7.5
100 mM MgCl₂
50 mM dithiothreitol
1 mM spermidine·Cl
1 mM EDTA

2× S1 nuclease buffer

0.56 M NaCl
0.1 M sodium acetate, pH 4.5
9 mM ZnSO₄
Filter sterilize and store at 4°C

S1 stop buffer

4 M ammonium acetate
20 mM EDTA, pH 8
40 µg/ml tRNA
Store at 4°C

10 mg/ml tRNA

Dissolve in H₂O at 10 to 20 mg/ml and extract repeatedly with buffered phenol.

COMMENTARY**Background Information**

S1 mapping can be used to do the following: (1) map 5' and 3' ends of a transcript using an end-labeled probe (Weaver and Weissman, 1979); (2) quantitate the level of a particular RNA; (3) determine the direction of transcription; and (4) map the location and size of introns in primary eukaryotic transcripts (see Fig. 4.6.1).

The S1 enzyme is a single-stranded endonuclease that will digest both single-stranded RNA and DNA. The principle of S1 analysis is first to hybridize a DNA probe fragment to cellular RNA. S1 nuclease is then added to digest all single-stranded regions: 5' overhangs, 3' overhangs, and introns, depending on the specific probe fragment used. The double-stranded RNA-DNA hybrid is resistant to cleavage. The labeled DNA fragment then reflects the amount and size of RNA in the hybrid that is homologous to the DNA probe.

The use of oligonucleotide probes as described in the alternate protocol is advantageous for several reasons. First, probe preparation is rapid as it is accomplished by a simple T4 polynucleotide kinase reaction. Second, by including equimolar amounts of two or more oligonucleotides in the kinase reaction, hybridization probes for the RNA of interest as well as a control RNA(s) can be prepared simultaneously. Hybridization of RNA to such a probe mixture makes it possible to measure the

amount of a given RNA with respect to an internal control(s). Third, the ability to obtain large amounts of a labeled oligonucleotide ensures that the probe is in considerable excess over RNA and that the hybridization reaction goes to completion, conditions that are essential for accurate quantitation. Fourth, as the probes are derived from synthetic oligonucleotides and hence completely single-stranded, complications due to variable amounts of the complementary strand are avoided. Fifth, as the probes are short, reasonable results can be obtained even when the RNA sample is somewhat degraded (this is not a recommendation).

Literature Review

Berk and Sharp (1977) developed the technique of S1 mapping and used it to examine early adenovirus transcripts using unlabeled RNA hybridized to high-specific-activity ³²P-labeled DNA. By using overlapping restriction fragments, they were able to map the sizes and endpoints of these transcripts. Weaver and Weissman (1979) modified the original protocol through the use of end-labeled probes, which allows both determination of which DNA strand is being transcribed and which part of the probe is protected from S1 degradation, yielding map distances directly. Favalaro et al. (1980) described a method of using

two-dimensional gel analysis to map introns more quickly. In addition, this paper provides a great deal of information about one-dimensional S1 gel analysis of RNAs both by direct autoradiography or by Northern analysis.

The protocol given here has the advantage of *not* requiring an empirical determination of the correct hybridization temperature. In the original Berk-Sharp protocol, a double-stranded probe was used; this dictated that a hybridization temperature be found in the “window” above the T_m for the DNA-DNA duplex, but below the T_m for RNA-DNA duplexes. This must be done to prevent probe renaturation while promoting hybridization to RNA (Casey and Davidson, 1977). As each probe is different, this empirical determination must be done for each probe. Hybridization is then carried out at the $T_m \text{ DNA-DNA} + 1^\circ\text{C}$. The protocol presented describes a simple method of isolating a strictly single-stranded probe, and hence probe renaturation is irrelevant.

Critical Parameters

There are two major problems, opposite in nature, that occur during S1 analysis: no signal and excessive background. There are three primary reasons for obtaining no signal. First, the specific activity of the probe is lower than it should be. One should obtain 1 to 2×10^7 cpm of single-stranded probe after starting with 0.1 μg of a 20-mer (specific activity $\sim 10^7/\mu\text{g}$ for a 400-base single-stranded probe). If the yield of probe is significantly lower than this, one possibility is that the kinase reaction has not worked efficiently. Some oligonucleotide preparations contain residual chemicals used in preparation that can inhibit kinase. These can sometimes be removed by adjusting the oligonucleotide solution to 10 mM MgSO_4 and precipitating with 5 vol ethanol.

Second, the RNA is degraded. Check the RNA by running the preparation on an agarose gel and staining with ethidium bromide.

Third, there is very little of the specific RNA present. If S1 analysis is being performed after transfection of mammalian cells, the level of specific RNA is frequently very low. The level can be enriched by oligo(dT) selection (UNIT 4.5). Alternatively, transfection efficiency may be lower than is possible, and the transfection protocol should be optimized (see introduction to Section I of Chapter 9). Finally, the use of SP6 probes (UNIT 4.7) will increase sensitivity ~ 10 -fold and may result in detection of a signal.

Problems with excessive background are not usually seen with the protocol presented here.

Generally, the final gel can be left on film for over 2 weeks without significant background appearing. One cause of background is inclusion of DNA homologous to the single-stranded probe in the probe preparation. This can be excluded by making probes that are much shorter than the single-stranded template (e.g., 400 bases or less) and being careful when cutting the probe out of the gel to avoid contamination by the unlabeled high-molecular-weight species. A second cause is incomplete S1 digestion. A control in which tRNA is hybridized to the probe will be informative—if there is background, one of these first two problems is probably the cause. It is also possible that the RNA sample could be extensively degraded.

Finally, if formamide hybridization was used, the RNA and probe mixture may never have been solubilized in the hybridization. This results in an intense probe band and very little signal. If RNA pellets dry completely, they become virtually impossible to resuspend. Allow pellets to remain moist, and try to monitor resuspension with a minimonitor. It is advisable to be overly zealous in resuspending pellets in buffers containing formamide.

See support protocol for controls that will help to ensure accurate quantitation.

Anticipated Results

One should obtain 1 to 2×10^7 cpm of single-stranded probe from a preparation that starts with 0.1 μg of oligo. As isolation of DNA from low gelling/melting temperature agarose is about 50% efficient, this should allow a rough estimate of specific activity at about 1 to 2×10^7 cpm/ μg of a completed 400-base probe. The probe can be used as long as the specific activity remains high enough to detect the transcript of interest (up to 6 weeks). Gels will become somewhat “noisier” with older probe due to decay of the probe during storage. The following percentage gel should be used:

<i>Polyacrylamide urea gel (%)</i> :	<i>Size of band (base pairs)</i> :
4	>200
5	80 to 200
8	40 to 100
12	10 to 50

Time Considerations

Labeling and digesting the probe takes 2 to 3 hr. Running the gel and isolating the finished

probe takes another 7 hr. It is advisable when using a double-stranded plasmid template for the probe to denature the plasmid DNA as a first step before kinasing the oligo.

Setting up the S1 hybridizations takes 2 hr at most; however, it is advisable to hybridize overnight to be sure that all the RNA sequences being analyzed have hybridized to the probe. Performing S1 digestion takes 1 hr; subsequent sample workup and running of the denaturing polyacrylamide gel should take no more than 5 hr total, and very little of this is hands-on time.

Literature Cited

- Berk, A.J. and Sharp, P.A. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* 12:721-732.
- Casey, J. and Davidson, N. 1977. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. *Nucl. Acids Res.* 4:1539-1552.
- Favaloro, J., Treisman, R., and Kamen, R. 1980. Transcription maps of polyoma virus-specific RNA: Analysis by two-dimensional nuclease S1 gel mapping. *Meth. Enzymol.* 65:718-749.

Weaver, R.F. and Weissman, C. 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: The 5' termini of 15S β -globin mRNA precursor and mature 10S β -globin mRNA have identical map coordinates. *Nucl. Acids Res.* 7:1175-1193.

Key Reference

Sharp, P.A., Berk, A.J., and Berget, S.M. 1980. Transcription maps of adenovirus. *Meth. Enzymol.* 65:750-768.

Contains a fairly detailed discussion of S1 mapping procedures using double-stranded probes as well as S1 endonuclease and its optimal digestion conditions.

Contributed by John M. Greene
Massachusetts General Hospital
Boston, Massachusetts

Kevin Struhl
Harvard Medical School
Boston, Massachusetts

Ribonuclease Protection Assay

UNIT 4.7

BASIC
PROTOCOL

Sequence-specific hybridization probes of high specific activity are prepared by cloning the probe sequence downstream of a bacteriophage promoter. The plasmid is cleaved with a restriction enzyme, and the plasmid DNA is transcribed with bacteriophage RNA polymerase, which efficiently transcribes the cloned sequence into a discrete RNA species of known specific activity and high abundance. The RNA is purified by removal of the DNA template with deoxyribonuclease, the protein with phenol/chloroform, and the unincorporated label by ethanol precipitation in the presence of 2 M ammonium acetate. Alternatively, the probe is purified by gel electrophoresis (support protocol). The probe RNA is hybridized to sample RNAs. The hybridization reactions are treated with ribonuclease to remove free probe, leaving intact fragments of probe annealed to homologous sequences in the sample RNA. These fragments are recovered by ethanol precipitation and analyzed by electrophoresis on a sequencing gel. The presence of the target mRNA in the samples is revealed by the appearance of an appropriately sized fragment of the probe.

Materials

Diethylpyrocarbonate (DEPC)
5× transcription buffer
200 mM dithiothreitol (DTT)
3NTP mix (ATP, UTP, and GTP at 4 mM each; *UNIT 3.4*)
[α - 32 P]CTP (10 mCi/ml, 400 to 800 Ci/mmol)
Placental ribonuclease inhibitor (e.g., RNasin from Promega Biotec)
0.5 mg/ml template DNA (support protocol)
Bacteriophage RNA polymerase (*UNIT 3.8*)
2.5 mg/ml RNase-free DNase I (*UNIT 4.1*)
10 mg/ml tRNA (*UNIT 4.6*)
25:24:1 phenol/chloroform/isoamyl alcohol
2.5 M ammonium acetate
100% ethanol
75% ethanol/25% 0.1 M sodium acetate, pH 5.2
Hybridization buffer
Ribonuclease digestion buffer
40 μ g/ml ribonuclease A
2 μ g/ml ribonuclease T1
20% (w/v) sodium dodecyl sulfate (SDS)
20 mg/ml proteinase K (store at -20°C)
RNA loading buffer

Additional reagents and equipment for phenol extraction (*UNIT 2.1*) and denaturing polyacrylamide gel electrophoresis (*UNIT 2.12* & *UNIT 7.6*)

Water and sodium acetate should be freshly treated with DEPC to inhibit RNase activity. See *UNIT 4.1*, reagents and solutions, for instructions. CAUTION: *DEPC is a suspected carcinogen and should be handled carefully.*

Prepare the probe

1. Mix in an autoclaved microcentrifuge tube:
 - 4 μ l 5× transcription buffer
 - 1 μ l 200 mM DTT
 - 2 μ l 3NTP mix
 - 10 μ l [α - 32 P]CTP (10 mCi/ml, 400 to 800 Ci/mmol)
 - 1 μ l placental ribonuclease inhibitor (20 to 40 U)

Preparation and
Analysis of RNA

4.7.1

Contributed by Michael Gilman

Current Protocols in Molecular Biology (1993) 4.7.1-4.7.8

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 24

- 1 μ l 0.5 mg/ml template DNA (25 μ g/ml final)
- 1 μ l bacteriophage RNA polymerase (5 to 10 U)

The first four ingredients should be added before the template DNA to avoid precipitation of the DNA by the spermidine present in the transcription buffer.

2. Incubate 30 to 60 min at 40°C for SP6 RNA polymerase or 37°C for T7 and T3 RNA polymerases.

Incubating for longer periods of time is not productive because the labeled nucleoside triphosphate is rapidly used up. Adding more enzyme will not help either. In fact, the presence of excess enzyme leads to random transcription of the template. Some of these transcripts will be complementary to the probe, resulting in high backgrounds in the hybridizations.

3. Add 5 μ g or 10 U RNase-free DNase I (typically 2 μ l of a 5000 U/ml or 2.5 mg/ml stock solution); incubate 15 min at 37°C.

This digestion removes the template DNA. It is critical to the success of the procedure (see commentary).

4. Add 2 μ l of 10 mg/ml tRNA as carrier and water to a final volume of 50 μ l.
5. Extract with phenol/chloroform/isoamyl alcohol.
6. Add to the aqueous phase 200 μ l of 2.5 M ammonium acetate and 750 μ l of 100% ethanol. Mix and precipitate the RNA by incubating 15 min on ice and centrifuging 15 min at 4°C.

If the RNA probe is to be gel purified (see commentary), proceed with support protocol.

7. Redissolve the pellet in 50 μ l water and add 200 μ l of 2.5 M ammonium acetate and 750 μ l of 100% ethanol; precipitate as in step 6.
8. Repeat step 7.

These three precipitations remove virtually all of the unincorporated label. Other methods for separating the polymerized RNA from the free nucleoside triphosphates (e.g., spin columns) may be substituted as long as care is exercised to avoid contamination with ribonuclease.

9. Rinse the pellet with 75% ethanol/25% 0.1 M sodium acetate, pH 5.2.
10. Dry and redissolve in 100 μ l hybridization buffer.
11. Count 1 μ l in liquid scintillation counter to determine incorporation.

The specific activity of the probe is $\sim 10^9$ cpm/ μ g, and the probe breaks down rapidly due to radiolysis. Thus, the probe is best if used the day it is prepared. Acceptable results are obtained with probes stored at 4°C for a few days. After a week, the probe is largely degraded.

Hybridize probe RNA to sample RNAs

12. Precipitate the sample RNAs with 100% ethanol or, if RNA is stored in water, lyophilize.

For total cellular or cytoplasmic RNA, 10 μ g is usually sufficient for most messages. It is often possible to use less for abundant species. Include a sample containing tRNA. This reaction serves as a control for background hybridization and completion of the ribonuclease digestion. This hybridization reaction should yield no protected probe.

13. Redissolve in 30 μ l hybridization buffer containing 5×10^5 cpm of probe RNA.

Care should be exercised to ensure that the RNA pellet is completely redissolved. This is best accomplished by repeated pipetting. Note that more than one probe may be included in the hybridization buffer, allowing multiple mRNAs to be assayed in a single sample. Background increases with the amount of probe added to the hybridization reaction. Therefore, do not add more than necessary to achieve a linear increase in signal with input RNA.

14. Incubate 5 min at 85°C to denature RNA.
15. Rapidly transfer to desired hybridization temperature. Incubate overnight (>8 hr).

Because the RNA probe cannot reanneal as can a double-stranded DNA probe in an S1 nuclease assay, hybridization temperature is not absolutely critical. For each probe, however, there is an optimal temperature, perhaps because of secondary structures that can form in the probe. A good temperature to try is 45°C, but it is advisable to test a range of temperatures from 30° to 60°C. For some RNAs, hybridization may be complete in as little as 4 to 6 hr.

Digest the ribonuclease

16. Add to each hybridization reaction 350 µl ribonuclease digestion buffer containing 40 µg/ml ribonuclease A and 2 µg/ml ribonuclease T1. Incubate 30 to 60 min at 30°C.

The final hybridization signal is relatively insensitive to changes in these incubation conditions. However, if problems are suspected with this step, incubation temperatures ranging from 15° to 37°C should be tested. If internal cleavage of the RNA duplexes is observed, omit ribonuclease A (see commentary).

17. Add 10 µl of 20% (w/v) SDS and 2.5 µl of 20 mg/ml proteinase K. Incubate 15 min at 37°C.
18. Extract once with 400 µl phenol/chloroform/isoamyl alcohol, removing the aqueous phase to a clean microcentrifuge tube containing 1 µl of 10 mg/ml yeast tRNA.
19. Add 1 ml ethanol and precipitate.
20. Dry pellet and redissolve in 3 to 5 µl RNA loading buffer.
21. Incubate 3 min at 85°C to denature.
22. Analyze on a denaturing polyacrylamide/urea (sequencing) gel.

RNA has a lower mobility in these gels than DNA of the same length. If DNA markers are used to estimate the size of a protected RNA fragment, the correct size is 5% to 10% smaller than this estimate. For example, if an RNA species runs with a DNA marker of 100 nucleotides, its actual length is 90 to 95 nucleotides.

GEL PURIFICATION OF RNA PROBES

The following protocol is used for purification of full-length probe, which may be necessary under certain conditions.

Additional Materials

TBE buffer (APPENDIX 2)
Elution buffer

1. Dry the RNA after the first ethanol precipitation and redissolve in 10 µl RNA loading buffer.

It is important that the RNA be completely dried and then fully redissolved.

**SUPPORT
PROTOCOL**

**Preparation and
Analysis of RNA**

4.7.3

2. Heat 5 min at 85°C to denature the RNA.
3. Load onto a gel containing 6% polyacrylamide (29:1 acrylamide/bisacrylamide) and TBE buffer. The gel is 0.4 mm thick and 14 cm long (sequencing-length gels may also be used). Run at 300 V (higher for longer gels) until the bromphenol blue dye has run one-half to two-thirds down the gel.

Nondenaturing gels are used in the interest of speed, but it is important that the RNA be fully denatured before loading. Denaturing gels containing urea may be used, if desired.

4. Disassemble the gel, leaving it on one plate. Wrap in plastic wrap, mark with hot ink, and expose to film for 30 sec. Using the film as a template, excise the full-length RNA band.

Don't be greedy. The RNA actually occupies a narrower band than the band on the film. Cleaner hybridizations result from smaller gel slices.

5. Elute the RNA in 400 µl elution buffer and shake 2 to 4 hr at 37°C.

The elution buffer generally precipitates at room temperature. Warm briefly before use. Small probes (<200 nucleotides) will elute in 90 min. Large probes (>400 nucleotides) require longer elution times, but amounts sufficient for the experiment may elute in 2 hr.

6. Remove the eluate to a fresh microcentrifuge tube and add 1 ml of 100% ethanol.
7. Incubate 15 min on ice and spin 15 min in a microcentrifuge.

Monitor the eluate and the gel slice. There should be more counts in the former. All counts in the eluate should precipitate with ethanol.

8. Redissolve the RNA pellet in 50 µl hybridization buffer and count 1 µl in a liquid scintillation counter.

Yields will be lower than those obtained without gel purification, but this procedure should yield sufficient probe for more than 50 hybridizations. Background will be substantially lower.

SUPPORT PROTOCOL

PREPARATION OF TEMPLATE DNA

Template DNA is prepared by inserting the sequences of interest into a plasmid vector carrying a bacteriophage promoter. Strategies for constructing such plasmids are discussed in the commentary. Vectors containing bacteriophage promoters are commercially available.

Digest the DNA with a restriction enzyme (Table 3.1.1) that cuts immediately downstream of the probe sequence. This allows the generation of a uniquely sized runoff transcript. The enzyme chosen may cut the plasmid in several places as long as it does not cut within the phage promoter, the probe sequence, or intervening vector DNA. Restriction enzymes that generate 5' overhangs are best. (Do not cut with enzymes that leave 3' overhangs because these overhangs serve as initiation sites for the polymerase, leading to synthesis of RNA complementary to the probe.) Cut the DNA to completion but do not grossly overdigest. Extract cut DNA with phenol/chloroform, precipitate with ethanol, and redissolve at 0.5 mg/ml in RNase-free TE buffer. As little as 100 ng of template DNA will yield a reasonable amount of probe.

REAGENTS AND SOLUTIONS

Bacteriophage RNA polymerase

SP6, T3, or T7, depending on the vector in which probe sequences are cloned. All are functionally equivalent.

[...-³²P]CTP

Purchased commercially as an aqueous solution at a concentration of 10 mCi/ml. Specific activities of 400 to 800 Ci/mmol should be used. Higher specific activities yield a probe too unstable to use.

Labeled GTP or UTP may be substituted (see commentary). Be sure to use an appropriate mix of the three unlabeled nucleoside triphosphates.

Elution buffer

2 M ammonium acetate
1% SDS
25 µg/ml tRNA

Hybridization buffer

5× stock solution:
200 mM PIPES, pH 6.4
2 M NaCl
5 mM EDTA

Working solution:
4 parts formamide
1 part 5× stock buffer

Prepare hybridization buffer fresh as needed from frozen 5× stock and formamide freshly deionized by vortexing with mixed bed resin beads. Alternatively, store in small aliquots at -70°C.

3NTP mix

4 mM each ATP, UTP, and GTP
0.5 mM EDTA, pH 8
Store at -20°C

Ribonuclease digestion buffer

10 mM Tris·Cl, pH 7.5
300 mM NaCl
5 mM EDTA
Add 1/50 vol of 50× ribonuclease mix:
2 mg/ml ribonuclease A
0.1 mg/ml ribonuclease T1

RNase digestion buffer may be stored at room temperature. Add RNases from frozen stocks as needed. Be sure to use disposable tubes to make up this reagent.

RNA loading buffer

80% (v/v) formamide 1 mM EDTA, pH 8.0
0.1% bromphenol blue
0.1% Xylene Cyanol

Do not use Maxam-Gilbert loading buffer or any buffer containing NaOH.

5× transcription buffers

<i>T7, T3 RNA polymerases:</i>	<i>SP6 RNA polymerase:</i>
200 mM Tris·Cl, pH 8	200 mM Tris·Cl, pH 7.5
40 mM MgCl ₂	30 mM MgCl ₂
10 mM spermidine	10 mM spermidine
250 mM NaCl	

Use freshly prepared DEPC-treated water (see UNIT 4.1).

COMMENTARY

Background Information

Bacteriophage RNA polymerases possess several properties that make them well suited for the preparation of high-specific-activity hybridization probes (for review, see Chamberlin and Ryan, 1982). They are single subunit enzymes that are relatively stable and easy to purify. Moreover, because their genes reside on phage genomes, they have been fairly straightforward to clone and express in uninfected *E. coli*, thereby increasing the ease and economy of purification. Second, they polymerize RNA at an exceedingly high rate—200 to 300 nucleotides per minute—approximately 10 times faster than *E. coli* RNA polymerase and faster than DNA polymerases. Thus, large amounts of probe are easily prepared. Third, they are very specific in their action, recognizing fairly long promoter sequences that are unlikely to appear fortuitously in other DNA. Therefore, the probes are very homogeneous in sequence and usually require little further purification.

To make a probe for use in this assay, it is first necessary to subclone a fragment containing the sequences of interest downstream of a phage promoter (see Fig. 4.7.1). The sequence to be analyzed must be cloned such that the RNA produced by the phage polymerase is complementary to the RNA to be analyzed. Ideally, this construct should be able to be digested with a restriction enzyme to produce a linear template that will be transcribed into a 100- to 300-base runoff transcript. For example, if one wished to analyze the level of appropriately initiated transcription from a promoter, a DNA fragment from that promoter with endpoints at +150 and -100 could be cloned such that the +150 site is immediately adjacent to the

phage promoter. Cleaving the resultant clone at a restriction site adjacent to the -100 site would then allow runoff synthesis of a 250-base probe that would give a 150-base signal in the assay.

Literature Review

This protocol is a substitute for the widely used S1 nuclease mapping technique (Berk and Sharp, 1977; Weaver and Weissman, 1979; UNIT 4.6, this manual). It was first reported in the literature by Zinn et al. (1983) and subsequently described in detail by Melton et al. (1984). These developments were made possible by the isolation and characterization of the phage-encoded RNA polymerase of the *S. typhimurium* phage SP6 (Butler and Chamberlin, 1982) and the mapping and cloning of a bacteriophage promoter sequence (Kassevetis et al., 1982; Melton et al., 1984). Subsequently, the RNA polymerase produced by the related *E. coli* phages T3 and T7 have been similarly employed. The advantages of this technique over classical S1 nuclease mapping are as follows: (1) the ease of probe preparation—gel purification is not usually required; (2) unlike end-labeled or nick-translated probes, the specific activities of these RNA probes are fixed by the specific activity of the labeled ribonucleoside triphosphate, not by the efficiency of the enzymatic reaction; (3) the probes are prepared in large quantities and at much higher specific activity than classical end-labeled S1 probes, dramatically increasing the sensitivity of detection; (4) the probes are single-stranded and therefore cannot reanneal, and they generate more stable duplexes than would a DNA probe; and (5) treatment of RNA-RNA duplexes with ribonuclease is a more

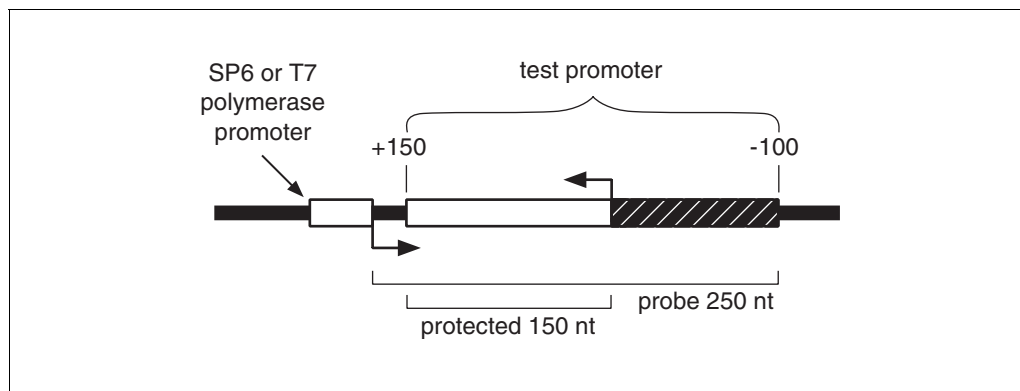


Figure 4.7.1 Structure of a typical transcription template for synthesizing an RNA probe.

reliable and reproducible reaction than is treatment of RNA-DNA hybrids with S1 nuclease, which can be sensitive to temperature and enzyme concentration. Many of these advantages may also be achieved, however, by carrying out S1 nuclease mapping with body-labeled single-stranded DNA probes (Ley et al., 1982).

RNA probes may also be substituted for DNA probes in blot hybridizations (UNIT 2.9).

Troubleshooting

The most common problems encountered with this procedure are probes that do not reach full length and high background in the hybridizations (i.e., excessive signal in the tRNA control). Incomplete transcripts may be caused by ribonuclease contamination or by pausing or termination of the RNA polymerase before completion of the transcript. Pausing and termination are sequence-specific phenomena which are further exacerbated by the low concentration of the labeled ribonucleoside triphosphate. They are best avoided by choosing a relatively small probe (100 to 300 nucleotides in length).

The appearance of incomplete transcripts can also depend on the choice of labeled ribonucleoside triphosphate. Therefore, changing the labeled nucleotide may help. In addition, specific activity may be sacrificed by adding unlabeled ribonucleoside triphosphate to increase the absolute concentration. If the problem persists, full-length probe may be gel purified (see support protocol), taking care to avoid ribonuclease contamination.

Background hybridization is most commonly caused by incomplete digestion of the template DNA in the transcription reaction. Residual DNA fragments copurify with the probe and will hybridize efficiently to the probe, generating a smear of bands which will appear in all lanes. This problem can usually be solved by trying a fresh preparation of DNase, by preparing a new batch of template DNA, or by gel purifying the probe. While the hybridization and digestion conditions listed in the protocol work satisfactorily for most probes, parameters that should be varied to optimize the signal-to-noise ratio include the hybridization temperature, amount of probe, and ribonuclease digestion conditions.

A second source of background is the presence of traces of sense RNA in the probe. This RNA can arise from excess enzyme in the transcription reaction or from the use of a template cleaved with a restriction enzyme that leaves 3' overhangs. These overhangs are in-

itiation sites for RNA polymerase. Even exceedingly small amounts of sense RNA in the antisense probe will create very high background in the hybridizations. Sense RNA contamination in the probe, whatever the cause, may be the explanation for persistent background problems. This situation is best resolved simply by gel purifying the probe (see support protocol).

Another problem that occasionally arises is internal cleavage of RNA duplexes during ribonuclease digestion. This leads to a loss of the full-length protected product and the generation of small subfragments. The bacterial CAT gene, for example, contains a long run of A residues near its 5' end. This region can "breathe" during ribonuclease treatment and then get cut by RNase A. This problem is solved simply by omitting RNase A from the digestions. Treatment with 2 $\mu\text{g}/\text{ml}$ RNase T1 alone is sufficient for most reactions, although it does not allow mapping of RNA ends at highest resolution.

Anticipated Results

Over half the label in the transcription reaction is typically incorporated into probe ($\sim 10^8$ cpm), often more. If substantially lower incorporation is achieved, the specific activity of the probe is just as high, but the yield is reduced. In principle, therefore, this probe should work just as well as probe from an efficient transcription reaction. In practice, however, greater backgrounds usually result when probes from low-incorporation reactions are used. When the probe is purified by gel electrophoresis (support protocol), yields are lower—typically $2\text{--}5 \times 10^7$ cpm—enough for at least 40 typical hybridizations.

The hybridization signal resulting from a fairly low abundance mRNA in 10 μg of total RNA should be detectable with an overnight exposure using an intensifying screen. Often, shorter exposures or exposures without a screen will suffice.

Time Considerations

Probe synthesis takes ~ 3 hr, mostly incubation time. With gel purification of probes, elapsed time—including probe electrophoresis and elution—is 6 to 8 hr. RNA samples for hybridization may be ethanol precipitated during this time. The probes are best used the day they are synthesized. Thus, a convenient arrangement is to prepare the probes and set up the hybridizations the same day. The following morning, nuclease digestions and gel

electrophoresis are performed. Allow 2 to 3 hr for the digestions and subsequent purification steps.

Literature Cited

- Berk, A.J. and Sharp, P.A. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* 12:721-732.
- Butler, E.T. and Chamberlin, M.J. 1982. Bacteriophage SP6-specific RNA polymerase. I. Isolation and characterization of the enzyme. *J. Biol. Chem.* 257:5772-5778.
- Chamberlin, M. and Ryan, T. 1982. Bacteriophage DNA-dependent RNA polymerase. In *The Enzymes*, Vol. XV (P. Boyer, ed.), pp. 87-108. Academic Press, NY.
- Kassevetis, G.A., Butler, E.T., Roulland, D., and Chamberlin, M.J. 1982. Bacteriophage SP6-specific RNA polymerase. II. Mapping of SP6 DNA and selective in vitro transcription. *J. Biol. Chem.* 257:5779-5787.
- Ley, T.J., Anagnou, N.P., Pepe, G., and Nienhuis, A.W. 1982. RNA processing errors in patients with β -thalassemia. *Proc. Natl. Acad. Sci. U.S.A.* 79:4775-4779.
- Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K., and Green, M.R. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids con-

taining a bacteriophage SP6 promoter. *Nucl. Acids Res.* 12:7035-7056.

Weaver, R.F. and Weissman, C. 1979. Mapping of RNA by modification of the Berk-Sharp procedure: The 5' termini of 15S β -globin mRNA and mature 10S β -globin mRNA have identical map coordinates. *Nucl. Acids Res.* 7:1175-1193.

Zinn, K., DiMaio, D., and Maniatis, T. 1983. Identification of two distinct regulatory regions adjacent to the human β -interferon gene. *Cell* 34:865-879.

Key References

Chamberlin and Ryan, 1982. See above.

Summarizes properties of bacteriophage RNA polymerases and provides references to the original literature.

Melton et al., 1984. See above.

Describes the construction of vectors carrying promoters for SP6 RNA polymerase, use of the polymerase for preparing RNA probes and large quantities of biologically active RNA, and biochemical properties of the transcription reaction.

Contributed by Michael Gilman
Cold Spring Harbor Laboratory
Cold Spring Harbor, New York

Primer Extension

UNIT 4.8

BASIC
PROTOCOL

This protocol can be used to map the 5' terminus of an RNA and to quantitate the amount of a given RNA by extending a primer using reverse transcriptase. The primer is an oligonucleotide (or restriction fragment) that is complementary to a portion of the RNA of interest. The primer is end-labeled, hybridized to the RNA, and extended by reverse transcriptase using unlabeled deoxynucleotides to form a single-stranded DNA complementary to the template RNA. The resultant DNA is analyzed on a sequencing gel. The length of the extended primer maps the position of the 5' end of the RNA, and the yield of primer extension product reflects the abundance of the RNA.

Materials

Diethylpyrocarbonate (DEPC; UNIT 4.1)
10× T4 polynucleotide kinase buffer (UNIT 3.4)
0.1 M and 1 M dithiothreitol (DTT; APPENDIX 2)
1 mM spermidine
50 to 100 ng/μl oligonucleotide primer (5 to 10 μM; UNIT 2.11)
10 μCi/μl [γ -³²P]ATP (3000 Ci/mmol)
20 to 30 U/μl T4 polynucleotide kinase (UNIT 3.10)
0.5 M EDTA, pH 8.0 (APPENDIX 2)
TE buffer, pH 8.0 (APPENDIX 2)
Cation-exchange resin (e.g., Bio-Rad AG 50W-X8), equilibrated in 0.1 M Tris-Cl (pH 7.5)/0.5 M NaCl
Anion-exchange resin (e.g., Whatman DE-52), equilibrated in TEN 100
TEN 100 buffer: 100 mM NaCl in TE buffer, pH 7.5 (APPENDIX 2)
TEN 300 buffer: 300 mM NaCl in TE buffer, pH 7.5 (APPENDIX 2)
TEN 600 buffer: 600 mM NaCl in TE buffer, pH 7.5 (APPENDIX 2)
Total cellular RNA (UNITS 4.1-4.3)
10× hybridization buffer
0.1 M Tris-Cl, pH 8.3 (APPENDIX 2)
0.5 M MgCl₂
1 mg/ml actinomycin D (store at 4°C protected from light; UNIT 1.4)
10 mM 4dNTP mix (UNIT 3.4)
25 U/μl AMV reverse transcriptase (UNIT 3.7)
RNase reaction mix
3 M sodium acetate (APPENDIX 2)
25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol (UNIT 2.1)
100% and 70% ethanol
Stop/loading dye (UNIT 7.4)
9% acrylamide/7 M urea gel (UNIT 2.12)
Silanized glass wool and 1000-μl pipet tip (APPENDIX 3)
65°C water bath
Additional reagents and equipment for denaturing gel electrophoresis (UNITS 2.12 & 7.6), phenol extraction and ethanol precipitation of DNA (UNIT 2.1), and autoradiography (APPENDIX 3)

NOTE: Water should be treated with DEPC to inhibit RNase activity. See UNIT 4.1, reagents and solutions, for instructions.

CAUTION: DEPC is a suspected carcinogen and should be handled carefully.

Preparation and
Analysis of RNA

4.8.1

Contributed by Steven J. Triezenberg

Current Protocols in Molecular Biology (1992) 4.8.1-4.8.5

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 20

Label and purify the oligonucleotide

1. Mix the following reagents in the order indicated (10 μ l final):

2.5 μ l H₂O
1 μ l 10 \times T4 polynucleotide kinase buffer
1 μ l 0.1 M DTT
1 μ l 1 mM spermidine
1 μ l 50-100 ng/ μ l oligonucleotide primer
3 μ l 10 μ Ci/ μ l [γ -³²P]ATP
0.5 μ l 20-30 U/ μ l T4 polynucleotide kinase.

Incubate 1 hr at 37°C.

In this step, 5 to 10 pmol of oligonucleotide are radiolabeled. Oligonucleotides used as primers should be 20 to 40 nucleotides long; shorter primers may not hybridize efficiently under the conditions described in step 7. Generally, primers should be selected to yield an extended product of <100 nucleotides to reduce the likelihood of premature termination of reverse transcriptase activity.

To avoid the possibility of precipitating the oligonucleotide, do not premix spermidine and oligonucleotide alone.

2. Stop reaction by adding 2 μ l of 0.5 M EDTA and 50 μ l TE buffer. Incubate 5 min at 65°C.
3. Prepare a small ion-exchange column by inserting a small plug of silanized glass wool into the narrow end of a silanized 1000 μ l pipet tip. Add 20 μ l of AG 50W-X8 resin and 100 μ l of DE-52 resin. Wash the column with 1 ml TEN 100 buffer.

This column will be used to purify the labeled oligonucleotide away from residual ATP. The plug for the column should consist of just enough glass wool to retain the chromatographic matrix.

4. Load the labeling reaction from step 2 onto the column. Collect flowthrough and reload it onto column.
5. Wash the column with 1 ml TEN 100, then with 0.5 ml TEN 300 (unincorporated nucleotide will be washed from the column; discard eluate as radioactive waste).
6. Elute radiolabeled oligonucleotide using 0.4 ml TEN 600. Collect eluate as a single fraction and store at -20°C in an appropriately shielded container until needed.

Other methods of purifying labeled oligonucleotides can be used, including repeated precipitation by ethanol in the presence of 2 M ammonium acetate (UNITS 2.12 & 8.2), gel-filtration chromatography using spin columns (UNIT 3.4), preparative gel electrophoresis (UNIT 2.12), and Sep-Pak chromatography (Sambrook et al., 1989).

Hybridize radiolabeled oligonucleotide and RNA

7. For each RNA sample, combine the following in a separate microcentrifuge tube (15 μ l final):

10 μ l total cellular RNA (10 to 50 μ g)
1.5 μ l 10 \times hybridization buffer
3.5 μ l radiolabeled oligonucleotide (from step 6).

Seal tubes securely and submerge 90 min in a 65°C water bath. Remove tubes and allow to cool slowly to room temperature.

Carry out primer extension reaction

8. For each sample, prepare the following reaction mix in a microcentrifuge tube on ice (multiply the indicated volumes by the number of samples plus one; 30.33 μ l final per sample):

0.9 μ l 1 M Tris-Cl, pH 8.3
0.9 μ l 0.5 M MgCl₂
0.25 μ l 1 M DTT
6.75 μ l 1 mg/ml actinomycin D
1.33 μ l 5 mM 4dNTP mix
20 μ l H₂O
0.2 μ l 25 U/ μ l AMV reverse transcriptase.

The 10 \times reverse transcriptase buffer defined in UNIT 3.4 is NOT used here because of the salt present in the hybridization reaction. Alternatively, the nucleic acids from this reaction may be precipitated and resuspended in 27 μ l water and 3 μ l of 10 \times RT buffer (UNIT 3.4).

Actinomycin D inhibits the initial DNA product from acting as both primer and template (as when portions of the product have complementary sequences that might hybridize), thus preventing synthesis of double-stranded "hairpin" DNA molecules.

9. To each tube containing RNA and oligonucleotide (from step 7), add 30 μ l reaction mix (from step 8). Incubate 1 hr at 42°C.

Reverse transcriptase stops less frequently when the reaction is carried out at 42°C than at lower temperatures.

Stop the reaction and analyze the product

10. Add 105 μ l RNase reaction mix to each primer extension reaction tube. Incubate 15 min at 37°C.

RNase digestion helps prevent aberrant electrophoresis of the primer extension products by reducing the amount of total RNA in the sample and by degrading the template RNA, leaving a cleanly labeled single-stranded DNA product.

11. Add 15 μ l of 3 M sodium acetate. Extract with 150 μ l phenol/chloroform/isoamyl alcohol, and remove aqueous (top) phase to a fresh tube.
12. Precipitate DNA by adding 300 μ l of 100% ethanol. Wash the pellet with 100 μ l of 70% ethanol. Remove all traces of ethanol using a pipet. Air dry the pellet with the cap open for 5 to 10 min.
13. Resuspend pellet in 5 μ l stop/loading dye. Heat tubes 5 min in a 65°C water bath.
14. Load samples on a 9% acrylamide/7 M urea gel and electrophorese until bromphenol blue reaches end of gel.
15. Dry gel and expose to X-ray film with an intensifying screen.

REAGENTS AND SOLUTIONS

10 \times hybridization buffer

1.5 M KCl
0.1 M Tris-Cl, pH 8.3
10 mM EDTA

RNase reaction mix

100 μ g/ml salmon sperm DNA
20 μ g/ml RNase A (DNase-free; UNIT 3.13) in TEN 100 buffer

COMMENTARY

Background Information

Primer extension is commonly used both to measure the amount of a given RNA and to map the 5' end of that RNA. The method employs reverse transcriptase, also known as RNA-dependent DNA polymerase. Like all DNA polymerases, this enzyme requires both a template to copy and a primer to be extended. Reverse transcriptase utilizes an RNA template for the synthesis of a complementary strand of DNA. In this analytical application of reverse transcriptase, a short piece of single-stranded DNA (the primer) hybridized to the RNA is extended until the 5' end of the RNA template is reached. The amount of DNA product is a reflection of the amount of the corresponding RNA present in the original sample. (Another major use for reverse transcriptase, the creation of cDNA libraries from mRNA, is covered in *UNIT 5.5*.)

Two other techniques, S1 protection (*UNIT 4.6*) and ribonuclease protection (*UNIT 4.7*), can also be used to define amounts and endpoints of RNA. Primer extension analysis is frequently used in conjunction with one of these protection methods to confirm 5'-end mapping results. Primer-extension differs from the protection assays, however, in one important aspect. With the protection methods, the defined DNA probe will be cleaved at any discontinuity between the probe and the RNA. Thus, either the end of the RNA or a splice junction will give discrete bands. Primer extension, on the other hand, will extend across splice sites or other discontinuities until the end of the RNA is reached. Thus, primer extension can be used to ensure that the signal observed in the protection assays truly reflects the 5' end of a transcript and not merely a splice junction.

Primer extension analysis is frequently employed for detecting transcripts produced during transient transfection assays or in vitro transcription assays (McKnight and Kingsbury, 1982; Jones et al., 1985). In such experiments, it is often easy to include internal controls to ensure that all technical aspects of the experiment—transfection, RNA isolation, and primer-extension reactions—are working properly. Such an internal control might be a heterologous RNA detected with its own specific primer. Alternatively, a single primer might be used for related RNA molecules, each containing sequences complementary to the primer, but differing in the distance between the

5' end and the primer binding site (Graves et al., 1985). In this latter case, variations in labeling efficiencies of the oligonucleotides can also be excluded to further reduce experimental error.

Additional applications of reverse transcriptase in primer extension experiments have been described in detail elsewhere. These include the determination of DNA sequences (Mierendorf and Pfeffer, 1987) and mapping sites on nucleic acids where proteins bind (Sasse-Dwight and Gralla, 1991).

Critical Parameters and Troubleshooting

The major difficulty in primer extension is encountering regions of RNA that cause the reverse transcriptase to pause or terminate. The resulting "short stops" show up as bands of intermediate size on the gel, both confusing interpretation and reducing the yield of fully extended primer. Such termination sites may be caused by extensive GC rich stretches of RNA or by secondary structures within the RNA. The protocol described here includes several measures to reduce such artifacts. The primer should be chosen so that the extension products are <100 nucleotides, to reduce the distance traversed by reverse transcriptase, thus minimizing the likelihood of encountering a significant secondary structure. Nucleotide concentrations well in excess of the apparent K_m of reverse transcriptase help to reduce pausing by the enzyme. A high incubation temperature (42°C) minimizes the effect of secondary structure.

Ideally, reverse transcriptase proceeds completely to the 5' terminus of the RNA, yielding a single discrete primer-extension product. Frequently, however, there is a cluster of primer-extension products which vary in size by increments of a single nucleotide, near the expected position. These incomplete extension products are probably caused by difficulties encountered by reverse transcriptase due to the methylated nucleotides of the 5' cap of eukaryotic mRNA.

Anticipated Results

Approximately 10^7 dpm of ^{32}P can be incorporated into 50 ng of a 30-base oligonucleotide. This primer should be able to detect messages that are 0.001% to 0.01% of the total RNA in a 10- to 50- μg sample. Primer-exten-

sion products should be detected after 15 to 30 hr of exposure to X-ray film with an intensifying screen.

Time Considerations

Labeling and purification of the oligonucleotide requires ~3 hr. The oligonucleotide can be used for up to 2 weeks, although best results are obtained within 1 week of labeling. The hybridization mixtures take ~15 min to prepare, with 90 min required for the hybridization itself. Primer extension reactions and sample preparation for electrophoresis require ~3 hr. Gel electrophoresis requires 2 hr, with another hour needed to dry the gel using a vacuum-type gel dryer.

Literature Cited

Graves, B.J., Eisenberg, S.P., Coen, D.M., and McKnight, S.L. 1985. Alternate utilization of two regulatory domains within the Moloney murine sarcoma virus long terminal repeat. *Mol. Cell. Biol.* 5:1959-1968.

Jones, K.A., Yamamoto, K.R., and Tjian, R. 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. *Cell* 42:559-572.

McKnight, S.L. and Kingsbury, R. 1982. Transcription control signals of a eukaryotic protein-encoding gene. *Science* 217:316-324.

Mierendorf, R.C. and Pfeffer, D. 1987. Sequencing of RNA transcripts synthesized in vitro from plasmids containing bacteriophage promoters. *Methods Enzymol.* 152:563-566.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed., p. 11.39. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Sasse-Dwight, S. and Gralla, J.D. 1991. Footprinting protein-DNA complexes in vivo. *Methods Enzymol.* 208:146-168.

Contributed by Steven J. Triezenberg
Michigan State University
East Lansing, Michigan

Analysis of RNA by Northern and Slot Blot Hybridization

Specific sequences in RNA preparations can be detected by blotting and hybridization analysis using techniques very similar to those originally developed for DNA (UNITS 2.9A, 2.9B, & 2.10). Fractionated RNA is transferred from an agarose gel to a membrane support (northern blotting); unfractionated RNA is immobilized by slot or dot blotting. The resulting blots are studied by hybridization analysis with labeled DNA or RNA probes. Northern blotting differs from Southern blotting largely in the initial gel fractionation step. Because they are single-stranded, most RNAs are able to form secondary structures by intramolecular base pairing and must therefore be electrophoresed under denaturing conditions if good separations are to be obtained. Denaturation is achieved either by adding formaldehyde to the gel and loading buffers or by treating the RNA with glyoxal and dimethyl sulfoxide (DMSO) prior to loading. Basic Protocol 1 describes blotting and hybridization of RNA fractionated in an agarose-formaldehyde gel. This is arguably the quickest and most reliable method for northern analysis of specific sequences in RNA extracted from eukaryotic cells. Alternate Protocol 1 gives details of the glyoxal/DMSO method for denaturing gel electrophoresis, which may provide better resolution of some RNA molecules. Alternate Protocol 2 describes slot-blot hybridization of RNA samples, a rapid method for assessing the relative abundance of an RNA species in extracts from different tissues. Stripping hybridization probes from blots can be done under three different sets of conditions; these methods are outlined in the Support Protocol.

Analysis of small noncoding RNA (microRNA, or miRNA) has received much attention as a new tool for analyzing gene expression. Because these miRNAs range from 20 to 30 nucleotides, traditional agarose gels will not separate the products adequately. Basic Protocol 2 describes a hybridization procedure using a polyacrylamide gel, adapted for these small RNAs. Alternate Protocol 3 describes a hybridization procedure for miRNAs that uses a non-formamide-containing hybridization solution.

NOTE: The ubiquity of contaminating RNases in solutions and glassware and the concomitant difficulties in ensuring that an RNA preparation remains reasonably undegraded throughout the electrophoresis, blotting, and hybridization manipulations can make it difficult to obtain good hybridization signals with RNA. To inhibit RNase activity, all solutions for northern blotting should be prepared using sterile deionized water that has been treated with diethylpyrocarbonate (DEPC) as described in UNIT 4.1. The precautions described in the introduction to Section I of this chapter (e.g., baking of glassware) should be followed religiously. In addition, RNA should not be electrophoresed in gel tanks previously used for DNA separations—a new tank plus accessories should be obtained and saved exclusively for RNA work. For full details on the establishment of an RNase-free environment, see Wilkinson (1991).

CAUTION: DEPC is a suspected carcinogen and should be handled carefully. Because DEPC reacts with ammonium ions to produce ethyl carbamate, a potent carcinogen, special care should be exercised when treating ammonium acetate solution with DEPC.

CAUTION: Investigators should wear gloves for all procedures involving radioactivity and should be careful not to contaminate themselves or their clothing. When working with ^{32}P , investigators should frequently check themselves and the working area for radioactivity using a hand-held radiation monitor. Any radioactive contamination should be cleaned up using appropriate procedures. Radioactive waste should be placed in appropriately designated areas for disposal. Follow the guidelines provided by the local radiation safety adviser; also see APPENDIX 1F.

Contributed by Terry Brown, Karol Mackey, and Tingting Du

Current Protocols in Molecular Biology (2004) 4.9.1-4.9.19

Copyright © 2004 by John Wiley & Sons, Inc.

**NORTHERN HYBRIDIZATION OF RNA FRACTIONATED BY
AGAROSE-FORMALDEHYDE GEL ELECTROPHORESIS**

The protocol is divided into three sections: electrophoresis of an RNA preparation under denaturing conditions in an agarose-formaldehyde gel, transfer of the RNA from the gel to a nylon or nitrocellulose membrane by upward capillary transfer, and hybridization analysis of the RNA sequences of interest using a labeled DNA or RNA probe. Hybridization is carried out in formamide solution, which permits incubation at a relatively low temperature, reducing degradation of the membrane-bound RNA. Nitrocellulose and nylon membranes are equally effective for northern hybridization analysis, although high backgrounds are likely with nylon membranes if the protocol is not followed carefully.

This protocol should be read in conjunction with *UNITS 2.9A & 2.10*, which describe the equivalent Southern procedures for DNA blotting and hybridization. Details of alternative transfer systems (upward capillary blots, electroblotting, and vacuum blotting) can be found in *UNIT 2.9A*. Modifications to the Southern hybridization procedure described in the Commentary to *UNIT 2.10* can also be used with northern blots, and the troubleshooting guide for DNA blotting and hybridization is also applicable to northern analysis. Other relevant units are located elsewhere in the manual: *UNIT 2.5A* covers the general features of agarose gel electrophoresis; *UNITS 3.18 & 3.19* describe the preparation of alternate non-radioactive probes and their use in hybridization analysis; and *UNIT 6.4* explains how to use labeled oligonucleotides as hybridization probes.

Materials

- 10× and 1× MOPS running buffer (see recipe for 10× buffer)
- 12.3 M (37%) formaldehyde, pH >4.0
- RNA sample: total cellular RNA (*UNITS 4.1-4.4*) or poly(A)⁺ RNA (*UNIT 4.5*)
- Formamide
- Formaldehyde loading buffer (see recipe)
- 0.5 μg/ml ethidium bromide in 0.5 M ammonium acetate *or* 10 mM sodium phosphate (pH 7.0; see recipe)/1.1 M formaldehyde with and without 10 μg/ml acridine orange
- 0.05 M NaOH/1.5 M NaCl (optional)
- 0.5 M Tris·Cl (pH 7.4; *APPENDIX 2*)/1.5 M NaCl (optional)
- 20×, 2×, and 6× SSC (*APPENDIX 2*)
- 0.03% (w/v) methylene blue in 0.3 M sodium acetate, pH 5.2 (optional)
- DNA suitable for use as probe *or* for in vitro transcription to make RNA probe (Table 2.10.1)
- Formamide prehybridization/hybridization solution (*UNIT 2.10*)
- 2× SSC/0.1% (w/v) SDS
- 0.2× SSC/0.1% (w/v) SDS, room temperature and 42°C
- 0.1× SSC/0.1% (w/v) SDS, 68°C
- 55°, 60°, and 100°C water baths
- Oblong sponge slightly larger than the gel being blotted
- RNase-free glass dishes (*UNIT 4.1*)
- Whatman 3MM filter paper sheets
- UV-transparent plastic wrap (e.g., Saran Wrap or other polyvinylidene wrap)
- Nitrocellulose or nylon membrane (see Table 2.9.1 for list of suppliers)
- Glass plate of appropriate size (Fig. 2.9.1)
- Vacuum oven
- UV transilluminator, calibrated (*UNIT 2.9A*)
- Hybridization oven (e.g., Hybridiser HB-1, Techne) and tubes

Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5A), radiolabeling of DNA by nick translation or random oligonucleotide priming (UNIT 3.5), RNA labeling by in vitro synthesis (UNIT 2.10), measuring specific activity of labeled nucleic acids and separating unincorporated nucleotides from labeled nucleic acids (UNIT 3.4), and autoradiography (APPENDIX 3A)

NOTE: All solutions should be prepared with sterile deionized water that has been treated with DEPC as described in UNIT 4.1; see unit introduction for further instructions and precautions regarding establishment of an RNase-free environment.

Prepare gel

1. Dissolve 1.0 g agarose in 72 ml water and cool to 60°C in a water bath (see UNIT 2.5A).

UNIT 2.5A provides details on preparing, pouring, and running the agarose gel; vary as described here.

This step will make a 1.0% gel, which is suitable for RNA molecules 500 bp to 10 kb in size. A higher-percentage gel (1.0 to 2.0%) should be used to resolve smaller molecules or a lower-percentage gel (0.7 to 1.0%) for longer molecules. The recipe may be scaled up or down depending on the size of gel desired; the gel should be 2 to 6 mm thick after it is poured and the wells large enough to hold 60 μ l of sample.

2. When the flask has cooled to 60°C, place in a fume hood and add 10 ml of 10 \times MOPS running buffer and 18 ml of 12.3 M formaldehyde.

CAUTION: Formaldehyde is toxic through skin contact and inhalation of vapors. All operations involving formaldehyde should be carried out in a fume hood.

The formaldehyde concentration in the gel is 2.2 M. Lower concentrations (down to 0.4 M) may be used; these result in less brittle gels but may not provide adequate denaturation for runs longer than 2 to 3 hr.

3. Pour the gel and allow it to set. Remove the comb, place the gel in the gel tank, and add sufficient 1 \times MOPS running buffer to cover to a depth of ~1 mm.

Prepare sample and run gel

4. Adjust the volume of each RNA sample to 11 μ l with water, then add:

5 μ l 10 \times MOPS running buffer
9 μ l 12.3 M formaldehyde
25 μ l formamide.

Mix by vortexing, microcentrifuge briefly (5 to 10 sec) to collect the liquid, and incubate 15 min at 55°C.

CAUTION: Formamide is a teratogen and should be handled with care.

5. Add 10 μ l formaldehyde loading buffer, vortex, microcentrifuge to collect liquid, and load onto gel.

0.5 to 10 μ g of RNA should be loaded per lane (see Commentary). Duplicate samples should be loaded on one side of the gel for ethidium bromide or acridine orange staining.

6. Run the gel at 5 V/cm until the bromphenol blue dye has migrated one-half to two-thirds the length of the gel.

This usually takes ~3 hr. Lengthy electrophoresis (>5 hr) is not recommended for northern transfers as this necessitates more formaldehyde in the gel (e.g., the recipe for a gel run overnight would be 1.0 g agarose, 60 ml water, 10 ml of 10 \times MOPS, and 30 ml of 12.3 M formaldehyde). Increasing the amount of formaldehyde causes the gel to become more brittle and prone to breakage during transfer and also increases the health hazard from volatilization of formaldehyde during electrophoresis.

Stain and photograph gel

- 7a. Remove the gel and cut off the lanes that are to be stained. Place this portion of the gel in an RNase-free glass dish, add sufficient 0.5 M ammonium acetate to cover, and soak for 20 min. Change solution and soak for an additional 20 min (to remove the formaldehyde). Pour off solution, replace with 0.5 $\mu\text{g/ml}$ ethidium bromide in 0.5 M ammonium acetate, and allow to stain for 40 min.

If necessary (i.e., if background fluorescence makes it difficult to visualize RNA fragments), destain in 0.5 M ammonium acetate for up to 1 hr.

- 7b. Alternatively, remove gel, cut off lanes, and stain 2 min in 1.1 M formaldehyde/10 mM sodium phosphate containing 10 $\mu\text{g/ml}$ acridine orange.

If necessary, destain 20 min in the same buffer without acridine orange.

8. Examine gel on a UV transilluminator to visualize the RNA and photograph with a ruler laid alongside the gel so that band positions can later be identified on the membrane.

*Molecular weight markers are not usually run on RNA gels as the staining causes rRNA molecules present in cellular RNA to appear as sharp bands that can be used as internal markers. In mammalian cells, these molecules are 28S and 18S (corresponding to 4718 and 1874 nucleotides respectively; Fig. 4.9.1). Mitochondrial rRNAs (16S and 12S in mammalian cells) may also be visible in some extracts, and plant extracts usually contain chloroplast rRNAs (23S and 16S). Bacterial rRNAs are smaller than the eukaryotic nuclear counterparts (*E. coli*: 23S and 16S, 2904 and 1541 nucleotides) and in some species one or both of the molecules may be cleaved into two or more fragments. If poly(A)⁺ RNA is being fractionated, commercial RNAs (e.g., 0.24- to 9.5-kb RNA ladder from Life Technologies) can be used as molecular weight markers.*

Prepare gel for transfer

9. Place unstained portion of gel in an RNase-free glass dish and rinse with several changes of sufficient deionized water to cover the gel.

The rinses remove formaldehyde, which would reduce retention of RNA by nitrocellulose membranes and hinder transfer onto nylon. The portion of the gel that will be blotted is not stained with ethidium bromide as this can also reduce transfer efficiency.

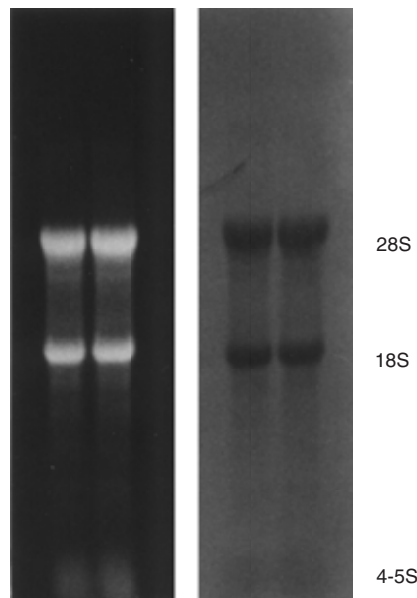


Figure 4.9.1 Rat liver RNA (5 μg) was electrophoresed on a formaldehyde 1% agarose gel containing ethidium bromide (left), transferred to a hybridization membrane and stained with methylene blue stain (Molecular Research Center; Herrin and Schmidt, 1988, right). Shown are 28S (4.7 kb) and 18S (1.9 kb) ribosomal RNAs, as well as 4S to 5S (0.10 to 0.15 kb) RNA containing mix of tRNA and 5S ribosomal RNA.

10. Add ~10 gel volumes of 0.05 M NaOH/1.5 M NaCl to dish and soak for 30 min. Decant and add 10 gel volumes of 0.5 M Tris·Cl (pH 7.4)/1.5 M NaCl. Soak for 20 min to neutralize.

This step is optional. It results in partial hydrolysis of the RNA which in turn leads to strand cleavage; the length reduction improves transfer of longer molecules. However, RNA is extremely sensitive to alkaline hydrolysis and smaller molecules may be fragmented into lengths too short for efficient retention by the membrane. Neutralization should be carried out only if efficient transfer of molecules >5 kb is required from a gel that has an agarose concentration of >1.0% and is >5 mm thick.

11. Replace solution with 10 gel volumes of 20× SSC and soak for 45 min.

This step is also optional but improves transfer efficiency with some brands of membrane.

Transfer RNA from gel to membrane

12. Place an oblong sponge slightly larger than the gel in a glass or plastic dish (if necessary, use two or more sponges placed side by side). Fill the dish with enough 20× SSC to leave the soaked sponge about half-submerged in buffer.

Refer to Figure 2.9.1A for a diagram of the transfer setup. The sponge forms the support for the gel. Any commercial sponge will do, but before a sponge is used for the first time, it should be washed thoroughly with distilled water to remove any detergents that may be present. Two or more sponges can be placed side by side if necessary. As an alternative, a solid support with wicks made out of Whatman 3MM paper (Fig. 2.9.1B) may be substituted. Do not use an electrophoresis tank, as the high-salt transfer buffer will corrode the electrodes.

If using a nylon membrane, a lower concentration of SSC (e.g., 10×) may improve transfer of molecules >4 kb; reduction of the SSC concentration is not recommended for a nitrocellulose membrane as the high salt is needed for retention of RNA.

13. Cut three pieces of Whatman 3MM paper to the same size as the sponge. Place them on the sponge and wet them with 20× SSC.
14. Place the gel on the filter paper and squeeze out air bubbles by rolling a glass pipet over the surface.
15. Cut four strips of plastic wrap and place over the edges of the gel.

This is to prevent buffer from “short-circuiting” around the gel rather than passing through it.

16. Cut a piece of nylon or nitrocellulose membrane just large enough to cover the exposed surface of the gel. Pour distilled water ~0.5 cm deep in an RNase-free glass dish and wet the membrane by placing it on the surface of the water. Allow the membrane to submerge. For nylon membrane, leave for 5 min; for nitrocellulose membrane, replace the water with 20× SSC and leave for 10 min.

Avoid handling nitrocellulose and nylon membranes even with gloved hands—use clean blunt-ended forceps instead.

17. Place the wetted membrane on the surface of the gel. Try to avoid getting air bubbles under the membrane; remove any that appear by carefully rolling a glass pipet over the surface.
18. Flood the surface of the membrane with 20× SSC. Cut five sheets of Whatman 3MM paper to the same size as the membrane and place on top of the membrane.
19. Cut paper towels to the same size as the membrane and stack on top of the Whatman 3MM paper to a height of ~4 cm.

20. Lay a glass plate on top of the structure and add a weight to hold everything in place. Leave overnight.

The weight should be sufficient to compress the paper towels to ensure good contact throughout the stack. Excessive weight, however, will crush the gel and retard transfer.

An overnight transfer is sufficient for most purposes. Make sure the reservoir of 20× SSC does not run dry during the transfer.

Prepare membrane for hybridization

21. Remove paper towels and filter papers and recover the membrane and flattened gel. Mark in pencil the position of the wells on the membrane and ensure that the up-down and back-front orientations are recognizable.

Pencil is preferable to pen, as ink marks may wash off the membrane during hybridization. With a nylon membrane only, the positions of the wells can be marked by slits cut with a razor blade (do not do this before transfer or the buffer will short-circuit). The best way to record the orientation of the membrane is by making an asymmetric cut at one corner.

22. Rinse the membrane in 2× SSC, then place it on a sheet of Whatman 3MM paper and allow to dry.

The rinse has two purposes: to remove agarose fragments that may adhere to the membrane and to leach out excess salt.

Immobilize the RNA and assess transfer efficiency

- 23a. *For nitrocellulose membranes:* Place between two sheets of Whatman 3MM filter paper and bake in a vacuum oven for 2 hr at 80°C.

Baking results in noncovalent attachment of RNA to the membrane; the vacuum is needed to prevent the nitrocellulose from igniting.

- 23b. *For nylon membranes:* Bake as described above or wrap the dry membrane in UV-transparent plastic wrap, place RNA-side-down on a UV transilluminator (254-nm wavelength), and irradiate for the appropriate length of time (determined as described in UNIT 2.9A, Support Protocol).

CAUTION: *Exposure to UV irradiation is harmful to the eyes and skin. Wear suitable eye protection and avoid exposure of bare skin.*

UV cross-linking is recommended for a nylon membrane as it leads to covalent attachment and enables the membrane to be reprobed several times. The membrane must be completely dry before UV cross-linking; check the manufacturer's recommendations, which may suggest baking for 30 min at 80°C prior to irradiation. The plastic wrap used during irradiation must be UV transparent—e.g., polyvinylidene (Saran Wrap). A UV light box (e.g., Stratagene Stratalinker) can be used instead of a transilluminator (follow manufacturer's instructions).

24. If desired, check transfer efficiency by either staining the gel in ethidium bromide or acridine orange as in steps 7 and 8 or (if using nylon membrane) staining the membrane in 0.03% (w/v) methylene blue in 0.3 M sodium acetate, pH 5.2, for 45 sec and destaining in water for 2 min.

If significant fluorescence is observed in the gel, not all the RNA has transferred. RNA bands on a nylon membrane will be stained by the methylene blue (Herrin and Schmidt, 1988).

Membranes can be stored dry between sheets of Whatman 3MM filter paper for several months at room temperature. For long-term storage they should be placed in a desiccator at room temperature or 4°C.

25. Prepare DNA or RNA probe labeled to a specific activity of $>10^8$ dpm/ μ g and with unincorporated nucleotides removed.

Probes are ideally 100 to 1000 bp in length. DNA for a double-stranded probe is obtained as a cloned fragment (Chapter 1) and purified from the vector by restriction digestion (UNIT 3.1) followed by recovery from an agarose gel (UNIT 2.6). The DNA is labeled by nick translation or random oligonucleotide priming (UNIT 3.5) to create the radioactive probe. A single-stranded DNA probe is created in the same fashion but using a single-stranded vector; the probe should be antisense so it will hybridize to the sense RNA strands that are bound to the membrane. An RNA probe, which should also be antisense, is created by in vitro synthesis from a single-stranded sense DNA fragment (UNIT 2.10).

26. Wet the membrane carrying the immobilized RNA (from step 23) in $6\times$ SSC.
27. Place the membrane RNA-side-up in a hybridization tube and add ~ 1 ml formamide prehybridization/hybridization solution per 10 cm^2 of membrane.

Prehybridization and hybridization are usually carried out in glass tubes in a commercial hybridization oven. Alternatively, a heat-sealable polyethylene bag and heat-sealing apparatus can be used. The membrane should be placed in the bag, all edges sealed, and a corner cut off. Hybridization solution can then be pipetted into the bag through the cut corner and the bag resealed.

28. Place the tube in the hybridization oven and incubate with rotation 3 hr at 42°C (for DNA probe) or 60°C (for RNA probe).

If using a bag, it can be shaken slowly in a suitable incubator or water bath. If using a nylon membrane, the prehybridization period can be reduced to 15 min.

29. If the probe is double-stranded, denature by heating in a water bath or incubator for 10 min at 100°C . Transfer to ice.
30. Pipet the desired volume of probe into the hybridization tube and continue to incubate with rotation overnight at 42°C (for DNA probe) or 60°C (for RNA probe).

The probe concentration in the hybridization solution should be 10 ng/ml if the specific activity is 10^8 dpm/ μ g or 2 ng/ml if the specific activity is 10^9 dpm/ μ g.

For denatured probe, add to hybridization tube as soon after denaturation as possible.

If using a bag, a corner should be cut, the probe added, and the bag resealed. It is very difficult to do this without contaminating the bag sealer with radioactivity. Furthermore, the sealing element (the part that gets contaminated) is often difficult to clean. Hybridization in bags is therefore not recommended.

Wash membrane and perform autoradiography

31. Pour off hybridization solution and add an equal volume of $2\times$ SSC/0.1% SDS. Incubate with rotation 5 min at room temperature, change wash solution, and repeat.

CAUTION: Hybridization solution and all wash solutions must be treated as radioactive waste and disposed of appropriately.

To reduce background, it may be beneficial to double the volume of the wash solutions. If using a bag, transfer the membrane to a plastic box for the washes.

32. Replace wash solution with an equal volume of $0.2\times$ SSC/0.1% SDS and incubate 5 min with rotation at room temperature. Change wash solution and repeat (this is a low-stringency wash; see UNIT 2.10 Commentary).
33. If desired, carry out two further washes using prewarmed (42°C) $0.2\times$ SSC/0.1% SDS for 15 min each at 42°C (moderate-stringency wash).

34. If desired, carry out two further washes using prewarmed (68°C) 0.1×SSC/0.1% SDS for 15 min each at 68°C (high-stringency wash).
35. Remove final wash solution and rinse membrane in 2× SSC at room temperature. Blot excess liquid and cover in UV-transparent plastic wrap.

Do not allow membrane to dry out if it is to be reprobbed.

36. Perform autoradiography.

If the membrane is to be reprobbed, the probe can be stripped from the hybridized membrane without removing the bound RNA (see Support Protocol). Do not add NaOH. The membrane must not be allowed to dry out between hybridization and stripping, as this may cause the probe to bind to the matrix.

ALTERNATE PROTOCOL 1

NORTHERN HYBRIDIZATION OF RNA DENATURED BY GLYOXAL/DMSO TREATMENT

In this procedure denaturation of the RNA is achieved by treating samples with a combination of glyoxal and DMSO prior to running in an agarose gel made with phosphate buffer. The glyoxal/DMSO method produces sharper bands after northern hybridization than do formaldehyde gels, but is more difficult to carry out as the running buffer must be recirculated during electrophoresis.

Additional Materials (also see Basic Protocol 1)

- 10 mM and 100 mM sodium phosphate, pH 7.0 (see recipe)
 - Dimethyl sulfoxide (DMSO)
 - 6 M (40%) glyoxal, deionized immediately before use (see recipe)
 - Glyoxal loading buffer (see recipe)
 - 20 mM Tris·Cl, pH 8.0 (APPENDIX 2)
- Apparatus for recirculating running buffer during electrophoresis
50° and 65°C water baths

NOTE: All solutions should be prepared with sterile deionized water that has been treated with DEPC as described in UNIT 4.1; see unit introduction for further instructions and precautions regarding establishment of an RNase-free environment.

Denature and carry out agarose gel electrophoresis

1. Prepare a 1.0% agarose gel by dissolving 1.0 g agarose in 100 ml of 10 mM sodium phosphate, pH 7.0. Cool to 60°C in a water bath, pour gel, and allow to set. Remove comb, place gel in gel tank, and add 10 mM sodium phosphate (pH 7.0) until gel is submerged to a depth of ~1 mm (see UNIT 2.5A).

UNIT 2.5A provides details on preparing, pouring, and running the agarose gel; vary as described here.

A 1.0% gel is suitable for RNA molecules 500 bp to 10 kb in size. A higher-percentage gel (1.0 to 2.0%) should be used to resolve smaller molecules or a lower percentage (0.7 to 1.0%) for longer molecules. The recipe may be scaled up or down depending on the size of gel desired; the gel should be 2 to 6 mm thick after it is poured and the wells large enough to hold 60 µl of sample.

2. Adjust volume of each RNA sample to 11 µl with water, then add:

- 4.5 µl 100 mM sodium phosphate, pH 7.0
- 22.5 µl DMSO
- 6.6 µl 6 M glyoxal.

Mix samples by vortexing, spin briefly (5 to 10 sec) in a microcentrifuge to collect the liquid, and incubate 1 hr at 50°C.

- Cool samples on ice and add 12 μ l glyoxal loading buffer to each sample. Load samples onto gel.

0.5 to 10 μ g of RNA should be loaded per lane (see Commentary). Duplicate samples should be loaded at one side of gel for ethidium bromide staining.

- Run the gel at 4 V/cm with constant recirculation of running buffer for ~3 hr or until bromphenol blue dye has migrated one-half to two-thirds the length of the gel.

Recirculation is needed to prevent an H^+ gradient forming in the buffer. If a gradient forms, the pH in parts of the gel may rise to >8.0, resulting in dissociation of the glyoxal from the RNA followed by renaturation. If no recirculation apparatus is available, electrophoresis should be paused every 30 min and the tank shaken to remix the buffer.

- Remove the gel, cut off lanes, and stain with ethidium bromide (see Basic Protocol 1, steps 7a and 8).

The RNA transfer (using the remaining portion of the gel) should be set up as soon as the gel is cut, before starting the staining.

Carry out northern transfer and hybridization analysis

- Transfer RNA (see Basic Protocol 1, steps 9 to 24).
- Immediately before hybridization, soak the membrane in 20 mM Tris·Cl (pH 8.0) for 5 min at 65°C to remove glyoxal.
- Continue with hybridization analysis (see Basic Protocol 1, steps 25 to 36).

NORTHERN HYBRIDIZATION OF UNFRACTIONATED RNA IMMOBILIZED BY SLOT BLOTTING

ALTERNATE PROTOCOL 2

RNA slot blotting is a simple technique that allows immobilization of unfractionated RNA on a nylon or nitrocellulose membrane. Hybridization analysis is then carried out to determine the relative abundance of target mRNA sequences in the blotted samples. The technique is based on the DNA dot- and slot-blotting procedure (UNIT 2.9B), the main difference being the way in which the samples are denatured prior to immobilization.

RNA dot blots can be prepared by hand but slot blots constructed using a manifold apparatus are preferable because the slots make it easier to compare hybridization signals by densitometry scanning.

Additional Materials (also see Basic Protocol 1)

0.1 M NaOH

10 \times SSC (APPENDIX 2)

20 \times SSC (APPENDIX 2), room temperature and ice-cold

Denaturing solution (see recipe)

100 mM sodium phosphate, pH 7.0 (see recipe)

Dimethyl sulfoxide (DMSO)

6 M (40%) glyoxal, deionized immediately before use (see recipe)

Manifold apparatus with a filtration template for slot blots (e.g., Bio-Rad Bio-Dot SF, Schleicher and Schuell Minifold II)

50° and 60°C water baths

NOTE: All solutions should be prepared with sterile deionized water that has been treated with DEPC as described in UNIT 4.1; see unit introduction for further instructions and precautions regarding establishment of an RNase-free environment.

Preparation and Analysis of RNA

4.9.9

Set up membrane for transfer

1. Clean the manifold with 0.1 M NaOH and rinse with distilled water.
2. Cut a piece of nylon or nitrocellulose membrane to the size of the manifold. Pour 10× SSC (for nylon membrane) or 20× SSC (for nitrocellulose membrane) into a glass dish; place membrane on top of liquid and allow to submerge. Leave for 10 min.

Avoid handling nitrocellulose and nylon membranes even with gloved hands—use clean blunt-ended forceps instead.

3. Place the membrane in the manifold. Assemble the manifold according to manufacturer's instructions and fill each slot with 10× SSC. Ensure there are no air leaks in the assembly.

Denature RNA samples

- 4a. Add 3 vol denaturing solution to RNA sample. Incubate 15 min at 65°C, then place on ice.

Up to 20 µg of RNA can be applied per slot. Total cellular RNA (UNITS 4.1-4.4) or poly(A)⁺ RNA (UNIT 4.5) can be used, although the latter is preferable (see Commentary).

- 4b. Alternatively, mix:

11 µl RNA sample
4.5 µl 100 mM sodium phosphate, pH 7.0
22.5 µl DMSO
6.6 µl 6 M glyoxal.

Mix by vortexing, spin briefly in a microcentrifuge to collect liquid, and incubate 1 hr at 50°C.

5. Add 2 vol ice-cold 20× SSC to each sample.

Pass samples through manifold

6. Switch on the suction to the manifold device and allow the 10× SSC added in step 3 to filter through. Leave the suction on.

The suction should be adjusted so that 500 µl buffer takes ~5 min to pass through the membrane. Higher suction may damage the membrane. Slots that are not being used can be blocked off by placing masking tape over them or by applying 500 µl of 3% (w/v) gelatin to each one. The former method is preferable as use of gelatin may lead to a background signal after hybridization. Alternatively, keep all slots open and apply 10× SSC instead of sample to the slots not being used.

7. Load each sample to the slots and allow to filter through, being careful not to touch the membrane with the pipet tip.
8. Add 1 ml of 10× SSC to each slot and allow to filter through. Repeat.
9. Dismantle the apparatus, place the membrane on a sheet of Whatman 3MM paper, and allow to dry.

Immobilize RNA and carry out hybridization

10. Immobilize the RNA (see Basic Protocol 1, step 23).

If glyoxal/DMSO denaturation has been used, immediately before hybridization soak the membrane in 20 mM Tris-Cl (pH 8.0) for 5 min at 65°C to remove glyoxal.

11. Carry out hybridization analysis as described in steps 25 to 36 of Basic Protocol 1.

REMOVAL OF PROBES FROM NORTHERN BLOTS

Hybridization probes can be removed from northern blots on nylon membranes without damage to the membrane or loss of the transferred RNA. Some probes (particularly RNA probes) are more resistant to stripping. In these cases, higher temperatures, longer incubation periods, or the inclusion of formamide may be necessary for complete probe removal. The following stripping procedures are appropriate for both radioactive and chemiluminescent probes. Begin with the mildest conditions (step 1a) and monitor results to determine the extent of stripping. If the hybridization signal is still evident, proceed with the more stringent treatments (steps 1b and 1c) until stripping is complete.

Materials

Northern hybridization membrane containing probe (see Basic Protocol 1, Alternate Protocol 1, or Alternate Protocol 2)

Stripping solution (see recipe)

Hybridization bags

65°, 80°, or 100° (boiling) water bath

UV-transparent plastic wrap (e.g., Saran Wrap or other polyvinylidene wrap)

Additional reagents and equipment for autoradiography (*APPENDIX 3A*)

CAUTION: If hybridization probes include a radioactive label, dispose of stripping solutions as radioactive waste. Observe appropriate caution when working with the toxic compound formamide.

- 1a. *To remove probes at 80°C:* Place membrane in a hybridization bag containing stripping solution without formamide. Place bag in water preheated to 80°C for 5 min. Pour out solution, then repeat this washing process three to four times.

Add sufficient solution to cover the membrane completely when using a bag; alternatively, stripping can be done in an open container, again with sufficient solution to cover the membrane.

- 1b. *To remove probes at 100°C:* Place membrane in a hybridization bag containing stripping solution without formamide. Place bag in boiling water for 5 min. Pour out solution, then repeat this washing process three to four times.

- 1c. *To remove probes with formamide:* Place membrane in a hybridization bag containing stripping solution with formamide. Place bag in water preheated to 65°C for 5 min. Pour out solution, then repeat this washing process three times using stripping solution with formamide and once using stripping solution without formamide.

2. Place membrane on filter paper to remove excess solution. Wrap membrane in plastic wrap and perform autoradiography to verify probe removal.

If a chemiluminescent probe was used, verify probe removal by chemiluminescent detection (UNIT 3.19). The membrane may be immediately rehybridized or air-dried and stored for future use.

NORTHERN HYBRIDIZATION OF SMALL RNA FRACTIONATED BY POLYACRYLAMIDE GEL ELECTROPHORESIS

This protocol is adapted for analysis of small RNAs. The major differences between this procedure and the traditional northern hybridization procedure are the fractionation system and the transfer system applied. Fractionation using denaturing polyacrylamide gel electrophoresis (PAGE) allows better separation of small RNAs. The introduction of a semidry transfer system reduces the time of the experimental procedure to 2 days.

Tissue or cell samples
TRIzol reagent (Invitrogen)
RNase-free H₂O (*UNIT 4.1*)
15% denaturing polyacrylamide sequencing (urea/TBE) gel (*UNIT 7.6*)
0.5× TBE electrophoresis buffer (*APPENDIX 2*)
Formamide loading dye (see recipe)
2× SSC (*APPENDIX 2*)
50 μM probe oligonucleotide (DNA or RNA; *UNIT 2.11*) in RNase-free H₂O
≥10 mCi/ml [γ -³²P]ATP (6000 Ci/mmol; ICN Biomedicals)
10× T4 polynucleotide kinase buffer (New England Biolabs)
200 U/μl T4 polynucleotide kinase (New England Biolabs)
Prehybridization/hybridization solution (see recipe), prewarmed to 37°C
2× SSC (*APPENDIX 2*) containing 0.1× (w/v) SDS, prewarmed to 37°C

95°C heating block or water bath
Hybond N+ Nylon Transfer Membrane (Amersham Biosciences)
Extra-thick blotting paper (Bio-Rad), slightly larger than the gel being blotted
Semi-dry transfer apparatus (e.g., Bio-Rad Trans-Blot SD cell)
Sephadex G-25 spin column
Hybridization oven with rotating glass hybridization bottles, 37°C
Image-analysis software (also see *UNIT 10.5*): e.g., QuantityOne (Bio-Rad) or
ImageGauge (Fuji)

Additional reagents and equipment for denaturing polyacrylamide gel
electrophoresis (*UNIT 7.6*), phosphor imaging (*APPENDIX 3A*), and digital
electrophoresis analysis (*UNIT 10.5*)

Prepare RNA sample and run the gel

1. Isolate total RNA from tissue or cell samples with TRIzol reagent according to the manufacturer's instructions. At the end of the procedure, dissolve the RNA to a final concentration of 10 μg/μl in RNase-free water.
2. Prerun 15% denaturing polyacrylamide sequencing (urea/TBE) gel for 15 min at 25 W in 0.5× TBE electrophoresis buffer. Mix 5 μl RNA sample (10 μg/μl) with equal volume of formamide loading dye. Heat 2 min at 95°C, and load onto gel (*UNIT 7.6*).
3. Run the gel at 25 W until the bromophenol blue dye has migrated to the bottom of the gel.

Transfer RNA from gel to membrane

4. Cut a piece of Hybond N+ nylon membrane slightly larger than the gel. Soak the membrane and four pieces of blotting paper of appropriate size in 0.5× TBE buffer for 10 min.
5. Stack two pieces of blotting paper on the anode platform of the transfer cell. Avoid getting air bubbles under or between the papers; remove any that appear by carefully rolling a glass pipet over the surface.
6. Place the membrane on top of the blotting paper and squeeze out air bubbles by rolling a glass pipet over the surface.
7. Carefully transfer the gel from glass plate to the top of the membrane and squeeze out air bubbles.
8. Stack another two pieces of blotting paper on the gel and squeeze out air bubbles.
9. Set the cathode assembly and the safety lid on the sandwich. Transfer for 1 hr at 300 mA.

Prepare membrane for hybridization

10. Disassemble the transfer cell. Remove the paper and the gel. Rinse the membrane in 2× SSC, then place it on a sheet of filter paper and allow it to air dry.
11. Place membrane RNA-side-down on a UV transilluminator (254-nm wavelength) or in a UV light box for the appropriate length of time to covalently attach the RNA to the membrane.

Prepare probe

The authors typically use a chemically synthesized 21–22 nt DNA or RNA oligonucleotide (see *UNIT 2.11* for oligonucleotide synthesis) perfectly complementary to the small RNA to be detected.

12. Set up the 5′-end-labeling reaction by combining the following reagents:

1 μl of 50 μM probe oligonucleotide (DNA or RNA)
1 μl of [γ -³²P]ATP (6000 Ci/mmol, ≥10 mCi/ml)
4 μl of 10× T4 polynucleotide kinase buffer
H₂O to a final volume of 40 μl
1 μl of 200 U/μl T4 polynucleotide kinase.

Incubate reaction 1 hr at 37°C.

13. Pass reaction mixture through Sephadex G-25 spin column (centrifuging per manufacturer's instructions) to remove unincorporated [γ -³²P]ATP.

Perform prehybridization and hybridization

14. Place membrane (from step 11) RNA-side-up in a hybridization bottle and add ~1 ml prewarmed (37°C) prehybridization/hybridization solution per 10 cm² of membrane. Place the bottle in a hybridization oven and incubate with rotation for 30 min at 37°C.
15. Pipet the entire reaction mix (from step 13) into the hybridization bottle and continue to incubate with rotation overnight at 37°C.

Wash membrane

16. Pour off hybridization solution and wash the membrane briefly with 2× SSC for 5 min. Add prewarmed (37°C) 2× SSC containing 0.1% SDS to the bottle. Incubate with rotation for 15 min at 37°C. Replace solution with fresh solution and repeat. Remove final wash solution, blot excess liquid, and wrap with plastic wrap.

Do not allow membrane to dry out if it is to be reprobbed.

Perform phosphor imaging and analyze the hybridization signals

17. Visualize the hybridization signals by phosphor imaging (*APPENDIX 3A*). Analyze hybridization result using appropriate software (*UNIT 10.5*). Subtract the background from the original signal to obtain the specific hybridization.
18. To compare the amount of small RNA in different samples, normalize the amount of small RNA detected to the nonspecific hybridization of the probe to 5S rRNA.

Alternatively, the blot can be reprobbed with a probe specific for 5S rRNA (for Drosophila 5′-CAA CAC GCG GTG TTC CCA AGC CG-3′) or, for Drosophila, the 2S RNA (5′-TAC AAC CCT CAA CCA TAT GTA GTC CAA GCA-3′).

The precise amount (pmol or molecules) of small RNA can be determined if concentration standards of synthetic RNA are included on the blot. The assay is typically linear with respect to concentration over a 10,000-fold range.

NORTHERN HYBRIDIZATION OF RNA USING CHURCH'S HYBRIDIZATION BUFFER

Church's hybridization buffer can be used as an alternative for the standard prehybridization/hybridization solution in this assay. It provides similar sensitivity.

Additional Materials (also see Basic Protocol 2)

Church's hybridization buffer without BSA (see recipe)

1. Perform northern blotting and prepare probe (see Basic Protocol 2, steps 1 through 13).
2. Perform prehybridization and hybridization (see Basic Protocol 2, steps 14 to 15) using Church's hybridization buffer (without BSA) in place of the prehybridization/hybridization solution.
3. Wash membrane and proceed with development and analysis (see Basic Protocol 2, steps 16 to 18).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Church's hybridization buffer without BSA

0.5 M sodium phosphate buffer, pH 7.2 (APPENDIX 2)

1 mM EDTA, pH 8.0 (APPENDIX 2)

7% (w/v) SDS

Store up to 1 year at room temperature

Denaturing solution

500 μ l formamide

162 μ l 12.3 M (37%) formaldehyde

100 μ l MOPS buffer (see recipe)

Make fresh from stock solutions immediately before use

If formamide has a yellow color, deionize as follows: add 5 g of mixed-bed ion-exchange resin (e.g., Bio-Rad AG 501-X8 or X8(D) resins) per 100 ml formamide, stir 1 hr at room temperature, and filter through Whatman #1 filter paper.

CAUTION: Formamide is a teratogen. Handle with care.

Formaldehyde loading buffer

1 mM EDTA, pH 8.0 (APPENDIX 2)

0.25% (w/v) bromphenol blue

0.25% (w/v) xylene cyanol

50% (v/v) glycerol

Store up to 3 months at room temperature

Formamide loading dye

98% (v/v) deionized formamide

10 mM EDTA pH 8.0 (APPENDIX 2)

0.025% (w/v) xylene cyanol

0.025% (w/v) bromphenol blue

Store indefinitely at -20°C

Glyoxal, 6 M, deionized

Immediately before use, deionize glyoxal by passing through a small column of mixed-bed ion-exchange resin (e.g., Bio-Rad AG 501-X8 or X8(D) resins) until the pH is >5.0 .

Glyoxal loading buffer

10 mM sodium phosphate, pH 7.0 (see recipe)
0.25% (w/v) bromphenol blue
0.25% (w/v) xylene cyanol
50% (v/v) glycerol
Store up to 3 months at room temperature

MOPS buffer

0.2 M MOPS [3-(*N*-morpholino)-propanesulfonic acid], pH 7.0
0.5 M sodium acetate
0.01 M EDTA
Store up to 3 months at 4°C
Store in the dark and discard if it turns yellow.

MOPS running buffer, 10×

0.4 M MOPS, pH 7.0
0.1 M sodium acetate
0.01 M EDTA
Store up to 3 months at 4°C

Prehybridization/hybridization solution

5× SSPE (see recipe)
5× Denhardt solution (*APPENDIX 2*)
50% (v/v) formamide
0.5% (w/v) SDS
72 µg/ml denatured herring sperm DNA (Promega)
Make fresh from stock solutions immediately before use
The herring sperm DNA is denatured by heating 10 min at 75°C just before it is added.

Sodium phosphate, pH 7.0, 100 mM and 10 mM

100 mM stock solution:
5.77 ml 1 M Na₂HPO₄
4.23 ml 1 M NaH₂PO₄
H₂O to 100 ml
Store up to 3 months at room temperature

10 mM solution:
Dilute 100 mM stock 1/10 with H₂O
Store up to 3 months at room temperature

SSPE, 10×

1.5 M NaCl
50 mM NaH₂PO₄·H₂O
5 mM EDTA
Store indefinitely at room temperature

Stripping solution

1% (w/v) SDS
0.1× SSC (*APPENDIX 2*)
40 mM Tris·Cl, pH 7.5 to 7.8 (*APPENDIX 2*)
Store up to 1 year at room temperature
Where formamide stripping is desired, prepare the above solution and add an equal volume of formamide just before use.

COMMENTARY

Background Information

The development of Southern blotting (*UNIT 2.9A*; Southern, 1975) was quickly followed by an equivalent procedure for the immobilization of gel-fractionated RNA (Alwine et al., 1977). The term northern blotting, initially used in a humorous fashion, has become enshrined in molecular biology jargon. Northern hybridization is a standard procedure for identification and size analysis of RNA transcripts and RNA slot blotting is frequently used to assess the expression profiles of tissue-specific genes (Kafatos et al., 1979).

Procedures for the removal of hybridization probes from northern blots are similar to those for Southern blots, except that NaOH is omitted to prevent hydrolysis of the RNA, and formamide may be included.

The recent discovery of microRNAs (miRNAs) revealed an entire new class of molecules that regulate gene expression. miRNAs are small noncoding RNAs that range from 20 to 30 nucleotides, making traditional formaldehyde-agarose gel electrophoresis unsuitable for their size fractionation. The modified northern protocol here (Basic Protocol 2) combines denaturing polyacrylamide gel electrophoresis (PAGE), which is ideal for the separation of small RNAs, with standard blotting and hybridization procedures.

Critical Parameters

Gel electrophoresis and northern blotting

The main distinction between northern and Southern blotting lies with the initial gel fractionation step. Because single-stranded RNA can form secondary structures, samples must be electrophoresed under denaturing conditions to ensure good separation.

A variety of denaturants for RNA gels have been used, including formaldehyde (Basic Protocol 1; Lehrach et al., 1977), glyoxal/DMSO (Alternate Protocol 1; Thomas, 1980), and the highly toxic methylmercuric chloride (Bailey and Davidson, 1976). Because of the substantial health risks, use of methylmercuric chloride is not advised. Formaldehyde gels are recommended, as they are easy to run and reasonably reliable. The formaldehyde must be rinsed from the gel before the transfer is set up, but this is a minor inconvenience compared to assembling the buffer recirculation system required for electrophoresis of glyoxal-denatured RNA.

Total cellular RNA (*UNITS 4.1-4.4*) or poly(A)⁺ RNA (*UNIT 4.5*) can be used for northern transfers and slot blots. Total RNA is less satisfactory because nonspecific hybridization, however slight, to one or both of the highly abundant rRNA molecules will lead to a substantial hybridization signal. Any hybridizing band that appears in the vicinity of an rRNA should be treated with suspicion and its identity confirmed by blotting with poly(A)⁺ RNA.

Under ideal conditions, a band that contains as little as 1 pg of RNA can be detected by northern hybridization with a probe labeled to a specific activity of 10⁹ dpm/μg. In practice, the effective detection limit with an overnight exposure is ~5 pg RNA. An mRNA is usually considered to be abundant if it constitutes >1% of the mRNA fraction. In a typical mammalian cell, the mRNA fraction makes up about 0.5% of total RNA, so >5 pg of an abundant mRNA should be present in just 100 ng of total RNA. If 10 μg of total RNA is transferred, abundant mRNAs should give strong hybridization signals and less abundant ones (down to 0.01% of the mRNA population) should be detectable with an overnight exposure. For rarer molecules, the poly(A)⁺ fraction must be prepared. In this sample 3 μg is sufficient for detecting an mRNA that makes up 0.0002% of the polyadenylated population.

Unlike probing for mRNA, which often requires enrichment by poly(A) selection prior to analysis, total cellular RNA can always be used for the detection of miRNA, because individual miRNA species can be present in thousands to tens of thousands of copies per cell. The highly abundant rRNA and microRNAs are well separated in a 15% denaturing polyacrylamide gel; thus, the nonspecific hybridization to rRNA will not affect the interpretation of the desired hybridizing signal.

RNA slot blots

Although easy to perform, RNA slot-blot hybridization is one of the most problematic techniques in molecular biology. A number of criteria must be satisfied if slot blotting is to be used to make meaningful comparisons of mRNA abundance in different extracts. The first requirement is that equal amounts of RNA must be loaded in each slot. In practice this is difficult to achieve, especially if RNA concentrations are estimated by absorbance spectroscopy (*APPENDIX 3D*), which is subject to errors

due to the small quantities being measured and the presence of contaminants such as protein and DNA.

Even if equal amounts of RNA are loaded, a difference in hybridization signal does not necessarily mean that the gene whose transcript is being studied is more active in a particular tissue. The analysis provides information on the abundance of an mRNA (i.e., the fraction of total RNA that it constitutes), not its absolute amount. To illustrate this point, consider a tissue in which a highly active gene is switched on at time t , where the transcripts of this gene constitute 0% of the mRNA at $t - 1$ but 20% of the mRNA at $t + 1$. If the slot blots of RNA from $t - 1$ and $t + 1$ are probed with the highly active gene, there will be a clear increase in hybridization signal after time t . In contrast, hybridization of the same slot blots with a second gene whose transcription rate is unchanged will show a decreased hybridization signal at $t + 1$. Transcripts of this gene are present in the same absolute amounts at $t - 1$ and $t + 1$, but their abundance decreases as the total mRNA population becomes larger due to activation of the highly expressed gene. To the unwary, the result of the hybridization analysis could appear to indicate down-regulation of a gene whose expression rate in fact remains constant.

Choice of membrane and transfer system

General information relating to the choice of membrane for a nucleic acid transfer is given in the Commentary to *UNIT 2.9A*. Because of the greater tensile strength of nylon, together with the fact that the RNA can be bound covalently by UV cross-linking, most transfers are now carried out using nylon rather than nitrocellulose. Nylon has the added advantage of being able to withstand the highly stringent conditions (50% formamide at 60°C) that may be required during hybridization with an RNA probe; nitrocellulose tends to disintegrate under these conditions.

For DNA transfer, a major advantage of positively charged nylon is that nucleic acids become covalently bound to the membrane if the transfer is carried out with an alkaline buffer. RNA can also be immobilized on positively charged nylon by alkaline transfer, but the procedure is not recommended as the alkaline conditions result in partial hydrolytic degradation of the RNA. This hydrolysis is difficult to control and smaller molecules are easily broken down into fragments too short for efficient retention by the membrane (see Table 2.9.1). This results in a loss of signal after hybridiza-

tion, a problem that is exacerbated by the increased background caused by lengthy exposure of the membrane to the alkaline solution. Only if the signals are expected to be strong should an alkaline transfer be considered. In this case, Basic Protocol 1 should be modified as follows: omit the pre-transfer alkaline hydrolysis (step 10), use 8 mM NaOH rather than 20× SSC as the transfer buffer, do not transfer for more than 6 hr, and rinse the membrane in 2× SSC/0.1% (w/v) SDS rather than plain SSC immediately after transfer (step 22). If using Alternate Protocol 1, omit step 7 as well, as the alkaline transfer buffer removes glyoxal from the RNA.

The standard northern transfer system can be modified as described for Southern blotting (*UNIT 2.9A*). Aqueous transfers onto nylon can be performed using a variety of buffers, although SSC is still most frequently used. Changes can be made to the transfer time and architecture of the blot (e.g., downward transfer; Chomczynski, 1992), and alternative methods such as electroblotting (Smith et al., 1984) and vacuum transfer (Peferoen et al., 1982) can be used.

Hybridization procedures

Hybridization analysis of an RNA blot is subject to the same considerations as DNA hybridization (see *UNIT 2.10* Commentary). The factors that influence sensitivity and specificity are the same, and incubation times, hybridization solutions, probe length, and mechanics of hybridization all have similar effects. There are just two additional points that need to be made with respect to RNA blots.

The first point is that formamide is almost always used in RNA hybridization solutions. The primary reason for this is to permit a lower hybridization temperature to be used, minimizing RNA degradation during the incubations.

The second point concerns the stability of the hybrids formed between the immobilized RNA and the probe molecules. For a DNA probe the relevant equation is (Casey and Davidson, 1977):

$$T_m = 79.8^\circ\text{C} + 18.5(\log M) + 0.58(\%GC) + 11.8(\%GC)^2 - 0.50(\%form) - \frac{820}{L}$$

and for an RNA probe (Bodkin and Knudson, 1985):

$$T_m = 79.8^\circ\text{C} + 18.5(\log M) + 0.58(\%GC) + 11.8(\%GC)^2 - 0.35(\%form) - \frac{820}{L}$$

where T_m is the melting temperature, M is the molarity of monovalent cations, %GC is the

percentage of guanosine and cytosine nucleotides in the DNA, %form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. What these equations indicate is that an RNA-RNA hybrid is more stable than a DNA-RNA hybrid: if %form is 50%, the T_m for an RNA-RNA hybrid is 7.5°C higher than that for an equivalent RNA-DNA hybrid. The greater stability of the RNA-RNA hybrid means that an RNA probe requires a more stringent hybridization and washing regime than a DNA probe (e.g., hybridization at 60°C in 50% formamide and final wash at 68°C in 0.1× SSC/0.1% SDS).

Reprobing conditions

Nitrocellulose presents problems regarding both membrane integrity and RNA retention (see UNIT 2.10). UV cross-linking of RNA to a neutral nylon membrane presents optimal conditions for northern blot reprobing.

Because of the sensitivity of RNA to alkaline hydrolysis, NaOH, which is included in protocols for probe removal from Southern blots, should not be used when removing probes from northern blots. It is recommended that probes be removed prior to membrane storage, because unstripped probes remain permanently attached if the blot dries.

Troubleshooting

The appearance of the agarose gel after staining gives a first indication of how successful a northern experiment is likely to be. If total RNA has been used, the rRNA bands should be clear and sharp (Fig. 4.9.1) with no “smearing” toward the positive electrode. The only exception is when the RNA has been prepared by the guanidinium isothiocyanate procedure (UNIT 4.2), in which case some smearing is normal. If the rRNA bands are not sharp, the RNA preparation may be of poor quality (usually because insufficient care has been taken in establishing an RNase-free environment) or the denaturing gel electrophoresis system may not have worked adequately. If the latter problem is suspected, make sure that the formaldehyde concentration in the gel is 2.2 M or, if glyoxal denaturation has been used, that the buffer recircularization is sufficient to maintain the gel pH at 7.0. Whatever the problem, if the rRNAs are not distinct, there is no point in proceeding with the transfer as the bands obtained after hybridization will also be fuzzy. In fact, even if the rRNA bands are clear there is no guarantee that the mRNAs are intact.

An indication of the efficiency of transfer onto nylon can be obtained by staining the membrane with methylene blue (see Basic Protocol 1, step 24), but often a problem with transfer is not recognized until after hybridization. If poor signals are obtained, the troubleshooting section of UNIT 2.10 (including Table 2.10.4) should be consulted to identify the likely cause. Note that it is relatively easy to detach RNA from a membrane before immobilization, so some loss may occur when the membrane is rinsed in 2× SSC to wash off agarose fragments and leach out salt (see Basic Protocol 1, step 22). If necessary, this rinse can be postponed until immediately before hybridization, after the RNA has been immobilized.

Other problems, such as high backgrounds, extra bands, and difficulties with probe stripping, should be dealt with by referring to Table 2.10.4.

To increase the sensitivity of the miRNA assay, RNA probes can be used instead of DNA probes. In some cases, RNA probes work better to detect miRNA precursors (~60 to 70 nt long) which contain the immature miRNAs in a stem-loop whose structure can prevent the hybridization of DNA probes.

Anticipated Results

Using either a nylon or nitrocellulose membrane and a probe labeled to $\geq 5 \times 10^8$ dpm/ μ g, it should be possible to detect transcripts that represent 0.01% of the mRNA population with a blot of 10 μ g total mammalian RNA or 0.0002% of the population with a blot of 3 μ g poly(A)⁺ RNA.

It should be possible to detect small RNAs at levels as low as 0.3 fmol miRNA by following the Basic Protocol 2. Small RNAs that differ by as little as 1 nt (or even by a single phosphate group) can be separated on 15% denaturing polyacrylamide gel (50 to 100 cm), particularly when a long gel is used. For long gels, only the lower portion of the gel is used for transfer to the membrane.

Time Considerations

Traditional blots

A northern experiment can be completed in 3 days. The agarose gel is prepared and electrophoresed during the first day and the transfer carried out overnight. On the second day the blot is prehybridized and then hybridized overnight. Washes are completed early on the third day. A slot-blot experiment takes only two days, as the blot can be prepared and prehybridized

on the first day, hybridized overnight, and washed on the second day.

The length of time needed for the autoradiography depends on the abundance of the target sequences in the blotted RNA. Adequate exposure can take anything from overnight to several days.

With blots that are intended for reprobing, stripping procedures can be completed in ~1 hr, not including the verification steps.

miRNA blots

Using a semidry transfer system, the northern experiment can be completed in 2 days.

The length of time needed for the autoradiography depends on the abundance of the small RNA sequence and the detection system used. Optimal exposure can range from overnight to several days.

Under ideal conditions, a band that contains as little as 0.3 fmol of small RNA can be detected by northern hybridization with a 1-day exposure to a phosphor imager plate when scanned at 25 μm resolution. The amount of RNA loaded and the exposure time may vary with the abundance of the individual miRNA species. For abundant miRNAs, loading of 5 μg total RNA and overnight exposure will give strong hybridization signals.

Literature Cited

- Alwine, J.C., Kemp, D.J., and Stark, G.R. 1977. Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proc. Natl. Acad. Sci. U.S.A.* 74:5350-5354.
- Bailey, J.M. and Davidson, N. 1976. Methylmercury as a reversible denaturing agent for agarose gel electrophoresis. *Anal. Biochem.* 70:75-85.
- Bodkin, D.K. and Knudson, D.L. 1985. Assessment of sequence relatedness of double-stranded RNA genes by RNA-RNA blot hybridization. *J. Virol. Methods* 10:45-52.
- Casey, J. and Davidson, N. 1977. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. *Nucl. Acids Res.* 4:1539-1552.
- Chomczynski, P. 1992. One-hour downward alkaline capillary transfer for blotting of DNA and RNA. *Anal. Biochem.* 201:134-139.
- Herrin, D.L. and Schmidt, G.W. 1988. Rapid, reversible staining of Northern blots prior to hybridization. *BioTechniques* 6:196-200.
- Kafatos, F.C., Jones, C.W., and Efstratiadis, A. 1979. Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. *Nucl. Acids Res.* 7:1541-1552.
- Lehrach, H., Diamond, D., Wozney, J.M., and Boedtker, H. 1977. RNA molecular weight determinations by gel electrophoresis under denaturing conditions: A critical reexamination. *Biochemistry* 16:4743-4751.
- Peferoen, M., Huybrechts, R., and De Loof, A. 1982. Vacuum-blotting: A new simple and efficient transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to nitrocellulose. *FEBS Lett.* 145:369-372.
- Smith, M.R., Devine, C.S., Cohn, S.M., and Lieberman, M.W. 1984. Quantitative electrophoretic transfer of DNA from polyacrylamide or agarose gels to nitrocellulose. *Anal. Biochem.* 137:120-124.
- Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
- Thomas, P.S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. U.S.A.* 77:5201-5205.
- Wilkinson, M. 1991. Purification of RNA. *In* Essential Molecular Biology: A Practical Approach, Vol. 1 (T.A. Brown, ed.) pp. 69-87. IRL Press, Oxford.

Contributed by Terry Brown
University of Manchester Institute of
Science and Technology
Manchester, United Kingdom

Karol Mackey (probe removal)
Molecular Research
Cincinnati, Ohio

Tingting Du (miRNA blots)
University of Massachusetts Medical School
Worcester, Massachusetts

Identification of Newly Transcribed RNA

Newly transcribed RNA can be identified using the nuclear runoff transcription assay. Isolated nuclei, free of membranes and cytoplasmic debris, are required for the assay. Cell lysis that does not allow the isolation of nuclei free of cell membranes and cytoplasmic material often results in poor incorporation of ^{32}P -labeled UTP into nascent transcripts. Although there is no way to predict what cell types present this problem, many adherent cell lines and lymphocytes isolated from murine spleen or thymus do. Interestingly, very few nonadherent cell lines have posed this problem. Isolating nuclei by detergent lysis of cells (basic protocol) works well for many tissue culture cell lines but may not be appropriate for all cell lines and many tissues. Detergent lysis and Dounce homogenization (first alternate protocol) or cell lysis in an isoosmotic solution and centrifugation through a sucrose cushion (second alternate protocol) are alternative methods for preparing nuclei. The support protocol describes preparation of the cDNA nitrocellulose filter strips that are used to detect the presence of specific transcripts in the nuclear runoff transcription assay.

NOTE: Experiments involving RNA require careful technique to prevent RNA degradation; see Chapter 4 Section I introduction.

BASIC PROTOCOL

NUCLEAR RUNOFF TRANSCRIPTION IN MAMMALIAN CELLS

Nuclear runoff transcription is currently the most sensitive procedure for measuring specific gene transcription as a function of cell state. Nuclei are first isolated from cultured cells or tissues and frozen in liquid nitrogen. Thawed nuclei are incubated with ^{32}P -labeled UTP and unlabeled NTPs to label nascent RNA transcripts. ^{32}P -labeled RNA is purified and used to detect specific RNA transcripts by hybridization to cDNAs immobilized on nitrocellulose membranes.

Materials

For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

- Cultures of mammalian cells (UNIT 9.0) or freshly isolated lymphoid cells
- Phosphate-buffered saline (PBS; APPENDIX 2), made fresh and ice cold
- Nonidet P-40 (NP-40) lysis buffer A (see recipe)
- Glycerol storage buffer (see recipe), ice cold
- 2× reaction buffer with and without nucleotides (see recipes)
- 10 mCi/ml [α - ^{32}P]UTP (760 Ci/mmol)
- 1 mg/ml DNase I (RNase-free; see recipe)
- HSB buffer (see recipe)
- SDS/Tris buffer (see recipe)
- 20 mg/ml proteinase K
- 25:24:1 (v/v/v) buffered phenol/chloroform/isoamyl alcohol (UNIT 2.1)
- 10% (v/v) trichloroacetic acid (TCA)/60 mM sodium pyrophosphate
- 10 mg/ml tRNA (UNIT 4.6)
- 5% (v/v) TCA/30 mM sodium pyrophosphate
- DNase I buffer (see recipe)
- 0.5 M EDTA, pH 8.0
- 20% (w/v) SDS
- Elution buffer (see recipe)
- 1 M NaOH
- 1 M HEPES (free acid)

Identification of Newly Transcribed RNA

4.10.1

Contributed by Michael E. Greenberg and Timothy P. Bender

Current Protocols in Molecular Biology (1997) 4.10.1-4.10.11

Copyright © 1997 by John Wiley & Sons, Inc.

3 M sodium acetate, pH 5.2
100% ethanol
TES solution (see recipe)
TES/NaCl solution (see recipe)
cDNA plasmid(s) immobilized on nitrocellulose membrane (support protocol)
2× SSC (APPENDIX 2)
10 mg/ml heat-inactivated RNase A (UNIT 7.3)
Rubber policeman
15- and 50-ml conical polypropylene centrifuge tubes
Beckman JS-4.2 and JA-20 rotors or equivalent
30° and 65°C shaking water baths
42° and 65°C water baths
0.45-µm HA filters (Millipore)
30-ml Corex tube, silanized (APPENDIX 3)
Whatman GF/F glass fiber filters
5-ml plastic scintillation vials
Whatman 3MM filter paper

NOTE: Keep cells and nuclei on ice until the nuclei are frozen.

Isolate nuclei

- 1a. *For cultures of adherent cells:* Remove medium from monolayer cultures (5×10^7 cells per assay) and place cells on ice. Rinse twice with 5 ml ice-cold PBS. Scrape flask with a rubber policeman and collect cells in a 15-ml centrifuge tube. Centrifuge 5 min at $500 \times g$ (1500 rpm in JS-4.2 rotor), 4°C. Remove supernatant.
- 1b. *For cultures of nonadherent cells:* Pipet up and down several times to resuspend cells (5×10^7 cells per assay) and transfer cells and medium to a 50-ml conical centrifuge tube. Centrifuge 5 min at $500 \times g$ (1500 rpm in JS-4.2 rotor), 4°C, and remove supernatant. Wash by gently resuspending pellet in 5 ml ice-cold PBS, adding 45-ml ice-cold PBS, and collecting cells by centrifuging 5 min at $500 \times g$. Remove supernatant. Wash cells once more with PBS and remove supernatant.
- 1c. *For freshly isolated lymphoid cells:* Transfer lymphoid cells (5×10^7 cells per assay) removed directly from organ to a 50-ml conical centrifuge tube. Centrifuge 5 min at $500 \times g$ (1500 rpm in JS-4.2 rotor), 4°C, and remove supernatant. Wash by resuspending pellet in 5 ml ice-cold PBS, adding 45 ml PBS, and collecting cells by centrifuging 5 min at $500 \times g$, 4°C. Remove supernatant. Wash cells once more with PBS and remove supernatant.

Lymphoid cell nuclei are more fragile than other cell types, so a gentle procedure (e.g., second alternate protocol) may be required to isolate intact nuclei. It is not necessary to eliminate erythrocytes from lymphoid cells prior to preparation of nuclei.

5×10^7 cells are required for each nuclear runoff transcription assay.

2. Loosen cell pellet by gently vortexing 5 sec. Add 4 ml NP-40 lysis buffer A, continuing to vortex as buffer is added. After lysis buffer is completely added, vortex cells 10 sec at half maximal speed.

Gentle vortexing (at a setting of six) uniformly resuspends cells and inhibits clumping. The same method is used to resuspend nuclei.

3. Incubate lysed cells 5 min on ice. Examine a few microliters of cell lysate on a hemacytometer with a phase-contrast microscope to ensure that cells have uniformly lysed and nuclei appear free of cytoplasmic material. Centrifuge 5 min at $500 \times g$, 4°C. Remove supernatant.

Supernatant contains cytoplasmic RNA that can be purified as described in UNITS 4.1 & 4.5, if desired.

4. Resuspend the nuclear pellet in 4 ml NP-40 lysis buffer A by vortexing as described in step 2. Centrifuge 5 min at $500 \times g$, 4°C . Discard supernatant and resuspend nuclei in 100 to 200 μl glycerol storage buffer by gently vortexing. Freeze resuspended nuclei in liquid nitrogen.

Nuclei are stable in liquid nitrogen for >1 year.

Perform nuclear runoff transcription

5. Thaw 200 μl frozen nuclei at room temperature and transfer to a 15-ml conical polypropylene centrifuge tube. Immediately add 200 μl of $2\times$ reaction buffer with nucleotides plus 10 μl of 10 mCi/ml [α - ^{32}P]UTP. Incubate 30 min at 30°C with shaking.

This reaction is done in a 15-ml polypropylene tube rather than a microcentrifuge tube to reduce the possibility of spilling radioactive materials.

6. Mix 40 μl of 1 mg/ml RNase-free DNase I and 1 ml HSB buffer. Add 0.6 ml of this solution to labeled nuclei and pipet up and down 10 to 15 times with a Pasteur pipet to mix thoroughly. Incubate 5 min at 30°C .
7. Add 200 μl SDS/Tris buffer and 10 μl of 20 mg/ml proteinase K. Incubate for 30 min at 42°C .

DNA and protein should be well digested and a fairly uniform solution should be obtained. The presence of a substantial amount of particulate matter usually indicates that either DNase I or proteinase K treatment was not effective and should be repeated. It may be necessary first to ethanol precipitate the RNA, then to repeat the treatment with DNase I and proteinase K with fresh reagents.

Extract and precipitate RNA

8. Extract sample with 1 ml 25:24:1 buffered phenol/chloroform/isoamyl alcohol. Centrifuge 5 min at $800 \times g$ (2000 rpm in JS-4.2 rotor), at or below room temperature. Transfer aqueous phase to a clean 15-ml polypropylene centrifuge tube.
9. Add 2 ml water, 3 ml of 10% TCA/60 mM sodium pyrophosphate, and 10 μl of 10 mg/ml *E. coli* tRNA carrier to aqueous phase. Incubate 30 min on ice.
10. Filter TCA precipitate onto 0.45- μm Millipore HA filter. Wash filter three times with 10 ml of 5% TCA/30 mM sodium pyrophosphate.

If the HA filter clogs and filters very slowly, a Whatman GF/A glass fiber filter can be used instead.

11. Transfer filter to a glass scintillation vial. Incubate with 1.5 ml DNase I buffer and 37.5 μl of 1 mg/ml RNase-free DNase I for 30 min at 37°C . Quench the reaction by adding 45 μl of 0.5 M EDTA and 68 μl of 20% SDS.
12. Heat sample 10 min at 65°C to elute the RNA. Remove supernatant and save. Add 1.5 ml elution buffer to filter and incubate 10 min at 65°C . Remove supernatant and combine with original supernatant.

This procedure removes >95% of the radioactivity from the filter.

13. Add 4.5 μl of 20 mg/ml proteinase K to 3 ml supernatant containing ^{32}P -labeled RNA. Incubate 30 min at 37°C .

14. Extract 3 ml RNA solution once with 3 ml of 25:24:1 buffered phenol/chloroform/isoamyl alcohol.
15. Remove aqueous phase to a silanized 30-ml Corex tube. Add 0.75 ml of 1 M NaOH to aqueous phase. Let stand 10 min on ice. Quench reaction by adding 1.5 ml of 1 M HEPES.
16. Precipitate RNA by adding 0.53 ml of 3 M sodium acetate and 14.5 ml of 100% ethanol. Incubate 30 min on dry ice or overnight at -20°C .
17. Centrifuge RNA 30 min at $10,000 \times g$ (9000 rpm in JA-20 rotor), 4°C . Remove ethanol and resuspend pellet in 1 ml TES solution. Shake 30 min at room temperature.
RNA should be completely dissolved.
18. Count a 5- μl aliquot of each sample in duplicate by spotting onto Whatman GF/F glass fiber filters. If necessary, dilute sample by adding TES solution to adjust ^{32}P -labeled RNA to $\geq 5 \times 10^6$ cpm/ml.

Hybridize RNA to cDNA

19. Mix 1 ml RNA solution with 1 ml TES/NaCl solution. In a 5-ml plastic scintillation vial, hybridize to cDNA immobilized on nitrocellulose membrane strip for 36 hr at 65°C with shaking.

Use a vial rather than a plastic bag for hybridization to ensure reproducible quantitative hybridization of the same number of cpm to each sample within a given experiment. Coil the nitrocellulose strip before inserting it into the scintillation vial and be sure the strip is completely immersed in hybridization solution.

20. After hybridization, transfer the strips to a 50-ml tube and wash filter in 25 ml of $2\times$ SSC 1 hr at 65°C . Repeat wash once with fresh $2\times$ SSC.
21. Remove filter to a glass scintillation vial containing 8 ml of $2\times$ SSC and 8 μl of 10 mg/ml RNase A. Incubate without shaking 30 min at 37°C .
22. Wash filter once more in 25 ml of $2\times$ SSC 1 hr at 37°C . Blot filter dry on Whatman 3MM filter paper. Unravel the strips, tape them to Whatman 3MM filter paper, and expose to X-ray film.

Appropriate exposure time and conditions will vary depending on the experiment.

ISOLATION OF NUCLEI BY DOUNCE HOMOGENIZATION

This protocol is used for isolation of nuclei from cell types that do not give clean nuclear preparations after lysis in NP-40 lysis buffer A. If cells are not lysed by treatment with NP-40 lysis buffer A (basic protocol), the addition of Dounce homogenization in NP-40 lysis buffer B is usually sufficient to lyse the cells.

Additional Materials

*For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.*

- Lysis buffer (see recipe), ice cold
- Nonidet P-40 (NP-40) lysis buffer B (see recipe)
- Glycerol storage buffer (see recipe), ice cold
- Dounce homogenizer with type B pestle, ice cold
- 1.5-ml microcentrifuge tubes, chilled on dry ice

NOTE: Keep cells and nuclei on ice until the nuclei are frozen.

ALTERNATE PROTOCOL

Preparation and Analysis of RNA

4.10.4

Harvest cells

1. Harvest and wash cells as in basic protocol step 1.
2. Remove supernatant and loosen cell pellet by gently vortexing 5 sec. Resuspend cell pellet to a single-cell suspension in 5 to 10 ml ice-cold lysis buffer. Add ice-cold lysis buffer to a total of 40 ml and rock tube back and forth for several seconds to distribute the cells.
3. Pellet cells at $500 \times g$ (1500 rpm in JS-4.2 rotor), 4°C . Remove and discard supernatant. Resuspend pellet in 1 ml lysis buffer per 5×10^7 cells and vortex gently to mix.

After centrifugation, the pellet should appear to be two to three times its initial size.

4. Add 1 ml NP-40 lysis buffer B per 5×10^7 cells and mix by gently rocking the tube.

Break cells and collect nuclei

5. Transfer cells to an ice-cold Dounce homogenizer and break them with ten strokes of a B pestle or until nuclei appear free of membrane components by phase-contrast microscopy.
6. Transfer homogenized cells to a plastic 50-ml conical centrifuge tube and pellet nuclei by centrifugation 5 min at $500 \times g$, 4°C .

Pellet should now be approximately one-third to one-half the starting volume and appear opaque white.

7. Carefully remove supernatant with a Pasteur pipet attached to a vacuum supply. Tilt the tube sideways so supernatant is pulled away from the pellet. Remove any bubbles or liquid that remain on the side of the tube. Return pellet to an ice bucket.
8. Loosen pelleted nuclei by gentle vortexing. Add 200 μl ice-cold glycerol storage buffer per 5×10^7 nuclei and resuspend pellet by pipetting up and down.

Nuclei will be clumped at first but will disperse with continued pipetting. Pipetting should be steady but not hard enough to cause bubbles.

9. Aliquot 210 μl ($\sim 5 \times 10^7$ nuclei) into chilled 1.5-ml microcentrifuge tube and immediately return tube to dry ice. Store nuclei at -70°C or in liquid nitrogen.

Frozen nuclei are stable for at least 1 year.

10. Proceed with nuclear runoff transcription assay starting with basic protocol step 5.

ISOLATION OF NUCLEI BY SUCROSE GRADIENT CENTRIFUGATION

Quality of nuclei used in nuclear runoff protocols is a major determinant in the success of the experiment. Normal lymphocytes in particular can present a problem because the nuclei are more fragile. In this protocol, cells are resuspended in an isoosmotic buffer containing nonionic detergent, then lysed by Dounce homogenization. Nuclei are collected by ultracentrifugation through a sucrose cushion and are quite clean and free of contaminating membranes and cytoplasmic components. Typically, 10% to 30% more [α - 32 P]UTP is incorporated into nascent transcripts of nuclei prepared by this method. The density of nuclei varies with cell type so a pilot experiment should be performed to verify that these conditions (which work well for murine splenic lymphocytes and other vertebrate cells) result in a nuclear pellet.

Additional Materials

For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

Sucrose buffer I (see recipe), ice cold

Sucrose buffer II (see recipe)

Dounce homogenizer with B pestle, ice cold

Polyallomer centrifuge tubes ($\frac{9}{16} \times 3\frac{3}{4}$ in., Beckman) for SW 40.1 rotor

Ultracentrifuge and SW 40.1 rotor or equivalent

1.5-ml microcentrifuge tubes, chilled on dry ice

NOTE: Keep cells and nuclei on ice until the nuclei are frozen.

Harvest and lyse cells

1. Harvest and wash cells as in basic protocol step 1.
2. Loosen cell pellet by gently vortexing 5 sec. Resuspend cell pellet in 4 ml ice-cold sucrose buffer I. Examine a small aliquot of cells for lysis with a phase-contrast microscope.

Many cell types will lyse at this point and do not require Dounce homogenization. If cells have lysed, proceed directly to step 4.

3. Transfer cells to an ice-cold Dounce homogenizer and break the cells with five to ten strokes of a B pestle or until the nuclei appear free of cytoplasmic tags. Check a few microliters of cells with a phase-contrast microscope to be sure they are uniformly lysed.
4. Transfer nuclei to a clean 50-ml conical polypropylene centrifuge tube and add 4 ml sucrose buffer II. Mix by gentle pipetting and inversion.

The final concentration of sucrose in cell homogenate should be sufficient to prevent a large buildup of debris at the interface between homogenate and sucrose cushion. The amount of sucrose buffer II added to cell homogenate may need to be adjusted.

Collect nuclei

5. Add 4.4 ml sucrose buffer II to polyallomer SW 40.1 tube.

Sucrose buffer II serves as the sucrose cushion. Unlysed cells will not sediment through the sucrose cushion. If these conditions do not result in a nuclear pellet, adjust the concentration of sucrose in sucrose buffer II.

6. Carefully layer nuclei (from step 4) onto the sucrose cushion. Use sucrose buffer I to top off the gradient.

Do not centrifuge more than 2×10^8 nuclei per tube.

7. Centrifuge the gradient 45 min at $30,000 \times g$ (15,500 rpm in SW 40.1 rotor), 4°C.
8. Remove supernatant by vacuum aspiration. Tilt the tube sideways so supernatant is pulled away from the pellet and remove any bubbles or liquid that remain on the side of the tube. Return tube to an ice bucket.

Nuclei should form a tight pellet at the bottom of the tube and there may be some debris caught at the interface between sucrose buffers I and II. If the cells did not lyse during Dounce homogenization, nuclei will not pellet. Thus, it is important to be sure that the majority of the cells are clearly lysed in step 3. If the pellet appears as a gelatinous mass, nuclei have lysed and the pellet should be discarded.

9. Loosen nuclear pellet by gently vortexing 5 sec. Add 200 μ l ice-cold glycerol storage buffer per 5×10^7 nuclei and resuspend nuclei by pipetting up and down.

Nuclei will be clumped at first but will disperse with continued pipetting. Pipetting should be steady but should not create air bubbles.

10. Aliquot 210 μ l ($\sim 5 \times 10^7$ nuclei) into chilled microcentrifuge tube and immediately return tube to dry ice. Store frozen nuclei at -70°C or in liquid nitrogen.

Frozen nuclei are stable for at least 1 year.

11. Proceed with nuclear runoff transcription assay starting at basic protocol step 5.

SUPPORT PROTOCOL

PREPARATION OF NITROCELLULOSE FILTERS FOR NUCLEAR RUNOFF TRANSCRIPTION ASSAY

cDNA plasmids are linearized and immobilized on nitrocellulose membrane filters for hybridization in the nuclear runoff transcription assay. Filters are prepared in advance and may be stored at least 6 months. Prior to their use for hybridization, filters are cut into strips that contain the cDNA plasmids of interest.

Materials

*For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.*

cDNA plasmid
1 M NaOH
6 \times SSC (**APPENDIX 2**)

0.45- μ m nitrocellulose membrane
Slot blot apparatus
80°C vacuum oven

Additional reagents and equipment for restriction endonuclease digestion (**UNIT 3.1**)

NOTE: Wear gloves and handle membranes with blunt-ended forceps.

1. Linearize 200 μ g cDNA plasmid by digestion with an appropriate restriction enzyme.

*It is usually not necessary to phenol extract or ethanol precipitate the DNA after digestion if BSA is absent from the restriction enzyme digestion buffer. If the buffer contains BSA, extract plasmid DNAs with phenol/chloroform/isoamyl alcohol, ethanol precipitate, and resuspend in TE or similar buffer prior to denaturation (**UNIT 2.1**).*

2. Add 49 μ l of 1 M NaOH to linearized DNA (200 μ g in 440 μ l). Incubate 30 min at room temperature to denature DNA.
3. Add 4.9 ml of 6 \times SSC to DNA and place on ice to neutralize the sample.

4. Set up slot blot apparatus with 0.45- μ m nitrocellulose membrane. Apply 125 μ l of sample (~5 μ g cDNA plasmid) to each slot under a low vacuum provided by a water aspirator. Rinse each slot with 500 μ l of 6 \times SSC.
5. Use a blue pencil to mark the location on the membrane of slots containing DNA.

It is difficult to detect the location of slots once nitrocellulose has dried. Mark the edge of the slot so that the nitrocellulose strip can be trimmed very close to the edge of the slot. It is possible to minimize the volume of hybridization solution if a narrow filter strip is used.
6. Air dry nitrocellulose filter overnight. Bake filter 2 hr in an 80°C vacuum oven. Store filter in a vacuum desiccator at either room temperature or 4°C.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

DNase I, RNase-free (1 mg/ml)

Adjust pH of 0.1 M iodoacetic acid/0.15 M sodium acetate to 5.3 and filter sterilize. Add sterile solution to lyophilized RNase-free DNase I (Worthington) to give a final concentration of 1 mg/ml. Heat 40 min at 55°C. Cool and add 1 M CaCl₂ to a final concentration of 5 mM. Store 0.3-ml aliquots at -20°C.

DNase I buffer

20 mM HEPES, pH 7.5
5 mM MgCl₂
1 mM CaCl₂
Sterilize by autoclaving

Elution buffer

1% (w/v) SDS
10 mM Tris-Cl, pH 7.5
5 mM EDTA
Sterilize by autoclaving

Glycerol storage buffer

50 mM Tris-Cl, pH 8.3
40% (v/v) glycerol
5 mM MgCl₂
0.1 mM EDTA

HSB buffer

0.5 M NaCl
50 mM MgCl₂
2 mM CaCl₂
10 mM Tris-Cl, pH 7.4
Sterilize by autoclaving

Lysis buffer

10 mM Tris-Cl, pH 7.4
3 mM CaCl₂
2 mM MgCl₂
Sterilize by autoclaving

Nonidet P-40 (NP-40) lysis buffers

Buffer A (basic protocol)

10 mM Tris·Cl, pH 7.4
10 mM NaCl
3 mM MgCl₂
0.5% (v/v) NP-40

Buffer B (alternate protocol)

10 mM Tris·Cl, pH 7.4
3 mM CaCl₂
2 mM MgCl₂
1% (v/v) NP-40

Autoclave first three components and cool before adding NP-40.

Reaction buffer, 2×

10 mM Tris·Cl, pH 8.0
5 mM MgCl₂
0.3 M KCl
Sterilize by autoclaving

Reaction buffer with nucleotides, 2×

1 ml 2× reaction buffer (see recipe)
10 μl 100 mM ATP
10 μl 100 mM CTP
10 μl 100 mM GTP
5 μl 1 M DTT
Prepare immediately prior to use

Separate 100 mM solutions of each nucleotide should be prepared in 0.5 M EDTA (pH 8.0) and the pH of each one checked to be sure it is between 7.0 and 8.0. The solutions should be stored in aliquots at -20°C.

SDS/Tris buffer

5% (w/v) SDS
0.5 M Tris·Cl, pH 7.4
0.125 M EDTA
Sterilize by autoclaving

Sucrose buffer I

0.32 M sucrose
3 mM CaCl₂
2 mM magnesium acetate
0.1 mM EDTA
10 mM Tris·Cl, pH 8.0
1 mM DTT
0.5% (v/v) Nonidet P-40 (NP-40)

Prepare without DTT and NP-40. Autoclave and cool to room temperature. Add DTT and NP-40 just prior to use.

Sucrose buffer II

2 M sucrose
5 mM magnesium acetate
0.1 mM EDTA
10 mM Tris·Cl, pH 8.0
1 mM DTT

Prepare without DTT. Autoclave buffer and cool to room temperature. Add DTT just prior to use.

TES solution

10 mM *N*-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid (TES), pH 7.4
10 mM EDTA
0.2% (w/v) SDS
Sterilize by autoclaving

TES/NaCl solution

10 mM *N*-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid (TES), pH 7.4
10 mM EDTA
0.2% (w/v) SDS
0.6 M NaCl
Sterilize by autoclaving

COMMENTARY

Background Information

The nuclear runoff transcription assay allows direct measurement and comparison of specific gene transcription in cells in various states of growth or differentiation. It takes advantage of the fact that newly synthesized RNA can be labeled to high specific activity in isolated nuclei, something that is difficult to accomplish in intact cells. The protocol described here has the advantage that it facilitates measurement of the level of transcription for many different genes in a single experiment. A point of controversy is whether some initiation of new RNA synthesis occurs in isolated nuclei during the runoff transcription reaction. What is clear is that transcripts that have initiated prior to cell lysis are faithfully elongated. Elongation of previously initiated transcripts most likely accounts for the bulk of the radioactivity incorporated into RNA in these reactions. Therefore, the method gives a reasonably accurate measure of the level of transcription occurring at the time of cell lysis. The runoff transcription assay is often used to assess whether changes in mRNA levels of a particular gene that occur as a function of cell state reflect a change in its synthesis as opposed to a change in mRNA degradation or transport from the nucleus to the cytoplasm.

A variety of different nuclear runoff transcription protocols have been described. The procedures differ primarily in the method of isolation of ³²P-labeled RNA. The protocol described here is advantageous because very low levels of background transcription are obtained. This is attributed in part to the TCA precipitation step, which allows effective removal of unincorporated ³²P. Limited digestion of ³²P-labeled RNA with NaOH appears to facilitate hybridization. Excellent reviews of

the nuclear runoff transcription method have been published by Marzluff (1978) and by Marzluff and Huang (1985).

Critical Parameters

The most critical step in the nuclear runoff assay is isolation of nuclei. For many different types of tissue culture cells, the procedure described in the basic protocol works well. However, the isolation protocol may have to be altered somewhat for isolating nuclei from tissues and lymphocytes (Marzluff and Huang, 1985). Poor incorporation of ³²P-labeled UTP into RNA in isolated nuclei may reflect damage to the nuclei during isolation or failure to isolate nuclei free of cytoplasmic and membrane contaminants. Lymphocyte nuclei are quite fragile, and it is necessary to employ an alternative isolation procedure to obtain nuclei that incorporate significant levels of radioactive isotope. In this alternative protocol, cells are lysed by Dounce homogenization in an isoosmotic buffer with nonionic detergent and nuclei are collected by centrifugation through a sucrose cushion. A minimum of 5×10^6 nuclei is required for a successful assay, but 5×10^7 is recommended. With fewer nuclei the level of incorporation of ³²P-labeled UTP into RNA drops significantly, and the level of specific transcription is usually difficult to distinguish from background radioactivity.

In this protocol the cpm/ml of ³²P-labeled RNA is equalized in each sample prior to hybridization based on the assumption that the overall level of RNA synthesis is not changing as a function of the cell state. When analyzing gene transcription under a new set of conditions, it is critical to determine if this is a reasonable assumption. Particular care must be taken in these experiments to ensure that the

same number of nuclei is used for each cell state analyzed.

Incorporation of ^{32}P -labeled UTP into RNA appears to increase when the number of nuclei is increased between 5 and 50×10^6 nuclei. Increasing the amount of ^{32}P -labeled UTP beyond 100 μCi per sample appears to be less useful. Increasing the incubation time of nuclei with ^{32}P -labeled UTP beyond 30 min is also ineffective for increasing the amount of ^{32}P incorporated into RNA.

Anticipated Results

Beginning with 5×10^6 nuclei and 100 μCi of ^{32}P -labeled UTP, the nuclear runoff transcription protocol allows incorporation of 1–10 $\times 10^6$ cpm into total RNA. Good results have been obtained when as few as 1×10^6 cpm were used in the hybridization reaction.

Time Considerations

Nuclei to be tested are isolated and stored in liquid nitrogen. NP-40 lysis and Dounce homogenization take <1 hr. Isoosmotic lysis and sucrose gradient centrifugation require ~2 hr. Ten to 20 samples can be processed at once through transcription. Labeling and isolation of RNA are accomplished in 1 day with an overnight RNA precipitation step. Preparation of nitrocellulose filters requires ~2 hr and can be done in advance. Hybridization of filters can be set up in a morning and requires a total of 2 days plus autoradiographic exposure.

Literature Cited

- Marzluff, W.F. 1978. Transcription of RNA in isolated nuclei. *Methods Cell Biol.* 19:317-331.
- Marzluff, W.F. and Huang, R.C.C. 1985. Transcription and Translation: A Practical Approach (B.D. Hames and S.J. Higgins, eds.) pp. 89-129. IRL Press, Oxford.

Key Reference

- Greenberg, M.E. and Ziff, E.B. 1984. Stimulation of 3T3 cells induces transcription of the *c-fos* proto-oncogene. *Nature* 311:433-438.

Describes the use of the nuclear runoff technique to measure the transcription rates of numerous genes as a function of a specific change in the cellular environment.

- Groudine, M., Peretz, M., and Weintraub, H. 1981. Transcriptional regulation of hemoglobin switching on chicken embryos. *Mol. Cell. Biol.* 1:281-288.

Describes the nuclear runoff transcription assay upon which the basic protocol is based.

Marzluff and Huang, 1985. See above.

Contains a great deal of useful information on the nuclear runoff transcription assay and describes in detail problems encountered in isolating the nuclei.

Contributed by Michael E. Greenberg
Harvard Medical School
Boston, Massachusetts

Timothy P. Bender (isolation of nuclei)
University of Virginia
Charlottesville, Virginia

CHAPTER 5

Construction of Recombinant DNA Libraries

INTRODUCTION

Construction of recombinant DNA molecules by simply ligating vector DNA and a fragment of interest is a straightforward process, discussed in Chapters 1 and 3. Special problems arise, however, when the fragment of interest represents only a very small fraction of the total target DNA. This is the case in two commonly encountered situations: isolation of single copy genes from a complex genome and isolation of rare cDNA clones derived from a complex mRNA population. This chapter describes techniques to generate recombinant DNA libraries which contain complete representation of genomic or cDNA sequences. The major differences between genomic and cDNA libraries are discussed in Section I (Overview of Recombinant DNA libraries). Insert DNA molecules are a necessary prerequisite for the construction of any DNA library and several approaches for producing insert DNA are discussed in Section II (Preparation of Insert DNA from Genomic DNA) and Section III (Preparation of Insert DNA from Messenger RNA). Finally, approaches to producing complete libraries are described in Section IV (Production of Genomic DNA and cDNA Libraries). The production of large libraries is a complicated process and all of the necessary steps are described here. However, a number of library-making kits are now commercially available which make this process much more feasible. These kits have the additional advantage that they utilize different cloning vectors that include proprietary features, which facilitate the use of library clones.

The central issue in library production is best presented in the context of considering the production of a genomic or cDNA library from a mammalian cell. The DNA of higher organisms is remarkably complex: a mammalian haploid genome contains approximately 3×10^9 base pairs. A particular 3000-bp fragment of interest thus comprises only 1 part in 10^6 of a preparation of genomic DNA. Similarly, a particularly rare mRNA species may comprise only 1 part in 10^5 or 10^6 of total poly(A) containing RNA, a ratio that is usually unaffected by the process of copying the RNA into cDNA. Clearly, the main problem in generating a useful recombinant DNA library from either genomic DNA or cDNA is the creation of the huge population of clones necessary to ensure that the library contains at least one version of every sequence of interest. The solutions to this problem are basically similar for genomic and cDNA libraries. As diagrammed in Fig. 5.0.1, the genomic DNA or cDNA are first prepared for insertion into the chosen vector. The vector and target DNA are then ligated together and introduced into *E. coli* by either packaging into phage λ heads in vitro or by direct transformation. In some aspects, however, strategies for isolation of individual genomic or cDNA clones can be quite different. The particular problems of creating these two different types of libraries will be discussed in detail separately (UNITS 5.1 & 5.2). Over the past several years, techniques for producing improved cDNA libraries have been optimized. These techniques will be described in the near future in an update to UNIT 5.8A.

Most cDNA library screening procedures involve positive identification of cDNA clones with either antibodies or hybridizing nucleic acid probes. Subtracted cDNA libraries provide a method for identifying mRNAs (as cDNAs) that are expressed in one cell but not another. cDNAs are synthesized from the cell expressing the desired mRNA and all of the sequences expressed in a cell not expressing the RNA are removed by hybridization and selection. The remaining sequences are cloned into a bacteriophage or plasmid vector

Contributed by J.G. Seidman

Current Protocols in Molecular Biology (2003) 5.0.1-5.0.3

Copyright © 2004 by John Wiley & Sons, Inc.

Construction of
Recombinant
DNA Libraries

5.0.1

Supplement 68

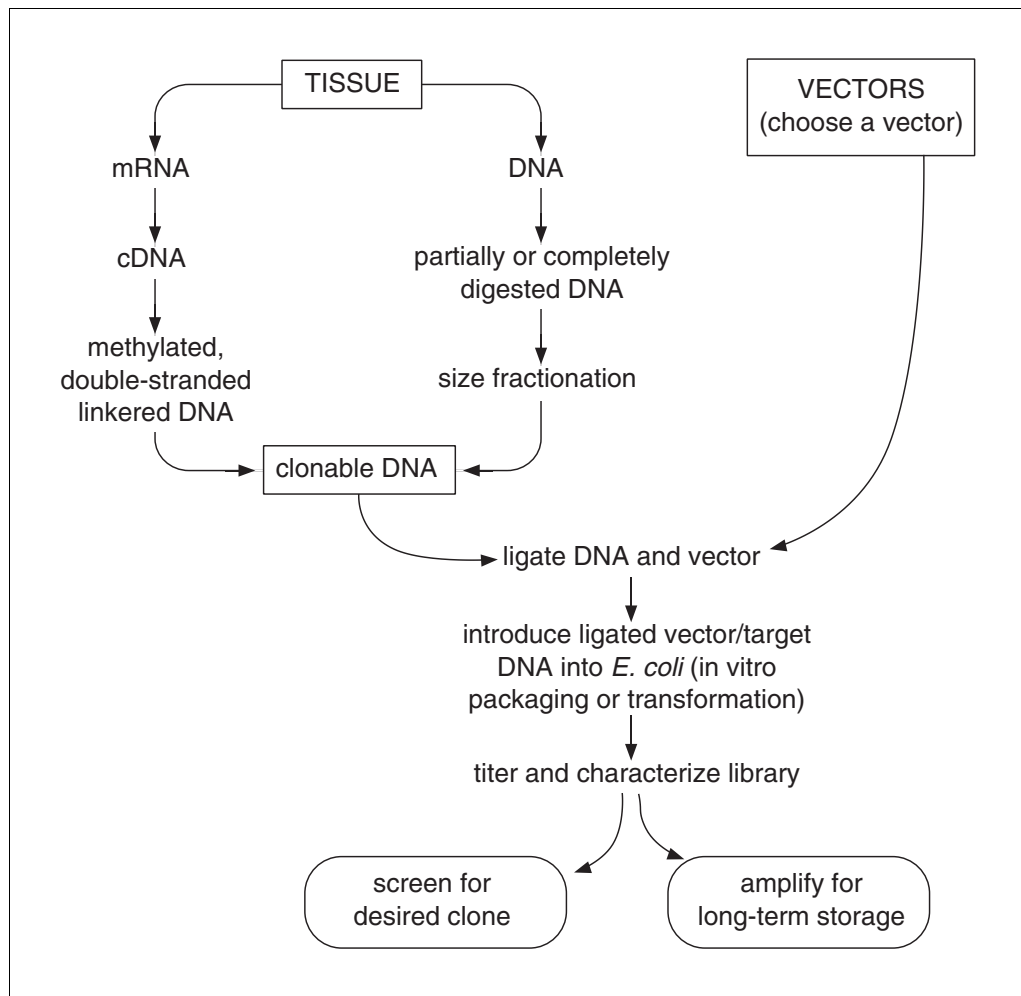


Figure 5.0.1 Steps involved in the construction of cDNA or genomic DNA libraries.

to produce a subtracted cDNA library. Estimation of the number of clones that must be screened is difficult, as it will vary with the cell type and the gene to be identified. Some cDNA libraries such as SAGE libraries and subtracted libraries do not capture full-length or near-full-length cDNAs. Construction of these libraries is described in Chapter 25, where they are used for assessing RNA expression.

The *E. coli* vectors described in this chapter are limited with regard to the size of insert DNA that can be accommodated (~20 kb for lambda and ~40 kb for cosmid vectors). The ability to clone much larger fragments of DNA, however, has become an essential requirement for many genome analysis projects. Yeast artificial chromosome (YAC) vectors, maintained in yeast hosts, typically carry inserts ranging from 0.3 to 1.2 Mb of genomic DNA. Both the size and complexity of YAC libraries pose special considerations for production, screening, and analysis, and these concerns are addressed in *UNITS 6.9 & 6.10* in the following chapter.

Two important general points pertain to both genomic and cDNA libraries. First, it is essential that both the vector DNA and target DNA used to create the library are not contaminated by exogenous sequences detectable by the probes that will be used to isolate the clones of interest. There are obvious, potentially disastrous effects of contaminated target DNA—for example, by only 1 part in 10^5 of a plasmid containing the cDNA sequences to be used as a probe. See “Going for the gene,” *The Boston Globe Magazine*,

Aug. 2, 1987, for an account of such a mistake. Common sense dictates care and use of absolutely clean and, where possible, disposable materials throughout.

Second, libraries can be stored indefinitely. Over the past 5 to 10 years, large numbers of genomic libraries have been made from a variety of organisms and an even larger number of cDNA libraries have been produced from an astonishing number of tissues and cell lines. Many useful libraries, including examples of human and other mammalian genomic or cDNA libraries, have been made over the years. In some cases, journals (e.g., *Cell*, *Science*, *Proceedings of the National Academy of Sciences U.S.A.*, and the publications of the American Society for Microbiology) require that libraries and individual clones discussed in their pages be freely available to other investigators. Both stock and custom-made libraries are also available from a variety of commercial sources. The ability to amplify these libraries without significantly altering the distribution of clones within the library allows investigators to utilize a library many times. Protocols for amplifying libraries are described in *UNITS 5.10 & 5.11*.

J.G. Seidman

OVERVIEW OF RECOMBINANT DNA LIBRARIES

The units in this section present strategies for generating genomic DNA and cDNA libraries (UNITS 5.1 & 5.2, respectively). Production of recombinant DNA libraries can be a very laborious procedure. We strongly recommend the purchase of genomic or cDNA libraries when possible; otherwise, we recommend purchasing reagent kits for producing libraries. These kits can save considerable time and effort.

Genomic DNA Libraries

Genomic DNA libraries are almost always screened by hybridization using a radioactive nucleic acid probe. Since this approach is essentially independent of a particular vector or type of target DNA, the main problem faced when considering creation of a genomic DNA library is simply generating a large enough number of recombinant DNA clones. The basic strategies used to address this problem have included both minimizing the number of clones necessary by incorporating large fragments of genomic DNA, and maximizing cloning efficiency by using vectors based on bacteriophage λ . This unit will discuss the appropriate numerical considerations for both ordinary genomic DNA libraries and subgenomic DNA libraries, and will then describe a limited number of appropriate vectors.

REPRESENTATION AND RANDOMNESS

The size of a library of completely random fragments of genomic DNA necessary to ensure representation of a particular sequence of interest is dictated by the size of the cloned fragments and the size of the genome. The likelihood that a sequence of interest is present in such a random library can be estimated by simple statistics based on the Poisson distribution (Clarke and Carbon, 1976). Specifically, the number of independent clones, N , that must be screened to isolate a particular sequence with probability P is given by

$$N = \ln(1 - P) / \ln[1 - (I/G)]$$

where I is the size of the average cloned fragment, in base pairs, and G is the size of the target genome, in base pairs.

For a 99% chance of isolating an individual sequence from a typical mammalian genome using a typical phage λ vector

$$N = \ln(1 - 0.99) / \ln[1 - (2 \times 10^4 / 3 \times 10^9)] = 690,000.$$

This equation can be used to define a useful rule-of-thumb by calculating the probability of isolating a fragment of interest as a function of (I/G). In general, to have a 99% chance of isolating a desired sequence, the number of clones screened should be such that the total number of base pairs present in the clones screened ($I \times N$) represents a 4.6-fold excess over the total number of base pairs in the genome (G) (Seed et al., 1982).

When the desired fragment can be purified, the size of the library can be reduced. The library size can then be estimated by

$$N \cong 3 \times 1/p$$

where p = the probability of isolating a particular fragment = 1/total number of fragments in the pool.

It is important to note that this simple analysis assumes that the cloned DNA segments randomly represent the sequences present in the genome. This assumption is true only if the target DNA was cleaved completely randomly prior to insertion into the vector. In the strictest sense, this level of randomness can be approached only by the relatively inconvenient means of shearing the target DNA.

With common sense and care, however, sufficiently random cleavage of target DNA can generally be obtained using partial digestions with restriction enzymes (UNIT 5.3). One simple limitation of this approach is that fragments which are larger than the capacity of the vector as complete digestion products will be excluded from the library. Clearly, it is best to use an enzyme that cuts the DNA of interest both frequently and without any bias in selection of one site over another (such bias is seen with *EcoRI*, for example). The enzyme *Sau3A*, which recognizes the 4-bp site GATC and generates fragments compatible with several convenient phage λ and cosmid vectors (see below), has proved useful for generating partial digestion libraries.

Given an enzyme that cleaves frequently and randomly, it is not so obvious which partial digestion protocol will lead to the most random products. As described in detail by Seed et al. (1982), randomness is maximized by partial digestion to an extent where the size of the numerically most abundant class of partial products equals the vector capacity. This is not reflected by an extent of digestion in which the location on a gel of the maximum staining intensity of partial digestion products equals the fragment size of interest (a consequence of the fact that larger fragments stain more intensely than an equal number of smaller fragments). To determine the optimal extent of digestion, resolve a series of partial digestions of increasing extent on a gel and examine the amount of staining in *only* the size class of interest. The lane in which the greatest amount of DNA is seen corresponds to *twice* the appropriate extent of digestion (Seed et al., 1982).

SUBGENOMIC DNA LIBRARIES

Sometimes only a small and relatively well characterized fragment is desired. For example, if a particular 1-kb *Bam*H1 fragment is of interest, it can be purified and used to generate a smaller, potentially easier to screen library. Such libraries, which represent only a fraction of the genome, are called subgenomic DNA libraries.

Numerical considerations show that maximizing the fold of purification of target DNA is crucial for subgenomic DNA libraries. One can use the equation described above to estimate the number of clones necessary by simply assuming that the genome size is reduced by the amount of purification. For example, if the desired 1-kb mammalian DNA fragment was purified 10-fold from the rest of the genomic DNA, and the resultant library was otherwise a random representation of the remaining 10% of the genome, then

$$N = \ln(1 - 0.99) / \ln[1 - (1 \times 10^3 / 3 \times 10^8)] = 1,382,000.$$

In this case the subgenomic approach has actually *increased* the number of clones necessary as compared to the random library, due to the large decrease in the size of the insert. Increasing purification by another factor of 10 decreases the number of clones needed by a factor of 10. As a minimum, the fold of purification must exceed the ratio of genomic DNA library insert size to subgenomic insert size.

A simple way to increase the fold of purification is to use multiple, sequential digestion

strategies in cases where details of the restriction map of the sequences of interest are known. After initial purification of a given fragment, redigestion with another enzyme that gives a smaller (clonable) fragment will generally yield significant further purification relative to the original DNA.

VECTORS FOR GENOMIC DNA LIBRARIES

Because of their combination of high cloning efficiency and relatively large insert size, either bacteriophage λ vectors or hybrid plasmid vectors called cosmids (which contain particular λ sequences that direct insertion of DNA into phage particles) are generally used to construct genomic DNA libraries. The biology and general properties of these two types of vectors are described in *UNIT 1.10*.

Briefly, the high cloning efficiency of both types of vectors is a consequence of the ability of simple extracts of phage λ -infected cells to insert exogenously added λ DNA, or recombinant DNA containing appropriate λ sequences, into preformed λ heads and tails, generating infectious phage particles. Up to 10% of added concatameric phage genomes can be packaged in this way, an efficiency significantly greater than that of introduced plasmid DNA into *E. coli* by transformation. In the case of cosmid vectors, the recombinant DNA inserted into the phage contains plasmid vector sequences and replicates as a plasmid after infecting bacteria (see *UNIT 1.10*).

The cloning capacity of both types of vectors is dependent on the size of DNA that can be accommodated by λ phage heads, approximately 35 to 50 kb. However, the vectors differ significantly in the fraction of this total taken up by vector sequences. Most phage vectors designed for genomic DNA libraries can accommodate foreign DNA fragments of 10 to 20 kb generated by a limited variety of restriction enzymes. Cosmids can generally accept 30- to 40-kb fragments generated by any of a number of restriction enzymes.

The choice between phage and cosmid vectors is generally based on the size of the desired genomic DNA segment. Most investigators feel that phage libraries are easier to handle, and choose a phage vector if the desired segment is less than ~20 kb. Larger segments require the use of cosmid vectors.

Bacteriophage λ Vectors

Significant design advances over the years have resulted in the development of several

easy-to-use phage λ vectors. These vectors have two basic features in common: ability to accept fragments generated by several restriction enzymes, and biochemical and/or genetic selection against the so-called stuffer sequences present in the original vector in the place of the exogenously added DNA. As described in *UNIT 1.10*, this stuffer fragment is necessary because vectors that contain the minimum segment of the λ genome necessary for phage propagation (approximately 30 kb) are too small to be packaged into normal phage heads. Earlier vectors required rather laborious biochemical separations of vector and stuffer fragments prior to insertion of foreign DNA.

The minimal λ genome contains restriction sites for a number of enzymes frequently used for cloning in plasmid vectors. Newer vectors have partially circumvented this problem by eliminating some of these sites (notably those for *Bam*HI) and addition of new, unique sites to polylinkers which flank the stuffer fragment.

The segment of the λ genome that can be replaced by exogenous DNA contains genes whose expression prevents phage growth in bacterial hosts containing P2 prophages (see *UNIT 1.10*). Engineering the stuffer fragment to express these genes regardless of orientation relative to the rest of the genome has resulted in vectors that will grow on such P2 lysogens only if they have incorporated foreign DNA in the place of the stuffer. Such vectors obviate the need to physically remove the stuffer prior to ligation to exogenous DNA.

An alternative biochemical strategy to prevent reinsertion of the stuffer segment has also been developed. In vectors designed for this approach, two identical polylinkers flank the stuffer in an inverted orientation relative to each other. In this arrangement, double digestion of the polylinkers with two appropriate enzymes generates vector and stuffer fragments with heterologous, nonligatable ends. Simple, preferential ethanol precipitations (see *UNIT 2.1*) remove the very small polylinker fragments. Since the vector segment by itself is too short to generate a viable phage, all plaques generated by in vitro packaging of a ligation of vector and insert DNA should be recombinant.

The vector λ EMBL3 allows both the genetic and the biochemical strategies to avoid purification of stuffer fragments, and includes several useful cloning sites in the polylinker.

This versatile and modern vector has been successfully used to create many libraries.

Cosmid Vectors

Any plasmid cloning vector that contains the λ *cos* site can be used as a cosmid. A number of cosmid vectors designed for particular applications include additional elements such as genes that allow transfer to non-*E. coli* bacterial cells or dominant markers for selection in mammalian cells. Such add-ons decrease the cloning capacity of the vector and should be avoided if possible. One useful, simple cosmid vector is pJB8 (Ish-Horowitz and Burke, 1981), a 5.4-kb plasmid that accepts genomic DNA digested with *Sau*3A and can be used with several cosmid cloning strategies.

Vectors for Subgenomic DNA Libraries

It is possible to use simple plasmid vectors for subgenomic DNA libraries if the level of purification and recovery of the target fragment is sufficient to overcome the relative inefficiency inherent in plasmid cloning. In general, however, phage λ vectors designed for direct insertion of foreign DNA rather than substitution for a stuffer fragment are used. A very large number of potential insertion vectors exist to accommodate fragments generated by a variety of restriction fragments: wild-type λ is a natural insertion vector which should accommodate *Sal*I or *Xho*I fragments up to 2 to 3 kb. λ gt10 (see *UNIT 1.10*) is the only vector in general use which allows selection against nonrecombinant phages, and is recommended for cloning *Eco*RI fragments.

LITERATURE CITED

- Clark, L. and Carbon, J. 1976. A colony bank containing synthetic *ColE1* hybrids representative of the entire *E. coli* genome. *Cell* 9:91-99.
- Ish-Horowitz, D. and Burke, J.F. 1981. Rapid and efficient cosmid vector cloning. *Nucl. Acids Res.* 9:2989-2999.
- Seed, B., Parker, R.C., and Davidson, N. 1982. Representation of DNA sequences in recombinant DNA libraries prepared by restriction enzyme partial digestion. *Gene* 19:201-209.

Contributed by David D. Moore
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

The most basic step in constructing a cDNA library is the process of generating a double-stranded DNA copy of the mRNA. In the last few years, preparation of this cDNA has been simplified by improved strategies and availability of higher quality enzymes. Thus, it should be straightforward to obtain essentially full-length cDNA copies for mRNAs up to the 3 to 4 kb range, and at least feasible for even larger mRNAs. As described in detail in the cDNA protocols (UNITS 5.5 & 5.6), the most important factor affecting quality of cDNA is the quality of the mRNA. Particularly for a large message, it is essential to start with the highest quality RNA available.

Two related issues dominate the strategies for constructing cDNA libraries. The first is the relative abundance of the clone of interest, which can vary over a wide range. Highly abundant messages can represent 10% or more the total mRNA, whereas very rare messages can be as low as 1 part in 10^6 , particularly if the gene of interest is only being expressed in a fraction of the cells used as a source of mRNA. The second issue is the screening method (see Chapter 6), which can range from simply sequencing several individual isolates until the desired clone is identified, through ordinary hybridization methods, to complex strategies involving expression of identifiable antigens or biological activities.

Obviously, the size of the library necessary to include the clone of interest is a direct reflection of the relative abundance of the mRNA of interest. In general, however, this abundance is not known with precision. In addition, the representation of some sequences in the cDNA library, particularly the 5' ends of large messages, will be less than expected from their mRNA abundance. It is sensible to aim for a library that contains at least 5 times more recombinants than the total indicated by the lowest abundance estimate. In some cases this number should be multiplied by various factors based on screening efficiency. If it is necessary to fuse a peptide-coding region to a vector in a particular reading frame, for example, the number of identifiable clones is only $\frac{1}{6}$ of those present in the library.

If the mRNA of interest is relatively abundant, efficiency of generating clones is not so important, and the choice of cloning strategy and vector should be based on the desired use

for the clone. If, for example, expression in *E. coli* is the object, the cDNA library can be inserted directly into an appropriate expression vector. This might involve choosing linkers or adaptors useful for insertion into the vector, and simple screening by hybridization.

In many cases, however, the mRNA of interest is relatively rare, and high cloning efficiency is of central importance. As with genomic libraries, this has led to development and use of phage λ vectors. In general, there are two types of λ vectors for cDNA cloning adapted for the two most common methods of library screening.

If the library is to be screened by hybridization with a nucleic acid probe (UNITS 6.1 & 6.3 or 6.4), any insertion vector is appropriate. A vector that is particularly good for this approach is λ gt10. As mentioned above, this insertion vector allows direct selection against nonrecombinant phages. A useful feature of this vector for cDNA cloning is that it accepts *EcoRI* inserts. Methylation of the double-stranded cDNA with *EcoRI* methylase and addition of *EcoRI* linkers is an efficient way to generate clonable cDNA (see UNITS 5.5 & 5.6).

If the library is to be screened by use of antibody probes (UNIT 6.7), it is necessary to use an appropriate *E. coli* expression vector. In general, such vectors are based on expression of a fusion protein in which a segment of the peptide of interest is fused to a highly expressed, stable *E. coli* protein. The most commonly used expression vector is λ gt11 (UNIT 1.10), in which the cloned peptide coding sequences are fused to coding sequences for β -galactosidase.

KEY REFERENCES

- Frischauf, A.-M., Lehrach, H., Poustka, A., and Murray, N. 1983. Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* 170:827.
- Huynh, T., Young, R., and Davis, R. 1984. Construction and screening cDNA libraries in λ gt10 and λ gt11. *In* DNA Cloning, Vol. 1: A Practical Approach (D. Glover, ed.) pp. 49-78. IRL Press, Oxford.

Contributed by David D. Moore
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

PREPARATION OF INSERT DNA FROM GENOMIC DNA

Completely or partially digested genomic DNA must be size fractionated before ligation to vector to remove irrelevant small and large fragments. If size fractionation is not performed, small fragments will ligate together and produce recombinants that are difficult to analyze. Large insert DNA fragments will not allow the vector to grow, but will ligate to vector DNA and alter vector DNA requirements. Size-purified fragments isolated from complete digests of genomic DNA are less complex than the entire genome and thus reduce the number of DNA clones that must be produced in order to obtain the desired subgenomic library.

Procedures for making insert fragments for genomic and subgenomic DNA libraries involve digestion of DNA followed by size fractionation. DNA is either partially digested for preparation of complete genomic libraries or completely digested for preparation of subgenomic libraries. Basic protocols for digesting DNA with restriction enzymes are presented in *UNIT 3.1*; however, *UNIT 5.3* presents support protocols with necessary modifications to ensure that large amounts of genomic DNA are properly digested.

Two methods for size fractionation of genomic DNA are presented. Both protocols are appropriate for the isolation of DNA fragments that will subsequently be used for library construction. These protocols aim to maximize DNA fragment yield while minimizing exposure of the DNA to reagents or conditions that inhibit subsequent ligation to vector and introduction of hybrid molecules into the bacterial cell. The sucrose gradient and preparative gel electrophoresis methods circumvent problems found in other fractionation protocols.

Sucrose gradient fractionation is generally faster than preparative gel electrophoresis. However, the latter procedure has a higher capacity for resolving large amounts of DNA, is applicable to a larger range of sizes, and has significantly better resolution.

These procedures emphasize the requirements for fractionating large quantities of DNA and producing DNA that will ligate to vector. Normally, large quantities of genomic DNA can be obtained from the species of interest so that when producing genomic libraries the investigator has the luxury of being able to work with more DNA than when producing cDNA libraries. The procedures used here all assume that genomic DNA is available (see *UNITS 2.1 & 2.4* for DNA prep procedures). The amount of DNA required varies depending on the complexity of the genome being used.

NOTE: For laboratories using recombinant DNA techniques and isolating large quantities of plasmid, bacteriophage, or cosmid DNA, remember that the smallest amount of contamination of genomic DNA with recombinant DNA is disastrous. Contamination of genomic DNA at 1 ppm with recombinant plasmid or bacteriophage will cause great difficulty because they may grow and be identified during screening procedures as the desired clones. Thus, all plasticware, glassware, and reagents used for the preparation of genomic DNA or mRNA and cDNA should be maintained separately from those used for plasmid or bacteriophage DNA preparation. The extensive use of disposable plasticware is strongly recommended.

Size Fractionation Using Sucrose Gradients

BASIC PROTOCOL

SUCROSE GRADIENT PREPARATION OF SIZE-SELECTED DNA

Partially or fully digested DNA consists of a population of DNA fragments ranging in size from hundreds of base pairs to over 100,000 bp in length. This protocol effectively separates such a mixture of DNA fragments into different size classes. To accomplish this the DNA solution is heated to dissociate aggregated DNA fragments and is then loaded onto a high-salt sucrose gradient. After centrifugation and gradient fractionation, the appropriate fractions are identified by agarose gel electrophoresis. This protocol can also be used to purify bacteriophage λ vector arms.

Materials

Completely or partially digested genomic DNA (support protocols)

STE buffer (APPENDIX 2)

10% and 40% sucrose solution

0.9% agarose gel

100% ethanol

TE buffer (APPENDIX 2)

Sucrose gradient maker

Beckman SW-28 or SW-41 rotor or equivalent

Additional reagents and equipment for ethanol precipitation (UNIT 2.1) and agarose gel electrophoresis (UNIT 2.5)

1. Begin with partially or fully restriction-enzyme digested DNA (see support protocol) at a concentration of about 1 mg/ml in STE buffer.

It is essential that the DNA be completely dissolved.

2. Prepare a linear 10% to 40% sucrose gradient in an SW-28 centrifuge tube (38-ml gradient) or SW-41 tube (12-ml gradient).

There are a variety of gradient makers; follow the manufacturer's instructions for preparing the gradient.

3. While the gradient is being poured, heat the digested DNA to 65°C for 5 min to dissociate any DNA aggregates.

4. Carefully layer the DNA solution on top of the sucrose gradient.

Do not exceed 0.5 mg genomic DNA per SW-28 centrifuge tube or 0.2 mg genomic DNA per SW-41 tube. For λ vector DNA, do not exceed 50 μ g and 20 μ g, respectively (see critical parameters).

5. Centrifuge at 20°C, 113,000 \times g (25,000 rpm in SW-28 rotor) or 154,000 \times g (30,000 rpm in SW-41 rotor) for 16 to 24 hr, depending upon the size of the desired DNA fragments. For cosmid-sized inserts (40,000 bp) centrifuge 16 to 18 hr. For phage-sized inserts (18,000 bp) centrifuge 24 hr.

6. Fractionate the gradient by carefully placing a capillary tube at the bottom of the centrifuge tube and pumping out the gradient, heavier fractions first. This prevents the mixing of smaller DNA fragments with the desired larger fragments. Collect 750- μ l fractions into microcentrifuge tubes (see Fig. 5.3.1A).

Note: 12-ml gradients can also be fractionated by removing 750- μ l aliquots from the top with a mechanical pipettor (e.g., P1000 Eppendorf pipet).

7. Determine the size of the collected DNA fractions by electrophoresing 40- μ l samples

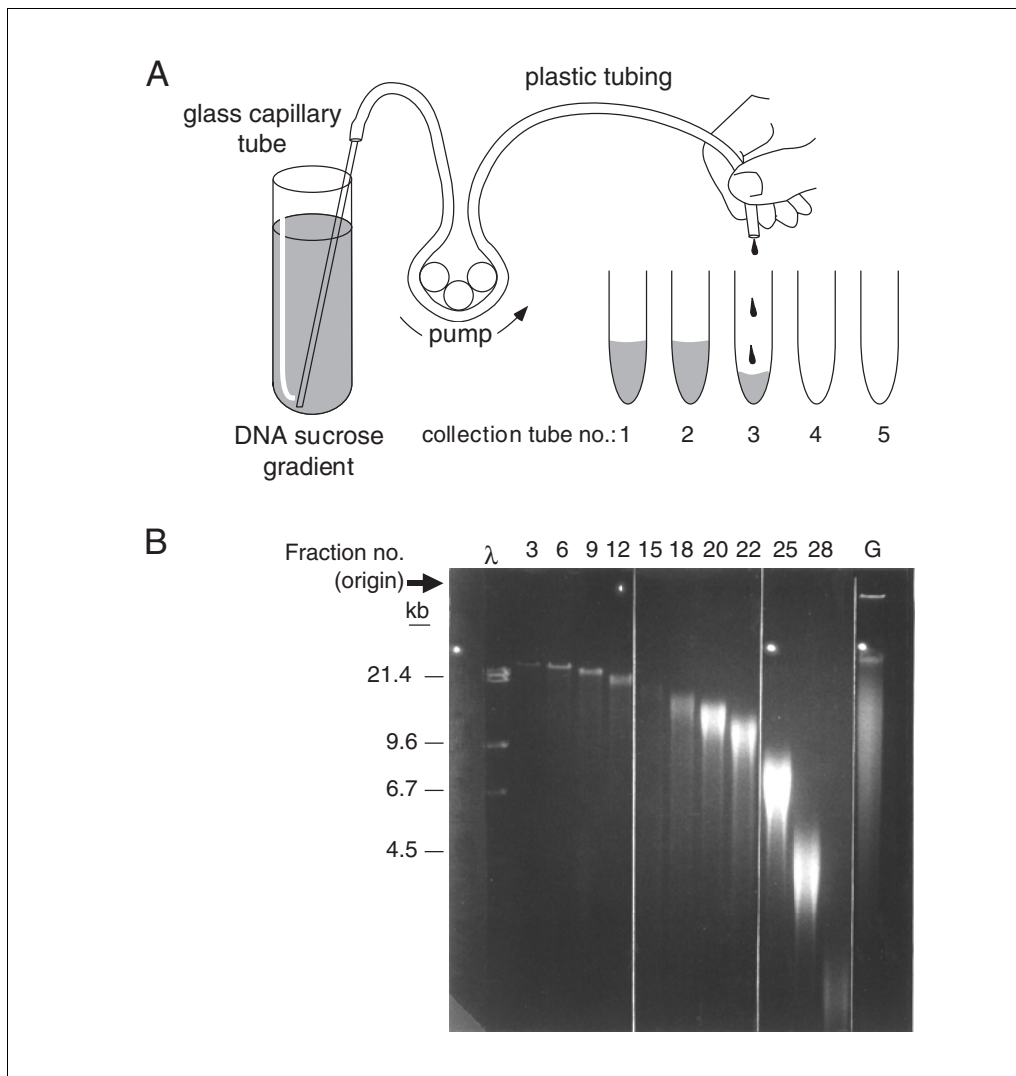


Figure 5.3.1 (A) Method for fractionating sucrose gradients. The gradient is pumped out of the tube from the bottom. The arrow indicates the flow of sucrose solution. The densest part of the gradient and the DNA of the highest molecular weight is pumped out first. The gradient is collected in 750- μ l aliquots. (B) Ethidium bromide-stained agarose gel used for characterization of sucrose gradient-fractionated genomic DNA, where G is the input genomic DNA.

of the gradient in a 0.9% agarose gel at high voltage (see *UNIT 2.5*).

8. Ethanol precipitate fractions containing correctly sized DNA. Divide each sucrose fraction into four microcentrifuge tubes, each containing \sim 180 μ l of sucrose/DNA solution. Add 320 μ l water and 1 ml ethanol to each tube and place at -20°C overnight.

For subgenomic libraries, appropriate fractions are identified by Southern blot analysis of the agarose gel using a DNA probe that will identify the desired fragment (see UNIT 2.9). For genomic libraries, the desired fractions are identified by the size of the DNA fragments. Cosmid vectors require inserts of \sim 30 to 50 kb, while bacteriophage vectors generally require inserts of 8 to 20 kb.

9. Resuspend and combine four ethanol-precipitated aliquots of DNA in a total volume of \sim 100 μ l TE buffer per sucrose gradient fraction.
10. Those fractions containing the desired size class can be used directly for ligations or can be ethanol precipitated again for long-term storage.

PARTIAL ENZYME DIGESTION

Genomic high-molecular-weight DNA is incubated with limiting amounts of a particular restriction enzyme for variable lengths of time. Samples from the digested DNA are removed at different time intervals and analyzed by agarose gel electrophoresis to determine the average length of the digested DNA. The time points of digestion that are most enriched for the desired size fractions are then used as the guide for the preparative digestion of the same DNA. In this way, size-selected DNA fragments can be prepared with a minimum of trial and error.

Additional Materials

High-molecular-weight genomic DNA (UNITS 2.2, 2.3, & 2.4)

10× restriction enzyme buffer (UNIT 3.1)

Stop solution

Restriction enzyme for DNA digestion (UNIT 3.1)

0.5% agarose gel (UNIT 2.5A)

DNA size markers (UNIT 2.5A)

1:1 phenol/chloroform (UNIT 2.1)

5 M NaCl

1. Start with genomic DNA suspended in STE or TE buffer at a concentration of ~0.1 to 0.5 mg/ml.

The size of the genomic DNA used for this procedure is critical to its success. The average size of the DNA must exceed 100 kb. If the DNA is smaller than 100 kb, very few partial digestion products will be obtained that have both ends produced by the restriction enzyme and thus be suitable for cloning into vector.

2. Transfer 100 µg DNA to a microcentrifuge tube, dilute to 900 µl, and add 100 µl of the appropriate 10× restriction enzyme buffer.

Sau3A is frequently used for mammalian genomic libraries and produces ends on the DNA that can ligate to BamHI cut vector (see commentary).

3. Incubate 10 min at 37°C.

This incubation prewarms the reagents to 37°C. Some investigators prefer to incubate at room temperature to slow down the reaction.

4. Remove 40 µl of solution, add to 10 µl stop solution, mix, and put on ice.

5. Add, in units, enough of the appropriate restriction enzyme to the remaining 960 µl of DNA solution to *fully* digest only 20% of the DNA in the tube in 60 min.

6. At 10-min intervals remove 40 µl digestion solution, add to 10 µl stop solution in a fresh microcentrifuge tube, and place on ice. Continue for 90 min.

There should be 9 separate aliquots of terminated reaction.

7. Prepare a 0.5% agarose gel and electrophorese the DNA aliquots using, as markers, uncut and *Hind*III-digested λ DNA (see UNIT 2.5).

The dye marker should be electrophoresed off the bottom of the gel because high resolution in the 20- to 50-kb region of the gel is required.

Handle 0.5% agarose gels carefully; they are extremely fragile.

8. Assess the amount of digestion. If the enzyme concentration is in the appropriate range, one should observe the high-molecular-weight genomic DNA being converted to lower-molecular-weight DNA with time. If there is very little high-molecular-weight material even in lanes representing the earliest time points, reduce the amount of restriction enzyme 10-fold and start over (step 1). If there is almost no digestion in the lane representing the longest time point, increase the amount of

enzyme 10-fold and start over (step 1). The desired time points are those in which a large amount (10 to 20%) of the DNA is in the 20- to 50-kb size range.

Most investigators prefer to choose earlier time points because they hope to find a better representation of partially digested fragments at these times than at later times. (See critical parameters for a further discussion of this subject.)

9. Once a correct time point and enzyme concentration has been determined, scale up digestion reaction to contain 1 mg DNA (10-ml reaction). Prewarm and digest.
10. Remove 3.3 ml of digest 5 min before the optimal time, add to 10 ml phenol/chloroform (1:1), and mix well. Remove 3.3 ml at the optimal time and 3.3 ml 5 min after the optimal time and add to phenol/chloroform mixture.

For unclear reasons, scaling up this reaction does not always work. If scale-up results in either over- or underdigestion, use 10 identical 1-ml reactions.

11. Centrifuge 5 min at 5000 to 10,000 \times g until the phases are separated and remove the supernatant.
12. Ethanol precipitate by adding 1 ml of 5 M NaCl and 30 ml 100% ethanol, chill, and centrifuge (see UNIT 2.1).
13. Resuspend DNA in 1 ml STE buffer and analyze the product on a 0.5% agarose gel (see step 8 this protocol, and UNIT 2.5).

Resuspend gently to avoid shearing the large DNA.

Assuming that a significant fraction (10 to 20%) of the DNA is in the 20- to 50-kb range it is suitable for library construction.

COMPLETE ENZYME DIGESTION

One approach to cloning a gene from a particular cell or tissue involves the production of a subgenomic library. A subgenomic library is constructed by completely digesting DNA with a particular restriction enzyme and then isolating the particular size class of DNA that contains the desired restriction fragment (see overview, UNIT 5.1).

0.1 to 1 mg of completely digested genomic DNA (see UNIT 3.1) is generally required to produce a subgenomic library. We suggest performing a trial digest of 1 to 10 μ g DNA

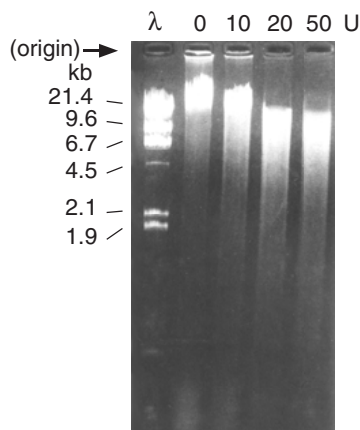


Figure 5.3.2 This 0.9% agarose minigel contains samples of genomic digests performed with increasing concentrations (left to right) of restriction enzyme. As the digest approaches completion, high-molecular-weight fragments become less prominent and faint satellite bands become visible. 10 μ g genomic DNA was incubated with 10, 20, and 50 U restriction enzyme for 2 hr, and 40 μ l (about 1 μ g) DNA was loaded on the gel.

SUPPORT PROTOCOL

Construction of Recombinant DNA Libraries

5.3.5

to determine the optimum enzyme concentration (Fig. 5.3.2), need for spermidine, etc., before attempting the large-scale digest. Overdigestion of the DNA is undesirable because the ends of the DNA fragments are sometimes destroyed by nuclease. It is desirable to limit the quantity of enzyme and length of incubation to prevent any nonspecific digestion and exonuclease activity which can interfere with ligation and cloning.

Occasionally, difficulty is encountered digesting genomic DNA. This is usually because of contamination of the DNA with proteins that inhibit the activity of the restriction enzyme. Also, even modest amounts of NaCl trapped in the DNA during ethanol precipitation can inhibit the activity of low-salt enzymes such as *SacI* and *KpnI*. If the DNA is resistant to digestion, it may be helpful to phenol extract, chloroform extract, and ethanol precipitate the DNA to remove contaminants.

When appropriate reaction conditions have been determined, scale up the digest (increase all reagent volumes proportionally) to include the entire genomic DNA sample to be used for cloning. Again, 0.1 to 1 mg mammalian genomic DNA should be sufficient for construction of a subgenomic library. Confirm completion of digestion by running ~1 µg digested DNA on a minigel. The completely digested DNA should be phenol/chloroform extracted and ethanol precipitated as described in *UNIT 2.1* to remove enzyme and salt.

REAGENTS AND SOLUTIONS

Stop solution

- 10 mM Tris·Cl, pH 7.5
- 20% glycerol
- 0.1% sodium dodecyl sulfate
- 0.1% bromphenol blue

10% sucrose solution

- 10% sucrose (10 g sucrose/100 ml solution)
- 1 M NaCl
- 20 mM Tris·Cl, pH 7.5
- 5 mM EDTA

40% sucrose solution

- Same as above except with 40% sucrose

COMMENTARY

Background Information

Sucrose gradients. DNA fragments migrate through a linear sucrose gradient at a rate that is dependent on their size. Other gradients (e.g., sodium chloride gradients) have also been proposed; however, we find that sucrose gradients are easier to prepare and provide equally good resolution. This procedure provides good resolution for DNA fragments 5 to 60 kb in size. Thus, partially digested genomic DNA can be fractionated for the production of cosmid or bacteriophage libraries and completely digested DNA can be fractionated for subgenomic DNA libraries. Sucrose gradients are also useful for purification of bacteriophage λ vector arms.

This is an extremely reliable, easy, and rapid method for fractionating digested genomic DNA. The technology has been used for many years. DNA fragments obtained by this procedure can normally be ligated to vector. The major weakness of the procedure is that it does not produce the resolution of some preparative agarose gel techniques and it does not have the capacity of those techniques (see *UNIT 5.4*). Despite these weaknesses this is probably the most widely used method for producing inserts for genomic and subgenomic libraries. Sucrose gradients can also be used for purification of bacteriophage λ vector arms.

Partial digestion. The preferred method

for producing the random collection of DNA fragments required to make a genomic library is to partially digest high-molecular-weight genomic DNA with a restriction enzyme. An extensive discussion of the theory of partial restriction enzyme digestion is provided in *UNIT 3.1*. Partial digestion can be achieved by limiting the reaction time or amount of restriction enzyme. *UNIT 3.1* presents a protocol in which the amount of enzyme is limited; the protocol described here involves limiting the time of digestion. Regardless of the method of creating partial digestion products, the products created earlier in the reaction are probably more random than the products produced later in the reaction (Seed et al., 1982).

Restriction enzymes that recognize 4-base sequences produce a more random collection of insert fragments than enzymes that recognize 6-base sequences. The most random collection of fragments is produced by shearing the DNA. However, because the techniques for cloning sheared DNA are not very efficient most investigators produce the random collection of DNA fragments by partial digestion with restriction enzymes. The theory of using restriction enzymes to produce inserts for genomic libraries is discussed by Seed et al. (1982). Restriction enzymes that produce the most random collection of fragments are those that cleave DNA most frequently.

A second reason for choosing a particular restriction enzyme is that it must produce fragments that can be ligated into the desired vector. Because they produce fragments that can be ligated into *Bam*HI-cut vectors and because they recognize a 4-base sequence (GATC), the enzymes *Sau*3A or *Mbo*I are normally used for this purpose.

Complete digestion. See second support protocol.

Critical Parameters

Sucrose gradients. The digested DNA must be completely dissolved and disaggregated. Undissolved or aggregated DNA will form a pellet at the bottom of the centrifuge tube.

The gradients must not be overloaded. If more than 0.5 mg of genomic DNA is loaded on a 38-ml sucrose gradient or 0.2 mg of genomic DNA on a 12-ml gradient, then the DNA will aggregate and fractionation will not occur. The amount of λ DNA that can be loaded is 10-fold less.

The distance the DNA migrates in the sucrose gradient affects the resolution of different size DNA fragments. The distance that the

DNA migrates is approximately proportional to the time of centrifugation and the rpm². The resolution in a particular size range can frequently be improved by altering the rate or time of centrifugation.

The gradients are relatively stable. Nevertheless, they must be handled with care and not jarred or mixed during preparation or fractionation. They should not be stored for long periods before or after use.

Partial digestion. The quality of the genomic DNA and the restriction enzymes used for this purpose are critical for success. The DNA must be high-molecular-weight. A useful insert fragment must have both ends generated by the restriction enzyme so it can be ligated to vector. In order to generate a majority of the random fragments with two “good” ends, the starting material must on average be at least 2, and preferably 3 or more times greater in length than the insert to be used in the cloning procedure. Furthermore, the genomic DNA must be cleavable with the restriction enzyme chosen to produce the partial digest. Procedures for producing high-molecular-weight DNA are described in *UNITS 2.2, 2.3, & 2.4*.

The restriction enzyme used for this procedure should be of high quality. Some batches of *Sau*3A or *Mbo*I do not produce DNA fragments with ends that can be ligated. New batches of these enzymes should be tested to ensure that the DNA fragments produced by digestion can be ligated.

In order for this procedure to be useful the partial digestion products should represent a set of random fragments that include the entire genome. Unfortunately, some restriction enzyme sites appear to be somewhat preferred by the enzyme to other sites. These preferred sites lead to the production of nonrandom partial digestion products and can eventually lead to an under- or overrepresentation of particular sequences in the library. Several methods have been proposed to ensure the randomness of the partial digestion. We recommend that conditions are chosen such that at least 30% of the genomic DNA remains undigested as high-molecular-weight DNA.

Complete digestion. See second support protocol.

Anticipated Results

Sucrose gradients. 0.5 mg of digested mammalian genomic DNA produces 5 to 10 μ g of size-fractionated DNA suitable for preparing bacteriophage or cosmid libraries. Nearly quantitative recovery of bacteriophage λ arms

can be expected.

Enzyme digestion. When available, 1 mg of mammalian genomic DNA is partially digested to produce enough DNA for a cosmid or bacteriophage library. Approximately 0.1 to 1 mg of genomic DNA is completely digested to produce a subgenomic library. By optimizing procedures, the outcome of digestion should be collection of random fragments with ends that can be ligated to vector.

Time Considerations

Partial or complete digestion of DNA requires 1 hr. However, preparation of the digested DNA for size fractionation usually requires several more hours.

Gradients can be poured in ~1 hr. Centrifugation of gradients is performed overnight (~16 hr) and the fractions collected in ~1 hr. Identifying the desired fractions can be very

rapid if the fractions are identified by mobility of DNA on agarose gel or considerably longer if the gel must be analyzed by Southern blot analysis (*UNIT 2.9*).

Literature Cited

Seed, B., Parker, R.C., and Davidson, N. 1982. Representation of DNA sequences in recombinant DNA libraries prepared by restriction enzyme partial digestion. *Gene* 19:201-209.

Contributed by John H. Weis (sucrose gradients and partial digestion)
Harvard Medical School
Boston, Massachusetts

Thomas Quertermous (complete digestion)
Massachusetts General Hospital
Boston, Massachusetts

Size Fractionation Using Agarose Gels

UNIT 5.4

ELECTROPHORESIS ON A SLAB AGAROSE GEL

BASIC
PROTOCOL

Digested genomic DNA can be size fractionated on a slab agarose gel (see also *UNIT 2.5*). The capacity of a 20 × 20 × 1 cm agarose gel is limited to <2 mg DNA. After fractionation, the appropriate region of the gel is defined by Southern blot analysis (for subgenomic libraries) or by size (for genomic libraries) and the DNA is eluted.

Materials

400 μg to 2 mg digested genomic DNA (*UNIT 5.3*)
DNA size markers

Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5*), Southern blot transfer (*UNIT 2.9*), ethanol precipitation (*UNIT 2.1*), and hybridization (*UNITS 2.9 & 6.3*)

1. Construct a 20 × 20 × 1 cm gel with a large preparative well. A small well is placed along one edge of the gel for running a size marker or a small quantity (10 μg) of the digested genomic DNA.
2. Identify the region of the gel with the fragment size of interest by comparing to the size marker or the autoradiogram produced by blotting and hybridizing the small genomic lane that has been removed from the gel.
3. Cut out and elute three adjacent 1-cm slices of gel spanning the region of interest as described in *UNIT 2.6*.
4. Concentrate the DNA by ethanol precipitation, run on a gel, blot, and hybridize to determine which fraction has the highest relative concentration of the fragment of interest. DNA from this fraction is used for constructing the subgenomic library.

ELECTROPHORESIS ON THE BULL'S-EYE AGAROSE GEL APPARATUS

ALTERNATE
PROTOCOL

Digested genomic DNA is loaded onto a large preparative circular agarose gel in a Bull's-eye apparatus. The DNA fragments are electrophoresed toward the center of the gel and eluted. Fragments leaving the gel are pooled and constitute a fraction. Fractions containing the gene of interest are identified by Southern blotting of a small aliquot of alternate fractions. This procedure allows the purification of large amounts of size-fractionated DNA that is particularly well suited for genomic library construction and normally allows creation of large numbers of recombinant clones.

Additional Materials

TAE buffer (*APPENDIX 2*)
Chloroform
0.5 M EDTA, pH 7.5
Buffered phenol (*UNIT 2.1*)
TE buffer (*APPENDIX 2*)

Bull's-eye electrophoresis apparatus (Hoefer)
Bull's-eye electronic control apparatus (Hoefer)
Dialysis tubing (flat width >1.69 in.)
Peristaltic pump, fixed rate (20 ml/min)
Peristaltic pump, variable rate
Fraction collector (capable of holding 16 × 150 mm tubes; fraction time adjustable up to 1.5 hr)

Construction of
Recombinant
DNA Libraries

Contributed by Thomas Quertermous

Current Protocols in Molecular Biology (1987) 5.4.1-5.4.4

Copyright © 2000 by John Wiley & Sons, Inc.

5.4.1

1. Place the gel mold in the electrophoresis apparatus (see manufacturer's directions) and seal joints with molten agarose. Size of gel is determined by the quantity of DNA to be run; ≥ 2 mg can be run on a 350-ml gel while 10 mg requires a maximum size 700-ml gel. Add agarose to TAE buffer, boil, and pour when it cools to 55°C (see UNIT 2.5A for discussion of percentage agarose). Set overnight.
2. Slip a segment of moist dialysis tubing over the central electrode assembly and fasten in place with O rings. Fill this chamber with water and check for leaks. If a leak occurs, adjust the O rings or replace the dialysis tubing.
3. Set up the Bull's-eye apparatus following manufacturer's instructions. Use the fixed-rate peristaltic pump in the recirculation circuit, and the adjustable peristaltic pump in the fraction collection circuit. Set the rate of the adjustable pump to allow all the fraction to elute during the collection period. Fill the apparatus with 3 liters of TAE buffer; set the level of the upper reservoir to the height of the gel by adjusting the drain height. This height also determines the sample size, which varies from 5 to 8 ml.
4. Apparatus settings:
 - Power supply—125 mA
 - Control unit:
 - Electrophoresis—30 min
 - Reverse—40 sec
 - Elution—5 min (depending on sample volume and peristaltic pump speed)
 - Fraction collector—38 min
5. Operate the machine for several cycles to identify difficulties before the sample is loaded. The DNA sample volume that can be loaded depends on the gel former used, but generally the sample slot can hold 5 ml (375-ml gel) to 10 ml (700-ml gel). Add 0.1 ml gel loading buffer per ml sample volume. Save ~ 50 μ g DNA to run as total genomic DNA on the fraction blots.

Fractions collected should range from 0.5 to 20 kb and, depending on the gel running length and gel percentage, a complete run may require 5 to 7 days. Large DNA fragments elute slowly from the gel, so it is advisable to increase the electrophoresis time to 60 or 90 min after ~ 50 fractions have been collected. The fraction collector intervals also have to be increased accordingly.
6. Samples can be conveniently stored in scintillation vials. To each vial, add several drops of chloroform and 50 μ l of 0.5 M EDTA, pH 7.5.
7. Run a preliminary minigel with 50 μ l of every third to fifth fraction to determine DNA concentration and size distribution. Then run a gel for Southern blot analysis that contains every second or third fraction in the size range of interest, including a lane of the total digested genomic DNA saved in step 5.

Of each fraction, 400 to 500 μ l generally provides enough DNA for an adequate blot and can be concentrated by ethanol precipitation or simply lyophilized to dryness.
8. Extract the fraction of interest with phenol and chloroform. Ethanol precipitate ~ 500 μ l. Resuspend the DNA in a small volume of TE buffer (10 to 20 μ l) and determine the concentration by fractionating a small portion on a minigel (UNIT 2.5A).

COMMENTARY

Background Information

Electrophoretic size fractionation of digested DNA permits enrichment of a restriction fragment over 100-fold (Carreira et al., 1980). DNA fragments obtained from a Bull's-eye gel are usually better resolved than DNA fractionated on a slab gel. This high resolution is particularly useful for the construction of subgenomic libraries. Use of the enriched fragment in constructing a subgenomic library allows one to clone even a single-copy gene by screening a fraction of the number of clones required to represent the entire genome.

This technology requires that one's interest is confined to a single restriction fragment (or at most 2 to 3), and thus requires some prior knowledge of the restriction map of the locus. It is possible to clone several restriction fragments and reconstruct a locus, but this is technically difficult. One very successful application of subgenomic libraries has been the cloning of multiple genomic rearranged immunoglobulin and T-cell receptor genes. In this case, knowledge of the restriction map in the joining region allows one to choose a restriction fragment that is likely to contain all the variable region sequences of interest.

Literature Review

The concept of simplifying cloning by gene enrichment has been reviewed by Edgell et al. (1979). They describe a two-step DNA fractionation procedure involving RPC5 column chromatography followed by preparative electrophoresis. Bott et al. (1980) have used preparative agarose electrophoresis to fractionate the *Bacillus subtilis* genome to produce DNA fragments for transformation analysis and cloning. The degree of enrichment achieved with preparative electrophoresis alone is adequate and allows one to start with a smaller quantity of DNA.

Methods for constructing an apparatus that allows collection of multiple samples of size-fractionated DNA from a slab horizontal gel have been described (Polisky et al., 1978; Southern, 1979a,b). The resolution on a preparative agarose slab gel is adequate (although not as good as the Bull's-eye apparatus). The primary difficulty with some preparative agarose gels is that the DNA that is eluted contains impurities that inhibit DNA ligase and prevent cloning. The Bull's-eye apparatus is discussed here because it is commercially available, yields DNA that can be ligated, and produces large numbers

of clones. Further, the Bull's-eye machine allows size fractionation and concentration of large amounts of DNA. The theory of design of the Bull's-eye apparatus and specifications for construction have been published (Southern, 1979a,b; Carreira et al., 1980).

Critical Parameters

Resolution of DNA in a preparative agarose gel is dependent on the conditions of electrophoresis, and can be defined by the thickness of the imaginary slice of gel that contains all of the fragment of interest but no additional larger or smaller fragment. In general, experience indicates that the thinnest slice is likely to be >3 mm thick. Resolution in the Bull's-eye apparatus is determined by several factors, including the concentration of the gel (Southern, 1979a,b). As the concentration of the gel increases, so does the time required for a DNA fragment of a certain size to elute. There is an upper limit to gel concentration, however, because with higher concentrations the time required for separation becomes unacceptably long, and the DNA concentration in each sample is reduced. Generally, a gel concentration of 1.2% is optimal for separation of fragments for cloning.

Troubleshooting

Using the preparative agarose gel technique, fragments can generally be identified and eluted. Problems sometimes arise in that not enough clones can be obtained to make a subgenomic library (see below). Generally, if this occurs, it is best to start over with another fractionation technique.

The Bull's-eye apparatus is not easy for the beginner. Setting up the apparatus can be difficult, and instruction from someone skilled in the use of this machine is advised. In particular, placement of the dialysis tubing over the central anode is a tricky and crucial step in assembly. However, once assembled correctly the machine rarely malfunctions.

Occasionally, bubbles will impede flow of buffer in one of the circuits. This is prevented by clearing the tubing of bubbles before starting a run. Failure of the tubing leading to the fraction collector to run dry in less than 7 min usually means that either the peristaltic pump in this circuit is too slow or that there is a leak into the sample collection chamber. Such a leak occurs through the dialysis tubing that covers the anode and is corrected by replacing the

dialysis tubing.

DNA fractionated in this fashion and cleaned by ethanol precipitation and chloroform extractions usually can be ligated efficiently into bacteriophage or plasmid vectors. If appropriate controls point to the fractionated DNA as the source of difficulty, it may help to repeat the phenol and chloroform extractions or to pass the DNA over an Elutip column (UNIT 2.6). If the DNA still cannot be ligated it probably means that the restriction fragment ends were damaged by exonuclease activity.

Anticipated Results

Up to 80% of the DNA applied to a gel (either a slab gel or a Bull's-eye gel) should be recovered. If 1 mg DNA is digested and loaded onto the gel, and 100 fractions collected, each fraction will contain several micrograms of DNA. This is enough DNA to produce a number of subgenomic libraries. The size distribution of the fragments will vary, depending on the enzyme used, but fractions around 4 kb will usually contain the most DNA.

Time Considerations

Fractionation in a preparative agarose slab gel is rapid, generally <24 hr. However, the identification of the desired region and elution of the fragment from the gel by Southern blotting will require 2 to 5 days.

Fractionation with a Bull's-eye apparatus is usually slower. Depending on the percentage of gel used, the current, and other parameters, the time required for a complete size fractionation is several days. It is easy to monitor the size of fragments coming off the gel by running a small quantity of eluted fractions on an

agarose minigel. The run may then be terminated as soon as the desired size DNA is eluted. Because the DNA fractions can be stored for years, however, it is wise to collect all the DNA for future cloning projects.

Literature Cited

- Bott, F., Moran, C.P., Edgell, M.H., Drew, B., and Charles, L. 1980. Direct fractionation of genes by preparative electrophoresis of *Bacillus subtilis* DNA. *Gene* 10:283.
- Carreira, L.H., Carlton, B.C., Bobbio, S.M., Nagao R.T., and Meagher, R.B. 1980. Construction and application of a modified "Gene Machine": A circular concentrating preparative gel electrophoresis device employing discontinuous elution. *Anal. Biochem.* 106:455-468.
- Edgell, M.H., Weaver, S., Haigwood, N., and Hutchinson, C.A. 1979a. Gene enrichment. In *Genetic Engineering*, Vol. I (J.K. Setlow and A. Hollander, eds.) pp. 37-49. Plenum, New York.
- Polsky, F., Edgell, M.H., Seidman, J.G., and Leder, P. 1978. High-capacity gel preparative electrophoresis for purification of fragments of genome DNA. *Anal. Biochem.* 87:397-410.
- Southern, E.M. 1979a. Gel electrophoresis of restriction fragments. *Meth. Enzymol.* 68:152-176.
- Southern, E.M. 1979b. A preparative gel electrophoresis apparatus for large-scale separations. *Anal. Biochem.* 100:304.

Key Reference

Southern, E.M. 1979a. See above.

Describes construction of a Bull's-eye electrophoresis apparatus and principles that govern separation of DNA fragments on this device.

Contributed by Thomas Quertermous
Massachusetts General Hospital
Boston, Massachusetts

PREPARATION OF INSERT DNA FROM MESSENGER RNA

Construction of cDNA libraries from mRNA requires a series of steps similar to those outlined in Figure 5.5.1. Conversion of mRNA into double-stranded DNA suitable for insertion into a vector requires the action of at least six different enzymes. This conversion is described in two separate units: (1) conversion of mRNA into double-stranded cDNA (UNIT 5.5) and (2) methylation and addition of linkers to double-stranded cDNA (UNIT 5.6). After completion of the protocols described in these units, the double-stranded DNA should be suitable for insertion into a vector. Although the procedure is divided into two parts, there are several convenient stopping places as noted in the Commentaries of the two units.

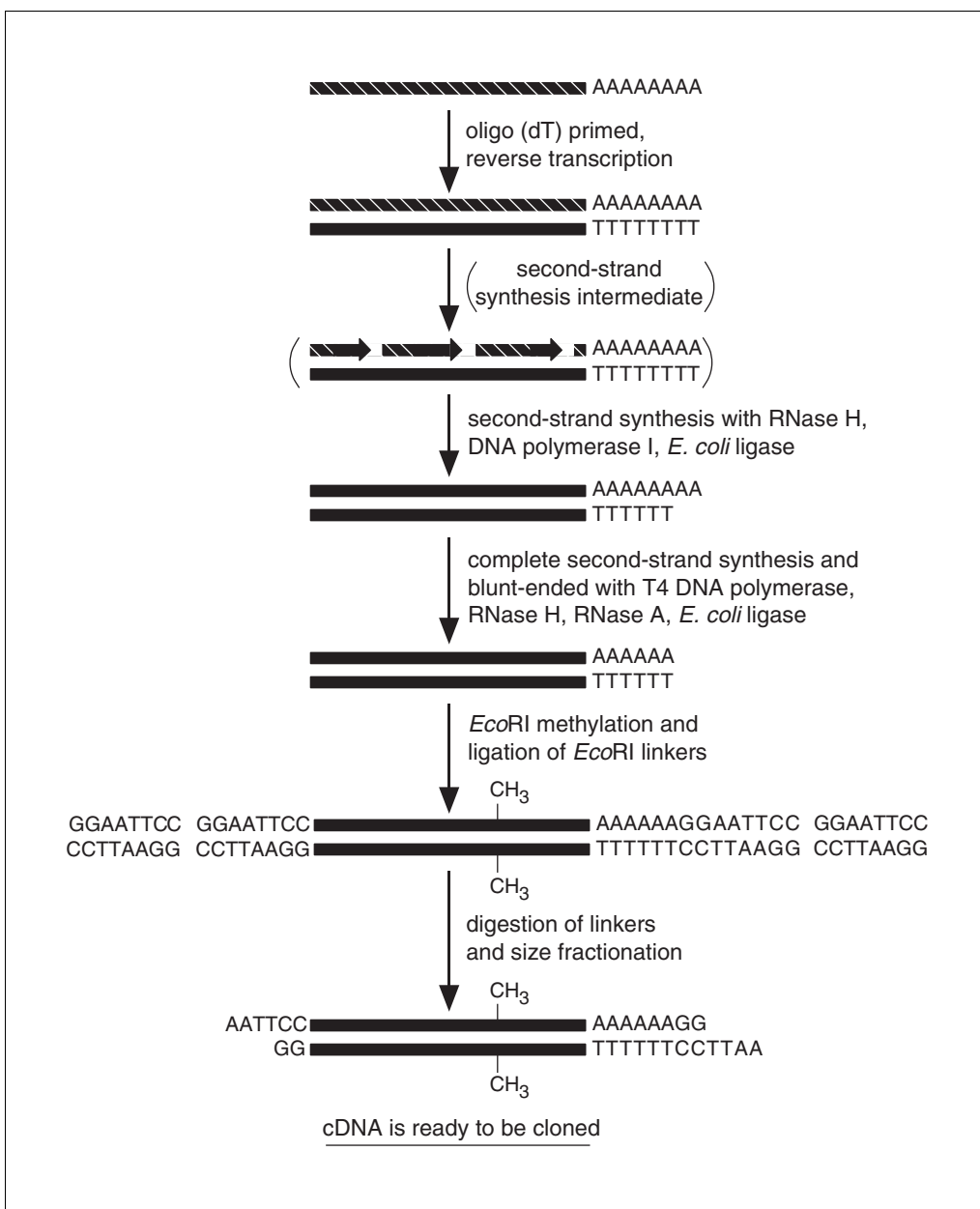


Figure 5.5.1 Outline of cDNA synthesis and preparation for insertion into a vector.

Conversion of mRNA into Double-Stranded cDNA

Enzymatic conversion of mRNA into double-stranded insert DNA can be accomplished by a number of different procedures. All of them involve the action of reverse transcriptase and oligonucleotide-primed synthesis of cDNA. After that, the procedures in common use diverge considerably. There are a number of methods for synthesizing the second strand and several procedures for producing suitable ends for making clonable DNA. The major goals of these procedures are to construct insert DNA that is as long as possible, with a high yield of conversion of mRNA into DNA that can ligate to vector DNA. The following protocols require only commercially available reagents and are usually successful in producing good cDNA libraries. The basic protocol describes a method for making blunt-ended cDNA that can then be ligated to linkers (UNIT 5.6) for subsequent cloning into a unique restriction site such as *EcoRI*. The alternate protocol is a variation that requires fewer enzymatic manipulations and allows construction of directional cDNA libraries, which are particularly desirable when the goal is to generate expression cDNA libraries. The alternate protocol takes advantage of a linker-primer consisting of (in order from 3' to 5') an oligo(dT) primer, a restriction site for the *XhoI* endonuclease, and a (GA)₂₀ repeat to protect the restriction site during generation of the blunt-ended cDNA. The internal *XhoI* sites on the individual cDNA molecules are protected by incorporation of 5-methyl-dCTP in the first-strand nucleotide mix. The resulting cDNAs having unique ends can be cloned into *EcoRI/XhoI*-digested vectors after ligation of *EcoRI* adaptors to the 5' end and digestion by *XhoI* to release the 3' *XhoI* sites that were incorporated into the cDNA by the linker-primer. These changes result in a considerably streamlined procedure that is substantially faster and easier than the basic protocol.

The quality and yield of insert DNA are in large part determined by the quality of the mRNA. An important feature of these procedures is that the yield at each step determines the yield at subsequent steps. Monitoring the efficiency of each reaction and determining the yield at each step is thus critical to the overall success of the project (and making yield determinations is strongly recommended). Guidelines for measuring the yield at each step are provided in the protocol. Because these protocols each consist of a series of sequential reactions, the quality of the enzymes used at each step is critical.

NOTE: For laboratories using recombinant DNA techniques and isolating large quantities of plasmid, bacteriophage, or cosmid DNA, remember that the smallest amount of contamination of genomic DNA or cDNA with recombinant DNA is disastrous. Contamination of cDNA at 1 ppm will cause great difficulty. Thus, all plasticware, glassware, and reagents used for the preparation of mRNA and cDNA should be maintained separately from those used for plasmid or bacteriophage DNA preparation. Extensive use of disposable plasticware is strongly recommended.

BASIC PROTOCOL

CONVERSION OF mRNA INTO BLUNT-ENDED DOUBLE-STRANDED cDNA

The conversion of mRNA into double-stranded cDNA for insertion into a vector is carried out in two parts. First, intact mRNA hybridized to an oligo(dT) primer is copied by reverse transcriptase and the products isolated by phenol extraction and ethanol precipitation. Then, in a single reaction vessel, the following reactions are carried out. The RNA in the RNA-DNA hybrid is removed with RNase H as *E. coli* DNA polymerase I fills in the gaps, similar to a nick translation. The second-strand fragments thus produced are ligated by *E. coli* DNA ligase. Second-strand synthesis is completed, residual RNA degraded, and cDNA made blunt with RNase H, RNase A, T4 DNA polymerase, and *E. coli* DNA ligase.

Conversion of
mRNA into
Double-Stranded
cDNA

5.5.2

Materials

- 5 mM 4dNTP mix (UNIT 3.4)
 - 5× reverse transcriptase (RT) buffer (see recipe)
 - 200 mM dithiothreitol (DTT)
 - 0.5 mg/ml oligo(dT)_{12–18} (Pharmacia Biotech; store at –80°C) *or* 15- to 40-mer antisense primer *or* random-hexamer primers
 - RNasin ribonuclease inhibitor (Promega; store at –20°C)
 - AMV (avian myeloblastosis virus) reverse transcriptase (Life Sciences; UNIT 3.7)
 - 10 μCi/μl [α -³²P]dCTP (10,000 Ci/mmol)
 - 0.5 M EDTA, pH 8.0
 - Buffered phenol (UNIT 2.1)
 - TE buffer, pH 7.5 (APPENDIX 2)
 - Diethyl ether or 24:1 chloroform/isoamyl alcohol
 - 7.5 M ammonium acetate
 - 95% and 70% ethanol, ice-cold
 - 10% trichloroacetic acid (TCA), ice-cold
 - 5× second-strand buffer I (see recipe)
 - 5 mM β-NAD⁺ (Sigma; store at –80°C)
 - RNase H (Pharmacia Biotech; UNIT 3.13)
 - E. coli* DNA ligase (New England Biolabs; UNIT 3.14)
 - E. coli* DNA polymerase I (New England Biolabs; UNIT 3.5)
 - 5× TA buffer (see recipe)
 - 2 μg/ml RNase A, DNase-free (see recipe and UNIT 3.13)
 - T4 DNA polymerase (Boehringer Mannheim; UNIT 3.5)
 - 10 mg/ml tRNA (store at –20°C)
 - 42° and 65°C water baths
 - Nitrocellulose membrane filter
 - 14°C incubator
- Additional reagents and equipment for preparation of poly(A)⁺ RNA (UNIT 4.5) and purification and concentration of DNA (UNIT 2.1)

Synthesize cDNA

1. Prepare ≥10 μg poly(A)⁺ RNA at a concentration of 1 μg/μl.

A small amount of ribosomal RNA will not interfere.

The quality of the mRNA is very important for full-length cDNA insert production. One approach to evaluating the mRNA is to perform northern blot analyses with a probe specific for a relatively abundant mRNA species known to be present (e.g., actin). The band in the northern blot should be distinct, with minimal trailing or smearing from the band toward the bottom of the gel. Other tests for determining the quality of the mRNA are suggested in UNIT 4.9.

2. Heat RNA (10 μg in 10 μl) in a tightly sealed microcentrifuge tube 5 min at 65°C, then place immediately on ice.

Alternatively, methylmercuric hydroxide denaturation of the RNA can be performed prior to first-strand synthesis. However, this is usually not necessary to achieve full-length synthesis of even long (>5 kb) cDNA.

3. In a separate tube add in the following order (180 μ l total):

20 μ l 5 mM dNTPs (500 μ M final each)
40 μ l 5 \times RT buffer (1 \times final)
10 μ l 200 mM DTT (10 mM final)
20 μ l 0.5 mg/ml oligo(dT)₁₂₋₁₈ (50 μ g/ml final)
60 μ l H₂O
10 μ l (10 U) RNasin (50 U/ml final).

Mix by vortexing, briefly microcentrifuge, and add the mixture to the tube containing the RNA. Add 20 μ l (200 U) AMV reverse transcriptase for a final concentration of 1000 U/ml in 200 μ l. Mix as above and remove 10 μ l to a separate tube containing 1 μ l of [α -³²P]dCTP. Leave both tubes at room temperature 5 min, then place both tubes at 42°C for 1.5 hr.

The aliquot is removed to determine incorporation and permit an estimation of recovery. The remainder of the cDNA will be labeled during second-strand synthesis. Labeling cDNA during first-strand synthesis to a high enough specific activity to permit easy detection with a hand-held radiation monitor during all subsequent steps requires a relatively large amount of label, which may then interfere with reverse transcription due to buffer effects.

Some investigators check the quality of the cDNA by fractionating the radiolabeled cDNA on an alkaline agarose gel and detecting it by autoradiography. Much of the cDNA should be >1000 bp long.

For a specifically primed library, substitute an equal weight of antisense 15- to 40-mer primer for the oligo(dT)₁₂₋₁₈. Expect a 100-fold enrichment of specific clones in the library. For a randomly primed library, substitute an equal weight of random-hexamer primers for oligo(dT)₁₂₋₁₈ (or use a 50:50 mix), and perform the reverse transcription at 37° instead of 42°C.

4. Add 1 μ l of 0.5 M EDTA, pH 8.0, to the radioactive reaction and freeze it at -20°C. It will be used later to estimate the amount of cDNA synthesized.
5. To the main reaction add 4 μ l of 0.5 M EDTA, pH 8.0, and 200 μ l buffered phenol. Vortex well, microcentrifuge at room temperature for 1 min to separate phases, and transfer the upper aqueous phase to new tube.

Save the tube containing the phenol layer, too.

6. Add 100 μ l TE buffer, pH 7.5, to the phenol layer and vortex and microcentrifuge as in step 5. Remove the aqueous layer and add it to the aqueous phase from the first extraction.

The volume of aqueous phase is now about 300 μ l; the phenol may be discarded.

Back extraction of the organic phase at each phenol extraction significantly improves the yield. See UNIT 2.1 for a discussion of phenol extraction and ethanol precipitation.

7. Add 1 ml diethyl ether, vortex, and microcentrifuge as in step 5. Remove and discard the upper (ether) layer with a glass pipet. Repeat the extraction with an additional 1 ml of ether.

A single chloroform/isoamyl alcohol extraction followed by back extraction of the organic phase may be substituted for the two ether extractions; however, the yield is normally slightly lower.

8. Add 125 μ l of 7.5 M ammonium acetate to the aqueous phase (final concentration 2.0 to 2.5 M) and 950 μ l of 95% ethanol. Place in dry ice/ethanol bath 15 min, warm

to 4°C, and microcentrifuge at 10 min at full speed, 4°C, to pellet nucleic acids. A small, yellow-white pellet may be visible.

Precipitation from ammonium acetate leaves short oligonucleotides in the supernatant, thus removing the oligo(dT) primer and enriching the pellet in longer cDNAs. Do not substitute sodium acetate for ammonium acetate.

9. Remove the supernatant with a pipet, fill the tube with ice-cold 70% ethanol, and microcentrifuge 3 min at full speed, 4°C. Again remove the supernatant, then dry the tube containing the precipitated nucleic acids briefly in a vacuum desiccator.
10. Thaw the tube containing the radioactive aliquot of the first-strand synthesis reaction and spot the sample onto a nitrocellulose membrane filter.
11. Wash the membrane with ice-cold 10% TCA and determine the radioactivity bound to the filter with a fluor and scintillation counter. Use the specific activity of the radiolabel in the reaction, the amount of RNA used, the counts incorporated, and the efficiency of the β -counter to calculate the amount of cDNA synthesized (see Sample Calculation for Determining Amount of cDNA Synthesized, in Commentary).

From 1 to 4 μ g is typical even though the theoretical maximum is 10 μ g.

Convert cDNA into double-stranded cDNA

12. Resuspend the pellet from the first-strand synthesis in 284 μ l water and add to the tube in the following order (400 μ l total):

4 μ l 5 mM dNTPs (50 μ M final each)
80 μ l 5 \times second-strand buffer I (1 \times final)
12 μ l 5 mM β -NAD⁺ (150 μ M final)
2 μ l 10 μ Ci/ μ l [α -³²P]dCTP (50 μ Ci/ml final).

Mix by vortexing, briefly microcentrifuge, and add:

4 μ l (4 U) RNase H (10 U/ml final)
4 μ l (20 U) *E. coli* DNA ligase, (not T4 DNA ligase; 50 U/ml final)
10 μ l (100 U) *E. coli* DNA polymerase I (250 U/ml final).

Mix by vortexing, briefly microcentrifuge, and incubate 12 to 16 hr at 14°C.

*Unrelated cDNA fragments are not ligated in this reaction because *E. coli* DNA ligase does not catalyze blunt ligations.*

13. After second-strand synthesis, remove 4 μ l of the reaction to a new tube and freeze at -20°C. Later, when time permits, determine the incorporation of radiolabel into acid-insoluble material as outlined in steps 10 and 11.

Expect 1-10 $\times 10^6$ cpm incorporated in the total reaction.

14. Phenol extract the second-strand synthesis reaction with 400 μ l buffered phenol and back extract the phenol phase with 200 μ l TE buffer, pH 7.5, as in steps 5 and 6.
15. Pool the two aqueous phases and extract twice with 900 μ l ether, as in step 7. The volume of the aqueous phase is now ~600 μ l.
16. Divide the aqueous phase evenly between two tubes, add ammonium acetate, and ethanol precipitate as in steps 8 and 9.

Unincorporated radioactive dCTP is removed by the ethanol precipitation and wash steps. From this point on, the cDNA may be followed with a hand-held radiation monitor.

Create blunt ends on double-stranded cDNA

17. Complete second-strand synthesis and blunt the double-stranded cDNA by resuspending the pooled pellets in 42 μ l water. Add in the following order (80 μ l total):

- 5 μ l 5 mM dNTPs (310 μ M final each)
- 16 μ l 5 \times TA buffer (1 \times final)
- 1 μ l 5 mM β -NAD⁺ (62 μ M final).

Mix by vortexing, microcentrifuge briefly, and add:

- 4 μ l of 2 μ g/ml RNase A (100 ng/ml final)
- 4 μ l (4 U) RNase H (50 U/ml final)
- 4 μ l (20 U) *E. coli* DNA ligase (250 U/ml final)
- 4 μ l (8 U) T4 DNA polymerase (100 U/ml final).

Mix as above and incubate 45 min at 37°C.

Volume of T4 DNA polymerase used to obtain 8 U may require adjustment depending on batch.

If the library is to be screened with an antiserum, some investigators (see Tamkun et al., 1986) digest the cDNA at this point with AluI or HaeIII to prepare small inserts that may produce more stable fusion proteins in λ gt11. Do this after the T4 polymerase step by diluting the 80- μ l reaction to 170 μ l with water; add 24 μ l of 5 \times TA buffer and 6 μ l of one of the above restriction enzymes. Incubate 1 hr at 37°C, then proceed to step 18, except do not add any TE buffer. An insert isolated from an immunological screen may then be used to screen a second library of full-length inserts.

18. Add 120 μ l TE buffer, pH 7.5, and 1 μ l of 10 mg/ml tRNA. Extract with 200 μ l buffered phenol and back extract the phenol phase with 100 μ l TE buffer as described in steps 5 and 6.

19. Pool the two aqueous phases and extract twice with 1 ml ether, as in step 7.

20. Ethanol precipitate the cDNA as in steps 8 and 9.

The cDNA is now ready to be tailed or linked to create compatible ends for subsequent cloning steps.

ALTERNATE PROTOCOL

CONVERSION OF mRNA INTO DOUBLE-STRANDED cDNA FOR DIRECTIONAL CLONING

Generation of cDNA that has unique ends for directional cloning is carried out in two parts, as in the basic protocol. First, the mRNA is hybridized to a linker-primer that incorporates a poly(dT) tract (at its 3' end) as well as a restriction site for *Xho*I (Fig. 5.5.2). The linker-primer is extended using an RNase H⁻ version of the Moloney murine leukemia virus reverse transcriptase (SuperScript) and a nucleotide mix in which dCTP is replaced with 5-methyl-dCTP. When first-strand synthesis is completed, the reaction mixture is transferred into a second tube that contains the prechilled second-strand mixture. The second strand is synthesized using RNase H and *E. coli* DNA polymerase I. Finally, a blunting step (consisting of treatments with mung bean nuclease and Klenow fragment) is carried out to prepare the cDNA for the *Eco*RI adaptor ligation.

Additional Materials (also see Basic Protocol)

- 5 \times SuperScript buffer (RNase-free; see recipe)
- 0.1 M DTT (RNase-free)
- 3dNTP/methyl-dCTP mix: 10 mM each dATP, dGTP, dTTP, and 5-methyl-dCTP (Pharmacia Biotech)

continued

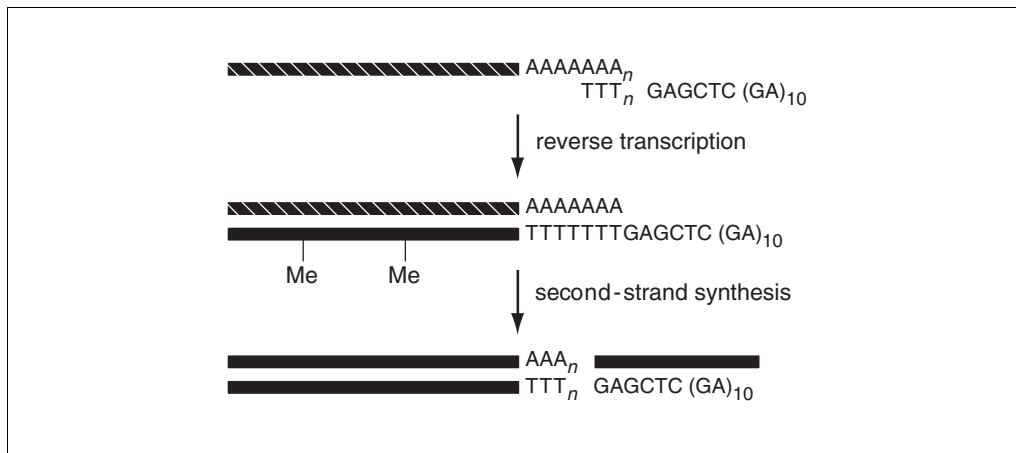


Figure 5.5.2 Synthesis of cDNA with unique ends for directional cloning.

0.25 $\mu\text{g}/\mu\text{l}$ oligonucleotide primer (*UNIT 2.11*) incorporating (from 5' to 3'):
 (dGdA)₁₀, *Xho*I restriction site, and (dT)₁₈

200 U/ μl SuperScript or SuperScript II (GIBCO/BRL)

5 \times second-strand buffer II (see recipe)

10 mM and 2 mM 4dNTP mix (Pharmacia Biotech)

10 $\mu\text{Ci}/\mu\text{l}$ [α -³²P]dATP (3000 Ci/mmol)

0.8 U/ μl RNase H (Pharmacia Biotech)

10 U/ μl *E. coli* DNA polymerase I (New England Biolabs)

1:1 (w/v) phenol/chloroform

100% ethanol ice-cold

10 \times Klenow buffer (see recipe)

5 U/ μl Klenow fragment of *E. coli* DNA polymerase I (New England Biolabs)

10 \times mung bean nuclease buffer (see recipe)

10 U/ μl mung bean nuclease (New England Biolabs)

1 M Tris·Cl, pH 8.0 (*APPENDIX 2*)

16°C incubator

Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5*)

Synthesize cDNA

1. Prepare 5 to 7 μg poly(A)⁺ RNA.

If mRNA is prepared as in UNIT 4.5, store under ethanol in a microcentrifuge tube until ready to proceed. If mRNA is in aqueous solution, add 3 M sodium acetate (RNase-free) to 0.3 M and 3 vol of 100% ethanol, chill ≥ 20 min at -20°C , and microcentrifuge 20 min at maximum speed, 4°C , to collect mRNA. Rinse the pellet with 70% ethanol (RNase-free), and invert the tube to air dry.

2. Prepare first-strand premix (45 μl final by mixing in the order listed):

10 μl 5 \times SuperScript buffer

5 μl 0.1 M DTT

2.5 μl 3dNTP/methyl-dCTP mix

1 μl RNasin

12 μl 0.25 $\mu\text{g}/\mu\text{l}$ oligonucleotide primer

14.5 μl H₂O.

3. Resuspend RNA pellet directly into premix. Incubate 15 min at room temperature, then add 5 μl of 200 U/ μl SuperScript and incubate 1 hr at 42°C .

4. Shortly before end of incubation, prepare second-strand premix (336 μ l final) containing:

80 μ l 5 \times second-strand buffer II
6 μ l 10 mM 4dNTP mix
2.5 μ l 10 μ Ci/ μ l [α -³²P]dATP
247.5 μ l H₂O.

Place on ice.

5. Continuing to work on ice, add the first-strand mix from step 3 (50 μ l total) to prechilled second-strand premix from step 4.
6. In a second microcentrifuge tube combine 4 μ l of 0.8 U/ μ l RNase H (3.2 U) and 10 μ l of 10 U/ μ l *E. coli* DNA polymerase I (100 U). Add combined enzymes to the tube from step 5 containing the cDNA and mix by rapidly inverting a few times.

In order to prevent formation of "snapbacks" during second-strand synthesis it is important to prechill the premix and add the RNase H and DNA polymerase simultaneously.

7. Microcentrifuge 5 sec at maximum speed, then incubate reaction 1 hr at 16°C and 1 hr at room temperature.

At this step, second-strand synthesis is completed.

8. Electrophorese ~2% to 5% of total reaction on a 0.7% agarose gel, and monitor size range of labeled cDNA by autoradiography.

Optimally, a smear ranging from ~200 bp to >8 kb will be seen, with a peak at 1 to 2 kb.

The reaction can be frozen overnight (or longer) at -20°C, or one can immediately proceed to blunting and linker ligation.

Create blunt-ends on double-stranded cDNA

9. Add 500 μ l of 1:1 phenol/chloroform to reaction. Microcentrifuge 5 min at maximum speed to separate phases.

If reaction was frozen, thaw by vortexing with the phenol/chloroform.

10. Remove upper phase to a fresh tube, add 200 μ l of 7.5 M ammonium acetate, and mix by inversion.
11. Divide sample evenly between two tubes (300 μ l in each). Add 600 μ l ice-cold 100% ethanol to each tube, place both tubes 5 min at -80°C, then microcentrifuge 20 min at maximum speed, 4°C.

This and the following series of steps involve transfer of labeled cDNA from tube to tube, and require a number of extractions and precipitations. After this initial precipitation, the bulk of the radiolabel is incorporated in the cDNA. It is important to monitor using a Geiger counter to ensure that all cDNA is transferred from step to step and not lost during the extractions.

12. Pour off supernatants and rinse pellets with ice-cold 70% ethanol. Dry briefly in a Speedvac evaporator. Add 30 μ l water to one of the two tubes, vortex to resuspend cDNA pellet, then transfer contents to the second tube.
13. Add an additional 10 μ l water, 5 μ l of 10 \times Klenow buffer, and 5 μ l of 2 mM 4dNTP mix to the first tube. Vortex and transfer to the second tube. Add 5 μ l of 10 U/ μ l Klenow fragment to the second tube and incubate 30 min at 37°C.

The aim of the first Klenow treatment is to ensure that the 3' end of the cDNA is

completely blunted so that the XhoI restriction site will not be harmed by the subsequent enzymatic manipulations.

14. Add 50 μl of 1:1 phenol/chloroform to reaction, extract, and transfer upper phase to new tube. Back extract lower phase with an additional 50 μl water and combine it with the first extraction.
15. Add 50 μl of 7.5 M ammonium acetate and 300 μl ice-cold 100% ethanol. Place at -80°C for 5 min, then microcentrifuge 20 min at maximum speed, 4°C .
16. Pour off supernatant and rinse pellet with cold 70% ethanol. Dry pellet briefly in a Speedvac evaporator. Resuspend cDNA in 225 μl water.
17. Add 25 μl of 10 \times mung bean nuclease buffer and 1 μl of 1 U/ μl mung bean nuclease. Incubate 15 min at 37°C .

This mung bean exonuclease treatment removes any residual single-stranded nucleic acid extensions (i.e., single-stranded DNA, RNA overhangs, or small hairpins) from the 5' end of the cDNA.

18. Add 25 μl of 1 M Tris-Cl, pH 8.0. Extract with 200 μl 1:1 phenol/chloroform and back extract with 50 μl water, pooling the upper phases in a new tube.
19. Add 175 μl of 7.5 M ammonium acetate and fill tube with ice-cold 100% ethanol. Place at -80°C for 5 min, then microcentrifuge 20 min at maximum speed, 4°C .
20. Rinse pellet with 70% ethanol and air dry briefly. Resuspend in 20 μl water, then add 2.5 μl of 10 \times Klenow buffer and 1 μl of 5 U/ μl Klenow fragment. Incubate 5 min at 37°C . Add 2.5 μl of 2 mM dNTP mix to the reaction and incubate an additional 25 min at room temperature.

Do not overdry pellet (i.e., if Speedvac evaporation is used, do not dry >5 min).

The goal of this second Klenow fragment treatment is to effectively blunt any short single-stranded extensions that might remain after the mung bean exonuclease treatment or be caused by the "sloppiness" of this enzyme. This series of steps effectively causes essentially all the cDNA to become blunt-ended.

21. Add 25 μl of 1:1 phenol/chloroform to the reaction and extract, removing upper phase to a new tube. Back extract lower phase with an additional 25 μl water and pool with first upper phase. Store sample frozen at -20°C until ready for the *EcoRI* adaptor ligation step.

EcoRI adaptor ligation is carried out as it is described in UNIT 5.6. Subsequently, XhoI digestion of the EcoRI adaptor-ligated cDNA results in a 5' end that has an EcoRI-compatible overhang and a 3' end with XhoI-compatible overhang.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

Buffer stock solutions

Evaluate water and buffers for nuclease activity by incubating 5 μ l solution with 3 μ g of intact total cellular RNA or 0.5 μ g of a DNA fragment, whichever is relevant, for 30 min. Check the RNA or DNA by agarose gel electrophoresis to be sure it was not degraded.

Prepare at least 10 ml of each of the following 1 M stock solutions. Use RNA-grade water (see *UNIT 4.1*) and pass each solution through a sterile 0.45- μ m filter. Store at room temperature unless otherwise indicated.

1 M Tris·Cl, pH 8.2, 42°C

1 M Tris·acetate: titrate aqueous Tris base to pH 7.8 with acetic acid

1 M KCl

1 M MgCl₂

1 M (NH₄)₂SO₄

1 M potassium acetate

1 M magnesium acetate

1 M DTT, -20°C, in tightly capped tube

5 mg/ml nuclease-free BSA, -20°C

The pH of Tris buffers varies considerably with temperature. Be sure to measure the pH at the indicated temperature and use an electrode that accurately measures the pH of Tris buffers.

Enzymes

Enzymes should be of the highest quality available. Before starting a cDNA cloning experiment with a rare or hard-to-obtain RNA prep, it is advisable to check the activities of all the enzymes used.

Klenow buffer, 10×

0.5 M Tris·Cl, pH 7.6

0.1 M MgCl₂

Prepare 1.0 ml buffer using high-quality water and buffer stock solutions (see recipe), then check for nuclease activity as described in buffer stock solution recipe.

Freeze in 200- μ l aliquots in screw-cap microcentrifuge tubes at -80°C.

Mung bean nuclease buffer, 10×

0.3 M sodium acetate, pH 4.5

0.5 M NaCl

10 mM ZnCl₂

50% (v/v) glycerol

Prepare 1.0 ml buffer using high-quality water and buffer stock solutions (see recipe), then check for nuclease activity as described in buffer stock solution recipe.

Freeze in 200- μ l aliquots in screw-cap microcentrifuge tubes at -80°C.

Reverse transcriptase (RT) buffer, 5×

250 μ l 1 M Tris·Cl, pH 8.2

250 μ l 1 M KCl

30 μ l 1 M MgCl₂

470 μ l H₂O

Prepare 1.0 ml buffer using high-quality water and buffer stock solutions (see recipe), then check for nuclease activity as described in buffer stock solution recipe.

Freeze in 200- μ l aliquots in screw-cap microcentrifuge tubes at -80°C.

RNase A (DNase-free)

Dissolve 1 mg/ml RNase A (Sigma) in TE buffer (*APPENDIX 2*); boil 10 min to remove contaminating DNase and store in aliquots at -20°C . Dilute to 2 $\mu\text{g}/\text{ml}$ in TE buffer when needed.

Second-strand buffer I, 5 \times

100 μl 1 M Tris-Cl, pH 7.5
500 μl 1 M KCl
25 μl 1 M MgCl_2
50 μl 1 M $(\text{NH}_4)_2\text{SO}_4$
50 μl 1 M DTT
50 μl 5 mg/ml BSA
225 μl H_2O

Prepare 1.0 ml buffer using high-quality water and buffer stock solutions (see recipe), then check for nuclease activity as described in buffer stock solution recipe. Freeze in 200- μl aliquots in screw-cap microcentrifuge tubes at -80°C .

Second-strand buffer II, 5 \times

94 mM Tris-Cl, pH 6.9
453 mM KCl
23 mM MgCl_2
50 mM $(\text{NH}_4)_2\text{SO}_4$

Prepare 1.0 ml buffer using high-quality water and buffer stock solutions (see recipe), then check for nuclease activity as described in buffer stock solution recipe. Freeze in 200- μl aliquots in screw-cap microcentrifuge tubes at -80°C .

SuperScript buffer, 5 \times

250 mM Tris-Cl, pH 8.3
375 mM KCl
15 mM MgCl_2

Prepare 1.0 ml buffer using high-quality water and buffer stock solutions (see recipe), then check for nuclease activity as described in buffer stock solution recipe. Freeze in 200- μl aliquots in screw-cap microcentrifuge tubes at -80°C .

TA buffer, 5 \times

200 μl 1 M Tris-acetate, pH 7.8
400 μl 1 M potassium acetate
60 μl 1 M magnesium acetate
3 μl 1 M DTT
105 μl 5 mg/ml BSA
432 μl H_2O

Prepare 1.2 ml buffer using high-quality water and buffer stock solutions (see recipe), then check for nuclease activity as described in buffer stock solution recipe. Freeze in 200- μl aliquots in screw-cap microcentrifuge tubes at -80°C .

COMMENTARY

Background Information

cDNA synthesis and subsequent library preparation are necessary for the study of most mRNAs because no technology for conveniently manipulating and propagating RNA sequences is currently available. Of fundamental importance is that the cDNAs faithfully represent the sequence, size, and complexity of the mRNA population. The high quality of commercially available enzymes and reagents has dramatically changed the major concerns of the scientist preparing a cDNA library. In the past, considerable attention was paid to the source of reverse transcriptase, the quality of other enzymes, methods for their further purification, and conditions for optimum activity (Maniatis et al., 1982). These issues have been successfully addressed and full-length cDNAs of 5 to 8 kb mRNAs are now obtainable. The major focus currently is on the preparation of libraries of full-length cDNAs that will contain copies of even the rarest mRNA.

There are several relatively recent innovations incorporated into these protocols. First, it is recommended that size fractionation of the mRNA be omitted in order to minimize the opportunities for degradation of the mRNA template. Size fractionation of the cDNA, which is much more stable, is discussed in *UNITS* 5.3 & 5.4. Second-strand synthesis employs RNase H, *E. coli* DNA ligase, and *E. coli* DNA polymerase I as described by Okayama and Berg (1982) and modified by Gubler and Hoffman (1983). More recently, Neve et al. (1986) introduced during the blunting process the additional second-strand synthesis step, which increases the yield and size of the cDNA. Using this modification, S1 nuclease cleavage of the 5' "hairpin loop" generated by reverse transcriptase, a problematic step in the past, is avoided. The cDNA produced is ready to be prepared for insertion into an appropriate vector. Variations such as omitting the DNA ligase, (e.g., as per Gubler and Hoffman, 1983), omitting the additional second-strand synthesis, or substituting Klenow fragment for the intact *E. coli* DNA polymerase I have been successfully employed by others. Finally, the basic protocol uses AMV (avian myeloblastosis virus) reverse transcriptase (RT), but the alternate protocol uses SuperScript, a genetically engineered derivative of Moloney murine leukemia viral (MoMuLV) RT that has had its RNase H activity removed (Kotewicz et al., 1988) and its polymerase activity enhanced, resulting in a

greater size and yield of cDNA from first-strand synthesis.

A number of variations of the protocols in this unit exist for particular applications. Random hexamers can be used to prime first-strand synthesis rather than oligo(dT)-containing primers: resulting cDNAs will be enriched for the 5'-end sequences of very long transcripts, although average cDNA length will be shorter than with the methods described here. Another method that has gained popularity for a variety of analytical techniques is to perform a standard first-strand synthesis from an mRNA of interest, and then amplify specific rare transcripts using PCR.

Over the last several years, a number of kits for cDNA synthesis have become commercially available (e.g., from GIBCO/BRL, Stratagene, Promega, Amersham, and Pharmacia Biotech), and generally yield excellent results.

Critical Parameters

RNA used to generate the cDNA must be undegraded and of the highest quality obtainable. It must also be DNA-free.

Buffers and solutions must be nuclease-free. These may be divided into aliquots, frozen, and stored for long periods; thus, it is worthwhile to check each by incubating intact RNA or DNA, whichever is appropriate, with a sample of the solution and checking for degradation by agarose gel electrophoresis.

Enzymes (and all reagents) should be the highest quality available. If experimenting with a rare or hard-to-isolate RNA preparation, it is advisable to check the activities of all enzymes used. Enzymes should be relatively fresh and stored at -20°C in a freezer that is not frost-free (because the warming and cooling cycles of a frost-free freezer cause the enzymes to lose activity). The RNasin added in the first-strand synthesis is a potent though reversible inhibitor of RNase and must be added after the DTT, as its activity is dependent on the presence of a reducing agent.

Troubleshooting

Once high-quality mRNA has been obtained, the cDNA synthesis usually proceeds well. Difficulties, if any, only occur later—when the cDNA is ligated to the vector. If the yield of cDNA is low, likely causes are inactive reverse transcriptase, inactive *E. coli* polymerase I, or deteriorated dNTP mixes. If the calculated yield is low, double-check the result

Sample Calculation for Determining Amount of cDNA Synthesized

1. Assume 2×10^{12} cpm/Ci; thus, if $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ is at a specific activity of 3000 Ci/mmol, there are 6×10^{15} cpm/mmol.
2. The second-strand synthesis aliquot representing 1% of the total contains 10^5 TCA-precipitable cpm, thus

$$\frac{10^5 \text{ cpm} \times 100}{6 \times 10^{15} \text{ cpm/mmol}} = 1.67 \times 10^{-9} \text{ mmol } [\alpha\text{-}^{32}\text{P}]\text{dCTP incorporated.}$$

3. The initial concentration of $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ was

$$\frac{50 \mu\text{Ci/ml}}{3000 \text{ Ci/mmol}} = 1.67 \times 10^{-2} \mu\text{M.}$$

4. Thus, the total dCTP incorporated is

$$\frac{50 \mu\text{M}}{1.67 \times 10^{-2} \mu\text{M}} \times 1.67 \times 10^{-9} \text{ mmol} \times \frac{0.3 \text{ g dCMP}}{\text{mmol}} = 1.5 \mu\text{g.}$$

5. There were four deoxynucleotides present in the reaction, so the total cDNA synthesized in the second-strand reaction is $6 \mu\text{g}$ and since the cDNA is double-stranded, there are $12 \mu\text{g}$ of cDNA total.

by running 5% of the total sample on an agarose minigel and staining with ethidium bromide, or by determining the amount of DNA in the sample by optical density or DAPI fluorescence. If the radiolabel was old or deteriorated, incorporation will be low but the amount of cDNA synthesis will be correct.

Very small cDNA is usually the result of DNase contamination. This is readily evaluated by adding a restriction fragment of DNA to the reaction mixes and checking the fragment after incubation by gel electrophoresis.

Anticipated Results

The typical amount of poly(A)⁺ RNA used to generate a library is $10 \mu\text{g}$; however, the protocol may be scaled down should RNA availability be limiting. The $10 \mu\text{g}$ of poly(A)⁺ RNA will yield 2 to $10 \mu\text{g}$ of double-stranded cDNA. An aliquot representing 2% to 5% of the total volume may be checked by agarose gel electrophoresis, ethidium bromide staining, and autoradiography. The aliquots removed from the first- and second-strand synthesis should contain 2 to 8×10^5 cpm and 2 to 10×10^4 cpm TCA-insoluble counts, respectively (basic protocol). 100 ng double-stranded cDNA should generate a full-complexity library (1×10^6 recombinants) if a phage vector such as $\lambda\text{gt}10$ is employed.

To determine the amount of cDNA synthesized, see Sample Calculation for Determining Amount of cDNA Synthesized.

Time Considerations

Either of these protocols can usually be completed in two days. In the basic protocol, first-strand synthesis should be started in the morning to permit the initial second-strand synthesis to proceed overnight. The final second-strand synthesis should be carried out the following morning. In the alternate protocol, 4 hr should be allowed for synthesis of the double-stranded cDNA and 4 to 5 hr for generation of blunt ends. The cDNA may be linker or tailed as described in UNIT 5.6 later the same day. If necessary, the protocol may be interrupted at any ethanol precipitation step.

Literature Cited

- Gubler, U. and Hoffman, B.J. 1983. A simple and very effective method for generating cDNA libraries. *Gene* 25:263-269.
- Kotewicz, M.L., Sampson, C.M., D'Alessio, J.M., and Gerard, G.F. 1988. Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. *Nucl. Acids Res.* 16:265-277.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. 1982. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Neve, R.L., Hanis, P., Kosik, K.S., Kurnit, D.M., and Donlon, T.A. 1986. Identification of the gene for the human microtubule-associated protein tau and chromosomal localization of the genes for tau and microtubule-associated protein 2. *Brain Res.* 387(3):271-80.

Okayama, H. and Berg, P. 1982. High-efficiency cloning of full-length cDNA. *Mol. Cell. Biol.* 2:161-170.

Tamkun, J.W., DeSimone, D.W., Fonda, D., Patel, R.S., Buck, D., Horwitz, A.F., and Hynes, R.O. 1986. Structure of integrin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin. *Cell* 46:271-282.

Key References

Gubler and Hoffman, 1983. See above.

Huse, W.D. and Hansen, C. 1988. cDNA cloning redefined. *Strategies (Stratagene)* 1:1-3.

First demonstrated use of linker primers and methylated nucleotides in cDNA synthesis.

Okayama and Berg, 1982. See above.

These authors first developed the RNase H/E. coli DNA polymerase I alternative to S1 nuclease and Klenow fragment for second-strand synthesis.

Contributed by Lloyd B. Klickstein
Brigham and Women's Hospital
Boston, Massachusetts

Rachael L. Neve
McLean Hospital
Belmont, Massachusetts

Erica A. Golemis
Fox Chase Cancer Center
Philadelphia, Pennsylvania

Jeno Gyuris
Mitotix, Inc.
Cambridge, Massachusetts

Ligation of Linkers or Adapters to Double-Stranded cDNA

UNIT 5.6

Linkers or adapters can be ligated to double-stranded cDNA (UNIT 5.5) to provide restriction endonuclease sites used in the production of a cDNA library (UNIT 5.8). For cloning purposes, only one linker or adapter must be present on each end of the cDNA. However, multiple linkers are usually ligated to the cDNA because both ends are phosphorylated and contain cohesive sequences. As a result, cDNA must first be methylated to protect it from a subsequent restriction digest designed to remove the multiple linkers (basic protocol). The procedure for ligating adapters to the cDNA is much simpler than that for linkers because only one end is phosphorylated, resulting in the ligation of just one adapter (alternate protocol). Linkered or adapted cDNA is then passed over a Sepharose CL-4B column to remove unligated linkers or adapters and other low-molecular-weight material (<350 bp) that would interfere with cloning (support protocol). This double-stranded cDNA is concentrated by ethanol precipitation and may be cloned directly (UNIT 5.8) or further fractionated by agarose gel electrophoresis (UNITS 2.5 & 2.6).

METHYLATION OF cDNA AND LIGATION OF LINKERS

BASIC
PROTOCOL

To convert blunt-ended, double-stranded cDNA into DNA suitable for ligation to a vector, it is methylated by *EcoRI* methylase, ligated to *EcoRI* linkers, and digested with *EcoRI*. The methylation step protects *EcoRI* sites in the cDNA from *EcoRI* digestion. Linkered cDNA is then purified as described above.

Materials

- Blunt-ended, double-stranded radiolabeled cDNA (UNIT 5.5)
- 2× methylase buffer (Table 3.1.3)
- 50× *S*-adenosylmethionine (SAM), freshly prepared (or from New England Biolabs with order of methylase)
- EcoRI* methylase (New England Biolabs; UNIT 3.1)
- TE buffer (APPENDIX 2)
- Buffered phenol (UNIT 2.1)
- Diethyl ether
- 7.5 M ammonium acetate
- 95% and 70% ethanol, ice-cold
- 10× T4 DNA ligase buffer (UNIT 3.4) containing 5 mM ATP
- 1 μg/μl phosphorylated *EcoRI* linkers, 8- or 10-mers (Collaborative Research)
- T4 DNA ligase (measured in cohesive-end units; New England Biolabs; UNIT 3.14)
- EcoRI* restriction endonuclease and 10× buffer (UNIT 3.1)
- 10× loading buffer without Xylene Cyanol (UNIT 2.5A)
- CL-4B column buffer
- Agarose, electrophoresis-grade
- TBE electrophoresis buffer (APPENDIX 3)
- Ethidium bromide solution (UNIT 2.5A)
- DNA molecular weight markers (UNIT 2.5A)
- 10 mg/ml tRNA
- 3 M sodium acetate
- 65°C water bath
- 5-ml CL-4B column (see support protocol)

Construction of
Recombinant
DNA Libraries

5.6.1

Additional reagents and equipment for cDNA synthesis (*UNIT 5.5*), quantitation of DNA (*APPENDIX 3*), agarose gel electrophoresis (*UNIT 2.5A*), and fragment purification (*UNIT 2.6*)

NOTE: See discussion of preparation of buffer stocks in reagents and solutions section.

Methylate blunt-ended, double-stranded cDNA

1. Dissolve blunt-ended, double-stranded cDNA pellet from *UNIT 5.5* in 23 μ l water and add in the following order (50 μ l final volume):

25 μ l 2 \times methylase buffer (1 \times final)
1 μ l 50 \times SAM (20 μ g/ml final).

Mix gently by pipetting up and down with a pipettor. Add 1 μ l (20 U) *EcoRI* methylase to 400 U/ml final, mix as above, and incubate 2 hr at 37°C.

2. Add 150 μ l TE buffer and extract with 200 μ l buffered phenol beginning with the vortexing as in step 5 of the cDNA synthesis protocol in *UNIT 5.5*. Back extract the phenol phase with 100 μ l TE buffer and pool the aqueous phases as described in step 6 of the cDNA synthesis protocol in *UNIT 5.5*.

The cDNA, if prepared as described in the cDNA synthesis protocol, may be followed at all stages with a hand-held radiation monitor.

3. Extract 300 μ l of aqueous phase twice with 1 ml diethyl ether as described in step 7 of the cDNA synthesis protocol in *UNIT 5.5*.
4. Ethanol precipitate with 125 μ l of 7.5 M ammonium acetate and 950 μ l of 95% ethanol, then wash with ice-cold 70% ethanol as described in steps 8 and 9 of the cDNA synthesis protocol in *UNIT 5.5*.

Ligate EcoRI linkers

5. Dissolve DNA in 23 μ l water and add in the following order (30 μ l final volume):

3 μ l 10 \times T4 DNA ligase buffer containing 5 mM ATP (1 \times and
0.5 mM ATP final)
2 μ l 1 μ g/ μ l phosphorylated *EcoRI* linkers (67 μ g/ml final).

Mix gently by pipetting up and down with a pipettor. Add 2 μ l (800 U) T4 DNA ligase to 27,000 U/ml final, mix as above, and incubate overnight at 4°C.

6. Microcentrifuge the ligation briefly and place tube in a 65°C water bath 10 min to inactivate the ligase.

Digest with EcoRI

7. Place tube on ice 2 min, then add in the following order:

95 μ l H₂O
15 μ l 10 \times *EcoRI* buffer (1 \times final).

Mix gently by pipetting up and down with a pipettor. Add 10 μ l (200 U) *EcoRI* restriction endonuclease to 1300 U/ml final, mix as above, and incubate 4 hr at 37°C.

During incubation, prepare the CL-4B column described in the support protocol.

8. Add an additional 3 μ l (60 U) *EcoRI* restriction endonuclease to the cDNA, mix, and incubate another hour at 37°C to ensure complete digestion of the linkers.
9. Place the tube containing the reaction mixture in a 65°C water bath 10 min to inactivate the endonuclease.

Remove excess linkers

10. Add 2 μl of 10 \times loading buffer without Xylene Cyanol to the reaction and load the cDNA onto a 5-ml CL-4B column prepared in a 5-ml disposable plastic pipet.
11. Allow the loaded sample to enter the column just until the top of the gel becomes dry. Fill the pipet with CL-4B column buffer and allow the column to flow by gravity, collecting $\sim 200\text{-}\mu\text{l}$ fractions manually. Follow the cDNA with a hand-held radiation monitor; the bromphenol blue indicates the position of digested linkers. Stop collecting fractions after the main peak of counts has eluted and before the dye begins to elute.
12. Count 2- μl aliquots of each fraction and plot the results—the elution profile should appear similar to what is shown in Figure 5.6.1.
13. Pool the first one-half of the peak (save the rest as an ethanol precipitate, just in case); add 2.5 vol ethanol (using two or three microcentrifuge tubes as necessary), mix, and place 15 min on dry ice.

The column buffer has sufficient NaCl for precipitation and no more should be added. If the cDNA is not to be ligated or further size-fractionated immediately, store it as an ethanol precipitate.

14. Remove the tubes from dry ice, let thaw, and microcentrifuge 10 min at full speed, 4°C; remove most of the supernatant, fill the tubes with ice-cold 70% ethanol, and microcentrifuge again. Remove most of the supernatant and dry the pellets under vacuum. Resuspend pellets in a total of 50 μl TE buffer.
15. Determine the cDNA concentration. Determine the amount of ^{32}P in 1 μl by scintillation counting and fractionate 2.5 μl on a 1% agarose minigel (see UNIT 2.5) to check the average cDNA size by ethidium fluorescence or autoradiography. A significant fraction of the double-stranded cDNA should be larger than 1.5 kb. If the double-stranded cDNA is to be cloned with no further size fractionation, proceed to ligation protocols in UNIT 5.8; otherwise, continue with step 16.

Approximately 50% to 70% of the starting radioactivity (1 to 3 μg of cDNA) should be recovered and most of the cDNA should be >1.5 kb.

Since only 50 to 100 ng of cDNA are required to produce a full complexity library with a phage vector, it is recommended that a library be produced at this stage in any event. A

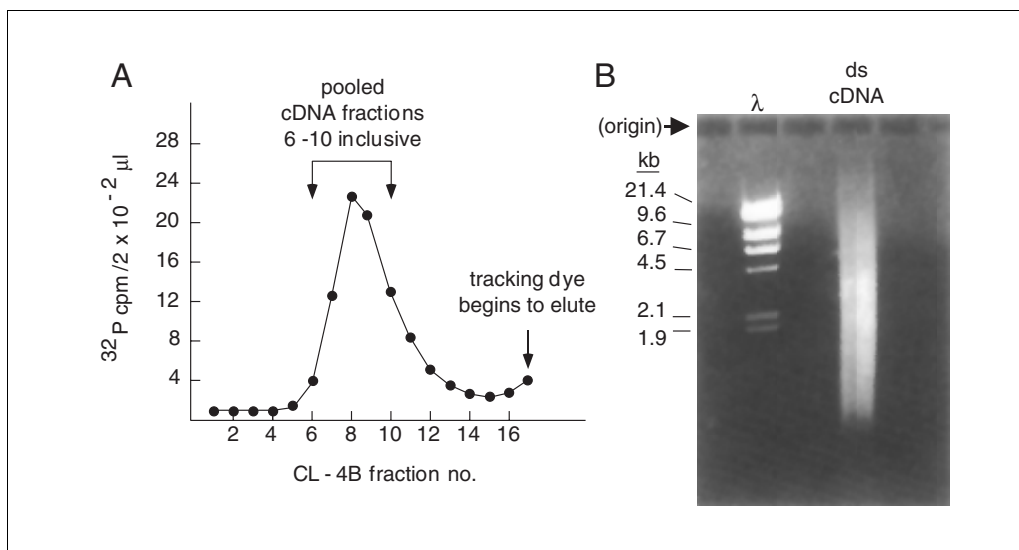


Figure 5.6.1 Fractionation of *Eco*RI-digested cDNA by (A) Sepharose CL-4B chromatography and (B) agarose gel electrophoresis.

library may be stored for years and may be useful in the future. For example, sequences related to the gene of interest may be identified that were excluded from a size-fractionated library.

Size-select the cDNA to obtain long inserts

16. Pour a 0.8% TBE agarose minigel—the gel should be thick enough such that all of the cDNA will fit into a single well. Rinse the gel box, tray, and comb thoroughly, and use fresh TBE electrophoresis buffer.

High-quality, nuclease-free agarose that does not inhibit ligation is essential. Most commercial agarose advertised as molecular biology grade is adequate. The agarose may be checked by first carrying an EcoRI fragment of a plasmid through the procedure and comparing its cloning efficiency in the cDNA vector, expressed as recombinants/ng insert, to the cloning efficiency of the same fragment prior to fractionation. Wash the gel box thoroughly afterward!

17. Add 10× loading buffer to 1× final to the cDNA and load it into a well near the center. Load DNA molecular weight standards (e.g., an HindIII digest of λ phage) two wells away from the sample. Electrophorese at 70 V until adequate resolution is achieved as determined by ethidium bromide fluorescence, usually 1 to 4 hr.

Be sure not to use standards with EcoRI ends.

18. Elute double-stranded cDNA of the desired size as estimated by comparison with the comigrated standards.

λgt10 and λgt11 have a maximum insert size of 7 kb, so collecting cDNA larger than this won't be useful unless a plasmid vector or a phage vector such as Charon 4A or EMBL 4 will be used.

19. Add 10 mg/ml tRNA to 20 μg/ml final, 1/10 vol of 3 M sodium acetate, pH 5.2 (APPENDIX 2), and 2.5 vol ice-cold 95% ethanol and place 15 min on dry ice. Microcentrifuge, wash and dry as in step 14, and resuspend pellet in 20 μl TE buffer.

Ethanol precipitation also extracts the ethidium from the DNA.

20. Determine radioactivity in 1 μl using a fluor and scintillation counter and then calculate the recovery of double-stranded cDNA (see commentary). Proceed to library construction protocols in UNIT 5.8.

ALTERNATE PROTOCOL

LIGATION OF BstXI SYNTHETIC ADAPTERS

Blunt-ended, double-stranded cDNA is ligated to phosphorylated *Bst*XI adapters and then purified as described in the basic protocol. Alternatively, *Eco*RI or *Eco*RI-*Not*I adapters may be used for cDNA to be cloned in vectors with the *Eco*RI site (Fig. 5.6.2). This protocol is simpler than that for linkers because the methylation and restriction digestion steps are unnecessary.

Additional Materials

*Bst*XI adapters (UNIT 2.11; Invitrogen), *Eco*RI adapters (New England Biolabs), or *Eco*RI-*Not*I adapters (Invitrogen)

1. Dissolve blunt-ended, double-stranded cDNA pellet in 23 μl water and add in the following order (30 μl final volume):
 - 3 μl 10× T4 DNA ligase buffer containing 5 mM ATP (1× and 0.5 mM ATP final)
 - 2 μl 1 μg/μl *Eco*RI, *Eco*RI-*Not*I, or phosphorylated *Bst*XI adapters (67 μg/ml final).

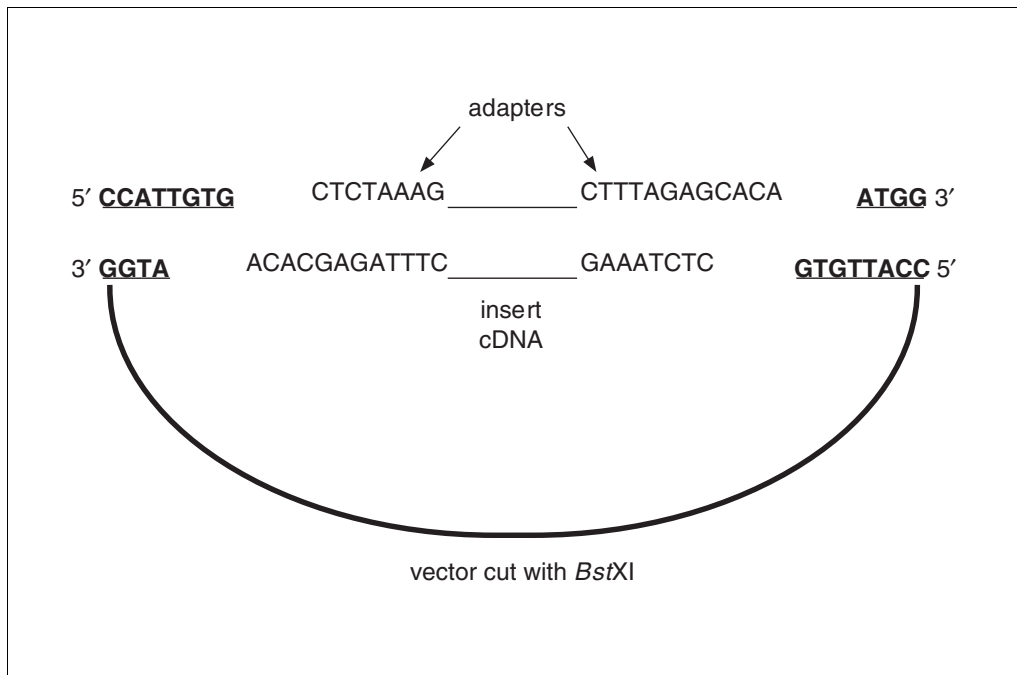


Figure 5.6.2 Noncomplementary adapter strategy. The insert and vector ends are compatible with each other but cannot self-ligate because the cohesive ends are not self-complementary. The *Bst*XI sites (underlined and bold) are not regenerated.

Mix gently by pipetting up and down with a pipettor. Add 2 μ l (800 U) T4 DNA ligase to 27,000 U/ml final, mix as above, and incubate overnight at 4°C.

It is helpful to use 32 P-labeled cDNA to follow the DNA on the subsequent CL-4B column. If the cDNA is not 32 P-labeled, it may be labeled at this step by using [32 P]labeled adapters, prepared with [γ - 32 P]ATP as in the T4 polynucleotide kinase exchange reaction (UNIT 3.10). Alternatively, if the adapters are not yet phosphorylated, they may be labeled with [γ - 32 P]ATP as in the T4 polynucleotide kinase forward reaction (UNIT 3.10).

2. Add 100 μ l TE buffer and remove excess adapters as in steps 10 to 15 of the basic protocol. If desired, size-select the cDNA as in steps 16 to 20 of the basic protocol. Resuspend purified cDNA pellet (obtained from either the CL-4B column or the gel) in 10 to 15 μ l TE buffer. Proceed to library construction protocols in UNIT 5.8.

Because adapter dimers formed during the ligation reaction will clone into the vector very efficiently, removal of the excess adapters is essential.

PREPARATION OF A CL-4B COLUMN

The CL-4B column (Fig. 5.6.3) effectively removes linkers or adapters that would otherwise interfere in subsequent cloning steps; it also allows selection of cDNA \geq 350 bp (see basic and alternate protocols). The column may be prepared while the cDNA is being digested with *Eco*RI, as described in steps 7 and 8 of the basic protocol.

Additional Materials

- Preswollen Sepharose CL-4B (Pharmacia), 4°C
- CL-4B column buffer
- Silanized glass wool
- Plastic tubing (new) with clamp

SUPPORT PROTOCOL

Construction of Recombinant DNA Libraries

5.6.5

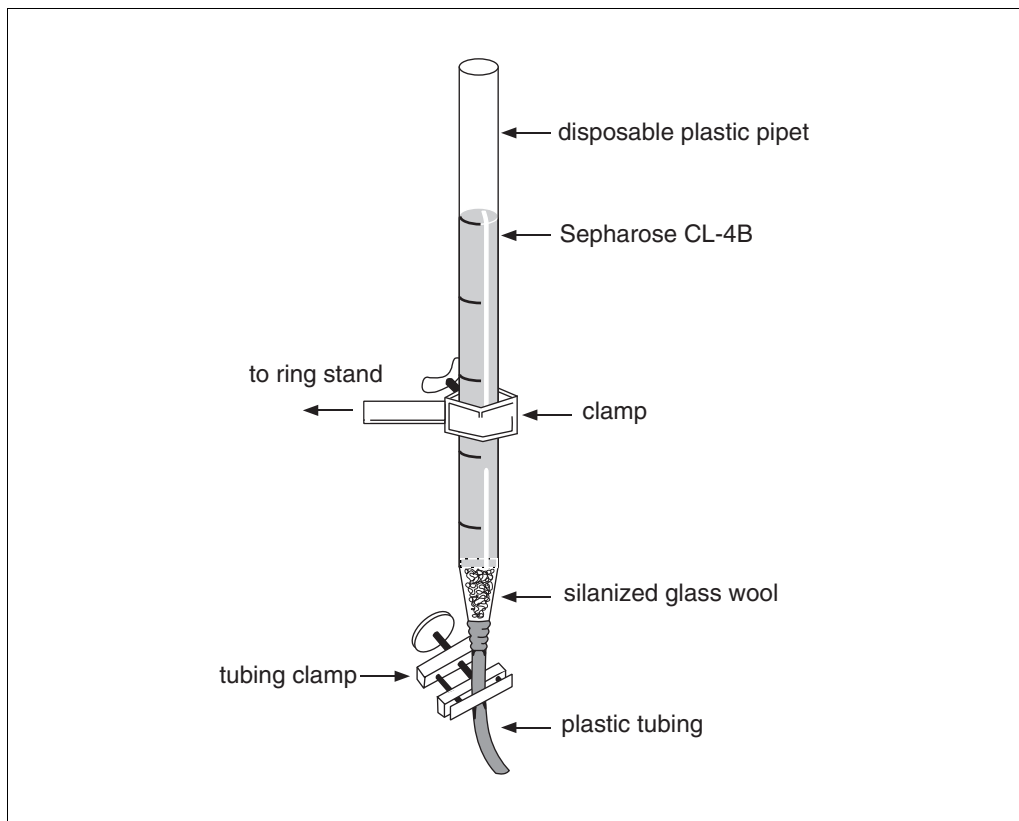


Figure 5.6.3 CL-4B column used for removal of *Eco*RI-digested linkers and selection of cDNA ≥ 350 bp.

1. Transfer 10 ml of preswollen Sepharose CL-4B to a 50-ml polypropylene tube and fill the tube with CL-4B column buffer. Mix by inverting several times and let the Sepharose CL-4B settle by gravity for 10 to 15 min. Aspirate the buffer above the settled gel, removing also the unsettled “fines.”
2. Fill the tube two times with CL-4B column buffer—allow the Sepharose CL-4B to settle each time and remove the fines as in step 1.
3. Add 10 ml CL-4B column buffer and mix by inverting several times. Incubate the tube 10 min at 37°C, then proceed at room temperature.

Outgassing may occur if the column is poured cold. The bubbles thus formed in the gel will interfere with the chromatography.

4. Break off the top of the 5-ml plastic pipet. Wearing gloves, use the 1-ml pipet to push a small piece (3- to 4-mm³) of silanized glass wool down to the tip of the 5-ml pipet. Push a 3-cm length of plastic tubing firmly onto the tip of the 5-ml pipet. Clamp the tubing and attach the column to the ring stand as shown in Figure 5.6.3.
5. With a pipet, carefully fill the column with the gel slurry from step 3. After a few minutes, release the clamp on the tubing and allow the column to flow. Periodically add more slurry to the column as the level drops until the volume of packed gel in the column is at the 5-ml mark.
6. Allow the level of buffer in the column to drop until it is just above the level of the gel and clamp the tubing to stop the flow. The column is ready to be loaded (see basic or alternate protocols).

REAGENTS AND SOLUTIONS

Buffer stock solutions

Prepare ≥ 10 ml of each of the following stock solutions. Use autoclaved water and pass each solution through a sterile 0.45- μm filter. Store at room temperature unless otherwise indicated.

1 M Tris·Cl, pH 8.0 and pH 7.5

0.5 M EDTA, pH 8.0

3 M sodium acetate, pH 5.2

5 M NaCl (prepare 100 ml)

1 M MgCl_2

7.5 M ammonium acetate

20% *N*-lauroylsarcosine (Sarkosyl)

1 M DTT; store at -20°C in tightly capped tube

0.1 M ATP, pH 7.0 (prepare 1.0 ml); neutralize as described in *UNIT 3.4*;
store at -20°C

From these stock solutions, prepare the following buffers, which should be checked for nuclease activity as described in *UNIT 5.5*. Enzyme buffers should be frozen in 200- μl aliquots at -80°C in screw-cap microcentrifuge tubes. Enzyme buffers prepared and stored as described will last for years.

Several of these solutions are routine and may be already available. Nonetheless, to help ensure success, it is best to prepare separate stocks for critical applications such as library preparation.

CL-4B column buffer, 500 ml

5 ml 1 M Tris·Cl, pH 8.0

60 ml 5 M NaCl

1 ml 0.5 M EDTA, pH 8.0

2.5 ml 20% Sarkosyl

431.5 ml H_2O

Filter sterilize and store at room temperature

50 \times S-adenosylmethionine (SAM)

1 mg SAM

1.0 ml 50 \times SAM dilution buffer (see below)

Prepare fresh just prior to use. Store dry SAM at -80°C for no longer than 2 months.

50 \times SAM dilution buffer, 7 ml

330 μl 3 M sodium acetate, pH 5.2

6.67 ml H_2O

Store in 1-ml aliquots

Silanized glass wool

Submerge the glass wool in a 1:100 dilution of a silanizing agent such as Prosil 28 (VWR) for 15 sec with shaking. Rinse the glass wool extensively with distilled H_2O . Autoclave the glass wool for 10 min and store at room temperature.

COMMENTARY

Background Information

The most common method currently employed to create compatible ends on cDNA prior to cloning is the attachment of synthetic linkers (basic protocol). This has for the most part superseded homopolymeric tailing (Maniatis et al., 1982), since linkerage is relatively efficient, and its use eliminates the need for the sometimes tricky procedure of titrating the tailing reaction conditions. Other methods, such as the sequential ligation of two different linkers (Maniatis et al., 1982), are now rarely used. The linked cDNA may be cloned into a plasmid or a phage vector. The use of synthetic adapters that ligate at high efficiency instead of linkers eliminates the methylation and restriction digestion steps (alternate protocol). The recently developed noncomplementary adapter strategy (see below) may ultimately supersede all the above methods as vectors with appropriate sites become available.

Size fractionation of cDNA is preferable to size fractionation of the initial mRNA primarily because DNA is considerably less susceptible to degradation than is RNA. In addition, the small fragments that are generated during the cDNA synthesis process (e.g., due to contaminating nucleases) are removed. These would otherwise be preferentially inserted and decrease the yield of long cDNA clones in the library. RNA enrichment procedures were undertaken in the past because cloning efficiencies then were not high enough to ensure the representation of very rare mRNAs in the library. Currently, the most common reason for using mRNA fractionation or enrichment is in preparation of subtracted libraries (UNIT 5.8; Hedrick et al., 1984).

The strategy of noncomplementary adapters (Fig. 5.6.2), developed by Brian Seed and co-workers (Seed, 1987), overcomes the loss of library complexity due to self-ligation of the cDNA inserts or the vector. This enables a high efficiency for the ligation of the vector to the insert since there are no competing reactions and the vector ends are similarly noncomplementary and do not self-ligate. The much higher yield of desirable ligation products ultimately results in a greater number of clones in the library. Furthermore, the occasional "scrambles" produced as a consequence of two unrelated cDNAs ligating together during the vector ligation step are eliminated; however, "scrambles" at the stage of adapter ligation

remain possible, though infrequent.

Adapters are favored over linkers in general because the methylation and restriction digestion steps are bypassed, thus simplifying the procedure. Adapters are required in the alternate protocol procedure to ensure that all of the cDNA will have the proper, noncomplementary "sticky ends." The use of *EcoRI-NotI* adapters for cloning with *EcoRI* compatible vectors introduces the rare *NotI* site at both ends of the insert, allowing the insert to be cut out of the vector as one fragment, which is not always possible with *EcoRI* linkers if there are internal *EcoRI* sites present.

The development of noncomplementary adapters has enabled the production of high-complexity cDNA libraries in multifunctional plasmid vectors such as CDM8 (Seed, 1987). These vectors permit library screening by functional expression in eukaryotic cells and production of single-stranded DNA for mutagenesis or subtraction, in addition to conventional hybridization methods. For some specialized applications and for selectable markers other than the *supF* present in CDM8, other vectors that employ the noncomplementary linker strategy (and use the same *BstXI* sequence) include several available from Invitrogen with different antibiotic resistances and eukaryotic selectable markers. Other available vectors include AprM8, an ampicillin-resistant version of CDM8 (L.B. Klickstein, unpublished results), and retroviral vectors in the pBabe series (Morgenstern and Land, 1990).

Critical Parameters

In order to maximize the length and cloning efficiency of the cDNA, it is essential that contamination by endo- and exonucleases be avoided. EDTA, an inhibitor of most nucleases, should be present whenever feasible. Reagents, solutions, and enzymes should be of the highest quality obtainable.

The SAM reagent in the methylation step is very unstable, must be freshly dissolved prior to use, and should not be kept >2 months, even if stored dry at -80°C and never thawed. Stabilized SAM is supplied free of charge with methylases purchased from New England Biolabs.

Methylation conditions may be checked by methylating λ DNA under the conditions described in step 1 of the basic protocol, digesting methylated and unmethylated DNA with

EcoRI, and comparing the two samples by agarose gel electrophoresis. Methylated λ DNA should be protected from *EcoRI* digestion.

Digested linkers must be completely removed. Because of their small size, even as a small weight percentage of the total DNA, they will comprise a large mole fraction; thus, if not completely removed, most clones in the library will contain only linkers. Some investigators remove digested linkers by three sequential precipitations from 2 M ammonium acetate. However, column chromatography provides better resolution of the cDNA from the linkers. The CL-4B column also removes small double-stranded cDNA molecules. Purification of cDNA by gel electrophoresis after ligation of the linkers or adapters may be performed as an alternative to the CL-4B column, but the yield usually is not as high and the separation of linkers or adapters from cDNA is not as complete.

If *noncomplementary* adapters are employed, both 5' ends should be phosphorylated for optimum results. If *complementary* adapters are used, only the blunt end should have a 5' phosphate. The other end is phosphorylated after ligation to the cDNA.

Because *BstXI* produces a cohesive end with a sequence not specified by the restriction site sequence, the adapters and vector must correspond—e.g., the cohesive ends of the adapters used for the cDNA must be complementary to that of the vector. The vectors CDM8, AprM8, pCDNAI, pCDNAII, and the pBabe series all use the same sticky ends (Fig. 5.6.2).

Troubleshooting

Methylation and addition of linkers to double-stranded DNA are important to the successful construction of a cDNA library, but cannot be evaluated immediately. Unmethylated cDNA will clone efficiently and the problem will only be detected if the isolated inserts all end at an internal *EcoRI* site. This possibility may be minimized by prior evaluation of the quality of the SAM used in the methylation reaction (see reagents and solutions), or by use of the stabilized SAM that accompanies New England Biolab's *EcoRI* methylase.

Incomplete addition of linkers or adapters will be detected as a poor cloning efficiency in the next step of ligating the inserts to the vector. The problem may be any of the following: (1) cDNA was not blunted properly; (2) linkers or adapters were not kinased well or not annealed; (3) multiple linkers were not cut off with

EcoRI; (4) ligation reaction did not work. However, the problem is usually improperly blunted cDNA or linkers that do not ligate well. Evaluate the linker cDNA (after the CL-4B column) by ligating 5% of the sample with no vector and running the reaction on a 1% agarose minigel next to an equal amount of unligated cDNA. Properly linker cDNA should significantly increase in size as determined by ethidium bromide staining or autoradiography of the dried gel. Poorly linker cDNA may often be salvaged by repeating the blunting and linking steps with fresh reagents.

Inadequate separation of linkers or adapters from cDNA will be detected as cloning efficiency that is too high and as clones with no inserts detectable by gel electrophoresis (white plaques with no inserts in λ gt11 or plaques on C600hflA with λ gt10 that have no insert). The remainder of this cDNA may be salvaged by repeating the CL-4B column chromatography.

Anticipated Results

Approximately 50% to 70% of the starting radioactivity present in the blunt-ended, double-stranded cDNA should be collected after the CL-4B column. This typically represents 1 to 3 μ g of cDNA. In a human tonsil library prepared in λ gt11 where the inserts were thoroughly evaluated, the mean insert size was 1.4 kb and actin clones represented 0.34% of all recombinants. Three-quarters of the actin cDNAs were nearly full length (actin mRNA = 2.1 kb).

Time Considerations

The methylation can be done on the same day as the blunting step from the cDNA synthesis protocol. The linker ligation is then set up overnight. *EcoRI* digestion and CL-4B chromatography are performed the following day, and the cDNA is either ligated to the vector overnight (UNIT 5.8) or stored as an ethanol precipitate overnight and further size selected by agarose gel electrophoresis the next day. At any ethanol precipitation in the procedure, the cDNA may be stored for several days as an ethanol precipitate.

The use of adapters rather than linkers eliminates the need for the methylation and linker digestion steps. The adapted cDNA is immediately loaded onto the CL-4B column after the adapter ligation and is used directly from the column after ethanol precipitation. The subsequent ligation of the adapted cDNA and the vector is performed overnight, requiring the same amount of time with adapters as with linkers.

Literature Cited

- Hedrick, S.M., Cohen, D.I., Nielsen, E.A., and Davis, M.M. 1984. Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. *Nature (Lond.)* 308:149-153.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. 1982. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Morgenstern, J.P. and Land, H. 1990. Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucl. Acids Res.* 18:3587-3596.
- Seed, B. 1987. An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. *Nature (Lond.)* 329:840-842.

Key References

- Neve, R.L., Hanis, P., Kosik, K.S., Kurnit, D.M., and Donlon, T.A. 1986. Identification of the gene for the human microtubule-associated protein tau and chromosomal localization of the genes for tau and microtubule-associated protein 2. *Brain Res.* 387:271-80.
- Seed, B. 1987. See above.

Contributed by Lloyd B. Klickstein
Brigham and Women's Hospital
Boston, Massachusetts

Rachael L. Neve
McLean Hospital
Belmont, Massachusetts

PRODUCTION OF GENOMIC DNA AND cDNA LIBRARIES

In this section the insert DNA (prepared in *UNIT 5.2* or *5.3*) is ligated to vector DNA and then the ligated DNA is propagated in *Escherichia coli*. The technical problems associated with joining DNA molecules together with ligase are described in *UNIT 3.14*. The creation of genomic and complete or subtracted cDNA libraries requires the production of large numbers of recombinant clones and this causes further complications. To recover large numbers of recombinant clones without using large amounts of packaging mixture, vector, or insert DNA, the optimal ratio of insert DNA to vector DNA must be determined prior to the production of the library. Without this optimization, the library will frequently not be large enough to contain the entire genome.

After ligation of vector DNA to insert DNA the ligated DNA must be introduced into *E. coli*. The introduction of bacteriophage or cosmid DNA into *E. coli* is carried out most efficiently by packaging the DNA into bacteriophage particles then allowing these bacteriophage particles to infect *E. coli* (see *UNIT 1.11* for a detailed description of this process). Unfortunately, plasmid DNAs cannot be packaged and thus must be introduced into *E. coli* by bacterial transformation (*UNIT 1.8*). Because of the high efficiency of introducing DNA into *E. coli* by the packaging procedure, we recommend producing genomic and complete or subtracted cDNA libraries using bacteriophage vectors whenever possible.

In this section the protocols required to ligate vector DNA to insert DNA and to introduce these molecules into *E. coli* are described. The protocols are divided into two parts—one for genomic DNA libraries and the other for complete or subtracted cDNA libraries—because the amount of insert DNA required for genomic libraries is usually considerably more than for cDNA libraries.

Production of a Genomic DNA Library

A number of small-scale ligations are performed using a set amount of vector and varying amounts of insert. Test ligations are transformed into bacteria (plasmid vectors) or packaged and plated on host bacteria (λ and cosmid vectors). The number of clones in the different ligations is compared, and the optimum ratio of vector to insert is indicated by the ligation with the most recombinant clones. A large-scale ligation is then set up using this optimum ratio. This protocol employs a bacteriophage vector; however, cosmid or plasmid vectors can be used with minor modifications (see commentary).

Determining the number of clones required to make a genomic DNA library is discussed in the chapter introduction. A library with a base of about 700,000 clones is required for a complete bacteriophage library of mammalian DNA.

Materials

Vector DNA (phage arms, cosmid arms, or linearized plasmid)

Insert fragment (*UNITS 5.3-5.6*)

10× DNA ligase buffer (*UNIT 3.4*)

T4 DNA ligase (measured in cohesive-end units; New England Biolabs; *UNIT 3.14*)

Packaging extract (phage and cosmid) or competent *E. coli* (plasmid)

LB or LB/ampicillin plates (*UNIT 1.1*)

Top agarose containing 10 mM MgSO₄ (*UNIT 1.1*)

Additional reagents and equipment for plating, packaging, and titrating bacteriophage (*UNIT 1.11*)

Contributed by Thomas Quertermous

Current Protocols in Molecular Biology (1988) 5.7.1-5.7.4

Copyright © 2000 by John Wiley & Sons, Inc.

Ligating insert DNA to vector DNA

1. Perform a series of test ligations that bracket equimolar concentrations of vector and insert, i.e., insert/vector molar ratios of 5:1, 2:1, 1:1, 0.5:1, and 0.2:1. (Keep the vector constant and vary the insert DNA.) Since only the relative quantities of recombinants obtained are important, a small quantity of vector can be used for each reaction (40 to 100 ng). A control tube containing only ligated vector (tube no. 6 in Table 5.7.1.) is important to assess the yield of recombinants (see below). Conditions of the ligation reaction are as described in *UNIT 3.14* (1× ligase buffer, 100 U T4 DNA ligase) and should be in a *total volume* of 5 to 10 μ l. See Table 5.7.1 for a sample set of reactions.

Vectors should be chosen as described in UNIT 5.1. Genomic DNA libraries should be constructed only with vectors that provide a selection for insert DNA.

Table 5.7.1 Sample Set of Reactions for Ligating λ Vector to Insert DNA

Reaction component (μ l)	Tube number					
	1	2	3	4	5	6
Vector DNA (100 ng/tube)	1	1	1	1	1	1
Insert DNA (20 kb, 100 ng/ μ l)	2	—	—	—	—	—
Insert DNA (20 kb, 50 ng/ μ l)	—	2	1	0.5	—	—
Insert DNA (20 kb, 10 ng/ μ l)	—	—	—	—	1	—
10× ligation buffer	0.5	0.5	0.5	0.5	0.5	0.5
Water	1	1	2	2.5	2	3
T4 DNA ligase	0.5	0.5	0.5	0.5	0.5	0.5

2. Package the phage recombinants as described in *UNIT 1.11*. Normally, commercially prepared packaging extract is designed for approximately 1 μ g of λ DNA, so it is possible to divide such an aliquot of packaging extract among several test ligations.
3. Dilute packaged extract to 0.5 ml and plate 1 μ l and 10 μ l of packaged bacteriophage per plate. Grow overnight.
4. Compare the number of plaques on the plates to determine the optimum relative concentrations of arms and insert. The yield of plaques with insert DNA (from ligation tubes 1 to 5) should be at least twice and preferably 5 times the yield of colonies on the control plate (from tube 6) containing vector DNA alone.
5. Prepare a large-scale ligation reaction using the optimum ratio of insert DNA to vector DNA determined in step 4. The amount of DNA is determined by the size of the library that is desired.

See UNIT 5.1 to determine total number of desired recombinants. If the expected number of recombinants that can be made uses too much insert DNA or vector DNA, either the insert DNA or the vector DNA is probably defective. The quality of the vector or insert DNA can be tested by ligating them together and fractionating the ligated DNA on a gel (see UNITS 3.14, 2.5A, & 2.5B) using commercially prepared extracts.

Introducing ligated DNA into *E. coli*

6. Package the ligated DNA. Remember that bacteriophage packaging extracts can be saturated by too much DNA.

Because producing extracts that efficiently package DNA can be quite difficult, we recommend the use of commercially available packaging kits (see UNIT 1.11).

The number of recombinants that can be obtained with 1 μg λ DNA from large-scale ligation will depend on the efficiency of the packaging extract or competent cells, but can be as high as 10^8 for phage vectors. For cloning from size-fractionated DNA, this represents a vast excess of recombinants. It is frequently possible to obtain over 10^6 recombinant phage from the test ligations, and this is often enough recombinants to clone a single-copy gene. Thus, it is useful to save the test ligations.

7. Package and titer the large-scale ligation and store at 4°C.

Plate the bacteriophage for amplification or screening as soon as possible after the titer is known, as the titer will drop 3- to 10-fold within 48 hr of packaging.

COMMENTARY

Background Information

The theory of the ligation reaction is presented in UNIT 3.16 and is well outlined by Williams and Blattner (1980). From these theoretical considerations it is possible to predict the optimum concentration of vector and insert. However, because of the difficulty of estimating small quantities of DNA and the inability of some ends to ligate, it is helpful to conduct a series of test ligations as outlined here to determine the ratio of reagents that will result in the maximum number of clones. The theory of ligation reactions suggests that the optimum concentration for a ligation reaction varies for different sized molecules. However, the conditions suggested here are optimal for most vectors and inserts used for making genomic and cDNA libraries.

The production of a cosmid library is identical to the procedure described here except that a cosmid vector is used rather than a bacteriophage vector. Different amounts of cosmid vector are used because cosmids are smaller (about 5 kb) than bacteriophages (50 kb). Thus the weight of vector DNA should be reduced 10-fold.

Production of a plasmid library involves the same procedure except that ligated DNA is introduced into *E. coli* cells by transformation (see UNIT 1.8) rather than by packaging.

Critical Parameters

The quality of the insert DNA and vector DNA are critical to the success of these experiments. The most frequent cause of failure is that one of these DNAs has defective ends and thus the molecules are not able to ligate efficiently.

Quality of packaging extracts is also critical. Normally, commercial packaging extracts are of high quality and yield large numbers of bacteriophage. However, an occasional bad lot

of packaging extract is obtained.

Troubleshooting

A large number of clones in all the test ligations, or a large number of clones with extremely small quantities of insert, may indicate a high background of nonrecombinant clones. Background is determined by ligating the vector in the absence of insert. A large number of clones in this control ligation indicates inadequate dephosphorylation, inadequate purification from stuffer, or possibly that the bacterial strain being used for genetic selection is not correct.

If few recombinants are obtained, there are a number of possible explanations. The packaging extract should be evaluated by appropriate control experiments. Usually, the commercially available packaging extract kits contain a suitable control fragment. At the same time a test ligation can be run on an agarose gel and compared to the unligated vector and insert. If both the insert and vector DNA appear intact (i.e., not degraded), either the sticky ends of the vector or insert have been damaged or one of the DNA samples (usually the insert) contains an inhibitor of the ligase reaction. Clonability of these two reagents can be evaluated by determining the ability of insert or vector DNA to ligate to itself or to an appropriate control. Inhibiting contaminants can sometimes be removed by phenol/ chloroform extraction followed by ethanol precipitation, or by passing the DNA over an Elutip column (UNIT 2.6).

Anticipated Results

The number of clones obtained will depend on a number of factors including the efficiency of ligation, transformation, and packaging. At best, in a λ vector, over 10^8 recombinants per μg of insert DNA can be achieved. Cloning into

plasmid usually produces fewer recombinants per quantity of insert because transformation into competent *E. coli* is 10 to 100 times less efficient than the in vitro packaging of λ phage. In a plasmid, it may be necessary to use 1 to 2 μg insert DNA to achieve 10^6 recombinant clones. Cosmid clones usually are obtained at 10^5 to 10^6 colonies per μg insert DNA.

Time Considerations

Test ligations are incubated overnight at 13°C (alternatively, incubate 2 hr at room temperature). The ligation reactions are packaged and then titered on the second day. Results are available late on the second day or the morning of the third day.

Literature Cited

- Kaiser, K. and Murray, N.E. 1984. The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries. *In* DNA Cloning, Vol. 1: A Practical Approach (D. Glover, ed.) pp. 1-47. IRL Press, Oxford.
- Rodriguez, R.L. and Tait, R.C. 1983. Recombinant Techniques: An Introduction. Addison-Wesley, Reading, MA.
- Williams, B.G. and Blattner, F.R. 1980. Bacteriophage lambda vectors for DNA cloning. *In* Genetic Engineering, Vol. 2 (J.K. Setlow and A. Mullander, eds.) p. 201. Plenum, NY.
- Describes the ligation reaction as applied to cloning with λ vectors.*

Contributed by Thomas Quertermous
Massachusetts General Hospital
Boston, Massachusetts

Production of a Complete cDNA Library

UNIT 5.8A

A complete cDNA library is one that contains at least one cDNA clone representing each mRNA in the cell. There are about 34,000 different types of mRNA in a mammalian cell and about 500,000 mRNA molecules per cell. For any cDNA library there is a probability, P , that each mRNA will be represented at least once. A useful method for calculating the required base of a cDNA library is to calculate the probability, q , that a single cDNA clone will not be found in the library after choosing, at random, N clones:

$$q = 1 - P = [1 - (n/T)]^N$$

where n is the number of molecules of the rarest mRNA in a cell and T is the total number of mRNA molecules in a cell. The desired base, B , is the number of clones that should be screened to achieve a 99% probability that a cDNA clone will exist in the library, defined as:

$$q = 1 - 0.99 = 0.01 = [1 - (n/T)]^B$$

where n is 8 and T is 562,800 (for one mammalian cell), and the base B , is 324,000. In general, if 500,000 to 1,000,000 independent cDNA clones are obtained, the library should contain at least one copy of every mRNA.

Either a phage or plasmid vector can be used in constructing a complete cDNA library (first basic and alternate protocols). Procedures for evaluating these cDNA libraries are described in the first support protocol. If necessary, the quality of the λ arms used in constructing complete phage cDNA libraries can be evaluated as described in the second support protocol.

A subtracted DNA library is useful when a specific DNA or antibody probe for the desired clone is unavailable. A method for constructing a subtracted cDNA library is presented in UNIT 25B.1.

LIGATION AND TRANSFECTION FOR A COMPLETE λ PHAGE LIBRARY

BASIC
PROTOCOL

Double-stranded cDNA with *EcoRI* ends is ligated to a 2-fold molar excess of phage vector DNA that has been digested with *EcoRI* and phosphatased to prevent ligation without a cDNA insert. The ligation mixture is packaged into infectious phage particles in vitro and transformed into host bacteria. Currently C600*hflA* is used to propagate λ gt10-based libraries and Y1088 is the host for λ gt11 libraries (see Table 1.4.5).

Materials

EcoRI-cut, phosphatased λ gt10, λ gt11, or λ gt11 DNA or other phage vector DNA

Double-stranded cDNA with *EcoRI* ends (UNIT 5.6)

10 \times DNA ligase buffer (UNIT 3.4)

T4 DNA ligase (measured in cohesive-end units; UNIT 3.14)

Cold 10 mM MgSO₄

λ phage packaging extract (Stratagene)

Suspension medium (SM; UNIT 1.11)

Chloroform

E. coli C600*hflA* (λ gt10) or Y1088 (λ gt11)

IPTG (λ gt11 only; UNIT 1.4)

Xgal (λ gt11 only; UNIT 1.4)

Additional reagents and equipment for plating and titering bacteriophage (UNIT 1.11)

Construction of
Recombinant
DNA Libraries

Contributed by Lloyd B. Klickstein

Current Protocols in Molecular Biology (1992) 5.8.1-5.8.8

Copyright © 2000 by John Wiley & Sons, Inc.

5.8.1

Supplement 18

1. Resuspend *Eco*RI-digested, phosphatased λ gt10 or λ gt11 DNA at 1 μ g/ μ l.

*λ gt10 and λ gt11 phosphatased phage arms are available commercially. Alternatively, prepare the phage DNA as described in UNIT 1.13 and digest with *Eco*RI and phosphatase (UNITS 3.1 & 3.10). Evaluate “homemade” phage arms as described in the second support protocol of this unit.*

Both λ gt10 and λ gt11 grow well and milligram quantities of DNA are routinely obtained from a 1-liter culture.

2. To a fresh tube add the following (final volume 19 μ l):

100 ng cDNA with *Eco*RI ends (5 μ g/ml final)
10 μ l (10 μ g) λ gt10 or λ gt11 DNA from step 1 (500 μ g/ml final)
2 μ l of 10 \times ligase buffer (1 \times final)
H₂O to 19 μ l.

Mix by gently pipetting up and down and add 1 μ l (400 U) T4 DNA ligase to 20 U/ μ l final. Mix gently and incubate overnight at 4°C. Inoculate an overnight culture of plating bacteria.

The objective is to maximize the number of recombinants obtained per microgram of cDNA and to minimize both the probability of two unrelated cDNA sequences ligating to each other and the probability of the cDNA forming an intramolecular loop. The 100-fold weight excess of vector used above is a 2-fold molar excess if the rough assumption is made that the average insert size is 1/50 that of the vector. Since the rate of intramolecular ligation is independent of cDNA concentration, the ligation is performed in a relatively small volume to favor intermolecular reactions.

3. Inoculate 50 ml broth with 0.5 ml of the overnight culture of host bacteria. When the OD₆₀₀ = 0.5, pellet the bacteria by centrifugation 2 to 4 hr at 28,000 \times g (3000 rpm in a tabletop centrifuge), 4°C, and resuspend in 5 ml cold 10 mM MgSO₄. Store at 4°C prior to use.
4. While the culture is growing, incubate the ligation for 30 min at –20°C, and thaw at room temperature. Package the DNA mixture with a λ phage packaging extract.

Use enough extracts such that the capacity of each is not saturated—typically 1 to 2 μ g ligated DNA per extract.

Commercially available packaging extracts are recommended. High-efficiency in vitro packaging extracts (>10⁸/ μ g intact phage DNA) are critical to successful preparation of the library as are freshly prepared plating bacteria. A test packaging is not recommended as the components of ligation cannot be recovered if ligation fails.

5. Add SM to the packaging mixtures and pool into a single 5-ml polypropylene tube such that the total volume is 2 ml. Add a few drops of chloroform and shake by hand 3 sec. Place the tube containing the unamplified library on ice.
6. Determine the titer of the library. Use *E. coli* C600*hflA* as the host for a λ gt10-based library or Y1088 as the host for a λ gt11-based library. Include IPTG and Xgal in the case of a λ gt11-based library to determine the fraction of the clones that represents recombinants. While the titer plates are incubating, store the library at 4°C.

In vitro packaged phage are unstable. If stored for more than a day or so, the titer of the stock will drop noticeably. Thus, the library must be amplified at this time or no later than the following day (see UNIT 25B.2).

7. Calculate the cloning efficiency as recombinants/ μ g cDNA.

Approximately 1–5 \times 10⁷ clones/ μ g cDNA are typical. For a λ gt11-based library, at least 90% of the plaques should be clear. The library is ready to be plated for amplification and/or screening.

LIGATION AND TRANSFORMATION FOR A COMPLETE PLASMID LIBRARY

ALTERNATE
PROTOCOL

Double-stranded cDNA with noncomplementary *Bst*XI ends is ligated to an equimolar amount of plasmid vector DNA that has been linearized at the two inverted *Bst*XI sites and purified by gel electrophoresis. The ligation is introduced into bacteria either by transfection of competent cells or by electroporation, and the bacterial cells are plated and grown overnight.

Additional Materials

CDM8 plasmid DNA (Invitrogen) or similar plasmid vector DNA, CsCl-purified *Bst*XI and 10× *Bst*XI buffer (UNIT 3.1)

Low melting temperature agarose for fragment purification (UNIT 2.6)

TE buffer (APPENDIX 2)

Double-stranded cDNA with *Bst*XI ends (UNIT 5.6)

Competent cells: MC1061/P3 (Invitrogen) is required for CDM8; MC1061, DK1, DH1, DH5, or HB101 are suitable for other vectors

LB medium (UNIT 1.1)

150-mm LB agar plates (UNIT 1.1) with appropriate antibiotics (10 µg/ml ampicillin + 7.5 µg/ml tetracycline for CDM8 in MC1061/P3; 50 to 100 µg/ml ampicillin for the other vectors and strains above)

80% glycerol in H₂O

Additional reagents and equipment for plating and growth of *E. coli* (UNITS 1.1-1.4), isolation of DNA from agarose gels (UNIT 2.6), alkaline lysis/CsCl gradient preparation of plasmid DNA (UNIT 1.7), and introduction of plasmid DNA into cells (UNIT 1.8)

Prepare vector DNA

1. Digest CDM8 plasmid DNA with *Bst*XI restriction enzyme by adding to a 500-µl tube (final volume 100 µl):

30 µg CDM8 DNA (300 mg/ml final)

10 µl 10× *Bst*XI buffer

H₂O to 92 µl

8 µl *Bst*XI (60 to 100 U).

Mix by vortexing; microcentrifuge at high speed briefly; incubate 3-4 hr at 55°C.

2. Purify the large vector fragment by electrophoresis through a low-melting-point agarose gel, by electroelution, or using DEAE-cellulose paper. Resuspend DNA at 1 µg/µl in TE buffer.

A purification step is essential, because all vectors employing the noncomplementary adapter strategy have two nearby (inverted) BstXI sites, and the fragment between them (several basepairs in pcDNA1 and pBabeNeo, ~300 bp in CDM8 and AprM8) must be completely removed for optimum cloning efficiency.

Ligate cDNA, transform of competent cells and store library

3. In a 500 µl microcentrifuge tube, add in the following order (final volume 20 µl):

1 µg cDNA with *Bst*XI ends (100 µg/ml final)

5 µl (5 µg) vector DNA from step 2 (500 µg/ml final)

1.5 µl of 10× DNA ligase buffer

H₂O to 19 µl.

Construction of
Recombinant
DNA Libraries

5.8.3

Mix by pipetting up and down with a micropipet and add 1 μl (400 U) T4 DNA ligase to 20 U/ μl final. Mix gently and incubate overnight at 4°C.

The above ligation employs equimolar amounts of vector and insert, assuming an average cDNA size of 1 kb and a vector size of 5 kb. An excess of vector is not required, because neither the vector nor the cDNA is able to self-ligate.

4. Transform 0.75 μl of the ligation mixture into each of twenty 100- μl aliquots of competent cells (as described in UNIT 1.8 or manufacturer's protocol).

High efficiency competent cells are critical to the successful production of a full complexity cDNA library. Whether competent cells are purchased commercially or are made in the laboratory, they should be tested prior to use. Efficiency yields should be $\geq 1 \times 10^8$ colonies per μg supercoiled pUC18 DNA. Electroporation is an excellent alternative if the special apparatus is available; it yields even higher efficiencies and has been employed successfully for cDNA libraries (UNIT 1.8).

5. Pool the transformed cells in a single 50-ml polypropylene tube, grow with shaking 1 hr at 37°C, and plate on 10 to 20 15-cm agar plates containing the appropriate antibiotic(s) (UNIT 1.3). Incubate overnight at 37°C.

The transformed bacteria may be grown in liquid culture, but there is a greater likelihood of altering the representation of the library due to overgrowth of relatively rapidly growing cells. For this reason, growth on agar plates is recommended.

If the library is to be screened by hybridization without amplification, a detergent-free nitrocellulose filter may be placed on the plate prior to plating the bacteria. Replica filters can then be prepared after incubation (UNIT 6.2)

6. Count the number of colonies in a sector of a plate to determine the library complexity expressed as colonies per microgram cDNA. Add 5 ml LB medium to each plate and gently elute the bacteria with a spreader (UNIT 1.3) to form a dense suspension. Wash each plate with an additional 1 ml LB medium to recover residual bacteria. Pool bacteria in a single sterile tissue culture flask.

Library complexities of 5×10^5 to 2×10^6 colonies per μg input cDNA are typical for libraries prepared with competent cells. Efficiencies 10-fold higher are possible for libraries prepared by electroporation (UNIT 1.8).

7. Take half of the pooled bacteria and perform an alkaline lysis/CsCl gradient purification of the plasmid DNA. This DNA represents the unamplified library and is suitable for transfection of mammalian cells. To the remainder of the pooled bacteria, add 0.2 vol of 80% glycerol and mix thoroughly by inverting the flask. Dispense 1-ml aliquots into sterile 1.5-ml microcentrifuge tubes and freeze at -80°C .

Up to 50 aliquots may be obtained with this procedure. Each aliquot will inoculate 500 ml of liquid culture for preparation of additional DNA.

8. Evaluate the library as described in the first support protocol. Note that the *Bst*XI sites are not reconstituted with the adaptors commonly used, thus flanking restriction sites must be used to cut out the inserts.

EVALUATION OF A cDNA LIBRARY

To ensure that a negative result upon screening is not the result of a technical error during the construction of the library, evaluate the library by (1) screening for an abundant mRNA sequence, such as actin, known to be present; (2) screening with a total mixed cDNA probe; or (3) examining the inserts from twelve or more randomly chosen clones. Evaluation of the library is carried out after amplification (*UNIT 25B.2*) or concurrently with screening if no amplification is done.

An amplified library is stable for years and may be screened thousands of times. Due to the expense and time involved in generating a cDNA library, amplification is recommended. If $>2 \times 10^6$ clones are obtained, some investigators plate $4-5 \times 10^5$ clones for immediate screening and amplify the remainder of the library, against the possibility that the target clones will be underrepresented as a result of the amplification procedure. A cDNA cloning efficiency within the expected range is a good initial indication that all is well; however, further evaluation by at least one of the following methods is necessary because although screening a library is conceptually simple, in practice, identification of a positive clone is not always straightforward. Thus, in the case of an evaluated library, a negative result upon screening implies problems with the probe, the screening method, or the source of the RNA.

Evaluate the library by screening duplicate lifts from a single plating of 10,000 to 50,000 clones (from a plasmid or phage library) with an actin hybridization probe, using the protocol in *UNIT 6.3*. Each duplicating signal on the film is scored as a positive. Table 5.8.1 gives a range of actin cDNAs for various tissues expressed as a percentage of total recombinant clones screened.

Actin is a good choice of probes with which to evaluate the library because its sequence is highly conserved across species and probes are widely available; however, any abundant mRNA known to be present, for which a probe is available, will do. This probe can be labeled by methods described in *UNIT 3.4*.

An alternative to a specific probe is to screen a single plating as above with a cDNA hybridization probe prepared from the same mRNA used to construct the library. Use the hybridization conditions for cloned probes outlined in the steps for screening libraries (*UNIT 6.3*). Approximately 50% to 90% of the clones in the plating should hybridize, depending in part upon the source of the RNA.

A third means of evaluating the library is to pick 12 single recombinant clones, often from the initial titer plate from step 6 of the basic protocol, and carry them through the phage DNA miniprep described in *UNIT 1.13*. Resuspend the miniprep DNA in 25 μ l water and digest 10 μ l with *EcoRI*. Analyze the digest by electrophoresis through a 1.5% agarose gel in TBE buffer. At least 10 of the 12 recombinant clones should have inserts visible by ethidium fluorescence, with an average size of ~ 1 kb or more.

Table 5.8.1 Frequency of Actin cDNA Clones in Various Mammalian cDNA Libraries

Tissue or cell type	Actin-hybridizing clones (% of clones screened)
Liver	0.02 to 0.06
Lymphoid tissue or cells	0.1 to 0.5
Fibroblastoid cell line	1 to 4

PREPARATION OF TEST INSERT FOR EVALUATION
OF PHOSPHATASED PHAGE ARMS

This protocol describes a procedure for testing the quality of λ arms. Bacteriophage DNA is ligated to test insert DNA and the efficiency of the vector is tested by packaging and plating the ligated DNA.

Additional Materials

Genomic DNA

1:1 phenol/chloroform (UNIT 2.1A)

Additional reagents and equipment for restriction endonuclease digestions (UNIT 3.1) and plating bacteriophage libraries (UNIT 6.1)

1. Digest 20 μg genomic DNA to completion with *EcoRI*. Bring to 300 μl with TE buffer.
2. Extract with 300 μl phenol/chloroform, add 30 μl of 3 M sodium acetate and 700 μl of 95% ethanol, and place on dry ice for 15 min to precipitate the DNA.
3. Thaw, microcentrifuge 10 min at high speed, wash the pellet with 70% ethanol, and resuspend in 1.7 ml TE buffer. The test insert DNA concentration is ~ 10 ng/ μl .
4. Ligate 1 μg of arms to 10 μl (10 ng) of test insert in 10 μl as described in step 2 of the first basic protocol. Also ligate 1 μg of arms without a test insert as a control. Package and titer 2 μl of each ligation (UNIT 1.11).
5. To determine the efficiency (no. PFU per microgram test insert), multiply the number of plaques from the packaging mixture by 500. Usable arms will give 1×10^7 to 2×10^8 PFU/ μg test insert. In the case of $\lambda\text{gt}11$ plaques $\leq 10\%$ of these should be blue on an Xgal plate. At least a 100-fold increase in the number of plaque-forming units should be observed in the presence of the test insert, compared with its absence from the ligation mixture.

Perhaps because of a large number of short fragments or because of a greater fraction of fragments with the correct ends, an EcoRI genomic digest gives a higher cloning efficiency than does cDNA.

COMMENTARY**Background Information**

Significant technological advances over the past few years permit the construction of cDNA libraries that closely reflect the size and complexity of the corresponding mRNA. Historically the two most troublesome steps were S1 nuclease digestion of the cDNA and G-C tailing of the cDNA and vector, both of which require laborious adjustment of reaction conditions (see Sambrook et al., 1989). These, for the most part, have been superseded by the RNase H/E. *coli* polymerase I second-strand synthesis method of Okayama and Berg (1982) and the ligation of synthetic linkers (Huynh et al., 1984), respectively. The commercial availability of high-quality enzymes and reagents has made these newer methods quite reliable. The introduction of $\lambda\text{gt}10$ and $\lambda\text{gt}11$ phage vectors (Young and

Davis, 1983) allowed the high efficiencies of in vitro packaging systems to be applied to the preparation of cDNA libraries. It is now nearly routine to generate libraries containing millions of independent recombinant clones, many of which will contain the entire protein-coding region of interest. A limited cDNA library may also be obtained by following the basic protocol for the production of subtracted cDNA libraries (UNIT 25B.2).

Compared with phage libraries, plasmid libraries usually contain fewer clones and are more complicated to screen by hybridization or with antibody. They are most useful when prepared for particular applications, such as functional screening in COS cells with the vector CDM8 (Seed, 1987; UNIT 16.12). One advantage of functional screening is that

clones may be identified when no antibody or hybridization probes are available. A second advantage is that, for the most part, only full-length clones are detected, although some clones may contain residual introns. A disadvantage is that cDNAs encoding large proteins (>140 kDa) are unlikely to be found, as are cDNAs for proteins requiring other subunits for function.

Critical Parameters

Vector DNA, whether phage or plasmid, must be completely digested and phosphatased. A small amount of intact vector will transform the host bacteria very efficiently and contribute to a high background of nonrecombinant clones in the library. Of course, phosphatasing is not necessary for vectors employing the noncomplementary linker strategy. Commercially available phosphatased λ gt10 or λ gt11 arms and *Eco*RI-cut, phosphatased pBR322 or *Bst*XI-cut CDM8 work well in these procedures. If the vector is prepared by the investigator, the steps outlined to evaluate the vector should be followed before it is used. High-efficiency in vitro packaging extracts are crucial to the successful preparation of a high complexity cDNA library. Commercially available extracts work well.

The CDM8 vector requires a bacterial host containing the P3 plasmid, e.g., MC1061/P3. P3 is a low copy number, ~60-kb plasmid containing an origin of replication able to coexist with ColE1. The *supF* gene carried by CDM8 suppresses the amber mutations in the structural sequence for tetracycline and ampicillin carried on P3 and allows expression of the antibiotic resistance phenotype (UNIT 1.4). P3 also carries a functional kanamycin resistance to allow for its selection. If a particular host is necessary for an experiment, the P3 plasmid can be readily recovered by an alkaline lysis miniprep of MC1061/P3 (UNIT 1.6) and transformed into the desired host to enable propagation of CDM8. The *supF* selectable marker minimizes the size of the vector, and may allow isolation of genomic clones by an elegant recombination-based screening method (Seed, 1983). Alternate vectors that utilize a conventional antibiotic resistance and contain the noncomplementary *Bst*XI cloning sites are mentioned above and in UNIT 5.6.

High-efficiency transformation of bacteria with the ligated cDNA is probably the single most important step. Commercial competent cells may be worthwhile, if the investigator has difficulty preparing competent cells $\geq 1 \times 10^8$

CFU/ μ g supercoiled pUC18 DNA. Transformation of bacteria by electroporation is superior to the use of competent cells.

Several methods of evaluating a library are presented; screening for a known abundant clone, such as actin, is probably the best of these. Thorough evaluation of a library is of particular value when a screen turns out negative. If a problem exists in an evaluated library it is usually due to the probe, the screening method, or the source of the RNA.

Troubleshooting

The major problem at this stage is usually too few clones and is often due to a problem in previous steps. The second most likely problem is inefficient packaging extracts or competent cells, which can be tested by packaging or transforming, respectively, with uncut vector. The best diagnostic test at this point is ligation of a test insert to the vector and preparation of a test insert library. For phage or plasmid vectors with an *Eco*RI cloning site, use the *Eco*RI digest of genomic DNA as a test insert (second support protocol). For vectors with a noncomplementary *Bst*XI cloning site, obtain new phosphorylated *Bst*XI adapters and self-ligate 1 μ g in 20 μ l to form linker dimers. Precipitate the DNA by dilution to 300 μ l with dH₂O, add 3 μ l of 1 M MgCl₂, 20 μ g of carrier tRNA and 4 vol ethanol, precipitate 30 min in dry ice and recover the oligonucleotide DNA by thawing, and then microcentrifuging at top speed for 15 min. Resuspend the adapter dimers (the test insert) in 10 μ l TE, and ligate 1 μ l (100 ng) to 1 μ g of vector. Transform *E. coli* (UNIT 1.11) and plate an aliquot to determine the efficiency, which should be $\geq 1 \times 10^5/100$ ng test insert.

Anticipated Results

A complete cDNA library constructed in either phage vector should contain in the range of 10^6 to 10^7 independent recombinants from 100 ng of cDNA. This is sufficiently complex to contain even the rarest mRNA. A cDNA library prepared in a plasmid vector employing the noncomplementary adapter strategy should contain 5×10^5 to 2×10^6 colonies per μ g input cDNA if prepared by transformation of competent cells. Ten-fold higher complexities should be obtained with electroporation.

Time Considerations

The initial ligation of cDNA to the phage arms is an overnight step. The ligation may be packaged or transformed the following day or stored at -20°C over the weekend. Once the

library is packaged in vitro it should be plated within 24 hr for amplification and/or screening since these in vitro packaged phage are unstable and the titer of the library will drop if stored at this stage. An amplified library is stable for years when stored at -80°C in 8% DMSO as outlined in the phage library amplification protocol (UNIT 25B.2), and evaluation and screening may be performed at the investigator's convenience.

For plasmid libraries, the ligation may be stored for several days at -20°C . Transformed competent cells should be plated the same day. As above, an amplified plasmid library may be stored for years at -80°C in glycerol.

Literature Cited

- Huynh, T.V., Young, R.A., and Davis, R.W. 1984. Constructing and screening cDNA libraries in $\lambda\text{gt}10$ and $\lambda\text{gt}11$. *In* DNA Cloning Techniques, Vol. 1: A Practical Approach (D. Glover, ed.) pp. 49-78. IRL Press, Oxford.
- Okayama, H. and Berg, P. 1982. High-efficiency cloning of full-length cDNA. *Mol. Cell. Biol.* 2:161-170.
- Sambrook, J., Maniatis, T., and Fritsch, E.F. 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Seed, B. 1983. Purification of genomic sequence from bacteriophage libraries by recombination and selection in vivo. *Nucl. Acids Res.* 12:5627-5633.
- Seed, B. 1987. An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. *Nature (Lond.)* 329:840-842.
- Young, R.A. and Davis, R.W. 1983. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. U.S.A.* 80:1194-1198.

Key References

Huynh et al., 1984. See above.

Describes production of a complete λ phage library.

Seed, B. 1987. See above.

Describes production and screening of a complete plasmid library.

Contributed by Lloyd B. Klickstein
Brigham and Women's Hospital
Boston, Massachusetts

Construction of Bacterial Artificial Chromosome (BAC/PAC) Libraries

Large-insert genomic libraries are important reagents for physical mapping of large chromosomal regions, for isolation of complete genes including all regulatory sequences, and for use as intermediates in DNA sequencing of entire genomes. There is a need for libraries of large insert clones in which rearrangements and cloning artifacts are minimal. With recent developments in host/vector systems, it is now possible to efficiently create large-insert libraries using low-copy-number bacterial plasmid vectors. Suitable vector systems have been developed using *Escherichia coli* F factor or P1 phage replicons, which are capable of producing large insert clones by bacterial transformation.

Construction of BAC and PAC libraries is detailed first (see Basic Protocol). Two vectors, pCYPAC2 (Ioannou et al., 1994) and pPAC4 (E. Frengen et al., unpub. observ.) have been used for preparing PAC libraries, and a new BAC vector pBACe3.6 (Frenge et al., 1999) has been developed for construction of BAC libraries. All vectors are illustrated in Figure 5.9.1. Support Protocol 1 describes preparation of PAC or BAC vector DNA for cloning by digestion with *Bam*HI or *Eco*RI, simultaneous dephosphorylation with alkaline phosphatase, and subsequent purification through pulsed-field gel electrophoresis (PFGE). For the preparation of high-molecular weight DNA for cloning, Support Protocols 2 and 3 provide procedures for embedding total genomic DNA from lymphocytes or animal tissue cells, respectively, in InCert agarose. Support Protocol 4 details the next steps for the genomic DNA: partial digestion with *Mbo*I or with a combination of *Eco*RI endonuclease and *Eco*RI methylase (including appropriate methods for optimizing the extent of digestion), and subsequent size fractionation by preparative PFGE. Finally, Support Protocol 5 covers the isolation of BAC and PAC plasmid DNA for analyzing clones. Isolation of BAC and PAC plasmid DNA is somewhat more difficult than standard plasmid extraction, because it requires more attention to cell concentration during extraction and resuspension of ethanol-precipitated DNA.

In order to construct, for example, a 10-fold redundant human library with 150-kb average insert size, 200,000 clones must be generated. The *Eco*RI-*Eco*RI ligation efficiency seems to be higher than the *Mbo*I-*Bam*HI ligation. A reasonable transformation efficiency from *Eco*RI-*Eco*RI ligation is ~2,000 colonies per transformation using 2 μ l ligation mixture and 20 μ l electrocompetent cells. In comparison, the efficiency of *Mbo*I-*Bam*HI ligation is ~800 colonies per transformation. Thus, it is necessary to perform at least 100 individual transformations for an *Eco*RI partially digested 10-fold redundant library, and 250 individual transformations for an *Mbo*I partially digested library. In order to perform 100 electroporations, 200 μ l of concentrated ligation mixture is required, thus five individual 250- μ l ligations must be prepared. For 250 transformations, 500 μ l of ligation mixture is needed, so thirteen individual 250- μ l ligations are required. Note that it will be necessary to repeat the Basic Protocol and Support Protocol 4 multiple times until enough clones are generated to complete the library. Steps 24 to 30 of the Basic Protocol are designed to handle the amount of size-fractionated genomic DNA that can reasonably and comfortably be generated from a single preparation as described in Support Protocol 4. It is advisable to prepare at least 20 DNA plugs (20 to 40 μ g/plug) for the entire library construction procedure; up to four plugs may be needed to optimize the partial digestion, and as many as sixteen plugs may be needed for library construction.

CAUTION: Radioactive, biological, and chemical substances require special handling; see APPENDIX 1F & 1H for guidelines.

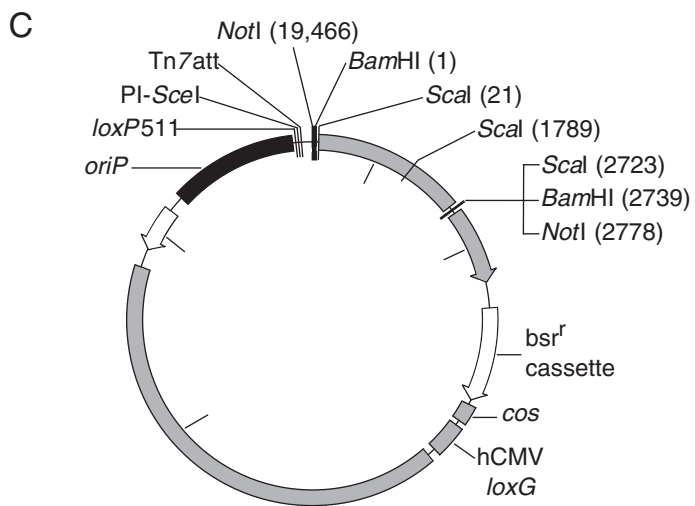
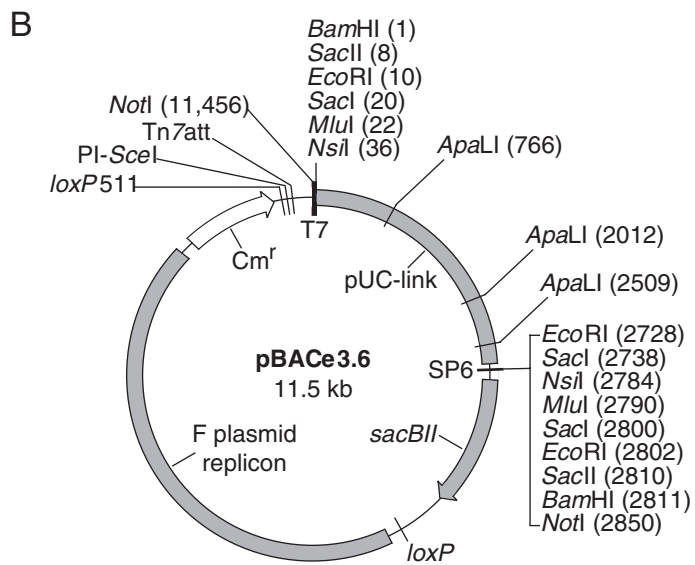
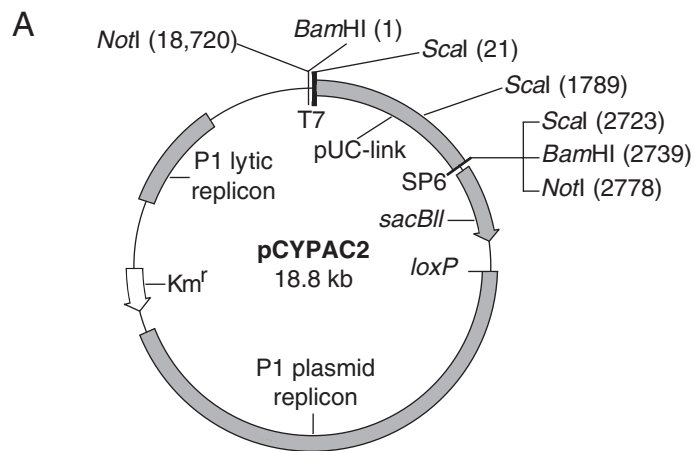


Figure 5.9.1 (at left) The (A) pCYPAC2, (B) pBACe3.6, and (C) pPAC4 vectors contain common features for positive selection of cloned inserts. The insert-containing *E. coli* cells can be grown in preference over cells containing plasmids without an insert by including sucrose in the medium. Sucrose is converted into toxic metabolites by levansucrase encoded by the *sacBII* gene present in all three vectors. The pUC-link segment contains a functional high-copy-number plasmid that includes the ampicillin resistance gene (Ap^r ; not shown). The pUC-link sequence is removed in the cloning procedure and is not present in the PAC and BAC clones. The digested genomic DNA fragment replaces the fragment containing the pUC-link in the PAC or BAC vectors. The *loxP* recombination sites are not used in cloning. A detailed description of vector elements is given elsewhere (see, Background Information, vector elements). (A) The pCYPAC2 vector contains the P1 plasmid replicon, which maintains recombinant clones at around one copy per cell. Alternatively, the P1 lytic replicon can be induced to provide higher copy numbers by adding the *lac* inducer IPTG into the medium. The kanamycin resistance gene (Km^r) is also present in this vector. (B) The pBACe3.6 vector contains the chloramphenicol resistance gene (Cm^r) and the F plasmid replicon, maintaining recombinant clones at one copy per cell. The multiple cloning sites flanking the pUC-link segment allow positive selection for cloned inserts using any of the restriction enzymes *Bam*HI, *Sac*II, *Eco*RI, *Sac*I, *Mlu*I, or *Nsi*I. The Tn7att sequence permits specific Tn7-based retrofitting of BAC clones. (C) The pPAC4 vector contains four elements from pCYPAC2: the P1 plasmid replicon, the kanamycin resistance gene, the *sacBII* gene, and the pUC-link. The P1 plasmid replicon ensures low-copy-number maintenance of recombinant clones; the P1 lytic replicon is removed. In addition, the *oriP* and *bsr*^r cassette have been included to facilitate the use of PAC clones for functional analysis of genes carried on the cloned inserts. The Tn7att sequence is also present, enabling specific retrofitting of PACs.

PREPARATION OF BAC/PAC CLONES USING pCYPAC2, pPAC4, OR pBACe3.6 VECTOR

Preparation of large-insert bacterial and P1-derived artificial chromosome (BAC and PAC) clones is conceptually similar to procedures used for preparing small plasmid clones. A single vector fragment with dephosphorylated *Bam*HI or *Eco*RI ends (see Support Protocol 1) is ligated to size-fractionated *Mbo*I- or *Eco*RI-digested genomic DNA fragments that are recovered from agarose slices using an electroelution procedure (see Support Protocol 4). Ligations are performed at low genomic DNA concentrations to favor the formation of circular molecules. A number of 120- to 300-kb fractions from the preparative gel are tested initially to identify the fraction giving the largest average insert size at an acceptable cloning efficiency. Once the optimal fraction is identified, a large-scale ligation is performed to make sufficient clones for a several-fold redundant library.

Size fractionation of the partially digested DNA is important to select for the desired size fragments and to remove smaller digestion products, which otherwise would result in an overwhelming background of small BAC or PAC clones. Circular plasmid ligation products are introduced directly into host cells by electroporation. Although this procedure is conceptually easy, there are several aspects of the Basic and Support protocols that are critical for successful preparation of BAC and PAC libraries (see Critical Parameters).

Materials

- ≥2 to 10 ng/μl size-fractionated *Mbo*I- or *Eco*RI-digested genomic DNA (see Support Protocol 4)
- 10 to 50 ng/μl pBACe3.6, pCYPAC2, or pPAC4 vector DNA prepared for cloning (see Support Protocol 1)
- 1 Weiss U/μl T4 DNA ligase (Life Technologies) and 5× buffer (see recipe)
- 0.5 M EDTA, pH 8.0 (APPENDIX 2)
- 10 mg/ml proteinase K

BASIC PROTOCOL

Construction of Recombinant DNA Libraries

5.9.3

100 mM phenylmethylsulfonyl fluoride solution (PMSF; see recipe)
 TE/PEG solution: 0.5× TE buffer, pH 8.0 (*APPENDIX 2*) containing 30% (w/v)
 polyethylene glycol 8000 (PEG 8000)
 Electrocompetent bacterial cells (ElectroMAX DH10B; Life Technologies)
 SOC medium (Life Technologies; also see *UNIT 1.8*, but reduce yeast extract to
 0.5%)
 LB plates (*UNIT 1.1*) containing 5% (w/v) sucrose and either 25 µg/ml kanamycin
 (for PAC clones) or 20 µg/ml chloramphenicol (for BAC clones):
 100 × 15–mm petri dishes for test transformation
 22 × 22–cm trays for picking colonies
 LB medium (*UNIT 1.1*) containing 20 µg/ml chloramphenicol (BAC clones) or 25
 µg/ml kanamycin (PAC clones)
 TE buffer, pH 8.0 (*APPENDIX 2*)
NotI restriction endonuclease and buffer (New England Biolabs)
 1% (w/v) ultrapure agarose solution (Life Technologies)
 0.5× TBE buffer (*APPENDIX 2*)
 Low-range PFG markers in agarose containing a mixture of lambda *HindIII*
 fragments and lambda concatemers (New England Biolabs)
 0.5 µg/ml ethidium bromide in 0.5× TBE buffer (*APPENDIX 2*)
 80% (v/v) glycerol, sterile

 Dry ice/ethanol bath
 16° and 37°C water baths
 0.025-µm-pore-size microdialysis filters (Millipore): 25-mm diameter for
 small-scale test ligation and 47-mm diameter for large-scale ligation
 Wide-bore pipet tips, sterile
 Disposable microelectroporation cuvettes with a 0.15-cm gap (Life Technologies
 or equivalent)
 Electroporator (Cell Porator equipped with a voltage booster; Life Technologies or
 equivalent)
 15-ml snap-cap polypropylene tubes, sterile
 Orbital shaker, 37°C
 Automated plasmid isolation system (AutoGen 740, Integrated Separation
 Systems, optional)
 Flexible plastic 96-well plate (Falcon or equivalent)
 Contour-clamped homogeneous electrical field (CHEF; *UNIT 2.5B*) apparatus
 (Bio-Rad) or field-inversion gel electrophoresis (FIGE; *UNIT 2.5B*) apparatus
 (Bio-Rad or equivalent)
 Digital imager (Alpha Innotech IS1000 or equivalent)
 50-ml disposable centrifuge tube (Corning or equivalent)

 Additional reagents and equipment for modified alkaline lysis preparation of BAC
 or PAC clone DNA (see Support Protocol 5; optional)

CAUTION: To prevent shearing, use sterile wide-bore pipet tips for all steps involving the handling of genomic DNA.

Perform small-scale test ligation

1. Mix ~50 ng of each 120- to 300-kb size fraction of *MboI*- or *EcoRI*-digested genomic DNA (see Support Protocol 4) with 25 ng *BamHI*- or *EcoRI*-digested and dephosphorylated pBACe3.6 vector DNA, or 50 ng *BamHI*-digested and dephosphorylated pCYPAC2 or pPAC4 vector DNA prepared for cloning (see Support Protocol 1).

Suitable ligation controls include reactions with vector only or without ligase.

The vector fragment is 8.7 kb for pBACe3.6, ~15.5 kb for pCYPAC2, and 16.8 kb for pPAC4. The average size of fractionated genomic DNA is ~150 to 200 kb. Thus, a 10/1 molar ratio of vector to insert is obtained by using approximately half quantity of pBACe3.6 vector to insert DNA in the ligation reaction, or equal quantities of vector and insert DNA for PAC cloning. The genomic DNA concentration in the final reaction is maintained at a low value (1 ng/ μ l) to favor circle formation over concatemer formation.

Ligation schedules will vary somewhat depending on the restriction endonuclease used in preparing the vector and insert DNA. For a typical EcoRI-EcoRI ligation schedule, perform step 1 through step 6 in a day and then dialyze the solution against TE/PEG solution at 4°C overnight. For BamHI-MboI cloning, incubate the ligation mixture in a thermal cycler at 16°C for 8 hr, then hold at 4°C overnight, and start step 4 the following day.

2. To each tube, add 10 μ l of 5 \times T4 DNA ligase buffer (1 \times final) and sterile water to bring the total volume to 50 μ l. Mix very gently.
3. Add 1 Weiss unit T4 DNA ligase and incubate at 16°C for 4 hr for EcoRI-EcoRI cloning or 8 hr for BamHI-MboI cloning.

Keep in mind that the vector preparation can always contain a very low fraction of shortened fragments caused by BamHI or EcoRI degenerate (star) activities. Ligation could occur at those unspecific sites after long incubation, which could result in a small portion (<10%) of noninsert clones in the library. The ligation time in this protocol has been optimized to generate a (near) maximum level of recombinant clones with a minimum ratio of nonrecombinant clones.

4. To stop the reaction, add 1 μ l of 0.5 M EDTA, pH 8.0, and 1 μ l of 10 mg/ml proteinase K and incubate at 37°C for 1 hr.

Inactivation of the T4 DNA ligase with proteinase K is preferable to that by heating. Heating high-molecular-weight DNA is not recommended.

5. To inactivate proteinase K, add 1 μ l of 100 mM PMSF solution and incubate at room temperature for 1 hr.

Generate test clones by electroporation

6. Spot the ligation mixture on the middle of a 25-mm-diameter, 0.025- μ m-pore-size microdialysis filter floating on sterile distilled deionized water in a disposable petri dish. Dialyze for \geq 2 hr at room temperature.

The salts in the ligation mixture must be removed to reduce heat development during the subsequent electrical pulse. After 2 hr dialysis, the volume will increase ~150 μ l due to the higher osmotic pressure in the ligation buffer (which contains PEG 8000) as compared to water (see Background Information, concentration of ligation mixture).

7. Carefully recover the dialyzed ligation mixture from the membrane using a sterile wide-bore pipet tip and transfer to a microcentrifuge tube.
8. Discard the water from the petri dish and pour ~15 ml TE/PEG solution in the dish. Transfer the empty membrane onto this solution.
9. Drop the recovered solution (step 7) onto the membrane again and continue dialysis for ~5 hr at room temperature until it is equilibrated.

Concentrating ligation solution reduces the number of transformations necessary (see Background Information, concentration of ligation mixture).

Dialysis that is to go longer than ~5 hr can be done at 4°C (see ligation/dialysis schedules in step 1).

10. Transfer the membrane (still containing the concentrated solution) onto a petri dish cover, recover the ligation mixture from the dialysis membrane using a sterile wide-bore pipet tip, and transfer it to a microcentrifuge tube. Keep on ice.

*After concentration, ~8 μ l should be recovered from 50 μ l ligation solution; thus, the solution will not be dispersed into the outside solution during the membrane transfer. The solution should be transformed into *E. coli* as soon as possible. Keeping it for long periods will cause transformation efficiency to drop significantly.*

11. Precool disposable microelectroporation cuvettes (0.15-cm gap) and microcentrifuge tubes on ice. The microelectroporation cuvette can be used up to 10 times as long as the same transformation is repeated.

The test transformation should be done in duplicate or triplicate to better estimate the average transformation efficiency.

12. Thaw the required amount of electrocompetent bacterial cells on ice and transfer 20- μ l aliquots into precooled microcentrifuge tubes. Keep on ice.

Thaw a total of $(n + 1) \times 20 \mu$ l cells, where n is the number of electroporations, to provide sufficient cells for the transformation.

13. Using a sterile wide-bore pipet tip, transfer 2 μ l ligation mixture to each aliquot and gently mix with the pipet tip.

14. Transfer the solution into a precooled electroporation cuvette, placing the droplet carefully between the electrodes and avoiding the formation of any air bubbles.

15. Place the cuvette into the electroporation chamber and deliver a pulse according to the following conditions:

voltage booster settings:
resistance on voltage booster, 4000 Ω

Cell-Porator settings:
voltage gradient, 13 kV/cm
capacitance, 330 μ F
impedance, low Ω
charge rate, fast.

16. Collect the droplet of cells and dilute immediately into 500 μ l SOC medium in a 15-ml snap-cap polypropylene tube. Incubate 1 hr at 37°C in an orbital shaker at 200 rpm.

This incubation allows time for the antibiotic resistance gene to be expressed.

Under these conditions the cell concentration is considerably higher than can be achieved at the stationary phase of growth, and thus there is less chance of cell division.

Analyze test clones

17. Spread the entire aliquot of cells on a 100 \times 15-mm LB plates containing 5% sucrose and either 20 μ g/ml chloramphenicol (for BAC clones) or 25 μ g/ml kanamycin (for PAC clones).

Alternatively, recover the cells by centrifugation and concentrate to a final volume of 100 to 200 μ l before plating. If the whole volume is to be plated without concentration, ensure that the plates are dried under a hood prior to plating. A thoroughly dried plate is used to speed up absorption of the solution.

18. Grow overnight at 37°C, count colonies, and estimate the titer per transformation.

If a series of size fractions are tested simultaneously, the cloning efficiency is seen to drop precipitously as the size of the fractions increases above 200 kb. The factors contributing to this sharp decline in cloning efficiency are not well understood at present.

Construction of a total genomic mammalian BAC or PAC library with 10-fold redundancy requires creation of >200,000 clones. It is important to achieve as high a cloning efficiency as possible with the largest possible insert size, because this will reduce the number of electroporations that are subsequently required.

Isolate and characterize individual test clones

19. Pick 40 clones with a sterile toothpick and grow the cells overnight in 1.5 ml LB medium containing 20 µg/ml chloramphenicol (BAC clones) or 25 µg/ml kanamycin (PAC clones).
20. Extract DNA using an automated plasmid isolation system (AutoGen 740) or the modified alkaline lysis procedure (see Support Protocol 5).
21. Dissolve DNA (0.5 to 1 µg) in 100 µl TE buffer. Digest 5 to 10 µl DNA with 0.1 U *NotI* in a 20-µl volume to separate the vector and insert DNA fragments.

It is convenient to perform the reaction in a flexible 96-well plate at 37°C for 2 hr.

22. Analyze digested DNA using a CHEF or a FIGE apparatus. Use a 1% agarose gel, 0.5× TBE buffer, low-range PFG markers, and the following conditions:

For CHEF: 14°C, 6 V/cm, 16 hr, 0.1 to 40 sec pulse time, and 120° angle.

For FIGE: Room temperature, 180 V forward voltage, 120 V reverse voltage, and 16 hr with 0.1 to 14 sec switch time linear shape.

The FIGE system is much less expensive than CHEF, and is able to obtain enough resolution for this purpose.

23. Stain the gel with 0.5 µg/ml ethidium bromide solution and calculate the molecular weight of inserts using a digital imager.

*Most of the clones will give two bands—a vector fragment (8.7 kb for pBACe3.6, 15.5 kb for pCYPAC2, and 16.8 kb for pPAC4) and the insert DNA of variable size. If the insert contains internal *NotI* sites, additional bands will be seen. Incomplete digestion usually leads to a characteristic doublet, where the two bands differ by the size of the vector.*

Create library

24. Repeat the ligation procedure using a size fraction of genomic DNA that gives the desired average insert size and cloning efficiency. Scale up the ligation reaction to 500 to 1000 µl using the entire remaining electroeluted insert DNA.

The ligation reaction is usually performed in several microcentrifuge tubes at 250 µl per tube. Otherwise, the increase in volume may cause the solution to be dispersed into the water during subsequent dialysis, even though a 47-mm-diameter filter is used.

25. Dialyze and concentrate the ligation mixture as described (steps 6 to 10), but use a 47-mm-diameter microdialysis filter for each 250-µl aliquot.

After concentration, ~40 µl should be recovered from each 250-µl ligation solution.

26. Add 12 µl concentrated DNA to tubes containing ~110 µl each of electrocompetent bacterial cells. Mix well by gentle stirring using a wide-bore pipet tip. Thaw enough electrocompetent cells to complete all transformations using the entire recovered ligation mixture.

27. Perform identical transformations at 20 μ l per electroporation and collect the transformed cells from 10 electroporations into a 50-ml disposable centrifuge tube containing 5 ml SOC medium at room temperature.

The transformation efficiency will drop if the medium containing transformed cells is kept at 4°C during the transformation experiment.

28. Incubate the cells on an orbital shaker at 200 rpm at 37°C for 1 hr.
29. Add 800 μ l of 80% sterile glycerol solution into the bacterial culture and mix very well. Spread 200 μ l on two individual 100 \times 15-mm LB plates containing 5% sucrose and the appropriate antibiotic to examine the titer from each tube.

The amount of culture spread per plate is determined based on the titer estimated from the test ligation and transformation (step 18). The number of colonies on a plate should be between 300 and 700 to estimate the accurate titer for colony picking.

It is easier to homogeneously spread 200 μ l culture when 300 μ l SOC medium is also added to the plate.

30. Freeze the remaining cells in a dry ice/ethanol bath and keep at -80°C until colony picking can be scheduled. To pick colonies, spread culture at 1600 clones/plate on 22 \times 22-cm LB plates containing 5% sucrose and the appropriate antibiotic. Grow overnight at 37°C.

The colony picking is usually scheduled within a couple of months. The frozen cells should be able to be stored for at least 1 year.

SUPPORT PROTOCOL 1

PREPARATION OF BAC/PAC VECTOR FOR CLONING

The integrity of the vector's cohesive ends and the absence of uncut vector molecules are prerequisites for success in preparing BAC and PAC libraries. This protocol describes steps for preparing the vectors to maximize cloning efficiency and to minimize the level of nonrecombinant clones.

Preparation of pBACe3.6, pCYPAC2, and pPAC4 for cloning involves a single restriction endonuclease digestion with *Bam*HI or *Eco*RI. These plasmids contain the pUC-link insert at the cloning site, thus inactivating the *sacB* gene and enabling clones containing the original plasmid to grow in the presence of sucrose. Clones containing uncut vectors cannot be distinguished from truly recombinant clones, and great care should be taken to ensure complete digestion of the original vector with *Bam*HI or *Eco*RI during vector preparation. Depending on the source and activity of the different enzymes used, it is necessary to optimize the conditions for each digestion. The following conditions have been found to yield good vector preparations using the recommended enzymes.

Additional Materials (also see Basic Protocol)

pBACe3.6, pCYPAC2, or pPAC4 stock in *E. coli* DH10B cells (P. de Jong; pieter@dejong.med.buffalo.edu)

LB plates (UNIT 1.1) containing:

25 μ g/ml kanamycin (for PAC) or 20 μ g/ml chloramphenicol (for BAC)

5% (w/v) sucrose and either kanamycin or chloramphenicol

5% (w/v) sucrose, 100 μ g/ml ampicillin, and either kanamycin or chloramphenicol

*Bam*HI and *Eco*RI restriction endonucleases and 10 \times buffers (New England Biolabs or equivalent)

0.7% (w/v) agarose gels (for standard electrophoresis)

Calf intestine alkaline phosphatase (AP; Boehringer Mannheim)

10 mg/ml proteinase K (Boehringer Mannheim) stock solution

95% (v/v) ethanol
1.0% (w/v) ultrapure agarose solution (Life Technologies; for CHEF system)
6× loading buffer (UNIT 2.5A), not containing xylene cyanol FF or SDS
1-kb ladder or lambda *Hind*III markers
T4 polynucleotide kinase (New England Biolabs)
30% (w/v) polyethylene glycol (PEG) 8000
1.5-mm-thick electrophoresis comb for CHEF apparatus
Dialysis tubing of $\frac{3}{4}$ in. diameter, mol. wt. exclusion limit of 12,000 to 14,000 daltons (Life Technologies or equivalent)
Dialysis clip
Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5A), plasmid extraction (e.g., UNIT 1.7), CsCl/ethidium bromide equilibrium centrifugation (e.g., UNIT 1.8), electroelution (see Support Protocol 4), and ethanol precipitation (UNIT 2.1A)

Test vector DNA

1. Streak recombinant DH10B cells harboring BAC or PAC vectors onto an LB plate containing the appropriate antibiotic (chloramphenicol or kanamycin, respectively). Incubate overnight at 37°C.
2. Isolate five single colonies and inoculate into separate 15-ml snap-cap polypropylene tubes containing 3 ml LB medium with the appropriate antibiotic. Grow cultures overnight at 37°C.
3. Use 1.5 ml of each culture to prepare DNA using an automated plasmid isolation system (AutoGen 740) or the modified alkaline lysis protocol (see Support Protocol 5). Store the remainder of each culture at 4°C.
4. Resuspend DNA in 100 µl TE buffer.
5. Digest 5 µl DNA from each preparation with *Not*I, 5 µl with *Bam*HI, and 5 µl with *Eco*RI, using the manufacturers' recommended conditions.
6. Analyze digested DNA in a 0.7% agarose gel (UNIT 2.5A). Discard clones that contain any rearrangements of the vector.

PAC vectors do not contain EcoRI sites. Otherwise all enzymes should liberate pUC19-link from the vectors, producing fragments of ~2.7 kb (the pUC19 stuffer fragment) and ~8.7 kb (pBACe3.6), 15.5 kb (pCYPAC2), or 16.8 kb (pPAC4).

Prepare vector for large-scale ligation

7. Dilute the selected culture into 1 liter LB medium containing the appropriate antibiotic. Grow to saturation at 37°C.

This culture is used for making a large-scale preparation of the PAC or BAC vectors.

8. Prepare a crude lysate of the culture by the alkaline lysis or cleared lysate method.

Any standard protocol for plasmid extraction may be used here (UNIT 1.7).

9. Purify vector plasmid from crude lysate by CsCl/ethidium bromide equilibrium centrifugation.

Any standard protocol for CsCl/ethidium bromide equilibrium centrifugation may be applied (see UNIT 1.8).

Digest vector with BamHI or EcoRI

10. To establish the minimal amount of restriction enzyme required for complete digestion of the vector, first digest 50 ng vector DNA with 0.1, 0.2, 0.5, and 1 U of either *Bam*HI or *Eco*RI for pBACe3.6, and *Bam*HI for PAC vectors, in separate 10- μ l reactions for 1 hr at 37°C. Perform conventional agarose gel electrophoresis and ethidium bromide staining to view the digestion.

It is essential to establish the minimal amount of enzyme required for complete digestion. If too little enzyme is used, many nonrecombinant clones containing undigested vector can be expected. On the other hand, excess enzyme results in a very high ratio of noninsert clones, which are most likely induced by star activity.

11. Next, to establish the maximal amount of DNA in the reaction (and the minimum scale up volume), digest 100, 200, and 400 ng vector DNA with 2, 4, and 8 times the minimal amount of enzyme in separate 10- μ l reactions (total 9 reactions) for 1 hr at 37°C. Perform conventional agarose gel electrophoresis and ethidium bromide staining to view the digestion.

*High-quality vector preparation is difficult. If necessary, the vector can be digested with *Apa*LI for pBACe3.6 or *Scal*I for pCYPAC2 and pPAC4 to destroy the pUC-stuffer fragment in the vectors nearly completely. Note that a precipitation step is needed in order to change the reaction buffer and that optimal enzyme concentration must be determined.*

12. Incubate ~30 μ g vector DNA with the defined amount of the appropriate enzyme at 37°C for 15 min.
13. Add 1 U calf intestine AP and continue incubating at 37°C for 1 hr.
14. To stop the reaction, add 0.5 M EDTA to a final concentration of 15 mM, add 10 mg/ml proteinase K to 200 μ g/ml final, and incubate at 37°C for 1 hr.
15. To inactivate proteinase K, add 100 mM PMSF solution to 2 mM final and incubate at room temperature for 1 hr.

Perform PFGE

16. During step 15, clean a 1.5-mm-thick electrophoresis comb with 95% ethanol and cover teeth with autoclave tape to create a large preparative slot with sufficient space to load the reaction mixture. Leave enough empty wells for a 1-kb ladder or lambda-*Hind*III markers on both sides and at least one empty well between the markers and the preparative slot.
17. Prepare a 1% agarose gel in a CHEF mold using the preparative comb. While the gel is solidifying, add 2 liters of 0.5 \times TBE buffer to the CHEF apparatus tank and equilibrate the unit at 14°C.
18. Place the gel in the precooled unit. Add 6 \times gel loading buffer to the samples (final 1 \times) and load samples in the large preparative well. Load 1-kb ladder or lambda *Hind*III markers in both side wells. Perform electrophoresis at 14°C and 6 V/cm for 16 hr with 0.1 to 40 sec pulse time at 120° angle.

Use of xylene cyanol FF in the gel loading buffer is not recommended because this dye migrates with the vector band under these electrophoresis conditions.

Recover vector DNA

19. Cut away two flanking gel slices that contain the markers as well as 1 to 2 mm from each side of the central vector lane. Wrap the remaining middle portion of the preparative lane in plastic wrap and store at 4°C.
20. Stain the flanking slices with ethidium bromide to detect the major vector fragment.

21. Cut linear vector DNA lacking the pUC-stuffer fragment from the central lane and recover by electroelution (see Support Protocol 4, steps 38 to 46).
22. Load sample into $\frac{3}{4}$ -in.-diameter dialysis tubing with a 12,000- to 14,000-Da mol. wt. exclusion limit, and dialyze once against 1 liter TE buffer at 4°C for >2 hr.
23. Recover solution from the dialysis tubing and precipitate DNA with sodium acetate and either ethanol or isopropanol (*UNIT 2.1A*). Dissolve DNA in 100 μ l TE buffer.

Check efficiency of dephosphorylation

24. Place four microcentrifuge tubes on ice and number them 1 to 4. Add 4 μ l of 5 \times T4 DNA ligase buffer and 50 ng of digested and AP-treated vector DNA into the tubes, and adjust the volume to 20 μ l with sterile distilled and deionized water.
25. Add 1 U T4 polynucleotide kinase to tubes 3 and 4. Incubate all four tubes at 37°C for 1 hr.

T4 polynucleotide kinase will be active in 1 \times ligase buffer.

26. Heat at 65°C for 20 min, then place on ice to inactivate the enzyme.
27. Add 1 U T4 DNA ligase to tubes 2 and 4. Incubate all four tubes at 16°C for >4 hr.
Thus, tube 1 contains no enzyme, tube 2 contains T4 DNA ligase, tube 3 contains T4 polynucleotide kinase, and tube 4 contains T4 polynucleotide kinase and T4 DNA ligase.
28. Run samples in a 0.7% agarose gel, stain the gel with ethidium bromide, and view.
Closed circular DNA should not be seen in samples 1, 2, and 3. Sample 4 should have closed circular DNA but no linear DNA.
29. Perform a ligation using all the remaining vector DNA solution in a 200- μ l reaction volume with 4 U T4 DNA ligase.
30. Purify the linear monomeric vector DNA away from the background of ligated vector by PFGE in a CHEF apparatus (see steps 16 through 23).

Check ratio of background clones

31. Perform test ligations using vector DNA alone (25 ng pBACe3.6, 50 ng PAC vectors) and vector plus 50 ng control DNA (see Basic Protocol, steps 1 to 5).
32. Dialyze against water and 30% PEG 8000, and perform transformation as described (see Basic Protocol, steps 6 to 16).
33. Spread transformed cells on LB plates containing 5% sucrose and either kanamycin or chloramphenicol, as appropriate. Grow overnight at 37°C.

Kanamycin/sucrose and chloramphenicol/sucrose plates are used to determine the level of recombinant clones and nonrecombinant clones. If the number of colonies obtained from vector DNA alone is high (>100), vector must be prepared again, beginning with enzyme titration (step 10). If the plasmid is isolated from these background clones and analyzed by PFGE after NotI digestion, smaller vector sizes will be observed. Presumably, star activity causes nonspecific digestion in the sacB gene (either promoter or coding sequence), thus inactivating the gene, and the nonspecifically digested, nondephosphorylated DNA is circularized by self-ligation. Thus, these colonies can grow on sucrose plates without insert DNA.

34. To examine the complete removal of stuffer fragment from the vector, spread the transformed cells with self-ligated vector on LB plates containing 5% sucrose, 100 μ g/ml ampicillin, and either kanamycin or chloramphenicol, as appropriate. Grow overnight at 37°C.

Kanamycin/ampicillin/sucrose and chloramphenicol/ampicillin/sucrose plates are used to determine the remaining level of undigested vector containing the pUC-link. Ideally, there should be zero Amp^R colonies.

PREPARATION OF HIGH-MOLECULAR-WEIGHT DNA FROM LYMPHOCYTES IN AGAROSE BLOCKS

Construction of a total genomic library demands that some consideration be given to the source of DNA to be used. Although DNA from cultured cell lines has been extensively used in constructing cosmid libraries, it is well known that passage in tissue culture may result in chromosomal rearrangements; therefore, it is considered more desirable to use DNA obtained directly from human or animal sources. A convenient source of DNA for this purpose is whole blood, and more specifically, the circulating lymphocytes. Assuming an average of $\sim 5 \times 10^6$ lymphocytes per milliliter of blood and 6 pg DNA per cell, as much as 30 μ g DNA can be recovered per ml of blood, making it a very convenient and inexpensive source of DNA. This method depends on preferential lysis of red blood cells (RBC) using a gentle procedure that does not affect lymphocytes. The method can be applied equally well to cultured cells by omitting the RBC-lysis step; however, the cells should be collected as quickly as possible, especially when using attached cells in culture. Other protocols for embedding high-molecular-weight mammalian DNA from tissues (see Support Protocol 3 and UNIT 2.5B) should also produce acceptable results.

Materials

Healthy human volunteer
PBS (APPENDIX 2), ice cold
1 \times RBC lysis solution (see recipe)
InCert agarose (FMC Bioproducts)
Proteinase K lysis solution (see recipe)
TE₅₀ buffer: 10 mM Tris·Cl, pH 8.0 (APPENDIX 2)/50 mM EDTA
0.1 mM PMSF solution: 100 mM PMSF solution (see recipe) diluted 1/1000 in
TE₅₀ buffer immediately before use
0.5 M EDTA, pH 8.0 (APPENDIX 2)
Blood-drawing equipment
Blood collection tubes containing EDTA
Automated hematology counter
50-ml conical screw-cap polypropylene tubes, sterile
Refrigerated centrifuge with rotor/adapters for 50-ml tubes (e.g., Sorvall RI6000D
centrifuge with H-1000B swinging-bucket rotor or equivalent)
Roller mixer (Robbins Scientific or equivalent)
50°C water bath
10 \times 5 \times 1.5-mm disposable DNA plug mold (Bio-Rad)

Collect blood and remove red blood cells

1. Use blood-drawing equipment to obtain 45 ml venous blood from a healthy human volunteer in blood collection tubes containing EDTA. Mix well to avoid clot formation.
2. Count the total number of lymphocytes using an automated hematology counter.
The total number should be 2.25 to 3.3 $\times 10^8$ in 45 ml.
3. Transfer blood to two 50-ml conical screw-cap polypropylene tubes, add 10 ml ice-cold PBS to each, and mix gently.
4. Centrifuge at 1876 $\times g$ for 5 min at 4°C. Discard the supernatant using a 10-ml disposable pipet, being careful not to remove any lymphocytes from the fuzzy coat layer.

Centrifugation results in a large pellet of red blood cells at the bottom of the tube and a very thin layer of lymphocytes (the fuzzy coat layer) at the interface with the plasma layer.

5. Repeat wash with ice-cold PBS ten times.
6. Discard the final supernatant, mix the cell suspension well, and divide into four 50-ml tubes.
7. Add 25 ml of 1× RBC lysis solution to each tube, mix samples gently on a roller mixer, and incubate on mixer ~20 min at room temperature.

During this time, an exchange process catalyzed by carbonic anhydrase results in accumulation of ammonium chloride inside the cells; this accumulation increases the internal osmotic strength and causes swelling and bursting of RBCs. Lymphocytes have much lower carbonic anhydrase activity and do not undergo lysis.

8. Visually monitor the progress of RBC lysis. When it appears complete (usually <30 min), centrifuge the tubes for 10 min at $208 \times g$, 4°C.

When examined against a light source, the suspension originally appears cloudy. The cloudiness will disappear as the red cells lyse. After centrifugation, a pellet of lymphocytes should be obvious at the bottom of each tube.

9. Discard the supernatant by gentle inversion, taking care not to disturb the lymphocyte pellet. Rinse the inside of the tubes with 2 ml ice-cold PBS, being careful not to disturb the pellet. Discard the supernatant with a micropipet tip.
10. Resuspend lymphocytes in 10 ml ice-cold PBS and combine in one 50-ml tube. Centrifuge 5 min at $208 \times g$, 4°C, and discard the supernatant carefully to remove most of the remaining lysate debris.

A variable amount of red cell debris may stay with the pellet.

11. Repeat wash with ice-cold PBS until most of the red color is removed.
12. Resuspend the pellet in ice-cold PBS at 1×10^8 cells/ml (~600 µg DNA/ml).

Embed cells in agarose

13. Dissolve 0.1 g InCert agarose in 10 ml PBS (1%) in a microwave oven and keep at 50° in a water bath.
14. Mix the cell suspension well and transfer 400 µl to a microcentrifuge tube. Warm the tube by hand for 3 min.
15. Add 400 µl of 1% molten InCert agarose and mix gently. Avoid making any bubbles.

The final agarose concentration is 0.5% and the cell concentration is 5×10^7 /ml, equivalent to ~300 µg DNA/ml.

16. Transfer the cell/agarose suspension as quickly as possible to $10 \times 5 \times 1.5$ -mm disposable DNA plug molds using a micropipet tip.

Each mold takes ~80 µl cell/agarose suspension, containing ~25 µg DNA. From 45 ml blood, it is possible to make ~45 plugs.

17. Place the molds on ice for 30 to 60 min to solidify the agarose.

Extract high-molecular-weight DNA

18. Extrude the DNA plugs from the mold directly into a 50-ml tube containing 50 ml proteinase K lysis solution.

The number of DNA plugs per tube should be <50.

- Stand the tube in a 50°C water bath, mix periodically, and incubate for 48 hr. Replace solution with fresh proteinase K lysis solution after the first 24 hr.

Residual amounts of hemoglobin from the RBC lysate remain trapped in the blocks and act as a convenient marker to monitor the digestion process. The residual color disappears within a few hours, but digestion is extended to 48 hr to ensure thorough protein removal.

- Rinse plugs several times with sterile distilled, deionized water.
- Add 50 ml TE₅₀ buffer and mix on a roller mixer at 4°C for 24 hr. Replace the solution with fresh TE₅₀ buffer at least twice during the 24 hr.
- To inactivate proteinase K, wash plugs twice for 2 hr each with 50 ml of 0.1 mM PMSF solution on the roller mixer at 4°C.
- Wash plugs with TE₅₀ buffer on the roller mixer at 4°C for 24 hr. Replace the solution with fresh TE₅₀ buffer at least once during the 24 hr.
- Store DNA plugs in 0.5 M EDTA at 4°C.

The DNA plugs can be stored under these conditions for more than a year.

SUPPORT PROTOCOL 3

PREPARATION OF HIGH-MOLECULAR-WEIGHT DNA FROM ANIMAL TISSUE CELLS IN AGAROSE BLOCKS

Animal tissues are also useful as sources of DNA for library preparation, since it is often difficult to obtain enough lymphocytes from small model organisms (e.g., mice, rats). Here, an alternate method is described for preparation of DNA plugs from animal tissue cells (e.g., spleen, kidney, and brain).

Additional Materials (also see Support Protocol 2)

Healthy animal (e.g., ~5-week-old mice, rats)

Sterile dissecting tools

Sterile Dounce homogenizer

15-ml conical screw-cap polypropylene tubes, sterile (Corning or equivalent)

Counting chambers (VWR)

Additional reagents and equipment for euthanasia with CO₂ (Donavan and Brown, 1995)

- Euthanize mice or rats in a plastic bag or desiccator containing CO₂ gas.
- Immediately dissect animal using sharp sterile scissors, remove desired tissue(s) from the animal, and transfer them to a petri dish on ice.

Spleen, kidney, and brain are all good sources for extraction of high-molecular-weight DNA.

- Rinse tissues with ice-cold PBS and remove fat and other associated tissues with sterile forceps.
- Transfer 1 to 2 spleens, kidneys, or brains to a precooled sterile Dounce homogenizer on ice.

Spleen, kidney, and brain should be homogenized separately.

- Add ~2.5 ml ice-cold PBS into the homogenizer and grind ~5 times on ice.
- Transfer the supernatant carefully from the homogenizer into a chilled 50-ml conical screw-cap polypropylene tube, taking care to leave behind as much tissue as possible.

7. Repeat steps 5 and 6 four times, adding the PBS to the chilled conical tube until most tissue is disrupted.
8. Remove debris from the homogenizer with forceps and repeat steps 4 to 7 for all spleens, kidneys, and brains.
9. Fill all tubes to 50 ml with ice-cold PBS.
10. Stand tubes on ice for 2 to 3 min and carefully transfer supernatants to new 50-ml tubes.

Large tissue debris sinks to the bottom of the old tubes and is removed in this step.

11. Centrifuge at $208 \times g$ (e.g., 1000 rpm in a Sorvall H1000B rotor) for 10 min at 4°C .
12. Discard supernatant by gentle inversion, taking care not to disturb the cell pellet.
13. Add ~ 1 ml ice-cold PBS, resuspend cells by pipetting gently, and remove large debris that cannot be suspended with a micropipet.
14. Fill tubes to 50 ml with ice-cold PBS, mix gently, and repeat centrifugation.
15. Decant the supernatant and resuspend the cells in the residual PBS on ice.
16. Add 1 ml ice-cold PBS and transfer to a 15-ml conical screw-cap polypropylene tube.
17. Rinse the 50-ml tubes with 1 ml ice-cold PBS and add to the 15-ml tubes.
18. Prepare 20 μl of a 20-fold dilution of the cell suspension in a 1.5-ml microcentrifuge tube.
19. Count the number of cells using a counting chamber.

Assume that 40% of the cells counted are red blood cells (erythrocytes), which do not contain chromosomal DNA.

20. Subtract the estimated number of red blood cells, then adjust the concentration of cells that contain chromosomal DNA to 1×10^8 cells/ml by adding ice-cold PBS.

A single cell carries 6 to 10 pg DNA; thus, 1×10^6 cells correspond to 6 to 10 μg DNA.

21. Embed cells in agarose and extract high-molecular-weight DNA (see Support Protocol 2, steps 13 to 24).

PARTIAL DIGESTION AND SIZE FRACTIONATION OF GENOMIC DNA

The PAC vectors utilize *Bam*HI restriction sites for cloning; however, it is not advisable to use *Bam*HI for partial digestion of genomic DNA, because there may be regions of DNA relatively devoid of the *Bam*HI site. Instead, it is preferable to use the four-base cutter *Mbo*I, which produces four-base overhangs that are compatible with the *Bam*HI overhangs on the vector. *Mbo*I sites are expected to occur on average every 256 bases, not taking into account the base composition of the genomic DNA. The sizes of inserts in PAC clones range from 120 to 300 kb, so it is reasonable to expect that it should be possible to generate overlapping clones from most, if not all, areas of the human genome. The pBACe3.6 vector has *Bam*HI and *Eco*RI cutting sites for cloning; therefore, it is possible to use either *Mbo*I or *Eco*RI for partial digestion. Although there are some other restriction sites in the vector, it is not advisable to use them for construction of highly redundant genomic libraries. These restriction sites are more useful for cloning a specific large fragment or region.

SUPPORT PROTOCOL 4

Construction of Recombinant DNA Libraries

5.9.15

In the following protocol, genomic DNA in agarose plugs (see Support Protocols 2 and 3) is partially digested using *Mbo*I or *Eco*RI. The *Mbo*I partial digestion is accomplished by allowing the enzyme to diffuse into the gel in the absence of MgCl₂, then allowing 10 mM MgCl₂ to diffuse into the gel at 4°C, and finally initiating digestion by transferring the plugs to 37°C. The concentration of enzyme and the length of incubation determine the degree of digestion. On the other hand, competition of *Eco*RI restriction enzyme and *Eco*RI methylase accomplish the *Eco*RI partial digestion. Both enzymes are permitted to diffuse into the gel simultaneously, and the degree of digestion is determined by the concentration of both enzymes. Alternative procedures for controlling *Mbo*I partial digestion (e.g., limiting Mg²⁺ concentration or *Mbo*I/dam methylase competition) may also be used; however, dam methylase is relatively expensive. After partial digestion is completed, the DNA is size fractionated by pulsed-field gel electrophoresis (PFGE; also see UNIT 2.5B) in high-melting-point agarose rather than low-melting-point agarose. The DNA is isolated from the appropriate gel slices by electroelution. Approximately 300 μl of a 2 to 20 ng/μl solution of size-fractionated DNA is obtained. The procedure must be performed multiple times to generate enough DNA for the construction of a library.

Materials

Agarose plugs with embedded high-molecular-weight genomic DNA stored in 0.5 M EDTA (see Support Protocols 2 and 3)
0.5× TBE buffer (APPENDIX 2), sterile
95% (v/v) ethanol
1% (w/v) ultrapure agarose gel (Life Technologies) in 0.5× TBE buffer
Low-range PFG markers in agarose containing a mixture of lambda *Hind*III fragments and lambda concatemers (New England Biolabs)
0.5 M EDTA, pH 8.0 (APPENDIX 2)
0.5 μg/ml ethidium bromide solution in 0.5× TBE buffer (APPENDIX 2)
1× *Mbo*I buffer (see recipe)
10 U/μl *Mbo*I restriction endonuclease (Life Technologies)
1 M MgCl₂
10 mg/ml proteinase K (Boehringer Mannheim) stock solution
10% (w/v) *N*-lauroylsarcosine
TE₅₀ buffer: 10 mM Tris·Cl, pH 8.0 (APPENDIX 2)/50 mM EDTA
100 mM PMSF solution (see recipe)
10 mg/ml BSA (New England Biolabs)
10× *Eco*RI endonuclease/methylase buffer (see recipe)
0.1 M spermidine
20 U/μl *Eco*RI endonuclease (New England Biolabs)
40 U/μl *Eco*RI methylase (New England Biolabs)
TE buffer, pH 8.0 (APPENDIX 2)
1× TAE buffer (APPENDIX 2; optional)
Lambda DNA
50-ml conical screw-cap polypropylene tubes, sterile
Contour-clamped homogeneous electrical field (CHEF) apparatus (Bio-Rad or equivalent; UNIT 2.5B) with 1.5-mm-thick, 20-well comb
Digital imager (Alpha Innotech IS1000 or equivalent)
Disposable γ-ray-sterilized inoculating loops
15-ml conical screw-cap polypropylene tubes, sterile
Dialysis tubing, 3/4-in. diameter, mol. wt. exclusion limit 12,000 to 14,000 daltons (Life Technologies or equivalent)
Dialysis clips

Submarine gel electrophoresis apparatus (Bio-Rad Sub-Cell GT DNA Electrophoresis Cell, 31-cm length × 16-cm width, or equivalent)

Wide-bore pipet tips

Additional reagents and equipment for analytical pulsed-field gel electrophoresis
(UNIT 2.5B)

CAUTION: To prevent shearing, use sterile wide-bore pipet tips in all steps involving the handling of genomic DNA.

Pre-electrophoresis DNA plugs

1. Remove six agarose plugs with embedded genomic DNA from 0.5 M EDTA and transfer the DNA plugs to a 50-ml conical screw-cap polypropylene tube containing 50 ml sterile 0.5× TBE buffer.
2. Dialyze the plugs against 0.5× TBE buffer at 4°C for ≥3 hr.
3. Clean a 1.5-mm-thick comb with 95% ethanol and cover teeth with autoclave tape to create a large preparative slot with sufficient space to place the DNA plugs. Leave enough empty wells for PFG markers on both sides and at least one empty well between the markers and the preparative slot.
4. Place the preparative comb along the short axis of a CHEF mold and prepare a 1% agarose gel of sufficient thickness to cover the sample gel blocks when laid along their length. While the gel is solidifying, add 2 liters of 0.5× TBE buffer to the CHEF apparatus tank and equilibrate the unit at 14°C.
5. Remove the comb gently and load the plugs in the large preparative slot. Load low-range PFG markers in the outer lanes on each side of the preparative lane.

Do not seal the wells with 1% agarose.

6. Place the gel in the precooled unit and run the gel along the long axis using 0.5× TBE buffer at 14°C and 4 V/cm for 10 hr with a 5-sec pulse time.

This step removes inhibitors and results in a more homogeneous partial digestion.

7. Remove the plugs from the preparative slot and place in 50 ml TE buffer and store at 4°C for short periods (up to 2 weeks). If needed, change storage solution to 0.5 M EDTA, pH 8.0.

To avoid DNA deterioration from nuclease, it is advisable to store the plugs in 0.5 M EDTA for longer periods. The DNA plugs can be stored in EDTA for >1 year.

8. Stain the gel in 0.5 µg/ml ethidium bromide solution and examine on a digital imager.

The embedded chromosomal DNA cannot migrate out of the agarose plug under these electrophoresis conditions. However, sheared DNA would be seen on the gel. Thus, it is possible to check the quality of the DNA plug.

Partially digest DNA

For MboI:

- 9a. Transfer pre-electrophoresed DNA plugs into 50 ml 1× *MboI* buffer and continue dialysis at 4°C overnight.
- 10a. Cut a DNA plug into four pieces, and transfer each piece into separate microcentrifuge tubes with a disposable γ-ray-sterilized inoculating loop.

- 11a. Add 400 μ l of 1 \times *Mbo*I buffer and 2.5 U *Mbo*I (diluted with 10 mM Tris·Cl, pH 8.0, just before use) to each tube. Keep the plugs on ice for 1 hr.
- 12a. Add 5 μ l of 1 M MgCl₂ to each tube and stand the tubes on ice for 15 min.
- 13a. Incubate the tubes for 20 to 40 min at 37°C. Test for optimal partial digestion by using various incubation times or amounts of enzyme in step 11a.
- 14a. Immediately place the tubes on ice and add 150 μ l of 0.5 M EDTA, 30 μ l of 10 mg/ml proteinase K, and 75 μ l of 10% *N*-lauroylsarcosine. Mix well and incubate at 37°C for 1 hr to inactivate *Mbo*I.
- 15a. Pour the solution and plugs to a petri dish and remove the solution with a micropipettor.
- 16a. Rinse plugs with TE buffer. Transfer plugs to a 15-ml conical screw-cap polypropylene tube using a disposable γ -ray-sterilized inoculating loop.
- 17a. Add 15 ml TE₅₀ buffer and 15 μ l of 100 mM PMSF solution and dialyze three times for 20 min each at 4°C.
- 18a. Dialyze twice for 30 min each with TE₅₀ buffer at 4°C. Proceed to step 19.

For EcoRI and EcoRI methylase:

- 9b. Transfer pre-electrophoresed DNA plugs into separate microcentrifuge tubes using a disposable γ -ray-sterilized inoculating loop.
- 10b. Add 25 μ l of 10 mg/ml BSA, 50 μ l of 10 \times *Eco*RI endonuclease/methylase buffer, and 13 μ l of 0.1 M spermidine and mix well.

The magnesium concentration is a critical parameter. EcoRI methylase retains only 50% activity in 4 mM Mg²⁺ and EcoRI endonuclease may not work well below 2 mM Mg²⁺. The commercial EcoRI methylase buffer does not contain Mg²⁺ and thus cannot be used for this reaction.

- 11b. Test for optimal partial digestion by varying 1 to 2 U *Eco*RI endonuclease and 0 to 200 U *Eco*RI methylase. Adjust the volume to 500 μ l with sterile distilled deionized water, and put the tubes on ice for 1 hr to allow the enzymes to penetrate into the plugs.
- 12b. Incubate the tubes at 37°C for 2.5 hr and then put on ice.
- 13b. Add 150 μ l of 0.5 M EDTA, 30 μ l of 10 mg/ml proteinase K, and 75 μ l of 10% *N*-lauroylsarcosine and mix well. Incubate the samples at 37°C for 1 hr.
- 14b. Wash plugs with TE₅₀ buffer and PMSF as described (steps 15a to 18a) and then proceed to step 19.

Analyze partial digestion results and optimize conditions

19. Determine the optimal partial digestion condition using a CHEF apparatus. Use the same conditions as in the Basic Protocol step 22.
20. Perform the same experiment using the optimal partial digestion conditions from steps 9 to 18. Use two entire plugs, cut into four pieces, and carry out eight identical partial digestion. Store plugs in TE₅₀ buffer at 4°C until use (up to a month).

Size fractionate DNA by CHEF

21. Clean a 1.5-mm-thick, 20-well comb with 95% ethanol and cover four to six teeth in the middle of the comb with autoclave tape to create a wide tooth for preparative use.

The overall arrangement of cells is seven to eight small wells, a large preparative well, and seven to eight small wells.

22. In a small CHEF gel mold, pour a 1% agarose gel of sufficient thickness to cover the sample gel blocks when laid in along their length. Allow the gel to solidify ~1 hr at room temperature. While the gel is solidifying, pour 2 liters of 0.5× TBE buffer in the CHEF tank and equilibrate the unit at 14°C.
23. Gently remove the comb and array two agarose plugs or eight small agarose pieces in the large preparative slot. Load low-range PFG markers on each side of the preparative lanes, and fill the remaining space in the preparative and marker slots with molten 1% agarose.

There should be two empty wells between the markers and the preparative slots.

24. Size fractionate the DNA by three sequential stages of gel electrophoresis with one exchange of the polarity, as outlined in Figure 5.9.2. In stage 1, orient the gel so that the field forces the DNA to migrate from the wells toward the nearest gel edge (1 cm away from the wells; Fig. 5.9.2). Perform gel electrophoresis at 14°C for 6 hr using 5.0 V/cm, with a 15-sec pulse time at 120° angle in 0.5× TBE buffer.

This is the most important step for successful large insert cloning (see Critical Parameters and Troubleshooting, discussion of size fractionation). It might be necessary to optimize the voltage (4.5 to 5.5 V/cm) depending on the concentration of insert DNA. The insert DNA usually migrates slower than the marker DNA because of its higher concentration. The conditions should be optimized so that fragments <120 kb run out from the gel.

If small-fragment contamination is a problem, due to the very high concentration of DNA, the agarose plugs in the preparative slot can be removed and discarded after this stage of electrophoresis. The preparative slot should then be sealed with 1% molten agarose.

25. Change the electrophoresis buffer.
26. In stage 2, rotate the gel 180° in the tank and use the same electrophoresis conditions to bring all fragments remaining in the gel back to the original starting wells.
27. Apply new marker DNA to flanking wells that were not previously used.
28. In stage 3, resolve high-molecular-weight fragments at 6 V/cm for 16 hr with 0.1 to 40 sec pulse time.
29. Cut the outer lanes containing markers as well as 1 to 2 mm from each side of the preparative lanes to assess the success of the partial digestion. Wrap the remaining portions of the preparative lanes in plastic wrap and store at 4°C.

The lanes containing DNA for cloning must never come into contact with ethidium bromide solution or be exposed to UV light.

30. Stain the outer portions of the gel with ethidium bromide and examine with a fluorescent ruler on a digital imager to ascertain the size ranges.

The stained portion of the preparative lanes of genomic DNA should contain a broad smear extending from ~120 kb to >1 Mb. Most of the smaller fragments (<120 kb) should be run out of the gel during size fractionation.

Prepare size-fraction blocks

31. Slice genomic DNA lanes by cutting horizontally at 0.5-cm intervals to obtain size-fraction blocks in the range of 150 to 500 kb.

The lowest cutoff point is normally 4.5 cm below the well.

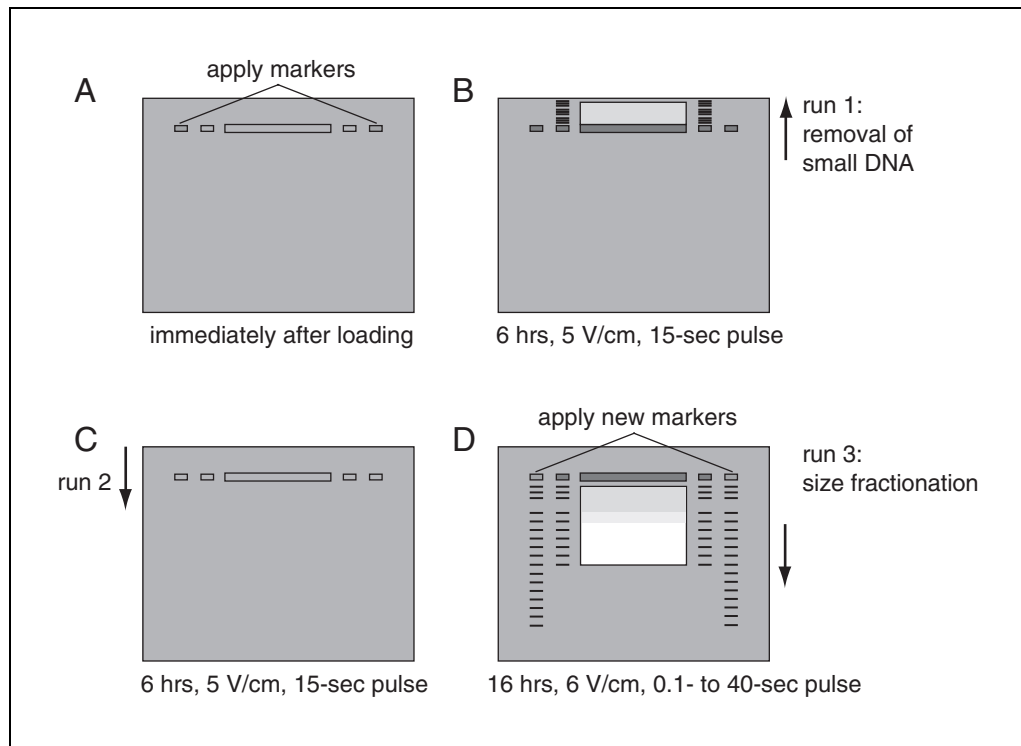


Figure 5.9.2 Isolation of insert DNA from agarose blocks by PFGE. **(A)** Agarose blocks containing partially digested DNA are applied to the center wells of a 1% agarose gel and low-range PFG markers are applied to flanking wells. **(B-D)** Size fractionation is performed in a CHEF apparatus. The DNA separation occurs in three stages, all with pulse directions at a 120° angle. **(B)** The initial direction of the field allows the DNA to migrate from the wells toward the nearest gel edge (1 cm away from the well), using 0.5× TBE at 14°C and 5.0 V/cm for 6 hr with 15-sec pulse time. **(C)** The same conditions are then used after turning the direction of the gel to bring all fragments remaining in the gel back to the original starting wells. **(D)** After the second run, new marker DNA is applied to additional flanking wells that were not previously used. The old markers are useful for checking the electrophoresis condition for the first run by comparing the new markers. The new markers are used to identify the size range after the final electrophoresis. High-molecular-weight fragments are then resolved in the third run, which is performed at 6 V/cm for 16 hr with 0.1- to 40-sec pulse time. After this procedure (not shown), the flanking marker lanes are removed from the gel and stained with ethidium bromide to indicate the location of the size ranges. Gel slices are then cut from the genomic DNA lanes by horizontal cuts at 0.5-cm intervals to obtain gel slices in the range of 150 to 500 kb.

32. Cut an ~1-mm-wide slice from each size-fraction block and store size-fractionated DNA blocks in 0.5 M EDTA at 4°C until use.

If eluting the DNA within a few days, it is possible to store the agarose slice in sterile 0.5× TBE at 4°C and omit the dialysis in step 36.

33. Load 1-mm slices directly into separate wells in 1% agarose gel and perform electrophoresis with 0.5× TBE buffer, using the following conditions in a CHEF apparatus: 120° angle, 6 V/cm for 16 hr with 0.1 to 40 sec pulse time at 14°C to assess the quality and size distribution of the DNA fragments in each block.
34. Stain the remaining gel pieces in 0.5 µg/ml ethidium bromide solution, and assemble the gel pieces on a digital imager. Take a picture with a fluorescent ruler to ascertain the size fractionation and cutoff point.

35. Stain the analytical gel from step 32 in 0.5 $\mu\text{g}/\text{ml}$ ethidium bromide solution and take a picture on the imager. Select one or two fractions from the blocks cut in step 33 that show 150- to 200-kb size range.

The slanted appearance of the DNA from left to right (or right to left) in each lane is a result of the orientation of high- and low-molecular-weight DNA in the 1-mm slice. The 1-mm slices contain higher-molecular-weight DNA at the top edge and lower-molecular-weight DNA at the bottom. As the 1-mm slices are rotated to load into the gel, the higher-molecular-weight DNA is on the left or right side of the lane, depending on which side of block the slice is taken from and which direction it is rotated.

36. Dialyze the selected size-fraction blocks against 15 ml sterile 0.5 \times TBE buffer for 3 hr.

If the sliced agarose block was stored in sterile 0.5 \times TBE buffer at step 32, omit this step.

Recover size-fractionated DNA by electroelution

37. Cut an \sim 10-cm-long piece of dialysis tubing and rinse it with sterile distilled deionized water. Close one end of the dialysis tubing with a dialysis clip and remove the residual water completely.

Use one piece of dialysis tubing per fraction block.

38. Place the fraction block and 300 μl sterile 0.5 \times TBE buffer in the dialysis tubing.
39. Completely remove air bubbles and seal the other end of the tubing with a clip. Orient the long axis of the gel parallel to the tubing.
40. Prepare 0.5 \times TBE buffer in a submarine gel electrophoresis tank and immerse the bag in a shallow layer of the 0.5 \times TBE. Place a plastic cover on the top of the clip to keep it down.

Additional dialysis bags can be electroeluted simultaneously.

41. Pass electric current through the dialysis bag(s) with 3 V/cm (100 V for the Sub-Cell GT DNA Electrophoresis Cell) for 3 hr.
42. Reverse the polarity of the current for 30 sec to release the DNA from the wall of the bag.
43. Transfer the bag (still containing the agarose block) to 1 liter TE buffer and dialyze at 4 $^{\circ}\text{C}$ for \geq 2 hr.

It is important to remove borate ions completely in this step, as they may inhibit the ligation reaction.

44. Open the bag and carefully transfer all of the solution to a fresh microcentrifuge tube using a wide-bore pipet tip.
45. Load 3 to 5 μl DNA into a 0.7% to 1.0% agarose gel in 0.5 \times TBE or 1 \times TAE buffer and perform electrophoresis with various amounts (5 to 50 ng) of lambda DNA. Estimate the DNA concentration using lambda DNA as a standard. Keep the eluted DNA at 4 $^{\circ}\text{C}$.

The partially digested genomic DNA fractions are ready to test in ligation reactions (see Basic Protocol). Use the eluted DNA within 10 days to prevent DNA deterioration and do not freeze the solution. The DNA concentration would be 2 to 20 ng/ μl in \sim 300 μl .

**MODIFIED ALKALINE LYSIS MINIPREP FOR RECOVERY OF DNA FROM
BAC/PAC CLONES**

It has been repeatedly observed that the standard alkaline lysis miniprep procedure gives a very low yield of plasmid DNA from BAC and PAC clones with large inserts. This procedure has thus been modified to improve yield. The most crucial factor appears to be increasing the dilution during lysis. Under the conditions described here, the yield of plasmid DNA is almost quantitative for clones with insert sizes ≤ 100 kb. The yield appears to be reduced for much larger clones, due either to trapping and loss in the debris or to the greater susceptibility to shearing. Clones with insert sizes ≤ 380 kb have been recovered in good yield with this protocol. This protocol yields DNA of sufficient quality for restriction endonuclease digestion and PCR analysis. Further purification may be necessary for other applications.

Materials

- BAC or PAC clones (see Basic Protocol)
- LB medium or terrific broth (*UNIT 1.1*) containing 25 $\mu\text{g/ml}$ kanamycin (for PAC clones) or 20 $\mu\text{g/ml}$ chloramphenicol (for BAC clones)
- Resuspension solution (see recipe)
- Alkaline lysis solution (see recipe)
- Precipitation solution (see recipe)
- Isopropanol
- 70% (v/v) ethanol
- TE buffer, pH 8.0 (*APPENDIX 2*)
- NotI* restriction endonuclease and buffer (New England Biolabs)
- Toothpicks, sterile
- 12- to 15-ml snap-cap polypropylene tubes
- Forceps, sterile
- Orbital shaker, 37°C
- 1.5-ml microcentrifuge tubes or 2-ml screw-cap tubes
- Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5A*)

Culture PAC or BAC clones

1. Using a sterile toothpick, inoculate a single isolated BAC or PAC clone into 2 ml LB medium or terrific broth containing either 25 $\mu\text{g/ml}$ kanamycin (PAC) or 20 $\mu\text{g/ml}$ chloramphenicol (BAC) in a 12- to 15-ml snap-cap polypropylene tube. Remove toothpick using sterile forceps. Grow overnight (≤ 16 hr) at 37°C in an orbital shaker at 200 rpm.
2. Remove snap cap from the tube and centrifuge 5 min at $1600 \times g$.

The temperature of this centrifugation is not critical.

Lyse cells

3. Discard supernatant and resuspend (vortex) pellet in 0.3 ml resuspension solution.
4. Add 0.3 ml alkaline lysis solution and gently shake tube to mix contents. Let stand at room temperature ~ 5 min.

The appearance of the suspension should change from very turbid to almost translucent.

5. While gently shaking, slowly add 0.3 ml precipitation solution. Stand tube on ice ≥ 5 min.

A thick white precipitate of protein and E. coli DNA will form.

6. Centrifuge 15 min at $16,000 \times g$, 4°C , to remove the white precipitate.

Recover plasmid DNA

7. Remove tube from centrifuge and place on ice. Using a micropipettor or disposable pipet, transfer supernatant to a 1.5-ml microcentrifuge tube containing 0.8 ml isopropanol. Mix by inverting a few times. Stand on ice ≥ 5 min.

Avoid transferring any of the white precipitate.

At this stage, samples can be stored overnight at -20°C .

8. Centrifuge 15 min at maximum speed in a microcentrifuge.

Temperature is not critical for this centrifugation, but better recovery may be achieved at room temperature.

9. Aspirate as much of the supernatant as possible. Add 0.5 ml of 70% ethanol and invert tube several times to wash the DNA pellet. Centrifuge 5 min at maximum speed, 4°C .

Occasionally, the pellet will become dislodged from the tube, so it is better to carefully aspirate the supernatant using a gel-loading pipet tip rather than pour it off.

10. Aspirate as much of the supernatant as possible. Air dry pellet at room temperature until the pellet becomes translucent.

When pellet changes from white to translucent, most of the ethanol is evaporated.

11. Resuspend in 100 μl TE buffer.

Do not use a narrow-bore pipet tip to mechanically resuspend DNA sample. Allow the solution to sit, occasionally tapping the bottom of the tube gently to mix the contents. For large BAC/PAC clones, resuspension may take >1 hr.

12. Use 5 to 10 μl DNA for digestion with 0.1 U *NotI*. Analyze insert size by pulsed-field gel electrophoresis.

*There are *NotI* sites flanking the SP6 and T7 promoter regions of the vectors; therefore, this is a very useful enzyme for analyzing insert size.*

*Use 18 μl DNA for digestion with more frequent-cutting enzymes such as *BamHI* or *EcoRI*.*

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Alkaline lysis solution

0.2 N NaOH

1% (w/v) SDS

Prepare before use

***EcoRI* endonuclease/methylase buffer, $10\times$**

100 μl 32 mM *S*-adenosylmethionine (New England Biolabs; final 0.8 mM)

80 μl 1 M MgCl_2 (final 20 mM)

800 μl 5 M NaCl (final 1 M)

2 ml 1 M Tris-Cl, pH 7.5 (APPENDIX 2; final 0.5 M)

40 μl 1 M dithiothreitol (DTT; final 10 mM)

980 μl sterile H_2O (total volume 4 ml)

Store up to 1 year at -80°C in small aliquots (0.21 ml)

Mbol buffer, 10× and 1×

For 10× buffer: Prepare 100 mM Tris·Cl, pH 8.0 (APPENDIX 2), containing 1 M NaCl and store up to 1 year at room temperature.

For 1× buffer: Dilute 1 part 10× buffer in 9 parts water and add 10 mM 2-mercaptoethanol. Prepare immediately before use.

Phenylmethylsulfonyl fluoride (PMSF) solution, 100 mM

Prepare a 100 mM stock solution in isopropanol and store up to 1 year at −20°C. Mix well by vortexing and dilute as appropriate immediately before use.

CAUTION: PMSF is hazardous. See APPENDIX 1H for guidelines on handling, storage, and disposal.

Precipitation solution

3 M potassium acetate, pH 5.5

Autoclave

Store indefinitely at 4°C

Proteinase K lysis solution

10 ml filtered 10% *N*-lauroylsarcosine (sodium salt; Sigma; final 2%)

40 ml 0.5 M EDTA, pH 8.0 (APPENDIX 2; final 0.4 M)

100 mg proteinase K (Boehringer Mannheim; final 2 mg/ml)

Prepare fresh and use immediately after adding proteinase K

RBC lysis solution, 10×, 1×

9.54 g NH₄Cl (1.78 M final)

0.237 g NH₄HCO₃ (0.03 M final)

Dissolve in 100 ml

Sterilize using 0.2-μm filter

Store up to 1 month at 4°C in a tightly closed container. Dilute to 1× concentration with sterile distilled deionized water before use.

Resuspension solution

15 mM Tris·Cl, pH 8.0 (APPENDIX 2)

10 mM EDTA

100 mg/ml DNase-free RNase A (UNIT 3.13)

Sterilize using a 0.2 μm filter

Store up to 6 months at 4°C

T4 DNA ligase buffer, 5×

2 M Tris·Cl, pH 7.6 (APPENDIX 2)

50 mM MgCl₂ (APPENDIX 2)

5 mM ATP

5 mM dithiothreitol (DTT; APPENDIX 2)

25% (w/v) polyethylene glycol 8000 (PEG 8000)

Store up to 1 year at −20°C

COMMENTARY**Background Information****Large-insert cloning**

Large-insert clones are of particular importance for preparing long-range physical maps consisting of clusters of contiguous clones (contigs). In this respect, YAC clones filled a

vital role in the beginning of various initiatives for the mapping of very complex genomes, such as that of humans. In recent years, it has become increasingly important to have clones available that represent the genome in a contiguous and nonrearranged format with better fidelity than YAC clones, that exhibit fewer

stability problems and fewer cloning artifacts. This has become important for two reasons: to have good intermediates for shotgun sequencing of large genomic regions and to have clones containing unaltered complete genes for functional analysis. In addition, much lower levels of cloning artifacts should facilitate positional cloning efforts, because the clones can be used without much verification for exon trapping, for hybrid selection, and for other gene-hunting procedures.

In view of these considerations, there has been an effort in various laboratories to develop bacterial plasmids with inserts considerably larger than can be contained in cosmids (35 to 45 kb). An elegant packaging system based on bacteriophage P1 and corresponding vectors was developed by Sternberg and collaborators (Sternberg, 1990; Pierce et al., 1992); however, the size of the bacteriophage particle restricts the maximum insert to ~95 kb, to fit within the total capacity of 110 kb for vector plus insert. A very useful expansion of this work was the cloning system developed by Shizuya et al. (1992), which combined the high transformation efficiencies obtainable by electroporation with a plasmid vector based on the F factor replicon. Insert sizes as large as 300 kb were originally reported; however, most clones have inserts closer to 100 kb in size. Because the F replicon provides for about one copy per cell, like the *Escherichia coli* chromosome, the resulting clones have been designated as bacterial artificial chromosome (BAC) clones. Some useful genomic libraries have been developed for human and mouse (see Chapter 5 introduction), various plant species (Wang et al., 1995; Woo et al., 1994), and hybrid cells (Wang et al., 1994).

The hybrid cloning system developed by Ioannou et al. (1994), which was based on a combination of the original BAC cloning procedures and P1 vectors, was thus designated P1-derived artificial chromosome (PAC) cloning. The PAC system permits generation of large-insert plasmids with insert sizes similar to BACs. In the BAC system, recombinant clones are picked from a mixture of colonies based on screening for recombinant clones using probe hybridization or a blue/white *lacZ*- α color assay. PAC clones are mostly recombinant as a consequence of positive selection for insert-containing clones, which is based on sucrose-sensitivity of clones lacking inserts. All of the procedures for screening and characterizing P1 clones are identical for PACs.

Since the pCYPAC vectors were directly derived from pAd10SacBII, these vectors share many features, including an inducible plasmid copy number and positive selection for inserts. The pCYPAC2 vectors have been used successfully to prepare a 16-fold redundant human PAC library (de Jong et al., unpub. observ.). The first 3.5-fold redundant segment of this library, designated RPCI-1, has been made available to over 40 genome centers, resource centers, and commercial screening companies (see Chapter 5 introduction). To facilitate the use of large-insert bacterial clones for functional studies, the authors have prepared a new PAC vector, designated pPAC4, with a dominant selectable marker gene (blasticidin gene), an EBV replicon, and modification sites (E. Frengen, P.J. de Jong, et al., unpub. observ.). The pPAC4 vector has been applied to generate several libraries, including a four-fold redundant human PAC library with average inserts of 135 kb (P.J. de Jong, et al., unpub. obser.), a 12-fold redundant mouse library (K. Osoegawa et al., unpub. observ.), and a 10-fold redundant rat library (Woon et al., 1998).

A new BAC vector, named pBACe3.6, has been developed at Roswell Park Cancer Institute (RPCI; E. Frengen et al., in press) and has the same system for selecting recombinant clones. BAC clones produced using the pBACe3.6 vector are mostly recombinant as a consequence of positive selection for insert-containing clones, which is based on sucrose-sensitivity of clones lacking inserts instead of using probe hybridization or a blue/white *lacZ*- α color assay. Sucrose is much less expensive than isopropyl- β -D-thiogalactopyranoside (IPTG) and Xgal. Also, it is unnecessary to distinguish recombinant and nonrecombinant clones for colony picking. Although it is possible to distinguish blue and white colonies by a colony-picking machine, the picking error is still high (~15%). From these points of view, it is considered that the pBACe3.6 is superior to the other BAC vectors.

It has been difficult to prepare highly redundant BAC or PAC libraries with average insert sizes larger than 150 kb. A new procedure for preparation of BAC and PAC libraries has been improved at RPCI (Osoegawa et al., 1998). Included in this unit is the improved protocol that enables the generation of highly redundant BAC/PAC libraries with larger and relatively uniform insert sizes. Better vector preparation protocols and improved DNA-sizing procedures minimize the generation of nonrecombinant clones and the number of small-insert

clones. Efficient methods for recovering insert DNA by electroelution and for concentrating ligation products without DNA deterioration result in higher transformation efficiencies. These factors are discussed in greater detail below. The improved methods have been applied at RPCI to prepare PAC and BAC libraries from human, murine, rat, canine, baboon, bovine, feline, silkworm, and fruit fly DNA, with average insert sizes ranging between 160 and 235 kb. It is now possible to construct highly redundant BAC/PAC libraries from various organisms applying this improved approach within a time period of a few months.

Choice of host strains

A critical aspect of any large-insert cloning endeavor is the choice of the host. It is important that inserts be cloned with little bias for sequence content, that most or all sequences be cloneable without artifacts, and that the sequences be maintained without rearrangements. Furthermore, clones should grow well and be easy to maintain or revive after freezing. Finally, the particular host strain(s) should be compatible with the cloning procedures and provide high transformation efficiencies with large-insert DNA ligation products, with minimal bias in favor of smaller inserts. Of course, this is a utopia and none of the available hosts exhibit all of these desirable features; however, it is important to understand why one particular host (DH10B) was selected for construction of PAC (Ioannou et al., 1994) and BAC libraries (Shizuya et al., 1992). This will allow users of this technology to make informed decisions for selecting alternative hosts. Extensive reviews on *E. coli* host strains for cloning have been published elsewhere (Hanahan et al., 1991; Raleigh, 1987; Wyman and Wertman, 1987).

The most important requirement for large-insert libraries is that the recombinant DNA be stable, so that multiple researchers can use the same clones over and over again. This is particularly important for libraries used as a reference resource by the scientific community or used for DNA sequence analysis of large genomic regions. The most important host gene for stabilizing inserts is *recA*. A defective *recA* allele is an absolute requirement for large-insert propagation with minimal chance of rearrangement. The *recA* gene product fulfills a central role in homologous recombination in *E. coli*, and a defective allele appears to stabilize, in particular, direct repeat structures in the insert DNA (Wyman and Wertman, 1987). The DH10B host used for PAC and BAC library

construction carries the stringent *recA1* allele (Hanahan et al., 1991).

Other stabilizing mutations in the host recombination pathways have been reported. In particular, defects in one or more of the *recB*, *recC*, or *recD* genes have been correlated with higher stability of palindromic DNA sequences (Wyman and Wertman, 1987); however, most of these studies have been done in double-mutant strains that are also mutated in *recA*. Such strains usually grow poorly and often accumulate revertants or pseudorevertants that grow better. These revertants have been designated as *sbc* mutants (suppressor of the *recBC* phenotype). A consequence of the reversion of the poor growth phenotype in *recBC* mutants is that some of the stability correlations previously attributed to *recB* or *recC* alleles may actually be the result of *sbc* mutations. Supporting this view, studies by Leach and co-workers (Chalker et al., 1988; Gibson et al., 1992) demonstrated that *sbcC* and *sbcD* mutations can stabilize palindromic DNA sequences irrespective of the *recBC* alleles present. All of the studies correlating the stability of particular DNA sequences with host mutations have been undertaken with earlier vector systems, mostly using the lambda cloning system. Thus, it remains to be demonstrated how these mutations affect the stability of large inserts in low-copy-number plasmid systems based on the P1 or F replicons (PACs or BACs).

Eukaryotic DNA sequences are methylated at the cytosine positions in many CpG dinucleotide sequences. The levels of methylation are highly variable and depend on sequence context and DNA source; therefore, any hosts that preferentially restrict DNA sequences containing 5-methylcytosine will result in a cloning bias with respect to sequence content. In addition, longer inserts will, in general, contain more methylated sequences, thus biasing against establishing larger-insert clones. Several host systems (*mcr*) restricting DNA containing 5-methylcytosine have been reported: *mcrA*, *mcrB*, and *mcrC* (for review, see UNIT 1.4 and Raleigh, 1987). These restriction pathways have different specificity with respect to the sequences flanking the methylated CpG sites (Hanahan et al., 1991). Indeed, mutations in these pathways greatly improve the cloning efficiencies for methylated DNA sequences and improve cloning large DNA inserts in cosmids and P1 clones (UNIT 5.3). With respect to BAC or PAC libraries, it is very likely that cloning success depends on the presence of the *mcr* mutations. The DH10B host contains an *mcrA*

mutation and a deletion of the *mcrBC* genes (Grant et al., 1990). The *mcr* mutations are important only for the initial entry of the foreign DNA sequences into *E. coli*. After propagation in *E. coli*, these sequences will lose their methylation patterns and the *mcr* status of the host becomes irrelevant. Hence, to stabilize an unstable clone, the PAC or BAC DNA can be transferred to any host strain with different recombination deficiencies, irrespective of the *mcr* status.

Large-insert cloning by the PAC or BAC systems requires that the host strain provide good transformation efficiencies for large-insert DNA. Not only should the host be highly electrocompetent, it should also have minimal bias in favor of shorter DNA inserts. The maximum efficiency of transformation for small plasmids is a factor of the genetic background of the host, which is not well understood. Derivatives of the MC1061 strain have, in general, been very good for use in chemical transformation procedures. The DH10B strain has been the best host for electrotransformation of small plasmid DNA molecules. A host mutation (*deoR*) that allows *E. coli* to grow on medium containing inosine as the sole carbon source appears to have a fortuitous effect on plasmid transformation using chemical procedures (Hananan et al., 1991). Small plasmid transformation for pUC and pBR322 was reportedly ~4-fold higher in a *deoR* derivative and ~30-fold higher for a larger 66-kb plasmid. The same *deoR* mutation also appears to have an important impact on cloning success using electroporation procedures, and it is, indeed, part of the DH10B genetic background (Grant et al., 1990).

In view of all previous considerations, DH10B appears to be a good host for general cloning with PAC and BAC vectors; however, a small percentage of PAC and BAC clones are nonetheless unstable in DH10B, and it would be desirable to have a strain available that exhibits increased stability while maintaining the viability and electrotransformation competency of the DH10B host.

Vector elements

The pCYPAC1 and pCYPAC2 vectors (Ioannou et al., 1994) are highly similar to the earlier P1 vector pAd10SacBII. The main differences of the pCYPAC vectors are the absence of the adenovirus-derived stuffer fragments and the presence, at a different location, of removable pUC plasmid sequences. Both pUC and the adenovirus sequences are re-

moved during the construction of recombinant clones. Hence, the differences in recombinant DNA structure disappear during cloning. As a consequence of the different cloning approaches, insert sizes for PACs can exceed the limits imposed on P1 clones by the packaging approach. In addition, as a consequence of the different transformation protocols (electroporation versus packaging and viral particle DNA injection), different host strains are used for P1 clones and PAC clones. In all other aspects, PAC clones can be used identically to P1 clones. Much as with P1 clones, PAC copy number can be enhanced using IPTG to inactivate the *lac* repressor, thus inducing the high-copy-number P1-lytic replicon mode. In the absence of IPTG, the low-copy-number P1-plasmid replicon is used for maintenance of the PAC DNA; however, it is advisable not to add IPTG in culture medium, because the PAC clone is maintained more stably as a single copy than in multiple copies.

The pPAC4 vector has been developed for the purpose of using PACs for functional analysis of genes carried on cloned inserts. Four elements are identical in the pPAC4 and the pCYPAC vectors: the P1 plasmid replicon, the kanamycin resistance gene, the *sacB* gene, and the pUC-link. The IPTG inducible P1 lytic replicon has been removed in this vector. The authors have observed that the quantities of recombinant DNA isolated from PAC clones constructed in the pPAC4 vector are 3- to 4-fold lower compared to yields from pCYPAC-based clones. This is probably because the P1 lytic replicon in the pCYPAC vectors is partly functional (due to a leaky *lac* repressor function), resulting in 3 to 4 copies of PAC clones per cell. In pPAC4-based recombinant molecules, however, the P1 plasmid replicon is the only replicon present, ensuring that the copy number is maintained at one per cell.

The pBACe3.6 vector carries the F plasmid replicon and the chloramphenicol resistance gene from the BAC vector pBAC108L (Shizuya et al., 1992). pBACe3.6, presenting the advantage of positive selection in a relatively small modular vector, is an excellent tool for the construction of the complex libraries needed for large-scale mapping and sequencing. Similar to the PAC vectors, this vector contains a modified pUC19 plasmid located within the *sacB* gene. The pUC-link is flanked by multiple cloning sites, allowing positive selection for recombinants by using any of the restriction enzymes *Bam*HI, *Sac*II, *Eco*RI, *Sac*I, *Mlu*I, or *Nsi*I (Fig. 5.9.1). Support Proto-

col 4 makes use of partial *EcoRI* digestion, controlled by *EcoRI* methylase, in the preparation of insert DNA for cloning in the pBACe3.6 vector.

The attachment site for the Tn7 transposon (Tn7att), which has been introduced into pPAC4 and pBACe3.6 vectors, facilitates highly specific Tn7-based retrofitting of BAC/PAC clones. This allows the engineering of clones—e.g., prior to transfection of BAC/PAC clones into mammalian cells for functional studies. In addition, sequences for linearization of the BAC/PAC clones facilitate large-scale mapping (e.g., optical mapping studies; Schwartz et al., 1993). The specific linearization of BAC/PAC clones can be achieved by the intron-homing endonuclease *PI-SceI*, or conceivably by using linear *loxP* fragments in Cre-mediated recombination with the corresponding *loxP* sites in the vector. Furthermore, the mutant *loxP* site *loxP*⁵¹¹, which recombines inefficiently with the wild-type *loxP* site, will also allow Cre-mediated exchange of the vector modules in selected BAC/PAC clones.

In order to facilitate the use of PAC clones for transfection into mammalian cells, two new genetic elements have been included in the pPAC4 vector. The *bsrr* cassette contains the blasticidin deaminase gene, which is placed in a eukaryotic expression cassette to enable selection of the desired mammalian clones by using the nucleoside antibiotic blasticidin. In addition, the Epstein-Barr virus latent origin of replication, *oriP*, has been included to ensure stable extrachromosomal maintenance of the PAC clones in mammalian cells. Furthermore, the pPAC4 vector has an hCMV-promoter-ATG-*loxP* targeting element that facilitates Cre-mediated integration of the cloned inserts into *loxP* target sites in mammalian chromosomes (Fukushige and Sauer, 1992). In vivo site-specific recombination with a complementary *loxP* site in the mammalian genome, coupled to a truncated promoterless *neo* gene, allows activation of G418 resistance under control of the hCMV promoter and ATG start codon. Thus, the PAC clones have the potential of being transferred either to mammalian cells or to the germ line of mutant mice as a replicon or integrated form in a specific chromosome site, allowing molecular complementation of a mutant phenotype. This would be a useful strategy for the identification of new genes or validation of genes or gene complexes present on the PACs by functional cloning.

The vectors contain ampicillin and kanamycin (the PAC vectors) or chloramphenicol (the pBACe3.6 vector) resistance genes, but in the recombinant clones only kanamycin or chloramphenicol resistance is maintained. Because of the low copy number of the P1 or F plasmid replicon (one copy per cell), including a removable high-copy-number replicon (pUC-link) in the BAC/PAC vectors facilitates obtaining large amounts of vector DNA. Single digestion of the vector with *BamHI* (PAC and BAC vectors) or *EcoRI* (BAC vector) is required to separate the P1 or F plasmid replicon fragment from the smaller stuffer fragment generated from the pUC-link plasmid. Adequate removal of the pUC-link sequences is essential to prevent interference with the library construction. The pUC-link plasmid also serves another role, which is the appropriate disruption of the *sacB* gene to keep strains containing the vector viable. Expression of the *sacB* gene in the presence of sucrose leads to the production of a toxic metabolite of sucrose; however, even in the absence of sucrose, considerable toxicity can be observed when the *sacB* gene is expressed (the probable mechanism is unknown). In recombinant clones, foreign DNA sequences are substituted for the pUC-link sequences. Hence, recombinant clones can be grown on agar plates with kanamycin and sucrose. Under such conditions, noninsert clones generated by direct ligation of the *BamHI* or *EcoRI* ends should not grow.

Another aspect of P1, PAC, and pBACe3.6 vectors is the presence of the *sacB* positive selection system. It has been observed that $\geq 30\%$ higher cloning efficiencies are obtained when sucrose selection is not applied. Even in the absence of sucrose selection, it is possible to obtain PAC libraries containing mostly recombinant clones; however, the number of noninsert clones certainly increases more than the number of recombinant clones, and sucrose selection is thus preferred. Does the increased number of recombinants reflect a cloning bias, which is removed when sucrose is not applied? Yes, many recombinant clones obtained without sucrose selection ($\sim 30\%$ in a particular human PAC library; P.A. Ioannou and P.J. de Jong, unpub. observ.) are sensitive to sucrose. Most likely, such clones are the result of fortuitous *E. coli* promoter-like sequences in the inserted DNA; these promoter-like sequences drive expression of the *sacB* gene. The occurrence in eukaryotic DNA of sequences that can function as promoter elements in bacteria has previously been reported (Gibson et al., 1987).

It is, however, quite unlikely that this bias presents a serious problem for cloning contiguous sections of eukaryotic genomes. It is likely that most promoter-like elements will be too far removed from the vector/insert junction and will, therefore, have no effect on *sacB* expression. Because of the approach used—random shotgun cloning by partial digestion—it is likely that each region has many opportunities to be cloned with different endpoints, and in most cases the promoter-like elements will not interfere. This notion is supported by the observation that it is quite easy to establish PAC contigs over large genomic regions. Gaps appear not to be present above statistical expectations.

Critical Parameters and Troubleshooting

Although BAC and PAC cloning is conceptually easy, there are several parts of the procedure that are critical for success in preparing BAC and PAC libraries.

Genomic DNA preparation

Genomic DNA fragments are generated by partial *Mbo*I or *Eco*RI digestion of DNA embedded in agarose. A pre-electrophoresis step before partial digestion removes inhibitors of the cloning process, which permits more homogeneous partial digestion. Over- or underdigestion of the genomic DNA can greatly affect the subsequent cloning efficiency of large DNA fragments, either by presenting many smaller and better cloneable fragments or by creating concentrations of desirable size fragments too low for efficient cloning.

Vector preparation

It is critical that some of the restriction and dephosphorylation steps for preparing the vector DNA be optimized carefully. The *Bam*HI or *Eco*RI digestion should be done quite carefully and overdigestion should be avoided. If a large fraction of noninsert clones is found upon library construction, then it is very probable that overdigestion with *Bam*HI or *Eco*RI resulted in cutting at a secondary *Bam*HI or *Eco*RI star site. If the alkaline phosphatase treatment is not complete, then the vector can religate between a *Bam*HI site and a *Bam*HI star site, or between an *Eco*RI site and an *Eco*RI star site, resulting in small deletions in the vector and affecting the expression of the *sacB* gene. Deletions around the *Bam*HI or *Eco*RI site thus result in sucrose resistance similar to that observed upon

insertion of foreign sequences into the cloning site. There are two ways to diminish the abundance of such undesirable noninsert clones in enzymatic manipulation: improving dephosphorylation of the restriction site (and restriction star) ends or lowering the level of enzyme in restriction endonuclease digestion. It has often been observed that excessive alkaline phosphatase treatment results in lower yields of desirable ligation products, perhaps through damage to the restriction ends; therefore, it is important to suppress *Bam*HI or *Eco*RI star activity as much as possible. Using minimal levels of *Bam*HI (two or three fold over digestion in optimal *Bam*HI buffer), it is possible to prepare libraries in which $\leq 1\%$ of clones lack inserts. Thus, each step in preparing the vector for cloning requires optimization by titration of restriction enzyme and alkaline phosphatase treatments, such that low nonrecombinant background levels at maximum cloning efficiency are obtained.

An alternative protocol to achieve the removal of deleted, still phosphorylated, vector molecules prior to the ligation step is implemented to circumvent this labor-intensive process. In this procedure, the vector DNA is simultaneously treated with suboptimized restriction enzymes and alkaline phosphatase. Then, vector preparation is accomplished by ligating the restricted, alkaline phosphatase-treated vector DNA and reisolating the majority of the nonligated monomeric vector DNA through pulsed-field gel electrophoresis. The background of ligated DNA is easily removed by exploiting the different mobility of circular and dimeric linear DNA from linear monomeric DNA on pulsed-field gels. The sizing step will also remove most of the nonligated, deleted vector molecules if the deletions are substantial.

Other clones lacking inserts found in the RPCI PAC and BAC libraries contain the complete vector sequence. Such clones are resistant to sucrose because the pUC-link sequences are present between the promoter and *sacB* open reading frame sequences and prevent *sacB* expression. The identity of these colonies can easily be confirmed by their ability to grow on ampicillin. If digestion and dephosphorylation are properly optimized, 0.01% to 0.05% of clones should be expected to be ampicillin resistant. The presence of such background clones has consequences for screening libraries by DNA probe hybridization. If the probes are derived from plasmid clones, these nonrecombinant clones will also frequently hybridize and

may initially mask the true probe-positive clones. If properly recognized, these nonrecombinants can facilitate rather than confuse screening by allowing better alignment between the autoradiograms and the colony patterns. To eliminate possible confusion, each library should be screened initially with a pUC-derived probe to establish the location of all nonrecombinant clones in the arrayed PAC or BAC library. These clones can then be inactivated or their hybridization results can be discounted in subsequent hybridization with different probes.

Size fractionation

Size fractionation of the DNA is the most critical factor affecting successful cloning of large inserts. RPCI human PAC libraries have been prepared using a single sizing procedure for partially digested genomic DNA. These libraries have a rather wide distribution of insert sizes, at or below the fragment size range selected prior to cloning. The average insert sizes are 110 to 120 kb. A paradoxical decrease of average clone insert sizes has been observed when larger fragment sizes were selected for cloning. The sharp decrease in cloning efficiency with increasing fragment size is likely the result of a conglomerate of different factors. It can be hypothesized that smaller-than-selected insert fragments result from contamination of small restriction fragments combined with preferential ligation, a lower transformation efficiency of larger ligation products, and a lower circularization rate for the larger plasmids. In addition, larger fragments are more susceptible to damage (nicks, gaps, or adducts) in the genomic DNA, thus decreasing their cloning success; therefore, it is critical to remove as many small DNA fragments as possible by using a size fractionation step. Small fragments that comigrate with large DNA fragments can be removed more completely by subjecting the DNA to a pseudo-double size fractionation procedure (Osoegawa et al., 1998). The double sizing procedure provides a more uniform insert DNA size than a single sizing procedure. The removal of small fragments and improved integrity of large fragments has made it feasible to routinely generate clone libraries with inserts averaging 160 to 180 kb, or even having inserts in the range of 220 to 280 kb.

Electroelution and electrophoresis buffer

Size-fractionated DNA fragments are recovered by electroelution at improved DNA integ-

riety and with higher recovery efficiency as compared to DNA extraction from gel slices by β -agarase treatment. These observations are consistent with a report by Strong et al. (1997). A further advantage of the electroelution method is the ability to use regular high-melting-point agarose, which facilitates gel handling and exhibits better resolving capabilities than the low-melting-point agarose required for enzyme-based DNA extraction. The TAE buffer system has previously been employed in the electrophoretic separation of partially digested DNA, as borate ions may inhibit the ligation reaction used in the construction of BAC and PAC libraries (Ioannou et al., 1994; Woo et al., 1994; Strong et al., 1997); however, large DNA fragments are optimally resolved on pulsed-field gels in 0.5 \times TBE buffer. Since it is desirable to achieve the best separation possible, 0.5 \times TBE buffer is used and the borate ions are easily removed after electroelution by dialysis against TE prior to ligation.

Concentration of ligation mixture

As described below, another important element is the transformation efficiency for construction of highly redundant libraries from complex genomes. A major limitation with respect to constructing BAC/PAC-based libraries for complex genomes is the number of transformations required to generate the library. For instance, to prepare a 20-fold redundant collection of PAC clones for the human genome, ~10,000 electroporation reactions were performed (P.J. de Jong et al., unpub. data). The laborious aspect of this work and the cost of commercial electrocompetent cells created an incentive to increase the number of colonies generated per electroporation to reduce the number of transformations. Previously, ligation products were dialyzed on a floating dialysis membrane against a low salt buffer to reduce heat development during the subsequent electrical pulse (Ioannou et al., 1994). This usually resulted in a two-fold volume increase of the ligation reaction from 50 μ l to ~100 μ l. The increase is due to the higher osmotic pressure in the ligation buffer, which contains PEG 8000, as compared to the dialysis buffer. The use of 30% PEG 8000 in the dialysate in order to obtain higher concentrations of ligation products allows higher transformation efficiencies, resulting in a 6- to 12-fold reduction in the number of electroporations required to generate a library compared to previous efforts (Osoegawa et al., 1998).

Electroporation

Another critical factor is the efficiency of transformation that can be obtained with a particular batch of DH10B host cells. It is the experience that the procedures to prepare competent host batches described by Hanahan et al. (1991) result in high transformation efficiencies. Nevertheless, it is hard to meet the level of electrocompetency obtained for DH10B by commercial suppliers, such as Life Technologies. It is advisable, therefore, to use commercially available electrocompetent DH10B cells if large libraries from complex genomes need to be constructed. It may be advisable to buy samples of an electrocompetent DH10B cell batch for testing in electroporation of large and small plasmids. If the sample passes the test, then a sufficient number of vials from the same batch should be bought, so that the entire library can be completed with a single batch of competent cells. In order to standardize the testing of such batches, it is good to have a stable mixture of small plasmids (pUC19, with ampicillin resistance) and large plasmids (a selected 150-kb PAC, with kanamycin resistance). The large plasmids degrade during storage in a refrigerator and even more during freezing and thawing procedures. It is thus best to prepare a mixture of the two plasmids and freeze many aliquots at -70°C and utilize each aliquot only once for testing. If the large plasmid degrades, then this will be evident from the change in the relative ratio of ampicillin- versus kanamycin-resistant colonies.

It has been reported by Sheng et al. (1995), and confirmed by the authors of this unit, that lower voltage gradients during electroporation decrease the transformation efficiencies of plasmids; however, the decrease is most significant for small plasmids, thus improving the relative transformation efficiency of large plasmids. As a result, the optimum voltage gradient should be in the range of 9 to 12 kV/cm and not in the higher range recommended by the suppliers of electrocompetent bacteria.

Anticipated Results

Under optimal conditions, one can obtain as many as 1800 transformants per 2 μl dialyzed solution with an average insert size of 160 kb, and as many as 250 transformants per 2 μl at 230 kb for BAC. Due to the low DNA concentrations in the ligation, a typical 10-fold redundant mammalian BAC or PAC library requires anywhere from a few hundred up to 2000 electroporations. Each electroporation should result in 300 to 2000 colonies for fragments in

the 150- to 180-kb range. Control plates generated with ligations lacking genomic DNA typically show less than ten colonies, provided that the vector is prepared well. A sharp dropoff in cloning efficiency for larger fragments (>200 kb) is always observed. This results from at least three factors: (1) DNA fragments and circular ligation products being increasingly more fragile, subject to shearing or nuclease action; (2) a lower efficiency in the formation of circular ligation intermediates; and (3) a dropoff in transformation efficiency of plasmids with increasing size, expressed as a molecular efficiency, and becoming more pronounced when expressed per DNA weight. It is imperative that highly competent bacterial cells are used to ensure that reasonable numbers of clones are obtained per electroporation. As a consequence of the dropoff in cloning efficiency with size, it is difficult to prepare BAC and PAC libraries with insert sizes averaging in 200 to 250 kb, and nearly impossible to construct highly redundant libraries with an average size >250 kb using the presented technology.

Once a provisional BAC or PAC library has been constructed, it is essential to characterize the clones and isolate plasmid DNA. The use of Support Protocol 5 or an automatic plasmid isolation machine (AutoGen 740) will result in ~ 0.5 to 1 μg PAC DNA (170 kb) per 1.5 ml of overnight culture (not induced with IPTG). When the pCYPAC2, pPAC4, or pBACe3.6 vector is used for library construction, two *NotI* sites flank the insert in recombinant plasmids; therefore, most clones should show only two bands when analyzed by PFGE: the vector (at ~ 15.5 kb for pCYPAC2, 16.8 kb for pPAC4, and 8.7 kb for pBACe3.6) and the insert of variable size from clone to clone. Additional bands may be seen if the insert contains internal *NotI* sites. Incomplete digestion usually leads to a very characteristic doublet, where the two bands differ by the size of the vector.

Time Considerations

It takes several days to prepare agarose-embedded high-molecular-weight DNA starting from a blood or animal tissue sample and at least another week for testing various conditions of partial digestion and analyzing the results using pulsed-field gel electrophoresis (weeks 1 and 2). Ligations and transformations take 3 to 5 days, and another 3 to 4 days are required to characterize candidate colonies by pulsed-field gel electrophoresis of *NotI*-restricted plasmids (weeks 2 and 3). If the par-

ticular ligation gives desirable results (good cloning efficiency, low nonrecombinant levels, and large inserts), another week (week 4) should be dedicated to the 200 to 2000 electroportations required for a 10-fold redundant mammalian BAC or PAC library. It is often convenient, perhaps necessary, and possibly desirable to take a break at this point. The arraying of the BAC or PAC libraries can be done at a different time and place, as it requires other facilities. Picking of colonies (e.g., 100,000) will take about one week and will require at least two people to prepare and label microtiter dishes and spread all the cells, provided that there is access to a facility equipped with microtiter plate fillers and a colony-picking robot. Hence, under the most optimal laboratory conditions and with the appropriate skills, a 10-fold redundant mammalian BAC or PAC library can be constructed in 1 to 3 months, requiring at least two people; however, it is more realistic to expect that most of the time will be invested in acquiring the skills and optimizing the various critical parameters. Anytime between 6 months and a year should be dedicated by a single person, who should be able to count on assistance from others from time to time.

Literature Cited

- Chalker, A.F., Leach, D.R., and Lloyd, R.G. 1988. *Escherichia coli* sbcC mutants permit stable propagation of DNA replicons containing a long palindrome. *Gene* 71:201-205.
- Chu, G., Vollrath, D., and Davis, R.W. 1986. Separation of large DNA molecules by contour-clamped homogeneous electric field. *Science* 234:1582-1585.
- Donovan, J. and Brown, P. 1995. Euthanasia. In *Current Protocols in Immunology*. (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevock, and W. Strober, eds.) pp. 1.8.1-1.8.4. John Wiley & Sons, New York.
- Frengen, E., Weichenhan, D., Zhao, B., Osoegawa, K., van Geel, M., and de Jong, P.J. 1999. A modular, positive selection bacterial artificial chromosome vector with multiple cloning sites. *Genomics*. In press.
- Fukushige, S. and Sauer, B. 1992. Genomic targeting with a positive-selection lox integration vector allows highly reproducible gene expression in mammalian cells. *Proc Natl Acad Sci U.S.A.* 89:7905-7909.
- Gibson, T.J., Coulson, A.R., Sulston, J.E., and Little, P.F. 1987. Lorist2, a cosmid with transcriptional terminators insulating vector genes from interference by promoters within the insert: Effect on DNA yield and cloned insert frequency. *Gene* 53:275-281.
- Gibson, F.P., Leach, D.R., and Lloyd, R.G. 1992. Identification of sbcD mutations as cosuppressors of recBC that allow propagation of DNA palindromes in *Escherichia coli* K-12. *J. Bacteriol.* 174:1221-1228.
- Grant, S.G., Jessee, J., Bloom, F.R., and Hanahan, D. 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U.S.A.* 87:4645-4649.
- Hanahan, D., Jessee, J., and Bloom, F.R. 1991. Plasmid transformation of *Escherichia coli* and other bacteria. *Methods Enzymol.* 204:63-113.
- Ioannou, P.A., Amemiya, C.T., Garnes, J., Kroisel, P.M., Shizuya, H., Chen, C., Batzer, M.A., and de Jong, P.J. 1994. A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nature Genet.* 6:84-89.
- Osoegawa, K., Woon, P.-Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J., and de Jong, P.J. 1998. An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 52:1-8.
- Pierce, J.C., Sauer, B., and Sternberg, N. 1992. A positive selection vector for cloning high molecular weight DNA by the bacteriophage P1 system: Improved cloning efficacy. *Proc. Natl. Acad. Sci. U.S.A.* 89:2056-2060.
- Raleigh, E.A. 1987. Restriction and modification in vivo by *Escherichia coli* K12. *Methods Enzymol.* 152:130-141.
- Schwartz, D.C., Li, X., Hernandez, L.I., Ramnarain, S.P., Huff, E.J., and Wang, Y.K. 1993. Ordered restriction maps of *Saccharomyces cerevisiae* chromosomes constructed by optical mapping. *Science* 262:110-114.
- Sheng, Y.L., Mancino, V., and Birren, B. 1995. Transformation of *Escherichia coli* with large DNA molecules by electroporation. *Nucl. Acids Res.* 23:1990-1996.
- Shizuya, H., Birren, B., Kim, U.-J., Mancino, V., Stepak, T., Tachiiri, Y., and Simon, M. 1992. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an f-factor-based vector. *Proc. Natl. Acad. Sci. U.S.A.* 89:8794-8797.
- Sternberg, N. 1990. Bacteriophage P1 cloning system for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs. *Proc. Natl. Acad. Sci. U.S.A.* 87:103-107.
- Strong, S.J., Ohta, Y., Litman, G.W., and Amemiya C.T. 1997. Marked improvement of PAC and BAC cloning is achieved using electroelution of pulsed-field gel-separated partial digests of genomic DNA. *Nucl. Acids Res.* 25:3959-3961.
- Wang, M., Chen, X.N., Shouse, S., Manson, J., Wu, Q., Li, R., Wrestler, J., Noya, D., Sun, Z.G., Korenberg, J., and Lai, E. 1994. Construction and characterization of a human chromosome 2-specific BAC library. *Genomics* 24:527-534.
- Wang, G.L., Holsten, T.E., Song, W.Y., Wang, H.P., and Ronald, P.C. 1995. Construction of a rice bacterial artificial chromosome library and iden-

- tification of clones linked to the Xa-21 disease resistance locus. *Plant J.* 7:525-533.
- Woo, S.S., Jiang, J., Gill, B.S., Paterson, A.H., and Wing, R.A. 1994. Construction and characterization of a bacterial artificial chromosome library of Sorghum bicolor. *Nucl. Acids Res.* 22:4922-4931.
- Woon, P.-Y., Osoegawa, K., Kaisaki, P.J., Zhao, B., Catanese, J.J., Gauguier, D., Cox, R., Levy, E. R., Lathrop, G.M., Monaco, A.P., and de Jong, P.J. 1998. Construction and characterization of a 10-fold genome equivalent rat P1-derived artificial chromosomal library. *Genomics* 50:306-316.
- Wyman, A.R. and Wertman, K.F. 1987. Host strains that alleviate underrepresentation of specific sequences: Overview. *Methods Enzymol.* 152:173-180.
- Osoegawa et al., 1998. See above.
An improved BAC/PAC cloning approach.
- Pierce et al., 1992. See above.
Development of an elegant and efficient P1 cloning vector, a direct predecessor to the pCYPAC vectors.
- Shizuya et al., 1992. See above.
The development of efficient procedures for constructing large-insert plasmids by bacterial transformation.
- Sternberg, 1990. See above.
The first large-insert P1 plasmids constructed using in vitro packaging and viral infection to transform E. coli.

Key References

- Albertsen, H.M., Abderrahim, H., Cann, H.M., Dausset, J., Le Paslier, D., and Cohen, D. 1990. Construction and characterization of a yeast artificial chromosome library containing seven haploid equivalents. *Proc. Natl. Acad. Sci. U.S.A.* 87:4256-4260.
- The first description of use of PFGE to size select DNA for ligation.*
- Ioannou et al., 1994. See above.
The first paper describing PAC cloning.
- O'Connor, M., Peifer, M., and Bender, W. 1989. Construction of large DNA segments in *Escherichia coli*. *Science* 244:1307-1312.
- Describes first large BAC clone constructed by a combination of in vivo recombination and recombinant DNA technology.*
- Contributed by Kazutoyo Osoegawa
and Pieter J. de Jong
Children's Hospital Oakland Research
Institute
Oakland, California
- Eirik Frengen
The Biotechnology Centre of Oslo
University of Oslo
Oslo, Norway
- Panayiotis A. Ioannou
The Murdoch Institute for Research into
Birth Defects
Royal Children's Hospital
Melbourne, Australia

AMPLIFICATION OF TRANSFORMED OR PACKAGED LIBRARIES

After libraries are packaged (or transformed) they can be amplified for storage. Amplification of a library involves replicating each clone in the library. This is frequently a worthwhile procedure for genomic or cDNA libraries because the libraries can then be used many times to isolate clones corresponding to different sequences. Procedures for amplifying bacteriophage, cosmid, and plasmid libraries are presented in this section.

Amplification of a Bacteriophage Library

UNIT 5.10

This protocol may be used for genomic DNA or cDNA libraries. A freshly packaged and titered library is adsorbed to log phase plating bacteria. The mixture is then plated at high density and allowed to grow until the plaques are just subconfluent. The phage are eluted from the plate by overnight incubation with phage buffer and the library is titered and stored at both 4°C and -80°C.

BASIC
PROTOCOL

Materials

LB medium containing 0.2% maltose and 10 mM MgSO₄ (UNIT 1.1)

Suitable host (Table 5.10.1)

In vitro packaged phage library (UNITS 5.7 & 5.8)

Top agarose (UNIT 1.1), warmed to 47°C

150-mm H plates (UNIT 1.1), warmed to 37°C

Suspension medium (SM; UNIT 1.11)

Chloroform

Dimethyl sulfoxide (DMSO)

Additional reagents and equipment for titering bacteriophage (UNIT 1.11)

1. Inoculate 250 ml LB medium containing 0.2% maltose and 10 mM MgSO₄ with 2.5 ml of a fresh overnight culture of host bacteria. Shake vigorously 2 to 4 hr at 37°C until the OD₆₀₀ ≅ 0.5.
2. Prepare *E. coli* cells for plating as described in UNIT 1.11.

The integrity of λ phage particles requires the presence of Mg²⁺. Maltose induces the expression of the λ phage receptor. To maximize the complexity of the library, it is important that the number of nonviable cells among the host bacteria be minimized; these represent “dead ends” for phage.

Table 5.10.1 Suitable *Escherichia coli* Host Strains for Amplifying Lambda-Constructed Libraries

Vector	<i>E. coli</i> host	Relevant host genotype
λgt10	C600 <i>hflA</i>	<i>hflA</i>
λgt11	Y1088	<i>SupF</i> , <i>lacI^q</i> (no antibiotics needed)
EMBL 3 or 4	P2392, Q359, NM539	P2 lysogen
Charon 4A	LE392	<i>SupF</i>

Contributed by Lloyd B. Klickstein

Current Protocols in Molecular Biology (1996) 5.10.1-5.10.3

Copyright © 2000 by John Wiley & Sons, Inc.

3. Shake by hand and then briefly microcentrifuge the tube containing the library to separate out the chloroform in the packaged library. Add the packaged, titered phage (but no chloroform) to plating bacteria, using 0.25 ml of host cells for every 1×10^5 phage. Incubate 15 min at 37°C to allow phage to adsorb to bacteria.
4. Add 0.5 ml of the host/phage mixture to 8 ml of 47°C top agarose, mix by inverting a few times, and pour onto a fresh 150-mm H plate, warmed to 37°C. Use as many plates as necessary to plate the entire mixture. Spread the top agarose by gently rotating the plate on the benchtop, and allow the agarose to harden for 5 min. Invert and incubate 6 to 7 hr at 37°C (39° to 41°C for λ gt11 libraries amplified in Y1088).

A myriad of tiny plaques should barely be visible at ~4 to 5 hr. These should expand and be just less than confluent at 6 hr. If the plaques are growing slowly, continue the incubation for up to 8 hr.

5. Remove plates from incubator, cover each lawn with 10 ml SM, and place on a level surface at 4°C for 2 to 16 hr to elute the phage.
6. Using a pipet, combine the SM from all plates into a single glass or polypropylene tube. Centrifuge 5 min at 3000 rpm (2800 \times g) and transfer the supernatant to a new tube. Add 0.5 ml chloroform and mix by inverting several times. Titer the amplified library as described in UNIT 1.11. Expect a titer of 10^{10} to 10^{11} /ml.

The volume of SM recovered from each plate will be nearly 1 ml less than what was added. If the plates were not fresh, greater losses of volume will occur. The total volume of the library will, of course, depend on the number of plates used in the amplification and may approach 100 ml in the case of large libraries.

7. Store the library at 4°C in Teflon-capped glass tubes with 0.5 ml chloroform. The library is stable in this form for years. (Storage in plastic results in a 100- to 1000-fold drop in titer.) For added security, transfer 930 μ l-aliquots of amplified library (no chloroform) to screw-cap microcentrifuge tubes, add 70 μ l of DMSO (final concentration 7%), mix thoroughly by inverting several times, and place at -80°C.

COMMENTARY

Background Information

Amplification is not essential to the successful creation and screening of a λ phage library; however, an unamplified library must be used right away and the number of times the library may be screened is limited by the life of the nitrocellulose filters or the primary plates which dry out in a few weeks to a month. In contrast, amplification allows the library to be stored nearly indefinitely and to be screened as many times as necessary. The principle of amplification is to allow each in vitro packaged phage to produce thousands of identical clones by a limited infection of a host. A potential disadvantage is that the composition of the library may change as a consequence of differences in growth rate during the amplification step, some clones being overrepresented compared to the corresponding mRNA abundance in the total cellular mRNA and others being underrepresented. The preadsorption of the li-

brary to the bacteria, the high plating density, and the relatively short incubation period all help minimize changes in the composition.

An alternative approach used by some is amplification in liquid culture. This method is not recommended, since rapidly growing phage are overrepresented to a greater degree than when plates are used. Furthermore, there are more steps involved, since the phage must be purified from the relatively large volume by PEG precipitation and CsCl-gradient centrifugation.

Critical Parameters

Amplification should be carried out as soon as possible after the library is packaged. Phage particles stored in the in vitro packaging mixture are unstable because the phage particles adsorb to bacterial debris in the mixture and are inactivated. A decrease in titer at this point decreases the complexity of the library. Fresh,

log-phase host cells are important to ensure that each phage infects a viable cell and makes a phage burst and thus is successfully amplified. A phage that injects its DNA into a dead cell is lost from the library.

For libraries constructed in vectors such as λ gt10 or the EMBL series, which allow selection for recombinants, amplification should be done under selective conditions since this eliminates nonrecombinants. Subsequent platings of the amplified library may be done on either selective or nonselective hosts. The titer of a library stored as described in the basic protocol will usually drop only 2- to 3-fold over the years. If stored in plastic, then the titer may drop as much as 100- to 1000-fold.

Troubleshooting

The only difficulties that can occur here are a failure of the phage to infect the host or a failure of the host to grow. This may be avoided if the plates, media, and host strain are first tested with nonrecombinant phage prior to amplifying the library. Causes of problems include top agarose that is too hot, incorrect host strain, incorrect plates, incorrect incubator temperature, and so forth. Losing a library because a tube was dropped or contaminated may be avoided by splitting the packaged ligation into two tubes and amplifying in duplicate. A library stored over chloroform at 4°C may occasionally drop in titer due to fungal contamination

that occurred during the elution of the plates. Storing the library in several tubes and freezing aliquots will circumvent loss of library due to this problem.

Anticipated Results

The titer of an amplified phage library will be in the range of $1-10 \times 10^{10}$ pfu/ml. The volume obtained will depend on the number of plates used, which in turn depends on the initial complexity of the library. For example, a library initially containing 2×10^6 clones in 1.5 ml requires ten 150-mm plates, so ~100 ml of SM will be recovered in total. Thus, the library will have been amplified by a factor of roughly 10^5 .

Time Considerations

After in vitro packaging, the library is stored overnight at 4°C while a titer plate is incubating. The library should be amplified the next day by starting step 1 early in the day. The plaques should develop in 6 to 10 hr. Suspension medium is added and the plates elute overnight at 4°C. Alternatively, step 1 may be started in the late afternoon and the plates grown overnight for 7 hr.

Contributed by Lloyd B. Klickstein
Brigham and Women's Hospital
Boston, Massachusetts

Amplification of Cosmid and Plasmid Libraries

UNIT 5.11

Bacteria containing the recombinant clones are grown on agar plates, washed off the plates, and stored in glycerol. This procedure produces about a million-fold amplification of the library.

**BASIC
PROTOCOL**

Materials

LB plates containing appropriate antibiotic (UNIT 1.1)

LB medium (UNIT 1.1)

Sterile glycerol

Nitrocellulose membrane filters (Millipore HATF)

Sterile rubber policeman

1. Plate drug-resistant bacteria on nitrocellulose filters that have been placed on LB plates (containing antibiotic to which the colonies are resistant), as described in UNIT 6.2.

The filters are used so that the colonies remain small and do not grow together. Be sure to use commercially available triton-free filters to ensure bacterial growth.

2. Grow bacteria until colonies are just confluent.
3. Flood each plate with LB medium. Use ~2 ml/10-cm plate and 4 ml/15-cm plate.
4. Using a sterile rubber policeman, rub colonies off from nitrocellulose filter, making a bacterial suspension.
5. Pool the suspensions from the different plates into one 50-ml plastic tube.
6. Add sterile glycerol to give a final glycerol concentration of 15%.
7. Mix the solution thoroughly and dispense 500 μ l into 1-ml tubes.
8. Freeze the aliquots of the library at -70°C . They should remain viable without appreciable loss of bacteria for over 1 year.
9. To screen the library, simply thaw an aliquot, titer the bacteria concentration (UNIT 1.3), and plate the appropriate numbers of cosmid- or plasmid-containing bacteria on the screening filters.

COMMENTARY

Background Information

Amplification of phage libraries is an established procedure, but the effective amplification of plasmid and especially cosmid libraries has proven to be more difficult. The major concern in amplifying a plasmid or cosmid library is that the doubling times of bacteria are often altered, depending upon the type of insert included within the plasmid or cosmid vector. This disproportionate growth of recombinant bacteria can easily be visualized by plating out a portion of a library and noting the variable colony size after 16 hr at 37°C . However, there is ample incentive to amplify the library. The advantages of obtaining thousands of copies of a cosmid or plasmid library from a single li-

brary are obvious. When amplifying a cosmid or bacterial library the individual must weigh the concern of the possible underrepresentation of particular clones because of overgrowth of some colonies, against the time required to create a library de novo for each screening or for the dissemination of the library to other researchers.

Critical Parameters

The major concern with any amplification step is that each original recombinant be equally represented. Differences in the duplication rate of any recombinant bacteria will result in over- or underrepresentation in the amplified library. This concern is less for sim-

**Construction of
Recombinant
DNA Libraries**

5.11.1

Contributed by John H. Weis

Current Protocols in Molecular Biology (1996) 5.11.1-5.11.2

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 34

ple plasmid libraries than for cosmid libraries, which contain bacteria that grow at markedly different rates.

Anticipated Results

The purpose of amplification is to provide a reagent library that can be used many times. An amplified cosmid or plasmid library will contain one library's equivalence in 10 to 100 μ l. By producing 100 μ l of such a library frozen in 100 1-ml tubes, over 1000 platings can be expected.

Time Considerations

After the transformation or transfection step in the preparation of the plasmid or cosmid libraries, respectively, the plating of the out-growths will require 1 to 2 hr. The plates are then incubated overnight at 30° to 37°C, depending upon the desired colony size and the density of the platings. The filter wash, glycerol mix, and aliquoting into freezer tubes require an additional 2 hr.

Contributed by John H. Weis
Harvard Medical School
Boston, Massachusetts

CHAPTER 6

Screening of Recombinant DNA Libraries

INTRODUCTION

The usual approach to isolating a recombinant DNA clone encoding a particular gene or mRNA sequence is to screen a recombinant DNA library. As described in Chapter 5, a recombinant DNA library consists of a large number of recombinant DNA clones, each one of which contains a different segment of foreign DNA. Since only a few of the thousands of clones in the library encode the desired nucleic acid sequence, the investigator must devise a procedure for identifying the desired clones. The optimal procedure for isolating the desired clone involves a positive selection for a particular nucleic acid sequence. If the desired gene confers a phenotype that can be selected in bacteria, then the desired clone can be isolated under selective conditions (*UNIT 1.4*). However, most eukaryotic genes and even many bacterial sequences do not encode a gene with a selectable function. Clones encoding nonselectable sequences are identified by screening libraries: the desired clone is identified either because (1) it hybridizes to a nucleic acid probe, (2) it expresses a segment of protein that can be recognized by an antibody, or (3) it promotes amplification of a sequence defined by a particular set of primers.

Screening libraries involves the development of a rapid assay to determine whether a particular clone contains the desired nucleic acid sequence. This assay is used first to identify the recombinant DNA clone in the library and then to purify the clone (see Fig. 6.0.1). Normally, this screening procedure is performed on bacterial colonies containing plasmids or cosmids or on bacteriophage plaques. To test a large number of clones at one time, the library is spread out on agarose plates (*UNIT 6.1*), then the clones are transferred to filter membranes (*UNIT 6.2*). The clones can be simultaneously hybridized to a particular probe (*UNITS 6.3 & 6.4*) or bound to an antibody (*UNITS 6.7 & 6.11*). When the desired clone is

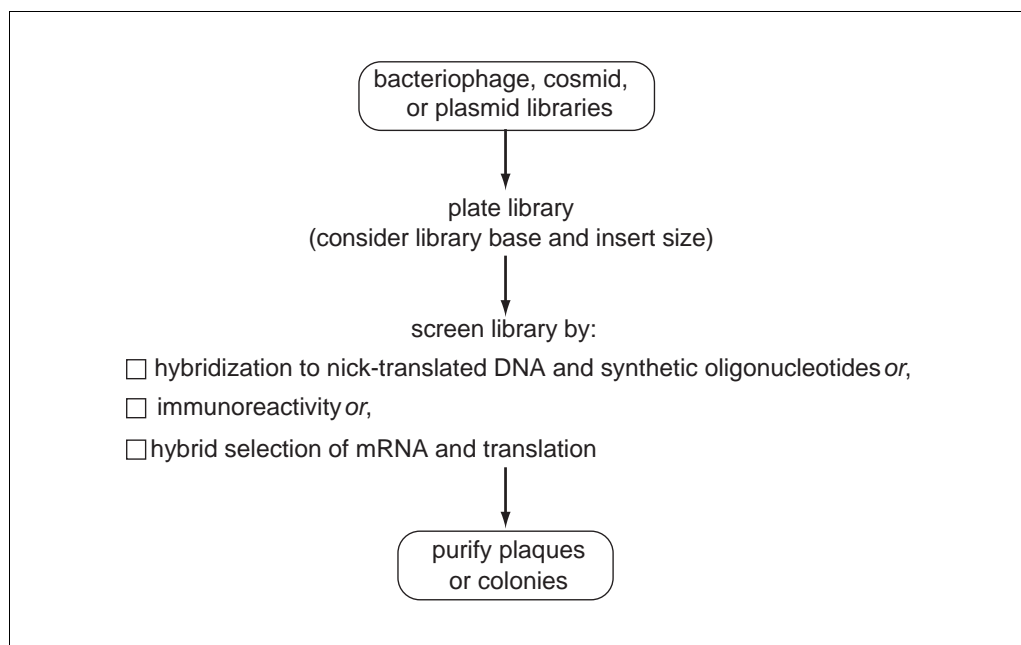


Figure 6.0.1 Flow chart for screening libraries.

Contributed by J.G. Seidman

Current Protocols in Molecular Biology (1994) 6.0.3-6.0.5

Copyright © 2005 by John Wiley & Sons, Inc.

first identified, it is usually found among many undesirable clones; an important feature of library screening is the isolation of the desired clones (*UNITS 6.5, 6.6 & 6.12*). Another method for identifying the desired clone involves hybrid selection (*UNIT 6.8*), a procedure by which the clone is used to select its mRNA. This mRNA is characterized by its translation into the desired protein. Libraries consisting of large genomic DNA fragments (~1 Mb) carried in yeast artificial chromosome (YAC) vectors have proven to be tremendously useful for genome analysis. In general, these libraries (which are usually produced by large “core” laboratories) are initially screened using a locus-specific PCR assay (*UNIT 6.9*); the clone resulting from the initial round of screening is subsequently analyzed by more conventional hybridization methods (*UNIT 6.10*).

To screen a DNA library, one must first devise the screening procedure. The next important choice is the selection of a recombinant DNA library. When choosing which library to screen the investigator should consider whether he or she wants to isolate clones encoding the gene or the mRNA sequence. cDNA clones encode the mRNA sequence and allow prediction of the amino acid sequence, whereas genomic clones may contain regulatory as well as coding (exon) and noncoding (intron) sequences. The differences between genomic and cDNA libraries are discussed in Chapter 5.

Another critical parameter to be determined before proceeding with a library screen is the number of clones in the library that must be screened in order to identify the desired clone. That is, what is the frequency of the desired clone in the library? This frequency is predicted differently for genomic and cDNA libraries, as described below.

Screening a genomic library. In general, genomic libraries can be made from DNA derived from any tissue, because only two copies of the gene are present per cell or per diploid genome. The predicted frequency of any particular sequence should be identical to the predicted frequency for any other sequence in the same genome. The formula for predicting the number of clones that must be screened to have a given probability of success is presented in *UNIT 5.1*. This number is a function of the complexity of the genome and the average size of the inserts in the library clones. For amplified libraries, the base (see *UNIT 5.1*) must exceed this number. Usually about 1 million bacteriophage clones or 500,000 cosmid clones must be screened to identify a genomic clone from a mammalian DNA library. Many of the clones that are screened from an amplified library will be screened more than once; the total number of clones that must be screened is 30 to 40% greater than the number calculated by the formula.

Screening a cDNA library. The optimal cDNA library is one made from a particular tissue or cell that expresses the desired mRNA sequence at high levels. In highly differentiated cells, a particular mRNA may comprise as many as 1 of 20 of the poly(A)⁺ mRNA molecules, while some mRNAs are either not present at all or comprise as low as 1 molecule in 100,000 poly(A)⁺ mRNA molecules. When choosing a cDNA library the investigator must make every effort to obtain a library from a cell where the mRNA is being expressed in large amounts. Of course, the number of clones that must be screened is determined by the abundance of the mRNA in the cell. The amount of protein that is found in the cell is frequently a good indicator of the abundance of the mRNA. Thus, proteins that comprise 1% of the total cell protein are made by mRNAs that usually comprise 1% of the total poly(A)⁺ mRNA, and the desired cDNA clones should comprise about 1% of the clones in the cDNA library.

Screening a YAC library. In the typical genomic libraries maintained in *E. coli* (described in Chapter 5), the size of the insert is limited to 20 to 25 kb for lambda vectors or to 40 to 45 kb for cosmid vectors. Yeast artificial chromosome (YAC) vectors, by contrast, are designed to carry much larger genomic DNA fragments and thereby facilitate genomic analysis, with inserts ranging from 0.3 to ~1 Mb in size. Conventional screening of YAC

libraries by hybridization is difficult, both because of the unfavorable signal-to-noise ratio and the sheer numbers of replica films required to represent an entire library.

For example, a standard YAC library representing 5 to 8 genome equivalents requires over 500 microtiter plates (and corresponding filters for screening by hybridization). Thus, most core laboratories screen YAC libraries using a locus-specific PCR assay whose primers define a particular sequence. The PCR screening is initially performed using pools (representing up to 4 microtiter plates or 384 YAC clones) or superpools (representing up to 20 microtiter plates or nearly 2000 clones), followed by subsequent rounds of screening to narrow down the possible candidates.

Specialized screening strategies. For particular applications, there exist specialized approaches to screening. For example, cloned cDNAs encoding cell surface or intracellular proteins can be identified by expression screening, involving rounds of transient expression of a library and subsequent screening by immunoselection (*UNIT 6.11*). The technique of recombination-based screening provides a rapid and efficient approach for screening a complex genomic library in bacteriophage lambda (*UNIT 6.12*). The library is screened for homology against a plasmid carrying a particular cloned target sequence. If homology exists, a recombination event occurs, resulting in integration of the plasmid into the phage, and the recombinant is isolated by genetic selection.

General considerations. When selecting the library it is critical that the base be larger than the number of clones to be screened. One problem with predicting the number of clones to screen is that most libraries are amplified and in the process of amplifying the library some clones are lost while others may grow more rapidly. Thus, if the desired clone is not found in a particular library, another independent library should be screened.

Having selected the library, the investigator is ready to begin screening for the desired clone. The technologies used to screen libraries are mostly extensions of the techniques that have been described earlier in the manual. Libraries are plated out, transferred to nitrocellulose filters, and hybridized to ^{32}P -labeled probes or bound to antibodies. The major problem associated with this technique is that “false” positives can be identified: the probe may hybridize to clones that do not encode the desired sequence. Approaches to minimize this problem are discussed in *UNIT 6.7*. A second source of undesired clones arises from the power of the screening procedures that are normally used to screen these libraries. The investigator will be screening as many as one million clones. If the library contains any contaminating recombinant DNA clones that have been previously grown in the laboratory, it will be identified in the screening procedure. Thus, extreme care must be exercised to prevent contamination of the library with previously isolated recombinant clones. Despite these problems the ability to screen large DNA libraries to isolate the desired clone provides a powerful tool for molecular biologists.

J.G. Seidman

PLATING LIBRARIES AND TRANSFER TO FILTER MEMBRANES

The basic principle of screening recombinant DNA libraries is that bacteriophage plaques, or bacterial colonies containing plasmids or cosmids, contain relatively large amounts of insert DNA that can be detected either directly by hybridization (see below) or indirectly by the protein that may be expressed from the cloned segment (UNIT 6.7). The first step in the nucleic acid hybridization screening procedure is to grow large numbers of colonies or plaques on agar plates. Replica copies of these colonies are transferred to nitrocellulose filters, where they can be screened. In this section the techniques for producing large numbers of colonies and plaques, and for transferring these to filter membranes, are discussed. Prerequisites to these procedures are that the library must already be chosen and the number of clones to be screened must be determined (see introduction to this chapter).

Plating and Transferring Bacteriophage Libraries

Bacteriophage are plated onto agar plates at high density so that as many as 1 million different plaques can be screened. The bacteriophage plaques are then transferred to nitrocellulose filters, denatured, and baked. The library and the number of clones to be screened are predetermined. Principles for choosing the plaque density and the number of plates to be used are outlined in the commentary.

Materials

- Host bacteria, selection strain if applicable (UNIT 1.10; Table 1.4.5; Table 5.10.1)
- Recombinant phage (UNIT 5.10)
- 0.7% top agarose (prewarmed; UNIT 1.1)
- 82-mm or 150-mm LB plates; or 245 × 245-mm Nunc bioassay LB plates (UNIT 1.1)
- 0.2 M NaOH/1.5 M NaCl
- 0.4 M Tris·Cl, pH 7.6/2× SSC
- 2× SSC (APPENDIX 2)
- Nitrocellulose membrane filters (or equivalent)
- 20-G needle
- 46 × 57-cm Whatman 3MM or equivalent filter paper
- 80°C vacuum oven or 42°C oven

Plating bacteriophage

1. Determine the titer of the library by serial dilution as described in UNITS 1.11 & 5.7.
For λ vectors that allow genetic selection against nonrecombinants, plating should be done on the appropriate bacterial strain (e.g., P2 lysogen for EMBL vectors). LB plates should be poured several days in advance to allow them to dry prior to plating. The large Nunc plates are particularly prone to condensation on the surface of the agar, but this can be alleviated by allowing them to sit on the benchtop with covers removed for a few minutes to several hours before use.
2. Mix recombinant phage and plating bacteria (prepared as described in UNIT 1.11) in a culture tube as outlined in Table 6.1.1 and incubate 20 min at 37°C.
3. Add 0.7% top agarose to culture tube and transfer mixture to LB plates. Disperse bacteria and agarose on plates by tilting the plates back and forth. Mix cells and agarose for the large Nunc plates by gently inverting several times in a capped 50-ml tube prior to plating.

Top agarose rather than top agar should be used as agar tends to lift off with the nitrocellulose filter.

Melt the top agarose and cool to 45° to 50°C before use. If top agarose is too hot it will kill the bacteria, while if it is too cold the library will solidify in the tube.

4. Incubate plates at 37°C until plaques cover the plate but are not confluent. Incubation time varies between 6 and 12 hr and depends on type of phage and bacteria used. Store at 4°C.

Do not incubate unattended overnight, but rather place at 4°C and allow to continue growth the next day. Allowing phage plaques to incubate for the correct amount of time is critical. The object is to optimize two parameters. First, the plaques must be large enough to contain sufficient DNA to give a good signal. Second, if the plaques are too large and become confluent they are difficult to purify in subsequent steps. Because most nucleic acid probes give a very strong signal, we tend to prefer having smaller plaques and weaker signals.

5. Incubate plates at 4°C for at least 1 hr before applying filters.

Transferring to nitrocellulose filters

6. Label nitrocellulose filters with a ballpoint pen and apply face down (ink side up) on cold LB plates bearing bacteriophage plaques. This is best accomplished by touching first one edge of the filter to the agarose and progressively laying down more of the filter as it wets. Bubbles should be avoided. If difficulties are encountered the filter should not be adjusted on the plate, but rather removed and replaced with a new filter.

Nitrocellulose filters should be handled only with forceps or gloved hands.

7. Leave filters on plates for 1 to 10 min to allow transfer of phage particles to the filter. During this transfer period the orientation of the filter to the plate is recorded by stabbing a 20-G needle through the filter into the agar at several asymmetric points around the edge of the plate. Up to five replicas can be made from each plate. Remove the filter slowly from the plate with blunt, flat forceps and place face up on paper towels or filter paper.

Some investigators dip the needle used to orient the filter in India ink to more clearly mark the filter and agar. Other investigators mark the back of the agar plate with a black marker.

Making two replicas from each filter, hybridizing both to the DNA probe, and comparing the autoradiographs of the replica filters eliminates many possible artifacts. This is particularly helpful when screening with an oligonucleotide probe.

8. Dry the filters on the benchtop for at least 10 min.

This drying process binds the plaques to the filter.

Table 6.1.1 Recommended Mixtures for Plating Bacteriophage Libraries

LB plate ingredient	Plate size		
	82 mm	150 mm	245 × 245 mm ^a
Bacteria ^b (ml)	0.2	0.5	2
Phage, pfu	5,000	20,000-30,000	150,000
Top agarose, ml	3	7	30

^aNunc Bioassay plates distributed by Vanguard International.

^bPlating bacteria are prepared as described in Chapter 1.

Denaturation and baking

9. Place 46 × 57-mm Whatman 3MM paper on the benchtop and saturate with 0.2 M NaOH/1.5 M NaCl. Place filters on the paper face up for 1 to 2 min.

The 3MM paper should be wet enough to allow immediate saturation of the filters, but not so wet that the solution pools on the surface.

10. Transfer filters (face up) to 3MM paper saturated with 0.4 M Tris·Cl, pH 7.6/2× SSC for 1 to 2 min and then to 3MM paper saturated with 2× SSC for 1 to 2 min.

Some investigators immerse the filters in all three solutions. This procedure can make the plaques detected by hybridization appear diffuse.

11. Dry filters in a vacuum oven 90 to 120 min at 80°C or overnight in a regular oven at 42°C. Store at room temperature in folded paper towels or other absorbent paper until needed for hybridization (described in *UNIT 6.3* or *6.4*).

COMMENTARY

Background Information

There are two parts to this protocol—plating the library and preparing filters. The number of bacteriophage per plate determines the number of plates that must be poured. This number is defined by the number of recombinants in the library (i.e., base of the library) and the frequency of the expected clone in the library. There is no advantage to screening more than 3 to 5 times the base of the library. The frequency of the clone in the library is determined as follows.

cDNA libraries: the expected frequency of the desired RNA among the total RNA of the cell, ranging from $1/100$ to $1/50,000$.

Genomic libraries: the size of the insert divided by the total genome size.

Subgenomic libraries: the size of insert per total genome size times the fold purification of the DNA fragment (usually 10- to 50-fold).

The usefulness of a recombinant phage library depends on the ability to screen a large number of phage and identify the clone that carries the DNA sequence of interest. This has been made possible by the technique of in situ plaque hybridization described by Benton and Davis (1977). The phage are allowed to multiply in host bacteria in a thin layer of agarose on regular bacterial plates. When nitrocellulose is applied to the agarose, phage particles and unpackaged DNA adsorb to the filter to produce a replica of the plate surface. If the agarose surface is not excessively wet, there will be little spreading of the phage on the filter. Subsequent treatment of the filter with sodium hydroxide destroys the phage particles and denatures the phage DNA which then binds to the nitrocellulose. Neutralization of the filters is required to maintain the integrity of the nitrocellulose. Hy-

bridization of these filters to a DNA or RNA probe will identify the location of the phage plaque of interest, which can then be recovered from the plate.

A common variation of this technique is the substitution of one of the nylon-based membranes for nitrocellulose (see *UNIT 2.9*). The advantage of nylon membranes is their durability, which allows multiple hybridizations to the same filter and allows one to sequentially clone several genes from the same library using a single set of filters. However, nylon filters do not offer an improvement in sensitivity and are often more expensive than nitrocellulose filter paper.

Literature Review

The molecular basis of λ phage replication and the adaptation of the λ genome for molecular cloning has been reviewed by Arber et al. (1983) and Williams and Blattner (1980). Principles governing the plating of λ phage have been outlined by Arber (1983); see also *UNIT 1.10*. Thorough understanding of these principles has led to a universal approach to plating phage libraries.

Critical Parameters

To prevent recombination between different phage, do not allow them to overgrow, and grow them in recombination-minus hosts where possible. Calculations of the amount of phage stock to be used per plate should be based on a recent titration, and plating cells should be fresh.

Filters must not become brittle during this procedure; brittle filters will be destroyed during the hybridization process. This can be avoided by limiting the time in the hydroxide solution to less than 5 min, making certain that

the 0.4 M Tris-Cl, pH 7.6/2× SSC brings the filters to neutral pH, and limiting the baking to 2 hr.

Troubleshooting

Plaques should be visible on the plate before filters are made. If there appears to be poor bacterial growth, it is possible that the top agarose was too warm and many bacteria were killed, or that the phage titer was higher than expected and most host cells were lysed. Lower than expected phage titer could be due to an inaccurate titration of the phage stock, poor host-cell preparation, or too little time for adsorption.

The preparation of the nitrocellulose filters will only be tested after hybridization is complete. Occasionally, hybridization to a plaque will produce a streak instead of a discrete circle on the autoradiograph, making location of the correct plaque difficult. Steps that will often correct this problem include: (1) drying plates with the cover removed for 1 to 2 hr before applying the filter, (2) drying the filters well before the hydroxide treatment, and (3) making certain that the face (phage side) of the filters is not directly in contact with the solutions.

Anticipated Results

This plating procedure characteristically produces plates with an even distribution of dense phage particles. It is sensitive enough to allow identification of a phage by hybridization even when the phage are plated at high density (>5000 plaques per 82-mm plate). A signal is easily visible after 18 to 24 hr, when filters are hybridized to a nick-translated DNA probe with activity of $>10^7$ counts/ μ g DNA.

Time Considerations

Usually plaques will become visible within 6 to 10 hr after plating. Bacteriophage should generally not be allowed to grow longer than necessary to visualize the plaques. Using the procedure outlined, even a large number of filters can be processed in a single day.

Literature Cited

- Arber, W. 1983. A beginner's guide to lambda biology. *In* Lambda II (R.W. Hendrix, J.W. Roberts, F.W. Stahl, and R.A. Weisberg, eds.) pp. 381-395. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Arber, W., Enquist, L., Hohn, B., Murray, N., and Murray, K. 1983. Experimental methods for use with lambda. *In* Lambda II (R.W., Hendrix, J.W. Roberts, F.W. Stahl, and R.A. Weisberg, eds.) pp. 433-466. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Williams, B.G. and Blattner, F.R. 1980. Bacteriophage lambda vectors for DNA cloning. *In* Genetic Engineering, Vol. 2 (J.K. Setlow and A. Mullander, eds.) p. 201. Plenum, NY.

Key References

- Benton, W.D. and Davis, R.W. 1977. Screening λ gt recombinant clones by hybridization to single plaques in situ. *Science* 196:180-182.
- Describes the method of plaque hybridization developed by the authors to allow isolation of phage possessing specific cloned DNA sequences.*

Contributed by Thomas Quertermous
Massachusetts General Hospital
Boston, Massachusetts

Plating and Transferring Cosmid and Plasmid Libraries

UNIT 6.2

**BASIC
PROTOCOL**

A bacterial suspension is suctioned through a porous membrane, leaving the bacteria bound to the membrane surface. The membrane is transferred, bacteria up, to an agar plate upon which the bacteria will receive enough nutrients to grow into colonies. These filters can then be used for replica platings and for hybridization with specific DNA probes.

Materials

- LB plates containing antibiotic (UNIT 1.1)
- LB medium (UNIT 1.1)
- LB plates containing 50 µg/ml chloramphenicol (UNIT 1.1)
- 0.5 M NaOH
- 1 M Tris·Cl, pH 7.5
- 0.5 M Tris·Cl, pH 7.5/1.25 M NaCl
- 10- or 15-cm Whatman 3MM or equivalent filter paper discs
- Sintered glass filter with vacuum
- Nitrocellulose membrane filters (10- or 15-cm, Millipore HATF)
- 20 × 20-cm Whatman 3MM or equivalent filter paper
- 20 × 20-cm glass plate
- 20-G needle
- 46 × 57-cm Whatman 3MM or equivalent filter paper
- 80°C vacuum oven

NOTE: All materials coming into contact with *E. coli* must be sterile.

Plating cosmids

1. Start with plasmid or cosmid library produced after transformation, transfection, or amplification (UNIT 5.7).
2. Determine titer of the library by serial dilutions using plates containing antibiotics (see UNIT 1.3).

Remaining library suspension can be held at 4°C overnight with only minimal loss of viable bacteria.

A 10-cm nitrocellulose filter can accommodate 10,000 to 20,000 colonies, while a 15-cm filter can hold up to 50,000.

3. Calculate the appropriate amount of the bacterial suspension for plating and dilute the suspension in LB medium such that there is the desired amount of bacteria in 5 ml (10-cm filter) or 10 ml (15-cm filter) of solution.
4. Meanwhile, prepare a layer of 10- or 15-cm Whatman 3MM paper discs on either the bottom part of a sintered glass Buchner funnel or on a porcelain filter funnel. Pour 10 to 20 ml LB medium over two or three layers of 3MM paper discs to make a level bed. The same pad of discs can be used for many filters.

Sterilize filter apparatus and filter paper before use. The 3MM and nitrocellulose filters can be sterilized by autoclaving them while wrapped in aluminum foil.

The purpose of this step is to spread the bacteria uniformly across the surface of a nitrocellulose filter. The filtering apparatus must be level, it must create a uniform suction to all the surface of the filter, and it should be easy to move the filters to and from the apparatus.

**Screening
Recombinant
DNA Libraries**

6.2.1

Contributed by Thomas Quertermous

Current Protocols in Molecular Biology (1987) 6.2.1-6.2.3

Copyright © 2004 by John Wiley & Sons, Inc.

Supplement 68

5. Label a nitrocellulose filter with a ballpoint pen on the side opposite that where the bacteria will be plated. Place the filter on the surface of the LB/antibiotic plate to wet it.

The antibiotic plate must be permissive for cosmid- or plasmid-bearing bacterial cells and usually is ampicillin or tetracycline.

Most ballpoint pen inks do not smudge during the hybridization reaction. If the one you choose runs, try another type.

6. Remove the wet filter from an antibiotic plate to the filtration apparatus.

The suction should be off.

Carefully pipet the 5 to 10 ml of bacterial suspension onto the surface of the nitrocellulose filter, leaving the outer 4 to 5 mm of the filter free of solution.

This outside bacteria-free ring leaves enough surface area to work with the filter without smearing or losing the colonies.

7. Slowly suction the solution down through the filter, taking care not to create any preferential suction pockets that would concentrate the bacteria. After suctioning all of the solution through the filter, transfer the filter back to the antibiotic plate on which it was wetted.

In laying the filter down on the agar surface, take care to avoid trapping any air bubbles between the surface of the plate and the filter.

8. Plate the entire library in this way and incubate the plates upside down (agar side up) at 37°C until the colonies are ~1 mm in diameter.

Do not overgrow the filters, as smaller colonies can be lost beneath larger, faster-growing recombinant bacteria.

Preparing replica filters

9. Label and wet another set of nitrocellulose filters, as described in step 5.

10. Remove the initial library filter from its plate and place on several sheets of 20 × 20 cm 3MM paper, bacteria side up. While wearing gloves, carefully position the wetted replica filter above the bacterial lawn. Lay the second filter upon the first, leaving the two filters offset by 2 to 3 mm.

This overlap will help in the separation of the two filters after the replica transfer.

Do not allow air bubbles to form between the two filters. These are excluded by touching the second filter to the first in the middle and then allowing the edges to fall.

11. Lay three sheets of 20 × 20-cm 3MM paper on the two filters, followed by a 20 × 20 cm glass plate. Using the palms of your hands, press with all your weight down on the glass plate, thus transferring the bacterial colonies from the library filter to the replica filter.
12. Remove the glass plate and the filter paper and, using a 20-G needle, punch holes 2 to 4 cm apart through both of the filters. These holes will allow the orientation of the film produced from the replica filter down on the library filter for the isolation of the correct clones.
13. Carefully peel the two filters apart, placing them both bacteria up, on their respective agar plates. Grow the replica colonies at 37°C overnight, leaving the library filters at 25°C overnight. After overnight growth, store the library filters on the agar plates at 4°C, while screening the replica filters.

Multiple replica filters can be made from the same library filter. Incubate library filters 2 to 4 hr at 37°C or overnight at 25°C to allow regrowth of the colonies.

5. Label a nitrocellulose filter with a ballpoint pen on the side opposite that where the bacteria will be plated. Place the filter on the surface of the LB/antibiotic plate to wet it.

The antibiotic plate must be permissive for cosmid- or plasmid-bearing bacterial cells and usually is ampicillin or tetracycline.

Most ballpoint pen inks do not smudge during the hybridization reaction. If the one you choose runs, try another type.

6. Remove the wet filter from an antibiotic plate to the filtration apparatus.

The suction should be off.

Carefully pipet the 5 to 10 ml of bacterial suspension onto the surface of the nitrocellulose filter, leaving the outer 4 to 5 mm of the filter free of solution.

This outside bacteria-free ring leaves enough surface area to work with the filter without smearing or losing the colonies.

7. Slowly suction the solution down through the filter, taking care not to create any preferential suction pockets that would concentrate the bacteria. After suctioning all of the solution through the filter, transfer the filter back to the antibiotic plate on which it was wetted.

In laying the filter down on the agar surface, take care to avoid trapping any air bubbles between the surface of the plate and the filter.

8. Plate the entire library in this way and incubate the plates upside down (agar side up) at 37°C until the colonies are ~1 mm in diameter.

Do not overgrow the filters, as smaller colonies can be lost beneath larger, faster-growing recombinant bacteria.

Preparing replica filters

9. Label and wet another set of nitrocellulose filters, as described in step 5.

10. Remove the initial library filter from its plate and place on several sheets of 20 × 20 cm 3MM paper, bacteria side up. While wearing gloves, carefully position the wetted replica filter above the bacterial lawn. Lay the second filter upon the first, leaving the two filters offset by 2 to 3 mm.

This overlap will help in the separation of the two filters after the replica transfer.

Do not allow air bubbles to form between the two filters. These are excluded by touching the second filter to the first in the middle and then allowing the edges to fall.

11. Lay three sheets of 20 × 20-cm 3MM paper on the two filters, followed by a 20 × 20 cm glass plate. Using the palms of your hands, press with all your weight down on the glass plate, thus transferring the bacterial colonies from the library filter to the replica filter.
12. Remove the glass plate and the filter paper and, using a 20-G needle, punch holes 2 to 4 cm apart through both of the filters. These holes will allow the orientation of the film produced from the replica filter down on the library filter for the isolation of the correct clones.
13. Carefully peel the two filters apart, placing them both bacteria up, on their respective agar plates. Grow the replica colonies at 37°C overnight, leaving the library filters at 25°C overnight. After overnight growth, store the library filters on the agar plates at 4°C, while screening the replica filters.

Multiple replica filters can be made from the same library filter. Incubate library filters 2 to 4 hr at 37°C or overnight at 25°C to allow regrowth of the colonies.

Then repeat steps 9 to 13. Normally, two copies of the cosmid are hybridized to each probe.

14. After the bacterial colonies have grown, the cosmids or plasmids on the replica filter are amplified by transferring them to an LB plate containing 50 µg/ml chloramphenicol and incubating at 37°C for 4 to 10 hr. This step will increase the signal produced by hybridization.

Preparing filters for hybridization

15. Remove the replica filters from the LB/chloramphenicol plates, place filters bacteria side up on a sheet of 46 × 57-cm 3MM paper soaked with 0.5 M NaOH, and leave them for 5 min.
16. Carefully transfer to a sheet of 46 × 57-cm 3MM paper soaked with 1 M Tris·Cl, pH 7.5. Allow neutralization to occur for 5 min.
17. Transfer to a third 46 × 57-cm filter soaked in 0.5 M Tris·Cl, pH 7.5/1.25 M NaCl. Neutralize 5 min.
18. Transfer filter to a dry sheet of 3MM paper to allow filter to dry.

After filters are completely dry, stack them on paper towels or other adsorbent paper. Each nitrocellulose filter should be separated by paper towels from other filters.

19. Transfer the stacked filters to a vacuum oven at 80°C for 90 min. Remove filters and hybridize with a nick-translated probe, as described in *UNITS 6.3* and *6.4*.

COMMENTARY

Background Information

There are two commonly used protocols for the screening of recombinant bacteria with hybridization probes. The first method involves the spreading of bacteria on the surface of agar using a sterile spreader (*UNIT 1.3*). A nitrocellulose membrane filter is then placed on top of the colonies and most of each colony is transferred to the filter. The filter is then treated as described in steps 15 to 19. This method works well when relatively small numbers of positive colonies are being selected (up to several thousand).

The second method employs a matrix of some type (here nitrocellulose filters are used) upon which bacteria can be plated and grown when the filter is placed on top of a nutrient agar surface. Once the plated bacteria have grown into visible colonies, the filters can be used for replica plating and in situ hybridization analysis.

Critical Parameters

In order to provide a uniform lawn of recombinant bacteria for screening, it is critical to ensure that the suction applied to the filters is uniform and not spotty. The best way to accomplish this is to suction the suspension through

the filter slowly and to avoid any preferential suction sites in the filter. Make sure that the apparatus is level and that adequate layers of LB-soaked chromatography paper are used. Air bubbles will prevent bacterial growth, so be certain that air is not trapped between the filter and the agar surface.

Time Considerations

Once the apparatus is set up, it takes ~5 min per filter to wet the filter, suction the bacteria, and transfer to an LB plate. The colonies take ~15 hr to grow at 37°C, after which they can be transferred to 4°C until ready for the replica platings. Replica plating also requires 5 min per filter, and resulting filters will be ready for denaturation and hybridization after 15 hr at 37°C.

Key Reference

Hanahan, D. and Meselson, M. 1983. Plasmid screening at high density. *Meth. Enzymol.* 100:333-342.

Contributed by John H. Weis
Harvard Medical School
Boston, Massachusetts

HYBRIDIZATION WITH RADIOACTIVE PROBES

After plaques or colonies have been transferred to nitrocellulose filters, the desired clone can be detected by its ability to hybridize to a DNA probe. This is a rapid, effective screening procedure that allows the identification of a single clone within a population of millions of other clones. The filters are hybridized with a ^{32}P -labeled nucleic acid probe, the excess and incorrectly matched probe is washed off the filter, and the filter is autoradiographed. Two features of the nucleic acid probe used for these experiments are critical to the successful screening of recombinant DNA libraries. First, the probe must hybridize only to the desired clones and not to any other clones. Thus, the nucleic acid sequence used for a probe must not contain any reiterated sequences or sequences that will hybridize to the vector. Second, the specific activity of the probe must be at least 10^7 cpm/ μg . Most of the procedures for labeling DNA or copy RNA molecules are described in Chapter 3, and a support protocol is presented here that allows the 5' end-labeling of a mixture of oligonucleotides.

The two basic protocols presented in this section describe steps required to hybridize labeled probes to recombinant DNA clones on filters. Two protocols are presented because conditions for hybridizing short oligonucleotide probes and longer nucleic acid probes to filters are different.

UNIT 6.3

BASIC PROTOCOL

Using DNA Fragments as Probes

HYBRIDIZATION IN FORMAMIDE

Bacteriophage plaques or bacterial colonies bound to a filter membrane are detected by hybridization with a radioactive probe. Hybridization proceeds on prewet filters placed in a sealable plastic bag. After hybridization the filters are removed from the sealed bag, excess probe is washed off, and the filters are autoradiographed to identify the clones that have hybridized with the probe.

Materials

Nitrocellulose membrane filters bearing plaques, colonies, or DNA (UNITS 6.1 & 6.2)

Hybridization solution I

Radiolabeled probe, 1 to 15 ng/ml (UNIT 3.5)

2 mg/ml sonicated herring sperm DNA

High-stringency wash buffer I

Low-stringency wash buffer I

Sealable bags

42°C incubator

Water bath adjusted to washing temperature (see commentary)

Glass baking dish

Additional reagents and equipment for autoradiography (APPENDIX 3)

Incubate filters with probe

1. Wet filters with hybridization solution I. Lay a filter membrane bearing plaques on top of 5 to 20 ml of hybridization solution I and allow solution to seep through filter. It is important to wet only one surface at a time to prevent trapping air in filter. Wet each filter in turn, producing a stack of wet filters.

When multiple filters are to be hybridized to the same probe, no more than twenty 8.2-cm discs or ten 20 × 20 cm square filters should be placed in one stack.

Using DNA Fragments as Probes

6.3.1

Contributed by William M. Strauss

Current Protocols in Molecular Biology (1993) 6.3.1-6.3.6

Copyright © 2000 by John Wiley & Sons, Inc.

Estimate the volume of hybridization solution used to wet the filters; this is a significant fraction of the volume of the hybridization reaction.

2. Transfer the stack of wetted filters to an appropriately sized sealable bag. Add enough hybridization solution to generously cover filters and seal.

Note the volume of hybridization solution used to cover the filters.

3. Prehybridize filters by placing the bag in a 42°C incubator for at least 1 hr.

Some investigators omit this step.

4. While filters are prehybridizing, pipet the radioactive probe into a screw-cap tube, add 2 mg (1 ml) sonicated herring sperm DNA, and boil 10 min. Place boiled probes directly into ice to cool.

The amount of probe used is important, and should be in the range of 1 to 15 ng/ml of hybridization reaction. The volume of the hybridization reaction can be assumed to be the amount of hybridization solution added to the filters.

5. Add 2 ml hybridization solution I to the boiled probe.
6. Remove bag containing filters from the 42°C incubator. Open bag, add probe mixture, exclude as many bubbles as possible, and reseal.

A good way to add the radioactive probe is to take it up in a syringe with an 18-G needle and then inject it into the bag. Reseal the bag after adding probe.

7. Mix probe in the bag so that filter is evenly covered. Replace bag in the 42°C incubator and let hybridize overnight.

Wash filters to remove nonhybridized probe

8. Warm 1 liter high-stringency wash buffer I to the “washing temperature” in a water bath.

The stability of washing temperature and salt concentrations are critical features of this experiment. See discussion in commentary.

9. Remove bag containing hybridizing filters from the 42°C incubator. Cut bag open and squeeze hybridization solution out of the bag.

CAUTION: Handle material carefully as it is extremely radioactive. This should be done on disposable paper bench covers.

10. Quickly immerse the filters in 500 ml low-stringency wash buffer I at room temperature in a glass baking dish. Separate all the filters, as they may stick together during hybridization.

The volume of the low-stringency wash buffer is not important as long as the filters are completely covered. The filters must not be allowed to dry as the radioactive probe will irreversibly bind the filters if the filters dry in contact with probe. (The type of container used to hold the filters is not important as long as it transfers heat well. Thus glass, metal, or enamel containers are better than plastic.)

The low-stringency wash only removes nonhybridized probe formamide and hybridization solution; it does not determine the stringency of the hybridization.

11. Rinse the filters three times with 500 ml low-stringency wash buffer. Let the filters sit 10 to 15 min at room temperature in low-stringency wash buffer with each rinse.
12. Pour off the low-stringency wash buffer and pour in 500 ml high-stringency wash buffer (prewarmed to washing temperature).

13. Replace the high-stringency wash buffer with another 500 ml of high-stringency wash buffer, then place the glass dish containing the filters in incubator at wash temperature. Make sure that the temperature in the glass dish reaches the desired washing temperature by placing a thermometer directly into the bath and measuring the temperature. Usually 15 to 20 min at the desired wash temperature is sufficient to remove most of the background radioactivity.

Of course, if the glass dish is placed in a water bath, be careful that the water from the water bath does not get into the filters.

Autoradiographing filters

14. Remove filters and mount them either wet or dry on a plastic backing. If the filter(s) is to be exposed wet, then isolate it from the film by covering it with plastic wrap.

Used X-ray film provides a good form of plastic backing for filters.

15. Mark the filters with radioactive ink to assist in alignment and autoradiograph.

An easy way to apply radioactive ink is to mark adhesive-backed paper labels with radioactive ink and then attach the stickers to the plastic wrap cover.

X-ray intensifying screens greatly decrease the amount of exposure time required.

ALTERNATE PROTOCOL

HYBRIDIZATION IN AQUEOUS SOLUTION

This method differs mainly in that formamide is not used in the hybridization solution. Follow the basic protocol except use the reagents and alternate parameters listed below.

Additional Materials

Hybridization solution II
Low-stringency wash buffer II
High-stringency wash buffer II
65°C incubator

1. Prehybridize as in basic protocol except that the filters are prehybridized at 65°C using hybridization solution II.

Hybridization solution II may have to be prewarmed to solubilize the SDS.

2. Prepare probe as in step 4 of basic protocol and dilute with 2 ml of hybridization solution II.
3. Hybridize overnight as in steps 6 and 7 of basic protocol except use a hybridization temperature of 65°C.
4. Remove bag containing hybridization from the 65°C incubator. Squeeze out the hybridization solution, taking care to avoid contamination with the excess radioactive hybridization solution.
5. Immediately rinse filters twice with low-stringency wash buffer II.

It is unnecessary to maintain a given temperature for this wash; just let the filters sit in wash buffer at room temperature until ready to proceed.

6. At 65°C, proceed to wash filters with high-stringency wash buffer II. Employ multiple quick washes (5 to 8) and immerse filter in a final wash for ~20 min. Check the radioactivity of the filters with a Geiger counter and be certain that they produce a signal only a fewfold above background levels.

REAGENTS AND SOLUTIONS

High-stringency wash buffer I

0.2× SSC (APPENDIX 2)
0.1% sodium dodecyl sulfate (SDS)

High-stringency wash buffer II

1 mM Na₂EDTA
40 mM NaHPO₄, pH 7.2
1% SDS

Hybridization solution I

Mix following ingredients for range of volumes indicated (in milliliters):

Formamide	24	48	72	120	240	480
20× SSC	12	24	36	60	120	240
2 M Tris-Cl, pH 7.6	0.5	1.0	1.5	2.5	5.0	10
100× Denhardt's solution	0.5	1.0	1.5	2.5	5.0	10
Deionized H ₂ O	2.5	5.0	7.5	12.5	25	50
50% dextran sulfate	10	20	30	50	100	200
10% SDS ^a	0.5	1	1.5	2.5	5	10
Total volume	50	100	150	250	500	1000

^aIn place of SDS, *N*-lauroylsarcosine (Sarkosyl) may be used.

Add the SDS last. The solution may be stored for prolonged periods at room temperature.

The dextran sulfate should be of high quality. Pharmacia produces acceptable-grade dextran sulfate. Recipes for SSC and Denhardt's solution are in APPENDIX 2.

Hybridization solution II

1% crystalline BSA (fraction V)
1 mM EDTA
0.5 M NaHPO₄, pH 7.2 (134 g Na₂HPO₄·7H₂O plus 4 ml 85% H₃PO₄/liter = 1 M NaHPO₄)
7% SDS

Low-stringency wash buffer I

2× SSC (APPENDIX 2)
0.1% SDS

Low-stringency wash buffer II

0.5% BSA (fraction V)
1 mM Na₂EDTA
40 mM NaHPO₄, pH 7.2
5% SDS

Sonicated herring sperm DNA, 2 mg/ml

Resuspend 1 g herring sperm DNA (Boehringer Mannheim #223636) in a convenient volume (about 50 ml of water) by sonicating briefly. The DNA is now ready to be sheared into short molecules by sonication. Place the tube containing the herring sperm DNA solution in an ice bath (the tube must be stable even if the ice begins to melt). The sonicator probe is placed in the DNA solution (without touching the bottom of the vessel). The sonicator is turned on to 50% power 20 min, or until there is a uniform and obvious decrease in viscosity. At no time should the tube containing the DNA become hot to the touch. After sonication, the DNA is diluted to a final concentration of 2 mg/ml, frozen in 50-ml aliquots, and thawed as needed.

COMMENTARY

Background Information

All hybridization methods depend upon the ability of denatured DNA to reanneal when complementary strands are present in an environment near but below their T_m (melting temperature). In a hybridization reaction involving double-stranded DNA on a filter and a single-stranded DNA probe there are three different annealing reactions occurring. First, there are the desired probe-DNA interactions which result in signal. Second, there are mismatch interactions that occur between related but non-homologous sequences; these mismatch hybrids are the ones that must be eliminated during the washing of the filters. Non-sequence-specific interactions also occur and these result in noise. The ability to extract information from a particular hybridization experiment is a function of the signal-to-noise ratio. High background or poor specific signal can both result in uninterpretable results.

Washing nitrocellulose filters is required to remove excess radioactive probe, as well as radioactive probe that has bound to the DNA on the filter as mismatch hybrids. Temperature and salt concentration dramatically affect the maintenance of specific hybrids. Detergents and other charged species can have a profound effect upon the nonspecific binding of species that contribute to background. In this protocol, hybridization is achieved in a solution containing 50% formamide. Excess probe is rinsed away under low-stringency conditions so that further hybridization will not occur. Once the hybridization solution is rinsed away, it is possible to proceed to a high-stringency wash without fear of further hybridization. When washing is complete, the filters should produce very little "noise" when monitored with a Geiger counter. Although single-copy sequence probe normally does not produce a signal that is detectable with a Geiger counter, a probe corresponding to more abundant sequences will produce a signal that can be "heard" with a Geiger counter.

Literature Review

Hybridization to filter membranes forms a basis of recombinant DNA technology and is described in detail earlier in the manual (UNIT 2.9). The protocols described here vary from those used for Southern blot filter hybridization in that the volume of the hybridization is usually larger and the washing conditions are different. Dextran sulfate is an important component of the hybridization

solution as it increases the rate of reassociation of the nucleic acids.

The protocols in this unit describe methods for hybridizing radioactive probes to membrane-bound plaques or colonies. These procedures for screening recombinant clones were first suggested by Grunstein and Hogness (1975) and by Benton and Davis (1977). The conditions of hybridization proposed in the basic protocol involving hybridization in formamide was originally described by Denhardt (1966) and Gillespie and Spiegelman (1965) while the alternate protocol using aqueous hybridization solution was introduced by Church and Gilbert (1984).

The method of washing filters under stringent conditions to remove background was first proposed by Southern (1975). Botchan et al. (1976) described the benefit of adding SDS to the wash solution. Jeffreys and Flavell (1977) first employed the wash conditions described in the protocols presented here.

Critical Parameters

Hybridization. Kinetically, the hybridization of DNA (or RNA) probes to filter-bound DNA is not significantly different from hybridization in solution. For single-stranded probes, the rate of hybridization follows first-order kinetics, since probe is available in excess. Under conditions of excess probe, the time for hybridization is inversely proportional to the probe concentration. For double-stranded probes the rate of hybridization displays a more complex relationship to the initial probe concentration. However, to a first approximation the initial probe concentration is inversely proportional to the rate of hybridization. To determine the actual time required for the successful hybridization of a given probe, either empirical data must be available or the following formula can be used to determine the length of time (in hours) required to achieve 50% hybridization (T_{50}):

$$\frac{1}{x} \times \frac{y}{5} \times \frac{z}{10} \times 2 = T_{50}$$

where x is the weight of probe in micrograms; y is the complexity of probe in kilobases; and z is the volume of hybridization solution in milliliters. The length of time T is given in hours. Maximum hybridization signal will be obtained if the reaction is allowed to proceed to $5 \times T_{50}$, although 1 to $2 \times T_{50}$ is often used.

It is also clear that nonspecific interactions

occur and that in any hybridization, sources of noise will be present. Therefore, from a practical standpoint one conventionally utilizes concentrations of nick-translated probe on the order of 1 to 15 ng/ml of hybridization, where the specific activity of the probe is from 5×10^7 cpm/ μ g to $>10^8$ cpm/ μ g. Too much probe in a hybridization is as bad as too little.

One important source of background hybridization to filters is due to the hybridization of the probe to vector sequences or to *E. coli* DNA. Be certain that there is no vector or *E. coli* DNA sequences in the probe. This can best be ensured by isolating the probe from one type of vector (e.g., plasmid) and screening a library made with a different type of vector (e.g., bacteriophage).

Washing temperature. Washing at low stringency is a straightforward proposition. Buffer is added at room temperature and washing proceeds at room temperature.

High-stringency wash is determined empirically. The relative homology between the probe and target sequence is a determining parameter. If the homology is 100%, a high temperature (65° to 75°C) can be used. As the homology drops, lower washing temperatures must be used. In general one starts at 37° to 40°C, raising the temperature by 3° to 5°C intervals until background is low enough not to be a major factor in the autoradiography.

The length of the probe is also important. Very short probes (<100 bp) must be washed at lower temperatures, even if the homology is 100%. Washing strategy is the same as for probes of differing homology.

Salt concentration. The lower the salt concentration, the higher the stringency. With this said, the protocols as outlined do not require adjustment of salt concentration for adjustment of stringency. Only the washing temperature is varied.

Probe. The nucleic acid probe must be of high specific activity and greater than 50 bp in length so that it can form stable hybrids.

Anticipated Results

After washing the filters the background

should be barely detectable with a Geiger counter.

With a high-specific-activity probe $>5 \times 10^7$ cpm/ μ g and an overnight hybridization reaction with a 1-kb unique sequence probe, hybridizing bacterial colonies or bacteriophage plaques can be visualized after a 1 to 18 hr exposure.

Time Considerations

Generally hybridizations are carried on overnight for 12 to 16 hr. This is sufficient for most probes and blots. However, with probes of increasing complexity longer hybridization times are required. This is preferable to increasing the probe concentration from experiment to experiment.

Autoradiography requires 1 to 18 hr.

Literature Cited

- Benton, W.D. and Davis, R.W. 1977. Screening λ gt recombinant clones by hybridization to single plaques in situ. *Science* 196:180.
- Botchan, M., Topp, W., and Sambrook, J. 1976. The arrangement of simian virus 40 sequences in the DNA of transformed cells. *Cell* 9:269-287.
- Church, G. and Gilbert, W. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 81:1991-1995.
- Denhardt, D. 1966. A membrane filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Commun.* 23:641-646.
- Gillespie, D. and Spiegelman, S. 1965. A quantitative assay for DNA-RNA hybrids with DNA immobilized on a membrane. *J. Mol. Biol.* 12:829-842.
- Grunstein, M. and Hogness, D. 1975. Colony Hybridization: A method for the isolating of cloned DNA's that contain a specific gene. *Proc. Natl. Acad. Sci. U.S.A.* 72:3961.
- Jeffreys, A.J. and Flavell, R.J. 1977. A physical map of the DNA region flanking the rabbit β globin gene. *Cell* 12:429-439.
- Southern, E.M. 1975. Detection of specific sequence among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.

Contributed by William M. Strauss
Harvard Medical School
Boston, Massachusetts

Using Synthetic Oligonucleotides as Probes

The protocols in this unit describe procedures for using mixtures of ^{32}P -labeled oligonucleotides to screen recombinant DNA clones bound to nitrocellulose filters. A partial amino acid sequence of a protein is used to predict the nucleotide sequence of the gene that would encode it. A mixture of oligonucleotides is chosen that includes all possible nucleotide sequences encoding that amino acid sequence. This mixture of oligonucleotides is then used to screen a recombinant DNA library for the corresponding clones. In some cases however, the exact nucleotide sequence of a desired clone is known and it is possible to use a unique oligonucleotide as a probe.

BASIC PROTOCOL

HYBRIDIZATION IN SODIUM CHLORIDE/SODIUM CITRATE (SSC)

This procedure outlines the steps necessary to screen nitrocellulose filters bearing DNA from bacteriophage or plasmids with mixtures of synthetic oligonucleotide probes. Hybridization and washing steps are carried out in solutions containing SSC. The washing temperature that produces the lowest background is determined empirically.

Materials

- Membrane filters bearing plasmid, bacteriophage, or cosmid libraries (UNITS 6.1 & 6.2)
- 3× SSC/0.1% SDS
- Prehybridization solution
- SSC hybridization solution
- 6× SSC/0.05% sodium pyrophosphate, prewarmed to wash temperature
- Filter forceps (e.g., American Scientific Products #2568-1)
- Sealable bags (or equivalent)
- Additional reagents and equipment for autoradiography (APPENDIX 3)

Prehybridize the filters

1. Prepare duplicate nitrocellulose filters of bacterial colonies or bacteriophage plaques. These should be processed and baked as described in UNITS 6.1 and 6.2.

Although some authors recommend wiping the wet filters prior to baking to remove bacterial debris, we do not advise this procedure because the hybridization signal may be reduced.

Filter forceps (i.e., without serrated tips) should be used to handle membrane filters to prevent marring the surface.

2. Wash the filters 3 to 5 times in 3× SSC/0.1% SDS at room temperature; about 50 82-mm filters can be washed in 500 ml. Then wash them once in the same solution at 65°C for at least 1.5 hr or overnight.

This step removes much of the bacterial debris from the filters.

3. Remove filters from 3× SSC/0.1% SDS and prehybridize them 1 hr at 37°C in prehybridization solution.

Herring sperm DNA in the prehybridization solution blocks nonspecific binding of probe to the filters and thus decreases the background level of radioactivity.

Hybridize oligonucleotides to the filters

4. Remove filters from the prehybridization solution and put them into sealable bags containing SSC hybridization solution. Place up to 20 filters and ≥20 ml SSC hybridization solution into each bag. Add 0.125 to 1.0 ng of each ^{32}P -labeled oligonucleotide per ml of hybridization solution to each bag. The mixed oligonucleotide probe is end-labeled with ^{32}P as described in the support protocol. For example, to 20 ml of hybridization solution that will contain a mixture of 128

17-base oligonucleotides, add 320 ng ($0.125 \text{ ng/ml} \times 128 \text{ oligonucleotides} \times 20 \text{ ml}$) of labeled probe. Hybridize filters 14 to 48 hr at the temperature indicated below:

14-base oligonucleotide	room temperature
17-base oligonucleotide	37°C
20-base oligonucleotide	42°C
23-base oligonucleotide	48°C

For bacterial colonies, adding much more than 0.125 ng of each oligonucleotide probe per ml of hybridization solution significantly increases the background on the autoradiogram. For bacteriophage plaques, there is less DNA per plaque than in a bacterial colony; as high backgrounds are not a problem with filters bearing bacteriophage plaques, more probe should be added to the hybridization mixture.

5. Remove filters from the hybridization bag and wash filters for 5 to 15 min, 3 to 5 times, in $6\times \text{SSC}/0.05\%$ pyrophosphate at room temperature.

It is important that the filters are well separated from each other and that the solution is occasionally or continuously gently agitated.

Wash the filters

6. Wash filters for 30 min in prewarmed $6\times \text{SSC}/0.05\%$ sodium pyrophosphate at the temperature indicated below:

14-base oligonucleotide	37°C
17-base oligonucleotide	48°C
20-base oligonucleotide	55°C
23-base oligonucleotide	60°C

Adjust the temperature of $6\times \text{SSC}/0.05\%$ pyrophosphate and filters. Measure the temperature of the filters and surrounding solution by putting the thermometer into the solution, not into the water bath. Make sure the filters are separated and are occasionally or continuously gently agitated.

7. Examine the filters with a Geiger counter; they should not exhibit above-background radioactivity. If the filters still show a significant degree of radioactivity above background, increase the temperature by 2° to 3°C for 15 to 30 min and reexamine the filters with the Geiger counter. Do not exceed the following temperatures:

14-base oligonucleotide	41°C
17-base oligonucleotide	53°C
20-base oligonucleotide	63°C
23-base oligonucleotide	70°C

The background level of bound radioactivity depends upon the amount of bacterial debris left on the filters, the amount of labeled oligonucleotides added to the hybridization mixture, and the guanosine-cytosine (G-C) content of the oligonucleotide mixture.

Perform autoradiography

8. When the filters exhibit a low level of radioactivity or the maximum temperatures referred to in step 7 have been reached, the filters should be removed from the wash solution and mounted wet on a solid support before exposure at -70°C to X-ray film, using an intensifying screen. Cover filters with plastic wrap. Do not allow the filters to dry out. Allow films to expose for 14 to 72 hr.

Autoradiograms made from filters with a high background may still yield interpretable results.

9. Develop the films; if a high background prevents proper interpretation of the films, rewash the filters at a higher temperature.
10. Number and mark the orientation of the films as described in *UNIT 6.3*. Spots that appear in precisely the same place on duplicate filters are “positives” (winners) and should be processed as described in *UNIT 6.5*.

*It is impossible to identify the characteristics of a true positive spot. Only colonies or plaques that produce evidence of hybridization on both filter copies should be processed as described below. Note that the intensity of the spot can vary dramatically between the duplicate filters. If a clear-cut spot appears on one filter and only a darkening of the background appears on the other, this should be considered positive and the plate should be processed as described in *UNIT 6.5*.*

Note that if two different oligonucleotide mixtures representing two different parts of the protein are available, either the positives obtained with one probe can then be hybridized with the other probe or four filter copies of the library can be made and hybridized to the two probes. Of course, depending on how far apart the sequences that hybridize to the two probes are, it is possible that neither will be present on a less than full-length cDNA clone.

BASIC PROTOCOL

HYBRIDIZATION IN TETRAMETHYLAMMONIUM CHLORIDE (TMAC)

This procedure is similar to the SSC protocol except that hybridization and washing are performed in solutions containing TMAC. In TMAC, the melting temperature of an oligonucleotide is a function of length and is independent of base composition; thus, spurious hybridization due to high G-C content of some of the oligonucleotides is reduced. Conditions are described for using 17-base oligonucleotides, but information is provided for determining the conditions when oligonucleotides of various lengths are employed.

Materials

Nitrocellulose or nylon membrane filters bearing plasmid, bacteriophage, or cosmid libraries (*UNITS 6.1* and *6.2*)

150-mm LB agarose plates (*UNIT 1.1*), prewarmed to 37°C

2× SSC/0.5% SDS/50 mM EDTA, pH 8.0, prewarmed to 50°C

TMAC hybridization solution, prewarmed to hybridization temperature

TMAC wash solution

2× SSC/0.1% SDS

15-cm glass crystallizing dishes

Filter forceps (e.g., American Scientific Products #2568-1)

Additional reagents and equipment for autoradiography (*APPENDIX 3*)

Process and prehybridize the filters

1. Process filters bearing bacterial colonies as described in *UNIT 6.2*. Produce filters bearing amplified bacteriophage plaques as follows:
 - a. Plate the bacteriophage from the library on LB agarose plates and transfer to nitrocellulose filters as described in *UNIT 6.1*, steps 1 to 7.

To obtain maximum sensitivity with oligonucleotide probes when the amplification procedure is used, plating density should be reduced to 8,000 to 10,000 plaques per 150-mm plate.

Either nitrocellulose or nylon (Colony/Plaque Screen Filters by New England Nuclear) filters can be used in this procedure. Nitrocellulose filters become fragile when hybridized in TMAC and must be handled very carefully. If this becomes a problem and nylon filters are substituted, the phage plaques must be amplified overnight. The rest of the protocol is unchanged.

- b. Amplify the bacteriophage by transferring the wet filter to a prewarmed (37°C) LB agarose plate so that the surface bearing the bacteriophage is faceup.

Refrigerate the master plates upon which the recombinant phage library were plated to prevent any further plaque expansion.
 - c. Incubate the plates at 37°C until the bacterial lawn re-forms on the surface of the nitrocellulose and plaques are evident. Plaque size will be somewhat larger than those on the original plate. This usually requires a 5- to 12-hr incubation period. Longer periods of growth will produce a dense bacterial lawn without significantly increasing plaque size or affecting hybridization signal.

Bacteriophage that produce small plaques (e.g., EMBL) are usually plated in the evening and allowed to grow overnight. The plaques are transferred to nitrocellulose filters the following morning and the phage are amplified on the filters by incubation for 5 to 7 hr during the day. Phage that produce large plaques (e.g., λgt10) are plated early in the morning, allowed to grow 5 to 7 hr, transferred to nitrocellulose filters (steps 6 and 7 of UNIT 6.1), transferred to fresh plates, and then incubated for amplification overnight.
 - d. Denature and bind the bacteriophage DNA to nitrocellulose filters as described in steps 8 to 11 of UNIT 6.1.
2. Wash filters bearing bacterial colonies as described in step 1 of the SSC protocol. Wet bacteriophage-bearing filters in a prewarmed (50°C) solution of 2× SSC/0.5% SDS/50 mM EDTA (pH 8.0). Float the filters on top of the solution (with the surface containing the dried bacteria and plaques faceup) to allow the nitrocellulose to wet completely. Submerge the filters and, with a gloved hand, gently rub the surface of the filters to remove the dried bacterial debris. Transfer the filters to a container of fresh solution of 2× SSC/0.5% SDS/50 mM EDTA to remove bacterial debris.

Alternatively, the filters can be incubated in this solution at 65°C for one to several hours and then scrubbed. Inadequate scrubbing of the filters results in an increase of nonspecific background hybridization, obscuring positive hybridization signals in the subsequent screening procedure.
 3. Transfer the filters to a 15-cm glass crystallizing dish containing 5 to 10 ml TMAC hybridization solution (per filter), which has been prewarmed to the appropriate hybridization temperature (48°C for 17-mer oligonucleotides; see Fig. 6.4.1 and commentary for other oligonucleotides) and seal the dishes with plastic wrap and rubberbands. Prehybridize 1 to 2 hr at the hybridization temperature, which is 5° to 10°C below the melting temperature.

Prehybridization and hybridization can be performed in glass crystallizing dishes that are slightly larger in diameter than the nitrocellulose filters. Gentle agitation on an orbital platform shaker will allow the solution to pass freely between the stacked filters and prevent the filters from sticking together. Place no more than 25 to 30 filters in each dish.

Alternatively, prehybridization and hybridization can be performed in a sealable bag (see SSC protocol) with <10 filters per bag.

Hybridize oligonucleotides to the filters

NOTE: The following hybridization and wash temperatures have been optimized for the use of oligonucleotide probes of 17 nucleotides in length. If different length oligonucleotide probes are used, these conditions should be adjusted based on the information presented in the commentary.

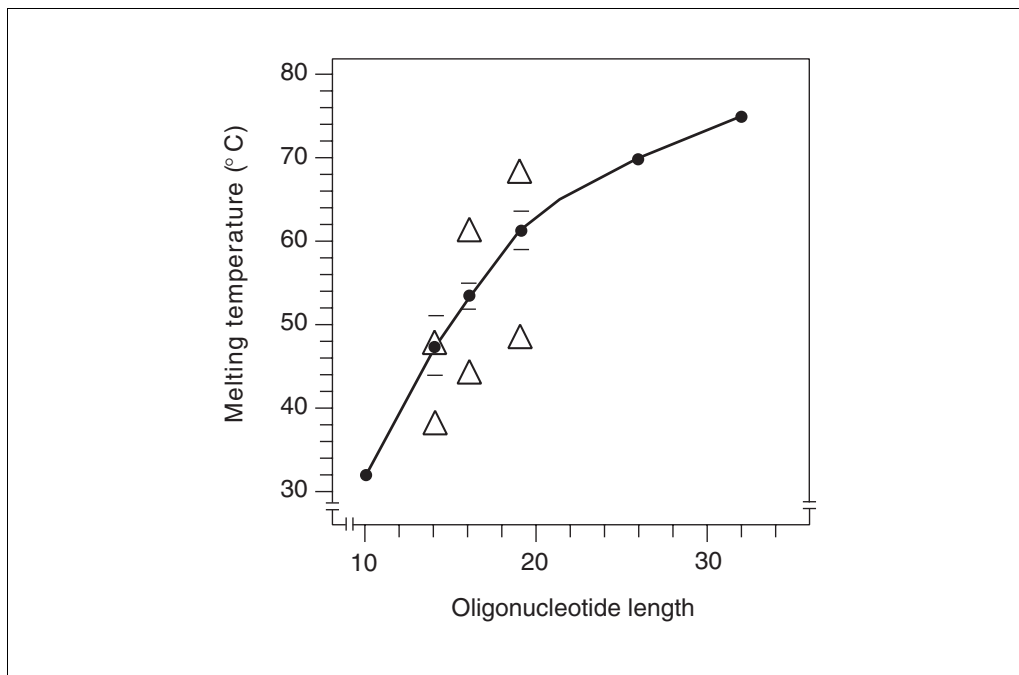


Figure 6.4.1 Melting temperatures of oligonucleotides of different length in TMAC and SSC hybridization solutions. Dots represent the average melting temperature of several different oligonucleotides of length 14, 16, or 19 bases in TMAC; bars represent the high and low melting temperatures for each length. Triangles represent the high and low melting temperatures for the same oligonucleotides in SSC. The melting temperature of only one oligonucleotide of length 10, 26, or 32 bases was determined. Hybridization temperature should be 5° to 10°C below the melting temperature, and washing temperature also should be 5° to 10°C below the melting temperature (Jacobs et al., 1988).

4. Transfer filters to a hybridization vessel (or bag) containing fresh, prewarmed (48°C) TMAC hybridization solution to remove residual SSC and bacterial debris from step 2 and to restore TMAC concentration to 3 M. Use 5 to 10 ml TMAC hybridization solution per filter.
5. Add 1 to 2×10^6 cpm of ^{32}P -labeled oligonucleotide probe per ml of hybridization solution directly to the solution of step 4 and incubate 40 to 60 hr at 48°C with gentle agitation on an orbital shaker.

^{32}P -label the oligonucleotide(s) with T4 polynucleotide kinase as described in the support protocol.

Wash the filters

6. Discard the hybridization solution containing radioactively labeled probe and rinse the filters with TMAC wash solution at room temperature. Use 5 to 10 ml TMAC wash solution per filter.
7. Transfer the filters individually to fresh TMAC wash solution (200 to 250 ml) and wash 15 min at room temperature with gentle agitation.

Individual transfer of filters reduces background.

8. Replace the room-temperature TMAC wash solution with a similar volume of prewarmed TMAC wash solution and incubate the filters 1 hr at the appropriate wash temperature (50°C for 17-mers; see Fig. 6.4.1 and commentary).
9. Replace the TMAC wash solution with a similar volume of $2\times$ SSC/0.1% SDS. Wash the filters 10 min at room temperature.

10. Repeat step 9 twice.

These washes remove residual traces of TMAC from the nitrocellulose filters and prevent crystallization of the TMAC salts upon drying.

Perform autoradiography

11. Autoradiograph the filters as described in steps 8 to 10 of the SSC protocol.

LABELING THE 5' ENDS OF MIXED OLIGONUCLEOTIDES

A mixture of synthetic oligonucleotides is ³²P-labeled using T4 polynucleotide kinase. This protocol is similar to that used to label a single oligonucleotide (*UNIT 3.10*).

Materials

2.5 to 250 pmol mixed oligonucleotides
[γ -³²P]ATP (>7000 Ci/mmol)
25 to 50 U T4 polynucleotide kinase (*UNIT 3.10*) and 10 \times kinase buffer (*UNIT 3.4*)
Ice-cold 10% trichloroacetic acid (TCA)

1. Set up reaction mixture on ice in microcentrifuge tube as follows:

2.5 to 250 pmol mixed oligonucleotides
7.5 μ l 10 \times T4 polynucleotide kinase buffer
66 pmol [γ -³²P]ATP (200 μ Ci)
25 to 50 U T4 polynucleotide kinase
H₂O to 75 μ l

Incubate 30 min at 37°C.

The reaction mixture should have either equimolar amounts of label and oligonucleotide ends, or the label should be in molar excess.

*1 mol deoxyribonucleotide \cong 330 g
1 OD₂₆₀ \cong 40 μ g/ml oligonucleotide
1 μ g 14-base oligonucleotide \cong 0.24 nmol
1 μ g 17-base oligonucleotide \cong 0.18 nmol
1 μ g 20-base oligonucleotide \cong 0.15 nmol*

2. At the end of the reaction, check for incorporation of label by precipitating 1 μ l of a diluted aliquot with ice-cold 10% TCA (acid precipitation, *UNIT 3.4*) and counting the incorporated radioactivity.

Using equimolar amounts of oligonucleotide and label, ~30% to 90% of the counts are incorporated.

*The labeled oligonucleotide can be further purified by a combination of phenol extraction and/or ethanol precipitation (*UNIT 2.1*). To remove unincorporated label, oligonucleotides of 17 bases or longer can be quantitatively precipitated from a solution of 2.5 M ammonium acetate containing 25 μ g carrier DNA plus 9 vol of 100% ethanol. The resulting pellets are washed with 70% ethanol, followed by 95% ethanol, air dried, and resuspended in 100 μ l TE buffer.*

3. Store mixture in appropriate container at -20°C.

REAGENTS AND SOLUTIONS

Prehybridization solution

6 \times SSC (*APPENDIX 2*)
5 \times Denhardtts solution (*APPENDIX 2*)
0.05% sodium pyrophosphate
100 μ g/ml boiled herring sperm DNA *continued*

SUPPORT PROTOCOL

Screening Recombinant DNA Libraries

6.4.6

0.5% sodium dodecyl sulfate (SDS)

SSC hybridization solution

6× SSC (APPENDIX 2)
1× Denhardt's solution (APPENDIX 2)
100 µg/ml yeast tRNA
0.05% sodium pyrophosphate

TMAC hybridization solution

3 M tetramethylammonium chloride (see recipe below for stock solution)
0.1 M NaPO₄, pH 6.8
1 mM EDTA, pH 8.0
5× Denhardt's solution (APPENDIX 2)
0.6% SDS
100 µg/ml denatured salmon sperm DNA

TMAC wash solution

3 M tetramethylammonium chloride (see recipe below for stock solution)
50 mM Tris-Cl, pH 8.0
0.2% SDS

Tetramethylammonium chloride (TMAC), 6 M stock solution

Dissolve 657.6 g TMAC (mol wt = 109.6) in H₂O and bring to 1 liter. Filter the solution through Whatman No. 1 filter paper and determine the precise concentration of the solution by measuring the refractive index (*n*) of a 3-fold diluted solution. The molarity (M) of the diluted solution = $55.6(n - 1.331)$ and the molarity of the stock solution = $3 \times M$. TMAC can be stored at room temperature in brown bottles.

CAUTION: TMAC can irritate eyes, skin, and mucous membranes. It should be used with adequate ventilation in a fume hood. Used TMAC solutions should be collected and discarded as hazardous and/or radioactive waste. Small amounts (<10 ml) can be flushed down the drain with a large quantity of tap water.

COMMENTARY

Background Information and Literature Review

Based on a study of the effect of single-bp mismatches on the hybridization behavior of oligonucleotides to ΦX174 DNA, Wallace et al. (1979) proposed that synthetic oligonucleotide mixtures representing all the possible coding sequences for a particular peptide sequence might be used as specific probes to identify cloned DNA. They demonstrated that duplexes with a single-bp mismatch—formed when 11-, 14-, or 17-base oligonucleotides were hybridized to ΦX174 DNA—were significantly less stable (dissociated at lower temperatures) than their perfectly matched counterparts. This difference in thermal stability made it possible, by the appropriate choice of hybridization conditions, to virtually eliminate the formation of mismatched duplexes without affecting the formation of perfectly matched ones. Mixed-sequence oligonucleotide probes were first used successfully for the isolation of a cloned cDNA encoding human β₂-microglobulin (Suggs et al., 1981). Many different

cDNA clones have been isolated using this approach.

Recombinant DNA libraries can be screened with probes consisting of single oligonucleotides or mixtures of oligonucleotides. Bacteriophage or plasmid clones that encode sequences perfectly complementary to the probe must be distinguished from clones that do not. Most often recombinant DNA libraries are screened with probes consisting of multiple oligonucleotides, chosen to cover all the coding possibilities of a particular amino acid sequence. Since any member of the oligonucleotide pool could match the target sequence, hybridization and wash temperatures are selected that allow the oligonucleotide with the lowest melting temperature (the lowest G-C content) to hybridize efficiently. In solutions containing SSC, oligonucleotides of high G-C content can potentially form stable duplexes with sequences to which they are not perfectly complementary.

Mixed probes can produce a high number of false positives using the SSC hybridization

conditions. This problem can be minimized by replacing SSC with TMAC in the hybridization and wash solutions. The melting temperature of long, native DNA in 3 M TMAC is independent of base composition (Melchior and von Hippel, 1973) and rate of renaturation is approximately the rate of renaturation in SSC (Wetmur, 1976). In solutions containing 3 M TMAC, the thermal stability of oligonucleotides 16 bases or longer is sequence-independent (Wood et al., 1985; Jacobs et al., 1988), and non-Watson-Crick base pairs decrease the thermal stability of oligonucleotide duplexes 1° to 1.5°C per percent mismatch (Jacobs et al., 1988). Thus, when the hybridization and wash are performed in 3 M TMAC the appropriate temperature is defined by the length of the oligonucleotide probes. All of the oligonucleotides in the pool will hybridize with equal efficiency to their complementary sequences and with reduced efficiency to sequences to which they are not perfectly complementary. The gene encoding erythropoietin was isolated using TMAC (Jacobs et al., 1985).

The support protocol describes a procedure for ³²P-labeling synthetic oligonucleotides and is adapted from Richardson (1971).

Critical Parameters

In choosing the stretch of amino acids to be used as the blueprint for the mixture of oligonucleotides, one finds a stretch of amino acids that is encoded by a minimum number of possible codons. Unfortunately, most amino acids are encoded by two or four codons. Only methionine and tryptophan—among the rarest of amino acids—are encoded by a single codon. In addition, three of the commonest amino acids have a total of six different codons each (see codon chart, APPENDIX 1). Computer programs are available to determine the optimal structure of oligonucleotide probes from amino acid sequences (Yang et al., 1984).

As the number of different sequences present in an oligonucleotide mixture increases, several possible problems arise. First, unfavorable signal-to-noise ratios may be a problem. However, mixtures of 17-bp oligonucleotides containing 300 to 600 different sequences have been used. Second, as the number of different sequences increases, the probability of completely matching an unrelated and undesired mRNA sequence increases. Third, the distinct possibility remains that the sequence of interest may not be present, and/or incorrect oligonucleotides may be present in the mixture.

Another strategy is to use a single, long,

synthetic oligonucleotide probe selected from codon usage frequencies (reviewed by Lathe, 1985). The uncertainty at each codon is ignored and increased probe length is used to confer probe specificity. The choice of the codon for each amino acid is based on codon utilization data, intercodon dinucleotide frequencies, and other rules. Knowledge of dinucleotide frequencies is important because of the decreased frequency of the dinucleotide pair, CpG. A list of codons recommended by Lathe (1985) is given in Table 6.4.1.

Determining the temperatures for hybridizing and washing is important to reduce the number of false positive plaques or colonies. Although there are examples of isolating clones that have a one-base mismatch with the oligonucleotide probe (<23 nucleotides), this usually is the result of a wrong sequence being

Table 6.4.1 Optimum Codon Choice When Deducing a Probe Sequence from Human Amino Acid Sequence Data

Amino acid	Optimum codon ^a when subsequent codon begins with	
	A or C or T	G
Methionine	ATG	nc ^b
Tryptophan	TGG	nc
Tyrosine	TAC	TAT
Cysteine	TGC	TGT
Glutamine	CAG	nc
Phenylalanine	TTC	TTT
Aspartic acid	GAC	GAT
Asparagine	AAC	AAT
Histidine	CAC ^c	CAT
Glutamic acid	GAG	nc
Lysine	AAG	nc
Alanine	GCC	GCT
Isoleucine	ATC	ATT
Threonine	ACC	ACA ^f
Valine	GTG ^d	nc
Proline	CCC ^e	CCT
Glycine	GGC	nc
Leucine	CTG	nc
Arginine	CGG	nc
Serine	TCC	TCT

^aThe optimum codon is the most frequent codon in all cases except Arg and Ser, where the indicated triplets generate a higher overall homology to all possible codons. Reprinted with permission from *Journal of Molecular Biology*.

^bNo change.

^cCAT when followed by C.

^dGTC when followed by T.

^eCCA when followed by T.

^fThese cases do not follow the “replace C by T” rule applied when the subsequent codon is headed by G.

present in the oligonucleotide mixture. Furthermore, one occasionally synthesizes a mixture of oligonucleotides that is missing the single correct sequence. This possibility should be remembered if no positive results are obtained after screening a library.

Using two copies of each filter to compare the signal obtained with both is an important method for eliminating false positive clones. Because of the variation in the appearance of positive signals and the frequently high backgrounds, it is virtually impossible to tell if a signal from one filter represents true oligonucleotide hybridization. Libraries can be screened in duplicate with one oligonucleotide pool or can be screened with two different oligonucleotides derived from amino acid sequences for the same protein.

The SSC hybridization procedure is similar to that presented by Woods et al. (1982). Other procedures use Tris buffers instead of sodium citrate buffers in the hybridization solution and employ a shorter period of time for the final stringent wash (Connor et al., 1983).

In the SSC protocol, a low temperature for hybridizing the oligonucleotide to the filters is critical to allow maximal binding of the perfectly complementary oligonucleotides. A high temperature in the stringent wash is essential to remove mismatch hybrids. The method for optimizing the signal-to-noise ratio is to wash the hybridized filters at an empirically determined temperature. Initial hybridization and stringent washing conditions can be determined by formulas (see Lathe, 1985) and improved upon by trial-and-error adjustments. Generally, probes ≥ 50 bp long that have $\geq 80\%$ homology will be specific and will hybridize to the sequence of interest. The probability of constructing an appropriate oligonucleotide probe is high but there is a possibility that a nonhybridizing probe will be made (Parker et al., 1986).

In the SSC protocol, background level of radioactivity can vary between none (unusual) to dark gray (usually not a problem). If the background is high, positive spots can often be identified by comparing duplicate filters. Filters can be re washed at a higher temperature and reexposed to film if the background is too dark. If this does not remove the high background, too little hybridization solution was used or the oligonucleotide probe is no good and should be remade. Another possibility is that the probe is hybridizing to *E. coli* DNA.

The TMAC protocol describes hybridization times and temperatures appropriate for screening libraries with 17-mers. In practice we

find that using both hybridization and wash temperatures in the range of 5° to 10°C below the melting temperature (T_m) for an oligonucleotide of particular length gives optimal results. With increasing degeneracy of the oligonucleotide, it is advisable to use the lower range of hybridization and wash temperatures (10°C below T_m) because of the reduced signal. For other applications or when using oligonucleotides of different lengths and complexity, the appropriate conditions can be determined using the information in Figure 6.4.1. A plot of T_m versus oligonucleotide length in TMAC and SSC is shown in Figure 6.4.1. Clearly, there is a broader melting range in SSC than in TMAC. By referring to this figure, the T_m for an oligonucleotide from 16 to 32 bases can be determined. Alternatively, the following equation can be used:

$$T_m = -682(L^{-1}) + 97$$

where L is the number of bases in the oligonucleotide and T_m is $^\circ\text{C}$.

With TMAC, background level due to radioactivity occurs sometimes with particular oligonucleotides but has not obscured duplicate positives. If oligos shorter than 17 bases are used, the general background will be higher, presumably because of the lower temperature used. If the background is too high, the filters can be washed at 5°C below the T_m , but only in preequilibrated buffer and for no more than 10 to 15 min (steps 7 to 10 of TMAC protocol).

When bacteriophage plaques were amplified and probes hybridized in TMAC, 17-mers that were 600-fold degenerate were used to isolate clones from a λ gt10 cDNA library; hybridization was performed for 3 days and exposure times of 24 hr were sufficient. Genomic libraries have been screened with 17-mers that were 512-fold degenerate, requiring hybridization times of 2 to 3 days and exposure times of 1 to 2 days. These times were not rigorously determined. For longer oligonucleotides or with less complex pools, shorter hybridization times may be used. Exposure times are usually determined by developing a few films and making a judgement based on the background and number of positives. When the bacteriophage plaques are amplified, stripping and screening the filters a second time is not recommended.

Anticipated Results

Any recombinant DNA library screened with oligonucleotides will yield "positives" because statistically there will be sequences that are

partially complementary to the probe and hybridization is not an all-or-nothing event. The number of positives will depend upon the completeness of the recombinant DNA library and, in the case of a cDNA library, the level of expression of the gene of interest. In fact, if no "positives" are obtained, there is a high probability that there was procedural error.

Using the TMAC hybridization protocol facilitates the process of identifying the clones of interest because true positives are darker than false positives and thus readily distinguished. If the correct clone is not identified after several attempts, consider remaking the oligonucleotides or making different oligonucleotides. There remain, of course, the unpleasant possibilities that there is an intron spanning the probe or the target sequence is underrepresented in the library.

Time Considerations

For the SSC protocol, 14 to 60 hr is recommended for hybridization and at least 3 hr for washing filters. The TMAC protocol requires 40 to 60 hr hybridization and 1 to 2 hr for washing filters.

Literature Cited

- Connor, B.J., Reyes, A.A., Morin, C., Itakura, K., Teplitz, R.L., and Wallace, R.B. 1983. Detection of sickle cell β^s -globin allele by hybridization with synthetic oligonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 80:278-282.
- Jacobs, K., Shoemaker, C., Rudersdorf, R., Neill, S.D., Kaufman, R.J., Mufson, A., Seehra, J., Jones, S.S., Hewick, R., Fritsch, E.F., Kawakita, M., Shimizu, T., and Miyake, T. 1985. Isolation and characterization of genomic and cDNA clones of human erythropoietin. *Nature* 313:806-810.
- Jacobs, K.A., Rudersdorf, R., Neill, S.D., Dougherty, J.P., Brown, E.L., and Fritsch, E.F. 1988. The thermal stability of oligonucleotide duplexes is sequence independent in tetraalkylammonium salt solutions: Application to identifying recombinant DNA clones. *Nucl. Acids Res.* In press.
- Lathe, R. 1985. Synthetic oligonucleotide probes deduced from amino acid sequence data. Theoretical and practical considerations. *J. Mol. Biol.* 183:1-12.

- Melchior, W.B. and von Hippel, P.H. 1973. Alteration of the relative stability of dA-dT and dG-dC base pairs in DNA. *Proc. Natl. Acad. Sci. U.S.A.* 70:298-302.
- Parker, P.J., Coussens, L., Totty, N., Rhee, L., Young, S., Chen, E., Stabe, S., Waterfield, M.D., and Ullrich, A. 1986. The complete primary structure of protein kinase C—the major phorbol ester receptor. *Science* 233:853-859.
- Richardson, C.C. 1971. Polynucleotide kinase from *Escherichia coli* infected with bacteriophage T4. *Nucl. Acids Res.* 2:815.
- Suggs, S.V., Wallace, R.B., Hirose, T., Kawashima, E.H., and Itakura, K. 1981. Use of synthetic oligonucleotides as hybridization probes: Isolation of cloned cDNA sequences for human β_2 -microglobulin. *Proc. Natl. Acad. Sci. U.S.A.* 78:6613.
- Wallace, R.B., Schaffer, J., Murphy, R.F., Bonner, J., Hirose, T., and Itakura, K. 1979. Hybridization of synthetic oligodeoxyribonucleotides to X174 DNA: The effect of single base pair mismatch. *Nucl. Acids Res.* 6:3543.
- Wetmur, J.G. 1976. Hybridization and renaturation kinetics of nucleic acids. *Ann. Rev. Biophys. Bioeng.* 5:337-361.
- Wood, W.I., Gitschier, J., Lasky, L.A., and Lawn, R.M. 1985. Base composition-independent hybridization in tetramethylammonium chloride: A method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl. Acad. Sci. U.S.A.* 82:1585-1588.
- Woods, D.E., Miarkham, A.F., Ricker, A.T., Goldberger, G., and Colten, H.R. 1982. Isolation of cDNA clones for the human complement protein factor B, a class III major histocompatibility complex gene product. *Proc. Natl. Acad. Sci. U.S.A.* 79:5661-5665.
- Yang, J.H., Ye, J.H., and Wallace, D.C. 1984. Computer selection of oligonucleotide probes from amino acid sequences for use in gene library screening. *Nucl. Acids Res.* 12:837-843.

Contributed by Allan Duby (SSC)
The University of Texas Health Science
Center at Dallas
Dallas, Texas

Kenneth A. Jacobs and Anthony Celeste
(TMAC)
Genetics Institute, Inc.
Cambridge, Massachusetts

PURIFICATION OF BACTERIOPHAGE, COSMID, AND PLASMID CLONES

After the screening procedure has identified potentially desirable clones they must be purified. The following protocols (which are similar to those in *UNITS 1.3* and *1.11*) minimize the number of steps because of the large number of clones on a plate and because of the difficulties associated with retesting these clones.

Purification of Bacteriophage Clones

UNIT 6.5

BASIC
PROTOCOL

Phage plates are correctly oriented to the autoradiograph film, and a region that should contain the clone of interest is sampled by toothpicking each phage plaque onto secondary plates containing a lawn of host cells. Alternatively, a plug of agarose can be taken from the primary plate, placed in SM, and this solution used to plate a small secondary library. Plaques on the secondary plates are transferred to nitrocellulose filters, hybridized to ³²P-labeled probe, and an isolated positive plaque is picked, diluted in SM, and regrown. This process is repeated until the desired plaque is purified.

Materials

- 0.7% top agarose (*UNIT 1.1*)
- Host bacteria (OD₆₀₀ 1.5 to 2 in 10 mM MgSO₄)
- LB plates (*UNIT 1.1*)
- Suspension medium (SM; *UNIT 1.11*)
- Chloroform
- Sterile round toothpicks (*UNIT 1.1*) or Pasteur pipet
- Nitrocellulose membrane filters
- Additional reagents and equipment for autoradiography (*APPENDIX 3*) and phage titering (*UNIT 1.11*)

Growth of secondary plaques

1. Plate 3 ml of 0.7% top agarose containing 200 µl host bacteria on 82-mm LB plates (one plate per clone) and allow to set 10 min.

Top agar cannot be used because it tends to lift off with the nitrocellulose filter.

2. Orient the autoradiograph to the primary library filters by radioactive tags that have been placed on the material used to support the filters. Then mark the autoradiograph at the points where the filters contain needle holes. Place plates containing the library on top of the autoradiograph on an X-ray view box and orient according to the needle marks.

To reduce the number of irrelevant clones, it is often helpful to produce two different exposures of the primary filters. One can then eliminate from consideration spots that do not appear on both autoradiographs.

- 3a. Insert toothpicks first into the primary plate in the area over the hybridization spot on the autoradiograph and then into the top agarose of one of the secondary plates prepared in step 1 above (see *UNIT 1.11*). A grid is helpful to guide spacing of the stabs 5 to 8 mm apart on the secondary plate. Initially attempt to pick individual plaques from the primary plate, then insert the toothpicks at random to ensure that an entire circular area with a 1-cm diameter has been well sampled. This usually requires 30 to 40 stabs per potential clone. It is advisable to make a duplicate secondary plate at the time of picking by simply sticking each toothpick into a second secondary plate in roughly the same pattern as the first.

- 3b. An alternate screening procedure to the above is as follows. Insert the large end of a Pasteur pipet into the top agarose of the primary plate to cut a circular plug corresponding to the region of the autoradiogram demonstrating radioactivity. Remove this plug with the Pasteur pipet or with a spatula and place into 1 ml SM with one drop of chloroform. Allow to sit 1 to 2 hr and then titer. As soon as the titer is known, make 3 to 6 plates with a density of <500 phage per plate. These secondary plaques are handled as in step 4 below.

Screening secondary plaques by hybridization

4. Grow secondary plates at 37°C overnight. Transfer plaques to nitrocellulose filters, process, hybridize, wash, and expose as outlined in *UNITS 6.1 to 6.4*. Mark filter orientation points on the autoradiograph and identify positive plaques on the secondary plates. Insert toothpick into the most strongly hybridizing plaque for each clone and placed into 1 ml SM for 5 min. Plate 1 µl of this phage stock and 1 and 10 µl of a 1:100 dilution onto tertiary LB plates.
5. Screen tertiary plates as above. Insert toothpick into an isolated hybridizing plaque and transfer to SM. This phage stock may be pure, but it is plated and evaluated by hybridization of these plates. If all plaques are positive, make a final SM stock from one of the plaques on these plates. Repeat these steps until the phage is pure. A high-titer stock solution can then be made, as outlined in *UNIT 1.12*.

COMMENTARY

Background Information

Careful purification of the clone of interest away from contaminating phage is required before growth and characterization of the clone can proceed. It is common for a “purified” clone to be contaminated by a second phage, leading to confusing results and wasted time. Several rounds of purification should be performed even if the phage appears pure as early as the secondary screening stage. This approach is similar to that presented by Kaiser and Murray (1984).

Critical Parameters

Plaque purification depends on the ability to go from a hybridization signal on the autoradiograph back to the correct plaque on the LB plate. Thus, attention to orientation of the filter to plate, filter to X-ray film, and film to plate is important.

If more than one potential clone is being plaque purified, it is important to prevent cross-contamination (e.g., via SM or agarose) between clones. Because only a single plaque is sampled at each step of purification, contamination could result in the loss of clones.

Troubleshooting

Occasionally, a clone will be lost during the purification process. When this happens it is usually best to go back as far as possible, i.e., to the secondary plates or to the SM stock.

Potential clones lost with the toothpick technique (step 3a) can sometimes be recovered by the plug procedure (step 3b). It is rarely worthwhile to attempt to purify a clone more than twice, since two unsuccessful attempts usually indicate a false primary clone.

Anticipated Results

This technique normally recovers >90% of clones confirmed by duplicate filters.

Time Considerations

A few hours are required every day for 5 to 8 days. Days picking and plating positive phage can alternate with days setting up filters and hybridization. Maintaining this schedule requires washing filters and obtaining an autoradiograph in 1 day.

Literature Cited

Kaiser, K. and Murray, N.E. 1984. The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries. *In* DNA Cloning: A Practical Approach, Vol. 1 (D.M. Glover, ed.) pp. 1-47. IRL Press, Oxford.

Contributed by Thomas Quertermous
Massachusetts General Hospital
Boston, Massachusetts

Purification of Cosmid and Plasmid Clones

UNIT 6.6

Cosmid- or plasmid-bearing colonies that are identified by hybridization are purified by spreading the cosmids or plasmids on an agar plate and repeating the colony hybridization.

**BASIC
PROTOCOL**

Materials

Cold LB medium containing antibiotic (UNIT 1.1)

LB plates containing antibiotic (UNIT 1.1)

Round toothpicks (UNIT 1.1)

Nitrocellulose membrane filters

Spreader (UNIT 1.3)

Additional reagents and equipment for plating bacteria (UNIT 1.3), plating and transferring plasmid/cosmid libraries (UNIT 6.2), plasmid minipreps (UNIT 1.6), and autoradiography (APPENDIX 3)

NOTE: All materials coming into contact with *E. coli* must be sterile.

1. Pick the positive clones, as detected by the in situ hybridization of nitrocellulose replica filters (UNITS 6.3 & 6.4), with a sterile toothpick.

If plate is dense with colonies, be sure to pick from a 3- to 5-mm circle to ensure selecting the correct clone.

2. Rinse the tip of the toothpick off into a microcentrifuge tube containing 1 ml cold LB medium with the appropriate antibiotic. Store these tubes at 4°C to inhibit any continued growth.

The vector must encode a gene conferring resistance to the appropriate antibiotic.

3. Plate out, using a sterile spreader, from 1 to 25 µl of the bacterial suspension onto an LB plate with the appropriate antibiotic. The correct number of clones to screen is from 25 to 250 per 100-mm plate. Allow the colonies to grow overnight at 37°C.
4. Make a replica copy of the bacterial lawn onto a nitrocellulose filter (UNIT 6.2). Denature, renature, bake, and hybridize as described in UNIT 6.2.
5. From the autoradiograph of the secondary plate, select the most isolated, positive colony. Grow the colony and isolate the DNA (UNIT 1.3).

COMMENTARY

Critical Parameters

If the colonies are too dense, purification of a single colony following a second round of hybridization is difficult. If the colonies are too sparse, many plates must be screened to identify a single hybridizing plaque. The number of clones required depends upon the original number of colonies picked in the original toothpicking. The purpose of rinsing the tip of the toothpick in cold LB medium and keeping the suspension cold is to stop the overgrowth of a single colony and to be able to reliably predict the titer of the suspension.

from 25 to 250 colonies per 100-mm plate—will allow the isolation of a single positive clone.

Time Considerations

Starting from a positive colony identification, this procedure requires one night for the colonies to grow, 10 min per plate for colony transfer to nitrocellulose, and one night for the second hybridization. If a probe with a high specific activity is used, an autoradiograph can be produced from the washed filters in a few hours.

Anticipated Results

Plating cosmid- or plasmid-bearing bacteria on an agar surface at the appropriate density—

Contributed by John H. Weis
Harvard Medical School
Boston, Massachusetts

Contributed by John H. Weis

Current Protocols in Molecular Biology (1989) 6.6.1

Copyright © 2000 by John Wiley & Sons, Inc.

**Screening
Recombinant
DNA Libraries**

6.6.1

Supplement 17

Antibodies that recognize a specific protein can be used to identify the desired clones in a recombinant DNA library. In this system the DNA inserts are cloned in an expression vector where the insert sequence should be expressed as part of a fusion protein, or the insert is used to select mRNA that can be translated *in vitro*. Both fusion proteins and *in vitro* translated mRNA can be detected with antibodies, as described in the two units of this section.

The basic requirement of the approaches described here is an antibody that is able to recognize the fusion form of the desired protein. Most investigators assume that the fusion form resembles its denatured form and thus use antibodies that are able to recognize the desired protein in a western blot (UNIT 10.8). A major problem with this technique is that false positive plaques are identified. A simple second test, independent of this technique (e.g., chromosomal assignment or tissue specificity), is extremely useful in establishing the validity of a positive clone.

UNIT 6.7

Immunoscreening of Fusion Proteins Produced in Lambda Plaques

Screening large numbers of plaques containing particular proteins is accomplished by techniques that are analogous to those described for screening with radioactive DNA probes. However, in the basic protocol described here the plaques are screened with antibodies specific to the desired proteins. The alternate protocol provides a method for increasing the amount of recombinant protein in each plaque by inducing expression from the *lac* promoter that directs its expression.

One requirement of the antibody screening procedure is that the recombinant sequences must be carried by expression vectors. A problem often encountered with this technique is that the cDNA library is not made in an expression vector.

BASIC PROTOCOL

SCREENING A λ gt11 EXPRESSION LIBRARY WITH ANTIBODIES

A particular cDNA clone in a bacteriophage λ gt11 library is identified by the protein that is produced after infection of *E. coli*. Bacteriophage from a λ gt11 library are plated. The proteins produced during phage growth and cell lysis are transferred to nitrocellulose filters. The filters are blocked with protein (nonfat milk) to prevent nonspecific interactions with antibody, incubated with antibody, washed, reacted with radiolabeled second antibody, washed again, and autoradiographed. The clones that are identified by autoradiography are isolated by plaque purification.

Materials

- λ gt11 cDNA expression library
- 150-mm LB plates (UNIT 1.1)
- E. coli* LE392 (Table 1.4.5)
- 1% LB top agar (UNIT 1.1)
- 0.05% (v/v) NaN_3 in India ink (optional)
- Immunoscreening buffer
- First-stage antibody
- ^{125}I -labeled second-stage reagent reactive with first-stage antibody

Immunoscreening of Fusion Proteins Produced in Lambda Plaques

6.7.1

Contributed by Thomas P. St. John

Current Protocols in Molecular Biology (1990) 6.7.1-6.7.6

Copyright © 2000 by John Wiley & Sons, Inc.

132-mm nitrocellulose membrane filters

Additional reagents and equipment for titrating and plating bacteriophage (UNITS 1.11 & 6.1) and autoradiography (APPENDIX 3)

NOTE: All materials coming into contact with *E. coli* must be sterile.

1. Titer and plate a λ gt11 cDNA library with *E. coli* LE392 on 150-mm LB plates, using 7 ml of 1% LB top agar per plate (see UNIT 6.1).

We have not found it necessary to use lysogenic inducing, protease-deficient (lon) strains, or late lytic inductions of the lacZ-cDNA fusion genes in order to detect antigenic material on a solid support. Normally E. coli LE392 is used for λ gt11.

2. Incubate plates 8 hr at 37°C.
3. Lay a numbered 132-mm nitrocellulose filter on the plate (UNIT 6.1, step 6).
4. Incubate the plates and filters overnight at 37°C.

Continued plaque growth in close contact with nitrocellulose strongly enhances the autoradiographic signal over that achieved by short-term contact at any time during plaque growth. This may be due to enhanced binding of fusion protein during lysis of infected bacteria that grow in the membrane rather than in the top agar.

5. Mark each filter asymmetrically with needle holes, as done for DNA plaque filter hybridization (UNIT 6.1, step 7).

It is helpful to add sodium azide to 0.05% to India ink to prevent contamination of the plaque plates.

6. Remove the nitrocellulose filters and block the protein binding sites by washing the filter in immunoscreening buffer at room temperature for 30 min. Repeat this wash 2 to 4 times to remove the bulk of the bacteria from the filter.
7. Incubate the filters with the first-stage antibody at a concentration of 0.5 to 10 μ g/ml diluted in immunoscreening buffer in a heat-sealed bag for 2 to 24 hr at 4°C on a horizontal shaker platform. Multiple filters may be placed in a single bag as long as there is sufficient liquid so that mixing between the filters occurs.

*In order to stabilize immune complexes, all reactions and washings should be done at 4°C with cold buffer. The dissociation rate of a high-affinity antibody is probably low enough that the antibody reaction could be performed at room temperature. Low-affinity sera, and especially monoclonal antibodies, may be more of a problem. Crude polyclonal rabbit serum, affinity-purified, or ammonium sulfate-precipitated sera have all been used with positive results. Monoclonal antibodies in a variety of forms including ascites fluid and tissue culture supernatants have also been used successfully. It is not generally necessary to remove immune reactivity to bacterial proteins prior to screening filters. However, if reaction to all of the plaques is observed, first-stage antibody can be absorbed with *E. coli* extracts. *E. coli* extracts are commercially available for the absorption of antibodies reactive to *E. coli* proteins. Another method for removing antibody specific for bacterial or bacteriophage proteins is to reuse first-stage antibody.*

8. Wash the filters 4 to 5 times in immunoscreening buffer at 4°C for 5 to 10 min per wash.

It is necessary to consider the dilution factor during washing to be sure that the concentration of residual first-stage antibody in the detection-stage probing is insignificant.

9. Incubate the filters with ^{125}I -labeled second-stage antibody, at a concentration of 0.5×10^6 cpm/ml diluted in immunoscreening buffer in heat-sealed bags for 2 to 6 hr at 4°C .

The second-stage antibody should react very specifically with the first-stage antibody. The use of protein A to detect those antibodies able to bind it is recommended. The use of anti-immunoglobulin as a second-stage reagent may result in the detection of cDNA clones by the anti-immunoglobulin alone. Several manufacturers produce enzyme-linked chromagenic detection kits for use in phage plaque immunological screening procedures. It has been this author's experience that the signals obtained with these kits do not match the absolute signal-to-noise ratio offered by ^{125}I -labeled second-stage reagent. Should the use of such kits be desirable, it is recommended that the researcher carefully follow the manufacturer's directions.

10. Wash 4 to 5 times in immunoscreening buffer at 4°C .

Be aware of local isotope disposal recommendations and treat waste accordingly.

11. Blot dry, wrap in plastic wrap, and expose to X-ray film with an intensifying screen at -70°C .

12. Purify λ cDNA fusion-protein clones by repeated dilutions until pure, and grow for DNA preparation (UNIT 6.5).

The techniques used for picking plaques are similar to those used in DNA hybridization screening.

SHORT PROTOCOL

SCREENING A λ gt11 EXPRESSION LIBRARY WITH ANTIBODIES

1. Titer and plate a λ gt11 cDNA library with *E. coli* LE392 on LB plates with 7 ml of 1% LB top agar/plate. Incubate 8 hr at 37°C .
2. Lay a numbered nitrocellulose filter on plate and incubate overnight at 37°C .
3. Mark each filter with needle holes and India ink, remove, and wash 3 to 5 times in immunoscreening buffer, 30 min per wash at room temperature.
4. Incubate filters with first-stage antibody (0.5 to 10 $\mu\text{g}/\text{ml}$ in immunoscreening buffer) in a heat-sealed bag, 2 to 24 hr at 4°C on a horizontal shaker platform.
5. Wash filters 4 to 5 times in cold immunoscreening buffer, 5 to 10 min per wash at 4°C . Incubate with ^{125}I -labeled second stage antibody (0.5×10^6 cpm/ml in immunoscreening buffer) 2 to 6 hr at 4°C .
6. Wash filters 4 to 5 times as described in step 5. Blot dry, wrap in plastic wrap, and expose to X-ray film at -70°C with intensifying screen.
7. Purify λ cDNA fusion-protein clones by repeated dilutions until pure, and grow for DNA preparation (UNIT 6.5).

INDUCTION OF FUSION PROTEIN EXPRESSION WITH IPTG PRIOR TO SCREENING WITH ANTIBODIES

ALTERNATE
PROTOCOL

The probability of success in screening a λ gt11 cDNA library with an antibody can sometimes be increased by preventing the expression of the fusion protein until the plaques are well established. The expression of the potential β -galactosidase-cDNA fusion proteins can be induced after 3 to 4 hr of plaque growth by placing a nitrocellulose filter containing the inducer IPTG onto the plate and continuing growth at 37°C. The nitrocellulose filters are then screened with antibodies as in the basic protocol.

Additional Materials

E. coli Y1090 (Table 1.4.5)
10 mM IPTG (Table 1.4.2)
42°C room or incubator

1. Absorb 10^4 to 5×10^4 cDNA-fusion λ phage (UNIT 1.11) with 0.5 to 1.0 ml *E. coli* Y1090. These cells should be from a fresh overnight culture grown at 37°C.

E. coli Y1090 cells express high levels of the lac repressor, ensuring that no fusion gene expression occurs until plaque growth is well established and the strain bears a mutation in a bacterial protease gene (*lon*), thus reducing degradation of the recombinant protein.

2. Plate on 150-mm LB plate with 7 ml LB top agar.

These plates should be reasonably fresh, poured 2 to 3 days before use. Plates that are too wet will have a tendency to lose the top agar layer to the first or second piece of nitrocellulose placed on the plate. Using 1% agar (instead of 0.7%) for the top layer will reduce this tendency.

3. Incubate the plates 3.5 hr at 42°C.

This higher temperature incubation should make any fusion protein produced as unstable as possible, as well as ensure that the temperature-sensitive λ cI857 repressor is completely denatured.

4. While the plates are incubating, soak a 132-mm nitrocellulose filter in 10 mM IPTG. Dry the filter.

5. Lay the filters on the plates bearing the bacteriophage library. Incubate 3.5 hr at 37°C. Mark each filter and remove from plate. Block the remaining protein-binding capacity and probe the filters as described in basic protocol (steps 6 to 12).

The IPTG in the filter induces expression of the phage lacZ-cDNA fusion gene. A second filter may be applied to the plate after removal of the first filter. In this case, incubate the plate at 37°C for an additional 3 hr before marking and probing the second filter.

REAGENTS AND SOLUTIONS

Immunoscreening buffer

Prepare in phosphate-buffered saline (PBS):
5% nonfat dry milk
0.1% Nonidet P-40 (NP-40)
0.05% sodium azide (made from 5% stock solution)

Screening
Recombinant
DNA Libraries

6.7.4

COMMENTARY

Background Information

The use of *in situ* immunoassays to isolate bacteriophage or plasmid recombinants expressing fusion proteins is straightforward in both concept and execution. Normally, production of a cDNA library of sufficient quality is more difficult than screening the library. The various techniques used to date are the generally standard immunoassays. The immunological screening of bacterially synthesized fusion proteins produced by either plasmids or phage involves two basic technical procedures: the synthesis and immobilization of antigenic material to a solid support followed by a sensitive detection procedure.

Demonstration that a cDNA encodes an antigenic determinant does not prove that it encodes the protein of interest. It is essential to note that the isolation of DNA sequences encoding an antigenic determinant (e.g., with a monoclonal antibody or an antisera directed, as most may be, against only a single epitope within any given cDNA sequence) carries the risk that the detected determinant is not encoded by the desired sequence, but by a sequence related only at the level of protein product structure. Additionally, it is important to determine early in the experiment that the autoradiographic signal of a positive clone is dependent upon the first antibody stage.

Antibody reactive sequences should be tested for the presence of at least one other distinguishing property that the desired sequence may reasonably be expected to have. These distinguishing properties may be unusual mRNA or encoded protein sizes, regulated expression in appropriate cell types, expression of the expected gene product or an appropriate genetic function by transfected genomic DNA or cDNA expression vector clones, or the expression of a second, different antigenic determinant in common with the known protein within the same clone. In general, clones isolated by monoclonal antibodies directed to different epitopes of the same protein would either cross-hybridize or be contained within the same full-length clone if the identified sequences were correct.

Literature Review

The development of *in situ* immunoassays for the detection of cloned gene products expressed on either plasmids or bacteriophage λ began no later than the report by Skalka and

Shapiro (1976) on immunological techniques for the detection of β -galactosidase protein in phage plaques and bacterial colonies. These techniques depended on the production of an immunoprecipitate within the plaque or colony and were superseded by more sensitive techniques based on the detection of immobilized antigen with ^{125}I -labeled detection stages. The immobilization techniques used in these procedures were either direct or required the binding of antigen to antibody followed by detection with a second antibody (Erlich et al., 1978; Kemp and Cowman, 1981; Helfman et al., 1983; Young and Davis, 1983).

Critical Parameters

It is important to ensure that the density of the plated phage is not too high. Most of the protein during phage growth is produced relatively late in the infection at the edge of the plaque when the largest number of cells have been infected. As a result, if such a large number of phage are present that plaques begin to touch one another, fewer cells will be infected in the area of contact, less protein will be produced, and a lower signal will result. The same consideration applies to DNA plaque filter hybridizations.

The antibody used as a probe must be as pure in immunological activity as possible. These techniques are exquisitely sensitive to immunological impurities and already suffer from the problems of cross-reactivity with other antigens produced as fusion proteins. The ideal probe would be a collection of monoclonal antibodies directed to different epitopes on the same protein. Used separately, these probes would each have their own special patterns of cross-reactivity with related protein structures. An immunologically very pure antisera may have activity directed to multiple epitopes on the same protein, but all potential cross-reactivities are present at once. It is possible that a sera may give a higher signal than a monoclonal antibody, but this does not seem to be routinely supported by experience. Signal strengths vary considerably. The higher the specific titer, the more likely it is that a positive signal will be detected. There is no obvious correlation as yet with the type of probe, monoclonal versus sera, or the isotype of the monoclonal. Be aware that occasional bacterial contaminants may give "signals."

It is the preference of this author to use ¹²⁵I-labeled reagent for reasons of increased sensitivity and for the ability to make multiple autoradiograms of each experiment.

Time Considerations

These procedures are well suited for 1- to 3-day cycle times. Generally, the phages are plated in the morning of the first day and the filters placed on the plates at the appropriate time. In the basic protocol, the filters are left on the plates overnight. The antibody probing may be completed the second day, or if the first-stage probing is also done overnight, the detection stage probing may be done on the third day. After probing with the first antibody, it is probably a good idea to finish the probing as quickly as possible to avoid any problem of dissociation of the primary antibody from the filter-bound antigen. In the alternate protocol, the filters may be either stored wet overnight or incubated in the first-stage antibody and the probing continued the following day.

Literature Cited

Erlich, H.A., Cohen, S.N., and McDevitt, H.O. 1978. A sensitive radioimmunoassay for detecting products translated from cloned DNA fragments. *Cell* 13:681-689.

Helfman, D.M., Feramisco, J.R., Fiddes, J.C., Thomas, G.P., and Hughes, S.H. 1983. Identification of clones that encode chicken troponin by direct immunological screening of a cDNA expression library. *Proc. Natl. Acad. Sci. U.S.A.* 80(1):31-35.

Kemp, D.J. and Cowman, A.F. 1981. Direct immunoassay for detecting *Escherichia coli* colonies that contain polypeptides encoded by cloned DNA segments. *Proc. Natl. Acad. Sci. U.S.A.* 78(7):4520-4524.

Skalka, A. and Shapiro, L. 1976. In situ immunoassays for gene translation products in phage plaques and bacterial colonies. *Gene* 1:65-79.

Young, R.A. and Davis, R.W. 1983. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. U.S.A.* 80(5):1194-1198.

Key Reference

Huynh, T.V., Young, R.A., and Davis, R.W. 1984. Construction and screening cDNA libraries in λ gt10 and λ gt11. In *DNA Cloning: A Practical Approach*, Vol. 1 (D.M. Glover, ed.) pp. 49-78. IRL Press, Oxford.

Provides an excellent description of immunological screening procedures.

Contributed by Thomas P. St. John
Fred Hutchinson Cancer Research Center
Seattle, Washington

Immunoscreening after Hybrid Selection and Translation

BASIC PROTOCOL

In this procedure plasmid cDNA clones are screened for their ability to select a particular mRNA. Plasmid DNA is bound to nitrocellulose filters, hybridized to mRNA, washed, and the selected mRNA is eluted from the filter. Eluted mRNA is characterized by translation into ³⁵S-labeled protein, which is identified by immunoprecipitation and denaturing (SDS) polyacrylamide gel electrophoresis. The desired clone is that which is able to select an mRNA that translates into the desired protein. This procedure can be modified to characterize cosmids and bacteriophage DNA.

Materials

Brain-heart-infusion (BHI) medium (37.5 g/liter, autoclaved) containing appropriate antibiotics
 Chloramphenicol (Table 1.4.1)
 TE buffer, pH 7.6 (APPENDIX 2)
 1 M NaOH
 Neutralization solution
 6× SSC (APPENDIX 2)
 Hybridization solution IV
 Poly(A)⁺ mRNA (UNIT 4.5)
 65°C TES buffer in 0.5% SDS
 65°C TES buffer
 10 mg/ml yeast tRNA
 Buffered phenol (UNIT 2.1)
 50:1 chloroform/isoamyl alcohol
 3 M sodium acetate, pH 5.2
 Ethanol
 Translation mixture
 [³⁵S]methionine (800 Ci/mmol)
 Immunoprecipitation buffer
 Nonimmune serum
 Protein A–Sepharose suspension
 Polyclonal or monoclonal antibodies (Chapter 11)
 High-salt immunoprecipitation buffer
 2× SDS/sample buffer (UNIT 10.2)
 96-well microtitration dish
 Beckman JS-4.2 rotor or equivalent
 Sterile 15-ml capped glass culture tubes
 0.45-μm nitrocellulose filters (2.5-cm diameter)
 Multifilter washing apparatus
 80°C vacuum oven
 Sterile 1.8-ml round-bottom plastic tubes (Nunc)
 Sterile silanized 1.5-ml microcentrifuge tubes (APPENDIX 3)
 Sterile needles
 Additional reagents and equipment for preparation of plasmid DNA (UNIT 1.6) and denaturing (SDS) polyacrylamide gel electrophoresis (UNIT 10.2)

Isolate plasmid DNA

1. Pick individual cDNA clones into wells of a microtiter dish containing 0.25 ml BHI medium plus the appropriate selective antibiotics. Grow overnight at 37°C.
2. Inoculate 50 ml BHI/antibiotic medium in a 250-ml flask, with 0.1 ml from each of 10 individual overnight clone cultures. Grow at 37°C in thermo-regulated shaker

until A_{590} is 0.7. Add chloramphenicol to a final concentration of 100 $\mu\text{g/ml}$ and grow overnight at 37°C. Harvest cells by centrifugation 10 min at 3000 rpm (2000 \times g), 4°C.

The number of clones in a pool that would allow the detection of an individual clone depends on the abundance of the specific mRNA and on the sensitivity of the immunodetection assay. If none of these parameters is known, 10 clones per pool are recommended. Manipulation of 10 to 20 different pools in parallel throughout the entire process is reasonable to handle.

3. Prepare plasmid DNA as described in UNIT 1.6.

Bind plasmid DNA to nitrocellulose filters

4. Dilute ~50 μg crude plasmid DNA in 1.5 ml TE buffer, transfer the solution into 15-ml sterile, capped glass culture tubes, incubate 10 min in a boiling water bath, immediately add 1.5 ml of 1 M NaOH, and leave at room temperature for 10 min. Add 9 ml neutralization solution, mix well, and place on ice while checking the pH.

pH after neutralization should be 6.5 to 7.5; otherwise, correct with NaOH or HCl. Do this as quickly as possible, as renaturation of DNA before binding to the filter should be prevented.

5. Place a 0.45- μm nitrocellulose filter (2.5-cm diameter) on porous support attached to a vacuum line. Pour denatured DNA solution (~50 μg plasmid DNA) over the filter at a flow rate of 1 ml/min. After all the DNA solution passes through the filter, continue suction for 3 min, increasing the vacuum to maximum. Wash each filter with 50 ml of 6 \times SSC. Fix the DNA to the filter by baking at 80°C in vacuum for 2 hr.

Binding of the DNA to the filter can be done on any of the commercially available filter washing devices. An apparatus allowing the manipulation of 10 filters simultaneously is preferable.

Once the filters have been baked at 80°C, be careful to handle and store in a sterile manner.

6. Punch out disks of 0.5 cm diameter from the filter using a sterile one-hole paper punch. Mark the filter disks with a ballpoint pen.

Select desired mRNA by hybridizing to plasmid DNA

7. Place 0.3 ml hybridization solution in a sterile, capped plastic tube (1.8-ml round bottom Nunc tubes are the most suitable for this purpose). Add 10 to 50 μg poly(A)⁺ mRNA, preheat 10 min at 70°C, then put up to ten filter disks in the tube and incubate 2 hr at 50°C.
8. Transfer filter disks to a 50-ml test tube (up to 20 filters per tube), wash with 25 ml TES/0.5% SDS buffer prewarmed to 65°C, hand vortex 0.5 min, and remove supernatant using a sterile Pasteur pipet connected to a vacuum line. Repeat nine more times and then wash twice with 25 ml prewarmed (65°C) TES buffer.
9. Transfer individual filters to sterile silanized 1.5-ml microcentrifuge tubes. Add to each tube 0.3 ml sterile water and 2 μl of 10 mg/ml yeast tRNA. Boil 60 sec then quick-freeze in a dry ice/ethanol bath and thaw at room temperature.

Commercial preparations of tRNA should be phenol extracted several times prior to use.

10. Remove filters with a sterilized needle, add to each eluted RNA solution 0.15 ml buffered phenol and 0.15 ml chloroform/isoamyl alcohol and extract once. Add 30 μl of 3 M sodium acetate to the aqueous solution and then precipitate the RNA with

2 vol ethanol. Pellet the RNA by a 15-min spin in a microcentrifuge. Wash the pellets with 0.5 ml ethanol, dry by lyophilization, and resuspend in 10 μ l water.

The RNA solutions should be stored at -70°C , or preferably used immediately in the translation reaction.

Translate selected mRNA and characterize by immunoprecipitation

11. To 5 μ l hybrid-selected RNA add 10 μ l translation mixture containing ^{35}S -labeled methionine. (Use amount of label recommended by the manufacturer of the translation mixture.) Incubate 60 min at 30°C .

At this stage the procedure can be discontinued by freezing the translation mixture at -70°C . The efficiency of the translation reaction can be monitored by determining ^{35}S -labeled methionine incorporation. To do so, apply 1 μ l translation mixture to 3-mm filter paper, place the filter in boiling 10% TCA for 10 min, rinse filter twice with ethanol, dry, and count incorporated label in a liquid scintillation counter.

12. To 15 μ l translation mixture add 15 μ l immunoprecipitation buffer and 1 μ l nonimmune serum (serum from a nonimmunized animal, or normal ascites fluid if monoclonal antibodies will be used for immunoprecipitation). Incubate 10 min at room temperature, then add 40 μ l protein A–Sepharose suspension and leave at room temperature for 30 min. Spin down Sepharose 2 min in microcentrifuge, carefully transfer the supernatant to another microcentrifuge tube, and discard the Sepharose pellet.

This step removes polypeptides from the translation mixture that bind nonspecifically to Sepharose or to antibodies.

13. Add 1 μ l polyclonal or monoclonal antibodies directed against the relevant gene product to the supernatant and incubate 10 min at room temperature. Add 40 μ l protein A–Sepharose suspension and incubate at room temperature for another 30 min. Spin down Sepharose and discard the supernatant.
14. Wash the Sepharose beads three times with 1 ml immunoprecipitation buffer, once with 1 ml high-salt immunoprecipitation buffer, and once with 1 ml water.
15. Resuspend pellet in 20 μ l $2\times$ SDS/sample buffer, boil 10 min to elute bound polypeptides from Sepharose protein-A beads. Spin down the Sepharose and put 15 to 20 μ l into slots of polyacrylamide gel.

Fractionate immunoprecipitated protein on denaturing gel

16. Run a denaturing SDS/polyacrylamide gel at the conditions optimal for separation of the polypeptide. Dry the gel and expose it to autoradiography.

Remember to run in parallel an appropriate control containing the polypeptide encoded by the gene for which selection is being made.

17. If one of the lanes contains the required polypeptide band, subject the individual clones of the positive pool to the same procedure.

REAGENTS AND SOLUTIONS

High-salt immunoprecipitation buffer

Same recipe as immunoprecipitation buffer, except NaCl is 0.5 M

Hybridization solution IV

65% deionized formamide
0.4 M NaCl
0.2% sodium dodecyl sulfate
30 mM PIPES, pH 6.5
50 µg yeast tRNA
50 to 500 µg poly(A)⁺ mRNA (see critical parameters)

Hybridization solution should be made fresh prior to use.

Immunoprecipitation buffer

10 mM Tris-Cl, pH 7.4
2 mM EDTA
0.15 M NaCl
10% (v/v) Nonidet P-40

Neutralization solution

200 ml 20× SSC (APPENDIX 2)
100 ml 1 N HCl
100 ml 1 M Tris-Cl, pH 8

Protein A–Sepharose suspension

1.5 g of Sepharose Protein-A (Pharmacia) is resuspended in:

10 mM Tris-Cl, pH 7.5
0.15 M NaCl
0.4% Triton X-100 (v/v)
0.5% Aprotinin (v/v) (Sigma)

Shake for 5 min. Spin down beads, wash 3 times with same buffer, then resuspend in 11 ml of the buffer.

TES buffer

0.15 M NaCl
10 mM Tris-Cl, pH 7.6
1 mM EDTA

Translation mixture

Any reticulocyte or wheat germ translation mixture system can be used; prepare as recommended by manufacturer.

COMMENTARY

Background Information

In hybrid selection screening, the various immobilized cDNA clones select homologous mRNA molecules from an mRNA pool. Individual mRNA species can then be eluted and used to direct the synthesis of the corresponding polypeptide in an *in vitro* translation system. This will yield a battery of polypeptides encoded by the various genes represented in a given cDNA library. The actual screening step in the procedure involves the specific identifi-

cation of the polypeptide molecule.

In the protocol described here, precipitation by homologous antibodies is suggested for the recognition of the desired polypeptide. The immunodiagnostic steps can be replaced by monitoring the specific biological activity of the polypeptide, providing the test of this activity is sensitive enough and the activity does not require interaction with multiple polypeptides (March et al., 1985).

The hybrid selection method can also be

used as a backup for other screening procedures. Subtractive hybridization screening procedures or hybridization with mixed oligonucleotides (see *UNITS 6.3* and *6.4*, respectively) are often ambiguous and result in the isolation of other clones in addition to the ones that are required. Here the hybrid selection system can be easily applied to distinguish between the various isolates and identify the correct one (Lemke and Axel, 1985).

This protocol is designed for screening a large number of cDNA clones and for isolating cDNAs corresponding to rare mRNAs.

Literature Review

Hybridization of mRNA to filter-immobilized DNA followed by translation of the eluted mRNA as a means for identification of specific DNA sequence was first suggested by Harpold et al. (1978) and Ricciardi et al. (1979). This approach was adapted for the screening of cDNA libraries by Parnes et al. (1981), upon which the protocol given here is based. Hybrid selection screening has been used for the identification of numerous genes during the past several years. Among the more recent examples are the Interleukin I gene (March et al., 1985) and structural protein of myelin cells (Lemke and Axel, 1985).

Critical Parameters

The number of cDNA clones that can be applied on one filter, the amount of mRNA that should be used for hybridization, and the temperature and duration of the hybridization can vary significantly from one clone to the other. These parameters depend on the abundance of the specific mRNA species in the mRNA pool, on the sensitivity of the assay used for polypeptide identification, and on the G:C content of the nucleic acids involved. The protocol given here is as "universal" as possible, but one should experiment with the above-mentioned parameters to obtain optimal results.

Degradation of the mRNA during hybridization or elution from filters is among the common problems of this method. RNA degradation can be easily monitored by incorporating into the assay a labeled SP6-derived mRNA transcript and its corresponding cDNA as an internal control. Such a control also allows a quantitative evaluation of each of the steps of the procedure.

Anticipated Results

A cDNA clone is considered positive if a filter carrying the cDNA binds an mRNA spe-

cies that can be translated into a polypeptide recognized by the specific antibody. The sequence of all positive clones should be determined for final verification. False positive clones may be detected by this method due to cross-reactivity of the antibodies or presence of additional specificities in the antiserum preparation.

Time Considerations

When planning hybrid selection screening it is advisable to separate the preparation of filters carrying the cloned DNA from the hybridization and translation steps. Filters can be prepared in advance and stored under sterile conditions for a couple of weeks. Preparation of 10 to 30 filters per day is very reasonable so that within a week, as many as 500 to 1500 individual clones can be transferred onto nitrocellulose. Hybridization to mRNA can be done in batches of 10 filters; 20 to 50 filters can be processed in one experiment and translated the same day. The procedure can be interrupted either after the translation by freezing the translation mixture or after immunoprecipitation by freezing the eluates recovered from the Sepharose beads.

Literature Cited

- Harpold, M.M., Dobner, P.R., Evans, R.M., and Bancroft, F.C. 1978. Construction and identification by positive hybridization translation of a bacterial plasmid containing a rat hormone structural gene sequence. *Nucl. Acids Res.* 5:2039.
- Lemke, G. and Axel, R. 1985. Isolation and sequence of a cDNA encoding the major structural protein of peripheral myelin. *Cell* 40:501.
- March, C.J., Mosley, B., Larsen, A., Cerret, D.P., Braedt, G., Price, V., Gillis, S., Henney, C.S., Krunheim, S.P., Grabstein, K., Canlon, P.J., Hopp, P., and Cosman, D. 1985. Cloning, sequence, and expression of two distinct human Interleukin I cDNAs. *Nature* 315:641.
- Parnes, J.R., Velan, B., Felsenfeld, A., Ramanathan, L., Ferrini, U., Appella, E., and Seidman, J.G. 1981. Mouse β 2-microglobulin cDNA clones: A screening procedure for cDNA clones corresponding to rare mRNAs. *Proc. Natl. Acad. Sci. U.S.A.* 78:2253.
- Ricciardi, R.P., Miller, J.S., and Roberts, B.E. 1979. Purification and mapping of specific mRNAs by hybridization selection and all free translation. *Proc. Natl. Acad. Sci. U.S.A.* 76:4921.

Contributed by Baruch Velan
Israel Institute for Biological Research
Ness Ziona, Israel

YEAST ARTIFICIAL CHROMOSOME LIBRARIES

Overview of Strategies for Screening YAC Libraries and Analyzing YAC Clones

Emphasis on identification of disease genes by positional cloning has underscored the need to clone fragments of genomic DNA >100 kb into a vector. The size of genomic inserts that can be carried in traditional cloning vectors has been limited to 20 to 25 kb for λ vectors and 40 to 45 kb for cosmid vectors. These vectors are of limited utility for analyzing very large genes or for “walking” to disease genes from DNA markers that may be 1 to 2 Mb away. Considerable progress has been made in cloning large DNA fragments in *Saccharomyces cerevisiae* using yeast artificial chromosome (YAC) vectors (see Fig. 13.4.6). YACs containing inserts that are >1 Mb have been produced and these are routinely propagated with apparent stability, suggesting that the major limitation to the size of YAC inserts is the quality of the starting genomic DNA. Large “core” laboratories that generate human YAC libraries—such as the Center for Genetics in Medicine, Washington University School of Medicine, St. Louis; the Centre d’Etude du Polymorphisme Humain (CEPH), Paris; and the Genome Analysis Laboratory, Imperial Cancer Research Fund, London—prepare human YACs with average insert sizes ranging from 0.3 to 1.2 Mb. Additional high-quality YAC libraries have been constructed using inserts from *Drosophila melanogaster*, *Caenorhabditis elegans*, *Schizosaccharomyces pombe*, and mouse (Burke et al., 1991; Rossi et al., 1992).

Anecdotal reports indicate YAC libraries may support the propagation of certain insert sequences that are poorly represented in *Escherichia coli*-based libraries. The YAC cloning system also offers the advantage that large genomic YAC inserts can be easily manipulated in yeast by homologous recombination. Thus, it is relatively simple to truncate a YAC insert or to introduce specific deletions, insertions, or point mutations with high efficiency using methods such as those described in UNIT 13.10.

This unit provides an introduction to the use of yeast artificial chromosome-bearing yeast clones (hereafter referred to as YAC clones) in genome analysis. It describes criteria for de-

signing a polymerase chain reaction (PCR) assay to be used in screening a YAC core library and discusses the rationale for verification and characterization of YAC clones obtained from these core laboratories. Protocols for maintaining YAC clones, analyzing YAC insert structure, preparing YAC DNA, and subcloning YAC inserts into other vectors are presented in UNIT 6.10. These protocols are outlined in the flow chart in Figure 6.9.1.

GENERATING YAC LIBRARIES

Although YAC cloning is the method of choice when insert sizes >100 kb are required, a number of features of the system have interfered with its rapid assimilation for routine cloning. Because the *S. cerevisiae* genome is at least an order of magnitude more complex than the *E. coli* genome and existing YACs are carried as only a single copy within yeast cells, the signal-to-noise ratio is less favorable for identifying a cognate clone in a YAC library than in a λ or cosmid library. Moreover, efforts to develop high-density screening methods for YACs have enjoyed only limited success. Most laboratories that maintain YAC libraries organize them as collections of individual clones in 96-well microtiter plates, which can be replicated faithfully and kept frozen for storage; in this form, a standard library representing 5 to 8 genome-equivalents comprises more than 500 microtiter plates. As a result, the effort and resources required to construct YAC libraries and prepare them for screening are enormous. Consequently, it is generally most practical for investigators wishing to obtain YACs carrying a specific DNA sequence to arrange for screening of a preexisting library maintained by a core laboratory.

Initially, YAC libraries were constructed with total genomic DNA (Burke et al., 1987). More recently, there has been interest in generating libraries from targeted DNA using somatic cell hybrids carrying a specific chromosome or portion of a chromosome. The feasibility of this approach has been demonstrated with the construction of a library carrying a portion of the human X chromosome (Abidi et

al., 1990). Additional targeted libraries are in the late stages of development and should reduce the cost and effort of screening for loci whose chromosomal location has been established.

YAC LIBRARY SCREENING BY A CORE LABORATORY

Methods used by YAC core laboratories for library screening evolve rapidly. It is possible to screen a library by hybridizing a single-copy probe to nylon filters stamped with a replica of one or more microtiter arrays. However, because of the low signal-to-noise ratio for hybridization and the substantial cost required to produce all of the nylon filter replicas, most laboratories perform library screening using PCR (Green and Olson, 1990).

At the time of this writing, most core facilities first extract DNA from pools of clones, usually representing 1 to 4 microtiter plates (96 to 384 YACs) per pool, and then combine this pooled DNA into more superpools of 1500 to 2000 YACs. The pools and superpools are screened by PCR to identify candidate microtiter plates containing at least one amplifying YAC clone. Final identification of the clone is most commonly performed either by colony hybridization using the PCR product as the probe or by screening pools of rows and col-

umns from the same microtiter plate using PCR. The time required for a YAC core laboratory to verify the specificity and parameters of the PCR assay and screen complex clone pools and subpools is usually 3 to 8 weeks.

As an example, the screening strategy used by one major core laboratory is described in the accompanying box. It should be noted that this procedure may change as technology advances; for instance, the recent advent of techniques providing reliable DNA extraction from small quantities of thousands of individual clones has made it feasible to screen individual wells on a plate and eliminate the laborious filter-hybridization step.

DESIGNING A LOCUS-SPECIFIC PCR ASSAY FOR SCREENING

An investigator arranging with a core laboratory for library screening is required to design a strategy for detecting the inserted genomic DNA and to provide the appropriate probe(s). It is worth investing considerable effort to create a convenient and reliable assay because the assay's success depends on its ability to detect the target sequence with high sensitivity while being insensitive to the presence of large excesses of yeast and plasmid sequences. Because a core laboratory must adopt PCR assays that have been imported

EXAMPLE: SCREENING OF HUMAN-GENOME YAC LIBRARY AT THE WASHINGTON UNIVERSITY SCHOOL OF MEDICINE

At this core facility, screening of the human-genome YAC library proceeds in three stages: (1) initial evaluation of the PCR assay; (2) screening of pools of YACs; and (3) identification of individual YACs from subpools to the single well by filter hybridization.

To permit pretesting of assays before they are sent to a screening core, new PCR assays are evaluated using four control DNA samples as templates: (1) CGM-1 human genomic DNA (33 ng/ μ l) from a lymphoblastoid cell line established from the donor whose DNA was used in preparing the YAC library; (2) YY212 DNA from a yeast strain carrying a YAC whose insert is yeast chromosomal DNA; (3) "single-membrane-pool" DNA (33 ng/ μ l) prepared from a pool of 396 YAC isolates; and "spiked-pool" DNA, which is single-membrane-pool DNA augmented with 5 ng/ μ l of CGM-1 DNA.

CGM-1 DNA serves as a positive control to demonstrate that the sensitivity of the PCR assay is adequate. YY212 DNA serves as a negative control, demonstrating that no product is amplified from either yeast host genomic DNA or the YAC vector. DNA from the single-membrane pool and the spiked pool provide additional negative/positive controls that more closely mimic library screening conditions. A negative signal from single-membrane-pool DNA demonstrates lack of cross-reactivity of the probe with the YAC vector, yeast genomic DNA, or common human repetitive sequences. A positive signal obtained from the spiked-pool DNA (containing only 5 ng/ μ l of CGM-1 DNA) is a strong indication that the assay possesses sufficient sensitivity against a yeast DNA background for successful library screening.

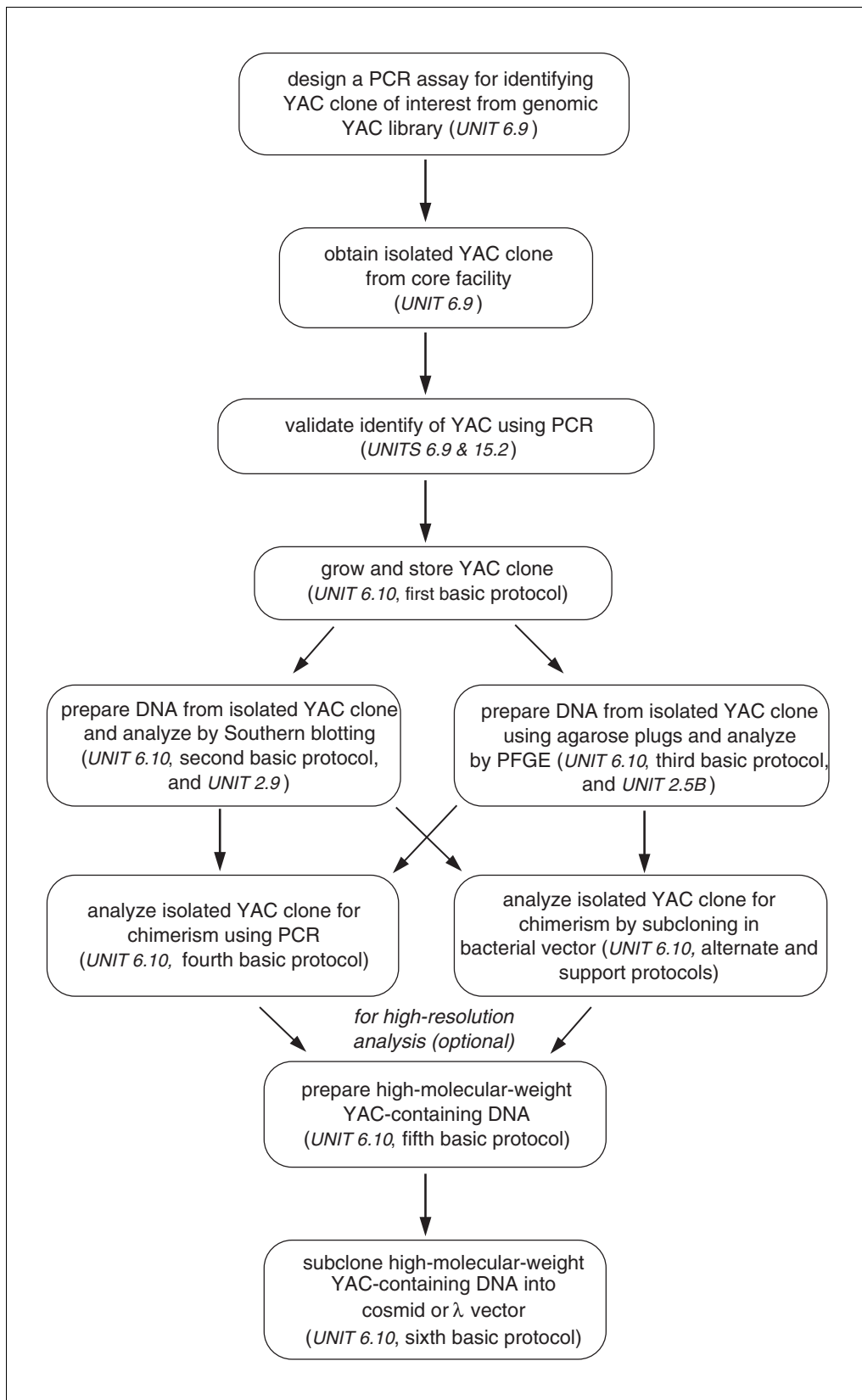


Figure 6.9.1 Flow chart showing protocols used to obtain and analyze YAC clones.

from outside laboratories, it is a good idea to inquire in advance about the protocols preferred by the specific core facility that will be performing the screening.

In general, any highly specific, sensitive, and robust PCR assay is suitable for screening a YAC library (see Chapter 15). Typically, two 18- to 30-mer oligonucleotide primers for use in amplifying a single-copy 75- to 750-bp product are satisfactory. Such primers define a landmark for genome mapping called an STS (sequence-tagged site; see introduction to Chapter 7 and Olson et al., 1989). When designing a PCR assay from scratch, it is useful to consider the following:

Fragment size

The STS should be 75 to 750 bp in length. Fragments in this range are most efficiently amplified by PCR and are easily detected by either polyacrylamide (UNIT 2.7) or standard agarose (UNIT 2.5A) gel electrophoresis.

Primer length

Each primer should ideally be 18 to 30 nucleotides long, be composed of 50% to 55% G + C, and be contained within a single-copy human-genomic-DNA segment. This ensures efficient priming and decreases the probability of false priming, enhancing the sensitivity and specificity of the assay. This also permits the amplified fragment to be used as a hybridization probe in the final hybridization-dependent steps of library screening (see below). If it is not possible to amplify a single-copy fragment, then some other single-copy probe (e.g., a synthetic oligonucleotide 30 nucleotides long) should also be prepared. Oligonucleotide design strategies are discussed further in UNITS 2.11 & 15.1.

Primer affinity

Primers should show little affinity for self-annealing or for annealing with each other. This prevents the production of small, template-independent PCR products that compete for primers in the reaction. A number of academic and commercial DOS-based and Macintosh software programs permit rapid selection of non-self-annealing primers from within a known DNA sequence (e.g., Oligo 4.0, National Biosciences; Primer, S. Lincoln and M. Daly, Whitehead Institute for Biomedical Research, Cambridge, Mass., and OSP, Hillier and Green, 1991; see UNIT 7.7). Although the use of these programs cannot remove all the uncertainty associated with designing a new PCR

assay, it does help eliminate some of the most trivial causes of assay failure.

ANALYZING INDIVIDUAL YAC CLONES

Once library screening has been successfully completed and the isolated YAC clone has been furnished to the investigator, attention should be directed to analyzing its structure. Initial studies should focus on determining whether the genomic insert is chimeric, checking for evidence of rearrangement within the insert, and verifying that the YAC is propagated in stable fashion in the yeast cell (see below). Simply analyzing several isolates of the same YAC in parallel may provide a means of recognizing instability, as each isolate serves as a control for the others. The following sections give an overview of strategies for analyzing YAC clones; specific protocols are given in UNIT 6.10.

Chimerism of the YAC Insert

A consistent problem in YAC cloning is chimerism of the YAC insert—i.e., the insert is composed of two or more separate genomic fragments joined in a single YAC. The mechanism(s) giving rise to chimeric YAC clones are currently not fully understood (Green et al., 1991). In most existing total genomic YAC libraries, chimeric clones represent from 5% to 50% of the total clones. Preliminary data suggest that targeted, chromosome-specific libraries may contain only 5% to 15% chimeric clones. Although future generations of YAC libraries are likely to contain lower frequencies of chimeric clones, chimerism will probably remain a significant problem requiring assessment for every new YAC clone being analyzed.

The most reliable way to determine if a YAC insert is chimeric is to isolate a small fragment from each end of the insert and determine its chromosome of origin and whether it shares sequences with overlapping YACs derived from the same chromosomal region. Many approaches have been suggested for isolating such YAC genomic insert end fragments, all of them relying upon the fact that end fragments are marked by their adjacent YAC vector sequences. Thus, it is possible to determine whether a YAC insert is chimeric by preparing probes from the two YAC vector arms and using these to demonstrate that both ends of the YAC map to the same general chromosomal region. This is generally done using hybridization or PCR analysis of a somatic hybrid cell line containing the appropriate human chromo-

some (or preferably a fragment thereof) as its sole human DNA. The appropriate end-fragment probes may be produced in several ways, two of which are presented in *UNIT 6.10*.

The most rapid and versatile approaches to producing end-fragment probes use PCR amplification (Riley et al., 1990; Green, 1992). The template is a restriction fragment produced by cutting the YAC DNA at sites near the two ends of the genomic insert. The YAC DNA is digested with a frequently cutting restriction endonuclease to produce a collection of small restriction fragments. One of the fragments contains the distal portion of the YAC insert still associated with a portion of the left vector arm, while another contains the other end of the insert associated with a portion of the right vector arm (the arm of the YAC vector containing the yeast centromere is arbitrarily designated the left vector arm and the arm containing the *ura3* selection marker is arbitrarily designated the right vector arm). These fragments are prepared for PCR amplification by ligation of a synthetic double-stranded DNA tag to both ends. This tag contains a 29-nucleotide “bubble” of noncomplementary sequence flanked by two 12-nucleotide complementary sequences (Fig. 6.10.2). The two YAC-insert end fragments can then be selectively amplified using one PCR primer derived from the YAC vector and one primer with the sequence of the noncomplementary portion of the bubble.

These methods are generally favored because of their speed, but they depend on the fortuitous placement of restriction sites close enough to the ends of the genomic insert that a fragment suitably sized for PCR amplification can be generated. Moreover, if highly repetitive sequences are present at the distal portions of the insert, the PCR method may fail to generate useful information.

A reliable but more time-consuming method of generating probes for end-fragment analysis is conventional subcloning of larger YAC-derived restriction fragments into plasmid or λ vectors (Bronson et al., 1991). Subcloning an end fragment several kilobases in size is time-consuming, but reliably assures identification of nonrepeated sequences for use as probes. The subcloning protocol given in *UNIT 6.10* involves double-digesting the YAC DNA to enrich for end fragments in the course of subcloning the insert into a pUC19-based vector. One of two specific enzymes that cut rarely in yeast and human genomic DNA, *ClaI* or *SalI*, is included in the double digestion mixture. A *ClaI* recognition sequence lies in the left arm

of the YAC vector, while a *SalI* recognition site lies in the right arm (Fig. 6.10.1). When one of these rarely cutting restriction enzymes is used together with a frequently cutting enzyme, doubly-digested fragments constitute only a small fraction of the total digested product. Ligation to a doubly-digested plasmid vector eliminates all of the single-digested fragments, resulting in a substantial enrichment for the YAC end fragment.

Internal Rearrangement or Instability of the YAC Insert

Internal rearrangement of a YAC insert is more difficult to identify than chimerism, and may become apparent only after high-resolution analysis of the clone. Existing reports of internal rearrangement of YAC inserts are anecdotal, infrequent, and usually identify only rather large-scale changes. It is likely, however, that subtle rearrangements will be recognized as more clones are analyzed. Nevertheless, the data suggest that important rearrangements will remain relatively infrequent and will not impede most YAC cloning efforts.

Although YACs are usually stable in culture, deletion or other rearrangements of the insert may occur months after the initial isolation of a clone. Thus, it is wise to verify the size of a YAC following prolonged passage in culture or after it has been thawed from a frozen stock. Several different colonies of the same YAC strain should be analyzed in parallel, using the protocols in *UNIT 6.10*, to confirm that the artificial chromosome is the same size in each of the isolates.

Because cytosine methylation, which is quite frequent in the DNA of higher eukaryotic species, does not occur in yeast (Proffitt et al., 1984), it is not possible to perform direct structural comparisons of the YAC inserts and the corresponding genomic DNA isolated from higher eukaryotic cells using infrequently cutting restriction enzymes to create large-scale restriction maps. Consequently, direct structural comparisons must be carried out using methylation-insensitive restriction enzymes and frequently spaced probes.

Evidence of internal rearrangement within a YAC clone can be obtained by preparing chromosomes from the clone (*UNIT 6.10*) and analyzing them by pulsed-field gel electrophoresis (PFGE; *UNIT 2.5B*). The CHEF gel system (Vollrath and Davis, 1987) is particularly useful in that it permits excellent resolution in the size range most common for individual YAC clones. Following electrophoresis, the artificial chro-

mosome can be visualized using ethidium bromide staining as an extra chromosome not present in the host yeast strain. Occasionally, the YAC is not immediately recognizable because it comigrates with one of the endogenous yeast chromosomes. As described in the third basic protocol in *UNIT 6.10*, the PFGE gel should be Southern blotted and analyzed by successive hybridization with probes specific for the locus used in the library screening and for one or both of the YAC vector arms [e.g., for pYAC4, the 351-bp *Clal/BamHI* fragment of pBR322 (left YAC arm) and the 276-bp *BamHI/SalI* fragment of pBR322 (right YAC arm)]. These blots should show hybridization to a single chromosome of the same size in all isolates from the same YAC strain as well as no hybridization to the AB1380 host strain.

CONSTRUCTION AND ANALYSIS OF A YAC-INSERT SUBLIBRARY

Although the large genomic DNA fragments provided by the YAC cloning system are easy to manipulate, it is often convenient to reduce a YAC to smaller fragments by subcloning it into a cosmid or λ vector. In particular, such smaller fragments are more amenable to high-resolution analysis; this is important because information concerning the specific content of the YAC insert is typically limited, and often the only internal probe that is available is the one used for YAC library screening. Protocols for preparing YAC insert DNA and constructing a cosmid sublibrary are provided in *UNIT 6.10*.

Two general strategies are available for preparing YAC insert DNA in order to create a saturating collection of subclones. The more elegant strategy is to purify the artificial chromosome itself by preparative CHEF gel electrophoresis (*UNIT 2.5B*). This permits isolation and analysis of the resulting recombinant clones without further selection, assuming that only a small amount of contaminating yeast DNA is present in the purified YAC, and that essentially all subclones isolated are derived from the human YAC insert. In practice, however, it is difficult to recover sufficient quantities of purified YAC DNA to permit construction of a cosmid or λ library. An alternate approach is to prepare a library from the total DNA of the YAC-carrying yeast strain. YAC-specific subclones must then be selected by hybridization. An initial round of screening is usually performed with total human genomic DNA (rich in repetitive sequences) as the probe.

This detects subclones that contain human repetitive elements and eliminates subclones consisting of yeast DNA. Additional analysis is performed to identify overlapping sequences and thereby establish an approximate map of the original YAC insert. Ultimately, one or more rounds of chromosome walking may be required to fill in gaps between contiguous groups of subclones.

Literature Cited

- Abidi, F.E., Wada, M., Little, R.D., and Schlessinger, D. 1990. Yeast artificial chromosomes containing human Xq24-Xq28 DNA: Library construction and representation of probe sequences. *Genomics* 7:363-376.
- Bronson, S.K., Pei, J., Taillon-Miller, P., Chorney, M.J., Geraghty, D.E., and Chaplin, D.D. 1991. Isolation and characterization of yeast artificial chromosome clones linking the HLA-B and HLA-C loci. *Proc. Natl. Acad. Sci. U.S.A.* 88:1671-1675.
- Burke, D.T., Carle, G.F., and Olson, M.V. 1987. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 236:806-812.
- Burke, D.T., Rossi, J.M., Leung, J., Koos, D.S., and Tilghman, S.M. 1991. A mouse genomic library of yeast artificial chromosome clones. *Mamm. Genome* 1:65-69.
- Green, E.D. and Olson, M.V. 1990. Systematic screening of yeast artificial chromosome libraries by use of the polymerase chain reaction. *Proc. Natl. Acad. Sci. U.S.A.* 87:1213-1217.
- Green, E.D., Riethman, H.C., Dutchik, J.E., and Olson, M.V. 1991. Detection and characterization of chimeric yeast artificial-chromosome clones. *Genomics* 11:658-669.
- Green, E.D. 1992. Physical mapping of human chromosomes: Generation of chromosome-specific sequence-tagged sites (STS). *Methods Mol. Genet.* In press.
- Hillier, L. and Green, P. 1991. A computer program for choosing PCR and DNA sequencing primers. *PCR Meth. Appl.* 1:124-128.
- Olson, M., Hood, L., Cantor, C., and Botstein, D. 1989. A common language for physical mapping of the human genome. *Science* 245:1434-1435.
- Proffitt, J.H., Davie, J.R., Swinton, D., and Hattman, S. 1984. 5-Methylcytosine is not detectable in *Saccharomyces cerevisiae* DNA. *Mol. Cell Biol.* 4:985-988.
- Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J.C., and Markham, A.F. 1990. A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucl. Acids Res.* 18:2887-2890.

Rossi, J.M., Burke, D.T., Leung, J.C., Koos, D.S., Chen, H., and Tilghman, S.M. 1992. Genome analysis using a yeast artificial chromosome library with mouse DNA inserts. *Proc. Natl. Acad. Sci. U.S.A.* 89:2456-2460.

Vollrath, D. and Davis, R.W. 1987. Resolution of DNA molecules greater than 5 megabases by contour-clamped homogeneous electric fields. *Nucl. Acids Res.* 15:7865-7876.

Key Reference

Burke, et al., 1987. See above.

Initial description of the YAC cloning system, covering general features of library construction.

Contributed by David D. Chaplin and
Bernard H. Brownstein
Howard Hughes Medical Institute and
Washington University School of Medicine
St. Louis, Missouri

Analysis of Isolated YAC Clones

The preceding unit gives an overview of methods involved in screening a YAC library to isolate a particular clone of interest (UNIT 6.9), with the sequence of methods illustrated in a flow chart (Fig. 6.9.1). This unit provides a series of protocols describing the analysis and manipulation of an isolated YAC clone. The procedures are based upon the use of the YAC vector pYAC4.

Once an isolated YAC clone has been obtained from a core laboratory (UNIT 6.9), the clone can be analyzed as described herein. As depicted in Figure 6.9.1, methods for analysis involve growing and storing YAC-containing yeast strains and purifying YAC DNA in a form suitable for assessing the size of the artificial chromosome and for conventional Southern blotting. Preparation of yeast chromosomes in agarose plugs for subsequent analysis by pulsed-field gel electrophoresis is also described. Additional protocols are provided for recovering DNA fragments from the ends of a YAC genomic insert to be used as probes for detecting chimerism and for chromosome walking. Finally, preparation of high-molecular-weight YAC DNA is described and a general method for subcloning YAC inserts into cosmid or λ vectors for higher-resolution analysis is provided.

NOTE: All solutions, media, glassware, and plasticware coming into contact with yeast or bacterial cells must be sterile, and sterile techniques should be followed throughout.

BASIC PROTOCOL

PROPAGATION AND STORAGE OF YAC-CONTAINING YEAST STRAINS

YACs prepared using the pYAC4 vector (Figs. 6.10.1 and 13.4.6; pYAC4 contains an *EcoRI* site within the *SUP4* gene in addition to the *SnaBI* site found in pYAC3, but is otherwise identical to pYAC3, carrying selectable markers *TRP1* and *URA3*) and the *S. cerevisiae* host strain AB1380 (*trp1⁻*, *ura3⁻*, *ade2-1*) are grown on AHC plates. They can be stored short-term on AHC plates or stored long-term (after growth in YPD medium) in YPD containing glycerol at -80°C .

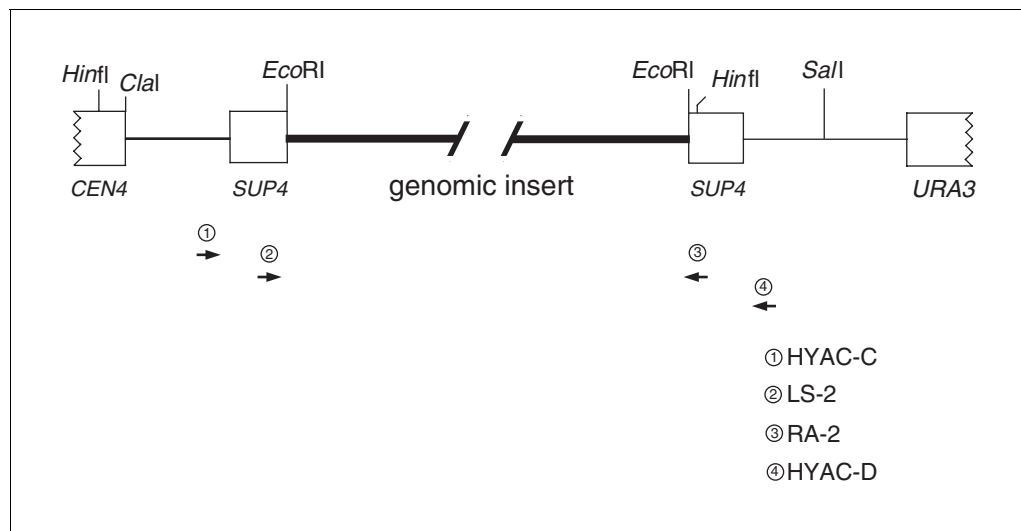


Figure 6.10.1 Structure of a representative pYAC4 clone at the vector/insert junction. Open boxes represent portions of the pYAC4 vector derived from yeast sequences (*CEN4*, the two halves of the *SUP4* element, and *URA3*). Thin lines represent sequences derived from pBR322 and the bold line represents the YAC genomic insert fragment. Sites of annealing of the HYAC-C, LS-2, RA-2 and HYAC-D oligonucleotides are indicated by arrows 1, 2, 3, and 4, respectively. The *EcoRI* cloning site, the *ClaI* and *SalI* sites (used for the end-fragment subcloning alternate protocol), and the *HinfI* sites that are utilized in the bubble linker end-fragment isolation protocol are indicated.

Analysis of Isolated YAC Clones

6.10.1

Contributed by David D. Chaplin and Bernard H. Brownstein

Current Protocols in Molecular Biology (1992) 6.10.1-6.10.19

Copyright © 2000 by John Wiley & Sons, Inc.

Materials

S. cerevisiae strain AB1380 containing pYAC4 with insert (from core facility; UNIT 6.9)

AHC plates (*ura*⁻, *trp*⁻)

YPD medium (UNIT 13.1)

80% (v/v) glycerol in YPD medium

30°C orbital shaking incubator (e.g., New Brunswick Scientific #G-24)

Cryovials

Additional reagents for preparation of yeast media (UNIT 13.1) and growth and manipulation of yeast (UNIT 13.2)

1. Streak strain AB1380 containing pYAC4 with insert onto AHC plates.

AHC medium selects for the presence of both arms of the YAC vector and thereby favors high stability of the YAC through successive passages.

2. Invert plate and incubate at 30°C until colonies are 1 to 3 mm in size.

*The AB1380 strain carries the *ade2-1* ochre mutation, a block in the purine biosynthetic pathway that leads to accumulation of red-hued intermediates. Because a genomic insert in the YAC interrupts the *SUP4* (ochre) gene in the YAC vector, colonies will have a red pigmentation.*

- 3a. *For short-term storage:* Seal plates with Parafilm and store at 4°C for 4 to 6 weeks.

- 3b. *For long-term storage:* Inoculate an individual colony into 3.2 ml YPD medium and shake overnight at 30°C. Add 1 ml of 80% glycerol in YPD medium, mix thoroughly, and transfer in 0.2- to 1.0-ml aliquots to cryovials. Store at -80°C.

YPD is a nonselective medium used to favor rapid growth and high cell viability.

Strains stored in this fashion are stable for ≥5 years. Before strains are used in an experiment, they should first be grown on selective medium (e.g., AHC plates) to avoid recovery of a contaminant clone or one that has lost its YAC.

PREPARATION OF YAC-CONTAINING DNA FROM YEAST CLONES FOR ANALYSIS BY SOUTHERN BLOTTING

Procedures used by core laboratories for isolating an individual clone from a YAC library ensure that the purified YAC supports amplification of an appropriately sized PCR product using the screening primer pair. However, it is best to confirm the identity of the clone by hybridization analysis. Various methods can be used to prepare DNA suitable for Southern blot analysis using frequently cutting restriction enzymes. This protocol yields substantial quantities of DNA in the size range of 50 to 200 kb; it involves growing and lysing a single red colony containing pYAC4 with the insert DNA, then obtaining the DNA from the supernatant after centrifugation and analyzing by Southern blotting. Yeast chromosomes prepared in agarose plugs or very-high-molecular-weight DNA prepared in solution (third and fifth basic protocols) may also be used.

Materials

Single colony of *S. cerevisiae* AB1380 containing pYAC4 with insert (first basic protocol)

AHC medium (*ura*⁻, *trp*⁻)

SCE buffer

SCEM buffer

50 mM Tris·Cl (pH 7.6)/20 mM EDTA (Tris/EDTA lysis buffer)

BASIC PROTOCOL

Screening of Recombinant DNA Libraries

6.10.2

10% (w/v) sodium dodecyl sulfate (SDS)
5 M potassium acetate, pH 4.8, ice-cold (*UNIT 1.6*)
95% ethanol, room temperature
TE buffer, pH 8.0 (*APPENDIX 2*)
1 mg/ml DNase-free RNase A (*UNIT 3.13*)
Isopropanol, room temperature
5 M NaCl
Total genomic DNA of the species or individual from which the library was made (e.g., *UNITS 2.2, 2.3 & 5.3*)
Appropriate single-copy probe designed to hybridize with the YAC insert (see *UNITS 2.9 & 6.9*)

Orbital shaker (e.g., New Brunswick Scientific #G-24)
50-ml conical plastic centrifuge tubes
Beckman JS-4.2 rotor or equivalent

Additional reagents and equipment for digestion of DNA with restriction endonucleases (*UNIT 3.1*), Southern blotting and hybridization (*UNIT 2.9*), and pulsed-field gel electrophoresis (*UNIT 2.5B*)

Culture and lyse cells from YAC clone

1. Inoculate a single red colony of a YAC-containing clone into 20 ml AHC medium in a 250-ml Erlenmeyer flask. Shake 24 hr at 250 rpm, 30°C, on an orbital shaker.

The culture should begin to turn pink. If not, continue incubation an additional 24 hr. If culture is still not pink, discard and start over with a new red colony.

Orbital shakers are preferred because they give much better aeration.

2. Inoculate 1 ml of culture from step 1 into 100 ml AHC medium in a 1-liter Erlenmeyer flask. Shake 24 hr at 250 rpm, 30°C.
3. Transfer culture to 50-ml plastic conical centrifuge tubes. Centrifuge 5 min at 2000 × g (2800 rpm in Beckman JS-4.2 rotor), 4°C.
4. Discard supernatants and resuspend cell pellets in a total of 5 ml SCE buffer. Pool into a single tube.
5. Add 1 ml SCEM buffer. Mix gently 1 to 2 hr at 100 rpm, 37°C, on an orbital shaker.
SCEM buffer contains lyticase, which will digest the cell wall.
6. Centrifuge 5 min at 2000 × g, 4°C. Discard supernatant and resuspend cell pellet in 5 ml Tris/EDTA lysis buffer.
7. Add 0.5 ml of 10% SDS and invert several times to mix. Incubate 20 min at 65°C.

Isolate nucleic acids

8. Add 2 ml of ice-cold 5 M potassium acetate, pH 4.8, and invert to mix. Keep 60 min on ice.
9. Centrifuge 10 min at 2000 × g, room temperature. Carefully pour nucleic acid-containing supernatant into a new tube. Add 2 vol room-temperature 95% ethanol and invert to mix.
10. Centrifuge 5 min at 2000 × g, room temperature. Discard supernatant and air-dry nucleic acid pellet 10 to 15 min. Add 3 ml TE buffer, pH 8.0, and dissolve overnight at 37°C.

Recover and analyze DNA

11. Add 0.1 ml of 1 mg/ml DNase-free RNase A and incubate 1 hr at 37°C.
12. Add 6 ml room-temperature isopropanol with swirling, then invert to mix.
13. Spool DNA using a capillary pipet and dissolve in 0.5 ml TE buffer, pH 8.0. Add 50 μ l of 5 M NaCl and 2 ml of room-temperature 95% ethanol. Mix by inverting.
14. Spool DNA again and dissolve in 0.5 ml TE buffer. Store at 4°C.

A yield of 1 to 1.5 μ g DNA/ 10^8 yeast cells can be expected.

15. Analyze 2- μ g aliquots of YAC DNA and 15- μ g aliquots of total genomic DNA from the species or individual from which the YAC library was made by digesting with several frequently cutting restriction enzymes. Proceed with Southern blotting and hybridization using a single-copy probe.

The product amplified by PCR screening (see UNIT 6.9) may be used as probe. Because the YAC donor may exhibit a restriction-fragment-length polymorphism for this probe, two restriction fragments may be observed in the donor DNA. One of these fragments should be identified in the isolated YAC DNA.

16. Once the YAC clone has been verified by Southern blotting, determine its size and obtain a preliminary assessment of its stability by preparing chromosomes in agarose plugs (third basic protocol) and analyzing by pulsed-field gel electrophoresis.

PREPARATION OF YEAST CHROMOSOMES IN AGAROSE PLUGS FOR PULSED-FIELD GEL ELECTROPHORESIS

BASIC PROTOCOL

In order to assess size, stability, and possible rearrangements within YACs, and to identify overlapping YACs, it is useful to isolate the YACs by embedding them in agarose plugs for subsequent analysis by pulsed-field gel electrophoresis (PFGE). Most methods of pulsed-field gel electrophoresis can be used (UNIT 2.5B); the CHEF (contour-clamped homogeneous electric-field electrophoresis) gel system is particularly suitable in that it reliably permits excellent resolution in the size range most common for YACs.

Materials

AHC medium (*ura*⁻, *trp*⁻)

Single colony of *S. cerevisiae* containing pYAC4 with insert (first basic protocol)

0.05 M EDTA, pH 8.0 (APPENDIX 2)

SEM buffer

10 mg/ml Lyticase (Sigma #L-8137 or ICN Biomedicals #190123)

2% InCert or SeaPlaque agarose (FMC Bioproducts), dissolved in SEM buffer and equilibrated to 37°C

SEMT buffer

Lithium lysis solution

20% (v/v) NDS solution

0.5 \times TBE (APPENDIX 2) or GTBE buffer (UNIT 2.5B)

30°C rotary platform shaking incubator

Beckman JS-4.2 rotor or equivalent

Gel sample molds (e.g., CHEF gel molds, Bio-Rad #1703622)

60-mm tissue culture plate

Additional reagents and equipment for pulsed-field gel electrophoresis (UNIT 2.5B)

Screening of Recombinant DNA Libraries

6.10.4

Prepare and lyse YAC clone

1. Inoculate 25 ml AHC medium with a single red colony of a YAC-containing clone. Shake 48 to 60 hr at 250 rpm, 30°C.

The culture should be pink. If it is not, discard and start over with a new red colony.

To assess the stability of an individual YAC and to facilitate distinguishing of the artificial chromosome from the native yeast chromosomes, it is useful to analyze 4 or 5 individual colonies from the same YAC strain as well as a colony of the untransformed yeast host.

2. Centrifuge 10 min at $2000 \times g$ (2800 rpm in a Beckman JS-4.2 rotor), 4°C. Discard supernatant and resuspend cell pellet in 10 ml of 0.05 M EDTA, pH 8.0.
3. Centrifuge 10 min at $2000 \times g$, 4°C. Remove all liquid from pellet and resuspend in 150 μ l SEM buffer.

Prepare agarose molds

4. Warm YAC sample to 37°C and add 25 μ l Lyticase. Add 250 μ l of 2% InCert or SeaPlaque agarose that has been melted in SEM buffer and equilibrated to 37°C.
5. Mix quickly and pour into CHEF gel sample molds. Chill 10 min at 4°C. Transfer solidified plugs to a 60-mm tissue culture plate.
6. Cover each plug with 4 ml SEMT buffer. Incubate 2 hr with gentle shaking at 37°C.
7. With a pipet, remove SEMT buffer and replace with 4 ml lithium lysis solution. Incubate 1 hr with gentle shaking at 37°C.
8. Remove and replace lithium lysis solution two or three times, shaking ≥ 1 hr each time. Shake the last change overnight.
9. Remove lithium lysis solution, replace with 4 ml of 20% NDS solution, and shake 2 hr at room temperature. Repeat once.

Electrophorese samples in individual agarose plugs

10. Cut into plugs of suitable size to fit into wells of a pulsed-field gel.

Store plugs individually in 20% NDS solution at 4°C. Plugs prepared and stored in this manner are usually stable for 4 to 8 weeks.

11. Soak each plug 30 min in 1 ml of 0.5 \times TBE or GTBE buffer. Change three times.
12. Analyze by pulsed-field gel electrophoresis.

Following electrophoresis, the artificial chromosome can be visualized in an ethidium bromide-stained gel as an extra chromosome not present in the host yeast strain. If desired, Southern blot hybridization (UNIT 2.9) with appropriate probes can be carried out.

END-FRAGMENT ANALYSIS USING PCR AMPLIFICATION

BASIC PROTOCOL

This protocol provides a means for recovering end fragments from the YAC insert using PCR amplification of end fragments. Digestion of YAC-containing DNA with a frequently cutting restriction enzyme produces a collection of small fragments: among these, one contains the distal portion of the YAC insert associated with part of the left vector arm, and another contains the other end of the insert associated with part of the right vector arm. Fragments encoding these vector sequences are prepared for PCR amplification by ligation of a double-stranded DNA tag containing a “bubble” of noncomplementary sequence flanked by short complementary sequences (Fig. 6.10.2). Selective amplification of these two end-fragment sequences is achieved using one PCR primer derived from the YAC vector (HYAC-C or LS-2 for the left arm or HYAC-D or RY-2 for the right arm) and one primer containing the sequence of the noncomplementary portion of the bubble (the 224 primer template, created by extension from the YAC-vector-specific primer; Fig. 6.10.3). Occasionally, nonspecific DNA fragments are amplified from the bubble PCR reaction. If this occurs, specificity may be restored using a hemi-nesting strategy. A small aliquot of the product of the initial PCR reaction (containing a mixture of the specific and nonspecific amplified fragments) is reamplified in a second round of PCR using an internal sequence from the vector arm as one of the primers. Because this sequence is not present in the nonspecific fragments, only the specific fragment will be amplified.

Materials

“Bubble-top” and “bubble-bottom” oligonucleotide primers (Fig. 6.10.2)

YAC-containing DNA (second basic protocol)

*Rsa*I and *Hinf*I restriction endonucleases and appropriate buffers (UNIT 3.1)

10× T4 DNA ligase buffer and 1 U/μl T4 DNA ligase (UNITS 3.4 & 3.14)

PCR reaction mix

PCR amplification primers HYAC-C, HYAC-D, 224, and RA-2, 4 μM each (Fig. 6.10.2)

Thermal cycling apparatus

65° and 68°C water baths

Additional reagents and equipment for phosphorylating synthetic oligonucleotides (UNIT 3.10), restriction endonuclease digestion (UNIT 3.1), PCR (UNIT 15.1), nondenaturing PAGE (UNIT 2.7), preparing radiolabeled oligonucleotide probes (UNITS 3.10, 4.6 & 15.2), and blunt-end ligation (UNIT 3.16)

Prepare bubble oligonucleotide tags

1. Phosphorylate the bubble-top oligonucleotide.

This step can usually be eliminated, but may modestly increase efficiency.

2. Adjust bubble-top and bubble-bottom oligonucleotide concentrations to 4 nmol/ml with water. Mix together 1 nmol of each, then anneal by heating 15 min at 68°C in a water bath, followed by slow cooling to room temperature over 30 to 60 min.

Digest YAC DNA and ligate to bubble oligonucleotides

3. Digest 2.5-μg aliquots of purified YAC-containing DNA to completion with *Rsa*I or *Hinf*I in 20 μl final volume, 37°C.

*Digestion of separate samples with *Rsa*I and *Hinf*I increases the chance of obtaining an end fragment of a size suitable for PCR amplification (<1.5 kb) and containing a substantial portion of insert sequence (>75 bp).*

4. Heat samples 15 min at 65°C to inactivate the restriction enzymes.



Figure 6.10.2 Oligonucleotides for amplification and sequencing of YAC insert end-fragments. **(A)** Annealing of the 53-mer universal “bubble-bottom” oligonucleotide to the 53-mer *RsaI* bubble oligonucleotide yields a blunt-ended DNA duplex in which 12-bp complementary sequences flank a 29-nucleotide “bubble” of noncomplementary sequence. This bubble linker can be ligated to any blunt-ended fragment (e.g., one generated by digestion with *RsaI*). The 224 primer does not anneal to either strand of the bubble, but is fully complementary to any DNA strand that is generated during PCR using the universal bubble-bottom strand as a template (see Fig. 6.10.3). **(B)** Annealing of the 53-mer universal bubble-bottom oligonucleotide to the 56-mer *HinfI* bubble-top oligonucleotide yields a DNA duplex with one blunt end and one cohesive end with the degenerate *HinfI* site. A mixture of all four nucleotides at a specific position is indicated by (N). **(C)** The HYAC-C, HYAC-D, RA-2, and LS-2 primers anneal to sequences in the pYAC4 vector (see Fig. 6.10.1). The bubble sequencing primer anneals to the *RsaI* and *HinfI* bubble-top sequences near their 5′ ends, permitting DNA sequencing from the bubble linker back into the YAC insert end-fragment.

5. Prepare the following ligation mix (50 μl total):

- 2 μl (250 ng) digested DNA
- 1 μl (2 pmol) annealed bubble oligonucleotides (from step 2)
- 5 μl 10× ligase buffer
- 2 μl (2 U) T4 DNA ligase
- 40 μl H₂O.

Incubate 2 hr at 37°C or overnight at room temperature.

The blunt-ended bubble composed of the universal bubble-bottom oligonucleotide and the RsaI bubble-top oligonucleotide should be used with RsaI-digested YAC DNA. Likewise, the HinfI cohesive bubble composed of the universal bubble-bottom oligo-

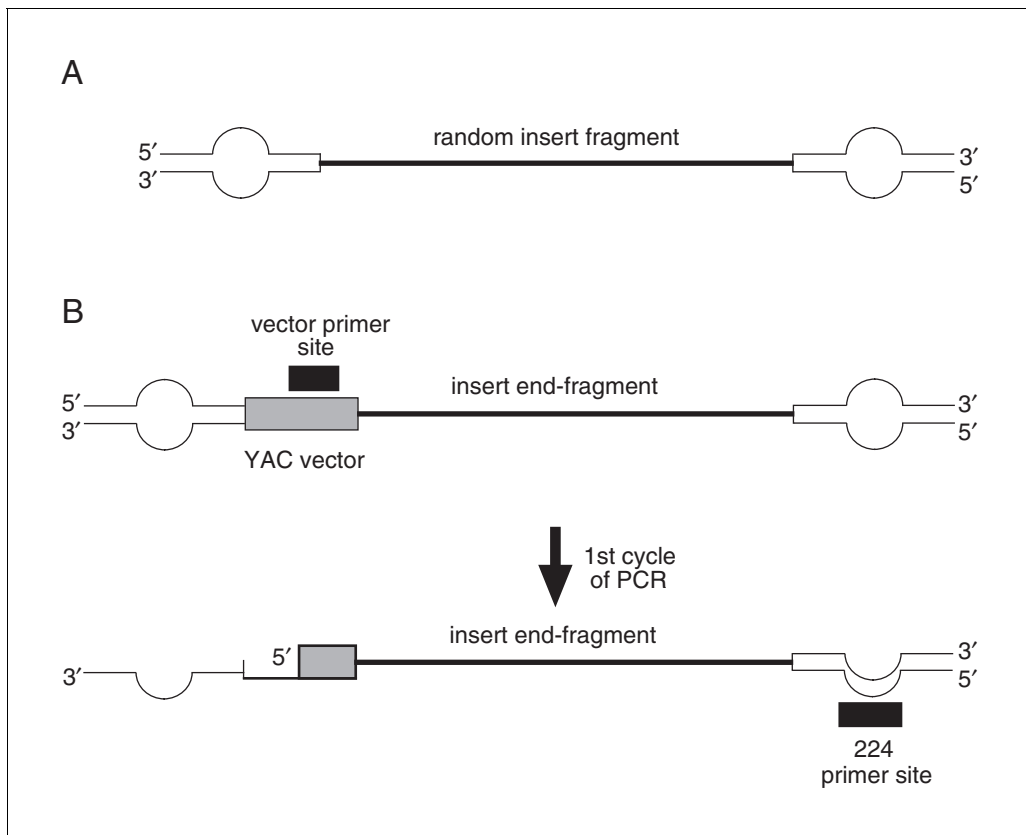


Figure 6.10.3 Selective PCR amplification from the YAC insert end-fragment. **(A)** Result of ligation of the bubble linker to a random fragment from the internal portion of the YAC insert. Because this fragment is not derived from the end of the YAC genomic insert, it contains no sequences from the YAC vector and has no site for annealing any of the HYAC-C, HYAC-D, RA-2, or LS-2 primers or for annealing of the 224 primer. Consequently, no fragment is amplified by PCR. **(B)** Result of ligation of the bubble linker to a fragment derived from the end of the YAC genomic insert and containing its associated YAC vector sequences. During the first cycle of PCR, extension from the YAC vector priming site produces sequences complementary to the universal bubble-bottom primer. This extended fragment provides a template for annealing of the 224 primer, thus permitting successful amplification of the insert end-fragment.

*nucleotide and the *Hin*I bubble-top oligonucleotide should be used with *Hin*I-digested YAC DNA.*

6. Add 200 μ l water to bring the DNA concentration to 1 ng/ μ l final.

Amplify fragments containing YAC-insert end sequences

7. Prepare the following PCR on ice (10 μ l total):
 - 8 μ l PCR reaction mix
 - 1 μ l (2 μ M each) PCR primer pair mix
 - 1 μ l (1 ng) digested, bubble-ligated YAC DNA.

Carry out 35 cycles of amplification as follows: 1 min at 92°C, 2 min at 65°C, and 2 min at 72°C.

To amplify left end of YAC insert, use primer pair mix made from equal amounts of primers 224 and HYAC-C (Fig. 6.10.1). To amplify right end of YAC insert, use primer pair mix made from equal amounts of primers 224 and RA-2 (specific for the SUP4 region of pYAC4; Fig. 6.10.1).

These parameters were optimized using the Perkin-Elmer TC1 thermal cycler. If another instrument is used, some adjustment of parameters may be required.

8. Analyze a 1- μ l aliquot of PCR product on a 5% polyacrylamide gel. A single, clearly visible amplified fragment should be observed after staining the gel with ethidium bromide.

Produce end fragments using hemi-nested amplification

To amplify the left end (either RsaI- or HinfI-digested DNA):

- 9a. Amplify 1 μ l digested, bubble-ligated YAC DNA (from step 6) with primers 224 and HYAC-C, using 20 cycles of 1 min at 92°C, 2 min at 62°C, and 2 min at 72°C.
- 10a. Dilute the amplification product 1:100 with water and add 1 μ l to a new PCR reaction containing primers 224 and LS-2 (specific for the *SUP4* region of pYAC4; see Figs. 6.10.1 and 13.4.6). Carry out 30 cycles of 1 min at 92°C, 2 min at 65°C, and 2 min at 72°C.

To amplify the right end using RsaI-digested DNA:

- 9b. Amplify 1 μ l digested, bubble-ligated YAC DNA (from step 6) with primers 224 and HYAC-D, using 20 cycles of 1 min at 92°C, 2 min at 62°C, and 2 min at 72°C.
- 10b. Dilute amplification product 1:100 with water and add 1 μ l to a new PCR reaction (see step 7) containing primers 224 and RA-2. Carry out 30 cycles of 1 min at 92°C, 2 min at 65°C, and 2 min at 72°C.

Hemi-nesting of the right end cannot be performed with HinfI-digested DNA, because there is a HinfI site only 24 bp from the EcoRI YAC vector cloning site.

11. Analyze a 1- μ l aliquot of each final PCR reaction on a 5% polyacrylamide gel.
12. End label amplified fragments with ³²P and use as hybridization probes or for nucleotide sequencing to produce an end-specific STS. Alternatively, subclone by blunt-end ligation to a plasmid vector prior to further manipulation.

Blunt-end subcloning of DNA fragments that have been amplified by PCR must be preceded by "polishing" of the ragged PCR ends with S1 nuclease or T4 DNA polymerase (see UNIT 15.7).

ALTERNATE PROTOCOL

END-FRAGMENT ANALYSIS BY SUBCLONING INTO A BACTERIAL PLASMID VECTOR

This method (an alternative to the previous protocol for recovering end fragments from the YAC insert) uses the strategy of double digesting the YAC-containing DNA to enhance the efficiency of subcloning into a pUC19-based vector; the *SUP4* element containing the YAC *EcoRI* cloning site is located within a portion of the YAC vector derived from pBR322 (Figs. 1.5.2 and 6.10.5). The cloning strategy enriches for end-fragment-containing subclones because the restriction endonuclease used for digestion of YAC-containing DNA is either *ClaI* (for the left arm) or *SalI* (for the right arm). Both enzymes cut rarely in human or yeast genomic DNA; therefore, when one is combined with a more frequently cutting enzyme, the resulting doubly digested fragments will represent a minor portion of the total DNA pool. Size fractionation on an agarose gel prior to subcloning affords a still further enrichment for end fragments. Note that after the first step of the protocol, all steps are performed in duplicate to identify both the right and left end fragments of the YAC insert.

Additional Materials

*Cla*I, *Sal*I, and other appropriate restriction endonucleases and digestion buffers (UNIT 3.1)

Left- and right-vector-arm probes (Fig. 6.10.4)

pUC19-ES and pUC19-HS plasmid vectors (support protocol and Fig. 6.10.5)

Transformation-competent *Rec*⁻ strain of *E. coli* (e.g., DH5; Table 1.4.5)

2× TY or LB agar plates (UNIT 1.1) containing 50 to 100 µg/ml ampicillin

Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5A), subcloning of DNA fragments (UNIT 3.16), transformation of *E. coli* (UNIT 1.8), Southern blotting and hybridization (UNIT 2.9), labeling by random-primed synthesis (UNIT 3.5), isolation and purification of DNA fragments from agarose gels (UNIT 2.6), replica plating (UNIT 1.3), and purification of plasmid DNA (UNITS 1.6 & 1.7)

Perform an appropriate double digest and analyze by hybridization

- 1a. For left arm: Digest 5-µg aliquots of YAC-containing DNA with *Cla*I and then with each possible second cloning enzyme—*Sac*I, *Kpn*I, *Sma*I, *Bam*HI, *Xba*I, and *Sph*I.

In addition to SmaI, other blunt-cutters not represented within the ClaI-EcoRI interval of the YAC vector may be tested.

- 1b. For right arm: Digest 5-µg aliquots of YAC-containing DNA with *Sal*I and then with each of the following possible second cloning enzymes: *Sac*I, *Kpn*I, *Sma*I, *Bam*HI, *Xba*I, *Sph*I, or *Hind*III.

All of the above restriction endonucleases have compatible cleavage sites within the polylinker of the modified pUC19; AccI provides a cohesive site for ClaI.

Carry out all remaining steps in parallel for the left- and right-arm probes:

2. Electrophorese doubly digested DNA on an agarose gel and transfer to a filter for Southern hybridization.
3. Prepare left- and right-vector-arm probes by PCR as described in Fig. 6.10.4.

Probes can also be obtained by digestion and fractionation of pBR322 DNA with subsequent labeling.

4. Hybridize each probe to the appropriate filter from step 2.

Left arm:	5' ATCGATAAGCTTTAATGCGGTAGT 3' (pBR322 bases 23-46)
	5' GATCCACAGGACGGGTGTGGTCGC 3' (pBR322 bases 379-356)
Right arm:	5' GATCCTCTACGCCGGACGCATCGT 3' (pBR322 bases 375-399)
	5' GTCGACGCTCTCCCTTATGCGACT 3' (pBR322 bases 656-632)

Figure 6.10.4 Generation of left- and right-vector-arm probes. The 351-bp *Cla*I-*Bam*HI and 276-bp *Bam*HI-*Sal*I fragments of pBR322, which hybridize to sequences immediately flanking the *sup4* sequences of the YAC vector, are appropriate probes for the YAC left and right vector arms. These probes can be obtained by restriction digestion and gel fractionation of pBR322 plasmid DNA or generated by PCR using 10 ng pBR322 as template for the primers illustrated here. Perform PCR using 25 cycles of 1 min at 92°C, 1 min at 50°C, and 2 min at 72°C. Extract the amplified material once with phenol and once with chloroform, then precipitate with ethanol (UNIT 2.1). Label directly by random priming (UNIT 3.5) without further purification.

5. Examine autoradiogram and choose an enzyme combination that yields a hybridizing DNA fragment in the 2- to 7-kb size range. Digest a 50- μ g aliquot of YAC-containing DNA with these two enzymes.

This should yield ~5 times more size-fractionated DNA than needed.

Isolate the DNA

6. Electrophorese doubly digested DNA on an agarose gel. Using a scalpel or razor blade, cut out the segment of gel that should contain the doubly digested DNA fragment.

To avoid missing the critical portion of the gel, it may be useful to excise adjacent gel slices containing fragments larger than and smaller than the expected size, and to process them in parallel.

7. Purify size-fractionated DNA from gel slice and resuspend in a final volume of 20 μ l TE buffer, pH 8.0.

For purifying the DNA, the best results have been obtained by using the GeneClean II kit (BIO 101, La Jolla, CA).

Subclone the end fragments

8. Ligate 20% of the purified YAC-derived insert DNA with 0.2 μ g of gel-purified, compatibly digested pUC19-HS or -ES vector DNA overnight in a total volume of 20 μ l.

Because the pUC19 plasmid from which they are derived has no homology with the portion of pBR322 detected by the ClaI/BamHI and BamHI/SalI probes (Fig. 6.10.4), these probes can be used to detect YAC-insert end-fragment-containing subclones in pUC19, and will not cross-hybridize to the pUC vector.

9. Transform the ligated DNA into a transformation-competent *Rec*⁻ host strain of *E. coli*. Plate sufficient transformation mix on 2 \times TY/ampicillin or LB/ampicillin plates to obtain ~200 colonies, a sufficiently low density that individual colonies may be recovered following hybridization. Invert plates and incubate overnight at 37°C.
10. Prepare colony-lift filters and hybridize overnight with ~1–2 \times 10⁷ cpm of appropriate ³²P-labeled left- or right-arm probes. Wash and autoradiograph.

Because of the enrichment afforded by double digestion, 1% to 4% of colonies will contain the end fragment.

11. Purify plasmid DNA from hybridizing colonies.
12. Verify the structure of the plasmid by comparing its restriction map to the data obtained during the initial analytical double digests of the YAC (steps 1 to 3).

SUPPORT PROTOCOL

DESIGN AND PREPARATION OF pUC19-ES and pUC19-HS SUBCLONING VECTOR

This protocol describes the construction of two vectors for subcloning YACs (previous basic protocol). pUC19 is modified by insertion of a “stuffer” fragment in both possible orientations (see UNIT 3.16; Fig. 6.10.5). If a double digest is performed on the resulting construct (UNIT 3.1), using *AccI* (cohesive with *ClaI*) or *SalI* and any of the other enzymes in the pUC polylinker, the presence of the stuffer makes it possible to visualize whether the vector has been fully cut. Complete double digestion is critical to the success of the end-fragment subcloning described in the previous protocol. For example, digestion with *AccI* or *SalI* will linearize the pUC19-ES vector. Subsequent digestion with *EcoRI*, *SacI*, *KpnI*, *SmaI*, *BamHI*, or *XbaI* will result in a shift in vector size from 3161 bp to 2686 bp

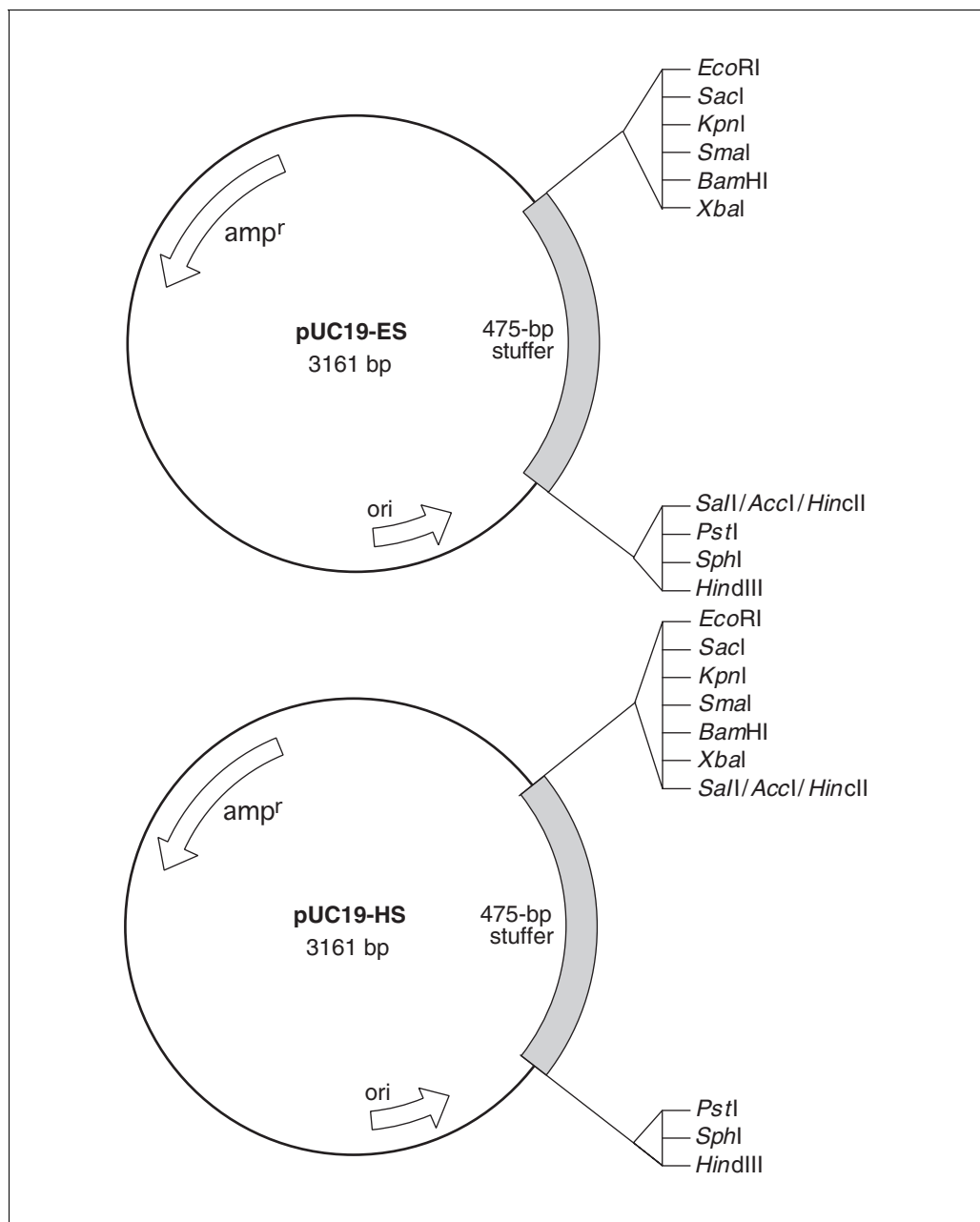


Figure 6.10.5 Structure of the pUC19-ES and pUC19-HS plasmids.

and free stuffer fragment will be generated. The doubly digested vector can then be isolated by fractionation in an agarose gel (UNIT 2.5A) and purified (UNIT 2.6).

pUC19-ES: Modify the pUC19 (see Fig. 1.5.2) vector by inserting a stuffer consisting of 475-bp *TaqI* fragment of pBR322 (positions 653-1128) into the pUC19 polylinker *AccI* (*HincII*) site. In the resulting plasmid, the *AccI* (and *SalI* and *HincII*) site adjacent to the polylinker *PstI* site is preserved, but the *AccI* site previously found next to the polylinker *XbaI* site (which would now be at the other end of the stuffer) is lost (Fig. 6.10.5).

pUC19-HS: Insert the 475-bp *TaqI* fragment stuffer described above into the same pUC19 *AccI* site but in the opposite orientation. In the resulting plasmid, the polylinker *AccI* site adjacent to the *XbaI* site is preserved, but the *AccI* site adjacent to the *PstI* site is lost (Fig. 6.10.5).

**PREPARATION OF HIGH-MOLECULAR-WEIGHT YAC-CONTAINING
YEAST DNA IN SOLUTION**

This protocol describes the purification of YAC-containing DNA of sufficiently high molecular weight to provide a source of YAC insert material for subcloning in λ or cosmid vectors. This DNA is also suitable for restriction mapping or other genetic manipulations. A cell lysate is fractionated on a sucrose gradient; the DNA-containing fraction is subsequently dialyzed, concentrated, and examined by electrophoresis through a pulsed-field gel.

Materials

- Single colony of *S. cerevisiae* containing pYAC4 with insert (first basic protocol)
- AHC medium (*ura*⁻, *trp*⁻)
- SCEM buffer
- Lysis buffer
- Step-gradient solutions: 50%, 20%, and 15% (w/v) sucrose
- TE buffer, pH 8.0 (*APPENDIX 2*)
- Dry granular sucrose
- 30°C orbital shaking incubator (e.g., New Brunswick Scientific #G-24)
- 250-ml conical centrifuge bottles (e.g., Corning #25350)
- 65°C water bath
- 25 × 89-mm tube (e.g., Beckman #344058)
- Beckman JS-4.2 and SW-27 rotors (or equivalents)
- Dialysis tubing (*APPENDIX 3*)
- Pyrex baking dish
- CHEF pulsed-field gel apparatus or equivalent (*UNIT 2.5B*)
- Additional reagents and equipment for size fractionation using a sucrose gradient (*UNIT 5.3*) and estimating DNA concentration (*UNIT 2.6*)

Grow and prepare the cells

1. Inoculate a single red colony of a YAC-containing clone into 25 ml AHC medium in a 250-ml flask. Shake at 250 rpm, 30°C, until culture reaches saturation (~3 days).
2. Transfer 1 ml of saturated culture to 100 ml AHC medium in a 1-liter flask. Shake 16 to 18 hr at 250 rpm, 30°C.
3. Harvest yeast cells by centrifuging 10 min at 2000 × *g* (2800 rpm in Beckman JS-4.2 rotor), room temperature, using a 250-ml conical centrifuge bottle. Discard supernatant.
4. Resuspend cells in 50 ml water. Centrifuge 5 min at 2000 × *g*, room temperature. Discard supernatant.
A cell pellet of ~4 g should be obtained.
5. Resuspend cells in 3.5 ml SCEM buffer.

Lyse the cells

6. Incubate 2 hr at 37°C with occasional gentle mixing. The mixture will become highly viscous.
7. Gradually add cell mixture to 7 ml lysis buffer in a 250-ml Erlenmeyer flask by allowing viscous cell suspension to slide down side of flask.
8. Gently mix by swirling flask until mixture is homogeneous and relatively clear.
9. Incubate 15 min at 65°C, then cool rapidly to room temperature in a water bath.

Fractionate cell contents

10. Fractionate on a sucrose step gradient. In a 25 × 89–mm tube, prepare a step gradient consisting of:

- 3 ml 50% sucrose
- 12 ml 20% sucrose
- 12 ml 15% sucrose
- 11 ml lysed sample.

Centrifuge 3 hr at 125,000 × *g* (26,000 rpm in a Beckman SW-27 rotor), room temperature.

11. Discard ~25 ml from top of gradient using a 10-ml pipet.

Dialyze and analyze DNA

12. Collect viscous DNA-containing solution at the 20% to 50% sucrose interface (~5 ml total volume) and place in dialysis tubing, leaving room for volume to increase ≥2- to 3-fold. Dialyze overnight against 2 liters TE buffer, pH 8.0, at 4°C.
13. Reconcentrate dialyzed DNA by placing dialysis tubing in an autoclaved Pyrex baking dish and covering with granular sucrose. Recover dialysis tubing when volume of contents has been reduced to ~2 ml.
14. Squeeze DNA solution to one end of dialysis tubing and tie an additional knot to keep DNA in a small volume. Dialyze overnight against 1 liter of TE buffer, pH 8.0, at 4°C.
15. Recover dialyzed DNA and check a small aliquot by electrophoresing in a CHEF pulsed-field gel. Stain with ethidium bromide and estimate DNA content by comparison to a known amount of λ DNA.

The DNA sample will contain a substantial amount of yeast RNA but should also contain a population of YAC DNA fragments migrating at a size of >100 kb. The presence of the RNA may make it difficult to determine the DNA concentration accurately; the concentration may be estimated by comparison to known DNA standards in an ethidium bromide-stained gel. The RNA will not affect restriction digestion of the DNA.

PREPARATION AND ANALYSIS OF A YAC-INSERT SUBLIBRARY

Construction of a sublibrary of fragments of the YAC insert facilitates high-resolution analysis of the insert sequence. This protocol details the steps required to produce a cosmid library, followed by a series of screenings to identify regions of interest and “walking” to establish a contiguous map of the insert.

Materials

High-molecular-weight YAC-containing DNA (fifth basic protocol)

Vector DNA (e.g., SuperCos 1, Stratagene #251301)

³²P-labeled (UNIT 3.10) probes: total genomic DNA of the individual or species from which the library was made (e.g., UNITS 2.2, 2.3 & 5.3), end-specific DNA (UNIT 3.10) or RNA (UNIT 3.8), and end fragment from YAC (fourth basic protocol or alternate protocol)

Additional reagents and equipment for restriction endonuclease digestion (UNIT 3.1), genomic DNA library production (UNIT 5.7), plating and transferrin a cosmid library (UNIT 6.2), and hybridization with radioactive probes (UNITS 6.3 & 6.4)

BASIC PROTOCOL

Screening of Recombinant DNA Libraries

6.10.14

Construct the library

1. Partially digest 1 to 2 μg of YAC-containing DNA with restriction endonuclease(s) appropriate for cosmid vector to be used. For example, to clone into the *Bam*HI site of SuperCos 1, digest the YAC DNA with either *Mbo*I or *Sau*3A.

The quantity of restriction endonuclease should be adjusted to produce digested fragments with an average size of ~40 kb. Although only a small fraction of the YAC-containing DNA used as starting material is actual YAC DNA (the rest being yeast genomic DNA), because of the low complexity of the yeast genome (e.g., compared to the human genome), only 3000 to 5000 cosmid clones are required to yield 3 yeast genome equivalents. Thus, only 1 to 2 μg of yeast DNA are required to make an adequate library.

2. Perform a series of test ligations as described in UNIT 5.7. Using optimal conditions, ligate insert DNA to vector DNA.
3. Package cosmid recombinants; dilute packaged extract and determine the titer.
4. Plate and transfer the sublibrary as appropriate for the vector, and prepare resulting filters for hybridization.

Screen the sublibrary

5. Perform a preliminary screen of the library using a ^{32}P -labeled probe of total genomic DNA of the individual or species from which the library was made.

This probe is a source of repetitive sequences. Because these repetitive sequences are spaced frequently throughout the source genome, and are absent from yeast, this probe will identify most of the source-DNA insert cosmids from the excess of yeast insert cosmids.

6. Organize this first set of cosmid clones into contigs by analyzing shared restriction fragments and by hybridizing with probes contained in the YAC insert or prepared from the ends of individual cosmid inserts.

Cosmid end-fragment-specific probes can be generated by digesting cosmid DNA and end-labeling the purified restriction fragment(s) that contain(s) the cloning site. If the cloning vector is SuperCos 1 (or a comparable vector), end-specific RNA probes may also be transcribed from the ends of the cosmid clones using T3 and T7 polymerase (see critical parameters).

7. Establish a complete contiguous collection of cosmid clones of the original YAC insert by screening the library with specific YAC-derived probes and cosmid end-specific probes.

Note that nitrocellulose filters may be reused for hybridization in subsequent steps without further washing or removal of probe. Repeated hybridization with sequential "walking probes" should reveal new hybridizing colonies at each step.

REAGENTS AND SOLUTIONS

AHC medium and plates (*ura⁻, trp⁻*)

1.7 g yeast nitrogen base *without* amino acids and *without* ammonium sulfate (Difco)
5 g ammonium sulfate
10 g casein hydrolysate-acid, salt-free and vitamin-free (U.S. Biochemical #12852)
50 ml (for medium) or 10 ml (for plates) of 2 mg/ml adenine hemisulfate (Sigma #A-9126)
Dissolve in a final volume of 900 ml H₂O
Adjust pH to 5.8
Autoclave 30 min, then add 100 ml sterile 20% (w/v) glucose. For AHC plates, add 20 g agar prior to autoclaving. Store at 4°C for ≤6 weeks.

Lithium lysis solution

1% lithium dodecyl sulfate (Sigma # L-4632)
100 mM EDTA
10 mM Tris·Cl, pH 8.0 (*APPENDIX 2*)
Filter sterilize and store indefinitely at room temperature

Lysis buffer

0.5 M Tris·Cl, pH 8.0 (*APPENDIX 2*)
3% (v/v) *N*-lauroylsarcosine (Sarkosyl)
0.2 M EDTA, pH 8.0 (*APPENDIX 2*)
Store indefinitely at room temperature. Add 1 mg/ml proteinase K just before use.

100% NDS solution

Mix 350 ml H₂O, 93 g EDTA, and 0.6 g Tris base. Adjust pH to ~8.0 with 100 to 200 pellets of solid NaOH. Add 5 g *N*-lauroylsarcosine (predissolved in 50 ml water) and adjust to pH 9.0 with concentrated NaOH. Bring volume to 500 ml with water. Filter sterilize and store indefinitely at 4°C. Dilute 1:5 with H₂O (20% final) just before use.

PCR reaction mix

1.5 mM MgCl₂
50 mM KCl
10 mM Tris·Cl, pH 8.3 (*APPENDIX 2*)
0.2 mM each dATP, dCTP, dGTP, and dTTP
0.05 U AmpliTaq polymerase (Perkin-Elmer/Cetus)/μl reaction mixture
0.03 μl Perfect Match Enhancer (Stratagene)/μl reaction mixture
Store all components at -20°C and mix just before use

SCE buffer

0.9 M sorbitol (Fisher, molecular biology grade)
0.1 M sodium citrate
0.06 M EDTA, pH 8.0
Adjust pH to 7.0
Store at room temperature ≤3 months

SCEM buffer

4.9 ml SCE buffer (see above)
0.1 ml 2-mercaptoethanol (2-ME)
Add 1 to 2 mg Lyticase (Sigma #L-8137 or ICN Biomedicals #190123) just before use.

SEM buffer

1 M sorbitol
20 mM EDTA, pH 8.0
14 mM 2-ME
Filter sterilize
Store at 4°C for ≤6 weeks

SEMT buffer

1 M sorbitol
20 mM EDTA
14 mM 2-ME
10 mM Tris·Cl, pH 8.0 (*APPENDIX 2*)
Filter sterilize
Add 1 mg/ml Lyticase (Sigma #L-8137 or ICN Biomedicals #190123) just before use.

COMMENTARY

Background Information

An overview of strategies for screening YAC libraries and analyzing YAC clones is presented in *UNIT 6.9*.

Critical Parameters and Troubleshooting

The protocols provided in this unit are intended to describe the analysis and characterization of particular YAC clones of interest. It is initially desirable to assure the integrity of the clone; that is, to ensure that the YAC indeed carries the proper insert, that the insert is not chimeric or rearranged, and that it is stably maintained and propagated in the yeast host.

Growth of YAC-containing strains. In YAC clones, genomic DNA is inserted into a cloning site carried within the *SUP4* gene of the vector (see Figs. 6.10.1 and 13.4.6). In the parent vector, the *SUP4* product complements the *ade2-1* ochre mutation carried in the host AB1380. This mutation causes a block in purine biosynthesis, resulting in accumulation of red pigment in the culture. Thus, disruption of *SUP4* by insertional inactivation prevents complementation of the ochre mutation in the host. Before a strain is used, it is important to check that upon growth the colonies or cultures exhibit a red pigmentation. If not, another isolate should be used. Additionally, growth on selective AHC medium requires the presence of both arms of the YAC vector and favors stability of the clone through passage.

Analysis of YAC DNA. Restriction analysis of purified YAC DNA (second basic protocol) can be used to assess YAC structure. If fragments of unanticipated sizes are detected in the YAC, then it is likely that the YAC contains

sequences homologous to, but different from, the desired clone, or that the YAC insert has undergone some sort of rearrangement during cloning. Alternatively, lack of methylation of the YAC DNA at the restriction enzyme recognition site may give a digestion pattern not seen in uncloned genomic DNA. Determination of size and stability of the YAC clone is made by preparing chromosomes in agarose plugs and subsequent pulsed-field gel electrophoresis. The PFGE gel can be blotted and analyzed by hybridization with sequence-specific probes and end-fragment probes. The results should reveal hybridization to the same size chromosome in all isolates of a given YAC clone. Variation in the size of the artificial chromosome between yeast isolates derived from one clone indicates YAC instability. Hybridization of a YAC vector arm probe to more than one artificial chromosome may indicate multiple transformation of the strain at the time of library construction. Alternatively, it may represent strain instability, with the smaller chromosome(s) representing deletion products of the original YAC. When two or more artificial chromosomes are identified with a single-copy genomic insert-specific probe, instability of the YAC insert is the most likely cause.

Assessing chimerism. In most existing YAC libraries, chimeric clones are observed to represent from 5% to 50% of the total clones. One of the most reliable ways to identify a chimeric YAC insert is to isolate a small fragment from each end of the YAC insert and define its chromosome of origin and whether it is contained in overlapping YACs from the same chromosomal region. Two procedures are given for analyzing the end fragments of

the YAC clone, which can be used to identify chimerism in the YAC insert—one based upon subcloning (which for large-scale mapping projects can require prohibitive amounts of time and effort) and one based upon PCR. With the PCR strategy, the resulting amplified product should migrate as a single fragment in a polyacrylamide gel. If multiple bands are present, it may be possible to demonstrate that one is an appropriately amplified fragment because it should be digested by *EcoRI* to yield the vector-linker fragment plus the insert end fragment (see Fig. 6.10.3). If digestion by *EcoRI* cannot be confirmed or if no amplified band is observed, it is useful to try a “hemi-nested” PCR amplification in which the initially amplified product is reamplified using another primer that should be contained only in the properly amplified fragment.

Each end-fragment probe should be shown to be single-copy by hybridization to a Southern blot of total DNA from the species used to prepare the YAC library. If a smear of hybridization is obtained, repetitive sequences are present within the probe. In the case of fragments obtained by subcloning, it is usually possible to identify a single-copy probe by digesting the fragment into several smaller pieces using selected restriction enzyme. The end fragments recovered by PCR are usually small, so that it is generally impossible to salvage a single-copy probe. It may be possible to suppress the repetitive DNA hybridization by including in the hybridization reaction an excess of unlabeled denatured repetitive DNA fragments from the species used to prepare the YAC library. If this is unsuccessful, the alternate protocol (subcloning into a bacterial vector) is usually necessary.

Construction of a cosmid sublibrary. For further high-resolution analysis and mapping of the YAC insert, it is desirable to construct a sublibrary (final basic protocol) from the YAC-containing DNA (fifth basic protocol).

A number of cosmid and phage vectors are available that are suitable for subcloning YACs into bacterial vectors (*UNITS 3.16 & 5.7*). One excellent candidate is the SuperCos-1 cosmid vector (Stratagene), which can accommodate inserts in the 35- to 42-kb range. It contains a neomycin-resistance cassette that permits selection of transfected clones in mammalian cells. It also contains T3 and T7 phage promoters flanking the genomic insert, which facilitate generation of RNA probes from the ends of the genomic inserts. This feature is useful for verifying overlaps of clones or to permit chromo-

some walking if a complete cosmid contig is not established in the first round of screening.

Once a sublibrary has been constructed, it should be screened with a probe consisting of ³²P-labeled total genomic DNA from the individual or species that was originally used to construct the library (a source of repetitive sequences). This will identify most of the cosmids containing DNA inserts from the source genome. Individual clones can be analyzed by Southern blotting with probes from the YAC insert or with genomic repetitive sequences. This data, together with results of Southern blots using probes derived from the ends of the cosmid clones using T3 or T7 polymerase, can be used to organize the cosmids into contigs. If a complete contig is not established, the cosmid library can be screened again with probes representing the ends of the YAC inserts, or derived from the ends of cosmid clones. Empirically, different portions of the YAC insert have been found to be nonrandomly represented in the cosmid library. Consequently, it is common for one or more rounds of chromosome walking to be required to fill in gaps between cosmid contigs.

Anticipated Results

The basic protocol for preparation of DNA from YAC clones can be expected to yield ~1 to 1.5 μg of DNA in the size range of 50 to 200 kb from 10⁸ yeast cells. The yield of DNA obtained by purification using preparative CHEF gel electrophoresis is ≥10-fold lower. The basic protocol for preparation of yeast chromosomes in agarose plugs for PFGE should yield sufficient material for ~40 lanes of a pulsed-field gel from a 25-ml yeast culture. The basic and alternate protocols for analysis of the YAC insert end fragments (by PCR and by subcloning into a plasmid vector) should each yield DNA fragments that identify single hybridizing bands in genomic DNA when used as probes for Southern blots. The basic protocol for preparation of high-molecular-weight YAC-containing yeast DNA in solution should yield ~25 to 50 μg of DNA ≥100 kb in size from a 100-ml culture. Basic protocol for preparation of a YAC insert cosmid sublibrary should yield ~1500 colonies/μg of starting yeast DNA. From these, ~30 to 50 genomic DNA-containing cosmids should be recovered.

Time Considerations

Purifying YAC DNA using the first basic protocol requires ~3 days to grow the culture, 3.5 hr to isolate the DNA, and after an overnight

resuspension, ~2 hr to remove RNA and reprecipitate the DNA.

Preparation of yeast chromosomes in agarose plugs takes ~4 days to grow the yeast culture, ~1½ hr to form the yeast-containing plugs, ~6 hr to lyse the yeast, an overnight incubation, and ~6 hr to wash and prepare the plugs for electrophoresis.

Preparation of YAC insert end fragments by PCR takes from 6 hr to overnight to anneal the bubble primers, digest the YAC DNA, and ligate it to the bubbles. Another 3 to 5 hr are needed to amplify the end fragment by PCR.

Isolation of YAC end fragments by subcloning requires 2 to 3 days to perform the preliminary analytical Southern blot to identify the enzyme combination of choice for subcloning. Once this is identified, 1 day is required for preparative isolation of the doubly digested DNA fragments, followed by an overnight ligation to the modified pUC 19 vector. An additional day is needed to transform bacteria and grow colonies, and 1 to 2 days are required to identify the specific end-fragment

subclones by hybridization. Finally, 2 days are needed to purify the subcloned DNA and to verify its structure by restriction enzyme analysis.

Preparation of high-molecular-weight YAC-containing DNA requires ~4 days to grow the culture, 7 hr to lyse the cells and perform the sucrose-gradient fractionation, an overnight dialysis, and 2 to 3 hr to concentrate the DNA. Following an additional overnight dialysis, the DNA is ready for use.

YAC insert cosmid sublibrary preparation and analysis takes 2 days to perform test digestions and ligations. Another 2 days are required to perform the preparative digestion, ligation, packaging, and plating of the library. Two more days are required for preparation of filters, hybridization, washing, and autoradiography.

Contributed by David D. Chaplin and
Bernard H. Brownstein
Howard Hughes Medical Institute and
Washington University School of Medicine
St. Louis, Missouri

SPECIALIZED STRATEGIES FOR SCREENING LIBRARIES

Use of Monoclonal Antibodies for Expression Cloning

This unit details the use of transient expression in mammalian cells to screen cDNA libraries with monoclonal antibodies (MAB) to isolate cDNA clones encoding cell-surface and intracellular proteins. The first basic protocol describes the cloning of cDNAs encoding cell-surface antigens. Several steps in this protocol involve transfection procedures that are described in greater detail in *UNIT 16.12*. The second basic protocol is a modification that facilitates isolation of cDNAs encoding antigens that are expressed intracellularly. Both protocols are designed for use with the expression vector CDM8, which contains a polylinker for subcloning double-stranded cDNA (Fig. 16.12.1).

ISOLATION OF cDNA CLONES ENCODING CELL-SURFACE ANTIGENS

This protocol is designed to isolate cDNAs encoding cell-surface proteins by screening cDNA libraries transiently expressed in mammalian cells. The procedure requires multiple rounds of transfection and immunoselection and is divided into four sections: (1) COS cell transfection by the DEAE-dextran method, (2) immunoselection by panning, (3) plasmid recovery and *E. coli* transformation, and (4) COS cell transfection by the spheroplast fusion method. A total of four rounds of transfection and immunoselection (one using DEAE-dextran, three using spheroplast fusions; Fig. 6.11.1) are used. After the final round of immunoselection, plasmid DNA is prepared from individual bacterial colonies. COS cells are then transfected with this DNA by the DEAE-dextran method and examined for their ability to express the foreign protein of interest by immunofluorescence microscopy (*UNIT 14.6*) or flow cytometry analysis (Holmes and Fowlkes, 1991).

DEAE-dextran transfection is a highly efficient means of introducing the cDNA library into COS cells to ensure that the transfected cells receive as complete a library representation as possible. Typically, ten 100-mm tissue culture plates of COS cells are transfected (Fig. 6.11.1). The subsequent panning steps allow rapid and efficient culling of cells expressing the protein of interest from the bulk of the transfected cells (each 60-mm, antibody-coated plate can be used to pan $1-2 \times 10^7$ transfectants). Plasmid DNA can be rescued from the panned cells by obtaining a Hirt supernatant (Hirt, 1967) and following amplification in *E. coli*, the plasmid DNA can be reintroduced into COS cells using spheroplast fusion. This transfection procedure ensures that a single plasmid type is delivered into each transfectant, allowing greater enrichment in subsequent rounds of transfection and immunoselection. Each round of screening usually requires a set of six fusions and each set of six fusions requires 100 ml of cells in broth.

NOTE: All incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise noted. See Chapter 9 introduction for critical parameters concerning media components and preparation.

Materials

- Complete Dulbeccos minimum essential medium containing 10% (v/v) NuSerum or 10% (v/v) calf serum (complete DMEM-10 NS or complete DMEM-10 CS; *APPENDIX 3F*)
- 100-mm tissue culture plates seeded with COS cells (~50% confluent)

cDNA library: plasmid expression vector DNA containing $>10^6$ of cDNA clones (UNIT 5.8; see background information), CsCl-purified (UNITS 1.7 & 9.1)

Phosphate-buffered saline (PBS; APPENDIX 2)

DEAE-dextran/chloroquine solution: PBS containing 10 mg/ml DEAE-dextran (Sigma) and 2.5 mM chloroquine (Sigma)

10% (v/v) DMSO in PBS

Trypsin/EDTA solution: PBS containing 0.5 mg/ml trypsin + 0.2 mg/ml EDTA

0.5 mM EDTA/0.02% (v/v) azide in PBS

0.5 mM EDTA/0.02% (v/v) azide/5% (v/v) calf serum in PBS

1 μ g/ml purified monoclonal antibody (MAb) or 1:100 dilution of ascites fluid (UNIT 11.1)

0.5 mM EDTA/0.02% (v/v) azide/2% (w/v) Ficoll

60-mm antibody-coated plates (first support protocol)

5% (v/v) calf serum in PBS

0.6% (w/v) SDS/10 mM EDTA

5 M NaCl (APPENDIX 2)

Phenol (extracted twice with 1 M Tris·Cl, pH 7.5)

2 μ g/ μ l linear polyacrylamide

TE buffer, pH 7.5 (APPENDIX 2)

Electroporation-competent *E. coli* cells (UNIT 1.8)

LB medium (UNIT 1.1)

100 mg/ml spectinomycin or 35 mg/ml chloramphenicol in ethanol

20% (w/v) sucrose/50 mM Tris·Cl, pH 8.0, ice cold

5 mg/ml lysozyme (Sigma #L6876), freshly prepared in 250 mM Tris·Cl, pH 8.0

250 mM EDTA, ice cold (APPENDIX 2)

50 mM Tris·Cl, pH 8.0 (APPENDIX 2)

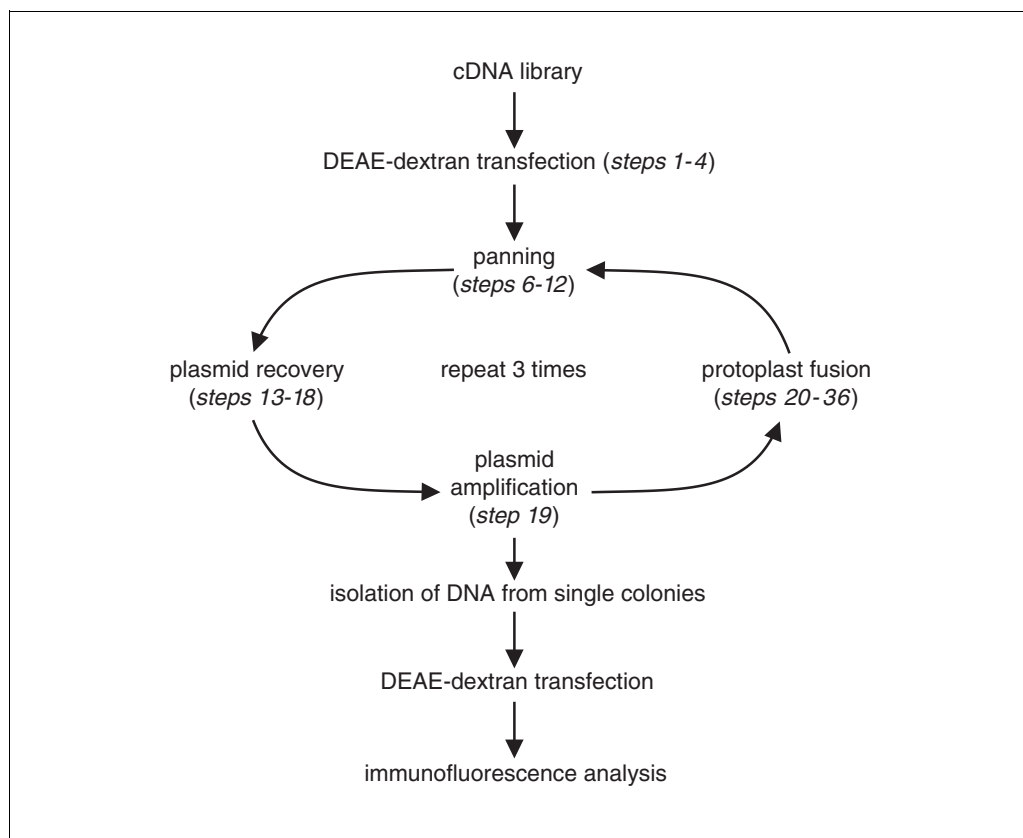


Figure 6.11.1 Isolation of a cDNA clone encoding a cell-surface antigen by transient expression in mammalian cells.

10% (w/v) sucrose/10 mM MgCl₂ in DMEM (GIBCO/BRL #320-1960AJ)
without serum, filter sterilized

60-mm tissue culture plates seeded with COS cells (~50% confluent)

50% (w/w) PEG 1000 or 1450 in DMEM (no serum), adjusted to pH 7 with
7.5% (w/v) sodium bicarbonate (Baker or Kodak)

DMEM without serum

Complete DMEM-10 CS (*APPENDIX 3F*) containing 15 µg/ml gentamycin sulfate

Nylon mesh, 100-µm pore size (Tetco)

Sorvall GSA rotor or equivalent

Swinging-bucket centrifuge (e.g., Sorvall RT-6000B)

Additional reagents and equipment for transformation of *E. coli* by electroporation
(*UNIT 1.8*), phenol extraction and ethanol precipitation (*UNIT 2.1*), alkaline lysis
miniprep (*UNIT 1.7*), and immunofluorescence (*UNIT 14.6*)

Transfect COS cells using DEAE-dextran

1. Add 5 ml complete DMEM-10 NS to each 100-mm plate of COS cells to be transfected.

Each 100-mm plate should be ~50% confluent the day of transfection (~5 × 10⁶ cells). This protocol is designed to be used with COS cells and is too harsh for WOP or MOP cells (see background information). If these murine lines must be used, it is important to reduce both the concentration of DEAE-dextran used to 200 µg/ml final and the time that the cells are exposed to the transfection medium to 2 hr, and to use IMDM (Iscoves modified Dulbeccos medium; GIBCO/BRL #430-2200) in place of DMEM (prepare complete IMDM media as for complete DMEM media, but omit amino acids).

2. To each dish, add 5 µg cDNA library and mix, then add 0.2 ml DEAE-dextran/chloroquine solution and mix. Incubate 4 hr.

Typically, libraries of >10⁶ clones are used to obtain plasmid DNA. It is important that the DNA and the DEAE-dextran form a fine, invisible precipitate. If the DNA is not diluted prior to addition of DEAE-dextran, a large DNA/DEAE-dextran precipitate forms (it is easily seen), which is not readily taken up by the cells.

Check the cells after ~3 hr exposure to the DEAE transfection mix, as their health can decline rapidly. This is particularly true of chloroquine transfections, and it is usually better to shorten the transfection than to allow too many cells to die.

3. Aspirate the medium and add 2 ml of 10% DMSO. Incubate ≥2 min at room temperature.

The time that the cells are exposed to the DMSO is not critical.

4. Remove DMSO and replace with 10 ml complete DMEM-10 CS. Incubate overnight.

5. Aspirate the medium, add PBS, then aspirate the PBS. Add 2 ml trypsin/EDTA to each plate and incubate 5 to 15 min until cells have lifted from the plate. Replate the cells on two new 100-mm plates and incubate overnight.

Replating the cells allows them to recover more effectively from the transfection. In addition, the DEAE-dextran transfection makes the cells sticky and replating allows them to be lifted from the plates with EDTA to initiate the panning step.

Immunoselect the cells by panning

6. Aspirate the medium, add 2 ml EDTA/azide solution, and incubate 10 to 20 min to detach cells from plates.

7. Pipet vigorously with a short Pasteur pipet to dislodge the cells, then transfer cells from each plate into a 15-ml centrifuge tube.

8. Centrifuge 4 min at $200 \times g$ (e.g., 1000 rpm in a Sorvall RT-6000B with GSA rotor or in a tabletop centrifuge) and discard supernatant.
9. Resuspend cells in 0.5 to 1.0 ml EDTA/azide/calf serum solution and add purified MAb to 1 $\mu\text{g}/\text{ml}$ final or ascites at a 1:100 dilution final. Incubate 30 to 60 min on ice.
10. Add an equal volume of EDTA/azide solution and carefully layer on 3 ml EDTA/azide/Ficoll solution. Centrifuge 4 min at $200 \times g$. Aspirate supernatant in one smooth movement.
11. Add 3 ml EDTA/azide/calf serum solution to each antibody-coated plate. Resuspend cells in 0.5 ml EDTA/azide solution, then add aliquots of the cells to these plates by pipetting them through a nylon mesh. Leave 1 to 3 hr at room temperature.

Four 6-mm antibody-coated plates are used in each round of panning. It is important to pass the cells through the nylon mesh to break up large clumps of cells which might contain both positive and negative cells. This ensures that individual antibody-coated cells bind to the panning plate.

12. Remove excess cells not adhering to the plate by gently washing two to three times with 3 ml of 5% calf serum (or complete DMEM-10 NS or complete DMEM-10 CS).

Washing gently means swirling the plate with a smooth motion for ~30 sec. The plate obtained after these washes is known as a panned plate.

Recover plasmid DNA and transform *E. coli*

13. Add 0.4 ml SDS/EDTA solution to the panned plate and leave 20 min at room temperature (to lyse the cells).

This incubation period can be as little as 1 min if there are only a few cells on the plate.

14. Pipet the viscous mixture into a microcentrifuge tube. Add 0.1 ml of 5 M NaCl, mix, and place ≥ 3 hr on ice or leave overnight at 4°C .

The viscosity is primarily due to the genomic DNA. It is important to avoid shearing the genomic DNA so that it will not contaminate the plasmid DNA. Keeping the mixture as cold as possible seems to improve the quality of the Hirt supernatant.

15. Microcentrifuge 4 min at top speed, 4°C , and remove supernatant carefully.
16. Extract with phenol (twice if the first interface is not clean) and add 5 μl (10 μg) of 2 $\mu\text{g}/\mu\text{l}$ linear polyacrylamide (or other carrier).
17. Fill the tube to the top with 100% ethanol and precipitate. Resuspend the pellet in 0.1 ml TE buffer, pH 7.5.
18. Add 10 μl of 3 M sodium acetate and 300 μl of 100% ethanol, and repeat precipitation. Resuspend the pellet in 0.1 ml TE buffer, pH 7.5.
19. Transform electroporation-competent *E. coli* cells by electroporation using DNA obtained from step 18. Incubate overnight at 37°C .

*Approximately 10^5 bacterial colonies should be obtained. It is advisable to transform *E. coli* with an aliquot of DNA to determine the amount necessary to obtain 10^5 colonies. Generally $1/10$ to $1/4$ of the recovered DNA will be used.*

Prepare the spheroplasts

20. Rinse the plate from step 19 several times with LB medium while scraping with a spreader to dislodge the bacteria. Use $\frac{1}{10}$ to $\frac{1}{5}$ of the pooled scrapings to inoculate 200 ml of LB medium. Grow to $OD_{600} = 0.5$ at 37°C with shaking.
21. Add 100 mg/ml spectinomycin to 100 $\mu\text{g}/\text{ml}$ or 35 mg/ml chloramphenicol to 150 $\mu\text{g}/\text{ml}$. Incubate with shaking 10 to 16 hr at 37°C .

Do not let the cells grow >16 hr or they will begin to lyse. If the cells lyse, do not proceed.

22. Centrifuge 100 ml of the culture in a 250-ml bottle, 5 min at $4000 \times g$ (e.g., 5000 rpm in a Sorvall with GSA rotor), room temperature or 4°C .
23. Drain well and resuspend pellet in 5 ml of ice-cold sucrose/Tris·Cl, pH 8.0.
24. Add 1 ml of 5 mg/ml lysozyme solution. Incubate 5 min on ice.
25. Add 2 ml of ice-cold 250 mM EDTA, pH 8.0, and incubate 5 min on ice.
26. Add 2 ml of 50 mM Tris·Cl, pH 8.0, and incubate 5 min in a 37°C water bath.
27. Place on ice. Check percent conversion to spheroplasts by microscopy.

A good spheroplast preparation gives about 80% to 90% conversion; anything <50% should not be used. Spheroplasts should appear as individual spheres, not as clumps.

28. In a tissue culture hood, slowly add 20 ml of ice-cold sucrose/ MgCl_2 solution dropwise at ~ 2 drops per second.

If this step is not carried out slowly, the spheroplasts will lyse. Lysis can easily be detected because the medium will become extremely viscous.

Transfect the COS cells by spheroplast fusion

29. Remove the medium from the COS cells plated the day before in 60-mm plates.

If the cells are confluent, they are more likely to lift from the plate.

30. Add 5 ml of the spheroplast suspension (from step 28) to each plate of cells (from step 29).
31. Place plates on top of tube carriers in a swinging-bucket centrifuge. Centrifuge 10 min at $100 \times g$ (e.g., setting 5.7 in a Sorvall RT-6000B with GSA rotor), room temperature. Aspirate the supernatant carefully from the plates.

Up to six plates can be prepared at once. Plates can be stacked on top of each other, but three in a stack is not advisable as the spheroplast layer on the top plate is often torn or detached after centrifugation.

32. Pipet 1.5 to 2 ml of 50% PEG solution into the center of the plate. If necessary, sweep the pipet tip around to ensure that the PEG spreads evenly and radially across the entire plate.

Use of PEG 1000 yields higher-efficiency transfections. This advantage is offset by the higher toxicity, requiring more careful timing of exposure; therefore, the use of PEG 1450 is advisable until the technique is mastered. WOP and MOP cells are more sensitive to PEG than are COS cells. When these cell lines are used, PEG 1450 is recommended and the procedure must be done as quickly as possible to prevent cell death.

33. After PEG has been added to the last plate, prop the plates up on their lids so that the PEG solution collects along the edge. Aspirate the PEG solution. Incubate 90 to 120 sec for PEG 1000 or 120 to 150 sec for PEG 1450 at room temperature.

The thin layer of PEG that remains on the cells is sufficient to promote fusion; the PEG is

easier to wash off and better cell viability is obtained than if the bulk of the PEG is left behind.

34. Pipet 1.5 ml of serum-free DMEM into the center of the plate to sweep the PEG layer radially across the plate. Tilt plates and aspirate medium. Repeat this DMEM wash.
35. Add 3 ml complete DMEM-10 CS containing gentamycin sulfate. Incubate 4 to 6 hr.
36. Remove medium and replace with 3 ml of complete DMEM-10 CS containing gentamycin sulfate. Incubate 2 to 3 days.
37. Repeat steps 6 to 36 two times using the transfected cells obtained in step 36.
38. Repeat steps 6 to 19 using the cells obtained in step 37. Prepare DNA from a single bacterial colony and transfect COS cells as in steps 1 to 4.

Approximately 1/4 of the DNA obtained from a 1.5-ml saturated bacterial culture is used for the transfection of a 100-mm plate.

39. Analyze the transfected cells by immunofluorescence microscopy or flow cytometry analysis.

One day after transfection, cells may be trypsinized (step 5) and replated in smaller plates to conserve materials used in staining. Flow cytometry analysis is beyond the scope of this book; the reader may consult Holmes and Fowlkes (1991) for a detailed description of this procedure.

SUPPORT PROTOCOL

PREPARATION OF ANTIBODY-COATED PLATES

This protocol describes a method for preparing plates coated with antibody to be used in the panning procedure of the first basic protocol.

Additional Materials

Anti-mouse affinity-purified antibody, sheep or goat (e.g., Cappel)

50 mM Tris·Cl, pH 9.5

0.15 M NaCl

1 mg/ml BSA in PBS (*APPENDIX 2*)

60-mm bacteriological plates

(e.g., Falcon #1007) or 100-mm plates (e.g., Fisher #8757-12)

1. Dilute anti-mouse antibody to 10 µg/ml in 50 mM Tris·Cl, pH 9.5. Add 3 ml diluted antibody per 60-mm plate or 10 ml per 100-mm plate and swirl. Incubate 1.5 hr at room temperature.

This antibody solution can be used on two more plates.

2. Wash plates three times with 0.15 M NaCl at room temperature.
3. Add 3 ml of 1 mg/ml BSA and incubate overnight at room temperature.
4. Aspirate the BSA and store plates at -20°C until used in step 11 of the first basic protocol.

ISOLATION OF cDNA CLONES ENCODING INTRACELLULAR ANTIGENS

One day after transfection with a cDNA library, COS cells are replated onto a polyvinylidene film. The following day, the cells are fixed with methanol. Cells expressing the target protein are radiolabeled by incubation with the appropriate antibody, followed by incubation with radioactive protein A or anti-Fc-antibody (alternatively, a radiolabeled primary antibody may be used). Autoradiography is used to identify the location of positive cells. Plasmid DNA is recovered by cutting the polyvinylidene wrap into squares and incubating them in SDS/EDTA and is then amplified by transforming *E. coli*. The COS cell transfection and screening procedure is repeated using plasmid DNA prepared from pools of bacterial colonies. Individual colonies from the pools enriched with the target gene are screened until the gene is clonally isolated. This strategy is illustrated in Figure 6.11.2.

**BASIC
PROTOCOL**

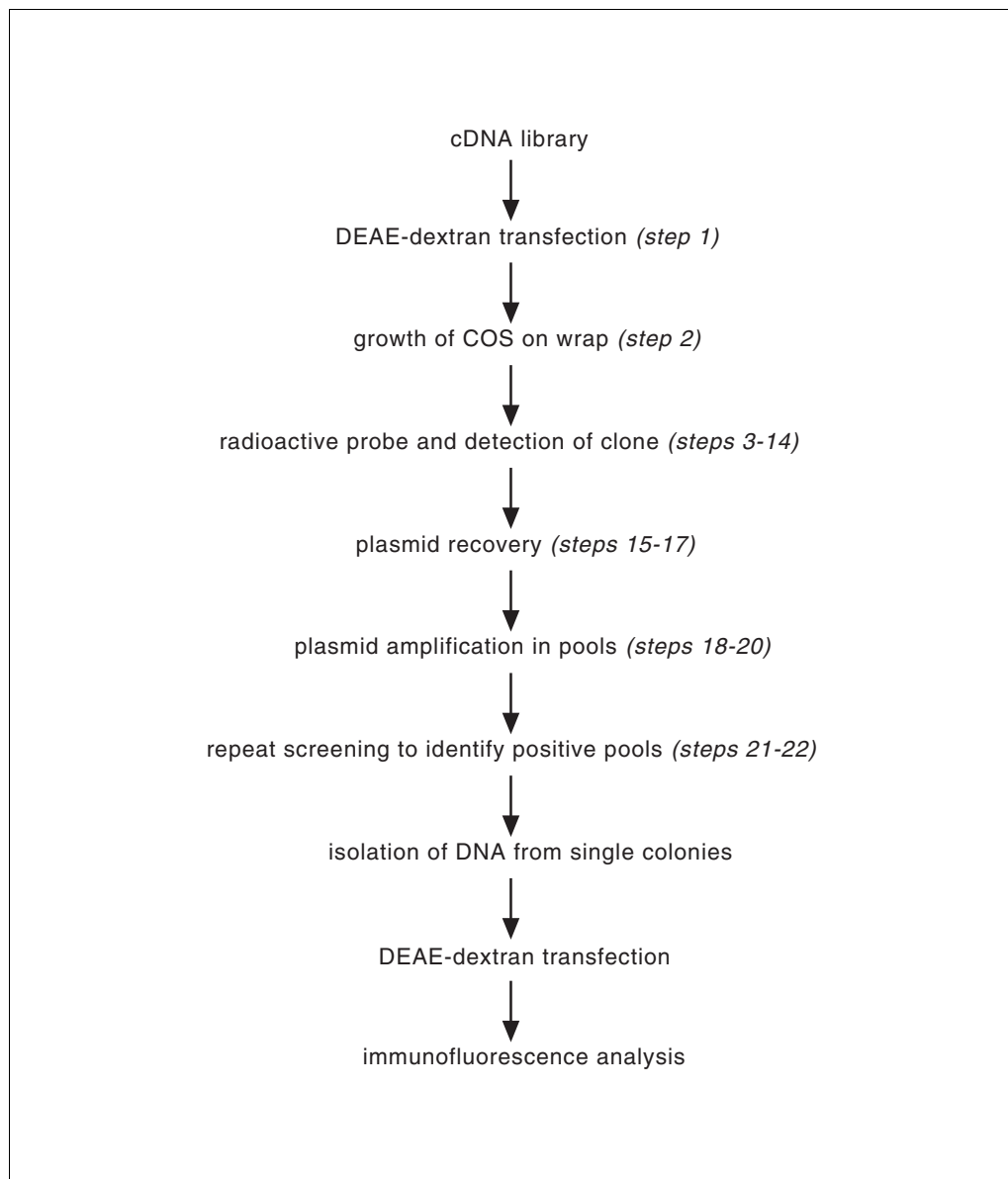


Figure 6.11.2 Sequence of steps for isolation of a cDNA clone encoding an intracellular antigen by transient expression in mammalian cells.

**Screening of
Recombinant
DNA Libraries**

6.11.7

Materials

100-mm tissue culture plates seeded with COS cells
cDNA library: plasmid DNA containing $>10^6$ cDNA clones (*UNIT 5.8*; see background information), CsCl purified (*UNITS 1.7 & 9.11*)
Trypsin/EDTA solution: PBS containing 0.5 mg/ml trypsin + 0.2 mg/ml EDTA
Phosphate-buffered saline (PBS; *APPENDIX 2*)
Methanol
1% (w/v) nonfat dry milk in PBS with and without monoclonal antibody (MAb)
1% (w/v) nonfat dry milk in PBS containing 0.25 μ Ci/ml of 125 I-labeled protein A
0.6% (w/v) SDS/10 mM EDTA buffer
LB medium (*UNIT 1.1*)

Polyvinylidene-wrapped plates (second support protocol)
X-ray film
Polyvinylidene wrap (e.g., Saran Wrap)
Rubber cement
Luminescent stickers

Additional reagents and equipment for alkaline lysis miniprep (*UNIT 1.7*) and autoradiography (*APPENDIX 3*)

Transfect cells and fix with methanol

1. Transfect ten 100-mm plates of COS cells with the cDNA library using the DEAE-dextran method as described in steps 1 to 4 of the first basic protocol.
2. Trypsinize each plate of transfected COS cells using trypsin/EDTA as described in step 5 of the first basic protocol and replate onto polyvinylidene-wrapped plates. Incubate 1 to 2 days.

Cells should be split at a ratio that will provide 50% to 75% confluency the following day. Complete DMEM containing penicillin and streptomycin should be used to avoid minor contamination that may occur because the plates are not sterile.

3. Remove medium from transfected COS cells. Wash plates by adding 5 ml PBS and aspirating at room temperature. Repeat wash one time.

Generally, 15 to 20 plates are used.

4. Add ~6 ml methanol and incubate 5 min at room temperature.
5. Wash plates three times with PBS, leaving the first addition of PBS on the plate for 2 to 3 min before aspirating.
6. Add 4 ml of 1% dry milk containing MAb to each plate. Incubate 30 to 60 min at room temperature.
7. Wash plates twice by adding 5 ml of 1% dry milk (without MAb), swirling gently, and removing solution.

Radiolabel and locate positive cells

8. Add 4 ml of 1% dry milk in PBS containing 125 I-labeled protein A to each plate. Incubate 30 min at room temperature.
9. Rinse plates four times with 1% dry milk in PBS and one time with PBS at room temperature. Remove all excess liquid.

It is helpful to prop the plates on edge to remove the excess liquid.

10. Completely cover a piece of X-ray film with polyvinylidene wrap and tape the edges

of the wrap to the film. Paint the wrap with a thin layer of rubber cement. Allow rubber cement to dry briefly.

The X-ray film acts as a support and may be a used piece that would otherwise be discarded. The presence of rubber cement does not appear to affect DNA recovery. Five plates will fit on an 8 × 10-inch (20.3 × 25.4-cm) film or 15 plates will fit on a 14 × 17-inch (35 × 43-cm) film.

11. Paint the bottom surface of the polyvinylidene-wrapped plate with a thin layer of rubber cement and allow the rubber cement to dry briefly.
12. Place the plates on the support. Lift the plate slightly and cut the wrap away from the edge of the plate.

The wrap with the cells will be left on the X-ray film. By slightly lifting the plate, a scalpel held nearly horizontal along the inside of the plate can be used to cut the wrap from the plate without cutting the wrap on the support. While this technique is not difficult, it is advisable to practice it on nonradioactive samples.

13. Affix luminescent stickers to the support. Cover the support and samples with polyvinylidene wrap.

There are now three layers of wrap on the support. Without the luminescent stickers, the film cannot be aligned over the samples to recover positive cells. Careful alignment is necessary to recover positive cells (there will be insufficient background to use the location of the samples for alignment).

14. Autoradiograph with an intensifying screen for 1 to 2 days at -70°C . Develop the film.
15. Align the film over the support using the luminescent stickers. Mark the location of positive cells by piercing the film alongside the spot with a needle (this leaves a mark on the polyvinylidene wrap).

A light box is helpful for locating the positive cells on the film.

16. Remove the film. Cut small squares of ~3 mm in the polyvinylidene wrap at the places it is marked.

Recover plasmid DNA and isolate the positive clone

17. Add 400 μl of SDS/EDTA buffer to a microcentrifuge tube. Place squares in tube (10 to 25 squares per tube) and incubate 30 min at room temperature.

All three layers of wrap are placed in the tube.

18. Recover plasmid DNA and transform *E. coli* as in steps 14 to 19 of the first basic protocol, plating transformed bacteria on separate LB plates to form pools of appropriate numbers of clones. Incubate overnight at 37°C .
19. Collect the bacteria from the LB plates by rinsing several times with LB medium while scraping with a spreader to dislodge the bacteria.
20. Prepare plasmid DNA from the scraped colonies using an alkaline lysis miniprep.
21. Transfect COS cells with the pools of plasmid DNA by the DEAE-dextran method as described in steps 1 to 4 of the first basic protocol.

In general, $1/10$ to $1/5$ of the DNA obtained from 1000 colonies is used to transfect one 100-mm plate of COS cells.

**SUPPORT
PROTOCOL**

22. Repeat steps 2 to 14 above to identify pools enriched with the gene of interest. Store the appropriate plasmid DNA in TE buffer at -20°C .

When screening pools, it is possible to use 60-mm plates (reduce all volumes to $\frac{1}{3}$).

23. Prepare DNA from a single bacterial colony, transfect COS cells, and analyze as in steps 38 and 39 of the first basic protocol.

PREPARATION OF POLYVINYLIDENE-WRAPPED PLATES

Plates are prepared in which the growth surface is a polyvinylidene wrap (to be used in the second basic protocol). The plates are quite sturdy and can be used in the same manner as standard tissue culture plates. Plates may be prepared a day or two in advance but the wrap will stretch and become floppy on prolonged storage.

Additional Materials

Chloroform
70% ethanol
0.1 mg/ml poly-L-lysine HCl (Sigma) in 50 mM Tris·Cl, pH 8.0, freshly prepared
100-mm or 60-mm tissue culture plates

1. Break the bottoms out of a 100- or 60-mm tissue culture plate with a blunt object.

Safety glasses are advisable. Strike the plate near the sides of the plate rather than in the center. If too much force is used, the sides of the plate will break as well. Structural stability is increased if the outer edges of the bottoms are not removed.

2. In a fume hood at room temperature, dip the top rim of the plate in chloroform to a depth of ~ 3 mm.
3. Shake off excess chloroform and place plate on a piece of polyvinylidene wrap laid flat. Place the plate and attached film into the lid of the plate to force the wrap into contact with the edges of the plate.
4. Remove the lid and gently but firmly pull the film tight to form a smooth surface.

Adhesion of the film to the outside of the plate helps maintain the strength of the seal.

5. Cut excess wrap from the plate with a razor blade.

The plate is now essentially inverted. The lid is placed over the opening that had been the bottom of the plate. When using 100-mm plates, a second lid is used to support the new wrap bottom.

6. Wash the plate two times with 70% ethanol. Allow the wrap to soak in the ethanol ~ 30 min.
7. Wash the plate with water.
8. Add 0.1 mg/ml poly-L-lysine to the plate—11 ml for a 100-mm plate and 4 ml for a 60-mm plate. Incubate 2 hr to overnight at room temperature.
9. In a tissue culture hood, rinse the plates twice with PBS.

Dishes should not be stored more than a few days because the wrap will stretch and become loose.

COMMENTARY

Background Information

Transient expression in mammalian cells has emerged as a powerful method for isolating cDNA clones that encode secreted cell-surface and intracellular proteins. It was first used to isolate cDNA clones encoding the lymphokine granulocyte/macrophage colony stimulating factor (GM-CSF; Lee et al., 1985; Wong et al., 1985). This cloning strategy is well suited for isolating any secreted proteins for which a rapid and sensitive bioassay exists, and has since been applied to isolate cDNA clones encoding a number of different lymphokines. Transient expression cloning was combined with the simple but powerful immunoselection technique of panning (Wysocki and Sato, 1978) to isolate a cDNA encoding the T cell-surface proteins CD2 and CD28 (Seed and Aruffo, 1987; Aruffo and Seed, 1987). This procedure has since been used to isolate cDNA clones encoding a number of different cell-surface proteins when antibodies against them were available. When antibodies against a cell-surface receptor of interest are not available, but its ligand is, transient expression in mammalian cells has been combined with ligand-binding assays to isolate cDNA clones encoding the receptor. This strategy was first used to isolate a cDNA clone encoding the receptor for the lymphokine interleukin 1 (Sims et al., 1988). Modifications to this protocol (Gearing et al., 1989) have since allowed the use of this strategy to isolate a number of receptors.

Recent improvements have allowed the use of transient expression in mammalian cells to isolate cDNA clones encoding intracellular proteins, including the major DNA-binding protein of the erythroid lineage (Tsai et al., 1989), the lysosomal membrane glycoprotein CD63 (Metzelaar, 1991), and fucosyltransferase, which adds fucose to *N*-acetylglucosamine with $\alpha(1,3)$ linkages, allowing the expression of the sialyl CD15 antigen (Goelz et al., 1990).

Transient expression in mammalian cells. Mammalian cells are ideal hosts for screening cDNA libraries prepared using mRNA isolated from higher eukaryotes. These cells are able to synthesize transcripts correctly from the cDNA clones in the library and are likely to process the proteins that they encode appropriately, thus maximizing the likelihood that the foreign proteins will be present in their native state and will be detectable using functional or immunological assays.

Mammalian cells were initially used as the host for isolation of genomic DNA fragments encoding oncogenes by stably introducing genomic DNA fragments derived from human tumors into murine cells (Goldfarb et al., 1982; Shih and Weinberg, 1982). The gene encoding the oncogene was then rescued from the transfected cells that had acquired the transformed phenotype. Stable transfections of mammalian cells were subsequently combined with immunoselection procedures to isolate genomic DNA fragments encoding the human HLA and β_2 microglobulin genes (Kavathas and Herzenberg, 1983) and cDNA clones encoding the T cell-surface proteins CD8 (Kavathas et al., 1984; Littman et al., 1985) and CD4 (Maddon et al., 1985).

The time involved in obtaining stable transfectants expressing the gene of interest and the difficulties associated with recovering the transfected DNA from the host's chromosomal DNA has limited the number of genes isolated using this cloning strategy. These difficulties prompted the development of transient expression systems for use in cloning with mammalian cells as the screening host. Unlike stable transfectants, these methods permit rapid preparation and detection of transfectants expressing the protein of interest and efficient recovery of the DNA encoding it.

Many technical advances have permitted efficient and routine screening of cDNA libraries in transiently expressing mammalian cells. These include the following developments: shuttle vectors that contain the appropriate eukaryotic transcription elements for high-level protein expression in transfected mammalian cells, mammalian cell lines that can act as effective heterologous expression hosts, and transfection protocols that allow efficient introduction of plasmid DNA into mammalian cells.

Expression vectors. A number of mammalian expression vectors permit screening of cDNA libraries by transient expression in mammalian cells (Chapter 16; Kaufman, 1990). These plasmids contain at least four basic elements: an efficient eukaryotic transcription unit, a viral-derived origin of replication, a prokaryotic origin of replication, and a prokaryotic selectable marker. One particular expression vector, CDM8 (Seed, 1987; Fig. 16.12.1; available from Invitrogen #V308-20) is especially engineered for these purposes, and is described in detail in *UNIT 16.12*. CDM8 con-

tains origins of replications derived from polyoma and SV40 viruses that allow for plasmid replication in cell lines expressing either the polyoma or SV40 large T antigens, respectively; usually, WOP (Dailey and Basilico, 1985) and COS (Gluzman, 1981) cells, respectively. Bacterial and M13 origins of replication (*ori*) are also present, allowing for plasmid amplification in bacteria and production of single-stranded DNA, respectively. The plasmid contains a *supF* gene as a prokaryotic selectable marker and a T7 RNA polymerase promoter for the preparation of mRNA in vitro from the subcloned cDNAs.

Mammalian cell lines. A number of cell lines have been developed that are excellent hosts for screening cDNA libraries prepared in the vector described above. Perhaps the most popular is the COS cell line (Gluzman, 1981), which was derived from the African green monkey kidney cell line CV-1 by transformation with an origin-defective SV40 virus. This cell line produces wild-type SV40 large T antigen but no viral particles. When plasmids containing an SV40 virus-derived *ori* are transfected into COS cells, the plasmid is replicated to a high copy number 48 hr posttransfection (10,000 to 100,000 copies/cell). This high-level replication has two important consequences. First, it allows for amplification of all DNA templates available for transcription. Second, it allows for recovery of the plasmid encoding the protein of interest from the immunoselected cells. This last point is of importance in the two basic cloning protocols because each cycle of transfection and immunoselection is followed by a plasmid-rescue step.

Other cell lines that have been used for expression cloning include the murine cell lines WOP (Dailey and Basilico, 1985) and MOP (Muller et al., 1984). These two cell lines express the polyoma large T antigen allowing the replication of plasmids containing a polyoma origin of replication in the transfected cells (1,000 to 10,000 copies/cell, 48 hr posttransfection). Another cell line, CV-1/EBNA, has been developed for screening cDNA libraries in conjunction with the expression vector pDC406, which contains an Epstein-Barr virus *ori* (McMahan et al., 1991).

The ability of COS cells to endure the transfection protocols and their ability to replicate the transfected plasmid to a very high copy number make them the cells of choice when screening a cDNA library with the methods described here. However, when the antibodies to be used in the immunoselection step cross-

react with proteins expressed by COS cells, another cell line must be used. In these cases, WOP or MOP cells can be used successfully, in spite of their more delicate nature and lower copy numbers of transfected plasmid.

Mammalian cell transfection. A number of transfection protocols have been developed for efficient introduction of foreign DNA into mammalian cells, including calcium phosphate, DEAE-dextran, spheroplast fusion, lipofection, and electroporation (*UNITS 9.1-9.5*). Two factors determine which transfection procedure should be used when screening a cDNA library by transient expression in mammalian cells: first, the efficiency of transfection, and second, the number of different plasmids that are introduced into each cell during transfection. Two transfection protocols, DEAE-dextran (McCutch and Pagano, 1968) and spheroplast fusion (Sandri-Goldin et al., 1981), are discussed below.

The mechanism by which DEAE-dextran transfections allow for introduction of foreign DNA into cells is poorly understood. It is believed that the positive charge of the DEAE-dextran polymer neutralizes the negative charge of the DNA polymer, forming a fine precipitate that can come into contact with the plasma membrane of the host cell. The DEAE-dextran/DNA complex is then internalized by pinocytosis. Some of this DNA makes its way to the host-cell nucleus, where it is replicated and transcribed. Because the foreign DNA enters the cell via endosomes, DNA integrity is enhanced by the addition of chloroquine to the transfection medium to prevent endosome acidification. DEAE-dextran transfections are very efficient, allowing for transfection of up to 70% of the host cells and delivery of up to 200 different plasmids into each transfected cell.

Introducing foreign DNA into mammalian cells by spheroplast fusion is very inefficient, allowing for transfection of only 1% to 2% of the host cells. Bacteria containing the foreign DNA are treated with lysozyme to remove their cell walls. The resulting spheroplasts are then fused with the host mammalian cell using polyethylene glycol (PEG), allowing introduction of the foreign DNA directly into the host cell cytoplasm. The DNA is then replicated and transcribed in the nucleus. Because of the inefficiency of the procedure, each host cell fuses with only one spheroplast, on average, introducing only a single plasmid type into each transfected cell.

The immunoselection screening method de-

scribed in the first basic protocol involves multiple rounds of mammalian cell transfection, immunoselection, and plasmid rescue steps. The rescued plasmids are amplified in *E. coli* and reintroduced into mammalian cells to initiate additional rounds of enrichment. To take full advantage of immunoselection, the two methods of transfection are used. When the cDNA library is first introduced into mammalian cells, it is important to obtain a complete representation of the library in the transfected host cells, ensuring that the protein of interest is expressed by the transfectants. For this reason, the first round of enrichment is initiated using DEAE-dextran transfection. In subsequent cycles of enrichment, it is important that a single plasmid type be delivered into each of the transfected cells to maximize the level of enrichment obtained in the subsequent immunoselection steps. This is accomplished using the spheroplast fusion transfection.

Immunoselection procedures. Two immunoselection techniques designed to rapidly select and enrich for plasmids encoding proteins of interest from a cDNA library transfected into mammalian cells are described. The first strategy is designed to isolate cDNA clones encoding surface proteins. A cDNA library prepared in a mammalian expression vector is transfected into COS cells using DEAE-dextran transfection. Forty-eight hours posttransfection, the cells are lifted from the plate and incubated with antibodies directed against the protein of interest. Cells expressing the foreign protein on their cell surface are easily culled from the bulk of the transfected cells by panning on plastic plates coated with anti-antibody antibodies (Wysocki and Sato, 1978). Plasmid DNA is then recovered from the transfected cells by the method of Hirt (Hirt, 1967), amplified in *E. coli*, and reintroduced into COS cells by spheroplast fusion. Two additional rounds of spheroplast fusion and panning are usually required to enrich for the plasmid encoding the protein of interest.

Panning has many advantages over other immunoselection procedures. It is fast, efficient (10^7 cells can easily be panned on two 60-mm plastic plates in 30 min), and very inexpensive. Other immunoselection techniques, such as sorting of fluorescence-labeled cells (Holmes and Fowlkes, 1991), may be used to screen cDNA libraries in transiently expressing mammalian cells (Yamasaki et al., 1988); but the greater demands on time, equipment, and technical expertise make these methods much less attractive.

The second strategy is designed to isolate cDNA clones encoding intracellular antigens. This method is a combination of the techniques described by Munro and Maniatis (1989) and Metzelaar et al. (1991). COS cells are transfected with a cDNA library by DEAE-dextran transfection. The day after transfection, they are replated onto poly-L-lysine-coated polyvinylidene wrap and allowed to grow for 1 to 2 additional days. They are then washed and fixed with methanol. The permeabilized cells are incubated with antibodies directed against the protein of interest, washed, and incubated with radiolabeled protein A (^{125}I). After washing, they are exposed to film to identify radiolabeled cells, which are then recovered by cutting the polyvinylidene wrap. The plasmid DNA is recovered from these cells by the method of Hirt, amplified in *E. coli*, and subjected to additional rounds of transfection and immunoselection.

Critical Parameters

The cloning strategy described in the first basic protocol is well suited for isolating cDNA clones encoding cell-surface proteins (see discussion above). If a cDNA library is thought to contain cDNA clones encoding a number of proteins of interest, it is possible to isolate all of them simultaneously by simply using a mixture of antibodies against all of the proteins of interest in the panning steps of the first three rounds of enrichment. The last cycle of immunoselection is carried out independently with antibodies against each of the different proteins.

Many times antibodies directed against the protein(s) of interest are of multiple isotypes. It is important, in this case, to use panning plates that have been coated with anti-Ig antibodies that bind to each of the isotypes present in the initial antibody pool. Alternatively, individual panning plates can be prepared for each of the antibody isotypes present in the initial antibody pool, but no significant advantage is achieved.

Although this cloning strategy has allowed isolation of a large number of cDNA clones encoding cell-surface proteins, it has some serious limitations. As with any expression cloning system, the gene of interest must initially be present in the library, the target protein must be functional or immunoreactive as a single chain, and the host system must posttranslationally modify the protein appropriately when these modifications are required for function or immunoreactivity. Compared with earlier bacterial expression cloning systems, the mam-

malian expression systems presented here are more likely to appropriately modify the gene products of higher eukaryotes. However, these methods require the host mammalian cell to express the target protein on its surface as a single molecule. This may not occur in cases where the target protein is part of a heterocomplex that requires more than one member for surface expression.

Several factors contribute to the successful application of this cloning strategy. The most critical parameter when screening a cDNA library by expression in mammalian cells is the quality of the cDNA library (for a more complete discussion, see *UNIT 5.8*). In addition, the quality of the COS cells and transfectability is of utmost importance. COS cells maintained in culture for prolonged periods tend to become refractory to transfection. For this reason, it is important to check the cells periodically for transfectability and replace them with cells from frozen stocks when necessary.

The competency level of the bacterial cells used to amplify the DNA rescued from the immunoselection step is very important and should be determined prior to the start of the experiment. If only a few positive cells are immunoselected by panning, it is of utmost importance that this plasmid DNA work its way into the bacteria so that it can potentially be amplified and thus be available for subsequent rounds of transfection and immunoselection. Ideally, only cells whose competency level is $\geq 10^9$ cfu/ μ g DNA should be used. For most bacterial strains, this can be achieved using electroporation.

When using panning to immunoselect the transfected cells, it is important to check that the antibody directed against the protein of interest does not bind to COS cells. If it does, another cell line should be used for screening. The researcher must also be mindful that DEAE-dextran treatment of cells changes their phenotype and thus the antibodies to be used in the panning step should be tested for cross-reactivity with mock-transfected cells.

The most difficult step, technically, is the spheroplast-fusion step. Careful timing of cell exposure to PEG is necessary to promote fusion while minimizing cell death. Although lower-molecular-weight PEG (PEG 1000) results in more efficient fusion, it is more toxic to the cells.

The cloning strategy described in the second basic protocol will not be effective if the antibodies used recognize epitopes that are sensitive to methanol treatment. Testing of the target

cells may reveal this limitation. If methanol sensitivity cannot be assayed, multiple antibodies against a given protein should be used when available.

Anticipated Results

In the first basic protocol, the authors typically use DNA prepared from a cDNA library of $\geq 10^6$ clones for screening. After four rounds of transfection, immunoselection, and plasmid rescue, 12 individual colonies are picked and plasmid DNA prepared from them. If the cloning procedure has been successful, at least one DNA preparation directs the expression of the protein of interest. In general, if the screening is unsuccessful, this indicates that the clone of interest may not be present in the library. In this case, screening of a new cDNA library may be successful. Unsuccessful screening may indicate that the target gene is refractory to cloning using this strategy. Possible alternative methods are described in Chapter 5.

In the second basic protocol, three to five spots are generally obtained on each 100-mm plate. It is possible to enrich a mixture that is 1:10,000 in the desired clone to 1:100 with a single immunoselection step.

Time Considerations

In the first basic protocol, the cloning strategy involves multiple steps. Each cycle of transfection, immunoselection, plasmid rescue, and amplification can be comfortably accommodated in 1 week. On this schedule, it is possible to screen a cDNA library in 1 month. However, the more ambitious can screen a library in 3 weeks.

In the second basic protocol, successful application allows identification of positive pools in 1½ weeks. The screening of subsets of the pool can be accelerated by using 60-mm plates. When screening individual clones, transfectants may be assayed by immunofluorescence. The length of time required to obtain a single isolated positive clone will depend on the pool sizes used.

Literature Cited

- Aruffo, A. and Seed, B. 1987. Molecular cloning of a CD28 cDNA by a high-efficiency COS cell expression system. *Proc. Natl. Acad. Sci. U.S.A.* 84:8753-8577.
- Dailey, L. and Basilico, C. 1985. Sequence in the polyomavirus DNA regulatory region involved in viral DNA replication and early gene expression. *J. Virol.* 54:739-749.

- Gearing, D.P., King, J.A., Gough, N.M., and Nicola, N.A. 1989. Expression cloning of a receptor for human granulocyte-macrophage colony-stimulating factor. *EMBO J.* 8:3667-3676.
- Gluzman, Y. 1981. SV40-Transformed simian cells support the replication of early SV40 mutants. *Cell* 23:175-182.
- Goelz, S.E., Hession, C., Goof, D., Griffiths, B., Tizard, R., Newman, B., Chi-Rosso, G., and Lobb, R. 1990. ELFT: A gene that directs the expression of an ELAM-1 ligand. *Cell* 63:1349-1356.
- Goldfarb, M., Schimizu, K., Perucho, M. and Wigler, M. 1982. Isolation and preliminary characterization of a human transforming gene from T24 bladder carcinoma cells. *Nature (Lond.)* 296:404-409.
- Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* 26:365-369.
- Holmes, K. and Fowlkes, B.J. 1991. Preparation of cells and reagents for flow cytometry. In *Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 5.3.1-5.3.11. Greene Publishing and John Wiley & Sons, New York.
- Kaufman, R.J. 1990. Overview of vectors used for expression in mammalian cells. *Methods Enzymol.* 185:487-511.
- Kavathas, P. and Herzenberg, L.A. 1983. Stable transformation of mouse L cells for human membrane T-cell differentiation antigens, HLA and 2-microglobulin: Selection by fluorescence-activated cell sorting. *Proc. Natl. Acad. Sci. U.S.A.* 80:524-528.
- Kavathas P., Sukhatme, V.P., Herzenberg, L.A., and Parnes, J.R. 1984. Isolation of the gene encoding the human T-lymphocyte differentiation antigen Leu-2 (T8) by gene transfer and cDNA subtraction. *Proc. Natl. Acad. Sci. U.S.A.* 81:7688-7692.
- Lee, F., Yokota, T., Otsuka, T., Gemmell, L., Larson, N., Luh, J., Arai, K.-I., and Rennick, D. 1985. Isolation of cDNA for a human granulocyte-macrophage colony-stimulating factor by functional expression in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 82:4360-4364.
- Littman, D.R., Thomas Y., Maddon, P.J., Chess, L., and Axel, R. 1985. The isolation and sequence of the gene encoding T8: A molecule defining functional classes of T lymphocytes. *Cell* 40:237-246.
- Maddon, P.J. Littman, D.R., Godfrey, M., Maddon, D.E., Chess, L., and Axel, R. 1985. The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: A new member of the immunoglobulin gene family. *Cell* 42:93-104.
- McCutchan, J.H. and Pagano, J.S. 1968. Enhancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethylaminoethyl-dextran. *J. Natl. Cancer Inst.* 40:351-357.
- McMahan, C.J., Slack, J.L., Mosley, B., Cosman, D., Lupton, S.D., Brunton, L.L., Grubin, C.E., Wignall, J.M., Jenkins, N.A., Brannan, C.I., Copeland, N.G., Huebner, L., Croce, C.M., Cannizzarro, L.A., Benjamin, D., Dower, S.K., Spriggs, M.K., and Sims, J.E. 1991. A novel IL-1 receptor, cloned from B cell by mammalian expression, is expressed in many cell types. *EMBO J.* 10:2821-2832.
- Metzelaar, M.J., Wijngaard, P.L.J., Peters, P.J., Sixma, J.J., Nieuwenhuis, H.K., and Clevers, H.C. 1991. CD63 antigen. *J. Biol. Chem.* 266:3239-3245.
- Muller, W.J., Naujokas, M.A., and Hassell, J.A. 1984. Isolation of large T antigen-producing mouse cell lines capable of supporting replication of polyomavirus-plasmid recombinants. *Mol. Cell. Biol.* 4:2406-2412.
- Munro, S. and Maniatis, T. 1989. Expression cloning of the murine interferon γ receptor cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 86:9248-9252.
- Sandri-Goldin, R.M., Goldin, A.L., Glorioso, J.C., and Levine, M. 1981. High-frequency transfer of cloned herpes simplex virus type I sequences to mammalian cells by protoplast fusion. *Mol. Cell. Biol.* 1:743-752.
- Seed, B. 1987. An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. *Nature (Lond.)* 329:840-842.
- Seed, B. and Aruffo, A. 1987. Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rapid immunoselection procedure. *Proc. Natl. Acad. Sci. U.S.A.* 84:3365-3369.
- Shih, C. and Weinberg, R.A. 1982. Isolation of a transforming sequence from a human bladder carcinoma cell line. *Cell* 29:161-169.
- Sims, J.E., March, C.J., Cosman, D., Widmer, M.B., MacDonald, H.R., McMahan, C.J., Grubin, C.E., Wignall, J.M., Jackson, J.L., Call, S.M., Friend, D., Alpert, A.R., Gillis, S., Urdal, D.L., and Dower, S.K. 1988. cDNA expression cloning of the IL-1 receptor, a member of the immunoglobulin superfamily. *Science* 241:585-589.
- Tsai, S.F., Martin, D.I., Zon, L.I., D'Andrea, A.D., Wong, G.G., and Orkin, S.H. 1989. Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression cloning. *Nature (Lond.)* 339:446-451.
- Wong, G.G., Witek, J.S., Tempel, P.A., Wilkens, K.M., Leary, A.C., Luxenberg, D.P., Jones, S.S., Brown, E.L., Kay, R.M., Orr, E.C., Shoemaker, C., Golde, D.W., Kaufman, R.J., Hewick, R.M., Wang, E.A., and Clark, S.C. 1985. Human GM-CSF: Molecular cloning of the complementary DNA and purification of the natural and recombinant proteins. *Science* 228:810-815.
- Wysocki, L.J. and Sato, V.L. 1978. Panning for lymphocytes: A method for cell selection. *Proc. Natl. Acad. Sci. U.S.A.* 75:2844-2848.

Yamasaki, K., Taga, T., Hirata, Y., Yawata, H., Kawanishi, Y., Seed, B., Taniguchi, T., Hirano, T., and Kishimoto, T. 1988. Cloning and expression of the human interleukin-6 (BSF-2/ IFN β 2) receptor. *Science* 241:825-828.

Key References

Aruffo and Seed, 1987; Seed and Aruffo, 1987. See above.

Contain original descriptions of cDNA library construction in CDM8 and isolation of cDNA clones encoding cell-surface antigens by expression cloning.

Metzelaar et al., 1991; Munro and Maniatis, 1989. See above.

Contain descriptions of growth of COS cells on wrap and screening for extracellular ligands.

Contributed by Diane Hollenbaugh and Alejandro Aruffo (cell-surface and intracellular antigens)

Bryan Jones and Peter Linsley
(intracellular antigens)
Bristol-Myers Squibb
Seattle, Washington

Recombination-Based Assay (RBA) for Screening Bacteriophage Lambda Libraries

The recombination-based assay represents a convenient way to screen a complex library constructed in bacteriophage λ for homology to a given sequence cloned into a specially designed plasmid. The technique serves to screen a bacteriophage library rapidly and efficiently with a sequence cloned into a plasmid; counterselection then yields the gene product of interest with its plasmid carrier deleted. Because 10^6 to 10^7 plaque-forming units (pfu) may be screened using several petri dishes, and the homology for crossing-over need only be >25 bp, the RBA represents an efficient way to screen complex λ libraries rapidly for homology to a given sequence.

In this procedure (outlined in Fig. 6.12.1), a λ library is screened using a specially designed R6K *supF* plasmid, pAD1 (Fig. 6.12.2), carrying the desired target sequence. Recombinants arising from cross-over events between the plasmid and a bacteriophage carrying a corresponding region of homology are selected by their ability to grow on strain DM21 (Fig. 6.12.3). Growth of λ on DM21 requires the presence of the *supF* allele encoded on the plasmid to suppress an amber mutation in the host strain that prevents λ propagation. Recovery of the original phage carrying the target sequence requires a reversal of the homologous recombination event. This reversal occurs spontaneously, and is detected by PCR amplification using primers that flank the cloning site in the λ vector (Fig. 6.12.4).

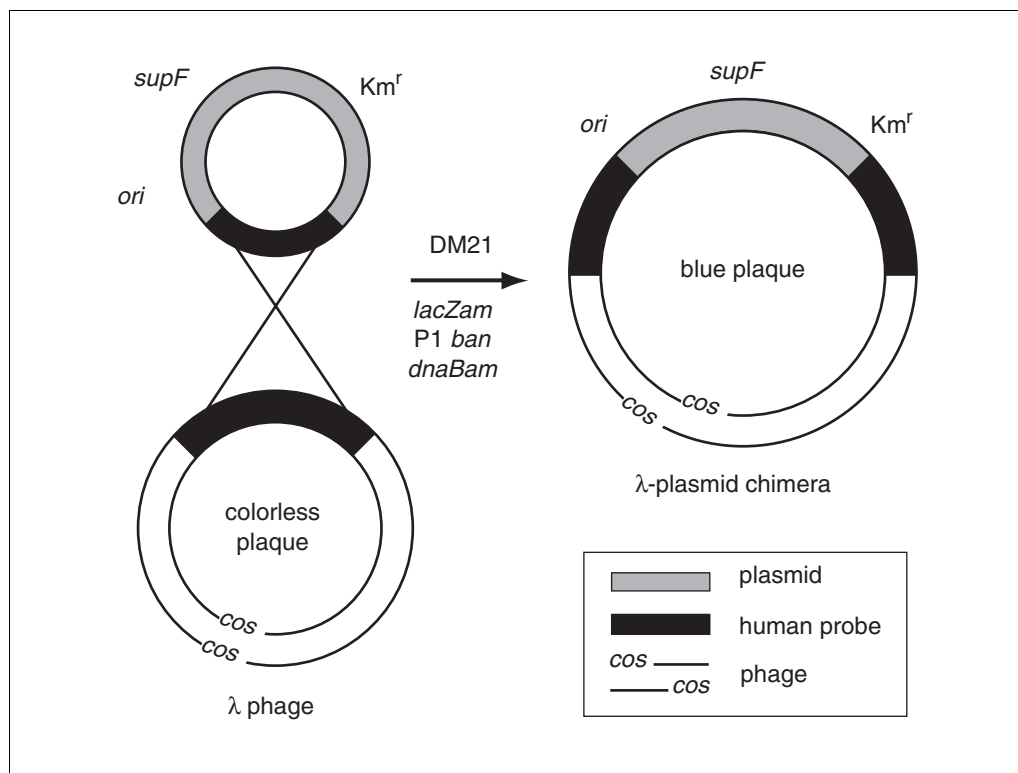


Figure 6.12.1 The recombination-based assay (RBA). Homology between sequences in a plasmid and a bacteriophage >25 bp long (Watt et al., 1985; Shen and Huang, 1986, 1989; King and Richardson, 1986) mediates a recombination event between the two vectors. As a result *supF* is integrated into the bacteriophage, allowing it to plate on the *dnaBam* host DM21 (see Table 6.12.1). The cointegrate yields a blue plaque in the presence of IPTG and Xgal on the *lacZam* host DM21, as *supF* suppresses the amber mutations in both the *dnaB* and *lacZ* genes. Different shadings indicate origins of DNA regions.

Contributed by David M. Kurnit

Current Protocols in Molecular Biology (1994) 6.12.1-6.12.12

Copyright © 2000 by John Wiley & Sons, Inc.

Materials

For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

DNA fragment encoding sequence of interest
Plasmid pAD1 (Fig. 6.12.2; available from Dr. D. Kurnit)
recA⁺ *E. coli* strain (Table 1.4.5 or commercial suppliers)
L broth (see recipe) with 50 µg/ml kanamycin (Table 1.4.1)
Bacteriophage λ library (UNIT 5.8)
Lambda top agar (see recipe)
Lambda plates (see recipe), some with 50 µg/ml kanamycin and some with 100 µg/ml streptomycin (Table 1.4.1)
Suspension medium (SM; see recipe)
Chloroform
E. coli DM21, DM75, DM392, and DM1061 (Fig. 6.12.3 and Table 6.12.1), saturated overnight cultures freshly grown in LB medium (UNIT 1.1) with 100 µg/ml streptomycin
100 mM IPTG (isopropyl thiogalactoside; Table 1.4.2)
2% Xgal in DMF (see recipe)
Additional reagents and equipment for subcloning DNA into plasmids (UNIT 3.16), culturing (UNIT 1.1) and transformation (UNIT 1.8) of bacteria, plating and titering λ phage (UNIT 1.11), β-galactosidase assay (UNIT 1.4), and PCR amplification (UNIT 15.1)

NOTE: All incubations are at 37°C unless otherwise specified.

Screen library and select recombinants

1. Clone the sequence of interest into a pAD1 plasmid and transform into *recA*⁺ *E. coli* strain yielding a kanamycin-resistant *recA*⁺ strain. Prepare a saturated overnight culture grown with aeration in L broth containing 50 µg/ml kanamycin.

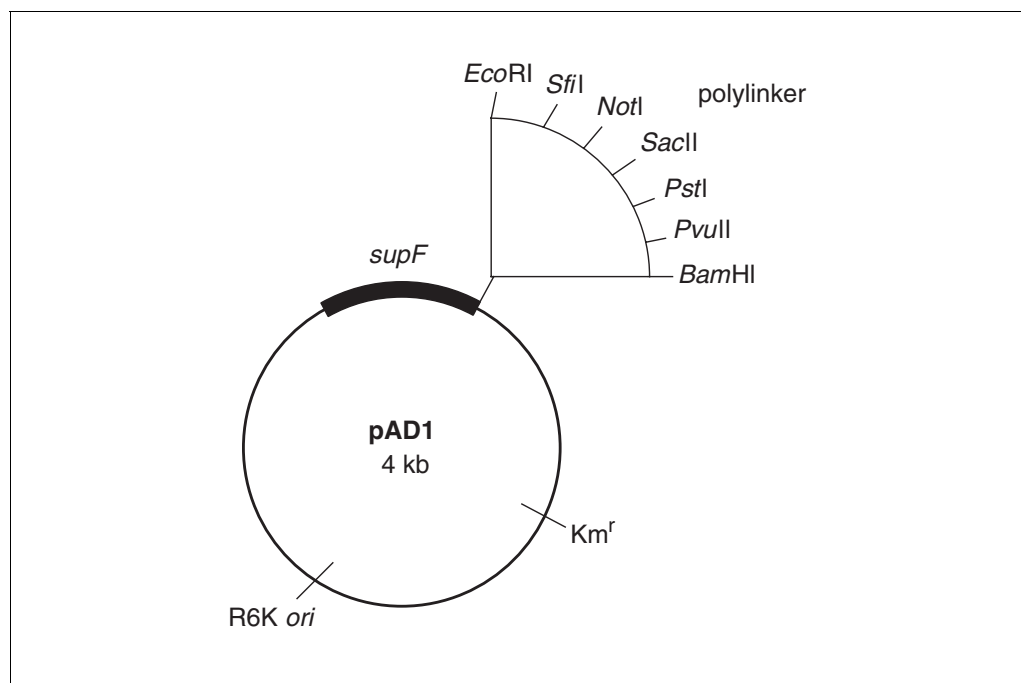


Figure 6.12.2 Structure of pAD1. This plasmid incorporates the R6K replicon, *Km*^r, *supF*, and a polylinker. It is not homologous to ColE1 plasmids.

- Mix 3 ml lambda top agar, 200 μ l of overnight culture, and 10^6 to 10^7 pfu of a bacteriophage λ library. Mix well and pour mixture onto a lambda/kanamycin plate. Incubate 7 hr to overnight until total lysis occurs.

If more convenient, incubation overnight is perfectly acceptable, because there is no need to harvest the plates just as lysis occurs.

- Add 3 ml SM and 0.5 ml chloroform to each plate. Swirl lightly. Incubate 2 hr to overnight at room temperature to allow the plates to elute.

SM and chloroform are immiscible; swirling them together ensures that the SM is saturated with chloroform, killing any eluted bacteria and minimizing phage adsorption to bacterial debris. The easiest method is to rotate a stack of plates slowly by hand after adding the liquid. Care should be taken not to get chloroform on the petri dish cover, as this can cause fusion of the cover and the plate bottom. If fusion occurs, the cover can be pried from the bottom (e.g., with a screwdriver).

- Using a nonsterile disposable transfer pipet, harvest the eluate from each plate into a 1.5-ml polypropylene microcentrifuge tube.

Although the transfer pipets are polyethylene, they hold chloroform-saturated SM for too short a time-span to be damaged by the solvent.

At this stage harvested eluates can be stored ≤ 1 week at 4°C before continuing the procedure.

- Add 50 μ l of eluate (5×10^8 to 1×10^9 pfu) to 200 μ l DM21 culture. Add 3 ml lambda top agar and pour mixture onto a lambda/streptomycin plate. Incubate 7 hr to overnight until plaques form.

*DM21 is selective (*dnaBam lacZam*) and resistant to streptomycin. DM75, DM1061, and DM392 (used in later steps) are also streptomycin-resistant, with growth and plating conditions identical to those for DM21.*

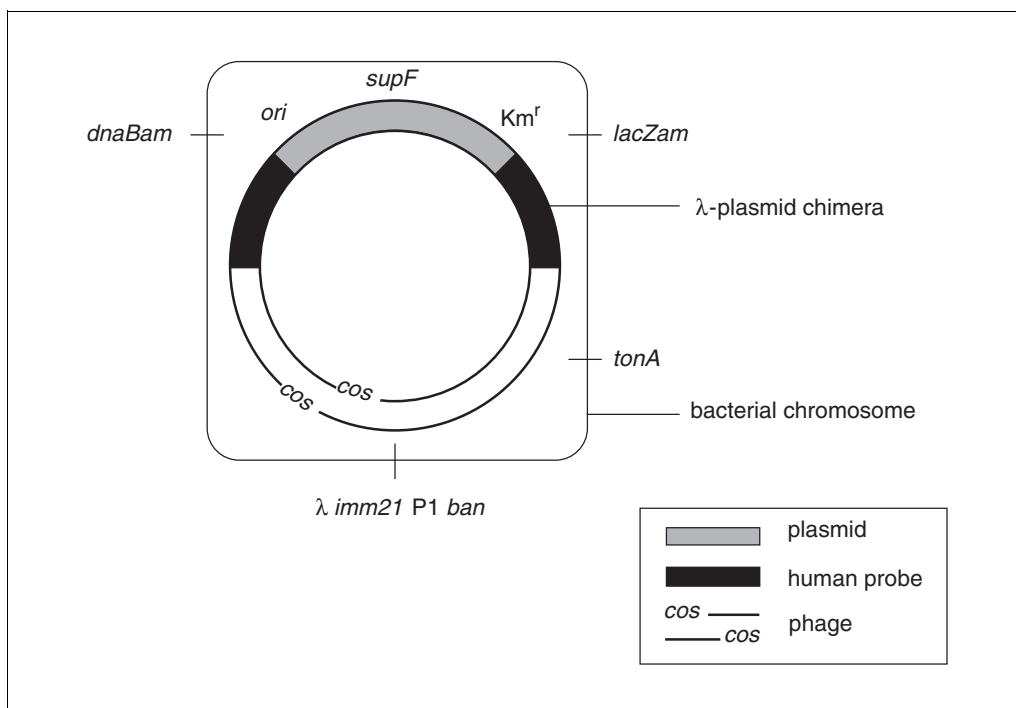


Figure 6.12.3 Bacterial strain DM21 (outer rectangle) containing λ plasmid chimera with *supF* integrated (inner circle). DM21 has the genotype *lacZ_{YA536}(am)*, *dnaB266(am)*, *Sm^r*, *hsdR⁺*, *hsdM⁺*, *tonA⁻* (λ *imm21 b515 b519 nin5 att⁺P1 ban*), *supO lacZ(am) dnaB(am)*. The *dnaB* amber allele selects for λ phage that have *supF* integrated as shown. *SupF* also suppress the *lacZ* amber mutation, yielding blue plaques. Different shadings indicate origins of DNA regions.

Titer eluates on permissive strain

6. Add 10 μl of each eluate to be titered to 990 μl SM to obtain a 1/100 dilution. Prepare a 100-fold dilution series (to 10^{-8}) in SM.

Several random eluates should be titered on the permissive (*supF*-bearing) strain DM392 to ensure that an appropriate number of phage have been added to the DM21 lawn.

7. Pour a lawn of DM392 (200 μl culture in 3 ml top agar) on a lambda/streptomycin plate. Drop 10- μl aliquots of each eluate dilution onto lawn. Dry 15 min in a forced-air hood (or for longer on bench or in incubator). Incubate 7 hr to overnight until total lysis occurs.

This drop-titer procedure is the most convenient method of titering the eluates.

8. Count plaques in the lowest dilution that yields plaques. Convert the result to pfu/ml by multiplying it by the appropriate dilution factor and by a factor of 10^2 .

Titration ensures that sufficient phage have been added to the DM21 lawn. Should too many be added (more pfu than cells), the lawn will not materialize due to lysis from without. This phenomenon occurs because every cell that is infected with a bacteriophage will die, even though only cells infected by a phage carrying *supF* will yield a productive burst that then goes on to infect other cells. In rare cases of lysis from

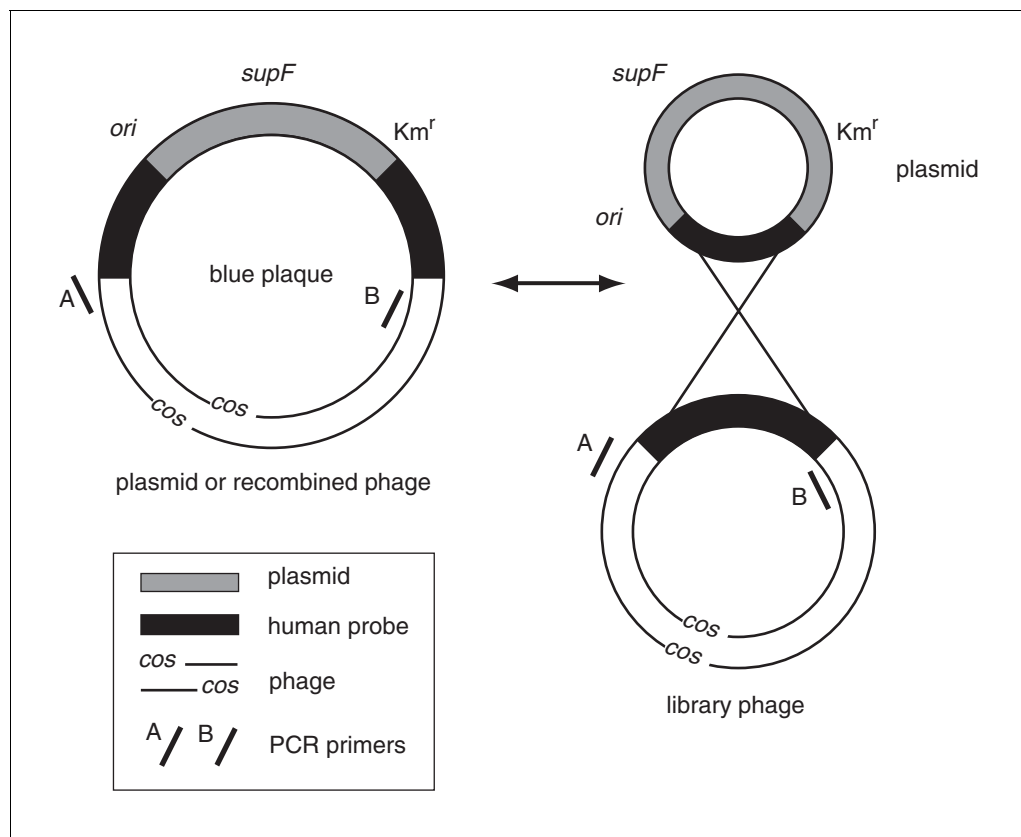


Figure 6.12.4 Counterselection. Reversal of the recombination event (which is an equilibrium event) occurs spontaneously. PCR using primers abutting the cloning site of the bacteriophage is employed preparatively to obtain the cDNA without the genomic sequence in pAD1 that was used to retrieve it. The cDNA insert + pAD1 + genomic insert is too large to be amplified by PCR; in contrast, the cDNA insert alone can be amplified. Because there is an equilibrium between the selected and the counterselected phage, the counterselected insert can be amplified directly from the selected blue plaque, which contains a mixture of the two phages. Different shadings indicate origins of DNA regions.

without, the plating should be repeated; either the eluate should be titered or less eluate used.

Plaques on DM21 are very small, because suppression of the *dnaBam* mutation (which is not fully efficient) is required for growth. This makes it difficult to confirm that *supF* is present via simultaneous suppression of the *lacZam* mutation by *supF*; therefore, phage must be transferred to another strain as described in the following steps.

Confirm phages have integrated *supF*

9. Elute plaques on DM21 (from step 5) into 100 μ l SM. Mix:

- 10 μ l eluate
- 200 μ l DM75 culture
- 3 ml lambda top agar
- 10 μ l 100 mM IPTG
- 100 μ l 2% Xgal in DMF.

Plate on lambda/streptomycin plates. Incubate 7 hr to overnight until total lysis occurs.

To mix water and DMF, the tubes of top agar must be inverted and righted several times, taking care not to create bubbles. It is best not to prepare more than several tubes at once, because cells do not tolerate the heating block for very long.

*Light blue plaques are the desired phage containing *supF*. A larger number of colorless plaques that have not integrated *supF* will also plate on this strain; these correspond to phage that were not adsorbed originally on DM21 and therefore remain viable. In addition, for a phage such as λ gt11, in which interruption of an intact *lacZ* gene serves as evidence of successful cloning, blue color can result from an intact *lacZ* gene in the phage. To differentiate between the two, note that the desired *supF* suppression of the single-copy chromosomal *lacZ* locus results in a light blue color that extends only to the plaque margins, whereas the high-copy-number *lacZ* gene on λ gt11 yields a dark blue halo that extends past the plaque margins.*

10. Elute each plaque thought to contain an integrated *supF* (from step 9) into 100 μ l SM. Pour lawns of DM75 and DM1061 (200 μ l culture in 3 ml top agar/IPTG/Xgal, as in previous step) onto separate lambda/streptomycin plates. Drop 10- μ l aliquots of each phage eluate onto a lawn of each strain. Incubate 7 hr to overnight until total lysis occurs.

*This serves to confirm that plaques result from phage with *supF* rather than *lacZ*. Phage with *supF* will be blue on DM75 (*lacZam*) but colorless on DM1061 (which contains a *lacZ* deletion), whereas phage carrying an intact *lacZ* gene will be blue on both strains.*

Counterselect with PCR

11. Pour a lawn of 200 μ l DM75 in 3 ml top agar onto lambda/streptomycin plate. Drop 10- μ l aliquots of phage eluate onto lawn. Incubate 7 hr to overnight, until a single large plaque (“macroplaque”) appears.

12. PCR amplify the cloned product from the macroplaque using primers that abut the *EcoRI* cloning site of the λ phage vector used to construct the library.

This reverses the selection process and accomplishes counterselection (see Fig. 6.12.4). Using the large macroplaque ensures that sufficient template is present.

*Because the recombination reaction is an equilibrium reaction, a small fraction of phage within a blue macroplaque represent colorless revertants that have excised the *pAD1* plasmid and its insert. In contrast, the major product in the macroplaque carries the phage insert, the plasmid, and the insert. Because this is too large to be amplified efficiently by PCR, the technique preferentially yields the desired genic insert from the phage without the unwanted plasmid and its insert.*

13. If desired, sequence the isolated genic clone (*UNITS 7.1-7.5*) and compare it to a database of known expressed sequences (*UNIT 7.7*) to obtain information about its possible significance, if available.

Repeatedly performing this protocol with different cDNA libraries allows determination of the timing of development and the tissue(s) in which the gene of interest is expressed. The latter can also be determined by using PCR primers specified by the sequence to see if amplification of different cDNA libraries occurs; given the sensitivity of this method, only cDNA library eluates, rather than DNA preparations, need be screened.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Lambda plates

10 g tryptone
5 g NaCl
13 g agar
3 ml 1 M MgCl₂
H₂O to 1 liter

Sterilize by autoclaving. Allow to cool until comfortable to touch. Add antibiotics as needed, mix gently to avoid bubbles, and pour plates. Store up to several months at 4°C.

Lambda top agar

10 g tryptone
5 g NaCl
8 g agar
3 ml 1 M MgCl₂
H₂O to 1 liter

Sterilize by autoclaving. Maintain ≤1 month molten at 60°C.

L broth

10 g tryptone
5 g NaCl
5 g yeast extract
5 g MgSO₄·7H₂O
1 g glucose
160 ml 12.5 M NaOH (to pH 7.2)
H₂O to 1 liter

Sterilize by autoclaving. Allow to cool until comfortable to touch. Add antibiotics as needed and mix. Store up to several months at 4°C.

Suspension medium (SM)

5.8 g NaCl
2 g MgSO₄·7H₂O
50 ml 1 M Tris·Cl, pH 7.5 (APPENDIX 2)
5 ml 2% (w/v) gelatin
H₂O to 1 liter

Sterilize by autoclaving. Store up to several months at 4°C.

Gelatin is prepared by adding 2 g gelatin to 100 ml H₂O, then autoclaving to dissolve when needed.

Xgal, 2% (v/v) in DMF

Dissolve 2% Xgal (5-bromo-4-chloro-3-indolyl-β-D-galactoside; see Table 1.4.2) in dimethylformamide (DMF). Place in polypropylene tube (not polystyrene, which will be dissolved by DMF), wrapped in aluminum foil. Store indefinitely at -20°C (solution will not freeze).

COMMENTARY

Background Information

The recombination-based assay (RBA) permits screening of a complex library or group of libraries with a given probe using only two petri dishes. As a result, the RBA is unparalleled in its efficiency and speed. The crux of the RBA is the insertion of a DNA fragment into a plasmid containing *supF*, followed by screening of a complex λ library (10⁶ to 10⁷ recombinants) for homology to the fragment. If such homology exists, a recombination event ensues between the inserts in the plasmid and homologous phage at a frequency of 10⁻² to 10⁻³ (see Fig. 6.12.1). As a result of this homology-mediated recombination event, the plasmid with *supF* is integrated into λ. Genetic selection for λ phage carrying the plasmid with *supF* results in the isolation of λ phage carrying an insert homologous to the insert in the plasmid. Given the high frequency of homologous recombination (10⁻² to 10⁻³), and the fact that 5 × 10⁸ to 10⁹ pfu can be plated onto a single petri dish, it is feasible to screen rapidly a λ library with a complexity of 10⁶ to 10⁷.

Bacterial host characteristics

This assay employs a bacterial strain, DM21 (see Figs. 6.12.1 and 6.12.3), that has been constructed to require the presence of *supF* in λ for phage propagation. As a result, sequences from a λ library that are homologous to a sequence cloned into the *supF*-bearing plasmid can be isolated on this strain. By screening a λ library carrying human genomic DNA sequences (Lawn et al., 1978), the copy number of a given sequence can be determined analytically. Plasmids carrying repetitive sequences rescue more phage clones from a human

genomic library than do plasmids carrying non-repetitive sequences (Neve and Kurmit, 1983). By screening a λ library corresponding to the genes encoded by a given tissue with single-copy sequences, the tissue and time in which a single-copy sequence is transcribed can be determined analytically. Selection for the desired *supF*-bearing phage is done using the *dnaB*/P1 *ban* balanced lethal system. In constructing the host, the *dnaB* unwinding protein that is normally essential for λ phage growth was replaced by the related, but not identical, P1 *ban* gene for *E. coli* growth. The resulting streptomycin-resistant *dnaBam* P1 *ban lacZam* host, DK21 (Kurmit and Seed, 1990), was then protected against a contaminating large (?T1) phage infection by a *tonA* mutation to yield the strain DM21 that is used in the protocol (the question mark notes characteristics that are likely but not definite). Analogously, strains LE392, LG75, and MC1061 have each been altered to carry a *tonA* mutation and resistance to streptomycin for use in the protocol, and have been renamed DM392, DM75, and DM1061, respectively (Table 6.12.1). DM21 selects for the plasmid-borne *supF* by requiring the suppression of an amber mutation in the *dnaB* gene to permit λ propagation. Furthermore, *supF* also suppresses the amber mutation in the *lacZ* gene of DM21, yielding a blue plaque upon addition of the chromogenic substrate Xgal in the presence of IPTG. This makes it possible to discard rare (<10⁻⁹) mutant λ phages (probably P-gene mutants) that lack *supF* but can be successfully plated on the *dnaBam* P1 *ban* host DM21, because these contaminating phages will yield colorless plaques.

Table 6.12.1 Bacterial Strains Used

Strain	Genotype ^a	Reference	Comment
DM21	<i>lacZ</i> _{YA536} (<i>am</i>), <i>dnaB266(am)</i> , <i>Sm</i> ^r , <i>hsdR</i> ⁺ , <i>hsdM</i> ⁺ , <i>?tonA</i> ⁻ (<i>λ imm21 b515 b519 nin5 att</i> ⁺ P1 <i>ban</i>)	Kurnit and Seed, 1990	<i>sup0 lacZam dnaBam</i>
DM75	<i>lacZ</i> _{YA536} (<i>am</i>), <i>Sm</i> ^r , <i>hsdR</i> ⁺ , <i>hsdM</i> ⁺ , <i>?tonA</i> ⁻	Guarente et al., 1980	<i>sup0 lacZam</i> ; strain used by the author and collaborators is <i>Sm</i> ^r , although the published genotype does not state this
DM392	<i>hsdR514 (hsdR</i> ⁻ , <i>hsdM</i> ⁺), <i>supE44</i> , <i>supF58</i> , <i>?lacY1</i> , <i>galK2</i> , <i>galT22</i> , <i>metB1</i> , <i>trpR55</i> , <i>Sm</i> ^r , <i>?tonA</i> ⁻	L. Enquist (unpub. observ.)	<i>sup</i> ⁺ ; made <i>Sm</i> ^r by the author and collaborators
DM1061	<i>araD139</i> , Δ (<i>ara</i> , <i>leu</i>)7697, Δ <i>lacX74</i> , <i>galU</i> ⁻ , <i>galK</i> ⁻ , <i>Sm</i> ^r , <i>hsdR</i> ⁻ , <i>hsdM</i> ⁺ , <i>mcrA</i> ⁻ , <i>mcrB</i> ⁻ , <i>?tonA</i> ⁻	Casadaban and Cohen, 1980	<i>sup0</i>

^aA question mark denotes characteristics that are likely, but not definite.

Counterselection

Regeneration of the phage as it existed before the recombination-based retrieval event requires deletion of the *supF*-bearing plasmid and its insert by reversal of the original homologous recombination event. Reversal of the selection event by PCR counterselection (Saiki et al., 1985; *UNIT 15.1*) allows preparative isolation of the transcribed sequence free of the sequence originally used to screen for it (Fig. 6.12.4). Fortunately, the frequency of this excision event is high in the presence of the *λ red* or *rap* genes: on the order of 10⁻² to 10⁻³ per generation where perfect homology exists between the genomic insert in the plasmid and the cDNA insert in the phage. Following selection on DM21 and amplification on DM75, counterselection is employed to delete the screening DNA sequence, leaving only the DNA sequence obtained by selection. This counterselection is achieved by PCR using primers that about the *EcoRI* cloning site in phage vectors (a specific set of primers is used for each phage vector; see Fig. 6.12.4). Each blue plaque contains both the selected sequence and a small proportion (~0.1%) of the counterselected sequence, which exists in equilibrium with the selected sequence (see Figs. 6.12.1 and 6.12.4). The selected sequence, which contains the R6K-derived *supF*-bearing pAD1 plasmid and its sequence as well as the (at least partially) homologous sequence in *λ*, is too large to be amplified efficiently by PCR, whereas the

shorter counterselected cDNA sequence can be amplified. Thus, PCR counterselection yields the desired cDNA sequence free of pAD1 and of the genomic sequence originally used to retrieve the cDNA sequence.

Another useful aspect of counterselection (Hanzlik et al., 1993) is that it distinguishes legitimate from illegitimate recombinants (Kurnit and Seed, 1990): legitimate recombinants will reverse the recombination reaction at a high frequency (10⁻²), whereas illegitimate recombinants will reverse the recombination reaction at a much lower frequency (10⁻⁹; Ikeda et al., 1982; Marvo et al., 1983). On an analytical basis, a rough indication of which category a recombinant belongs to is provided by the number of plaques arising during selection: the existence of multiple plaques (indicating multiple events) rules out rare nonhomologous events as the cause of recombination. This differential in reversal frequency allows recombination mediated by full homology to be distinguished from that mediated by partial or no homology, as the latter reverses at a much lower frequency due to mismatching or absence of matching (Watt et al., 1985; King and Richardson, 1986; Shen and Huang, 1986; 1989). When counterselection is performed, this large difference results in the isolation of PCR product in the case of recombination mediated by legitimate homology versus no product in the case of poorly-matched or illegitimate recombination. Thus, in addition to yield-

ing a desired clone, reversal of the recombination reaction (counterselection) allows one to distinguish legitimate from the rarer and undesired illegitimate recombinants.

Avoiding plasmid–phage library homology

For recombination-based screening, there can be no homology between the screening plasmid and the library to be screened lest this homology yield false positives. As a result, the plasmid used must have no homology to sequences present in a recombinant library to be screened. To permit recombination-based screening of common cDNA libraries that contain ColE1 sequences, *supF* has been inserted into a R6K plasmid origin of replication that is not homologous to ColE1. The ColE1 origin is present in a number of common plasmid vectors, including pBR– and pUC–derivatives and the author's π *supF* vectors (Bolivar et al., 1977; Seed, 1983; Yanisch-Perron et al., 1985; Lutz et al., 1987; Kurnit and Seed, 1990; Stewart et al., 1991). The R6K replicon chosen in this case lacks homology with ColE1 and therefore with λ libraries carrying ColE1 sequences (Poustka et al., 1984; Stewart et al., 1991).

ColE1 sequences are present in a variety of desirable cDNA libraries that have been constructed to date. By necessity, λ gt11 is propagated on a strain that contains *lac* sequences on the ColE1 replicon plasmid pMC9 to ensure repression of sequences downstream from the *lac* promoter in the λ phage. pMC9 thereby shares *lac* homology with λ gt11, fostering recombination between the two mediated by that shared homology. As a result, phage incorporate pMC9 at a low, but for these purposes appreciable, frequency of $\sim 10^{-4}$. Once internalized, the pMC9 integrated in the phage can recombine with ColE1-derived *supF*-bearing plasmids because they share homology at the ColE1 *ori*. This shared homology prevents the use of ColE1-derived plasmids for background-free recombination-based screening of λ gt11 libraries. Although theoretically λ gt10 libraries do not suffer this problem, many λ gt10 libraries do contain ColE1-derived sequences (Jankowski et al., 1990), indicating that these libraries have been passaged in strains intended for λ gt11 or that accidental contamination with ColE1 sequences has occurred. The vector pYAC4 contains pBR322 (ColE1) sequences, which can be used to clone end fragments. This prevents background-free screening by recombination with a probe cloned in a ColE1 origin vector of λ libraries made from pYAC4 recombinants. Furthermore, the ColE1 replicon in

phasmid vectors (i.e., recombinants arising from phage and plasmids) such as λ ZAP (Short et al., 1988) and CharonBS prevents recombination-based screening of libraries constructed in these vectors with inserts cloned in ColE1-based *supF*-bearing plasmids due to the shared ColE1 homology. To avoid these difficulties, we constructed a plasmid, pAD1 (Stewart et al., 1991; Fig. 6.12.2), based on the R6K replicon that is not homologous with ColE1 plasmids (Poustka et al., 1984). The sequences cloned in this plasmid may be used to screen all of the above λ libraries regardless of the presence of ColE1 sequences.

Other plasmid characteristics

Construction of the 4-kb plasmid pAD1 (Stewart et al., 1991) entailed cloning the R6K γ *ori*, *supF*, a kanamycin-resistance gene, and a polylinker containing sites for *Bam*HI, *Eco*RI, *Mlu*I, *Not*I, *Pst*I, *Pvu*II, *Sal*I, and *Sfi*I (Fig. 6.12.2). To confirm that the final vector, pAD1, indeed lacks homology to ColE1 sequences, it has been demonstrated by recombination (Seed, 1983; Kurnit and Seed, 1990) that this plasmid does not recombine with ColE1 DNA sequences. This lack of homology is consistent with the known sequence of R6K and with the finding that R6K-based cosmids do not recombine with ColE1-based plasmids (Poustka et al., 1984).

Recombination genes in host strain and phage

For recombination to function at a useful level of 10^{-2} to 10^{-3} if there is perfect homology, there must be genes promoting recombination in both the bacterial host and the bacteriophage. Thus, the bacterial host must be *recA*⁺ and the bacteriophage must be either *red*⁺ or *rap*⁺. In the case where the bacteriophage is neither *red*⁺ nor *rap*⁺, the *rap* gene can be supplied in *trans* from the pACYC-derived plasmid, pOM-PRAP2, constructed by Kurnit and Seed (1990). Most λ vectors are *rap*⁻, because *rap* lies within the *nin5* region deleted in most λ vectors.

Uses of the RBA

The RBA has special utility for two purposes:

1. *Isolating single-copy sequences.* A fragment whose copy number in the genome is to be elucidated is cloned into a plasmid with *supF*. Because cloning sequences into *supF*-bearing plasmids is also required for step (2) below, this procedure satisfies both require-

ments simultaneously. Sequence repetitiveness is assayed by the frequency with which a given insert in a *supF*-bearing plasmid mediates recombination between the plasmid and a recombinant bacteriophage library constructed from large random human genomic fragments (Neve et al., 1983; Neve and Kurnit, 1983). This author uses the library of Lawn et al. (1978) for this purpose, because it was constructed fortuitously (and unbeknownst at the time) in a Charon 4A λ vector that contains the $\phi 80$ *rap* gene (Kurnit and Seed, 1990). Although this older human genomic library is incomplete for single-copy sequences, this does not pose a problem because it is used merely to assay sequence repetitiveness, for which it is adequate. Subsequent human genomic libraries have been constructed in *red⁻ rap⁻* vectors, which cannot be screened without the awkward placement of a *rap* gene in *trans* (Kurnit and Seed, 1990). The amount of recombination correlates with the degree of repetitiveness in the genome, with the understandable and benign caveat that more highly repetitive sequences, which manifest significant mismatching, show some depression of the recombination frequency (Neve and Kurnit, 1983). This depression does not interfere with the ability of the assay to sort out the desired single-copy sequences; it merely results in the finding that *Alu* sequences (Rubin et al., 1980), which are actually reiterated 10^6 times in the genome, behave in the RBA as if they are repeated only 10^3 to 10^4 times (Neve et al., 1983; Neve and Kurnit, 1983). The salient point is that the methodology allows rapid analysis and isolation of sequences of a given copy number in the genome: "single" (1 to 10 copies), low-order-repeated (10 to 100 copies), and more highly repeated (>100 copies; Neve and Kurnit, 1983).

2. *Determining tissue- and time-specific transcriptional activity of single-copy fragments and isolating genes.* Gene libraries containing $>10^6$ independent recombinants are constructed: each corresponds to the totality of genes made in a given tissue at a given time in development. Screening a pool of 10^6 recombinants from a cDNA library requires only two petri dishes. The phage are first plated on a bacterial lawn carrying the sequence to be tested cloned in a *supF*-bearing plasmid. Following confluent lysis, 5×10^8 to 5×10^9 pfu are eluted and plated on DM21 to select for phage that have integrated *supF*. If no phage plaques are observed on DM21, this indicates that the sequence is not transcribed in the tissue

at the developmental stage present when the cDNA libraries were made. If plaques are observed on DM21, this indicates that the sequence is transcribed at that stage. The transcribed sequence is isolated free of the genomic sequence initially used to screen for it by reversing the recombination event (Fig. 6.12.4). In all the libraries used to date— λ gt10 (Huynh et al., 1985), λ gt11 (Young and Davis, 1983), and Sumo 15A (Kurachi et al., 1989)—the desired sequence is liberated as an *EcoRI* fragment that can be subcloned. As well as liberating the sequence, the reversal also makes it possible to discard rare nonhomologous (or imperfect) recombination events, which are identified by the fact that they reverse at the same low 10^{-9} frequency that they occur (for nonhomologous events) and at an intermediate frequency (for partially homologous events). In contrast, homologous recombination events, which can occur in a forward direction at a similar 10^{-8} frequency (assuming a worst case where a sequence is present only once per genome equivalent in a phage library of 10^6 recombinants, which is multiplied by a 10^{-2} chance of recombining if there is homology), reverse at a much higher 10^{-2} frequency. Thus, reversal of the recombination reaction will yield the cDNA free of the genomic sequence and will simultaneously allow rarer nonhomologous or partially homologous exchange events to be identified and discarded.

The RBA can be employed to determine the tissue and time of transcription of candidate genes discovered by other technologies as well as to obtain the gene of interest (in the form of the larger gene sequence that is transcribed). The technique is useful either alone or in combination with other methods for defining single-copy transcribed sequences. If DNA sequencing (as part of the genome initiative) or techniques to define transcribed sequences are used to identify genes, the RBA is still useful for determining the tissue and developmental timing of transcription, as well as for isolating a larger gene of interest. Technologies for defining transcribed sequences include exon trapping/amplification (Nisson and Watkins, 1994; Duyk et al., 1990; Buckler et al., 1991), use of somatic cell hybrids (Liu et al., 1989), and the use of hybridization-based schemes (Hochgeschwender and Brennan, 1994; Hochgeschwender et al., 1989; Kao and Yu, 1991), including hybrid selection (Lovett, 1994; Lovett et al., 1991; Parimoo et al., 1991). The RBA will proceed cooperatively, rather than competitively, with these other methods be-

cause it efficiently accomplishes two necessary tasks: identifying the timing and tissue of gene transcription and isolating a large transcribed sequence.

Critical Parameters

Plaque size is a major issue in this assay because plaques on the *dnaB am* strain DM21 are so small. Fresh λ plates should be used to maximize plaque size, because plaques will be smaller on older (drier) plates; likewise, it is important to plate cells on lambda plates, because plaques will be smaller on richer (e.g., LB) plates.

It is essential that there be no homology between the screening plasmid and sequences in the λ libraries (see Background Information). Therefore, screening should be performed solely with R6K *supF* plasmids, not with ColE1 *supF* plasmids.

Although titering all eluates would be too time-consuming, a few eluates should be titered to ensure that lysis and elution are occurring as expected. This is especially important because a lysed plate may vary from clear to grainy, rendering it difficult to determine visually whether complete lysis has occurred. Eluates should be saved until the DM21 plates have been scored as a precaution in case too many phage have been added, resulting in lysis from without. If this happens, the eluate may be titered or a lesser amount plated on DM21.

Anticipated Results

The abundance of sequences in screened λ libraries should be reflected in the number of phage that plate on DM21. Assuming a recombination rate of 1/500 (the exact number that will depend on the extent of homology), a sequence abundance of 1/10⁶ should yield one plaque on DM21 per 5 × 10⁸ phage plated. A higher abundance should yield a correspondingly greater number of plaques on DM21. If mismatching occurs in an interspersed "salt-and-pepper" manner (as for *Alu* sequences), recombination will be depressed (e.g., ~1000-fold for *Alu* sequences; Neve et al., 1983).

Time Considerations

The major advantage of the RBA is its rapidity: selection can be completed in four days using the following schedule. Day 1, grow bacterial cultures; day 2, add λ library and perform lysis; day 3, elute and plate on DM21; and day 4, identify plaques on DM21.

Counterselection takes an additional four days. One day is necessary for elution of

plaques from DM21 that are plated on DM75 with IPTG and Xgal in top agar. A second day is required for elution of putative light blue plaques and confirmatory macroplaque plating on DM75 and DM1061 with IPTG and Xgal. PCR counterselection of macroplaques that are blue on DM75 and colorless on DM1061 takes one day and a final day is necessary to isolate the counterselected PCR band from the gel.

Literature Cited

- Bolivar, F., Rodriguez, R., Green, P.J., Betlach, M., Heyneker, H.L., Boyer, H.W., Crosa, J., and Falkow, S. 1977. Construction and characterization of new cloning vehicles. *Gene* 2:95-113.
- Buckler, A.J., Chang, D.D., Graw, S.L., Brook, J.D., Haber, D.A., Sharp, P.A., and Housman, D.E. 1991. Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl. Acad. Sci. U.S.A.* 88:4005-4009.
- Casadaban, M.J. and Cohen, S.N. 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.* 138:179-207.
- Duyk, G.M., Kim, S., Myers, R.M., and Cox, D.R. 1990. Exon trapping: A genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. *Proc. Natl. Acad. Sci. U.S.A.* 87:8995-8999.
- Guarente, L., Lauer, G., Roberts, T.M., and Ptashne, M. 1980. Improved methods for maximizing expression of a cloned gene: A bacterium that synthesizes rabbit β -globin. *Cell* 20:543-553.
- Hanzlik, A.J., Hauser, M.A., Osemlak-Hanzlik, M.M., and Kurnit, D.M. 1993. The recombination-based assay demonstrates that the fragile X sequence is transcribed widely during development. *Nature Genet.* 3:44-48.
- Hochgeschwender, U. 1994. Identifying transcribed sequences in arrayed bacteriophage or cosmid libraries. In *Current Protocols in Human Genetics* (Dracopoli, N., Haines, J.L., Korf, B., Moir, D.T., Morton, C.M., Seidman, C.E., Seidman, J.G., and Smith, D.R., eds.) pp. 6.2.1-6.2.15. John Wiley & Sons, New York.
- Hochgeschwender, U., Sutcliffe, J.G., and Brennan, M.D. 1989. Construction and screening of a genomic library specific for mouse chromosome 16. *Proc. Natl. Acad. Sci. U.S.A.* 86:8482-8486.
- Huynh, T., Young, R.A., and Davis, R.W. 1985. Constructing and screening cDNA libraries in λ gt10 and λ gt11. In *DNA cloning*, Vol. II (D. Glover, ed.). IRL Press, Eynsham, U.K.
- Ikeda, H., Aoki, K., and Naito, A. 1982. Illegitimate recombination mediated in vitro by DNA gyrase of *Escherichia coli*: Structure of recombinant DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.* 79:3724-3728.
- Jankowski, S., Stewart, G.D., Buraczynska, M., Galt, J., Van Keuren, M., and Kurnit, D.M. 1990. Molecular approaches to trisomy 21. *Prog. Clin. Biol. Res.* 360:79-88.

- Kao, F.-T. and Yu, J.-W. 1991. Chromosome microdissection and cloning in human genome and genetic disease analysis. *Proc. Natl. Acad. Sci. U.S.A.* 88:1844-1848.
- King, S.R. and Richardson, J.P. 1986. Role of homology and pathway specificity for recombination between plasmids and bacteriophage λ . *Mol. Gen. Genet.* 204:141-147.
- Kurachi, S., Baldori, N., and Kurnit, D.M. 1989. Sumo 15A: A lambda plasmid that permits easy selection for and against cloned inserts. *Gene* 85:35-43.
- Kurnit, D.M. and Seed, B. 1990. Improved genetic selection for screening bacteriophage libraries by homologous recombination in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 87:3166-3169.
- Lawn, R.M., Fritsch, E.H., Parker, R.C., Blake, G., and Maniatis, T. 1978. The isolation and characterization of linked δ - and β -globin genes from a cloned library of human DNA. *Cell* 15:1157-1174.
- Liu, P., Legerski, R., and Siciliano, M.J. 1989. Isolation of human transcribed sequences from human-rodent somatic cell hybrids. *Science* 246:813-815.
- Lovett, M. 1994. Direct selection of cDNAs using genomic contigs. In *Current Protocols in Human Genetics* (Dracopoli, N., Haines, J.L., Korf, B., Moir, D.T., Morton, C.M., Seidman, C.E., Seidman, J.G., and Smith, D.R., eds.) pp. 6.3.1-6.3.15. John Wiley & Sons, New York.
- Lovett, M., Kere, J., and Hinton, L.M. 1991. Direct selection: A method for the isolation of cDNAs encoded by large genomic regions. *Proc. Natl. Acad. Sci. U.S.A.* 88:9628-9632.
- Lutz, C.T., Hollifield, W.C., Seed, B., Davie, J.M., and Huang, H.V. 1987. Syrinx 2A: An improved λ phage vector designed for screening DNA libraries by recombination in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 84:4379-4383.
- Marvo, S.L., King, S.R., and Jaskunas, S.R. 1983. Role of short regions of homology in intermolecular illegitimate recombination events. *Proc. Natl. Acad. Sci. U.S.A.* 80:2452-2456.
- Neve, R.L. and Kurnit, D.M. 1983. Comparison of sequence repetitiveness of human cDNA and genomic DNA using the miniplasmid vector piVX. *Gene* 23:355-367.
- Neve, R.L., Bruns, G.A.P., Dryja, T.P., and Kurnit, D.M. 1983. Retrieval of human DNA from rodent-human genomic libraries by a recombination process. *Gene* 23:343-354.
- Nisson, P.E. and Watkins, P.C. 1994. Isolation of exons from cloned DNA by exon trapping. In *Current Protocols in Human Genetics* (Dracopoli, N., Haines, J.L., Korf, B., Moir, D.T., Morton, C.M., Seidman, C.E., Seidman, J.G., and Smith, D.R., eds.) pp. 6.1.1-6.1.14. John Wiley & Sons, New York.
- Parimoo, S., Patanjali, S.R., Shukla, H., Chaplin, D.D., and Weissman, S.M. 1991. cDNA selection: Efficient PCR approach for the selection of cDNAs encoded in large chromosomal DNA fragments. *Proc. Natl. Acad. Sci. U.S.A.* 88:9623-9627.
- Poustka, A., Rackwitz, H.-R., Frischauf, A., Hohn, B., and Lehrach, H. 1984. Selective isolation of cosmid clones by homologous recombination in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 81:4129-4133.
- Rubin, C.M., Houck, C.M., Deininger, P.L., and Schmid, C.W. 1980. Partial nucleotide sequence of the 300 nucleotide interspersed repeated human DNA sequences. *Nature* 284:372-374.
- Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G., Erlich, H.A., and Arnheim, N. 1985. Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230:1350-1354.
- Seed, B. 1983. Purification of genomic sequences from bacteriophage libraries by recombination and selection in vivo. *Nucl. Acids Res.* 11:2427-2445.
- Shen, P. and Huang, H.V. 1986. Homologous recombination in *Escherichia coli*: Dependence on substrate length and homology. *Genetics* 112:441-457.
- Shen, P. and Huang, H.V. 1989. Effect of base pair mismatches on recombination via the *recBCD* pathway. *Mol. Gen. Genet.* 218:358-360.
- Short, J.M., Fernandez, J.M., Sorge, J.A., and Huse, W.D. 1988. λ ZAP: A bacteriophage λ expression vector with in vivo excision properties. *Nucl. Acids Res.* 16:7583-7599.
- Stewart, G.D., Hauser, M.A., Kang, H., McCann, D.P., Osemlak, M.M., Kurnit, D.M., and Hanzlik, A.J. 1991. Plasmids for recombination-based screening. *Gene* 106:97-101.
- Watt, V.M., Ingles, C.J., Urdea, M.S., and Rutter, W.J. 1985. Homology requirements for recombination in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 82:4768-4772.
- Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103-119.
- Young, R.A. and Davis, R.W. 1983. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. U.S.A.* 80:1194-1198.

Contributed by David M. Kurnit
University of Michigan Medical Center
Ann Arbor, Michigan

CHAPTER 7

DNA Sequencing

INTRODUCTION

For many recombinant DNA experiments, knowledge of a DNA sequence is a prerequisite for its further manipulation. DNA sequencing followed by computer-assisted searching for restriction endonuclease cleavage sites is often the fastest method of obtaining a detailed restriction map (*UNITS 3.1-3.3*). This information is particularly useful when vectors designed to overexpress proteins or to generate protein fusions are used to subclone a gene of interest (Chapter 16). Computer-assisted identification of protein-coding regions (open reading frames, or ORF) within the DNA sequence, followed by computer-assisted similarity searches of DNA and protein data bases, can lead to important insights about the function and structure of a cloned gene and its product (*UNIT 7.7*). In addition, the DNA sequence is a prerequisite for a detailed analysis of the 5' and 3' noncoding regulatory regions of a gene. DNA sequence information is essential for site-directed mutagenesis (*UNIT 8.1*). Small amounts of DNA sequence information (sequence tagged sites, or STS; or expressed sequence tags, or ESTs) are the basis of methods for mapping and ordering large DNA segments cloned into yeast or bacterial artificial chromosomes (CYACs; BACs) or cosmids (Olson et al., 1989; Stephens et al., 1990; Green and Olson, 1990; Adams et al., 1991; Shizuya et al., 1992). EST databases are extremely valuable in gene discovery.

DNA sequencing techniques are primarily based on electrophoretic procedures using high-resolution denaturing polyacrylamide gels. These so-called sequencing gels are capable of resolving single-stranded oligonucleotides up to 800 bases in length that differ in size by a single deoxynucleotide. In practice, for a given region to be sequenced, a set of labeled, single-stranded oligonucleotides is generated whose members have one fixed end, and which differ at the other end by each successive deoxynucleotide in the sequence. The key to determining the sequence of deoxynucleotides is to generate, in four separate enzymatic or chemical reactions, all oligonucleotides that terminate at the variable end in

A, T, G, or C. The oligonucleotide products of the four reactions are then resolved on adjacent lanes of a sequencing gel. Because all possible oligodeoxynucleotides are represented among the four lanes, the DNA sequence can be read directly from the four "ladders" of oligonucleotides as shown in Figure 7.0.1. In automated four-color fluorescent DNA sequencers, the oligonucleotide products terminated from each of the four bases (A, C, G, T) are run in a single lane and resolved on the basis that the DNA fragments ending with each of the four bases are labeled with four different fluorescent tags.

The practical limit on the amount of information that can be obtained from a set of sequencing reactions is the resolution of the sequencing gel (see *UNIT 7.6* for protocols on setting up and running sequencing gels). Current technology allows ~500 nucleotides of sequence information to be reliably obtained in one set of sequencing reactions, although more information (up to 800 nucleotides) is often obtained using an automated sequencer. Thus, if the region of DNA to be sequenced is <500 nucleotides, a single cloning into the appropriate vector is all that is usually necessary to produce a recombinant molecule that can easily be sequenced in a single set of reactions.

For a larger region of DNA, it is generally necessary to break a large fragment into smaller ones that are then individually sequenced. This can be done in a random or an ordered fashion. *UNIT 7.1* contains a discussion of strategies for sequencing large regions of DNA. Two protocols for subdividing large regions of DNA are provided in *UNIT 7.2*. These protocols are used to create a set of ordered, or nested, deletions for DNA sequencing using exonuclease III or *Bal* 31 nuclease.

The two methods that are widely used to determine DNA sequences, the enzymatic dideoxy method and the chemical method, differ primarily in the technique used to generate the ladder of oligonucleotides. In the enzymatic dideoxy sequencing method, a DNA polymerase is utilized to synthesize a labeled, complementary copy of a DNA template. In the

chemical sequencing method, a labeled DNA strand is subjected to a set of base-specific chemical reagents. These two techniques are further described below.

DIDEOXY (SANGER) SEQUENCING

Sequencing Method

The dideoxy or enzymatic method, originally developed by F. Sanger and co-workers (1977, 1980), utilizes a DNA polymerase to synthesize a complementary copy of a single-stranded DNA template. DNA polymerases cannot initiate DNA chains; rather, chain elongation occurs at the 3' end of a primer DNA that is annealed to "template" DNA (Fig. 7.0.2). The deoxynucleotide added to the growing primer chain is selected by base-pair matching to the template DNA. Chain growth involves the formation of a phosphodiester bridge between the 3' hydroxyl group at the growing end of the primer and the 5' phosphate group of the incoming deoxynucleotide (see Fig. A.1.7 for line

drawings of the nucleotides). Thus, overall chain growth is in the 5'→3' direction.

Dideoxy sequencing capitalizes on the ability of the DNA polymerase to use 2',3'-dideoxynucleoside triphosphates (ddNTPs) as substrates. When a ddNMP is incorporated at the 3' end of the growing primer chain, chain elongation is terminated at G, A, T, or C because the primer chain now lacks a 3' hydroxyl group. To generate the four sequencing ladders shown in Figure 7.0.1, only one of the four possible ddNTPs is included in each of the four reactions (see below). The ddNTP/dNTP ratio in each reaction is adjusted such that a portion of the elongating primer chains terminates at each occurrence of the base in the template DNA corresponding to the included complementary ddNTP. In this way, each of the four elongation reactions contains a population of extended primer chains, all of which have a fixed 5' end determined by the annealed primer and a variable 3' end terminating at a specific dideoxynucleotide.

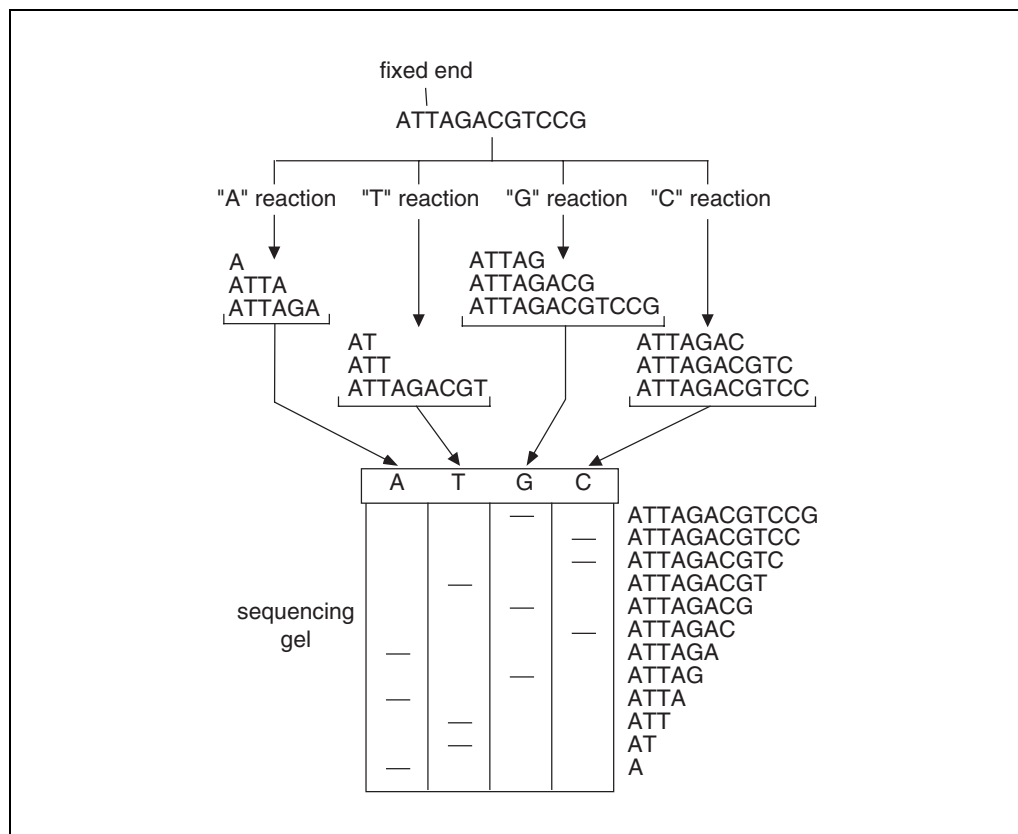


Figure 7.0.1 General strategy for DNA sequencing. To sequence a fragment of DNA, a set of radiolabeled single-stranded oligonucleotides is generated in four separate reactions. In each of the four reactions, the oligonucleotides have one fixed end and one end that terminates sequentially at each A, T, G, or C, respectively. The products of each reaction are fractionated by electrophoresis on adjacent lanes of a high-resolution polyacrylamide gel. After autoradiography, the DNA sequence can be "read" directly from the gel.

Two protocols for dideoxy sequencing using radiolabeled ddNTPs are provided in *UNIT 7.4A*. The original dideoxy method, which in this chapter is referred to as the Sanger procedure (Sanger et al., 1977, 1980), was developed for use with the large fragment of *E. coli* DNA polymerase I known as the Klenow fragment. A synthetic oligonucleotide primer is annealed to the 3' end of the region to be sequenced on a single-stranded DNA template (Fig. 7.0.2). The annealed template + primer is then divided into four reaction mixtures, each of which contains DNA polymerase, all four deoxynucleoside triphosphates (dNTPs)—one of which is radiolabeled—and one of the four ddNTPs (Fig. 7.0.2, right side). Under these conditions, the primer is extended and labeled until incorporation of a specific dideoxynucleotide causes termination. In a subsequent chase reaction, high concentrations of all four dNTPs are added so that all chains not terminated by a dideoxynucleotide will be elongated into high-molecular-weight DNA that remains unresolved at the top of the sequencing gel.

In the “labeling/termination” method, developed for use with modified T7 DNA polymerase (Sequenase; *UNIT 3.5*; Tabor et al., 1987; Tabor and Richardson, 1987a, 1987b, 1989a, 1989b, 1990), labeling of the primer and termination by incorporation of a dideoxynucleotide occur in two separate reactions (Fig. 7.0.2, left side). After annealing of the primer to the template, the labeling reaction occurs when the primer is elongated and labeled in the presence of a low concentration of all four dNTPs, one of which is radiolabeled. DNA synthesis continues until one or more of the nucleotide pools is exhausted, leading to almost complete incorporation of the labeled nucleotide. The termination step takes place in four separate reactions, each of which contains additional dNTPs and one of the four ddNTPs. In the termination step, a high concentration of dNTPs ensures processive DNA synthesis until the growing chains are terminated by the incorporation of a ddNMP.

In the Sanger procedure, the average length of the sequencing products is controlled by the ddNTP/dNTP ratio, where a higher ratio leads to shorter products. In the labeling/termination protocol, the average length of the sequencing products can be modulated either by the concentration of dNTPs in the labeling reaction (a higher concentration leads to longer products) or by the ddNTP/dNTP ratio in the termination reaction.

If Sequenase is used, the labeling/termination method is capable of yielding longer sequencing products, on average, than those obtained using the original Sanger protocol. Therefore, the labeling/termination method is advantageous for obtaining the maximum amount of sequence information per template. For applications where large amounts of sequence information are not needed (such as verifying constructions or sequencing small regions of DNA), the Sanger procedure is usually adequate. The Sanger procedure may be more reliable for obtaining the first few nucleotides of sequence information after the primer.

A variety of DNA polymerases are commercially available for sequencing. A description of these polymerases and their appropriate uses can be found in *UNIT 7.4A*. Thermostable DNA polymerases are the newest class of enzymes available for DNA sequencing. They are useful because they can carry out a sequencing reaction at high temperatures, allowing thermal cycling to increase the yield of sequencing fragments, and therefore enhanced detection sensitivity. This property provides a way of destabilizing secondary structures of the DNA template, which can interfere with the elongation reaction.

Vectors and Templates for Dideoxy Sequencing

Dideoxy sequencing requires a single-stranded template to which the primer can anneal. Single-stranded templates can be easily generated using specialized vectors derived from M13, an *E. coli* filamentous phage that contains a single-stranded circular DNA molecule (*UNITS 1.14 & 1.15*; Messing, 1983, 1988). Dideoxy sequencing can also be readily carried out using double-stranded DNA with a cycle sequencing procedure involving denaturation of the double-stranded DNA in each cycle (Chen and Seeburg, 1985; Haltiner et al., 1985; Zagursky et al., 1985; Hattori and Sakaki, 1986). Dideoxy sequencing of a double-stranded template is particularly useful when DNA sequencing is the only rapid method available for verifying a particular plasmid construction. For large-scale sequencing projects, the use of a single-stranded DNA vector system, such as the M13mp series, is recommended, because the preparation of high-quality generation of sequencing-quality, single-stranded DNA templates is somewhat more reliable than for plasmid DNA templates. However, plasmids offer the advantage of being able to obtain sequence from both the 5' and 3' ends.

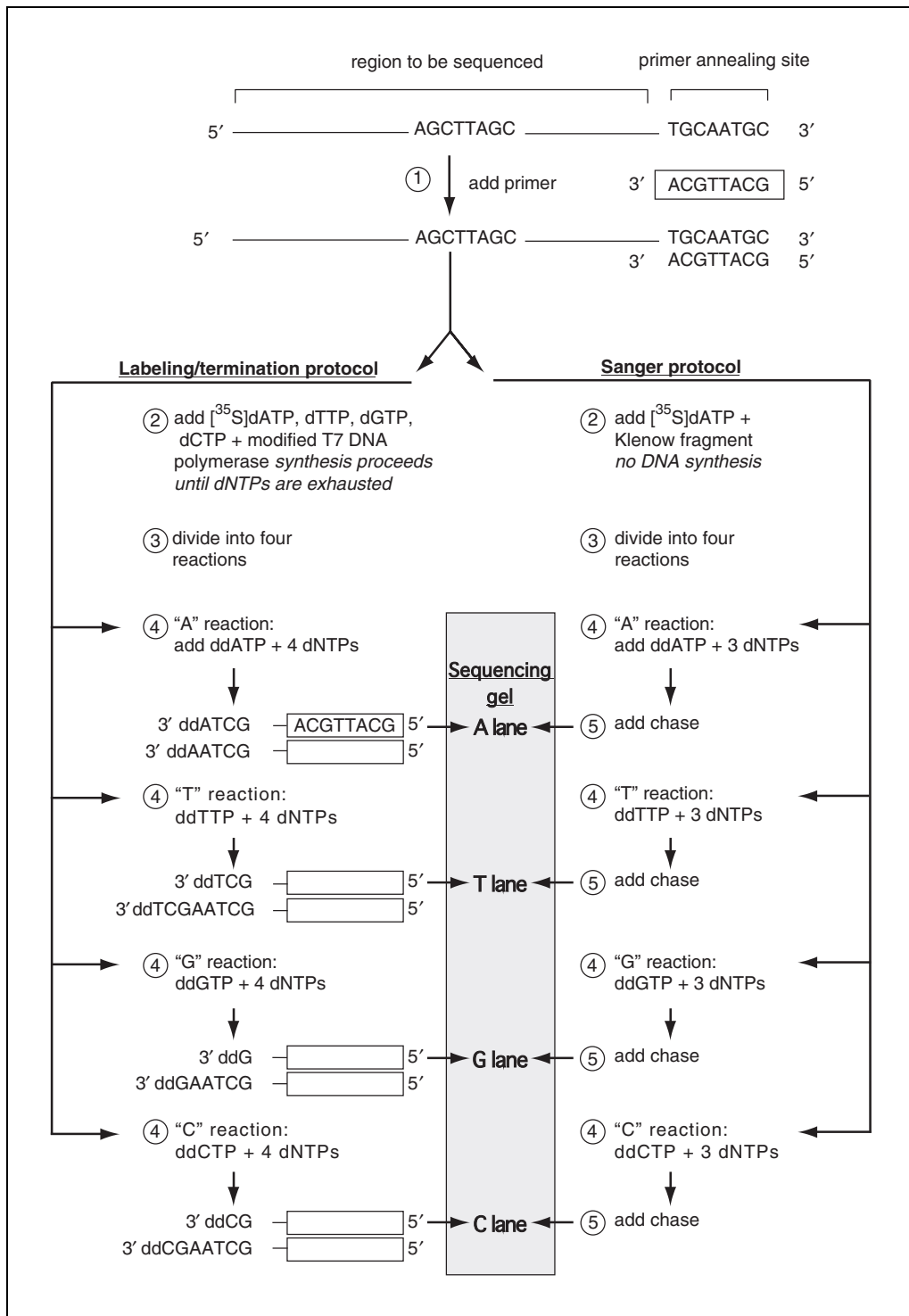


Figure 7.0.2 Dideoxy sequencing methods. In each method, a single-stranded DNA fragment is annealed to an oligonucleotide primer for polymerization (**step 1**). In the Sanger protocol (right side), the Klenow fragment and radiolabeled dATP are added (**step 2**). The reaction is divided into four aliquots (**step 3**) and the other three dNTPs and either ddATP, ddTTP, ddGTP, or ddCTP are added (**step 4**). DNA synthesis occurs until terminated by the incorporation of a ddNTP. A “chase” of all four dNTPs (**step 5**) elongates chains not terminated by a ddNMP into higher-molecular-weight DNA. In the labeling/termination protocol (left side), after the first step, a limiting amount of the four dNTPs—one of which is radiolabeled—and Sequenase are added (**step 2**). DNA synthesis proceeds until the dNTPs are exhausted. The reaction mix is divided into four aliquots (**step 3**) and all four dNTPs plus either ddATP, ddTTP, ddGTP, or ddCTP are added (**step 4**). Synthesis resumes, but termination specifically occurs when a ddNMP is incorporated. In each method, after the reactions are terminated, samples are loaded on adjacent lanes of a sequencing gel.

This feature is very valuable for genomic sequencing using randomly generated DNA fragments (the “shotgun” approach). A more detailed discussion of vectors used for dideoxy sequencing is provided in *UNIT 7.1*, and protocols for preparation of DNA templates derived from M13, plasmid, and bacteriophage λ vectors are provided in *UNIT 7.3*. The products of the polymerase chain reaction (PCR) can also be sequenced by the dideoxy method, and several protocols for generating these templates are provided in *UNITS 15.2 & 15.5*.

Radiolabels for Dideoxy Sequencing Reactions

$[\alpha\text{-}^{35}\text{S}]\text{dATP}$. The dideoxy sequencing protocols in *UNIT 7.4A* involve radiolabeling nascent DNA chains with $[\alpha\text{-}^{35}\text{S}]\text{dATP}$ rather than with $[\alpha\text{-}^{32}\text{P}]\text{dATP}$, for the following reasons. First, the low-energy β emissions of ^{35}S result in sharper autoradiographic bands compared to those generated by ^{32}P , allowing more sequence to be read from the upper portion of a gel. Second, the lower-energy emissions of ^{35}S cause fewer breaks in the sugar-phosphate backbone of the DNA. In practice, this means that ^{35}S -labeled reaction products can be stored at -20°C for several weeks without significant degradation; by contrast, ^{32}P products should be electrophoresed within a day. Third, users receive a lower radiation dose with ^{35}S than with ^{32}P .

$[\alpha\text{-}^{32}\text{P}]\text{dATP}$. In contrast to ^{35}S however, ^{32}P offers the advantage of short exposure times and is particularly useful in situations, such as verification of plasmid constructions, where maximizing resolution in the higher region of the sequencing gel is not a priority.

$[\alpha\text{-}^{33}\text{P}]\text{dATP}$. ^{33}P has a maximum β -emission energy that is 50% stronger than ^{35}S , but 5-fold weaker than ^{32}P . Sequences generated using $[\alpha\text{-}^{33}\text{P}]\text{dATP}$ have short exposure times like ^{32}P but band resolution comparable to that of ^{35}S (Zagursky et al., 1991).

5' end labeling. An alternative to labeling the nascent oligonucleotide with $[\alpha\text{-}^{35}\text{S}]\text{dATP}$ is to use a 5'-end-labeled primer generated by T4 polynucleotide kinase and $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ or $[\gamma\text{-}^{35}\text{S}]\text{ATP}$ (*UNIT 3.10*). Sequencing of large double-stranded DNA templates (such as $\lambda\text{gt}11$) with 5'-end-labeled primers has been found to give better results than standard labeling techniques. Protocols for sequencing using end-labeled primers are provided in *UNIT 7.4A*.

CHEMICAL (MAXAM-GILBERT) SEQUENCING

Sequencing Method

In the chemical method of DNA sequencing developed by A. Maxam and W. Gilbert (Maxam and Gilbert, 1977, 1980; Rubin and Schmid, 1980; Ambrose and Pless, 1987), the four sets of deoxyoligonucleotides are generated by subjecting a purified 3'- or 5'-end-labeled deoxyoligonucleotide to a base-specific chemical reagent that randomly cleaves DNA at one or two specific nucleotides. Because only end-labeled fragments are observed following autoradiography of the sequencing gel, four DNA ladders are observed as shown in Figure 7.0.3.

The Maxam and Gilbert chemical sequencing method is based on the ability of hydrazine, dimethyl sulfate (DMS), or formic acid to specifically modify bases within the DNA molecule. Piperidine is then added to catalyze strand breakage at these modified nucleotides. The specificity resides in the first reaction with hydrazine, DMS, or formic acid, which react with only a few percent of the bases. The second reaction, piperidine strand cleavage, must be quantitative. The chemical mechanisms of the first reactions are as follows:

G: DMS methylates nitrogen 7 of G, which then opens between carbon 8 and nitrogen 9. Piperidine then displaces the modified guanine from its sugar.

G+A: Formic acid weakens A and G glycosidic bonds by protonating purine-ring nitrogens. The purines can then be displaced by piperidine.

T+C: Hydrazine splits the rings of T and C. The fragments of these bases can then be displaced by piperidine.

C: In the presence of NaCl, only C reacts with hydrazine. The modified C can then be displaced with piperidine.

In all four reactions, piperidine also catalyzes phosphodiester bond cleavage at the position where the modified base has been displaced by piperidine.

Vectors for Chemical Sequencing

Chemical sequencing reactions can be performed on either single- or double-stranded DNA, as long as only one end is labeled. Specialized vectors have been developed (Eckert, 1987; Volckaert et al., 1984; Arnold and Puhler, 1988) that allow unique labeling of only one

strand of the cloned target DNA using *Tth*1111 or other restriction enzymes that have asymmetric recognition sites adjacent to the cloned DNA. These vectors are described more fully in *UNIT 7.1* and a protocol for their use is provided in *UNIT 7.5*.

CHOOSING BETWEEN DIDEOXY AND CHEMICAL SEQUENCING METHODS

As described above, the dideoxy chain-termination method is based on the ability of DNA polymerase to synthesize DNA 5'→3' from a defined primer annealed to the vector DNA at a site adjacent to the DNA being sequenced. Each reaction contains one of the four ddNTPs, which terminates synthesis selectively at G, A, T, or C.

Dideoxy sequencing is rapid. The primer-annealing and sequencing reactions can be completed within 60 to 90 min. A large number of single- or double-stranded samples can be

prepared for sequencing simultaneously. The method also offers excellent band resolution if ³⁵S-labeled nucleoside triphosphates are used to label the DNA.

The major disadvantage of dideoxy sequencing is that composition or secondary structure of the template can sometimes cause premature termination by DNA polymerase. Klenow fragment is more prone to this problem than are T7 DNA polymerase or the other alternative polymerases discussed in *UNIT 7.4A*. Despite these alternative polymerases, DNA is sometimes encountered that cannot be accurately sequenced by the dideoxy method.

The second method, chemical cleavage, is based on the ability of various chemicals to cleave DNA with a high specificity. The major advantage of the chemical method is that problems associated with polymerase synthesis of DNA (i.e., premature termination due to DNA sequence or structure) are eliminated, permitting sequencing of stretches of DNA that cannot

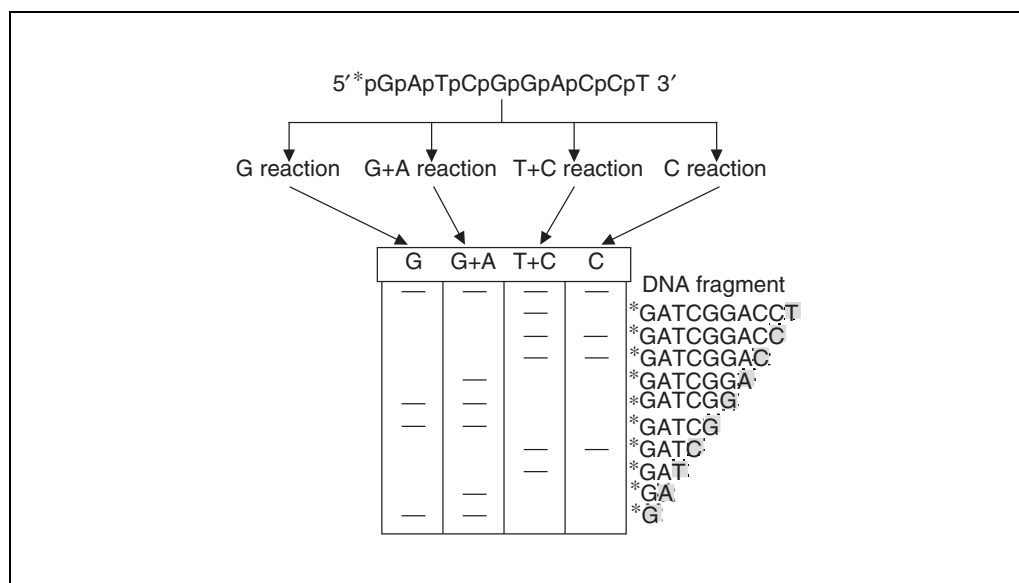


Figure 7.0.3 Chemical sequencing strategy. The ladder of oligonucleotides after gel electrophoresis of the products from the four chemical cleavage reactions is shown. The asterisk (*) indicates the position in the DNA fragment of the ³²P label, which is placed on the 5' end in this example. The direction of fragment migration is downward; smaller DNA oligonucleotides migrate faster in the sequencing gel than larger oligonucleotides. The shaded bases at the 3' end of the fragments to the right of the gel indicates bases that have been chemically modified and then displaced from the oligonucleotide during piperidine-mediated strand scission. For example, after a limited reaction with dimethyl sulfate (DMS), which is specific for G's, followed by quantitative release of the modified G residue by piperidine, a set of oligonucleotides are generated that terminate at the base immediately 5' of each G in the sequence. In this example, the oligonucleotide products are *pGpApTpCpGp and *pGpApTpCp. Each of these products forms a band in the G lane. For *pG, the product is *p, which would most likely run off the gel, making it difficult to determine the identity of the 5'-terminal base. Because formic acid is specific for purines (G's and A's), a fragment that terminates in G or A will produce a band in the G + A lane. Hydrazine in the absence of NaCl cleaves T's and C's resulting in a band in the T + C lane. Hydrazine in the presence of NaCl cleaves only C's; thus, a band is observed in the C lane.

be sequenced by the enzymatic method. In addition, obtaining the sequence of shorter regions of DNA using the chemical method does not require subcloning into an appropriate sequencing vector, such as is required for dideoxy sequencing. Finally, chemical cleavage is the only sequencing method available for small oligonucleotides.

Before the development of specialized cloning vectors, such as pSP64CS and pSP65CS, the major disadvantage of the chemical sequencing method was that preparation of the DNA prior to sequencing was very time-consuming. The fragment to be sequenced had to be end labeled and then the end of interest had to be isolated from all other labeled ends. This required several gel electrophoresis steps and often entailed a significant loss of the ^{32}P -labeled DNA fragment. However, vectors such as pSP64CS or pSP65CS allow direct subcloning of the DNA fragment of interest and end-labeling of a single predetermined terminus adjacent to the sequence, and they eliminate the time-consuming gel purification steps otherwise necessary. These vectors also make it possible to sequence a large number of samples simultaneously, because each recombinant plasmid to be sequenced is processed in a systematic manner (UNIT 7.1).

ALTERNATIVES TO RADIOLABELED SEQUENCING REACTIONS

Chemiluminescence

Chemiluminescence is a newer detection method that is comparable in sensitivity to traditional radiolabeling. Detection of the sequencing products occurs by a chemiluminescent reaction that can be monitored by autoradiography (UNIT 7.4B). A biotinylated primer is used in the dideoxy sequencing reactions. After electrophoresis of the biotinylated sequencing products on a sequencing gel, the products are transferred and cross-linked to a nylon membrane. The sequencing products are detected using streptavidin, biotinylated alkaline phosphatase, and a dioxetane substrate for alkaline phosphatase. The multivalent streptavidin cross-links the biotinylated sequencing product to the biotinylated alkaline phosphatase, immobilizing the phosphatase. Upon dephosphorylation by alkaline phosphatase, the dioxetane substrate emits light that is detected by autoradiography (Beck et al., 1989; Tizard et al., 1990; Creasey et al., 1991; Evans, 1991; Martin et al., 1991). Technology for end labeling DNA

fragments with biotin allows this detection method to also be used in chemical sequencing reactions (Tizard et al., 1990).

Alternatively, the dideoxy sequencing reactions can be carried out with a standard, nonbiotinylated primer without radioactivity. After electrophoresis, transfer, and cross-linking to a membrane, the sequencing products are hybridized with a biotinylated probe complementary to the primer before performing the detection described above. This method can be used with the products of chemical sequencing.

Multiplex Sequencing

Multiplex sequencing uses hybridization to a specific probe to detect an individual sequencing ladder in a mixture of ladders. In this method, sequencing products derived from a mixture of templates are subjected to electrophoresis on a sequencing gel, transferred to a membrane, and hybridized with a probe specific for one template. After hybridization and detection of the sequencing ladder derived from a single template, the probe is removed at a high temperature and the membrane is rehybridized to a different probe, complementary to an independent set of sequencing products that have been subjected to electrophoresis in the same lanes of the sequencing gel. Thus, the amount of sequence information available from one gel is multiplied by the number of times the membrane can be rehybridized (in practice, up to 20 times). Multiplex sequencing originally used radioactive probes and chemical sequencing technology (Church and Gilbert, 1984; Church and Kiefer-Higgins, 1988) but it is equally well adapted to chemiluminescent detection and/or dideoxy reactions.

A set of sequencing vectors is commercially available (Millipore) for multiplex sequencing using the dideoxy method. They have the identical priming region but contain unique sequences between the primer locus and the cloning site for the target DNA. These unique sequences are incorporated into the sequencing reaction products. Thus, sequencing reactions using several templates and the same primer can be performed in one set of reactions and later differentiated by successive hybridizations with probes for each unique sequence.

DEVELOPMENTS IN SEQUENCING TECHNOLOGY

Commercial Kits for Sequencing

Commercially available kits eliminate the need to assemble and calibrate numerous mixes

and can save a significant amount of startup time, although they are somewhat less flexible and can limit the ability to troubleshoot reactions when necessary. Nevertheless, these kits offer an excellent option for the novice at a reasonable cost. Kits are available for constructing nested deletions (see Table 7.2.1 for suppliers) as well as for dideoxy sequencing reactions using radiolabeling or chemiluminescent detection (Table 7.4.1) and for chemical sequencing using radiolabeling (UNIT 7.5).

Automated Sequencers and Automation of Sequencing Reactions

Automated sequencing machines automate the gel electrophoresis step, detection of DNA band pattern, and analysis of bands. Currently, all commercially available automated sequencers are designed for enzymatic sequencing reactions. The four sets of oligonucleotides generated by the sequencing reactions are loaded onto the gel manually and electrophoresis is then controlled automatically. Detection occurs at a point near the bottom edge of the gel in one of two ways. In one method, applicable to either fluorescently or radioactively labeled DNA, the bands of DNA moving sequentially past a detector are recorded. In the second method, the banding pattern of fluorescently labeled DNA is detected using an imaging camera. All automated sequencers possess data-collection capabilities, and either include further analysis programs or provide portability to external data-analysis software programs.

Fluorescent label can be incorporated into the sequencing products either through the primer or the ddNTPs. In the simplest system, a single fluorescently labeled primer is used and the four reaction products are run in separate lanes (Ansorge et al., 1987; Brumbaugh et al., 1988; Kambara et al., 1988; Middendorf et al., 1988; Fujita et al., 1990; Rosenthal et al., 1990; Sears et al., 1992). Other systems use either identical primers labeled with four different fluors (Smith et al., 1985, 1986; Connel et al., 1987; Johnston-Dow et al., 1987; Wilson et al., 1988; Mardis and Roe, 1989), or a different fluorescent tag for each ddNTP (Prober et al., 1987; Zagursky and McCormick, 1990). In these latter systems, because the primers or ddNTPs fluoresce at different wavelengths, all four reaction products are run in a single lane and the fragments terminating at the four different bases (A, C, G, T) are identified by four distinct emission colors of the fluorescent tags upon laser excitation. This quadruples the throughput compared to sequencing with radi-

olabels. Figure 7.0.4 shows a four-color sequencing gel image from the automated ABI 377 DNA sequencer. Among the various fluorescence detection methods, four-color DNA sequencing based on the Sanger dideoxy chain-termination method is now accepted as the technique of choice for large-scale sequencing projects.

Another promising approach is automation of the sequencing reaction by the use of robotics (Martin et al., 1985; Frank et al., 1988; Wilson et al., 1988; Mardis and Roe, 1989; Zimmerman et al., 1989; D’Cunha et al., 1990; Fujita et al., 1990; Smith et al., 1990). As this technology and that of automated sequencers is further developed, more rapid and less tedious sequence-data acquisition for large projects should be possible. For example, a fully automated sequencing system has been reported for genomic-scale sequencing (Hawkins et al., 1997).

Capillary Array DNA Sequencer

The most recent innovation in automated DNA sequencers utilizes capillary electrophoresis (CE; UNIT 2.8) and laser-induced fluorescence detection. CE using gel-filled narrow-bore (10- to 100- μm internal diameter) capillaries provides rapid, high-field, high-resolution separation of DNA fragments without heating artifacts. The use of multiple capillaries coupled with confocal laser-induced fluorescence for DNA sequencing was first reported in 1992 (Huang et al., 1992). Capillary array DNA sequencers (MegaBACE 1000 and ABI 3700) allow automated loading of 96 samples simultaneously as well as automated sample tracking and analysis (Khetarpal and Mathies, 1999). Currently, the separation matrix linear polyacrylamide (LPA) provides the best resolution and longer read-length in CE sequencers. Using LPA on CE sequencers with close to single-base resolution, read-length of 1000 bp with an electrophoresis time of 1 hr has been reported (Salas-Solano et al., 1998). This is an order-of-magnitude increase in speed compared to traditional slab gel sequencing. This new generation of automated DNA sequencer coupled with separation matrix development and improved fluorescent tags has the potential to substantially improve the throughput, speed, and overall process of larger sequencing projects.

Thermal Cycle Sequencing

Thermal cycle sequencing is a method of dideoxy sequencing in which a small number of template DNA molecules are repetitively

utilized to generate a sequencing ladder. A dideoxy sequencing reaction mixture (consisting of template, primer, dNTPs, ddNTPs, and a thermostable DNA polymerase) is subjected to repeated rounds of denaturation, annealing, and synthesis steps, similar to PCR (Chapter 15), using a commercially available thermal cycling machine. In this manner, linear amplification of the sequencing products occurs, allowing much less template DNA to be used than is usually required. In addition, thermal cycle sequencing eliminates the requirements for a separate annealing reaction preceding the sequencing reaction itself and for denaturing double-stranded DNA templates, and is compatible with automation processes. Various protocols have been developed for thermal cycle sequencing reactions (Applied Biosystems, 1989; Carothers et al., 1989; Murray, 1989; Adams and Blakesley, 1991; Craxton, 1991; Krishnan et al., 1991; Young and Blakesley, 1991; Sears et al., 1992); two such protocols are included in *UNIT 7.4A*. In addition, several thermal cycle sequencing kits are now commercially available utilizing each of the detection methods described above (Table 7.4A.1).

Energy Transfer Fluorescent Labeling Technology for Four-Color DNA Sequencing

Ideally, the fluorophores used for four-color DNA sequencing should have a similar high

molar absorbance at a common excitation wavelength as well as high fluorescence quantum yields, exhibit strong and well-separated fluorescence emissions, and introduce the same relative mobility shift. These criteria cannot be met optimally by the spectroscopic properties of single fluorescent dye molecules, and indeed were poorly satisfied by the fluorescent tags initially used for automated DNA sequencing. By exploiting resonance energy transfer (ET), the constraints imposed by the use of single dyes were overcome and fluorescent tags for DNA sequencing that met the performance criteria set out above were developed (Ju et al., 1995). Figure 7.0.5 shows how the energy transfer principle is utilized to construct ET primers that exhibit much enhanced fluorescence compared to single-dye fluorescent tags. The ET primer is labeled with two fluorophores that are separated by a spacer and coupled by resonance fluorescence energy transfer. The excitation energy that is captured by the fluorophore that absorbs at the shorter wavelength (the “donor”) is transferred to the longer wavelength-absorbing chromophore (the “acceptor”) located some distance away. This transfer results in a loss of donor fluorescence and the appearance of enhanced acceptor fluorescence emission even though the acceptor has only weak absorbance at the excitation wavelength. For resonance energy transfer to take place, the emission spectrum of the donor must

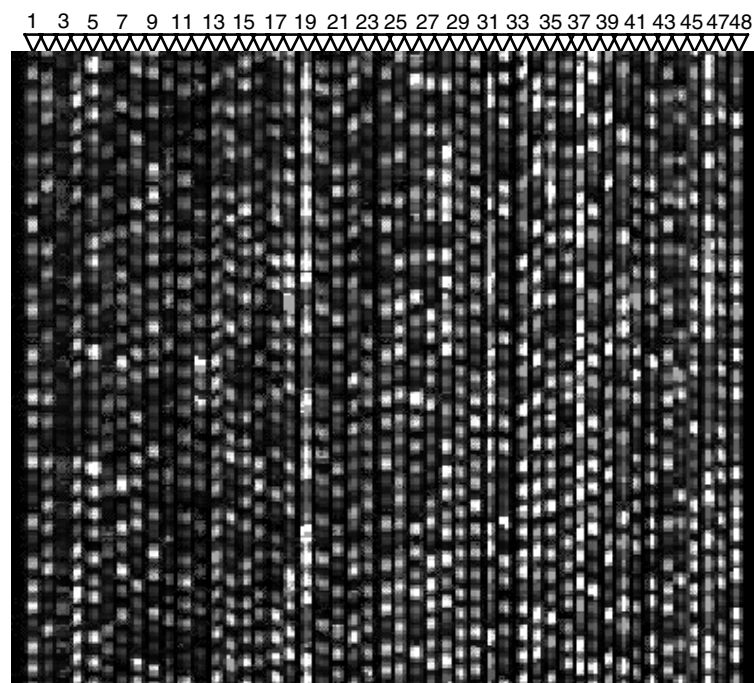


Figure 7.0.4 Gel image of 48 sequencing samples from the four-color ABI 377 automated DNA Sequencer. *This black and white facsimile of the figure is intended only as a placeholder; for full-color version of figure go to <http://www.currentprotocols.com/colorfigures>*

overlap the absorption spectrum of the acceptor. The efficiency of energy transfer is proportional to the inverse sixth power of the distance separating the donor and acceptor. Thus, the transfer efficiency and consequently the ratio of acceptor to donor emission can be controlled through changes in the spacing between the donor and acceptor. The ET dye-labeled primers and terminators can be efficiently excited at a common wavelength and exhibit strong and distinct fluorescent emissions (Ju et al., 1996; Lee et al., 1997). They are markedly superior to single dye-labeled primers and terminators for DNA sequencing and PCR fragment analysis. The ability to sequence directly from large-insert clones (>30 kb), such as BAC clones, is very important for closing the gaps in large genomic sequencing and mapping projects. Such large templates are very difficult to

sequence due to the smaller number of available priming sites (fewer molecules) compared with plasmid templates of the same mass. The higher sensitivity provided by primers/terminators tagged with ET dyes make it possible to sequence these large-template DNAs directly, which is a significant advantage for large-scale sequencing and mapping projects (Marra et al., 1996; Heiner et al., 1998). The sequencing protocol for large DNA templates are provided in *UNIT 7.4*

Solid-Phase Sequencing

Another recent innovation that is applicable to both manual and automated DNA sequencing is the use of solid-phase capture strategy to generate single-stranded DNA templates (Hultman et al., 1989, 1991; Jones et al., 1991; Kaneoka et al., 1991; Zimmerman et al., 1992).

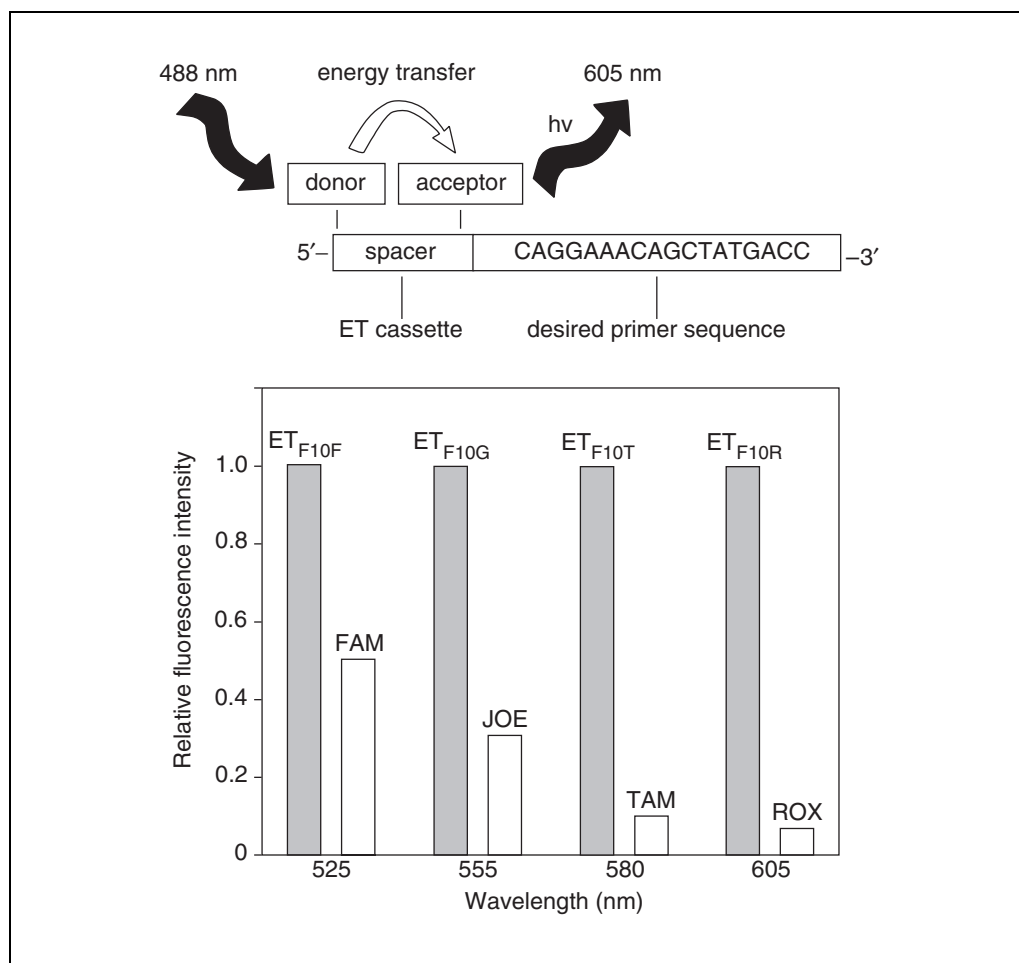


Figure 7.0.5 In ET primers, a common donor with a high absorbance at the excitation wavelength harvests energy and transmits it efficiently to acceptor fluorophores that emit in distinct wavelength regions. To avoid fluorescence quenching, the donor and acceptor are separated by a spacer that can be an oligonucleotide or other chemical functionality. The fluorescence emission intensity of current ET primers is 2- to 24-fold higher than that of conventional single dye-labeled primers, leading to high-quality DNA sequencing data.

In this approach, one strand of a double-stranded DNA molecule is biotinylated (e.g., by amplification using PCR in which one of the two primers is biotinylated; Chapter 15). The hemibiotinylated DNA molecule is then bound to streptavidin-ferromagnetic beads. The strands are denatured by treating the beads with alkali and the biotinylated strands are separated from the nonbiotinylated strands using a magnet that traps the bead complex to which the biotinylated strands are bound. Sequencing reactions can be performed using either the biotinylated strand-bead complex or the nonbiotinylated strand preparation as the template.

Fluorescent sequencing procedures (both dye primer and dye terminator methods) have disadvantages, most notably false termination and background noise. In the dye primer method, all the extended DNA fragments—including false-terminated fragments—from the primer carry a fluorescent dye, and thus all are detected by the fluorescent sequencer. This causes background noise and results in inaccurate sequencing data. In the dye terminator method, the excess dye-labeled dideoxynucleotides need to be cleaned up completely. Furthermore, if RNAs and nicked DNAs are present in the DNA templates, they will act as primers to generate false termination or high background noise. Thus, a DNA sequencing method that overcomes these disadvantages is desirable.

A sequencing chemistry using solid-phase-capturable dideoxynucleotides was recently developed that produces much cleaner sequencing data on both slab gel and capillary array sequencers, eliminating the disadvantages of current dye primer and dye terminator chemistries (Ju et al., 1997; Ju, 1999). The procedure involves coupling fluorescent ET primers that produce high fluorescent signals with solid phase-capturable terminators such as biotinylated dideoxynucleotides. After the sequencing reaction, the extension DNA fragments are captured with magnetic beads coated with streptavidin, while the other components in the sequencing reaction are washed away. Only the pure dideoxynucleotide terminated extension products are released from the magnetic beads and loaded on the sequencing gel, producing high-quality data.

Sequencing with Mass Spectrometry

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS; *UNITS 10.21 & 10.22*) has recently been explored for DNA sequencing (Koster et al.,

1996; Monforte et al., 1997; Fu et al., 1998). The Sanger dideoxy procedure is used to generate the DNA sequencing fragments and no labels are required. The fragments are mixed with a matrix compound, 3-hydroxypicolinic acid, forming microcrystals on a flat plate. The sample plate is then placed in the vacuum chamber of the mass spectrometer. Upon irradiation by a laser light, a microscopic portion of the matrix molecules desorbs off the plate, entraining the DNA fragments. In an electric field, the charged DNA molecules (ions) are accelerated toward the detector, which measures the mass of the fragments based on their time of flight, which is inversely proportional to their mass. Since the mass of each nucleotide (A, C, G, T) is different, the mass difference between adjacent peaks in the mass spectrum is used to establish the sequence of the DNA templates. For example, if the mass difference between the two adjacent peaks in a mass spectrum of an unknown DNA is the mass of “A” base, then the peak with the higher mass is identified as an “A”. Thus simple computation software is all that is needed to assemble the sequence of the DNA from its mass spectrum. Compared to gel-based sequencing systems, MS produces very high resolution of the sequencing fragments (sometimes as good as 1 Da) and extremely fast separation, on a time scale of microseconds. The high resolution allows accurate detection of mutations and heterozygosity.

Another advantage of sequencing with mass spectrometry is that the compressions associated with gel-based systems are completely eliminated. However, in order for accurate measurements of the masses of the sequencing DNA fragments to be obtained, the sample must be free from alkaline and alkaline earth salts. The samples must therefore be desalted before the MS analysis. Solid-phase procedures using either biotinylated primer or immobilized templates are generally used for desalting (Koster et al., 1996; Monforte and Becker, 1997). Both approaches introduce false-terminated DNA fragments into the mass detector. An elegant method to obtain pure sequencing fragments is to use solid phase-capturable dideoxynucleotides—such as biotinylated terminators—in the Sanger reactions to generate sequencing fragments. In this procedure, only the correctly terminated DNA fragments are isolated by streptavidin-coated beads, which are subsequently released and loaded on the mass spectrometer, resulting in accurate sequencing data (Ju et al., 1997; Ju, 1999). The current limit of mass spectrometry for sequenc-

ing is in the 100-bp range. Optimized solid-phase sequencing chemistry plus improvement in detector sensitivity for large DNA fragments will further improve mass spectrometry for DNA sequencing.

Sequencing by Hybridization

Sequencing by hybridization (SBH) makes use of an array of all possible short oligonucleotides to identify a segment of sequences present in an unknown DNA (Drmanac et al., 1989, 1993). This can be clearly explained by the following example. A pentanucleotide 5' CAGTA-3', with a complementary sequence of 5'-TACTG-3' is the sequence that need to be determined from a pool of all possible trinucleotides ($4^3 = 64$). This pentanucleotide will specifically hybridize only with the complementary trinucleotides TAC, ACT, and CTG, revealing the presence of these blocs in the complementary sequence. From this the sequence 5'-TACTG-3' can be reconstructed. Thus with a library of 8- to 10-mer oligonucleotides, much larger segments of DNA sequences can be established. Computational approaches are then used to assemble the complete sequence. In the current state of the art of this technology, SBH has been used for detecting mutations and for resequencing a genome as well as for detecting polymorphisms (Chee et al., 1996; Drmanac et al., 1998). Robust de novo sequencing has not yet been demonstrated. Potential applications of SBH include physical mapping (ordering) of overlapping DNA clones, sequence checking, DNA fingerprinting comparisons of normal and disease-causing genes, and identification of DNA fragments with particular sequence motifs in complementary DNA and genomic libraries.

COMPUTER ANALYSIS

Once the gels are run and autoradiograms are obtained, computer software is practically indispensable for analysis of the sequence information. Computer software can assist at three stages. First, DNA sequence data can be entered into a computer data base either by "reading" the sequencing gels manually with a digitizer system or by using an automated gel scanner. Second, several software packages are available for detecting overlaps in sequence data and then assembling contiguous DNA sequences (contigs) from individual templates. Third, computer assistance is indispensable for analyzing final sequence data, e.g., in finding open reading frames or finding homologies to other sequences present in the nucleotide (UNIT

19.2) or protein databases (UNIT 19.3). UNIT 7.7 provides an overview of software and technology currently available.

LITERATURE CITED

- Adams, M.D., Keller, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiaio, H., Merrill, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R., and Venter, J.C. 1991. Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science* 252:1651-1656.
- Adams, S. and Blakesley, R. 1991. Linear amplification sequencing. *Focus (BRL)* 13:56-57.
- Ambrose, B. and Pless, R. 1987. DNA sequencing: Chemical methods. *Methods Enzymol.* 52:522-538.
- Ansorge, W., Sproat, B., Stegemann, J., Schwager, C., and Zenke, M. 1987. Automated DNA sequencing: Ultrasensitive detection of fluorescent bands during electrophoresis. *Nucl. Acids Res.* 15:4593-4602.
- Applied Biosystems. 1989 to present. Model 370 and 370A Automated Sequencer User Bulletin. Applied Biosystems, Foster City, Calif.
- Arnold, W. and Puhler, A. 1988. A family of high copy number plasmid vectors with single end-label sites in rapid nucleotide sequencing. *Gene* 70:171-179.
- Beck, S., O'Keefe, T.O., Coull, J.M., and Koster, H. 1989. Chemiluminescent detection of DNA: Application for DNA sequencing and hybridization. *Nucl. Acids Res.* 17:5115-5123.
- Brumbaugh, J., Middendorf, L., Grone, D., and Ruth, J. 1988. Continuous on-line DNA sequencing using oligonucleotide primers with multiple fluorophores. *Proc. Natl. Acad. Sci. U.S.A.* 85:5610-5614.
- Carothers, A.M., Urlab, G., Mucha, J., Grunburger, D., and Chasin, L.A. 1989. Point mutation analysis in a mammalian gene: Rapid preparation of total RNA, PCR amplification of cDNA and *Taq* sequencing by a novel method. *BioTechniques* 7:494-499.
- Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X.C., Stern, D., Winkler, J., Lockhart, D.J., Morris, M.S., and Fodor, S.P. 1996. Accessing genetic information with high-density DNA arrays. *Science* 274:610-614.
- Chen, E.Y. and Seeburg, P.H. 1985. Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. *DNA (N.Y.)* 4:165-170.
- Church, G. and Gilbert, W. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 81:1991-1995.
- Church, G. and Kiefer-Higgins, S. 1988. Multiplex DNA sequencing. *Science* 240:185-188.
- Connel, C., Fung, S., Heiner, C., Bridgeham, J., Chakerian, V., Heron, E., Jones, B., Menchen, S., Mordan, W., Raff, M., Racknor, M., Smith, L., Springer, J., Woo, S., and Hunkapiller, M. 1987. Automated DNA sequence analysis. *BioTechniques* 5:342-348.

- Craxton, M. 1991. Linear amplification sequencing: A powerful method for sequencing DNA. *Methods, a companion to Methods Enzymol.* 3:20-26.
- Creasey, A., D'Angio, L.M., Dunne, T., Kissinger, C., O'Keefe, T., Perry-O'Keefe, H., Moran, L., Roskey, M., Shildkraut, I., Sears, L., and Slatko, B. 1991. Application of a novel chemiluminescent-based DNA detection method to single-vector and multiplex DNA sequencing. *BioTechniques* 11:102-109.
- D'Cunha, J., Berson, B.J., Brumly, R.L., Wagner, P.R., and Smith, L.M. 1990. An automated instrument for the performance of enzymatic DNA sequencing reactions. *BioTechniques* 9:80-90.
- Drmanac, R., Labat, I., Brukner, I., and Crkvenjakov, R. 1989. Sequencing of megabase-plus DNA by hybridization: Theory of the method. *Genomics* 4:114-128.
- Drmanac, R., Drmanac, S., Strezoska, Z., Paunesku, T., Labat, I., Zeremski, M., Snoddy, J., Funkhouser, W.K., Koop, B., Hood, L., et al. 1993. DNA sequence determination by hybridization: A strategy for efficient large-scale sequencing. *Science* 260:1649-1652.
- Drmanac, S., Kita, D., Labat, I., Hauser, B., Schmidt, C., Burczak, J.D., and Drmanac, R. 1998. Accurate sequencing by hybridization for DNA diagnostics and individual genomics. *Nature Biotechnol.* 16:54-58.
- Eckert, R. 1987. New vectors for rapid sequencing of DNA fragments by chemical degradation. *Gene* 51:245-252.
- Evans, S. 1991. Millipore's system speeds up DNA sequencing and eliminates radioactivity. *Genet. Eng. News* 14:29-41.
- Frank, R., Bosserhoff, A., Boulin, C., Epstein, A., Gausepohl, H., and Ashman, K. 1988. Automation of DNA sequencing reactions and related techniques: A workstation for micromanipulation of liquids. *BioTechnology* 6:1211-1213.
- Fu, D.J., Tang, K., Braun, A., Reuter, D., Darnhofer-Demar, B., Little, D.P., O'Donnell, M.J., Cantor, C.R., and Koster, H. 1998. Sequencing exons 5 to 8 of the p53 gene by MALDI-TOF mass spectrometry. *Nature Biotechnol.* 16:381-384.
- Fujita, M., Usui, S., Kiyama, M., Kambara, H., Murakawa, K., Suzuki, S., Sambe, H., and Takachi, K. 1990. Chemical robot for enzymatic reactions and extraction processes of DNA in DNA sequence analysis. *BioTechniques* 9:584-591.
- Green, E.D. and Olson, M.V. 1990. Chromosomal region of the cystic fibrosis gene in yeast artificial chromosomes: A model for human genome mapping. *Science* 250:92-98.
- Haltiner, M., Kempe, T., and Tjian, R. 1985. A novel strategy for constructing clustered point mutations. *Nucl. Acids Res.* 13:1015-1025.
- Hattori, M. and Sakaki, Y. 1986. Dideoxy DNA sequencing method using denatured plasmid templates. *Anal. Biochem.* 152:232-238.
- Hawkins, T.L., McKernan, K.J., Jacotot, L.B., MacKenzie, J.B., Richardson, P.M., and Lander, E.S. 1997. A magnetic attraction to high-throughput genomics. *Science* 276:1887-1889.
- Heiner, C.R., Hunkapiller, K.L., Chen, S.M., Glass, J.I., and Chen, E.Y. 1998. Sequencing multimegabase-template DNA with BigDye terminator chemistry. *Genome Res.* 8:557-561.
- Huang, X.C., Quesada, M.A., and Mathies, R.A. 1992. DNA sequencing using capillary array electrophoresis. *Anal. Chem.* 64:2149-2154.
- Hultman, T., Stahl, S., Hornes, E., and Uhlen, M. 1989. Direct solid phase sequencing of genomic and plasmid DNA using magnetic beads as solid support. *Nucl. Acids Res.* 17:4937-4946.
- Hultman, T., Bergh, S., Moks, T., and Uhlen, M. 1991. Bidirectional solid phase sequencing of in vitro amplified plasmid DNA. *BioTechniques* 10:84-93.
- Johnston-Dow, E., Mardis, E., Heiner, C., and Roe, B.A. 1987. Optimized methods for fluorescent and radiolabeled DNA sequencing. *BioTechniques* 5:754-765.
- Jones, D. S., Schofield, J.P., and Vaudin, M. 1991. Fluorescent and radioactive solid phase dideoxy sequencing of PCR products in microtitre plates. *J. DNA Seq. Map.* 1:279-283.
- Ju, J. 1999. Nucleic Acid Sequencing with Solid Phase Capturable Terminators. U.S. Patent no. 5,876,936.
- Ju, J., Ruan, C., Fuller, C.W., Glazer, A.N., and Mathies, R.A. 1995. Energy transfer fluorescent dye-labeled primers for DNA sequencing and analysis. *Proc. Natl. Acad. Sci. USA.* 92:4347-4351.
- Ju, J., Glazer, A.N., and Mathies, R.A. 1996. Energy transfer primers: A new fluorescence labeling paradigm for DNA sequencing and analysis. *Nature Med.* 2:246-249.
- Ju, J., Yan, H., Zaro, M., Doctolero, M., Goralski, T., Konrad, K., Lachenmeier, E., and Cathcart, R. 1997. DNA sequencing with solid phase capturable terminators *Microb. Comp. Genomics* 2:223.
- Kambara, H., Nishikawa, T., Katayama, K., and Yamaguchi, T. 1988. Optimization of parameters in a DNA sequenator using fluorescence detection. *BioTechnology* 6:816-821.
- Kaneoka, H., Lee, D.R., Hsu, K.-C., Sharp, G.C., and Hoffman, R.W. 1991. Solid phase DNA sequencing of allele specific polymerase chain reaction amplified HLA-DR genes. *BioTechniques* 10:30-40.
- Kheterpal, I. and Mathies, R.A. 1999. Capillary array electrophoresis DNA sequencing. *Anal. Chem.* 71:31A-37A.
- Koster, H., Tang, K., Fu, D.J., Braun, A., van den Boom, D., Smith, C.L., Cotter, R.J., and Cantor, C.R. 1996. A strategy for rapid and efficient DNA sequencing by mass spectrometry. *Nature Biotechnol.* 14:1123-1128.

- Krishnan, B.R., Blakesley, R.W., and Berg, D.E. 1991. Linear amplification DNA sequencing directly from single phage plaques and bacterial colonies. *Nucl. Acids Res.* 19:1153.
- Lee, L.G., Spurgeon, S.L., Heiner, C.R., Benson, S.C., Rosenblum, B.B., Menchen, S.M., Graham, R.J., Constantinescu, A., Upadhy, K.G., and Cassel, J.M. 1997. New energy transfer dyes for DNA sequencing. *Nucl. Acids Res.* 25:2816-2822.
- Mardis, E.R. and Roe, B.A. 1989. Automated methods for single-stranded DNA isolation and dideoxynucleotide DNA sequencing reactions on a robotic workstation. *BioTechniques* 7:840-850.
- Marra, M., Weinstock, L.A., and Mardis, E.R. 1996. End sequence determination from large insert clones using energy transfer fluorescent primers. *Genome Res.* 6:1118-1122.
- Martin, W., Warmington, J., Galinski, B., Gallager, M., Davies, R., Beck, M., and Oliver, S. 1985. A system to perform the Sanger dideoxy sequencing reactions. *BioTechnology* 3:911-915.
- Martin, C., Bresnick, L., Juo, R.-R., Voyta, J.C., and Bronstein, I. 1991. Improved chemiluminescence DNA sequencing. *BioTechniques* 11:110-113.
- Maxam, A.M. and Gilbert, W. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A.* 74:560-564.
- Maxam, A.M. and Gilbert, W. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499-559.
- Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* 101:20-78.
- Messing, J. 1988. M13, the universal primer and the polylinker. *Focus (BRL)* 10:21-26.
- Middendorf, L., Brumbaugh, J., Grone, D., Morgan, C., and Ruth, J. 1988. Large scale DNA sequencing. *Am. Biol. Lab.* (August) 14-22.
- Monforte, J.A. and Becker, C.H. 1997. High-throughput DNA analysis by time-of-flight mass spectrometry. *Nature Med.* 3:360-362.
- Murray, V. 1989. Improved double-stranded DNA sequencing using the linear polymerase chain reaction. *Nucl. Acids Res.* 17:8889.
- Olson, M., Hood, L., Cantor, C., and Botstein, D. 1989. A common language for physical mapping of the human genome. *Science* 245:1434-1435.
- Prober, J., Trainor, G., Dam, R., Hobbs, F., Robertson, C., Zagursky, R., Cocuzza, R., Jensen, M., and Baumeister, K. 1987. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* 238:336-341.
- Rosenthal, A., Sproat, B., Voss, H., Stegemann, J., Schwager, C., Erfle, H., Zimmerman, J., Courelle, C., and Ansorge, W. 1990. Automated sequencing of fluorescently labeled DNA by chemical degradation. *J. DNA Seq. Map.* 1:63-71.
- Rubin, J. and Schmid, C. 1980. Pyrimidine-specific chemical reactions useful for DNA sequencing. *Nucl. Acids Res.* 8:4613-4619.
- Salas-Solano, O., Carrilho, E., Kotler, L., Miller, A.W., Goetzinger, W., Susic, Z., and Karger, B.L. 1998. Routine DNA sequencing of 1000 bases in less than one hour by capillary electrophoresis with replaceable linear polyacrylamide solutions. *Anal. Chem.* 70:3996-4003.
- Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74:5463-5467.
- Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.M., and Roe, B.A. 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* 143:161-178.
- Sears, L., Moran, L., Kissinger, C., Creasey, T., Perry-O'Keefe, H., Roskey, M., Sutherland, E., and Slatko, B. 1992. Circumvent thermal cycle sequencing and alternative manual and automated DNA sequencing protocols using the highly thermostable Ventr (exo⁻) DNA polymerase. *BioTechniques.* 13:626-633.
- Shizuya, H., Birren, B., Kim, U.-J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. 1992. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Nat. Acad. Sci. U.S.A.* 89:8794-8797.
- Smith, L., Fung, S., Hunkapiller, M., Hunkapiller, T., and Hood, L. 1985. The synthesis of oligonucleotides containing an aliphatic amino group at the 5' terminus: Synthesis of fluorescent DNA primers for use in DNA sequencing analysis. *Nucl. Acids Res.* 13:2399-2412.
- Smith, L., Sanders, J., Kaiser, R., Hughes, P., Dodd, C., Heiner, C., Kent, S., and Hood, L. 1986. Fluorescence detection in automated DNA sequence analysis. *Nature* 321:674-679.
- Smith, V., Brown, C.M., Banker, A.T., and Barrell, B.G. 1990. Semiautomated preparation of DNA template for large scale sequencing projects. *J. DNA Seq. Map.* 1:73-78.
- Stephens, J.C., Cavanaugh, M.L., Gradie, M.I., Mador, M.L., and Kidd, K.K. 1990. Mapping the human genome: Current status. *Science* 250:237-244.
- Tabor, S. and Richardson, C.C. 1987a. Selective oxidation of the exonuclease domain of bacteriophage T7 DNA polymerase. *J. Biol. Chem.* 262:15330-15333.
- Tabor, S. and Richardson, C.C. 1987b. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 84:4767-4771.
- Tabor, S. and Richardson, C.C. 1989a. Selective inactivation of the exonuclease activity of bacteriophage T7 DNA polymerase by in vitro mutagenesis. *J. Biol. Chem.* 264:6447-6458.

- Tabor, S. and Richardson, C.C. 1989b. Effect of manganese ions on the incorporation of dideoxynucleotides by bacteriophage T7 DNA polymerase and *Escherichia coli* DNA polymerase I. *Proc. Natl. Acad. Sci. U.S.A.* 86:4076-4080.
- Tabor, S. and Richardson, C.C. 1990. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase: Effect of pyrophosphorylation and metal ions. *J. Biol. Chem.* 265:8322-8328.
- Tabor, S., Huber, H., and Richardson, C.C. 1987. *Escherichia coli* thioredoxin confers processivity of the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J. Biol. Chem.* 262:16212-16223.
- Tizard, R., Cate, R.L., Ramachandran, K.L., Wysek, M., Voyta, J.C., Murphy, O.J., and Bronstein, I. 1990. Imaging of DNA sequences with chemiluminescence. *Proc. Natl. Acad. Sci. U.S.A.* 87:4514-4518.
- Volckaert, G., de Vleeschouwer, E., Blocker, H., and Frank, R. 1984. A novel type of cloning vector for ultrarapid chemical degradation sequencing of DNA. *Gene Anal. Tech.* 1:52-59.
- Wilson, R., Yuen, A., Clark, S., Spence, C., Arakelian, P., and Hood, L. 1988. Automation of dideoxynucleotide DNA sequencing reactions using a robotic workstation. *BioTechniques* 6:776-787.
- Young, A. and Blakesley, R. 1991. Sequencing plasmids from single colonies with the dsDNA cycle sequencing system. *Focus (BRL)* 13:137.
- Zagursky, R. and McCormick, R. 1990. DNA sequencing separations in capillary gels on a modified commercial DNA sequencing instrument. *BioTechniques* 9:74-79.
- Zagursky, R., Baumeister, K., Lomax, N., and Berman, M. 1985. Rapid and easy sequencing of large double-stranded DNA and supercoiled plasmid DNA. *Gene Anal. Tech.* 2:89-94.
- Zagursky, R.J., Conway, P.S., and Kashdan, M.A. 1991. Use of ^{33}P for Sanger DNA sequencing. *BioTechniques* 11:36-38.
- Zimmerman, J., Voss, H., Schwager, C., Stegemann, J., Erfle, H., and Ansorge, W. 1989. Automated preparation and purification of M13 templates for DNA sequencing. *Methods Mol. Cell Biol.* 1:29-34.
- Zimmerman, J., Dietrich, T., Voss, H., Erfle, H., Schwager, C., Stegemann, J., Hewitt, N., and Ansorge, W. 1992. Fully automated Sanger sequencing protocol for double-stranded DNA. *Methods Mol. Cell Biol.* 3:39-42.

KEY REFERENCES

- Maxam and Gilbert, 1977. See above.
Describes the chemical cleavage method.
- Sanger et al., 1977. See above.
Describes the traditional Sanger procedure.
- Tabor and Richardson, 1987b. See above.
Describes the labeling/termination method.
- Sears et al., 1992. See above.
Describes the thermal cycle sequencing procedure.

Frederick M. Ausubel, Lisa M. Albright,
and Jingyue Ju

The chapter editors wish to acknowledge the substantial assistance of Richard L. Eckert of Case Western Reserve and Barton E. Slatko of New England Biolabs.

This unit contains a general discussion of factors that should be considered before embarking on a DNA sequencing project. In general, any sequencing strategy should include plans for sequencing both strands of the DNA fragment. Complementary strand confirmation leads to higher accuracy, especially when sequencing regions where artifacts such as “compressions” are a problem. Sequencing the opposite strand is often required to obtain accurate data for such regions.

The most commonly used methods for generating appropriately sized DNA fragments for dideoxy and chemical sequencing are discussed below. The biochemistry underlying these procedures, as well as how to choose between these and alternative sequencing methods, are discussed in the introduction to this chapter.

DIDEOXY SEQUENCING

Planning for Dideoxy Sequencing

Sequencing determination of a fragment of <500 nucleotides is usually straightforward because this amount of sequence information can reliably be determined from a single set of sequencing reactions. Fragments of this size can usually be subcloned directly into an appropriate single- or double-stranded DNA sequencing vector. If the vector generates single-stranded DNA, such as the M13mp vectors described below, the fragment should be cloned in both orientations so that both strands of the insert are produced as single-stranded DNA. A primer that hybridizes to a site on the vector adjacent to the insert DNA is then used to sequence both clones, generating the sequence of each strand. When sequencing double-stranded plasmid DNA, there are two options for obtaining the sequence of each strand. A single primer can be used if the insert DNA is cloned in both orientations. Alternatively, two primers that hybridize to plasmid sequences on opposite sides (and opposite strands) of the insert DNA can be used to sequence a single clone.

To sequence larger regions of DNA completely, it is generally necessary to subdivide a large fragment into smaller ones that can then be individually sequenced. Three general approaches are currently used. In the first approach, known as “shotgun cloning,” random

fragments are created from longer DNA fragments by physical shear, digestion by nucleases, (e.g., DNase I) or by restriction digests with endonucleases that make frequent cuts in the fragment (e.g., those with four-base recognition specificity; Frischauf et al., 1980; Anderson, 1981; Bankier and Barrell, 1983; Bankier et al., 1988; Hong, 1982; Messing, 1983, 1988; Deininger, 1983a, 1983b; Bankier, 1984; Lin et al., 1985). These fragments are combined and the entire pool is ligated into the appropriate sequencing vector. After the DNA sequence of the various cloned fragments has been determined, the final sequence is compiled by computer using overlapping information from the individual fragments (UNIT 7.7). However, with more complex (i.e., longer) sequences, this approach becomes tedious since it requires purifying, ligating, and cloning numerous individual fragments. In addition, finding the final few percent of a sequence by this procedure can consume a disproportionately large amount of time.

A second subcloning strategy for sequencing large DNA fragments is to generate an ordered set of subclones from a large DNA molecule. This is usually done by making progressive (nested) sets of deletions from a clone containing the entire DNA fragment to be sequenced. Numerous protocols exist for making nested deletions by enzymatic means; two such protocols using exonuclease III (Henikoff, 1984; Guo and Wu, 1982; Okita, 1985; Ozkaynak and Putney, 1987; Smith, 1979, 1980; Strauss and Zagurski, 1991) and nuclease *Bal* 31 (Guo and Wu, 1982; Guo et al., 1983; Vocke and Bastia, 1983; Yanisch-Perron et al., 1985; Poncz et al., 1982; Misra, 1985) are presented in UNIT 7.2. Another enzymatic method for making nested deletions utilizes T4 DNA polymerase (Dale et al., 1985). Other methods for isolating nested sets of deletion fragments include size-selection of inserts (Barnes, 1987; Barnes and Bevan, 1983; Barnes et al., 1983; Vocke and Bastia, 1983), oriented restriction fragment subcloning (Yanisch-Perron et al., 1985; Lee and Lee, 1989; Hoheisel and Pohl, 1986), transposon-mediated deletions (Ahmed, 1987a, 1987b; Nag et al., 1988; Peng and Wu, 1986), progressive synthesis (Burton et al., 1988; Liu and Hackett, 1989), and oligonucleotide-directed mutagenesis (Shen and Wayne, 1988; Wang et al., 1989). Commercial kits for

making nested deletions are currently available from a variety of sources (see Table 7.2.1 for a list of suppliers).

A third strategy for sequencing large DNA fragments, known as “primer walking,” is uniquely suited to dideoxy DNA sequencing and bypasses the need for subcloning smaller pieces of DNA. Initial sequence data is obtained using a vector-based primer. As new sequence is ascertained, an oligonucleotide is synthesized that hybridizes near the 3′ end of the newly obtained sequence and primes synthesis in a subsequent set of dideoxy reactions (Strauss et al., 1986). The widespread introduction of automated oligonucleotide synthesizers and advances in oligonucleotide purification techniques have made this approach widely accepted and cost effective. Primer walking is also effective as a way to fill gaps in sequence information left by the shotgun cloning or nested deletion approach.

Vectors for Dideoxy Sequencing

A wide variety of vectors are suitable for preparing template DNA for dideoxy sequencing. The decision as to vector choice is subject to the following general constraints:

1. Protocols involving construction of nested deletions generally require particular restriction endonuclease sites in the vector that permit generation of deletions or subsequent cloning of deletion products. For instance, the *exo III* deletion protocol (UNIT 7.2) generally requires the presence of restriction endonuclease sites which allow the generation of a 3′ overhang and a 5′ overhang or blunt end at either end of a DNA fragment.

2. For deletion strategies that require subcloning of deletion products into a secondary vector, such as the *Bal 31* procedure (UNIT 7.2), it is essential that the secondary vector utilize a positive screening method for inserts. For example, many vectors utilize a blue-versus-white screen for inserts into the *E. coli lacZ* gene using Xgal and α -complementation of β -galactosidase (UNIT 1.4 and below).

3. Synthetic oligonucleotides suitable for priming dideoxy sequencing reactions must be available either commercially or through custom synthesis. Primers are also commercially available for many vectors that are not used primarily for sequencing.

4. For large sequencing projects, it is most efficient to use a vector designed to produce single-stranded DNA, such as the phage-based M13mp vectors described below. For projects that require a plasmid-based vector (usually for

reasons other than sequencing), it is often still advantageous to have the capability to generate single-stranded DNA for sequencing templates. Many plasmid vectors contain an M13, *f1*, or *fd* origin of replication that can be activated for production of single-stranded DNA using helper phage (UNIT 1.14 and below).

5. If the sequencing strategy involves sequencing double-stranded templates, a high-copy-number plasmid is advantageous for sequencing minipreps of plasmid DNA.

M13 vectors. The dideoxy sequencing method has been greatly facilitated by the development of the filamentous *E. coli* phage M13 as a cloning vector (Sanger et al., 1980; Messing, 1988). Each M13 phage contains a single-stranded circular DNA molecule that replicates as a double-stranded circular replicative form (RF; UNIT 1.14). The RF can be isolated as a plasmid which can then be transformed into competent *E. coli* cells to yield either phage, which contain single-stranded DNA molecules, or additional double-stranded circular RF molecules. Therefore, a DNA fragment of interest can be cloned into the RF of M13 and then single-stranded DNA for sequencing can be readily produced in large quantities.

The M13 vectors most widely used for dideoxy sequencing are a series called M13mp constructed by J. Messing and his collaborators (Messing and Vieira, 1982; Messing, 1983, 1988; Norrander et al., 1983; Yanisch-Perron et al., 1985). The M13mp series contains the *lacZ* promoter and a partial *lacZ* gene, encoding the α -fragment of β -galactosidase (UNIT 1.14). After infection of an *E. coli* F′ host containing another partial *lacZ* gene encoding the ω -fragment of β -galactosidase and induction by IPTG, M13mp phage produce blue plaques on Xgal agar (see UNIT 1.4 for a discussion of α -complementation and the use of Xgal for assaying β -galactosidase). In addition, each vector in the M13mp series contains a synthetic polylinker inserted into the fifth codon of *lacZ*. Because these polylinkers do not alter the *lacZ* reading frame, M13mp vectors still produce blue plaques on Xgal agar. However, insertion of a DNA fragment into one of the unique polylinker cloning sites has a high probability of disrupting the *lacZ* reading frame, thus generating a recombinant phage that produces colorless plaques.

The more recent members of the M13mp series come as pairs with the polylinker inserted in each orientation. Using these vectors, restriction fragments with different ends (generated by restriction endonucleases whose sites are

present in the polylinker) can be readily cloned in either orientation with respect to the *lacZ* gene (i.e., forced cloning). This procedure permits preparation of opposite strands of the insert from the two recombinant phages, a feature that is particularly important for dideoxy sequencing.

Thus, the use of the M13mp vectors for sequencing involves the following steps. A DNA fragment of interest is ligated into the polylinker in opposite orientations or into an appropriate pair of M13mp vectors. Phage that produce white plaques are picked and used individually to infect a culture of *E. coli* F' cells. Following an incubation period, the supernatant containing the recombinant M13 phage is collected and single-stranded phage DNA is purified for use directly as template DNA in dideoxy sequencing. Because the polylinker is inserted into the same site in *lacZ* in all M13mp

derivatives, a synthetic oligonucleotide—complementary to a region of *lacZ* just downstream on the 3' side of the polylinker—is used as a “universal” primer for all sequencing reactions. If the insert is cloned in both orientations, both strands can be sequenced from the same primer site. It is routine to sequence 300 to 400 nucleotides in M13 vectors in a single set of reactions. The M13mp series of vectors are amenable to all deletion strategies described above. For dideoxy sequencing using single-stranded DNA templates we recommend M13mp18 or M13mp19 (see Table 7.1.1 for suppliers).

Plasmid sequencing vectors. Numerous plasmid vectors, most available commercially, can be used for double-stranded dideoxy sequencing (see Table 7.1.1). Each of these plasmids replicates to a high copy number in *E. coli*, each features a blue-versus-white screen for inserts, and for each plasmid, primers are com-

Table 7.1.1 Dideoxy Sequencing Vectors

Vector ^a	Supplier ^b	Vector ^a	Supplier ^b
<i>M13 phage vectors^c</i>			
M13mp18/19	APB, ATCC, BR, LT, IBI, ICN, NEB, USB	M13BM20/21	BM
<i>Hybrid phage/plasmid vectors^d</i>			
Bluescript II series	ST	pSPORT1	LT
pBC series	ST	pSVK3	APB
pBS [±]	ST	pT3/T7-3	CL
pcDNAII	IN	pT7 T3 18U/19U	APB
pEMBL18/19	ATCC, BM	pT7/T3 α -18/19	LT
pfdA/B	ATCC	pT71/2	USB
pGEM9z/11z/13z f [±]	PR	pTZ18/19	USB
pIBI24(-25)	IBI	pTZ18R	IN
pICEM18/19	ATCC	pTZ18U/19U	BR
pSELECT1	PR	pUC118/119	ATCC
<i>Plasmid vectors^e</i>			
pAM18/19	APB	pUC18/19	APB, BM, BR, CL, LT, IBI, ICN, NEB, USB
pAT153	ICN		
pT3/T7/lac	BM		
pTTQ	APB	pUCBM20/21 SP72/73	BM PR

^aList of vectors modified from Slatko (1991). All vectors have commercially available primers, extensive polylinkers, and blue-versus-white screening for inserts using α -complementation of β -galactosidase.

^bAbbreviations: APB, Amersham Pharmacia Biotech; ATCC, American Type Culture Collection; BM, Boehringer Mannheim; BR, Bio-Rad; CL, Clontech; IBI, International Biotechnologies; ICN, ICN Biomedicals; IN, Invitrogen; LT, Life Technologies; NEB, New England Biolabs; PR, Promega; ST, Stratagene; USB, United States Biochemical. See APPENDIX 4 for locations and phone numbers. Suppliers' catalogues contain citations of original literature describing these vectors, if available.

^cThese vectors produce both double-stranded RF DNA and packaged single-stranded DNA.

^dThese vectors replicate as plasmids but produce single-stranded DNA upon infection with helper phage.

^eThese vectors produce only double-stranded DNA.

mercially available for sequencing both strands of insert DNA. Several of these vectors also contain an M13, f1, or fd origin of replication that can be activated to produce single-stranded copies of the plasmid by infection with helper phage (*UNITS 1.14 & 1.15*). Several of these hybrid phage/plasmid vectors (designated “ ϕ ” in Table 7.1.1) come as pairs with the phage origin of replication in both orientations such that helper phage directs the synthesis of opposite strands on the hybrid vector.

The design of the polylinker is the most important feature of the vectors listed in Table 7.1.1. The M13mp18/19 polylinker is found in many of these vectors, a feature that is often but not always indicated by the “18/19” designation in the plasmid name. Several other vectors, including Bluescript (which is commonly used in procedures for making nested deletions) have polylinkers containing several useful sites not present in the M13mp18/19 polylinker, as well as a different configuration of these sites. We recommend obtaining information about specific polylinkers from the supplier to determine which vector is best suited for a particular sequencing project.

CHEMICAL SEQUENCING

Planning for Chemical Sequencing

Chemical sequencing requires a DNA fragment that is labeled at only one end (*UNITS 3.5 & 3.10*). Traditionally, chemical sequencing strategies required relatively detailed knowledge of the restriction map for the fragment to be sequenced. This permitted generation of a series of subfragments that could be labeled at both ends by polynucleotide kinase or by the *E. coli* DNA polymerase I large fragment (Klenow fragment; *UNIT 3.5*). These labeled fragments would then be cut asymmetrically to produce two fragments labeled at each end, which could be purified individually by gel electrophoresis. However, specially designed vectors (described below) circumvent the requirement for purifying and labeling individual restriction

fragments and greatly facilitate chemical sequencing projects. These vectors contain asymmetric sites (see below) that permit labeling only one end of the target DNA to be sequenced; this strategy can be thought of as the chemical sequencing analog to the universal priming site. These vectors are particularly suited for cloning a set of nested deletions generated by *Bal 31*. We recommend using the nested deletion strategy for large chemical sequencing projects. For small fragments of DNA, the traditional end-labeling strategy is as efficient as subcloning into specialized vectors.

Vectors for Chemical Sequencing

Eckert (1987) has described the construction of specialized vectors that make it possible to rapidly and simultaneously sequence a large number of samples by the chemical cleavage method. These vectors, pSP64CS and pSP65CS (CS refers to chemical sequencing) are high-copy-number plasmids that confer ampicillin resistance and carry a synthetic polylinker containing two *Tth111I* sites flanking a *SmaI* site (Table 7.1.2). *Tth111I* cleaves at the sequence GACNNGTC leaving a single protruding 5' base. The two *Tth111I* sites were synthesized so it would be possible to selectively label one end of a *Tth111I* fragment using Klenow fragment in an end-labeling reaction as shown in Figure 7.1.1, bottom (Volckaert et al., 1984). End labeling of DNA fragments using Klenow fragment is simple and efficient and, since gel purification of labeled fragments is eliminated, the DNA can be sequenced directly after labeling. Since the plasmids are high-copy-number, a small (1.5 ml) bacterial culture yields enough plasmid DNA to repeat the sequencing reactions five to ten times, once the DNA is radiolabeled. The vectors allow for sequencing of a large number of fragments bordered by any of the unique restriction sites in the polylinker provided that at least one of the *Tth111I* labeling sites is retained. The sequence of the polylinker region of pSP64CS and pSP65CS is shown in Figure 7.1.1, top. The

Table 7.1.2 Chemical Sequencing Vectors

Vector	Supplier ^a
pSP64CS	ATCC
pSP65CS	ATCC
pUR222	BM
pCSV series	Volckaert et al., 1984
pSVB series	Arnold and Puhler, 1988

^aSuppliers' abbreviations are listed in Table 7.1.1.

two vectors differ only in the orientation of the polylinker relative to an SP6 promoter element, a feature of the plasmids not relevant to DNA sequencing (UNIT 3.8). Thus, it is necessary to use only one of the plasmids for DNA sequencing. (The SP6 promoter can be utilized to synthesize RNA complementary to the cloned DNA; thus, exon/intron mapping and other types of nuclease protection/mapping experiments can be carried out using the same plasmid construction as for sequencing.)

To use these vectors, the DNA fragment to be sequenced is subcloned into pSP64CS or pSP65CS and recombinant colonies are selected on ampicillin plates. Recombinant plasmid is prepared, digested with *Tth111I*, and the 5' single-base overhang created by *Tth111I* is filled in using the appropriate ³²P-labeled dNTP and Klenow fragment, such that the label is incorporated only at the end adjacent to the DNA to be sequenced (Fig. 7.1.1). Protocols for preparing pSP64/65CS-derived clones are presented in UNIT 7.3; protocols for *Tth111I* end labeling and chemical sequencing are provided in UNIT 7.5.

The *Bal* 31 procedure (UNIT 7.2) can be used to generate a set of nested deletions for subcloning into pSP64CS or pSP65CS. The blunt end generated by *Bal* 31 is ligated to the *Sma*I site in pSP64/65CS. This places the deleted end of each fragment next to a *Tth111I* site for end labeling (see Figs. 7.1.1 & 7.2.3 for strategy and protocol for constructing *Bal* 31 deletions).

Tth111I is an ideal site to utilize in an end-labeling vector because (1) it leaves a single protruding 5' base, assuring that the filling reaction will add only a single base per end; (2) *Tth111I* sites are rare, and thus sites present in the DNA being sequenced rarely complicate the labeling reaction; (3) the protruding base is G, A, T, or C, allowing a single DNA terminus to be selectively labeled (Fig. 7.1.1); and (4) the presence of two *Tth111I* sites in the vector makes it possible to selectively label either side of a fragment cloned between them. Together with *Tth111I*, other asymmetric sites (*Bst*EII and *Eco*81I) are incorporated in an additional family of plasmid vectors for chemical sequencing that utilize the same basic end-labeling strategy as pSP64CS/65CS (Table 7.1.1; Arnold and Puhler, 1988).

LITERATURE CITED

Ahmed, A. 1987a. A simple and rapid procedure for sequencing long (40kb) DNA fragments. *Gene* 61:363-372.

Ahmed, A. 1987b. Use of transposon-promoted deletions in DNA sequence analysis. *Methods Enzymol.* 155:177-204.

Anderson, S. 1981. Shotgun sequencing strategy using cloned DNase I-generated fragments. *Nucl. Acids Res.* 9:3015-3027.

Arnold, W. and Puhler, A. 1988. A family of high-copy-number plasmid vectors with single end-label sites for rapid nucleotide sequencing. *Gene* 70:171-179.

Bankier, A.T. 1984. Advances in dideoxy sequencing. *BioTechniques* 2:72-77.

Bankier, A.T. and Barrell, B.G. 1983. Shotgun DNA sequencing. In *Techniques in Life Sciences, B5: Nucleic Acid Biochemistry* (R.A. Flavell, ed.) pp. 1-34. Elsevier/North Holland, Limerick, Ireland.

Bankier, A.T., Weston, K.M., and Barrell, B.G. 1988. Random cloning and sequencing by the M13/dideoxynucleotide chain termination method. *Methods Enzymol.* 155:52-93.

Barnes, W.M. 1987. Sequencing DNA with dideoxyribonucleotides as chain terminating inhibitors: Hints and strategies for big projects. *Methods Enzymol.* 152:538-556.

Barnes, W.M. and Bevan, M. 1983. Kilo-sequencing: An ordered strategy for rapid DNA sequence data acquisition. *Nucl. Acids Res.* 11:349-368.

Barnes, W.M., Bevan, M., and Son, P.H. 1983. Kilo-sequencing: Creation of an ordered nest of asymmetric deletions across a large target sequence carried on phage M13. *Methods Enzymol.* 101:98-122.

Burton, F.H., Loeb, D.D., McGraw, R.A., Edgell, M.H., and Hutchison, C.A. 1988. A directed nucleotide-sequencing approach for single-stranded vectors based on recloning intermediates of a progressive DNA synthesis reaction. *Gene* 67:159-168.

Dale, R.M.K., McClure, B.A., and Houchins, J.P. 1985. A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: Application to sequencing the corn mitochondrial 18 S rDNA. *Plasmid* 13:31-40.

Deininger, P. 1983a. Random subcloning of sonicated DNA: Application to shotgun DNA sequence analysis. *Anal. Biochem.* 129:216-223.

Deininger, P. 1983b. Approaches to rapid DNA sequence analysis. *Anal. Biochem.* 135:247-263.

Eckert, R. 1987. New vectors for rapid sequencing of DNA fragments by chemical degradation. *Gene* 51: 245-252.

Frischauf, A.M., Garoff, H., and Lerach, H. 1980. A subcloning strategy for DNA sequence analysis. *Nucl. Acids Res.* 8:5541-5549.

Guo, L.-H. and Wu, R. 1982. New rapid methods for DNA sequencing based on exonuclease III digestion followed by repair synthesis. *Nucl. Acids Res.* 10:2065-2084.

Guo, L.-H., Yeng, R.C.A., and Wu, R. 1983. An improved strategy for rapid sequencing of both

- strands of long DNA molecules cloned in a plasmid. *Nucl. Acids Res.* 11:5521-5539.
- Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28:351-359.
- Hoheisel, J. and Pohl, F.M. 1986. Simplified preparation of unidirectional deletion clones. *Nucl. Acids Res.* 14:3605.
- Hong, G.F. 1982. A systematic DNA sequencing method. *J. Mol. Biol.* 158:539-549.
- Lee, Y.-M. and Lee, S.-C. 1989. A DNA sequencing strategy. *Anal. Biochem.* 175:521-524.
- Lin, H., Lei, S., and Wilcox, G. 1985. An improved DNA sequencing strategy. *Anal. Biochem.* 147:114-119.
- Liu, Z. and Hackett, P.B. 1989. Rapid generation of subclones for DNA sequencing using the reverse cloning procedure. *BioTechniques* 7:722-728.
- Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* 101:20-78.
- Messing, J. 1988. M13, the universal primer and the polylinker. *Focus (BRL)* 10:21-26.
- Messing, J. and Vieira, J. 1982. A new pair of M13 cloning vectors for selecting either DNA strand of double-digest restriction fragments. *Gene* 19:269-276.
- Misra, T.K. 1985. A new strategy to create ordered deletions for rapid nucleotide sequencing. *Gene* 34:263-268.
- Nag, D., Huang, H., and Berg, D. 1988. Bidirectional chain termination nucleotide sequencing: Transposon Tn5seq1 as a mobile source of primer sites. *Gene* 64:135-145.
- Norlander, J., Kempe, T., and Messing, J. 1983. Construction of improved M13 vectors using oligodeoxyribonucleotide-directed mutagenesis. *Gene* 26:101-106.
- Okita, T. 1985. Non-random sequencing of exonuclease III-deleted complementary DNA. *Anal. Biochem.* 144:207-211.
- Ozkaynak, E. and Putney, S.D. 1987. A unidirectional deletion technique for the generation of clones for sequencing. *BioTechniques* 5:770-773.
- Peng, Z. and Wu, R. 1986. A simple and rapid nucleotide sequencing strategy and its application in analyzing a rice histone 3 gene. *Gene* 45:135-145.
- Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E., and Surrey, S. 1982. "Nonrandom" DNA sequence analysis in bacteriophage M13 by the dideoxy chain-termination method. *Proc. Natl. Acad. Sci. U.S.A.* 79:4298-4302.
- Sanger, F., Coulson, A. R., Barrell, B.G., Smith, A.J.H., and Roe, B.A. 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* 143:161-178.
- Shen, W. and Waye, M. 1988. A novel method for generating a nested set of unidirectional deletion mutants using mixed oligodeoxynucleotides. *Gene* 70:205-211.
- Slatko, B. 1991. Sources of reagents and suppliers for dideoxy DNA sequencing and other applications. *In Methods in Nucleic Acids Research* (J. Karam, L. Chao, and G. Warr, eds.) pp. 379-392. CRC Press, Boca Raton, Fla.
- Smith, A.J.H. 1979. The use of exonuclease III for preparing single-stranded DNA for use as a template in the chain termination sequencing method. *Nucl. Acids Res.* 6:831-848.
- Smith, A. 1980. DNA sequencing analysis by primed synthesis. *Methods Enzymol.* 65:560-578.
- Strauss, N.A. and Zagurski, R.J. 1991. In vitro production of large single-stranded templates for DNA sequencing. *BioTechniques*. 10:376-384.
- Strauss, E., Kobori, J., Siu, G., and Hood, L. 1986. Specific-primer directed DNA sequencing. *Anal. Biochem.* 154:353-360.
- Vocke, C. and Bastia, D. 1983. Primary structure of essential replicon of the plasmid pSC101. *Proc. Natl. Acad. Sci. U.S.A.* 80:6557-6561.
- Volckaert, B., de Vleeschouwer, E., Blocker, H., and Frank, R. 1984. A novel type of cloning for ultra-rapid chemical degradation of sequencing of DNA. *Gene Anal. Tech.* 1:52-59.
- Wang, L.-M., Geihl, D.K., Choudhury, G., Minter, A., Martinez, L., Weber, D., and Sakaguchi. 1989. Site-directed deletion mutagenesis using phagemid vectors and genetic selection. *BioTechniques* 7:1000-1010.
- Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103-109.

Contributed by Barton E. Slatko
New England Biolabs
Beverly, Massachusetts

Richard L. Eckert
Case Western Reserve University
Cleveland, Ohio

Lisa M. Albright
Reading, Massachusetts

Frederick M. Ausubel
Massachusetts General Hospital
Boston, Massachusetts

Constructing Nested Deletions for Use in DNA Sequencing

Nested deletions useful for dideoxy DNA sequencing (UNIT 7.4) are a set of deletions originating at one end of a target DNA fragment and extending various lengths along the target DNA. Each successively longer deletion brings “new” regions of the target DNA into sequencing range (about 300 bp for normal sequencing gels) of the primer site (see UNIT 7.1 for a general discussion of nested deletions in DNA sequencing).

Two protocols for generating nested subclones via enzymatic digestion are included in this unit. In the basic protocol, a set of nested deletions is generated by exonuclease III. The primary advantage of this method is that the deletion products generated from the original clone can be recircularized to generate functional plasmids and thus do not require subcloning into another vector. An alternate method utilizes *Bal* 31 nuclease to generate the deletions. This method requires subcloning of the deletion fragments into a separate vector for subsequent use. Both the *exo* III and *Bal* 31 methods require the presence of unique restriction sites in the vector that are not present in the insert DNA. *Bal* 31 can also be used to generate nested deletions for chemical sequencing (UNIT 7.5) in conjunction with specialized chemical sequencing vectors (UNIT 7.1).

Commercially available kits for making nested deletions using a number of methods are described in UNIT 7.1 and listed in Table 7.2.1.

Table 7.2.1 Commercial Kits for Making Nested Deletions^a

Deletion method	Supplier ^b
<i>Exo</i> III	NEB, PH, PR, ST
<i>Bal</i> 31	NEB
T4 DNA polymerase	IB
Oligonucleotide-directed	B101

^aModified from Slatko, 1991.

^bAbbreviations: B101, Bio101; IB, International Biotechnologies; NEB, New England Biolabs; PH, Pharmacia LKB; PR, Promega; ST, Stratagene.

USING EXONUCLEASE III TO CONSTRUCT UNIDIRECTIONAL DELETIONS

BASIC PROTOCOL

This method is based on the enzymatic properties of *exo* III, a 3′ exonuclease specific for double-stranded DNA (UNIT 3.11). Exonuclease III can initiate digestion at blunt ends or ends with a 5′ overhang, but cannot efficiently initiate digestion at a 3′ overhanging end. The protocol contains the following steps: (1) cloning of target DNA into a polylinker in a suitable sequencing vector (see strategic planning, below, and UNIT 7.1 for a discussion of useful vectors); (2) double digestion of this plasmid at a site adjacent to the cloned target DNA with an enzyme that leaves a 5′ overhanging or blunt end and an enzyme that leaves a 3′ overhanging end of four bases; (3) exonucleolytic attack with *exo* III for varying lengths of time to create unidirectional digestion of the target sequence; (4) treatment with a single-stranded nuclease (*S*1 or mung bean nuclease) to remove the 5′ single strand, and repair of ends with Klenow fragment of *E. coli* DNA polymerase I; and (5) circularization and ligation with T4 DNA ligase and transformation of competent *E. coli* cells.

DNA Sequencing

7.2.1

Contributed by Barton Slatko, Peter Heinrich, B. Tracy Nixon, and Daniel Voytas

Current Protocols in Molecular Biology (1991) 7.2.1-7.2.20

Copyright © 2000 by John Wiley & Sons, Inc.

Materials

DNA fragment to be sequenced (insert DNA)
Appropriate sequencing vector (strategic planning and Table 7.1.1)
Restriction endonucleases and corresponding buffers (UNIT 3.1)
Buffered phenol (UNIT 2.1)
25:24:1 phenol/chloroform/isoamyl alcohol (UNIT 2.1)
Ice-cold 100% and 70% ethanol
3 M sodium acetate
Exonuclease III (UNIT 3.11) and 1× exonuclease III buffer
1 U/μl S1 nuclease (UNIT 3.12) and S1 nuclease buffer
S1 nuclease stop buffer
10× loading buffer (UNIT 2.5A)
2 U/μl Klenow fragment of *E. coli* DNA polymerase I (UNIT 3.5)
0.25 mM 4dNTP mix (dATP, dCTP, dGTP, and dTTP at 0.25 mM each; UNIT 3.4)
Exo III ligation buffer
10 mM ATP
1 U/μl T4 DNA ligase (measured in Weiss units; UNIT 3.14)

If plasmid vector is used

E. coli DK1 (*recA*; available from BRL) or DH5αF' (*recA*) made competent for transformation (Table 1.4.5 and UNIT 1.8)
LB plates and medium (UNIT 1.1) containing appropriate antibiotic

If one of the M13mp plasmids is the cloning vector

E. coli JM109 (*recA*) or DH5αF' (*recA*; BRL) made competent for transformation (Table 1.4.5 and UNIT 1.8)
LB plates (UNIT 1.1), prewarmed
H top agar (UNIT 1.1)
Xgal/IPTG mixture [30 μl Xgal stock and 30 μl IPTG stock (Table 1.4.2) for each 3 ml H top agar]
2× TY medium (UNIT 1.1)
Phage loading buffer

Additional reagents and equipment for subcloning DNA fragments (UNIT 3.16), CsCl/EtBr purification of plasmid DNA (UNIT 1.7), restriction enzyme digestion (UNIT 3.1), phenol extraction and ethanol precipitation (UNIT 2.1), agarose gel electrophoresis (UNIT 2.5A), transforming *E. coli* (UNIT 1.8), plasmid minipreps (UNIT 1.6), preparation of frozen glycerol stocks (UNITS 1.3 & 1.6), and dideoxy DNA sequencing (UNIT 7.4)

Strategic Planning

Figure 7.2.1 gives an overview of the experimental strategy for constructing a nested set of unidirectional deletions with exo III. Between the target DNA and the sequencing primer site there must be a recognition site for a restriction enzyme that generates a 5' overhanging end or a blunt end; between that restriction site and primer site there must be a recognition site for a restriction enzyme that generates a four-base 3' overhanging end that will be protected from exo III digestion. Incorporation of α-thio nucleotide analogs by filling in a 5' overhanging end adjacent to the primer site provides an alternative means of protection (see support protocol in this unit). Recently it has been shown that the *lac* repressor bound at *lac* operator site can also protect DNA from exo III digestion (Johnson et al., 1990).

To sequence both strands of the target DNA, deletions must be made from each end. Since exo III deletions of target DNA do not require subcloning into another vector, the vector into which the target DNA is initially cloned is also the one that will eventually

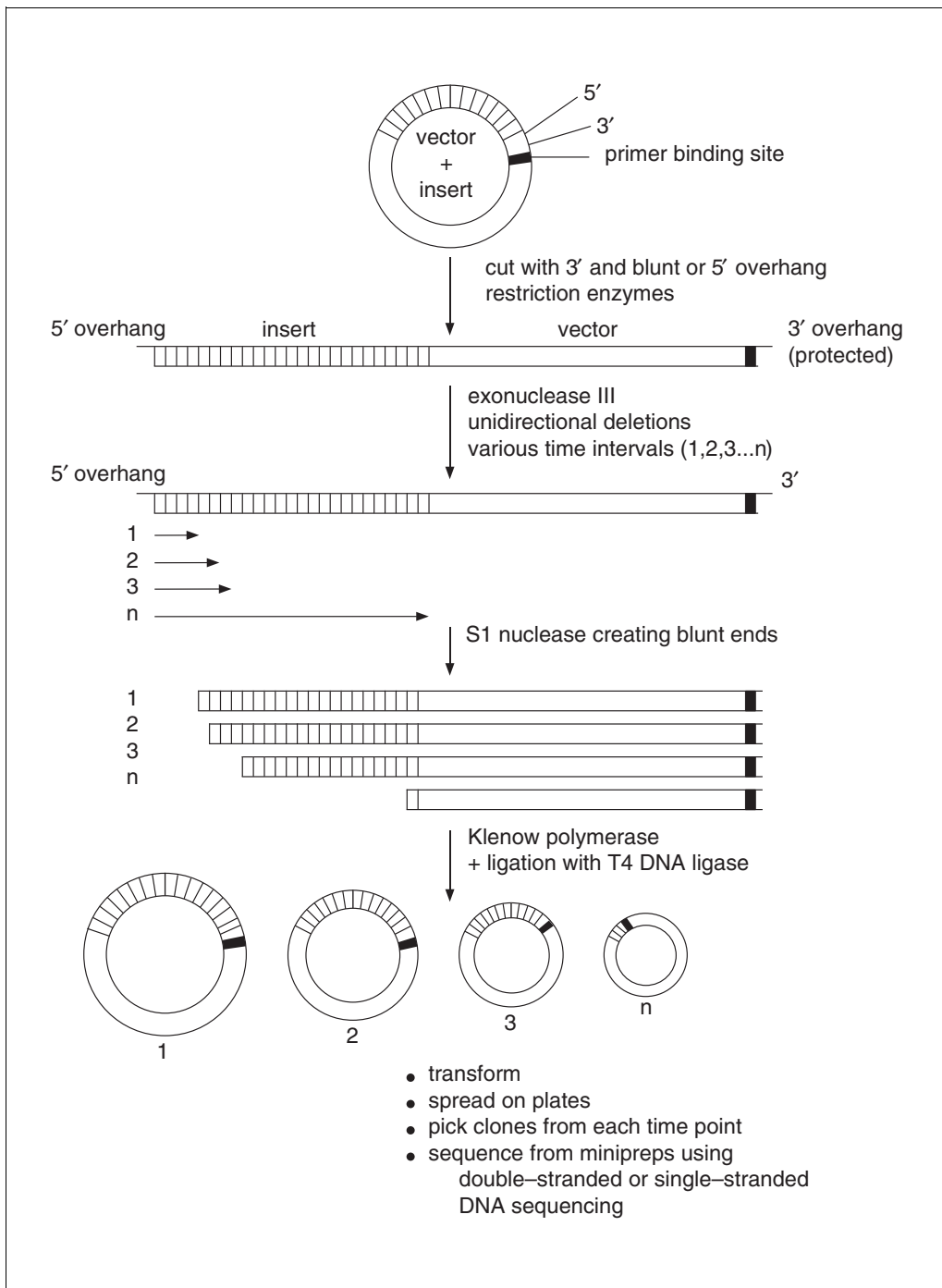


Figure 7.2.1 Construction of unidirectional deletions using *exo III*. The vector::insert DNA is digested with restriction endonucleases which generate a 3' overhang next to the sequencing primer site and a 5' overhang next to the insert DNA. Digestion with *exo III* generates unidirectional deletions from the end with the 5' overhang. Treatment with S1 nuclease and repair with Klenow fragment create blunt ends from both the *exo III*-treated ends and the 3' overhang which was protected from *exo III* digestion. The deleted plasmids are treated with DNA ligase and used to transform *E. coli*. Plasmids which have deletions are identified by gel electrophoresis of minipreparations of DNA.

provide sequencing templates. Thus, the DNA to be sequenced should be situated adjacent to restriction sites suitable for creating exo III deletions; in addition, priming sites for dideoxy sequencing should be located next to each end of the target DNA (Fig. 7.2.1). If the target DNA will be sequenced using double-stranded DNA templates, it is possible to construct deletions from each end of the fragment if both sequencing priming sites and suitable restriction sites are present in the vector on both sides of the target fragment. If the target fragment will be sequenced using single-stranded M13mp DNA templates, the target DNA must be cloned in both orientations with respect to the primer sequencing site using a pair of M13mp vectors. Refer to *UNIT 7.1* and Table 7.1.1 for information about sequencing vectors.

This protocol is written for constructing deletions in a ~2.0-kb fragment. However, we have successfully constructed nested sets of deletions in fragments as long as 7.0 kb, and it is likely that even longer insert fragments can be used. The protocol should be scaled up proportionately for longer fragments.

Preliminary Steps

1. Clone the DNA fragment to be sequenced (insert DNA) in the polylinker of an appropriate sequencing vector (see Fig. 7.2.1).

Refer to strategic planning above and to UNIT 7.1 for information regarding vector requirements for this procedure.

2. Prepare DNA twice-purified by CsCl/EtBr equilibrium centrifugation of the recombinant plasmids constructed in step 1.

It is important that the DNA be supercoiled; nicks in the vector DNA may cause degradation in the exonucleolytic step. Approximately 5 µg of plasmid DNA are required per experiment.

Exo III Procedure

DAY ONE: Linearize vector::insert DNA

3. Completely linearize 5 µg vector::insert DNA by double digestion with restriction enzymes that leave a 3' overhang adjacent to the primer site and a 5' overhang or blunt end adjacent to the insert DNA (see Fig. 7.2.1).

The 5' or blunt restriction site must be positioned between the 3' restriction site and the insert. Verify that digestion is complete by running a small aliquot on an agarose minigel (UNIT 2.5A).

4. Extract with buffered phenol, extract with phenol/chloroform/isoamyl alcohol, add $\frac{1}{10}$ vol of 3 M sodium acetate, and precipitate with ethanol (see *UNIT 2.1*). Wash the pellet in ice-cold 70% ethanol and dry it.
5. Dissolve the dried pellet in 1× exo III buffer for a final concentration of 0.1 µg/µl.

Create a nested set of unidirectional deletions

The extent of digestion of exo III is regulated by the reaction temperature and the time of incubation. Rates of 250 bp/min at 37°C and 120 bp/min at 30°C were obtained using the conditions outlined in the following steps.

6. Remove 25 µl (2.5 µg) of the linearized DNA (from step 5) and incubate 2 min at 37°C.
7. Add 150 U exo III per pmol susceptible 3' ends. Continue the 37°C incubation.

Conversion factor: 1 µg of 1-kb DNA = 1.5 pmol susceptible 3' ends.

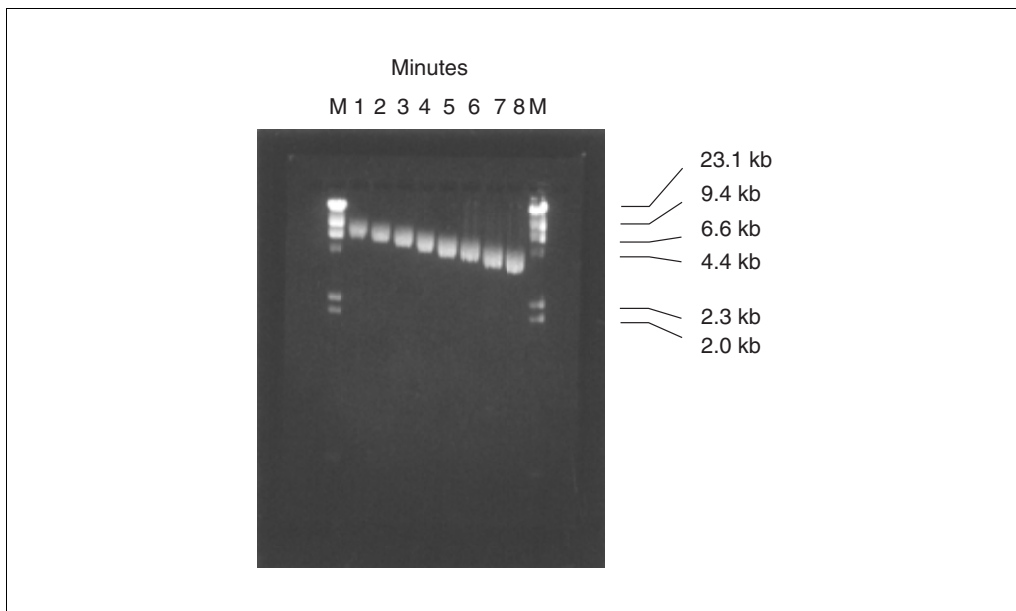


Figure 7.2.2 Ethidium bromide–stained agarose gel depicting exo III digestion. Bluescript vector (10 μg with a 3-kb insert) was cut with *Xba*I, which leaves a 5' overhanging end, and with *Bst*XI, which leaves a 3' overhanging end. The vector was then subjected to exo III digestion for eight different time intervals, followed by S1 nuclease digestion. Approximately 1 μg of each DNA sample was loaded on the gel. The DNA fragments in the lanes labeled “M” are molecular weight markers.

8. Remove 3- μl aliquots at 1-min intervals to individual microcentrifuge tubes (giving a total of eight samples) and place immediately on dry ice for 5 min. Add 3 μl water and inactivate the exo III by incubating 10 min at 70°C. Put the samples on ice.
9. Add 15 μl S1 nuclease buffer and 4 μl (4 U) S1 nuclease to each sample. Incubate 20 min at room temperature.
10. Stop the S1 nuclease reaction (via pH shift) by adding 5 μl S1 nuclease stop buffer to each sample.

Subject deleted DNA samples to electrophoresis

11. Add 2 μl of 10 \times loading buffer to an 8- μl aliquot of each sample and subject to electrophoresis on a 1% agarose gel, including ethidium bromide in the gel and buffer. Also run a lane of molecular weight markers (e.g., bacteriophage λ DNA, cut with *Hind*III) to determine the actual rate of exonuclease degradation. Decide which aliquots are worth saving.

Figure 7.2.2 illustrates the extent of deletions created on a Bluescript vector with a 3-kb insert starting with 10 μg DNA.

Create blunt ends for ligation

12. Add to each remaining 22 μl of the selected samples 1 μl (2 U) Klenow fragment. Incubate ~2 min at 37°C. Then add 1 μl of 0.25 mM 4dNTP mix and continue incubation at 37°C for 10 min to create blunt ends.

Recircularize the deleted molecules by ligation

13. Add 20 μl exo III ligation buffer, 3 μl of 10 mM ATP, 14 μl water, and 1 μl (1 U) T4 DNA ligase. Incubate 5 hr at room temperature or overnight at 15°C.

If using a plasmid vector and planning to carry out dideoxy sequencing with double-stranded plasmid template, proceed with steps 14 to 18. If using an M13mp vector and planning to carry out dideoxy sequencing with single-stranded templates, proceed to step 19.

DAY TWO (DOUBLE-STRANDED TEMPLATE): Transform ligated DNAs into competent *E. coli*

14. Transform 100 μ l of competent cells (*E. coli* DK1 or DH5 α F'; Table 1.4.5) with 10 μ l of each ligation reaction. Following heat-shock in the transformation protocol, spread $\frac{1}{5}$ of each of the cultures on LB plates containing antibiotic appropriate for the plasmid and incubate overnight at 37°C.

*Any competent *E. coli* recA strain can be used for transformation. About 2000 colonies are expected for each sample.*

DAY THREE (DOUBLE-STRANDED TEMPLATE): Characterize deleted plasmid clones

15. Pick two to four colonies from each plate and inoculate 5 ml LB medium containing appropriate antibiotic and grow overnight. Isolate DNA from 1.5 ml of each overnight culture using one of the miniprep procedures described in UNIT 1.6. Make a frozen glycerol stock with 1 ml of each culture (see UNIT 1.3) and store the remainder of each overnight culture at 4°C for DNA sequencing template preparation.

16. Add 2 μ l of 10 \times loading buffer and 10 μ l water to each 2.5- μ l aliquot of the plasmid DNA from step 15 and run on a 1% agarose gel. As size markers, also load 0.1 μ g each of the undeleted plasmid and parent vector on the gel. It should be possible to discern a ladder of deletions from full insert size to ~300 bp.

If it is difficult to size double-stranded circular molecules in increments of 200 bp, linearize the DNA at a common polylinker site before electrophoresis.

DAY FOUR (DOUBLE-STRANDED TEMPLATE): Prepare double-stranded plasmid templates for DNA sequencing

17. Choose a set of clones (from step 16) that differ from each other by intervals of 250 to 300 bp, spanning the complete region of the target DNA. Prepare plasmid DNA from the overnight cultures stored at 4°C (step 15) suitable as template for dideoxy sequencing (UNIT 7.3).

18. Subject the plasmid DNA of the deleted clones to nucleotide sequencing analysis, as described in UNIT 7.4.

The clones can initially be characterized by T-track analysis, in which each template is used in a dideoxy-T sequencing reaction. Whether or not overlapping regions occur in the chosen clones can be determined by the sequencing pattern.

Follow steps 19 to 26 if using an M13mp vector and planning to carry out dideoxy sequencing using single-stranded templates.

DAY TWO (SINGLE-STRANDED TEMPLATE): Transform ligated DNAs into competent *E. coli*

19. Transform 100 μ l of competent JM109 or DH5 α F' *E. coli* cells with 10 μ l of each reaction mix (from step 13). Also transform with 0.5 ng of the M13mp RF vector DNA and 0.5 ng of the M13mp::insert RF DNA that carries the nondeleted fragment.

**E. coli* DH5 α F' (Table 1.4.5) can be maintained on rich media without loss of the F' factor, is suppressor plus, restriction minus, modification plus, recombination deficient (recA1), and exhibits a higher transformation efficiency than JM109.*

20. Following the heat shock step in the transformation protocol, add 0.2 ml of an overnight culture of *E. coli* JM109 or DH5 α F', pipet contents into 3 ml of room

temperature H top agar containing 60 μ l Xgal/IPTG mixture, mix, and pour on LB plates. Incubate overnight at 37°C.

About 2000 plaques are expected for each sample. All plaques should be white except for those from the M13mp RF vector DNA, which should be blue.

DAY THREE (SINGLE-STRANDED TEMPLATE): Characterize deleted phages

21. Picking two to four M13 plaques from each transfection plate (i.e., two to four per exonuclease time point), grow individual M13mp phage stocks as described in steps 1 through 4 in the protocol for preparing single-stranded templates in UNIT 7.3. Grow stocks of the M13mp vector and M13mp::insert as controls. Store supernatants at 4°C.

Following sequencing, particular deletion intervals may be missing. For this reason, it is useful to be able to examine additional phage from a particular time point. Because phage diffuse within the agar, we advise picking several additional colorless plaques from each time point into individual wells in a 96-well microtiter plate containing 200 μ l of 2 \times TY medium plus 15% glycerol, growing them at 37°C for 6 hr, and freezing at -20°C for future use, if necessary.

If further small-scale manipulations with RF DNA from the deleted phage are desired, harvest cells in addition to the phage supernatants and isolate RF DNA by the alkaline lysis method described in UNIT 1.6.

22. Place 20 μ l of each phage supernatant into 4 μ l of phage loading buffer and load on a 1% agarose gel. Also load M13mp vector(s) and M13mp::inserts (undeleted) supernatants as size markers. It should be possible to discern a ladder of deletions from full insert size to ~300 bp.

Optional test for detecting small inserts: "C testing" entails mixing 10 μ l of a phage supernatant carrying a presumptive insert with 10 μ l of the supernatant from the M13 control carrying the full-length, undigested opposite strand. Add 4 μ l phage loading buffer and incubate 30 min at 65°C prior to electrophoresis. Annealing between the phage DNAs causes significant retardation in the gel compared to the single-stranded circular forms.

C testing can also be used to determine the orientation of an insert.

DAY FOUR (SINGLE-STRANDED TEMPLATE): Prepare single-stranded phage templates for DNA sequencing, choosing phage that contain appropriate deletions from previous step

23. Choose clones from step 22 that differ from each other by intervals of 250 to 300 bp, spanning the complete insert DNA.
24. Respin the stored phage supernatants from step 21 (there will be some cell growth, even at 4°C) and transfer supernatants to fresh 1.5-ml microcentrifuge tubes.

The presence of chromosomal DNA will interfere with sequencing reactions.

25. Proceed immediately to step 5 of the protocol in UNIT 7.3 for preparing single-stranded template DNA for sequencing.
26. Subject the prepared single-stranded DNA templates to dideoxy sequencing as described in UNIT 7.4. Also perform sequencing reactions on the vector containing the undeleted fragment.

See annotation to step 18 concerning characterization of clones by T-track analysis.

**PROTECTION OF DNA FROM EXONUCLEASE III DIGESTION
USING [α -³⁵S]dNTPs**

Creating unidirectional deletions for DNA sequencing using exonuclease III requires that the end of the molecule adjacent to the sequencing primer site be protected from digestion. This is typically accomplished using restriction endonucleases that provide four-base, 3' overhanging ends. Incorporating α -thio nucleotide analog, [α -³⁵S]dNTP, adjacent to the primer site provides an alternative means of protecting against exo III digestion (Putney et al., 1981). Restriction sites adjacent to the primer site that leave a 5' overhang can be filled in with Klenow fragment using the appropriate thio nucleotide analog to cap the 3' end. This increases flexibility in designing a deletion strategy, as any restriction site that leaves a 5' overhang can be appropriately modified to provide the necessary protection from digestion by exo III (Ozkaynak and Putney, 1987).

Additional Materials

- 5 mM appropriate dNTPs (UNIT 3.4)
- 5 mM appropriate [α -³⁵S]dNTPs (Pharmacia)

1. In a 20- μ l reaction mixture, completely linearize 5 μ g vector::insert DNA by digesting with a restriction endonuclease that leaves a 5' overhanging end adjacent to the sequencing primer site.

Verify that the digestion is complete by running a small aliquot on an agarose minigel (UNIT 2.5A).

2. Add 1 μ l of each of the appropriate 5 mM [α -³⁵S]dNTPs and 5 mM dNTPs, and 1 U Klenow fragment (UNIT 3.5). Incubate 30 min at room temperature.

The [α -³⁵S]dNTPs and dNTPs required depend upon the sequence of the 5' overhang left by the restriction endonuclease. Adequate protection from exo III digestion will be obtained for both partially and completely filled-in restriction sites, providing that the 3'-most nucleotide added by Klenow fragment contains the thio group.

Ozkaynak and Putney (1987) used higher concentrations of [α -³⁵S]dNTPs; however we have found that the above concentration will provide adequate incorporation.

3. Stop the Klenow polymerase reaction by incubating 10 min at 75°C. Digest the DNA with a second restriction endonuclease that generates a 5' overhanging or blunt end adjacent to the insert DNA.

Be sure to adjust the salt concentration of the reaction mix to that appropriate for the second enzyme (see UNIT 3.1). Do not exceed 20% glycerol in the final restriction endonuclease reaction mix.

4. Proceed with step 4 of the protocol for constructing nested deletions with exo III.

USING BAL 31 NUCLEASE TO CONSTRUCT NESTED DELETIONS

This method is based on the properties of Bal 31 nuclease, a single strand-specific endonuclease that can degrade the ends of duplex linear DNA from both the 5' and 3' termini (see UNIT 3.12). As illustrated in Figure 7.2.3 (left side), the Bal 31 procedure involves linearizing vector::insert DNA at one end of the region to be sequenced. In a second step, a series of deletions are created that extend various amounts into the desired region using Bal 31. Because Bal 31 will shorten both ends of the linearized vector::insert, the third step separates the deleted insert from the vector component by digestion with a second restriction endonuclease at the opposite end of the insert DNA. Finally, the deleted insert fragments are subcloned into an appropriate dideoxy sequencing vector such that the deleted Bal 31 ends are adjacent to the sequencing priming site. Alternatively, for chemical sequencing, the deleted fragments can be cloned into

chemical sequencing vectors, such as pSP64CS/65CS, so that the deleted *Bal* 31 ends are adjacent to a universal site for end-labeling. To sequence both strands of the insert DNA, a nested set of deletions from each side of the insert must be generated. To do this, the *Bal* 31 digestion and subcloning procedures are repeated using a separate aliquot of the plasmid that has been linearized at the other end of the region to be sequenced (Fig. 7.2.3, right side).

A nested set of deletions spanning a 1.5- to 2.0-kb region can be reliably generated using this procedure. If a larger region is to be sequenced, it is best to choose two large deletions (one from each end of the fragment) from the first set of *Bal* 31 digestions and then to repeat the procedure with these deleted fragments. Alternatively, a series of ~2.0-kb fragments spanning the region to be sequenced can be cloned individually. If this latter procedure is adopted, choose a set of overlapping rather than adjacent fragments to avoid the problem of overlooking very small restriction fragments.

This protocol is written for generating a set of deletions from only one end of the DNA (i.e., the left half of Fig. 7.2.3). The protocol can be duplicated and performed simultaneously to generate the second set of deletions (i.e., the right half of Fig. 7.2.3).

Materials

DNA fragment to be sequenced (insert DNA)
Appropriate “source” and “sequencing” vectors (see strategic planning below and Table 7.1.1)
Buffered phenol (UNIT 2.1)
Ice-cold 70% ethanol
3 M sodium acetate
TE buffer (APPENDIX 2)
10× and 1× *Bal* 31 nuclease buffer
Bal 31 nuclease (UNIT 3.12)
200 mM EGTA
1 mg/ml yeast tRNA
Restriction endonucleases and corresponding buffers (UNIT 3.1)
10× loading buffer (UNIT 2.5A)
Low gelling/melting temperature agarose (UNIT 2.6)
Ethidium bromide
50× TAE buffer (optional; APPENDIX 2)
95% ethanol (optional)
10× T4 DNA ligase buffer (UNIT 3.4)
400 U/μl T4 DNA ligase (measured in cohesive-end units; UNIT 3.14)

If plasmid vector is used

E. coli JM109, DH5αF', or equivalent made competent for transformation (Table 1.4.5 and UNIT 1.8)
LB plates and medium (UNIT 1.1) containing appropriate antibiotic

If one of the M13mp plasmids is the cloning vector

E. coli MC1061, DH5αF', JM109, or equivalent made competent for transformation (Table 1.4.5 and UNIT 1.8)
Overnight cultures of *E. coli* JM101, JM107, JM109 or equivalent (Table 1.4.5) in LB or 2× TY medium (UNIT 1.1)
H top agar (UNIT 1.1)
Xgal/IPTG solution [30 μl Xgal stock and 30 μl IPTG (Table 1.4.2) for each 3 ml of H top agar]
LB plates (UNIT 1.1), prewarmed
2× TY medium (UNIT 1.1)
Phage loading buffer

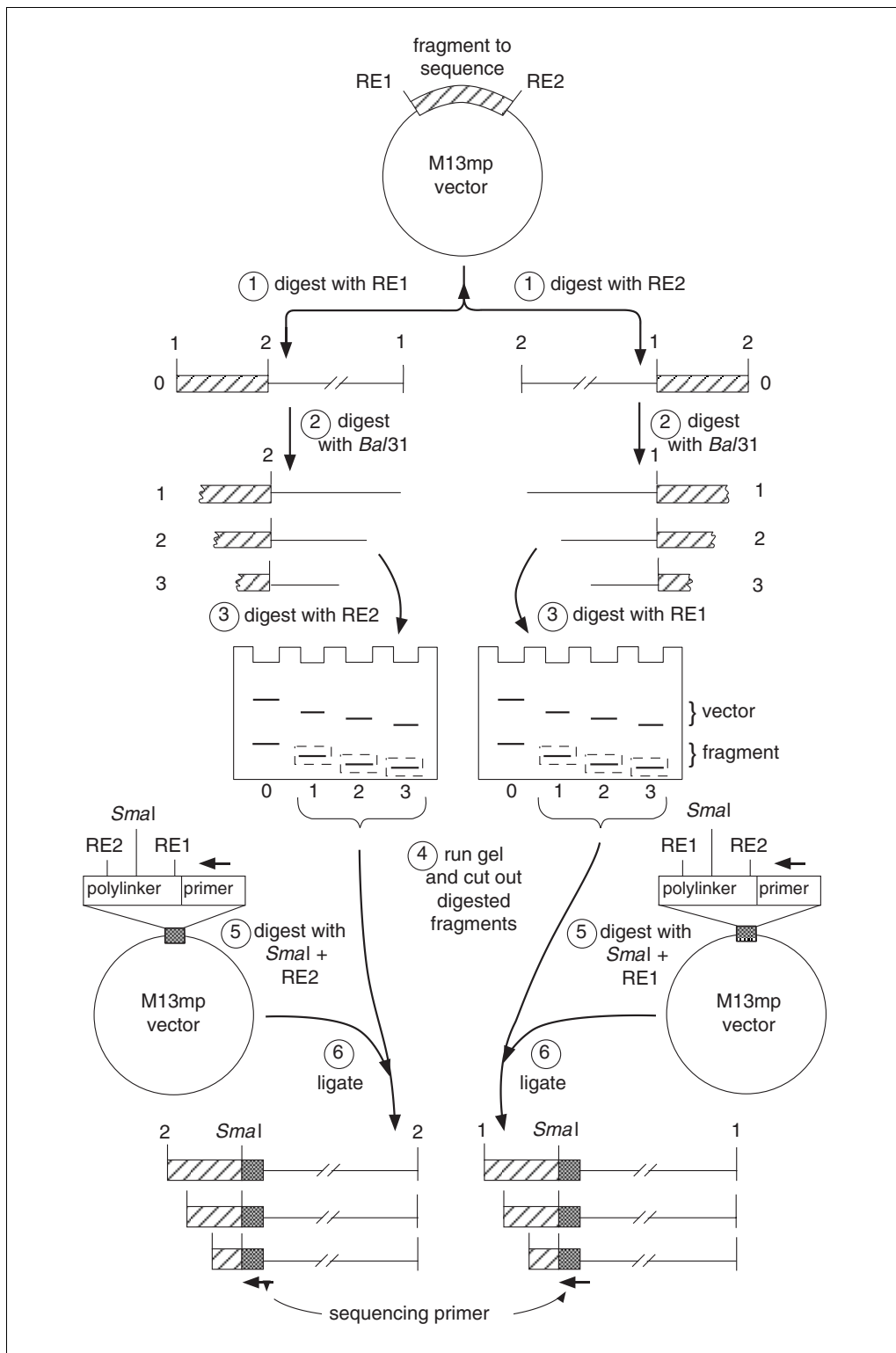


Figure 7.2.3 Strategy for making a nested set of deletions with *Bal*31 nuclease. The source vector containing a fragment to be deleted (in this example RF DNA of an M13mp vector) is linearized at each end of the fragment in separate reactions (**step 1**) and a nested set of deletions is generated with *Bal* 31 nuclease (**step 2**). The source vector and insert DNA are separated by restriction endonuclease digestion (**step 3**) and agarose gel electrophoresis (**step 4**). The deleted fragments are eluted from the gel and ligated to a sequencing vector (in this example, an M13mp vector) that has been digested with a restriction enzyme generating a blunt end (compatible with ends digested by *Bal* 31; in this example, *Sma*I) and another restriction enzyme corresponding to the site at the other end of the insert (RE1 or RE2). "RE2" and "RE2" refer to two different restriction endonucleases or their respective recognition sites.

Additional reagents and equipment for subcloning DNA fragments (*UNIT 3.16*), CsCl/EtBr purification of plasmid DNA (*UNIT 1.7*), phenol extraction and ethanol precipitation (*UNIT 2.1*), preparing and transforming competent *E. coli* (*UNIT 1.8*), agarose gel electrophoresis (*UNIT 2.5A*), isolating fragments using low gelling/melting temperature agarose (*UNIT 2.6*), plasmid minipreps (*UNIT 1.6*), and preparing single-stranded phage DNA templates (*UNIT 7.3*)

Strategic Planning

Refer to Figure 7.2.3 for an overview of the experimental strategy that can be employed for constructing a nested set of deletions with *Bal* 31 nuclease.

Because *Bal* 31–deleted fragments are subcloned into a separate sequencing vector, two vector selections must be made prior to generating the nested set of deletions: one that will serve as a source of DNA to be subjected to *Bal* 31 digestion (“source” vector), and one that will accept the deleted fragments and ultimately generate sequencing templates (“sequencing” vector).

Source vector. Two factors are important in selecting the source vector for the initial cloning of the insert to be subjected to *Bal* 31 digestion:

- The source vector::insert must be constructed such that the region to be deleted has a unique restriction site at each end. For each set of *Bal* 31 deletions, one site is required for DNA cleavage where the deletions will begin and a second site (at the opposite end of the insert) serves to separate the deleted insert from the vector. These restriction sites flanking the insert must be compatible with polylinker sites in the sequencing vector.
- In this procedure, the *Bal* 31–digested insert DNA is purified from vector DNA by agarose gel electrophoresis. In theory, any insert that is minimally 80% different (either larger or smaller) than the vector can be differentiated on a standard agarose gel. For example, if the source vector is one of the pUC series of plasmids that are ~2.7 kb, the insert should be ≤ 2.3 kb or ≥ 3.2 kb.

A wide variety of vectors meet the above criteria (see Table 7.1.1). One advantage of using an M13mp vector as source vector is that M13 is generally considerably larger than the region to be sequenced and can be readily separated from the insert by gel electrophoresis.

NOTE: Inserts >2.5 kb are often unstable in M13mp vectors and may delete spontaneously. JM109 (*recA*; Yanisch-Perron et al., 1985) may stably propagate larger inserts.

An alternative strategy for subcloning the *Bal* 31–deleted fragments is to omit the gel electrophoresis step after the second restriction enzyme digestion (Fig. 7.2.3, step 4) and simply subclone everything into an appropriate sequencing vector. Those subclones that contain insert DNA can be differentiated from those containing source vector DNA via plaque lift or colony blot analysis (*UNITS 6.1-6.4*). In this case, size difference between source vector and insert is not important. It is advantageous to use a high-copy-number source vector, such as the pUC series, to easily prepare sufficient quantities of DNA for *Bal* 31 digestion (5 μ g of DNA is required per 150 to 200 bp of DNA to be deleted).

Sequencing vector. Choice of sequencing vector depends on the method that will be used to sequence the nested set of deletions. See *UNIT 7.1* and Tables 7.1.1 and 7.1.2 for a description of vectors suited to dideoxy and chemical sequencing and the factors involved in their selection. The major constraint on the sequencing vector is that it have a polylinker-site configuration that can accept the *Bal* 31–deleted fragments so that the blunt end created by *Bal* 31 digestion is next to a priming site for dideoxy sequencing (or an end-labeling site for chemical sequencing).

If M13mp sequencing vectors are chosen, (UNIT 7.1) it may be necessary to use a pair of complementary M13mp vectors, each of which will accept fragments with deletions from one end. This will be the case if, in the sequencing vector polylinker, the two restriction sites flanking the insert DNA to be sequenced are located on opposite sides of the restriction enzyme site used to provide a blunt end. This case is illustrated in Figure 7.2.3.

Preliminary Steps

Prepare source vector::insert double-stranded DNA as starting material

1. Clone the fragment to be sequenced in an appropriate source vector, choosing a vector according to the criteria in strategic planning, above.

If an M13mp vector is chosen as source vector, clone the insert DNA in both orientations with respect to the sequencing priming site so that the sequence at each end of the undeleted insert can be obtained. However, prepare CsCl-purified DNA for only one of the orientations to use for Bal 31 digestion.

2. Prepare 5 μ g of CsCl-purified double-stranded DNA of the clone constructed in step 1 for each 150 bp to be deleted (e.g., 50 μ g for a 1.5-kb fragment; this is a two-fold excess to allow for repetition if necessary).

RNA will interfere with digestion by Bal 31 nuclease; if RNA contaminates the DNA preparation after one CsCl gradient, purify the DNA through a second gradient.

Prepare sequencing vector to accept Bal 31-digested fragments

Two separate preparations of sequencing vector DNA are required, each for a set of deletions from one end (see Fig. 7.2.3, step 5). For each set of deletions, the sequencing vector should have a blunt end (compatible with *Bal* 31-digested ends) adjacent to a dideoxy sequencing priming site (or chemical sequencing end-labeling site) and a sticky end compatible with the restriction site on the opposite side of the insert from which *Bal* 31 digestion was initiated in that set of deletions (and which was therefore protected from *Bal* 31 digestion). It can be prepared anytime in advance and stored at -20°C .

3. For each set of *Bal* 31 deletions, digest ≥ 2 μ g of CsCl-purified double-stranded DNA of the chosen sequencing vector with the appropriate restriction endonucleases.

A typical protocol for preparing M13mp sequencing vector DNA is described in the second support protocol.

Prepare competent recipient cells

4. For subcloning *Bal* 31-deleted fragments, prepare cells competent for transformation and store at -70°C . If fragments are to be subcloned into sequencing templates that feature *lac* α -complementation or into M13mp vectors, use *E. coli* JM109, DH5 α F', or equivalent host strain. Alternatively, MC1061 can be initially transformed, then mixed with DH5 α F' or equivalent host, which will allow formation of plaques.

Strain MC1061 helps the efficiency of this procedure because of its high transformation frequency, but cannot be used for M13mp7, 8, 9, 10, and 11 because amber mutations in these vectors are not suppressed by MC1061.

Bal 31 Procedure

DAY ONE: Linearize source vector::insert DNA (Fig. 7.2.3, step 1)

5. Completely linearize 2 μg of source vector::insert DNA for each 150 bp to be deleted plus an additional 2 μg (e.g., 22 μg for a 1.5-kb fragment) using a restriction endonuclease that cuts at one end of region to be deleted.

Two separate preparations of linearized source vector::insert DNA are required, each for deletions from one end.

6. Verify that digestion went to completion by running a small aliquot on an agarose minigel.

For each set of deletions, the maximum amount of DNA that should be linearized is 28 to 30 μg , i.e., sufficient to construct deletions in a 2.0-kb fragment.

7. Clean up the restriction digest from step 5 by extracting with buffered phenol; add $\sim 1/10$ vol 3 M sodium acetate and precipitate with ethanol. Wash the pellets with ice-cold 70% ethanol, and dry. Resuspend the pellets in TE buffer at 1 $\mu\text{g}/\mu\text{l}$ and store at 4°C.

The final concentration of the DNA can be checked by reading the A_{260} of a sample diluted $\sim 1:200$ (APPENDIX 3).

Determine the sensitivity of each batch of linear source vector::insert DNA to digestion using Bal 31 nuclease

The following steps are necessary because base composition of the DNA and impurities (such as RNA) can affect the rate of Bal 31 digestion and because Bal 31 may lose activity during storage.

8. Mix 2 μl (2 μg) linearized DNA, 38 μl water, and 5 μl of 10 \times Bal 31 nuclease buffer, and aliquot 9 μl of resulting mixture into five microcentrifuge tubes. To each tube add 1 μl of either 1 \times Bal 31 buffer or Bal 31 nuclease diluted $1/5$, $1/10$, $1/20$, or $1/40$ in 1 \times Bal 31 buffer.
9. Incubate each reaction 30 min at 37°C, then stop the reactions by adding 1 μl of 200 mM EGTA to each and heating 5 min at 65°C.

EGTA chelates Ca^{++} required by Bal 31.

10. Analyze the samples by agarose gel electrophoresis (a minigel is acceptable; see UNIT 2.5A) to determine the dilution of Bal 31 nuclease that completely digests the DNA in the 30-min incubation.

DNA samples typically require 0.03 to 0.10 U Bal 31 nuclease per μg linearized DNA. As the digestion proceeds, the DNA bands first decrease uniformly in size, but gradually become increasingly diffuse before becoming completely digested. Electrophoresing in the presence of ethidium bromide saves time.

DAY TWO: Create a nested set of deletions (Fig. 7.2.3, step 2)

Perform Bal 31 digestions on each of the linearized source vector::insert DNA preparations (from step 5). Use 2 μg DNA of each for each 150 bp to be deleted. Use the amount of Bal 31 nuclease previously determined (steps 8 to 10) to completely digest the DNA. The protocol assumes that 1-min time points will be taken.

11. For each 150 bp to be deleted, mix 2 μl (2 μg) linearized DNA, 43 μl water, and 5 μl of 10 \times Bal 31 nuclease buffer.
12. Preheat diluted DNA by incubating 10 min at 37°C.
13. For each linearized DNA preparation, set aside a 1.5-ml microcentrifuge tube for each time point (i.e., one tube for each 2 μg DNA to be subjected to Bal 31 digestion).

Add 5 μ l of 200 mM EGTA to each tube and label tubes to indicate the linearized species and the *Bal* 31 digestion time.

14. Add *Bal* 31 nuclease of the determined amount to the tube of linearized DNA, vortex gently, and incubate at 37°C. Every minute, transfer 45 μ l to the appropriately labeled tube from step 13 containing 200 mM EGTA and place at room temperature.

The digestion described should generate 150- to 200-bp deletions; the time interval or the amount of enzyme can be adjusted. We recommend taking time points for no longer than 10 to 12 min. At longer digestion times, the DNA bands that form following electrophoresis (see step 18 annotations below and Fig. 7.2.4) become too diffuse to be eluted efficiently from an agarose gel.

Repeat steps 11 to 14 for each preparation of linearized source vector::insert DNA, one set of deletions from each end of the insert DNA.

15. After all *Bal* 31 digestions have been collected, incubate them 5 min at 65°C.

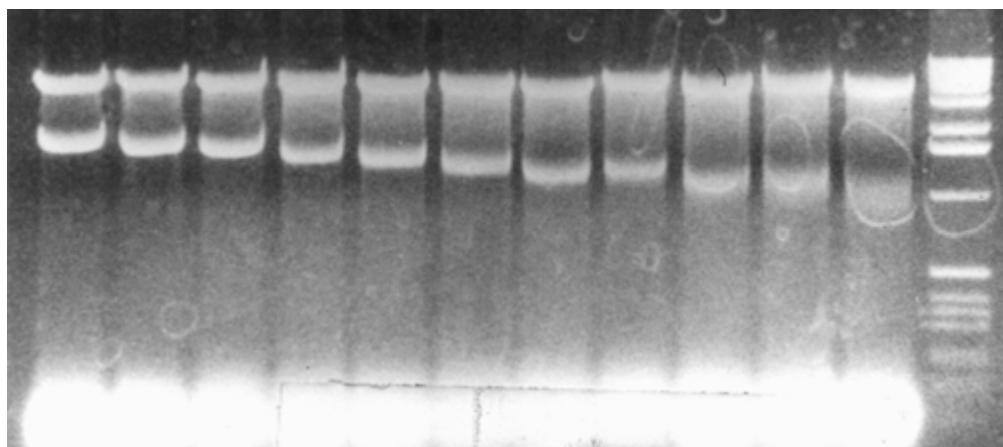


Figure 7.2.4 Ethidium bromide–stained agarose gel depicting successive digestion of a linearized M13mp vector::insert with *Bal* 31 nuclease. Following *Bal* 31 digestion for increasing amounts of time (1-min time points, left to right), the M13mp vector::insert was digested with a restriction enzyme that separated the vector (upper band) from the DNA fragment of interest (lower band). Digested fragments appear as increasingly diffuse bands with increased digestion time. The molecular weight markers on the right are the 1-kb ladder available from Gibco/BRL; the size of the deleted insert varied from ~2.0 kb to ~1 kb.

Separate digested insert and vector sequences (Fig. 7.2.3, step 3)

16. Add 2 μ g of 1 mg/ml yeast tRNA to each tube and ethanol precipitate the samples using 3 M sodium acetate. Dissolve the washed and dried pellets in 20 μ l of the appropriate restriction endonuclease buffer for the next step.

*Optional: Before digestion with the second restriction endonuclease, the *Bal* 31–treated ends can be repaired with Klenow fragment of DNA polymerase to create blunt ends (see UNIT 3.5). This will increase the efficiency of ligations for subcloning. The reaction can take place in most restriction endonuclease buffers. Inactivate the Klenow fragment by heating the reaction mix at 75°C for 10 to 15 min before proceeding to the next step.*

17. Digest each of the DNA samples with the restriction endonuclease whose recognition site was protected from *Bal* 31 digestion (i.e., flanking the insert opposite the site used for linearization).

Subject digested DNA samples to electrophoresis (Fig. 7.2.3, step 4)

18. Add 2 μ l of 10 \times loading buffer to each digested DNA sample and subject to electrophoresis on low gelling/melting temperature agarose, including ethidium bromide in the gel and buffers. Also run a lane of molecular weight markers spanning the size of the fragment to be sequenced.

This gel can be run overnight at low voltage. If an M13mp source vector has been used, run a 1% gel and allow bromphenol blue dye to migrate ~4 cm. This gives good separation between vector and insert.

Because electrophoresis buffer and chambers can develop nuclease contamination, when fragments will be isolated from gels for subsequent cloning it is desirable to use autoclaved 50 \times TAE for preparing buffers and casting gels.

The digested fragments should appear as increasingly diffuse bands with increased digestion time (see Fig. 7.2.4).

DAY THREE: Elute Bal 31–digested fragments from gel

19. Using a longwave UV (354-nm) lamp and the molecular weight markers as a guide (and minimizing exposure to UV light), cut out a series of DNA fragments (usually the entire band of digested insert DNA) differing in molecular weight by 150 to 200 bp from each lane (Fig. 7.2.4) and place in individual microcentrifuge tubes.
20. Add TE buffer to ~400 μ l total volume for each tube, incubate at 65 $^{\circ}$ C until all the agarose melts (about 5 to 10 min), and extract with 1 ml buffered phenol by vortexing at low speed for 30 sec.

When extracting agarose, it is important to obtain sufficient mixing for equilibration between the phenol and water (i.e. vortex vigorously for 2 full min). Do not use phenol/chloroform or phenol/chloroform/isoamyl alcohol; they will not extract efficiently.

21. Repeat the phenol extraction on the aqueous phase combined with its milky interface which, after the second extraction, will become much less abundant. If a significant amount of interface remains, repeat extraction a third time.
22. Precipitate with 95% ethanol. Add 2 μ g of 1 mg/ml yeast tRNA, if desired. Dissolve the washed and dried pellets in 30 μ l TE buffer and store at 4 $^{\circ}$ C.

Low gelling/melting temperature agarose is very reliable for isolating fragments for subsequent ligations. It is sometimes possible to save time by diluting melted slices to <0.1% agarose and ligating without extraction and precipitation of the DNA. Refer to UNIT 2.6 for alternative methods of purifying DNA from agarose gels.

Ligate eluted Bal 31–digested fragments to sequencing vector (Fig. 7.2.3, step 6)

23. Add 3 to 8 ml of each DNA sample eluted from the agarose gel to appropriate sequencing vector DNA prepared in step 3. Include as a control one sample which contains vector DNA alone. Add 1 ml of 10 \times T4 DNA ligase buffer and adjust the final reaction volume to 10 ml. Add ~0.5 ml T4 DNA ligase and incubate overnight at 4 $^{\circ}$ C.

20 to 40 ng of M13mp and pUC sequencing vectors is sufficient. A convenient concentration of sequencing vector DNA is 20 μ g/ml; therefore, 1 to 2 μ l would be added to the Bal 31–digested DNA. The sample containing vector DNA alone controls for the amount of vector DNA that was not fully digested with both restriction endonucleases; such molecules can recircularize and be ligated.

The amount of Bal 31–digested DNA used in this ligation should be chosen according to the recovery of the DNA from the agarose gel. Three microliters would be used for >90% recovery.

A large amount of ligase is required for the blunt-end ligation.

If planning to sequence double-stranded templates, follow steps 24 and 25. If using M13mp vector, proceed to step 26.

DAY FOUR (DOUBLE-STRANDED TEMPLATE): Transform ligated DNAs into competent *E. coli*

24. Transform 100 μ l of competent cells (JM109, DH5 α F' or equivalent from step 4) with 5 to 10 μ l of each ligation reaction (from step 23). In addition, transform with sequencing vector DNA untreated with ligase, as described in step 25. Following heat-shock of transformation protocol, spread $\frac{1}{5}$ of each of the cultures on LB plates containing the appropriate antibiotic and incubate overnight at 37°C.

The hosts described above are appropriate for sequencing vectors that allow a positive screen for inserts using alpha-complementation of β -galactosidase. In this case, include Xgal/IPTG in the LB plates (UNIT 1.1).

DAYS FIVE AND SIX (DOUBLE-STRANDED TEMPLATE): Characterize recombinant plasmid clones and prepare double-stranded plasmid templates for DNA sequencing.

25. If there are many colorless (not blue) colonies, proceed with steps 15 to 18 in the first basic (exo III) protocol. Pick two to four colorless colonies from each time point and grow in LB medium containing the appropriate antibiotic.

If using an M13mp sequencing vector and planning to carry out dideoxy sequencing using single-stranded templates, proceed with steps 26 to 30.

DAY FOUR (SINGLE-STRANDED TEMPLATE): Transform ligated DNAs into *E. coli* MC1061 Strain MC1061 is recommended because of its high transformation frequency; however, other hosts can be used (see step 4).

26. Mix each of the ligation reactions with 40 μ l of competent cells.

If the vector was not treated with phosphatase, many more plaques—predominantly blue—may result. Under these circumstances, transform with $\leq \frac{1}{4}$ of each ligation reaction.

27. In addition, transform with M13mp sequencing vector (Fig. 7.2.3, step 6) cut with *Sma*I and additional enzyme, untreated with ligase.

The sample not treated with ligase controls for the amount of undigested M13mp sequencing vector molecules that will form blue plaques.

Other controls: Also transform with 0.5 ng of the M13mp RF sequencing vector used to subclone the Bal 31-deleted fragments and the source vector::insert clones carrying the undeleted fragment. Expect nearly a lawn of M13 plaques with the RF control DNA.

28. Following the heat shock step in the transformation protocol, add 0.2 ml of an overnight culture of *E. coli* JM109, DH5 α F', or equivalent in LB medium or 2 \times TY medium. Pipet contents into 3 ml of room temperature H top agar containing 60 μ l Xgal/IPTG solution, mix, and pour on LB plates.

Prior to pouring top agar, it is essential to prewarm LB plates to at least room temperature to retard setting of the top agar.

29. After the top agar hardens, incubate plates upside down overnight at 37°C.

It is essential to incubate plates at 37°C for plaques to form.

Each transformation should result in ~100 colorless plaques. Depending on the fidelity of vector preparation, these will be accompanied by zero to many blue plaques.

DAYS FIVE AND SIX (SINGLE-STRANDED TEMPLATE): Characterize recombinant phage and prepare single-stranded templates for DNA sequencing

30. If there are many colorless (not blue) plaques, proceed as in steps 21 to 26 in the previous basic (exo III) protocol. Picking two plaques for each time point and two plaques from each of the control plates is usually sufficient.

If only a few colorless plaques are present, they may not be the desired recombinants. The presence of inserts detected by electrophoresis will resolve the issue, as will "C testing" (see step 22 in the exo III protocol).

PREPARATION OF M13mp SEQUENCING VECTOR DNA FOR SUBCLONING OF BAL 31-DIGESTED DNA FRAGMENTS

**SUPPORT
PROTOCOL**

This protocol provides an example of the preparation of M13mp 18 or M13mp 19 vector DNA for subcloning *Bal* 31-digested DNA fragments. The vector DNA is first digested with a restriction endonuclease generating a blunt end (in this example *Sma*I) which is compatible with *Bal* 31-digested ends. In a second step, the vector DNA is digested with a restriction endonuclease generating a sticky end which is compatible with the sticky end on the opposite end of the *Bal* 31-digested fragment.

Materials

M13mp18 or M13mp19 (UNITS 1.14 & 1.15)

*Sma*I restriction endonuclease and buffer (UNIT 3.1)

Buffered phenol

Ice-cold 100% ethanol

70% ethanol

Appropriate restriction endonuclease (Fig. 7.2.3) and buffer

Calf intestine phosphatase (CIP; UNIT 3.10; optional)

TE buffer (APPENDIX 2)

Additional reagents and equipment for restriction enzyme digestion (UNIT 3.1) and phenol extraction and ethanol precipitation (UNIT 2.1)

1. For each 10 µg of DNA that will be subjected to *Bal* 31 digestion, digest 1 µg of mp18 and/or mp19 with *Sma*I.

It is important to determine the minimal amount of SmaI needed to linearize the vector; too much SmaI can damage ends and prevent subsequent ligation. Usually, ~2 U SmaI per microgram of mp18 or mp19 DNA for 60 min at 37°C is sufficient.

Alternatively, DNA polymerase (UNIT 3.5) or single-stranded nuclease (UNIT 3.12) can be used to convert a sticky end to a blunt end.

2. Extract with buffered phenol and precipitate with ethanol as in UNIT 2.1.

If the DNA preparations are really clean, it may be necessary to add 1 to 2 µg yeast tRNA before precipitation to see the pellets.

3. Digest the *Sma*I-linearized sequencing vector with the restriction enzyme whose site flanks the insert DNA and which will be protected from *Bal* 31 digestion in that set of deletions (Fig. 7.2.3, step 5). Use the minimum amount of enzyme necessary. If desired, remove phosphate groups by adding ≤0.2 µl CIP (~5 U) for the last 20 min of the digestion with the chosen restriction endonuclease.

Two separate preparations of sequencing vector DNA are required, each for a set of deletions from one end.

The following restriction buffers work for this procedure: 10 mM Tris·Cl (pH 7.5), 10 mM MgCl₂, and 10 to 150 mM NaCl. A vast excess of phosphatase is used, allowing the use of a suboptimal buffer. Phosphatase treatment sometimes results in a nonfunctional vector. If

DNA Sequencing

7.2.17

dephosphorylation is omitted, transform with less DNA in the second basic protocol (steps 25 or 29).

4. Extract with buffered phenol, precipitate with ethanol, and dissolve in 50 μ l TE buffer. Stored at -20°C , such vectors last months.

REAGENTS AND SOLUTIONS

1 \times exonuclease III buffer

15 mM Tris·Cl, pH 8.0
0.66 mM MgCl_2

Exo III ligation buffer

80 mM Tris·Cl, pH 7.5
30 mM dithiothreitol (DTT)
20 mM MgCl_2
Store at -20°C

Phage loading buffer

0.25% bromphenol blue
15.0% Ficoll 400
2.0% sodium dodecyl sulfate (SDS)
10 mM EDTA
Store at room temperature

S1 nuclease buffer

16 mM sodium acetate, pH 4.6
400 mM NaCl
1.6 mM ZnSO_4
8% glycerol

S1 nuclease stop buffer

0.8 M Tris·Cl, pH 8.0
20 mM EDTA, pH 8.0
80 mM MgCl_2

10 \times Bal 31 buffer

2.0 M NaCl
0.2 M Tris·Cl, pH 8.0
0.12 M MgCl_2
0.12 M CaCl_2
0.02 M EDTA

COMMENTARY

Background Information

Exonuclease III. The experimental strategy of the exo III method is based on the unique properties of *E. coli* exo III, which is a 3'→5' double-stranded, specific exonuclease that catalyzes release of 5' nucleotides from the 3' hydroxy end of double-stranded DNA (see UNIT 3.11 and Fig. 3.11.3). Blunt DNA ends (e.g., created by *Sma*I) or overhanging 5' ends (e.g., created by *Hind*III, *Xba*I, *Bam*HI) are substrates for an exo III attack. However, exo III will not efficiently initiate digestion at an overhanging 3' end of four bases created

by restriction enzymes such as *Apa*I, *Sac*I, *Kpn*I, and *Bst*XI. The 3' overhang generated by *Pst*I digestion is not adequately protected from exo III digestion. Exo III also will not digest ends of DNA that contain an [α - ^{35}S]dNMP (see first support protocol). Thus, a cloned fragment of DNA can be specifically digested with exo III if it is adjacent to a 5' overhanging end or blunt end, whereas a primer sequencing site and vector DNA can be protected from digestion if they are adjacent to a 3' overhanging end or an end "capped" with a-thio nucleotide analog. A third method of protecting DNA from exo III digestion by *lac* repressor bound at *lac* operator site has recently

been shown (Johnson et al., 1990).

Exo III is particularly suited for generating nested sets of deletions because of the synchrony and uniformity of its exonucleolytic activity (see Fig. 7.2.2).

The exo III method has one major advantage over the *Bal* 31 nuclease method (Poncz et al., 1982; Guo and Wu, 1982) for constructing nested sets of deletions. The *Bal* 31 procedure requires subsequent recloning of the set of digested fragments; in contrast, the exo III procedure is carried out on a single linearized plasmid and is followed by relegation. Double digestion of the DNA, exo III treatment, and relegation can be accomplished in 3 hr (Hoheisel and Pohl, 1986). Another advantage of the exo III procedure over the *Bal* 31 procedure is that larger DNA fragments (up to 14 kb; Henikoff, 1984) can be easily subjected to construction of progressive deletions. On the other hand, the exo III procedure depends on the availability of appropriate restriction sites between the insert and primer site. If the required sites are not available for the exo III procedure, the *Bal* 31 procedure can be used.

***Bal* 31 exonuclease.** The basis of this method is the ability of *Bal* 31 nuclease to progressively shorten a DNA restriction fragment, generating blunt ends that can be ligated with phage T4 DNA ligase. Because *Bal* 31 digestion continues in a nearly linear fashion for up to 10 to 15 min, a nested set of deletions can be easily created. This method is appropriate for DNA fragments that do not have the required arrangement of restriction sites for exo III digestion. All that is needed is that the DNA fragment be flanked by unique restriction sites.

Critical Parameters and Troubleshooting

For phage preparations of M13mp clones and DNA preparations of M13mp or plasmid sequencing templates, refer to *UNIT 7.3*. When examining M13mp phage DNA for inserts, optimum resolution is achieved when large gels are run for 10 to 15 hr at low voltage.

Exonuclease III. It is important to start with a highly purified DNA preparation that has been subjected to two rounds of CsCl/ethidium bromide centrifugation. Preexisting nicks in the DNA will create a substrate for exo III and S1 nuclease, leading to nonspecific degradation. During the double-digestion step, it is very important that the DNA is completely digested. Choosing restriction sites as far apart as possible increases the probability of obtaining a complete digestion.

If the yield of DNA from longer digestion times turns out to be disproportionately low, the exo III is probably contaminated with other nucleases. A high background of randomly deleted clones is another sign of nuclease contamination in the exo III, the restriction enzymes, or the S1 nuclease. Because the activity of exo III is very sensitive to temperature, the temperature of the reaction can be varied between 30° and 40°C to achieve a satisfactory rate of digestion. It is advantageous to fix time points (30 to 60 sec) and adjust the temperature to obtain the desired deletion interval. The rate of exo III digestion can also be regulated by including NaCl in the digestions (Tomb and Barcak, 1989).

The final concentration of NaCl in the ligation reaction is somewhat inhibitory for T4 ligase, but a sufficient number of deleted clones should be obtained. If poor ligation efficiency results in few transformants, the DNA can be precipitated from 3.75 M ammonium acetate and two vol of ethanol. Alternatively, the ligation volume can be increased to dilute the NaCl.

***Bal* 31 exonuclease.** It is important to start with a highly purified DNA preparation, because *Bal* 31 is very sensitive to RNA contamination. If RNA contamination is suspected after one round of CsCl-ethidium bromide centrifugation, the DNA should be purified through a second gradient. Complete linearization of the plasmid DNA—by excessive treatment with the appropriate restriction endonuclease to the extent that the ends are damaged—should theoretically not be a problem, as *Bal* 31 will convert a damaged end into a blunt end suitable for ligation.

The number of nucleotides separating the ends of the deletions can be varied by adjusting the amount of *Bal* 31 used and the time of digestion, and by the use of commercially available preparations of *Bal* 31 that are characterized as “fast” or “slow.” If insufficient digestion occurs with *Bal* 31, repeat the procedure using a higher concentration of enzyme.

Some researchers improve the quality of the *Bal* 31-generated ends by filling out with Klenow fragment (see annotation to step 16) or by digestion with S1 nuclease or mung bean nuclease before digestion with the second restriction enzyme (*UNIT 3.12*). Using MC1061 seems to eliminate this requirement, because of its high transformation frequency. If low numbers of transformants are obtained, that optional step can be performed.

EGTA very effectively stops *Bal* 31 digestion by chelating Ca⁺⁺; however, it also inhibits

subsequent digestion with a variety of restriction enzymes. Therefore, it is very important to ethanol precipitate samples carefully after the addition of EGTA. This will allow digestion with minimal amounts of restriction enzymes, thus providing high-quality ends for subsequent ligation reactions.

Successful cloning of *Bal 31*-digested fragments into sequencing vectors also depends on using the minimal amount of restriction endonuclease treatment to prepare the sequencing vector for ligation. Unfortunately, because the sequencing vector is first cut with *Sma*I (or other restriction enzyme generating a blunt end) and then with a second enzyme that cuts only a few bases away in the polylinker, it is difficult to determine the extent of cleavage of the second enzyme. Sequencing vectors suitable for ligation are best prepared by individually titrating each restriction enzyme with uncut sequencing vector to determine the minimal amount to completely digest the DNA. If necessary, check the ability of ligase to circularize the cut sequencing vector following *Sma*I digestion and again following digestion with the second enzyme. There should be a dramatic decrease in the sequencing vector's ability to be self-ligated following treatment with the second enzyme. The doubly-digested sequencing vector can be treated with calf intestine phosphatase (*UNIT 3.10*) to further reduce self-ligation. Alternatively, it can be separated from the small piece of DNA between the two restriction sites by electrophoresis on low gelling/melting temperature agarose (see steps 18 to 22). Excess prepared sequencing vector should last for months or years if kept frozen.

Time Considerations

Exonuclease III. Starting with CsCl/ethidium bromide-purified plasmid DNA, it takes 3 or 4 days to generate and characterize a set of 30 overlapping fragments of a 3.0-kb target DNA.

Bal 31 exonuclease. Starting with CsCl/ethidium bromide-purified plasmid DNA, it takes ~6 days to generate a set of 20 ordered fragments for a 2.0-kb insert of interest.

Literature Cited

- Guo, L.-H. and Wu, R. 1982. New rapid methods for DNA sequencing based on exonuclease III digestion followed by repair synthesis. *Nucl. Acids Res.* 10:2065-2084.
- Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28:351-359.

- Hoheisel, J. and Pohl, F.M. 1986. Simplified preparation of unidirectional deletion clones. *Nucl. Acids Res.* 14:3605.
- Johnson, D.F., Nierlich, D.P., and Lysis, A.J. 1990. The use of *lac* repressor in constructing sequential deletions and a new sequencing vector. *Gene* 94:9-14.
- Ozkaynak, E. and Putney, S.D. 1987. A unidirectional deletion technique for the generation of clones for sequencing. *BioTechniques* 5:770-773.
- Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E., and Surrey, S. 1982. "Nonrandom" DNA sequence analysis in bacteriophage M13 by the dideoxy chain-termination method. *Proc. Natl. Acad. Sci. U.S.A.* 79:4298-4302.
- Putney, S.D., Benkovic, S.J., and Schimmel, P. 1981. A DNA fragment with an α -phosphorothioate nucleotide at one end is asymmetrically blocked from digestion by exonuclease III and can be replicated in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 78:7350-7354.
- Slatko, B. 1991. Sources of reagents and supplies for dideoxy DNA sequencing and other applications. In *Methods in Nucleic Acids Research* (J. Karam, L. Chao, and G. Warr, eds.) pp. 379-392. CRC Press, Boca Raton, Fla.
- Tomb, J.F., and Barcak, G.J. 1989. Regulating the 3'→5' activity of exonuclease III by varying the NaCl concentration. *BioTechniques* 7:932-933.
- Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103-109.

Key References

- Henikoff, S. 1987. Unidirectional digestion with exonuclease III in DNA sequence analysis. *Methods Enzymol.* 155:156-165.
- Hoheisel and Pohl. 1986. See above.
These papers describe the original procedure from which the exo III protocol is adapted.
- Poncz et al. 1982. See above.
Describes the original procedure from which the Bal 31 protocol is adapted.

Contributed by Barton Slatko
New England Biolabs
Beverly, Massachusetts

Peter Heinrich (exo III)
Consortium für Elektrochemische Industrie
Munich, Federal Republic of Germany

B. Tracy Nixon (*Bal 31*)
Pennsylvania State University
University Park, Pennsylvania

Daniel Voytas ($[\alpha\text{-}^{35}\text{S}]\text{dNTPs}$)
Iowa State University
Ames, Iowa

Preparation of Templates for DNA Sequencing

UNIT 7.3

This unit contains protocols for preparing DNA suitable for use as dideoxy sequencing templates and as material for end labeling and chemical sequencing. In all protocols, the starting material contains the recombinant molecule to be sequenced. DNA from M13mp-derived phage is easily prepared (first basic protocol) and is currently the most reliable source of template for large-scale dideoxy sequencing projects. Because it is occasionally necessary or convenient to use a λ -derived phage as a source of DNA, a protocol for preparing λ phage DNA from plate lysates is provided (second basic protocol). Two protocols for minipreps of plasmid DNA are provided, one intended for dideoxy sequencing, the other for end labeling and chemical sequencing (third and fourth basic protocols); they differ primarily in the way in which cellular RNA is removed. Protocols for preparing the products of the polymerase chain reaction (PCR) for dideoxy sequencing are provided in *UNIT 15.2*.

Any double-stranded template used in dideoxy sequencing should be denatured before being annealed to the primer. In general practice, alkali denaturation of plasmid DNA works better for sequencing than heat denaturation (Chen and Seeburg, 1985; Zagursky et al., 1985; Hattori and Sakaki, 1986; Zhang et al., 1988); thus, alkali denaturation of double-stranded DNA is described in this unit (fifth basic protocol). In contrast, heat denaturation of PCR products can be important in generating reliable sequencing templates (*UNIT 15.2*). A final protocol describes the preparation of template for thermal cycle sequencing from a single phage plaque or bacterial colony.

PREPARATION OF SINGLE-STRANDED M13 PHAGE DNA

Single-stranded M13mp DNA to be sequenced by the dideoxy method is prepared from *E. coli* DH5 α F' (or equivalent *E. coli* F' host) infected with an M13mp plaque. Because M13mp clones are sometimes susceptible to DNA rearrangements, we recommend preparing the template DNA from a single isolated plaque. The plaque should be purified shortly before use by transformation of competent *E. coli* cells with a ligation mixture (such as would come from construction of nested deletions, *UNIT 7.2*) or M13 RF DNA or single-stranded DNA. A stock of M13 phage can be replated to give individual plaques. An alternative to plaque purification is to infect the *E. coli* host with a phage suspension. Because this method bypasses the plaque purification step, it should be kept in mind as a source of background bands in sequencing ladders generated from such DNA templates (see *UNIT 7.4*, troubleshooting).

Cells are harvested and phage are precipitated from the culture supernatant with polyethylene glycol and high salt. Single-stranded viral DNA is then extracted from the phage with phenol and ethanol precipitated. If the single-stranded M13mp DNA obtained using the procedure described below is not clean enough (i.e., insufficient label is incorporated during the sequencing reactions or the sequencing gels have a high background), the optional annotations to the steps should be followed.

Materials

- E. coli* DH5 α F' (or equivalent; Table 1.4.5), competent (*UNIT 1.8*) and overnight cultures in LB medium (*UNITS 1.1 & 1.2*)
- Recombinant M13mp RF DNA or M13mp single-stranded DNA (*UNIT 1.15*)
- H top agar (*UNIT 1.1*)
- LB plates (*UNIT 1.1*), 37°C
- 2 \times TY medium (*UNIT 1.1*)
- M13 polyethylene glycol (PEG) solution

**BASIC
PROTOCOL**

DNA Sequencing

7.3.1

TE buffer (*APPENDIX 2*)
Buffered phenol (*UNIT 2.1*)
3 M sodium acetate, pH 5.2 (*APPENDIX 2*)
100% ethanol and 70% ethanol, ice-cold
Pasteur pipets, regular and drawn-out (sterile)
18 × 150–mm tubes
Additional reagents and equipment for transformation (*UNIT 1.8*), cell growth (*UNIT 1.2*), and agarose gel electrophoresis (*UNIT 2.5*)

Grow M13mp phage

1. If starting with M13 RF DNA or M13 single-stranded DNA, transform 40 μ l competent *E. coli* DH5 α F' with ~0.5 ng recombinant M13mp RF DNA or with 5 to 10 ng single-stranded DNA. Following the heat-shock step, add 0.2 ml of an overnight *E. coli* DH5 α F' culture to the transformed cells. Pipet the cells into 3 ml of H top agar, mix, and pour onto 37°C LB plates. Incubate upside-down overnight at 37°C.

It is essential to incubate plates at 37°C for plaques to form. Xgal and IPTG can be added to the top agar to detect contaminating wild-type plaques (which will be blue; recombinant plaques will be clear).

Single-stranded DNA transforms much less efficiently than double-stranded DNA.

*If transforming with a ligation mixture for creating nested deletions (UNIT 7.2), follow steps 19 and 20 in the exo III protocol or steps 26 through 29 in the Bal 31 procedure, using appropriate *E. coli* host.*

*To purify individual plaques from a previously prepared phage stock, titer stock on an overnight culture of *E. coli* F' cells (UNIT 1.15).*

2. Dilute an overnight *E. coli* DH5 α F' culture 100-fold with 2× TY medium and dispense 1.5-ml aliquots into loosely capped 18 × 150–mm tubes.
3. Pick one M13mp plaque with a sterile Pasteur pipet and expel it into one of the cultures. Be sure that a plug of agar is transferred. Repeat for each plaque.

Alternatively, each culture can be inoculated with a “loopful” of phage from a previously prepared M13 phage stock that has been stored in 15% glycerol at –20°C.

4. Grow cells 5 to 6 hr at 37°C (preferably on a roller drum).

To decrease the chance of defective phage production, do not incubate >8 hr. Defective phage, which arise spontaneously, are secreted into the medium and can reinfect cells in which the recombinant clone is growing, thereby creating a mixed population of template DNA.

Harvest cells and precipitate phage

5. Transfer cells to a 1.5-ml microcentrifuge tube and microcentrifuge at high speed for 5 min at 4°C to room temperature. Transfer supernatant to a clean 1.5-ml microcentrifuge tube.

*Supernatants can be stored at 4°C for several weeks. If this is done, microcentrifuge the stored supernatants at high speed (there will be some cell growth, even at 4°C) before proceeding to the next step, and transfer supernatant to a clean 1.5-ml microcentrifuge tube. The presence of *E. coli* chromosomal DNA in the stored supernatants interferes with sequencing reactions.*

6. Add 200 μ l of M13 PEG solution to the phage supernatant, mix, and incubate 15 min at room temperature.
7. Microcentrifuge 5 min at high speed and discard the supernatant.

A white pellet should be visible. The pellet will not be disturbed by vigorous shaking.

8. Microcentrifuge 30 sec at high speed and remove any remaining supernatant with a drawn-out Pasteur pipet.

It is critical to remove all traces of the PEG supernatant.

It is optional to resuspend pellet in 500 μ l TE buffer; this provides further purification of phage particles from cellular contaminants. Add 100 μ l of M13 PEG solution, mix well, and leave 15 min at room temperature. Repeat steps 7 and 8.

Prepare single-stranded DNA from phage

9. Resuspend phage pellet in 100 μ l TE buffer.

If many samples are being prepared, PEG pellets can be resuspended in TE buffer by placing all tubes in a suitable rack and vigorously shaking rack by hand.

10. Extract with 100 μ l buffered phenol by vortexing 15 to 20 sec at low speed. Microcentrifuge at high speed 5 min at 4°C to room temperature and transfer upper (aqueous) phase to a clean microcentrifuge tube.

It is optional to extract upper phase with 100 μ l chloroform.

11. Add 11 μ l of 3 M sodium acetate, pH 5.2, and 250 μ l of 100% ethanol. Freeze 20 min at -70°C.

12. Microcentrifuge at high speed 10 min, 4°C, and discard supernatant.

13. Add 100 μ l ice-cold 70% ethanol, microcentrifuge, and remove supernatant. Repeat this 70% ethanol wash once.

It is optional to wash the ethanol precipitate with 100% ethanol to remove phenol remaining in the aqueous phase.

14. Dry DNA pellet in a Speedvac evaporator.

Single-stranded DNA pellets as an arched film, most readily seen after tubes are dry.

15. Resuspend pellet (2 to 3 μ g) in 20 μ l TE buffer. Electrophorese 0.5 μ l on a minigel to confirm the DNA concentration. Store the remainder at -20°C until used as template for sequencing reactions (UNIT 7.4).

Dissolve the DNA film in TE buffer by simply pipetting buffer over the film until the drop moves freely over the tube surface.

PREPARATION OF λ DNA FROM SMALL-SCALE LYSATES

Phage are prepared from a plate lysate using agarose in place of agar in both the top agarose and plates. The phage are precipitated with polyethylene glycol and high salt, then lysed in SDS and EDTA. Phage DNA is then extracted with phenol and chloroform and precipitated with isopropanol.

Materials

Recombinant λ phage

E. coli strain appropriate for λ phage of interest

Lambda top agarose and agarose plates (i.e., substitute agarose for agar; UNIT 1.1), freshly prepared on the day of use

1 mg/ml DNase I (UNIT 3.12)

1 mg/ml RNase A, heat-inactivated (see reagents and solutions)

Lambda polyethylene glycol (PEG) solution

SM medium (UNIT 1.11)

10% sodium dodecyl sulfate (SDS)

Buffered phenol (UNIT 2.1)

0.5 M EDTA

BASIC PROTOCOL

DNA Sequencing

7.3.3

1:1 phenol/chloroform
Chloroform
3 M sodium acetate, pH 7.0
Isopropanol
TE buffer, pH 8.0 (*APPENDIX 2*)
Pasteur pipets, drawn out (sterile)

Additional reagents and equipment for preparing and titering a λ phage plate lysate (*UNITS 1.11 & 1.12*)

Prepare λ phage stock

1. Prepare the recombinant λ phage stock on an *E. coli* strain appropriate for the λ phage by plate lysis using freshly prepared lambda top agarose and lambda agarose plates.

It is critical to use agarose in place of agar to yield DNA suitable for sequencing and restriction endonuclease analysis. Use one freshly prepared lambda agarose plate per phage stock. If freshly prepared agarose plates are unavailable or inconvenient, older plates may be substituted. Add 5 ml sterile SM buffer to the plate and allow the buffer to soak into the plate for 15 min. Remove excess buffer with a sterile pipet. Take care after pouring the top agarose, since it is loosely attached to the bottom agarose; incubate the plates right-side up.

2. Titer the λ phage stock—the titer should be $>10^9$ pfu/ml.
3. Place 700 μ l of phage stock in a sterile 1.5-ml microcentrifuge tube. Add 1 μ l of 1 mg/ml DNase I and 1 μ l of 1 mg/ml heat-inactivated RNase A. Incubate 30 min at 37°C.

This step removes cellular DNA and RNA that might remain in the medium.

Precipitate the phage

4. Add 700 μ l lambda PEG solution and incubate 1 hr on ice.

Do not incubate >4 hr or contaminating chromosomal DNA will precipitate.

5. Microcentrifuge at high speed 10 min at 4°C and pour off supernatant.
6. Microcentrifuge at high speed 2 min at 4°C and remove any remaining supernatant with a drawn-out Pasteur pipet.
7. Resuspend phage pellet in 500 μ l SM medium.

If the pellet remains undissolved, it will dissolve in the next step.

Prepare DNA from phage

8. Add 5 μ l of 10% SDS and 0.5 μ l of 0.5 M EDTA to the solution and mix gently. Incubate 15 min at 65°C. Periodically mix gently to dissolve any remaining pellet.
9. Extract with 500 μ l buffered phenol by vigorous vortexing. Microcentrifuge at high speed 2 min at room temperature and transfer upper (aqueous) phase to a clean microcentrifuge tube.
10. Repeat step 9 using 500 μ l of 1:1 phenol/chloroform.
11. Repeat step 9 using 500 μ l chloroform.
12. Add 50 μ l of 3 M sodium acetate, pH 7.0, followed by 550 μ l isopropanol. Freeze 30 min at -20°C .
13. Microcentrifuge at high speed 15 min at 4°C and aspirate and discard supernatant.

14. Air dry the DNA pellet.

The pellet will be invisible.

15. Resuspend pellet (~20 µg) in 50 µl TE buffer, pH 8.0.

16. For sequencing, denature 0.1 to 0.3 pmol (2 to 4 µg) in alkali as described in the fifth basic protocol.

MINIPREP OF RECOMBINANT pSP64CS OR pSP65CS PLASMID DNA FOR CHEMICAL SEQUENCING

**BASIC
PROTOCOL**

Recombinant DNA from pSP64CS or pSP65CS is prepared from a saturated culture after lysis, phenol extraction, and isopropanol precipitation. This miniprep procedure yields DNA suitable for end labeling and chemical sequencing. The method is simple, rapid, requires a minimum number of reagents (lysozyme digestion is eliminated), and conserves tubes. Other miniprep methods such as the boiling protocol (*UNIT 1.6*) may be used if the final plasmid preparation is free of bacterial genomic DNA (which can interfere with the end-labeling reaction).

Materials

LB medium containing 40 µg/ml ampicillin (*UNIT 1.1*)
E. coli strain carrying recombinant pSP64CS or pSP65CS
Tris/EDTA/glucose buffer
2 N NaOH
4% SDS
3 M sodium acetate, pH 6.0
10 µg/ml RNase A, heat-inactivated (see reagents and solutions)
25:24:1 phenol/chloroform/isoamyl alcohol (*UNIT 2.1*)
Isopropanol
70% ethanol
18 × 150–mm tubes

1. Inoculate 5 ml LB medium containing 40 µg/ml ampicillin in 18 × 150–mm tubes with an *E. coli* strain carrying recombinant pSP64CS or pSP65. Grow until saturated (usually overnight) at 37°C.
2. Transfer 1.5 ml to a microcentrifuge tube. Microcentrifuge 2 min and discard supernatant.
3. Resuspend pellet in 120 µl Tris/EDTA/glucose buffer by vigorously vortexing.

It is important to resuspend the entire pellet because the reagents used in subsequent steps will not penetrate clumps of bacteria.

4. To the pellet, add 8 µl of 2 N NaOH and 8 µl of 4% SDS and shake vigorously (but do not vortex).
5. Immediately add 22 µl of 3 M sodium acetate, pH 6.0, and shake as in step 4. Immediately microcentrifuge 10 min.
6. Use a pipettor tip to remove and discard the pellet. Microcentrifuge supernatant 2 min and transfer supernatant to a clean microcentrifuge tube.
7. Add 2 µl of 10 µg/ml heat-inactivated RNase A and incubate 10 min at 37°C.
8. Extract twice with 100 µl phenol/chloroform/isoamyl alcohol.
9. Add 1 vol isopropanol to the upper (aqueous) phase, mix, and wait 5 min.

DNA Sequencing
7.3.5

10. Microcentrifuge 5 min at high speed and discard supernatant.
11. Wash pellet twice with 1 ml of 70% ethanol, microcentrifuging each time before removing ethanol, then dry the pellet.
12. Resuspend in 20 μ l sterile water for use in end labeling and sequencing (UNIT 7.5).

**BASIC
PROTOCOL**

**MINIPREP OF DOUBLE-STRANDED PLASMID DNA FOR
DIDEOXY SEQUENCING**

CsCl/ethidium bromide-purified closed-circular DNA (UNIT 1.7) can be used directly as a template for double-stranded sequencing. If such purified DNA is available, proceed directly to the fifth basic protocol for denaturation of plasmid DNA.

Plasmid DNA prepared from log-phase cultures by various miniprep procedures is also suitable as a template if treated to remove excess RNA. This can be accomplished by precipitating the RNA and single-stranded DNA with LiCl as described in this protocol. Alternatively, plasmid DNA can be treated with RNase A (UNIT 3.13), followed by phenol extraction and ethanol precipitation (UNIT 2.1), or treated as in the previous protocol.

Materials

- LB medium containing appropriate antibiotic (UNIT 1.1)
- E. coli* strain carrying recombinant plasmid
- Tris/EDTA/glucose buffer
- 1.0% SDS/0.2 N NaOH
- 3 M sodium acetate, pH 4.8 (APPENDIX 2)
- Isopropanol
- TE buffer, pH 8.0 (APPENDIX 2)
- 4 M LiCl
- Buffered phenol (UNIT 2.1)
- Chloroform
- Isopropanol
- 70% ethanol, ice-cold
- TE buffer, pH 8.0 (APPENDIX 2)

Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5)

1. Inoculate 5 ml LB containing appropriate antibiotic with *E. coli* strain carrying the recombinant plasmid. Grow until culture is in mid-log phase at 37°C.

It is very important to use a mid-log-phase culture ($OD_{560} \cong 1$). Do not use saturated or overnight cultures.
2. Transfer 1.5 ml to a microcentrifuge tube. Microcentrifuge 2 min and discard supernatant.
3. Resuspend pellet in 150 μ l Tris/EDTA/glucose buffer and leave 5 min at room temperature.
4. Add 300 μ l of 1.0% SDS/0.2 N NaOH and mix by inverting ~15 times (do not vortex). Leave 5 min at room temperature.
5. Add 225 μ l of 3 M sodium acetate, pH 4.8, and mix by inverting ~15 times (do not vortex). Leave 45 min on ice.
6. Microcentrifuge at high speed 5 min. Transfer 650 μ l of supernatant to a clean microcentrifuge tube.
7. Add 650 μ l isopropanol. Mix and let sit 10 min at room temperature.

8. Microcentrifuge at high speed 5 min at room temperature, then decant and discard supernatant. Dry pellet in desiccator.
9. Resuspend pellet in 125 μ l TE buffer, pH 8.0, by vortexing. Add 375 μ l of 4 M LiCl and let stand 20 min on ice.
10. Microcentrifuge 5 min at 4°C. Transfer supernatant to a clean microcentrifuge tube.
11. Extract with 500 μ l buffered phenol by vigorous vortexing. Microcentrifuge at high speed 2 min at room temperature and transfer upper (aqueous) phase to a clean microcentrifuge tube. Repeat this extraction using 500 μ l chloroform.
12. Precipitate DNA with 2 vol isopropanol for 30 min at room temperature.
13. Microcentrifuge 5 min at high speed and decant and discard supernatant.
14. Vortex pellet in ~1 ml ice-cold 70% ethanol. Microcentrifuge at high speed 5 min, then decant and discard supernatant. Dry pellet.
15. Resuspend pellet in 50 μ l TE buffer, pH 8.0. Electrophorese 5 to 10 μ l on an agarose gel to verify purity and approximate DNA concentration.
16. For sequencing, denature ~0.5 pmol (10 to 45 μ l) in alkali as described in the fourth basic protocol.

ALKALI DENATURATION OF DOUBLE-STRANDED PLASMID DNA FOR DIDEOXY SEQUENCING

**BASIC
PROTOCOL**

Theoretically, it should be possible to denature double-stranded templates either by treating with alkali or by boiling and achieve equal results in sequencing. In practice, however, alkali denaturation of closed-circular double-stranded templates and λ templates usually gives superior results for dideoxy sequencing. In contrast, for PCR products heat denaturation seems to yield better results in dideoxy sequencing reactions (see *UNIT 15.2*).

In the procedure described below, a recombinant plasmid is denatured using NaOH. After adjusting the pH to ≤ 7.0 , the DNA is precipitated with ethanol, washed, and dried. The dried pellet is suitable to be added to annealing reactions prior to DNA sequencing reactions (*UNIT 7.4*).

Materials

Recombinant plasmid DNA (third and/or fourth basic protocol)
 2 M NaOH/2 mM EDTA
 3 M sodium acetate, pH 6.0
 95% and 70% ethanol
 0.5-ml microcentrifuge tubes

1. Add ~0.5 pmol of recombinant plasmid DNA to a 0.5-ml microcentrifuge tube. If the volume is >20 μ l, ethanol precipitate the DNA (*UNIT 2.1*) and redissolve in 20 μ l water. If the volume is ≤ 20 μ l, add water to bring the volume to 20 μ l.
0.5 pmol of a 5-kbp plasmid is ~1.6 μ g.
2. Add 2 μ l of 2 M NaOH/2 mM EDTA and gently mix by drawing up and down with a pipet. Incubate 5 min at 25° to 37°C.
3. Place sample on ice, add 7 μ l water, and mix thoroughly by drawing solution up and down with the pipettor.
4. Add 7 μ l of 3 M sodium acetate, pH 6.0 (to neutralize DNA solution). Mix thoroughly by drawing solution up and down with the pipettor. Check the pH of the solution by

DNA Sequencing
7.3.7

spotting 1 μ l on pH paper. Add 3 M sodium acetate, pH 6.0, in 1- μ l increments until the pH is \leq 7.0.

5. Add 75 μ l of 95% ethanol and place 10 min on dry ice.
6. Microcentrifuge 10 min, 4°C, and carefully remove and discard the supernatant.
Use caution when removing the ethanol because the DNA pellet may also be removed.
7. Add 400 μ l 70% ethanol. Microcentrifuge 10 min at 4°C, then carefully remove and discard the ethanol layer.
8. Dry pellet 10 min in a Speedvac evaporator. Store the pellet at 20°C (up to several weeks) until used as template for dideoxy sequencing reactions (UNIT 7.4).

BASIC PROTOCOL

PREPARATION OF PLASMID DNA FROM AN *E. COLI* COLONY OR PHAGE DNA FROM A PLAQUE FOR THERMAL CYCLE SEQUENCING

Plasmid DNA is prepared from a single *E. coli* colony or phage DNA from a single plaque after resuspension in Tris/EDTA/proteinase K. The supernatant contains enough DNA for one set of thermal cycle sequencing reactions using a 5'-end-labeled primer (Krishnan et al., 1991; Young and Blakesley, 1991; Sears et al., 1992).

Materials

Agar plate containing *E. coli* colonies carrying a recombinant plasmid *or* λ
or M13 recombinant phage plaques
Tris/EDTA/proteinase K
Sterile toothpick or glass rod

- 1a. *For E. coli colony:* add 12 μ l of Tris/EDTA/proteinase K mix to a 1.5-ml microcentrifuge tube. With a sterile toothpick or glass rod, transfer a single *E. coli* colony carrying a recombinant plasmid to this solution and vortex 15 sec.
- 1b. *For phage plaque:* add 12 μ l of Tris/EDTA/proteinase K mix to a 1.5-ml microcentrifuge tube. With a sterile toothpick, transfer a single plaque to this solution and vortex 15 sec.

Try to remove as little top agar as possible.

2. Incubate 15 min at 55°C, then 15 min at 80°C.
3. Place 1 min on ice.
4. Microcentrifuge 3 min and immediately transfer supernatant to a clean 1.5-ml microcentrifuge tube.
5. Use 9 μ l as the template DNA in the end-labeled thermal cycle sequencing protocols (UNIT 7.4). Store remainder at -20°C.

REAGENTS AND SOLUTIONS

Lambda PEG solution

20% (w/v) PEG 6000
2.5 M sodium acetate, pH 6.0

M13 PEG solution

20% (w/v) PEG 800
2.5 M NaCl

RNase A, heat-inactivated

Dissolve RNase A at 10 mg/ml in 10 mM Tris·Cl (pH 7.5)/15 mM NaCl. Heat 10 min at 80°C, slowly cool to room temperature, and store in aliquots at 20°C. Dilute to 1 mg/ml or 10 µg/ml as required.

Tris/EDTA/glucose buffer

25 mM Tris·Cl, pH 8.0
10 mM EDTA
50 mM glucose

Tris/EDTA/proteinase K

10 mM Tris·Cl, pH 7.5
1 mM EDTA
50 µg/ml proteinase K

Prepare fresh for each use at the final concentrations listed above using stock solutions of 1 mg/ml proteinase K and TE buffer, pH 7.5 (APPENDIX 2).

COMMENTARY

Background Information

Each protocol in this unit provides a specific application of a general method for preparing DNA that has been proven to be effective for DNA sequencing. *UNITS 1.15, 1.13 & 1.6* provide information about M13-, λ-, and plasmid-based vector systems, respectively, and preparation of each type of vector DNA.

Critical Parameters and Troubleshooting

General troubleshooting information is provided in the units referred to above. When preparing phage from M13mp clones, vigorous aeration during phage growth is required to obtain high titers. If low phage titers are a problem, 100 mM MOPS buffer, pH 6.5, may be added to the 2× TY medium used for growing the phage. High titers of λ phage (>10⁹ pfu/ml) are also crucial for a good DNA yield.

Following precipitation of the M13mp or λ phage with PEG, it is very important to remove as much PEG as possible before resuspending the phage pellet. Contaminating PEG will result in high background in sequencing gels. Phenol remaining in the aqueous phase is readily removed by washing the ethanol precipitates with 100% ethanol. This also hastens drying

and reveals the single-stranded DNA as an arched film.

In addition to PEG, many other cellular and reagent contaminants can interfere with dideoxy sequencing reactions and with the end-labeling reaction for chemical sequencing. In general, it is crucial to avoid contamination by cellular chromosomal DNA and RNA and to thoroughly wash all ethanol pellets. The M13mp protocol contains an optional PEG precipitation and chloroform extraction, which should be added to the procedure if DNA of insufficient quality for sequencing is obtained.

A common problem with double-stranded DNA preparations is the occurrence of nicks in the DNA, which appear to act as priming sites for DNA polymerase in dideoxy sequencing reactions. This may result in a background of “shadow” or “ghost” bands. Plasmid preparations can be checked for nicked or linear molecules by agarose gel electrophoresis (*UNIT 2.5*). By using an *E. coli* host, which is EndoI⁻, and being careful to prevent nuclease contamination during and after the DNA preparation, nicks in the DNA may be avoided. However, some plasmids, because of their size or origin of replication, appear to be more susceptible to nicking. If this occurs during the preparations

described in this unit, CsCl/ethidium bromide centrifugation (UNIT 1.7) can be used to purify the DNA. An alternative procedure, particularly useful for λ DNA templates, is to use 5'-end-labeled primers so that the only labeled oligonucleotides are those primed by the end-labeled primer (see UNIT 7.4). Linearization of supercoiled plasmid preparations before alkali denaturation has been reported to improve the quality of the sequencing ladder (Wang and Sodja, 1991).

When preparing DNA from a plaque for use in thermal cycle sequencing, it is important to use agar that does not contain inhibitors of thermostable DNA polymerases (GIBCO/BRL #M29450 Luria agar has given satisfactory results). In addition, plaque size is critical for obtaining a strong sequencing signal from one plaque. A large plaque (e.g., 2 mm from λ gt10) is required for an overnight exposure; a small plaque (e.g., from λ gt22A) may require a longer exposure.

For use as templates in thermal cycle sequencing, plasmid DNA should be purified by CsCl gradient, miniprep methods (UNITS 1.6 & 1.7), or mini-column chromatography (UNITS 2.1 & 2.10). DNA obtained from single-stranded M13 should be purified by PEG precipitation (UNIT 1.15), mini-column chromatography, or solid support purification procedures (Hultman et al., 1991; Kaneoka et al., 1991; Zimmerman et al., 1989). λ and cosmid DNA should be purified using CsCl gradient protocols or miniprep methods (UNIT 1.13). To use PCR products as templates, remove excess primers and nucleotides from double- and single-stranded products after completing the reaction by glass bead procedures or exclusion methods (UNIT 2.1), gel purification (UNIT 2.6), or the drop dialysis method (Silhavy et al., 1984).

Anticipated Results and Time Considerations

UNITS 1.15, 1.13 & 1.6 provide general information about anticipated results and time considerations for preparation of M13, λ , and plasmid DNA, respectively. Double-stranded plasmid DNA can be denatured with alkali and prepared for annealing reactions in ~1 hr.

Literature Cited

- Adams, S. and Blakesley, R. 1991. Linear amplification sequencing. *Focus (BRL)* 13:56-57.
- Chen, E.Y. and Seeburg, P.H. 1985. Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. *DNA (N.Y.)* 4:165-170.

- Hattori, M. and Sakaki, Y. 1986. Dideoxy DNA sequencing method using denatured plasmid templates. *Anal. Biochem.* 152:232-238.
- Hultman, T., Stahl, S., Hornes, E., and Uhlen, M. 1989. Direct solid phase sequencing of genomic and plasmid DNA using magnetic beads as solid supports. *Nucl. Acids Res.* 17:4937-4946.
- Kaneoka, H., Lee, D.R., Hsu, K.-C., Sharp, G.C., and Hoffman, R.W. 1991. Solid phase DNA sequencing of allele specific polymerase chain reaction amplified HLA-DR genes. *BioTechniques* 10:30-40.
- Krishnan, B.R., Blakesley, R.W., and Berg, D.E. 1991. Linear amplification DNA sequencing directly from single phage plaques and bacterial colonies. *Nucl. Acids Res.* 19:1153.
- Sears, L.B., Moran, L.S., Kissinger, C., Creasey, T., Perry-O'Keefe, H., Roskey, M., Sutherland, E., and Slatko, B.E. 1992. CircumVent thermal cycle sequencing and alternative manual and automated DNA sequencing protocols using the highly thermostable Ventr (exo⁻) DNA polymerase. *BioTechniques.* 13:626-633.
- Silhavy, T., Berman, M., and Enquist, L. 1984. Experiments with Gene Fusions. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, N.Y.
- Wang, B. and Sodja, A. 1991. Alternate approach to sequencing double-stranded template DNAs. *BioTechniques* 10:198-201.
- Zagursky, R.J., Baumeister, K., Lomax, N., and Berman, M.L. 1985. Rapid and easy sequencing of large double-stranded DNA and supercoiled plasmid DNA. *Gene Anal. Tech.* 2:89-94.
- Zhang, H., Scholl, R., Browse, J., and Sommerville, C. 1988. Double-stranded DNA sequencing as a choice for DNA sequencing. *Nucl. Acids Res.* 16:1220.
- Zimmerman, J., Dietrich, T., Voss, H., Erfle, H., Schwager, C., Stegemann, J., Hewitt, N., and Ansorge, W. 1992. Fully automated sequencing protocol for double-stranded DNA. *Methods Mol. Cell. Biol.* 3:39-42.

Contributed by Barton E. Slatko
(λ templates, alkali denaturation, thermal cycle sequencing templates)
New England Biolabs
Beverly, Massachusetts

Peter Heinrich (double-stranded templates)
Consortium für Elektrochemische Industrie
Munich, Federal Republic of Germany

B. Tracy Nixon (single-stranded templates)
Pennsylvania State University
University Park, Pennsylvania

Richard L. Eckert (pSP64CS/pSP65CS)
Case Western Reserve University
Cleveland, Ohio

DNA Sequencing by the Dideoxy Method

In the basic dideoxy sequencing reaction, an oligonucleotide primer is annealed to a single-stranded DNA template (prepared by one of the methods described in *UNIT 7.3*) and extended by DNA polymerase in the presence of four deoxyribonucleoside triphosphates (dNTPs), one of which is ^{35}S -labeled. The reaction also contains one of four dideoxyribonucleoside triphosphates (ddNTPs), which terminate elongation when incorporated into the growing DNA chain (see introduction to Chapter 7 for a general overview of dideoxy sequencing). After completion of the sequencing reactions, the products are subjected to electrophoresis on a high-resolution denaturing polyacrylamide gel and then autoradiographed to visualize the DNA sequence (protocols for pouring, running, and processing sequencing gels are provided in *UNIT 7.6*). Reading and interpretation of sequencing gels are discussed (see Commentary).

Three variations of the dideoxy sequencing procedure are currently in use and are presented in this unit. In the “labeling/termination” procedure (see Basic Protocol 1), primer chains are initially extended and labeled in the absence of terminating ddNTPs, whereas in the traditional “Sanger” procedure (see Basic Protocol 2), labeling and termination of primer chains occur in a single step (see Fig. 7.0.2). The Commentary provides a more detailed comparison of the two protocols and recommendations for their use. A variety of DNA polymerases, including modified T7 DNA polymerase (Sequenase), *E. coli* DNA polymerase I large fragment (Klenow fragment), and thermostable enzymes are commercially available and may be used with either protocol (see Commentary). Instructions for substituting one polymerase for another are provided in Alternate Protocols 2 and 3.

A recent variation of the dideoxy sequencing method is thermal cycle sequencing (see Basic Protocol 3 and Alternate Protocol 6). In this procedure, the reaction mixture, containing template DNA, primer, thermostable DNA polymerase, dNTPs, and ddNTPs, is subjected to repeated rounds of denaturation, annealing, and elongation steps. The resulting linear amplification of the sequencing products allows much less template DNA to be used and eliminates independent primer annealing and template denaturation steps, which are required for the labeling/termination or Sanger procedures (see Commentary for a detailed discussion of thermal cycle sequencing).

In each of the basic protocols, substitution of a nucleotide analog, such as 7-deaza-dGTP, 7-deaza-dATP, or, in some cases, dITP, for dGTP in the nucleotide mixes can help to destabilize structures that can otherwise form in the sequencing products during electrophoresis to cause gel “compressions” (see Commentary). In addition, in some protocols manganese can be substituted for magnesium in the labeling/termination reaction (Alternate Protocol 1). Manganese increases the band uniformity exhibited by Sequenase and can increase the intensity of the sequencing ladder near the primer.

Each of the basic protocols utilizes [α - ^{35}S]dATP to label the nascent chain. A 5'-end-labeled primer can be used in either a one-step noncycling procedure (Alternate Protocol 4) or in a thermal cycling procedure (Alternate Protocols 5 and 6). Nonradioactive detection methods are described in the Chapter 7 introduction.

Commercially available DNA sequencing kits provide most of the reagents required for the abovementioned protocols. These kits save a significant amount of startup time, although they may limit flexibility in troubleshooting and are somewhat more expensive than assembling components individually (see Table 7.4A.1). Automated fluorescent sequencers, which are used in four-color dideoxy DNA sequencing (Alternate Protocols 7 and 8), are also commercially available (see Table 7.4A.1).

An alternative to radioisotope labeling (chemiluminescent detection) can be found in *UNIT 7.4B*.

**BASIC
PROTOCOL 1**

**LABELING/TERMINATION SEQUENCING REACTIONS
USING SEQUENASE**

The labeling/termination sequencing protocol involves two steps (Tabor and Richardson, 1987b, 1990). In the labeling step, primed DNA synthesis is initiated in the presence of limiting concentrations of all four dNTPs, including [α - ^{35}S]dATP, and continues until one of the dNTP pools is depleted. At this point, the uniformly labeled DNA chains are distributed randomly in length from a few nucleotides to hundreds of nucleotides. In the second step, synthesis resumes in the presence of additional dNTPs and one ddNTP. Elongation of the DNA chains in this step is rapid and processive until termination occurs at specific bases after incorporation of the corresponding dideoxynucleotide. In this protocol, the average length of the radioactively labeled oligonucleotide products is modified by altering the concentration of dNTPs in the first step; however, it can also be regulated by altering the dNTP:ddNTP ratio in the termination reaction (see Critical Parameters).

This protocol uses Sequenase. The labeling/termination procedure can also be used with other polymerases; however, because each polymerase has different buffer and Mg^{2+} concentration optima, and each discriminates to a different extent against ddNTPs, the concentrations of these components must be modified in each case (see Commentary). Conditions for substitution of *Taq* DNA polymerase and Klenow fragment are provided in the alternate protocols that follow (see Alternate Protocols 2 and 3).

Materials

- 0.5 pmol single-stranded *or* denatured double-stranded DNA template (*UNIT 7.3*)
- 0.5 to 1 pmol/ μl oligonucleotide primer in water (store at -20°C ; *UNIT 2.11*)
- 10 \times Sequenase buffer (see recipe)
- Sequenase termination mixes (see recipe)
- Sequenase/pyrophosphatase mix (see recipe)
- Sequenase diluent (see recipe)
- Labeling mixes (see recipe)
- 10 mCi/ml [α - ^{35}S]dATP (500 to 1200 Ci/mmol)
- Stop/loading dye (see recipe)
- 65 $^\circ$ and 95 $^\circ\text{C}$ water baths

- 0.5-ml microcentrifuge tubes
- Heat-resistant microtiter plates (optional; Table 7.4A.1)

Anneal primer and template

For each single-stranded DNA template

1a. Mix the following in a 0.5-ml microcentrifuge tube:

- 0.5 pmol single-stranded DNA template
- 0.5 pmol primer
- 1 μl 10 \times Sequenase buffer
- H_2O to 10 μl .

Mix gently by pipetting up and down (avoid creation of bubbles). Incubate 6 min at 65 $^\circ\text{C}$, then 20 min at 37 $^\circ\text{C}$. Proceed to step 2.

1.6 μg of single-stranded DNA 10,000 nucleotides long corresponds to 0.5 pmol of template molecules. Approximately 2.8 ng of a 17-mer is 0.5 pmol of primer molecules. Incubation

at 65°C destabilizes secondary structures that may be present in the single-stranded DNA template preparation. For more stringent annealing conditions, incubate at 42° to 50°C after the 65°C incubation (see Critical Parameters).

For each double-stranded DNA template

1b. Resuspend a dried pellet containing 0.5 pmol denatured double-stranded DNA in the following mixture:

1 pmol primer
1 µl 10× Sequenase buffer
H₂O to 10 µl.

Mix gently by pipetting up and down (avoid creation of bubbles). Incubate 30 min at 37°C, then keep at this annealing temperature until ready to proceed to step 2.

1.6 µg of 5000-bp double-stranded DNA molecule corresponds to 0.5 pmol of template molecules. For a λgt11-derived template (50,000 bp), 0.2 to 0.3 pmol (6 to 10 µg) is sufficient (see Critical Parameters). Double-stranded DNA templates require twice as much primer as single-stranded DNA templates. For more stringent annealing conditions, incubate at 42° to 50°C instead of 37°C (see critical parameters).

Set up termination reactions

2. While the primer is being annealed to the template, label four microcentrifuge tubes A, C, G, and T for each template to be sequenced.

For simultaneous sequencing of a large number of templates, the reactions can be carried out in heat-resistant, round-bottom, 96-well microtiter plates with heat-resistant lids (Table 7.4A.1).

3. Add 2.5 µl each of A, C, G, and T Sequenase termination mixes to the bottom of the A, C, G, and T tubes, respectively.

Keep tubes closed because the small volumes evaporate rapidly when open to the air.

Carry out labeling reaction

4. Immediately before use, dilute the Sequenase/pyrophosphatase mix in Sequenase diluent to 1 to 2 U Sequenase/µl and keep on ice.

5. Add the following to the annealed primer and template:

2 µl labeling mix
0.5 to 1.5 µl 10 mCi/ml [α -³⁵S]dATP
2 µl diluted Sequenase/pyrophosphatase mix.

Table 7.4A.1 Suppliers of DNA Sequencing Kits, Reagents, and Equipment^a

Product	Supplier ^b
Automated fluorescent DNA sequencers	APB, BK, LI, MD, MJ, PE, VG
Heat-resistant microtiter plates (round-bottom, 96-well)	APB, BK, CO, SA, ST
Nucleotides, nucleotide analogs	BM, BR, ICN, LT, NEB, PR, USB
Radiolabeled nucleotides	APB, ICN, NEN
Sequencing kits	APB, BM, BR, IBI, LT, NEB, NEN, PE, PR, ST, USB

^aModified from Slatko, 1991b.

^bAbbreviations: APB, Amersham Pharmacia Biotech; BK, Beckman; BM, Boehringer-Mannheim; BR, Bio-Rad; CO, Costar; IBI, International Biotechnologies; ICN, ICN Biomedicals; LI, LICOR; MD, Molecular Dynamics; MJ, MJ Research; NEB, New England Biolabs; NEN, NEN Life Sciences; PE, Perkin-Elmer; PR, Promega; SA, Sarstedt; ST, Stratagene; USB, U.S. Biochemical; VG, Visible Genetics. Addresses of suppliers are listed in APPENDIX 4.

The total volume is 14.5 to 16 μl . Incubate 5 min at 25°C (room temperature).

Choose the labeling mix (see Reagents and Solutions) appropriate for the lengths of sequencing products that are desired (see Critical Parameters). The short labeling mix is formulated to generate sequences closer to the primer annealing site, while the long labeling mix is formulated to generate sequences further away from the primer annealing site. Reaction times can be extended to 20 min without detriment, although the reaction is complete within 5 min, by which time the nucleotide pools are exhausted. This step can be performed at 37°C; however, 25°C limits the processivity of the Sequenase, an advantage during the labeling step, and maximizes its half-life during the sequencing reaction.

At least 3 pmol of [α -³⁵S]dATP are required for reactions containing the short labeling mixes whereas 15 pmol must be added to reactions containing the long labeling mixes. Amounts greater than these are not incorporated during the labeling step, because the other three dNTPs become limiting. The concentration of [α -³⁵S]dATP at 10 mCi/ml with a specific activity of 1000 Ci/mmol is 10 μM . Thus for such a preparation, 0.3 and 1.5 μl are sufficient for the labeling step with the short and long labeling mixes, respectively (see Critical Parameters).

Carry out termination reaction

6. Add 3.5 μl of the labeling reaction mixture (from step 5) to the tube containing Sequenase termination mix A (from step 3). Mix the solution by gently pipetting up and down. Repeat this addition to the C, G, and T tubes, changing pipet tips each time. Incubate 5 to 10 min at 37°C.

Reactions are complete within 2 to 3 min but incubations can be extended to 30 min without problems, except when using dITP under some conditions (see discussion of pyrophosphorolysis in the Commentary).

7. Add 4 μl stop/loading dye.

³⁵S sequencing reactions may be stored for up to one week at -20°C before electrophoresis.

8. Heat samples 2 min in a 95°C water bath, then place on ice. Load 2 to 3 μl of each sample on a sequencing gel. Electrophorese the gel (UNIT 7.6) and read the sequence (see commentary).

Excessive boiling of the completed reactions in formamide/dye solution may cause DNA chain breakage and smeared bands on the sequencing gel. If repeated loadings are planned, remove a 3- μl aliquot of each reaction to heat before each loading.

ALTERNATE PROTOCOL 1

USING Mn²⁺ IN THE LABELING/TERMINATION REACTIONS

In the presence of Mn²⁺, Sequenase incorporates deoxynucleotides and dideoxynucleotides at the same rate (Tabor and Richardson, 1989b, 1990). The addition of Mn²⁺ to a sequencing reaction has two effects. First, dideoxynucleotides are incorporated at the same rate at each position, resulting in bands of nearly uniform intensity. This high degree of band uniformity is most advantageous for automated sequencing procedures. Second, the relative incorporation of dideoxynucleotides to deoxynucleotides is about four times higher in the presence of Mn²⁺ than in the presence of Mg²⁺. The effect is the same as increasing the ddNTP:dNTP ratio 4-fold, resulting in an increase in the intensity of bands near the primer.

The Mn buffer specified in this protocol should be used only with Sequenase. Using the short labeling mixes and Mn buffer, it should be possible to read from near the primer up to ~200 nucleotides. To read further, it is necessary to increase the amount of dNTPs in the labeling reaction (by using the long mixes) or to increase the ratio of dNTP:ddNTP in the termination mixtures (see Critical Parameters).

Mn buffer is effective with the mixtures described in the labeling/termination basic protocol when dGTP or the analog 7-deaza-dGTP are employed. However, when sequencing with dITP, the lane will be faint near the bottom of the gel, as Mn^{2+} reduces the discrimination against dITP to a greater extent than it does ddGTP. To use Mn buffer with the dITP mixes, the amount of ddGTP should be increased 5-fold (to 8 μM); all the other nucleotides should be kept at the same concentration. The bromphenol blue in the lanes containing Mn buffer will be more diffuse than in the absence of Mn buffer. This does not interfere with the resolution of the sequencing ladder.

Additional Materials (also see *Basic Protocol 1*)

Mn buffer: 0.15 M sodium isocitrate/0.1 M $MnCl_2$ (store up to 3 months at $-20^\circ C$)

To sequence using Mn buffer, carry out the labeling/termination basic protocol, making the addition at the step indicated below.

5. After adding labeling mix and label to the annealed primer and template, add 1 μl Mn buffer to the primer/template/labeling mixture prior to adding Sequenase.

10 \times Sequenase buffer is still used; the effect of Mn^{2+} is dominant. The amount of Mn buffer added is not critical; comparable results will be observed if 0.2 to 2.0 μl are added (i.e., the effects are “all-or-none”; intermediate results are not observed). Do not mix Mn buffer with other reagents prior to use because Mn^{2+} is prone to oxidation, forming a yellow-brown precipitate.

USING OTHER POLYMERASES IN THE LABELING/TERMINATION REACTIONS

Labeling/Termination Reactions Using *Taq* DNA Polymerase

Taq DNA polymerase is useful for templates exhibiting secondary structures that may inhibit elongation by Sequenase (see Troubleshooting). Using *Taq* DNA polymerase, the reactions can be performed at temperatures high enough to destabilize many secondary structures (see Background Information).

Additional Materials (also see *Basic Protocol 1*)

10 \times *Taq* sequencing buffer: 500 mM Tris·Cl, pH 9.0 (APPENDIX 2)

Taq termination mixes (see recipe)

Taq DNA polymerase

45 $^\circ C$ water bath

To use *Taq* DNA polymerase in the labeling/termination reactions, modify the steps in Basic Protocol 1 as indicated below.

1. Substitute 10 \times *Taq* sequencing buffer for Sequenase buffer in the annealing reaction.
3. Add A, C, G, and T *Taq* termination mixes to the labeled tubes.
4. Immediately before use, dilute *Taq* DNA polymerase to 1 U/ μl in 1 \times *Taq* sequencing buffer and keep on ice.
5. Substitute *Taq* DNA polymerase for Sequenase; incubate labeling reactions at 45 $^\circ C$.

The initial 30 sec of the labeling reaction should not be carried out at a higher temperature than the maximum annealing temperature for the specific primer (see Critical Parameters and see Troubleshooting).

6. After adding aliquots of the labeling reaction mixture to each tube containing the termination mix, incubate the termination reactions at 65 $^\circ C$.

**ALTERNATE
PROTOCOL 2**

DNA Sequencing

7.4A.5

Labeling/Termination Reactions Using Klenow Fragment

Klenow fragment can be substituted for Sequenase (Stambaugh and Blakesley, 1988; Mardis and Roe, 1989) in the labeling/termination protocol (see Basic Protocol 1) as described in the modified steps below.

Additional Materials (also see Basic Protocol 1)

- 10× Klenow sequencing buffer (see recipe)
- Klenow termination mixes (see recipe)
- Klenow fragment (UNIT 3.5)
- 37° to 42°C water bath

1. Substitute 10× Klenow sequencing buffer for Sequenase buffer in the annealing reaction.
3. Add A, C, G, and T Klenow termination mixes to the labeled tubes.
4. Immediately before use, dilute Klenow fragment to 1 to 2 U/μl in 1× Klenow sequencing buffer and keep on ice.
5. Substitute Klenow fragment for Sequenase and incubate the labeling reactions at 37° to 42°C for 5 to 10 min.

Both the labeling reactions and the termination reactions (step 6) can be carried out at temperatures up to 50°C. The initial 30 sec of the labeling reaction should not be greater than the maximum annealing temperature for the specific primer (see Critical Parameters and see Troubleshooting).

6. Incubate the termination reactions at 37° to 42°C for 5 to 10 min.

SEQUENCING BY THE SANGER PROCEDURE USING KLENOW FRAGMENT

The traditional Sanger method proceeds in two stages. First, primed DNA synthesis occurs in the presence of a mixture of dNTPs and ddNTPs; termination occurs when a ddNMP is incorporated. Second, a chase with high concentrations of dNTPs ensures that oligonucleotides that have not specifically terminated by incorporation of a dideoxynucleotide are elongated past the region of interest. This protocol is written for use with Klenow fragment. Other DNA polymerases can be substituted along with appropriate buffer and reaction mixes (see Commentary). Conditions for *Taq* DNA polymerase are provided in Alternate Protocol 2.

Materials

- 0.5 to 1 pmol single-stranded *or* denatured double-stranded DNA (UNIT 7.3)
- 0.5 to 1 pmol/μl oligonucleotide primer in water (store at -20°C; UNITS 2.11 & 2.12)
- 10× and 1× Klenow sequencing buffer (see recipe)
- Klenow Sanger mixes (see recipe)
- Klenow fragment (UNIT 3.5)
- 10 mCi/ml [α -³⁵S]dATP (500 to 1200 Ci/mmol)
- dNTP chase (see recipe)
- Stop/loading dye (see recipe)
- 0.5-ml microcentrifuge tubes
- 37° to 42°C water bath
- Additional reagents and equipment for labeling/termination sequencing reactions (see Basic Protocol 1)

1. Anneal primer to single-stranded or denatured double-stranded DNA template as described in steps 1a and 1b in the labeling/termination protocol (see Basic Protocol 1), using 10× Klenow sequencing buffer in place of 10× Sequenase buffer.
2. Prepare labeled tubes as in step 2 of the labeling/termination protocol (see Basic Protocol 1).
3. Add 3 μl each of A, C, G, and T Klenow Sanger mixes to the bottom of the A, C, G, and T tubes, respectively.
4. Immediately before use, dilute Klenow fragment in 1× Klenow sequencing buffer to 2.5 to 5 U/μl and keep on ice.
5. Add the following to the annealed primer and template:

2 μl 10 mCi/ml [α -³⁵S]dATP
1 μl diluted Klenow fragment.

Mix by pipetting the solution up and down. The total volume is 13 μl.

6. Add 2.5 μl of primer/template/Klenow polymerase mixture to the tube containing the A Klenow Sanger mix (from step 3). Mix by gently pipetting the solution up and down. Repeat the addition to the C, G, and T tubes with the appropriate Klenow Sanger mix, changing pipet tips each time. Incubate 10 min at 37° to 42°C.
7. Add 1.0 μl dNTP chase to each A, C, G, and T reaction and mix as above. Incubate 10 min at 37° to 42°C.
8. Carry out steps 7 and 8 of the labeling/termination protocol (see Basic Protocol 1), except add 6 μl stop/loading dye.

USING *TAQ* DNA POLYMERASE IN THE SANGER PROCEDURE

Sequencing reactions using *Taq* DNA polymerase can be carried out at temperatures high enough to destabilize many secondary structures that would otherwise inhibit elongation of the polymerase (see Commentary).

Additional Materials (also see Basic Protocol 2)

10× *Taq* sequencing buffer: 500 mM Tris·Cl, pH 9.0 (APPENDIX 2)
Taq Sanger mixes (see recipe)
Taq DNA polymerase
50° to 75°C water bath

To use *Taq* DNA polymerase in the Sanger procedure (see Basic Protocol 2), modify the steps in that protocol as indicated below.

1. Use 10× *Taq* sequencing buffer in the annealing reaction.
3. Add A, C, G, and T *Taq* Sanger mixes to the labeled tubes.
4. Immediately before use, dilute *Taq* DNA polymerase to 2.5 U/μl in 1× *Taq* sequencing buffer and keep on ice.
5. Substitute *Taq* DNA polymerase for Klenow fragment in the reaction mixture.
6. Incubate the sequencing reactions 10 min at 50° to 75°C.
The initial 30 sec of this sequencing reaction should be at a temperature no greater than the maximum annealing temperature of the primer to the template (see Critical Parameters).
7. Incubate the chase reactions 10 min at 50° to 75°C.

**ALTERNATE
PROTOCOL 4**

DNA Sequencing

7.4A.7

ONE-STEP SEQUENCING REACTIONS USING 5'-END-LABELED PRIMERS

5'-end-labeled primers are used primarily for sequencing very large double-stranded DNA templates (such as λ gt11) or for templates that have given less than optimum results with nascent chain labeling (Slatko, 1991a). Nicks in larger or partially degraded templates can act as priming sites in the labeling reaction, resulting in a high background on the sequencing gel. When a 5'-end-labeled primer is used, only bona fide products from elongation of the primer will be detected by autoradiography.

Because the primer is labeled prior to the sequencing reaction, only one step is required for extension of the primer in the presence of ddNTPs. Nucleotide mixes vary depending on the DNA polymerase used in the reaction. A chase reaction containing high levels of all four dNTPs is recommended when utilizing Klenow fragment. Sequencing ladders from 5'-end-labeled primers generally have a clean background and band uniformity that is limited by variations caused by polymerase-specific artifacts. Because radiolysis of a 5'-end-labeled primer results only in a nonradioactive fragment, the sequencing reactions and end-labeled primers are relatively stable, and can be stored for up to 1 month with good results.

Primers may be 5' end-labeled with ^{32}P or ^{35}S using T4 polynucleotide kinase. The following protocol provides enough end-labeled primer for ten double-stranded DNA templates. If more end-labeled primer is desired, add at least as many pmol of ATP as pmol of primer to the end-labeling reaction.

Additional Materials (also see Basic Protocol 1)

$[\gamma\text{-}^{32}\text{P}]\text{ATP}$ (3000 Ci/mol) or $[\gamma\text{-}^{35}\text{S}]\text{ATP}$ (1300 Ci/mmol)

Sequenase termination mixes (see recipe), *Taq* termination mixes (see recipe), or Klenow one-step sequencing mixes (see recipe)

DNA polymerase and appropriate 10 \times buffer: Sequenase (see Basic Protocol 1), *Taq* DNA polymerase (see Alternate Protocol 2), or Klenow fragment (see Alternate Protocol 3)

dNTP chase (for Klenow fragment; see recipe)

37 $^{\circ}$ to 42 $^{\circ}$ C water bath (for Klenow fragment or 50 $^{\circ}$ to 75 $^{\circ}$ C water bath (for *Taq* polymerase)

Additional reagents and equipment for end labeling (UNIT 3.10)

End label primer

Carry out 5'-end-labeling of the primer using $[\gamma\text{-}^{32}\text{P}]\text{ATP}$

- 1a. End label ~10.5 pmol primer in a 25- μl reaction as described in UNIT 3.10 for labeling by the forward reaction. Use ~11.5 pmol $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ (3000 Ci/mmol) and 10 U of T4 polynucleotide kinase (BSA may be omitted from reaction). Incubate 30 min, 37 $^{\circ}$ C.
- 2a. Terminate the reaction by incubating 5 min at 95 $^{\circ}$ C. Briefly microcentrifuge at high speed, room temperature. Proceed to step 3.

It is not necessary to remove the unused ATP from the reaction; it will not interfere with the sequencing reaction.

Carry out 5'-end-labeling of the primer using $[\gamma\text{-}^{35}\text{S}]\text{ATP}$

- 1b. End label ~10.5 pmol primer in a 25- μl reaction as described in UNIT 3.10 for labeling by the forward reaction. Use ~13 pmol $[\gamma\text{-}^{35}\text{S}]\text{ATP}$ (1300 Ci/mmol) and 5 U of T4 polynucleotide kinase (BSA may be omitted from the reaction). Incubate 4 hr at 37 $^{\circ}$ C, adding an additional 5 U of T4 polynucleotide kinase every hour.

$[\gamma\text{-}^{35}\text{S}]\text{ATP}$ is not used efficiently by T4 kinase.

- 2b. Terminate the reaction by incubating 5 min at 95°C. Briefly microcentrifuge at high speed, room temperature. Proceed to step 3.

It is not necessary to remove the unused ATP from the reaction.

Carry out sequencing reactions using 5'-end-labeled primers

3. Anneal the 5' end-labeled primer to the DNA template as described in steps 1a or 1b of the labeling/termination protocol (see Basic Protocol 1), using ~2.4 µl of the 25-µl end-labeling reaction (~1 pmol primer) per denatured double-stranded DNA template, and 10× buffer appropriate for the DNA polymerase being utilized.
4. Place 3.5 µl each of the appropriate A, C, G, and T mixes in the bottom of four microcentrifuge tubes labeled A, C, G, and T, respectively—use Sequenase termination mixes, *Taq* termination mixes, or Klenow one-step sequencing mixes.

For Sequenase, these termination mixes will yield a readable sequencing ladder of up to ~150 nucleotides from the primer. To sequence further (up to ~300 nucleotides), increase the ratio of dNTP:ddNTP in the termination mixes from 10:1 (as described) to 20:1, by either increasing the concentration of all four dNTPs or reducing the concentration of ddNTPs by one-half and keeping the dNTP concentration the same. To sequence even further, the dNTP:ddNTP ratio should be increased to 40:1 (see Critical Parameters for further discussion of factors affecting extension lengths).

5. Add ~2.5 µl of the annealed labeled primer/DNA mix from step 4 to each tube. Incubate 30 sec to 1 min at 37°C.
6. Dilute enzymes in appropriate diluent as follows:
 - Sequenase—1 to 2 U/µl final
 - Taq* DNA polymerase—1 to 2.5 U/µl final
 - Klenow fragment—2.5 to 5 U/µl final.
7. Add 1 to 2 µl diluted DNA polymerase to each of the tubes from step 5. Incubate 5 to 10 min at the temperature appropriate for the DNA polymerase being utilized:
 - Sequenase—37°C
 - Taq* DNA polymerase—50° to 75°C
 - Klenow fragment—37° to 42°C.
8. If using Klenow fragment, add 1 µl dNTP chase to each tube after carrying out step 7. Incubate 5 min at 37° to 42°C.

*A chase step is not necessary when using Sequenase or *Taq* DNA polymerase.*
9. Carry out steps 7 and 8 of the labeling/termination protocol (see Basic Protocol 1).

THERMAL CYCLE SEQUENCING REACTIONS USING α-LABELED NUCLEOTIDES

This method is useful for situations where 100 to 200 ng of M13-derived or plasmid DNA or 2 µg of λ DNA is available. The DNA is combined with an oligonucleotide primer, all four dNTPs (one of which is labeled), and appropriate ddNTP for each reaction. Thermal cycle sequencing is then carried out in the presence of Vent_R (exo⁻) DNA polymerase, and the products are resolved on a sequencing gel.

This protocol and Alternate Protocol 6 were developed for use with the thermostable Vent (exo⁻) DNA polymerase (Sears et al., 1992). *Taq* DNA polymerase or modified *Taq* DNA polymerase can also be used in a thermal cycle sequencing protocol with the appropriate buffer and nucleotide mixes. Kits are commercially available with each of these DNA polymerases (Table 7.4A.1).

**BASIC
PROTOCOL 3**

DNA Sequencing

7.4A.9

Materials

Vent_R (exo⁻) sequencing mixes (see recipe)
0.04 to 0.1 pmol single-stranded or double-stranded DNA template (UNIT 7.3)
0.6 to 1.2 pmol oligonucleotide primer (UNIT 2.11)
10× Vent_R (exo⁻) DNA sequencing buffer (see recipe)
3% (v/v) Triton X-100
10 mCi/ml [α -³⁵S]dATP, [α -³²P]dATP, or [α -³³P]dATP (500 to 3000 Ci/mmol)
2000 U/ml Vent_R (exo⁻) DNA polymerase
Mineral oil, sterile
TCS stop/loading dye (see recipe)
Thermal cycling apparatus

Set up sequencing reactions

1. Label four microcentrifuge tubes A, C, G, and T for each template to be sequenced.
2. Add 3 μ l each of A, C, G, and T Vent_R (exo⁻) sequencing mixes to the bottom of the A, C, G, and T tubes, respectively.

Combine template and primer

For each single-stranded DNA template

- 3a. Mix the following in a 0.5-ml microcentrifuge tube:

0.04 pmol single-stranded DNA template
0.6 pmol primer
1.5 μ l 10× Vent (exo⁻) sequencing buffer
1 μ l Triton X-100
H₂O to 12 μ l.

Mix the solution gently by pipetting up and down.

For each double-stranded DNA template

- 3b. Mix the following in a 0.5-ml microcentrifuge tube:

0.1 pmol double-stranded DNA template
1.2 pmol primer
1.5 μ l 10× Vent_R (exo⁻) sequencing buffer
1 μ l 3% Triton X-100
H₂O to 12 μ l.

Mix the solution gently by pipetting up and down.

It is not necessary to denature a double-stranded DNA template in this procedure.

Carry out sequencing reactions

Individually process each template/primer mixture through step 6. When all sets of reaction mixtures are complete, proceed to step 7.

4. Add to each template/primer mix:

1 to 2 μ l 10 mCi/ml [α -³⁵S]dATP, [α -³²P]dATP, or [α -³³P]dATP

Mix the solution gently by pipetting up and down.

5. Add 1 μ l of 2000 U/ml Vent_R (exo⁻) DNA polymerase to the tube from step 4.
6. Immediately add 3.2 μ l of the mix from step 5 to the tube containing Vent (exo⁻) sequencing mix A from step 2. Mix the solution by gently pipetting up and down.

Repeat this addition to the C, G, and T tubes, changing pipet tips each time. Place the reaction mixtures on ice.

- Overlay each reaction mixture with 1 drop of sterile mineral oil.
- Set thermal cycle apparatus for the following thermal cycle conditions:

20 cycles:	20 sec	95°C
	20 sec	55°C
	20 sec	72°C

Place the tubes in the thermal cycler.

See Commentary for more discussion of thermal cycle temperature settings.

Terminate sequencing reactions

- Add 4 μl TCS stop/loading dye to each tube beneath the mineral oil.

³⁵S and ³³P reactions may be stored at -20°C for several weeks, ³²P reactions up to 1 week.

- Heat samples 3 min at 85°C . Insert pipet tip underneath the oil covering the reaction and load 2 μl on a sequencing gel. Electrophorese the gel (UNIT 7.6) and read the sequence (see Commentary).

A (5-sec) microcentrifugation will help to fully separate the aqueous layer from the oil layer. If necessary, briefly touch the pipet tip to a tissue to remove residual oil before loading. Alternatively, remove the oil by using silicone oil to cover the reactions followed by precipitation of the reaction products from the oil (Ross and Leavitt, 1991) or by using Parafilm to remove the oil (Whitehouse and Spears, 1991). Another alternative is to use a "hot-top" apparatus, set 15°C hotter than the warmest step in the cycle reaction ($\sim 115^{\circ}\text{C}$), which eliminates evaporation and precludes the requirement for oil on the reactions. A "hot-top" apparatus may be a clothes iron (set just to "steam" setting) or an upside-down heat-block or heating-stirring apparatus. The authors emphasize that no sequence aberration has been observed due to overlay oil contaminating the reaction in the sequencing gel lane.

THERMAL CYCLE SEQUENCING REACTIONS USING 5'-END-LABELED PRIMERS

ALTERNATE PROTOCOL 6

This method allows detection of a few nanograms of M13-derived or plasmid DNA or 300 ng of λ DNA, much less than required in Basic Protocol 3. PCR products must have excess primers and nucleotides removed after completion of the PCR reaction by drop dialysis, gel purification, glass bead procedures, or by spin columns.

Additional Materials (also see Basic Protocol 3)

10 mCi/ml [γ -³²P]ATP or [γ -³³P]ATP (3000 Ci/mmol), or biotin or fluorescent conjugates (see Commentary)

Additional reagents and equipment for end-labeling as described for one-step sequencing reactions (see Alternate Protocol 5, steps 1a and 2a)

Preliminary step

- End label primer using [γ -³²P]ATP or [γ -³³P]ATP (3000 Ci/mmol) as described in steps 1a and 2a of Alternate Protocol 5.

Set up sequencing reactions

- Set up tubes and sequencing mixes as described in steps 1 and 2 of the basic thermal cycle sequencing protocol (see Basic Protocol 3).

Carry out sequencing reactions

Individually process tubes as described in the thermal cycle sequencing protocol (see Basic Protocol 3).

For each single-stranded DNA template

3a. Mix the following in a 0.5-ml microcentrifuge tube:

- 0.004 pmol single-stranded DNA template
- 0.6 pmol 5'-end-labeled primer
- 1.5 μ l 10 \times Vent_R (exo⁻) sequencing buffer
- 1 μ l 3% Triton X-100 solution
- H₂O to 14 μ l.

Mix gently by pipetting up and down.

For each double-stranded DNA template

3b. Mix the following in a 0.5-ml microcentrifuge tube:

- 0.01 pmol double-stranded DNA template
- 1.2 pmol 5'-end-labeled primer
- 1.5 μ l 10 \times Vent_R (exo⁻) sequencing buffer
- 1 μ l 3% Triton X-100 solution
- H₂O to 14 μ l.

Mix gently by pipetting up and down.

Carry out and terminate sequencing reactions

4. Add 1 μ l of 2000 U/ml Vent_R (exo⁻) DNA polymerase to the tube from step 3a or 3b. Mix gently by pipetting up and down.
5. Complete the sequencing reactions as described in steps 6 to 10 of the basic protocol for thermal cycle sequencing (see Basic Protocol 3).

CYCLE SEQUENCING USING FLUORESCENCE DYE-LABELED PRIMER (DYEPRIMER) OR TERMINATOR (DYETERMINATOR)

The following two protocols are provided as examples of automated fluorescent sequencing. Although the methods described here are for direct sequencing of large template DNA, such as bacterial artificial chromosome clones (BAC), protocols for sequencing M13 clones, plasmids, or PCR products are performed in the same format with reduced amounts of sequencing reagents and DNA templates. The protocols for these latter applications are readily available from the sequencing reagent provider with nucleotide termination mix, DNA polymerase, and reaction buffer already mixed together as a cocktail (see Table 7.4A.1).

ALTERNATE PROTOCOL 7

Sequencing with Energy Transfer Dye-Labeled Primers (ET Primer)

The DyePrimer sequencing system uses four reactions, each for a particular dideoxynucleotide, generating data with uniform peaks. Since all the DNA fragments extended from the primer carry a fluorescent tag, the fragments prematurely terminated at dNTPs will be detected by the automated sequencer as false stops in the DyePrimer procedure.

The ET Primer Sequencing Kit is available from Amersham Pharmacia Biotech.

Materials

ThermoSequenase reaction buffer: 260 mM Tris·Cl/65 mM MgCl₂, pH 9.5
ThermoSequenase nucleotide mix: 750 μM each dNTP + 2.5 μM of specific ddNTP
ET primer: 0.5 μM dye-labeled A or C primer; 1.0 μM dye-labeled G or T primer
1.5 U/μl ThermoSequenase (Amersham Pharmacia Biotech)
0.25 μg/μl BAC DNA
Formamide loading buffer (see recipe)
0.2-ml strip tubes
Thermal cycler
Automated DNA sequencer (e.g., Perkin-Elmer)

1. Combine the following reagents (per reaction) to make a cocktail:

0.4 μl ThermoSequenase reaction buffer
1.6 μl ThermoSequenase nucleotide mix
1.0 μl ET primer
1.0 μl ThermoSequenase.

Multiply the amount of reagents per reaction by the number of reactions planned.

2. Dispense 4 μl of each cocktail into 0.2-ml microcentrifuge tubes with rows labeled A, T, G, and C. Add 4 μl (~1 μg) of DNA/reaction to each tube.
3. Microcentrifuge samples 30 sec at 2500 × g to bring the reagents to the bottom of the tube. Set up the following thermal cycles:

1 cycle:	2 min	95°C
20 cycles:	5 sec	95°C
	10 sec	55°C
	60 sec	72°C

Cool to 4°C and hold.

4. Remove samples from thermal cycler and add 1 μl of glycogen and 100 μl of 100% ethanol to the “A” row only. Pool samples from the A, C, and G reactions down into the “T” row reaction for precipitation.
5. Centrifuge samples 30 min at 3000 × g, 4°C. Decant ethanol supernatant and vacuum-dry samples, then resuspend in 3 μl formamide loading buffer for running on the automated DNA sequencer.

For large scale sequencing, carry out cycle sequencing reactions in 96-well or 384 microplates (Automated DNA Sequencing: Chemistry Guide, Perkin-Elmer Applied Biosystems, 1998).

Sequencing with Energy Transfer Dye-Labeled Terminators (BigDye Terminators)

The DyeTerminator procedure uses a one-tube reaction, and false stops (i.e., the fragments terminated at dNTPs) do not carry a fluorescent tag and therefore are not detected by the automated sequencer. The sequencing peaks, however, are less uniform as compared to DyePrimer data, due to the discrimination of the polymerase with respect to the dye-labeled dideoxynucleotides.

The BigDye terminator sequencing kit is available from Perkin-Elmer Applied Biosystems and the ET terminator kit is available from Amersham Pharmacia Biotech.

**ALTERNATE
PROTOCOL 8**

DNA Sequencing

7.4A.13

Materials

Termination nucleotide mix (see recipe)
5× sequencing buffer: 400 mM Tris·Cl, pH 9.0/10 mM MgCl₂
AmpliTaq FS for terminators (Perkin-Elmer)
Primer
0.25 µg/µl BAC DNA
Centri-Sep spin columns (Princeton Separations)
Thermal cycler
Automated DNA sequencer (e.g., Perkin-Elmer)

1. Combine the following reagents (per reaction) to make a cocktail:

8.0 µl termination nucleotide mix
4.0 µl 5× sequencing buffer
0.5 µl AmpliTaq FS for terminators
4.0 pmol primer
~400 ng BAC DNA
Water to 40.0 µl final volume.

Multiply the amount of reagents per reaction by the number of reactions planned.

2. Set up the following thermal cycles:

1 cycle:	5 min	95°C
30 cycles:	30 sec	95°C
	20 sec	55°C
	4 min	60°C

Cool to 4°C and hold.

3. Remove excess terminators with Centri-Sep spin columns. Vacuum dry samples and resuspend in 2 or 4 µl formamide loading buffer for running on the automated DNA sequencer.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

NOTE: Unless otherwise specified, the sources for recipes for each protocol are as follows: Klenow labeling/termination reactions, modified from Stambaugh and Blakesley (1988) and Mardis and Roe (1989); Sequenase labeling/termination reactions, U.S. Biochemical; *Taq* labeling/termination reactions, modified from Promega Biotech; Klenow Sanger and Klenow one-step reactions, modified from New England Biolabs; *Taq* Sanger reactions, Slatko, 1991a; Vent_R thermal cycle sequencing reactions, modified from New England Biolabs (Sears et al., 1992). See APPENDIX 4 for addresses and phone numbers of suppliers.

dNTP chase

Prepare 0.25 mM each dATP, dCTP, dGTP, and dTTP in 10 mM Tris·Cl, pH 7.5 (APPENDIX 2)/0.1 mM EDTA using nucleotide stock solutions (see recipe). Store up to 1 year at -20°C

Formamide loading buffer

95% (v/v) formamide
0.09% (w/v) bromphenol blue
0.09% (w/v) xylene cyanol FF
Store up to 3 months at 4°C

Klenow one-step sequencing mixes

Concentrations of dNTPs and ddNTPs in the A, C, G, and T mixes are listed below in μM . Prepare all mixes in 10 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)/0.1 mM EDTA using nucleotide stock solutions (see recipe). Store up to 1 year at -20°C .

	A	C	G	T
dATP	4.4	44	44	44
dCTP	44	4.4	44	44
dGTP	44	44	4.4	44
dTTP	44	44	44	4.4
ddATP	50	—	—	—
ddCTP	—	50	—	—
ddGTP	—	—	50	—
ddTTP	—	—	—	50

Klenow sequencing buffer, 10 \times

100 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)

150 mM MgCl_2

75 mM dithiothreitol (DTT)

Store up to 1 year at -20°C

Klenow Sanger mixes

Concentrations of dNTPs and ddNTPs in the A, C, G, and T mixes are listed below in μM . Prepare all mixes in 10 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)/0.1 mM EDTA using nucleotide stock solutions (see recipe). Store up to 1 year at -20°C .

	A	C	G	T
dATP	—	—	—	—
dCTP	56	5.6	56	56
dGTP	56	56	5.6	56
dTTP	56	56	56	5.6
ddATP	22	—	—	—
ddCTP	—	300	—	—
ddGTP	—	—	300	—
ddTTP	—	—	—	450

To substitute dITP for dGTP, Bankier and Barrell (1983) recommend replacing dGTP with 4 \times as much dITP in each of the Klenow Sanger mixes and reducing ddGTP to 20 μM in the G mix. However, the authors have successfully substituted an equimolar concentration of dITP with no change in the ddGTP concentration (B.E. Slatko, unpub. observ.).

To substitute 7-deaza-dGTP for dGTP, replace dGTP with an equimolar concentration of 7-deaza-dGTP in each of the Sanger mixes (Mizusawa et al., 1986). The ddGTP concentration remains unchanged.

Klenow termination mixes

Concentrations of dNTPs and ddNTPs in the A, C, G, and T mixes are listed below in μM . Prepare all mixes in 10 mM Tris·Cl, pH 7.5 (APPENDIX 2)/0.1 mM EDTA using nucleotide stock solutions (see recipe). Store up to 1 year at -20°C .

	A	C	G	T
dATP	25	250	250	250
dCTP	250	25	250	250
dGTP	250	250	25	250
dTTP	250	250	250	25
ddATP	240	—	—	—
ddCTP	—	100	—	—
ddGTP	—	—	120	—
ddTTP	—	—	—	400

To substitute dITP or 7-deaza-dGTP for dGTP, replace dGTP with an equimolar concentration of dITP or 7-deaza-dGTP, respectively, in each termination mix; do not alter the concentration of ddGTP in the G mix.

Labeling mixes

To obtain sequencing reaction products that are relatively short (<350 bases), use the short mix. Prepare the long mix when relatively long (>150 bases) products are desired. Prepare labeling mixes for Sequenase and Klenow fragment in 10 mM Tris·Cl, pH 7.5 (APPENDIX 2)/0.1 mM EDTA using nucleotide stock solutions (see below). Prepare labeling mixes for *Taq* DNA polymerase in 0.1 mM EDTA. Store mixes up to 1 year at -20°C .

Short mix: 1.5 μM each dCTP, dGTP, and dTTP

Long mix: 7.5 μM each dCTP, dGTP, and dTTP

To substitute 7-deaza-dGTP for dGTP, replace dGTP with an equimolar concentration of 7-deaza-dGTP in each mix.

*To substitute dITP for dGTP in Sequenase labeling mixes, replace dGTP with 3 μM dITP in the short mix and 15 μM dITP in the long mix. To substitute dITP for dGTP in the Klenow labeling mixes, replace dGTP with an equimolar concentration of dITP in each mix (B. Slatko, unpub. observ.). Substitution of dITP is not recommended when using *Taq* DNA polymerase.*

Nucleotide stock solutions

Purchase (Table 7.4A.1) or prepare sequencing grade dNTPs (including dITP and 7-deaza-dGTP) and ddNTPs as concentrated 20 mM stock solutions in 0.1 mM EDTA. Determine the actual concentrations spectrophotometrically using the extinction coefficients in Table 3.4.3; the extinction coefficient for each ddNTP is the same as that of the corresponding dNTP. Store up to 1 year at -20°C .

10 \times Sequenase buffer

400 mM Tris·Cl, pH 7.5 (APPENDIX 2)

200 mM MgCl_2

500 mM NaCl

50 mM dithiothreitol (DTT)

Store up to 1 year at -20°C

Sequenase diluent

10 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
5 mM dithiothreitol (DTT)
0.5 mg/ml bovine serum albumin (BSA)
Store up to 1 year at -20°C .

Sequenase/pyrophosphatase mix

Pyrophosphorolysis, the reversal of the polymerase reaction, is a problem when sequencing with Sequenase under some conditions, particularly when dITP is used (see Commentary). Inorganic pyrophosphatase prevents pyrophosphorolysis, and thus should be included in all sequencing reactions using Sequenase, particularly when dITP is present.

Pyrophosphatase is included with U.S. Biochemical's Sequenase sequencing kit. A mixture of the two enzymes can be prepared and stored in 20 mM KPO_4 buffer (pH 7.4)/0.1 mM DTT/0.1 mM EDTA/50% glycerol at -20°C . Mix 10 ng of sequencing grade yeast inorganic pyrophosphatase (0.005 U) per 2 to 4 U of Sequenase (i.e., per set of reactions), keeping the Sequenase stock as concentrated as possible. These enzymes are stable together at least 1 year.

Sequenase termination mixes

Concentrations of dNTPs and ddNTPs in the A, C, G, and T mixes are listed below in μM . Prepare all mixes in 10 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)/0.1 mM EDTA using nucleotide stock solutions (see recipe). Store up to 1 year at -20°C .

	A	C	G	T
dATP	80	80	80	80
dCTP	80	80	80	80
dGTP	80	80	80	80
dTTP	80	80	80	80
ddATP	8	—	—	—
ddCTP	—	8	—	—
ddGTP	—	—	8	—
ddTTP	—	—	—	8
NaCl (mM)	50	50	50	50

To substitute dITP for dGTP, replace dGTP with 160 μM dITP in each termination mix and use 1.6 μM ddGTP in the G termination mix. Inorganic pyrophosphatase should be used in reactions with dITP (see Commentary).

To substitute 7-deaza-dGTP for dGTP, replace dGTP with 80 μM 7-deaza-dGTP in each termination mix; do not alter the concentration of ddGTP in G termination mix.

Stop/loading dye

Prepare in deionized formamide:

0.05% (w/v) bromphenol blue
0.05% (w/v) xylene cyanol
20 mM EDTA

A high-quality grade of formamide that does not require deionization is available from Fluka.

Taq Sanger mixes

Concentrations of dNTPs and ddNTPs in the A, C, G, and T mixes are listed below in μM . Prepare all mixes in 0.1 mM EDTA using nucleotide stock solutions (see recipe). Store up to 1 year at -20°C .

	A	C	G	T
dATP	2	2	2	2
dCTP	20	20	20	20
dGTP	20	20	20	20
dTTP	20	20	20	20
ddATP	150	—	—	—
ddCTP	—	500	—	—
ddGTP	—	—	222	—
ddTTP	—	—	—	520

To substitute 7-deaza-dGTP for dGTP, replace dGTP with an equimolar concentration of 7-deaza-dGTP in each termination mix (Brow, 1990; B.E. Slatko, unpub. observ.); do not alter the concentration of ddGTP in the G mix.

Taq termination mixes

Concentrations of dNTPs and ddNTPs in the A, C, G, and T mixes are listed below in μM . Prepare all mixes in 0.1 mM EDTA using nucleotide stock solutions (see recipe). Store up to 1 year at -20°C .

	A	C	G	T
dATP	8	80	80	80
dCTP	80	8	80	80
dGTP	80	80	8	80
dTTP	80	80	80	8
ddATP	120	—	—	—
ddCTP	—	55	—	—
ddGTP	—	—	16	—
ddTTP	—	—	—	100

To substitute 7-deaza-dGTP for dGTP, replace dGTP with an equimolar concentration of 7-deaza-dGTP in each termination mix; reduce the concentration of ddGTP in the G mix to 8 μM .

TCS stop/loading dye

Prepare in deionized formamide:

0.3 (w/v) bromphenol blue

0.3 (w/v) xylene cyanol

12 mM EDTA

Store up to 1 year at -20°C .

A high-quality grade of formamide that does not require deionization is available from Fluka.

Termination nucleotide mix for cycle sequencing

For use with BigDye terminator (Perkin-Elmer Applied Biosystems). Concentrations are 100 μM for dATP, dCTP, and dUTP (replace dTTP), 500 μM for dITP (replace dGTP). Concentrations for dye-labeled terminators are 0.11 μM ddATP, 0.16 μM ddCTP, 0.10 μM ddGTP, and 1.12 μM ddTTP.

10× Vent_R (exo⁻) sequencing buffer

100 mM KCl
100 mM (NH₄)₂SO₄
200 mM Tris·Cl, pH 8.8 (APPENDIX 2)
50 mM MgSO₄
Store up to 1 year at -20°C

Vent_R (exo⁻) sequencing mixes

Concentrations of dNTPs and ddNTPs in the A, C, G, and T mixes are listed below in μM. Prepare all mixes in 1× Vent_R (exo⁻) sequencing buffer using nucleotide stock solutions (see recipe). Store up to 1 year at -20°C.

	A	C	G	T
dATP	30	30	30	30
dCTP	100	37	100	100
dGTP	100	100	37	100
dTTP	100	100	100	33
ddATP	900	—	—	—
ddCTP	—	480	—	—
ddGTP	—	—	400	—
ddTTP	—	—	—	720

To substitute 7-deaza-dGTP for dGTP, replace dGTP with an equimolar concentration of 7-deaza-dGTP in each of the Vent_R (exo⁻) sequencing mix (Sears et al., 1992); do not alter the ddGTP concentration.

dITP is not recommended for use with Vent_R (exo⁻) DNA polymerase.

COMMENTARY

Background Information

Refer to the introduction to Chapter 7 and to UNIT 7.1 for a general discussion of dideoxy sequencing. For automated fluorescent sequencing (Alternate Protocols 7 and 8), refer to Automated DNA Sequencing: Chemistry Guide 1998, from Perkin-Elmer Applied Biosystems, which is available at <http://www2.perkin-elmer.com/ab/techsupp/377.html>. This guide contains detailed steps for performing fluorescent sequencing.

Important differences between the labeling/termination and Sanger procedures

Labeling/termination sequencing reactions are carried out in two steps. In the first step, the primer is extended in the presence of low concentrations of dNTPs, including [α -³⁵S] dATP, until one or more of the dNTP pools is depleted (labeling step). The limiting levels of dNTPs and a low reaction temperature reduce the processivity (see below) of Sequenase, increasing the number of chains that are extended and labeled. At the end of the labeling reaction the uniformly labeled fragments range in size from a few to several hundred nucleotides. In the

second step, synthesis resumes in the presence of additional dNTPs and ddNTPs (termination step). In the termination reaction, the high dNTP concentration and an increased reaction temperature render Sequenase processive, ensuring that the polymerase extends each chain without dissociation until the incorporation of a dideoxynucleotide (an average of ~120 nucleotides).

Two steps are also employed in the Sanger dideoxy sequencing reaction, but the purposes of the steps are different. In the first step, a pulse reaction extends the primer in the presence of [α -³⁵S]dATP, unlabeled dNTPs, and ddNTPs. Thus, both labeling and termination occur in the pulse. The second step, a chase, employs a high concentration of all four dNTPs and ensures that all extended primers that have not incorporated a ddNMP are extended past the region to be sequenced.

In the labeling/termination protocol, extension lengths can be modulated by altering the dNTP concentrations in the first step independently of reaction conditions in the termination step. As a consequence, this method provides a capability of producing longer prod-

ucts, on average, than the Sanger protocol, particularly when Sequenase is employed. This is an advantage when trying to maximize the amount of sequence information obtained from each template. It can be a disadvantage, however, when only the first few nucleotides of sequence information after the primer are desired, although this problem can be overcome by altering labeling mixes and/or sequencing buffer (see Critical Parameters).

Another advantage of the labeling/termination protocol is that only a single reaction is used for the labeling step. Two conditions of the labeling step (low temperature and a low concentration of dNTPs) reduce the processivity of the polymerase (see below), resulting in some dissociation of the enzyme from the primer. When dissociation occurs at a site where reinitiation during the termination step is impeded, an artifactual band will occur at an equal intensity in all four lanes, and is readily identifiable as an artifact. In the Sanger protocol, by contrast, because four separate sequencing reactions are used, artifactual bands due to pausing of the polymerase sometimes occur in only one or two lanes, due to slight differences in the amount of enzyme or mixture in each reaction. This may result in artifactual bands being erroneously interpreted as dideoxy termination sites.

For most sequencing projects, where maximizing the amount of sequence information obtained per template is desirable, the labeling/termination protocol is recommended. For situations where limited amounts of sequence information are required (such as verifying constructions, confirming mutations from oligonucleotide-directed mutagenesis, or confirming overlapping clones), the Sanger protocol is appropriate.

Thermal cycle sequencing

In thermal cycle sequencing, the reaction mixture containing template, a large excess of primer, dNTPs, ddNTPs, and a thermostable DNA polymerase is subjected to 20 to 30 cycles of denaturation, annealing, and elongation steps. Because the template DNA is utilized repeatedly, linear amplification of the sequencing products occurs, and only femtomoles of template DNA are required in the reaction. Thus, thermal cycle sequencing is particularly useful in situations where minimal amounts of DNA template are available. Sequence information can be obtained from individual bacterial colonies or phage plaques using 5'-end-la-

beled primers in the thermal cycle sequencing procedure (UNIT 7.3). Thermal cycle sequencing has been particularly useful for automated sequencing, which normally requires larger reaction mixtures than manual sequencing.

Thermal cycle sequencing eliminates the requirement for a separate annealing reaction as well as the requirement for denaturing double-stranded DNA templates prior to the sequencing reaction, saving considerable amounts of time. In addition, much less template DNA is required in the sequencing reaction than in traditional methods, because the template is amplified in the reaction. Finally, the rapid thermal cycling helps prevent sequencing problems due to reannealing of linear double-stranded templates such as PCR products. Because a thermostable DNA polymerase is required, this method offers the option of performing the elongation reaction at a high temperature to destabilize secondary structures that may be present in the DNA template (see discussion of premature termination below).

Thermal cycle sequencing is compatible not only with traditional radiolabeling methods as described in this unit, but also with nonradioactive labeling methods such as chemiluminescence (using 5'-end-biotinylated primers; see UNITS 3.18 & 3.19) or fluorescence (in automated sequencers using fluorescently end-labeled primers or ddNTPs; Sears et al., 1992). In the latter case, the mutant form of *Taq* DNA polymerase for DNA sequencing is the preferred enzyme to use, because it incorporates fluorescently tagged ddNTPs more efficiently than does Vent_R (exo⁻) DNA polymerase.

DNA polymerases for sequencing

The original Sanger protocol was developed utilizing the Klenow fragment of *E. coli* DNA polymerase I (Sanger and Coulson, 1978; Sanger et al., 1977, 1980). The labeling/termination protocol is a modification of the Sanger procedure designed to take advantage of the high processivity and elongation rate of Sequenase (Tabor and Richardson, 1987a, 1987b, 1989a, 1989b, 1990). Other DNA polymerases can be used in either protocol. The following discussion and Table 7.4A.2 and Table 3.5.1 contain a brief summary of the relevant properties of the DNA polymerases used for DNA sequencing. In addition to Sequenase and Klenow fragment, these include *Taq* and Vent DNA polymerases and other thermophilic DNA polymerases, and avian or murine reverse transcriptase.

Properties of DNA polymerases relevant for sequencing

3'→5' exonuclease. Some DNA polymerases, such as *E. coli* DNA polymerase I, native T7 DNA polymerase, and Vent DNA polymerase, have an inherent 3'→5' exonuclease activity that acts as an editor to remove misincorporated nucleotides or nucleotide analogs (Table 3.5.1). A 3'→5' exonuclease activity can be a problem for several reasons. First, it can make it difficult to label efficiently the synthesized fragments, since in the presence of low concentrations of dNTPs the exonuclease activity can operate more efficiently than the polymerase activity, resulting in degradation of the primer. Second, because the 3'→5' exonuclease activity serves a proofreading function, it can increase the discrimination against nucleotide analogs (e.g., dITP and 7-deaza dGTP) that are used in DNA sequencing by preferentially degrading primers that have incorporated the analogs. Third, it can increase the variability in band intensity because a polymerase will pause at some sequences, cycling between its polymerase and exonuclease activities, resulting in a turnover of dNTPs. This gives ddNMPs repeated opportunities to be incorporated, increasing the probability that they will be incorporated at such sites. Fourth, the 3'→5' exonuclease activity has sequence preferences for where it will remove a dideoxynucleotide, which will also result in an increase in the variability of bands on a sequencing gel.

5'→3' exonuclease. Some DNA polymerases, such as *Taq* and *E. coli* DNA polymerase I, also have a 5'→3' exonuclease activity. This can be a problem for DNA sequencing because with uniform labeling it will result in a heterogeneous 5' end (thus multiple bands). With 5' end-labeling, it will result in the loss of some labeled material altogether. The 5'→3' exonuclease activity of *E. coli* DNA polymerase I is not present in the Klenow fragment. In addition, several commercial versions of *Taq* DNA polymerase have the 5'→3' exonuclease activity removed by mutagenesis or posttranslational modification.

Processivity. Processivity refers to the ability of the enzyme to polymerize nucleotides on a DNA chain without dissociating from the chain (or the number of nucleotides polymerized per polymerase binding event). A nonprocessive polymerase undergoes many dissociation and reassociation cycles during chain growth; this can lead to a population of fragments terminated not by the incorporation of a dideoxynucleotide but rather by dissociation of

the polymerase at a site where reinitiation is impeded. This results in a background of radioactivity on the gel at positions that do not correspond to incorporation of a dideoxynucleotide. In contrast, with a highly processive polymerase, a greater percentage of chains will terminate with the incorporation of a dideoxynucleotide, increasing the signal-to-noise ratio in the sequencing gel.

Elongation rate. Elongation rate is the rate at which DNA synthesis occurs. A rapid rate of elongation is advantageous because if a polymerase synthesizes DNA at a slow rate (as does AMV reverse transcriptase), pause sites that are difficult to synthesize through are more likely to be a problem in the time frame of the sequencing reactions, leading to background bands on the sequencing gel.

Incorporation of nucleotide analogs. To be useful for sequencing, the DNA polymerase must be able to incorporate dideoxynucleotides. Polymerases differ dramatically in their ability to use ddNTPs. The degree of discrimination against ddNTPs determines the ddNTP:dNTP ratio required in the nucleotide mixes to achieve a sequencing ladder in the region of the gel that resolves fragments that differ by a single nucleotide. Klenow fragment and *Taq* DNA polymerase discriminate about 100-fold more strongly against ddNTPs than does Sequenase. As a consequence, these two enzymes require proportionally greater amounts of ddNTPs in the sequencing reactions. This is a factor to be considered in comparing the cost of DNA sequencing using the different protocols and enzymes. Other nucleotide analogs, such as dITP, 7-deaza-dATP (Pharmacia, 1991; Barr et al., 1986; Seela et al., 1989), 7-deaza-dITP, and 7-deaza-dGTP, are useful in destabilizing secondary structures that can cause “compressions” on the sequencing gel (see below); therefore the ability to efficiently incorporate these analogs is an advantage. In addition, the DNA polymerase must be able to incorporate α -thio-analogs (such as [α -³⁵S]dATP) to label the sequencing products with ³⁵S (see introduction to Chapter 7 for a discussion of the advantages of labeling sequencing products with ³⁵S compared to ³²P).

Band uniformity. If ddNTPs are discriminated against in a sequence-dependent manner, the intensity of adjacent bands in a sequencing ladder will vary. This is a problem for three reasons. First, it decreases the sensitivity of the sequencing method, which is limited by the ability to detect the faintest bands. Second, it can make it difficult to interpret whether a

Table 7.4A.2 Characteristics of DNA Sequencing Enzymes^a

Characteristic	T7	Klenow	<i>Taq</i>	Bst	Vent	AMV RT
3'→5' exonuclease	none	low ^b	none	none	none	none
5'→3' exonuclease	none	none	present ^c	none	none	none
Processivity	high	low	inter.	inter.	low	inter.
Elongation rate	high	inter.	inter.	high	inter.	low
Use of nucleotide analogs ^d						
dITP	Y	Y	N	Y	Y	ND
7-deaza-dGTP	Y	Y	Y	Y	Y	Y
Band uniformity	very good	poor	good	very good	good	fair
Sequencing reaction temperature (°C)	<55 ^e	<50	<70	<65	<75	<42

^aModified from U.S. Biochemical. This information is meant to serve as a guideline, not as a reference source. Some characteristics depend strongly on reaction conditions. Enzyme abbreviations: T7, genetically modified T7 DNA polymerase (Sequenase Version 2.0); Klenow, *E. coli* DNA polymerase I, large fragment; *Taq*, *Thermus aquaticus* DNA polymerase; Bst, *Bacillus stearothermophilus* DNA polymerase; Vent, *Thermococcus litoralis* DNA polymerase, exo⁻; AMV RT, avian myeloblastosis virus reverse transcriptase. Other abbreviations used in the table are: inter., intermediate; Y, yes; N, no; ND, not determined.

^bMutants of Klenow are available that lack a 3'→5' exonuclease (Derbyshire et al., 1988); however, sequencing results identical to that of normal Klenow fragment are obtained.

^cSome versions of *Taq* DNA polymerase have the 5'→3' exonuclease genetically or chemically removed.

^dThis entry refers to whether the enzymes use dITP and dGTP efficiently.

^eTermination reactions; labeling reactions <37°C.

relatively weak band results from artifactual dissociation of the polymerase from the growing chain before a dideoxynucleotide is incorporated or from a correct stop due to inefficient incorporation of a ddNMP. Third, it makes it more difficult to read DNA sequences several hundred bases from the primer, where bands become increasingly close together. A large difference in radioactive intensity of adjacent dideoxy-terminated fragments increases the difficulty in determining whether the bands are due to single or multiple termination sites.

Temperature range. There are several situations in which it is advantageous for a DNA polymerase to be active at high temperatures. First, performing sequencing reactions at high temperatures makes it possible to increase the stringency of the primer annealing reaction to reduce priming at secondary hybridization sites on the template. Second, if the secondary structure of the template impedes the progress of the polymerase, resulting in incorrectly terminated chains, sequencing reactions can be performed at a high temperature to destabilize the secondary structure. Third, protocols for “thermal cycle sequencing,” which allow sequencing of nanogram amounts of template (see introduction to Chapter 7), require a thermostable DNA polymerase.

Sequenase

A high level of 3'→5' exonuclease activity in the native form of T7 DNA polymerase makes it ineffective for DNA sequencing. However, the exonuclease activity can be selectively removed without affecting the polymerase activity, either by a chemical reaction that oxidizes amino acid residues essential for exonuclease but not polymerase activity (Sequenase Version 1; Tabor and Richardson, 1987a) or genetically by deleting 28 amino acids essential for the exonuclease activity (Sequenase Version 2; Tabor and Richardson, 1989a). The chemically modified T7 DNA polymerase has 0.1% to 1.0% the level of exonuclease activity of the native enzyme, while the genetically modified T7 DNA polymerase has <10⁻⁷% the exonuclease activity of the native enzyme. While either can be used, the genetically modified enzyme has a higher specific activity than the chemically modified enzyme and yields more stable dideoxynucleotide-terminated fragments. T7 DNA polymerase does not have a 5'→3' exonuclease activity.

Sequenase synthesizes DNA at a rapid elongation rate by a highly processive mechanism (Table 7.4A.2). Under the termination reaction conditions, it discriminates against ddNTPs by ~4-fold compared to dNTPs. Thus, although Sequenase produces bands of higher uniformity than those produced by Klenow fragment or

Taq DNA polymerase, there is about a 10-fold variation in the intensity of adjacent bands (Fig. 7.4A.1). When Mn^{2+} is included in the reaction mixture, Sequenase incorporates dideoxynucleotides at the same rate as deoxynucleotides and the bands are almost completely uniform (see Alternate Protocol 1; Tabor and Richardson, 1989b, 1990). This high degree of uniformity is most advantageous for automated sequencing procedures.

Sequenase utilizes dITP and 7-deaza-dGTP efficiently (Fig. 7.4A.1). Sequenase has a tendency to stall in sequencing reactions using dITP, resulting in sequencing ladders which have a higher frequency of bands in all four lanes, typically at G positions. (However, this should not be used as a criterion for identifying G residues in ambiguous cases.) One group has reported successful removal of these inappropriately terminated products by reaction with terminal deoxynucleotidyl transferase (see troubleshooting; Fawcett and Bartlett, 1990). Labeling reactions using modified T7 DNA polymerase should be kept below 25°C, both to reduce the processivity of the enzyme and to maintain its activity in the termination reaction.

However, termination reactions using Sequenase can be performed from 37° to 55°C.

Klenow fragment

Klenow fragment has a 3'→5' exonuclease activity. However, it is a much weaker activity than that found in the native T7 DNA polymerase (Tabor and Richardson, 1989a) and does not interfere with the use of Klenow fragment for DNA sequencing. The 5'→3' exonuclease activity present in native *E. coli* DNA polymerase I is absent from Klenow fragment. Klenow fragment is relatively nonprocessive and has an intermediate elongation rate (Table 7.4A.2). The low processivity contributes to a somewhat higher background and lower signal-to-noise ratio than that found in reactions carried out using Sequenase. Occasionally, Klenow fragment encounters difficulty in reinitiating DNA synthesis at particular sites, resulting in "ghost" or "shadow" bands. The most common examples of this are: authentic T bands accompanied by faint shadow bands in the G and C tracks; authentic A bands accompanied by faint shadow bands in the G and C tracks; authentic A bands accompanied by faint shadow bands in the T track; and occasionally,

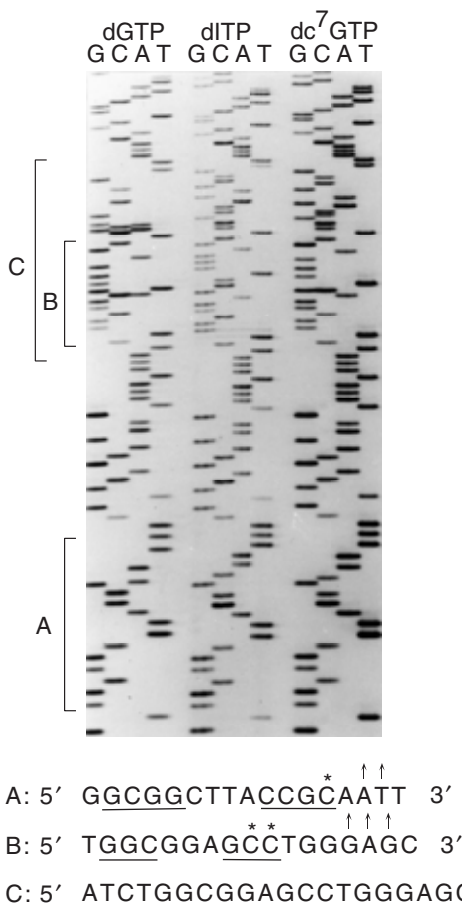


Figure 7.4A.1 Sequencing reactions using the labeling/termination protocol, Sequenase, [α -³⁵S]dATP labeling, and dGTP, dITP, or 7-deaza-dGTP (dc⁷GTP). Several compressions are present in the sequencing ladder produced with dGTP; each compression is bracketed, and the corresponding sequence is shown at the bottom of the figure. Regions of dyad symmetry are underlined, asterisks are above nucleotides corresponding to bands compressed in the presence of dGTP, and arrows are above "expanded" nucleotides. Electrophoresis was in a 7% acrylamide sequencing gel with wedge spacers (from Tabor and Richardson, 1987b; reprinted with permission of the authors).

authentic A bands accompanied by dark shadow bands in the other three lanes.

Klenow fragment discriminates strongly against ddNTPs, having a several-thousand-fold preference for a dNTP over the corresponding ddNTP. In addition, Klenow fragment exhibits a sequence-dependent variability in discrimination: adjacent dideoxy-terminated fragments can differ by as much as 50-fold in intensity. Listed below are the three most common examples of sequence-dependent discrimination in reactions using Klenow fragment: (1) The first A in a run of A's is usually the darkest (Fig. 7.4A.2, position 115). (2) The second C in a run of C's is always the most intense and the first C is very light; the third C is frequently dark as well (Fig. 7.4A.2, positions 59, 78, 90, and 106). (3) The second G in a run of G's is always very dark when the run of G's is preceded by a T (Fig. 7.4A.2, position 130).

Although the discrimination against ddNTPs is reduced over 100-fold by inclusion of Mn^{2+} (Tabor and Richardson, 1989b), a significant degree of sequence-dependent discrimination persists, producing bands that are still more variable in intensity than those resulting from reactions with Sequenase. 7-deaza-dGTP and dITP can each be used in sequencing

reactions with Klenow fragment. Sequencing reactions using Klenow fragment can be performed as high as 50°C.

Thermus aquaticus (Taq) DNA polymerase

Taq DNA polymerase lacks a 3'→5' exonuclease activity. Native *Taq* DNA polymerase has a 5'→3' exonuclease activity; however, a number of genetically engineered or posttranslationally modified versions of *Taq* DNA polymerase have recently been made commercially available in which the 5'→3' exonuclease activity has been removed. *Taq* DNA polymerase synthesizes DNA at an intermediate elongation rate with a moderate degree of processivity (Table 7.4A.2). It discriminates strongly against ddNTPs and requires ddNTP/dNTP ratios comparable to Klenow fragment. It exhibits band uniformity better than that of Klenow fragment but worse than that of Sequenase. Sequencing reactions carried out with *Taq* DNA polymerase generally produce a very clean background on the sequencing gel (Fig. 7.4A.3). 7-deaza-dGTP or 7-deaza-dATP can each be used in sequencing reactions with *Taq* DNA polymerase; however, dITP is usually not recommended for use with *Taq* DNA polymerase because of an unacceptably high frequency of inappropriately terminated chains

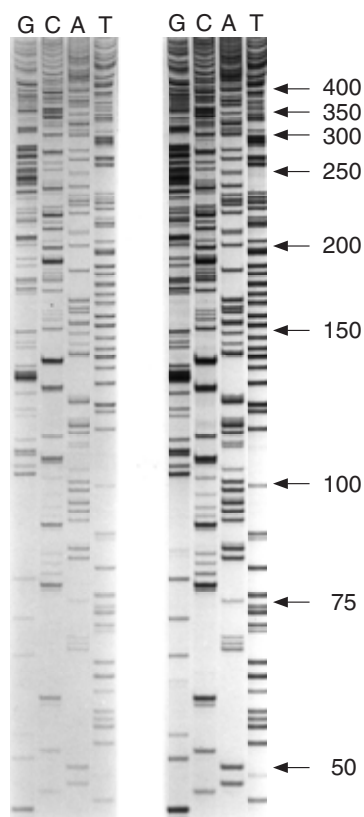


Figure 7.4A.2 Sequencing reactions using the Sanger procedure, Klenow fragment, and [α - 35 S]dATP labeling. The right panel was exposed for three times as long as the left panel. Electrophoresis was in an 8% sequencing gel at 60 W constant power for 1 hr (from Williams et al., 1986; reprinted with permission of *BioTechniques*).

(Fig. 7.4A.3). Sequencing reactions using *Taq* DNA polymerase can be performed at 55° to 70°C in low salt-buffer, conditions which promote destabilization of many template secondary structures (see troubleshooting; Innis et al., 1988; Peterson, 1988; Gyllensten, 1989; Brown, 1990).

Recent improvements on thermostable DNA polymerase have led to the development of mutant forms of *Taq* DNA polymerase—ThermoSequenase and AmpliTaq DNA polymerase FS (Tabor and Richardson, 1995)—which provide uniform termination efficiency and high-quality sequencing data in fluorescent cycle sequencing. ThermoSequenase contains a point mutation at the active site, in which phenylalanine is replaced with tyrosine at residue 667 (F667Y). This mutation results in less discrimination against dideoxynucleotides, and leads to a more even peak-intensity pattern. AmpliTaq DNA polymerase FS also contains a point mutation in the amino-terminal domain, in which glycine is replaced with aspartate at residue 46 (G46D), which removes almost all of the 5'→3' nuclease activity. This eliminates artifacts that arise from exonuclease activity.

Vent DNA polymerase

Vent_R DNA polymerase is purified from *E. coli* carrying the Vent DNA polymerase gene from the archaeobacterium *Thermococcus litoralis* (Perler et al., 1992; Kong et al., 1993). It lacks a 5'→3' exonuclease activity but has a 3'→5' exonuclease activity, giving it a high degree of fidelity that is useful for cloning. Vent_R (exo⁻) DNA polymerase is a genetically modified form of Vent_R DNA polymerase that has the 3'→5' exonuclease activity removed. This is the preferred form of the enzyme for sequencing. Like *Taq*, Vent_R (exo⁻) DNA polymerase synthesizes DNA at an intermediate elongation rate with a low degree of processivity (Table 7.4A.2; Jack et al., 1992). Vent (exo⁻) DNA polymerase has a relatively high *K_m* for dNTPs (~50 μM) and a relatively low *K_m* for DNA (~0.1 nM). Although the processivity of the enzyme is relatively low, under conditions of enzyme excess and appropriate concentrations of dNTPs and ddNTPs, the enzyme quickly binds and extends primers to create nascent dideoxy-terminated DNA chains (Sears et al., 1992; Jack et al., 1992; Kong et al., 1993).

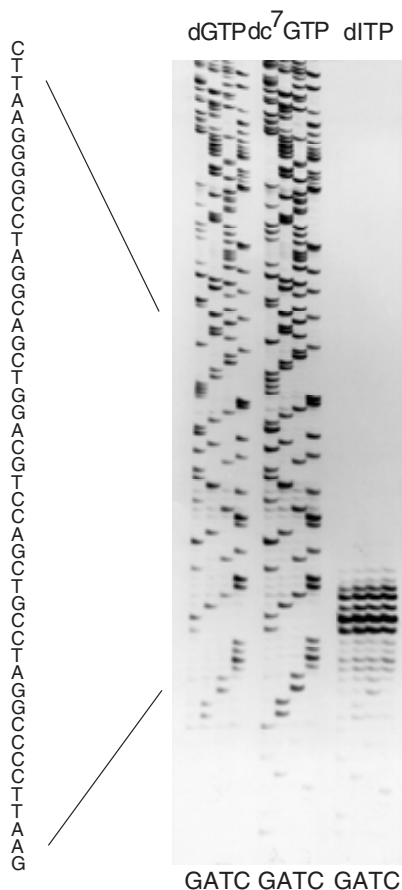


Figure 7.4A.3 The left side of the figure shows the sequence of the polylinker region of M13mp7, which includes a 48-base inverted repeat (palindrome). The right side of the figure shows the products of sequencing reactions using M13mp7 DNA template, *Taq* DNA polymerase (Δ *Taq* Version 2.0), the protocol and reagents from the TAQuence Version 2.0 DNA sequencing kit (U.S. Biochemical), and dGTP, 7-deaza-dGTP (dc⁷GTP), or dITP. The secondary structure of the template DNA may inhibit elongation of the DNA polymerase under some conditions; synthesis in the presence of dITP is stopped completely within the palindromic region. No such stopping is seen on other templates lacking inverted repeat sequences. The palindromic region can also exhibit strong compression artifacts on sequencing gels; electrophoresis in these examples was in the presence of 40% formamide and 7 M urea. A mild compression near the 3' end of the palindrome (GGGGAATTC) is still evident in lanes with samples from dGTP-containing reactions (compare to samples from 7-deaza-dGTP-containing reactions). (Photo courtesy of Carl Fuller, U.S. Biochemical.)

Vent_R (exo⁻) DNA polymerase discriminates strongly against ddNTPs and requires ddNTP/dNTP ratios comparable to those of Klenow fragment. Although it exhibits better band uniformity than with Klenow fragment, the following effects of sequence-dependent discrimination against ddNTPs (illustrated in Fig. 7.4A.4) are generally seen: (1) The second A of a run of A's is darker than the preceding (and/or following) A's. (2) The first C of a run of C's is darker than the following C's. (3) G following an A tends to be darker than other G's. (4) T following an A tends to be darker than other T's.

In sequencing reactions, 7-deaza-dGTP or 7-deaza-dATP can each be used with Vent_R (exo⁻) DNA polymerase; however, dITP is not efficiently utilized by the enzyme. Sequencing reactions using Vent_R (exo⁻) DNA polymerase can be performed at temperatures up to 80°C, and the enzyme is stable for several hours at temperatures up to 100°C (Kong et al., 1993; Perler et al., 1992).

Vent_R (exo⁻) activity is distinguished from that of *Taq* by three major differences. First, although both enzymes exhibit a comparable temperature optimum between 75° and 80°C, Vent_R demonstrates increased thermostability at higher temperature over extended periods of time compared with *Taq* (e.g., at 95°C, the

half-life of Vent_R is 7 hr versus 1.6 hr for *Taq*), thereby allowing the use of high denaturation temperatures and extended reaction times without significant loss of enzyme activity. Second, Vent_R exhibits lower processivity than *Taq*, although limiting the size of extension products in sequencing reactions with Vent_R has not been a problem. Third, the 3'→5' exonuclease activity associated with Vent_R (and lacking in *Taq*) confers a "proofreading" activity that allows the removal of a misincorporated nucleotide at the 3' end of the growing strand. Polymerases such as *Taq* that lack this activity make many (up to ten-fold) more errors during DNA synthesis. However, even without the 3'→5' exonuclease activity, Vent_R (exo⁻) exhibits an error rate that is two-fold lower than that of *Taq*.

Bst polymerase

Bst polymerase is isolated from the mesophilic bacterium *Bacillus stearothermophilus*. The large fragment of the enzyme, used for sequencing, is prepared by subtilisin digestion of the *B. stearothermophilus* holoenzyme, DNA polymerase I (Ye and Hong, 1987). The enzyme rapidly synthesizes dideoxy-terminated fragments and is very efficient at primer and template utilization; polymerization rate is rapid, and while processivity rates have not yet been determined, they are likely low to moder-

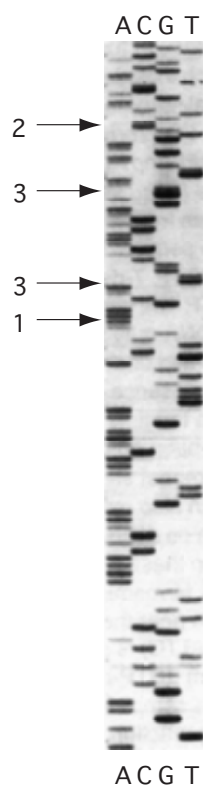


Figure 7.4A.4 The results of thermal cycle sequencing of a double-stranded PCR DNA template utilizing the Vent_R (exo⁻) DNA polymerase followed by chemiluminescent detection. The lane order is A, C, G, T. The effects of sequence-dependent discrimination against ddNTPs are shown: (1) in the A lane, position marker shows an example of the second A in a run of A's darker than the preceding (and/or following) A's; (2) in the C lane, position marker shows that the first C in a run of C's is darker than the following C's; (3) in the G, C, or T lanes, position markers show that a nucleotide following an A tends to be darker than other nucleotides. (Photo courtesy of Millipore Corp./New England Biolabs, Inc.)

ate, as has been observed for other thermostable polymerases (Jack et al, 1992). *Bst* polymerase has also been shown to be useful in both robotic and automated sequencer formats (Mardis and Roe, 1989). *Bst* polymerase, similar to other thermostable enzymes and *E. coli* DNA polymerase I Klenow (or large) fragment, discriminates strongly against ddNTPs, requiring an excess of ddNTPs relative to dNTPs in the respective sequencing mixes. The heat-stable enzyme permits DNA sequencing at elevated temperatures and provides a highly uniform distribution of band intensities, using either a Sanger (“one-step”) protocol or a labeling-termination type protocol (Mardis and Roe, 1989; McClary et al., 1991; Mead et al., 1991). While the temperature maximum for the sequencing reaction is about 70°C, the enzyme does not have the inherent heat stability to be useful for thermal cycling protocols. Nevertheless, *Bst* polymerase has been shown to be useful for sequencing templates with difficult G+C and “hairpin” palindromic regions (Ye and Hong, 1987). The higher reaction temperature helps alleviate problems due to premature termination (nondideoxy terminated fragments; see above). *Bst* polymerase is also able to incorporate base analogs such as 7-deaza dGTP and dITP to aid in the relief of gel compressions (see above), although the ratios of these nucleotides to ddGTP must be doubled, relative to that of dGTP.

Reverse transcriptase

Avian myeloblastosis virus (AMV) reverse transcriptase has been used for DNA sequencing in the past but is not now widely used. This enzyme lacks both 3'→5' and 5'→3' exonuclease activities. Compared to the other polymerases used for sequencing (Table 7.4A.2), it has an intermediate level of processivity and a low rate of elongation. The slow rate of elongation is a disadvantage for DNA sequencing as it is more difficult to obtain as incorporation of radioactivity into the fragments and background bands due to sites where the polymerase has paused are more frequent. AMV reverse transcriptase uses ddNTPs efficiently and requires ddNTP:dNTP ratios comparable to those for Sequenase. It exhibits slightly better band uniformity than does Klenow fragment. 7-deaza-dGTP can be used in sequencing reactions with reverse transcriptase. Sequencing reactions using AMV reverse transcriptase can be performed at 37° to 42°C (Mierendorf and Pfeffer, 1987; Stoflet et al., 1988; Zagursky et al., 1985; Bartlett et al., 1986; Krawetz, 1987).

Choosing a DNA polymerase for sequencing

Deciding which DNA polymerase to use for a sequencing project is sometimes as much a matter of convenience or personal preference as it is a decision based upon enzyme characteristics; in general, more than one polymerase can give reliable sequence information when used properly in a given protocol with clean templates. Conversely, since any of the DNA polymerases can produce artifactual sequence data under certain conditions, as the size of a sequencing project increases, so does the likelihood that no single enzyme will give 100% unambiguous data. A strategy that is usually effective consists of choosing one enzyme to generate the bulk of the sequence data and then switching to an alternate enzyme and/or protocol to resolve ambiguities.

For sequencing projects that do not require thermal cycling, use of Sequenase with the labeling/termination basic protocol is recommended because of its high degree of band uniformity, low background, and efficient use of ddNTPs and other nucleotide analogs. The only drawback is that Sequenase is not thermophilic. In situations where a thermostable DNA polymerase is required, *Taq* DNA polymerase or one of the other thermophilic DNA polymerases should be used.

Klenow fragment can also be used in the labeling/termination protocol. Klenow fragment is readily available from a large number of sources and has a long (and strong) track record of use in DNA sequencing.

In the Sanger protocol, Klenow fragment is generally the most appropriate enzyme to use; *Taq* DNA polymerase is a good alternative when template secondary structure causes premature termination of Klenow fragment. Template secondary structure most commonly occurs in stretches of DNA that are rich in G+C or A+T or that are extensively palindromic (see Troubleshooting). Reverse transcriptase has proven to be effective at sequencing through G-C-rich regions, while Klenow fragment has been effective at A-T-rich regions.

Using *Taq* DNA polymerase with a 5'-end-labeled primer in the one-step alternate protocol at 70°C is a good method for sequencing templates that are the products of the polymerase chain reaction (PCR; Chapter 15). In this case, the high temperature of the sequencing reactions not only destabilizes template secondary structure but also provides increased priming specificity. Vent_R (exo⁻) DNA polymerase, with the appropriate buffer and nucleotide mixes, can be used in place of *Taq* DNA po-

lymerase for sequencing reactions that require a high reaction temperature. [Nonthermal cycling protocols using Vent_R (exo⁻) DNA polymerase are available (Sears et al., 1992).]

For large sequencing projects, use of fluorescent cycle sequencing (Alternate Protocols 7 and 8) is recommended because of the simplicity of the procedure and the automation of data analysis.

Thermal cycle sequencing protocols obviously require a thermostable polymerase such as Vent_R (exo⁻)—as described in this unit—or the mutant forms of *Taq* DNA polymerase ThermoSequenase or AmpliTaq FS. Thermal cycle sequencing is particularly useful for sequencing small amounts of DNA or templates that tend to have secondary structures, such as PCR products (UNIT 15.2) or large double-stranded templates such as λ gt10, λ gt11, or cosmids.

Critical Parameters

For automated fluorescent sequencing (Alternate Protocols 7 and 8), refer to Automated DNA Sequencing: Chemistry Guide 1998, from Perkin-Elmer Applied Biosystems, which is available at <http://www2.perkin-elmer.com/ab/techsupp/377.html>. This guide contains detailed steps for performing fluorescent sequencing, as well as Critical Parameters and Troubleshooting information.

Primer/template ratio

In the labeling/termination and Sanger procedures, the optimal primer/template molar ratios are 1:1 for a single-stranded DNA template and 2:1 for a double-stranded DNA template. Increasing the ratio to >5:1 can result in the primer annealing to secondary sites on the template, resulting in extra “ghost” or “shadow” bands. However, a higher ratio of primer to template can be used with λ templates without detriment, possibly because the complexity of λ -derived vectors is generally greater than that of most plasmid-based templates, providing a larger pool of potential secondary annealing sites.

In the thermal cycle sequencing procedure, less template is required than in the labeling/termination or Sanger protocols because it is repetitively utilized during the reaction. However, primers are elongated into sequencing products and thus removed from the reaction, so primer is required in amounts comparable to those in a noncycling protocol. Thus the initial molar ratio of primer to template is much higher (15:1).

Annealing temperature

The maximum annealing temperature for primers that are 15 to 25 nucleotides long and that are used in the sequencing buffers in this unit can be approximated as $4 \times (\text{number of G and C residues}) + 2 \times (\text{number of A and T residues}) - 5$ (Van Zeeland et al., 1990). Thus, the maximum annealing temperature for a 20-mer, 50% G+C, is 55°C. If annealing times are longer than recommended in the sequencing protocols, the authors recommend annealing close to the maximum annealing temperature to prevent hybridization of the primer at secondary sites on the template. When sequencing reactions are performed at high temperatures with thermophilic DNA polymerases, the initial 30 sec of the labeling reaction in the labeling/termination protocol, or the sequencing reaction in the Sanger protocol, should not be at a temperature greater than the maximum annealing temperature for the specific primer.

In the thermal cycle sequencing procedure, the standard annealing temperature setting is 55°C; however, for primers that have a high A+T content or that are <20 bases, it may be necessary to adjust the primer annealing temperature to 50°C. Conversely, for primers have a high G+C content or that are >24 bases, it may be possible to skip the annealing step.

Thermal cycling conditions

If the amount of template DNA used in the thermal cycling protocol is larger than recommended, the number of cycles in the reaction may be reduced. However, when smaller amounts of DNA are used, increasing the number of cycles to >30 has not been successful; often “ghost” or “shadow” bands become more prominent.

The cycling regime may also be varied. The authors have performed the thermal cycling protocol using a two-step cycle, a 95°C denaturation step, and a 72°C annealing and extension step using primers with apparent melting temperatures as low as 55°C. The signal intensities under these conditions may be weaker than with the standard three-step protocol. Another way to perform two-step cycling is to lower the extension temperature to that of the annealing step, i.e., a 95°C denaturation and 55°C annealing and extension. Because of their high efficiency at ramping from one temperature to another, some thermal cycler machines may need to be set for longer reaction times (increasing each step from 20 sec to 1 min in each cycle) to achieve optimal results.

Factors affecting extension lengths in the labeling/termination protocol

For a given polymerase, three major parameters determine the average extension length of the synthesis reactions. The first factor is the concentration of dNTPs in the labeling reaction. Longer sequencing reaction products can be obtained simply by increasing the concentrations of nucleotides in the labeling mix. The two types of labeling mixes that have been provided in this unit are calibrated to give shorter (<350 bases) or longer (>150 bases) sequencing products.

The second factor determining the average extension length is the dNTP:ddNTP ratio in the termination reaction mixtures. Thus, the ratio of dNTP:ddNTP will determine in a predictable, statistical manner the average length of the extensions, depending upon the ability of the DNA polymerase to discriminate against the incorporation of dideoxynucleotides. Basic Protocol 1 uses a 10:1 ratio of dNTP to ddNTP. In the presence of Mg^{2+} , which is the metal ion used in the labeling/termination basic protocol, Sequenase discriminates ~4-fold against ddNTPs (Tabor and Richardson, 1989b). Thus, each chain is extended during the termination reaction an average of 120 nucleotides before incorporation of a dideoxynucleotide, although the range of extension lengths is quite wide. A higher dNTP/ddNTP ratio will result in correspondingly longer extensions while a lower ratio will result in terminations nearer the start of the primer.

The third factor is the amount of template DNA. During the labeling reaction, the primers are extended until the nucleotide pool is exhausted. Thus, the smaller the amount of template present in the reaction mixture, the further each primer will be extended on average. For M13 DNA templates or double-stranded plasmid DNA templates, ~0.5 pmol of template gives a broad distribution of radioactive fragments from near the primer to several hundred bases in length.

In summary, for normal labeling/termination reactions, the short labeling mixes are usually adequate. If only the first 100 nucleotides from the primer are of interest, four factors can be varied, to produce a greater percentage of terminations in this region: (1) the concentration of the three dNTPs in the short labeling mix can be reduced (e.g., 3- to 5-fold); (2) the time of the labeling reaction can be reduced (e.g., to 2 min); (3) the dNTP/ddNTP ratio in the termination mixes can be decreased (e.g., ~5-fold) by raising the ddNTP concentration);

and (4) the DNA concentration can be increased (e.g., 2-fold). It is not necessary to vary all of these factors; in most cases the size range of sequencing products can be closely regulated by varying just the concentration of the three dNTPs in the labeling reaction.

If the sequence just after the primer is the major region of interest, the intensity of the bands corresponding to this region can be increased by the addition of Mn^{2+} to the labeling/termination reactions (Alternate Protocol 1; Tabor and Richardson, 1989b, 1990). Use of Mn^{2+} reduces the discrimination against ddNTPs by Sequenase by ~4-fold and thus has the same effect as increasing the ddNTP to dNTP ratio by a factor of four. As a consequence, the average extension of each radioactive fragment in the termination reaction is reduced proportionately.

If sequences 300 to 1000 nucleotides from the primer are of primary interest, then either the concentration of the four dNTPs in the labeling reaction can be raised (by using the long labeling mixes and increasing the amount of [α - ^{35}S]dATP) or the dNTP:ddNTP ratio in the termination mixes can be increased (e.g., by a factor of 2 to 3).

Using the short and long labeling mixes, respectively, 3 pmol and 15 pmol of each nucleotide are added to the labeling reaction. The amount of [α - ^{35}S]dATP included should not be lower than these values so that it will not become the limiting nucleotide. Amounts greater than these are not incorporated in the labeling reaction since the other three dNTPs become limiting. However, the authors have found empirically that adding an excess of [α - ^{35}S]dATP (such as 40 pmol, or 2 μ l of 500 Ci/mmol at 10 mCi/ml) to the labeling reactions containing the short mixes results in darker bands on the sequencing gel. This is probably because the [α - ^{35}S]dATP unused in the labeling reaction contributes a significant proportion of the dATP in the termination reactions. It is recommended that only the appropriate amount of [α - ^{35}S]dATP be used in the labeling reactions to avoid generating more radioactive waste than necessary, and because the labeled nucleotide is a significant fraction of the cost of sequencing reagents.

Factors affecting extension length in the Sanger and thermal cycle sequencing procedures

Because extension and termination occur in the same step in the Sanger and thermal cycle sequencing procedures, the average length of

Table 7.4A.3 General Troubleshooting Guide for Dideoxy Sequencing

Problem	Possible cause	Solution
Blank autoradiogram	Inactive polymerase	Test using control primer, template, and reagents.
	Old label Old reagents Incorrect or defective primer Incorrect or defective template	Replace defective reagent.
Entire autoradiogram too light/poor incorporation of label	Incorrect exposure procedure (plastic wrap left on, gel facing backwards, etc.)	Correct error
	Old label	Obtain new label
	Low enzyme activity	Obtain fresh preparation of enzyme. Dilute immediately before use and keep on ice.
	Primer did not anneal well to template	See critical parameters for calculating maximum annealing temperature for primers. If template is double-stranded, remember to denature it in alkali before annealing (<i>UNIT 7.3</i>).
	Not enough primer	Test using control DNA and primer. Check concentration of primer. Use 0.5 pmol/single-stranded template, 1 pmol/double-stranded template.
	Not enough template	Test using control DNA and primer. Check template concentration on gel or by other method. Use ~0.5 pmol template DNA per set of reactions.
High background in all lanes	Incorrect exposure or development of autoradiogram	Reexpose gel. Change developing reagents, if necessary.
	Failure to remove plastic wrap from gel if using ³⁵ S	Remove plastic wrap
	Impure template DNA	See if problem persists with control DNA. If not, make new template DNA. For single-stranded templates, try optional annotations in <i>UNIT 7.3</i> . For double-stranded templates, treat with RNase A (<i>UNIT 3.13</i>), phenol extract and ethanol precipitate (<i>UNIT 7.3</i>). If necessary, CsCl-purify DNA.

continued

dideoxy-terminated fragments can only be adjusted by changing the ddNTP/dNTP ratio in the sequencing mixes. Changing the extension length in the Sanger reaction is most common when sequence information close to the primer is of interest. In this case the ddNTP:dNTP ratios can be increased ~5-fold, usually by increasing ddNTP concentrations. Conversely, decreasing the ddNTP concentration will effectively lengthen the sequence that can be read, while decreasing the frequency of stops at the bottom part of the sequence. If the template is particularly rich in one nucleotide, depletion of

the complementary nucleotide in that reaction can occur (because the dNTP concentration is lower in that reaction mix) and the sequencing ladder for that reaction may “fade out” at the top. In these cases, the ddNTP/dNTP ratio can be individually adjusted (e.g., by raising the dNTP concentration in that mix).

Troubleshooting

For automated fluorescent sequencing (Alternate Protocols 7 and 8), refer to Automated DNA Sequencing: Chemistry Guide 1998, from Perkin-Elmer Applied Biosystems, which

Table 7.4A.3 General Troubleshooting Guide for Dideoxy Sequencing, continued

Problem	Possible cause	Solution
Bands throughout the lanes are diffuse or fuzzy	Impure template DNA	See if problem persists with control DNA. If not, make new template DNA. For single-stranded templates, try optional annotations in protocol, <i>UNIT 7.3</i> . For double-stranded templates, treat with RNase A, phenol extract and ethanol precipitate as described in <i>UNIT 7.3</i> . If necessary, CsCl-purify the DNA.
	Poor quality of acrylamide gel	Prepare fresh acrylamide, bisacrylamide, and buffers using only high-quality reagents. Store stock solutions at 4°C in dark (<i>UNIT 7.6</i>).
	Gel used too soon after pouring	Let gel set up longer
	Buffer concentration in gel differs from concentration in reservoirs	Prepare gel and reservoir solutions from same 10× TBE stock
	Excessive boiling of samples	If samples are to be loaded on multiple gels, remove 3-μl aliquot of each reaction for heating and loading on gel
	Samples not denatured before running on gel	Heat samples in 95°C water bath 2 to 3 min
	Gel electrophoresis at too high a temperature	Monitor gel temperature with thermometer, keep temperature below 65°C. If necessary, run gel at lower wattage to keep temperature lower.
Areas on gel where bands are fuzzy	Gel or plastic wrap on top of gel dried with wrinkle	See <i>UNIT 7.6</i> , processing sequencing gels, for suggestions to avoid wrinkles
	Film was not clamped tightly to the gel	Insert filter paper behind gel to take up any extra space in X-ray cassette. Use more clamps, if necessary.
Distortion of all bands in 450-550-nucleotide region of the gel	>0.5% glycerol in samples (Tabor and Richardson, 1987b)	Dilute DNA polymerase in diluent without glycerol
Bands in all four lanes over entire gel	Low enzyme activity	Use fresh enzyme preparation; dilute immediately before use in appropriate diluent or 1× sequencing buffer
	Incorrect buffer or nucleotide mixes	Prepare new mixes
	Contaminated reagents	Prepare fresh reagents
Anomalous spacing of bands, missing bands, or bands at the same position in two or three lanes only at specific regions	Compression (due to secondary structure of newly synthesized DNA strands under conditions of gel electrophoresis)	See Commentary for a discussion of various methods for eliminating compressions
Bands at the same position in more than one lane throughout the gel	DNA preparation contains two different DNAs that are producing overlapping sequences	Prepare new DNA starting from a single plaque or colony
	Primer has annealed to secondary sites on template	See Critical Parameters for discussion of primer/template ratio and annealing temperature

Table 7.4A.4 Guide to Troubleshooting in the Labeling/Termination Protocol

Problem	Possible cause	Solution
Bands at the same position in all four lanes, especially near the primer	Impure template DNA ^a	See if problem persists with control DNA. If not, make new template DNA.
	Impure or old reagents; not enough dNTPs ^a	Prepare fresh reagents
	DNA polymerase activity low ^a	(1) Use fresh enzyme preparation. (2) Add more enzyme, e.g., four times as much per reaction. (3) Dilute enzyme immediately before use. (4) Reduce labeling reaction to 2 min to minimize the chance of T7 DNA polymerase dissociating from template.
Bands below a certain site on the gel are very dark, and bands above that are very faint	Secondary structure of the template impeding the extensions in the labeling step	Alter the conditions of the labeling reaction so that the extensions in the labeling reaction do not reach the secondary structure: (1) lower the concentration of the 3 dNTPs in the labeling mix, e.g., dilute 4-fold; (2) reduce the time of the labeling reaction to 1 or 2 min; (3) raise the DNA concentration 2- to 3-fold Use higher reaction temperature with alternate polymerase (and maybe alternate protocol) such as thermophilic polymerase. Refer to Commentary for discussion of alternate polymerases. Add 0.5 µg <i>E. coli</i> single-stranded DNA binding protein to the labeling reaction to help the polymerase synthesize through regions of secondary structure. The binding protein must be removed by treating with 0.1 µg of proteinase K at 65°C for 20 min after adding formamide/dye stop solution.
Intensity of bands at the bottom of gel is too low	Proportion of extensions that have terminated near the bottom of the gel is too low	Reduce the concentration of the labeling mix in the labeling reaction, e.g., dilute an additional 3- to 5-fold Reduce labeling time to 1 to 2 min Raise DNA concentration 2- to 3-fold Decrease ratio of dNTP to ddNTP in the termination reaction, e.g., raise the ddNTP concentration 2-fold Use a primer that lies further upstream of the site to be sequenced
Intensity of bands at the top of the gel is too low	The polymerase is incorporating a ddNMP before the DNA fragments have reached the appropriate length	Raise the concentration of the labeling mix in the labeling reaction (e.g. by using “long” mixes) Increase the dNTP/ddNTP ratio in the termination, e.g., reduce the ddNTP concentration in each termination mix 2-fold
Range of bands that can be sequenced is narrow (radioactivity is restricted to bands spanning 200 nucleotides)	DNA concentration in the sequencing reaction is low	Increase the DNA concentration in the sequencing reaction to 2 µg

^aThis may occur because of enzyme pausing during the labeling reaction and not reinitiating synthesis efficiently.

Table 7.4A.5 Guide to Troubleshooting in the Sanger Procedure

Problem	Possible cause	Solution
High background in “A” track, especially when [³⁵ S]dATP is older than one half-life	Old label	Replace with fresh label, or reduce label in reactions
Intensity of bands is too low at top of gel	Ratio of ddNTP to dNTP is too high	If problem is specific to one mix make up new nucleotide mix; otherwise decrease ddNTP/dNTP ratio by lowering ddNTP in each mix
Intensity of bands is too low at bottom of gel	Ratio of ddNTP to dNTP is too low	If problem is specific to one mix, make up new mix; otherwise increase ddNTP/dNTP ratio by increasing ddNTP in each mix

is available at <http://www2.perkin-elmer.com/ab/techsupp/377.html>. This guide contains detailed steps for performing fluorescent sequencing, as well as Critical Parameters and Troubleshooting information.

Refer to Table 7.4A.3 for a guide to troubleshooting general problems that are encountered with dideoxy sequencing, and to Table 7.4A.4, Table 7.4A.5, and Table 7.4A.6 for problems specific to the labeling/termination, Sanger, and thermal cycle sequencing procedures, respectively.

DNA preparations of poor quality are the most common problem encountered with sequencing reactions. Template impurities reduce the effectiveness of DNA polymerase in the sequencing reactions, resulting in higher levels of artifactual bands caused by premature termination. It is important to have highly purified control DNA for testing or troubleshooting reactions.

Denatured double-stranded DNA templates can give particularly troublesome backgrounds because these preparations are often contaminated with residual salts, proteins, and RNA. Moreover, nicks in the DNA can act as priming sites during the sequencing reaction. Using a 5′-end-labeled primer rather than uniform labeling with [α -³⁵S]dATP often overcomes the background problems associated with nicked templates because only elongated primers are detected (see Alternate Protocol 5). Empirically, the authors have found that the greater the percentage of supercoiled molecules in a plasmid preparation, the lower the background on the sequencing gel. Thus, for example, it is recommended that an *E. coli* strain which is

EndoI⁻ be used to propagate the plasmid. Somewhat paradoxically, there are several reports in the literature that linearization of circular double-stranded DNA templates reduced artifactual bands (Gravel et al., 1985; Mierendorf and Pfeffer, 1987; Wallace et al., 1981).

Interpretation of sequencing gels

Refer to Figures 7.4A.1 and 7.4A.3 for examples of how to “read” dideoxy sequencing gels. Autoradiograms are read from bottom to top, i.e., in the direction of synthesis (5′→3′). The spacing between oligodeoxynucleotides that differ in length by only one nucleotide gradually becomes narrower towards the top of the gel. To read more sequence data at the top of the gel, a shorter exposure time is often best, as this results in finer bands and a cleaner background. Use of a higher-contrast film such as Kodak OM-1, rather than the Kodak XAR-5 recommended in *UNIT 7.6* can also help; however, the exposure time is about three times longer for OM-1 than for XAR-5 for a comparable exposure.

Dark “shadow” or “ghost” bands may become more of a problem with long exposure times. By paying careful attention to spacing between oligodeoxynucleotides and considering enzymatic artifacts described in the previous section, it is usually possible to determine what is a shadow band and what is not.

Compressions

The largest single problem encountered in both the dideoxy and chemical sequencing methods is that of compressions in the banding pattern. Compressions are electrophoretic arti-

Table 7.4A.6 Guide to Troubleshooting in the Thermal Cycle Sequencing Protocols

Problem	Possible cause	Solution
Bands are seen at same position in all four lanes, especially near primer.	Impure template DNA	See if problem exists with control DNA. If not, repurify DNA.
	Incorrect primer annealing temperature	Reduce primer annealing temperature; use longer (more stable) primer; increase annealing and extension steps to 1 min each
	Impure or old reagents; ddNTP amount too high or dNTP amount too low	Prepare fresh reagents and readjust dNTP/ddNTP ratios
Bands below a certain site on gel are very dark, and bands above that are very faint.	Secondary structure of template impeding extensions	Use higher reaction temperatures
A dark band is present across all 4 lanes of gel; bands below and above it are equally dark.	Secondary structure of reaction products causing anomalous electrophoresis position of products in gel	Use higher gel temperatures or formamide gel; use base analogs or TdT in reaction; for PCR products, check purification procedure
Intensity of bands at bottom of gel is too low.	Proportion of reaction products terminating near bottom of gel too low	Increase ddNTP:dNTP ratio in sequencing mixes
Intensity of bands at top of gel is too low.	Proportion of reaction products terminating near top of gel too low	Decrease ddNTP:dNTP ratio in sequencing mixes
One lane of reaction failed or is weak or smeary.	Possible cyclor fault in one reaction slot	Check cyclor performance
	Pipetting error	Review chemistry procedure
All lanes are smeary or show high backgrounds.	Template impurity	Review template preparation; run control DNA template
	Primer impurity	Check primer purity
	Cyclor problem	Check cyclor function
	Reagent problem	Review chemistry procedure; make fresh reagents
	Primer/template ratio too high Incorrect priming temperature	Review chemistry procedure Calculate T_m as guideline for reaction
Entire sequence is light.	Insufficient template or primer in reaction	Double DNA and primer quantities in reaction
	Cyclor error; incorrect or inadequate cycling conditions	Check cyclor performance; make steps 1 min instead of 30 sec; increase cycles to 30
Oil covering reaction makes it difficult to load gel.	Excess oil present	Remove oil or use hot top-apparatus (see commentary); microcentrifuge reaction before loading gel

facts that occur when the bases of the single-stranded product of the sequencing reaction interact to form secondary structures, such as hairpins, that are stable even in the 7 M urea that is included in the sequencing gel. Secondary structures are most likely to form in regions of DNA that contain extended palindromic sequences or that are rich in G and C residues.

In practice, the migration of an oligonucleotide in a sequencing gel is affected by secondary structure only when the region of DNA capable of forming that structure is found at the 3' end of the oligonucleotide. This explains the observation that the migration of only a few adjacent fragments in the sequencing ladder is altered by a sequence that can form a secondary structure. The result is that any given compression occurs only in a limited region of the gel. Anomalous spacing of bands, extra bands, or missing bands are all important clues that a compression has occurred. A typical compression pattern consists of comigration of bands in different lanes and anomalously wide spacing of bands above this region (Fig. 7.4A.1). However, it is sometimes nearly impossible to identify a compression from the sequence of one strand alone.

A compression does not typically occur at the same site on both strands. When a compression does occur, the region that will be compressed on the opposite strand (if the complementary secondary structure is stable) will be on the other side of the secondary structure. This is because the secondary structure cannot form until the nucleotides involved in forming the secondary structure have been added to the growing oligonucleotide chain. Thus, the best criterion for identifying compressions is any discrepancy in the sequence between the two strands. Analysis of the discrepancies and the potential secondary structure of the neighboring sequence often reveals the site of compression and the true sequence of the region.

A variety of methods have been used to eliminate compressions. A very simple solution that may resolve some compressions is to run the gels at the highest possible temperature. This approach is limited, however, by the tendency of the glass gel plate to crack at high temperatures and by reduction in the resolution of the gel at high temperatures, leading to "fuzzy" bands.

Inclusion of formamide in the sequencing gel promotes denaturation of secondary structures and often solves many compression problems (Fig. 7.4A.3; Martin, 1987; Brown, 1984;

U.S. Biochemical, 1990). If this approach is tried, include 25% to 40% formamide in the acrylamide gel solutions (see UNIT 7.6).

A third method for eliminating compressions, not widely used, is to chemically modify C residues so that they can no longer engage in the formation of G-C base pairs. This can be accomplished by treating the synthesized oligonucleotides with either bisulfite or a mixture of bisulfite and methoxyamine (Ambartsumyan and Mazo, 1980; Hayatsu, 1976). These three procedures for eliminating compressions are equally applicable to enzymatic or chemical sequencing methods (UNIT 7.5).

For enzymatic sequencing, the use of nucleotide analogs that have a lower tendency to form secondary structures is frequently an effective method for eliminating compressions (Fig. 7.4A.1). Deoxyinosine-5'-triphosphate (dITP) has been used traditionally for this purpose (Mills and Kramer, 1979). Inosine is an analog of guanosine that bonds considerably more weakly to cytidine. The concentrations in the nucleotide mixes must be adjusted to compensate for the fact that DNA polymerases use dITP less efficiently than dGTP. dITP is substituted for dGTP in the nucleotide mixes but ddGTP is still used (not ddITP; see Alternate Protocols; Bankier and Barrell, 1983).

When Sequenase is used in sequencing reactions containing dITP, some fragments are susceptible to degradation by a reaction called pyrophosphorolysis (see below). The presence of inorganic pyrophosphatase prevents this degradation (Tabor and Richardson, 1990). Pyrophosphatase should always be present when dITP is used in the labeling/termination procedure with modified T7 DNA polymerase. A mixture of modified T7 DNA polymerase and yeast inorganic pyrophosphatase can be mixed together and stored in 50% glycerol at -20°C (see reagents and solutions). These enzymes are stable together for several years. Sequencing reactions using dITP should be performed in parallel with reactions containing dGTP because dITP-containing reactions tend to have banding artifacts caused by premature termination of the polymerase.

An alternative to dITP that is effective in reducing band compressions in some cases is 7-deaza-dGTP (7-deaza-2'-deoxyguanosine-5'-triphosphate; Seela et al., 1982; Mizusawa et al., 1986; Barr et al., 1986). 7-deaza-dGTP, which has a methine moiety at the N-7 position of the guanine nucleus, has been suggested to reduce compressions by disrupting the formation of alternative Hoogsteen base pairs

(Mizusawa et al., 1986; Sarecchi et al., 1970). However, it is possible that 7-deaza-dGTP also interferes with standard Watson-Crick base pairing by base pairing more weakly with cytidine. 7-deaza-dGTP is used by Sequenase, Klenow fragment, *Taq* polymerase, Vent_R polymerase, and reverse transcriptase more efficiently than dITP; it can be substituted at a 1:1 molar ratio for dGTP in the sequencing reactions (see alternate protocols). However, it may not be as effective as dITP at removing many compressions (Fig. 7.4A.1; Tabor and Richardson, 1987b). 7-deaza-dATP in combination with 7-deaza-dGTP may also be useful in resolving compressions (Pharmacia, 1991).

Premature termination

Stretches of single-stranded DNA that are rich in G + C or A + T or that are extensively palindromic can be particularly difficult to sequence accurately, not only because of stable secondary structure formation in the products on the gel (compressions) but also because of stable secondary structure in the template during the sequencing reaction. When bands in all four lanes are seen at or shortly before such a sequence and the sequencing ladder is very faint after that point, it is likely that the template secondary structure in the sequencing reaction has prevented the DNA polymerase from advancing past the “foldback” position (Fig. 7.4A.3). In such cases, using more polymerase can be helpful. Another strategy is to perform sequencing reactions at a higher temperature to destabilize the template secondary structure. The thermophilic DNA polymerases, such as *Taq* DNA polymerase, are the enzymes of choice in this situation (see above). An alternative strategy is to include *E. coli* single-stranded DNA binding protein in the sequencing reactions to destabilize the secondary structures (Table 7.4A.4). For templates with severe secondary structure problems, chemical sequencing (UNIT 7.5) can be employed.

Premature termination can also result in “ghost” or “shadow” bands or bands in all four lanes. This can arise from a variety of causes, such as impure template DNA, use of dITP (Fig. 7.4A.3), or performing reactions at a temperature at which the polymerase rapidly loses activity. One group has reported improvement of “shadow” bands when using Sequenase in the labeling/termination protocol by addition of terminal deoxynucleotidyl transferase (TdT; UNIT 3.6) to the sequencing reactions (Fawcett and Bartlett, 1990). TdT, in the presence of dNTPs, extends chains which have not been

terminated by incorporation of a ddNMP in a template-independent process to higher-molecular-weight DNA.

Pyrophosphorolysis

During DNA synthesis, the incorporation of dNMPs results in the accumulation of inorganic pyrophosphate. At a high enough level, the inorganic pyrophosphate can drive the reverse reaction of polymerization, called pyrophosphorolysis. Pyrophosphorolysis can lead to the disappearance of specific fragments on a DNA sequencing gel with some DNA polymerases (e.g., Sequenase; Tabor and Richardson, 1990). The removal of a dideoxynucleotide by pyrophosphorolysis leaves a chain with a normal 3'-hydroxyl group that is then rapidly extended by the DNA polymerase. This results in bands that are missing from the sequencing ladder or whose intensity is weak. Pyrophosphorolysis is most noticeable when using Sequenase, and particularly in reactions in which dITP has been substituted for dGTP. A simple solution to this problem is the inclusion of inorganic pyrophosphatase in the sequencing reaction. It is strongly recommended that inorganic pyrophosphatase be present in all sequencing reactions using Sequenase; under these conditions, all bands are stable to incubation of the sequencing reactions for at least 1 hr (Tabor and Richardson, 1990). Alternatively, when using Sequenase in the absence of inorganic pyrophosphatase, incubation of the termination reactions should be kept to a minimum (5 min), particularly when using dITP. Pyrophosphorolysis has also been observed with AMV reverse transcriptase-catalyzed sequencing reactions (Ruan et al., 1990).

Time Considerations

The dideoxy sequencing protocol can be carried out in a single long day. In general, the authors recommend that if a large region of DNA is to be sequenced, the sequencing should be carried out in a concerted effort. DNA sequencing is easier, more efficient, and more successful when a daily rhythm has been established.

A day's work usually starts with pouring the sequencing gels. Then, template-primer annealing reactions are set up. During the annealing reactions, there is usually time to develop the autoradiograms of the sequencing gels from the previous day. Next, the primer extension reactions are started. When these are complete, the reactions are loaded onto the gels. After lunch, the gels are stopped, dried down, and put on film overnight. The day ends by cleaning the

gel plates and setting them up for the next day. The whole procedure takes 7 to 8 hr for 10 templates, 9 to 10 hr for 20 templates.

It is possible to split the protocol into 2 days. On day 1, the sequencing reactions are carried out, stop/loading dye is added, and the mixtures are stored overnight at -20°C . The sequencing gels can also be poured and stored overnight (see UNIT 7.6). The next day, the samples are defrosted, heated to 95°C for 2 min, and loaded on the sequencing gels. Samples containing uniform-labeled ^{35}S can be stored up to one week. Samples containing uniform-labeled ^{32}P should be run on the gel the same day; in this case only gels can be prepared on day 1.

The same schedule may be followed for thermal cycle sequencing. Alternatively, the gel may be poured in the afternoon and the cycling reactions set up to be run overnight. The following morning, the gel is run and processed and the procedure from the previous day is repeated as the gel is analyzed.

Literature Cited

- Ambartsumyan, N.S. and Mazo, A.M. 1980. Elimination of the secondary structure effect in gel sequencing of nucleic acids. *FEBS Lett.* 114:265-268.
- Bankier, A.T. and Barrell, B.G. 1983. Shotgun DNA sequencing. In *Techniques in Life Sciences, B5: Nucleic Acid Biochemistry* (R.A. Flavell, ed.) pp. 1-34. Elsevier/North Holland, Limerick, Ireland.
- Barr, P., Thayer, R., Najarian, R., Seela, F., Laybourn, P., and Tolan, D. 1986. 7-deaza-2'-guanosine triphosphate: Enhanced resolution in M13 dideoxy sequencing. *BioTechniques* 4:428-432.
- Bartlett, J., Gaillard, P., and Joklik, W. 1986. Sequencing of supercoiled plasmid DNA. *BioTechniques* 4:208-210.
- Brow, M.A.D. 1990. Sequencing with *Taq* DNA polymerase. In *PCR Protocols: A Guide to Methods and Applications* (M. Innis, D.H. Gelfand, J.J. Snisky and T.J. White, eds.) pp. 189-196. Academic Press, San Diego.
- Brown, N. 1984. DNA sequencing. *Methods Microbiol.* 17:259-313.
- Derbyshire, V., Freemont, P.S., Sanderson, M.R., Beese, L., Friedman, J.M., Joyce, C.M., and Steitz, T.A. 1988. Genetic and crystallographic studies of the 3', 5'-exonucleolytic site of DNA polymerase I. *Science* 240:199-201.
- Fawcett, T.W. and Bartlett, S.G. 1990. An effective method for eliminating "artifact banding" when sequencing double-stranded DNA templates. *BioTechniques* 9:46-48.
- Gravel, R.A., Quan, F., and Korneluk, R.G. 1985. Rapid and reliable dideoxy sequencing of double-stranded DNA. *Gene* 40:317-323.
- Gyllenstein, U.B. 1989. PCR and DNA sequencing. *BioTechniques* 7:700-708.
- Hayatsu, H., 1976. Bisulfite modification of nucleic acids and their constituents. *Prog. Nucleic Acid Res. Mol. Biol.* 16:75-124.
- Innis, M., Myambo, K., Gelfand, D., and Brown, M. 1988. DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc. Natl. Acad. Sci. U.S.A.* 85:9436-9440.
- Jack, W., Kucera, R., and Kong, W. 1992. Biochemical characterization of Vent α and Deep Vent α DNA polymerase. *NEB Transcript (New England Biolabs)* 4:4-5.
- Kong, H., Kucera, R., and Jack, W. 1993. Characterization of a DNA polymerase from the hyperthermophile Archae *Thermococcus litoralis*. *J. Biol. Chem.* 268:1965-1975.
- Krawetz, S. 1987. DNA sequencing with AMV (avian myeloblastosis virus) reverse transcriptase and a modified T7 polymerase (Sequenase). *BioTechniques* 5:620-627.
- Mardis, E.R. and Roe, B.A. 1989. Automated methods for single-stranded DNA isolation and dideoxynucleotide DNA sequencing reactions on a robotic workstation. *BioTechniques* 7:840-850.
- Martin, R. 1987. Overcoming DNA sequencing artifacts: Stops and compressions. *Focus (BRL)* 9(1):8-10.
- McClary, J., Ye, S., Hong, G.F., and Whitney, F. 1991. Sequencing with the large fragment of DNA polymerase I from *Bacillus stearothermophilus*. *J. DNA Map. Seq.* 1:173-180.
- Mead, D.A., McClary, J.A., Luckey, J.A., Kostichka, A.J., Whitney, F.R., and Smith, L.M. 1991. *Bst* polymerase permits rapid sequence analysis from nanogram amounts of template. *BioTechniques* 11:76-87.
- Mierendorf, R.C. and Pfeffer, D. 1987. Direct DNA sequencing of denatured plasmid DNA. *Methods Enzymol.* 152:556-562.
- Mills, D.R. and Kramer, F.R. 1979. Structure independent nucleotide sequence analysis. *Proc. Natl. Acad. Sci. U.S.A.* 76:2232-2235.
- Mizusawa, S., Nishimura, S., and Seela, F. 1986. Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. *Nucl. Acids Res.* 14:1319-1324.
- Perler, F., Comb, D., Jack, W., Moran, L., Qiang, B.-Q., Kucera, R., Benner, J., Slatko, B., Nwankwo, D., Hempstead, K., Carlow, P., and Jannasch, H. 1992. Intervening sequences in an Archae DNA polymerase gene. *Proc. Natl. Acad. Sci. U.S.A.* 89:5577-5581.
- Peterson, M. 1988. DNA sequencing using *Taq* polymerase. *Nucl. Acids Res.* 16:10915.
- Pharmacia. 1991. New! 7-deaza-2'-deoxyadenosine triphosphate (c7 dATP) *Analecets* 19(4):3-4.
- Ross, J. and Leavitt, S. 1991. Improved sample recovery in thermo cycle sequencing protocols. *BioTechniques* 11:618-619.

- Ruan, C.C., Samols, S.B., and Fuller, C.W. 1990. Role of pyrophosphorolysis in DNA sequencing. *Comments (U.S. Biochemical)* 17(1):1-27.
- Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74:5463-5467.
- Sanger, F. and Coulson, A.R. 1978. The use of thin acrylamide gels for DNA sequencing. *FEBS Lett.* 87:107-110.
- Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H., and Roe, B.A. 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* 143:161-178.
- Sarecchi, M.-T., Courtois, Y., and Guschlbauer, W. 1970. Protonated polynucleotide structures: Specific complex formation between polycytidylic acid and guanosine or guanylic acids. *Eur. J. Biochem.* 14:411-421.
- Sears, L.E., Moran, L.S., Kissinger, C., Creasey, T., Perry-O'Keefe, H., Roskey, M., Sutherland, E., and Slatko, B.S. 1992. CircumVent thermal cycle sequencing and alternative manual and automated DNA sequencing protocols using the highly thermostable Vent_R (exo⁻) DNA polymerase. *BioTechniques*. 13:626-633.
- Seela, F., Tran-Thi, G.-H., and Franzen, D. 1982. Poly (7-deazaquanylic acid), the homopolynucleotide of the parent nucleoside of queuosine. *Biochemistry* 21:4338-4343.
- Seela, F., Berg, H., and Rosemeyer, H. 1989. Bending of oligonucleotides containing an isosteric nucleobase: 7-deaza-2'-deoxyadenosine replacing dA within d(A)₆ tracts. *Biochemistry* 28:6193-6198.
- Slatko, B.E. 1991a. Protocols for manual dideoxy DNA sequencing. In *Methods in Nucleic Acids Research* (J. Karam, L. Chao, and G. Warr, eds.) pp. 83-129. CRC Press, Boca Raton, Fla.
- Slatko, B.E. 1991b. Sources of reagents and supplies for dideoxy DNA sequencing and other applications. In *Methods in Nucleic Acids Research* (J. Karam, L. Chao, and G. Warr, eds.) pp. 379-392. CRC Press, Boca Raton, Fla.
- Stambaugh, K. and Blakesley, R. 1988. Extended DNA sequencing with Klenow fragment: The Kilobase Sequencing system. *Focus (BRL)* 10(2):29-51.
- Stofflet, E., Koeber, L., Sarkar, G., and Sommer, S. 1988. Genomic amplification with transcript sequencing. *Science* 239:491-494.
- Tabor, S. and Richardson, C.C. 1987a. Selective oxidation of the exonuclease domain of bacteriophage T7 DNA polymerase. *J. Biol. Chem.* 262:15330-15333.
- Tabor, S. and Richardson, C.C. 1987b. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 84:4767-4771.
- Tabor, S. and Richardson, C.C. 1989a. Selective inactivation of the exonuclease activity of bacteriophage T7 DNA polymerase by in vitro mutagenesis. *J. Biol. Chem.* 264:6447-6458.
- Tabor, S. and Richardson, C.C. 1989b. Effect of manganese ions on the incorporation of dideoxynucleotides by bacteriophage T7 DNA polymerase and *E. coli* DNA polymerase I. *Proc. Natl. Acad. Sci. U.S.A.* 86:4076-4080.
- Tabor, S. and Richardson, C.C. 1990. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase: Effect of pyrophosphorolysis and metal ions. *J. Biol. Chem.* 265:8322-8328.
- Tabor, S. and Richardson, C.C. 1995. A single residue in DNA polymerases of the *Escherichia coli* DNA polymerase I family is critical for distinguishing between deoxy and dideoxyribonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 92:6339-6343.
- U.S. Biochemical. 1990. Formamide gels (40%) for sequencing DNA. *Comments* 17(1):31.
- Van Zeeland, A., Vrieling, H., Simons, J., and Lohman, P. 1990. Sequence determination of point mutations at the HPRT locus in mammalian cells following in vitro amplification of HPRT cDNA prepared from total cellular RNA. In *Current Communications in Molecular Biology: Polymerase Chain Reaction* (H. Ehrlich, P. Gibbs, and H. Kazazian, eds.) pp. 119-124. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Wallace, R.B., Johnson, M.J., Suggs, S.V., Miyoshi, K., Bhatt, R., and Itakura, K. 1981. A set of synthetic oligodeoxyribonucleotide primers for DNA sequencing in the plasmid vector pBR322. *Gene* 16:21-26.
- Whitehouse, E. and Spears, T. 1991. A simple method for removing oil from cycle sequencing reactions. *BioTechniques* 11:616-618.
- Williams, S.A., Slatko, B.E., Moran, L.S., and DeSimone, S.M. 1986. Sequencing in the fast lane: A rapid protocol for [α -³⁵S]dATP dideoxy sequencing. *BioTechniques* 4:138-147.
- Ye, S. and Hong, G. 1987. Heat stable DNA polymerase I large fragment resolves hairpin structure in DNA sequencing. *Sci. Sin.* 30:503-506.
- Zagursky, R.J., Baumeister, K., Lomax, N., and Berman, M.L. 1985. Rapid and easy sequencing of large double-stranded DNA and supercoiled plasmid DNA. *Gene Anal. Tech.* 2:89-94.

Key References

- Tabor and Richardson, 1987b, 1990. See above.
Describes the procedure for DNA sequence analysis with Sequenase.
- Sanger et al., 1977. See above.
Describes the Sanger protocol using Klenow fragment.
- Sears et al., 1992. See above.
Describes the thermal cycle reaction.

Contributed by Barton E. Slatko
New England Biolabs
Beverly, Massachusetts

Lisa M. Albright
Reading, Massachusetts

Stanley Tabor
Harvard Medical School
Boston, Massachusetts

*Stanley Tabor is also a consultant for U.S.
Biochemical.*

Jingyue Ju (fluorescent sequencing)
Incyte Pharmaceuticals
Palo Alto, California

Dideoxy DNA Sequencing with Chemiluminescent Detection

Standard dideoxy DNA sequencing (UNIT 7.4A) can be performed easily and efficiently with nonisotopic, chemiluminescent detection by utilizing primers labeled with biotin in the sequencing reactions. Reaction products are separated by denaturing gel electrophoresis (UNIT 7.6), transferred to a nylon membrane, and detected by first binding a streptavidin–alkaline phosphatase conjugate, then incubating with a chemiluminescent 1,2-dioxetane substrate. The emitted light signal is imaged on standard X-ray film, producing high-resolution DNA sequencing ladders. This procedure is presented below (see Basic Protocol). Indirect alkaline phosphatase–labeling of biotinylated DNA with free streptavidin and biotinylated alkaline phosphatase is also detailed (see Alternate Protocol 1). Finally, the detection of sequencing reactions labeled with other haptens using specific antibody–alkaline phosphatase conjugates is described (see Alternate Protocol 2).

Most of the information in the Chapter 7 introduction and UNITS 7.1–7.3, and the methods described in UNIT 7.4A, are relevant to chemiluminescent sequencing methods. Primers derivatized with nonisotopic labels such as biotin can be utilized in protocols for DNA sequencing reactions designed for 5′-end-labeled primers. The products generated are subsequently detected by chemiluminescent methods. DNA sequencing reactions performed with Sequenase, Klenow fragment, *Taq* DNA polymerase, as well as thermal cycle reactions using Vent_R(exo[−]) polymerase, are described in UNIT 7.4A; sequencing using 5′-end-labeled primers and *Bst* polymerase is described here. Chemiluminescent DNA sequencing kits incorporating nonisotopically labeled primers and chemiluminescent detection reagents are commercially available (Table 7.4B.1). Biotinylation of oligonucleotide primers is possible using biotinylating phosphoramidites; alternatively, the labeled primers can be acquired from commercial sources at a relatively low cost. With standard cloning vectors, biotinylated versions of universal and many other primers are available and are usually included in nonisotopic DNA sequencing reaction kits.

DNA SEQUENCING USING BIOTINYLATED PRIMERS WITH CHEMILUMINESCENT DETECTION

BASIC PROTOCOL

In this protocol, DNA sequencing reactions are performed using biotinylated primers and *Bst* polymerase. Sequencing products are electrophoresed and DNA is transferred from the gel to a nylon membrane, then UV cross-linked to the membrane. It is incubated first with a streptavidin–alkaline phosphatase conjugate, then with a dioxetane detection solution before being exposed to X-ray film.

Alternatively, the 5′-biotinylated primers can be used in the end-labeled primer protocols of UNIT 7.4A; subsequent electrophoresis, transfer, and chemiluminescent detection is as described here.

Materials

- 1 μg single-stranded DNA (~0.5 pmole) or 1 to 3 μg denatured double-stranded DNA template
- Biotinylated DNA sequencing primer
- 5× *Bst* reaction buffer: 100 mM Tris·Cl, pH 8.5 (APPENDIX 2)/100 mM MgCl₂
- Primer termination mixes (A, C, G, T; see recipe)
- 1 U/μl *Bst* polymerase (Bio-Rad)
- 20× chase solution: 10 mM each of dATP, dCTP, dGTP, dTTP in H₂O

Contributed by Chris S. Martin

Current Protocols in Molecular Biology (1999) 7.4B.1–7.4B.14
Copyright © 1999 by John Wiley & Sons, Inc.

DNA Sequencing

7.4B.1

Supplement 47

Stop solution (see recipe)
0.2- to 0.4-mm-thick sequencing gel containing 8 M urea (*UNIT 7.6*)
1× TBE buffer (*APPENDIX 2*)
Blocking buffer I (see recipe)
Conjugate solution (see recipe)
Wash buffer I (see recipe)
Assay buffer (see recipe)
Dioxetane detection solution (see recipe)
65°, 70°, and 80°C water baths or heating blocks *or* thermal cycler
Filter paper (Whatman 3MM or equivalent)
Nylon membrane (Table 7.4B.1)
UV light source: UV cross-linking apparatus, UV transilluminator, *or* hand-held UV lamp
Large heat-sealable bags (~40 × 53-cm; Table 7.4B.1)
X-ray film (Kodak XAR-5 or equivalent)
Additional materials for DNA sequencing gel electrophoresis (*UNIT 7.6*) and exposure and processing of X-ray films (*APPENDIX 3*)

Perform DNA sequencing reactions

1. Prepare the DNA template mix by combining the following in a microcentrifuge tube:
 - 1 µg single-stranded DNA (~0.5 pmol) *or* 1 to 3 µg denatured double-stranded DNA template
 - 0.5 pmol biotinylated DNA sequencing primer
 - 2 µl 5× *Bst* reaction buffer
 - H₂O to 11 µl.

Heat to 70°C, then cool slowly over 30 min to 30°C.

This step is optional for most single-stranded templates. Sufficient primer annealing occurs at the 65°C reaction temperature (step 4).

2. While the template mix is cooling, aliquot 2 µl each of the A, C, G, and T primer termination mixes into four separate microcentrifuge tubes or microtiter plate wells, and label them A, C, G, and T. Warm the tubes or plate containing termination mixes to 65°C.
3. Add 1 µl *Bst* polymerase to the cooled template mixture from step 1.
4. Add 2.5 µl polymerase/template mixture to each primer termination mix set up in step 2 and incubate 5 min at 65°C.

This is the elongation-termination reaction.
5. Prepare a 1× chase solution by diluting the 20× chase solution stock with water. Add 2 µl of 1× chase solution to each sample tube or well and incubate 5 min at 65°C.

Eight microliters of diluted chase solution are needed per reaction set.
6. Add 4 µl stop solution to each tube or well.

At this point, reactions may be stored at -20°C for future use.

Run the DNA sequencing gel and transfer samples to nylon membrane

7. Set up a standard 0.2- to 0.4-mm-thick sequencing gel containing 8 M urea.

The width of the gel used must be <30 cm to accommodate the average width of commercially available nylon membranes (see step 9). Alternatively, several pieces of membrane

Table 7.4B.1 Suppliers of Nonisotopic DNA Sequencing Reaction Kits, Chemiluminescent Detection Reagents, and Supplies

Product	Supplier ^a
<i>General</i>	
Biotin DNA sequencing reaction kit	MI, NEB, TR, USB
Digoxigenin DNA sequencing reaction kit	BM
<i>Bst</i> polymerase	BR, ET
Biotin DNA sequencing chemiluminescent detection kit	MI, NEB, TR, USB
Digoxigenin chemiluminescent detection kit	BM
Chemiluminescent 1,2-dioxetane substrate ^b	
CSPD	TR
AMPPD	BR, PR, TR
Lumigen-PPD	BM, LU, MI, NEB
Membrane-blocking reagent	
I-Block	TR
Genius blocking agent	BM
Streptavidin–alkaline phosphatase conjugate	LT, TR, USB
Antibody–alkaline phosphatase conjugate	
anti-digoxigenin	BM
anti-fluorescein	DK, NEN
anti-DNP	DK
Large heat-sealable bags	NEB, MI, TR, USB
Nylon membranes	LT, MI, MS, NEB, PA, ST, TR, USB
<i>Specialized</i>	
Large bag sealer, 60 cm (24 in.)	NBC
Horizontal electroblotter	
Panther HEP-3	OSP
TE 90	HPB
Direct transfer electrophoresis (DTE) apparatus;	HPB
TwoStep	
Rotating-bottle apparatus; Navigator	BI
Automated blot development apparatus; PR 1000	HPB

^aAbbreviations: BI, Biocomp Instruments; BM, Boehringer Mannheim; BR, Bio-Rad; DK, Dako; ET, Epicentre Technologies; HPB, Hoefer Pharmacia Biotech; LT, Life Technologies; LU, Lumigen; MI, Millipore; MS, Micron Separations; NBC, National Bag Company; NEB, New England Biolabs; NEN, NEN Life Sciences; OSP, Owl Scientific Plastics; PA, PALL; PR, Promega; ST, Stratagene; TR, Tropix; USB, U.S. Biochemical. Addresses of suppliers are listed in APPENDIX 4.

^bAbbreviations: AMPPD, disodium 3-(4-methoxyspiro[1,2-dioxetane-3,2'-tricyclo[3.3.1.1^{3,7}]decan]-4-yl)phenyl phosphate; CSPD, disodium 3-(4-methoxyspiro[1,2-dioxetane-3,2'-(5'-chloro)tricyclo [3.3.1.1^{3,7}]decan]-4-yl)phenyl phosphate; Lumigen-PPD, 4-methoxy-4-(3-phosphate phenyl)-spiro-(1,2-dioxetane-3,2'-adamantane), disodium salt.

may be placed on the gel or the transfer may be accomplished by direct transfer electrophoresis (DTE; see Critical Parameters).

- Immediately before loading, heat reactions for 2 min at 80°C and place on ice. Load 2 to 3 μ l. Run gels as in UNIT 7.6, monitoring the migration of the marker dyes (see Table 7.6.2).

Do not perform the post-electrophoresis processing described in UNIT 7.6; DNA must be transferred from the gel to nylon membrane to perform the chemiluminescent detection.

- Cut three pieces of Whatman 3MM filter paper to size of gel or slightly larger. Cut a piece of nylon membrane that will cover the desired region of gel.

- Disassemble gel apparatus and separate glass plates.
- Place one piece of dry Whatman 3MM filter paper on the gel, gently rub the paper, and carefully lift the gel up by peeling back the paper. Place the paper with gel attached, gel-side-up, on the glass plate.

The glass plate provides support for the filter paper with the attached gel. It is not necessary to fix the gels or remove the urea.

- Wet nylon membrane thoroughly with 1× TBE buffer, either by immersion or by squirting the membrane with buffer from a squeeze bottle.
- Carefully place wet membrane on gel. Carefully roll a pipet or smooth rod over membrane to remove air bubbles.
- Place two pieces dry Whatman 3MM filter paper on top of membrane. Place another glass plate on top of filter paper and cover with ~2 to 4 kg of weight. Allow transfer to proceed 1 hr.

It is important to assemble the gel/membrane/filter paper/glass plate sandwich immediately after removing the gel from the plate, as the liquid in the gel will cause the filter paper to wrinkle over time. This may lead to uneven contact between the gel and membrane.

- Separate gel/membrane sandwich. Carefully peel membrane back from gel. Mark side of membrane that was in contact with gel with a pencil. Place membrane on filter paper, DNA-side-up, and air dry ~10 min.

Slightly wetting one corner of the membrane with TBE buffer may aid in the separation.

UV-cross-link DNA to membrane

- Using a UV cross-linking apparatus, UV transilluminator, or hand-held UV lamp, achieve a total UV exposure of the membrane of 120 mJ/cm². Proceed with step 17 or store membrane between two pieces of plastic wrap ≤6 months at 4°C.

UV exposure time may be calculated as in the following example: 1.2 mW/cm² × 100 sec = 120 mJ/cm². With most UV sources, irradiation times of 1 to 10 min are adequate. Overexposure to UV light is not deleterious to signal strength because DNA hybridization is not required for detection. Hand-held lamps can be used to expose the whole membrane by clamping the UV source to a support and consecutively exposing individual portions of the membrane. If necessary, the optimum exposure time for a particular UV light source can be calibrated by performing control experiments in which a number of membrane strips containing biotinylated DNA are exposed to varying amounts of irradiation.

Alternatively, baking 1 hr at 80°C may be used to immobilize the DNA with most nylon membranes, according to the manufacturer's recommendations.

Perform chemiluminescent detection

Perform the following steps at room temperature with moderate shaking (120 to 140 rpm). Add each solution either by cutting the bag and resealing with heat or by folding the cut corner of the bag several times and clamping it with binder clips (see Fig. 3.19.2). Reagent volumes given apply to a single 1000- to 1200-cm² membrane. For smaller membranes, add proportionately less of each solution (e.g., 250 ml blocking buffer to a 600-cm² membrane). Only one membrane per bag is recommended. Solutions should not be reused.

- Place the membrane in large heat-sealable bag. Add 500 ml blocking buffer I, carefully remove air bubbles and seal the bag, then incubate 10 min.

A large tray or a roller-bottle machine can be used in place of heat-sealable bags.

- Incubate 20 min in 200 ml conjugate solution.

19. Wash 5 min in 500 ml blocking buffer I.
20. Wash three times, 10 min each time, in 500 ml wash buffer I.
21. Wash two times, 2 min each time, in 500 ml assay buffer.
22. Incubate 5 min in 50 ml dioxetane detection solution.

Manual agitation by rubbing or rocking the bag will facilitate uniform distribution of the solution across the membrane.

23. Drain excess detection solution from bag, smooth out any wrinkles, and reseal.

The membrane must remain wet and extreme care should be taken to avoid any wrinkles during this step, as these will cause distortions in the imaged DNA sequence data. If bag sealing causes wrinkles, the sealed areas should be cut off and excess solution should be pressed out before the membrane is exposed to film. Alternatively, the membrane may be placed between two pieces of plastic wrap or thin Mylar film.

Expose and develop film

24. Place wrapped membranes in direct contact with X-ray film at room temperature.

Exposures are best performed in hard-sided cassettes with a clamping mechanism to ensure even contact of the film with the wrapped membrane.

An initial 60-min exposure is normally adequate. Additional shorter or longer exposures can be performed to optimize signal intensity and resolution. Light output will increase over a period of 6 to 8 hr due to the kinetics of the 1,2-dioxetane light emission on nylon membranes. Incubation of the wrapped membrane at room temperature for several hours before exposure will result in much shorter film exposure times (see Commentary).

TWO-STEP (INDIRECT) DETECTION USING STREPTAVIDIN AND BIOTINYLATED ALKALINE PHOSPHATASE

ALTERNATE PROTOCOL 1

An alternative method for binding alkaline phosphatase to biotinylated DNA takes advantage of the multiple biotin binding sites of streptavidin. Biotinylated primers are used in the sequencing reaction, and the DNA is electrophoresed and transferred to a nylon membrane as described in the basic protocol. Free streptavidin is then bound to the biotin-labeled DNA, followed by incubation with biotinylated alkaline phosphatase, which links the enzyme label to the DNA through a streptavidin bridge. This procedure generates chemiluminescent signal intensities similar to those obtained with streptavidin–alkaline phosphatase conjugates. The following solution volumes apply to one 1000-cm² membrane.

Additional Materials (also see Basic Protocol)

- Two-step blocking solution (see recipe)
- Streptavidin solution (see recipe)
- Two-step wash solution I (see recipe)
- Biotinylated alkaline phosphatase solution (see recipe)
- Two-step wash solution II (see recipe)

1. Carry out DNA sequencing reactions, electrophoresis, DNA transfer, and UV cross-linking (see Basic Protocol, steps 1 to 16).
2. Place membrane in a bag or container. Add 100 ml of two-step blocking solution and incubate 5 min.

Perform all incubations and washes from this step on with moderate shaking (120 to 140 rpm) at room temperature. Solutions should not be reused.

A large tray or a roller bottle machine can be used in place of heat-sealable bags.

DNA Sequencing

7.4B.5

3. Incubate 5 min in 50 ml streptavidin solution.
4. Wash two times, 5 min each time, in 500 ml of two-step wash solution I.
5. Incubate 5 min in 50 ml biotinylated alkaline phosphatase solution.
6. Wash two times, 5 min each time, in 500 ml of two-step wash solution II.
7. Wash 5 min in assay buffer.
8. Incubate 5 min with 50 ml dioxetane detection solution.

Manual agitation by rubbing or rocking the bag will facilitate uniform distribution the solution across the membrane.

9. Drain dioxetane detection solution, process membrane, and expose to X-ray film (see Basic Protocol, steps 23 and 24).

**ALTERNATE
PROTOCOL 2**

**SEQUENCING WITH HAPTEN-LABELED PRIMERS AND DETECTION
WITH ANTIBODY-ALKALINE PHOSPHATASE CONJUGATES**

As an alternative to biotin, other haptens such as digoxigenin, fluorescein, and 2,4-dinitrophenyl (DNP) may be used to label the products of the DNA sequencing reactions by incorporating primers labeled with these haptens. Primers 5'-end-labeled with the appropriate haptens are used in DNA sequencing reactions as described in the basic protocol or in the end-labeled primer protocols of *UNIT 7.4A*. Following gel electrophoresis, DNA transfer, and UV cross-linking, an antibody-alkaline phosphatase conjugate is substituted for the streptavidin conjugate used in the basic protocol; subsequent washing and detection steps are unchanged. Chemiluminescent signal intensity is highly dependent on the conjugate utilized and superior quality conjugates that have been tested for blotting applications are highly recommended (see Table 7.4B.1 for suppliers). Background levels can be higher with antibody detections using certain membranes; several membranes (e.g., Pall Biotodyne A, MSI MagnaCharge, Tropilon-Plus, and Boehringer Mannheim nylon membranes; Table 7.4B.1) are recommended for use with this protocol.

Additional Materials (also see Basic Protocol)

- 5'-haptent-end-labeled primer
- Blocking buffer II (see recipe)
- Wash buffer II (see recipe)
- Antibody-conjugate solution (see recipe)

1. Prepare the DNA template mix (see Basic Protocol, step 1), using a 5'-haptent-end-labeled primer in place of the biotinylated sequencing primer.
2. Perform the DNA sequencing reactions, run the sequencing gel, and transfer the DNA to a nylon membrane (see Basic Protocol, steps 2 through 17).

Sequencing reactions described in UNIT 7.4A may also be carried out.

3. Place the membrane in a bag or container and add 200 ml antibody-conjugate solution. Incubate 30 min.

Perform all incubations and washes from this step on with moderate shaking (120 to 140 rpm) at room temperature.

4. Wash 5 min in 500 ml blocking buffer II.
5. Wash three times, 10 min each time, in 500 ml wash buffer II.
6. Wash two times, 2 min each time, in 500 ml assay buffer.

7. Incubate 5 min in 50 ml dioxetane detection solution.

Manual agitation by rubbing or rocking the bag will facilitate uniform distribution of the solution across the membrane.

8. Drain dioxetane detection solution and expose to X-ray film as described in steps 23 and 24 of the basic protocol.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Antibody-conjugate solution

Prepare antibody-conjugate (e.g., anti-digoxigenin, anti-fluorescein, or anti-DNP alkaline phosphatase; Table 7.4B.1) dilutions in blocking buffer II (see recipe). Follow conjugate manufacturer's recommendations for the appropriate dilutions; these are typically 1/2500 to 1/10,000. Prepare fresh.

Assay buffer

Dissolve DEA (99% purity) in H₂O to 0.1 M (final), adjust pH to 10.0 with concentrated HCl, and add MgCl₂ to 1 mM (final).

Add sodium azide to 0.02% (w/v) to store solution for 2 to 7 days at 4°C.

Biotinylated alkaline phosphatase solution

Dilute biotinylated alkaline phosphatase to 0.5 µg/ml in two-step blocking solution (see recipe). Prepare fresh.

Biotinylated alkaline phosphatase that has been tested for chemiluminescent detection, available as part of kits from Millipore and New England Biolabs, is highly recommended.

Blocking buffer I

1× phosphate-buffered saline (PBS, see recipe)

0.2% (w/v) purified casein or membrane-blocking reagent (Table 7.4B.1)

0.5% (w/v) SDS

Prepare 1× PBS solution from 10× stock. Add casein or membrane-blocking reagent, heat to 70°C with constant stirring, then add SDS. Do not boil.

Add sodium azide to 0.02% (w/v) to store solution for 2 to 7 days at 4°C.

Blocking buffer II

1× phosphate-buffered saline (PBS, see recipe)

0.2% (w/v) purified casein or membrane-blocking reagent (Table 7.4B.1)

0.1% (v/v) Tween 20

Prepare 1× PBS solution from 10× stock. Add casein or membrane-blocking reagent, heat to 70°C with constant stirring, then add Tween 20. Do not boil.

Add sodium azide to 0.02% (w/v) to store solution for 2 to 7 days at 4°C.

Conjugate solution

Prepare streptavidin–alkaline phosphatase (Table 7.4B.1) in blocking buffer I (see recipe). Follow manufacturer's recommendations for appropriate dilution; 1/5000 is typically used. Prepare fresh.

Dioxetane detection solution

Prepare a 1/100 dilution of concentrated (100×) dioxetane (Table 7.4B.1) in assay buffer (see recipe). Make fresh and do not reuse solution.

Phosphate-buffered saline (PBS), 10×82.3 g Na₂HPO₄ (0.58 M)23.5 g NaH₂PO₄ (0.17 M)

40.0 g NaCl (0.69 M)

H₂O to 1 liter

Sterilize and store indefinitely at room temperature

*The pH of a 1× solution should be 7.2.***Primer termination mixes**

Concentrations of dNTPs and ddNTPs in the A, C, G, and T mixes are listed below in mM. Prepare all mixes in highly purified H₂O. Store indefinitely at −20°C.

	A	C	G	T
dATP	4	80	80	80
dCTP	80	4	80	80
dGTP	80	80	4	80
dTTP	80	80	80	4
ddATP	200	—	—	—
ddCTP	—	100	—	—
ddGTP	—	—	60	—
ddTTP	—	—	—	225

Stop solution95% (v/v) deionized formamide (*UNIT 14.3*)

10 mM EDTA

0.1% (w/v) xylene cyanol

0.1% (w/v) bromphenol blue

Store indefinitely at −20°C

Streptavidin solution, pH 7.2

1 mg/ml streptavidin

6.8 mM Na₂HPO₄3.2 mM NaH₂PO₄

150 mM NaCl

0.05% (w/v) sodium azide

Stable ≤6 months at 4°C

Streptavidin that has been tested for chemiluminescent detection, available from Millipore and New England Biolabs, is highly recommended.

Two-step blocking solution, pH 7.2

5% (w/v) SDS

17 mM Na₂HPO₄8 mM NaH₂PO₄

Add sodium azide to 0.02% (w/v) to store solution for 2 to 7 days at 4°C.

Two-step wash solution I

Prepare a 1/10 dilution of two-step blocking solution (see recipe) in water. Make fresh and do not reuse.

Two-step wash solution II

10 mM Tris·Cl, pH 9.5

10 mM NaCl

1 mM MgCl₂

Autoclave to sterilize, then store indefinitely at room temperature.

Wash buffer I

Prepare 0.5% (w/v) SDS in 1× PBS (see recipe). Make fresh if possible and do not reuse.

If necessary, add sodium azide to 0.02% (w/v) to store solution 2 to 7 days at 4°C.

Wash buffer II

Prepare 0.1% (v/v) Tween 20 in 1× PBS (see recipe). Make fresh if possible and do not reuse.

If necessary, add sodium azide to 0.02% (w/v) to store solution 2 to 7 days at 4°C.

COMMENTARY

Background Information

Originally, Maxam-Gilbert and Sanger dideoxy procedures involved the detection of DNA fragments labeled with radioactive isotopes. Recently, however, nonisotopic labeling alternatives have become widely used for DNA sequencing as well as for other applications in molecular biology (UNITS 3.18, 3.19 & 10.8). These nonisotopic methods have proven to be rapid, highly sensitive, and quantitative; furthermore, concerns associated with the use and disposal of radioactivity are eliminated.

Nonisotopic labeling techniques include colorimetric (BCIP and NBT; Richterich et al., 1989), fluorescent (Prober et al., 1987), and 1,2-dioxetane chemiluminescent methods (Beck et al., 1989; Tizard et al., 1990; Creasey et al., 1991; Martin et al., 1991; Hölte et al., 1992; Richterich and Church, 1993; Olesen et al., 1993). In standard isotopic-based sequencing, the label may be attached either to the sequencing primer or incorporated into the extension products as part of a modified nucleotide. For nonisotopic sequencing, the nonisotopic label is usually attached to the 5'-end of the sequencing primer. Although fluorescent detection methods generally require the use of relatively expensive automated equipment, chemiluminescent methods can be performed manually with standard sequencing gel electrophoresis and no specialized equipment.

The basic protocol for DNA sequencing with chemiluminescence described here is based on standard dideoxy methods. Through the incorporation of biotinylated primers, DNA fragments resulting from the sequencing reactions are biotin-labeled; they can then be detected by binding a streptavidin–alkaline phosphatase conjugate and incubating with a 1,2-dioxetane chemiluminescent substrate for alkaline phosphatase. The quality of the DNA sequence data is comparable to results obtained with radioactive isotopes. In fact, the method's high sensitivity has made the use of 1,2-di-

oxetane substrates increasingly popular. In addition to DNA sequencing, chemiluminescent techniques for performing Southern and western hybridization analyses have been developed (UNITS 3.18, 3.19 & 10.8).

Detection is based on the generation of a light signal following dephosphorylation of a 1,2-dioxetane substrate by the enzyme alkaline phosphatase. For example, enzymatic dephosphorylation of the 1,2-dioxetane CSPD (Table 7.4B.1) results in a strongly electron-releasing, destabilized dioxetane anion, which breaks down further, emitting light (Fig. 7.4B.1; Bronstein et al., 1991; Martin et al., 1991). The dephosphorylated dioxetane anion has a half-life of ~40 min on nylon membrane and the emission maximum is 460 nm. The kinetics of light emission of the 1,2-dioxetane substrates AMPPD and CSPD are shown in Figure 7.4B.2 (Martin et al., 1991). The relatively long half-life of the anion on nylon membrane results in the delay before maximum light emission. The presence of a chlorine atom, a remote substituent on the adamantyl group in CSPD, decreases the half-life of the anion, which results in greater intensity light output at early time-points. The shorter half-life, together with the more hydrophobic nature of the chloro-appended adamantyl group, also decreases the rate of diffusion of dephosphorylated anion and limits the loss of imaged band resolution observed in next-day exposures (Martin et al., 1991).

Other luminescent systems have also been used for the detection of nucleic acids (for a survey, see Kricka, 1992). Most of these systems have not been investigated in DNA sequencing applications. The other commonly used luminescent detection method for nucleic acids is based upon the activation of luminol with antibody–horseradish peroxidase conjugates (Matthews et al., 1985; Pollard-Knight et al., 1990). The use of luminol-based detection methods for DNA sequencing was reported by Richterich and Church (1993).

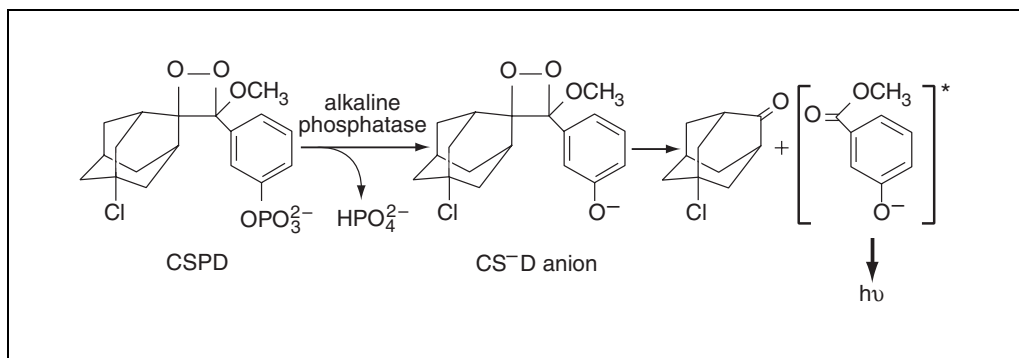


Figure 7.4B.1 Pathway of decomposition of CSPD chemiluminescent substrate. Asterisk indicates compound in its excited state. Reprinted with permission of Tropix.

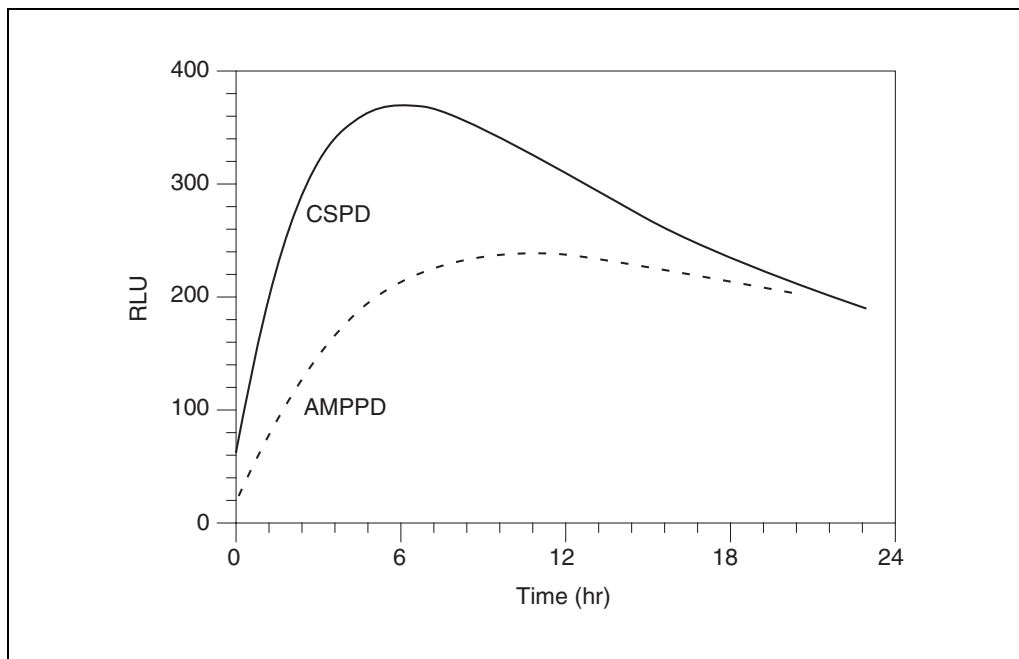


Figure 7.4B.2 Chemiluminescence kinetics on nylon membrane: CSPD and AMPPD. Abbreviations: RLU, relative light units on a Turner luminometer; for chemical names, see Table 7.4B.1. Reprinted with permission of Tropix.

Standard chemiluminescent dideoxy sequencing with biotinylated primers using 1,2-dioxetane-alkaline phosphatase substrates is also described in Beck et al. (1989), Creasey et al. (1991), and Martin et al. (1991). Other sequencing/detection systems include the detection of digoxigenin-labeled DNA sequencing reactions with antibody-alkaline phosphatase conjugates (Höltke et al., 1992; Olesen et al., 1993) and the use of fluorescein- and dinitrophenyl (DNP)-labeled primers with detection by antibody-alkaline phosphatase conjugates (Olesen et al., 1993). DNA sequencing reactions can also be detected on nylon membranes by the hybridization of nonisotopically labeled probes followed by chemiluminescent detection using methods similar to genomic

sequencing procedures originally described for ^{32}P -labeled probes by Church and Gilbert (1984). These probes can either be labeled with biotin (Creasey et al., 1991) or alkaline phosphatase directly (Tizard et al., 1990; Karger et al., 1992, 1993).

Multiplex sequencing (Church and Kieffer-Higgins, 1988), which involves the sequential detection of overlapping sets of DNA reactions on one membrane by the successive hybridization of sequence-specific probes, has been performed with chemiluminescent detection (Creasey et al., 1991; Lakey et al., 1993). A multiplex-labeling procedure can also be used with chemiluminescent detection (Richterich and Church, 1993; Olesen et al., 1993). With this method, overlapping sets of DNA sequenc-

ing reactions, performed with primers labeled with different haptens (e.g., biotin, DNP, digoxigenin, or fluorescein) at the 5'-end, are detected sequentially using hapten-specific alkaline phosphatase conjugates.

The adaptability of the 1,2-dioxetane chemiluminescent detection system is demonstrated by the wide variety of DNA sequencing applications performed with the system. Simple adaptations of the basic protocol described here should be able to be used to perform other techniques requiring high resolution gels, including primer extension experiments and the detection of simple sequence repeat alleles by polymerase chain reaction (PCR) amplification (UNIT 15.1).

Critical Parameters and Troubleshooting

DNA sequencing reactions

Any type of sequencing reaction chemistry can be performed if 5'-end-labeled primers and protocols are utilized. An alternative labeling system based upon the incorporation of biotinylated nucleotides during the extension reactions is available as a kit from U.S. Biochemical (APPENDIX 4). It is important to use high-quality DNA sequencing reactions to give an adequate signal-to-noise ratio. If high backgrounds are seen on long exposures, a sequencing reaction using a control template, provided in many sequencing kits, can be used to distinguish between failure of the sequencing reactions and failure of the chemiluminescent detection procedure.

DNA transfer to membrane

The efficiency of DNA transfer to nylon membrane is highly dependent on the method and the type of membrane used. Capillary transfer seems to be more efficient with positively-charged nylon than with neutral nylon membranes (C.S.M., unpub. observ.). However, the use of neutral nylon normally results in lower background with antibody-based detections. Slightly sharper band resolution is exhibited when the gel is removed from the glass plate prior to setting up the transfer, as compared to placing the membrane directly on the gel without removing it from the plate. A similar membrane/filter paper sandwich is assembled using wet filter paper when the transfer is performed with a large-format electroblotting apparatus (Table 7.4B.1). The efficiency of transfer with an electroblotter is >90%, compared to ~20%

with the capillary method (C.S.M., unpub. observ.), although greater band resolution is achieved with the capillary transfer. This is most likely due to liquid-induced gel expansion during the electrophoretic transfer procedure. For most DNA sequencing experiments, the amount of DNA on the membrane transferred by the capillary method results in a strong chemiluminescent signal, but the more efficient electrophoretic transfer method may be necessary when smaller amounts of DNA are loaded on the gel.

As an alternative to standard gel electrophoresis and horizontal DNA transfer, direct-transfer electrophoresis (DTE) instruments are commercially available (Table 7.4B.1). They are modifications of the original design of Pohl and Beck (1987) and perform simultaneous electrophoretic separation and transfer to nylon membrane. They function by depositing the DNA fragments onto a membrane that is moved slowly along the bottom of the electrophoresis plates. Highly efficient DNA transfer and much better resolution of high-molecular-weight fragments are achieved with this apparatus, resulting in longer sequence readings.

UV cross-linking

Fixing the DNA fragments to nylon membrane prior to chemiluminescent detection is necessary with most membranes. Strongly positively charged nylon membranes such as Biodyne B (Pall) do not require UV cross-linking. However, other positively charged membranes, such as MSI MagnaCharge (Micron Separations, APPENDIX 4), appear to lose slight amounts of DNA during the detection procedures when UV cross-linking has not been performed. Without this cross-linking, neutral nylon membranes consistently lose DNA during the detection procedure. No reduction in the chemiluminescent signal intensity from UV irradiation when using biotinylated primers has been noted (C.S.M., unpub. observ.); therefore, UV cross-linking is recommended if the membrane has not been previously tested for DNA retention. Although overexposure to UV light does not decrease the ability to detect biotinylated DNA, it will interfere with the ability to hybridize probes to bound DNA. Procedures involving multiplex sequencing and detection of DNA sequences with oligonucleotide-alkaline phosphatase conjugates do require DNA hybridization, and it is important to cross-link DNA to the membrane with the optimum UV dose (generally 30–120 mJ/cm²).

Chemiluminescent detection

Chemiluminescent detection steps are easily performed in large plastic bags. Large trays can be used but it is often difficult to find the appropriate size and shape; in addition, larger solution volumes are usually necessary with the trays. Bags should be sealed close to the membrane with enough space at one end to permit resealing or folding. Large (60-cm) bag sealers that facilitate uniform sealing are available commercially (Table 7.4B.1). Folding the membrane can be avoided by adding solutions slowly and not using excessive volumes of solutions. When small volumes of solutions are added, as during the conjugate and substrate incubation steps, manual agitation by rubbing or rocking the bag will facilitate uniform distribution of the solution across the membrane. Membranes and bags should be handled with gloves.

Alternatively, a large format rotating-bottle apparatus (Table 7.4B.1) can be used to perform the chemiluminescent detection steps; in addition, an automated system is available that may be programmed to pump the detection reagents in and out of the rotating bottle (Table 7.4B.1). Both of these instruments utilize large bottles to accommodate full-size sequencing gel blots without overlap. Longer incubation times and washes are necessary with the roller-bottle assemblies (C.S.M., unpub. observ.).

To avoid bacterial alkaline phosphatase contamination, detection solutions should be prepared with ultrapure water. Diluted alkaline phosphatase conjugate solutions should be used immediately, and their reuse is not recommended. Reusing the substrate solution is possible but may result in much lower signal intensity if contamination by alkaline phosphatase occurs. Commercially available chemiluminescent DNA sequencing kits are convenient because they include membrane-blocking agents and enzyme conjugates that have been specifically optimized for use with chemiluminescence. The quality of commercially available enzyme conjugates varies greatly. Using a conjugate not specifically optimized for chemiluminescent detection on nylon membranes may lead to poor results. Casein preparations can be contaminated with biotin or alkaline phosphatase, leading to high background. Therefore, the use of highly purified casein or blocking agents optimized for membrane-based chemiluminescent detection is highly recommended.

Film exposure

The membrane should not be allowed to dry during film exposure. It may be left in the plastic bag and exposed to film after the substrate solution has been removed. However, nonuniform bag sealing may cause wrinkles, which will prevent even contact between the membrane and film and result in blurry regions or dark lines. To avoid this problem, a margin of ~2 cm should be left on all sides of the membrane when sealing the bag. The sealing seams should be trimmed from the bag after performing the chemiluminescent detection steps and excess solution should be squeezed out of the bag, using firm pressure from a paper towel, to avoid leakage during film exposure. Alternatively, the membrane may be placed in a new bag or between two pieces of plastic wrap (thicker wrap will wrinkle less). Two pieces of Mylar film, similar to the material used in report covers, works very well, but may be difficult to find in the right size. Thicker plastic sheets or previously used X-ray film may be used as a membrane support, with a layer of plastic wrap placed over the membrane (this results in fewer wrinkles than found when using two pieces of plastic wrap). It is important to note that the use of any thick transparent material will result in some loss of in-band resolution, because of the diffusion of light through a thicker layer.

Due to the kinetics of light emission of 1,2-dioxetane substrates on nylon membranes, the required exposure time will decrease over a 6- to 8-hr period (see Fig. 7.4B.2). However, delaying exposure to film after substrate addition is unnecessary. Generally, an initial 60-min film exposure will be ~2-fold lighter than a subsequent 60-min exposure performed immediately after this initial one. Film exposures obtained the next day may exhibit more diffuse bands, depending on the specific dioxetane substrate and membrane used. After 2 to 3 days, the signal will disappear from the high-signal areas of the membrane due to substrate depletion.

Table 7.4B.2 lists problems particular to dideoxy sequencing with chemiluminescent detection. General troubleshooting guidelines for dideoxy sequencing can be found in Tables 7.4A.3 to 7.4A.5.

Anticipated Results

Any template DNA can be sequenced using this chemiluminescent detection protocol with biotinylated primers. The technique is sensitive and flexible. The data generated are visualized on X-ray film and are similar to those achieved with radioisotopic detection. In general, band

Table 7.4B.2 Troubleshooting Guide for DNA Sequencing with Chemiluminescent Detection^a

Problem	Possible cause	Solution
High background	Contamination of reagents with alkaline phosphatase	Prepare all buffers daily; use ultrapure deionized water and other reagents free of alkaline phosphatase contamination. Avoid cross-contamination of wash solutions by conjugate solution—i.e., cover solutions during storage; wash funnel used for liquid transfers; clean outside of bag after each solution addition.
	Glove dust	Use powder-free gloves
	Insufficient agitation during incubations	Increase shaker speed; increase volume of liquid; remove air bubbles
	Nonspecific binding of conjugate to membrane	Increase incubation time with blocking buffer or increase number of washes after enzyme/conjugate incubation
Spotchy images	Bacterial or alkaline phosphatase contamination of membrane	Determine that all buffers are free of contamination prior to use and that membrane, blotting paper, and hybridization bags are clean and fingerprint-free
Spots or lines	Mechanical damage to the membrane	Repeat with new membrane
Poor band resolution	Exposure begun too long after substrate removal (depending on specific dioxetane substrate and membrane used)	Perform film exposures soon after substrate addition
	Membrane not in good contact with gel during transfer	Carefully press membrane onto gel with pipet; add heavier weight during capillary transfer procedure
	X-ray film not in good contact with membrane	Reseal membrane with plastic wrap and avoid wrinkles; use exposure cassette with clamps
Low signal	Incorrect assay buffer pH	Adjust assay buffer to pH 10
	Inefficient DNA transfer	Review transfer procedure and repeat experiment. When using capillary transfer method, it is very important to have gel sandwich on an absolutely flat surface. Increase amount of weight to provide more uniform transfer.
Low intensity signal from low-molecular-weight fragments	Low-molecular-weight fragments bind poorly, especially to neutral nylon	Use charged nylon membrane; always cross-link DNA to membrane

^aGeneral troubleshooting information on DNA sequencing reactions and gel electrophoresis is described in Critical Parameters and UNIT 7.4A.

resolution is as good as or better than ³²P detection methods.

Time Considerations

Due to the short film-exposure times, the chemiluminescent DNA sequencing procedure can be performed in a 7- to 10-hr period. The gel pouring and DNA sequencing reactions require 1 hr, and gel electrophoresis takes 3 hr. DNA transfer takes 1 hr, drying and UV cross-linking requires 0.5 hr, detection requires 1.5 hr, and film exposure takes 1 hr. It is advisable to reserve time at the end of the day to perform

any necessary extra film exposures. Alternatively, the procedure may be easily performed over 2 or more days or blocks of time.

Sequencing reactions and gel preparation can be performed in advance. Biotinylated DNA sequencing reactions are stable at -20°C for long periods of time. After electrophoresis and DNA transfer, the nylon membrane can be stored dry between layers of plastic wrap for several weeks at 4°C. UV cross-linking can be performed before or after storage. Any of the chemiluminescent detection steps, except the substrate incubation step, can be increased to

permit an experimental break. Long incubations in excess substrate solution can result in signal "bleeding" from the bands. The length of time suggested for each chemiluminescent detection step may be shortened, generally with satisfactory results. However, if higher background or lower signal occurs, the standard procedure should be followed.

Literature Cited

- Beck, S., O'Keefe, T., Coull, J.M., and Koster, H. 1989. Chemiluminescent detection of DNA: Application for DNA sequencing and hybridization. *Nucl. Acids Res.* 17:5115-5123.
- Bronstein, I., Juo, R.-R., Voyta, J.C., and Edwards, B. 1991. Novel chemiluminescent adamantyl 1,2-dioxetane enzyme substrates. In *Bioluminescence and Chemiluminescence: Current Status*, (P. Stanley and L.J. Kricka, eds.) pp. 73-82. John Wiley & Sons, Chichester.
- Church, G.M. and Gilbert, W. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 81:1991-1995.
- Church, G.M. and Kieffer-Higgins, S. 1988. Multiplex DNA sequencing. *Science* 240:185-188.
- Creasey, A., D'Angio Jr., L., Dunne, T.S., Kissinger, C., O'Keefe, T., Perry-O'Keefe, H., Moran, L.S., Roskey, M., Schildkraut, I., Sears, L.E., and Slatko, B. 1991. Application of a novel chemiluminescence-based DNA detection method to single-vector and multiplex DNA sequencing. *BioTechniques* 11:102-109.
- Höltke, H.J., Sanger, G., Kessler, C., and Schmitz, G. 1992. Sensitive chemiluminescent detection of digoxigenin-labeled nucleic acids: A fast simple protocol and its applications. *BioTechniques* 12:104-113.
- Karger, A.E., Weiss, R., and Gesteland, R.F. 1992. Electronic chemiluminescence imaging of DNA sequencing blots using a charge-coupled device camera. *Nucl. Acids Res.* 20 (24):6657-6665.
- Karger, A.E., Weiss, R., and Gesteland, R.F. 1993. A line scanning system for direct digital chemiluminescence imaging of DNA sequencing blots. *Anal. Chem.* 65:1785-1793.
- Kricka, L.J. (ed.) 1992. *Nonisotopic DNA Probe Techniques*. Academic Press, San Diego.
- Lakey, N.D., Ghazizadeh, H., Jaehn, L., Richterich, P., Robinson, K., and Church, G. 1993. Automated multiplex sequencing of *E. coli* genomic DNA. *Genome Science and Technology I*, p. 41.
- Martin, C., Bresnick, L., Juo, R.-R., Voyta, J.C., and Bronstein, I. 1991. Improved chemiluminescent DNA sequencing. *BioTechniques* 11:102-109.

- Mathews, J.A., Batki, A., Hynds, C., and Kricka, L.J. 1985. Enhanced chemiluminescent method for the detection of DNA dot-hybridisation assays. *Anal. Biochem.* 151:205-209.
- Olesen, C.E.M., Martin, C.S., and Bronstein, I. 1993. Chemiluminescent DNA sequencing with multiplex labeling. *BioTechniques* 15:480-485.
- Pohl, F.M. and Beck, S. 1987. Direct transfer electrophoresis used for DNA sequencing. *Methods Enzymol.* 155:250-259.
- Pollard-Knight, D., Read, C.A., Downes, M.J., Howard, L.A., Leadbetter, L.A., Pheby, S.A., McNaughton, E., Syms, A., and Brady, M.A.W. 1990. Nonradioactive nucleic acid detection by enhanced chemiluminescence using probes directly labeled with horseradish peroxidase. *Anal. Biochem.* 185:84-89.
- Prober, J.M., Trainor, G.L., Dam, R.J., Hobbs, F.W., Robertson, C.W., Zagursky, R.J., Cocuzza, A.J., Jensen, M.A., and Baumeister, K. 1987. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* 238:336-341.
- Richterich, P. and Church, G.M. 1993. DNA sequencing with direct transfer electrophoresis and nonradioactive detection. *Methods Enzymol.* 218:187-223.
- Richterich, P., Heller, C., Wurst, H., and Pohl, F.M. 1989. DNA sequencing with direct blotting electrophoresis and colorimetric detection. *BioTechniques* 5:52-59.
- Tizard, R., Cate, R.L., Ramachandran, K.L., Wysk, M., Voyta, J.C., Murphy, O.J., and Bronstein, I. 1990. Imaging of DNA sequences with chemiluminescence. *Proc. Natl. Acad. Sci. U.S.A.* 87:4514-4518.

Key References

- Creasey et al., 1991. See above.
Describes the two-step (indirect) detection protocol.
- Martin et al., 1991. See above.
Describes the basic chemiluminescent detection protocol.
- Olesen et al., 1993. See above.
Describes the basic chemiluminescent detection protocol, as well as the hapten-labeled primer and antibody-conjugate detection methods.

Contributed by Chris S. Martin
Tropix, Inc.
Bedford, Massachusetts

DNA Sequencing by the Chemical Method

This procedure is written for a novice with no chemical sequencing experience. It is assumed, however, that the investigator has experience with DNA fragment isolation (UNIT 2.6) and DNA labeling with [^{32}P]dNTPs (UNITS 3.5 & 3.10). Table 7.5.1 is provided for quick reference during the sequencing reactions. Commercial kits for chemical sequencing are available from Du Pont and E. Merck (APPENDIX 4).

A DNA fragment labeled at only one end with ^{32}P or ^{35}S is divided into four aliquots and subjected to base-specific modification reactions. Piperidine is then added to catalyze strand scission at the modified bases. Finally, the four reactions are subjected to electrophoresis on adjacent lanes of a high-resolution denaturing polyacrylamide gel. The gel is then autoradiographed and the sequence is read from the film.

STRATEGIC PLANNING

The chemical (Maxam-Gilbert) sequencing procedure relies on DNA fragments labeled at *one* end with ^{32}P by a variety of labeling procedures. The DNA fragment can be either double- or single-stranded. Approximately 50,000 Cerenkov cpm are required. End-labeling of DNA at the 5' end can be achieved with T4 polynucleotide kinase and [γ - ^{32}P]ATP (UNIT 3.10). Labeling at the 3' end is achieved with terminal transferase and [α - ^{32}P]ATP followed by alkaline hydrolysis to remove all but the first adenylic acid residue (UNIT 3.6) or by "filling in" the complementary strand of a 5' single-stranded protruding end created by digestion with a restriction endonuclease (UNIT 3.5).

The methods described above always label both ends of the DNA fragment; there are several procedures to obtain a fragment labeled at only one end. In one approach, a restriction fragment labeled at both ends can be digested with a restriction endonuclease that cleaves the fragment asymmetrically. The two labeled fragments are then fractionated by gel electrophoresis and eluted from the gel (UNIT 2.6). In another approach, a fragment labeled at both ends is subjected to electrophoresis under denaturing conditions in which the two complementary single-stranded fragments are resolved. This latter method is no longer widely used and is not treated directly in this manual.

A third way to label only one end of a DNA fragment is to fill in an asymmetric restriction enzyme cleavage site as shown in Figure 7.1.1. Specialized cloning vectors pSP64CS and pSP65CS, described in UNIT 7.1, have been designed for this use. Vectors of this type are particularly convenient for cloning and labeling a set of nested deletions generated by nuclease *Bal* 31 digestion (UNIT 7.2). These vectors are strongly recommended when undertaking large-scale chemical sequencing projects. A support protocol for *Tth*111I digestion and end labeling is included in this unit. A protocol for small-scale plasmid preparation is provided in UNIT 7.3.

In the dideoxy sequencing method, larger oligodeoxynucleotides are much better resolved in the autoradiograms of sequencing gels when ^{35}S is used in place of ^{32}P . There are two reasons, however, why most investigators use ^{32}P for the chemical sequencing method. First, in the chemical method, each oligodeoxynucleotide contains a single radiolabeled nucleotide; because both the specific activity and the energy of decay is higher for ^{32}P than for ^{35}S , considerably shorter exposure times can be obtained with the ^{32}P label. Moreover, it is possible to shorten exposure time with an intensifying screen when using ^{32}P . Second, the oligodeoxynucleotides in the chemical sequencing method are present in approximately equimolar amounts and therefore all bands on an autoradiograph of a chemical sequencing gel have the same intensity. This results in excellent resolution with

Contributed by Richard L. Eckert

Current Protocols in Molecular Biology (1992) 7.5.1-7.5.11

Copyright © 2000 by John Wiley & Sons, Inc.

³²P. In contrast, in the dideoxy method bands often differ considerably in intensity. If a high-intensity band is adjacent to a low-intensity band at the top of the gel, an exposure that allows the low-intensity band to be seen may overexpose the higher-intensity band, obscuring the lower-intensity band.

Despite the advantages of using ³²P labeling for chemical sequencing, it is sometimes convenient to use ³⁵S labeling. DNA fragments can be labeled at the 3' end using terminal transferase (*UNIT 3.6* and Ornstein and Kashdan, 1985).

CHEMICAL SEQUENCING USING ³²P-LABELED DNA

Materials

- ³²P-labeled DNA (see strategic planning)
- 3 M sodium acetate, pH 5.0
- Isopropanol
- 70% ethanol prepared with sterile H₂O
- Dimethyl sulfate (DMS) reaction buffer
- 5 M NaCl
- DMS
- Formic acid
- Hydrazine
- DMS stop buffer
- Hydrazine stop buffer
- 100% ethanol, prechilled to -20°C
- 5 M NaOH
- 3 M ferric chloride
- 10% piperidine, freshly prepared in sterile H₂O
- Formamide loading buffer

- 1.5-ml snap-top and screw-top (Sarstedt #72.692) microcentrifuge tubes
- 90°C water bath or oven

- Additional reagents and equipment for preparing and running sequencing gels (*UNIT 7.6*) and for autoradiography (*APPENDIX 3*)

Preparation of Salt-Free DNA

1. Precipitate [³²P]DNA by adding 3 M sodium acetate, pH 5.0, to 0.3 M plus 1 vol isopropanol, and place 5 min in a dry ice/ethanol bath (*UNIT 2.1*).
2. Microcentrifuge 5 min at 15,000 × g, room temperature, and discard supernatant.
3. Rinse the pellet twice with 1 ml of 70% ethanol, and dry.
³²P-labeled DNA must be rinsed thoroughly with 70% ethanol prior to sequencing to remove excess salt, which can suppress the reaction of hydrazine with thymine and result in faint T bands in the T + C lane.
4. Resuspend the salt-free DNA pellet in 40 μl sterile water.
5. Count the [³²P]DNA using Cerenkov counting by placing the entire tube in a 20-ml scintillation vial (no scintillation fluid), and counting on the ¹⁴C channel. Multiply the result by 4 to obtain the counts per minute.

If there are too many counts in the tube, 1 μl of the material can be spotted onto a small piece of Whatman 3MM paper, dried, and counted in a scintillation vial without fluid.

It is ideal to have ≥10,000 Cerenkov counts of [³²P]DNA per tube; however, it is possible to obtain a satisfactory sequence with as few as 1000 cpm.

- Distribute 10 μl [^{32}P]DNA into each of four 1.5-ml snap-top microcentrifuge tubes labeled G, G+A, T+C, and C.

If the same amount of DNA is used for each of the four base-specific modification reactions, the intensity of the bands in the G+A and T+C lanes will be approximately half the intensity of the bands in the G and C lanes. If this proves to be a problem, which might be the case if the amount of labeled fragment is limiting, use twice as much DNA in the G+A and T+C lanes.

When sequencing large numbers of samples, it is useful to perform each reaction sequentially. For example, when sequencing 10 samples, number the tubes consecutively from 1 to 40. Tubes 1 to 4 are G, G+A, T+C, and C, respectively, for sample 1; tubes 5 to 8 are G, G+A, T+C, and C, respectively, for sample 2, and so forth. The tubes containing each of the four reactions are color coded (e.g., the G reaction tubes 1, 5, 9, etc. all have the same color code). This allows each set of reactions for each sample to be readily distinguished when adding reagents. The G reactions are performed as a group, followed by the G+A group, etc. While the second group is reacting, the first group can be precipitating, and so on.

Sequencing Reactions

A detailed description of the sequencing protocol is given below and is designed to complement Table 7.5.1, which contains an outline of the sequencing reactions. The outline should be conveniently accessible when doing the sequencing reactions.

Prepare the DNA samples (Table 7.5.1A)

- Prepare the four DNA samples by adding DMS reaction buffer, water, and 5 M NaCl.

NaCl in the C reaction suppresses the reaction of hydrazine with thymine so that only cytosines react and are observed as bands on the gel.

Carry out the base-specific modification reactions (Table 7.5.1B)

CAUTION: Conduct these reactions in a fume hood wearing plastic gloves. DMS is volatile and toxic; it can be inactivated with 5 M NaOH. Hydrazine is toxic and in the anhydrous state, explosive; it can be inactivated with 3 M ferric chloride.

- Add DMS and formic acid to the G and G+A reaction tubes, respectively, and immediately close the tubes and mix.
- Add hydrazine to the T+C and C reaction tubes, and immediately close the tubes and mix.
- Incubate the four tubes at 25°C for the times indicated in Table 7.5.1B—use a stopwatch to time the reactions.

The concentrations of the reagents should not be modified (except for DMS; see commentary) since this may result in unwanted side reactions. The extent of reaction can be controlled by changing the reaction time. The times indicated in Table 7.5.1 are intended as a starting point and will usually allow from 1 to 300 bases to be read, provided adequate counts are used. Shorter reaction times will shift more counts into larger fragments and longer times will result in more counts in shorter fragments.

Stop reactions and collect modified DNA by ethanol precipitation (Table 7.5.1C)

- Add either DMS or hydrazine stop buffer plus 100% ethanol, prechilled to -20°C , and immediately immerse in a dry ice/ethanol bath for 5 min.

The dilution plus the low temperature terminates the reactions.

12. Microcentrifuge 5 min at $15,000 \times g$ and discard the DMS supernatant into 5 M NaOH and the hydrazine supernatant into 3 M ferric chloride.

It is convenient to keep 200 ml of each of these inactivating agents in a dark, 1-liter glass vessel in the hood for the purpose of collecting waste.

13. Rinse the pellets twice with 70% ethanol (to remove excess reagents) and discard the supernatants as in step 12.

Reprecipitate the DNA (Table 7.5.1D)

14. Resuspend each pellet in 200 μ l sterile water, add 20 μ l of 3 M sodium acetate and 500 μ l 100% ethanol, and mix.

15. Immerse 5 min in a dry ice/ethanol bath and repeat steps 12 and 13.

These steps remove residual DMS and hydrazine from the samples, which can complicate the piperidine reaction.

Table 7.5.1 Outline of Chemical Sequencing Reactions

Reaction	Reaction specificity			
	G	G + A	T + C	C
A. Prepare the DNA samples				
DNA, μ l	10	10	10	10
DMS reaction buffer, μ l	200	—	—	—
H ₂ O, μ l	—	—	10	5
5 M NaCl, μ l	—	—	—	5
B. Carry out the base-specific modification reactions				
DMS, μ l	1	—	—	—
Formic acid, μ l	—	25	—	—
Hydrazine, μ l	—	—	30	30
Time (min) at 25°C	4	5	8	8
C. Stop the reactions by diluting with stop buffer and -20°C 100% ethanol and immersing in a bath of dry ice/ethanol for 5 min				
DMS stop buffer, μ l	50	—	—	—
Hydrazine stop buffer, μ l	—	200	200	200
100% ethanol, μ l	750	750	750	750
Collect DNA by centrifuging 5 min at $15,000 \times g$; rinse twice with 70% ethanol; drain				
D. Remove any remaining reagents by reprecipitation				
H ₂ O, μ l	200	200	200	200
3 M sodium acetate, μ l	20	20	20	20
100% ethanol, μ l	500	500	500	500
E. Catalyze the strand scission at modified bases by incubating 30 min at 90°C with 10% piperidine; remove the piperidine by drying under vacuum				
10% piperidine, μ l	70	70	70	70
F. Remove traces of piperidine by evaporating the sample twice from water				
Sterile H ₂ O, μ l (first evaporation)	30	30	30	30
Sterile H ₂ O, μ l (second evaporation)	20	20	20	20

16. Dry the pellets in a Speedvac evaporator.

Catalyze strand-scission reaction with piperidine (Table 7.5.1E)

17. Resuspend pellets in 70 μ l of 10% piperidine.
18. Transfer samples to 1.5-ml screw-top microcentrifuge tubes, tightly cap the tubes, and incubate 30 min at 90°C.

During this reaction, the tubes must be tightly sealed to prevent piperidine loss. If loss occurs, the piperidine reaction with hydrazine-modified nucleosides may not go to completion, resulting in smeared bands in the pyrimidine sequencing lanes. Although snap-top microcentrifuge tubes are most convenient for steps 5 to 11, they do not seal well enough to contain piperidine heated to 90°C. Other materials to seal the tubes, such as Teflon tape, do not provide a dependable seal. Transferring the reactants to Sarstedt screw-top microcentrifuge tubes at this stage provides absolute assurance that these reactions proceed normally.

Remove traces of piperidine from reactions by evaporation (Table 7.5.1F)

19. Microcentrifuge tubes briefly to collect any condensate from the sides of the tubes. Transfer the tubes to a Speedvac evaporator and evaporate the piperidine to dryness.
20. Resuspend dried samples in 30 μ l sterile water, transfer to new tubes, and dry in a Speedvac evaporator.
21. Repeat step 20 using 20 μ l sterile water.

Failure to completely remove piperidine will result in smearing of bands on the sequencing gel. The samples are now ready for gel fractionation.

Denaturing Polyacrylamide Gel Electrophoresis

The following steps are described in full detail in the basic protocol of *UNIT 7.6* (see also commentary of that unit for an extensive discussion of critical parameters).

Prepare the sequencing gels

22. Prepare 6%, 8%, or 12% sequencing gels, preheated with wells rinsed.

To “read” the sequencing gel at the extreme high- and low-molecular-weight regions (1 to 300 bases), it is necessary to run a set of the four sequencing reactions on either two 6% gels run for different lengths of times or on both a 6% and an 8% gel (see annotation to step 26). Running two 6% gels allows the sequence to be read from 20 to 350 bases. To read the first few bases, it is necessary to use an 8% or 12% gel.

Load and run the sequencing gels

23. Resuspend dried samples in 10 μ l formamide loading buffer and vortex to dissolve any material adhering to the tube sides.
24. Heat the samples 2 to 3 min at 90°C.

It is recommended that a burner be used to heat a pan of water to 90°C. This provides a rapid supply of 90°C water and eliminates the need to maintain a water bath at this temperature.

25. Place the samples on ice and immediately layer onto the preheated sequencing gel.

Draw the sample up into a 20- μ l pipettor and place the tip (5 to 200 μ l) down on the surface of the lower (back) plate pointing into the well. The sample can easily be extruded without any special care and will sink to the bottom of the well. The loaded sample will not be a discrete band and will appear to be smeared throughout the well; however, this does not affect the quality of the separation. It is not necessary to change pipet tips between samples. Loading by this method can be accomplished much more quickly than by using a Hamilton syringe.

We suggest layering the samples in the following order: G, A, T, C (grouping purines and pyrimidines). This also facilitates reading the gels, as the lanes will actually contain G, G+A, T+C, C; thus, the related lanes G, G+A and T+C, C are adjacent.

26. Electrophorese the samples 1.5 to 4.5 hr at 40 to 45 V/cm (~70 W for a standard 30 × 40-cm gel.)

If two 6% gels are used, run one for 1.5 to 2 hr and the second one for 4 to 5 hr. This allows the sequence to be read from 20 to 350 bases. To read the first few bases, it is necessary to use an 8% or 12% gel electrophoresed for 1 hr. If a 6% gel and an 8% gel are used, run the 6% gel 3 hr and the 8% gel 1.5 hr.

Process the sequencing gels

27. Dry the gels.
28. Autoradiograph the dried gels on X-ray film with an intensifying screen 1 to 4 days at -70°C.

It is not necessary to fix the gel prior to drying as described in UNIT 7.6; however it will improve the clarity of gels containing ³²P.

29. Read the sequence from the film (see Fig. 7.5.1 and commentary).

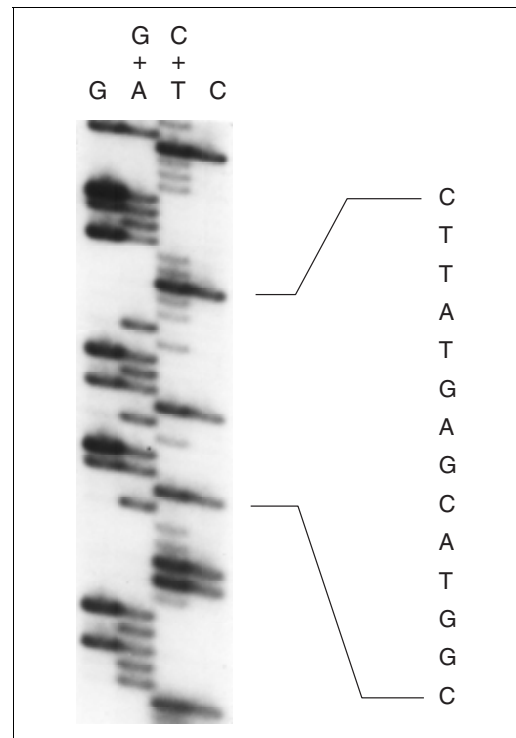


Figure 7.5.1 Autoradiogram of a chemical sequencing gel with a portion of the sequence indicated. This example does not contain any compressions.

***Tth*111I DIGESTION AND END LABELING**

This protocol describes end labeling DNA fragments cloned into pSP64CS or pSP65CS in preparation for chemical sequencing (Eckert, 1987). Because the *Tth*111I sites allow labeling one end of a fragment, no gel purification of labeled fragments is necessary, and the DNA can be used directly for chemical sequencing.

Additional Materials

- pSP64CS or pSP65CS recombinant plasmid DNA (prepared as in *UNIT 7.3*)
- Tth*111I restriction endonuclease and 10× buffer (*UNIT 3.1*)
- Klenow fragment labeling mix
- [α -³²P]dNTP (dGTP, dATP, dTTP, or dCTP with specific activity of 600 or 3000 Ci/mmol)
- Klenow fragment of *E. coli* DNA polymerase I (*UNIT 3.5*)
- TE buffer (*APPENDIX 2*)
- 25:24:1 phenol/chloroform/isoamyl alcohol (*UNIT 2.1*)
- Sephadex G-50 slurry, equilibrated in 5 mM Tris·Cl, pH 7
- Additional reagents and equipment to prepare Sephadex G-50 spin column (*UNIT 3.4*)

1. Mix the following in a microcentrifuge tube:

- 5 μ l pSP64CS or pSP65CS recombinant plasmid DNA
- 2 μ l 10× *Tth*111I buffer
- 12 μ l H₂O
- 1 μ l *Tth*111I.

Incubate 20 min at 60°C.

2. Cool tube to 15°C and add the following:

- 20 μ l Klenow fragment labeling mix
- 2 μ l desired [α -³²P]dNTP
- 1 U Klenow fragment.

Incubate 10 min at 15°C.

3. Add 70 μ l TE buffer to terminate reaction.
4. Extract with 50 μ l of 25:24:1 phenol/chloroform/isoamyl alcohol and save the top aqueous phase (*UNIT 2.1*).
5. Prepare two spin columns as follows (see also *UNIT 3.4*):
 - a. Fill two 1-ml syringe barrels (from plastic syringes) with Sephadex G-50 slurry, equilibrated in 5 mM Tris·Cl, pH 7.
 - b. Spin in tabletop centrifuge at 1500 rpm for exactly 4 min.
 - c. Add additional slurry to fill the syringes and centrifuge again at the same speed and time.
6. Apply the sample to one of the columns and centrifuge at 1500 rpm for exactly 4 min.
7. Save the column overflow and pass over the second column as in step 6.
8. Add 1 ml of 100% ethanol to the column overflow and place for 5 min in a dry ice/ethanol bath.
9. Microcentrifuge 5 min at 15,000 × *g* and discard supernatant.
10. Dry the sample and resuspend pellet in 40 μ l water. Use salt-free samples in step 5 of the sequencing (basic) protocol.

REAGENTS AND SOLUTIONS

Dimethyl sulfate (DMS) reaction buffer

50 mM sodium cacodylate, pH 8.0
1 mM EDTA
Store 6 months to 1 year at -20°C

DMS stop buffer

1.5 M sodium acetate, pH 7.0
1.0 M β -mercaptoethanol
Filter sterilize, then add tRNA to 100 $\mu\text{g}/\text{ml}$ final
Store indefinitely at -20°C

Formamide loading buffer

80% formamide
10 mM NaOH
1 mM EDTA
0.1% xylene cyanol
0.1% bromphenol blue
Store at room temperature or 4°C (prepare fresh periodically)

High-quality formamide that does not require deionization is available from Fluka, IBI, and American Bioanalytical.

Hydrazine stop buffer

0.3 M sodium acetate, pH 7.0
0.1 mM EDTA
Filter sterilize, then add tRNA to 25 $\mu\text{g}/\text{ml}$ final
Store indefinitely at -20°C

Klenow fragment labeling mix

40 mM Tris-Cl, pH 7.6
20 mM MgCl_2
20 mM dithiothreitol
Store 6 months to 1 year at -20°C

COMMENTARY

Background Information

Background information and a general discussion of chemical sequencing are provided in the introduction to Chapter 7 and *UNIT 7.1*. Detailed information about pSP64CS pSP65CS can be found in these locations, particularly in *UNIT 7.1*. It is also useful to refer to *UNIT 7.4*, which provides detailed protocols and troubleshooting guides for dideoxy sequencing, because some of this material is also applicable to the chemical sequencing method.

In the original method of Maxam and Gilbert (1980), 1 M piperidine formate was employed in the G+A reaction and the reaction took 60 min at 20°C . In the method presented here, undiluted formic acid is added followed

by a 5-min reaction. In addition, the piperidine strand cleavage of the G+A reaction is carried out later, along with the piperidine cleavage for the G, T+C, and C reactions. These modifications represent a considerable time savings when sequencing a large number of samples. In addition, the use of screw-top tubes for the piperidine reaction is a simple means of guaranteeing that these reactions work well.

Critical Parameters

General guidelines concerning critical aspects of DNA sequencing and the preparation of sequencing gels can be found in Table 7.5.2 and *UNITS 7.4 & 7.6*.

Two critical features unique to the chemical sequencing method are (1) removal of excess

Table 7.5.2 Troubleshooting Guide for Chemical Sequencing

Problem	Possible cause	Solution
Bands fade away at top of gel in one or more lanes	Chemical reactions proceeded too far	Decrease base-specific modification reaction times (do not alter the concentrations of the reagents, except DMS; see critical parameters)
Bands fade away at bottom of gel in one or more lanes	Chemical reactions did not proceed far enough	Increase base-specific modification reaction times (do not alter the concentrations of the reagents, except DMS; see critical parameters)
Shadow bands in all lanes	Incomplete filling-in by Klenow fragment of a 5' overhang created by digestion with a restriction endonuclease (usual cause) Two or more labeled nucleotides have been incorporated at the end of a fragment; decay of the "outer" nucleotide will produce a labeled fragment one base shorter	Repeat end-labeling reaction (<i>UNIT 3.5</i>)
Bands in all 4 lanes	Site-specific nicking of DNA fragment by restriction endonuclease or contaminating nuclease	Increase stringency of restriction endonuclease reaction by altering salt concentration and/or repurify DNA fragment prior to labeling
Bands in all 4 lanes and/or reduced or irregular spacing	Compressions due to secondary structure of oligodeoxynucleotide	Refer to commentary in <i>UNIT 7.6</i> for a discussion of how to recognize and resolve compressions in sequencing gels
Extraneous bands of lower intensity and irregular spacing	Impure DNA fragment. Irregular spacing results from different mobilities of oligos of same length but different base compositions	Repurify DNA fragment prior to labeling
Smearing in pyrimidine (T, T+C lanes)	Incomplete scission by piperidine; residual hydrazine present during piperidine reaction	Ensure that hydrazine is removed before piperidine addition and that piperidine reaction tubes are tightly sealed
Weak T reaction	Residual salt present in DNA sample	Wash DNA fragment thoroughly with 70% ethanol before sequencing
Bands in G, G+A, and T+C lanes	Hydrazine reaction with guanine residues at neutral pH in hydrazine stop mix if sample is not kept chilled	Make sure samples are well-chilled in step 11
Gap in sequence due to missing band	May be due to 5-methyl cytosine modification at <i>Eco</i> RII site [C ^{Me} C(A/T)GG] in DNA from a <i>dcm</i> ⁻ <i>E. coli</i> strain	Verify presence of <i>Eco</i> RII site or prepare DNA from a <i>dcm</i> ⁻ strain of <i>E. coli</i>

salt (steps 1 to 3) before initiating the base-specific modifications and (2) removal of piperidine (steps 19 to 21) following the strand-scission reaction and before loading the samples on the sequencing gels. Failure to remove salt will interfere with the reaction of hydrazine with thymidine residues and will result in weak T bands in the T+C sequencing lane. Failure to remove piperidine following the reactions will result in smearing of the bands in the sequencing gel and inability to read the sequence.

Another critical step is the addition of DMS to the G reaction (step 8). It is important to measure the small 1- μ l aliquot carefully and to be sure that the DMS mixes thoroughly into the reaction. It is sometimes necessary to use shorter reaction times for the DMS reaction than indicated in Table 7.5.1 and if there are many samples, it is difficult to add the 1- μ l DMS aliquots accurately. This problem can be overcome by adding an appropriate amount of DMS to the DMS reaction buffer (the DMS concentration may be lowered if the reaction is proceeding too fast) just before the base-specific modification reactions will be carried out. Then 200 μ l of the DMS reaction buffer containing DMS are added to each sample to start the reaction.

It is very important that the piperidine strand-scission reactions go to completion. This is facilitated by ensuring that the reaction tubes containing piperidine are tightly sealed during the 30-min 90°C incubation (step 18).

Finally, it may be convenient to add additional tRNA to the DMS and hydrazine stop buffers to facilitate recovery of the DNA pellets in the subsequent precipitation steps. Add tRNA to the DMS stop buffer to 600 μ g/ml and to the hydrazine stop buffer to 150 μ g/ml. The tRNA will not interfere with the piperidine reactions.

Troubleshooting

Some of the problems encountered in chemical sequencing are also encountered in the dideoxy method. For example, if the DNA fragment to be sequenced has not been sufficiently purified or is degraded, poor incorporation of label or a high background level in the sequencing gel may occur. When sequencing recombinant pSP64CS or pSP65CS, it is important that the plasmid DNA preparation method efficiently removes the bacterial genomic DNA. Refer to Table 7.5.2 and Kuebbing (1983) for brief troubleshooting guides to

chemical sequencing, and to Maxam and Gilbert (1980) for a more comprehensive discussion of chemical sequencing.

Anticipated Results

Reading a chemical sequencing gel pattern for the first time can be time-consuming; the following is a quick reference guide. As mentioned, it is helpful to layer the samples onto the gel by grouping purines and pyrimidines. Thus, a sample film resulting from the sequence of *G-A-T-C-G-G-A-C-C-T would appear as shown in Figure 7.0.3.

Reading the sequence can be thought of as a problem of subtraction (i.e., G+A minus G is equal to A; T+C minus C is equal to T). The G+A lane will contain all bands resulting from cleavage at Gs and As, while the G lane will contain only G-specific cleavages; therefore, any band in the G+A lane that is not paired with a band in the G lane are As. Likewise, any band in the T+C lane that is not paired with a band in the C lane is a T. In this way, the sequence can be assembled by reading from the bottom of the sequencing gel to the top. The largest fragment will be the uncleaved original fragment. This fragment will frequently be so large (≥ 1 kb) that it will not be resolved on the gel. In most cases, 1 to 100 bases can be comfortably read on a 12% gel, 20 to 200 bases on an 8% gel, and 50 to 400 on a 6% gel. A photograph of a chemical sequencing gel is shown in Figure 7.5.1.

Time Considerations

It is best to try the four reactions (G, G+A, T+C, C) initially with a single DNA practice sample before working with important DNA fragments. After gaining experience, sequencing reactions on as many as ten DNA fragments can be carried out in 4 hr. Generally, sequence reactions are carried out the first day, gels are run the second day, and sequence data are recorded on the third or fourth day.

Literature Cited

- Eckert, R. 1987. New vectors for rapid sequencing of DNA fragments by chemical degradation. *Gene* 51:245-252.
- Kuebbing, D. 1983. Diagnostic chemical DNA sequencing. *Focus* 5(2):1-3.
- Maxam, A. and Gilbert, W. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499-560.

Ornstein, D.L. and Kashdan, M.A. 1985. Sequencing DNA using ^{35}S labeling: A troubleshooting guide. *BioTechniques* 3:476-483.

Key Reference

Maxam and Gilbert 1980. See above.

A superb discussion of troubleshooting chemical sequencing reactions.

Contributed by Richard L. Eckert
Case Western Reserve University
Cleveland, Ohio

Denaturing Gel Electrophoresis for Sequencing

The accuracy of DNA sequence determination depends largely upon resolution of the sequencing products in denaturing polyacrylamide gels. This unit provides a detailed description of the setup, electrophoresis, and processing of such gels.

In general, the gels required for DNA sequencing are 40-cm long, of uniform thickness, and contain 4% to 8% acrylamide and 7 M urea (basic protocol). Modifications of the basic protocol increase the length of readable sequence information which can be obtained from a single gel. These include forming the gel with wedge-shaped spacers to create a field gradient (Ansorge and Labeit, 1984; Olsson et al., 1984), or incorporating a buffer gradient (first alternate protocol), an electrolyte gradient (second alternate protocol), or an acrylamide step gradient (Isfort and Ihle, 1988; Williams et al., 1986) into the gel. It is also possible to use longer gels (80 to 100 cm), although these can be technically challenging to pour and handle.

Another modification to the basic protocol—inclusion of formamide in the sequencing gel—is designed to overcome gel compressions arising from secondary structure in the sequencing products during gel electrophoresis (third alternate protocol). Formamide concentration ranges of 25% (Brown, 1984) to 40% (USB, 1990) can be used to alleviate gel compressions. See critical parameters and troubleshooting for a discussion of acrylamide concentrations and electrophoresis conditions.

CAUTION: Acrylamide and bisacrylamide are neurotoxins and should be handled with gloves. Dimethyldichlorosilane should also be handled with gloves and plate treatment with this solution should be carried out in a fume hood (see silanization, *APPENDIX 3*). TEMED and formamide should also be handled with care.

BASIC PROTOCOL

POURING, RUNNING, AND PROCESSING SEQUENCING GELS

This protocol describes preparing the sequencing plates, pouring the sequencing gels, and loading, running, and processing the gels to analyze a set of sequencing reactions (*UNITS 7.4 & 7.5*).

The steps are the same for all acrylamide concentrations; formulations for 4%, 6% and 8% gels are provided in reagents and solutions. Refer to critical parameters for a discussion of choosing the appropriate acrylamide concentration.

Materials

- 70% ethanol *or* isopropanol in squirt bottle
- 5% dimethyldichlorosilane (diluted in CHCl₃; Sigma #D-3879)
- Denaturing acrylamide gel solution
- TEMED
- 10% (w/v) ammonium persulfate (made fresh weekly and stored at 4°C)
- 1× TBE buffer, pH 8.3-8.9 (*APPENDIX 2*)
- Sequencing samples in formamide/dye solution (*UNITS 7.4 or 7.5*)
- 5% acetic acid/5% methanol (vol/vol) fixer solution
- 30 × 40-cm front and back gel plates
- 0.2- to 0.4-mm uniform-thickness spacers
- Large book-binder clamps
- 60-ml syringe
- Pipet tip rack or stopper
- 0.2- to 0.4-mm sharktooth or preformed-well combs
- Sequencing gel electrophoresis apparatus

Denaturing Gel Electrophoresis for Sequencing

7.6.1

Contributed by Barton E. Slatko and Lisa M. Albright

Current Protocols in Molecular Biology (1992) 7.6.1-7.6.13

Copyright © 2000 by John Wiley & Sons, Inc.

Pasteur pipet *or* Beral thin stem
 Power supply with leads
 Sequencing pipet tip
 Gel dryer
 Shallow fixer tray
 46 × 57–cm gel blotting paper (e.g., Whatman 3MM)
 Kodak XAR-5 X-ray film

NOTE: Many companies provide equipment needed for sequencing experiments; a list of suppliers is provided in Table 7.6.1.

Assemble gel sandwiches

1. For each gel, meticulously wash a pair of 30 × 40–cm front and back gel plates with soap and water; rinse well with deionized water and dry.

A thorough washing and rinsing is critical. Soap residue decreases polymerization time, thus risking polymerization while pouring the gel. Dust and other particulates (e.g., residue from previous gels) left on the plates will make it extremely difficult to pour a gel that does not contain bubbles. Particulates in the gel will also distort the banding pattern.

2. Wet plates with 70% ethanol or isopropanol in a squirt bottle and wipe dry with a Kimwipe or other lint-free paper towel.
3. Wearing gloves, apply a film of 5% dimethyldichlorosilane in CHCl₃ to one side of each plate by wetting a Kimwipe with the solution and wiping the whole plate carefully. After the film dries, wipe with 70% ethanol or isopropanol and dry with a Kimwipe. Check plates a final time for dust and other particulates.

Perform this step in a fume hood.

There are two reasons for silanizing the plates: to facilitate pouring a gel without bubbles and to prevent the gel from sticking to the plates during postelectrophoresis processing (see critical parameters for further discussion of silanization strategies).

Table 7.6.1 Suppliers of Sequencing Gel Electrophoresis Equipment^a

Equipment	Supplier ^b
Sequencing gel apparatus, including necessary clamps, combs and spacers	AAP, ABA, BR, GB, HO, IBI, JS, OSP, PH, SS
Gel tape	ABA, GB, HS, IBI, OSP
Power supplies	ABA, ACS, BR, EC, FD, GB, HO, IBI, IS, OSP, PH, SS, ST
Beral thin stem	BE
Ultrathin pipet tips for loading gels	ABA, BR, CO, DR, DY, HO, IBI, IS, MB, PH, SS, ST
Gel thermometer	BR
Shallow fixer tray	OSP
Transfer paper	ABA, SS, WH
Gel dryers	ATR, BR, HO, SV

^aModified from Slatko, 1991a.

^bAbbreviations: AAP, Ann Arbor Plastics; ABA, American Bioanalytical; ACS, Accurate Chemical and Scientific; ATR, ATR; BE, Beral Enterprises; BR, Bio-Rad; CO, Costar; DR, Drummond; DY, Dynalab; EC, EC Apparatus; FD, Fotodyne; GB, Gibco/BRL; HO, Hoefer; IBI, International Biotechnologies; IS, Integrated Separation Systems; JS, Jordan Scientific; MB, Marsh Biomedical; OSP, Owl Scientific Plastics; PH, Pharmacia LKB; SS, Schleicher & Schuell; ST, Stratagene; SV, Savant; WH, Whatman. See APPENDIX 4 for addresses of suppliers.

4. Assemble gel plates according to the manufacturer's instructions with 0.2- to 0.4-mm uniform-thickness spacers and large book-binder clamps, making certain the side and bottom spacers fit tightly together.

A good fit between side and bottom spacers is crucial to prevent leaks. If necessary, use a small amount of acrylamide or 1% agarose to seal the edges and bottom of the gel. Alternatively, before the top gel plate is placed on the spacers and bottom plate, put a very small dab of high-temperature vacuum grease at the point where the bottom spacer meets the side spacers (the grease can be removed later with 70% ethanol).

Gel sandwiches may also be prepared by using two side spacers only (no bottom spacer) and taping the plates together along the sides and bottom using yellow vinyl electrical tape. To prevent leakage, tape the bottom of the sandwich a second time forming "hospital bed corners." It is important that the tape make good tight contact along its length with the sides of the gel plates and that any bubbles are smoothed out.

Prepare gel solution and pour sequencing gel

5. For each gel, prepare 60 ml of the desired denaturing acrylamide gel solution in a 100-ml beaker.

The gel mix can be heated to speed dissolution of the urea; however, to prevent degradation of the acrylamide, do not heat over 55°C. Allow the solution to cool to room temperature ($\leq 25^\circ\text{C}$) to prevent polymerization while pouring the gel.

If particulate matter remains after mixing, filter the solution through Whatman No. 1 filter paper in a funnel.

6. Thoroughly mix 60 μl TEMED, then 0.6 ml of 10% ammonium persulfate, into each acrylamide solution. Add these components immediately before pouring each gel (step 7).

These reagents initiate polymerization and should not be added earlier than this stage. To achieve slower polymerization (an advantage when learning to pour gels), the amounts of TEMED and ammonium persulfate can be reduced to 40 μl and 0.4 ml, respectively.

To check the freshness of the ammonium persulfate solution and TEMED, test the polymerization time. Transfer 0.6 ml of the gel mix to a 1.5-ml microcentrifuge tube and add 0.6 μl TEMED and 6 μl of 10% ammonium persulfate. Mix the solution by inverting the tube several times. Polymerization should occur within 5 min. If it does not, make fresh ammonium persulfate and test another 0.6 ml sample. The amount of TEMED can also be adjusted.

7. Pour the gel immediately. Gently pull the acrylamide solution into a 60-ml syringe, avoiding bubbles. With the short plate on top, raise the top of each gel sandwich to a 45° angle from the benchtop and slowly expel the acrylamide between the plates along one side. Adjust the angle of the plates so that the gel solution flows slowly down one side.

Other methods of delivering the gel solution include using a 25-ml pipet or pouring the solution through the sidearm of a 125-ml sidearm flask. Air bubbles can be prevented by pouring the gel solution at a steady pace while maintaining constant contact between the fluid being poured and the fluid already between the plates. Any air bubbles that form can be removed by tapping the glass plates behind the bubbles or by rocking the plates until the bubbles move to the top of the gel solution.

8. When the gel solution reaches the top of the short plate, lay the gel sandwich down so that the top edge is ~5 cm above the benchtop; place an empty disposable pipet tip rack or stopper underneath the sandwich to maintain the low angle.
9. Insert the flat side of a 0.2- to 0.4-mm sharktooth comb into the gel solution 2 to

3 mm below the top of the short plate, being very careful to avoid bubbles. Use book-binder clamps to pinch the combs between the plates so that no solidified gel forms between the combs and the plates. Layer extra acrylamide gel solution onto the comb to ensure full coverage. Rinse the syringe with water to remove excess acrylamide.

The flat side of the sharktooth comb creates a flat surface at the top of the gel. Be careful not to insert the comb too far into the gel. This will make it difficult to insert the teeth all the way into the gel when the comb is reversed in step 18. Combs with preformed wells are also commercially available. If this type of comb is used, insert the teeth 2 to 3 mm below the top of the short plate after the gel is poured and pinch with book-binder clamps as described above.

If tape was used to seal the gel sandwiches, clamp the sides of the gel together with book-binder clamps in the positions where the gel will be clamped to the electrophoresis apparatus. Place the clamps over the side spacers and not over the gel itself. This minimizes distortion of the gel in the apparatus (see critical parameters).

10. Check for leaks in the gels and slow or stop them by placing book-binder clamps at these spots. Observe polymerization of the gel solution.

Polymerization requires 10 min to 2 hr, depending on the temperature of the room (i.e., cold acrylamide polymerizes more slowly) and the amount of TEMED and ammonium persulfate added. When polymerization is complete, a Schlieren line will be visible at the border of the comb and the surface of the gel.

11. Use immediately or store gels ≤ 48 hr at room temperature.

To store gels, place a paper towel dampened with 1× TBE buffer over the comb. Put entire gel sandwich in a sealed plastic bag or wrap top of the gel tightly with plastic wrap.

Set up the sequencing gel

12. Remove the bottom spacer or the tape at the bottom of each gel sandwich.
13. Remove extraneous polyacrylamide from around the combs with a razor blade. Clean spilled urea and acrylamide solution from plate surfaces with water.

This step helps to remove the comb cleanly and prevents particles of polymerized acrylamide from falling onto the gel surface.

14. Remove sharktooth combs gently from each gel sandwich, avoiding pulling or stretching the top of the gel. Clean combs with water so they will be ready to be reinserted in step 18.

If a preformed-well comb was used, take care to prevent tearing of the polyacrylamide wells. This comb will not be reinserted.

15. Fill the bottom reservoir of each gel apparatus with 1× TBE buffer such that the gel plates will be submerged 2 to 3 cm in the buffer.

16. Place each gel sandwich in a sequencing electrophoresis apparatus and clamp the plates to the support per manufacturer's instructions.

If a bottom spacer was used, a large air bubble will often get trapped at the bottom of the gel. Place the gel sandwich on one bottom corner in the apparatus and slowly lower the sandwich to chase out a large portion of this air bubble. Squirt buffer under the plates using a syringe with a bent 20-G needle to remove any remaining bubbles.

17. Pour 1× TBE buffer into the top reservoir to ~3 cm above the top of the gel. Rinse the top of the gel with 1× TBE buffer using a Pasteur pipet or Beral thin stem to remove fragments of excess polyacrylamide and urea leached from the gel.

Check for leaks from the top buffer tank. Pieces of clay or agarose can be used to seal leaks

from the top buffer tank gasket. To prevent “smiling” (distortion) of the gel bands, check that the buffer level is horizontal and parallel to the bottom of the gel (see critical parameters/troubleshooting).

18. Reinsert the teeth of the cleaned sharktooth combs into the gel sandwich with the points just barely sticking into the gel. Using a Pasteur pipet or Beral thin stem, rinse the wells thoroughly with 1×TBE buffer to remove stray fragments of polyacrylamide.
19. Preheat the gels by turning on the power supplies to 45 V/cm, 1700 V, 70 W constant power ~30 min before loading sequencing samples.

Each commercial gel apparatus will have its own recommendations about voltage and power settings. In general, power settings should be 45 W to 70 W with constant power.

Load and run the gel

20. Rinse wells just prior to loading the gels to remove urea that has leached into them.
21. Cover and heat completed sequencing samples in formamide/dye solution for 2 min at 95°C, then place on ice.
22. Load 2 to 3 µl of each sample per well on each gel to be run. Rinse sequencing pipet tip twice in the lower reservoir after dispensing from each reaction tube.

It is helpful to mark the outside glass plate with individual lanes or sets of four lanes to keep track of the loading order.

Sample loading using ultrathin DNA sequencing pipet tips is recommended (Table 7.6.1). With practice, all samples can easily be loaded in ≤5 min, eliminating the need for electrophoresing one set of samples into the gel before loading the next set.

23. Run gels at 45 to 70 W constant power according to the manufacturer’s recommendations. Maintain a gel temperature of ~65°C.

Temperatures higher than this can result in cracked plates or smeared bands, while too low a temperature can lead to incomplete denaturation of the sequencing products. Monitor the temperature by taping a round-faced gel thermometer with a flat back surface (Table 7.6.1) to the glass plate with paper tape. Dab a spot of vacuum grease between the glass plate and the thermometer to assist in heat conduction.

24. Observe the migration of the marker dyes (Table 7.6.2) to determine the length of the electrophoresis.

If sequencing reactions are loaded on multiple gels, the gels must be coordinated such that sequences in common from the long and short runs will allow overlapping of the sequence, yet still maximize the amount of sequence obtained from the long run (see critical parameters/troubleshooting for examples of electrophoresis times).

Table 7.6.2 Migration of Oligodeoxynucleotides (Bases) in Denaturing Polyacrylamide Gels in Relation to Dye Markers

Polyacrylamide (%)	Bromphenol blue	Xylene cyanol
5	35 b	130 b
6	26 b	106 b
8	19 b	75 b
10	12 b	55 b

Process and dry the gel

Many mishaps occur when processing sequencing gels because they are very thin and fragile and can easily be torn. Minimal handling is essential in the following steps.

25. Fill the dry ice traps attached to the gel dryer (if required) and preheat the dryer to 80°C.
26. After electrophoresis of each gel is complete, drain the buffer from the upper and lower reservoirs of the sequencing apparatus and discard the liquid as radioactive waste.

If two gels are run for a set of sequencing reactions, the difference in electrophoresis times allows processing and drying of each gel as soon as it is finished.

27. Remove the gel sandwich from the electrophoresis apparatus and place under cold running tap water until the surfaces of both glass plates are cool.

This facilitates handling of the gels and prevents the gels from “curling” as they cool.

28. Lay the sandwich flat on paper towels with the notched (short) plate up. Remove excess liquid and remaining clamps or tape.

29. Remove one side spacer and insert a long metal spatula between the glass plates where the spacer had been. Pry the plates apart with a gentle rocking motion of the spatula. The gel should stick to the bottom plate. If it sticks to the top plate, flip the sandwich over. Slowly lift the top plate from the side with the inserted spatula, gradually increasing the angle until the top plate is completely separated from the gel.

If the gel sticks to the top plate in an isolated spot, a stream of water from a squirt bottle can be sprayed at the spot to aid separation. Should gels recurrently stick to both glass plates during processing, more thorough cleaning of the glass plates prior to electrophoresis and/or resilanization of the plates may be necessary.

30. Once the plates are separated, remove the second side spacer and any extraneous bits of polyacrylamide around the gel.

31. If samples contain ³²P, the fixing steps (31 and 32) are optional; proceed to step 33. If samples contain ³⁵S, transfer the gel on the glass plate to a shallow fixer tray and gently cover to a depth of 2 cm with 5% acetic acid/5% methanol fixer solution. Soak the gel 10 to 15 min. Gently rock the tray periodically to gradually loosen the gel from the glass plate.

Urea quenches the ³⁵S signal and must be removed from ³⁵S-containing gels as described here. Although not required, removing the urea will also increase the clarity of ³²P-containing gels.

Occasionally, one or two spots on the gel will stubbornly stick to the glass plate. In this case, free the gel from the plate by moving it gently with a gloved hand. Extreme care must be taken at this stage to avoid stretching or tearing the gel.

A tray with a hole in the bottom through which the fix solution can be drained is useful (Table 7.6.1).

32. Reposition the gel over the plate and remove the fixer solution by aspiration or gravity. Take care to keep the gel centered over the glass plate at this point. Carefully lift the plate with the gel on top from the tray.

33. Place gel on a benchtop. Hold two pieces of dry blotting paper together as one piece. Beginning at one end of the gel and working slowly towards the other, lay the paper on top of the gel. Take care to prevent air bubbles from forming between the paper and the gel.

34. Peel the blotting paper up off the plate—the gel should come with it. Gradually curl the paper and gel away from the plate as it is being pulled away.

Whatman 3MM seems to work best for this procedure.

Optional: this method is more reliable but slightly more tedious. Set another glass plate on top of Schleicher & Schuell #GB002 blotting paper and center it over the gel. Flip the entire sandwich over so that the gel is now resting on top of the blotting paper. Position the sandwich with the glass plates extended ~10 cm past the edge of the benchtop. Slowly slide only the bottom plate back onto the benchtop. Without the support of the bottom plate, the weight of the gel and blotting paper causes them to slowly peel away from the top plate. Reposition the sandwich so that the bottom glass plates extend ~10 cm over the edge of the benchtop (the top glass plate is now extended ~20 cm over the edge) and repeat the process. Lift the top plate from the portion of the gel still in contact with it. If the gel persists in sticking to the top plate, use a stream of water from a squirt bottle to loosen the gel.

35. Place the paper and gel on the preheated gel dryer. Cover with plastic wrap.

Remove any bubbles between the plastic wrap and the gel by gently rubbing the covered surface of the gel from the middle toward the edges with a Kimwipe.

36. Dry the gel thoroughly 20 min to 1 hr at 80°C. Peel the plastic wrap away.

The time required for drying depends on the efficiency of the dryer vacuum. When the gel is completely dry, the plastic will easily peel off without sticking. The plastic wrap will absorb ³⁵S emissions and must be removed prior to autoradiography. Autoradiography without plastic wrap will improve the clarity of sequencing ladders from reactions containing ³²P.

If a gel dryer is not available, place the paper and gel on a glass plate. Clip the paper to the plate to prevent curling. Dry in a forced-air oven or overnight at room temperature.

37. Place each dried gel in a separate X-ray cassette with Kodak XAR-5 film in direct contact with the gel and autoradiograph at room temperature.

Autoradiography is described in APPENDIX 3. No intensifying screen is used. Equivalent films from other manufacturers can be used.

If necessary, several sheets of paper towels or filter paper can be layered in the cassette before enclosing the dried gel and X-ray film. This ensures that the X-ray film and gel are tightly abutted during exposure in the X-ray cassette. Take care not to overpack the cassette which may cause light leaks around the cassette edges.

38. After sufficient exposure time (usually overnight), remove the X-ray film and process according to manufacturer's instructions.

ALTERNATE PROTOCOL

BUFFER-GRADIENT SEQUENCING GELS

Increasing the ionic strength in a sequencing gel decreases the voltage gradient and therefore decreases the rate of migration of DNA fragments (Biggin et al., 1983). In this protocol, the sequencing gel contains a higher concentration of buffer at the bottom of the gel than at the top and as a result, the migration of shorter oligodeoxynucleotides (at the bottom) are slowed down relative to the longer oligodeoxynucleotides (at the top). This allows the gel to be run longer without losing the shorter oligodeoxynucleotides off the bottom and improves the resolution between longer oligodeoxynucleotides.

Additional Materials

Buffer-gradient gel solutions (containing 0.5× and 2.5× TBE; see reagents and solutions)

25-ml pipet equipped with a rubber pipet-filler bulb

1. Assemble one gel plate sandwich as in steps 1 to 4 of basic protocol.
2. In two 100-ml beakers, prepare buffer-gradient gel solutions: 50 ml using 0.5× TBE (clear solution) and 25 ml using 2.5× TBE (blue solution). Heat gently at $\leq 50^{\circ}\text{C}$ while stirring until all the solids have dissolved. Cool the solutions until they are room temperature ($\leq 25^{\circ}\text{C}$).
3. Add 20 μl TEMED and 200 μl of 10% ammonium persulfate to the 0.5× TBE/gel solution and mix gently.
4. Add 10 μl TEMED and 100 μl of 10% ammonium persulfate to the 2.5× TBE/gel solution and mix gently.
5. Using a 25-ml pipet equipped with a rubber pipet-filler bulb, pull up 12.5 ml of the clear 0.5× TBE/gel solution followed gently by 12.5 ml of the blue 2.5× TBE/gel solution.
6. Allow three or four air bubbles to be taken up into the pipet by gently squeezing the suction inlet on the rubber bulb.
This causes the two layers to gently mix at the interface.
7. Release the solution down the glass plate sandwich in a gentle, even manner.
8. Using the same 25-ml pipet, fill the rest of the gel with the 0.5× TBE/gel solution.
9. Insert combs, position clamps, and observe polymerization as in steps 9 and 10 of basic protocol.
10. Run and process the gel as in steps 11 to 38 of basic protocol.

ELECTROLYTE-GRADIENT SEQUENCING GELS

An alternative way to generate an ionic-strength gradient in a gel has been developed by Sheen and Seed (1988). In this protocol, the ionic strength of the bottom of the gel is increased simply by increasing the salt concentration in the bottom buffer chamber. During the run, the salt is electrophoresed into the gel and generates a reproducible and effective gradient.

Additional Materials

0.5× and 1× TBE buffer (*APPENDIX 2*)
3 M sodium acetate unbuffered

1. Pour and prerun the sequencing gel as described in steps 1 to 19 of basic protocol. Use 0.5× TBE buffer in the top and 1× TBE buffer in the bottom reservoir.
2. Prepare and load sequencing samples as in steps 20 to 22 of basic protocol.
- 3a. If ≤ 400 bases of sequence information is needed, add 3 M sodium acetate to the bottom reservoir to a final concentration of 1 M. Proceed to step 4.
- 3b. If ≥ 400 bases of sequence information is desired, proceed to step 4 and wait 2 to 3 hr after beginning the electrophoresis to add 3 M sodium acetate to the bottom reservoir to a final concentration of 1 M.
4. Run gels at 60 W constant power.
It will take about 75% longer than usual for the dye markers to migrate to the same location on the gel.
The temperature of the gel should be monitored carefully. If the gel becomes significantly hotter at the top than the bottom, reduce the power.
5. Process the gel as described in steps 25 to 38 of basic protocol.

**ALTERNATE
PROTOCOL**

DNA Sequencing

7.6.8

FORMAMIDE-CONTAINING SEQUENCING GELS

In this protocol, formamide is added to the acrylamide gel solution (U.S. Biochemical, 1990) to destabilize the secondary structures of sequencing products that can cause the anomalous migration of bands known as compressions (*UNIT 7.4*).

Additional Materials

Formamide gel solution
5% acetic acid/20% methanol (v/v) fixer solution

1. Assemble gel sandwich as in steps 1 to 4 of basic protocol.
2. Prepare formamide gel solution according to the acrylamide concentration desired. Heat gently with stirring until dissolved. Cool solution until it is $\leq 30^{\circ}\text{C}$.

Do not let the temperature of gel solution exceed 55°C .

Filter solution through Whatman No. 1 filter paper if particulate matter is visible.

3. Add 0.15 ml TEMED and 1 ml of 10% ammonium persulfate. Immediately pour the solution into the gel sandwich. Observe polymerization.

The gel solution is viscous. Hold plates at a nearly vertical angle while pouring the solution—this is easiest if the solution is in a beaker.

The gel should polymerize within 30 min.

4. Run the gel 45 to 70 W, constant power.

This requires a higher voltage (60% higher) than nonformamide-containing gels. The DNA migrates about half as fast.

5. Process and dry the gel as described in steps 25 to 38 of basic protocol, except fix the gel 15 min in 5% acetic acid/20% methanol fixer solution in step 31a.

20% methanol prevents the gel from swelling during the fixing step.

REAGENTS AND SOLUTIONS

38% acrylamide/2% bisacrylamide

Mix 38 g of acrylamide (ultrapure) and 2 g *N,N'*-bismethylene acrylamide (ultrapure) in ~60 ml H_2O . Heat if necessary to dissolve, but do not heat above 55°C . Adjust volume to 100 ml. Store in a brown jar for up to several months at 4°C . Discard if an ammonia smell is detected.

CAUTION: Acrylamide is a neurotoxin. Handle with gloves.

Ultrapure acrylamide and bisacrylamide are available from numerous commercial sources. Many of these suppliers also offer premixed acrylamide/bisacrylamide in both solid and liquid form.

Deionization and filtering are usually not needed when using ultrapure reagents. If deionization is desired, add 5 g Amberlite MB-1 (Sigma) or equivalent mixed bed resin and stir 30 min at room temperature. Filter through Whatman No. 1 paper.

Buffer-gradient gel solutions

0.5× TBE (clear solution):

Reagent	Acrylamide concentration (%)		
	4	6	8
Urea (ultrapure) (g)	21	21	21
38% acrylamide/2% bisacrylamide (ml)	5.0	7.5	10.0
10× TBE (APPENDIX 2) (ml)	2.5	2.5	2.5
H ₂ O (ml)	26.5	24	21.5
Total volume (ml)	50.0	50.0	50.0

2.5× TBE (blue solution):

Reagent	Acrylamide concentration (%)		
	4	6	8
Urea (ultrapure) (g)	10.5	10.5	10.5
38% acrylamide/2% bisacrylamide (ml)	2.5	3.75	5.0
10× TBE (APPENDIX 2) (ml)	6.25	6.25	6.25
Sucrose (g)	2.5	2.5	2.5
1% (wt/vol) bromphenol blue (μl)	250	250	250
H ₂ O (ml)	6.25	5.0	3.75
Total volume (ml)	25.0	25.0	25.0

Denaturing acrylamide gel solution

Reagent	Acrylamide concentration (%)		
	4	6	8
Urea (ultrapure) (g)	25.2	25.2	25.2
38% acrylamide/2% bisacrylamide (ml)	6.0	9.0	12.0
10× TBE (APPENDIX 2) (ml)	6.0	6.0	6.0
H ₂ O (ml)	27	24	21
Total volume (ml)	60	60	60

Filter solution through Whatman No. 1 filter paper. Quantities are for a single sequencing gel. If gels are poured daily, make gel solutions in quantity—e.g., make 1 liter of gel solution by multiplying the above quantities by 16.7. Store 2 to 4 weeks at 4°C. Solutions of acrylamide deteriorate quickly, especially when exposed to light or left at room temperature.

Formamide gel solution

Reagent ^a	Acrylamide concentration (%)		
	4	6	8
Urea (ultrapure) (g)	42	42	42
38% acrylamide/2% bisacrylamide (ml)	10	15	20
10× TBE (APPENDIX 2) (ml)	10	10	10
Deionized formamide (ml) ^b	40	40	40
H ₂ O (ml)	10	5	0
Total volume (ml)	100	100	100

^aHeat the ingredients in a beaker while stirring. When dissolved, add H₂O to 100 ml total volume.

^bDeionize formamide by stirring with Amberlite MB-1 (Sigma) or equivalent mixed-bed resin 30 min at 4°C. Filter through Whatman No. 1 filter paper and store at -20°C. Formamide that has been deionized is available from American Bioanalytical.

COMMENTARY

Background Information

Refer to the Chapter 2 introduction for a general discussion of gel electrophoresis of nucleic acids and gels as electric circuits.

Critical Parameters and Troubleshooting

Silanizing the plates

After electrophoresis, the gel is supported by one of the glass plates during processing. As the gel sandwich is disassembled, however, the gel often sticks to both plates because polymerized acrylamide is inherently sticky, especially after being heated during electrophoresis. Silanization helps to prevent this common problem by aiding the release of the gel from one of the plates when the gel sandwich is taken apart; annotations to the steps describing processing of the sequencing gel offer suggestions for coping when the problem does occur. The basic protocol recommends silanizing both gel plates; an alternative strategy is to silanize one plate (usually the short, notched plate) to increase the chance that gel will stick to untreated plate.

An alternate silanizing agent is Sigmacote (Sigma #SL-2), which requires overnight drying or drying at 90°C (see manufacturer's instructions). The advantage of Sigmacote is that it is relatively permanent and only has to be reapplied when the gel starts sticking to the glass plates during subsequent processing. The untreated side of each plate should be marked with a piece of tape to make sure that the silanized side always faces the gel.

Choosing acrylamide gel concentrations

As the percentage of acrylamide in the gel decreases, the pore size increases, allowing resolution of larger oligonucleotides. Thus, lower acrylamide percentages (4% to 6%) allow more sequence to be read from a single gel than does an 8% gel. However, these gels are harder to handle due to increased fragility and softer texture, and the banding pattern may not be quite as sharp as those observed in 8% gels.

We recommend running 6% acrylamide, nongradient gels as a starting point for the novice. Two gels can be run, one for a relatively short time to retain the shorter oligonucleotides, the second for a longer time to maximize separation of the longer oligonucleotides. This two gel procedure, using 6% acrylamide, al-

lows 300 to 350 bases of sequence information to be obtained from a single set of sequencing reactions. If fewer than five sets of sequencing reactions have been performed, one 6% gel can be used for both the short and the long runs (the typical gel format has room for ten sets of sequencing reactions, or 40 samples). An alternative practice is to use an 8% gel for the shorter oligonucleotides and a 6% gel for the longer oligonucleotides (see below for a discussion of electrophoresis times for each of the above examples). As experience is gained with manipulation of gels, modifications of the basic protocol, as described in the introduction, can be incorporated to increase the amount of sequence information gained from each set of reactions.

Precipitating TBE

10× TBE buffer at pH 8.3 to 8.6 (as described in APPENDIX 2) tends to precipitate during long-term storage. Adjusting the pH to 8.9 with NaOH prevents this precipitation without affecting the resolution of the sequencing gel (Mayeda and Krainer, 1991). The TBE buffer in the gel and buffer reservoirs should be prepared from the same 10× stock.

Leaking combs

Sharktooth combs have a smooth edge that forms the top of the gel during polymerization and a "sharktooth" edge containing teeth that form the sample wells when the points of the teeth are inserted into the top of the gel after polymerization. Lanes created by sharktooth combs are spaced more closely together than those made with a preformed-well comb, thus allowing more samples to be loaded on a gel and yielding a sequencing ladder that is easier to read. Care must be taken to preserve the points of the teeth since they will make the wells. Even when all the teeth are intact, sharktooth combs have a tendency to leak between the wells because variations in the comb's thickness can result in a space between the teeth and the glass plates when the combs are reinserted to form wells. Leakage can be tested before loading samples by applying a small amount of loading dye into each well and observing whether the dye flows into adjacent lanes. Those wells that leak can then be omitted from use. This test dye should be run into the gel a short distance before applying the actual samples so that loading can be tracked. Alternatively, for minor leaking, a set of four reactions can be

loaded and run into the gel before the next set is applied. Another trick is to apply a tiny amount of Vaseline on the points of the teeth before reinserting them into the gel. This seals the bottoms of the wells.

“Smiling” gels

Level, “nonsmiling” (nondistorted) bands require even heat dissipation across the gel. To achieve this, both buffer surfaces and the bottom and top of the gel must be aligned parallel to each other. The alignment of the bottom buffer can be a particular problem because it tends to climb up between the back glass plate and the electrophoresis apparatus due to capillary action when the gel sandwich is clamped to it. This can be prevented by inserting a small piece of folded filter paper between the gel sandwich and the electrophoresis apparatus, ~5 cm above the level of the bottom buffer. This forces the bottom of the sandwich away from the apparatus, allowing the bottom buffer to form a level surface.

In addition, deformation of the gel in the apparatus can contribute to uneven heat dissipation. If the gel sandwich is assembled with tape rather than clamps, distortion can be prevented during polymerization by clamping the gel plates with book-binder clamps at the same positions in which they will attach the gel sandwich to the electrophoresis apparatus. This maintains the same pressure points on the gel when it is polymerizing. Some commercially available gel setups allow the gel to be poured in the gel apparatus, avoiding this complication.

To ensure even conduction of the heat generated during electrophoresis, an aluminum plate (0.4 cm thick, 34 × 22 cm) can be clamped onto the front glass plate with the same book-binder clamps used to hold the gel sandwich to the apparatus. The aluminum plate must be positioned so that it does not touch any buffer during electrophoresis. Some commercially available gel setups have this feature, or have a top buffer reservoir that extends down the rear of the apparatus to evenly distribute heat over the gel.

Reading beyond 350 nucleotides

To unambiguously read DNA sequence beyond 350 nucleotides, it is essential to prevent distorted bands as described above. In addition, a concentration of ≥5% glycerol in the gel samples results in distortion of bands in the 450- to 550-nucleotide range (Tabor and

Richardson, 1987); this can be avoided by diluting the sequencing polymerase in diluent without glycerol (UNIT 7.4). The modifications to the basic protocol resulting in various types of gradient gels (see background information) will generally increase the ability to read the sequence in the 350-nucleotide range. A rather infrequently used tactic is to load duplicate sets of each reaction in the order GATCGTAC; this facilitates correct ordering of bands ≥350 nucleotides because each reaction is run on the gel next to each of the other three reactions.

Electrophoresis times and coordinating multiple gels

To decide how long to run the gel that will retain the shortest oligonucleotides (“short” gel), the length of the primer and how far away from the beginning of the target DNA synthesis will be initiated must be known. Add those two numbers together and use Table 7.6.2 as a guideline for bromphenol blue dye migration. Adjust the electrophoresis time so that some polylinker sequences can be retained on the gel and identified. The “long” gel should be run so that the shortest oligonucleotides remaining on this gel overlap by at least ten bases with the longest oligonucleotides on the short gel.

Experience is the best guide in the coordination of the gels. The following examples may serve as initial references.

If two 6% gels and the –40 sequencing primer (a 24-mer that initiates DNA synthesis 40 nucleotides from the M13 polylinker; New England Biolabs) are used, run the first gel 10 to 15 min after the bromphenol blue runs off the bottom of the gel. Run the second gel until the xylene cyanol is almost at the bottom of the gel (~5 hr).

If an 8% gel is used for the shorter oligonucleotides and a 6% gel for the longer oligonucleotides, run the 8% gel until the bromphenol blue runs off and the xylene cyanol has run through ~40% of the gel (~2 hr). Run the 6% gel until the xylene cyanol is at the bottom of the gel (~5 hr).

If the same 6% gel is used for both the short and long runs, load one set of reactions and run the gel until the bromphenol blue is at the bottom. Turn off the power to the gel and load a duplicate set of reactions on the unused part of the gel. Continue running the gel until the duplicate bromphenol blue is at the bottom of the gel. The xylene cyanol dye from the first set of reactions should be at the bottom of the first (long) run.

Time Considerations

Pouring, running, and processing of a sequencing gel is usually interspersed with performing sequencing reactions (UNIT 7.4). The typical gel format has room for ten sets of reactions generating 40 samples. Larger gel formats (up to 78 lanes) are available commercially but we recommend that the novice run and process no more than two 40-sample format gels at one time. This can be accomplished in a single long day, typically 7 to 8 hr for two gels. An experienced sequencer can typically perform the process for four gels in 9 to 10 hr. Sequencing gels can be prepared up to 48 hr in advance of use; assembling and pouring the gel takes ~1 hr.

Literature Cited

- Ansorge, W. and Labeit, S. 1984. Field gradients improve resolution in DNA sequencing gels. *J. Biochem. Biophys. Methods* 10:237-243.
- Biggin, M.D., Gibson, T.J., and Hong, G.F. 1983. Buffer gradient gels and ³⁵S-label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. U.S.A.* 80:3963-3965.
- Brown, N.L. 1984. DNA sequencing. *Methods Microbiol.* 17:259-13.
- Isfort, R. and Ihle, J. 1988. The 4-6-8 method of sequence analysis. *BioTechniques* 6:138-141.
- Mayeda, A. and Krainer, A.R. 1991. Long-term storage of concentrated Tris borate-EDTA electrophoresis buffers without precipitation. *BioTechniques* 10:182.
- Olsson, A., Moks, T. Muhlen, J., and Gaal, A.B. 1984. Uniformly spaced banding patterns in

DNA sequencing gels by the use of a field-strength gradient. *J. Biochem. Biophys. Methods* 10:83-90.

- Sheen, J.-Y. and Seed, B. 1988. Electrolyte gradient gels for DNA sequencing. *BioTechniques* 6:942-944.
- Slatko, B. 1991a. Sources of reagents and supplies for dideoxy DNA sequencing and other applications. *In Methods in Nucleic Acids Research* (J. Karam, L. Chao, and G. Warr, eds.) pp. 379-392. CRC Press, Boca Raton, Fla.
- Tabor, S. and Richardson, C.C. 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 84:4767-4771.
- U.S. Biochemical. 1990. Formamide gels (40%) for sequencing DNA. *Comments* 17(1):31.
- Williams S.A., Slatko, B.E., Moran, L.S., and DiSimone, S.M. 1986. Sequencing in the fast lane: A rapid protocol for [α -³⁵S]dATP dideoxy DNA sequencing. *BioTechniques* 4:138-147.

Key Reference

- Slatko, B. 1991b. Protocols for manual dideoxy DNA sequencing. *In Methods in Nucleic Acids Research* (J. Karam, L. Chao, and G. Warr, eds.) pp. 83-129. CRC Press, Boca Raton, Fla.

Contains references for numerous modifications to the basic sequencing gel protocol.

Contributed by Barton E. Slatko
New England Biolabs
Beverly, Massachusetts

Lisa M. Albright
Reading, Massachusetts

Computer Manipulation of DNA and Protein Sequences

The ability to determine DNA sequences is now commonplace in many molecular biology laboratories. As the amount of DNA sequence data available to researchers has increased, the use of computers to manipulate, compare, and analyze this data has grown accordingly. Growth rates of DNA databases have entered an exponential phase as determination of DNA sequence becomes more automated and less expensive (see Fig. 19.2.1). The cost of a computer system appropriate for an individual laboratory has also decreased, allowing essentially all researchers Internet access to electronic databases and computerized sequence analysis (see UNIT 19.1).

This unit outlines a variety of methods by which DNA sequences can be manipulated by computers. To begin analysis of a DNA sequence, the information must be in a form that computers can understand. Generally this requires that the DNA sequence information be contained in a file (a file on the computer is simply a collection of information in computer-readable form). Procedures for entering sequence data into the computer and assembling raw sequence data into a contiguous sequence are described first. This is followed by a description of methods of analyzing and manipulating sequences—e.g., verifying sequences, constructing restriction maps, designing oligonucleotides, identifying protein-coding regions, and predicting secondary structures. This unit also provides information on the large amount of software available for sequence analysis.

In addition to the information concerning computer manipulations of DNA sequences in this unit, related units can be found in Chapter 19; these units explain how to use the worldwide Internet computer network to access information from or submit information to the major sequence databases and other electronic resources, and how to conduct DNA and protein homology searches. The chapter also includes general information about electronic information handling.

Many of the electronic resources described in this unit and in Chapter 19 are available at no cost on the Internet. The Appendix to this unit lists some of the commercial software, shareware (whereby the author asks for a donation from satisfied users), and free software

(public domain software as well as copyrighted software that is distributed without charge by the author) related to DNA sequence manipulation. Additional software related to carrying out DNA and protein homology searches is described in Chapter 19.

The discussion of computer-mediated DNA sequence analysis in this unit is intended to be a generalized overview; specific instructions for using particular sequence analysis software packages are beyond the scope of this unit. However, recommendations of software packages for particular tasks have been included wherever possible. Our goal is to serve as a starting point for researchers interested in utilizing the tremendous sequencing resources available to the computer-knowledgeable molecular biology laboratory.

NOTE: The Appendix to this unit, located at the end of the unit but before the references, lists the addresses and phone numbers of all software and hardware vendors discussed below. Additional information about relevant journals and databases is also provided in the Appendix.

SEQUENCE DATA ENTRY

For small sequencing projects, DNA sequence data can easily be entered and manipulated using a word processor or text editor on a microcomputer such as an Apple Macintosh or IBM-compatible. As the size of a project increases, however, a specialized editor for data entry becomes more and more desirable. All software packages for DNA sequence analysis provide some form of sequence editor—a program that acts like a word processor for sequence data. Such an editor may be no more than a small window into which sequence data is entered or it may contain specialized features that aid in proofreading and facilitate input.

Currently there is no standard format for DNA or peptide sequence files. In the early days of sequence analysis, most software programs had their own unique formats for storing sequence information. Although these format differences still exist today (see Fig. 7.7.1), most available software can import several different sequence formats. Thus, a sequence file can be converted into the default format for a particular program by importing (reading) the original sequence file into the program and then

Contributed by J. Michael Cherry

Current Protocols in Molecular Biology (1995) 7.7.1-7.7.23
Copyright © 2000 by John Wiley & Sons, Inc.

```

A) EMBL:
ID hummycc.primate
DE hummycc.primate
SQ          100 BP
AGCTTGTTTGGCCGTTTTAGGGTTTGTGGAAATTTTTTTTCGTCTATGT
ACTTGTGAATTATTTACAGTTTGCCATTACCGGTTCTCCATAGGGTGATG
//

B) GenBank:
LOCUS          HUMMYCC      10996 bp ds-DNA      PRI      25-JUL-1994
DEFINITION    Human (Lawn) c-myc proto-oncogene, complete coding ...
ACCESSION    J00120 K01908 M23541 V00501 X00364
KEYWORDS     Alu repeat; c-myc proto-oncogene; myc oncogene; ...
SOURCE       Human DNA (genomic library of Lawn et al.), clones ...
  ORGANISM    Homo sapiens ...
REFERENCE    1 (bases 3507 to 7559)
AUTHORS      Colby, W.W., Chen, E.Y., Smith, D.H., and Levinson, A.D.
TITLE        Identification and nucleotide sequence of a human ...
JOURNAL      Nature 301 (5902), 722-725 (1983)
MEDLINE      83141777
COMMENT      The myc gene is the cellular homologue of the ...
FEATURES     Location/Qualifiers
BASE COUNT   2747 a 2723 c 2733 g 2793 t
ORIGIN       198 bp upstream of Sau96A site, on chromosome 8 (q24)
              1 agcttgtttg gccgttttag ggtttggttg aatTTTTTTT tcgtc ...
//

C) GCG:
Human (Lawn) c-myc proto-oncogene, complete coding sequence and flanks.
  hummycc.primate Length: 100 Sun, Mar 17, 1991 9:48 PM Check: 6864 ..
      1 AGCTTGTTG GCCGTTTTAG GGTTGTGG AATTTTTTTT TCGTCTATGT
      51 ACTTGTGAAT TATTTACAGT TTGCCATTAC CGGTTCTCCA TAGGGTGATG

D) Intelligenetics:
; hummycc.primate, 100 bases.
; Human (Lawn) c-myc proto-oncogene, complete coding sequence and flanks.
;
hummycc.primate
AGCTTGTTTGGCCGTTTTAGGGTTTGTGGAAATTTTTTTTCGTCTATGT
ACTTGTGAATTATTTACAGTTTGCCATTACCGGTTCTCCATAGGGTGATG1

E) NBRF:
>DL;hummycc.primate
hummycc.primate, 100 bases.
      1 AGCTTGTTG GCCGTTTTAG GGTTGTGG AATTTTTTTT TCGTCTATGT
      51 ACTTGTGAAT TATTTACAGT TTGCCATTAC CGGTTCTCCA TAGGGTGATG*

F) DNA Strider Text:
; DNA sequence hummycc.primate, 100 b.p.
;
AGCTTGTTTGGCCGTTTTAGGGTTTGTGGAAATTTTTTTTCGTCTATGT
ACTTGTGAATTATTTACAGTTTGCCATTACCGGTTCTCCATAGGGTGATG
//

G) FASTA:
>hummycc Human (Lawn) c-myc proto-oncogene
AGCTTGTTG GCCGTTTTAG GGTTGTGG AATTTTTTTT TCGTCTATGT
ACTTGTGAAT TATTTACAGT TTGCCATTAC CGGTTCTCCA TAGGGTGATG

```

Figure 7.7.1 Commonly used sequence file formats having specific defined elements and defining codes. **(A)** EMBL comment lines begin with two-letter codes: ID, short sequence name; DE, description; and SQ, sequence length. DNA or protein sequence follows; sequence end is denoted by two slashes (//) on a separate line. **(B)** GenBank comments precede the sequence and are separated from it by the code "ORIGIN." Sequence end is denoted by two slashes on a separate line. The actual text of this entry has been abbreviated; see Fig. 19.2.3 for a more complete example of a GenBank file. **(C)** GCG comments precede the sequence and are separated from it by two dots (.). **(D)** Intelligenetics comment lines begin with semicolons (;). A single description line follows, and then the sequence begins on a separate line. Sequence end is denoted by a numeral one (1). **(E)** NBRF (also called PIR format) first line starts with four required characters: a greater-than sign (>); either "D" for DNA or "P" for protein; either "L" for linear or a "C" for circular; and a semicolon. The short sequence name follows on the same line. The next line is a description line. Sequence starts on a new line and its end is denoted by an asterisk (*). **(F)** DNA Strider Text is similar to the Intelligenetics format, but lacks description line. **(G)** FASTA (also called Pearson format) first line begins with a greater-than sign (>), followed by sequence name and a short description. Sequence data starts on a separate line. *Note:* Some formats (including GenBank, GCG, and NBRF) allow numbers to be included within the sequence for ease of reading (the numbers are ignored during sequence analysis).

saving the file in the desired format. However, it is generally simplest to use the sequence editor or entry program provided with the software to be used for subsequent analysis.

Manual Entry Using Word Processors and Sequence Editors

Word processors. A DNA or peptide sequence can be entered into a computer using a word processor or text editor program, simply by creating a new document and then typing in the sequence data. This is sufficient to put sequence data in computer-readable format and is useful if no specialized sequence entry program is available or if the sequence file is to be transferred to a different computer for analysis.

It is imperative that DNA or peptide sequence documents be saved in “text” or ASCII format, because sequence analysis programs cannot translate the default files used by word processor programs; however, most packages provide an easy one-step method of reformatting text files containing DNA or peptide sequences into a format appropriate to the computer at hand. Saving documents as text may require a slightly different Save command or option, as specified in the word processor manual.

Sequence files often contain more information than just the DNA or peptide sequence data. Comments and reference information about the sequence and the history and dates of changes made to the sequence are commonly included. Such comments must be distinguishable by the program from the sequence data in order for the sequence analysis software to function properly. However, it is important to determine what sequence and comment-file formats are allowed by the particular analysis software being used. The most commonly used file formats are shown in Figure 7.7.1. Note that some—e.g., the Intelligenetics and DNA Strider text formats—use specific characters to identify the comments before the sequence file; other formats simply require a string of characters that is used to denote the end of the comments and the beginning of the sequence—e.g., the GenBank flatfile format uses a line beginning with “ORIGIN” and GCG uses “..” (two periods). The NBRF software package (also known as PIR) allows only a single line of comments. Some formats (e.g., NBRF and Intelligenetics) require a specific terminator character for the sequence entry.

It is a good idea to test the file format before investing a large amount of time typing in data with the word processor. This can be done by

creating a small DNA sequence, then experimenting with the comment and sequence formats as well as the word processor’s Save option. The resulting file can then be checked in the analysis program.

Sequence editors. All analysis programs include some form of sequence editor. This is the most effective tool for entering DNA or peptide sequence data because it produces a sequence file in the correct format with the comments separate from the sequence as they should be. There are several free or relatively inexpensive programs available for Macintosh and IBM-compatible computers that simplify the task of entering sequence data. One example for the Macintosh is DNA Strider, which is built around a very easy-to-use sequence editor. It also makes restriction maps, finds open reading frames, and translates DNA sequences into amino acid sequences.

Recently, several programs have become available for microcomputers that “speak” the sequence as it is entered. This may seem like a frivolous feature; however, it is quite useful for entering data by hand from an autoradiogram. The audio feature permits verification of the sequence, thus reducing the need to continually look back and forth between the film and the computer screen. Some of these talking sequence editors use a digitized human voice, but others use synthetically generated speech. The digitized voice, which is generally easier to understand, is more common in programs for Macintosh computers—e.g., SeqSpeak, designed for Macintosh equipped with HyperCard version 2.0 and distributed free of charge (see Table 7.7.2).

Semiautomated Entry Using Digitizing Hardware and Software

Several sequence analysis software packages can accommodate the connection of a relatively inexpensive digitizing pad to the computer; a digitizing pad permits DNA sequences to be read directly from an autoradiogram. Digitizing pads increase both the speed and the accuracy of input. Digitizing devices designed for DNA sequence entry consist of a light box for viewing the autoradiogram and a stylus that resembles a ball-point pen with an attached wire. When the stylus is pressed at the location of a lane or band on the autoradiogram, the digitizing pad detects a signal and converts it to an *x-y* coordinate value that is recorded by the computer. Most digitizing pad entry software requires the basic shape (boundaries) of the autoradiogram lanes to be

entered initially. Subsequently, when the stylus is pressed on an individual band in a lane, the computer utilizes this boundary information to determine which lane was selected, then displays the appropriate nucleotide letter on the computer screen.

The result of the digitized entry process is a file containing the newly determined DNA sequence. In some cases the file must be reformatted before it can be used with an analysis program—this can be determined by comparing the format defined by the entry software with the list of acceptable formats in the sequence analysis software.

DNA sequence entry using a digitizing program can be very quick and accurate when the gel is of good quality. However, if the lanes are irregular due to curving or excessive smiling, the program may incorrectly identify which lane contains a selected band. Commercial gel-reading programs are designed to handle some types of common gel problems. For example, programs avoid smiling errors by requiring the order of the bands to be determined by the user so that the program determines only which lane was selected. The SEQED program of the GCG package for UNIX and VMS and the DNASTar package for IBM-compatibles and Macintosh can both be used with a digitizing pad for semiautomatic DNA sequence determination.

Automated Entry Using Gel Readers and Automated Sequencers

Automated gel readers. Several automated autoradiogram readers are now available that include a scanner and a computer. Automated readers use a high-resolution gray-scale scanner or a digital video camera to digitize an autoradiogram of a sequencing gel produced by conventional methods. The digitized image of the film is then analyzed by the computer using image-analysis techniques. The software is designed to identify first the locations of lanes and then bands within the lanes. This process can be confounded by smudges and random spots on the original film, as well as by gel smiling and curved lanes. Indeed, a major challenge to designers of automatic gel-reading programs is the elimination of such noise in scanned autoradiograms. Automated gel-reader systems (sold by BRL and Milligen, among others) are generally priced for use by large laboratories and core facilities.

When using automated image analysis, it may be advantageous to employ the multiplex method of DNA sequencing (introduction to Chapter 7 and UNIT 15.2; Church and Kieffer-Hig-

gins, 1988), because it is then possible to incorporate a known sequence of DNA into each gel; the resulting sequence can then be analyzed to determine the particular characteristics of the gel. Having this sort of an internal standard greatly simplifies the image analysis required to read more than one unknown sequence from the same gel using the automatic scanner.

Automated DNA sequencers. Automated DNA sequencing machines (discussed in UNIT 7.0) determine the sequence of a DNA fragment and then place the DNA sequence into a file. The most common automated sequencers use special fluorescence-tagged oligonucleotide primers in a dideoxy primer-extension reaction. Reaction products are separated by electrophoresis through a polyacrylamide gel and the fluorescent tag is detected after excitation by a laser at the bottom of the gel. Because each specific nucleotide reaction uses a different fluorescent tag, all four reactions can be loaded in the same lane. The output from the sequencer is a trace of the amount of fluorescence observed as the tagged reaction products pass off the gel. An attached computer analyzes the trace and determines the DNA sequence it represents. Thus, entry of sequence data is completely automated; the user merely has to define or identify the particular sequencing run.

Automated sequencers are quick and currently provide about the same level of accuracy as manual sequencing. Automated and manual sequencing employ similar sequencing technologies—e.g., primer extension and polyacrylamide gel electrophoresis. Although automated sequencing machines are expensive, for large sequencing projects, the automated sequencers may be more economical than manual sequencing because of the reduction in labor costs.

Editing automated sequence data. Because automated autoradiogram scanners are not absolutely accurate, the software packages that accompany them typically permit the user to edit the automatically entered sequence. Generally, the editing process is greatly simplified for projects where overlapping randomly chosen segments or overlapping nested sets of deletions are sequenced because each nucleotide is determined multiple times. A sequence-assembly program can then be used to align the sequences based on overlapping regions. For DNA sequences that have been determined multiple times, inconsistencies between different versions are readily identified by the software, allowing the user to go back to the original data (digitized image or autoradiogram) to assess why the inconsistency occurred.

Automatic sequencing machines (such as those produced by Pharmacia Biotech, Li-Cor, and Applied Biosystems) provide a similar editing feature whereby the user can view the fluorescent trace on a computer screen and override the computer's choice of base if it appears to be incorrect. Some vendors of automated sequencing machines use a statistical means of correcting for discrepancies between regions that have been sequenced multiple times. That is, if a particular region is sequenced several times and the number of differences is low, the machine will automatically adopt the majority consensus sequence.

SEQUENCE DATA VERIFICATION

As with any experimental result, it is important to verify DNA sequence data. This is typically accomplished by sequencing a given region more than once and by sequencing both strands. The computer can be used to compare sequences, find overlapping regions, highlight differences, and provide a graphic overview of the overlapping regions.

Comparison of Multiple Entries

For small projects, the amount and quality of data can be optimized. To identify errors introduced during manual or automated data entry, each gel can be read more than once and the independent readings compared. The comparison can be carried out automatically using an alignment program (see Homology Searching) and any differences can then be investigated. Similarly, several of the digitizing gel readers provide a confirmation option in which each segment of a sequence can be checked automatically, with the machine simply re-entering the band locations. These programs, such as SEQED (included in the GCG package; see Table 7.7.1), alert the user as differences occur. In this way the entry process can be quite fast without sacrificing accuracy.

For large sequencing projects that involve sequencing a number of random clones from a library covering the region of interest, it is usually possible to achieve a sufficient level of redundancy that a majority consensus sequence can be determined.

Comparison to Known Restriction Maps

If a restriction map for the DNA region of interest already exists, a restriction map generated from the DNA sequence (as described below) can serve as a check on the accuracy of the sequence. This is particularly useful for

large regions that have been subcloned for sequencing, where it may not be completely clear how the shorter pieces of DNA that were actually sequenced fit together within the larger region.

Test Translations to Detect Shifts in Reading Frame

One of the most widespread features of sequencing programs is the capacity to translate a DNA sequence into amino acids, thereby generating a putative protein sequence from an open reading frame (ORF); many restriction mapping programs also include a translation feature. When the DNA sequence in question is a prokaryotic coding region or the sequence from a cDNA clone, or when it is believed to be related to a gene from another organism, translating the sequence in this fashion provides a simple check for deletions and insertions. The process will find any stop codons within the sequence; if a stop codon is found within an area that is actually known or thought to be a coding region, the sequence is automatically suspect and should be rechecked. (It is important to remember that in some DNA—such as mitochondrial DNA—the stop codons may be used to specify an amino acid. In yeast mitochondrial DNA, for example, the codon TGA specifies isoleucine. Some software is able to take into account variant genetic codes; for instance, DNA Strider allows the user to select from a list of variant codes, and the GCG package allows the user to reset individual codons.) Although this technique will detect only a subset of possible errors, it can often quickly identify common problems such as simple typing mistakes or miscounting of the number of the same nucleotides in a run of identical nucleotides.

Most DNA sequence packages allow the user to specify a range and reading frame to be used in translating a DNA sequence. DNA Strider (Table 7.7.1) has a useful feature that hunts for ORFs in a DNA sequence and highlights the putative coding sequence, thus doing all the necessary work.

Detecting Overlap with Other Sequenced Fragments

With increasing worldwide interest in genome sequencing projects, sequence assembly packages now provide very effective automatic DNA sequence assembly, connecting shorter pieces of DNA to build the longest continuous sequence possible. Once the sequences to be assembled are identified, they are

compared, overlaps identified, and contiguous sequences (contigs) constructed. Typically, parameters are available to allow adjustment of the alignment process. If the default parameter settings do not result in sequence assembly, the minimum amount of overlap required to identify a match (i.e., the number of base pairs involved) can be reduced or the stringency of the required overlap decreased (i.e., the number of allowable mismatches increased). A good commercial assembly program is LaserGene, available from DNASTar for both Macintosh and IBM-compatible computers (Table 7.7.1). Once a sequence contig has been assembled, the LaserGene program creates a graphic overview of the sequencing project, highlighting regions that need further verification. Another commercial program is Sequencher (Gene Codes). This program provides expanded features for the assembly, processing, and editing of DNA sequences determined with the ABI sequencer; however, it is only available for the Macintosh computer.

If a sophisticated assembly program is not available, contigs can be constructed “by hand” using a comparison program and a multiple sequence editor. First, the comparison program is used to analyze two suspected overlapping sequences; then, if an overlap is found, a consensus sequence is generated that can then be compared with other sequences. In this manner, an overall consensus sequence (contig) can be progressively assembled. A potential pitfall of this process that should be mentioned is that although some DNA sequence analysis programs will consider both the DNA sequence that has actually been entered and its complement sequence (the sequence from the complementary strand of DNA, sometimes called the reverse complement sequence), others analyze only the sequence entered. With these latter programs, unless the DNA sequence of interest contains an inverted repeat (palindromic region), only one of the strands of DNA is likely to be recognizable as being similar to a previously known sequence. It is therefore important to determine which way a particular program operates and, when using a program that does not automatically consider the complement sequence, to conduct a separate search of that sequence.

When conducting sequence comparisons to identify overlaps, it is important to take into account the fact that different programs take slightly different approaches to this process. Some comparison programs that use the Needleman and Wunsch algorithm (e.g., the GCG

GAP program; Needleman and Wunsch, 1970) are designed to find the maximum number of matches between two sequences over the entire length of the two sequences, with the minimum number of gaps. This type of program is best suited for aligning two sequences along their entire lengths. If the two sequences differ greatly in size, however, the result may not satisfactorily represent the similarity between them. For identifying regions of similarity between two pieces of DNA that are largely different in sequence or of significantly different lengths, programs employing algorithms such as those described by Smith and Waterman (1981) or Wilbur and Lipman (1983)—e.g., the GCG BESTFIT program—are more suitable. These algorithms do not try to align the two sequences being compared in their entirety, but instead search for short matches within the sequences.

Editing a Contig and Verifying the Sequence

Generally, software packages that provide sequence assembly programs include multiple sequence editors that display the individual sequences of the aligned contig together one on top of the other, one sequence per line (see Fig. 7.7.2), and can generate the consensus sequence automatically. A multiple sequence editor differs from the sequence editors mentioned earlier in that several sequences can be manipulated at once. However, the most significant feature of a multiple sequence editor is its ability to produce a consensus sequence for an aligned set of sequences. If the consensus sequence is not satisfactory, the alignment can instantly be changed and a new consensus sequence generated.

Another use of a multiple sequence editor is for manually comparing a group of overlapping sequences by aligning common regions vertically on the screen (Fig. 7.7.2), which makes it easy to identify differences in the aligned sequences. It is generally useful to go back to the original gel to determine whether these differences result from misreading, sequence compression, or gel defect. Multiple sequence editors will often have the ability to display the reverse complement of a particular sequence, irrespective of which strand was sequenced.

Some very useful programs are available free of charge; these programs provide automatic multiple sequence alignment functions as well as other types of analysis on a set of sequences. Examples of these are the MACAW software for Microsoft Windows or Macintosh,

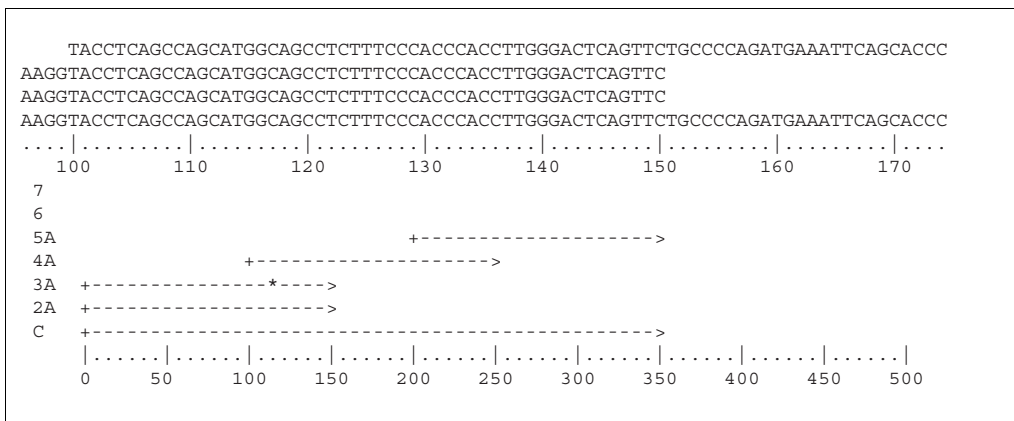


Figure 7.7.2 Multiple sequence editor. The GCG program GELASSEMBLE displays the aligned sequences on the top of the screen and a schematic of the sequenced fragments on the bottom. Arrows indicate the direction of sequencing; the asterisk in the lower part of the display indicates the position of the cursor in the sequence alignment as the user edits the sequence.

developed by Greg Schuler of the NCBI (Schuler et al., 1991), and GDE for X-Windows, developed by Steven Smith of the Harvard Genome Laboratory (Table 7.7.2).

RESTRICTION MAPPING

Once a sequence contig is generated, a map of restriction endonuclease cleavage sites can be a useful aid in further analysis of the region. As stated above, construction of a restriction map based on newly determined sequence information allows rapid visual comparison of the sequence with a known restriction map. Restriction maps also identify sites for subcloning or other molecular genetic manipulations and can provide useful summaries of newly determined DNA sequences.

Mapping All Known Commercial Restriction Enzyme Recognition Sites

Programs are available that identify on a strand of DNA the sites of all known restriction enzymes, according to either the first nucleotide of the recognition site or the location of cleavage. Most of these programs produce one list of the sizes of fragments that would be created by cutting with each restriction enzyme and a separate list of the recognition sites identified. They can also display the cleavage sites of specified sets of restriction enzyme of interest, such as those that generate 3'- or 5'-strand overhangs or blunt ends. All restriction mapping programs contain a file or files in which restriction site data are stored; this information can be updated as new restriction enzymes and their cleavage sites are identified. Restriction mapping programs are one of the most common types of molecular biology software; virtually

all commercial software provides excellent restriction mapping features.

An up-to-date list of restriction enzymes is maintained by Richard Roberts (New England Biolabs). It is accessible in the form of a text file called rebase and contains all known type II restriction enzyme recognition sites, sites of cutting, a complete cross-reference of isoschizomers, reference citations, and commercial sources for the restriction enzyme with addresses and phone numbers. The latest version of rebase can be obtained free from a variety of electronic mail (e-mail) and network servers (see UNIT 19.1).

Mapping to Predict Band Sizes

Some software programs can model double digests, predicting the band sizes that would result from cleavage of an entered sequence with a given pair of restriction enzymes. Such double-digest band-size analyses simplify the reading of restriction fragment patterns on gels and can also be useful for planning subsequent cloning strategies. A few programs can also predict the results of partial digests. Software packages that provide restriction fragment pattern analysis features are available from Textco's Gene Construction Kit (for Macintosh only) and Pro-RFLP from DNA ProScan (for Macintosh and IBM-compatible computers).

Graphical Restriction Mapping

Several very elegant programs are available that produce graphical restriction-site maps that are useful for searching for possible sites to be used in further recombinant DNA manipulations. Graphical restriction maps are often represented as collections of horizontal lines,

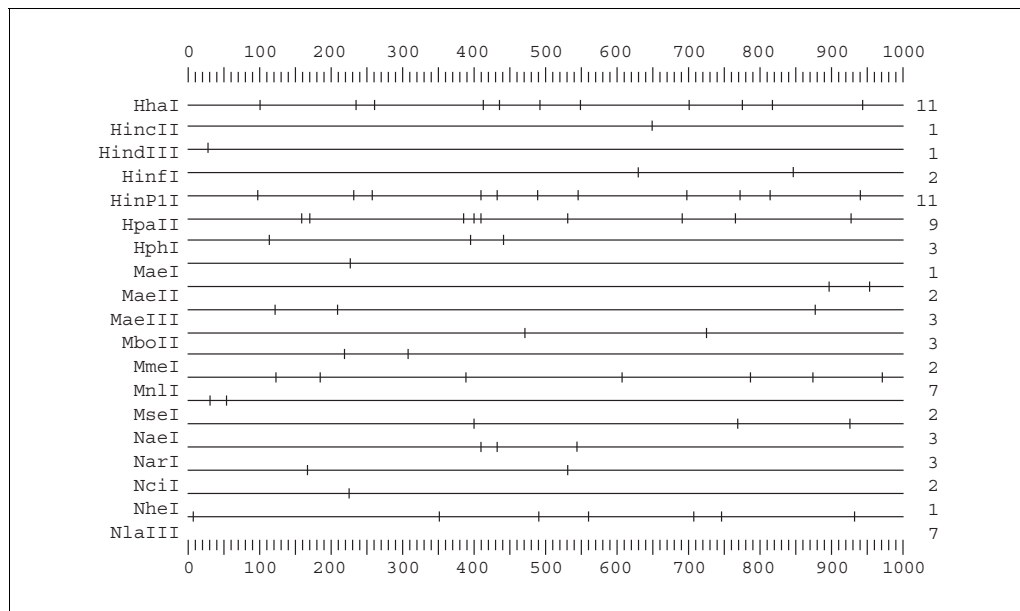


Figure 7.7.3 One type of graphical restriction map. This figure was produced by the free PlotZ program; GCG MAPPLOT produces similar output.

one for each restriction enzyme, with the recognition sites represented by short vertical marks at the appropriate locations (Fig. 7.7.3). A limited number of programs produce pictures of sequences that look like standard circular plasmid maps, including coding regions and other features of interest within the sequence, as well as the vector DNA and polylinker. These graphical maps can be saved and then manipulated in a graphics editor or drawing program, greatly simplifying the task of preparing figures for presentation; an excellent example of such a program is the Textco Gene Construction Kit. This program reads a variety of sequence-file formats and generates a picture of the DNA sequence in standard restriction map form. The pictures can be edited to include arrows and information boxes that are of publishable quality. In contrast to most graphical mapping programs, the Gene Construction Kit and other commercially available programs such as MacVector and LaserGene use DNA sequence data directly to generate the pictorial maps. Many shareware mapping programs, such as Jingdong Liu's MacPlasMap, do not work with the sequence data directly; rather, recognition site locations must be input by the user. One shareware program that does have the capacity to create graphical maps directly from sequence data, and provides many analysis features as well, is DNA Strider, developed by Christian Marck (Table 7.7.2).

Recent versions of many mapping programs can actually simulate the electrophoretic-gel

banding patterns that will be observed following particular restriction digests. The programs allow the user to specify the type of gel medium being used, making it possible to determine visually whether the banding pattern observed on a gel could be produced by a given sequence. Some programs will even simulate partial digests. The Gene Construction Kit provides the gel simulation feature as does the ACEDB genome database software (see Genetic Sequence Databases); the latter is limited in that it can simulate the gel banding patterns only for sequences already contained within the database being used.

PREDICTION OF NUCLEIC ACID STRUCTURE

Once the DNA sequence of a region has been identified, a number of analyses can be performed to identify interesting features such as repeats, areas of atypical base composition, and RNA secondary structure. These in turn can help to define functional regions within the sequence of interest.

Base Repeats

Direct and inverted repeats are often part of transcriptional or translational control regions. Most sequence analysis software can identify repeats and provide an optional graphic display of their location. The standard "dot matrix" plot is a simple and effective method of identifying repeated regions. The two-dimensional dot matrix represents one sequence on the *x* axis and

the other on the *y* axis. Such programs generate a “dot” at each *x-y* intersection of all GG, AA, TT, and CC pairs; thus, repeats are revealed as diagonal lines. Some dot matrix programs show inverted repeats—which may indicate potential stem-loop structures in the corresponding DNA or RNA—as lines with negative slopes relative to direct repeats. To identify repeats within a single sequence, the same sequence is used for both the *x* and *y* axes.

Analysis packages that do not produce a graphical presentation of the repeated regions usually include a program that lists the repeats and their locations. This involves searching a sequence against itself to find direct repeats or against its complement to detect inverted repeats. The repeats must then be plotted by hand or the distances between them calculated to determine if a periodicity is present. Needless to say, the graphical presentation is preferred because patterns can be readily seen that it would take some time and effort to reveal by the comparison method.

GC and AT Content

The GC/AT content of a sequence may provide some insight into structural features such as Z DNA and bent DNA (see APPENDIX 1). GC/AT content can also serve as a reasonable indicator of a coding region in many invertebrate and plant species. A few sequence-analysis packages contain specialized programs, such as GCG STATPLOT, that show the GC and AT content of a sequence in graphical form—allowing AT- or GC-rich regions to be readily identified. However, many packages provide little more than a tally of nucleotide composition, i.e., the program only lists the number of A, G, C, and T residues in a sequence. A simple method to determine the local GC/AT content of a large sequence is to divide the sequence into several small (e.g., 100-base) segments, then determine the GC/AT content of these small segments individually.

RNA Secondary Structure

Although folding programs are available that predict RNA secondary structure, this type of analysis is still an art. RNA-folding programs help identify possible stable stems in an RNA molecule, but a trial-and-error process is required to determine the biological significance of these results for a given RNA molecule. Even with this limitation, secondary structure predictions can be useful for identifying mRNA control regions as well as possible stable folded regions of an RNA molecule.

The greatest problem with predicting secondary structure is modeling the interactions present in a tertiary structure and then relating those back to the primary sequence for use in a folding program. Indeed, current RNA-folding programs do not take into account possible tertiary structures of a nucleic acid molecule. These programs determine the energetics of a limited number of two-dimensional folded structures. The most stable structure predicted by the program in a two-dimensional world may be far from the most stable structure in three dimensions where loops can interact with loops, helical regions can stack, and various non-Watson-Crick base-pairing structures can occur (see illustrations in APPENDIX 1).

Currently, the most sophisticated RNA-folding program is MFOLD (for Multifold, an extension of an earlier program known as RNA-Fold or FOLD in the GCG package), designed by Michael Zuker of the National Research Council of Canada (Zuker, 1989a,b; Table 7.7.2). In addition to standard analysis of base-pairing energetics, MFOLD takes into account base-pair-stacking energies (Freier et al., 1986; Turner et al., 1987) and single-base-stacking (dangle) enthalpies (Turner et al., 1988). Another major feature of MFOLD is that in addition to predicting a single structure with the lowest predicted free energy, it also depicts many suboptimal structures. The number of suboptimal structures displayed can be varied within a percentage or an absolute value of the best structure's energy. VMS, UNIX, DOS, and Macintosh versions of this program are available from many of the software archives (see Table 7.7.2). Although the output of MFOLD is text-based (Fig. 7.7.4A), several programs are available that generate graphic representations of the predicted structure (e.g., LoopViewer, developed by Don Gilbert; see Fig. 7.7.4B).

OLIGONUCLEOTIDE DESIGN STRATEGY

Increased use of polymerase chain reaction (PCR) methods has stimulated the development of many programs to aid in the design or selection of oligonucleotides used as primers for PCR. Four such programs that are freely available via the Internet (see Table 7.7.2) are: PRIMER by Mark Daly and Steve Lincoln of the Whitehead Institute (UNIX, VMS, DOS, and Macintosh), Oligonucleotide Selection Program (OSP) by Phil Green and LaDeana Hiller of Washington University in St. Louis (UNIX, VMS, DOS, and Macintosh), PGEN by Yoshi (DOS only), and Amplify by Bill Engels

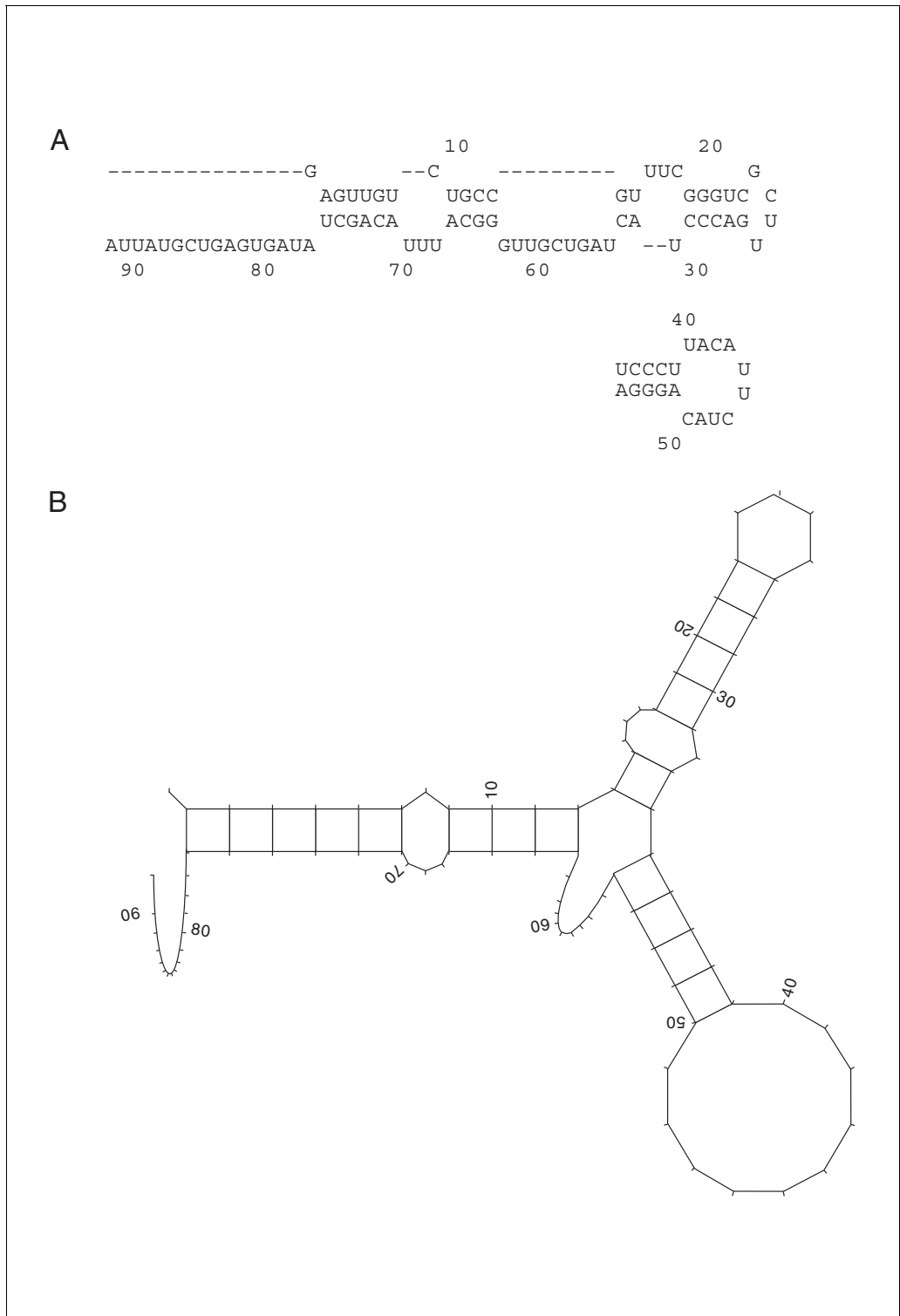


Figure 7.7.4 (A) Text-based output from Zuker's RNA-folding program, available from GCG under the name of FOLD. This type of representation is difficult to visualize, but acceptable when only a quick view of the possible folded structures is desired. (B) Graphic representation of the structure shown in part A, produced by the GCG Squiggles program. The free LoopViewer program (for Macintosh) produces similar representations.

of the University of Wisconsin (Macintosh only). Generally these programs help in the design of PCR primers by searching for bits of known repeated-sequence elements and then optimizing the T_m by analyzing the length and GC content of a putative primer. Commercial software is also available and primer selection procedures are rapidly being included in most general sequence analysis packages.

Sequencing and PCR Primers

Designing oligonucleotides for use as either sequencing or PCR primers requires selection of an appropriate sequence that specifically recognizes the target, and then testing the sequence to eliminate the possibility that the oligonucleotide will have a stable secondary structure. Inverted repeats in the sequence can be identified using a repeat-identification or RNA-folding program such as those described above (see Prediction of Nucleic Acid Structure). If a possible stem structure is observed, the sequence of the primer can be shifted a few nucleotides in either direction to minimize the predicted secondary structure. The sequence of the oligonucleotide should also be compared with the sequences of both strands of the appropriate vector and insert DNA. Obviously, a sequencing primer should only have a single match to the target DNA. It is also advisable to exclude primers that have only a single mismatch with an undesired target DNA sequence. For PCR primers used to amplify genomic DNA, the primer sequence should be compared to the sequences in the GenBank database to determine if any significant matches occur. If the oligonucleotide sequence is present in any known DNA sequence or, more importantly, in any known repetitive elements, the primer sequence should be changed.

Degenerate Probes for Detecting Related Genes

Once a conserved protein sequence has been identified, a degenerate oligonucleotide can be designed for use as a hybridization probe to screen a library to identify additional members of the protein family (see UNIT 6.4). To design this oligonucleotide, the conserved protein sequence must be translated into a degenerate DNA sequence. Most software packages provide this feature; their output is a DNA sequence that is produced using the IUPAC degenerate-nucleotide codes. A degenerate oligonucleotide is then synthesized to correspond to the back-translated protein sequence. Most DNA synthesizers will create an oligonu-

cleotide with more than one nucleotide at any position within the sequence except for the 3' nucleotide.

For efficiency of synthesis and hybridization, the following guidelines, designed to yield oligonucleotides with the lowest possible level of degeneracy, should be followed. First, it is not necessary to incorporate both G and A to match a consensus sequence position containing C and T, because G pairs with both C and T. Second, inosine (I) pairs with G, C, and A. Third, a pyrimidine-pyrimidine mismatch does not disrupt base pairing, but a purine-purine mismatch is destabilizing.

As a general rule, create the minimum sequence that hybridizes to the consensus sequence. For each species (unique sequence) of oligonucleotide in the synthesis, the concentration of oligo used in hybridizations must be increased to achieve an equivalent C_{0t} value.

IDENTIFICATION OF PROTEIN-CODING REGIONS

Identification of potential protein-coding regions, especially in genomic DNA sequences from higher eukaryotes, is still not a completely automated process. It is helpful to simply translate the region in all six reading frames (three on each strand) and then identify all possible exon regions as uninterrupted open reading frames (ORFs). Although identifying the AUG initiation codon is a simple task, determining the location of introns is not as straightforward (unless the sequence is from yeast, where the rules for splicing are simple and seemingly absolute). Although several research groups are working on techniques to identify exon-intron boundaries, the process requires thoughtful consideration of several types of analyses. One useful technique available with some software packages (e.g., GCG TestCode) uses a purely statistical method to determine the nonrandomness of the triplet code characteristic of an ORF (Fickett, 1982). This method works best with large windows (~200 nucleotides), although it may be of limited use for identifying small ORFs. The basic principle is that introns are evolving without any restraint and are thus more random in sequence than exons, which are subject to stabilizing selection.

Prediction of ORFs is a very active research area. The rules used by the cellular machinery to define splice sites for higher eukaryotic sequences are still illusive. Four projects are trying to develop truly automatic gene-identification methods. Although there is some disagreement about the current exact reliability of these

methods, they all claim to locate about 90% of known (i.e., previously well-characterized) exons, and they are continuing to be improved. The first, Gene Recognition and Analysis Internet Link (GRAIL), is available via an electronic mail (e-mail) server and as a UNIX application; it is being developed by a group led by Edward Uberbacher of the Oak Ridge National Laboratory. GRAIL utilizes an artificial intelligence technique called a neural network that learns by example. The GRAIL software was not programmed to recognize a specific set of characteristics of human coding regions. Rather, it is given a set of well-characterized human sequences for which the locations of exons and introns have been experimentally identified. The neural network is programmed to search for particular types of simple features and then to correlate these features with the input set's exon-intron boundaries.

In contrast to GRAIL, the second program, GeneID, developed by Steen Knudsen and Kathleen Klose of Temple Smith's group at Boston University, utilizes many features of coding regions, including exon/intron consensus sequences and codon preferences. Because its rules are generally based on human sequences, GeneID's usefulness with nonmammalian sequences may be limited. GeneID is available from BMERC at Boston University via an e-mail server (see Appendix and *UNIT 19.1*).

The third gene-finding resource, BCM Gene Finder, is provided by the Baylor College of Medicine via e-mail and the World Wide Web (see Appendix and *UNIT 19.1*). A set of analysis programs is available to aid in the identification of genes in human DNA sequences. The analysis method involves predicting all possible internal exons based on the combination of characteristics describing a potential splice site. Then the set of potential exons is analyzed to determine the optimal combination and a model for the putative gene is constructed.

The final service was announced in 1992 and uses a neural network program to identify coding regions. The NetGene service is available through an e-mail server and is provided by the Department of Physical Chemistry at the Technical University of Denmark. This server appears to be changing the slowest, but because no one has yet produced the definitive gene-finding software, it is suggested that all of these servers should be tried to determine which resource is the most useful.

HOMOLOGY SEARCHING

Searching for homology between a newly obtained sequence and a sequence already listed in one of the DNA or protein databases can be very informative. Similarity to a known sequence can suggest the function of the new protein or indicate that no similar sequence has yet been deposited in the database. Because of the size of the databases and the speed with which they are expanding (see Fig. 19.2.1), the task of searching the database is not always easy to accomplish using an isolated laboratory microcomputer. Searching sequence databases for similarities is one of the few sequence-analysis tasks that is still best performed on a larger computer system. To search a single sequence against the entire GenBank, European Molecular Biology Laboratory (EMBL), or DNA Data Bank of Japan (DDBJ) databases requires about an hour on a smaller VAX or microcomputer. For this reason, many laboratories obtain an account on a large computer system or modern workstation that provides access to the large genetic sequence databases (see *UNIT 19.1*). Several sources, including the European Bioinformatics Institute (EBI), the National Center for Biotechnology Information (NCBI), and the University of Houston, provide free databases and homology searches over the Internet to anyone with access to e-mail (see *UNIT 19.1*). Protocols for carrying out these homology searches are described in *UNIT 19.3*.

Comparison of Two Sequences

Many programs are able to align two DNA or protein sequences. Such programs are often used to format an alignment for publication or simply to identify regions of similarity between two input sequences. In addition, these programs also introduce gaps into a sequence to optimize the alignment. Because alignment programs assign numerical penalties to gaps and mismatches, the alignment can be influenced by varying gap and mismatch parameters.

Protein sequences are aligned using a scoring matrix developed by Margaret Dayhoff (Schwartz and Dayhoff, 1978) known as PAM250, which represents the evolutionary change that takes place in a protein sequence over time. A PAM (rearranged acronym for Accepted Point Mutations) is a measure of the number of individual amino acid changes occurring per 100 amino acid residues as a result of evolution. A PAM of 250, therefore, represents 250 mutations occurring within 100 resi-

dues. For two proteins to be separated by 250 PAM requires some amino acids in the sequence to have mutated multiple times. However, as calculated by Dayhoff, two such protein sequences would still have ~20% of their amino acids in common, assuming both a completely random distribution of mutations and lack of selection for conserved functional sequences. The so-called PAM250 log odds matrix, the log of the probability that a given amino acid could mutate in the evolutionary time equal to 250 PAM, was created by analyzing many families of protein sequences that were available in the late 1970s. The PAM250 matrix is used to determine the score of an aligned pair of sequences by summing the matrix values corresponding to each aligned pair of amino acids. These matrices are still commonly used today; however, the PAM matrix table has been recalculated (Gonnet et al., 1992). The Gonnet tables are the first recalculation of the PAM tables since their initial formulation. Subsequently a new amino acid substitution matrix was derived directly from sequence or three-dimensional structural alignments of distantly related proteins. This matrix, called BLOSUM62, is reported to perform better than the previous matrix at detecting distant relationships using either BLAST or FASTA (Henikoff and Henikoff, 1993). The BLAST servers at the NCBI use the BLOSUM62 matrix by default for protein sequence comparisons.

An example of the way information contained within a substitution matrix can be expressed is given here. The standard PAM250 matrix states that a methionine opposite an methionine in an aligned pair of sequences has a score of 6, a methionine opposite a valine has a score of 2, and a methionine opposite a cysteine has a score of -5. These scores reflect the fact that in the data used to generate the matrix, methionine was observed to change frequently to valine (thus, this is considered to be a conservative substitution), whereas methionine very rarely changed to cysteine. Thus, some amino acid changes have positive effects on the alignment score, and others have negative effects. The larger the number, the better the match, and thus a negative value between a pair of amino acids in the matrix means that that combination actually takes away from the alignment.

Comparison of Multiple Sequences

Programs for simultaneously aligning multiple sequences require more computer power and are less common than the alignment pro-

grams discussed above. Three standard multiple-alignment algorithms have been developed for UNIX and VMS-based systems: Des Higgins' Clustal (Higgins and Sharp, 1988, 1989; see Table 7.7.2), the Feng and Doolittle algorithm implemented in the GCG PILEUP program (Feng and Doolittle, 1987; see Table 7.7.1), and PIRAlign from NBRF, which is based on a variation of the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970; see Table 7.7.1). A powerful multiple-alignment tool for Microsoft Windows and Macintosh is the MACAW program created by Greg Schuler (Schuler et al., 1991). These programs identify common regions (or segments) in all the sequences that have been input, and then use the common regions as a starting point for building an overall alignment. They generally work best if the extent of the sequences being aligned is limited to the regions that are conserved among them.

Database Searches

Most researchers currently carry out both DNA and protein database searches over the Internet using the program Basic Local Alignment Search Tool (BLAST), developed by NCBI (Altschul et al., 1990). Detailed protocols for carrying out BLAST searches are given in an upcoming supplement. Here, we simply present a brief overview of the most important features of the BLAST family of programs.

The BLAST family of programs allows rapid similarity searching of nucleic acid or protein databases. The basic BLAST algorithm is used in several different programs that are each specific for a particular database and a particular type of input sequence. BLASTN is used to search a nucleic acid sequence against a nucleic acid database, BLASTP is used to search an amino acid sequence against a protein database, and TBLASTN is used to search an amino acid sequence against a nucleic acid database. In the latter program, the database is translated in all six reading frames prior to the search. The converse analysis is performed by BLASTX, which takes a nucleic acid input sequence and translates it in all six reading frames before searching against a protein-sequence database. Finally, if BLASTP does not identify significant sequence similarities, a more extensive program, BLAST3, is available that also compares an amino acid input sequence against a protein database. Like BLASTP, BLAST3 identifies regions of similarity between the input sequence and sequences present in the database, but the initial search is at lower stringency. This search

produces a collection of pairwise matching sequences that are then compared to each other, resulting in three-way matches where the component two-way matches were not significant. BLAST3 can be useful in identifying divergent members of a common gene family.

If the sequence of interest contains a protein-coding region, it is more informative to search the predicted protein sequence against one of the protein databases than to search the DNA sequence against a nucleic acid database. Because protein sequences evolve at a slower rate than DNA sequences, a distant homology between protein sequences may be missed at the DNA level. If no coding region has been defined, BLASTX can be used to translate the DNA sequences in all six reading frames before searching against a protein database. Because the protein-sequence databases only contain identified proteins, it is also important to check a newly defined amino acid sequence, or translated DNA sequence, against the current GenBank, EMBL, or DDBJ DNA sequence databases using TBLASTN. Such a search may identify a significant similarity to a DNA sequence that was not previously known to encode a protein.

An important feature of BLAST is a statistical significance score of the reported match. This statistical significance score is determined using an implementation of Karlin's significance formula (Karlin and Altschul, 1990), which calculates the Poisson probability that the observed sequence similarity will occur by chance based on the size and composition of the sequence database as well as on the size and quality of the match.

Another frequently used program for searching both protein and DNA sequence databases is FASTA (Pearson and Lipman, 1988), an updated version of the FASTN and FASTP programs. FASTA, which has been included in many commercial analysis packages, uses an initial fast search through the database to identify sequences with a high degree of identity to the test sequence. This fast search is performed by limiting the search to short regions of identity between the test sequence and the database. The word size (or k-tuple) parameter used in the program can be varied; it is the size of the initial match (sequence segment) that is used. The size of the k-tuple parameter indirectly affects the speed and sensitivity of the initial search through the database. Generally, the default value is the most efficient value to use; however, if distant homologies are desired, the

k-tuple value can be decreased (decreasing the k-tuple value causes the search to be more sensitive and take much more computer time).

FASTA builds a list or dictionary of all possible sequences of the size specified by the k-tuple value. The test sequence and all sequences in the database are then processed to find the locations of all segments in the sequence of a length equal to the k-tuple value that are present in the dictionary. For example, if the k-tuple was set to four, the sequence AGTCCTG would only have four entries (AGTC, GTCC, TCCT, and CCTG) in the dictionary of 256 (4^4) different four- (k-tuple value) letter words. The dictionaries of the two sequences can be more quickly compared than the sequences themselves, allowing for efficient identification of regions that contain small similarities. Once a list of the highest-scoring sequences is produced using the initial fast search, a second comparison is performed on just the top-scoring sequences. This secondary alignment uses the algorithm of Needleham and Wunsch (1970) to produce an alignment with gaps and is the output at the conclusion of the analysis. If no good homologies are observed after running FASTA, it is sometimes helpful to repeat the analysis using a smaller k-tuple value or an alternate scoring matrix.

A convenient method of performing a FASTA search on a local computer is to use a free e-mail server (UNIT 19.1). A computerized service provided by several institutions automatically accepts requests for FASTA searches via e-mail, searches the sequence contained within the mail message against a variety of databases, and then returns the results via e-mail. A mail server for FASTA is available from the University of Houston and the European Bioinformatics Institute. A sample mail message is shown in Figure 7.7.5. The FASTA program is currently available for most computers in common use: UNIX, VMS, IBM-compatible, and Macintosh versions of the program are available via the Internet or from Bill Pearson at the University of Virginia. The DOS and Macintosh versions of FASTA can search the CD-ROM version of GenBank or EMBL databases in a few hours.

Both FASTA and TFASTA (which, like T-BLAST, checks an amino acid sequence against DNA sequence databases) produce a score representing the quality of the match. This score is not the same as the significance score determined by one of the BLAST programs. FASTA and TFASTA scores are for each matching pair

```
TITLE A test search of the EMBL Other Mammalian DNA sequences
LIB EMAM
WORD 4
LIST 100
ALIGN 20
SEQ
tgcttggtgaggagccataggacgagagcttctggtgaagtgtgtttcttgaatcat
caccacatggacagcaaa
END
```

Figure 7.7.5 Text of a message sent to the EBI FASTA mail server (e-mail address: fasta@ebi.ac.uk). This message requests that the sequence be searched against the Other Mammalian section of the EMBL database. The answer will include the top 100 matching sequences and alignments of the top 20 matching sequences.

of sequences generated using the PAM250 matrix described earlier (see Comparison of Two Sequences). Thus, the score is a numerical value representing the quality of the observed similarity between two sequences. A high-scoring similarity is generally a significant match; however, FASTA does not give a significance value like that provided by BLAST programs. With FASTA output, it is up to the user to recognize whether a particular observed match is significant or not, and this determination is not always an easy task. Moreover, with either FASTA or BLAST, determining whether a match is biologically significant is up to the investigator. If the BLAST mathematical significance value indicates that a match is not the result of chance, then the biological significance is very likely to be high. However, if the mathematical significance indicates a large probability that the match is just a random event, this does not necessarily mean that the biological significance is also low. Unfortunately, there is no systematic way to determine biological significance of an apparent homology in cases where mathematical similarity is low. In making this determination, the known or presumed function of the protein, the consensus match to known active sites or sequence motifs, and the number of distinct sequences that fit the proposed homologous group must be taken into account.

Because BLAST and FASTA use different algorithms, it is advisable to perform searches on a given sequence using both of these programs. If a significant match is not identified using one of the programs, use the other.

The BLAST programs are available from e-mail servers and as an Internet service (UNIT

19.1) using either a command-line interface or graphical interface (NetBLAST). UNIX and VMS versions are available via network file servers (see Table 7.7.2). A public domain program called MAILFASTA is now available; this program will take a DNA or protein sequence, reformat it, and e-mail it to any or all of the following e-mail servers: BLAST, BLITZ, BLOCKS, FASTA, GeneID, GRAIL, PredictProtein, and Pythia.

GENETIC SEQUENCE DATABASES AND OTHER ELECTRONIC RESOURCES AVAILABLE TO MOLECULAR BIOLOGISTS

Different gene sequence databases can be searched via the Internet or are available to anyone free of charge either via network file transfer protocol (FTP; see UNIT 19.1) or for the price of the distribution media. Some of the larger sequence analysis packages (e.g., GCG) provide tools that reformat the database data files as they are received from the database distributors. Reformatting is generally performed by a computer systems manager. Many users currently access the protein and DNA databases on the Internet. Detailed descriptions of the databases and how to access them via the Internet can be found in UNIT 19.2. UNIT 19.1 is a general guide to the Internet and includes a list of electronic resources available to molecular biologists. Detailed protocols to carry out homology searches using the BLAST family of programs are provided in UNIT 19.3.

Acknowledgement: We wish to thank Rose Marie Woodsmall and her colleagues at the National Center for Biotechnology Information for their assistance in updating this unit.

LITERATURE CITED

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Church, G.M. and Kieffer-Higgins, S. 1988. Multiple DNA sequencing. *Science* 240:185-188.
- Feng, D.F. and Doolittle, R.F. 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* 25:351-360.
- Fickett, J. 1982. Recognition of protein coding regions in DNA sequences. *Nucl. Acids Res.* 10:5303-5318.
- Freier, S.M., Kierzek, R., Jaeger, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T., and Turner, D.H. 1986. Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci. U.S.A.* 83:9373-9377.
- Gonnet, G.H., Cohen, M.A., and Benner, S.A. 1992. Exhaustive matching of the entire protein sequence database. *Science* 256:1443-1445.
- Henikoff, S. and Henikoff, J.G. 1993. Performance evaluation of amino acid substitution matrices. *Proteins* 17:49-61.
- Higgins, D.G. and Sharp, P.M. 1988. Clustal: A package for performing multiple sequence alignment on a microcomputer. *Gene* 73:237-244.
- Higgins, D.G. and Sharp, P.M. 1989. Fast and sensitive multiple sequence alignments on a microcomputer. *Comp. App. Biosci.* 5:151-153.
- Karlin, S. and Altschul, S.F. 1990. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci. U.S.A.* 87:2264-2268.
- Needleman, S.B. and Wunsch, C.D. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48:443-453.
- Pearson, W.R. and Lipman, D.J. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A.* 85:2444-2448.
- Schuler, G.D., Altschul, S.F., and Lipman, D.J. 1991. A workbench for multiple alignment construction and analysis. *Proteins Struct. Funct. Genet.* 9:180-190.
- Schwartz, R.M. and Dayhoff, M.O. (eds.) 1978. Matrices for Detecting Distant Relationships: Atlas of Protein Sequence and Structure. National Biomedical Research Foundation, Washington, D.C.
- Smith, T.F. and Waterman, M.S. 1981. Identification of common molecular subsequences. *J. Mol. Biol.* 147:195-197.
- Turner, D.H., Sugimoto, N., Jaeger, J.A., Longfellow, C.E., Freier, S.M., and Kierzek, R. 1987. Improved parameters for prediction of RNA structure. *Cold Spring Harbor Symp. Quant. Biol.* 52:123-133.
- Turner, D.H., Sugimoto, N., and Freier, S.M., and Kierzek, R. 1988. RNA structure prediction. *Annu. Rev. Biophys. Chem.* 17:167-192.
- Wilbur, W.J. and Lipman, D.J. 1983. Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci. U.S.A.* 80:726-730.
- Zuker, M. 1989a. On finding all suboptimal foldings of an RNA molecule. *Science* 244:48-52.
- Zuker, M. 1989b. The use of dynamic programming algorithms in RNA secondary structure prediction. *In* Mathematical Methods for DNA Sequences (M.S. Waterman, ed.) p. 159-184. CRC Press, Boca Raton, Fla.

Contributed by J. Michael Cherry
Stanford University
Palo Alto, California
cherry@genome.stanford.edu

APPENDIX

This section lists products and other resources designed for DNA and protein sequence analysis—e.g., databases, software, and journals—and how to obtain them. Some but not all of these products have been described or cross-referenced earlier in this unit.

Commercially Available Software

Table 7.7.1 is a brief description of sequence analysis software packages and some comments concerning their notable features.

Shareware and Free Software

Table 7.7.2 is a brief description of software that is distributed without charge. See also FTP archive and electronic mail listings below.

Electronic Mail Servers

Electronic mail is available at most research institutions, and these servers provide useful information, software, and analysis at no charge. With most servers, sending the simple message “HELP” will cause the latest documentation to be transmitted via e-mail (see *UNIT 19.1*).

FTP Archives Available on Internet

A vast amount of software and databases applicable to molecular biology is available at no charge from anonymous FTP servers as described in *UNIT 19.1*.

Free Databases

A brief description of databases that are available without charge is in *UNIT 19.1*.

Addresses of Databases for Retrieval and Submission

DNA Data Bank of Japan (DDBJ)

Submissions

Laboratory of Genetic Information

Analysis

Center for Genetic Information Research

National Institute of Genetics (NIG)

Mishima, Shizuoka 411 Japan

E-mail: ddbjsub@ddbj.nig.ac.jp

WWW URL: <http://www.nig.ac.jp>

EMBL Data Library

EMBL Nucleotide Sequence Submissions

European Bioinformatics Institute

Hinxton Hall, Hinxton

Cambridge CB10 1RQUK

E-mail (retrieval): datilib@ebi.ac.uk

E-mail (submission): datasub@ebi.ac.uk

WWW URL: <http://www.ebi.ac.uk>

GenBank

National Center for Biotechnology
Information

National Library of Medicine

National Institutes of Health

Building 38A, Room 8N-803

8600 Rockville Pike

Bethesda, MD 20894

(301) 496-2475

E-mail: info@ncbi.nlm.nih.gov

E-mail (submission): gb-sub@ncbi.nlm.nih.gov

International Protein Information

Database in Japan

Research Institute for Biosciences

Science University of Tokyo

2669 Yamazaki

Noda 278 Japan

Martinsried Institute for Protein

Sequences

Am MPI für Biochemie

8033 Martinsried, Germany

49-89-8578-2656

E-mail: datasub@mips.embnet.org

Protein Identification Resource (PIR)

National Biomedical Research Foundation

Georgetown University Medical Center

3900 Reservoir Road, N.W.

Washington, DC 20007

E-mail: pirmail@nbrf.georgetown.edu

Publications

The Applications of Computers to Research on Nucleic Acids—published annually by IRL Press, P.O. Box Q, McLean, VA 22101. (703) 356-4031; FAX: (703) 356-4303.

BioTechniques—published monthly by Eaton Publishing, 154 E. Central Street, Natick, MA 01760. (508) 655-8282; FAX (508) 655-9910.

BioTechnology Software—published bimonthly by Mary Ann Liebert, 1651 Third Avenue, New York, NY 10128. (212) 289-2300; FAX (212) 289-4697.

Computer Applications in the Biosciences (CABIOS)—published monthly by Oxford University Press, 2001 Evans Road, Cary, NC 27513. (919) 677-0977; FAX: (919) 677-8877.

Table 7.7.1 Commercially Available Sequence-Analysis Software

Name	Description	Source
Ball & Stick	Molecular graphics display, printing, and manipulation for the Macintosh; a helpful demo is available via ftp.bio.indiana.edu.	Cherwell Scientific Publishing 27 Park End Street Oxford, OX1 1HU, UK (44) 865 774 800 FAX (44) 865 794 664
ChemDraw Chem3D	Desktop publishing for chemical structures in two and three dimensions; very useful for creating journal figures or instructional illustrations of molecular structures.	Cambridge Scientific Computing 875 Massachusetts Ave., Suite 61 Cambridge, MA 02139 (617) 491-6862 FAX (617) 491-8208
DNA Strider	A simple and very useful Macintosh sequence analysis program; includes restriction mapping and circular plasmid maps generated from DNA sequence.	Christian Marck Service de Biochimie—Bat 142 Centre d'Etudes Nucléaires de Saclay 91191 Gif-sur-Yvette Cedex France
EUGENE & SAM	Extensive nucleic and amino acid analysis package for Sun Microsystems SPARCstations; includes FASTA for sequence searches and very quick keyword searching on DNA and protein database.	Lark Sequencing Technologies 9545 Katy Freeway, Suite 200 Houston, TX 77024 (713) 464-7488; (800) 288-3720 FAX (713) 464-7492
GCG	Package including nucleic and amino acid sequence analysis, sequencing project management, database searching, RNA folding, protein secondary structure prediction, and sequence motif generation and searching for VMS and UNIX multi-user systems.	Genetics Computer Group University Research Park 575 Science Drive, Suite B Madison, WI 53711 (608) 231-5200 FAX (608) 231-5202 E-mail: help@gcg.com
Gene Construction Kit DNA Inspector	The ultimate plasmid database, design, and presentation tool for the Macintosh; DNA Inspector is a basic analysis program for the Macintosh; demo disks available.	Textco 27 Gilson Road West Lebanon, NH 03784 (603) 643-1471
GENEPRO	Complete sequence-analysis software for DOS, including nucleic and amino acid analysis; GenBank and EMBL DNA databases or PIR and SWISS-PROT protein databases can be searched on floppy disks; demo disk available.	Riverside Scientific Enterprises 15705 Point Monroe Drive N.E. Bainbridge Island, WA 98110 (206) 842-9498 FAX (206) 842-9534
HIBIO DNASIS PROSIS	Complete DNA and protein analysis packages for DOS; includes restriction analysis, secondary structure prediction, sequencing project management, digitizer and speech synthesizer support, and database searches from CD-ROM drive.	Hitachi Software Engineering America Computer Division 1111 Bayhill Drive, Suite 395 San Bruno, CA 94066 (800) 624-6176 In CA (800) 225-9925 FAX (415) 615-7699
Intelligenetics Suite PC Gene GeneWorks	Multifunction sequence analysis package—Intelligenetics Suite is for VMS and Sun Microsystems, PC Gene is for DOS, and GeneWorks is for Macintosh.	Intelligenetics 700 East El Camino Real Mountain View, CA 94040 (415) 962-7300 FAX (415) 962-7302

continued

Table 7.7.1 Commercially Available Sequence-Analysis Software, continued

Name	Description	Source
LaserGene Protean	Excellent nucleic and amino acid sequence analysis packages for DOS and Macintosh computers; programs for sequence comparison and alignment, editing and analysis, sequencing project management, restriction analysis and mapping, and database searching; demo available.	DNASStar, Inc. 1228 South Park Street Madison, WI 53715 (608) 258-7420 FAX (608) 258-7439
MacMolly	Good basic sequence analysis package for Macintosh.	Soft Gene Berlin Offenbacher Str 5, D-100 Berlin, 33 Germany 030-821-1407
MacVector	Complete package for the Macintosh including restriction map presentation, sequence comparison and database searching from hard disk or CD-ROM drive, and protein secondary structure analysis.	International Biotechnologies P.O. Box 9558 New Haven, CT 06535 (203) 786-5600; (800) 243-2555 FAX (203) 786-5694
Pearson Sequence Analysis package FASTA	UNIX, VMS, Macintosh, and DOS-based sequence analysis software and fast database searching; may be stored on either hard disk or CD-ROM.	William R. Pearson Department of Biochemistry Box 440 Jordan Hall University of Virginia Charlottesville, VA 22908
Plasmid Artist	Publication-quality restriction map diagram production program for the Macintosh.	Clontech Laboratory 4030 Fabian Way Palo Alto, CA 94303 (415) 424-8222 FAX (415) 424-1352
Pro-RFLP	Macintosh and DOS gel analysis software for use with a scanner or TV camera; provides a number of features including calibration, database for unknowns, searching for matches to unknowns, complete printer output.	DNA ProScan, Inc. P.O. Box 12185 Nashville, TN 37212 (800) 841-4362
Rodger Staden programs	Several sequence assembly and analysis packages and complete UNIX analysis package; includes features that can analyze the fluorescence trace output files from ABI automatic sequencer machines.	Rodger Staden MRC Lab of Molecular Biology Hills Road Cambridge, England CB2 2QH
Sequencher	Macintosh sequencing support software including vector screening, ORF reports, translations, restriction maps, and a good interface with fluorescent sequencers.	Gene Codes Corporation 2901 Hubbard Road Ann Arbor, MI 48105 (313) 769-7249; (800) 497-4939 FAX (313) 930-0145
XQS ALIGN PSQ NAQ ATLAS	Database management and searching software featuring very fast and flexible keyword searches for VMS.	National Biomedical Research Foundation. Georgetown Univ. Medical Center 3900 Reservoir Road, N.W. Washington, DC 20007 (202) 687-2121 FAX (202) 687-1662

Table 7.7.2 Free Sequence-Analysis Software Programs

Name	Description	Source ^a
Amplify	Utility to aid in designing, analyzing, and even simulating experiments for PCR reactions.	FTP: sumex-aim.stanford.edu
BinHex 4.0	Essential utility required to decode files retrieved from FTP archives or received from mail servers for the Macintosh.	Available from most FTP archives; however, it is generally easiest to ask a colleague for a copy
BLAST	Ultra-fast database searching program for UNIX and VMS C source.	FTP: ncbi.nlm.nih.gov EM: netserv@ebi.ac.uk
BioSCAN	Forms-based search and retrieval resource from the University of North Carolina, Chapel Hill.	WWW URL: http://genome.cs.unc.edu/
BLOCKS Search	Forms-based submission tool via WWW for comparisons to BLOCKS database of highly conserved regions in protein sequences.	WWW URL: http://www.blocks.fcrc.org/
BoxShade	Prepares publication-quality figures of aligned sequences; resulting PostScript output contains shading over regions of similarity.	EM: netserv@ebi.ac.uk
Clustal	Multiple sequence alignment and production of phylogenetic dendrograms; available for DOS, Macintosh, UNIX, and VMS.	FTP: ftp.bio.indiana.edu FTP: ch.embnet.org
Covariation	HyperCard stack to analyze an aligned set of sequences to aid in the identification of covariations (two or more sites in an RNA or DNA molecule that are evolving together).	FTP: ftp.bio.indiana.edu
Disinfectant	Excellent Macintosh virus detection and removal application.	FTP: ftp.acns.nwu.edu FTP: sumex-aim.stanford.edu
Entrez	Application for access to genetics and biomolecular subset of MEDLINE bibliographic database; available for Macintosh, Windows, UNIX, and VMS; computer must be registered with NCBI to obtain access; registration is free.	FTP: ncbi.nlm.nih.gov WWW URL: http://atlas.nlm.nih.gov:5700/Entrez/index.html/
Enzyme Kinetics	HyperCard stack to analyze and plot enzyme kinetic experimental data.	FTP: ftp.bio.indiana.edu
FASTA	Fast sequence database searching program for UNIX, and VMS.	FTP: uvaarpa.virginia.edu
GDE	Superb analysis package for Sun Microsystems SPARCstations; includes multiple sequence alignment, sequencing project management, database searching, and more.	FTP: golgi.harvard.edu
GenBank Search	HyperCard stack that prepares a mail message for the GenBank FASTA and sequence retrieval mail servers; requires MacTCP software on Macintosh computers.	FTP: ftp.bio.indiana.edu

continued

Table 7.7.2 Free Sequence-Analysis Software Programs, continued

Name	Description	Source ^a
GeneMapper	Full-featured coding region identification and more for Sun Microsystems SPARCstations.	FTP: haywire.nmsu.edu
GenoBase	An experimental WWW resource provided by NIH; provides tables and query capabilities to an object-oriented molecular biology database.	WWW URL: http://specter.dcrn.nig.gov:8004/
GenQuest	Forms-based database searching via WWW from Genome Data Base; provides BLAST, FASTA, and Smith-Waterman searching SWISS-PROT, Genome Sequence Database (GSDB), Protein Databank (PDB), and Prosite.	WWW URL: http://www.gdb.org
Gopher	Client application to the growing number of Internet Gopher servers.	FTP: ftp.bio.indiana.edu FTP: boombox.micro.umn.edu
Linkage	Performs linkage analysis on genetic markers; available for UNIX, DOS, and VMS computers.	FTP: corona.med.utah
LoopViewer	Displays results of Zuker RNA folding programs on a Macintosh; produces output similar to that shown in Figure 7.7.4B.	FTP: ftp.bio.indiana.edu
MACAW	Excellent multiple sequence alignment and editing tool for Microsoft Windows and Macintosh.	FTP: ncbi.nlm.nih.gov
MacMolecule	Displays and rotates molecular structures of nucleic acids and proteins; very useful as a teaching aid.	FTP: ftp.bio.indiana.edu
MacPattern	Macintosh application that searches for patterns in protein sequences utilizing the PROSITE motif database; allows documentation from PROSITE to be directly accessed as well as entry of user-defined patterns.	FTP: ftp.bio.indiana.edu EM: netserv@ebi.ac.uk
MacPlasMap	Plasmid map drawing program for the Macintosh; the cutting sites of restriction enzymes must be entered by the user.	FTP: ftp.bio.indiana.edu EM: netserv@ebi.ac.uk
MandM	Materials and methods HyperCard stack for organizing laboratory protocols.	EM: netserv@ebi.ac.uk
MapMaker	Linkage analysis of F ₂ and CEPH populations; available for Sun Microsystems and VMS.	EM: mapm@genome.wi.mit.edu FTP: genome.wi.mit.edu
Mase	Multiple sequence alignment editor for UNIX.	FTP: mbcr.harvard.edu
MFOLD (or LRNA & CRNA)	VMS version for the prediction of Zuker RNA secondary structures.	FTP: amber.mgh.harvard.edu FTP: ftp.bio.indiana.edu
MulFold	Macintosh version of the MFOLD for prediction of Zuker RNA secondary structures, created by Don Gilbert.	FTP: ftp.bio.indiana.edu

continued

Table 7.7.2 Free Sequence-Analysis Software Programs, continued

Name	Description	Source ^a
NCSA GelReader	Automates the measurement of DNA fragment lengths using a scanned image of an autoradiogram or EtBr-stained gel on a Macintosh; identifies lanes and bands, then exports the predicted band sizes as a text file for use with a spreadsheet.	FTP: ftp.ncsa.uiuc.edu
NCSA Mosaic	First WWW browser for Macintosh, Windows, and UNIX; requires an Internet connection.	FTP: ftp.ncsa.uiuc.edu
Netscape	Commercial WWW browser, available to educational users for free; provides fast displays and inline images in either GIF or JPEG encoding formats; available for Macintosh, Windows, and UNIX.	FTP: ftp.netscape.com
NIH Image	General image-analysis program for biologists that uses scanned images on a Macintosh.	FTP: alw.nih.gov
OSP	Oligonucleotide selection program to create appropriate PCR primers; selects primers using the sequence of the target region determining GC content, possible secondary structure, primer and amplified product length.	FAX: (314) 362-2985 c/o Paula Kassos (Phil Green, Washington University, St. Louis) EM: pg@genome.wustl.edu
PAUP	Phylogenetic analysis software for DOS and Macintosh.	EM: swofford@uxh.cso.uiuc.edu (David Swofford, Illinois Natural History Survey, Champaign, IL)
PGEN	Oligonucleotide design and analysis software for DOS; requires EGA or VGA graphics adaptors; will work from a 360K floppy drive.	FTP: ftp.bio.indiana.edu
Phylip	Phylogenetic analysis software for UNIX.	FTP: ftp.bio.indiana.edu
Plot/A	Protein analysis software with a variety of secondary structure predictions.	FTP: ftp.bio.indiana.edu
PlotZ	Graphical restriction mapping for UNIX and VMS.	FTP: amber.mgh.harvard.edu
PLSearch	Database of primary protein sequence patterns determined by analyzing the SWISS-PROT database, for UNIX.	FTP: mbcr.harvard.edu
Primer	Primer selection software for UNIX, VMS, DOS, and Macintosh; selects primers using the sequence of the target region by determining the GC content, possible secondary structure, primer, and amplified product length.	FTP: genome.wi.mit.edu
SeqApp	Analyzes DNA and protein sequences while also allowing easy access to WAIS and Internet Gopher resources; developed by Don Gilbert; requires a Macintosh and MacTCP.	FTP: ftp.bio.indiana.edu

continued

Table 7.7.2 Free Sequence-Analysis Software Programs, continued

Name	Description	Source ^a
SeqSpeak	Simple DNA sequence editor; developed by Keith Conover for the Macintosh; provides audio feedback using a digitized human voice.	FTP: ftp.bio.indiana.edu
SLINK and FASTSLINK	Linkage analysis software for UNIX computers	FTP: watson.hgen.pitt.edu
Virtual Library of BioSciences	Expanding collection of Internet resources available for biology; continually updated and an excellent starting point for exploration of information available on the Internet.	WWW URL: http://golgi.harvard.edu/biopages.html
WAIS	Wide area information service, providing effective access to large collections of information.	FTP: think.com (NeXT, Motif, and Macintosh software) FTP: samba.oit.unc.edu (DOS, SunView, and VMS software)

^aSource abbreviations: FTP, file transfer protocol (all are anonymous); EM, electronic mail; WWW URL, World Wide Web Uniform Resource Locator. See *UNIT 19.1* for descriptions of these sources. Contact person is listed in parentheses.

Polony DNA Sequencing

This unit provides protocols for performing polony DNA sequencing, a nonelectrophoretic sequencing method that eliminates *in vivo* cloning artifacts and affords a lower cost per base than conventional Sanger sequencing (Shendure et al., 2005). The first step in polony DNA sequencing is to construct an *in vitro* paired-tag library from genomic DNA (Basic Protocol 1; Fig. 7.8.1). Next, library molecules are clonally amplified on microbeads by emulsion PCR (Basic Protocol 2; Dressman et al., 2003). This clonal amplification yields polymerase colonies, or polonies, that can be sequenced. A Support Protocol describes a method used to optimize the amount of library DNA used as the ePCR template. Amplicon-bearing beads are then enriched (Basic Protocol 3) and two-dimensionally arrayed within an acrylamide gel matrix on a microscope coverglass (Basic Protocol 4). Finally, millions of short reads are generated in parallel from the microbeads via a cyclic DNA sequencing strategy that utilizes T4 DNA ligase to selectively tag each microbead with fluorescent labels that correlate with the unique nucleotide sequence present on any given bead (Basic Protocol 5).

The combination of protocols can be thought of as a modular platform, with each major step (library construction, emulsion PCR, sequence readout) independent enough to be replaced by an alternate. For example, if paired tags are not necessary to place reads (i.e., to match sequences to a reference genome sequence), single-tag libraries can be constructed (e.g., for barcode sequencing or RNA quantitation). The emulsion PCR protocol presented in this unit is a variation of that published by Dressman et al. (2003), but should not be thought of as the only applicable amplification method. Different sequencing biochemistries can also be used; in fact, during development of the polony sequencing platform, a number of alternate polymerase- and ligase-driven approaches were evaluated.

To implement the protocols presented in this unit, various instruments (e.g., fluorescent microscope, autosampler, and flow cell) and computer hardware and software are required. A discussion of the critical features of the setup is provided (see Critical Parameters and Troubleshooting), but the reader is directed to <http://arep.med.harvard.edu/Polonator> for more detailed information, including full schematics and drawings. Instructions are available there for using the hardware and software, as are binary and source codes for the software. All software is free and open-source, and all hardware is commercially available.

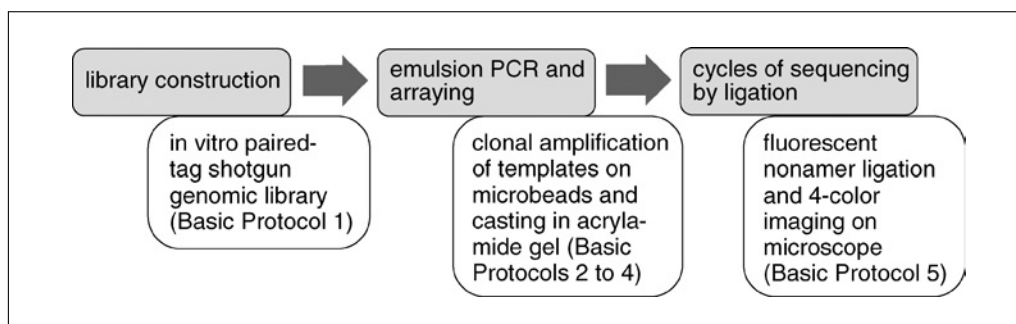


Figure 7.8.1 General procedure for polony sequencing. A paired-tag genomic library is used as template for emulsion PCR on microbeads to generate polymerase colonies (polonies). The beads are cast on a coverslip in a thin layer of polyacrylamide, and put through iterative cycles of single-base sequencing.

CONSTRUCTION OF A SHOTGUN PAIRED-TAG GENOMIC LIBRARY

Polony genomic resequencing is generally performed on a shotgun paired-tag library (Shendure et al., 2005). Each library molecule is 135 bp in length, and has two 17- to 18-bp paired genomic tags separated and flanked by common sequences (Fig. 7.8.2). The key step in this protocol is the circularization of randomly sheared and size-selected genomic DNA around a synthetic insert-linker. This links sequences that are separated on the genome by a defined distance distribution.

Materials

- Genomic DNA
- Buffer EB (Qiagen)
- End-It DNA end repair kit (Epicentre), including:
 - 10× buffer
 - 10× ATP
 - 10× dNTP mix
 - Enzyme mix
- 10× PCR buffer without MgCl₂ (Invitrogen)
- 50 mM MgCl₂
- 100 mM dATP
- 5 U/μl *Taq* DNA polymerase
- 20 mg/ml glycogen
- Oligonucleotides:
 - 100 μM T30-T: 5'-phosphorylated-GTCGGAGGCCAAGGCGGCCGTACGT CCAACT-3' (purified by HPLC)
 - 100 μM T30-B: 5'-phosphorylated-GTTGGACGTACGGCCGCCTTGGCC TCCGACT -3' (purified by HPLC)
 - 1 mM N6 oligonucleotides: 5'-NNNN*N*N-3' (*signifies phosphorothioate linkage; IDT)
 - 100 μM FDV-B: 5'-ATCACCGACTGCCCATAGAGAGGAAAGCGGAGGC GTAGTGGTT-3' (purified by HPLC)
 - 100 μM FDV-T: 5'-AACCACTACGCCTCCGCTTTCCTCTCTATGGGCAGT CGGTGAT-3' (purified by HPLC)
 - 100 μM RDV-B: 5'-AGAGAATGAGGAACCCGGGGCAGTT-3' (purified by HPLC)
 - 100 μM RDV-T: 5'-AACTGCCCCGGGTTTCCTCATTCTCT-3' (purified by HPLC)
- Quick ligation kit (NEB), including:
 - 2× Quick ligation buffer
 - Quick T4 DNA ligase
- 20 U/μl Exonuclease I (NEB)
- 100 U/μl Exonuclease III (NEB)
- TE buffer, pH 8.0 (APPENDIX 2)
- 25 mM dNTP mix (25 mM each nucleotide)
- 10× RepliPhi phi29 reaction buffer (Epicentre Technologies)
- 100 U/μl RepliPhi phi29 polymerase (Epicentre)
- 10× NEBuffer 4 (NEB)
- 1.6 mM *S*-adenosylmethionine (SAM; diluted from 32 mM stock in 1× NEBuffer 4; NEB)
- 2 U/μl *Mme*I (NEB)
- 2000 U/μl T4 DNA ligase and 10× buffer (NEB)
- 40% (w/v) polyethylene glycol 8000 (PEG) in H₂O
- 10× NEBuffer 2 (NEB)
- 10 U/μl *E. coli* DNA polymerase I (NEB)
- 5 U/μl Platinum *Taq* DNA polymerase (Invitrogen)

Qiaquick columns (Qiaquick PCR cleanup kit; Qiagen)
NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies)
Microcon-30 column (Millipore)
Thermal cycler

Additional reagents and equipment for DNA quantitation (*APPENDIX 3D*),
phenol/chloroform extraction and ethanol precipitation (*UNIT 2.1A*), and agarose
and polyacrylamide gel electrophoresis (*UNITS 2.5A & 2.7*)

Prepare genomic DNA for circularization

1. Shear 15 µg genomic DNA to desired size distribution.

The Hydroshear (<http://www.genomicsolutions.com>) is the recommended instrument for this purpose since, when the manufacturer's instructions are followed, it produces a population of fragments having a relatively tight size distribution and ends that can be efficiently blunted. The authors typically shear DNA at a concentration of ~60 ng/µl.

2. Purify DNA on Qiaquick columns (≤ 10 µg per column) as per manufacturer's instructions.

Use as few columns as possible.

3. To repair the DNA ends (i.e., blunt them following shearing), combine the DNA with the components of the End-It DNA end repair kit as follows:

10 µg column-purified, sheared genomic DNA in buffer EB
25 µl 10× buffer
25 µl 10× ATP
25 µl 10× dNTP mix
5 µl enzyme mix
H₂O to a final volume of 250 µl.

Incubate 1 hr at room temperature.

4. Purify DNA on a Qiaquick column per manufacturer's instructions, except elute with 90 µl buffer EB.

5. Quantitate DNA using a NanoDrop ND-1000 spectrophotometer (see *APPENDIX 3D*).

Concentration should be ~100 ng/µl.

6. Incubate 15 min at 70°C to eliminate any residual enzyme activity.

7. Add an adenosine tail by combining the following components:

2.2 µg end-repaired DNA
10 µl 10× PCR buffer without MgCl₂
6 µl 50 mM MgCl₂
0.5 µl 100 mM dATP
0.5 µl 5 U/µl *Taq* DNA polymerase
H₂O to 100 µl.

Incubate 30 min at 70°C then cool to 4°C. Transfer to ice.

*In the A-tailing reaction, *Taq* polymerase adds a single A to the 3' ends of the blunted DNA in a template-independent fashion.*

8. Extract DNA from the reaction mixture with phenol/chloroform and precipitate in ethanol using 20 μg glycogen as carrier (*UNIT 2.1A*). Resuspend pellet in 40 μl buffer EB.

Glycogen should be used as a carrier in all ethanol precipitation steps in this protocol.

9. Separate the sheared, A-tailed DNA on a 6% polyacrylamide gel with TBE buffer (*UNIT 2.7*). Stain with ethidium bromide.

For Invitrogen precast gels, mix 20 μl DNA with 5 μl of 5 \times loading buffer (Novex) and load across two 0.5-cm lanes.

10. Cut out the desired bands and extract DNA using the crush-and-soak method (*UNIT 2.7*). Precipitate in ethanol using glycogen as a carrier, and resuspend in 10 μl buffer EB.

For a paired-tag library with an average intertag distance of 1 kb, excise and pool gel bands centered at 1 kb with a total width of no more than 500 bp. Exposure of the DNA to UV light during visualization should be minimized.

11. Quantitate the size-selected DNA by separating on an agarose or polyacrylamide gel (*UNIT 2.5A or 2.7*) and comparing the brightness of the smear to that of a known amount of a molecular weight standard.

The concentration should be ~ 20 to 40 ng/ μl .

Circularize genomic material around synthetic oligonucleotides

12. Anneal oligonucleotides T30-T and T30-B in a thermal cycler by combining equal amounts and heating to 95°C for 10 min and then slowly cooling to room temperature over the course of 1 hr.

The final concentration of the T30 insert-linker should be 100 μM . T30 has T 3' overhangs designed to be complementary to the A 3' overhangs of the genomic DNA.

13. Ligate genomic DNA fragments in the presence of annealed T30 using the Quick ligation kit as follows:

170 ng size-selected DNA (0.25 pmol at ~ 1 kb)
0.8 μl 1 μM annealed T30 (0.8 pmol)
40 μl 2 \times Quick ligation buffer
4.0 μl Quick T4 DNA ligase
H₂O to a final volume of 80 μl .

Incubate reaction 10 min at room temperature then heat-inactivate 10 min at 65°C. Use the heat-inactivated reaction mixture directly in the next step without further purification.

The ligation should be performed under conditions favoring formation of monomeric recombinant circles (i.e., genomic DNA–T30). To this end, T30 should be present at a three-fold molar excess to genomic fragments.

14. Eliminate all noncircularized material by exonucleolysis as follows:

1.0 μl 20 U/ μl Exonuclease I
0.1 μl 100 U/ μl Exonuclease III
80 μl heat-inactivated reaction mix
TE buffer, pH 8.0, to a final volume of 90 μl .

Incubate 45 min at 37°C, and then heat-inactivate 20 min at 80°C. Use this reaction mixture directly in the next step without further purification.

Amplify circular DNA

15. Prepare the master mix for hyper-branched rolling-circle amplification of the circular DNA:

12 μ l 25 mM dNTP mix
30 μ l 10 \times RepliPhi phi29 reaction buffer
15 μ l 1 mM N6 oligonucleotides
213 μ l H₂O
30 μ l circularized DNA.

Split into 6 tubes of 50 μ l each.

16. Denature material by heating to 95°C for 5 min, then anneal by rapidly cooling to 4°C. Add 2.5 μ l phi29 polymerase to each tube, keeping on ice. Incubate overnight at 30°C.
17. Pool amplified circular DNA into a single tube. Purify DNA using a Microcon-30 column according to the manufacturer's instructions, washing with 1 ml TE buffer, pH 8.0. Wash the membrane several times to maximize recovery.
18. Resuspend the DNA using 750 μ l buffer EB, preheated to 50°C. Quantitate the DNA using a NanoDrop ND-1000 spectrophotometer.

Concentration should be ~230 ng/ μ l.

Release paired tags

19. Digest the amplified DNA with *MmeI* to release the 70-bp T30-linked paired tags:

40 μ g amplified circular DNA
100 μ l 10 \times NEBuffer 4
1.6 μ l of 1.6 mM SAM
60 μ l *MmeI* (2 U/ μ l)
H₂O to 1000 μ l.

Split into eight tubes of 125 μ l each on ice, and then incubate at 37°C for 30 min.

MmeI cuts 18 to 19 bp from its recognition site, which is 1 bp from the start of the genomic fragment. The released fragments contain the common T30 sequence flanked by two 17- to 18-bp tags of genomic DNA.

20. Immediately extract the DNA with phenol/chloroform and precipitate with ethanol using glycogen as a carrier. Resuspend pellet in a total volume of 80 μ l TE buffer, pH 8.0.
21. Purify the 70-bp paired-insert library as in steps 9 and 10, resuspending precipitated DNA in 20 μ l TE buffer, pH 8.0.

The sample should be split across four 0.5-cm lanes of a precast 6% polyacrylamide gel in TBE buffer.

22. Quantitate the gel-purified DNA by electrophoresis (step 11).

DNA concentration should be ~12 ng/ μ l.

Add emulsion PCR adapters

23. To perform the blunting reaction (to remove 2-nt 3' overhangs from the *MmeI* digestion), combine the DNA with components of the End-It DNA end repair kit as follows:

8.5 μ l paired-insert library (~100 ng DNA)
1.25 μ l 10 \times buffer
1.25 μ l 10 \times ATP
1.25 μ l 10 \times dNTP mix
0.25 μ l enzyme mix.

Incubate at room temperature for 45 min.

24. Put on ice and add TE buffer, pH 8.0 to bring volume to 50 μ l. Extract the DNA with phenol/chloroform and precipitate in ethanol using 20 μ g glycogen as carrier. Resuspend DNA in 8 μ l TE, pH 8.0.

25. Prepare emulsion PCR amplification primer sequences by annealing oligonucleotides FDV-T and FDV-B and oligonucleotides RDV-T and RDV-B as in step 12.

The final concentration of each oligonucleotide should be 50 μ M. FDV and RDV are not 5'-phosphorylated, so they will not ligate to one another.

26. Prepare the ligation reaction mixture by mixing the following components on ice:

8.0 μ l blunt-ended library DNA (~100 ng, 2 pmol)
1.0 μ l 50 μ M FDV (50 pmol)
1.0 μ l 50 μ M RDV (50 pmol)
2.5 μ l 10 \times T4 DNA ligase buffer (~0.5 \times)
21.1 μ l 40% PEG
12.3 μ l H₂O.

27. Mix on ice, then add 2.0 μ l of 2000 U/ μ l T4 DNA ligase and incubate overnight at 16°C to ligate primer sequences to the library molecules.

After ligation, several species will be present: FDV-insert-FDV, FDV-insert-RDV, and RDV-insert-RDV.

28. Increase volume to 100 μ l with TE buffer, pH 8.0. Extract the DNA with phenol/chloroform, precipitate in ethanol using glycogen as a carrier, and resuspend in 10 μ l buffer EB.

29. Purify the 135-bp FDV-RDV paired-insert library as in steps 9 and 10, resuspending the precipitated DNA in 10 μ l TE buffer, pH 8.0.

30. Assemble on ice the reaction mixture used to remove nicks in double-stranded library molecules by nick translation:

10.0 μ l library DNA
0.5 μ l 25 mM dNTP mix
2.5 μ l 10 \times NEBuffer 2
1.0 μ l 10 U/ μ l *E. coli* DNA polymerase I
11.0 μ l H₂O.

Incubate 30 min at 16°C.

31. Increase volume to 100 μ l with TE buffer, pH 8.0. Extract the DNA with phenol/chloroform, precipitate in ethanol using glycogen as a carrier, and resuspend in 10 μ l TE buffer, pH 8.0.

Amplify FDV/RDV-adapted library

32. Prepare the master mix for PCR amplification (for eight tubes of 50 μ l each):

50 μ l 10 \times PCR buffer without MgCl₂
4.0 μ l 25 mM dNTP mix
15 μ l 50 mM MgCl₂
426 μ l H₂O
1.0 μ l 100 μ M FDV-T
1.0 μ l 100 μ M RDV-T
2.0 μ l 5 U/ μ l Platinum *Taq* DNA polymerase.

33. Split the master mix into eight tubes of 50 μ l each, and add 0.5 μ l FDV/RDV-adapted library material to each tube.
34. Carry out PCR using the following amplification cycles:

Initial step:	2 min	94°C	(denaturation)
12 cycles:	30 sec	94°C	(denaturation)
	30 sec	55°C	(annealing)
	90 sec	72°C	(extension).

The number of cycles is kept low to minimize PCR amplification bias. Due to suppression PCR effects (Siebert et al., 1995), FDV/FDV- and RDV/RDV-adapted molecules are not amplified.

35. Purify DNA using Qiaquick columns according to the manufacturer's instructions for PCR cleanup.
36. Purify the amplified 135-bp FDV-RDV paired-insert library as in steps 9 and 10, resuspending the precipitated DNA in 10 μ l TE buffer, pH 8.0.

When excising the 135-bp library band, be careful to avoid contamination by marker DNA, if used.

37. Quantitate the final DNA library by electrophoresis (step 11).

The concentration should be \sim 2 ng/ μ l.

EMULSION PCR OF PAIRED-TAG LIBRARY ON MICROBEADS

Emulsion PCR (ePCR) is used to generate clonally amplified DNA fragments suitable for polony sequencing. First, a biotinylated PCR primer is immobilized on streptavidin-coated beads. Next, billions of microreactors are formed by the emulsification of a PCR mix in mineral oil (Dressman et al., 2003). When one bead (bearing an immobilized PCR primer) and one template molecule are trapped in a compartment, amplification will proceed and result in tens of thousands of copies of the template molecule attached to the bead. The result is a complex mixture of clonally amplified templates (polonies) anchored to beads. Each bead bears numerous copies of the same template, but different beads bear different templates.

Materials

MyOne C1 1- μ m paramagnetic, streptavidin-coated beads (Dynal)

Bind and wash (B&W) buffer (see recipe)

Oligonucleotides:

1 mM PR1-F-2BIO: 5'-dual-biotin-CCACTACGCCTCCGCTTTCCTCTCTAT
GGGCAGTCGGTGAT-3'

2 mM PR1-R: 5'-CTGCCCCGGGTTTCCTCATTCTCT-3'

10 μ M PR1-3LF: 5'-CCTCTCTATGGGCAGTCGGTGAT-3'

TE buffer, pH 8.0 (APPENDIX 2)

Light mineral oil (Sigma)

10% Span 80 (see recipe)

Tween 80

Triton X-100

10 \times PCR buffer without MgCl₂ (Invitrogen)

50 mM MgCl₂

25 mM dNTP mix (25 mM each nucleotide)

5 U/ μ l Platinum Taq DNA polymerase (Invitrogen)

Template DNA (library DNA at appropriate concentration; see Basic Protocol 1 and Support Protocol)

BASIC PROTOCOL 2

DNA Sequencing

7.8.7

NXS buffer (see recipe)
0.1 M NaOH
1.5-ml microcentrifuge tubes
Magnetic particle concentrator (MPC; Dynal)
2-ml cryogenic vials (Corning no. 430661)
Stir bar, flea-size (VWR no. 58948-353)
Magnetic stirrer (closed-loop; VWR)
200- μ l eight-tube PCR strips
Thermal cycler

Bind forward PCR primer to microbeads

1. To 30 μ l MyOne beads in a 1.5-ml microcentrifuge tube, add 30 μ l B&W buffer.
2. Mix and then remove all liquid using a magnetic particle concentrator (MPC).
Vortex to mix in all steps except those where enzyme activity is required. In those cases mix gently by flicking or rotating the tube.
3. Wash two times in 60 μ l B&W buffer using an MPC, and resuspend in 60 μ l B&W buffer.
4. Add 1.2 μ l PR1-F-2BIO, mix, and incubate 20 min at room temperature.
Mix periodically by pipetting to prevent settling of beads.
PR1-F-2BIO is synthesized with a dual biotin at the 5' end to tightly bind the streptavidin-coated microbeads. Dual biotin is necessary to ensure the template remains on the bead for the duration of the sequencing run.
5. Remove all liquid using an MPC, and then wash twice in 100 μ l B&W buffer and once in 100 μ l TE buffer, pH 8.0.
6. Remove all liquid using an MPC and resuspend in 60 μ l TE buffer.
Concentration should now be $\sim 5 \times 10^9$ beads/ml.

Perform emulsion PCR reaction

7. Prepare the oil phase by assembling six 1.5-ml microcentrifuge tubes containing the following:
545 μ l light mineral oil
450 μ l 10% Span 80
4.0 μ l Tween 80
0.5 μ l Triton X-100.

Vortex each tube individually for at least 20 sec to thoroughly mix all components.
Watch the pipet tips carefully to ensure that viscous reagents are delivered into the tube. Use reverse pipetting (see http://www.rainin.com/pdf/rainin_classic_manual.pdf) for all oil-phase components.
8. Prepare the aqueous phase by combining the following:
96 μ l 10 \times PCR buffer without MgCl₂
359 μ l 50 mM MgCl₂
135 μ l 25 mM dNTP mix
12.0 μ l 2 mM PR1-R
4.8 μ l 10 μ M PR1-3LF
60 μ l microbeads from step 6
246 μ l H₂O
54 μ l 5 U/ μ l Platinum *Taq* DNA polymerase
1.0 μ l template DNA.

9. Place twelve 2-ml cryogenic vials in the centers of three magnetic stirrers set to 1400 rpm.
10. Place a flea-size stir bar in each vial.
11. Add 400 μ l oil phase to each vial.
12. Add 75 μ l aqueous phase, dropwise over 1 min, to each vial.

Be sure each drop of liquid hits the bottom of the tube, not the walls.

13. Stir for 30 min using a magnetic stirrer.
14. Pipet the contents of each cryogenic vial into a separate eight-tube PCR strip (50 μ l per tube) for a total of twelve strips, or 96 tubes.

Aspirate and dispense slowly to minimize the amount of emulsion that sticks to the walls of the pipet tip.

15. Carry out PCR using the following amplification cycles:

Initial step:	2 min	94°C	(denaturation)
120 cycles:	15 sec	94°C	(denaturation)
	30 sec	57°C	(annealing)
	75 sec	70°C	(extension)
Finish:	2 min	72°C	
Hold:	indefinitely	4°C.	

16. Pool contents of each strip of eight 200- μ l tubes into a separate 1.5-ml microcentrifuge tube (twelve tubes). Immediately proceed with recovery.

Recover beads from emulsion

17. Add 800 μ l NXS buffer to each tube.
 18. Vortex each tube individually at maximum speed for at least 20 sec.
 19. Centrifuge all tubes 90 sec at 16,110 \times g, room temperature.
 20. Aspirate all but 100 μ l supernatant (being sure to remove all of the top oil phase).
 21. Repeat steps 17 to 20 two more times.
 22. Pellet beads in tubes using an MPC, and aspirate all liquid.
 23. Resuspend each pellet in 25 μ l TE buffer, pH 8.0, and transfer all to a single new tube. Mix by pipetting up and down.
- All twelve tubes should be pooled into this one new tube. Try to minimize the amount of oil and detergent carried over into the new tube from the walls of the old tube.*
24. Pellet the beads using an MPC, aspirate all liquid, and add 250 μ l TE buffer, pH 8.0. Vortex briefly, then pulse centrifuge to pull the liquid into the bottom of the tube. Repeat two more times.
 25. Pellet beads using an MPC, and aspirate all liquid.
 26. Remove the nonbiotinylated strand from the beads by incubating in 100 μ l of 0.1 M NaOH for 20 min at room temperature. Periodically mix by pipetting to prevent beads from settling.
 27. Pellet beads using an MPC, and remove all liquid.
 28. Wash once with 100 μ l 0.1 M NaOH and three times with 150 μ l TE buffer. Mix each wash by vortexing to ensure removal of residual NaOH.
 29. Pellet beads using an MPC, aspirate all liquid, and resuspend in 60 μ l TE buffer.

ENRICHMENT FOR AMPLICON-BEARING BEADS

The result of ePCR is a population of beads, most (~80%) of which do not bear an amplified template. In order to maximize the number of amplicon-bearing beads on the sequencing array, an enrichment step is required. Fluorescence-activated cell sorting can be employed to select for these beads by hybridizing a fluorescently labeled probe specific to amplicons. Alternately, the centrifugation technique presented below can be used for more rapid and cost-effective enrichment.

In this protocol, the dense paramagnetic ePCR beads are incubated with low-density polystyrene beads bearing an oligonucleotide complementary to a common amplicon sequence. Amplicon-bearing beads hybridize to these capture beads, while non-amplicon-bearing beads do not. The density differential between captured and uncaptured ePCR beads is exploited for enrichment by centrifugation through a viscous 60% glycerol solution; the former remain in the supernatant while the latter form a pellet.

Materials

Spherotech particles (3- μ m streptavidin-coated polystyrene particles; Spherotech no. SVP-30-5)

Bind and wash (B&W) buffer (see recipe)

1 mM PR1-BIOXL:

5'-biotinTEG-CGTACCCCGCTTGGTCTTTCTCCCGTACCCCGCTTGG
TCTTTCTCCCTGCCCGGGTTCCTCATTCTCT-3'

TE buffer, pH 8.0 (*APPENDIX 2*)

ePCR beads (Basic Protocol 2)

Glycerol

0.1 M NaOH

Magnetic particle concentrator (MPC; Dynal)

1.5-ml microcentrifuge tubes

Bind capture oligonucleotide to capture beads

1. Centrifuge 400 μ l Spherotech particles for 30 sec at $16,110 \times g$, room temperature.
2. Remove all liquid, and wash pellet with 400 μ l B&W buffer.
3. Centrifuge for 30 sec at $16,110 \times g$.
4. Remove all liquid, and resuspend in 400 μ l B&W buffer.
5. Add 8 μ l of 1 mM PR1-BIOXL and incubate at room temperature for 20 min.
6. Wash three times in 800 μ l TE buffer, pH 8.0, pelleting each time by centrifugation. Resuspend final pellet in 80 μ l B&W buffer.

Anneal capture beads to emulsion PCR beads

7. Add 60 μ l ePCR beads to 80 μ l capture beads and incubate at 56°C for 10 min.
8. During the incubation, prepare 1 ml of fresh 60% (v/v) glycerol in H₂O. Vortex to mix thoroughly.
Must be made fresh each day.
9. Add 150 μ l of 60% glycerol to each of four 1.5-ml microcentrifuge tubes.
10. Layer 35 μ l bead mixture from step 7 onto the top of the glycerol solution in each of the four tubes.

Be careful to pipet gently so the bead solution does not mix with glycerol solution.

Separate bead populations by centrifugation

11. Centrifuge tubes for 1 min at $16,110 \times g$.

The amplicon-bearing ePCR beads are hybridized to the complementary sequences carried on the low-density polystyrene capture beads and will remain in the supernatant; the ePCR beads without amplified sequences will form a pellet in the bottom of the tube.

12. Transfer all the supernatant to four new microcentrifuge tubes, being careful not to disturb the pellets.

If necessary, wash the walls of the tubes with the supernatant to collect any beads bound there.

13. To each recovered supernatant tube, add 1 ml H₂O and vortex until thoroughly mixed.

14. Centrifuge for 3 min at $16,110 \times g$, room temperature, to pellet all beads.

15. Aspirate, being careful to not disturb the pellet, and then place tubes on an MPC to hold the pellet firmly in place. Remove any remaining liquid.

16. Separate ePCR beads from capture beads by adding 120 μ l of 0.1 M NaOH and incubating 10 min at room temperature to denature the hybridized DNA strands.

17. Using the MPC, collect the amplicon-bearing ePCR beads into a pellet. Remove and discard all liquid.

The liquid contains the released capture beads.

18. Wash the pellet twice by adding 120 μ l of 0.1 M NaOH, vortexing briefly, centrifuging briefly to pull the liquid to the bottom of the tube, and collecting the pellet using the MPC.

19. Resuspend in 6.5 μ l TE buffer, pH 8.0.

CASTING A POLONY BEAD ARRAY

The amplified, enriched beads are cast in a two-dimensional array to fix their positions during the subsequent sequencing reactions by mixing with acrylamide and pouring into a shallow mold formed by a Teflon-masked microscope slide. The coverslip placed on top of the gel is treated with Bind Silane, which covalently attaches the gel to the glass coverslip. The acrylamide gel mixture is formulated such that it takes 30 min or more to polymerize, during which time the beads settle into a single monolayer before their positions become fixed. This allows efficient imaging of all beads in the gel, since they all reside in the same focal plane of the optics used for imaging.

Once polymerized, the gel is approximately 30 microns thick, while the beads are only 1 micron in diameter. To ensure the DNA-bearing beads are accessible to reagents for enzymatic sequencing reactions, the gel is cast such that the beads will be fixed at the exposed surface (i.e., opposite the side attached to the Bind Silane-treated glass).

Materials

1% Triton X-100 in H₂O

Glacial acetic acid

Bind Silane (Promega)

100% ethanol

Enriched ePCR beads in TE buffer, pH 8.0 (Basic Protocol 3)

40% acrylamide/bisacrylamide (19:1) solution

5% TEMED

0.5% ammonium persulfate (APS)

Wash 1 (see recipe)

BASIC PROTOCOL 4

DNA Sequencing

7.8.11

Round coverslips, 40 mm diameter, no. 1.5 (Bioprotechs no. 40-1313-0319)
1000-ml plastic beaker
Shaker
Vacuum desiccator
Teflon-masked microscope slides (Erie Scientific, no. ER-203W)

Treat coverslips with Bind Silane

1. Wash round coverslips in 1% Triton X-100 for 20 min with gentle shaking.
2. Prepare silane solution in a 1000-ml plastic beaker:
 - 500 ml H₂O
 - 110 μ l glacial acetic acid
 - 2 ml Bind Silane.

Stir for 15 min until Bind Silane droplets have completely disappeared.

3. Immerse coverslip in silane solution and incubate 1 hr at room temperature with gentle shaking.
4. Wash thoroughly with H₂O, then rinse with 100% ethanol.
5. Allow to dry completely in vacuum desiccator. Store coverslips up to 1 month under vacuum.

Cast enriched ePCR beads

6. Prepare bead/acrylamide mixture:
 - 6.5 μ l enriched ePCR beads in TE buffer
 - 1.25 μ l 40% acrylamide/bisacrylamide solution (19:1)
 - 0.5 μ l 5% TEMED
 - 0.5 μ l H₂O.

Mix by pipetting, then add 0.75 μ l of 0.5% APS.

7. Immediately pipet 9.5 μ l onto a Teflon-masked microscope slide and cover with a silane-coated coverslip. Allow to polymerize for 45 min.

The coverslip should spread the liquid across the entire surface of the circular well on the Teflon-masked slide. During polymerization, the beads will settle to the surface of the gel closest to the microscope slide.

8. Invert the slide/coverslip stack and remove microscope slide from gel by gently lifting with a razor blade near the edge of the coverslip.

The gel should remain attached to the coverslip.

9. Immerse the coverslip, with polony bead array attached, in wash 1 until ready for use, up to ~1 month at 4°C.

BASIC PROTOCOL 5

DNA SEQUENCING BY LIGATION WITH DEGENERATE FLUORESCENT NONAMERS

An anchor primer is first hybridized to a known sequence, or priming site, within the single-stranded template. This primer is synthesized with deoxyuridine in place of thymidine. A ligation reaction is then performed with a pool of fully degenerate, fluorescently labeled nonanucleotides. During the ligation reaction, each bead becomes tagged with a fluorophore that indicates the identity of the base present at the position being interrogated. After imaging the array to determine which fluorophore labels each bead, the anchor-primer/fluorescent-nonamer complex is removed from the template by reacting

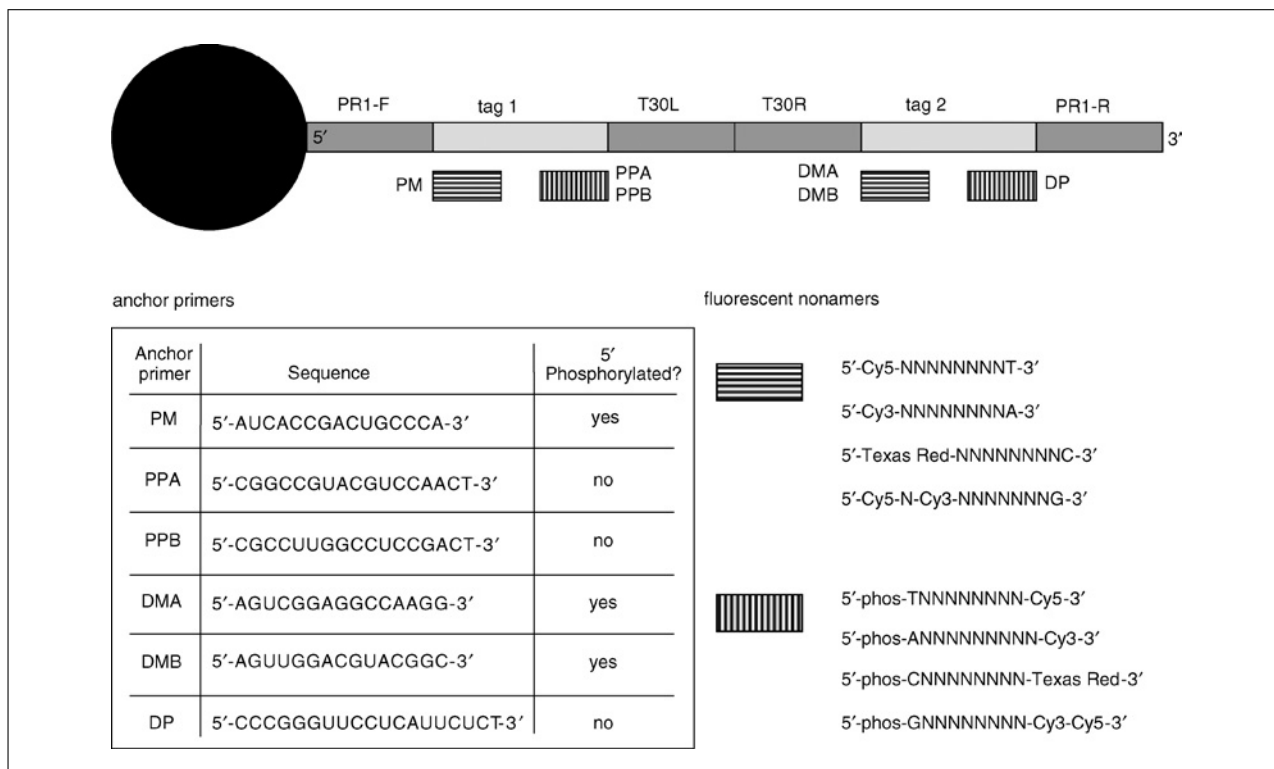


Figure 7.8.2 Schematic of the 135-bp bead-bound template and sequences of oligonucleotides used in sequencing reactions (Basic Protocol 5). For each reaction, the combination of anchor primer(s) and fluorescent nonamers is specified by the position in the tag being sequenced. All anchor primers should be synthesized with deoxyuridine to allow enzymatic removal of the primers. Fluorescent nonamers should be synthesized by Integrated DNA Technologies (IDT) with fully degenerate positions hand mixed to ensure equal ratios of each nucleotide at each position. The T30 sequence is shown as T30L and T30R to emphasize that the sequencing primers do not bind the full length of T30. Abbreviations: DMA and DMB, distal minus A and B strands; DP, distal plus; PM, proximal minus; PPA and PPB, proximal plus A and B strands; PR1-F and PR1-R, forward and reverse PR1 primers.

with the uracil-specific excision reagent (USER) enzyme. The USER enzyme excises all deoxyuridines from the anchor primer, resulting in short fragments that dissociate at moderate temperatures ($<60^{\circ}\text{C}$) due to their low melting points (T_m), and leaving the bead-bound template ready for another position to be interrogated.

The combination of degenerate nonamer pool and anchor primer specifies which position in the template the reaction will interrogate (see Fig. 7.8.2). Whether the anchor primer or the nonanucleotides should be synthesized with a 5' phosphate is dependent on which position in the template is being queried with a given combination of anchor primer and nonamer. When sequencing the 5' side of a tag, the anchor is phosphorylated, and when sequencing the 3' side, the nonamers are phosphorylated.

At the start of a sequencing run, all 3' ends of unextended forward ePCR primers and templates are capped by the addition of dideoxynucleotides using terminal deoxytransferase. This capping reaction is necessary to prevent nonspecific ligation of labeled nonamers to bead-bound ePCR primers and templates.

This protocol describes the ligation and stripping steps required to query the base at a specific position in an iterative fashion. When using the autosampler and flow cell system, multiple solutions can be prepared for one or several cycles and used with a computer program preset with the appropriate times, temperatures, and volumes indicated in the protocol steps. Imaging and analysis procedures are discussed in Critical Parameters and Troubleshooting.

Materials

Bead array with template DNA (Basic Protocol 4)
5× tailing buffer (Invitrogen)
1.25 mM ddNTP mix (1.25 mM each dideoxynucleotide)
15 U/μl terminal deoxytransferase (recombinant; Invitrogen)
Wash 1 (see recipe)
6× SSPE (*APPENDIX 2*)/0.01% Triton X-100
Anchor primer (see Fig. 7.8.2)
2000 U/μl T4 DNA ligase and 10× buffer (NEB)
100 μM degenerate fluorescent nonamer mix (Integrated DNA Technologies)
TE buffer, pH 8.0 (*APPENDIX 2*)
1 U/μl USER enzyme mix (NEB)
Automated fluorescent microscope with flow cell (see Table 7.8.1 for parts list)
Autosampler (see Table 7.8.1)
1-ml disposable syringe
0.45-μm pore size, 4-mm cellulose acetate syringe filter (VWR)

Perform dideoxynucleotide capping reaction

1. Prepare the capping reaction mix:

114 μl H₂O
30 μl 5× tailing buffer
4 μl 1.25 mM ddNTP mix
2 μl 15 U/μl terminal deoxytransferase.

2. Place the bead array with template DNA on a coverslip into the flow cell mounted on the microscope stage. Prime fluid lines with wash 1.
3. Add the reaction mix to the appropriate well of the autosampler. Incubate 60 min at 35°C.

Hybridize anchor primer to array

4. Prepare the hybridization mix by combining:

150 μl 6× SSPE with 0.01% Triton X-100
1.5 μl 1 mM anchor primer.

5. Add hybridization mix to the appropriate well of the autosampler. Incubate 5 min at 56°C, then 2 min at 42°C.

Ligate fluorescent nonamers to visualize complexes

6. Prepare ligation reaction on ice:

15 μl 10× T4 DNA ligase buffer
12.0 μl 100 μM degenerate fluorescent nonamer mix
3.0 μl 2000 U/μl T4 DNA ligase
120 μl H₂O.

When preparing multiple ligation reaction mixtures (for sequential cycles) at once, be sure they are consumed within 6 hr, since T4 DNA ligase is unstable once mixed with reaction buffer. In addition, only one mix containing 5'-phosphorylated nonamers should be prepared, and it should be used in the first sequencing cycle performed.

7. Aspirate the solution into a 1-ml syringe, attach a 0.45- μm filter to the Luer syringe adapter, and pass the contents through the filter into the appropriate well of the autosampler. Incubate 30 min at 32°C.

The nonamer solutions contain large aggregates of fluorescent dye. Filtering is essential to ensure these aggregates are not introduced into the flow cell, since they will gradually accumulate by sticking to the gel, interfere with imaging and data extraction, and poison subsequent reactions.

Incubation temperature and duration can be optimized for variations in signal. If signal is low, a slightly longer time or higher temperature can be used; if signal is very high, cycle time can be decreased by decreasing incubation time appropriately. Although increased concentrations of nonamer may also drive the reaction to completion faster, it should be noted that cost will increase and fidelity may be affected.

8. Image using fluorescence microscopy to determine the nonamer ligated to each bead and a single position of one tag.

The automated microscope used for polony sequencing includes a number of add-on components both for imaging (e.g., motorized stage, piezo objective positioner, illumination wheels) and for fluidics (e.g., flow cell assembly, autosampler, digital I/O modules). Complete details are beyond the scope of this unit. All major components are listed in Table 7.8.1, and a discussion is provided in Critical Parameters and Troubleshooting. More detailed information on assembly (including fittings, tubing, and more) can be found at <http://arep.med.harvard.edu/Polonator>.

Strip fluorescent complexes from the beads

9. Prepare stripping reaction mix by combining:

300 μl TE buffer, pH 8.0
6 μl USER enzyme.

Stripping is performed by reacting the array twice with the stripping solution, hence the volume is double that of a standard reaction. This can be adjusted according to the specific performance observed. If stripping is observed to be complete after one round of reaction and heating, the second stripping can be eliminated.

10. Add wash 1 and the stripping mix to the appropriate wells of the autosampler. Perform stripping protocol as follows:

incubate at 37°C for 5 min
incubate at 56°C for 1 min
wash
incubate at 37°C for 5 min
wash.

The autosampler software for stripping implements this protocol by default.

Query next position

11. Repeat steps 4 to 8 using a different combination of either anchor primer or fluorescent nonamer or both for up to 26 cycles, stripping the fluorescent complexes as in steps 9 and 10 between sequencing cycles.

The read length limit of T4 DNA ligase is 6 or 7 bp from the ligation junction, for a total of 13 bp per tag.

TITRATION OF TEMPLATE FOR CLONAL AMPLIFICATION BY ePCR

Before performing emulsion PCR on a new sample for the first time, the appropriate amount of template must be determined empirically. Since emulsion PCR generates clonally amplified beads by a limiting dilution of template, one can be sure most beads are clonal by finding the template concentration at which approximately 15% to 20% of beads bear an amplicon. Generally, the correct amount of template is somewhere near 1 fmol per reaction. Sample four different concentrations and, if necessary, interpolate to determine the final amount. For example, given that the concentration of the library material is known, amounts of 5 fmol, 1 fmol, 0.2 fmol, and 0.05 fmol per reaction should be tried. Somewhere in this range, a linear increase in amplified beads will be observed with a linear increase in template.

After amplification of the four template samples, bead arrays are cast on microscope slides instead of coverslips (see Basic Protocol 4) for ease of handling. Amplified beads are identified by hybridizing a fluorescently labeled probe complementary to a sequence present in the central (T30) sequence. The percent of amplified beads can then be determined by fluorescence microscopy.

Additional Materials (also see Basic Protocols 2 and 4)

Fluorescently labeled probes:

100 μ M T30-P2-Cy5-A: 5'-Cy5-AGUUGGACGUACGGCC-3'

100 μ M T30-P2-Cy5-B: 5'-Cy5-AGUCGGAGGUCAAGGC-3'

Heat block or slide thermal cycler

Microscope for bright-field and fluorescence microscopy, with CCD camera and filter set for Cy5

Amplify template dilutions

1. Prepare four emulsion PCR reactions as in Basic Protocol 2, using four different amounts of template DNA in the aqueous phase (step 8). Scale down each reaction twelve-fold, so that each template amount is prepared in a single cryogenic vial (step 9) and amplified in a single eight-tube strip (step 14).

This will result in four strips of beads, one per template amount tested. Be sure to scale down the amount of DNA used because each reaction is one-twelfth the size of a standard emulsion PCR reaction.

2. When recovering the beads from the emulsion, be sure to not combine the four strips as in step 23.

The result will be 4 tubes of beads.

3. Resuspend each tube in 6.5 μ l TE buffer, pH 8.0, and proceed to labeling of amplicons without performing enrichment.

Cast bead arrays

4. Cast each sample as a separate bead array on a Teflon-coated microscope slide. Use the same procedure as in Basic Protocol 4, but treat the microscope slide with Bind Silane instead of the coverslip, and invert the stack during polymerization so that the beads are immobilized next to the coverslip instead of the slide.

Hybridize fluorescent probe

5. Prepare the hybridization mixture:

100 μ l of 6 \times SSPE with 0.01% Triton X-100

1.0 μ l of 100 μ M T30-P2-Cy5-A (5'-Cy5-AGUUGGACGUACGGCC-3')

1.0 μ l of 100 μ M T30-P2-Cy5-B (5'-Cy5-AGUCGGAGGUCAAGGC-3')

6. Dry the Teflon surface of the slide with a Kimwipe, being careful to not touch the gel.
7. Pipet the hybridization mixture onto the gel.
The mixture will stay contained on top of the gel if the slide was dried properly.
8. Place the slide face up on a heat block or slide thermal cycler and incubate at 56°C for 5 min.
9. Wash for 5 min in fresh wash 1 to remove free fluorescent probe.

Identify amplified beads

10. Place an array under a fluorescence microscope and acquire an image under transmitted bright-field illumination.

Magnification of 20× is sufficient for the pixel size of most CCD cameras. Long working distance optics with relatively low NA are sufficient when imaging bead arrays at low density (~20,000 beads per field of view). When higher densities are desired (~150,000 to 200,000 per field of view) high-NA dry optics should be used.

Depending on the particular optics used, the beads will appear as black circles or as white points surrounded by dark halos.

11. Acquire a fluorescence image at the same position using appropriate excitation and emission filters for Cy5.
12. Superimpose the two images in false color and determine the percentage of beads that are labeled.
13. Calculate the amount of library material to use in each ePCR based on the percentage of beads that bear labeled amplicons.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Bind and wash buffer

5 mM Tris·Cl, pH 7.5 (APPENDIX 2)
0.5 mM EDTA, pH 8.0 (APPENDIX 2)
1.0 M NaCl
Store up to 6 months at room temperature

NXS buffer

10 mM Tris·Cl, pH 7.5 (APPENDIX 2)
1 mM EDTA, pH 8.0 (APPENDIX 2)
100 mM NaCl
1% Triton X-100
1% (w/v) SDS
Mix ingredients thoroughly
Store up to 6 months at room temperature

Span 80, 10% (v/v)

10% (v/v) Span 80 in light mineral oil

Measure ingredients by positive-displacement syringe or pipet. Vortex to mix thoroughly and incubate under vacuum to remove air bubbles. Store up to 1 week at room temperature.

Wash 1

10 mM Tris-Cl, pH 7.5 (APPENDIX 2)
50 mM KCl
2 mM EDTA, pH 8.0 (APPENDIX 2)
0.01% (v/v) Triton X-100
Mix ingredients thoroughly
Store up to 6 months at room temperature

COMMENTARY

Background Information

Since the dideoxy Sanger method (Sanger et al., 1977) was introduced almost 30 years ago, DNA sequencing has become a standard (albeit expensive) tool in the biologist's repertoire. With the list of reference genome sequences growing, attention is shifting toward the study of how a reference genome differs from derivative experimental genomes of interest (e.g., adaptive evolution of bacterial strains, drug resistance/sensitivity). In many cases, to fully enable these studies, the cost per quality base of DNA sequence must be lower than that obtained with state-of-the-art Sanger sequencing.

Polony sequencing aims to address this need by leveraging highly multiplex processes during both sample preparation and sequencing itself (Shendure et al., 2005). This allows an effective decrease in reagent volumes per reaction from the microliter scale to the femtoliter scale, while also dramatically reducing data acquisition time per base. These reductions in time and materials translate to a lower cost per base and, as an added bonus, higher throughput. As a cyclic array method (for a more thorough discussion, see Shendure et al., 2004), sequence is generated base-by-base through successive cycles of ligation and imaging. Each cycle of ligation and imaging sequences a single base in the template. Each 17- to 18-bp genomic tag in the sequencing template yields 13 bp of sequence when queried by T4 DNA ligase (with the central 4 to 5 bp of the tag remaining unsequenced). The result of a sequencing run is one or more million discontinuous reads of 26 bp each. While each is much shorter than a Sanger read, for many applications this length is sufficient.

Polony sequencing derives its name from the sequencing of DNA molecules that are clonally amplified in or on a solid matrix. Each amplified clone represents a polymerase colony or "polony." Polonies can be amplified from template molecules immobilized in a polyacrylamide matrix (Mitra and Church, 1999) or on beads (Diehl et al., 2006). In this

unit, polonies are generated from a genomic paired-tag library on magnetic beads and are then enriched by hybridization to a common sequence. The enriched beads are cast in an array in a polyacrylamide matrix, where they are sequenced.

The focus of this unit is genomic resequencing. By using a reference genome as a scaffold onto which reads are placed and interrogated for differences, the genome in question can be effectively resequenced using much shorter reads than if a reference did not exist. Additionally, paired tags allow for placement of reads too short to otherwise span repetitive regions. All that is required is that a read-pair be unique enough, given its intertag distance distribution, that it places to only one location on the reference genome. Shendure et al. (2005) demonstrated that 90% coverage of the *E. coli* genome is feasible with paired 13-bp reads having an intertag distance distribution of 1 kb \pm 250 bp. In addition to genomic resequencing, polony technology in general has been used for haplotyping (Mitra et al., 2003; Turner et al., 2006; Zhang et al., 2006) and splicing analysis (Zhu et al., 2003).

Critical Parameters and Troubleshooting

Paired-tag library

The key to successful execution of the library protocol is to achieve yields at each step at least as high as indicated in the protocols. In particular, gel purifications, ethanol precipitations, and column purifications must be reproducibly efficient. DNA quantification should be performed as indicated in the protocol, because NanoDrop readings are not reliable in the authors' hands following ethanol precipitation.

The most critical step in the paired-tag library protocol is the circularization of the genomic fragment around the synthetic linker-insert. Shendure et al. (2005) found this to be the rate-limiting step in terms of the number of unique sequencing templates that were

generated in the library construction. This library protocol is inefficient from a DNA mass standpoint because it represents the product of a number of inefficient individual steps. Additionally, the release of a pair of 18-bp tags from each fragment following circularization results in a nominal mass loss of ~96% for 1000-bp fragments (linearly worse for longer fragments). This loss necessitates the hyperbranched rolling-circle amplification performed after circularization.

Emulsion PCR

When performing emulsion PCR, it is critical not to deviate in any way from the protocol provided until it works to the point of producing clonal, strongly amplified beads from stable emulsions. Once reliable, the method can be used as a control in further optimizations. Inconsistent stability from one batch to the next is usually due to either inconsistent measuring of reagents (especially detergents) or using oil phase components (e.g., 10% Span 80) that were not freshly prepared.

Reverse amplification primer (nonbiotinylated) is present at a very high concentration in the reaction. This primer must be HPLC-purified to prevent formation of primer-dimer product on beads (which may also bear library molecule amplicons). Although primer-dimer formation is also a function of sequence, the PR1 sequences yield especially strong ePCR amplicons relative to several other primer pairs.

The interested reader is directed to excellent papers (Dressman et al., 2003; Diehl et al., 2005, 2006; Li et al., 2006) which cover the technique of emulsion PCR on beads in more detail.

Casting a polony bead array

When preparing Bind Silane-treated slides or coverslips, be sure to use Bind Silane which has not reached its expiration date. Use Erie Teflon-masked slides as the gel molds. Their Teflon coating is unusually hydrophobic and uniform in surface flatness, both of which are necessary to keep the gel from leaking out the top of the circular well before polymerization.

Be certain that no bubbles are present in the gel after it has been spread between the slide and the coverslip. A small bubble noticed in the gel after lowering the coverslip onto the slide can be removed by gently lifting the edge of the coverslip slightly with the edge of a razor blade to move the bubble to the edge of the gel and slowly lowering the coverslip back onto the slide. If this process is performed carefully, the gel will not spread outside the Teflon well.

A jig can be used to hold the coverslip in the correct position relative to the Teflon well of the microscope slide. A 0.5-in.-thick block of polyethylene can be machined with a shallow pocket that will constrain movement of both the slide and the coverslip. This ensures that the array always forms in the center of the coverslip, and thus is always properly placed in the center of the flow cell chamber.

Software, hardware, and sequence analysis

To perform polony sequencing using the protocols presented here, repetitive cycles of sequencing biochemistry and imaging are performed until the desired read length is reached (e.g., 26 cycles for 26 bp). Then, fluorescence data must be extracted from the ~50,000 2-MB images generated by the instrument, normalized, and converted to sequence. In the case of genomic resequencing, these short reads must be placed onto (or matched to) a reference genome to determine which positions differ. The process is summarized below.

The polony sequencing instrument is an inverted epifluorescence microscope which has motorized, encoded *x*, *y*, and *z* axes, as well as motorized fluorescence excitation and emission wavelength selection. The motorized features are controlled by a series of C++ binary files (available with source codes at <http://arep.med.harvard.edu/Polonator>). Attached to the microscope is the fluidics/temperature control system, which performs the biochemistry in the flow cell attached to the automated stage. It is controlled by a series of Matlab scripts, also available at <http://arep.med.harvard.edu/Polonator>.

A list of the major components necessary to build a duplicate of the instrument used by the authors is shown in Table 7.8.1; a full parts list (including fittings, tubing, and details for assembly) is available at <http://arep.med.harvard.edu/Polonator>. In building an instrument, the recommendation is to make no substitutions for parts, even when the particular components seem to have similar performance characteristics. Subtle features present in each component listed are in many cases necessary for the system to function as expected. For example, Nikon is the only microscope manufacturer that manufactures a 20× plan apochromatic objective with both a phase contrast ring (necessary for imaging densely packed bead arrays) and a 1.0-mm working distance (necessary for adequate clearance around the flow cell).

Table 7.8.1 Parts List for Automated Microscope Used in Polony Sequencing

Quantity	Part no.	Manufacturer	Description
1	86022-SPR	Chroma	Cy3/Cy5 polychroic filter w/matched individual exciters/emitters
1	41004*	Chroma	Custom TxRed filter set (S572/23x, Q595LP, S630/60m)
2	CFL	Chroma	Empty filter cube for TE2000
1	91002-02	Hamamatsu	Camera, 1000x1000, electron multiplied, 30 fps readout
1	500-H117E1T4	Prior	Stage, motorized encoded <i>x,y</i> for TE2000; 1-mm screw pitch
1	500-H30XY-E	Prior	Controller for <i>x,y</i> stage w/encoder connectors
1	500-CS152V2	Prior	Joystick for manual <i>x,y</i> control
1	LB-LS/OF30	Sutter	Lambda light source, 300 W xenon
1	LLG	Sutter	Liquid light guide for LS, 2 meters
1	LG-N27	Sutter	Adapter to attach light guide to TE2000 (no epi arm)
1	LB10-3	Sutter	Control unit for 3 wheels and 2 shutters
1	LB10-NW1Q	Sutter	10-position wheel for 25-mm filters with smart shutter
1	LB10-NEW	Sutter	10-position wheel for 25-mm filters - setup for emission
1	10-N27-EM	Sutter	Adapter to attach wheel to side port of TE2000
1	IQ25-SA	Sutter	Smart shutter for transmitted light
1	10-N27-EC	Sutter	Adapter to attach transmitted light shutter to Nikon HMX-II
1	O661301	Sutter	Extra xenon bulb, 300 W
1			Pifoc 400- μ m travel piezo positioner w/controller
1	MEA51001	Nikon	TE2000-E2 inverted microscope
1	MEF55010	Nikon	T-HUBC hub controller
1	MEF55000	Nikon	T-RP remote control pad
1	MPF52061	Nikon	Universal power supply, 110-240V
1	MAE15001	Nikon	Lamphouse HMX2
1	MEE59900	Nikon	T-DH dia-illuminator, 100W
1	MAE15002	Nikon	Collector lens for QH100
1	MBF13240	Nikon	Lamp socket F/HMX 100W QI
1	MXA20425	Nikon	FXA-HMX2 adpt f/halogen
1	84125	Nikon	Microphot-FXA L.L. 12V-100W BU
1	91141-IN	Nikon	T-BP R100 optical path prism
1	MEF42252	Nikon	TE2-PS 100W power supply

continued

Table 7.8.1 Parts List for Automated Microscope Used in Polony Sequencing, *continued*

Quantity	Part no.	Manufacturer	Description
2	79035	Nikon	Power cord
1	91155	Nikon	Null modem cable
1	MEP51300	Nikon	T-N6 sextuple nosepiece
1	MBN11710	Nikon	Filter 45mm NCB11
1	MEL3000	Nikon	Diaphot condenser turret
1	MEL36200	Nikon	Diaphot condenser Lwd lens
1	MEH41200	Nikon	Te-C Lwd Ph2 module
1	MEV51100	Nikon	T-FLMC motorized cassette holder
1	MRD30200	Nikon	Cfi plan apo dm 20× objective
1	060319-2-0303-20	Bioptechs	Temperature controller, no alarm
1	060319-2-03-30	Bioptechs	FCS2 chamber only, 30-mm aperture
1	060319-2-1242HG	Bioptechs	Low dead volume top, 3-in. (0.020-in. i.d.) stainless-steel tubing, drilled
1	060319-2-0049	Bioptechs	FCS2 connector assembly, modified for slide only
5	130119-5-292	Bioptechs	Specialty microaqueduct slide - bus bars, 2 holes, no grooves
20	1907-250	Bioptechs	Shiny gasket, harvard die 452458
1	40-1313-0319	Bioptechs	40-mm round coverslips, 250/pk
1	A	HMS machine shop	Fluidics bulkhead ^a
2	B	HMS machine shop	Flow cell support arms ^a
1	C	HMS machine shop	A/D I/O module ^a
1	D	HMS machine shop	Round coverslip caddy ^a
1	AL719-P	Alcott Chromatography	Harvard autosampler
1	3000XL	Cavro	Syringe pump: RS232 1/4-28 3-port, 500- μ l

^aDrawings and schematics available at <http://arep.med.harvard.edu/Polonator/2006>.

Neither of these specifications is advertised as important by most vendors, but for this particular application, both are indeed important.

At the start of the instrument run, a C++ binary file is executed on the computer attached to the microscope to determine the proper focal position for each raster location. In this way, during each subsequent cycle of imaging, the microscope can quickly acquire images from all 500 to 700 raster locations without the need to refocus.

At the conclusion of the sequencing run, images from each cycle must be aligned to each other such that bead pixels from one image correspond to the same bead-pixels in all other images from that same raster location. Bright-field images are then thresholded and used to define the image pixels which coincide with beads. The mean intensity of each bead in each image is computed and saved to disk. These fluorescence values are normalized and converted to four-dimensional unit vectors. When plotted in four-dimensional space, the

data for each cycle will naturally form four clusters corresponding to the four base identities (A, C, G, or T). Bases are called by assigning each bead to the base identity of the cluster to which it is nearest by Euclidean distance. The quality score of each call is the Euclidean distance from the bead to the center of its assigned cluster.

To determine which sequenced bases differ from bases in the reference genome, the polony reads must be placed on the reference sequence while allowing for mismatches. The output of this algorithm is the list of input polony reads annotated with the genomic position to which each places, if any. The critical constraint is that a given read only place to one location on the reference genome, such that bases in the experimental genome can be unambiguously assigned. Once a list of placements has been computed, the experimental genome sequence can be easily derived by sorting the reads by location and reporting the call at each location. When more than one base is called at a given location (e.g., from two different reads), quality scores of each raw base call can be used to resolve the ambiguity.

Anticipated Results

The raw data produced by each sequencing run is ~50,000 2-MB images. These are processed automatically by software available at <http://arep.med.harvard.edu/Polonator> to yield several million 26-bp reads, each annotated with quality scores for each raw base. However, the number of reads produced depends on the number of microbeads arrayed on the coverglass and the fraction of those beads that bear useable amplicons.

Time Considerations

Library construction takes ~1 week if long incubations are performed overnight. The protocol can be stopped, if necessary, after any purification step. Titration of template concentration for emulsion PCR takes 2 days. Emulsion PCR takes part of two days. In experienced hands, two plates of emulsion PCR can be prepared, recovered, and enriched in 2 days by one person. The sequencing run takes 3 hours to set up. Each sequencing cycle takes 1.75 hr, for a total of 49 hours if performed continuously.

Literature Cited

Diehl, F., Li, M., Dressman, D., He, Y., Shen, D., Szabo, S., Diaz, L.A. Jr., Goodman, S.N., David, K.A., Juhl, H., Kinzler, K.W., and Vogelstein, B. 2005. Detection and quantification of muta-

tions in the plasma of patients with colorectal tumors. *Proc. Natl. Acad. Sci. U.S.A.* 102:16368-16373.

Diehl, F., Li, M., He, Y., Kinzler, K.W., Vogelstein, B., and Dressman, D. 2006. BEAMing: Single-molecule PCR on microparticles in water-in-oil emulsions. *Nat. Methods* 3:551-559.

Dressman, D., Yan, H., Traverso, G., Kinzler, K.W., and Vogelstein, B. 2003. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc. Natl. Acad. Sci. U.S.A.* 100:8817-8822.

Li, M., Diehl, F., Dressman, D., Vogelstein, B., and Kinzler, K.W. 2006. BEAMing up for detection and quantification of rare sequence variants. *Nat. Methods* 3:95-97.

Mitra, R.D. and Church, G.M. 1999. In situ localized amplification and contact replication of many individual DNA molecules. *Nucleic Acids Res.* 27:e34.

Mitra, R.D., Butty, V.L., Shendure, J., Benjamin, R., Williams, B.R., Housman, D.E., and Church, G.M. 2003. Digital genotyping and haplotyping with polymerase colonies. *Proc. Natl. Acad. Sci. U.S.A.* 100:5926-5931.

Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74:5463-5467.

Shendure, J., Mitra, R.D., Varma, C., and Church, G.M. 2004. Advanced sequencing technologies: Methods and goals. *Nat. Rev. Genet.* 5:335-344.

Shendure, J., Porreca, G.J., Reppas, N.B., Lin, X., McCutcheon, J.P., Rosenbaum, A.M., Wang, M.D., Zhang, K., Mitra, R.D., and Church, G.M. 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309:1728-1732.

Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A., and Lukyanov, S.A. 1995. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* 23:1087-1088.

Turner, D.J., Shendure, J., Porreca, G., Church, G., Green, P., Tyler-Smith, C., and Hurler, M.E. 2006. Assaying chromosomal inversions by single-molecule haplotyping. *Nat. Methods* 3:439-445.

Zhang, K., Zhu, J., Shendure, J., Porreca, G.J., Aach, J.D., Mitra, R.D., and Church, G.M. 2006. Long-range polony haplotyping of individual human chromosome molecules. *Nat Genet.* 38:382-387.

Zhu, J., Shendure, J., Mitra, R.D., and Church, G.M. 2003. Single molecule profiling of alternative pre-mRNA splicing. *Science* 301:836-838.

Contributed by Gregory J. Porreca, Jay Shendure, and George M. Church
Harvard Medical School
Boston, Massachusetts

CHAPTER 8

Mutagenesis of Cloned DNA

INTRODUCTION

Genes and other genetic elements are frequently characterized by correlating specific changes in DNA sequence with effects on function. The classical approach to this problem has been to obtain mutations by selecting for organisms having new properties. The relevant wild-type and mutant genes can then be cloned and subjected to DNA sequence analysis. Although this approach has been and continues to be useful, it suffers from several disadvantages. First, the methods of mutagenesis severely constrain the kinds of mutations that are obtained. Second, since the entire organism is subjected to mutagenesis, mutations occurring in the gene of interest are relatively rare. Third, since mutations are identified by virtue of their phenotype, it is essentially impossible to obtain “mutants” that behave indistinguishably from the wild-type gene; these are particularly valuable for determining which parts of the gene are not important for function.

The advent of recombinant DNA technology has made it possible to reverse the procedure of classical mutagenesis. Mutations are first generated in cloned segments of DNA by using a variety of chemical and enzymatic methods. These methods can produce mutations at an extremely high frequency (approaching 100% in some cases), and essentially all possible mutations can be generated. Once generated, the mutant DNAs are subjected to DNA sequence analysis and then analyzed for the specific function of interest. In this way, mutations can be obtained in a systematic manner *without* regard to their phenotype. The end result is that the functions of a given region of DNA can be investigated in much more detail.

This chapter describes six basic protocols for altering the nucleotide sequence of cloned DNA segments. The first method, oligonucleotide-directed mutagenesis (*UNIT 8.1*), makes it possible to alter the DNA sequence in a defined way. This is accomplished by synthesizing an oligonucleotide whose sequence contains the mutation of interest, hybridizing the oligonucleotide to a template containing the wild-type sequence, and extending the primer with T4 DNA polymerase. The resulting product is a heteroduplex molecule containing a mismatch due to the mutation in the oligonucleotide. The mutation is “fixed” upon repair of the mismatch in *Escherichia coli* cells. This method is extremely valuable for situations in which it is desired to determine the effects of particular changes in the DNA, and it is also useful for introducing restriction sites at specific positions within a given stretch of DNA. However, it is relatively expensive (one oligonucleotide per mutation) and hence is limited to circumstances where one or a few specific mutations are desired.

The second protocol describes a method for generating a large number of mutations within a small region of DNA (*UNIT 8.2A*). It makes use of synthetically derived mixtures of oligonucleotides that are obtained by adding small, defined amounts of “incorrect” precursors at each step of the DNA synthesis. Each oligonucleotide molecule in the mixture thus has a defined probability of being altered from the wild-type sequence. This degenerate oligonucleotide mixture is converted to double-stranded DNA whereupon individual oligonucleotide molecules are isolated by molecular cloning. In principle, mutations occur at the frequency that was programmed into the DNA synthesis, and they occur at random positions throughout the region of interest. The major limitation of this method is the size of the oligonucleotide; thus, it is valuable for mutagenizing regions of DNA as large as 80 bases in length. However, it should be mentioned that larger regions can be mutagenized by using a set of contiguous or overlapping oligonucleotides that cover the region of interest.

The third procedure makes it possible to synthesize any desired gene segment by combining long oligonucleotides (*UNIT 8.2B*). The region of interest is subdivided into pairs of long single-stranded oligonucleotides that can be annealed at their 3' ends. Using methods similar to those described in *UNIT 8.2A* these pairs of oligonucleotides are converted into double-stranded DNA suitable for cloning. The final product, a custom-designed gene, is obtained by correctly assembling the double-stranded oligonucleotides. It is possible to generate desired sequences up to 400 bp in a single step, and longer regions can be obtained by combining the products of individual steps. The ability to synthesize long regions of any desired sequence is extremely valuable. For example, the introduction of restriction sites throughout a region greatly facilitates further analysis, and extensive modification of codons throughout a protein-coding sequence may result in an increased production of the protein. The only disadvantage of this procedure is its relative expense. It is most useful for solving a particular problem or for creating a modified gene that is more amenable for further study.

The fourth method makes it possible to generate many mutations within larger regions of DNA (up to 1 to 3 kb; *UNIT 8.3*). Here, single-stranded DNA containing the region of interest is treated with a variety of chemicals. By using an appropriate oligonucleotide primer, the mutated region is copied and then cloned. The mutation frequency can be set by the severity of the chemical treatment, and essentially all possible base substitutions can be obtained. This method is particularly valuable when mutagenizing regions of DNA that are larger than can be accommodated in a single or a few oligonucleotides (for shorter regions, the second method is preferred). However, since there are many possible mutations in such relatively large regions, this method is less useful for saturating a region with mutations. Instead, it is best suited for obtaining mutations that confer phenotypes of interest.

The fifth protocol, linker scanning (*UNIT 8.4*), describes a method for creating clustered point mutations in a relatively short region (typically 4 to 10 bp) such that a restriction site is located at the site of mutation. By creating and analyzing a series of linker-scanning mutations throughout a region, it is possible to quickly determine which sequences are functionally important. In addition, the presence of a common restriction site in these mutations allows for the generation of precise deletion or duplication mutations. For an initial functional dissection of a region, linker-scanning mutations have several advantages. Unlike deletion mutations, the changes are highly localized and do not alter any spacing relationships. Compared to point mutations, they are more disruptive and many fewer derivatives are necessary to cover the region; however, they are much less appropriate for analyzing the sequence requirements of a genetic element. Linker-scanning mutations are often used for dissecting transcriptional regulatory signals; they are rarely employed for analyzing protein-coding sequences.

The sixth protocol is a version of oligonucleotide-directed mutagenesis that utilizes the polymerase chain reaction. Oligonucleotides containing the desired changes are used as primers for amplification and the resulting products are cloned by standard methods. The specific procedure differs depending on whether or not the introduced changes generate a restriction site. The major advantages of this PCR-based mutagenesis method are technical simplicity, high frequency of incorporating mutations, and the relative lack of constraints regarding location of restriction sites in the DNA of interest. The method is relatively expensive because it requires two oligonucleotides that are specific for each introduced mutation. In addition, at least two other oligonucleotides are necessary for amplification of PCR, but these can be used for many different experiments if the relevant DNA fragments are cloned into a common vector.

Oligonucleotide-Directed Mutagenesis without Phenotypic Selection

UNIT 8.1

**BASIC
PROTOCOL**

A DNA sequence can be specifically altered by synthesizing the desired sequence change within an oligonucleotide, and then converting this into a biologically active circular DNA strand by using the oligonucleotide to prime *in vitro* synthesis on a single-stranded circular template. This protocol uses a DNA template containing a small number of uracil residues in place of thymine, as shown in Figure 8.1.1 (see further commentary on p. 8.1.5). Use of the uracil-containing template allows rapid and efficient recovery of mutants; in principle this same template can be applied to most of the other mutagenesis protocols in use.

Materials

Single-stranded bacteriophage vector with insert
TY medium containing 0.25 µg/ml uridine (UNIT 1.5)
E. coli CJ236 or alternative *dut⁻ ung⁻ F'* strain (Invitrogen and Table 1.4.5)
5× PEG/NaCl solution
TE buffer (APPENDIX 2)
T4 polynucleotide kinase (UNIT 3.10) and 10× kinase buffer (UNIT 3.4)
10 mM ATP (UNIT 3.4)
Mutagenic oligonucleotide primer
100 and 500 mM EDTA, pH 8.0 (APPENDIX 2)
20× SSC (APPENDIX 2)
5× polymerase mix
T4 OR T7 DNA polymerase (*not* Sequenase; see UNIT 3.5)
T4 DNA ligase (measured in Weiss units; UNIT 3.14)

Additional reagents and equipment for phage titering (UNIT 1.11), phenol extraction (UNIT 2.1), ethanol precipitation (UNIT 2.1), agarose gel electrophoresis (UNIT 2.5), preparation and transfection of competent cells (UNIT 1.8), and DNA sequence analysis (UNIT 7.4)

Preparing uracil-containing DNA template

1. With a sterile Pasteur pipet, remove one plaque produced by single-stranded phage containing the DNA of interest and place in 1 ml sterile TY medium in a 1.5-ml microcentrifuge tube.
2. Incubate 5 min at 60°C to kill cells, vortex vigorously to release phage from the agar, then pellet cells and agar with a 2-min spin in microcentrifuge.
3. Place 100 µl of resulting supernatant into 1-liter flask containing 100 ml TY medium supplemented with uridine to 0.25 µg/ml; add 5 ml of a midlog culture of *E. coli* CJ236 (*dut⁻ ung⁻ F'*) (Kunkel et al., 1987).

*These proportions result in a multiplicity of infection of considerably less than one per cell. Thus, essentially all of the input phage infect cells and are "passaged" through the *dut⁻ ung⁻* strain. Since few uracil-lacking phage remain, a single cycle of growth results in a sufficient survival difference (as measured by titers on *ung⁺* and *ung⁻* hosts) to make the DNA suitable for the *in vitro* mutagenesis protocol. Neither the thymidine nor the adenosine supplementation of medium originally described (Kunkel, 1985) is necessary.*

4. Incubate with vigorous shaking at 37°C for 6 to 18 hr.

Good aeration provided by vigorous shaking is important for high-phage titers.

**Mutagenesis of
Cloned DNA**

8.1.1

5. Pellet the cells by centrifugation at $5000 \times g$ for 30 min. The clear supernatant should contain phage at a titer of $\sim 10^{10}$ to 10^{11} pfu/ml.

*The phage yield may vary, depending on the vector and strain used. For example, experience with *E. coli* RZ1032 for growth of M13 derivatives containing amber mutations suggests that phage titers of 2 to 5×10^{10} pfu/ml are not unusual.*

6. Titer the phage on any *E. coli ung*⁻ (e.g., CJ236) versus *ung*⁺ strain (e.g., JM105, JM107, or JM109). Phage containing uracil in the DNA have normal biological activity in the *ung*⁻ host, but >100,000-fold lower survival in the *ung*⁺ host.
7. Precipitate phage by adding 1 vol of 5× PEG/NaCl solution to 4 vol supernatant. Mix and incubate 1 hr at 0°C.
8. Collect precipitated phage by centrifugation at $5000 \times g$ for 15 min. Drain the pellet well, resuspend in 5 ml TE buffer in a 15-ml Corex tube, and vortex vigorously.
9. Place the resuspended phage solution on ice for 1 hr and then centrifuge as above.

This step removes residual cell debris and has proven useful in reducing the level of endogenous low-molecular-weight DNA, which can nonspecifically prime in subsequent in vitro DNA polymerase reactions.

10. Purify the single-stranded phage DNA by phenol extraction and ethanol precipitation as described in UNIT 2.1. Determine the DNA concentration spectrophotometrically at 260 nm ($1 \text{ OD}_{260} = 36 \mu\text{g/ml}$).

Normally, further purification of the DNA is unnecessary in order to achieve high efficiencies of in vitro mutagenesis. If problems related to template purity are encountered or if mutant production approaching 100% is needed, the DNA can be subjected to any standard purification procedure, since the substitution of a small percentage of thymidine residues by deoxyuridine will not affect the physical properties of the DNA.

*In principle, any cloning vector that can be passaged through an *E. coli dut*⁻ *ung*⁻ strain can be used with the uracil selection technique. Once the uracil-containing DNA is prepared, it can be used as a standard template in a variety of in vitro methodologies for altering DNA sequences (Smith, 1986). Presented below is a typical oligonucleotide-directed mutagenesis experiment.*

Primer extension and product analysis

11. To a 1.5-ml microcentrifuge tube add the following:

2 μl 10× T4 polynucleotide kinase buffer
2 μl 10 mM ATP
Mutagenic oligonucleotide (15 to 50 nucleotides long)
H₂O to 20 μl

The amount of oligonucleotide used depends upon the desired molar ratio of oligonucleotide to single-stranded template. Ratios from 2:1 to 10:1 are routine, but in some instances higher ratios may be useful. For primers that are 15 to 20 bases in length, this corresponds to 4 to 30 ng/ μg single-stranded M13 template.

If the oligonucleotide has already been phosphorylated, proceed directly to step 13.

12. Add 2 U of T4 polynucleotide kinase and incubate 60 min at 37°C. Terminate the reaction by adding 3 μl of 100 mM EDTA, and heat to 70°C to denature the enzyme.

Phosphorylation of the oligonucleotide is required for the subsequent ligation step.

13. To the phosphorylated oligo (20 μ l) add the single-stranded circular uracil-containing DNA template (typically 1 μ g in 1 μ l); add 1.25 μ l of 20 \times SSC.
14. Mix thoroughly, spin 5 sec in microcentrifuge, then place tube in 500-ml beaker of water at 70°C. Allow to cool to room temperature. After another 5-sec spin to collect condensation, place tube on ice.

The precise hybridization conditions depend upon the stability of the heteroduplex formed between the mutagenic oligonucleotide and the template DNA. The ability of the oligonucleotide to prime in vitro DNA synthesis efficiently and at the appropriate position under the chosen hybridization conditions can be determined by DNA sequence analysis.

15. To the hybridization mixture add the following:

20 μ l 5 \times polymerase mix
2.5 U T7 or T4 DNA polymerase
2 U T4 DNA ligase
H₂O to 100 μ l

Since the T4 DNA polymerase activity in commercial preparations can vary with source and age of the enzyme, it may be necessary to add more than 2.5 U.

Enzymes are added last. T7 or T4 DNA polymerase is preferred over Klenow fragment because it will not "strand displace" the mutagenic oligonucleotide after synthesis is completed; this permits efficient ligation and expression of the mutation. T7 DNA polymerase (not Sequenase) is ideal for complete synthesis due to its high processivity. With some oligonucleotide template combinations, this enzyme does not work well, in which case 1 U Klenow fragment can be substituted.

16. Mix thoroughly, then incubate 5 min at 0°C, 5 min at room temperature, and 2 hr at 37°C.

The reaction is begun at lower temperatures to polymerize a small number of bases onto the 3' end of the oligonucleotide, thus stabilizing the initial duplex between the template DNA and the mutagenic oligonucleotide primer. However, since T4 DNA polymerase does not utilize long stretches of single-stranded DNA template well at low temperature, synthesis is then completed at 37°C. The use of a high concentration of dNTPs (500 μ M each) serves to optimize DNA synthesis and to reduce the 3'-exonuclease activity of the T4 DNA polymerase.

17. Add 3 μ l of 500 mM EDTA to terminate the reaction.
18. Analyze 20 μ l of the reaction mixture (200 ng of DNA) by agarose gel electrophoresis in a 0.8% agarose gel. For comparison, adjacent lanes should contain the following standards: single-stranded circular viral DNA; double-stranded, closed circular DNA (RFI); and nicked double-stranded circular DNA (RFII).

A successful reaction should convert essentially all the single-stranded template DNA into RFI and RFII DNA. In the absence of strand displacement (e.g., using T4 DNA polymerase), both RFI and RFII products yield mutations at high efficiency.

Transfection and DNA sequence analysis

19. Based upon an estimate from the gel analysis, use 1 to 100 ng of double-stranded DNA product to transfect any desired *ung*⁺ strain of *E. coli* cells.
20. Resulting clones (as phage plaques or colonies) can be selected or, if no phenotype is known, chosen randomly for isolation of pure genetic stocks. These can be analyzed by sequencing the DNA.

REAGENTS AND SOLUTIONS

5× PEG/NaCl solution

15% (wt/vol) polyethylene glycol 8000
2.5 M NaCl

5× polymerase mix

100 mM Tris·Cl, pH 8.0
10 mM dithiothreitol
50 mM MgCl₂
2.5 mM each of dATP, dTTP, dGTP, dCTP
5 mM ATP

Use of highly purified (e.g., by HPLC) dNTPs is critical.

COMMENTARY

Background Information

The basis of this method is use of a DNA template containing a small number of uracil residues in place of thymine (Fig. 8.1.1). The uracil-containing DNA is produced within an *E. coli dut⁻ ung⁻* strain. *Escherichia coli dut⁻* mutants lack the enzyme dUTPase and therefore contain elevated concentrations of dUTP which effectively competes with TTP for incorporation into DNA. *Escherichia coli ung⁻* mutants lack the enzyme uracil *N*-glycosylase which normally removes uracil from DNA. In the combined *dut⁻ ung⁻* mutant, deoxyuridine is incorporated into DNA in place of thymidine and is not removed. Thus, standard vectors containing the sequence to be changed can be grown in a *dut⁻ ung⁻* host to prepare uracil-containing DNA templates for site-directed mutagenesis.

For the in vitro reactions typical of site-directed mutagenesis protocols, uracil-containing DNA templates are indistinguishable from normal templates. Since dUMP in the template has the same coding potential as TMP, the uracil is not mutagenic, either in vivo or in vitro. Furthermore, the presence of uracil in the template is not inhibitory to in vitro DNA synthesis. Thus, this DNA can be used in vitro as a template for the production of a complementary strand that contains the desired DNA sequence alteration, but contains only TMP and no dUMP residues (Fig. 8.1.1).

After completing the in vitro reactions, uracil can be removed from the template strand by the action of uracil *N*-glycosylase. Glycosylase treatment can be carried out with purified enzyme, but it is most easily achieved by introducing the unfractionated products of the in vitro incorporation reaction into wild-type (i.e., *ung⁺*) *E. coli* cells. Treatment with the glycosylase, either in vitro or in vivo, releases uracil,

producing apyrimidinic (AP) sites in the template strand. These AP sites are lethal lesions, presumably because they block DNA synthesis and are sites for incision by AP endonucleases which produce strand breaks.

Thus, the template strand is rendered biologically inactive and the majority of progeny arise from the infective complementary strand which contains the desired alteration (Fig. 8.1.1). The resulting high efficiency of mutant production (typically >50%) allows one to screen for mutants by DNA sequence analysis, thus identifying mutants and confirming the desired alteration in a single step. This feature is particularly advantageous when no selection for the desired mutants is available.

Literature Review

Several variations of in vitro mutagenesis by primer extension that yield mutants with high efficiency have been developed (for review, see Smith, 1986). The procedure described here is the simplest site-directed mutagenesis protocol, but applied to a special uracil-containing template which allows rapid and efficient recovery of mutants (Kunkel, 1985; Kunkel et al., 1986). In principle, this same template can be applied to most of the other protocols in use.

Critical Parameters

Highly efficient mutagenesis depends upon a good template and a successful DNA polymerase reaction. The template should exhibit a strong difference in survival in *ung⁻* versus *ung⁺* hosts. In addition, it should be relatively free of endogenous low-molecular-weight DNA which could prime in vitro DNA synthesis in the absence of added oligonucleotide; such aberrant priming will yield nonmutant clones. A successful in vitro polymerase reac-

tion should convert the oligonucleotide into a fully double-stranded DNA molecule. In the absence of strand displacement of the oligonucleotide (e.g., using T4 DNA polymerase), high efficiency can even be achieved without ligation (Kunkel et al., 1986).

Finally, the oligonucleotide should be of high quality; i.e., purified away from lower molecular weight contaminants that arise from incomplete DNA synthesis. In some cases, especially for oligonucleotides larger than 40 nucleotides, purification by polyacrylamide gel electrophoresis may be necessary (see UNIT 2.7).

Troubleshooting

Incomplete synthesis can result from several factors, including inefficient hybridization of the oligonucleotide primer, inactive (or excess) DNA polymerase, contaminants in the DNA, the polymerase, or the reagents, or a DNA template that contains structures (e.g., hairpin loops) which block polymerization. Such problems must be dealt with on an individual basis, e.g., by varying hybridization conditions, by ensuring the activity of the DNA polymerase, by repurifying the DNA, or by using alternative incubation temperatures or single-stranded DNA binding protein to assist the polymerase in synthesis on unusually difficult templates.

Low biological activity could also result from dNTP contamination by dUTP (e.g., by deamination of dCTP) which, when incorporated in vitro, provides targets for the production of lethal AP sites in the newly synthesized strand containing the desired mutant. For this reason, high-quality dNTP substrates should be used. This eliminates the need for dUTPase pretreatment of the deoxynucleoside triphosphates (dUTPase is not commercially available).

Anticipated Results

With a high-quality template and polymerase reaction, mutants are recovered at 50 to 80%. Thus, direct sequence analysis of random clones to identify the desired mutant(s) is feasible.

Time Considerations

Starting from a single plaque, it should take 1 day to prepare single-stranded template DNA containing uracil residues. Enzyme reactions can be done in 1 day, and reaction products can be introduced into competent *E. coli* cells at the end of that day. Plaques that arise are purified (1 day) and single-stranded DNA ready for DNA sequence analysis is prepared (1 to 2 additional days). DNA sequencing reactions take 1 day, and the results can be seen on the next day. Thus, the entire procedure can take 5 to 7 days. The procedure can be stopped at essentially any step without significant effect on the mutant frequency.

Literature Cited

- Kunkel, T.A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. U.S.A.* 82:488-492.
- Kunkel, T.A., Roberts, J.D., and Zakour, R.A. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Meth. Enzymol.* 154:367-382.
- Smith, M. 1985. In vitro mutagenesis. *Ann. Rev. Genet.* 19:423-463.

Key Reference

Kunkel, 1985. See above.

Presents the original description of this technique.

Contributed by Thomas A. Kunkel
National Institute of Environmental
Health Sciences
Research Triangle Park, North Carolina

Mutagenesis with Degenerate Oligonucleotides: Creating Numerous Mutations in a Small DNA Sequence

UNIT 8.2A

**BASIC
PROTOCOL**

A collection of mutagenic oligonucleotides is synthesized by including low concentrations of the three non-wild-type precursors at each step of the synthesis. The frequency of mutation is set by choosing the appropriate amounts of non-wild-type precursors at each step of nucleotide addition. The oligonucleotide also contains an 8-nucleotide palindromic sequence at its 3' end that encompasses a restriction endonuclease site. The palindromic sequence allows nonidentical, but related, mutant DNAs to hybridize at their 3' ends, thereby serving as substrates for DNA polymerase. The resulting molecule is fully double-stranded homoduplex DNA. After digestion with the appropriate restriction endonuclease(s), the double-stranded DNA fragments can be readily cloned to generate a collection of single and multiple point mutants. The protocol is divided into four stages: (1) oligonucleotide design (which includes determination of the mutation frequency, the nucleotides to be mutated, and the structure of the 5' and 3' ends) and subsequent purification after synthesis; (2) mutually primed synthesis to convert the single-stranded oligonucleotide to double-stranded homoduplex form; (3) endonuclease cleavage of the homoduplexes and purification of the final product; and (4) cloning of the double-stranded oligonucleotide into suitable vectors and identification of the mutagenic sequence.

UNIT 8.2B describes related methods for synthesizing long gene segments of any desired sequence.

Materials

- 10× *E. coli* DNA polymerase I buffer (UNIT 3.4)
 - 10 mM dNTP mix of 2.5 mM 4dNTP mix (UNIT 3.4)
 - [α -³²P]dNTP (400 to 800 Ci/mmol; see UNIT 3.4)
 - Klenow fragment of *E. coli* DNA polymerase I (UNIT 3.5)
 - 0.5 M EDTA, pH 8.0 (APPENDIX 2)
 - TE buffer, pH 7.5 (APPENDIX 2)
 - 3 M sodium acetate, pH 5.2
 - Buffered phenol (UNIT 2.1)
 - Ethanol
 - Gel elution buffer: 0.5 M ammonium acetate/1 M EDTA (store protected from light up to several months at room temperature)
 - T4 DNA ligase (measured in cohesive-end units; UNIT 3.14)
- Additional reagents and equipment for restriction endonuclease digestion (UNIT 3.1), denaturing polyacrylamide gel electrophoresis (UNIT 7.4), phenol extraction and ethanol precipitation (UNIT 2.1), nondenaturing polyacrylamide gel electrophoresis (UNIT 2.7), transformation of *E. coli* (UNIT 1.8), and DNA sequence analysis (Chapter 7)

Design and synthesis of the degenerate oligonucleotide (see Fig. 8.2.1)

1. The oligonucleotide is designed such that the 3' end contains an 8-nucleotide palindromic sequence encompassing a restriction endonuclease cleavage site. The site can be one that gives cohesive or blunt ends.

The palindromic sequence is necessary for hybridization of two oligonucleotides to generate substrates for DNA polymerase. See critical parameters.

2. If possible, the 5' end of the oligonucleotide should consist of sequences encompassing a restriction endonuclease site.

A 5' restriction site is not crucial, since the polymerase reaction will generate blunt ends

Contributed by David E. Hill

Current Protocols in Molecular Biology (1989) 8.2.1-8.2.7

Copyright © 2000 by John Wiley & Sons, Inc.

**Mutagenesis of
Cloned DNA**

8.2.1

Supplement 46

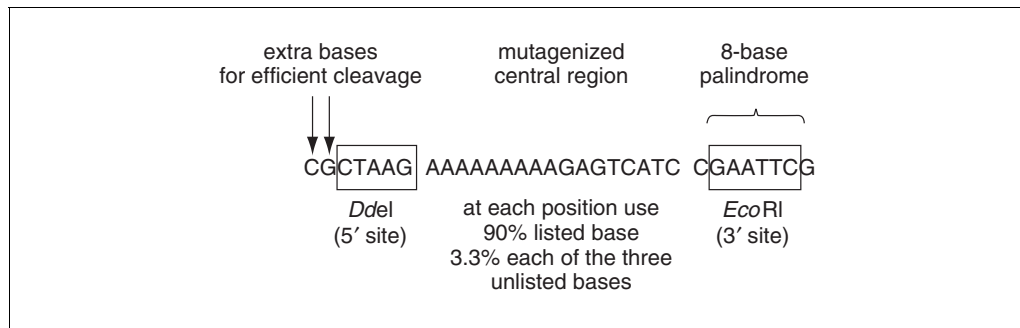


Figure 8.2.1 Design of degenerate oligonucleotide. A 17 base-pair region located between *EcoRI* and *Ddel* sites is mutagenized at a frequency of 10% per position.

suitable for cloning. See critical parameters for further details.

3. The central region contains the mutagenized sequence of interest. Depending on the specific experiment, the length and the mutation frequency of the central region will vary.

The overall limitation on length depends on the efficiency of oligonucleotide synthesis. The phosphite triester method of DNA synthesis as performed by an automated DNA synthesizer can easily produce degenerate oligonucleotides up to at least 100 nucleotides in length.

4. Synthesize the oligonucleotide. At positions where mutations are not desired, such as those comprising the restriction endonuclease cleavage sites, the synthesis is programmed to use homogenous solutions of individual nucleotide precursors. Where mutations are desired, the synthesis is programmed to use defined mixtures of nucleotide precursors.

The frequency of mutation at a given position is determined simply by the relative molarities of the precursors present in the solution. For example, a 10% mutation rate would be achieved with a mixture of 90% wild-type nucleotide and 3.3% each of the three "incorrect" nucleotides. Although newer models of DNA synthesizers are capable of programmed mixing, the most reproducible way to achieve a desired mixture of precursors is to combine appropriate amounts of solid nucleoside phosphoramidites prior to solubilization in anhydrous acetonitrile.

5. Following synthesis, the oligonucleotide is purified by high-performance liquid chromatography and/or electrophoresis in denaturing polyacrylamide gels containing 7 M urea. After purification, the oligonucleotide concentration is adjusted to 1 mg/ml in water (~50 to 100 μ M).

Conversion to double-stranded DNA by mutually primed synthesis (Fig. 8.2.2)

6. Transfer 200 pmol (roughly 1 to 2 μ g) of oligonucleotide to a 500- μ l microcentrifuge tube. Adjust the volume to 7 μ l with water.

The oligonucleotide does not need to be phosphorylated if there will be a restriction site at the 5' end after conversion to double-stranded form. If the 5' end will not contain a restriction site, the oligonucleotide should be phosphorylated by T4 polynucleotide kinase (UNIT 3.10).

7. Incubate the oligonucleotide for 5 min at 70°C; add 1 μ l of 10 \times DNA polymerase I buffer, cool to room temperature, and incubate for at least 60 min.

The second incubation can be done at any temperature that permits hybridization of the 3' end palindromic sequences (typically 23°C).

8. Add 2 μ l of 10 mM 4dNTP mix (2.5 mM each 4dNTP), 5 U Klenow fragment, and 10 μ Ci of any one [α -³²P]dNTP.

A low level of radioactively labeled dNTP is included in the reaction to facilitate purification of the DNA.



Figure 8.2.2 Mutually primed synthesis for cloning degenerate oligonucleotides. A degenerate oligonucleotide (top line; see Fig. 8.2.1.) is self-annealed via the 8-nucleotide palindrome on the 3' end that encompasses the *EcoRI* site, and then treated with the Klenow fragment. The resulting double-stranded DNA contains two oligonucleotide units each derived from an original oligonucleotide molecule (nucleotides that deviate from the "wild-type" sequence are depicted as large letters). Upon cleavage by *EcoRI* and *DdeI*, the double-stranded oligonucleotides are cloned into an appropriate vector.

9. Incubate 1 hr at 23°C. Add 5 more units of Klenow fragment and continue incubation for 2 hr more (or overnight).
10. Stop the reaction with 1 µl of 0.5 M EDTA, adjust the volume to 50 µl with TE buffer, and add sodium acetate to 0.3 M. Extract once with buffered phenol and precipitate the DNA with ethanol. Resuspend the DNA in 20 µl TE buffer.
11. Remove 2 µl for analysis on denaturing polyacrylamide gels (see step 17).
12. In a 30-µl reaction, digest the now double-stranded oligonucleotide with the restriction endonuclease recognizing the outside sites (originally the 5' site of the single-stranded oligonucleotide).

Because of the high concentration of restriction sites, complete digestion will probably require incubation for at least 2 hr with 10 to 40 U of enzyme per original microgram of oligonucleotide. If the oligonucleotide does not contain a 5' restriction site, cleave the DNA with the enzyme recognizing the internal restriction site (originally at the 3' end of the single-stranded oligonucleotide).

13. After digestion, remove 2 µl for analysis on denaturing polyacrylamide gels (see step 17), then extract the remainder of the reaction mixture with buffered phenol, and concentrate the DNA by ethanol precipitation.
14. Purify the double-stranded oligonucleotide mixture by electrophoresis on non-denaturing polyacrylamide gels. To prevent denaturation of the DNA, do not allow the gel to heat up above room temperature.

This step effectively removes the small DNA fragments generated by the restriction digest as well as unreacted, single-stranded oligonucleotides from the desired product. These

contaminants will reduce the ligation efficiency. In general, it is better to perform the gel purification step after cleavage of the outside restriction sites, but prior to cleavage at the internal site (originally the 3' site of the single-stranded oligonucleotide).

15. After electrophoresis, the desired double-stranded molecule is excised as a band from the gel and eluted in gel elution buffer, as described in UNIT 2.7. Resuspend the double-stranded oligonucleotide mixture in 20 μ l TE buffer and store at -20°C .

Since the double-stranded oligonucleotide is a heterogenous mixture of different sequences, the DNAs might not migrate as a sharp band, even though they are of identical length. Electroelution of the double-stranded oligonucleotide is an alternate method of extraction (see UNIT 2.6). Be sure that the molecular weight cutoff of the dialysis bag is below that of the oligonucleotide.

16. Digest the double-stranded oligonucleotide with the enzyme recognizing the internal restriction site (the original 3' site) to produce the final product—a double-stranded, homoduplex version of the oligonucleotide mixture with 5' and 3' ends suitable for ligation into standard vectors. Remove 2 μ l for analysis on denaturing polyacrylamide gels (see step 17). Phenol extract the remainder, ethanol precipitate, and resuspend in 20 μ l TE buffer.

17. Analyze the various 2- μ l aliquots (see steps 11, 13, 16) by electrophoresis on denaturing polyacrylamide gels to confirm that the individual reactions have produced the desired products.

Depending on the initial size of the oligonucleotide, 6 to 12% polyacrylamide gels similar to those used for DNA sequencing are employed. ^{32}P end-labeled oligonucleotides of various lengths are useful as size standards.

Cloning of oligonucleotide

18. Serially dilute the double-stranded oligonucleotide in TE buffer by successive factors of 10 until a 10,000-fold dilution is reached.

It is necessary to empirically determine the optimal oligonucleotide concentration for each cloning experiment. However, the empirical test will often generate a sufficient number of colonies that can be subjected to DNA sequence analysis.

19. Set up a series of ligation reactions, each containing a constant amount of vector and a portion of each dilution of oligonucleotide (UNIT 3.16).

The double-stranded oligonucleotides can be ligated directly into either bacteriophage M13-based vectors (UNIT 1.15) or into plasmid vectors. Due to reliable methods of sequencing double-stranded DNA molecules (UNIT 7.4), it is possible to ligate the oligonucleotide into the vector that will be used to determine directly the phenotypes of the mutants.

20. Introduce the ligation mixtures into an appropriate strain of *E. coli* by standard transformation procedures.

In order to avoid DNAs containing multiple oligonucleotide insertions, it is best to analyze colonies that were generated with the lowest amount of oligonucleotide. If the vector was generated by cleavage with two separate restriction enzymes, the background of colonies without inserted oligonucleotides should be low (UNIT 3.16).

21. Analyze the transformants by restriction enzyme digestion and DNA sequence analysis.

If the background is too high, colonies containing inserted oligonucleotides can be identified by colony filter hybridization using 5'- ^{32}P -labeled oligonucleotide as a probe. The procedure for hybridization using oligonucleotide probes is described in UNIT 6.4.

COMMENTARY

Background Information

The procedure described here is an efficient method for using the products of a single oligonucleotide synthesis to create numerous mutations in a small region of DNA (20 to 80 nucleotides). Mutant oligonucleotides are generated during the organic synthesis reaction by including low concentrations of the three non-wild-type nucleotide precursors with the wild-type precursor at each step of the synthesis (see Fig. 8.2.1). The product of such a DNA synthesis is a degenerate oligonucleotide, i.e., a complex mixture of related molecules, each of which has a defined probability of being altered from the wild-type sequence. The frequencies and types of zero, single, double, and higher order mutations can be set simply by choosing the appropriate amounts of non-wild-type precursors at each step of nucleotide addition during the synthesis.

An important feature of the method is that the single-stranded degenerate oligonucleotide is converted to double-stranded homoduplex molecules that can be cloned directly into standard vectors (see Fig. 8.2.2). The palindromic character of the 3' end of the oligonucleotide allows nonidentical DNAs to hybridize; thus, the oligonucleotides serve as mutual primers for extension by the Klenow fragment of *E. coli* DNA polymerase I. The resulting product is a homoduplex of length $2A + 2N + B$ (where A is the length of the 5' flanking sequence, N is the length of the heterogeneous, mutagenized central region, and B is the length of the 3' palindrome flanking region). By cloning homoduplex DNAs, mismatch repair in vivo is avoided, potential bias against particular mismatched nucleotides is prevented, and mutations located in all possible positions of a sequence can be obtained.

Since the desired rate of mutation is programmed into the oligonucleotide synthesis and hence is predetermined, the expected results from any particular cloning of degenerate oligonucleotides can be calculated simply using the laws of probability. For example, to maximize the number of single and double point mutants within a 15-bp region of DNA, a 10% mutation rate can be employed. This should give, on average, 1.5 nucleotide substitutions per oligonucleotide. Most of the oligonucleotides will have one or two mutations, while some will have none or more than two. The oligonucleotides lacking any mutations

(i.e., the wild-type sequence) provide a necessary control without requiring an extra synthesis.

For many experiments where it is desired to generate a large number of single base pair substitutions within a given region, low mutation rates (1 to 20% per nucleotide) are typically used. For such experiments, eight mixtures of nucleotide precursors are needed during the synthesis. Four of the mixtures are homogeneous solutions of pure nucleotide precursors such as would be used for conventional oligonucleotides. The remaining four mixtures are heterogeneous solutions composed of one major precursor and predetermined amounts of the three "mutagenic" precursors.

Higher mutation frequencies can be used to saturate a particular site with all possible nucleotide changes. In this case, a separate oligonucleotide is synthesized for each single nucleotide that is to be changed to the three possible non-wild-type nucleotides. At the relevant step in the synthesis, an equimolar mixture of the three non-wild-type nucleotide precursors (i.e., 33% each) replaces the wild-type precursor.

Finally, it is sometimes useful to synthesize "custom-designed" degenerate oligonucleotides. For example, when carrying out a low-frequency mutagenesis as described above, it is extremely difficult to obtain all possible mutations. Indeed, although it is relatively easy to obtain about 70% of the possible changes, continued random sequencing will result in many "repeat" mutations before new ones are obtained. In this case, it is better to design a degenerate oligonucleotide that optimizes the chances of obtaining any "missing" mutations of interest. Many, if not all, of the missing mutations can often be obtained with one or a few additional oligonucleotides.

Literature Review

The use of degenerate oligonucleotides for mutagenesis of small regions of DNA was first described by Matteucci and Heyneker (1983). The protocol described here differs from this and alternative procedures basically in the method of cloning the degenerate oligonucleotides. Previous cloning methods have required the synthesis of complementary DNA strands followed by hybridization. Such hybrids necessarily have one or more mismatches which must be repaired in vivo following trans-

formation. The use of mismatch hybrids can reduce overall cloning efficiency and limit the spectrum of single and multiple base changes that occur in a given region of DNA.

The method of mutually primed synthesis described here eliminates cloning bias and positional effects of mismatches because all DNAs are cloned as homoduplex molecules (Hill et al., 1986a,b; Oliphant et al., 1986). In addition, since only a single oligonucleotide synthesis is necessary for obtaining both the wild-type control DNA and a large collection of mutants, significant cost savings can be realized.

Critical Parameters

Since the method relies on efficient organic DNA synthesis, it is essential that the oligonucleotide be synthesized with the freshest possible precursors. This is especially important for the mixed solutions in order to minimize any bias in nucleotide addition. In particular, G residues are sometimes added relatively poorly if the reagents are not fresh. The overall limitation on the length of the region to be mutagenized depends critically on the efficiency of the oligonucleotide synthesis. It is now possible to obtain oligonucleotides that are 100 bases in length, and technological improvements are likely to extend this further. Since 8 to 20 nucleotides at the 5' and 3' ends are constrained by the method, this makes it possible to mutagenize a region as large as 90 bp with a single oligonucleotide. Moreover, larger regions can be mutagenized by using a set of contiguous or overlapping oligonucleotides.

The main critical parameters are the structures at the 5' and 3' ends. The only absolute requirement of the method is that the 3' end be a palindromic sequence that can be cleaved by a restriction endonuclease. The palindrome is necessary for the mutual priming reaction, and the restriction site is necessary for cleavage of the initial product, an oligonucleotide dimer, to oligonucleotide units suitable for cloning. It is preferable if the 5' end also contains a sequence that is recognized by a restriction endonuclease. In the ideal case, cleavage at the 5' site will generate cohesive ends that are different from those generated at the 3' site; this facilitates directional cloning (UNIT 3.16) and makes it possible to perform complex ligation reactions using two different fragments to provide the relevant 5' and 3' joining sites. However, the 5' terminal sequences can be anything; in this

case, the oligonucleotides are cloned via the blunt ends generated by the mutually primed synthesis procedure.

If a restriction site is to be used at both the 5' and 3' end of the oligonucleotide, there is a choice as to which site should be on the different ends. In this case, the restriction site having the highest GC composition should be used at the 3' site in order to facilitate the annealing reaction. For this reason, the palindrome at the 3' end should usually be extended to 8 bases by flanking the site with an additional G or C residue; this has the consequence that the inserted oligonucleotide will contain an extra base between the 3' restriction site and the sequence of interest. Six-base palindromes at the 3' end can suffice for mutually primed synthesis if all residues are GC. However, the reaction is considerably less efficient if other 6-base palindromes are used.

If the 5' end contains a sequence that is to be cleaved with a restriction endonuclease, it is useful to include 1 to 3 extra nucleotides beyond the recognition sequence at the 5' end to facilitate cleavage. In addition, it is advantageous to minimize the length of the palindrome at the 5' end in order to disfavor hybridization that might block complete extension. This can be easily arranged by choosing additional 5'-terminal nucleotides that are not complementary to those immediately flanking the 5' site on the other side.

The choice of vector depends on what is to be done with the mutations once they are obtained. Cloning into standard M13 vectors such as M13mp18 or M13mp19 (UNIT 1.15) has the advantage of easy DNA sequencing; however, the mutated fragments will generally have to be transferred to another molecule for phenotypic analysis. This can be rather cumbersome as it requires preparation of double-stranded M13 DNA (for cloning) in addition to single-stranded DNA (for sequencing). If possible, it is better to clone the degenerate oligonucleotides directly into the molecule that will be used for the phenotypic analysis. By sequencing the double-stranded DNA (UNIT 7.4), only one DNA preparation has to be performed. However, it should be noted that sequencing using the final molecule usually requires a special oligonucleotide primer adjacent to the region of interest.

Anticipated Results

The organic synthesis reaction should easily produce several micrograms of oligonucleo-

tion. After conversion to the double-stranded form, one should have at least 100 ng of DNA. With standard transformation efficiencies, it is possible to generate numerous different mutants and to effectively saturate a small region of DNA with all possible single base changes.

Time Considerations

Typically, oligonucleotides are produced on automated synthesizers that are operated by trained personnel. The actual synthesis and initial purification of the oligonucleotide takes only a few days (waiting for the synthesis facility to make the oligonucleotide can sometimes take a lot longer).

Given the oligonucleotide, conversion to double-stranded form by mutually primed synthesis, the first restriction endonuclease digestion, and nondenaturing gel electrophoresis will require 1 to 2 days. The gel can be run overnight at the end of day 1. Elution of the double-stranded oligonucleotide from the gel, the second restriction digest, and ethanol concentration will require ~1 day. The ligation reaction can be performed overnight at the end of day 2. Transformation and plasmid DNA preparation will require 2 to 3 days. Thus, at the end of 5 days, plasmid or phage DNA can be ready for sequencing to identify the exact

mutation. DNA sequencing will require 1 day with an overnight autoradiographic exposure.

Thus, in 1 week, it is possible to start with a degenerate oligonucleotide and finish with identified mutant DNAs already cloned into vectors suitable for phenotypic analysis.

Literature Cited

- Hill, D.E., Hope, I.A., Macke, J.P., and Struhl, K. 1986a. Saturation mutagenesis of the yeast *His3* regulatory site: Requirements for transcriptional induction and binding by GCN4 protein. *Science* 234:451-457.
- Hill, D.E., Oliphant, A.R., and Struhl, K. 1987. Mutagenesis with degenerate oligonucleotides: An efficient method for saturating a defined DNA region with base pair substitutions. *Meth. Enzymol.* 155:558-568.
- Matteucci, M.D. and Heyneker, H.L. 1983. Targeted random mutagenesis: The use of ambiguously synthesized oligonucleotides to mutagenize sequences immediately 5' of an ATG initiation codon. *Nucl. Acids Res.* 11:3113-3121.
- Oliphant, A.R., Nussbaum, A.L., and Struhl, K. 1986. Cloning of random-sequence oligodeoxynucleotides. *Gene* 44:177-183.

Contributed by David E. Hill
Applied Biotechnology
Cambridge, Massachusetts

Gene Synthesis: Assembly of Target Sequences Using Mutually Priming Long Oligonucleotides

BASIC PROTOCOL

The most straightforward way to generate a desired sequence is simply to synthesize it. In the long run, synthesis of appropriately modified versions of a wild-type sequence can also simplify approaches to obtaining many types of mutant versions. This protocol uses pairs of oligonucleotides annealed at a short duplex segment at their 3' ends as both templates and primers (mutually primed synthesis) to generate desired sequences up to 400 bp in a single step (see Fig. 8.2.3).

As with the procedure described in UNIT 8.2A, this protocol is divided into three sections: design of the oligonucleotides, extension by mutually primed synthesis, and cloning of the extension products. The strategy for design of oligonucleotides (steps 1 to 5 in the previous protocol) is discussed in the commentary following this protocol. The procedures for using mutually primed synthesis to extend and clone two partially complementary oligonucleotides are analogous to those described in UNIT 8.2A using a single small oligonucleotide, but have been modified significantly.

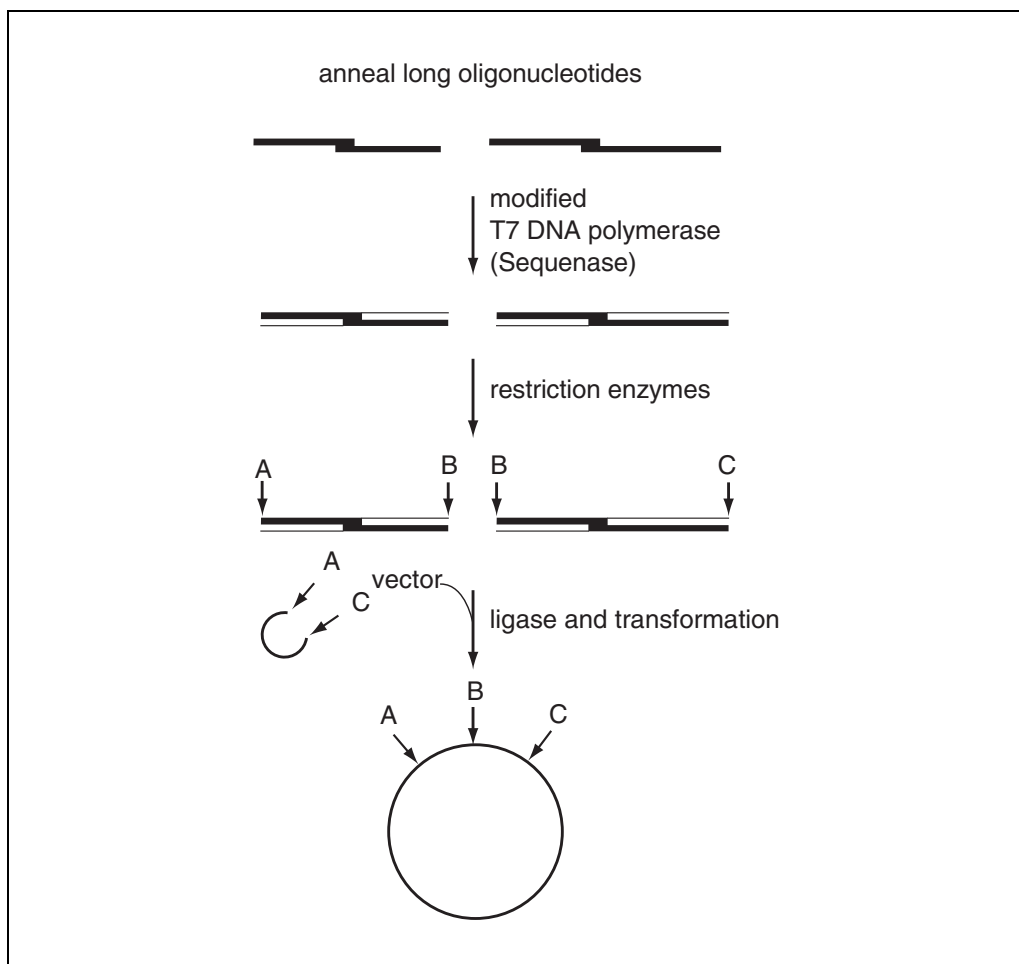


Figure 8.2.3 Strategy for constructing a synthetic gene using two pairs of oligonucleotides. Lines represent two pairs of oligonucleotides, corresponding to the four sequences designated by arrows in Figure 8.2.5. Note the region of duplex in each pair which makes mutually primed synthesis possible. A, B, and C represent restriction endonuclease sites needed for cloning the annealed, extended oligonucleotide pairs into the vector. In Figure 8.2.5 sites A, B, and C are *Xba*I, *Hind*III, and *Pst*I, respectively.

Materials

10× T7 DNA polymerase (Sequenase) buffer (UNIT 3.4)
4dNTP mix (2.5 mM each dNTP; UNIT 3.4)
Modified T7 DNA polymerase (Sequenase, U.S. Biochemical; UNIT 3.5)
4 M ammonium acetate
100% and 95% ethanol
TE buffer

Additional reagents and equipment for restriction endonuclease cleavage (UNIT 3.1), phenol extraction and ethanol precipitation (UNIT 2.1), sieving agarose gel electrophoresis (UNIT 2.8), subcloning (UNIT 3.1), and DNA sequencing (UNIT 7.4)

The following steps replace steps 6 to 21 of the basic protocol in UNIT 8.2A.

1. Transfer 1 μg of each of the two oligonucleotides to a microcentrifuge tube, adjust to 17 μl water, and add 2 μl of 10× Sequenase buffer.
2. Heat 5 min at 70°C, followed by 5 min at a temperature appropriate to anneal the oligonucleotides. Remove a 2- μl aliquot for later analysis.

The temperature may be crudely estimated by adding 2°C for each A-T base pair plus 4°C for each G-C base pair in the duplex segment (see UNIT 6.4).

3. Extend the annealed oligonucleotides by adding 2 μl dNTP mix (2.5 mM each dNTP) and 10 U Sequenase. Incubate 30 min at 37°C.

*Because of its processivity, high activity, and relative insensitivity to secondary structure (see UNIT 3.5), Sequenase is particularly appropriate for extensions of longer oligonucleotides, although other DNA polymerases (such as the Klenow fragment of *E. coli* DNA polymerase I) should be adequate. It is not necessary to add ^{32}P -labeled deoxynucleotide triphosphate since there should be enough material to check the single-stranded starting material and the double-stranded product by ethidium bromide staining of a sieving agarose gel (Fig. 8.2.4).*

4. Heat the extended oligonucleotides 10 min to 70°C to inactivate the DNA polymerase. Remove a 2- μl aliquot.
5. Increase the volume to 100 μl with restriction enzyme buffer and 20 to 100 U of restriction enzymes appropriate for cloning. Digest ≥ 2 hr.
6. Phenol extract. Add 100 μl of 4 M ammonium acetate and 400 μl of 100% ethanol, chill to -70°C for 15 min, and spin 5 min in a microcentrifuge. Redissolve pellet in 100 μl TE buffer, add 100 μl of 4 M ammonium acetate and 400 μl of 100% ethanol, chill, and spin again. Rinse pellet with 95% ethanol and dry. Resuspend in 20 μl TE buffer. Remove 2 μl for analysis.

This ammonium acetate differential precipitation is only appropriate if the extended product is larger than approximately 50 bp (see UNIT 2.1). Shorter extended duplexes can be purified by electrophoresis on nondenaturing polyacrylamide gels (UNIT 2.7) or used directly for cloning after recovery using two ethanol precipitations from 0.3 M sodium acetate (UNIT 2.1).

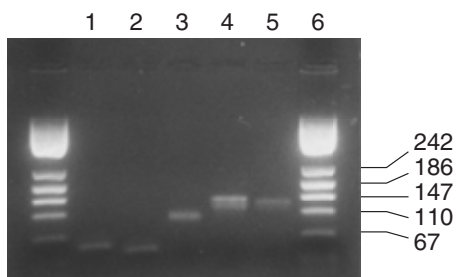


Figure 8.2.4 Sieving agarose gel electrophoresis of starting oligonucleotides and extended, restriction endonuclease-digested product. The oligonucleotides used are 80 nucleotides long, with a 20-bp duplex for self-priming, and the digested product is 135 bp. The sizes of the MW marker (bp) are indicated. **Lanes 1 and 2**, single-stranded oligonucleotides; **lane 3**, annealed oligonucleotides; **lane 4**, extended oligonucleotides; **lane 5**, restriction enzyme-digested extended oligonucleotides; **lane 6**, molecular weight marker.

7. Check the starting material, the extended product, and the digested product by electrophoresis on sieving agarose (see Fig. 8.2.4).

The extended, digested material should give a sharp band of the expected molecular weight.

8. Estimate the amount of extended oligonucleotide present from the analytical gel and subclone using desired vector (see UNIT 3.16 for discussion of subcloning).

Target sequences of 300 to 400 nucleotides are conveniently cloned in three fragment ligations containing two extended, digested target oligonucleotides (see UNIT 3.16, Example 3.16.2). In this case it is important to use equal concentrations of each double-stranded fragment.

9. Sequence several appropriate subclones.

See UNIT 7.4 for a protocol for direct sequencing of minipreps of double-stranded plasmids.

COMMENTARY

Background Information

The strategy for large synthetic projects used here is diagramed in Figure 8.2.3. Overall, this approach is simpler and substantially less expensive than most previous approaches (e.g., Wosnick et al., 1987)—which have relied on annealing relatively large numbers of shorter oligonucleotides—and treatment with DNA ligase to generate the desired fragment. The use of pairs of relatively long oligonucleotides (>100 nucleotides) for mutually primed synthesis minimizes the number of base pairs synthesized chemically, thus minimizing cost. The simple priming reaction is easier to set up and control than the multioligonucleotide annealing reaction. In general, it is also more convenient to synthesize and purify two relatively long oligonucleotides than a larger number of shorter ones. A somewhat similar strategy using single, self-priming long oligonucleotides has recently been described. (Uhlmann, 1988).

It is difficult to describe an average synthesis project, since essentially any sequence could be a target for such an approach. The strategy described here is appropriate for synthesis of *cis*-acting regulatory elements such as enhancers, promoters, and origins of DNA replication. However, it is useful to discuss in some detail the strategy for one relatively common, simple problem: synthesis of the coding region for a protein of modest size. The example described here is the synthesis of a modified version of the coding region for the *E. coli trp* repressor appropriate for further mutagenesis and for expression in mammalian cells. Although this example deals with a wild-type

sequence, any desired amino acid changes could obviously be included, and mutagenesis of a particular region could be carried out using partially degenerate oligonucleotides, as described in the basic protocol.

Generating the target sequence

The synthetic version of the coding region of the wild-type *E. coli trp* repressor is diagramed in Figure 8.2.5. Briefly, two types of changes were incorporated into the synthetic version. (1) A number of changes were introduced to facilitate expression in mammalian cells, including extensive alteration in the area of the initiator AUG to conform to appropriate rules for translation initiation (Kozak, 1987), and alteration of several codons poorly represented in mammalian coding sequences to more commonly used forms (Aota et al., 1988). (2) Restriction endonuclease sites convenient for cloning were placed at the 5' ends of the oligonucleotides, and a number of changes that generated useful new restriction enzyme sites but did not affect amino acid sequence were included. The newly added sites were chosen to facilitate both analysis of primary subclones of the oligonucleotides and future mutagenesis. Since not all cloned oligonucleotides end up with precisely the desired sequence, such additional sites also allow assembly of secondary constructs with the desired sequence from various mutant primary isolates, and are a potentially useful aspect of any large synthesis strategy.

Identification of potential sites for nucleotide sequence changes that generate restric-

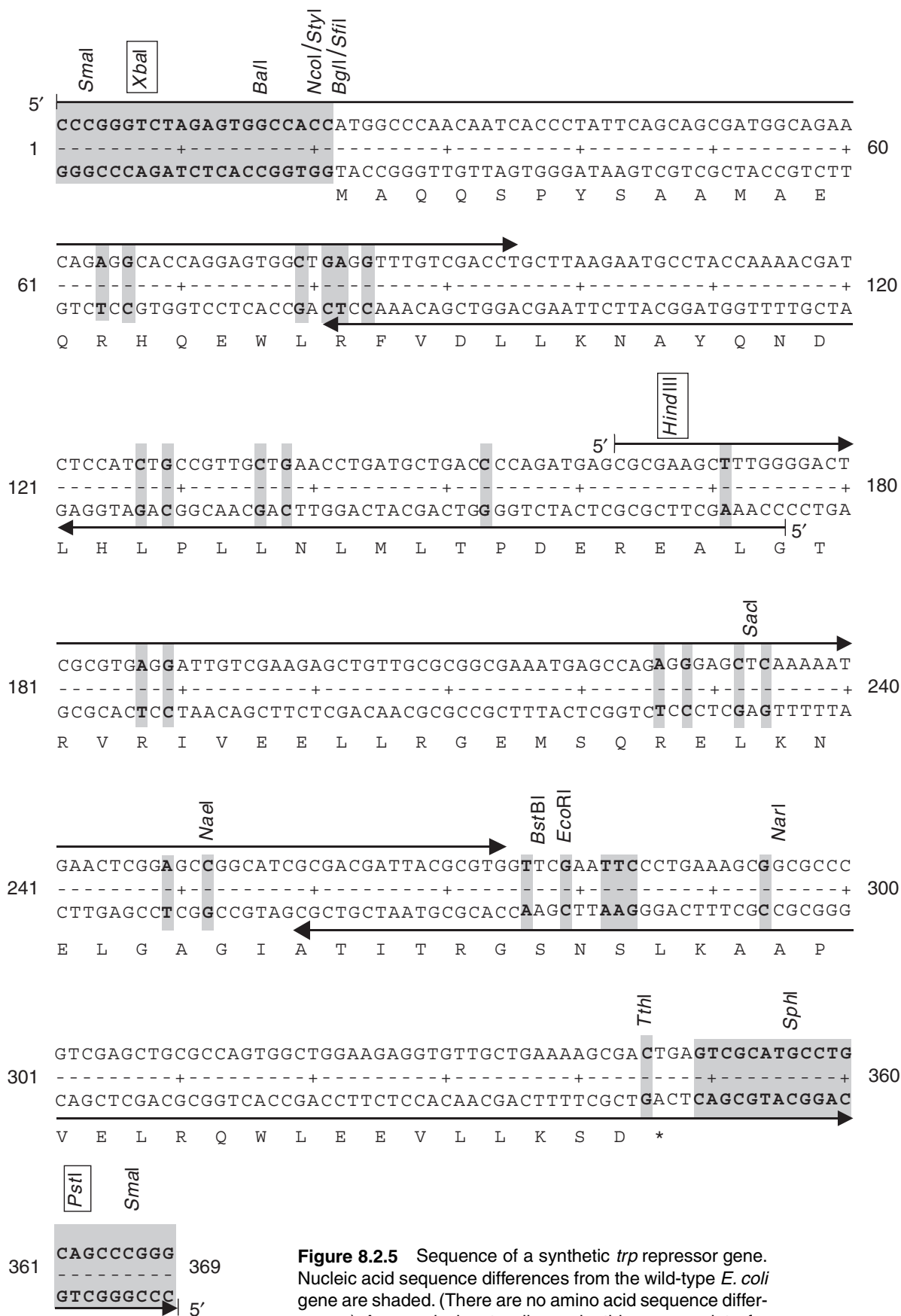


Figure 8.2.5 Sequence of a synthetic *trp* repressor gene. Nucleic acid sequence differences from the wild-type *E. coli* gene are shaded. (There are no amino acid sequence differences.) Arrows designate oligonucleotides appropriate for synthesis. Several newly introduced restriction enzyme sites are indicated; sites used for cloning are boxed.

tion sites but do not alter amino acid sequence is greatly aided by programs specifically designed for this purpose, such as MAP (using the option /SIL) from the University of Wisconsin computer group (see Devereaux et al., 1984). It is worth noting that the opposite strategy—removal of restriction sites without changing coding regions—is quite straightforward and often useful.

Dividing the target sequence into synthetic segments

As indicated by arrows in Figure 8.2.5, two pairs of oligonucleotides (2 of 95 nucleotides and 2 of 111) were chosen, based on the strategy diagrammed in Figure 8.2.3, to cover the 369 bp of the desired fragment. The two pairs include a 15- or 16-bp duplex segment for mutually primed synthesis. This length allows rapid and efficient annealing at convenient temperatures and the resulting duplex is stable during the extension reaction. Additional residues were added beyond sites used for cloning to allow efficient enzyme cleavage.

This example describes a strategy for synthesis of a 369-bp fragment in one step by cloning two pairs of extended oligonucleotides at once. While the use of a three-fragment ligation lowers efficiency of subcloning somewhat, it should not create difficulties if the fragments have heterologous ends (subcloning strategies are discussed in *UNIT 3.16*). For longer target sequences, even more complex ligations could be imagined. However, longer synthetic segments may be substantially less convenient to sequence than several shorter subclones, and are more likely to contain undesired mutations necessitating reconstruction of the desired sequence from individual primary clones. Although other strategies are certainly possible, it may be simpler to divide larger projects into segments of 300 to 400 base pairs, which can then be assembled in secondary ligations. Thus, the 1600-bp coding region of an approximately 50-kD protein might be synthesized as four such segments and assembled by sequentially combining 400- and 800-bp fragments. Of course, careful choice of restriction sites and vectors is particularly important in such projects.

Critical Parameters

Clearly, it is essential to use high-quality oligonucleotides and deoxynucleotide triphos-

phates. The oligonucleotides must be purified away from contaminating shorter chains and unwanted byproducts of synthesis by gel electrophoresis (*UNIT 2.12*) or HPLC (*UNITS 10.12-10.14*).

As described in the example, a ~15-bp duplex is a convenient length for mutual priming. In general, longer sequences are not necessary, and duplexes <10 base pairs should be avoided to ensure stability, particularly if they are A·T rich. A potentially important design problem not encountered in the example described here concerns secondary structures that might interfere with the mutually primed synthesis. In particular, to avoid undesirable extension products, the 3' ends of the oligonucleotides should be designed so they are not self-complementary and do not form stable duplexes with internal sequences. Large, stable internal hairpins could also inhibit extension and, if possible, should be avoided.

Judicious choice of restriction sites can greatly ease analysis of primary clones and their assembly into larger target sequences. Particularly for larger sequences, it may be necessary not only to add desired sites but also to remove undesirable sites by alteration of third positions of codons. Mutations are commonly observed using any synthetic approach, and it is essential to sequence the primary subclones carefully. Including appropriate restriction sites can facilitate reassembly of the desired sequence, from combinations of mutant primary clones.

At various stages of the project, it may be necessary to manually enter rather long sequences into the computer and/or the synthesizer. Thus, there is a realistic possibility of human error in getting the right sequence from design to clone, and sequences must be checked carefully.

Troubleshooting

The main problems likely to be encountered with this strategy are lack of extension of the oligonucleotides and high mutation frequencies. In both cases, low-quality oligonucleotides are the most likely cause. It is essential to purify the oligonucleotides before use, but experience indicates that even gel-purified oligonucleotides produced by syntheses of relatively poor yield do not extend reliably.

Other reasons for failure to extend the oligonucleotides include unanticipated secondary structures or problems with deoxynucleotide triphosphates or DNA polymerase.

As rather large quantities of oligonucleotides are used, one should be sure to include sufficient amounts of triphosphates.

Anticipated Results

The mutually primed synthesis approach should generate a large amount of double-stranded, extended oligonucleotide, and it should be straightforward to use it to generate a large number of primary subclones. Depending on the quality of the oligonucleotides, it should be feasible to obtain desired target sequences of ~400 bp in one step. With judicious choice of restriction sites for cloning, it is possible to start with a number of 400-bp segments at once and assemble larger target sequences rapidly.

Time Considerations

It should take 2 to 3 days to generate primary clones from oligonucleotides and another 2 to 3 days to verify their sequence. (In many cases, of course, one may have to wait much longer for the oligonucleotides to be synthesized and purified.) Starting with a number of 400-bp segments at once, and assuming perfect oligonucleotides and triphosphates, it is theoretically possible for a diligent worker to generate the coding region for a large protein (50 kD) in

<2 weeks. More realistically, one might estimate that assembling 700 to 800 bp of desired sequence in 2 weeks is a fair goal.

Literature Cited

- Aota, S.I., Gojobi, T., Ishibashi, F., Maruyama, T., and Ikemura, T. 1988. Codon usage compiled from the GenBank sequence data. *Nucl. Acids Res.* 16:r315-r403.
- Devereaux, J., Haberli, P. and Smithies, O. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* 12:387-395.
- Kozak, M. 1987. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J. Mol. Biol.* 196:947-950.
- Uhlmann, E., 1988. An alternate approach in gene synthesis: Use of long selfpriming oligodeoxynucleotides for construction of double-stranded DNA. *Gene* 71:29-40.
- Wosnick, M.A., Barnett, R.W., Vicentini, A.M., Erfle, H., Elliott, R., Summner-Smith, M., Mantei, N., and Davies, R.W. 1988. Rapid construction of large synthetic genes: Total chemical synthesis of two different versions of the bovine pro-chymosin gene. *Gene* 60:115-127.

Contributed by David D. Moore
Massachusetts General Hospital
Boston, Massachusetts

Random Mutagenesis by PCR

Error-prone PCR (EP-PCR) is the method of choice for introducing random mutations into a defined segment of DNA that is too long to be chemically synthesized as a degenerate sequence. For shorter (<100 bp) DNA segments, the most powerful method is to chemically synthesize the sequence with a low degree of nucleotide degeneracy (UNIT 8.2A). This method, however, is too expensive and time-consuming to be used for mutagenizing long stretches of DNA. Using EP-PCR, the 5' and 3' boundaries of the mutated region may be defined by the choice of PCR primers. Accordingly, it is possible to mutagenize an entire gene or merely a segment of a gene. The average number of mutations per DNA fragment can be controlled as a function of the number of EP-PCR doublings performed.

The EP-PCR technique described here is based on the protocol of Cadwell and Joyce (1992). The Basic Protocol is for a 400-bp sequence; the Alternate Protocol is for a library. EP-PCR takes advantage of the inherently low fidelity of *Taq* DNA polymerase, which may be further decreased by the addition of Mn^{2+} , increasing the Mg^{2+} concentration, and using unequal dNTP concentrations.

STRATEGIC PLANNING

After choosing a region of DNA to randomly mutagenize, one must decide on the desired level of mutagenesis that is best suited to the project. If the mutation rate is too low, it may not be possible to find the potentially rare variants of interest. If the mutation rate is too high, nearly all of the resulting library molecules will carry multiple mutations and may therefore be inactive. The desired extent of mutation depends on the type of activity one is attempting to generate and the number of library members that can be screened. For example, if one is attempting to generate an activity that may require multiple mutations, such as a novel binding or catalytic activity, then it may be necessary to generate molecules with several mutations per template. In this case it would also be important to have a screening or selection protocol that would allow for the evaluation of a large number of variants. On the other hand, if one is looking for mutations causing loss of protein function or decrease in thermal stability, for example, then 1 to 2 mutations per template may be ideal. In this case it may be sufficient to screen a smaller number of

Table 8.3.1 Average Number of Mutations per DNA Template as a Function of Template Length and Number of EP-PCR Doublings

EP-PCR doublings	Mutations per nucleotide position	Template length				
		100 bp	200 bp	400 bp	800 bp	1600 bp
5	0.0033	0.33	0.66	1.3	2.6	5.3
10	0.0066	0.66	1.3	2.6	5.3	11
20	0.013	1.3	2.6	5.3	11	21
30	0.020	2.0	4.0	7.9	16	32
50	0.033	3.3	6.6	13	26	53

variants. A reasonable approach in many instances is to generate a library such that a few unmutagenized molecules will be present in the collection of screened clones.

The average number of mutations per template increases as a function of the number of doublings in the EP-PCR reaction (Table 8.3.1). It should be noted that it is the number of doublings that is the determining factor, rather than the number of EP-PCR cycles. Each cycle of EP-PCR generally increases the amount of DNA by a factor of 1.7 to 1.9 until the DNA concentration reaches a plateau and then stops increasing altogether. The point at which this plateau occurs depends on the template and primer lengths and sequences, but is generally in the range of 5 to 50 ng/μl. It is not advisable to continue thermal cycles beyond the plateau point.

Prior to performing the actual EP-PCR reaction, it is important to run a pilot reaction to determine the amplification efficiency under the EP-PCR reaction conditions. This should be evaluated for two reasons. First, if the amplification per cycle is too low (<1.7-fold increase in product DNA concentration per cycle), DNA fragments that contain one or both of the primer-binding sites, but are shorter than the desired product, may have a strong selective advantage for amplification. These shorter fragments are produced by mispriming during normal or error-prone PCR. After several cycles, these shorter sequences may “take over” the EP-PCR reaction. This can be an especially severe problem when many cycles (>15) are to be performed. Second, the amplification per cycle must be known in order to calculate the number of EP-PCR cycles necessary to achieve the desired number of doublings.

The amount of DNA amplification per EP-PCR cycle can be determined by diluting a known amount of the unmutagenized PCR product, then amplifying it using the EP-PCR protocol and occasionally removing portions of the reaction for quantitation on an ethidium bromide-stained agarose gel (UNIT 2.7). The amplification per cycle should generally be >1.7. The yield per cycle can be optimized by altering the annealing temperature. Also, the extension time should last for at least 3 min, to ensure complete extension. If this does not bring satisfactory results, longer or shorter primers may have to be used. Primer lengths of 20 to 40 nucleotides usually produce acceptable results. Other parameters may also be optimized to improve the amplification, as described in UNIT 15.1 (the Mg²⁺ concentration should not be reduced, however, as this may lead to increased fidelity of DNA synthesis).

Table 8.3.2 Fraction of Nonmutated DNA Templates as a Function of Template Length and Number of EP-PCR Doublings

EP-PCR doublings	Mutations per nucleotide position	Template length				
		100 bp	200 bp	400 bp	800 bp	1600 bp
5	0.0033	0.72	0.52	0.27	0.071	0.0050
10	0.0066	0.52	0.27	0.071	0.0050	2.5 × 10 ⁻⁵
20	0.013	0.26	0.070	0.0049	2.4 × 10 ⁻⁵	5.8 × 10 ⁻¹⁰
30	0.020	0.14	0.018	0.00033	1.1 × 10 ⁻⁷	1.3 × 10 ⁻¹⁴
50	0.033	0.035	0.0012	1.5 × 10 ⁻⁶	2.2 × 10 ⁻¹²	4.8 × 10 ⁻²⁴

Table 8.3.1 outlines the average number of nucleotide substitutions per template as a function of the number of EP-PCR doublings and the length of the template. Table 8.3.2 shows what fraction of the resulting products will be completely free from mutation.

In most cases, the mutagenized DNA of interest will encode a protein. The fraction of mutated amino acids will be higher than the fraction of mutated nucleotides by a factor of ~2.2. This is because a mutation in any of the three positions of a codon may result in an amino acid substitution. If the initial template is a random open reading frame (i.e., equal probability of each nucleotide at each position in each codon), mutation at the first position of a codon will cause an amino acid change 96% of the time; mutation at the second and third positions will cause amino acid changes 100% and 23% of the time, respectively (as calculated using the mutation frequencies in Table 8.3.6; see Anticipated Results).

MUTAGENESIS OF A DNA SEQUENCE

In this protocol, a 400-bp DNA sequence is mutagenized for ten doublings, to achieve a mutation rate of 0.66% per nucleotide position.

Materials

- 100 mM Tris·Cl, pH 8.3 (*APPENDIX 2*)
- 2 M KCl
- 200 mM MgCl₂
- 25 mM dCTP, pH ~7
- 25 mM dTTP, pH ~7
- 5 mM dATP, pH ~7
- 5 mM dGTP, pH ~7
- 100 μM each 5' and 3' PCR primers
- 200 pg/μl DNA template (400 bp in length)
- 25 mM MnCl₂
- 5 U/μl *Taq* DNA polymerase
- 100-μl PCR tubes (Sarstedt)
- Thermal cycler (see *UNIT 15.1*)
- TOPO T/A cloning kit (Invitrogen)
- QIAprep kit (Qiagen)
- Additional reagents and equipment for PCR amplification (*UNIT 15.1*) and agarose gel electrophoresis (*UNIT 2.7*)

1. Make up the following PCR reaction mixture in a 100-μl PCR tube on ice:

- 51 μl water
- 10 μl 100 mM Tris·Cl, pH 8.3 (10 mM final)
- 2.5 μl 2 M KCl (50 mM final)
- 3.5 μl 200 mM MgCl₂ (7 mM final)
- 4 μl 25 mM dCTP (1 mM final)
- 4 μl 25 mM dTTP (1 mM final)
- 4 μl 5 mM dATP (0.2 mM final)
- 4 μl 5 mM dGTP (0.2 mM final)
- 2 μl 100 μM 5' primer (2 μM final)
- 2 μl 100 μM 3' primer (2 μM final)
- 10 μl 200 pg/μl template DNA (20 pg/μl final)
- 2 μl 25 mM MnCl₂ (0.5 mM final)
- 1 μl 5 U/μl *Taq* DNA polymerase (0.05 U/μl final)
- Total volume, 100 μl.

BASIC PROTOCOL

Mutagenesis of Cloned DNA

8.3.3

The MnCl₂ should be added immediately before the thermal cycling is initiated. The Taq DNA polymerase should be added when the thermal cycling reaction has reached the first annealing step.

2. Place the tube in the thermal cycler and perform ~12 PCR cycles (UNIT 15.1), or enough to obtain a 1000-fold (10 doublings) increase in the amount of PCR product relative to the input template.

The cycling conditions will vary depending on the template and primers, but reasonable starting conditions are: 94°C for 1 min (denaturation), 60°C for 1 min (annealing), and 72°C for 3 min (extension).

The annealing temperature should be kept >50°C, if possible, to avoid mispriming, the frequency of which increases at the high divalent cation concentration used for EP-PCR. The 3-min extension time reduces the selective amplification of shorter, undesirable sequences produced by mispriming.

3. Run an ethidium bromide-containing agarose gel to confirm the amount and correct molecular weight of the product (UNIT 2.7).
4. Clone and sequence a sample of the resulting PCR DNA to determine the frequency of mutations in the product using the TOPO T/A cloning kit and the QIAprep kit.

For information on DNA sequencing, see Chapter 7.

To achieve higher levels of mutagenesis, the template in the initial reaction will need to be diluted to a greater extent, and more EP-PCR cycles will need to be performed. Also, if more than ~15 cycles of EP-PCR are to be performed, a fresh aliquot of Taq polymerase should be added after the 15th cycle.

One problem that often occurs when attempting to achieve a large number of doublings is that PCR products that are smaller than the desired one “take over” the PCR reaction. If this happens, one should first make sure that the EP-PCR conditions are optimized, resulting in an increase in DNA product of at least 1.7-fold per cycle. This may require increasing the extension time to over 3 min, especially when the desired product is >1 kb. Another way to avoid conditions that selectively amplify shorter templates is to increase the denaturing time (up to 75 sec). Also, the annealing temperature should be as high as possible to minimize the occurrence of mispriming events. The highest annealing temperature that gives efficient amplification must be determined empirically. UNIT 15.1 describes other parameters that can be varied to optimize the amplification reaction. If these measures still fail to eliminate the problem, it may be necessary to perform a smaller number of cycles (using a higher starting template concentration), and then gel purify the full-length PCR product before continuing with more thermal cycling. In some cases, it may be necessary to perform this gel purification step periodically (e.g., every 8 cycles). One can use agarose gel purification (UNIT 2.7) which is very easy, sensitive, and convenient. However, polyacrylamide gel electrophoresis (PAGE; UNIT 2.7) can accomplish a higher degree of purification.

MUTAGENIZING A LIBRARY OF SEQUENCES

Sometimes it is desirable to mutagenize an entire collection of sequences simultaneously. The Basic Protocol is appropriate in cases where the starting template is a unique sequence, but the following modifications are recommended when the starting template is itself a library.

The Basic Protocol calls for the EP-PCR reaction to be initiated with a very small amount of template, but this amount may be insufficient to preserve the initial library complexity. To avoid complexity loss before and during the amplification process, one can start with a comparatively large template concentration and perform only four EP-PCR cycles, and then transfer ~10% of the resulting material into a fresh EP-PCR reaction. These “serial transfers” are continued until the desired number of doublings is achieved. One additional advantage of this method is that the progress of the EP-PCR reaction can be monitored throughout the entire procedure, a luxury that is not possible using the Basic Protocol.

This protocol will give ~50 EP-PCR doublings and results in mutations in ~3.5% of the nucleotide positions in the DNA template. However, the actual mutagenic rate may vary with conditions and template.

Additional Materials (also see Basic Protocol)

30 ng/μl DNA template (library)

1. Make up the following EP-PCR reaction mixture on ice:

960 μl water
150 μl 100 mM Tris-Cl, pH 8.3 (10 mM final)
37.5 μl 2 M KCl (50 mM final)
52.5 μl 200 mM MgCl₂ (7 mM final)
60 μl 25 mM dCTP (1 mM final)
60 μl 25 mM dTTP (1 mM final)
60 μl 5 mM dATP (0.2 mM final)
60 μl 5 mM dGTP (0.2 mM final)
30 μl 100 μM 5' primer (2 μM final)
30 μl 100 μM 3' primer (2 μM final)
Total volume, 1500 μl.

2. Divide the EP-PCR reaction mixture into 16 labeled tubes suitable for 100 μl PCR reactions (i.e., into 90-μl aliquots).

These may be stored at 4°C for a few hours.

3. Add 7 μl of the 30 ng/μl DNA library to tube 1 to give ~2 ng/μl. Place tube 1 in the thermal cycler; once it has reached the annealing temperature, add the following and mix:

2 μl 25 mM MnCl₂ (0.5 mM final)
1 μl 5 U/μl *Taq* DNA polymerase (0.05 U/μl final).

ALTERNATE PROTOCOL

4. Perform four cycles of EP-PCR amplification (see Basic Protocol, step 2 annotations, for discussion of conditions). During the final extension at 72°C, place tube 2, containing the fresh EP-PCR, mixture into the same PCR block. Before the final extension is completed on tube 1, but ensuring that tube 2 has reached the extension temperature, transfer 10 µl of EP-PCR reaction mixture from tube 1 into tube 2, and then add the following to tube 2 and mix:

2 µl 25 mM MnCl₂ (0.5 mM final)
1 µl 5 U/µl *Taq* DNA polymerase (0.05 U/µl final).

Remove tube 1 from the block and store at 4°C.

The numbers given here for starting DNA template concentration and transfer volume may need to be modified in accordance with results from pilot EP-PCR reactions, which serve to determine the amplification efficiency.

5. Repeat step 4 for the remaining 14 tubes. Analyze the PCR reaction using agarose gel electrophoresis (UNIT 2.7) after every fourth transfer, and quantitate the bands in successive PCR amplifications.

The DNA amplification per EP-PCR cycle should not decrease to below 1.7, even for the fourth cycle. It is also important that the amount of DNA at the end of the four EP-PCR cycles not increase from transfer to transfer. If this does occur, reduce the transfer volume.

Before the entire EP-PCR protocol is attempted, it is important to pilot the EP-PCR conditions to ensure that low-molecular-weight PCR products are not “taking over” the reaction, and that the amplification per cycle is at least 1.7. The optimal PCR amplification conditions may be different from normal PCR amplification performed upon the same library.

This serial transfer approach yields a succession of samples with increasing levels of mutagenesis. If one is uncertain about the optimal level of mutagenesis for a particular application, the samples from different stages of the EP-PCR procedure can be mixed prior to screening or selection.

COMMENTARY

Background Information

The simplest and most versatile method for introducing random point mutations into a DNA sequence is to take advantage of the inherently low fidelity of *Taq* DNA polymerase. This enzyme can misincorporate with a frequency as high as 0.02% per position (Eckert and Kunkel, 1991). By performing exponential amplification through PCR, these errors will accumulate. Leung et al. (1989) showed that the fidelity of copying by *Taq* polymerase can be further decreased by skewing the relative dNTP concentrations, using a high Mg²⁺ concentration, and including Mn²⁺ in the reaction mixture. Cadwell and Joyce (1992) used a similar strategy to obtain an error rate per nucleotide position of about 0.066% per PCR cycle. This protocol, which forms the basis of that described here, has many useful qualities: (1) it is

extremely simple in that it does not require any unusual reagents such as nucleotide analogs; (2) the EP-PCR efficiency is comparable to that for standard PCR conditions; (3) in many instances the desired level of mutagenesis can be obtained in ~1 hr with a single PCR reaction; and (4) a wide spectrum of nucleotide substitutions is generated. Alternative, more complex procedures also exist that utilize multiple parallel EP-PCR reaction conditions with different dNTP ratios, or use nucleotide analogs (Spee et al., 1993; Fromant et al., 1995; Xu et al., 1999). Vartanian et al. (1996) have described a protocol that can result in very high mutagenic levels.

Critical Parameters

The key to success in this technique is to characterize the amplification efficiency of the

Table 8.2.3 Troubleshooting Guide for Random Mutagenesis by PCR

Problem	Possible cause	Solution
No EP-PCR product	EP-PCR conditions need optimizing	Optimize pilot EP-PCR reaction, especially with regard to annealing temperature. Use long extension times (at least 3 min). If this fails, test primers, template, and other reagents under “normal” PCR conditions to ensure that there has not been a primer design or synthesis error, or a degeneration in reagent quality.
Multiple EP-PCR products of incorrect lengths	EP-PCR conditions need optimizing	Optimize pilot EP-PCR reaction, as above. If this fails to solve the problem, periodically gel purify the product of correct length.
Brown precipitate observed in EP-PCR reaction mixture	Manganese salts are precipitating out of solution	Add manganese to EP-PCR reaction mixture just prior to thermal cycling.
Successive transfers contain decreasing amounts of DNA when visualized on an agarose gel	Transfer volume is too small	Increase transfer volume
Successive transfers contain increasing amounts of DNA when visualized on an agarose gel	Transfer volume is too large, possibly because the efficiency of the PCR reaction is increasing as the most easily amplified sequences begin to dominate the mixture	Decrease transfer volume

EP-PCR reaction and confirm that the DNA concentration is increasing by at least 1.7-fold per cycle. Once the amplification efficiency is known, it will be easy to calculate the number of EP-PCR cycles that must be performed in order to achieve the desired level of mutagenesis. These issues are discussed in detail above (see Strategic Planning).

Troubleshooting

Guidelines for troubleshooting experimental difficulties are presented in Table 8.2.3.

Anticipated Results

The degree of mutagenesis achieved is a function of the number of EP-PCR doublings performed, and therefore can be controlled by the investigator. Anticipated levels of mutagenesis at the DNA, and when applicable, the protein level, are described in Tables 8.3.1, 8.3.2, 8.3.3, 8.3.4, and 8.3.5, and discussed above in Strategic Planning. It is important,

however, to verify that the intended degree of mutagenesis has been achieved by sequencing a few of the resulting, cloned DNA molecules.

The EP-PCR protocol produces all types of substitution mutations, but the distribution is highly nonrandom. The results obtained by the authors in one experiment are shown in Table 8.3.6. It should be noted that in the EP-PCR reaction, both top and bottom DNA strands are equally subject to mutagenesis, so mutations from G to A and from C to T, for example, cannot be distinguished from each other; consequently these are combined together in the table.

Time Considerations

The amount of time required to determine the amplification efficiency of the EP-PCR is 3 hr. If necessary, optimizing the conditions to improve the EP-PCR efficiency requires 1 day. Performing the actual EP-PCR and verifying the total number of doublings achieved takes

Table 8.3.4 Average Number of Amino Acid Mutations per Open Reading Frame (ORF) as a Function of ORF Length and Number of EP-PCR Doublings

EP-PCR doublings	Mutations per codon	ORF length				
		100 bp	200 bp	400 bp	800 bp	1600 bp
5	0.0076	0.25	0.50	1.0	2.0	4.0
10	0.015	0.50	1.0	2.0	4.0	8.0
20	0.030	1.0	2.0	4.0	8.1	16
30	0.045	1.5	3.0	6.0	12	24
50	0.076	2.5	5.0	10	20	40

Table 8.3.5 Fraction of ORF's Encoding Wild-Type Polypeptide as a Function of ORF Length and the Number of EP-PCR Doublings

EP-PCR doublings	Mutations per codon	ORF length				
		100 bp	200 bp	400 bp	800 bp	1600 bp
5	0.0076	0.78	0.60	0.36	0.13	0.017
10	0.015	0.60	0.36	0.13	0.017	0.00030
20	0.030	0.36	0.13	0.017	0.00028	7.8×10^{-8}
30	0.045	0.21	0.045	0.0021	4.3×10^{-6}	1.8×10^{-11}
50	0.076	0.073	0.0053	2.8×10^{-5}	7.9×10^{-10}	6.3×10^{-19}

Table 8.3.6 Observation of Each Type of Substitution in a Collection of 97 Mutations Generated by EP-PCR

Type of mutation	Number times observed
A→T and T→A	34
G→A and C→T	26
A→G and T→C	24
A→C and T→G	6
G→C and C→G	5
G→T and C→A	2

anywhere between 2 hr and 1 day, depending on the number of EP-PCR cycles performed. It may take one extra day if periodic gel purification of full-length EP-PCR product must be performed. Cloning a sample of the EP-PCR product may be carried out in 1 day. DNA minipreps for sequencing may be prepared in 1 day. Sequencing the DNA to determine the level of mutagenesis may require 1 to 5 days, depending on the sequencing facility. Analyzing the sequencing results and calculating the degree of mutagenesis takes 1 day. If the level of mutagenesis is significantly different from what was desired, one may have to repeat the EP-PCR, then resequence and determine the level of mutagenesis once again.

Literature Cited

- Cadwell, R.C. and Joyce, G.F. 1992. Randomization of genes by PCR mutagenesis. *PCR Meth. Appl.* 2:28-33.
- Eckert, K.A. and Kunkel, T.A. 1991. DNA polymerase fidelity and the polymerase chain reaction. *PCR Meth. Appl.* 1:17-24.
- Fromant, M., Blanquet, S., and Plateau, P. 1995. Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal. Biochem.* 224:347-353.
- Leung, D.W., Chen, E., and Goeddel, D.V. 1989. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* 1:11-15.
- Spee, J.H., de Vos, W.M., and Kuipers, O.P. 1993. Efficient random mutagenesis method with adjustable mutation frequency by use of PCR and dITP. *Nucleic Acids Res.* 21:777-778.
- Vartanian, J.-P., Henry, M., and Wain-Hobson, S. 1996. Hypermutagenic PCR involving all four transitions and a sizeable proportion of transversions. *Nucleic Acids Res.* 24:2627-2631.
- Xu, H., Petersen, E.I., Petersen, S.B., and El-Gewely, M.R. 1999. Random mutagenesis libraries: Optimization and simplification by PCR. *BioTechniques* 27:1102-1108.

Contributed by David S. Wilson
Zyomyx
Hayward, California

Anthony D. Keefe
Massachusetts General Hospital
Boston, Massachusetts

Linker-Scanning Mutagenesis of DNA

Two protocols are described in which clusters of point mutations are introduced throughout a sequence of interest that has been cloned into a plasmid vector. The first protocol uses complementary oligonucleotides and requires a unique restriction site adjacent to the region that is to be mutagenized. A nested series of deletion mutations is first generated in the region. A pair of complementary oligonucleotides are synthesized to fill in the gap in the sequence of interest between the linker at the deletion endpoint and the nearby restriction site (Fig. 8.4.1). The linker sequence actually provides the desired clusters of point mutations as it is moved or “scanned” across the region by its position at the varied endpoints of the deletion mutation series. An alternate protocol makes use of site-directed mutagenesis procedures (UNIT 8.1) to introduce smaller clusters of point mutations throughout the target region.

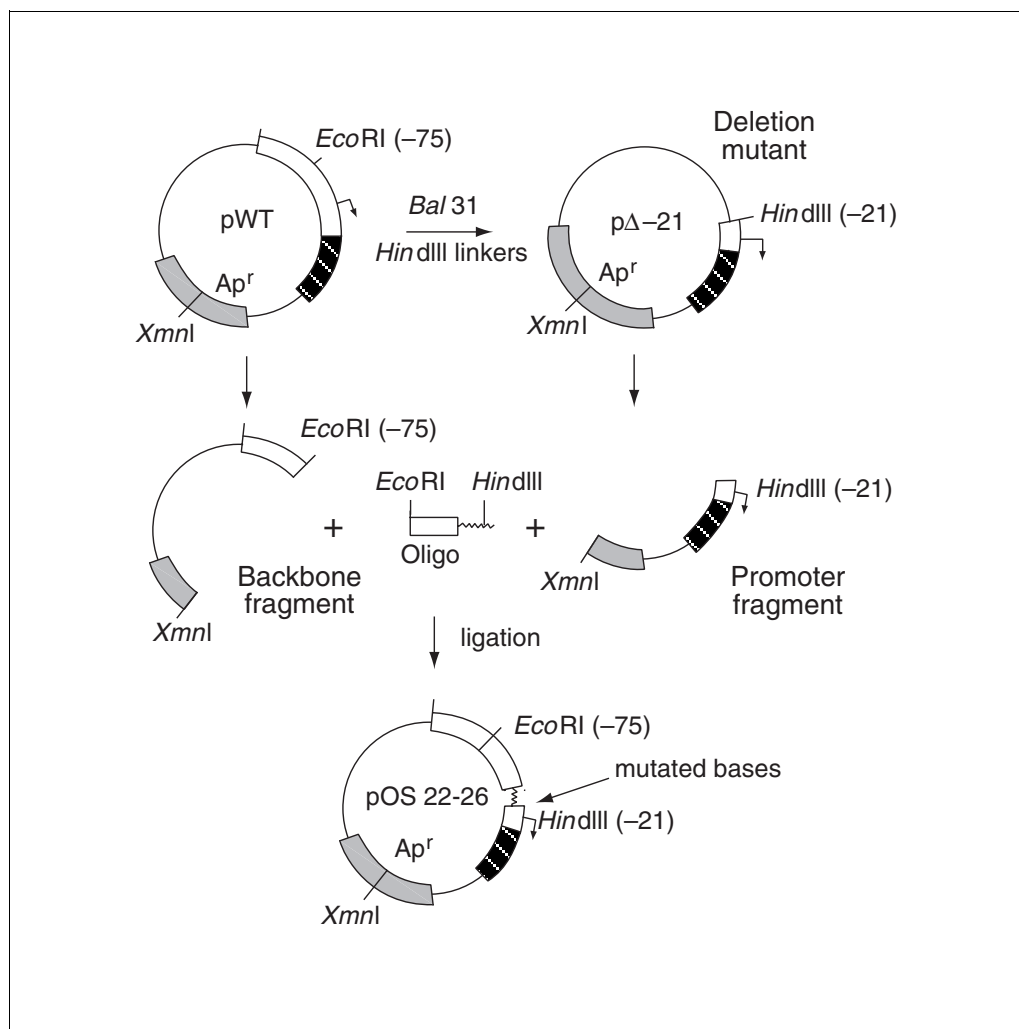


Figure 8.4.1 Assembly of linker-scan mutations using basic protocol. pWT is linearized near the region of interest at a site such as the *Eco*RI site at -75. *Bal* 31 is used to create deletions and *Hind*III linkers are added to the deletion endpoints. The DNA is digested with *Hind*III and *Xmn*I, and then is ligated to the complementary *Xmn*-*Hind*III vector fragment from pWT to form the deletion mutation pΔ -21. The backbone fragment is prepared from pWT by an *Xmn*-*Eco*RI digest and the promoter fragment is prepared from pΔ -21 by a *Hind*III-*Xmn* digest. After isolation, the two fragments are ligated together with the hybridized oligonucleotides to create the linker-scan mutation pOS 22-26. This plasmid is identical to pWT except for the mutated bases at the site of the *Hind*III linker at -21.

LINKER SCANNING USING NESTED DELETIONS AND COMPLEMENTARY OLIGONUCLEOTIDES

This protocol can be used to generate clusters of point mutations throughout a promoter region or throughout any other sequence element involved in replication, splicing, or protein structure. The concept is to make a series of deletion mutations whose nested ends will serve to “move” a linker sequence throughout the region of interest. The deleted sequences are then filled in using synthetic oligonucleotides to regenerate the original sequence at every position except for the bases mutated by the linker (Fig. 8.4.2).

Materials

- High-copy-number plasmid vector (UNIT 1.5)
- Linkers (UNIT 3.16, Example 3.16.6)
- Bal 31 nuclease (UNIT 3.12) or exonuclease III and S1 nuclease (UNITS 3.11 & 3.12, respectively)
- Restriction endonucleases and buffers (UNIT 3.1)
- 0.5 M EDTA, pH 8 (optional; APPENDIX 2)
- 100% ethanol
- Low gelling/melting temperature agarose (UNIT 2.6)
- Competent *E. coli* cells (UNIT 1.8)
- 29:1 acrylamide/bisacrylamide (UNIT 2.7) or sieving agarose (UNIT 2.8)
- Synthetic oligonucleotides for the gene of interest

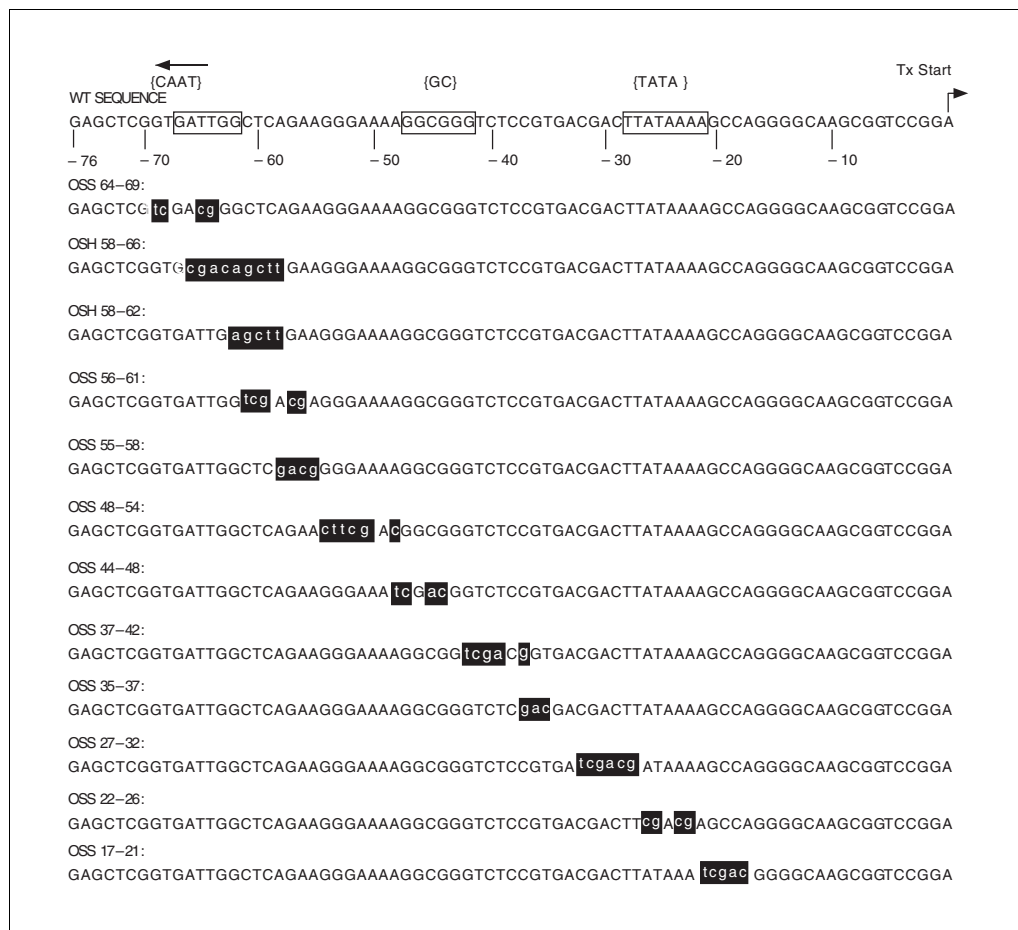


Figure 8.4.2 Linker-scanning mutagenesis of a human *hsp70* gene promoter region (Greene et al., 1987). Wild-type sequence is shown at top. Bases shown in black boxes are those changed by mutagenesis.

TE buffer (*APPENDIX 2*) containing 150 mM NaCl
T4 DNA ligase (measured in cohesive-end units; *UNIT 3.14*)
10× T4 DNA ligase buffer (*UNIT 3.4*)
Selective plates

Additional reagents and equipment for creating nested sets of deletions (*UNIT 7.2*), ethanol precipitation (*UNIT 2.1*), ligating DNA fragments (*UNIT 3.16*), isolating DNA fragments using low gelling/melting temperature agarose (*UNIT 2.6*), preparing and transforming competent *E. coli* cells (*UNIT 1.8*), minipreps of plasmid DNA (*UNIT 1.6*), nondenaturing polyacrylamide gel electrophoresis (*UNIT 2.7*) or sieving agarose gel electrophoresis (*UNIT 2.8*), and DNA sequencing (Chapter 7)

Create a nested set of deletions

1. Generate a nested set of either 5' or 3' deletion mutations in the region of interest in your plasmid. Either *Bal* 31 or *exo* III and S1 nuclease can be used (see *UNIT 7.2*). First, linearize the plasmid with a restriction enzyme for subsequent nuclease digestion; any restriction site that cuts near the sequence of interest (such as the *Eco*RI site at -75 in Fig. 8.4.1) can be used (it is not necessary to use the site defining one end of the insert). For *Bal* 31 digestion, it is necessary to determine the rate of nuclease digestion and adjust conditions (temperature, time of digestion, and amount of nuclease; see *UNIT 3.12*) in trial runs so that the time points taken contain deletion endpoints every six to eight base pairs and span the region to be scanned in order to make a complete set of scan mutations. The extent of nuclease digestion can be monitored by running aliquots of each time point on an agarose gel with appropriate size markers. Stop the digestions by heating to 65°C or by adding EDTA to 50 mM. Precipitate the DNA with 100% ethanol.

You may wish to repair the ends of the DNA as mentioned in UNIT 3.12 to improve subsequent ligation frequencies.

Any vector can be used in cloning the sequence of interest, although a vector such as pUC19 is preferable because of its ease of sequencing and high copy number. See UNIT 7.2 for a full discussion and complete steps for making nested sets of deletions.

2. Ligate a synthetic linker to the ends of the DNA obtained from the appropriate time points (see Example 3.16.6), preferably choosing one denoting a unique restriction site in the sequence of interest but present in the vector. A site in the polylinker is ideal if cloning with pUC19 (*Hind*III in Fig. 8.4.1). Digest with an excess of the restriction enzyme that recognizes the linker in order to leave a cohesive end as well as with another enzyme cutting in the vector sequences—thus yielding fragments with the linker at the deleted end and a vector site at the other.
3. Isolate the desired fragments using low gelling/melting temperature agarose or another method of choice (*UNIT 2.6*), then ligate them to a vector fragment to create intact plasmids containing the deletion mutations (*UNIT 3.16*).
4. Transform competent *E. coli* cells with the ligation mix (*UNIT 1.8*), pick colonies, and prepare minipreps of plasmid DNA (*UNIT 1.6*).
5. Digest the miniprep DNA with appropriate restriction enzymes to yield a small DNA fragment with one terminus at the deletion endpoint. Run the digest on a nondenaturing polyacrylamide gel (*UNIT 2.7*) or sieving agarose gel (*UNIT 2.8*) to determine the fragment size. A small fragment (200 to 300 bp) is necessary to accurately estimate the location of the endpoint of the deletion for each individual mutation.

6. When a satisfactory set of deletions has been obtained that spans the region of interest, sequence each deletion mutation to determine the exact deletion endpoint (Chapter 7).

This is most easily done if the deletion is adjacent to a vector sequence for which an oligonucleotide primer is available for dideoxy sequencing (e.g., the polylinker of pUC19).

Design complementary oligonucleotides

7. Design complementary oligonucleotides to restore the wild-type sequence between the linker present at each deletion endpoint and a nearby restriction site (again, preferably unique to the plasmid) upstream of the region for a 5' deletion series or downstream of the region for a 3' deletion series (*EcoRI* was chosen in Fig. 8.4.1).

*The site must be located so that the oligonucleotides are of a reasonable length, i.e., <100 bases. Remember to include the cohesive ends for both sites in the oligos (*EcoRI* and *HindIII* in Fig. 8.4.1).*

Synthesis of all of the oligonucleotides required to make a complete set of linker-scan mutations can be very expensive. This cost can be reduced almost in half by using mutually primed synthesis (UNIT 8.2; Fig. 8.2.2) to produce the oligonucleotides.

Isolate DNA fragments

8. Make the following fragments for ligation to the oligonucleotides: (a) a fragment spanning the linker site and a convenient site in the plasmid's drug marker ("promoter" fragment; *Hind-Xmn* in Fig. 8.4.1), cut from each deletion mutation; and (b) a fragment spanning the site in the drug marker and the restriction site near the region to be scanned ("backbone" fragment; *Xmn-Eco* in Fig. 8.4.1), cut from the wild-type parental plasmid. Isolate the necessary fragments by one of the methods in UNIT 2.6.

A site in the drug marker is preferred because it provides "forced cloning"—i.e., the site must be regenerated to allow expression of the selective marker.

Hybridize complementary oligonucleotides

9. Resuspend each oligonucleotide strand at ~100 µg/ml in TE buffer containing 150 mM NaCl. Mix equimolar amounts of each strand and heat 10 min at 65°C.
10. Cool slowly to room temperature by placing the tubes at room temperature for 20 min. Check the OD₂₆₀ to determine final concentration. Also, check to be sure that the oligonucleotides have completely hybridized by running aliquots of each complementary strand as well as the hybrid on a 4% sieving agarose gel.

Assemble linker-scan mutations

11. Set up the three-part ligation (Fig. 8.4.1), as exemplified by the following: Mix equimolar amounts of the promoter and backbone fragments with a 50-fold molar excess of the hybridized oligonucleotides relative to the plasmid fragments (UNIT 3.16); use ~400 ng total DNA in 10 µl, for a final DNA concentration of 40 µg/ml. Add 0.4 µl (160 cohesive-end units) T4 DNA ligase and incubate 2 hr at 15°C. If a blunt site is being used at the promoter-backbone fragment junction (as *Xmn* is used in Fig. 8.4.1), add 1 µl (400 U) ligase and incubate an additional 2 hr at 30°C. Dilute 40-fold with TE buffer to a DNA concentration of 1 µg/ml to promote intramolecular ligation in the next step.

Note that during the ligation the only 5' terminal phosphate groups present are on the backbone and promoter fragments. These phosphate groups are essential for covalent linkage of each DNA strand by ligase (see UNIT 3.14).

- The hybridized oligos do not have 5' terminal phosphates. Therefore, each end of the backbone and promoter fragment will be covalently attached to one strand of a double-stranded oligonucleotide. It is necessary to melt the noncovalently bound strand of each oligonucleotide away from the ligated fragments to allow circularization of the plasmid. To do this, heat the diluted ligation reaction to 65°C for 10 min, then cool to room temperature slowly.

This step should result in the hybridization of the two complementary strands that are covalently bound to the plasmid fragments. The net effect is to incorporate one copy of the hybridized oligos into the ligation product, with one nick on each strand. Tandem repeats of the oligo sequence are thus prevented as all noncovalently bound oligos melt away.

- Transform competent *E. coli* cells with 20 to 25 µl of the diluted ligation reaction (UNIT 1.8) and plate on selective media. Pick colonies and prepare minipreps of plasmid DNA to be checked by restriction digests and by DNA sequencing to be sure the final sequence is as desired—i.e., a restoration of the wild-type sequence except for the six bases changed by the linker sequence at the deletion endpoint.

ALTERNATE PROTOCOL

LINKER SCANNING USING OLIGONUCLEOTIDE-DIRECTED MUTAGENESIS

It is also possible to construct a set of linker-scan mutations using the methods of Kunkel (see UNIT 8.1). In this protocol, mutations are introduced into a sequence by annealing a synthetic oligonucleotide containing one or more mismatches to the sequence of interest cloned into a single-stranded M13 vector. This template is grown in an *E. coli dut⁻ ung⁻* strain, which allows the incorporation of uracil into the template strand. The oligo is annealed to the template and extended with T4 DNA polymerase to create a double-stranded heteroduplex. Finally, the heteroduplex is introduced into a wild-type *E. coli* strain which will prevent replication of the template strand due to the presence of apurinic sites (generated where uracil is incorporated), thereby resulting in plaques containing only mutated DNA.

Additional Materials

- M13 template (UNIT 1.15)
- E. coli dut⁻ ung⁻* strain (UNIT 8.1 and Table 1.4.5)
- T4 DNA polymerase (UNIT 3.5)

Additional reagents and equipment for oligonucleotide-directed mutagenesis without phenotypic selection (UNIT 8.1)

- Design a set of oligonucleotides, each of which has several adjacent bases mutated as a group flanked by wild-type sequences that will allow hybridization to the appropriate region of the gene. As the group of adjacent mismatches are “scanned” through the oligo sequences, they will result in a set of exact linker-scan mutations.

This protocol has the major advantage that it does not require an adjacent restriction site or the construction of a nested set of deletion mutations. One difficulty with using this approach to create linker scans is that it is unclear how many adjacent bases can be stably mismatched at one time. Linker scans have been made via this approach with mismatched bases in groups of three (T. Kadesch, personal communication).

A critical issue in designing the oligonucleotides is the length of the mismatched region. Long mismatches (e.g., 8 to 10 bases) will necessitate fewer constructions to fully mutate a given stretch of DNA. However, these mismatches may form less stable hybrids with the template, and thus may be more difficult to construct. In addition, the longer the mismatched region, the greater the possibility of having

one linker-scan mutation disrupting more than one element. In order to identify all of the functional elements of a region, it is necessary to have null mutations interdigitated between mutations that disrupt function. This allows detection of individual functional elements. The larger the mismatched, or linker-scanned, region is, the lower the probability that this result will be obtained.

2. Hybridize the oligonucleotides to the M13 template, extend with T4 DNA polymerase, and transform wild-type *E. coli* as described by Kunkel (UNIT 8.1). Plate and isolate the double-stranded RF DNA from colonies of each transformation.
3. Sequence the DNA purified from the transformants to verify that the desired mutations have been correctly introduced.

COMMENTARY

Background Information

In mapping sequences important for any function, the first step is usually to make deletion mutations to try to find critical sequences. For more detailed analysis, however, this approach suffers from the drawback that the deletion often removes elements important for function and thus can obscure the individual contribution of each sequence element. By putting in clusters of point mutations, elements outside of the clustered mutations are left intact. Use of clustered mutations minimizes potential artifacts caused by creating new sequences at the junction between vector and the deletion endpoint. In addition, linker-scan mutations do not alter the spatial relationships between sequence elements in the region of interest. Although advanced methods are now available to saturate a region with point mutations (see UNIT 8.3), such protocols require a large amount of time and effort, including DNA sequencing. In many cases it may be preferable to locate and analyze important sequence elements by linker scanning before resorting to point mutagenesis of such elements.

Literature Review

The basic protocol is an adaptation of the original linker scanning protocol developed by McKnight and Kingsbury (1982) to locate sequence elements important for transcription of the HSV *tk* gene. Their procedure uses two sets of nested 5' and 3' deletion mutations with linkers at the endpoints. "Matching" deletions are then combined at the linker to create the scan mutations. The problem with their approach is that it is very hard to get perfectly matching deletions, resulting in the insertion or deletion of one or more bases.

The basic protocol presented here uses only one such set of deletions. Oligos are then used to fill in the gap between the nearby restriction

site and the linker at the deletion endpoint, resulting in exact scan mutations—no bases are inserted or deleted. Exact substitutions are also made by the procedure described in the alternate protocol.

The alternate protocol, based on the procedure by Kunkel (UNIT 8.1), has two major advantages. Fewer oligonucleotides are needed, since it is not necessary to synthesize two complementary oligos as with the basic protocol. Also, it is not necessary to have a restriction site adjacent to the region to be mutagenized. It is not yet clear, however, how many adjacent mismatched bases can be tolerated in the Kunkel procedure. Once this parameter has been optimized, the alternate protocol will most probably be more rapid and less expensive than the basic protocol. Finally, the alternate protocol requires special *E. coli* strains, fine-tuning of hybridization and elongation steps, and—because the procedure is only 50% to 80% efficient—somewhat more screening for correctly introduced mutations than is required for the basic protocol.

Critical Parameters

The main requirement for the use of the basic protocol is to have a restriction site in the sequence nearby the region of interest. The site should leave cohesive ends upon digestion and should preferably be unique. However, as long as the site is not duplicated within any one fragment needed for the construction, this does not present a problem because the fragments are prepared in separate digests. In Figure 8.4.1, the *EcoRI* site used to define one end of the backbone fragment is not unique—two other *EcoRI* sites are also present further downstream. Fortunately, these sites are contained in the promoter fragment, which is prepared by an *Xmn-HindIII* digest, and thus do not present a problem in the construction.

Although Figure 8.4.1 depicts a linker-scan mutation with a 5' deletion used in the construction, there is no reason why a 3' deletion cannot be used as well. All that is required is a restriction site nearby but downstream of the region of interest, instead of upstream as for the 5' deletion.

Selection of the linker or of the mismatched bases in the two protocols must be made very carefully. It is possible for a linker sequence to mimic a functional sequence element upon fusion to certain sequences at the deletion endpoint, which could introduce artifacts into an entire set of constructs. A linker should be chosen that does not resemble any known functional element. See also the discussion in the annotation to step 1 of alternate protocol.

The amount of sequence that can be scanned using the basic protocol is limited by the length of oligos that can be synthesized. Very little oligo is required for this protocol, but yields of oligos greater than 100 bases in length are very low and the expense of synthesizing these oligos is very high. Conversely, if an oligo is too short (<6 bases) it will not stably hybridize to its complementary oligo. As long as one strand can "bridge" the two restriction sites on either side of the gap to be filled, the resulting single-stranded gap can be filled in after the ligation reaction with the Klenow fragment of *E. coli* DNA polymerase I (*UNIT 3.5*). The alternate protocol has no such limit, as all the required oligonucleotides differ in sequence but not in length.

Most of the problems that occur in making linker scans with either protocol can be traced to incorrect oligonucleotides. It is best to have the oligos gel-purified and as clean as possible for optimal ligations. Accurate sequencing of both the deletion endpoints and of the final constructs is critical to obtain accurate data.

Anticipated Results

The three-part ligation reaction usually results in 20 to 30 colonies after transforming and plating 5% of the ligation reaction. Even if the resulting linker-scan mutation appears correct

upon restriction digest, it is best to sequence the constructs to be sure the structure is completely accurate. With a well spaced set of deletion mutations, it is possible to change every base of a region in 4- to 6-base groups.

As noted above, the alternate protocol is only 50% to 80% efficient in introducing the desired mutations into the construct. In contrast, virtually all colonies obtained using the basic protocol have the correct structure.

Time Considerations

Generation of a deletion series and sequencing the endpoints can be accomplished in ~2 weeks using the procedures cross-referenced in this manual. The time required to synthesize all of the necessary oligonucleotides is usually the rate-limiting step for either protocol. Once all of the oligonucleotides are obtained and hybridized, the ligations, minipreps, and final sequencing require only 3 to 4 days.

The alternate protocol will require approximately the same amount of time as the basic protocol, depending on the size of the region to be scanned and on the actual efficiency of the procedure. Both of these factors contribute to the number of colonies that must be screened by DNA sequencing.

Literature Cited

- Greene, J.M., Larin, Z., Taylor, I.C.A., Prentice, H., Gwinn, K.A., and Kingston, R.E. 1987. Multiple basal elements of a human *hsp70* promoter function differently in human and rodent cell lines. *Mol. Cell. Biol.* 7:3646-3655.
- McKnight, S.L. and Kingsbury, R. 1982. Transcriptional control signals of a eukaryotic protein-coding gene. *Science* 217:316-324.

Key Reference

- McKnight and Kingsbury, 1982. See above.
Describes the basic concept of a linker scan.

Contributed by John M. Greene
Massachusetts General Hospital
Boston, Massachusetts

Directed Mutagenesis Using the Polymerase Chain Reaction

UNIT 8.5

The polymerase chain reaction (PCR) is most often used for the enzymatic amplification and direct sequencing of small quantities of nucleic acids (see Chapter 15). This technology can also be used as a quick and efficient method for introducing any desired sequence change into the DNA of interest.

This unit contains two basic protocols for introducing base changes into specific DNA sequences. Basic Protocol 1 describes the incorporation of a restriction site and Basic Protocol 2 details the generation of specific point mutations. An alternate protocol describes generating point mutations by sequential PCR steps. Although the general procedure is the same in all three protocols, there are differences in the design of the synthetic oligonucleotide primers and in the subsequent cloning and analyses of the amplified fragments.

INTRODUCTION OF RESTRICTION ENDONUCLEASE SITES BY PCR

**BASIC
PROTOCOL 1**

In this protocol, synthetic oligonucleotides incorporating the desired restriction site are used in conjunction with the appropriate flanking sequence primers to amplify two fragments of DNA. Each of these amplified fragments will contain the new restriction site at one end. Following enzymatic digestion at both the new and flanking restriction sites, the amplified fragments are ligated and subcloned into a vector to facilitate sequence analysis. The procedure is summarized in Figure 8.5.1.

Materials

- DNA sample to be mutagenized
- pUC19 plasmid vector (Figure 1.5.2) or similar high-copy-number plasmid having M13 flanking primer sequences
- TE buffer (*APPENDIX 2*)
- 10× MgCl₂-free PCR amplification buffer (*UNIT 15.1*) supplemented with MgCl₂ as appropriate (see step 5)
- 2 mM 4dNTP mix (*UNIT 15.1*)
- 500 ng/μl (100 pmol/μl) M13 forward and reverse flanking sequence primers (New England Biolabs)
- 5 U/μl *Taq* DNA polymerase (*UNITS 15.1 & 3.5*)
- Mineral oil
- Chloroform (*UNIT 2.1*)
- Buffered phenol (*UNIT 2.1*)
- 100% ethanol
- Appropriate restriction endonucleases (Table 8.5.1)
- 500-μl microcentrifuge tube
- Automated thermal cycler
- Additional reagents and equipment for subcloning DNA (*UNIT 3.16*), plasmid DNA miniprep (*UNIT 1.6*), synthesis and purification of oligonucleotides (*UNITS 2.11 & 2.12*), PCR amplification (*UNIT 15.1*), DNA extraction and precipitation (*UNIT 2.1*), quantitation of DNA by absorbance spectrometry (*APPENDIX 3D*), restriction endonuclease digestion (*UNIT 3.1*), agarose and polyacrylamide gel electrophoresis of DNA (*UNITS 2.5 & 2.7*), purification of DNA from low gelling/melting agarose gels (*UNIT 2.6*), ligation of DNA fragments (*UNIT 3.16*), transformation of *E. coli* (*UNIT 1.8*), and DNA sequence analysis (*UNIT 7.4*)

**Mutagenesis of
Cloned DNA**

8.5.1

Contributed by Brendan Cormack

Current Protocols in Molecular Biology (1997) 8.5.1-8.5.10

Copyright © 1997 by John Wiley & Sons, Inc.

Supplement 37

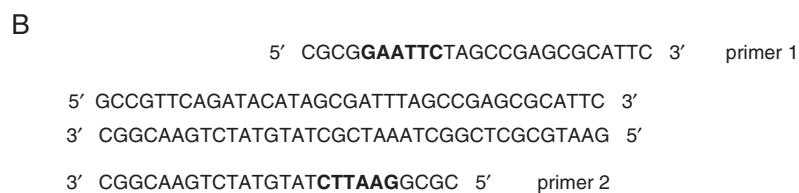
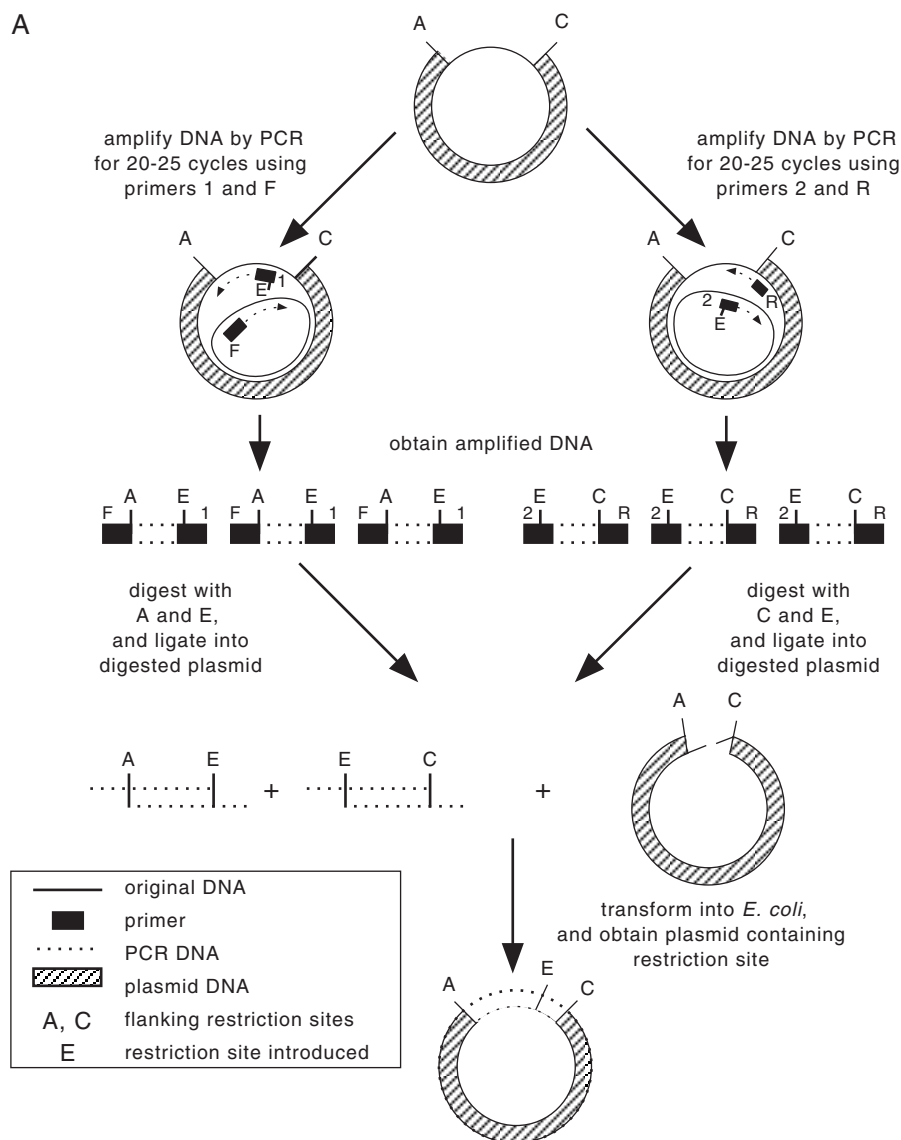


Figure 8.5.1 Introduction of a restriction site into a specific DNA fragment. **(A)** The fragment of interest is cloned into a high-copy-number vector. Sites for two oligonucleotide primers, such as the M13 forward and reverse primers (F and R), flank the cloning site. In two separate reactions, fragments upstream (A—E) and downstream (E—C) of the introduced site are PCR-amplified using the flanking primers and oligonucleotides containing the site to be introduced (primers F and 1, R and 2). The amplified fragments are digested with the appropriate restriction endonucleases, then ligated and subcloned into an appropriately cut vector and transformed into *E. coli*. The resultant plasmid contains an inserted fragment identical to the original DNA except for the introduced restriction site. **(B)** The oligonucleotides needed to change the primary sequence to an *EcoRI* restriction site are indicated. Note the four-base “clamp” sequence at the 5' end of the primers.

Table 8.5.1 Relative Efficiencies of Restriction Enzyme Cleavage When Restriction Site is Near End of DNA Fragment^a

Enzyme	Oligo sequence	% cleavage	
		2 hr	20 hr
<i>Afl</i> III	CCACATGTGG	>90	>90
<i>Asc</i> I	GGCGCGCC	>90	>90
<i>Ava</i> I	CCCCCGGGGG	>90	>90
<i>Bam</i> HI	CGGGATCCCCG	>90	>90
<i>Bgl</i> II	GAAGATCTTC	75	>90
<i>Bss</i> HII	TTGGCGCGCCAA	50	>90
<i>Bst</i> EII	GGGT(A/T)ACCC	0	10
<i>Bst</i> XI	CTGCAGAACCAATGCATTGGATGCAT	25	>90
<i>Cla</i> I	CCATCGATGG	>90	>90
<i>Eco</i> RI	GGAATTCC	>90	>90
<i>Hae</i> III	GGGGCCCC	>90	>90
<i>Hind</i> III	CCCAAGCTTGGG	10	75
<i>Kpn</i> I	GGGGTACCCC	>90	>90
<i>Mlu</i> I	CGACGCGTCG	25	50
<i>Nco</i> I	CATGCCATGGCATG	50	75
<i>Nde</i> I	GGAATTCATATGGAATTCC	75	90
<i>Nhe</i> I	CTAGCTAGCTAG	10	50
<i>Not</i> I	AAGGAAAAAAGCGGCCGCAAAGGAAAA	25	>90
<i>Nsi</i> I	CCAATGCATTGGTTCTGCAGTT	>90	>90
<i>Pac</i> I	CCTTAATTAAGG	0	>90
<i>Pme</i> I	AGCTTTGTTTAAACGGCGCGCCGG	75	>90
<i>Pst</i> I	AAAAC TGCAGCCAATGCATTGGAA	>90	>90
<i>Pvu</i> I	ATCGATCGAT	10	25
<i>Sac</i> I	CGAGCTCG	10	10
<i>Sac</i> II	TCCCCGCGGGGA	50	90
<i>Sal</i> I	ACGCGTCGACGTCGGCCATAGCGGCCGCGGAA	10	75
<i>Sca</i> I	AAAAGTACTTTT	75	75
<i>Sma</i> I	TCCCCGCGGGGA	>90	>90
<i>Spe</i> I	GACTAGTC	10	>90
<i>Sph</i> I	ACATGCATGCATGT	10	50
<i>Stu</i> I	AAGGCCTT	>90	>90
<i>Xba</i> I	GCTCTAGAGC	>90	>90
<i>Xho</i> I	CCGCTCGAGCGG	10	75
<i>Xma</i> I	TCCCCCGGGGGGA	>90	>90

^aReprinted with permission from New England Biolabs.***Prepare the template DNA and oligonucleotide primers***

1. Subclone the DNA to be mutagenized into pUC19 or similar high-copy-number vector, using restriction sites flanking the area to be mutated (see *UNIT 3.16*).

A different restriction site should be used for each end of the fragment to facilitate cloning at various steps in the procedure.

2. Prepare template DNA by plasmid miniprep (*UNIT 1.6*). Resuspend 100 ng of the DNA in TE buffer at a final concentration of 1 ng/μl.

3. Synthesize oligonucleotide primers (primers 1 and 2 in Fig. 8.5.1B; also see *UNIT 2.11*).

The 5' end of each primer consists of the restriction site to be introduced, preceded by three or four extra nucleotides (a "clamp" sequence). The restriction site sequences are followed by ≥ 15 bases that are homologous to the template DNA.

4. Purify synthesized oligonucleotides by denaturing polyacrylamide gel electrophoresis (*UNIT 2.12*). Recover the oligonucleotides (*UNIT 2.1*) and resuspend in 500 μl TE buffer. Determine absorbance at A_{260} (*APPENDIX 3D*) and adjust concentration to 500 ng/ μl (100 pmol/ μl), if necessary.

Because the purity of the oligonucleotides does not seem to affect the PCR reaction significantly, primers may also be used in a reaction immediately after synthesis.

Amplify by PCR

5. Combine the following in each of two 500- μl microcentrifuge tubes, adding oligonucleotides 1 and 2 to separate tubes:

10 μl (10 ng) template DNA
10 μl 10 \times amplification buffer
10 μl 2 mM 4dNTP mix
1 μl (500 ng) oligonucleotide 1 or 2 (1 μM final)
1 μl (500 ng) appropriate M13 flanking sequence primer, forward or reverse (1 μM final)
H₂O to 99.5 μl
0.5 μl *Taq* DNA polymerase (5 U/ μl).

Overlay reaction with 100 μl mineral oil.

The first tube will contain oligonucleotide 1 and its appropriate flanking sequence primer. The second tube will contain oligonucleotide 2 and its appropriate sequence primer (see Fig. 8.5.1A). If the oligonucleotides are not purified, add $1/500$ to $1/2000$ of the oligonucleotide preparation (from step 3) to the reaction. This should represent 500 to 1000 ng (100 to 200 pmol) full-length oligonucleotide.

*A 15-mM concentration of MgCl_2 in the amplification buffer works for many primers; the optimal concentration for a given sequence and primer can be determined as outlined in *UNIT 15.1*. It should be kept in mind, however, that the fidelity of *Taq* DNA polymerase decreases with increasing Mg^{2+} concentration.*

6. Carry out PCR (*UNIT 15.1*) in an automated thermal cycler under the following conditions:

20 to 25 cycles:	45 sec	93°C (denaturation)
	2 min	50°C (annealing)
	2 min	72°C (extension)
1 cycle:	10 min	72°C (extension)

7. Analyze 4 μl of the reaction mix by nondenaturing agarose (*UNIT 2.5A*) or acrylamide gel electrophoresis (*UNIT 2.7*) to verify that the amplification has yielded the predicted product.

*See *UNITS 2.5 & 2.7* to choose the type of gel best suited for the size of the predicted amplification product.*

8. Pipet most of the mineral oil from the reaction mix and extract once with chloroform to remove remaining oil. Extract with buffered phenol and concentrate by precipitation with 100% ethanol (*UNIT 2.1*).

Subclone fragments and check the restriction-site mutation

9. Digest half the amplified DNA with the restriction endonucleases for the flanking and introduced sites. Purify digested fragments on a low gelling/melting agarose gel (UNIT 2.6).
10. Ligate and subclone (UNIT 3.16) both fragments into an appropriately digested vector to obtain a recombinant plasmid containing a single DNA fragment incorporating the new restriction site.

Alternatively, subclone them individually to facilitate the next step.

11. Transform plasmid into *E. coli* (UNIT 1.8). Prepare DNA by plasmid miniprep (UNIT 1.6).
12. Analyze the amplified fragment portion of the plasmid DNA by DNA sequencing (UNIT 7.4) to confirm the addition of the restriction-site mutation.

This is critical because the Taq DNA polymerase can introduce additional mutations into the fragment (see Critical Parameters).

INTRODUCTION OF POINT MUTATIONS BY PCR

In this protocol, synthetic oligonucleotides are designed to incorporate a point mutation at one end of an amplified fragment. Following PCR, the amplified fragments are made blunt-ended by treatment with Klenow fragment. These fragments are then ligated and subcloned into a vector to facilitate sequence analysis. This procedure is summarized in Figure 8.5.2.

Materials

DNA sample to be mutagenized
Klenow fragment of *E. coli* DNA polymerase I (UNIT 3.5)
Appropriate restriction endonuclease (Table 8.5.1)

Additional reagents and equipment for synthesis and purification of oligonucleotides (UNITS 2.11 & 2.12), phosphorylation of oligonucleotides (UNIT 3.10), electrophoresis of DNA on nondenaturing agarose and low gelling/melting agarose gels (UNITS 2.5A & 2.6), restriction endonuclease digestion (UNIT 3.1), ligation of DNA fragments (UNIT 3.16), transformation of *E. coli* (UNIT 1.8), plasmid DNA miniprep (UNIT 1.6), and DNA sequence analysis (UNIT 7.4)

Prepare the template DNA and oligonucleotide primers

1. Prepare template DNA (see Basic Protocol 1, steps 1 and 2).
2. Synthesize (UNIT 2.11) and purify (UNIT 2.12) the oligonucleotide primers (primers 3 and 4 in Fig. 8.5.2B).

The oligonucleotide primers must be homologous to the template DNA for ≥ 15 bases. No four-base "clamp" sequence is added to these primers.

3. Phosphorylate the 5' end of the oligonucleotides (UNIT 3.10).

This step is necessary because the 5' end of the oligonucleotide will be used directly in cloning.

Amplify DNA and prepare blunt-end fragments

4. Amplify the template DNA (see Basic Protocol 1, steps 5 and 6).
5. After the final extension step, add 5 U Klenow fragment to the reaction mix and incubate 15 min at 30°C.

**BASIC
PROTOCOL 2**

**Mutagenesis of
Cloned DNA**

8.5.5

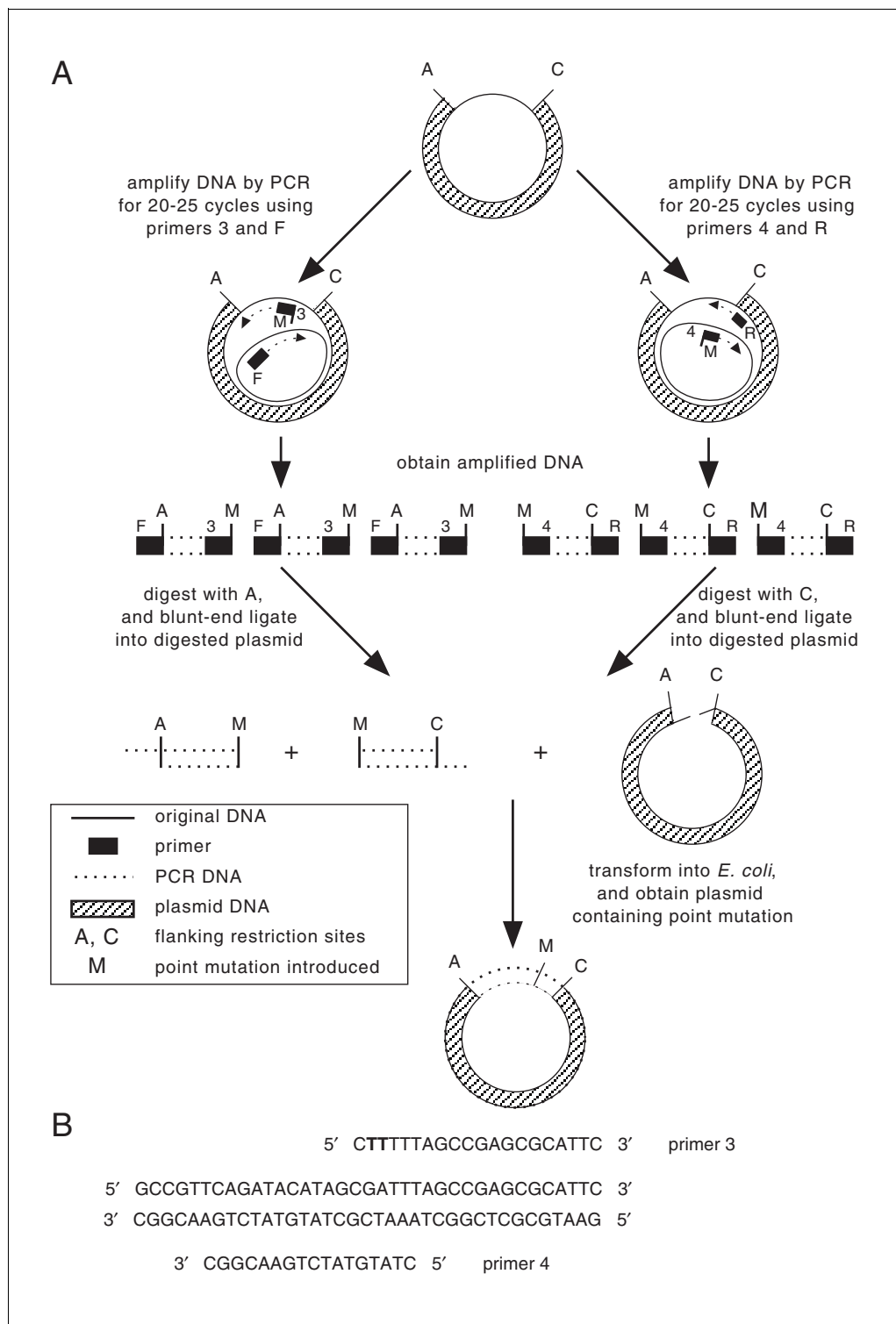


Figure 8.5.2 Introduction of a point mutation into a specific DNA fragment. **(A)** The fragment of interest is cloned into a high-copy-number plasmid vector. Sites for two oligonucleotide primers, such as the M13 forward and reverse primers (F and R), flank the cloning site. In two separate reactions, fragments upstream (A—M) and downstream (M—C) of the point mutation are PCR-amplified using the flanking primers and oligonucleotides containing the point mutation to be introduced (primers F and 3, R and 4). The fragments are digested with the appropriate restriction endonucleases and made blunt-ended. These fragments are ligated and subcloned into an appropriately cut vector and transformed into *E. coli*. The resultant plasmid contains an inserted fragment identical to the original DNA except for the introduced point mutation. **(B)** The oligonucleotides needed to generate a point mutation in the primary sequence are shown. Note that there are no “clamp” sequences at the 5′ ends of the primers.

During PCR, the Taq polymerase adds an extra nontemplated nucleotide to the 3' end of the fragment. The 3' → 5' exonuclease activity of the Klenow fragment is required to make the ends flush and suitable for blunt-end cloning (UNIT 3.5).

6. Analyze and process the reaction mix (see Basic Protocol 1, steps 7 and 8).
7. Digest half the amplified fragments with the restriction endonucleases for the flanking sequences (UNIT 3.1). Purify digested fragments on a low gelling/melting agarose gel (UNIT 2.6).

Subclone fragments and check the point mutation

8. Subclone the two amplified fragments into an appropriately digested vector by blunt-end ligation (UNIT 3.16).
9. Transform recombinant plasmid into *E. coli* (UNIT 1.8). Prepare DNA by plasmid miniprep (UNIT 1.6).
10. Analyze the amplified fragment portion of the plasmid DNA by DNA sequencing to confirm the point mutation (UNIT 7.4).

This is critical because the Taq DNA polymerase can introduce additional mutations into the fragment (see Critical Parameters).

INTRODUCTION OF A POINT MUTATION BY SEQUENTIAL PCR STEPS

ALTERNATE PROTOCOL

In this procedure, the two fragments encompassing the mutation are annealed with each other and extended by mutually primed synthesis; this fragment is then amplified by a second PCR step, thereby avoiding the blunt-end ligation required in Basic Protocol 2. This strategy is outlined in Figure 8.5.3. For materials, see Basic Protocols 1 and 2.

Prepare template DNA and oligonucleotide primers

1. Prepare template DNA (see Basic Protocol 1, steps 1 and 2).
2. Synthesize (UNIT 2.11) and purify (UNIT 2.12) the oligonucleotide primers (primers 5 and 6 in Fig. 8.5.3B).

The oligonucleotides must be homologous to the template for 15 to 20 bases and must overlap with one another by ≥10 bases. The 5' end does not have a "clamp" sequence.

Carry out first PCR amplification

3. Amplify the template DNA and generate blunt-end fragments (see Basic Protocol 2, steps 4 and 5).
4. Purify the fragments by nondenaturing agarose gel electrophoresis (UNIT 2.5A). Resuspend in TE buffer at 1 ng/μl.

Carry out second PCR amplification

5. Combine the following in a 500-μl microcentrifuge tube:
 - 10 μl (10 ng) each amplified fragment
 - 1 μl (500 ng) each flanking sequence primer (each 1 μM final)
 - 10 μl 10× amplification buffer
 - 10 μl 2 mM 4dNTP mix
 - H₂O to 99.5 μl
 - 0.5 μl *Taq* DNA polymerase (5 U/μl).

Overlay with 100 μl mineral oil.

Mutagenesis of Cloned DNA

8.5.7

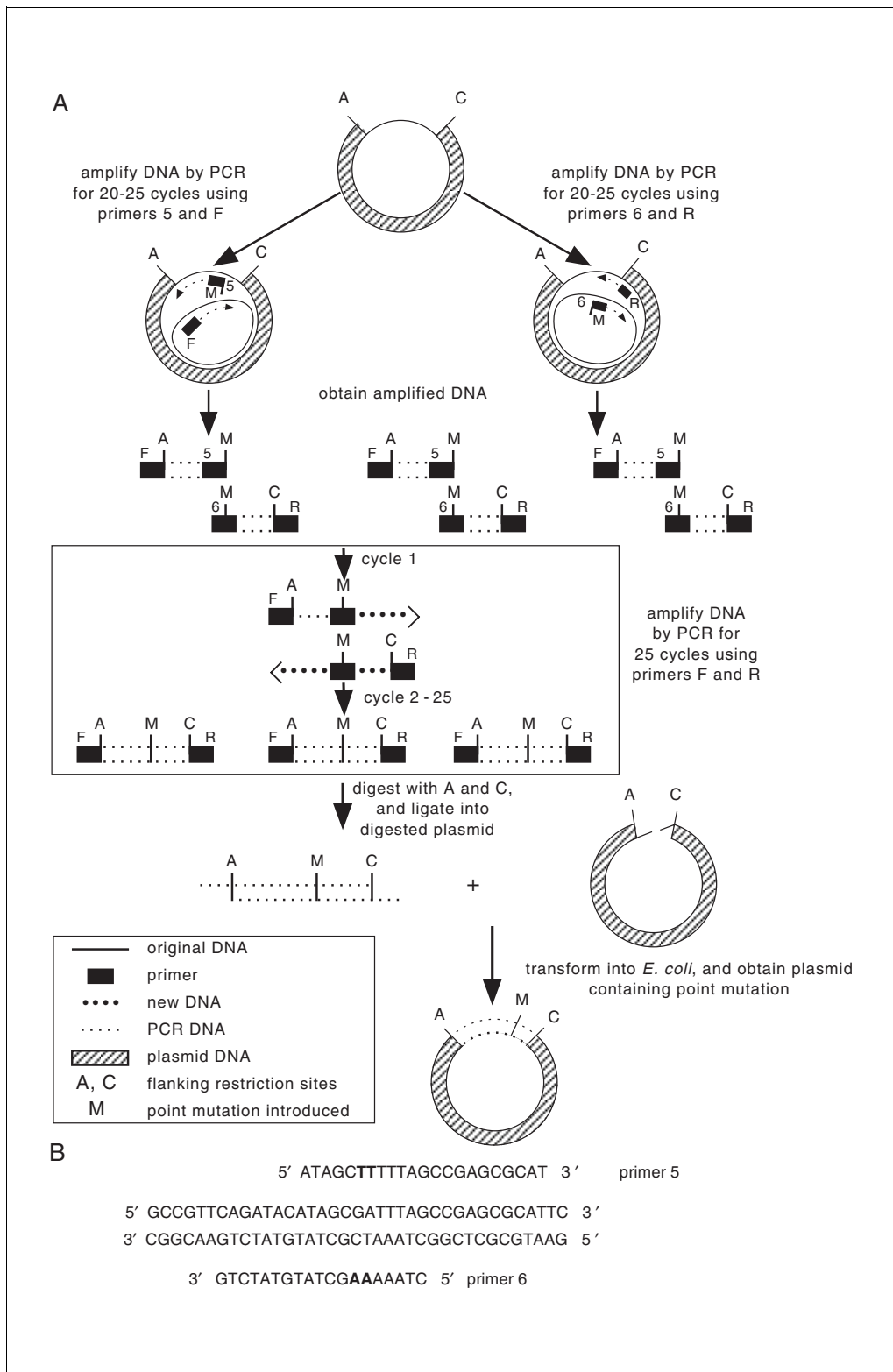


Figure 8.5.3 Introduction of a point mutation by sequential PCR steps. **(A)** The first steps of this protocol are as described in Figure 8.5.2A except primers 5 and 6 are used. The amplified fragments are then placed in the same tube and amplified in a second PCR step (using primers F and R only). This second PCR step obviates the need for the blunt-end ligation. The full-length fragment generated here is digested with the appropriate restriction endonucleases, subcloned into an appropriately cut vector, and transformed into *E. coli*. The resultant plasmid contains an inserted fragment identical to the original DNA except for the point mutation. **(B)** The oligonucleotide primers needed to generate a point mutation in this procedure are shown. Note that there are 12 bases of overlap between the two primers.

6. Carry out PCR for 20 to 25 cycles, using the conditions for introduction of restriction endonuclease sites by PCR (see Basic Protocol 1, step 6).
7. Analyze and process the reaction mix (see Basic Protocol 1, steps 7 and 8).

Subclone fragments and check the mutation

8. Digest the DNA fragment with the appropriate restriction endonuclease for the flanking sites (*UNIT 3.1*). Purify the digested fragment on a low gelling/melting agarose gel (*UNIT 2.6*). Subclone into an appropriately digested vector.
9. Transform recombinant plasmid into *E. coli* (*UNIT 1.8*). Prepare DNA by plasmid miniprep (*UNIT 1.6*).
10. Analyze the amplified fragment portion of the plasmid DNA by DNA sequencing (*UNIT 7.4*) to confirm the point mutation.

This is critical because the Taq DNA polymerase can introduce additional mutations into the fragment (see Critical Parameters).

COMMENTARY

Background Information

The PCR procedure described here will rapidly, efficiently, and reproducibly introduce any desired change into a DNA fragment. It is similar to the oligonucleotide-directed mutagenesis method described in *UNIT 8.1*, but does not require the preparation of a uracil-substituted DNA template.

The main disadvantage of PCR-generated mutagenesis is related to the fidelity of the *Taq* DNA polymerase. The mutation frequency for *Taq* DNA polymerase was initially estimated to be as high as $1/5000$ per cycle (Saiki et al., 1988). This means that the entire amplified fragment must be sequenced to be sure that there are no *Taq*-derived mutations. To reduce the amount of sequencing required, it is best to introduce the mutation by amplifying as small a fragment as possible. With rapid and reproducible methods of double-stranded DNA sequencing (*UNIT 7.4*), the entire amplified fragment can usually be sequenced from a single primer. If the fragment is somewhat longer, it is best to subclone the fragment into an M13-derived vector, so that both forward and reverse primers can be used to sequence the amplified fragment.

If there are no convenient restriction sites flanking the fragment of interest, the utility of this method is somewhat reduced. Many researchers prefer the mutagenesis procedure in *UNIT 8.1* to avoid excessive sequencing.

Critical Parameters

The most important requirement for this protocol is that the melting temperature of the oligonucleotide be sufficiently high to allow

efficient annealing and amplification. The T_m can be estimated according to the equation:

$$T_m = nGC(4^\circ\text{C}) + mAT(2^\circ\text{C})$$

where n is the number of G and C residues and m is the number of A and T residues in the oligonucleotide.

For a 15-mer of 50% GC content, the calculated T_m will be 42°C. However, after a single round of extension, there are six to ten additional bases of homology, and the T_m increases to between 60° and 72°C. Generally, primers with 15 bases of homology to the template are sufficient for amplification. It may be necessary, however (especially in the case of AT-rich sequences), to increase the length of homology or to lower the hybridization temperature accordingly. This must be determined empirically.

If a mismatch is present in the oligonucleotide, it is better that it be as close as possible to the 5' end of the oligonucleotide. There, the mismatch has relatively little effect on the T_m . By contrast, a point mutation in the middle of an oligonucleotide will have a large effect on the T_m . When designing oligonucleotides, it is desirable to have the longest uninterrupted stretch of homology possible.

An additional problem is that not all restriction endonucleases are capable of efficiently digesting sites at the end of a fragment; this problem is mitigated to some extent by the inclusion of the extra four bases as a "clamp." But even with this precaution, it will be difficult to clone with certain enzymes. Table 8.5.1 is a partial list of enzymes and their relative abilities

to cleave sites on short oligonucleotides, which should be used as a guide when choosing between potential restriction enzymes. An alternative is to blunt-end clone the amplified fragment and subsequently subclone from that cloned insert using the enzyme of interest.

A full discussion of critical parameters for PCR amplification can be found in *UNIT 15.1*.

Anticipated Results

Each of the procedures presented here has a 100% efficiency rate. All cloned, amplified fragments will contain the mutation corresponding to the synthesized oligonucleotide.

Time Considerations

Once the oligonucleotides are synthesized, the relevant fragments can be amplified and

subcloned within a single day. Thus, in 2 to 3 days it is possible to proceed from oligonucleotides to a fully sequenced fragment with the incorporated mutation.

Literature Cited

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487-491.

Contributed by Brendan Cormack
Harvard Medical School
Boston, Massachusetts

CHAPTER 9

Introduction of DNA into Mammalian Cells

INTRODUCTION

After cloning a gene, many researchers wish to analyze its characteristics by reintroducing normal and mutant variants into various cell types. There are many reasons to do this. The genetic elements responsible for regulation of expression can be determined by mutating regulatory sequences and subsequently examining activity under a variety of physiological conditions. The effects that the gene has on cellular growth can be ascertained by characterizing the phenotypes of cell lines that either overexpress the gene, or that express mutant forms of the gene. Overexpressing cell lines can also be used for purification of the product for biochemical characterization, or large-scale production of a product for use as a drug. All these goals require the ability to introduce DNA into a cell efficiently. The protocols in this chapter cover methods that are used to introduce genes into mammalian cells, as well as techniques used to investigate their regulation, to express the gene products, and to delete the endogenous form of the gene.

Many researchers must transfect DNA into a cell line that is chosen not for ease of transfection but for its biological properties. A frequent reason for introducing DNA into mammalian cells is to analyze gene expression. It is critical that an appropriate cell line be used in these studies. For example, in HeLa cells the β -globin gene is expressed at a low level that is not inducible, but it is inducible upon differentiation of MEL cells. Clearly, MEL cells should be used to investigate the mechanism of the induction. Many researchers must therefore transfect DNA into a cell line that is chosen not for ease of transfection but for its biological properties. Analysis of phenotypes of overexpression or deletion usually must also be done in appropriate cell lines.

The first section of the chapter presents four techniques for introducing DNA into mammalian cells: calcium phosphate transfection (*UNIT 9.1*), DEAE-dextran transfection (*UNIT 9.2*), electroporation (*UNIT 9.3*), and liposome-mediated transfection (*UNIT 9.4*). The first two procedures produce a chemical environment that results in DNA attaching to the cell surface; the DNA is then endocytosed by as yet uncharacterized pathways. The parameters for transfecting cells by these techniques vary for each different cell type and therefore need to be carefully optimized; these issues are discussed in the introduction to this first section.

Electroporation uses an electric field to open up pores in the cell. The DNA presumably diffuses into the cell through the pores. This technique therefore is not dependent upon special characteristics of the cell and can be used with virtually any cell type. The optimal amplitude and length of pulse will vary for each cell type, so this procedure too requires fine-tuning.

In liposome-mediated transfection, the mechanism by which liposomes containing cationic and neutral lipids mediate transfection of DNA into animal cells is not well understood. Negatively charged phosphate groups on DNA bind to the positively charged surface of the liposome, and the residual positive charge then presumably mediates binding to negatively charged sialic acid residues on the cell surfaces.

Contributed by Robert E. Kingston

Current Protocols in Molecular Biology (2003) 9.0.1-9.0.5

Copyright © 2003 by John Wiley & Sons, Inc.

**Introduction
of DNA into
Mammalian Cells**

9.0.1

Supplement 64

The simplest way to optimize parameters for transfection is to use an easily assayed reporter gene as the transfected DNA during the optimization. Reporter systems are covered in the second section of this chapter (*UNITS 9.6 & 9.7*), with descriptions of how these genes can be used to optimize transfection efficiency and to analyze promoter function. Reporter genes allow indirect measurement of promoter activity, so the researcher will need to analyze RNA levels and the structure of the RNA produced from the transfected gene. Strategies for using the techniques of RNA isolation and analysis (Chapter 4) in conjunction with transfection procedures to analyze directly the regulation of a gene are described in *UNIT 9.8*.

Reporter genes are used not only to measure the activity of a gene, but also to visualize the location of the protein that is expressed. *UNIT 9.7C* describes the use of the Green Fluorescent Protein (GFP) to directly visualize proteins in living cells. This reporter can provide a simple method for evaluating the percentage of cells that have been transfected and can serve as a marker for cells within a population that have received transfected DNA. GFP is also used to examine the dynamics of protein movement and expression within a cell over time by creating a fusion protein between GFP and a protein of research interest.

CHOICE OF TRANSFECTION METHOD

There are two types of transfections that are routinely done in mammalian systems—*transient* and *stable* or *permanent*. In a transient transfection, transcription or replication of the transfected gene can be analyzed between 1 and 4 days after introduction of the DNA, generally by harvesting the transfected cell. Alternatively, many experiments require formation of cell lines that contain gene(s) that are integrated into chromosomal DNA resulting in a stable or permanent transfection; *UNIT 9.5* describes selectable markers that are used in this type of transfection.

All four methods described above can be used for transient transfections. Three of these—calcium phosphate transfection, electroporation, and liposome-mediated transfection—can be used to efficiently produce cell lines containing stably integrated DNA. Electroporation is most easily done using suspension cultures, as cells must be transferred into cuvettes, whereas calcium phosphate- and liposome-mediated transfections are most easily done using adherent cells. Two methods for calcium phosphate-mediated transfection are presented in the first unit. The HEPES-buffered system can be used for producing either transiently or stably transformed lines. The second method, based on a BES buffer system, is now widely used for the stable transformation of most common fibroblast and epithelial cells, and for these applications is at least 10-fold more efficient than other methods. DEAE-dextran transfection does not work well when producing stable cell lines, but is more reproducible than calcium phosphate transfection when used in transient protocols. Electroporation is also very reproducible, but requires more cells than either of the chemical procedures do. Although these considerations are important in choosing a transfection protocol, the most critical parameter may be the efficiency with which DNA is introduced into the recipient cell, and this may have to be determined experimentally.

There are numerous techniques used in mammalian cell culture that are not discussed in this chapter. In particular, it is assumed that the conditions needed for optimal growth and passage of the cell lines used in a specific research project will already have been determined. In order for transfection to work efficiently, it is necessary that the recipient cells be healthy. A brief discussion at the end of this introduction describes the reagents used in mammalian cell culture.

VIRAL VECTORS

DNA can also be introduced into eukaryotic cells by using viral vectors. Viral vectors can allow introduction of DNA into virtually all cells in a given experiment. Certain viral systems will result in overexpression of proteins (e.g., baculovirus and vaccinia virus; see Chapter 16) or will allow single-copy stable integrants of expressed genes (e.g., retroviral vectors, *UNITS 9.9-9.14*). Viral systems therefore can be much more powerful than transfection systems—however, constructing a recombinant virus and packaging that virus involves a significant effort. It is therefore prudent to plan experiments unusually carefully when using viral vectors to ensure that a recombinant virus is really needed and that the optimal virus is constructed.

Retroviruses are not generally used for overproducing a protein, but are instead used to introduce a gene in a stable fashion into a cell line or into cells in an animal. These vectors can therefore be used either to express a gene product in a set of cells or to express a marker gene that will allow identification of an infected cell and its progeny. An overview of the retroviral life cycle and of retroviral vectors is provided in *UNIT 9.9*. Standard cloning procedures are used to create a retroviral vector, after which a packaging line must be made to produce infectious particles (*UNIT 9.10*). An alternative approach for the transient transfection and production of retroviral supernatants is presented in *UNIT 9.11*. These cell lines can then be used to produce large amounts of recombinant retroviruses (*UNIT 9.12*). The resultant retrovirus stocks may then be used in a variety of experiments; however, it is critical to ensure that they are free of wild-type “helper” retroviruses that may arise from the packaging cell lines. These helper viruses are replication-competent, and therefore can create serious complications in certain experiments. Assays that can be used to detect helper virus in a preparation are described in *UNIT 9.13*. Detailed protocols for introducing recombinant retroviruses into animals are beyond the scope of this chapter, although *UNIT 9.14* provides a beginning point for these *in vivo* experiments.

Until recently, the introduction of mutations into mammalian systems relied upon random mutagenesis followed by selection for a particular phenotypic change. This approach works only with genes whose loss of or change in function results in a selectable phenotype. In contrast, gene targeting in somatic cells by homologous recombination (*UNIT 9.15*) allows mutation of any desired gene, regardless of phenotype. Mutating the gene in somatic cells that have defined behavioral characteristics allows the researcher to assess the impact of the mutation in a defined and clonal cellular background. Mutant cells can be examined in culture, or can be examined after injection into animals. While this protocol can only work with cells that are diploid and efficiently transfectable, it has proven to be very useful in a variety of situations.

REAGENTS AND SOLUTIONS

Unlike *Escherichia coli*, mammalian cells are extremely sensitive to growth conditions. In addition, because their doubling time is usually 12 to 48 hr, mammalian cultures are susceptible to contamination by bacteria and fungi, and the necessity to maintain sterile conditions during experimental procedures involving cell growth cannot be overemphasized. The four transfection protocols described in this chapter are stressful to the cell and thus can cause extensive cell death if the transfected cells are not healthy prior to transfection. This means that cell culture conditions that are marginal for normal growth and passage of cells will probably not suffice for transfection purposes. The following precautions and ingredients are crucial to success in culturing mammalian cells:

1. **Glassware.** Numerous solutions commonly used in the laboratory are cytotoxic when present even in trace amounts. Many laboratories therefore keep separate glassware for tissue culture purposes. This glassware is never mixed with glassware used in the laboratory. Tissue culture glassware should be left soaking in water after use in order to prevent drying of residual chemicals onto the glass. For these reasons, plasticware is frequently used for pipets and for storage of reagents used in tissue culture.
2. **Water.** Only distilled and deionized water should be used for preparation of solutions used in tissue culture. Some laboratories use a glass still to produce this water, whereas in some newer facilities, house purification systems are sufficient.
3. **Media.** It is possible to purchase sterile media in solution. Alternatively, packaged, premixed powders can be purchased. These can be made in 10- to 20-liter batches in the laboratory. The resulting solutions must then be filter sterilized. This is most conveniently done using a high-capacity pump (500 ml/min) and filter apparatus.

Media are supplemented shortly before use with serum. Frequently, fetal bovine serum (FBS; also known as fetal calf serum) is used, but for some uses, less expensive sera may be employed (see below). Other supplements can include glutamine, nonessential amino acids, 2-mercaptoethanol (2-ME), penicillin, and streptomycin sulfate to formulate “complete” media; however, the exact formulation will vary with different cell types. Nomenclature of media employed in this manual indicates the base medium as well as the percentage of serum. For example, “complete DMEM-10” means that DMEM is supplemented with 10% serum plus the other ingredients noted above (see recipe, *UNIT 9.4*).

4. **Serum.** Horse, calf, and fetal bovine serum can be purchased. In using serum, it is accepted that a partially undefined material is being used, which may show considerable variation in the ability to support growth of particular cells. Thus, for every culture application, serum should be “screened” to determine if, in fact, it supports the growth of the cells of interest. Such screening can also be done to determine if the serum is unduly stimulatory to the cells of interest, thereby leading to high background responses. There is no easy way to screen lots of serum because the suitability of a given lot can only be determined in the application at hand. It is recommended that for extensive studies of a given cell population, small amounts of different lots of serum be obtained from one or more suppliers with the proviso that the supplier put aside a much larger amount of the lot(s) for possible delivery. The various lots are then tested in an assay of interest (a proliferation assay, a transient transfection assay, etc.) to determine which lot provides the most vigorous response and the least background response (response in the absence of specific stimulus). The lot with this characteristic is then bought in large amounts and the other lots are “released” (allowed to be sold to other purchasers). Reliable suppliers of serum are Armour and Hyclone.

Generally, heat-inactivated (1 hr at 56°C) serum is required to support long-term cultures. The heat treatment is thought by some labs to reduce the number of viral and other adventitious contaminants and by others to be required to inactivate complement. It therefore has developed into a standard procedure before the use of a particular batch of serum, despite alleged disadvantages (i.e., inactivation of certain growth factors and production of heat-labile serum components that are toxic to cells). If desired, test whether the growth-supporting function of a certain lot of serum can be improved by not heating it, particularly when culturing fastidious cells.

Given the high cost of serum, a lower percentage (5%) can be used for short-term procedures such as washing cells. In addition, bovine serum, which is about ten times

less expensive than fetal bovine serum, can sometimes be substituted in short-term procedures. Store all serum at -20°C .

5. **Supplements.** Many researchers use antibiotics in mammalian culture media to prevent contamination. The most effective addition is gentamicin sulfate, which is purchased as a 50 mg/ml solution and used at 20 to 50 $\mu\text{g/ml}$. A somewhat less expensive (and less stable) antibiotic combination that is also somewhat less effective is a cocktail of penicillin (50 to 100 U/ml) and streptomycin sulfate (100 $\mu\text{g/ml}$). Note that use of antibiotics in the media can mask sloppy tissue culture technique. Sloppy technique can result in contamination of a culture by agents that are not susceptible to the antibiotics, such as mycoplasma. Some laboratories therefore do not use antibiotics in order to ensure that strict sterility is being maintained. Some experiments will require other additions to the media, such as nucleosides or drugs. These solutions should be made up using tissue culture glassware or plasticware, and care should be taken to avoid the use of stir bars or other items that might have trace amounts of laboratory chemicals on them. The solutions should be filter sterilized through Nalgene filters.

2-mercaptoethanol is often included in media but the precise mode(s) of action of 2-ME has not been defined. The presence of 2-ME during growth of established cell lines appears not to be required, but it has proven to be a critical ingredient in primary cultures. 2-mercaptoethanol must be added to media immediately before use because its activity rapidly declines in diluted form. Thus, concentrated 2-ME (14.3 M) is diluted to 50 mM in HBSS or PBS and this solution is added to media at a final concentration of 50 μM . The 50 mM stock solution is stored at 4°C and should be replaced after ~ 4 months.

Other reagents are added to media immediately before use; these are typically made in large batches at 100 \times or 200 \times concentrations and stored in aliquots of convenient size until use. Stock solutions containing L-glutamine, nonessential amino acids, and antibiotics are stored at -20°C . Stock solutions containing HEPES are stored at room temperature.

6. **Trypsin/EDTA.** A buffered salt solution containing 0.5% (w/v) trypsin and 0.2% (w/v) EDTA is used to dissociate adherent cells from tissue culture dishes. This solution can be prepared in the laboratory or purchased.

A more complete description of the preparation of culture media and techniques for mammalian cell tissue culture can be found in *APPENDIX 3F*.

Robert E. Kingston

TRANSFECTION OF DNA INTO EUKARYOTIC CELLS

OPTIMIZATION OF TRANSFECTION

When embarking upon any transfection procedure, a critical first step is to optimize conditions. Every mammalian cell type has a characteristic set of requirements for optimal introduction of foreign DNA; there is a tremendous degree of variability in the transfection conditions that work, even among cell types that are very similar to one another. Often, an experimenter must screen a wide variety of cell types for a desired regulatory trait, such as an appropriate response to a particular effector molecule. It is thus helpful to have a straightforward, systematic approach to optimizing transfection efficiency. Transient assay systems are particularly useful for this purpose. A fusion gene that is known to function in mammalian cells can be transfected into cells under a variety of conditions, and transfection efficiency can be easily monitored by assaying for the fusion gene product. The human growth hormone (hGH; see *UNIT 9.7A*) assay system is particularly useful for this purpose because both harvest and assay take very little time. However, any reporter system can be used to optimize transfection efficiency.

The single most important factor in optimizing transfection efficiency is selecting the proper transfection protocol. This usually comes down to a choice among calcium phosphate-mediated gene transfer (*UNIT 9.1*), DEAE-dextran-mediated gene transfer (*UNITS 9.2 & 16.12*), electroporation (*UNIT 9.3*), and liposome-mediated transfection (*UNIT 9.4*). Fusion techniques such as protoplast fusion and microinjection may also be considered. Cells are variable with respect to which transfection protocol is most efficient. It is recommended that any adherent cell line under investigation be tested for transfection ability with DEAE-dextran, calcium phosphate, and liposome-mediated transfection. Nonadherent cell lines can be transfected by electroporation and liposome-mediated transfection. Generally, if a cell can be grown in culture, it can be transfected.

Calcium Phosphate Transfection

The primary factors that influence efficiency of calcium phosphate transfection (*UNIT 9.1*) are the amount of DNA in the precipitate, the length of time the precipitate is left on the cell, and the use and duration of glycerol or DMSO shock. A calcium phosphate optimization is shown in Table 9.1.1. Generally, higher concentrations of DNA (10 to 50 μg) are used in calcium phosphate transfection. Total DNA concentration in the

Table 9.1.1 Optimization of Calcium Phosphate Transfection

Dish (10-cm)	pXGH5 (μg)	pUC13 (μg)	Exposure to precipitate (hr)	Glycerol shock (min)
1	5	5	6	—
2	5	15	6	—
3	5	35	6	—
4	5	5	16	—
5	5	15	16	—
6	5	35	16	—
7	5	5	6	3
8	5	15	6	3
9	5	35	6	3
10	5	5	16	3
11	5	15	16	3
12	5	35	16	3

Contributed by Robert E. Kingston, Claudia A. Chen, Hiroto Okayama, and John K. Rose
Current Protocols in Molecular Biology (1996) 9.1.1-9.1.11
Copyright © 2003 by John Wiley & Sons, Inc.

precipitate can have a dramatic effect on efficiency of uptake of DNA with calcium phosphate-mediated transfection. With some cell lines, more than 10 to 15 μg of DNA added to a 10-cm dish results in excessive cell death and very little uptake of DNA. With other cell types, such as primary cells, a high concentration of DNA in the precipitate is necessary to get any DNA at all into the cell on a routine basis. For example, with human foreskin fibroblasts, transfection of 5 μg of a reporter plasmid with 5 μg of carrier DNA (e.g., pUC13) gives significantly less expression than does transfection of 5 μg of reporter plasmid with 35 μg of carrier DNA. Presumably, this is because the amount of DNA affects the nature of the precipitate and thus alters the fraction of the applied DNA that is taken up into cells.

The optimal length of time that the precipitate is left on cells varies with cell type. Some cell types, such as HeLa or BALB/c 3T3, are efficiently transfected by leaving the precipitate on for 16 hr. Other cell types cannot survive this length of exposure to the precipitate. Transfection efficiency of some cell types, such as CHO DUKX BII, is dramatically increased by glycerol or DMSO shock (*UNIT 9.1*). The pilot experiment listed will indicate whether the cell type is tolerant to long exposure to a calcium phosphate precipitate and whether glycerol shock should be used. Once the results of this experiment are in hand, finer experiments can be done to further optimize conditions. For example, if shocking with 10% glycerol for 3 min enhances transfection efficiency, an experiment varying the time of glycerol shock or also trying 10% and 20% DMSO shock might be done.

Once optimal conditions for transfection are found, extensive DNA curves varying the amount of reporter plasmid should be prepared. The total amount of DNA should be kept constant at the optimal level determined in the first experiment. The amount of reporter plasmid DNA (e.g., pXGH5) should be varied, and carrier DNA (e.g., pUC13) should be used to make up the difference. This is to ensure that transfections are performed under conditions where the amount of reporter plasmid in the cell is not saturating the cellular transcription and translation machinery.

DEAE-Dextran Transfection

There are several factors that can be varied in DEAE-dextran transfection (*UNIT 9.2*). The number of cells, concentration of DNA, and concentration of DEAE-dextran added to the dish are the most important to optimize. To a first approximation, most cell types that can be transfected using DEAE-dextran will have a preference for 1 to 10 μg DNA/10-cm dish and for 100 to 400 μg DEAE-dextran/ml of medium. Table 9.1.2 shows how the

Table 9.1.2 Optimization of DEAE-Dextran Transfection

5×10^5 cells/10-cm dish:			2×10^6 cells/10-cm dish:		
Dish	pXGH5 (μg)	DEAE-dextran ($\mu\text{g}/\text{ml}$)	Dish	pXGH5 (μg)	DEAE-dextran ($\mu\text{g}/\text{ml}$)
1	1	400	11	1	400
2	1	200	12	1	200
3	1	100	13	1	100
4	4	400	14	4	400
5	4	200	15	4	200
6	4	100	16	4	100
7	10	400	17	10	400
8	10	200	18	10	200
9	10	100	19	10	100
10	0	200	20	0	200

dishes in an optimization might be chosen. The 20-dish experiment consists of two sets of 10 dishes; one set is plated at 5×10^5 cells/dish, the other is plated at 2×10^6 cells/dish. Each set contains dishes that will be transfected with 1 to 10 μg of a reporter plasmid and 100 to 400 $\mu\text{g}/\text{ml}$ DEAE-dextran. If an hGH expression vector such as pXGH5 is used, a time course of expression under each condition can be determined by removing 100- μl aliquots of the medium 2, 4, and 7 days posttransfection (with a medium change after the day 4 aliquot is removed).

With the results of this pilot experiment in hand, a second experiment using a narrower range of DEAE-dextran concentrations and a wider range of DNA doses should be undertaken. For example, if the cells appear to express more hGH at 100 $\mu\text{g}/\text{ml}$ DEAE-dextran than at higher concentrations in the pilot experiment, the second experiment should cover from 25 to 150 $\mu\text{g}/\text{ml}$ DEAE-dextran. Because DEAE-dextran is toxic to some cells, a brief exposure to small concentrations may be optimal. The wide range of added DNA in this experiment is crucial in two respects. First, it is valuable to know the smallest amount of the transfected reporter gene that can give a readily detectable signal. Second, the linearity of the dose of DNA with the amount of reporter gene expression generally decays for large amounts of input DNA. When excessive (i.e., nonlinear) amounts of DNA are used in transfection experiments, it is possible that the effects observed are dose-response effects rather than the phenomenon intended for study. This serious and common problem can be eliminated by doing a careful DNA dose-response curve as above.

Electroporation

Perhaps because it is not a chemically based protocol, electroporation (*UNIT 9.3*) tends to be less affected by DNA concentration than either DEAE-dextran- or calcium phosphate-mediated gene transfer. Generally, DNA amounts in the range of 10 to 40 $\mu\text{g}/10^7$ cells work well, and there is a good linear correlation between the amount of DNA present and the amount taken up. The parameter that can be varied to optimize electroporation is the amplitude and length of the electric pulse, the latter being determined by the capacitance of the power source. The extent to which this can be varied is determined by the electronics of the power supply used to supply the pulse. The objective is to find a pulse that kills between 20% and 60% of the cells. This generally is in the range of 1.5 kV at 25 μF . If excessive cell death occurs, the length of the pulse can be lowered by lowering the capacitance. Settings between 3 and 25 μF can be tried.

Liposome-Mediated Transfection

Three primary parameters—the concentrations of lipid and DNA and incubation time of the liposome-DNA complex—affect the success of DNA transfection by cationic liposomes (*UNIT 9.4*). These should be systematically examined to obtain optimal transfection frequencies.

Concentration of lipid. In general, increasing the concentrations of lipid improves transfection of four cell lines examined (CV-1 and COS-7 with Lipofectin, and HeLa and BHK-21 with TransfectACE; see *UNIT 9.4*). However, at high levels ($>100 \mu\text{g}$), the lipid can be toxic. For each particular liposome mixture tested, it is important to vary the amount as indicated in Table 9.1.3.

Concentration of DNA. In many of the cell types tested, relatively small amounts of DNA are effectively taken up and expressed. In fact, higher levels of DNA can be inhibitory in some cell types with certain liposome preparations. In the optimization protocol outlined in Table 9.1.3, the standard reporter vector pSV2CAT is used; however, any plasmid DNA whose expression can be easily monitored would be suitable.

Table 9.1.3 Optimization of Liposome-Mediated Transfection

Dish (35-mm)	pSV2CAT (μg)	Liposomes (μl)	Dish (35-mm)	pSV2CAT (μg)	Liposomes (μl)
1	0.1	1	11	5	5
2	0.1	2	12	5	10
3	0.1	4	13	5	15
4	0.1	8	14	5	20
5	0.1	12	15	5	30
6	0.5	1	16	10	5
7	0.5	2	17	10	10
8	0.5	4	18	10	15
9	0.5	8	19	10	20
10	0.5	12	20	10	30

Time of incubation. When the optimal amounts of lipid and DNA have been established, it is desirable to determine the length of time required for exposure of the liposome-DNA complex to the cells. In general, transfection efficiency increases with time of exposure to the liposome-DNA complex, although after 8 hr, toxic conditions can develop. HeLa or BHK-21 cells typically require ~3 hr incubation with the liposome-DNA complex for optimal transfection, while CV-1 and COS-7 cells require 5 hr of exposure.

UNIT 9.1

Calcium Phosphate Transfection

This unit presents two methods of calcium phosphate-based eukaryotic cell transfection that can be used for both transient and stable (UNIT 9.5) transfections. In these protocols, plasmid DNA is introduced to monolayer cell cultures via a precipitate that adheres to the cell surface. A HEPES-buffered solution is used to form a calcium phosphate precipitate that is directly layered onto the cells (see Basic Protocol). For some cells, shocking the cells with glycerol or DMSO (see Support Protocol) improves transfection efficiency. In the alternate high-efficiency method, a BES-buffered system is used that allows the precipitate to form gradually in the medium and then drop onto the cells (see Alternate Protocol). The alternate method is particularly efficient for stable transformation of cells with circular plasmid DNA. For transformation with linear plasmid or genomic DNA, or for transient expression, however, the Alternate Protocol is comparable to the Basic Protocol. Both methods of transfection require very high-quality plasmid DNA. Factors that can be optimized for calcium phosphate transfections are presented in the introduction to Section I, and protein expression strategies are discussed in Chapter 16. Additional details of mammalian cell culture are given in APPENDIX 3F.

BASIC PROTOCOL

TRANSFECTION USING CALCIUM PHOSPHATE-DNA PRECIPITATE FORMED IN HEPES

A precipitate containing calcium phosphate and DNA is formed by slowly mixing a HEPES-buffered saline solution with a solution containing calcium chloride and DNA. This precipitate adheres to the surface of cells and should be visible in the phase contrast microscope the day after transfection. Depending on the cell type, up to 10% of the cells on a dish will take up the DNA precipitate through an as yet undetermined mechanism. Glycerol or dimethyl sulfoxide shock increases the amount of DNA absorbed in some cell types (see Support Protocol).

Calcium Phosphate Transfection

9.1.4

Materials

- Exponentially growing eukaryotic cells (e.g., HeLa, BALB/c 3T3, NIH 3T3, CHO, or rat embryo fibroblasts)
- Complete medium (depending on cell line used)
- CsCl-purified plasmid DNA (10 to 50 μg per transfection)
- 2.5 M CaCl_2 (see recipe)
- 2 \times HEPES-buffered saline (HeBS; see recipe)
- PBS (APPENDIX 2)
- 10-cm tissue culture dishes
- 15-ml conical tube
- Additional reagents and equipment for ethanol precipitation (UNIT 2.1A) and mammalian cell tissue culture (APPENDIX 3F)

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

NOTE: All culture incubations are performed in a humidified 37°C, 5% CO_2 incubator unless otherwise specified.

1. Split exponentially growing eukaryotic cells into 10-cm tissue culture dishes the day before transfection. Feed cells with 9.0 ml complete medium 2 to 4 hr prior to precipitation.

When transfecting adherent cells that double every 18 to 24 hr, a 1:15 split from a confluent dish generally works well. On the day of the transfection, it is important that cells are thoroughly separated on the dish, as the ability to take up DNA is related to the surface area of the cell exposed to the medium. Cells should be split in a manner that accomplishes this.

The desired density of cells on dishes to be transfected will vary with cell type and the reason for doing the transfection. The optimal density is that which produces a near confluent dish when the cells are harvested or split into selective medium.

2. Ethanol precipitate the DNA to be transfected and air dry the pellet by inverting the microcentrifuge tube on a fresh Kimwipe inside a tissue culture hood. Resuspend the pellet in 450 μl sterile water and add 50 μl of 2.5 M CaCl_2 .

The amount of DNA that is optimal for transfection varies from 10 to 50 μg per 10-cm plate, depending on the cell line to be transfected.

DNA to be transfected should be purified twice by CsCl gradient centrifugation (UNIT 1.7). DNA can also be prepared using column procedures (UNIT 2.1B). Some column procedures produce DNA that does not transfect well, so column-purified DNA should be tested and compared to CsCl-purified DNA for transfection efficiency. Supercoiled DNA works well in transfections. Impurities in the DNA preparation can be deleterious to transfection efficiency. A description of how to optimize the amount of DNA to transfect and other parameters of calcium phosphate-mediated transfection is provided in the discussion of Optimization of Transfection.

Ethanol precipitation sterilizes the DNA to be transfected. For transfections that will be harvested within 3 to 4 days (transient analysis), this is not necessary. For transient experiments, many researchers make a 450- μl aqueous solution containing the DNA directly, without ethanol precipitation. If this is done, care should be taken to keep the amount of Tris in the solution to a minimum, as Tris may alter the pH of the precipitate and therefore reduce transfection efficiency.

3. Place 500 μl of 2 \times HeBS in a sterile 15-ml conical tube. Use a mechanical pipettor attached to a plugged 1- or 2-ml pipet to bubble the 2 \times HeBS and add the DNA/ CaCl_2

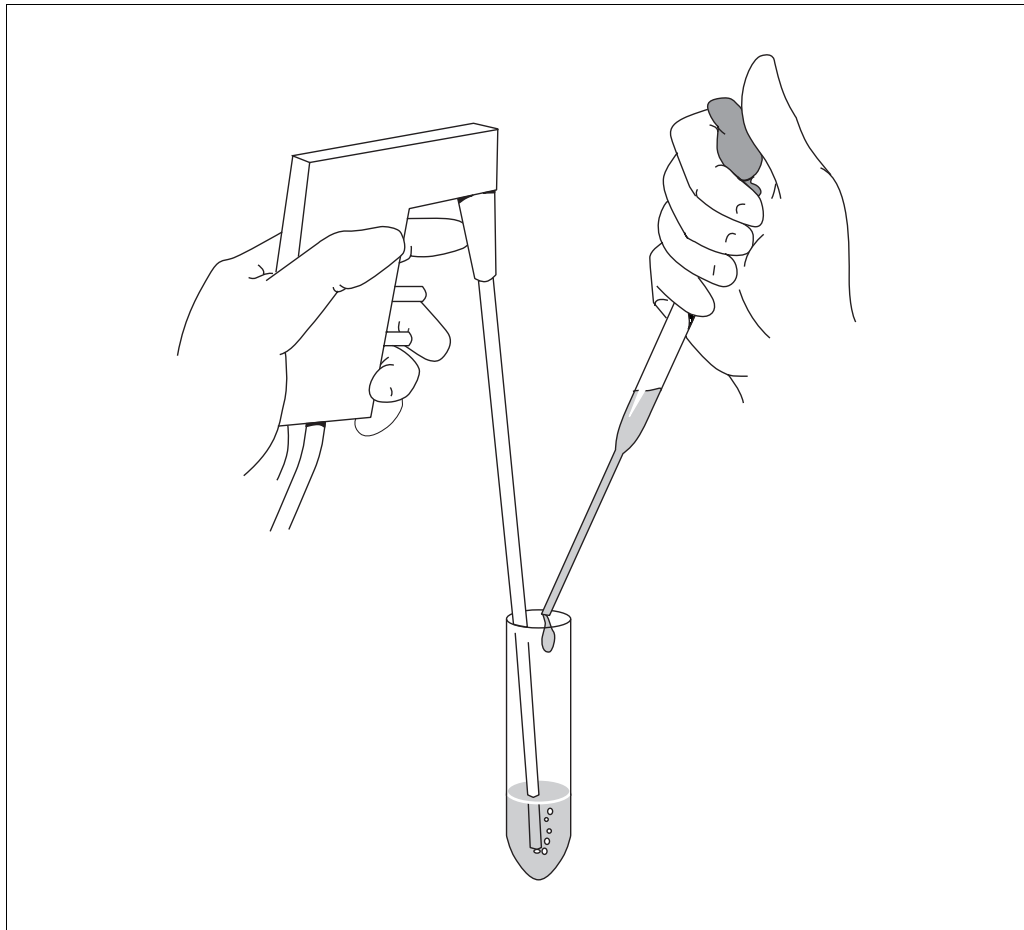


Figure 9.1.1 Formation of calcium phosphate precipitate.

solution dropwise with a Pasteur pipet (see Fig. 9.1.1). Immediately vortex the solution for 5 sec.

If no mechanical pipettor is available, the solution can be bubbled by blowing through rubber tubing that is attached to a pipet via a filter. The filter is necessary to maintain sterility. This does not give as reproducible results as the mechanical pipettor.

4. Allow precipitate to sit 20 min at room temperature.
5. Use a Pasteur pipet to distribute the precipitate evenly over a 10-cm plate of cells and gently agitate to mix precipitate and medium.
6. Incubate the cells 4 to 16 hr under standard growth conditions. Remove the medium. Wash cells twice with 5 ml PBS and feed cells with 10 ml complete medium.

The amount of time that the precipitate should be left on the cells will vary with cell type. For hardy cells such as HeLa, NIH 3T3, and BALB/c 3T3, the precipitate can be left on for 16 hr. Other cell types will not survive this amount of exposure to the precipitate. See discussion of Optimization of Transfection for optimization of this step as well as for a discussion of how to determine whether glycerol shock is useful.

7. For transient analysis, harvest the cells at the desired time point (UNITS 9.6-9.8 & 14.6). For stable transformation, allow the cells to double twice before plating in selective medium (UNIT 9.5).

GLYCEROL/DMSO SHOCK OF MAMMALIAN CELLS

The Basic Protocol works well for cell lines such as HeLa, BALB/c 3T3, NIH 3T3, and rat embryo fibroblasts. Transfection efficiency in some cell lines, such as CHO DUKX, is dramatically increased by “shocking” the cells with either glycerol or DMSO. Precipitates are left on the cell for only 4 to 6 hr, and the cells are shocked immediately after removal of the precipitate.

Additional Materials (also see *Basic Protocol*)

10% (v/v) glycerol solution or DMSO in complete medium, sterile
PBS (*APPENDIX 2*), sterile

Replace step 6 of the Basic Protocol with the following:

- 6a. Incubate the cells 4 to 6 hr and remove the medium. Add 2.0 ml of a sterile 10% glycerol solution. Let the cells sit 3 min at room temperature.

Alternatively, 10% or 20% DMSO can be used. DMSO tends to be somewhat less harmful to the cells, but also may not work as well.

- 6b. Add 5 ml of PBS to the glycerol solution on the cells, agitate to mix, and remove the solution. Wash twice with 5 ml of PBS. Feed the cells with complete medium.

It is important to dilute the glycerol solution on the cells with PBS before removing the glycerol solution so the cells do not stay in glycerol too long. Excessive exposure to glycerol will kill cells.

HIGH-EFFICIENCY TRANSFECTION USING CALCIUM PHOSPHATE–DNA PRECIPITATE FORMED IN BES

A solution of calcium chloride, plasmid DNA, and *N,N*-bis(2-hydroxyethyl)-2-aminoethanesulfonic acid (BES) buffer, pH 6.95, is added to a plate of cells containing culture medium. The plates are incubated overnight while a calcium phosphate–DNA complex forms gradually in the medium under an atmosphere of 3% CO₂. With this method, 10% to 50% of the cells on a plate stably integrate and express the transfected DNA. Transient expression under these conditions is comparable to that obtained with the Basic Protocol. Glycerol or DMSO shock does not increase the number of cells transformed.

Materials

Exponentially growing mammalian cells (see *Critical Parameters*)
Complete medium: Dulbecco modified Eagle medium containing 10% (v/v) fetal bovine serum (FBS)
CsCl-purified plasmid DNA
TE buffer, pH 7.4 (*APPENDIX 2*)
2.5 M CaCl₂ (see recipe)
2× BES-buffered solution (BBS; see recipe)
PBS (*APPENDIX 2*)
Selection medium (*UNIT 9.5*; optional)
10-cm tissue culture dishes
35°C, 3% CO₂ humidified incubator
35° to 37°C, 5% CO₂ humidified incubator
Fyrite gas analyzer (optional; Fisher Scientific or Curtin Matheson)

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

1. Seed exponentially growing mammalian cells at 5×10^5 cells/10-cm tissue culture dish in 10 ml complete medium the day prior to transfection.

There should be $<10^6$ cells/dish just prior to infection. Enough surface area should remain on the plate for at least two more doublings.

2. Dilute CsCl-purified plasmid DNA with TE buffer to $1 \mu\text{g}/\mu\text{l}$. Store the DNA solution at 4°C .

Purity of the plasmid DNA is critical.

The optimum amount of plasmid to use can be determined by transfecting three dishes of cells with 10, 20, and $30 \mu\text{g}$ of plasmid DNA and incubating overnight. The dishes should then be examined with a microscope at $100\times$. A coarse, clumpy precipitate will form at DNA concentrations that are too low, a fine (almost invisible) precipitate will form at concentrations that are higher than optimal, and an even, granular precipitate will form with optimal DNA concentrations.

3. Prepare 0.25 M CaCl_2 from 2.5 M stock. Mix 20 to $30 \mu\text{g}$ plasmid DNA with $500 \mu\text{l}$ of 0.25 M CaCl_2 . Add $500 \mu\text{l}$ of $2\times$ BBS, mix well, and incubate 10 to 20 min at room temperature.

4. Add the calcium phosphate–DNA solution dropwise onto the medium-containing plate while swirling the plate. Incubate 15 to 24 hr in a 35°C , 3% CO_2 incubator.

Level of carbon dioxide is critical. Use a Fyrite gas analyzer to measure percent CO_2 prior to incubation.

5. Wash the cells twice with 5 ml PBS, and add 10 ml complete medium. For stable transformation, incubate overnight in a 35° to 37°C , 5% CO_2 incubator. For studies involving transient expression (UNITS 9.6-9.8 & 14.6), incubate the cells for 48 to 72 hr after adding the DNA.

6. Split the cells 1:10 to 1:30, depending on the growth rate of the host cell, before beginning to select stable transformants. Incubate overnight in a 35° to 37°C , 5% CO_2 incubator.

7. Start selection by changing the medium to selection medium or by incubating cells under appropriate selection conditions (UNIT 9.5).

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

BES-buffered solution (BBS), $2\times$

50 mM *N,N*-bis(2-hydroxyethyl)-2-aminoethanesulfonic acid (BES;

Calbiochem-Novabiochem)

280 mM NaCl

1.5 mM Na_2HPO_4 , pH 6.95

800 ml H_2O

Adjust to pH 6.95 with 1 N NaOH, room temperature

Add H_2O to 1 liter

Filter sterilize through a $0.45\text{-}\mu\text{m}$ nitrocellulose filter (Nalgene)

Store in aliquots at -20°C

As discussed in the Commentary, the pH of this solution is critical (pH 6.95 to 6.98). When a new batch of $2\times$ BES buffer is prepared, its pH should be checked against a reference stock prepared (and tested) earlier.

This solution can be frozen and thawed repeatedly.

CaCl₂, 2.5 M

183.7 g CaCl₂·2H₂O (Sigma; tissue culture grade)

H₂O to 500 ml

Filter sterilize through a 0.45-μm nitrocellulose filter (Nalgene)

Store at -20°C in 10-ml aliquots

This solution can be frozen and thawed repeatedly.

HEPES-buffered saline (HeBS) solution, 2×

16.4 g NaCl (0.28 M final)

11.9 g HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid; 0.05 M final)

0.21 g Na₂HPO₄ (1.5 mM final)

800 ml H₂O

Titrate to pH 7.05 with 5 N NaOH

Add H₂O to 1 liter

Filter sterilize through a 0.45-μm nitrocellulose filter

Test for transfection efficiency

Store at -20°C in 50-ml aliquots

An exact pH is extremely important for efficient transfection. The optimal pH range is 7.05 to 7.12.

There can be wide variability in the efficiency of transfection obtained between batches of 2× HeBS. Efficiency should be checked with each new batch. The 2× HeBS solution can be rapidly tested by mixing 0.5 ml of 2× HeBS with 0.5 ml of 250 mM CaCl₂ and vortexing. A fine precipitate should develop that is readily visible in the microscope. Transfection efficiency must still be confirmed, but if the solution does not form a precipitate in this test, there is something wrong.

COMMENTARY

Background Information

Calcium phosphate transfection was first used to introduce adenovirus DNA into mammalian cells by Graham and van der Eb (1973). It was later found to be possible to integrate exogenous DNA into mammalian chromosomes using this technique (Wigler et al., 1978). The protocol described here has evolved from these methods; in particular, the order of addition of the HEPES-buffered saline solutions (HeBS) to calcium chloride is different. In addition, many investigators have found that it is not necessary to add chromosomal DNA as carrier when doing a transfection. Transfections work well if only plasmid DNA, such as pUC vectors, is used.

This HEPES-based approach has been used to analyze replication and promoter function using “transient” protocols, in which cells are harvested 48 to 60 hr after the transfection is started. It also is presently the most widely used technique for producing cell lines in which transfected DNA is stably integrated into the chromosome, although the BES-based technique (see Alternate Protocol) is gaining widespread use for production of stable transfor-

ants (UNIT 9.5). Calcium phosphate-mediated transfection tends to work much better than DEAE-dextran-mediated transfection (UNIT 9.2) in formation of stable cell lines. It is believed that this is because cells in a calcium phosphate transfection pick up more DNA than DEAE-dextran-transfected cells. Electroporation (UNIT 9.3) and liposome-mediated transfection (UNIT 9.4) can also be successfully used to produce stable cell lines.

Another major strength of calcium phosphate transfection is that transfected cells generally contain a representative sampling of the various plasmids in the precipitate. Hence, one can prepare a 10:1 ratio of two plasmids and expect that the plasmids will be present in that ratio in the transfected cells.

The BES-based high-efficiency calcium phosphate transfection (see Alternate Protocol) is a modification of the standard calcium phosphate method that employs a buffer system originally developed for phage particle-mediated gene transfer (Ishiura et al., 1982). With this buffer system the calcium phosphate-DNA precipitate forms gradually in the culture medium, dropping gently onto the cells over a 15-

to 24-hr period. This method can stably transform most common mammalian fibroblast and epithelial cell lines 10- to 100-fold more efficiently than other methods. For transient expression, it is no better than the standard method.

Critical Parameters and Troubleshooting

Calcium phosphate transfections are finicky—they are not difficult to do, they just do not always work—even in the hands of people who routinely do them. In the Basic Protocol, the most common reason for failure is a 2× HeBS solution that is no longer at the appropriate pH. The optimum pH range for transfection is extremely narrow (between 7.05 and 7.12; Graham and van der Eb, 1973). The pH of the solution can change during storage and an old 2× HeBS solution may not work well. Some investigators also have noticed that the 2.5 M CaCl₂ solution can go bad over time. Both solutions should be made fresh if transfections have stopped working well.

A second problem is that the pH of the medium can turn acidic while the transfection is in progress. This results in an extremely heavy precipitate (making the medium resemble orange juice) and generally results in cell death. Care should be taken to maintain a pH of 7.2 to 7.4 and CO₂ concentrations in the incubator as listed in the protocols. Incubator and medium conditions that are fine for routinely growing cells may not suffice for calcium phosphate transfection.

Many factors can influence the efficiency of HEPES-buffered calcium phosphate-mediated transfection. A description of experiments that can be done to optimize transfection efficiency using both this procedure and the BES procedure can be found in the discussion of Optimization of Transfection.

Several parameters are crucial to achieve high efficiency with the Alternate Protocol: pH of the 2× BES-buffered saline (BBS), percentage of carbon dioxide in the incubator during formation of precipitate, and form and amount of DNA used.

A pH curve of the 2× BBS buffer should be made because minor variations in pH can have substantial effects on transfection efficiency. Perform pilot experiments with buffers of varying pH. The optimal pH is within a very narrow range (6.95 to 6.98). Once the optimal buffer is found, use it as a reference to prepare buffer stocks. If no precipitate forms, the concentration of calcium chloride or 2× BES solution

(BBS) may be wrong. Be sure to mix reagents thoroughly before adding them to the DNA. Crystal formation upon the addition of calcium chloride indicates incorrect calcium chloride concentration, and the transfection must be repeated.

The first overnight incubation should be at 3% CO₂, but a variation of 0.5% may be acceptable. After overnight incubation at 3% CO₂ the culture medium should be alkaline (pH 7.6). Measure the CO₂ levels of the incubator before using. A Fyrite device is recommended for this.

Only plasmid DNA gives high efficiencies of gene transfer with the Alternate Protocol. Efficiency is also dependent on the purity and concentration of the DNA. Toxicity that often occurs with common fibroblasts and epithelial cell lines is usually caused by impure DNA, not calcium phosphate. DNA prepared by cesium chloride gradient centrifugation is rarely toxic to cells. The optimal DNA concentration varies among plasmid preparations as well as with different cells and media. Each new plasmid preparation and each new cell line being transfected should be tested for optimum DNA concentration.

If it is suspected that a particular plasmid preparation is toxic, use a control plasmid—one known not to be toxic to these cells—to test for toxicity. If the plasmid DNA is toxic, prepare new DNA.

Cotransfection efficiency is 10- to 20-fold better with the BES method (see Alternate Protocol) than with the Basic Protocol, although efficiencies vary with the plasmid and marker gene used. A 1:10 ratio of selectable marker to nonselected gene is recommended, but the efficiency of transfection will depend on optimum DNA dose. Glycerol shock or DMSO treatment will not increase the number of cells transfected with this method.

The BES protocol (see Alternate Protocol) has been optimized for use with cells that grow in Dulbecco modified Eagle medium containing 10% (v/v) fetal bovine serum (FBS). RPMI and minimal essential medium (MEM α) have also been demonstrated to give good results under the stated conditions. The FBS must be tested before use by examining the growth, plating efficiency, and transformation efficiency of at least two cell lines. (Serum is tested without heat inactivation.) Because FBS is very costly, 10% (v/v) newborn calf serum and 5% (v/v) FBS can be added to the medium (a total of 75 ml serum/500 ml medium) instead of 10% FBS. This gives equivalent efficien-

cies of stable transformation. The lots of newborn calf serum also must be tested for growth and plating efficiencies, as these have been found to vary. If horse or other serum or medium is required, optimum conditions may need to be rechecked.

The Alternate Protocol works well for most established cell lines that grow as monolayers, including mouse, rat, hamster, monkey, and human. It does not seem to work well for neuronal lines. This may be due to the deleterious effect of the calcium on these cells. Cells that grow in suspension are transfected rather poorly by this method, but their stable transformation frequencies seem to be better than those obtained by the Basic Protocol and almost comparable to those obtained by electroporation.

Anticipated Results

The efficiency that can be obtained with calcium phosphate-mediated transfection varies with cell type and other parameters as described. Methods for optimizing these parameters are presented in the discussion of Optimization of Transfection. Up to 10^3 colonies can be obtained by transfecting 1 μg of a plasmid containing a dominant selectable marker (*UNT 9.5*) into 10^6 cells by the HEPES-based protocol. Efficiencies for stable transformants are generally >10- to 100-fold higher when the BES-based protocol is used, with 10% to 50% of the cells on a plate stably transformed (Chen and Okayama, 1987). Transient expression is comparable in the Basic and Alternate Protocols (Chen and Okayama, 1988).

Time Considerations

For the Basic Protocol, preparation of twelve DNA precipitates and addition of the precipitates to the cells takes 1 to 2 hr. Without the ethanol precipitation step, the procedure can be done in 1 hr. With practice, the actual mixing of the CaCl_2 and $2\times$ HeBS solutions will take ~ 1 min. This means that up to eighteen precipitates can be made before the first precipitate is ready to apply to the cells.

For the Alternate Protocol, no ethanol precipitation step has been necessary for either transient or stable transfections. It takes slightly less time than the Basic Protocol because the $2\times$ BBS does not need to be added dropwise to the calcium chloride–DNA solution.

Literature Cited

- Chen, C. and Okayama, H. 1987. High efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* 7:2745-2752.
- Chen, C. and Okayama, H. 1988. Calcium phosphate-mediated gene transfer: A highly efficient system for stably transforming cells with plasmid DNA. *BioTechniques* 6:632-638.
- Graham, F.L. and van der Eb, A.J. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* 52:456.
- Ishiura, M., Hirose, S., Uchida, T., Hamada, Y., Suzuki, Y., and Okada, Y. 1982. Phage particle-mediated gene transfer to cultured mammalian cells. *Mol. Cell. Biol.* 2:607-616.
- Wigler, M., Pellicer, A., Silverstein, S., and Axel, R. 1978. Biochemical transfer of single-copy eucaryotic genes using total cellular DNA as donor. *Cell* 14:725.

Key References

- Chen and Okayama, 1987. See above.
- Ishiura et al., 1982. See above.
- Provides the basis for BES-mediated transfection.*

Contributed by Robert E. Kingston
(HEPES method)
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

Claudia A. Chen (BES method)
National Institute of Mental Health
Bethesda, Maryland

Hiroto Okayama (BES method)
Osaka University
Osaka, Japan

John K. Rose (optimization)
Yale University School of Medicine
New Haven, Connecticut

Transfection Using DEAE-Dextran

Transfection of cultured mammalian cells using diethylaminoethyl (DEAE)-dextran/DNA can be an attractive alternative to other transfection methods in many circumstances. The major advantages of the technique are its relative simplicity and speed, limited expense, and remarkably reproducible interexperimental and intraexperimental transfection efficiency. Disadvantages include inhibition of cell growth and induction of heterogeneous morphological changes in cells. Furthermore, the concentration of serum in the culture medium must be transiently reduced during the transfection. Any of these factors may adversely affect or be incompatible with some bioassays or experimental goals. In addition, for nonstandard cell types there may be a requirement for extensive preliminary investigation of optimal transfection conditions. Together, these factors influence the suitability of this technique to specific purposes. In general, DEAE-dextran DNA transfection is ideal for transient transfections with promoter/reporter plasmids in analyses of promoter and enhancer functions, and is suitable for overexpression of recombinant protein in transient transfections or for generation of stable cell lines using vectors designed to exist in the cell as episomes. The procedure may also be used for expression cloning (Aruffo and Seed, 1987; Kluxen and Lubbert, 1993; Levesque et al., 1991), although electroporation is usually preferred for this purpose (Puchalski and Fahl, 1992).

This unit presents a general description of DEAE-dextran transfection (see Basic Protocol) as well as two more specific protocols for typical experimental applications (see Alternate Protocols 1 and 2).

The Basic Protocol is suitable for transfection of anchorage-dependent (attached) cells. For cells that grow in suspension, electroporation (UNIT 9.3) or lipofection (UNIT 9.4) is usually preferred, although DEAE-dextran-mediated transfection can be used (Fregeau and Bleackley, 1991). For suspension cells, the transfection step should be performed on collected cells that have been resuspended at 10^7 cells/ml in transfection medium, using reagents and conditions that are otherwise similar to those of the Basic Protocol.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and proper sterile technique should be followed accordingly.

NOTE: All culture incubations are performed in a 37°C, 5% CO₂ incubator unless otherwise specified.

GENERAL PROCEDURE FOR DEAE-DEXTRAN TRANSFECTION

Cultured cells are incubated in medium containing plasmid DNA and DEAE-dextran, which form complexes that are taken up by cells via endocytosis. Chloroquine can be included to inhibit degradation of plasmid DNA. Cells are exposed transiently to DMSO or another permeabilizing agent to increase DNA uptake (DMSO “shock”). Important variables include the concentration of DEAE-dextran, the ratio of DNA to DEAE-dextran, the duration of transfection, and the presence and timing of chloroquine exposure (see Critical Parameters). This protocol is suitable for transfection of COS and CV1 cells; Alternate Protocols 1 and 2 describe two examples of transfection experiments.

Materials

- Cells to be transfected and appropriate culture medium (e.g., complete DMEM; APPENDIX 3F) with and without 10% FBS
- 100 mM (1000×) chloroquine diphosphate in PBS, filter-sterilized (store at 4°C)
- Plasmid DNA(s), prepared by CsCl density-gradient centrifugation or affinity chromatography (UNIT 1.7)

BASIC PROTOCOL

Introduction of DNA into Mammalian Cells

9.2.1

Contributed by Tod Gulick

Current Protocols in Molecular Biology (1997) 9.2.1-9.2.10

Copyright © 1997 by John Wiley & Sons, Inc.

TE buffer (*APPENDIX 2*)
 10 mg/ml DEAE-dextran stock solution (see recipe)
 10% (v/v) dimethyl sulfoxide (DMSO) in PBS, filter-sterilized (store up to 1 month at room temperature)
 Phosphate-buffered saline (PBS; *APPENDIX 2*)
 Appropriate-sized tissue culture vessels (Table 9.2.1)
 Inverted microscope
 Additional reagents and equipment for mammalian cell culture (*APPENDIX 3F*)

1. Plate cells at a density to achieve 50% to 75% confluence on the target day for transfection. For COS or CV1 cells, perform a 1:10 split 2 days prior to transfection.

The surface area of various cell culture vessels given in Table 9.2.1 can be used to determine how to split cells to the desired density.

Some cell types including many primary cells show particular sensitivity to the toxicity of DEAE-dextran. These cells should be plated at higher density or transfected after reaching near-confluence.

2. Determine the total volume of medium to be used in the transfection based on the number of culture vessels containing cells to be transfected and the volume per vessel shown in Table 9.2.1. Make up this amount of medium (plus some excess) to contain 2.5% FBS by combining 1 part medium containing 10% FBS with 3 parts serum-free medium.

DEAE-dextran can precipitate in the presence of high medium protein, necessitating use of a low FCS concentration. Alternatively, NuSerum (Collaborative Research), which contains only ~30% serum, can be used at a final concentration of 10%.

3. Add 100 mM (1000×) chloroquine diphosphate stock solution to the 2.5%-FBS-containing transfection medium prepared in step 2 to achieve a final concentration of 100 μM. Warm transfection medium to 37°C.

Table 9.2.1 Surface Areas of Commonly Used Tissue Culture Vessels and Corresponding Appropriate DEAE-Dextran Transfection Medium Volumes

Vessel	Area (cm ²)	Appropriate vol. DEAE-dextran medium ^a (ml)
T175 flask	175	
T150 flask	150	
T75 flask	75	
T25 flask	25	
150-mm dish	148 ^b	10
100-mm dish	55 ^b	4
60-mm dish	21 ^b	2
35-mm dish	8 ^b	1
6-well plate (35-mm wells)	9.4 ^b	1
12-well plate (22-mm wells)	3.8 ^b	0.5
24-well plate (15.5-mm wells)	1.9 ^b	0.25

^aThese volumes are roughly a linear function of vessel surface area. To ensure that cells are completely covered by medium during the transfection, small wells require proportionately larger volumes due to annular sequestration of medium because of surface tension at the periphery.

^bCostar; other manufacturer products may deviate slightly.

Chloroquine is toxic to all cells, so exposure time should be limited to <4 hr. If longer transfection times are required for optimal transfection of a particular cell type, chloroquine should be added during the final hours of the transfection.

4. Dilute plasmid DNA in TE buffer or distilled water to between 1.0 and 0.1 $\mu\text{g}/\mu\text{l}$, depending on the quantity to be transfected. Add the DNA solution directly to the warmed transfection medium to a final concentration of 1.0 $\mu\text{g}/\text{ml}$.

DNA solution(s) should comprise <1% of the total volume of transfection medium, so that the concentration of medium components is not significantly altered. Optimal DNA concentration in the transfection medium may have to be determined experimentally.

Maintaining dilute stock DNA solutions for dedicated use in transfections reduces interexperimental variation as well as the time required to set up transfection experiments.

5. Warm the 10 mg/ml DEAE-dextran stock solution to 37°C and mix thoroughly by inversion. Add to the DNA-supplemented transfection medium to a final concentration of 100 $\mu\text{g}/\text{ml}$ DEAE-dextran and mix by inversion.

The order of addition to the transfection medium is critical. Adding plasmid DNA to medium that has already been supplemented with DEAE-dextran can result in precipitation, seen as a rosy white glob. Optimal DEAE-dextran concentration in the transfection may have to be determined experimentally.

6. Aspirate medium from the 50% to 70% confluent cell cultures (see step 1) and replace with the appropriate volume of 37°C DEAE-dextran/DNA-supplemented transfection medium (see Table 9.2.1). Incubate 4 hr.

Uniformity of transfection efficiency may be improved by placing culture vessels on a rocker platform within the incubator during the transfection to ensure even exposure of cells to DEAE-dextran/DNA in the medium and to avoid desiccation of cells in the center of the vessel. Optimal transfection time may have to be determined experimentally.

7. Examine cells with an inverted microscope.

Cells may appear granular, some cell nuclei may appear pyknotic, and some cell borders may be somewhat ragged. An efficient DEAE-dextran transfection is usually associated with 25% to 75% cell death.

8. Warm the 10% DMSO/PBS to 37°C. Aspirate the transfection medium, note the volume, and replace with 2 to 3 volumes of 37°C DMSO/PBS. Incubate at room temperature for >2 but <10 min. Aspirate the DMSO/PBS and wash the cell layer with a volume of PBS equal to the amount of DMSO/PBS removed. Aspirate and replace with a standard amount of complete medium containing 10% FBS.

Loss of firm cellular anchorage to the culture vessel may occur. Medium exchange and cell washing should therefore involve careful aspiration and pipetting, perhaps by holding the tip of the pipet against a wall of the culture dish or well. It is sometimes advisable to omit the PBS wash (as in the experiment described in Alternate Protocol 1, step 8) and simply add the complete medium, then change the medium a second time several hours after the DMSO shock when cells have recovered and are more firmly adherent.

9. Continue incubating the cells and analyze at times appropriate to the bioassay or intended purpose of the experiment.

The onset and duration of expression of the transfected gene varies from one cell type to another, and especially with the expression vector used. It is advisable to perform a parallel transfection with a readily assayable reporter gene in the identical vector to assess the temporal features of expression. A reporter that is secreted by the cell into the culture medium, such as human growth hormone or secreted alkaline phosphatase, is ideal for this purpose, since aliquots of medium from a single transfection sample can be collected at serial time points. This parallel transfection can also be used in preliminary experiments to optimize transfection conditions.

SAMPLE EXPERIMENT: TRANSFECTION TO TEST PROMOTER FUNCTION

In this protocol for a typical application of DEAE-dextran transfection, a thyroid hormone–response element in a hormone-responsive gene promoter will be mapped. Thyroid hormone (T_3) modulates transcription by activating the thyroid hormone receptor (TR), a transcription factor that binds to specific response elements in target-gene promoters as a heterodimer with the retinoid X receptor (RXR). Expression of TR, but not RXR, is minimal and limiting in fibroblasts, such that overexpression of TR using an expression vector significantly increases transcriptional response to T_3 in these cells. In this example, four promoter/reporter constructs will be tested, representing a promoter 5' deletion series. The hormones, retinoids, and fatty acids in fetal bovine serum (FBS) can interfere with or cause high background in transcription assays in transfected cells. These moieties can be removed from FBS by charcoal treatment (see Reagents and Solutions).

This protocol corresponds step-for-step with Basic Protocol; variations from the original procedure and reagents specific to this particular experiment are noted.

Additional Materials (also see Basic Protocol)

- CV-1 cells (ATCC #CCL 70) growing in 100-mm dish
- Complete DMEM medium (APPENDIX 3F) with and without 10% FBS
- Complete DMEM medium (APPENDIX 3F) with 10% charcoal-treated FBS (see Support Protocol)
- Plasmid DNAs:
 - Control reporter plasmid (e.g., β -galactosidase, secreted alkaline phosphatase, or growth hormone, driven by a viral promoter)
 - Four test promoter constructs (promoter/CAT or promoter/luciferase; UNIT 9.6)
 - Expression plasmid with TR gene insert (pTR)
 - No-insert expression plasmid (p[–])
- Complete DMEM medium (APPENDIX 3F) with 10% charcoal-treated FBS (see Support Protocol), supplemented with 10 nm thyroid hormone (T_3)
- 12-well tissue culture plates
- 100-ml tissue culture dishes
- Additional reagents and equipment for trypsinizing and subculturing monolayer cells (APPENDIX 3F) and analyzing reporter gene activity (UNITS 9.6 & 9.7)

1. Two days prior to the transfection, trypsinize and suspend CV1 cells from a confluent 100-mm dish in 36 ml complete DMEM medium/10% FBS. Place 1 ml complete DMEM/10% FBS in each well of four 12-well plates, then add 250 μ l of the cell suspension to each of these wells. Plate the residual cells in 100-mm dishes for later use, or discard. 12 to 24 hr before the transfection, change medium to complete DMEM/10% charcoal-treated FBS.

The sensitivity of most reporter assays permits use of many fewer cells/transfection conditions than are generally used, with consequent cost savings. CV1 cells in a well of a 12-well plate will provide sufficient reporter-gene activity for most promoter/reporters. In this protocol, triplicate wells for each condition will be analyzed. Four promoter/CAT or promoter/luciferase reporters (UNITS 9.6 & 9.7) will be tested. Cells will include or exclude pTR cotransfection. Transfected cells will be incubated in the presence or absence of T_3 . Thus, four 12-well plates are required for the experiment—i.e., 3 wells (triplicates) \times 4 wells (four reporters) \times 2 wells (with and without pTR cotransfection) \times 2 wells (with and without T_3) = 48 wells. These 48 wells include \sim 180 cm^2 (see Table 9.2.1) such that approximately one-third of the cells on one 100-mm (55 cm^2) dish will be used in order to achieve a 1:10 split.

Addition of suspended cells to empty tissue culture plates or wells results in an uneven distribution of adherent cells, which can introduce undesirable intersample variability; hence medium is added to the wells first.

Preincubation in medium supplemented with charcoal-stripped FBS (see Support Protocol) ensures that any hormones of interest are absent from medium bathing the control cells during subsequent experiments.

2. Add 7 ml DMEM/10% charcoal-treated FBS to 21 ml serum-free DMEM to make 28 ml of DMEM/2.5 % FBS.

In this experiment, the FBS in the DMEM/10 has been stripped of low molecular weight hydrophobic moieties, including T_3 , using charcoal.

The amount here was calculated as 48 wells multiplied by 0.5 ml/well, and a small excess was added. In experiments in which the transfection medium will be divided into multiple aliquots carrying different plasmids or plasmid combinations, it is useful to carry a volume excess throughout the preparation of the separate transfection media to adjust for pipet-calibration errors.

3. Add 28 μ l 100 mM chloroquine diphosphate to the medium and place the tube in a 37°C water bath.
4. Dilute the reporter plasmids (control and test promoter constructs) in TE buffer to 1 μ g/ μ l; dilute the TR (pTR) and no-insert (p[-]) expression plasmids in TE buffer to 0.2 μ g/ μ l. Add 14 μ l of the diluted reporter plasmid to the transfection medium (final concentration, 0.5 μ g/ml). Divide this medium into four equal 6.6-ml aliquots and add 6.6 μ l of a test promoter construct to each separate tube (final concentration, 1 μ g/ml). Divide each of these into two tubes, each containing 3.2 ml. Add 8 μ l pTR to one set of four tubes and 8 μ l p[-] to the other set of four tubes (final concentration, 0.5 μ g/ μ l). Use each of these eight transfection medium samples to transfect cells within six wells in step 6, below.

All cells will be transfected with a control reporter—e.g., β -galactosidase, secreted alkaline phosphatase, or growth hormone—driven by a viral promoter. Each of the four test promoter/reporters will be transfected into cells within 12 wells. A TR expression vector will be transfected into 6 of each of these sets of 12 wells. Cells in triplicate wells in each of these conditions will be cultured in medium supplemented with T_3 , while those in the other triplicate wells will be cultured in medium devoid of T_3 .

Plasmids that are included in more than one transfection condition are added prior to division of medium into separate aliquots to ensure that these samples receive equivalent amounts of plasmid DNA. The final concentration of DNA in each transfection medium sample should be equivalent. For this and other reasons, an “empty vector” should be used as described for p[-] above.

5. Warm the stock DEAE-dextran to 37°C, mix by inversion, and add 32 μ l to each of the eight tubes containing DNA-supplemented transfection medium (final DEAE-dextran concentration, 100 μ g/ml). Mix by gentle inversion.
6. Aspirate medium from six wells on one plate, and replace with one of the eight transfection-medium samples prepared in step 4 at 500 μ l per well. Repeat for each of the different transfection-medium samples. Incubate 4 hr.
7. Examine cells with inverted microscope.
8. Aspirate medium from wells in one plate and add 1 ml DMSO/PBS per well. Repeat for each plate. Return to the first plate, aspirate the DMSO/PBS and replace with 1 ml/well complete DMEM/10% charcoal-treated FBS. Incubate cells for 4 to 12 hr. Aspirate medium from wells and replace the medium from one set of triplicate wells for each transfection condition with 1 ml/well T_3 -supplemented DMEM/10% char-

coal-treated FBS and the other set of triplicate wells for each transfection condition with (T_3 -unsupplemented) DMEM/10% charcoal-treated FBS.

9. Incubate cells 24 to 48 hr. Aspirate 500 μ l of medium from each separate well and save for control reporter activity determinations in order to normalize test reporter activities. Wash wells with PBS and harvest cells for CAT or luciferase activity measurements (UNITS 9.6 & 9.7).

Reporter activity is used to confirm and quantitate promoter T_3 responsiveness ($+T_3/-T_3$), to verify a direct transcriptional response mediated by the TR (augmentation of response to T_3 in pTR-cotransfected cells), and to map the T_3 -responsive region of the promoter (region present in a T_3 -responsive promoter/reporter and absent in an unresponsive one).

The use of a control reporter permits normalization for transfection efficiency, and for nonspecific-stimulus effects on gene expression. In many cases, when using this transfection technique, adequate and informative preliminary experiments can be conducted without this control (since transfection efficiency is so uniform), thereby saving time and expense. This is seldom possible when using other transient-transfection techniques, especially calcium phosphate coprecipitation (UNIT 9.1), where transfection efficiency varies markedly within an experiment. Of course, initial experiments should be performed to exclude nonspecific stimulus effects.

ALTERNATE PROTOCOL 2

SAMPLE EXPERIMENT: TRANSFECTION TO TEST ENZYME STRUCTURE/ACTIVITY RELATIONSHIPS

In this experimental application of DEAE-dextran transfection, an enzyme and several enzyme mutants are overexpressed in COS cells to provide material for kinetic analyses in a structure/activity analysis. A vector designed for high-level expression that replicates in SV40-transformed cells (e.g., CDM8) will be used.

This protocol corresponds step-for-step with Basic Protocol; variations from the original procedure and reagents specific to this particular experiment are noted.

Additional Materials (also see Basic Protocol)

- COS cells (ATCC #1650) growing in 100-mm dish
- Complete DMEM medium (APPENDIX 3F) with and without 10% FBS
- Control plasmid containing reporter gene (e.g., luciferase, CAT, or secreted alkaline phosphatase)
- CDM8 vectors containing gene for wild-type enzyme and genes for four mutant enzymes
- 100-ml tissue culture dishes

Additional reagents and equipment for analyzing reporter gene activity (UNITS 9.6 & 9.7) and analysis of recombinant proteins (Chapter 10)

1. Two days prior to the transfection, split five 100-mm dishes of confluent COS cells into 50 100-mm dishes.

Ten 100-mm dishes of COS cells will provide sufficient recombinant enzyme activity for kinetic analyses. Recombinant wild-type enzyme and four mutant enzymes will be overexpressed. Thus, 50 dishes are required for this experiment.

2. Add 50 ml DMEM/10% FBS to 150 ml serum-free DMEM to make 200 ml DMEM/2.5% FBS.

This was calculated as 50 dishes at 4 ml/100-mm dish (see Table 9.2.1).

3. Warm the transfection medium in a 37°C water bath. Do not add chloroquine diphosphate.

Chloroquine treatment increases DEAE-dextran transfection efficiency, but may reduce the amount of recombinant protein produced by the transfected cells. It is advisable to test this in early pilot experiments.

4. Dilute the reporter gene plasmid and each of the CDM8/enzyme expression vectors to 1 µg/µl in TE buffer. Add 20 µl reporter plasmid to the transfection medium and mix (final concentration, 0.1 µg/ml). Divide the medium into five aliquots of 40 ml. Add 160 µl of one CDM8/enzyme expression plasmid to one aliquot (final concentration, 4 µg/ml) and repeat this for each of the CDM8/enzyme expression plasmids.

All cells will be transfected with a control plasmid containing a reporter gene such as luciferase, CAT, or secreted alkaline phosphatase to evaluate transfection efficiency. Each of the five enzyme expression vectors will be used to transfect cells in ten dishes.

If the COS cells have an endogenous activity identical or similar to the activity of the recombinant protein to be overexpressed, it may be prudent to include cells that are transfected with “empty vector” in experiments of this type to permit parallel assays of endogenous COS cell activities for “background” subtraction.

Plasmids may compete for replication and/or transcription factors. The control reporter plasmid can be included at low concentration because the reporter has a very sensitive assay.

5. Warm the stock DEAE-dextran to 37°C, mix by inversion, and add 800 µl to each of the tubes containing DNA-supplemented transfection medium (final DEAE-dextran concentration, 200 µg/ml). Mix by gentle inversion.
6. Aspirate medium from ten dishes and replace with 4 ml of the appropriate DEAE-dextran/DNA-supplemented transfection medium. Repeat for each set of ten dishes. Incubate 3 hr.
7. Examine cells with inverted microscope. Continue the transfection until some cells appear slightly granular.
8. Aspirate transfection medium from one of the sets of ten dishes and replace with 10 ml/dish DMSO/PBS. After 2 min, aspirate the DMSO/PBS and wash gently with 10 ml PBS. Aspirate PBS and add 10 ml DMEM/10 FBS. Repeat for each set of ten dishes.
9. Incubate cells 48 to 96 hr. Harvest cells as appropriate for recombinant protein assays. Save aliquots of culture medium (for secreted control reporter) or cell extract (intracellular reporter) to determined transfection efficiency.

Normalization for transfection efficiency may not be necessary, particularly in circumstances where an independent assay for recombinant protein production is available, as might be provided by a specific antibody.

CHARCOAL STRIPPING OF FETAL BOVINE SERUM

Activated charcoal is used to remove low-molecular-weight lipophilic compounds from serum including hormones, retinoids, and fatty acid ligands of nuclear receptor transcription factors.

Materials

Fetal bovine serum (FBS), heat-inactivated (*APPENDIX 3F*)

Activated charcoal, acid-washed (Sigma)

Ultracentrifuge with Beckman SW 28, JA-20.1, or equivalent swinging-bucket rotor

SUPPORT PROTOCOL

Introduction of DNA into Mammalian Cells

9.2.7

1. Add 2 g activated charcoal per 100 ml heat-inactivated FBS. Add a stir-bar and place on a stir plate in a 4°C cold room or refrigerator. Stir for 2 hr.

Although it is difficult to maintain sterile conditions during charcoal stripping, care should be taken to avoid flagrant contamination and to keep the serum at 4°C.

2. Collect serum in centrifuge tubes and centrifuge 30 min at $72,000 \times g$ (20,000 rpm in SW 28 rotor) or 60 min at $51,500 \times g$ (20,000 rpm in JA-20.1), 4°C.

3. Gently pour off serum from each centrifuge tube into a sterile beaker, divide into aliquots in sterile conical tubes, and store frozen at -20°C .

Some residual charcoal may be present but the serum should be only lightly peppered with charcoal after centrifugation.

4. Prior to use, thaw a tube of charcoal-stripped medium and use immediately to supplement complete medium. Filter sterilize the medium using a 0.22- μm bottle-top filter and collect in a sterile bottle. Store at 4°C.

The serum is filtered after it is added to the medium because the undiluted serum is too viscous to filter readily.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

DEAE-dextran stock solution, 10 mg/ml (100 \times)

Make a 10 mg/ml stock solution of diethylaminoethyl (DEAE)-dextran (mol. wt. $\sim 500,000$ Da; Sigma) in tissue culture-quality PBS (APPENDIX 2). Mix well, filter sterilize using an 0.22- μm filter, mix again, divide into aliquots, and store up to 3 months at 4°C.

Warm to 37°C and mix well by inversion immediately before each use.

COMMENTARY

Background Information

The mechanism by which cells take up DNA in DEAE-dextran-mediated transfection appears to involve endocytosis after adsorption of DNA/DEAE-dextran complexes onto cells. The advantages of the technique relate largely to the relative simplicity, limited expense, and lack of interexperimental and intraexperiment variability in transfection efficiency.

Critical Parameters

There are several crucial parameters in the transfection procedure, and the weight of importance of each differs for different cell types. Furthermore, maximizing efficiency of a transfection (percent of cells transfected) does not necessarily correspond with optimizing the desired goal of the transfection, such as production of recombinant protein (Kluxen and Lubbert, 1993) or maintenance of colony-forming potential of cells containing transfected DNA (Puchalski and Fahl, 1992). Systematic analyses of variables in DEAE-dextran/DNA trans-

fection for particular cell types have been reported (Fregeau and Bleackley, 1991; Puchalski and Fahl, 1992; Kluxen and Lubbert, 1993; Yang and Yang, 1997), and these studies provide useful information. However, the existence of numerous variables and their mutual codependence makes interpolation, extrapolation, or guessing transfection conditions toward any specific end for any particular cell type difficult to extract from literature reports. Thus, the investigator should carry out a modest set of pilot experiments to optimize conditions.

The major variables that influence DEAE-dextran/DNA transfection include: (1) DEAE-dextran concentration; (2) DNA concentration and the ratio of DNA concentration to DEAE-dextran concentration; (3) duration of transfection; (4) use of chloroquine; (5) use of permeabilizing agents; and (6) serum concentration. The influence of each of these factors on the various goals of DNA transfection will be addressed.

The concentration of DEAE-dextran used in transfections varies from 50 to 500 $\mu\text{g/ml}$. There is an inverse relationship between the concentration used and the duration of exposure prior to onset of cytotoxicity for all cell types. In initial experiments, it may be best to use a low concentration (100 or 200 $\mu\text{g/ml}$) to provide an adequate temporal window to evaluate cytotoxicity microscopically. Lethal cytotoxicity can precede microscopic evidence of the same, such that transfections with high concentrations of DEAE-dextran can result in near-complete cell loss even when most cells appear healthy at the end of the transfection period.

The DNA concentration used depends in part on the vector, cell type, and the purpose of the transfection. For maximal cell adsorption of complexes and maximal transfection efficiency, the ratio of transfection-medium DEAE-dextran concentration to DNA concentration should be 40:1 to 50:1, perhaps because of the existence of an optimal electrostatic cell membrane/DEAE-dextran/DNA-complex interaction (Yang and Yang, 1997). Thus, at a DEAE-dextran concentration of 200 $\mu\text{g/ml}$, plasmid DNA should be included at 4 to 5 $\mu\text{g/ml}$. Since this ratio influences transfection efficiency, "empty vector" should be included in controls at the same concentration as test vectors. This ideal [DEAE-dextran]/[DNA] ratio may depend on transfection-medium serum concentration.

The amounts of DNA to be used in cotransfection experiments again depend on the intended purpose of each component transfection. When a control reporter plasmid is used to simply normalize for transfection efficiency, it is often possible and desirable to include a low concentration of the plasmid, since reporter enzyme assays are generally exquisitely sensitive. In experiments in which simultaneous cotransfection of single cells with two plasmids is a goal, it is appropriate to add the plasmids at the same or similar concentrations. Furthermore, transfection efficiency is of paramount importance in this circumstance, since the percentage of cells that take up both plasmids is a fraction of those that take up each alone. This is particularly important in a case like that presented in Alternate Protocol 1, where coexistence of the transcription-factor expression vector and the promoter/reporter plasmid in a single cell is important.

DEAE-dextran-mediated cytotoxicity is a function of exposure time. In general, efficient transfections can be achieved with a 4-hr incu-

bation if other parameters are adjusted appropriately. This is a convenient transfection duration since it corresponds to the maximal allowable period of cell exposure to chloroquine. Shorter transfection times may be appropriate for some cell types.

The utility of chloroquine in increasing the efficiency of DEAE-dextran transfection is well documented. However, there may be an attenuation in the amount of transfected gene expression (Kluxen and Lubbert, 1993) and increased cell loss and disruption of cellular morphology (Puchalski and Fahl, 1992) when this reagent is included. Thus, use of chloroquine is appropriate for purposes where transfection efficiency is a dominant priority. In other cases, preliminary experiments should evaluate the impact of this agent on the bioactivity desired. While some investigators add chloroquine in the final stage of (or after) a longer transfection, it is generally easier to increase the DEAE-dextran concentration in this circumstance to permit a single 4-hr transfection in medium containing all components.

Using a final cell "shock" with a permeabilizing agent markedly increases DEAE-dextran transfection efficiency without additional cytotoxicity (Lopata, et al, 1984) and should be used universally, unless this manipulation somehow interferes with the desired bioactivity. DMSO at a concentration of 10% is generally used, although 15% glycerol may be more effective for some cell types. Enhancement of transfection efficiency increases as a function of permeabilizing-agent exposure time up to 2 min, after which there is no additional impact (Sussman and Milman, 1984). Since longer exposure, up to 5 to 10 min, produces no negative effect, the "shock" should be for >2 min, but need not be rigorously timed or reproduced, even between transfected samples within an experiment.

Use of transfection medium supplemented with 10% FBS results in formation of macroscopic protein/DEAE-dextran/DNA complexes that are not compatible with efficient transfection. Serum-free medium can be used during the transfection, as can medium supplemented with either 2.5% FBS or 10% NuSerum. The concentration of serum in the transfection medium affects the extent of DEAE-dextran-mediated and chloroquine-mediated cytotoxicity, with a protective effect provided by the serum. FBS concentration may also influence the optimal ratio of [DEAE-dextran]/[DNA].

Anticipated Results

The efficiency of DEAE-dextran-mediated transfection varies considerably among cell types. One can expect to achieve 20% to 60% transfection efficiency with many cells using this procedure if proper attention is paid to optimization of transfection conditions. Following transfection using this technique, there are generally significant morphological changes in cells, and some of these may weigh against the use of this technique for certain purposes.

Time Considerations

Splitting cells into the required number of plates several days prior to the transfection may take ~30 min. Depending on the complexity of the transfection, preparing transfection media may take 5 min to 1 hr. If stock solutions of DNA, DEAE-dextran, chloroquine, and DMSO/PBS are prepared in advance, 30 min should be adequate for this phase, even in a moderately complex transfection experiment like that presented in Alternate Protocol 1. During the subsequent transfection, cells should be monitored using a microscope periodically after 3 hr. The DMSO or glycerol shock, cell wash, and medium replacement requires only ~10 min. Thus, with minimal but careful planning of experimental design, a transfection can readily be completed in 5 hr.

Literature Cited

- Aruffo, A. and Seed, B. 1987. Molecular cloning of a CD28 cDNA by a high-efficiency COS cell expression system. *Proc. Natl. Acad. Sci. U.S.A.* 84:8573-8577.
- Fregeau, C.J. and Bleackley, R.C. 1991. Factors influencing transient expression in cytotoxic T cells following DEAE-dextran-mediated gene transfer. *Somatic Cell Mol. Genet.* 17:239-257.
- Kluxen, F.-W. and Lubbert, H. 1993. Maximal expression of recombinant cDNAs in COS cells for use in expression cloning. *Anal. Biochem.* 208:352-356.
- Levesque, J.P., Sansilvestri, P., Hatzfeld, A., and Hatzfeld, J. 1991. DNA transfection in COS cells: A low-cost serum-free method compared to lipofection. *Biotechniques* 11:313-318.
- Lopata, M.A., Cleveland, D.W., and Sollner-Webb, B. 1984. High level transient expression of a chloramphenicol acetyl transferase gene by DEAE-dextran mediated DNA transfection coupled with a dimethyl sulfoxide or glycerol shock treatment *Nucl. Acids Res.* 12:5707-5717.
- Puchalski, R.B. and Fahl, W.E. 1992. Gene transfer by electroporation, lipofection, and DEAE-dextran transfection: Compatibility with cell-sorting by flow cytometry. *Cytometry* 13:23-30.
- Sussman, D.J. and Milman, G. 1984. Short-term, high-efficiency expression of transfected DNA. *Mol. Cell. Biol.* 4:1641-1643.
- Yang, Y.-W. and Yang, J.-C. 1997. Studies of DEAE-dextran-mediated gene transfer *Biotechnol. Appl. Biochem.* 25:47-51.

Contributed by Tod Gulick
Massachusetts General Hospital
Charlestown, Massachusetts

Transfection by Electroporation

UNIT 9.3

Electroporation—the use of high-voltage electric shocks to introduce DNA into cells—is a procedure that is gaining in popularity. It can be used with most cell types, yields a high frequency of both stable transformation and transient gene expression, and, because it requires fewer steps, can be easier than alternate techniques (UNITS 9.1, 9.2, 9.4 and introduction to Section I).

The basic protocol describes the electroporation of mammalian cells. The alternate protocol outlines modifications for preparation and transfection of plant protoplasts.

ELECTROPORATION INTO MAMMALIAN CELLS

**BASIC
PROTOCOL**

Electroporation can be used for both transient and stable (UNIT 9.5) transfection of mammalian cells. Cells are placed in suspension in an appropriate electroporation buffer and put into an electroporation cuvette. DNA is added, the cuvette is connected to a power supply, and the cells are subjected to a high-voltage electrical pulse of defined magnitude and length. The cells are then allowed to recover briefly before they are placed in normal growth medium. Factors that can be varied to optimize electroporation effectiveness are discussed in introduction to Section I, and protein expression strategies are discussed in Chapter 16.

Materials

Mammalian cells to be transfected

Complete medium (APPENDIX 3F) without and with appropriate selective agents (UNIT 9.5)

Electroporation buffer, ice-cold

Linear or supercoiled, purified DNA preparation (see step 7)

Beckman JS-4.2 rotor or equivalent

Electroporation cuvettes (Bio-Rad #165-2088) and power source

Additional reagents and equipment for stable transformation in selective medium (UNIT 9.5) and for harvesting transfected cells (UNITS 9.6-9.8 & 14.6)

Prepare the cells for electroporation

1. Grow cells to be transfected to late-log phase in complete medium. Each permanent transfection will usually require 5×10^6 cells to yield a reasonable number of transfectants. Each transient expression may require $1-4 \times 10^7$ cells, depending on the promoter.
2. Harvest cells by centrifuging 5 min at $640 \times g$ (1500 rpm in a JS-4.2 rotor), 4°C.
Adherent cells are first trypsinized (introduction to Chapter 9) and the trypsin inactivated with serum.
3. Resuspend cell pellet in half its original volume of ice-cold electroporation buffer.
The choice of electroporation buffer may depend on the cell line used. See critical parameters for a complete discussion.
4. Harvest cells by centrifuging 5 min as in step 2.
5. Resuspend cells at 1×10^7 /ml in electroporation buffer at 0°C for permanent transfection. Higher concentrations of cells (up to 8×10^7) may be used for transient expression.

**Introduction of
DNA into
Mammalian Cells**

9.3.1

Contributed by Huntington Potter

Current Protocols in Molecular Biology (1996) 9.3.1-9.3.6

Copyright © 2003 by John Wiley & Sons, Inc.

Supplement 62

6. Transfer 0.5-ml aliquots of the cell suspension into desired number of electroporation cuvettes set on ice.

Add DNA and electroporate the cells

7. Add DNA to cell suspension in the cuvettes on ice.

For stable transformation, DNA should be linearized by cleavage with a restriction enzyme (UNIT 3.1) that cuts in a nonessential region and purified by phenol extraction and ethanol precipitation (UNIT 2.1). For transient expression, the DNA may be left supercoiled. In either case, the DNA should have been purified through two preparative CsCl/ethidium bromide equilibrium gradients (UNIT 1.7) followed by phenol extraction and ethanol precipitation. The DNA stock may be sterilized by one ether extraction (UNIT 2.1); the (top) ether phase is removed and the DNA solution allowed to dry for a few minutes to evaporate any remaining ether.

For transient expression, 10 to 40 μg is optimal. For stable transformation, 1 to 10 μg is sufficient. Cotransfection (UNIT 9.5), although not recommended because of the work required to select and test transformants, can be done with 1 μg of a selectable marker containing DNA and 10 μg of the DNA containing the gene of interest.

8. Mix DNA/cell suspension by holding the cuvette on the two “window sides” and flicking the bottom. Incubate 5 min on ice.
9. Place cuvette in the holder in the electroporation apparatus (at room temperature) and shock one or more times at the desired voltage and capacitance settings.

The number of shocks and the voltage and capacitance settings will vary depending on the cell type and should be optimized (critical parameters; see also introduction to Section I).

10. After electroporation, return cuvette containing cells and DNA to ice for 10 min.

Culture and harvest the transfected cells

11. Dilute transfected cells 20-fold in nonselective complete medium and rinse cuvette with this same medium to remove all transfected cells.
- 12a. For stable transformation: Grow cells 48 hr (about two generations) in nonselective medium, then transfer to antibiotic-containing medium.

Selection conditions will vary with cell type. For example, neo selection generally requires ~400 $\mu\text{g}/\text{ml}$ G418 in the medium. XGPRT selection requires 1 $\mu\text{g}/\text{ml}$ mycophenolic acid, 250 $\mu\text{g}/\text{ml}$ xanthine, and 15 $\mu\text{g}/\text{ml}$ hypoxanthine in the medium (see UNIT 9.5).

It is often convenient to plate adherent cells at limiting dilution (see UNIT 11.8) immediately following the shock, or suspension cells at the time of antibiotic addition.

- 12b. For transient expression: Incubate cells 50 to 60 hr, then harvest cells for transient expression assays.

Transfected cells can be visualized by standard transient expression assays (UNITS 9.6A-9.7).

ALTERNATE PROTOCOL

ELECTROPORATION INTO PLANT PROTOPLASTS

This is a modification of the basic protocol that is intended for use with plant cells. Plant cells are stripped of their cell walls and DNA is introduced into the resulting protoplasts.

Additional Materials

- 5-mm strips (1 g dry weight) sterile plant material
- Protoplast solution
- Plant electroporation buffer
- 80- μm -mesh nylon screen
- Sterile 15-ml conical centrifuge tube

Additional reagents and equipment for plant RNA preparation (UNIT 4.3)

1. Obtain protoplasts from carefully sliced 5-mm strips of sterile plant material by incubating in 8 ml protoplast solution for 3 to 6 hr at 30°C on a rotary shaker.
2. Remove debris by filtration through an 80- μ m-mesh nylon screen.
3. Rinse screen with 4 ml plant electroporation buffer. Combine protoplasts in a sterile 15-ml conical centrifuge tube.
4. Centrifuge 5 min at $300 \times g$ (1000 rpm in a JS-4.2 rotor). Discard supernatant, add 5 ml plant electroporation buffer, and repeat wash step. Resuspend in plant electroporation buffer at $1.5\text{--}2 \times 10^6$ protoplasts/ml.

Protoplasts can be counted with a hemacytometer (UNIT 1.2).

5. Carry out electroporation as described for mammalian cells (steps 6 to 11 of the basic protocol). Use one or several shocks at 1 to 2 kV with a 3- to 25- μ F capacitance as a starting point for optimizing the system.

Alternatively, use 200 to 300 V with 500 to 1000 μ F capacitance if the phosphate in the electroporation buffer is reduced to 10 mM final.

6. Harvest cells after 48 hr growth and isolate RNA, assay for transient gene expression, or select for stable transformants.

Protoplasts can also be selected and grown into full transgenic plants (Rhodes et al., 1988).

REAGENTS AND SOLUTIONS

Electroporation buffers

Choice of electroporation buffer depends on the cells being used in the experiment (see critical parameters). The following buffers (stored at 4°C) can be used:

1. PBS (APPENDIX 2) without Ca^{++} or Mg^{++}
2. HEPES-buffered saline (HeBS; UNIT 9.1)
3. Tissue culture medium without FCS (introduction to Chapter 9)
4. Phosphate-buffered sucrose: 272 mM sucrose/7 mM K_2HPO_4 (adjusted to pH 7.4 with phosphoric acid)/1 mM MgCl_2

Plant electroporation buffer

Prepare in PBS (APPENDIX 2):

0.4 M mannitol
5 mM CaCl_2
Store at 4°C

Protoplast solution

2% (w/v) cellulase (Yakult Honsha)
1% (w/v) macerozyme (Yakult Honsha)
0.01% (w/v) pectylase
0.4 M mannitol
40 mM CaCl_2
10 mM 2-[*N*-morpholino]ethanesulfonic acid (MES), pH 5.5
Prepare fresh before use

COMMENTARY

Background Information

DNA transfection by electroporation is a technique that is applicable to perhaps all cell types. It yields a high frequency of stable transformants and has a high efficiency of transient gene expression. Electroporation makes use of the fact that the cell membrane acts as an electrical capacitor that is unable to pass current (except through ion channels). Subjecting membranes to a high-voltage electric field results in their temporary breakdown and the formation of pores that are large enough to allow macromolecules (as well as smaller molecules such as ATP) to enter or leave the cell. The reclosing of the membrane pores is a natural decay process that is delayed at 0°C.

During the time that the pores are open, nucleic acid can enter the cell and ultimately the nucleus. Linear DNA with free ends is more recombinogenic and more likely to be integrated into the host chromosome to yield stable transformants. Supercoiled DNA is more easily packaged into chromatin and is generally more effective for transient gene expression.

The use of high-voltage electric shocks to introduce DNA into cells was first performed by Wong and Neumann using fibroblasts (Wong and Neumann, 1982; Neumann et al., 1982). The technique was then generalized (Potter et al., 1984) to all cell types—even those such as lymphocytes that, unlike fibroblasts, cannot be transfected with other procedures (e.g., calcium phosphate or DEAE-dextran DNA coprecipitates). Although whole plants or leaf tissue have been reported to be transfectable by electroporation, plant cells must generally be made into protoplasts before DNA can be easily introduced into them (alternate protocol; Fromm et al., 1985; Ou-Lee et al., 1986). Like mammalian cells, plant protoplasts may be electroporated under a variety of electrical conditions (critical parameters). Both high voltage with low capacitance (short pulse duration) or low voltage with high capacitance (long pulse duration) have been used to achieve successful gene transfer (Chu et al., 1991).

The wide use of electroporation has been made possible in large part by the availability of commercial apparatuses that are safe and easy to use and that give extremely reproducible results. Designs of these machines vary substantially, but fall into two basic categories that use different means of controlling pulse duration and voltage (the two electrical parameters that govern pore formation). One kind

uses a capacitor discharge system to generate an exponentially decaying current pulse, and the other generates a true square wave (or an approximation thereof). The capacitor discharge instruments charge their internal capacitor to a certain voltage and then discharge it through the cell-DNA suspension. Both the size of the capacitor and the voltage can be varied. Because the current pulse is an exponentially decaying function of (1) the initial voltage, (2) the capacitance setting of the instrument, and (3) the resistance of the circuit (including the sample), changing the capacitor size to allow more (or less) charge to be stored at the voltage will result in longer (or shorter) decay times and hence a different effective pulse duration. In contrast, square wave generators control both the voltage and pulse duration with solid-state switching devices. They also can produce rapidly repeating pulses.

Most of our electroporation experiments have used the Bio-Rad Gene Pulser, a capacitor discharge device, but are directly applicable to other capacitor discharge devices, and with some adjustment to square wave generators. Capacitor discharge devices are also available from GIBCO/BRL, BTX, Hoefer Scientific, and International Biotechnologies (see *APPENDIX 4* for suppliers' addresses). These machines, either in a single unit or through add-on components, can deliver a variety of electroporation conditions suitable for most applications. Square wave generators are available from BTX or Baekon and offer great control over pulse width, allow multiple, rapid pulses, and can be more effective for cells that are very sensitive or otherwise difficult to transfect. These machines are generally more expensive. It has become apparent that alternating current pulses at ~100 kHz may be the most effective wave form for electroporation and possibly electrofusion (Chang, 1989). However, dedicated electroporation devices utilizing such waves are not yet commercially available and must be constructed from components. For a complete discussion of electroporation instruments, see Potter (1988).

Electroporation can be easier to carry out than alternative techniques, which is why it is becoming popular. Its drawback for use with transient analysis is that almost five-fold more cells and DNA are needed than with either calcium phosphate- or DEAE-dextran-mediated transfection (*UNITS 9.1, 9.2, & 16.12*). The main difference between electroporation and

calcium phosphate coprecipitation procedures is the state of the integrated DNA after selection in appropriate antibiotic media. In the case of calcium phosphate, the amount of DNA taken up and integrated into the genome of each transfected cell is in the range of 3×10^6 bp. As a result, the transfected DNA often integrates as large tandem arrays containing many copies of the transfected DNA. This would be an advantage when transfection of genomic DNA into recipient cells and selection for some phenotypic change such as malignant transformation is desired; here a large amount of DNA integrated per recipient cell is essential. In contrast, electroporation can be adjusted to result in one to many copies of inserted DNA per recipient cell. This would be an advantage for gene expression studies, as the identity of the particular copy responsible for the gene expression can be controlled.

Critical Parameters

As discussed above, the two parameters that are critical for successful electroporation are the maximum voltage of the shock and the duration of the current pulse (see also introduction to Section I). The voltage and capacitance settings must be optimized for each cell type, with the resistance of the electroporation buffer being critical for choosing the initial instrument settings. The guidelines presented in this unit are meant to be adapted according to the manufacturers' instructions and the individual investigator's needs. Optimal stable and transient transformation occur at about the same instrument settings, so transient expression can be used to optimize conditions for a new cell type.

For low-resistance (high-salt) buffers such as PBS, HeBS, or tissue culture medium, start with a capacitor setting of 25 μ F and a voltage of 1200 V for 0.4-cm cuvettes, then increase or decrease the voltage until optimal transfection is obtained (usually at ~40% to 70% cell viability as detected by trypan blue exclusion; *UNIT 9.6A*). For many cell types, the choice between PBS, HeBS, and tissue culture medium is arbitrary. However, some cells (especially primary cells) are very easily killed and thus electroporate poorly at the high voltages needed for PBS or HeBS electroporation buffers. Particularly sensitive cells seem to prefer tissue culture medium, though it has been shown that the calcium and magnesium ions in the medium lower the electroporation efficiency (Neumann et al., 1982). Phosphate-buffered sucrose has the advantage that it can be optimized at voltages several hundred volts below those used

with PBS or HeBS. Alternatively, Chu et al. (1991) found many sensitive cells were electroporated more effectively in HeBS with a low voltage/high capacitance setting that resulted in at least 10-fold longer pulse duration. For these conditions, start at 250 V/960 μ F and change the voltage up to 350 V or down to 100 V in steps to determine optimal settings.

Keeping cells on ice (at 0°C) often improves cell viability and thus results in higher effective transfection frequency, especially at high power which can lead to heating (Potter et al., 1984). However, Chu et al. (1991) found that under low voltage/high capacitance conditions, some cell lines electroporate with higher efficiency at room temperature. Therefore, steps 6 to 10 of the basic protocol should be carried out separately at both temperatures to determine the optimum conditions for a new cell line.

Another factor contributing to cell death appears to be the pH change that results from electrolysis at the electrodes. This problem can be reduced by replacing some of the ionic strength of the PBS with extra buffer (e.g., 20 mM HEPES, pH 7.5).

Optimal parameters for plant electroporation differ depending on whether tissue culture cells or various parts of the whole plant are used as a source of protoplasts. In particular, the high salt in PBS can be damaging to protoplasts freshly produced from plant tissue. Replacing the NaCl in PBS with 135 mM LiCl may increase CAT transient gene expression (*UNIT 9.6A*) in electroporated plant protoplasts 4- to 70-fold (Saunders et al., 1989). Alternatively, an electroporation buffer of 0.6 M mannitol/25 mM KCl for leaf cells, or 0.7 M mannitol/40 mM KCl/4 mM MES (pH 5.7)/1 mM 2-ME added for root and stem cells, is recommended (Sheen, 1990). In addition, 0.1% BSA/15 mM 2-ME/1 mM MgCl₂ can be added to either protoplast isolation buffer and the CaCl₂ reduced to 1 mM final. A low salt concentration in the electroporation buffer reduces the optimal capacitance setting to 200 μ F.

Anticipated Results

The efficiency of transfection by electroporation is dependent upon cell type. For fibroblasts, which are easily transfected by calcium phosphate or DEAE-dextran coprecipitation (*UNITS 9.1 & 9.2*), electroporation gives a stable transformation frequency of 1 in $\sim 10^3$ to 10^4 live cells—approximately that obtainable by the above traditional procedures. For cells refractory to traditional methods, electroporation gives a stable transformation frequency be-

tween 1 in 10^4 to 10^5 for most cell types. Occasionally a cell line (e.g., some T lymphocytes) will transfect poorly under our standard conditions (1 in 10^6), and even this frequency is sufficient to obtain significant numbers of transfectants. In general, cells that transfect efficiently for stable transformants also do so for transient gene expression. Increasing the number of cells and the amount of DNA used in the electroporation for studying transient gene expression can circumvent problems of low transfection efficiency and low promoter/enhancer efficiency.

For plant protoplast electroporation, the frequency of stable transformants is between 1 in 10^2 and 1 in 10^3 dividing cells.

Time Considerations

The entire process of electroporation of mammalian cells will take <1 hr. Electroporation of plant cells requires ≤ 6 hr to prepare the protoplasts and <1 hr for the actual electroporation process. As with other transfection procedures, the experiment should be planned to allow for harvest or splitting of the cells 1 to 2 days after transfection.

Literature Cited

- Chang, D.C. 1989. Cell poration and cell fusion using an oscillating electric field. *Biophys. J.* 56:641-652.
- Chu, G., Hayakawa, H., and Berg, P. 1987. Electroporation for the efficient transfection of mammalian cells with DNA. *Nucl. Acids Res.* 15:1311-1326.
- Fromm, M., Taylor, L.P., and Walbot, V. 1985. Expression of genes transferred into monocot and dicot plant cells by electroporation. *Proc. Natl. Acad. Sci. U.S.A.* 82:5824-5828.

Neumann, E., Schaefer-Ridder, M., Wang, Y., and Hofschneider, P.H. 1982. Gene transfer into mouse lyoma cells by electroporation in high electric fields. *EMBO J.* 1:841-845.

Ou-Lee, T.M., Turgeon, R., and Wu, R. 1986. Uptake and expression of a foreign gene linked to either a plant virus or *Drosophila* promoter in protoplasts of rice, wheat and sorghum. *Proc. Natl. Acad. Sci. U.S.A.* 83:6815-6819.

Potter, H. 1988. Electroporation in biology: Methods, applications, and instrumentation. *Anal. Biochem.* 174:361-373.

Potter, H., Weir, L., and Leder, P. 1984. Enhancer-dependent expression of human κ immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. *Proc. Natl. Acad. Sci. U.S.A.* 81:7161-7165.

Rhodes, C.A., Pierce, D.A., Mettler, I.J., Mascarenhas, D., and Detmar, J.J. 1988. Genetically transformed maize plants from protoplasts. *Science* 240:204-207.

Saunders, J.A., Matthews, B.F., and Miller, P.D. 1989. Plant gene transfer using electrofusion and electroporation. *In* Electroporation and Electrofusion in Cell Biology (E. Neumann, A.E. Sowers, and C.A. Jordan, eds.) p. 343-354. Plenum, New York.

Sheen, J. 1990. Metabolic repression of transcription in higher plants. *Plant Cell* 2:1027-1038.

Wong, T.K. and Neumann, E. 1982. Electric field mediated gene transfer. *Biochem. Biophys. Res. Commun.* 107:584-587.

Key Reference

Potter et al., 1984. See above.

The original paper from which the basic protocol is adapted.

Contributed by Huntington Potter
Harvard Medical School
Boston, Massachusetts

Transfection of Cultured Eukaryotic Cells Using Cationic Lipid Reagents

The development of high-efficiency methods for the introduction of functional genetic material into eukaryotic cells using cationic lipid-based transfection reagents has accelerated biology research in the studies of gene expression, control of cell growth, and cell lineage. Cationic lipid-mediated transfection techniques are commonly used in industrial protein production as well as in some clinical gene therapy protocols.

Most natural lipids are either neutral or negatively charged (anionic). Positively charged (cationic) lipids were first introduced in 1987 (Felgner et al., 1987). The cationic lipids function by spontaneous electrostatic interaction of their positive charges with the negative charges in the backbone of DNA, RNA, or oligonucleotides, condensing the extended macromolecules to a compact structure. The cationic charges and the lipophilic nature of the cationic lipids then allow the condensed aggregates to interact with and cross the negatively charged and hydrophobic cell membrane and enter the target cells. It is also possible to deliver some proteins into cells using cationic lipids (Sells et al., 1995). Most (but not all) cationic lipid reagents used for transfection consist of mixtures of cationic and neutral lipids (e.g., DOPE, cholesterol) that are formulated in water to yield noncovalent structures called liposomes—hollow spheres with aqueous cores and a diameter of 100 to 400 nm (Felgner et al., 1987). Some are dissolved in ethanol and form micelles (Behr et al., 1989). There are many cationic lipid reagents available (see Background Information and Table 9.4.1). They perform with various efficiencies in different applications and target cells.

This unit describes cationic lipid-mediated transfection of a variety of cell types. DNA transfection is presented for adherent mammalian cells (cell lines as well as primary cultures; see Basic Protocol 1), along with a modified protocol for enhanced transfection (see Alternate Protocol). DNA transfection of suspension mammalian cells (lymphoid, myeloid, and leukemic-derived cells) is also presented (see Basic Protocol 2). For adherent mammalian cells, RNA transfection is also covered (see Basic Protocol 3). Finally, DNA transfection is presented for insect cells (see Basic Protocol 4). To determine

Table 9.4.1 Partial Listing of Commercially Available Cationic Lipid Transfection Reagents

Supplier	Reagent	Uses
Invitrogen	PerFect Lipid Transfection Kit (8 lipids)	Test for different cell types
Life Technologies	LipofectAmine 2000	General, rapid
	LipofectAmine Plus	Difficult-to-transfect adherent cells
	LipofectAmine	General
	Lipofectin	Liver and endothelial cells, oligonucleotides
	DMRIE-C	Suspension cells, RNA
Promega	CellFectin	Insect cells
	Transfectam	General
	Tfx Transfection Trio	Test for different cell types
Qiagen	Effectene	General
Roche/BMB	DOTAP	General
	DOSPER	General
	FuGENE 6	General

Contributed by Pamela Hawley-Nelson and Valentina Ciccarone
Current Protocols in Molecular Biology (1999) 9.4.1-9.4.17
Copyright © 1999 by John Wiley & Sons, Inc.

the best transfection conditions for a particular reagent, use the optimization/fine-tuning procedure (see Support Protocol).

NOTE: There are minor differences among the recommended protocols for the various commercially available cationic lipid reagents. The appropriate product profile sheets should be consulted before beginning an experiment.

NOTE: All culture incubations for mammalian cells should be performed in a humidified 37°C, 5% CO₂ incubator. Some media may require altered levels of CO₂ (e.g., for growth of CHO cells in suspension, 8% CO₂ is preferable) to maintain pH 7.4. Insect cells are cultured at 27°C.

BASIC PROTOCOL 1

CATIONIC LIPID-MEDIATED TRANSFECTION OF ADHERENT MAMMALIAN CELLS WITH DNA

This protocol describes the procedure for transfection of DNA into most adherent mammalian cell lines or cultures of primary cells grown attached to culture vessels (Fig. 9.4.1). DNA and lipid reagent are diluted into separate aliquots of serum-free medium, and are then mixed together and allowed to form complexes. Complexes and transfection medium (which may contain serum) are added to the cells in one of two ways. Either the complexes are diluted with transfection medium and this mixture is added to cells that have been rinsed and aspirated (as shown in Fig. 9.4.1), or complexes are added directly to transfection medium that has already been added to the cells (with or without washing). The cells and complexes are incubated together for several hours. After transfection, the volume of the medium is increased if necessary to prevent drying, serum is added if the transfection was serum free, and cells are incubated an additional day or two to allow expression of the transgene.

It is advisable to optimize the conditions for transfection using the Support Protocol. Suggested starting ranges for optimizing the various components of these mixtures are given in Table 9.4.2 for six different popular culture vessel sizes. Transfections can be scaled up to other vessel sizes by increasing the amounts of lipid, DNA, and medium in

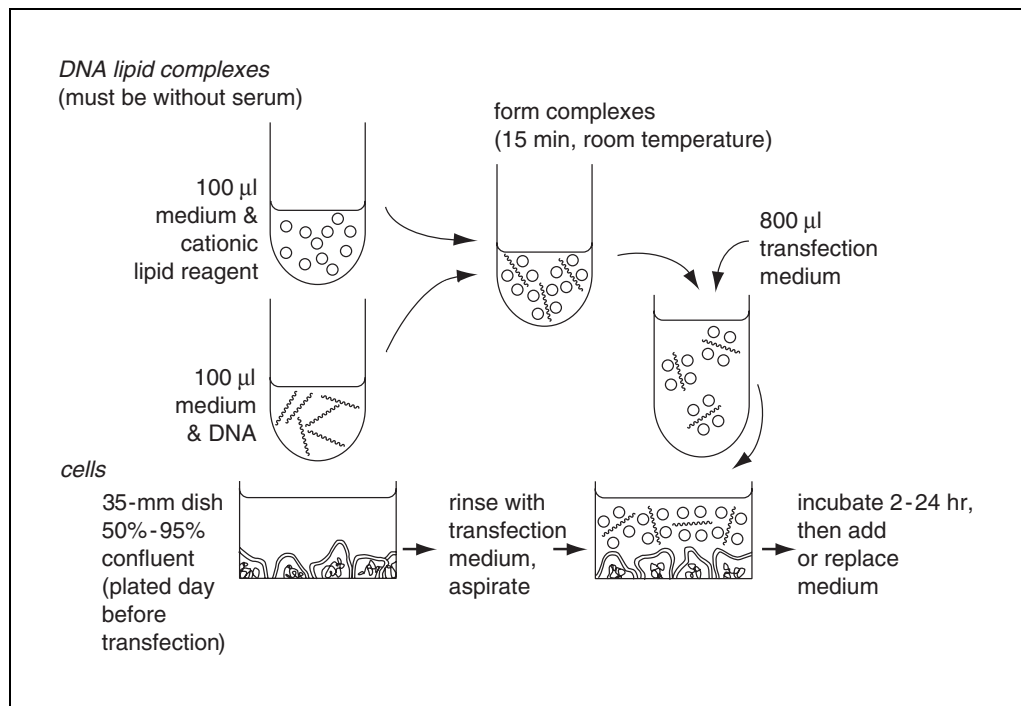


Figure 9.4.1 Diagram of cationic lipid-mediated transfection procedure.

proportion to the difference in surface area. Table 9.4.3 shows the surface areas of several popular culture vessels.

Materials

Adherent cells

Cell culture medium with serum (e.g., complete DMEM, *APPENDIX 3F*)

Dilution medium: serum-free cell culture medium or specialized medium for lipid-mediated transfection (e.g., Opti-MEM I, Life Technologies)

Plasmid DNA, purified by anion-exchange chromatography (e.g., Concert High Purity columns, Life Technologies; or see *UNIT 2.1B* or Goldsborough et al., 1998), cesium chloride density gradient (*UNIT 1.7*), or alkaline lysis (*UNIT 1.6*)

Cationic lipid reagent (see Table 9.4.1)

Polystyrene or polypropylene tubes

Additional reagents and equipment for trypsinization and counting of cells (*APPENDIX 3F*), Xgal staining (*UNIT 9.11*), and selection of stable transformants (*UNIT 9.5*)

1. The day before transfection, trypsinize and count adherent cells (*APPENDIX 3F*). Plate cells in cell culture medium with serum so that they are 50% to 95% confluent on the day of transfection. Avoid antibiotics at the time of plating and during transfection.

The single most important factor in reproducible, high-efficiency transfection is a consistent number of healthy, proliferating cells. Transfection is most efficient when the cells are

Table 9.4.2 Suggested Starting Ranges of Reagents for Lipid-Mediated Transfection^a

Culture vessel	DNA (μg)	Dilution medium (μl)	Cationic lipid reagent (μl)	Transfection medium (ml)	Transfection volume (ml)
	Step 2	Steps 2 & 3	Step 3	Step 5	Step 6
96-well	0.05-0.4	10-25	0.075-1.5	0.08-0.1	0.1-0.15
24-well	0.2-1.6	25-50	0.5-10	0.2-0.5	0.25-0.6
12-well	0.4-3.2	50-100	1-20	0.4-1	0.5-1.2
6-well	1-8	100-250	2.5-50	0.8-2.5	1-3
60-mm	2-16	250-500	5-100	2-5	2.5-6
100-mm	6-48	750-1500	15-300	5-15	6.5-18

^aStep numbers are indicated from Basic Protocol 1. Volumes have been optimized for LipofectAmine 2000. See Support Protocol for additional optimization strategies.

Table 9.4.3 Surface Areas of Commonly Used Culture Vessels

Culture vessel	Surface area (cm ²)
96-well	0.3
24-well	2
12-well	4
6-well	10
35-mm	8
60-mm	20
100-mm	60
150-mm	140
T25	25
T75	75
T150	150

maintained in mid-log growth. Because transfection efficiency is sensitive to culture confluency, cultures should be maintained carefully and passaged frequently, and a standard seeding protocol should be followed from one experiment to the next. Antibiotics may cause some toxicity if present during transfection. For some transfection reagents, the higher cell density is recommended (90% to 95%).

Multiwell cell culture dishes are the easiest to use. Some cells are weakly adherent. If necessary, increase adherence by plating cells on polylysine-coated wells (0.1 mg/ml poly-L-lysine, Sigma; poly-D-lysine-precoated plates, Becton Dickinson).

This protocol is written without specific amounts of reagents. Table 9.4.2 gives the recommended starting amounts for several popular vessels. See Support Protocol for fine tuning the reagent volumes for highest transfection efficiency.

2. On the day of transfection, dilute plasmid DNA into dilution medium in a polystyrene or polypropylene tube and mix. Prepare in bulk for multiple transfections.

Although DNA prepared by anion-exchange chromatography (UNIT 2.1B) or CsCl gradients (UNIT 1.7) yield the best results, DNA prepared by alkaline lysis (miniprep; UNIT 1.6) will work with lower efficiency.

Commercial medium that is specialized for lipid-mediated transfection (e.g., Opti-MEM I) yields the highest efficiency, but other serum-free media may be used.

Polystyrene or polypropylene tubes work well for dilutions. For small-scale transfections, round-bottom 96-well plates with covers are suitable.

3. Dilute cationic lipid reagent into dilution medium in a second tube and mix. Prepare in bulk for multiple transfections.

See Table 9.4.1 for a partial listing of commercially available reagents. A preparation of the cationic lipid DDAB with the neutral lipid DOPE can be made in the laboratory (Rose et al., 1991), although efficiency may be lower than with some commercial preparations.

If using Lipofectin, dilute into Opti-MEM I and allow to incubate at room temperature for 30 to 45 min (Ciccarone and Hawley-Nelson, 1995). If using LipofectAmine 2000, dilute into Opti-MEM I and allow to incubate no more than 30 min (Ciccarone et al., 1999).

4. Combine diluted DNA and diluted cationic lipid reagent, mix, and incubate for 15 min at room temperature.

Incubation times >15 min (up to 6 hr for some reagents) work just as well.

5. While complexes are forming, replace medium on the cells with the appropriate volume of fresh transfection medium.

This step is not necessary if the complexes are diluted with transfection medium or if using LipofectAmine 2000 (see step 6).

This medium can be the same as the dilution medium. It is possible to use serum in the transfection medium at this step.

6. Add DNA-lipid complexes to each well containing cells. Mix complexes into the medium gently, holding the plate at a slant. Incubate 5 hr at 37°C in 5% CO₂.

Adding transfection medium directly to the cells (as described in the steps) helps prevent cells in multiwell plates from drying out, as transfection medium can be added to multiple wells rapidly. Alternatively, complexes can be diluted with the appropriate volume of fresh transfection medium and added to cells from which the cell culture medium has been aspirated (with or without washing, see Fig. 9.4.1). If using LipofectAmine 2000, fresh transfection medium is not required. Undiluted complexes can be added directly to the cell culture medium.

The exposure time may be >5 hr (up to overnight). Be sure that there is sufficient medium to prevent the cells from drying out (it is not necessary to increase other components if this

is done). If necessary to maximize cell growth, replace the medium containing complexes with fresh complete medium after 5 hr incubation.

7. After 5 hr incubation, add cell culture medium to reach normal volume and add serum to bring the final concentration to that of normal cell culture medium.

This step is omitted in some protocols, especially when serum is present during transfection and volumes are adequate to prevent drying.

- 8a. *For transient expression analysis:* Assay cell extracts or stain cells (UNIT 9.11) in situ 24 to 48 hr after the start of transfection, depending on cell type and promoter activity.
- 8b. *For stable expression analysis:* Passage cells into fresh culture medium 1 day after the start of transfection. At 2 days posttransfection, add the appropriate antibiotic to select for expression of the transfected antibiotic-resistance gene (UNIT 9.5).

Several days or weeks of selection are required for stable expression.

ENHANCED CATIONIC LIPID-MEDIATED TRANSFECTION OF ADHERENT MAMMALIAN CELLS WITH DNA

**ALTERNATE
PROTOCOL**

This is an efficient and reproducible protocol for transfection of DNA into most adherent mammalian cell lines or cultures of primary cells grown attached to culture vessels (Dube, 1997; Shih et al., 1997; Fig. 9.4.2). The procedure is essentially as described in Basic Protocol 1, except that DNA is diluted into serum-free medium along with a proprietary enhancing reagent and incubated for 15 min, allowing the formation of precomplexed DNA before addition to the diluted cationic lipid reagent.

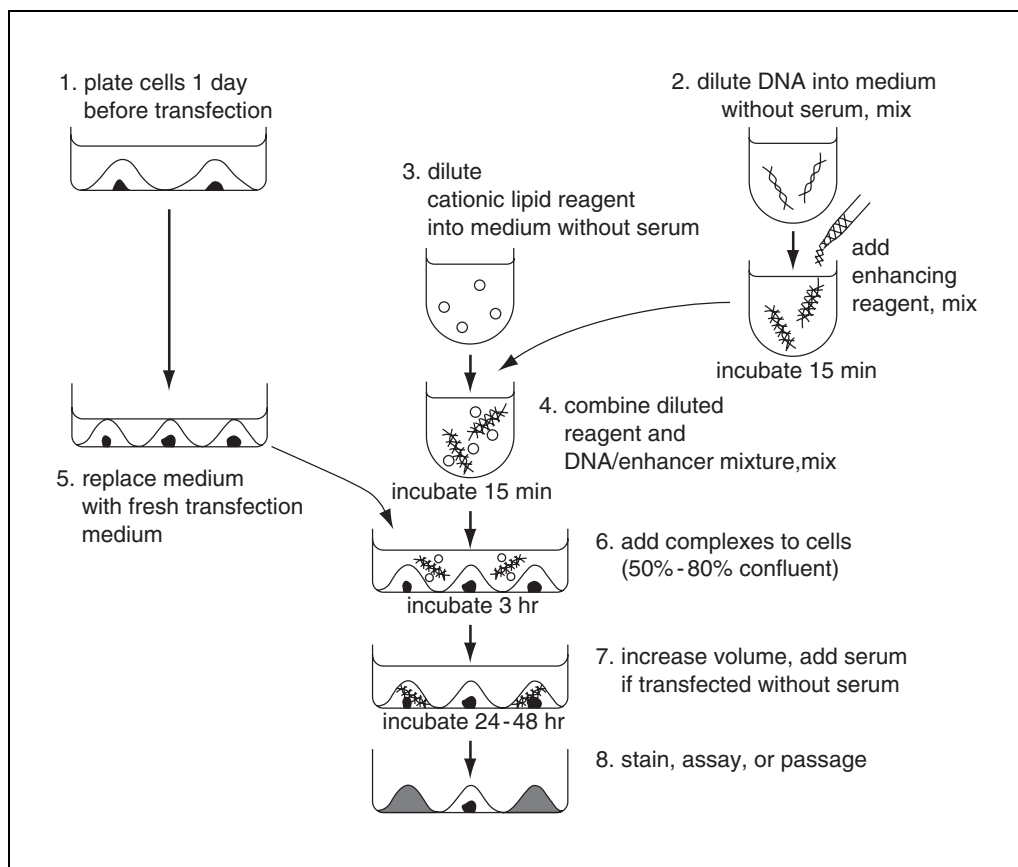


Figure 9.4.2 Diagram of enhanced cationic lipid-mediated transfection procedure.

**Introduction of
DNA into
Mammalian Cells**

9.4.5

Table 9.4.4 Suggested Starting Amounts of Reagents for Transfection with LipofectAmine Plus^a

Culture vessel	DNA (μg)	Plus reagent (μl)	Dilution medium (μl)	LipofectAmine (μl)	Transfection medium (ml)	Transfection volume (ml)
	Step 2	Step 2	Steps 2 & 3	Step 3	Step 5	Step 6
96-well	0.1	1	10	0.5	0.08	0.1
24-well	0.4	4	25	1	0.2	0.250
12-well	0.7	5	50	2	0.4	0.5
6-well	1	6	100	4	0.8	1.0
60-mm	2	8	250	12	2	2.5
100-mm	4	20	750	30	5	6.5

^aStep numbers indicated from Alternate Protocol. It is possible to fine tune transfections by testing a range of lipid and DNA concentrations to obtain optimal efficiency; however, the peak activity is usually a broad plateau with this reagent.

The protocol that follows was developed for LipofectAmine Plus, which is composed of two reagents: LipofectAmine (the cationic lipid reagent) and Plus reagent (the enhancer). One of the main advantages of the Plus enhancer is a high plateau of transfection activity across a broad range of lipid and DNA concentrations. This decreases the necessity for optimization and allows the recommendation of specific starting conditions (Table 9.4.4). Transfections can be scaled up to other vessel sizes by increasing the amounts of lipid, DNA, enhancer, and medium in proportion to the difference in surface area (see Table 9.4.3).

The other commercially available enhanced cationic lipid reagent is Effectene (Qiagen, see Table 9.4.1). If using this product, consult the appropriate product profile sheets before beginning the experiment.

NOTE: Consult annotations of the standard transfection procedure (see Basic Protocol 1) for additional details, which also apply to this protocol.

Additional Materials (also see Basic Protocol 1)

Cationic lipid reagent and enhancer (e.g., LipofectAmine Plus, including LipofectAmine and Plus reagents; Life Technologies)

1. The day before transfection, trypsinize and count adherent cells (*APPENDIX 3F*). Plate cells in cell culture medium with serum so that they are 50% to 80% confluent on the day of transfection. Avoid antibiotics at the time of plating and during transfection.

For this protocol, refer to Table 9.4.4 for recommended amounts of reagents in different vessels.

2. On the day of transfection, dilute plasmid DNA into dilution medium in a polystyrene or polypropylene tube and mix well. Add enhancer, mix, and incubate 15 min at room temperature.

DMEM is preferred over Opti-MEM 1 in this protocol. It is important to add the DNA first and mix well before adding the Plus reagent to avoid precipitation of the DNA. Incubation times >15 min (up to an hour) work just as well.

3. Dilute cationic lipid reagent into dilution medium in a second tube and mix.
4. Combine precomplexed DNA and diluted cationic lipid reagent, mix, and incubate for 15 min at room temperature.

Incubation times >15 min (up to an hour) work just as well when LipofectAmine is the cationic lipid and Plus reagent is the enhancer.

5. While complexes are forming, replace medium on the cells with the appropriate volume of fresh transfection medium.

The medium can be the same as the dilution medium. It is possible to use serum in the transfection medium at this step. In some cells (e.g., HeLa and NIH 3T3), transfection in medium containing serum is as efficient or more efficient than in medium without serum.

6. Add DNA-enhancer-lipid complexes to each well containing cells. Mix complexes into the medium gently, holding the plate at a slant. Incubate at 37°C in 5% CO₂ for several hours.

The exposure time with LipofectAmine Plus may be as short as 3 hr or up to overnight. Be sure there is sufficient medium to prevent the cells from drying out (it is not necessary to increase other components if this is done).

7. After incubation, add cell culture medium to reach normal volume and add serum to bring the final concentration to that of normal cell culture medium.

8. Perform transient or stable expression analysis (see Basic Protocol 1, steps 8a and 8b).

CATIONIC LIPID-MEDIATED TRANSFECTION OF SUSPENSION CELLS WITH DNA

BASIC PROTOCOL 2

This protocol is essentially the same as for adherent cells (see Basic Protocol 1) in that lipid and DNA are diluted separately into dilution medium, mixed, and allowed to form complexes before exposing to cells. However, complexes are formed in the wells of multiwell culture plates, and cells are then distributed into the wells containing complexes and allowed to transfect.

Materials

Dilution medium: cell culture medium without serum or specialized medium for transfection (e.g., Opti-MEM I, Life Technologies)

Cationic liposome reagent (e.g., DMRIE-C or LipofectAmine 2000, Life Technologies; also see Table 9.4.1)

Plasmid DNA, purified by anion-exchange chromatography (UNIT 2.14 or Goldsborough et al., 1998), cesium chloride density gradient (UNIT 1.7), or alkaline lysis (UNIT 1.6)

Cell suspension: 1 × 10⁷ cells/ml in normal cell culture medium without serum or antibiotics

Cell culture medium (e.g., complete DMEM; APPENDIX 3F)

Serum

6-well tissue culture plates

1. To each well of a 6-well tissue culture plate add 0.5 ml dilution medium.

Commercial medium that is specialized for lipid-mediated transfection (e.g., Opti-MEM I), without serum or antibiotics, gives the best results. However, other serum-free media may also be used.

When transfecting in different-sized culture plates, change the amounts of DNA, cationic lipid reagent, and medium in proportion to the difference in surface area (see Table 9.4.3).

2. Add 0, 2, 4, 6, 8, or 12 μl cationic lipid reagent to each well and mix gently by swirling the plate.

DMRIE-C was found to give high efficiency transfection of DNA in Jurkat (human T cell lymphoma), K562 and KG-1 (human myelogenous leukemia), and MOLT-4 (human lymphoblastic leukemia) cell lines. It is a lipid suspension that may settle with time. To ensure that a homogenous sample is taken, mix thoroughly by inverting the tube 5 to 10 times before removing a sample for transfection.

Introduction of DNA into Mammalian Cells

9.4.7

3. Add 0.5 ml dilution medium containing 4 μg plasmid DNA to each well. Mix by swirling plate.

The amount of DNA should be optimized for each cell line.

4. Incubate at room temperature for 15 to 45 min to allow formation of lipid-DNA complexes.

5. Add 0.2 ml cell suspension (2×10^6 cells) to each well and mix gently.

The single most important factor in reproducible, high-efficiency transfection is a consistent number of healthy, proliferating cells. Transfection is most efficient when the cells are maintained in mid-log growth.

6. Incubate several hours at 37°C in a 5% CO₂ incubator.

A 4-hr incubation is adequate for DMRIE-C transfections.

7. To each well add 2 ml cell culture medium containing 1.5 \times the usual amount of serum.

For Jurkat and MOLT-4 cells, addition of 1 $\mu\text{g}/\text{ml}$ phytohemagglutinin (PHA) and 50 ng/ml phorbol myristate acetate (PMA) enhances promoter activity and gene expression. For K562 and KG-1 cells, PMA alone enhances promoter activity.

8. Assay the cells at 24 or 48 hr posttransfection for transient or stable expression (see Basic Protocol 1, step 8a or 8b).

BASIC PROTOCOL 3

CATIONIC LIPID-MEDIATED TRANSFECTION OF ADHERENT CELLS WITH RNA

In this protocol, lipid is first diluted into dilution medium and mixed. RNA is then mixed directly into the diluted lipid and immediately added to cells (which have been rinsed with serum-free medium), and cells are incubated for transfection.

Materials

Adherent cells

Cell culture medium with serum (e.g., complete DMEM; *APPENDIX 3F*)

Dilution medium: serum-free cell culture medium or specialized medium for transfection (e.g., Opti-MEM I, Life Technologies)

Cationic lipid reagent (e.g., DMRIE-C, Life Technologies; also see Table 9.4.1)
mRNA (*UNIT 4.5*)

6-well or 35-mm tissue culture plate

12 \times 75-mm polystyrene tubes

Additional reagents and equipment for trypsinizing, counting, and plating cells

(*APPENDIX 3F*)

1. The day before transfection, trypsinize and count adherent cells (*APPENDIX 3F*). In each well of a 6-well tissue culture plate, or in six 35-mm tissue culture plates, seed $\sim 2\text{--}3 \times 10^5$ cells in 2 ml of the appropriate cell culture medium supplemented with serum.

Transfection is most efficient when the cells are growing rapidly. Cultures should be maintained carefully and passaged frequently. As transfection efficiency may be sensitive to culture confluency, it is important to maintain a standard seeding protocol from experiment to experiment.

2. Incubate at 37°C in a 5% CO₂ incubator until the cells are $\sim 80\%$ confluent.

This will usually take 18 to 24 hr, but the time will vary among cell types.

3. On the day of transfection, wash the cells in each well with 2 ml dilution medium at room temperature.

Commercial medium that is specialized for lipid-mediated transfection (e.g., Opti-MEM I), without serum or antibiotics, gives the best results. However, other serum-free media may also be used.

4. Add 1.0 ml dilution medium to each of six 12 × 75–mm polystyrene tubes.
5. Add 0, 2, 4, 6, 8, or 12 µl cationic lipid reagent to each tube and mix or vortex briefly.

DMRIE-C was found to give high-efficiency transfection of RNA in adherent cell lines (Ciccarone et al., 1995). It is a lipid suspension that may settle with time. To ensure that a homogenous sample is taken, mix thoroughly by inverting the tube 5 to 10 times before removing a sample for transfection.

6. Add 2.5 to 5.0 µg RNA to each tube and vortex briefly.

mRNA that is capped and polyadenylated is translated more efficiently and is more stable within the cell.

7. Immediately add lipid-RNA complexes to washed cells and incubate 4 hr at 37°C in a 5% CO₂ incubator.

The time of exposure of cells to lipid-RNA complexes, as well as the amount of RNA added to the cells, should be adjusted for each cell type.

8. Replace transfection medium with cell culture medium containing serum.
9. Allow cells to express the RNA for 16 to 24 hr and analyze them for expression of the transfected RNA as appropriate for the transgene used.

CATIONIC LIPID–MEDIATED TRANSFECTION OF ADHERENT Sf9 AND Sf21 INSECT CELLS WITH BACULOVIRUS DNA

BASIC PROTOCOL 4

As for transfecting mammalian cells (see Basic Protocol 1), cationic lipid reagent and nucleic acid are diluted separately into serum-free medium and then mixed and allowed to form complexes. Complexes are then diluted with fresh transfection medium and added to the cells for transfection. After the cells are fed and incubated, budded virus can be isolated from the medium.

Materials

Insect cells: Sf9 or Sf21 cells (UNIT 16.9)

Insect medium (UNIT 16.9; e.g., Sf-900 II SFM, Life Technologies) with and without serum and antibiotics

Baculovirus DNA: purified DNA or bacmid DNA miniprep (UNITS 16.9 & 16.10; Anderson et al., 1995)

Cationic lipid reagent (Table 9.4.1)

6-well tissue culture plate

27°C incubator

12 × 75–mm polystyrene tubes, sterile

Additional reagents and equipment for culturing insect cells and harvesting baculovirus from cell supernatants (UNIT 16.10)

1. In each well of a 6-well tissue culture plate, seed $\sim 9 \times 10^5$ insect cells in 2 ml insect medium without serum or antibiotics (UNIT 16.10).

Insect cells must be plated when they are in mid-log growth phase. Cells that have reached stationary phase transfect and infect at very low efficiency. Therefore, it is advisable to

Introduction of DNA into Mammalian Cells

9.4.9

maintain a standard cell passage protocol that keeps the cells in log growth. For Sf9 or Sf21 cells adapted in Sf-900 II SFM, cells are passaged twice weekly to a density of 3×10^5 cells/ml in suspension, and plated for transfection on the third day postseeding, when they are in mid-log phase. For other cell culture media and growth conditions, adjust conditions to maintain similar growth characteristics.

For culture of insect cells, use 50 units/ml penicillin and 50 μ g/ml streptomycin (half the usual final concentration). For transfection, it is preferable to omit antibiotics from the medium to avoid toxicity.

2. Allow cells to attach at 27°C for ≥ 1 hr.
3. For each transfection, dilute 1 to 2 μ g baculovirus DNA into 100 μ l insect medium without serum or antibiotics in a 12 \times 75-mm polystyrene tube.
4. For each transfection, dilute 1.5 to 9 μ l cationic lipid reagent into 100 μ l insect medium without serum or antibiotics in a separate 12 \times 75-mm polystyrene tube.

The suggested amount is 6 μ l, but this should be optimized for each system.

CellFectin gives high-efficiency transfection of DNA in insect cell lines (Anderson et al., 1995). It is a lipid suspension that may settle with time. To ensure that a homogenous sample is taken, mix thoroughly by inverting the tube 5 to 10 times before removing a sample for transfection.

5. Combine the two solutions, mix gently, and incubate at room temperature for 15 to 45 min to form lipid-DNA complexes.
6. For each transfection, add 0.8 ml insect medium without serum or antibiotics to each tube containing lipid-DNA complexes and mix gently.
7. Aspirate medium from cells and overlay diluted lipid-DNA complexes onto the washed cells.

Alternatively, the medium on the cells can be replaced with 0.8 ml fresh insect medium and the undiluted complexes can be added directly to the fresh medium on the cells.
8. Incubate cells for 5 hr in a 27°C incubator. Protect plates from evaporation by putting them in a humidified container.
9. Remove transfection mixture and add 2 ml insect medium containing antibiotics and serum, if desired. Incubate cells in a 27°C incubator for 72 hr.
10. Harvest baculovirus from cell supernatants (UNIT 16.10).

Gene expression may also be evaluated in the cells after removal of virus-containing medium.

SUPPORT PROTOCOL

FINE TUNING OR OPTIMIZING CONDITIONS FOR CATIONIC LIPID REAGENT TRANSFECTIONS

This protocol provides an example of a simple one-step procedure for determining conditions conducive to high-efficiency transfections using cationic lipid reagents in any target cell type. A matrix of DNA and lipid reagent concentrations is used on transfections performed in a multiwell plate (Fig. 9.4.3). Once the best conditions have been determined, the transfections may be scaled up to larger vessels using the relative surface area (see Table 9.4.3) to increase the amounts of all reagents proportionately. This protocol can be modified for use with any transfection protocol.

Additional Materials (also see *Basic Protocol 1* and *Alternate Protocol*)

- 24-well tissue culture plates
- 96-well round-bottom plates (sterile, with lid)

1. The day before transfection, trypsinize and count cells (*APPENDIX 3F*). Plate cells in each well of a 24-well tissue culture plate using normal cell culture medium with serum, so that they are 50% to 95% confluent on the day of transfection. Avoid antibiotics at the time of plating and during transfection.

The single most important factor in high-efficiency transfection is healthy, proliferating cell cultures. Antibiotics may cause some toxicity if present during transfection.

In a 24-well plate, seeding 4×10^4 to 2×10^5 cells per well will usually give good plating density. Any type of plate may be used by scaling the reagent and cell amounts in proportion to the relative surface area (see Table 9.4.3).

2. Dilute DNA into dilution medium (appropriate for the lipid being optimized) without serum or antibiotics in four microcentrifuge tubes. Use a range of DNA concentrations, and use a volume of dilution medium that is 7× the appropriate protocol volume (see *Basic Protocol 1*, step 2, and Table 9.4.1). Mix gently after each addition.

This makes enough DNA per tube for seven wells on a 24-well plate. Good ranges include 0.2 to 1.6 μg DNA per well.

If the Plus enhancer is being used, include it in the diluted DNA tubes, using 10 μl Plus reagent per μg DNA. Add the Plus reagent to the diluted DNA after mixing well. If the Plus reagent is added first, precipitation may occur.

3. Dilute cationic lipid reagent into dilution medium without serum or antibiotics in six microcentrifuge tubes. Use a range of DNA concentrations, and use a volume of dilution medium that is 5× the appropriate protocol volume (see *Basic Protocol 1*, step 3, and Table 9.4.1). Mix gently after each addition.

Be sure to observe timing that works best for the cationic lipid reagent being used.

This makes enough diluted lipid per tube for five wells on a 24-well plate. Good ranges include 0.5 to 5 μl lipid reagent per well.

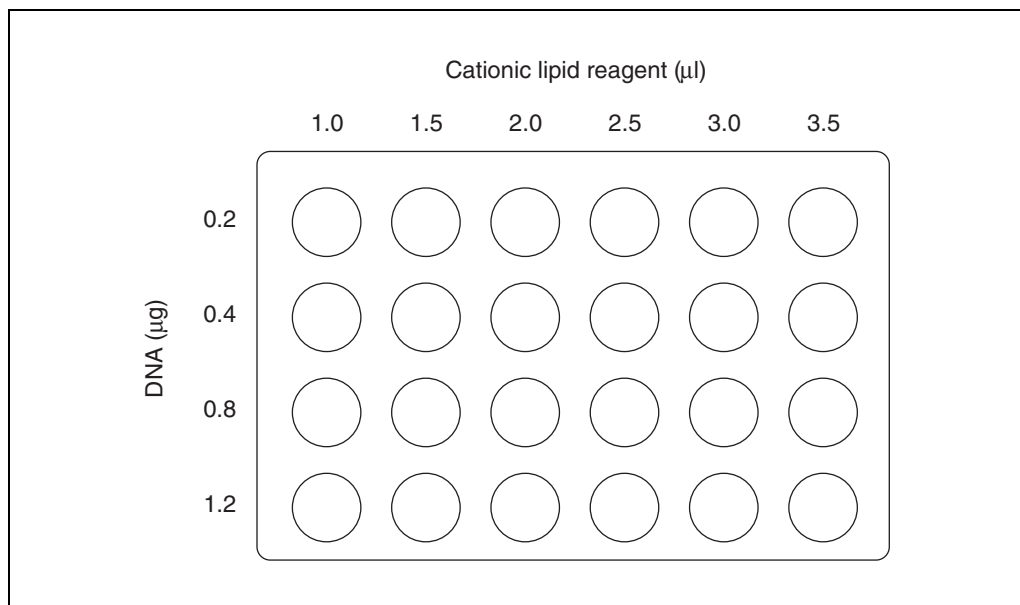


Figure 9.4.3 A sample matrix for fine tuning (optimizing) transfection reagent efficiencies using cationic lipid reagents.

4. Pipet equal per-well volumes of diluted DNA and diluted cationic lipid reagent into the wells of a 96-well plate in a matrix corresponding to the wells on the 24-well plate (Fig. 9.4.3). Mix the complexes with the pipet tip by triturating. Cover the plate and incubate for 15 min at room temperature.

Incubation times >15 min (up to an hour) work just as well, but be sure the complexes do not dry by covering them well.

5. While complexes are forming, replace medium on the cells with fresh transfection medium.

See Basic Protocol 1, steps 5 and 6, for alternate procedures for combining complexes, medium, and cells. The medium can be the same as the serum-free dilution medium. It is possible to use serum in the transfection medium at this step. It is also possible to omit this step when using LipofectAmine 2000.

6. Add aliquots of DNA-lipid complexes (total volume from wells in step 4) to each well containing cells with fresh transfection medium. Mix complexes into the medium gently, holding the plate at a slant. Incubate at 37°C in 5% CO₂ for 5 hr.

The exposure time may be >5 hr (up to overnight). Be sure there is sufficient medium to prevent the cells from drying out (it is not necessary to increase other components if this is done). If using the Plus enhancer, a 3-hr exposure is sufficient.

7. After 5 hr incubation, add cell culture medium to reach normal volume and add serum to bring the final concentration to that of normal cell culture medium.

If necessary to maximize cell growth, replace the medium containing the complexes with fresh complete medium after 5 hr incubation. This step may be omitted entirely for some protocols.

8. Check expression as described (see Basic Protocol 1, steps 8a and 8b).

If peak activity is found to occur on the edge of the matrix of concentrations tested, adjust the concentrations to include the observed peak at the center of a new matrix and repeat the transfection.

COMMENTARY

Background Information

There are currently at least eight companies that market cationic lipid-based transfection reagents. A partial listing of companies and products may be seen in Table 9.4.1. Many companies offer more than one type of reagent. Among the more popular ones are LipofectAmine 2000 and LipofectAmine Plus (Life Technologies), DOTAP and FuGENE 6 (Roche), and Effectene (Qiagen). Some of the structures are proprietary. The structures that are published can be classified into two general categories based on the number of positive charges in the lipid headgroup. The first cationic lipid (DOTMA) has a single positive charge per molecule and is used in Lipofectin (Life Technologies; Felgner et al., 1987). Several other cationic lipid-based transfection reagents such as DOTAP liposomal transfection reagent (Roche) and DMRIE-C (Life Technologies) also make use of singly charged cationic lipid mole-

cules. Increasing the number of positive charges per cationic lipid molecule to as many as five improved transfection efficiency dramatically in most cell types. This can be seen in the examples of DOGS, the cationic lipid in Transfectam (Promega; Behr et al., 1989); DOSPA in LipofectAmine (Life Technologies; Hawley-Nelson et al., 1993); and TMTPS in CellFectin (Life Technologies; Anderson et al., 1995). Further increase in transfection efficiency can sometimes be achieved by precomplexing DNA with a proprietary enhancer. Two commercially available transfection kits with enhancers are LipofectAmine Plus (Life Technologies; Shih et al., 1997) and Effectene (Qiagen).

Life Technologies has designed cationic lipid reagents with specialized applications such as high-efficiency transfection of insect cells (see Basic Protocol 4) or delivery of RNA (see Basic Protocol 3). Lipofectin has high

activity for endothelial cell transfection (Tilkens et al., 1994).

Basic Protocol 1 and the Alternate Protocol described in this unit are the procedures with the highest potential for efficient DNA transfection of adherent mammalian cells (Shih et al., 1997; Ciccarone et al., 1999). LipofectAmine 2000 has a simple protocol that yields the highest transfection efficiencies in many cell types. Using the enhancer reagent results in more reproducible transfections without extensive optimization because of the overall high activity. Prior to the availability of enhanced cationic liposome transfections (e.g., using LipofectAmine 2000 and Effectene), the most effective procedure for transfection of adherent mammalian cells with DNA was with other polycationic reagents (e.g., LipofectAmine; Hawley-Nelson et al., 1993) following Basic Protocol 1. In order to achieve high-efficiency transfections with Basic Protocol 1, it is necessary to optimize lipid and DNA concentrations with the target cells at the desired plating density using a procedure similar to that described in the Support Protocol. Many cationic lipid reagents, as well as transfection reagents based on other chemistries, are available that can be used in Basic Protocol 1 for adherent mammalian cell DNA transfection, but they may yield lower efficiencies than the Alternate Protocol with the enhancer. Optimization using the Support Protocol is highly recommended when not using the enhancer, and the protocol can be modified for use with any cationic lipid reagent.

Critical Parameters

The most critical parameter for successful transfection is cell health. Cells should be proliferating as rapidly as possible at the time they are plated for transfection. On the day of transfection, mitoses should be abundant in healthy cultures. Fresh cultures with a finite life span should be used at the earliest possible passage.

For reproducible transfection results, it is critical to plate the same number of healthy cells for each transfection. Cells should always be counted, preferably using a hemacytometer and trypan blue (APPENDIX 3F).

Although optimization is not required for high-efficiency transfection when using an enhancer (see Alternate Protocol), it is essential for success without the enhancer, and may improve efficiency even with the enhanced method.

The medium used to dilute and form complexes between the cationic liposomes and the DNA must not contain serum. Serum contains

sulfated proteoglycans and other proteins, which compete with the DNA for binding to the cationic lipids. The medium should also not contain antibiotics. There is toxicity to the cells when cationic lipid reagents are used in the presence of antibiotics.

The dilution medium/plating medium for the cells may have some influence on transfection efficiency. Some proprietary serum-free media contain components that inhibit transfection and should be replaced with Opti-MEM I, DMEM, or other media without serum during transfection (Hawley-Nelson and Ciccarone, 1996).

Serum may be present in the medium on the cells during transfection. For most cationic lipid reagents, on most cell types, transfection activity is not inhibited in the presence of serum provided the complexes were formed in serum-free medium (Brunette et al., 1992; Ciccarone et al., 1993, 1999; Shih et al., 1997).

The specific serum-free medium used to dilute the lipid and DNA can have a slight effect on the efficiency of transfection. For the enhanced protocol (Alternate Protocol), normal culture medium such as DMEM is recommended. For the standard procedure (Basic Protocol 1, 2, and 3), Opti-MEM I medium works best. The improvements resulting from using the recommended media are less than two fold. When using Lipofectin, dilution in Opti-MEM I followed by a 30- to 45-min incubation is recommended (Ciccarone and Hawley-Nelson, 1995). With LipofectAmine 2000, the reverse is true: extended incubation (>30 min) of LipofectAmine 2000 in Opti-MEM prior to addition of DNA results in lower transfection activity (Ciccarone et al., 1999).

High-purity DNA will increase transfection efficiency. Miniprep DNA does work, however, when efficiency is not critical. A wide range of sizes of polynucleotides may be transfected, from 18-mer single-stranded oligonucleotides (Chiang et al., 1991; Bennett et al., 1992; Yeoman et al., 1992; Wagner et al., 1993) to 400-kb YAC DNAs (Lamb et al., 1993). Excess vortexing of complexes or DNA solutions may result in some shearing, especially with larger molecules. The concentration of EDTA in the diluted DNA should not exceed 0.3 mM.

Transgene expression may be increased in some cell types by inducing the promoter. This is observed in Jurkat cells when phytohemagglutinin and phorbol myristate acetate are added following transfection to activate the cytomegalovirus promoter (Schifferli and Ciccarone, 1996).

Troubleshooting

The most common complaints surrounding transfections include decreased transfection efficiency and low cell yield. Decreases in efficiency often result from changes in the target cell line. Cultured cell lines are usually aneuploid and often consist of a mixture of genotypes and phenotypes that can be subject to selection in the laboratory environment. Primary cultures, although usually genotypically uniform, often consist of a mixture of phenotypes from different tissues and can change their population characteristics in response to their environment. Whenever a decrease in transfection efficiency is observed, the first

thing to try is to work with a freshly thawed culture or isolate (Hawley-Nelson and Shih, 1995). Be sure the same number of cells is plated in each experiment, since plating density affects efficiency and peak position (Hawley-Nelson et al., 1993).

Low cell yield often results from the use of too much DNA or cationic lipid reagent. Use lower concentrations of these two components and examine the results for transfection efficiency as well as cell yield. Acceptable efficiency can usually be obtained with higher cell yield by using lower concentrations of lipid and DNA (Hawley-Nelson et al., 1993; Life Technologies, 1999).

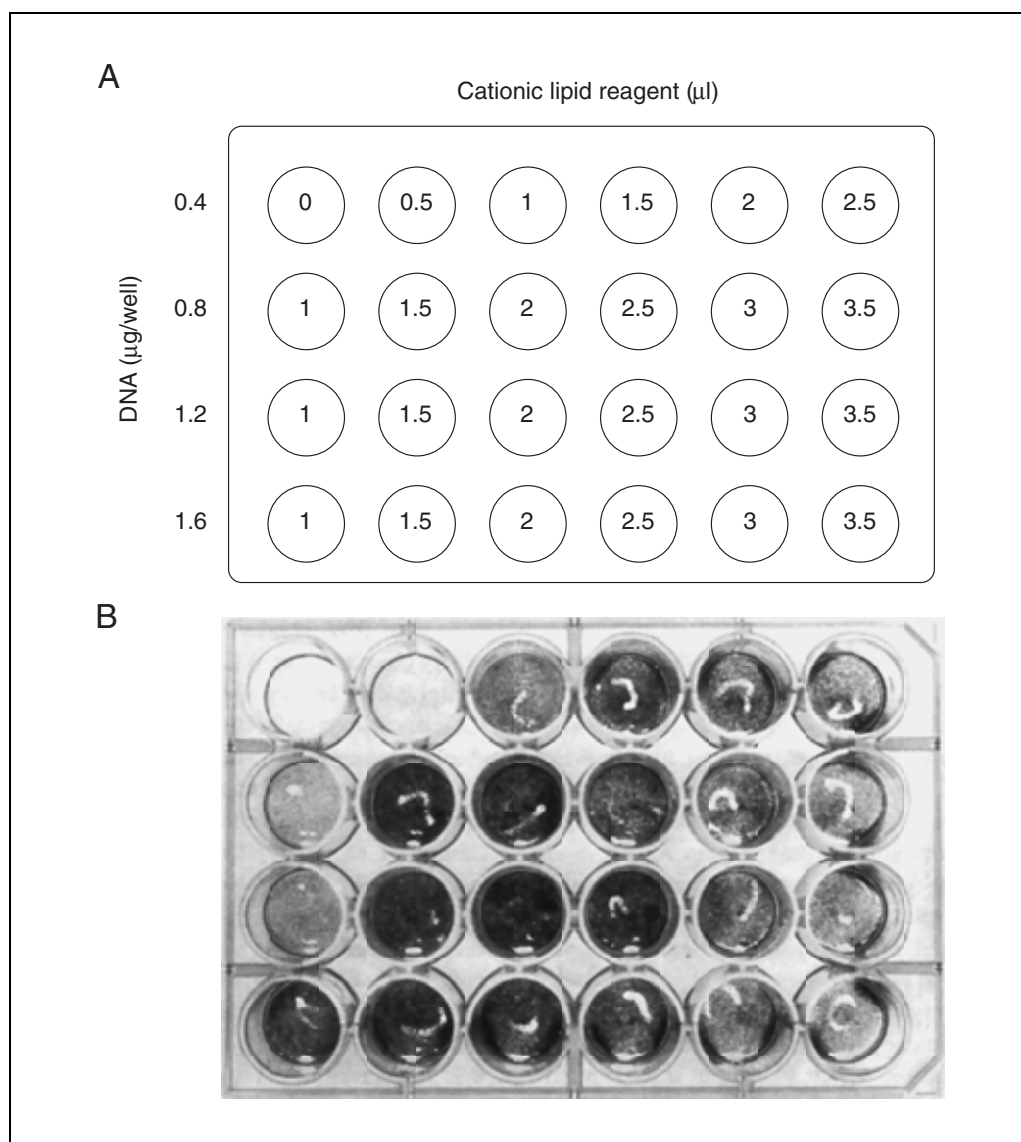


Figure 9.4.4 Results of fine-tuning or optimizing conditions for transfection. Before transfection, 293 H cells were plated at 2×10^5 /well in a 24-well plate precoated with poly-D-lysine. The following day, cells were transfected with pCMV-SPORT β gal DNA using LipofectAmine 2000 as described (see Support Protocol). One day posttransfection, cells were fixed and stained with Xgal. **(A)** Amounts of DNA and lipid reagent used. **(B)** Results of Xgal staining.

Table 9.4.5 Activity for a Scaled-up Transfection Using LipofectAmine Plus in BHK-21 Cells

Plate	Surface area ratio to 24-well plate	Cells/well ($\times 10^4$)	DNA/well (μg)	Plus reagent (μl)	LipofectAmine reagent (μl)	β -gal (ng/cm^2) ^a
24-well	1	4	0.4	2	2	188 \pm 5
12-well	2	8	0.8	4	4	193 \pm 12
6-well	5	20	2	10	10	179 \pm 27
60-mm	10	40	4	20	20	171 \pm 16
100-mm	28	112	11.2	560	56	157

^aResults are the mean of three transfections \pm the standard deviation.

Cell yield can also be improved in several other ways. (1) Increasing the plating density. This usually requires adjustment of lipid and DNA amounts, but often the transfection efficiency as well as the cell yield increases with higher plating input (Life Technologies, 1999). (2) Decreasing time of exposure of the cells to cationic lipid–DNA complexes. This can be done by increasing volume and adding back serum at earlier times or by removing the complexes from the cells at the end of transfection. (3) Performing the transfection in the presence of serum. Most cationic lipid reagents work well in transfection medium containing serum (Brunette et al., 1992; Ciccarone et al., 1993, 1999; Shih et al., 1997). One exception is LipofectAmine without the Plus enhancer.

Some cationic lipid solutions are naturally cloudy. Sometimes cloudiness is observed when complexes are made with DNA. Usually this does not interfere with transfection efficiency. Most cationic liposome solutions (especially DMRIE-C and CellFectin) should be mixed gently by inversion just before use to produce a uniform suspension. With Plus reagent, it is possible to precipitate the DNA when the Plus reagent is diluted first into the DMEM and DNA is added second. Always dilute the DNA into DMEM and mix well before adding Plus reagent.

Anticipated Results

Transfection should be observed for most adherent mammalian cells transfected with DNA using Basic Protocol 1. Efficiencies vary with cell type. For example, 293, COS-7, and CHO-K1 cells yield 95% or more blue cells following Xgal staining of cells transfected with pCMVSPORT β gal plasmid DNA using LipofectAmine 2000. The authors have noted efficiencies of other cell types as high as 49% for SK-BR3 breast cancer cell lines, 77% for BE(2)C human neuroblastoma cells, and 43% for MDCK canine kidney cells (Ciccarone et

al., 1999). Efficiencies also vary for suspension cells. The authors note that while LipofectAmine Plus is relatively inefficient for transfecting Jurkat cells, DMRIE-C can yield up to 85% blue cells following pCMV·SPORT β gal plasmid DNA transfection, gene activation with phytohemagglutinin and phorbol myristate acetate, and Xgal staining (Ciccarone et al., 1995, Schifferli and Ciccarone, 1996).

The result of a typical fine-tuning/optimization protocol is shown in Figure 9.4.4. The transfection reagent was LipofectAmine 2000, the DNA was pCMV·SPORT β gal, the cells were 293 H. Cells were stained the day following transfection and were allowed to stain overnight at 37°C. A selection of transfection conditions can be made.

Conditions found to be advantageous for transfection in small wells may be scaled up, the results of a typical scale-up are given in Table 9.4.5.

The results of a stable transfection of NIH 3T3 cells with pSV2*neo* DNA using an enhancer reagent (LipofectAmine Plus) are shown in Figure 9.4.5. The transfection was done on 24-well plates, the cells were passaged onto 6-well plates the following day, and selection with geneticin was done for 10 days. The figure shows optimization of conditions and the generally high efficiency that can be achieved using this method.

Time Considerations

Counting and plating the cells should be done the day before transfection and will usually require <1 hour. Transfection is usually done in the morning. Depending on the number of conditions being evaluated, it may require all morning plus a short period in the afternoon to increase medium volume or feed the cells. The total time required is <1 day.

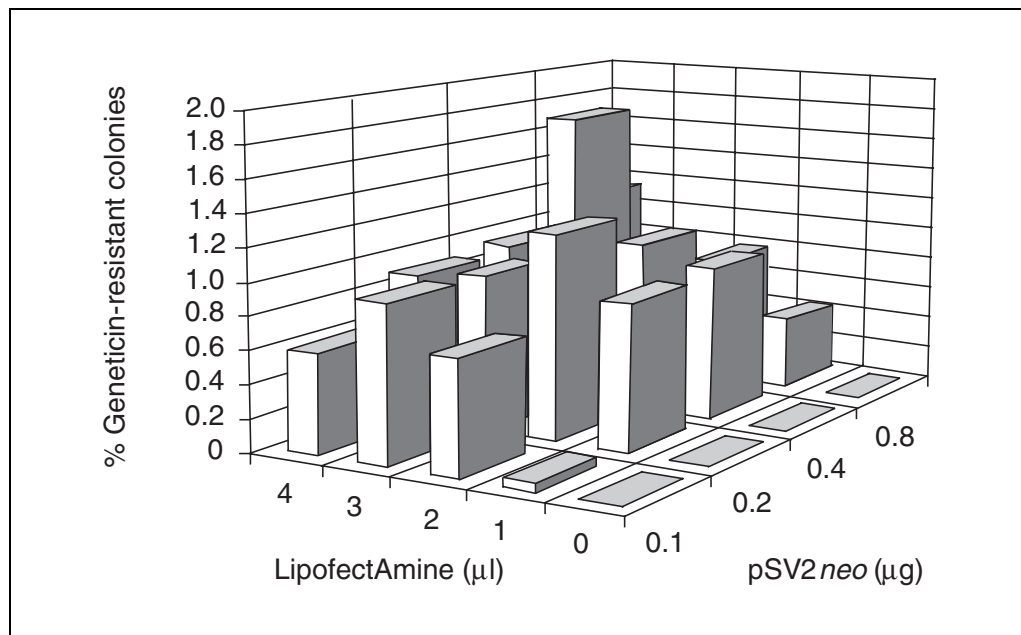


Figure 9.4.5 Stable transfection of NIH 3T3 cells. Cells were plated at 6×10^4 cells/well in 24-well plates. The day after plating, cells were transfected with LipofectAmine Plus complexed with pSV2neo DNA. The following day, cells were passaged at a total dilution of 1/150. Cells were exposed to 0.6 mg/ml Geneticin antibiotic from day 3 to day 13, and were then washed once with PBS and stained with 0.2% toluidine blue in PBS with 10% formalin.

Literature Cited

- Anderson, D., Harris, R., Polayes, D., Ciccarone, V., Donahue, R., Gerard, G., Jessee, J., and Luckow, V. 1995. Rapid generation of recombinant baculovirus and expression of foreign genes using the Bac-to-Bac baculovirus expression system. *Focus* 17.2:53-58.
- Behr, J.-P., Demeneix, B., Loeffler, J.-P., and Perez-Nutul, J. 1989. Efficient gene transfer into mammalian primary endocrine cells with lipopolyamine-coated DNA. *Proc. Natl. Acad. Sci. U.S.A.* 86:6982-6986.
- Bennett, C.F., Chiang, M.-Y., Chan, H., Shoemaker, J.E.E., and Mirabelli, K. 1992. Cationic lipids enhance cellular uptake and activity of phosphorothioate antisense oligonucleotides. *Mol. Pharmacol.* 41:1023.
- Brunette, E., Stribling, R., and Debs, R. 1992. Lipofection does not require the removal of serum. *Nucl. Acids Res.* 20:1151.
- Chiang, M.-Y., Chan, H., Zounes, M.A., Freier, S.M., Lima, W.F., and Bennett, C.F. 1991. Antisense oligonucleotides inhibit intercellular adhesion molecule 1 expression by two distinct mechanisms. *J. Biol. Chem.* 266:18162.
- Ciccarone, C. and Hawley-Nelson, P. 1995. Lipofectin transfection activity increased by protocol improvement. *Focus* 17:103.
- Ciccarone, V., Hawley-Nelson, P., and Jessee, J. 1993. Cationic liposome-mediated transfection: Effect of serum on expression and efficiency. *Focus* 15:80-83.
- Ciccarone, V., Anderson, D., Lan, J., Schifferli, K., and Jessee, J. 1995. DMRIE-C reagent for transfection of suspension cells and for RNA transfections. *Focus* 17.3:84-87.
- Ciccarone, V., Chu, Y., Schifferli, K., Pichet, J.-P., Hawley-Nelson, P., Evans, K., Roy, L., and Bennett, S. 1999. LipofectAmine 2000 Reagent for rapid, efficient transfection of eukaryotic cells. *Focus* 21.2:54-55.
- Dube, S. 1997. Transfection using LipofectAmine Plus Reagent. *Focus* 19.3:57.
- Felgner, P.L., Gadek, T.R., Holm, M., Roman, R., Chan, H.W., Wenz, M., Northrop, J.P., Ringold, G.M., and Danielsen, M. 1987. Lipofection: A highly efficient, lipid-mediated DNA-transfection procedure. *Proc. Natl. Acad. Sci. U.S.A.* 84:7413-7417.
- Goldsborough, M.D., Evans, K., Xu, L., and Young, A. 1998. High purity plasmid DNA from anion exchange chromatography. *Focus* 20:68-69.
- Hawley-Nelson, P. and Shih, P.-J. 1995. Sensitivity of transfection efficiency to culture age. *Focus* 17:62.
- Hawley-Nelson, P. and Ciccarone, V. 1996. Transient transfection efficiency of human keratinocytes in two serum-free media. *Focus* 18.2:43-44.
- Hawley-Nelson, P., Ciccarone, V., Gebeyehu, G., Jessee, J. and Felgner, P. 1993. LipofectAmine reagent: A new, higher efficiency polycationic liposome transfection reagent. *Focus* 15.3:73-79.

Lamb, B.T., Sisodia, W.W., Lawler, A.M., Slunt, H.H., Kitt, C.A., Kearns, W.G., Pearson, P.L., Price, D.L., and Gearhart, J.D. 1993. Introduction of the 400 kilobase *precursor amyloid protein* gene in transgenic mice. *Nature Genet.* 5:22-30.

Life Technologies. 1999. Guide to eukaryotic transfections with cationic lipid reagents, 2nd ed. Life Technologies, Inc., Rockville, Md.

Rose, J.K., Buonocore, L., and Whitt, M. 1991. A new cationic liposome reagent mediating nearly quantitative transfection of animal cells. *BioTechniques* 10:520-525.

Schifferli, K.P. and Ciccarone, V. 1996. Optimization of cationic lipid reagent-mediated transfection for suspension cell lines. *Focus* 18:45-47.

Sells, M.A., Li, J., and Chernoff, J. 1995. Delivery of proteins into cells using cationic liposomes. *BioTechniques* 19:72-78.

Shih, P.J., Evans, K., Schifferli, K., Ciccarone, V., Lichaa, F., Masoud, M., Lan, J., and Hawley-Nelson, P. 1997. High efficiency transfection with minimal optimization using the Lipofect-Amine Plus Reagent. *Focus* 19.3:52-56.

Tilkins, M.L., Hawley-Nelson, P., and Battista, P. 1994. Transient transfection of endothelial cells. *Focus* 16.4:117-119.

Yeoman, L.C., Danels, Y.J., and Lynch, M.J. 1992. Lipofectin enhances cellular uptake of antisense DNA while inhibiting tumor cell growth. *Antisense Res. Dev.* 2:51.

Wagner, R.W., Matteucci, M.D., Lewis, J.G., Gutierrez, A.J., Moulds, C., and Froehler, B.C. 1993. Antisense gene inhibition by oligonucleotides containing C-5 propyne pyrimidines. *Science* 260:1510.

Key References

Kriegler, M. 1990. Gene Transfer and Expression: A Laboratory Manual. Stockton Press, New York.

Good general review on gene expression.

Life Technologies, 1999. See above.

Describes history and gives protocols and considerations for cationic lipid reagent transfections.

Tilkins, M.L., Hawley-Nelson, P., and Ciccarone, V. 1998. Transfection of mammalian and invertebrate cells using cationic lipids. *In Cell Biology: A Laboratory Handbook*, Vol. 4, 2nd ed. (J.E. Celis, ed.) pp. 145-154. Academic Press, New York.

Review of cationic lipid transfection procedures.

Felgner et al., 1987. See above.

The original description of cationic lipid transfection.

Contributed by Pamela Hawley-Nelson
and Valentina Ciccarone
Life Technologies, Inc.
Rockville, Maryland

Selection of Transfected Mammalian Cells

Analysis of gene function frequently requires the formation of mammalian cell lines that contain the studied gene in a stably integrated form (Chapter 16). Approximately one in 10^4 cells in a transfection will stably integrate DNA (the efficiency can vary depending on the cell type). Therefore, a dominant, selectable marker is used to permit isolation of stable transfectants. In the first part of this unit, the procedure for determining selection conditions and the resulting stable transfection is presented (see Basic Protocol 1) and the most commonly used selectable markers are discussed (see Basic Protocol 2). Basic Protocol 2 includes conditions for thirteen markers commonly used for selection of mammalian cells.

In order to study the function of a particular gene, it is often necessary to work with a homogeneous population of cells overexpressing either wild-type or mutant variants of the corresponding protein. Two points must be considered, however, when designing such experiments. First, if the gene product of interest has a negative or toxic effect on cell growth, selection of cells stably expressing that gene cannot be achieved using a dominant selection marker. Cells carrying an antibiotic-resistance marker lose their growth advantage when expressing a toxic gene. Second, isolation of a homogeneous pool of cells stably expressing a gene by antibiotic selection often takes a month or more. By the time a stably transfected pool of cells is isolated, many important, temporally dependent biochemical events may have already occurred. One solution to both of these problems is to select transfected cells from the total population soon after transfection (see Basic Protocol 3). To do so, cells are cotransfected with a plasmid (pHook-1 from Invitrogen) that expresses a membrane-anchored selection tag, along with a second plasmid expressing the gene of interest. Alternatively, cells can be transfected with pHook-2 or pHook-3, plasmids that express both the gene of interest and the selection tag. To isolate a homogeneous pool of transfected cells, the total cell population is harvested and incubated with magnetic beads that bind to the selection tag displayed on the transfected cells. After exposure to a magnet, bead-bound transfected cells are centrifuged down into a pellet while unbound cells remain in the supernatant and can be discarded. Optimization of transfection conditions can be facilitated by a simple staining assay (see Support Protocol).

STRATEGIC PLANNING

There are several aspects of the vector to be considered for use in stable transfection. The following parameters should be considered when designing vectors to express either the marker gene or a gene of interest that is introduced with the marker gene.

Construct Design

A typical construct is depicted in Figure 9.5.1.

Promoters

The choice of promoter can be critical for efficient expression, as it is the promoter and accompanying enhancers that drive the expression of the coding regions of the gene of interest and the selectable marker. Promoters are usually short DNA sequences (<1 kb) that bind endogenous transcription factors. They may or may not have a TATA box. Cytomegalovirus and simian virus 40 (SV40) promoters have been favorite general mammalian promoters for years, and many derivatives have been made and are commercially available. More recently, other promoters such as pMC1 and PGK (promoter from phosphoglycerate kinase) have been popular because of the high levels of expression obtainable. It is also possible to use a promoter that is specific for the cell type.

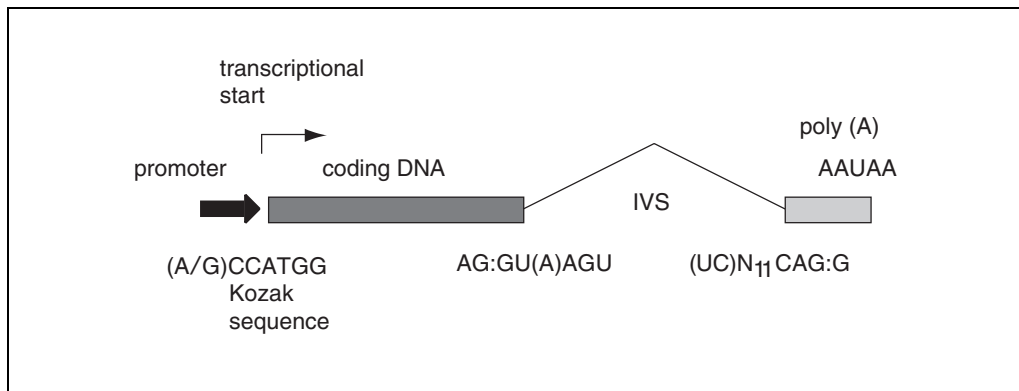


Figure 9.5.1 Generalized construct for expression of genes in mammalian cells. Each integrated construct must contain a promoter region with a transcriptional start site and a coding sequence. Features that increase expression include translational start sites according to Kozak's rules, intervening sequences with splice donor and acceptor sites, and polyadenylation signals. Abbreviation: IVS; intervening sequence.

Conditional-expression promoters have also been introduced recently, increasing the utility and flexibility of stably transfected lines. These promoters require the introduction of a specific transactivator, which is sensitive to an administered compound. Two transcription factors have been developed based on the tet operon which use tetracycline to turn expression off (tet-off) or on (tet-on). For details of tetracycline-regulated systems, see *UNIT 16.14*, Gossen and Bujard (1992), and Gossen et al. (1995). In addition, the insect hormone ecdysone has been used to activate expression based on the ecdysone receptor and responsive element (No et al., 1996).

Translational start site

The translational start site can dramatically affect expression and should be considered both in designing the expression construct for the gene of interest and in choosing the specific version of a selectable marker. Most selectable markers are based on a bacterial resistance gene, and in their native state may give low levels of expression. One major reason for this is the differences in translational start sites. Whereas bacterial systems translate polycistronic messages well, translation of most mammalian mRNA is greatly decreased (~10-fold) by upstream ATG start sites. In addition, the mammalian start site is most efficient if it follows "Kozak's rules" (Kozak, 1989). Many commonly used selectable markers have been modified to give good mammalian translation.

Polyadenylation site

A sequence in the 3' end of the expression vector that signals the addition of a poly(A) tail often stabilizes the mRNA and leads to better expression. The polyadenylation signal consists of a conserved sequence (AAUAA) 11 to 30 nucleotides upstream of the polyadenylation site, and a GU- or U-rich region downstream. Many different sources of polyadenylation signals are used. A common one is derived from the SV40 virus, but constructs using the PGK promoter often use the phosphoglycerate kinase polyadenylation signal as well.

Intervening sequence and splice sites

Addition of an intervening sequence (intron) that is spliced out in the final mRNA can also increase expression. A 5' and 3' splice site flanking the intron are necessary for splicing out of the intron. The mechanism by which addition of an intron might increase expression is not well understood.

Introduction of DNA into Cells

For the creation of stable lines, the results can vary depending on the method of DNA introduction used. Calcium phosphate (*UNIT 9.1*) and dextran (*UNIT 9.2*), as well as liposome-mediated and positively charged liposome-mediated (*UNIT 9.4*) transfection methods, all tend to introduce multiple copies of transfected genes, usually at a single site. These methods aggregate the DNA and are suitable for cotransfection of separate plasmids, typically in the ratios of 5 to 10 parts DNA of interest to one part DNA of selectable marker. These multiple copies can integrate as head-to-head, head-to-tail, or a combination of the two orientations. Because of the tandem array of similar sequences, intrachromosomal rearrangement can occur, leading to loss of copies and a decrease in the level of expression.

Electroporation (*UNIT 9.3*) will usually introduce DNA in a single copy in a single site, depending on the DNA concentration (i.e., high DNA concentrations will introduce multiple copies at a single site in ~25% of cells, although 75% will still contain single copies). If a single copy is desired (as for removal of selection markers as described below) electroporation of colinear constructs is required.

Separate plasmids for the gene of interest and the selectable marker can be used if multiple copies of genes are desired. Both aggregation methods and electroporation will yield resistant colonies with multiple copies of the gene. If only a single copy is desired, then electroporation should be used with a single plasmid that contains both the gene of interest and the selectable marker.

Removal of Selectable Markers

The selectable marker can be removed using recombinases (Fig. 9.5.2). This is desirable if more than one sequential genetic manipulation is planned. Because tandem repeats would be eliminated by the recombinase expression, single-copy integration is desirable; therefore the preferred method of transfection is electroporation. Removal of the selectable marker is accomplished by transiently expressing the Cre recombinase, which removes the sequence between the identical recognition sequences (*lox* sites).

Amplification of Expressed Gene

Some selectable markers (most notably DHFR) can be used to amplify the expression of the integrated DNA. Amplification is accomplished by gradually increasing the selection pressure with higher concentrations of selection medium. This results in tandem duplication of the integrated DNA (see *UNIT 23.5*).

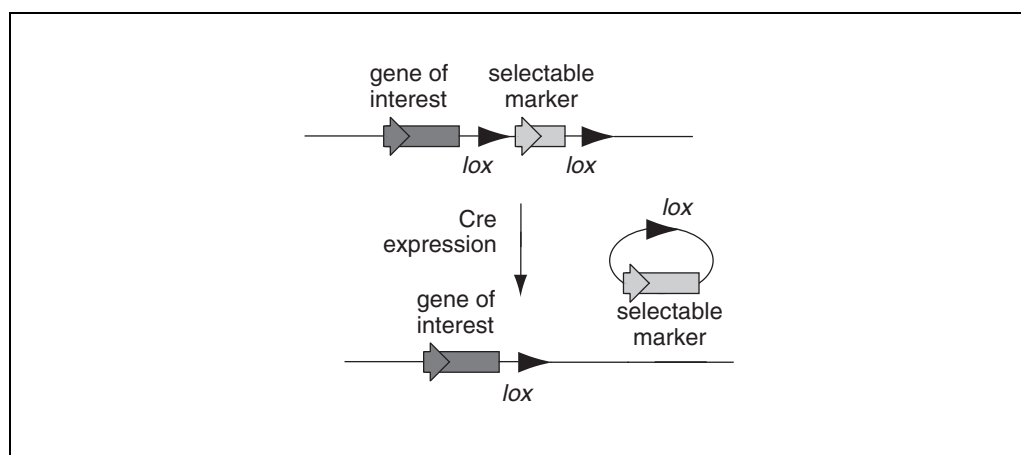


Figure 9.5.2 Using the Cre-*lox* system to remove a selectable marker after integration at a unique site. Transient expression of Cre recombinase causes the excision of sequences between the *lox* sites.

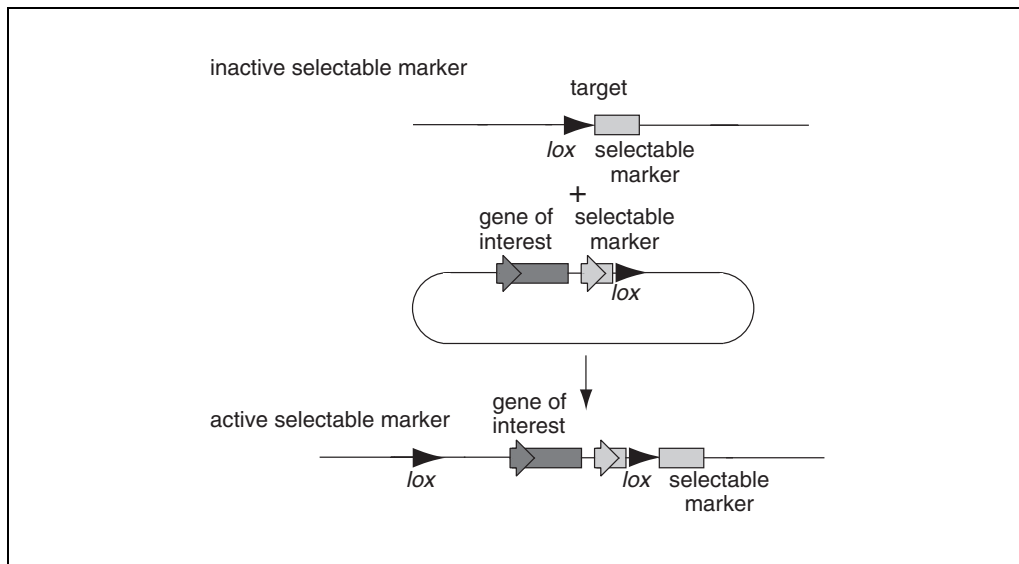


Figure 9.5.3 Using the Cre-lox system to integrate an expression construct at a specific reproducible site. The target site is a previous integrated defective form of a selectable marker (lacking a promoter and ATG start codon). Transfection with a plasmid containing the gene of interest leads to integration at the *lox* site and adds a promoter and translational start site, restoring expression of the selectable marker.

Introduction of DNA at Specific Sites

One of the problems with production of stable lines is that the transfected DNA integrates at a random site. Many of the integrations will not give adequate expression. To address this problem, two different approaches can be taken.

It is possible to use homologous recombination to introduce DNA at a specific site. While this approach has largely been applied to transgenic animals, the same principles can be applied to a cell line. It is possible then to create a cell line where the gene of interest is expressed by an endogenous promoter in its normal genomic location.

Sauer and coworkers developed the Cre recombinase system to introduce DNA at a selected site, shown in Figure 9.5.3 (Fukushige and Sauer, 1992). Initially an inactive modified *neo* is integrated into the DNA which contains an in-frame *loxP* site at the 5' end of *neo* without an ATG initiation codon or a promoter. This cell line can then be transfected with a plasmid containing the gene of interest and a promoter in front of a modified *lox* site. When electroporated into the previously established cell line with a Cre expression vector, recombination is targeted to the *loxP* site in the genome and can be selected in G418 because of the reconstituted *neo* activity. This constant-site integration gives consistent expression of the gene of interest in different clones.

STABLE TRANSFER OF GENES INTO MAMMALIAN CELLS

In this protocol, appropriate selection conditions for the parental cell line are first determined. The gene to be studied is then cotransfected into that cell line with a gene that expresses a selectable marker. The cells are allowed to grow under selection for approximately ten doublings before individual colonies are picked and expanded into cell lines.

Materials

- Complete medium (APPENDIX 3F)
- Selective medium (see Basic Protocol 2)
- Cloning cylinders (UNIT 16.23 and Fig. 16.23.2)

Additional reagents and equipment for mammalian cell culture and counting cells (APPENDIX 3F) and transfection (UNITS 9.1, 9.3 & 9.4)

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and proper sterile technique should be followed accordingly.

NOTE: All culture incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

1. Ensure that the cell line to be transfected can grow as an isolated colony. For adherent cells, plate ~100 cells on a 10-cm tissue culture dish and feed every 4 days for 10 to 12 days. Count the number of viable colonies (see APPENDIX 3F for basic cell culture techniques).

Selection of stably transfected colonies requires that cells be grown essentially in isolation. A cell line that does not grow as an isolated cell will not be able to be stably transfected. The simplest way to determine how many colonies are present on a tissue culture dish is to stain with methylene blue. Aspirate off the medium, then place ~2 ml of 2% methylene blue (made up in 50% ethanol) on the dish. Wait 2 min, pour off methylene blue, and wash residual dye off in a bucket of cold water.

Cell lines that survive poorly as single cells may be selected by using a feeder layer—i.e., a monolayer of cells treated to prevent cell division (Robertson, 1987).

2. Determine the selection conditions for the parental cell line. For example, if hygromycin-B selection is to be used, determine the minimum level of hygromycin-B that must be added to the medium to prevent cell growth. Split a confluent dish of cells 1:15 into medium containing various levels of the drug. Incubate the cells for 10 days, feeding with the appropriate selective medium every 4 days, and examine the dishes for viable cells.

Mammalian cells can divide once or twice under selective conditions that will eventually kill them. For example, DHFR-deficient cells will divide in selective medium until they deplete their endogenous reserve of nucleotides. For this reason, it is essential to split cells at least 1:15 when testing selection conditions, or the cells may reach confluence before the selection can take effect. Several selection markers commonly used in mammalian cell culture are described below. The choice of a marker is determined both by the cell type to be transfected and the reason for doing the transfection (see Basic Protocol 2).

3. Determine the most efficient means of transfecting the parental cell type (introduction to Section I).

Calcium phosphate- and liposome-mediated transfection (UNITS 9.1 & 9.4) and electroporation (UNIT 9.3) are the methods of choice for producing stable transfectants, as both techniques introduce a large amount of DNA into cells that pick up DNA. The presence of a large amount of DNA increases the probability that some of it will become stably integrated. Transfection mediated by DEAE-dextran does not work well for production of stable cell lines.

4. If calcium phosphate transfection is used, split the parental cell line 1:15 into complete medium the day before the DNA is applied to the cell. If electroporation is used, refer to UNIT 9.3, Basic Protocol, step 7. If liposome-mediated transfection is used, see UNIT 9.4.
5. Cotransfect desired gene and marker gene into parental cell line. Use a molar ratio 5:1 of the plasmid containing the gene of interest to the plasmid containing the selective marker. Include a control where carrier DNA (e.g., pUC13) is used instead of the plasmid containing the gene to be studied. Do several separate transfections containing the gene of interest in case one of the transfections fails.

It is not necessary to physically link the selective marker to the gene of interest prior to transfection if a 5:1 ratio is used. If selection requires transfection of a large amount of the selection plasmid, then a 5:1 ratio will not be achievable in some cases, and a single plasmid should be constructed that contains both the selective marker and the gene of interest. If no colonies result from the transfection containing the gene of interest, but colonies do appear in the control containing pUC13, then the gene of interest may be cytotoxic.

Efficient transfection by electroporation requires linearization of the construct.

6. Allow the cells to double twice under nonselective conditions. Split the cells 1:15 into selective medium.
7. Place at least five dishes into selective medium from each transfected dish to maximize the number of colonies that can be picked and expanded into cell lines. Feed the cells with selective medium every 4 days (or as needed). After 10 to 12 days, inspect the plates for colonies by holding the plates up to a light at an angle.

One of the plates can be stained (see annotation to step 1) to facilitate counting the colonies. Note that staining kills the cells.

8. Pick large, healthy colonies.

Colonies should contain ~500 to 1000 cells when they are picked.

In some cases, it may be wise to continue selection after picking. Some large colonies can "protect" nonresistant cells.

BASIC PROTOCOL 2

SELECTABLE MARKERS FOR MAMMALIAN CELLS

For many of the selection conditions outlined below, the desired concentration of one or more of the drugs used for selection will vary depending on susceptibility of a particular cell line to the drug. For these drugs, approximate ranges needed for selection are given. Precise ranges for the parental cell line to be used can be determined by characterizing cell viability across a range of concentrations of the drug (see Basic Protocol 1, step 2).

Positive Selection

Adenosine deaminase (ADA)

Selection conditions. Medium supplemented with 10 µg/ml thymidine, 15 µg/ml hypoxanthine, 4 µM 9-β-D-xylofuranosyl adenine (Xyl-A), and 0.01 to 0.3 µM 2'-deoxycoformycin (dCF). Fetal bovine serum (FBS) contains low levels of ADA which will detoxify the medium, so serum should be added immediately prior to use.

Basis for selection. Xyl-A can be converted to Xyl-ATP and incorporated into nucleic acids, resulting in cell death. Xyl-A is detoxified to its inosine derivative by ADA. dCF is a transition state analog inhibitor of ADA, and is needed to inactivate ADA endogenous to the parental cell type. As the level of endogenous ADA varies with cell type, the appropriate concentration of dCF for selection will vary as well.

Comments. High levels of expression of the transfected ADA gene will be necessary to achieve selection in cells with high endogenous ADA levels. ADA-deficient CHO cells are available. ADA can be used in amplification systems (Kaufman et al., 1986), as increasing the level of dCF selects for cells that have amplified the ADA gene.

Reference. Kaufman et al. (1986).

Aminoglycoside phosphotransferase (neo, G418, APH)

Selection conditions. 100 to 800 µg/ml G418 in complete medium. G418 should be prepared in a highly buffered solution (e.g., 100 mM HEPES, pH 7.3) so that addition of the drug does not alter the pH of the medium.

Basis for selection. G418 blocks protein synthesis in mammalian cells by interfering with ribosomal function. It is an aminoglycoside, similar in structure to neomycin, gentamycin, and kanamycin. Expression of the bacterial *APH* gene (derived from *tn5*) in mammalian cells therefore results in detoxification of G418.

Comments. Varying concentrations of G418 should be tested, as cells differ in their susceptibility to killing by G418. Different lots of G418 can have different potencies, causing many investigators to buy a large amount of one lot to standardize selection conditions. Cells will divide once or twice in the presence of lethal doses of G418, so the effects of the drug take several days to become apparent.

Reference. Southern and Berg (1982).

Bleomycin (phleo, bleo, zeocin)

Selection conditions. Complete medium supplemented with 0.1 to 50 µg/ml phleomycin.

Basis for selection. This gene encodes a 13,000-Da protein that stoichiometrically binds the drug (bleomycin, phleomycin, or zeocin) and inactivates it.

Comments. Three homologous genes that confer bleomycin resistance have been isolated from the gram-positive bacterial plasmid pUB110, from the central region of the gram-negative transposon Tn5, and from resistant strains of *Actinomyces*.

References. Mulsant et al. (1988); Sugiyama et al. (1994).

Cytosine deaminase (CDA, CD)

Selection conditions. Medium containing 1 mM *N*-(phosphonacetyl)-L-aspartate, 1 mg/ml inosine, and 1 mM cytosine.

Basis for selection. Cytosine deaminase converts cytosine to uracil. *N*-(Phosphonacetyl)-L-aspartate blocks the de novo synthesis pathway of pyrimidines and forces the cells to rely on a cytosine deaminase scavenger pathway.

Comment. Originally defined as a negative selectable marker.

Reference. Wei and Huber (1996).

Dihydrofolate reductase (DHFR)

Selection conditions. α - medium supplemented with 0.01 to 300 µM methotrexate (MTX) and dialyzed fetal bovine serum (FBS).

Basis for selection. DHFR is necessary for purine biosynthesis. In the absence of exogenous purines this enzyme is required for growth. Dialysis of serum to remove endogenous nucleosides and use of media devoid of nucleosides is therefore necessary for selection. MTX is a potent competitive inhibitor of DHFR, so increasing MTX concentration selects for cells that express increased levels of DHFR.

Comments. Extremely high levels of expression of the transfected normal *DHFR* gene are needed for selection in cell lines with high endogenous DHFR levels. A mutant *DHFR* gene is available that encodes an enzyme resistant to MTX (Simonsen and Levinson, 1983). This gene can be used for dominant selection in most cell types. DHFR can be used to amplify transfected genes. This is most efficiently accomplished using a DHFR-deficient CHO cell line and a normal *DHFR* gene for selection.

Reference. Simonsen and Levinson (1983).

Histidinol dehydrogenase (hisD)

Selection conditions. Complete medium supplemented with 2.5 mM histidinol or medium lacking histidine and containing 0.125 mM histidinol.

Basis for selection. This gene encodes an enzyme that catalyzes the oxidation of L-histidinol to L-histidine. Thus, medium deficient in histidine but containing histidinol will not support cell growth. In addition, histidinol is toxic in the presence of histidine because of inhibition of histidyl-tRNA synthase, so *hisD*-containing cells can be selected by their ability to detoxify histidinol.

Comment. Histidinol in complete medium usually works well.

Reference. Hartman and Mulligan (1988).

Hygromycin-B-phosphotransferase (HPH)

Selection conditions. Complete medium supplemented with 10 to 400 µg/ml hygromycin-B.

Basis for selection. Hygromycin-B is an aminocyclitol that inhibits protein synthesis by disrupting translocation and promoting mistranslation. The *HPH* gene (isolated from *E. coli* plasmid pJR225; Gritz and Davies, 1983) detoxifies hygromycin-B by phosphorylation.

Comments. While the level of hygromycin-B needed for selection can vary from 10 to 400 µg/ml, many cell lines require 200 µg/ml.

References. Gritz and Davies (1983); Palmer et al. (1987).

Puromycin-N-acetyl transferase (PAC, puro)

Selection conditions. Complete medium supplemented with 0.5 to 10 µg/ml puromycin.

Basis for selection. Puromycin inhibits protein synthesis. The gene encodes an enzyme for the inactivation of puromycin by acetylation.

Comment. This selectable marker is comparable to *neo* in selection efficiency. Many cell types select well at 2.0 µg/ml.

Reference. de la Luna et al. (1988).

Thymidine kinase (TK)

Selection conditions. Forward (TK⁻ to TK⁺): Complete medium supplemented with 100 µM hypoxanthine, 0.4 µM aminopterin, 16 µM thymidine, and 3 µM glycine (HAT medium).

Basis for selection. Under normal growth conditions, cells do not need thymidine kinase, as the usual means for synthesizing dTTP is through dCDP. Selection of TK⁺ cells in HAT medium is primarily due to the presence of aminopterin, which blocks the formation of dTDP from dCDP. Cells therefore need to synthesize dTTP from thymidine, a pathway that requires TK.

Comments. Thymidine kinase is widely used in mammalian cell culture because both forward and reverse selection conditions exist. Unlike markers such as ADA and DHFR, however, it is not possible to select for variable levels of TK, so the gene cannot be used for amplification. Like ADA and DHFR, most mammalian cell lines express TK, removing the possibility of using the marker in those lines unless BUdr is used to select a TK⁻ mutant.

Reference. Littlefield (1964).

Xanthine-guanine phosphoribosyltransferase (XGPRT, gpt)

Selection conditions. Medium containing dialyzed fetal bovine serum, 250 µg/ml xanthine, 15 µg/ml hypoxanthine, 10 µg/ml thymidine, 2 µg/ml aminopterin, 25 µg/ml mycophenolic acid, and 150 µg/ml L-glutamine.

Basis for selection. Aminopterin and mycophenolic acid both block the de novo pathway for synthesis of GMP. Expression of XGPRT allows cells to produce GMP from xanthine, allowing growth on medium that contains xanthine but not guanine. It is therefore necessary for selection to use dialyzed fetal calf serum and a medium that does not contain guanine.

Comments. XGPRT is a bacterial enzyme that does not have a mammalian homolog, allowing XGPRT to function as a dominant selectable marker in mammalian cells. The amount of mycophenolic acid necessary for selection varies with cell type and can be determined by titration in the absence and presence of guanine.

Reference. Mulligan and Berg (1981).

Negative Selection

Cytosine deaminase (CDA, CD)

Selection conditions. Complete medium supplemented with 50 to 250 µg/ml 5-fluorocytosine.

Basis for selection. Cytosine deaminase converts the 5-fluorocytosine to 5-fluorouracil, resulting in inhibition of proliferation.

Comment. If the CD-positive cells are a minority (<1%), there is no detectable bystander killing; however, if the majority of cells contain CD then CD-negative cells survive only at higher dilutions (≤ 10,000 plated cells per 100-mm plate).

Reference. Mullen et al. (1992).

Diphtheria toxin (DT)

Selection conditions. The expression of the diphtheria toxin gene is itself toxic.

Basis for selection. Diphtheria toxin inhibits protein synthesis by the NAD-dependent ADP-ribosylation of elongation factor 2.

Comment. Since this selection is not conditional, this marker is only useful for elimination of cells expressing DT. Therefore, stable cell lines expressing DT can not be isolated. It has been used as a substitution for HSV-TK for homologous recombination or for the elimination of tissues in transgenic animals.

Reference. Yagi et al. (1990).

Herpes simplex virus thymidine kinase (HSV-TK), + to –

Selection conditions. Complete medium supplemented with 2 µM ganciclovir {9-[(1,3-dihydroxy-2-propoxy)methyl]guanine; DHPG} or 0.2 µM FIAU [1-(2'-deoxy-2'-fluoro-1-β-D-arabinofuranosyl-5-iodo)uracil].

Basis for selection. The selection drugs (ganciclovir or FIAU) are nucleoside analogs and are phosphorylated by the HSV-TK gene product. This phosphorylation leads to incorporation of the drug into the growing DNA chain during S phase, and subsequent cell death. Although the related compound 9-[(2-hydroxyethoxy)methyl]guanine (acyclovir) causes chain termination, ganciclovir does not. These drugs are ~1000-fold poorer substrates for mammalian TK and are thus not harmful to non-HSV-TK containing cells at these concentrations.

Comment. This system has been widely used for increasing the selection of homologous recombinants by including the *HSV-TK* gene outside the regions of homology (see *UNIT 23.1*). It can be used for selection of recombination after Cre recombinase expression if the *TK* is placed between the *lox* sites.

References. Cheng et al. (1983); Staschke et al. (1994).

RAPID SELECTION OF TRANSFECTED MAMMALIAN CELLS

This protocol is for mammalian cells that have been transfected in one of two ways. Either they have been cotransfected with pHook-1 and a preexisting expression plasmid, or transfected with pHook-2 (or pHook-3) with the gene of interest subcloned into the multicloning site. Cells are incubated to allow for expression of the gene of interest and the selection tag, then they are harvested. Transfected cells are subsequently selected from culture by incubation with Capture-Tec magnetic beads (Invitrogen).

Materials

Gene of interest
Cell line for transfection and appropriate complete medium (*APPENDIX 3F*)
Capture-Tec system (Invitrogen) consisting of pHook-1, pHook-2, or pHook-3 plasmid (three different kits are sold) and Capture-Tec magnetic beads
3 mM EDTA in PBS
Phosphate-buffered saline (PBS; *APPENDIX 2*)

Cell scraper
60-mm tissue culture plates
Magnetic stand (e.g., Invitrogen; other models may be used) or strong magnet
End-over-end rotating mixer

Additional reagents and equipment for mammalian cell culture and counting cells (*APPENDIX 3F*), transfection (*UNITS 9.1-9.4*), and subcloning genes (*UNIT 3.16*)

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and proper sterile technique should be followed accordingly.

NOTE: All culture incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

Choose vector system and transfect cells

1a. *If an expression construct containing the gene of interest already exists:* Cotransfect cells with construct containing gene of interest and pHook-1 (*UNITS 9.1-9.4* and Strategic Planning in this unit). Mock transfect several plates without added DNA to determine the amount of background selection (if any).

Also cotransfect with pHook-1 and control plasmid pcDNA3.1/His/lacZ (provided with the pHook-1 version of the Capture-Tec kit) in place of the construct containing the gene of interest, to assess cotransfection efficiency (see Support Protocol). Carry these control cells through all subsequent steps.

Transfection conditions for each cell line should be optimized to ensure the highest efficiency (see UNITS 9.1-9.4).

1b. *If it is desirable to link expression of the gene of interest with the sFv hook (see Commentary):* Subclone gene of interest into the multicloning site of either pHook-2 or pHook-3 (*UNIT 3.16*), then transfect cells with the plasmid (*UNITS 9.1-9.4* and Strategic Planning in this unit). Mock transfect several plates without added DNA to determine the amount of background selection (if any).

Also transfect cells with pHook-2/lacZ or pHook-3/lacZ control plasmids (provided, respectively, with the pHook-2 and pHook-3 versions of the Capture-Tec kit) to assess transfection efficiency (see Support Protocol). Carry these control cells through all subsequent steps.

Transfection conditions for each cell line should be optimized to ensure the highest efficiency (see UNITS 9.1-9.4).

2. After transfection (or mock transfection), return cells to incubator for 2 to 48 hr.

Harvest the transfected cell population

3. After incubation, replace medium with 3 mM EDTA in PBS. Harvest cells, using a cell scraper if necessary, and transfer to a centrifuge tube.

If the cell line being used tends to form aggregates while in culture, low levels of trypsin (0.05%) can be added to disperse them at this point. Keep in mind, however, that under these conditions the sFv membrane tag (see Background Information) may be cleaved, thereby lowering the number of cells selected overall.

4. Centrifuge cells 5 to 10 min at 800 to 1000 × g, room temperature. Decant supernatant.
5. Resuspend cells in 1 ml complete medium per 60-mm plate (~1 × 10⁶ cells). Transfer each sample to a single 1.5-ml microcentrifuge tube.

It is important to achieve a single-cell suspension to ensure efficient selection. For many cell lines this can be done by pipetting the sample up and down in a small-bore pipet. A Pasteur pipet with a flame-polished tip works well for this purpose.

Select transfected cells with magnetic beads

6. Resuspend the Capture-Tec magnetic bead slurry by vortexing, then place 10 μl (1.5 × 10⁶ beads) into each of a series of microcentrifuge tubes.

Prepare sufficient beads for the three magnetic separations described below.

7. Wash beads in each tube by resuspending in 1 ml complete medium, then separating them out into a pellet with a magnetic stand or strong magnet and removing the medium.
8. Add each of the cell suspensions from step 5 to one of the microcentrifuge tubes containing washed Capture-Tec beads. Incubate 30 min at 37°C on an end-over-end rotator at 5 to 10 rpm.
9. Place the tubes containing the bead/cell mixture in a magnetic stand and mix for 30 sec to 1 min with gentle rocking. Remove supernatant containing unbound cells.

If working with pHook-1, retain the supernatants with the unselected cells from the pcDNA3.1/His/lacZ-transfected controls for use in the Support Protocol. If working with pHook-2 or -3, retain the supernatants from the pHook-2/lacZ or pHook-3/lacZ controls.

10. Remove the tubes from the magnetic stand and resuspend the beads and cells in 1 ml complete medium. Vortex gently to resuspend cells.

It is important to resuspend cells thoroughly to disperse any aggregates that may have formed during the selection procedure. Cell aggregates can trap beads and be selected nonspecifically.

11. Repeat magnetic separation on the cells (steps 8 to 10) twice more.
12. Resuspend selected cells in a small volume (50 to 100 μl) of complete medium or PBS and count cells (APPENDIX 3F).

If the cells are to be lysed, the beads may be removed after lysis by using a magnet to pellet the beads. The supernatant may then be transferred to a new tube.

When observing cells microscopically after selection, the magnetic beads can be easily seen. Unbound beads should be visible, indicating that the beads are in excess over the number of transfected cells. If unbound beads are not visible, it may indicate that not all the transfected cells were selected. Increasing the number of beads to 20 μ l (3×10^6 beads) the next time cells are selected will ensure that all of the transfected cells have an ample chance to interact with the beads and be selected.

OPTIMIZATION OF COTRANSFECTION CONDITIONS

The positive control plasmid pcDNA3.1/His/lacZ, provided with pHook-1, results in the expression of β -galactosidase when successfully cotransfected into cells along with pHook-1. Similarly, the plasmids pHook-1/lacZ and pHook-2/lacZ result in expression of β -galactosidase when successfully transfected. Hence, these plasmids can be used to check for cotransfection in selected cells and to assess transfection efficiencies. To do so, cells transfected with one of the above plasmids are magnetically selected (see Basic Protocol 3) then selected cells, nonselected cells, and mock-transfected cells are stained with 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (Xgal) and counted. Cells expressing β -galactosidase will turn blue in the presence of Xgal. Comparison of the number of blue nonselected cells versus blue selected cells will allow the investigator to determine the selection efficiency. Untransfected cells (mock transfected) should not stain blue with Xgal. Optimal cotransfection conditions are defined as the pHook-1 to pcDNA3.1/His/lacZ ratio that gives the greatest enrichment of blue-stained cells in the selected population. For pHook-2/lacZ and pHook-3/lacZ controls, optimum transfection conditions are defined as the DNA concentrations that yield the highest number of blue cells selected.

Materials

- Magnetically selected and unselected (supernatant) control cells (see Basic Protocol 3) transfected with:
 - pHook-1 and control plasmid pcDNA3.1/His/lacZ (if using pHook-1)
 - pHook-2/lacZ (if using pHook-2)
 - pHook-3/lacZ (if using pHook-3)
- Magnetically selected and unselected (supernatant) mock transfected cells
- Phosphate-buffered saline (PBS; APPENDIX 2)
- Xgal staining solution (see recipe)
- Tissue culture dishes
- Inverted microscope

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and proper sterile technique should be followed accordingly.

NOTE: All culture incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

1. Centrifuge the suspensions of cells 5 min at 800 to 1000 \times g, room temperature. Decant supernatants.
2. Wash by resuspending each pellet in 1 ml PBS, centrifuging again as in step 1, then decanting the supernatants.
3. Resuspend each pellet of selected and unselected cells in 100 μ l Xgal staining solution. Transfer to tissue culture dishes.
 - Depending on cell type, cells may be incubated in tubes, then counted on a hemacytometer.*
4. Incubate cells 2 to 24 hr at 37°C.

5. Examine dishes under an inverted microscope for development of blue color. Count total cells and blue-stained cells.

Remember to normalize to total cell number.

In pHook-1 cotransfections, nearly all selected cells should express β -galactosidase. If there are a large number of nonselected cells that are blue, increasing the amount of pHook-1 relative to pcDNA3.1/His/lacZ could increase the number of cells selected, but may lower the percentage of selected cells that express β -galactosidase. By design, selected cells expressing pHook-2 and pHook-3 should all express the gene of interest.

REAGENTS AND SOLUTIONS

Xgal staining solution

Inorganic salt mix

PBS, pH 7.4 (APPENDIX 2) containing:

4 mM $K_3Fe(CN)_6$

4 mM $K_4Fe(CN)_6 \cdot 3H_2O$

2 mM $MgCl_2 \cdot 6H_2O$

Prepare fresh

Working solution: Immediately before use, add 20 mg/ml Xgal stock solution in dimethylformamide (store at $-20^\circ C$) to a final concentration of 1 mg/ml.

The authors prepare the inorganic salt mix fresh for each experiment by adding 400 mM $K_3Fe(CN)_6$ stock, 400 mM $K_4Fe(CN)_6$ stock, and 200 mM $MgCl_2 \cdot 6H_2O$ stock (all of which may be stored frozen at $-20^\circ C$) to 10 \times PBS (APPENDIX 2) at ten times the final concentrations listed above, then diluting this 10 \times mix to 1 \times with water. 20 mg/ml Xgal stock solution is then added immediately before use to a concentration of 1 mg/ml, as described above.

For further description of Xgal, see Table 1.4.2.

COMMENTARY

Background Information

Isolation of a cell line stably transfected with DNA encoding a particular gene is dependent on a selection strategy that affords the transfected cell a growth advantage in culture. To implement this approach, an antibiotic-resistance marker is transfected into cells along with the gene of interest. The antibiotic-resistance genes allow the transfected cells to survive in a culture containing the corresponding antibiotic, while cells that do not have that marker stably integrated into their genome perish after several weeks in culture.

Thymidine kinase (*TK*) was the first gene to be extensively used for selection in mammalian cells. The use of 5-bromodeoxyuridine (BrdU) to provide an easy selection for TK^- cells resulted in establishment of several such lines, one of which was used to demonstrate that calcium phosphate-mediated transfection of a viral *TK* gene could lead to establishment of TK^+ cells that contained a stably integrated copy of the viral gene (Wigler et al., 1977). This advance made it possible to introduce foreign genes into mammalian cells by transfecting

them with a *TK*-expressing plasmid (cotransfection; Perucho et al., 1980; Robins et al., 1981).

The difficulty with using *TK* as a selectable marker is that mammalian cells are TK^+ and only a very low level of *TK* expression is needed for cells to survive in HAT medium, necessitating the use of a TK^- parental cell type. This motivated the development of vectors that express bacterial drug resistance genes in mammalian cells. These genes can act as dominant selectable markers in virtually every mammalian cell type, as no endogenous mammalian gene can confer resistance. Four commonly used genes of this sort are *APH* (commonly referred to as *neo*), *XGPRT* (*gpt*), *HPH*, and *puro* (see Basic Protocol 2). These genes allow introduction of exogenous DNA into any cell that can be transfected.

While bacterial-derived markers allow facile introduction of genes into mammalian cells, they do not allow selection over a wide range of selective conditions. The ability to select for expression over a greater than thousand-fold range in drug concentration provides the ability

to overexpress newly introduced genes. Mammalian cells that achieve increased resistance to a drug generally have increased the copy number of the gene conferring drug resistance, a process termed amplification. Any gene that is cotransfected with the marker gene will integrate near that gene, and will thus also become amplified. Amplification of the *DHFR* gene through use of increasing amounts of the drug methotrexate has been widely used to overexpress exogenous genes. The adenosine deaminase (*ADA*) gene can also be amplified by using increasing concentrations of 2'-deoxycoformycin (dCF; Kaufman et al., 1986).

The antibiotic-resistance marker method works well when the gene of interest has either a neutral or a stimulatory effect upon cell growth. However, if a gene has a negative or toxic effect on cells, any growth advantage afforded transfected cells by an antibiotic-resistance marker may be negated. The magnetic-bead selection method (see Basic Protocol 3) helps overcome this problem. By using the magnetic-bead system, a homogeneous population of transfected cells can be isolated from culture within hours of introduction of DNA rather than weeks or months later as is the case with standard antibiotic-selection technologies. The system described in Basic Protocol 3 employs a single-chain antibody (sFv) directed against a hapten (phenyloxazalone; phOx) to isolate transiently transfected cells from total populations in culture. Three plasmids can be used: pHook-1, pHook-2, and pHook-3 (Fig. 9.5.4). Each of these encodes a fusion protein between the anti-phOx sFv (NQ10/12.5; Griffiths et al., 1984; Hoogenboom et al., 1991), and the transmembrane domain of the platelet derived growth factor receptor (PDGFR; Gronwald et al., 1988). When expressed in transfected cells, the fusion protein anchors to the membrane via the transmembrane domain of the PDGFR, and the functional phOx sFv (sFv hook) is displayed on the cell surface. Cells transiently transfected with the pHook vectors can then be selected from culture using antigen-coated magnetic beads (Fig. 9.5.5). This system bypasses the need for creating stably transfected cell lines in order to generate homogeneous populations of cells expressing an exogenous gene.

pHook-2 and pHook-3 are designed to constitutively express both the sFv hook (driven by the RSV promoter) for cell selection and a second gene (driven by the CMV promoter) from a single plasmid. These vectors can be used to enrich for populations in which 100%

of selected cells express the exogenous gene (Chesnut et al., 1996). pHook-2 and pHook-3 both have the same promoters driving expression of the sFv hook and the multicloning site. The differences between pHook-1 and pHook-2 lie in the antibiotic resistance system employed, plasmid size, and unique restriction sites in the multicloning site. pHook-2 expresses the ampicillin-resistance gene (β -lactamase) for selection in bacteria and a gene encoding neomycin resistance (aminoglycoside phosphotransferase) for stable selection in mammalian cells. pHook-3 which is significantly smaller, contains the resistance gene for Zeocin (the *Sh ble* gene product; see Basic Protocol 2 and Drocourt et al., 1990), which acts as a selectable marker in both bacteria and mammalian cells.

The systems described above are useful in studying a wide variety of molecules and their effect on cellular biochemistry. Transiently transfected cells expressing novel proteins or RNA molecules, even those that exert a negative or toxic effect on cell growth, can be selected from culture as homogeneous populations at specific time points and subjected to biochemical characterization. In addition to measuring the effect of gene expression on cell morphology or the modification of specific cellular proteins, the temporally dependent activation and inactivation of downstream genes can be determined by techniques such as differential-display PCR (*UNIT 25B.3*; Liang et al., 1993).

Critical Parameters

General considerations

The efficiency with which the parental cell line can be transfected and the degree of expression of the selective marker determine the success of an attempt to produce stable cell lines. As well, it is important to carefully consider whether the parental cell line is appropriate to the final goal of the research. At least 1 month is required to prepare stable cell lines, and many researchers have reached the end of that time realizing that the resultant line would be much more useful if the initial transfection had been planned more carefully. For example, if stable cell lines are being produced to analyze how a promoter is induced during differentiation, a parental line should be used that differentiates well and an internal reference promoter might be included in the initial transfection.

In order to produce stable cell lines, it is obviously necessary that the selective marker

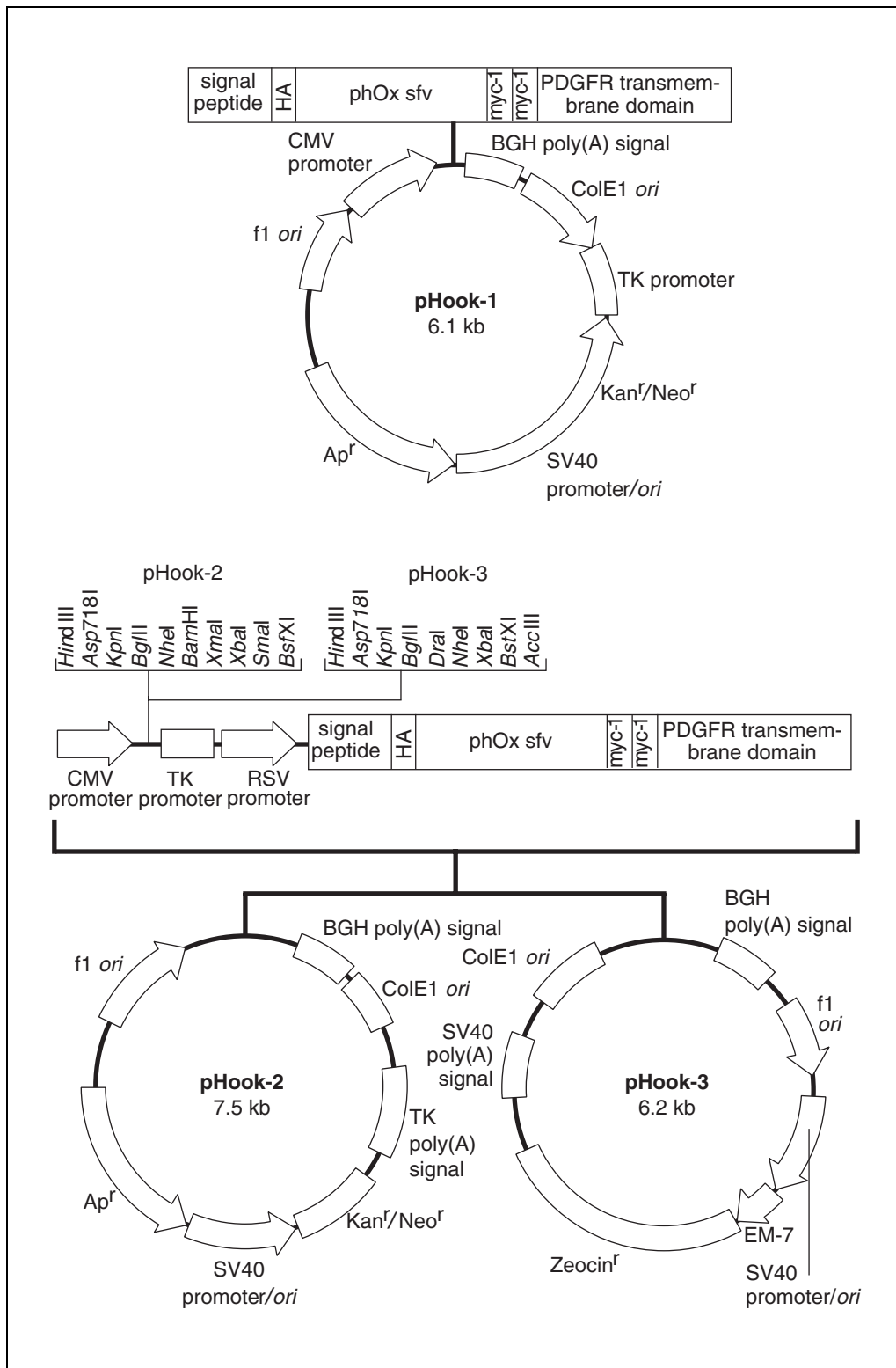


Figure 9.5.4 Plasmid maps of pHook-1, pHook-2, and pHook-3. pHook-1 expresses a fusion protein consisting of the signal peptide from Ig κ V-J2-C region fused with the anti-phOx sFv and the transmembrane domain from the PDGF receptor. pHook-2 and pHook-3 express the same fusion protein but also contain a separate expression cassette consisting of the CMV promoter, a multiple cloning site, and a polyadenylation signal. Definitions: EM-7, synthetic bacterial promoter; HA, hemagglutinin epitope tag; myc-1, myc epitope tag (see text for additional definitions).

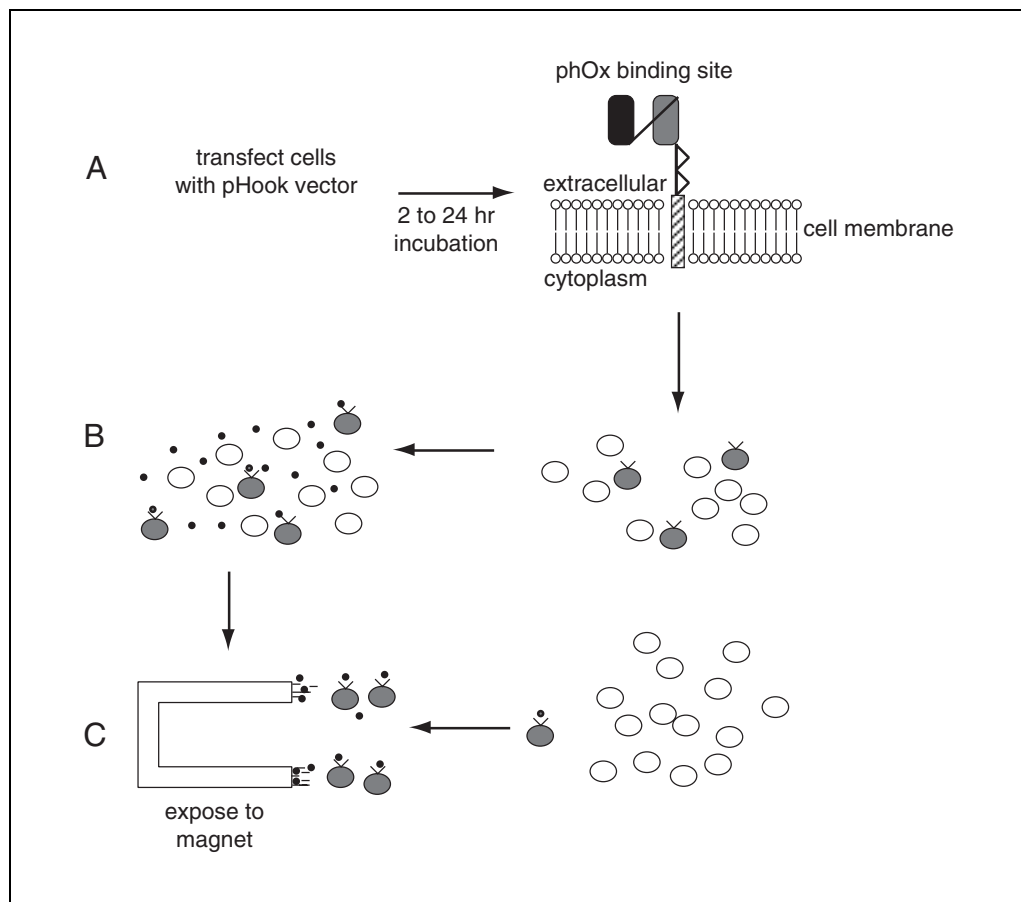


Figure 9.5.5 Magnetic cell-selection protocol. **(A)** Cells are transfected with one of the pHook vectors and returned to culture for 2 to 24 hr. **(B)** The cell population, including transfected cells displaying the sFv hook (dark circles) and untransfected cells (open circles), is then harvested and incubated with antigen-coated magnetic beads (small dots) for 30 min. **(C)** Selection of transfected cells is achieved by exposing the cell/bead mixture to a magnet and discarding cells that are not retained by the magnet.

be expressed at a level adequate to detoxify the selective drug. This is most easily accomplished using *TK* as a marker when transfecting TK^- cells, or when using any bacterial gene (*APH*, *XGPRT*, *HPH*, or *puro*) as a dominant selectable marker. This is because the recipient cells contain no endogenous gene that detoxifies the selective drugs, so a small level of synthesis from the transfected gene will give the transfected cells a significant advantage over the parental cells. This is much harder to accomplish when using DHFR or ADA to transfect a normal cell type. Mammalian cells can have fairly high endogenous levels of ADA and DHFR, and therefore extremely high levels of expression from the transfected gene may be necessary to enable the recipient cells to grow under selective conditions. While it has proven possible to select stable transfectants of normal cells using both ADA and DHFR, it is difficult. If these markers are used, a parental line that is ADA- or DHFR-deficient greatly increases the

yield of transfectants, as a much lower level of selection can be used. If amplification of the transfected genes is desired, use of the deficient lines is recommended as there is no risk of amplifying the endogenous gene when the selective pressure is applied.

If a cell line cannot be efficiently transfected, it can prove extremely difficult to select stable cell lines using any selective marker. It is essential, then, to have a protocol that provides efficient transfection of the parental cell type. Optimization of transfection conditions for transient expression (see introduction to Section I) usually produces conditions that work well for production of stable cell lines. The exception to this is DEAE-dextran-mediated transfection, which generally works less efficiently for production of stable cell lines than either calcium phosphate transfection or electroporation even with a cell type that is transfected efficiently by DEAE-dextran as judged at transient times after transfection.

Finally, when setting up a transfection to produce stable lines, it is important to remember that the results of the transfection will not be known for two weeks. Having one dish of a transfection fail can therefore be extremely frustrating if a duplicate dish was not done. Many investigators therefore do transfections in duplicate or triplicate to ensure at least one set of successful transfections when trying to produce a particular stable cell line.

Magnetic bead selection

To ensure the highest number of selected cells in a particular experiment, it is important to optimize the transfection conditions for that cell line. Once the transfection conditions are optimized, the choice of pHook vector must be made. If the gene of interest is already cloned into an expression vector, pHook-1 may be the best choice. This vector is designed as a cotransfection plasmid, and since the sFv hook is expressed from the strong CMV promoter, it offers the highest efficiency selection of the pHook vectors. Cells cotransfected with pHook-1 and a reporter vector have been selected at high efficiency where 95% of the selected cells express the gene of interest (Chesnut et al., 1996). If the analysis of gene activity requires very low or no untransfected cell background, then one may wish to subclone the gene of interest into either pHook-2 or pHook-3. These vectors are designed to express both the sFv membrane tag and the gene of interest from the same construct. While these vectors allow the isolation of cells which all express the cloned gene, the overall selection efficiency is lower than with cells transfected with pHook-1. This is most likely due to the fact that the sFv-hook expression is driven by the weaker RSV promoter in these vectors.

One of the major obstacles to efficient selection is aggregation of cells in culture. The efficiency of selection decreases in cases where cells aggregate in such a way that untransfected cells are brought down into the pellet along with those that are transfected. The severity of aggregation varies from cell line to cell line. In some cases, however, limited trypsinization prior to the selection step can increase efficiency. This treatment should be undertaken only as a last resort since the exposed sFv fusion protein contains several exposed lysines and arginines that may be cleaved by trypsin, thereby reducing cell selection.

Anticipated Results

Approximately 10^3 colonies can be produced by transfecting 1 μg of a plasmid expressing a selective marker into 10^6 cells from a highly transfectable cell line (e.g., HeLa, CHO DUKX, NIH 3T3). Approximately 10% of these will be large, vigorously growing colonies that can readily be expanded into cell lines. The precise number of stable transfectants to be expected can vary dramatically depending on the cell type and the efficiency with which the marker gene is expressed.

In the magnetic-bead selection system, the number of cells that are selected in a particular experiment will vary with the transfection efficiency and the ability of the cells to secrete and display the sFv membrane tag. In general, selection efficiencies range from 2% to 25% in most adherent human cell lines. The overall number of cells selected using pHook-1 is usually higher than with pHook-2 or pHook-3. In many cases, selection is up to 50% higher with pHook-1. As mentioned above, this decrease in selection with pHook-2 and pHook-3 is accompanied by an increase of the percentage of selected cells expressing the gene of interest, from ~95% to 100%.

Time Considerations

It takes 1 day to set up the DNA and perform the initial transfections. The transfected cells should double twice before being split into selective medium, and after being placed in selective medium generally take 10 to 12 days to form colonies. Another week or so is required to expand the colonies into cell lines. The entire procedure takes ~1 month.

The overall time investment required for the magnetic bead selection procedures comprises mostly the incubations required for tissue culture. Once the cells have been transfected and allowed to express the gene of interest and the sFv hook, the actual selection procedure can be completed in <1.5 hr.

Literature Cited

- Cheng, Y.C., Huang, E.S., Lin, J.C., Mar, E.C., Pagano, J.S., Dutschman, G.E., and Grill, S.P. 1983. Unique spectrum of activity of 9-[(1,3-dihydroxy-2-propoxy)methyl]guanine against herpesviruses in vitro and its mode of action against herpes simplex virus type 1. *Proc. Natl. Acad. Sci. U.S.A.* 80:2767-2770.
- Chesnut, J.D., Baytan, A.R., Russell, M., Chang, M.-P., Bernard, A., Maxwell, I.H., and Hoeffler, J.P. 1996. Selective isolation of transiently transfected cells from a mammalian cell population

- with vectors expressing a membrane anchored single-chain antibody. *J. Immunol. Methods* 193:17-27.
- de la Luna, S., Soria, I., Pulido, D., Ortin, J., and Jimenez, A. 1988. Efficient transformation of mammalian cells with constructs containing a puromycin-resistance marker. *Gene* 62:121-126.
- Drocourt, D., Calmels, T., Reynes, J.P., Baron, M., and Tiraby, G. 1990. Cassettes of the *Streptoloteichus hindustanus ble* gene for transformation of lower and higher eukaryotes to phleomycin resistance. *Nucl. Acids Res.* 18:4009.
- Fukushige, S. and Sauer, B. 1992. Genomic targeting with a positive-selection *lox* integration vector allows highly reproducible gene expression in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 89:7905-7909.
- Gossen, M. and Bujard, H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U.S.A.* 89:5547-5551.
- Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W., and Bujard, H. 1995. Transcriptional activation by tetracyclines in mammalian cells. *Science* 268:1766-1769.
- Griffiths, G.M., Berek, C., Kaartinen, M., and Milstein, C. 1984. Somatic mutation and the maturation of immune response to 2-phenyl oxazolone. *Nature* 312:271-275.
- Gritz, L. and Davies, J. 1983. Plasmid-encoded hygromycin-B resistance: The sequence of hygromycin-B-phosphotransferase gene and its expression in *E. coli* and *S. cerevisiae*. *Gene* 25:179-188.
- Gronwald, R.G., Grant, F.J., Haldeman, B.A., Hart, C.E., O'Hara, P.J., Hagen, F.S., Ross, R., Bowen-Pope, D.F., and Murray, M.J. 1988. Cloning and expression of a cDNA coding for the human platelet-derived growth factor receptor: Evidence for more than one receptor class. *Proc. Natl. Acad. Sci. U.S.A.* 85:3435-3439.
- Hartman, S.C. and Mulligan, R.C. 1988. Two dominant-acting selectable markers for gene transfer studies in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 85:8047-8051.
- Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P., and Winter, G. 1991. Multi-subunit proteins on the surface of filamentous phage: Methodologies for displaying antibody (Fab) heavy and light chains. *Nucl. Acids Res.* 19:4133-4137.
- Kaufman, R.J., Murtha, P., Ingolia, D.E., Yeung, C.-Y., and Kellems, R.E. 1986. Selection and amplification of heterologous genes encoding adenosine deaminase in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 83:3136-3140.
- Kozak, M. 1989. The scanning model for translation: An update. *J. Cell Biol.* 108:229-241.
- Liang, P., Averboukh, L., and Pardee, A.B. 1993. Distribution and cloning of eukaryotic mRNAs by means of differential display: Refinements and optimization. *Nucl. Acids Res.* 21:3269-3275.
- Littlefield, J.W. 1964. Selection of hybrids from matings of fibroblasts in vitro and their presumed recombinants. *Science* 145:709-710.
- Mullen, C.A., Kilstrup, M., and Blaese, R.M. 1992. Transfer of the bacterial gene for cytosine deaminase to mammalian cells confers lethal sensitivity to 5-fluorocytosine: A negative selection system. *Proc. Natl. Acad. Sci. U.S.A.* 89:33-37.
- Mulligan, R.C. and Berg, P. 1981. Selection for animal cells that express the *E. coli* gene coding for xanthine-guanine phosphoribosyltransferase. *Proc. Natl. Acad. Sci. U.S.A.* 78:2072-2076.
- Mulsant, P., Gatignol, A., Dalens, M., and Tiraby, G. 1988. Phleomycin resistance as a dominant selectable marker in CHO cells. *Somatic Cell Mol. Genet.* 14:243-252.
- No, D., Yao, T.P., and Evans, R.M. 1996. Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* 93:3346-3351.
- Palmer, T.D., Hock, R.A., Osborne, W.R.A., and Miller, A.D. 1987. Efficient retrovirus-mediated transfer and expression of a human adenosine deaminase gene in diploid skin fibroblasts from an adenosine-deficient human. *Proc. Natl. Acad. Sci. U.S.A.* 84:1055-1059.
- Perucho, M., Hanahan, D., and Wigler, M. 1980. Genetic and physical linkage of exogenous sequences in transformed cells. *Cell* 22:309-317.
- Robertson, E.J. 1987. Embryo-derived stem cell lines. In *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach* (E.J. Robertson, ed.) pp. 71-112. IRL Press, Oxford and New York.
- Robins, D.M., Ripley, S., Henderson, A.S., and Axel, R. 1981. Transforming DNA integrates into the host chromosome. *Cell* 23:29-39.
- Simonsen, C.C. and Levinson, A.D. 1983. Isolation and expression of an altered mouse dihydrofolate reductase cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 80:2495-2499.
- Southern, P.J. and Berg, P. 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *J. Mol. Appl. Gen.* 1:327-341.
- Staschke, K.A., Colacino, J.M., Mabry, T.E., and Jones, C.D. 1994. The in vitro anti-hepatitis B virus activity of FIAU [1-(2'-deoxy-2'-fluoro-1-β-D-arabinofuranosyl-5-iodo)uracil] is selective, reversible, and determined, at least in part, by the host cell. *Antivir. Res.* 23:45-61.
- Sugiyama, M., Thompson, C.J., Kumagai, T., Suzuki, K., Deblaere, R., Villarroel, R., and Davies, J. 1994. Characterisation by molecular cloning of two genes from *Streptomyces verticillus* encoding resistance to bleomycin. *Gene* 151:11-16.
- Wei, K. and Huber, B.E. 1996. Cytosine deaminase gene as a positive selection marker. *J. Biol. Chem.* 271:3812-3816.

Wigler, M., Silverstein, S., Lee, L-S., Pellicer, A., Cheng, Y.-C., and Axel, R. 1977. Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. *Cell* 11:223-232.

Yagi, T., Ikawa, Y., Yoshida, K., Shigetani, Y., Takeda, N., Mabuchi, I., Yamamoto, T., and Aizawa, S. 1990. Homologous recombination at c-fyn locus of mouse embryonic stem cells with use of diphtheria toxin A-fragment gene in negative selection. *Proc. Natl. Acad. Sci. U.S.A.* 87:9918-9922.

Contributed by Richard Mortensen
Brigham and Women's Hospital
Boston, Massachusetts

Jonathan D. Chesnut and James P. Hoeffler
(magnetic bead selection)
Invitrogen Corporation
Carlsbad, California

Robert E. Kingston
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

USES OF FUSION GENES IN MAMMALIAN TRANSFECTION

USING FUSION GENES TO MEASURE PROMOTER ACTIVITY

Transfection is the most commonly used procedure for analyzing mammalian gene expression *in vivo* (introduction to Section I, *UNITS 9.1-9.4, 16.12 & 16.13*). Early studies of gene expression relied on the creation of clonal cell lines that contained the promoter of interest. Creation of such lines is extremely time-consuming, however, and it was found that chromosomal location can have significant effects on gene expression. Alternatively, expression can be measured by simply introducing the intact gene of interest into cells and measuring the level of RNA produced from this gene (*UNIT 9.8*). This method of analysis is direct, difficult, not very sensitive, and time-consuming.

Many experimenters therefore use transient assay systems that are based on the use of fusion genes and assess gene expression within 48 hr after introduction of the DNA. The fusion gene usually consists of the promoter activator binding site or enhancer sequence under study attached to a gene directing the synthesis of a reporter molecule; the amount of reporter protein synthesized under various conditions is presumed to reflect the ability of the inserted sequence to direct and/or promote transcription. It follows that a useful transient expression system should be based on the synthesis of an easily assayed reporter protein that has minimal or no effects on the physiology of the transfected cell. Ideally, the assay for this reporter molecule would be extremely sensitive. This section presents a general overview of reporter systems (*UNIT 9.6*), followed by specific protocols that describe the use of selected reporter systems that fulfill all of the above requirements. These examples include two isotopic assays using chloramphenicol acetyltransferase (CAT) and human growth hormone (hGH) expression plasmids (*UNIT 9.7A*), two chemiluminescent assays using firefly luciferase and β -galactosidase expression plasmids (*UNIT 9.7B*), and fluorescent assays using expression plasmids carrying green fluorescent protein (*UNIT 9.7C*).

CAT has been the most frequently used reporter enzyme, and thus a major advantage of this system is that it is an extensively validated and widely accepted measure of promoter activity. Another advantage is that the enzyme assays are relatively easy and quite reliable but can be time-consuming. A disadvantage of this system is that, for reasons that are not completely clear, CAT mRNA accumulates to only very low levels. This can make it difficult to demonstrate that the promoter of interest is initiating transcription at the appropriate position.

Unlike the enzymatic assays for CAT, β -galactosidase, and luciferase, the hGH reporter assay is immunologically based, employing commercially available radioimmunoassay kits that are relatively quick and convenient to use. Moreover, unlike the enzymatic or mRNA assays, hGH is assayed in the medium, rather than in cell extracts. This makes it possible to monitor expression directed by a single population of transfected cells over time and also makes hGH useful as an internal control for any of the enzymatic assays. Such internal controls can be very helpful in comparing expression levels between different test plasmids. Based on ease of detecting expression from various comparable plasmids, the hGH assay seems to be a slightly more sensitive than the CAT assay. This is at least partly a consequence of the substantially greater accumulation of hGH mRNA.

Luciferase is another widely used reporter enzyme. The luciferase assay is much faster and more sensitive than the CAT assay and does not use radioactivity. The increased

sensitivity is particularly useful if the cell line of interest is hard to transfect, or if the promoter of interest is very weak. A disadvantage of the luciferase assay is that it requires a rather expensive piece of equipment, the luminometer, to measure enzyme activity.

The use of β -galactosidase coupled with a chemiluminescent detection system provides a quick and sensitive alternative nonisotopic reporter system. In this assay (see *UNIT 9.7B*), the 1,2-dioxetane-substituted AMPGD (Galacton, Tropix) is used as a substrate for the β -galactosidase assay. Following a short incubation of cell extract with enzyme, light emission is measured using either a luminometer or a scintillation counter. The sensitivity of the chemiluminescent β -galactosidase assay is comparable to that of the luciferase assay.

The green fluorescent protein (GFP), cloned from the jellyfish *Aequorea victoria*, serves as an easily detected fluorescent tag requiring no substrates or associated cofactors. GFP—along with its variously engineered versions—has increasingly become widely used as a reporter of gene expression and for studies of protein localization in vivo (see *UNIT 9.7C*). A number of GFP expression systems are now available, and have been used successfully to introduce GFP into a wide variety of organisms and cell types, and into virtually any subcellular region.

A potentially serious problem with using reporter genes to analyze promoters is that sequences within a gene can play a significant role in regulation of its promoter. Thus, subcloning a promoter into a vector that expresses a reporter gene may separate the promoter from sequences necessary for the transcriptional response under study. Similarly, the reporter gene may contain sequences that affect the ability of particular promoter sequences to function normally. It is therefore important to verify that any regulatory event observed with a fusion gene also affects expression of the gene in vivo, both in terms of kinetics and magnitude of the response. It is also a good idea to assess the mRNA by S1 analysis (*UNIT 4.6*) to make sure that the subcloned promoter still initiates transcription at the correct start site.

Vectors are available that contain a reporter gene and mammalian processing signals adjacent to restriction sites into which a promoter of interest can be inserted (see Figs. 9.7.1 and 9.7.2). When fusing a promoter to a reporter gene, it is essential to ensure that there are no ATG codons between the transcriptional start site and the ATG of the reporter gene. In particular, it is important to check that no ATG is created when the fusion is engineered.

CHARACTERIZATION OF TRANSFECTION EFFICIENCY THROUGH USE OF FUSION GENES

Fusion genes that express an easily assayed protein are critical in establishing optimal transfection conditions. The use of such genes allows rapid and simple assessment of the efficient introduction of DNA into cells under a variety of conditions (introduction to Section I). The ability to monitor transfection efficiency easily is extremely useful in establishing transfection conditions and in ensuring that a transfection has worked prior to extensive workup of the transfected cells, e.g., when doing S1 analysis (*UNIT 4.6*).

Overview of Genetic Reporter Systems

A central question in molecular and cell biology is how *cis*-acting DNA sequences and *trans*-acting factors act in unison to control eukaryotic gene expression. These interactions are mediated by specific binding of a transcription factor or a complex of factors to enhancer and promoter elements generally found upstream of the transcription start site. Direct quantitation of changes in gene expression requires the measurement of specific mRNAs using techniques such as northern blot hybridizations (UNIT 4.9) or nuclease protection assays (UNIT 4.7). These procedures can be time-consuming and are not always practical for analysis of many different gene constructs. Furthermore, such techniques require that the analyzed gene be modified in some manner to distinguish it from the native gene in transfected cells.

An alternate approach to gauge changes in transcription is to link the presumed *cis*-acting sequence(s) from the gene of interest to the

coding sequence for an unrelated reporter gene (Alam and Cook, 1990). To test for complete promoters, the DNA fragment is placed directly upstream of the reporter gene in a vector lacking endogenous promoter activity (Fig. 9.6.1). Similarly, for activator binding sites, the heterologous sequence is placed in a vector proximal to a basal promoter that contains sequences required for recognition by the basic transcription machinery (e.g., RNA polymerase II). Following introduction of the chimeric reporter construct into an appropriate cell type or animal, measurement of reporter-gene product provides an indirect estimate of the induction in gene expression directed by the regulatory sequences. It is important to recognize that these assays use reporter constructs to measure protein level or activity and not RNA level. The two levels are frequently, but not always, correlated with one another. Any treatment of the cells or extracts that may affect message translation will alter the accuracy of measurements

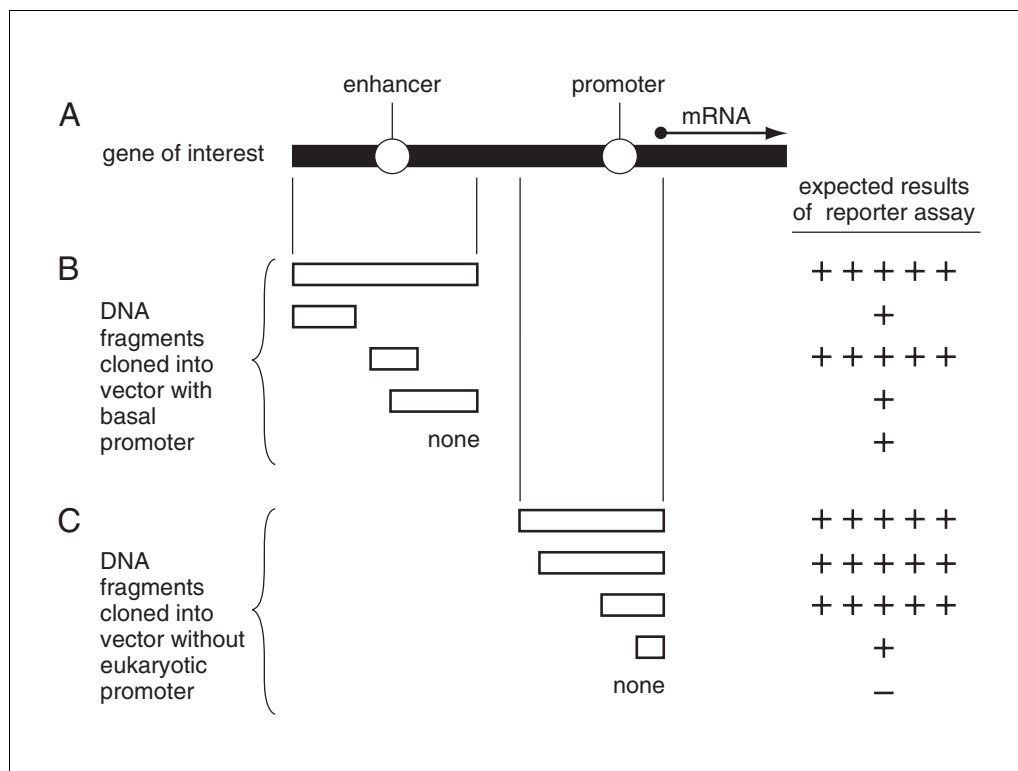


Figure 9.6.1 Schematic of promoter and enhancer definition studies. (A) A hypothetical eukaryotic gene with promoter and enhancer elements indicated by open circles. (B) A series of DNA fragments that could be cloned into the MCS of a reporter vector containing a basal promoter for the purpose of identifying the position of an enhancer. (C) Cloning of DNA into a promoterless reporter vector to define the minimal functional promoter. The expected results of the subsequent reporter assays for these constructs are shown to the right of the corresponding DNA fragments. + and +++++ indicate lesser and greater expression, and - indicates no expression.

of the reporter construct. Furthermore, when dealing with an inducible system, the stability of the reporter must be taken into account. For example, if the reporter-gene product is very stable, high levels of protein can build up in the cell prior to induction. Upon induction following addition of the effector, the protein level may well increase, but the observed fold increase may be less dramatic because of the high basal level of reporter protein that has already accumulated in the cell.

With these considerations, the reporter can be used to characterize the cellular, tissue, or temporal specificity of the *cis*-acting element or its responsiveness to external stimuli such as hormones. Furthermore, *cis*-regulatory elements can be functionally characterized by constructing successive deletion mutants (Fig. 9.6.1), by altering sequence orientations, or by site-directed mutagenesis. Moreover, the requirements for *trans*-acting factor(s) can be investigated by expression of the reporter construct in different cell types and organisms or by biochemical manipulation of the source of the factors. For example, involvement of specific signaling pathways in the response to external stimuli can be investigated by treating cells with inhibitors and/or activators of protein phosphorylation, glycosylation, or fatty acylation. Kinetic studies with protein synthesis inhibitors may be employed to correlate the requirement for new protein synthesis with induction of the reporter gene.

There are several criteria for selection of a transcription reporter gene. (1) The reporter protein should be absent from the host, or easily distinguished from endogenous versions. (2) A simple, rapid, sensitive, and cost-effective assay should be available to detect the reporter protein. (3) The assay for the reporter protein should have a broad linear range to facilitate analysis of both large and small changes in promoter activity. (4) Expression of the gene must not alter the physiology of the recipient cells or organism. These criteria are met to varying degrees by the reporter systems outlined in Table 9.6.1. Each of these reporter genes and the corresponding assay systems have specific features and limitations that must be considered in choosing a system tailored to the particular question being studied. For example, the reporter protein may be expressed intracellularly or secreted from the cell and assayed in the culture medium. There may be limitations in the cell types suitable for a particular reporter system due to low expression levels or the presence of endogenous activities

that contribute to high background. The reporter proteins that are commonly used vary widely in their relative stabilities, which is an important consideration in the study of inducible reporter constructs. Finally, the reporter protein may be measured by an *in vitro* activity or immunological assay or detected *in vivo* via histochemical procedures. These and other issues will be addressed for each reporter gene in the sections that follow.

DESIGN OF REPORTER VECTORS

The general design of a reporter vector referred to as “pGENERIC,” is shown in Figure 9.6.2. The vector backbone typically contains a bacteriophage origin of replication (*f1 ori*) for single-stranded DNA production to facilitate sequencing and mutagenesis studies, and a bacterial origin of replication (often from pBR322 or pUC-derived plasmids) for propagation of the vector in *Escherichia coli*. Additional origins of replication are often added to permit stable propagation in other host species. Production of the vector in bacteria is also facilitated by an antibiotic resistance gene, in most cases coding for β -lactamase, which imparts ampicillin resistance to *E. coli* cells transformed with the vector. A polyadenylation signal downstream of the reporter gene ensures proper and efficient processing of the reporter transcript in eukaryotic cells. A second polyadenylation site may be placed upstream of the reporter to prevent background transcription from cryptic promoters located in the flanking DNA (Araki et al., 1988). Lastly, a multiple cloning site (MCS) upstream of the reporter gene allows for the insertion of foreign DNA containing putative promoter and/or enhancer elements. For cloning purposes, the MCS typically contains five to seven closely aligned unique restriction endonuclease sites. In addition to the MCS, reporter vectors often contain one or more unique restriction sites downstream of the reporter gene for testing putative enhancer elements or for excision of the reporter gene.

Some vectors already contain a basal promoter immediately upstream of the reporter gene. Such promoters are frequently of viral origin (Wenger et al., 1994), and ideally provide a constitutive level of transcription in a broad range of cell types. These reporter vectors are used to identify and characterize activator binding sites and enhancer sequences, and also to study the process of transcription activation. A reporter vector containing a basal promoter, and possibly an upstream enhancer,

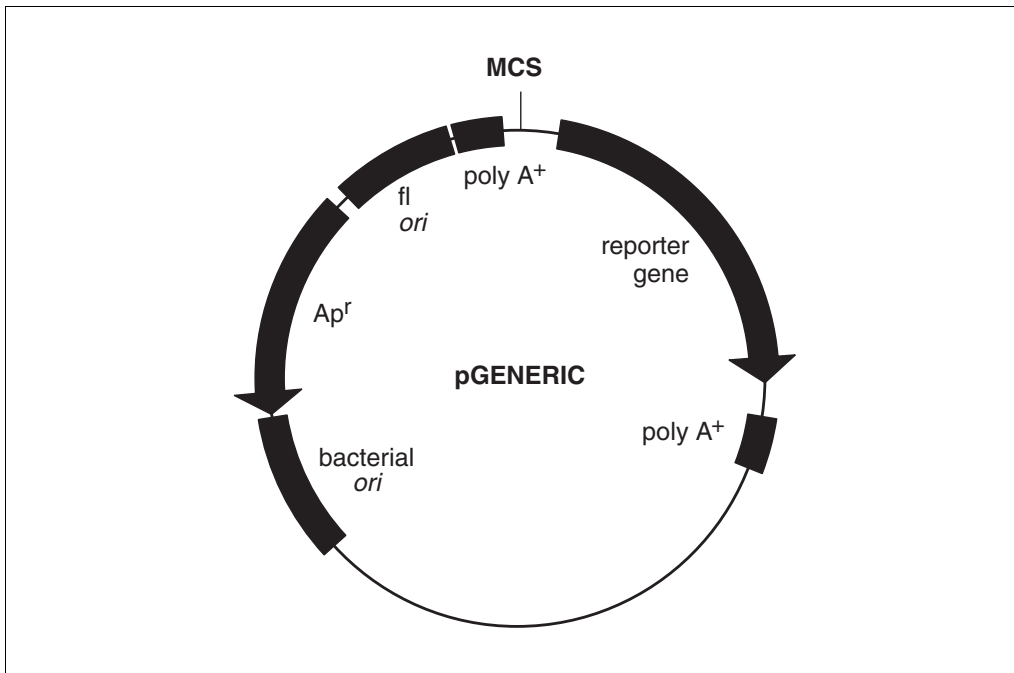


Figure 9.6.2 Plasmid map of a representative reporter vector. MCS denotes the multiple cloning site used for insertion of foreign DNA into the reporter vector. Poly A⁺ indicates the positions of polyadenylation signals. *Ori* refers to the bacteriophage (f1) and bacterial plasmid origins of replication. Arrows indicate the direction of transcription for the reporter gene and an antibiotic resistance gene (Ap^r).

may also be used as a positive control to normalize transfection efficiencies of experimental reporter constructs.

IN VITRO REPORTER ASSAYS

In vitro reporter assays refer to procedures in which the reporter protein is quantified using either cell or tissue lysates containing the reporter, or (for secreted reporter proteins) the culture medium from transfected cells. These assays utilize direct quantitation of the reporter protein by enzymatic or immunological means to provide an indirect estimate of the transcriptional activity of the reporter vector encoding the protein. Such assays are distinguished from the in vivo reporter systems described below in that numerical data suitable for comparative studies of different promoters, enhancers, and cell-type requirements are more readily obtained.

Chloramphenicol Acetyltransferase (CAT)

Many of the above criteria for genetic reporters are satisfied by the enzyme chloramphenicol acetyltransferase (CAT; *UNIT 9.7A*; Gorman, 1982). This enzyme catalyzes the transfer of acetyl groups from acetyl-coenzyme A to chloramphenicol. Because CAT is a prokaryotic enzyme, there are minimal competing ac-

tivities in most eukaryotic cells, resulting in high signal-to-background ratios in these reporter assays. However, most formats of the CAT assay require a relatively expensive radioactive substrate, the assays are very time-consuming to perform, and the sensitivity of CAT assays is inferior to that of recently developed nonisotopic reporter systems. CAT is quite stable (half-life 50 hr in mammalian cells; Thompson et al., 1993), making it well-suited as a reporter in transient transfection experiments, but less desirable for stable expression studies. Moreover, the stability exhibited by CAT can be a disadvantage in inducible systems (e.g., inducible enhancers), where accumulation of the reporter protein may mask the induction brought about by addition of the effector. For applications of this type, a reporter protein with a relatively short half-life, such as firefly luciferase, is more desirable.

The most common CAT assay is based on incubation of cell lysates prepared from transfected cells with ¹⁴C-labeled chloramphenicol. Acetylated and unacetylated forms of this compound are separated by thin-layer chromatography on siliconized glass plates. The plates are exposed to X-ray film for a qualitative estimate of CAT activity. For more quantitative data, the spots corresponding to the positions of the mono- and diacetylated forms of chloram-

phenicol can be scraped from the plate and counted in a liquid scintillation counter. This assay can be particularly tedious with several samples and is difficult to interpret due to the presence of two reaction products. Differential extraction techniques have been developed for the quantitative separation of chloramphenicol from its acetylated derivatives, eliminating the need for time-consuming chromatography (UNIT 9.7A). In this format, the radiolabel resides on acetyl-coenzyme A, and ^3H can be substituted for ^{14}C , making the assay somewhat less hazardous and less expensive. Use of an alternate solvent system results in enhanced sensitivity by improving the signal-to-background ratio (Cassinotti and Weitz, 1994).

There are two nonisotopic detection methods for quantifying CAT. A CAT ELISA procedure performed in 96-well microtiter plates is available to directly quantify CAT using an anti-CAT antibody (from Boehringer Mannheim). This assay has the advantage of directly measuring CAT protein levels rather than CAT activity. A nonisotopic CAT activity assay has also been developed; this uses a fluorescent derivative of chloramphenicol (from Molecular Probes) that undergoes a single acetylation reaction. Because only a single reaction product is produced from this substrate, quantitative analysis of CAT activity is more reliable than with chloramphenicol. Moreover, this assay is more sensitive than those employing a radioactive substrate and has a broader linear range.

Firefly Luciferase

Cloning of the *luc* gene from the firefly *Photinus pyralis* (de Wet et al., 1987) provided the first nonisotopic genetic reporter system with widespread utility in mammalian cells. The bioluminescent reaction catalyzed by luciferase requires luciferin (the substrate), ATP, Mg^{2+} , and molecular O_2 . Mixing these reagents with cell lysates containing luciferase results in a flash of light that decays rapidly (in <1 sec). The light signals are detected using a luminometer equipped with an autoinjection device to facilitate rapid mixing of reaction components. Alternatively, the light signal can be recorded using a liquid scintillation counter (Nguyen et al., 1988). The total light emission is proportional to the luciferase activity of the sample, which in turn provides an indirect estimate of the transcription of the luciferase reporter gene. Luciferase is sensitive to degradation by proteases, and therefore has a half-life of ~ 3 hr in transfected mammalian cells

(Thompson et al., 1993). The rapid turnover of luciferase protein makes this reporter a good candidate for studying inducible systems where the increase above basal expression levels needs to be maximized.

The original luciferase assay provides good sensitivity (10 to 1000 times as sensitive as the radioactive CAT assay), but is somewhat intricate to perform and lacks reproducibility between samples, largely due to the rapid “flash” kinetics of the photon emission. An improved assay for luciferase that includes coenzyme A in the reaction mixture produces a more sustained light signal because of the favorable reaction of luciferase with luciferyl-CoA (luciferyl-coenzyme A; Bronstein et al., 1994). Moreover, the longer duration of the light signal increases luminescence intensity by tenfold, thereby making the luciferase assay more sensitive. Future prospects for improving the assay further include the production of modified luciferase enzymes with shifted emission wavelengths which would permit simultaneous analysis of two or more genetic reporters in the same population of cells via detection of different colors of light.

β -Galactosidase

The *lacZ* gene from *E. coli*, which encodes the enzyme β -galactosidase, is among the most versatile genetic reporters, having both in vitro and in vivo assay formats employing a variety of different substrates. The enzyme catalyzes the hydrolysis of various β -galactosides, including several specialized substrates tailored to different assay formats. In addition to its use as a reporter for uncharacterized *cis*-regulatory sequences, expression of β -galactosidase under the control of a constitutive promoter is frequently used as an internal control to normalize the variability of other reporter assays. β -galactosidase is particularly useful for normalizing CAT and firefly luciferase, as cell lysates prepared for these reporter assays are also suitable for measurement of β -galactosidase activity (Alam and Cook, 1990). Both a colorimetric assay using *o*-nitrophenyl- β -D-galactopyranoside (ONPG) and a fluorometric assay using 4-methylumbelliferyl- β -D-galactoside (MUG) have been developed for β -galactosidase (UNIT 16.20 and Young et al., 1993, respectively), but these have received limited use, due primarily to their poor sensitivity. The development of chemiluminescent 1,2-dioxetane substrates for β -galactosidase has greatly improved the utility of *lacZ* as a transcriptional reporter by increasing the sen-

sitivity of the assay and extending the linear dynamic range of detection (*UNIT 9.7B*; Bronstein et al., 1994). When a luminometer is used to detect the chemiluminescent signal, the assay is 50,000-fold more sensitive than the colorimetric assay. Sensitivity can also be enhanced by using assay conditions that minimize endogenous enzyme activity contributed by eukaryotic β -galactosidase (Young et al., 1993).

Secreted Alkaline Phosphatase (SEAP)

Secreted alkaline phosphatase (SEAP; Berger et al., 1988) differs from the abovementioned reporter proteins in that SEAP is secreted from transfected cells and can thus be assayed using a small aliquot of the culture medium. The SEAP gene encodes a truncated form of human placental alkaline phosphatase (PLAP), which lacks a critical membrane-anchoring domain, thereby allowing the protein to be efficiently secreted from transfected cells. Levels of SEAP activity detected in culture medium are directly proportional to changes in intracellular concentrations of SEAP mRNA and protein (Cullen and Malim, 1992). SEAP has the unusual properties of being extremely heat-stable and resistant to the phosphatase inhibitor L-homoarginine. Therefore, endogenous alkaline phosphatase activity can be eliminated by pretreatment of samples at 65°C and incubation with the inhibitor. The secreted nature of SEAP provides several advantages for its use as a genetic reporter: (1) preparation of cell lysates is not required; (2) the kinetics of gene expression can be easily studied by repeated collection of medium from the same cultures; (3) transfected cells are not disturbed during measurement of SEAP activity and remain intact for further investigations; (4) background from endogenous alkaline phosphatase activity in the culture medium is almost absent; and (5) sample collection and assay can be automated using 96-well microtiter plates.

The original SEAP assay was a colorimetric procedure using the alkaline phosphatase substrate *p*-nitrophenyl phosphate (Cullen and Malim, 1992). This procedure is fast, simple to perform, and very inexpensive. However, the assay's sensitivity is poor, and it has a narrow linear dynamic range. Sensitivity can be improved by using a two-step bioluminescent assay for SEAP based on hydrolysis of D-luciferin-*O*-phosphate (Berger et al., 1988). The dephosphorylation reaction catalyzed by SEAP yields free luciferin, which in turn serves as the substrate for firefly luciferase. The sensitivity

of this assay is roughly equivalent to that of the conventional bioluminescent assay for luciferase. The most sensitive SEAP assays use chemiluminescent alkaline phosphatase substrates such as 1,2-dioxetane CSPD (Bronstein et al., 1994; see Table 10.8.1). Dephosphorylation of CSPD results in a sustained "glow"-type luminescence that remains constant up to 60 minutes and is readily detected using a luminometer or scintillation counter. In addition to enhancing sensitivity, the chemiluminescence assay for SEAP greatly increases the linear dynamic range of detection.

Human Growth Hormone (hGH)

Human growth hormone (hGH) is normally secreted exclusively from the somatotrophic cells of the anterior pituitary gland. The restricted pattern of hGH expression makes this protein an attractive choice as a genetic reporter for most mammalian cell types. hGH reporter assays have many of the same advantages as SEAP (*UNIT 9.7A*; Selden et al., 1986), but are less desirable due to the relatively low sensitivity of the procedure and the need for a hazardous radioimmunoassay to quantitate hGH. Measurement of hGH has been used as an internal control to normalize transfection efficiency and is well suited in this regard to normalize expression from experimental SEAP reporter constructs.

β -Glucuronidase (GUS)

The bacterial β -glucuronidase (GUS) gene is the predominant reporter used to study gene expression in plants (Gallagher, 1992). GUS is used as a reporter in both plant and mammalian cells, but is particularly useful in higher plants due to the absence of endogenous GUS activity in most species. Higher plants transformed with GUS are healthy, develop normally, and are fertile. As with β -galactosidase, one of the principle advantages of GUS as a reporter is the wide diversity of assays available for the enzyme. Several different colorimetric assays for GUS have been developed using a variety of β -glucuronides as substrates. The most popular of these uses X-Gluc (Jefferson et al., 1987), a substrate that can also be used for histochemical staining of tissues and cells expressing GUS activity. A more sensitive fluorescence assay for GUS uses the substrate 4-MUG (Jefferson et al., 1987). A chemiluminescent assay for GUS has also been developed; this is very similar to the procedure used to quantify β -galactosidase, and employs an adamantyl 1,2-dioxetane aryl glucuronide sub-

Table 9.6.1 Genetic Reporter Systems

Reporter gene	In vitro assay	In vivo assay	Strengths	Weaknesses	Limit of detection (molecules) ^a	Cost/assay (U.S. dollars) ^b	Available reporter vectors ^c
Chloramphenicol acetyltransferase (CAT)	Chromatography, differential extraction, fluorescence, or immunoassay	None commonly available	Minimal endogenous activity in mammalian cells; stable protein; various assay formats available	Assays are time-consuming and laborious; most formats require expensive radioactive substrate; relatively low sensitivity and narrow linear range (2 orders of magnitude); fluorescent assays offer better sensitivity, but require a fluorometer for quantitation; short half-life of CAT mRNA	$5-10 \times 10^7$	2.00-7.00	pSV0CAT and pSV2CAT (Gorman et al., 1982); pA ₁₀ CAT2 (Laiminis et al., 1984); pCAT-Basic, pCAT-Enhancer, pCAT-Promoter, and pCAT-Control (Promega)
Firefly luciferase	Bioluminescence (luminometer or scintillation counter)	Bioluminescence assay in live cells with luciferin esters as substrate	Nonisotopic; good sensitivity and broad linear range (4 orders of magnitude); minimal endogenous activity in mammalian cells; relatively inexpensive	Short half-life of protein; assay requires a luminometer or scintillation counter; conventional assay lacks reproducibility	$1-2.5 \times 10^5$	0.10-0.25	pT8luc, pSluc2, pXP1/pXP2, and pOLUC (Nordeen, 1988) pMAMneo-LUC (Clontech) pGL2-Basic, pGL2-Enhancer, pGL2-Promoter, and pGL2-Control (Promega)
β -galactosidase	Colorimetric, fluorescence, chemiluminescence (luminometer or scintillation counter)	Histochemical staining with X-gal substrate; bioluminescence assay in live cells with fluorescein di- β -D-galactopyranoside (FDG)	Nonisotopic; various assay formats available for different applications; chemiluminescent assay is very sensitive, and has an extremely broad linear range (5-6 orders of magnitude); well suited as internal control for normalizing other reporters	Many cell types have high endogenous β -galactosidase activity; assays require fluorometer or luminometer	10^4-10^5 (chemiluminescence assay)	0.01-0.02	p β gal-Basic, p β gal-Enhancer, p β gal-Promoter, and p β gal-Control (Clontech) pSV- β -gal (Promega)
Secreted alkaline phosphatase (SEAP)	Colorimetric, bioluminescence, or chemiluminescence (luminometer or scintillation counter)	None commonly available	Nonisotopic; secreted reporter protein; various assay formats available; chemiluminescent assay is very sensitive, and has a broad linear range (4 orders of magnitude); useful for high throughput assays performed in 96-well microtiter plates	May not be suitable for cells expressing low levels of placental-type alkaline phosphatase (lung, testes, and cervix); may not be suitable if experimental design affects secretory capacity of target cells	10^4-10^5 (chemiluminescence assay)	0.01-0.04	pSEAP-Basic, pSEAP-Enhancer, pSEAP-Promoter, and pSEAP-Control; (Clontech) pBC12/RSV/SEAP and pBC12/HIV/SEAP (Berger et al., 1988)

continued

Table 9.6.1 Genetic Reporter Systems, continued

Reporter gene	In vitro assay	In vivo assay	Strengths	Weaknesses	Limit of detection (molecules) ^a	Cost/assay (U.S. dollars) ^b	Available reporter vectors ^c
Human growth hormone (hGH)	Radioimmunoassay	None commonly available	Secreted reporter protein; direct detection of protein levels; simple assay, easy to perform; high-throughput assay performed in 96-well microtiter plates; useful as internal control for normalizing other reporters (e.g., SEAP)	Radioactive assay requiring ¹²⁵ I-labeled antibody; no signal amplification as with activity assays; relatively low sensitivity and narrow linear range; expensive	3 × 10 ⁸	1.00-2.00	pXGH, pOGH, and pTKGH (Selden et al., 1986)
β-glucuronidase (GUS)	Colorimetric, fluorescence, or chemiluminescence (luminometer or scintillation counter)	Histochemical staining with X-Gluc substrate	Nonisotopic; various assay formats available for different applications; GUS protein is stable; chemiluminescent assay is very sensitive, and has an extremely broad linear range (6 orders of magnitude); predominant reporter used in plant genetic research	Assays with best sensitivity require fluorometer or luminometer (scintillation counter can substitute)	N/A	N/A	pBI101, pBI101.2, pBI101.3, pBI121, pBI221, and pGUSN358-S (Clontech)
Green fluorescent protein (GFP)	Fluorescence (UV light box or fluorescence imaging device)	Fluorescence (UV light box, fluorescence microscopy, or FACS analysis)	Reporter of gene expression and protein localization in live cells; fluorescence occurs in a species-independent fashion; fluorescence is an intrinsic property of the protein and does not require additional gene products, substrates, or cofactors; fluorescent signal is highly resistant to photobleaching; no apparent toxic effects of GFP expression in bacteria or eukaryotes	Signal intensity may be too weak for some applications; this limitation is largely overcome by brighter GFP mutants such as EGFP	N/A	Essentially free, no additional reagents necessary	pEGFP, pEGFP1, pEGFP-N1,2,3, and pEGFP-C1,2,3 (Clontech)

^{a,b,c}Much of this information was taken from Alam and Cook, 1990. The lists of available reporter vectors may not be complete. N/A: information not available.

strate (Bronstein et al., 1994). This assay is approximately 100-fold more sensitive than the fluorometric assay using 4-MUG, and has a wider linear dynamic range. In addition to the uses of GUS as a genetic reporter, GUS fusion proteins have been used extensively to study protein localization and transport in both plant and animal species (Gallagher, 1992).

IN VIVO REPORTER ASSAYS

In vivo reporter assays may be defined as those in which the reporter protein is detected in either live cells or tissues (e.g., of transgenic animals), or in cells or tissues that have been fixed for histochemical staining. The data provided by this approach is less quantitative than in vitro reporter systems, but provides important information regarding the cell-type specificity of promoter/enhancers and the tissue distribution of specific transcription factors. Because this method focuses on localization of reporter proteins to the site of their production, it is well suited to reporter assays that employ either precipitating substrates or substrates that generate a luminescent or fluorescent signal. Conversely, secreted proteins such as SEAP and hGH are of less use as in vivo reporter genes.

Green Fluorescent Protein (GFP)

Light is produced in several bioluminescent jellyfish as a result of energy transfer to green fluorescent proteins (GFPs). One such GFP, from *Aequorea victoria*, fluoresces in vivo after receiving energy from the Ca²⁺-activated photoprotein aequorin (Inouye and Tsuji, 1994). This process occurs without requirement for a substrate or cofactor and proceeds via direct transfer of energy between the two proteins. Purified GFP has similar spectral properties to the protein expressed in vivo, absorbing blue light and emitting green light that is detectable using a fluorescence microscope or a UV light box or by fluorescence-activated cell sorting (FACS). The absorption and emission wavelengths of GFP are similar to those of fluorescein, and conditions used for visualizing fluorescein are also suitable for GFP. Full-length GFP appears to be required for fluorescence; however, the minimal chromophore responsible for light emission can be assigned to a hexapeptide within the protein (Cody et al., 1993). This region of the protein contains a Ser-dehydroTyr-Gly trimer, which cyclizes to yield a chromophore that emits light by an as yet unknown mechanism.

Cloning and sequencing of the *A. victoria* GFP gene (Prasher et al., 1992) allowed expression of the protein in heterologous systems. GFP expressed in either prokaryotic or eukaryotic cells yields a bright green fluorescence when the cells are excited by blue or UV light. GFP fluorescence does not require additional gene products from *A. victoria* and occurs in a species-independent fashion. Recent studies using the nematode *Caenorhabditis elegans*, an organism often used for developmental studies, have demonstrated the utility of GFP as a reporter of gene expression in vivo (Chalfie et al., 1994). These researchers expressed GFP in *C. elegans* under the control of a neuron-specific promoter and used GFP fluorescence to monitor the formation of neuronal processes in real time as the worms developed. GFP has also been expressed in yeast, mammalian cells, and *Drosophila*. Moreover, both N- and C-terminal protein fusions with GFP have been constructed and shown to maintain the fluorescence properties of native GFP (Wang and Hazelrigg, 1994).

A critical aspect of GFP-based reporter systems is the ability to monitor transcriptional changes in real time. Future directions for improvement of GFP as a genetic reporter include efforts to increase the intensity of the fluorescent signal and thereby enhance detection sensitivity, and to reduce the variability of the signal between replicate samples. As with firefly luciferase, variant GFPs with shifted emission wavelengths may allow the use of this reporter to study two or more transcriptional events in the same cell population. The availability of multiple fluorescent proteins derived from wild-type GFP will not only facilitate the simultaneous detection of several transcription reporters, but also permit a similar analysis of protein trafficking events using GFP fusion proteins.

Firefly Luciferase

Recent studies have shown that firefly luciferase activity can be detected in live cells by means of soluble luciferase substrates capable of crossing the plasma membrane (Bronstein et al., 1994). These compounds are generally uncharged luciferin esters that readily penetrate and cross lipid bilayers. Once inside the cell, these substrates are converted by endogenous esterases into firefly luciferin, which in turn serves as the substrate for the firefly luciferase encoded by the reporter gene. A photolyzable "caged" luciferin compound has been described that appears to enter cells more efficiently. This

compound can be hydrolyzed to yield free luciferin either by endogenous esterases or by visible light (Yang and Thomason, 1993). In its present form, the *in vivo* luciferase assay is somewhat insensitive and lacks the capability for real time analysis provided by GFP. Through the development of alternate luciferase substrates having a higher quantum yield and more sophisticated detection instruments, it is possible that luciferase may develop into a useful reporter system for *in vivo* analyses.

β -Galactosidase

In vivo levels of the β -galactosidase reporter protein can be determined in prokaryotic and eukaryotic cells, tissue sections, and intact embryos using the precipitating substrate X-Gal (Alam and Cook, 1990). The reaction with X-Gal produces a rich blue color that can easily be scored against background in most applications. Histochemical staining with X-Gal requires fixation of the cells or tissue prior to the enzymatic assay. Detection of *lacZ* expression in whole embryos has been particularly useful for characterizing tissue-specific gene expression during early development. Detection of β -galactosidase activity in live cultured cells is achieved with the substrate fluorescein di- β -D-galactopyranoside (FDG; Bronstein, 1994). FDG is delivered into live cells by hypotonic loading, and after cleavage by β -galactosidase is trapped inside the cell due to the hydrophobic structure of the reaction product. Cells expressing the reporter protein are detected via fluorescence from the fluorescein moiety of the metabolized substrate.

Another important application of *in vivo* expression of β -galactosidase are so called "enhancer trap" studies, in which positional effects of reporter integration into the chromosomes of transgenic organisms is used as a means of locating enhancer sequences (Alam and Cook, 1990). In one approach, a vector containing a weak basal promoter cloned directly upstream of the *lacZ* gene is introduced into mouse eggs and integrates into the mouse genome. Stable transgenic animals are bred from the progeny of the resulting adult animals. A sufficiently weak promoter is chosen such that expression of β -galactosidase occurs only when the promoter is proximal to an enhancer sequence near the site of integration. The expression pattern of β -galactosidase in transgenic lines indirectly reflects both the tissue specificity and strength of nearby enhancer(s). Experiments of this type utilize expression of a reporter to identify novel genes and regulatory flanking sequences that

have potentially interesting patterns of expression during early development.

LITERATURE CITED

- Alam, J. and Cook, J.L. 1990. Reporter genes: Application to the study of mammalian gene transcription. *Anal. Biochem.* 188:245-254.
- Araki, E., Shimada, F., Shichiri, M., Mori, M., and Ebina, Y. 1988. pSV00CAT: Low background CAT plasmid. *Nucl. Acids Res.* 16:1627-1630.
- Berger, J., Hauber, J., Hauber, R., Geiger, R., and Cullen, B.R. 1988. Secreted placental alkaline phosphatase: A powerful new quantitative indicator of gene expression in eukaryotic cells. *Gene* 66:1-10.
- Bronstein, I., Fortin, J., Stanley, P.E., Stewart, G.S., and Kricka, L.J. 1994. Chemiluminescent and bioluminescent reporter gene assays. *Anal. Biochem.* 219:169-181.
- Cassinotti, P. and Weitz, M. 1994. Increasing the sensitivity of a common CAT assay. *BioTechniques* 17:36-38.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. 1994. Green fluorescent protein as a marker for gene expression. *Science* 263:802-805.
- Cody, C.W., Prasher, D.C., Westler, W.M., Prendergast, F.G., and Ward, W.W. 1993. Chemical structure of the hexapeptide chromophore of the *Aequorea* green fluorescent protein. *Biochemistry* 32:1212-1218.
- Cullen, B.R. and Malim, M.H. 1992. Secreted placental alkaline phosphatase as a eukaryotic reporter gene. *Methods Enzymol.* 216:362-368.
- De Wet, J.R., Wood, K.V., DeLuca, M., Helinski, D.R., and Subramani, S. 1987. Firefly luciferase gene: Structure and expression in mammalian cells. *Mol. Cell. Biol.* 7:725-737.
- Gallagher, S.R. 1992. GUS Protocols: Using the GUS Gene as a Reporter of Gene Expression. Academic Press, San Diego, Calif.
- Gorman, C.M., Moffat, L.F., and Howard, B.H. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* 2:1044-1051.
- Inouye, S. and Tsuji, F.I. 1994. *Aequorea* green fluorescent protein: Expression of the gene and fluorescence characteristics of the recombinant protein. *FEBS Lett.* 341:277-280.
- Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. 1987. GUS fusions: β -glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* 6:3901-3907.
- Laiminis, L.A., Gruss, P., Pozzatti, R., and Khoury, G. 1984. Characterization of enhancer elements in the long terminal repeat of Moloney murine sarcoma virus. *J. Virol.* 49:183-189.
- Nguyen, V.T., Morange, M., and Bensaude, O. 1988. Firefly luciferase luminescence assays using scintillation counters for quantitation in transfected mammalian cells. *Anal. Biochem.* 171:404-408.

- Nordeen, S.K. 1988. Luciferase reporter gene vectors for analysis of promoters and enhancers. *BioTechniques* 6:454-457.
- Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G., and Cormier, M.J. 1992. Primary structure of the *Aequora victoria* green fluorescent protein. *Gene* 111:229-233.
- Selden, R.F., Howie, K.B., Rowe, M.E., Goodman, H.M., and Moore, D.D. 1986. Human growth hormone as a reporter gene in regulation studies employing transient gene expression. *Mol. Cell. Biol.* 6:3173-3179.
- Thompson, J.F., Hayes, L.S., and Lloyd, D.B. 1993. Modulation of firefly luciferase stability and impact on studies of gene regulation. *Gene* 103:171-177.
- Wang, S. and Hazelrigg, T. 1994. Implications for bcd mRNA localization from spatial distribution of exu protein in *Drosophila* oogenesis. *Nature* 369:400-403.
- Wenger, R.H., Moreau, H., and Nielson, P.J. 1994. A comparison of different promoter, enhancer, and cell type combinations in transient transfections. *Anal. Biochem.* 221:416-418.
- Yang, J. and Thomason, D.B. 1993. An easily synthesized, photolyzable luciferase substrate for in vivo luciferase activity measurement. *BioTechniques* 15:848-850.
- Young, D.C., Kingsley, S.D., Ryan, K.A., and Dutko, F.J. 1993. Selective inactivation of eukaryotic β -galactosidase in assays for inhibitors of HIV-1 TAT using bacterial β -galactosidase as a reporter enzyme. *Anal. Biochem.* 215:24-30.

Contributed by Steven R. Kain (overview and table)

Clontech Laboratories, Inc.
Palo Alto, California

Subinay Ganguly (table)
SmithKline Beecham
King of Prussia, Pennsylvania

Isotopic Assays for Reporter Gene Activity

UNIT 9.7A

This unit describes two widely used reporter systems that are based on radioactive detection assays. The first assay uses chloramphenicol acetyltransferase (CAT) activity as a measure of the level of expression of a transfected gene. This bacterial enzyme catalyzes the transfer of an acyl group from acetyl CoA (or any of several other acyl CoA cofactors) to chloramphenicol. In the assays described here, transfected cells are harvested and lysed by freeze-thaw steps (first basic protocol) or simply lysed in the tissue culture plates using Triton X-100 (first alternate protocol). Acyl CoA and radioactively labeled chloramphenicol are added to cell lysate, and modified derivatives of the antibiotic are separated from the starting material using either thin-layer chromatography (TLC; basic protocol) or phase-extraction (second alternate protocol). The chromatographic approach, which uses ^{14}C -labeled chloramphenicol, is the most frequently encountered. The phase-extraction assay, which uses ^3H -labeled chloramphenicol, is substantially cheaper and easier. The second reporter system (second basic protocol) uses a kit to perform a simple two-site radioimmunoassay to quantitate the amount of human growth hormone (hGH) secreted into culture medium by transfected cells. Medium is incubated with ^{125}I -labeled antibody specific for hGH, and immune complexes are collected by an avidin-coated bead. The quantity of hormone is determined based on comparison with a standard curve. *UNIT 9.6* provides an overview of reporter gene systems.

CHROMATOGRAPHIC ASSAY FOR CAT ACTIVITY

Cells are harvested 48 to 72 hr after transfection with chloramphenicol acetyltransferase (CAT) expression plasmids (Fig. 9.7.1), and a lysate is prepared by freeze-thaw cycles. A cytoplasmic extract of the lysate is incubated with radioactive chloramphenicol and acetyl CoA. Acetylated chloramphenicol is separated from the nonacetylated form by thin-layer chromatography (TLC) and autoradiographed.

Materials

Cells transfected with CAT expression plasmid (*UNITS 9.1-9.4 & UNIT 16.12*; Fig. 9.7.1), in 100-mm petri plates

Phosphate-buffered saline (PBS; *APPENDIX 2*)

TEN solution (see recipe)

1 M and ice-cold 0.25 M Tris·Cl, pH 7.5 (*APPENDIX 2*)

200 $\mu\text{Ci/ml}$ [^{14}C]chloramphenicol (35 to 55 mCi/mmol)

4 mM acetyl CoA (store ≤ 2 weeks at -20°C)

Ethyl acetate

19:1 (v/v) chloroform/methanol

Rubber policeman or equivalent

Thin-layer chromatography (TLC) tank

Whatman 3MM filter paper

Thin-layer chromatography (TLC) sheets (plastic-backed, silica gel 1B; J.T. Baker)

Pen or marker with radioactive ink

Additional reagents and equipment for autoradiography (*APPENDIX 3*)

Prepare extracts by freeze-thaw lysis

1. Wash 100-mm plate of cells transfected with CAT expression plasmid twice with 5 ml PBS per wash. Add 1 ml TEN solution to each plate. Let cells sit 5 min on ice.

TEN solution helps adherent cells lift off the plate. If cells in suspension are used, simply wash the cells once with PBS and once with TEN solution, collecting each time by centrifugation, and proceed to step 3.

**BASIC
PROTOCOL 1**

**Introduction of
DNA into
Mammalian Cells**

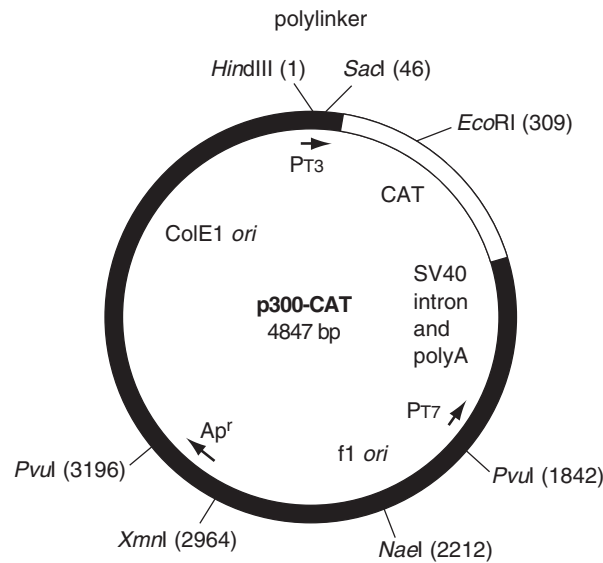
9.7.1

Contributed by Robert E. Kingston, Jen Sheen, and David Moore

Current Protocols in Molecular Biology (1995) 9.7.1-9.7.11

Copyright © 2003 by John Wiley & Sons, Inc.

Supplement 63



polylinker :

```

HindIII   SphI   PstI   SalI   XbaI   BamHI   SmaI   KpnI   SacI
1  AAGCTTGCAT GCTGCAGGT CGACTCTAGA GGATCCCCGG GTACCGAGCT CGAATTGGGA
      |
      start CAT
61  GCTTGGCGAG ATTTTCAGGA GCTAAGGAAG CTAATG GAG AAA AAA ATC ACT

```

Figure 9.7.1 Expression vector p300-CAT. This vector is designed for insertion of the desired control elements. It was constructed by cutting the CAT cassette out of pUC-CAT (Gilman et al., 1986) and blunt-end coning it into the *EcoRI* site of pBluescript M13⁻. Every restriction endonuclease site in the polylinker, which is pUC-like, is unique and therefore available for cloning. The vector is a phagemid that produces the strand that permits sequencing of promoter inserts in the 5' to 3' direction using the reverse primer. In addition, the design of mutagenic oligonucleotides is facilitated by use of the sequence of the top strand of a double-stranded sequence.

2. Scrape the cells off the plate with a rubber policeman. Transfer them to a 1.5-ml microcentrifuge tube on ice.
3. Microcentrifuge cells 1 min at maximum speed, 4°C.

Do not centrifuge cells too long or they will form a pellet that is difficult to resuspend.

4. Resuspend cell pellet in 100 µl ice-cold 0.25 M Tris·Cl, pH 7.5.

One hundred microliters of solution is sufficient for ~10⁷ cells.

5. Freeze cells 5 min in dry ice/ethanol. Transfer to 37°C and thaw 5 min. Repeat this freeze-thaw process twice more, for a total of three freeze-thaw cycles.

Successive freezing and thawing lyses the cells.

6. Cool cell lysate on ice, then microcentrifuge 5 min at maximum speed, 4°C. Remove and save the supernatant.

The supernatant is a cytoplasmic extract and contains the CAT enzyme. The supernatant can be stored by freezing in dry ice/ethanol and storing at -20°C. It is important to freeze the supernatant rapidly to maintain full activity.

Assay the extract

7. To assay 20 μl of cell extract, make the following cocktail (130 μl per reaction):

2 μl 200 $\mu\text{Ci/ml}$ [^{14}C]chloramphenicol (35 to 55 mCi/mmol)
20 μl 4 mM acetyl CoA
32.5 μl 1 M Tris·Cl, pH 7.5
75.5 μl H₂O.

One to fifty microliters of extract can be assayed. When assaying different amounts of extract, adjust the cocktail so that the final concentration of Tris·Cl is 0.25 M and the final volume of each reaction is 150 μl .

8. For each assay, add 130 μl cocktail and 20 μl extract to a microcentrifuge tube and mix gently. Incubate 1 hr at 37°C.

Assay time can be extended to several hours if very little activity is present. If a long assay is performed, replace the 4 mM acetyl CoA with 40 mM acetyl CoA because acetyl CoA is not stable under the assay conditions. It is critical that the assay be done in the linear time range. This can be verified by setting up a 600- μl reaction, removing 145- μl aliquots at various time points, and analyzing conversion to acetylated chloramphenicol, as in steps 9 to 14. A plot of activity versus time must be linear through the time point chosen for routine assay for the assay to be valid.

9. Add 1 ml ethyl acetate to the reaction and vortex. Microcentrifuge 1 min and remove the top (ethyl acetate) layer.

The chloramphenicol and acetylated chloramphenicol are extracted into the ethyl acetate.

10. Dry the ethyl acetate 45 min in a Speedvac evaporator.

Alternatively, the ethyl acetate can be allowed to evaporate overnight in a fume hood. This sometimes leads to extra spots on the final TLC because of breakdown products.

Perform thin-layer chromatography

11. Equilibrate the thin-layer chromatography tank: add 190 ml chloroform, 10 ml methanol, and a piece of Whatmann 3MM filter paper approximately the size of the TLC sheet to the tank and let stand 2 hr.

The ratio of methanol to chloroform has a major effect on the separation. Methanol can be lost quickly by evaporation, so it is best to use the solvent mixture for only 24 hr or less.

12. Resuspend each sample in 30 μl ethyl acetate.
13. Spot sample, 5 μl at a time, 2 cm above the edge of a plastic-backed TLC sheet.
14. Develop the chromatogram in an equilibrated chromatography tank containing 200 ml of 19:1 chloroform/methanol.
15. Allow the chromatography to run 2 hr or until the solvent front is close to the top of the sheet. Remove the TLC sheet and air dry. Mark the TLC sheet with radioactive ink, cover with plastic wrap, and place on film for autoradiography.

The final autoradiogram will have up to five spots for each sample. They are, in ascending order, a weak spot at the origin, nonacetylated chloramphenicol, the two forms of acetylated chloramphenicol, and diacetylated chloramphenicol (Fig. 9.7.2). If the diacetylated spot is present, the assay is out of the linear range (conversion to acetyl chloramphenicol $\geq 20\%$ to 30%). If this occurs, dilute sample or reduce assay time.

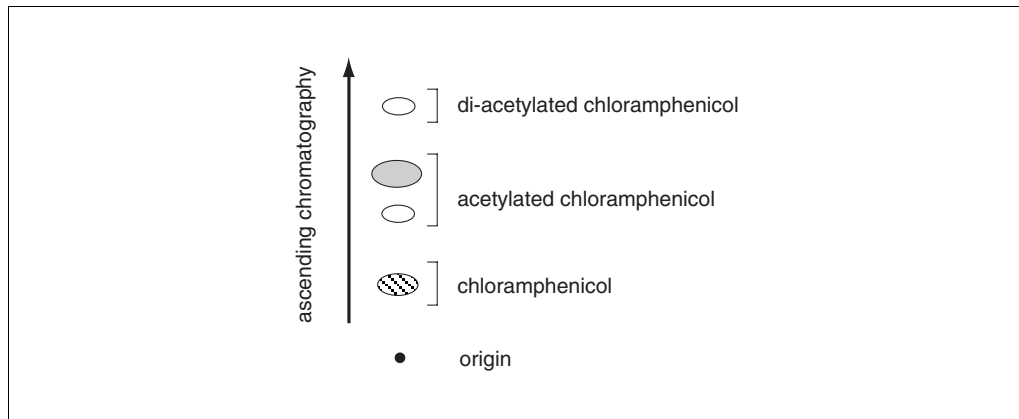


Figure 9.7.2 Schematic depiction of results of a CAT assay.

16. Calculate the activity of the extract by first determining the percentage of counts that are in the monoacetylated chloramphenicol species. To determine the amount of label in each spot, align the chromatogram and the autoradiogram using a light box and the marks made with radioactive ink. Outline the area of each spot on the chromatogram with a pencil, cut out each spot, add to scintillation fluid in a scintillation vial, and count in a scintillation counter. Calculate CAT activity as follows:

$$\% \text{ acetylated} = \frac{\text{counts in acetylated species}}{\text{counts in acetylated species} + \text{counts in nonacetylated chloramphenicol}}$$

ALTERNATE PROTOCOL 1

IN SITU LYSIS OF CELLS FOR CAT ASSAY

This protocol—in situ lysis of cells transfected with CAT expression plasmid, based on an approach developed by Brian Seed of Massachusetts General Hospital—is faster and easier than the freeze-and-thaw lysis in the basic protocol, but it may lead to somewhat more variable results.

Additional Materials (also see *Basic Protocol 1*)

Cells transfected with CAT expression plasmid (*UNITS 9.1-9.4 & 16.13*), in 60-mm petri plates
 Hypotonic buffer (see recipe)
 Triton lysis buffer (see recipe)

1. Wash 60-mm plate of cells transfected with CAT expression plasmid once with 2 ml PBS.
2. Add 2 ml hypotonic buffer per plate. Incubate at room temperature 2 to 5 min.
During this incubation the cells will swell noticeably.
3. Aspirate hypotonic buffer and add 400 μ l Triton lysis buffer.
Cells will lyse almost immediately after they come in contact with Triton lysis buffer.
4. Scrape the plate with a rubber policeman and transfer lysate to a 1.5-ml microcentrifuge tube using a 1000- μ l pipettor.
5. Microcentrifuge 1 min to remove nuclei and insoluble proteins.
6. Transfer supernatant to a clean microcentrifuge tube and proceed with the assay for CAT activity (first basic protocol, step 7 or second alternate protocol, step 5).

PHASE-EXTRACTION ASSAY FOR CAT ACTIVITY

ALTERNATE
PROTOCOL 2

Extracts of cells transfected with CAT expression plasmid are incubated with ^3H - or ^{14}C -labeled chloramphenicol and unlabeled butyryl CoA. Butyrylated chloramphenicol is separated from the unmodified form by a simple phase extraction. The assay is simple, fast (hours versus days), and inexpensive (50 cents versus 7 to 8 dollars per sample with ^3H -labeled chloramphenicol).

Additional Materials (also see Basic Protocol 1)

- Mammalian cells transfected with CAT expression plasmid (UNITS 9.1-9.4 & 16.13) or protoplasts transfected with CAT expression plasmid (UNIT 9.3)
- 0.01 $\mu\text{Ci}/\mu\text{l}$ [^3H]chloramphenicol solution (see recipe)
- 5 mg/ml butyryl CoA (store ≤ 4 months at -20°C)
- 100 mg/ml unlabeled chloramphenicol
- 2 M Tris-Cl, pH 8.0 (APPENDIX 2)
- 2:1 (v/v) tetramethylpentadecane (TMPD)/xylenes

Prepare cell extracts

From mammalian cells:

- 1a. To prepare cell extracts from mammalian cells, follow either the freeze-thaw method in steps 1 to 6 of the first basic protocol or the in situ lysis method in the first alternate protocol. If cells are harvested using trypsin or by scraping, replace steps 1 to 4 of the in situ lysis method with steps 2a to 4a below.
- 2a. Spin cells in a clinical centrifuge 5 min at $300 \times g$, room temperature, and discard supernatant.
- 3a. Add 2 ml hypotonic buffer per 10^7 cells and repeat step 2a.
- 4a. Add 200 μl Triton lysis buffer per 10^7 cells.

From plant protoplasts:

- 1b. Harvest plant protoplasts (10^5 cells) in 1.5-ml microcentrifuge tubes by microcentrifuging 5 min at 2000 to 4000 rpm, room temperature. Discard supernatant.
- 2b. Add 50 μl hypotonic buffer to the pellet and vortex vigorously.
- 3b. Freeze the sample by incubating the tube ≥ 15 min at -70°C or 5 min in a dry ice/ethanol bath.
- 4b. Thaw and microcentrifuge 2 min at 13,000 rpm, room temperature. Transfer supernatant (containing the CAT extract) to a clean tube.

It is sometimes necessary to heat the sample 10 min at 65°C to inactivate potential inhibitors (see Critical Parameters).

Assay cell extract

5. To assay 50 μl cell extract, make the following CAT assay mix (50 μl per reaction):
 - 20 μl 0.01 $\mu\text{Ci}/\mu\text{l}$ [^3H]chloramphenicol solution
 - 5 μl 5 mg/ml butyryl CoA
 - 5 μl 2 M Tris-Cl, pH 8.0
 - 20 μl H_2O .

For each reaction, use 0.2 μCi [^3H]chloramphenicol. If [^{14}C]chloramphenicol is used as substrate, also use 0.2 μCi /reaction.

Introduction of
DNA into
Mammalian Cells

9.7.5

6. For each assay, add 50 μ l CAT assay mix to 50 μ l cell extract. Incubate 30 to 90 min at 37°C.

The range of linearity is between 0.01% and 50% conversion, making this an excellent assay for measuring the activity of very weak promoters. For measuring the activity of strong promoters that give 30% to 100% conversion, incubate <90 min to obtain data in the linear range, or add additional substrates (unlabeled chloramphenicol and butyryl CoA) to the CAT assay mix.

7. Extract acylated chloramphenicol with 200 μ l of 2:1 TMPD/xylenes by vigorous shaking. Centrifuge and remove the top (organic) phase to a scintillation vial.

A 2:1 mixture of TMPD/xylenes provides both efficient extraction and a low background (see Critical Parameters). Alternatively, samples may be extracted with 200 μ l of 100% xylenes. If xylenes are used, it is necessary to back extract twice to reduce the extremely high background before transferring the organic phase to a scintillation vial. Back-extract the sample by adding 100 μ l water or TE buffer to 200 μ l of xylenes that contain acylated chloramphenicol and shaking vigorously. Centrifuge, remove the top xylenes phase to a new tube, and repeat.

8. Add 3 to 5 ml scintillation fluid to the sample and count to determine CAT activity.

BASIC PROTOCOL 2

RADIOIMMUNOASSAY FOR HUMAN GROWTH HORMONE

Growth hormone secreted into culture medium by transfected cells is measured by a simple two-site radioimmunoassay using a kit.

Materials

Cells transfected with hGH expression plasmid (*UNITS 9.1-9.4 & UNIT 16.12*; see Fig. 9.7.3)

Human Growth Hormone Radioimmunoassay Kit (Allegro Human Growth Hormone, Nichols Institute Diagnostics) containing:

¹²⁵I-labeled antibody solution

Wash solution

Human growth hormone (hGH) standards

Avidin-coated beads

12 \times 75-mm round-bottom test tube

γ counter

1. Remove 100 to 500 μ l medium from mammalian cells transfected with hGH expression plasmid (Fig. 9.7.3).

hGH is secreted by virtually all mammalian cells after transfection with the appropriate vector. In medium, hGH is stable several months at -20°C and several days at 4°C. Avoid repeated thawing and freezing of samples.

2. Pipet 100 μ l medium or standards into a 12 \times 75-mm round-bottom test tube.
3. Add 100 μ l of ¹²⁵I-labeled antibody solution to the tube and mix.
4. Add one avidin-coated bead and cap the tube or cover it with Parafilm. Incubate 90 min at room temperature on a horizontal rotating shaker set at ~170 rpm.

Alternatively, the assay can be done in a stationary tube with a 4-hr incubation.

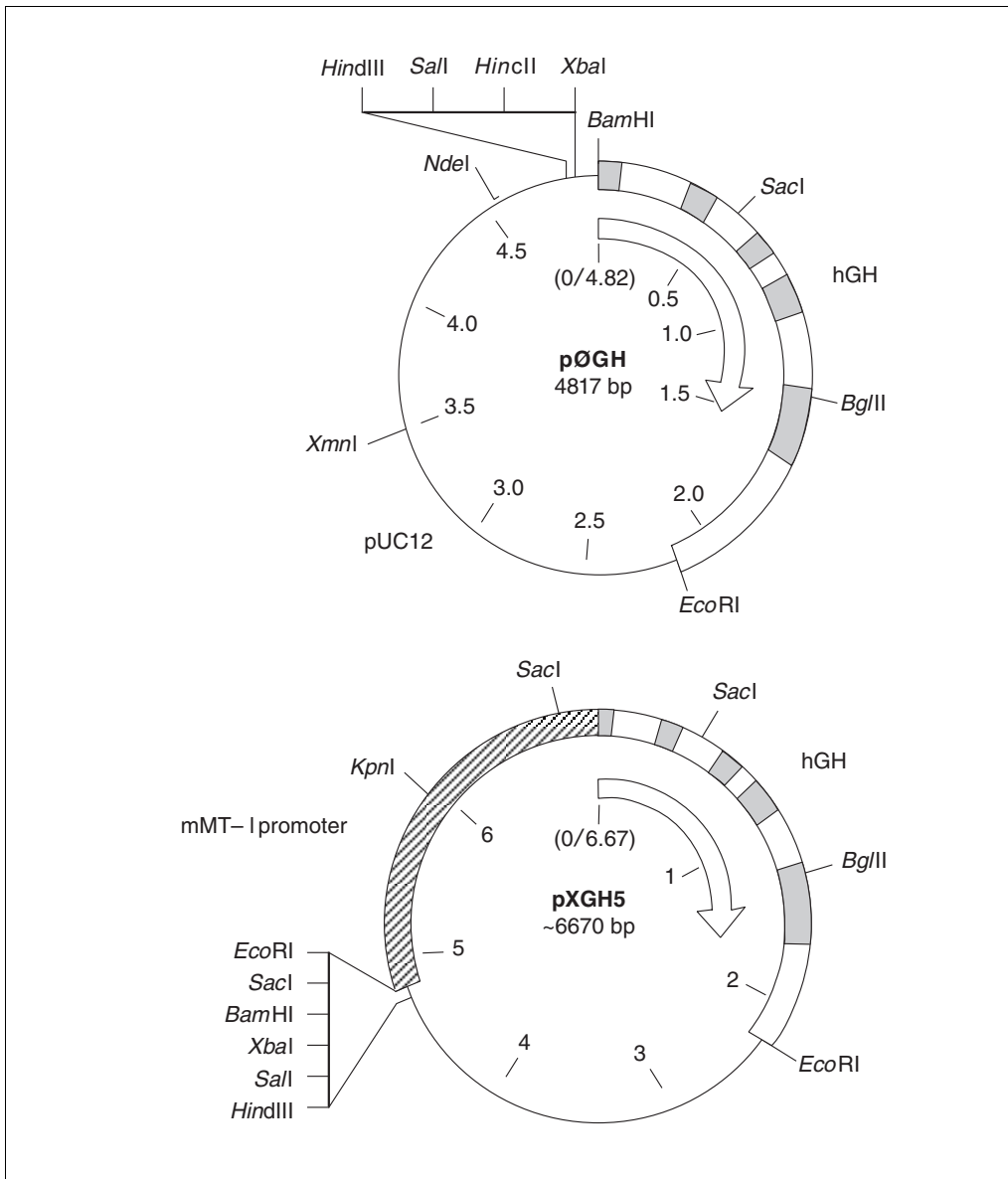


Figure 9.7.3 Expression vectors for human growth hormone (hGH; Selden et al., 1986).

5. Wash the bead twice by adding 2 ml wash solution and aspirating completely.

The wash solution should be added rapidly so that the bead is raised from the bottom of the tube.

6. Count the tube in a γ counter 1 min.
7. Plot a standard curve with the data from the hGH standards supplied with growth hormone assay kit. Calculate the values of the unknowns using this standard curve.

The assay is not linear for values >50 ng/ml. Samples with values >50 ng/ml should be diluted and reassayed.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

[³H]chloramphenicol solution

Prepare a 0.2 μCi/μl [³H]chloramphenicol stock by adding 960 μl of 100% ethanol and 40 μl of 100 mg/ml unlabeled chloramphenicol to 250 μl of [³H]chloramphenicol (250 μCi/250 μl in ethanol; 42.0 to 58.2 Ci/mmol).

Preextract the 0.2 μCi/μl [³H]chloramphenicol stock by first diluting it 20-fold in water, and then extracting this mixture with an equal volume of xylenes. Shake vigorously. Microcentrifuge 2 min at 13,000 rpm, at room temperature in a clinical centrifuge or microcentrifuge at maximum speed to separate the phases, and discard the top xylenes phase. Extract the aqueous phase one more time, centrifuge, and discard the top xylenes phase. This creates a working solution of 0.01 μCi/μl [³H]chloramphenicol.

When using ³H-labeled instead of ¹⁴C-labeled chloramphenicol, preextraction with xylenes is necessary to reduce the background. Preextraction of [¹⁴C]chloramphenicol reduces background, but not to the same extent as with [³H]chloramphenicol.

Hypotonic buffer

25 mM Tris·Cl, pH 7.5 (APPENDIX 2)

2 mM MgCl₂

Store up to 6 months at room temperature

TEN (Tris/EDTA/NaCl) solution

40 mM Tris·Cl, pH 7.5 (APPENDIX 2)

1 mM EDTA, pH 8.0 (APPENDIX 2)

150 mM NaCl

Store up to 6 months at room temperature

Triton lysis buffer

0.25 M Tris·Cl, pH 7.8 (APPENDIX 2)

0.5% (v/v) Triton X-100

Store up to 6 months at 4°C

COMMENTARY

Background Information

CAT reporter gene assay

Chloramphenicol acetyltransferase (CAT) is a bacterial drug-resistance gene that inactivates chloramphenicol by acetylating the drug at one or both of its two hydroxyl groups. This gene is not found in eukaryotes; thus, there is no background activity in a normal mammalian cell. This characteristic, along with the ease and sensitivity of the assay for CAT activity, has made the CAT gene the most widely used reporter gene for studies of mammalian gene expression. Another commonly encountered reporter system uses firefly luciferase. That assay, which is more sensitive than the CAT assay, is described in UNIT 9.7B.

The CAT gene was first adapted for use in mammalian systems by Gorman et al. (1982).

The bacterial gene was fused to SV40 splice and polyadenylation signals to create a form of the gene that is expressed in mammalian cells. The original vectors contained either a *Hind*III or *Sma*I site for insertion of the promoter. Modified vectors, such as pUC-CAT, have been made that contain a polylinker for promoter insertion (Fig. 9.7.1; Gilman et al., 1986; Prost and Moore, 1986).

The chief advantage of using the CAT gene as a reporter in mammalian cells is that it can be very easily and sensitively detected. It is possible to detect CAT activity driven by a promoter whose expression is several orders of magnitude too weak to be detected directly by RNA analysis. This can create a problem in interpretation for the investigator, as it is possible that the level of CAT expression from a fusion gene reflects the level of a message that

is not originating at the promoter under study but is instead originating from an upstream promoter. Thus, without being able to analyze directly the structure of the CAT message, it is not possible to state unequivocally that the level of CAT enzyme reflects appropriately initiated message from the promoter under study.

There is a wide body of anecdotal evidence to suggest that the CAT mRNA is present at low levels and thus is unusually difficult to analyze directly. It is clearly possible, however, to detect appropriately initiated transcription from mammalian promoters fused to CAT genes by S1 analysis as early as 30 hr after introducing a nonreplicating plasmid into cells by transfection procedures. It remains possible, if not likely, that fusion of a test promoter to a mammalian fusion gene, such as human growth hormone (hGH), may create a more easily analyzed fusion RNA. As the ability to determine whether the level of the reporter gene actually reflects the amount of appropriately initiated message is crucial to the long-term success of most projects, it is worthwhile to characterize the RNA made from any fusion gene early on.

The first basic protocol uses an assay that requires [¹⁴C]chloramphenicol and thin-layer chromatography, making this a costly assay. The phase-extraction assay (Seed and Sheen, 1988) presented in the second alternate protocol is much less expensive, and exploits the relatively low specificity of CAT for the acyl donor, allowing butyryl CoA to be substituted for acetyl CoA. The greater hydrophobicity of butyrylated chloramphenicol allows an effective discrimination between free and acylated chloramphenicol by a simple phase extraction. The assay permits [³H]chloramphenicol to be substituted for the 10-fold more expensive [¹⁴C]chloramphenicol and affords substantial savings compared to alternatives currently available. Preextraction of the radiolabeled chloramphenicol before the reaction and either the use of TMPD/xylenes or back-extraction of the xylenes phase after the reaction allow a significant increase in sensitivity (25-fold) by reducing the background. The assay is linear with enzyme concentration over two to three orders of magnitude, without dilution of the enzyme or adjustment of the duration of incubation.

Human growth hormone reporter gene assay

The human growth hormone (hGH) transient assay system (Selden et al., 1986) is based on a simple immunological detection of hGH

secreted by the transfected cells. Human growth hormone is a 191-amino acid protein synthesized and secreted by the somatotroph cells of the anterior pituitary. The hGH gene (Fig. 9.7.3; DeNoto et al., 1981) contains five exons and is the best characterized of the five members of the hGH gene family.

Secretion of hGH into the medium begins ~24 hr posttransfection, but the time of onset varies somewhat depending on the cell type. In several cell lines, levels of secreted hGH have been shown to be proportional to levels of cytoplasmic hGH mRNA. Thus, secretion does not appear to be the rate-limiting step for appearance of hGH in the medium. Several characteristics of hGH make it useful as a reporter protein.

First, the medium of transfected cells is assayed for secreted hGH. The transfected cells are not destroyed, so this allows continuous monitoring of transient expression from the same cell population.

Second, hGH messenger RNA is quite stable in most mammalian cells. Changes in level of the reporter protein can therefore be correlated with changes in the level of hGH mRNA.

Third, the hGH assay is rapid, sensitive, and easy to perform. All cell types tested so far (including AtT-20, CV-1, HeLa, BALB/c 3T3, REF-52, rat embryo fibroblasts, human foreskin fibroblasts, GC, GH⁴, H35, JEG, L, and primary pituitary cells) were able to secrete hGH into the medium.

These properties suggest several related applications for the hGH transient systems:

1. Use of hGH as a reporter gene to analyze gene expression, analogous to the use of the CAT assay system. Vectors are available that contain a polylinker immediately adjacent to the hGH gene for insertion of eukaryotic promoters (pØGH; Fig. 9.7.3; Selden et al., 1986).
2. Use of a well-characterized hGH fusion gene, such as pXGH5 (Fig. 9.7.3; Selden et al., 1986), to characterize or optimize transfection efficiency (introduction to Section I).
3. Use of a well-characterized hGH fusion gene as an internal control for normalizing the efficiency of expression between different transfections. Transfection efficiency can vary dramatically between dishes in an individual experiment. Thus, it is valuable to transfect a constant amount of an internal standard plasmid in each transfection and use the level of expression from that plasmid to normalize for transfection efficiency. For example, an hGH fusion gene can be used for normalization when

analyzing a set of mutant promoters that have been fused to CAT.

A fundamental difference between the hGH transient assay system and other systems such as CAT or β -galactosidase is that hGH is a mammalian gene. This may contribute to the apparent stability of the hGH message, a valuable characteristic in analyzing promoter expression. It should be noted, however, that the sequences encoding the gene may not be entirely "silent" in terms of regulatory signals. In particular, the hGH gene encodes sequences that can allow response to glucocorticoids in certain fusion genes (Selden et al., 1986).

Critical Parameters

It is essential that the assay for CAT activity be within the linear range both for time (step 8 of the first basic protocol and step 6 of the second alternate protocol) and for activity. When conversion of chloramphenicol to acetyl chloramphenicol is >20% to 30%, the assay ceases to be linear. For example, a sample giving 90% conversion contains greater than nine times more activity than a sample giving 10% conversion. Thus, the sample giving 90% conversion must be diluted or assayed for a shorter time period to give an accurate measure of CAT activity.

Preparation of plant protoplasts using the second alternate protocol varies for the type of plant used. Some investigators have found that certain plant extracts (e.g., *Arabidopsis*) contain inhibitors of the CAT enzyme. A comparison of CAT activities from matched extracts that have been either heated 10 min at 65°C or not heated will indicate whether this heat inactivation step is needed.

Xylenes efficiently extract the acylated chloramphenicol (second alternate protocol) but also extract some of the nonacylated chloramphenicol, leading to a high background. While tetramethylpentadecane (TMPD) does not extract the nonacylated chloramphenicol, it does not efficiently extract the acylated form either. When xylenes alone are used to extract the acylated chloramphenicol, back-extraction of the xylenes phase can be performed to reduce the background before counting. The combination of preextraction and back-extraction (when xylenes alone are used) can increase the sensitivity of the assay ~25-fold, as compared to a single extraction with TMPD/xylenes.

As with any assay, results for the hGH radioimmunoassay must fall within the linear

range of the assay to be meaningful. The standards supplied with the growth hormone kit used in this protocol range from 0.5 to 50 ng/ml growth hormone. Usually the assay is linear within this range. The standards should be used in every growth hormone assay to ensure that the values obtained experimentally fall within the linear range.

Troubleshooting

There are three possible reasons why an extract might give no CAT activity. First, the fusion construct may simply be inactive in the cell type being transfected. Plasmids such as pSV2-CAT (Gorman et al., 1982), which are known to efficiently produce CAT, can be used as controls. Second, a bad reagent may affect the assay, although this rarely happens. This problem can be assessed by using an extract previously known to have good activity, or by using a commercial CAT preparation (e.g., Promega) as a positive control. Finally, the most common reason for a lack of CAT activity is poor transfection efficiency. Transfection solutions can be remade, or transfection efficiency can be tested using an expression vector known to be functional.

There are reports that extracts from some cell lines contain uncharacterized inhibitors of CAT that can be inactivated by heating 20 min at 70°C. This will generate a substantial amount of precipitated denatured protein that should be removed by microcentrifuging 1 min. The fully active CAT enzyme remains in the supernatant.

Anticipated Results

The amount of CAT enzyme produced in a transfection varies widely with transfection efficiency and the particular fusion construct. It is necessary to empirically establish standard conditions for various combinations of promoters and cell types. The chromatographic and phase-extraction assays should be approximately equal in sensitivity. The lower limit of this assay is 0.2 to 1.0×10^{-6} U; there is no upper limit because cell extracts with higher activity can be diluted and reassayed. 50% conversion of [3 H]chloramphenicol should yield 100,000 to 124,000 cpm.

The amount of hGH produced after a transfection will vary according to the fusion construct used and the cell type transfected. Transfection of pXGH5 into HeLa or BALB/c 3T3 cells by CaPO₄ transfection should produce levels near 50 ng/ml in the medium 48 hr after transfection.

Time Considerations

Harvest and preparation of extracts from twelve samples for the CAT assay with the basic freeze-thaw procedure takes 1 to 2 hr. In situ lysis takes ~30 min. With most transfection systems, a 1-hr incubation time is sufficient to detect CAT activity. If a longer incubation time is used, the amount of acetyl CoA in the reaction mix should be increased 10-fold. Thin-layer chromatography takes 2 hr to set up and after completing the incubation, 3 to 4 hr to run the chromatogram. A signal should be detectable on a 24-hr autoradiogram. Starting from the termination of the CAT reaction, the phase-extraction alternate protocol takes ≤30 min to obtain quantitative results from twelve samples.

The hGH radioimmunoassay is rapid and requires very little hands-on time. Harvest is accomplished by removing an aliquot of medium, so it takes only a few minutes. Assay of twenty-four samples takes no more than 30 min hands-on time and 2 hr running time.

Literature Cited

- DeNoto, F.M., Moore, D.D., and Goodman, H.M. 1981. Human growth hormone DNA sequence and mRNA structure: Possible alternative splicing. *Nucl. Acids Res.* 9:3719-3730.
- Gilman, M.Z., Wilson, R.N., and Weinberg, R.A. 1986. Multiple protein-binding sites in the 5'-flanking region regulate c-fos expression. *Mol. Cell. Biol.* 6:4305-4316.
- Gorman, C.M., Moffat, L.F., and Howard, B.H. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* 2:1044-1051.

Prost, E. and Moore, D.D. 1986. CAT vectors for analysis of eukaryotic promoters and enhancers. *Gene* 45:107-111.

Seed, B. and Sheen, J.-Y. 1988. A simple phase-extraction assay for chloramphenicol acetyltransferase activity. *Gene* 67:271-277.

Selden, R.F, Burke-Howie, K., Rowe, M.E., Goodman, H.M., and Moore, D.D. 1986. Human growth hormone as a reporter gene in regulation studies employing transient gene expression. *Mol. Cell. Biol.* 6:3173-3179.

Key References

Gorman et al., 1982. See above.

Describes the initial CAT vectors and the harvest and chromatographic assay.

Seed and Sheen, 1988. See above.

Describes the phase-extraction protocol for CAT assays.

Selden et al., 1986. See above.

Describes the hGH assay, its use to characterize regulatory events, and vectors that can be used in the assay system.

Contributed by Robert E. Kingston and
Jen Sheen (phase-extraction)
Massachusetts General Hospital and
Harvard Medical School
Boston, Massachusetts

David Moore (hGH assay)
Massachusetts General Hospital
Boston, Massachusetts

Nonisotopic Assays for Reporter Gene Activity

This unit describes two nonisotopic systems for reporter gene activity in cells transfected with the firefly luciferase expression plasmid (first basic protocol and alternate protocol) or the β -galactosidase expression plasmid (second basic protocol). Both of these chemiluminescent assays have the advantages of high sensitivity and broad linear range. In the chemiluminescent detection procedures given in this unit, both luciferase activity and β -galactosidase activity can be measured with either a luminometer or a scintillation counter. Additional discussion and comparison of nonisotopic systems can be found in *UNIT 9.6*.

BASIC PROTOCOL 1

FIREFLY LUCIFERASE REPORTER GENE ASSAY

This protocol describes a reporter system in which firefly luciferase activity is measured to determine the level of expression of a transfected gene. This nonradioactive system has become popular because it is more sensitive than the chloramphenicol acetyltransferase (CAT) system (*UNIT 9.7A*), and individual assays are much quicker and easier to carry out. Cells transfected with luciferase expression plasmids (Fig. 9.7.4) are lysed using either Triton X-100 or freezing and thawing (alternate protocol) to release the reporter protein luciferase. Both ATP and the substrate luciferin are added to the lysate in a luminometer. The enzyme catalyzes a rapid, ATP-dependent oxidation of the substrate, which then emits light. Total light output is measured and is proportional to the amount of luciferase present over a wide range of enzyme concentrations.

Materials

- Cells transfected with luciferase expression plasmid (*UNITS 9.1-9.4 & 16.13*; Fig. 9.7.4), in 60-mm petri plates
- PBS (*APPENDIX 2*), ice-cold
- Triton/glycylglycine lysis buffer (see recipe)
- Luciferase assay buffer (see recipe)
- Luciferin stock solution (see recipe)
- 25 mM glycylglycine, pH 7.8 (free base, crystalline; Sigma)
- Firefly luciferase of known activity (Sigma) for use as standard
- Rubber policeman or equivalent
- Luminometer (manual or automated) with printer or chart recorder and cuvettes

1. Wash three 60-mm plates of cells transfected with luciferase expression plasmid three times with 4 ml ice-cold PBS per wash, aspirating medium between each wash.

The luciferase enzymatic reaction is inhibited by traces of calcium; therefore, calcium-containing medium must be washed out of the plate thoroughly before cell harvest. Triplicate plates are needed for statistical manipulations.

2. Add 350 μ l Triton/glycylglycine lysis buffer to each plate. Scrape with a rubber policeman.
3. Transfer solubilized cells to a 1.5-ml microcentrifuge tube. Microcentrifuge 5 min at maximum speed, 4°C. Transfer supernatant (cell lysate) to a clean microcentrifuge tube and store on ice for assay.

Reporter luciferase activity is efficiently extracted by Triton; <10% of reporter activity should remain in the insoluble pellet after centrifugation.

4. Gently vortex cell lysate and place 100 μ l in a luminometer cuvette. Add 360 μ l luciferase assay buffer. Place cuvette in luminometer chamber.

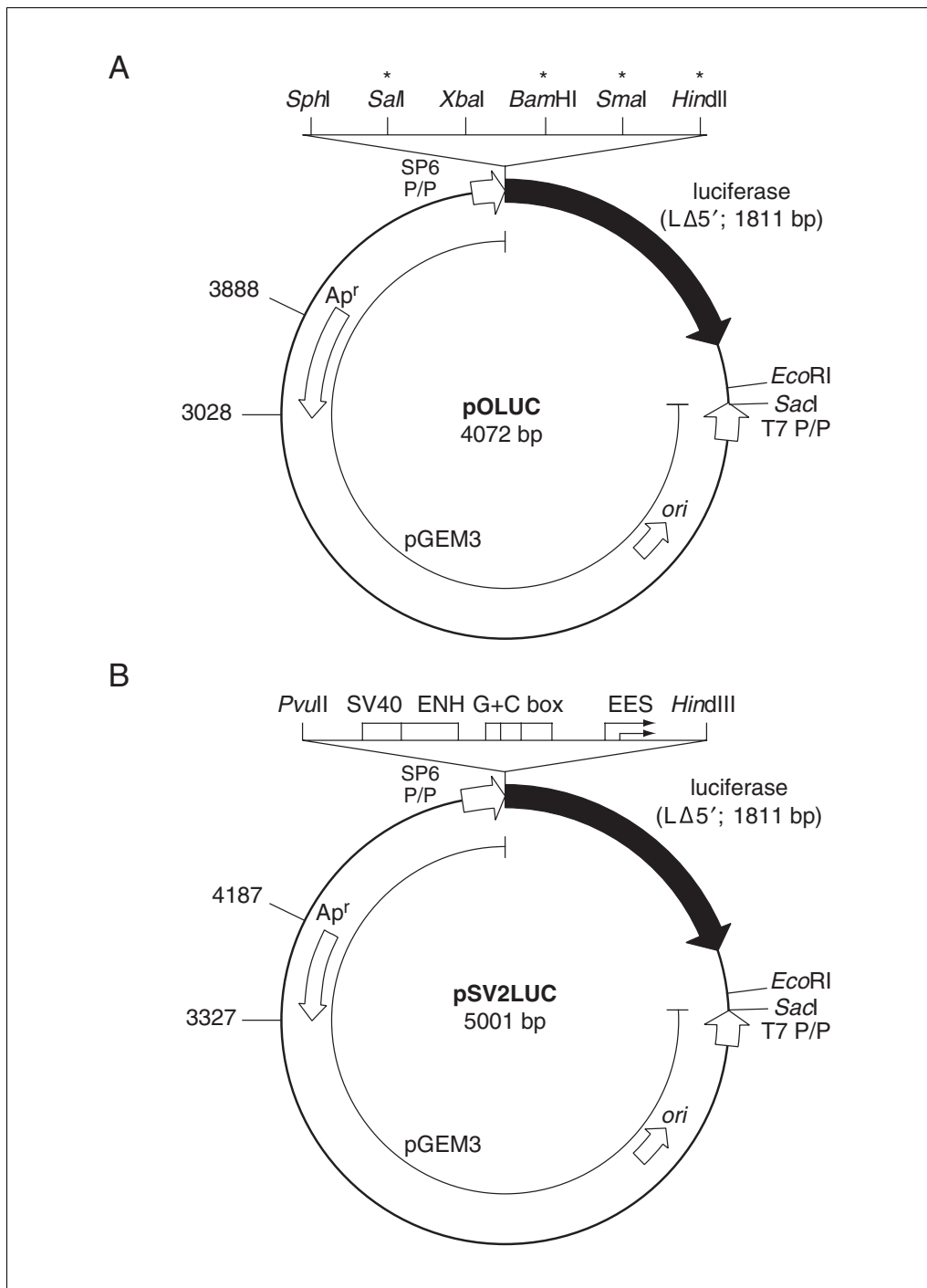


Figure 9.7.4 Luciferase expression vectors. **(A)** pOLUC is a promoterless vector containing a luciferase gene. It is derived from pGEM3, a high-copy ampicillin-resistance plasmid whose multiple cloning sites are flanked by SP6 promoter/primer sequences (SP6 P/P) and T7 promoter/primer sequences (T7 P/P). Unique restriction sites for insertion of sequences of interest are indicated by asterisks (*). **(B)** pSV2LUC contains the strong eukaryotic simian virus-40 (SV40) promoter/enhancer sequences inserted into pOLUC for use as a positive control for luciferase expression or as a standard for normalization between transfections. SV40 ENH designates the 72-base-pair repeats of the SV40 enhancer, G+C box(es) are the SP1-binding sites of the SV40 promoter, and EES designates the early-early transcription start sites.

5. Dilute luciferin stock solution to 200 μ M in 25 mM glycylglycine, pH 7.8.

Luciferin is rapidly oxidized by exposure to light; discard unused diluted solution.

6. Inject 200 μ l diluted luciferin solution into the sample in the luminometer and measure light output for 20 sec at 25°C. Quantitate by subtracting machine background (determined by measuring light output in a cuvette that has no cellular lysate) and by comparing bioluminescence of each sample with luciferase standards obtained commercially.

Enzymatic activity is proportional to light activity over three orders of magnitude. Because the luciferase reaction is complete in 20 sec, samples containing reporter activity that exceeds the linear range of the photomultiplier tube can be immediately reassayed after appropriate dilution in glycylglycine buffer.

ALTERNATE PROTOCOL

LUCIFERASE ASSAY IN FREEZE-THAW-LYSED CELLS

This protocol uses cycles of freezing and thawing transfected cells in a potassium phosphate buffer to produce a cell lysate. Although this method of extraction takes longer than Triton X-100 lysis, the lysates give more CAT enzymatic activity for experiments using cotransfected CAT reporter vectors and are stable for prolonged periods at 4°C.

Additional Materials (also see *Basic Protocol 1*)

Extraction buffer (see recipe)

Additional reagents and equipment for freeze-thaw lysis of cells as for chromatographic CAT assay (*UNIT 9.7A*)

1. Wash three 60-mm plates of cells transfected with luciferase plasmid in PBS as described in step 1 of the first basic protocol.
2. Add 1 ml extraction buffer to each plate. Scrape cells immediately with a rubber policeman.
3. Transfer cells to a 1.5-ml microcentrifuge tube. Microcentrifuge 15 to 30 sec.
4. Remove and discard supernatant. Add 100 μ l extraction buffer to the cell pellet.
5. Lyse cell membranes by three cycles of freeze-thaw.
6. Microcentrifuge lysed cells 5 min at maximum speed, 4°C. Remove and assay the supernatant as in the luciferase assay basic protocol (steps 4 to 6).

CHEMILUMINESCENT β -GALACTOSIDASE REPORTER GENE ASSAY

This protocol describes a chemiluminescence detection procedure for measuring β -galactosidase activity in extracts of cells transfected with a β -galactosidase expression plasmid (Fig. 9.7.5). The assay requires <30 min and has a sensitivity equivalent to or greater than the luciferase assay. Cells are lysed with a lysis solution containing Triton X-100. After a short incubation with substrate, a light-emission accelerator is added to enhance light production and detection. Light emission is measured with a luminometer or a scintillation counter. When the detection is performed in a luminometer, the dynamic range of the assay is approximately six orders of magnitude (20 fg to 20 ng of β -galactosidase). Lysates of cells cotransfected with luciferase reporter plasmids can also be assayed for luciferase activity (first basic protocol).

**BASIC
PROTOCOL 2**

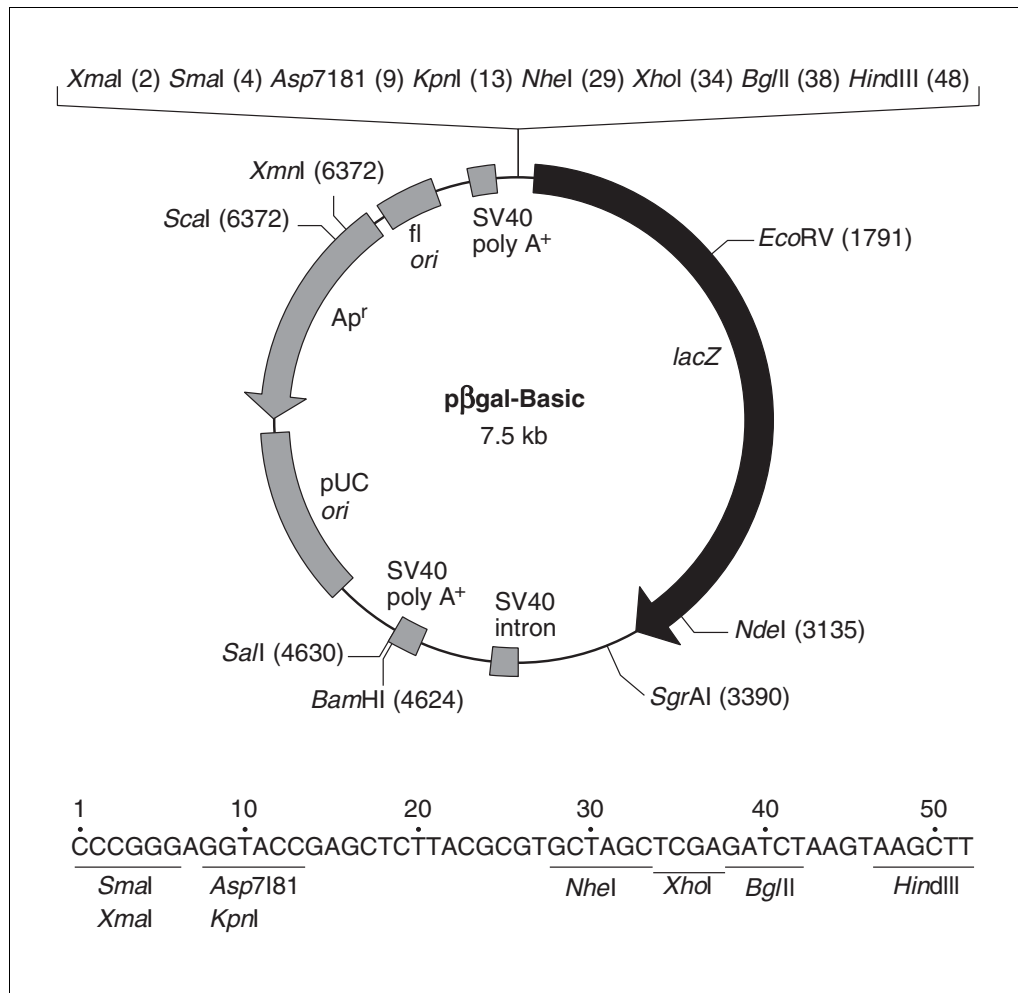


Figure 9.7.5 β -galactosidase expression vector. p β -gal-Basic lacks eukaryotic promoter and enhancer sequences and allows insertion of promoter-containing DNA fragments into the polylinker upstream of the *lacZ* gene. Enhancer sequences can be cloned into the polylinker or into unique sites downstream of *lacZ*. All four vectors in this series contain an SV40 intron and SV40 polyadenylation signal downstream of *lacZ* to ensure proper and efficient processing of the transcript. The vector backbone contains a f1 origin for single-stranded DNA production and a pUC-19 origin of replication and ampicillin-resistance gene for propagation in *Escherichia coli*. The multiple cloning site (polylinker) region is identical in all four vectors in the series; in p β -gal-Promoter and in p β -gal-Control a 202-bp promoter-containing fragment has been inserted between *BglII* and *HindIII*. The sequence of p β -gal-Basic has been deposited in GenBank (Accession #U13184).

**Introduction of
DNA into
Mammalian Cells**

9.7.15

Materials

Cells transfected with a β -galactosidase expression plasmid (*UNITS 9.1-9.4 & 16.13*; Fig. 9.7.5), in 60-mm petri plates
Mock-transfected control cells, in 60-mm petri plates
PBS (*APPENDIX 2*)
Triton lysis solution (see recipe)
 β -galactosidase reaction buffer (see recipe)
Light-emission accelerator solution (see recipe)
Rubber policeman (or equivalent)
Luminometer with chart recorder and tubes *or* scintillation counter

1. Rinse one 60-mm plate of cells transfected with β -galactosidase expression plasmid and one 60-mm plate of mock-transfected cells twice with PBS.

Mock-transfected cells are processed in parallel with transfected cells and the resulting lysate is used for positive and negative controls.

2. Add 250 μ l Triton lysis solution.
3. Detach cells from plate using a rubber policeman. Transfer cell extract to a 1.5-ml microcentrifuge tube and microcentrifuge 2 min at maximum speed, 4°C.
4. Transfer supernatant to a clean microcentrifuge tube.

As a negative control, assay a volume of mock transfected cell extract equivalent to the volume of experimental cell extract used. As a positive control, add 1 μ l β -galactosidase (Sigma Grade VIII, 10 U/ml stock) to a volume equivalent to that of experimental cell extract.

Cell extracts may be used immediately or frozen several months at -70°C before continuing with the procedure.

5. Place 2 to 20 μ l cell extract into a luminometer tube.

Assays should be performed in triplicate.

The amount of cell extract required may vary depending on the level of enzyme expression. For accurate comparisons, the signal should remain within the linear range of the assay. Use 5 μ l extract if β -galactosidase is an internal control and 10 to 20 μ l extract for experiments where low levels of enzyme expression are expected. Smaller volumes of extract may increase the coefficient of variation as a result of increased pipetting errors.

6. Add 200 μ l β -galactosidase reaction buffer to the luminometer tube and incubate 60 min at room temperature.

Incubation time may be decreased to 15 min; however, sensitivity may decrease slightly as a result.

Chemiluminescent signal intensity varies with time. The incubation time with β -galactosidase reaction buffer should be identical for each sample. For example, if it takes 10 sec to complete each measurement, the reaction buffer should be added to consecutive tubes at 10-sec intervals.

7. Place tube into luminometer and inject 300 μ l light-emission accelerator solution. Wait 2 to 5 sec after injection and take the measurement for 5 sec.

Injections may be performed manually if the accelerator is added in the same time interval as the reaction buffer. See Background Information for a discussion of how to use a scintillation counter to measure activity.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

β-galactosidase reaction buffer

100 mM sodium phosphate, pH 8.0

1 mM magnesium chloride

1× 3-(4-methoxyspiro[1,2-dioxetane-3,2'-(5'-chloro)-tricyclo[3.3.3.1^{3,7}]decan]-4-yl)phenyl β-D-galactopyranoside (Galacton) chemiluminescent substrate (Tropix)

Store up to one month at 4°C

This solution is available as part of the Galacto-Light assay kit (Tropix).

Extraction buffer

100 mM potassium phosphate, pH 7.8

1 mM DTT (add just before use)

Store at room temperature; stable several months without DTT

Light-emission accelerator solution

10% (v/v) Emerald luminescence amplifier (Tropix)

0.2 N sodium hydroxide

Store up to 1 year at 4°C

A modified version of this solution that produces lower, nonenzymatic background is available as part of the Galacto-Light assay kit (Tropix).

Luciferase assay buffer

25 mM glycylglycine, pH 7.8

15 mM potassium phosphate, pH 7.8

15 mM MgSO₄

4 mM EGTA

2 mM ATP (Sigma)

1 mM DTT (add just before use)

Store at 4°C; stable several months without DTT

Luciferin stock solution

1 mM D-luciferin, synthetic crystalline (Sigma)

25 mM glycylglycine, pH 7.8

10 mM DTT (add just before use)

Divide into 1-ml aliquots and place in a light-tight box

Store at -70°C; stable several months without DTT

Triton/glycylglycine lysis buffer

1% (v/v) Triton X-100

25 mM glycylglycine, pH 7.8

15 mM MgSO₄

4 mM EGTA

1 mM DTT (add just before each use)

Store at 4°C; stable several months without DTT

Triton lysis solution

100 mM potassium phosphate, pH 7.8

0.2% (v/v) Triton X-100

1 mM DTT (add fresh immediately prior to use)

Store at 4°C; stable several months without DTT

This solution is available as part of the Galacto-Light assay kit (Tropix).

COMMENTARY

Background Information

Firefly luciferase assay

Firefly luciferase is an enzyme that produces light by ATP-dependent oxidation of the substrate luciferin in a two-step process. The heterocyclic luciferin is first adenylated and then undergoes oxidative decarboxylation, producing AMP, CO₂, and light from an activated luciferin intermediate. Light emission is maximal at 562 nm; the time course consists of an initial flash within seconds of adding luciferin, followed by a plateau of light output. When the assay is performed in the presence of excess substrates, light output is proportional to enzyme concentration. The luciferase reporter assay offers a sensitive, rapid, nonradioactive assay for measurement of promoter activity (Gould and Subramani, 1988; Brasier et al., 1989; Williams et al., 1989). A discussion of this and other widely used reporter systems is presented in the introduction to Section II and *UNIT 9.6* of this chapter.

With the cloning and expression of the luciferase gene from the firefly *Photinus pyralis* by de Wet et al. (1987), luciferase cDNA was cloned into derivatives of the pSV2 vector. The pSVOALA5' plasmid is a promoterless vector that contains both a 5' deleted cDNA missing an upstream ATG initiation codon and tandem polyadenylation signals between the pBR322 portion of the vector and the luciferase 5' untranslated tract.

Because of the limited number of unique restriction sites available for cloning test promoters, several investigators have engineered multiple cloning sites into the same or distinct vectors. Nordeen et al. (1988) generated the PXP-(1,2) series of vectors that contain multiple cloning sites in a modified pSV2 vector. Brasier et al. (1989) generated a promoterless luciferase reporter vector, pOLUC, containing LA5' sequences in a pGEM3-derived plasmid (Fig. 9.7.4). The pGEM3-derived vectors have the advantage of being high-copy-number plasmids, which allows production of larger amounts of plasmid DNA for transfection studies. Because of the potential problem with luciferase activation of cryptic promoters within reporter vectors, it is always wise to map the mRNA start site of any new promoter. With the pGEM3-derived luciferase vectors, no transcripts were detected originating upstream of a minimal promoter from the gene encoding angiotensinogen when it was transfected into HepG2 cells (Ron et al., 1990).

Chemiluminescent β -galactosidase assay

β -galactosidase is one of the most widely used reporter genes. It is commonly quantitated using the colorimetric substrate *o*-nitrophenyl- β -D-galactopyranoside (ONPG), but this assay has poor sensitivity with a detection limit of ~100 pg. It has been reported that as little as 2 fg of β -galactosidase can be detected with a chemiluminescent assay incorporating a 1,2-dioxetane chemiluminescent substrate for β -galactosidase (Galacton substrate; Jain and Magrath, 1991). Therefore, the chemiluminescent β -galactosidase assay is over three orders of magnitude more sensitive than the chloramphenicol acetyltransferase (CAT) assay and two orders of magnitude more sensitive than the fluorometric assay utilizing 4-methylumbelliferyl- β -D-galactopyranoside (MUG; Alam and Cook, 1990; Jain and Magrath, 1991). Upon cleavage with β -galactosidase, Galacton substrate breaks down and light is generated at a maximum wavelength of 470 nm (Bronstein et al., 1989). The signal-to-noise ratio of the reaction is further increased by the addition of a quaternary ammonium polymer formulation such as the Emerald luminescence amplifier (Tropix). These polymers increase the emission efficiency of the reaction by reducing aqueous quenching of the excited-state emitter in the polymer's hydrophobic domains. This increases the radiative decay of the excited state, thus producing a higher-intensity signal. Energy transfer to fluorescein, which possesses a higher quantum yield of fluorescence, provides further signal enhancement and a shift in emission maximum to 530 nm.

Higher sensitivity is achieved by performing the enzyme reaction at a different pH than is present during the measurement of the light signal. Bacterial β -galactosidase operates at a pH optimum between 6.5 and 8.0. In contrast, light production is increased at a pH > 9.5. Therefore, purified enzyme or cell extract containing β -galactosidase is incubated 15 min to 1 hr with the chemiluminescent substrate at pH 8 to allow the enzymatic reaction to occur. A light-emission accelerator containing alkali is added and light emission is measured immediately.

This chemiluminescent β -galactosidase assay has also been used successfully to measure β -galactosidase levels in bacteria (Nelis and Van Poucke, 1993) and yeast (D. Nathan and S. Linqvist, personal communication).

Critical Parameters

Measurements for luminescent reporter assays require specialized equipment—either a luminometer or a scintillation counter. Choice of equipment to quantitate bioluminescence must include consideration of several variables. Luciferase activity is proportional either to peak light output or integrated light output between 1 and 16 sec. The intra-assay variation is lower with integrated values, but some manual luminometers are equipped with chart recorders that only register peak luminescence. Although automated equipment is generally more expensive than manual, automation can be cost-effective for laboratories that rely heavily on data generated by cell transfection. Sensitivity is a vexing problem in measurement of bioluminescence as no standards exist for comparisons among different equipment. In practice, most available luminometers are adequately sensitive for assaying reporter gene activity; however, before investing in an expensive piece of equipment, it is wise to test a transfection on a demonstration model. More extensive discussion of luminometers and suppliers is found in VanDyke (1985) and Stanley (1992). A luminometer with a low-noise photomultiplier tube is recommended to achieve maximum sensitivity and dynamic range in the chemiluminescent β -galactosidase assay and the bioluminescent luciferase assay. Alternatively, a scintillation counter may be used for each of these assays (Fulton and Van Ness, 1993), but lower sensitivity may result. In the chemiluminescent β -galactosidase assay, the light emission maximum occurs ~2 sec after the addition of accelerator and decays with a half-life of approximately 4.5 min. Alternatively, Galacton-Plus chemiluminescent substrate exhibits a half-life of greater than 180 minutes and is better suited for use with luminometers without injectors or with scintillation counters. In contrast, the luciferase assay can be formatted in two ways that yield different kinetics: an unenhanced assay characterized by a flash of light followed by a rapid decay in light emission, and an enhanced assay that exhibits greater light output and slower signal decay. Nguyen describes the use of a scintillation counter to measure an unenhanced luciferase assay (Nguyen et al., 1988). Enhanced luciferase assays, available from Analytical Luminescence Laboratory and Promega, allow easier measurement in luminometers without injectors and in scintillation counters because light emission half-lives for these assays are 5 and 10 min, respectively.

Many luminometers are equipped with automatic injection devices. Such devices are optional with the β -galactosidase and the enhanced luciferase assays. The volumes recommended in the β -galactosidase detection protocol for both the reaction buffer and the light-emission accelerator may be proportionally reduced to use the smaller-volume injectors often recommended in luciferase protocols or microtiter plates, with only a slight reduction in sensitivity.

Some scintillation counters can measure chemiluminescence directly if the coincidence circuit is turned off. If chemiluminescence is measured without turning off the coincidence circuit, a linear relationship can be established by calculating the square root of the cpm minus background (Nguyen et al., 1988). The usefulness of a scintillation counter for measuring luminescence assays depends on the instrument background noise and ease with which the sample may be introduced into the counting chamber.

Firefly luciferase assay

Cell type. Luciferase enzyme in transfected mammalian cells is targeted to membrane-bound peroxisomes, apparently by information encoded by carboxy terminal amino acids (Gould and Subramani, 1988). This targeting may confound application of the assay to all cell types—e.g., it may affect the optimal time of harvest for transfected cell lines. In HepG2 cells transiently transfected with SV40 enhancer/promoter-driven luciferase plasmids, stable luciferase enzymatic activity was seen after 48 hr. In striking contrast, with rat hepatoma (H4 II E) and rat islet (RIN 1056A) cells, peak expression of luciferase reporter activity was seen 12 to 24 hr after transfection. Thus, this reporter system may not be ideal for use in these particular cell lines. Cell lines that have been successfully transfected include HepG2, JEG, HeLa, BHK, COS, mouse L cells, CV-1, and Jurkat. Clearly, a time course for assay of luciferase is an important parameter to examine when transfecting cell types other than these.

Cell lysis. Luciferase is commercially available for use as a positive control for assay reagents. ATP concentration is one important parameter to optimize in cellular lysates because ATP levels may vary within biological samples. In HepG2 cells, a concentration of 1 to 2 mM ATP gives optimal luminescence; concentrations >2 mM inhibit the activity of the transfected enzyme (Nguyen, 1988; Brasier et al., 1989).

Extracts prepared by Triton lysis are more unstable than extracts prepared by freeze-thaw lysis (alternate protocol), and must be assayed immediately or stored at -70°C to maintain full activity (Brasier et al., 1989).

Hormonal manipulation. The luciferase system has successfully been used to analyze the regulatory effects of several hormones, including glucocorticoids and thyroid hormone, on various promoters. However, an artifactual 2- to 5-fold induction of luciferase activity by dibutyryl cyclic AMP in pOLUC-transfected HepG2 cells has been observed (A.B., unpub. observ.). Thus, this vector is probably not appropriate for studies of cyclic AMP response elements.

Cotransfection with control plasmid. Changes in transfection efficiency can be evaluated by transfecting with both a luciferase test plasmid and an internal control plasmid such as the chloramphenicol acetyltransferase (CAT) expression plasmid. Normalizing data for transfection efficiency lowers scatter within triplicate plates and makes measurement of weak promoters reproducible. The inclusion of a cotransfected reporter may influence the choice of a lysis strategy. CAT enzymatic activity is detectable with the Triton lysis protocol, but the sensitivity of the CAT assay can be reduced 10-fold by the concentration of detergent used in the luciferase lysis procedure. Lower concentrations of detergent (e.g., 0.5%) are adequate to lyse cells, but freeze-thaw lysis should be considered if the activity of the cotransfected CAT reporter is weak. The CAT assay is compatible with the potassium phosphate lysis buffer, and both enzymes exhibit the same pH optimum. Alkaline phosphatase expressed by a cotransfected control plasmid can also be extracted with the luciferase Triton lysis procedure. In the spectrophotometric assay of Triton X-100 extracts, alkaline phosphatase activity is linear, but the blank and alkaline phosphatase standards must have the same final concentration of Triton as the cellular lysates. Use of the hGH reporter system (UNIT 9.7A) or the chemiluminescent secreted alkaline phosphatase assay (Bronstein et al., 1994) avoids the problem of matching lysis strategies.

Chemiluminescent β -galactosidase assay

Cotransfection. The chemiluminescent β -galactosidase substrate is ideal for cotransfection with the luciferase reporter gene because both assays are rapid, exhibit high sensitivity, and can be performed in the same cell extracts

(Fulton and Van Ness, 1993). Luciferase assay measurements can be performed during the β -galactosidase assay steps. Detection of the β -galactosidase samples follows the luciferase measurements.

Controls. Purified β -galactosidase should be added to mock-transfected cell extracts and used as a positive control. Mock-transfected cell extracts are important as a negative control to establish an assay baseline by measuring the endogenous enzyme background.

Cell lysis. Cell extracts may be prepared using a Triton X-100 lysis buffer such as that listed in the protocol. Alternatively, the cell culture lysis reagent recommended by Promega for use with luciferase can be used in cotransfections of the luciferase and β -galactosidase reporter genes.

Endogenous enzyme activity. It has been shown that endogenous β -galactosidase activity can be reduced by increasing the pH of the reaction buffer from pH 7 to pH 8 (Jain and Magrath, 1991). Inhibition of the endogenous enzyme enables greater discrimination between the transfected bacterial enzyme and endogenous activity. Heating extracts at 50°C for one hour has been shown to significantly reduce endogenous enzyme activity (Young et al., 1993). A modified version of this procedure has been used with the luminescent detection assay to detect β -galactosidase in tissue (Shaper et al., 1994).

Anticipated Results

Luciferase activity is linear up to 200 μg protein in cell extracts; using integrated light output between 1 and 16 sec, the intra-assay variation should be $\leq 5\%$. In a representative experiment from transfected HepG2 cells with assays of triplicate plates under the conditions described in the basic protocol, activity from pOLUC-transfected cells was 1572 ± 60 integrated light units (ILU) and activity from pSV2LUC-transfected cells was $83,040 \pm 1400$ ILU, where the luminometer background was 100 ILU. The sensitivity of the luciferase reporter system is estimated to be 10 to 1000 times that of the CAT reporter system (de Wet et al., 1987; Williams et al., 1989).

The sensitivity of the chemiluminescent β -galactosidase assay varies widely and is highly dependent on the level of endogenous β -galactosidase. Background attributed to the chemiluminescent reagents alone is generally very low; however, the range may vary somewhat depending upon the circumstances and instrument used. For example, background measured

using a Berthold Luminometer Model LB952T is usually 400 to 800 relative light units measured in a 5-sec interval. In cells with relatively high levels of endogenous enzyme activity, this assay can still be quite useful as an internal control for transfection efficiency.

Time Considerations

Harvesting twelve samples and preparing extracts for the luciferase assay should take 30 min. Each assay point is finished within 20 sec of injecting luciferin solution. With an automated luminometer, a 24-point assay is completed within 12 min. Data manipulation requires ~1 hr.

Extract preparation for 12 samples for the chemiluminescent β -galactosidase assay may be accomplished in <30 min. A 15-min to 1-hr substrate incubation is followed by accelerator addition and measurement in a luminometer, which requires ~15 sec per sample. When this assay is used as an internal control for luciferase, both assays can be performed concurrently.

Literature Cited

- Alam, J. and Cook, J.L. 1990. Reporter genes: Application to the study of mammalian gene transcription. *Anal. Biochem.* 188:245-254.
- Brasier, A.R., Tate J.E., and Habener, J.F. 1989. Optimized use of the firefly luciferase assay as a reporter gene in mammalian cell lines. *BioTechniques* 7:1116-1122.
- Bronstein, I., Edwards, B., and Voyta, J.C. 1989. 1,2-Dioxetanes: Novel chemiluminescent enzyme substrates. *J. Biolum. Chemilum.* 4:99-111.
- Bronstein, I., Fortin, J.J., Voyta, J.C., Juo, R.R., Edwards, B., Olesen, C.E.M., Lijam, N., and Kricka, L.J. 1994. Chemiluminescent reporter gene assays: Sensitive detection of the GUS and SEAP gene products. *BioTechniques* 17:172-178.
- de Wet, J.R., Wood, K.V., DeLuca, M., Helinski, D.R., and Subramani, S. 1987. Firefly luciferase gene: Structure and expression in mammalian cells. *Mol. Cell. Biol.* 7:725-737.
- Fulton, R. and Van Ness, B. 1993. Luminescent reporter gene assays for luciferase and β -galactosidase using a liquid scintillation counter. *BioTechniques* 14(5):762-763.
- Gould, S.J. and Subramani, S. 1988. Firefly luciferase as a tool in molecular and cell biology. *Anal. Biochem.* 7:5-13.
- Jain, V. and Magrath, I. 1991. A chemiluminescent assay for quantitation of β -galactosidase in the femtogram range: Application to quantitation of β -galactosidase in lacZ-transfected cells. *Anal. Biochem.* 199:119-124.

- Nelis, H.J. and Van Pouke, S.O. 1993. Comparison of chemiluminogenic, fluorogenic, and chromogenic substrates for the detection of total coliforms in drinking water. *Proc the Water Quality Technology Conference (AWWA)* 1663-1673.
- Nguyen, V.T., Morange, M., and Bensusade, O. 1988. Firefly luciferase luminescence assays using scintillation counters for quantitation in transfected mammalian cells. *Anal. Biochem.* 171:404-408.
- Nordeen, S.K. 1988. Luciferase reporter gene vectors for analysis of promoters and enhancers. *BioTechniques* 6:454-457.
- Ron, D., Brasier, A.R., Wright, K.A., Tate, J.E., and Habener, J.F. 1990. An inducible 50 kilodalton NFB-like protein and a constitutive protein both bind the acute-phase response element of the angiotensinogen gene. *Mol. Cell. Biol.* 10:1023-1032.
- Shaper, N. Harduin-Lepers, A. and Shaper, J.H. 1994. Male germ cell expression of murine β -4-galactosyltransferase. A 796-base pair genomic region containing two cAMP-responsive elements (CRE)-like elements, mediates expression in transgenic mice. *J. Biol. Chem.* 269:25165-25171.
- Stanley, P.E. 1992. A survey of more than 90 commercially available luminometers and imaging devices for low-light measurement of chemiluminescence and bioluminescence, including instruments for manual, automatic, and specialized operations, for HPLC, LC, GLC, and microtiter plates. Part I: Description. *J. Biolum. Chemilum.* 7:77-108.
- VanDyke, K. 1985. *In* Bioluminescence and Chemiluminescence: Instruments and Applications, Vol I (K. VanDyke, ed.) pp. 83-128. CRC Press, Boca Raton, Fla.
- Williams, T.M., Buerlein, J.E., Ogden, S., Kricka, L.J., and Kant, J.A. 1989. Advantages of the firefly luciferase as a reporter gene: Application to the interleukin-2 gene promoter. *Anal. Biochem.* 176:28-32.
- Young, D.C., Kingsley, S.D, Ryan, K.A., and Dutko, F.J. 1993. Selective inactivation of eukaryotic β -galactosidase in assays for inhibitors of HIV-1 TAT using bacterial β -galactosidase as a reporter enzyme. *Anal. Biochem.* 215:24-30.

Key Reference

Jain and Magrath, 1991. See above.

Describes optimization of the chemiluminescent β -galactosidase assay.

Contributed by Allan R. Brasier
University of Texas
Galveston, Texas

John J. Fortin
Tropix, Inc.
Bedford, Massachusetts

**Introduction of
DNA into
Mammalian Cells**

9.7.21

Use of the *A. Victoria* Green Fluorescent Protein to Study Protein Dynamics in Vivo

Fluorescent molecules serve as valuable tools for the detection of a variety of biochemical phenomena. Such reagents have been employed for protein localization, quantitation of gene expression, detection of nucleic acids, cell sorting, and determination of chemical concentrations. However, the use of such techniques generally requires significant nonphysiological perturbations to the biological system being studied. For example, cells being processed for immunofluorescence must be fixed and permeabilized. Therefore, although this method is relatively simple and efficacious, it is not appropriate for the observation of dynamic phenomena.

A variety of fluorescent compounds exist that can efficiently localize within living cells. For instance, the karyophilic dye Hoechst 33258 is often used to detect the nuclei of living cells in culture, whereas the lipophilic dye DiI (1, 1'-dioctadecyl-3, 3, 3', 3'-tetramethylindocarbocyanine perchlorate) is commonly used to label cell membranes within living tissues. However, such compounds cannot be used to localize specific proteins, and the range of cellular constituents that can be labeled is somewhat limited.

Methods for transiently permeabilizing cells to allow uptake of fluorescent macromolecules have been successfully employed in conjunction with in vivo detection. For instance, microinjection of fluorophore-labeled antibodies can be used to track the dynamics of cellular antigens. However, such techniques are restricted to a limited variety of cell types, and permeabilization may perturb the cells' behavior. Furthermore, such procedures are technically difficult and may require specialized equipment.

In sum, although fluorescence is a useful tool for detecting molecules within cells, its application in vivo has heretofore been limited. The ideal vital fluorescent tag should (1) be detectable without causing cytological damage, (2) be able to label a wide variety of cell types readily, and (3) be able to be targeted to virtually any subcellular region. The recently cloned (Prasher et al., 1992) *Aequorea victoria* green fluorescent protein (GFP) is potentially such a molecule. The object of this section is to describe the use of this proteinaceous fluorophore for in vivo observation of cellular phenomena.

OVERVIEW OF GFP FLUORESCENCE

In the jellyfish *A. victoria*, the chemiluminescent protein aequorin emits blue light upon binding calcium ions. In vivo, however, the jellyfish glows green. This is because a second protein, GFP, absorbs the (short-wavelength) blue light and concomitantly emits (longer-wavelength) green light (Morin and Hastings, 1971). Although the 28-kDa protein was isolated over twenty years ago, its sequence has only recently been determined (Prasher et al., 1992). The utility of this protein as a molecular probe was first demonstrated by Chalfie et al. (1994), who expressed GFP in both *Escherichia coli* and *Caenorhabditis elegans*. In both cases, upon excitation with long-wave ultraviolet or blue light, the organism glowed green, indicating that *A. victoria*-specific factors were not required for GFP fluorescence. Furthermore, the spectral properties of GFP produced in *E. coli* are essentially indistinguishable from those of the protein isolated directly from the jellyfish (Morise et al., 1974; Chalfie et al., 1994; Inouye and Tsuji, 1994a).

UTILIZATION OF GFP

Because the cDNA sequence of GFP is known, it can in principle be expressed in any organism for which a gene delivery system is available. To date, GFP has been expressed and detected in a wide variety of organisms, including bacteria (Chalfie et al., 1994), fungi (Flach et al., 1994), plants (Haseloff and Amos, 1995), cultured mammalian cells (Olson et al., 1995; Pines, 1995), nematodes (Chalfie et al., 1994), and fruit flies (Wang and Hazelrigg, 1994). It has been expressed and detected both in intact form and as fusions to other proteins. In at least one reported case, both N- and C-terminal fusions localized properly and retained their function in vivo (Wang and Hazelrigg, 1994).

The fluorescence properties of GFP have been extensively studied (see Ward, 1981). The fluorophore is excited maximally at a wavelength of 395 nm with a secondary peak (of lower absorption) at 475 nm. The emission spectrum has a single peak at 509 nm. Thus, GFP can be detected (suboptimally) using a microscope equipped for fluorescence with filters for the commonly used fluorophore fluorescein isothiocyanate (FITC). Figure 9.7.6A illustrates the GFP excitation and emission

Use of the
A. Victoria Green
Fluorescent
Protein to Study
Protein Dynamics
in Vivo

9.7.22

Contributed by Jason A. Kahana and Pam A. Silver
Current Protocols in Molecular Biology (1996) 9.7.22-9.7.28
Copyright © 2000 by John Wiley & Sons, Inc.

spectra along with the spectra of a typical FITC filter set. Notice that a typical FITC filter set fails to excite the major peak of GFP and excludes much of the light emitted by the protein. Two filter sets more suited to the excitation and emission spectra of GFP have recently become commercially available (Figs. 9.7.6B and C).

Isolated GFP is extremely stable; it retains its native fluorescence characteristics even when incubated with 6 M guanidinium hydrochloride, 8 M urea, or 1% SDS (Ward et al., 1980). Fluorescence is observed over a pH range of 7 to 12.2 and at temperatures $>65^{\circ}\text{C}$ (Ward, 1981). The protein can be denatured using a combination of chaotropic salt with strong acid (pH <4), strong base (pH >11), or high temperature. However, the protein will almost fully renature when such conditions are alleviated (Ward and Bokman, 1982). Thus, the

protein should retain its fluorescence under most physiological conditions.

Nascent GFP does not fluoresce immediately after translation; an intramolecular oxidation reaction that forms the chromophore must first occur. This chromophore is believed to be derived from the sequence Ser-Tyr-Gly at amino acid positions 65 to 67 (Cody et al., 1993). When GFP is expressed in bacteria under anoxic conditions, fluorescence is not observed. However, when the protein is exposed to atmospheric oxygen at room temperature, half-maximal fluorescence is observed after 1.5 to 2 hr (Heim et al., 1995). When purified protein is exposed to strong reducing agents such as sodium dithionite, fluorescence is completely lost. However, this state is reversible upon removal of reducing agent in the presence of atmospheric oxygen (Inouye and Tsuji, 1994b).

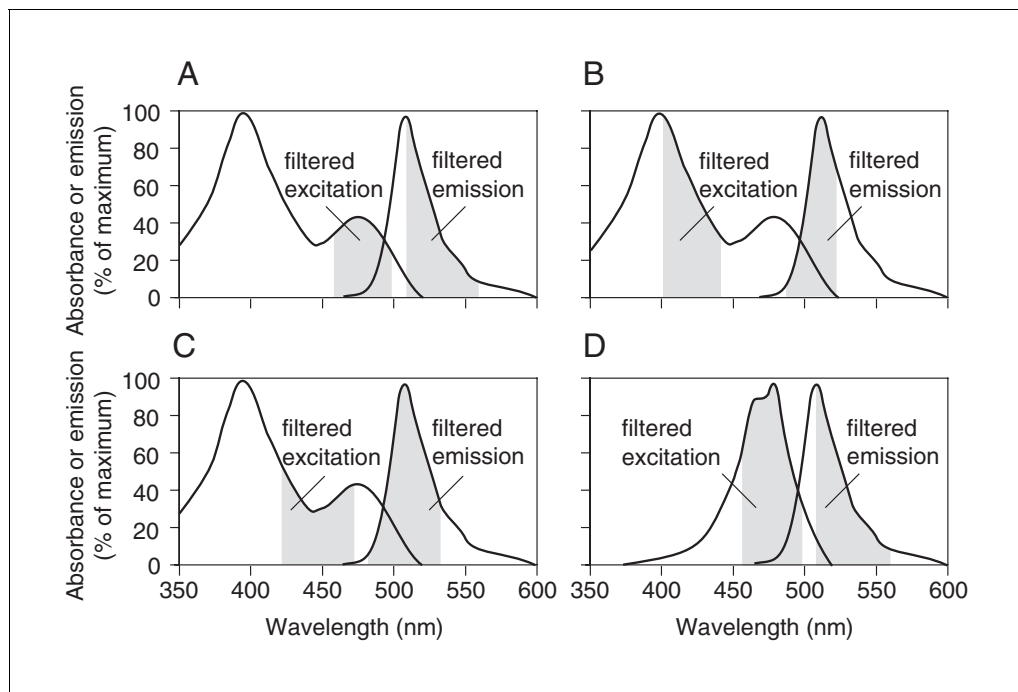


Figure 9.7.6 Selected spectra of wild-type and mutant GFPs juxtaposed with the spectra of several GFP epifluorescence filter sets (all from Chroma Technology). The left curve in each part is the excitation spectrum and the right is the emission spectrum. The grayed areas indicate the regions of the GFP spectra that are excited (“filtered excitation”) and passed (“filtered emission”) by the filter set. Note, however, that some excitation sources (e.g., mercury arc lamps) do not provide appreciable amounts of light over the entire filtered excitation region. **(A)** Spectra of wild-type GFP with the “Hi-Q FITC” filter set. **(B)** Wild-type GFP with the “Standard GFP” filter set, which takes advantage of the strong light output of a mercury lamp at 405 and 436 nm as well as the excitation and emission characteristics of GFP. **(C)** Wild-type GFP with the “Hi-Q GFP” filter set, which takes advantage of the strong light output of the mercury lamp at 436 nm as well as the high-amplitude ($\lambda_{\text{max}} = 509 \text{ nm}$) emission peak of GFP. **(D)** A spectrally shifted mutant [S65T, V163A] with the “Hi-Q FITC” filter set. Although the spectra of this mutant coincide well with those of the FITC filter set, a mercury arc lamp lacks strong excitation peaks over the 450 to 500 nm range. Thus, when using this mutant, a xenon arc lamp may be a more appropriate excitation source.

A number of GFP expression systems are currently available. The authors have developed a series of C-terminal fusion vectors that promote GFP expression in the budding yeast *Saccharomyces cerevisiae* (Fig. 9.7.7). Vectors designed to drive GFP expression in mammalian cells are commercially available (Clontech). Bacterial expression vectors have been described by Chalfie et al. (1994) and Heim et al. (1994). In these cases, high-level GFP expression is driven by a strong transcriptional promoter. For accurate expression (and localization) of fusion proteins, it may be advantageous to use the fusion gene-specific promoter if it provides a sufficient level of protein. For instance, a fusion of a yeast centrosomal protein, Nuf2p, to GFP was produced using the

NUF2 promoter to drive its expression (Kahana et al., 1995).

PROBLEMS WITH GFP

A frequent complaint about GFP is that it does not fluoresce brightly enough to be detected when expressed at low to moderate levels. However, high-level expression of GFP fusion proteins may result in mislocalization or cellular damage. In this section, three potential causes of this apparent “dimness” are identified: low “capacity,” “delayed” chromophore formation, and poor expression.

Low fluorescence capacity

Molecule for molecule, GFP is “dimmer” than most commonly used chemical fluoro-

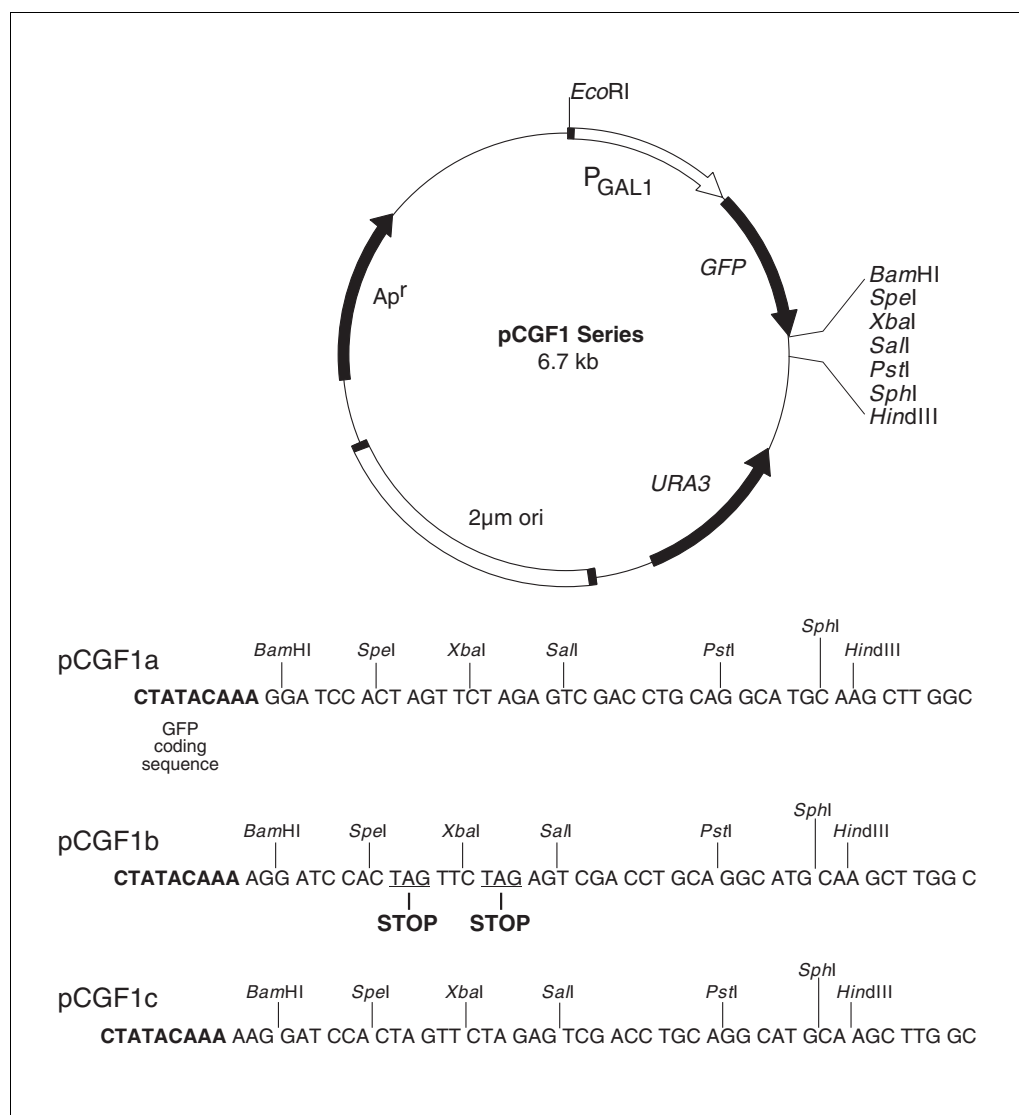


Figure 9.7.7 pCGF (C-terminal GFP Fusion) vectors for use in the yeast *S. cerevisiae*. The galactose-inducible *GAL 1* promoter (Yocum et al., 1984) is used to drive expression of a GFP fusion protein. A protein may be fused in any of three frames to the C-terminus of GFP. All restriction sites shown are unique.

phores. This is generally because GFP has a lower capacity to absorb excitation light than most commonly used chemical fluorophores. Basically, the phenomenon of fluorescence involves two activities: the absorption of light of one wavelength and emission of light of a longer wavelength. The ability of a material to fluoresce can therefore be expressed in terms of these two activities. Two quantities that describe them can be defined: extinction coefficient (ϵ) and quantum yield (QY). Extinction coefficient is a measure of the ability of a material to absorb radiation of a certain wavelength. Quantum yield reflects the ratio of (longer-wavelength) photons emitted to (shorter-wavelength) photons absorbed. In layman's terms, ϵ describes the "capacity" of the fluorophore, whereas QY represents its "efficiency." Thus, fluorescence "intensity" may be described (to a first approximation) as the product of ϵ and QY.

For wild-type GFP at a wavelength of 395 nm (the major excitation peak), ϵ is $24,000 \text{ M}^{-1} \text{ cm}^{-1}$ and QY is ~ 0.77 —that is, 77 photons are emitted per 100 photons absorbed (Ward, 1981; Heim et al., 1995). If a standard FITC filter set is used, this peak is not excited (see Fig. 9.7.6A). At the longer-wavelength excitation peak of 475 nm (which is excited using a FITC set), ϵ is diminished to $7150 \text{ M}^{-1} \text{ cm}^{-1}$. In contrast, FITC exhibits a similar QY but an ϵ of $76,000 \text{ M}^{-1} \text{ cm}^{-1}$ at its optimal excitation wavelength of 495 nm. Hence, a protein observed by immunofluorescence with a FITC-conjugated secondary antibody (assuming one FITC molecule per protein molecule) will appear roughly ten times brighter than a GFP fusion protein when viewed through a FITC filter set. In theory, GFP being detected with a GFP-specific filter set (see Fig. 9.7.6B) should fluoresce about one-third to one-fourth as brightly as FITC observed with a FITC filter set. In conclusion, GFP is an inherently dimmer fluorophore than the spectrally similar FITC.

"Delayed" fluorescence

Because GFP must undergo a posttranslational oxidation step to form its chromophore, it may not be detectable in certain situations. As mentioned, half-maximal fluorescence is observed in bacteria ~ 1.5 hr after exposure to atmospheric oxygen at room temperature. Similar measurements have not been made for GFP expressed in eukaryotic cells, but the reducing potential of cytoplasm may inhibit chromophore formation. When a GFP-cyclin fusion is expressed in COS-7 cells, there is a 3-hr lag

between the time it can be detected by immunofluorescence and the time it can be detected by *in vivo* fluorescence (Pines, 1995). Thus, if a fusion protein were to exhibit a rapid turnover rate *in vivo*, it would never exist in a fluorescent form. Based upon these observations, the use of GFP for a marker of transient gene expression may be infeasible, because the protein is not detectable immediately after translation.

Poor expression

A variety of GFP expression systems are now available. If fluorescence is not observed, however, it may be difficult to determine if a protein is being expressed efficiently. Because GFP cDNA is derived from jellyfish mRNA, it is not clear if the codons used are efficiently translated in other cell types. Furthermore, it is clear that the *A. victoria*-derived mRNA is not stable in all organisms. For instance, it has been shown that GFP mRNA is aberrantly spliced in *Arabidopsis thaliana* (Haseloff and Amos, 1995). Thus, the only reliable way to assay expression is by immunoblotting (UNIT 10.8). If a fusion protein is being expressed, antibodies raised against the fused protein (i.e., the original protein of interest) should recognize a band shifted by ~ 28 kDa. The authors have found that commercially available anti-GFP antibodies (Clontech) recognize fusion proteins in both immunoblotting and immunoprecipitation experiments. Thus, it should be possible to determine whether an observed lack of fluorescence is due to an actual lack of expression.

MUTANTS OF GFP

To ameliorate some of the problems discussed here, a variety of mutant forms of GFP have been synthesized. These mutants offer increased fluorescence intensity, shifted spectra, and more rapid posttranslational chromophore formation. Heim et al. (1995) and Delagrave et al. (1995) have found that a variety of mutations in and around the chromophore sequence can cause changes in the spectral characteristics of the protein. A mutation of Tyr⁶⁶ to histidine causes the protein to absorb long-wave ultraviolet light and emit blue light. However, the intensity is diminished by $\sim 50\%$ compared to that of the wild-type GFP (Heim et al., 1994). Mutation of Ser⁶⁵ to tyrosine (S65T) results in a protein that absorbs light maximally at 490 nm but continues to emit maximally at 509 nm, making it more suited to the spectra of a standard FITC filter set. Furthermore, the extinction coefficient of the S65T mutant is increased to $39,200 \text{ M}^{-1} \text{ cm}^{-1}$ at 490 nm exci-

tation (QY = 0.68) and the half-time of chromophore formation is reduced by ~66% (Heim et al., 1995). Such a mutant should be extremely useful, especially when a standard FITC filter set is used. The authors have determined that mutation of Val¹⁶³ to alanine (V163A) results in a GFP that is significantly more soluble in bacteria and seems to form its chromophore up to 25% faster than does wild-type GFP, but has the same spectral characteristics as the wild-type protein. A double mutant of S65T and V163A (which has the characteristic spectral shift of S65T) has been expressed in yeast and found to provide a signal that is significantly more detectable than that of wild-type GFP when a FITC filter set is used (see Fig. 9.7.6D).

MICROSCOPY SETUP

The authors routinely observe GFP fusion proteins in yeast using a standard microscope equipped for epifluorescence. Several aspects of the experimental setup have been optimized for detection of the GFP signal.

Objective lenses

Light intensity observed through a lens system may be described by the following relation (Hecht, 1987):

$$\text{intensity} \propto \frac{(\text{numerical aperture})^4}{(\text{magnification})^2}$$

Thus, for optimal signal, high-numerical-aperture (high-NA) objectives and the lowest possible magnification should be used. For example, using a 1.4-NA lens instead of a 1.3-NA lens of the same magnification results in nearly a 35% increase in intensity. Reduction in magnification from 100× to 60× (at constant NA) results in a roughly two-fold increase in intensity. For yeast, the authors have found that a 60× to 63× 1.4-NA lens with 10× eyepieces gives optimal signal and resolution for most applications. For mammalian cells (which are considerably larger), a 40×, 1.3-NA lens may be more appropriate.

Excitation sources

Most fluorescence microscopes are equipped with mercury arc lamps. In general, such lamps are available in 50-, 100-, and 200-W versions. However, there is not necessarily a direct correlation between power and intensity. Although a 200-W mercury lamp gives off more total light energy than a 100-W lamp, this is distributed over a much wider area than that emitted by the lower-power lamp. In practice, the overall intensity (total light output per unit

area) of a 100-W lamp is greater than that of a 200-W lamp. Furthermore, mercury arc lamps do not produce equal amounts of light over the entire spectrum; they tend to emit in strong “peaks” of high output at specific wavelengths. For instance, a mercury lamp provides most of its light in strong peaks at 337, 365, 405, 435, 546, and 577 nm (see Reichman, 1994). However, due to its high overall intensity, a 100-W mercury lamp is suitable for most routine epifluorescence applications.

Xenon arc lamps are also used as light sources for fluorescence microscopes. Overall, they produce much less light than mercury arc lamps at similar power levels. However, their light output is more spectrally uniform than that of the mercury sources. In practice, a 75-W xenon arc lamp will produce roughly as much (GFP-exciting) blue light as a 100-W mercury lamp (see Reichman, 1994). One benefit of the xenon lamp is that it produces less short-wavelength light than does a mercury lamp, which may minimize photodamage to the sample by stray light that passes fluorescence filters. The authors have found that a 75-W xenon arc lamp works extremely well for observing GFP fusions in yeast cells.

Often, lasers are utilized as light sources for microscopy (i.e., in confocal scanning microscopes). It has been reported that the 488-nm line of a krypton/argon laser is suitable for exciting GFP (Wang and Hazelrigg, 1994; Olson et al., 1995).

Fluorescence filters

Several types of filter sets are currently available that are appropriate for observing GFP. Before discussing these, it is important to delineate the three commonly used types of optical filters: band-pass, short-pass, and long-pass. Band-pass filters allow a finite range of light to pass through while excluding both longer and shorter wavelengths; short-pass filters exclude only light longer than a particular wavelength; and long-pass filters exclude only light shorter than a certain wavelength. Band-pass filters are typically used to specify excitation illumination, whereas both band-pass and long-pass filters are commonly used to specify emission spectra. Short-pass filters are sometimes used as excitation filters, but generally not for routine epifluorescence experiments.

Band-pass emission filters are generally more useful for observing a particular fluorophore when others are present or when nonspecific background fluorescence is present in the sample. Long-pass (LP) filters are typically

more useful when signal strength is low and maximal signal must be detected. The authors have used both standard FITC long-pass (excitation 450 to 490 nm, emission 520 LP) and FITC band-pass filters (excitation 460 to 500, emission 510 to 560; see Figs. 9.7.6A and D) for detecting GFP fluorescence in yeast. The authors have also used the GFP-specific (band-pass) filter set (excitation 400 to 440 nm, emission 480 to 520 nm; see Fig. 9.7.6C), but have observed that photobleaching of GFP is greatly accelerated when excitation is with this shorter-wavelength light. When a digital camera is used to record multiple images over extended periods of time, a FITC long-pass filter gives the most usable signal in terms of maximizing detectability and minimizing photobleaching.

Placed in between the excitation and emission filters is the third moiety of the filter set, known as the dichroic mirror (or beam splitter). This mirror serves to direct excitation (short-wavelength) light from the lamp to the sample and emission (long-wavelength) light from the sample to the eyepieces (or detector). Typically, the dichroic mirror is defined by a single wavelength: all light below this wavelength is reflected (to the sample), while all longer-wavelength light passes through the mirror (to the detector). Thus, the wavelength of a dichroic mirror is chosen to be between the longest excitation-filter wavelength and the shortest emission-filter wavelength. For example, when using the standard FITC long-pass filter set (excitation 450 to 490 nm, emission 520 LP), a dichroic mirror defined at a wavelength of 510 nm is used.

CONCLUSIONS

The existence of a completely proteinaceous fluorophore gives researchers a unique opportunity to observe dynamic cellular phenomena under physiological conditions. Because it requires no exogenous factors for fluorescence, the *A. victoria* green fluorescent protein has made such observations more feasible in a number of biological systems. The fact that GFP has been expressed and detected in a wide variety of organisms in both intact and (recombinantly) fused forms illustrates its versatility as a reporter molecule. Although difficulties have been reported in detecting GFP in some systems, advances in technology will certainly improve the situation. As spectrally enhanced mutants, expression vectors, antibodies, and GFP-specific filters become more widely available, the use of this unique in vivo fluorophore should become more common among a wide

variety of researchers. As GFP technology becomes more accessible, the end result should be a greater understanding of a variety of dynamic biological phenomena.

LITERATURE CITED

- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. 1994. Green fluorescent protein as a marker for gene expression. *Science* 263:802-805.
- Cody, C.W., Prasher, D.C., Westler, W.M., Pendergast, F.G., and Ward, W.W. 1993. Chemical structure of the hexapeptide chromophore of the *Aequorea* green-fluorescent protein. *Biochemistry* 32:1212-1218.
- Delagrave, S., Hawtin, R.E., Silva, C.M., Yang, M.M., and Youvan, D.C. 1995. Red-shifted excitation mutants of the green fluorescent protein. *BioTechnology* 13:151-154.
- Flach, J., Bossie, M., Vogel, J., Corbett, A.H., Jinks, T., Willins, D.A., and Silver, P.A. 1994. A yeast RNA-binding protein shuttles between the nucleus and the cytoplasm. *Mol. Cell. Biol.* 14:8399-8407.
- Haseloff, J. and Amos, B. 1995. GFP in plants. *Trends Genet.* 11:328-329.
- Hecht, E. 1987. Optics, 2nd ed. Addison-Wesley Publishing, Reading, Mass.
- Heim, R., Cubitt, A.B., and Tsien, R.Y. 1995. Improved green fluorescence. *Nature* 373:663-664.
- Heim, R., Prasher, D.C., and Tsien, R.Y. 1994. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A.* 91:12501-12504.
- Inouye, S. and Tsuji, F.I. 1994a. *Aequorea* green fluorescent protein: Expression of the gene and fluorescence characteristics of the recombinant protein. *FEBS Lett.* 341:277-280.
- Inouye, S. and Tsuji, F.I. 1994b. Evidence for redox forms of the *Aequorea* green fluorescent protein. *FEBS Lett.* 351:211-214.
- Kahana, J.A., Schnapp, B.J., and Silver, P.A. 1995. Kinetics of spindle pole body separation in budding yeast. *Proc. Natl. Acad. Sci. U.S.A.* 92:9707-9711.
- Morin, J.G. and Hastings, J.W. 1971. Energy transfer in a bioluminescent system. *J. Cell Physiol.* 77:313-318.
- Morise, H., Shimomura, O., Johnson, F.H., and Winant, J. 1974. Intermolecular energy transfer in the bioluminescent system of *Aequorea*. *Biochemistry* 13:2656-2663.
- Olson, K.R., McIntosh, J.R., and Olmstead, J.B. 1995. Analysis of MAP4 function in living cells using green fluorescent protein (GFP) chimeras. *J. Cell Biol.* 130:639-650.
- Pines, J. 1995. GFP in mammalian cells. *Trends Genet.* 11:326-327.
- Prasher, D.C., Eckenrode, V.K., Ward, W.W., Pendergast, F.G., and Cormier, M.J. 1992. Primary

- structure of the *Aequorea victoria* green-fluorescent protein. *Gene* 111:229-233.
- Reichman, J. 1994. Handbook of Optical Filters for Fluorescence Microscopy. Chroma Technology Corporation, Brattleboro, Vt.
- Wang, S. and Hazelrigg, T. 1994. Implications for *bcd* mRNA localization from spatial distribution of *exu* protein in *Drosophila* oogenesis. *Nature* 369:400-402.
- Ward, W.W. 1981. Properties of the coelenterate green fluorescent proteins. In *Bioluminescence and Chemiluminescence: Basic Chemistry and Analytical Applications* (M.A. DeLuca and W.D. McElroy, eds.) pp. 225-234. Academic Press, San Diego.
- Ward, W.W. and Bokman, S.H. 1982. Reversible denaturation of *Aequorea* green fluorescent protein: Physical separation and characterization of the renatured protein. *Biochemistry* 21:4535-4540.
- Ward, W.W., Cody, C.W., Hart, R.C., and Cormier, M.J. 1980. Spectrophotometric identity of the energy transfer chromophores in *Renilla* and *Aequorea* green-fluorescent proteins. *Photochem. Photobiol.* 31:611-615.
- Yocum, R.R., Hanley, S., West, R., and Ptashne, M. 1984. Use of LacZ fusions to delimit regulatory elements of the inducible divergent GAL1-GAL10 promoter in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 4:1985-1998.

Contributed by Jason A. Kahana and
Pam A. Silver
Dana-Farber Cancer Institute
Boston, Massachusetts

It is possible to transfect mammalian cells with a gene of interest and directly detect the RNA made from that gene 24 to 48 hr later (UNITS 9.1-9.4, 16.12, & 16.13; see Treisman, 1986, and Kingston et al., 1986, for examples). This type of analysis circumvents the need for fusion genes, but takes many more hours to do than an experiment done using fusion genes. The ability to analyze the RNA directly allows mutation and functional analysis of an intact gene, as the promoter does not have to be subcloned. This technique is also essential when fusion genes are used. In order to be sure that the level of the reporter protein produced by a fusion gene provides an accurate measure of appropriately initiated RNA from the promoter under study, it is necessary to determine the amount and 5' end of the fusion message. An investigator who verifies this using direct RNA analysis can proceed with some confidence that the level of the reporter protein is a measure of promoter activity.

The difficulty in directly analyzing RNA is that the sensitivity of detection is at the limits of present technology. Consequently, the transfection protocol, RNA preparation, and RNA analysis techniques all have to be working near optimum in order for an experiment to work. The cell type used must be readily transfectable, and the promoter being studied must be relatively efficiently expressed or the amount of RNA produced will be below the limits of detection. Direct analysis of RNA is several orders of magnitude less sensitive than the commonly used reporter gene systems. These difficulties make it clear why fusion genes containing a readily assayed reporter protein are presently so popular.

The following parameters should be considered when directly analyzing RNA after a transfection.

TRANSFECTION EFFICIENCY

It is imperative that the cell type and transfection protocol used result in efficient transfection. The efficiencies that result from CaPO₄-mediated transfection of HeLa, BALB/c 3T3, or NIH 3T3 cells are sufficient to do direct analysis of RNA expression after transient transfection. Optimization of transfection efficiency in the cell type used (introduction to Section I) should be done prior to attempting to analyze RNA from transfected genes.

Even after optimization, transfection efficiencies vary from day to day. One way to avoid the extensive workup of an experiment that is doomed to failure due to inefficient transfection is to include an easily assayed reporter gene in the transfection as an indicator of transfection efficiency. For example, an hGH expression plasmid can be included in each dish, and the level of secreted hGH can be rapidly measured at the time of harvest. The experiment can be aborted if the transfection efficiency is too low (we expect hGH levels of greater than 50 ng/ml 48 hr after transfecting 1 μg of pXGH5 into a 10-cm dish of BALB/c 3T3 cells). Alternatively, a reporter gene such as a CAT construct can be transfected into one or two plates and that activity determined, although use of a non-secreted reporter protein is more cumbersome due to the necessity of making a cell lysate.

RNA PREPARATION

RNA must be efficiently made in an intact form. The total cytoplasmic (UNIT 4.1) and guanadinitium isothiocyanate (UNIT 4.2) protocols are both well suited to this purpose. Two 10-cm dishes of adherent cells should produce 100 μg of total RNA (this number will be lower for particularly flat cell types). Expression from a transfected gene can be detected by analyzing 40 μg of total RNA. Sensitivity will be increased by using poly(A)-selected RNA (UNIT 4.5), but this is time-consuming. In addition, numerous dishes would have to be transfected to produce the amount of RNA required for the poly(A) selection.

ANALYSIS OF RNA

Both S1 analysis using single-stranded probes (UNIT 4.6) and ribonuclease protection analysis (UNIT 4.7) have the sensitivity necessary to detect gene expression at transient times after transfection. S1 analysis tends to have a lower background, allowing longer exposure times to be used. Ribonuclease protection analysis uses a probe of higher specific activity and thus should be significantly more sensitive if optimized. It can be more difficult to optimize the ribonuclease protection protocol, but it is worthwhile if the added sensitivity results in routine ability to detect expression from a transfected gene.

PROMOTER STRENGTH

Clearly, it is difficult to detect expression of RNA from an extremely weak promoter. If the promoter strength is more than an order of magnitude below that of, for example, the SV40 early promoter in HeLa cells, heroic efforts may be necessary in order to analyze the RNA. In such cases, poly(A) selection may be necessary as well as the use of body-labeled probes.

LITERATURE CITED

- Kingston, R.E., Cowie, A., Morimoto, R.I., and Gwinn, K.A. 1986. Binding of polyomavirus large T antigen to the human hsp70 promoter is not required for trans-activation. *Mol. Cell. Biol.* 6:3180-3190.
- Treisman, R. 1986. Transient accumulation of c-fos RNA following serum stimulation requires a conserved 5' element and c-fos 3' sequences. *Cell* 42:889-902.

Contributed by Robert E. Kingston
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

TRANSDUCTION OF GENES USING RETROVIRUS VECTORS

Overview of the Retrovirus Transduction System

A retrovirus vector is an infectious virus used to introduce a nonviral gene into mitotic cells *in vivo* or *in vitro*. Originating from replication-competent viruses isolated from rodents or chickens, the vectors are modified in various ways to serve in the transduction process. The efficient and precise integration machinery of naturally occurring retroviruses is utilized to produce either a single copy or a few copies of the viral genome stably integrated into a host chromosome.

Retroviral vectors are useful in achieving stable and efficient transduction of a gene or genes into cells that are not easily transfected, such as primary cells of various types and cells *in vivo* (which are usually not susceptible to transfection at all). Transfection refers to gene transfer (usually for expression) by processes that make a cell transiently permeable to DNA. Transfection methods include calcium phosphate- and DEAE-dextran-mediated transfection, and electroporation (*UNITS 9.1-9.3*). Not all cells can be transfected (especially primary cells), and the optimal transfection procedure varies for different cell types. Transfection methods may result in stable integration, in which the introduced DNA is integrated into the host chromosomal DNA, or transient transfection, in which there is no selection for integration and most of the introduced DNA has not integrated into the host chromosomes. In contrast to transfection, retroviral infection is dependent on binding to specific receptors on the cell surface. Some of these receptors, such as ecotropic and amphotropic receptors, are expressed on many cell types, including most primary cells, and as a result, it is possible to infect most cells with retroviruses. For the purposes of these units, transduction refers to the ability of a retroviral vector to insert itself into the host DNA as a provirus. Because most commonly used retroviral vectors are replication-defective, transduction usually refers to the ability of a replication-defective retroviral vector to undergo one round of infection before stable integration.

Retroviral transduction may not be the method of choice when working with cells that

can be efficiently transfected or in experiments in which high plasmid copy number is desirable (for transient expression protocols, see *UNITS 9.1-9.4* and *UNITS 16.13 & 16.15-16.18*). Because the size of retroviral RNA that can be efficiently packaged is limited, retroviral gene transfer is unable to transduce large (>11 kb) DNA fragments. Another limitation to retroviral transduction is that postmitotic cells cannot be transduced because mitosis is required for entry of the viral integration complex into the nucleus (Miller et al., 1990; Roe et al., 1993). New technologies that may circumvent some of these limitations are being developed—e.g., visna virus-based retroviral vectors (Naldini et al., 1996).

This unit presents an overview of the retrovirus life cycle and a description of vector designs and packaging cell lines (reviewed in Boris-Lawrie and Temin, 1993; Riviere and Sadelain, 1996; Miller, 1996). Production of retroviral stocks by either transient transfection (*UNIT 9.11*) or stable transduction (*UNIT 9.10*) is then detailed, with protocols for titering and assaying the resultant virus stock. For various applications, it is desirable to produce large quantities of the virus, to concentrate the virus stock (*UNIT 9.12*), and/or to check the stock for contamination by helper viruses (particularly when animals are to be infected; *UNIT 9.13*). Finally, outlines for infection of cells *in vivo* and *in vitro* are suggested (*UNIT 9.14*).

RETROVIRUS LIFE CYCLE

Although retroviral genomes are made of RNA, retroviruses produce a DNA copy of their genome immediately after infection of the host cell through a series of steps involving both retroviral and host factors (Fig. 9.9.1). The viral DNA then integrates (mediated by the viral *int* gene product) into the host genome where it is referred to as a provirus (Fig. 9.9.2). Most experimental retrovirus vectors were originally derived from cloned proviral sequences.

Infectious retroviral stocks are made by transfecting a retroviral vector plasmid into cell lines called packaging lines for replication-incompetent vectors (Fig. 9.9.3), or into host lines

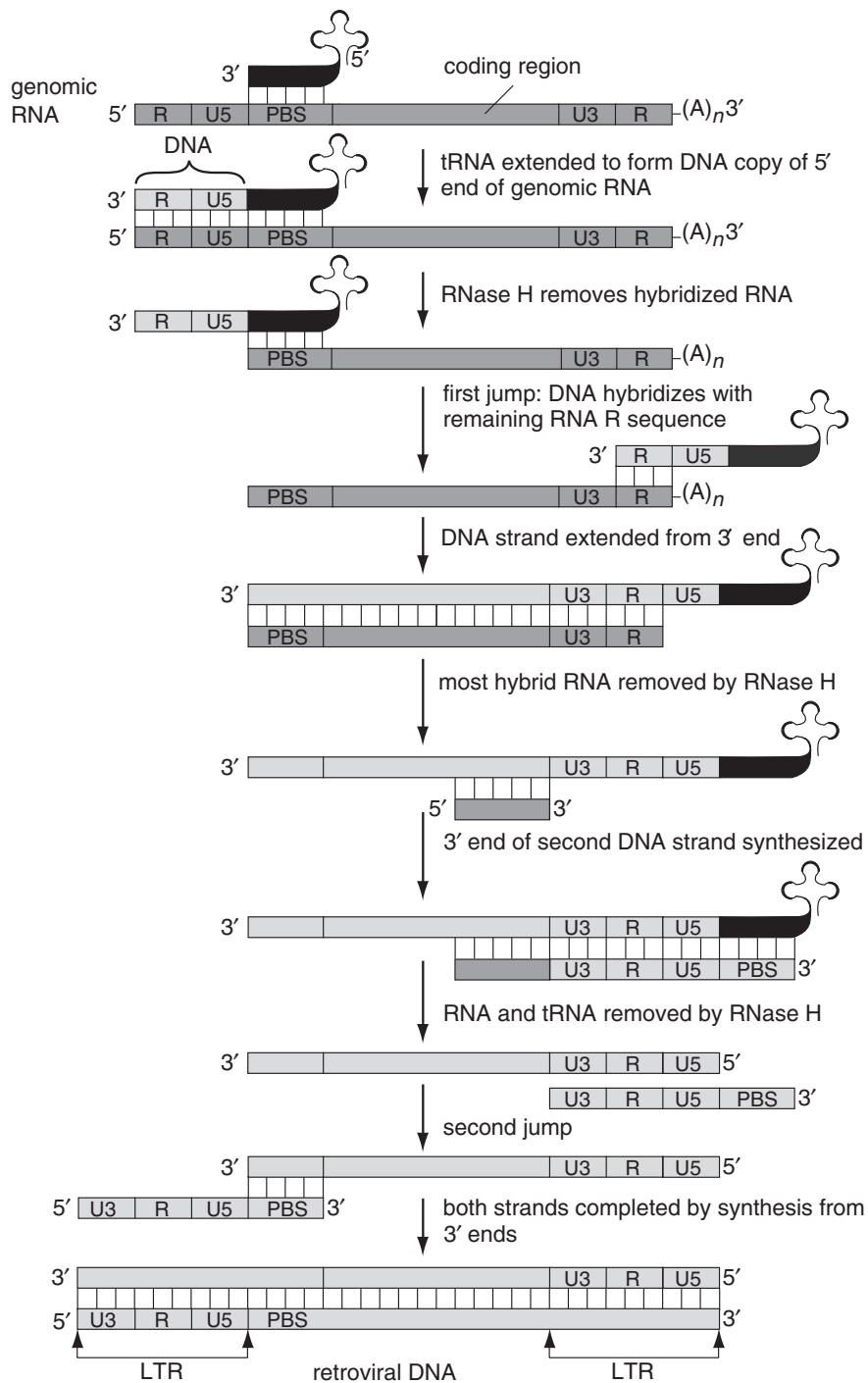


Figure 9.9.1 Generation of a double-stranded DNA copy of the RNA genome of a retrovirus. The genomic RNA is packaged in the virion with a retrovirus-specific cellular tRNA hybridized to a complementary sequence near its 5' end called the primer-binding site (PBS). The retroviral RNA has a short direct-repeat terminal sequence (R) at each end. The overall reaction is catalyzed by reverse transcriptase, which has both deoxynucleotide-polymerizing and RNase H activities and which digests the RNA strand in a DNA-RNA hybrid. The entire process yields a double-stranded DNA molecule that is longer than the template RNA and has a long terminal repeat (LTR) at each end. Note that the 5' LTR in the double-stranded DNA is derived from the 3' LTR. This has important implications for vector design and control of retroviral expression from transfected plasmids. (Adapted from *Molecular Cell Biology* by Darnell, Lodish, and Baltimore. Copyright 1986 by Scientific American Books, Inc. Used with permission of W.H. Freeman and Company.)

that support replication for replication-competent vectors (see discussion of Replication-Competent Vectors). Once in a cell line, transcription proceeds from the viral long terminal repeat (LTR) promoter encoded by the plasmid, generating an RNA viral genome. The viral genome is then encapsidated by viral structural

proteins, and infectious viral particles are produced by budding from the surface of cells (Fig. 9.9.3). The supernatant produced by such cells comprises a virus stock. The number of infectious viral particles per ml in a stock, referred to as colony forming units/ml (CFU/ml), is determined by a bioassay (UNITS 9.10 & 9.11). The

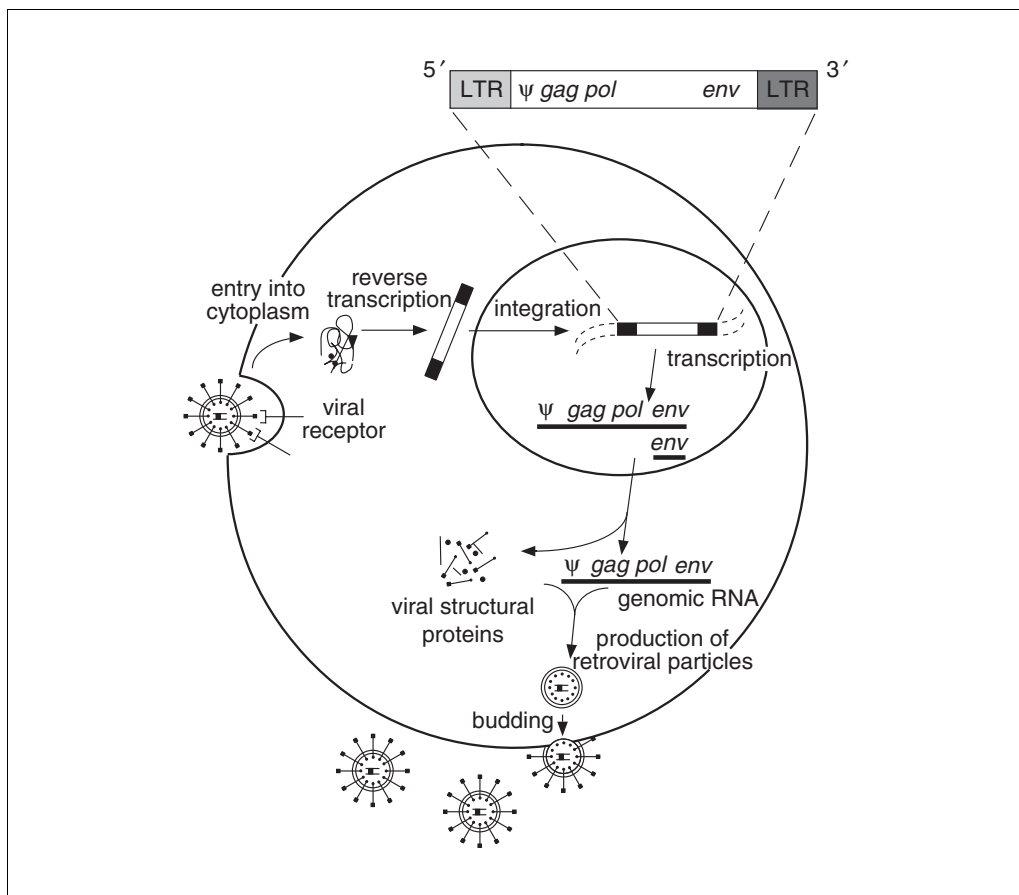


Figure 9.9.2 Life cycle of replication-competent retrovirus. All retroviruses infect host cells through an interaction with a specific viral receptor on the host-cell surface. They thereby enter the host cell and initiate reverse transcription of the RNA viral genome using reverse transcriptase contained within the viral particle. After synthesis of a linear duplex DNA copy of the viral genome, the viral DNA integrates into the host genome. Integration is mediated by the viral *int* gene product (encoded at the 3' end of the *pol* region) and usually occurs at the ends of the viral genome. Within the host genome, integration can occur at a very large number of chromosomal locations with little or no apparent specificity. One infectious virion produces one integrated copy of the viral genome. This “provirus” is subsequently replicated with the host DNA and is passed on to all progeny cells. After integration, the viral LTR promoter at the 5' end of the genome is usually active and directs synthesis of a full-length (i.e., unspliced) copy of the viral genome, which terminates in the 3' LTR.

In replication-competent viruses, the viral genome encodes the viral proteins *gag*, *pol*, and *env* required to make viral particles. The full-length viral transcript serves several functions. It is the mRNA for the *gag* proteins (which comprise the bulk of the viral particle), reverse transcriptase, and the *int* gene product, and it is the precursor RNA that is used to generate a spliced mRNA for the *env* protein, found on the surface of the viral particle. Finally, in its unspliced form, it is also the RNA that encodes Ψ, the recognition sequence for encapsidation. Because it contains the Ψ sequence, it can be recognized by the packaging proteins and become packaged into a viral particle. Due to this role, the unspliced full-length transcript is also referred to as the “genomic RNA.” Infection by a replication-competent virus thus leads to production of infectious retroviral particles identical to the original infecting particle. These particles are budded from the host cell and can go on to initiate more rounds of infection in other host cells.

bioassay, or “titration,” is carried out on a convenient target cell line that is easy to grow and is very susceptible to infection by a particular virus, such as NIH 3T3 cells for viruses with an ecotropic, amphotropic, or polytropic host range. The titered stock is then used to infect cells *in vivo* or *in vitro*. Once integrated into the host cell the vector provirus makes mRNA(s) for the gene(s) of interest (Fig. 9.9.3).

REPLICATION-INCOMPETENT VECTORS

Replication-incompetent retrovirus vectors were derived from proviruses that were cloned from cells infected with a naturally occurring retrovirus. Deletion of some or all of the genes encoding virion structural proteins was carried out to create room for inserted genes and to cripple the virus (see Fig. 9.9.4). Although extensively deleted, the vectors retain the *cis*-

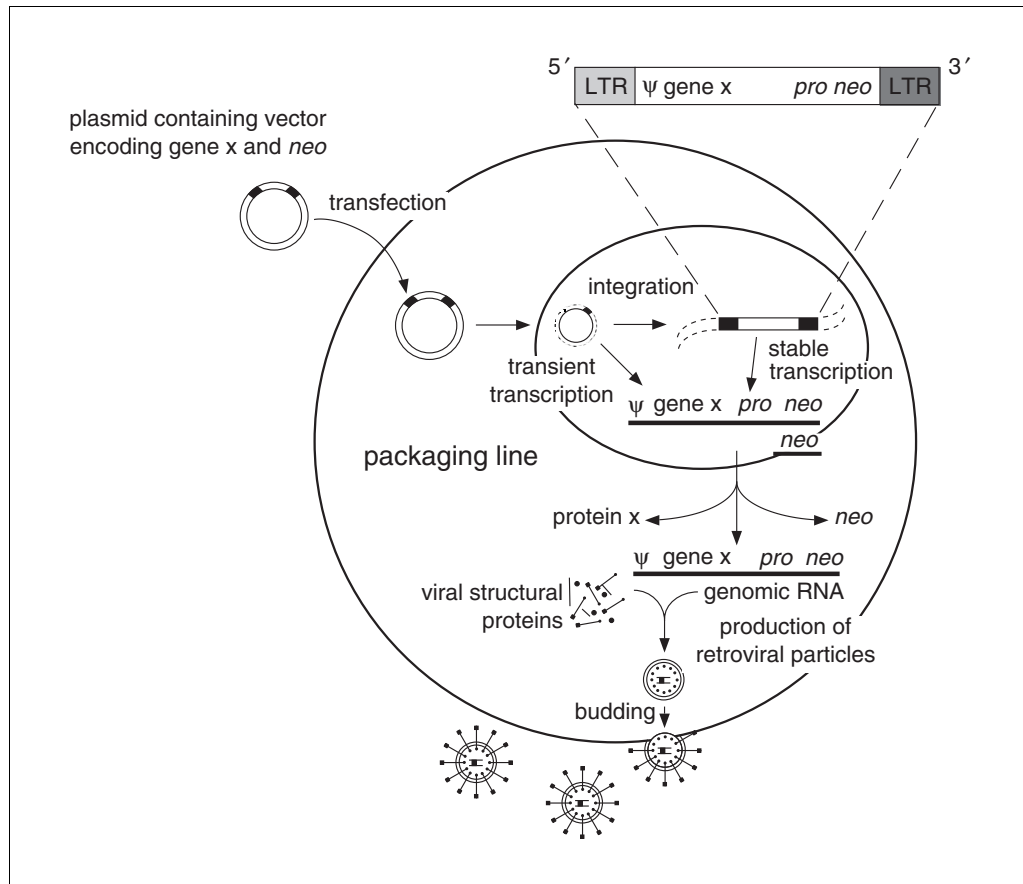


Figure 9.9.3 Production of an infectious retroviral particle from a bacterial plasmid that encodes a retroviral genome requires the introduction of the bacterial plasmid into mammalian or avian cells. This is true for both replication-competent and replication-incompetent vectors. In the example shown here, a replication-incompetent vector encoding *neo* and a gene of interest (*x*) are introduced, via CaPO_4 -mediated transfection, into a packaging line. The packaging line is a mouse or quail cell line (see Table 9.9.1) that encodes the proteins necessary for production of viral particles (i.e., gag, pol, and env). The bacterial plasmid can be transcribed either transiently (for a few days after transfection) from nonintegrated plasmid molecules or stably from integrated plasmid molecules. The viral transcript is initiated in the 5' LTR and terminated in the 3' LTR, and is thus a full-length viral transcript. It contains the packaging sequence Ψ , which is recognized by the capsid proteins and allows it to be packaged into viral particles. A fully infectious viral particle containing the vector genome is thus budded from the packaging cell. The culture supernatant is removed from cells and used as the source of virus for future experiments (e.g., see Fig. 9.9.4).

Also shown in this figure is the production of another transcript, initiated at an internal promoter (*pro*) and encoding the *neo* gene. The *neo* transcript produced by the internal promoter does not contain Ψ and thus does not become encapsidated. It will, however, serve as an mRNA for *neo*, thereby allowing for growth of these cells in the selective drug G418. As described in *UNIT 9.10*, this drug-resistance feature can be exploited to select cells that have been stably transfected with the retroviral plasmid.

acting viral sequences necessary for transmission. These sequences include the Ψ packaging sequence (necessary for recognition of the viral RNA for encapsidation into the virus particle), reverse transcription signals, integration signals, and viral promoter, enhancer, and polyadenylation sequences (Fig. 9.9.4). The viral promoter, enhancer, and polyadenylation signals are contained in the retroviral LTR. The former two elements, together with the retroviral envelope, determine the cell-type specificity of the virus. For example, Moloney murine leukemia virus (MoMuLV), whose LTR is the most common one used for retroviral vectors, is expressed in most cell types. Murine stem cell virus (MSCV), a derivative of murine embryonic stem cell virus (MESV; Hawley et al., 1994), is reported to give better expression in stem cells, whereas spleen focus-forming virus

(SFFV) gives better expression in erythroid cells (Boris-Lawrie and Temin, 1993).

Using the transcription regulatory sequences provided by the virus, any cDNA can be expressed in the vector. Alternate constructs that lack the viral promoter and/or enhancer have also been made. These constructs are used for applications such as studies of alternative promoter regulation, or when trying to avoid the repression of transcription that sometimes can occur when using the viral LTR. Recently, retroviral vectors have been described that replace the viral promoter/enhancer with inducible elements (Hofmann et al., 1996; Paulus et al., 1996), such as the tetracycline-responsive elements (Gossen and Bujard, 1992; reviewed in Shockett and Schatz, 1996). These vectors may circumvent some of the difficulties in expressing toxic genes, as well as providing the

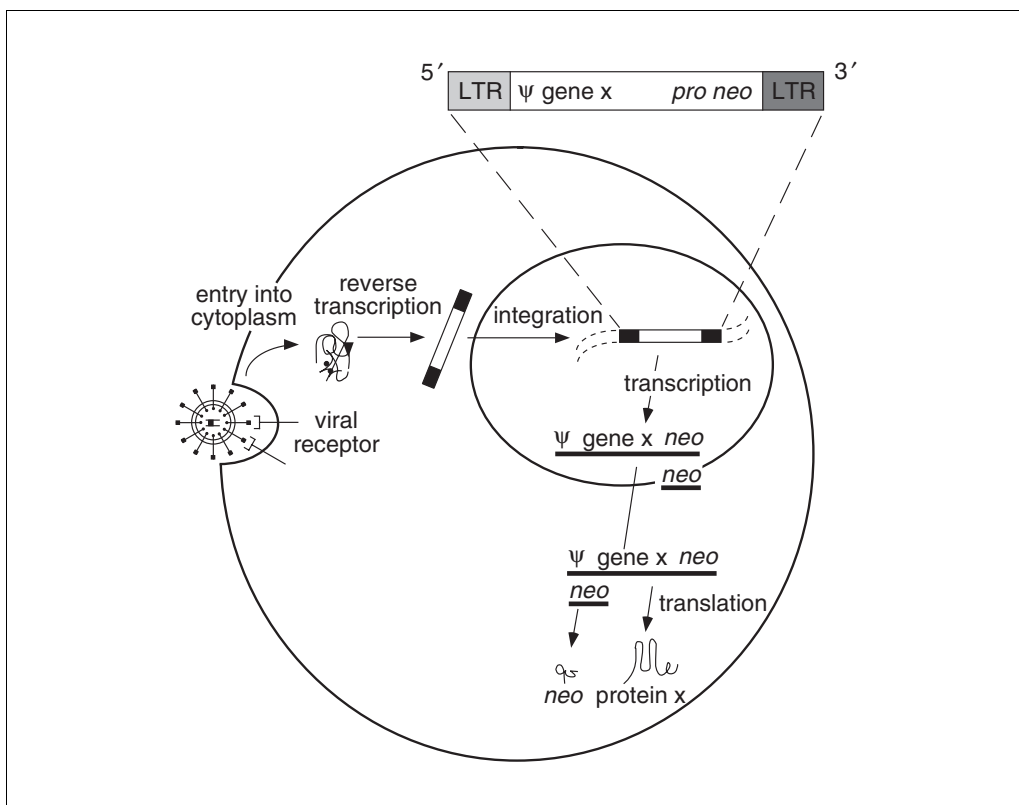


Figure 9.9.4 Infection of a target cell and expression by a replication-incompetent retrovirus vector occurs through interaction with a host-cell receptor. The virus then reverse transcribes and integrates, just like a replication-competent virus, as described in Figure 9.9.2. A full-length viral transcript can also be initiated in the 5' LTR and end in the 3' LTR, exactly as in the case of the replication-competent virus. However, the viral transcript typically does not encode any of the proteins required to make a viral capsid. It usually encodes a gene of interest (x). In addition, many replication-incompetent vectors encode another promoter (pro) within the viral genome for synthesis of an mRNA for another gene (e.g., neo). Because such a vector does not encode viral capsid proteins, cells infected with a replication-incompetent vector cannot make more viral particles. Thus, there is no spread of the viral genome from an infected cell to other cells. However, the viral genome does pass to all progeny cells through the normal processes used by the host to pass on all host genes.

ability to regulate gene expression from retroviral vectors. One drawback of using retroviral vectors that lack the viral promoter and/or enhancer is that retroviral titers drop by several orders of magnitude (Yu et al., 1986).

Vector designs are continually evolving, and it is best to obtain vectors from a laboratory that has experience with several different designs and that will advise on their use. There are many more murine replication-incompetent vectors available than avian ones. (Perhaps because more effort has gone into the design and use of murine replication-incompetent vectors, murine versions are also in general more successful than their avian counterparts.) In many instances, a vector will have to be modified by cloning a gene(s) of interest into an appropriate site, and subsequently introducing the recombinant vector into an appropriate packaging line (UNITS 9.10 & 9.11). Other vectors will be ready for use without any additional cloning or transfection steps. Examples of several different types of retroviral vectors that have been successfully used for gene transfer are shown in Figures 9.9.5 and 9.9.6. A large number of retroviral vectors have been developed and a review of all of these is beyond the scope of this chapter. For a more detailed explanation of individual retroviral vectors, the reader should consult the specific references.

REPLICATION-COMPETENT VECTORS

Vectors that retain all of the genes encoding the virion structural proteins, as well as all of the cis-acting viral elements necessary for transmission, are "replication-competent." Vectors of this type have been designed to carry and express additional, nonviral sequences. This frequently has been successful for genes ≤ 2 kb using avian replication-competent vectors (Hughes et al., 1987), but few successes have been reported for the murine replication-competent vectors. Replication-competent vectors have not been as popular as replication-incompetent vectors due to the greater popularity of mammals as experimental organisms. However, when working with avian species, they are very useful when it is desirable to transduce nearly all of the cells in an animal or in a tissue culture dish.

PACKAGING LINES AND VIRUS PRODUCTION

Replication-incompetent retrovirus vectors do not encode the structural genes whose products comprise the viral particle (see Fig. 9.9.4). To produce infectious viral particles from a

retrovirus plasmid, the viral structural proteins encoded by *gag*, *pol*, and *env* are supplied in *trans* by packaging cell lines (Fig. 9.9.3 and Table 9.9.1). These cells may be derived from a variety of cell lines that contain the viral *gag*, *pol*, and *env* genes as a result of introduction by transfection. The most commonly used cells are human 293 cells, primate COS cells, mouse NIH 3T3 fibroblasts, and quail QT6 cells. In contrast to retroviral vectors, the viral RNA that encodes the structural proteins does not contain the retroviral packaging sequence Ψ . Thus, the packaging lines do not efficiently package the RNA that encodes the genes *gag*, *pol*, or *env*. Prior to the introduction of vector DNA into these lines, cellular RNAs are randomly encapsidated and budded as normal viral particles. To produce viral particles that contain the vector genome, the vector DNA is introduced into the packaging cell line. The vector RNA is preferentially packaged because it contains the packaging sequence Ψ (Fig. 9.9.3).

Until recently, most packaging cell lines were developed using NIH 3T3 cells and used for the creation of stable retroviral producer cell lines. Stable retroviral producer cell lines contain a single (or few) integrated provirus and are capable of long-term production of retroviral supernatants. Although stable retroviral producer cell lines are able to produce high-titer retroviral supernatants expressing some cDNAs, with many retroviral constructs it has proven difficult to obtain high-titer retroviral supernatants. This is thought to result from outgrowth of clones that express low levels of retrovirus during the prolonged selection process necessary for creation of stable cell lines (see Fig 9.9.7). To overcome this problem, several groups have developed strategies based on transient retroviral production (Landau and Littman, 1992; Pear et al., 1993; Finer et al., 1994; Soneoka et al., 1995; Naviaux et al., 1996; see Fig. 9.9.7 and UNIT 9.11). These strategies, based on the use of either COS cells or 293 cells, have proven very successful for the production of high-titer retroviral supernatants from a large number of cDNAs, including some that were unable to produce high viral titers in NIH 3T3-based packaging lines.

As with naturally occurring retroviruses, packaged retrovirus vectors enter the host cell via interaction of a viral envelope glycoprotein (a product of the viral *env* gene) with a host-cell receptor (Table 9.9.2). The murine and avian viruses have several classes of envelope glycoprotein that interact with different host-cell receptors. Packaging lines that express various

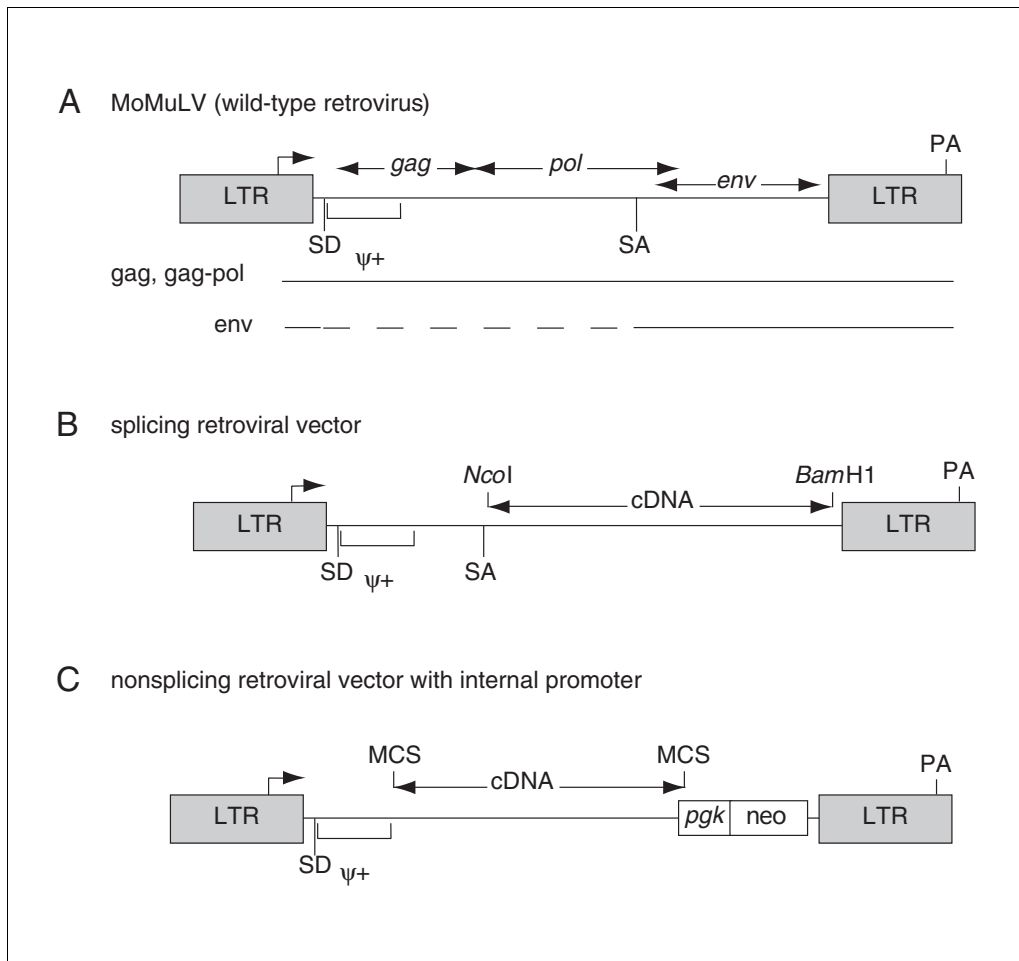


Figure 9.9.5 Overview of some recent developments in retroviral vector design. **(A)** Structure of wild-type Moloney murine leukemia virus (MoMuLV). Shown are the retroviral long terminal repeat (LTR), consisting of U3, R, and U5 regions (transcription initiation begins in the R region); splice donor (SD) and splice acceptor (SA) sites; extended retroviral packaging site (Ψ^+); and retroviral structural genes *gag*, *pol*, and *env*. Because Ψ^+ extends into *gag* (Bender et al., 1987), to prevent translation of *gag* product most retroviral vectors that use Ψ^+ contain mutations in the *gag* start codon. The *gag* and *pol* messages are expressed from an unspliced transcript, whereas the *env* message is spliced as shown below the vector. **(B)** Structure of a splicing retroviral vector. These vectors express the cloned DNA from a spliced message, as though it were the *env* gene. The example shows MFG, a vector that tries to closely emulate *env* expression for the cloned cDNA by expressing it from a spliced transcript and cloning it into an *Nco*I restriction site so that the translation initiation site is identical to *env* (Dranoff et al., 1993). **(C)** Structure of a nonsplicing retroviral vector that also contains an internal promoter. These vectors express the cloned cDNA as an unspliced message. There is also an internal promoter that expresses a marker for antibiotic resistance, which may be expressed as either an unspliced or a spliced message. One potential drawback to this strategy is competition between the internal promoter and the retroviral LTR. Because selection is frequently based on the marker and not the gene of interest, it is not uncommon to obtain drug-resistant colonies that do not express the gene of interest. This particular example shows murine stem cell virus (MSCV), which in addition to the above features contains a multiple cloning site (MCS) to facilitate insertion of the cDNA, an LTR from murine stem-cell virus to enhance expression in stem cells, and other modifications to enhance viral titer and gene expression (Hawley et al., 1994).

Table 9.9.1 Retroviral Packaging Cells

Cell line ^a	Drug resistance genes ^b	Helper virus production	Cell type	Features ^c	Source ^d	Reference
<i>Murine ecotropic</i>						
Ψ-2	<i>gpt</i>	Yes	NIH 3T3	—	—	Mann et al. (1983)
Ψ-CRE	<i>hph, gpt</i>	No	NIH 3T3	—	—	Danos and Mulligan (1988)
GP+E-86	<i>gpt</i>	No	NIH 3T3	—	—	Markowitz et al. (1988a)
ΩE	<i>gpt</i>	No	NIH 3T3	—	—	Morgenstern and Land (1990)
ampli-GPE	<i>neo</i>	No	NIH 3T3	—	—	Takahara et al. (1992)
COS, SV-Ψ-E-MLV ^e	None	No	COS	Titers ~10 ⁵ CFU/ml and dependent on amplification	—	Landau and Littman (1992)
Bosc23	<i>neo, hph, gpt</i>	No	293	Retroviral packaging cell line for transient production	ATCC CRL 11270 ^f (293T/17, ATCC CRL 11268; Anjou, ATCC CRL 11269)	Pear et al. (1993)
φNX-eco	<i>neo</i>	No	293	Retroviral packaging cell line for transient production; <i>gag-pol</i> construct has an IRES-CD8 surface marker to follow gag-pol expression	—	P. Achacoso and G. Nolan (unpub. observ.)
293T, pHIT60, pHIT123 ^e	<i>neo</i>	No	293	—	ATCC CRL 11268 ^f	Soneoka et al. (1995)
pCL-eco ^e	None	No	293	—	—	Naviaux et al. (1996)
<i>Murine amphotropic</i>						
Ψ-Am	<i>gpt</i>	Yes	NIH 3T3	—	—	Cone and Mulligan (1984)
PA317	<i>tk</i>	Rare	NIH 3T3	—	ATCC CRL 9078	Miller and Buttimore (1986)
Ψ-CRIP	<i>hph, gpt</i>	No	NIH 3T3	—	—	Danos and Mulligan (1988)
GP+env/Am12	<i>hph, gpt</i>	No	NIH 3T3	—	—	Markowitz et al. (1988b)

continued

Table 9.9.1 Retroviral Packaging Cells, continued

Cell line ^a	Drug resistance genes ^b	Helper virus production	Cell type	Features ^c	Source ^d	Reference
FLYA4	<i>bsr, bleo</i>	No	HT1080 (human fibrosarcoma)	Resistant to human serum	—	Cosset et al. (1995)
Bing	<i>hph, gpt</i>	No	293	Releases 100 CFU/ml hph ^r retrovirus per ml retroviral supernatant	ATCC CRL 11554 ^f	Pear et al. (1996)
ϕNX-Am	<i>neo</i>	No	293	Amphoteric counterpart to ϕNX-eco	—	P. Achacoso and G. Nolan (unpub. observ.)
293, kat ^e	None	No	293	—	—	Finer et al. (1994)
pCL-ampho ^e	None	No	293	—	—	Naviaux et al. (1996)
<i>GALV</i>						
PG13	<i>tk, dhfr</i>	No	3T3	—	ATCC CRL 10686	Miller et al. (1991)
VSV						
293 gag/pol+CMV VSV-G ^g	<i>dhfr</i>	No	293	Supernatants can be concentrated by centrifugation	—	Burns et al. (1993)
GP7C-tTA-G10	<i>hph, pac</i>	No	3T3	—	—	Yang et al. (1995)
<i>Other</i>						
RD114	<i>bsr, bleo</i>	No	HT1080	Feline virus that uses the simian retrovirus receptor on human cells; virions resistant to human serum	—	Cosset et al. (1995)
10A1	<i>tk, dhfr</i>	No	NIH 3T3	Capable of infecting by either amphoteric or GALV receptor; in human cells, the amphoteric receptor is more efficiently utilized	—	Miller and Chen (1996)

continued

Table 9.9.1 Retroviral Packaging Cells, continued

Cell line ^a	Drug resistance genes ^b	Helper virus production	Cell type	Features ^c	Source ^d	Reference
pCL-10A1 ^e	None	No	293	Capable of infecting by either amphoteric or GALV receptor; in human cells, the amphoteric receptor is more efficiently utilized	—	Naviaux et al. (1996)
293T, pCMVΔR9, pHR ^{e1,h}	<i>neo</i>	No	293	HIV-based vectors, titers ~10 ⁵ CFU/ml for mitotic cells; able to infect neurons and other non-mitotic cells at lower efficiencies	—	Naldini et al. (1996)
<i>Avian</i> Q2bn	—	Yes	—	Works well for transient expression but does not give high titers for stable production; uses subgroup A env	—	Stoker and Bissell (1988)
Isolde	—	No	—	Gives good titers from stably transfected lines; uses subgroup A env	—	Cosset et al. (1990)

^aAbbreviations: CMV, cytomegalovirus; GALV, gibbon ape leukemia virus; IRES, internal ribosome entry site; RD114, RD114 endogenous cat retrovirus; SV, simian virus; VSV, vesicular stomatitis virus; 10A1, 10A1 murine leukemia virus.

^bDrug resistance genes that are present in the packaging cells as a result of coinfection with the retroviral packaging constructs. These markers cannot be used for subsequent selection. *gpt*, xanthine-guanine phosphoribosyltransferase; *tk*, herpes simplex virus thymidine kinase; *lpph*, hygromycin phosphotransferase; *diffr*, a mutant dihydrofolate reductase gene; *ble*, a bacterial gene that confers resistance to bleomycin and phleomycin in mammalian cells; *pac*, puromycin *N*-acetyl phosphotransferase; *bsr*, a bacterial gene that confers resistance to blasticidin S; and *neo*, neomycin phosphotransferase.

^cExcept where noted, all lines are capable of titers >10⁶ CFU/ml.

^dUnless otherwise noted, contact the authors of the specified reference.

^eThese are not packaging cell lines. Rather than being stably introduced into the 293 cells, the retroviral structural genes are cotransfected with the retroviral vector.

^fTo receive these cell lines, obtain a Material Transfer Agreement from and return it to: Office of the General Counsel, Attn: Teresa L. Solomon, Esq., The Rockefeller University, 1230 York Ave., New York, NY 10021-6399, (212) 327-7598, Fax (212) 327-7688.

^gBecause of its cytotoxicity, VSV G protein is not stably integrated into the gag-pol-producing cell line.

^hIn the original report, the retroviruses were pseudotyped with both MLV-amphotropic envelope and VSV G protein.

viral envelope proteins are listed in Table 9.9.1, along with a summary of some of their properties. The abilities of these viruses to transduce mouse and/or human cells is listed in Table 9.9.2. It should be noted that a given murine vector can be encapsidated by packaging proteins produced by ecotropic, amphotropic, or polytropic cell lines. The gag and pol proteins encoded by the ecotropic and amphotropic packaging lines are virtually identical. The only significant differences among the proteins encoded by the various lines are in the env protein.

MURINE RETROVIRUSES

Ecotropic Receptors

The most commonly used type of murine viral glycoprotein is the ecotropic type, which binds to the *ecotropic receptor* found only on

rat and mouse cells. The ecotropic receptor has been cloned (MCAT-1; Albritton et al., 1989) and is a cationic transporter (Kim et al., 1991; Wang et al., 1991). The ecotropic glycoprotein, gp70, does not allow infection of human cells, and thus viruses bearing ecotropic receptors are considered relatively safe for gene-transfer experiments. The Ψ 2 line (Mann et al., 1983) was one of the most commonly used packaging cell lines of this class. Although capable of high-titer retrovirus production, this cell line can also produce helper virus (*UNIT 9.13*), a replication-competent virus (RCV) that can promote the horizontal spread of marker viruses. Two more recently developed ecotropic packaging lines— Ψ CRE (Danos and Mulligan, 1988) and GP+E-86 (Markowitz et al., 1988a)—have not shown helper-virus production to date and have been widely used for the creation of retro-

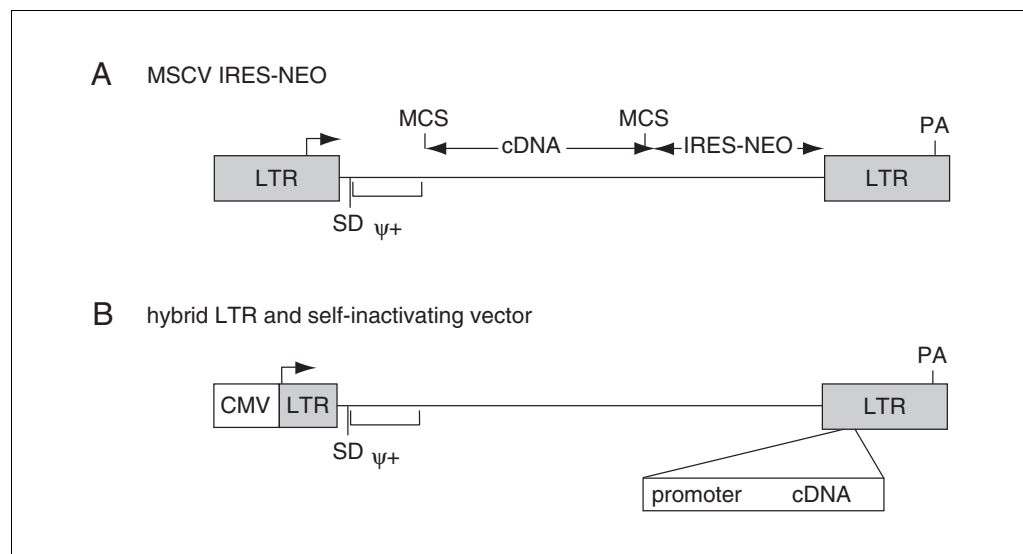


Figure 9.9.6 Expression elements in retroviral design. **(A)** Structure of MSCV IRES-NEO, a retroviral vector that uses an internal ribosomal entry site (IRES) from murine stem cell virus (MSCV) to express a second gene. An alternative to expressing a second gene from an internal promoter is to use an IRES. An IRES provides *cis*-acting sequences that allow for cap-independent translation initiation at internal initiation codons (reviewed in Boris-Lawrie and Temin, 1993). When used in a retroviral vector, this may alleviate some of the problems associated with promoter competition (Ghattas et al., 1991). Use of an IRES results in a polycistronic message that is translated independently. It is possible to use multiple IRES constructs to express more than two genes; compared to the gene closest to the LTR, all subsequent genes are expressed at 50% of this level (Zitvogel et al., 1994). IRES constructs that have been successfully used in retroviral vector have been cloned from poliovirus or encephalomyocarditis virus (EMCV). **(B)** Structure of a retroviral vector that contains a hybrid 5' LTR and is also a self-inactivating vector. Because the 5' LTR drives expression of the retroviral RNA destined to become encapsidated but does not become the 5' LTR in the provirus (see Fig. 9.9.1), it is possible to replace nonessential U3 sequences in the 5' LTR with heterologous sequences. For example, vectors that fuse the cytomegalovirus (CMV) promoter to the R-U5 region of the MoMuLV LTR are associated with higher titers in 293 cells (Finer et al., 1994; Soneoka et al., 1995; Naviaux et al., 1996). It is also possible to insert genes into the 3' U3 region outside the retroviral transcription unit. Following reverse transcription this region will be copied in both LTRs, allowing for expression of the cloned DNA from a cloned promoter. Although this strategy is associated with lower retroviral titers, it has recently been successfully used for the creation of tetracycline-regulatable retroviruses (Hofmann et al., 1996; Paulus et al., 1996).

Table 9.9.2 Host Range of Retroviral Vectors^a

Vector pseudotype	Cells that can be transduced	
	Mouse	Human
Ecotropic	Yes	No
Amphotropic	Yes	Yes
GALV	No	Yes
VSV	Yes	Yes
RD114	No	Yes
10A1	Yes	Yes

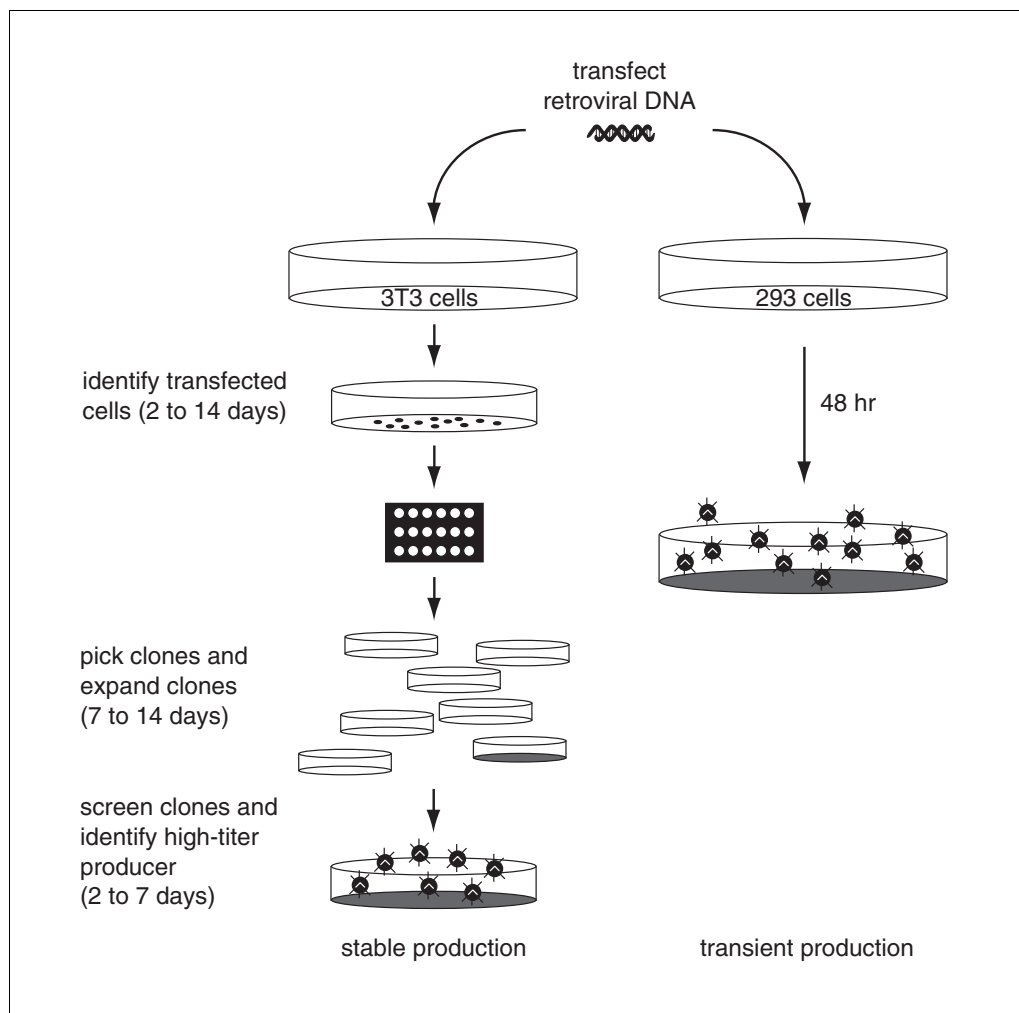
^aData kindly provided by A. D. Miller.

Figure 9.9.7 Comparison of methods used for production of retroviral supernatants by stable (left) and transient (right) producer cell lines. In both methods, the first step is introduction of the retroviral plasmid DNA into the packaging cell line. Because packaging cell lines used for stable production are derived from NIH 3T3 cells, very few cells in the population release high-titer retroviral supernatants (see *UNIT 9.11*); it is therefore necessary to identify these cells. This is usually accomplished in two steps. First, those cells that have been successfully transfected are identified and subsequently tested for high-titer retroviral production. Under optimal circumstances, stable retroviral producer lines can be isolated after 2 to 3 weeks of selection and screening. In the transient methods, it is possible to harvest high-titer retroviral supernatants 48 hr after transfection. The high efficiency of 293-based packaging cell lines is thought to be due to both their high transfection efficiency and their ability to express transfected DNA at high levels.

viral producer cell lines. Several alternative transient transfection approaches give equivalent results for production of high-titer, helper-free viruses with ecotropic host range by transient transfection. Either the retroviral vector is introduced into the 293-derived packaging cell line, Bosc23 (Pear et al., 1993), which contains stably integrated cDNAs expressing the retroviral structural genes; or the vector is transiently introduced, together with the packaging constructs, into 293 cells (Soneoka et al., 1995; Naviaux et al., 1996).

Amphotropic Receptors

The *env* genes of the amphotropic class of viral glycoproteins endow a murine virus with a very broad host range, including rodent, mouse, human, chicken, dog, cat, and mink. The gene for the amphotropic virus receptor, Ram1, which localizes to human chromosome 8, has been cloned and encodes a sodium-dependent phosphate symporter (Kavanaugh et al., 1994; Miller et al., 1994; van Zeijl et al., 1994). Packaging cell lines, similar to those developed for production of retroviruses with an ecotropic host range, have been created for both stable and transient production of amphotropic retroviruses. Commonly used NIH 3T3-based packaging lines for the stable production of viruses with an amphotropic host range include PA317 (Miller and Buttimore, 1986), Ψ CRIP (Danos and Mulligan, 1988), and GP+Am (Markowitz et al., 1988b). Production of low levels of replication-competent virus has been reported for PA317 cells, whereas helper virus production has not been reported for either the Ψ CRIP or GP+Am cell lines. Strategies for producing high titer by transient transfection have also been described utilizing either a 293-packaging cell line, Bing (Pear et al., 1996), or cotransfection of the retroviral vector and structural proteins into 293 cells (Finer et al., 1994; Naviaux et al., 1996). Although neither of the transient strategies releases helper virus, the Bing cells release 100 infectious units/ml of a retrovirus that encodes hygromycin phosphotransferase resistance (Pear et al., 1996). An amphotropic packaging cell line similar to Bing, ϕ NX-A does not release hygromycin-resistant retrovirus (P. Achacoso & G. Nolan, pers. comm.).

A receptor that shares close homology with the receptor for murine amphotropic retroviruses is the receptor for gibbon ape leukemia virus (GALV; Miller et al., 1994; van Zeijl et al., 1994). The GALV receptor

(GLVR1) is located on human chromosome 2 and, similar to the amphotropic receptor, is a sodium-dependent phosphate symporter (Miller et al., 1994; van Zeijl et al., 1994). GLVR1 is also the receptor for simian sarcoma-associated virus (SSAV) and feline leukemia virus subgroup B (Takeuchi et al., 1992). Closely related to the amphotropic and GALV envelopes is the 10A1 envelope. Viruses that contain the 10A1 envelope can enter cells by both the amphotropic and GALV receptors (Miller and Miller, 1994a).

Although these vector systems work well in vitro, their potential in vivo use for human gene therapy is limited by their rapid inactivation by human serum. Packaging cell lines derived from the human fibrosarcoma cell line HT1080 (Cosset et al., 1995) produce retroviruses that are resistant to human complement (especially when pseudotyped with the feline virus RD114 envelope).

The expression of a specific receptor differs among both species and cell types. For optimal infection efficiency, it is important that the retrovirus express the appropriate envelope protein.

Polytropic and Other Receptors

Receptors other than the ecotropic and amphotropic receptors can be utilized for retroviral entry. Pseudotyping the retroviral envelope with the vesicular stomatitis virus (VSV) G protein (Burns et al., 1993; reviewed in Yee et al., 1994) endows viruses of one class with characteristics of another class. VSV G protein-pseudotyped viruses enter cells by interacting with widely expressed cell membrane components and thus are able to infect a wide variety of cell types, including insects, fish, frogs, and humans. Although VSV G protein is toxic to cells, it is possible to transiently express this protein in 293 cells with a retroviral vector and obtain high-titer retroviral supernatants that use VSV G protein for cell entry (Burns et al., 1993).

Retroviral pseudotyping can also be done using cell-surface molecules (Kasahara et al., 1994; Han et al., 1995) or antibodies against cell-surface molecules (Somia et al., 1995; Marin et al., 1996). To date, retroviral titers obtained using these approaches have been low; however, they offer the hope that it will be possible to target retroviral entry. Also in development are packaging lines based on lentiviruses, which offer the prospect of transducing nonmitotic cells (Naldini et al., 1996).

AVIAN RETROVIRUSES

Avian retroviruses can use six different subgroups of *env*: A, B, C, D, E, and F. Initially it was thought that each *env* subgroup used a different receptor on avian cells. However, definition of the relationship between each avian *env* subgroup and its receptor is presently not clear. Two avian leukemia virus packaging lines—Q2bn (Stoker and Bissell, 1988) and Isolde (Cosset et al., 1990)—are used for production of avian replication-incompetent vectors. They were made using the same strategies employed in constructing Ψ 2 and Ψ CRE (see Ecotropic Receptors), respectively. These avian lines employ the avian subgroup A envelope, which does not allow infection of mammalian cells; viruses produced by these lines are therefore considered relatively free of hazard to the investigator.

SAFETY ISSUES

There are important biological safety issues that should be carefully considered when working with retrovirus vectors. These issues will vary with a given experiment, so detailed protocols should be discussed with the appropriate biological safety committee at the experimenter's company or institution. In particular, extreme precautions and containment levels may be necessary when working with amphotropic and polytropic viruses.

LITERATURE CITED

- Albritton, L.M., Tseng, L., Scadden, D., and Cunningham, J.M. 1989. A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection. *Cell* 57:659-66.
- Bauer, T.R. Jr., Miller, A.D., and Hickstein, D.D. 1995. Improved transfer of the leukocyte integrin CD18 subunit into hematopoietic cell lines by using retroviral vectors having a gibbon ape leukemia virus envelope. *Blood* 86:2379-2387.
- Bender, M.A., Palmer, T.D., Gelinis, R.E., and Miller, A.D. 1987. Evidence that the packaging signal of Moloney murine leukemia virus extends into the *gag* region. *J. Virol.* 61:1639-1646.
- Boris-Lawrie, K.A. and Temin, H.M. 1993. Recent advances in retrovirus vector technology. *Curr. Opin. Genet. Dev.* 3:102-9.
- Burns, J.C., Friedmann, T., Driever, W., Burrascano, M., and Yee, J.K. 1993. Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: Concentration to a very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 90:8033-8037.
- Cone, R.D. and Mulligan, R.C. 1984. High-efficiency gene transfer into mammalian cells: Generation of helper-free recombinant retrovirus with broad mammalian host range. *Proc. Natl. Acad. Sci. U.S.A.* 81:6349-6353.
- Cosset, F.L., Legras, C., Chebloune, Y., Savatier, P., Thoraval, P., Thomas, J.L., Samarut, J., Nigon, V.M., and Verdier, G. 1990. A new avian leukosis virus-based packaging cell line that uses two separate transcomplementing helper genomes. *J. Virol.* 64:1070-8.
- Cosset, F.L., Takeuchi, Y., Battini, J.L., Weiss, R.A., and Collins, M.K.. 1995. High-titer packaging cells producing recombinant retroviruses resistant to human serum. *J. Virol.* 69:7430-7436.
- Danos, O. and Mulligan, R.C. 1988. Safe and efficient generation of recombinant retroviruses with amphotropic and ecotropic host ranges. *Proc. Natl. Acad. Sci. U.S.A.* 85:6460-6464.
- Dranoff, G., Jaffee, E., Lazenby, A., Golumbek, P., Levitsky, H., Brose, K., Jackson, V., Hamada, H., Pardoll, D., and Mulligan, R.C. 1993. Vaccination with irradiated tumor cells engineered to secrete murine granulocyte-macrophage colony-stimulating factor stimulates potent, specific, and long-lasting anti-tumor immunity. *Proc. Natl. Acad. Sci. U.S.A.* 90:3539-3543.
- Finer, M.H., Dull, T.J., Qin, L., Farson, D., and Roberts, M.R. 1994. kat: A high-efficiency retroviral transduction system for primary human T lymphocytes. *Blood* 83:43-50.
- Ghattas, I.R., Sanes, J.R., and Majors, J.E. 1991. The encephalomyocarditis virus internal ribosome entry site allows efficient coexpression of two genes from a recombinant provirus in cultured cells and in embryos. *Mol. Cell. Biol.* 11:5848-59.
- Gossen, M. and Bujard, H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U.S.A.* 89:5547-5551.
- Han, X., Kasahara, N., and Kan, Y.W. 1995. Ligand-directed retroviral targeting of human breast cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* 92:9747-9751.
- Hawley, R.G., Lieu, F.H., Fong, A.Z., and Hawley, T.S. 1994. Versatile retroviral vectors for potential use in gene therapy. *Gene Ther.* 1:136-138.
- Hofmann, A., Nolan, G.P., and Blau, H.M. 1996. Rapid retroviral delivery of tetracycline-inducible genes in a single autoregulatory cassette. *Proc. Natl. Acad. Sci. U.S.A.* 93:5185-5190.
- Hughes, S.H., Greenhouse, J.J., Petropoulos, C.J., and Suttrave, P. 1987. Adaptor plasmids simplify the insertion of foreign DNA into helper-independent retroviral vectors. *J. Virol.* 61:3004-12.
- Kasahara, N., Dozy, A.M., and Kan, Y.W. 1994. Tissue-specific targeting of retroviral vectors through ligand-receptor interactions. *Science* 266:1373-1376.

- Kavanaugh, M.P., Miller, D.G., Zhang, W., Law, W., Kozak, S.L., Kabat, D., and Miller, A.D. 1994. Cell-surface receptors for gibbon ape leukemia virus and amphotropic murine retrovirus are inducible sodium-dependent phosphate symporters. *Proc. Natl. Acad. Sci. U.S.A.* 91:7071-7075.
- Kim, J.W., Closs, E.I., Albritton, L.M., and Cunningham, J.M. 1991. Transport of cationic amino acids by the mouse ecotropic retrovirus receptor. *Nature* 352:725-8.
- Landau, N.R. and Littman, D.R. 1992. Packaging system for rapid generation of murine leukemia virus vectors with variable tropism. *J. Virol.* 66:5110-5113.
- Lodish, H., Baltimore, D., and Darnell, J. 1995. *Molecular Cell Biology*. W.H. Freeman, New York.
- Mann, R., Mulligan, R.C., and Baltimore, D. 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* 33:153-159.
- Marin, M., Noel, D., Valsesia-Wittman, S., Brockly, F., Etienne-Julan, M., Russell, S., Cosset, F.K., and Piechaczyk, M. 1996. Targeted infection of human cells via major histocompatibility complex class I molecules by Moloney murine leukemia virus-derived viruses displaying single-chain antibody fragment-envelope fusion proteins. *J. Virol.* 70:2957-2962.
- Markowitz, D., Goff, S., and Bank, A. 1988a. A safe packaging line for gene transfer: Separating viral genes on two different plasmids. *J. Virol.* 62:1120-1124.
- Markowitz, D., Goff, S., and Bank, A. 1988b. Construction and use of a safe and efficient amphotropic packaging cell line. *Virology* 167:400-406.
- Miller, A.D. 1996. Retroviral vectors. In *Current Protocols in Human Genetics* (N. Dracopoli, J. Haines, B.R. Korf, D.T. Moir, C.C. Morton, C.E. Seidman, J.G. Seidman, and D.R. Smith, eds.) pp. 12.5.1-12.5.19. John Wiley & Sons, New York.
- Miller, A.D. and Buttimore, C. 1986. Redesign of retrovirus packaging cell lines to avoid recombination leading to helper virus production. *Mol. Cell Biol.* 6:2895-2902.
- Miller, A.D. and Chen, F. 1996. Retrovirus packaging cells based on 10A1 murine leukemia virus for production of vectors that use multiple receptors for cell entry. *J. Virol.* 70:5564-5571.
- Miller, D.G. and Miller, A.D. 1994. A family of retroviruses that utilize related phosphate transporters for cell entry. *J. Virol.* 68:8270-8276.
- Miller, D.G., Adam, M.A., and Miller, A.D. 1990. Gene transfer by retrovirus vectors occurs only in cells that are actively replicating at the time of infection. *Mol. Cell Biol.* 10:4239-4242.
- Miller, A.D., Garcia, J.V., von Suhr, N., Lynch, C.M., Wilson, C., and Eiden, M.V. 1991. Construction and properties of retrovirus packaging cells based on gibbon ape leukemia virus. *J. Virol.* 65:2220-4.
- Miller, D.G., Edwards, R.H., and Miller, A.D. 1994. Cloning of the cellular receptor for amphotropic murine retroviruses reveals homology to that for gibbon ape leukemia virus. *Proc. Natl. Acad. Sci. U.S.A.* 91:78-82.
- Morgenstern, J.P. and Land, H. 1990. Advanced mammalian gene transfer: High-titer retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucl. Acids Res.* 18:3587-3596.
- Naldini, L., Blomer, U., Gally, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M., and Trono, D. 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272:263-267.
- Naviaux, R.K., Costanzi, E., Haas, M., and Verma, I.M. 1996. The pCL vector system: Rapid production of helper-free, high-titer, recombinant retroviruses. *J. Virol.* 70:5701-5705.
- Paulus, W., Baur, I., Boyce, F.M., Breakefield, X.O., and Reeves, S.A. 1996. Self-contained, tetracycline-regulated retroviral vector system for gene delivery to mammalian cells. *J. Virol.* 70:62-67.
- Pear, W., Nolan, G., Scott, M., and Baltimore, D. 1993. Production of high-titer helper-free retroviruses by transient transfection. *Proc. Natl. Acad. Sci. U.S.A.* 90:8392-8396.
- Pear, W.S., Scott, M.L., and Nolan, G.P. 1996. Generation of high-titer, helper-free retroviruses by transient transfection. In *Methods in Molecular Biology: Methods in Gene Therapy* (P. Robbins, ed.), in press. Humana Press, Totowa, N.J.
- Riviere, I. and Sadelain, M. 1996. Methods for the construction of retroviral vectors and the generation of high-titer producers. In *Methods in Molecular Medicine: Gene Therapy Protocols* (P. Robbins, ed.) in press. Humana Press, Totowa, N.J.
- Roe, T., Reynolds, T.C., Yu, G., and Brown, P.O. 1993. Integration of murine leukemia virus DNA depends on mitosis. *EMBO J.* 12:2099-108.
- Shockett, P.E. and Schatz, D.G. 1996. Diverse strategies for tetracycline-regulated inducible gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 93:5173-5176.
- Somia, N.V., Zoppe, M., and Verma, I.M. 1995. Generation of targeted retroviral vectors by using single-chain variable fragment: An approach to in vivo gene delivery. *Proc. Natl. Acad. Sci. U.S.A.* 92:7570-7574.
- Soneoka, Y., Cannon, P.M., Ramsdale, E.E., Griffiths, J.C., Romano, G., Kingsman, S.M., and Kingsman, A.J. 1995. A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucl. Acids Res.* 23:628-633.
- Stoker, A.W. and Bissell, M.J. 1988. Development of avian sarcoma and leukemia virus-based vector-packaging cell lines. *J. Virol.* 62:1008-1015.
- Takahara, Y., Hamada, K., and Housman, D.E. 1992. A new retrovirus packaging cell for gene transfer constructed from amplified long terminal repeat-free chimeric proviral genes. *J. Virol.* 66:3725-32.

- Takeuchi, Y., Vile, R.G., Simpson, G., O'Hara, B., Collins, M.K., and Weiss, R.A. 1992. Feline leukemia virus subgroup B uses the same cell surface receptor as gibbon ape leukemia virus. *J. Virol.* 66:1219-22.
- van Zeijl, M., Johann, S.V., Closs, E., Cunningham, J., Eddy, R., Shows, T.B., and O'Hara, B. 1994. A human amphotropic retrovirus receptor is a second member of the gibbon ape leukemia virus receptor family. *Proc. Natl. Acad. Sci. U.S.A.* 91:1168-72.
- Wang, H., Kavanaugh, M.P., North, R.A., and Kabat, D. 1991. Cell-surface receptor for ecotropic murine retroviruses is a basic amino-acid transporter. *Nature* 352:729-31.
- Yang, Y., Vanin, E.F., Whitt, M.A., Fornerod, M., Zwart, R., Schneiderman, R.D., Grosveld, G., and Nienhuis, A.W. 1995. Inducible, high-level production of infectious murine leukemia retroviral vector particles pseudotyped with vesicular stomatitis virus G envelope protein. *Hum. Gene Ther.* 6:1203-1213.
- Yee, J.K., Friedmann, T., and Burns, J.C. 1994. Generation of high-titer pseudotyped retroviral vectors with very broad host range. *Methods Cell Biol.* 43:99-112.
- Yu, S.F., von Ruden, T., Kantoff, P.W., Garber, C., Seiberg, M., Ruther, U., Anderson, W.F., Wagner, E.F., and Gilboa, E. 1986. Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 83:3194-3198.
- Zitvogel, L., Tahara, H., Cai, Q., Storkus, W.J., Muller, G., Wolf, S.F., Gately, M., Robbins, P.D., and Lotze, M.T. 1994. Construction and characterization of retroviral vectors expressing biologically active human interleukin-12. *Hum. Gene Ther.* 5:1493-1506.

Contributed by Constance Cepko
Harvard Medical School
Boston, Massachusetts

Warren Pear
University of Pennsylvania
Philadelphia, Pennsylvania

Preparation of a Specific Retrovirus Producer Cell Line

UNIT 9.10

Establishing a cell line that produces high levels of a specific retrovirus construct is a lengthy process that involves several steps. First, the retroviral construct must be stably introduced into an appropriate packaging cell line. This can be accomplished either directly by transfection or by transfection to produce a transient virus stock followed by cross-infection of this stock into a separate packaging line (see Basic Protocol 1 for descriptions of both of these methods). After stable lines are produced, they must be characterized to identify lines that produce high titers of virus with an appropriate structure. The most common method for determining virus titer is using a drug selection protocol (see Basic Protocol 2); a sample calculation of BAG virus titer is also presented (see Support Protocol 1). Sometimes, a short and direct method of estimating virus titer is available when the virus encodes a histochemically detectable gene such as *lacZ*. In such a case, infected cells can be stained with Xgal (see Support Protocol 2). Similarly, it is possible (regardless of the genes encoded by the vector) to estimate rapidly the levels of virus being produced by a given cell line by quantitating any specific product of the virus (e.g., RNA) present within the producer line (see Alternate Protocol).

Although the protocols described in this unit for generating stable retroviral producer cell lines allow expression at high titer for some constructs, their creation can be laborious and may not be effective for all constructs (e.g., due to a gene that is toxic to producer cells). Alternatively, the generation of high-titer retrovirus by transient production, described in *UNIT 9.11*, has facilitated the production of high-titer retroviral supernatants from cDNAs that have proved problematic with stable producer lines.

INTRODUCTION OF A RETROVIRUS VECTOR INTO A PACKAGING CELL LINE

BASIC
PROTOCOL 1

Producing an infectious virus from a bacterial plasmid involves introduction of the plasmid into a packaging line. The transfection protocol most commonly used, which is quite reliable and easy to perform, is from Parker and Stark (1979) and is a modification of the CaPO_4 transfection protocol of Graham and van der Eb (1973; also see *UNIT 9.1*). This protocol has been successfully used with every packaging line we have tried ($\Psi 2$, ΨCRE , ΨCRIP , PA317, Q2bn, and Isolde; see Table 9.9.1). Stable producers can be selected among the stably transfected cells, or by cross-infection using transiently produced virus from one class of packaging line to infect a packaging line bearing a different class of envelope (Fig. 9.10.1). Protocols for production of replication-competent avian vectors are also given in Morgan and Fekete (1996).

Materials

- Appropriate packaging cell line(s) (*UNIT 9.9*) with appropriate medium
- Retrovirus plasmid DNA
- HEPES-buffered saline (HeBS; see recipe)
- 2 M CaCl_2
- Medium, without and with serum
- HeBS containing 15% (v/v) glycerol (HeBS/glycerol)
- 800 $\mu\text{g}/\text{ml}$ (100 \times) polybrene in distilled H_2O , filter sterilized (store at -20°C)
- 10% to 15% (v/v) dimethylsulfoxide (DMSO)
- 10-cm or 6-cm dishes
- 24-well or 6-well tissue culture dishes
- Cloning cylinders (*UNIT 16.23*)

Introduction
of DNA into
Mammalian Cells

Contributed by Constance Cepko

Current Protocols in Molecular Biology (1996) 9.10.1-9.10.13

Copyright © 2003 by John Wiley & Sons, Inc.

9.10.1

Supplement 63

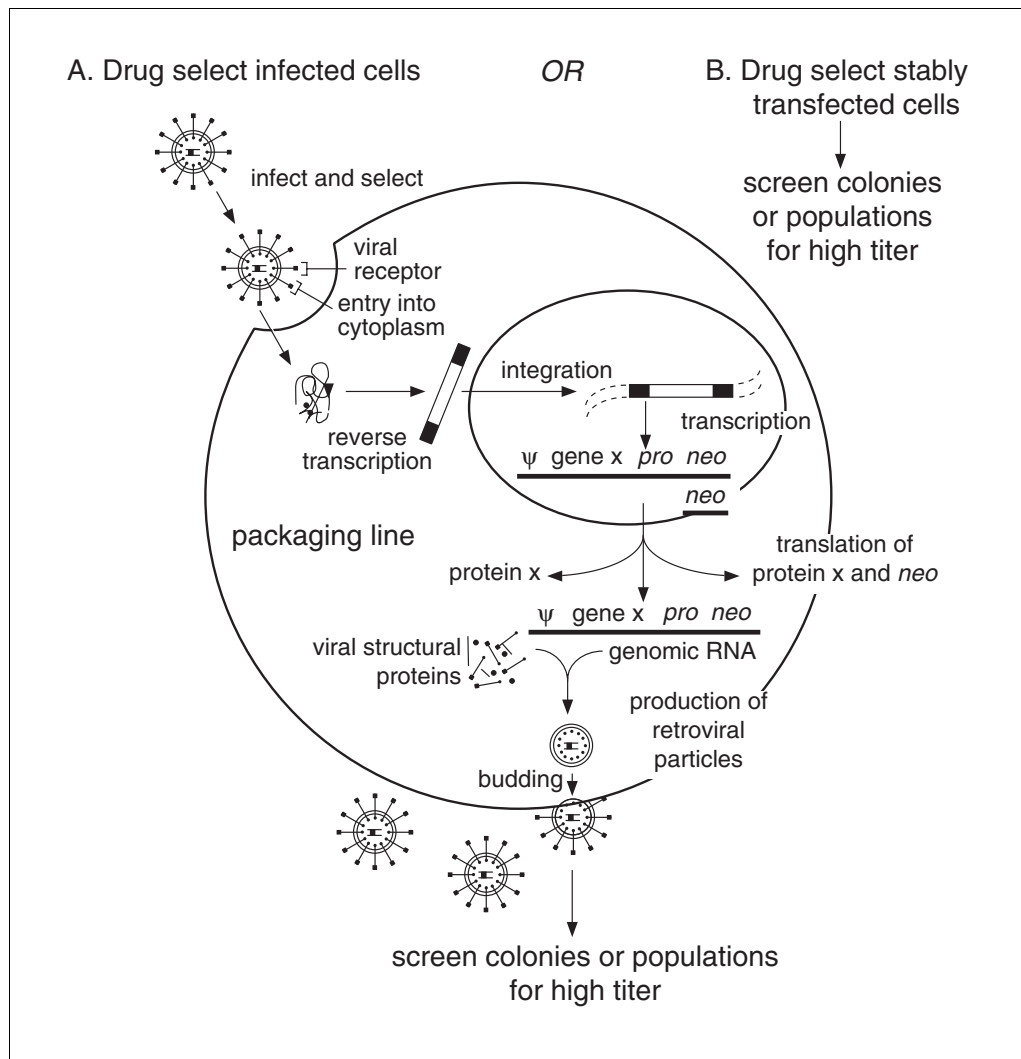


Figure 9.10.1 Production of a stable producer line for a replication-incompetent vector. To have a continual source of high-titer virus stocks, a stable producer line is made by stable transduction of the viral genome into packaging cells. One of two basic methods for transduction is used. The starting point for both methods is to introduce a bacterial plasmid encoding the vector into packaging cells as shown in Figure 9.9.3. The virus produced transiently by the transfected cells is then harvested and used to infect a packaging line bearing a different *env* protein (see text) as shown in **A**. Alternatively, selection of the transfected packaging cells in which integration of the bacterial plasmid has occurred can be carried out as shown in **B**. In either method, drugs (e.g., G418 for the virus shown here) are used to select those cells that have stably integrated the viral sequences. Drug-resistant colonies can then be screened for production of high-titer virus stocks.

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

NOTE: All incubations involving tissue culture cells should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise noted.

Transfecting Cells and Obtaining Viral Supernatant

1. Plate packaging cells to ~10% to 20% the density of confluent cells (referred to hereafter as a 1:10 or a 1:5 split) in a 10-cm dish the day before transfection.

A 10-cm dish is typically used but a 6-cm dish can also be used, in which case all volumes given here and the amount of DNA should be scaled down by a factor of 3.

2. Place 10 μg of retrovirus plasmid DNA into 0.5 ml HeBS.

The DNA need not be sterile. Falcon tubes (5-ml; #2054) provide for good visibility of the precipitate. A selectable drug-resistance gene must be used if the purpose of this procedure is to make a stable packaging line. The drug marker can be encoded by the vector, or by an independent, nonretroviral plasmid that can be cotransfected. If cotransfection is being performed, use a molar ratio of viral to nonviral (or drug marker) plasmid of 10:1 and mix them together in the HeBS (see UNIT 9.5 for discussion of selectable markers).

3. Add 32 μl of 2 M CaCl_2 to DNA while gently shaking the tube. Tap the tube for ~30 sec and incubate 45 min at room temperature.

A fine, hazy blue precipitate should develop. Large, clumpy precipitates do not work well.

4. Remove medium from packaging cells and gently pipet the HeBS-DNA precipitate onto the center of the dish. Expose cells to DNA 20 min in a tissue culture hood; after ~10 min, gently rock the dish to redistribute the solution from the edges of the dish to the entire surface of the dish.

5. Add 10 ml medium and incubate cells 4 hr at 37°C.

6. Aspirate the medium completely, and gently add 2.5 ml room temperature HeBS/glycerol. Return dish to incubator for 3.5 min, 90 sec, or a time that has been determined to be optimal for the cells.

An incubation time of 3.5 min in HeBS/glycerol is suitable for several lines, including $\Psi 2$, ΨCRE , ΨCRIP , and PA317. However, only 90 sec was optimal for some avian lines, including Q2bn. If problems are encountered with glycerol overexposure, such as cells lifting off the dish during the rinse, titrate the amount of time in the HeBS/glycerol solution.

7. Quickly remove HeBS/glycerol and gently rinse with 10 ml medium. Repeat medium rinse and add 5 ml medium with serum (use medium and serum recommended for the particular packaging line). Incubate 18 to 24 hr.

8. Remove the medium and filter through a 0.45- μm filter (the type that easily fits onto a syringe) to obtain supernatant containing the transiently produced virus.

This “transient” harvest can be stored at -70° or -80°C or used immediately for an infection of another packaging line bearing a different env class in order to make an infected producer line (steps 10 to 12 below). If the goal is to produce virus transiently for some other purpose, the experiment will end at step 8 and the original transfected cells can be discarded.

Generating Stable Producer Lines

If the purpose of the transfection is to make stable producer lines, there are two options at this point. The first option (step 9 followed by steps 13 to 18) is to select among the transfected cells for those that have stably integrated the vector plasmid. They can be placed under drug selection and the resulting drug-resistant cells screened for virus production (see Support Protocol 2). The second option is to use “cross-infected” packaging cells. The transiently produced virus harvested in step 8 above can be used to infect another packaging line, as outlined in steps 10 to 12. The *infected* packaging cells can then be placed under drug selection (steps 13 to 19). Obviously, only viruses that encode drug resistance genes can be used for this latter protocol. Regardless of the method used, the goal is to isolate packaging cells that have stably integrated the viral genome and that make the highest possible titer.

Prepare transfected cells for selection

9. After harvesting the transiently produced virus, add 10 ml medium to the transfected cells. Incubate 2 to 3 days, then proceed to step 13.

Prepare infected cells for selection

This route can only be used if the virus carries a selectable marker. This method may be preferable over using transfected cells, as the provirus that results from infection is usually stably associated with the host genome and does not rearrange or delete after integration. Transfection often leads to rearrangements, including deletion, of integrated plasmids. Moreover, for some retroviruses it has been observed that single-copy integrated proviruses can lead to a more uniform level of transcription among individual colonies, whereas transfection leads to widely varied levels of gene products among different colonies. To use this method, two packaging lines expressing different classes of viral glycoproteins must be used. When a packaging line produces a particular *env* glycoprotein (e.g., murine amphotropic *env*), the receptors on the surface of the packaging cells become blocked by the binding of that viral glycoprotein. Thus, if one tries to infect such a line with a virion bearing the same *env* type (e.g., amphotropic *env*), there are no free receptors on the cell surface and the added virion cannot gain entry to the cell and thus a packaging line bearing a different *env* type must be “cross-infected.” When using cross-infection, the two packaging lines must be homologous (i.e., both must be murine or both must be avian).

10. Split a packaging line 1:10 to 1:20 the day before the infection (i.e., the day of the transfection). Use a packaging line with a different envelope class than the cell line used in step 1.
11. Remove the medium from packaging line to be infected. Add virus (from step 8) as follows: for a 10-cm dish, dilute 0.1 to 1 ml virus stock to 3 to 5 ml final volume and add 800 $\mu\text{g/ml}$ polybrene to 8 $\mu\text{g/ml}$; for a 6-cm dish, dilute 0.1 to 1 ml virus stock to 2 ml final and add 800 $\mu\text{g/ml}$ polybrene to 8 $\mu\text{g/ml}$ final. Incubate ≥ 1 hr.

The amount of virus stock to use is difficult to predict if the vector is a novel construct. Use several amounts to infect several dishes to ensure that at least one dish will yield enough producer colonies after the selection process. If working with a published vector, use an amount of virus suggested by prior experience.

Include polybrene for murine viruses and all avian subgroups other than A. Longer incubation times do not hurt the cells if they are insensitive to polybrene (as most packaging cells are).

12. Add medium to 10 ml final volume for a 10-cm dish, and to 4 ml for a 6-cm dish. Incubate 2 to 3 days.

If the cells are insensitive to polybrene, as well as to the medium and serum that are present in the viral inoculum, leave the virus on the cells and simply add medium to the cells to increase the volume as indicated above. However, when this is not the case, discard the viral inoculum and replace it with the appropriate volume of the normal medium and serum used for growth of the cells.

Select stable producers

13. Prepare 1:10 or 1:20 splits of the transfected (step 9) or infected (step 12) cells 2 to 3 days after transfection or infection and plate in selective medium. Incubate 3 days.
14. Replace the medium with fresh selective medium. Incubate an additional 4 to 7 days (7 to 10 days after cells have been placed under selection) until colonies are visible.

15. Pick well-isolated colonies using cloning cylinders and transfer into two wells for each clone in 24-well or 6-well tissue culture dishes. Grow until 50% to 90% confluent.

A population of colonies can be pooled, rather than carried as individual lines. Pooling producers can be satisfactory if high titer is not required. However, when transfection is being used to generate stable producers, the way to isolate the highest-titer virus is to pick colonies and screen them individually. When cross-infection is being used, quite often pooled producers make as high a titer as individual colonies (see Commentary).

Cultivate producer clones and harvest virus stocks

16. Remove the medium and replace it with one-half the normal volume. Incubate 2 to 3 days.

The incubation time depends on the packaging line. The authors determine the time empirically and use 2 to 3 days for Ψ 2 and 1 to 2 days for Isolde or Ψ CRIP. One can also remove the medium every 24 hr for 2 to 3 days and pool the individual harvests.

17. Remove the medium and either titer it immediately (see Basic Protocol 2) or store at -70° or -80°C for future titration (see Support Protocol 2) and use.

18. Continue to passage the producer clones until they can be frozen, and/or until the best producers are identified.

Use the same passaging protocol that is suggested by the originators of the particular packaging line.

19. When a good producer clone is identified, store many vials (20 to 25) of these cells in 10% to 15% DMSO in liquid nitrogen, using aliquots of 1 to 2 ml.

Freezing many vials is done to guard against the typical cell culture hazards, as well as the problem of recombination that can generate helper virus, and to allow for a backup supply of cells in the event that there is loss of titer (some clones reduce their virus output over time, for unknown reasons).

DETERMINATION OF VIRAL TITER: IDENTIFICATION OF PRODUCER CLONES MAKING HIGH-TITER VIRUS

After isolating clones of virus producers or after generating pools of producers, those making high titers need to be identified. In addition, it is necessary to determine the amount of virus in a given stock for use in experiments. The usual method is to perform a bioassay using the virus to infect target cells as described here, although rough quantitative analyses can alternatively be performed on the producer cells to quantitate the presence of vector RNA or protein (see Alternate Protocol and Support Protocol 2).

It should be noted that if the protocol described here for determining viral titer is used, one is simply measuring the amount of virus that produces either a selectable marker or a histochemical marker. Therefore, it is also necessary to ensure that the titered virus has not rearranged or deleted—i.e., that it still expresses the gene of interest as well as the markers that are being scored in the titration assay. The final step of this protocol addresses the issue of subsequent functional and/or structural assays that should be carried out.

Additional Materials (also see Basic Protocol 1)

Target cell line (e.g., NIH 3T3 fibroblasts)

Virus stock (see Basic Protocol 1; e.g., Ψ 2 BAG supernatant)

Xgal (see Support Protocol 2), or G418 or other selection drug (e.g., UNIT 9.5)

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

BASIC PROTOCOL 2

Introduction of DNA into Mammalian Cells

9.10.5

NOTE: All incubations involving tissue culture cells should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise noted.

1. Split NIH 3T3 target cells 1:10 to 1:20 into 6-cm dishes the day before infection.

NIH 3T3 cells are mouse fibroblasts that are frequently used for murine retrovirus infections (see Table 9.9.1). If using avian viruses, substitute chicken embryo fibroblasts split 1:4 or 1:5, or use quail cell line QT6 split 1:6 for subgroups A or E (but not subgroup B).

When titering on CEFs, make sure that the genotype of the CEFs will allow infection of the avian env subgroup that is being titered (see Morgan and Fekete, 1996).

2. On the day of infection, remove medium from the target cells and add medium containing virus stock. Use 1 to 2 ml medium containing from 0.01 µl to 0.1 ml of virus stock, depending upon the expected titer, to infect a 6-cm dish of cells, or 3 to 5 ml for a 10-cm dish; add 800 µg/ml polybrene to 8 µg/ml final for murine or avian subgroups other than A. Incubate cells (in 8 µg/ml polybrene concentration) 1 to 3 hr at 37°C.

It is best to titer at least two dilutions that differ by 10-fold to obtain a countable number of colonies (20 to 100). Some cells can stay overnight in the high (8 µg/ml) polybrene whereas other cells are bothered by this. Because the murine virus only has a half-life of 4 hr, longer incubations do not usually result in much more infection. Virus absorption takes place fairly rapidly, within ~1 hr.

If the cells are insensitive to polybrene, as well as to the medium and serum present in the virus stock, the medium containing virus can remain on the cells and more medium can simply be added to the cells to bring the volume to the level required to dilute the polybrene to 2 µg/ml. However, when this is not the case, discard the medium containing the polybrene and virus and replace it with the appropriate volume of the normal medium and serum that are used for growth of the cells.

3. Add medium to dilute polybrene to 2 µg/ml. Incubate infected cells at least 2 or 3 times the length of a cell cycle.

This period allows for integration and expression of the viral genes and is 2 to 3 days for NIH 3T3 and QT6 cells. For primary CEF cells, it may be shorter (e.g., 1.5 to 2 days).

- 4a. *If the virus carries a histochemical marker gene such as lacZ:* Stain the infected cells with Xgal (see Support Protocol 2) or using other marker-appropriate method; no drug selection is needed. Proceed to step 7 to calculate titer by counting the number of blue colonies.

- 4b. *If the virus carries a drug resistance gene:* Split infected cells into selection conditions. If the resistance gene is *neo*, split the cells 1:10 to 1:20 and set up two 10-cm dishes containing G418 for each. Incubate 3 days.

The amount of G418 to use depends upon the cells being selected. For murine fibroblasts (e.g., NIH 3T3) use 1 mg/ml G418 (i.e., 1 mg/ml as weighed out from the bottle, not corrected for percent of active compound).

5. Change the medium, including appropriate selection drug(s). Incubate a total of 7 to 10 days (for NIH 3T3 under G418 selection) after which time colonies should be obvious.
6. Count colonies before they spread (usually before 12 days under selection).
7. Calculate the titer as follows:

$$\text{G418-R CFU/ml} = \frac{\text{no. of colonies}}{\text{virus volume (ml)} \times \text{replication factor} \times \text{fraction of infected cells plated}}$$

or, if the virus carries *lacZ*,

$$\text{Xgal CFU/ml} = \frac{\text{no. of Xgal}^+ \text{ colonies}}{\text{virus volume (ml)}}$$

As with most bioassays where there are several steps, perfect reproducibility in this assay should not be expected. A rule of thumb is that titers varying >3-fold are significantly different; titers within a factor of 3 are for most intents and purposes about the same. If greater accuracy is necessary, perform the assay several times and average the results. When it is important to compare two stocks, titer them side by side in the same experiment.

Calculation of titer is done differently in different labs, partly because it is difficult to account for the number of cell divisions that occurred after viral integration and before plating the cells into selective medium. For example, if there are two to three cell divisions after integration and before G418 selection, a single viral particle could result in two to four G418-resistant colonies. Recently, with the advent of screenable markers, such as *lacZ*, it is possible to titer the CFU using a direct measure of the number of Xgal⁺ colonies present 2 days after inoculation of NIH 3T3 cells. A direct method such as this does not require a subcultivation after infection and before assessment of the number of colonies. There is thus no ambiguity due to the number of cell divisions, as each colony is scored as 1, regardless of the number of cells in the colony.

The BAG virus (Fig. 9.10.2) encodes both *neo* and *lacZ* and has allowed direct comparison of two methods of scoring positive cells—i.e., resistance to G418 and Xgal staining. This comparison has been made with several BAG stocks, infecting with the same stock at the same time in parallel dishes. In this case, the number of G418-resistance (R) CFU/ml can be corrected for the number of cell divisions by determining the average

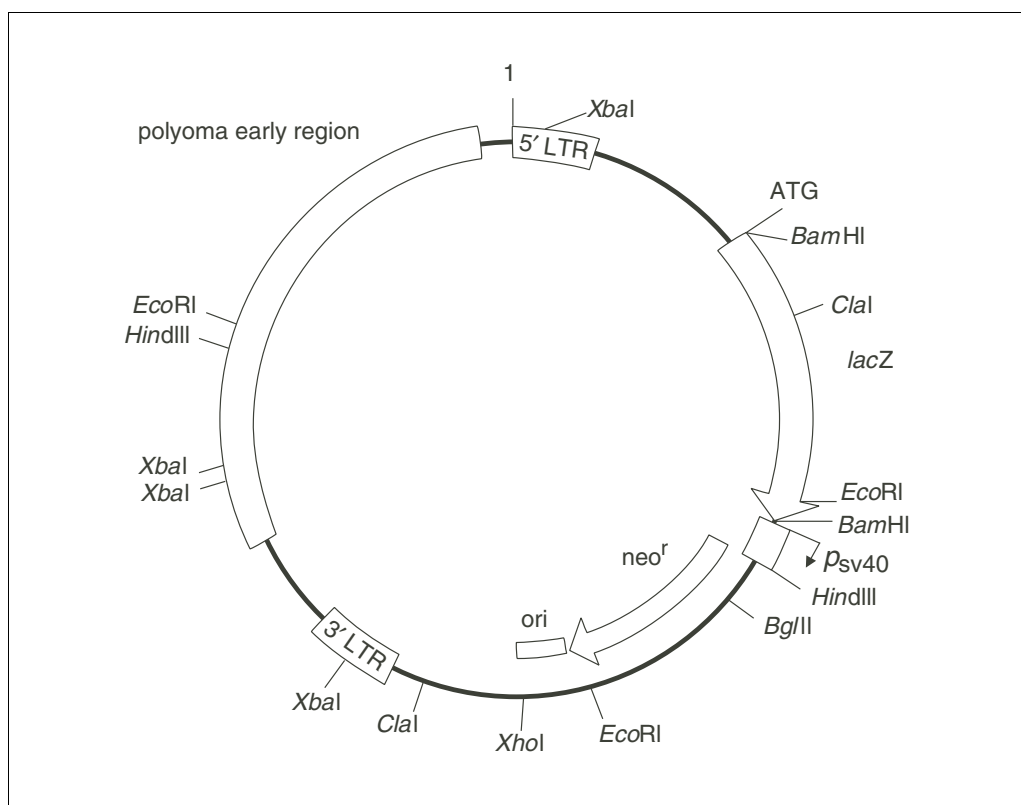


Figure 9.10.2 The BAG retrovirus (Price et al., 1987) was created from the Moloney murine leukemia virus–derived vector, DO-L of Korman et al. (1987). The *lacZ* gene of *E. coli* was fused to the start codon for Moloney virus pr65 *gag*. The *neo* gene of Tn5 is used to provide for selection in G418. Transcription occurs from two promoters: the viral 5' LTR promoter for *lacZ* mRNA, and the SV40 early promoter for *neo* mRNA. Polyadenylation for both mRNAs is provided by sequences in the 3' LTR.

number of cells per Xgal⁺ colony and using that factor as the replication factor (shown above) in the calculation of G418-R titer. When such corrections have been made, the Xgal CFU/ml and the G418-R CFU/ml are in agreement usually within a factor of 2 (see Support Protocol 1). When the vector under titration does not have a histochemical marker gene, then a guess should be made for the replication factor. Usually, 2 to 4 are reasonable estimates of the replication factor if two to three cell cycle times were allowed after infection and prior to selection.

8. Run an assay to determine that the producer clone is making a virus that is not rearranged or deleted—i.e., a virus that has the intended structure.

Because the titration is usually performed using a drug marker such as neo, or a histochemical marker such as lacZ, the titration assay does not prove that the virus is transducing and expressing any gene other than neo or lacZ. Presumably, for many vectors, there is an additional gene of interest that one hopes to express. Thus, some assay capable of determining whether the protein of interest is being expressed must be carried out. If such an assay is not straightforward (e.g., the function of the gene of interest has not been ascertained), at least the structure of the viral genome, and the presence of viral RNAs, should be investigated. A Southern blot analysis (UNITS 2.9 & 2.10) can be performed on cells infected with virus from the producer clone. Alternatively, or additionally, the proviral genome can be cloned directly from infected cells into bacteria using one of several rescue schemes (Cepko et al., 1984). The cloned provirus can then be examined in great detail. In addition, it is important to analyze RNA structure using northern blot (UNIT 4.9) or RNase protection assays (UNIT 4.7) when the mRNA for the gene of interest is a subgenomic RNA.

SUPPORT PROTOCOL 1

SAMPLE CALCULATION OF TITER FOR BAG VIRUS

Experimental setup

Each of two 10-cm dishes of NIH 3T3 cells were infected with 1 µl of BAG virus (which encodes both *neo* and *lacZ*) diluted into 3 ml medium for the initial step of the infection protocol. Following a 2-day incubation, the two parallel dishes were treated two different ways to assess *neo* and *lacZ* expression:

1. Expression of *neo* was assessed by splitting one dish of infected cells 1:20 onto two 10-cm G418 selection dishes (thus, fraction of infected cells plated = 0.05) and incubating 10 days (see Support Protocol 2, steps 4b, 5, and 6).
2. Expression of *lacZ* was assessed by staining the other dish of infected cells with Xgal (see Support Protocol 2, step 4a). The number of Xgal⁺ colonies was determined that same day.

Experimental results

Ten days post-initiation of G418 selection, an average of twenty G418-resistant (G418-R) colonies/dish were observed in the G418 selection dishes. On the dish stained with Xgal, there were 200 Xgal⁺ colonies. [The average number of cells per colony after Xgal staining was four (i.e., replication factor = 4).]

Calculation of titer (see Basic Protocol 2, step 7)

$$\text{G418-R CFU/ml} = \frac{20}{0.001 \times 4 \times 0.05} = 10^5$$

$$\text{Xgal CFU/ml} = \frac{200}{0.001} = 2 \times 10^5$$

RAPID EVALUATION OF PRODUCER COLONIES

Rather than titrate the virus stock for vector activity, one can quantitatively analyze the producer cells for the presence of vector RNA or protein. There is often (though not always) a correlation between amount of viral product within the packaging cell and the titer of virus vector in the supernatant. For example, the amount of β -galactosidase protein present in a producer line that carries a *lacZ* vector is usually directly correlated with the amount of *lacZ* virus in the supernatant. Thus, each clone can be screened by ONPG hydrolysis (Table 1.4.2) or by Xgal staining (see Support Protocol 2). Alternatively, RNA dot blots (UNIT 4.9) can be used to quickly assess the amount of vector RNA and this has also been correlated with titer. Because Xgal staining, as well as RNA analysis methods, are invasive and lethal, duplicate wells containing the colonies picked from the drug-selection dishes (see Basic Protocol 2, step 5) are made so that a dish of live cells is available after the analysis. Direct titration of the culture supernatants (see Basic Protocol 2) is the more typical method to screen producer clones, but requires a longer period of time and thus more subcultivations of the producers while awaiting the results.

XGAL STAINING OF CULTURED CELLS

This protocol can be used to stain infected cells to obtain a direct measure of titer. It can also be used to stain producer lines themselves (see Alternate Protocol) to discover which producer clones make high levels of β -galactosidase as an indication of those that may make high titers of *lacZ*-bearing viruses. Cultured cells are stained with Xgal using either of two fixatives. The choice of fixatives is based primarily on convenience as both work quite well, although the glutaraldehyde fixative gives slightly bluer cells.

Materials

- Cells infected or transfected with a *lacZ*-encoding virus (see Basic Protocol 1)
- PBS (APPENDIX 2)
- Fixative solution: 0.05% glutaraldehyde *or* 2% paraformaldehyde solution (see recipes)
- Xgal solution (see recipe)

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

NOTE: All incubations involving tissue culture cells should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise noted.

1. Discard medium and add fixative solution (2 ml for 6-cm dish and 5 ml for 10-cm dish). Incubate at room temperature with 2% paraformaldehyde for 60 min, or with 0.05% glutaraldehyde for 5 to 15 min.

It is best to perform all steps involving fixatives in a fume hood. Longer incubation times may inhibit enzyme activity in subsequent steps.

2. Discard fixative solution using proper chemical disposal protocol for your institution and rinse cells thoroughly three times in PBS at room temperature. Leave the second rinse on for ~10 min and perform first and third washes quickly.

Residual fix will inhibit the enzyme reaction.

3. Add Xgal solution.

Use a minimal volume to cover the cells because Xgal is very expensive.

4. Incubate 1 hr to overnight at 37°C. Positive cells will stain blue.

The dish can be stored indefinitely at 4°C.

**ALTERNATE
PROTOCOL**

**SUPPORT
PROTOCOL 2**

**Introduction
of DNA into
Mammalian Cells**

9.10.9

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Glutaraldehyde fixative (0.05%)

Glutaraldehyde (obtained from Sigma) is typically supplied as a 25% solution and can be stored frozen at -20°C in aliquots. Aliquots of the 25% solution can be frozen and thawed many times. Immediately before use, thaw an aliquot and dilute 500-fold in PBS to 0.05%.

CAUTION: *Glutaraldehyde is a carcinogen and may cause an allergic reaction. Use in a fume hood and discard waste as directed by your institution.*

HEPES-buffered saline (HeBS)

137 mM NaCl

5 mM KCl

0.7 mM Na_2HPO_4

6 mM dextrose

21 mM HEPES

Adjust pH to 7.05; check carefully, as this is a critical parameter

Store indefinitely at room temperature

Paraformaldehyde fixative (2%)

Dissolve 2 g paraformaldehyde (e.g., from BDH) in 100 ml of 0.1 M PIPES, pH 6.9 [30.24 g piperazine-*N,N'*-bis(2-ethanesulfonic acid) per liter] containing 2 mM MgCl_2 (200 μl of a 1 M solution) and 1.25 mM EGTA (250 μl of a 0.5 M solution, pH 8.0). Heat in the fume hood with stirring 5 to 10 min (do not boil). Cool to 4°C . Always use a paraformaldehyde solution that is well buffered and <1 week old.

CAUTION: *Formaldehyde is a carcinogen and may cause an allergic reaction. Use in a fume hood and discard waste as directed by your institution.*

Xgal solution

Prepare in PBS:

5 to 35 mM $\text{K}_3\text{Fe}(\text{CN})_6$ (potassium ferricyanide)

5 to 35 mM $\text{K}_4\text{Fe}(\text{CN})_6 \cdot 3\text{H}_2\text{O}$ (potassium ferrocyanide)

1 to 2 mM MgCl_2 or MgSO_4

Just before use, add 40 \times Xgal in *N,N*-dimethyl formamide (Table 1.4.2) to 1 mg/ml final

CAUTION: *Avoid contact and inhalation of cyanide. Discard waste as directed by your institution.*

The amount of ferric and ferrous cyanide to use is a matter of preference. The higher amount causes precipitation of indole to occur more quickly and thus reduces diffusion. However, it may cause a greenish background upon prolonged incubation (overnight or longer) in some tissues, although not usually on cell lines.

The first three ingredients (i.e., excluding Xgal) can be stored at least a few months at room temperature. 40 \times Xgal can be stored in a glass container covered with foil for months at -20°C .

COMMENTARY

Background Information and Critical Parameters

Introduction of a vector into a packaging line can be accomplished by transfecting a vector plasmid or by using virus produced by one line to infect a different packaging line (“cross-infection”; Fig. 9.10.1). Transfection is a nonspecific entry technique that allows one to introduce any DNA into some types of cells (UNITS 9.1-9.5). Transfected DNA can be transcribed transiently in ~10% to 50% of the packaging cells, within a period of a few hours to a few days after transfection. However, only ~1 in 5000 cells will stably integrate the DNA. Although the transfected DNA bears the retroviral sequences necessary for the precise integration that is the hallmark of the retroviral lifecycle, transfected retroviral DNA does not undergo precise or efficient integration. Transfected retroviral DNA integrates as nonspecifically and with as poor an efficiency as any other plasmid DNA. Key features of the viral structure and enzymatic activities are present only following infection and are absent from transfected plasmid DNA.

Transiently or stably produced vector RNA (which contains Ψ) is efficiently packaged and budded from the surface of packaging cells (Fig. 9.9.3). Thus, the infectious retrovirus vector can be harvested by removal of the supernatant of the packaging line. A given line that has stably integrated the virus may produce high amounts of virus indefinitely. Cells that have stably integrated the viral genome are usually selected by application of a drug that selects for a gene encoded by the virus. If no such dominant, selectable gene (UNIT 9.5) is encoded by the virus, the transfected cells can be cotransfected with a nonviral plasmid that does encode such a gene. By using a 10:1 molar ratio of viral to nonviral plasmid, cells can be selected that have integrated both plasmids. If instead a producer is generated by cross-infection, a very small amount of virus should be used for the infection. If a high multiplicity of infection (MOI) is used, there is a risk of transferring defective genomes or wild-type helper recombinants (see UNIT 9.13). In some instances, with some vectors, it appears that cross-infection results in producer clones that have a high probability of producing the maximal titer for a given vector. Producer clones isolated following transfection produce quite variable titers, presumably because transfection results in varied configurations of the transfected DNA. Infec-

tion, on the other hand, results in a single-copy provirus quite reliably.

Infection of a packaging line by virus produced by the same packaging line is very low in efficiency (although it is not blocked entirely). The viral glycoprotein produced by a packaging line apparently binds the host-cell receptor for that *env*-encoded glycoprotein class. A virion bearing that same *env* class thus cannot easily enter. However, because the different classes of viral glycoproteins use different cellular receptors, virions produced by transfections of one class of packaging line can be used to “cross-infect” another (e.g., Ψ CRE-produced virus can infect Ψ CRIP quite effectively). There is also an alternative method that allows infection of a packaging line with virions produced by that packaging line class (Rein et al., 1982). An inhibitor of glycosylation, such as tunicamycin, can be used to transiently block the production of the *env* glycoprotein. During the application of this drug block, the cells can be infected with virions bearing the same class of *env* glycoprotein. The drug is then washed out. This method requires careful titration of the drug parameters (e.g., time and dose) as these drugs can be toxic to the cells. As it is quite easy to introduce viral genomes using standard transfection methods, these methods have been outlined in this unit.

Troubleshooting

The titer (concentration, expressed as colony forming units per ml, CFU/ml) of virus varies widely and depends on a combination of factors. The vector, the insert(s), the packaging line, and the method of harvesting all contribute to the final titer. Although titers from stable producers can be as high as 10^8 CFU/ml, 10^6 CFU/ml is considered to be a good titer for a murine ecotropic stock, and 10^5 CFU/ml for an avian replication-incompetent stock. Titers of transiently produced stocks (i.e., those stocks harvested from packaging cells 18 to 24 hr posttransfection) are generally lower, from 10^4 to 10^5 CFU/ml.

If no high-titer producers are identified among at least 20 individual producer clones, it may be due to poor transfection, poor production of virions by the packaging line, poor infectibility of the cells upon which the virus is titered, or problems with transmission or expression of the virus. All but the latter can be controlled for by transfecting and titering a control vector that is known to produce a high

titer. If the control vector works well, but the new vector does not, the problem is in the design of the new vector and can be difficult to resolve. Try using different vectors to express the gene of interest. Variables relating to the vector that can be important include the viral sequences present within the vector (e.g., the more *gag*, the better; Bender et al., 1987), the promoter used for expression of the gene of interest and/or the marker gene whose activity is being followed in the titration assay, and the position of the gene of interest within the vector. Within the sequence of the gene of interest, the presence of splice or polyadenylation sites can cause problems in virus transcription and transmission. The inserted gene can be trimmed to encode only the protein-coding region to alleviate some problems. Trying different vector designs can sometimes improve the situation, but there are some genes which, for obscure reasons, do not express and/or transmit at a high level within any existing retrovirus vector.

When attempting to isolate stable producer lines for a particular construct, difficulties can also be encountered when large numbers of colonies are being handled at once, prior to identification of the few colonies that are producing a high titer. It is definitely an advantage if either RNA or some other vector gene product can be scored using a short-term assay (see Alternate Protocol) in order to minimize the number of subcultivations prior to discarding the poor producers. The other problems that can be encountered in making stable producers are loss of titer from a clone that previously had demonstrated high titer, and helper-virus contamination (the latter is covered in UNIT 9.13).

For loss of titer, two explanations seem most likely. The first is loss of capsid production by the packaging cells. There sometimes is a loss of virion production with some producer lines (Isolde, PA317, and possibly others). In these cases, it is suggested that the packaging lines and producer clones be carried under conditions of drug selection for maintenance of the *gag*, *pol*, and *env* genes (Cosset et al., 1990; Miller and Buttimore, 1986). The other problem that leads to loss of titer is loss of vector RNA production, presumably due to some modulation of the LTR promoter activity (C. Cepko, unpub. observ.). Clones that were originally 100% blue when stained with Xgal became a mixture of blue and white, and even 100% white, upon subcultivation over a short period of time (weeks). When *lacZ* RNA levels were checked on a northern blot, it was found

that white clones, or subclones of previously mixed clones, had low levels of *lacZ* RNA. This phenomenon has been seen with many different promoters and with nonretroviral constructs. Unfortunately, there is no known method of making the promoter turn back on. Even if one can select for a linked drug-resistance gene, this does not always select for the LTR activity. Of course, producer clones with a large percentage of white cells do not make high viral titer. The same phenomenon probably occurs with other vectors that do not encode *lacZ*, as other producers prepared in the author's laboratory have lost titer as well. When such clones were checked for production of virions using a reverse transcriptase assay (UNIT 9.13), there did not appear to be any loss of virion production, and thus loss of vector RNA production is the likely problem.

Anticipated Results

Using a good murine vector, the titer of transiently produced stocks from $\Psi 2$ or Ψ CRIP should be 10^3 to 10^4 CFU/ml. A lower titer—e.g., 10 to 100 CFU/ml—usually indicates that the stable titers will also be low. For transient titers from Q2bn, a good titer is 10^4 CFU/ml; in the author's experience Isolde does not yield good transient titers—only ~10 to 100 CFU/ml.

When selecting stable colonies, most packaging lines transfect well. Expect ~30 to 100 colonies per 10-cm dish under selection with a 1:20 split into drug selection. Good stable titers are $\sim 10^6$ CFU/ml for $\Psi 2$ stocks. For all other packaging lines (with the exception of Bosc, as discussed in UNIT 9.11), the author has not achieved stable titers as high as those made by $\Psi 2$. However, other groups report that for some vectors, they have made producers with titers as high as those made by $\Psi 2$. If stable transfection is used to make the producers, expect ~25% of the producer clones to have the optimal titer for that vector. If cross-infection is used, 25% to 100% of the clones might have the optimal titer for that vector. If pooled producers from cross-infected cells are made, the optimal titer from the pool may be achieved. For Isolde, stable titers can be 10^5 CFU/ml; Q2bn seems to give ~10-fold-lower stable titers. Replication-competent avian vectors have stable titers of $>10^6$ CFU/ml (Hughes et al., 1987; and see Morgan and Fekete, 1996).

Time Considerations

During the first 2 weeks following transfection, there is minimal work. However, during

the period of isolating colonies, titering supernatants, and freezing producer clones, there can be nearly constant tissue culture work. The amount of work will depend on how many colonies are isolated.

Literature Cited

- Bender, M.A., Palmer, T.D., Gelinis, R.E., and Miller, A.D. 1987. Evidence that the packaging signal of Moloney murine leukemia virus extends into the *gag* region. *J. Virol.* 61:1639-1646.
- Cepko, C.L., Roberts, B.E., and Mulligan, R.E. 1984. Construction and applications of a highly transmissible murine retrovirus shuttle vector. *Cell* 37:1053-1062.
- Cosset, F.-L., Legras, C., Chebloune, Y., Savatier, P., Thoraval, P., Thomas, J.L., Samarut, J., Nigon, V.M., and Verdier, G. 1990. A new avian leukosis virus (ALV)-based packaging cell line using two separate transcomplementing helper genomes. *J. Virol.* 64:1070-1078.
- Graham, F.L. and van der Eb, A.J. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* 52:446-467.
- Hughes, S.H., Greenhouse, J.J., Petropoulos, C.J., and Suttrave, P. 1987. Adaptor plasmids simplify the insertion of foreign DNA into helper-independent retroviral vectors. *J. Virol.* 61:3004-3012.
- Korman, A.J., Frantz, J.D., Strominger, J.L., and Mulligan, R.C. 1987. Expression of human class II major histocompatibility complex antigens using retrovirus vectors. *Proc. Natl. Acad. Sci. U.S.A.* 84:2150-2154.
- Miller, A.D. and Buttimore, C. 1986. Redesign of retrovirus packaging cell lines to avoid recombination leading to helper virus production. *Mol. Cell. Biol.* 6:2895-2902.
- Morgan, B.A. and Fekete, D.M. 1996. Manipulating gene expression with replication competent retroviruses. *Methods Cell Biol.* 51:185-218.
- Parker, B.A. and Stark, G.R. 1979. Regulation of simian virus 40 transcription: Sensitive analysis of the RNA species present early in infections by virus or viral DNA. *J. Virol.* 31:360-369.
- Price, J., Turner, D., and Cepko, C. 1987. Lineage analysis in the vertebrate nervous system by retrovirus-mediated gene transfer. *Proc. Natl. Acad. Sci. U.S.A.* 84:156-160.
- Rein, A., Schultz, A.M., Bader, J.P., and Bassin, R.H. 1982. Inhibitors of glycosylation reverse retroviral interference. *Virology* 119:185-192.

Contributed by Constance Cepko
Harvard Medical School
Boston, Massachusetts

Transient Transfection Methods for Preparation of High-Titer Retroviral Supernatants

Although stable retroviral producer cell lines are useful for expressing some genes at high titer and the creation of such lines is the method of choice for producing large quantities of retroviral supernatant, this is a lengthy and laborious process, and for many genes it has failed to yield high retroviral titers. As a result of these deficiencies, transient production of retroviral supernatants has become the preferred method of producing retrovirus for most laboratory uses. In comparison to the production of stable retroviral producer cell lines, generation of high-titer retrovirus by transient production not only is less laborious, but also has allowed the production of high-titer retroviral supernatants from cDNAs that cannot be produced by stable producer cell lines. Transient transfection has also increased the versatility of retrovirus-mediated gene transfer to include the rapid testing of different constructs, viral pseudotyping, and construction of retroviral cDNA libraries.

The most significant change allowing the use of transient transfection is the choice of packaging cell lines. NIH 3T3 cells, which are the choice for stable production (UNIT 9.10), produce relatively low titers (10^3 to 10^4 infectious U/ml) following transient transfection. Development of packaging cell lines derived from either primate (COS) or human (293) cells made it possible to obtain transient titers from one (COS) to several logs (293) higher than NIH 3T3 cells. The increase in titer with these latter cell lines is thought to be due to several properties of these cell lines—high transfectability, few endogenous proviruses, high-level expression from a number of promoters, and expression of additional proteins (such as adenoviral E1A or polyoma large T antigen) that may enhance gene expression and retroviral titer.

Although transient retroviral production has been described in COS cells (Landau and Littman, 1992), systems based on human 293 cells, an adenovirus-transformed human embryonic kidney cell line (Graham et al., 1977), have produced the highest retroviral titers and are the most widely used. Two approaches have been taken for producing high-titer retroviral supernatants in 293 cells (Pear et al., 1993; Finer et al., 1994; Soneoka et al., 1995; Naviaux et al., 1996). One approach (Pear et al., 1993) has created a stable retroviral packaging cell line in 293 cells by sequentially introducing constructs encoding the retroviral structural genes. Using this strategy, a cell line with an ecotropic host range, Bosc23, and a cell line with an amphotropic host range, Bing, have been developed. To create retroviral supernatants, only the retroviral vector needs to be introduced into the cell lines. In the approach of Finer et al. (1994), Soneoka et al. (1995), and Naviaux et al. (1996), plasmids encoding the retroviral vector and retroviral structural genes are coin-troduced into 293 or 293T cells. Both approaches have consistently yielded retroviral titers of $>10^6$ for a wide variety of genes, and neither produces replication-competent retrovirus.

Although retroviruses have been most widely used for infecting cells through the ecotropic or amphotropic receptors, methods have recently been described for expanding the host range of retroviral vectors. To date, this has been best exploited for vesicular stomatitis virus (VSV) G protein, which confers a polytropic host range capable of infecting mammalian, fish, and insect cells (Burns et al., 1993; reviewed in Yee et al., 1994). VSV G protein is thought to mediate cell entry by interacting with phosphatidyl serine and possibly other components of lipid bilayers present in the cell membranes of most mammalian and nonmammalian eukaryotic cells (Schlegel et al., 1983). In addition to their wide host range, retroviruses pseudotyped with VSV G protein can be concen-

Contributed by Warren Pear

Current Protocols in Molecular Biology (1996) 9.11.1-9.11.18

Copyright © 2004 by John Wiley & Sons, Inc.

trated to achieve titers $>10^9$ infectious U/ml. An important breakthrough for VSV G pseudotyping was the use of transient transfection in 293 cells for expressing VSV G protein. This is because VSV G protein is fusogenic and leads to syncytium formation and cell death. By expressing VSV G protein by transient transfection in 293 cells, it is possible to successfully pseudotype retroviral virions during a window between the onset of protein expression and massive cell death.

This unit describes methods for optimizing retroviral production from the 293-based systems (see Basic Protocol 1) and for growing and freezing 293 cells (see Support Protocol). Methods are included for pseudotyping the virus with VSV G protein by sequential transfection (see Basic Protocol 2) or cotransfection (see Alternate Protocol 1). Virus produced by transiently transfected cells can be used to infect cells. Adherent cells can be infected directly with retroviral supernatant (see Basic Protocol 3) or by spin infection (see Alternate Protocol 2). Nonadherent cells can be infected by addition of retrovirus supernatant (see Basic Protocol 4), cocultivation with producer cells (see Alternate Protocol 3), or spin infection (see Alternate Protocol 4). These infection methods are also applicable to retrovirus produced by any of the stable producer cell lines (*UNIT 9.10*).

CAUTION: The viral supernatants produced by these methods may contain potentially hazardous recombinant constructs. The user must exercise due caution in the production, use, and storage of recombinant retroviral virions, especially those with amphotropic and polytropic host ranges, for all genes expressed by the retroviral vectors. Appropriate NIH and regional guidelines should be followed in the use of these recombinant retrovirus production systems.

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

NOTE: All culture incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

BASIC PROTOCOL 1

TRANSIENT TRANSFECTION OF A RETROVIRUS VECTOR INTO 293 CELLS

The first step in producing infectious virus from a bacterial plasmid involves introduction of the plasmid into the retroviral packaging cells. Transfection is the method of choice for this step because it causes many copies of the retroviral vector plasmid DNA to be taken up by each cell. This results in a high level of expression with consequent high retroviral titer. Although other equally efficacious methods exist (see Critical Parameters), the least expensive and simplest method is a modification of the calcium phosphate-mediated transfection protocol of Graham and van der Eb (1973; see *UNIT 9.1* and discussion of Optimization of Transfection).

Although it is possible to scale up, highest retroviral titers are obtained using 60-mm dishes. This protocol can be used to cotransfect retroviral structural genes together with the retroviral vector into 293 cells or to transfect retroviral vector into 293-derived packaging cell lines.

Materials

- 293 cells *or* appropriate packaging cell line (see Table 9.9.1)
- 293 cell growth medium (see recipe)
- 25 mM chloroquine in PBS (1000×; store at -20°C; optional)
- Retroviral plasmid DNA
- 2 M CaCl₂
- 2× HEPES-buffered saline (HeBS), pH 7.05 (see recipe)

Transient Production of Retroviral Supernatants

9.11.2

500 mM sodium butyrate, pH 7.0 (50×; optional)

60-mm tissue culture dishes

0.45- μ m filter Sorvall RT-3000B centrifuge and rotor (or equivalent)

Additional reagents and equipment for culture of mammalian cells (APPENDIX 3F) and calcium phosphate transfection (UNIT 9.1)

NOTE: All solutions should be filtered through a 0.2- μ m filter or autoclaved.

1. Twenty-four hours prior to transfection, plate 2.5×10^6 293 cells in 4 ml of 293 cell growth medium (6.25×10^5 cells/ml) in a 60-mm tissue culture dish (see Critical Parameters).

The culture should be ~80% confluent prior to transfection and cell clumping should be minimized.

2. Just prior to transfection, remove the medium and gently add fresh 293 cell growth medium.

Addition of 25 μ M chloroquine to the medium before transfection may boost the retroviral titer by 2-fold. If chloroquine is used, it is important that it not be left on the cells for >12 hr (see Critical Parameters).

3. Prepare a transfection cocktail by adding 6 to 10 μ g retroviral vector plasmid DNA to water and diluting to 438 μ l final volume. Add 62 μ l of 2 M CaCl_2 to the DNA solution.

Plasmid DNA prepared by cesium chloride gradients or by using a commercial kit appears to work equally well. It is unnecessary to perform additional phenol extraction or precipitation steps prior to using the DNA, and the DNA may be stored in either water or TE buffer.

When cotransfecting retroviral vector with retroviral structural genes on multiple plasmids, the total amount of transfected DNA should be 10 to 15 μ g and the plasmids should be introduced in approximately equimolar ratios. These amounts and ratios may have to be altered to optimize retroviral titers (see Critical Parameters).

4. Add 500 μ l of 2× HeBS and mix by shaking the tube or pipetting. Within 1 to 2 min, add this solution to the cells and gently swirl the plate to ensure uniform mixing. Incubate until transfected cells are close to 100% confluent (24 to 48 hr post transfection). Remove the medium and add fresh medium at ~24 hr post transfection (see step 5).

HeBS may also be added to the DNA/ CaCl_2 solution by bubbling. Transfection efficiency is highest if DNA/ CaCl_2 /HeBS mix is added immediately to cells rather than waiting for a visible precipitate to form (30 to 60 min). Although a precipitate is usually visible on the plate within 1 to 2 hr after transfection, the quantity of precipitate appears to be related to the amount of input DNA (W. Pear; unpub. observ.), and the amount of precipitate is usually an unreliable indicator of transfection efficiency.

If chloroquine is included in the medium, remove medium ~10 hr after transfection and gently replace with fresh 293 cell growth medium.

Addition of 10 mM sodium butyrate to the cells may enhance retroviral titers by several-fold (Soneoka et al., 1995). It should be added to the 293 growth medium 12 to 24 hr after transfection at a final concentration of 10 mM. Cells should be treated for 12 to 14 hr with sodium butyrate and washed with PBS prior to proceeding to step 5.

5. Approximately 16 to 24 hr prior to harvesting the retroviral supernatant, remove the medium and gently replace with 4 ml of 293 cell growth medium.

The transfected cells should be close to 100% confluent at this time. The relative retroviral titer may be increased by adding 3 ml rather than 4 ml of medium. If the cells that will be infected require special medium (e.g., RPMI medium, growth factors), the special medium,

rather than 293 cell growth medium, should be added to the cells. Except for the presence of fetal bovine serum, 293 cells appear to tolerate a wide variety of culture conditions.

6. Harvest retroviral supernatant ~48 hr following transfection by gently removing the supernatant from the cells. Either filter the supernatant through a 0.45- μ m filter or centrifuge it 5 min at $500 \times g$ (1000 rpm in Sorvall RT-3000B rotor), 4°C, to remove live cells.

At harvest, the supernatant will be yellow. Acidity does not appear to affect retroviral titer. Once retroviral supernatant is harvested, the transfected cells may be assayed for transfection efficiency by staining for a histochemical marker (such as lacZ, UNIT 9.10) or cell surface marker, by preparing DNA from the cells and assaying by Southern hybridization (UNIT 2.9A), or by lysing cells for protein immunoblot analysis (UNIT 10.8). Optimal transfection efficiencies should range from 30% to 60%. If the supernatant is to be tested for the presence of replication-competent (helper) virus, see UNIT 9.13.

It is necessary to wait at least 36 hr post transfection to obtain high-titer retroviral supernatants. At this point, it may be possible to serially, harvest and replenish supernatant every 12 hr up to 72 hr post transfection without a significant drop in retroviral titer. If the cells are sparse at 24 hr post infection, it may be necessary to wait until 72 hr to obtain high-titer supernatant.

7. If the retroviral supernatant is to be used within 1 to 2 hr, store it on ice. For longer intervals, freeze the supernatant on dry ice and transfer it to -80°C.

Retroviral half-life is 3 to 6 hr at 37°C (Sanes et al., 1986). To maintain high titers, supernatants should be left on ice or frozen following harvest. Freezing does not appear to cause >2-fold drop in titer, as long as the supernatant does not undergo more than one freeze/thaw cycle; exposure to more than one such cycle causes a marked drop in retroviral titer. For this reason, supernatant should be aliquoted in appropriate volumes for subsequent infection. To thaw frozen supernatants, warm for a minimal period at 37°C and use immediately.

SUPPORT PROTOCOL

GROWTH AND STORAGE OF 293 CELLS

The two most important variables in determining the success of retroviral production are the state of the packaging cells and the quality of the transfection reagents. This protocol discusses methods for optimizing growth of 293 cells, passaging the cells, and freezing stocks. Further details for the culture, freezing, and storage of mammalian cell lines can be found in APPENDIX 3F.

Materials

293 cells or packaging cell lines (e.g., Bosc23; see Table 9.9.1) derived from 293 cells

293 cell growth medium (see recipe)

Freezing medium: 90% (v/v) heat-inactivated FBS/10% (v/v) DMSO

Trypsin solution: 0.05% (v/v) trypsin/0.53 mM EDTA

PBS without Ca²⁺ or Mg²⁺ (APPENDIX 2)

10-cm tissue culture dishes

Sorvall RT-3000B centrifuge and rotor (or equivalent)

2-ml cryogenic vials

Culture and passage 293 cells

1. Passage and grow 293 cells or the packaging cell lines derived from them in 10 ml of 293 cell growth medium in 10-cm tissue culture dishes.

Cells are typically grown in 10-cm dishes; however, other sizes can be used and the protocols scaled appropriately. It is recommended that cells be maintained at relatively high density and passaged at a split ratio $\leq 1:5$ to maintain uniformity of the cells in culture. This is especially true for packaging cell lines. It is also recommended that when cells are first received they be expanded to prepare 50 to 100 vials of frozen cells to provide a uniform source of cells for virus production.

Cells should not be allowed to become overconfluent, as this leads to formation of cell clumps that can cause uneven cell distribution after replating and result in less efficient transfection. Cells should be split when they reach 90% confluence and the medium is acidic. If clumping occurs, growing cells for 1 to 2 passages at high density (1:3 split) should remove the clumps.

293 cells and packaging lines derived from those cells appear to tolerate a wide variety of growth conditions and addition of factors without drop in retroviral titer. One variable that is important for maximal retroviral titers is to use heat-inactivated fetal bovine serum (FBS) in the tissue culture medium, rather than horse serum (HS) or newborn calf serum (NCS). To date, both DMEM and RPMI have given good results, and the addition of IL-2, IL-3, IL-6, and stem cell factor (SCF) has not caused the retroviral titers to decrease (W. Pear, unpub. observ.).

2. When the cells are confluent, remove the medium and gently rinse once with PBS without Ca^{2+} and Mg^{2+} .

293 cells are much less adherent than murine fibroblasts. As a result, washes must be performed very gently to avoid removing cells.

3. Add 1 ml trypsin solution and incubate 2 to 3 min at 37°C.

Although cells may be detached by tapping the plate after 30 sec, a more efficacious method to obtain a single-cell suspensions is to return the cells to the incubator for 2 to 3 min. The cells should then detach by themselves.

4. Quench trypsin with 4 ml of 293 growth medium and transfer the cell suspension to a conical centrifuge tube. Centrifuge 3 min at $500 \times g$ (1000 rpm in Sorval RT-3000B rotor), 4°C or room temperature.

To obtain single-cell suspensions, it is recommended that cells be vigorously pipetted in a conical tube rather than on the tissue culture dish.

5. Resuspend pellet vigorously in 5 ml of 293 cell growth medium. Let the suspension stand 1 min to allow the largest clumps to settle. Aliquot the cells without using the last 200 μl of cell suspension (clumps).

Freeze 293 cells

6. Incubate a 10-cm dish of 293 cells in 293 growth medium until the cells are 80% confluent.

Maximal viability is achieved if cells are frozen prior to confluence. One 10-cm dish that is 80% confluent will provide enough cells for 2 to 3 vials of frozen cells.

7. Using trypsin solution, remove the cells from the dish (see steps 2 to 4). Centrifuge the cells 3 min at $500 \times g$, 4°C or room temperature.
8. Remove the medium and resuspend the cells in 2 to 3 ml freezing medium per 10-cm dish.

9. Add 1 ml cell suspension to each 2-ml cryogenic vial and store overnight at -80°C . On the following day, transfer the vials to liquid nitrogen.

Thaw 293 cells

10. Remove a vial of cells from liquid nitrogen and thaw rapidly, warming by hand or at 37°C .
11. Add 1 ml 293 growth medium to the cryogenic vial and add the cell suspension to a tube containing 10 ml growth medium. Centrifuge 3 min at $500 \times g$, 4°C or room temperature.
12. Remove the supernatant and resuspend the cell pellet in 10 ml growth medium. Transfer the cells to a 10-cm tissue culture dish. Incubate.

BASIC PROTOCOL 2

PSEUDOTYPING A STABLE CELL LINE SEQUENTIALLY WITH VSV G PROTEIN

Pseudotyping retroviruses with the vesicular stomatitis virus (VSV) G protein creates virus with a polytropic host range and makes it possible to concentrate the virus by centrifugation. Calcium phosphate-mediated transfection yields retroviral titers in the range of 10^5 to 10^6 CFU/ml; however, concentration of the viruses can increase the titer by several log factors. Because concentration requires larger initial volumes of retroviral supernatants, best results are obtained if a stable cell line that expresses gag-pol and the gene of interest is prepared first. A VSV G protein expression vector is then transiently expressed in this cell line for production of retroviral supernatants. This method is adapted from Burns et al. (1993) and Matsubara et al. (1996). An alternative method is to introduce retroviral vector and VSV G protein-expressing plasmids by cotransfection (see Alternate Protocol 1); it should be possible to introduce *gag-pol* construct at the same time. Although the latter approach may be laborious for obtaining large quantities of retroviral supernatants for concentration, it may yield higher titers for proteins that are detrimental to the long-term growth of 293 cells.

Materials

- Retroviral packaging cell line with amphoteric host range (e.g., PA317, Ψ CRIP, GP+Am12, or Bing; see Table 9.9.1) and appropriate culture medium
- Retroviral vector DNA
- gag-pol-expressing cell line (e.g., Anjou65, ATCC CRL 11268; see Table 9.9.1) and appropriate culture medium
- 293 cell growth medium (see recipe)
- VSV G protein expression plasmid: e.g., pHCMV-VSV-G (Matsubara et al., 1996)
- TNE buffer (see recipe) or 1% (v/v) Hanks basic salt solution (HBSS; APPENDIX 2)
- 60-mm tissue culture dishes
- Sorvall RT-3000B centrifuge and rotor (or equivalent)
- Beckman L3-50 centrifuge and SW 41 rotor (or equivalent)
- Additional reagents and equipment for culturing of mammalian cells (APPENDIX 3F), transfecting cells and harvesting virus (see Basic Protocol 1), infecting cells (see Basic Protocols 3 or 4 or Alternate Protocols 2, 3, or 4), selecting clones with drugs (UNIT 9.5), assaying clones for the gene of interest (e.g., UNITS 9.6-9.8 & 9.10), and freezing cells (see Support Protocol)

Create a stable cell line

1. Transfect a retroviral packaging cell line with an amphotropic host range with 6 to 20 μg retroviral vector DNA (see Basic Protocol 1). Incubate 48 hr.

2. Twenty-four hours before harvesting retroviral supernatant, plate 2.5×10^6 gag-pol-expressing cells into the appropriate culture medium in 60-mm tissue culture dish.
3. Collect retroviral supernatant at 48 hr post transfection. Filter through a 0.45- μ m filter or centrifuge 5 min at $500 \times g$ (1000 rpm in Sorvall RT-3000B rotor), 4°C, to remove cells.
4. Use 1 ml retroviral supernatant to infect the gag-pol-expressing cell line.

For production of stable producer cell lines, infection is preferred to transfection because there is a reduced risk of rearrangement.

5. Select individual clones of stable producer cell lines—i.e., by drug selection (UNIT 9.5), FACS sorting, or other method.
6. Identify clones that express the gene of interest (e.g., UNITS 9.6-9.8 & 9.10).

These are the stable producer cell line(s) that will be used for subsequent experiments.

7. Aliquot and prepare frozen stocks of the stable producer cell line (see Support Protocol).

Pseudotype the producer cell line

8. Twenty-four hours prior to transfection, plate 2.5×10^6 stable producer cells in appropriate culture medium in a 60-mm dish.

If desired, the procedure can be scaled up for larger volumes. It is even possible to transfect spinner cultures (J.C. Burns, pers. comm.).

9. Transfect stable producer cells with 3 to 6 μ g VSV G protein expression plasmid by modified calcium phosphate-mediated transfection (see Basic Protocol 1). Incubate 24 hr.

Adjust the volumes proportionally if larger volumes are used in step 8.

10. Replace the medium with fresh 293 cell growth medium and continue incubation 24 to 48 hr.
11. Harvest retroviral supernatants between 48 and 72 hr by collecting the medium from the culture. Concentrate supernatant immediately (proceed to step 12) or store at -80°C .

Syncytium formation is evident ~36 hr following transfection. Syncytia are identified by the formation of multinucleated giant cells which are the result of membrane fusion caused by VSV G protein.

Concentrate VSV G protein-pseudotyped viral supernatant

12. Concentrate retroviral supernatant by centrifuging 90 min at $50,000 \times g$ (17,000 rpm in an SW 41 rotor), 4°C.
13. Resuspend the pellet overnight at 4°C in 20 μ l of TNE buffer or 0.1% HBSS.

Pellet is resuspended in 1/200 of viral culture volume.

14. Use concentrated retroviruses immediately for infection, store them frozen at -80°C , or concentrate the stock by another round of centrifugation (steps 12 and 13).

PSEUDOTYPING COTRANSFECTIONALLY WITH VSV G PROTEIN

It is also possible to produce a retrovirus pseudotyped with vesicular stomatitis virus (VSV) G protein by transient cotransfection. This method is adapted from J. Burns (pers. comm.) and Pear et al. (1996).

Materials

293 cells or 293-derived cells expressing retroviral gag-pol: e.g., Anjou65 (ATCC CRL 11269)
293 cell growth medium (see recipe)
Retrovirus vector DNA
VSV G protein expression plasmid: e.g., pHCMV-VSV-G (Matsubara et al., 1996)
60-mm tissue culture dishes

Additional reagents and equipment for culture of mammalian cells (*APPENDIX 3F*), modified calcium phosphate–mediated transient transfection of cells (see Basic Protocol 1), and concentrating pseudotyped viral supernatants (see Basic Protocol 2)

1. Twenty-four hours prior to transfection, plate 2.5×10^6 293 cells, or cells of a 293 derivative stably expressing the retroviral gag-pol genes, into 293 cell growth medium in a 60-mm dish.
2. Cotransfect 3 to 5 μg retroviral vector DNA and 3 to 5 μg of a VSV G protein expression plasmid by modified calcium phosphate–mediated transfection (see Basic Protocol 1). Incubate.

If 293 cells were plated in step 1, also cotransfect a gag-pol expression plasmid.

3. On the following day, replace the medium with 4 ml fresh 293 growth medium.
4. Harvest retroviral supernatants between 48 and 72 hr by collecting the culture medium. Concentrate the virus immediately (see Basic Protocol 2) or store at -80°C .

Syncytium formation is evident ~36 hr following transfection.

INFECTION OF ADHERENT CELLS WITH RETROVIRAL SUPERNATANT

This protocol gives methods for infecting adherent cells with retroviral supernatant. This is the most common method for determining the retroviral titer (*UNIT 9.10*) and testing for the presence of helper virus (*UNIT 9.13*). Infectious titer can be quantitated based on marker staining—e.g., using *lacZ* (*UNIT 9.10*), or green fluorescent protein (GFP; *UNIT 9.7C*)—or drug resistance (*UNIT 9.5*). Recent reports suggest that viral transduction by centrifugation (spin infection; see Alternate Protocol 2) will increase infectious titers 2- to 4-fold (Kotani et al., 1994). These methods are written for cells of the commonly used murine fibroblast line NIH 3T3; however, they should be adaptable to other adherent cell lines with minor modifications.

Materials

Retroviral supernatant, fresh or frozen (see Basic Protocols 1 or 2)
Target cells: e.g., NIH 3T3 cells
4 mg/ml polybrene in PBS (*APPENDIX 2*), filtered and stored at 4° or -20°C
Fibroblast growth medium
Antibiotic for drug selection (*UNIT 9.5*; optional)
10-cm tissue culture dishes

Additional reagents and equipment for culture of mammalian cells (*APPENDIX 3F*) and for assaying of reporter gene expression (e.g., *UNITS 9.6-9.8 & 9.10*) or drug selection (*UNIT 9.5*).

1. Twelve to eighteen hours prior to transfection, plate 5×10^5 NIH 3T3 cells into appropriate tissue culture medium in a 10-cm tissue culture dish.

Infection protocols can be scaled up or down. For a 60-mm plate, plate 2×10^5 cells and infect in a volume of 1 ml.

2. Prepare a 3-ml infection cocktail containing:

Retroviral supernatant: ≤ 1.5 ml

4 $\mu\text{g/ml}$ polybrene

Fibroblast growth medium to a final volume of 3 ml.

The retroviral supernatant should be $\leq \frac{1}{2}$ the final volume of the infection cocktail. 293 supernatants appear to contain a cytostatic factor that can be diluted out by changing the transfection medium 18 to 24 hr before retroviral harvest and by avoiding infectious stocks that contain viral supernatant at ratios $>1:1$ (Naviaux et. al., 1996).

3. Remove the medium from the dish of cells and add the infection cocktail. Incubate ≥ 3 hr.

If infection continues >6 hr, a 5-ml infection cocktail is recommended to prevent dehydration.

4. After the infection incubation, add 7 ml fibroblast growth medium to the dish. Incubate 36 to 48 hr.

*Drug selection (e.g., G418, puromycin) can begin 24 hr post infection (*UNIT 9.5*). If the cells are not over-confluent (i.e., if they are $<50\%$ confluent) at this time, it is unnecessary to split the cells prior to adding drug. If the cells are $>50\%$ confluent, it is necessary to split the cells as described in *UNIT 9.5*.*

5. Assay the cells for reporter gene expression (*UNITS 9.6-9.8 & 9.10*).

INFECTION OF ADHERENT CELLS BY SPIN INFECTION

Adherent cells can also be infected by centrifuging the cells in the presence of the infection cocktail (spin infection). This protocol is modified from the methods of Kotani et al. (1994).

Materials

Target cells: e.g., NIH 3T3 cells

Fibroblast growth medium

Retroviral supernatant

4 mg/ml polybrene in PBS (*APPENDIX 2*), filtered and stored at 4° or -20°C

6-well tissue culture plates

Beckman GS-6KR or Sorvall RT-3000B centrifuge with microplate carriers (or equivalent)

Additional reagents and equipment for culture of mammalian cells (*APPENDIX 3F*) and for assaying of reporter gene expression (e.g., *UNITS 9.6-9.8 & 9.10*) or drug selection (*UNIT 9.5*)

ALTERNATE PROTOCOL 2

1. Twelve to eighteen hours prior to infection, plate 1×10^5 target cells per well into fibroblast growth medium in a 6-well tissue culture plate.
2. For each well to be transfected, prepare a 4-ml infection cocktail containing:
 - Retroviral supernatant: ≤ 2 ml
 - 4 $\mu\text{g/ml}$ polybrene
 - Fibroblast growth medium up to 4 ml.

An infection cocktail of ≥ 4 ml is necessary to prevent dehydration of the cells.

The retroviral supernatant should be $\leq 1/2$ the final volume of the infection cocktail. 293 supernatants appear to contain a cytostatic factor that can be diluted out by changing the transfection medium 18 to 24 hr before retroviral harvest and by avoiding infectious stocks that contain viral supernatant at ratios $>1:1$ (Naviaux et. al., 1996).
3. Remove the medium from the cells and add the infection cocktail.
4. Centrifuge the cells 1.5 to 2 hr at $1000 \times g$ (2500 rpm in Beckman GS-6KR or Sorvall RT-3000B rotor), room temperature.
5. Return the cells to the incubator.
6. Change medium the next day and incubate until 48 hr post infection.
7. Assay for reporter gene expression (e.g., *UNITS 9.6-9.8 & 9.10*), add antibiotic for drug selection (*UNIT 9.5*), or test for effect of the gene of interest.

**BASIC
PROTOCOL 4**

**INFECTION OF NONADHERENT CELLS BY RETROVIRAL
SUPERNATANT**

Nonadherent cells can be infected by direct addition of retroviral supernatant or by cocultivation of nonadherent cells with retroviral producer cells (see Alternate Protocol 3). The advantage of cocultivation is that there is continuous retroviral production; however, this must be weighed against the disadvantage of harvesting producer cells together with target cells. This effect may be minimized by irradiating producer cells prior to cocultivation. Alternatively, nonadherent cells can be infected by spin infection (see Alternate Protocol 4). In general, infection by cocultivation is equivalent to spin infection.

Materials

- Retroviral supernatant
- 4 mg/ml polybrene in PBS (*APPENDIX 2*), filtered and stored at 4° or -20°C
- Target cell growth medium
- Exponentially growing nonadherent target cells
- 15-ml centrifuge tube
- 60-mm tissue culture dishes
- Sorvall RT-3000B centrifuge and rotor (or equivalent)
- Additional reagents and equipment for culture of mammalian cells (*APPENDIX 3F*) and for assaying of reporter gene expression (*UNITS 9.6-9.8 & 9.10*) or drug selection (*UNIT 9.5*)

1. Prepare a 3-ml infection cocktail containing:
 - Retroviral supernatant: ≤ 1.5 ml
 - 2 $\mu\text{g/ml}$ polybrene
 - Target cell growth medium up to 3 ml.

The retroviral supernatant should be $\leq 1/2$ the final volume of the infection cocktail. 293 supernatants appear to contain a cytostatic factor that can be diluted out by changing

the transfection medium 18 to 24 hr before retroviral harvest and by avoiding infectious stocks that contain viral supernatant at ratios >1:1 (Naviaux et. al., 1996).

Many nonadherent cells are more stringent in their growth requirements than fibroblasts. For this reason, it may be necessary to use a medium other than 293 cell growth medium during the 24 hr prior to viral harvest and/or add growth factors or cytokines to the medium. It is essential, however, that heat-inactivated fetal bovine serum (FBS) be used as the serum source.

2. Centrifuge 3×10^5 to 3×10^6 exponentially growing target cells 5 min at $500 \times g$ (1000 rpm in Sorvall RT-3000B rotor), 4°C or room temperature.
3. Remove supernatant and resuspend the cells in the infection cocktail at a final density of 10^5 to 10^6 cells/ml.
4. Add the cell suspension to a 60-mm dish and incubate 24 hr.
5. Twenty-four hours post transfection, centrifuge the cells 5 min at $500 \times g$, 4°C or room temperature, and resuspend the pellet in target cell growth medium. Incubate cells another 24 hr.
6. Assay for reporter gene expression (e.g., *UNITS 9.6-9.8 & 9.10*), add antibiotic for drug selection (*UNIT 9.5*), or determine infectious titer (*UNIT 9.10*).

INFECTION OF NONADHERENT CELLS BY COCULTIVATION

Nonadherent cells can also be infected by cocultivating them with retroviral producer cells for 48 hr.

Materials

Transfected packaging cells (see Basic Protocol 1) in 60-mm tissue culture dishes

Retroviral supernatant

4 mg/ml polybrene in PBS (*APPENDIX 2*), filtered and stored at 4° or -20°C

10^5 to 10^6 cells/ml nonadherent target cells

Target cell growth medium

15-ml conical centrifuge tubes

60-mm tissue culture dishes

Sorvall RT-3000B centrifuge (or equivalent)

Additional reagents and equipment for culture of mammalian cells (*APPENDIX 3F*) and for assaying of reporter gene expression (e.g., *UNITS 9.6-9.8 & 9.10*) or drug selection (*UNIT 9.5*)

1. Twenty-four hours after packaging cells have been transfected, prepare a 3-ml infection cocktail containing:
 - Retroviral supernatant: ≤ 1.5 ml
 - 2 $\mu\text{g/ml}$ polybrene
 - 3×10^5 to 3×10^6 nonadherent target cells
 - Target cell growth medium to 3 ml.

Replication of transfected retroviral packaging cells without loss of retroviral production can be achieved by irradiating them with 1500 rads (C. Zent, Univ. of Chicago, unpub. observ.) prior to addition of infection cocktail.

2. Aspirate the medium from the transfected packaging cells and gently add the infection cocktail to the cells. Incubate 24 hr.

To prevent detachment of adherent transfected packaging cells, add cocktail to the side of the dish rather than directly onto the cells.

**ALTERNATE
PROTOCOL 3**

**Introduction
of DNA into
Mammalian Cells**

9.11.11

3. Twenty-four hours after the infection cocktail is added (48 hr post transfection), gently remove 2 ml medium (which will contain many nonadherent cells) and transfer to a conical tube. Centrifuge 5 min at $500 \times g$, 4°C or room temperature.

It is not necessary to remove all of the nonadherent cells from the dish because the purpose of this step is to replenish the medium (which is acidic). Sufficient residual medium should remain on the dish to prevent dehydration during the short centrifugation step.

4. Prepare a fresh infection cocktail containing:

Retroviral supernatant: ≤ 1.5 ml (optional)
2 $\mu\text{g/ml}$ polybrene
Target cell growth medium to 3 ml.

Adding virus in this step is optional, but it may increase infection efficiency.

5. Remove the supernatant and gently resuspend the cell pellet in freshly prepared infection cocktail.
6. Using extreme care to avoid disruption of the adherent cell layer, add the fresh infection cocktail to the wall of the plate. Incubate 24 hr.
7. At 72 hr post transfection, collect nonadherent cells from the dish by gentle pipetting.

It is necessary to sufficiently wash the plate with FBS or medium to remove most nonadherent cells while minimizing contamination with adherent cells. Typically, there is about 10% contamination of nonadherent cells. Because adherent cells detach in sheets, rather than as individual cells, many of the cells can be removed by allowing the large clumps of cells to settle out and removing the others (see Support Protocol, step 5). Alternatively, cells can be transferred to a new 60-mm dish and given sufficient time for the adherent cells to reattach (~ 3 hr), after which the nonadherent cells can be removed.

8. Centrifuge the suspension of nonadherent cells 5 min at $500 \times g$, 4°C or room temperature. Resuspend the cells in target cell growth medium. Plate the target cells in 60-mm tissue culture dishes. Incubate an additional 24 to 48 hr.
9. Assay for reporter gene expression (e.g., *UNITS 9.6-9.8 & 9.10*) or add antibiotics for drug selection (*UNIT 9.5*), or assay for the gene of interest.

ALTERNATE PROTOCOL 4

INFECTION OF NONADHERENT CELLS BY SPIN INFECTION

Alternatively, nonadherent target cells can be infected by centrifuging the cells in the presence of infection cocktail, a method modified from Kotani et al. (1994).

Materials

Retroviral supernatant
4 mg/ml polybrene in PBS (*APPENDIX 2*), filtered and stored at 4° or -20°C
Target cell growth medium
Exponentially growing nonadherent target cells
24-well tissue culture plates
Sorvall RT-3000B centrifuge with microplate carrier (or equivalent)
10-cm tissue culture dishes

Additional reagents and equipment for culture of mammalian cells (*APPENDIX 3F*)
and for assaying of reporter gene expression (e.g., *UNITS 9.6-9.8 & 9.10*) or drug
selection (*UNIT 9.5*)

- For each well to be transfected, prepare a 2-ml infection cocktail containing:
 - Retroviral supernatant: ≤ 1 ml
 - 4 $\mu\text{g/ml}$ polybrene
 - Target cell growth medium to 2 ml.
- Add $1\text{--}2 \times 10^6$ exponentially growing nonadherent target cells per well of a 24-well tissue culture plate. Centrifuge 5 min at $500 \times g$ (1000 rpm in Sorvall RT-3000B rotor), 4°C or room temperature.

The cell number should be adjusted so that cells completely cover the surface of the well and form a monolayer following centrifugation.
- Remove the supernatant and add infection cocktail. Centrifuge the cells 1.5 to 2 hr at $1000 \times g$ (2500 rpm in Sorvall RT-3000B rotor), room temperature.
- Either immediately after centrifugation or after ≤ 6 hr of incubation, transfer the cells in the infection medium to a 10-cm dish. Dilute to 10 ml final volume with target cell growth medium. Incubate to 48 hr post infection.

If the cells are left overnight in only 1 to 2 ml of infection cocktail, the medium will be very acidic by the following morning.
- Assay cells for reporter gene expression (e.g., *UNITS 9.6-9.8 & 9.10*) or add antibiotic for drug selection (*UNIT 9.5*).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

293 cell growth medium

Dulbecco's modified Eagle medium (high glucose) containing:

- 10% (v/v) heat-inactivated fetal bovine serum (FBS)
- 100 U/ml penicillin
- 100 U/ml streptomycin
- 2 mM L-glutamine
- Store at 4°C

This medium should be stable for at least 3 months except that L-glutamine has a half-life of 1 month.

HEPES-buffered saline (HeBS), 2 \times

- 50 mM *N*-2-hydroxyethylpiperazine-*N'*-ethanesulfonic acid (HEPES)
- 10 mM KCl
- 12 mM dextrose
- 280 mM NaCl
- 1.5 mM Na_2HPO_4
- Adjust pH to 7.05 ± 0.05
- Filter through 0.2- μm filter
- Store in 50- μl aliquots at -20°C

Avoid multiple freeze/thaw cycles. To thaw, warm to room temperature and invert or vortex the tube to achieve uniform mixing.

Although it is unclear why, the ability of the 2 \times HeBS solution to produce working calcium phosphate precipitates deteriorates after 6 months to 1 year, even when the solution is stored at -20°C . Fresh 2 \times HeBS should therefore be prepared every 6 months.

TNE buffer

50 mM Tris-Cl, pH 7.8

130 mM NaCl

1 mM EDTA

Store at room temperature indefinitely

COMMENTARY

Background Information

Creation of high-titer retroviral supernatants by transient transfection offers several advantages over the creation of stable cell lines, especially for experimental purposes. The transient methods described in this unit take advantage of the high transfectability and high protein expression in 293 cells. When retroviral vectors (e.g., see Figs. 9.9.5 & 9.9.6) are expressed with the retroviral packaging constructs in these cells, the vector RNA (containing the Ψ site) is efficiently packaged and budded from the surface of the packaging cells. The infectious retrovirus vector may be harvested by removing the packaging cell supernatant. Because the transient methods are not dependent on stable integration of the introduced plasmid(s), high-titer viral supernatant may be harvested from the packaging cells from 36 to 96 hr following transfection. Two different approaches, both utilizing 293 cells, have successfully produced helper-free high-titer retroviral supernatants expressing a wide variety of genes. In the method of Pear et al. (1993), three cell lines (Anjou, Bosc23, and Bing) were made by stably transfecting plasmids expressing gag-pol (Anjou), gag-pol and ecotropic envelope (Bosc23), or gag-pol and amphotropic envelope (Bing) into 293T cells. 293T cells are a derivative of 293 cells that express the SV40 large T antigen (DuBridge et al., 1987). These cell lines were selected for high levels of expression of reverse transcriptase and the appropriate envelope protein. All of the retroviral structural genes are expressed from the Moloney murine leukemia virus (MoMuLV) long terminal repeat (LTR). High-titer retroviral supernatants may be made from these cell lines by introducing either a retroviral vector alone (Bosc and Bing) or a retroviral vector and a plasmid-expressing envelope (Anjou). Because the retroviral structural genes are expressed from two different stably integrated plasmids that contain multiple mutations (Danos and Mulligan, 1988), more than two recombination events are necessary for the creation of replication-competent retrovirus. To date, generation of helper virus by these cell lines has not been reported.

A second generation retroviral packaging cell line, termed ϕ NX, has recently been developed for high-titer, transient production (P. Achacoso and G.P. Nolan, pers. comm.). The major addition is that the *gag-pol* construct has an IRES-CD8 surface marker, so it is possible to follow gag-pol expression in living cells on a cell-by-cell basis. This facilitates checking the line for continued high-level gag-pol expression using CD8 as proxy. Additionally, non-MoMuLV-based promoters are used to express *gag-pol* (and *env*) genes, further minimizing the potential for recombination with vector LTRs. Otherwise, the construction of these cell lines is similar to Bosc23. Amphotropic and ecotropic versions of these have been prepared, and they stably express gag-pol and env proteins over several months. These lines are helper-free and produce retroviral supernatants that have titers similar to Bosc23 and Bing. The methods described in this unit work well with the ϕ NX cell line.

In the strategies of Finer et al. (1994), Soneoka et al. (1995) and Naviaux et al. (1996), both the retroviral vector and the retroviral packaging constructs are cotransfected into 293 cells and retroviral supernatants are harvested 48 hr later. These groups found that titers were highest when the packaging constructs were expressed from a cytomegalovirus (CMV) promoter and a retroviral vector in which the 5' LTR was a hybrid between CMV and MoMuLV sequences. Although previous reports suggested that cotransfection of retroviral vector and packaging constructs could lead to replication-competent retrovirus (Danos and Mulligan, 1988; Pear et al., 1993), these groups have engineered their constructs to minimize regions of homology that might lead to helper virus formation. In particular, deletion of 3' LTR sequences appears important for preventing the formation of replication-competent retrovirus (Soneoka et al., 1995).

Critical Parameters

Several parameters are important to optimize for high-titer retroviral production. Cells should be at optimal density (usually ~80% confluent) with minimal clumping at the time

of transfection. In addition to cell density, production of high-titer retroviral supernatants also requires that the cells continue to grow for ≥ 24 hr following transfection. A rough indicator is that the plate should be nearly 100% confluent ~ 24 hr after transfection. Several factors that may account for failure to obtain this degree of confluence are the nature of the protein expressed by the retroviral vector, the quantity of input DNA, and the length of chloroquine and/or sodium butyrate treatment. Expression of some proteins (such as P210bcr/abl) appears detrimental to 293 cell growth. High-titer retroviral supernatants can be obtained with vectors expressing such genes by increasing the number of cells plated prior to transfection and/or decreasing the quantity of transfected retroviral plasmid DNA. For expression of some genes, 293 cells may need to be nearly 100% confluent at the time of transfection. In addition to the protein product, it appears that the quantity of transfected DNA may have a toxic effect upon the cells. For up to 10 to 15 μg transfected DNA per 60-mm plate, increasing the amount of input DNA increases retroviral titer; however, as more DNA is added beyond that point, the quantity of DNA has a toxic effect on the cells. As discussed in the protocols, to prevent toxicity from chloroquine or sodium butyrate, cells should not be exposed to either of these reagents for ≥ 12 to 14 hr. When trying to optimize titers from a new retroviral construct, it is recommended that cell density, input DNA, and length of chloroquine and/or sodium butyrate treatment all be optimized. If the cells are not confluent at 24 hr post transfection, but reach confluence by 48 hr post transfection, it may be possible to obtain high-titer supernatants between 48 and 72 hr post transfection. Supernatants harvested >96 hr post transfection do not produce high retroviral titers.

Another variable that must be optimized is the transfection reaction. Of particular importance is the pH of the $2\times$ HEPES-buffered saline (HeBS), which should be 7.05 ± 0.05 . For reasons that are not clear, $2\times$ HeBS loses its effectiveness between 6 months and 1.5 years. Unfortunately, there is no way to detect this except for a bad experimental result. For this reason, it is recommended that fresh stocks of $2\times$ HeBS be prepared every 6 months.

Although the modified calcium phosphate-mediated transfection method is inexpensive and gives consistently good results, other methods of gene transfer, including lipid-mediated

transfection (*UNIT 9.4*), work equally well. A recent report suggests that receptor-mediated, adenovirus-augmented transfection is superior to calcium phosphate-mediated transfection for some cell types (von Ruden et al., 1995). It has not yet been determined if this method is superior for gene transfer into 293 cells.

There are several techniques that may increase retroviral titer—adding factors to the transfection or cell growth medium, amplifying the plasmid DNA, decreasing the incubation temperature, and infecting by centrifugation. Addition of $25 \mu\text{M}$ chloroquine during transfection may increase titers 2-fold; presumably this is due to the lysosomal neutralizing activity of the chloroquine. Sodium butyrate has been reported to increase retroviral titers from several-fold to 1 log by acting as a transcriptional activator of the MoMuLV LTR, as well as other promoters including CMV (Soneoka et al., 1995). Increasing the extracellular concentration of dNTPs may increase retroviral titers up to 10-fold by enhancing endogenous reverse transcription (Zhang et al., 1995). 293T cells, a 293 cell derivative, contain a temperature-sensitive large T antigen, whose permissive temperature is 32°C (DuBridge et al., 1987). Thus, incubation at this temperature in conjunction with the use of constructs that contain the SV40 origin of replication may result in plasmid replication and higher retroviral titers. This may come at a cost, however, as large T-induced replication is associated with higher recombination rates (St.-Onge et al., 1993). Shifting the cells to lower incubation temperatures (32° to 34°C) following transfection has been reported to increase retroviral titers by 5- to 15-fold (Kotani et al., 1994). This is hypothesized to result from both improved retroviral survival and increased production. Similarly, infecting by centrifugation may increase retroviral infection by 4- to 10-fold (Kotani et al., 1994). It is important to note that all of the above methods have given variable results in different investigators' hands, and each should be optimized (see Troubleshooting).

Unlike retroviruses expressing the amphotropic and ecotropic envelopes, VSV G protein-pseudotyped retroviruses can be efficiently concentrated 100- to 1000-fold because they can better withstand shearing forces during centrifugation. This attribute together with their polytropic host range makes them a very useful laboratory tool, especially for stable gene transfer into nonmammalian cells.

Troubleshooting

To maximize retroviral production, it is recommended that conditions be optimized using a nontoxic reporter protein such as lacZ (*UNIT 9.10*), green fluorescent protein (GFP; *UNIT 9.7C*), placental alkaline phosphatase (PLAP), or a cell-surface marker. During optimization, both transfection frequency of the producer cells and infection rate of the target cells should be measured. A number of variables that can affect retroviral production have been previously discussed (see Critical Parameters).

To avoid drift of the packaging cell lines or highly transfectable 293 cells, it is recommended that they be maintained for ~20 passages (2 months) before establishing a fresh stock by thawing a frozen vial of cells. It may be possible to carry the cells for longer periods without loss of efficacy, and it is recommended that the cells be assayed at various times to check the ability to produce high-titer supernatants. Bosc and Bing cells were made using 293T cells that contain selectable markers for neomycin (1 mg/ml), hygromycin (400 µg/ml), and mycophenolic acid (50 µg/ml). There does not appear to be any advantage in growing the cells in selection medium. If retroviral titers drop when using these cells, their growth may be tested by growing in hygromycin- and xanthine-guanine phosphoribosyltransferase-containing selection medium. If a majority of the cells are not resistant to these antibiotics, a fresh vial of earlier-passage cells should be thawed rather than attempting to reselect the cells. Similarly, when cells are initially received, aliquots should be tested for resistance to these antibiotics. 293 cells do not contain a selectable marker.

It should be noted that the amphotropic envelope-expressing Bing cells release ~100 units of infectious retrovirus per ml of retroviral supernatant expressing hygromycin resistance which results from the introduction of the gene for hygromycin resistance into these cells by a retroviral vector (W. Pear, unpub. observ.). Replication-competent retrovirus has not been detected in this cell line. The φNX-A cells do not produce a hygromycin-resistant retrovirus (G.P. Nolan, pers. comm.).

For efficient infection, addition of a polycation increases infection efficiency by >100-fold (reviewed in Burns et al., 1993). Polybrene gives the best results at concentrations between 2 and 8 µg/ml. Polybrene may, however, be toxic to some cell types. For example, NIH 3T3 cells tolerate 4 µg/ml polybrene for at least 18

hr; however, treatment of these cells with 8 µg/ml for this period is detrimental. If high polybrene concentrations are chosen, the medium should be changed after 3 to 5 hr. For nonadherent cells, including primary bone marrow cultures, 1 to 2 µg/ml polybrene is well tolerated for ≤48 hr. It is possible to use other cations, such as protamine sulfate or poly-L-lysine; however, infectious titers are diminished by 2- to 4-fold with these agents (Burns et al., 1993).

Efficient infection of nonadherent cells has been cumbersome due to the need to cocultivate the retroviral producer cells with the target cells. It is difficult to separate the producer cells from target cells. An additional difficulty is imposed by 293 cells which, because of their high density, tend to rapidly acidify medium. For this reason, it is important to replenish medium daily when performing cocultivations. An alternative to cocultivation is spin infection. This technique appears to produce equivalent results as cocultivation for several nonadherent cell lines (W. Pear, unpub. observ.). Although the original protocols of Kotani et al. (1994) indicate that infections should be performed at 32°C, many centrifuges are unable to maintain that temperature; equivalent results are obtained by spin infecting at room temperature (W. Pear, unpub. observ.).

One major problem with retroviral gene transfer is that, despite producing high-titer retroviruses, it is difficult to maintain gene expression following infection. Although there are a number of potential mechanisms at the transcriptional and posttranscriptional levels, a teleologic explanation is that the retrovirally expressed protein is detrimental to cell growth resulting in outgrowth of clones that express lower levels of the protein. This is a particular problem with vectors that contain an internal promoter expressing a selectable marker, because even though the marker is selected there is no selective pressure to express the gene of interest. One method to increase the likelihood that the gene of interest will be expressed (although there may still be posttranscriptional mechanisms that affect the level of expression) is to use a polycistronic retroviral vector (see *UNIT 9.10*). Other methods of overcoming gene toxicity are being developed. Particularly promising are the tetracycline-inducible promoters originally described by Gossen and Bujard (1992); these promoters have shown some degree of success when incorporated into retro-

viral vectors (Hoffman et al., 1996; Paulus et al., 1996).

Anticipated Results

Transfection efficiencies of 293 cells should be in the range of 30% to 60%. Infectious titers of 10^6 to 10^7 infectious U/ml should be expected from most retroviral vectors when titered on murine NIH 3T3 cells. Titers may be slightly reduced with large cDNA inserts (>4 kb). Titers may be lower on other cell types, particularly cells expressing only the amphotropic envelope receptor. For VSV G protein pseudotyping followed by concentration by centrifugation, titers of 10^9 infectious U/ml should be expected.

Time Considerations

Plating cells, transfections, infections, and medium changes may take from <1 hr/day to several hours per day depending on the number of dishes and transfections/infections. Production of viral supernatant and infection of target cells each require 3 to 4 consecutive days.

Literature Cited

- Burns, J.C., Friedmann, T., Driever, W., Burrascano, M., and Yee, J.K. 1993. Vesicular stomatitis virus G glycoprotein-pseudotyped retroviral vectors: Concentration to a very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 90:8033-8037.
- Danos, O. and Mulligan, R.C. 1988. Safe and efficient generation of recombinant retroviruses with amphotropic and ecotropic host ranges. *Proc. Natl. Acad. Sci. U.S.A.* 85:6460-6464.
- DuBridg, R.B., Tang, P., Hsia, H.C., Phaik-Mooi, L., Miller, J.H., and Calos, M.P. 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol. Cell. Biol.* 7:379-387.
- Finer, M.H., Dull, T.J., Qin, L., Farson, D., and Roberts, M.R. 1994. kat: A high-efficiency retroviral transduction system for primary human T lymphocytes. *Blood* 83:43-50.
- Gossen, M. and Bujard, H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U.S.A.* 89:5547-5551.
- Graham, F.L. and van der Eb, A.J. 1973. Transformation of rat cells by DNA of human adenovirus 5. *Virology* 54:536-539.
- Graham, F.L., Smiley, J., Russell, W.C., and Naish, R. 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* 36:59-72.
- Hofmann, A., Nolan, G.P., and Blau, H.M. 1996. Rapid retroviral delivery of tetracycline-inducible genes in a single autoregulatory cassette. *Proc. Natl. Acad. Sci. U.S.A.* 93:5185-5190.
- Kotani, H., Newton, P.B.R., Zhang, S., Chiang, Y.L., Otto, E., Weaver, L., Blaese, R.M., Anderson, W.F., and McGarrity, G.J. 1994. Improved methods of retroviral vector transduction and production for gene therapy. *Hum. Gene Ther.* 5:19-28.
- Landau, N.R. and Littman, D.R. 1992. Packaging system for rapid generation of murine leukemia virus vectors with variable tropism. *J. Virol.* 66:5110-5113.
- Matsubara, T., Beeman, R.W., Shike, H., Besansky, N.J., Mukabayire, O., Higgs, S., James, A.A., and Burns, J.C. 1996. Pantropic retroviral vectors integrate and express in cells of the malaria mosquito, *Anophelex gambiae*. *Proc. Natl. Acad. Sci. U.S.A.* 93:6181-6185.
- Naviaux, R.E., Costanzi, E., Haas, M., and Verma, I.M. 1996. The pCL vector system: Rapid production of helper-free, high-titer recombinant retroviruses. *J. Virol.* 70:5701-5705.
- Paulus, W., Baur, I., Boyce, F.M., Breakefield, X.O., and Reeves, S.A. 1996. Self-contained, tetracycline-regulated retroviral vector system for gene delivery to mammalian cells. *J. Virol.* 70:62-67.
- Pear, W., Nolan, G., Scott, M., and Baltimore, D. 1993. Production of high-titer helper-free retroviruses by transient transfection. *Proc. Natl. Acad. Sci. U.S.A.* 90:8392-8396.
- Pear, W.S., Scott, M.L., and Nolan, G.P. 1996. Generation of high titer, helper free retroviruses by transient transfection. *In* Methods in Molecular Biology: Methods in Gene Therapy, (P. Robbins, ed.) Humana Press: Totawa, NJ.
- Sanes, J.R., Rubenstein, J.L.R., and Nicolas, J.F. 1986. Use of recombinant retrovirus to study post-implantation cell lineage in mouse embryos. *EMBO J.* 5:3133-3142.
- Schlegel, R., Tralka, T.S., Willingham, M.C., and Pastan, I. 1983. Inhibition of VSV binding and infectivity by phosphatidylserine: Is phosphatidylserine a VSV-binding site? *Cell* 32:639-646.
- Soneoka, Y., Cannon, P.M., Ramsdale, E.E., Griffiths, J.C., Romano, G., Kingsman, S.M., and Kingsman, A.J. 1995. A transient three-plasmid expression system for the production of high-titer retroviral vectors. *Nucl. Acids Res.* 23:628-633.
- St-Onge, L., Bouchard, L., and Bastin, M. 1993. High-frequency recombination mediated by polyomavirus large T antigen defective in replication. *J. Virol.* 67:1788-95.
- von Ruden, T., Stingl, L., Cotten, M., Wagner, E., and Zatloukal, K. 1995. Generation of high-titer retroviral vectors following receptor-mediated, adenovirus-augmented transfection. *Biotechniques* 18:484-489.
- Yee, J.K., Friedmann, T., and Burns, J.C. 1994. Generation of high-titer pseudotyped retroviral vectors with very broad host range. *Methods Cell Biol.* 43:99-112.

Zhang, H., Duan, L.X., Dornadula, G., and Pomerantz, R.J. 1995. Increasing transduction efficiency of recombinant murine retrovirus vectors by initiation of endogenous reverse transcription: Potential utility for genetic therapies. *J. Virol.* 69:3929-3932.

Contributed by Warren Pear
University of Pennsylvania
Philadelphia, Pennsylvania

Large-Scale Preparation and Concentration of Retrovirus Stocks

UNIT 9.12

For some applications, such as infection of cells *in vivo*, it is necessary to concentrate retrovirus stocks in order to increase their titer. This is typically due to a limitation in the volume that can be injected at any one site. In addition, if one wishes to infect a population of cells uniformly with a virus that does not encode a selectable marker, high-titered stocks are necessary. This is because the initial infection must be relied upon, without selection, to generate as many infected cells as possible. In order to prepare the highest-titered stock for multiple experiments, several hundred milliliters to a few liters of producer cell supernatant can be prepared, titered, and concentrated (and tested for helper-virus contamination if this is required for the application at hand).

Because viruses are macromolecular structures, they can be concentrated fairly easily by a relatively short centrifugation step. The viral particles can either be pelleted (see Basic Protocol) or centrifuged onto a sucrose density step gradient. Alternatively, virions can be precipitated using polyethylene glycol or ammonium sulfate, and the resulting precipitate collected by centrifugation (see Alternate Protocol 1). After resuspension of the precipitates, the virions can be used directly, or further purified either by sedimentation onto sucrose density gradients, or by column filtration as described here. Perhaps most simply of all, small volumes of virus stock can be concentrated by centrifugation through a filter that allows only small molecules to pass (see Alternate Protocol 2).

NOTE: Retroviral particles are fragile, with short half-lives even under optimum conditions. When handled too roughly, they lose infectivity. Thus, resuspension of pellets and other procedures must be carried out gently, and materials should be kept cold.

PREPARATION OF VIRUS STOCK AND CONCENTRATION BY CENTRIFUGATION

**BASIC
PROTOCOL**

In this procedure, the virus stock is harvested after producer cells are split from newly confluent cultures and grown for 2 to 3 days. The stock is titered and then concentrated by centrifugation. Upon suspension of the pellet, the concentrated virus is again titered and frozen in small aliquots. The example given below is for a murine replication-incompetent Ψ 2 stock.

Materials

- Identified high-titer Ψ 2 producer cells, 50% to 90% confluent (UNIT 9.10)
- Complete DMEM containing 10% calf serum (DMEM-10, APPENDIX 3F, prepared with calf serum instead of FBS)
- NIH 3T3 cells
- 800 μ g/ml polybrene
- 0.45- μ m filters for large volumes (e.g., Nalgene 115-ml or 500-ml filters)
- Beckman JA-14 rotor with 250-ml centrifuge bottles (or equivalent) if concentrating large volumes *or* Beckman SW-27 or SW-41Ti rotor (or equivalent) if concentrating small volumes

NOTE: All incubations involving tissue culture cells should be performed in a humidified incubator 37°C, 5% CO₂ unless otherwise noted.

**Introduction of
DNA into
Mammalian Cells**

Contributed by Constance Cepko

Current Protocols in Molecular Biology (1997) 9.12.1-9.12.6

Copyright © 1997 by John Wiley & Sons, Inc.

9.12.1

Supplement 37

Prepare virus stock

1. Split producer cells 1:10 or 1:20 from newly confluent cultures into twenty to fifty 10-cm dishes. Incubate cells until they are 50% to 90% confluent.
2. Discard medium and add half the normal volume of medium (e.g., add 5 ml for a 10-cm dish). Make sure that the medium is not alkaline and add the medium gently as the entire monolayer may peel off if stressed. Incubate 2 to 3 days.

The degree of confluence suggested here is based upon both theoretical and empirical considerations. The greater the number of producer cells, the greater the virus production; however, if cells are past confluence for too long, they may reduce their production of virus and, in some cases, the monolayer of cells will peel off the dish. In our experience with Ψ2 producers, the best titers are obtained as described here. However, for different packaging lines, optimum conditions may vary—e.g., the supernatant may need to be harvested just as the cells reach confluency, or the medium may need to be changed to half volume the day before confluency and harvested the following day.

3. Harvest the supernatant, filter through a 0.45- μ m filter, and store at -70° or -80° C, or titer and concentrate immediately (steps 4 to 8).

The cells will be extremely densely packed when the harvest is made and the medium will be yellow. Nonetheless, this method of harvest seems to work well for several producers (e.g., Isolde and Ψ2).

4. Titer the virus stock as described in UNIT 9.10.

It is customary to titer before concentration. If the titer is sufficiently high, it can be aliquoted and used without concentration. However, if the highest possible titer is desirable, stocks should be concentrated. Stocks can be concentrated before being titered, although sometimes this is wasted effort as the stock may be very low in titer before concentration and thus not worth concentrating.

Concentrate virus stock

5. For large volumes, centrifuge the virus stock 20 min at $25,000 \times g$ (e.g., 14,000 rpm in a JA-14 rotor), 4° C, in 250-ml bottles that have been sterilized by autoclaving. For smaller volumes, centrifuge 10 min in an SW-27 or SW-41Ti rotor at 20,000 rpm using tubes that have been rinsed with 70% ethanol (to reduce contamination with bacteria).

This step serves to pellet nonviral, cellular debris.

6. Pour the supernatant directly into sterile bottles and centrifuge 5 to 16 hr at 14,000 rpm, 4° C, if using a JA-14 rotor or 2 hr at 20,000 rpm if using an SW-27 or SW-41Ti rotor.

The longer the spin, the harder the pellet, and the easier it is to pour off the supernatant (step 7) without dislodging the pellet. However, it is harder to resuspend the pellet after the longer spins. With centrifugation of large-volume stocks (e.g., >200 ml), we recommend the longer spins and the use of a Beckman JA-14 rotor or its equivalent. With smaller volumes, we recommend the SW-27 or SW-41Ti swinging bucket rotors with shorter spins (2 hr at 20,000 rpm, 4° C). Swinging-bucket rotors make a firm pellet at the bottom of the tube, so that it is easy to pour off the supernatant without dislodging the pellet; in addition, it is easier to see the pellet in tubes spun in swinging-bucket rotors.

7. Discard supernatant while carefully avoiding discarding the pellet, which sometimes dislodges when using the large volumes.

If using the JA-14 rotor, when removing the bottles from the rotor, mark the side of the bottle that should contain the pellet. The pellet will be smeared down the side of the bottle and will be difficult to see. Although normally the supernatant is discarded, a

small volume of the supernatant may be saved to titer to determine whether the virus was indeed pelleted.

8. Gently resuspend the pellet in DMEM-10 using 0.1% to 1% of the original volume.

Resuspension may take 2 hr if the pellet is fairly sticky. It is convenient to leave the centrifuge bottle in an ice bucket in the tissue culture hood and pipet the suspension approximately every 15 min. A 1- or 2-ml pipet can be used if large volumes were concentrated. If so, leave the original pipet in the bucket and use this pipet for the complete resuspension. This is because the virus pellet often sticks to the first pipet that touches it. Sometimes the pellet can be seen as an almost translucent, gooey yellow pellet. If the JA-14 rotor was used, be sure to wash down the sides of the bottle, as this is where much of the stock will be located. To resuspend pellets made in the SW-27 or SW-41 rotors, use a pipettor suitable for small volumes. In either case, avoid making bubbles, as this can denature the viral proteins.

Alternatively, an easier way to resuspend pellets made in the swinging-bucket rotors is to pour off almost all of the supernatant, e.g., leaving only 0.10 to 0.15 ml in an SW-27 large bucket; this may require keeping a Pasteur pipet handy to remove the last few drips of supernatant from the mouth of the tube as one pours it off. Return the tube to the bucket and replace the bucket cap. Place the buckets on a shaker at low speed in a 4°C cold room, or place them in an ice bucket and slowly shake 1 to 2 hr on the benchtop. After shaking, gently pipet up and down several times and transfer to a smaller tube.

Titer and freeze concentrated virus

9. Store the virus at -70° or -80°C in small aliquots (typically 5 to 50 µl).

Systematic tests in our laboratory have indicated that, in general, stocks can be frozen and thawed several times with no loss of titer. However, we have occasionally observed loss of titer over time in the freezer, or upon freeze-thaw. Thus, the most conservative approach is to make small aliquots and to minimize the number of times they are thawed and refrozen.

We have also found that small volumes can desiccate over long-term storage. To combat this, we place a drop of the oil used to cover PCR reactions onto each small aliquot before freezing.

10. Titer an aliquot of concentrated and unconcentrated virus as described in UNIT 9.10.

The titer can also be determined using an aliquot before freezing the stock.

CONCENTRATION BY PEG PRECIPITATION AND CHROMATOGRAPHY

A method that is rapid and easy is to precipitate virus with polyethylene glycol, with an optional step of chromatography on Sepharose (Aboud et al., 1982).

Additional Materials

- 5 M NaCl, filter sterilized
- Polyethylene glycol (PEG) 6000, filter sterilized
- NTE buffer (see recipe)
- Sepharose CL-4B or CL-2B (Pharmacia)
- Savant high-speed centrifuge *or* Beckman SW-41Ti rotor
- Additional equipment for preparation of Sepharose column (UNIT 5.6)

Precipitate virus with PEG

1. Prepare a virus stock (see Basic Protocol, steps 1 to 4)
2. Add 5 M NaCl to the virus stock to 0.4 M final while stirring at 4°C.

ALTERNATE PROTOCOL 1

Introduction of DNA into Mammalian Cells

9.12.3

3. Slowly add PEG 6000 to 8.5% final (w/v) and continue stirring 1 to 1.5 hr at 4°C.

For reasons that are not clear, some batches of PEG do not work well as they lead to problems in resuspending the pellet.

4. Collect the precipitate by centrifuging 10 min at 7000 × g (e.g., at 7500 rpm in a SW-41Ti rotor or 10,000 rpm in a Savant high-speed centrifuge.)
5. Dissolve the pellet in NTE buffer in 1% of the original volume. If necessary, proceed to step 6 for column chromatography.

The stock can be used directly after this step or stored at -70° or -80°C. It can also be further purified on a Sepharose CL-4B column to remove PEG.

Carry out Sepharose chromatography

6. Prepare a column of Sepharose CL-4B or CL-2B as described in UNIT 5.6 (Support Protocol), equilibrating in NTE buffer.

Column size will depend upon the amount of virus pellet; 10-ml Econo-Columns from Bio-Rad work well for pellets prepared from 100 ml of virus stock.

7. Apply the virus to the top and chromatograph at 1 ml/min with NTE buffer.
8. Collect 0.3- to 0.5-ml fractions.
9. Assay fractions for virus by measuring absorbance at 280 or 260 nm (see Aboud et al., 1982) or by reverse transcriptase (RT) assay (UNIT 9.13).
10. Pool appropriate fractions and titer as in UNIT 9.10.

For examples of virus recovery as determined by both Xgal CFU bioassay and RT assay, see Figure 9.12.1.

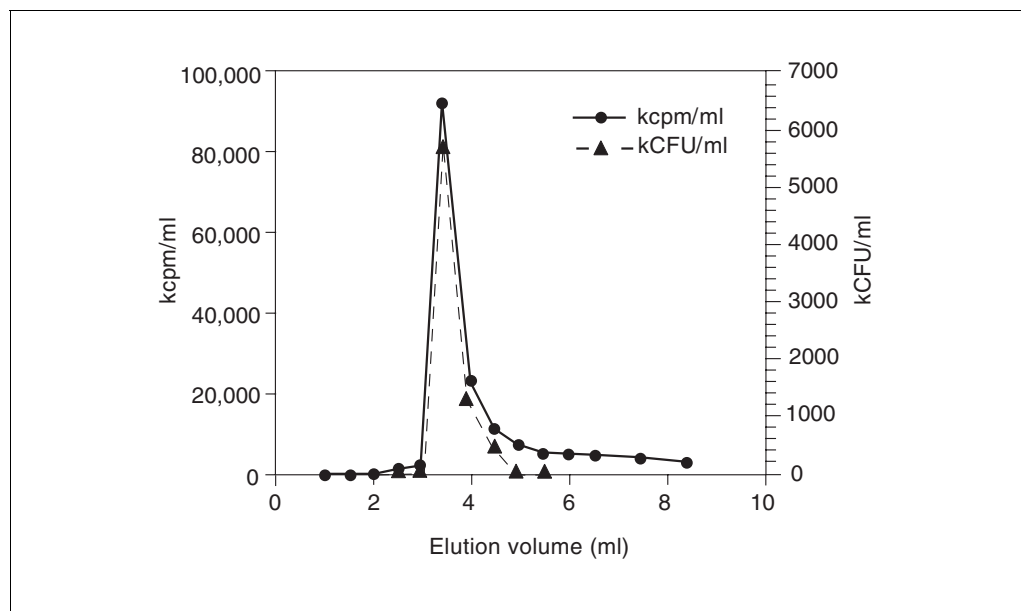


Figure 9.12.1 Ψ2 BAG supernatant (8 ml) was concentrated 100-fold by PEG precipitation. The resulting 80 μl was loaded onto 10 ml of Sepharose CL-2B in a Bio-Rad Econo-Column. Fractions of 0.5 ml were collected and tested for both reverse transcriptase activity and Xgal CFU.

CONCENTRATION USING MOLECULAR-WEIGHT-CUTOFF FILTERS

This method is the simplest means of concentrating small to medium volumes of retrovirus stocks. Supernatants prepared as in steps 1 to 4 of the Basic Protocol are centrifuged through filters, either the Centricon-30 microconcentrator from Amicon or the CentriCell 60 from Polysciences, essentially following the manufacturers' instructions. These filters allow passage of molecules that are much smaller than the virus (e.g., Centricon-30 filters allow passage of molecules of <30 kDa). For example, supernatants can be centrifuged at $5000 \times g$ in a Sorvall SS-34 centrifuge (Stoker et al., 1990); the material that does not flow through the filter is removed and used as the concentrated stock. The amount of time that the stock is centrifuged is chosen according to the degree that one wishes to concentrate the stock. The longer the centrifugation, the more concentrated the stock.

REAGENTS AND SOLUTIONS

NTE buffer

100 mM NaCl
10 mM Tris·Cl, pH 7.4 (APPENDIX 2)
1 mM EDTA
Filter sterilize

COMMENTARY

Troubleshooting

When resuspending the pellet that results from centrifugation alone or from PEG precipitation followed by centrifugation, complete dissolution may not be possible; there are some stocks that have small clumps remaining, even after much effort is spent in resuspending them. (The clumps are most likely composed of viral particles and perhaps cellular debris.) Do not overdo the pipetting, as the solution has a very high protein concentration and will degenerate into a foamy mess with inactivated virus. The virus is not very hardy and can lose infectivity rapidly if mishandled. If the resuspended pellet is very clumpy, spin out the clumps at a low speed (e.g., 1000 rpm for 5 min in a tabletop centrifuge, or 1 min in a microcentrifuge). Store the supernatant at -70° or -80°C and then continue to attempt resuspension of the remaining clumps in pellet if more recovery is necessary.

The PEG protocol can work well with murine viruses; it has not been attempted with avian viruses in this laboratory. However, as indicated above, lots of PEG can apparently vary, causing problems in resuspension of the pellet. Changing the PEG lot and reducing the amount of serum in the virus stock may solve the resuspension problem. Producer cells can be grown in the usual 10% serum until they reach confluency, and then can be changed to reduced protein—either 2% serum or 10% NuSerum (Col-

laborative Biomedical Products) until the supernatant is harvested 2 to 3 days later.

If a concentrated virus is used to infect cells *in vitro*, a toxicity problem may be encountered, which is quite obvious within 12 hr of infection. Apparently, some cells fuse and form multinucleated syncytia when incubated with high concentrations of virions. This is probably due to virions causing membrane fusion between the plasma membranes of two or more cells, presumably via the virion, which has its own membrane and may fuse with the cellular membranes after binding to the viral receptor. To minimize this problem, it may be necessary to dilute the virus somewhat, reduce the time of cells' exposure to the inoculum, and/or plate the target cells such that very few are touching each other.

Anticipated Results

The recovery of virus using the centrifugation protocol is usually not 100%, although it may be this high on rare occasions. Usually, it is 10% to 30%. If the reduction in volume is $100\times$, and the recovery is 10%, then the titer increases $10\times$, which is useful for many experiments. If greater increases in titer are needed, when resuspending the pellet from the centrifugation protocol, use a volume for resuspension that is $\geq 200\times$ less than the original volume of the virus stock. One can also attempt to centrifuge the first concentrated stock and resuspend in $0.1\times$ the volume. However, the stock that

results is extremely viscous and clumpy and difficult to use for most experiments.

When the PEG protocol works (i.e., there is no problem in resuspending the pellet), it works easily and well, yielding highly infectious stocks of 100× concentration. This is a rapid method that can be easier than the centrifugation method. It can yield virus stocks that are less clumpy than the centrifugation method, especially when column chromatography is used.

Time Considerations

The pelleting/centrifugation protocol can be completed in 3 to 4 hr, including labeling tubes, aliquoting, etc. The PEG protocol, from the time of precipitating the stock to resuspending the pellet, takes ~2 hr. Column chromatography takes ~1 hr to set up and run the column.

Assaying column fractions can take 2 hr if using the RT assay, and 2 to 14 days if performing a bioassay such as Xgal or G418-R CFU determination.

Literature Cited

- Aboud, M., Wolfson, M., Hassan, Y., and Huleihel, M. 1982. Rapid purification of extracellular and intracellular Moloney murine leukemia virus. *Arch. Virol.* 71:185-195.
- Stoker, A.W., Hatier, C., and Bissell, M.J. 1990. The embryonic environment strongly attenuates v-src oncogenesis in mesenchymal and epithelial tissues, but not in endothelia. *J. Cell Biol.* 111:217-228.

Contributed by Constance Cepko
Harvard Medical School
Boston, Massachusetts

Helper virus is a replication-competent virus that is sometimes present in stocks of replication-incompetent virus. There are several types of applications in which the presence of helper virus can be problematic. If animal infections are being done, helper virus can lead to leukemia, particularly if the infection is carried out pre- or neonatally. If retroviruses are being used for lineage analysis, helper virus may cause horizontal spread of the marker virus, creating false lineage relationships.

This unit describes protocols for the detection of helper virus by a selectable marker assay (see Basic Protocol), by rescue of an integrated provirus (see Alternate Protocol 1), or by measuring reverse transcriptase activity (see Alternate Protocol 2).

DETECTION OF HELPER VIRUS THROUGH HORIZONTAL SPREAD OF DRUG RESISTANCE

BASIC PROTOCOL

The following is a very sensitive method of examining the ability of a virus supernatant to promote the horizontal spread of a viral genome from an infected cell to neighboring nonsibling cells. Target cells, such as NIH 3T3 or CEF cells, are infected with the test virus stock, and after several generations of growth the supernatant is harvested. This supernatant is then assayed for the presence of virus carrying the marker of the initial stock—e.g., neomycin resistance (*neo*). The presence of virus that can transmit this marker indicates that the original stock contained helper virus. The assay is easy to perform for a variety of vectors and also allows the host range of the helper virus to be tested (see Background Information). The example below is for a murine virus. An alternative, equally sensitive assay to detect helper virus is rescue of an integrated provirus that expresses an easily assayable marker, such as lacZ (see Alternate Protocol 1).

For controls in this assay, a helper virus-free stock and a helper-containing stock are used. The former has usually been verified in previous experiments as helper-free. A helper-containing stock (wild-type or replication-competent) is usually obtained from a retrovirologist. Alternatively, a helper-containing stock can be made by obtaining a plasmid encoding a wild-type replication-competent provirus and transfecting it (UNIT 9.10) into NIH 3T3 cells (for murine virus) or CEF cells (for avian virus). After transfection, the cells are passaged for 1 to 2 weeks (with 2 µg/ml polybrene for the murine virus and avian subgroup E virus) to allow for spread. The supernatant is then harvested and stored at -70° or -80°C. Titration of a wild-type stock can be accomplished by the XC plaque assay for murine viruses (Rowe, 1970), by the RT assay (see Alternate Protocol 2), or by immunohistochemical staining of *gag* for avian viruses (Stoker and Bissell, 1987). If the RT assay is used, a stock of a known titer is needed for comparison.

Materials

NIH 3T3 cells (used only with murine viruses)

Three titered virus stocks: test virus containing *neo* marker, helper virus-free control containing *neo* marker (negative control), and wild-type helper virus
800 µg/ml polybrene, filter sterilized (stored at -20°C)

Complete DMEM containing 10% calf serum (DMEM-10, APPENDIX 3F; prepared with calf serum instead of FBS)

6-cm dishes

0.45-µm filter

Additional reagents and equipment for determination of viral titer (UNIT 9.10)

Introduction of DNA into Mammalian Cells

9.13.1

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

NOTE: All incubations involving tissue culture cells should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise noted.

1. The day before the assay is initiated, split NIH 3T3 cells 1:50 into three 6-cm dishes. Label one dish as a positive control, one dish as a negative control, and one dish for the stock to be tested.

A 1:50 split is used instead of the standard 1:10 or 1:20 split to allow time for the virus to spread before the cells become confluent. Viruses can only integrate into mitotic cells.

2. Prepare test virus stock: Add 0.01 ml of 800 µg/ml polybrene (8 µg/ml final) to 1 ml of test virus supernatant. Filter this solution through a 0.45-µm filter.

The test virus stock carries the neo or other selective marker.

If a concentrated virus stock is used (UNIT 9.12), 1 to 30 µl of the concentrated stock is diluted in 1 ml medium, and 0.01 ml polybrene (8 µg/ml final) is added. This solution is passed through a 0.45-µm filter immediately before use.

The volume of test virus to be used for the infection depends on the degree to which the stock must be scrutinized, although typically 1 ml of an unconcentrated stock is used.

3. Prepare a positive control stock: Prepare a mixture of 1 ml helper-free stock and 10 to 100 CFU of helper stock. Add 0.01 ml of 800 µg/ml polybrene (8 µg/ml final) and filter through a 0.45-µm filter.

The helper-free virus stock carries the neo or other selective marker. It is used for detection purposes. The helper stock is used to allow horizontal spread of the neo genome of the helper-free stock.

The volumes of helper-free and helper stocks to be used depends on the titers of these stocks and the titer of the test stock. The amount (CFU) of helper-free virus should equal the amount (CFU) of the test virus. Only a small amount (10 to 100 CFU) of helper virus is required so that one can monitor the sensitivity of the assay for small amounts of helper-virus contamination.

4. Prepare a negative control stock: Prepare 1 ml of helper-free stock alone by adding 0.01 ml of 800 µg/ml polybrene (8 µg/ml final) and filtering through a 0.45-µm filter.

5. Infect one dish (step 1) with each virus stock.

6. Place all three infected dishes 1 to 3 hr in a CO₂ incubator at 37°C to permit absorption. Add 4 ml DMEM-10 and continue incubation until cells reach confluency (~3 to 4 days).

Polybrene concentration must be maintained at 2 µg/ml for the length of the assay to allow spread of both helper and marker-containing viruses.

7. Split cells 1:50 into new 6-cm dishes. Set up only one new dish for each of the three dishes, and discard the unused portion of originally infected cells. Include polybrene at 2 µg/ml (final). Incubate until cells reach 50 to 90% confluency.

Although not necessary, it is advisable to save some of the initial supernatants (e.g., remove ~2 ml supernatant and store at -80°C) prior to splitting in this step. If there is a helper virus present, the relative titers of these initial supernatants and the final supernatants (step 9) may be informative concerning the amount of helper present in the stock.

8. Discard medium and replace with half the volume of fresh DMEM-10. Incubate an additional 2 to 3 days.

At this point, the supernatant of the dish inoculated with the test virus stock will contain virus only if the original test stock contained helper virus.

9. Harvest the final supernatant from the confluent cells, filter through a 0.45- μ m filter, and add polybrene to 8 μ g/ml final. Store at -70° or -80° C or use immediately to titer and assay the virus for G418 resistance (*neo* marker).
10. Split fresh, uninfected NIH 3T3 cells 1:10 or 1:20 into three 6-cm dishes the day before titering the supernatants.
11. Use 1 ml of each supernatant (from step 9) to infect NIH 3T3 cells and carry out the remaining titration steps as described in *UNIT 9.10* (Basic Protocol 2, steps 1 to 6).

If the test stock is an avian replication-incompetent virus stock, use CEF cells instead of NIH 3T3 cells. In steps 1 and 7 split them 1:10 (they do not grow well if split more dilute than 1:10). The helper-free and helper-containing negative and positive control stocks must also be avian. Polybrene is required for avian viruses of subgroups other than A. Make sure that the CEF cells are from a strain that can support infection with the subgroup under study (Morgan and Fekete, 1996).

PROVIRAL RESCUE TO DETECT REPLICATION-COMPETENT RETROVIRUS

ALTERNATE PROTOCOL 1

This alternate protocol describes another assay, equal in sensitivity to the drug resistance assay presented in the Basic Protocol, for examining the ability of a virus supernatant to promote the horizontal spread of a viral genome from an infected cell to neighboring nonsibling cells. This assay uses rescue of an integrated provirus containing a marker gene to detect the presence of replication-competent retrovirus. In the example below, an indicator cell line that contains a single integrated copy per cell of a retroviral vector expressing β -galactosidase (*lacZ*) is infected with a retroviral stock. If the stock contains the retroviral packaging genes, the integrated provirus will be packaged and it will thereby be possible to transfer this genome to sibling cells. The presence of helper virus can be easily detected by assaying for *lacZ* activity by histochemical staining (*UNIT 9.10*). In order to obtain sufficient retroviral titers to detect the rescued virus, it is necessary to passage the secondarily infected cells for several weeks in the presence of polybrene. In this example, the *lacZ*-bearing indicator provirus is integrated into NIH 3T3 cells, and thus only retroviruses capable of infecting these cells can be assayed for helper activity. Similar indicator cells can be constructed for retroviruses with other host ranges by using cell lines appropriate for this host range and selecting a clone that contains only a single integrated provirus.

Additional Materials (also see *Basic Protocol*)

BAG-3T3 indicator cells: NIH 3T3 cells containing a single copy per cell of an integrated BAG provirus (Fig. 9.10.2)

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

NOTE: All incubations involving tissue culture cells should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise noted.

1. The day before the assay is initiated, split BAG-3T3 cells 1:50 into three 6-cm dishes. Label one dish as a positive control, one dish as a negative control, and one dish for the stock to be tested.

A 1:50 split is used instead of the standard 1:10 or 1:20 split to allow time for the virus to spread before the cells become confluent. Viruses can only integrate into mitotic cells.

2. **Prepare test virus stock:** Add 0.01 ml of 800 µg/ml polybrene (8 µg/ml final) to 1 ml of test virus supernatant. Filter this solution through a 0.45-µm filter.

Unlike in the Basic Protocol, the test virus stock does not need to carry a selectable marker.

If a concentrated virus stock is used (UNIT 9.12), 1 to 30 µl of the concentrated stock is diluted in 1 ml medium, and 0.01 ml polybrene (8 µg/ml final) is added. This solution is passed through a 0.45-µm filter immediately before use.

The volume of test virus to be used for the infection depends on the degree to which the stock must be scrutinized, although typically 1 ml of an unconcentrated stock is used.

3. **Prepare a positive control stock:** Prepare a stock that contains 10 to 100 CFU of helper virus in 1 ml. Add 0.01 ml of 800 µg/ml polybrene (8 µg/ml final) and filter through a 0.45-µm filter.

The helper stock is used to allow horizontal spread of the integrated lacZ-expressing provirus. It may be desirable to assay several dilutions of the helper stock to determine the sensitivity of the marker rescue assay.

4. **Prepare a negative control stock:** Prepare 1 ml of helper-free stock alone by adding 0.01 ml of 800 µg/ml polybrene (8 µg/ml final) and filtering through a 0.45-µm filter.

5. Infect one dish (step 1) with each virus stock.

6. Place all three infected dishes 1 to 3 hr in a CO₂ incubator at 37°C to permit absorption. Add 4 ml DMEM-10 and continue incubation until cells reach confluency (~3 to 4 days).

Polybrene concentration must be maintained at 2 µg/ml for the length of the assay to allow spread of both helper and marker-containing viruses.

7. Split cells 1:50 into fresh 6-cm dishes. Set up only one new dish for each of the three dishes, and discard the unused portion of originally infected cells. Include polybrene at 2 µg/ml final. Incubate until cells reach 50% to 90% confluency.

Although not necessary, it is advisable to save some of the initial supernatants (e.g., remove ~2 ml supernatant and store at -80°C) prior to splitting in this step. If there is a helper virus present, the relative titers of these initial supernatants and the final supernatants (step 9) may be informative concerning the amount of helper present in the stock.

8. Discard medium and replace with ½ vol fresh DMEM-10. Incubate an additional 2 to 3 days.

At this point, the supernatant of the dish inoculated with the test virus stock will contain virus only if the original test stock contained helper virus.

It may be necessary to maintain the cells for 1 to 2 additional passages (i.e., repeat step 7) in order to detect helper virus.

9. Harvest the final supernatant from the confluent cells, filter through a 0.45- μm filter, and add polybrene to 8 $\mu\text{g}/\text{ml}$ final. Store at -70° or -80°C or use immediately to titer and assay the virus for β -galactosidase (*UNIT 9.10*, Support Protocol 2).
10. Split fresh, uninfected NIH 3T3 cells 1:10 or 1:20 into three 6-cm dishes the day before titering the supernatants.
11. Use 1 ml of each supernatant (from step 9) to infect NIH 3T3 cells and carry out the remaining titration steps as described in *UNIT 9.10* (Basic Protocol 2, steps 1 to 6).

REVERSE TRANSCRIPTASE ASSAY TO DETECT HELPER VIRUS

Virus supernatants generated in the Basic Protocol (step 9) can be monitored for the incorporation of radioactive dTTP. This serves as an assay for reverse transcriptase activity that will only be observed when helper virus is in the original stock. This assay also can be used for other purposes—e.g., to test packaging lines for maintenance of virus production.

Additional Materials (also see *Basic Protocol*)

- Reverse transcriptase (RT) reaction cocktail (see recipe)
- Virus supernatants: test and control samples (see Basic Protocol, step 9)
- 2 \times SSC (*APPENDIX 2*)
- 95% ethanol
- 96-well microtiter dish
- Whatman DE52 or DE81 paper, precut to 2.5-cm circles or sheets cut to accommodate the number of samples

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

NOTE: All incubations involving tissue culture cells should be performed in a humidified 37 $^\circ\text{C}$, 5% CO_2 incubator unless otherwise noted.

1. Add 50 μl RT reaction cocktail to separate wells of a 96-well microtiter dish. Set up one well for each test or control sample.
2. Add 10 μl of each virus supernatant to its designated well.
3. Cover microtiter dish and place 1 to 2 hr in a CO_2 incubator at 37 $^\circ\text{C}$.
4. Spot 10 μl of each reaction onto a 2.5-cm circle of DE52 or DE81 paper, set up on a piece of plastic wrap. Label the paper with a soft (no. 2) lead pencil.

Samples can also be spotted on a large sheet of paper, which is then cut into individual sample pieces after the washing process.

5. Place paper in a tray and cover with 2 \times SSC. Wash 20 min with gentle shaking on a shaker at room temperature. Discard SSC and repeat wash step twice.
6. Soak 1 min in 95% ethanol.
7. Air dry \sim 10 min.
8. Count in scintillation counter or expose to film.

The higher the level of helper virus the more virions in the supernatant—and thus the higher the level of reverse transcriptase activity (see Background Information).

**ALTERNATE
PROTOCOL 2**

**Introduction
of DNA into
Mammalian Cells**

9.13.5

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Reverse transcriptase (RT) reaction cocktail

Prepare an appropriate quantity of RT reaction cocktail as indicated by Table 9.13.1.

Table 9.13.1 Preparation of RT Reaction Cocktail

Reagent ^a	Number of assays			
	20	40	80	100
1 M Tris·Cl, pH 8.3, μ l	50	100	200	250
0.2 M DTT, μ l	100	200	400	500
0.02 M MnCl ₂ , μ l	30	60	120	150
3 M NaCl, μ l	20	40	80	100
100 μ g/ml oligo(dT), μ l	50	100	200	250
100 μ g/ml poly(rA), μ l	100	200	400	500
0.2 mM dTTP, μ l	50	100	200	250
1% NP-40, μ l	50	100	200	250
[α - ³² P]dTTP, μ Ci ^b	100	200	400	500
Final volume (add dH ₂ O), ml	1	2	4	5

^aCocktail components are for assay of murine RT (from Goff et al., 1981). For avian RT, substitute 1% 2-mercaptoethanol (final) for DTT; 10 mM MgCl₂ (final) for MnCl₂; and 100 μ M dTTP (final) for 10 μ M dTTP (Omer and Faras, 1982).

^bSpecific activity >500 Ci/mmol. Because volume will vary with different lots, the volume of distilled H₂O to add should be varied to compensate.

COMMENTARY

Background Information

The production of wild-type, replication-competent helper virus by packaging cell lines can be an issue of concern when using replication-incompetent vectors. If the vectors are being used for lineage analysis, the virus stock must be free of helper virus. Similarly, for most infections of young animals, where the wild-type viruses are leukemogenic, helper viruses must not be present.

The genome that supplies the *gag*, *pol*, and *env* genes in the Ψ 2, Ψ am, PA12, PA317, and Q2bn packaging cell lines (UNIT 9.9) does not encode the Ψ sequence, but can still become packaged—although at a low frequency. If the helper genome is coencapsidated with the vector genome, recombination in the next cycle of reverse transcription can occur. If recombination allows the helper genome to acquire the Ψ sequence from the vector genome, a recombinant that is capable of autonomous replication will result. This recombinant can spread through an entire culture (although slowly due to envelope

interference). Once this occurs, it is best to discard the producer clone, as there is no convenient way to eliminate the wild-type virus.

As would be expected, helper virus contamination happens with a greater frequency in stocks with high titer. The PA317 packaging line has a safer design than Ψ 2, Ψ am, PA12, and Q2bn, while the Ψ CRE, GP + E-86, Ω E, Ψ CRIP, and Isolde lines probably have the best design for not producing helper virus. Recombination leading to helper virus creation has not yet been observed in Ψ CRE, GP + E-86, Ω E, Ψ CRIP, or Isolde, but one must be cautious and assay stocks produced by these lines (for a review of lines, see Table 9.9.1).

These protocols can also be used to determine the host range of the helper virus. This can be tested by infection of cells of different species [e.g., dog (D17) or human (HeLa) cells cannot be infected with ecotropic virus, only with amphotropic virus].

Alternative methods to detect helper virus include the XC plaque assay (Rowe et al.,

1970) and S+L⁻ assay (Eckner and Hettrick, 1977; Miller and Buttimore, 1986). These assays require additional cell lines. The assay listed here for horizontal spread is quite sensitive but very time-consuming. The RT assay is sensitive, but does not always detect crippled helper (virus that can replicate, but is mutant and does not replicate as well as wild-type).

Troubleshooting

If supernatants are assayed by titration of selectable markers and many (e.g., greater than several thousand) marker-resistant colonies are observed, the original test stock contained wild-type helper virus. Once this occurs, it is best to discard the producer clone, as there is no convenient way to eliminate the helper virus. If a smaller number of colonies result than expected (see Anticipated Results), there are several possible explanations. There may be “crippled” helper in the original test stock generated by an imperfect recombination event. For many applications, crippled helper is not a confounding influence. Alternatively, some of the Ψ^- genome (e.g., the genome carrying *gag*, *pol*, and *env*) can transfer to the first set of infected NIH 3T3 cells and lead to a low titer, without true helper virus contamination (e.g., see Danos and Mulligan, 1988).

Anticipated Results

Virus stock that contains wild-type helper virus should produce more than several thousand colonies carrying the *neo* marker (G418-resistant) colonies on the assay dishes generated by step 9 of the Basic Protocol or several thousand *lacZ*⁺ colonies generated by step 11 of Alternate Protocol 1. If the reverse transcriptase assay is used, the RT activity from the test supernatants should be compared to that in the control supernatants (cells infected with a known helper-free virus stock, and cells infected with a stock of known wild-type helper virus). Stocks contaminated with helper virus should show a >1000-fold enhancement of RT activity when compared to helper-free controls. A truly helper-free stock will be equal in RT activity to that generated by the helper-free control.

Time Considerations

Virus supernatant generation (Basic Protocol) requires 10 days. To titer the stocks for presence of viruses carrying the selectable marker or *lacZ*, another 10 to 13 days are needed (see UNIT 9.11). To test supernatants for reverse transcriptase activity (Alternate Protocol 2), 50 samples can be processed in 2 to 3 hr.

Literature Cited

- Danos, O. and Mulligan, R.C. 1988. Safe and efficient generation of recombinant retroviruses with amphotropic and ecotropic host ranges. *Proc. Natl. Acad. Sci. U.S.A.* 85:6460-6464.
- Eckner, R.J. and Hettrick, K.L. 1977. Defective Friend spleen focus-forming virus: Interfering properties and isolation free from standard leukemia-inducing helper virus. *J. Virol.* 24:383-396.
- Goff, S., Trakman, P., and Baltimore, D. 1981. Isolation and properties of murine leukemia virus mutants: Use of a rapid assay for release of virion reverse transcriptase. *J. Virol.* 38:239-248.
- Miller, A.D. and Buttimore, C. 1986. Redesign of retrovirus packaging cell lines to avoid recombination leading to helper virus production. *Mol. Cell. Biol.* 6:2895-2902.
- Morgan, B.A. and Fekete, D.M. 1996. Manipulating gene expression with replication competent retroviruses. *Methods Cell Biol.* 51:185-218.
- Omer, C.A. and Faras, A.J., 1982. Mechanism of release of the avian retrovirus tRNA^{TIP} primer molecule from viral DNA by ribonuclease H during reverse transcription. *Cell* 30:797-805.
- Rowe, W.P., Pugh, W.E., and Hartley, J.W. 1970. Plaque assay techniques for murine leukemia viruses. *Virology* 42:1136-1139.
- Stoker, A.W. and Bissell, M.J. 1987. Quantitative immunocytochemical assay for infectious avian retroviruses. *J. Gen. Virol.* 68:2481-2485.

Contributed by Constance Cepko
Harvard Medical School
Boston, Massachusetts

Warren Pear (proviral rescue)
University of Pennsylvania
Philadelphia, Pennsylvania

Retrovirus Infection of Cells In Vitro and In Vivo

UNIT 9.14

There are many applications in which retrovirus vectors are used as transduction agents. In some cases, the vector carries a gene that one wishes to express in a target cell in order to study the function of that gene. In other cases, the virus is used to introduce a histochemical marker gene into cells in order to follow their fate. Retrovirus vectors can also be used in a variety of cells type to investigate regulatory sequences in which a reporter gene and regulatory sequences are carried by the vector and to immortalize or transform primary cells by transduction of oncogenes.

For each application, the infection protocol may vary and must often be optimized. Guidelines for infection of cells in some typical *in vivo* and *in vitro* experiments are presented here. In addition to these guidelines, the cited literature contains protocols that have been used successfully to infect cells under specific circumstances and may provide a good starting point for similar cells. To optimize infection of a particular type of cell, it is often advantageous to use vectors that are easy to score. These include vectors carrying selectable or screenable markers (such as BAG; Figs. 9.10.1 and 9.10.2).

CAUTION: When working with human blood, cells, or infecting agents, strict biosafety practices must be followed (see UNIT 16.15).

INFECTION OF CELLS IN VITRO

Infection of target cells *in vitro* is accomplished by simply incubating the virus with the cells (e.g., UNIT 9.10, Basic Protocol 2). For most *in vitro* applications using a murine virus, a polycation such as polybrene is used to aid viral infection. These polycations apparently promote virus binding to the host cell surface by reducing electrostatic repulsion between the negatively charged surfaces of the cell and virion. When incubating target cells with virus *in vitro*, problems may occur if undiluted, high-titer virus is used directly on some cell types—i.e., fusion of target cells can occur with subsequent death of most fused cells within a few days. This is presumably due to the virus serving as a fusogen (i.e., binding to the surface of more than one cell and promoting membrane fusion).

Alternatively, cells to be infected *in vitro* can be incubated with the packaging line that produces the desired vector (cocultivation

method). This method is used to infect hematopoietic cells and appears to greatly increase the infection efficiency (Williams et al., 1984; Lemischka et al., 1986). In some cases, it may be desirable to prevent cell division of the producer cells during or after cocultivation. This can be achieved by using protocols that allow virus production, but that prevent further cell division of the producer cells. For example, prior to cocultivation, confluent or nearly confluent producer cells can be killed by irradiation (2800 rad for NIH 3T3 cells, 1500 rad for 293 cells) or a 3-hr treatment with mitomycin C (10 μ g/ml in medium) followed by several rinses with medium. After this treatment, the cells to be infected are plated onto producer cells. Target cells should continue to grow, but producer cells will die as they can no longer divide. If the target cells are nonadherent, they can be removed by simply washing the producer monolayer gently after ~48 hr of cocultivation. An alternative to cocultivation is spin infection (UNIT 9.11; Kotani et al., 1994), which achieves similar infectious titers without the difficulties caused by the presence of producer cells.

It is difficult to generalize about the efficiency of infection, although it can approach 100%. For the most part, the variables that influence infectability of a given cell are unknown. However, it is clear that for optimal results, the cells should be as mitotically active as possible. Although virus will enter cells and undergo partial reverse transcription in nonmitotic cells, there will be no integration or expression of the viral genes. If virus is applied to cells that are not dividing, it may remain competent to finish reverse transcription and integration for 1 to 2 days, although this has not been carefully determined for most cell types. One approach that may overcome the difficulties in transducing non-dividing cells is the use of retroviral packaging cell lines and vectors based on human immunodeficiency virus (HIV; Naldini et al., 1996). This approach has successfully produced high-titer, helper-free viruses that are able to infect cell lines arrested at various phases of the cell cycle at frequencies that far exceed those attainable using Moloney-based packaging systems. In addition, these vectors are able to infect adult rat neurons at low frequency.

Expression of viral gene products is usually

**Introduction
of DNA into
Mammalian Cells**

9.14.1

assayed two or three cell cycles after infection. When assaying *lacZ* expression in NIH 3T3 cells via Xgal staining (UNIT 9.10), we have found that the number of clones peaks at ~48 hr postinfection (three cell cycles). However, Xgal staining is visible within 24 hr in a subset of the infected cells.

Infections of tissue explants or cultured embryos can also be performed essentially as described above. A tissue explant can be bathed in as much virus stock as is desired for a few hours (when using murine viruses or avian viruses of any subgroup other than A, include polybrene at 8 µg/ml), or it can be cultured over a monolayer of producer cells (as above in the cocultivation method).

Creation of Retroviral cDNA Libraries

Expression cloning has been extremely useful for isolating a wide variety of genes, including cell surface markers, cytokines, and multiple receptors (see UNIT 6.11). This methodology relies on high levels of gene expression in the transfected cell line and the ability to recover the appropriate cDNA by plasmid rescue. To accomplish the former, COS cells are frequently used for expression cloning, as the expression of SV40 large T antigen causes high levels of protein expression from vectors that contain the SV40 origin of replication.

Although expression cloning in COS cells has been extremely useful for the expression cloning of a number of cDNAs, this approach is limited to cell lines that are efficiently transfected and to phenotypic assays with a short readout period. These limitations have made it difficult to use expression cloning strategies in primary hematopoietic cells due to their poor transfectability and the long-term culture required for phenotypic selection. In contrast to expression cloning in COS cells, retroviral cDNA libraries offer the possibility of introducing cDNAs into a wide variety of cell types, including primary cells. As a result, it is possible to devise novel assays to select cDNAs that are based on specific properties of the cells themselves, such as transformation, growth factor dependence, or in vitro differentiation.

Retroviral expression cloning has been successfully used to clone cDNAs using NIH 3T3-based packaging lines. These strategies have been useful for identifying genes involved in factor-dependent growth (Rayner and Gonda, 1994; Wong et al., 1994) and transformation (Whitehead et al., 1995), as well as for the selection of peptides and antisense messages that induce etoposide resistance (Gudkov et al.,

1993, 1994). Although overcoming some of the limitations of COS-based expression cloning, use of NIH 3T3-based packaging lines has two major disadvantages for expression cloning. These are that (1) it is difficult to obtain high titers by transient transfection, limiting the complexity of the library, and (2) the frequency of each cDNA may change during selection of the transfected population and the resulting retroviral stock may not represent the original cDNA library. An advance in the ability to prepare cDNA expression libraries has been the methodology to prepare high titer retroviral supernatants by transient transfection (UNIT 9.11; reviewed in Onishi et al., 1996). These approaches overcome the major disadvantages of the NIH 3T3-based packaging cells—i.e., it is possible to obtain high transient transfection efficiencies and it is unnecessary to select the cells prior to harvest of the retroviral stock. An application of the use of transient retroviral production for expression cloning has been described for cloning cDNAs from factor-dependent cell lines (Kitamura et al., 1995). In this report, the authors were able to develop expression libraries with complexities of $>10^6$ and detect cDNAs at a frequency of 1 in 10^7 (Kitamura et al., 1995).

The technology for retroviral expression cloning is presently evolving, and the reader should consult the most recent literature for the best choice of packaging cell line and retroviral vector. With the methods for transient retroviral production, the basic strategy is to clone a cDNA library into a retroviral vector, transfect the library into the packaging cells, harvest the retroviral stock between 48 and 72 hr post-transfection, and infect the target cells. These methods are described in UNIT 9.11. The target cells are subsequently assayed for the phenotype of interest. Some considerations in the choice of retroviral vectors are the presence of multiple cloning sites for efficient cloning of the cDNAs into the vector, the presence of the extended packaging signal for improved titer (Bender et al., 1987), the use of IRES (internal ribosome entry site) elements rather than internal promoters to increase the likelihood that the cloned cDNA will be expressed, and the presence of primers that surround the multiple cloning site so that the cDNA can be amplified by PCR following phenotypic selection. Another consideration is the infection efficiency of the target cells. For example, in the studies of Onishi et al. (1996), the authors were able to infect Ba/F3 and BW5147 cells with $>50\%$ efficiency but infection of LG3 myeloid cells

resulted in an infection efficiency of only 8.3%. To obtain adequate infection efficiencies, it may be necessary to pseudotype the retroviruses (UNIT 9.11) or introduce a retroviral receptor into the cell line to be assayed (UNIT 9.9). One drawback to high infection efficiencies, however, is that with higher infection efficiency, the infected cells are more likely to contain multiple integrated proviruses (Onishi et al., 1996). This may increase the difficulty in identifying the cDNA that confers the phenotype of interest. Another consideration in the generation of retroviral cDNA libraries is whether the library should be normalized prior to screening. An advantage of normalization is that there will be an “enrichment” of rare messages; however, the cost is that the resulting cDNA library is no longer representative of initial population of cDNAs.

INFECTION OF RODENTS AND CHICKS IN VIVO

Transduction of genes via retrovirus vectors can be performed *in vivo*. The following paragraphs describe *in vivo* infection of rodents, the laboratory animal most often used in experimental procedures; however, the techniques apply to other species (e.g., avian) as well. For more detailed information on infection of chick embryos, see Fekete and Cepko (1993a,b), Homburger and Fekete (1996), and Morgan and Fekete (1996).

NOTE: Detailed protocols for the care and handling of laboratory animals are beyond the scope of this unit. The reader is referred to *Current Protocols in Immunology*, Chapter 1 (Donovan and Brown, 1995) for instructions on proper animal restraint, anesthesia, injections, and euthanasia techniques that are essential to the infection methods described.

Infection of Postnatal Animals In Situ

Volume and titer of virus stock. The volume that can safely be delivered to a tissue *in vivo* is generally quite small—0.1 to 1 μ l. It is therefore important to prepare high-titer (usually concentrated) virus stock (UNIT 9.12). In general, virus titer should range from 10^6 to 10^8 CFU/ml. Because of these limitations, it is quite important to deliver the virus directly to the area containing the highest percentage of mitotic target cells. It is unclear if there are factors in tissue fluids (e.g., ventricular fluids or cerebral spinal fluids) that inhibit or destroy viral infectivity.

Delivery of virus. Virus delivery to postnatal animals is fairly straightforward. It is possible

to use a hand-held Hamilton syringe with a 33-G needle or use a drawn-out glass pipet. If glass pipets are used, the size of the needle tip should be determined empirically for the tissue under study. The skin, and even the skull, are soft enough on the first few days after birth for direct injection into the tissue. For tougher injection sites on older animals, it may be necessary to first make a hole using a stronger needle, such as one made of steel, prior to inserting a more delicate needle at a specific site. Coinjection with a dye such as 0.05% (w/v) trypan blue or 0.025% fast green aids in detecting the accuracy of injections and does not impair viral infectivity. Rodents ≤ 7 days old can be anesthetized by simply cooling on ice for a few minutes. Landmarks (e.g., sutures or blood vessels) near the area to be injected can be visualized using a fiber-optics light source. The injection is made directly into the desired area using a hand-held pipet. It is best to practice a series of injections with dye alone and then immediately dissect the animal for examination of the injection site. After the injection method is perfected, virus can be injected. Animals can then be examined for evidence of viral infection at any time, depending upon the experimental design.

Setup of pilot experiments to optimize injection and expression efficiency. When performing infections of tissue for the first time, or when infection and expression efficiency of the target cells is unknown, it is useful to perform pilot experiments using a retrovirus vector carrying a histochemical marker gene such as *lacZ*. Infection with a *lacZ*-coding virus allows for examination of the injected animal a few days after infection to determine if the target cells have been infected and are capable of expressing the viral genes. For such experiments, it is necessary to optimize conditions of the Xgal staining reaction (UNIT 9.10) for the particular tissue under study.

An excellent way to simultaneously determine the accuracy of injection, and whether the Xgal histochemistry is working, is to inject cells that contain the *lacZ* gene, such as the $\Psi 2$ BAG producer cells (obtainable from ATCC, #CRL 9560; UNIT 9.10). The cells can be pre-labeled with a fluorescent dye such as carboxy-fluorescein diacetate succinimyl ester (CFSE, Molecular Probes; prepare 10 mM stock in DMSO and store in foil at 4°C). As detailed by Bronner-Fraser (1985), cells suspended in PBS are exposed to 0.3 mM CFSE solution at 37°C for 30 min just before trypsin treatment. After dissociation, the cells are pelleted and washed

twice with medium or PBS, and resuspended to a concentration of 10^8 cells/ml. A few minutes after injection of the CFSE-labeled cells, the animal should be killed, the tissue fixed, then stained with Xgal (UNIT 9.10). The injected cells can then be monitored independently from the Xgal staining by viewing the CFSE fluorescence under UV illumination. If the Xgal histochemistry is working properly, no fluorescent cells should be visible because Xgal-stained cells are usually so full of indigo dye that all fluorescence is absorbed. Once the injection method and Xgal histochemistry have been optimized, BAG virus is injected into the site and the tissue is examined for viral infection several days later.

If the Xgal histochemistry is *not* working well, fluorescent cells will be visible, indicating the site of the injection and the failure of the Xgal histochemistry. In this case, one should vary fixation and tissue preparation methods until good Xgal staining is achieved (Cepko, 1989). Alternatively, the DAP retrovirus vector (Fields-Berry et al., 1992) encoding human placental alkaline phosphatase can be used. A stable $\Psi 2$ line that makes DAP at titers of $1-2 \times 10^6$ CFU/ml (more reliably than $\Psi 2$ BAG) can be obtained from ATCC (CRL #1949). Histochemical detection of the alkaline phosphatase enzyme is as simple as Xgal histochemistry (Fields-Berry et al., 1992). For further discussion of both Xgal and alkaline phosphatase assays, see Cepko et al. (in press).

Infection of Prenatal Rodents In Utero

Injections made in utero (through uterine wall) are performed with drawn-out glass pipets (e.g., see Austin and Cepko, 1990). The actual diameter and shape of the tip should be determined empirically (e.g., for injections of the lateral ventricle through the uterine wall of rats at embryonic day 15—E15—the outside diameter of the tip is ~ 50 μm and the inside diameter is ~ 15 μm). Animals are anesthetized with a mixture of ketamine (20 to 40 mg/kg) and xylazine (3 to 5 mg/kg) and are opened via an incision along the midline. The orientation of an embryo can be determined by visualizing the head with fiber optics. Injections of 0.1 to 1.0 μl are made through the uterine wall into the area of interest. When the target of the injection is the lateral ventricle, the inoculum can be seen filling the lateral ventricle when properly delivered to this area. Other areas may not be as visible and practice injections followed by immediate dissection and examination are again recommended. The mother is

then closed with suture. The injected animals can be delivered prenatally, or allowed to finish gestation and be born normally, to be sacrificed at a postnatal age, depending upon the design of the experiment.

Infection of Prenatal Rodents Exo Utero

It is difficult to make precisely directed injections into many of the mitotic zones of prenatal mammals through the uterine wall. The exo utero surgical procedure developed by Muneoka et al. (1986) at least partially circumvents this problem. In this procedure, mouse embryos at embryonic days 11 to 19 (E11 to E19) are released from the uterus by cutting the uterine wall, but remain attached to the uterus via the placenta. The abdominal cavity of the mother is filled with a buffered saline solution to protect the embryos. An incision can be made in the extra-embryonic membranes that surround the embryo so that embryo can be directly manipulated or injected. Subsequently, the extra-embryonic membranes are closed with fine sutures. The embryos can be brought to term in the abdominal cavity and delivered by Caesarean section.

Injections can be made through an incision in the extra-embryonic membranes that are subsequently closed with 10-0 suture. However, in our hands (Turner et al., 1990), embryonic and neonatal survival was improved by injecting directly through the extra-embryonic membranes without an incision or suture. Although injections made by this approach were more difficult to target, 25% of the embryos injected at E13 survived to adulthood and all injected animals contained clones of retrovirally infected cells. Additional factors that influence the success of this method are choice of mouse strain and health of the mouse colony. Outbred mouse strains such as CD-1 or Swiss Webster appear to be best, but even these strains may have different embryo survival rates when obtained from different suppliers or colonies. Apparently healthy mouse colonies can harbor subclinical infections that do not affect unoperated embryos but that can stress operated embryos beyond their ability to survive.

Ex Vivo Infection of Murine Bone Marrow Cells and Bone Marrow Transplant

Ex vivo retroviral infection of murine bone marrow has been used for in vitro experiments as well as for subsequent transfer and rescue of

lethally irradiated animals. The goal of many of these experiments is to transduce hematopoietic stem cells; however, the methods discussed below can be altered to optimize infection of a particular hematopoietic cell type. As discussed below, the conditions for infecting hematopoietic stem cells are still being developed, and the reader should consult the most recent literature to ascertain the best conditions. Optimal conditions for irradiating mice need to be established for each irradiator and are also dependent on weight, age, and mouse strain.

Donor mice (age 5 to 8 weeks) are treated with 250 μg per gram (mouse weight) 5-fluorouracil (5-FU) by tail vein injection in a volume of $\leq 200 \mu\text{l}$. 5-FU treatment kills dividing cells, and in doing so, enriches for hematopoietic progenitor cells (van Zandt, 1984; Bodine et al., 1991). Five days later, donor mice are sacrificed and the tibia and fibias are removed. These bones are flushed with 2 to 3 ml per bone of PBS or DMEM using a 25- to 27-G needle attached to a syringe. The flushed marrow cells (including red blood cells) are pooled, centrifuged 10 min at $500 \times g$, 4°C , and resuspended in 10^6 cells/ml of a prestimulation cocktail consisting of DMEM, 15% fetal bovine serum, antibiotics (penicillin and streptomycin), 6 ng/ml IL-3, 10 ng/ml IL-6, and 100 ng/ml stem cell factor. The bone marrow cells are incubated in the prestimulation cocktail for 24 to 48 hr prior to infection. A number of other cytokines have been tried, including IL-1 α , and the concentrations of the cytokines used vary widely from lab to lab (Bodine et al., 1989, 1994; Daley et al., 1990; Fraser et al., 1992; Pear et al., 1996).

The cells are infected by either spin infection or cocultivation (UNIT 9.11) in the above infection cocktail with 4 $\mu\text{g}/\text{ml}$ polybrene added. In addition, viral supernatant may replace one-third of the volume of DMEM. Following infection, nonadherent cells are removed and injected into the tail vein of lethally irradiated 4- to 8-week-old recipient mice in a volume of $\leq 200 \mu\text{l}/\text{mouse}$. Approximately $2\text{--}5 \times 10^5$ unfractionated bone marrow cells per mouse are necessary to ensure long-term reconstitution of most mouse strains. The animals are subsequently monitored for bone marrow reconstitution.

LITERATURE CITED

- Austin, C.P. and Cepko, C.L. 1990. Cellular migration patterns in the developing mouse cerebral cortex. *Development* 110:713-732.
- Bender, M.A., Palmer, T.D., Gelinas, R.E., and Miller, A.D. 1987. Evidence that the packaging signal of Moloney murine leukemia virus extends into the gag region. *J. Virol.* 61:1639-1646.
- Bodine, D.M., Karlsson, S., and Nienhuis, A.W. 1989. Combination of interleukins 3 and 6 preserves stem cell function in culture and enhances retrovirus-mediated gene transfer into hematopoietic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 86:8897-8901.
- Bodine, D.M., McDonagh, K.T., Seidel, N.E., and Nienhuis, A.W. 1991. Survival and retrovirus infection of murine hematopoietic stem cells in vitro: effects of 5-FU and method of infection. *Exp. Hematol.* 19:206-212.
- Bodine, D.M., Seidel, N.E., Gale, M.S., Nienhuis, A.W., and Orlic, D. 1994. Efficient retrovirus transduction of mouse pluripotent hematopoietic stem cells mobilized into the peripheral blood by treatment with granulocyte colony-stimulating factor and stem cell factor. *Blood* 84:1482-1491.
- Bronner-Fraser, M. 1985. Alterations in neural crest migration by a monoclonal antibody that affects cell adhesion. *J. Cell Biol.* 101:610-617.
- Cepko, C.L. 1989. Lineage analysis and immortalization of neural cells via retrovirus vectors. In *Neuromethods*, Vol. 16: Molecular Neurobiological Techniques (A.A. Boulton, G.B. Baker, and A.T. Campagnoni, eds.) pp. 367-392. Humana Press, Clifton, N.J.
- Cepko, C.L., Ryder, E., Fekete, D.M., and Bruhn, S. In press. Detection of β -galactosidase and alkaline phosphatase activities in tissue. In *Methods in Cell Biology* (D. Spector, L. Leinwand, and R. Goldman, eds.). Cold Spring Harbor Laboratory Press. Cold Spring Harbor, N.Y.
- Daley, G.Q., Van Etten, R.A., and Baltimore, D. 1990. Induction of chronic myelogenous leukemia in mice by the P210 bcr/abl gene of the Philadelphia chromosome. *Science* 247:824-830.
- Donovan, J. and Brown, P. 1995. Care and handling of laboratory animals. In *Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 1.0.1-1.8.4. John Wiley & Sons, New York.
- Fekete, D.M. and Cepko, C.L. 1993a. Retroviral infection coupled with tissue transplantation limits gene transfer in the chick embryo. *Proc. Natl. Acad. Sci. U.S.A.* 90:2350-2354.
- Fekete, D.M. and Cepko, C.L. 1993b. Replication competent retroviral vectors encoding alkaline phosphatase reveal spatial restriction of viral gene expression/transduction in the chick embryo. *Mol. Cell. Biol.* 13:2604-2613.

- Fields-Berry, S.C., Halliday, A.L., and Cepko, C.L. 1992. Novel recombinant retrovirus encoding alkaline phosphatase confirms clonal boundary assignment in lineage analysis of murine retina. *Proc. Natl. Acad. Sci. U.S.A.* 89:693-697.
- Fraser, C.C., Szilvassy, S.J., Eaves, C.J., and Humphries, R.K. 1992. Proliferation of totipotent hematopoietic stem cells in vitro with retention of long-term competitive in vivo reconstituting ability. *Proc. Natl. Acad. Sci. U.S.A.* 89:1968-1972.
- Gudkov, A.V., Zelnick, C.R., Kazarov, A.R., Thimmapaya, R., Suttle, D.P., Beck, W.T., and Roninson, I.B. 1993. Isolation of genetic suppressor elements, inducing resistance to topoisomerase II-interactive cytotoxic drugs, from human topoisomerase II cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 90:3231-3235.
- Gudkov, A.V., Kazarov, A.R., Thimmapaya, R., Axenovich, S.A., Mazo, I.A., and Roninson, I.B. 1994. Cloning mammalian genes by expression selection of genetic suppressor elements: association of kinesin with drug resistance and cell immortalization. *Proc. Natl. Acad. Sci. U.S.A.* 91:3744-3748.
- Homburger, S.A. and Fekete, D.M. 1996. High efficiency gene transfer into the embryonic chick CNS using B-subgroup retroviruses. *Dev. Dyn.* 206:112-120.
- Kitamura, T., Onishi, M., Kinoshita, S., Shibuya, A., Miyajima, A., and Nolan, G.P. 1995. Efficient screening of retroviral cDNA expression libraries. *Proc. Natl. Acad. Sci. U.S.A.* 92:9146-9150.
- Kotani, H., Newton, P.B.R., Zhang, S., Chiang, Y.L., Otto, E., Weaver, L., Blaese, R.M., Anderson, W.F., and McGarrity, G.J. 1994. Improved methods of retroviral vector transduction and production for gene therapy. *Hum. Gene Ther.* 5:19-28.
- Lemischka, I.R., Raulet, D.H., and Mulligan, R.C. 1986. Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell* 45:917-927.
- Morgan, B.A. and Fekete, D.M. 1996. Manipulating gene expression with replication competent retroviruses. *Methods Cell Biol.* 51:185-218.
- Muneoka, K., Wanek, N., and Bryant, S.V. 1986. Mouse embryos develop normally exo utero. *J. Exp. Zool.* 239:289-293.
- Naldini, L., Blomer, U., Gally, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M., and Trono, D. 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272:263-267.
- Onishi, M., Kinoshita, S., Morikawa, Y., Shibuya, A., Phillips, J., Lanier, L.L., Gorman, D.M., Nolan, G.P., Miyajima, A., and Kitamura, T. 1996. Applications of retrovirus-mediated expression cloning. *Exp. Hematol.* 24:324-329.
- Pear, W.S., Aster, J.C., Scott, M.L., Hasserjian, R.P., Soffer, B., Sklar, J., and Baltimore, D. 1996. Exclusive development of T cell neoplasms in mice transplanted with bone marrow expressing activated Notch alleles. *J. Exp. Med.* 183:2283-2291.
- Rayner, J.R. and Gonda, T.J. 1994. A simple and efficient procedure for generating stable expression libraries by cDNA cloning in a retroviral vector. *Mol. Cell Biol.* 14:880-887.
- Turner, D.L., Snyder, E.Y., and Cepko, C.L. 1990. Lineage-independent determination of cell type in the embryonic mouse retina. *Neuron* 4:833-845.
- van Zandt, G. 1984. Studies of hematopoietic stem cells spared by 5-fluorouracil. *J. Exp. Med.* 159:679-690.
- Whitehead, I., Kirk, H., and Kay, R. 1995. Expression cloning of oncogenes by retroviral transfer of cDNA libraries. *Mol. Cell Biol.* 15:704-710.
- Williams, D.A., Lemischka, I.R., Nathan, D.G., and Mulligan, R.C. 1984. Introduction of a new genetic material into pluripotent haematopoietic stem cells of the mouse. *Nature* 310:476-480.
- Wong, B.Y., Chen, H., Chung, S.W., and Wong, P.M. 1994. High-efficiency identification of genes by functional analysis from a retroviral cDNA expression library. *J. Virol.* 68:5523-5531.

Contributed by Constance Cepko
Harvard Medical School
Boston, Massachusetts

Warren Pear
University of Pennsylvania
Philadelphia, Pennsylvania

INACTIVATION OF GENES IN MAMMALIAN CELLS

Human Somatic Cell Gene Targeting

Human somatic cell gene targeting provides a powerful tool to scientists studying gene function in cultured human cells. This technology allows scientists to knock out genes in human somatic cells in a fashion analogous to the creation of knockout mice. Human somatic cell gene targeting brings the power of genetics to the study of human genes in human cells by making it possible to compare cells or individuals that are genetically identical except for a single, well-defined mutation in an endogenous gene. A partial list of genes successfully targeted in human somatic cells is presented in Table 9.15.1.

Once this “molecular scalpel” has been employed, a researcher has at his or her disposal a set of cultured human cells that are genetically identical except for a specific, targeted change—the presence (in the parental cells) or absence (in the knockout cells) of a functional human gene. Since the cells are otherwise genetically the same, any biological or biochemical difference between the cells sheds light on the function of the targeted gene. This type of classical genetic analysis—the study of gene function by comparing cells or organisms that are genetically identical except for a single, well-defined genetic change in an endogenous gene—has been a mainstay of genetics in model systems for decades and has yielded innumerable critical discoveries. Until relatively recently, however, this capability has been restricted to model organisms and was not applicable to human cells.

The biological and biochemical characteristics of such isogenic sets of human cells can be analyzed in many different ways. For example, the growth properties of the cells can be studied *in vitro* to assess differences in their morphology and other growth characteristics (Shirasawa et al., 1993). The cells can be analyzed using flow cytometry to assess their cell-cycle profile during exponential growth or cell-cycle arrest, or after treatment with growth factors, small-molecule therapeutics, and lead compounds (Waldman et al.,

Table 9.15.1 Gene Targeting in Human Cells

Gene targeted	Reference
<i>Smad4</i>	Zhou et al., 1998
<i>p53</i>	Bunz et al., 1998
<i>14-3-3σ</i>	Chan et al., 1999
<i>DNMT1</i>	Rhee et al., 2000
<i>BAX</i>	Zhang et al., 2000a
<i>PPARδ</i>	Park et al., 2001
<i>Securin</i>	Jallepalli et al., 2001
<i>ORC2</i>	Dhar et al., 2001
<i>Ferredoxin reductase</i>	Hwang et al., 2001
<i>ATR</i>	Cortez et al., 2001
<i>DEC1</i>	Zawel et al., 2002
<i>DNMT3b</i>	Rhee et al., 2002
<i>β-Catenin</i>	Kim et al., 2002a; Chan et al., 2002; Sekine et al., 2002

Contributed by Todd Waldman, Carolyn Lee, Tagvor G. Nishanian, and Jung-Sik Kim

Current Protocols in Molecular Biology (2003) 9.15.1-9.15.20

Copyright © 2003 by John Wiley & Sons, Inc.

1995). The propensity of the cells to undergo apoptosis in vitro can be assessed using a variety of assays (Waldman et al., 1996). Microarrays can be employed to identify novel putative effectors of the targeted gene (Kim et al., 2002a).

The cells can also be studied in vivo. Since the cell lines commonly used for gene targeting are derived from human cancers, they will form tumors when injected subcutaneously into immunodeficient mice. By creating and comparing so-called “isogenic tumors,” it is possible to examine the role of the targeted gene in a wide variety of processes—e.g., one can study the effects of oncogenes and tumor-suppressor genes in tumor formation and on the sensitivity of tumors to anticancer agents (Shirasawa et al., 1993; Waldman et al., 1997).

This unit presents protocols for human somatic cell gene targeting. The Strategic Planning section presents a number of considerations that must be made when planning such an experiment. Basic Protocol 1 provides details for vector construction. Basic Protocol 2 describes the transfection procedures. Basic Protocol 3 is a method for excising the antibiotic-resistance gene from a heterozygous knockout to make it possible to target the remaining allele; the Alternate Protocol describes how to switch selectable markers for the same purpose. Basic Protocol 4 describes how to target the remaining allele, converting a heterozygous knockout cell line to a homozygous knockout cell line.

STRATEGIC PLANNING

Selection of a Parental Cell Line

To date, six different human cell lines have been successfully used for human somatic cell gene targeting (listed in Table 9.15.2). The list is short because suitable cell lines must possess each of three distinct characteristics. The first of these, which is very important, is that suitable cell lines should be diploid (or near-diploid), such that there are two (and only two) copies of the gene to be targeted. Second, the cells should transfect with high efficiency, since the promoterless enrichment built into human gene targeting vectors leads to the formation of relatively few colonies (since randomly integrated promoterless targeting vectors generally fail to express *neo^r* and therefore do not form drug-resistant colonies). Third, the cells should grow quickly and clonally, since one will need to grow individual clones in 96-well plates, preferably in 2 to 3 weeks. Of the cell lines listed in Table 9.15.1, HCT116 cells have proven to be the most robust and therefore have been used in the majority of human somatic cell gene targeting projects. As such, the authors of this unit tend to focus initial targeting efforts on HCT116 cells, since it is preferable to test new targeting vectors in a cell line that has been well established as suitable for human somatic cell gene targeting. Once the efficacy of targeting in HCT116 cells has been demonstrated, only then is the attempt made to target the gene in other cell lines as well.

Table 9.15.2 Human Cell Lines for Gene Targeting

Cell line	Reference
HCT116	Shirasawa et al., 1993
DLD1	Shirasawa et al., 1993
HT1080	Porter and Itzhaki, 1993
SW48	Kim et al., 2002b
HaCaT	Zawel et al., 2002
Human fibroblasts	Brown et al., 1997

In addition to meeting these three requirements, it is important for both experimental and technical reasons that the chosen cell line express the gene to be targeted. Experimentally, there is not much point in deleting a gene in somatic cells that is not expressed. Technically, homologous integration of a promoterless targeting vector into a transcriptionally silent locus will lead to the absence of *neo^r* expression and the failure to form a drug-resistant colony. Thus, the authors recommend first confirming that the gene of interest is expressed in candidate cell lines prior to initiating a targeting project (via immunoblotting, northern blotting, or RT-PCR).

Architecture of a Promoterless Targeting Vector

Promoterless targeting vectors require that the selectable marker gene be inserted in-frame into the gene being targeted. Homologous integration of such a targeting vector into the human genome then leads to the expression of a functionally active drug-resistance protein. There are several general factors to consider when planning to construct such a targeting vector.

One early concern was that targeting vectors for human somatic cell gene targeting would need to be composed of isogenic DNA—i.e., DNA derived from the actual cell line in which the gene would be targeted. These concerns were based on early gene-targeting studies performed in mouse embryonic stem cells, which indicated that the polymorphisms present in genomic DNAs derived from different strains of mice could affect the targeting frequency in a dramatically adverse fashion (Deng and Capecchi, 1992; van Deursen and Wieringa, 1992). If this were true for gene targeting in human cells, it would be difficult to target a gene in multiple cell lines using the same targeting vector. Furthermore, since humans (unlike mice) are not inbred, the targeting vector for each allele of a gene would need to be composed of different genomic DNA. As such, the need for isogenic DNA would substantially increase the technical complexity of human somatic cell gene targeting. Fortunately, experience has shown that targeting vectors composed of nonisogenic genomic DNA are able to create knockouts in a wide variety of genetically unrelated human cell lines (Sedivy et al., 1999).

In the authors' experience, the *neo^r* protein is frequently unable to tolerate the fusion of unrelated amino acids onto its amino terminus (C. Lee., J.S. Kim., and T. Waldman, unpub. observ.). This being the case, it is inadvisable to design a targeting strategy which leads to the formation of a *neo^r* fusion protein. There are two effective strategies that circumvent this limitation. First, one can design the targeting vector such that the initiating methionine of the gene being targeted is replaced precisely with the initiating methionine of *neo^r* gene. Second, one can replace internal exon(s) with an IRES-*neo^r* gene.

Finally, it is not possible to design a true promoterless targeting vector for deleting the first exon of a gene, since the left arm of such a targeting vector would contain promoter elements of the gene being targeted. If the initiating methionine of a gene is, in fact, in the first exon, the authors recommend designing the targeting vector to delete internal exon(s) using an IRES-*neo^r* strategy.

Building a Targeting Vector

Building promoterless targeting vectors is a nontrivial exercise in recombinant DNA technology. There are two factors that complicate construction of these vectors. First, there is virtually no sequence flexibility in the junction between the left arm of the targeting vector and the selectable marker cassette, making the use of standard restriction enzyme-based recombinant DNA strategies problematic. Second, it is frequently necessary to add several sequence features between the end of the selectable marker gene and

the right homology arm of the targeting vector, such as restriction sites and a PCR priming site.

The authors have found that the most straightforward way to circumvent these difficulties is to build the targeting vectors in two steps, exploiting the high efficiency of homologous recombination in *S. cerevisiae* (Storck et al., 1996; Oldenburg et al., 1997). First, a 5- to 10-kb fragment containing the genomic region of interest is cloned (generally from a bacterial artificial chromosome, or BAC; *UNIT 5.9*) into a yeast shuttle vector, which contains selectable markers and origins of replication for propagation in both *E. coli* and *S. cerevisiae*. Next, the selectable marker gene (generally either *neo^r* or IRES-*neo^r*) is PCR-amplified from a plasmid using long primers that add both the needed sequence features, along with 30 to 50 nucleotides of homology to the cloned genomic region for directing the desired recombination event. This PCR product is then cotransformed into *S. cerevisiae* together with linearized recombinant shuttle vector, and recombinants are identified and used as the final targeting vector. This strategy makes it possible to design any junction and to add desired sequence features all in one step, and does not rely on the serendipitous presence of needed restriction sites.

It should be noted that it is also theoretically possible to exploit homologous recombination in *E. coli* to accomplish the same goals (Zhang et al., 2000b). However, unlike in yeast, in *E. coli* it is generally necessary to *select* for homologous recombination between the vector and the PCR product. Therefore, the PCR-amplified *neo^r* cassette would need to contain a selectable marker for growth in bacteria as well.

Screening for Knockouts in Human Cells

The simplest approach for identifying gene-targeted cell lines is to conduct an initial PCR-based screen, then confirm the putative knockouts with a Southern blot. There are two general factors to consider when planning the details of such a screening strategy.

Historically, PCR screening for knockouts has relied on one primer in the selectable marker gene (generally *neo^r*), and another primer located in the genomic region just outside one of the homology arms of the targeting vector. Homologous integration of the targeting vector creates a template for PCR, whereas random integration does not. As such, diagnostic PCR products are produced only from gene-targeted clones. The fundamental difficulty with this classic PCR-based screening strategy is the lack of a positive control for PCR—i.e., by definition there exists no template for PCR unless a knockout is created. As such, it is virtually impossible to optimize the PCR reaction prior to screening for knockouts, and therefore difficult to definitively interpret a negative result. If none of the clones are PCR-positive, the question remains—were there truly no knockouts, or did the PCR not work?

Recently, a strategy has been devised to overcome this limitation, providing an internal positive control for PCR in each reaction (Chan et al., 1999; for a schematic of this approach, see Fig. 9.15.1). This approach is very similar to the classic approach, except that the PCR primer generally located in the *neo^r* gene is substituted by a PCR primer located in the genomic region being *deleted by the targeting vector*. Importantly, this PCR priming site is built into the targeting vector such that homologous integration of the targeting vector simultaneously *deletes the endogenous priming site, and reintroduces it in a slightly different location*. As such, clones in which the targeting vector has integrated randomly create a PCR product of a certain size, and clones in which the PCR product has integrated via homologous recombination create a PCR product of a different size. Therefore, it is no longer the presence or absence of a PCR product that determines a knockout; instead, a PCR product is produced in all the PCR reactions and its size

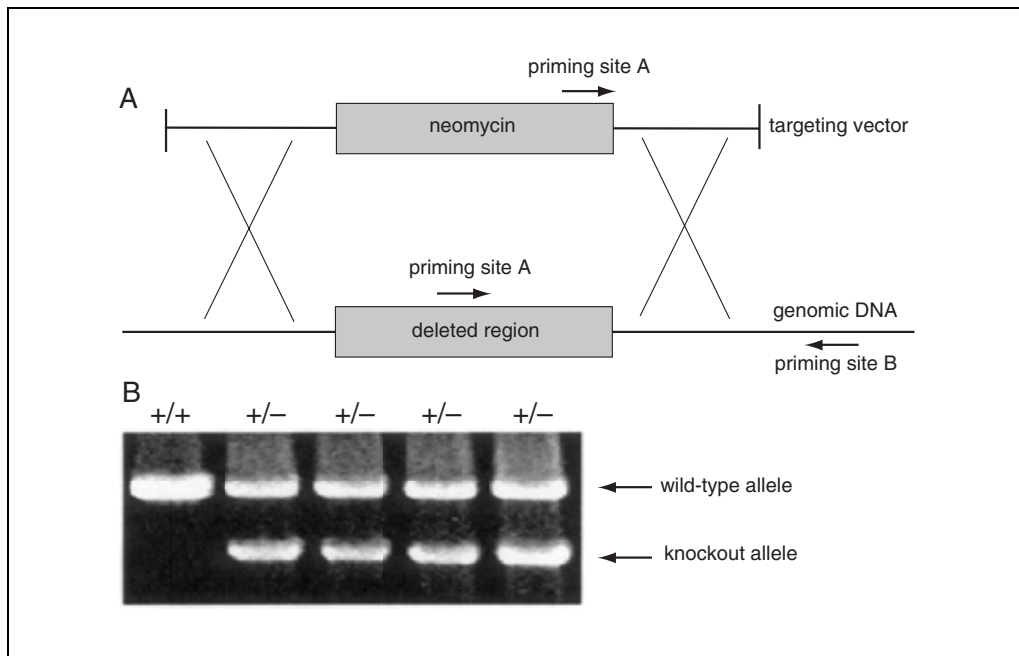


Figure 9.15.1 PCR screening for knockouts.

determines whether a knockout is present. This elegant modification of the classic PCR screening approach has significant implications for interpreting PCR screening data. The other practical implication is that the diagnostic PCR must be optimized prior to completion of the targeting vector, since the PCR priming site must be built into the targeting vector itself.

In addition to altering the size of a PCR product, homologous integration of the targeting vector invariably modifies the restriction map of the local genomic region. By using a Southern blot–based approach, it is possible to discriminate between wild-type alleles and targeted alleles. The details of choosing restriction sites and probes are described in Basic Protocol 1.

Targeting the Second Allele

One particularly challenging aspect of human somatic cell gene targeting is the creation of homozygous knockouts by disruption of the second, remaining allele. Since heterozygous knockout cells are already drug-resistant, it is necessary to modify the targeting vector to change the selectable marker gene. While swapping selectable markers is not technically demanding (especially if homologous recombination in *S. cerevisiae* was employed to create the targeting vector), experience has demonstrated that selectable markers other than *neo^r* (e.g., *hyg^r* and *puro^r*, among others) are suboptimal for somatic cell gene targeting. In particular, targeting vectors relying on these selectable markers are frequently either unable to generate drug-resistant colonies at all or require such low concentrations of drug that it is difficult to kill the background cells (Hanson and Sedivy, 1995; J.S. Kim, C. Lee, and T. Waldman, unpub. observ.). While the reasons for these difficulties are unknown, some have speculated that endogenous promoters are unable to generate a high enough level of gene expression to confer resistance to these specific drugs.

The most straightforward way to circumvent difficulties with selectable markers other than *neo^r* is to build a Cre-*loxP* recyclable *neo^r* targeting vector and use it to target each allele sequentially. In such a targeting vector, the *neo^r* gene is flanked by *loxP* sites. After

creation of heterozygous knockout cells, Cre is added to the cells (generally via infection with a recombinant Cre-containing adenovirus), and drug-sensitive derivatives are identified. These cells are then retransfected with the same targeting vector for deletion of the remaining allele.

The second inherent difficulty in targeting the second allele is that the frequency of homologous integration is exactly half that of the first allele. This is because half the second allele homologous integration events will disrupt the already targeted allele.

BUILDING A TARGETING VECTOR

The first step in the creation of a human promoterless targeting vector is to obtain human genomic DNA that will form the basis of the homology arms of the targeting vector. To do this, the authors generally identify BAC clone(s) that contain the needed genomic region (UNIT 5.9).

Once BACs have been identified and prepped, the next step is to subclone a smaller fragment (generally 5 to 10 kb) into a yeast shuttle vector. This subclone will be used as raw material for final assembly of the targeting vector via homologous recombination in *S. cerevisiae* (Storck et al., 1996; Oldenburg et al., 1997). Alternatively, it is possible to consecutively subclone each homology arm into a vector. However, subcloning a single larger piece that will ultimately form *both* homology arms saves a subcloning step.

The final step in construction of the targeting vector is to use homologous recombination in *S. cerevisiae* to replace critical exon(s) in the cloned genomic region with an in-frame *neo^r* or IRES-*neo^r* gene. The decision whether to use *neo^r* or IRES-*neo^r* is made according to the guidelines in step 6, below, and Critical Parameters and Troubleshooting. There are two feasible alternative approaches. First, it is possible to add the promoterless *neo^r* gene via conventional restriction enzyme–based cloning. However, the authors find that such an approach often requires complex multistep strategies because of the lack of convenient restriction sites. Second, it is possible to conduct similar homologous recombination reactions in strains of *E. coli* that overexpress recombinase proteins (Zhang et al., 2000b). In this case the disadvantage is that such an approach generally requires that the insert include an *E. coli* selectable marker gene.

Before beginning the homologous recombination, it is necessary to identify a unique restriction site in the cloned genomic region for linearization of the recombinant shuttle vector prior to transformation of *S. cerevisiae*. The recombination junctions will need to flank this restriction site.

Once such a restriction site is identified, the *neo^r* gene (or IRES-*neo^r* gene) is PCR-amplified with long primers designed to add needed sequence features to the ends (e.g., a priming site for PCR-based diagnosis of knockouts, restriction site or sites for Southern blot-based diagnosis of knockouts, and sequence homologies for recombination in *S. cerevisiae*). This PCR product is then cotransformed into *S. cerevisiae* with the linearized recombinant yeast shuttle vector. After plating the transformation mix on the appropriate selective medium, individual yeast colonies are screened via whole-cell PCR to identify those harboring the desired recombinant plasmid. PCR products derived from recombinant plasmids are then sequenced to exclude plasmids harboring mutations in the recombination junctions. These plasmids are then transferred into *E. coli* and prepped. Finally, the integrity of the finished targeting vector is confirmed by restriction analysis and sequencing of critical junctions.

Materials

Genomic DNA (UNIT 2.2) derived from the parental cell line (e.g., HCT116; ATCC #CCL 247)

BAC DNA purification kit: e.g., Nucleobond Plasmid Maxi kit (Clontech) or similar kit from Qiagen (also see UNIT 2.1B)

Yeast shuttle vectors (e.g., YEp24, pRS423, pRS424, pRS425, or pRS426, available from ATCC)

E. coli cells (DH10B)

pMC1neopolyA (Stratagene) or an IRES-*neo*^r plasmid (available from John Sedivy at Brown University upon request)

EXPAND High Fidelity PCR System (Roche)

QIAX Gel Extraction Kit (Qiagen) or equivalent

Appropriate selective yeast media (Qbiogene)

Yeast PCR kit (Qbiogene)

Yeast RPM kit (Qbiogene)

Facility for BAC library screening (e.g., Research Genetics)

30°C incubator

Additional reagents and equipment for primer design (UNIT 7.7), synthesis of oligonucleotides (UNIT 2.11), PCR (UNIT 15.1), gel purification of DNA fragments (UNIT 2.7), preparation and screening of a BAC library (<http://www.chori.org/bacpac/home.htm>), plating bacteria (UNIT 1.3), single-cell PCR (UNIT 25A.3), restriction digestion (UNIT 3.1), agarose gel electrophoresis (UNIT 2.5A), restriction mapping (UNITS 3.2, 3.3, & 7.7), finding sequence information on the Web (UNITS 19.2 & 19.3), dephosphorylation of DNA by calf intestinal phosphatase (UNIT 3.10), DNA subcloning and ligation (UNIT 3.16), introduction of plasmid DNA into *E. coli* (UNIT 1.8), colony hybridizations (UNIT 6.3), preparation of plasmid DNA (UNIT 1.7), DNA sequencing (Chapter 7), phenol extraction and ethanol precipitation of DNA (UNIT 2.1A), gel purification of DNA fragments (UNITS 2.5A, 2.7, or 2.8), introduction of DNA into *S. cerevisiae* (UNIT 13.7), growth of *S. cerevisiae* (UNIT 13.2), and DNA preparation from *S. cerevisiae* (UNIT 13.11)

Identify and isolate BAC clones

1. Design (UNIT 7.7) and synthesize (as described in UNIT 2.11; or order from synthesis facility) a PCR primer pair for amplification of the genomic region of interest from human genomic DNA derived from the parental cell line. Optimize and perform PCR amplification (UNIT 15.1; the PCR product should be 500 to 1000 nucleotides in length) to produce a probe for BAC library screening (step 2). Gel purify the DNA fragment obtained from PCR (UNIT 2.7).

The primer pair will be used both to create a probe for library screening and to confirm that the BACs identified harbor the correct insert. Ideally, the probe will be derived from intronic sequence, since probes derived entirely from exons can hybridize to pseudogenes.

2. Perform a hybridization-based BAC library screening using the probe for the region of interest created in step 1. For protocol details, see the laboratory Web page of Pieter De Jong at the Children's Hospital Oakland Research Institute (<http://www.chori.org/bacpac/home.htm>).

Since the physical map and sequence of the human genome is by now largely completed, it is often possible to search GenBank (UNIT 19.2) and identify the needed BAC, then order it from a BAC repository. However, if such information is unavailable, or if the BAC or BACs ordered from the repository do not contain the correct insert, it is necessary to screen a library to identify the needed BACs. In fact, the authors often screen a library even if a single correct BAC is available, since it is generally advantageous to have several different BAC clones with the needed insert, each with different boundaries and junctions.

Although the materials are available to screen BAC libraries in academic laboratories), the authors generally contract out the screening to Research Genetics, which performs this procedure routinely. Screening BAC libraries involves hybridization of radiolabeled probes to very large commercially available membranes that contain hundreds of thousands of gridded BAC clones. Performing the actual hybridizations is tricky because of the size of the membranes. Furthermore, identification of the correct clones on the huge filters takes practice.

If sending the probe to a commercial facility (e.g., Research Genetics), it generally takes 2 to 6 weeks to receive the final product, which will be stabs inoculated with bacteria harboring the desired BACs.

Alternatively, it is possible to skip this BAC step by generating the needed genomic DNA via PCR amplification of the needed homology arms from a genomic DNA template. However, the authors prefer to use BACs as the source of genomic DNA for building targeting vectors, since it can be difficult to create large homology arms using PCR. Also, PCR can introduce mutations into the homology arms and therefore reduce the targeting efficiency of the completed targeting vector.

3. Once the BACs are received, confirm that they contain the correct insert as follows.
 - a. Streak out bacteria from the stabs onto the appropriate selective medium to generate single colonies (UNIT 1.3).
 - b. Grid individual colonies (generally ten or so) to a grid plate (UNIT 1.3).
 - c. Using the PCR primer pair optimized in step 1 above, test each of the grid positions with whole-cell PCR (UNIT 25A.3) to identify bacteria harboring the desired insert.

Confirming the BAC clone is important since the bacterial stabs are created by directly inoculating bacteria stored as glycerol stocks in 384-well plates. It is possible (and in the authors' experience not uncommon) for the bacteria in a particular well to be a mixed population caused by spillage from adjacent wells, or the wrong clone entirely.

4. Prepare BAC DNA from positive clones using the Nucleobond Plasmid Maxi or similar DNA purification kit.

The authors generally use Nucleobond columns from Clontech, though kits available from other manufacturers (e.g., Qiagen; UNIT 2.1B) will undoubtedly also work well.

5. Digest the various BACs with several different restriction enzymes (UNIT 3.1) and compare the restriction digestion patterns by agarose gel electrophoresis (UNIT 2.5A).

This step confirms the large size of the purified DNAs. Furthermore, it is generally expected that different BACs share some bands but differ in others, reflecting the fact that they are derived from the same genomic region but have different boundaries.

Subclone genomic DNA from the BAC into the yeast shuttle vector

To choose a restriction fragment to subclone, one must first decide what exon(s) the completed vector will be designed to target. There are several factors to consider when choosing an exon to target. If possible, target the exon containing the initiating methionine, since it is then possible to build a promoterless targeting vector in which the initiating methionine of *neo^r* precisely replaces the initiating methionine of the targeted gene. However, avoid targeting the first exon, since the left arm of such a targeting vector would contain promoter elements of the gene itself, reducing the efficiency of the promoterless selection. If the initiating methionine is in the first exon, choose an internal exon and use an IRES-*neo^r* based strategy (Sedivy and Dutriaux, 1999). There are two main issues to consider when targeting an internal exon with an IRES-*neo^r*. First, it is desirable to target an early exon and therefore create a protein that is as dramatically altered as possible. Second, it is desirable to target an exon which, if it were skipped by the splicing machinery, would change the reading frame and result in a premature stop codon.

- Using published or Web-based human genomic sequence information (*UNITS 19.2 & 19.3*), make a sequence-based restriction map of the desired genomic region (see *UNITS 3.2, 3.3 & 7.7*). Identify several suitable restriction fragments for subcloning into the yeast shuttle vector.

When choosing potential restriction fragments, keep in mind that one arm of the final targeting vector will need to be short enough to PCR amplify across (i.e., less than ~2 kb) when identifying gene-targeted cell lines.

- Using the RepeatMasker Web site (<http://repeatmasker.genome.washington.edu>), analyze the sequence within and surrounding these restriction fragments to identify repeat elements such as *Alu* and *LINE-1*.

It is important to avoid designing targeting vectors composed largely of repeat elements, since the arms of such targeting vectors will not provide the unique homologies required to guide homologous recombination in human cells.

- To perform the actual subcloning, first digest both the BAC DNA and the yeast shuttle vector with the appropriate restriction enzyme or enzymes (see *UNIT 3.1* and documentation for shuttle vector).

*There are numerous yeast shuttle vectors available for use in this subcloning step. The authors generally use one of five different yeast shuttle vectors available from ATCC: *YEp24*, *pRS423*, *pRS424*, *pRS425*, and *pRS426* (Botstein et al., 1979; Christianson et al., 1992). *YEp24* is a *pBR322*-based, low-copy-number vector, whereas the *pRS* series are *pBluescript*-based, high-copy-number vectors. All work well for creating targeting vectors via homologous recombination; however, in the authors' experience it is slightly easier to clone larger fragments (>5 kb) into low-copy-number vectors such as *YEp24*.*

- Treat the linearized vector with calf intestinal alkaline phosphatase (CIP; *UNIT 3.10*) to limit recircularization of the vector during ligation.
- Ligate the linearized, CIP-treated vector to the digested BAC DNA fragments (*UNIT 3.16*). Transform or electroporate appropriate *E. coli* cells with the ligated plasmid vector (*UNIT 1.8*). After plating on the appropriate selective medium, grid ~500 colonies (*UNIT 1.3*). Also grid bacteria harboring various BACs to serve as positive and negative controls for the next step.
- Identify bacteria harboring the desired recombinant plasmid by performing colony hybridizations with a radiolabeled oligonucleotide probe (*UNIT 6.3*).

*It is often possible to use one of the PCR primers from in step 1, above, as the probe. Note that the positive signals produced by bacteria harboring high-copy-number recombinant plasmids (e.g., *pRS423-6*) will be much stronger than the signals produced by bacteria harboring low-copy-number recombinant plasmids (e.g., BACs, *YEp24*).*

- Streak out the bacteria (*UNIT 1.3*) in positive grid positions to purify individual colonies. Prepare plasmid DNA using the Nucleobond Plasmid kit or equivalent (also see *UNIT 1.7*).
- Confirm the integrity of the subclones by restriction analysis and sequencing of junctions (Chapter 7).

It is important to identify and avoid subclones with multiple, concatamerized inserts.

Identify priming sites and restriction sites for PCR and Southern blot-based identification of gene-targeted cell lines

It is important to identify these sequence features before completing the targeting vector, since the sequence comprising a PCR primer and possibly a restriction site will be built into the targeting vector during the final step of construction.

14. To identify suitable PCR primers, design (UNIT 7.7) and synthesize (as described in UNIT 2.11; or order from synthesis facility) three sense and three antisense primers that meet the following criteria:
 - a. By convention, one primer is located in the genomic DNA just adjacent to the shorter homology arm of the targeting vector.
 - b. The other primer is not located in the *neo^r* gene (or IRES-*neo^r* gene) itself, but is instead located in the genomic region deleted by the targeting vector.

*Importantly, this priming site will be built into the targeting vector itself at the junction of the short homology arm and the *neo^r* gene. Homologous integration of the targeting vector deletes the endogenous site and moves it, altering the size of the PCR product (generally shortening it).*

15. Test all nine permutations of these primers on genomic DNA derived from the parental cell line (e.g., HCT116) at several different annealing temperatures, as described in UNIT 15.1. Choose the most robust primer pair for eventual screening of putative gene-targeted clones.

As described above (see Strategic Planning), the authors use a strategy for PCR-based diagnosis of knockouts that is slightly different from the conventional strategy used to identify knockouts in mouse embryonic stem cells. This strategy is preferable to the conventional PCR-based screen, since it includes a positive internal control with each PCR reaction.

16. To identify candidate restriction sites for Southern blot–based diagnosis of knockouts, make a computer-generated restriction map of the genomic region using published or Web-based genomic sequences (UNIT 7.7). Identify suitable sites based on the following criteria:
 - a. The site should be *present* at a defined location in the genomic DNA at least 500 nucleotides outside one homology arm of the completed targeting vector (to allow for placement of a Southern blot probe; also see UNIT 10.8).
 - b. The site should be *absent* in that homology arm itself.
 - c. The site should be *present* at a defined location either in the other homology arm or in the genomic DNA adjacent to the other homology arm.
 - d. Furthermore, the site should either be present in the *neo^r* gene itself, or should be added to the targeting vector (generally at a junction of a homology arm and *neo^r*) during the final step in vector construction.

By following these guidelines, homologous integration of the targeting vector will alter the genomic restriction map in a way that should be easily measurable by Southern blot analysis (UNIT 10.8).

Replace critical exon(s) with a promoterless *neo^r* or IRES-*neo^r* gene via homologous recombination in *S. cerevisiae*

17. Using sequence analysis software (UNIT 7.7), identify candidate restriction sites for linearization of the recombinant yeast shuttle vector created in steps 8 to 13. Confirm location and uniqueness of candidate site(s) via routine restriction analysis (UNIT 3.1).
18. Once an appropriate site has been chosen, perform a larger-scale preparative digest using approximately 2.5 µg of the recombinant shuttle vector plasmid DNA from step 12. Phenol extract and ethanol precipitate the linearized plasmid (UNIT 2.1A).
19. Design (UNIT 7.7) and synthesize (as described in UNIT 2.11; or order from synthesis facility) PCR primers for amplification of the *neo^r* or IRES-*neo^r* gene with appropriate ends to add the PCR priming sites and restriction sites and to add 40 to 50 nucleotides of sequence homology to direct homologous recombination in *S. cerevisiae*.

The authors order primers PAGE-purified from Integrated DNA Technologies. If using the neo^r gene as the template, the authors generally use the pMC1neoPolyA plasmid as the template and use the following PCR priming sites: sense 5'-ATGGGATCGGCCATTGAA-CAA-3', antisense 5'-GTCGACGGATCCGAACAAACG-3'. Of note, the first three nucleotides of the sense primer comprise the initiating methionine of neo^r. If using IRES- neo^r (available from John Sedivy at Brown University) as the template, the authors use the following PCR priming sites: sense 5'-AACGTTACTGGCCGAAGCCGC-3', antisense 5'-TCCCAACTCATCCCGGCCTC-3'.

20. Perform high-fidelity PCR with the pMC1neopolyA plasmid or IRES-*neo^r* as template, to create the *neo^r* PCR product, using the EXPAND High Fidelity PCR System or equivalent (also see *UNIT 15.1*).
21. Gel purify the PCR product using the QIAX Gel Extraction Kit or equivalent (also *UNITS 2.5A, 2.7, or 2.8*).
22. Cotransform (*UNIT 13.7*) 150 ng of linearized recombinant shuttle vector (from step 18) together with an estimated 500 ng of gel-purified PCR product (from step 21) into *S. cerevisiae*. Include negative controls consisting of vector alone and PCR product alone.

*The authors generally use either commercially available competent yeast (strain MaV203; Invitrogen) or prepare competent yeast from strain INVSc1 (Invitrogen) using the Frozen-EZ Yeast Transformation II kit (Zymo Research). The choice of yeast strain depends on which auxotrophic marker is needed. Specific protocols for transformation are available in each kit. Also see *UNIT 13.7* for additional information relating to yeast transformation.*

23. Plate the transformation mixes on the appropriate selective medium (available from Qbiogene), and incubate at 30°C for 2 to 4 days.

*General considerations for the growth of S. cerevisiae are presented in *UNIT 13.2*. Expect to see ~5-fold more colonies in the experimental reactions than in vector alone. There should be no colonies in insert alone.*

24. Grid approximately 20 colonies from the vector + insert transformation mix (see *UNIT 13.2*). Also grid some vector-only colonies to serve as negative controls for the subsequent PCR reaction.
25. Design and synthesize (as described in *UNIT 2.11*; or order from synthesis facility) a set of PCR primers that flank the recombination event.

Such a primer pair will create PCR products of different sizes in recombined and nonrecombined plasmids.

26. Test yeast from individual grid positions by PCR to identify those in which a recombination event has occurred.

The authors generally use the yeast itself as a PCR template after pretreatment with zymolyase to destroy the cell wall. The authors use the Yeast PCR Kit (Qbiogene) for this purpose.

27. Once yeast-harboring recombinant plasmids have been identified, sequence the ends of the PCR products to demonstrate that the recombination junctions are intact and free of mutations. Design (*UNIT 7.7*) and synthesize (as described in *UNIT 2.11*; or order from synthesis facility) sequencing primers for the purpose of sequencing the junctions from “inside out,” making it possible to generate sequence from the very end of the PCR products. Sequence the PCR products (see Chapter 7).

Transfer completed targeting vectors into E. coli for large-scale DNA preparation

28. Prepare a miniprep of the recombinant yeast DNA using the Yeast RPM kit.
29. Electroporate the miniprep DNA into *E. coli* as described in *UNIT 1.8*. Plate on the appropriate selective medium (*UNIT 1.3*). Streak out individual colonies to purify.
30. Perform large-scale plasmid preps as described in *UNIT 1.7*.

The authors generally use Nucleobond kits (Clontech).

31. Confirm the integrity of the resultant plasmid via restriction analysis and resequencing of the critical junctions (see Chapter 7) using the sequencing primers used in step 27 above.
32. Linearize the vector for transfection (Support Protocol).

SUPPORT PROTOCOL

LINEARIZATION OF THE COMPLETED TARGETING VECTOR PRIOR TO TRANSFECTION INTO HUMAN CELLS

The completed targeting vector must be linearized prior to transfection, since homologous recombination is more efficient with linear DNA. As such, it is necessary to identify a suitable restriction enzyme that cuts only in the vector backbone. There are generally sites for rare-cutting enzymes (e.g., *SalI* or *NotI*) in the polylinker of the yeast shuttle vector that are useful for this purpose

Materials

Completed targeting vectors (Basic Protocol 1)
TE buffer (*APPENDIX 2*)

Additional reagents and equipment for restriction digestion (*UNIT 3.1*), agarose gel electrophoresis (*UNIT 2.5A*), and phenol extraction and ethanol precipitation of DNA (*UNIT 2.1A*)

1. Identify an appropriate restriction enzyme (*UNIT 3.1*) for linearization of the completed targeting vector in the vector backbone. Confirm that the enzyme cuts the targeting vector singly and in the predicted location by performing small-scale restriction digests and analyzing an aliquot by agarose gel electrophoresis (*UNIT 2.5A*).
2. Perform a preparative digest of 50 to 100 μg of the targeting vector with the chosen enzyme (*UNIT 3.1*). Phenol extract and ethanol precipitate the digest (*UNIT 2.1A*), then resuspend in TE buffer to a final concentration of 0.5 $\mu\text{g}/\mu\text{l}$.
3. Confirm the efficiency of digestion by analyzing an aliquot of the digest by agarose gel electrophoresis (*UNIT 2.5A*).

The authors generally prepare linearized DNA from three identical targeting vectors derived from different yeast colonies, in the unlikely event that one or several of the vectors has a PCR-generated inactivating mutation in the neo' (or IRES-neo') gene.

BASIC PROTOCOL 2

TRANSFECTION OF THE LINEARIZED TARGETING VECTOR INTO CULTURED HUMAN CELLS AND IDENTIFICATION OF HETEROZYGOUS KNOCKOUT CELLS

The authors generally use HCT116 cells for the initial gene-targeting experiments, since these cells divide rapidly, grow well as individual clones in 96-well plates, and are generally well established as suitable for human somatic cell gene targeting. Once the targeting vector has been shown to work efficiently in HCT116 cells, attempts can be made to target the gene in other cell lines as well.

Two sequential transfections are performed in this step. Initially, the cells are transfected, and colonies are selected directly in 25-cm² tissue culture flasks and stained with crystal violet. This preliminary experiment demonstrates whether the new targeting vector(s) can create G418-resistant colonies, and identifies any targeting vectors that may harbor inactivating mutations in their *neo*^r (or IRES-*neo*^r) gene. Also, based on the number of colonies produced, it is possible to estimate the dilutions needed for creation of single colonies via limiting dilution in 96-well plates.

Next, the HCT116 cells are retransfected and set up in a large number (50 to 100) of 96-well plates for obtaining individual drug-resistant clones by limiting dilution. Once single colonies have formed (first identifiable ~10 days post-transfection), they are expanded (at ~20 days post-transfection) for cryopreservation and preparation of genomic DNA. The authors generally work up 200 to 500 colonies in this step.

Finally, the genomic DNAs are tested by PCR to identify heterozygous knockout cell lines. Knockouts are then thawed, and are grown up to create more frozen vials and to reprepare genomic DNA. The genotype of the cell lines is then reconfirmed by PCR and Southern blotting.

Materials

- Complete McCoy's 5A medium (see recipe) with and without 0.6 µg/ml G418 (geneticin, Life Technologies; see *UNIT 9.5*)
- HCT116 cells (ATCC #CCL 247)
- Lipofectamine reagent (Invitrogen; see *UNIT 9.4*)
- Linearized targeting vector (Support Protocol)
- Plasmid unrelated to targeting vector and lacking *neo*^r gene (negative control for colony formation)
- PMC1neopolyA plasmid (Stratagene)
- CMV β-gal plasmid (Clontech)
- Hanks' Balanced Salt Solution (HBSS, Life Technologies; also see *APPENDIX 2*)
- 0.05% (w/v) trypsin/0.53 mM tetrasodium EDTA (Life Technologies)
- 25-cm² and 75-cm² tissue culture flasks
- 8-channel pipettor (e.g., Easy Step with 8-port manifold from Continental Laboratory Products) with 50-ml reservoir (Brinkmann 22-49-614-0)
- 96-well tissue culture plates
- 24-well tissue culture plates
- 15-ml conical centrifuge tubes
- Additional reagents and equipment for culture of mammalian cells (*APPENDIX 3F*), transfection of mammalian cells (*UNIT 9.4*), Xgal staining of cultured cells (*UNIT 9.10*), crystal violet staining of cultured cells (as for plaque assay; see *UNIT 16.16*, Support Protocol), preparation of genomic DNA (*UNIT 2.2*), PCR (*UNIT 15.1*), and Southern blotting (*UNIT 2.9A*)

NOTE: All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

Determine if target vector is capable of creating *neo*^r clones

1. Grow HCT116 cells (see *APPENDIX 3F* for culture techniques) to 60% to 70% confluence in complete McCoy's 5A medium (without selective agent). Transfect HCT116 cells with linearized targeting vectors using Lipofectamine reagent as described in

UNIT 9.4 and in the manufacturer's instructions, using 3 µg DNA per 25-cm² tissue culture flask of 60% to 70% confluent cells. Also transfect:

- a. A plasmid harboring a *neo^r* gene driven by a heterologous promoter (e.g., pMC1neopolyA) as a positive control for colony formation.
- b. An unrelated plasmid lacking a *neo^r* gene as a negative control for colony formation.
- c. CMV β-Gal plasmid as a control for transfection efficiency.

Do not add selective medium to the cells yet.

The authors recommend transfecting ~10 to 20 25-cm² flasks in this step.

2. The day after transfection, stain the CMV β-Gal transfected cells with Xgal (UNIT 9.10; expect to see 20% to 40% blue cells). Change the medium on the remaining flasks to complete McCoy's 5A medium containing 0.6 µg/ml G418.

The authors generally use the β-Gal Staining Kit (Roche).

3. Replace the medium with fresh complete McCoy's 5A medium containing 0.6 µg/ml G418 every 3 days (colonies should begin to appear at approximately day 10). Stain the colonies with crystal violet (see UNIT 16.16, Support Protocol) on approximately day 15.

*In the authors' experience, high-efficiency promoterless targeting vectors generally create between 30 and 200 colonies per transfected 25-cm² flask of HCT116 cells, whereas vectors harboring *neo^r* genes driven by heterologous promoters generally produce thousands of colonies per transfected 25-cm² flask.*

Perform selection on *neo^r* clones

4. Once the pilot experiment described in steps 1 to 3 above demonstrates that targeting vector is capable of creating *neo^r* clones, repeat to generate hundreds of individual *neo^r* clones. To do this, repeat the transfection strategy described above, including the positive and negative controls. At 18 to 24 hr after transfection, stain the CMV β-Gal-transfected control cells and feed the other control cells with complete McCoy's 5A medium containing 0.6 µg/ml G418.

The remainder of the flasks will be trypsinized and transferred to 96-well plates, as described is step 5.

5. Wash the cell monolayers in the 25-cm² flasks with 5 ml HBSS, then add 1 ml of 0.05% trypsin/0.53 mM EDTA and swirl gently to detach the cells. In a sterile, disposable beaker, mix the trypsinized cells with complete McCoy's 5A medium containing 0.6 µg/ml G418.

The volumes to be used depend on the dilution and the number of 96-well plates to be created from the cell suspension (25 ml of medium per 96-well plate). Initially, try three different "dilutions" of cells to increase the likelihood of obtaining wells with single colonies: 1, 1/5, and 1/20th of a 25-cm² flask in a 96-well plate.

6. Place an aliquot of the cell suspension in each well of the appropriate number of 96-well plates using an 8-channel pipettor with a sterile 50-ml reservoir. Stack plates, cover with plastic wrap to prevent evaporation, and begin incubating.
7. Identify wells containing single colonies as early as possible (after 10 days to 2 weeks).

As colonies become larger, cells tend to detach and form "satellite" colonies. Also, over time, two initially distinct colonies may merge together and appear to be a single colony.

8. When colonies cover at least 50% of the surface area of the well, transfer to 24-well plates. To do this, wash the well once with 250 μ l HBSS, then add 100 μ l of 0.05% trypsin/0.53 mM EDTA. Transfer each colony to a separate well of a 24-well plate containing 1 ml of complete McCoy's 5A medium containing 0.6 μ g/ml G418.
9. As the wells become confluent, expand the cultures to 25-cm² flasks. To do this, wash the well once with 1 ml HBSS, then add 250 μ l 0.05% trypsin/0.53 mM EDTA. Transfer the trypsinized cells to a 25-cm² flask containing 5 ml of complete McCoy's 5A medium containing 0.6 μ g/ml G418.

Harvest cells and identify knockouts

10. As the 25-cm² flasks become confluent, harvest the cells: use 1/3 to prepare genomic DNA; cryopreserve the remaining 2/3. To harvest:
 - a. Wash the monolayer once with 2 ml HBSS, then add 1 ml of 0.05% trypsin/0.53 mM EDTA.
 - b. Transfer 1/3 and 2/3 of the cells to separate 15-ml conical tubes, each containing 5 ml of complete McCoy's 5A medium (without selective agent).
 - c. Centrifuge 7 min at 250 \times g, room temperature, to create a cell pellet.
11. Prepare genomic DNA from the tube containing 1/3 of the cells (*UNIT 2.2*) and cryopreserve the cells in the tube containing 2/3 of the cells (*APPENDIX 3F*).
12. Test genomic DNAs by PCR as described in Basic Protocol 1 and *UNIT 15.1* to identify knockouts.
13. Thaw PCR-positive clones into a 75-cm² tissue culture flask containing 20 ml of complete McCoy's 5A medium with 0.6 μ g/ml G418. Expand into a 225-cm² flask for cryopreservation of nine cryovials (*APPENDIX 3F*) and reparation of genomic DNA (*UNIT 2.2*).
14. Retest by PCR (*UNIT 15.1*) and confirm via Southern blotting (*UNIT 2.9A*).

RESTORATION OF G418 SENSITIVITY VIA Cre-MEDIATED EXCISION OF THE *neo*^r GENE

In order to recycle the targeting vector to disrupt the remaining allele, it is necessary to excise the *neo*^r gene (or IRES-*neo*^r gene) from heterozygous knockouts via expression of Cre recombinase. Generally, a recombinant Cre-expressing adenovirus is used for this purpose.

Materials

- High-titer Cre-expressing adenovirus (e.g., Microbix Biosystems)
- Heterozygous knockout cell lines (see Basic Protocol 2) growing at 50% confluence in 25-cm² tissue culture flasks
- Complete McCoy's 5A medium (see recipe) with and without 0.6 μ g/ml G418 (geneticin, Life Technologies; see *UNIT 9.5*)
- Hanks' Balanced Salt Solution (HBSS, Life Technologies; also see *APPENDIX 2*)
- 0.05% (w/v) trypsin/0.53 mM tetrasodium EDTA (Life Technologies)
- 25-cm² tissue culture flasks
- 96-well tissue culture plates
- 24-well tissue culture plates
- Additional reagents and equipment for culture of mammalian cells (including counting and cryopreservation; *APPENDIX 3F*), plating cells at limiting dilution (as for hybridoma cells; *UNIT 11.8*), and crystal violet staining of cultured cells (as for plaque assay; see *UNIT 16.16*, Support Protocol)

BASIC PROTOCOL 3

Introduction of DNA into Mammalian Cells

9.15.15

NOTE: All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

1. Add 1 µl recombinant adenovirus to heterozygous knockout cells at 50% confluence in a 25-cm² tissue culture flasks containing 5 ml complete McCoy's 5A medium without selective agent (G418). Incubate 24 hr.
2. Wash cells twice with HBSS, then trypsinize (see Basic Protocol 2) and count cells (*APPENDIX 3F*).
3. Plate out at limiting dilution (*UNIT 11.8*) in 96-well plates in complete McCoy's 5A medium without selective agent, to obtain single colonies. To be sure to obtain single colonies, set up a variety of 96-well plates, e.g., with 1 cell/well, 5 cells/well, and 10 cells/well.
4. Identify single colonies after 10 days, and trypsinize at ~2 weeks as described in Basic Protocol 2. Expand each clone to a well in a 24-well plate, and then to two 25-cm² tissue culture flasks.

One of the 25-cm² flasks will be used to test the G418 sensitivity, and the other will be used for cryopreservation.
5. Add complete McCoy's 5A medium containing 0.6 µg/ml G418 to one of the 25-cm² flasks when it reaches 50% confluence. Continue incubating for ~10 to 14 days, feeding every 3 days. Cryopreserve the cells in the other 25-cm² flask when it reaches confluence, as described in *APPENDIX 3F*.
6. After ~10 to 14 days, stain the 25-cm² flask containing the cells exposed to selective agent with crystal violet as described in *UNIT 16.16*. Identify those cell lines that are now completely G418-sensitive, and use them to disrupt the remaining allele (see Basic Protocol 4).

ALTERNATE PROTOCOL

SWAPPING SELECTABLE MARKERS IN THE TARGETING VECTOR

An alternative to recycling the *neo^r* targeting vector (see Basic Protocol 3) is to modify the targeting vector to switch selectable markers, replacing *neo^r* with *hyg^r*. It should be noted that the authors have been consistently unable to generate drug-resistant colonies when using *puro^r*, *IRES-hyg^r*, or *IRES-puro^r* (C. Lee, J. S. Kim, T. Waldman, unpub. observ.).

To do this, repeat Basic Protocol 1 with PCR amplification of a *hyg^r* gene. Perform Basic Protocol 2 as written, except substitute 0.1 µg/ml hygromycin for G418.

BASIC PROTOCOL 4

TARGETING OF THE REMAINING ALLELE

The principles and protocols needed for targeting the remaining allele are virtually identical to those already described for targeting the first allele. If recycling the *neo^r* targeting vector, one should repeat Basic Protocol 2, using G418-sensitive heterozygous knockout cell lines (see Basic Protocol 3) as the "parental" cells. If swapping selectable markers in the targeting vector, use *neo^r* heterozygous knockout cell lines as the "parental" cells. Of the homologous integration events obtained, approximately half should represent homozygous knockouts, and the other half should have targeted the already targeted allele. After confirmation of the homozygous knockout with PCR (*UNIT 15.1*) and Southern blotting (*UNIT 2.5A*), the authors generally perform an immunoblot (*UNIT 10.8*) to confirm the absence of protein.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Complete McCoy's 5A medium

McCoy's 5A medium (Life Technologies cat. no. 16600-082)

10% fetal bovine serum (FBS, Life Technologies; also see APPENDIX 3F)

1× penicillin/streptomycin (Life Technologies cat. no. 15140-122)

Store up to 4 months at 4°C

COMMENTARY

Background Information

History of human somatic cell gene targeting

The first descriptions of successful gene targeting in cultured human cells appeared in 1991, when two groups reported heterozygous targeting of different genes. Heterozygous targeting of the *6-16* gene, an interferon-inducible transcript, was reported by Itzhaki and Porter (1991). The frequency of homologous integration was extremely low (estimated at 1 per 1550 colonies screened), leading these authors to employ an unusual screen involving integration of a human growth hormone gene (*hGH*) and the screening of conditioned media from putative clones via hGH ELISA. That same year, heterozygous targeting of the *HPRT* gene was reported (Zheng et al., 1991). That publication was a proof-of-principle study in which the authors were able to directly select for heterozygous knockout cell lines. In 1993, the first homozygous (or complete) knockout of a gene in human cells was reported (Porter and Itzhaki, 1993). Using a promoterless targeting vector, Porter and Itzhaki were able to disrupt the remaining allele of *6-16* in the heterozygous gene-targeted cells described in their 1991 paper. These three early studies demonstrated the feasibility of gene targeting in human cells; they also demonstrated that it was a very labor-intensive process.

In 1993, a landmark paper in the history of human somatic cell gene targeting was published, reporting the heterozygous gene targeting of the K-ras oncogene in two human colon cancer cell lines (Shirasawa et al., 1993). Although this was not the first manuscript to describe human somatic cell gene targeting, it pointed out the potential utility of somatic cell gene targeting for studying human cancer. In particular, these authors created derivatives of both HCT116 and DLD1 cells in which either the wild-type or the oncogenic allele of K-ras

had been deleted. They went on to show that the K-ras oncogene was absolutely required for the tumorigenic properties of both cell lines. They reported an extremely high frequency of knockouts (8/93 in HCT116; 7/24 in DLD1), attributing it to the use of a promoterless *neo^r* gene. Furthermore, they introduced the use of HCT116 and DLD1 cells for gene targeting.

In 1995, a careful study was published that examined the merits of promoterless versus conventional targeting vectors for somatic cell gene targeting (Hanson and Sedivy, 1995). Their conclusion bolstered the work done by Shirasawa et al. (1993) showing that promoterless targeting vectors were virtually essential to achieve high frequencies of homologous recombination. They also reported that negative selectable markers commonly used for gene targeting in embryonic stem cells were of limited utility for somatic cell gene targeting. Furthermore, they pointed out that while the *neo^r* gene was extremely useful for gene targeting, the *hyg^r* gene seemed to work poorly. Subsequent studies, mostly unpublished, confirmed these seminal observations.

Also in 1995, homozygous deletion of the p21 cdk inhibitor in human HCT116 cells was reported (Waldman et al., 1995). This study brought together several of the most recent technical advances, such as the use of promoterless targeting vectors, the lack of a need for negative selectable markers, and the value of HCT116 cells. Importantly, this study demonstrated the feasibility and utility of human somatic cell gene targeting for the study of tumor suppressor gene pathways.

Other important incremental technical gains have been made as well. Advances have been made that simplify the construction of human promoterless targeting vectors, such as the use of homologous recombination in *Saccharomyces cerevisiae* (see UNIT 13.10). An important modification to the standard PCR screen for knockouts has made identification of gene tar-

ged cell lines simpler and more robust (see *UNIT 13.10*). Cre-*loxP* recombination (see *UNITS 9.5 & 13.3*) has been successfully employed to overcome difficulties associated with creation of homozygous knockouts by making it possible to recycle the same targeting vector for both alleles of a gene.

Today, through the implementation of these advances, human somatic cell gene targeting has emerged as an important tool for the study of gene function in human cells.

Critical Parameters and Troubleshooting

The frequency of homologous recombination in human cells is 3 to 5 orders of magnitude lower than the frequency of nonhomologous recombination. As such, the vast majority of drug-resistant clones formed after transfection of a conventional targeting vector represent random integration events. Identification of the tiny fraction of gene-targeted clones can be extremely challenging. It should be noted that the frequency of homologous integration in human cells is thought to be significantly lower than the frequency of homologous integration in murine embryonic stem (ES) cells. This fact, coupled with the fact that creation of homozygous knockouts in human cells requires sequential targeting of both alleles (unlike in mice, in which heterozygous knockouts are mated to homozygosity), renders promoter-containing targeting vectors unsuitable for human somatic cell gene targeting. Instead, promoterless targeting vectors are required. In such targeting vectors, the selectable marker gene (generally *neo^r*) lacks its own promoter and is instead fused in-frame with the gene being targeted.

The most important critical parameter is the efficiency with which the targeting vector creates gene-targeted cell lines. Vectors vary in their efficiency from <0.1% to 40%. Unfortunately, it is not known why some targeting vectors work extremely well and others do not. In general, the authors recommend building targeting vectors in which the bigger homology arm is as long as possible, and which lack repeat elements in the homology arms. Unfortunately, it is difficult to troubleshoot a vector that either does not create knockouts (yet creates G418-resistant colonies) or does so with very low efficiency. In general, the authors attempt to create another targeting vector based on different homology arms that target a different exon. Studies have demonstrated that different targeting vectors for the same gene can work with

dramatically different efficiencies (Park et al., 2001).

Another critical parameter that is often difficult to predict ahead of time is whether the gene being targeted is an essential gene. One often does not learn this until after the homozygous targeting step, when each of the homologous integration events has disrupted the already targeted allele (and none has disrupted the remaining wild-type allele). For an example of this, see Kim et al. (2002b). One way to solve this problem is to modify the targeting vector to create a conditional knockout. Though creation of conditional knockouts is beyond the scope of this protocol, it is described in Cortez et al. (2001).

Anticipated Results

Upon completion of this protocol, one should have generated an isogenic set of human cells that differs only in the presence or absence of the targeted gene.

Time Considerations

Creation of gene-targeted human cell lines is a technically demanding and time-consuming process. It can take anywhere from 3 to 9 months to create a new targeting vector, and another 6 to 12 months to create homozygous knockout cell lines.

Literature Cited

- Botstein, D., Falco, S.C., Stewart, S.E., Brennan, M., Scherer, S., Stinchcomb, D.T., Struhl, K., and Davis, R.W. 1979. Sterile host yeasts (SHY): A eukaryotic system of biological containment for recombinant DNA experiments. *Gene* 8:17-24.
- Brown, J.P., Wei, W., and Sedivy, J.M. 1997. Bypass of senescence after disruption of p21^{CIP1/WAF1} gene in normal diploid human fibroblasts. *Science* 277:831-834.
- Bunz, F., Dutriaux, A., Lengauer, C., Waldman, T., Zhou, S., Brown, J.P., Sedivy, J.M., Kinzler, K.W., and Vogelstein, B. 1998. Requirement for p53 and p21 to sustain G2 arrest after DNA damage. *Science* 282:1497-1501.
- Chan, T.A., Hermeking, H., Lengauer, C., Kinzler, K.W., and Vogelstein, B. 1999. 14-3-3 σ is required to prevent mitotic catastrophe after DNA damage. *Nature* 401:616-620.
- Chan, T.A., Wang, Z., Dang, L.H., Vogelstein, B., and Kinzler, K.W. 2002. Targeted inactivation of CTNNB1 reveals unexpected effects of β -catenin mutation. *Proc. Natl. Acad. Sci. U.S.A.* 99:8265-8270.
- Christianson, T.W., Sikorski, R.S., Dante, M., Shero, J.H., and Hieter, P. 1992. Multifunctional

- yeast high-copy-number shuttle vectors. *Gene* 110:119-122.
- Cortez, D., Guntuku, S., Qin, J., and Elledge, S.J. 2001. ATR and ATRIP: Partners in checkpoint signaling. *Science* 294:1713-1716.
- Deng, C. and Capecchi, M.R. 1992. Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol. Cell Biol.* 12:3365-3371.
- Dhar, S.K., Yoshida, K., Machida, Y., Khaira, P., Chaudhuri, B., Wohlschlegel, J.A., Leffak, M., Yates, J., and Dutta, A. 2001. Replication from oriP of Epstein-Barr virus requires human ORC and is inhibited by geminin. *Cell* 106:287-296.
- Hanson, K.D. and Sedivy, J.M. 1995. Analysis of biological selections for high-efficiency gene targeting. *Mol. Cell. Biol.* 15:45-51.
- Hwang, P.M., Bunz, F., Yu, J., Rago, C., Chan, T.A., Murphy, M.P., Kelso, G.F., Smith, R.A., Kinzler, K.W., and Vogelstein, B. 2001. Ferredoxin reductase affects p53-dependent, 5-fluorouracil-induced apoptosis in colorectal cancer cells. *Nat. Med.* 7:1111-1117.
- Itzhaki, J.E. and Porter, A.C. 1991. Targeted disruption of a human interferon-inducible gene detected by secretion of human growth hormone. *Nucleic Acids Res.* 19:3835-3842.
- Jallepalli, P.V., Waizenegger, I.C., Bunz, F., Langer, S., Speicher, M.R., Peters, J.M., Kinzler, K.W., Vogelstein, B., and Lengauer, C. 2001. Securin is required for chromosomal stability in human cells. *Cell* 105:445-457.
- Kim, J.S., Crooks, H., Dracheva, T., Nishanian, T.G., Singh, B., Jen, J., and Waldman, T. 2002a. Oncogenic β -catenin is required for bone morphogenetic protein 4 expression in human cancer cells. *Cancer Res.* 62:2744-2748.
- Kim, J.S., Crooks, H., Foxworth, A., and Waldman, T. 2002b. Proof-of-principle: Oncogenic β -catenin is a valid molecular target for the development of pharmacological inhibitors. *Mol. Cancer Ther.* 1:1355-1359.
- Oldenburg, K.R., Vo, K.T., Michaelis, S., and Paddon, C. 1997. Recombination-mediated PCR-directed plasmid construction in vivo in yeast. *Nucleic Acids Res.* 25:451-452.
- Park, B.H., Vogelstein, B., and Kinzler, K.W. 2001. Genetic disruption of PPAR δ decreases the tumorigenicity of human colon cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* 98:2598-2603.
- Porter, A.C. and Itzhaki, J.E. 1993. Gene targeting in human somatic cells: Complete inactivation of an interferon-inducible gene. *Eur. J. Biochem.* 218:273-281.
- Rhee, I., Jair, K.W., Yen, R.W., Lengauer, C., Herman, J.G., Kinzler, K.W., Vogelstein, B., Baylin, S.B., and Schuebel, K.E. 2000. CpG methylation is maintained in human cancer cells lacking DNMT1. *Nature* 404:1003-1007.
- Rhee, I., Bachman, K.E., Park, B.H., Jair, K.W., Yen, R.W., Schuebel, K.E., Cui, H., Feinberg, A.P., Lengauer, C., Kinzler, K.W., Baylin, S.B., and Vogelstein, B. 2002. DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature* 416:552-556.
- Sedivy J.M. and Dutriaux, A. 1999. Gene targeting and somatic cell genetics—a rebirth or a coming of age? *Trends Genet.* 15:88-90.
- Sedivy, J.M., Vogelstein, B., Liber, H., Hendrickson, E., and Rosmarin, A. 1999. Gene targeting in human cells without isogenic DNA. *Science* 283:9 [Technical Comments].
- Sekine, S., Shibata, T., Sakamoto, M., and Hirohashi, S. 2002. Targeted disruption of the mutant β -catenin gene in colon cancer cell line HCT116: Preservation of its malignant phenotype. *Oncogene* 21:5906-5911.
- Shirasawa, S., Furuse, M., Yokoyama, N., and Sasazuki, T. 1993. Altered growth of human colon cancer cell lines disrupted at activated Ki-ras. *Science* 260:85-88.
- Storck, T., Kruth, U., Kolhekar, R., Sprengel, R., and Seeburg, P.H. 1996. Rapid construction in yeast of complex targeting vectors for gene manipulation in the mouse. *Nucleic Acids Res.* 24:4594-4596.
- van Deursen, J. and Wieringa, B. 1992. Targeting of the creatine kinase M gene in embryonic stem cells using isogenic and nonisogenic vectors. *Nucleic Acids Res.* 20:3815-3820.
- Waldman, T., Kinzler, K.W., and Vogelstein, B. 1995. p21 is necessary for the p53-mediated G1 arrest in human cancer cells. *Cancer Res.* 55:5187-5190.
- Waldman, T., Lengauer, C., Kinzler, K.W., and Vogelstein, B. 1996. Uncoupling of S phase and mitosis induced by anticancer agents in cells lacking p21. *Nature* 381:713-716.
- Waldman, T., Zhang, Y., Dillehay, L., Yu, J., Kinzler, K., Vogelstein B., and Williams, J. 1997. Cell-cycle arrest versus cell death in cancer therapy. *Nat. Med.* 3:1034-1036.
- Zawel, L., Yu, J., Torrance, C.J., Markowitz, S., Kinzler, K.W., Vogelstein, B., and Zhou, S. 2002. DEC1 is a downstream target of TGF- β with sequence-specific transcriptional repressor activities. *Proc. Natl. Acad. Sci. U.S.A.* 99:2848-53.
- Zhang, L., Yu, J., Park, B.H., Kinzler, K.W., and Vogelstein, B. 2000a. Role of BAX in the apoptotic response to anticancer agents. *Science* 290:989-992.
- Zhang, Y., Muyrers, J.P., Testa, G., and Stewart, A.F. 2000b. DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.* 18:1314-1317.
- Zheng, H., Hastay, P., Brenneman, M.A., Grompe, M., Gibbs, R.A., Wilson, J.H., and Bradley, A. 1991. Fidelity of targeted recombination in human fibroblasts and murine embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 88:8067-8071.

Zhou, S., Buckhaults, P., Zawel, L., Bunz, F., Rig-
gins, G., Dai, J.L., Kern, S.E., Kinzler, K.W., and
Vogelstein, B. 1998. Targeted deletion of Smad4
shows it is required for transforming growth
factor beta and activin signaling in colorectal
cancer cells. *Proc. Natl. Acad. Sci. USA* 95:2412-
2416.

Contributed by Todd Waldman, Carolyn Lee,
Tagvor G. Nishanian, and Jung-Sik Kim
Georgetown University School of Medicine
Washington, D.C.

CHAPTER 10

Analysis of Proteins

INTRODUCTION

The purification and analysis of proteins is integral to designing oligonucleotide probes for gene cloning, confirming DNA sequence data, and synthesizing peptides for eliciting anti-peptide antibodies. This chapter presents protocols for the detection, separation, and purification of proteins. Additional aspects of protein analysis are covered elsewhere in this manual—e.g., in Chapter 12 (DNA-protein interactions), Chapter 16 (protein expression), Chapter 17 (protein glycosylation), Chapter 18 (protein phosphorylation), Chapter 19 (bioinformatics and sequence databases), Chapter 20 (protein-protein interactions), Chapter 21 (chromatin assembly and analysis), Chapter 24 (generation and use of combinatorial libraries), and Chapter 27 (RNA-protein interactions).

WHAT IS THE PROTEIN PROBLEM?

Briefly put, the protein problem is to try to understand the diversity, distribution, and function of proteins based on specific characteristics that can be determined using current technology. Whether one studies a solitary protein (if there is such a reality) or protein complexes involved in a biological function, one should ask the following questions: What is the molecular mass of the protein being studied? What are the shape and charge of the protein and its various molecular forms? Is the protein membrane-bound? Does the protein have more than one biological function? What is the copy number of the protein in the cell? What are the instantiation differences among cell types and physiological state of the cells from which the protein is being isolated? Is the protein found in protein complexes? If so, in how many different complexes? What are the kinetics and thermodynamics of protein complex formation? Do the structure and function of the protein complex depend on cell physiology? Is the protein present in more than one cellular compartment? How many molecular forms of the protein are there (i.e., variations in co- and post-translation, functional degra-

dition products, and complex inhibition or activation)?

Why is it important to answer these questions? Protein molecular mass varies between low (<6,000 Da) and high (>1,000,000 Da). A protein's molecular mass will affect the choice of purification method, since chromatographic matrices and polyacrylamide gels vary in pore size (defining molecular exclusion limits). The shape of an isolated protein is either globular (usually, an oblate ellipsoid) or fibrillar. Most proteins are globular in their native state (although this generalization may not apply intracellularly for all proteins), and most purification and analysis methods are tailored to working with globular proteins. The charge of most proteins is negative, although some are positively charged (e.g., histones). This is important to remember when choosing between anionic and cationic ion-exchange matrices (in general, choose anionic exchange). Proteins are arrays of hydrophilic or hydrophobic amino acids, and subtle changes in the distribution of charge and hydrophobicity can lead to changes in biological function and intracellular distribution. Proteins can be free or membrane-bound (i.e., bound to a lipid bilayer). Free proteins allow the use aqueous buffers during purification and analysis (occasionally requiring addition of mild nonionic detergents, e.g., for high-molecular-mass proteins or protein complexes); membrane-bound proteins always require the use of ionic detergents (i.e., SDS) or nonionic detergents (e.g., Tween, Triton X-100) to keep the protein in solution.

The function of any protein (or protein complex) is defined by a specific biological assay. In describing protein function, it is essential to remember that a protein characterized by one assay may function quite differently in another, and that an assay might reveal an artifact rather than the "true" biological function of a protein (or protein complex). Further, an assay is generally designed to comply with current biological dogma, and a class-specific protein assay will only reveal the expected

protein function. The number of protein-coding genes is agreed upon for many organisms. However, the number of distinct proteins contained in any cell and their copy numbers have not been determined unequivocally, since purification methods are innately destructive and detection methods can at best resolve proteins present at intermediate (100 to 1000 copies per cell) and high (>1000 copies per cell) copy numbers. Low-copy-number proteins (<100 copies per cell) are lost to non-specific binding on surfaces (tubes, tubing, and glass) or to natural or nonspecific protein complex interactions, where they remain undetectable by contemporary analytical methods. Protein instantiation (i.e., the variation among expressed protein types) represents the diversity between cell types and between cells of a given population. Until unicellular analysis of complete protein ensembles is possible, the complexity of intracellular protein diversity will remain a daunting challenge. Clearly, it is simplest to isolate a protein known to be enriched in a single organ (e.g., insulin from pancreas; erythropoietin from kidney). However, identifying each of the proteins that differentiate a pancreatic cell from a renal cell is beyond current protein purification and analysis methods.

Are all proteins associated with protein complexes? If one had a molecular device that could selectively pick a member of a protein complex and gently pull it from a cell, would every other protein in the complex come along? Would that complex be tethered to others that would in turn be recoverable? If so, what are the boundaries to such a network of complexes—i.e., a compartment, an entire cell? What are the relationships of a protein complex and its associated protein complexes with a defined biological pathway? How are the thermodynamics and kinetics of the complex regulated?

Proteins may be present simultaneously in more than one cellular compartment, as is known from regulatory pathways (e.g., the MAP kinase system). Does a given protein have the same biological function in each compartment? Are its molecular form (i.e., mass, charge, shape, and modification) and interactions with other proteins identical? If not, how do they differ?

The determination of the number of discrete molecular forms of a protein is beyond the capability of current methods of purification and analysis. This limitation is the result of variations in copy number, co- and post-translational modifications, folding

states, compartmental and complex distributions, degradation products, and combinations of these causes. In a higher eukaryotic cell, there may be >500,000 discrete proteins with subtle and possibly duplicative biological function, depending on their involvement in different pathways. There is no currently available purification or analysis method that would allow the deconvolution of this molecular system.

Answering these questions is at the heart of solving the protein problem. This chapter deals with what is known about protein purification and characterization. New innovations in protein science will be required to address the questions posed above.

The increasing expansion and sophistication of technologies for protein production and purification have encouraged the editors of *Current Protocols* to continue to expand the coverage of these issues. Thus, not only do several chapters of this manual now deal with specific aspects of protein analysis, but a companion volume in this series, *Current Protocols in Protein Science (CPPS)* (Coligan et al., 2006), is wholly devoted to strategies and protocols for purification and characterization of proteins. *CPPS* provides specific strategies and recommendations for experimental design of protein purification schemes for both native and recombinant proteins. In addition, it covers in depth and detail such topics as detection and separation of proteins from complex mixtures; computational, structural, and functional analyses; and detection and analysis of post-translational modifications. Throughout this introduction, reference is made, wherever appropriate, to the relevant chapters or units within *Current Protocols in Molecular Biology* and *CPPS*.

OVERVIEW OF PROTEIN PURIFICATION AND CHARACTERIZATION

Aims and Objectives

Most of the methods used in protein purification were established by the 1960s and 1970s, at least in their principles. More recent developments in protein purification have been mainly in instrumentation designed to optimize the application of each methodology. Developments in instrumentation have been stimulated by the rapid progress in molecular biology, because gene isolation has often been preceded by isolation of the gene product. Because such products can now be characterized sufficiently (i.e., partially sequenced)

using minute amounts of protein, the need for large-scale or even moderate-scale procedures has decreased. Hence there has been an explosive development of modern equipment designed specifically for dealing with amounts of protein in the milligram to microgram range. On the other hand, structural studies using X-ray crystallography and nuclear magnetic resonance (NMR) require hundreds of milligrams of pure protein, so larger-scale equipment and procedures are still needed in the research laboratory.

As more proteins, and particularly enzymes, were purified and crystallized, they started to be used increasingly in diagnostic assays and enzymatic analyses, as well as in the large-scale food, tanning, and detergent industries. Many enzymes used in industry are not in fact very pure, but as long as they do the job needed, that is sufficient. "Process" enzymes such as α -amylase, proteases, and lipases are produced in ton quantities, mainly as secretion products in bacterial cultures, and to minimize costs may undergo only limited purification processes. At the other extreme, enzyme products for research and analysis require a high degree of purification to ensure that contaminating activities do not interfere with the intended use. Anyone familiar with molecular biology enzymes will appreciate how minute levels of contamination with DNase or RNase can completely destroy carefully planned experiments.

The nature of the proteins studied has also changed substantially. Whereas enzymes were once the most favored subjects for research, they have now been superseded by nonenzymatic proteins such as growth factors, hormone receptors, viral antigens, and membrane transporters. Many of these are present in minute amounts in their natural source, and their purification can be a major task. Heroic efforts in the past have used kilogram quantities of rather unpleasant starting materials, such as human organs, and ended up with a few micrograms of pure product. It is now more usual, however, to take the genetic approach: clone the gene before the protein has been isolated or even properly identified, then express it in a suitable host cell culture or organism. The expression level may be orders of magnitude higher than in the original source, which will make purification a relatively simple task. It can be useful to know beforehand some physical properties of the protein, to facilitate the development of a suitable purification protocol from the recombinant source. On the other hand, there are now several ways of

preparing fusion proteins, which can be purified by affinity techniques without any knowledge of the properties of the target protein (see Chapter 16). Moreover, there are ways of modifying the expressed product to simplify purification further (*UNITS 10.11B & 20.2*).

Thus, the approach to protein purification must first take into account the reason it is being done, as the methods will vary greatly with different requirements. At one extreme is the one-of-a-kind purification, in a well-financed and equipped laboratory, that is carried out to obtain a small amount of product for sequencing so that gene isolation can proceed. Here, the expense of equipment and reagents may pose no problem, and a very low overall recovery of product can be acceptable, provided the product is pure enough. At the other extreme are the requirements of commercial production of a protein in large amounts on a continuing basis, where high recovery and economy of processing are the chief parameters to be considered. There are many intermediate situations as well.

Many publications in the area of protein research are entitled "Purification and characterization of . . .," and describe a purification procedure in sufficient detail that it can be reproduced in another laboratory. The characterization section may include structural, functional, and genetic information, and carrying out such studies is likely to require at least milligram quantities of pure protein. Ideally the purification should involve a small number of steps, with good recovery at each step. If the recovery is poor (<50% at any step), however, there should be some indication of what happened to the missing activity. Has it been discarded in the other fractions for the sake of purity, or does it represent a true loss of activity? If the latter, then the end-product may be less than fully active despite apparent homogeneity indicated by standard analysis. The choice between recovery and purification at each step can be problematical; taking a narrow cut of a chromatographic peak may provide a very pure fraction at the expense of losing a good deal of less pure active component on either side. In making such decisions, the objective of the exercise must be kept in mind: if yield is not important, then the tradeoff of poor yield for the sake of purity may be logical.

By far the most important requirement of a publication is reproducibility of the method reported. It is not sufficient to have carried out a process only once if it is expected that other investigators will want to repeat it. There are always factors that influence the process

but may be overlooked at first, and that can have a major effect on the purification procedure even if varied only slightly. The process being reported should always be repeated exactly as described before a manuscript is submitted for publication. There is one exception, namely, the case where purification was conducted simply to obtain enough protein for sequencing and gene isolation; if those objectives were achieved, there should be no need to provide instructions for repetition.

Sources of Material for Protein Purification

For many people embarking on a protein purification project, there is no choice of material. They are studying a particular biological tissue or organism, and the objective is to purify a protein from that source. However, there may be approaches that can make the project simpler. If, for instance, the source is difficult to obtain in large amounts, it may be best to carry out at least preliminary trials on a source species more readily obtained. The most obvious and relevant example is when the species being studied is *Homo sapiens*, and tissue samples are not readily available for practical or ethical reasons, or both. In this case, it is usual to go to where mammalian tissue is readily available (i.e., an abattoir) and work with bovine, ovine, or porcine sources. Alternatively, if quantity of tissue is not a problem, the humble laboratory rat may suffice. Once a protocol for purifying the protein from substitute sources has been worked out, it will be much easier to develop one using human material—the identical procedure may work satisfactorily. Proteins differ to a fairly small extent between species that have diverged within ~100 million years, a time frame that groups together most higher mammals. Thus, the behavior of proteins derived from different animals with respect to the various fractionation procedures is likely to be similar, and a protocol worked out for pig tissues is likely to need only minor adjustments for application to human tissues.

A second example is where the interest is mainly in the function of a protein, especially an enzyme, for which functions and actions have generally been strongly conserved through evolution. In that case, a preliminary screening of potential sources or, better still, the literature should provide a raw material that is best suited to the investigator's purposes. Considerations should include the following: (1) What functions are required of the end product? For instance, an enzyme having a

low K_m may be needed, so selecting the source with the highest activity may not suffice. (2) How convenient is it to grow or obtain the raw material, and are there problems relating to pathogenicity or extractability? (3) Does the quantity of the protein vary with growth conditions or age, and does the protein deteriorate in situ if left too long? Obviously one requires a source that reliably produces the highest amount of the desired protein per unit volume to maximize the chances of developing a good purification procedure. (4) What storage conditions are required for the raw material? It is important to consider that fresh raw material may not be immediately available whenever a purification is attempted.

The above considerations are relevant to the traditional situation for commencing a protein purification project. It is becoming increasingly common, however, for proteins to be purified as recombinant products using techniques in which the gene is expressed in a host organism or in cultured cells. This of course requires that the gene encoding the protein of interest be available. Until the mid-1980s, such material was usually obtained by hybridization of an oligonucleotide synthesized according to amino acid sequence information. This required the protein to have been purified first, so the initial task of protein purification still needed to be done at least once. More recently, genetic techniques have permitted the isolation of many genes encoding known proteins, even though the proteins may never have been studied directly. Moreover, with the completion of the Human Genome Project and related DNA sequencing efforts, many genes for both known and unknown proteins have become available and can be expressed in recombinant form without ever being purified from the host species. As a result some completely new considerations for protein purification come into play, including the possibility of modifying the gene structure not only to increase expression level and alter the protein product itself to enhance a desired function, but, equally importantly, to aid in purification. Recombinant proteins may be expressed in bacteria, yeasts, insect cells, and animal tissue cultures. Further details may be found in Chapter 16.

Detection and Assay of Proteins

During a protein purification procedure there are two measurements that need to be made, preferably for each fraction: total protein and amount of the desired protein (usually assessed in terms of bioactivity). Details of the most commonly used assay methods are

given in *UNIT 10.1A*. It is not possible to isolate a protein without a method of determining whether it is present; an assay, either quantitative or at least semiquantitative, indicating which fraction contains the most of the desired protein is essential.

Assays may range from the quick-and-easy type (e.g., instantaneous spectrophotometric measurement of enzyme activity) to long and tedious bioassays that may take days to produce an answer. The latter situation is very difficult, because by the time one knows where the protein is, it may be “was,” owing to degradation or inactivation. Moreover, this may not become clear until the next step has been completed and its products assayed. Any assay that is quick is therefore advantageous, even if this means sacrificing accuracy for speed.

Measurement of total protein is useful, as it indicates the degree of purification at each step. However, unless the next step critically depends on how much protein is present, measurement of total protein is not extremely important: a small sample can be put aside and measured later, when the purification is complete. It is, however, very important to know how much protein is present in the final, presumed pure sample, as this will indicate the specific activity (if the protein has an activity), which can be compared with that of other preparations. The general object is to obtain as high a specific activity as possible (taking into account recovery considerations), which means retaining as much of the desired protein as possible while ending up with as little total protein as possible.

Methods for Separation and Purification of Proteins

The methods available for protein purification range from simple precipitation procedures used since the nineteenth century to sophisticated chromatographic and affinity techniques that are constantly undergoing development and improvement. Methods can be classified in several alternative ways; perhaps one of the best is based on the properties of the proteins that are being exploited. Thus, the methods can be divided into four distinct but interrelated groups depending on protein characteristics: surface features, size and shape, net charge, and bioproperties.

Methods based on surface features

Surface features include charge distribution and accessibility, surface distribution of hydrophobic amino acid side chains, and, to a lesser extent, net charge at a given pH (see

discussion of net charge, below). Methods exploiting surface features depend mainly on solubility properties. Differences in solubility result in precipitation by various manipulations of the solvent in which the proteins are solubilized. Methods for obtaining an extract containing the desired protein in soluble form are given in Chapter 4 of *CPPS*. The solvent, nearly always water containing a low concentration of buffer salts, can be treated to alter properties such as ionic strength, dielectric constant, pH, temperature, and detergent content, any of which may selectively precipitate some of the proteins present. Conversely, proteins may be selectively solubilized from an insoluble state by manipulation of the solvent composition. The surface distribution of hydrophobic residues is an important determinant of solubility properties; it is also exploited in hydrophobic chromatography, both in the reversed-phase mode (*UNITS 10.12-10.14*) and in aqueous-phase hydrophobic-interaction chromatography.

Also included in this category is the highly specific technique of immunoaffinity chromatography (*UNIT 10.11A*), in which an antibody directed against an epitope on the protein surface is used to pull out the desired protein from a mixture.

Methods based on whole structure:

Size and shape

Although the size and shape of proteins can have some influence on solubility properties, the chief method of exploiting these properties is gel-filtration chromatography (*UNIT 10.9*). In addition, preparative gel electrophoresis makes use of differences in molecular size. Proteins range in size from the smallest classified as proteins rather than polypeptides (~5000 Da) up to macromolecular complexes of many million daltons. Many proteins in their bioactive states are oligomers of more than one polypeptide, and these can be dissociated, though normally with loss of overall structure. Thus, many proteins have two “sizes”: that of the native state and that (or those) of the polypeptides in the denatured and dissociated state. Gel-filtration procedures normally deal only with native proteins, whereas electrophoretic procedures commonly involve separation of dissociated and denatured polypeptides.

Methods based on net charge

Two separation techniques exploit the overall charge of proteins: ion-exchange chromatography (by far the most important) and

electrophoresis (*UNITS 10.2A, 10.3 & 10.4*; see also Chapter 10 of *CPPS*). Ion exchangers bind charged molecules, and there are essentially only two types of ion exchangers, anion and cation. The net charge of a protein depends on the pH—positive at very low pH, negative at high pH, and zero at some specific point in between, termed the isoelectric point (pI). It should be stressed that at its pI a protein has a great many charges; it just happens that at this pH the total negatives exactly equals the total positives. The most charged state (disregarding the charge sign) is in the pH range 6.0 to 9.0. This is the most stable pH range for most proteins, as it encompasses common physiological pH values. Ion exchangers consist of immobilized charged groups and attract oppositely charged proteins. They provide the mode of separation that has the highest resolution for native proteins. High-performance reversed-phase chromatography has equivalent or even better resolution, but it generally involves at least partial denaturation during adsorption and is not recommended for sensitive proteins such as enzymes. Protein purification using ion-exchange chromatography has mainly employed positively charged anion exchangers, for the simple reason that the majority of proteins are negatively charged at neutral pH (i.e., have a low isoelectric point). Details of this methodology are found in *UNIT 10.10*.

Methods based on bioproperties (affinity)

A powerful approach for separating the desired protein from others is to use a biospecific method in which the particular biological property of the protein is exploited. One such property is the affinity between a protein's binding site and its ligand: the ligand, when immobilized, attracts the protein from a mixture, while other molecules are washed away (*UNITS 10.11B & 24.5*). This general approach is known as affinity chromatography; there are also nonchromatographic methods of exploiting the same concept (*UNITS 10.20, 10.21, 10.24 & 10.27*). Although this direct affinity is limited to proteins with a natural binding activity, most proteins of interest do have specific ligands: enzymes have substrates and cofactors, and hormone-binding proteins, growth factors, and receptor molecules are designed to bind specifically and tightly to their natural partners. There are also other related affinity methods for nonbinding proteins. One is immunoaffinity chromatography (*UNIT 10.11A*), in which the "ligand" is an antibody to the protein in question. An equally important and

now widely used method is the creation of a fusion protein whose fused portion has a strong affinity toward an immobilized ligand (Chapter 16). For this latter technique it is not necessary for the protein of interest to have any binding property of its own; the fusion portion possesses a strong affinity for standardized immobilized ligands.

Characterization of the Protein Product

Once a pure protein is obtained, it may be employed for a specific purpose such as enzymatic analysis (e.g., glucose oxidase and lactate dehydrogenase) or as a therapeutic agent (e.g., insulin and growth hormone). However, it is normal, when a protein has been isolated for the first time, to characterize it in terms of structure and function. Several features are generally expected to be addressed (and reported in the literature) in the characterization of a new protein. One is the molecular weight, or at least the size of the subunit(s), which is determined by SDS-PAGE (*UNITS 10.2A-10.5*) and/or gel filtration (*UNIT 10.9*). Spectral properties such as the UV spectrum (Trp and Tyr content) and circular dichroism (CD) spectrum (secondary structure), and special characteristics of proteins with prosthetic groups (e.g., quantitation and spectra), may also be described. For glycoproteins, the number and nature of substituent carbohydrates should be determined (Chapter 17). Furthermore, if the gene has not already been reported, some amino-terminal sequence analysis should be given, if at all possible, along with the results of a database search for similar sequences (*UNIT 19.3*). Functional proteins should be demonstrated to have the appropriate function, and detailed kinetic characterization is appropriate for enzymes. Ultimately the full three-dimensional structure of the protein may be determined, which will require crystals; any successful crystallization attempts should be reported.

The Protein Purification Laboratory

The requirements for a protein purification laboratory cannot be exactly formulated because they depend greatly on the types and amounts of proteins being isolated. To cover all eventualities, it would be necessary to have one set of equipment to deal with submicrogram quantities and another set to deal with multi-gram quantities—a range of $\sim 10^8$! One laboratory dedicated to protein purification may not need small-scale equipment if, for example, its work involves plasma proteins that are

always available in large quantities. Another may have all the latest in high-performance equipment but not be able (nor need) to handle quantities of protein in excess of a few milligrams.

If it is assumed that neither extreme in quantity is to be attempted, and that the laboratory is handling a variety of protein types and sources, then certain basic pieces of equipment are needed. Obtaining the starting material and making an extract of it require homogenization equipment and centrifuges to remove insoluble residues. Preliminary fractionation starting from a crude extract of tissue or cells requires equipment and materials that will not become clogged by particulates. Adsorbents and similar materials used at the first step should be relatively inexpensive so that when performance falls off after a few uses, owing to intransigent impurity buildup, they can be discarded. It is also relevant that a larger amount is handled at the initial step than later steps; therefore, reagent expense can be an important consideration. After the first one or two steps, the sample should be sufficiently clean and clear to permit use of high-performance equipment.

High-performance liquid chromatography (HPLC) is a term with a variety of meanings (see *UNITS 10.12-10.14*). To some it refers exclusively to reversed-phase chromatography (here distinguished as RP-HPLC); to others it includes all sorts of chromatography, provided that the equipment is fully automated and high-performance adsorbents are used. A high-performance system designed specifically for proteins—produced by Amersham Pharmacia Biotech (see *APPENDIX 4*) under the name Fast Protein Liquid Chromatography or FPLC—uses standard protein chromatographies such as ion exchange, hydrophobic interaction, and gel filtration. Scaleup is possible with larger equipment based on the FPLC design, so that laboratory development can be quickly translated to large-scale production. FPLC is designed to separate proteins in their native active configuration, whereas RP-HPLC often causes at least transient denaturation during adsorption and elution. RP-HPLC has a high resolving power, but it is best suited to peptides and proteins smaller than ~30 kDa. Chromatography run with older-style low-pressure adsorbents is sometimes referred to as “low-performance” or “open-column” chromatography; neither of those descriptions is necessarily accurate. These methods require simple fraction collector and monitoring equipment. This equipment will be used for larger-scale

operations (tens of milligrams of protein and upward), probably at an earlier stage in the protocol than with HPLC.

Various columns, both prepacked with proprietary adsorbents and empty for self-packing, will be needed, with the sizes and types depending on the scale of operations. Several anion-exchange columns of different sizes, one or two cation-exchange columns, and gel-filtration media are all essential, along with a range of alternative adsorbents such as hydrophobic interaction materials, dyes, hydroxyapatite, and chromatofocusing and specialist affinity media.

Fully equipped protein purification laboratories should also have preparative electrophoresis and isoelectric focusing apparatuses for the rare occasions when other techniques fail to give sufficient separation.

In addition to equipment used in the actual fractionation processes, a variety of other items are needed. In particular it is important to be able to change buffers quickly and to concentrate protein solutions with ease. These operations require such items as dialysis membranes (*UNIT 10.5 & APPENDIX 3C*), ultrafiltration cells (*APPENDIX 3C*), and gel-filtration columns of various sizes (*UNIT 10.9*).

Finally, equipment for assaying and analyzing the preparations is needed. Most such equipment is fairly standard in biochemical laboratories and includes spectrophotometers, scintillation counters, analytical gel and capillary electrophoresis apparatuses, immunoblotting materials, and immunochemical reagents. A listing of standard equipment is found in *APPENDIX 2*.

STRATEGIES FOR PROTEIN PURIFICATION

Classification of Proteins

As with most heterogeneous collections of things, proteins can be classified in several different ways, such as by function, by structure, or by physicochemical characteristics. Each protein species consists of identical molecules with exactly the same size, amino acid sequence, and three-dimensional shape. In this way a solution of a mixture of proteins differs from a solution of synthetic polymers or sheared DNA, both of which contain a complete spectrum of possible sizes centered around the average. The protein mixture has only discrete sizes of molecules corresponding to each type of protein present. Although it would be possible to classify proteins by size, it would be of limited use, as there is

Table 10.0.1 Classification of Proteins by Structural Characteristics

Structural characteristic	Examples	Comments
Monomeric	Lysozyme, growth hormone	Usually extracellular; often have disulfide bonds
Oligomeric		
Identical subunits	Glyceraldehyde-3-phosphate dehydrogenase, catalase, alcohol dehydrogenase, hexokinase	Mostly intracellular enzymes; rarely have disulfide bonds
Mixed subunits	Aspartate carbamoyltransferase, pertussis toxin	Allosteric enzymes; different subunits have separate functions
Membrane-bound		
Peripheral	Mitochondrial ATPase, alkaline phosphatase	Readily solubilized by detergents
Integral	Porins, cytochromes P ₄₅₀ , insulin receptor	Require lipid for stability
Conjugated	Glycoproteins, lipoproteins, nucleoproteins	Many extracellular proteins contain carbohydrate

Table 10.0.2 Classification of Proteins by Function

Function	Examples
Amino acid storage	Seed proteins (e.g., gluten), milk proteins (e.g., casein)
Structural	
Inert	Collagen, keratin
With activity	Actin, myosin, tubulin
Binding	
Soluble	Albumin, hemoglobin, hormones
Insoluble	Surface receptors (e.g., insulin receptor), antigens (e.g., viral coat proteins)
With activity	Enzymes, membrane transporters (e.g., amino acid uptake systems, ion pumps)

usually no obvious relationship between size and function.

A more useful structural classification takes into consideration shape and oligomeric structure (Table 10.0.1). In part, structure reflects biological location and origin. Simple, fairly rigid protein molecules occur in the extracellular environment, more complex and readily deactivated molecules are found intracellularly, and hydrophobic proteins are associated with membranes.

Classification by function is even more relevant (Table 10.0.2). Proteins can be simply stores of amino acids, can be structural, or can have specific binding functions. The most

“functional” proteins are enzymes, which have both binding and catalytic roles. In part, this reflects the degree to which the detailed structure is a requirement for the protein’s function, which in turn relates to conservation of structure through evolution. But as with every attempt at classification, there will always be examples that do not fit the pattern well. Most proteins of interest to the pharmaceutical industry belong to the general class of binding proteins: for instance, hormones (e.g., insulin and bovine somatostatin), viral antigens (e.g., hepatitis B antigen), growth factors [e.g., interferons, interleukins and colony-stimulating factors (CSFs)], and antibodies.

Strategies Based on Classification

Soluble extracellular proteins

The source of soluble extracellular proteins is the extracellular medium, whether it be an animal source such as blood or spinal fluid, or a culture medium in which bacterial, fungal, animal, or plant cultures have been grown. Generally these do not contain a large number of different proteins (blood is an exception), and the desired protein may be a major component, especially if produced as the result of recombinant expression. Nonetheless, the protein in the starting material may be quite dilute, and a large volume may therefore need to be processed. The starting fluid may also contain many compounds other than proteins, whose behavior must be taken into account. The first stage should aim mainly to reduce the volume and remove as much nonprotein material as possible; some protein-protein separation is also useful, but not essential. No general rules can be given, but a batch adsorption method using an inexpensive material such as hydroxyapatite, ion-exchange resin, immobilized metal affinity chromatography (IMAC) medium, or affinity adsorbent is best, if feasible. Following the first step, the sample should be in a form that is amenable to standard purification processes such as precipitation and column chromatography (see Chapter 8 of *CPPS*).

Intracellular (cytoplasmic) proteins

To obtain soluble intracellular proteins (which are mainly enzymes), cells must be broken open or lysed to release their soluble contents. The ease with which cell disruption can be accomplished varies considerably; animal cells are readily broken, as are many bacteria, but plants and fungi have tough cell walls. Methods for obtaining cell extracts are given in Chapter 4 of *CPPS*. The macromolecular soluble contents of cells are mainly proteins, with nucleic acids as a minor but significant component. Bacterial extracts may be viscous unless DNase is added to break down the long DNA molecules. Although chromatographic procedures can be applied to crude extracts, valuable high-performance materials should not be employed in the first step, as there are always compounds, including unstable proteins, that may bind to them and be difficult to remove.

Membrane-associated proteins

There are two approaches to isolating a membrane-associated protein. In one method, the relevant membrane fraction can first be prepared and then used to isolate the protein. Alternatively, whole tissue can be subjected to

an extraction that solubilizes the membranes and releases the cytoplasmic contents as well. The former approach is much better in that purification is accomplished by isolating the membranes: the specific activity of the solubilized membrane fraction will be much higher than with the second method. However, the process of purifying the membrane fraction may lead to substantial losses, and it may be difficult to scale up. If total recovery of the protein is more important than purity, a whole-tissue extract is likely to be more appropriate. Although this means that a greater degree of purification is needed, the fact that membrane proteins, by definition, have properties somewhat different from those of cytoplasmic proteins permits some very effective purification steps (e.g., hydrophobic chromatography).

Peripheral membrane proteins are only loosely attached and may be released by gentle conditions such as high pH, EDTA, or low (nonionic) detergent concentrations. Once in solution, some peripheral proteins no longer require the presence of detergent to maintain their solubility. Integral membrane proteins are much more demanding—they require high concentrations of detergent for solubilization (i.e., complete solubilization of the membrane is needed to release them) and generally are neither soluble nor stable in the absence of detergent. It is sometimes necessary to maintain natural phospholipids in association with the proteins in order to maintain activity. Even when the final objective does not require activity (e.g., protein sequencing), it is generally necessary to maintain bioactivity so that an assay can be used to monitor the presence of the protein during the purification process. If a particular band on a gel is known to be the desired protein, then no other assay is needed and loss of bioactivity can be allowed.

Purification processes may be affected by the presence of detergents. The problem of association with detergent micelles makes purifying integral membrane proteins difficult; the close association of the different proteins originating from membranes often results in very poor separation in conventional fractionation procedures.

Insoluble proteins

Natural proteins that are insoluble in normal solvents are generally structural proteins, which are sometimes cross-linked by post-translational modification. The first stage of purification is obvious—it involves extracting and washing away all proteins that are soluble, leaving the residue containing the desired

material. Further purification in a native state, however, may be impossible; extracting away other proteins using more vigorous solvents or attempting to solubilize the target protein may destroy the natural structure. Cross-linked proteins such as elastin or collagen cannot be dissolved without breaking the cross-links, and the individual proteins may even be cross-linked together.

Insoluble recombinant proteins (inclusion bodies)

A major new class of insoluble proteins are recombinant proteins expressed (usually in *E. coli*) as inclusion bodies. These are dense aggregates found inside cells that consist mainly of a desired recombinant product, but in a nonnative state. Inclusion bodies may form for a variety of reasons, such as insolubility of the product at the concentrations being produced, inability to fold correctly in the bacterial environment, or inability to form correct, or any, disulfide bonds in the reducing intracellular environment. Their purification is simple, as the inclusion bodies can be separated by differential centrifugation from other cellular constituents, giving almost pure product; the problem is that the protein is not in a native state and is insoluble. Some methods for obtaining an active product from inclusion bodies are described in *CPPS* (Palmer and Wingfield, 2004).

Soluble recombinant proteins

Recombinant proteins that are not expressed in inclusion bodies either will be soluble inside the cell or, if an excretion vector is being used, will be extracellular (or, if *E. coli* is the host, possibly periplasmic). They can be purified by conventional means. In some systems, expression is so good that the desired product is the major protein present and its purification is relatively simple. In systems where the expression level is low, the purification process can be tedious—though easier, it can be hoped, than isolation from the natural source. It should be remembered that a procedure developed for isolating a protein from natural sources may not work successfully with the recombinant product, because the nature of the other proteins present influences many fractionation procedures.

Many recombinant proteins are now produced in a fusion form in which the fusion portion is designed to facilitate purification by affinity chromatography (*UNIT 10.11B*). A variety of “fusion tags” are available as part of the expression vector (see *UNITS 16.6–16.8*

and Sassenfeld, 1990, for specific examples). These include whole proteins such as protein A, glutathione-S-transferase, and maltose-binding protein, as well as short peptide segments that may be recognized by a specific antibody, become biotinylated in the host cell, or bind tightly to an immobilized metal ion adsorbent. Nearly complete purification can be achieved in one step by passing the crude extract through an appropriate adsorbent and eluting specifically. A further advantage of the fusion approach is that expression levels are likely to be determined more by the transcription and translation signals for the fusion peptide (which have been optimized in the vector) than by the structural features of the inserted gene. The only remaining task is to remove the fusion peptide, though for some purposes this is not necessary (see *UNIT 16.4B*). This requires a proteolytic step: highly specific proteases such as blood clotting factors are generally used, their recognition sites having been incorporated in the sequence joining the fusion peptide. It is sometimes convenient to “digest” the fusion protein from the adsorbent using the protease.

Unstable proteins may be modified by the molecular biological technique of site-directed mutagenesis to remove the site of instability—for instance, an oxidizable cysteine. Such techniques are appropriate for commercial production of proteins, but may of course alter natural functioning parameters. One modification that can be useful is increased thermostability, although it is not easy to predict mutations that will improve that parameter. Thermostable proteins originating from thermophilic bacteria do not need structural modification and, if expressed in large amounts, can be purified satisfactorily in one step by simply heat-treating the extract at 70°C for 30 min, which will denature virtually all the host proteins (e.g., see Oka et al., 1989).

The host bacteria used for production of recombinant proteins are usually *E. coli* or *Bacillus subtilis*; they may express proteins at anywhere from 1% to >50% of the cellular protein, depending on such variables as the source, promoter structure, and vector type. Generally the proteins are expressed intracellularly, but leader sequences for excretion may be included. In the latter case, the protein is generally excreted into the periplasmic space, which limits the amount that can be produced. Excretion from gram-positive species such as *B. subtilis* sends the product into the culture medium, with little feedback limitation on the total expression level.

PROTEIN PURIFICATION FLOW CHARTS

Protein purification flow charts are presented to give a broad outline of the methods used for different types of proteins. They cannot give any detail, as the process appropriate for each protein will have its own variations at each stage. In most cases, the first stage is to obtain a solution containing the desired protein, after which it can be dealt with by the many separation techniques described in the following chapters. In some cases the insolubility of the desired protein can be exploited by removing soluble fractions. Purification procedures are commonly divided into three stages: (1) the primary steps deal with crude mixtures of proteins and other molecules present in the raw material; (2) the secondary processing generates a product near homogeneity; and (3) the polishing steps remove minor contaminants, a process that is especially important for therapeutic proteins.

Soluble Recombinant Proteins

Proteins expressed in a recombinant manner may be (1) soluble in the cytoplasm, (2) insoluble as inclusion bodies, (3) excreted from the cells into the culture medium, (4) excreted into the periplasmic space (e.g., in gram-negative bacteria), or (5) associated with organelles or membrane fractions. In addition they may be expressed (6) as the normal, mature, naturally occurring protein, (7) containing a natural leader peptide that would normally be processed, (8) as a fusion protein with a peptide that is not natural to the protein, or (9) lacking glycosylation or other posttranslational modification, or modified incorrectly. Possibilities (1) to (5) affect the method of extraction used to obtain the starting material for purification. Cases (6) to (9) can affect the methods used for purification.

The scheme for purifying soluble recombinant proteins is outlined in Figure 10.0.1. The first stage is to obtain a clarified solution containing the desired protein, with as little in the way of unwanted proteins as possible. For soluble cytoplasmic proteins, case (1), it is not normally possible to exclude any significant amount of unwanted soluble proteins, but in cases (2) to (5) the compartmentalization away from the cytoplasm allows such separation in the initial stage.

It may be necessary to carry out a concentration step before proceeding, especially if the protein has been excreted into the culture medium. Normally ultrafiltration is used (see APPENDIX 3C), although other tech-

niques are possible, especially if the extract contains particulates that block ultrafiltration membranes.

Recombinant expression in the cytoplasm of bacteria followed by extraction via total cell disruption results in large amounts of nucleic acids being solubilized with the protein. A number of treatments to remove nucleic acids are possible. Streptomycin is used to precipitate ribosomal material, and cationic polymers such as protamine (a basic protein) and polyethylenimine will form insoluble complexes (at low ionic strength) with nucleic acids. In addition, viscosity caused by DNA can be reduced by adding small amounts of DNase.

Insoluble Recombinant Proteins

It has been found that many proteins do not fold correctly when expressed in bacteria (mainly in *E. coli*), and as a result aggregation occurs, leading to large insoluble inclusion bodies within the cytoplasm of the cells (see Chapter 6 of CPPS). Although this creates major difficulties in obtaining satisfactory amounts of active native product, it greatly simplifies the initial stage of purification.

The purification scheme for recombinant insoluble proteins is outlined in Figure 10.0.2. After cell disruption, inclusion bodies can be obtained in a fairly pure state by differential centrifugation. They must then be solubilized, however, and the active protein generated by encouraging correct folding. Solubilization is usually accomplished with guanidine hydrochloride and/or urea, and thiols such as 2-mercaptoethanol or glutathione are included to disrupt any disulfides that have formed and prevent more from forming. Folding the protein correctly may require a variety of additions to the solution, as well as slow removal of the denaturant. The latter can be carried out by simple dilution or by dialysis (APPENDIX 3C). Folding occurs best at low protein concentrations, so dilution may be adequate. If the native protein does contain disulfides, then it is important to create redox conditions such that some (but not excessive) oxidation of thiols can occur. A combination of oxidized and reduced glutathione is commonly used. In addition, the action of the enzyme protein disulfide isomerase, which can make and unmake disulfides by exchange reactions, has been found to be beneficial in many cases. If the native protein is of intracellular origin, it probably will not contain disulfides; it will, however, contain cysteines, so a full reducing potential should be maintained.

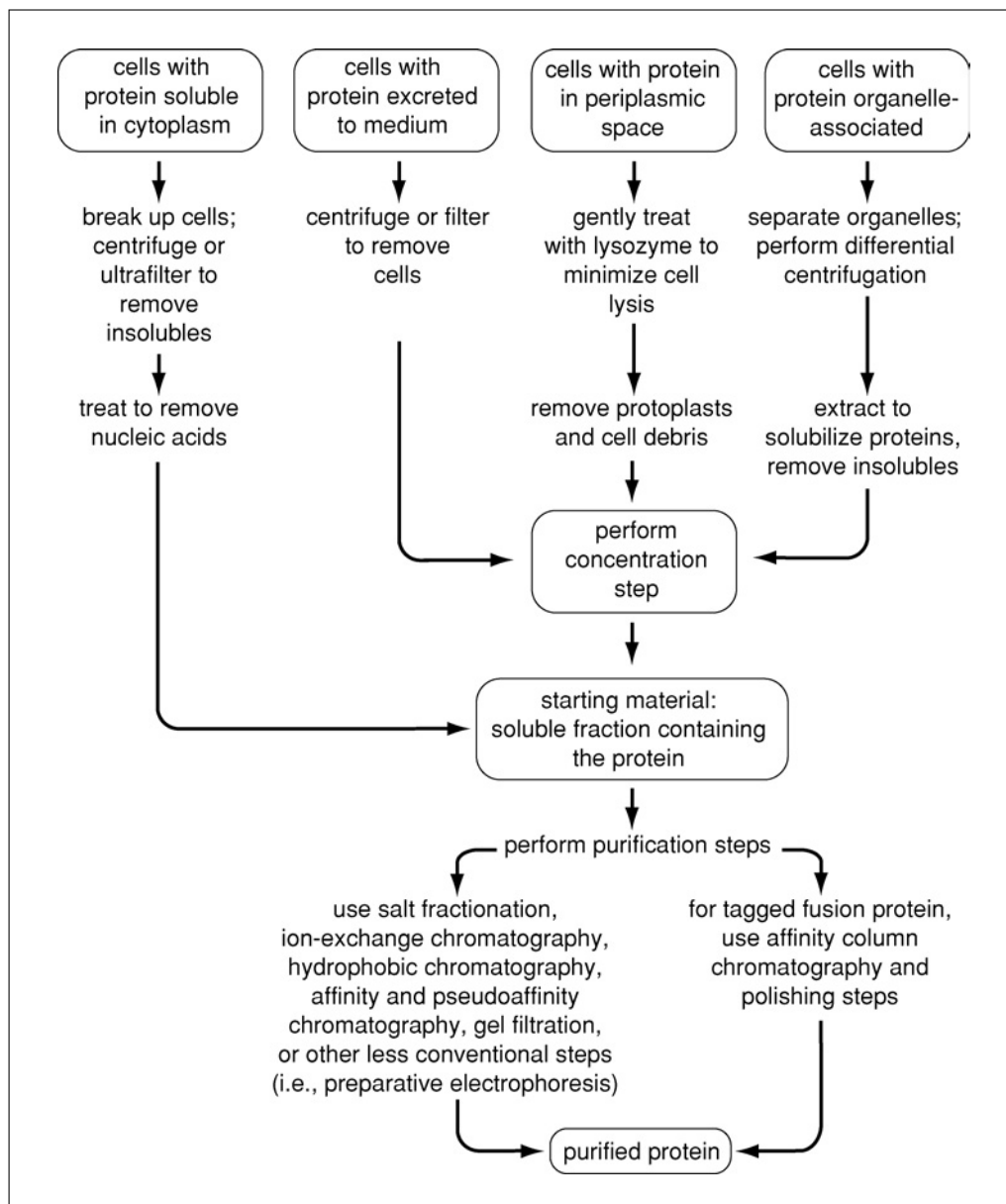


Figure 10.0.1 Purification scheme for soluble recombinant proteins, which may be excreted or located in the periplasm, the membrane fraction, or most commonly the cytoplasm. The first step is to obtain an extract containing the desired protein in soluble form. After this, conventional purification steps may be carried out, or affinity purification of tagged fused proteins can be performed.

Specific methodology is discussed in *CPPS* (Wingfield, 1995; Wingfield et al., 1995).

Not all proteins can fold unassisted by other cellular components. Chaperonins are proteins whose role is to assist in folding proteins including those unfolded by heat shock (Zeilstra-Ryalls et al., 1991). The ones most studied, which are just becoming commercially available as of 1996, are the *E. coli* chaperonins GroEL and GroES, both of which are needed, together with ATP, to renature many proteins. Proline residues can adopt two isomeric conformations in proteins, and the wrong con-

formation is switched to the correct one by the enzyme prolyl isomerase, aiding the process of protein folding. At present these are not large-scale prospects, both because of the cost of the chaperonins and because the agents operate best in vitro at very low protein concentrations.

Once the proteins are folded, the purification process consists of removing small amounts of still incorrectly folded protein plus any other host proteins that were trapped with the original inclusion bodies. The former may be difficult, as incorrectly folded species have

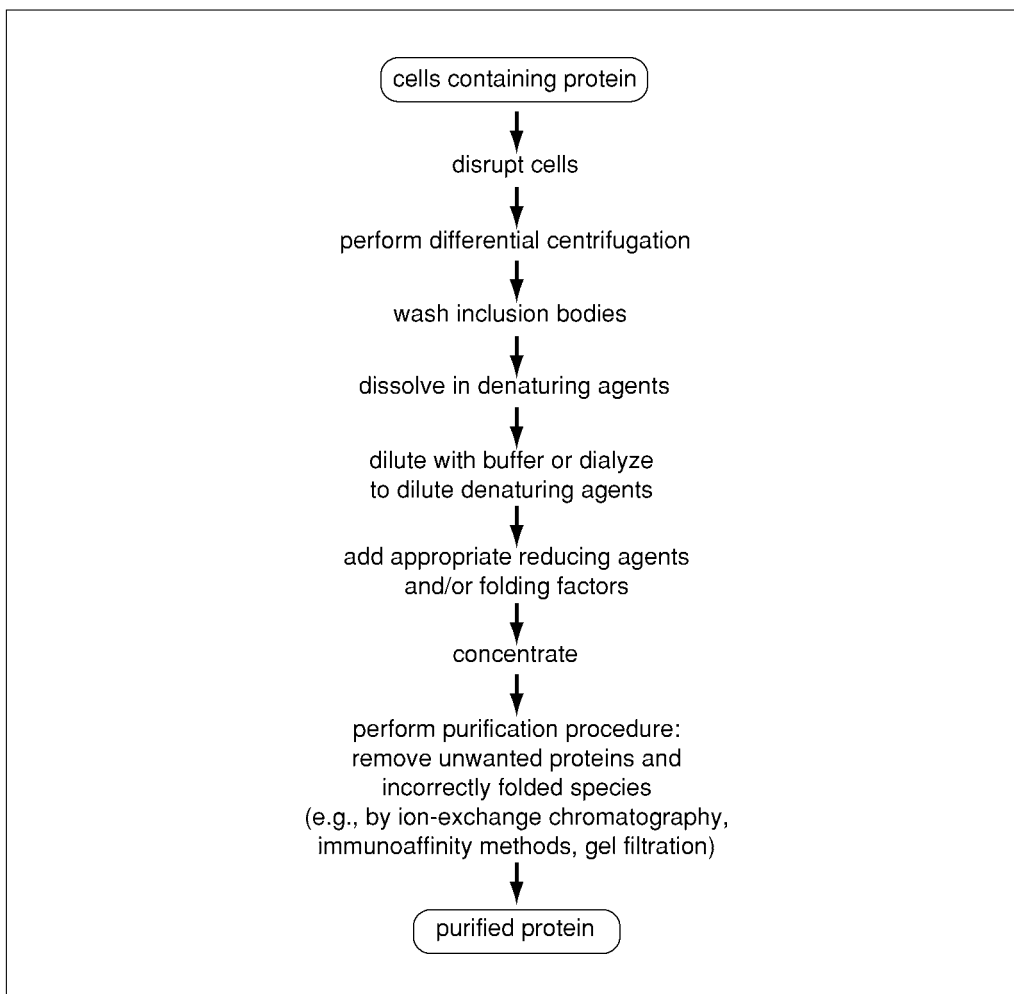


Figure 10.0.2 Purification scheme for insoluble recombinant proteins that are produced as inclusion bodies in the cytoplasm of host cells. The cells must be broken open and the insoluble inclusion bodies separated by differential centrifugation. Solubilization is achieved by the use of denaturing solvents, and renaturation of the dissolved protein occurs on removal of the denaturant. Further polishing steps will be needed to remove small amounts of contaminating proteins as well as incorrectly folded species. Additional information can be found in *CPPS* (see Pain, 1995; Wingfield et al., 1995; Palmer and Wingfield, 2004).

a size and charge similar to those of the correct product. However, subtle differences arising from the folded conformation can be exploited by chromatographic techniques. In ideal cases immunoaffinity techniques using antibodies specific for either the incorrectly folded form or the correct one can be used to resolve the mixture.

Soluble Nonrecombinant Proteins

There are so many sources of soluble proteins that it is not possible to give a complete overview of methods used to obtain starting extracts from which a desired protein can be isolated. The sources can be classified as either microorganisms, plants, or animals, as shown in Figure 10.0.3, but these in turn should be subdivided according to how the starting ex-

tract is obtained. In particular there is a distinction between extracellular and intracellular proteins. With the latter it is necessary to disrupt the cells and release the proteins, whereas with the former, if the extracellular fluid can be obtained directly, there need be no contamination with intracellular proteins. Extracellular sources include microorganism culture medium, plant and animal tissue culture medium, venoms, milk, blood, and cerebrospinal fluids. Soluble proteins may also occur within organelles such as mitochondria; these may best be obtained by first isolating the organelle, then disrupting it to release the contents.

The starting extract normally contains between 5 and 20 mg protein per milliliter, though lesser concentrations can be dealt with,

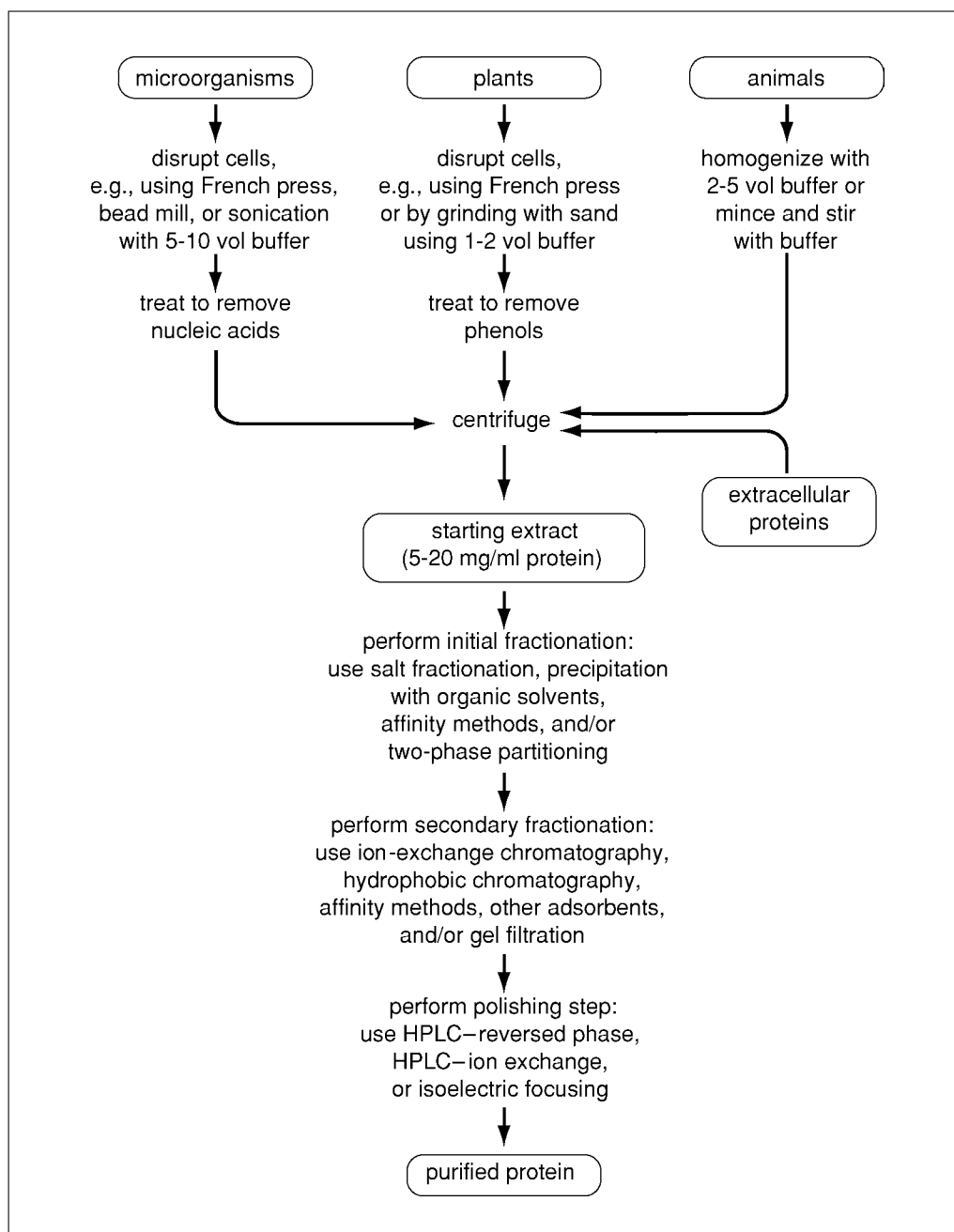


Figure 10.0.3 Purification scheme for soluble proteins present in their natural host cells. Cells must be disrupted to release the proteins, usually in the presence of 2 to 10 ml of a suitable buffer per gram weight. After removal of insoluble material, the process will generally require several steps, using various standard fractionation procedures in a suitable order. For production of highly pure protein, a final polishing step may be required to remove final trace contaminants. Additional information can be found in *CPPS* (see Wingfield, 2005).

especially if working on a small scale. It may be necessary to include a concentration step before starting the purification process in order to approach that level. There are exceptions to every rule, however, and very high protein concentrations can be handled, for example, with two-phase partitioning (Walter and Johansson, 1994). When isolating proteins on

a large scale, the volumes being manipulated become of increasing concern, so maximizing protein concentration can be an important aim. The starting extract should be clarified, usually by centrifugation; on a large scale, ultrafiltration methods (*APPENDIX 3C*) are becoming more widely used. Pretreatment of certain extracts to remove excessive amounts of nucleic

acids, phenolics, and lipids may be necessary in order to obtain an extract that is amenable to standard fractionation procedures.

Fractionation procedures can somewhat arbitrarily be divided into three steps: initial fractionation, secondary fractionation, and polishing. In initial processing, which deals with a large amount of extract that is not all protein, materials may become soiled and may be unable to be used many times. Consequently, methods that do not require expensive reagents or adsorbents are preferred. Classic salt fractionation and the less-used organic solvent fractionation can achieve, if not a high degree of purification, a useful level of concentration and removal of much unwanted nonproteinaceous material. Alternatively, a highly selective affinity procedure may be used as the first step, but only if the affinity material is inexpensive to make and/or the extract is a simple, clear solution as opposed to a turbid whole-cell homogenate.

Secondary processing achieves the main purification, and in difficult situations may involve two or more steps. Ion-exchange (*UNIT 10.10*) and hydrophobic-interaction chromatography, gel filtration (*UNIT 10.9*), and affinity techniques (*UNITS 10.11B & 20.2*) are the main procedures. Finally, it may be necessary to remove traces of contaminants by “polishing” using high-resolution procedures such as RP-HPLC (*UNITS 10.13 & 10.14*) and isoelectric focusing (IEF; *UNITS 10.3 & 10.4*). Because every protein has unique characteristics, it is impossible to make general statements about procedures to be followed.

Membrane-Associated and Insoluble Nonrecombinant Proteins

Proteins that are not physiologically soluble can be purified after extracting and removing soluble proteins, thereby achieving a substantial degree of purification at the extraction step (Fig. 10.0.4; also see Wingfield, 2002, 2005). To carry out a purification it is nearly always necessary to obtain the desired protein in a soluble form, which will often require the addition of solubilizing agents such as detergents. Some proteins remain insoluble even with detergent treatment, and so can be substantially purified by removing the soluble fractions. Some membrane-associated proteins become partly solubilized during breaking up of the tissue, and recovery in the particulate fraction may be poor. In such cases it may be best to solubilize the whole tissue by including detergent in the homogenizing buffer. Extraction of insoluble residues using

detergents can be done differentially; some proteins are released at low detergent concentration, whereas others require complete solubilization of the membrane fraction. Suitable detergents include nonionic (e.g., Triton) and weakly acidic types (e.g., cholic acid derivatives). Strongly acidic detergents such as sulfate esters (e.g., sodium dodecyl sulfate) usually cause denaturation.

Detergents can be removed either by adsorption of the protein on a column and subsequent elution without detergent, by use of special detergent-adsorbing beads, or even by extraction with nonmiscible organic solvents in which the detergent partitions. On the other hand, many membrane proteins require the presence of detergent at all times in order to remain in solution and in a native conformation. These include most integral membrane proteins, for example cytochrome P₄₅₀, transmembrane receptors, and transporters. The most sensitive proteins require a particular combination of natural lipids (in addition to the detergent) to maintain structural integrity. Purification methods include most of those used for soluble proteins, but some techniques are not recommended if detergent is needed at all times. For instance, ammonium sulfate precipitation will often cause a detergent-protein complex to come out of solution and float rather than sink on centrifugation; this can be useful, but the “floatate,” when redissolved, may have a high detergent content. Hydrophobic chromatography can be very useful, as membrane proteins are naturally hydrophobic.

Integral membrane proteins that are completely insoluble in normal detergents may be solubilized by denaturation using compounds such as sodium dodecyl sulfate and guanidine hydrochloride. Some cross-linked proteins such as elastin are not soluble without disruption of the covalent linkages.

WHAT THIS CHAPTER COVERS

This chapter consists of protocols that permit researchers to obtain answers to the questions described above: How much protein is there? Is the protein pure? Does the protein have subunits? How many protein subunits are there? How is the protein isolated? How can the protein be synthesized *in vitro*? How can a protein be biosynthetically labeled? How can amino-terminal sequences of scarce proteins be determined? How can internal sequences be derived from N-terminally blocked proteins?

As indicated in Figure 10.0.5, the answer to the first question is provided by colorimetric and spectrophotometric methods (*UNIT 10.1A*)

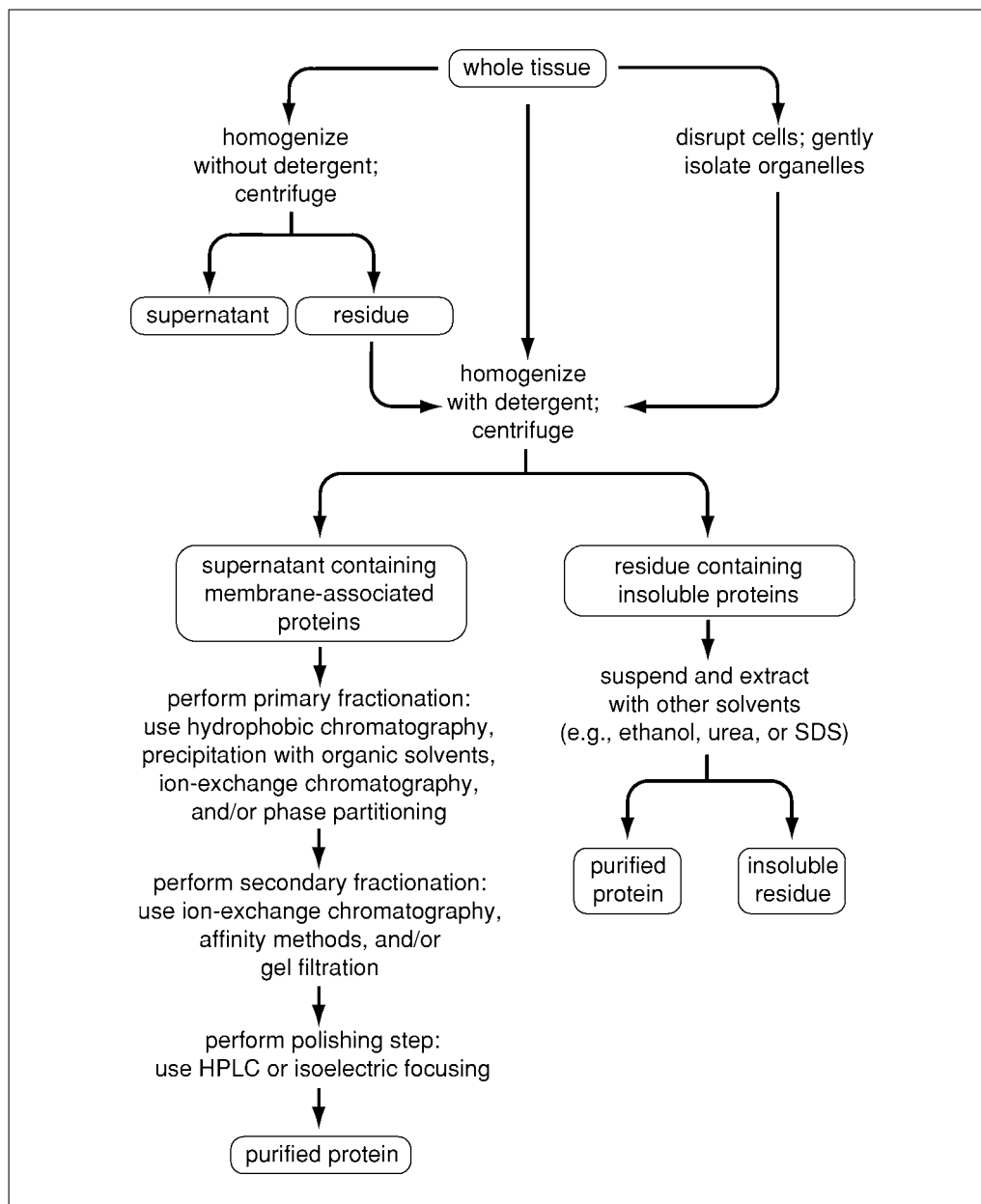


Figure 10.0.4 Purification scheme for membrane-associated and poorly soluble proteins (nonrecombinant). An initial purification can be achieved by isolation of organelles containing the desired protein. Membrane proteins are normally solubilized with a nonionic detergent, although loosely associated proteins may be extracted without detergent at high pH, with EDTA, or with small amounts of an organic solvent such as *n*-butanol. Normal fractionation procedures may need some modification if the detergent is required throughout to maintain the integrity of the protein.

and by amino acid analysis (UNIT 10.1B). The amount of protein can also be determined by comparing the staining intensity of an unknown protein to the staining intensity of protein standards separated by either one- or two-dimensional polyacrylamide gel electrophoresis (PAGE; UNITS 10.2A-10.4). UNIT 10.5 describes the process required for digital imaging of one- and two-dimensional gels from image capture using scanners to image manipulation and analysis using advanced imag-

ing software. Digital imaging is an immensely useful technique for analyzing changes in protein expression and is a mainstay for functional genomic analysis. Proteins can also be stained while they are still within the polyacrylamide gel (UNIT 10.6) or after they have been transferred to a blot transfer membrane (UNITS 10.7 & 10.19). In addition, protein-specific monoclonal or polyclonal antibodies can be used for detection by immunoblotting (UNIT 10.8).

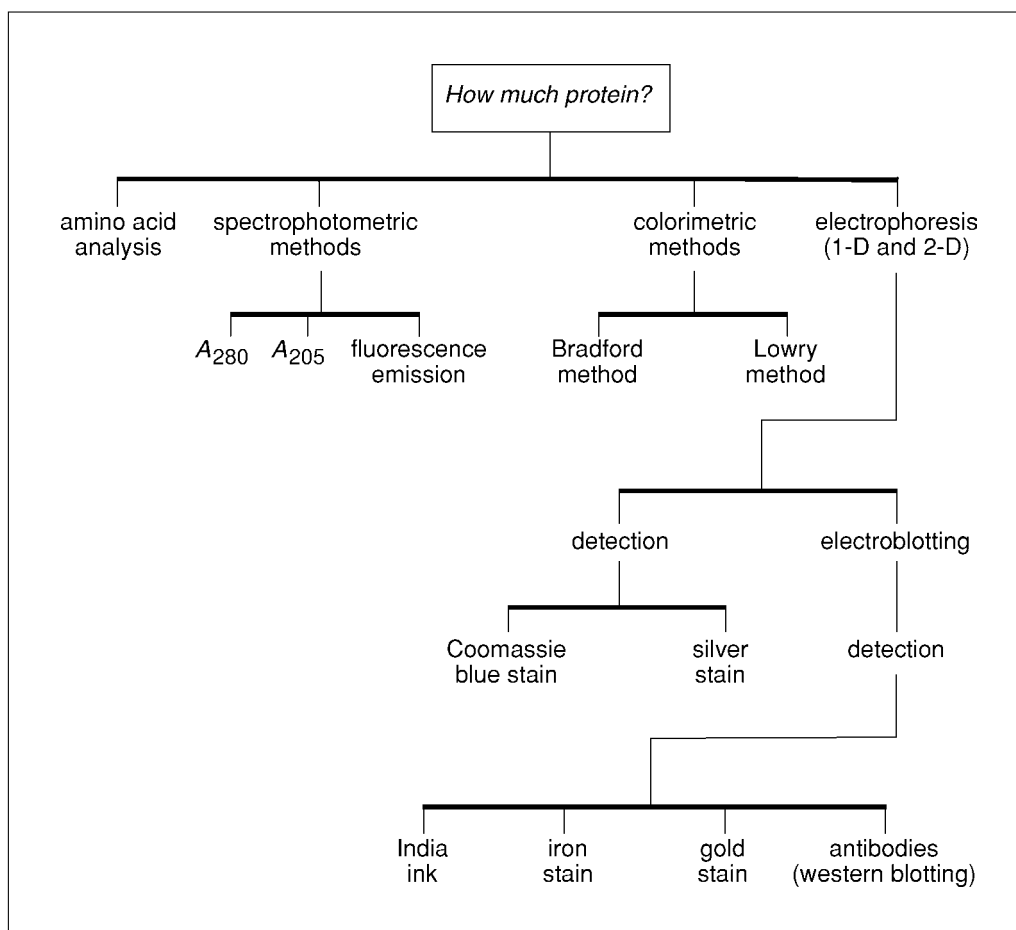


Figure 10.0.5 Quantitation of proteins.

Another method for quantifying a protein (or peptide) in a sample is the radioimmunoassay (RIA), which is presented as a solution-based assay in *UNIT 10.24*. A solution RIA determines the amount of a specific protein in an unknown sample by comparison to a standard curve developed using known amounts of the same protein. The amount of protein in the unknown sample and standards is determined based on the protein's ability to compete with radiolabeled protein (the tracer) for binding to an antibody against the protein. A secondary antibody and PEG are used to precipitate the antibody-protein complexes. Because this is a competition assay, the amount of radioactivity in the complex is inversely proportional to the amount of protein in the sample. Under ideal conditions, an RIA is capable of measuring picograms of antigen, but the tracer must be radiolabeled at a high specific activity. Radioiodine (either ^{125}I or ^{131}I) is the label of choice. The solid-phase RIA (*UNIT 11.17*) may also be used for proteins and polypeptides, but is not appropriate for peptides, since the solid-phase method requires that two antigenic sites be present in the antigen being measured. Al-

though two or more antigenic sites are likely in the case of proteins and polypeptides, they are unlikely in the case of short peptides.

The answers to the next three questions come from an analysis of the data derived from a combination of electrophoresis and chromatography—i.e., conventional gel filtration or size-exclusion high-performance liquid chromatography (HPLC; see Fig. 10.0.6). For a protein without subunits or a protein with identical subunits, detection of a single protein band after one-dimensional gel electrophoresis under denaturing conditions or a single spot after two-dimensional gel electrophoresis indicates that the protein is pure. If the protein consists of multiple subunits of different molecular sizes, purity is confirmed by detecting a single stainable band after gel electrophoresis under nondenaturing conditions. Once the protein is demonstrated to be pure, an estimate of the molecular size of the protein is made by comparing the elution volume of the protein from a conventional gel-filtration or high-performance size-exclusion column to the elution volumes of standard proteins. An estimate of the size of the subunits can

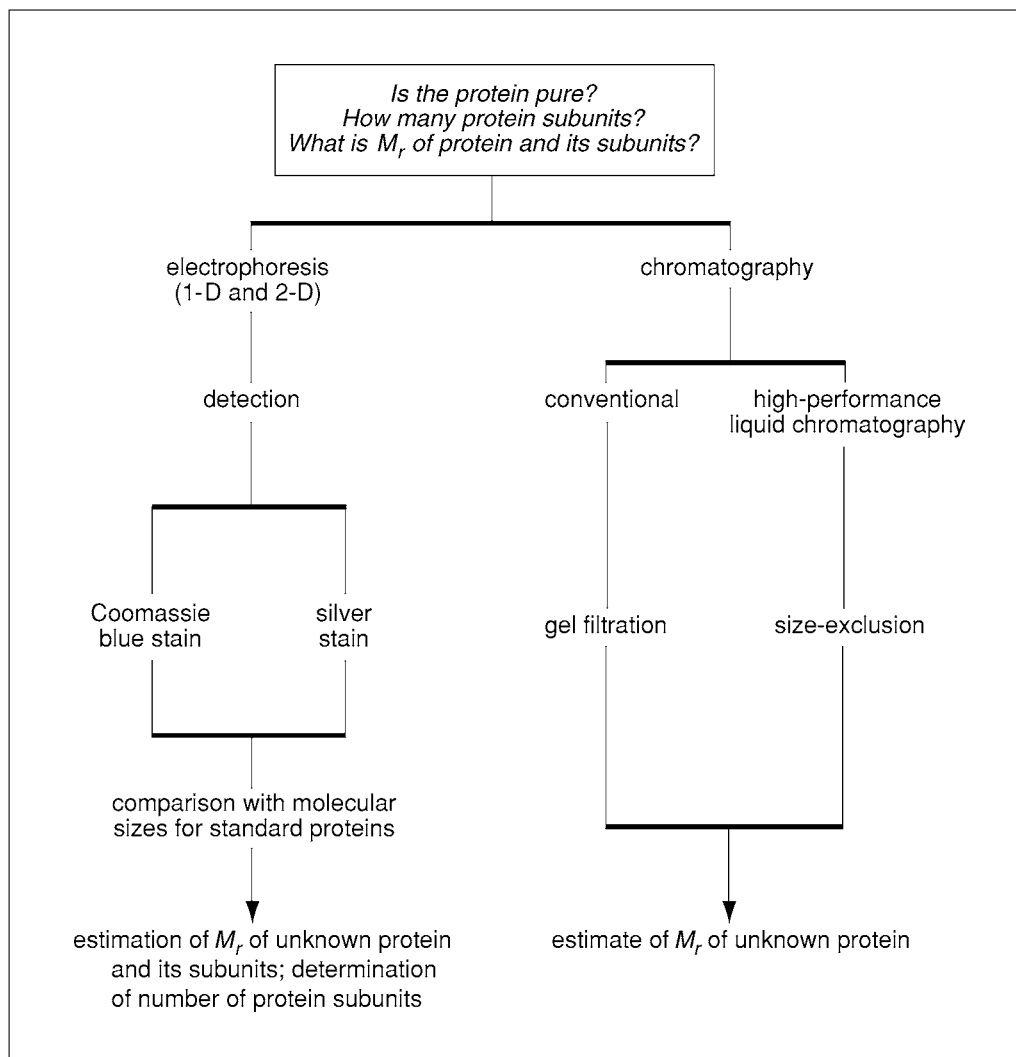


Figure 10.0.6 Analysis of protein purity, molecular weight, and subunit structure.

be determined by subsequent electrophoresis under denaturing conditions. The number of each subunit is then deduced by comparing the molecular size of “native” (i.e., nondenatured) protein and the molecular size(s) of the subunit(s).

To determine how a protein or protein fragment should be isolated, the following factors must be considered: (1) the amount of a protein in the available starting material, (2) the cost of preparing starting material (e.g., cell culture, fermentation, or organs) and the cost of labor, (3) the molecular size of the protein, and (4) the physical properties of the protein. In most cases a protein is being isolated and purified in order to ascertain partial protein sequence information by automated Edman degradation using a commercially available protein sequencer. However, almost all proteins are isolated by a combination of conventional chromatography, HPLC, and electrophoresis (Fig. 10.0.7).

Each of these isolation methods is discussed in detail in the following sections of this introduction.

Both one- and two-dimensional gel electrophoresis are high-resolution separation methods, yielding protein whose sequence can be determined after electroblotting onto polyvinylidene difluoride (PVDF) membrane filters (which are compatible with a gas-phase protein sequencer; UNIT 10.19) or capillary electrophoresis (UNITS 10.14 & 10.20). In most cases, electrophoretic methods are used after several successive modes of conventional chromatography or HPLC have been used to purify progressively a given protein from a crude protein mixture. However, if the protein is separated under denaturing conditions, the biological activity of a desired protein will likely be lost. This is a reason for utilizing gel electrophoresis last when purifying a protein whose identity is based on a functional assay.

advantages of HPLC over conventional chromatography are that small amounts (<1 nmol) of a given protein can be handled and purified, less time is required for a separation (<3 hr), and reduced chromatographic peak volume is achieved. RP-HPLC is primarily useful for purifying small proteins (molecular weight <20,000) and protein fragments (discussed below).

UNIT 10.15 describes epitope tagging, an innovative technique whereby an epitope (i.e., an antigenic site) is genetically engineered into a protein to provide a convenient tag to isolate the expressed protein from complex protein mixtures (such as cell lysates). Antibodies directed at the antigenic site interact with the tagged protein to form immune complexes, known as immunoprecipitates. Immunoprecipitation is another method of protein purification (*UNIT 10.16*) by which a specific protein can be selectively precipitated from a complex protein mixture, provided that specific antibodies directed against the protein are available. Both conventional and monoclonal antibodies may be employed for immunoprecipitation. Studies involving protein-protein interactions (*UNITS 20.1–20.6*), far western analysis (*UNIT 20.6*), chromatin assembly analysis (*UNIT 21.3*), aptamer-protein interactions (*UNITS 24.2–24.4*), and RNA-protein interactions (*UNITS 27.3 & 27.4*) are variations of these affinity approaches.

Another approach is the *in vitro* synthesis of proteins by transcription and translation of cloned genes (*UNIT 10.17*). Such *in vitro*-synthesized proteins are extremely useful for a variety of purposes including analysis of DNA-protein interactions and studies of mutant proteins obtained from mutagenesis of cloned DNA. In this approach, the protein-coding sequences are cloned into a vector containing a promoter for SP6 or T7 RNA polymerase. Messenger RNA encoding the protein is generated by transcribing the DNA template with the appropriate bacteriophage RNA polymerase. The essentially pure mRNA is then translated *in vitro* using wheat germ extracts or reticulocyte lysates. By using [³⁵S]methionine during the translation reaction, the protein is synthesized as a radiolabeled species. For most applications, *in vitro*-synthesized proteins can be used directly without any further purification. A major advantage of this method is that any desired mutant protein can be generated simply by altering the DNA template.

Biosynthetic labeling techniques are commonly used in the study of biochemical properties, processing, intracellular transport, secretion, and degradation of proteins. Methods for labeling many secreted and membrane proteins are presented *UNIT 10.18*.

Many proteins, however, are normally synthesized with the α -N₂ group blocked by an acyl (e.g., acetyl) moiety and are consequently refractory to Edman degradation. Therefore, it is wise to fragment a protein (by chemical and/or enzymatic methods) after it has been purified to homogeneity by a combination of separation methods including SDS-PAGE, and to isolate the individual protein fragments by RP-HPLC (*UNIT 10.19*). The partial sequence analysis of several peptides may then be completed. These data may be used collectively to confirm the identity of an unknown gene or to form the basis for the synthesis of oligodeoxynucleotides (i.e., primers or probes) or peptides.

UNIT 10.20 presents capillary zone electrophoresis of proteins and peptides, an extraordinarily high-resolution technique that may be used to separate a variety of proteins of differing properties. Capillary electrophoresis is ideally suited to preparation of proteins for protein sequence analysis (*UNITS 10.14 & 10.20*) and for separating proteins while the capillary is interfaced to a mass spectrometer (*UNITS 10.14 & 10.20*).

UNIT 10.21 provides an overview of peptide and protein analysis by mass spectrometry, an analytical technique that all molecular biologists who are interested in defining the primary structure, disulfide bonding, and translational modifications of proteins need to understand. *UNIT 10.22* extends the principles and techniques of the previous unit and describes state-of-the-art methods for identifying proteins by mass spectrometry using peptide-mass data analysis, peptide-sequence tags, uninterpreted fragment-ion searching, and *de novo* sequencing by tandem mass spectrometry (MS/MS). The incredible resolving power of two-dimensional gel electrophoresis (*UNITS 10.3 & 10.4*), RP-HPLC (*UNITS 10.13 & 10.14*), and capillary electrophoresis (*UNITS 10.14 & 10.20*), coupled with the analytical power of mass spectrometry, is among the drivers of functional genomic research, and direct mass spectroscopic analysis of gel-separated proteins is thoroughly described. The extent and complexity of translational modification of proteins remains poorly

understood for most proteins, and over 400 distinct chemical modifications have been described. The detection and chemical identification of such modifications is one of the frontiers of cell biology.

UNIT 10.23 describes differential gel electrophoresis (DIGE), a technique for labeling proteins with three unique cyanine fluorescent dyes (called Cy2, Cy3, and Cy5), which differ in their fluorescence spectral properties. The Cy dyes are covalently attached to the proteins in a mixture using an active ester coupling procedure (using *N*-hydroxysuccinimide ester derivatives of the various Cy dyes), which results in minimal (1% to 2%) labeling of the ϵ -amino group of lysine residues of the various proteins in a mixture. Because the net charge of a lysine residue or a Cy dye molecule is -1 , there is no appreciable change in the isoelectric point. In addition, there is no detectable change in the molecular mass of the labeled proteins, unless they are very low molecular mass, since the mass of a Cy molecule is ~ 500 . The labeled proteins from different protein samples are analyzed by 2-D electrophoresis using a multiplex analysis format (i.e., two or three samples are mixed and applied to the same gel) and the individual separation patterns (between two or among three Cy-labeled mixtures) are revealed by scanning the gels with excitation wavelengths corresponding to the maxima of the different dyes. Because the Cy dyes are detected with the sensitivity of silver-stained gels, the DIGE method is also useful for comparing the expression of recombinant proteins under different expression conditions, protein mixtures fractionated by conventional chromatography or HPLC, and membrane protein mixtures from different strains. Importantly, Cy dye-labeled proteins eluted from gels or separated chromatographically are compatible with subsequent mass spectrometric analysis.

UNIT 10.24 describes the use of solution radioimmunoassays to determine picomolar amounts of proteins in complex biological samples. The method requires the initial purification of a sufficient quantity of the unknown protein so that it may be used as an immunogen for generating polyclonal or monoclonal anti-protein antibodies, as radioiodinated protein tracer, and as control protein for constructing a standard curve. Proteins are routinely radioiodinated in trace amounts by either Bolton-Hunter reagent (labeling lysyl residues) or by lactoperoxidase (labeling tyrosine residues).

CONCLUSION

Through proteomics, structural genomics, and cell biology, the post-genomic era is addressing aspects of the protein problem that are now tractable. The currently intractable aspects will only be elucidated when new technologies are developed that move us beyond the limitations of purification and analysis methods that are the state of the art described in this chapter. Without a concerted effort to move beyond the limitations of contemporary methods, which were developed to isolate industrial quantities of proteins (beginning in the 1940s) and were premised upon proteins existing as solitary molecules with invariant function (the “one gene, one protein” hypothesis, which is largely the realm of molecular biology), a real understanding of the diversity, distribution, and functions of proteins will remain an illusory goal.

LITERATURE CITED

- Coligan, J.E., Dunn, B.M., Speicher, D.W., and Wingfield, P.T. (eds.). 2006. *Current Protocols in Protein Science*. John Wiley & Sons, Hoboken, N.J.
- Oka, M., Yang, Y.S., Nagata, S., Esaki, N., Tanaka, M., and Soda, K. 1989. Overproduction of thermostable leucine dehydrogenase of *Bacillus stearothermophilus* and its one-step purification from recombinant cells of *Escherichia coli*. *Biotechnol. Appl. Biochem.* 11:307-316.
- Pain, R.H. 1995. Overview of protein folding. *In Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 6.4.1-6.4.7. John Wiley & Sons, New York.
- Palmer, I. and Wingfield, P.T. 2004. Preparation and extraction of insoluble (inclusion-body) proteins from *Escherichia coli*. *In Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, D.W. Speicher, and P.T. Wingfield, eds.) pp. 6.3.1-6.3.18. John Wiley & Sons, Hoboken, N.J.
- Sassenfeld, H.M. 1990. Engineering proteins for purification. *Trends Biotechnol.* 8:88-93.
- Walter, H. and Johansson, G. (eds.) 1994. Aqueous two-phase systems. *Methods Enzymol.* 228:1-725.
- Wingfield, P.T. 1995. Use of protein folding reagents. *In Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, D.W. Speicher, and P.T. Wingfield, eds.) pp. A.3A.1-A.3A.4. John Wiley & Sons, New York.
- Wingfield, P.T. 2002. Overview of the purification of recombinant proteins produced in *Escherichia coli*. *In Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 6.1.1-6.1.37. John Wiley & Sons, New York.

- Wingfield, P.T. 2005. Preparation of soluble proteins from *Escherichia coli*. In Current Protocols in Protein Science (J.E. Coligan, B.M. Dunn, D.W. Speicher, and P.T. Wingfield, eds.) pp. 6.2.1-6.2.22. John Wiley & Sons, Hoboken, N.J.
- Wingfield, P.T., Palmer, I., and Liang, S.-M. 1995. Folding and purification of insoluble (inclusion-body) proteins from *Escherichia coli*. In Current Protocols in Protein Science (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 6.5.1-6.5.27. John Wiley & Sons, New York.
- Zeilstra-Ryalls, J., Fayet, O., and Georgopoulos, C. 1991. The universally-conserved GroE (Hsp60) chaperonins. *Annu. Rev. Microbiol.* 45:301-325.
- Kennedy, J.F. and Cabral, J.M. (eds.) 1993. Recovery Processes for Biological Materials. John Wiley & Sons, New York.
- A useful introduction to the problems of large-scale methods.*
- Kenny, A. and Fowell, S. (eds.) 1992. Practical protein chromatography. *Methods Mol. Biol.* 11:1-327.
- Extensive descriptions of affinity chromatographic techniques with protocols and recipes.*
- Scopes, R.K. 1993. Protein Purification, Principles and Practice, 3rd ed. Springer-Verlag, New York and Heidelberg.
- General principles of all the main techniques used in purifying proteins. A useful laboratory handbook; does not include recipes or procedures for specific proteins.*

KEY REFERENCES

- Coligan et al., 2006. See above.
Companion to this manual giving detailed protocols for many procedures commonly used in protein chemistry, including general handling, purification, and analysis.
- Deutscher, M.P. (ed.) 1990. Guide to protein purification. *Methods Enzymol.* 182:1-894.
Extensive collection of purification methods, with some general protocols and examples.
- Janson, J.-C. and Ryden, L.G. 1989. Protein Purification: Principles, High Resolution Methods, and Applications. VCH Publishers, New York.
A useful collection of methods and examples.

Contributed by R.K. Scopes
La Trobe University
Bundoora, Australia

John A. Smith
University of Alabama
Birmingham, Alabama

QUANTITATION OF PROTEINS

The success or failure of protein-centered projects can frequently be traced to the quality of the analytical procedures used to characterize the sample at different stages. Qualitative and quantitative analysis can aid in definition of the sample for the purpose of designing separations. A knowledge of the properties of the desired protein (e.g., whether it has a high aromatic amino acid content) can suggest methods of analysis that will help locate the desired protein in a complex mixture. Establishing the properties of an isolated protein creates benchmarks that future researchers can use to evaluate their protocols and final product. Accurate quantitation of the amount of protein at the beginning, middle, and end of a series of steps is the only valid way to evaluate the overall yield of a procedure. The observation of significant loss of protein, without a substantial increase in the purity of the desired protein, following a particular purification procedure would indicate that the procedure should be omitted or revised.

Several spectroscopic procedures for characterizing protein samples are described in *UNIT 10.1A*. Measuring the absorbance of the aromatic amino acids in a protein at difference wavelengths yields a very useful measure of protein concentration. This is non-destructive and requires very little sample or time. A more qualitative, but much more sensitive, evaluation is provided by fluorescence spectroscopy. Quantitation of the amount of protein contained in a solution also can be conveniently accomplished using colorimetric methods. The Bradford and the Lowry methods are the most frequently used and reliable procedures. The method of choice is the Bradford method, which is easy and rapid to complete.

Another approach is amino acid analysis—qualitative analysis to determine purity and quantitative analysis to provide concentrations—both of which are presented in *UNIT 10.1B*. Procedures are also given for calculating amino acid composition from primary analytical data. Significant advances that have improved the precision and sensitivity of amino acid analysis have reinvigorated this method, which had for some years been neglected.

Spectrophotometric and Colorimetric Determination of Protein Concentration

UNIT 10.1A

This unit describes spectrophotometric and colorimetric methods for measuring the concentration of a sample protein in solution. In Basic Protocol 1, absorbance measured at 280 nm (A_{280}) is used to calculate protein concentration by comparison with a standard curve or published absorptivity values for that protein (a_{280}). In Alternate Protocol 1, absorbance measured at 205 nm (A_{205}) is used to calculate the protein concentration. The A_{280} and A_{205} methods can be used to quantitate total protein in crude lysates and purified or partially purified protein. Both of these methods are simple and can be completed quickly. The A_{280} method is the most commonly used. The A_{205} method can detect lower concentrations of protein and is useful for dilute protein samples, but is more susceptible to interference by reagents in the protein sample than the A_{280} method. Basic Protocol 2 uses a spectrofluorometer or a filter fluorometer to measure the intrinsic fluorescence emission of a sample solution; this value is compared with the emissions from standard solutions to determine the sample concentration. The fluorescence emission method is used to quantitate purified protein. This simple method is useful for dilute protein samples and can be completed in a short amount of time. The Bradford colorimetric method, based upon binding of the dye Coomassie brilliant blue to an unknown protein, is presented in Basic Protocol 3; the Lowry method, which measures colorimetric reaction of tyrosyl residues in an unknown, is given in Alternate Protocol 2.

USING A_{280} TO DETERMINE PROTEIN CONCENTRATION

Determination of the concentration of protein in a solution by absorbance measurements at 280 nm (A_{280}) is based on the inherent absorbance of UV light by the aromatic amino acids tryptophan and tyrosine, as well as by cystine (disulfide-bonded cysteine residues). The measured absorbance of a protein sample solution is used to calculate the concentration either from its published absorptivity at 280 nm (a_{280}) or by comparison with a calibration curve prepared from measurements with standard protein solutions. This assay can be used to quantitate solutions with protein concentrations of 20 to 3000 $\mu\text{g/ml}$.

Materials

- 3 mg/ml spectrophotometric standard protein solution (see recipe; optional)
- Sample protein
- Spectrophotometer with UV lamp
- Quartz cuvette

1. For calibrating with standards, use the 3 mg/ml standard protein solution to prepare dilutions of 20, 50, 100, 250, 500, 1000, 2000, and 3000 $\mu\text{g/ml}$ in the same solvent used to prepare the sample protein. Prepare a blank consisting of solvent alone.

Ideally, for purified or partially purified protein, the protein standard should have an aromatic amino acid content similar to that of the sample protein. For the total protein of a crude lysate, an average absorptivity (step 5a) can be used to calculate the protein concentration directly from the measured A_{280} without standards. Bovine serum albumin (BSA) is commonly used as a standard for spectrophotometric quantitation of protein concentration of crude lysates. However, the a_{280} for BSA is 0.6 ml/mg, which is less than one half the calculated average a_{280} (1.3 ml/mg) from 80 proteins (Pace et al., 1995). Therefore, use of BSA as a standard for protein quantitation could result in an underestimation of protein concentration of a given sample that has a greater aromatic amino acid content than BSA.

**BASIC
PROTOCOL 1**

**Analysis of
Proteins**

10.1A.1

2. Turn on the UV lamp of the spectrophotometer and set the wavelength to 280 nm. Allow the instrument to warm up 30 min before taking measurements.

3. Zero the spectrophotometer with the solvent blank.

When removing solutions from the cuvette, do not remove the cuvette from the spectrophotometer. Leaving the cuvette in the spectrophotometer between readings will result in greater reproducibility of the absorbance readings. Use of an aspirator pump and thin tubing is the ideal method for completely removing the solutions from the cuvette. If this is not available, use a pipettor and tip to remove the solutions from the cuvette.

4. Measure the absorbance of the protein standard and unknown solutions.

If the A_{280} of the sample protein is >1.5 , dilute the sample further in the same solvent and measure the A_{280} again.

5a. *If the a_{280} of the protein is known:* Calculate the unknown sample concentration from its absorbance value using the following equation, where a_{280} has units of ml/mg cm and b is the path length in cm.

$$\text{concentration (mg/ml)} = \frac{A_{280}}{a_{280} \times b}$$

Because most cuvettes have a 1-cm path length, the b term can usually be ignored in this equation. For calculation of the protein content of crude lysates, an average a_{280} of 1.3 ml/mg cm can be used. This value was determined from absorptivities of 80 proteins (Pace et al., 1995).

5b. *If standard solutions are used for quantitation:* Create a calibration curve by either plotting or performing regression analysis of the A_{280} versus concentration of the standards. Use the absorbance of the sample protein to determine the concentration from the calibration curve.

ALTERNATE PROTOCOL 1

USING A_{205} TO DETERMINE PROTEIN CONCENTRATION

Determination of protein concentration by measurement of absorbance at 205 nm (A_{205}) is based on absorbance by the peptide bond. The concentration of a protein sample is determined from the measured absorbance and the absorptivity at 205 nm (a_{205}). This assay can be used to quantitate protein solutions with concentrations of 1 to 100 $\mu\text{g/ml}$ protein.

Additional Materials (also see Basic Protocol 1)

Brij 35 solution: 0.01% (v/v) Brij 35 (Sigma) in an aqueous solution appropriate for dissolving or diluting the sample protein

1. Dissolve or dilute the protein sample in Brij 35 solution.
2. Turn on the UV lamp of the spectrophotometer and set the wavelength to 205 nm. Allow the instrument to warm up 30 min before taking measurements.
3. Zero the spectrophotometer with the Brij 35 solution alone.
4. Measure the absorbance of the sample protein.

When removing solutions from the cuvette, do not remove the cuvette from the spectrophotometer. Leaving the cuvette in the spectrophotometer between readings will result in greater reproducibility of the absorbance readings. Use of an aspirator pump and thin tubing is the ideal method for completely removing the solutions from the cuvette. If this is not available, use a pipettor and tip to remove the solutions from the cuvette.

- 5a. *If the a_{205} of the protein is known:* Use the equation relating A_{280} and protein concentration (see Basic Protocol 1, step 5a) to calculate the concentration of the sample protein, *except* substitute the appropriate values for A_{205} and a_{205} .
- 5b. *If the a_{205} is not known:* Estimate the concentration of the sample protein from its measured absorbance using the following equation, where the absorptivity value, 31, has units of ml/mg cm and b is the path length in cm.

$$\text{concentration (mg/ml)} = \frac{A_{205}}{31 \times b}$$

The absorptivity value of 31 ml/mg cm is an average derived from measurement of ten purified proteins (Scopes, 1974). The proteins were first dried and then several dilutions were made in buffer; the a_{205} for each protein was calculated from the absorbance readings of the dilutions.

USING FLUORESCENCE EMISSION TO DETERMINE PROTEIN CONCENTRATION

BASIC PROTOCOL 2

Protein concentration can also be determined by measuring the intrinsic fluorescence based on fluorescence emission by the aromatic amino acids tryptophan, tyrosine, and/or phenylalanine. Usually tryptophan fluorescence is measured. The fluorescence intensity of the protein sample solution is measured, and the concentration of the protein sample solution calculated from a calibration curve based on the fluorescence emission of standard solutions prepared from the purified protein. This assay can be used to quantitate protein solutions with concentrations of 5 to 50 $\mu\text{g/ml}$.

Materials

Spectrophotometric protein standard solution (see recipe) prepared using the purified protein
Sample protein
Spectrofluorometer *or* filter fluorometer with an excitation cutoff filter at ≤ 285 nm and an emission filter at > 320 nm
Quartz fluorometer cuvette

1. Prepare dilutions of the purified protein at 5, 7.5, 10, 25, and 50 $\mu\text{g/ml}$ in the same solvent as the sample protein. Prepare a blank consisting of solvent alone.
2. Turn on the lamp of the instrument and allow it to warm up 30 min before taking measurements.

If a spectrofluorometer is used, set the excitation wavelength to 280 nm and the emission wavelength to between 320 and 350 nm. If the exact emission wavelength is not known, determine it empirically by scanning the standard solution with the excitation wavelength set to 280 nm. If the instrument is a filter fluorometer, use an excitation cutoff filter at ≤ 285 nm and an emission filter at > 320 nm.

3. Zero the instrument with the solvent blank.
4. Measure the fluorescence of the protein standard and sample protein solutions.
5. Create a calibration curve by either plotting or performing regression analysis of the fluorescence intensity versus concentration of the standards. Using the fluorescence intensity of the sample protein, determine the concentration from the calibration curve.

Fluorescence emission is a linear function of concentration only over a limited range.

Analysis of Proteins

10.1A.3

**USING THE BRADFORD METHOD TO DETERMINE PROTEIN
CONCENTRATION**

The Bradford method depends on quantitating the binding of a dye, Coomassie brilliant blue, to an unknown protein and comparing this binding to that of different amounts of a standard protein, usually BSA. It is designed to quantify 1 to 10 μg protein. Protein determinations in the range of 10 to 100 μg may be carried out by increasing the volume of the dye solution 5-fold and using larger tubes.

Materials

Colorimetric standard protein solution (0.5 mg/ml BSA; see recipe)
0.15 M NaCl
Coomassie brilliant blue solution (see recipe)
1 ml, 1-cm-path-length microcuvette

1. Into eight microcentrifuge tubes place duplicate aliquots of 0.5 mg/ml BSA (5, 10, 15, and 20 μl) and dilute each to 100 μl with 0.15 M NaCl. Into two more microcentrifuge tubes, place 100 μl each of 0.15 M NaCl; these are blank tubes.
2. Add 1 ml Coomassie brilliant blue solution and vortex. Allow to stand 2 min at room temperature.
3. Determine the A_{595} using a 1-cm-path-length (1 ml) microcuvette and make a standard curve by plotting absorbance at 595 nm versus protein concentration. Determine the absorbance for the unknown and use the standard curve to determine the concentration of protein in the unknown.

If the unknown protein concentration is too high, dilute the protein, assay a smaller aliquot of the unknown, or generate another standard curve in a higher concentration range (e.g., 10 to 100 μg).

**USING THE LOWRY METHOD TO DETERMINE PROTEIN
CONCENTRATION**

The Lowry method depends on quantitating the color obtained from the reaction of Folin-Ciocalteu phenol reagent with the tyrosyl residues of an unknown protein and comparing this color value to the color values derived from a standard curve of a standard protein, usually BSA. This assay is designed to quantify 1 to 20 μg protein. Protein determinations in the range of 5 to 100 μg may be carried out by increasing all the volumes 5-fold.

Additional Materials (also see Basic Protocol 3)

0.15% (w/v) sodium deoxycholate
72% (w/v) trichloroacetic acid (TCA)
Copper tartrate/carbonate (CTC) solution (see recipe)
20% (v/v) Folin-Ciocalteu reagent

1. Into eight microcentrifuge tubes place duplicate aliquots of 0.5 mg/ml BSA (10, 20, 30, and 40 μl) and dilute each to 200 μl with water. Into two more microcentrifuge tubes, place 200 μl water; these are blank tubes.
2. Add 20 μl of 0.15% sodium deoxycholate to each tube and vortex. Allow to stand 10 min at room temperature.
3. Add 20 μl of 72% TCA to each tube and vortex.
4. Microcentrifuge 15 min at $3000 \times g$, then decant supernatant (protein will be in the precipitate).
5. Dissolve protein pellets in 200 μl water. Add 200 μl water to the blank tubes.

6. Add 200 μl CTC solution, vortex, and allow to stand 10 min at room temperature.
7. Add 100 μl of 20% Folin-Ciocalteu reagent, *immediately* vortex, and allow to stand 30 min at room temperature.
8. Determine the A_{750} and compare against standard curve (see Basic Protocol 3, step 3).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Colorimetric standard protein solution (0.5 mg/ml BSA)

Prepare a 10 mg/ml solution of BSA and measure its A_{280} in a 1-cm-path-length microcuvette. The A_{280} should equal 6.61 (i.e., a 0.5 mg/ml solution will have an $A_{280} = 0.33$).

Coomassie brilliant blue solution

In a 1-liter volumetric flask, dissolve 100 mg Coomassie Brilliant Blue G-250 in 50 ml of 95% ethanol. Add 100 ml of 85% phosphoric acid. Bring to volume with water. Filter through Whatman no. 1 filter paper. Store at 4°C.

Commercial kits are available from Pierce and Bio-Rad.

Copper tartrate/carbonate (CTC) solution

Stock solution: To 50 ml of a solution of 0.2% (w/v) $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ and 0.2% (w/v) potassium tartrate, slowly add, while vigorously stirring, 50 ml of a solution of 20% sodium carbonate. This solution is stable >8 weeks at room temperature.

Working solution: Mix equal volumes of CTC stock solution, 0.8 M NaOH solution, 10% (w/v) SDS solution, and water. The reagent is stable 1 to 2 weeks at room temperature.

Spectrophotometric standard protein solution, 3 mg/ml

Weigh out dry protein and prepare a stock solution at a concentration of 3 mg/ml in the same solvent as used for the sample protein. Store up to 3 months at -20°C .

To prepare calibration standard solutions, dilute the stock solution in solvent to give the desired final concentrations for the standard curve.

Bovine serum albumin (BSA, fraction V; Sigma) is frequently used as a protein standard solution. A 3 mg/ml solution of BSA should have an A_{280} of 1.98, based on an A_{280} of 6.61 for a 1% (w/v) solution.

For quantitation of a purified or partially purified protein, the protein standard should, if possible, have an aromatic amino acid content similar to that of the sample protein.

COMMENTARY

Background Information

Measuring absorbance at 280 nm (A_{280}) is one of the oldest methods for determining protein concentration (Warburg and Christian, 1942; Layne, 1957). This method is still widely used, because it is simple and does not require incubating the sample with exogenous chromophores. However, the detection limit is higher than with colorimetric methods and therefore higher concentrations of protein are necessary. To reduce the amount of

sample required for the A_{280} method, either microcuvettes, which use 50 μl of sample, can be used or the method can be adapted to microtiter plate readers capable of UV absorbance detection. The A_{280} method requires that the protein being quantitated contain aromatic amino acids, primarily tryptophan and tyrosine. Because of the variability in aromatic amino acid content among different proteins, their absorptivity at 280 nm (a_{280}) also varies. Therefore, if calibration standards are

Table 10.1A.1 Absorption Maxima and Molar Absorptivity (ϵ) of Amino Acids^a

Amino acid	Wavelength maxima (nm)	$\epsilon \times 10^{-3}$ ($M^{-1} \text{ cm}^{-1}$)
Cysteine	250	0.3
Histidine	211	5.9
Phenylalanine	188	60.0
	206	9.3
	257	0.2
Tryptophan	219	47.0
	279	5.6
Tyrosine	193	48.0
	222	8.0
	275	1.4

^aValues are for aqueous solutions at pH 7.1 (Freifelder, 1982; Fasman, 1989).

Table 10.1A.2 Molar Absorptivity (ϵ) of Amino Acids at 280 nm^a

Amino acid	$\epsilon \times 10^{-3}$ ($M^{-1} \text{ cm}^{-1}$)
Cystine	0.110
Phenylalanine	0.0007
Tryptophan	5.559
Tyrosine	1.197

^aValues are for aqueous solutions at pH 7.1 except for cystine, which is for water (Fasman, 1989).

used for quantitation, the aromatic amino acid content of the standard must be similar to that of the sample protein for accurate results. Accordingly, the quantitation of proteins by peptide bond absorption at 205 nm (A_{205}) is more universally applicable. Furthermore, the absorptivity for a given protein at 205 nm is several-fold greater than that at 280 nm (Scopes, 1974; Stoscheck, 1990). Thus, lower concentrations of protein can be quantitated with the A_{205} method. The disadvantage of this method is that some buffers and other components absorb at 205 nm (Stoscheck, 1990).

In addition to the aromatic amino acids, several others have absorption maxima in the UV range. Table 10.1A.1 shows the wavelengths of absorption maxima and corresponding molar absorptivity (ϵ) for the amino acids with appreciable absorbance in the UV range. Only tryptophan has an absorption maximum near 280 nm, although tyrosine will absorb somewhat at that wavelength. The ϵ_{280} for tryptophan is nearly 5-fold greater than that

for tyrosine and 50-fold greater than that for cystine (Table 10.1A.2). The contribution to the absorbance at 280 nm of the third aromatic amino acid, phenylalanine, is negligible. Several amino acids other than those in Table 10.1A.1 absorb light below 205 nm (Fasman, 1989), but either the molar absorptivities are too low to be significant or the wavelengths are too short for practical absorbance measurements.

The aromatic amino acids also exhibit fluorescence emissions when excited by light in the UV range. Table 10.1A.3 gives the excitation wavelength, fluorescence emission wavelength, and quantum yield (Q) for tryptophan, tyrosine, and phenylalanine. The quantum yield is the ratio of photons emitted to photons absorbed. Typically, phenylalanine fluorescence is not detected in the presence of tyrosine and tryptophan due to low Q . Furthermore, tyrosine fluorescence is nearly completely quenched if the tyrosine residue is ionized or near an amino group, a carboxyl group, or a tryptophan residue (Teale, 1960;

Table 10.1A.3 Fluorescence Properties of Aromatic Amino Acids^a

Amino acid	Excitation wavelength	Emission wavelength	Quantum yield
Phenylalanine	260 nm	283 nm	0.04
Tryptophan	285 nm	360 nm	0.20
Tyrosine	275 nm	310 nm	0.21

^aValues are for aqueous solutions at pH 7 and 25°C (Hawkins and Honigs, 1987; Fasman, 1989).

Freifelder, 1982). Therefore, tryptophan fluorescence is what is customarily measured.

Measurement of intrinsic fluorescence by aromatic amino acids is primarily used to obtain qualitative information (Freifelder, 1982). However, with a protein standard whose aromatic amino acid content is similar to that of the sample, intrinsic fluorescence can be used for quantitation (Hawkins and Honigs, 1987). An additional consideration is that the tertiary structure of a protein will influence the fluorescence, e.g., adjacent protonated acidic groups in a protein molecule will quench tryptophan fluorescence (Freifelder, 1982).

The most frequently employed colorimetric methods for determining protein concentration are those of Bradford (1976) and Lowry et al. (1951). The Bradford method, which is faster than the Lowry method, is the method of choice for determining protein concentration. It is possible to miniaturize the Bradford assay by using a 96-well microtiter plate assay as described by Brogdon and Dickinson (1983).

The Lowry method (as modified by Peterson, 1977) is dependent on the unknown and the standard protein having a similar content of tyrosine per microgram protein. If the protein has fewer (or more) tyrosines/ μg than BSA, then the protein concentration determined will be too low (or too high). This method is not appropriate for proteins without tyrosine residues, since the assay relies on the reaction of these residues with the reagent.

Critical Parameters and Troubleshooting

A 1-cm-path-length quartz cuvette is most often used for spectrophotometric detection. However, quartz cuvettes with shorter path lengths (0.01 to 0.5 cm) are available (e.g., from Hellma Cells or Beckman); these allow higher concentrations of protein solutions to be measured. The equations relating absorbance and protein concentration (see Basic Protocol 1 and Alternate Protocol 1) assume the cuvette has a path length of 1 cm; when cuvettes of shorter path length are used, the

correct value for b must be substituted in the equation.

Glass cuvettes can be used for colorimetric detection. Quartz cuvettes are not recommended for the Bradford assay because the dye adheres more readily to quartz than to glass.

Spectrophotometric detection

The solvent pH and polarity will affect the absorbance and fluorescence properties of a protein. A notable example of pH effects on absorbance is seen with tyrosine residues, where a change in pH from neutral to alkaline results in a shift of the absorbance maximum to a longer wavelength and an increase in absorptivity due to dissociation of the tyrosine phenolic hydroxyl group (Freifelder, 1982; Fasman, 1989). An example of solvent polarity effects on fluorescence is that observed with tryptophan, where a decrease in solvent polarity results in a shift in fluorescence emission to shorter wavelengths and an increase in intensity (Freifelder, 1982). Because of these effects, the following precautions should be taken for accurate results: (1) when calibration curves are used for quantitation by absorbance or fluorescence, standards must be in the same solvent as the samples; and (2) when a published absorptivity at a given wavelength is used for quantitation, the solvent composition of the sample must be the same as that used in obtaining the published data.

Many buffers and other reagents can interfere with A_{280} and A_{205} spectrophotometric measurements. Stoscheck (1990) lists the concentration limits for many such reagents used in these spectrophotometric methods. The more commonly used reagents that absorb at 280 and 205 nm are listed in Table 10.1A.4. In addition, reagents that contain carbon-carbon or carbon-oxygen double bonds can interfere with the A_{205} method.

Because stray light can affect the linearity of absorbance versus concentration, absorbance values >2.0 should not be used for sample proteins measured by the A_{280} or A_{205} method. Samples with absorbance >2.0

Table 10.1A.4 Concentration Limits of Interfering Reagents for A_{205} and A_{280} Protein Assays^a

Reagents ^b	A_{205}	A_{280}
Ammonium sulfate	9% (w/v)	>50% (w/v)
Brij 35	1% (v/v)	1% (v/v)
DTT	0.1 mM	3 mM
EDTA	0.2 mM	30 mM
Glycerol	5% (v/v)	40% (v/v)
KCl	50 mM	100 mM
2-ME	<10 mM	10 mM
NaCl	0.6 M	>1 M
NaOH	25 mM	>1 M
Phosphate buffer	50 mM	1 M
SDS	0.10% (w/v)	0.10% (w/v)
Sucrose	0.5 M	2 M
Tris buffer	40 mM	0.5 M
Triton X-100	<0.01% (v/v)	0.02% (v/v)
TCA	<1% (w/v)	10% (w/v)
Urea	<0.1 M	>1 M

^aValues from Stoscheck (1990).

^bAbbreviations: DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; 2-ME, 2-mercaptoethanol; SDS, sodium dodecyl sulfate; TCA, trichloroacetic acid.

should be diluted further in the appropriate buffer to obtain absorbances <2.0.

Nucleic acids have substantial absorbance at 280 nm and can interfere with A_{280} quantitation of protein in crude samples. To resolve the protein concentration in such samples, measure the absorbance at 260 nm and 280 nm and calculate the protein concentration as follows (Warburg and Christian, 1942; Layne, 1957): protein concentration (mg/ml) = $1.55 \times A_{280} - 0.76 \times A_{260}$.

This estimation of protein concentration is valid up to 20% (w/v) nucleic acid or an A_{280}/A_{260} ratio <0.6.

Colorimetric detection

Many substances will interfere with the Bradford protein assay, including glycerol, detergents, 2-mercaptoethanol, acetic acid, ammonium sulfate, Tris, and certain alkaline buffers, but the appropriate controls can be used to correct this interference.

Many substances will interfere with the Lowry protein assay (Peterson, 1979). If detergents, denaturants, organic buffers, and/or thiols are present in the unknown protein solution, it is important to assess their effects on absorbance by creating standard curves

based on data taken in both the presence and the absence of these compounds. Precipitation of a protein with deoxycholate/trichloroacetic acid will eliminate many of these interfering substances while allowing a quantitative recovery of cytosolic and membranous proteins (Peterson, 1979). Although the standard curve for BSA is nonlinear, the simplicity, sensitivity, and reproducibility of the method make it worthwhile. The A_{750} will decrease at a rate of 1% to 2% per hour at room temperature.

If the unknown protein solution contains detergents such as 1% Triton X-100, 1% Tween 20, 1% Nonidet P-40, or 0.75% cetyltrimethyl-ammonium bromide, the addition of ~4% SDS will prevent precipitation of the Folin reagent (Cadman et al., 1979).

Anticipated Results

Depending on the protein, the concentration range for the A_{280} method is 20 to 3000 $\mu\text{g/ml}$, for the A_{205} method is 1 to 100 $\mu\text{g/ml}$, and for the fluorescence emission method is 5 to 50 $\mu\text{g/ml}$.

Published absorptivities of proteins at 280 nm are usually given as the absorbance for a 1% (w/v) protein solution per cm, $A^{1\%}$, or as the molar absorptivity, ϵ , which has units

of liters/mol cm (also expressed as $M^{-1} \text{ cm}^{-1}$). To convert these published coefficients to units of mg/ml, use one of the following equations:

$$\text{concentration (mg/ml)} = \frac{A_{280} \times 10}{A^{1\%} \times b} \text{ or}$$

$$\text{concentration (mg/ml)} = \frac{A_{280} \times \text{molecular weight}}{\epsilon_{280} \times b}$$

Depending on the reagent volumes used, it is possible to quantitate protein in the range of 1 to 100 μg using the Bradford or Lowry assay.

Time Considerations

When the absorptivity for a protein is known, the A_{280} and A_{205} measurements require <30 min depending on the number of samples. When standards are used for quantitation with these assays or for intrinsic fluorescence quantitation, 1 hr is required.

The Bradford method requires ~30 min. The Lowry method requires 2 hr to generate a standard curve and to run several unknown protein samples.

Literature Cited

- Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* 72:248-254.
- Brogdon, W.G. and Dickinson, C.M. 1983. A microassay system for measuring esterase activity and protein concentration in small samples and in high-pressure liquid chromatography eluate fractions. *Anal. Biochem.* 131:499-503.
- Cadman, E., Bostwick, J.R., and Eichberg, J. 1979. Determination of protein by a modified Lowry procedure in the presence of some commonly used detergents. *Anal. Biochem.* 96:21-23.
- Fasman, G.D. 1989. *Practical Handbook of Biochemistry and Molecular Biology*. CRC Press, Boca Raton, Fla.
- Freifelder, D. 1982. *Physical Biochemistry: Applications to Biochemistry and Molecular Biology* 2nd ed. W.H. Freeman, New York.
- Hawkins, B.K. and Honigs, D.E. 1987. A comparison of spectroscopic techniques for protein quantification in aqueous solutions. *Am. Biotechnol. Lab.* 5:26-37.
- Layne, E. 1957. Spectrophotometric and turbidimetric methods for measuring proteins. *Methods Enzymol.* 3:447-454.
- Lowry, O.H., Rosebrough, N.J., Farr, A.L., and Randall, R.J. 1951. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* 193:265-275.

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. 1995. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* 4:2411-2423.

Peterson, G.L. 1977. A simplification of the protein assay method of Lowry et al. which is more generally applicable. *Anal. Biochem.* 83:346-356.

Peterson, G.L. 1979. Review of the Folin phenol protein quantitation method of Lowry, Rosebrough, Farr and Randall. *Anal. Biochem.* 100:201-220.

Scopes, R.K. 1974. Measurement of protein by spectrophotometry at 205 nm. *Anal. Biochem.* 59:277-282.

Stoscheck, C.M. 1990. Quantitation of protein. *Methods Enzymol.* 182:50-68.

Teale, F.W.J. 1960. The ultraviolet fluorescence of proteins in neutral solutions. *Biochem. J.* 76:381-388.

Warburg, O. and Christian, W. 1942. Isolierung und Kristallisation des Gärungsferments Enolase. *Biochem. Z.* 310:384-421.

Key References

Chen, R.F. 1990. Fluorescence of proteins and peptides. In *Practical Fluorescence*, 2nd ed. (G.G. Guilbault, ed.) pp. 575-682. Marcel Dekker, Inc., New York.

Detailed discussion of intrinsic fluorescence of proteins and what factors affect fluorescence emission by the aromatic amino acids (see pp. 618-663).

Darbre, A. 1986. Analytical methods. In *Practical Protein Chemistry: A Handbook* (A. Darbre, ed.) pp. 227-335. John Wiley & Sons, New York.

Describes colorimetric methods for protein determination including the Lowry method and several modifications, the biuret method and several modifications, and the Bradford method (see pp. 284-295).

Fasman, 1989. See above.

Contains tables with absorptivities for UV spectrophotometric detection and tables with data on excitation and emission wavelengths for fluorescence detection of many proteins. Also includes a table with molecular weights for many characterized proteins.

Stoscheck, 1990. See above.

Contains list of substances that can interfere with 205- and 280-nm spectrophotometric and colorimetric measurements of proteins and of concentration limits for these substances.

Contributed by Michael H. Simonian
(spectrophotometric methods)

Beckman Coulter, Inc.
Fullerton, California

John A. Smith (colorimetric methods)
University of Alabama at Birmingham
Birmingham, Alabama

Quantitative Amino Acid Analysis

UNIT 10.1B

The use of quantitative amino acid analysis in the characterization of protein samples is often overlooked in many laboratories. This is typically due to the difficulties involved in obtaining reliable amino acid composition data, as well as to the advent of colorimetric dye-binding assays as alternative methods to determine protein concentration. However, many proteins do not bind dyes in the same way as the proteins normally used as standards, making quantitation by those techniques difficult. Fortunately, at many institutions, the creation of protein chemistry core laboratories devoted mainly to protein sequence analysis and peptide synthesis has provided new instrumentation that yields the necessary precision. This unit describes the information that can be derived from quantitative amino acid analysis, presents some details on sample preparation, and gives examples of the calculations that are needed.

INFORMATION OBTAINED FROM AMINO ACID ANALYSIS

The benefits of amino acid composition data are two-fold. First, identifying the nature of the amino acids present in a sample can provide an important qualitative measure of the purity of the preparation. Amino acid analysis does not depend on the shape, charge, or function of a protein because the hydrolysis completely destroys structure and reduces the protein to its constituent amino acids. Although large proteins typically contain a mixture of all the commonly occurring amino acids, certain residues are relatively rare: Trp, Met, His, Cys, and occasionally Arg are present at lower than average ratios compared to Gly, Ala, Ser, and Leu. For a pure protein, the ratios of the amino acids should approach the theoretical levels to within ~10%, with a few caveats. For a sample from a column fraction, for example, the presence of a higher than expected ratio of a certain amino acid is a very good indication that the sample is not pure and that further fractionation is indicated. In the same way, a lower than expected ratio could also indicate impurity—for example, the presence of a second protein that has a lower ratio of some significant amino acid present in higher amounts in the protein of interest. Because the expression of a protein in bacteria or other hosts utilizes a gene of known sequence, the expected amino acid composition is also known (see Chapter 16 for expression and purification of recombinant proteins).

Second, averaged composition yields an excellent measure of the amount of protein in a sample. On the ion-exchange-based Beckman 6300 amino acid analyzer running with ninhydrin detection chemistry, the minimum amount of the standard (a mixture of amino acids used to derive integration parameters as well as elution times) that is required is typically 1 nmol of each amino acid. Amounts as low as 250 pmol (0.25 nmol) of an amino acid can be measured with confidence from protein hydrolyzates. For a protein of 50 kDa mass, 1 nmol is 50 μ g, an amount that can be obtained readily by preparation of the protein from a high-expression system followed by purification by standard methods. For isolation of a trace protein from a natural source, this required amount may be too large to be feasible. However, a protein of 50 kDa mass contains ~400 amino acids, or roughly 20 copies of each amino acid on average. Thus, the actual amount of protein required for detection and semiquantitative analysis could be as low as 0.05 nmol, or 2.5 μ g. At that level, rare amino acids cannot be accurately quantitated, but the more common ones can be measured easily. Obviously, the more material available, the more accurate the analysis. Users should check with their local facility for specific advice on the amount of sample that can be processed by existing technologies.

The advent of advanced instrumentation, such as the Applied Biosystems 420 amino acid analyzer and the Waters PicoTag system, has provided another increment in sensitivity. These instruments utilize precolumn derivatization followed by separation based on reversed-phase high-performance liquid chromatography (HPLC). With such analyzers, 100 pmol of each amino acid, or 5 pmol of the 50-kDa protein described above, can be quantitated accurately.

Another important consideration in the design of a purification scheme that can be elucidated through the use of amino acid quantitation is the compilation of the mass balance—the relationship between how much protein is applied to the chromatography column and how much is recovered in a particular step. Amino acid analysis of a measured aliquot of the sample taken before application (typically $\leq 1\%$), along with aliquots of column fractions or pools, can be used to calculate the amounts of protein loaded and recovered. These calculations can provide essential information to as-

Contributed by Ben Dunn

Current Protocols in Molecular Biology (1996) 10.1.11-10.1.13
Copyright © 2000 by John Wiley & Sons, Inc.

Analysis of
Proteins

10.1.11

Supplement 35

sess the success of a particular purification procedure.

SAMPLE PREPARATION

The success or failure of quantitative amino acid analysis is strongly dependent upon the quality of the sample, as well as upon the exact composition of the sample matrix (i.e., the aggregate of all components including salts, detergent, and dirt). The following section gives a few key recommendations for successful sample preparation.

Sample Concentration

An ideal sample for amino acid analysis would contain 1 to 10 nmol protein in a volume of 10 to 100 μ l water. A sample that has been dialyzed against a low concentration of buffer (if needed to maintain solubility) or pure water is optimal, because the presence of salt, detergent, high concentrations of buffers, or other agents may interfere with either hydrolysis or subsequent analysis.

A sample that has been lyophilized to dryness is also usually acceptable, with some important considerations. First, lyophilization should be done in such a way as to concentrate the lyophilized residue at the bottom of a conical tube. Drying a sample in a round-bottom flask may leave the protein sample spread over a large surface area; such samples are extremely difficult to handle and will not give satisfactory results. Second, lyophilization of a buffered solution will concentrate the salts as well as the protein. Therefore, it is recommended that the total ionic strength of the solution be $<50 \mu$ M before lyophilization.

Estimate of Total Protein

An estimate of the total amount of protein present is very helpful in planning an accurate analysis. This can be based on a colorimetric dye-binding assay, an A_{280} measurement, or an activity assay in which the results are compared to a standard value.

All information about the sample should be communicated to the individual charged with conducting the analysis to ensure the maximum return on the sample. Many facilities provide a form with a series of questions to be answered concerning the sample.

Samples from PVDF Blots

With the advent of more sensitive analyzers based on phenylthiohydantoin (PTH) chemistry, it is now possible to obtain compositional data on samples separated by electrophoresis

and blotted onto membranes (UNITS 10.8 & 10.19) such as are used for protein sequence analysis. This is especially helpful in cases where a protein sample fails to yield any sequence during gas-phase Edman degradation. Acid hydrolysis and analysis of sample blot pieces recovered from the sequencer sample compartment yield a rough estimate of the amount of protein present. If the amount is larger than that normally required for sequence analysis (10 pmol), it may then be concluded that the protein sample was N-terminally blocked.

CALCULATION OF THE AVERAGE COMPOSITION

Table 10.1B.1 presents the results of analysis of a 24-kDa recombinant protein. The initial sample volume was 1.0 ml and it was expected to contain ~ 4 mg total protein. A 30- μ l aliquot was taken for analysis and was estimated to contain 120 μ g, or ~ 5 nmol. The aliquot was transferred to a glass tube for hydrolysis and dried by application of a vacuum at room temperature. One milliliter of 6 M HCl was added, the tube was cooled in a dry ice bath, then a vacuum was applied and the tube sealed by heating in a gas flame. Hydrolysis was conducted for 24 hr at 124°C, followed by removal of the HCl under vacuum. The sample was taken up in 1.0 ml application buffer, 50 nmol norleucine was added as an internal standard, and a 50- μ l aliquot was analyzed by ion-exchange chromatography on a Beckman 6300 system with ninhydrin detection.

1. The total number of nanomoles of amino acids detected is first calculated by summing the second column of Table 10.1B.1, yielding 1446 nmol.

2. The total nanomoles is then divided by the total number of expected amino acids (based on protein sequence; in this case, 212) to obtain the average number of nanomoles per residue; $1446 \text{ nanomoles} / 212 \text{ residues} = 6.82 \text{ nmol per residue}$. If the total number of amino acids is not known, it may be estimated by dividing the mol. wt. of the protein by 110.

3. The number of nanomoles of each amino acid measured is divided by the average number of nanomoles per residue (step 2) to give the ratio¹ observed for that amino acid. For Asp, $155 \div 6.82 = 22.7$.

The resulting ratios are given in the fourth column and may be compared to the ratio expected listed in the third column. For most of the amino acids, the ratio¹ observed agrees with the ratio expected.

The value of average nanomoles per residue

Table 10.1B.1 Data Used to Calculate Amino Acid Composition for a 24-kDa Protein

Amino acid	Nanomoles measured	Ratio expected ^a	Ratio ¹ observed	Ratio ² observed
Asp	155	23	22.7	21.8
Thr	78	13	11.4	10.9
Ser	63	13	9.2	8.9
Glu	153	21	22.4	21.5
Pro	52	7	7.6	7.3
Gly	157	22	23.0	22.1
Ala	90	12	13.1	12.6
Val	159	23	23.3	22.3
Met	24	5	3.5	3.4
Ile	88	13	12.9	12.3
Leu	138	19	20.2	19.4
Tyr	43	6	6.3	6.1
Phe	56	8	8.2	7.9
His	16	3	2.3	2.3
Lys	121	17	17.7	17.0
Arg	53	7	7.8	7.4

^aRatio expected based on amino acid sequence of protein.

calculated in step 2 is based on the analysis of a 30- μ l sample; the total sample of 1 ml contains ~5.46 mg of protein (average moles/residue \times grams/mole protein \times 1000/30). This analysis utilized only 50 μ l of the 1.0 ml final solution, so a sample size smaller than 3% of the original sample could have been used. Using an excess, however, provides the flexibility to repeat the determination if the first analysis gives peaks that are either too large or too small for accurate integration.

The ratio¹ observed for Ser, Thr, His, and Met are in this sample are ~30%, 13%, 24%, and 30% below the expected values, respectively. These results are typical for these four amino acids, but for different reasons. The Ser and Thr results are low because these amino acids undergo partial degradation during acid hydrolysis. Ser and Thr losses of 10% to 40% are normal. The lower than expected values for His and Met are an anomaly that is most likely caused by the relatively low copy number of those two amino acids in the protein. Due to these known problems, most investigators choose to delete the values for sensitive residues (Ser and Thr) as well as the values for those amino acids appearing in low amounts (in this case, His and Met) and calculate the average number of nanomoles per residue independently of those amino acids. This gives the

corrected ratios shown in the last column of Table 10.1B.1 (ratio² observed). In this example, the recalculated values are changed by as much as 0.9 due to the higher average number of nanomoles per residue (7.11).

The deviation of ratio¹ observed from the expected ratio ranges from 1% to 32% for the various amino acids, with an average error of 8.8%. When the high variation for Ser (32%), Met (32%), and His (24%) removed from consideration, the average deviation of ratio² observed from expected ratio drops to 4.1%. By standard criteria, this sample would be judged to be pure; of course, this result needs to be combined with electrophoretic separation (*UNIT 10.2*), chromatographic purification (*UNITS 10.9-10.11B*), and perhaps mass spectroscopic analyses before a final assessment of the purity of the sample can be made.

Key Reference

Juris, O. 1990. Amino acid analysis. *Methods Enzymol.* 182:587-601.

Excellent discussion of sample preparation issues and hydrolysis methods.

Contributed by Ben Dunn
University of Florida
Gainesville, Florida

ELECTROPHORETIC SEPARATION OF PROTEINS

Protein separations in vertical slab gels are performed in a variety of formats. Typically, the separating gel size is in the range of 14 × 14-cm. Smaller formats (6 × 8-cm) are also popular and have been termed minigels. Due to the limited separation area, minigels do not offer the same level of resolution when compared to larger-format gels. However, any procedure performed in the standard format works equally well with minigels, and minigels use less reagent and separate proteins much faster.

One-dimensional gel electrophoresis of proteins can provide information about the molecular size and purity of proteins, as well as the number and molecular size of subunits. By analyzing proteins before and after deglycosylation, it is possible to learn about the carbohydrate content. Information about sulfate, phosphate, and carbohydrate content can be determined for proteins analyzed after various radiolabeling procedures. Furthermore, separated proteins can be recovered from polyacrylamide gels by electroelution (*UNIT 10.5*) for subsequent studies.

The purity and molecular size of a peptide or protein can be assessed by analytical separation in the presence of 0.1% SDS on one-dimensional gels (*UNIT 10.2*) or two-dimensional gels (*UNITS 10.3 & 10.4*). By adding a reducing agent (e.g., 2-mercaptoethanol or dithiothreitol) to the sample, it is possible to determine the number and size of subunits in a pure protein. Nondenaturing gels can be used to examine and isolate the “native” protein.

Two-dimensional gel electrophoresis separates proteins in the first dimension by isoelectric focusing and in the second dimension by electrophoresis in the presence of SDS. By separating proteins in this manner, information is obtained not only about size, as in one-dimensional gels, but also about the charge of a protein. Two-dimensional gels are superior for resolving complex mixtures and for assessing protein purity.

Both one- and two-dimensional gel electrophoreses are high-resolution separation methods that yield protein whose sequence can be determined after either electroelution (*UNIT 10.5*) or electroblotting onto polybrene-coated, derivatized glass fiber sheets or polyvinylidene difluoride (PVDF) membrane filters; all are compatible with gas-phase protein sequencers). In many cases, electrophoretic methods are used after successive modes of conventional types of column chromatography (*UNITS 10.9-10.11*) or HPLC (*UNITS 10.12-10.15*) have been used to increasingly purify a given protein from a crude protein mixture. If the protein is separated under denaturing conditions, any biological activity will likely be lost and, therefore, gel electrophoresis should definitely be the last purification step for a protein whose identity is based on a functional assay.

One-Dimensional SDS Gel Electrophoresis of Proteins

Electrophoresis is used to separate complex mixtures of proteins, (e.g., from cells, subcellular fractions, column fractions, or immunoprecipitates) to investigate subunit compositions, and to verify homogeneity of protein samples. It can also serve to purify proteins for use in further applications. In polyacrylamide gel electrophoresis, proteins migrate in response to an electrical field through pores in the gel matrix; pore size decreases with higher acrylamide concentrations. The combination of gel pore size and protein charge, size, and shape determines the migration rate of the protein.

The standard Laemmli method (see Basic Protocol 1) is used for discontinuous gel electrophoresis under denaturing conditions, that is, in the presence of sodium dodecyl sulfate (SDS). The standard method for full-size gels (e.g., 14 × 14 cm) can be adapted for the minigel format (e.g., 7.3 × 8.3 cm; see Basic Protocol 2). Minigels provide rapid separation but give lower resolution.

Several alternate protocols are provided for specific applications. The first two alternate protocols cover electrophoresis of peptides and small proteins, separations that require modification of standard buffers: either a Tris-tricine buffer system (see Alternate Protocol 1), or modified Tris buffer in the absence of urea (see Alternate Protocol 2). Continuous SDS-PAGE is a simplified method in which the same buffer is used for both gel and electrode solutions and the stacking gel is omitted (see Alternate Protocol 3). Other protocols cover the preparation and electrophoresis of various types of gels: ultrathin gels (see Alternate Protocol 4), multiple single-concentration gels (see Support Protocol 1), gradient gels (see Alternate Protocol 5), multiple gradient gels (see Support Protocol 2), and multiple gradient minigels (see Support Protocol 3). Proteins separated on gels can be subsequently analyzed by immunoblotting (UNIT 10.8), autoradiography or phosphor imaging (APPENDIX 3A), or staining with protein dyes (UNIT 10.6).

CAUTION: Before any protocols are used, it is extremely important to read the following section about electricity and electrophoresis.

ELECTRICITY AND ELECTROPHORESIS

Many researchers are poorly informed concerning the electrical parameters of running a gel. It is important to note that the voltages and currents used during electrophoresis are dangerous and potentially lethal. Thus, safety should be an overriding concern. A working knowledge of electricity is an asset in determining what conditions to use and in troubleshooting the electrophoretic separation, if necessary. For example, an unusually high or low voltage for a given current (milliamperes) might indicate an improperly made buffer or an electrical leak in the chamber.

Safety Considerations

1. Never remove or insert high-voltage leads unless the power supply voltage is turned down to zero and the power supply is turned off. Always grasp high-voltage leads one at a time with one hand only. Never insert or remove high-voltage leads with both hands. This can shunt potentially lethal electricity through the chest and heart should electrical contact be made between a hand and a bare wire. On older or homemade instruments, the banana plugs may not be shielded and can still be connected to the

power supply at the same time they make contact with a hand. Carefully inspect all cables and connections and replace frayed or exposed wires immediately.

2. Always start with the power supply turned off. Have the power supply controls turned all the way down to zero. Then hook up the gel apparatus: generally, connect the red high-voltage lead to the red outlet and the black high-voltage lead to the black outlet. Turn the power supply on with the controls set at zero and the high-voltage leads connected. Then, turn up the voltage, current, or power to the desired level. Reverse the process when the power supply is turned off: i.e., to disconnect the gel, turn the power supply down to zero, wait for the meters to read zero, turn off the power supply, and then disconnect the gel apparatus one lead at a time.

CAUTION: If the gel is first disconnected and then the power supply turned off, a considerable amount of electrical charge is stored internally. The charge will stay in the power supply over a long time. This will discharge through the outlets even though the power supply is turned off and can deliver an electrical shock.

Ohm's Law and Electrophoresis

Understanding how a gel apparatus is connected to the power supply requires a basic understanding of Ohm's law: voltage = current \times resistance, or $V = IR$. A gel can be viewed as a resistor and the power supply as the voltage and current source. Most power supplies deliver constant current or constant voltage. Some will also deliver constant power: power = voltage \times current, or $VI = I^2R$. The discussion below focuses on constant current because this is the most common mode in vertical SDS-PAGE.

Most modern commercial equipment is color-coded so that the red or positive terminal of the power supply can simply be connected to the red lead of the gel apparatus, which goes to the lower buffer chamber. The black lead is connected to the black or negative terminal and goes to the upper buffer chamber. This configuration is designed to work with vertical slab gel electrophoreses in which negatively charged proteins or nucleic acids move to the positive electrode in the lower buffer chamber (an anionic system).

When a single gel is attached to a power supply, the negative charges flow from the negative cathode (black) terminal into the upper buffer chamber, through the gel, and into the lower buffer chamber. The lower buffer chamber is connected to the positive anode (red) terminal to complete the circuit. Thus, negatively charged molecules, such as SDS-coated proteins and nucleic acids, move from the negative cathode attached to the upper buffer chamber toward the positive anode attached to the lower chamber. SDS-PAGE is an anionic system because of the negatively charged SDS.

Occasionally, proteins are separated in cationic systems. In these gels, the proteins are positively charged because of the very low pH of the gel buffers (e.g., acetic acid/urea gels for histone separations) or the presence of a cationic detergent (e.g., cetyltrimethylammonium bromide, CTAB). Proteins move toward the negative electrode (cathode) in cationic gel systems, and the polarity is reversed compared to SDS-PAGE: the red lead from the lower buffer chamber is attached to the black outlet of the power supply, and the black lead from the upper buffer chamber is attached to the red outlet of the power supply.

Most SDS-PAGE separations are performed under constant current (consult instructions from the manufacturer to set the power supply for constant current operation). The resistance of the gel will increase during SDS-PAGE in the standard Laemmli system. If the current is constant, then the voltage will increase during the run as the resistance goes up.

Power supplies usually have more than one pair of outlets. The pairs are connected in parallel with one another internally. If more than one gel is connected directly to the outlets

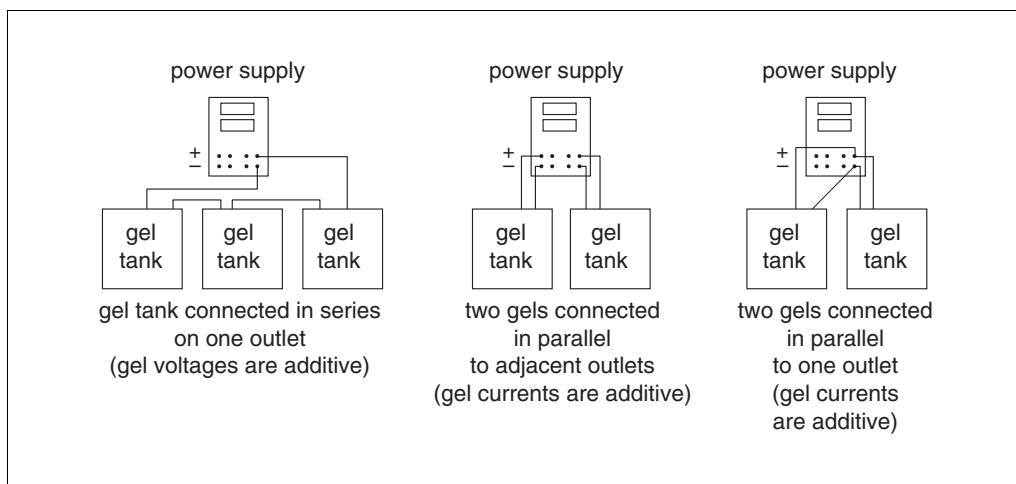


Figure 10.2A.1 Series and parallel connections of gel tanks to power supply.

of a power supply, then these gels are connected in parallel. In a parallel circuit, the voltage is the same across each gel. In other words, if the power supply reads 100 V, then each gel has 100 V across its electrodes. The total current, however, is the sum of the individual currents going through each gel. Therefore, under constant current it is necessary to increase the current for each additional gel that is connected to the power supply. Two identical gels require double the current to achieve the same starting voltages and electrophoresis separation times.

Multiple gel apparatuses can also be connected to one pair of outlets on a power supply. This is useful with older power supplies that have a limited number of outlets. When connecting several gel units to one outlet, make certain the connections between the units are shielded and protected from moisture. The gels can be connected in parallel or in series (Fig. 10.2A.1). In the case of two or more gels running off the same outlet in series, the current is the same for every gel. If 10 mA is displayed by the power supply meter, for example, each gel in series will experience 10 mA. The voltage, however, is additive for each gel. If one gel at a constant 10 mA produces 100 V, then two identical gels in series will produce 200 V (100 V each) and so on. Thus, the voltage can limit the number of units connected in series on low-voltage power supplies.

Gel thickness affects the above relationships. A 1.5-mm gel can be thought of as consisting of two 0.75-mm-thick gels run in parallel. Because currents are additive in parallel circuits, a 0.75-mm gel will require half the current of the 1.5-mm gel to achieve the same starting voltage and separation time. If a gel thickness is doubled, then the current must also be doubled. There are limits to the amount of current that can be applied. Thicker gels require more current, generating more heat that must be dissipated. Unless temperature control is available in the gel unit, a thick gel should be run more slowly than a thin gel.

NOTE: Milli-Q-purified water or equivalent should be used throughout the protocols.

**DENATURING (SDS) DISCONTINUOUS GEL ELECTROPHORESIS:
LAEMMLI GEL METHOD**

One-dimensional gel electrophoresis under denaturing conditions (i.e., in the presence of 0.1% SDS) separates proteins based on molecular size as they move through a polyacrylamide gel matrix toward the anode. The polyacrylamide gel is cast as a separating gel (sometimes called resolving or running gel) topped by a stacking gel and secured in an electrophoresis apparatus. After sample proteins are solubilized by boiling in the presence of SDS, an aliquot of the protein solution is applied to a gel lane, and the individual proteins are separated electrophoretically. 2-Mercaptoethanol (2-ME) or dithiothreitol (DTT) is added during solubilization to reduce disulfide bonds.

This protocol is designed for a vertical slab gel with a maximum size of 0.75 mm × 14 cm × 14 cm. For thicker gels, or minigels (see Basic Protocol 2 and Support Protocol 3), the volumes of stacking and separating gels and the operating current must be adjusted. Additional protocols describe the preparation of ultrathin gels (see Alternate Protocol 4) and gradient gels (see Alternate Protocol 5), as well as the use of gel casters to make multiple gels, both single-concentration gels (see Support Protocol 1) and gradient gels (see Support Protocol 2).

Materials

- Separating and stacking gel solutions (Table 10.2A.1)
- H₂O-saturated isobutyl alcohol
- 1× Tris·Cl/SDS, pH 8.8 (dilute 4× Tris·Cl/SDS, pH 8.8; Table 10.2A.1)
- Protein sample to be analyzed
- 2× and 1× SDS sample buffer (see recipe)
- Protein molecular-weight-standards mixture (Table 10.2A.2)
- 6× SDS sample buffer (see recipe; optional)
- 1× SDS electrophoresis buffer (see recipe)

- Electrophoresis apparatus: Protean II 16-cm cell (Bio-Rad) or SE 600/400 16-cm unit (Amersham Pharmacia Biotech) with clamps, glass plates, casting stand, and buffer chambers
- 0.75-mm spacers
- 0.45-μm filters (used in stock solution preparation)
- 25-ml Erlenmeyer side-arm flask
- Vacuum pump with cold trap
- 0.75-mm Teflon comb with 1, 3, 5, 10, 15, or 20 teeth
- 25- or 100-μl syringe with flat-tipped needle
- Constant-current power supply (see Electricity and Electrophoresis, above)

Pour the separating gel

1. Assemble the glass-plate sandwich of the electrophoresis apparatus according to manufacturer's instructions using two clean glass plates and two 0.75-mm spacers.
If needed, clean the glass plates in liquid Alconox or RBS-35 (Pierce). These aqueous-based solutions are compatible with silver and Coomassie blue staining procedures.
2. Lock the sandwich to the casting stand.
3. Prepare the separating gel solution as directed in Table 10.2A.1, degassing using a rubber-stoppered 25-ml Erlenmeyer side-arm flask connected with vacuum tubing to a vacuum pump with a cold trap. After adding the specified amount of 10% ammonium persulfate and TEMED to the degassed solution, stir gently to mix.

Table 10.2A.1 was prepared as a convenient summary to aid in the preparation of separating and stacking gels. The stacking gel is the same regardless of the separating gel used.

The desired percentage of acrylamide in the separating gel depends on the molecular size of the protein being separated. Generally, use 5% gels for SDS-denatured proteins of 60 to 200 kDa, 10% gels for SDS-denatured proteins of 16 to 70 kDa, and 15% gels for SDS-denatured proteins of 12 to 45 kDa (Table 10.2A.1).

4. Using a Pasteur pipet, apply the separating gel solution to the sandwich along an edge of one of the spacers until the height of the solution between the glass plates is ~11 cm.

Use the solution immediately; otherwise it will polymerize in the flask.

Sample volumes <10 μ l do not require a stacking gel. In this case, cast the resolving gel as you normally would, but extend the resolving gel into the comb to form the well. The proteins are then separated under the same conditions as used when a stacking gel is present. Although this protocol works well with single-concentration gels, a gradient gel is recommended for maximum resolution (see Alternate Protocol 5).

5. Using another Pasteur pipet, slowly cover the top of the gel with a layer (~1 cm thick) of H₂O-saturated isobutyl alcohol, by gently layering the isobutyl alcohol against the edge of one and then the other of the spacers.

Be careful not to disturb the gel surface. The overlay provides a barrier to oxygen, which inhibits polymerization, and allows a flat interface to form during gel formation.

The H₂O-saturated isobutyl alcohol is prepared by shaking isobutyl alcohol and H₂O in a separatory funnel. The aqueous (lower) phase is removed. This procedure is repeated several times. The final upper phase is H₂O-saturated isobutyl alcohol.

6. Allow the gel to polymerize 30 to 60 min at room temperature.

A sharp optical discontinuity at the overlay/gel interface will be visible on polymerization. Failure to form a firm gel usually indicates a problem with the ammonium persulfate, TEMED (N, N, N', N'-tetramethylethylenediamine), or both. Ammonium persulfate solution should be made fresh before use. Ammonium persulfate should "crackle" when added to the water. If not, fresh ammonium persulfate should be purchased. Purchase TEMED in small bottles so, if necessary, a new previously unopened source can be tried.

Pour the stacking gel

7. Pour off the layer of H₂O-saturated isobutyl alcohol and rinse with 1 \times Tris-Cl/SDS, pH 8.8.

Residual isobutyl alcohol can reduce resolution of the protein bands; therefore, it must be completely removed. The isobutyl alcohol overlay should not be left on the gel longer than 2 hr.

8. Prepare the stacking gel solution as directed in Table 10.2A.1.

Use the solution immediately to keep it from polymerizing in the flask.

9. Using a Pasteur pipet, slowly allow the stacking gel solution to trickle into the center of the sandwich along an edge of one of the spacers until the height of the solution in the sandwich is ~1 cm from the top of the plates.

Be careful not to introduce air bubbles into the stacking gel.

10. Insert a 0.75-mm Teflon comb into the layer of stacking gel solution. If necessary, add additional stacking gel to fill the spaces in the comb completely.

Again, be careful not to trap air bubbles in the tooth edges of the comb; they will cause small circular depressions in the well after polymerization that will lead to distortion in the protein bands during separation.

11. Allow the stacking gel solution to polymerize 30 to 45 min at room temperature.

A sharp optical discontinuity will be visible around wells on polymerization.

Table 10.2A.1 Recipes for Polyacrylamide Separating and Stacking Gels^a**SEPARATING GEL**

Stock solution ^b	Final acrylamide concentration in separating gel (%) ^c									
	5	6	7	7.5	8	9	10	12	13	15
30% acrylamide/ 0.8% bisacrylamide	2.50	3.00	3.50	3.75	4.00	4.50	5.00	6.00	6.50	7.50
4× Tris·Cl/SDS, pH 8.8	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75
H ₂ O	8.75	8.25	7.75	7.50	7.25	6.75	6.25	5.25	4.75	3.75
10% (w/v) ammonium persulfate ^d	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
TEMED	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Preparation of separating gel

In a 25-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution, 4× Tris·Cl/SDS, pH 8.8 (see reagents, below), and H₂O. Degas under vacuum ~5 min. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

STACKING GEL (3.9% acrylamide)

In a 25-ml side-arm flask, mix 0.65 ml of 30% acrylamide/0.8% bisacrylamide, 1.25 ml of 4× Tris·Cl/SDS, pH 6.8 (see reagents, below), and 3.05 ml H₂O. Degas under vacuum 10 to 15 min. Add 25 μl of 10% ammonium persulfate and 5 μl TEMED. Swirl gently to mix. Use immediately. Failure to form a firm gel usually indicates a problem with the persulfate, TEMED, or both.

REAGENTS USED IN GELS**30% acrylamide/0.8% bisacrylamide**

Mix 30.0 g acrylamide and 0.8 g *N,N'*-methylenebisacrylamide with H₂O in a total volume of 100 ml. Filter the solution through a 0.45-μm filter and store at 4°C in the dark. The 2× crystallized grades of acrylamide and bisacrylamide are recommended. Discard after 30 days, as acrylamide gradually hydrolyzes to acrylic acid and ammonia.

CAUTION: *Acrylamide monomer is neurotoxic. A mask should be worn when weighing acrylamide powder. Gloves should be worn while handling the solution, and the solution should not be pipetted by mouth.*

4× Tris·Cl/SDS, pH 6.8 (0.5 M Tris·Cl containing 0.4% SDS)

Dissolve 6.05 g Tris base in 40 ml H₂O. Adjust to pH 6.8 with 1 N HCl. Add H₂O to 100 ml total volume. Filter the solution through a 0.45-μm filter, add 0.4 g SDS, and store at 4°C up to 1 month.

4× Tris·Cl/SDS, pH 8.8 (1.5 M Tris·Cl containing 0.4% SDS)

Dissolve 91 g Tris base in 300 ml H₂O. Adjust to pH 8.8 with 1 N HCl. Add H₂O to 500 ml total volume. Filter the solution through a 0.45-μm filter, add 2 g SDS, and store at 4°C up to 1 month.

^aThe recipes produce 15 ml of separating gel and 5 ml of stacking gel, which are adequate for a gel of dimensions 0.75 mm × 14 cm × 14 cm. The recipes are based on the SDS (denaturing) discontinuous buffer system of Laemmli (1970).

^bAll reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

^cUnits of numbers in table body are milliliters. The desired percentage of acrylamide in the separating gel depends on the molecular size of the protein being separated. See annotation to step 3, Basic Protocol 1.

^dBest to prepare fresh.

Prepare the sample and load the gel

12. Dilute a portion of the protein sample to be analyzed 1:1 (v/v) with 2× SDS sample buffer and heat 3 to 5 min at 100°C in a sealed screw-cap microcentrifuge tube. If the sample is a precipitated protein pellet, dissolve the protein in 50 to 100 μl of 1× SDS sample buffer and boil 3 to 5 min at 100°C. Dissolve protein-molecular-weight standards mixture in 1× SDS sample buffer according to supplier's instructions; use these standards as a control (Table 10.2A.2).

For dilute protein solutions, consider adding 5:1 protein solution/6× SDS sample buffer to increase the amount of protein loaded. Proteins can also be concentrated by precipitation in acetone, ethanol, or trichloroacetic acid (TCA), but losses will occur.

For a 0.8-cm-wide well, 25 to 50 μg total protein in <20 μl is recommended for a complex mixture when staining with Coomassie blue, and 1 to 10 μg total protein is needed for samples containing one or a few proteins. If silver staining is used, 10- to 100-fold less protein can be applied (0.01 to 5 μg in <20 μl depending on sample complexity).

To achieve the highest resolution possible, the following precautions are recommended. Prior to adding the sample buffer, keep samples at 0°C. Add the SDS sample buffer (room temperature) directly to the 0°C sample (still on ice) in a screw-top microcentrifuge tube. Cap the tube to prevent evaporation, vortex, and transfer directly to a 100°C water bath for 3 to 5 min. Let immunoprecipitates dissolve for 1 hr at 56°C in 1× SDS sample buffer prior to boiling. DO NOT leave the sample in SDS sample buffer at room temperature without first heating to 100°C to inactivate proteases (see Critical Parameters and Troubleshooting). Endogenous proteases are very active in SDS sample buffer and will cause severe degradation of the sample proteins after even a few minutes at room temperature. To test for possible proteases, mix the sample with SDS sample buffer without heating and leave at room temperature for 1 to 3 hr. A loss of high-molecular-weight bands

Table 10.2A.2 Molecular Weights of Protein Standards for Polyacrylamide Gel Electrophoresis^a

Protein	Molecular weight
Cytochrome <i>c</i>	11,700
α-Lactalbumin	14,200
Lysozyme (hen egg white)	14,300
Myoglobin (sperm whale)	16,800
β-Lactoglobulin	18,400
Trypsin inhibitor (soybean)	20,100
Trypsinogen, PMSF treated	24,000
Carbonic anhydrase (bovine erythrocytes)	29,000
Glyceraldehyde-3-phosphate dehydrogenase (rabbit muscle)	36,000
Lactate dehydrogenase (porcine heart)	36,000
Aldolase	40,000
Ovalbumin	45,000
Catalase	57,000
Bovine serum albumin	66,000
Phosphorylase <i>b</i> (rabbit muscle)	97,400
β-Galactosidase	116,000
RNA polymerase, <i>E. coli</i>	160,000
Myosin, heavy chain (rabbit muscle)	205,000

^aProtein standards are commercially available in kits (e.g., Amersham Pharmacia Biotech, Life Technologies, Bio-Rad, or Sigma).

and a general smearing of the banding pattern indicate a protease problem. Once heated, the samples can sit at room temperature for the time it takes to load samples.

- Carefully remove the Teflon comb without tearing the edges of the polyacrylamide wells. After the comb is removed, rinse wells with 1× SDS electrophoresis buffer.

The rinse removes unpolymerized monomer; otherwise, the monomer will continue to polymerize after the comb is removed, creating uneven wells that will interfere with sample loading and subsequent separation.

- Using a Pasteur pipet, fill the wells with 1× SDS electrophoresis buffer.

If well walls are not upright, they can be manipulated with a flat-tipped needle attached to a syringe.

- Attach gel sandwich to upper buffer chamber following manufacturer's instructions.

- Fill lower buffer chamber with the recommended amount of 1× SDS electrophoresis buffer.

- Place sandwich attached to upper buffer chamber into lower buffer chamber.

- Partially fill the upper buffer chamber with 1× SDS electrophoresis buffer so that the sample wells of the stacking gel are filled with buffer.

Monitor the upper buffer chamber for leaks and if necessary, reassemble the unit. A slow leak in the upper buffer chamber may cause arcing around the upper electrode and damage the upper buffer chamber.

- Using a 25- or 100- μ l syringe with a flat-tipped needle, load the protein sample(s) into one or more wells by carefully applying the sample as a thin layer at the bottom of the wells. Load control wells with molecular weight standards. Add an equal volume of 1× SDS sample buffer to any empty wells to prevent spreading of adjoining lanes.

Preparing the samples at approximately the same concentration and loading an equal volume to each well will ensure that all lanes are the same width and that the proteins run evenly. If unequal volumes of sample buffer are added to wells, the lane with the larger volume will spread during electrophoresis and constrict the adjacent lanes, causing distortions.

The samples will layer on the bottom of the wells because the glycerol added to the sample buffer gives the solution a greater density than the electrophoresis buffer. The bromphenol blue in the sample buffer makes sample application easy to follow visually.

- Fill the remainder of the upper buffer chamber with additional 1× SDS electrophoresis buffer so that the upper platinum electrode is completely covered. Do this slowly so that samples are not swept into adjacent wells.

Run the gel

- Connect the power supply to the cell and run at 10 mA of constant current for a slab gel 0.75 mm thick, until the bromphenol blue tracking dye enters the separating gel. Then increase the current to 15 mA.

For a standard 16-cm gel sandwich, 4 mA per 0.75-mm-thick gel will run ~15 hr (i.e., overnight); 15 mA per 0.75-mm gel will take 4 to 5 hr. To run two gels or a 1.5-mm-thick gel, simply double the current. When running a 1.5-mm gel at 30 mA, the temperature must be controlled (10° to 20°C) with a circulating constant-temperature water bath to prevent "smiling" (curvature in the migratory band). Temperatures <5°C should not be used because SDS in the running buffer will precipitate. If the level of buffer in the upper chamber decreases, a leak has occurred.

22. After the bromphenol blue tracking dye has reached the bottom of the separating gel, disconnect the power supply.

Refer to Safety Considerations under Electricity and Electrophoresis.

Disassemble the gel

23. Discard electrode buffer and remove the upper buffer chamber with the attached gel sandwich.
24. Orient the gel so that the order of the sample wells is known, remove the sandwich from the upper buffer chamber, and lay the sandwich on a sheet of absorbent paper or paper towels.
25. Carefully slide one of the spacers halfway from the edge of the sandwich along its entire length. Use the exposed spacer as a lever to pry open the glass plate, exposing the gel.
26. Carefully remove the gel from the lower plate. Cut a small triangle off one corner of the gel so the lane orientation is not lost during staining and drying. Proceed with protein detection.

The gel can be stained with Coomassie blue or silver (UNIT 10.6), or proteins can be electroeluted, electroblotted onto a polyvinylidene difluoride (PVDF) membrane for subsequent staining or sequence analysis, or transferred to a membrane for immunoblotting UNIT 10.8. If the proteins are radiolabeled, they can be detected by autoradiography (APPENDIX 3A).

ELECTROPHORESIS IN TRIS-TRICINE BUFFER SYSTEMS

Separation of peptides and proteins under 10 to 15 kDa is not possible in the traditional Laemmli discontinuous gel system (see Basic Protocol 1). This is due to the comigration of SDS and smaller proteins, obscuring the resolution. Two approaches to obtain the separation of small proteins and peptides in the range of 5 to 20 kDa are presented: the following Tris-tricine method and a system using increased buffer concentrations (see Alternate Protocol 2). The Tris-tricine system uses a modified buffer to separate the SDS and peptides, thus improving resolution. Several precast gels are available for use with the tricine formulations (Table 10.2A.3).

**ALTERNATE
PROTOCOL 1**

Table 10.2A.3 Vertical Format Precast Gel Compatibility

Gel type and compatibility	Gel supplier			
	Bio-Rad	ISS/Daiichi	Jule	Novex
<i>SDS-PAGE gel type offered</i>				
Peptide (tricine)	×	×	×	×
Single concentration	×	×	×	×
Gradient	×	×	×	×
Minigel size	×	×	×	×
Standard gel size		×	×	
<i>Compatibility of gel with equipment manufactured by</i>				
Amersham Pharmacia Biotech		×	×	×
Bio-Rad	×	×	×	
Life Technologies	×	×	×	×
Novex		×		×
ISS/Daiichi		×	×	

**Analysis of
Proteins**

10.2A.9

Additional Materials (also see *Basic Protocol 1*)

Separating and stacking gel solutions (Table 10.2A.4)
2× tricine sample buffer (see recipe)
Peptide molecular-weight-standards mixture (Table 10.2A.5)
Cathode buffer (see recipe)
Anode buffer (see recipe)
Coomassie blue G-250 staining solution (see recipe)
10% (v/v) acetic acid

1. Prepare and pour the separating and stacking gels (see *Basic Protocol 1*, steps 1 to 11), using Table 10.2A.4 in place of Table 10.2A.1.
2. Prepare the sample (see *Basic Protocol 1*, step 12), but make the following changes for tricine gels. Substitute 2× tricine sample buffer for the 2× SDS sample buffer. Dilute an aliquot of the protein or peptide sample to be analyzed 1:1 (v/v) with 2× tricine sample buffer. Treat the sample at 40°C for 30 to 60 min prior to loading.

If proteolytic activity is a problem, heating samples to 100°C for 3 to 5 min before loading the wells may be required (see Basic Protocol 1, annotation to step 12). Use the peptide molecular-weight-standards mixture for peptide separations (Table 10.2A.5).
3. Load the gel and set up the electrophoresis apparatus (see *Basic Protocol 1*, steps 13 to 20) with the following alterations. Remove comb and, using the tricine-containing

Table 10.2A.4 Recipes for Tricine Peptide Separating and Stacking Gels^a

SEPARATING AND STACKING GELS

Stock solution ^b	Separating gel	Stacking gel
30% acrylamide/0.8% bisacrylamide	9.80 ml	1.62 ml
Tris·Cl/SDS, pH 8.45	10.00 ml	3.10 ml
H ₂ O	7.03 ml	7.78 ml
Glycerol	4.00 g (3.17 ml)	—
10% (w/v) ammonium persulfate ^c	50 μl	25 μl
TEMED	10 μl	5 μl

Prepare separating and stacking gel solutions separately.

In a 50-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (Table 10.2A.1), Tris·Cl/SDS, pH 8.45 (see reagents, below), and H₂O. Add glycerol to separating gel only. Degas under vacuum 10 to 15 min. Add 10% ammonium persulfate and TEMED. Swirl gently to mix, use immediately. Failure to form a firm gel usually indicates a problem with the persulfate, TEMED, or both.

ADDITIONAL REAGENTS USED IN GELS

Tris·Cl/SDS, pH 8.45 (3.0 M Tris·Cl containing 0.3% SDS)

Dissolve 182 g Tris base in 300 ml H₂O. Adjust to pH 8.45 with 1 N HCl. Add H₂O to 500 ml total volume. Filter the solution through a 0.45-μm filter, add 1.5 g SDS, and store at 4°C up to 1 month.

^aThe recipes produce 30 ml of separating gel and 12.5 ml of stacking gel, which are adequate for two gels of dimensions 0.75 mm × 14 cm × 14 cm. The recipes are based on the Tris-tricine buffer system of Schagger and von Jagow (1987).

^bAll reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

^cBest to prepare fresh.

Table 10.2A.5 Molecular Weights of Peptide Standards for Polyacrylamide Gel Electrophoresis^a

Peptide	Molecular weight (Da)
Myoglobin (polypeptide backbone)	16,950
Myoglobin 1-131	14,440
Myoglobin 56-153	10,600
Myoglobin 56-131	8,160
Myoglobin 1-55	6,210
Glucagon	3,480
Myoglobin 132-153	2,510

^aPeptide standards are commercially available from Sigma.

cathode buffer, or water, rinse once and fill wells. Fill the lower buffer chamber with anode buffer, assemble the unit, and attach the upper buffer chamber. Fill the upper buffer chamber with cathode buffer and load the samples.

4. Connect the power supply to the cell and run 1 hr at 30 V (constant voltage) followed by 4 to 5 hr at 150 V (constant voltage). Use heat exchanger to keep the electrophoresis chamber at room temperature.
5. After the tracking dye has reached the bottom of the separating gel, disconnect the power supply.

Refer to Safety Considerations under Electricity and Electrophoresis.

Coomassie blue G-250 is used as a tracking dye instead of bromphenol blue because it moves ahead of the smallest peptides.

6. Disassemble the gel (see Basic Protocol 1, steps 23 to 26). Stain proteins in the gel for 1 to 2 hr in Coomassie blue G-250 staining solution. Follow by destaining with 10% acetic acid, changing the solution every 30 min until background is clear (3 to 5 changes). For higher sensitivity, use silver staining as a recommended alternative.

Prolonged staining and destaining will result in the loss of resolution of the smaller proteins (<10 kDa). Proteins diffuse within the gel and out of the gel, resulting in a loss of staining intensity and resolution.

NONUREA PEPTIDE SEPARATIONS WITH TRIS BUFFERS

A simple modification of the traditional Laemmli buffer system presented in Basic Protocol 1, in which the increased concentration of buffers provides better separation between the stacked peptides and the SDS micelles, permits reasonable separation of peptides as small as 5 kDa.

Additional Materials (also see Basic Protocol 1)

Separating and stacking gel solutions (Table 10.2A.6)

2× SDS electrophoresis buffer (see recipe)

2× Tris·Cl/SDS, pH 8.8 (dilute 4× Tris·Cl/SDS, pH 8.8; Table 10.2A.1)

1. Prepare and pour the separating gel (see Basic Protocol 1, steps 1 to 6) using Table 10.2A.6 in place of Table 10.2A.1.
2. Prepare and pour the stacking gel (see Basic Protocol 1, steps 7 to 11), using 2× Tris·Cl/SDS, pH 8.8, rather than 1× Tris·Cl/SDS buffer, for rinsing the gel after removing the isobutyl alcohol overlay.

**ALTERNATE
PROTOCOL 2**

**Analysis of
Proteins**

10.2A.11

- Prepare the sample and load the gel (see Basic Protocol 1, steps 12 to 20) and substitute 2× SDS electrophoresis buffer for the 1× SDS electrophoresis buffer.

Table 10.2A.5 lists the standards for small protein separations.

- Run the gel (see Basic Protocol 1, steps 21 and 22).

Note that the separations will take ~25% longer than those using Basic Protocol 1. The increased buffer concentrations lead to faster transit through the stacking gel but lower mobility in the resolving gel.

- Disassemble the gel (see Basic Protocol 1, steps 23 to 26).

Proteins in the gel may now be stained.

Table 10.2A.6 Recipes for Modified Laemmli Peptide Separating and Stacking Gels^a

SEPARATING AND STACKING GELS

Stock solution ^b	Separating gel	Stacking gel
30% acrylamide/0.8% bisacrylamide	10.00 ml	0.65 ml
8× Tris·Cl, pH 8.8	3.75 ml	—
4× Tris·Cl, pH 6.8	—	1.25 ml
10% (w/v) SDS	0.15 ml	50 μl
H ₂ O	1.00 ml	3.00 ml
10% (w/v) ammonium persulfate ^c	50 μl	25 μl
TEMED	10 μl	5 μl

Prepare separating and stacking gel solutions separately

In a 25-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.2A.1), 8× Tris·Cl, pH 8.8 (separating gel) or 4× Tris·Cl, pH 6.8 (stacking gel), 10% SDS (see reagents, below), and H₂O. Degas under vacuum 10 to 15 min. Add 10% ammonium persulfate and TEMED. Swirl gently to mix; use immediately. Failure to form a firm gel usually indicates a problem with the persulfate, TEMED, or both.

ADDITIONAL REAGENTS USED IN GELS

4× Tris·Cl, pH 6.8 (0.5 M Tris·Cl)

Dissolve 6.05 g Tris base in 40 ml H₂O. Adjust to pH 6.8 with 1 N HCl. Add H₂O to 100 ml total volume. Filter the solution through a 0.45-μm filter and store up to 1 month at 4°C.

8× Tris·Cl, pH 8.8 (3.0 M Tris·Cl)

Dissolve 182 g Tris base in 300 ml H₂O. Adjust to pH 8.8 with 1 N HCl. Add H₂O to 500 ml total volume. Filter the solution through a 0.45-μm filter and store up to 1 month at 4°C.

10% (w/v) SDS

Mix 1 g SDS in 10 ml of H₂O. Use immediately.

^aThe recipes produce 15 ml of separating gel and 5 ml of stacking gel, which are adequate for one gel of dimensions 0.75 mm × 14 cm × 14 cm. The recipes are based on the modified Laemmli peptide separation system of Okajima et al. (1993).

^bAll reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

CONTINUOUS SDS-PAGE

With continuous SDS-PAGE, the same buffer is used for both the gel and electrode solutions. Although continuous gels lack the resolution of the discontinuous systems, they are extremely versatile, less prone to mobility artifacts, and much easier to prepare. The stacking gel is omitted.

Additional Materials (also see *Basic Protocol 1*)

- Separating gel solution (Table 10.2A.7)
- 2× and 1× phosphate/SDS sample buffer (see recipe)
- 1× phosphate/SDS electrophoresis buffer (see recipe)

1. Prepare and pour a single separating gel (see *Basic Protocol 1*, steps 1 to 4), except use solutions in Table 10.2A.7 in place of those in Table 10.2A.1 and fill the gel sandwich to the top. Omit the stacking gel. Insert the comb (see *Basic Protocol 1*, step 10) and allow the gel to polymerize 30 to 60 min at room temperature.
2. Mix the protein sample 1:1 with 2× phosphate/SDS sample buffer and heat to 100°C for 2 min.

*For large sample volumes or samples suspended in high ionic strength buffers (>50 mM), dialyze against 1× sample buffer prior to electrophoresis. Note that the precautions about proteases (see *Basic Protocol 1*, step 12) apply here.*

Table 10.2A.7 Recipes for Separating Gels for Continuous SDS-PAGE^a**SEPARATING GEL**

Stock solution ^b	Final acrylamide concentration in separating gel (%) ^c											
	5	6	7	8	9	10	11	12	13	14	15	
30% acrylamide/ 0.8% bisacrylamide	2.5	3.00	3.50	4.00	4.50	5.00	5.50	6.00	6.50	7.00	7.50	
4× phosphate/SDS, pH 7.2	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	
H ₂ O	8.75	8.25	7.75	7.25	6.75	6.25	5.75	5.25	4.75	4.25	3.75	
10% (w/v) ammonium persulfate ^d	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	
TEMED	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	

Preparation of separating gel

In a 25-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.2A.1), 4× phosphate/SDS, pH 7.2, and H₂O. Degas under vacuum about 5 min. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

ADDITIONAL REAGENTS USED IN GELS**4× phosphate/SDS, pH 7.2 (0.4 M sodium phosphate/0.4% SDS)**

Mix 46.8 g NaH₂PO₄·H₂O, 231.6 g Na₂HPO₄·7 H₂O, and 12 g SDS in 3 liters H₂O.

Store at 4°C for up to 3 months.

^aThe recipes produce 15 ml of separating gel, which is adequate for one gel of dimensions 0.75 mm × 14 cm × 14 cm. The recipes are based on the original continuous phosphate buffer system of Weber et al. (1972). The stacking gel is omitted.

^bAll reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

^cUnits of numbers in table body are milliliters. The desired percentage of acrylamide in the separating gel depends on the molecular size of the protein being separated. See *Basic Protocol 1*, annotation to step 3.

^dBest to prepare fresh.

**ALTERNATE
PROTOCOL 4**

3. Assemble the electrophoresis apparatus and load the sample (see Basic Protocol 1, steps 13 to 20), using the phosphate/SDS electrophoresis buffer and loading empty wells with 1× phosphate/SDS sample buffer.
4. Connect the power supply and start the run with 15 mA per 0.75-mm-thick gel until the tracking dye has entered the gel. Continue electrophoresis at 30 mA for 3 hr (5% gel), 5 hr (10% gel), 8 hr (15% gel), or until the dye reaches the bottom of the gel.
Use temperature control if available to maintain the gel at 15° to 20°C. SDS will precipitate below 15°C in this system.
5. Disassemble the gel (see Basic Protocol 1, steps 23 to 26).
See Safety Considerations in introduction. Proteins in the gel may now be stained.

CASTING AND RUNNING ULTRATHIN GELS

Ultrathin gels provide superb resolution but are difficult to handle. In this application, gels are cast on Gel Bond, a Mylar support material. Silver staining is recommended for the best resolution. Combs and spacers for gels <0.5 mm thick are not readily available for most protein electrophoresis units. However, by adapting combs and spacers used for DNA sequencing, casting gels from 0.2 to 0.5 mm thick is straightforward.

Additional Materials (also see Basic Protocol 1)

- 95% (v/v) ethanol
 - Gel Bond (FMC BioProducts) cut to a size slightly smaller than the gel plate dimensions
 - Glue stick
 - Ink roller (available from art supply stores)
 - Combs and spacers (0.19 to 0.5 mm; sequencing gel spacers and combs can be cut to fit)
1. Wash gel plates with water-based laboratory detergent followed by successive rinses with hot tap water, deionized water, and finally 95% ethanol. Allow to air dry.
Gel plates must be extremely clean for casting thin gels.
Gloves are needed throughout these procedures to prevent contamination by proteins on the surface of skin.
 2. Apply a streak of adhesive from a glue stick to the bottom edge of the glass plate. Quickly position the Gel Bond with the hydrophobic side down (a drop of water will bead up on the hydrophobic surface). Apply pressure with Kimwipe tissue to attach the Gel Bond firmly to the plate. Finally, pull the top portion of the Gel Bond back, place a few drops of water underneath, and roll flat with an ink roller.
Make sure the Gel Bond does not extend beyond the edges of the upper and lower sealing surface of the plate. This will cause it to buckle on sealing. Reposition the Gel Bond if needed to prevent it from extending beyond the glass plate. Material may also be trimmed to fit flush with the plate edge.
 3. Assemble the gel cassette according to the manufacturer's instructions (also see Basic Protocol 1, steps 1 and 2). Just prior to assembly, blow air over the surface of both the Gel Bond and the opposing glass surface to remove any particulate material (e.g., dust).
Sequencing gel spacers can be easily adapted. First, cut the spacers slightly longer than the length of the gel plate. Position a spacer along each edge of the glass plate and assemble the gel sandwich, clamping in place. With a razor blade, trim the excess spacer at top and bottom to get a reusable spacer exactly the size of the plate.

4. Prepare and pour the separating and stacking gels (see Basic Protocol 1, steps 3 to 9). In place of the Teflon comb, insert a square well sequencing comb cut to fit within the gel sandwich. Allow the stacking gel to polymerize 30 to 45 min at room temperature.

Less solution is needed for ultrathin gels. For example, a 0.5-mm-thick gel requires 33% less gel solution than a 0.75-mm gel.

5. Prepare the sample and load the gel (see Basic Protocol 1, steps 12 to 20).

When preparing protein samples for ultrathin gels, 3 to 4 μ l at 5 μ g protein/ μ l is required for Coomassie blue R-250 staining, whereas 10-fold less is needed for silver staining.

6. Run the gel (see Basic Protocol 1, steps 21 and 22), except conduct the electrophoresis at 7 mA/gel (0.25-mm-thick gels) or 14 mA/gel (0.5-mm-thick gels) for 4 to 5 hr.

7. When the separation is complete, disassemble the unit and remove the gel (see Basic Protocol 1, steps 23 to 26) with the Gel Bond still attached. With a gloved hand, wash away the adhesive material from the back of the Gel Bond under a stream of water before proceeding to protein detection.

Either Coomassie blue or silver staining may be used, but silver staining produces particularly fine resolution with thin Gel Bond-backed gels. Compared to staining thicker (>0.75 mm) gels, thin (<0.75 mm) gels stain and destain more quickly. Although the optimum staining times must be empirically determined, all steps in Coomassie blue and silver staining procedures are generally reduced by half.

CASTING MULTIPLE SINGLE-CONCENTRATION GELS

Casting multiple gels at one time has several advantages. All the gels are identical, so sample separation is not affected by gel-to-gel variation. Furthermore, casting ten gels is only slightly more difficult than casting two gels. Once cast, gels can be stored for several days in a refrigerator.

Additional Materials (also see Basic Protocol 1)

Separating and stacking gels for single-concentration gels (Table 10.2A.8)

H₂O-saturated isobutyl alcohol

Multiple gel caster (Bio-Rad, Amersham Pharmacia Biotech)

100-ml disposable syringe and flat-tipped needle

Extra plates and spacers

14 × 14-cm acrylic blocks or polycarbonate sheets

250- and 500-ml side-arm flasks (used in gel preparation)

Long razor blade *or* plastic wedge (Wonder Wedge, Amersham Pharmacia Biotech)

Resealable plastic bags

Pour the separating gel

1. Assemble the multiple gel caster according to the manufacturer's instructions.

With the Amersham Pharmacia Biotech unit make sure to insert the large triangular space filler plug in the base of the caster. The plug is removed when casting gradient gels (see Support Protocol 2).

2. Assemble glass sandwiches and stack them in the casting chamber. Stack up to ten 1.5-mm gels and fill in extra space with acrylic blocks or polycarbonate sheets to hold the sandwiches tightly in place. Make sure the spacers are straight along the top, right, and left edges of the glass plates and that all edges of the stack are flush.

The presence of loosely fitting sandwiches in the caster will lead to unevenly cast gels, creating distortions during electrophoresis. Polycarbonate inhibits gel polymerization.

SUPPORT PROTOCOL 1

Analysis of Proteins

10.2A.15

Table 10.2A.8 Recipes for Multiple Single-Concentration Polyacrylamide Gels^a**SEPARATING GEL**

Stock solution ^b	Final acrylamide concentration in separating gel (%) ^c										
	5	6	7	8	9	10	11	12	13	14	15
30% acrylamide/0.8% bisacrylamide	52	62	72	83	93	103	114	124	134	145	155
4× Tris·Cl/SDS, pH 8.8	78	78	78	78	78	78	78	78	78	78	78
H ₂ O	181	171	160	150	140	129	119	109	98	88	78
10% (w/v) ammonium persulfate ^d	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
TEMED	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21

Preparation of separating gel

In a 500-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.2A.1), 4× Tris·Cl/SDS, pH 8.8 (Table 10.2A.1), and H₂O. Degas under vacuum ~5 min. Add 10% ammonium persulfate and TEMED. Swirl gently to mix; use immediately.

STACKING GEL

In a 250-ml side-arm flask, mix 13.0 ml 30% acrylamide/0.8% bisacrylamide solution, 25 ml 4× Tris·Cl/SDS, pH 6.8 (Table 10.2A.1), and 61 ml H₂O. Degas under vacuum ~5 min. Add 0.25 ml 10% ammonium persulfate and 50 μl TEMED. Swirl gently to mix. Use immediately. Failure to form a firm gel usually indicates a problem with the persulfate, TEMED, or both.

^aThe recipes produce about 300 ml of separating gel and 100 ml of stacking gel, which are adequate for ten gels of dimensions 1.5 mm × 14 cm × 14 cm. Volumes were measured using 1.5-mm spacers. For thinner spacers or fewer gels, calculate volumes using the equation in the annotation to step 4. The recipes are based on the SDS (denaturing) discontinuous buffer system of Laemmli (1970).

^bAll reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

^cUnits of numbers in table body are milliliters. The desired percentage of acrylamide in separating gel depends on the molecular size of the protein being separated. See Basic Protocol 1, annotation to step 3.

^dBest to prepare fresh.

Therefore, if polycarbonate sheets are placed in the caster before and after the set of glass sandwiches, the entire set will slide out as one block after polymerization. Placing polycarbonate sheets between each gel sandwich makes them easier to separate from one another after polymerization.

- Place the front sealing plate on the casting chamber, making sure the stack fits snugly. Secure the plate with four spring clamps and tighten the bottom thumb screws.
- Prepare the separating (resolving) gel solution (Table 10.2A.8).

A 12-cm separating gel with a 4-cm stacking gel is recommended.

If fewer than ten gels are prepared (Table 10.2A.8), use the following formula to estimate the amount of separating gel volume needed:

$$\text{Volume} = \text{gel number} \times \text{height (cm)} \times \text{width (cm)} \times \text{thickness (cm)} + 4 \times \text{gel number} + 10 \text{ ml.}$$

- Using a 100-ml disposable syringe with flat-tipped needle, inject the resolving gel solution down the side of one spacer into the multiple caster. A channel in the silicone plug distributes the solution throughout the whole caster. Avoid introducing bubbles by giving the caster a quick tap on the benchtop once the caster is filled.

6. Overlay the center of each gel with 100 μ l H₂O-saturated isobutyl alcohol and let polymerize for 1 to 2 hr.
7. Drain off the overlay and rinse the surface with 1 \times Tris-Cl/SDS, pH 8.8. If the gels will not be used immediately, skip to step 12.

Pour the stacking gel

8. Prepare the stacking gel solution either singly (see Basic Protocol 1, step 8) or for all the gels at once (Table 10.2A.8).

The stacking gel solution should be prepared just before pouring the gel.

9. Fill each sandwich in the caster with stacking gel solution.
10. Insert a comb into each sandwich and let the gel polymerize for 2 hr.

Insert the combs at a 45° angle to avoid trapping air underneath the comb teeth. Air bubbles will inhibit polymerization and cause dents in the wells and a distorted pattern of protein bands.

11. Remove the combs and rinse wells with 1 \times SDS electrophoresis buffer.

Remove the gels from the caster

12. Remove the gels from the caster and separate by carefully inserting a long razor blade or knife between each gel sandwich.

A plastic wedge (Amersham Pharmacia Biotech's Wonder Wedge) also works well.

13. Clean the outside of each gel plate with running water to remove the residual polymerized and unpolymerized acrylamide.
14. Overlay gels to be stored with 1 \times Tris-Cl/SDS, pH 8.8, place in a resealable plastic bag, and store at 4°C until needed (up to 1 week).

SEPARATION OF PROTEINS ON GRADIENT GELS

Gels that consist of a gradient of increasing polyacrylamide concentration resolve a much wider size range of proteins than standard uniform-concentration gels (see Critical Parameters and Troubleshooting). The protein bands, particularly in the low-molecular-weight range, are also much sharper. Unlike single-concentration gels, gradient gels separate proteins in a way that can be represented easily to give a linear plot from 10 to 200 kDa. This facilitates molecular weight estimations.

The quantities given below provide separating gel solution sufficient for two 0.75-mm gels (~7 ml of each concentration) or one 1.5-mm gel (~14 ml of each concentration). Volumes can be adjusted to accommodate gel sandwiches of different dimensions.

Additional Materials (also see *Basic Protocol 1*)

Light and heavy acrylamide gel solutions (Table 10.2A.9 and Table 10.2A.10)
 Bromphenol blue (optional; for checking practice gradient)
 TEMED

Gradient maker (30 to 50 ml, Amersham Pharmacia Biotech SG30 or SG50; or 30 to 100 ml, Bio-Rad 385)

Tygon tubing with micropipet tip

Peristaltic pump (optional; e.g., Markson A-13002, A-34040, or A-34105 minipump)

Whatman 3MM filter paper

***ALTERNATE
 PROTOCOL 5***

Set up the gradient maker and prepare the gel solutions

1. Assemble the magnetic stirrer and gradient maker on a ring stand as shown in Figure 10.2A.2. Connect the outlet valve of the gradient maker to Tygon tubing attached to a micropipet tip that is placed over the vertical gel sandwich. If desired, place a peristaltic pump in line between the gradient maker and the gel sandwich.

A peristaltic pump will simplify casting by providing a smooth flow rate.

2. Place a small stir-bar into the mixing chamber of the gradient maker (i.e., the chamber connected to the outlet).
3. Using the recipes in Table 10.2A.9 and Table 10.2A.10, prepare light and heavy acrylamide gel solutions. Do not add ammonium persulfate until just before use (step 7).
4. With the outlet port and interconnecting valve between the two chambers closed, pipet 7 ml of light (low-concentration) acrylamide gel solution into the reservoir chamber for one 0.75-mm-thick gradient gel.

Recommended gradient ranges are 5% to 20% for most applications (to separate proteins of 10 to several hundred kilodaltons).

A practice run with heavy and light solutions is recommended. Bromphenol blue should be added to the heavy solution to demonstrate linearity of the practice gradient.

5. Open the interconnecting valve briefly to allow a small amount (~200 μ l) of light solution to flow through the valve and into the mixing chamber.

The presence of air bubbles in the interconnecting valve may obstruct the flow between chambers during casting.

Deaeration is not recommended for either the light or heavy solution. Omitting the deaeration will allow polymerization to proceed more slowly, letting the gradient establish itself in the gel sandwich before polymerization takes place.

6. Add 7 ml of heavy (high-concentration) acrylamide gel solution to the mixing chamber.

Keep the heavy solution on ice until use. Once the ammonium persulfate is added to the heavy solution, it will polymerize without TEMED, albeit more slowly; keeping the solution on ice prevents this. The gel solution will come to room temperature during casting. The higher the percentage of acrylamide, the more severe the problem of premature polymerization.

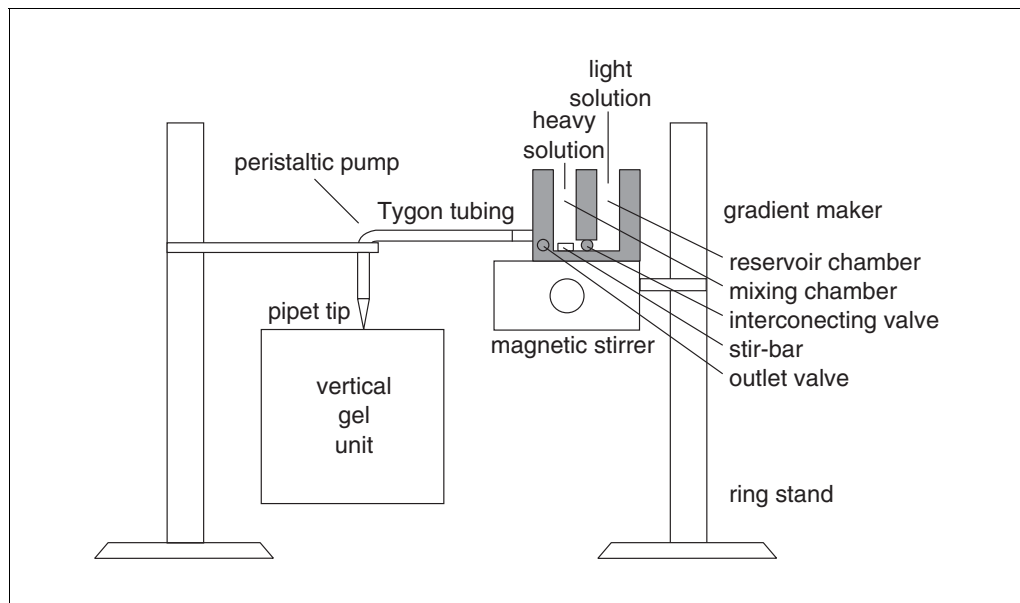


Figure 10.2A.2 Gradient gel setup. A peristaltic pump, though not required, will provide better control.

Table 10.2A.9 Light Acrylamide Gel Solutions for Gradient Gels^a

Stock solution	Acrylamide concentration of light gel solution (%) ^b									
	5	6	7	8	9	10	11	12	13	14
30% acrylamide/ 0.8% bisacrylamide ^c	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
4× Tris·Cl/SDS, pH 8.8 ^c	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75
H ₂ O	8.75	8.25	7.75	7.25	6.75	6.25	5.75	5.25	4.75	4.25
10% (w/v) ammonium persulfate ^d	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05

^aTo survey proteins ≥10 kDa, 5%-20% gradient gels are recommended. To expand the range between 10 and 200 kD, a 10%-20% gel is recommended.

^bNumbers in body of table are milliliters of stock solution. Deaeration is not required. Keep solution at room temperature prior to adding TEMED no longer than 1 hr.

^cSee Table 10.2A.1 for preparation.

^dBest to prepare fresh.

Table 10.2A.10 Heavy Acrylamide Gel Solutions for Gradient Gels^a

Stock solution	Acrylamide concentration of heavy gel solution (%) ^b											
	10	11	12	13	14	15	16	17	18	19	20	
30% acrylamide/ 0.8% bisacrylamide ^c	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0	
4× Tris·Cl/SDS, pH 8.8 ^c	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	
H ₂ O	5.0	4.5	4.0	3.5	3.0	2.5	2.0	1.5	1.0	0.5	0	
Sucrose (g)	2.25	2.25	2.25	2.25	2.25	2.25	2.25	2.25	2.25	2.25	2.25	
10% (w/v) ammonium persulfate ^{b,d}	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	

^aDeaeration is not recommended for gradient gels.

^bNumbers in body of table are milliliters of stock solution (except sucrose). Do not add the ammonium persulfate until just before use. The heavy acrylamide will polymerize, albeit more slowly, without the addition of TEMED. Keep the heavy solution on ice after adding ammonium persulfate.

^cSee Table 10.2A.1 for preparation.

^dBest to prepare fresh.

7. Add the specified amount of 10% ammonium persulfate and ~2.3 μl TEMED per 7 ml acrylamide solution to each chamber. Mix the solutions in each chamber with a disposable pipet.

Form the gradient and cast the gel

8. Open the interconnecting valve completely.

Some of the heavy solution will flow back into the reservoir chamber containing light solution as the two chambers equilibrate. This will not affect the formation of the gradient.

9. Turn on the magnetic stirrer and adjust the rate to produce a slight vortex in the mixing chamber.

10. Open the outlet of the gradient maker slowly. Adjust the outlet valve to a flow rate of 2 ml/min. If using a peristaltic pump, calibrate the flow rate with a graduated cylinder prior to casting the gel.

Some adjustment of the flow rate may be necessary during casting. If the light solution is not flowing into the mixing chamber, a bubble may be caught in the interconnecting valve.

Quickly close the outlet and cover the top of the reservoir chamber with a gloved thumb. Push down with the thumb to increase the pressure in the chamber and force the air bubble out of the center valve.

11. Fill the gel sandwich from the top. Place the pipet tip against one side of the sandwich so the solution flows down one plate only. The heavy solution will flow into the sandwich first, followed by progressively lighter solution.
12. Watch as the last of the light solution drains into the outlet tube and adjust the flow rate to ensure that the last few milliliters of solution do not flow quickly into the gel sandwich and disturb the gradient.
13. Overlay the gradient gel with H₂O-saturated isobutyl alcohol. Allow the gel to polymerize ~1 hr.

In this gel system, the gel will polymerize from the bottom (i.e., heavy solution) up. Because polymerization is an exothermic reaction, heat can be felt evolving from the bottom of the gel sandwich during polymerization. A sharp optical discontinuity at the gel-overlay interface indicates that polymerization has occurred. In general, 1 hr is adequate for polymerization.

14. Remove the H₂O-saturated isobutyl alcohol and rinse with 1× Tris-Cl/SDS, pH 8.8. Cast the stacking gel (see Basic Protocol 1, steps 8 to 11).

The gel can be covered with 1× Tris-Cl/SDS, pH 8.8, sealed in a plastic bag, and stored for up to 1 week.

Load and run the gel

15. Prepare the protein sample and protein molecular-weight-standards mixture. Load and run the gel (see Basic Protocol 1, steps 13 to 26).

The gel can be stained with Coomassie blue or silver (UNIT 10.6).

16. After staining, dry the gels onto Whatman 3MM or equivalent filter paper.

Gradient gels >0.75 mm thick require special handling during drying to prevent cracking. The simplest approach to drying gradient gels is to use thin gels; ≤0.75-mm gradient gels with ≤20% acrylamide solutions will dry without cracking as long as the vacuum pump is working properly and the cold trap is dry at the onset of drying. For gradient gels >0.75 mm thick, add 3% (w/v) glycerol to the final destaining solution to help prevent cracking. Another method is to dehydrate and shrink the gel in 30% methanol for up to 3 hr prior to drying. Then place the gel in distilled water for 5 min before drying.

SUPPORT PROTOCOL 2

CASTING MULTIPLE GRADIENT GELS

Casting gradient gels in a multiple gel caster has several advantages. In addition to the time savings, batch casting gives gels that are essentially identical. This is particularly important for gradient gels, where slight variations in casting technique can cause variations in protein mobility. The gels may be stored for up to 1 week after casting to ensure internal consistency from run to run during the week. Furthermore, gels with several ranges of concentrations (e.g., 5% to 20% and 10% to 20% acrylamide) can be cast and stored, giving much more flexibility to optimize separations.

Additional Materials (also see Alternate Protocol 5)

Plug solution (see recipe)

Light and heavy acrylamide gel solutions for multiple gradient gels (Table 10.2A.11 and Table 10.2A.12)

TEMED

H₂O-saturated isobutyl alcohol

Multiple gel caster (Bio-Rad, Amersham Pharmacia Biotech)
Peristaltic pump (25 ml/min)
500- or 1000-ml gradient maker (Bio-Rad, Amersham Pharmacia Biotech)
Tygon tubing

Set up system and pour separating gel

1. Assemble the multiple caster as in casting multiple single-concentration gels (see Support Protocol 1, steps 1 to 3), making sure to remove the triangular space filler plugs in the bottom of the caster.

The plug is used only when casting single-concentration gels.

2. Set up the peristaltic pump (Fig. 10.2A.3). Using a graduated cylinder and water, adjust the flow rate so that the volume of the gradient solution plus volume of plug solution is poured in ~15 to 18 min (~25 ml/min).
3. Set up the gradient maker. Close all valves and place a stir-bar in the mixing chamber, which is the one with the outlet port. Attach one end of a piece of Tygon tubing to the outlet of the gradient maker. Run the other end of the tubing through the peristaltic pump and attach it to the red inlet port at the bottom of the caster.

Choose a gradient maker that holds no more than four times the total volume of the gradient solution to be poured (i.e., a 1000- or 500-ml gradient maker).

4. Prepare solutions for the gradient maker (Table 10.2A.11 and Table 10.2A.12).
5. Add the TEMED to both heavy and light solutions (54 μ l/165 ml) and immediately pour the light (low-concentration) solution into the mixing chamber (the one with the port). Open the mixing valve slightly to allow the tunnel to fill and to avoid air bubbles. Close the valve again and pour the heavy (high-concentration) acrylamide solution into the reservoir chamber.
6. Start the magnetic stirrer and open the outlet valve; then start the pump and open the mixing valve.

In units for casting multiple gels, acrylamide solution flows in from the bottom. To use a multiple casting unit, the light solution is placed in the mixing chamber and the heavy solution in the reservoir. This is the reverse of casting a single gel (see Alternate Protocol 5).

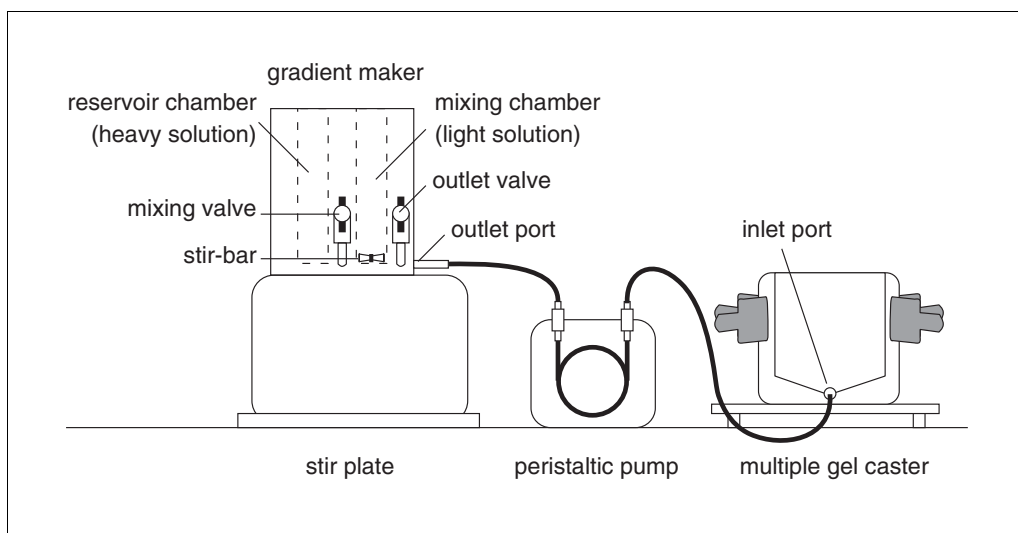


Figure 10.2A.3 Setup for casting multiple gradient gels. Casting multiple gradient gels requires a peristaltic pump and a multiple gel caster. Gel solution is introduced through the bottom of the multiple caster.

Table 10.2A.11 Light Acrylamide Gel Solutions for Multiple Gradient Gels^{a,b}

Stock solution	Acrylamide concentration of light separating gel solution (%) ^c									
	5	6	7	8	9	10	11	12	13	14
30% acrylamide/ 0.8% bisacrylamide ^d	28	33	39	44	50	55	61	66	72	77
4× Tris·Cl/SDS, pH 8.8 ^d	41	41	41	41	41	41	41	41	41	41
H ₂ O	96	91	85	80	74	69	63	58	52	47
10% (w/v) ammonium persulfate ^e	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55

^aTo survey proteins ≥10 kDa, 5%-20% gradient gels are recommended. To expand the range between 10 and 200 kDa, a 10%-20% gel is recommended.

^bRecipes produce ten 1.5-mm-thick gradient gels with 10 ml extra solution to account for losses in tubing.

^cNumbers in body of table are milliliters of stock solution. Deaeration is not required. Keep solution at room temperature prior to adding TEMED no longer than 1 hr.

^dSee Table 10.2A.1 for preparation.

^eBest to prepare fresh.

Table 10.2A.12 Heavy Acrylamide Gel Solutions for Multiple Gradient Gels^{a,b}

Stock solution	Acrylamide concentration of heavy gel solution (%) ^c										
	10	11	12	13	14	15	16	17	18	19	20
30% acrylamide/ 0.8% bisacrylamide ^d	55	61	66	72	77	83	88	94	99	105	110
4× Tris·Cl/SDS, pH 8.8 ^d	41	41	41	41	41	41	41	41	41	41	41
H ₂ O	55	50	44	39	33	28	22	17	11	5.5	0
Sucrose (g)	25	25	25	25	25	25	25	25	25	25	25
10% (w/v) ammonium persulfate ^e	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55

^aDeaeration is not recommended for gradient gels.

^bRecipes produce 10 ml extra solution to account for losses in tubing.

^cNumbers in body of table are milliliters of stock solution (except sucrose). Do not add the ammonium persulfate until just before use. The heavy acrylamide will polymerize, albeit more slowly, without the addition of TEMED. Keep the heavy solution on ice after adding ammonium persulfate.

^dSee Table 10.2A.1 for preparation.

^eBest to prepare fresh.

Thus, light solution enters the multiple caster first, followed by progressively heavier solution. Finally, the acrylamide solution is stabilized in the multiple caster with a heavy plug solution and allowed to polymerize (see step 8 and manufacturer's instructions).

- When almost all the acrylamide solution is gone from the gradient maker, stop the pump and close the mixing valve. Tilt the gradient maker toward the outlet side and remove the last milliliters of the mix. Do not allow air bubbles to enter the tubing.
- Add the plug solution to the mixing chamber and start the pump. Make sure that no bubbles are introduced. Continue pumping until the bottom of the caster is filled with plug solution to just below the glass plates; then turn off the pump. Clamp the tubing close to the red port of the casting chamber.
- Quickly overlay each separate gel sandwich with 100 μl H₂O-saturated isobutyl alcohol. Use the same amount on each sandwich.

Failure to use the same amount of overlay solution will cause the gel sandwiches to polymerize at different heights.

10. Drain off the overlay and rinse the surface of the gels with 1× Tris-Cl/SDS, pH 8.8.

Pour stacking gel and remove gels from caster

11. Prepare and cast the stacking gel as in casting multiple single-concentration gels (see Support Protocol 1, steps 8 to 11).
12. Remove gels from the caster and clean the gel sandwiches (see Support Protocol 1, steps 12 and 13). Store gels, if necessary, according to the instructions for multiple single-concentration gels (see Support Protocol 1, step 14).

ELECTROPHORESIS IN SINGLE-CONCENTRATION MINIGELS

Separation of proteins in a small-gel format is becoming increasingly popular for applications that range from isolating material for peptide sequencing to performing routine protein separations. The unique combination of speed and high resolution is the foremost advantage of small gels. Additionally, small gels are easily adapted to single-concentration, gradient, and two-dimensional SDS-PAGE procedures. The minigel procedures described are adaptations of larger gel systems.

This protocol describes the use of a multiple gel caster. The caster is simple to use, and up to five gels can be prepared at one time with this procedure. Single gels can be prepared using adaptations in the manufacturer's instructions. A multiple gel caster is the only practical way to produce small linear polyacrylamide gradient gels (see Support Protocol 3).

Materials

- Minigel vertical gel unit (Amersham Pharmacia Biotech Mighty Small SE 250/280 *or* Bio-Rad Mini-Protean II) with glass plates, clamps, and buffer chambers
- 0.75-mm spacers
- Multiple gel caster (Amersham Pharmacia Biotech SE-275/295 *or* Bio-Rad Mini-Protean II multicasting chamber)
- Acrylic plate (Amersham Pharmacia Biotech SE-217 *or* Bio-Rad 165-1957) *or* polycarbonate separation sheet (Amersham Pharmacia Biotech SE-213 *or* Bio-Rad 165-1958)
- 10- and 50-ml syringes
- Combs (Teflon, Amersham Pharmacia Biotech SE-211A series *or* Bio-Rad Mini-Protean II)
- Long razor blade
- Micropipet
- Additional reagents and equipment for standard denaturing SDS-PAGE (see Basic Protocol 1)

Pour the separating gel

1. Assemble each gel sandwich by stacking, in order, the notched (Amersham Pharmacia Biotech) *or* small rectangular (Bio-Rad) plate, 0.75-mm spacers, and the larger rectangular plate. Be sure to align the spacers properly, with the ends flush with the top and bottom edge of the two plates, when positioning the sandwiches in the multiple gel caster (Fig. 10.2A.4).

The protocol described is basically for the Amersham Pharmacia Biotech system. For other systems, make adjustments according to the manufacturer's instructions. Alternatively, precast minigels can be purchased from a number of suppliers (see Table 10.2A.3).

The multiple casters from Amersham Pharmacia Biotech have a notch in the base designed for casting gradient gels. A silicone rubber insert fills up this space when casting

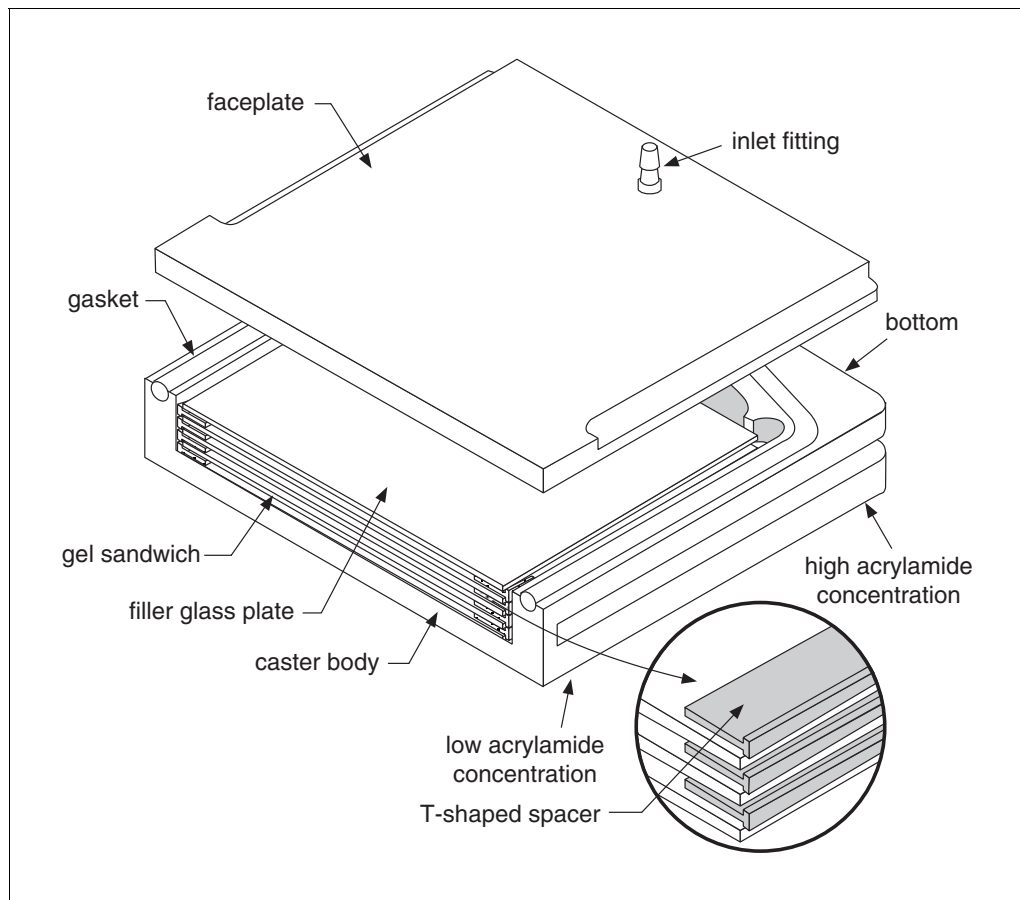


Figure 10.2A.4 Minigel sandwiches positioned in the multiple gel caster. Extra glass or acrylic plates or polycarbonate sheets are used to fill any free space in the caster and to ensure that the gel sandwiches are held firmly in place.

single-concentration gels. The Amersham Pharmacia Biotech spacers are T-shaped to prevent slipping. The flanged edge of the spacer must be positioned against the outside edge of the glass plate. Placing a sheet of wax paper between the gel sandwiches will help separate the sandwiches after polymerization.

2. Fit the gel sandwiches tightly in the multiple gel caster. Use an acrylic plate or polycarbonate separation sheet to eliminate any slack in the chamber.

Loosely fitting sandwiches in the caster will lead to unevenly cast gels, creating distortions during electrophoresis.

3. Place the front faceplate on the caster, clamp it in place against the silicone gasket, and verify alignment of the glass plates and spacers.
4. Prepare the separating gel solution as directed in Table 10.2A.1. For five 0.75-mm-thick gels, prepare ~30 ml solution (i.e., double the volumes listed).

To compute the total gel volume needed, multiply the area of the gel (e.g., 7.3×8.3 cm) by the thickness of the gel (e.g., 0.75 mm) and then by the number of gels in the caster. If needed, add ~4 to 5 ml of extra gel solution to account for the space around the outside of the gel sandwiches.

Do not add TEMED and ammonium persulfate until just before use.

5. Fill a 50-ml syringe with the separating gel solution and slowly inject it into the caster until the gels are 6 cm high, allowing 1.5 cm for the stacking gel.

- Overlay each gel with 100 μ l H₂O-saturated isobutyl alcohol. Allow the gels to polymerize for ~1 hr.

Pour the stacking gel

- Remove the isobutyl alcohol and rinse with 1 \times Tris-Cl/SDS, pH 8.8.

Stacking gels can be cast one at a time with the gel mounted on the electrophoresis unit, or all at once in the multiple caster.

- Practice placing a comb in the gel sandwiches before preparing the stacking gel solution. Press the comb against the rectangular or taller plate so that all teeth of the comb are aligned with the opening in the gel sandwich, then insert into the sandwich. Remove combs after practicing.
- Prepare the stacking gel solution (2 ml per gel) as directed in Table 10.2A.1. Fill a 10-ml syringe with stacking gel solution and inject the solution into each gel sandwich.
- Insert combs, taking care not to trap bubbles. Allow gels to polymerize 1 hr.
- Remove the front faceplate. Carefully pull the gels out of the caster, using a long razor blade to separate the sandwiches.

If the gels are left to polymerize for prolonged periods, they will be difficult to remove from the caster.

The gels can be stored tightly wrapped in plastic wrap with the combs left in place inside a sealable bag to prevent drying for ~1 week. Without the stacking gel, the separating gel can be stored for 2 to 3 weeks. Keep gels moist with 1 \times Tris-Cl/SDS, pH 8.8, at 4°C. Do not store gels in the multiple caster.

Prepare the sample, load the gel, and conduct electrophoresis

- Remove the combs and rinse the sample wells with 1 \times SDS electrophoresis buffer. Place a line indicating the bottom of each well on the front glass plate with a marker.
- Fill the upper and lower buffer chambers with 1 \times SDS electrophoresis buffer. The upper chamber should be filled to 1 to 2 cm over the notched plate.
- Prepare the protein sample and protein-standards mixture (see Basic Protocol 1, step 12).
- Load the sample using a micropipet. Insert the pipet tip through the upper buffer and into the well. The mark on the glass plate will act as a guide. Dispense the sample into the well.

For a complex mixture, 20 to 25 μ g protein in 10 μ l SDS sample buffer will give a strongly stained Coomassie blue pattern. Much smaller amounts (1 to 5 μ g) are required for highly purified proteins, and a 10- to 100-fold smaller amount of protein in the same volume (e.g., 10 μ l) is required for silver staining.

- Electrophorese samples at 10 to 15 mA per 0.75-mm gel until the dye front reaches the bottom of the gel (~1 to 1.5 hr).
- Disassemble the gel (see Basic Protocol 1, steps 23 to 26). Proceed with detection of proteins.

PREPARING MULTIPLE GRADIENT MINIGELS

Polyacrylamide gradients not only enhance the resolution of larger-format gels but also greatly improve protein separation in the small format. Casting gradient minigels one at a time is not generally feasible because of the small volumes used, but multiple gel casters make it easy to cast several small gradient gels at one time. The gels are cast from the bottom in multiple casters, with the light acrylamide solution entering first. This is the opposite of casting one gel at a time, in which the heavy solution enters from the top of the gel sandwich and flows down to the bottom.

Additional Materials (also see *Basic Protocol 2*)

Plug solution (see recipe)

Additional reagents and equipment for preparing gradient gels (see *Alternate Protocol 5*)

Set up the system and prepare the gel solutions

1. Assemble minigel sandwiches in the multiple gel caster as described for single-concentration minigels (see *Basic Protocol 2*, steps 1 to 3).
2. Set up the 30-ml gradient maker, magnetic stirrer, peristaltic pump (optional), and Tygon tubing as in Figure 10.2A.3. Connect the outlet of the 30-ml gradient maker to the inlet at the base of the front faceplate of the caster.

The monomer solution will be introduced through the inlet at the bottom of the front faceplate of the caster first, followed by progressively heavier solution.

3. Prepare light (Table 10.2A.9) and heavy (Table 10.2A.10) acrylamide gel solutions. Use ~12 ml of each solution for five 0.75-mm-thick minigels.

Adjust volumes if a different thickness or number of gels is needed. Do not add ammonium persulfate until just before use. Deaeration is not recommended for gradient gels.

4. With the outlet and interconnecting valve closed, add the heavy solution to the reservoir chamber. Briefly open the interconnecting valve to let a small amount of heavy solution through to the mixing chamber, clearing the valve of air.
5. Fill the mixing chamber with light solution. Add 4 μ l TEMED per 12 ml acrylamide solution to each chamber and mix with a disposable pipet.

Form the gradient and cast the gels

6. Turn on the magnetic stirrer. Open the interconnecting valve and allow the chambers to equilibrate. Then slowly open the outlet port to allow the solution to flow from the gradient maker to the multiple caster by gravity (a peristaltic pump may be used for better control). Adjust the flow rate to 3 to 4 ml/min.

Faster flow rates are possible and will also produce good gradients. However, a fast flow increases the potential for introduction of bubbles into the caster.

7. Close the outlet port as the last of the gradient solution leaves the mixing chamber, just before air enters the outlet tube. Fill the two chambers with plug solution and slowly open the outlet once again.
8. Allow the plug solution to push the acrylamide in the caster up into the plates. Close the outlet when the plug solution reaches the bottom of the plates.

A discontinuity between the bottom of the gels and the plug solution will be obvious.

9. Quickly add 100 μ l H₂O-saturated isobutyl alcohol to each gel sandwich. Let the gels polymerize undisturbed for ~1 hr.
10. Prepare and pour the stacking gel (see *Basic Protocol 2*, steps 9 and 10).

Disassemble the system

11. Disconnect the gradient maker, place the caster in a sink, and remove the front faceplate. The plug solution will drain out from the bottom of the caster.
12. Remove the gels (see Basic Protocol 2, step 11).

Gradient minigels can be stored as described for single-concentration minigels (see Basic Protocol 2, step 11 annotation). For instructions on preparing, loading, and running the gels, see Basic Protocol 2, steps 12 to 17.

REAGENTS AND SOLUTIONS

Use Milli-Q water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Anode buffer

121.1 g Tris base (0.2 M final)
500 ml H₂O
Adjust to pH 8.9 with concentrated HCl
Dilute to 5 liters with H₂O
Store at 4°C up to 1 month

Final concentration is 0.2 M Tris·Cl, pH 8.9.

Cathode buffer

12.11 g Tris base (0.1 M final)
17.92 g tricine (0.1 M final)
1 g SDS [0.1% final; recrystallization (see recipe) optional]
Dilute to 1 liter with H₂O
Do not adjust pH
Store at 4°C up to 1 month

Coomassie blue G-250 staining solution

200 ml acetic acid (20% final)
1800 ml H₂O
0.5 g Coomassie blue G-250 (0.025% final)
Mix 1 hr and filter (Whatman no. 1 paper)
Store at room temperature indefinitely

Phosphate/SDS electrophoresis buffer

Dilute 500 ml of 4× phosphate/SDS, pH 7.2 (Table 10.2A.7) with H₂O to 2 liters.
Store at 4°C up to 1 month.

Final concentrations are 0.1 M sodium phosphate (pH 7.2)/0.1% (w/v) SDS.

Phosphate/SDS sample buffer, 2× (for continuous systems)

0.5 ml 4× phosphate/SDS, pH 7.2 (Table 10.2A.7; 20 mM sodium phosphate final)
0.2 g SDS [2% final; recrystallization (see recipe) optional]
0.1 mg bromphenol blue (0.001% final)
0.31 g DTT (0.2 M final)
2.0 ml glycerol (2% final)
Add H₂O to 10 ml and mix

Plug solution

0.125 M Tris·Cl, pH 8.8 (APPENDIX 2)
50% (w/v) sucrose
0.001% (w/v) bromphenol blue
Store at 4°C up to 1 month

Recrystallized SDS (optional)

High-purity SDS is available from several suppliers, but for some sensitive applications (e.g., protein sequencing) recrystallization is useful. Commercially available electrophoresis-grade SDS is usually of sufficient purity for most applications.

Add 100 g SDS to 450 ml ethanol and heat to 55°C. While stirring, gradually add 50 to 75 ml hot H₂O until all SDS dissolves. Add 10 g activated charcoal (Norit 1, Sigma) to solution. After 10 min, filter solution through Whatman no. 5 paper on a Buchner funnel to remove charcoal. Chill filtrate 24 hr at 4°C and 24 hr at -20°C. Collect crystalline SDS on a coarse-frit (porosity A) sintered-glass funnel and wash with 800 ml -20°C ethanol (reagent grade). Repeat crystallization without adding activated charcoal. Dry recrystallized SDS under vacuum overnight at room temperature. Store in a desiccator over phosphorous pentoxide (P₂O₅) in a dark bottle.

If proteins will be electroeluted or electroblotted for protein sequence analysis, it may be desirable to crystallize the SDS twice from ethanol/H₂O (Hunkapiller et al., 1983).

SDS electrophoresis buffer, 5×

15.1 g Tris base (0.125 M final)
72.0 g glycine (0.96 M final)
5.0 g SDS [0.5% final; recrystallization (see recipe) optional]
H₂O to 1000 ml
Dilute to 1× or 2× for working solution, as appropriate

Do not adjust the pH of the stock solution, as the solution is pH 8.3 when diluted. Store at 0° to 4°C until use (up to 1 month).

SDS sample buffer, 2× (for discontinuous systems)

25 ml 4× Tris·Cl/SDS, pH 6.8 (Table 10.2A.1)
20 ml glycerol (20% final)
4 g SDS [4% final; recrystallization (see recipe) optional]
2 ml 2-ME or 3.1 g DTT (0.2% 2-ME or 0.2 M DTT final)
1 mg bromphenol blue (0.001% final)
Add H₂O to 100 ml and mix
Store in 1-ml aliquots at -70°C

To avoid reducing proteins to subunits (if desired), omit 2-ME or DTT (reducing agent) and add 10 mM iodoacetamide to prevent disulfide interchange.

SDS sample buffer, 6× (for discontinuous systems)

7 ml 4× Tris·Cl/SDS, pH 6.8 (Table 10.2A.1)
3.0 ml glycerol (30% final)
1 g SDS [10% final; recrystallization (see recipe) optional]
0.93 g DTT (0.6 M final)
1.2 mg bromphenol blue (0.012% final)
Add H₂O to 10 ml (if needed)
Store in 0.5-ml aliquots at -70°C

Tricine sample buffer, 2×

2 ml 4× Tris·Cl/SDS, pH 6.8 (Table 10.2A.1; 0.1 M)
2.4 ml (3.0 g) glycerol (24% final)
0.8 g SDS [8% final; recrystallization (see recipe) optional]
0.31 g DTT (0.2 M final)
2 mg Coomassie blue G-250 (0.02% final)
Add H₂O to 10 ml and mix

COMMENTARY

Background Information

Polyacrylamide gels form after polymerization of monomeric acrylamide into polymeric polyacrylamide chains and cross-linking of the chains by *N,N'*-methylenebisacrylamide. The polymerization reaction is initiated by the addition of ammonium persulfate, and the reaction is accelerated by TEMED, which catalyzes the formation of free radicals from ammonium persulfate. Because oxygen inhibits the polymerization process, deaerating the gel solution before the polymerization catalysts are added will speed up polymerization; deaeration is not recommended for the gradient gel protocols because slower polymerization facilitates casting of gradient gels.

Precast gels for commonly used vertical minigel and standard-sized SDS-PAGE apparatuses are available from several manufacturers (Table 10.2A.3). Flatbed (horizontal) isoelectric focusing (IEF) and SDS-PAGE gels are not listed. Amersham Pharmacia Biotech supplies a range of horizontal gels for a variety of applications and should be consulted for further information. When using precast gels, pay strict attention to shelf life. In general, manufacturers overrate the shelf life, and the sooner the gels are used, the better. When reasonably fresh, precast gels provide excellent resolution that is as good or better than a typical gel cast in the laboratory.

The most widely used method for discontinuous gel electrophoresis is the system described by Laemmli (1970). This is the denaturing (SDS) discontinuous method used in Basic Protocol 1. A discontinuous buffer system uses buffers of different pH and composition to generate a discontinuous pH and voltage gradient in the gel. Because the discontinuous gel system concentrates the proteins in each sample into narrow bands, the applied sample may be more dilute than that used for continuous electrophoresis.

In the discontinuous system the sample first passes through a stacking gel, which has large pores. The stacking gel buffer contains chloride ions (called the leading ions) whose electrophoretic mobility is greater than the mobility of the proteins in the sample. The electrophoresis buffer contains glycine ions (called the trailing ions) whose electrophoretic mobility is less than the mobility of the proteins in the sample. The net result is that the faster migrating ions leave a zone of lower conductivity between themselves and the migrating protein.

The higher voltage gradient in this zone allows the proteins to move faster and to “stack” in the zone between the leading and trailing ions. After leaving the stacking gel, the protein enters the separating gel. The separating gel has a smaller pore size, a higher salt concentration, and higher pH compared to the stacking gel. In the separating gel, the glycine ions migrate past the proteins, and the proteins are separated according to either molecular size in a denaturing gel (containing SDS) or molecular shape, size, and charge in a nondenaturing gel.

Proteins are denatured by heating in the presence of a low-molecular-weight thiol (2-ME or DTT) and SDS. Most proteins bind SDS in a constant-weight ratio, leading to identical charge densities for the denatured proteins. Thus, the SDS-protein complexes migrate in the polyacrylamide gel according to size, not charge. Most proteins are resolved on polyacrylamide gels containing from 5% to 15% acrylamide and 0.2% to 0.5% bisacrylamide (see Table 10.2A.1). The relationship between the relative mobility and log molecular weight is linear over these ranges (Fig. 10.2A.5). With the use of plots like those shown in Figure 10.2A.5, the molecular weight of an unknown protein (or its subunits) may be determined by comparison with known protein standards (Table 10.2A.2). In general, all of the procedures in this unit are suitable for radiolabeled and biotinylated proteins without modification.

Basic Protocol 1 relies on denaturing proteins in the presence of SDS and 2-ME or DTT. Under these conditions, the subunits of proteins are dissociated and their biological activities are lost. A true estimate of a protein's molecular size can be made by comparing the relative mobility of the unknown protein to proteins in a calibration mixture (Fig. 10.2A.5). Gradient gels (Alternate Protocol 5) simplify molecular-weight determinations by producing a linear relationship between log molecular weight of the protein and log % T over a much wider size range than single-concentration gels. Although percent acrylamide monomer is a more common measure of gel concentration, % T, the percentage of total monomer (acrylamide plus bisacrylamide) in the solution or gel, is used for molecular weight calculations in gradient gels. The % T of a stained protein is estimated assuming the acrylamide gradient is linear. For example, proteins in the gel shown in Figure 10.2A.6 were separated in a 5.1% to 20.5% T acrylamide gradient. The % T of the point

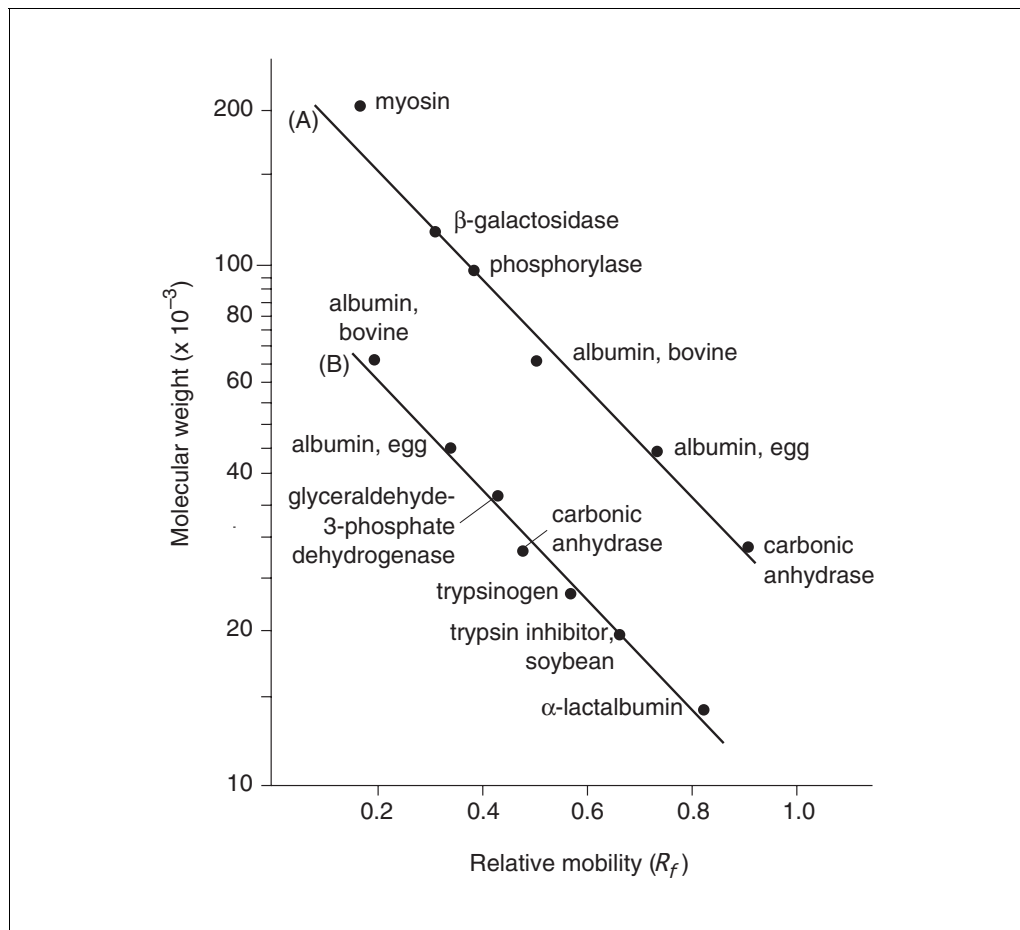


Figure 10.2A.5 Typical calibration curves obtained with standard proteins separated by nongradient denaturing (SDS) discontinuous gel electrophoresis based on the method of Laemmli (1970). (A) Gel with 7% polyacrylamide. (B) Gel with 11% polyacrylamide. (Redrawn with permission from Sigma.)

halfway through the resolving gel is 12.5% T. Simply plotting log molecular mass versus distance moved into the gel (or R_f) also produces a relatively linear standard curve over a fairly wide size range.

If two proteins have identical molecular sizes, they more than likely will not be resolved with one-dimensional SDS-PAGE, and two-dimensional SDS-PAGE should be considered. Unusual protein compositions can cause anomalous mobilities during electrophoresis (see Critical Parameters and Troubleshooting), but similar-sized proteins of widely different amino acid composition or structure may still be resolved from one another using one-dimensional SDS-PAGE. Purified protein complexes or multimeric proteins consisting of subunits of different molecular size will be resolved into constituent polypeptides. Comparison of the protein bands obtained under nonreducing and reducing conditions provides information about the molecular size of disulfide cross-linked com-

ponent subunits. The individual polypeptides can be isolated by electroelution or electroblotting, and the amino acid sequences can be determined.

Both the tricine (Schagger and von Jagow, 1987) and the modified Laemmli (Okajima et al., 1993) peptide separation procedures presented here (Alternate Protocols 1 and 2) are simple to set up and provide resolution down to 5 kDa. To separate peptides below 5 kDa, the tricine procedure must be modified by preparing a 16.5% T, 2.7% C resolving gel that uses a 10% T spacer gel between the stacking and resolving gel (Schagger and von Jagow, 1987). % C is the percentage of cross-linker (bisacrylamide) in the total monomer (acrylamide plus bisacrylamide).

Continuous electrophoresis, where the same buffer is used throughout the tank and gel, is popular because of its versatility and simplicity. The phosphate system described in Alternate Protocol 3 is based on that of Weber et al.

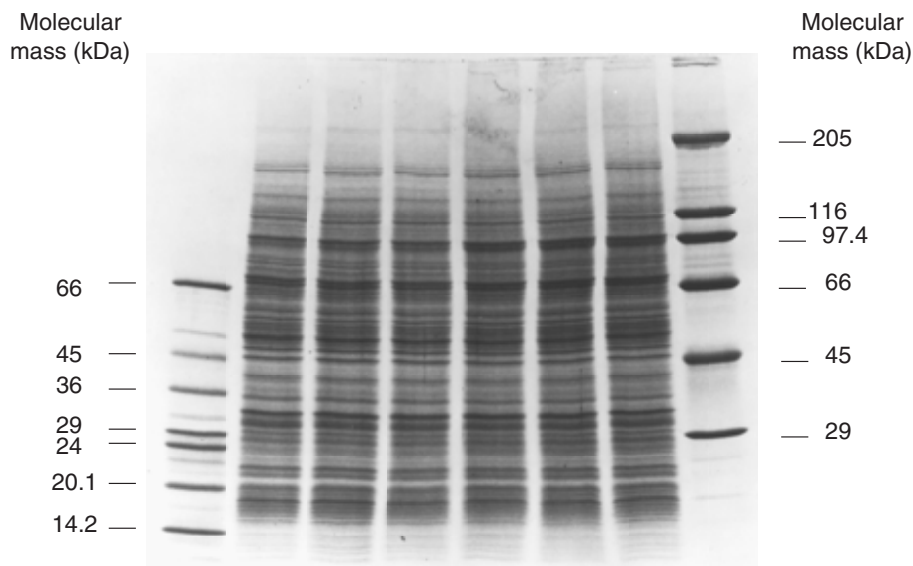


Figure 10.2A.6 Separation of membrane proteins by 5.1% to 20.5% T polyacrylamide gradient SDS-PAGE. Approximately 30 μ l of 1 \times SDS sample buffer containing 30 μ g of Alaskan pea (*Pisum sativum*) membrane proteins was loaded in wells of a 14 \times 14-cm, 0.75-mm-thick gel. Standard proteins were included in the outside lanes. The gel was run at 4 mA for ~15 hr.

(1972). Although unable to produce the high-resolution separations of the discontinuous SDS-PAGE procedures, continuous SDS-PAGE uses fewer solutions with one basic buffer and no stacking gel. Artifacts are also less likely to occur in continuous systems. Pepsin, for example, migrates anomalously on Laemmli-based discontinuous SDS-PAGE but has the expected mobility after electrophoresis in the phosphate-based continuous system described here. This is also true of cross-linked proteins.

Multiple gel casting (Support Protocols 1 to 3) is appropriate when gel-to-gel consistency is paramount or when the number of gels processed exceeds five a week. The variety of multiple gel casters, gradient makers, and inexpensive pumps available from major suppliers simplifies the process of casting gels in the laboratory. Alternatively, precast gradient gels are available for most major brands of gel apparatuses (Table 10.2A.3).

Minigels (Basic Protocol 2) are generally considered to be in the 8 \times 10-cm size range, although there is considerable variation in exact size. Every technique that is used on larger systems can be translated with little difficulty into the minigel format. This includes standard and gradient SDS-PAGE and separations for immunoblotting and peptide sequencing. Two-dimensional SDS-PAGE electrophoresis also adapts well, but here the limitation of separa-

tion area becomes apparent; for high-resolution separations, large-format gels are required. Gradient minigels (Support Protocol 3) are popular due to the combination of separation range and resolution (Matsudaira and Burgess, 1978). They are particularly useful for separation of proteins prior to peptide sequencing.

Mylar support (Gel Bond) provides a practical way of casting, running, and, staining extremely thin gels. When gels <0.75 mm thick are used, reagents have much better access both into and out of the gel, reducing staining time in both Coomassie blue and silver staining. Double and broadened images caused by differential migration of the protein across the thickness of the gel are minimized, improving resolution.

Critical Parameters and Troubleshooting

If an electrophoretically separated protein will be electroeluted or electroblotted for sequence analysis, the highest-purity reagents available should be used. If necessary, SDS can be purified by recrystallization following the procedure given in Reagents and Solutions.

If the gels polymerize too fast, the amount of ammonium persulfate should be reduced by one-third to one-half. If the gels polymerize too slowly or fail to polymerize all the way to the top, use fresh ammonium persulfate or increase the amount of ammonium persulfate by one-

third to one-half. The overlay should be added slowly down the spacer edge to prevent the overlay solution from crashing down and disturbing the gel interface.

After a separating gel is poured, it may be stored with an overlay of the same buffer used in the gel. Immediately prior to use, the stacking gel should be poured; otherwise, there will be a gradual diffusion-driven mixing of buffers between the two gels, which will cause a loss of resolution.

The protein of interest should be present in 0.2- to 1- μ g amounts in a complex mixture of proteins if the gel will be stained by Coomassie blue (UNIT 10.6). Typically, 30 to 50 μ g of a complex protein mixture in a total volume of <20 μ l is loaded on a 0.75-mm-thick slab gel (16 cm, 10 wells).

When casting multiple gradient gels, eliminate all bubbles in the outlet tubing of the gradient maker. If air bubbles get into the outlet tube, they may flow into the caster and then up through the gradient being poured, causing an area of distortion in the polymerized gel. Air bubbles are not so great a problem when casting single gradient gels from the top. As the gels are cast, the stirrer must be slowed so that the vortex in the mixing chamber does not allow air to enter the outlet.

Uneven heating of the gel causes differential migration of proteins, with the outer lanes moving more slowly than the center lanes (called smiling). Increased heat transfer eliminates smiling and can be achieved by filling the lower buffer chamber with buffer all the way to the level of the sample wells, by maintaining a constant temperature between 10° to 20°C, and by stirring the lower buffer with a magnetic stirrer. Alternatively, decrease the heat load by running at a lower current.

If the tracking dye band is diffuse, prepare fresh buffer and acrylamide monomer stocks. If the protein bands are diffuse, increase the current by 25% to 50% to complete the run more quickly and minimize band diffusion, use a higher percentage of acrylamide, or try a gradient gel. Lengthy separations using gradient gels generally produce good results (Fig. 10.2A.6). Check for possible proteolytic degradation that may cause loss of high-molecular-weight bands and create a smeared banding pattern.

If there is vertical streaking of protein bands, decrease the amount of sample loaded on the gel, further purify the protein of interest to reduce the amount of contaminating protein applied to the gel, or reduce the current by 25%.

Another cause of vertical streaking of protein bands is precipitation, which can sometimes be eliminated by centrifuging the sample or by reducing the percentage of acrylamide in the gel.

Proteins can migrate faster or slower than their actual molecular weight would indicate. Abnormal migration is usually associated with a high proportion of basic or charged amino acids (Takano et al., 1988). Other problems can occur during isolation and preparation of the protein sample for electrophoresis. Proteolysis of proteins during cell fractionation by endogenous proteases can cause subtle band splitting and smearing in the resulting electrophoretogram (electrophoresis pattern). Many endogenous proteases are very active in SDS sample buffers and will rapidly degrade the sample; thus, first heating the samples to 70° to 100°C for 3 min is recommended.

In some cases, heating to 100°C in sample buffer will cause selective aggregation of proteins, creating a smeared layer of Coomassie blue-stained material at the top of the gel (Gallagher and Leonard, 1987). To avoid heating artifacts and also prevent proteolysis, the use of specific protease inhibitors during protein isolation and/or lower heating temperatures (70° to 80°C) have been effective (Dhugga et al., 1988).

Although continuous gels suffer from poor band sharpness, they are less prone to artifacts caused by aggregation and protein cross-linking. If streaking or aggregation appear to be a problem with the Laemmli system, then the same sample should be subjected to continuous SDS-PAGE to see if the problem is intrinsic to the Laemmli gel or the sample.

If the protein bands spread laterally from gel lanes, the time between applying the sample and running the gel should be reduced in order to decrease the diffusion of sample out of the wells. Alternatively, the acrylamide percentage should be increased in the stacking gels from 4% to 4.5% or 5% acrylamide, or the operating current should be increased by 25% to decrease diffusion in the stacking gel. Use caution when adding 1 \times SDS electrophoresis buffer to the upper buffer chamber. Samples can get swept into adjacent wells and onto the top of the well arm.

If the protein bands are uneven, the stacking gel may not have been adequately polymerized. This can be corrected by deaerating the stacking gel solution thoroughly or by increasing the ammonium persulfate and TEMED concentrations by one-third to one-half. Another cause

of distorted bands is salt in the protein sample, which can be removed by dialysis, gel filtration, or precipitation. Skewed protein bands can be caused by an uneven interface between the stacking and separating gels, which can be corrected by starting over and being careful not to disturb the separating gel while overlaying with isobutyl alcohol.

If a run takes too long, the buffers may be too concentrated or the operating current too low. If the run is too short, the buffers may be too dilute or the operating current too high.

If double bands are observed, the protein may be partially oxidized or partially degraded. Oxidation can be minimized by increasing the 2-ME concentration in the sample buffer or by preparing a fresh protein sample. If fewer bands than expected are observed and there is a heavy protein band at the dye front, increase the acrylamide percentage in the gel.

Anticipated Results

Polyacrylamide gel electrophoresis done under denaturing and reducing conditions should resolve any two proteins, except two of identical size. Resolution of proteins in the presence of SDS is a function of gel concentration and the size of the proteins being separated. Under nondenaturing conditions, the biological activity of a protein will be maintained.

Comparison of reducing and nonreducing denaturing gels can also provide valuable information about the number of disulfide cross-linked subunits in a protein complex. If the subunits are held together by disulfide linkages, the protein will separate in denaturing gels as a complex or as smaller-sized subunits under nonreducing or reducing conditions, respectively. However, proteins separated on nonreducing denaturing gels appear more diffuse and exhibit less overall resolution than those separated on reducing gels.

Gradient gels provide superior protein-band sharpness and resolve a larger size range of proteins, making them ideal for most types of experiments in spite of being more difficult to prepare. Molecular-weight calculations are simplified because of the extended linear relationship between size and protein position in the gel. Increased band sharpness of both high- and low-molecular-weight proteins on the same gel greatly simplifies survey experiments, such as gene expression studies where the characteristics of the responsive protein are not known. Furthermore, the increased resolution dramatically improves autoradiographic analysis. Preparation of gradient gels is straightfor-

ward, although practice with gradient solutions containing dye is recommended. The gradient gels can be stored for several days at 0° to 4°C before casting the stacking gel.

Time Considerations

Preparation of separating and stacking gels requires 2 to 3 hr. Gradient gels generally take 5 min to cast singly. Casting multiple single-concentration gels requires an additional 10 min for assembly. Casting multiple gradient gels takes 15 to 20 min plus assembly time. It takes 4 to 5 hr to run a 14 × 14-cm, 0.75-mm gel at 15 mA (70 to 150 V), and 3 to 4 hr to run a 0.75-mm gel at 20 mA (80 to 200 V). Overnight separations of ~12 hr require 4 mA per 0.75-mm gel. It takes 4 to 5 hr to run a 1.5-mm gel at 30 mA. Electrophoresis is normally performed at 15° to 20°C, with the temperature held constant using a circulating water bath. For air-cooled electrophoresis units, lower currents and thus longer run times are recommended.

It takes ~1 hr to run a 0.75-mm minigel at 20 mA (100 to 120 V). Separation times are not significantly different for gradient minigels.

Literature Cited

- Dhugga, K.S., Waines, J.G., and Leonard, R.T. 1988. Correlated induction of nitrate uptake and membrane polypeptides in corn roots. *Plant Physiol.* 87:120-125.
- Gallagher, S.R. and Leonard, R.T. 1987. Electrophoretic characterization of a detergent-treated plasma membrane fraction from corn roots. *Plant Physiol.* 83:265-271.
- Hunkapiller, M.W., Lujan, E., Ostrander, F., and Hood, L.E. 1983. Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. *Methods Enzymol.* 91:227-236.
- Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.
- Matsudaira, P.T. and Burgess, D.R. 1978. SDS microslab linear gradient polyacrylamide gel electrophoresis. *Anal. Biochem.* 87:386-396.
- Okajima, T., Tanabe, T., and Yasuda, T. 1993. Nonurea sodium dodecyl sulfate-polyacrylamide gel electrophoresis with high-molarity buffers for the separation of proteins and peptides. *Anal. Biochem.* 211:293-300.
- Schagger, H. and von Jagow, G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* 166:368-379.
- Sigma. Molecular weight markers for proteins kit (Technical Bulletin MWS-877L). Sigma Chemical Company, St. Louis, Mo.

Takano, E., Maki, M., Mori, H., Hatanaka, N., Marti, T., Titani, K., Kannagi, R., Ooi, T., and Murachi, T. 1988. Pig heart calpastatin: Identification of repetitive domain structures and anomalous behavior in polyacrylamide gel electrophoresis. *Biochemistry* 27:1964-1972.

Weber, K., Pringle, J.R., and Osborn, M. 1972. Measurement of molecular weights by electrophoresis on SDS-acrylamide gel. *Methods Enzymol.* 26:3-27.

Key Reference

Hames, B.D. and Rickwood, D. (eds.) 1990. Gel Electrophoresis of Proteins: A Practical Approach, 2nd ed. Oxford University Press, New York.

An excellent book describing gel electrophoresis of proteins.

Contributed by Sean R. Gallagher
Motorola Corporation
Tempe, Arizona

One-Dimensional Electrophoresis Using Nondenaturing Conditions

Nondenaturing or “native” electrophoresis—i.e., electrophoresis in the absence of denaturants such as detergents and urea—is an often-overlooked technique for determining the native size, subunit structure, and optimal separation of a protein. Because mobility depends on the size, shape, and intrinsic charge of the protein, nondenaturing electrophoresis provides a set of separation parameters distinctly different from mainly size-dependent denaturing sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE; *UNITS 10.2A*) and charge-dependent isoelectric focusing (IEF; *UNITS 10.3 & 10.4*). Two protocols are presented below. Continuous PAGE (see Basic Protocol) is highly flexible, permitting cationic and anionic electrophoresis over a full range of pH. The discontinuous procedure (see Alternate Protocol) is limited to proteins negatively charged at neutral pH but provides high resolution for accurate size calibration.

CONTINUOUS ELECTROPHORESIS IN NONDENATURING POLYACRYLAMIDE GELS

BASIC PROTOCOL

Separation of proteins by nondenaturing electrophoresis requires the same type of equipment used for denaturing slab gels (*UNIT 10.2A*) and is adaptable to a range of gel sizes (e.g., from 7.3 × 8.3–cm minigels to 14 × 16–cm full-size gels) and matrix types (e.g., single-concentration and gradient gels). This protocol outlines straightforward procedures for making acrylamide solutions, casting separating gels (stacking gels are omitted), loading samples, and conducting electrophoresis. Continuous systems, although flexible, do not give the high-resolution separation found in discontinuous systems (see Alternate Protocol).

Separation in a continuous system (i.e., in which the same buffer is used for preparing acrylamide solutions and filling electrophoresis chambers) is governed by pH, and this protocol describes four types of buffers useful over discrete ranges from pH 3.7 to pH 10.6. Use of unadjusted acetic acid gel buffer can extend the range to pH 2.0. The choice of pH and thus the buffer system depends on the protein being studied (i.e., its isoelectric point) and often must be determined empirically. In general, the system should be between pH 5.0 and 8.0 for optimal results. Extremes of pH can lead to precipitation or denaturation of the protein. Acrylamide concentrations are empirically determined, but the higher the percent acrylamide, the sharper the protein bands.

It is important to include native protein standards in the electrophoresis runs. Several manufacturers supply standards for isoelectric focusing that are also suitable for native electrophoresis. The standards have a range of isoelectric points and will carry a net positive, negative, or zero charge depending on the pH of the gel system. Alternatively, Sigma supplies a standard kit that is useful for calculating molecular weights under neutral pH, nondenaturing conditions. The samples and standard proteins, should be used at concentrations of ~1 to 2 μg/μl.

Materials

- 4× acetic acid gel buffer (200 mM acetic acid, pH 3.7 to 5.6; see recipe)
- 4× phosphate gel buffer (400 mM sodium phosphate, pH 5.8 to 8.0; see recipe)
- 4× Tris gel buffer (200 mM Tris·Cl, pH 7.1 to 8.9; see recipe)
- 4× glycine gel buffer (200 mM glycine, pH 8.6 to 10.6; see recipe)
- 300 mM sodium sulfite (0.38 g in 10 ml H₂O; used in acetic acid gel preparation)
- Protein samples to be analyzed

Contributed by Sean R. Gallagher

Current Protocols in Molecular Biology (1999) 10.2B.1-10.2B.11

Copyright © 1999 by John Wiley & Sons, Inc.

Two-Dimensional Gel Electrophoresis Using the ISO-DALT System

Two high-resolution electrophoretic procedures (isoelectric focusing and SDS-polyacrylamide gel electrophoresis) are combined to provide much greater resolution than either of the individual procedures (O'Farrell, 1975). Solubilized proteins are first separated according to their isoelectric point by isoelectric focusing in a 1.5×140 -mm tube gel. This first-dimension gel is then applied to the top of an SDS-polyacrylamide slab gel and electrophoresed. The proteins in the first-dimension gel migrate into the second-dimension gel where they are further separated on the basis of their molecular size. The ISO-DALT system was specifically designed for running multiple high-resolution two-dimensional gels at one time.

FIRST-DIMENSION (ISOELECTRIC FOCUSING) GELS

BASIC PROTOCOL

Materials

Urea
Ampholytes
30% acrylamide/1.8% bisacrylamide
Nonidet P-40 (NP-40)
TEMED (*N, N, N', N'*-tetramethylethylenediamine)
10% ammonium persulfate in H₂O
0.85% phosphoric acid in H₂O
10 M NaOH in H₂O
Bromphenol blue solution
25-ml, 10-ml, 1-ml, and 50- μ l syringes
Vacuum pump with cold trap
0.2- or 0.45- μ m filter capsule (Acrodisk, Gelman Sciences)
Single-edge razor blades
22-G hypodermic needle (2-in. long)
1-dram vials
First-dimension gel apparatus (ISS or Hoefer Scientific)

NOTE: Distilled, deionized water should be used throughout this protocol.

Pour isoelectric focusing gels

1. Assemble parts (clean and dry) of a 20-gel first-dimension gel apparatus.
2. In a vacuum flask (~100 ml) place 8.25 g urea, 6 ml water, 0.75 to 0.8 ml ampholytes (BDH pH 4.0 to 8.0), and 2 ml of 30% acrylamide/1.8% bisacrylamide. Swirl under warm tap water (~37°C) until all the urea dissolves. This makes 15 ml of gel solution.
3. Degas by applying vacuum for 3 to 4 min, add 0.3 ml NP-40, and swirl until dissolved.
4. Pour solution into a 25-ml syringe fitted with a 0.2- or 0.45- μ m filter capsule and force the liquid through the filter.
5. Add 10 μ l TEMED and 70 μ l of 10% ammonium persulfate; mix by swirling.
This solution will polymerize in ~2¹/₂ min, so work quickly!
6. Fill gel trough with gel solution and carefully overlay the gel solution with water from a pipet.
7. Slowly lower the gel tube assembly into the lower buffer tank containing ~2 liters of water. As gel tube assembly is lowered into water, gel tubes will fill uniformly with

acrylamide solution. Add water to the lower buffer tank to bring the acrylamide solution to the desired height in gel tubes.

For future reference, place a mark on the lower buffer tank to indicate the final water level. After the gel tube assembly has been removed, place another mark on the lower buffer tank to indicate initial water level.

Prepare isoelectric focusing gels

8. Allow at least 1 hr for polymerization, then remove the gel tube assembly from the lower buffer tank. Carefully remove the gel trough (but not the gel retainer), and using a new single-edge razor blade, cut between the retainer and the tubes. Cut flush with the bottoms of the gel tubes. The gel retainer can now be removed, and excess acrylamide rinsed from the gel tubes with deionized tap water.
9. In a vacuum flask degas 2 liters of water. To 1800 ml of the degassed water, add 200 ml of 0.85% phosphoric acid and pour the solution into the lower buffer tank. Place the gel tube assembly in the lower buffer tank.
10. Add 0.3 ml of 10 M NaOH to 150 ml degassed water, mix, and pour into the upper buffer tank. The gel tubes can be filled with the upper buffer from a 10-ml syringe fitted with a 2-in., 22-G needle.

Take care that no air bubbles remain in the gel tubes.

11. Connect the power supply (negative lead goes to upper tank) and pefocus the gels 1 hr at 200 V.

Load and run isoelectric focusing gels

12. Disconnect the power supply and load solubilized samples (see support protocol) by layering them on the top of each gel with a 50- μ l syringe.

Up to 30 μ l of sample can be applied.

13. Reconnect the power supply and set output for 800 V constant voltage. Run 16 to 18 hr (12,800 to 14,400 V-hr).

Optimum run time and voltage should be determined experimentally, but 12,800 V-hr usually gives satisfactory results.

14. End isoelectric focusing (IEF) by turning power supply off and disconnecting leads.

Extrude isoelectric focusing gels

15. Remove gel tube assembly from lower buffer tank and pour out upper buffer. Using a 50- μ l syringe, place \sim 1 μ l bromphenol blue solution on the top of each gel.

The dye marks the basic end of the gel and acts as a tracking dye in the second-dimension gel.

16. Remove the gel tube apparatus from its holder and set it on two short pieces of 4 \times 4 timber or similarly shaped objects in order to elevate the unit enough to allow easy access to the bottoms of the tubes.

17. The IEF gels can be extruded into numbered vials (1 dram) by water pressure from a 1-ml syringe fitted with a small (200- μ l) Eppendorf pipet tip.

Approximately $\frac{1}{4}$ in. of the large end of the tip must be removed so that it will fit on the syringe. With a little practice the gels can be extruded into the vials with almost no liquid entering the vials. The gels can be stored at -70°C for later use or can be used immediately.

18. Clean IEF tubes by soaking overnight in chromic acid cleaning solution; then rinse under deionized tap water 10 to 15 min.

A beaker large enough for the IEF tubes, but not the plastic legs, to fit in can be used as a chromic acid bath. Chromic acid can severely damage the plastic.

CAUTION: Face and eye protection should be worn when pouring chromic acid solutions.

19. Remove excess water from IEF tubes with vacuum suction and allow tubes to air dry.

SECOND-DIMENSION (GRADIENT) GELS

BASIC PROTOCOL

Materials

30% acrylamide/0.8% bisacrylamide
L10 buffer
L20 buffer
10% sodium dodecyl sulfate (SDS)
TEMED
10% ammonium persulfate solution
Blue glycerol (working solution)
Isobutyl alcohol, H₂O saturated
DALT tank buffer salts
Equilibration buffer
Agarose solution
Protein molecular weight standards (Table 10.2.2)
Gel plates
Gel identification tags
Gel casting box (ISS or Hoefer Scientific)
Carpenter's level
Gradient maker (ISS or Hoefer Scientific)
DALT tank (ISS or Hoefer Scientific)
Loading lectern (ISS or Hoefer Scientific)
Nylon screen

The ISO-DALT gradient maker and gel casting box are shown in Figure 10.3.1.

NOTE: Distilled, deionized water should be used throughout this protocol.

Prepare the casting box and gradient maker

1. Remove the front plate from the gel casting box and fill the box with 24 gel plates.
2. Replace the front plate and place a gel identification tag into the right-hand corner of each plate.

Consecutive numbers typed on Whatman No. 1 filter paper make convenient identification tags.

3. Mount the casting box on its stand, level it, and then tilt the box ~30° toward the side the acrylamide enters.

In new versions of the casting box, the gels are poured with the box level at all times.

4. Check the entire gel casting system to be sure that it is clean and free of polymerized acrylamide.
5. Close 20% acrylamide chamber (valve A) and place clamps at points 1, 2, and 3 on the delivery line from the gradient maker.

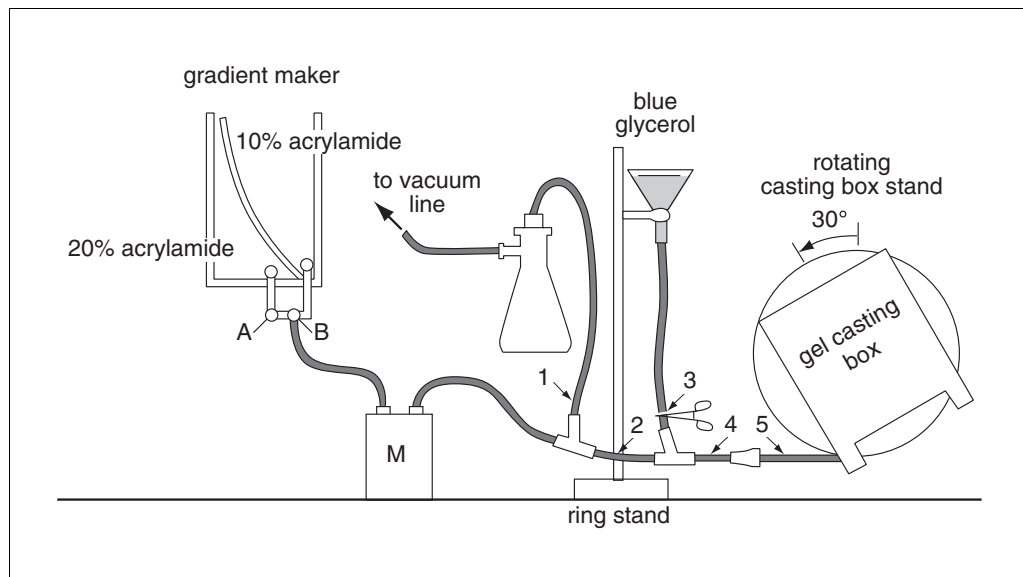


Figure 10.3.1 Gradient maker and gel casting box. The gradient maker is shown on the left. Valve A controls the flow of 20% acrylamide solution; valve B controls the flow of 10% acrylamide. The acrylamide solutions flow through the delivery line, in which a mixer (M) is used to ensure complete mixing of the two solutions before entering the casting box. Clamping points on the delivery line are indicated by arrows 2, 4, and 5. Arrow 1 indicates the clamping point on the vacuum line and arrow 3 the clamping point (clamp in place) on the blue glycerol line.

Prepare acrylamide solutions

6. Acrylamide solutions for pouring 24 plates are prepared in 1-liter vacuum flasks by making additions as follows:

20% acrylamide solution:
 400 ml 30% acrylamide/
 0.8% bisacrylamide
 200 ml L20 buffer
 6 ml 10% SDS
 30 μ l TEMED
 4 ml 10% ammonium persulfate

10% acrylamide solution:
 247 ml 30% acrylamide/
 0.8% bisacrylamide
 503 ml L10 buffer
 7.5 ml 10% SDS
 140 μ l TEMED
 9 ml 10% ammonium persulfate

7. Add acrylamide/bisacrylamide and L buffer to each flask and degas before adding SDS.
8. Add TEMED and swirl.

The amount of TEMED may need to be adjusted, depending on temperature and acrylamide supplier.

9. Add ammonium persulfate and swirl.

Pour gradient gels

10. Pour the 10% acrylamide solution into the right side of the gradient maker.
11. Open valve A just enough to allow 10% solution to displace the air from the tube between the two sides of the gradient maker.
12. Fill the left side of the gradient maker with the desired amount of 20% acrylamide solution.

13. Place the end of the delivery line in the flask containing the 10% acrylamide solution, release clamp at point 2 in order to flush all air bubbles from the tubing, and replace clamp.

Pinching the line at the site of air bubbles will aid in their removal.

14. Connect the delivery line to the casting box.
15. Refill the right side of the gradient maker with desired amount of 10% solution.
Normally more 10% acrylamide solution than 20% acrylamide solution is used in order to form a layer of 10% acrylamide at the top of the gels.
16. Turn on mixer and close 10% acrylamide chamber (valve B).
17. Release the clamp at point 2.
18. Open valve B to deliver desired flow.
19. As the level of the 10% solution approaches that of the 20% solution, open valve A (20% acrylamide chamber) so that the level of both solutions stays the same.

The gradient maker should empty in 4 to 6 min.

20. Begin bringing the casting box back to level when the acrylamide solution nears the upper corner of the gel plates.

The box should be brought to level in small increments. The final position should be checked with a carpenter's level.

21. When the level of acrylamide in the casting box reaches the desired height (allow ~1.5 cm for addition of blue glycerol), clamp the delivery line at point 2.
22. Release the clamp on the vacuum line (point 1) to remove any remaining acrylamide solution from the gradient maker.
23. Slowly release the clamp on the blue glycerol line (point 3). Bring the level of the blue glycerol to slightly below the bottom of the gel plates.

While the blue glycerol is flowing into the casting box, release the clamp at point 2 just enough to displace the acrylamide from the delivery line.

24. Final adjustment of the blue glycerol level in the casting box is made by raising or lowering the reservoir funnel. Do not clamp the line between the funnel and the casting box before polymerization of the acrylamide is complete.
25. Apply a thin layer of isobutyl alcohol (water saturated) to the tops of the gels (a glass plant spritzer works well).

The gels will be ready to be removed from the box in 1½ hr.

26. Turn off the vacuum and place the waste line in sink.
27. Rinse the gradient maker with at least 2 liters of water.
28. Dispose of the acrylamide in the vacuum flask trap.

Remove plates from casting box

29. Place clamps at points 2, 3, 4, and 5.
30. Disconnect the delivery line from the casting box between points 4 and 5.
31. Place the casting box in the sink and release the clamp at point 5.
32. Remove the wing nuts and front plate from the casting box.

33. After all liquid has drained from the box, set it on the side of the sink.
34. Remove the plates from the casting box by inserting a single-edge razor blade between the back of the first plate and the front of the second plate, then prying forward.

Rubber gloves such as those used for dishwashing should be worn.

35. As each plate is removed, wash it under tap deionized water to remove excess acrylamide and isobutyl alcohol from the top of the gel.
36. Place washed gels on their side in a dish drainer so that excess water will drain from the top of the plates.

Gels may be used immediately or stored refrigerated for several days in an airtight plastic container.

Load and run second-dimension gels

37. Add DALT tank buffer salts to 3 liters of water and stir on a magnetic stirrer.
38. Pour the concentrated DALT tank buffer into a DALT tank and fill to appropriate level with water.
39. If frozen first-dimension (IEF) gels are to be used, they should be thawed on the laboratory bench before adding equilibration buffer.

It is best not to thaw more first-dimension gels at one time than can be loaded in ~20 min, because the protein bands will diffuse.

40. Add 1.5 ml equilibration buffer to each vial and rock for 0 to 30 min.

Some proteins are extracted during long (30 min) equilibration, while others run better. For most samples, 1 to 2 min of exposure to equilibration buffer is sufficient.

41. Pour the gel out of the vial onto a piece of stiff nylon screen large enough to cover a 200-ml beaker.
42. Transfer the gel to the top of the loading lectern and, using a metal spatula, lay the gel out straight (blue-stained—i.e., basic—end on your right) near the front edge of the lectern.

Have a second-dimension gel in place on the front of the lectern.

43. With one hand, dispense a small amount (~100 μ l) of agarose solution from a Pasteur pipet onto one corner of the top of the gel. With the other hand, immediately grasp both the lectern and the gel and tip the lectern ~45° so that the agarose will flow across the entire surface of the gel; then set the lectern back on the laboratory bench.
44. Move the first-dimension gel onto the top of the second-dimension gel by scooting it forward with a spatula.

Take care that no air bubbles get trapped between the first- and second-dimension gels and that the first-dimension gel does not get stretched or compressed.

45. Pipet a thin layer of agarose solution over the first-dimension gel.

Protein molecular weight markers may be run as one-dimensional separations on the sides of the second-dimension gel. Mix marker proteins 1:1 with 1% agarose and draw the hot solution up in a 0.2-ml (1.5 mm diameter) glass pipet. Extrude a short piece of the solidified agarose and apply it the same as the first-dimension gel.

46. After the agarose has solidified, place the gel in the DALT tank for electrophoresis.

For best results, put the rubber hinge on the gel plate in the down position; sample on the cathode (-) side.

47. Attach power supply leads to the DALT tank and turn on the power supply.

At constant voltage of 220 to 260 V, the run will require 4 to 6 hr depending on actual voltage used and tank buffer temperature. For an overnight run of ~16 hr, apply 100 to 150 V with the tank temperature near 10°C.

48. Remove the gels when the tracking dye reaches the end of the gel.

49. Pry the plates open with a spatula and carefully remove the gel.

The gels may stick along the spacers. Thus, it is advisable to free the gel from the spacers by running a thin spatula (or razor blade) between the gel and spacers before peeling the gel from the plate.

50. Place gels in fixative or staining solution (UNIT 10.6) or use autoradiography (APPENDIX 3) to detect radiolabeled protein spots. A representative two-dimensional gel pattern is shown in Figure 10.3.2.

ISOELECTRIC FOCUSING OF BASIC AND VERY ACIDIC PROTEINS

First-dimension isoelectric focusing gels run in the standard way using broad-range ampholytes will resolve proteins with pI values between ~3.8 and 8. Proteins that are more basic or more acidic may be resolved by modifying the first-dimension gels as described below.

ALTERNATE
PROTOCOLS

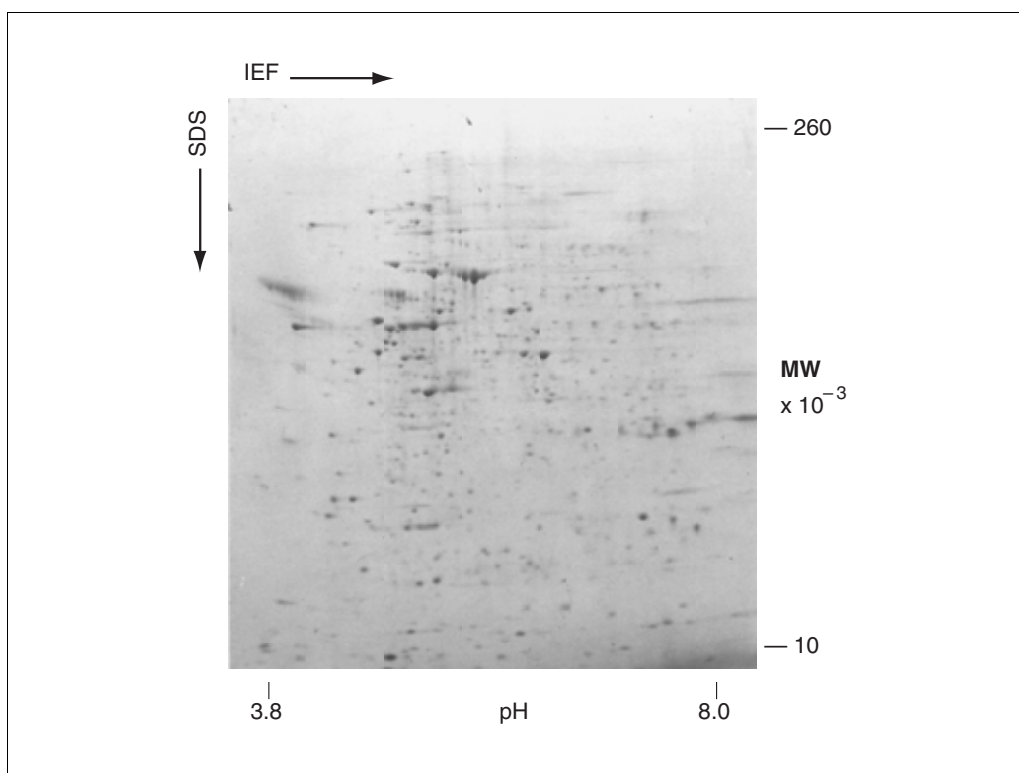


Figure 10.3.2 Coomassie-stained proteins separated by two-dimensional gel electrophoresis. In this gel, a broad range pH gradient (pH 3.5 to 9) was used for isoelectric focusing in the first dimension, and a 10 to 20% gradient SDS/polyacrylamide gel was used in the second dimension.

Analysis of
Proteins

10.3.7

For Basic Proteins

The procedure is the same as described in the basic protocol for first-dimension gels with the following exceptions in indicated steps. *Samples must be prepared in urea solubilization buffer for basic proteins.*

2. Use 2 to 11 pH ampholytes in gel solution.
9. Fill lower buffer tank with 2 liters of degassed 20 mM NaOH.
10. Fill upper tank with 250 ml of degassed 10 mM phosphoric acid.
11. Do not refocus gels. The anodic and cathodic solutions are reversed, so the negative lead now goes to the lower tank and the positive to the upper tank.
12. Overlay samples with degassed 4 M urea to protect proteins from phosphoric acid.
13. Focus 1 hr at 400 V and then 4 to 5 hr at 800 V for a total of ~4000 V-hr.

The blue glycerol applied to the top of the gel at the end of the run will now mark the acidic end, so load first-dimension gels on the second-dimension gel with the blue end to the left.

For Very Acidic Proteins

The procedure is the same as described in the basic protocol for first-dimension gels with the following exceptions in indicated steps:

2. Prepare the gel solution as follows:
 - 8.25 g urea
 - 2 ml 30% acrylamide/1.8% bisacrylamide
 - 1 ml of pH 2.5 to 4 ampholyte
 - 0.3 ml of pH 2 to 11 ampholyte
 - 5.5 ml H₂O
5. Degas the gel solution and add the following:
 - 0.3 ml NP-40
 - 90 µl 10% ammonium persulfate
 - 10 µl TEMED
9. Make lower buffer by adding 3 ml concentrated H₂SO₄ to 2 liters of water.
10. Make upper buffer by adding 1 ml of 2 to 11 pH ampholyte to 40 ml of degassed water.
13. Run 3600 to 4000 V-hr.

SUPPORT PROTOCOL

SAMPLE SOLUBILIZATION

Materials

Urea solubilization buffer (standard)
Urea solubilization buffer (for basic proteins)
SDS solubilization buffer

Dounce homogenizer
Pestles A and B

1. Weigh and place tissue samples in a Dounce homogenizer.
2. Add 1.5 ml SDS or urea solubilization buffer for each 100 mg of tissue.

3. Homogenize 50 strokes with pestle B, then 50 strokes with pestle A.

Cultured cells and body fluids do not normally require homogenization.

Enzymatic digestion may be required in addition to homogenization in order to remove cell walls from bacteria, yeast, or plant cells.

4. Let stand a few minutes, then transfer an aliquot to a centrifuge tube. Centrifuge 1 hr at $>200,000 \times g$.

If using SDS solubilization buffer, the sample should be placed in boiling water for 5 min before centrifugation.

Centrifugation should be done just prior to loading the sample on the first-dimension gel. Failure to centrifuge tissue samples is likely to cause plugging of the IEF gels, resulting in poor entry of proteins. The Beckman 42.2 Ti rotor, which accommodates 74 tubes of 200 μ l each, works well.

REAGENTS AND SOLUTIONS

NOTE: Distilled, deionized water should be used to prepare all reagents.

Acrylamide/bisacrylamide solutions

<i>30% acrylamide/1.8% bisacrylamide (first dimension):</i>	<i>30% acrylamide/0.8% bisacrylamide (second dimension):</i>
30 g acrylamide	300 g acrylamide
1.8 g bisacrylamide	8 g bisacrylamide
H ₂ O to 100 ml total volume	H ₂ O to 1 liter total volume

Store at 4°C in tightly capped amber bottle after filtering through 0.2- to 0.45- μ m filter. Discard after 30 days since acrylamide gradually hydrolyzes to acrylic acid and ammonia.

CAUTION: Acrylamide monomer is neurotoxic. Gloves should be worn while handling the solution, and the solution should not be pipetted by mouth.

Agarose solution

0.25 g agarose (Bio-Rad standard low- M_r)
50 ml DALT tank buffer

Add dry agarose to the DALT tank buffer in a 125-ml Erlenmeyer flask and place in boiling water bath.

10% ammonium persulfate in H₂O

10 g ammonium persulfate (Bio-Rad)
H₂O to 100 ml total volume
Store refrigerated ~2 weeks

Bromphenol blue solutions

<i>Concentrated:</i>	<i>Blue glycerol (working solution):</i>
50% aqueous glycerol (v/v) saturated with bromphenol blue	50% aqueous glycerol (v/v) + 5 to 10 ml/l concentrated bromphenol blue

DALT tank buffer salts

<i>For 36.7-liter tank:</i>	<i>For 26-liter tank:</i>
110.0 g Tris base	84.5 g Tris base
528.5 g glycine	375 g glycine
36.7 g SDS	26 g SDS (Bio-Rad)

It is convenient to use 1-quart plastic boxes to store measured buffer salt.

Equilibration buffer

200 ml H₂O
7.5 g Tris base
50 ml glycerol
10.5 g SDS

Dissolve Tris base in H₂O and bring to pH 6.8 with 6 N HCl. Add other ingredients and bring to 500 ml final volume with H₂O. To use, add 2 ml 2-mercaptoethanol (Kodak) or 130 mg of dithiothreitol (DTT) to 98 ml equilibration buffer. Can be stored refrigerated for ~3 weeks after addition of reducing agent.

L buffer

400 g Tris base
200 g Tris·Cl
~120 ml 6 N HCl
Adjust pH between 8.5 and 8.6 at room temperature with 6 N HCl
Add H₂O to 3 liters

L10 buffer

3 parts L buffer plus 5 parts water.

L20 buffer

3 parts L buffer plus 1 part glycerol.

SDS solubilization buffer

1.0 g CHES (2-[*N*-cyclohexylamino]ethanesulfonic acid)
2.0 g SDS
1.0 g dithiothreitol
10 ml glycerol
H₂O to 100 ml total volume
Adjust to pH 9.5 with NaOH
Store aliquots at -20°C

Standard urea solubilization buffer

54 g urea
4 ml NP-40
10 ml 20% (w/v) stock ampholyte (pH 9 to 11)
2 ml 2-mercaptoethanol (Kodak)
H₂O to 100 ml total volume
If pH is less than 9.5, add NaOH
Store aliquots at -20°C

Urea solubilization buffer, for basic proteins

54 g urea
4 ml NP-40
5 ml ampholyte (pH 3-10)
2 ml 2-mercaptoethanol (Kodak)
H₂O to 100 ml total volume
Adjust to pH 3.0 with H₃PO₄

COMMENTARY

Background Information

Two-dimensional gel electrophoresis is the combination of two independent electrophoretic techniques, isoelectric focusing and SDS-polyacrylamide gel electrophoresis (O'Farrell, 1975). It is an extremely powerful method for examining complex mixtures of proteins (Celis and Bravo, 1984). It is possible to resolve more than 1500 proteins on a single two-dimensional gel. In addition to analytical applications, two-dimensional gels can provide a means of collecting small amounts of extremely pure proteins for amino acid sequence analysis or antibody production.

The ISO-DALT system (Anderson and Anderson, 1978a,b) was designed to facilitate running large numbers of highly reproducible, high-resolution, two-dimensional gels. Investigators who run small numbers of two-dimensional gels can still take advantage of some of the features of this system. The first-dimension apparatus quickly and inexpensively produces 20 (or 40) identical IEF gels. Even if only a small number of samples (e.g., 3 to 5) are to be run, one may apply duplicate loads or a range of concentrations of each sample. Second-dimension gels are run in multiples of 10 with the original ISO-DALT system (available from Hoefer Scientific and ISS Enprotech).

Proteins may be visualized by staining procedures or by autoradiography of radiolabeled proteins (Celis and Bravo, 1984). Computer analysis of two-dimensional protein gel patterns (e.g., the QUEST system of Garrels et al., 1984, and the Visage system, marketed by Protein Databases and Bio Image, respectively) may be useful for data analysis.

Critical Parameters

Several steps in this procedure may require optimization for a particular sample. Sample solubilization, for example, may require some experimentation to determine the best solubilization buffer and the optimum ratio of buffer to sample. For most tissue samples 1.5 ml of urea solubilization buffer per 100 mg of tissue will give good results. If samples must be stored, it is better to store (-70°C) the original sample than the solubilized sample.

The quality of the first-dimension separation depends on the ampholytes used, voltage, volt-hours, and reagent quality. For good resolution over a broad pH range, blends of ampholytes from different suppliers may give better results than broad range ampholytes from a

single supplier; however, pH 4.0 to 8.0 ampholytes from BDH have been found to perform well in most applications. If possible, test all available lots of ampholytes and buy a large supply of the best lot. Optimum voltage and volt-hours must be determined experimentally. Once suitable conditions have been established, all steps must be done in exactly the same way for each run in order to have the best possible reproducibility.

Troubleshooting

If all steps in this protocol are done correctly, the resulting two-dimensional gels should contain numerous round or elliptical protein spots. Mammalian tissue or whole cell samples typically yield two-dimensional patterns of >1500 spots distributed throughout the second-dimension gel with only a few spots overlapping. Problems arising from the first-dimension separation usually result in spots that are elongated horizontally (poorly focused) or a spot pattern in which most of the proteins are crowded into a narrowed zone along the x axis of the gel.

Vertically elongated spots and spot patterns that are compressed along the y axis of the gel indicate problems with the second-dimension separation. Spot patterns that appear wavy are nearly always the result of disturbances in the acrylamide gradient.

The following checklists will help the beginner locate common problems associated with two-dimensional electrophoresis.

General. (1) Were all of the reagents fresh? (2) Were all of the reagents properly prepared? (3) Were the samples properly solubilized?

First-dimension. (1) Were the samples ultracentrifuged just prior to loading? (2) Were the desired number of volt-hours applied? (3) Were the correct ampholytes used? (4) Was the salt content of the sample low?

Second-dimension. (1) Were the tops of the gels flat and smooth? (2) Was the acrylamide gradient smooth and uniform from top to bottom? (3) Was the tank buffer fresh and the pH proper?

Anticipated Results

A two-dimensional gel electrophoretic separation of proteins similar to that shown in Figure 10.3.2 will be obtained. Proteins of 200,000 molecular weight migrate 1 to 2 cm into the slab gel while those of 10,000 molecular weight migrate 14 to 15 cm. The resulting

two-dimensional pattern is a collection of protein spots in which proteins between 250,000 and 10,000 molecular weight and with isoelectric points between pH 3.8 and 8 are clearly separated from each other.

Time Considerations

First-dimension gels can be used as soon as 1 hr after casting or they can be kept for >24 hr before use. It is convenient to start the first-dimension gels in the afternoon, so they can run overnight and be ready to take off early the following day. Second-dimension gels can be poured the day before they are to be used. The next day the electrophoresis is completed at high voltage (220 to 280 V) for ~5 hr. The two-dimensional gels will be ready to put in fixative or staining solution at the end of the day and can remain in that solution overnight or even over the weekend. Usually 3 working days are required to complete the entire procedure, including staining.

Literature Cited

Anderson, N.G. and Anderson, N.L. 1978a. Analytical techniques for cell fractionations. XXI. Two-dimensional analysis of serum and tissue proteins: Multiple isoelectric focusing. *Anal. Biochem.* 85:332-340.

Anderson, N.L. and Anderson, N.G. 1978b. Analytical techniques for cell fractionations. XXII. Two-dimensional analysis of serum and tissue proteins: Multiple gradient-slab gel electrophoresis. *Anal. Biochem.* 85:341-354.

Celis, J.E. and Bravo, R. (eds.) 1984. Two-Dimensional Gel Electrophoresis of Proteins. Academic Press, Orlando, Fla.

Garrels, J.I., Farrar, J.T., and Burwell, C.B. 1984. The QUEST system for computer-analyzed two-dimensional electrophoresis of proteins. *In* Two-Dimensional Electrophoresis of Proteins (J.E. Celis and R. Bravo, eds.) Academic Press, Orlando, Fla., pp. 37-91.

O'Farrell, P.H. 1975. High-resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250:4007-4021.

Key Reference

Tollaksen, S.A., Anderson, N.L., and Anderson, N.G. 1984. Operation of the ISO-DALT system (ANL-BIM-84-1). Argonne National Laboratory, Argonne, Ill.

A detailed guide to the ISO-DALT method for two-dimensional gel electrophoresis.

Contributed by Lonnie D. Adams
The Upjohn Company
Kalamazoo, Michigan

Native protein standards

Electrophoresis buffer: appropriate 4× gel buffer diluted to 1× with H₂O

75-ml side-arm flask (used in gel preparation)

Additional reagents and equipment for gel electrophoresis (UNIT 10.2A) and staining proteins in gels (UNIT 10.6)

Prepare the gel

1. Assemble the glass-plate sandwich of the gel electrophoresis unit and secure it to the casting stand.

Either single-concentration or gradient gels can be used in the minigel or standard-size format. Gradient gels will enhance the band sharpness of the separated proteins.

2. Prepare acrylamide solutions according to the recipes in Table 10.2B.1, Table 10.2B.2, Table 10.2B.3, or Table 10.2B.4, adding the ammonium persulfate and TEMED just before use.

Deaeration of the solution before the polymerization catalysts are added will speed polymerization by removing inhibitory oxygen, but is not generally required. The pH used depends on many factors. The most important are the pI values of both the protein of interest and any contaminants, as well as protein mobility and protein solubility. Determining which pH and thus which buffer system to use is largely empirical. However, extremes of pH (<4.0 and >9.0) can lead to denaturation and should be avoided. Prior knowledge of the pI of a protein (UNITS 10.3 & 10.4) allows determination of the net charge under the

Table 10.2B.1 Recipes for Acetic Acid Nondenaturing Polyacrylamide Gels^a: pH range 3.7 to 5.6^b

Stock solution ^c	Final acrylamide concentration in gel (%) ^d						
	5	7.5	10	12.5	15	17.5	20
30% acrylamide/0.8% bisacrylamide	6.7	10	13.3	16.8	20	23.32	26.6
300 mM sodium sulfite ^e	0.4	0.4	0.4	0.4	0.4	0.4	0.4
4× acetic acid gel buffer	10	10	10	10	10	10	10
H ₂ O	22.58	19.28	15.98	12.48	9.28	5.96	2.68
10% (w/v) ammonium persulfate ^e	0.3	0.3	0.3	0.3	0.3	0.3	0.3
TEMED ^f	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Preparation of gel

In a 75-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.2A.1), 300 mM sodium sulfite, 4× acetic acid gel buffer (see Reagents and Solutions), and H₂O. If desired to speed polymerization, degas under vacuum ~5 min. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

^aThe recipes produce 40 ml gel solution, which is adequate for one gel of dimensions 1.5 mm × 14 cm × 16 cm or two gels of dimensions 0.75 mm × 14 cm × 16 cm.

^bThe pH range can be extended to ~2.0 (the pH of acetic acid) by using unadjusted acetic acid in place of 4× acetic acid gel buffer, although there is little buffering capacity at this pH.

^cAll reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

^dUnits of numbers in table body are milliliters. The desired percentage of acrylamide in the gel solution depends on the molecular size of the protein being separated.

^eMust be freshly made. Sodium sulfite is needed for efficient polymerization at acid pH.

^fAdded just before polymerization.

Table 10.2B.2 Recipes for Phosphate Nondenaturing Polyacrylamide Gels^a: pH range 5.8 to 8.0

Stock solution ^b	Final acrylamide concentration in gel (%) ^c						
	5	7.5	10	12.5	15	17.5	20
30% acrylamide/0.8% bisacrylamide	6.7	10	13.3	16.8	20	23.32	26.6
4× phosphate gel buffer	10	10	10	10	10	10	10
H ₂ O	23.08	19.78	16.48	12.98	9.78	6.46	3.18
10% (w/v) ammonium persulfate ^d	0.2	0.2	0.2	0.2	0.2	0.2	0.2
TEMED ^e	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Preparation of gel

In a 75-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.2A.1), 4× phosphate gel buffer (see Reagents and Solutions), and H₂O. If desired, degas under vacuum ~5 min to speed polymerization. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

^aThe recipes produce 40 ml gel solution, which is adequate for one gel of dimensions 1.5 mm × 14 cm × 16 cm or two gels of dimensions 0.75 mm × 14 cm × 16 cm.

^bAll reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

^cUnits of numbers in table body are milliliters. The desired percentage of acrylamide in the gel solution depends on the molecular size of the protein being separated.

^dMust be freshly made.

^eAdded just before polymerization.

separation conditions (i.e., if gel pH < protein pI, the protein will have a net positive charge; if gel pH > protein pI, the protein will be negatively charged).

3. Pour gel to 2 cm from the top of the gel mold and insert the comb. Avoid trapping air bubbles under the comb teeth.

Air bubbles will cause small semicircular depressions in the well and lead to distortions in the protein banding pattern.

4. Allow gel solution to polymerize 1 to 2 hr.

Polymerization is indicated by a sharp optical discontinuity around the wells.

Prepare samples and load the wells

5. Solubilize the protein sample to be analyzed using 5% (w/v) sucrose in water or dilute (1 to 5 mM) gel buffer if possible. Also prepare native protein standards.

The concentration of protein will vary depending on the sample complexity and detection method. For Coomassie blue staining of highly enriched samples such as the standards, use 1 to 2 mg/ml (1 to 2 μg/μl). For more complex mixtures, use 5 to 10 mg/ml (5 to 10 μg/μl). Load 10- to 100-fold less for silver staining. In general, samples should be loaded in a minimum volume, preferably 10 to 20 μl for 0.75- and 1.5-mm-thick gels, respectively. With thin gels, this means using a more concentrated protein sample.

6. Remove comb carefully and rinse wells with electrophoresis buffer (appropriate 4× gel buffer diluted to 1×).

Rinsing with electrophoresis buffer is needed to remove residual unpolymerized acrylamide monomer, which will continue to polymerize after comb removal, creating uneven wells that may interfere with sample loading.

Table 10.2B.3 Recipes for Tris Nondenaturing Polyacrylamide Gels^a: pH range 7.1 to 8.9

Stock solution ^b	Final acrylamide concentration in gel (%) ^c						
	5	7.5	10	12.5	15	17.5	20
30% acrylamide/0.8% bisacrylamide	6.7	10	13.3	16.8	20	23.32	26.6
4× Tris gel buffer	10	10	10	10	10	10	10
H ₂ O	23.08	19.78	16.48	12.98	9.78	6.46	3.18
10% (w/v) ammonium persulfate ^d	0.2	0.2	0.2	0.2	0.2	0.2	0.2
TEMED ^e	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Preparation of gel

In a 75-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.2A.1), 4× Tris gel buffer (see Reagents and Solutions), and H₂O. If desired, degas under vacuum ~5 min to speed polymerization. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

^aThe recipes produce 40 ml gel solution, which is adequate for one gel of dimensions 1.5 mm × 14 cm × 16 cm or two gels of dimensions 0.75 mm × 14 cm × 16 cm.

^bAll reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

^cUnits of numbers in table body are milliliters. The desired percentage of acrylamide in the gel solution depends on the molecular size of the protein being separated.

^dMust be freshly made.

^eAdded just before polymerization.

7. Fill wells with electrophoresis buffer. If desired, prerun gel.

The gel can be prerun at this point to remove any charged material such as ammonium persulfate from the gel prior to loading the sample. Assemble the electrophoresis unit and fill the buffer chambers with electrophoresis buffer. Run the gel at 300 V until the current no longer drops. This should take ~30 min. Disassemble the unit, discard the buffer, and proceed to the next step.

8. Carefully load up to 10 µl (0.75-mm gels) or 20 µl (1.5-mm gels) sample per lane as a thin layer at the bottom of the wells. Load control wells with native protein standards. Add an equal volume of electrophoresis buffer to any empty wells to prevent spreading of adjoining lanes.

Mobility (R_f) markers require special consideration in nondenaturing gel systems. For cationic systems, cytochrome c (pI ~9 to 10, 5 to 10 µg/lane) works well as an R_f marker. Bromphenol blue (10 µg/ml) is a suitable marker for anionic systems. The marker should be included in the solubilization buffer with the sample.

Perform the separation

9. Assemble the gel unit, fill the upper and lower buffer chambers with electrophoresis buffer, and connect the unit to the power supply. Set current to 30 mA for a 1.5-mm-thick gel (15 mA for a 0.75-mm-thick gel).

If the protein is negatively charged under the separation conditions, then the standard SDS-PAGE electrode polarity should be used (proteins will migrate to the anode or positive electrode; see UNIT 10.2A). If the protein is positively charged, then the electrodes should be reversed at the power supply (i.e., red high-voltage cable to the black output and black high-voltage lead to the red output) so the positively charged protein migrates to the negative cathode.

Table 10.2B.4 Recipes for Glycine Nondenaturing Polyacrylamide Gels^a:
pH range 8.6 to 10.6

Stock solution ^b	Final acrylamide concentration in gel (%) ^c						
	5	7.5	10	12.5	15	17.5	20
30% acrylamide/0.8% bisacrylamide	6.7	10	13.3	16.8	20	23.32	26.6
4× glycine gel buffer	10	10	10	10	10	10	10
H ₂ O	23.08	19.78	16.48	12.98	9.78	6.46	3.18
10% (w/v) ammonium persulfate ^d	0.2	0.2	0.2	0.2	0.2	0.2	0.2
TEMED ^e	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Preparation of gel

In a 75-ml side-arm flask, mix 30% acrylamide/0.8% bisacrylamide solution (see Table 10.2A.1), 4× glycine gel buffer (see Reagents and Solutions), and H₂O. If desired, degas under vacuum ~5 min to speed polymerization. Add 10% ammonium persulfate and TEMED. Swirl gently to mix. Use immediately.

^aThe recipes produce 40 ml gel solution, which is adequate for one gel of dimensions 1.5 mm × 14 cm × 16 cm or two gels of dimensions 0.75 mm × 14 cm × 16 cm.

^bAll reagents and solutions used in the protocol must be prepared with Milli-Q-purified water or equivalent.

^cUnits of numbers in table body are milliliters. The desired percentage of acrylamide in the gel solution depends on the molecular size of the protein being separated.

^dMust be freshly made.

^eAdded just before polymerization.

10. Continue electrophoresis until the R_f marker reaches the bottom of the gel.

For minigels, electrophoresis will take 1 to 2 hr. Standard gels require 4 to 6 hr runs.

11. Turn off power supply, disassemble the unit, and remove gel from sandwich.

12. Stain the gel according to UNIT 10.6.

NATIVE DISCONTINUOUS ELECTROPHORESIS AND GENERATION OF MOLECULAR WEIGHT STANDARD CURVES (FERGUSON PLOTS)

ALTERNATE PROTOCOL

One straightforward approach to discontinuous native electrophoresis is to leave out the SDS and reducing agent (DTT) from the standard Laemmli SDS-PAGE protocol (UNIT 10.2A). The gels are prepared as described in UNIT 10.2A except that the sample buffer contains no SDS or DTT (samples are not heated), and the gel and electrophoresis solutions are prepared without SDS. This protocol illustrates the separation of standard proteins at four different concentrations of acrylamide and how the results are used to construct a molecular weight standard curve (Ferguson plot) without the need for SDS. By plotting relative mobility against %T (percentage weight per volume of acrylamide plus bisacrylamide in the gel), the presence of isoforms and multimeric proteins can also be detected.

Materials

4× Tris·Cl, pH 8.8 (1.5 M Tris·Cl; APPENDIX 2)

4× Tris·Cl, pH 6.8 (0.5 M Tris·Cl; APPENDIX 2)

Protein sample of interest

2× Tris/glycerol sample buffer (see recipe)

Analysis of Proteins

10.2B.5

Native protein standards (e.g., Sigma nondenatured protein molecular weight kit)
Tris/glycine electrophoresis buffer (see recipe)

1. Assemble the glass-plate sandwich of the gel electrophoresis unit and place it in the casting stand.
2. Prepare and cast the gels, using 4× Tris-Cl, pH 8.8, for the separating gel and 4× Tris-Cl, pH 6.8, for the stacking gel instead of the SDS-containing counterparts (Table 10.2A.1). Prepare a minimum of four separate gels at different acrylamide concentrations.

A typical range of concentrations is from 5% to 12.5% (e.g., 5%, 7.5%, 10%, 12.5% acrylamide). As with SDS-PAGE, typical gel thickness ranges from 0.75 to 1.5 mm. The 0.75-mm-thick gels are recommended because they offer a combination of fast staining and high resolution.

3. Mix protein sample of interest 1:1 with 2× Tris/glycerol sample buffer to attain a 1 to 2 μg/μl final concentration. Also prepare native protein standards. Remove comb, rinse wells, and load 10 to 20 μl per well for Coomassie brilliant blue staining, and 1 to 2 μl for silver staining.

Some proteins must be dissolved in 50 mM NaCl or water to become fully solubilized prior to mixing with the sample buffer (Sigma, 1986).

4. Assemble gel electrophoresis unit, using Tris/glycine electrophoresis buffer to fill both lower and upper buffer chambers. Connect power supply and conduct electrophoresis.

Conditions for separation are the same as for discontinuous SDS-PAGE (i.e., 30 mA for 1.5-mm-thick gels, 15 mA for 0.75-mm-thick gels). For standard-size gels the separation takes 4 to 5 hr; for minigels, 1 to 2 hr is required. Alternatively, standard gels can be run at 4 to 6 mA/gel overnight.

5. After the bromphenol blue R_f marker has reached the bottom of the gel, fix and stain the proteins in the gels according to UNIT 10.6. Estimate relative mobilities of the proteins.

An example of a stained gel is shown in Figure 10.2B.1. A minimum of four gel concentrations is recommended. In Figure 10.2B.1, Sigma native molecular weight standards were separated on 5%, 7.5%, 10%, and 12.5% acrylamide gels (5.1%, 7.7%, 10.3%, and 12.8% T, respectively).

6. Plot $\log R_f$ against gel concentration (% T) (Fig. 10.2B.2). Determine the slope of K_r using linear regression.
7. Plot $-\log K_r$ of the curves from step 6 against \log molecular weight of the standards (Fig. 10.2B.3). Determine the slope using linear regression.
8. Estimate the size of the standards and unknowns from the generated curve (Ferguson plot).

Use the curve generated by linear regression to estimate the predicted size of the standard for comparison to the actual size stated by the supplier. This indicates the accuracy of the curve. The $-\log K < D_r$ value (y) of the unknown is then used to predict the molecular weight (x).

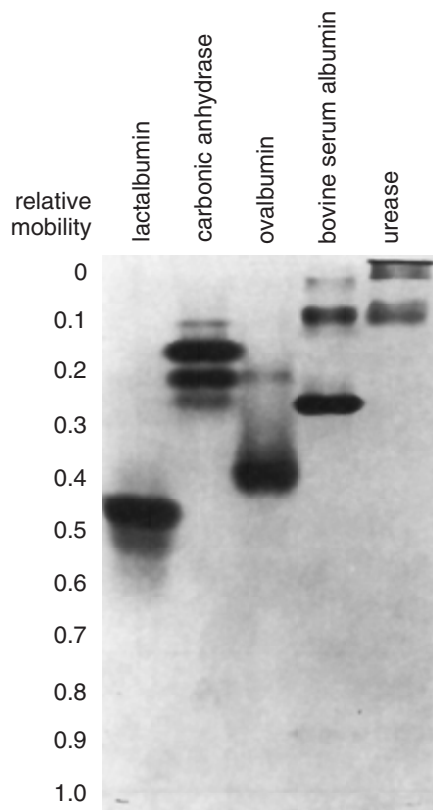


Figure 10.2B.1 Separation of native protein standards under nondenaturing conditions by discontinuous polyacrylamide gel electrophoresis at 12.8% T. Approximately 20 μg protein was loaded per lane on a 1.5-mm-thick, 16-cm-long gel. The gel and samples were prepared according to the Alternate Protocol and were electrophoresed 16 hr at 6 mA. Proteins were stained with Coomassie blue.

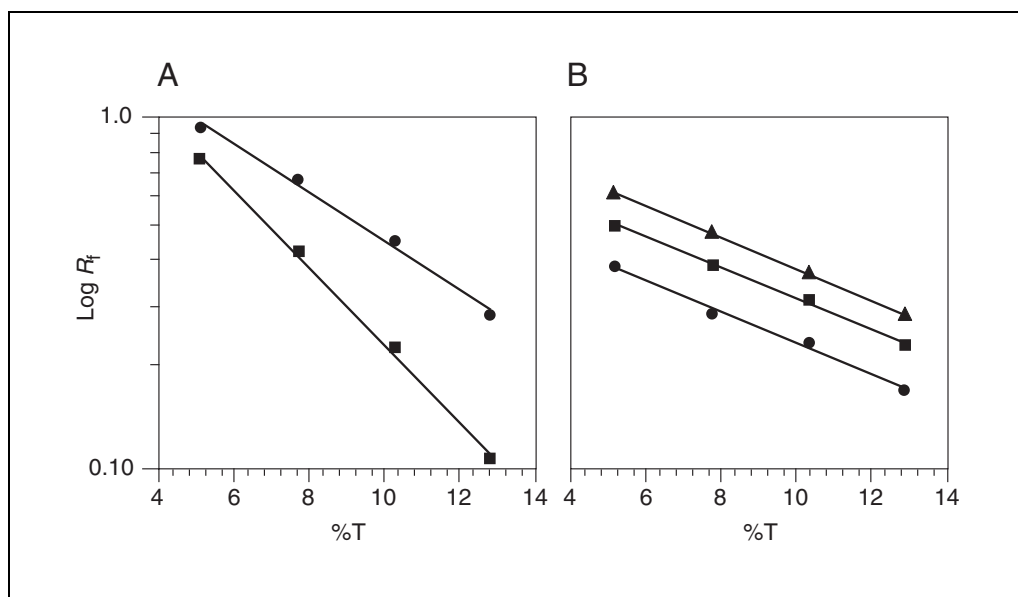


Figure 10.2B.2 Effect of %T on the relative mobility of several native proteins. The relative mobility (R_f) of the standard proteins shown in Figure 10.2B.1 was determined at four different gel concentrations and plotted as $\log R_f$ against %T. See text for details. **(A)** BSA monomer (squares) and dimer (circles); **(B)** carbonic anhydrase isoforms.

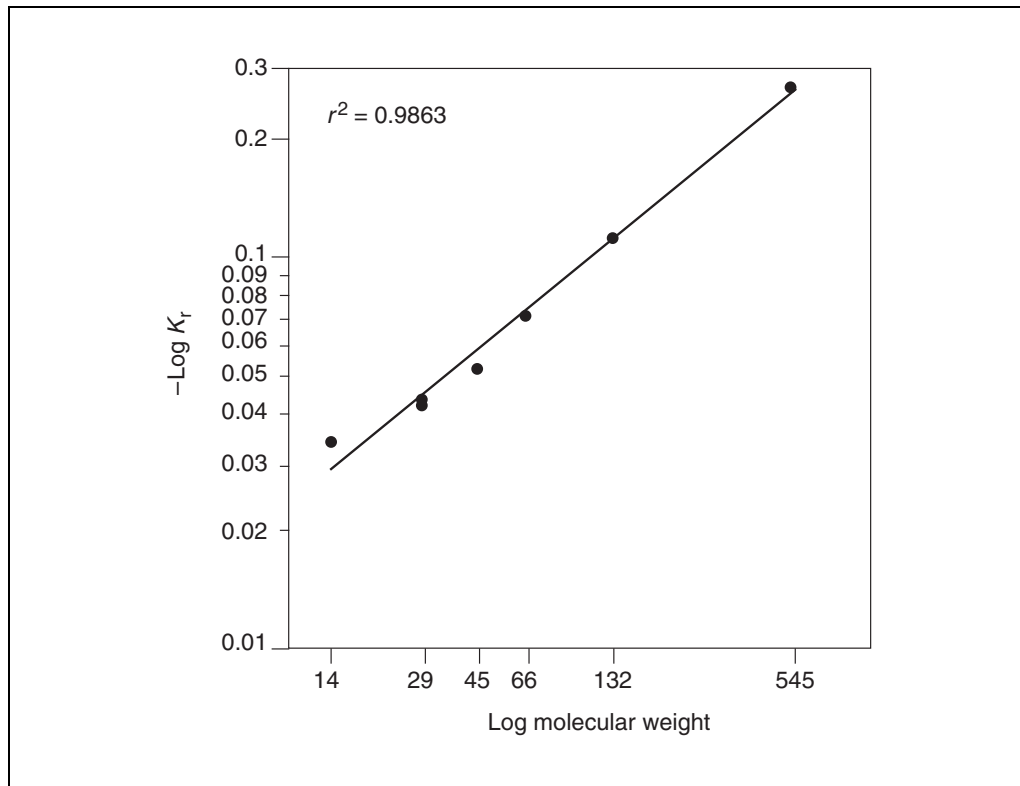


Figure 10.2B.3 Native molecular weight standard curve. The $-\log$ slope of the line (K_r) from Figure 10.2B.2 is plotted against log molecular weight of the standards.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Acetic acid gel buffer, 4× (200 mM acetic acid, pH 3.7 to 5.6)

- 11.49 ml glacial acetic acid
- Add to 500 ml H₂O
- Adjust to pH 3.7 to 5.6 with 1 M NaOH
- Add H₂O to 1000 ml
- Store up to 1 month at 4°C

Glycine gel buffer, 4× (200 mM glycine, pH 8.6 to 10.6)

- 15.01 g glycine
- Add to 500 ml H₂O
- Adjust to pH 8.6 to 10.6 with 1 M NaOH
- Add H₂O to 1000 ml
- Store up to 1 month at 4°C

Phosphate gel buffer, 4× (400 mM sodium phosphate, pH 5.8 to 8.0)

- 55.2 g NaH₂PO₄·H₂O
- Add to 500 ml H₂O
- Adjust to pH 5.8 to 8.0 with 1 M NaOH
- Add H₂O to 1000 ml
- Store up to 1 month at 4°C

Tris gel buffer, 4× (200 mM Tris·Cl, pH 7.1 to 8.9)

24.23 g Tris base
Add to 500 ml H₂O
Adjust to pH 7.1 to 8.9 with 1 M HCl
Add H₂O to 1000 ml
Store up to 1 month at 4°C

2× Tris/glycerol sample buffer

25 ml 0.5 M Tris·Cl, pH 6.8
20 ml glycerol
1 mg bromphenol blue
Add H₂O to 100 ml and mix
Store in 1-ml aliquots up to 6 months at -70°C

Tris/glycine electrophoresis buffer

15.1 g Tris base
72.0 g glycine
H₂O to 5000 ml
Store up to 1 month at 4°C

COMMENTARY

Background Information

Under nondenaturing conditions, in which protein activity, native charge, and conformation are sustained, electrophoretic separation depends on many factors, including size, shape, and charge. Characteristics such as intrinsic molecular weight (i.e., in the absence of denaturation), the number of isoforms, and the presence of multimeric proteins can be determined with nondenaturing electrophoresis (often called native electrophoresis).

The most important application of nondenaturing electrophoresis is in the determination of native protein size (Alternate Protocol). Ferguson plots (reviewed by Andrews, 1986) were first described for starch gels (Ferguson, 1964) and then for polyacrylamide gels (Hedrick and Smith, 1968). Ferguson plots are prepared by separating proteins under nondenaturing conditions at several different gel concentrations. As the acrylamide concentration (%T) is increased, the relative mobility (R_f) of the protein decreases. This is plotted as log relative mobility (on the y axis) versus %T (on the x axis) to produce a straight line. The slope of this line is referred to as the retardation coefficient (K_r) and measures how effectively a protein is slowed by the increase in %T. Large proteins will be retarded much more significantly than small proteins with increasing gel concentration, with the size of the protein being proportional to the slope of the curve. Once the K_r plots for several size standards are generated (Fig. 10.2B.2), the K_r values are plotted against

the molecular weight of the standard proteins using a log-log graph (Fig. 10.2B.3). The retardation coefficient also depends on a large number of other variables including temperature, pH, buffer type, ionic strength, and %C (percent bisacrylamide cross-linker). All these factors should be kept constant for a given experiment.

In addition to estimated size, other types of information are available from the Ferguson plots (Rodbard and Chrambach, 1971; Andrews, 1986). For example, if two components differ in size but have the same charge per unit size (e.g., for a multimeric protein with identical subunits), curves similar to those illustrated by BSA monomer and dimer (Fig. 10.2B.2A) will result. Note that when the curve is extrapolated back to 0% T, it is evident that the monomer and the dimer have similar free solution mobilities. Furthermore, as the acrylamide concentration is increased, the separation between the two also increases. However, if two proteins have similar sizes but different amounts of charge, the curves will be parallel on the log plot. This is illustrated by the carbonic anhydrase isoforms (Fig. 10.2B.2B). In this example, optimal separation of the isoforms occurs at the lower concentrations of acrylamide as this is a log plot.

Further applications of nondenaturing electrophoresis include preparative purification. The pH of the gel determines the net charge on the protein. Below its isoelectric point (pI) a protein will have a net positive charge, whereas above its pI it will have a net negative charge.

In general, most proteins will be positively charged at pH 2.0 to 4.0; above pH 8.0, most proteins will be negatively charged. As these general guidelines imply, the majority of proteins have isoelectric points between pH 4.0 and 8.0. There are, however, many exceptions. A protein with a highly acidic isoelectric point (e.g., pepsin, with a pI of 2.2) will remain negatively charged at a pH down to its pI. Although a full range of pH options are given, extremes of pH (<4.0 and >9.0) should be avoided, if possible, to minimize denaturation or inactivation. By picking an appropriate electrophoresis pH, it is possible to ensure that the protein of interest will be either positively or negatively charged so that it can be selectively run into the gel, excluding a large proportion of contaminants that have the opposite or no charge. Furthermore, the pH conditions determine the mobility and can be adjusted to ensure a difference in mobility between the protein of interest and contaminants.

Continuous gel systems (see Basic Protocol) offer the most flexibility in terms of separation design. The pH can be tailored so that a given protein has a net positive, neutral, or negative charge. Depending on the polarity of the gel, the protein can then be excluded from or electrophoresed into the gel. Discontinuous gels have a fixed pH and gel polarity. For the nondenaturing Laemmli gel presented in the Alternate Protocol, the proteins of interest should have an isoelectric point of ≤ 7.0 in order to be negatively charged so that they move into the gel. Other more basic and more acidic discontinuous gel systems can be found in Hames (1990) and Schägger (1994).

Critical Parameters

The success of a gel separation under nondenaturing conditions depends on many factors, and two of the most important are protein solubility and isoelectric point. The protein must be soluble at the pH and the ionic strength of the gel, and it must be charged at that pH in order to move into the gel. If the protein experiences a pH below its isoelectric point, then it will have a net positive charge and will move to the negative electrode. Note that this is the reverse of typical SDS-PAGE. If the protein experiences a pH above its isoelectric point, it will have a net negative charge and will migrate to the positive electrode.

Solubility is a complex issue. Membrane-associated and other hydrophobic proteins are difficult to separate by nondenaturing electrophoresis (Schägger, 1994). Nonionic deter-

gents at concentrations up to 1% and solubilizing reagents such as urea (4 to 8 M) can be used, but these reagents, especially urea, are likely to alter the protein's conformation and most likely the isoelectric point by exposing previously hidden charged groups. If detergent or urea must be included for solubilization, the minimum required to solubilize the protein should be used. Schägger (1994) lists several nonionic detergents suitable for solubilization. Among the more popular are octylglucoside and CHAPS. In general, detergents should be used near the critical micelle concentration (CMC; 0.001% to 1%, depending on the detergent).

The gel concentration has a dramatic effect on resolution and should be optimized in order to achieve the best separation and band sharpness. In general, increasing the %T will improve band sharpness.

Troubleshooting

Gel polymerization at acid pH can be problematic, and sodium sulfite is needed for efficient polymerization (Andrews, 1986). Both the ammonium persulfate and the sodium sulfite must be freshly made, and the highest quality reagents available should be used. Furthermore, the gel solutions should be at room temperature for effective polymerization.

If the protein does not enter the gel and no stained material is present at the well surface, try reversing the polarity of the electrode. If material concentrates at the top of the gel, try lowering the acrylamide concentration. Stained material at the top of the gel may also indicate poor solubilization, and increasing the ionic strength of the solubilization buffer or adding a small amount of urea and/or nonionic detergent may be required.

Anticipated Results

Proteins will resolve depending on their solubility and native charge at the chosen pH. Ideally, a distinct band representing the protein of interest will be visible. If the band is diffuse, then increasing the gel concentration or using a gradient gel will improve resolution. If the band is not visible, then the protein may be at its isoelectric point or may have moved out of the gel because it had the wrong charge. Continuous gel systems, although more versatile, will give lower resolution than discontinuous gels. Detergents or other solubilizing agents such as urea may be needed to fully solubilize and resolve the protein. Once the conditions that resolve the protein are determined, Ferguson plots will give indications of multiple

subunit structure, native size, and potential isoform relationships.

Time Considerations

Separations will be complete when the tracking dye or protein reaches the bottom of the gel. For minigels, this generally takes 1 to 2 hr, using 15 or 30 mA for 0.75- or 1.5-mm-thick gels, respectively. Standard-format gels require 4 to 5 hr at 15 or 30 mA for 0.75- or 1.5-mm-thick gels, respectively. Standard-format gels can also run overnight at 4 to 6 or 8 to 12 mA for 0.75- or 1.5-mm-thick gels, respectively.

Literature Cited

- Andrews, A.T. 1986. Electrophoresis: Theory, Techniques and Biochemical and Clinical Applications, 2nd ed. Oxford University Press, New York.
- Ferguson, K.A. 1964. Starch-gel electrophoresis—application to the classification of pituitary proteins and polypeptides. *Metabolism* 13:985-1002.
- Hames, D. 1990. One-dimensional polyacrylamide gel electrophoresis. *In* Gel Electrophoresis of Proteins: A Practical Approach, 2nd ed. (B.D. Hames and D. Rickwood, eds.) pp. 1-147. Oxford University Press, New York.

Hedrick, J.L. and Smith, A.J. 1968. Size and charge isomer separation and estimation of molecular weights of proteins by disc gel electrophoresis. *Arch. Biochem. Biophys.* 126:155-164.

Rodbard, D. and Chrambach, A. 1971. Estimation of molecular radius, free mobility, and valence using polyacrylamide gel electrophoresis. *Anal. Biochem.* 40:95-134.

Schägger, H. 1994. Native gel electrophoresis. *In* A Practical Guide to Membrane Protein Purification (G. Von Jagow and H. Schägger, eds.) pp. 81-104. Academic Press, San Diego.

Sigma. 1986. Nondenatured protein molecular weight marker kit (Technical Bulletin No. MKR-137). Sigma Chemical Company, St. Louis, Mo.

Key Reference

Andrews, 1986. See above.

Covers a variety of electrophoretic techniques, including nondenaturing electrophoresis and Ferguson plots.

Contributed by Sean R. Gallagher
Motorola Corporation
Tempe, Arizona

Two-Dimensional Gel Electrophoresis

Two-dimensional gel electrophoresis is the combination of two high-resolution electrophoretic procedures (isoelectric focusing and SDS–polyacrylamide gel electrophoresis) to provide much greater resolution than either procedure alone. In the first-dimension gel, solubilized proteins are separated according to their isoelectric point (pI) by isoelectric focusing. This gel is then applied to the top of an SDS-slab gel and electrophoresed. The proteins in the first-dimension gel migrate into the second-dimension gel where they are separated on the basis of their molecular weight.

The resulting two-dimensional gel contains numerous round or elliptical protein spots well separated from each other. A two-dimensional gel of whole cell or tissue lysate may easily contain 1500 protein spots if a sensitive protein detection method, such as silver staining (UNIT 10.6) or autoradiography (APPENDIX 3A), is used. One-dimensional separation of a similar sample would result in only about 100 protein bands, with many of the bands containing more than one protein. Therefore, two-dimensional electrophoresis is the method of choice when maximum resolution of proteins in a complex mixture is required. Even when dealing with purified protein samples such as immunoprecipitates (UNIT 10.16), two-dimensional electrophoresis offers the advantage of being able to simultaneously monitor changes in charge and/or molecular weight of proteins of interest. This can be especially helpful when working with recombinant proteins (which may vary slightly from those naturally expressed) and with processed proteins. Immunoprecipitated proteins can be solubilized in either the SDS or the urea solubilization buffer described in the reagents and solutions section of this unit using the procedures described in UNIT 10.16. Remember that urea solubilization buffer is never put in boiling water.

Basic Protocols 1 and 2 are based on the type of equipment originally described by O'Farrell (1975). Such equipment is relatively inexpensive and can produce satisfactory results in situations that do not require running large numbers of two-dimensional gels. Many variations of the original procedure have been described in the literature and by equipment suppliers. The procedures below are appropriate, with minor modifications, for different equipment available from commercial suppliers or custom-made equipment similar to that described by O'Farrell (1975); detailed instructions are usually supplied by the manufacturer. Alternate Protocol 1 describes modifications for very basic proteins and Alternate Protocol 2 describes modifications for very acidic proteins. Alternate Protocol 3 describes how two-dimensional electrophoresis can be performed using a minigel system. Protein sample preparation is presented in Support Protocol 1.

IEF systems are gaining in popularity, and a procedure for using immobilized pH gradients and tube gels with thin rectangular IEF gels supported by a plastic backing is included (see Alternate Protocol 4 and Support Protocol 2).

FIRST-DIMENSION (ISOELECTRIC-FOCUSING) GELS

This protocol is for broad-range first-dimension gels which resolve proteins with isoelectric points between about pH 4 to 8. Very basic proteins can be focused using Alternate Protocol 1 and very acidic proteins can be focused using Alternate Protocol 2. If increased resolution in a narrow pH range is needed, narrow-range ampholytes can be added to the broad-range ampholytes in a 2:1 ratio. For example, 0.50 ml of pH 5 to 7 ampholytes, plus 0.25 ml of pH 4 to 8 ampholytes, can be used instead of 0.75 ml of pH 4 to 8 ampholytes alone to increase the separation distance of proteins in the pH 5 to 7 range.

The first-dimension gels are poured in gel tubes, which are placed in a tube electrophoresis cell. After filling the reservoirs, the samples are loaded and the gels are run. The gels are

BASIC PROTOCOL 1

Analysis of Proteins

10.4.1

Contributed by Lonnie D. Adams and Sean R. Gallagher

Current Protocols in Molecular Biology (2004) 10.4.1-10.4.23

Copyright © 2004 by John Wiley & Sons, Inc.

extruded using water pressure and can then be stored at -70°C or used immediately for the second-dimension gel (Fig. 10.4.1).

Materials

Urea (ultrapure)
30% acrylamide/1.8% bisacrylamide (see recipe)
Ampholytes, pH 4 to 8 (Bio-Rad, Serva, Invitrogen, Sigma-Aldrich)
Nonidet P-40 (NP-40)
TEMED (*N,N,N',N'* tetramethylethylenediamine)
10% ammonium persulfate (see recipe)
0.085% phosphoric acid (see recipe)
0.02 M NaOH (see recipe)
Protein samples (see Support Protocol 1)
Concentrated bromphenol blue (see recipe)
Chromic acid cleaning solution (see recipe)

1.0- to 3.0-mm-inner-diameter glass gel tubes (~1.5 in. longer than the width of the second-dimension gel; 4- to 6-mm outer diameter)
2.5- to 3.0-cm-inner-diameter gel-casting glass tube, ~2 cm shorter than gel tubes
Small vacuum flask
50- μl , 1-ml, and 20-ml syringes
0.2- or 0.45- μm filter capsule (Acrodisk; Gelman)
Single-edge razor blade
Rubber grommets
Tube cell (Bio-Rad, Topac, CBS Scientific, Scie-Plas)
22-G hypodermic needle (2-in. long)
200- μl pipettor tip
1-dram gel vials

NOTE: Distilled, deionized water should be used throughout this protocol.

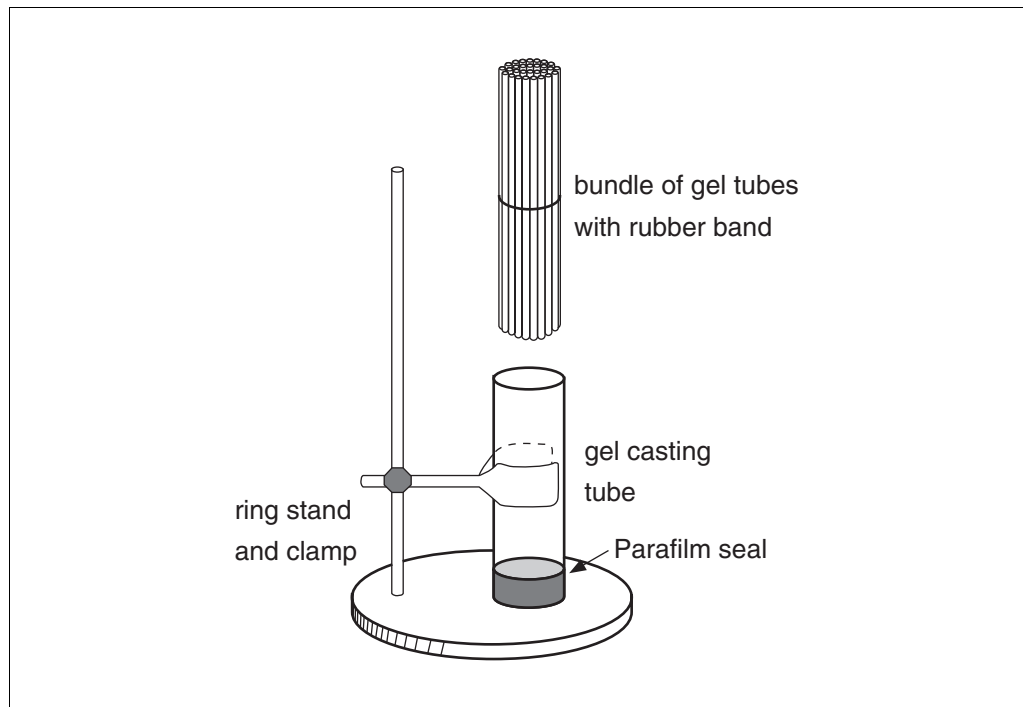


Figure 10.4.1 Assembly of gel tubes in casting tube.

Prepare and pour the gels

1. Mark clean, dry 1.0- to 3.0-mm-i.d. gel tubes to indicate the desired height of the gel (usually the same as the width of the second-dimension gel).

Cleaning instructions are provided in step 26.

The most commonly used type of isoelectric focusing apparatus is the cylindrical unit (Bio-Rad or Hoefer). Gel tubes are available with a wide range of inside diameters; however, for the highest resolution, tubes of 1.5-mm i.d. are preferred. Tubes up to 3-mm i.d. may be used if it is necessary to load large sample volumes or larger amounts of protein.

2. Place a rubber band around the gel tubes so that they form a tight bundle (~12 tubes fit into a bundle; Fig. 10.4.1).
3. Hold the bundle vertically on a flat surface and push down on the tops of the tubes so that the bottoms are even.
4. Carefully seal one end of the 2.5- to 3.0-cm-i.d. gel-casting glass tube with three or four layers of Parafilm to form a strong, water-tight seal.
5. Place the bundle of gel tubes inside the gel casting tube and support the glass tube in a vertical position with a ring stand and clamp to allow the sealed end of the glass tube to rest on a solid surface.
6. Add 8.25 g urea, 6.0 ml water, 2.0 ml of 30% acrylamide/1.8% bisacrylamide, and 0.75 ml ampholytes, pH 4 to 8, to a small vacuum flask. Add a small stir bar to the flask. Place the flask in a warm water bath on a magnetic stirrer and stir just until the urea is in solution; do not heat the solution to $>30^{\circ}\text{C}$.

This makes 15 ml acrylamide solution, sufficient for twenty 1.5×160 -mm gels.

Although the BDH pH 4 to 8 Resolytes are highly recommended, one may wish to use a blend of ampholytes as a substitute. For a broad-range blend, mix 1 part ampholytes, pH 2 to 11 (Serva) and 2 parts ampholytes, pH 3.5 to 10 (Amersham Biosciences). For an alternative broad-range blend, use equal volumes of ampholytes, pH 3 to 10 (Bio-Rad) and pH 2 to 11 (Serva). Mix enough ampholytes at one time to last ≥ 1 year. This will improve run-to-run reproducibility. This alternative can be used if Resolytes are unavailable or for a custom pH gradient.

7. Deaerate the solution by applying a strong vacuum for 2 to 3 min.
 8. Add 0.3 ml NP-40 and swirl until dissolved.
- Do not add the NP-40 before deaeration because it will foam.*
9. Pour solution into a 20-ml syringe fitted with a 0.2- or 0.45- μm filter capsule and force through the filter.
 10. Add 10 ml TEMED and swirl. Add 70 ml of 10% ammonium persulfate and swirl. Immediately pipet the gel solution into the space between the gel tubes and the large glass tube.

Work quickly, as the gel solution will start to polymerize in ~3 min.

11. Gently run water down the outside of the gel tubes using a wash bottle. Add water until the level of the acrylamide solution inside the tubes reaches the desired height.

As water is layered on top of the gel solution it will force the solution up the tubes from the bottom. There is no need to overlay the gel solution inside the tubes. Allow ≥ 1 hr for polymerization.

Set up the gels in the tube cell

12. Remove the Parafilm from the bottom of the gel casting tube and push the gel tubes, containing the polymerized gel, out the bottom. Cut across the gel-tube bottoms to remove excess acrylamide with a single-edge razor blade. Rinse the bottom of the gel tubes under running deionized water to remove residual acrylamide.
13. Place rubber grommet on the top of each tube, making sure that the top surface of the gel is visible below the grommet; ~5 mm of the gel tube should be visible above the grommet.
14. Seat the tube and grommet assemblies in the holes of the upper buffer reservoir of the tube cell. Plug any unused holes with rubber stoppers.
15. Fill the lower reservoir with ~3 liters of 0.085% phosphoric acid.
16. Place the upper reservoir into lower reservoir and adjust lower buffer level if needed.

The buffer should cover the entire gel for good heat dissipation.
17. Fill the upper buffer reservoir with 250 ml of 0.02 M NaOH.
18. Fill the gel tubes to the top with 0.02 M NaOH using a 1-ml syringe equipped with a 22-G hypodermic needle. Be careful to eliminate any air bubbles in the gel tubes.
19. Connect the tube cell to the power supply. The black (–) lead goes to the upper reservoir. Prefocus the gel 1 hr at 200 V constant voltage (see *UNIT 10.2A* introduction for a discussion of electricity and electrophoresis). Disconnect the tube cell from the power supply.

Load the samples and run the gels

20. Layer protein samples on top of the gels through the upper buffer with a 50- μ l syringe.

The maximum amount of total protein that can be loaded onto a first-dimension gel varies depending upon the nature of the sample. Samples such as whole-cell lysates which contain a large number of proteins of widely varied isoelectric point (pI) can contain much more total protein than a sample of a single highly purified protein. As much as 100 to 150 μ g of a protein mixture can be loaded on a first-dimension gel 1.5-mm in diameter, while a tenth that amount of highly purified protein would probably be an overload. For gels 1.5-mm in diameter, 10 to 20 μ l of sample is preferred; however, up to 30 μ l can be applied.

Alternatively, place the samples directly on the gel, then overlay with 10 to 20 μ l half-strength solubilization buffer (diluted 1:2 with water). This will protect samples from exposure to the basic upper buffer. Fill the remainder of the tube with 0.02 M NaOH to eliminate any bubbles.

21. Place the lid on the upper reservoir and attach the electrical leads to a power supply.
22. Turn on the power supply and adjust to the desired settings at constant voltage.

For 1.5-mm-i.d., 16-cm-long gels, 700 to 800 V (constant voltage) for 16 hr (11,000 to 13,000 V hr) works well for most samples.

23. Reduce the voltage setting to zero and turn off the power supply to end the run. Add ~1 μ l concentrated bromphenol blue to the top of each gel with a 50- μ l syringe.

The bromphenol blue will quickly diffuse into the gel. At the end of the run, the gels may be noticeably shorter and the bromphenol blue may stop short of the top of the gel. In this case, the power can be applied for a few minutes and the dye will migrate into the top of the gels. The blue dye will mark the basic end of the first-dimension gel and serve as the tracking dye in the second-dimension separation.

Extrude the gels

24. Extrude the gels from the tubes using water pressure from a 1-ml syringe fitted with a 200- μ l pipettor tip (cut off ~1 cm of the large end of the tip so it fits on the syringe).

Gels with a diameter >2.0 mm may have to be loosened by inserting a blunt 26-G, 9-cm-long needle between the gel and the wall of the tube while injecting water. Rotate the tube so the needle will pass over the entire circumference of the gel while continuously injecting water. Repeat this operation at the opposite end of the tube and the gel should slide out easily.

25. Place each gel in a labeled gel vial. The gels can be stored at -70°C for many weeks or used immediately.
26. Soak gel tubes overnight in chromic acid cleaning solution, then rinse thoroughly under running deionized tap water for 15 min. Remove excess water from the gel tubes with suction and allow them to dry.

SOLUBILIZATION AND PREPARATION OF PROTEINS IN TISSUE SAMPLES

Tissue samples are solubilized in either SDS or urea solubilization buffer (see commentary) by Dounce homogenization. After centrifugation, the protein sample can be loaded onto the first-dimension gel.

Materials

Tissue samples
SDS or urea solubilization buffer (see recipes)
Dounce homogenizer with pestles A and B
200- μ l centrifuge tubes
Beckman 42.2-Ti rotor (or equivalent)

1. Weigh and place tissue samples in a Dounce homogenizer.
2. Add 1.5 to 2.0 ml SDS or urea solubilization buffer per 100 mg tissue.
3. Homogenize using 50 strokes with pestle B, then 50 strokes with pestle A.

Cultured cells and body fluids do not normally require homogenization.

4. Let stand a few minutes, then transfer an aliquot to a 200- μ l centrifuge tube. Centrifuge ≥ 2 hr at $100,000 \times g$, or 1 hr at $>200,000 \times g$, 20°C . Save the supernatant (protein sample) to load onto the first-dimension gel as described in Basic Protocol 1.

If using SDS solubilization buffer, place the sample in boiling water 5 min before centrifugation. Never heat samples in urea solubilization buffer. Heating proteins in a urea solution will cause carbamylation of proteins, and thus result in charge variants. Centrifugation should be done just prior to loading the sample onto the first-dimension gel. Failure to centrifuge tissue samples may cause plugging of the first-dimension gels, resulting in poor entry of proteins. The Beckman 42.2-Ti rotor, which accommodates seventy-four 200- μ l tubes, is recommended.

SECOND-DIMENSION GELS

A wide variety of equipment suitable for casting and running second-dimension gels is available. Nearly any SDS-PAGE slab-gel system will work. For the best possible resolution over a wide molecular-weight range, a 10% to 20% acrylamide gradient gel is recommended (UNIT 10.2A); however, for many applications, a conventional slab gel as described below is sufficient. The procedure described uses the Bio-Rad PROTEAN II system, but it can be easily adapted to similar equipment, such as the vertical slab-gel units from Hoefer.

SUPPORT PROTOCOL 1

BASIC PROTOCOL 2

Analysis of Proteins

10.4.5

The second-dimension gel is poured between glass plates. The first-dimension gel is then loaded onto the second-dimension gel and sealed in place with agarose. The gels are mounted in the cell and run. After disassembling the electrophoresis unit, the gel is stained and analyzed.

Materials

30% acrylamide/0.8% bisacrylamide (see recipe for acrylamide/bisacrylamide solutions)
Gel buffer (see recipe)
10% (w/v) SDS
TEMED
10% ammonium persulfate (see recipe)
Isobutyl alcohol, H₂O-saturated
Stacking gel buffer (optional; see recipe)
First-dimension gel (see Basic Protocol 1)
Equilibration buffer (see recipe)
Hot 0.5% and 1% agarose (see recipe; keep in boiling water bath)
Protein molecular weight standards (Table 10.2A.2; available as kits Bio-Rad or Amersham Biosciences)
SDS solubilization buffer (see recipe)
Reservoir buffer (see recipe), prechilled to 10° to 20°C
Coolant (from running tap water or circulating refrigerated water bath)

Gel plates, one long and one short
1.5-mm spacers (~14 cm × 14 cm × 0.75 mm)
Casting stand
Gel identification tag (e.g., typed consecutive numbers on filter paper)
Nylon screen
5 × 15-cm glass plate
PROTEAN II electrophoresis cell (Bio-Rad)

Prepare and pour the gel

1. Assemble the gel plates by placing 1.5-mm spacers vertically between a long and short gel plate. The side where the long plate protrudes is the top.

Gel thickness is usually equal to the diameter of the isoelectric focusing gels, but this is not absolutely necessary.
2. Position clamps on each side of the gel sandwich over the spacers and place on the casting stand. Be sure the plates and spacers are properly aligned, then tighten the clamps and cams to get a leak-proof seal. Make adjustments so that plates are level and vertical.
3. Place the gel identification tag between the glass plates so that it rests in the lower right hand corner.
4. Prepare the gel solution by combining 30% acrylamide/0.8% bisacrylamide, gel buffer, and water in a vacuum flask (Table 10.4.1).
5. Deaerate the solution by applying vacuum for 5 min.
6. Add 10% SDS and TEMED and swirl, then add 10% ammonium persulfate and swirl (Table 10.4.1).
7. Fill the gel sandwich to 5 mm below the top of the short plate and overlay with H₂O-saturated isobutyl alcohol or water.

Table 10.4.1 Solutions for Second-Dimensional Gels^a

Stock solutions	Final acrylamide concentration (%)			
	7.5	10	12.5	15
30% acrylamide/ 0.8% bisacrylamide	12.5	16.7	20.8	25.1
Gel buffer	12.5	12.5	12.5	12.5
H ₂ O	24.6	20.3	16.2	12.0
10% SDS	0.5	0.5	0.5	0.5
TEMED	0.026	0.026	0.026	0.026
10% ammonium persulfate	0.185	0.185	0.185	0.185

^aValues in table body are milliliters of stock solution. Preparation of solutions is described in steps 4 to 6 of protocol.

A stacking gel is not usually required when running second-dimension gels. If a stacker is desired, stop the gel 1.5 cm from the top of the short plate, overlay with H₂O-saturated isobutyl alcohol, and allow to polymerize 1 hr. Remove the overlay and fill the remaining space with stacking gel solution made by combining 4.8 ml stacking gel buffer, 3.8 ml water, and 1.4 ml 30% acrylamide/0.8% bisacrylamide (all deaerated) with 200 μ l of 10% ammonium persulfate and 8 μ l TEMED. Pipet onto the separating gel, overlay with H₂O-saturated isobutyl alcohol, and allow to polymerize 30 min.

8. Allow the second-dimension gel to polymerize \geq 1.5 hr.

Load first-dimension gel onto second-dimension gel

9. If first-dimension gel is frozen, thaw at room temperature. Add equilibration buffer to completely cover the gel.

The time the gel is in equilibration buffer containing SDS can vary from seconds to several minutes depending upon the sample. Most proteins run well without the introduction of SDS from equilibration buffer; however, a few proteins do not. Leaving gels in equilibration buffer for >10 to 15 min can result in loss of proteins.

10. Pour the gel and equilibration buffer onto a nylon screen placed over a beaker.
11. Place the first-dimension gel on a 5 \times 15-cm glass plate (Parafilm is not rigid enough). Using a spatula, lay the gel out straight along one edge of the glass plate.
12. Pipet a very thin layer of hot 0.5% agarose on the top of the slab gel to be loaded.

This can easily be done by putting ~0.1 ml hot agarose on the upper left corner of a gel and quickly tilting the gel to the right so that the agarose will flow across the gel surface.

13. Using a spatula, slide the first-dimension gel off the glass plate and place it across the top of the slab gel. Orient the first dimension gel with the blue (basic) end to your right (Fig. 10.4.2).

Take care that no air bubbles get trapped between the first- and second-dimension gels and that the first-dimension gel does not get stretched or compressed.

14. Pipet a thin layer of hot 0.5% agarose over the first-dimension gel to seal it in place. Allow the agarose to solidify.

Protein molecular-weight markers may be run as one-dimensional separations on the sides of the second-dimension gel. Solubilize marker proteins in SDS solubilization buffer by boiling for 5 min, then dilute 1:1 with hot 1% agarose solution and draw up the hot solution in a glass tube the same diameter as the first-dimension gel. A short piece of the solidified agarose can be applied to one or both sides of the second-dimension gel and held in place with 0.5% agarose.

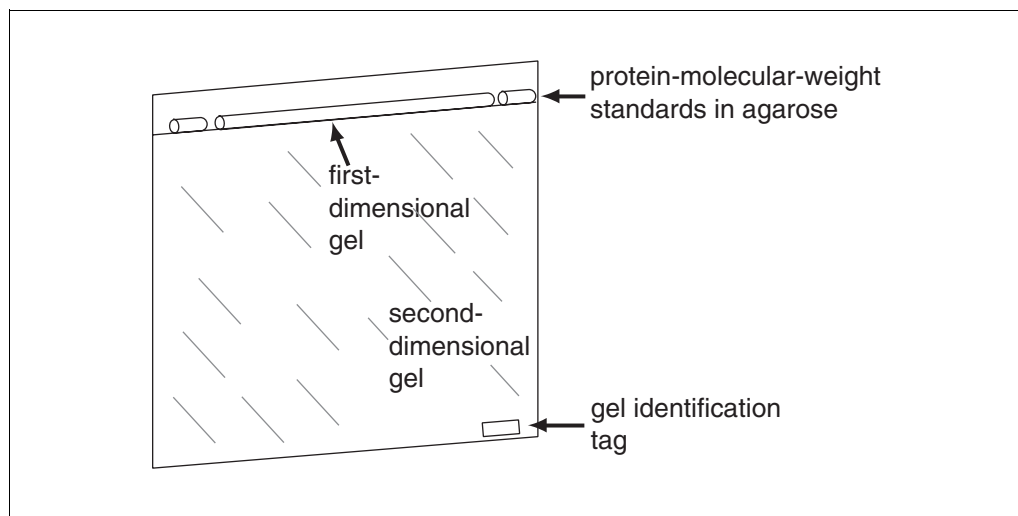


Figure 10.4.2 Assembly of second-dimension gel.

Run and analyze the gel

15. Mount the gels on the PROTEAN II electrophoresis cell (the short plate goes toward the center of the unit).
16. Fill the upper and lower reservoirs with prechilled reservoir buffer.
17. Attach tubing for coolant to the in and out ports and start the flow of coolant to maintain the temperature of the tank buffer at 10° to 20°C during the run to ensure that the gels are adequately cooled.

The tank buffer should be the same temperature for each run.

18. Attach the electrical leads to the power supply (the upper reservoir is connected to the negative lead). Electrophorese at 15 to 20 mA/gel until the tracking dye reaches the end of the gel (or 3 to 5 mA/gel overnight).
19. Reduce the voltage setting to zero and turn off the power supply at the end of the run. Remove the gels from the electrophoresis unit and take off the clamps. Pry the glass plates apart with a spatula.
20. Stain the gels (*UNIT 10.6*), process the gels for immunoblotting (*UNIT 10.8*), or autoradiograph the gels (*APPENDIX 3A*).

***ALTERNATE
PROTOCOL 1***

ISOELECTRIC FOCUSING OF VERY BASIC PROTEINS USING NEPHGE

First-dimension IEF gels run in the standard way using broad-range ampholytes will resolve proteins with isoelectric points (pI) between 3.8 and 8. More basic or more acidic proteins may be resolved by modifying the first-dimension gels using nonequilibrium pH gradient electrophoresis (NEPHGE) as described below (Anderson, 1988).

Additional Materials (also see *Basic Protocol 1*)

- Ampholytes, pH 2 to 11 (Serva)
- 0.01 M phosphoric acid, deaerated
- 4 M urea, deaerated

To analyze very basic proteins, the procedure is the same as described in *Basic Protocol 1* for first-dimension gels with the following exceptions in the indicated steps:

- 6a. Use 0.75 ml ampholytes, pH 2 to 11, in gel solution.
- 15a. Fill the lower reservoir with ~3 liters deaerated 0.02 M NaOH.

- 17a. Fill the upper reservoir with 250 ml deaerated 0.01 M phosphoric acid.
- 19a. Do not refocus gels.
- 20a. Overlay samples with deaerated 4 M urea to protect proteins from phosphoric acid.
- 21a. Attach the negative lead to the lower tank and the positive to the upper tank (this is the reverse of the usual setup).
- 22a. Focus 1 hr at 400 V and then 4 to 5 hr at 800 V for a total of 4000 V-hr.

Because the bromphenol blue applied to the top of the gel at the end of the run now marks the acidic end, load first-dimension gels onto the second-dimension gel with the blue end to the left.

ISOELECTRIC FOCUSING OF VERY ACIDIC PROTEINS USING NEPHGE

Additional Materials (also see Basic Protocol 1)

Ampholytes, pH 2.5 to 4 (Pharmacia LKB)
Concentrated sulfuric acid
Water, deaerated

To analyze very acidic proteins, the procedure is the same as described in Basic Protocol 1 for first-dimension gels with the following exceptions in the indicated steps:

6b. Prepare the gel solution as follows:

8.25 g urea
5.5 ml H₂O
2.0 ml 30% acrylamide/1.8% bisacrylamide
1.0 ml ampholytes, pH 2.5 to 4 (Pharmacia LKB)
0.3 ml ampholytes, pH 2 to 11 (Serva).

- 10b. Add 90 μ l of 10% ammonium persulfate and swirl, then add 10 μ l TEMED and swirl.
- 15b. Prepare lower buffer by adding 4.5 ml concentrated sulfuric acid to 3 liters of water and fill the lower reservoir.
- 17b. Prepare upper buffer by adding 3 ml ampholyte, pH 2 to 11, to 120 ml deaerated water and place in the upper reservoir.
- 22b. Run at 800 V for 4.5 to 5.0 hr or 250 V for 16 hr for a total of 3600 to 4000 V-hr.

TWO-DIMENSIONAL MINIGELS

Though limited in resolving area, small two-dimensional SDS-PAGE is a quick way to separate proteins for a variety of applications. These include peptide sequencing, purity checks, and protocol development. Where a large-format two-dimensional SDS-PAGE separation might take two days, a minigel separation (*UNIT 10.2A*) takes 4 to 5 hr, including the first-dimension (IEF) gel. The solutions and procedures in the basic protocols directly translate to the small format, except that much less reagent is required and electrophoresis times are much shorter.

To perform two-dimensional SDS-PAGE in the small format, a few simple changes from the basic protocol are needed. Tube gel adaptors for IEF, such as the Hoefer SE 220 or the Bio-Rad Mini-Protean II Tube Cell, fit in the minigel unit (*UNIT 10.2A*). Because the tube gels are much shorter (6 to 8 cm), the isoelectric focusing time is less; 2 to 4 hr at 500 V is usually adequate. Furthermore, less protein is generally required. Stacking gels

**ALTERNATE
PROTOCOL 2**

**ALTERNATE
PROTOCOL 3**

**Analysis of
Proteins**

10.4.9

**ALTERNATE
PROTOCOL 4**

are not normally required, so the mini slab gel should be cast to within 0.5 to 1 cm of the top. The second-dimension gel is processed in the same way as a first-dimension minigel (*UNIT 10.2A*).

**TWO-DIMENSIONAL (2-D) ELECTROPHORESIS WITH IMMOBILIZED
pH GRADIENTS**

In Basic Protocol 1, first-dimension separation is performed in carrier ampholyte-containing polyacrylamide gels cast in narrow tubes. In this protocol, the carrier ampholyte-generated pH gradients have been replaced with immobilized pH gradients and the tube gels have been replaced with thin rectangular IEF gels supported by a plastic backing (Westermeier et al., 1983; Berkelman and Stenstedt, 1998; Görg et al., 2000). Such gradients produce superior resolution and reproducibility (Bjellqvist et al., 1982). This protocol describes the use of the Ettan IPGphor isoelectric focusing system (Amersham Biosciences), but the PROTEAN IEF System from Bio-Rad is also popular. The cost is higher than when Basic Protocols 1 and 2 are used, but many laboratories have adopted these systems. The use of commercial quality-controlled and standardized reagents and instrumentation greatly simplifies the process of 2-D analysis and makes it possible to obtain much more consistent results.

Materials

- Protein sample for analysis
- Sample solution: default sample solution, hydrophobic sample solution, or tissue sample solution (see recipes)
- Tissue homogenization solution (see recipe)
- Rehydration stock solution (see recipe)
- IPG buffer or Pharmalyte (same range as the IPG strip; Amersham Biosciences)
- Cleaning solution (e.g., Ettan IPGphor Strip Holder Cleaning Solution; Amersham Biosciences)
- IPG Dry Strip Cover Fluid (Amersham Biosciences)
- Vertical gel for SDS-PAGE (Basic Protocol 2 in *UNIT 10.2A*)
- SDS equilibration buffer (see recipe)
- SDS electrophoresis buffer (*UNIT 10.2A*)
- Molecular weight markers (*UNIT 10.2A*)
- Agarose sealing solution (see recipe)
- Isoelectric focusing system (Ettan IPGphor, Amersham Biosciences; PROTEAN IEF System, Bio-Rad)
- IPG strips (Amersham Biosciences)
- Toothbrush
- Platform rocker
- 100°C heating block
- Thin plastic ruler
- Additional reagents and equipment for second-dimension gel electrophoresis (see Basic Protocol 2)

Prepare the sample

- 1a. *For an unknown protein:* Dissolve sample in default sample solution.
- 1b. *For large and more hydrophobic proteins:* Dissolve sample in hydrophobic sample solution.
- 1c. *For proteins from tissues that are dilute sources of protein and contain high levels of interfering substances (e.g., plant tissues):* Grind tissue in mortar and pestle with liquid nitrogen. Suspend powder in tissue homogenization solution. Keep at -18°C

Table 10.4.2 Suitable Sample Loads for Silver and Coomassie Staining Using Rehydration Loading

Immobiline DryStrip		Suitable sample load (μg of protein)		Rehydration solution volume, per strip
Size	pH range	Silver stain	Coomassie stain	
7 cm	4–7	4–8	20–120	125 μl
	6–11	8–16	40–240	
	3–10, 3–10 NL	2–4	10–60	
11 cm	4–7	10–20	50–300	200 μl
	6–11	20–40	100–600	
	3–10 L	4–8	20–120	
13 cm	4–7	15–30	75–450	250 μl
	6–11	30–60	150–900	
	3–10, 3–10 NL	8–15	40–240	
18 cm	4–7	30–60	150–900	340 μl
	6–11, 6–9, narrow interval ^a	60–120	300–1500	
	3–10, 3–10 NL	15–30	75–450	
24 cm	4–7, 3–7	45–90	200–1300	450 μl
	6–9, narrow interval ^a	80–170	400–2000	
	3–10, 3–10 NL	20–40	100–600	

^aImmobiline DryStrip narrow intervals pH: 3.5–4.5, 4.0–5.0, 4.5–5.5, 5.0–6.0, and 5.5–6.7.

overnight and centrifuge. Wash pellet with acetone. Dry and suspend in tissue sample solution.

This method produces protein solutions substantially free of salts, nucleic acids, and other contaminants.

Rehydrate IPG strips

IPG strips must be rehydrated prior to IEF. With integrated IEF systems such as the Amersham IPGphor Isoelectric Focusing and Bio-Rad PROTEAN system, both rehydration of the IPG strip and IEF occur in individual strip holders or trays. Different strip holder lengths are available for different IPG strip lengths. If the strip holders are in use for IEF, rehydration may be carried out in the Immobiline DryStrip Reswelling Tray (see Support Protocol 2).

2. Just prior to use, slowly thaw a 2.5-ml aliquot of rehydration stock solution. Add the appropriate amount (see Reagents and Solutions) of IPG Buffer or Pharmalyte, if it is not already included in the rehydration stock solution, and add 7 mg DTT.

Sample may be applied now (“rehydration loading”), or at step 10. If rehydration loading is desired, refer to Table 10.4.2 for suitable sample loads.

DTT and the sample must be added fresh, just prior to use.

3. Select the strip holder(s) corresponding to the IPG strip length chosen for the experiment.

Handle the ceramic strip holders with care, as they are fragile.

4. Wash each strip holder with detergent to remove residual protein, using a neutral-pH detergent (e.g., Ettan IPGphor Strip Holder Cleaning Solution) as follows.

- a. Rinse off the strip holder.

A mild liquid soap may be used to remove any residual DryStrip cover fluid.

Table 10.4.3 Recommended Final Concentrations and Volumes of IPG Buffer/Pharmalyte for the Rehydration Solution^a

Strip holder	Rehydration and sample well loading	Vertical gels, flatbed	0.5% IPG Buffer (12.5 μ l per 2.5 ml)
Cup loading	Cup holder	Vertical gels, flatbed	2% IPG Buffer (50 μ l per 2.5 ml)

^aThe recommended IPG Buffer / Pharmalyte concentration for the IPGphor system is 0.5%, but up to 2% can be added if sample solubilization remains a problem.

- b. Squeeze a few drops of Ettan IPGphor Strip Holder Cleaning Solution into the strip holder slot. Use a toothbrush and vigorous agitation to clean the strip holder.
- c. Rinse well with distilled or deionized water. Thoroughly air dry the strip holders or dry well with a lint-free tissue prior to use.

The Ettan IPGphor Strip Holder Cleaning Solution has been specifically formulated for removing protein deposits and will not damage the strip holder.

Clean strip holders after each first-dimension IEF run. Do not let solutions dry in the strip holder. Cleaning may be more effective if the strip holders are first soaked a few hours overnight in a solution of 2% to 5% Ettan IPGphor Strip.

Recalcitrant or dried-on protein deposits may be removed with hot (up to 95°C) 1% (w/v) SDS. Add 1% (w/w) DTT for complete removal of sticky proteins. Rinse completely with distilled or deionized water after cleaning. Handle clean strip holders with gloves to avoid contamination. Strip holders may be baked, boiled, or autoclaved. Do not expose them to strong acids or bases, including alkaline detergents.

5. Pipet the appropriate volume of rehydration solution into each strip holder as indicated in Table 10.4.3. Deliver the solution slowly at a central point in the strip holder channel away from the sample application wells. Remove any larger bubbles.

To ensure complete sample uptake, do not exceed the recommended volume of rehydration solution (see Table 10.4.3).

6. Place the IPG strip in the strip holder as follows.
 - a. To prevent damage to the basic (square) end of the IPG strip, which is generally softer, remove the protective cover foil from the IPG strip starting at the acidic (pointed) end.
 - b. Position the IPG strip with the gel side down and the pointed (anodic) end of the strip directed toward the pointed end of the strip holder.
 - c. Pointed end first, lower the IPG strip onto the solution. To help coat the entire strip, gently lift and lower the strip and slide it back and forth along the surface of the solution, tilting the strip holder slightly as needed to assure complete and even wetting.
 - d. Finally, lower the cathodic (square) end of the IPG strip into the channel, making sure that the gel contacts the strip holder electrodes at each end.

The gel can be visually identified once the rehydration solution begins to dye the gel. Be careful not to trap air bubbles under the IPG strip.

7. Apply IPG Dry Strip Cover Fluid to minimize evaporation and urea crystallization. Pipet the fluid dropwise into one end of the strip holder until one half of the IPG strip is covered, then pipet the fluid dropwise into the other end of the strip holder, adding fluid until the entire IPG strip is covered.

8. Place the cover on the strip holder.

Pressure blocks on the underside of the cover assure that the IPG strip maintains good contact with the electrodes as the gel swells.

9. Allow the IPG strip to rehydrate.

For rehydration loading of sample (see annotation to step 2), there are two possible rehydration conditions: passive rehydration with no electric field applied during rehydration, and rehydration under voltage. In some cases, rehydration under a low voltage (30 to 120 V) facilitates the entry of high-molecular weight proteins.

Rehydration can proceed on the bench top or on the Ettan IPGphor unit platform. Ensure that the strip holder is on a level surface. A minimum of 10 hr is required for rehydration; overnight is recommended. The rehydration period can be programmed as the first step of an Ettan IPGphor protocol. This is especially convenient if temperature control during rehydration is a concern.

Perform first-dimension electrophoresis

10. If the sample was not applied as a part of the rehydration solution in step 2, pipet ≥ 7.5 μl of sample solution prepared in step 1a, b, or c (i.e., 15 μl per well or 30 μl total if both sides of both wells are used) to the lateral wells at either end of the strip holder, below the IPG DryStrip Cover Fluid. Replace cover on strip holder.

When using the alternative “cup loading” procedure (also see instrumentation manual) an increased sample concentration will lead to an increased risk of protein precipitation in the sample cup. A maximum concentration of 100 μg protein/100 μl sample solution (100 μl is the volume of the cup) is recommended. This is a general recommendation, which will work for most samples, but the maximum concentration that can be used varies greatly between sample types. For larger sample loads, rehydration loading (see annotation to step 2 and Table 10.4.2) is recommended.

The IPG strip backing is impermeable; do not apply the sample to the back of the strip.

11. Perform first-dimension IEF separation per manufacturer’s instructions. See Table 10.4.4 for guidelines on different experimental setups.

Equilibrate strips prior to second-dimension electrophoresis

The second-dimension vertical gel must be ready for use prior to IPG strip equilibration. See Basic Protocol 2 for preparation of vertical gels for 2-D SDS PAGE.

12. Place the IPG strips in individual tubes with the support film toward the wall of the tube. Add 10 ml SDS equilibration buffer (containing DTT) to each tube. Cap the tube, and place it on its side on a platform rocker. Equilibrate for 15 min.
13. If desired, perform a second equilibration with SDS equilibration buffer containing 250 mg/ml iodoacetamide (without DTT). Add 10 ml of solution per tube. Cap the tube, place it on its side on a platform rocker, and equilibrate for 15 min.

This second equilibration step reduces point streaking and other artifacts.

14. Dip the equilibrated IPG strip in SDS electrophoresis buffer to lubricate it.
15. *Optional:* Apply molecular weight marker proteins to a separate small strip.

Best results are obtained when the molecular weight marker protein solution is mixed with an equal volume of a hot 1% agarose solution prior to application to the IEF sample application piece. The resultant 0.5% agarose will gel and prevent the marker proteins from diffusing laterally prior to the application of electric current. Other alternatives are to apply the markers to a paper IEF sample application piece in a volume of 15 to 20 μl . For less volume, cut the sample application piece proportionally. Place the IEF application piece on a glass plate and pipet the marker solution onto it, then pick up the application piece with forceps and apply to the top surface of the gel next to one end of the IPG strip. The markers should contain 200 to 1000 ng of each component for Coomassie staining and about 10 to 50 ng of each component for silver staining.

16. Melt each aliquot of agarose sealing solution needed (each gel will require 1 to 1.5 ml) in a 100°C heat block (it takes ~10 min to fully melt the agarose; an ideal time to carry out this step is during IPG strip equilibration). Allow the agarose to cool until

Table 10.4.4 Guidelines for Ettan IPGphor with Rehydration Loading/IEF for Immobiline Dry Strips^a

Immobiline DryStrip	pH range(s)	Step and voltage mode	Step duration ^b (h:min)	Kilovolt-hours (kV-hr)
7 cm	3–10	1. Step and hold 500	0:30	0.25
	3–10 NL	2. Step and hold 1000	0:30	0.5
	4–7	3. Step and hold 5000 ^c	1:40	7.5
		Total	2:40	8.0
11 cm	3–10	1. Step and hold 500	1:00	0.5
	4–7	2. Step and hold 1000	1:00	1.0
		3. Step and hold 8000 ^c	1:50	12.5
		Total	3:50	14.0
13 cm	3–10	1. Step and hold 500	1:00	0.5
	3–10 NL	2. Step and hold 1000	1:00	1.0
	4–7	3. Step and hold 8000 ^c	2:00	14.5
		Total	4:00	16.0
18 cm	3–10	1. Step and hold 500	1:00	0.5
	3–10 NL	2. Step and hold 1000	1:00	1.0
	4–7	3. Step and hold 8000 ^c	4:00	30.5
		Total	6:00	32.0
	Narrow intervals ^d	1. Step and hold 500	1:00	0.5
		2. Step and hold 1000	1:00	1.0
		3. Step and hold 1000 ^c	7:30	58.5
		Total	9:30	60.0
24 cm	3–10	1. Step and hold 500	1:00	0.5
	3–10 NL	2. Step and hold 1000	1:00	1.0
	4–7	3. Step and hold 8000 ^c	8:20	62.5
		Total	10:20	64.0
	Narrow intervals ^d	1. Step and hold 500	1:00	0.5
		2. Step and hold 1000	1:00	1.0
		3. Step and hold 8000 ^c	10:30	94.5
		Total	12:30	96.0

^aVoltage step and hold mode, 50 μ A/IPG strip; 0.5% IPG buffer; 20°C for both rehydration and IEF. Rehydration time, 12 hr. The total rehydration time can be adjusted somewhat for convenience, but must be >10 hr.

^bThe sample entry phase step 1 and 2 should be extended for high protein loads, or for convenience, if the strips are to be run overnight.

^cThis voltage may not be reached within the suggested step duration.

^dNarrow intervals, 3.5–4.5, 4.0–5.0, 4.5–5.5, 5.0–6.0, and 5.5–6.7.

the tube can be held with fingers (60°C) and then slowly pipet the amount required to seal the IPG strip in place, avoiding introduction of bubbles. Only apply the minimum quantity of agarose sealing solution required to cover the IPG strip. Allow a minimum of 1 min for the agarose to cool and solidify.

In any vertical gel system, the agarose sealing serves to prevent the IPG strip from moving or floating in the electrophoresis buffer. If precast Ettan DALT gels are used for the second-dimension electrophoresis, the agarose sealing also serves to block the narrow gap(s) between the gel edge(s) and the lateral spacer(s) to prevent leakage of the upper buffer.

17. Center the IPG strip between the plates on the surface of the second-dimension gel with the plastic backing against one of the glass plates. With a thin plastic ruler, gently push the IPG strip down so that the entire lower edge of the IPG strip is in contact with the top surface of the slab gel. Ensure that no air bubbles are trapped between the IPG strip and the slab gel surfaces or between the gel backing and the glass plate.
18. Perform second-dimension electrophoresis (see Basic Protocol 2).

REHYDRATION OF IPG STRIPS USING THE IMMOBILINE DRYSTRIP RESWELLING TRAY

The IPG strips are rehydrated in the Immobiline DryStrip Reswelling Tray if strip holders are in use for IEF. The Immobiline DryStrip Reswelling Tray has 12 independent reservoir slots that can each hold a single IPG strip up to 24 cm long. Separate slots allow the rehydration of individual IPG strips in a minimal volume of solution.

Additional Materials (also see Alternate Protocol 4)

Immobiline DryStrip Reswelling Tray (Amersham Biosciences)

1. Slide protective lid of reswelling tray completely off the tray and level tray by turning the leveling feet until the bubble in the spirit level is centered. Ensure that tray is clean and dry.
2. Pipet appropriate volume of rehydration solution into each slot as indicated in Table 10.4.3. Mix with sample solution (see Alternate Protocol 2, step 1) for rehydration loading (see Alternate Protocol 4, step 2 annotation). Deliver solution slowly at a central point in the slot. Remove any larger bubbles.

To ensure complete fluid (and sample) uptake, do not apply excess rehydration solution.

3. Remove protective cover from IPG strip starting at the acidic (pointed) end. Position IPG strip with gel side down and pointed end of strip against the sloped end of the slot. Lower IPG strip onto solution. To coat the entire IPG strip, carefully lift and lower the strip and slide it back and forth along the surface of the solution. Do not trap bubbles under the IPG strip.
4. Overlay the IPG strip with 3 ml DryStrip Cover Fluid to minimize evaporation and prevent urea crystallization.
5. Slide lid onto reswelling tray and allow IPG strips to rehydrate at room temperature.

A minimum of 10 hr is required for rehydration; overnight is recommended.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Acrylamide/bisacrylamide solutions

30% acrylamide/1.8% bisacrylamide (first dimension):

30 g acrylamide

1.8 g bisacrylamide

H₂O to 100 ml

30% acrylamide/0.8% bisacrylamide (second dimension):

300 g acrylamide

8.0 g bisacrylamide

H₂O to 1 liter

Filter through 0.2- to 0.45- μ m filter. Store in tightly capped amber bottles at 4°C. Discard after 30 days, as acrylamide gradually hydrolyzes to acrylic acid and ammonia.

CAUTION: *Acrylamide monomer is neurotoxic. Gloves should be worn while handling the solution, and the solution should not be pipetted by mouth. Wear a mask when weighing out the solid acrylamide.*

Agarose sealing solution

100 ml SDS electrophoresis buffer (*UNIT 10.2A*)
0.5 g agarose (NuSieve GTG, Cambrex)
200 μ l 1% bromphenol blue stock (see recipe; 0.002% w/v final)

Combine all ingredients in a 500-ml Erlenmeyer flask. Swirl to disperse. Heat in a microwave oven on “low” or a stirring hotplate until the agarose is completely dissolved. Do not allow the solution to boil over. Dispense 2-ml aliquots into screw-cap tubes and store up to 1 week at 4°C.

Agarose solution, 0.5%

0.25 g agarose (standard low- M_r ; Bio-Rad)
50 ml reservoir buffer (see recipe)

Heat in a boiling water bath to dissolve the agarose and keep in a boiling water bath. Prepare fresh each time.

Agarose solution, 1%

Prepare as for hot 0.5% agarose solution (see recipe), substituting 0.5 g agarose for 0.25 g agarose.

Ammonium persulfate, 10%

10 g ammonium persulfate (Bio-Rad)
H₂O to 100 ml
Store refrigerated \leq 2 weeks

Bromphenol blue stock solution, 1%

100 mg bromphenol blue (Sigma; 1% w/v final)
60 mg Tris base (50 mM final)
Double-distilled H₂O to 10 ml
Prepare fresh

Chromic acid cleaning solution

Add a 25 ml bottle of Chromerge (Fisher) to a 9-lb bottle of concentrated sulfuric acid. Add ~5 ml at a time and stir.

CAUTION: *This is very corrosive and toxic. Carefully read and observe package instructions.*

Concentrated bromphenol blue

50% aqueous glycerol (v/v)
0.01 mg/ml bromphenol blue
Prepare fresh

Default sample solution

8 M urea
4% (w/v) CHAPS
60 mM DTT
2% (v/v) Pharmalyte 3–10 (Amersham Biosciences)
0.002% (w/v) bromphenol blue (add from 1% stock; see recipe)
Store up to 6 months at –20°C

Equilibration buffer

3.75 g Tris base
25 ml glycerol
5.25 g SDS
333 mg dithiothreitol (DTT)

Dissolve Tris base in H₂O and adjust pH to 6.8 with 6 M HCl. Add other ingredients. Add H₂O to 250 ml final volume. This buffer can be stored for up to 1 week at room temperature.

Gel buffer

Dissolve 90.8 g Tris base in 300 ml H₂O. Adjust to pH 8.6 with 6 M HCl. Add H₂O to 500 ml. This buffer can be stored for several weeks in the refrigerator.

Hydrophobic sample solution

7 M urea
2 M thiourea
4% (w/v) CHAPS
60 mM DTT
2% (v/v) Pharmalyte pH 3–10 (Amersham Biosciences)
0.002% (w/v) bromphenol blue (add from 1% stock; see recipe)
Store up to 6 months at –20°C

NaOH, 0.02 M

Just before using, add 0.5 ml of 10 M NaOH to 250 ml freshly deaerated water. Store up to 1 month at room temperature.

It is especially important to make 0.02 M NaOH with deaerated water.

Phosphoric acid, 0.085%

Just before using, dilute 300 ml of 0.85% phosphoric acid to 3.0 liters with deaerated water. Store up to 1 month at room temperature.

Rehydration stock solution

12 g urea (FW 60.06; 8 M final)
0.5 g CHAPS (2% w/v final)
50 µl 1% bromphenol blue stock (see recipe; 0.002% w/v final)
Double distilled H₂O to 25 ml
Store stock with above components in 2.5-ml aliquots up to 1 month at –20°C
Just prior to use add 7 mg DTT per 2.5-ml aliquot and IPG Buffer or Pharmalyte (same range as IPG strip) to 0.5% (for IPGphor apparatus) or 2% (for Multiphor II and Immobiline DryStrip kit). Use 125 µl IPG Buffer for a 0.5% concentration and 500 µl IPG Buffer for a 2% concentration. Use Pharmalyte 3–10 for Immobiline DryStrip 3–10 or 3–10 NL and Pharmalyte 5–8 for Immobiline DryStrip 4–7. For rehydration loading (see Alternate Protocol 4, step 2 annotation), also add sample to the 2.5 ml aliquot of rehydration solution just prior to use (see Table 10.4.2 for suitable sample loads).

If necessary, the concentration of urea can be increased to 9 or 9.8 M. Other detergents (Triton X-100, NP-40, and other nonionic or zwitterionic detergents) can be used instead of CHAPS.

Reservoir buffer

15.0 g Tris base
72.0 g glycine
5.0 g SDS
H₂O to 5 liters
For convenience, make up as a 10× stock or store the preweighed dry ingredients in packets for future use. The 10× stock can be stored for several weeks in the refrigerator.

SDS equilibration buffer

10 ml 1.5 M Tris·Cl, pH 8.8 (APPENDIX 2; 50 mM final)
72.07 g urea (FW 60.06; 6 M final)
69 ml 87% (v/v) glycerol (30% v/v final)
4 g SDS (FW 288.38; 2% w/v final)

continued

400 μ l 1% bromphenol blue stock (see recipe; 0.002% w/v final)
Double-distilled H₂O to 200 ml
Store solution with above components at -20°C
Just before use add:
100 mg DTT or 250 mg iodoacetamide

SDS solubilization buffer

0.1 g 2-(*N*-cyclohexylamino)ethanesulfonic acid (CHES)
0.2 g SDS
0.1 g DTT
1.0 ml glycerol
H₂O to 10 ml total volume
Store aliquots at -70°C

Stacking gel buffer

15.0 g Tris base
1.0 g SDS
Dissolve Tris base and SDS in 200 ml H₂O. Adjust to pH 6.8 with 6 M HCl. Add H₂O to 250 ml. The 5 \times stock can be stored for several weeks in the refrigerator.

Tissue homogenization solution

Prepare in acetone:
10% (w/v) trichloroacetic acid (TCA)
0.3% (w/v) DTT
Prepare fresh

Tissue sample solution

9 M urea
2% (w/v) CHAPS
1% (w/v) DTT
2% (v/v) Pharmalyte 3–10
Store up to 1 month at -20°C

Urea solubilization buffer

5.4 g urea
0.4 ml NP-40
1.0 ml 20% (w/v) stock ampholyte (pH 9 to 11)
0.2 ml 2-ME (Kodak)
H₂O to 10 ml total volume; if pH is <9.5 , add 2 M NaOH
Store aliquots up to 1 month at -70°C

COMMENTARY

Background Information

Two-dimensional gel electrophoresis is an extremely powerful analytical tool for characterization of complex protein mixtures (O'Farrell, 1975). No other procedure can resolve so many proteins in a single operation. Therefore, two-dimensional electrophoresis is an excellent way to examine protein changes between control and experimental samples. Because charge and molecular weight are monitored simultaneously, small alterations in proteins can be detected. Increased resolution with two-dimensional gels facilitates quantitative analy-

sis of proteins. With very few exceptions, each spot on a two-dimensional gel contains only one protein, whereas a band on a one-dimensional gel (UNIT 10.2A) may contain more than one protein. (Single proteins can yield multiple spots due to variability in sialylation. Treatment of samples with neuraminidase eliminates this type of heterogeneity; follow instructions from Genzyme for use of neuraminidase.) It is impossible to accurately assess quantitative changes with more than one protein in a band because individual proteins can change independently.

In addition to analytical applications, two-dimensional PAGE can provide a means of collecting small amounts of extremely pure protein for amino acid sequence analysis (Hirano, 1989) or antibody production (Dunbar, 1987). Preparative and analytical studies using two-dimensional gels require running a large number of gels. The ISO-DALT system (UNIT 10.3; Anderson, 1988) is specifically designed for running numerous two-dimensional gels.

Minigels (UNIT 10.2) have become popular for one-dimensional separation of proteins. The small size (8 × 10-cm, or 10 × 10-cm) of minigels makes it possible to run and stain the gels in much less time than standard sized gels. It is possible to use mini slab gels as the second dimension of a two-dimensional separation when small first-dimension gels (equipment available from Bio-Rad and Hoefer) are used. The mini two-dimensional systems can save a considerable amount of time and reagents, but they cannot resolve large numbers of proteins present in complex mixtures such as a whole-cell lysate. Therefore, it is recommended that the use of minigels be restricted primarily to simple protein mixtures.

Proteins on two-dimensional gels may be visualized by staining procedures (UNIT 10.6) or by autoradiography (APPENDIX 3A) of radiolabeled proteins (UNIT 10.18). A large number of extremely sensitive silver staining methods have been published. The most sensitive of these stains (Ochs et al., 1981) can detect ≤1 ng of most proteins. Although silver staining is a very sensitive method, impurities in reagents and water are often detected as well as proteins. The staining sensitivity can vary from one protein to another. This can be particularly troublesome when doing quantitative analysis of two-dimensional patterns.

A large and growing application of 2-D electrophoresis is “proteome analysis.” Proteome analysis is “the analysis of the entire PROTEin complement expressed by a genOME” (Wilkins et al., 1996). The analysis involves the systematic separation, identification, and quantification of many proteins simultaneously from a single sample. Two-dimensional electrophoresis is used in this technique due to its unparalleled ability to separate thousands of proteins simultaneously. Two-dimensional electrophoresis is also unique in its ability to detect post- and co-translational modifications, which cannot be predicted from the genome sequence. Applications of 2-D electrophoresis include proteome analysis, cell differentiation, detection of disease markers, moni-

toring therapies, drug discovery, cancer research, purity checks, and microscale protein purification. New mass spectrometry techniques have been developed that allow rapid identification and characterization of very small quantities of peptides and proteins extracted from single 2-D spots.

In the newer Immobiline systems (see Alternate Protocol 4), the 2-D process begins with sample preparation. Proper sample preparation is absolutely essential for a good 2-D result. The next step in the 2-D process is IPG strip rehydration. IPG strips are provided dry and must be rehydrated with the appropriate additives prior to IEF. First-dimension IEF is performed on a flatbed system at very high voltages with active temperature control. Next, strip equilibration in SDS-containing buffer prepares the sample for the second-dimension separation. Following equilibration, the strip is placed on the second-dimension gel for SDS-PAGE. The final steps are visualization and analysis of the resultant two-dimensional array of spots.

Principles of isoelectric focusing (IEF)

IEF is an electrophoretic method that separates proteins according to their isoelectric point (pI). Proteins are amphoteric molecules; they carry either positive, negative, or zero net charge, depending on the pH of their surroundings. The net charge of a protein is the sum of all the negative and positive charges of its amino acid side chains and amino and carboxyl-termini. The isoelectric point (pI) is the specific pH at which the net charge of the protein is zero. Proteins are positively charged at pH values below their pI and negatively charged at pH values above their pI. If the net charge of a protein is plotted versus the pH of its environment, the resulting curve intersects the x axis at the isoelectric point.

The presence of a pH gradient is critical to the IEF technique. In a pH gradient, under the influence of an electric field, a protein will move to the position in the gradient where its net charge is zero. A protein with a positive net charge will migrate toward the cathode, becoming progressively less positively charged as it moves through the pH gradient until it reaches its pI. A protein with a negative net charge will migrate toward the anode, becoming less negatively charged until it also reaches zero net charge. If a protein should diffuse away from its pI, it immediately gains charge and migrates back. This is the focusing effect of IEF, which concentrates proteins at their pIs and allows

proteins to be separated on the basis of very small charge differences.

The resolution is determined by the slope of the pH gradient and the electric field strength. IEF is therefore performed at high voltages (typically >1000 V). When the proteins have reached their final positions in the pH gradient, there is very little ionic movement in the system, resulting in a very low final current (typically below 1 mA). IEF of a given sample in a given electrophoresis system is generally performed for a constant number of volt-hours (V-hr), this unit being the integral of the volts applied over the time.

IEF performed under denaturing conditions gives the highest resolution and the cleanest results. Complete denaturation and solubilization is achieved with a mixture of urea and detergent, ensuring that each protein is present in only one configuration and aggregation and intermolecular interaction is minimized.

The original method for first-dimension IEF depended on carrier ampholyte-generated pH gradients in polyacrylamide gel rods in tubes (O'Farrell, 1975). Carrier ampholytes are small, soluble, amphoteric molecules with a high buffering capacity near their pI. Commercial carrier ampholyte mixtures are comprised of hundreds of individual polymeric species with pIs spanning a specific pH range. When a voltage is applied across a carrier ampholyte mixture, the carrier ampholytes with the highest pI (and the most negative charge) move toward the anode and the carrier ampholytes with the lowest pI (and the most positive charge) move toward the cathode. The other carrier ampholytes align themselves between the extremes, according to their pIs, and buffer their environment to the corresponding pHs. The result is a continuous pH gradient.

Although the technique described in Basic Protocol 1 is commonly used in 2-D electrophoresis studies, it has several limitations that have prevented its more widespread application. Carrier ampholytes are mixed polymers that are not well characterized and suffer from batch-to-batch manufacturing variations. These variations can reduce the reproducibility of the first-dimension separation. Carrier ampholyte pH gradients are unstable and have a tendency to drift, usually toward the cathode, over time. Gradient drift adversely affects reproducibility by introducing a time variable. Gradient drift also causes a flattening of the pH gradient at each end, particularly above pH 9, rendering the 2-D technique less useful at pH

extremes. Finally, the soft polyacrylamide tube gels used in this procedure have low mechanical stability. The gel rods may stretch or break, affecting reproducibility. Results are often dependent on the skill of the operator.

As a result of the limitations and problems with carrier ampholyte pH gradients, immobilized pH gradients (Alternate Protocol 4) were developed and Immobiline chemicals were introduced for the generation of this type of pH gradient. Westermeier (1983) and Görg et al. (2000) pioneered the development and use of IPG IEF for the first-dimension of 2-D electrophoresis. An immobilized pH gradient (IPG) is created by covalently incorporating a gradient of acidic and basic buffering groups into a polyacrylamide gel at the time it is cast. The buffers are composed of a single acidic or basic buffering group linked to an acrylamide monomer.

The general structure of Immobiline reagents is:



where R = weakly acidic or basic buffering group. Immobilized pH gradients are formed using two solutions, one containing a relatively acidic mixture of acrylamide buffers and the other containing a relatively basic mixture. The concentrations of the various buffers in the two solutions define the range and shape of the pH gradient produced. Both solutions contain acrylamide monomers and catalysts. During polymerization, the acrylamide portion of the buffers copolymerize with the acrylamide and bisacrylamide monomers to form a polyacrylamide gel. To help in handling and performance, the IPG gel is cast onto a plastic backing. It is then washed to remove catalysts and unpolymerized monomers, which could otherwise modify proteins and interfere with separation. Finally the gel is dried and cut into 3-mm-wide strips. The resulting IPG strips can be rehydrated with a rehydration solution containing the necessary components for first-dimension IEF.

In the Immobiline systems, IEF is performed horizontally on a flatbed electrophoresis unit. Because IEF requires high field strengths to obtain sharply focused bands, high voltages must be applied. A flatbed design is the most economical way to meet the necessary safety standards required to operate at such high voltages. It also allows for cooling and close temperature control, which can be effectively achieved on a horizontal ceramic cooling plate

connected to a thermostatic circulator or a Peltier cooling plate. The first-dimension separation is more reproducible because the covalently fixed gradient cannot drift.

Critical Parameters

Two-dimensional PAGE is not difficult to perform but it is easy for problems to arise. Careful attention to detail is essential.

Several steps in the procedure need to be optimized for particular samples. Sample solubilization, for example, may require some experimentation to determine the best solubilization buffer and the optimum ratio of buffer to sample. For most tissue samples, 1.5 to 2.0 ml of urea solubilization buffer per 100 mg of tissue will give good results; however, the ratio of tissue to buffer may have to be altered in some instances. If solubilized samples are stored frozen (-70°C), avoid thawing and re-freezing.

In general, keep the sample preparation strategy as simple as possible to avoid protein losses. Additional sample preparation steps may improve the quality of the final 2-D result, but at the possible expense of selective protein loss. The cells or tissue should be disrupted in such a way as to minimize proteolysis and other modes of protein degradation. Cell disruption should be done at as low a temperature as possible and with a minimum of heat generation. Cell disruption should ideally be carried out directly by placing the cells into a strongly denaturing solution containing protease inhibitors (*UNIT 10.2*). Preserve sample quality by preparing the sample just prior to IEF or storing samples in aliquots at -80°C . Do not expose samples to repeated thawing. Remove all particulate material by ultracentrifugation. Solid particles and lipids must be removed because they will block the gel pores.

The sample solubilization solutions are in general optimized for either Basic Protocol 1 or Alternate Protocol 4. For Basic Protocol 1, either urea- or SDS-based sample solubilization may be used. Most proteins are soluble in either SDS or urea solubilization buffer. However, some proteins are more soluble in one or the other; therefore, it may be necessary to try both types of solubilization for a particular sample. In general, urea is more effective than SDS for tissue solubilization. Samples that have been solubilized in SDS can only be used with isoelectric focusing (IEF) gels containing urea and NP-40. Otherwise, the SDS remains associated with the proteins, giving them a strong negative charge and causing them to

migrate to the acidic end of the gel. The IPG technique outlined in Alternate Protocol 4 is particularly sensitive to residual SDS, which must be removed or reduced in concentration to $<0.25\%$ by dilution of the sample into rehydration solution (see Reagents and Solutions) containing nonionic detergent (e.g., CHAPS) prior to isoelectric focusing (Berkelman and Stenstedt, 1998). To avoid modification of proteins, never heat a sample after adding urea. When the sample contains urea, it must not be heated over 37°C . Elevated temperatures cause urea to hydrolyze to isocyanate, which modifies proteins by carbamylation.

Immunoprecipitates prepared as described in *UNIT 10.16* and solubilized in either SDS or urea solubilization buffer can be analyzed by two-dimensional electrophoresis. The sample should not contain $>2\%$ SDS. Samples in SDS solubilization buffer may be heated in a boiling water bath for 5 min to aid in protein solubilization. Heating is not always required and should never be done with samples in urea solubilization buffer.

The quality of the first-dimension separation depends on the ampholytes used, voltage, volt-hours, and reagent quality. For good resolution over a broad pH range, blends of ampholytes from different suppliers generally give better results than broad-range ampholytes from a single supplier. However, ISO-DALT-grade Resolyte, pH 4 to 8, made by BDH Chemicals and sold in the United States by Hoefer, works very well. Optimum voltage and volt-hours must be determined experimentally. Once suitable conditions have been established, the same conditions must be used for each run to achieve reproducibility.

When using homogeneous slab gels, the acrylamide concentration of the gel must be matched to the sample. Low-molecular-weight proteins can be resolved on high-percentage (18% to 20% acrylamide) gels while high-molecular-weight proteins must be run on low-percentage (7.5% to 12.5% acrylamide) gels. Gradient gels, such as 10% to 20% acrylamide gradients, are able to resolve proteins over a wide size range (250 to 10 kDa) on a single gel. Pouring gradient gels is more difficult than pouring homogeneous slabs, but with some practice the technique can be mastered (*UNIT 10.2*). One alternative is to buy precast gradient gels from a commercial supplier, such as Integrated Separation Systems.

IEF in the Ettan IPGphor system is conducted at very high voltages (up to 8000 V) and very low currents (typically $<50\ \mu\text{A}$ per IPG

strip) due to the low ionic strength within IPG strips. During IEF, the current decreases while the voltage increases as proteins and other charged components migrate to their equilibrium positions. A typical IEF protocol generally proceeds through a series of voltage steps that begin at a relatively low value. Voltage is gradually increased to the final desired focusing voltage, which is held for several hours. A low initial voltage minimizes sample aggregation and allows the parallel separation of samples with differing salt concentrations. A gradual increase in voltage is particularly advised for higher protein loads (100 μg or more per IPG strip).

Many factors affect the amount of time required for complete focusing, and each specific set of conditions (e.g., sample and rehydration solution composition, IPG strip length, and pH gradient) requires an empirical determination for optimal results. An approximate time for complete focusing is given in the example protocols provided in Table 10.4.4. Factors that increase the required focusing time include residual ions, which must move to the ends of the IPG strips before protein focusing can occur, and the presence of IPG Buffers or Pharmalyte, which contribute to the ionic strength of the electrophoresis medium. A higher IPG buffer concentration increases the conductivity of the IPG strip, resulting in a lower final voltage when the system is limited by the maximum current setting. Longer focusing times may therefore be required at IPG Buffer/Pharmalyte concentrations higher than 0.5%.

For higher protein loads (up to 1 mg or more) the final focusing step of each protocol can be extended with an additional 20% of the total recommended V-hr (see Background Information) if necessary. Using preparative protein loads and high salt concentrations in the Ettan IPGphor Cup Loading Strip Holder may cause the cover oil to redistribute and leak out of the strip holder. This can be counteracted by using a longer voltage ramp-up time. Note that exceeding the current limit of 50 μA /IPG strip is not recommended, as this may result in excessive heat generation and may damage the IPG strip and/or strip holder. Under extreme circumstances, the IPG strip may burn. Overfocusing can sometimes occur on longer runs and may contribute to horizontal streaking, visible in the 2-D result.

Troubleshooting

If all steps in this protocol are done correctly, the resulting two-dimensional gels should con-

tain numerous round or elliptical protein spots. Mammalian tissue or whole-cell samples typically yield two-dimensional patterns of >1000 silver-stained protein spots distributed throughout the second-dimension gel, with only a few spots overlapping. Purified protein samples such as immunoprecipitation may have only one major protein and a few minor ones.

The following checklists will help the beginner locate common problems associated with two-dimensional electrophoresis.

General. (1) Were all of the reagents fresh? (2) Were all of the reagents properly prepared? (3) Were the samples properly solubilized?

First dimension. Spots that are elongated horizontally (poorly focused), or a spot pattern in which most of the proteins are crowded into a narrowed zone along x axis of gel, should trigger the following questions: (1) Were samples ultracentrifuged just prior to loading? (2) Were desired number of volt-hours applied? (3) Were correct ampholytes used? (4) Was salt content of the sample low?

Second dimension. Vertically elongated spots and spot patterns that are vertically compressed along the y axis of the gel should trigger the following questions: (1) Was the gel solution properly deaerated? (2) Were the tops of the gels flat and smooth? (3) Was the reservoir buffer at the desired temperature throughout the run? (4) Was the reservoir buffer fresh and the proper concentration?

Anticipated Results

A two-dimensional electrophoretic separation of proteins similar to that shown in Figure 10.3.2 will be obtained. In this case, a 10% to 12% acrylamide second-dimension gel was used. Proteins between 10 and 250 kDa and pI 3.8 to 8.0 can be seen on this gel. Protein spots on a two-dimensional gel should be round or elliptical and separated from each other. A complex protein mixture such as mammalian tissue or whole cells should give a two-dimensional pattern of ≥ 1000 silver-stained spots distributed over most of the gel area. Fewer spots will be seen if stains less sensitive than silver stain are used (e.g., Coomassie blue; UNIT 10.6). Few of the protein spots should touch each other.

The 2-D gel procedures described will be suitable for subsequent staining via protocols listed in UNIT 10.6 and Lilley (2002), which allow both pre- and post-staining of samples for documentation, quantitation, and analysis by tandem mass spectrometry (MS/MS; UNIT 10.22) techniques.

Time Considerations

Sample preparation takes 10 to 20 min/sample. First-dimension gels can be used as soon as 1 hr after casting or they can be kept for at least 24 hr before use. It is convenient to start the first-dimension gels in the afternoon, so they can run overnight and be ready for the second-dimension separation early the following day. Second-dimension gels can be poured up to a week before they are to be used, provided they are stored refrigerated in a humid, air-tight container to prevent drying. The second dimension run takes ~5 hr. The two-dimensional gels will be ready to put in fixative or staining solution at the end of the day and can remain in that solution overnight. Usually, 3 working days are required to complete the entire procedure, including sample preparation and gel staining. Two-dimensional minigels take 4 to 5 hr.

Literature Cited

- Anderson, N.L. 1988. Two-Dimensional Electrophoresis Operation of the ISO-DALT (R) System. Large Scale Biology Press, Washington, D.C.
- Berkelman, T. and Stenstedt, T. 1998. 2-D Electrophoresis. Principles and Methods. Amersham Biosciences, San Francisco.
- Bjellqvist, B., Ek, K., Righetti, P.G., Gianazza, E., Görg, A., Westermeier, R., and Postel, W. 1982. Isoelectric focusing in immobilized pH gradients: Principle, methodology and some applications. *J. Biochem. Biophys. Methods* 6:317-339.
- Dunbar, B.S. 1987. Two-Dimensional Electrophoresis and Immunological Techniques. Plenum, New York.
- Görg, A., Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R., and Weiss, W. 2000. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 21:1037-1053.
- Hirano, H. 1989. Microsequence analysis of winged bean seed proteins electroblotted from two-dimensional gel. *J. Prot. Chem.* 8:115-130.
- Lilley, K.S. 2002. Protein profiling using two-dimensional difference gel electrophoresis (2-D DIGE). In *Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, D.W. Speicher, and P.T. Wingfield, eds.) pp. 22.2-22.14. John Wiley & Sons, Hoboken, N.J.
- Ochs, D.C., McConkey, E.H., and Sammons, D.W. 1981. Silver stains for proteins in polyacrylamide gels: A comparison of six methods. *Electrophoresis* 2:304-307.
- O'Farrell, P.H. 1975. High-resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250:4007-4021.
- Westermeier, R., Postel, W., Weser, J., and Görg, A. 1983. High-resolution two-dimensional electrophoresis with isoelectric focusing in immobilized pH gradients. *J. Biochem. Biophys. Methods* 8:321-330.
- Wilkins, M.R., Pasquali, C., Appel, R.D., Ou, K., Golaz, O., Sanchez, J.C., Yan, J.X., Gooley, A.A., Hughes, G., Humphery-Smith, I., Williams, K.L., Hochstrasser, D.F. 1996. From proteins to proteomes: Large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (N.Y.)* 14:61-65.

Key References

Berkelman and Stenstedt, 1998. See above.

Manual for the GE/Amersham family of 2-D electrophoresis equipment; details the background and protocols for use of immobilized pH gradient (IPG) strips in high-resolution analysis.

Celis, J.E. and Bravo, R. (eds.) 1984. Two-Dimensional Gel Electrophoresis of Proteins. Academic Press, San Diego, Calif.

Provides detailed information on two-dimensional gels and their applications.

Dunbar, 1987. See above.

A comprehensive guide to two-dimensional electrophoresis. It covers basic principles of electrophoresis, gives instructions for performing two-dimensional electrophoresis and associated procedures such as protein detection, photography, and preparation of antibodies from proteins excised from two-dimensional gels.

Garfin, D. and Heerdt, L. (eds.) 2000. 2D Electrophoresis for Proteomics. A Methods and Product Manual. Bio-Rad Laboratories, Inc., Hercules, Calif.

Manual for the Bio-Rad family of 2-D electrophoresis equipment; details the background and protocols for use of immobilized pH gradient (IPG) strips in high-resolution analysis.

Görg et al., 2000. See above.

An important update on the latest developments in 2-D electrophoresis.

Contributed by Lonnie D. Adams
The Upjohn Company
Kalamazoo, Michigan

Sean R. Gallagher
UVP, Inc.
Upland, California

Gel electrophoresis has become a ubiquitous method in molecular biology for separating biomolecules. This prominence is the result of several factors, including the robustness, speed, and potential throughput of the technique. The results of this method are traditionally documented using silver-halide-based photography followed by manual interpretation. While this remains an excellent method for qualitative documentation of single-gel results, digital capture offers a number of significant advantages when documentation requires quantitation and sophisticated analysis. Digital images of gel electropherograms can be obtained rapidly using an image-capture device, and the images can be easily manipulated using image analysis software.

REASONS FOR DIGITAL DOCUMENTATION AND ANALYSIS

There are several reasons to consider digital documentation and analysis of electrophoresis results. These justifications can usually be categorized into issues of ease of handling, accuracy, reproducibility, and cost.

Ease of Handling

A major advantage of the digital revolution has been in storage and retrieval of information. Storage in notebooks and filing cabinets previously meant that searching for specific data or experiments was a tedious manual process. With digital information, modern search engines can quickly find specific information in a fraction of the time usually required for a manual search. Making backup copies of nondigital data can be difficult, expensive, and time consuming since it requires copying, retyping, or photographic reproduction. Copies of digital data can be generated more easily and at reduced costs.

Manipulation of information is also easier when it is in a digital format. While the cut-and-paste analogy comes from physical documentation, it takes on a new perspective when applied digitally. Electrophoresis images can be resized, cropped, and inserted into reports. Data can be passed to spreadsheets and statistical packages for analysis and later insertion into notebooks and reports. These reports can be distributed via the Internet to colleagues throughout the world. A single individual can do all this in a few hours.

Digital analysis also provides an easier method for handling the data when comparing large numbers of results or large numbers of separate experiments. Research that requires comparing the banding patterns on 1000 gels containing 50 lanes each can be an undertaking of heroic proportions if the analysis is performed manually. Database software can dramatically speed the analysis and handle the more mundane tasks, leaving the researcher free to interpret the data.

Accuracy

The human eye is an extremely versatile measuring instrument. It can handle light intensities covering a range of nearly nine orders of magnitude and is sensitive to a fairly wide spectrum of light (Russ, 1995). Yet the eye cannot accurately and reproducibly quantitate density and patterns, nor can it deal with large numbers of bands or spots. Accuracy of measurement is a primary reason for using digital analysis on electrophoretically separated proteins and nucleic acids. Two categories of accuracy are key to digital analysis: positional accuracy, which is important for mobility determinations such as molecular weight, and quantitative accuracy.

Positional accuracy is based on both resolution of the recording medium and measuring accuracy. Silver-halide-based recording has a theoretical resolution based on ~2000 imaging elements (silver grains) per inch. Measurement traditionally occurs using a ruler, with an accuracy of ~20 to 40 elements (50 to 100 elements per inch). In comparison, typical digital systems have 200 to 600 picture elements (pixels) per inch. The advantage that digital systems have is in measuring accuracy, which can occur at the level of a single imaging element.

Quantitative accuracy is also an issue. The amount of material represented by a band or spot is difficult to determine accurately from an image of a gel unless it is a digital image. On a digital image, the amount present is directly correlated with the derived volume of the band or spot—the volume is calculated using the intensity values of the pixels within the object.

Reproducibility

Any technique or measurement is only as good as its ability to be faithfully replicated. With software-defined routines, measurements are performed in the same manner every time. Allowing the computer to do repetitive tasks and complicated calculations minimizes the chance for individual errors. This does not imply that such measurements are correct, just that they are reproducible. An incorrect routine or algorithm can also invalidate data.

Cost

A consideration when evaluating any laboratory method is cost. Digital electrophoresis analysis equipment can be expensive. In many cases however, it offers the only method for achieving acceptable analysis performance. In other cases, equal performance can be achieved using silver-halide technology. However, traditional photography can also be expensive when the costs of consumable supplies such as film and developers and other expensive requirements such as developing tanks and dark rooms are included. Often, digital methods can be a good choice when all costs are considered.

KEY TERMS FOR IMAGING

There are several specialized terms encountered during digital image analysis. The most commonly encountered are contrast, brightness, gamma, saturation, resolution, and dynamic range. They describe controls on how the light detectors report a range of light intensities. Below is a brief description of each.

Contrast

Contrast describes the slope of the light intensity response curve. An increase in the contrast increases the slope of the curve. The result is a more detailed display over a narrowed range of intensities with less detail in the remaining portions of the intensity range. This is depicted in Figure 10.5.1A and 10.5.1B, where a normal, unadjusted image and a contrast-adjusted image are displayed, respectively. The contrast was increased on midrange intensity values in Figure 10.5.1B to highlight band intensity differences at the expense of background information. Images with a narrow range of informative intensities can benefit from increasing the contrast since it effectively increases the scale and improves detection of minor differences in intensity. Contrast settings should be lowered if information is being lost outside of the contrast range. For example, in Figure 10.5.1B, loss of background information between peaks indicates that this image should not be used for quantitation.

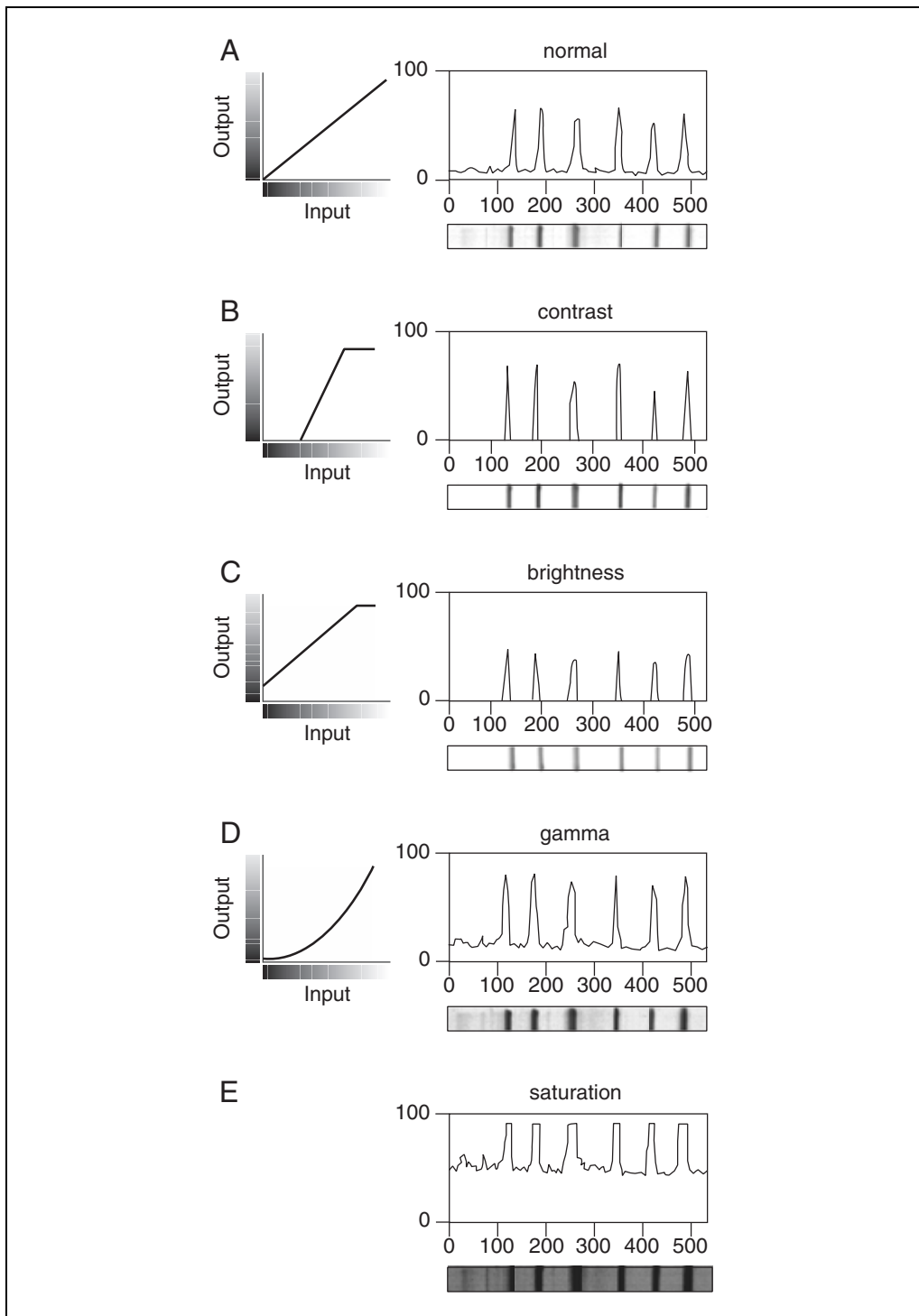


Figure 10.5.1 Examples of how altering image capture settings affects the image and the analysis. The graph on the left displays the light intensity response curve used for image capture while the image and resulting lane profile on the right display how the setting affects the image. The lane profile displays pixel position versus normalized pixel intensity (**A**). In this case, the output has not been altered, giving a straight line with a slope of 1 on the response curve. (**B**) The image acquisition was adjusted to increase the contrast of the displayed image. Although useful for images with a narrow range of informative intensity values, increasing the contrast can lead to a loss of low and high values. (**C**) Decreasing the brightness reduces peak values but also leads to a loss of the weak bands and original background. (**D**) Gamma adjusts raw data to appear more visually accurate. Note that this leads to a loss of fidelity between the adjusted image and the original. (**E**) Saturation indicates that the detector is reporting its maximum value or that the dynamic range for the visualization method has been exceeded.

Brightness

While brightness can have many different definitions, only one will be considered here. Brightness shifts the light intensity response curve without changing its slope as is shown in Figure 10.5.1C. Another name for brightness is black level since it is commonly used to control the number of black picture elements (pixels) in an image. Incorrect brightness levels can lead to either high background and potential image saturation or, as is illustrated in Figure 10.5.1C, to loss of background information entirely and partial loss of band information.

Gamma

Nonlinear corrections are often applied to images to compensate for how the eye perceives changes in intensity, how display devices reproduce images, or both. The most common correction is an exponential one with the exponent in the equation termed the gamma. A typical gamma value is 0.45 to 0.50 for camera-based systems and is illustrated in Figure 10.5.1D. This is a compromise value that compensates for the 2.2 to 2.5 gamma present in most video monitors and the print dynamics of most printers. Since it is a nonlinear correction, special care must be taken if quantitation is desired. Unless otherwise directed by the manufacturer, gamma values other than 1.0 should be avoided when quantitating. More information on gamma correction can be found on Poynton's Gamma FAQ (<http://www.inforamp.net/~poynton/Poynton-color.html>).

Dynamic Range

Dynamic range describes the breadth of intensity values detectable by a system and is usually expressed in logarithmic terms such as orders of magnitude, decades, or optical density (OD) units. A large dynamic range is important when trying to quantitate over a wide range of concentrations. The most accurate quantitation occurs in the linear part of the dynamic range, which is usually not the complete dynamic range of the system. An additional consideration is the dynamic range of the visualization method. Many popular visualization methods have linear dynamic ranges of 1 to 2.5 orders of magnitude. An imaging system with greater dynamic range analyzing the results of such a visualization method will not improve the dynamic range.

Saturation

Saturation occurs when a detector or visualization method receives input levels beyond the maximum end of the dynamic range. This results in a loss of detail and quantitative information from those data points that are saturated. For fluorescent and chemiluminescent samples, reduction in the sampling time can sometimes correct saturation problems. Optical density-based visualization techniques can also generate saturated images, as is illustrated in Figure 10.5.1E; this can sometimes be avoided with longer sampling times or increased detection source intensities. More often, it will be necessary to perform another electrophoresis with more dilute samples or to alter the visualization process to generate a less optically dense material.

Resolution

Resolution is the ability of a system to distinguish between two closely placed or similar objects. Three types of resolution are important for analysis—spatial resolution, intensity resolution, and technique-dependent resolution.

Spatial resolution is the ability to detect two closely placed objects in one-, two-, or three-dimensional space. It is most accurately described as the closest distance two objects can be placed and still be detected as separate objects. In practice, it is often defined

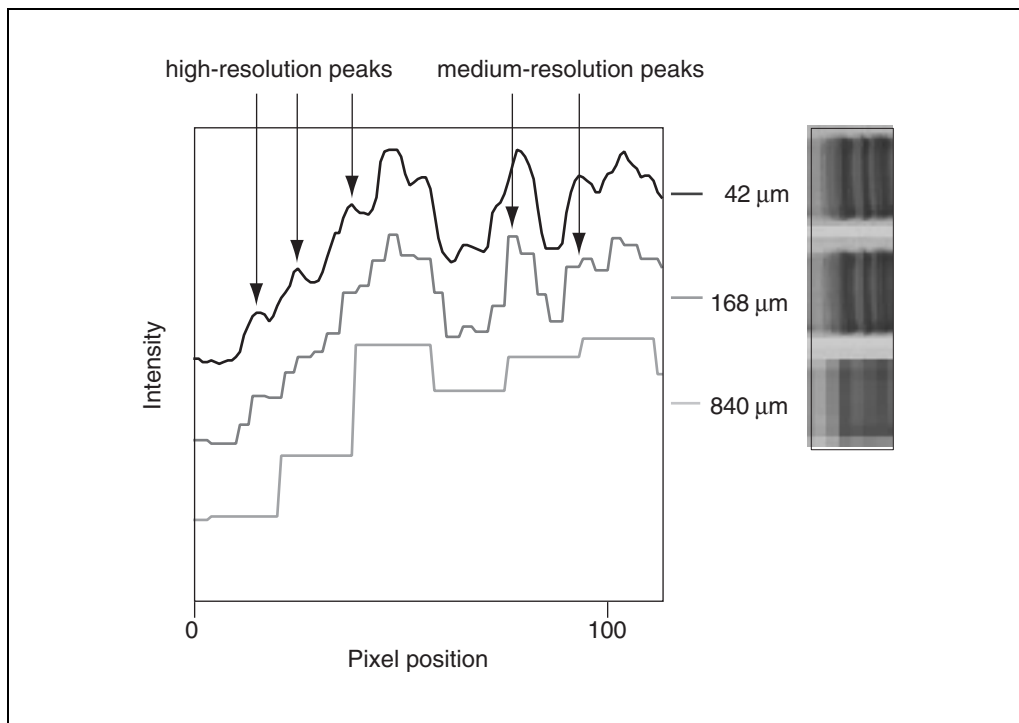


Figure 10.5.2 The effect of spatial resolution on the ability to detect closely spaced objects. Whole-cell protein lysates from *E. coli* were separated using SDS-PAGE and visualized with Coomassie blue staining. An image was captured at 42 μm (600 dpi), 168 μm (150 dpi), and 840 μm (30 dpi) from a segment of the lane, and a lane profile was generated for each image. The lane profiles have offset intensities to allow for comparison. Only major bands can be detected with the low-resolution image (840 μm); at higher resolutions more bands are detectable.

nominally in terms of the number of detectors per unit area such as dots per inch (dpi) or the number of detectors present in total or in each dimension such as 512×512 (262,144 total detectors). Actual resolution is less than half the nominal resolution due to the need for two detectors for every resolvable object (one for the object and one for the separation space) and the effects of optical resolution. Figure 10.5.2 demonstrates how spatial resolution can affect detection of objects. The 42- μm resolution image allows detection of closely spaced bands, the 168- μm resolution image detects fewer bands, and the 840- μm resolution image detects only major bands. For instruments with on-line detection systems, a pseudo spatial resolution is often reported in units of time from the start of the separation or the time interval between two objects crossing the detection path.

Intensity resolution is the ability to identify small changes in intensity. It is a function of both the dynamic range of a detector and the number of potential values that detector can report. Greater dynamic range decreases the intensity resolution of a given detector. The number of potential values a detector reports is described by its bit depth. An 8-bit detector can report 256 (2^8) different possible values, while a 12-bit detector can report 4,096 (2^{12}) values, and a 16-bit detector can report 65,536 (2^{16}) values. The higher the bit depth, the greater the intensity resolution.

Technique-dependent resolution directly affects the spatial and intensity resolution. Electrophoretic separation techniques that generate overlapping objects or that have object separation distances shorter than the spatial resolution will fail to provide reliable data. Many factors, including the amount of sample loaded, gel pore size, buffer constituents, and electrophoresis field strengths, can dramatically affect separation and resolution of biomolecules. Likewise, detection methods that can only generate a small range of discrete intensity values will not benefit from systems with improved intensity resolution.

IMAGE CAPTURE

Systems

Several components make up a digital imaging system (Fig. 10.5.3). These include: (1) a light source for illuminating the sample, either for nonfluorescent white-light imaging or for fluorescent excitation of a sample stained with a fluorescent dye; (2) a filter on the lens to act as a contrast or signal-to-noise (S/N) enhancer for nonfluorescent applications, or an emission wavelength isolation filter for fluorescent detection applications; (3) a fixed or zoom lens; (4) the CCD camera for acquiring the image; and (5) the computer and software for analysis.

Filters used for digital imaging have a number of functions depending on the application (Table 10.5.1). For nonfluorescent imaging, colored glass or wavelength-isolation filters are typically used to enhance the contrast and S/N. For example, a Coomassie blue-stained protein gel shows enhanced contrast over background illumination with use of an orange filter. This is very useful for acquiring faintly stained bands on a gel. With fluorescent applications, the lens filter acts to isolate and detect the emission wavelength from the fluorescent sample or gel. These filters are typically very accurate band-pass interference filters.

Devices

Capturing digital images involves a detection beam or source, a sensor for that beam or source, and some method of assembling a two-dimensional image from the data generated. Most systems use a light source for detection. The light wavelengths used range from ultraviolet (UV) to infrared (IR) and can be broad spectrum or narrow wavelength.

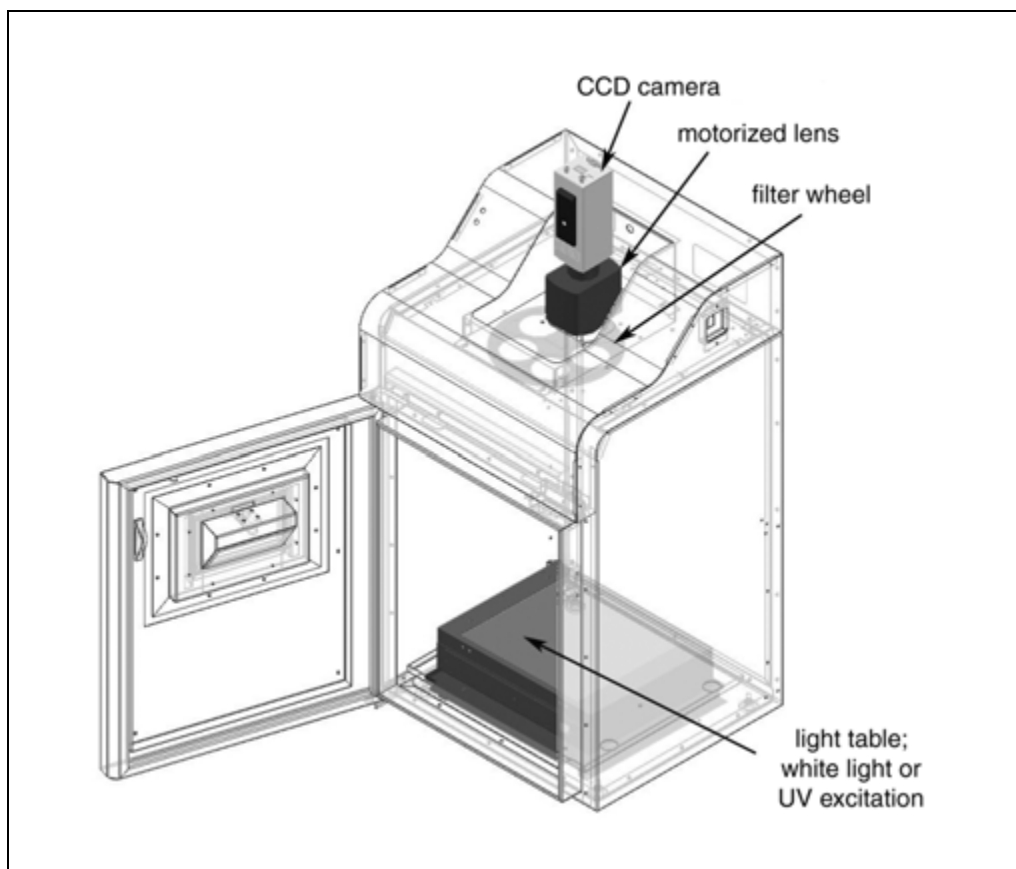


Figure 10.5.3 Typical light-tight desktop system for UV/visible colorimetric, chemiluminescent/bioluminescent, and fluorescent imaging.

Table 10.5.1 Filters for Digital Imaging Applications

Application	Desired effect	Typical filters	Source
Optical protein stains (e.g., Coomassie Blue)	Contrast enhancement	Colored glass Broad band-pass Short-pass Long-pass	Edmund Scientific, Omega Optical, Schott Glass
Fluorescent nucleic acid and protein stains, green fluorescent protein	Signal to noise (S/N) enhancement Excitation and emission wavelength isolation	Interference filters Broad and short band-pass Short-pass Long pass	Edmund Scientific, Omega Optical, Schott Glass

Broad-spectrum detection is more versatile since it can often be used for more than one detection wavelength. A typical UV light table will emit an broad excitation light that peaks at 302 nm, useful for excitation of many fluorescent stains. However, when compared to narrow-wavelength sources such as lasers, broad-spectrum detection suffers from reduced sensitivity and reduced dynamic range. While many types of light sensors have been used, including charge-coupled devices (CCDs), charge-injection devices (CIDs), and photon multiplier tubes (PMTs), technology advances in CCDs have led to their dominance. CCDs are semiconductor imaging devices that convert photons into charge. This charge is then read and converted into a digital format via an analog-to-digital converter (ADC).

The method of image assembly depends on the light source and detector geometry. One method is to capture the image all at once using a two-dimensionally arrayed CCD detector similar to the detectors found in digital and video cameras (see Advances in CCD Technology). Typically a camera-type sensor is paired with a light source that evenly illuminates the sample. This same sensor is often used with fluorescent and chemiluminescent detection methods as its ability to detect light continuously over the entire sample reduces image capture times. Another method of image assembly is to capture the image a line at a time. This typically involves a linearly arrayed CCD scanning slowly across the sample in conjunction with the detection beam of light. The data from each line are then compiled into a composite image. Spatial resolution in this method can be significantly better on large-format samples compared to the resolution of a camera-based system. This method is also advantageous when OD-based detection is used, since the more focused light beam is usually of higher intensity and can penetrate denser material. A third method of image assembly is to use a point light source and single-element detector on each point on a sample. The image is then compiled from each point sampled. This method is slower than the others but can offer extremely high resolution and sensitivity. A fourth commonly encountered method is that of generating a pseudoimage of electrophoresis results through the use of a finish line type of detection system. This is comprised of a light source positioned at the bottom of the gel (i.e., the end opposite of the site of sample loading) and light detectors positioned next to each lane to detect the transmitted light or emitted fluorescence. A lane trace is generated using time on the *x*-axis and light intensity on the *y*-axis. The pseudoimage is then generated from this data (Sutherland et al., 1987).

Capture Process

Prior to image capture, electrophoretic separation and any visualization steps are performed. To calibrate the separation process, standards are usually run at the edges of the gel and often at internal positions. If quantitation of specific proteins or nucleic acids is to be performed, a dilution series of standards with similar properties to the experimental

Table 10.5.2 Compatibility of Popular Image Capture Devices with Common Visualization Methods^a

Visualization method	Image capture device				
	Silver-halide photography	CCD camera	Desktop scanner	Storage phosphor	Fluorescent scanner
Optical density ^b	+	+	++	–	–
Fluorescence	+	+	–	–	++
Chemiluminescence	+	++	–	±	–
Radioactivity	+	–	–	++	–

^aThe device with the highest sensitivity and greatest dynamic range for a visualization method is marked with a ++, other devices that can detect this visualization method are indicated with a +, and devices that are not suitable for a visualization method are indicated with a –. A ± indicates that only some devices of this type can be used with this visualization method.

^bOptical density methods include Coomassie blue staining.

samples should also be included. After separation, the protein or nucleic acid is visualized if necessary. Visualization can include binding of a fluorochrome or chromophore (e.g., Coomassie blue), precipitation of metal ions (e.g., copper, silver, or gold), enzymatic reactions, and exposure of film or phosphor screens to radiant sources. These methods can be grouped based on the type of detection into optical density, fluorescence, chemiluminescence, and radioactivity. The suitability of popular detection devices with these methods is described in Table 10.5.2. Once visualization has occurred, image capture consists of the following steps: previewing the image while adjusting capture parameters, capturing the image, and saving the image for later analysis.

During the preview process, capture parameters are optimized for data content and for ease and rapidity of later processing steps. Typically, the first step is to place the sample so that when the image is captured, the rectangular edges of the gel are horizontal and vertical on the monitor and any lanes are either horizontal or vertical. Since band and spot detection will be much easier if the image is properly oriented, this eliminates the need to later rotate the image digitally. Image rotation is time consuming and can result in spatial linearity errors (a change in the size and shape of objects in image) resulting from rectangular image-capture device geometries. The next step for camera-based systems is to adjust magnification and to focus the sample image. For thicker samples, it might be necessary to reduce the aperture on camera-based systems to get a sufficient depth-of-field to focus the entire sample. Often at this point image imperfections—e.g., dust, liquid, or other foreign objects that will detract from later analysis—are detected, and they need to be removed. Next, image intensity is set. Within the area of interest on the image, band or spot peaks should have values less than the maximum saturated value, and the background should have nonzero values. This is usually accomplished through adjustment of the light-source intensity or the sensor signal integration. If the device allows precapture optimization of other parameters such as spatial resolution, contrast, brightness, gain, or gamma, they are adjusted next. Note that this only applies to controls that affect the response of the sensor or processing of the image prior to a data reduction step and not to controls that affect the image at later stages. The latter process can enhance visualization of specific features but is best left to adjustments in look-up tables (LUTs) in later analysis steps than during image capture since there is a risk of data loss during postacquisition image processing. LUTs are indexed palettes or tables where each index value corresponds to color or gray-scale intensity values present in an image. Many image analysis programs alter LUTs instead of image values directly since it is both faster and does not change the original image data.

Once all the capture parameters are optimized, the image capture process is initiated. This might take less than a second for images captured with camera-based detectors and up to hours for scanning single-point detectors. When the image has been captured, it should be carefully examined for content. It should fully capture the area of interest and the parameters should have been set so that all necessary information is detectable. Furthermore, it should be in a form that will allow for easy analysis. Extra time spent on optimizing the capture parameters will often result in a reduction in total image analysis time and in an increase in data quality. When the best possible image has been captured, it often contains information outside the area of interest. While this is unlikely to cause problems with later analysis, it is often advantageous to crop the image so that the only portion that is saved contains the area of interest. This reduces the amount of disk space necessary to store the image, and the image usually will load and analyze faster with the analysis software.

The last step in image capture is to save the image. Several options are available at this point, including choosing the location at which to save the image, what file type or format to use, and whether to use some form of compression.

The location where the image is saved is not as trivial a question as it might seem if the image will need to be transferred to another computer at some point. File sizes can easily exceed 15 megabytes on high-resolution images. This is a manageable size for hard drives but exceeds current floppy drive sizes by an order of magnitude. There are software utilities available that will subdivide files into disk-size chunks and then reassemble them at the next computer, but this is an inconvenient and slow method. If the computer used to help capture the image is connected to a network, the image files can easily be transferred this way or potentially saved on a central server. Alternatively, several types of high-capacity removable media are available (e.g., Zip, or Jazz). This usually requires the installation of additional hardware onto two or more computers but does make backing up data easier.

Since image files can be very large, compression techniques are sometimes used to reduce disk space requirements. Compression algorithms use several methods, typically by replacing frequent or repetitive values or patterns with smaller reference values and by replacing pixel values with the smaller difference values describing the change in adjacent pixels. When the file is later decompressed, the compressed values are then replaced with the original information. Not all images compress equally, with simple images containing mostly repetitive motifs compressible by $\geq 90\%$, while complex images will benefit much less from compression. Because compression is a much slower method of saving files and not every file will benefit from it, compression is not used to save all files. Several different forms of compression are available but are separable into two main classes, lossless and lossy. Lossless methods faithfully and completely restore the image when it is decompressed (no loss of data) but offer only moderate file compression. Compression values range from $\sim 10\%$ to 90% , depending on the image. Examples of lossless compression include Huffman coding (Huffman, 1952), RLE (Run Length Encoding), and LZW (Lempel, Ziv, and Welch; Welch, 1984). In comparison, lossy methods such as JPEG (Joint Photographic Experts Group), MPEG (Moving Picture Experts Group), or fractal compression schemes can reach compression values of $\geq 98\%$ (Russ, 1995). The trade-off is that not all information from the original file is recovered during decompression. Lossy compression is sometimes necessary for applications with extremely large image files such as real time video capture, but it usually represents an unacceptable loss of data if used with electrophoresis image capture.

Many different file types have been developed to store digital images. Some of these file types are proprietary or are hardware specific. For example, PICT is a Macintosh format

and BMP is a PC-compatible format (see descriptions of both systems, below). Each file type has its own structure. Some types do not allow compression, for others it is optional, and for some it is mandatory. They vary in the types of images they support, particularly in the number of colors or gray levels. Below is a brief description of a few of the more prevalent file types.

TIFF (Tagged-Image File Format) is one of the most commonly used formats. It is particularly versatile since it is an open format that can be modified for specific applications. One reason for its versatility is the ability to attach or tag data to the image. The tags can include information such as optical density calibration, resolution, experimenter, date of capture, and any other data that the application software supports. TIFF images can be monochrome, 4-, 8-, or 16-bit gray scale, or one of many color-image formats. Compression is optional, with LZW, RLE, and JPEG often supported (Russ, 1995). Since TIFF is supported by both Macintosh and PC computers, it is a good choice for multiple-platform environments. The versatility of TIFF can also be a weakness. Since there are many different tagging schemes and since not all programs support all possible compression and color schemes, it is sometimes not possible for one program to access the information in a TIFF file generated by a different program.

GIF (Graphics Interchange Format) is a file format that is widely encountered on the Internet due to its compactness and standardization. Its compactness is attributable to a mandatory modified LZW compression. Another feature of GIF is the use of a LUT to index the values in the image. One interesting ability of GIF is that it supports storing multiple images within a single file. This can offer some advantages for applications such as time-lapse image capture. A GIF image can contain no more than 256 individual color or gray levels and does not support intensity resolutions higher than 8 bit. In addition, since the image is implemented as a LUT, it also is not a true gray-scale image. Due to these limitations and others, alternative formats such as PNG have been developed to replace GIF.

PICT is a file format and graphics metafile language (it contains commands that can be played back to recreate an image) designed for the Apple Macintosh. It can contain both bitmap images and vector-based objects such as polygons and fonts. It supports a ≤ 256 -gray-level LUT, and monochrome images can be RLE compressed. Because it only offers a 256-gray-level LUT, it has the same weaknesses that GIF does with true gray-scale and high-intensity-resolution images. In addition, any vector objects in the image are difficult to translate on a PC since they are designed to be interpreted by Macintosh QuickDraw routines.

BMP is the native bitmap file format present on Windows-based PCs. It supports 2-, 16-, 256-, or 16-million level images. With images of ≤ 256 gray levels, it implements a LUT, while the highest-resolution image is implemented directly. RLE compression is optional for 16- and 256-gray-level images. Since compression is prohibited on 16 million-gray-level images and there is no intermediate level supported beyond 256 levels, BMP is not a good choice for images with high-intensity resolution requirements.

ADVANCES IN CCD TECHNOLOGY

Currently there is a growing emphasis on using CCDs to capture the intensity or spectral data produced in electrophoretic assays. A CCD offers great flexibility in data capture, since it provides both location and intensity information at the same time. Once the data are captured, they must be understood and interpreted. The most difficult component in this scheme is the experimental design necessary to meet the requirements of this type of detector, or the choice of the appropriate detector for a particular experimental design.

Then, it is necessary to know the influence of the detector on the data set. To make good choices in the experimental setup and proper interpretation of the data, a thorough and intimate knowledge of the detector is essential.

CCD detectors operate on a simple principle. When a photon of acceptable wavelength interacts with the CCD substrate (a silicon crystal lattice chip), the energy of the photon generates an electron and a “hole pair.” This effect produces a negative charge in that region of lattice, and the accumulation of these electrons and hole pairs is directly proportional to the number of photons that successfully interact with the silicon lattice. Short-wavelength photons interact near the surface of the lattice, and progressively longer wavelengths interact at progressively deeper depths of the lattice.

A CCD, substrate is electronically segregated into discrete regions called pixels by creating electrical currents and gates at the surface of a silicon wafer. This is the same process used in other semiconductor manufacturing processes to produce memory chips and other microelectronic devices. When appropriate voltages and timing signals are passed through the microcircuits, a CCD chip becomes divided into pixels and is able to convert incoming photons into an electron signal, create a two-dimensional map of this signal, and then transfer this signal to an amplifier. Each pixel in the CCD has an *X* and *Y* value, measured in microns plus a depth in the substrate. These values determine a volume of the pixel and how many electron hole pairs can be generated and segregated in each pixel. This capacity is referred to as the full-well capacity or saturation level of a pixel.

When the CCD chip is used to create a camera, the amplified signal is passed to a monitor or recording device for viewing or documentation. If the camera is designated as a digital camera, the analog signal is passed to an analog-to-digital converter (ADC) before being sent to a computer. A key point here is to remember that all CCDs are analog devices. Cameras may be analog or digital depending on how the camera is designed, but the CCD is always analog.

Why Use CCDs?

CCD technology is currently at the point where the conversion efficiency of photons to electron hole pairs can exceed 90%. This value is called the quantum efficiency of the CCD, and is the key factor in generating the signal in the detector (Fig. 10.5.4). On the opposite side of the signal level is how much noise is generated in the detector and the camera during the signal collection and transfer process. This relationship between the amount of signal collected and the amount of noise generated is called the signal-to-noise ratio, and is one of the key terms necessary to understand the choice of detector and choice of experimental design. In a CCD, this signal-to-noise ratio has reached the point where in most cases the limiting factor of the data accuracy is the signal itself. Few other detectors offer an equivalent signal-to-noise ratio at the data collection speeds of a CCD when a two-dimensional data array is required. CCD imagers range from relatively low cost for simple white light and fluorescent imaging to more sophisticated and costly cooled cameras for chemiluminescent and *in vivo* imaging (Table 10.5.3).

Glossary of CCD Terms

CCD Charge-coupled device, the basis of operation for each detector. A CCD is a silicon-based detector in which an electronic charge (an electron and a hole pair) is produced in the silicon lattice by a photon.

Pixel An electronically segregated area in the silicon lattice produced by microelectronic circuits on the surface of the silicon substrate, or chip.

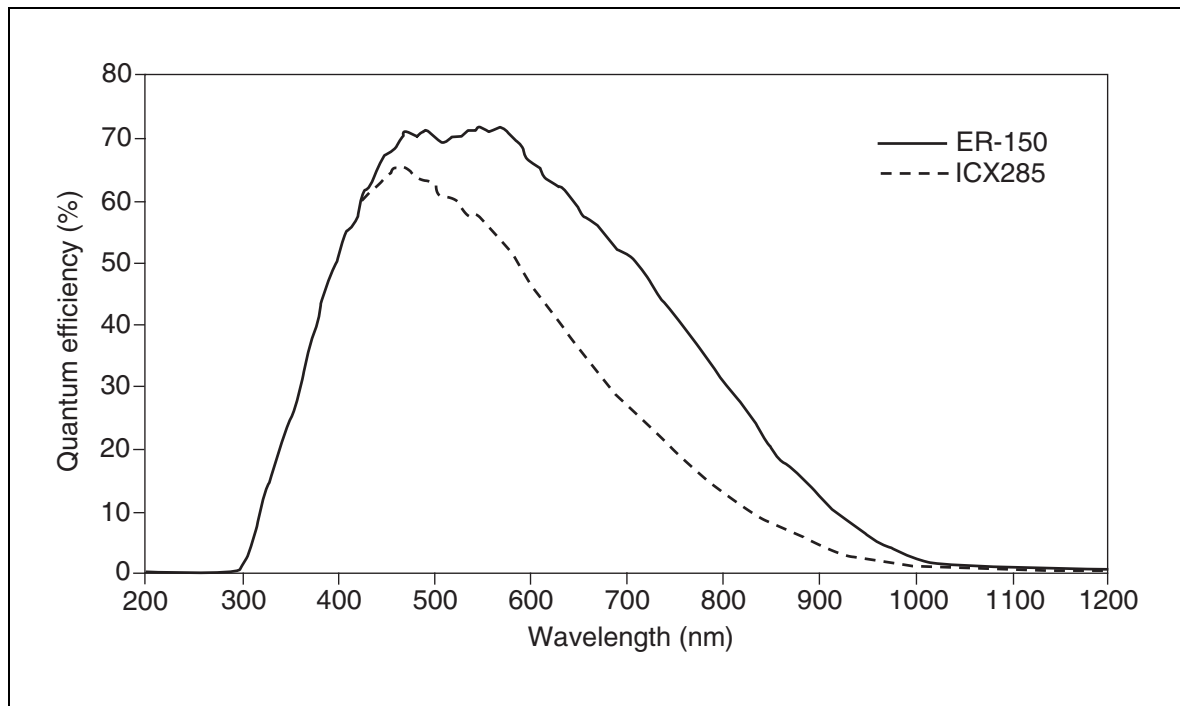


Figure 10.5.4 Quantum efficiency of an extended range (ER-150; solid line) and standard (ICX285; broken line) CCD.

Photon The minimal unit of light, an electromagnetic energy with a particular wavelength (frequency) that interacts with the silicon lattice in a CCD to produce an electron and a hole pair.

Full-well capacity The maximum number of electron hole pairs that can be contained in a pixel. Also defined as the upper limit of detection, the point at which additional signals can no longer be detected, and the saturation level of a detector.

Noise A statistical description of the fluctuation in an otherwise stable current. Generally, noise is expressed as an RMS (root mean square) value. Noise terms are added or subtracted in quadrature; it is necessary to square each noise term, add or subtract the terms, then take the square root of the result.

Shot noise See Signal noise

Signal noise The fluctuation in the number of photons in a given amount of time from a stabilized, uniform light source. Under these conditions, the signal noise should be equal to the square root of the signal level, expressed in electrons. Signal noise is not a property of a CCD or a camera, only the signal itself.

Readout noise The fluctuation in a nominal zero level signal as it is transferred through the serial register and readout amplifier of a CCD. This noise is related to the speed of the readout (pixel clock speed) and the related circuit design. In modern cameras this is almost always the factor that limits the low-light detection capability.

Dark current The number of electrons that appear in a pixel when no light is falling on the detector. Dark current is caused by migration of electrons from areas outside the sensing area of the CCD into the sensitive area and is reduced by 50% with every 8°C reduction of the CCD temperature. Dark current is expressed as the number of electrons (e)/pixel (p)/sec (s). In modern cameras, this is generally so small that it can be disregarded at short exposure times.

Example: Specification for Dark Current = $25e/p/s = 25$ electrons per pixel per sec.

Table 10.5.3 CCD Camera Specifications

Price range	Low-end	Mid-range	High-end	Premium
Model/manufacture	C9260-001, Hamamatsu	C8484-51-03G, Hamamatsu	ORCA AG, Hamamatsu	ORCA 2ER, Hamamatsu
Pixel array/size	1344 × 1024/6.45 × 6.45 μm	1344 × 1024/6.45 × 6.45 μm	1344 × 1024/6.45 × 6.45 μm	1344 × 1024/6.45 × 6.45 μm
CCD size	2/3 in.	2/3 in.	2/3 in.	2/3 in.
Readout noise / frame rate	20 electrons/4 Hz	6 electrons/8.9 Hz	6 electrons/8.9 Hz	6 electrons at 5.6 Hz; 3 electrons at 0.8 Hz
Dark current (electrons/pixel/sec) at indicated temperature	8 electrons/pixel/sec at 20°C	0.1 electrons/pixel/ sec at -10°C	0.03 electrons/pixel/sec at 30°C	0.0045 electrons/pixel/sec at -60°C
Camera noise	21.5 electrons	6 electrons	6 electrons	6 electrons/3 electrons
Full-well capacity (1×1)/(2×2)	18,000/23,000 electrons	15,000 electrons	18,000 electrons	18,000 electrons/40,500 electrons
Dynamic range (1×1)/(2×2)	832:1/1070:1	2500:01:00	3000:01:00	3000:1/13,500;1
Bit depth/gray levels	12 bit/4096	12 bit/4096	12 bit/4096	12 bit/4096 and 14 bit/16,384
Digitizer gain (1×1)/(2×2)	4.3 electrons/5.6 electrons	3.7 electrons	4.4 electrons	4.4 electrons and 2.5 electrons
Maximum quantum efficiency (QE) at indicated wavelength	50% at 525 nm	72% at 550 to 570 nm	72% at 550 to 570 nm	72% at 550 to 570 nm
Binning	1 × 1, 2 × 2, 4 × 4	1 × 1, 2 × 2, 4 × 4, 8 × 8	1 × 1, 2 × 2, 4 × 4, 8 × 8	1 × 1, 2 × 2, 4 × 4, 8 × 8

Dark noise The amount of fluctuation in the dark current. This should equal the square root of the dark current.

Example 1: dark current (D) = 25 $e/p/s$; exposure time 1 sec:

$$\text{dark noise } (N_D) = \sqrt{25e} = 5e$$

Example 2: dark current (D) = 25 $e/p/s$; exposure time 0.5 sec:

$$\text{dark noise } (N_D) = \sqrt{\frac{25e}{0.5}} = \sqrt{12.5e} = 3.53e$$

Camera noise The sum of the readout noise and the dark noise. This term refers only to a camera, not to an image.

Example: camera with read noise = 10 e and dark noise = 5 e :

$$\text{camera noise} = C_N = \sqrt{(N_R)^2 + (N_D)^2} = \sqrt{(10e)^2 + (5e)^2} = \sqrt{100e + 25e} = 11.18e$$

Readout noise The fluctuation in a nominal zero level signal as it is transferred through the serial register and readout amplifier of a CCD. Often referred to as simply “read noise,” the noise is related to the speed of the readout (pixel clock speed) and the related circuit design. In modern cameras, this is almost always the factor that limits the low-light-detection capability.

Total noise The sum of the camera noise and the signal noise. This term only applies to an image.

$$\text{total noise} = \sqrt{(N_R)^2 + (N_D)^2 + (N_S)^2}$$

where N_R is read noise, N_D is dark noise and N_S is signal noise.

Example: 1-sec exposure, camera with $10e$ read noise, 5 e/p/s dark current, average gray value in one region of the image equals 3800, digitizer gain equals 4 electrons per count. To find the camera noise, the signal noise, and the total noise in the image:

1. read noise (N_R) = $10e$ (from camera data sheet or measured);
2. dark noise (N_D)

$$= \sqrt{D_s} = \sqrt{25e} = 5e$$

(from camera data sheet or measured);

3. camera noise (N_C)

$$= \sqrt{(R_N)^2 + (D_N)^2} = \sqrt{(10e)^2 + (5e)^2} = \sqrt{100e + 25e} = 11.18e$$

4. signal noise (N_S)

$$= \sqrt{S} = \sqrt{3800 \times 4e} = \sqrt{15,200e} = 123.2e$$

5. total noise (N_T)

$$= \sqrt{(N_R)^2 + (N_D)^2 + (N_S)^2} = \sqrt{(10e)^2 + (5e)^2 + (123e)^2} = \sqrt{15,325e} = 123.7e$$

Please note the difference the camera noise made in this image, only 0.5 electrons!

Dynamic range A calculation of the biggest difference in brightness detectable by a device. To obtain this value, divide the Full-Well Capacity (FWC) of a detector by the camera noise to get this range as a ratio.

Example: Camera with 20,000 electrons FWC and 10 electrons camera noise:

$$20,000e/10e = 2000:1 \text{ dynamic range.}$$

For dynamic range in decibels (dB), take the log of $2000 = 3.3 \times 20$ to get 66 dB.

Bit depth The number of gray levels that the digitizer will create from the maximum signal level of a device. Bits refers to the exponent of 2 used to create the number of gray levels.

Examples:

10-bit = $2^{10} = 1024$ gray levels,

12-bit = $2^{12} = 4096$ gray levels.

A 10-bit digitizer will divide the maximum signal (full-well capacity of a CCD) by 1024 gray levels to produce 1024 gray levels in the image. Bit depth is determined by the digitizer in a digital camera; not the CCD. See example under “digitizer gain,” below.

Digitizer gain/digitizer count The number of electrons in the signal used to represent each gray level in an image. This is an essential number for making quantitative calculations and comparisons of digital camera data. A close approximation may be determined by dividing the full-well capacity by the number of gray levels in the digitizer. Camera data sheets should include the exact number.

Example: Camera with FWC = 20,000 electrons; what is the digitizer gain?

10-bit = $2^{10} = 1024$ gray levels = 20,000 (electrons)/1024 (gray levels) = $19.5e$

12-bit = $2^{12} = 4096$ gray levels = 20,000 (electrons)/4096 (gray levels) = $4.9e$

14-bit = $2^{14} = 16384$ gray levels = 20,000 (electrons)/16384 (gray levels) = $1.2e$

Always compare data and results by number of electrons in the signal, not by gray level, since different cameras have different digitizer gains. In many cases, larger gray values may actually represent smaller data values when making comparisons or calculating the signal noise in an image.

Digitizer offset A value established in the analog-to-digital converter that represents the average zero value in an analog signal. Since a digitizer only has positive numbers but an analog signal will fluctuate both above and below an average value, a digitizer must have some positive values that can represent the negative portion of the analog signal to prevent clipping of the full data set. This range of positive values in the digitizer is called the digitizer offset. A digitizer with an offset of 200 means that the values from 1 to 199 are used to represent values less than zero in the analog signal. Gray level 200 represents the zero level signal and values above 200 represent signal levels above zero.

List of Typical Camera Features and Benefits

CCD camera features and their advantages and disadvantages that should be considered are listed below.

Number of pixels

A CCD camera has more spatial resolution when compared to another camera with the same size pixel, but fewer pixels. This means there is less sensitivity when compared to another camera with fewer, but larger, pixels. In addition, a camera has a potentially slower frame rate when compared to another camera with fewer pixels.

Size of pixels

Smaller pixels provide higher spatial resolution. Cameras with larger pixels provide larger full-well capacity, bigger dynamic range, and better accuracy, and collect signal faster than small pixels.

Size of detector

Larger detector dimensions mean a larger field of view, but require more expensive optics.

Cooling temperature

Temperature alone is not meaningful. The dark current at that temperature must be determined. Cooling reduces the dark current. Cooling temperature may not be relevant for fast exposures

Readout noise specification

Lower values improve low-light sensitivity, but slower readout speeds may be required to achieve lower values. Always be sure the readout noise specification is valid at the desired frame rate.

Quantum efficiency (QE) value

Higher QE is better if it is in the same wavelength range as the signal, but is a problem if it is in a range other than the signal—i.e., near-infrared! QE value must be considered along with other factors, such as camera noise.

Dark current

Lower dark current is better for dim signals and for dynamic range calculations. The cost of necessary cooling may not be worthwhile for very short exposures.

Dark noise

Dark noise is the square root of the dark current. It is important to use the dark noise to calculate camera noise, not dark current.

Major Considerations in Camera Selection

Detectability

Detectability is determined by the camera noise and QE as illustrated in the Figure 10.5.3. Once the camera noise is minimized, the largest gain in this area is obtained through greater QE.

Accuracy of a single measurement

Accuracy is determined by the noise of the signal itself in most applications. Since this noise is equal to the square root of the signal level, several points are worth noting. The signal level must always be calculated in electrons, using the digitizer gain value to multiply the gray levels in the image. The more signal that is collected, the better. Since the square root of a larger signal is a larger number, this concept is not always obvious. The importance is found in the signal-to-noise calculation (S/N). As the signal becomes bigger, the noise becomes a smaller percentage of the signal. This percentage is a good approximation of the limit of accuracy.

Examples:

$$\sqrt{100e} = 10e = 10\% \text{ accuracy}$$

$$\sqrt{500e} = 22e = 4\% \text{ accuracy}$$

The accuracy of data in different areas of an image is dependent upon the signal level in each area. The full-well capacity is the limit to accuracy for the CCD. In absorption studies (measuring dark regions in a bright background), the need for accuracy means that devices with extremely large full-well capacities are needed. The dark regions need to have high signal and the background areas must still not be saturated.

It is not uncommon for a typical CCD camera to have 20,000 electrons full-well capacity and a 12-bit digitizer. Under these conditions, each of the 4096 gray values will represent

~5 electrons of signal. With a digitizer offset of 200 gray levels, any signal from a sample will be superimposed on this 200 level.

In this example, we will assume that some region of the image has an average intensity of 500 gray levels and an adjacent region has an intensity of 600 gray levels. We want to find out how much confidence we can place in each of these measurements and how much different they are.

We first must subtract the digitizer offset from the indicated signal which leaves us with gray levels of 300 and 400 respectively.

Then we multiply each data set by 5 electrons to get the actual signal values: $1500e$ in one region and $2000e$ in the adjacent region.

Now we take the square root of each to get $39e$ and $45e$. These values represent the total noise in each region of the image. This is the amount of fluctuation we can expect in each signal. We can now make these numbers a percentage of the signal (2.6% and 2.3%) or convert them back to gray levels (8 gray levels and 9 gray levels) to establish the range of fluctuation, or accuracy, that we can expect from this data. If the expectation is accuracy of ~2%, we have enough data to make this statement. If the need for accuracy is 1%, we will have to find some way to improve our data set. We would need about 10,000 electrons in one image to get close to 1% accuracy.

Improving accuracy

Accuracy can be improved in two ways. First, one should collect more signals by increasing the QE, exposing longer, or increasing the signal intensity. These are the most common choices, but are not always possible. Second, measurements should be averaged. If it is possible to repeat the data collection process and mathematically average more than one image of the same data set, then the accuracy can be improved by the square root of the number of frames averaged. Even averaging two frames will increase the accuracy by 1.4 times.

ANALYSIS

Once the image has been captured, the data must be analyzed and distilled into information about the results of the electrophoresis experiment. Through the use of standards and experimenter input, this software-driven process can estimate mass and quantity of objects in an image and detect relationships between objects within one image and between similar images. The type of software used depends on the analysis to be performed. Images from single electrophoretic separations are examined by one-dimensional analysis software optimized for lane-based band detection. Images from two-dimensional electrophoresis are best handled by specific programs designed to detect spots and to assign two mobility values and a quantity value to the spot. After the initial characterization of bands and spots, comparisons are often made between bands or spots from different experiments through the use of database programs and matching algorithms.

Software for One-Dimensional Analysis

Table 10.5.4 lists popular sources of electrophoresis imaging software.

Lane positioning

For one-dimensional analysis, the first activity is to detect the lanes on the image. One of three different methods is commonly employed for this. For images with straight, well-defined lanes with a large number of bands, automatic lane-detection algorithms can quickly and accurately place the lanes. On images with very well-defined lanes, such as

Table 10.5.4 Popular Sources of One- and Two-Dimensional Electrophoresis Analysis Software

Company	Web site	Software
<i>1-D gels</i>		
UVP, Inc.	http://www.uvp.com	LabWorks, DocIt
Nonlinear Dynamics	http://www.nonlinear.com	Phoretix series
MediaCybernetics	http://www.mediacy.com	Gel-Pro, ImagePro
Scanalytics, Inc.	http://www.scanalytics.com	1Dscan EX
NIH (free, with gel tutorial)	http://rsb.info.nih.gov/nih-image/	NIH Image
<i>2-D gels</i>		
Genebio	http://www.genebio.com	Melanie
Nonlinear Dynamics	http://www.nonlinear.com	Phoretix series
Amersham	http://www.amersham.com	Melanie
Scanalytics, Inc.	http://www.scanalytics.com	Gellab II+
Bio-Rad	http://www.biorad.com	PDQuest
Compugen	http://www.2Dgels.com	Z3, Z4000

pseudoimages from finish-line type electrophoresis equipment, automated lane calling based on image position is possible. For images with “smiling,” bent, or irregular lanes, manual positioning of the lanes is often the fastest and most accurate method of lane definition. Regardless of the method of identifying the lanes, the lane boundaries need to be carefully set for accurate quantitation and mass determinations. Lane widths should be wide enough so that the entire area of all bands in that lane are included, but they should not be so wide as to include bands from adjacent lanes. To accomplish this, curved or bent lanes might need to be used in order to follow the electrophoresis lane pattern. Lane length and position also must be adjusted as necessary so that all bands of interest are included. If mass determinations are necessary, the sample loading point should probably also be included in the lane or be the start of the lane. At this point, lines of equal mobility (often called R_f or iso-molecular-weight lines) are added to the image as necessary. These lines allow for correction of lane-to-lane deviations in the mobility of reference bands and generate more accurate measurements of mass. A similar form of correction is also possible for within-lane correction of mobilities. This correction is important for accurate detection and quantitation of closely spaced bands.

Band detection

Once the lanes have been defined, the bands present in each lane need to be detected. There are many methods for detecting bands. One method is to systematically scan the lane profile from one end to the other, identifying regions of local maxima as bands. Another common method is to use first and second order derivatives of the lane image or lane profile in order to find inflection points in the change of slope in pixel intensity values (Patton, 1995). Regardless of the method used, it is often necessary to alter the search parameters so that they perform reliably under a given experimental condition. Typical search parameters include ones for detection sensitivity, smoothing, minimum interband gaps, and minimum or maximum band peak size. Smoothing reduces the number of bands detected due to noise in the image. A minimum interband gap is often used to avoid detection of false secondary bands on the shoulders of primary bands. Limitations on peak sizes, especially for within-lane comparison to the largest band’s peak, can be a useful way to allow sensitive detection of bands in underloaded lanes without detecting false bands in overloaded lanes.

Band edges are often detected in addition to band peaks in order to further define bands or to quantitate band amounts. This can be accomplished by using local minima, derivatives of the lane profile, or fixed parameters such as image distances or a percentage of band peak height. The band edges can be applied as edges perpendicular to the long axis of the lane or as a contour of equal intensity circling the band. The perpendicular method is advantageous for bands with uneven distribution of material across the face of the band, while the latter method is better for “smiling” or misshapen bands.

Background subtraction

With nearly all electrophoresis procedures, the most informative images have a low level of signal intensity at each pixel that does not result from protein or nucleic acid. Instead, this background intensity is attributable to the gel medium, the visualization method, electronic noise, and other factors. Since this background tends to be nonuniformly distributed throughout the image, failure to subtract it can make band detection and quantitation less accurate. Many methods of background subtraction are possible. Sometimes it is possible to generate a second image under conditions that do not detect the protein or nucleic acid. The second image is then digitally subtracted from the data-containing image to remove the background. More often, background information must be obtained from a single image. If the background varies uniformly across the image, a line that crosses the variation can be defined at a point where no bands are present. The intensity values at each point on the line can be used as the background value for the pixels perpendicular to the line at that point. Commonly, background is also present as variations in intensity along the long axis within each lane. One simple method is to take the lowest point in the lane profile as the background. Another method is to use an average value of the edge of each band as the background for that band. More complicated methods such as valley-to-valley and rolling-disk use local minima points in the lane profile to define a variable background along the length of the lane. Because there can be many different causes and distributions of background, no single method of background determination can be recommended for all experiments.

Characterization

Once lanes and bands have been detected, it is possible to interpret the mobility of the nucleic acid or protein bands. Depending on the method of electrophoretic separation, information on mass (length or size), pI, or relative mobility (R_f) can be inferred from mobility information. The mobility is characterized using a standard curve with internal standards of known properties. The type of curve depends on several parameters. By definition, with R_f -based separation, a linear first-order curve is used since it represents the linear relationship between mobility and R_f . Similarly, pI and mobility are generally linear in isoelectric focusing separations. For separations based on size, a curve generated from mobility versus the log of the molecular weight provides a relatively good fit as measured by the correlation coefficient (R^2). Several other curves have been suggested for size-based separations, including modified hyperbolic curves and curves of mobility versus (molecular weight)^{2/3} that have good correlation coefficients (Plikaytis et al., 1986). In some cases, no single curve equation can adequately represent the data, and methods of fitting smooth contiguous curves using only neighboring points such as a Lagrange or spline fit (described in Hamming, 1973) are necessary. This is most common for size separations with a very large range of separation sizes and with nonlinear gradient gels. Care must be taken with multiple-curve techniques since they rely on only a few data points for any one part of the composite curve, and outlying data points can drastically affect the outcome.

For size and R_f determination, a uniform position must be found in each lane as a point from which to measure the mobility of each band. Many software analysis packages use

the end of the lane as the measuring start point, so for them it is important to position each lane start point at an iso-molecular-weight or iso- R_f point. A convenient point is the well or sample-loading position since it is usually easily detectable and at an equal mobility position in each lane. A consistent point on each band must also be chosen to measure mobility. A band's peak is easily defined in digital image analysis and is commonly used. Since peak positions are harder to detect visually than edges on silver-halide images, the leading band edge is sometimes used when comparing digital results with silver-halide-based results.

Once lanes and bands have been detected, it is also possible to quantify the amount or at least relative amount of nucleic acid or protein present in each band. The amount in a band is related to the sum total of the intensity values of each pixel subtracted by the background value for each pixel in a band. For absorptively detected bands, intensity values are converted to OD values. The total value that is calculated is equivalent to the volume of the band and can be directly compared to other bands that are within the linear range for the visualization method. If standards of known amounts are loaded onto the same gel, they can be used to generate a standard curve that converts band volume into standard units such as micrograms. For greatest accuracy, it is important to be able to generate multiple standard curves when using visualization methods, such as Coomassie blue staining, that are affected by band or spot composition.

Quantitation becomes more complicated when bands are not fully resolved. In this case, material from one band is contributing to the volume measurement of an adjacent band and vice versa. The simplest method for handling this is to partition into each band only the volume within its edges. Alternatively, a Gaussian curve can be fitted to each band and the volume contained within the curve used to estimate the amount of the band. Since most electrophoresis bands have a pronounced skew towards the leading edge of the band, modified Gaussian curves have also been used (Smith and Thomas, 1990). In either case, the curve-fitting process is calculation intensive and can significantly increase analysis times for images with many bands.

Software for Two-Dimensional Analysis

In two-dimensional analysis, the first-dimension separation is performed in a single column or lane followed by a second separation performed perpendicular to the first. The result after visualization is a rectangular image of up to 10,000 spots. The most common two-dimensional gel type is one in which protein is separated first by apparent pI and second by molecular weight, although two-dimensional separation of nucleic acids is also possible. While many of the concepts and analysis techniques used with one-dimensional gels are applicable to two-dimensional gels, the complex nature of most two-dimensional gels requires somewhat different methodology. For example, spots are more difficult to detect since they are not conveniently arranged in lanes and can vary more in shape and overlap than bands. In addition, two-dimensional experiments usually require some method of comparing between two images whereas one-dimensional images usually contain all of the information from an experiment.

See Table 10.5.4 for sources of two-dimensional electrophoresis imaging software.

Spot detection

Probably the most difficult aspect of two-dimensional analysis is efficient and accurate spot detection. If it is incorrectly done, it can lead to hours of manual editing. Due to the complexity and computational intensity of some algorithms, the detection process itself can last hours on relatively fast desktop computers. One theoretically effective but computationally intense method is to treat the image as essentially a three-dimensional

image with spots treated as hills and background as valleys. A large number of Gaussian curves are then combined to describe the topology of the image. Many other methods make use of a digital-imaging technique known as filtering. In essence, filtering is a way to weight the value of a pixel and its neighbors in order to generate a new value for a pixel. By passing a filter across an image pixel by pixel, a secondary image is generated. Filters can be designed for many tasks, including sharpening an image or removing high-frequency noise. Filters can also be generated to help detect spots by making images that are first and second derivatives of the original image. The derivative images indicate inflection points in the intensity pattern and can be used to detect spot centers and edges. In a different method, called thresholding, filters can be used to detect the edges of objects. Instead of looking for inflection points, threshold filters identify intensities above a set level or at ratios between central and edge pixels above a set value. Since the edges on two-dimensional spots tend to be diffuse, sharpening filters are sometimes used prior to the thresholding filter. In some cases, multiple techniques are used to detect spots (Glasbey and Horgan, 1994).

Unlike one-dimensional detection, detection on images of two-dimensional experiments usually requires secondary processing to get acceptable performance. One example of a secondary process is to discard spots with sizes below a set minimum or above a set maximum. Another is to analyze spots that are oval for possible splitting into two spots. Even after secondary processing, it is likely that a small amount of manual editing will be necessary. When manually editing an image, care must be taken to use as objective criteria as possible, especially if two or more images are to be matched and spot volumes compared.

Characterization

In a two-dimensional system, determination of protein or nucleic acid mobility is complicated by the fact that there are two mobilities to account for and that the second-dimension separation tends to make estimation of the separation which occurred in the first dimension more difficult. One method for dealing with this is to have a series of markers in the sample that, after both separations are completed, are evenly distributed within the gel and image. It is also possible to estimate separation characteristics from calibration points located at the periphery of the gel. For example, distance measurements can be used to pass calibration data from the first dimension separation, and standards can be separated at the ends of the gel to calibrate for mobility in the second dimension. Regardless of the method used, in many instances, a series of related images will be examined and similar spots in each image will be matched. When this occurs, it is possible to calibrate one image and then pass the calibration information via the matches to the related images.

Quantity determination is similar in many regards to that which occurs in one-dimensional analysis, but there are some differences. If spot edges are detected, a simple method of determining spot volume is to take the sum of the intensity value of each pixel in the spot reduced by a background intensity value. Multiple Gaussian curves can also be fitted to the spot to approximate the volume (Garrels, 1989). More difficult is attempting to compensate for a skewed distribution in a size-separating dimension while trying to use a regular Gaussian fit for a pI separation, such as is encountered with the most common form of two-dimensional protein separations. The distribution of background makes quantitation more difficult in two-dimensional gels. There is no lane-dependent component, so it is necessary to use other methods such as image stripes, finding local minima values or using values derived from the spot edges to determine background values.

Matching

Matching is the process in which proteins or nucleic acids with similar separation properties are linked or clustered together. Matching can occur within one image or between multiple images as long as a frame of reference is established. Matching allows for comparisons between samples. It also makes annotation and data entry easier, since if one spot or band is matched to others and is characterized or annotated, this information is easily passed to all the other matches. An underlying assumption of matching is that objects with similar separation properties are actually similar. Care must be taken to confirm the identity of matched spots or bands by other methods on critical experiments.

A simple form of matching is to link bands or spots at similar positions on the gel images. This works well when separation and imaging conditions are uniform. This is very seldom the case, since slight differences in the electrophoresis, visualization, and imaging conditions across a gel and between gels generates incorrect matching with this method. Since bands on one-dimensional gels are relatively easy to calibrate for mobility, matching can occur along contours of equal mobility. This dramatically decreases but does not eliminate the variability in detecting similar bands. Much of the remaining variability can be attributed to calibration errors. This error can often be compensated for by allowing a small tolerance in mobility values in determining whether a band is matched or not. Because of the difficulties in calibrating mobility in two-dimensional gels, it is often more practical to use matched spots for calibrating mobility than vice versa. Spot matching between two two-dimensional images starts by finding a small number of landmark spots that are used as seeds for subsequent matches (Appel et al., 1991; Monardo et al., 1994). There are many methods for finding the landmarks in both images, including finding the highest intensity spots, spots in unique clusters, and manual positioning. The most common procedure from this point is to derive a vector that describes the direction and extent of the path from one matched spot to the other when the two images are superimposed. The vector is used as the basis for finding more matches near the landmark matches. To allow for error, the area within a small radius is searched extending from the end of the vector. Once another match is found, its vector is computed and used as the starting point for finding neighboring matches. From this progression, the entire gel is matched. If all vectors are displayed graphically when matching is complete, questionable matches can often be identified as vectors that are significantly different from neighboring vectors.

One specialized use of matching is as an estimator of the similarity and potential genetic relatedness of organisms. For example, on a one-dimensional gel image, a ratio of the matched to unmatched bands for each pairwise combination of lanes can be calculated. This ratio can be used as an indicator of similarity, with values near 1 indicating a pair of highly similar lanes, and values near zero indicating very dissimilar lanes. Assuming that the contents of the lanes are valid samples of the originating organism's genetic makeup, the information on lane similarity can be converted to estimates of genetic similarity. A convenient way to display this similarity data graphically is to generate a dendrogram with similar objects close to each other and less similar objects more distantly placed. An example of such a dendrogram is presented in Figure 10.5.5, where samples from *Listeria* isolates are arranged based on banding pattern.

Databases

In many cases, image analysis is not the last step in the process. The image and analysis data need to be archived in a searchable format. There may be a need to analyze the data from multiple experiments conducted at different sites or in laboratories around the world. Bioinformatic links to diverse data sources might be desired to help develop a unified understanding of the biology behind particular phenomena. When these situations arise, database programs can be utilized to store, link, and search image analysis results.

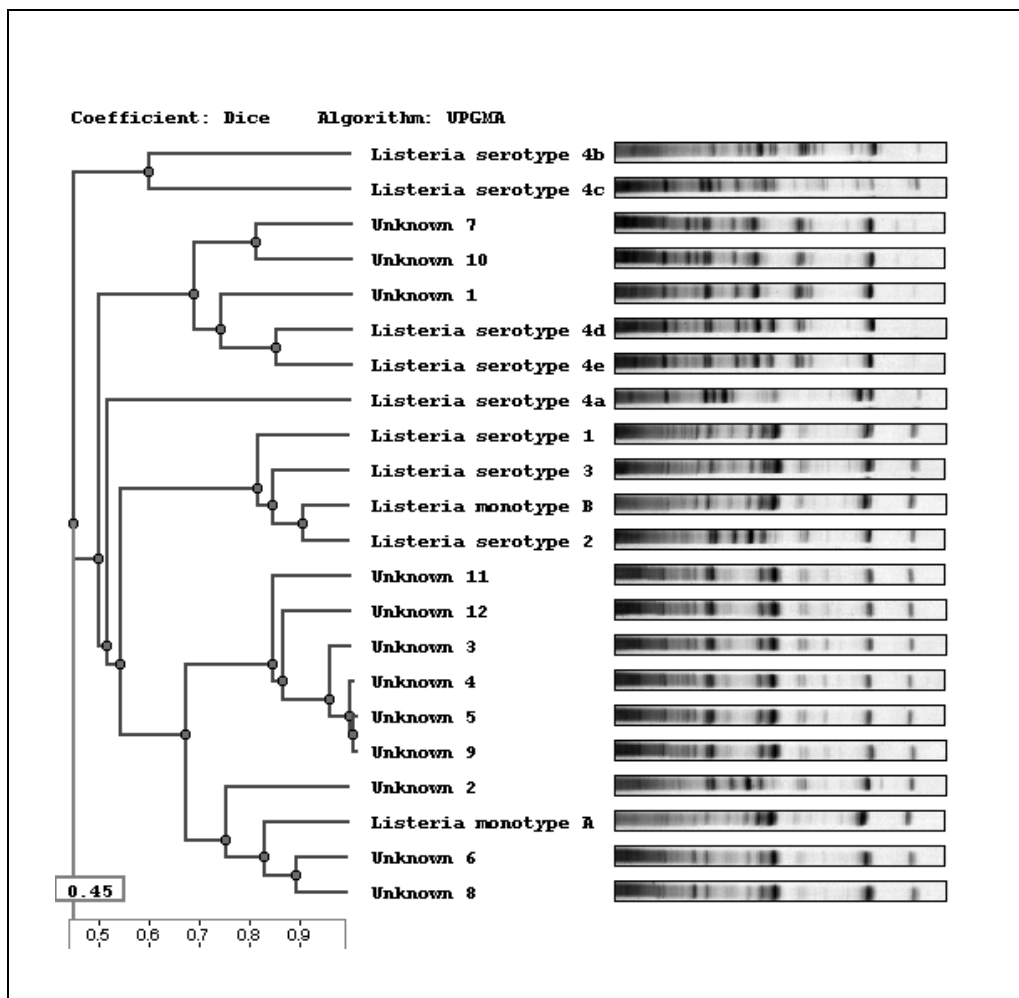


Figure 10.5.5 Example of a dendrogram generated from similarity data on band matching between lanes. DNA samples from 22 isolates of *Listeria* were subjected to Random Amplification of Polymorphic DNA (RAPD) analysis and the resulting electrophoresis image was analyzed with ImageMaster software (Amersham Pharmacia Biotech). Clustering was performed using the Dice coefficient with a tree structure based on the Unweighted Pair-Group Method using Arithmetic Averages (UPGMA). Similarity values between isolates can be determined by locating the node that connects the isolates and reading the value from the scale on the lower left edge of the dendrogram.

As the number of images that are captured and analyzed grows, it becomes increasingly more difficult to find particular information from the large number of files that are stored. Relatively simple databases can be used if the major requirement is to find previously analyzed images and associated data. Such databases often display a miniaturized version of each image to aid in visual scanning for the file as well as simple searching for image-specific information such as date of analysis, file name, or other information that was entered at the time of image capture. More powerful database products are also available that can perform complex searches on data generated during the analysis. For an example, a search on a two-dimensional database might include finding proteins exhibiting a specific expression profile and having a molecular weight >20 kDa with a pI between 3 and 5 or 8 and 10 with an amount <50 ng in a series of experiments conducted <1 year ago. Such searches can quickly target potentially interesting molecules for further analysis.

With the increasing ease of transferring data through the Internet as well as local- and wide-area networks, it has become practical to quickly find and examine data from distant

locations. Of course, great care must be taken to ensure that similar experimental conditions are employed, otherwise the results will be difficult to compare. In this manner it is sometimes possible to dramatically increase the sample size and statistical accuracy as well as the probability of detecting rare events. In addition, if one data set is more completely characterized, this extra information can be extracted and applied to the other data set. For example if there is a band in common in two databases and there is sequence information for it in one database, that sequence information can be added to the other database. Currently most public electrophoresis database sites are two-dimensional protein databases. A list with links to many of these Internet database sites can be found at <http://www-lmmb.ncifcrf.gov/EP/table2Ddatabases.html>.

With biological questions becoming more complicated and the answers to the questions often requiring information from a variety of sources, it is becoming increasingly important to be able to move easily between information sources. A relational type of database can help achieve this. Unlike a conventional database with a fixed arrangement of data, relational databases have links between related files that allow for easy movement from one file to another. Another approach to interconnecting electrophoresis data with data from other sources is to generate a series of hypertext links between data sets, similar to what occurs on the Internet. Selecting a specific link moves the search to the related network site and the related information. Regardless of the method, the end goal is similar. An example of what is possible: a researcher selects a protein spot on a two-dimensional gel image which triggers accessing of related information on this protein. The protein sequence is accessed from mass spectroscopy analysis of the spot on a separate gel. The sequence of the gene and the cDNA that generated the protein is retrieved. The expression pattern of the gene in various tissues and conditions as well as information on similar genes in other organisms is incorporated. Citations and annotations to this are retrieved as well. All of this information is compiled automatically into an interactive report about the protein. From this report, the researcher can formulate a more refined hypothesis and plan the most appropriate experiments to test it.

LITERATURE CITED

- Appel, R.D., Hochstrasser, D.F., Funk, M., Vargas, J.R., Muller, A.F., and Scherrer, J.-R. 1991. The MELANIE project: From a biopsy to automatic protein map interpretation by computer. *Electrophoresis* 12:722-735.
- Garrels, J.I. 1989. The QUEST system for quantitative analysis of two-dimensional gels. *J. Biol. Chem.* 264:5269-5282.
- Glasbey, C.A. and Horgan, G.W. 1994. *Image Analysis for the Biological Sciences*. John Wiley & Sons, Chichester, England.
- Hamming, R.W. 1973. *Numerical methods for scientists and engineers*, 2nd ed. Dover Publications, New York.
- Huffman, D.A. 1952. A method for the construction of minimum-redundancy codes. *Proc. Inst. Elect. Radio Eng.* 40:9-12.
- Monardo, P.J., Boutell, T., Garrels, J.I., and Latter, G.I. 1994. A distributed system for two-dimensional gel analysis. *Comput. Appl. Biosci.* 10:137-143.
- Patton, W.F. 1995. Biologist's perspective on analytical imaging systems as applied to protein gel electrophoresis. *J. Chromatogr. A.* 698:55-87.
- Plikaytis, B.D., Carlone, G.M., Edmonds, P., and Mayer, L.W. 1986. Robust estimation of standard curves for protein molecular weight and linear-duplex DNA base-pair number after gel electrophoresis. *Anal. Biochem.* 152:346-364.
- Russ, J.C. 1995. *The Image Processing Handbook*. CRC Press, Boca Raton, Fla.
- Smith, J.M. and Thomas, D.J. 1990. Quantitative analysis of one-dimensional gel electrophoresis profiles. *Comput. Appl. Biosci.* 6:93-99.
- Sutherland, J.C., Lin, B., Monteleone, D.C., Mugavero, J., Sutherland, B.M., and Trunk, J. 1987. Electronic imaging system for direct and rapid quantitation of fluorescence from electrophoretic gels: Application to ethidium bromide-stained DNA. *Anal. Biochem.* 163:446-457.
- Welch, T.A. 1984. A technique for high performance data compression. *IEEE Computer.* 17:21-32.

KEY REFERENCES

Glasbey and Horgan, 1994. See above.

Describes general image-processing techniques as they are applied to biological images.

Russ, 1995. See above.

A general reference book on digital image capture and analysis.

Sutherland, J.C. 1993. Electronic imaging of electrophoretic gels and blots. *In Advances in Electrophoresis*, Vol. 6. (A. Chrambach, M.J. Dunn, and B.J. Radola eds.) pp. 1-41. VCH Verlagsgesellschaft mbH, Weinheim, Germany.

Provides an overview of image capture with particular emphasis on types of capture equipment.

INTERNET RESOURCES

<http://rsb.info.nih.gov/nih-image>

NIH Image is free software that provides basic image analysis tools for the Macintosh.

<http://www.inforamp.net/~poynton/Poynton-color.html>

Contains an excellent description of gamma correction in the Gamma FAQ.

<http://www-immb.ncifcrf.gov/EP/table2Ddatabases.html>

A list of links to many two-dimensional databases that are available via the Internet.

Contributed by Scott Medberry
Amersham Pharmacia Biotech
San Francisco, California

Sean Gallagher
UVP, Inc.
Upland, California

Butch Moomaw
Hamamatsu Photonic Systems
Spring Branch, Texas

DETECTION OF PROTEINS

The protocols in this section require that detectable proteins be previously separated using either one-dimensional (UNIT 10.2) or two-dimensional (UNITS 10.3 & 10.4) electrophoresis. In UNIT 10.6, detection by Coomassie blue and silver staining of protein-containing bands or spots in a gel is described. Alternatively, proteins in a gel may be electrophoretically transferred to a blot transfer membrane and detected by staining with India ink or colloidal gold (UNIT 10.7) or by immunoblotting (UNIT 10.8). Another method is via biosynthetic labeling of the protein of interest using [³⁵S]methionine (UNIT 10.18), followed by autoradiography (APPENDIX 3).

Staining Proteins in Gels

UNIT 10.6

The location of a protein in a gel can be determined by either Coomassie blue staining (see Basic Protocol 1) or silver staining (see Basic Protocol 2). The former is easier and more rapid; however, silver staining methods are considerably more sensitive and thus can be used to detect smaller amounts of protein. Rapid staining procedures are provided for each method in Alternate Protocols 1, 2, and 3. Support Protocol 1 describes how to photograph stained gels.

Fluorescent staining has become a popular alternative to traditional staining procedures, mainly because it is more sensitive than Coomassie staining, and often as sensitive as silver staining. The unit therefore includes a protocol describing SYPRO Orange or SYPRO Red staining of proteins in SDS-polyacrylamide gels (see Basic Protocol 3); variations on that procedure for proteins in nondenaturing gels are included as well (see Alternate Protocol 4). SYPRO Ruby staining of two-dimensional gels is also described (see Basic Protocol 4). Support Protocol 2 describes the photography of the fluorescently stained proteins.

Additional fluorescent techniques have been developed to monitor phosphorylated and glycosylated proteins. Alternate Protocol 5 describes the staining of phosphoproteins in minigels and larger, two-dimensional gels with Pro-Q Diamond phosphoprotein gel stain, while Support Protocol 3 tells how to image and document the findings from the phosphoprotein staining. Support Protocol 4 details a method for selectively detecting phosphotyrosine residues.

Alternate Protocol 6 details the differential fluorescent staining of glycosylated and nonglycosylated proteins, and Support Protocol 5 describes how to photograph these gels.

COOMASSIE BLUE STAINING

Detection of protein bands in a gel by Coomassie blue staining depends on nonspecific binding of a dye, Coomassie brilliant blue R, to proteins. The detection limit is 0.3 to 1 μg/protein band. In this procedure, proteins separated in a polyacrylamide gel are precipitated using a fixing solution containing methanol/acetic acid. The location of the precipitated proteins is then detected using Coomassie blue (which turns the entire gel blue). After destaining, the blue protein bands appear against a clear background. The gel can then be stored in acetic acid or water, photographed, or dried to maintain a permanent record.

BASIC PROTOCOL 1

Analysis of Proteins

10.6.1

Materials

Polyacrylamide gel (UNIT 10.2A)
Fixing solution for Coomassie blue and silver staining (see recipe)
Coomassie blue staining solution (see recipe)
Methanol/acetic acid destaining solution (see recipe)
7% (v/v) aqueous acetic acid
Whatman 3MM filter paper (optional)
Gel dryer (optional)

1. Place the polyacrylamide gel in a plastic container and cover with 3 to 5 gel volumes of fixing solution. Agitate slowly 2 hr at room temperature on an orbital shaker or rocking platform.

If agitation is too rapid, the gel may break apart. Use fixing solution only once.

2. Pour out fixing solution. Cover the gel with Coomassie blue staining solution for 4 hr and agitate slowly.

Use staining solution only once.

3. Pour out staining solution. Rinse the gel briefly with ~50 ml fixing solution.
4. Pour out fixing solution. Cover the gel with methanol/acetic acid destaining solution for 2 hr and agitate slowly.
5. Pour out destaining solution. Add fresh methanol/acetic acid destaining solution and continue destaining until blue bands and a clear background are obtained. Store the gel in 7% aqueous acetic acid or water.

Do not store gels in fixing solution because protein bands will eventually disappear. Gels can be stored up to 1 year at 4°C in a sealable plastic bag.

6. If desired, photograph the gel (see Support Protocol 1).
7. If desired, dry the gel to maintain a permanent gel record. Place the gel on two sheets of Whatman 3MM filter paper and cover top with plastic wrap. Dry in a conventional gel dryer 1 to 2 hr at ~80°C.

Alternatively, place gel on filter paper and dry according to instructions supplied with gel dryer.

ALTERNATE PROTOCOL 1

RAPID COOMASSIE BLUE STAINING

Protein bands stained using this protocol can be detected within 5 to 10 min after adding rapid Coomassie staining solution. Because the Coomassie blue concentration is lower than that used in Basic Protocol 1, the gel background never stains very darkly and the bands can be seen even while the gel remains in the staining solution. Another difference is that isopropanol is substituted for methanol in the fixing solution. This method is slightly less sensitive than the standard procedure (see Basic Protocol 1).

Additional Materials (also see Basic Protocol 1)

Isopropanol fixing solution (see recipe)
Rapid Coomassie blue staining solution (see recipe)
10% (v/v) acetic acid

1. Place a polyacrylamide gel in a plastic or glass container. Cover the gel with 3 to 5 gel volumes isopropanol fixing solution and shake gently at room temperature. For a 0.7-mm-thick gel, shake 10 to 15 min; for a 1.5-mm-thick gel, shake 30 to 60 min.

2. Pour out fixing solution. Cover the gel with rapid Coomassie blue staining solution and shake gently until desired intensity is reached, 2 hr to overnight at room temperature.

Bands will become visible even in the staining solution within 5 to 30 min, depending on gel thickness. The gel background will never stain very darkly.

3. Pour out staining solution. Cover the gel with 10% acetic acid to destain, shaking gently ≥ 2 hr at room temperature until a clear background is obtained.
4. If necessary, pour out 10% acetic acid and add more. Continue destaining until clear background is obtained. Store gel in 7% acetic acid or water, or in plastic wrap at 4°C.

It is usually unnecessary to add additional destaining solution.

5. If desired, photograph (see Support Protocol 1) or dry (see Basic Protocol 1, step 7) the gel.

SILVER STAINING

Detection of protein bands in a gel by silver staining depends on binding of silver to various chemical groups (e.g., sulfhydryl and carboxyl moieties) in proteins. The detection limit is 2 to 5 ng/protein band. In this procedure, proteins separated in a polyacrylamide gel are successively fixed in methanol/acetic acid and glutaraldehyde. After exposure to silver nitrate, the gel is treated with developer to control the level of staining. When the desired staining intensity is reached, the gel is fixed, photographed, and dried.

Materials

Polyacrylamide gel (UNIT 10.2)
Fixing solution for Coomassie blue and silver staining (see recipe)
Methanol/acetic acid destaining solution (see recipe)
10% (v/v) glutaraldehyde (freshly prepared from 50% stock; Kodak)
Silver nitrate solution (see recipe)
Developing solution (see recipe)
Kodak Rapid Fix Solution A
Whatman 3MM filter paper (optional)
Gel dryer (optional)

NOTE: Wear gloves at all times to avoid fingerprint contamination.

1. Place a polyacrylamide gel in a plastic container and add 5 gel volumes of fixing solution. Agitate slowly ≥ 30 min at room temperature on an orbital shaker.
2. Pour out fixing solution. Fix the gel with 5 gel volumes of methanol/acetic acid destaining solution for ≥ 60 min, agitating slowly.

No actual destaining takes place in this step; fixation continues using the same solution used for destaining in Basic Protocol 1.

3. Pour out destaining solution. Add 5 gel volumes of 10% glutaraldehyde and agitate slowly 30 min in a fume hood.

CAUTION: *Wear gloves and work only in a fume hood.*

4. Pour out the glutaraldehyde. Wash the gel at least four times with water, ≥ 30 min for each wash and preferably overnight for the last wash. Agitate slowly with each wash.

BASIC PROTOCOL 2

Analysis of Proteins

10.6.3

5. Pour out the last wash. Stain the gel with ~5 gel volumes silver nitrate solution (to cover the gel) for 15 min with vigorous shaking.

CAUTION: Dispose of ammoniacal silver solution immediately by flushing with copious amounts of water, as it becomes explosive upon drying.

6. Transfer the gel to another plastic box and wash five times with deionized water, exactly 1 min for each wash. Agitate slowly with each wash.
7. Dilute 25 ml developing solution with 500 ml water. Transfer gel to another plastic box, add enough diluted developer to cover the gel during agitation, and shake vigorously until the bands are as intense as desired. If the developer turns brown, change to fresh developer.

Development should be stopped immediately when gel background starts to appear.

8. Transfer to Kodak Rapid Fix Solution A for 5 min.

If necessary, swab gel surface with soaked cotton to remove residual silver deposits.

9. Pour off Rapid Fix Solution and wash the gel exhaustively in water (four to five times).
10. Photograph the gel (see Support Protocol 1).

Gels should be photographed as soon as possible because there may be slight changes in color intensity and increases in nonspecific background. The silver-stained proteins remain clearly visible for at least 18 hr.

11. Dry the gel to maintain a permanent record (see Basic Protocol 1, step 7) or store in sealable plastic bag (will last 6 to 12 months).

ALTERNATE PROTOCOL 2

NONAMMONIACAL SILVER STAINING

This nonammoniacal silver staining procedure uses more stable solutions and detects certain proteins not stained using the preceding protocol (Morrissey, 1981).

Additional Materials (also see Basic Protocol 2)

- 5 µg/ml dithiothreitol (DTT; APPENDIX 2)
- 0.1% (w/v) silver nitrate solution (store in brown bottle at room temperature up to ~1 month)
- Carbonate developing solution (see recipe)
- 2.3 M citric acid
- 0.03% (w/v) sodium carbonate (optional)

1. Place a polyacrylamide gel in a glass or polyethylene container and add 100 ml fixing solution. Agitate slowly 30 min at room temperature on an orbital shaker.

Times are appropriate for all gel sizes.

2. Pour out fixing solution. Immerse gel in methanol/acetic acid destaining solution and agitate slowly 30 min.
3. Pour out destaining solution. Cover gel with 50 ml 10% glutaraldehyde and agitate slowly 10 min in a fume hood.

CAUTION: Wear gloves and work only in a fume hood.

Fixing in glutaraldehyde ensures that the gel will not bleach (and will probably last longer) and is important for retention and detection of very small proteins.

4. Pour out glutaraldehyde. Wash the gel *thoroughly* in several changes of water (or in running water) for 2 hr to ensure low background levels.

5. Pour out water. Soak the gel in 100 ml 5 $\mu\text{g}/\text{ml}$ DTT for 30 min.

Treating proteins with DTT results in more reproducible silver staining.

6. Pour out DTT. Without rinsing, add 100 ml 0.1% silver nitrate solution and agitate slowly 30 min.
7. Pour out silver nitrate. Wash the gel once quickly with a small amount of water, then twice rapidly with a small amount of carbonate developing solution.
8. Soak the gel in 100 ml carbonate developing solution and agitate slowly until desired level of staining is achieved.
9. Stop staining by adding 5 ml 2.3 M citric acid per 100 ml carbonate developing solution for 10 min and agitate slowly.

Pay attention to the volumes of carbonate and citric acid solutions (steps 8 and 9). These must be balanced carefully to bring the pH to neutrality. If the pH is high, the reaction will not stop; if the pH is too low, the gel will bleach.

10. Pour off the solution. Wash the gel several times in water, agitating slowly 30 min.
11. Photograph the gel (see Support Protocol 1) and/or store by soaking 10 min in 0.03% sodium carbonate and wrapping in plastic wrap or sealing in a heat-sealable bag.

RAPID SILVER STAINING

This protocol (Bloom et al., 1987) is based upon the preceding nonammoniacal silver staining method. It is rapid and gives low background but may not be quite as sensitive in detecting very small proteins because there is no glutaraldehyde fixation.

Additional Materials (also see Alternate Protocol 2)

Formaldehyde fixing solution (see recipe)
0.2 g/liter sodium thiosulfate ($\text{Na}_2\text{S}_2\text{O}_3$)
Thiosulfate developing solution (see recipe)
Drying solution (see recipe)
Dialysis membrane soaked in 50% methanol
Glass plates

1. Place a polyacrylamide gel in a plastic container and add 50 ml formaldehyde fixing solution. Agitate slowly 10 min at room temperature on an orbital shaker.
Times indicated are flexible and are appropriate for a 0.75-mm \times 5.5-cm \times 8-cm, 12.5% acrylamide slab gel. Each gel is placed in an 8 \times 14-cm plastic container.
2. Pour out fixing solution. Wash the gel twice with water, 5 min for each wash. Agitate slowly for each wash.
3. Pour out water. Soak gel 1 min in 50 ml 0.2 g/liter $\text{Na}_2\text{S}_2\text{O}_3$, agitating slowly.
4. Pour out $\text{Na}_2\text{S}_2\text{O}_3$. Wash the gel twice with water, 20 sec for each wash.
5. Pour out water. Soak gel 10 min in 50 ml 0.1% silver nitrate solution, agitating slowly.
6. Pour out the silver nitrate. Wash the gel with water and then with a small volume of thiosulfate developing solution.
7. Soak the gel in 50 ml fresh thiosulfate developing solution and agitate slowly until band intensities are adequate (~1 min).

Development continues a little after stopping (next step), so do not overdevelop here.

**ALTERNATE
PROTOCOL 3**

**Analysis of
Proteins**

10.6.5

8. Add 5 ml of 2.3 M citric acid per 100 ml thiosulfate developing solution and agitate slowly 10 min.

Pay attention to the volumes of thiosulfate developing and citric acid solutions (see Alternate Protocol 2, step 9).

9. Pour off the solution. Wash the gel in water, agitating slowly 10 min.
10. Pour off the water. Soak the gel 10 min in 50 ml drying solution.
11. Sandwich gel between two pieces of wet dialysis membrane on a glass plate. Clamp edges of the plate with notebook clamps and dry overnight at room temperature.

SUPPORT PROTOCOL 1

GEL PHOTOGRAPHY OF COOMASSIE- OR SILVER-STAINED GELS

Many research facilities have photographic services that are well versed in gel photography and can provide excellent documentation of experiments. The following guidelines and equipment should be used to achieve optimal results. Any good single-lens reflex (SLR) camera attached to a copy stand will work. If larger-format cameras are available, sheet film will give spectacular resolution. The light box must produce relatively even lighting.

Kodak T-Max 400 film is a fine-grained panchromatic half-tone film that provides extremely high resolution and that works well for 35-mm gel photography. For contrast enhancement of Coomassie blue-stained gels, photographing the gels through a deep-yellow to yellow-orange filter (Wratten #8 or 9 or a yellow-orange Cokin filter) is recommended. Silver-stained gels are photographed with a blue-green filter (Wratten #58 or a blue Cokin filter). Develop according to manufacturer's instructions. Medium-contrast resin-coated paper is recommended for printing.

The instant films from Polaroid are ideal for fast, high-quality photographs of gels for laboratory notebooks and publications. Typically, type 57 (4 × 5-in., MP4 camera) or type 667 (3¼ × 4¼-in., DS 34 camera) black and white film is used for documentation at a shutter speed of 1/60 to 1/30 sec with an aperture of *f*16. Type 55 and 665 positive/negative films provide not only a high-quality print, but also a fine-grain, medium-format negative that can be used to produce multiple photographs of the gel in any size; for these films, use a shutter speed of 1/4 to 1/2 sec and an aperture of *f*16. Aperture settings of *f*11 or higher should be used for sharp photographs. If the picture is too light, increase the shutter speed or go to a higher *f* number. If the picture is too dark, decrease shutter speed or go to a lower *f* number (adapted from *Polaroid Guide to Instant Imaging*).

A simple test for the effectiveness of a filter is to place the gel on a light box and observe the gel through the filter. Increased contrast of the bands should be obvious. Silver staining can produce bands of various colors, which can present a problem in filter selection. Kodak X-Omat Duplicating Film is a useful alternative because it yields black banding in spite of the coloration caused by the silver staining (Anderson, 1988).

BASIC PROTOCOL 3

FLUORESCENT STAINING USING SYPRO ORANGE OR RED

Fluorescent dyes have a number of advantages over traditional protein stains. SYPRO Orange and Red protein gel stains, outlined in the procedure below, can detect 1 to 2 ng protein/band. This is more sensitive than Coomassie brilliant blue staining and as sensitive as many silver staining techniques. Staining is straightforward, with less hands-on time than typical silver staining protocols, and is complete in <1 hr. Stained proteins can be visualized using a standard 300-nm UV transilluminator or a laser scanner. Although the protocol below is limited to one-dimensional SDS-PAGE, alternative kits for native,

isoelectric focusing (IEF), and two-dimensional SDS-PAGE applications are available from the supplier (Molecular Probes). Coomassie Fluor Orange protein gel stain, a premixed fluorescent stain, is also available from Molecular Probes.

The staining properties of the two SYPRO dyes are similar, and both are equally suitable for use in most procedures. The SYPRO Orange gel stain is slightly brighter, whereas the SYPRO Red gel stain has somewhat lower background fluorescence. For those using a laser-excited gel scanner, the authors recommend the SYPRO Orange stain for argon-laser-based instruments and the SYPRO Red stain for instruments that employ green He-Ne or Nd/YAG (neodymium:yttrium-aluminum-garnet) lasers. Both dyes are efficiently excited by UV or broad-band illumination and, with the proper filters, work nicely with CCD camera archiving systems.

Materials

SYPRO Orange or Red fluorescent staining solution (see recipe)

7.5% (v/v) acetic acid

0.1% (v/v) Tween-20

Additional reagents and equipment for one-dimensional SDS-PAGE (UNIT 10.2A)

1. Prepare and separate proteins using SDS-PAGE (UNIT 10.2A), but use 0.05% SDS in the running buffer instead of the usual 0.1% SDS.

Gels run in 0.1% SDS show the same sensitivity for staining as those run with lower SDS concentrations, but require either more time in staining solution or a 10-min rinse in water before staining to reduce the background fluorescence that is produced by dye interaction with SDS. Gels run in 0.05% SDS show no change in the migration of proteins and can be photographed sooner because they require less time in the staining solution to clear the SDS from the gel. Gels run in SDS concentrations <0.05% or in old running buffer exhibit poor resolution of bands and other problems, so it is essential that the SDS stock solution used to prepare the running buffer be fresh and be at the proper SDS concentration.

Do not fix the proteins in the gel with methanol-containing solutions. Methanol removes the SDS coat from proteins, strongly reducing the signal from SYPRO Orange or Red stains.

2. Pour SYPRO Orange or Red fluorescent staining solution into a small plastic dish. For one or two standard-size minigels, use ~50 ml staining solution. For larger gels, use between 500 and 750 ml staining solution.

Staining dishes should be cleaned and rinsed well before use, as detergent will interfere with staining.

Alternatively, use Coomassie Fluor Orange protein gel stain which is available as a premixed staining solution.

3. Place gel into the staining solution. Cover the container with aluminum foil to protect the dye from bright light.

The staining solution may be reused up to four times. However, due to reduced sensitivity, the use of fresh staining solution is recommended.

Several alternate methods of staining can be used. (1) Gels may be stained in sealable plastic bags. However, it is still important to use the proper amount of staining solution. (2) For low-percentage gels and for very small proteins, increasing the staining solution from 7.5% to 10% acetic acid will result in better retention of the protein in the gel without compromising sensitivity. (3) Protein gel stains can be dissolved 5000-fold into the cathode (top) running buffer to stain proteins as the gel runs. The dye moves through the gel with the SDS front, so that all sizes of protein are stained. Staining does not influence relative migration of proteins through the gel. This method results in poorer protein staining than the standard poststaining method, and requires the same amount of time because the gel must be destained for 15 to 40 min in 7.5% acetic acid to reduce background fluorescence.

4. Gently agitate at room temperature for 10 to 60 min.

The staining time depends on the thickness and percentage of the gel. For 1-mm-thick 15% polyacrylamide gels, the signal is typically optimal after 40 to 60 min staining. Once the optimal signal is achieved, additional staining time (several hours to overnight) does not enhance or degrade the signal. Gels can be left in stain for up to a week with only a small loss in sensitivity; the detection limits under these conditions are ~2 to 4 ng/band.

5. Rinse briefly (<1 min) with 7.5% acetic acid.

This brief rinse removes excess stain from the gel surface to reduce background fluorescence on the surface of the transilluminator or gel scanner (see Support Protocol 2).

6. Store gel in staining solution, protected from light.

The signal decreases somewhat after several days, but, depending on the amount of protein in the bands, gels may retain a usable signal for many weeks. Gels may be left in staining solution overnight without losing sensitivity. However, fixation in acetic acid is relatively mild, so, for low-percentage gels or very small proteins, photographs should be taken as soon as possible after staining, before the proteins begin to diffuse.

Gels may be dried between cellophane membrane backing sheets (Bio-Rad), although there is sometimes a slight decrease in sensitivity. If the gels are dried onto paper, the light will scatter and the sensitivity will decrease. Other types of plastic sheets are not typically transparent to UV light. Store dried gels in the dark to prevent photobleaching.

7. Destain gel by incubating overnight in 0.1% Tween-20.

Alternatively, incubating in several changes of 7.5% acetic acid will eventually remove all of the stain. Incubating in methanol will strip off dye and SDS, but will also precipitate proteins.

8. Photograph gel (see Support Protocol 2).

ALTERNATE PROTOCOL 4

FLUORESCENT STAINING ON NONDENATURING GELS USING SYPRO ORANGE OR RED

Proteins can be stained after native gel electrophoresis (UNIT 10.2B) by dissolving SYPRO dyes in water and then following Basic Protocol 3. Staining proteins in nondenaturing gels is highly protein selective and will generally be less sensitive than staining proteins in SDS gels; however, because there is essentially no background fluorescence, photographic exposures can be very long. If it is not necessary to maintain the protein in a native state after electrophoresis, the best sensitivity can be achieved if the gel is soaked in 0.05% SDS for ~30 min and then stained with a solution of SYPRO dye diluted in 7.5% acetic acid.

BASIC PROTOCOL 4

FLUORESCENT PROTEIN STAIN USING SYPRO RUBY FOR 2-D GEL ANALYSIS

Fluorescent gel stains are gaining popularity due to the combination of simplicity and sensitivity that they offer. SYPRO Ruby protein gel stain is an ultrasensitive, fluorescent stain for the specific detection of proteins separated by polyacrylamide gel electrophoresis (PAGE). This stain, designed especially for use in 2-D PAGE (Fig. 10.6.1; UNIT 10.4), has proven to be the most sensitive protein gel stain for standard 1-D SDS-PAGE (UNIT 10.2A) and an excellent stain for isoelectric focusing (IEF) gels (UNIT 10.3) as well. SYPRO Ruby protein gel stain achieves comparable sensitivity to that of the silver-staining techniques (see Basic Protocol 2 and Alternate Protocols 2 and 3) and offers substantial processing advantages including the absence of overstaining, a linear quantitation range of over three orders of magnitude, and less protein-to-protein variability. In addition, the SYPRO Ruby

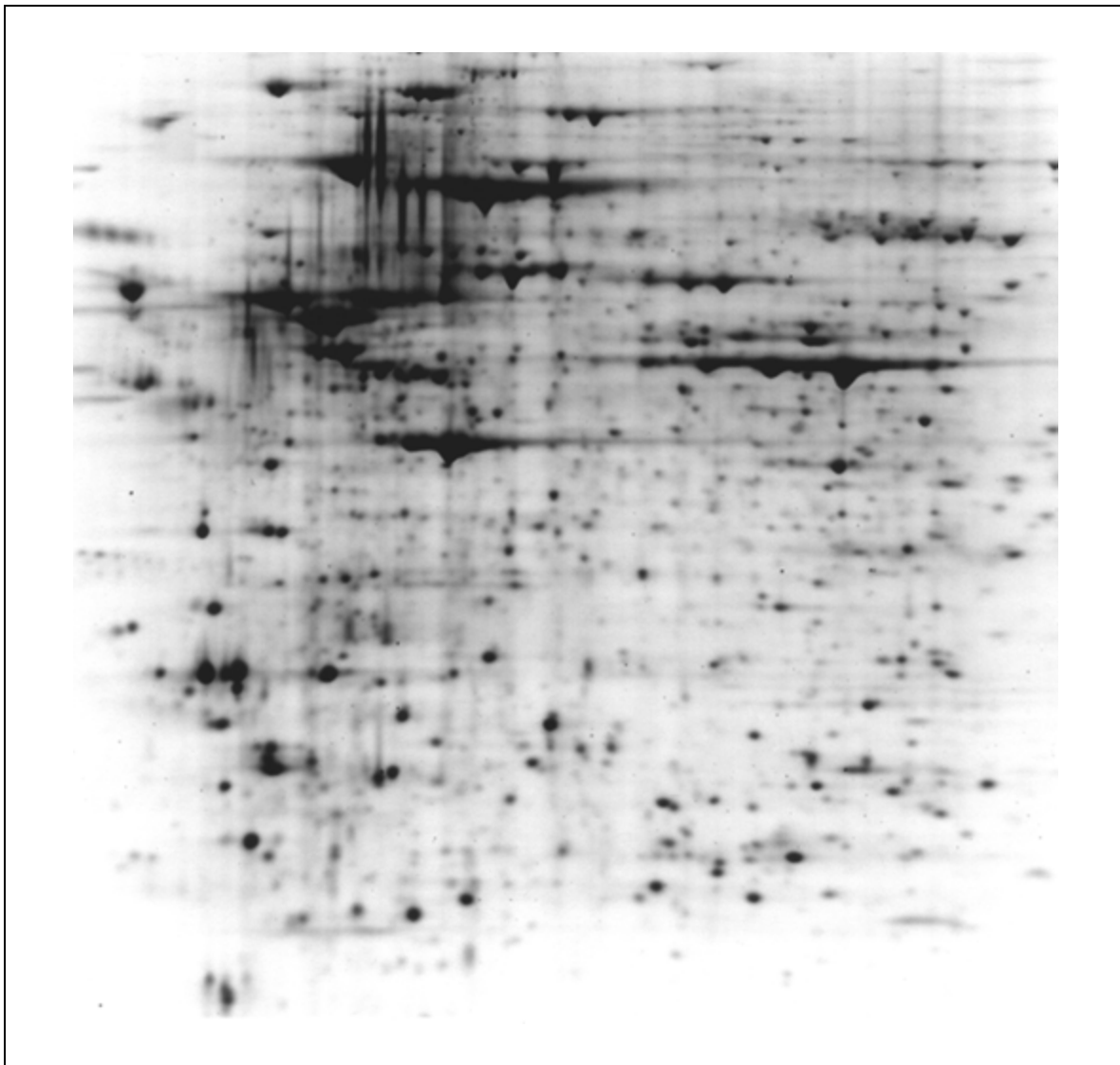


Figure 10.6.1 Typical two-dimensional gel stained with SYPRO Ruby gel stain.

staining protocol below stains glycoproteins, lipoproteins, calcium-binding proteins, fibrillar proteins, and other difficult-to-stain proteins, and does not interfere with subsequent analysis of proteins by Edman-based sequencing or mass spectrometry. The stain can be used with many types of gels, including 2-D gels (*UNITS 10.3 & 10.4*), Tris-glycine SDS gels (*UNIT 10.2A*), and Tris-tricine precast SDS gels and nondenaturing gels (*UNIT 10.2B*). SYPRO Ruby stain is also compatible with gels adhering to plastic backings.

Materials

1-D or 2-D polyacrylamide *or* IEF gels (*UNITS 10.2-10.4*)

Fixing solution for SYPRO Ruby staining of 2-D polyacrylamide gels (see recipe)

Fixing solution for IEF gels: 40% (v/v) methanol/10% (w/v) trichloroacetic acid in H₂O

SYPRO Ruby protein gel stain (Molecular Probes; see recipe)

10% (v/v) methanol (or ethanol)/7% (v/v) acetic acid

2% (v/v) glycerol

Plastic staining dishes of appropriate size for gels: polypropylene (e.g., Rubbermaid Servin' Savers) or PVC photographic staining trays (e.g., Photoquip Cesco-Lite 8-in. × 10-in., for large-format 2-D gels)

Orbital shaker

Additional reagents and equipment for photography of fluorescently stained gels (see Support Protocol 2)

Fix, stain, and wash the gel

- 1a. *For 1-D polyacrylamide gels:* Proceed directly to step 2 (no fixation required).
- 1b. *For 2-D polyacrylamide gels:* Fix for 30 min in an appropriate fixing solution for SYPRO Ruby staining of polyacrylamide gels.

It is also possible to use serial combinations of these fixatives. Note that the combination of ethanol and acetic acid can result in the formation of ethyl acetate, which is not only toxic but may interfere with identification of proteins by mass spectrometry.

- 1c. *For IEF gels:* Fix for 3 hr in 40% methanol/10% trichloroacetic acid, and then perform three washes, each for 10 min in distilled water, before proceeding to the staining step.
2. Transfer the gel to a clean polypropylene or PVC dish of appropriate size. Incubate the gel in the appropriate amount of undiluted SYPRO Ruby protein gel stain with continuous gentle agitation, e.g., on an orbital shaker at 50 rpm. For maximum sensitivity in 1-D or 2-D gels, incubate the gel for at least 3 hr. For IEF gels, incubate the gel overnight.

The authors have found that polypropylene dishes, such as Rubbermaid Servin' Savers, are the optimal containers for staining because the high-density plastic adsorbs only a minimal amount of the dye. For small gels, circular staining dishes provide the best fluid dynamics on orbital shakers, resulting in less dye aggregation and better staining. For large-format 2-D gels, polyvinyl chloride (PVC) photographic staining trays, such as Photoquip Cesco-Lite 8 inch × 10 inch photographic trays also work well. Glass dishes are not recommended. The minimal staining volumes for typical gel sizes are: 50 ml, for 8 cm × 10 cm × 0.75-mm gels; 330 ml, for 16 cm × 20 cm × 1-mm gels; 500 ml, for 20 cm × 20 cm × 1-mm gels, or ~10× the volume of the gel for other gel sizes.

Using too little stain will lower the sensitivity. For convenience, gels may be left in the dye solution overnight or longer without overstaining. Do not dilute the stain, as diluted stain will result in decreased sensitivity. Do not reuse the staining solution, as this will result in a significant loss of sensitivity.

3. To reduce background fluorescence and increase sensitivity, transfer the gel to a clean staining dish and wash it in 10% methanol (or ethanol)/7% acetic acid for 30 min. For polyacrylamide gels, perform this wash once; for IEF gels, perform the wash three times.

This transfer step helps to minimize the deposition of stain speckles on the gel. To reduce organic waste, stained gels may alternatively be washed in distilled water, although this method does not reduce background fluorescence to the same extent. The gel may be monitored periodically using UV epi-illumination to determine the level of background fluorescence.

4. *Optional (for permanent storage):* Incubate the gel in a solution of 2% (v/v) glycerol at room temperature for 30 min. Dry the stained gel using a gel dryer.

Note that proteins present at very low levels may no longer be detectable after gel drying.

5. View and photograph the gel (see Support Protocol 2).

SYPRO Orange or Red Stains

Photographing and archiving the gel is essential to obtain high sensitivity. The camera's integrating effect can make bands visible that are not visible to the eye. Place the fluorescently labeled gel directly on a standard 300-nm UV transilluminator or a blue-light transilluminator (e.g., Clare Chemical Dark Reader). Plastic wraps, such as Saran Wrap, should not be used, as they fluoresce naturally and will fluoresce even more when exposed to SYPRO Orange or Red stain. This results in a large background signal, making it impossible to achieve good sensitivity. Pharmacia PhastGels have a polyester backing material (Gelbond) that is not only highly autofluorescent, but also binds the SYPRO Orange and Red protein gel stains, producing additional background fluorescence. Consequently, the plastic backing should be removed before trying to visualize bands. Pharmacia markets a gel backing remover for use with their Phast Transfer system. The surface of the transilluminator should be cleaned with water and a soft cloth after use, to minimize the buildup of fluorescent dyes.

When using a Polaroid camera, use Polaroid 667 black-and-white print film and a SYPRO protein gel stain photographic filter (Molecular Probes) to obtain the highest sensitivity. Do not use standard ethidium bromide filters, as they will block much of the light and lead to lower sensitivity. Supplemental UV blocking filters are not usually required. Polaroid 667 film is a fast film with an ISO rating of ASA 3000. The use of different film types may require longer exposure times or different filters. Exposure time will vary with the intensity of the illumination source; with an *f*-stop of 4.5, 2 to 5 sec is typical for SYPRO Orange, and 3 to 8 sec is typical for SYPRO Red. Noticeable photobleaching can occur after several minutes of exposure to ultraviolet light. If a gel becomes photobleached, it can be restained by simply returning it to the staining solution.

Charge-coupled device (CCD) cameras provide good sensitivity. Contact the camera manufacturer for the optimal filter sets to use. For those using a laser-excited gel scanner, the SYPRO Orange stain is recommended for argon laser-based instruments, and the SYPRO Red stain is recommended for instruments that employ green He-Ne or Nd/YAG lasers.

SYPRO Ruby Stain

The SYPRO Ruby protein gel stain has two excitation maxima, one at ~280 nm and one at ~450 nm, and has an emission maximum near 610 nm. Proteins stained with the dye can be visualized using a 300-nm UV transilluminator, a blue-light transilluminator, or a laser scanner. The stain has exceptional photostability, allowing long exposure times for maximum sensitivity.

UV or blue-light transilluminator

Proteins stained with SYPRO Ruby protein gel stain are readily visualized using a UV or blue-light source. The use of a photographic camera or CCD camera and the appropriate filters is essential to obtain the greatest sensitivity. The instrument's integrating capability can make bands visible that cannot be detected by eye. It is important to clean the surface of the transilluminator after each use with deionized water and a soft cloth (like cheese-cloth); otherwise fluorescent dyes such as SYPRO stains, SYBR stains, and ethidium bromide will accumulate on the glass surface and cause a high background fluorescence. The authors use a 300-nm transilluminator with six 15-W bulbs. Excitation with different light sources may not give the same sensitivity. Using a Polaroid camera and Polaroid 667 black-and-white print film, the highest sensitivity is achieved with a 490-nm long-pass filter, such as the SYPRO protein gel stain photographic filter (S-6656), available

from Molecular Probes. The authors typically photograph minigels using an *f*-stop of 4.5 for 1 sec. Using a CCD camera, images are best obtained by digitizing at $\sim 1024 \times 1024$ pixels resolution with 12- or 16-bit grayscale levels per pixel. The camera manufacturer should be consulted for recommendations on filter sets to use. A CCD camera-based image analysis system can gather quantitative information that will allow comparison of fluorescence intensities between different bands or spots. Using such a system, the authors have found that the SYPRO Ruby gel stain has a linear dynamic range over three orders of magnitude. The polyester backing on some premade gels is highly fluorescent. For maximum sensitivity using a UV transilluminator, the gel should be placed polyacrylamide-side-down and an emission filter, such as the SYPRO protein gel stain photographic filter, should be used to screen out the blue fluorescence of the plastic. The use of a blue-light transilluminator (UVP Visi-Blue Transilluminators) or laser scanner will reduce the amount of fluorescence from the plastic backing so that the gel may be placed polyester-side down.

Laser-scanning instruments

Gels stained with the SYPRO Ruby protein gel stain can be visualized using imaging systems equipped with lasers that emit at 450, 473, 488, or 532 nm. For the analysis of stained proteins, note that SYPRO Ruby dye sometimes generates small speckles of precipitated dye on the gel. The speckles have diameters $\sim 20\%$ the size of the smallest stained protein spot, making them very easy to distinguish. Analysis software for 2-D gels will ignore small speckles if the minimum spot size of the program is set appropriately (determined by trial and error).

For the identification of individual protein spots, it should be noted that SYPRO Ruby protein gel stain does not bind covalently to proteins. Edman-based sequencing or mass spectrometry data can be obtained after staining, with no interference from the stain. Accurate mass spectrometry has been performed on a spot containing as little as 75 fmol of stained protein.

ALTERNATE PROTOCOL 5

FLUORESCENT PHOSPHOPROTEIN GEL STAINING FOR SELECTIVELY STAINING PHOSPHOPROTEINS IN POLYACRYLAMIDE GELS

Fluorescent phosphoprotein gel staining provides a method for selectively staining phosphoproteins in polyacrylamide gels. It is ideal for identification of kinase targets in signal transduction pathways (see Chapter 18) and for phosphoproteomic studies. This fluorescent stain allows direct, in-gel detection of phosphate groups attached to tyrosine, serine, or threonine residues. The stain can be used with standard SDS-polyacrylamide gels or with 2-D gels. Blotting is not required, and there is no need for phosphoprotein-specific antibodies or immunoblot detection reagents. The protocol delivers results in 4 to 5 hr. The stain is also compatible with mass spectrometry, allowing analysis of the phosphorylation state of entire proteomes with detection of as little as 1 to 16 ng of phosphoprotein per band. For individual phosphoproteins, the strength of the signal correlates with the number of phosphate groups and is linear over three orders of magnitude. The fluorescent stain with its $\sim 555/580$ nm excitation/emission maxima can be detected by use of a visible-light scanning instrument, a visible-light transilluminator, or a 300-nm transilluminator.

Materials

- Protein-containing sample of interest
- Methanol, spectroscopy grade
- Chloroform, spectroscopy grade
- Appropriate $1\times$ sample buffer for electrophoresis (*UNITS 10.2-10.4*)

Table 10.6.1 Examples of Commercially Available Phosphorylated and Nonphosphorylated Proteins for Use as Controls

Protein	Mol. wt. (Da)	Number of phosphate residues	Lower limit of detection
Riboflavin-binding protein	29,200	8	1-3 ng
α -Casein	23,600	8	1-2 ng
β -Casein	24,500	5	1-2 ng
Ovalbumin	45,000	2	4-8 ng
Pepsin	35,500	1	8-16 ng
Carbonic anhydrase	30,000	0	Not applicable
Bovine serum albumin (BSA)	66,000	0	Not applicable

Phosphoprotein standards (PeppermintStick Phosphoprotein Molecular Weight Standards, Molecular Probes; also see Table 10.6.1)

Fixing solution for phosphoprotein gels (see recipe)

Pro-Q Diamond phosphoprotein gel stain (Molecular Probes; see recipe)

Phosphoprotein gel destain solution (see recipe, or purchase from Molecular probes)

Polystyrene staining dish (e.g., a weighing dish)

Orbital shaker

Additional reagents and equipment for polyacrylamide gel electrophoresis (*UNITS 10.2-10.4*), SYPRO Ruby protein gel staining (see Alternate Protocol 5), Coomassie blue stain (see Basic Protocol 1), silver staining (see Basic Protocol 2), and imaging and documenting phosphoprotein-stained gels

Prepare samples by desalting and delipidating

1. Place a 150- μ l sample containing ~150 to 300 μ g of protein in a 1.5-ml microcentrifuge tube. Add 600 μ l of methanol and mix well by vortexing, then add 150 μ l of chloroform and mix well by vortexing. Finally, add 450 μ l of distilled water and mix well by vortexing.

A delipidated and desalted sample is essential for adequate separation of the proteins by electrophoresis and subsequent staining by Pro-Q Diamond phosphoprotein gel stain.

2. Microcentrifuge 5 min at ~12,000 rpm, room temperature. Discard the upper phase, keeping the white precipitation disc that forms between the upper and lower phases. Add 450 μ l of methanol and mix well by vortexing.
3. Microcentrifuge 5 min at ~12,000 rpm, room temperature. Discard the supernatant and dry the pellet in a Speedvac evaporator for 10 min. Resuspend the pellet in standard 1 \times sample buffer for electrophoresis.

Separate by electrophoresis

4. Separate the proteins using standard polyacrylamide electrophoresis techniques (*UNITS 10.2-10.4*) along with phosphoprotein standards. To ensure detection of less abundant phosphoproteins, use approximately the same mass of protein that would be used for a typical Coomassie blue dye-stained gel (see Basic Protocol 1).

Use known phosphorylated and nonphosphorylated proteins as controls to help verify the phosphorylation status of the unknown protein. Table 10.6.1 lists several commonly available proteins that can serve as positive and negative controls

Fix the gel

- 5a. *For minigels:* Transfer gel to a staining dish (plastic container or large plastic weighing dish). Immerse the gel in ~100 ml of fixing solution for phosphoprotein minigels (containing acetic acid) and incubate at room temperature with gentle agitation, e.g., on an orbital shaker at 50 rpm, for at least 30 min. If needed, repeat the fixation step once more to ensure that all of the SDS is washed out of the gel. If desired, leave gels in the fixing solution overnight.

If reusing a plastic container, clean thoroughly and rinse it with 70% ethanol.

Adhere strictly to the volumes and times specified in this protocol for fixation, washing, staining, and destaining. Replicating the protocol is essential for consistent gel-to-gel and day-to-day comparisons.

- 5b. *For larger 2-D gels:* Transfer gel to a staining dish (plastic container or large plastic weighing dish). Immerse the gel in ~500 ml fix solution for phosphoprotein 2-D gels (containing trichloroacetic acid) and incubate at room temperature overnight with gentle agitation, e.g., on an orbital shaker at 50 rpm. Perform a second fixation, for 1 hr, to ensure that all of the SDS is washed out of the gel.

In both of the above fixation steps, the second fixation is especially important if non-electrophoresis-grade SDS has been used.

Wash the gel

- 6a. *For minigels:* Incubate the gel in ~100 ml water with gentle agitation for 10 min. Repeat this step for a total of two washes.
- 6b. *For larger 2-D gels:* Incubate the gel in ~500 ml water with gentle agitation for 15 min. Repeat this step three times for a total of four washes.

In both of the above washing steps, it is important that the gel be completely immersed in the water in order to remove all of the methanol and acetic acid from the gel. Residual methanol or acetic acid will interfere with Pro-Q Diamond phosphoprotein staining.

Stain the gel

- 7a. *For minigels:* Incubate the gel in the dark in 50 ml of Pro-Q Diamond phosphoprotein gel stain with gentle agitation for 75 to 120 min.
- 7b. *For larger 2-D gels:* Incubate the gel in the dark in 500 ml of Pro-Q Diamond phosphoprotein gel stain with gentle agitation for 3 to 4 hr.

In both of the above staining steps, if directly comparing multiple gels, it is important that the incubation time be the same for each gel. Under no circumstances should the gel be stained overnight, as this will result in higher background staining.

For selective detection of phosphotyrosine residues, see Support Protocol 5.

Destain the gel

- 8a. *For minigels:* Incubate the gel in the dark in 80 ml of phosphoprotein gel destain solution at room temperature for a total of ~3 hr with two changes of destain solution (e.g., three incubations of 60 min each).

The first destaining step may be as short as 45 min; the last destaining step may be as long as overnight.

- 8b. *For larger 2-D gels:* Incubate the gel in the dark in 500 ml of phosphoprotein gel destain solution at room temperature for a total of ~4 hr with three changes of destain solution (e.g., four incubations of 60 min each).

If desired, the last destaining step may be overnight.

Image and document the gel

9. Image and document the phosphoprotein-stained gel using the appropriate instrumentation (see Support Protocol 3).

The phosphoprotein staining pattern must be viewed and documented before proceeding with total-protein staining (steps 10 a or b), because the stain will be washed away during the staining procedure for total protein.

Stain the gel for total protein

- 10a. *For minigels:* After obtaining results with the Pro-Q Diamond phosphoprotein gel stain, stain the gel with a quantitative total-protein stain, such as SYPRO Ruby protein gel stain (see Alternate Protocol 5), to ascertain the relative phosphorylation state of proteins.

In this way, an abundant nonphosphorylated protein that exhibits low nonspecific staining with Pro-Q Diamond stain can be distinguished from a less abundant highly phosphorylated protein. Note that nonquantitative total-protein stains, such as silver stains, are much less useful in this application.

- 10b. *For larger 2-D gels:* After staining with Pro-Q Diamond stain, treat the gel with a total-protein stain, such as SYPRO Ruby protein gel stain (see Alternate Protocol 5), Coomassie blue stain (see Basic Protocol 1), or a silver stain (see Basic Protocol 2).

A quantitative stain such as Coomassie blue or SYPRO Ruby protein gel stain will be the most useful, because it will aid in determining the relative phosphorylation state of a given protein. Furthermore, for 2-D gels, total protein staining makes it easier to localize a protein to a particular spot within a complex protein pattern.

IMAGING AND DOCUMENTING THE PHOSPHOPROTEIN-STAINED GEL

The phosphoprotein stain used in Alternate Protocol 5 has an excitation maximum at ~555 nm and an emission maximum at ~580 nm. Imaging instruments with light sources and filters that match the excitation and emission maxima will result in the highest sensitivity.

Visible Light–Based Scanners

Stained gels are best visualized using excitation at 532 to 560 nm, such as that provided with a visible light laser–based or xenon arc lamp–based gel-scanning instrument. For most instruments, a ~580 nm long-pass or a ~600 nm band-pass emission filter is recommended.

Transillumination

Stained gels can be visualized on a blue-light transilluminator, such as the Visi-Blue series of transilluminators (UVP), Dark Reader transilluminator (Clare Chemical Research), or on a 300-nm UV transilluminator; however, the sensitivity will be lower than with a scanning instrument. Images can be documented using either conventional or digital photography. With a Polaroid camera and Polaroid 667 black-and-white film, use an appropriate long-pass filter, such as the SYPRO photographic filter (S-6656), and exposure times of ~15 to 30 sec. The red-orange filters typically used to photograph gels stained with ethidium bromide will not work well. For digital cameras, use a filter that corresponds closely to the emission characteristics of the stain, such as a 600 nm band-pass filter.

SUPPORT PROTOCOL 3

Analysis of Proteins

10.6.15

SELECTIVE DETECTION OF PHOSPHOTYROSINE RESIDUES

Phosphotyrosine residues can be selectively detected by removing phosphoserine and phosphothreonine residues through a β -elimination reaction. To detect only phosphotyrosine residues, one must perform the β -elimination reaction before carrying out Pro-Q Diamond staining as in Alternate Protocol 5. To obtain data on all phosphorylation sites, stain with Pro-Q Diamond as in Alternate Protocol 5, then document the gel image. Next, perform the β -elimination reaction and stain the gel with Pro-Q Diamond stain again. All spots or bands appearing in the first staining but not appearing in the second staining will be due to phosphoserine or phosphothreonine residues.

Additional Materials (also see *Alternate Protocol 5*)

Gel for phosphoprotein fluorescent staining (either before fixation/staining or after fixation/staining/destaining; see *Alternate Protocol 5*)

Barium hydroxide octahydrate

Argon source

Glacial acetic acid

50°C shaking water bath

1. If the gel has not already been fixed and stained, perform fix and wash steps (see *Alternate Protocol 5*, steps 5a or b and 6a or b). If the gel has already been fixed, stained, and destained, then perform only the wash procedure (see *Alternate Protocol 5*, step 6a or b).
2. Prepare a saturated solution of barium hydroxide by dissolving 12.6 g barium hydroxide octahydrate in 40 ml of degassed distilled water. Mix for 15 to 20 min and centrifuge 10 min at $10,000 \times g$, room temperature, to pellet any insoluble barium hydroxide. Store under argon gas.
3. Incubate 40 ml of the saturated barium hydroxide solution in a 50°C water bath for 30 min. At the same time, warm 40 ml of degassed distilled water to 50°C in the water bath.
All solutions should be treated with argon gas to remove atmospheric carbon dioxide and prevent the formation of insoluble barium carbonate.
4. Mix 40 ml of the warmed barium hydroxide solution with 40 ml of the warmed degassed water; and incubate the gel in the diluted solution at 50°C for 30 min with gentle agitation.
5. Stop the reaction by lowering the pH to 4.0 with addition of ~6 ml glacial acetic acid.
6. Wash and stain the gel (see *Alternate Protocol 6*, steps 6a or b and 7 a or b).
7. Document the results (see *Support Protocol 3*).

FLUORESCENT STAINING FOR DIFFERENTIALLY STAINING GLYCOSYLATED AND NONGLYCOSYLATED PROTEINS IN THE SAME GEL

Fluorescent staining provides a powerful method for differentially staining glycosylated and nonglycosylated proteins in the same gel. The technique combines a highly sensitive glycoprotein stain with ultrasensitive SYPRO Ruby protein gel stain. Both stains provide simple, sensitive, and robust detection. The Pro-Q Emerald 300 glycoprotein stain reacts with periodate-oxidized carbohydrate groups, creating a bright green fluorescent signal on glycoproteins. Using this stain, it is possible to detect as little as 0.5 ng of glycoprotein per band, depending upon the nature and the degree of glycosylation, making it about 50-fold more sensitive than the standard periodic acid–Schiff base method using acidic fuchsin dye. The green fluorescent signal can be visualized with 300-nm UV illumination. The second staining uses SYPRO Ruby protein gel stain (see *Basic Protocol 4*) to detect

total protein. This easy-to-use fluorescent stain provides the same sensitivity as silver staining, but has the advantage that it does not require glutaraldehyde, which can produce false positive responses when glycoproteins are stained. The use of SYPRO Ruby stain makes it possible to detect contaminating proteins and to easily compare the sample with molecular weight standards. For 2-D gels, total-protein staining makes it easier to localize a protein to a particular spot in the complex protein pattern. Proteins show orange-fluorescent staining when illuminated with a 300-nm UV transilluminator or a laser scanner with a 473-nm, 488-nm, or 532-nm light source. Glycosylated molecular weight standards are also highlighted containing a mixture of glycosylated and nonglycosylated proteins, which, when separated by electrophoresis, provide alternating positive and negative controls.

Materials

Protein-containing sample of interest

8-cm × 10-cm SDS- polyacrylamide minigel (*UNIT 10.2A*)

Sample buffer (*UNIT 10.2A*)

Pro-Q Emerald glycoprotein gel staining kit (Molecular Probes) including:

Pro-Q Emerald 300 staining reagent (component A), 50× concentrate in DMF (store at −20°C up to 6 months, protected from light)

Pro-Q Emerald 300 staining buffer (component B; store at room temperature up to 6 months)

Oxidizing reagent (component C): 2.5 g of periodic acid (add 250 ml of 3% v/v acetic acid and store at room temperature up to 6 months)

CandyCane glycoprotein molecular weight standards (store at −20°C up to 6 months)

Fixing solution: 50% (v/v) methanol in H₂O

Wash solution: 3% (v/v) glacial acetic acid in H₂O

Polystyrene staining dish (e.g., large weighing dish)

Orbital shaker

Additional reagents and equipment for SDS-PAGE (*UNIT 10.2A*), viewing and documenting glycoprotein-stained gels (see Support Protocol 5), and SYPRO Ruby staining (see Basic Protocol 4)

NOTE: The following procedure is optimized for staining 0.5- to 0.75-mm thick, 8-cm × 10-cm minigels. Large 2-D gels (20 cm × 20 cm) require much larger volumes and longer fixation and staining times, as indicated in the annotations to the respective steps.

1. Dilute protein sample to ~10 to 100 μg/ml with sample buffer and load 5 to 10 μl of the diluted sample per lane of an 8-cm × 10-cm polyacrylamide gel. Also dilute 0.5 μl of the CandyCane standard mixture (from the Pro-Q Emerald glycoprotein gel staining kit) with 7.5 μl of sample buffer, vortex, and load in a lane of the gel. Perform standard SDS-PAGE as described in *UNIT 10.2A*.

In the standard lane there will be ~250 ng of each standard protein included in the CandyCane standard mix; for larger gels increase the amount of standard and buffer used. Larger gels typically require twice as much material for sample and standards.

2. Transfer gel to a polystyrene staining dish and immerse the gel in 100 ml of fixing solution (50% methanol) and incubate at room temperature with gentle agitation (e.g., on an orbital shaker at 50 rpm) for 45 min. Repeat this wash step to ensure that all of the SDS is washed out of the gel.

For large (20 × 20-cm) 2-D gels, use 700 ml of fixing solution and incubate at room temperature overnight.

3. Wash the gel by incubating in 50 ml of wash solution (3% acetic acid) with gentle agitation for 10 min. Repeat this step once.

Use 700 ml of wash solution for large 2-D gels.

4. Oxidize the carbohydrates by incubating the gel in 25 ml of oxidizing solution (component C in the Pro-Q Emerald 300 kit; 2.5 g periodic acid in 250 ml of 3% acetic acid) with gentle agitation for 30 min.

Large 2-D gels require 500 ml of oxidizing solution and should be incubated for 1 hr. The 250-ml volume of oxidizing solution can be diluted with 250 ml of 3% acetic acid in order to have an adequate volume for large 2-D gels.

The Pro-Q Emerald glycoprotein gel staining kit provides sufficient materials to stain ten 8-cm × 10-cm, 0.5 to 0.75-mm thick gels.

5. Wash the gel by incubating in 50 ml (700 ml for large 2-D gels) of wash solution (3% acetic acid) with gentle agitation for 5 to 10 min. Repeat this step two more times for a total of three washes.

For large 2-D gels, 700 ml wash solution should be used for each wash and three additional washes (for a total of four washes) should be performed.

6. Just before use, dilute the 50× concentrate of Pro-Q Emerald 300 staining reagent (component A in the kit) 50-fold into Pro-Q Emerald 300 staining buffer (component B in the kit); e.g., dilute 500 µl of component A into 25 ml component B. Stain the gel by placing it in the dark in 25 ml of the mixed staining solution while gently agitating for 90 to 120 min.

The signal can be seen after about 20 min and maximum sensitivity is reached at about 120 min. The authors do not recommend staining overnight. A 200-ml volume of staining solution and staining period of 2.5 hr are required for large 2-D gels.

7. Wash the gel in 50 ml wash solution (3% acetic acid) at room temperature for 15 min. Repeat this wash once for a total of two washes. Do not leave the gel in wash solution >2 hr, as the staining will start to decrease.

Use 700 ml of wash solution for large 2-D gels.

8. View and photograph the gel to document the staining pattern (see Support Protocol 5).

For best results, the Pro-Q Emerald 300 stain should be used first and the glycoprotein staining pattern documented before proceeding with SYPRO Ruby staining. After SYPRO Ruby staining, the fluorescent signal from the Pro-Q Emerald 300 glycoprotein stain can still be seen, but the sensitivity will be somewhat decreased.

9. Stain the gel with SYPRO Ruby (see Basic Protocol 4).

10. View and photograph the gel (see Support Protocol 5).

SUPPORT PROTOCOL 5

VIEWING AND PHOTOGRAPHING THE GLYCOPROTEIN-STAINED GEL

The green-fluorescent Pro-Q Emerald 300 staining should be viewed and documented before staining total proteins with SYPRO Ruby protein gel stain. The Pro-Q Emerald 300 stain has an excitation maximum at ~280 nm and an emission maximum near 530 nm. Stained glycoproteins can be visualized using a 300-nm UV transilluminator (UVP). The use of a photographic camera or CCD camera and the appropriate filters is essential to obtain the greatest sensitivity. The instrument's integrating capability can make bands visible that cannot be detected by eye. It is important to clean the surface of the transilluminator after each use with deionized water and a soft cloth (e.g., cheesecloth); otherwise fluorescent dyes can accumulate on the glass surface and cause a high back-

ground fluorescence. Some fluorescent speckling may occur, especially near the edges of the gel. This speckling is an intrinsic property of the stain and does not affect sensitivity. When analyzing amounts of glycoprotein near the limit of detection, the authors advise that samples be run in the middle lanes of the gel. The authors use a 300-nm transilluminator with six 15-W bulbs. Excitation with different light sources may not give the same sensitivity. Using a Polaroid camera and Polaroid 667 black-and-white print film, the highest sensitivity is achieved with a 490-nm long-pass filter, such as the SYPRO protein gel stain photographic filter, available from Molecular Probes. The authors typically photograph minigels using an *f*-stop of 4.5 for 2 to 4 sec, using multiple 1-sec exposures. Using a CCD camera, images are best obtained by digitizing at $\sim 1024 \times 1024$ pixels resolution with 12-, 14-, or 16-bit grayscale levels per pixel. The camera manufacturer should be consulted for recommendations on filters to use. A CCD camera-based image-analysis system can gather quantitative information that will allow comparison of fluorescence intensities between different bands or spots. The polyester backing on some precast gels is highly fluorescent. For maximum sensitivity using a UV transilluminator, the gel should be placed polyacrylamide-side-down and an emission filter should be used to screen out the blue fluorescence of the plastic.

SYPRO Ruby staining for total protein is described in Basic Protocol 4. Viewing and photographing SYPRO Ruby-stained protein gels is described in Support Protocol 2. One should keep in mind that SYPRO Ruby protein gel stain has two excitation peaks and can be viewed using either UV illumination or blue-light illumination with a laser scanner. With UV illumination, both stains can be visualized simultaneously, although the signal from the green fluorescent glycoprotein stain may be somewhat reduced, compared to what it was before SYPRO Ruby staining. For documentation, the orange-red fluorescent SYPRO Ruby staining can be separated from the green-fluorescent Pro-Q Emerald 300 staining in one of two ways. If using UV illumination, use either a long-pass filter with a cutoff between 620 and 650 nm, or a band-pass filter with a center wavelength at about 645 nm, to document the SYPRO Ruby stain alone. Filters with cutoffs at wavelengths shorter than 620 nm may show some bleed-through of the Pro-Q Emerald 300 signal. Alternatively, the gel can be imaged using visible-light excitation, such as used in a laser scanner. Visible light will excite the SYPRO Ruby stain, but not the Pro-Q Emerald 300 stain. The fluorescent signal from the SYPRO Ruby stain can then be documented as described.

REAGENTS AND SOLUTIONS

Use high-quality deionized, distilled water (≥ 18 M Ω) in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Carbonate developing solution

- 0.5 ml 37% formaldehyde per liter solution
- 3% (w/v) sodium carbonate
- Prepare fresh before use

Coomassie blue staining solution

- 50% (v/v) methanol
- 0.05% (w/v) Coomassie brilliant blue R-250 (Bio-Rad or Pierce)
- 10% (v/v) acetic acid
- 40% H₂O

Dissolve Coomassie brilliant blue R in methanol before adding acetic acid and water. Store for up to 6 months at room temperature. If precipitate is observed following prolonged storage, filter to obtain a homogeneous solution.

Developing solution

0.5 g sodium citrate
0.5 ml 37% formaldehyde solution (Kodak)
H₂O to 100 ml
Store up to ~1 month at room temperature

Drying solution

10% (v/v) ethanol
4% (v/v) glycerol
86% H₂O
Store up to ~1 month at room temperature

Fixing solution for Coomassie blue and silver staining

50% (v/v) methanol
10% (v/v) acetic acid
40% H₂O
Store up to ~1 month at room temperature

Fixing solution for phosphoprotein gels

For minigels:

50% (v/v) methanol
10% (v/v) acetic acid

For 2-D gels:

50% (v/v) methanol
10% (w/v) trichloroacetic acid

One 6-cm × 9-cm × 0.75-mm minigel will require ~200 ml of fix solution; one 20-cm × 20-cm × 1-mm 2-D gel will require ~1 liter of fix solution. In order to improve the specificity of phosphoprotein staining in 2-D gels, the authors recommend fixing them in 10% trichloroacetic acid/50% methanol instead of 10% acetic acid/50% methanol.

Store fixing solutions up to 1 month at room temperature

Fixing solution for SYPRO Ruby staining of 2-D polyacrylamide gels

Fix with any of the following solutions (all prepared in H₂O):

10% (v/v) methanol
7% (v/v) acetic acid/25% (v/v) ethanol
12.5% (w/v) trichloroacetic acid/10% (v/v) ethanol
7% (v/v) acetic acid/50% (v/v) ethanol
3% (v/v) acetic acid/40% (v/v) ethanol
10% (v/v) acetic acid

Store up to 1 month at room temperature

These fixative solutions are used in 5- to 10-fold excess of the gel volume. Thin (<0.75-mm) gels will equilibrate with fixative much more quickly than thick gels. Larger-format gels require more fixative to effectively exchange the gel buffers for the fixative.

Formaldehyde fixing solution

40% (v/v) methanol
0.5 ml 37% formaldehyde per liter solution
60% H₂O
Store up to ~1 month at room temperature

Isopropanol fixing solution

25% (v/v) isopropanol
10% (v/v) acetic acid
65% H₂O
Store indefinitely at room temperature

Methanol/acetic acid destaining solution

5% (v/v) methanol

7% (v/v) acetic acid

88% H₂O

Store up to ~1 month at room temperature

Phosphoprotein gel destain solution

15% (v/v) 1,2-propanediol (propylene glycol)

50 mM sodium acetate, pH 4.0

For each liter of destain solution to be prepared, add 50 ml of 1 M sodium acetate, pH 4.0, to 800 ml of distilled water. Add 150 ml of 1,2-propanediol (or 40 ml of acetonitrile; see below). Bring the volume to 1 liter with distilled water and mix thoroughly. Alternatively, this destain Solution is available from Molecular Probes as a separate product. Store up to 1 month at room temperature.

If 1,2-propanediol (propylene glycol) is not available, 4% acetonitrile may be substituted. However, note that acetonitrile is a hazardous compound and its use will involve waste disposal restrictions.

The destaining step is very important for maximizing detection sensitivity for phosphoproteins while minimizing nonspecific staining of nonphosphorylated proteins. One 6-cm × 9-cm × 0.75-mm minigel will require ~250 ml of destain solution; one 20-cm × 20-cm × 1-mm 2-D gel will require ~2 liters of destain solution.

Pro-Q Diamond phosphoprotein gel stain

Purchase Pro-Q Diamond phosphoprotein gel stain from Molecular Probes (available in 200-ml, 1-liter, and 5-liter quantities). 200 ml provides sufficient material to stain ~4 minigels; 1 liter provides sufficient material to stain ~20 minigels or two large-format 2-D gels; 5 liters provide sufficient material to stain ~100 minigels or 10 large-format gels. Upon receipt, store the stain at room temperature, protected from light. For long-term storage, store the stain at 2° to 6°C, protected from light. When stored properly, the stain should be stable for at least 6 months.

Rapid Coomassie blue staining solution

10% (v/v) acetic acid

0.006% (w/v) Coomassie brilliant blue G-250 (Bio-Rad)

90% H₂O

Store indefinitely at room temperature

Silver nitrate solution (ammoniacal)

Add 3.5 ml concentrated NH₄OH (~30%) to 42 ml 0.36% NaOH and bring the volume to 200 ml with H₂O. Mix with a magnetic stirrer and slowly add 8 ml of 19.4% (1.6 g/8 ml) silver nitrate. Use within 20 min.

If the solution is cloudy, carefully add NH₄OH until it clears. Alternatively, use NH₄OH that is <3 months old.

CAUTION: *This solution is potentially explosive when dry and therefore should be precipitated by the addition of an equal volume of 1 M HCl. The resultant silver chloride can be washed down a drain with a large volume of cold water.*

SYPRO Orange or Red fluorescent staining solution

Allow stock vial of SYPRO Orange or Red protein gel stain (Molecular Probes) to warm to room temperature and then briefly microcentrifuge to deposit the dimethyl sulfoxide (DMSO) solution at the bottom of the vial. If particles of dye are present, dissolve by briefly sonicating the tube or vortexing it vigorously after warming. Dilute stock 1:5000 (v/v) in 7.5% (v/v) acetic acid and mix vigorously. Store in very

continued

clean, detergent-free glass or plastic bottles, protected from light, at 4°C (stable ≥3 months).

SYPRO Orange: 300 and 470 nm excitation, 570 nm emission; SYPRO Red: 300 and 550 nm excitation, 630 nm emission.

The stock solutions should be stored protected from light at room temperature, 4°C, or 20°C. When stored properly, they are stable for 6 months to 1 year.

SYPRO Ruby protein gel stain

Purchase SYPRO Ruby protein gel stain from Molecular Probes (available in 200-ml, 1-liter, and 5-liter quantities; 200 ml provides sufficient material to stain ~4 minigels). Store at room temperature protected from light (stable for at least 9 months). For convenient storage and dispensing, the 5-liter unit size is packaged in a cubical box with a spigot. Once opened, the box can be stored on its side with the top flap closed to protect the stain from light.

Thiosulfate developing solution

3% (w/v) sodium carbonate

0.0004% (w/v) sodium thiosulfate

0.5 ml 37% formaldehyde per liter solution (add immediately before use)

Store indefinitely without formaldehyde at room temperature

COMMENTARY

Background Information

Coomassie brilliant blue (Basic Protocol 1) binds nonspecifically to proteins (Wilson, 1983). Because the dye does not bind to the polyacrylamide gel, proteins will be detected as blue bands surrounded by clear gel zones.

Silver staining relies on differential reduction of silver ions, which is the basis for photographic processes. A highly sensitive photochemical silver staining technique (Switzer et al., 1979; Merrill et al., 1984) permits the detection of polypeptides in gels at more than 100× lower concentrations than Coomassie brilliant blue (i.e., femtomole levels of protein). The basic silver staining protocol described here (Basic Protocol 2) is derived from a modified technique developed by Oakley et al. (1980), which is simpler and less expensive than the original procedures. The first alternate silver staining protocol presents a very popular method described by Morrissey (1981). The second alternate silver staining protocol is a very rapid method described by Bloom et al. (1987).

Fluorescent protein gel stains provide a number of advantages over conventional colorimetric stains. The SYPRO Orange and Red protein gel stains described here can detect 1 to 2 ng protein per minigel band, more sensitive than Coomassie brilliant blue staining and as sensitive as many silver staining techniques. In addition, staining is complete in <1 hr. After electrophoresis, the gel is simply stained,

rinsed, and photographed; no separate fixation or destaining step is required and there is no fear of overstaining the gel (Steinberg et al., 1996a,b, 1997). In addition, stained proteins can be visualized using a standard 300-nm UV transilluminator or a laser scanner (Fig. 10.6.2). Because the dyes interact with the SDS coat around proteins in the gel, they give more consistent staining between different types of proteins compared to Coomassie or silver staining and do not exhibit negative staining. Furthermore, the dyes detect a variety of proteins down to ~6500 Da without staining nucleic acid or lipopolysaccharide contaminants that are sometimes found in protein preparations derived from cell or tissue extracts.

Critical Parameters

The high sensitivity of the silver staining technique renders it susceptible to impurities and staining artifacts. It is mandatory that the polyacrylamide gels and all staining solutions be prepared from high-quality reagents in order to avoid staining artifacts. Especially important is the use of high-quality water (glass-distilled or deionized, carbon-filtered). The glassware used for gel polymerization and plastic containers should be cleaned thoroughly, and gels should be handled with vinyl, powder-free gloves. To avoid uneven staining of the gel surface, the polyacrylamide gel should be covered with a sheet of Parafilm in order to uniformly wet the gel surface during staining, and

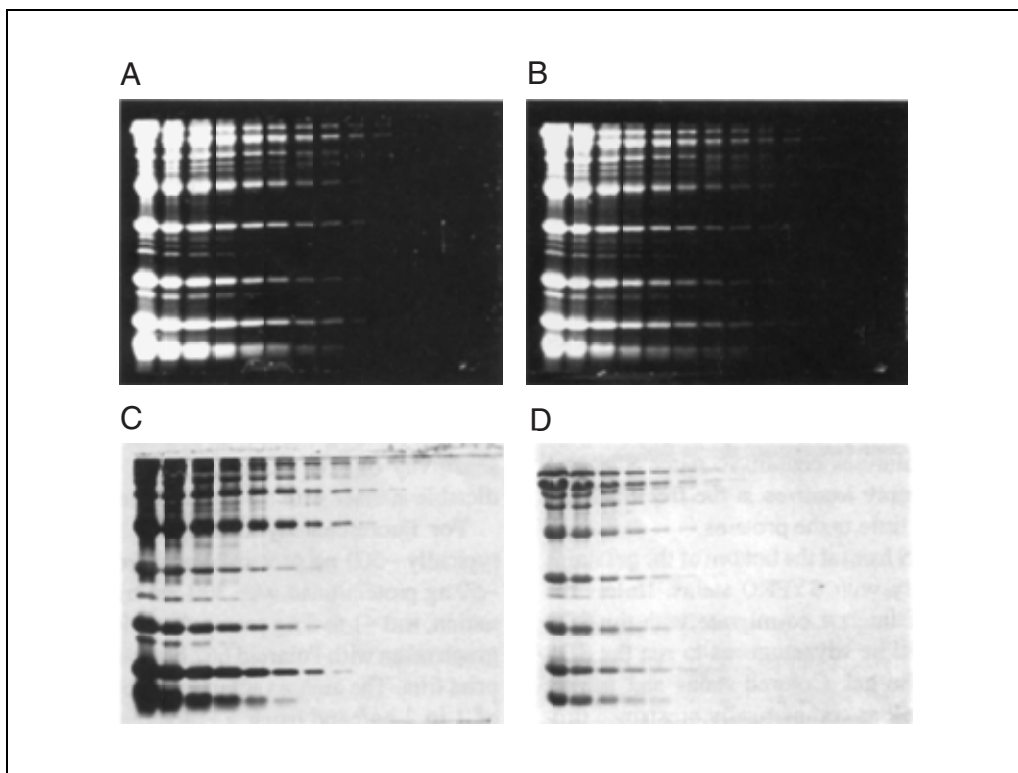


Figure 10.6.2 Identical polyacrylamide minigels stained with (A) SYPRO Orange gel stain, (B) SYPRO Red gel stain, (C) silver stain, and (D) Coomassie brilliant blue stain according to standard protocols. The SYPRO-stained gels were photographed using 300-nm transillumination, a SYPRO Orange/Red protein gel stain photographic filter, and Polaroid 667 black-and-white print film. The Coomassie- and silver-stained gels were photographed using transmitted white light and Polaroid 667 black-and-white print film; no optical filter was used.

should be touched only very gently with gloved hands. If silver staining is performed infrequently, commercial silver staining kits should be used; those distributed by Bio-Rad and Pierce have been tested and found to be reliable and sensitive.

For immunoblotting and other blotting techniques, fluorescent stains can be diluted in standard transfer buffer. However, staining the gel in transfer buffer will result in lower sensitivity. Therefore, for blotting techniques, staining the gel with SYPRO Tangerine protein gel stain, which does not require acetic acid fixation, or staining the blot directly with SYPRO Ruby protein blot stain is recommended.

Diluting fluorescent stain below the recommended concentration will result in reduced staining sensitivity. Using higher staining concentrations than recommended will not result in better detection, but will instead result in increased background and quenching of the fluorescence from dye molecules crowded around the proteins.

SYPRO Red and Orange stains cannot be used to pre-stain protein samples for SDS gels. Loading solutions contain so much SDS that

the dye simply localizes in the free SDS and binds very little to the proteins.

The SDS front at the bottom of the gel stains very heavily with SYPRO stains. Unless the proteins of interest co-migrate with the SDS front, it will be advantageous to run the SDS front off the gel. Colored stains and marker dyes, as well as commercially prestained protein markers, interfere with SYPRO dye staining and quench fluorescence.

Highly colored prosthetic groups (e.g., heme) that remain bound in native gels will quench fluorescence of the SYPRO Orange and Red stains. Odd marks on stained gels can be caused by several factors. If the gel is squeezed, a mark appears that stains heavily with the SYPRO dyes. This is probably due to a localized high concentration of SDS that has difficulty diffusing out. Glove powder can also give background markings, so rinsing or washing gloves is recommended prior to handling gels. Staining with the SYPRO Orange dye occasionally results in gels with scattered fluorescent speckles.

Due to different staining properties of proteins, dual staining procedures can reveal pro-

teins with one procedure that the other has not visualized. SYPRO-stained gels can be restained with either Coomassie brilliant blue or silver stain procedures. In fact, for some silver staining methods, the authors have found that prestaining with SYPRO dyes actually increases the rate of staining and the sensitivity for detection. To fluorescently stain gels that have previously been stained with Coomassie, the Coomassie stain must be completely removed, as it will quench the fluorescence of SYPRO dyes. Soaking the gel in either 30% methanol or 7.5% acetic acid with several changes of destaining solution is effective at removing Coomassie. Once the Coomassie has been removed, the gel should be incubated in 0.05% SDS for 30 min before staining with the SYPRO stain as usual.

Triton X-100 at $\geq 0.1\%$ interferes with SYPRO dye staining. If Triton X-100 is used in the gel, the authors recommend soaking the gel in two to three changes of buffer to be sure the Triton X-100 is diluted out, and then incubating the gel in 0.05% SDS for 30 min before staining as usual.

Staining for glycoproteins

The overall specificity of glycoprotein detection by the Pro-Q Emerald 300 reagent method depends greatly upon the specificity of the oxidation reaction, which is governed in turn by the reaction conditions used (e.g., periodic acid concentration, pH, temperature, and exposure to light). Careful attention to the protocol is required to avoid oxidation of serine, threonine, and hydroxylysine residues to form aldehyde groups, which result in false-positive signals. Residual SDS in the gel will also lead to nonspecific staining. This can be avoided by adhering strictly to the fixation and wash volumes and times indicated in Alternate Protocol 6. In addition, it is advisable to stain a duplicate gel, eliminating the oxidation step (step 4), as a negative control.

Anticipated Results

The sensitivity of Coomassie blue gel staining is 0.3 to 1 μg /protein band; the sensitivity of silver staining is 2 to 5 ng/protein band. The sensitivity of both stains varies in an unpredictable manner with the protein being stained.

For fluorescent dyes, detection limits are typically ~ 500 ng protein/band in room light, ~ 50 ng protein/band with 300-nm transillumination, and ~ 1 to 2 ng protein/band in a photograph taken with Polaroid 667 black-and-white print film. The authors achieve detection limits

of 1 to 2 ng/band using a Fotodyne Foto/UV 450 ultraviolet transilluminator, which has six 15-watt bulbs that provide peak illumination at 312 nm. When using weaker illumination sources, exposures must be correspondingly longer. Although the authors' detection limits are 1 to 2 ng/band for most proteins, it should be emphasized that bands containing 5 to 10 ng/protein are more readily detected. Bands containing less than 5 to 10 ng protein require longer exposures and sharp bands for good visualization. Longer exposures can result in higher background.

Time Considerations

Coomassie blue staining requires 8 to 12 hr. Silver staining requires ~ 5 hr. Fixation may be extended for several days before Coomassie blue staining. Fixation may be extended for longer periods—up to several weeks—before silver staining.

Use of either of the rapid staining protocols considerably reduces the time required to visualize proteins. Detection of protein bands by rapid Coomassie blue staining requires ≤ 90 min from the time a minigel is run (30 to 60 min) until the gel is fixed (10 min) and placed in staining solution (5 to 10 min); however, additional time may be necessary for larger gels. Separated proteins stained with the rapid silver stain method can be visualized in ~ 35 min.

The staining time for SYPRO dyes is 10 to 60 min, depending on the thickness and percentage of the gel. For 1-mm-thick 15% polyacrylamide gels, the signal is typically optimal at 40 to 60 min of staining.

Literature Cited

- Anderson, N.L. 1988. Two Dimensional Electrophoresis Operation of the ISO-DALT (R) System; Large Scale Biology Press, Washington, D.C.
- Bloom, H., Beier, H., and Gross, H.S. 1987. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* 8:93-99.
- Merril, C.R., Goldman, D., and Van Keuren, M.L. 1984. Gel protein stains: Silver stain. *Methods Enzymol.* 104:441-447.
- Morrissey, J.H. 1981. Silver stain for proteins in polyacrylamide gels: A modified procedure with enhanced uniform sensitivity. *Anal. Biochem.* 117:307-310.
- Oakley, B.R., Kirsch, D.R., and Morris, N.R. 1980. A simplified ultrasensitive silver stain for detecting proteins in polyacrylamide gels. *Anal. Biochem.* 105:361-363.
- Steinberg, T.H., Haugland, R.P., and Singer, V.L. 1996a. Applications of SYPRO orange and

- SYPRO red protein gel stains. *Anal. Biochem.* 239:238-245.
- Steinberg, T.H., Jones, L.J., Haugland, R.P., and Singer, V.L. 1996b. SYPRO orange and SYPRO red protein gel stains: One-step fluorescent staining of denaturing gels for detection of nanogram levels of protein. *Anal. Biochem.* 239:223-237.
- Steinberg, T.H., White, H.M., and Singer, V.L. 1997. Optimal filter combinations for photographing SYPRO orange or SYPRO red dye-stained gels. *Anal. Biochem.* 248:168-172.
- Switzer, R.C., Merrill, C.R., and Shifrin, S. 1979. A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. *Anal. Biochem.* 98:231-237.
- Wilson, C.M. 1983. Staining of proteins on gels: Comparison of dyes and procedures. *Methods Enzymol.* 91:236-247.

Contributed by Joachim Sasse
Shriners Hospital for Crippled Children
Tampa, Florida

Sean R. Gallagher
UVP, Inc.
Upland, California

Detection of Proteins on Blot Transfer Membranes

This unit presents protocols for the staining of proteins that have been electroblotted from polyacrylamide gels to blot transfer membranes (UNIT 10.8). If the samples of interest are electrophoresed in duplicate and transferred to a membrane, half the membrane can be stained to determine the efficiency of transfer to the membrane and the other half can be used for immunoblotting (i.e., western blotting; see UNIT 10.8). Procedures are described for India ink staining (see Basic Protocol 1) and gold staining of the blots (see Alternate Protocol). Support Protocol 1 describes a method for alkali treatment that enhances subsequent staining of bound proteins. Detection limits of each staining method and compatible blot transfer membranes and gels are presented in Table 10.7.1.

Fluorescent staining methods have gained popularity because of their high sensitivity and direct compatibility with colorimetric, fluorogenic, and chemiluminescent techniques. A method is described (see Basic Protocol 2) for staining blots with SYPRO Ruby, a sensitive fluorescent stain that works equally well on both nitrocellulose and PVDF blots. Support Protocol 2 provides detail for photographing fluorescently stained blots and recording the results.

Table 10.7.1 Properties and Compatibilities of “Protein Transfer” Membrane Stains^a

Stain	Detection limit (ng)	Membrane types			Gel types	
		Nitro-cellulose	Nylon	PVDF	SDS-PAGE	Native PAGE
India ink	50	+	–	+	+	+
Gold	3	+	–	+	+	+
SYPRO Ruby	2	+	+	+	+	+

^aAbbreviations: n.d., not determined; PVDF, polyvinylidene difluoride; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

INDIA INK STAINING

India ink is used to stain proteins on blot transfer membranes. The transferred proteins (≥ 50 ng) appear as black bands on a gray background.

Materials

Protein sample transferred onto blot transfer membrane, nitrocellulose or PVDF (UNIT 10.8)
Tween 20 solution (see recipe)
India ink solution (see recipe)
Plastic boxes

NOTE: Deionized, distilled water should be used throughout this protocol.

1. If desired, pretreat proteins transferred onto nitrocellulose membrane with alkali (see Support Protocol 1).
2. Place blot transfer membrane(s) in a plastic box on an orbital shaker and wash in Tween 20 solution at 37°C three times for 30 min each.

**ALTERNATE
PROTOCOL**

3. Continue to wash the membrane in Tween 20 solution at room temperature two times for 30 min each.
4. Stain the membrane in India ink solution 3 hr or overnight.
5. Rinse the membrane twice in Tween 20 solution, destain in Tween 20 solution until an acceptable background is obtained, then air dry.

Black bands appear against a gray background.

GOLD STAINING

A colloidal gold sol is used to stain proteins on blot transfer membranes. The transferred proteins (≥ 3 ng) will appear as red bands on an almost white background.

Materials

Membragold colloidal gold solution (Diversified Biotech)
Heat-sealable plastic bag

1. If desired, pretreat proteins transferred onto nitrocellulose membrane with alkali (see Support Protocol 1).
2. Place blot transfer membrane(s) in a plastic box on an orbital shaker and wash in Tween 20 solution three times at 37°C for 30 min each.

Do not attempt to stain nylon membranes.

3. Continue to wash the membrane in Tween 20 solution at room temperature three times for 5 min each.
4. Rinse well with water.
5. Stain the membrane 30 min to 1 hr at room temperature with continuous shaking in Membragold colloidal gold sol.

A heat-sealed plastic bag is a convenient container for the staining. Extended staining periods do not impair visualization.

6. Rinse the membrane briefly in water and air dry on filter paper.

**SUPPORT
PROTOCOL 1**

ALKALI ENHANCEMENT OF PROTEIN STAINING

A brief pretreatment of nitrocellulose-bound protein with alkali enhances subsequent staining with either India ink or colloidal gold.

Materials

1% (w/v) KOH
PBS (APPENDIX 2)
Glass or Pyrex dish

1. Place nitrocellulose plot or dot membrane with transferred proteins in a glass dish containing enough 1% KOH to cover the membranes. Soak 5 min at 20°C.
2. Rinse twice each time for 10 min, in PBS.

FLUORESCENT PROTEIN BLOT STAINING

The fluorescent SYPRO Ruby protein blot stain provides a rapid, simple, and highly sensitive method for detecting proteins on nitrocellulose or polyvinylidene difluoride (PVDF) membranes (blots). Staining total protein before applying specific protein detection techniques provides an assessment of protein transfer efficiency, and makes it possible to detect contaminating proteins in the sample and to compare the sample with molecular weight standards. For blots of 2-D gels, total protein staining makes it easier to localize a protein to a particular spot in the complex protein pattern. The bright, orange-red fluorescent stain can be easily visualized using UV illumination or a laser scanner. The staining procedure is simple to perform and can be completed within 1 hr. SYPRO Ruby has a sensitivity limit of 2 to 8 ng/band, making it about 60 times more sensitive than reversible stains like Ponceau S and 20 to 30 times more sensitive than Amido Black or Coomassie Brilliant Blue stains. The SYPRO Ruby protein blot stain will not stain nucleic acids and is compatible with immunodetection and colorimetric, fluorogenic, and chemiluminescent detection techniques, as well as with Edman sequencing and mass spectrometry.

Materials

Proteins electroblotted onto PVDF or nitrocellulose membrane (*UNIT 10.8*)
7% acetic acid/10% methanol
SYPRO Ruby protein blot stain (purchase from Molecular Probes; also see recipe)
Small polypropylene staining dish
Orbital shaker

NOTE: Perform all washing, staining, and other incubation steps with continuous, gentle agitation (e.g., on an orbital shaker at 50 rpm). For PVDF membranes, be sure to float the membrane face down on the solution.

After electroblotting to nitrocellulose membranes

- 1a. Completely immerse the membrane in 7% acetic acid/10% methanol and incubate at room temperature for 15 min in a small polypropylene staining dish.
- 2a. Incubate the membrane in four changes of deionized water, each time for 5 min.
- 3a. Completely immerse the membrane in SYPRO Ruby protein blot stain for 15 min.
- 4a. Wash the membrane 4 to 6 times, for 1 min each time, in deionized water, to remove excess dye.

Membranes stained with SYPRO Ruby blot stain should be periodically monitored using UV epi-illumination to determine if background fluorescence has been washed away.

After electroblotting to PVDF membranes

- 1b. Allow the membrane to dry completely.
- 2b. Float the membrane face down in 7% acetic acid/10% methanol and incubate for 15 min.
- 3b. Float the membrane for 5 min each time in four changes of deionized water, then float the membrane in SYPRO Ruby protein blot stain for 15 min.
- 4b. Wash the membrane 2 to 3 times in deionized water, each time for 1 min, to remove excess dye.

For Edman-based microsequencing, use a PVDF membrane, stain with SYPRO Ruby protein blot stain as described above, and then partially destain the blot, first washing it by placing it face down on a solution of 150 mM Tris·Cl, pH 8.8 (APPENDIX 2)/20% methanol for 10 min, with gentle agitation, then rinsing the blot four times, each time for 1 min in distilled water. Finally, air dry the blot.

5. Allow the membranes to air dry.

After staining, wet membranes should not be touched, since residue found on latex gloves may destroy the staining pattern. Use forceps to handle wet blots. Once dry, the membranes can be handled freely.

VIEWING AND PHOTOGRAPHING THE BLOT

SYPRO Ruby protein blot stain (also see Basic Protocol 2) has two excitation maxima, one at ~280 nm and one at ~450 nm, and has an emission maximum near 618 nm. Proteins stained with the dye can be visualized using a 300-nm UV transilluminator, a blue-light transilluminator, or a laser scanner. The stain has exceptional photostability, allowing long exposure times for maximum sensitivity.

UV Epi- or Transilluminator

Proteins stained with SYPRO Ruby protein blot stain are readily visualized using UV illumination. The front face of the membrane can be illuminated using a hand-held UV-B (~300 nm) light source. Alternatively, a UV light box can be placed on its side to illuminate the blots, or a top-illuminating system such as the UVP AutoChemi System or the Bio-Rad Fluor-S imager can be used to visualize the stain. Satisfactory results can also be obtained from direct UV transillumination through the blotting membrane. In either case, the use of a photographic or CCD camera and the appropriate filters is essential to obtain the greatest sensitivity. The instrument's integrating capability can make bands visible that cannot be detected by eye. It is important to clean the surface of the transilluminator after each use with deionized water and a soft cloth (like cheesecloth); otherwise, fluorescent dyes, such as SYPRO stains, SYBER stains, and ethidium bromide will accumulate on the glass surface and cause a high background fluorescence.

Using a Polaroid camera and Polaroid 667 black-and-white print film, the highest sensitivity is achieved with a 490-nm long-pass filter, such as the SYPRO protein gel stain photographic filter (cat. no. S-6656, Molecular Probes). In the authors' laboratory, blots are typically photographed using an f-stop of 4.5 for 1 sec.

Using a CCD camera, images are best obtained by digitizing at ~1024 × 1024-pixel resolution with 12- or 16-bit gray scale levels per pixel. The manufacturer of the particular camera should be contacted for recommendations on filter sets to use.

Laser-Scanning Instruments

Blots stained with the SYPRO Ruby protein blot stain can be visualized using imaging systems equipped with lasers that emit at 450, 473, 488, or 532 nm. Proteins may be analyzed immediately after staining by immunostaining, glycoprotein staining, or mass spectrometry. It is important to photograph or otherwise document the SYPRO Ruby stain before immunostaining, as over 90% of the stain is washed off the blot during the blocking step.

REAGENTS AND SOLUTIONS

Use deionized, distilled water or equivalent in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

India ink solution

0.1% India ink (e.g., Pelikan 17 black) in Tween 20 solution (see recipe)

SYPRO Ruby protein blot stain

SYPRO Ruby protein blot stain (Molecular Probes) is provided in a unit size of 200 ml. The 200-ml volume is sufficient for staining 10 to 40 minigel electroblots or four large-format electroblots (20 × 20 cm). SYPRO Ruby protein blot stain may

continued

be reused up to four times with little loss in sensitivity. The reagent is stable for at least 6 months to 1 year when stored at room temperature, protected from light.

Tween 20 solution

0.3% (v/v) Tween 20 in phosphate-buffered saline (PBS; *APPENDIX 2*), pH 7.4.

COMMENTARY

Background Information

The introduction of electrophoretic transfer of separated proteins has been paralleled by the use of extremely sensitive immunoblotting techniques to detect the binding of antibodies and lectins to the transferred protein bands. "On-blot" protein staining procedures help one control the electrophoretic transfer and correlate the transferred polypeptide pattern with that detected in the overlay. An additional advantage of on-blot protein staining as compared to gel staining is that the pattern of the protein membrane stains can be directly compared to the results of the overlay assays. Comparing overlay assays with stained gels is cumbersome due to swelling or shrinking of the polyacrylamide gels.

Membranes used for the electrophoretic transfer are manufactured from nitrocellulose, nylon, or polyvinylidene difluoride (PVDF). Depending on the transfer membrane used and other considerations (e.g., the use of SDS-polyacrylamide gels or polyacrylamide gels without SDS and the level of sensitivity required), different staining procedures must be selected (Table 10.7.1).

SYPRO Ruby, in contrast to India ink and colloidal gold, is fully compatible with colorimetric, fluorogenic, and chemiluminescent immunodetection techniques. Stains such as colloidal gold often block epitopes required for subsequent immunodetection. In addition, the dark color of the colloidal gold or India ink stains makes it difficult to visualize colorimetric or fluorogenic immunodetection reagents in subsequent immunodetection assays on the membrane.

Staining of membrane-bound polypeptides by gold sol is mediated by hydrophobic interactions and by ionic interaction of negatively charged gold particles with positive groups on the proteins. The protocol listed is that described by Moeremans et al. (1985). The gold stain is very easy to perform and is the most sensitive blot membrane stain.

The protocol for India ink staining of polypeptides is based on the procedure described by Hancock and Tsang (1983).

Brief exposure to alkali described in the support protocol significantly enhances staining of proteins by either India ink or colloidal gold. Sutherland and Skerritt (1986) suggest that the alkali treatment enhances retention of proteins on the nitrocellulose surface during extensive washing of the membrane; hence more protein is available for staining.

Critical Parameters

It is mandatory that all glassware and plasticware used with the gels and blot membranes be thoroughly cleaned in order to avoid staining artifacts. All blot membranes should be handled by forceps only. The stains cannot be used for detecting polypeptides blotted onto nylon membranes, since these membranes are positively charged. Furthermore, due to its very high sensitivity, the gold stain is susceptible to impurities present in buffers and on the surface of staining boxes.

Anticipated Results

Detection limits and compatibilities of transfer membranes and gel types are summarized in Table 10.7.1.

Time Considerations

The India ink staining procedure requires ~6 hr, while gold staining can be performed in ~3 hr. An additional 30 min is required at the beginning of either protocol to perform the alkali pretreatment. SYPRO fluorescent staining is completed in <1 hr.

Literature Cited

- Hancock, K. and Tsang, V.C.M. 1983. India ink staining of proteins on nitrocellulose paper. *Anal. Biochem.* 133:157-162.
- Moeremans, M., Daneels, G., and De Mey, J. 1985. Sensitive colloidal metal (gold or silver) staining of protein blots on nitrocellulose membranes. *Anal. Biochem.* 145:315-321.
- Sutherland, M.W., and Skerritt, J.H. 1986. Alkali enhancement of protein staining on nitrocellulose. *Electrophoresis* 7:401-406.

Key References

Moeremans et al., 1985. See above.

For investigators wishing to prepare their own reagents, this paper include descriptions of how to make gold and iron sols.

Patton, W.F. 2000. A thousand points of light: The application of fluorescence detection technologies to two-dimensional gel electrophoresis and proteomics. *Electrophoresis* 21:1123-1144.

Contributed by Joachim Sasse
Shriners Hospital for Crippled Children
Tampa, Florida

Sean R. Gallagher
UVP, Inc.
Upland, California

Immunoblotting and Immunodetection

UNIT 10.8

Immunoblotting (often referred to as western blotting) is used to identify specific antigens recognized by polyclonal or monoclonal antibodies. Protein samples are solubilized, usually with sodium dodecyl sulfate (SDS) and reducing agents such as dithiothreitol (DTT) or 2-mercaptoethanol (2-ME). Following solubilization, the material is separated by SDS-PAGE (UNITS 10.2-10.4). The antigens are then electrophoretically transferred in a tank (see Basic Protocol 1) or a semidry transfer apparatus (see Alternate Protocol 1) to a nitrocellulose, PVDF, or nylon membrane, a process that can be monitored by reversible staining (see Support Protocol 1) or by Ponceau S staining (see Support Protocol 2). Previously stained gels may also be blotted (see Alternate Protocol 2).

The transferred proteins are bound to the surface of the membrane, providing access for reaction with immunodetection reagents. All remaining binding sites are blocked by immersing the membrane in a solution containing either a protein or detergent blocking agent. After probing with the primary antibody, the membrane is washed and the antibody-antigen complexes are identified with horseradish peroxidase (HRPO) or alkaline phosphatase enzymes coupled to the secondary anti-IgG antibody (e.g., goat anti-rabbit IgG). The enzymes are attached directly (see Basic Protocol 2) or via an avidin-biotin bridge (see Alternate Protocol 3) to the secondary antibody. Chromogenic or luminescent substrates (see Basic Protocol 3 and Alternate Protocols 4 and 5) are then used to visualize the activity. Finally, membranes may be stripped and reprobed (see Support Protocol 3).

PROTEIN BLOTTING WITH TANK TRANSFER SYSTEMS

In this procedure, blotting is performed in a tank of buffer with the gel in a vertical orientation, completely submerged between two large electrode panels. In some systems up to four gels can be transferred at one time. For difficult-to-transfer proteins (>100 kDa or hydrophobic; e.g., myosin), tank blotting is preferable to semidry systems (see Basic Protocol 2) because prolonged transfers are possible without buffer depletion. However, transfers >1 hr at high power require cooling using a heat exchanger and a circulating water bath that can maintain a constant transfer temperature of 10° to 20°C.

Materials

Samples for analysis

Protein molecular weight standards (UNIT 10.2A): prestained (Sigma or Bio-Rad), biotinylated (Vector Labs or Sigma), fluorescent (e.g., BenchMark Fluorescent Protein Standards; Invitrogen), or compatible with other colorimetric and fluorescent detection method (e.g., MagicMark and MagicMark XP Western Protein standards; Invitrogen)

Transfer buffer (see recipe)

Powder-free gloves

Scotch-Brite pads (3M) or equivalent sponge

Whatman 3MM filter paper or equivalent

Transfer membrane: 0.45- μ m nitrocellulose (Millipore or Schleicher & Schuell), PVDF (Millipore Immobilon P), neutral nylon (Pall Biotodyne A), or positively charged nylon (Pall Biotodyne B; Bio-Rad Zetabind) membrane

Electroblotting apparatus (Bio-Rad, Invitrogen, Amersham, or Harvard Bioscience)

Indelible pen (e.g., Paper-Mate) or soft lead pencil

Additional reagents and equipment for one- and two-dimensional gel electrophoresis (UNITS 10.2-10.4) and staining proteins in gels (UNIT 10.6) and on membranes (see Support Protocol 1 and UNIT 10.7)

NOTE: Deionized, distilled water should be used throughout this protocol.

Contributed by Sean Gallagher, Scott E. Winston, Steven A. Fuller, and John G.R. Hurrell

Current Protocols in Molecular Biology (2004) 10.8.1-10.8.24

Copyright © 2004 by John Wiley & Sons, Inc.

**BASIC
PROTOCOL 1**

**Analysis of
Proteins**

10.8.1

Supplement 66

Electrophorese the protein sample

1. Prepare antigenic samples and separate proteins using small or standard-sized one- or two-dimensional gels (UNITS 10.2-10.4). Include prestained or biotinylated protein molecular weight standards in one or more gel lanes.

The protein markers will transfer to the membrane and conveniently indicate membrane orientation and sizes of proteins after immunostaining.

MagicMark Western Protein standards allow direct visualization of protein size standards on membrane blots without the need for protein modification or special detection reagents. The standard proteins are derived from E. coli cells containing a construct with repetitive units of a fusion protein forming the size variation and an IgG binding site. The proteins can be visualized with the colorimetric, chemiluminescent, or fluorescent detection system of choice simply by processing the membrane for the specific protein. The IgG binding site will allow all the standard proteins to react with the specific primary and secondary antibodies. These protein standards do not have to be heated or reduced—they are in a ready-to-use format. Alternatively, BenchMark Fluorescent Protein Standards are visualized directly via UV transillumination on a UV transilluminator (available from UVP Inc.) when wet, or via overhead UV illumination (apparatus also available from UVP, Inc.) when dry.

A variety of gel sizes and percentages of acrylamide can be used (UNIT 10.2A). Most routinely used are either 14 cm × 14 cm × 0.75-mm gels or 8 cm × 10 cm × 0.75-mm minigels. Acrylamide concentrations vary from 5% to 20%, but are usually in the 10% to 15% range.

Assemble the immunoblot sandwich

2. When electrophoresis is complete, disassemble gel sandwich and remove stacking gel. Equilibrate gel 30 min at room temperature in transfer buffer.

IMPORTANT NOTE: Use gloves when manipulating filter papers, gels, and membranes. Oil from hands blocks the transfer.

Match the appropriate transfer buffer to the membrane (see Reagents and Solutions).

Gel equilibration is required to prevent a change in the size of the gel during transfer. Any shift in gel dimension will result in a blurred transfer pattern.

3. Assemble transfer sandwich in a tray large enough to hold the plastic transfer cassette. Fill with transfer buffer so that cassette is covered.

The transfer cassette should be assembled under buffer to minimize trapping of air bubbles. Use Figure 10.8.1 as a guide to assembly.

4. On bottom half of plastic transfer cassette, place Scotch-Brite pad or sponge, followed by a sheet of filter paper cut to same size as gel and prewet with transfer buffer.

5. Place gel on top of filter paper; the side of the gel touching the paper arbitrarily becomes the cathode side of the gel (i.e., ultimately toward the negative electrode when positioned in the tank). Remove any air bubbles between gel and filter paper by gently rolling a test tube or glass rod over surface of gel.

Any bubbles between the filter paper, gel, and membrane will block current flow and prevent protein transfer. This problem is indicated on the membrane by sharply defined white areas devoid of transferred protein.

Proteins are usually negatively charged in transfer buffer and move toward the positive anode. However, some proteins may be positively charged. An additional membrane placed on the cathode side of the gel will bind these proteins.

6. Prepare transfer membrane. Cut membrane to same size as gel plus 1 to 2 mm on each edge. Place into distilled water slowly, with one edge at a 45° angle; the water will wick up into the membrane, wetting the entire surface (if it is inserted too quickly into the water, air gets trapped and will appear as white blotches in the membrane; protein will not transfer onto these areas). Equilibrate 10 to 15 min in transfer buffer.

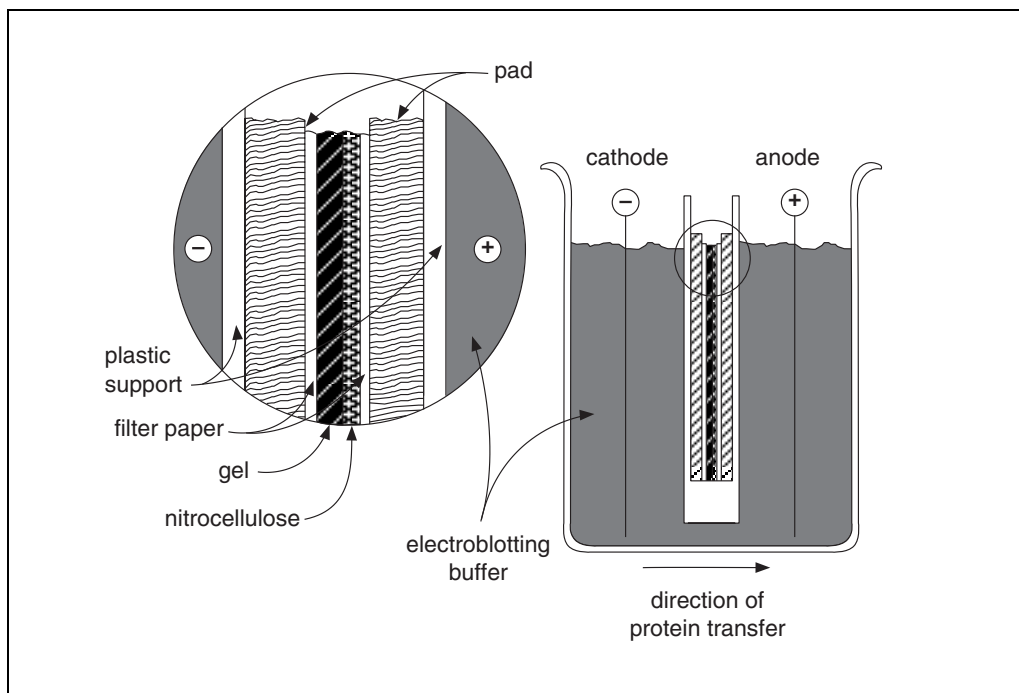


Figure 10.8.1 Immunoblotting with a tank blotting unit. The polyacrylamide gel containing the protein is laid on a sheet of filter paper. The uncovered side of the gel is overlaid with a sheet of membrane precut to the size of the gel plus 1 to 2 mm on each edge, then this membrane is overlaid with another sheet of filter paper. The filter paper containing the gel and membrane is sandwiched between Scotch-Brite pads. This sandwich is placed in a plastic support, and the entire assembly is placed in a tank containing transfer buffer. For transfer of negatively charged protein, the membrane is positioned on the anode side of the gel. For transfer of positively charged protein, the membrane is placed on the cathode side of the gel. Charged proteins are transferred electrophoretically from the gel onto the membrane. Transfer is achieved by applying a voltage of 100 V for 1 to 2 hr (with cooling) or 14 V overnight.

This wetting procedure works for nitrocellulose and nylon membranes only. PVDF membranes are hydrophobic and will not wet simply from being placed into distilled water or transfer buffer. For these membranes, first immerse 1 to 2 sec in 100% methanol, then equilibrate 10 to 15 min with transfer buffer. Do not let membrane dry out at any time. If this occurs, wet once again with methanol and transfer buffer as described above.

7. Moisten surface of gel with transfer buffer. Place prewetted membrane directly on top side of gel (i.e., anode side) and remove all air bubbles as in step 5.

Poor contact between the gel and membrane will cause a swirled pattern of transferred proteins on the membrane. Some proteins will transfer as soon as the gel is placed on the membrane; repositioning the gel or membrane can result in a smeared or double image on the developed blot.

The use of 0.2- μ m membranes may improve retention of smaller-molecular-weight proteins.

8. Wet another piece of Whatman 3MM filter paper, place on anode side of membrane, and remove all air bubbles. Place another Scotch-Brite pad or sponge on top of this filter paper.
9. Complete assembly by locking top half of the transfer cassette into place (Fig. 10.8.1).

It is important to orient the sandwich so that the membrane faces the anode (positively charged) side of the tank.

Transfer proteins from gel to membrane

10. Fill tank with transfer buffer and place transfer cassette containing sandwich into electroblotting apparatus in correct orientation. Connect leads of power supply to corresponding anode and cathode sides of electroblotting apparatus.

Transfer buffer should cover the electrode panels but should not touch the base of the banana plug.

11. Electrophoretically transfer proteins from gel to membrane for 30 min to 1 hr at 100 V with cooling or overnight at 14 V (constant voltage), in a cold room.

Transfer time is dependent on the thickness and the percent acrylamide of the gel, as well as the size of the protein being transferred. In general, proteins are transferred within 1 to 6 hr, but high-molecular-weight molecules may take longer. Overnight transfer at low voltage is reliable and convenient. Cooling (at 10° to 20°C) is required for transfers >1 hr at high power. Heat exchanger cooling cores using a circulating water bath are placed into the transfer unit for cooling.

12. Turn off the power supply and disassemble the apparatus. Remove membrane from blotting apparatus and note orientation by cutting a corner or marking with a soft lead pencil or Paper-Mate pen.

Many ballpoint inks come off, but Paper-Mate stays on the membrane.

Membranes can be dried and stored in resealable plastic bags at 4°C for 1 year or longer at this point. Prior to further processing, dried PVDF membranes must be placed into a small amount of 100% methanol to wet the membrane, then in distilled water to remove the methanol.

13. Stain gel for total protein with Coomassie blue (UNIT 10.6) to verify transfer efficiency. If desired, stain membrane reversibly to visualize transferred proteins (see Support Protocol 1), or irreversibly with Coomassie blue, India ink, naphthol blue, or colloidal gold (UNIT 10.7).

These staining procedures are incompatible with nylon membranes.

If membrane shows significant staining on the backside, either the gel was heavily overloaded or the membrane has poor protein-binding capacity (see Troubleshooting). In either case, protein-binding sites on the side facing the gel are saturated, allowing protein to migrate to the other side of the membrane. Nitrocellulose, in particular, will show diminished binding capacity with age or poor storage conditions (e.g., high temperature and humidity). In addition, some proteins simply do not bind well to a particular matrix. By using several membrane sheets in place of one, the protein can be detected as it passes through each consecutive sheet. This will give an indication of how efficiently the membrane binds to a particular protein.

14. Proceed with immunoprobng and visual detection of proteins (see Basic Protocols 2 and 3 and Alternate Protocols 3 and 4).

ALTERNATE PROTOCOL 1

PROTEIN BLOTTING WITH SEMIDRY SYSTEMS

Even and efficient transfer of most proteins is also possible with semidry blotting, a convenient alternative to tank transfer systems. Instead of being placed vertically into a tank filled with transfer buffer, the gel is held horizontally between buffer-saturated blotting paper that is in contact with the electrodes (Fig. 10.8.2), greatly reducing the amount of buffer required. The electrodes are close together, giving high field strengths and rapid transfer with a standard electrophoresis power supply. Prolonged transfers (>1 hr) are not recommended; tank blotting (see Basic Protocol 1) should be used for proteins that require long blotting times for efficient transfer.

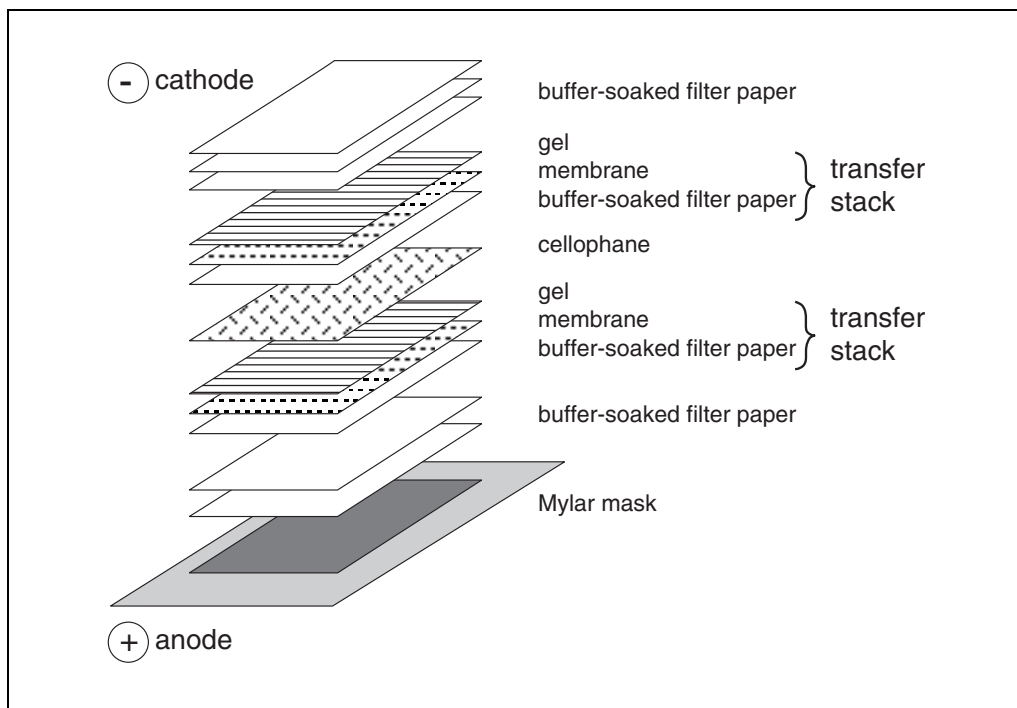


Figure 10.8.2 Immunoblotting with a semidry transfer unit. Generally, the lower electrode is the anode, and one gel is transferred at a time. A Mylar mask (optional in some units) is put in place on the anode. This is followed by three sheets of transfer buffer-soaked filter paper, the membrane, the gel, and finally, three more sheets of buffer-soaked filter paper. To transfer multiple gels, construct transfer stacks as illustrated, and separate each with a sheet of porous cellophane. For transfer of negatively charged protein, the membrane is positioned on the anode side of the gel. For transfer of positively charged protein, the membrane is placed on the cathode side of the gel. Transfer is achieved by applying a maximum current of 0.8 mA/cm^2 of gel area. For a typical minigel ($8 \times 10 \text{ cm}$) and standard-sized gel ($14 \times 14 \text{ cm}$), this means 60 and 200 mA, respectively.

Additional Materials (also see Basic Protocol 1)

- Six sheets of Whatman 3MM filter paper or equivalent, cut to size of gel and saturated with transfer buffer
- Semidry transfer unit (Hoefer, Bio-Rad, or Sartorius)

1. Prepare samples and separate proteins using small or standard-sized one- or two-dimensional gels (UNITS 10.2-10.4).

Because transfer efficiency depends on many factors (e.g., gel concentration and thickness, protein size, shape, and net charge) results may vary. Below is a guideline for 0.75-mm-thick SDS-PAGE gels transferred by semidry blotting.

Percent acrylamide (resolving gel)	Size range transferred (~100% efficiency)
5–7	29–150 kD
8–10	14–66 kD
13–15	<36 kD
18–20	<20 kD

2. Prepare transfer membrane (see Basic Protocol 1, step 6).
3. Disassemble gel sandwich. Remove and discard stacking gel.

Equilibration of the separating gel with transfer buffer is not normally required for semidry blotting, but it may improve transfer in some cases.

4. Place three sheets of filter paper saturated with transfer buffer on the anode (Fig. 10.8.2).

Most transfer units are designed so that negatively charged proteins move downward toward either a platinum or graphite positive electrode (anode).

CAPS transfer buffer, pH 10.5 (see recipe for transfer buffer) can be used in place of the Tris/glycine/methanol transfer buffer of Basic Protocol 1. CAPS buffer should be used if the protein is to be sequenced right on the membrane (Moos, 1992), as glycine will interfere with this procedure.

The filter paper should be cut to the exact size of the gel. This forces the current to flow only through the gel and not through overlapping filter paper. Some manufacturers (e.g., Hoefer) recommend placing a Mylar mask on the lower platinum anode. With an opening that is slightly less than the size of the gel, the mask forces the current to flow through the gel and not the surrounding electrode area during transfer.

5. Place equilibrated transfer membrane on top of filter paper stack. Remove all bubbles between membrane and filter paper by rolling a test tube over surface of membrane.

Any bubbles in the filter paper stack or between the filter paper, membrane, and gel will block current flow and prevent protein transfer. This problem is indicated on the membrane by sharply defined white areas devoid of transferred protein.

6. Place gel on top of membrane. Gently roll a test tube over surface of gel to insure intimate contact between gel and membrane and to remove any interfering bubbles.

Poor contact between the gel and membrane will cause a swirled pattern of transferred proteins on the membrane. Some proteins will transfer as soon as the gel is placed on the membrane; repositioning the gel or membrane can result in a smeared or double image on the developed blot.

7. Complete the transfer stack by putting the three remaining sheets of filter paper on top of gel. Roll out bubbles as described above.

Multiple gels can be transferred using semidry blotting. Simply put a sheet of porous cellophane (Hoefer) or dialysis membrane (Bio-Rad or Sartorius) equilibrated with transfer buffer between each transfer stack (Fig. 10.8.2). Transfer efficiency is dependent on the position of the transfer stack in the blotting unit and for critical applications transferring one gel at a time is recommended. The gel next to the anode tends to be more efficiently transferred when blotting more than one gel at a time.

Transfer proteins from gel to membrane

8. Place top electrode onto transfer stack.

Most units have safety-interlock features and can only be assembled one way. Consult manufacturer's instructions for details.

Once the transfer stack has been assembled with both electrodes, do not move the top electrode. This can shift the transfer stack and move the gel relative to the membrane. Some transfer will occur as soon as the gel contacts the membrane, and any shifting of the transfer stack after assembly will distort the transfer pattern.

9. Carefully connect high-voltage leads to the power supply (see UNIT 10.2 for safety precautions). Apply constant current to initiate protein transfer. Transfers of 1 hr are generally sufficient.

In general, do not exceed 0.8 mA/cm² of gel area. For a typical minigel (8 × 10 cm) and standard-sized gel (14 × 14 cm) this means ~60 and 200 mA, respectively.

Monitor the temperature of the transfer unit directly above the gel by touch. The unit should not exceed 45°C. If the outside of the unit is warm, too much current is being applied. Note that units with graphite electrodes are more prone to heating, because graphite has much more resistance to current flow than platinum or steel electrodes.

10. After transfer, turn off power supply and disassemble unit. Remove membrane from transfer stack, marking orientation as in step 12 of Basic Protocol 1. Proceed with staining and immunoprobng (see Basic Protocol 1, steps 13 and 14).

BLOTTING OF STAINED GELS

Gels stained with Coomassie blue R250 (*UNIT 10.6*) can be effectively immunoblotted by the following procedure, based on Perides et al. (1986) and Dionisi et al. (1995). Briefly, the stained gel is soaked in a series of solutions designed to increase the solubility of the proteins after staining and fixation. After transfer, the membranes are treated with 45% or 100% methanol to decrease the Coomassie blue bound to the membrane prior to processing for chromogenic development. For chemiluminescent development, removal of the Coomassie blue is generally not needed.

Materials

Destained gel containing proteins of interest (*UNIT 10.6*)
25 mM Tris base/192 mM glycine/1% SDS
25 mM Tris base/192 mM glycine/0.1% SDS

1. Soak destained gel containing proteins of interest in distilled water for 15 min.
2. Equilibrate gel with 25 mM Tris base/192 mM glycine/1% SDS for 1 hr with gentle agitation.
3. Transfer gel to 25 mM Tris base/192 mM glycine/0.1% SDS and equilibrate 30 min with gentle agitation.

To increase transfer efficiency of larger proteins, the gel should be transferred to the above solution with 6 M urea for an additional 30 min.

4. Proceed with transfer (see Basic Protocol 1, steps 2 to 12).

For the most efficient transfer and binding to the membrane, the transfer buffer should contain SDS.

5. After transfer, soak membranes for 10 to 30 min in 45% methanol (nitrocellulose) or 100% methanol (nylon or PVDF) to remove the bound Coomassie blue.

This step is not needed if using chemiluminescent reactions or radiolabeled protein A for immunodevelopment. Destaining of the nitrocellulose membrane is enhanced by adding a small ball of laboratory tissue to the methanol to absorb the Coomassie blue.

6. Proceed with immunoprobng and visual detection of proteins (see Basic Protocols 2 and 3 and Alternate Protocols 3 and 4).

REVERSIBLE STAINING OF TRANSFERRED PROTEINS

To verify transfer efficiency, nitrocellulose and PVDF membranes can be reversibly stained. This method will not work on nylon membranes.

Additional Materials (also see Basic Protocol 1)

Ponceau S solution (see recipe)

Additional reagents and equipment for photographing membranes (*UNIT 10.6*)

1. Following protein transfer to nitrocellulose or PVDF (see Basic Protocol 1 or Alternate Protocol 1), place membrane in Ponceau S solution 5 min at room temperature.

ALTERNATE PROTOCOL 2

SUPPORT PROTOCOL 1

Analysis of Proteins

10.8.7

**SUPPORT
PROTOCOL 2**

2. Destain 2 min in water. Photograph membrane if required (*UNIT 10.6*) and mark the molecular-weight-standard band locations with indelible ink.
3. Completely destain membrane by soaking an additional 10 min in water.

QUANTITATION OF PROTEIN WITH PONCEAU S

In addition to qualitatively visualizing proteins on membranes after blotting, Ponceau S provides a convenient method for quantitating the amount of protein in a given lane. By eluting the dye off the strip and reading in a spectrophotometer (OD_{525}), an internal control value of protein on a lane is obtained. This value is used to correct for any differences in protein loading from lane to lane. Comparison of the Ponceau S value to the chemiluminescent or chromogenic immunodetection value determined by densitometry provides a straightforward correction for lane-to-lane variation. This method works best for complex mixtures where the immunodetected protein represents a small proportion of the total protein (Klein et al., 1995).

Additional Materials (also see Basic Protocol 1)

Spectrophotometer and 2-ml cuvette

1. Following protein transfer to nitrocellulose, PVDF, or nylon (see Basic Protocol 1 or Alternate Protocol 1), stain membrane, photograph, and destain (see Support Protocol 1).

Membranes should be destained until the background becomes white.

2. Mark lanes with a soft pencil and cut lanes into strips.
3. Place each strip into 7 ml of distilled water for 7 min and remove the resulting solution. If any particulates are visible, centrifuge 30 min at 2000 rpm to remove them.
4. Read OD_{525} in a 2-ml cuvette.

Any variation in gel-to-gel sample loading and blotting efficiency will be reflected in a change in OD of the sample lanes when compared to the control.

The change in OD can be calibrated to a known amount of protein loaded on the control lane. This will be a relative value, however, since the transfer out of the gel and binding to the membrane is rarely 100%.

**BASIC
PROTOCOL 2**

**IMMUNOPROBING WITH DIRECTLY CONJUGATED SECONDARY
ANTIBODY**

Immobilized proteins are probed with specific antibodies to identify and quantitate any antigens present. The membrane is immersed in blocking buffer to fill all protein-binding sites with a nonreactive protein or detergent. Next, it is placed in a solution containing the antibody directed against the antigen (primary antibody). The blot is washed and exposed to an enzyme-antibody conjugate directed against the primary antibody (secondary antibody; e.g., goat anti-rabbit IgG). Antigens are identified by chromogenic or luminescent visualization (see Basic Protocol 3 and Alternate Protocol 4) of the antigen/primary antibody/secondary antibody/enzyme complex bound to the membrane. Tween 20 is a common alternative to protein-blocking agents when using nitrocellulose or PVDF filters.

Materials

Membrane with transferred proteins (see Basic Protocol 1 or Alternate Protocol 1)
Blocking buffer appropriate for membrane and detection protocol (see recipe)
Primary antibody specific for protein of interest
TTBS (nitrocellulose or PVDF) or TBS (nylon; see APPENDIX 2 for recipes)
Secondary antibody conjugate: horseradish peroxidase (HRPO)–or alkaline phosphatase (AP)–anti-Ig conjugate (Cappel, Vector Labs, Kirkegaard & Perry, or Sigma; dilute as indicated by manufacturer and store frozen in 25- μ l aliquots until use)

Heat-sealable plastic bag

Powder-free gloves

Plastic box

1. Place membrane in heat-sealable plastic bag with 5 ml blocking buffer and seal bag. Incubate 30 min to 1 hr at room temperature with agitation on an orbital shaker or rocking platform.

Usually 5 ml buffer is sufficient for two to three membranes (14 \times 14-cm size). If membrane is to be stripped and reprobed (see Support Protocol 3), blocking buffer must contain casein (for AP systems) or nonfat dry milk.

Plastic incubation trays are often used in place of heat-sealable bags, and can be especially useful when processing large numbers of strips in different primary antibody solutions.

2. Dilute primary antibody in blocking buffer.

Primary antibody dilution is determined empirically but is typically 1/100 to 1/1000 for a polyclonal antibody (Fig. 10.8.3; UNIT 11.14), 1/10 to 1/100 for hybridoma supernatants (UNIT 11.8), and \geq 1/1000 for murine ascites fluid containing monoclonal antibodies (UNIT 11.10). Ten- to one-hundred-fold higher dilutions can be used with alkaline phosphatase–or luminescence–based detection systems. Both primary and secondary antibody solutions can be used at least twice, but long-term storage (i.e., >2 days at 4°C) is not recommended.

3. Open bag and pour out blocking buffer. Replace with diluted primary antibody and incubate 30 min to 1 hr at room temperature with constant agitation.

Usually 5 ml diluted primary antibody solution is sufficient for two to three membranes (14 \times 14-cm size). Incubation time may vary depending on conjugate used.

When using plastic trays, the primary and secondary antibody solution volume should be increased to 25 to 50 ml. For membrane strips, incubation trays with individual slots are recommended. Typically, 0.5 to 1 ml solution/slot is needed.

4. Remove membrane from plastic bag with gloved hand. Place in plastic box and wash four times by agitating with 200 ml TTBS (nitrocellulose or PVDF) or TBS (nylon), 10 to 15 min each time.

5. Dilute secondary antibody HRPO- or AP-anti-Ig conjugate in blocking buffer.

Commercially available enzyme–conjugated secondary antibody is usually diluted 1/200 to 1/2000 prior to use (Harlow and Lane, 1988).

6. Place membrane in new heat-sealable plastic bag, add diluted HRPO- or AP-anti-Ig conjugate, and incubate 30 min to 1 hr at room temperature with constant agitation.

When using plastic incubation trays, see step 3 annotation for proper antibody solution volumes.

7. Remove membrane from bag and wash as in step 4. Develop according to appropriate visualization protocol (see Basic Protocol 3 or Alternate Protocol 4).

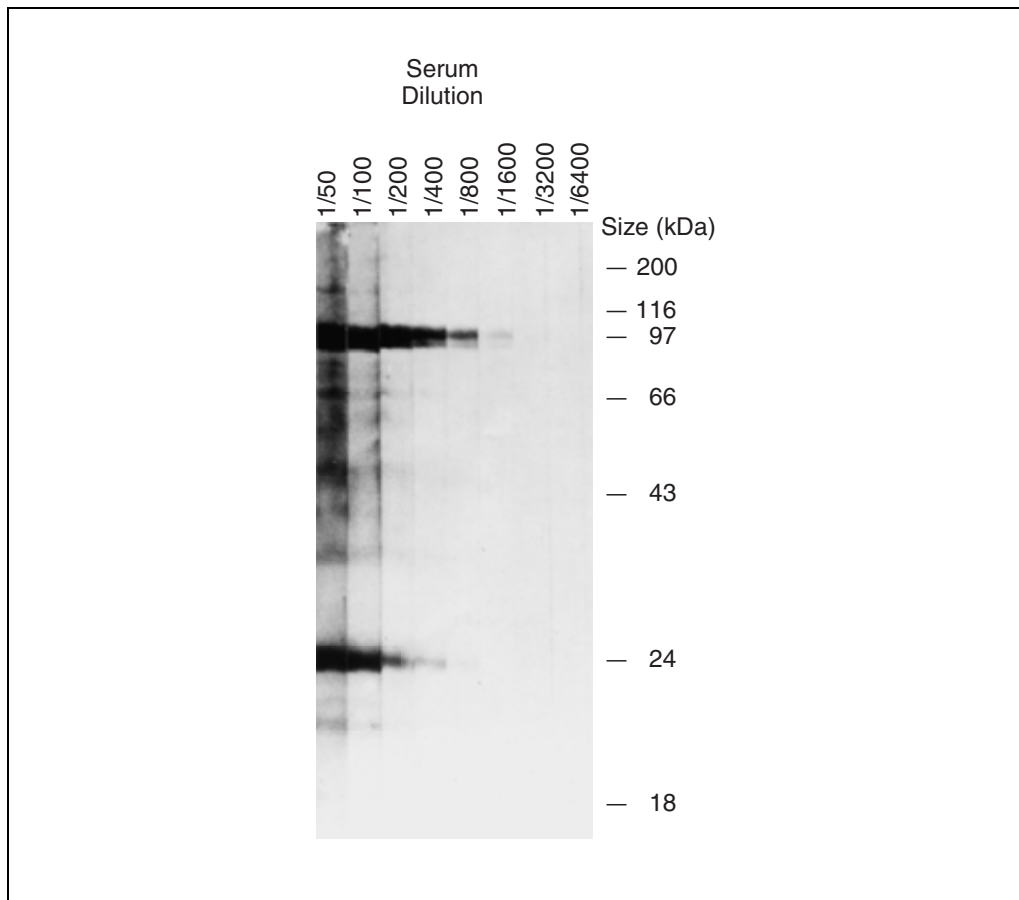


Figure 10.8.3 Serial dilution of primary antibody directed against the 97-kDa catalytic subunit of the plant plasma membrane ATPase. Blot was developed with HRPO-coupled avidin-biotin reagents according to the second alternate protocol and visualized with 4-chloro-1-naphthol (4CN). Note how background improves with dilution.

ALTERNATE PROTOCOL 3

IMMUNOPROBING WITH AVIDIN-BIOTIN COUPLING TO SECONDARY ANTIBODY

The following procedure is based on the Vectastain ABC kit from Vector Labs (*APPENDIX 4*). It uses an avidin-biotin complex to attach horseradish peroxidase (HRPO) or alkaline phosphatase (AP) to the biotinylated secondary antibody. Avidin-biotin systems are capable of extremely high sensitivity due to the multiple reporter enzymes bound to each secondary antibody. In addition, the detergent Tween 20 is a popular alternative to protein-blocking agents when using nitrocellulose or PVDF filters.

Additional Materials (also see Basic Protocol 2)

Blocking buffer appropriate for membrane and detection protocol (see recipe)
TTBS (nitrocellulose or PVDF) *or* TBS (neutral or positively charged nylon; see *APPENDIX 2* for recipes)

Vectastain ABC (HRPO) or ABC-AP (AP) kit (Vector Labs) containing the following: reagent A (avidin), reagent B (biotinylated HRPO or AP), and biotinylated secondary antibody (request membrane immunodetection protocols when ordering)

1. Equilibrate membrane in appropriate blocking buffer in heat-sealed plastic bag with constant agitation using an orbital shaker or rocking platform. For nitrocellulose and PVDF, incubate 30 to 60 min at room temperature. For nylon, incubate ≥ 2 hr at 37°C.

TTBS is well suited for avidin-biotin systems. For nylon, protein-binding agents are recommended. Because nonfat dry milk contains residual biotin, which will interfere with the immunoassay, it must be used in the blocking step only. If membrane is to be stripped and reprobed (see Support Protocol 3), blocking buffer must contain casein (for AP systems) or nonfat dry milk.

Plastic incubation trays are often used in place of heat-sealable bags, and can be especially useful when processing large numbers of strips in different primary antibody solutions.

2. Prepare primary antibody solution in TTBS (nitrocellulose or PVDF) or TBS (nylon).

Dilutions of sera containing primary antibody generally range from 1/100 to 1/100,000. This depends in large part on the sensitivity of the detection system. With high-sensitivity avidin-biotin systems, dilutions from 1/1000 to 1/100,000 are common. Higher dilutions can be used with AP- or luminescence-based detection systems. To determine the appropriate concentration of the primary antibody, a dilution series is easily performed with membrane strips. Separate antigens on a preparative gel (i.e., a single large sample well) and immunoblot the entire gel. Cut 2- to 4-mm strips by hand or with a membrane cutter (Schleicher and Schuell; Inotech) and incubate individual strips in a set of serial dilutions of primary antibody. The correct dilution should give low background and high specificity (Fig. 10.8.3).

3. Open bag, remove blocking buffer, and add enough primary antibody solution to cover membrane. Incubate 30 min at room temperature with gentle rocking.

Usually 5 ml diluted primary antibody solution is sufficient for two to three membranes (14 × 14-cm size). Incubation time may vary depending on conjugate used.

When using plastic trays, the primary and secondary antibody solution volume should be increased to 25 to 50 ml. For membrane strips, incubation trays with individual slots are recommended. Typically, 0.5 to 1 ml solution/slot is needed.

4. Remove membrane from bag and place in plastic box. Wash membrane three times over a 15-min span in TTBS (nitrocellulose or PVDF) or TBS (nylon). Add enough TTBS or TBS to fully cover the membrane (e.g., 5 to 10 ml/strip or 25 to 50 ml/whole membrane).

5. Prepare biotinylated secondary antibody solution by diluting two drops biotinylated antibody with 50 to 100 ml TTBS (nitrocellulose or PVDF) or TBS (nylon).

This dilution gives both high sensitivity and enough volume to easily cover a large 14 × 14-cm membrane.

6. Transfer membrane to fresh plastic bag containing secondary antibody solution. Incubate 30 min at room temperature with slow rocking, then wash as in step 4.

When using plastic incubation trays, see step 3 annotations for proper antibody solution volumes.

7. While membrane is being incubated with secondary antibody, prepare avidin-biotin-HRPO or -AP complex. Mix two drops Vectastain reagent A and two drops reagent B into 10 ml TTBS (nitrocellulose or PVDF) or TBS (nylon). Incubate 30 min at room temperature, then further dilute to 50 ml with TTBS or TBS.

Diluting the A and B reagents to 50 ml expands the amount of membrane that can be probed without greatly affecting sensitivity. Sodium azide is a peroxidase inhibitor and should not be used as a preservative. Casein, nonfat dry milk, serum, and some grades of BSA may interfere with the formation of the avidin-biotin complex and should not be used in the presence of avidin or biotin reagents (Gillespie and Hudspeth, 1991; Vector Labs).

8. Transfer membrane to avidin-biotin-enzyme solution. Incubate 30 min at room temperature with slow rocking, then wash over a 30-min span as in step 4.

9. Develop membrane according to the appropriate visualization protocol (see Basic Protocol 3 or Alternate Protocol 4).

VISUALIZATION WITH CHROMOGENIC SUBSTRATES

Bound antigens are typically visualized with chromogenic substrates. The substrates 4CN, DAB/NiCl₂, and TMB are commonly used with horseradish peroxidase (HRPO)–based immunodetection procedures, while BCIP/NBT is recommended for alkaline phosphatase (AP)–based procedures (see Table 10.8.1). After incubation with primary and secondary antibodies, the membrane is placed in the appropriate substrate solution. Protein bands usually appear within a few minutes.

Materials

Membrane with transferred proteins and probed with antibody-enzyme complex
(see Basic Protocol 2 or Alternate Protocol 3)

TBS (*APPENDIX 2*)

Chromogenic visualization solution (Table 10.8.1)

Additional reagents and equipment for gel photography (*UNIT 10.6*)

1. If final membrane wash (see Basic Protocol 2, step 7, or see Alternate Protocol 3, step 8) was performed in TTBS, wash membrane 15 min at room temperature in 50 ml TBS.
The Tween 20 in TTBS interferes with 4CN development (Bjerrum et al., 1988).
2. Place membrane into chromogenic visualization solution. Bands should appear in 10 to 30 min.
3. Terminate reaction by washing membrane in distilled water. Air dry and photograph (*UNIT 10.6*) for a permanent record.

VISUALIZATION WITH LUMINESCENT SUBSTRATES

Antigens can also be visualized with luminescent substrates. Detection with light offers both speed and enhanced sensitivity over chromogenic and radioisotopic procedures. After the final wash, the blot is immersed in a substrate solution containing luminol for horseradish peroxidase (HRPO) systems or dioxetane phosphate for alkaline phosphatase (AP) systems, sealed in thin plastic wrap, and placed firmly against film. Exposures range from a few seconds to several hours, although typically strong signals appear within a few seconds or minutes.

Additional Materials (*also see Basic Protocol 3*)

Luminescent substrate buffer: 50 mM Tris·Cl, pH 7.5 (for HRPO; *APPENDIX 2*) *OR*
dioxetane phosphate substrate buffer (for alkaline phosphatase; see recipe)

Nitro-Block solution (AP reactions only): 5% (v/v) Nitro-Block (Applied
Biosystems) in dioxetane phosphate substrate buffer, prepared just before use

Luminescent visualization solution (Table 10.8.1)

Clear plastic wrap

Additional reagents and equipment for autoradiography (*APPENDIX 3A*)

NOTE: See Troubleshooting section for suggestions concerning optimization of this protocol, particularly when employing AP-based systems.

1. Equilibrate membrane in two 15-min washes with 50 ml substrate buffer.
For blots of whole gels, use 50 ml substrate buffer; for strips, use 5 to 10 ml/strip.
2. *For AP reactions using nitrocellulose or PVDF membranes:* Incubate 5 min in Nitro-Block solution, followed by 5 min in substrate buffer (volumes as in step 1).

Nitro-Block enhances light output from the dioxetane substrate in reactions using AMPPD, CSPD, or Lumigen-PPD concentrate. It is required for nitrocellulose and recommended for PVDF membranes. It is not needed for Lumi-Phos 530, AP reactions on nylon

Table 10.8.1 Chromogenic and Luminescent Visualization Systems^a

System	Reagent ^b	Reaction/Detection	Comments ^c
Chromogenic			
HRPO-based	4CN	Oxidized products form purple precipitate	Not very sensitive (Tween 20 inhibits reaction); fades rapidly upon exposure to light
	DAB/NiCl ₂ ^d	Forms dark brown precipitate	More sensitive than 4CN but potentially carcinogenic; resulting membrane easily scanned
	TMB ^e	Forms dark purple stain	More stable, less toxic than DAB/NiCl ₂ ; may be somewhat more sensitive ^e ; can be used with all membrane types; kits available from Kirkegaard & Perry, TSI, Moss, and Vector Labs
AP-based	BCIP/NBT	BCIP hydrolysis produces indigo precipitate after oxidation with NBT; reduced NBT precipitates; dark blue-gray stain results	More sensitive and reliable than other AP-precipitating substrates; note that phosphate inhibits AP activity
Luminescent			
HRPO-based	Luminol/H ₂ O ₂ / <i>p</i> -iodophenol	Oxidized luminol substrate gives off blue light; <i>p</i> -iodophenol increases light output	Very convenient, sensitive system; reaction detected within a few seconds to 1 hr
AP-based	Substituted 1,2-dioxetane-phosphates (e.g., AMPPD, CSPD, Lumigen-PPD, Lumi-Phos 530 ^f)	Dephosphorylated substrate gives off light	Protocol described gives reasonable sensitivity on all membrane types; consult instructions of reagent manufacturer for maximum sensitivity and minimum background (see Troubleshooting)

^aAbbreviations: AMPPD or Lumigen-PPD, disodium 3-(4-methoxy-1,2-dioxetane-3,2'-tricyclo[3.3.1.1^{3,7}]decan-4-yl)phenyl phosphate; AP, alkaline phosphatase; BCIP, 5-bromo-4-chloro-3-indolyl phosphate; 4CN, 4-chloro-1-naphthol; CSPD, AMPPD with substituted chlorine moiety on adamantane ring; DAB, 3,3'-diaminobenzidine; HRPO, horseradish peroxidase; NBT, nitroblue tetrazolium; TMB, 3,3',5,5'-tetramethylbenzidine.

^bRecipes and suppliers are listed in Reagents and Solutions except for TMP, for which use of a kit is recommended.

^cSee Commentary for further details.

^dDAB/NiCl₂ can be used without the nickel enhancement, but it is much less sensitive.

^eMcKimm-Breschkin (1990) reported that if nitrocellulose filters are first treated with 1% dextran sulfate for 10 min in 10 mM citrate-EDTA (pH 5.0), TMB precipitates onto the membrane with a sensitivity much greater than 4CN or DAB, and equal to or better than that of BCIP/NBT.

^fLumi-Phos 530 contains dioxetane phosphate, MgCl₂, CTAB, and fluorescent enhancer in a pH 9.6 buffer.

membranes, or HRPO-based reactions on any type of membrane. Lumi-Phos 530 is not recommended for nitrocellulose membranes.

3. Transfer membrane to visualization solution. Soak 30 sec (HRPO reactions) to 5 min (AP reactions; volumes as in step 1).

Alternatively, lay out a square of plastic wrap and pipet 1 to 2 ml visualization solution into the middle. Place membrane on the plastic so that the visualization solution spreads out evenly from edge to edge. Fold wrap back onto membrane, seal, and proceed to step 5.

4. Remove membrane, drain, and place face down on a sheet of clear plastic wrap. Fold wrap back onto membrane to form a liquid-tight enclosure.

To insure an optimal image, only one layer of plastic should be between the membrane and film. Sealable bags are an effective alternative. Moisture must not come in contact with the X-ray film.

5. In a darkroom, place membrane face down onto film.

Do this quickly and do not reposition; a double image will be formed if the membrane is moved while in contact with the film. A blurred image is usually caused by poor contact between membrane and film; use a film cassette that insures a tight fit.

6. Expose film for a few seconds to several hours.

Typically, immunoblots produce very strong signals within a few seconds or minutes. However, weak signals may require several hours to an overnight exposure. If no image is detected, expose film 30 min to 1 hr, and if needed, overnight (see Troubleshooting).

7. If desired, wash membrane in two 15-min washes of 50 ml TBS and process for chromogenic development (see Basic Protocol 3).

Chemiluminescent and chromogenic immunoblotting can be easily combined on a single blot to provide a permanent visual marker of a known protein. First probe membrane with the chemiluminescent reactions to record on film. If stripping and reprobing are needed, then process by wetting and NaOH treatment (see Support Protocol 3). For the last reaction, use chromogenic development to produce a permanent visual record of the blot. Alternatively, once the film record of the chemiluminescent blot is recorded, the blot can be rinsed briefly with distilled water and placed in the appropriate chromogenic solution for chromogenic development of the blot. This results in a permanent reference stain on the blot for comparison to the more easily scanned and quantitated film record.

ALTERNATE PROTOCOL 5

VISUALIZATION WITH CHEMILUMINESCENT, FLUORESCENT, OR CHROMOGENIC SUBSTRATES

Specialized kits are available that provide significant signal amplification and can accommodate the chemiluminescent, fluorescent, or chromogenic development of a substrate. The Vectastain ABC-AmP Western Blotting Immunodetection Kit (Vector Laboratories) features a reagent that is a preformed complex between streptavidin and biotinylated alkaline phosphatase. The substrate is either the chromogenic substrate for alkaline phosphatase, BCIP/NBT (5-bromo-4-chloro-3-indolyl phosphate/nitroblue tetrazolium), which produces a blue-purple precipitate, or a chemiluminescent/fluorescent substrate that provides very high sensitivity and has sustained light-emission characteristics. Blots can be exposed to film several times over an 8-hr period for the optimization of band intensities or resolution. In addition to providing an amplified signal, the reagent produces a very low background on both nitrocellulose and PVDF membranes. These kits have been developed to produce maximum sensitivity with minimal background staining.

Materials

- Membrane with transferred proteins (see Basic Protocol 1 or Alternate Protocol 1)
- Primary antibody (from mouse or rabbit) specific for protein of interest
- Vectastain ABC-AmP Chromogenic *or* Chemiluminescent Western Blotting Immunodetection Kit (Vector Laboratories) containing:
 - 10× casein solution (250 ml)
 - Vector biotinylated secondary antibody (0.25 ml): anti-mouse IgG (for mouse primary antibodies) *or* anti-rabbit IgG (for rabbit primary antibodies)
 - Vectastain ABC-AmP reagent A (0.5 ml)
 - Vectastain ABC-AmP reagent B (0.5 ml)
 - Substrate: chemiluminescent/fluorescent substrate (100 ml) *and* BCIP/NBT chromogenic substrate kit (stock reagents for 200 ml)
- PBS: 10 mM sodium phosphate buffer, pH 7.5 (APPENDIX 2) containing 150 mM NaCl
- 0.1 M Tris·Cl, pH 9.5 (APPENDIX 2)

Staining trays

X-ray film (e.g., Kodak BioMax)

UV transilluminator or UV imaging trans- and epi-CCD acquisition system (UVP, Inc.)

NOTE: The components supplied in each Vectastain ABC-AmP Western Blotting Immunodetection Kit provide sufficient reagents to develop approximately twenty 100-cm² blots. The volumes of the reagents in the protocol below are optimized for the development of a 100-cm² membrane. Volumes may be proportionally adjusted for blots of a different size. All kit reagents may be used immediately following dilution. For optimal results, it is recommended that all diluted reagents from the kit be used the same day that they are prepared. Vectastain ABC-AmP Kit stock reagents should be stored under refrigeration and kept in the box in which they are supplied.

Block membrane

1. For a 100-cm² (10 × 10-cm) blot, prepare 120 ml of 1× casein solution by adding 12 ml of 10× casein solution (provided with Vectastain kit) to 108 ml distilled water. Block membrane by incubating in 10 ml of 1× casein solution 5 min at room temperature with gentle agitation. Save remaining 1× casein to be used in subsequent steps.

In rare cases, some nonspecific bands may develop. This can usually be eliminated by increasing the concentration of NaCl to 0.3 to 0.5 M in the Vectastain ABC-AmP reagent (see step 5, below). Some enzymes isolated from tissues may have covalently attached biotin as a cofactor. If high salt does not prevent the Vectastain ABC-AmP reagent from binding to particular bands, use an avidin/biotin blocking step (Vector Laboratories, cat. No. SP-2001) between steps 1 and 2 in this protocol.

2. Incubate the membrane for 30 min (or for a time established to be optimal for the concentration of primary antibody used) with gentle agitation at room temperature in an appropriate concentration of primary antibody diluted in PBS.
3. Wash the membrane in 10 ml of 1× casein solution three times, each time for 4 min at room temperature with gentle agitation.
4. Prepare the biotinylated secondary antibody solution by adding 10 μl of biotinylated anti-mouse IgG or anti-rabbit IgG (provided with the Vectastain kit; choice depends on species from which primary antibody was obtained) to 10 ml of 1× casein solution (final concentration 1.5 μg/ml). Incubate membrane 30 min at room temperature with gentle agitation in 10 ml of the biotinylated secondary antibody solution. At end of incubation, wash in 1× casein solution as in step 3.
5. Prepare the Vectastain ABC-AmP reagent by adding 20 μl of reagent A and 20 μl of reagent B (both provided with the kit) to 10 ml of 1× casein solution. Incubate the membrane in 10 ml of the Vectastain ABC-AmP reagent for 10 min at room temperature with gentle agitation. At end of incubation, wash with 1× casein solution as in step 3.

Incubate membrane in substrate solution

For chemiluminescent and/or fluorescent signal detection, follow steps 6a to 10a. For chromogenic signal detection, follow steps 6b to 8b. It is recommended that the membrane be transferred to a different staining tray for the substrate development step.

To detect chemiluminescent (and/or fluorescent) signal

- 6a. Equilibrate membrane for 5 min in 0.1 M Tris·Cl, pH 9.5. Remove excess buffer by holding membrane vertically and touching edge of membrane to absorbent paper.
- 7a. Place membrane target-side-up on plastic wrap on a level surface. Pipet 5 ml of substrate on to membrane surface. Incubate 5 min under subdued light or in dark.
- 8a. Rinse membrane in 0.1 M Tris·Cl, pH 9.5, for a few seconds, then remove excess buffer by holding membrane vertically and touching edge of membrane to absorbent paper.
- 9a. Place membrane between two pieces of acetate, plastic wrap, or clear sheet protector. Expose membrane to X-ray film (e.g., KODAK BioMax film) for the appropriate time.

If necessary, background may be further reduced by washing membrane in 0.1 M Tris·Cl, pH 9.5, for 5 min at room temperature and removing excess buffer before exposure to film.
- 10a. *Optional:* To obtain fluorescent signal, following step 9a, wash membrane in 0.1 M Tris·Cl, pH 9.5, for 5 min at room temperature. Remove excess buffer. Place membrane target-side-up on a UV transilluminator (254 to 365 nm) or UV imaging trans- and epi-CCD acquisition system, and observe results.

To detect chromogenic signal

- 6b. Equilibrate membrane from step 6 for 5 min in 0.1 M Tris·Cl, pH 9.5.
- 7b. Prepare chromogenic reagent by adding 4 drops of reagent from each of the three dropper bottles in the BCIP/NBT substrate kit (provided with Vectastain kit) into 10 ml of 0.1 M Tris·Cl, pH 9.5. Incubate membrane in the substrate solution at room temperature with gentle agitation until the appropriate density of colored bands develops.

Incubation times may vary from 30 min up to several hours.
- 8b. Briefly rinse the membrane in PBS and air dry. Store blot protected from light.

**SUPPORT
PROTOCOL 3**

STRIPPING AND REUSING MEMBRANES

This stripping procedure works with blotted membranes from one- and two-dimensional gels as well as with proteins blotted from previously stained gels (Suck and Krupinska, 1996). Reprobing PVDF membranes that have been developed with chemiluminescent reagents is simple and straightforward. All residual antibodies are removed from the membrane by first rewetting it in water and then briefly treating it with NaOH. Although repeated reprobing can lead to loss of signal, up to five reprobings generally are feasible. The blot should have been previously blocked with 5% nonfat dry milk prior to treatment.

Materials

0.2 M NaOH

1. Wash blot 5 min in distilled water.

In order to effectively reprobe the membranes, casein (for AP systems) or nonfat dry milk must be used as the blocking agent. Chromogenic development leaves a permanent stain on the membrane that is difficult to remove, and should not be used when reprobing. The stain can interfere with subsequent analysis if reactive bands from sequential immunostainings are close together.

2. Transfer to 0.2 M NaOH and wash 5 min.
3. Wash blot 5 min in distilled water.
4. Proceed with immunoprobng procedure (see Basic Protocol 2 and Alternate Protocol 3).

Casein or nonfat dry milk is recommended as blocking agent when reprobing membranes.

REAGENTS AND SOLUTIONS

Deionized, distilled water should be used to prepare all solutions. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4. For selection of appropriate chromogenic or luminescent solutions, and for definition of abbreviations, see Table 10.8.1.

Alkaline phosphate substrate buffer

100 mM Tris·Cl, pH 9.5 (APPENDIX 2)
100 mM NaCl
5 mM MgCl₂

BCIP/NBT visualization solution

Mix 33 μl NBT stock (100 mg NBT in 2 ml at 70% DMF, stored <1 year at 4°C) and 5 ml alkaline phosphate substrate buffer (see recipe). Add 17 μl BCIP stock (100 mg BCIP in 2 ml of 100% DMF, <1 year at 4°C) and mix. Stable 1 hr at room temperature.

Recipe is from Harlow and Lane (1988). Alternatively, BCIP/NBT substrates may be purchased from Sigma, Kirkegaard & Perry, Moss, and Vector Labs.

Blocking buffer

Colorimetric detection

For nitrocellulose and PVDF: 0.1% (v/v) Tween 20 in TBS (TTBS; APPENDIX 2).

For neutral and positively charged nylon: Tris-buffered saline (TBS; also see APPENDIX 2) containing 10% (w/v) nonfat dry milk. Prepare just before use.

TTBS can be stored ≤1 week at 4°C.

Luminescence detection

For nitrocellulose, PVDF, and neutral nylon (e.g., Pall Biodyne A): 0.2% casein (e.g., Hammarsten grade or I-Block; Applied Biosystems) in TTBS (APPENDIX 2). Prepare just before use.

For positively charged nylon: 6% (w/v) casein/1% (v/v) polyvinyl pyrrolidone (PVP) in TTBS (APPENDIX 2). Prepare just before use.

With constant mixing, add casein and PVP to warm (65°C) TTBS. Stir for 5 min. Cool before use.

4CN visualization solution

Mix 20 ml ice-cold methanol with 60 mg 4CN. Separately mix 60 μl of 30% H₂O₂ with 100 ml TBS (APPENDIX 2) at room temperature. Rapidly mix the two solutions and use immediately.

DAB/NiCl₂ visualization solution

5 ml 100 mM Tris·Cl, pH 7.5 (APPENDIX 2)
100 μl DAB stock (40 mg/ml in H₂O, stored in 100-μl aliquots at -20°C)
25 μl NiCl₂ stock (80 mg/ml in H₂O, stored in 100-μl aliquots at -20°C)
15 μl 3% H₂O₂
Mix just before use

CAUTION: *Handle DAB carefully, wearing gloves and mask; it is a carcinogen.*

Suppliers of peroxidase substrates are Sigma, Kirkegaard & Perry, Moss, and Vector Labs.

Dioxetane phosphate substrate buffer

1 mM MgCl₂
0.1 M diethanolamine
0.02% sodium azide (optional)
Adjust to pH 10 with HCl and use fresh

Traditionally, the AMPPD substrate buffer has been a solution containing 1 mM MgCl₂ and 50 mM sodium carbonate/bicarbonate, pH 9.6 (Gillespie and Hudspeth, 1991). The use of diethanolamine results in better light output (Applied Biosystems Western Light instructions).

Alternatively, 100 mM Tris·Cl (pH 9.5)/100 mM NaCl/5 mM MgCl₂ can be used (Sandhu et al., 1991).

Dioxetane phosphate visualization solution

Prepare 0.1 mg/ml AMPPD or CSPD (Applied Biosystems) or Lumigen-PPD (Lumigen; see Table 10.8.1) substrate in dioxetane phosphate substrate buffer (see recipe). Prepare just before use. Lumi-Phos 530 (Boehringer Mannheim or Lumigen) is a ready-to-use solution and can be applied directly to the membrane.

This concentration (240 μM) of AMPPD substrate is the minimum recommended by Applied Biosystems Western Light. Ten-fold lower concentrations can be used but require longer exposures.

Luminol visualization solution

0.5 ml 10× luminol stock [40 mg luminol (Sigma) in 10 ml DMSO]
0.5 ml 10× *p*-iodophenol stock [optional; 10 mg (Aldrich) in 10 ml DMSO]
2.5 ml 100 mM Tris·Cl, pH 7.5 (APPENDIX 2)
25 μl 3% H₂O₂
H₂O to 5 ml
Prepare just before use

*Recipe is from Schneppenheim et al. (1991). Premixed luminol substrate mix (Mast Immunosystems; Amersham ECL; Du Pont NEN Renaissance; Kirkegaard & Perry LumiGLO) may also be used. *p*-Iodophenol is an optional enhancing agent that increases light output. Luminol and *p*-iodophenol stocks can be stored for ≤6 months at −20°C.*

Ponceau S solution

Dissolve 0.5 g Ponceau S in 1 ml glacial acetic acid. Bring to 100 ml with water. Prepare just before use.

Transfer buffer

Add 18.2 g Tris base and 86.5 g glycine to 4 liters of water. Add 1200 ml methanol and bring to 6 liters with water. The pH of the solution is ~8.3 to 8.4. For use with PVDF filters, decrease methanol concentration to 15%; for nylon filters, omit methanol altogether.

CAPS transfer buffer can also be used. Add 2.21 g cyclohexylaminopropane sulfonic acid (CAPS; free acid), 0.5 g DTT, 150 ml methanol, and water to 1 liter. Adjust to pH 10.5 with NaOH and chill to 4°C. For proteins >60 kDa, reduce methanol content to 1% (Moos, 1992).

COMMENTARY

Background Information

Immunoprecipitation has been widely used to visualize the antigens recognized by various antibodies, both polyclonal and monoclonal. However, there are several problems inherent with this method, including the requirement for radiolabeling of antigen, co-

precipitation of tightly associated macromolecules, occasional difficulty in obtaining precipitating antibodies, and insolubility of various antigens (Talbot et al., 1984).

To circumvent these problems, electroblotting (Towbin et al., 1979)—subsequently popularized as western blotting or im-

munoblotting (Burnette, 1981)—was conceived. Immunoblotting is a rapid and sensitive assay for the detection and characterization of proteins that works by exploiting the specificity inherent in antigen-antibody recognition. It involves the solubilization and electrophoretic separation of proteins, glycoproteins, or lipopolysaccharides by SDS-PAGE (UNIT 10.2) or urea-PAGE, followed by quantitative transfer and irreversible binding to nitrocellulose, PVDF, or nylon. This technique has been useful in identifying specific antigens recognized by polyclonal or monoclonal antibodies and is highly sensitive (1 ng of antigen can be detected).

Protein blotting has important clinical applications—it is the confirmatory test for human immunodeficiency virus Type 1 (HIV-1). SDS-PAGE-separated virus proteins are blotted onto nitrocellulose and processed with patient sera. After washing and incubating (using a Tween 20/nonfat dry milk diluting and washing solution) with an anti-human IgG coupled to HRPO or alkaline phosphatase, the antigens are identified by chromogenic development. Typically, the prepared blots and all reagents needed for the test are purchased commercially (Bio-Rad Novapath; Organon Teknika Cappel HIV-1 Western Blot Kit; Du Pont/Biotech HIV Western Blot Kit).

Electroblotting of previously stained gels is a convenient way to visualize and document the gel prior to immunoblotting. Transfer efficiencies at all molecular weights will be lower with fixed and stained gels. This is particularly true of proteins >50 kDa (Perides et al., 1986). The additional incubation in 6 M urea will significantly increase transfer efficiency of all proteins and is required for proteins >50 kDa.

Ponceau S staining provides an easy method for calibrating and quantitating the amount of material on a nitrocellulose or PVDF blot. An alternative to this method is to use an internal protein control with a separate antibody probe, but these tend to be expensive and time-consuming to use. Other applications for Ponceau S calibration include monitoring transfer efficiency under varied conditions for optimization of tank and semidry blotting.

Immunoblotted proteins can be detected by chromogenic or luminescent assays (see Table 10.8.1 for a description of the reagents available for each system, their reactions, and a comparison of their advantages and disadvantages).

Luminescent detection methods offer several advantages over traditional chromogenic procedures. In general, luminescent substrates increase the sensitivity of both HRPO and phosphatase systems without the need for radioisotopes. Substrates for the latter have only recently been applied to protein blotting (see Gillespie and Hudspeth, 1991; Sandhu et al., 1991; Bronstein et al., 1992). Luminescent detection can be completed in as little as a few seconds; exposures rarely go more than 1 hr. Depending on the system, the luminescence can last for 3 days, permitting multiple exposures of the same blot. Furthermore, the signal is detected by film, and varying the exposure can result in more or less sensitivity. Luminescent blots can be easily erased and reprobed because the reaction products are soluble and do not deposit on the membrane (see below). Compared to chromogenic development, the luminescent image recorded on film is easier to photograph and to quantitate by densitometry.

Alkaline phosphatase-based luminescent protocols that achieve maximum sensitivity with minimum background can be complex, and the manufacturer's instructions should be consulted (see Reagents and Solutions). The procedure described in Alternate Protocol 4 gives reasonable sensitivity on nitrocellulose, PVDF, and nylon membranes with a minimum of steps.

Alternate Protocol 5 describes a new product that can be used in chromogenic, chemiluminescent, or fluorescent detection schemes. Vectastain's DuoLuX Chemiluminescent/Fluorescent Substrate is a novel acridan-based substrate that can be used not only in immunoblots, as described here, but also in nucleic acid detection methods (e.g., Southern, northern, western or dot blotting, colony lifts, and ELISA). The substrate is available for either alkaline phosphatase (AP) or horseradish peroxidase (HRP) development. The choice of enzyme will depend on the application. Generally, signal development is faster when using HRP. Thus, HRP may be preferred when digital imaging systems are used or when abundance of target reduces the need for a high signal-to-noise ratio. However, for applications where optimal sensitivity is required, AP is recommended, as it will provide a higher signal-to-noise ratio than HRP.

The DuoLuX Chemiluminescent/Fluorescent Substrate has very high sensitivity and prolonged light-emission characteristics. This enables image documentation with either

film or digital imaging systems. Unlike some chemiluminescent substrates, blots can be re-exposed to film as often as necessary over many hours. Because many digital imaging systems require a longer exposure time than film, the faster signal development of HRP relative to AP may be preferred when using these systems. PVDF, nitrocellulose, or nylon membranes can be used, although the chemiluminescent signal develops faster on nylon and PVDF.

In addition to its chemiluminescent properties, the reaction product is also fluorescent. Fluorescence can be recorded with a digital imaging system or a conventional camera months after chemiluminescence has faded. For fluorescence detection, nitrocellulose is recommended. Acquisition of fluorescent signal requires a much shorter exposure time than chemiluminescence, often a fraction of a second.

Critical Parameters

First and foremost, the antibody being used should recognize denatured antigen. Nonspecific binding of antibodies can occur, so control antigens and antibodies should always be run in parallel. Time of transfer and primary antibody and conjugate dilutions should always be optimized.

A variety of agents are currently used to block binding sites on the membrane after blotting (Harlow and Lane, 1988). These include Tween 20, PVP, nonfat dry milk, casein, BSA, and serum. A 0.1% (v/v) solution of Tween 20 in TBS (TTBS), a convenient alternative to protein-based blocking agents, is recommended for chromogenic development of nitrocellulose and PVDF membranes (Blake et al., 1984). In contrast to dry milk/TBS blocking solution (BLOTTO), TTBS is stable and has a long shelf life at 4°C. Furthermore, TTBS generally produces a clean background and permits subsequent staining with India ink. However, even with the application of such standard blocking procedures as 5% to 10% milk protein or 0.05% to 0.1% Tween 20, background can still be a significant problem. If this happens, using a blocking protein (e.g., goat, horse, or rabbit normal serum) from the same species as the primary antibody can reduce the background, presumably by reducing cross-reactivity between the primary antibodies and the blocking agent. If a secondary antibody detection is used, the blocking protein (i.e., normal serum) should be from the same species as the second-

dary antibody. Combinations of blocking agents can also be effective. Thus, 0.1% human serum albumin (HSA) and 0.05% Tween 20 in TBS is recommended when probing Immobilon-P membranes with human serum (Craig et al., 1993). However, this can also lead to overall loss of antigen signal, requiring a ten-fold increase in the primary antibody (serum) concentration to achieve an adequate background free antigen signal.

When using chemiluminescent detection for immunoblotting, high background frequently occurs, particularly for strong signals (Pampori et al., 1995). Several methods are available for reducing the background from chemiluminescent reactions. These include changing the type and concentration of blocking agents (see above), optimizing antibody concentrations, letting the reaction proceed for several minutes before exposing to film, or simply limiting the exposure time of the film on the blot. These procedures are not always successful, however, and can lead to inconsistent results. An alternative approach is to reduce the concentration of reagents ten-fold. This effectively removes the background and has a number of advantages which include lower cost, increased signal-to-noise ratio, and reduced detection of cross-reacting species.

Two types of nylon membrane are used for western transfer—neutral (e.g., Pall Biodyne A) and positively charged (e.g., Pall Biodyne B). Although the positively charged membranes have very good protein-binding characteristics, they tend to give a higher background. These membranes remain positively charged from pH 3 to pH 10. Neutral nylon membranes are also charged, having a mix of amino and carboxyl groups that give an isoelectric point of 6.5. Because of their high binding capacity, positively charged membranes are popular for protein applications using luminescence.

Nylon membranes require more stringent blocking steps. Here 10% nonfat dry milk in TBS is recommended for chromogenic development. During luminescence development, however, background is a more significant problem. Compared to dry milk, purified casein has minimal endogenous alkaline phosphatase activity (AP activity leads to high background) and is therefore recommended as a blocking agent for nitrocellulose, PVDF, and nylon membranes. Positively charged nylon requires much more stringent blocking with 6% (w/v) casein and 1% (v/v) poly-

vinylpyrrolidone (PVP). Because nonfat dry milk and casein may contain biotin that will interfere with avidin-biotin reactions, subsequent steps are done without protein-blocking agents when using these systems. If background is a problem, highly purified casein (0.2% to 6%) added to the antibody incubation buffers may help.

If reprobing is desired, blots can be air dried and stored at 4°C for 3 months after chemiluminescence detection. After drying, store in a sealed freezer bag until use. Repeated probing will lead to a gradual loss of signal and increased background. However, this will depend in part on the properties of the sample.

If the primary procedure is problematic due to loss of sensitivity or an increase in the background, then two possible alternative procedures for stripping membranes are recommended. The first uses 2-mercaptoethanol and SDS (Kaufmann et al., 1987; Tesfaigzi et al., 1994). Briefly, the membranes are incubated in 2% SDS/100 mM Tris·Cl, pH 7.4/100 mM 2-mercaptoethanol for 30 min at 70°C, effectively removing primary and secondary antibodies. As with the primary procedure recommended above, the repeated probing should be done with caution due to the potential loss of detection signal, and 5% nonfat dry milk is required as a blocking agent. The milk blocking agent facilitates antibody removal from the blot (Kaufmann et al., 1987). The second uses guanidine·HCl. For nylon and PVDF membranes (do *not* use with nitrocellulose), incubate the immunoblot in 7 M guanidine·HCl for 10 min at room temperature. (The short wash time is critical, as guanidine·HCl is a very strong denaturant, so do not leave the filter in this solution >15 min.) Pour off excess guanidine·HCl and then rinse the membrane several times in 1× TTBS (UNIT 10.8). Reblock the membrane and proceed with your standard immunoblotting procedure. Membranes stripped using this procedure can generally be reused three or four times.

In Alternate Protocol 5, extensive washing will reduce signal strength; do not extend the wash time unless high background is experienced. If background is excessive, repeat steps 6a through 9a with a wash time of 5 to 10 min in step 8a (optimal wash time is dependent on the degree of background previously detected and, therefore, may require optimization). The long emission lifetime of the substrate allows the user to re-expose the same blot until optimal signal to noise is achieved. Blotting

can be done onto either nylon or nitrocellulose. Nylon requires shorter exposure times and is, therefore, preferred for chemiluminescent applications. However, because of nylon's intrinsic fluorescence, nitrocellulose is preferred for fluorescence detection.

Troubleshooting

There are several problems associated with immunoblotting. The antigen is solubilized and electrophoresed in the presence of denaturing agents (e.g., SDS or urea), and some antibodies may not recognize the denatured form of the antigen transferred to the membrane. The results observed may be entirely dependent on the denaturation and transfer system used. For example, zwitterionic detergents have been shown to restore the antigenicity of outer membrane proteins in immunoblotting (Mandrell and Zollinger, 1984). Gel electrophoresis under nondenaturing conditions can also be performed (UNIT 10.2).

Other potential problems include high background, nonspecific or weak cross-reactivity of antibodies, poor protein transfer or membrane binding efficiency, and insufficient sensitivity. For an extensive survey and discussion of immunoblotting problems and artifacts, see Bjerrum et al. (1988).

If no transfer of protein has occurred, check the power supply and electroblot apparatus to make sure that the proper electrical connections were made and that power was delivered during transfer. In addition, check that the correct orientation of filter and gel relative to the anode and cathode electrodes was used.

If the transfer efficiency using the tank system appears to be low, increase the transfer time or power. Cooling (using the unit's built-in cooling cores) is generally required for transfers >1 hr. At no time should the buffer temperature go above 45°C. Prolonged transfers (>1 hr) are not possible in semidry transfer units due to rapid buffer depletion.

Alternatively, the transfer buffer can be modified to increase efficiency. Adding SDS at a concentration of 0.1% to the transfer buffer improves the transfer of all proteins out of the gel, particularly those above 60 to 90 kD in size. Lowering the concentration of methanol will also improve the recovery of proteins from the gel. These procedures are tradeoffs. Methanol improves the binding of proteins to PVDF and nitrocellulose, but at the same time hinders transfer. With SDS present, transfer efficiency is improved, but

the SDS can interfere with protein binding to the membrane. Nylon and PVDF membranes are particularly sensitive to SDS interference. If needed, 0.01% to 0.02% SDS may be used in PVDF membrane transfer buffers (Milipore, 1990). SDS and methanol should not be used in the transfer buffer for nylon (Peluso and Rosenberg, 1987).

Gel cross-linking and thickness also have a profound effect on the transfer efficiency. In general, 0.5- to 0.75-mm-thick gels will transfer much more efficiently than thicker gels (e.g., 1.5 mm thick). Gels with a higher acrylamide percentage will also transfer less efficiently. Proteins can be particularly difficult to transfer from gradient gels, and a combination of longer transfer times, thin gels, and the addition of SDS to the transfer buffer may be needed.

If the protein bands are diffuse, check the transfer cassette. The gel must be held firmly against the membrane during transfer. If the transfer sandwich is loose in the cassette, add another thin sponge or more blotter paper to both sides.

Occasionally, a grid pattern will be apparent on the membrane after tank transfer. This is caused by having either the gel or the membrane too close to the sides of the cassette. Correct this by adding more layers of filter paper to diffuse the current flowing through the gel and membrane. Use a thinner sponge and more filter paper if necessary.

If air bubbles are trapped between the filter and the gel, they will appear as clear white areas on the filter after blotting and staining. Take extra care to make sure that all bubbles are removed.

Insufficient blocking or nonspecific binding of the primary or secondary antibody will cause a high background stain. A control using preimmune serum or only the secondary antibody will determine if these problems are due to the primary antibody. Try switching to another blocking agent; protein blocking agents may weakly cross-react. Lowering the concentration of primary antibody should decrease background and improve specificity (Fig. 10.8.3).

Due to the nature of light and the method of detection, certain precautions are warranted when using luminescent visualization (e.g., Harper and Murphy, 1991). Very strong signals can overshadow nearby weaker signals on the membrane. Because light will pipe through the membrane and the surrounding plastic wrap, overexposure will produce a

broad diffuse image on the film. The signal can also saturate the film, exposing the film to a point whereby increased exposure will not cause a linear increase in the density of the image on the film.

With the alkaline phosphatase substrate AMPPD, nitrocellulose, PVDF, and nylon membranes require 2, 4, and 8 to 12 hr, respectively, to reach maximum light emission. In addition, PVDF is reported to give a stronger signal than nitrocellulose (Applied Biosystems Western Light instructions). Positively charged nylon requires special blocking procedures to minimize background (Gillespie and Hudspeth, 1991). These procedures include using a blocking and primary antibody solution containing 6% casein, 1% polyvinylpyrrolidone-40 (PVP-40), 3 mM NaN₃, 10 mM EDTA, and PBS, pH 6.8. Prior to use, the casein must be heated to 65°C to reduce alkaline phosphatase activity in the casein itself. In addition, maximum sensitivity has been observed when free biotin or biotinylated proteins are removed by pretreating the casein with avidin-agarose (Sigma).

Anticipated Results

Immunoblotting should result in the detection of one or more bands. Although antibodies directed against a single protein should produce a single band, degradation of the sample (e.g., via endogenous proteolytic activity) may cause visualization of multiple bands of slightly different size. Multimers will also form spontaneously, causing higher-molecular-weight bands on the blot. If simultaneously testing multiple antibodies directed against a complex protein mixture (e.g., using patient sera against SDS-PAGE-separated viral proteins in AIDS western blot test), multiple bands will be visualized.

For immunoblot or protein dot blot chemiluminescent applications, the sensitivity using HRP is ~1 pg of target protein. For chemiluminescent western blots, DuoLuX Chemiluminescent/Fluorescent Substrate (see Alternate Protocol 5) can be used on either nitrocellulose or PVDF membranes.

Time Considerations

The entire immunoblotting procedure can be completed in 1 to 2 days, depending on transfer time and type of gel. Gel electrophoresis requires 4 to 6 hr on a regular gel and 1 hr on a minigel. Transfer time can be 1 hr (high-power transfer) to overnight. Blocking, conjugate incubation, and washing each take

30 min to 1 hr. Finally, substrate incubation requires 10 to 30 min (chromogen) and a few seconds to several hours (luminescence).

For the DuoLuX Chemiluminescent/Fluorescent Substrate (see Alternate Protocol 5), optimal exposure times can vary from 5 sec to 15 min. It is recommended that initial exposures be taken between 1 and 5 min. Band intensities or resolution can then be optimized by lengthening or shortening exposure times based on the initial results.

Literature Cited

- Bjerrum, O.J., Larsen, K.P., and Heegaard, N.H.H. 1988. Nonspecific binding and artifacts-specificity problems and troubleshooting with an atlas of immunoblotting artifacts. *In* CRC Handbook of Immunoblotting of Proteins, Vol. I: Technical Descriptions (O.J. Bjerrum and N.H.H. Heegaard, eds.) pp. 227-254. CRC Press, Boca Raton, Fla.
- Blake, M.S., Johnston, K.H., Russell-Jones, G.J., and Gotschlich, E.C. 1984. A rapid, sensitive method for detection of alkaline phosphatase-conjugated anti-antibody on Western blots. *Anal. Biochem.* 136:175-179.
- Bronstein, I., Voyta, J.C., Murphy, O.J., Bresnick, L., and Kricka, L.J. 1992. Improved chemiluminescent Western blotting procedure. *BioTechniques* 12:748-753.
- Burnette, W.N. 1981. Western blotting: Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* 112:195-203.
- Craig, W.Y., Poulin, S.E., Collins, M.F., Ledue, T.B., and Ritchie, R.F. 1993. Background staining in immunoblot assays. Reduction of signal caused by cross-reactivity with blocking agents. *J. Immunol. Methods* 158:67-76.
- Dionisi, H.M., Checa, S.K., and Viale, A.M. 1995. Protein immunoblotting of stained gels. *BioTechniques* 19:348-350.
- Gillespie, P.G. and Hudspeth, A.J. 1991. Chemiluminescence detection of proteins from single cells. *Proc. Natl. Acad. Sci. U.S.A.* 88:2563-2567.
- Harlow, E. and Lane, D. 1988. Immunoblotting. *In* Antibodies: A Laboratory Manual, pp. 471-510. CSH Laboratory, Cold Spring Harbor, N.Y.
- Harper, D.R. and Murphy, G. 1991. Nonuniform variation in band pattern with luminol/horseradish peroxidase Western blotting. *Anal. Biochem.* 192:59-63.
- Kaufmann, S.H., Ewing, C.M., and Shaper, J.H. 1987. The erasable Western blot. *Anal. Biochem.* 161:89-95.
- Klein, D., Kern, R.M., and Sokol, R.Z. 1995. A method for quantification and correction of proteins after transfer to immobilization membranes. *Biochem. Mol. Biol. Int.* 36:59-66.

Mandrell, R.E. and Zollinger, W.D. 1984. Use of zwitterionic detergent for the restoration of antibody-binding capacity of electroblotted meningococcal outer membrane proteins. *J. Immunol. Methods* 67:1-11.

McKimm-Breschkin, J.L. 1990. The use of tetramethylbenzidine for solid phase immunoassays. *J. Immunological Methods* 135:277-280.

Millipore. 1990. Protein blotting protocols for Immobilon-P transfer membrane. Bedford, Mass.

Moos, M. 1992. Isolation of proteins for microsequence analysis *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 8.7.1-8.7.12. Greene Publishing Associates and John Wiley & Sons, New York.

Pampori, N.A., Pampori, M.K., and Shapiro, B.H. 1995. Dilution of the chemiluminescence reagents reduces the background noise on Western blots. *BioTechniques* 18:588-590.

Peluso, R.W. and Rosenberg, G.H. 1987. Quantitative electrotransfer of proteins from sodium dodecyl sulfate polyacrylamide gels onto positively charged nylon membranes. *Anal. Biochem.* 162:389-398.

Perides, G., Plagens, U., and Traub, P. 1986. Protein transfer from fixed, stained, and dried polyacrylamide gels and immunoblot with protein A-gold. *Anal. Biochem.* 152:94-99.

Sandhu, G.S., Eckloff, B.W., and Kline, B.C. 1991. Chemiluminescent substrates increase sensitivity of antigen detection in Western blots. *BioTechniques* 11:14-16.

Schneppenheim, R., Budde, U., Dahlmann, N. and Rautenberg, P. 1991. Luminography—a new, highly sensitive visualization method for electrophoresis. *Electrophoresis* 12:367-372.

Suck, R.W.L. and Krupinska, K. 1996. Repeated probing of Western blots obtained from Coomassie Brilliant Blue-stained or unstained polyacrylamide gels. *BioTechniques* 21:418-422.

Talbot, P.V., Knobler, R.L., and Buchmeier, M. 1984. Western and dot immunoblotting analysis of viral antigens and antibodies: Application to murine hepatitis virus. *J. Immunol. Methods* 73:177-188.

Tesfaigzi, J., Smith-Harrison, W., and Carlson, D.M. 1994. A simple method for reusing western blots on PVDF membranes. *BioTechniques* 17:268-269.

Towbin, H., Staehelin, T., and Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* 76:4350-4354.

Key References

Bjerrum, O.J. and Schafer-Nielsen, C. 1986. Buffer systems and transfer parameters for semidry electroblotting with a horizontal apparatus. *In* Electrophoresis '86 (M.J. Dunn, ed.) pp. 315-327. VCH Publishers, Deerfield Beach, Fla.

Describes the semidry blotting system.

Gillespie and Hudspeth, 1991. See above.

Describes alkaline phosphatase-luminescent detection methods.

Harlow and Lane, 1988. See above.

Details alternative detection methods.

Salinovich, O. and Montelaro, R.C. 1986. Reversible staining and peptide mapping of proteins transferred to nitrocellulose after separation by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Anal. Biochem.* 156:341-347.

Describes the use of Ponceau S staining for immunoblotting.

Schneppenheim et al., 1991. See above.

Details peroxidase-based luminescent detection methods.

Contributed by Sean Gallagher
UVP, Inc.
Upland, California

Scott E. Winston and Steven A.
Fuller (tank transfer systems)
Nabi
Rockville, Maryland

John G.R. Hurrell (tank transfer systems;
reversible staining of proteins)
FluorRx
Carmel, Indiana

PURIFICATION OF PROTEINS BY CONVENTIONAL CHROMATOGRAPHY

Purification of proteins by conventional chromatography is usually achieved by using a combination of chromatographic methods, including gel filtration, ion-exchange, immunoaffinity, and affinity chromatography. With the possible exceptions of immunoaffinity and affinity chromatography, it is rarely possible to purify a protein from a mixture of proteins in a single chromatographic step. Hence, it will usually be necessary to utilize separate, sequential steps involving different modes of chromatography (e.g., gel filtration *and* ion-exchange). In turn, each type of chromatography requires an informed choice of matrix—considering bead shape, size, and porosity, in addition to functional group type, charge, distribution, and density—as well as elution conditions (i.e., pH, ionic strength, and gradient shape). The goal is to isolate a specific protein, and simultaneously reduce contaminants to levels that will not interfere with structural or functional analyses, in the minimum number of chromatographic steps.

Gel filtration (*UNIT 10.9*) separates proteins on the basis of differences in molecular size. Gel filtration matrices differ in porosity and chemical composition (i.e., hydrophobicity and charge). The former property leads to selective exclusion of proteins that are not small enough to penetrate the matrix pores, which is the physical basis of gel filtration. The latter property leads to mixed-mode separation effects—glycoproteins might interact with sugar-based polymeric matrices or negatively charged proteins might be repulsed by a matrix with a partial negative charge. The choice of gel filtration matrix is based on the molecular size of the protein being purified, as well as the expected range of molecular sizes of the contaminating proteins, since the effect of composition is unpredictable. Gel filtration is ideally suited to separating a desired protein from proteins that are either much larger or much smaller. It is a poor choice for separating proteins with similar molecular sizes, and therefore is never the method of choice for beginning a multistep purification from a complex mixture of proteins. However, gel filtration is excellent for removing low-molecular-size contaminants, such as salts and unbound detergents, from a purified protein in preparation for structural or functional analysis.

In general terms, the most versatile chromatographic separation method is ion-exchange chromatography (*UNIT 10.9*), since it relies on subtle differences between both the number of charges and the distribution of charge groups, under a given set of pH and solvent conditions, on the surface of the desired protein and all other contaminating proteins. Most proteins carry a net negative charge at pH values above their isoelectric point, and a net positive charge when below it. As a general rule, one may assume that most proteins are negatively charged at neutral pH and that anion-exchange chromatography is the method of choice for their purification; in the occasional cases when a desired protein is positively charged, cation-exchange chromatography should be used. It is important to realize that all chromatographic separations are mixed-mode separations, because chromatographic matrices have different porosities, leading to molecular sieving whereby large proteins may be excluded from the matrix while small molecules are included. Hence, small molecules will elute more slowly than larger molecules. In addition, matrices differ in chemical composition and in their hydrophilicity or hydrophobicity. The selection process for a matrix is largely empirical, but a thorough understanding of ion-exchange chromatography, as described in the Strategic Planning section of *UNIT 10.10*, will help an investigator make a judicious initial choice of chromatographic matrix and elution conditions.

Immunoaffinity chromatography (*UNIT 10.11A*) is a highly selective chromatographic method for purifying proteins to homogeneity, or markedly simplifying a protein-con-

Contributed by John Smith

Current Protocols in Molecular Biology (1998) 10.9.1-10.9.2

Copyright © 1998 by John Wiley & Sons, Inc.

taining mixture, prior to gel filtration or ion-exchange chromatography. The selectivity of immune recognition makes it possible, under favorable conditions, to purify a protein in a single step. However, this form of chromatography relies on the existence of a monoclonal or polyclonal antibody specific for the native structure of the protein being purified. It is also important to keep in mind the likelihood of nonspecific protein binding of contaminating proteins to the immunoaffinity matrix, to its coupled antibody, and to the protein-antibody complex. An empirical choice of elution buffer—with regard to pH, ionic strength, and detergent concentration—must be made for each purification to minimize these effects.

Metal-chelate affinity chromatography (*UNIT 10.11B*) is another highly selective method, which relies on the binding affinity of short sequences of amino acids (such as polyhistidine) to nickel chelates tethered to a chromatographic matrix. Such polyamino acid sequences, or tags, are incorporated into expressed proteins by genetic engineering of the protein genes as a means of simplifying purification. As in other forms of chromatography, mixed-mode separation will occur simultaneously and may complicate the separation.

Chromatographic purification of membrane proteins may require the addition of low concentrations of nonionic detergents (*APPENDIX 11*) to solubilize the protein and to decrease its hydrophobic interaction with chromatographic matrices. The choice of detergent depends on its critical micelle concentration, UV properties, and cost.

The simplest way of monitoring the effectiveness of a purification is to assess the electrophoretic patterns before and after a given purification by gel electrophoresis (*UNIT 10.2*), and the presence of the desired protein by either a biological or immunological assay. It is important to think carefully about the design of the assay or mode of detection, since an unnecessarily complicated assay or analysis greatly augments the tedium of chromatographic separations.

John Smith

Gel-Filtration Chromatography

Gel filtration (GF) separates proteins solely on the basis of differences in molecular size. Separation is achieved using a porous matrix to which the molecules, for steric reasons, have different degrees of access (smaller molecules have greater access and larger molecules less). The matrix is packed into a chromatographic column, the sample mixture applied, and the separation accomplished by passing an aqueous buffer (the mobile phase) through the column. Protein molecules that are confined in the volume outside the matrix beads will be swept through the column by the mobile phase and the sample is thus eluted in decreasing order of size. The protein zones eluted are detected by an in-line UV monitor and fractions are collected for subsequent specific analysis or further preparation steps.

GF is often used to condition samples prior to further purification by adsorptive techniques. It is also frequently used in the final polishing stage to remove a few remaining contaminants with surface properties similar to the target protein (e.g., for polishing monomers from oligomers).

Three basic protocols are provided in this unit, corresponding to the major applications of gel filtration. GF is used for desalting or group separation (see Basic Protocol 1), in which the target protein and contaminating solutes differ substantially in molecular size. It is also used for protein fractionation (see Basic Protocol 2), where proteins with only a small size difference must be separated. A third application is in characterizing the molecular dimensions of proteins (see Basic Protocol 3). A Support Protocol is also provided for calibrating GF columns to be used in estimating molecular size.

NOTE: Among the items referred to in this unit, Sephadex, Sephacryl, Superdex, Superose, and HiLoad are trademarks of Amersham Pharmacia Biotech; Bio-Gel and Bio-Sil are trade marks of Bio-Rad; and Toyopearl and TSK SW are trademarks of Toyo Soda.

STRATEGIC PLANNING

Selecting Matrix and Column for Desalting

The exclusion limit—i.e., the smallest-sized protein molecule that will be excluded from the pores of the matrix—is the most important factor in selecting a matrix for desalting. If the protein is able to permeate the matrix, flow-dependent zone broadening will occur, reducing the resolution. On the other hand, gels of very small pore size generally contain more matrix, and thus have smaller separating volumes. A gel of sufficiently small exclusion limit, but not smaller, should therefore be selected. Another factor to take into consideration is the particle size of the matrix. This should not be too small because the pressure drop over the column is inversely proportional to the square of the particle size; this factor is especially important for gravity flow columns.

The selection of column dimensions (i.e., the bed volume of the column) is influenced by the amount of sample that is to be processed. Columns for desalting are typically short and wide to accommodate large sample volumes and keep the processing time short. Because the zone broadening of the column will have a limited effect on the peak width in desalting as compared to other applications of gel filtration—e.g., protein fractionation—short columns run at high flow rates are used for large-scale desalting. The typical length of columns for laboratory use is ~10 cm.

Columns for desalting vary greatly in size and sophistication, ranging from Pasteur pipets plugged with glass wool to prepacked high-performance GF columns. The scale of

Table 10.9.1 Gel Filtration Matrices Suitable for Desalting^a

Gel type	Exclusion limit		Supplier ^d
	M_r^b	Radius (Å) ^c	
Sephadex G-10	700	7	PB
Bio-Gel P-2	1,800	10	BR
Sephadex G-25	5,000	14	PB
Bio-Gel P6DG	6,000	15	BR
Sephadex G-50	30,000	25	PB
Bio-Gel P-30	40,000	28	BR

^aData as reported by manufacturers.

^bMolecular mass of smallest protein excluded from matrix.

^cApparent radius of smallest spherical protein excluded from matrix, calculated as $R = 0.808 M_r^{1/3}$, where R is the radius and M_r the molecular mass (Hagel, 1989).

^dAbbreviations: BR, Bio-Rad; PB, Amersham Pharmacia Biotech. Addresses and phone numbers of suppliers are provided in APPENDIX 4.

operation for which they are used also varies widely, with applications ranging from micropreparative cleanup of a 50- μ l reduction/alkylation mixture using a 0.9-ml column to industrial de-ethanolization of 12 liters of human serum albumin using a 75-liter column. Incidentally, columns for both these procedures may be packed with Sephadex G-25—of course, of different particle size.

Some gel-filtration media suitable for laboratory-scale desalting are listed in Table 10.9.1, and some empty columns that can be used with these media are listed in Table 10.9.2. A partial compilation of disposable prepacked columns for desalting is found in Table 10.9.3. This table also contains information about spin columns for use with a centrifuge. Spin columns require special techniques, and they are presently not widely used for desalting proteins; therefore no protocols for their use have been included in this unit. More information about spin columns can be obtained from the suppliers.

Selecting Matrix and Column for Protein Fractionation

Suitability of a gel for protein fractionation may be judged from “typical runs” made with synthetic protein mixtures, the results of which are published by suppliers. The solute of interest should ideally be eluted in the first part of the chromatogram, but not too close to the void volume. A matrix with a distribution coefficient (see Basic Protocol 3) of ~0.2 to 0.4 has been reported to yield maximum resolution (Hagel, 1989). This corresponds to an elution volume of ~1.5 to 2 times the void volume. A long bed and a matrix with a large pore volume will increase the peak-to-peak distance between the proteins and thus increase the resolution of the system (also see Critical Parameters). Another way to increase resolution is to employ a GF matrix of small particle size, which will reduce the zone width of the protein bands. The flow rate will also affect resolution. As in the case of desalting, the maximum sample volume that can be processed depends upon the bed volume and the pore volume of the packed bed. Thus, there are many parameters to consider when optimizing protein fractionation by GF. In the worst case, a suitable matrix may have to be found by trial and error.

Commercial prepacked columns for protein fractionation often sacrifice column length and pore volume in favor of small particle size. This results in columns that are optimized for fast separation at the expense of some reduction in the maximum sample volume that can be loaded. Therefore, traditional columns packed in the laboratory are often still a

good alternative for purification of large volumes of sample. If time is critical in the separation (e.g., in proteinase removal), fast column operation is of course advantageous. Matrices of large particle size will permit protein fractionation at low operating pressure. Some matrices suitable for protein fractionation are listed in Table 10.9.4, and some empty columns suitable for use with these media are listed in Table 10.9.5. Table 10.9.6 lists some prepacked columns for protein fractionation. Columns and media are selected on the basis of the difficulty of the separation (i.e., how small the size difference is between the proteins to be separated) and whether scaleup will be required (i.e., the availability of

Table 10.9.2 Empty Columns Suitable for Desalting^a

Column	Minimum bed volume (ml)	Maximum bed volume (ml)	Maximum operating pressure (MPa)	Supplier ^b
PD-10 ^c	1	15	—	PB
Econo-Pac ^c	1	20	—	BR
C series ^d	0	520	0.1	PB
Econo-Column ^d	1	1374	0.1	BR
XK series ^d	2	1884	0.5	PB

^aData as reported by the manufacturers.

^bAbbreviations: BR, Bio-Rad; PB, Amersham Pharmacia Biotech. Addresses and phone numbers of suppliers are provided in APPENDIX 4.

^cPlastic, disposable.

^dBorosilicate glass.

Table 10.9.3 Prepacked Columns Suitable for Desalting^a

Prepacked column	Gel type	Exclusion limit		Maximum sample volume (ml)	Supplier ^d
		M_r^b	Radius (Å) ^c		
<i>Columns for use with a pump</i>					
Hi Trap Desalting	Sephadex G-25 SF	5,000	14	1.5	PB
Fast Desalting HR 10/10	Sephadex G-25 SF	5,000	14	3.0	PB
Econo-Pac P6	Bio-Gel P-6	6,000	15	3.0	BR
<i>Gravity-flow columns</i>					
NAP-5	Sephadex G-25 M	5,000	14	0.5	PB
NAP-10	Sephadex G-25 M	5,000	14	1.0	PB
PD-10	Sephadex G-25 M	5,000	14	2.5	PB
Econo-Pac 10DG	Bio-Gel P-6DG	6,000	15	3.0	BR
<i>Spin columns</i>					
MicroSpin S-200	Sephacryl S-200 HR	250,000	51	0.01-0.05	PB
Bio-Spin 6	Bio-Gel P-6	6,000	15	0.05-0.1	BR
Nick Spin Column	Sephadex G-25 SF	5,000	14	0.075-0.15	PB
Bio-Spin 30	Bio-Gel P-30	40,000	28	0.05-0.1	BR

^aData reported by the manufacturer.

^bMolecular mass of smallest protein excluded from matrix.

^cApparent radius of smallest spherical protein excluded from matrix, calculated as $R = 0.808 M_r^{1/3}$, where R is the radius and M_r is the molecular mass (Hagel, 1989).

^dAbbreviations: BR, Bio-Rad; PB, Amersham Pharmacia Biotech. Addresses and phone numbers of suppliers are provided in APPENDIX 4.

a variety of columns with different diameters that can be packed with bulk media facilitates linear scaleup).

A matrix of small particle size will permit preparation of columns of high efficiency (i.e., low zone broadening) at the expense of high pressure drops and greater cost. Also, smaller particles generally have smaller pore volumes. The net effect of the relationships between these parameters is that the prepacked columns listed in Table 10.9.6 have resolving power of the same order of magnitude, although protein fractionation may be carried out faster on the columns that have particles of small size.

A final factor to consider in choosing a matrix for protein fractionation is the likelihood of matrix-solute interactions (e.g., as in the case of lectins on gels containing glucose-type structures).

Selecting Matrix and Column for Molecular Size Determination

A matrix with high resolving power is recommended for determination of molecular size as well as for protein fractionation (see Table 10.9.4, Table 10.9.5, and Table 10.9.6 for

Table 10.9.4 Gel-Filtration Matrices for Protein Fractionation or Size Determination^a

Gel type	Fractionation range (M_r)	Particle size (μm)	Supplier ^b
Toyopearl HW 40S	100-10,000	25-40	TH
Superdex 30 prep grade	200-10,000	34	PB
Toyopearl HW 50S	500-80,000	25-40	TH
Sephacryl S-100 HR	1,000-100,000	47	PB
Toyopearl HW 55S	1,000-700,000	25-40	TH
Superdex 75 prep grade	3000-70,000	34	PB
Sephacryl S-200 HR	5,000-250,000	47	PB
Superose 6 prep grade	5,000-5,000,000	34	PB
Superdex 200 prep grade	10,000-600,000	34	PB
Sephacryl S-300 HR	10,000-1,500,000	47	PB
Sephacryl S-400 HR	20,000-8,000,000	47	PB
Sephacryl S-500 HR	20,000-30,000,000	47	PB
Toyopearl HW 65S	50,000-5,000,000	25-40	TH
Toyopearl HW 75S	500,000-50,000,000	25-40	TH

^aData as reported by manufacturers.

^bAbbreviations: PB, Amersham Pharmacia Biotech; TH, Toso Haas. Addresses and phone numbers of suppliers are provided in APPENDIX 4.

Table 10.9.5 Empty Columns Suitable for Protein Fractionation or Size Determination^a

Columns	Dimensions ^b (length \times inner diameter, cm)	Maximum operating pressure (MPa)	Supplier ^c
HR series	20 \times 0.5–50 \times 1.6	3–5	PB
G series	25 \times 1.0–100 \times 4.4	3-7	AM
XK series	20 \times 1.6–100 \times 5	0.5	PB

^aData as provided by manufacturers. All columns are made of borosilicate glass.

^bColumn dimensions given here are those suitable for protein fractionation and size determination. Other column dimensions may also be available.

^cAbbreviations: AM, Amicon; PB, Amersham Pharmacia Biotech. Addresses and phone numbers of suppliers are provided in APPENDIX 4.

Table 10.9.6 Prepacked Columns for Protein Fractionation or Size Determination^a

Column type ^b	Fractionation range (M_r) ^c	Particle size (μm)	Dimensions available (length \times inner diameter, cm)	Supplier ^d
Superdex Peptide	100-7,000	13	30 \times 1	PB
HiLoad Superdex 30 prep grade	200-10,000	34	60 \times 1.6 60 \times 2.6	PB
TSK SW 2000	500-60,000	10	30 \times 0.75 60 \times 0.75	TH
HiLoad Sephacryl S-100 HR	1,000-100,000	47	60 \times 1.6 60 \times 2.6	PB
TSK SW 3000	1,000-300,000	10	30 \times 0.75 60 \times 0.75	TH
Protein-Pak 60	2,000-30,000	10	30 \times 0.78	WA
Superdex 75 HR 10/30	3,000-70,000	13	30 \times 1	PB
HiLoad Superdex75 prep grade	3,000-70,000	34	60 \times 1.6 60 \times 2.6	PB
Bio-Sil SEC 125	5,000-100,000	10	60 \times 0.75 60 \times 2.15	BR
G2000SW _{XL}	5,000-150,000	5	30 \times 0.75	TH
HiLoad Sephacryl S-200 HR	5,000-250,000	47	60 \times 1.6 60 \times 2.6	PB
TSK SW 4000	5,000-1,000,000	13	30 \times 0.75 60 \times 0.75	TH
Superose 6 HR 10/30	5,000-40,000,000	13	30 \times 1	PB
Protein-Pak 125	10,000-80,000	10	30 \times 0.78	WA
Glass GF-250	10,000-250,000	4	30 \times 1	DU
Bio-Sil SEC 250	10,000-300,000	10	60 \times 0.75 60 \times 2.15	BR
G3000SW _{XL}	10,000-500,000	5	30 \times 0.75	TH
Superdex 200 HR 10/30	10,000-600,000	13	30 \times 1	PB
HiLoad Superdex 200 prep grade	10,000-600,000	34	60 \times 1.6 60 \times 2.6	PB
HiLoad Sephacryl S-300 HR	10,000-1,500,000	47	60 \times 1.6	PB
G4000SW _{XL}	20,000-10,000,000	7	30 \times 0.75	TH
Glass GF-450	25,000-900,000	6	30 \times 1	DU

^aData as reported by the manufacturers.

^bIn general, matrices of natural polymers (e.g., Superdex, Sephacryl, or Superose) have larger pore volumes than silica-based materials (e.g., TSK SW, Protein-Pak, Bio-Sil SEC, or GF-250/450).

^cSelectivity of materials decreases with increased width of the fractionation range and also with decreased pore volume.

^dAbbreviations: BR, Bio-Rad; DU, DuPont; PB, Amersham Pharmacia Biotech; TH, Toso Haas; WA, Waters. Addresses and phone numbers of suppliers are provided in APPENDIX 4.

matrices and columns suitable for both these applications). However, certain other criteria that are important for selecting a matrix and column for size determination differ from those for protein fractionation. First, when analyzing samples containing several proteins of varying molecular size (e.g., protein digests) a matrix of sufficient separation range is needed. As selectivity inherently limits separation range, optimal separation range may not be achieved using a matrix of high selectivity. Second, because peak broadening is not critical as long as the peaks are resolved, short (e.g., 30-cm) columns operated at high flow rates may be used. However, in situations where the effluent is continuously monitored with a mass-sensitive detector (see Basic Protocol 3), complete resolution between peaks is essential. In this case the criteria used for protein fractionation may be applied.

**BASIC
PROTOCOL 1**

DESALTING (GROUP SEPARATION)

Desalting, or group separation, is used to separate the target protein from low-molecular-mass contaminants. The purpose might be to remove salt prior to lyophilization, or to remove low-molecular-mass reagents—e.g., a reductive cleavage mixture from a preparation step or polybuffer from chromatofocusing. Desalting is also used to change from one buffer to another—e.g., to decrease ionic strength prior to ion-exchange chromatography (UNIT 10.10). The pore size of the matrix is selected to exclude the protein of interest while allowing maximum permeation of the contaminants. This is beneficial in three ways. First, because the protein does not penetrate into the pores, little dilution of the sample will take place except where the sample volume is very small. Second, because the protein is unable to diffuse into the pores, no flow-sensitive zone broadening that results from slow mass-transfer will take place. Hence, the flow rate may be kept high during the desalting. Third, because the protein and the low-molecular-mass solutes are maximally separated, very large sample volumes may be applied before contamination of the protein band becomes evident. Thus, very large sample volumes may be applied and desalted in a short time without dilution of the sample.

The procedure described here uses a closed column system and a pump; annotations explain variations needed in some of the steps if an open-bed column and gravity flow are used. The GF matrix is swelled, the column is packed, and the void volume and total volume of the column are determined in preliminary runs using high- and low-molecular-weight solutes. This information is used to determine the maximum sample volume to apply for the unknown protein to be desalted. Guidelines for estimating approximate sample volumes to use for desalting are also given. Alternative steps are provided for using this procedure either to determine elution volume or to desalt a protein for preparative purposes (or further analysis). Elution volumes must be determined to ascertain the exact void volume and total volume of a column in preparation for desalting. It is also necessary to measure the elution volume for certain protein analyses—e.g., for determining molecular size (also see Basic Protocol 3).

Materials

- GF matrix (Table 10.9.1) or prepacked GF column (Table 10.9.3) with appropriate exclusion limit for protein of interest (also see Strategic Planning and Critical Parameters)
- GF desalting buffer (see recipe)
- Colored marker: 0.2 mg/ml Blue Dextran 2000 or 0.2 mg/ml vitamin B₁₂
- Void marker (see recipe)
- Total liquid volume marker (see recipe)
- Protein sample to be desalted
- 90°C water bath (optional)

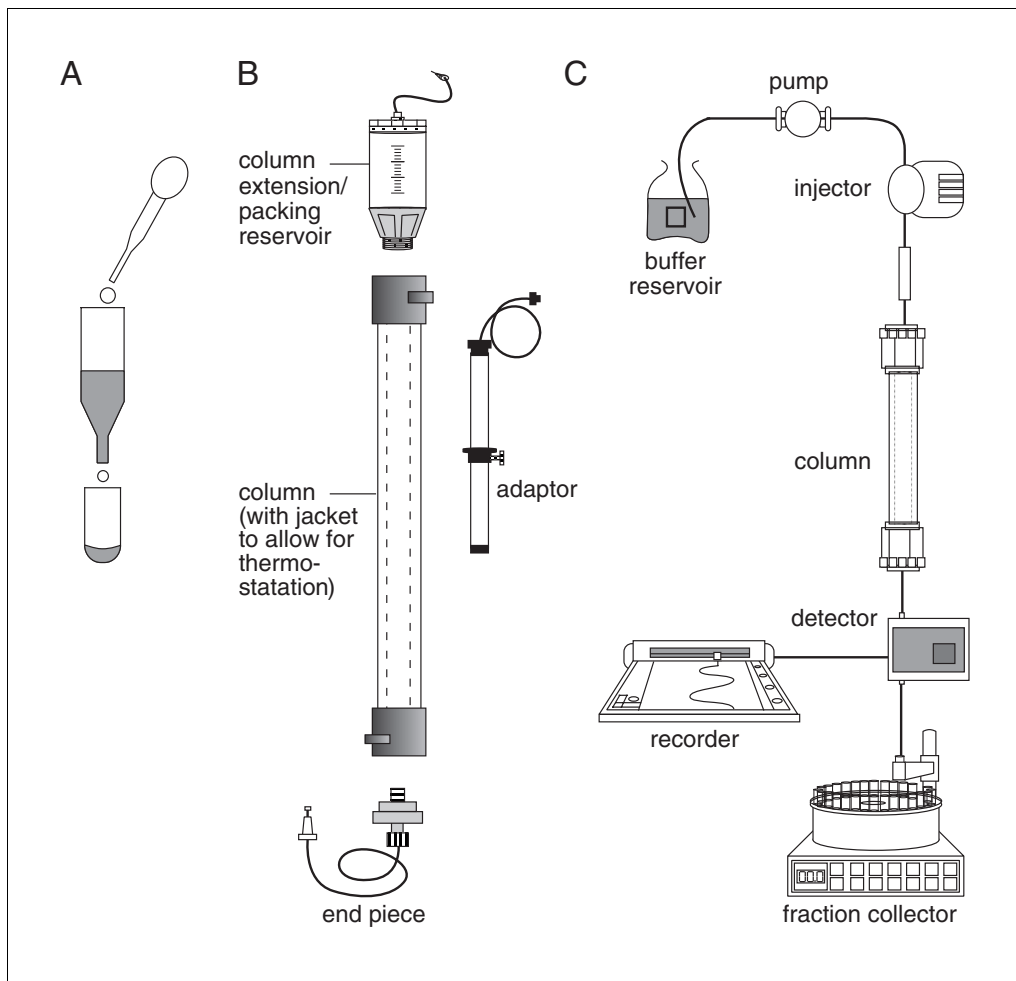


Figure 10.9.1 Equipment used for gel filtration. (A) Simple setup for desalting using an open-bed column made with a Pasteur pipet. (B) Column and attachments. (C) Complete automated chromatography system. Courtesy of Amersham Pharmacia Biotech.

- Buchner or sintered-glass funnel
- GF chromatography column (see Critical Parameters and Table 10.9.2) with column extension (optional), adaptors, and buffer reservoir (Fig. 10.9.1)
- Carpenter's level
- Peristaltic pump (Fig. 10.9.1)
- 0.22- μm filter: any 0.22- μm filter for buffer; protein-compatible for sample
- Detector (optional; Fig. 10.9.1)
- Chart recorder (optional; Fig. 10.9.1)
- Sample applicator: sample application loop (e.g., Superloop, Amersham Pharmacia Biotech) or syringe
- Fraction collector (optional; Fig. 10.9.1)

NOTE: Small-scale desalting of multiple samples may be carried out using spin columns and a centrifuge. At present, spin columns are primarily used for preparing microliter-sized samples of oligonucleotides; however, they are equally useful for small-volume protein samples. As the procedure is different from that described here, manufacturer's instructions should be consulted (see Table 10.9.3 for suppliers).

Prepare gel

If GF matrix is supplied dry:

- 1a. Calculate the amount of dry matrix needed from the swelling factor given by the manufacturer and the approximate bed volume of the column to be packed. Add the calculated amount of dry gel to a volume of GF desalting buffer equal to twice the calculated final gel volume (i.e., twice the approximate bed volume of the column).
- 2a. Carefully stir suspension with a glass rod. Let gel swell overnight at room temperature or for 3 hr in a 90°C water bath, stirring as needed to keep gel in suspension.

Do not use a magnetic stirrer, as this may break fragile gel particles.

- 3a. Allow gel bed to settle, then remove fines and broken beads by decanting the hazy solution. Repeat decantation (4 or 5 times if necessary) until bed settles as a sharp zone, then dilute so that final slurry is 50% (v/v) settled gel and 50% (v/v) GF buffer.

Careful removal of fine particles will prevent the nets that support the column end pieces from becoming blocked, and will also permit a higher flow rate through the column.

If GF matrix is supplied preswollen:

- 1b. Wash gel with an excess of GF buffer on a Buchner or sintered-glass funnel to remove storage buffer.

With some products, the gel slurry is obtained by simply shaking the gel bottle, and this can be used to pack the column without further preparation. If this type of gel is used, skip steps 1b to 3b and proceed directly to step 4.

Gel particles may adhere to the glass wall of the funnel; therefore, to avoid contamination from other types of gels (e.g., ion-exchange media), it is important to ensure that glassware used is clean.

- 2b. If gel contains fines, remove by decantation as in step 3a.
- 3b. Dilute so that final slurry is 50% (v/v) settled gel and 50% (v/v) GF buffer.

Pack column

4. Examine column to ensure that it is clean and that the support nets are undamaged. If necessary, supply column with new support nets.
5. Mount column on a stable laboratory stand, using a carpenter's level to ensure that it is vertical. Equip column with an extension that together with the column will hold the complete volume of the gel slurry (i.e., twice the volume of the settled gel).

If <50% of the column volume is to be packed, there is no need for a column extension.
6. Remove air from outlet tubing and end piece of column by injecting GF buffer into outlet tubing (using syringe or peristaltic pump) until support net is covered with 0.5 cm of buffer. Close outlet of column.
7. Inject GF buffer into inlet tubing of adaptor until net is wetted, while holding outlet upwards to ensure that air easily escapes from net. Place adaptor in a beaker containing GF buffer until it is time to use it.
8. Swirl gel slurry from step 3a or b and pour entire slurry into column along a glass rod tilted toward the inner wall of the column or column extension. Fill the remaining space with GF buffer by carefully pouring buffer along glass rod so as to cause minimum disturbance of the gel layer. Put lid on column extension (or put top adaptor on column).

Small columns (e.g., 1-cm i.d., 5-cm length) to be used for preliminary experiments where optimum resolution is not critical may be quickly prepared by adding the gel slurry to the

chromatographic tube, letting the gel settle by gravity, and washing the bed with one column volume of GF buffer before applying the sample.

9. Fill buffer reservoir with GF buffer and place at a level higher than the pump. Connect reservoir to pump using a large-i.d. tube.

This step is to ensure a free flow of buffer to the pump head.

10. Purge pump with GF buffer and attach outlet from pump to inlet of column or column extension. Open outlet of column and start pump at flow rate recommended by manufacturer for the gel/column combination used. Continue flow until height of gel bed becomes constant (typically 1 to 4 hr).

If no manufacturer's recommendation is available, use a flow rate 50% higher than that to be used in the separation.

If column is to be packed by gravity flow—i.e., by attaching the buffer reservoir directly to the inlet of the column—the actual flow rate must be monitored (e.g., by continuously weighing the effluent, or in the simplest case, by using a graduated cylinder) or controlled by a pump placed after the column. In the latter case it is necessary to check that the free gravity flow exceeds the pump flow—i.e., that the pump is serving to reduce the flow to a constant value.

The flow rate caused by gravity depends upon the hydrostatic pressure, determined by the difference in height between the air inlet in the buffer reservoir and the outlet tubing from the column. However, because the pressure drop over the column increases with the height of the packed bed, the free gravity flow rate will decrease during packing if the hydrostatic pressure is kept constant. Gravity flow rate may be increased by lowering the outlet tube or raising the buffer reservoir.

11. Turn pump off, close column outlet, remove column extension, adjust bed height if necessary, and adjust inlet adaptor so that it is flush with the bed surface.

Avoid trapping air under the support net.

Adjustment of the bed height is recommended only to remove excess gel. This is done by carefully swirling the upper part of the gel surface with a spatula and sucking the excess slurry out with a Pasteur pipet or tubing attached to a pump. If a crater or uneven surface is produced, the gel should be swirled so that a sharp bed surface is obtained before adjusting the inlet adaptor atop the bed. If the bed is too short the column should be repacked, although in most cases the old gel can be reused and added to new slurry prepared as in steps 1a to 3a or 1b to 3b before repacking.

12. Reattach column to pump, open column outlet, and resume flow conditions used in step 10 for 1 hr to stabilize column bed height. Readjust the adaptor so that it is flush with the new bed surface.

In some cases it is desirable to lower the inlet adaptor a few millimeters into the gel surface to make sure that there is contact between adaptor and gel.

Stabilizing the bed at a higher flow rate than that to be used during the separation will prevent the bed from shrinking as a result of “after-packing” during the run. If a gap appears between the adaptor and the bed surface, step 12 must be repeated at an increased flow rate before experiments are carried out. A space between the adaptor and the bed surface will act as a mixing chamber; the effect on separation will depend on the separation problem at hand.

13. Inspect packed bed visually for cracks, trapped air, and particle aggregates by observing the light scattered from a light source held behind the column. Chromatograph a colored marker (2 mg/ml Blue Dextran 2000 or 0.2 mg/ml vitamin B₁₂) as in steps 20 to 23, and ascertain that the zone produced is horizontal and sharp.

Quality of the packed bed need not be as high for desalting as for protein fractionation (see Basic Protocol 2); therefore evaluation of the packed bed may be performed by visual

inspection. However, it is important for it to be free of cracks and for a colored sample to produce a sharp horizontal zone.

Air bubbles would be a major problem and should be avoided at all costs (see Troubleshooting).

If column is to be stored, GF buffer containing 0.02% (w/w) sodium azide or 20% (v/v) ethanol as antibacterial agents should be prepared, then 2 column volumes of this mixture should be pumped through the column. The outlet is then closed and the column stored in the buffer/antibacterial agent at room temperature for several months. See recipe for GF buffer for precautions concerning azide.

Prepare and test the GF system

14. Calculate amount of GF buffer necessary for the run and filter this amount plus a 50% excess through a 0.22- μm filter. Adjust pH if necessary and pour filtered buffer into buffer reservoir.

If the column is going to be used for a prolonged time, it is good practice to continuously deaerate the GF buffer by heating the buffer slightly with heat from a light bulb, or to use a degasser.

15. Assemble GF system according to manufacturer's instructions or Figure 10.9.1, placing detector and chart recorder (if used) in line but leaving column off line. Attach buffer reservoir to pump and purge pump with GF buffer with column off line.

If the system is run by gravity flow alone, the risk of the column running dry may be eliminated by placing the end of the outlet tubing higher than the inlet adaptor or placing a loop of the tubing from the reservoir with its bend at a lower level than the outlet.

Detection of salt may be accomplished in line by conductivity or off line by flame photometry or conductimetry. Colorimetric detection of proteins is readily done at 280 or 254 nm; the latter wavelength provides higher sensitivity. Refer to UNITS 10.6-10.8 for specific methods of protein detection.

16. Connect outlet of pump to column via the injection valve, run system with GF buffer at flow-rate settings to be used for separation, collect fractions with fraction collector, and note actual flow rate of the pump (e.g., by weighing collected fractions or measuring them with a graduated cylinder).

If the concentration of the solute is detected by in-line monitoring, use of a two-channel recorder set at different sensitivities (e.g., 100% and 10% full-scale) is recommended to ensure that the peaks of unknown solutes are not off-scale.

The actual flow rate of the system must be determined with the column installed, because the flow rate from a peristaltic pump or similar device depends upon the back-pressure of the system.

Determine separating volume of column

17. Determine the void volume of the system (V_0) by running a void marker and obtaining the elution volume according to steps 20 to 23.

The sample volume for determination of void volume and total liquid volume (steps 17 and 18) is not critical and may typically be 1% to 2% of the bed volume. However, correction of the calculated elution volume is needed if large sample volumes are applied (see annotation to step 23a). See Figure 10.9.2 for graphical depiction of void and total liquid volume.

The markers for void volume and total liquid volume may be mixed and run together.

Void volume (V_0) is typically 30% to 33% of the bed volume of a nonrigid (e.g., agarose) gel and 36% to 40% of the bed volume of a rigid (e.g., silica) gel.

18. Determine the total liquid volume of the system (V_l) by running a total liquid volume marker and obtaining the elution volume according to steps 20 to 23.

Total liquid volume (V_l) is typically 85% to 95% of the geometric bed volume ($\pi r^2 L$, where r is the inner radius of the column) for a nonrigid gel and 70% to 80% of the bed volume for a rigid gel. Because solute-matrix interactions may occur, it is advisable to check that the total liquid volume is smaller than the bed volume and perhaps run several different total liquid volume markers to confirm the result.

19. Calculate the separating volume of the column (V_i) as $V_t - V_o$.

The separating volume (V_i) is the amount of liquid that will pass through the column from the point where large proteins begin to emerge to the point where small solutes are eluted. It is equal to the pore volume of the bed, which is typically 50% to 65% of the bed volume of a nonrigid gel and 30% to 50% of the bed volume of a rigid gel.

Prepare, apply, and chromatograph the sample

20. Dissolve sample to be desalted in GF buffer. Filter through a 0.22- μm protein-compatible filter.

If possible, let sample solution reach room temperature before being applied to column to reduce viscosity effects. This is more important for large sample volumes.

The influence of the sample volume on the resolution (see Critical Parameters) is related to the need for column efficiency (see Basic Protocol 2). Too large a sample volume may ruin the separation of closely eluting proteins in a separation where high efficiency and narrow peaks are required. This is of primary interest in protein fractionation (see Basic Protocol 2), whereas in desalting, conditions are chosen to yield maximum separation of the protein from contaminants. To this end, column efficiency is less important, and very large sample volumes may be applied. See Critical Parameters for guidelines on optimal sample concentration.

When GF is performed for analytical purposes, the sample concentration should be as small as possible within the sensitivity of the detection method employed. For preparative purposes, on the other hand, the concentration should be as high as possible. The limiting factor in the latter case is the viscosity of the sample divided by the viscosity of the GF buffer (relative viscosity). At relative viscosities >1.5 , the hydrodynamic instability of the trailing edge of the zone may cause distortion of the protein zone (resulting from invasion of less viscous eluant into the more viscous zone; called "viscous fingering"), which will ruin the separation. It is generally safe to use protein concentrations up to 50 mg/ml and in some cases up to 70 mg/ml, whereas concentrations of 100 mg/ml often result in viscous

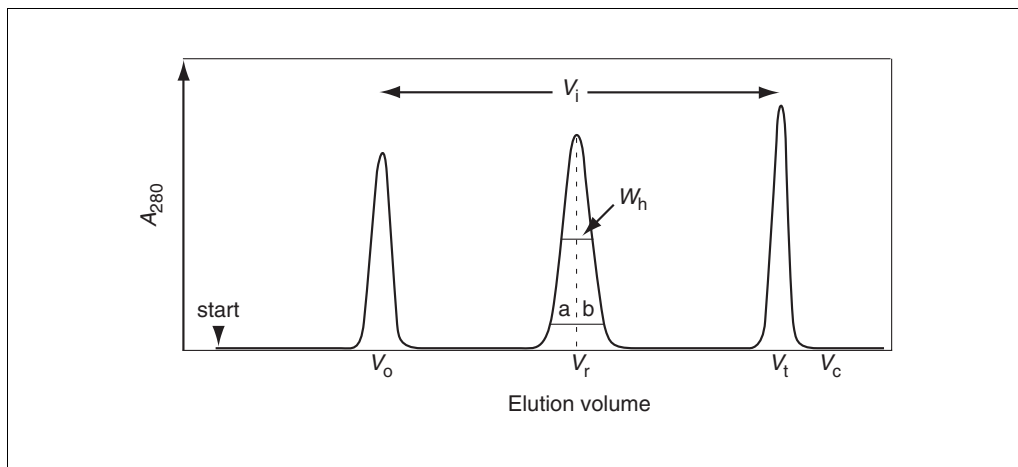


Figure 10.9.2 Chromatogram illustrating the separating volume (V_i) of gel filtration as determined by the void volume (V_o) and the total liquid volume (V_l). The total bed volume (V_c) is equal to the total liquid volume plus the matrix volume. Column efficiency is calculated as $N = 5.54(V_r/W_h)^2$, where N is the efficiency of the column in terms of number of theoretical plates, V_r is the elution volume of the peak, and W_h is the peak width at half peak height. The symmetry of the peak is calculated at 10% peak height as b/a , where b is the width of the trailing part of the peak and a is the width of the leading part of the peak.

fingering. However, when large sample volumes are applied—e.g., in desalting—the low dilution factor will result in a high concentration of protein in the peak, hence lower concentrations than these may be required to avoid viscosity effects.

21. Open outlet from column, turn on pump, and let two bed volumes of GF buffer pass through the column. Turn on detector and recorder and allow baseline to stabilize.
22. Load sample applicator with the appropriate volume of sample for desalting, not exceeding the separating volume determined in step 19. Switch the sample application valve to the load position and put a start mark on the chart recorder paper (or start a stopwatch if no chart recorder is used).

Smaller sample volumes are used for protein fractionation than for desalting, and much smaller sample volumes are used for analysis.

Before injecting, flush injector loop with excess sample to avoid trapping air bubbles, which otherwise may end up in the top of the column.

If the separating volume of the column has not been determined, the maximum sample volume may be estimated as ~50% of the bed volume for nonrigid media and ~30% of the bed volume for rigid media. For a first attempt at desalting, 50% of these volumes should be applied (i.e., 25% and 15% of the bed volume for nonrigid and rigid media, respectively).

If an open-bed column is used, gently add GF buffer to the top of the column with a Pasteur pipet or similar device, then open the outlet from the column and continue adding GF buffer until two bed volumes of buffer have passed through. Let the last portion of buffer penetrate the gel without the gel surface becoming dry; this is facilitated by using a top filter on the bed. Gently apply the sample to the clear gel surface or the top filter using a micropipet (or a Pasteur pipet if large volumes of sample are to be desalted).

- 23a. *To determine elution volume:* Pass GF buffer through the system at the appropriate flow rate and follow detector response with recorder while collecting fractions. Construct a chromatogram and calculate elution volume as the time from the start point to the apex of the peak for the protein (or other solute) multiplied by the flow rate of the pump (as determined in step 16).

If large sample volumes are being applied, the correct elution volume of the peak is obtained by subtracting half the applied sample volume from the apparent elution volume.

For determination of void volume as in step 17, the fractions that emerge between the point where a volume of buffer equal to 25% of the bed volume has passed through the column and the point where buffer volume equal to 50% of the bed volume has passed through the column are analyzed (either in line by monitoring detector response or off line by another method; see annotation below). For determination of the total liquid volume as in step 18, fractions emerging between the point where a buffer volume equal to 75% of the bed volume has passed through and that where 125% has passed through are analyzed. To determine the elution volume of an unknown protein sample (as in determining molecular size; see Basic Protocol 3), fractions emerging between the point where a buffer volume equal to 25% of the bed volume has passed and that where 125% has passed through are analyzed. Allowing 125% to pass through will permit the entire peak to be traced.

If fractions are to be collected for subsequent off-line analysis to determine where elution occurs (e.g., by photometry at 280 nm to detect Blue Dextran 2000 or proteins, or conductivity to trace the salt peak), start and stop the fraction collector to sample fractions at the volumes stated above. The fraction size should be large enough to permit accurate assay of the solute of interest; often 1 ml or 2% of the bed volume (whichever is smallest) is a suitable fraction size. The prefraction (i.e., the amount of liquid eluted from the point of sample application to the start of fraction collection for analysis) should be sampled in a vessel of known weight. The elution volume is calculated from the volume of the fractions by adding half the volume of the peak fraction and the volume of the prefraction, then subtracting half the sample volume. For dilute aqueous buffers (e.g., 0.5 M) a density of 1

may be assumed for converting weight to volume. If the experiment will take a long period of time, it is advisable to minimize evaporation by sealing the fraction cups with Parafilm.

If an open-bed (gravity flow) column is used, carefully add a volume of GF buffer that corresponds to the void volume of the column reduced by half the sample volume—unless the sample volume is <4% of the bed volume, in which case add 1 void volume of buffer. If the column is operating correctly, the effluent collected should contain no protein. Add GF buffer in ≤ 1 -ml increments if the column has a bed volume <50 ml, collect fractions, assay for proteins, and calculate elution volumes as described above.

- 23b. *To desalt a protein:* Pass a volume of GF buffer equal to the void volume reduced by half the applied sample volume through the column, and collect a waste fraction. Place a vessel that can hold 1.5 times the applied sample volume under the column, then add a volume of GF buffer equal to the applied sample volume. Collect the desalted protein.

Depending upon the efficiency of the desalting system, these quantities may have to be adjusted (i.e., the volume collected in the second fraction may need to be decreased in order to preclude contamination from substances that elute later). If system effects (e.g., mixing after the column) are substantial or the column is badly packed (e.g., contains cracks in the bed), it will be necessary to pass more GF buffer through the column than the void volume minus half the sample volume before sampling the protein (see Critical Parameters). Figure 10.9.3 shows a sample chromatogram of optimized desalting of 400 ml hemoglobin.

The desalting step may be continuously monitored by simultaneous in-line detection of the protein zone by absorbance at 280 nm and in-line detection of the salt peak by conductivity.

24. Wash column with ≥ 1 column volume of GF buffer containing an antibacterial agent. Close column outlet and store column.

See recipe for GF buffer for acceptable antibacterial agents.

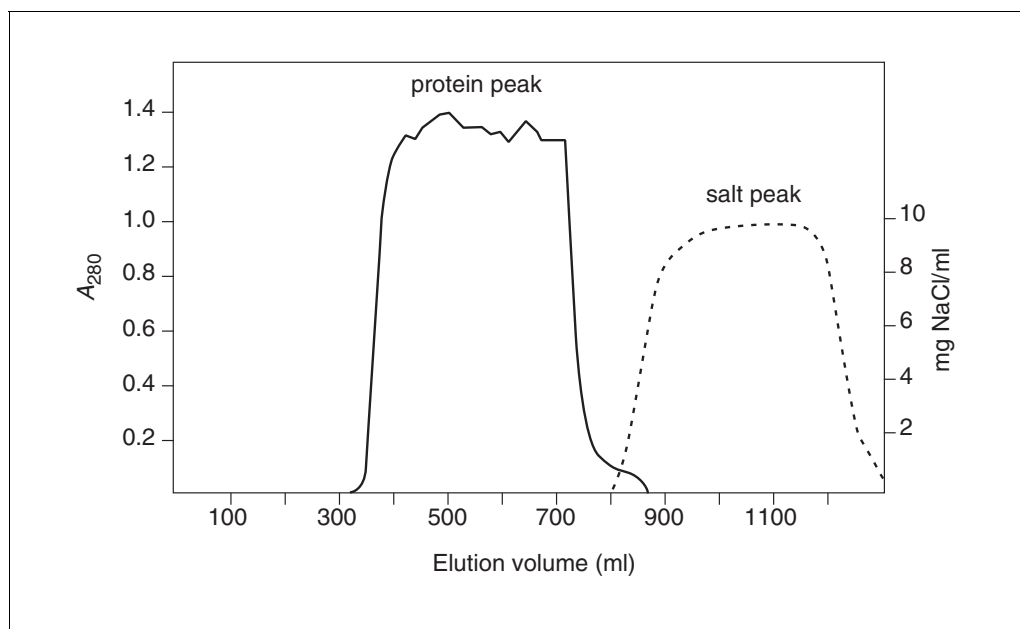


Figure 10.9.3 Chromatogram illustrating desalting of proteins by gel filtration. A 4×85 -cm column packed with Sephadex G-25 is used. The sample consists of 400 ml hemoglobin (protein peak; solid line) in NaCl (salt peak; broken line). Note that the sample volume is close to the pore volume of the packed matrix, i.e., 490 ml. No correction for sample volume has been made. Calculation from the figure yields $V_0 = 560 \text{ ml} - 200 \text{ ml} = 360 \text{ ml}$, which is 31% of the geometric column volume (V_c). Reproduced from Flodin (1961) with permission of the *Journal of Chromatography*.

PROTEIN FRACTIONATION

Protein fractionation refers to the separation of proteins of similar molecular size for purification purposes. The objective may be to separate dimer and higher aggregates from a pharmaceutically active protein or peptide. Requirements for the column and system are therefore substantially higher for protein fractionation than for desalting.

The same materials and methodology described for desalting are used for protein fractionation. However, selection of matrix and column (see Strategic Planning) is more critical in this procedure, as are the quality of the column packing and control over experimental conditions (e.g., temperature and flow rate). The protocol includes steps for determining the column efficiency to evaluate the fitness of a packed column for use in protein fractionation.

Materials

GF matrix (see Table 10.9.4) or prepacked GF column (see Table 10.9.6) with appropriate selectivity for protein of interest (also see Strategic Planning and Critical Parameters)

GF fractionation buffer (see recipe)

Colored marker: 0.2 mg/ml Blue Dextran 2000 or 0.2 mg/ml vitamin B₁₂

Low-molecular-weight marker (e.g., 5 mg/ml acetone or 2 M sodium chloride)

Protein sample to be fractionated

GF chromatography column (see Critical Parameters and Table 10.9.5)

Prepare column

1. Prepare GF matrix as in steps 1a to 3a or 1b to 3b of Basic Protocol 1 using GF fractionation buffer wherever GF buffer is indicated.

It is generally difficult to prepare columns of maximum efficiency using a GF medium with a small (e.g., 10- μ m) particle size. The need to scale up the procedure should be considered when selecting medium and column (see Strategic Planning and Critical Parameters).

The GF fractionation buffer must be selected in light of the method planned for testing purity of the final protein fraction (step 11), as this may put special restrictions on the buffer (e.g., sodium azide absorbs light at 254 nm and will also interfere with the anthrone reaction used for assaying carbohydrates).

2. Pack the column (see Basic Protocol 1, steps 4 to 12).

Because protein fractionation puts high demands on the quality of the packed column, care must be taken in selecting columns of proper design. Columns for protein fractionation are typically 30 to 100 cm long, depending upon the difficulty of the separation and the particle size of the medium used. Columns used must be constructed so that parts coming into contact with buffer and sample are made of appropriate materials. Long tubings and other sources of dead space or mixing chambers should be avoided. The pressure resistance of the column should be compatible with the expected pressure drop of the bed. See Strategic Planning and Critical Parameters for details concerning column selection.

3. Check quality of column packing visually (see Basic Protocol 1, step 13).

4. Assemble and test the system (see Basic Protocol 1, steps 14 to 16).

Place system and column in an environment not subject to temperature variations (e.g., the column should not be exposed to direct sunlight). If the system is to be used in a cold room, the influence of temperature on viscosity must be considered. Zone broadening of proteins and the pressure drop of the system are proportional to the viscosity of the buffer, and the flow rate will have to be adjusted accordingly.

Successful, reproducible protein fractionation requires a constant flow rate. This may be achieved using a peristaltic or high-precision pump whose flow rate is regularly checked. Continuous monitoring of the effluent is necessary to ensure proper sampling of the fraction

of interest. Detection may be carried out at 280 or 254 nm to trace proteins containing aromatic amino acids, or at lower wavelength (e.g., 214 or 206 nm) to measure absorption by peptide bonds. Other parts of a protein fractionation system that might be of use include a fraction collector (although in the simplest case it is possible to collect the fraction of interest by using a switch valve) and a chromatography controller, which permits automatic fractionation of the protein of interest.

Evaluate column

5. Chromatograph a colored marker (2 mg/ml Blue Dextran 2000 or 0.2 mg/ml vitamin B₁₂; see Basic Protocol 1, steps 20 to 23) and ascertain that the zone produced is horizontal and sharp.

A volume of marker solution 1% to 4% of the bed volume should be applied. The zone produced by a colored sample must be horizontal and sharp for packing to be considered acceptable.

If the supplier recommends a different method than this for evaluating column packing, that method should ordinarily be used.

6. Chromatograph a low-molecular-weight marker (e.g, 5 mg/ml acetone or 2 M sodium chloride; see Basic Protocol 1, steps 20 to 23) and construct a chromatogram to determine the column efficiency.

The sample volume applied in this step may contribute considerably to the peak width and result in a falsely low value for column efficiency. It is therefore recommended that the sample volume not exceed 0.9%, 0.5%, and 0.4% of the bed volume for columns packed with GF matrices having average particle diameter of 100 μm , 30 μm , and 10 μm respectively (Hagel, 1985).

7. Calculate the number of theoretical plates per column according to the equation $N = 5.54(V_r/W_h)^2$, where N = number of theoretical plates per column, V_r = elution (retention) volume of peak, and W_h = width of the peak at half peak height. See Figure 10.9.2 for a graphical depiction of these variables.

The efficiency of the column corresponds to the number of theoretical plates under given experimental conditions.

8. Calculate the asymmetry factor of the peak according to the equation $A_s = (b/a)$, where a is the width of the leading part of the peak and b is the width of the tailing part of the peak at 10% peak height.

9. Compare the plate number and asymmetry factor obtained for the column with the acceptance limits for these parameters in the manufacturer's documentation.

If no data for expected column performance is given, calculate the plate height (H) according to the equation $H = L/N$, where L is the length of the packed bed and N is the number of theoretical plates calculated in step 7. As a rule of thumb, H should be about three times the average particle diameter of the matrix for a good column and twice the average particle diameter for an excellent column. Plate heights of five times the average particle diameter or more are generally not accepted for high-resolution chromatography. The asymmetry factor should be 1.0 ± 0.2 .

The flow rate used will influence the broadening of the solute peak; suggested flow rates for GF matrices of different particle size are given in Table 10.9.7.

Perform chromatography and evaluate results

10. Dissolve, apply, and chromatograph protein sample to be fractionated (see Basic Protocol 1, steps 20 to 23).

Start by applying a small fraction of the sample to get a rough idea of the separation and to calculate optimum conditions for the actual run. A sample volume of ~1% to 4% of the bed volume has been found to be optimal for protein fractionation using a medium of 30- μm

particle size, depending upon the processing rate (i.e., amount of substance processed per unit time) of the system (Hagel et al., 1989). As a general rule, a sample volume of 2% of the bed volume can be applied. A chromatogram illustrating protein fractionation where dimers and higher aggregates are separated from the target monomeric form of the protein is given in Figure 10.9.4.

Optimum sample volume depends upon many experimental factors; discussion of optimization strategies have been published (Hagel et al., 1989; Hagel and Janson, 1992). In general, the larger the bead size used, the larger the sample volume that may be applied before the initial resolution is lost. This is due to the fact that smaller bead sizes normally result in more efficient columns that produce narrower peaks; hence the influence of sample volume on peak width is more detrimental.

Table 10.9.7 Practical Velocities for Protein Fractionation or Size Determination^a

Solute molecular mass	Nominal eluant velocities (cm/hr) at a support particle size of (μm):						
	2	5	10	15	30	50	100
1,000	507	203	101	68	34	20	10
5,000	293	117	59	39	20	12	6
10,000	234	94	47	31	16	9	5
15,000	195	78	39	26	13	8	4
50,000	137	55	27	18	9	6	3
75,000	121	48	24	16	8	5	2
100,000	109	44	22	15	7	4	2
200,000	86	34	17	11	6	3	2
300,000	76	30	15	10	5	3	2

^aAdapted from Hagel (1989) with permission from VCH Publishers.

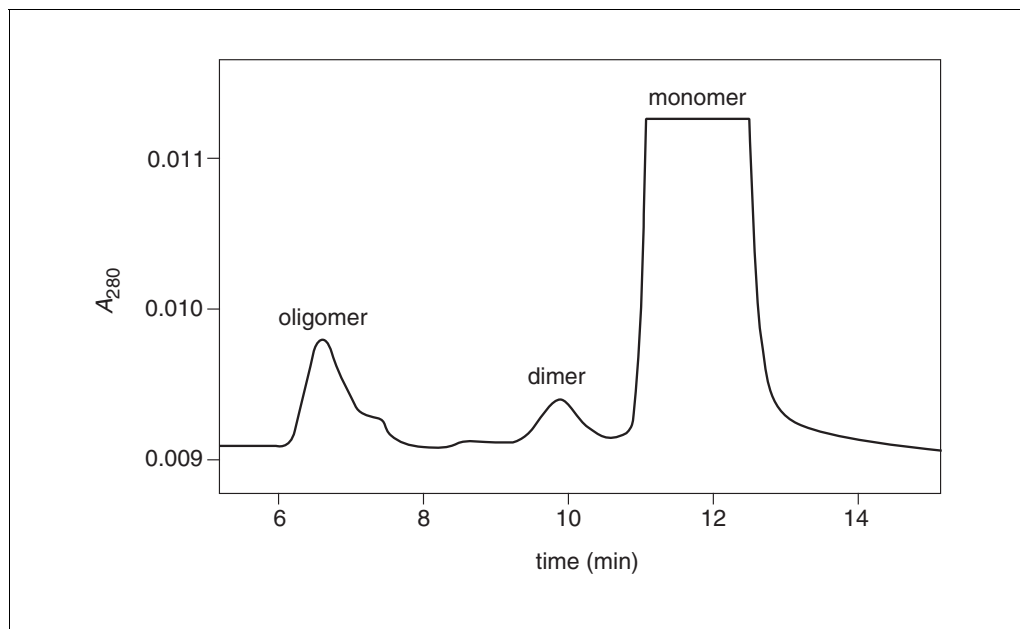


Figure 10.9.4 Chromatogram illustrating protein fractionation by gel filtration. A Superdex 75 HR 10/30 column is used. The sample consists of recombinant human growth hormone. This figure is an enlarged section of the original chromatogram displayed to illustrate complete resolution between monomer and dimer. Reproduced from Hagel (1993) with permission of the publisher. Courtesy of B. Pavlu, Kabi-Pharmacia Peptide Hormones, and H. Lundström, Pharmacia Biotech.

11. Check the purity of the collected fraction(s).

Suitable methods of determining purity include HPLC (see UNITS 10.12-10.14) and electrophoresis (see UNITS 10.2-10.4).

DETERMINATION OF MOLECULAR SIZE

The third important application of gel filtration is in determining molecular size. To accomplish this, a GF column is calibrated with reference samples of known molecular size (see Support Protocol). Gel filtration is then carried out on an unknown protein solute as in Basic Protocols 1 and 2, and the molecular size of that solute is then inferred from the calibration curve, which relates elution volumes or distribution coefficients on the calibrated column to the molecular sizes of the different reference samples. If the relationship between the molecular size and mass is known, the mass of the solute may be calculated. However, because the shapes of biological molecules vary considerably (e.g., proteins may be spherical or ellipsoidal, or in some cases more or less rod-shaped), the calculation of molecular mass (M_r) from elution volume must be done with great care. If the shape of the molecule under study is not known, the experiment will yield only its estimated solvated size. The influence of solute shape on size may be exemplified by the fact that the radius of gyration (one estimate of size) is proportional to $M_r^{1/3}$ for spherical molecules (e.g., some proteins), to $M_r^{1/2}$ for molecules shaped like flexible coils (e.g., dextrans and denatured proteins), and to M_r for rod-shaped molecules (e.g., fibrous proteins and DNA fragments of intermediate length). Thus, gel filtration is very suitable for estimating the molecular size but generally not the molecular mass of unknown solutes.

Materials

- GF matrix or prepacked GF column (see Table 10.9.4, Table 10.9.6, Strategic Planning, and Critical Parameters)
- GF fractionation buffer (see recipe; note variations for size determination)
- GF chromatography column (Table 10.9.5)
- High-precision pump
- Protein sample for size determination

1. Prepare gel, pack column, then assemble and test GF system (see Basic Protocol 1, steps 1 to 16) using GF fractionation buffer wherever GF buffer is indicated.

For high-precision analytical gel filtration, the temperature of the chromatographic bed is often kept constant by using a jacketed column and circulating water from a temperature-controlled water bath. This makes the instrument setup more complicated, and the need for this precaution must be assessed on a case-by-case basis (e.g., if the column is calibrated at every other run, as is possible with high-speed gel filtration, there is little need for temperature control).

A high-precision pump must be used to ensure accurate and constant flow rate, which is of utmost importance for achieving reliable results in molecular size determinations. In addition to a UV detector for determining protein concentration, a detector for viscosity and/or light scattering—e.g., one employing multiple-angle laser light scattering (MALLS)—may be used for assay of molecular size or (indirectly) molecular mass. A mass spectrometry detector—e.g., one employing electrospray ionization mass spectrometry (ESI-MS)—may be used for assay of molecular mass.

If denatured standards are to be used in the calibration to eliminate the influence of molecular shape (see Support Protocol), 6 M guanidine hydrochloride (see recipe for GF buffers in Reagents and Solutions) is used as the GF fractionation buffer in this and subsequent steps.

Once a column has been calibrated (see Support Protocol), it is important to check column performance by determining the elution volume of one of the reference samples used for calibration as part of the routine testing of the system.

2. Determine the void volume (V_0) and the total liquid volume (V_t) for the system (see Basic Protocol 1, steps 17 and 18).

3. Calibrate column (see Support Protocol). Check flow rate during calibration by sampling effluent and weighing fractions.

Calibration of the column (and chromatography of the sample) may be carried out under denaturing conditions to eliminate the effects of molecular shape. Variations in procedure are described in the Support Protocol.

There are no general restrictions on the flow rate as long as the protein of interest separates from contaminating solutes (ideally, baseline separation should be achieved). A high flow rate will not influence the qualitative information sought—i.e., the position of the peak maximum. Thus, qualitative information may be obtained by very fast gel filtration; however, at high flow rates the danger of shear effects on elongated solutes must be considered. If a high-precision pump is used, there is less need to check the flow rate for every run (though in the author's experience this is a good practice because any pump may malfunction). If constant flow rate can be taken for granted, time may be substituted for volume in the calculations.

For both the calibration (see Support Protocol) and the run to determine molecular size of the unknown sample, the elution volume should generally be corrected for the contribution from the sample volume (see Basic Protocol 1, step 23a). However, if the sample volume is kept constant throughout the entire investigation it may arbitrarily be regarded as a part of the dead volume of the system, and no correction is necessary.

4. Apply, elute, and chromatograph unknown sample (see Basic Protocol 1, steps 20 to 23). Prepare a chromatogram and calculate elution volume. Check flow rate by sampling the effluent and weighing fractions.

The same applied sample volume should be used for the unknown sample as for the size standards (see Support Protocol).

If denatured standards were used in the calibration to eliminate the effects of molecular shape (see Support Protocol), the sample must be denatured. Elution is then carried out using 6 M guanidine hydrochloride as the GF buffer.

Often the calculations are performed on a time basis using an integrator or chromatography software (Amersham Pharmacia Biotech).

5. Calculate molecular size of unknown sample components using the calibration graph constructed in performing the support protocol.

If two different detectors were used for the reference standards and the sample (e.g., a refractive index detector for calibration with dextran standards and a UV detector for the protein sample), a correction factor for the difference in volume between the two detectors must be used in the calculations.

The estimate of molecular size will be correct only if there are no sample-matrix or sample-sample interactions (in fact, gel filtration has been used to study these self-association processes). The size estimate will be an apparent gel-filtration size which, together with data about the molecular mass (e.g., from mass spectrometry), will yield information about molecular shape.

6. Continue to run GF buffer through column until the next run is to be made.

It is important to continue flushing the column to maintain column characteristics and avoid the need for recalibration (see Critical Parameters). If the next run is not to be made for a long time, the flow rate may be reduced and buffer recirculated through a filter to the buffer reservoir (see Critical Parameters). Before reusing the column after a long period of rest, the validity of the calibration curve should be checked by running two or three reference samples covering the range of interest and calculating their molecular size from the old calibration curve. The apparent size should agree with the assigned size within $\pm 5\%$ error.

COLUMN CALIBRATION

Molecular size standards for column calibration should ideally be of the same shape and type as the molecules under study (e.g., globular proteins, fibrous proteins, or peptides). In cases where this requirement cannot be met, the column can be calibrated with other suitable reference substances of known size (e.g., dextrans). An apparent size may thus be assigned to the unknown and the actual size then inferred from the relationship between the shape of the different molecules. It is thus possible to use dextran as a reference for the estimation of Stokes radius for globular proteins (Hagel, 1993). However, a general shape parameter governing separation of all types of molecules by gel filtration has not yet been found (Dubin et al., 1990).

When protein standards are used, the calibration curve will reflect small natural variations in the molecular shapes of proteins, which will decrease the precision of the curve. However, the procedure described here is very simple and will readily yield an apparent molecular mass or molecular weight, which in many cases is quite satisfactory.

The influence of protein shape on the calibration curve may be eliminated by disruption of noncovalent bonds with denaturing solvents (e.g., guanidine hydrochloride). Disulfide bonds are broken by reductive cleavage with a reagent such as DTT, and reoxidation is prevented by carboxymethylation. This treatment converts the shape of sample and reference proteins to random coils. The physical size of a random coil is larger than that of a compact molecule of equal mass; therefore a more porous gel is needed to separate the former species. Detergents may also be used to denature proteins. However, it must be noted that a protein-detergent complex is larger than the protein and that ionic detergents will introduce charged groups. Also, information about size or molecular mass obtained using detergent-denatured proteins may not be relevant to the native protein. As these types of treatment are frequently used to prepare proteins for analytical electrophoresis, the reader is referred to *UNITS 10.3 & 10.4* for further details.

The influence of variation in shape on the calibration curve may also be eliminated by using as molecular size standards reference substances from a homologous class of nonprotein polymers—e.g., dextran or polyethylene glycol. Dextran has been used extensively as reference substance for calibration of gel-filtration columns. Calibration can be done by running a number of narrow fractions or one sample of broad size distribution.

Additional Materials (also see *Basic Protocol 3*)

- Calibration standards (see recipe and Table 10.9.8)
- 6 M guanidine hydrochloride (see recipe for GF buffer)
- In-line refractive-index detector

For calibration with native proteins

- 1a. Calibrate the column using protein standards (see Table 10.9.8) as in steps 20 to 23 of *Basic Protocol 1* and construct a chromatogram.

It is practical to mix a number of the calibration standards and chromatograph them simultaneously so that the calibration may be completed in a few runs (see example in Figure 10.9.5).

- 2a. Plot the elution volume (V_r ; i.e., the volume on the chromatogram from the injection point to the apex of the peak for a particular size standard) versus the logarithm of the molecular size, $\log_{10}R$, where R is the molecular radius of the calibration standard.

Table 10.9.8 gives the molecular radii of commonly used calibration standards. A sigmoid curve should be obtained that is approximated by a straight line in the middle section (see Fig. 10.9.5).

If the sample volumes are not held constant, the elution volume must be normalized by setting the injection point at half injected volume.

The calibration curve is often made system-independent by plotting distribution coefficient (K_D) versus $\log_{10}R$. The distribution coefficient is calculated as

$$K_D = (V_r - V_o)/(V_i - V_o),$$

where V_r is the elution volume for a particular peak, V_o is the void volume for the column, and V_i is the total liquid volume for the column.

For substances belonging to a homologous series (e.g., dextrans), it is meaningful to plot elution volume versus the logarithm of the molecular mass ($\log_{10}M_r$). Where substances of

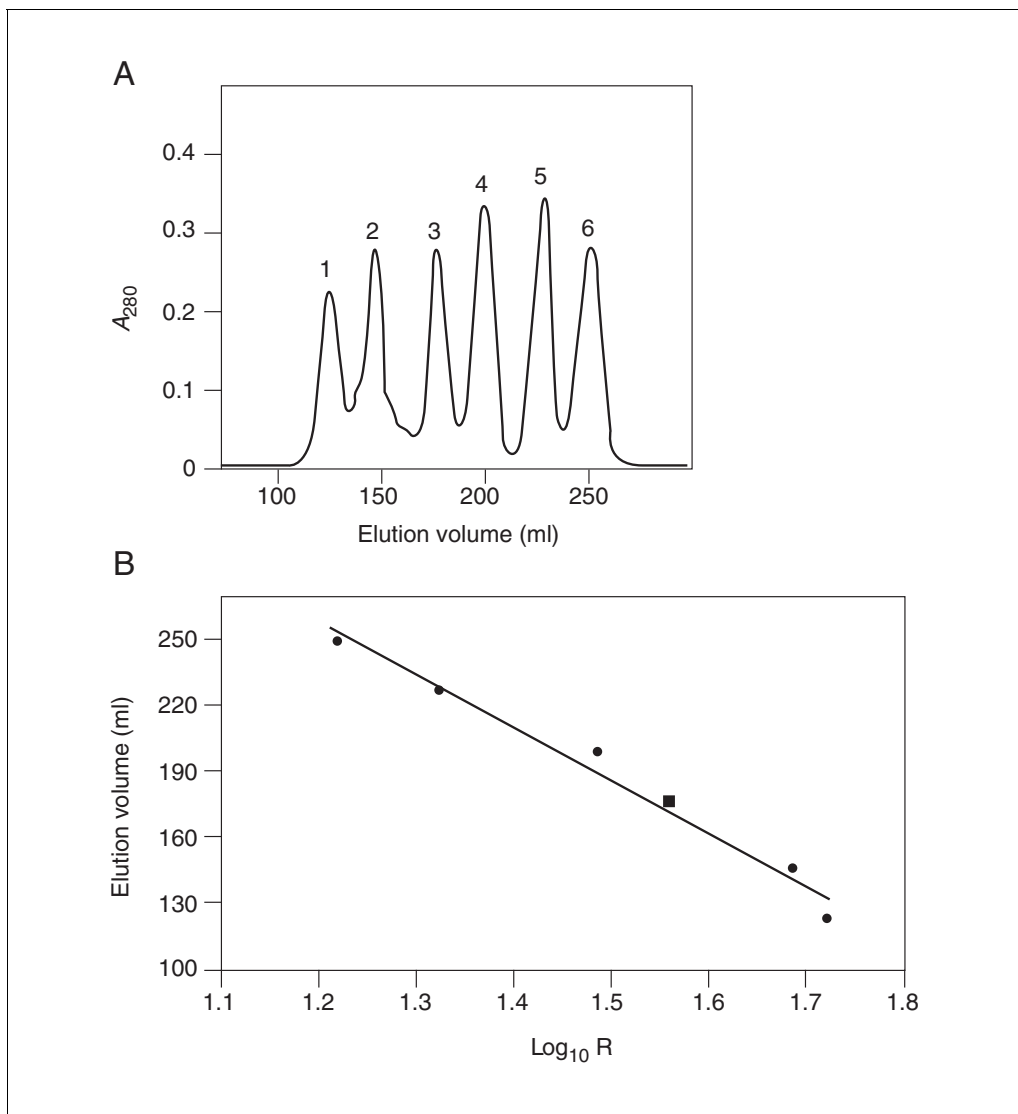


Figure 10.9.5 Calibration of GF column for determination of protein size. A K 26/70 column packed with Sephadex G-200 SF is used. **(A)** Chromatogram obtained for six size standards: 1, catalase; 2, aldolase; 3, bovine serum albumin; 4, ovalbumin; 5, chymotrypsinogen A; 6, ribonuclease A. **(B)** Calibration curve obtained by plotting the logarithms of the molecular radii ($\log_{10}R$) of the standards (x axis) versus the elution volumes for peaks 1, 2, 4, 5, and 6 from panel A (filled circles). The curve was used for evaluating the size of bovine serum albumin (filled square; elution volume corresponds to peak 3 in panel A), giving an apparent size of 34.8 Å. The excellent agreement with the nominal molecular size value of 35.5 Å may be incidental. Repeating the procedure for peak 4 in panel A showed that a deviation of ~3 Å from the nominal value may be expected for proteins of globular shape. Panel A reproduced from Pharmacia Biotech (1991) with permission from the publisher.

different chemical nature are being used to calibrate the column, the preferable way to calibrate the column is by using the hydrodynamic volume, V_h . This is calculated according to the equation $V_h = (\eta)(M_r)/Nv$, where η is the intrinsic viscosity, M_r is the relative molecular mass, N is Avogadro's number (6.02×10^{23} molecules per mole), and v is a shape factor, which for a spherical solute is 2.5.

It is wise to check the accuracy of the calibration by deleting one point at a time from the calibration data set and calculating the size or mass of the protein represented by that point from the calibration curve obtained with the remaining data. The average absolute difference from the nominal value indicates the accuracy of the calibration (see Fig. 10.9.5).

For calibration with denatured standards

- 1b. Denature calibration standards according to a standard denaturing protocol.
- 2b. Carry out steps 1a and 2a above, using a column equilibrated in 6 M guanidine hydrochloride and using 6 M guanidine hydrochloride as the GF buffer in the cross-referenced steps of Basic Protocol 1.

Where denatured standards are used, denaturing conditions must be maintained in chromatography of the sample (see Basic Protocol 3).

For calibration with nonprotein polymer standards

- 1c. Using an in-line refractive index detector, run system until a stable baseline is obtained. Chromatograph a number of dextran calibration standards (see Basic Protocol 1, steps 20 to 23) that will give calibration points surrounding the protein under study, to obtain their elution volumes or distribution coefficients.

The proper number and size of calibration standards to be run may be quickly found by running first the unknown sample and then the necessary calibration samples that will be eluted close to the unknown protein.

Apply each standard individually or as a mixture of two dextrans of high and low molecular mass.

Using a two-channel recorder set at two different sensitivities may prevent the peak from running off scale.

- 2c. Plot the elution volume (V_r) or distribution coefficient (K_D) versus the logarithm of the viscosity radius of dextran ($\log_{10}R_{vis}$).

The viscosity radius of dextran, R_{vis} , is related to the molecular mass (M) according to the equation $R_{vis} = 0.271 \times M^{0.498}$. Because dextran is a polydisperse polymer, the value of M used here is the molecular mass of dextran that corresponds to the peak apex, sometimes denoted M_p .

When two different detectors are used for calibration and chromatography of the sample—e.g., a refractive index detector for dextran standards and a UV detector for the protein sample—the difference in volume between the two detectors must be taken into account in calculation of elution volume.

Another procedure where the entire separation range of interest was calibrated in one run, using an integral calibration method, has been described (Hagel and Andersson, 1991; Hagel, 1993). Unfortunately, this simple and powerful method cannot yet be performed on routine basis because it requires data for the molecular mass distribution of the sample, and this information is, at the moment, not readily available in most cases.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

GF buffers

For desalting: Use a volatile buffer—e.g., 0.05 M ammonium hydroxide, 0.05 M acetic acid, 0.05 M ammonium acetate, or 0.5 M ammonium bicarbonate—unless the step is for buffer exchange, in which case use the desired final buffer.

Distilled (Milli-Q-purified or equivalent) water may be preferred to buffer as an eluant in some cases. This will yield adequate results with noncharged solutes; however, ionic interactions with the small amount of charged groups on the gel surface may disturb the purification of slightly charged molecules when water alone is used. Addition of a volatile bacteriostatic agent—e.g., chlorhexidine—may sometimes provide enough ionic strength to suppress such interactions.

For protein fractionation: Use 0.05 M sodium or potassium phosphate ($pK_a = 2.1, 7.2, 12.4$), 0.05 M Tris·Cl ($pK_a = 8.1$), or 0.05 M sodium or potassium acetate ($pK_a = 4.8$).

The buffering capacity of these buffers is adequate at a pH within ± 0.5 units of the pK_a value. If the protein is to be freeze-dried, use of a volatile buffer or water is advantageous.

It is sometimes necessary to add 0.1 M sodium chloride to completely eliminate ionic interactions. If parts of the system that come into contact with the buffer are composed of steel, the sodium chloride should be replaced with 0.05 M sodium sulfate to prevent halide-induced corrosion. In some cases the ionic strength of the buffer will affect the sample, e.g., by disrupting aggregates. This effect can be eliminated only by changing the ionic strength.

For determining molecular size: Use the same buffers as for protein fractionation but select one that will preclude solute-matrix interactions (see Support Protocol). If the column is to be used for a long period, add preservatives to prevent bacterial growth. Make sure the buffer is compatible with the detection step (e.g., use a buffer with low salt content if mass spectrometry will be used for detection).

If calibration is to be conducted under denaturing conditions, 6 M guanidine hydrochloride is used as the GF buffer. Guanidine hydrochloride is a better denaturing agent than urea.

For all buffers: Add antimicrobial agent if buffer will be used for a long time, adjust the pH of the buffer, filter through a 0.22- μ m filter, then recheck pH and readjust if necessary.

It is also recommended to degas the buffer prior to prolonged use. It is common practice to prepare fresh buffer once a week, but if an antimicrobial agent has been added it may be stored for 1 month or more. Antimicrobial agents used for GF buffers include ethanol added to a concentration of 20% (v/v), sodium azide added to a concentration of 0.02% to 0.05% (w/v), or chlorhexidine added to a concentration of 0.002% (w/v). Ethanol is not recommended for use with Sephadex gels as organic solvents cause the bed to shrink when used with this matrix.

IMPORTANT NOTE: Sodium azide is a very potent antimicrobial agent and therefore widely used. However, handling of azide, especially the dry powder, requires special attention; please refer to information from the supplier for safety precautions. Frequent exposure to the dry powder may be avoided by preparing a stock solution (e.g., 20% w/v) from which aliquots (e.g., 1 ml azide stock solution to 1 liter GF buffer) may be drawn conveniently with a micropipet. Azide may form explosive salts in lead-pipe waste disposal systems; collect the effluent in waste bottles.

Calibration standards

Table 10.9.8 provides a compilation of commercially available molecular size calibration standards and their suppliers. Some of these are available as kits (e.g., from Amersham Pharmacia Biotech). Protein standards should be prepared in the GF buffer that will be used in the molecular size determination. To prolong the life of the column, samples should be filtered through a protein-compatible filter prior to injection. The exact concentration to be prepared depends on the additives present (e.g., sucrose as stabilizer), detector sensitivity, and the zone broadening of the column. Concentrations of proteins in the range from 1 to 5 mg/ml are frequently used for calibration. It is possible to mix several standards to calibrate the separation range of interest in a few runs (it is seldom necessary to calibrate the entire separation range of the column).

Table 10.9.8 also lists some dextran fractions that are useful for calibrating the column in hydrodynamic volume, which makes comparing solutes of different shape easier.

Total liquid volume marker

5 mg/ml acetone, 10 mg/ml sodium azide, 0.1 mg/ml cytidine, 10 mg/ml sodium nitrate, and 2 M sodium chloride have been used for determination of total volume. A negative pulse from injection of water can also be used. In cases where interactions between matrix and solute must be eliminated, deuterium oxide can be used.

Acetone and cytidine are detected by UV absorbance at 280 nm, azide at 254 nm, and nitrate at 254 nm. Sodium chloride can be detected by flame photometry, conductivity, a sodium or chloride electrode, or precipitation of chloride with silver nitrate solution. A negative pulse from injection of water is detected by measuring the absorbance or refractive index of the buffer or by monitoring the conductivity of the effluent. Deuterium oxide is detected with a refractive-index detector.

Void marker

A suitable substance for determining void volume is a protein of the correct size (not too large and not too small), which may be selected by running a series of large proteins and observing which of them start to permeate the matrix. Other substances that have been used include 2 mg/ml Blue Dextran 2000 and DNA.

Blue Dextran 2000 has the advantage of being visible to the eye. Disadvantages are that small amounts of free dye may stick to the gel and a broad permeating peak may be visible in addition to the sharp excluded peak if the pore size of the gel is large enough (e.g., as with Sepharose 4B). This occurs because of the polydispersity of dextran. If the molecule is too large, so-called secondary exclusion effects, in which the sample is excluded from the narrow space where beads intersect, will yield a falsely low void volume. However, this phenomenon is only important in high-precision gel filtration and rarely has to be considered in desalting.

COMMENTARY

Background Information

Although separation of molecules according to size had been noticed in the mid-1950s (Lindqvist and Storgårds, 1955; Lathe and Ruthven, 1956), the discovery of gel filtration (GF) may be attributed to Porath and Flodin (1959), who realized the potential of the technique and were the first to report using it to desalt proteins. Their discovery provided the basis for a new separation technology that rapidly gained widespread acceptance owing to the commercial availability of a suitable cross-linked gel-filtration matrix—Sephadex—in-

roduced in 1959. Remarkably enough, after 35 years this matrix is still one of the premier choices for desalting of proteins (see Table 10.9.1). The use of gel filtration for determination of molecular size of proteins was described later (Andrews, 1970). Gel filtration was also rapidly established as an important tool for determining the molecular mass distribution of polymers in aqueous (Granath and Flodin, 1961) and organic solutions (Moore, 1964), for which suitable methods were lacking at that time. Moore used cross-linked polystyrene-based media for his application, naming the

Table 10.9.8 Molecular Size Standards for Calibration of Gel Filtration Columns

Standard	Molecular mass (M_r)	$\log_{10}(M_r)$	Molecular size (radius, Å) ^a	Supplier ^a
<i>Proteins^b</i>				
Thyroglobulin, bovine thyroid	669,000	5.825	85.0	PB
Ferritin, horse spleen	440,000	5.643	61.0	PB
Catalase, bovine liver	232,000	5.365	52.2	PB
Gamma globulin, bovine	158,000	5.199	—	BR
Aldolase, rabbit muscle	158,000	5.199	48.1	PB
Transferrin	81,000	4.908		SI
Albumin, bovine serum	67,000	4.826	35.5	PB
Ovalbumin, chicken	44,000	4.643	—	BR
Ovalbumin, hen egg	43,000	4.633	30.5	PB
Chymotrypsinogen A, bovine pancreas	25,000	4.398	20.9	PB
Myoglobin, equine	17,500	4.243	—	BR
Ribonuclease A, bovine pancreas	13,700	4.137	16.4	PB
<i>Non-protein standard</i>				
Vitamin B ₁₂	1,350	3.130	—	BR
<i>Dextrans^c</i>				
Dextran T-10	8,100		23	PB
Dextran T-40	23,600		40	PB
Dextran T-70	33,000		50	PB
Dextran T-500	370,000		150	PB
Dextran 1	1,080		9	PC
Dextran 5	4,440		18	PC
Dextran 12	9,980		26	PC
Dextran 25	21,400		39	PC
Dextran 50	43,500		55.3	PC
Dextran 80	66,700		68.4	PC
Dextran 150	123,600		93.1	PC

^aAbbreviations and symbols: —, Not stated by manufacturer; BR, Bio-Rad; PB, Amersham Pharmacia Biotech; PC, Pharmacosmos; SI, Sigma. Addresses and phone numbers of suppliers are provided in APPENDIX 4.

^bThe size of proteins may vary according to the source of information and the methodology used for determination—e.g., a value of 67.1 Å has also been assigned to ferritin (de Haen, 1987).

^cDextran is a polydisperse polymer and values given here—e.g., peak molecular mass—may vary from lot to lot. In cases where peak molecular mass is not stated by the supplier, the column is calibrated as elution volume versus $M = (M_w/M_n)^{1/2}$, where M_w is the weight-average molecular mass and M_n is the average molecular mass. The molecular size of dextran is calculated according to the equation $R_{vis} = 0.271 \times M^{0.498}$, where R_{vis} is the viscosity radius and M is the molecular mass (Hagel, 1989).

technique gel-permeation chromatography. Thus, within a few years after the discovery of gel filtration, applications of the technique basically covered the present range.

During the following years, the focus was directed to investigating the mechanism behind gel filtration. Theories suggesting that the separation was regulated exclusively by diffusion of molecules were soon replaced by the explanation, accepted today, that the process is driven by loss of conformational entropy when free

permeation of the matrix by a solute is hindered by steric factors (Casassa, 1967). Of course, diffusion affects peak width and thus also the degree of separation, but diffusion does not influence the elution volume of a solute. The definite impact of steric exclusion on the separation has led to the proposal to rename the technique “steric exclusion chromatography” (ASTM, 1978), but this term has not gained wide acceptance. Another generic designation sometimes used in the literature is size-exclu-

sion chromatography (SEC). It is important to realize that all of these designations refer to the same basic separation principle, which the majority of protein chemists refer to as gel filtration.

Development of gel-filtration media was focused for a long period on natural polymers or hydrophilic polymers. The late 1970s saw the introduction of silica-based materials, which combined rigidity and small particle size to enable fast gel filtration. A few years later, natural polymers of enhanced rigidity were introduced, which made it possible to pack columns with matrices of 10- μm bead size. To distinguish these media from the traditional type, the designation of high-performance size-exclusion chromatography (HPSEC) was introduced. However, it has been recently pointed out that the advantage of HPSEC media lies not in performance (i.e., traditional media packed in long columns may have equal or even better resolving capacity) but in speed of operation (Hagel, 1992). In addition, none of these materials have the outstanding selectivity shown by some of the traditional matrices (e.g., Sephadex and Bio-Gel P). Therefore, development of matrices for gel filtration has now been directed toward increasing their selectivity, which is done by decreasing the apparent width of the pore-size distribution of the matrix. Development is also geared toward tailoring the apparent pore size of matrices for the separation of molecules in specific interesting areas of application (e.g., for polishing of recombinant peptides). Recently, some of these modern media have become available for separation of peptides and proteins of moderate size (Hagel, 1993).

Even though the theoretical framework for the separation mechanism in gel filtration is well understood, there are still some uncertainties as to which molecular size estimate the elution volume reflects (Dubin et al., 1990). It is generally agreed that Stokes radius is a poor measure of molecular size, especially if different types of molecules are being studied. The hydrodynamic volume seems to be a general size parameter for polymers, globular proteins, and DNA of a certain length (Potschka, 1987). It appears that more work needs to be done to resolve this fundamental issue.

Until recently, development of instrumentation for gel filtration has been modest. High-precision pumps and automated systems for injection, fraction collection, and calculation were introduced more than a decade ago. New equipment and technology being introduced includes multiple-angle laser light scattering

(MALLS) detectors and in-line viscosity detectors, which provide information about the shape and size of molecules, and electrospray ionization mass spectrometry (ESI-MS), which has the power to provide in-line detection of molecular mass. ESI-MS combined with high-resolution gel-filtration columns provides a very powerful system for characterizing protein dimensions. Molecular size of solutes may also be determined with an in-line light-scattering or viscosity detector combined with a concentration-sensitive detector. It is now possible to obtain the exact molecular mass of small-to-medium-sized proteins by connecting an in-line mass spectrometer to the GF column. The advantages of using a mass-sensitive detector are that the system does not need to be calibrated, fluctuations in flow rate are not detrimental, and solute-matrix interactions or solute-solute interactions will not affect the molecular mass obtained. In this case, the column acts as a high-resolution separating device to resolve the sample components prior to determination of molecular mass by mass spectrometry.

Critical Parameters

Optimization and scaleup of desalting

Optimization of desalting starts with selection of a matrix of suitable exclusion limit with maximum pore volume and proper particle size. The bed volume is dictated by the sample volume to be purified. The actual elution protocol to be used depends upon the dispersion of the chromatography system, which in turn is related to the particle size and the column length. The particle size will influence the pressure drop over the column; this may be a limiting factor determining the maximum possible flow rate, the maximum column length, and the possibility of scaling up to columns of large diameters (i.e., the pressure specification is often reduced for columns of large diameter). However, in this case the column length can be traded off for an increase in column diameter to keep the bed volume high (also see Strategic Planning, Selecting Matrix and Column for Desalting).

To maximize the sample volume that can be applied to a column

A gel matrix should be selected from which the protein is excluded (see Strategic Planning) and in which the contaminants are eluted close to the total liquid volume (see Basic Protocol 1). It should also have a small matrix volume.

The maximum sample volume that can be applied to a column is directly related to the pore volume of the packed bed (V_i). The pore volume is proportional to the geometric column volume. As a rule of thumb, the maximum sample volume that may be applied for desalting is 25% to 40% of the bed volume for the materials listed in Table 10.9.1. The lower figure applies to small columns (e.g., <10-ml bed volume) and the higher figure to large columns (e.g., 100-ml bed volume). After the bed length is determined on basis of the acceptable pressure drop (which is proportional to the bed length), the sample capacity is maximized by increasing the diameter of the column. This is done because the applicable volume is directly proportional to the cross-sectional area of the column. Limitations in applicable column length may be eliminated by using columns in series ("stacked columns").

The maximum volume that theoretically can be applied for desalting is equal to the separating volume of the system, which is calculated by subtracting the void volume from the total volume (see Basic Protocol 1 for steps to determine these quantities). To account for extra zone-broadening effects, it is generally advisable to apply not more than 80% of the separating volume or even less—e.g., 50%—for small columns in which eddy dispersion (see discussion below) will contribute proportionally more to the zone broadening.

To optimize the yield of desalting

Sampling of the protein fraction must be adjusted according to the zone broadening of the sample (i.e., the width of the protein zone). For a well-packed column with a plate height (see Basic Protocol 2) roughly equal to two particle diameters, this parameter may be estimated as the sample volume plus the dispersion of the zone that occurs through different pathways in the packed bed—i.e., eddy dispersion. The peak width (w_b) is described by Equation 10.9.1, where V_t is the total liquid volume of the column, d_p is the particle size of the matrix, and L is the length of the packed bed. Thus, for a 5-ml NAP-5 column (0.9 × 2.8-cm bed dimensions) packed with Sephadex G-25 medium-particle matrix, which has a particle size of 175 μm, eddy dispersion adds ~0.6 ml to the sample volume. This is roughly equal to the maximum sample volume (i.e., the separation volume; see Basic Protocol 1) and thus the eluted fraction is twice the sample volume. However, for a much larger 4 × 85-cm column packed with Sephadex G-25 coarse-particle

matrix, which has an estimated particle size of 300 μm, the eddy dispersion adds ~100 ml to the originally applied sample volume of 400 ml, giving a protein zone of 500 ml (see Fig. 10.9.3). If G-25 medium-particle matrix were used, the smaller particle size would in this case theoretically decrease the protein zone from 500 to 475 ml, at the expense of a pressure drop over the column three times higher than with the coarser medium.

$$w_b = 4(V_t) \left(\frac{2d_p}{L} \right)^{\frac{1}{2}}$$

Equation 10.9.1

To minimize the time of the separation

Use the highest possible flow rate for the desalting step. High flow rates are suitable for desalting because eddy dispersion effects are rather insensitive to flow rate. The limiting factors on the flow rate are the rigidity of the matrix and the pressure limitations of the system used. Because media with small pore volume are generally more rigid than those with large pore volume, the disadvantage of a smaller separating volume may be offset by increasing the applicable flow rate to keep productivity high when working with the former type of medium. Incomplete removal of fine particles from the medium prior to packing may reduce the flow rate of the column after use as a result of partial blockage of the outlet filter. Operating the column with the flow direction reversed may help to increase the applicable flow rate in such situations.

The total process time for repetitive runs to desalt large volumes may be further minimized by utilizing the dead time of the void fraction. To do this, samples are added, not after total liquid volume + sample volume + eddy dispersion volume, but after subtracting the void volume from the total process volume (i.e., the next sample is added before the salt peak of the previous sample has emerged). The saving in time may be in the range of 20% to 30%, but this practice can only be recommended for well-understood processes.

$$R_s = \frac{1}{4} \log_{10} \left(\frac{M_{r1}}{M_{r2}} \right) b \left[\frac{1}{\left(\frac{V_o}{V_i} + K_D \right)} \right] \left(\frac{L}{H} \right)^{\frac{1}{2}}$$

Equation 10.9.2

Optimization and scaleup of protein fractionation

There are two basic ways to optimize the resolution of a separation. One is to increase the peak-to-peak distance and the other is to decrease the peak widths. The resolution for gel filtration is described by Equation 10.9.2, where M_{r1} and M_{r2} are the molecular masses of the two sample components being separated, b is the slope of the selectivity curve ($-\partial K_D/\partial \log M_r$), V_o is the void volume of the column, V_i is the pore volume of the bed, L the bed height, and H the plate height of the column (i.e., the term $L/H = N$, the plate number). The impact of these parameters will be described below.

To increase the peak-to-peak distance

1. A matrix of high selectivity should be chosen. The most important factor for optimizing protein fractionation is selectivity, which is the inherent ability of the matrix to resolve proteins of similar size. This factor is a function of the pore-size distribution of the matrix. Selectivity is often illustrated by a plot of the elution volumes of a series of reference substances versus the logarithm of the molecular size or molecular mass of those substances. A function of the elution volume (e.g., the distribution coefficient; see Basic Protocol 3) may be substituted for the elution volume in constructing a selectivity curve. The steeper the selectivity curve (i.e., the greater the absolute value of the slope b , defined above), the better the resolvability of proteins differing in size. The selectivity of different gels can be judged from the fractionation range stated by the manufacturer. Gels with a small fractionation range (in terms of $\log M_r$) will generally have higher selectivity than gels with a large fractionation range.

2. A gel of high pore volume should be selected. The pore volume of a matrix is often related to the particle size. Smaller particles need to resist a higher friction force from the high flow rates that are used with matrices of smaller bead size; therefore the matrix volume is generally larger and pore volume smaller when using these matrices. A traditional soft-gel matrix may have up to 98% pore volume, whereas rigid matrices such as silica may have as low a pore volume as 52% (of the bead volume).

3. A column of sufficient length (L) should be selected. The resolution between peaks is proportional to the square root of the column length. The column length chosen for protein

fractionation is usually optimized with respect to the particle size of the matrix—i.e., a 100-cm column is often chosen where the particle size is 100 μm , a 60-cm column where the bead size is 30 μm , and a 30-cm column where the bead size is 10 μm .

To decrease peak width

1. A material that yields a small void volume (V_o) should be selected. Reducing the void volume will decrease the dispersion of the protein band (as will minimization of all extra column dead volumes). Irregularly shaped or rigid particles generally yield beds with larger void volumes than spherical, nonrigid beads.

2. A material of small particle size should be used, as this will affect the plate height (H ; also see Basic Protocol 3). Zone broadening of proteins within a column is directly proportional to the square of the particle size. Therefore, smaller particles will yield smaller peak widths unless very large sample volumes are applied or large system effects are causing dispersion of zones.

3. The separation should be run at optimum flow rate, as this will affect the plate height (see Basic Protocol 3). Zone broadening is caused by dispersion within the void volume of the chromatographic bed, or by nonequilibrium effects that occur in the separation process itself when the mobile-phase velocity in the extra-particle space is too high as compared with the diffusivities of the proteins and the diffusion distances in the matrix. Flow rates that will allow minimum zone broadening and complete equilibrium will in most cases be impractically low. Table 10.9.7 lists practical flow rates for different solute molecular weights and matrix particle sizes, which represent compromises with regard to these parameters.

Dissolving the sample

Care must be taken to assure that the sample is completely dissolved in the GF buffer. Conditions must be chosen so that spontaneously formed aggregates are dissolved and aggregates that are to be removed are not dissolved merely to reaggregate in the fraction tube. Because gel filtration permits rather free choice of pH, ionic strength, and additives, it is generally not difficult to find suitable conditions.

Solute-matrix interactions

Solute-matrix interactions are generally to be avoided. Even though mixed-mode separations (i.e., separations on the basis of both molecular size and the solute-matrix interac-

tion) have sometimes given good resolution, one must be concerned about the reproducibility of such mechanisms. Solute-matrix interactions may be specific or nonspecific, and the chemical basis may be an ionic or hydrophobic interaction. In the case of specific solute-matrix interactions, the only thing to do is to change the matrix to one based on a different chemistry. Ionic interaction can be suppressed by addition of an electrolyte; hydrophobic interaction can be reduced by adding an organic modifier (e.g., methanol, acetonitrile, or isopropanol) or by reducing the ionic strength of the buffer. Of course, avoidance of solute-matrix interactions is a prerequisite for estimating molecular size using a calibrated column (see Basic Protocol 3 and Support Protocol). Solute-matrix interactions can be uncovered by varying the composition of the mobile phase or temperature and noting the effect. Because ideal gel filtration is driven by entropy, temperature should not affect the elution volume of proteins (unless the temperature change affects protein size or the pore size).

Column maintenance

Column life is drastically reduced by introduction of air into the column, deposition of material on the bed surface, and use of flow rates that exceed the maximum recommended flow rate. To protect valuable columns from air it is advisable to use an air trap, to degas the buffer, or to use a deaerator. Deposition of material is minimized by filtering both sample and buffer through a 0.22- μm filter prior to application. Exceeding the maximum flow rate or pressure (e.g., by failing to compensate for viscosity effects at low temperatures) will result in the bed becoming compressed. This may be avoided by using a flow or pressure sensor to shut down the pump if the limit is exceeded. To maintain the calibration of a column for molecular size determinations (see Basic Protocol 3 and Support Protocol), the flow through the column should never be turned off completely, but may be reduced (e.g., to one-tenth) and the effluent recycled to the buffer reservoir through a filter. The buffer should be treated with preservatives and must be discarded when the column is put back in service. According to the author's experience, the lifetime of a column may be expected to exceed several years and calibration may remain valid for up to 1 year if the appropriate precautions are taken.

A parameter that reflects the possibility of bed compression is the void fraction—i.e., the void volume divided by the bed volume. This

parameter is 0.26 for beds made up of tetrahedral, closely packed spheres and 0.40 for beds made up of randomly packed spheres. Experimentally determined void fractions vary from 0.30 to 0.47 because of the surface and packing properties of the matrix and the particle shape. Void fractions <0.30 may indicate compression of beads.

Flow rate

Flow rate must be carefully controlled, especially if the column is used for characterization of molecular size as in Basic Protocol 3. The actual flow rate of a pump seldom agrees with the nominal flow rate, so it is essential that the flow rate be checked regularly.

Troubleshooting

Gel filtration is a comparatively uncomplicated separation technique and troubleshooting is straightforward. Provided that the critical aspects of the technique (see Critical Parameters) are kept under control, gel filtration can be expected to be relatively trouble-free. Problems that may arise most frequently are compression of the bed and consequent decrease in flow rate, air bubbles in the column, tailing or leading peaks, and peaks that elute in the "wrong place."

Bed compression and decreased flow rate

Compression of the bed most likely results from the maximum flow rate being exceeded during packing (see Critical Parameters). It may also arise from excessive packing of the column, which creates an inhomogeneously packed bed with a denser layer in the lower part of the column that will eventually collapse even at a moderate flow rate. A column found to have a compressed bed must be repacked.

Compression may also result from partial clogging of the inlet filter and the upper part of the bed. To save the column in such a case, the adaptor should be removed, the filter cleaned, contaminated upper matrix removed, the upper part of the bed swirled, and the bed restabilized (see Basic Protocol 1, step 11). If the column still does not perform as expected, it must be repacked.

Decrease in the flow rate often accompanies compression of the bed. If the gel contains fines, these may block parts of the outlet net, causing a decrease in flow rate. Reversing the direction of flow may restore normal conditions. If this is not successful, the column should be repacked after exhaustive decantation of fines (see Basic Protocol 1, steps 3a and 2b).

Air bubbles in column

Small air bubbles in the column may be due to failure to deaerate the buffer or to the introduction of bubbles during sample injection. Moving the column from a cold room directly to a warmer laboratory will also release air bubbles in the column and should be avoided. An attempt may be made to remove the bubbles by pumping deaerated buffer through the column in the reverse direction. Heating the buffer slightly is sometimes successful, as the temperatures of the buffer and gel will equilibrate, allowing the warmer buffer to dissolve some air. This may have to be done at least overnight and perhaps for several days to be fully successful. A large air pocket situated right under the adaptor that has not been taken up by the buffer may be removed using the procedures for readjusting the adaptor to the bed surface (see Basic Protocol 1, step 11).

Tailing and leading peaks

Tailing peaks indicate a large dead volume somewhere in the system, solute-matrix interactions, or a bed that is packed too loosely. To check the system for extra dead volumes, another column or a dummy column of small volume should be packed and a sample run to see if the problem reproduces itself. If it does not, the problem can be attributed to the column, and a sample of a different substance should then be run to ensure that the effect is not caused by solute-matrix interactions. If the tailing problem is still noted, the defect probably lies in the column packing, and an attempt should be made to stabilize the bed at a higher flow rate (see Basic Protocol 1, step 11). If this does not help, the column should be repacked at a higher flow rate. Before doing this, however, it should first be ascertained that the tailing is really seen with the sample to be run under the actual conditions under which it is to be run. Slight tailing of a small solute may be negligible when compared to a large diffusion-regulated zone broadening of a protein. Tailing may also be caused by partial separation of a small amount of a solute of slightly smaller size or slightly different surface properties.

Leading (fronting) peaks indicate solute-matrix interactions or cracks in the bed. Solute-matrix interactions (e.g., ionic-exclusion effects) are detected by running a different substance and/or using a different buffer. If the problem is related to the column packing (i.e., cracks in the bed), the column must be repacked at a lower flow rate. As with tailing peaks, however, the actual sample should first be run

under the actual conditions to be used to determine whether the effect of fronting on the separation is serious. Fronting may also be caused by partial separation of a small amount of slightly larger contaminant.

Peaks that elute in the wrong place

Peaks that elute prior to the void volume indicate ionic-exclusion effects. To suppress these effects, the ionic strength of the GF buffer should be increased. Solutes that elute after the total liquid volume indicate adsorptive effects. These may be masked by adding an organic modifier or by increasing or decreasing the ionic strength of the buffer (see Critical Parameters for discussion of solute-matrix effects).

Anticipated Results

Results from gel filtration are in general very predictable. Guidelines for optimization and scaleup are presented in Hagel et al. (1989) and Hagel and Janson (1992).

Time Considerations

Gel filtration using media of small particle size may generally be carried out in half an hour. Gel filtration using traditional larger-particle matrices are typically run for 5 hr. Desalting is very fast: e.g., 1.4 ml of sample may be desalted in 30 sec.

Preparation (i.e., swelling) of the gel will take 3 hr to overnight, depending on the procedure chosen. Packing of the column takes 1 to 4 hr, and the subsequent stabilization of the packed bed takes 1 hr. Thus, a column is typically prepared in a day, and testing of the packing quality is carried out overnight.

Literature Cited

- ASTM (American Society for Testing and Materials). 1978. Annual Book of ASTM Standards. ASTM D 3016-78. American Society for Testing and Materials, Philadelphia.
- Andrews, P. 1970. Estimation of molecular size and molecular weights of biological compounds by gel filtration. *In* Methods of Biochemical Analysis, Vol. 18 (D. Glick, ed.) pp. 1-53. Wiley-Liss, New York.
- Casassa, E.F. 1967. Equilibrium distribution of flexible polymer chains between a macroscopic solution phase and small voids. *J. Polym. Sci. Polymer Letters Part B* 5:773-778.
- de Haen, C. 1987. Molecular weight standards for calibration of gel filtration and sodium dodecyl sulfate-polyacrylamide gel electrophoresis: Ferritin and apoferritin. *Anal. Biochem.* 166:235-245.

- Dubin, P.L., Kaplan, J.I., Tian, B.-S., and Metha, M. 1990. Size-exclusion chromatography dimensions for rod-like macromolecules. *J. Chromatogr.* 515:37-42.
- Flodin, P. 1961. Methodological aspects of gel filtration, with special reference to desalting operations. *J. Chromatogr.* 5:103-115.
- Granath, K.A. and Flodin, P. 1961. Fractionation of dextran by the gel filtration method. *Makromol. Chem.* 48:160-171.
- Hagel, L. 1985. Effect of sample volume on peak width in high-performance gel filtration chromatography. *J. Chromatogr.* 324:422-427.
- Hagel, L. 1989. Gel filtration chromatography. In *Protein Purification: Principles, High Resolution Methods, and Applications* (J.-C. Janson and L. Rydén, eds.) pp. 63-106. VCH Publishers, New York.
- Hagel, L. 1992. Peak capacity of columns for size-exclusion chromatography. *J. Chromatogr.* 591:47-54.
- Hagel, L. 1993. Size-exclusion chromatography in an analytical perspective. *J. Chromatogr.* 648:19-25.
- Hagel, L. and Andersson, T. 1991. A simple procedure for calibrating a column in solute radius. Presented at the 11th International Symposium on HPLC of Proteins, Peptides and Polynucleotides, October 20-23, Washington D.C.
- Hagel, L. and Janson, J.-C. 1992. Size-exclusion chromatography. In *Chromatography*, 5th ed., part A: fundamentals and techniques (E. Heftmann, ed.) pp. A267-A307. Elsevier/North-Holland, Amsterdam.
- Hagel, L., Lundström, H., Andersson, T., and Lindblom, H. 1989. Properties, in theory and practice, of novel gel filtration media for standard liquid chromatography. *J. Chromatogr.* 476:329-344.
- Lathe, G.H. and Ruthven, C.R.J. 1956. The separation of substances and estimation of their relative molecular size by the use of columns of starch in water. *Biochem. J.* 62:665-674.
- Lindqvist, B. and Storgårds, T. 1955. Molecular-sieving properties of starch. *Nature* 175:511-512.
- Moore, J.C. 1964. Gel permeation chromatography. I. A new method for molecular weight distribution of high polymers. *J. Polym. Sci. Part A* 2:835-843.
- Pharmacia Biotech. 1991. *Gel Filtration Principles and Methods* (5th ed.), Lund, Sweden.
- Potschka, M. 1987. Universal calibration of gel permeation chromatography and determination of molecular shape in solution. *Anal. Biochem.* 162:47-64.
- Porath, J. and Flodin, P. 1959. Gel filtration: A method for desalting and group separation. *Nature* 183:1657-1659.

Key References

Yau, W.W., Kirkland, J.J., and Bly, D.D. 1979. *Modern Size-Exclusion Liquid Chromatography, Practice of Gel Permeation and Gel Filtration Chromatography*. John Wiley & Sons, New York.

Comprehensive overview of theory and use of gel filtration, covering the entire spectrum of applications (of which protein separation is only a small part). Describes in detail analytical determination of molecular size and size distribution of aqueous and nonaqueous polymers.

Hagel, L. 1998. Gel filtration, In *Protein Purification: Principles, High Resolution Methods, and Applications*, 2nd ed. (J.-C. Janson and L. Rydén, eds.) pp. 19-144. John Wiley & Sons, New York.

Describes practical implications of gel filtration theory, with special reference to proteins; a good complement to Yau et al. (1979), and particularly for protein chemists.

Contributed by Lars Hagel
Amersham Pharmacia Biotech AB
Uppsala, Sweden

Ion-Exchange Chromatography

Ion-exchange chromatography separates biomolecules on the basis of charge characteristics. Charged groups on the surface of a protein interact with oppositely charged groups immobilized on the ion-exchange medium. As illustrated in Figure 10.10.1, the charge of a protein depends on the pH of its environment (the operating pH). The pH at which the net charge of a protein is zero (i.e., where the number of positive charges equals the number of negative charges) is known as the isoelectric point (pI). When the operating pH is greater than the pI, the protein will have a net negative charge, and should bind to anion-exchange media, which are positively charged. When the operating pH is less than the pI, the protein will have a net positive charge, and should bind to cation-exchange media, which are negatively charged.

The Strategic Planning section outlines the basic steps in planning and carrying out ion-exchange chromatography to separate proteins. Basic Protocol 1 describes batch adsorption of protein to an ion-exchange medium followed by elution using a step gradient of increasing salt concentration. This technique accommodates a wide range of sample volumes and is most often used in the initial capture stage of protein purification. Batch techniques have minimal system requirements. The Alternate Protocol describes use of a buffer of a different pH to elute via step gradient. Basic Protocol 2 describes adsorption of protein to an ion-exchange medium in a column, followed by elution with a linear gradient. Basic Protocol 2 provides a higher resolution than Basic Protocol 1, and is therefore used in the intermediate purification and final polishing stages of protein separation. Support Protocol 1 describes a pilot experiment to determine initial conditions for batch or column chromatography (i.e., pH required for binding, change in pH or salt concentration required for elution, and available capacity of a medium). Support Protocol 2 describes a means of calculating the dynamic capacity of an ion-exchange column, Support Protocol 3 describes methods for producing continuous gradients of pH and salt concentration to elute proteins from ion-exchange columns, Support Protocol 4 describes

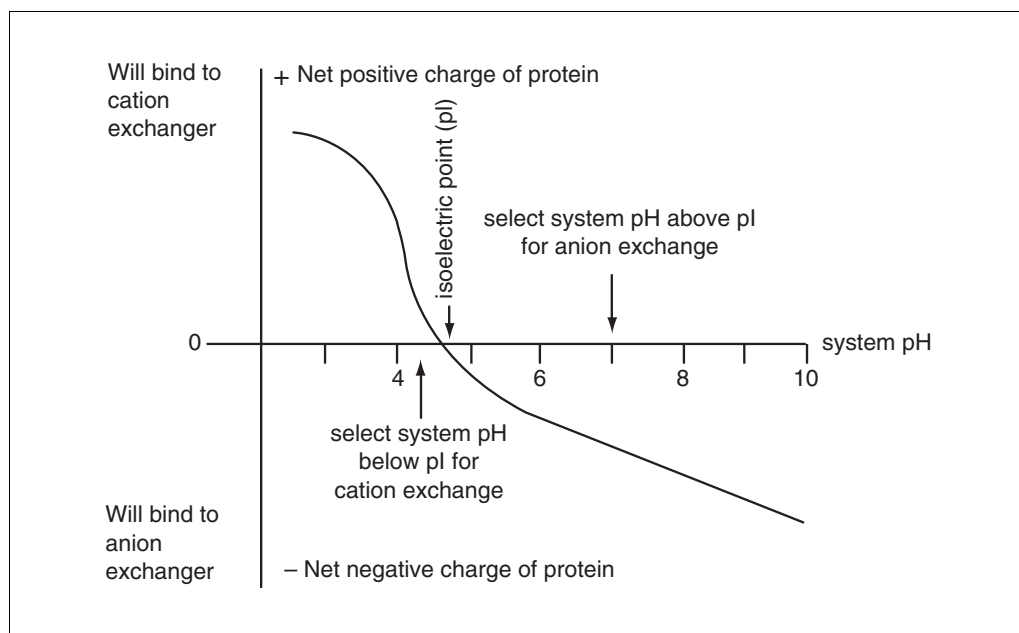


Figure 10.10.1 Net charge of a protein as a function of pH, showing the pH ranges in which protein is bound to anion or cation exchangers. The pH range over which the protein is stable may be only a small fraction of the binding range; this must also be taken into consideration when choosing an ion-exchange medium.

regeneration of used ion-exchange media, and Support Protocol 5 details storage of ion-exchange media.

STRATEGIC PLANNING

There are a number of basic steps in developing an ion-exchange method for protein purification. First, an appropriate ion-exchange medium must be selected, as well as the optimal operating pH and buffer system for the medium and sample (see following discussions on *Selecting an Ion-Exchange Medium* and *Selecting a Buffer System*). It is then necessary to decide whether batch or column chromatography is appropriate given the purity, protein concentration, and physical characteristics of the sample, the sample volume to be used, and the availability of suitable equipment (see following discussion on *Selecting Batch Versus Column Purification*). Next, pilot experiments are conducted to determine conditions for binding and eluting the protein (see Support Protocol 1). It is also necessary to make sure that the capacity of the medium is sufficient to isolate the desired quantity of protein (see Support Protocol 1 and Support Protocol 2). When these initial conditions have been determined, the medium and sample are prepared, the sample is bound to the medium, unbound sample components are washed away, and bound sample components are selectively eluted and collected (see Basic Protocol 1, Basic Protocol 2, or Alternate Protocol). The medium may then be cleaned and regenerated (see Support Protocol 5). Finally, results are evaluated and conditions optimized if necessary (see *Critical Parameters*).

Selecting an Ion-Exchange Medium

A wide variety of ion-exchange media are commercially available, but no miracle medium exists that is best for every protein purification. Criteria for selecting an ion-exchange medium include the specific requirements of the application, the pI and molecular size of the sample components (i.e., target protein and contaminants), and the available equipment (e.g., pumps and columns).

It is necessary to begin by selecting either anion or cation exchange. This requires knowledge of the pI and pH stability of the target protein. If the pI of the target protein is known, an anion-exchange medium with an operating pH above the pI of the target protein or a cation-exchange medium with an operating pH below the pI of the target protein should be selected. If the pI of the target protein is unknown, it is desirable to determine it before beginning. The optimal operating pH can be determined empirically (see Support Protocol 1). Because the pI for most proteins is below pH 7 (Gianazza and Righetti, 1980), it is reasonable to select an anion-exchange medium and an operating pH of 8.5 to start, then evaluate the results and optimize conditions as necessary.

It is also useful to know the pI and binding characteristics of the contaminants present in the protein solution. For example, if the pH of the binding buffer is higher than the pI of a major contaminant, that contaminant will not bind to an anion-exchange medium. If the starting material is crude cell lysate, DNA will be present and generally needs to be removed from the protein. Since DNA is anionic, it binds tightly to anion-exchange media and typically is not eluted with the salt and pH conditions used during protein purification.

Selecting a Buffer System

A buffer system must be selected for the desired pH range. As with selection of an ion-exchange medium, there are several factors that must be considered in selecting a buffer system, including the type of ion exchange to be performed, the pH stability of the sample and the pH range to be used, the required buffering capacity, and, finally, the cost.

Table 10.10.1 Buffers for Ion-Exchange Chromatography^a

pK _a (25°C)	pH range	Buffer ^b	Working concentration (mM)	Temperature factor ^c
<i>Anion exchange</i>				
4.75	4.5-5.0	<i>N</i> -Methylpiperazine	20	-0.015
5.68	5.0-6.0	Piperazine	20	-0.015
5.96	5.5-6.0	L-Histidine	20	
6.46	5.8-6.4	bis-Tris	20	-0.017
6.80	6.4-7.3	bis-Tris propane	20	
7.76	7.3-7.7	Triethanolamine	20	-0.02
8.06	7.6-8.5	Tris-Cl	20	-0.028
8.52	8.0-8.5	<i>N</i> -Methyldiethanolamine	50	-0.028
8.88	8.4-8.8	Diethanolamine	20 (pH 8.4)	-0.025
8.64	8.5-9.0	1,3-Diaminopropane	20	-0.031
9.50	9.0-9.5	Ethanolamine	20	-0.029
9.73	9.5-9.8	Piperazine	20	-0.026
10.47	9.8-10.3	1,3-Diaminopropane	20	-0.026
11.12	10.6-11.6	Piperidine	20	-0.031
<i>Cation exchange</i>				
2.00	1.5-2.5	Maleic acid	20	
2.88	2.4-3.4	Malonic acid	20	
3.13	2.6-3.6	Citric acid	20	-0.0024
3.81	3.6-4.3	Lactic acid	50	
3.75	3.8-4.3	Formic acid	50	+0.0002
4.21	4.3-4.8	Butanedioic acid	50	-0.0018
4.76	4.8-5.2	Acetic acid	50	+0.0002
5.68	5.0-6.0	Malonic acid	50	
7.20	6.7-7.6	Phosphate	50	-0.0028
7.55	7.6-8.2	HEPES	50	-0.0140
8.35	8.2-8.7	BICINE	50	-0.0180

^aInformation from Pharmacia Biotech (1995).

^bAbbreviations: BICINE, *N,N*-bis[2-hydroxyethylglycine]; bis-Tris, bis[2-hydroxyethyl]iminotris[hydroxymethyl]methane; bis-Tris propane, 1,3-bis[tris(hydroxymethyl)methylamino]-propane; HEPES, *N*-[2-hydroxyethyl]-piperazine-*N'*-[2-ethanesulfonic acid].

^cChange in pK_a per °C (i.e., $\partial\text{pK}_a/\partial T$).

Anionic buffers (e.g., acetate and phosphate) are preferred for cation exchange and cationic buffers (e.g., Tris-Cl, ethanolamine, and piperazine) are preferred for anion exchange. It is important to ensure that the buffering ion will have the same charge as the ion exchanger, and hence will not be bound. A constant buffering capacity and pH will thus be maintained during the ion-exchange experiment. Table 10.10.1 lists a variety of buffers useful for anion and cation exchange. Additives to be included in the buffers (e.g., detergents or protease inhibitors) should also carry the same charge as the ion-exchange medium, to preclude binding.

Selecting Batch Versus Column Purification

If the sample volume is large in relation to the size of the pumps and columns available in the laboratory, batch adsorption techniques (see Basic Protocol 1) are appropriate for the capture stage of purification. Batch methods, in which the sample and medium are directly mixed without use of a column, are employed to reduce the sample volume and

Table 10.10.2 Ion-Exchange Media for Batch Adsorption

Medium	Type	Supplier ^a
<i>Dextran bead matrix</i>		
QAE Sephadex A-25	Strong anion	Amersham Pharmacia Biotech
QAE Sephadex A-50	Strong anion	Amersham Pharmacia Biotech
SP Sephadex C-25	Strong cation	Amersham Pharmacia Biotech
SP Sephadex C-50	Strong cation	Amersham Pharmacia Biotech
<i>Microgranular cellulose matrix</i>		
QA-52	Strong anion	Whatman
QA-53	Strong anion	Whatman
SE-52	Strong cation	Whatman
SE-53	Strong cation	Whatman
<i>Polymer-coated silica matrix</i>		
Accell Plus QMA	Strong anion	Waters
Accell Plus CM	Weak cation	Waters
<i>Agarose bead matrix</i>		
DEAE Bio-Gel A	Weak anion	Bio-Rad
CM Bio-Gel A	Weak cation	Bio-Rad

^aFor addresses and telephone numbers of suppliers, see *APPENDIX 4*.

total protein content prior to column chromatography. Batch techniques are also very useful for capture of the target protein from extremely crude samples (e.g., cell lysates and samples containing particulates and/or aggregates). In batch methods the flow and packing characteristics of the medium are of minor importance, but economy and capacity are of primary importance. Ion-exchange media suitable for batch purification at the capture stage of protein purification are presented in Table 10.10.2.

High resolution is required in the intermediate purification and polishing stages of protein purification, and this can only be achieved using column chromatography. However, higher-than-necessary resolution is frequently traded off for greater throughput (i.e., amount of material processed in a defined time). The packing and flow characteristics of the medium are of primary importance when using column chromatography. Columns packed with smaller beads usually require higher operating pressures than columns packed with larger beads. However, smaller beads offer higher resolution, resulting from the increased efficiency that comes with decreasing bead size. Larger beads are desirable for applications requiring higher throughput. A medium with the smallest bead size that will provide the necessary throughput (and whose operating pressure can be accommodated by available equipment) should be selected. Table 10.10.3 lists a variety of media suitable for column purification at various stages in the purification process. See *APPENDIX 4* for contact information of suppliers of these media, all of whom provide technical support for selection and use of specific media.

Capacity of Ion-Exchange Media

The amount of ion-exchange medium required for a particular application will be determined by the protein capacity of the medium (i.e., the amount of protein that can be bound to the matrix) at the chosen pH. The amount of medium required will in turn determine the size of column required. The capacity of an ion-exchange medium is dependent on a variety of factors, including charge and molecular size of the components in the sample and experimental conditions employed.

Table 10.10.3 Ion-Exchange Media for Column Purification^a

Medium ^b	Matrix	Avg. bead diameter (μm)	Use	Form available	Ionic capacity (meq/ml)	Available capacity (mg/ml)	Dynamic capacity (mg/ml)	Conditions for determining dynamic capacity
<i>Amersham Pharmacia Biotech</i>								
Mono Q	PS/DVB	10	I/P	PP	320	—	65	4.5 mg/ml BSA; 50 mM Tris-Cl, pH 8.0; 100 cm/hr; 0.5 × 5-cm column
Mono S	PS/DVB	10	I/P	PP	160	—	75	Human IgG; 100 mM sodium acetate, pH 5.0; 300 cm/hr; 0.5 × 5-cm column
SOURCE 15Q	PS/DVB	15	I/P	PP/BU	ND	—	54	4.5 mg/ml BSA; 50 mM Tris-Cl, pH 8.0; 100 cm/hr; 0.5 × 5-cm column
SOURCE 15S	PS/DVB	15	I/P	PP/BU	ND	—	85	55 mg/ml lysozyme; 20 mM phosphate, pH 6.8; 1 ml/min; 0.63 × 3-cm column
Q Sepharose High Performance	Agarose	34	C/I/P	PP/BU	170	—	108	4.5 mg/ml BSA; 5 mM Tris-Cl, pH 8.0; 100 cm/hr; 0.5 × 5-cm column
SP Sepharose High Performance	Agarose	34	C/I/P	PP/BU	170	—	90	RNase; 100 mM sodium acetate, pH 5.0; 150 cm/hr; 0.5 × 5-cm column
Q Sepharose Fast Flow	Agarose	90	C/I	PP/BU	215	—	120	Human serum albumin; 50 mM Tris-Cl, pH 8.3
SP Sepharose Fast Flow	Agarose	90	C/I	PP/BU	215	—	50	Human IgG; 100 mM sodium acetate, pH 5.0; 300 cm/hr; 0.5 × 5-cm column
SP Sepharose Big Beads	Agarose	200	C	BU	210	—	80	2 mg/ml BSA; 100 mM sodium acetate, pH 5.0; 300 cm/hr
STREAMLINE DEAE	Agarose	200	B/C	BU	205	—	55	BSA; 50 mM Tris-Cl, pH 7.5; 300 cm/hr; 5-cm-i.d. expanded-bed column
STREAMLINE SP	Agarose	200	B/C	BU	170	—	70	Lysozyme; 50 mM phosphate, pH 7.5, 300 cm/hr; 5-cm-i.d. expanded-bed column
Hi-Trap Q ^c								
HiTrap SP ^d								

continued

Table 10.10.3 Ion-Exchange Media for Column Purification^a, continued

Medium ^b	Matrix	Avg. bead diameter (µm)	Use	Form available	Ionic capacity (meq/ml)	Available capacity (mg/ml)	Dynamic capacity (mg/ml)	Conditions for determining dynamic capacity
<i>Bio-Rad</i>								
Macro-Prep high Q	PM	10	I/P	PP/BU	115	—	20	5 mg/ml BSA; 50 mM Tris-Cl, pH 8.3; 100 cm/hr; 1.5 × 10-cm column
Macro-Prep high S	PM	10	I/P	PP/BU	127	—	50	5 mg/ml human IgG; 20 mM sodium acetate, pH 5.0; 100 cm/hr; 1.5 × 10-cm column
Macro-Prep high Q	PM	50	B/C/I	PP/BU	400	—	40	5 mg/ml BSA; 50 mM Tris-Cl, pH 8.3; 100 cm/hr; 1.5 × 10-cm column
Macro-Prep high S	PM	50	B/C/I	P/B	160	—	60	5 mg/ml human IgG; 20 mM sodium acetate, pH 5.0; 100 cm/hr; 1.5 × 10-cm column
<i>Econo-Pac high Q^c</i>								
<i>Econo-Pac high S^f</i>								
<i>TosoHaas</i>								
Toyopearl QAE-550C	PM	100	C/I	PP/BU	330	70 ^g	ND	—
Toyopearl SP-550C	PM	100	C/I	PP/BU	160	110 ^h	ND	—
Toyopearl SUPER Q-650C	PM	100	C/I	PP/BU	280	129 ^g	ND	—
Toyopearl SUPER Q-650M	PM	65	C/I/P	PP/BU	240	143 ^g	ND	—
Toyopearl SUPER Q-650S	PM	35	I/P	PP/BU	240	126 ^g	ND	—
Toyopearl SP-650C	PM	100	C/I	PP/BU	150	45 ^h	ND	—
Toyopearl SP-650M	PM	65	C/I/P	PP/BU	150	50 ^h	ND	—
Toyopearl SP-650S	PM	35	I/P	PP/BU	150	50 ^h	ND	—
<i>Waters</i>								
Protein-Pak Q HR	PM	8	P	PP	200	—	65	BSA; 20 mM Tris-Cl, pH 8.2
Protein-Pak SP HR	PM	8	P	PP	225	—	40	Cytochrome c; 25 mM MES, pH 5.0
Protein-Pak Q HR	PM	15	I/P	PP/BU	200	—	75	BSA; 20 mM Tris-Cl, pH 8.2
Protein-Pak SP HR	PM	15	I/P	PP/BU	225	—	40	Cytochrome c; 25 mM MES, pH 5.0
Protein-Pak Q HR	PM	40	B/C/I	BU	200	—	60	BSA; 20 mM Tris-Cl, pH 8.2
Protein-Pak SP HR	PM	40	B/C/I	BU	225	—	20	Cytochrome c; 25 mM MES, pH 5.0

continued

Table 10.10.3 Ion-Exchange Media for Column Purification^a, continued

Medium ^b	Matrix	Avg. bead diameter (µm)	Use	Form available	Ionic capacity (meq/ml)	Available capacity (mg/ml)	Dynamic capacity (mg/ml)	Conditions for determining dynamic capacity
<i>BioSeptra</i> Q-Hyper D	CP	10	P	PP	250	—	90	5 mg/ml BSA; 50 mM Tris-Cl, pH 8.6; 150 cm/hr; determined at 10% breakthrough
S-Hyper D	CP	10	P	PP	110	—	105	5 mg/ml lysozyme; 50 mM acetate, pH 4.5; 150 cm/hr; determined at 10% breakthrough
Q-Hyper D	CP	20	I/P	PP/BU	250	—	90	5 mg/ml BSA; 50 mM Tris-Cl, pH 8.6; 150 cm/hr; determined at 10% breakthrough
S-Hyper D	CP	20	I/P	PP/BU	100	—	105	5 mg/ml lysozyme; 50 mM acetate, pH 4.5; 150 cm/hr; determined at 10% breakthrough
Q-Hyper D	CP	35	I/P	PP/BU	250	—	40	5 mg/ml BSA; 50 mM Tris-Cl, pH 8.6; 150 cm/hr; determined at 10% breakthrough
S-Hyper D	CP	35	I/P	PP/BU	100	—	100	5 mg/ml lysozyme; 50 mM acetate, pH 4.5; 150 cm/hr; determined at 10% breakthrough
Q-Hyper D	CP	60	C/I	BU	200	—	100	5 mg/ml BSA; 50 mM Tris-Cl, pH 8.6; 150 cm/hr; determined at 10% breakthrough
S-Hyper D	CP	60	C/I	BU	100	—	100	5 mg/ml lysozyme; 50 mM acetate, pH 4.5; 150 cm/hr; determined at 10% breakthrough

^aAbbreviations: B, batch; BSA, bovine serum albumin; BU, bulk; C, capture; CP, composite polymer, I, intermediate purification; i.d., inner diameter; MES, 2-[N-morpholino]ethanesulfonic acid; ND, not determined; P, polishing; PM, polymethacrylate; PP, prepacked; PS/DVB, polystyrene/divinylbenzene.

^bFor addresses and telephone numbers of suppliers, see *APPENDIX 4*.

^cPrepacked column suitable for use with syringe. For medium specification, see Q Sepharose High Performance.

^dPrepacked column suitable for use with syringe. For medium specification, see SP Sepharose High Performance.

^ePrepacked column suitable for use with syringe. For medium specification, see Macro-Prep high Q.

^fPrepacked column suitable for use with syringe. For medium specification, see Macro-Prep high S.

^gConditions for determining available capacity: 5 mg/ml BSA; 50 mM Tris-Cl, pH 8.7.

^hConditions for determining available capacity: 5 mg/ml lysozyme; 20 mM phosphate, pH 6.0.

Analysis of Proteins

10.10.7

The number of charged groups on a specified amount of matrix that can participate in ion exchange is termed **ionic** or **total** capacity. However, ion-exchange media are porous and the ion-exchange groups are usually distributed throughout the bead. The amount of protein that will bind is dependent on the ability of the sample components to diffuse into the bead, and this ability is related to the molecular weight and shape of the components and the pore structure of the matrix (see *UNIT 10.9*). In addition, large molecules such as proteins and nucleic acids can bind to multiple ion-exchange groups (or sterically hinder other molecules from binding) depending on the shape and surface-charge distribution of the protein, as well as the spatial orientation of the exchange ligands on the matrix. These factors are taken into account in the available and dynamic capacity of a matrix or column.

Available capacity is defined as the amount of a specific protein that will bind to an ion-exchange medium at a defined pH, salt concentration, and sample concentration in a batch purification process. When medium is packed in a column, the amount of a specific protein that will bind is dependent on many of these factors, as well as on the column dimensions and flow rate. The capacity of a column operating under defined conditions is termed the **dynamic capacity**. The dynamic capacity is typically less than the available capacity for the same sample under the same operating conditions.

Ion-exchange media in which the charge state of the ligand is sensitive to the operating pH are termed weak ion exchangers. The usable pH range of weak ion exchangers is narrower than that of strong ion-exchange media, in which the charge on the ligand used for exchange is constant (i.e., insensitive to pH). Ion-exchange media with DEAE (diethylaminoethyl) or CM (carboxymethyl) groups as the ligand for exchange are considered weak ion exchangers. Because the ionic capacity of weak ion-exchange media is variable, strong ion-exchange media are recommended for pilot experiments (see Support Protocol 1).

When selecting ion-exchange media for batch or column purifications, the amount of medium should be estimated from the manufacturer-supplied capacity data for a protein similar in size to the target protein to be purified. Be aware of what type of capacity (i.e., ionic, available, or dynamic) is being reported and the experimental conditions under which the reported capacity was determined (see Table 10.10.3). A general recommendation is to begin pilot experiments at ~10% of the published capacity data, then determine the capacity for your system.

Preparing Solutions

All aqueous solutions for use in ion-exchange chromatography should be prepared with distilled or deionized water (Milli-Q purified or equivalent). Solutions and samples for column chromatography should be filtered, degassed, and equilibrated to the appropriate temperature prior to use. Special attention must be given to any additives to be included in the buffer to be sure that they will not bind to the column (see preceding discussion on Selecting a Buffer System). If the chosen additives increase the viscosity of the mobile phase, the flow rate should be decreased.

Chromatography Systems

Column techniques typically result in higher resolution than batch techniques, and should be used during the intermediate purification and polishing stages of a purification procedure. However, to achieve high resolution, an appropriate chromatography system is required. A chromatography system may be described simply by the functional operations required to perform a chromatographic separation. The minimum functions required (Fig. 10.10.2) include a buffer delivery system capable of gradient formation, a mechanism for introducing sample into the system, the column itself (properly packed with the appro-

Table 10.10.4 Scale-Up Calculations for Ion-Exchange Column Chromatography

Scale	1×	10×	100×
<i>Sample</i>			
Sample concentration (mg/ml)	1	1	1
Sample load (mg)	10	100	1000
Sample volume (ml)	10	100	1000
<i>Column</i>			
Column diameter (cm)	0.5	1.6	5
Cross-sectional area (cm ²)	0.2	2	20
Bed height (cm)	5	5	5
Bed volume (ml)	1	10	100
<i>Flow rate</i>			
Volumetric flow (ml/min) ^a	2	20	200
Linear velocity (cm/hr)	300	300	300
<i>Gradient</i>			
Gradient volume (ml)	20	200	2000

^aLinear flow (cm/hr) multiplied by the cross-sectional area of the column (cm²) equals the volumetric flow rate (ml/hr); cm/hr × cm² = cm³/hr = ml/hr.

appropriate ion exchanger), an in-line UV monitor suitable for use at 280 nm, a chart recorder, and a fraction collector to preserve the separation. Additional unit operations may include system controllers for automated operation as well as data-acquisition systems. The equipment required to perform each of these unit operations is ultimately determined by the chromatography medium and column selected. Specifically, the operating pressure and flow characteristics of the packed column will be the most important factors for selecting equipment to fit the column, or for selecting columns to fit existing equipment.

It is highly recommended that prepacked columns be used and that the manufacturer be consulted prior to purchase as to equipment requirements for proper operation. This will save time and money and increase the probability of success. If columns are to be packed in the laboratory, the manufacturer's packing instructions for the medium should be followed and column efficiency and function should be tested with representative standards prior to use with the actual sample. There is no single packing methodology that is optimal for all types of media. Table 10.10.3 lists a variety of ion-exchange media available in prepacked columns and/or in bulk for packing in the laboratory. Several manufacturers offer inexpensive disposable ion-exchange columns, suitable for many applications, that can be operated using only a syringe.

Scale-Up Conditions

Ion-exchange separations are typically developed and optimized on a small scale before the entire sample is committed. It is thus important to choose an ion exchanger that will allow simple and convenient scale-up, so that methods established on a small column can be applied more or less directly to a larger one.

Scaling up in ion-exchange chromatography can be approached in two different ways—either the same ion-exchange medium can be used in a larger-diameter column with the same bed height, or a different medium having the same charged group immobilized on a matrix with a higher throughput can be substituted for the medium used in the small-scale separation.

The former approach, using the same medium throughout, is the simplest. Scaling up a separation with the same medium and bed height used in the small-scale optimization is achieved by employing simple scale factors to adapt the flow, gradient volume, and sample load to the increased volume of a larger-diameter column. The scale factor for the required increase in flow rate is the ratio of the cross-sectional area of the larger column to that of the smaller column (i.e., the factor by which the total flow of solution through the column must be increased so that the same linear velocity used in the small-scale experiment is maintained). The increase in gradient volume and sample loading are directly proportional to the increase in column volume. Table 10.10.4 shows an example of changes that must be made to various parameters for a 10- and 100-fold increase in scale.

BATCH ADSORPTION AND STEP-GRADIENT ELUTION WITH INCREASING SALT CONCENTRATION

This protocol describes adsorption of a protein to an ion-exchange medium by direct mixing of the protein-containing sample and the medium (in this case, an anion-exchange gel). This batch purification is in contrast to column purification (see Basic Protocol 2), in which the gel is packed in a column and the sample is passed through it. Batch adsorption is particularly useful where large volumes of sample are involved, which is usually the case in the capture stage of protein purification (*UNIT 8.1*). The equipment used here is sufficient for sample volumes up to 1.5 liter. Larger samples may be divided and processed with parallel apparatuses, or processed sequentially on a single apparatus. Different-sized funnels and side-arm flasks may be used when appropriate.

Once the starting conditions have been determined (see Support Protocol 1), the range of pH and salt concentration where binding and elution occur is well defined. Conditions for batch adsorption and elution used in this procedure are based on the results, presented in Figure 10.10.3, of a test tube pilot experiment performed as described in Support Protocol 1. Buffer volumes and equipment were chosen on the basis of a sample volume of 1 liter containing 50 to 1000 mg of total protein.

Elution of sample components adsorbed to an ion-exchange medium can be achieved either by increasing the salt concentration of the mobile phase, as in this protocol, or by changing its pH (see Alternate Protocol). In batch techniques, discontinuous (step) gradients are used for elution. Column chromatography (see Basic Protocol 2) allows the use of either step or continuous gradients. Formation of continuous gradients requires special equipment (see Basic Protocol 2 and Support Protocol 3).

The number of incremental steps to include in a step gradient must be judged empirically based on the resolution achieved. Ideally (as in the procedure described here), three steps should be used in which all weakly bound materials are desorbed and eluted in the first gradient step using wash buffer, the target protein is eluted in the second gradient step using elution buffer, and all strongly adsorbed materials are eluted in the third gradient step using regeneration buffer. A minimum increment of 50 mM NaCl is recommended for experiments using step-gradient elution with increasing salt concentration. Steps in the gradient may be added, omitted, or modified as determined empirically.

Materials

QAE Sephadex A-25 (Amersham Pharmacia Biotech) or equivalent
anion-exchange gel

Binding buffer: 20 mM Tris·Cl, pH 7.5 (or other buffer as determined empirically;
see Support Protocol 1)

Protein sample to be purified

Wash buffer: 20 mM Tris·Cl (pH 7.5)/100 mM NaCl (or other buffer/salt solution
as determined empirically; see Support Protocol 1)

Elution buffer: 20 mM Tris-Cl (pH 7.5)/350 mM NaCl (or other buffer/salt solution as determined empirically; see Support Protocol 1)

Regeneration buffer: 20 mM Tris-Cl (pH 7.5)/2 M NaCl (also see Support Protocol 5)

Boiling water bath (optional)

500-ml sintered-glass filter funnel, medium porosity

Three 2000-ml side-arm flasks

Conductivity meter

Swell and equilibrate ion-exchange gel

1. Add 10 g QAE Sephadex A-25 to 1 liter binding buffer and swell 2 days at room temperature or 2 hr in a boiling water bath.

The volume of swollen gel will be ~70 ml.

2. Mount 500-ml sintered-glass filter funnel on 2000-ml side-arm flask and apply suction.

Centrifugation (~5000 × g, 1 min) may be used instead of filtration to collect the gel, and may be preferred for processing small sample volumes. It is also possible to simply allow the gel to settle under gravity, then decant or aspirate the supernatant; this may be preferable for very large-scale operations.

3. Gently swirl swollen gel to resuspend. Pour gel slurry into funnel as fast as the fluid level in funnel permits, allowing buffer to collect in flask. Continue until all gel has been collected. Release suction after all buffer has been removed.

IMPORTANT NOTE: *Never use a magnetic stir-bar to resuspend chromatography medium. This can damage beads by grinding them against the bottom of the vessel, resulting in fine particles that can slow the filtration process.*

4. Add 200 ml binding buffer to funnel and resuspend gel using a stirring rod. Allow resuspended gel to stand for 5 min, then apply suction to remove buffer.

5. Repeat step 4 at least three times to ensure equilibration.

The gel is considered equilibrated when the pH and salt concentration of the eluant are the same as those of the binding buffer. If the procedures here are followed, there is little need to measure pH and salt concentration of the eluant because the gel is swollen and washed in binding buffer and the chance of error is very small. The worst-case scenario is that the target protein will not bind (see Troubleshooting). The repeated washings are more for the removal of fines than for adjustment of pH or salt concentration.

6. Add enough binding buffer (~100 ml) to produce an ~50% (v/v) slurry and allow gel to stand in funnel until sample has been prepared.

Adsorb sample to gel

7. Adjust pH and salt concentration of protein sample to initial optimal values.

For a pilot experiment to determine optimal initial values for pH and salt concentration, see Support Protocol 1.

Samples prepared by salt precipitation may require desalting prior to ion exchange (see UNIT 10.9). The sample can be dialyzed against the binding buffer. Also, the salt concentration of a sample can be reduced by adding distilled water until a desired salt concentration is achieved, as measured by a conductivity meter.

8. Combine gel and sample in 2000-ml beaker or wide-mouth flask.

The ~50% gel slurry is swirled and poured from the funnel; transfer may be aided with a rubber policeman, spoon, stirring rod, or wash bottle containing binding buffer.

9. Gently mix sample and gel by swirling every 15 min or shake on a platform shaker at sufficient rate to maintain gel in suspension. Allow 1 to 2 hr for binding at room temperature or 3 to 4 hr at 4°C.

IMPORTANT NOTE: Excessive shaking may result in foaming and possible denaturation of the target protein.

10. Collect gel by filtration as in step 3. Save filtrate for assay.

This filtrate is saved in case the target protein did not bind. All filtrates should be kept until the fraction containing the target protein has been identified. Storage conditions depend on the stability of the target protein; as a general rule they should be refrigerated to minimize microbial growth. Bacteriostatic agents may be added for prolonged storage.

11. Add 100 ml binding buffer to funnel and resuspend gel using a stirring rod. Allow resuspended gel to stand for 5 min, then apply suction to remove buffer. Repeat three or four times, pooling all washings with the filtrate from step 10.

The gel containing the bound sample may be packed in a column for subsequent elution steps (see Basic Protocol 2).

Elute with step gradient of increasing salt concentration

12. Transfer the sintered-glass funnel to the top of a clean 2000-ml side-arm flask.

13. Add an equal volume of wash buffer to gel in funnel and resuspend using a stirring rod. Allow resuspended gel to stand for 5 min, then apply suction to remove buffer. Save filtrate.

The wash buffer is the first step of the gradient. The filtrate will contain sample components that were weakly adsorbed to the gel (see annotation to step 10).

14. Measure absorbance of filtrate at 280 nm using wash buffer as blank.

15. Repeat steps 13 and 14 until the last filtrate shows no significant absorbance at 280 nm. Pool and save the filtrates.

16. Transfer the sintered-glass funnel to the top of a clean 2000-ml side-arm flask.

17. Add an equal volume of elution buffer to gel in funnel and resuspend using a stirring rod. Allow resuspended gel to stand 5 min, then apply suction to remove buffer. Save filtrate.

The elution buffer is the second step of the gradient. The filtrate will contain the sample components that were more strongly adsorbed to the gel and should include the target protein.

18. Measure the absorbance of the filtrate at 280 nm using elution buffer as a blank.

19. Repeat steps 17 and 18 until the last filtrate shows no significant absorbance at 280 nm. Pool and save the filtrates for assay and subsequent purification of the target protein.

If gel is to be reused, continue with remaining steps. Otherwise, discard the gel.

Regenerate ion-exchange medium for reuse

20. Transfer the sintered-glass funnel to the top of a clean 2000-ml side-arm flask.

21. Add an equal volume of regeneration buffer to gel in funnel and resuspend using a stirring rod. Allow resuspended gel to stand for 5 min, then apply suction to remove buffer. Repeat five times and discard filtrates.

The regeneration buffer removes strongly bound materials from the gel. See Support Protocol 5 for additional considerations in regenerating ion-exchange media.

22. Repeat step 11 for a total of five washes to reequilibrate the gel with binding buffer.
23. Check the pH and conductivity of the last filtrate to ensure that the gel is properly equilibrated.

See Support Protocol 5 for information on storing ion-exchange media.

pH-BASED STEP-GRADIENT ELUTION

The net charge of a protein is pH dependent. Therefore, altering the pH of the mobile phase to make it closer to the pI of a protein can change its net charge, causing it to desorb and elute from an ion exchanger (see Figure 10.10.1). Continuous pH gradients are very difficult to produce at constant salt concentration because charge characteristics of sample components, buffering ions, and ion-exchange media depend on the pH of the system. pH-based step gradients are much easier to produce and more reproducible. For anion-exchange media, elution occurs when the pH is decreased. For cation exchange, the pH is raised for elution.

The first basic protocol can be modified for pH-gradient elution by selecting a buffer at a pH suitable for elution, based on results obtained in a pilot experiment (see Support Protocol 1). This elution buffer is then used in steps 17 to 19 of the first basic protocol. Incubation times and buffer volumes should be increased by 20% when using pH elution.

When scouting methods for ion exchange, it is best to start with changes in salt concentration (see Basic Protocol 1) as this technique is simpler and more reproducible. Elution with pH change is only recommended when the ion-exchange behavior of the sample is well known and the resolution required cannot be attained with a change in salt concentration. pH elution may also be appropriate in cases where it provides a superior ionic/pH environment for loading in the next chromatographic step. A combination of increasing salt concentration and pH change may also be used for elution.

COLUMN CHROMATOGRAPHY WITH LINEAR GRADIENT ELUTION

Column chromatography is generally preferred to batch chromatography (see Basic Protocol 1), especially where high resolution is required, although it accommodates smaller sample volumes. Column chromatography requires specialized equipment to achieve high resolution, including a column containing the medium and a fluid-delivery system appropriate for the pressure and flow rate at which the column is to be operated. It is the technique preferred for the intermediate purification and polishing stages of protein purification. Also in contrast to the first basic protocol, this technique uses a linear gradient for elution. Although more difficult to produce than step gradients (see Basic Protocol 1), linear gradients lead to better resolution (see Support Protocol 3).

The example presented here uses a RESOURCE Q anion-exchange column from Amersham Pharmacia Biotech, which may be used with either FPLC (fast protein liquid chromatography) or HPLC (high-performance liquid chromatography) systems. Optimal conditions for other columns should be obtained from the manufacturer. Procedures mentioned in the protocol steps should be carried out according to the manufacturer's instructions for the particular chromatography system being used.

NOTE: All buffers, media, and other system components should be filtered, degassed, and equilibrated to the same temperature before use.

**ALTERNATE
PROTOCOL**

**BASIC
PROTOCOL 2**

**Analysis of
Proteins**

10.10.13

Materials

Liquid chromatography system (FPLC or HPLC)

Elution buffer: binding buffer (see Basic Protocol 1 and Support Protocol 1) containing 1 M NaCl

Binding buffer (see Basic Protocol 1 and Support Protocol 1)

RESOURCE Q chromatography column (1-ml packed bed volume; Amersham Pharmacia Biotech)

Protein sample to be purified

Conductivity meter

0.22- μ m filter

Prepare chromatography system

1. Set up the liquid chromatography system according to manufacturer's instructions, without the column in-line.
2. Test system performance by running a blank gradient ranging from 0% to 100% elution buffer in 20 ml at a constant flow rate of 5 ml/min.

See Figure 10.10.2 for a diagram of a typical column chromatography system. The gradient is formed by placing binding and elution buffer in the appropriate buffer reservoirs (see Support Protocol 3 for detailed discussion of continuous gradient formation).

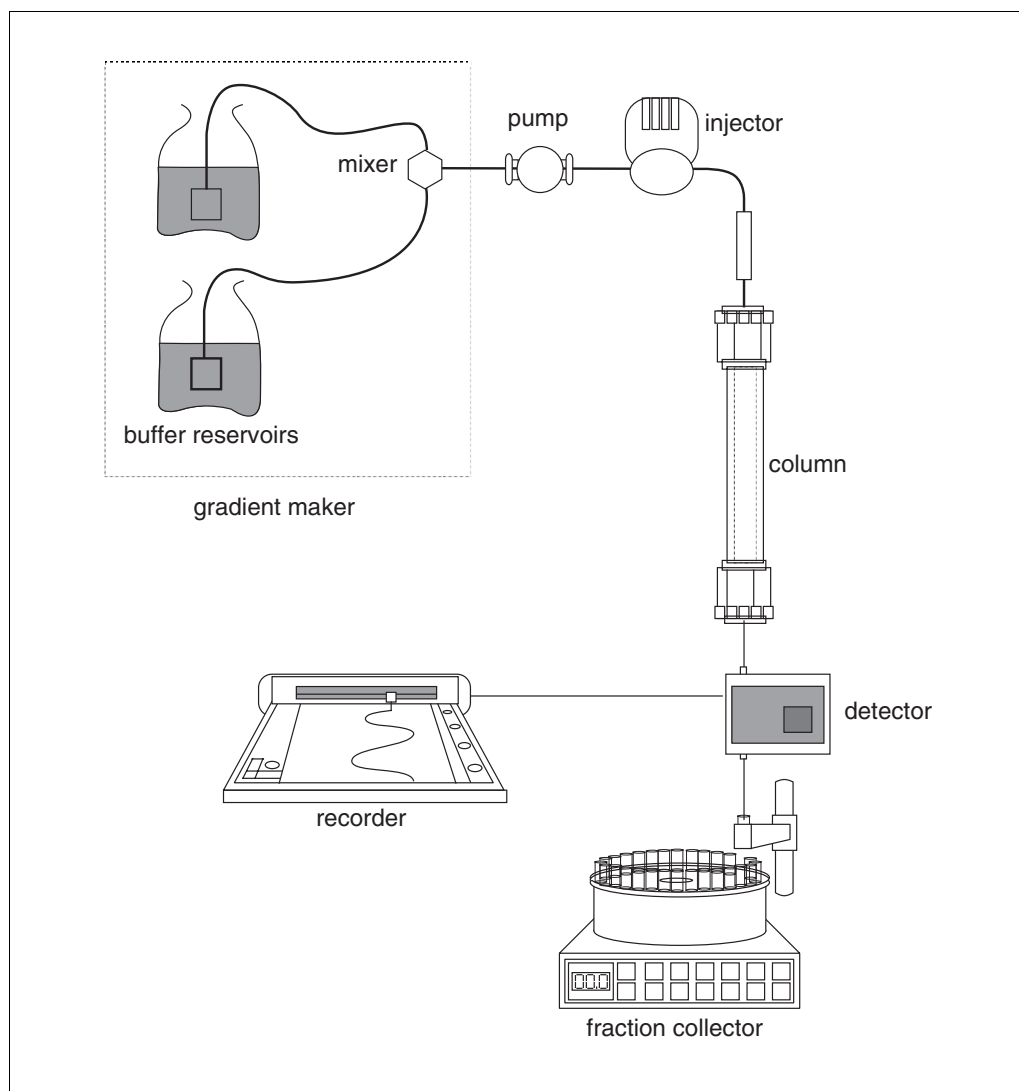


Figure 10.10.2 Liquid column chromatography system with gradient maker.

Stability of the column at extremes of pH must be considered when choosing binding and elution buffers. The working stability of the RESOURCE Q column is pH 2 to 12, and the stability for cleaning is pH 1 to 14.

The flow rate for the test should be appropriate for the column to be connected. For the RESOURCE Q column, the recommended flow rate is 1 to 10 ml/min and the maximum operating pressure is 4 bar.

3. Test for system leaks and ascertain monitor stability according to system documentation. Test accuracy of gradient composition using conductivity meter.

Adding 0.1% (v/v) acetone to the elution buffer used in step 2 will allow observation of gradient composition at 280 nm using the UV monitor. Do not include acetone in the elution buffer used for the chromatographic run.

Prepare column

4. Purge system with binding buffer to remove any air, then install column in system.

New columns may require precycling prior to use; consult product documentation. Used columns may require cleaning or removal of storage solution prior to use (see Support Protocol 4).

5. Wash column with $5 V_c$ (5 ml for RESOURCE Q column) of elution buffer at a flow rate of 5 ml/min and check for leakage.

V_c refers to the packed bed volume of a column and is calculated according to the equation $V_c = \pi r^2 \times L$, where r is the radius of the column and L is the bed height (i.e., the height of the packed medium in the column). As the RESOURCE Q column used here has a packed bed volume of 1 ml, the column is washed with 5 ml of elution buffer. Consult product literature to find V_c for other columns.

6. Equilibrate column with 5 to 10 V_c of binding buffer at 5 ml/min. Collect one fraction at the end of the equilibration stage and measure pH and conductivity to ensure the pH and salt concentration of the fraction are the same as that of binding buffer.

If the pH and salt concentration of the fraction are not the same as that of the binding buffer, continue passing binding buffer through the column until equilibration is complete.

Purify protein sample and regenerate column

7. Adjust pH and salt concentration of sample to binding conditions determined in Support Protocol 1. Filter sample prior to injection using a 0.22- μ m filter to remove particulates that may clog system or column.

For initial experiments, a sample containing 25 mg total protein is appropriate.

For columns packed with beads of average diameter $\geq 90 \mu\text{m}$, a 1- μm filter may be used; for columns packed with beads of average diameter 34 to 90 μm , a 0.45- μm filter may be used. The 0.22- μm filter is appropriate for columns packed with beads of average diameter $< 34 \mu\text{m}$ (e.g., the RESOURCE Q). Alternatively, the sample may be centrifuged 5 min at $20,000 \times g$ to remove particulates.

8. Inject sample into column by opening sample injection valve. Begin fraction collection.

Fractions of 1-ml volume are generally sufficient.

The flow rate during sample injection should be reduced for concentrated samples. It may be increased for dilute samples.

9. After all sample is injected, wash column with 3 to 5 V_c of binding buffer at 5 ml/min to remove unbound or weakly bound materials.

The monitor signal should be allowed to approach the baseline before continuing.

10. Elute with a linear gradient of 0% to 100% elution buffer in 20 V_c (20 ml).
Sample-injection valves should be closed to reduce the system dead volume before beginning gradient.
11. Regenerate column by washing with 5 V_c of elution buffer.
12. Reequilibrate column by washing with 5 to 10 V_c of binding buffer.
13. Assay fractions for target protein. Pool fractions containing target protein for further processing.

TEST TUBE PILOT EXPERIMENT TO DETERMINE STARTING CONDITIONS FOR ION-EXCHANGE CHROMATOGRAPHY

Before beginning an ion-exchange separation (Basic Protocol 1 or 2), the pH and salt concentration required for binding the target protein and the change in pH or salt concentration required for elution must be determined empirically. Once the initial binding conditions have been determined, the available protein capacity (for batch chromatography) is determined to predict the amount of medium that must be used for a given amount of sample. Dynamic capacity (for column chromatography) may be determined as in Support Protocol 2. This protocol is most appropriate when the sample is not strictly limited—e.g., when milligram quantities of a recombinant protein are available. If sample is scarce, a column method should be substituted.

As illustrated in Figure 10.10.3, aliquots of the protein-containing sample to be purified are mixed in test tubes with small quantities of ion-exchange gel that have been equilibrated with a series of buffers of different pH. The optimal binding pH is chosen on the basis of disappearance of target protein in the supernatant at a certain point in the series, then the process is repeated with another series of test tubes containing gel equilibrated with a series of salt solutions of different concentrations. The optimal salt concentrations for binding and elution are chosen, respectively, on the basis of disappearance, then reappearance at another salt concentration, of protein in the supernatant. Finally, aliquots containing increasing protein concentrations are mixed with the gel in a buffer under the optimal binding pH and salt concentrations previously determined. The available capacity of the gel is then determined on the basis of the point where unbound protein appears in the supernatant.

The stability of the target protein at extremes of pH must be considered—i.e., the pH range chosen for the purification should be compatible with retention of structure and activity. In general, this can be evaluated with a crude preparation by incubating aliquots at a series of pH values, then performing the structural or functional assay at the optimal pH value.

NOTE: The method presented here is for anion exchange with increasing salt concentration for elution. Differences for cation exchange are noted where appropriate. Select cation-exchange buffers from Table 10.10.1.

Materials

- 20 mM piperazine, pH 5.0, 5.5, and 6.0 (prepare from 100 mM stock)
- 20 mM 1,3-bis[tris(hydroxymethyl)methylamino]propane (bis-Tris propane), pH 6.5 and 7.0 (prepare from 100 mM stock)
- 20 mM Tris·Cl, pH 7.5, 8.0, and 8.5 (*APPENDIX 2*; prepare from 100 mM stock)
- Q Sepharose Fast Flow (50% slurry in 20% ethanol; Amersham Pharmacia Biotech) or equivalent anion exchange resin in appropriate buffer
- Protein sample to be purified, containing known quantity of target and total protein

4 M NaCl

15-ml test tubes

Centrifuge with rotor accommodating 15-ml test tubes (optional)

Determine optimal binding pH

1. Set up a series of test tubes, numbered 1 to 8, containing different buffers as follows:

- tube 1: 5 ml 20 mM piperazine, pH 5.0
- tube 2: 5 ml 20 mM piperazine, pH 5.5
- tube 3: 5 ml 20 mM piperazine, pH 6.0
- tube 4: 5 ml 20 mM bis-Tris propane, pH, 6.5
- tube 5: 5 ml 20 mM bis-Tris propane, pH 7.0
- tube 6: 5 ml 20 mM Tris·Cl, pH 7.5
- tube 7: 5 ml 20 mM Tris·Cl, pH 8.0
- tube 8: 5 ml 20 mM Tris·Cl, pH 8.5.

2. Shake bottle of Q Sepharose Fast Flow to resuspend gel. Pour 25 ml of gel slurry into a graduated cylinder, then allow gel to settle. Adjust volume of liquid above gel to equal the volume of settled gel (i.e., prepare 50% slurry), then resuspend with a glass stirring rod.

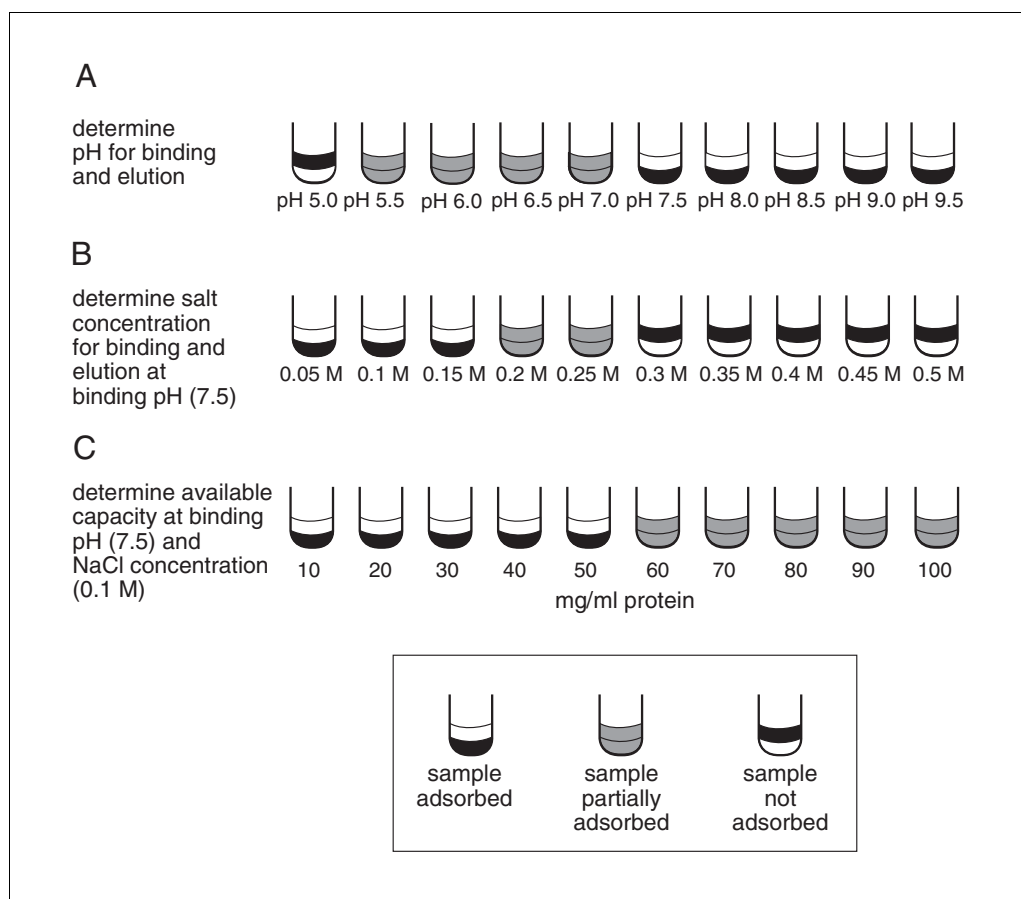


Figure 10.10.3 Results of a typical test tube pilot experiment for selecting initial conditions for anion-exchange chromatography. Series of tubes with protein and ion-exchange medium at **(A)** varying pH (to select binding and elution pH); **(B)** constant pH, varying NaCl concentration (to select binding and elution salt concentrations); **(C)** constant pH and NaCl concentration, varying protein concentration (to determine available capacity). The binding pH here will be ≥ 7.5 ; the elution pH will be ≤ 5.0 ; the binding salt concentration will be ≥ 0.15 M; the elution salt concentration will be > 0.3 M; and the available capacity will be > 50 mg/ml.

3. Add 2 ml of 50% slurry of Q Sepharose Fast Flow to each tube and mix.
SP Sepharose Fast Flow (Amersham Pharmacia Biotech) or equivalent may be used for cation exchange.
4. Allow gel to settle to bottom of tube or centrifuge 1 min at $\sim 5000 \times g$, room temperature.
5. Decant supernatant from each tube and replace with 5 ml of the same buffer added in step 1. Gently shake or vortex to resuspend gel, then let tubes stand 2 min.
6. Repeat steps 4 and 5 three times to equilibrate gels in each tube.
7. Decant supernatants and resuspend each equilibrated gel in 1 ml of the same buffer added in step 1.
8. Set up eight identical aliquots (≥ 1 ml) of the protein to be purified in tubes numbered 1 to 8, each containing 0.1 to 1 mg total protein. Adjust the pH of each to obtain eight different pH values corresponding to those of tubes 1 to 8 in step 1.

Quantity of protein to be used depends on availability of sample and extinction coefficient of target protein. pH may be adjusted by titrating with acid or base using pH paper or by diluting 1:1 with buffer of desired pH.

9. Add the corresponding pH-adjusted protein aliquot to each of the 8 tubes from step 7. Mix gently for 10 min by periodic swirling, then allow gel to settle.

If the salt concentration of the sample is too high, no binding will occur. It is therefore recommended that the sample be desalted into the appropriate buffer as in UNIT 10.9.

10. Assay supernatant from each tube for the target protein.

Measuring the absorbance of the supernatant at 280 nm is usually sufficient. Figure 10.10.3 shows a set of possible results. A decrease or absence of target protein activity (or absorbance) in the supernatant indicates binding. Specific assays for the target protein may also be used.

11. Select the lowest pH at which the highest quantity of target protein is bound to use as the binding pH for further work. Select the highest pH where no binding occurs as the elution pH.

For anion exchange, it is best to select a binding pH no more than 0.5 to 1 pH unit (pH 7.5 to 8.0 in Fig. 10.10.3) above the highest pH at which no binding occurs. If too high a pH is chosen, elution becomes difficult and high salt concentrations may be required.

For cation exchange, select a binding pH 0.5 to 1 unit below the lowest pH where no binding occurs. Select the lowest pH where no binding occurs as the elution pH.

Determine salt concentration for binding and elution

12. Set up a series of test tubes, numbered 1 to 10. Using the binding buffer selected in step 10 in place of the different buffers in step 1, perform steps 1 to 7 to equilibrate the Q Sepharose Fast Flow with buffer in all ten tubes.
13. To each tube, add the same amount of the protein to be purified that was added to the tubes in step 8, equilibrated in the binding buffer selected in step 10.
14. Mix contents of tubes gently for 10 min, then allow gel to settle.
15. Decant supernatants from each tube and discard. Wash gel two times with 5 ml of the binding buffer, allowing the gel to settle and decanting the supernatants after each addition of buffer.

16. Add 2 ml of the binding buffer to each tube, then sequentially add water and 4 M NaCl to each tube as follows:

tube 1: 1.90 ml H₂O and 0.10 ml 4 M NaCl (0.10 M NaCl final)
tube 2: 1.80 ml H₂O and 0.20 ml 4 M NaCl (0.20 M NaCl final)
tube 3: 1.70 ml H₂O and 0.30 ml 4 M NaCl (0.30 M NaCl final)
tube 4: 1.60 ml H₂O and 0.40 ml 4 M NaCl (0.40 M NaCl final)
tube 5: 1.50 ml H₂O and 0.50 ml 4 M NaCl (0.50 M NaCl final)
tube 6: 1.40 ml H₂O and 0.60 ml 4 M NaCl (0.60 M NaCl final)
tube 7: 1.30 ml H₂O and 0.70 ml 4 M NaCl (0.70 M NaCl final)
tube 8: 1.20 ml H₂O and 0.80 ml 4 M NaCl (0.80 M NaCl final)
tube 9: 1.10 ml H₂O and 0.90 ml 4 M NaCl (0.90 M NaCl final)
tube 10: 1.00 ml H₂O and 1.00 ml 4 M NaCl (1.00 M NaCl final).

17. Mix contents of tubes gently for 10 min, then allow gel to settle.
18. Assay supernatant from each tube for the target protein.

Figure 10.10.3 shows a set of possible results. Presence of target protein in the supernatant indicates elution of the target protein.

19. Select the highest salt concentration at which no elution occurs as the maximum salt concentration allowable for binding target protein and for washing away unbound sample components. Select a salt concentration at least 0.05 M above the concentration where no protein binds to the gel as the salt concentration for elution.

Determine available capacity

20. Prepare a buffer/salt solution with the binding pH selected in step 10 and the binding salt concentration selected in step 19.
21. Set up a series of test tubes numbered 1 to 10. Using the buffer/salt solution prepared in step 20 in place of the different buffers in step 1, perform steps 1 to 7 to equilibrate the Q Sepharose Fast Flow with buffer/salt solution in all ten tubes.
22. Add aliquots of sample to each tube to obtain the following amount of target protein:

tube 1: 10 mg
tube 2: 20 mg
tube 3: 30 mg
tube 4: 40 mg
tube 5: 50 mg
tube 6: 60 mg
tube 7: 70 mg
tube 8: 80 mg
tube 9: 90 mg
tube 10: 100 mg.

23. Mix gently for 10 min, then allow gel to settle.
24. Assay supernatant from each tube for the target protein.

Figure 10.10.3 shows a set of possible results. Presence of target protein in supernatant indicates that the available capacity has been exceeded.

25. Select the highest protein concentration at which the target protein does not appear in the supernatant as the available capacity.

Calculate the amount of ion-exchange medium needed for batch purification based on 50% of the available capacity. If a column system is used in pilot experiments, however, it is best to determine the dynamic capacity of the medium (see Support Protocol 2). However, 20% of the available capacity is a safe and reasonable starting point for pilot experiments using column chromatography.

**MEASUREMENT OF DYNAMIC (COLUMN) CAPACITY AND
BREAKTHROUGH CAPACITY OF ION-EXCHANGE COLUMNS**

For chromatographic separations that will be performed on a routine basis as well as separations that will be scaled up, it is necessary first to optimize conditions with respect to resolution and to determine the capacity of the medium with respect to the target protein (see Support Protocol 1). Knowledge of the capacity allows optimal use of the gel medium in terms of cost and yield, and also allows working limits to be set on the sample load to ensure robustness of a particular purification step. After capacity has been determined, the separation can be optimized with respect to time. There are two capacity terms that commonly appear in the literature (in addition to available capacity; also see Support Protocol 1): dynamic capacity and breakthrough capacity.

The dynamic (or column) capacity is the capacity of an ion-exchange column under defined conditions of flow rate, pH, salt concentration and sample concentration. The breakthrough capacity (Q_B) for a system is obtained by calculating the amount of protein that has been absorbed by the column when the sample is first detected in the effluent, or when the recorder signal reaches some arbitrarily defined percent of full-scale deflection. Breakthrough capacity values at 50% full-scale deflection (Q_{B50}) are commonly reported in the literature. However, it is generally recommended to use only 10% to 20% of the published Q_B as the practical capacity of the column to ensure high recovery of the target protein.

Additional Materials (also see Basic Protocol 2)

Ion-exchange gel of unknown capacity, and column
Elution buffer capable of eluting target protein in single step of salt concentration or pH (e.g., 2 M NaCl; see Support Protocol 1)

1. Pack a defined amount of the ion-exchange gel into the column.

See UNIT 10.9 for guidelines on column packing. Normally, a quantity of gel sufficient to give a packed bed volume of ~1 ml is sufficient. In the case of a prepacked column, the amount of gel is predetermined. The packed bed volume (V_c) is calculated according to the equation $V_c = \pi r^2 \times L$, where r is the radius of the column and L is the bed height (i.e., the height of the packed medium in the column).

2. Prepare chromatography system and column (see Basic Protocol 2, steps 1 to 6) using the optimal binding and elution conditions determined in Support Protocol 1.

Do not exceed 75% of the flow rate used for packing. The effect of flow rate on capacity should be determined in a true optimization.

3. Prepare protein solution and inject into column (see Basic Protocol 2, steps 7 and 8). Apply solution continually until the recorder shows >50% full-scale deflection, then stop.

The absorbance of the binding buffer corresponds to 0% and the absorbance of the sample corresponds to 100% full-scale deflection.

4. Wash column with binding buffer until 0% full-scale deflection is approached.

5. Elute proteins with a buffer providing a single-step increase in salt concentration or pH. Continue collecting eluant until recorder shows $\leq 2\%$ full-scale deflection, then pool fractions.

A typical chromatogram for this procedure is presented in Figure 10.10.4.

6. Assay pooled fractions from step 5 for target protein. Calculate maximum amount of protein (A) that can be bound to the column according to the equation $A = C_p \times V$,

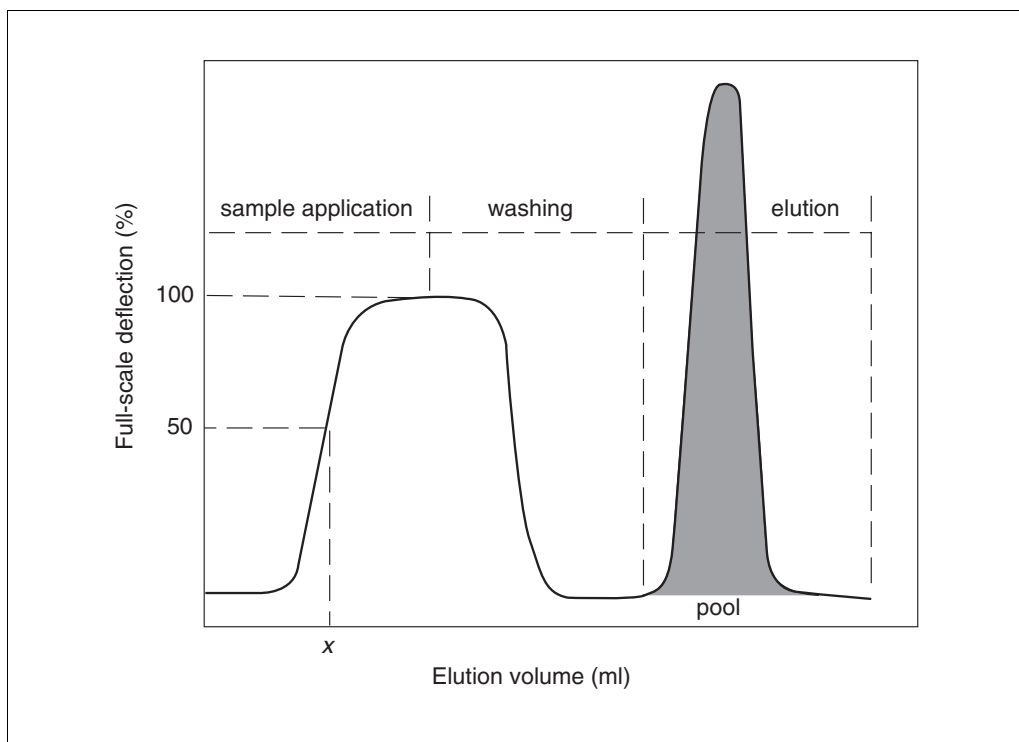


Figure 10.10.4 Typical chromatogram for determining the capacity of an ion-exchange column. The volume of eluant (x) at 50% full-scale deflection is used to calculate the breakthrough capacity (Q_{B50}).

where C_p is the protein concentration of the pooled fractions (mg/ml) and V is the volume of the pooled fractions (ml).

7. Calculate the column (dynamic) capacity (mg/ml) using the equation dynamic capacity = A/V_c .
8. Calculate breakthrough capacity (Q_{B50}) using the equation $Q_{B50} = (C_s \times x)/V_c$, where C_s is the concentration of protein in the original sample and x is the volume of sample that has been applied to the column at the point where 50% full-scale deflection has been attained.

If Q_{B50} greatly exceeds the dynamic capacity, then the target material may still be adsorbed on the column, and the elution conditions should be reexamined.

See Figure 10.10.4 for graphical depiction of these variables.

GRADIENT-FORMATION TECHNIQUES

Simple Gradient Mixers

Gradient mixers of the type shown in Figure 10.10.5 can generate linear pH or salt gradients in gravity-based or one-pump chromatography systems. Similarly designed gradient mixers used for casting electrophoresis gels may also be used for chromatography. The reservoir closest to the column is filled with the binding buffer and the second reservoir is filled with an equal volume of the elution buffer. The gradient begins when the valve between the two reservoirs is opened. Continuous mixing is required in the reservoir closest to the column to ensure gradient linearity and reproducibility.

**SUPPORT
PROTOCOL 3**

**Analysis of
Proteins**

10.10.21

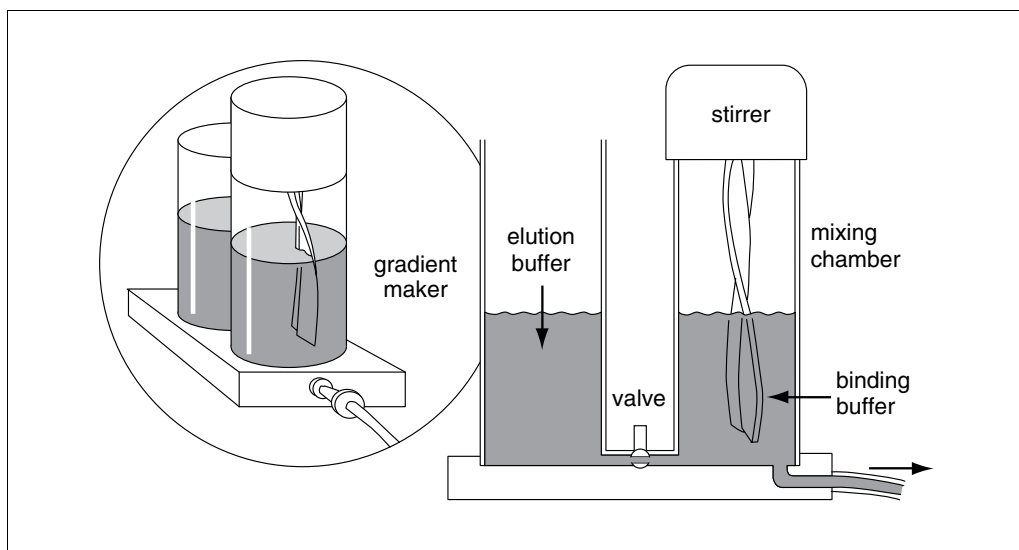


Figure 10.10.5 Gradient mixer for forming gradients of pH and salt concentration used during ion-exchange chromatography. The apparatus shown is an Amersham Pharmacia Biotech Gradient Mixer GM-1, which can be used for preparing gradients up to 500 ml. A linear gradient of salt concentration is produced by filling the reservoirs with buffers at the same pH and different salt concentration. The mixing chamber contains the lower-salt-concentration (binding) buffer, and the other chamber contains the higher-salt-concentration (elution) buffer. A continuous linear or nonlinear pH gradient may be produced by filling the reservoirs with buffers at the same salt concentration and different pH. A gradient of increasing pH is used for cation-exchange separations, and a gradient of decreasing pH is used for anion-exchange separations. The buffer chambers are joined by a channel controlled by a valve, and outflow from the mixing chamber is controlled by another valve.

Gradient Mixing with Multichannel Peristaltic Pumps

Peristaltic pumps that can accommodate three or more pumping channels simultaneously (Fig. 10.10.6) can be used to form linear or complex gradients (see Critical Parameters). Linear gradients are easily formed by using the same size tubing for all channels, pumping with one channel from the elution buffer reservoir to the binding buffer reservoir, and pumping with the other two channels from the binding buffer reservoir to the column. Continuous mixing is required in the binding buffer reservoir closest to the column to ensure gradient linearity and reproducibility.

Switch-Valve-Based Gradient Mixing

Switch-valve-based gradient formation is useful for forming step, linear, and complex gradients (see Critical Parameters). Gradients are formed by proportioning variable amounts of binding buffer and elution buffer through a single three-way valve using a system controller. When programmed to deliver 10% elution buffer, the controller will open the valve port for binding buffer 90% of the time and the port for elution buffer 10% of the time. As higher switching rates are required for lower flow rates to ensure gradient accuracy, high-quality system controllers will vary the rate at which the valve switches (i.e., the number of switching events per unit time) as a function of flow rate. A dynamic mixer should be included in the flow path just after the proportioning valve to ensure gradient accuracy. Switch-valve-based gradients are generally not very accurate from 1% to 10% elution buffer (90% to 99% binding buffer) or from 1% to 10% binding buffer (90% to 99% elution buffer). Therefore, the ionic strength of the binding and elution buffers may be adjusted to allow a gradient of the desired composition to be formed by mixing the two buffers in the range of 20% to 80% elution buffer (80% to 20% binding buffer).

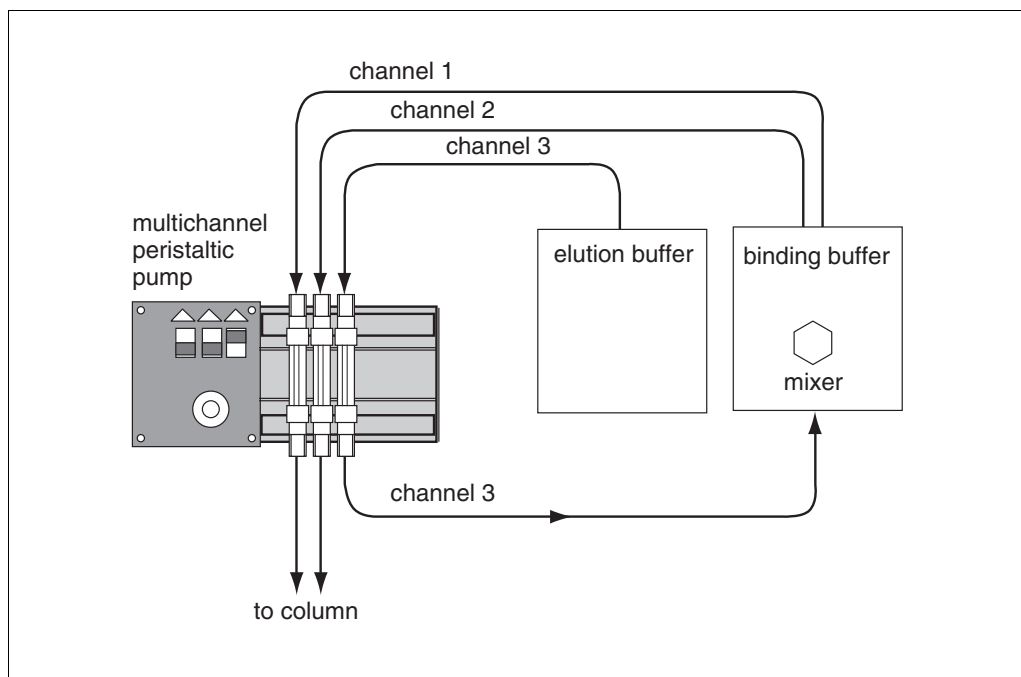


Figure 10.10.6 Peristaltic pump accommodating three pumping channels for continuous gradient formation.

Gradient Formation by Autoblending

Some chromatography systems allow gradient formation over a range of both pH and salt concentration. Four solvents—a high-pH buffer, a low-pH buffer, a concentrated salt solution, and water—are blended via four two-way valves. A system controller proportions the high- and low-pH buffers to achieve the desired pH gradient while simultaneously proportioning the concentrated salt solution and water to achieve the desired salt-concentration gradient. Gradients formed in this way are usually difficult to reproduce but may be useful for scouting initial binding and elution conditions.

Gradient Formation with Twin Pumps

Use of two positive-displacement pumps, as in HPLC and FPLC systems, provides the highest degree of accuracy, precision, and reproducibility in gradient formation. One pump delivers binding buffer while the other delivers elution buffer. A system controller allows gradient formation simply by controlling the flow rate of each pump. A dynamic mixer should be included in the flow path, after the pumps and before the column, to ensure gradient accuracy.

CLEANING AND REGENERATION OF ION-EXCHANGE MEDIA

Proper maintenance of chromatography media is as much an art as protein purification itself. Knowing what kind of contaminating material from the sample is bound to the column helps in selecting a regime for maintaining good column hygiene. It is best to practice preventative maintenance by regenerating the medium after every run, using a high-salt buffer or large change in pH, and periodically cleaning the column and checking the system flow characteristics. If any increase in column back-pressure is noted, the column should be cleaned as soon as possible. Once enough contaminant has bound to a column to reduce or block the flow, it is probably too late to save the medium without investing considerable time and effort.

**SUPPORT
PROTOCOL 4**

**Analysis of
Proteins**

10.10.23

Routine Washing of Columns and Media

Columns or bulk medium should be washed after every run with salt solution until an ionic strength of ≥ 2 M is reached. This should remove any substances bound by ionic forces. This is recommended after every run.

Removal of Contaminants

Alkali-soluble contaminants. Contaminants such as lipids, proteins, and nucleic acids can usually be removed by washing with 2 to 3 column volumes of 0.1 M NaOH, followed by 2 to 3 column volumes of Milli-Q-purified (or equivalent) distilled water and then 2 to 3 column volumes of binding buffer. Many media can tolerate brief exposure to 1 M NaOH, if harsher conditions are required to remove contaminants; however, manufacturer's guidelines should be consulted regarding pH stability of the medium.

Hydrophobic contaminants. Lipids and other hydrophobic materials may be removed by washing with alcohol solutions (e.g., 70% ethanol) or nonionic detergents. When using alcohols or organic solvents to remove hydrophobic materials, it is often effective to wash first with a gradient from 0% to 100% and then with a gradient from 100% to 0%. Repeat washing with the alternating gradients until no contaminants are detected in the eluant.

Metallic contaminants. Contamination with metals may cause blue or gray discoloration at the top of an ion-exchange column. These metals can originate from impurities in buffer salts or water, or they can leach from metallic system components in contact with buffer solutions. Metal contaminants may usually be removed by treating the column with several column volumes of 10 mM HCl (i.e., pH 2) saturated with EDTA. Manufacturer's guidelines should be checked for pH stability of medium and column before exposure to mild or strong acid solutions.

Precipitated materials. Precipitated materials that have accumulated in a column are very difficult to remove. Unless the precipitated material can be physically removed by removing the gel at the top of the column, the material will have to be resolubilized. Detergents, urea, and guanidine-Cl can be introduced to help dissolve contaminants. The viscosity of concentrated solutions used for cleaning may require very low flow rates. The column may be equilibrated and incubated in 6 M urea, then washed with distilled water and buffer. Degradative enzymes (e.g., proteases and lipases) may also be introduced in an attempt to degrade precipitates on the column; however, enzymatic cleaning can be quite expensive. It may be useful to reverse the direction of flow through the column (i.e., from bottom to top) when attempting to remove particulate material from the top of the bed.

SUPPORT PROTOCOL 5

STORAGE OF ION-EXCHANGE MEDIA

Bacterial and microbial growth can seriously interfere with the chromatographic properties of any chromatography medium and endanger the sample as well. Prior to storage, all media should be cleaned and sanitized. During prolonged experiments or storage (i.e., >24 hr at room temperature or 48 hr at 4°C), a bacteriostatic or antimicrobial agent should be added to the ion exchanger. Antimicrobials chosen to be added for storage should not bind to the ion exchanger and should be easily removed when the gel is to be reused. Azide is anionic and will bind to anion-exchange media.

Treatment with 1 M NaOH for 1 hr will sanitize most media by lowering the bacterial count 100-fold, and will help solubilize dead cells (Pharmacia Biotech, 1997). Storage of sanitized media in 0.1 M NaOH will not leave any toxic materials on the column—a significant consideration in pharmaceutical production. Manufacturer's guidelines should

be checked for stability of the medium and column in NaOH. Silica-based media will not tolerate high pH.

Exposure to 70% ethanol for 3 to 4 hr can be used to sanitize most media. Most ion-exchange media and columns can be stored for prolonged periods of time in 20% ethanol. Manufacturer's guidelines should be checked for stability of the medium and column in ethanol.

Effective antimicrobial agents appropriate for use with anion-exchange media include either 0.001% phenyl mercuric salts or 0.002% chlorhexidine, in weakly alkaline solution.

An effective antimicrobial agent for use with cation-exchange media is 0.005% merthiolate in weakly acidic solution. An agent that is effective for use with either anion or cation exchangers is 0.05% trichlorobutanol in weakly acidic solution.

COMMENTARY

Background Information

The first ion exchangers were synthetic resins designed for applications such as demineralization, water treatment, and recovery of ions from waste. Early ion-exchange resins were tightly cross-linked hydrophobic polymers, highly substituted with ionic groups, and had very high capacity for small ions. The high degree of cross-linking provided mechanical strength, but the limited porosity restricted use of these media with large molecules. In addition, the high charge density resulted in very strong binding, and the hydrophobic matrix tended to adsorb and denature labile biological materials. The first ion exchangers designed for use with biomolecules were developed by Peterson and Sober (1956). These cellulose-matrix ion exchangers were very hydrophilic and had little tendency to denature proteins, but had low capacities and poor flow characteristics. Modern ion-exchange media have very high capacity, often >100 mg/ml. Modern particle technology has significantly increased the mechanical strength and flow characteristics of ion-exchange media so that separations that once required hours or days may now be completed in a few minutes.

Ion-exchange can easily be coupled to other methods of protein separation in a multistep purification. Samples purified by ion exchange can typically be applied to a hydrophobic-interaction chromatography (HIC) medium simply by adding NaCl (or other salts) to a concentration sufficient for binding. Samples to be further purified by chromatofocusing or an additional round of ion exchange will require desalting (UNIT 10.9). Samples may usually be applied directly to gel-filtration columns (UNIT 10.9) with no intervening treatment.

Critical Parameters

The most critical aspects of any ion-exchange experiment are selection of the appropriate ion-exchange medium and of the initial conditions for binding and elution of the target protein or contaminants. Once these parameters are selected, conditions can be optimized with respect to the three primary parameters of any chromatographic separation—i.e., resolution, capacity, and speed. These three parameters are mutually exclusive. Resolution will typically decrease when the sample load is increased and when the flow rate is increased (to increase speed). It is generally recommended that resolution be optimized first, then capacity, and finally speed.

Optimization of resolution

This parameter is most easily optimized by controlling the selectivity of the medium through manipulation of binding and elution conditions. Construction of chromatographic titration curves followed by optimization of the shape of the elution gradient is a common approach to achieving maximum resolution.

Chromatographic titration curves can be constructed by performing a series of anion- and cation-exchange experiments at different pH values as in Figure 10.10.7. For each of these experiments, the elution salt concentrations (in mM NaCl) for sample components A, B, and C are plotted on the y axis and the operating pH is plotted on the x axis. It is then simple to observe the pH ranges within which optimal resolution between the various sample components is achieved. The different sample components are usually detected by UV absorption, but may also be viewed by other means (e.g., fluorescence or bioactivity). Data for chromatographic titration curves can often

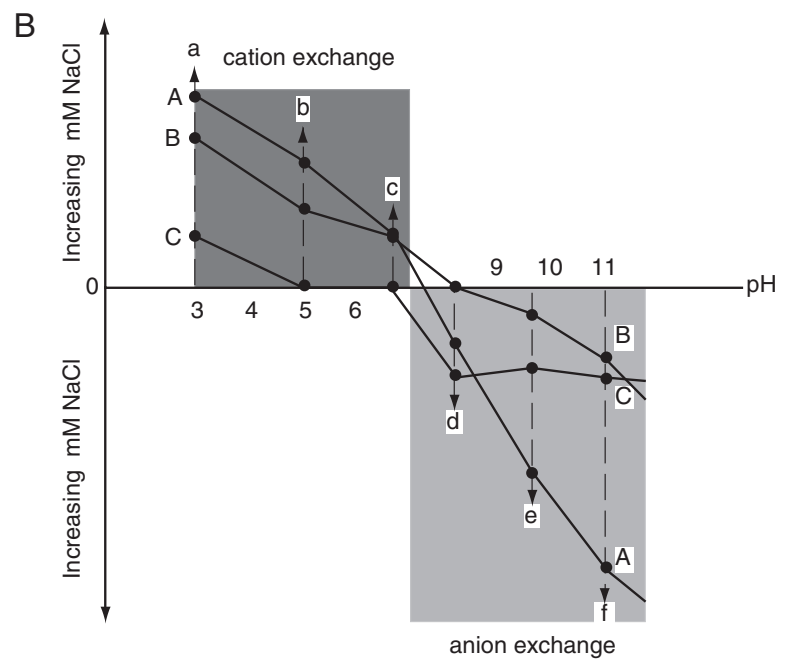
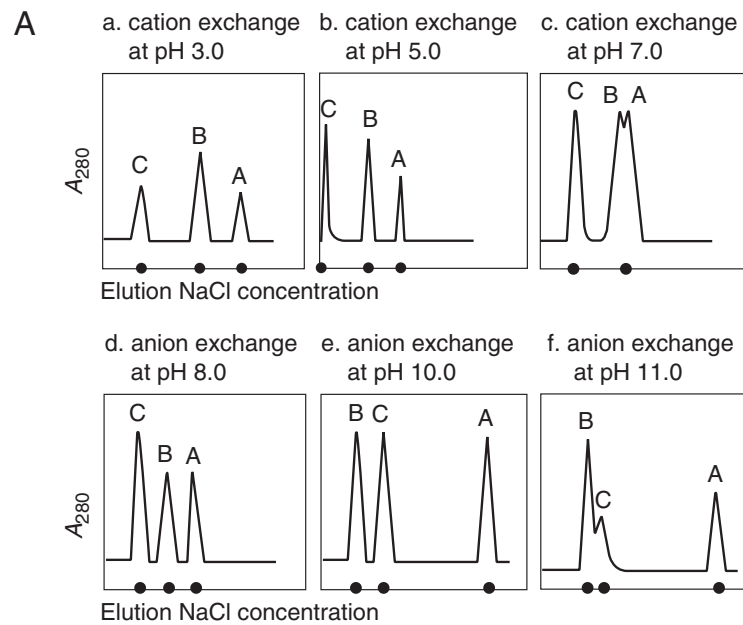


Figure 10.10.7 Construction of a chromatographic titration curve. **(A)** Six chromatographic separations (a to f) of a sample containing proteins A, B, and C are carried out at six different operating pHs using a linear gradient of increasing salt concentration. **(B)** The elution ionic strengths for components A, B, and C (y axis) are plotted against the operating pH (x axis) for experiments a to f. Note that the salt concentration in mM NaCl (y axis) is ascending in both directions. The optimal resolution with respect to component A occurs below pH 5 with cation exchange and above pH 9 with anion exchange.

be generated by fast screening with low-volume gradients (e.g., 5 to 10 V_c) and low sample mass (e.g., <1 mg).

After optimization of resolution with respect to pH, the shape of the elution gradient can be optimized. A gradient can be conceptually divided into segments, and each segment may have a different gradient shape. Gradient shapes include linear, concave, convex, and step. Complex gradients are constructed by using different gradient shapes in consecutive segments. The slope (e.g., mM NaCl per minute) of a linear segment is constant, that of a concave segment continually increases, that of a convex gradient continually decreases, and that of a step gradient approaches infinity. Resolution between sample components in any segment can usually be increased by decreasing the gradient slope in that segment. However, decreasing the gradient slope will increase the peak volume during elution, the gradient volume, and the time of the separation. Step gradients can provide the most concentrated fractions, and are desirable when minimizing sample volume is of primary importance. Use of concave gradients in the initial part of a segment and convex gradients in the later part of a segment can improve resolution. The capabilities of available equipment must be considered in the construction of complex gradients. Use of simple gradients that do not stress the performance limits of the system will result in robust methods that are insensitive to small changes in sample or buffer composition and gradient shape, thereby increasing probability of success.

To construct a complete gradient profile, it is first necessary to identify the gradient range where desorption of the target protein occurs, as well as the gradient slope that provides optimal resolution. A single-step elution is then done at a salt concentration just below the range of where the target protein elutes, to remove weakly bound material. The target protein is then desorbed and eluted. A single-step elution at a high salt concentration will then remove tightly bound sample components and prepare the ion-exchange gel for regeneration. This approach, using step gradients before and after the range of interest, can provide simple, robust methods that are easy to scale up, and will save time and reagents.

Different ion-exchange media offer different selectivities. If adequate resolution for a specific stage in a purification cannot be obtained with a specific ion-exchange medium, it

may be desirable to select an alternate ion exchanger with a different functional group, or to use a different chromatographic technique at this stage in the separation. Salts other than NaCl may be used in elution, but there are no specific effects from different salts on the separation in ion-exchange chromatography, other than differences in concentration. Karlsson et al. (1998) provides a good review of the effects of nonbuffering ionic species in ion exchange.

Optimization of capacity

When optimum resolution has been achieved, the effect of increasing sample load on resolution should be evaluated to determine the optimal balance of resolution and capacity utilization. The procedure for determining the dynamic (column) capacity is described in Support Protocol 3 and can be performed using the conditions for optimal resolution (as determined above) to elute the target protein. Once the capacity is determined under the optimal elution conditions, the effect of changes in sample load on resolution can be investigated to identify optimal conditions. It is generally recommended that only 10% to 20% of the dynamic capacity be used to ensure high recoveries of the target protein.

Optimization of throughput

Once optimal conditions for resolution and capacity have been determined, the flow rate in each gradient segment may be increased and the effect on resolution and capacity monitored to complete the optimization process.

Troubleshooting

Common problems encountered during ion-exchange experiments, their potential causes, and suggested solutions are outlined in Table 10.10.5.

Anticipated Results

A chromatogram illustrating typical results of an ion-exchange column chromatography experiment (i.e., Basic Protocol 1) is shown in Figure 10.10.8.

The goal of any purification step is the pure target protein obtained at the end of the day. The chromatogram is only a record for comparison and is deceptive at best. Great care should be exercised in interpreting a chromatogram—i.e., the results from assays performed on fractions collected must be used in interpreting the chromatogram and inferring the identity of any given chromatographic peak.

Table 10.10.5 Troubleshooting Guide for Ion-Exchange Chromatography

Problem	Possible cause	Solution
Protein does not elute in the salt gradient	Buffer pH is incorrect	Use buffer pH closer to pI of the protein
	Salt concentration is too low	Use more concentrated elution buffer
Protein does not elute in pH gradient	Buffer pH is incorrect	Calibrate pH meter and prepare new solutions
Protein elutes in wash phase	Salt concentration of binding buffer is too high	Decrease salt concentration of binding buffer
	Salt concentration of sample is too high or pH is wrong	Perform buffer exchange on sample
	Column is not properly equilibrated	Repeat or prolong equilibration until conductivity is constant
	Ionic detergents or other additives are absorbed to column	Clean column
Resolution is less than expected	Gradient slope is too steep	Use shallower gradient or include plateau in gradient
	Flow rate is too high	Run separation at lower flow rate
	Proteins or lipids have precipitated on columns	Clean and regenerate column
	Sample has not been filtered properly before application to the column	Regenerate column, filter sample, and repeat chromatography
	Proteins in sample have formed aggregates that bind strongly to gel	Use urea or detergents to clean column
	Column is poorly packed	Check packing by running a colored compound and observing the band. Repack column if necessary.
	Too much sample mass has been loaded onto column	Clean and regenerate column. Decrease sample load and repeat run.
	Detector flow cell volume or mixing spaces in or after column are too large	Change flow cell or reduce all post-column volumes
Leading or very rounded peaks observed in chromatogram	Column is overloaded	Decrease sample load and repeat run
	Column is poorly packed	Check packing by running a colored compound and observing the band. Repack column if necessary.
	Column needs regeneration	Clean and regenerate column. If this does not help, replace with new one.
Tailing of peak is observed in the chromatogram	Sample is too viscous	Reduce amount of protein, or remove nucleic acids from sample.
	Precipitation of protein has occurred in column filter and/or at top of the gel bed	Clean column; exchange or clean filter.
	Column is poorly packed	Check packing by running a colored compound and observing band. Repack column if necessary.
Previous elution profile cannot be reproduced	Buffer pH and/or ionic strength are incorrect	Prepare new solutions
	Sample has deteriorated during storage	Prepare fresh sample
	Proteins or lipids have precipitated on column	Clean and regenerate column

continued

Table 10.10.5 Troubleshooting Guide for Ion-Exchange Chromatography, continued

Problem	Possible cause	Solution
	Sample has not been filtered properly	Regenerate column, filter sample carefully, and repeat step
	Equilibration was incomplete	Repeat or prolong equilibration until conductivity is constant
Recovery of activity is low but recovery of protein is normal	Target protein is not stable in the elution buffer and is therefore inactivated	Change elution conditions
	Enzyme separated from cofactor or other substance needed for activity	Test by pooling all fractions and repeating assay
	Microbial growth has occurred in the column	Clean column and store gel in 20% ethanol or other antimicrobial agent ^a
Amount of protein in the eluted fractions is much less than expected	Protein was degraded by proteinases	Add proteinase inhibitors to buffers
	Protein adsorbed to filter during sample preparation	Use another type of filter or add detergents to buffer
	Microbial growth occurred in the column	Clean column and store gel in 20% ethanol or other antimicrobial agent ^a

^aMicrobial growth rarely occurs in columns during use, but steps should always be taken to prevent infection of packed columns, buffers, and gel suspensions; see Support Protocol 5.

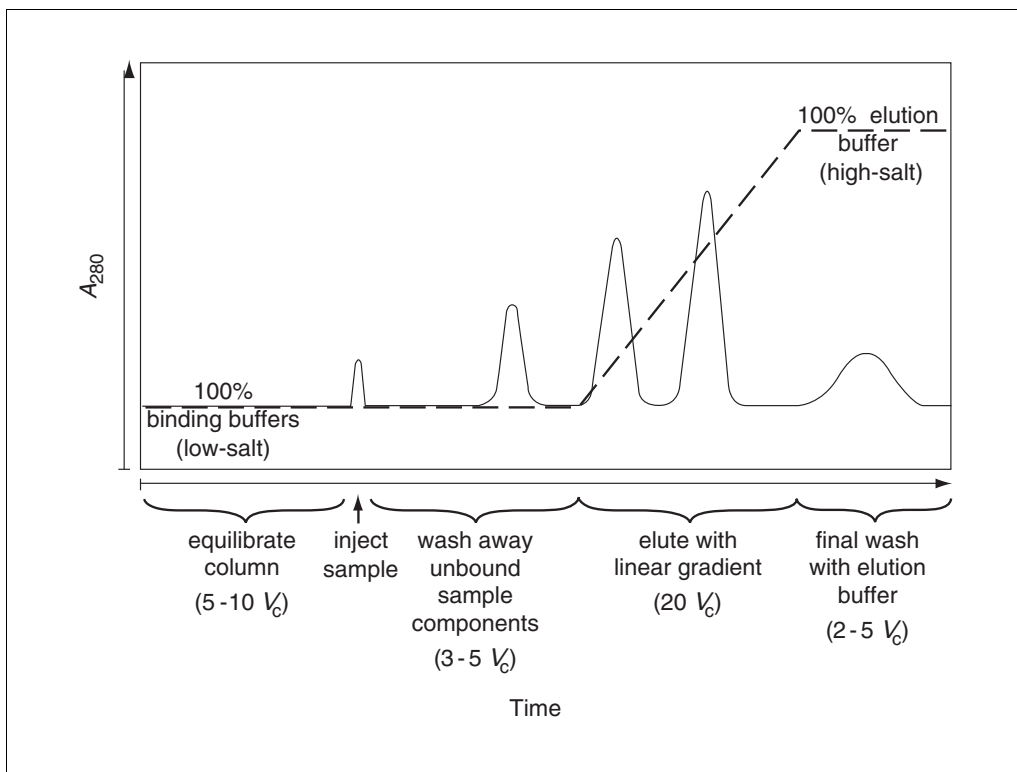


Figure 10.10.8 Chromatogram representing typical results of stages of an ion-exchange column chromatography experiment. A linear gradient of 0% to 100% elution buffer is used as in Basic Protocol 2. The broken line superimposed on the chromatogram represents the composition of the elution gradient (100% binding buffer to equilibrate column and remove unbound sample components, followed by a linear gradient from 100% binding buffer to 100% elution buffer to elute bound proteins and a final wash with 100% elution buffer). V_c represents the packed bed volume of the column.

Time Considerations

If all necessary equipment is available and functioning properly, it should be possible to develop a chromatographic method in a day. With automated systems (e.g., FPLC and HPLC) and modern ion-exchange media, method development may require just a few hours. However, the most time-consuming aspect in developing a purification step is typically the determination of protein purity in the fractions collected, which will always depend on the specific protein of interest. The key to minimizing the time required for such determinations is to strive for high resolution and a minimum number of fractions.

Literature Cited

- Karlsson, E., Rydén, L., and Brewer, J. 1998. Ion exchange chromatography. *In* Protein Purification: Principles, High Resolution Methods and Applications, 2nd ed. (J.C. Janson and L. Rydén, eds.) pp. 145-205. John Wiley & Sons, New York.
- Gianazza, E. and Righetti, P.G. 1980. Size and charge distribution of macromolecules in living systems: *J. Chromatog.* 193:1-8.
- Peterson, E.A. and Sober, H.A., 1956. Chromatography of proteins. I. cellulose ion exchange adsorbents. *J. Amer. Chem. Soc.* 78:751-755.
- Pharmacia Biotech. 1995. Ion Exchange Chromatography: Principles and Methods, ed. AA. Pharmacia Biotech AB, Uppsala, Sweden.

Pharmacia Biotech. 1997. Application Note 18-1124-57: Use of Sodium Hydroxide for Cleaning and Sanitizing Chromatography Media and Systems. Pharmacia Biotech, Uppsala, Sweden.

Key References

Cooper, E.H., Turner, R., Webb, J.R., Lindblom, H., and Fägerstam, L. 1985. Fast liquid protein chromatography scale-up procedures for the preparation of low molecular weight proteins from urine. *J. Chromatogr.* 327:269-277.

Good example of methods development with respect to optimization of resolution.

Pharmacia Biotech. 1985. FPLC Ion Exchange and Chromatofocusing: Principles and Methods. Pharmacia Biotech AB, Uppsala, Sweden.

Contains detailed discussions of experimental approach, methodology, and applications for protein purification.

Pharmacia Biotech, 1995. See above.

Concise descriptions of theory and practice in planning and implementing ion-exchange purification.

Contributed by Alan Williams and
Verna Frasca
Amersham Pharmacia Biotech
Piscataway, New Jersey

Immunoaffinity Chromatography

UNIT 10.11A

Immunoaffinity purification is a powerful technique for isolating proteins. Purifications of 10,000-fold or more can often be achieved in one step. Immunoaffinity columns most often employ monoclonal antibodies covalently attached to a solid-phase matrix; polyclonal antibodies have been used successfully, but they usually lack the specificity required for a single-step protein purification. The amount of protein that can be purified is solely dependent on the amount and affinity of the antibody employed.

This unit describes the isolation of soluble or membrane-bound protein antigens from cells or homogenized tissue by immunoaffinity chromatography. The technique involves the elution of a single protein from an immunoaffinity column after prior elution of nonspecifically adsorbed proteins. Specifically, antibodies are coupled to Sepharose (an insoluble, large-pore-size chromatographic matrix). High-molecular-weight antigens pass freely into and out of the pores and bind to antibodies covalently bound to the matrix. To elute the bound antigen from the immunoaffinity matrix, the antibody-antigen interaction is destabilized by brief exposure to high-pH (see Basic Protocol) or low-pH (see Alternate Protocol 1) buffer. The use of batch purification of antigens shortens the column loading time (see Alternate Protocol 2). The detergent octyl β -D-glucoside can be used instead of Triton X-100 for elution. Because octyl β -D-glucoside has a high critical micelle concentration (CMC), it can be readily removed by dialysis (see Alternate Protocol 3).

ISOLATION OF SOLUBLE OR MEMBRANE-BOUND ANTIGENS

Two different Sepharose columns in series—a precolumn to remove nonspecifically binding material and a specific column—are used to isolate antigens from a cell or tissue lysate. Column fractions are analyzed by SDS-PAGE and silver staining to detect the antigens.

Materials

- Antibody (Ab)-Sepharose (UNIT 10.16)
- Activated, quenched (control) Sepharose, prepared as for Ab-Sepharose (Support Protocol, UNIT 10.16) but eliminating Ab or substituting irrelevant Ab during coupling
- Cells or homogenized tissue
- Tris/saline/azide (TSA) solution (see recipe), ice cold
- Lysis buffer (see recipe), ice cold
- 5% (w/v) sodium deoxycholate (Na-DOC; filter sterilize and store at room temperature)
- Wash buffer (see recipe)
- Tris/Triton/NaCl buffers, pH 8.0 and 9.0 (see recipe), ice-cold
- Triethanolamine solution (see recipe), ice cold
- 1 M Tris·Cl, pH 6.7 (APPENDIX 2), ice cold
- Column storage solution (see recipe), ice cold
- Chromatography columns
- Ultracentrifuge
- Quick-seal centrifuge tubes (Beckman)
- Additional reagents and equipment for column chromatography (APPENDIX 3A), preparing antibody-Sepharose (UNIT 10.16), SDS-PAGE (UNIT 10.2), and silver staining (UNIT 10.6)

NOTE: Carry out all procedures involving antigen in a 4°C cold room or on ice.

**BASIC
PROTOCOL**

**Analysis of
Proteins**

10.11.1

Contributed by Timothy A. Springer

Current Protocols in Molecular Biology (1996) 10.11.1-10.11.9

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 36

Prepare the columns

1. Prepare an Ab-Sepharose immunoaffinity column (5 ml; 5 mg/ml antibody per milliliter packed Sepharose) and an activated, quenched (control) Sepharose precolumn (5 ml packed bed volume) linked in series (Fig. 10.11A.1).

Irrelevant antibody can be coupled to the Sepharose in the precolumn.

Column size can vary; adjust amounts of Sepharose and cells proportionally.

Prepare the lysate

2. Suspend 50 g of cells at $1-5 \times 10^8$ cells/ml in ice-cold TSA solution, or add 1 to 5 vol ice-cold TSA per vol packed cells or homogenized tissue. Add an equal volume of ice-cold lysis buffer and stir 1 hr at 4°C.

For glycosylphosphatidylinositol (GPI)-anchored proteins, incubate 10 min at 20°C for efficient solubilization.

3. Centrifuge 10 min at $4000 \times g$, 4°C, to remove nuclei. Decant supernatant and save.

For purification of cytoplasmic (soluble) antigens, it is not necessary to add detergents to the solutions and buffer used in subsequent steps. Detergent is needed only for cell lysis and solubilization of integral membrane proteins.

4. For purification of membrane antigens, add 0.2 vol of 5% Na-DOC to the postnuclear supernatant, and leave 10 min at 4°C or on ice. Transfer to quick-seal centrifuge tubes and centrifuge 1 hr at $100,000 \times g$, 4°C. Carefully remove supernatant and save.

Set up and wash the columns

5. Attach Sepharose precolumn to immunoaffinity column (Fig. 10.11A.1).

6. Wash both columns using the following regimen:

10 column volumes of wash buffer

5 column volumes of Tris/Triton/NaCl buffer, pH 8.0

5 column volumes of Tris/Triton/NaCl buffer, pH 9.0

5 column volumes of triethanolamine solution

5 column volumes of wash buffer.

Isolate the antigen

7. Apply the supernatant from steps 3 or 4 (reserving some for analysis as described in step 15 below) to the precolumn and allow it to flow through the precolumn and specific column linked in series at a flow rate of 5 column volumes/hr. Collect the flowthrough fractions, each $\frac{1}{10}$ to $\frac{1}{100}$ the volume of the applied supernatant.

“Fat” chromatography columns, filled with Sepharose to a height of $\sim 2 \times$ column diameter, are used to maximize flow rates. A 10- to 20-ml syringe is used for 5 ml of Sepharose. The flow rate is adjusted with a hydrostatic head of up to 250 cm (Fig. 10.11A.1). Sample loading can routinely take up to 2 days with no deleterious effect, but longer periods would suggest the column is clogged or the lysate is too viscous. The latter is usually due to the presence of DNA.

8. Wash with 5 column volumes of wash buffer, then close the stopcocks on both columns and disconnect the precolumn from the immunoaffinity column. Open the stopcock of the immunoaffinity column and allow fluid above the top of the column to drain out to bed level.

The Sepharose has some elasticity and draining can continue until there is no buffer above the Sepharose bed. Draining until cracks appear in the Sepharose should be avoided.

Fractions of this wash and washes obtained below should be saved.

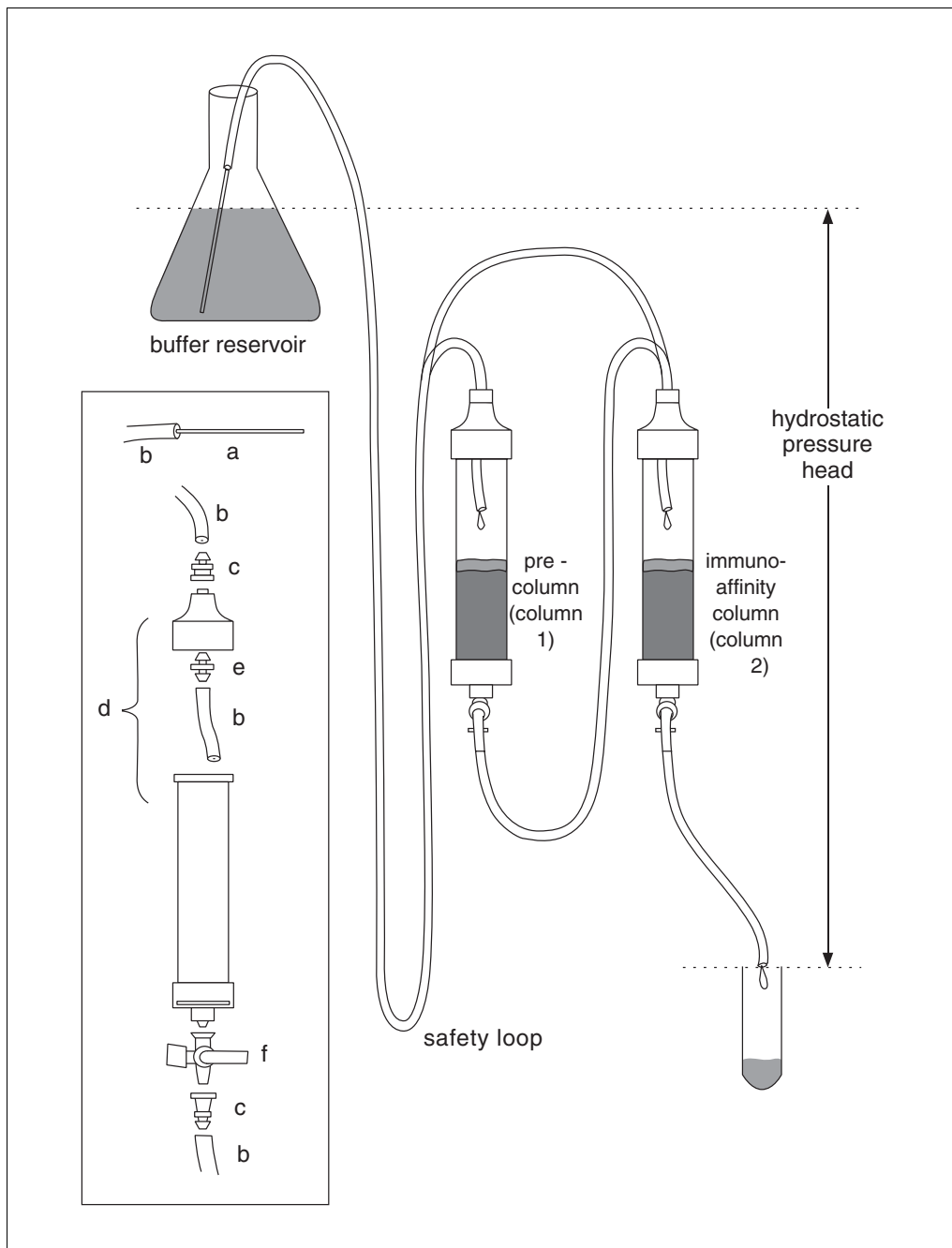


Figure 10.11A.1 Immunoaffinity chromatography. During the application of the sample, two Sepharose columns, a Sepharose precolumn (without covalently bound specific antibody or with a covalently bound irrelevant antibody) and an immunoaffinity column (with covalently bound antibody), are attached in series to a buffer reservoir containing the sample. After the sample has been washed through, the precolumn is removed and the tubing of the safety loop is connected to the immunoaffinity column. The hydrostatic pressure head is the distance between the top of the solution in the buffer reservoir and the tip of the tubing at the bottom of the immunoaffinity column. When the elution reservoir is emptied, the hydrostatic head becomes zero when the fluid level reaches the safety loop, preventing columns from running dry. Fluid remaining above the column beds can be removed by raising the safety loop. After rinsing the tubing, the next elution is begun by placing the end of the safety loop in another reservoir containing the next elution buffer.

Inset: Schematic diagram of an immunoaffinity column. (a) 50- μ l disposable capillary micropipet. (b) Tubing: Tygon S-54-HL Microbore, 0.05-in. i.d., or Tygon R-3603, $\frac{1}{16}$ -in. i.d. (softer tubing). (c) Female Luer fitting, white nylon (Value Plastics), $\frac{1}{16}$ in. (d) Kontes Flex-column (Kontes Glass). (e) Barbed nipple connector, polypropylene, $\frac{3}{32}$ -in. top, $\frac{1}{16}$ -in. bottom (Value Plastics, Series AD). (f) Luer-Lok two-way stopcock (Kontes Glass).

9. Wash the immunoaffinity column between each change of buffers (steps 10 to 14) as follows. Close the stopcock and remove the end cap of the column. With a syringe connected to the outlet of the tubing from the buffer reservoir, aspirate all buffer from the tubing. Place tubing into the next buffer contained in another reservoir. Aspirating with a syringe, fill the tubing from the reservoir and remove the syringe. Crimp the tubing to regulate flow and rinse the inside wall of the column with the buffer. Open the column stopcock and drain the buffer to bed level. Put end cap loosely on the column and allow buffer to drain into the column to a level several centimeters above the bed. Secure end cap and commence washes or elution.
10. Wash with 5 column volumes of wash buffer.
11. Wash with 5 column volumes of Tris/Triton/NaCl buffer, pH 8.0.
12. Wash with 5 column volumes of Tris/Triton/NaCl buffer, pH 9.0.

Some nonspecifically bound proteins may be eluted at this step. In addition, some monoclonal antibodies may release some ligand at this pH. This should be checked.
13. Elute the antigen with 5 column volumes of triethanolamine solution. Collect fractions of 1 column volume into tubes containing 0.2 vol of 1 M Tris·Cl, pH 6.7, to neutralize the fractions collected.

In some cases it may be desirable to lower the pH of the triethanolamine solution to preserve the functional activity of the ligand. The ideal pH gives complete release of the ligand, as verified by SDS-PAGE evaluation of a sample (~20 μ l) of the eluted column bed (Ab-Sepharose) and eluate (50 μ l).
14. Wash the column with 5 column volumes of TSA solution.

A column may be reused many times and remain active for several years after storage at 4°C in TSA solution. It is important to prevent drying out of a column during storage. The use of column storage solutions inhibits the growth of microorganisms.
15. Analyze fractions for the presence of antigen—50- μ l aliquots of each eluate fraction should be analyzed by SDS-PAGE and silver staining. Analyze 0.5- to 1-ml aliquots of the sample applied to the column and representative flowthrough and wash fractions by immunoprecipitation with Ab-Sepharose, and detect by silver staining to determine whether the column was saturated.

If antibody leaches off the column during elution, it may be removed from the eluate by passage through protein A-Sepharose (Ey et al., 1978). Even the weakly binding mouse IgG1 subclass can be quantitatively removed at pH 8 (M. Dustin, pers. comm.).

Preliminary analysis of fractions can be done using A_{280} readings when octyl β -D-glucoside or sodium deoxycholate are used as detergents.

**ALTERNATE
PROTOCOL 1**

LOW-pH ELUTION OF ANTIGENS

Some protein antigens may be eluted more completely, with greater retention of native conformation and with fewer contaminants, by employing low-pH buffers.

Additional Materials (also see Basic Protocol)

- Sodium phosphate buffer, pH 6.3 (see recipe)
- Glycine buffer (see recipe)
- 1 M Tris·Cl, pH 9.0 (APPENDIX 2)

1. Prepare the columns and lysate, wash the columns, and isolate the antigen (see Basic Protocol, steps 1 to 11).

It is essential to remove sodium deoxycholate from the column before acid elution, because it precipitates or forms a gel at acid to neutral pH.

TSA and triethanolamine solutions used here may also be modified by substituting 1% octyl β -D-glucoside for 0.1% Triton X-100 (see Alternate Protocol 3).

2. Wash with 5 column volumes of sodium phosphate buffer, pH 6.3.
3. Elute with 5 column volumes of glycine buffer. Collect fractions into tubes containing 0.2 vol of 1 M Tris-Cl, pH 9.0.

Mix each fraction immediately after collection.

4. Analyze fractions for antigen (see Basic Protocol, step 15).

BATCH PURIFICATION OF ANTIGENS

Batch purification of antigens shortens the column loading time. This technique is valuable for viscous lysates that take too long to load on a column and for antigens especially susceptible to proteolysis, because less time is required to complete the steps. A precolumn is not utilized because the supernatant is mixed with Ab-Sepharose and poured into a column. The antigen is then eluted (see Basic Protocol). The drawbacks of this protocol are that more “hands-on” time is required by the investigator and that nonspecifically binding material is not removed by a precolumn.

1. Suspend and centrifuge the cells and purify the membrane antigens (see Basic Protocol, steps 2 to 4) to obtain the postnuclear supernatant.
2. Suspend Ab-Sepharose in the supernatant in a flask. Shake gently on a rotary shaker for 3 hr. Stop shaking and allow the Sepharose to settle.
3. Decant most of the supernatant. Pour the Ab-Sepharose and the remainder of the supernatant into a column and open the stopcock. Continue draining the column until all the Sepharose has been added. Allow the fluid to drain to bed level and close the stopcock.
4. Wash the immunoaffinity column, elute the antigen, and analyze the fractions (see Basic Protocol, steps 9 to 15).

ELUTION IN OCTYL β -D-GLUCOSIDE

The detergent octyl β -D-glucoside has a high critical micelle concentration (CMC) of 0.73% and can therefore be readily removed by dialysis. Also, adsorption of membrane proteins to surfaces for ELISA or adhesion assays is much more efficient when solutions containing membrane proteins are diluted so that the concentration of octyl β -D-glucoside is below the CMC. Because octyl β -D-glucoside is expensive, initial steps requiring large volumes may be done as described in the Basic Protocol; the initial detergent is then replaced by octyl β -D-glucoside in the wash step prior to elution.

Additional Materials (also see Basic Protocol)

TSA solution (see recipe) containing 1% octyl β -D-glucoside

1. Prepare the columns and lysate, wash the columns, and isolate the antigen (see Basic Protocol, steps 1 to 11).
2. Wash with 5 column volumes of TSA solution containing 1% octyl β -D-glucoside.
3. Elute with 5 column volumes of triethanolamine solution with 1% octyl β -D-glucoside substituted for 0.1% Triton X-100.
4. Collect fractions of 1 column volume into tubes containing 0.2 vol of 1 M Tris-Cl, pH 6.7, to neutralize the fractions collected.

**ALTERNATE
PROTOCOL 2**

**ALTERNATE
PROTOCOL 3**

**Analysis of
Proteins**

10.11.5

5. Wash the column and analyze the fractions (see Basic Protocol, steps 14 and 15).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

Column storage solution

Prepare in TSA solution (see recipe) either 1 mM EDTA/20 µg/ml gentamicin *or* 0.01% thimerosal (Aldrich).

Detergent stock solutions

Prepare 10% Triton X-100 *or* 5% sodium deoxycholate in water. Sterilize either solution by Millipore filtration. Both solutions remain stable for 5 years at room temperature; Triton X-100 solution should be stored in the dark to prevent photo-oxidation.

NP-40 can be used in place of Triton X-100.

Glycine buffer

50 mM glycine-HCl, pH 2.5
0.1% Triton X-100 (see recipe for detergent stock solutions)
0.15 M NaCl

Lysis buffer

TSA solution (see recipe) containing:
2% Triton X-100 (see recipe for detergent stock solutions)
5 mM iodoacetamide
Aprotinin (0.2 trypsin inhibitor U/ml)
1 mM phenylmethylsulfonyl fluoride (add fresh from 100 mM stock solution prepared in absolute ethanol)

IMPORTANT NOTE: *Iodoacetamide is a protease inhibitor and prevents oxidation of free cysteines to disulfide-bonded cysteines. It should be omitted for enzymes that require cysteines for activity.*

NP-40 can be used in place of Triton X-100.

Sodium phosphate buffer, pH 6.3

50 mM sodium phosphate, pH 6.3
0.1% Triton X-100 (see recipe for detergent stock solutions)
0.5 M NaCl

Triethanolamine solution

50 mM triethanolamine, pH ~11.5
0.1% Triton X-100 (see recipe for detergent stock solutions)
0.15 M NaCl

Other organic solutions, such as diethylamine, may be used in place of triethanolamine. The pH of this solution should be determined for each antibody as elution conditions may vary (see Critical Parameters).

Tris/saline/azide (TSA) solution

0.01 M Tris-Cl, pH 8.0 (at 4°C)
0.14 M NaCl
0.025% NaN₃

CAUTION: *Sodium azide (NaN₃) is poisonous; wear gloves and handle cautiously.*

Tris/Triton/NaCl buffer, pH 8.0 and 9.0

50 mM Tris·Cl, pH 8.0 or pH 9.0

0.1% Triton X-100 (see recipe for detergent stock solutions)

0.5 M NaCl

Wash buffer

0.01 M Tris·Cl, pH 8.0 (at 4°C)

0.14 M NaCl

0.025% NaN₃

0.5% Triton X-100 (see recipe for detergent stock solutions)

0.5% sodium deoxycholate (see recipe for detergent stock solutions)

CAUTION: *Sodium azide (NaN₃) is poisonous; wear gloves and handle cautiously.*

COMMENTARY

Background Information

The review of affinity chromatography by Wilchek et al. (1984) discusses available methods for activation of solid supports, coupling of ligands, adsorption of proteins, and elution of protein from affinity columns. Table III of that review lists numerous examples of proteins that have been purified by immunoaffinity chromatography and the elution conditions for each purification.

Traditionally, purification of membrane proteins started with a membrane purification step, and in some cases this is still desirable. However, it is difficult to achieve more than a 5-fold purification of plasma membranes and yields are usually only 10% to 40%. Omission of membrane purification in this protocol (Williams and Barclay, 1986; Johnson et al., 1985) results in increased yield and decreased experiment time.

Purification to homogeneity or near-homogeneity can usually be achieved for protein antigens present in $\geq 10,000$ molecules per eukaryotic cell. This protocol can be used for both membrane and intracellular antigens. However, for soluble antigens, immunoaffinity chromatography is completed without detergent.

Critical Parameters

Binding capacities of Ab-Sepharose columns (coupled at 5 mg monoclonal antibody/ml Sepharose) have been found to be 2% to 20% of the theoretical binding capacity. The lowest and highest binding capacity values were found for antigens of 150,000 and 18,000 M_r , respectively, suggesting that increased antigen size may constrain access to antibody in the pores of the affinity matrix. Because of its larger pore size, Sepharose CL-4B (a 4% cross-linked agarose) is far preferable to Sepharose 6B (a 6% agarose), which is the usual commer-

cially available preactivated agarose. The cross-links in the CL series yield mechanically more robust beads and do not appreciably decrease activation with cyanogen bromide (CNBr).

A binding capacity of 40% was reported for coupling at 2 to 3 mg antibody/ml Sepharose. Successful purification has been achieved using monoclonal antibodies with affinity constants ranging from 2×10^7 to 4×10^8 M⁻¹. The column should be saturated with antigen by allowing some of the antigen to flow through the column during loading. This will result in the highest antigen purity and a mass yield, and will diminish the relative level of antibody eluted along with the antigen when antibody is leaching off the column.

Sodium deoxycholate is used in the solubilization protocol (Johnson et al., 1985) because it dissociates proteins from the membrane more effectively than Triton X-100. However, because sodium deoxycholate releases DNA from nuclei, it must be added to the lysate after the nuclei are removed. Sodium deoxycholate forms a mixed micelle with Triton X-100. Although sodium deoxycholate can be substituted for Triton X-100 during high-pH elution, it gels at low pH and in high salt. Both sodium deoxycholate and nonionic detergents may dissociate subunits of protein complexes, which interact within the membrane. Addition of phospholipid, low concentrations of Triton X-100, and mild detergents (e.g., digitonin, octylglucoside, and CHAPS) preserves membrane protein complexes (Helenius et al., 1979; Rivnay et al., 1982; Tsuchiya and Saito, 1984).

Glycosylphosphatidylinositol (GPI)-anchored proteins are resistant to detergent solubilization at 4°C (Schroeder et al., 1994). Preparation of the lysate (see Basic Protocol, step 2) should be conducted for 10 min at 20°C

for these antigens. Additional information regarding detection, extraction, and partitioning of GPI-anchored proteins can be found in *Current Protocols in Protein Science* (Doering et al., 1995).

Octyl β -D-glucoside is an expensive detergent that is readily removable by dialysis. Furthermore, adsorption of proteins to substrates is enhanced by dilution below a CMC of 0.73%. In Alternate Protocol 3, the initial detergent is replaced with octyl β -D-glucoside prior to elution of antigen from the affinity column.

Protein antigens eluted by acid or base can frequently be renatured by neutralization. However, some protein antigens are irreversibly denatured. The structure of certain antigens is preserved after acid, but not base, elution (Plunkett and Springer, 1986), whereas the structure of other antigens is preserved after base, but not acid, elution (Johnson et al., 1985). Some antigens are eluted at low pH, but not at high pH; for others the reverse is true. For each antibody-antigen combination, the optimal pH for elution of specific antigens as well as contaminants must be empirically defined. Antibody binding capacity is usually retained after repeated exposure to low- and high-pH elution buffers (see Alternate Protocol 1 and see Basic Protocol, respectively). The use of chaotropic agents (e.g., potassium thiocyanate and urea) for elution is a seldom-employed alternative (Johnson et al., 1985).

Troubleshooting

Immunoaffinity chromatography relies on the elution of a single protein from an immunoaffinity column after prior elution of all other nonspecifically adsorbed proteins. Thus, depending on the elution conditions used, the desired protein antigen may be contaminated with other proteins. It is very important to determine whether a contaminant is present if the protein is to be analyzed for amino acid composition or protein sequence. One-dimensional gel electrophoresis (UNIT 8.4) should be used to verify elution of contaminating proteins during washing of the immunoaffinity column, as well as the purity of the protein in the final eluate. If the protein is not pure, the wash steps must be optimized to ensure that other contaminating proteins are removed.

Anticipated Results

Antigen yield is typically 40% to 70% of starting material (Kürzinger and Springer, 1982; Johnson et al., 1985) and purification factors of 1,000- to 10,000-fold may be achieved (Kürz-

inger and Springer, 1982; Williams and Barclay, 1986; Plunkett and Springer, 1986). Further purification can usually be achieved by a second cycle of immunoaffinity chromatography. Monoclonal antibodies are most convenient, but affinity-purified polyclonal IgG can also be used.

Time Considerations

Pouring the column takes a few minutes and lysate preparation requires ~6 hr. Purification proceeds over 1 to 2 days depending on the flow rate of the immunoaffinity column. The majority of this time involves loading the sample on the column, which may be done from a reservoir and requires little hands-on time. The use of batch purification (see Alternate Protocol 2) reduces the sample application time to 3 hr. Elution of an immunoaffinity column requires 5 to 6 hr.

Literature Cited

- Doering, T.L., Englund, P.T., and Hart, G.W. 1995. Detection of glycopospholipid anchors on proteins. *In Current Protocols in Protein Science* (J. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 12.5.1-12.5.14. John Wiley & Sons, New York.
- Ey, P.L., Prowse, S.J., and Jenkin, C.R. 1978. Isolation of pure IgG₁, IgG_{2a} and IgG_{2b} immunoglobulins from mouse serum using protein A-Sepharose. *Immunochemistry* 15:429-436.
- Helenius, A., McCaslin, D.R., Fries, E., and Tanford, C. 1979. Properties of detergents. *Methods Enzymol.* 56:734-749.
- Johnson, P., Williams, A.F., and Woollett, G.R. 1985. Purification of membrane glycoproteins with monoclonal antibody affinity columns. *In Hybridoma Technology in the Biosciences and Medicine* (T.A. Springer, ed.) pp. 163-175. Plenum, New York.
- Kürzinger, K. and Springer, T.A. 1982. Purification and structural characterization of LFA-1, a lymphocyte function-associated antigen, and Mac-1, a related macrophage differentiation antigen. *J. Biol. Chem.* 257:12412-12418.
- Plunkett, M.L. and Springer, T.A. 1986. Purification and characterization of the lymphocyte function-associated-2 (LFA-2) molecule. *J. Immunol.* 136:4181-4187.
- Rivnay, B., Wank, S.A., Poy, G., and Metzger, H. 1982. Phospholipids stabilize the interaction between the alpha and beta subunits of the solubilized receptor for immunoglobulin E. *Biochemistry* 21:6922-6927.
- Schroeder, R., London, E., and Brown, D. 1994. Interactions between saturated acyl chains confer detergent resistance on lipids and glycosylphosphatidylinositol (GPI)-anchored proteins: GPI-anchored proteins in liposomes and cells

- show similar behavior. *Proc. Natl. Acad. Sci. U.S.A.* 91:12130-12134.
- Tsuchiya, T. and Saito, S. 1984. Use of *n*-octyl- β -D-thioglucoside, a new nonionic detergent, for solubilization and reconstitution of membrane proteins. *J. Biochem.* 96:1593-1597.
- Wilchek, M., Miron, T., and Kohn, J. 1984. Affinity chromatography. *Methods Enzymol.* 104:3-55.
- Williams, A.F. and Barclay, A.N. 1986. Glycoprotein antigens of the lymphocyte surface and their purification by antibody affinity chromatography. In *Immunological Methods in Biomedical Sciences* (D.M. Weir, L.A. Herzenberg, C.C. Blackwell, and L.A. Herzenberg, eds.) pp. 22.1-22.24. Blackwell Scientific, Oxford.
- Hjelmeland, J.M. and Chrambach, A. 1984. Solubilization of functional membrane proteins. *Methods Enzymol.* 104:305-318.
- Johnson et al., 1985. See above.
Describes the critical parameters involved in immunoaffinity chromatography.
- Wilchek et al., 1984. See above.
Describes the mechanism of activation of Sepharose by CNBr and alternative activation procedures, and lists numerous examples of proteins purified by affinity chromatography.

Key References

- Harlow, E. and Lane, D. 1988. *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Contributed by Timothy A. Springer
Center for Blood Research
Harvard Medical School
Boston, Massachusetts

Metal-Chelate Affinity Chromatography

Recombinant proteins engineered to have six consecutive histidine residues on either the amino or carboxyl terminus can be purified using a resin containing nickel ions (Ni^{2+}) that have been immobilized by covalently attached nitrilotriacetic acid (NTA). This technique, known as metal-chelate affinity chromatography (MCAC), can readily be performed with either native or denatured protein. The Strategic Planning section discusses techniques for creating a fusion protein consisting of the protein of interest with a histidine tail attached (for purification by MCAC). The Basic Protocol describes expression of histidine-tail fusion proteins and their purification in native form by MCAC. Two alternate protocols describe purification of histidine-tail fusion proteins by MCAC under denaturing conditions and their renaturation by either dialysis or solid-phase renaturation. Support protocols are provided for analysis of the purified product and regeneration of the NTA resin. All of these protocols are easily adaptable to any protein expression system.

Prior to large-scale preparation, the cells should be tested for expression of the protein in soluble form (see UNIT 16.2). Even if protein is expressed mostly in insoluble form (i.e., bacterial inclusion bodies), there may be a small fraction that remains soluble and can therefore be purified using the Basic Protocol. If little or no soluble protein is observed, one of the denaturing protocols (see Alternate Protocols 1 and 2) should be used.

NOTE: All solutions and equipment coming into contact with live cells must be sterile, and proper sterile technique should be used accordingly.

STRATEGIC PLANNING

MCAC is designed to purify a specific protein whose complementary DNA (cDNA) sequence is available. A protein-expression system is selected and the cDNA is inserted into the appropriate expression vector (see Chapter 16). The cDNA sequence must encode a minimum of six histidines at either the amino or carboxyl terminus and must include an initiator methionine at the amino terminus and a termination codon at the carboxyl terminus. This can be accomplished by polymerase chain reaction (PCR) using primers containing unique restriction-site sequences at the 5' end (Fig. 10.11B.1). Properly

A	5'-GGGNNNNNNATGCATCATCATCATCAT ... N ₁₅₋₃₀ -3' RE site – MetHisHisHisHisHisHis ...
B	5'-GGGNNNNNNNTTAATGATGATGATGATGATG ... N ₁₅₋₃₀ -3' RE site – ENHisHisHisHisHisHis ...

Figure 10.11B.1 Sequences of primers required to create histidine tails at protein termini. Functions (e.g., protein sequence) are marked below. Three guanines are included at the 5' ends of each primer to facilitate restriction enzyme digestion of the PCR product prior to subcloning. NNNNNN represents a unique restriction enzyme site compatible with the selected vector. N₁₅₋₃₀ represents 15 to 30 additional nucleotides specific to the cDNA beginning with the second codon. Met represents an initiator methionine and END represents a termination codon. **(A)** Sequence of 5' primer used to create a histidine tail at the amino terminus. The 3' primer should include a second unique restriction enzyme site and the final 5 to 10 codons (including a stop codon) of the cDNA sequence. **(B)** Sequence of 3' (antisense) primer used to create a histidine tail at the carboxyl terminus. The 5' primer should contain a second unique restriction site and the first 5 to 10 codons of the cDNA sequence.

designed primers will permit insertion and expression of the cDNA in the correct reading frame (see *UNITS 3.17, 15.1 & 15.7*).

An alternative to PCR is to subclone the cDNA into an existing vector, synthesize complementary oligonucleotides containing the hexahistidine sequence with compatible restriction enzyme site overhangs, and insert them into the subcloned cDNA (*UNIT 3.16*). It is important to insert the oligonucleotide into the cDNA near one of the termini in the correct reading frame.

Several companies (e.g., Qiagen, Novagen, and Invitrogen; see *APPENDIX 4*) sell expression systems that permit direct subcloning of cDNAs into vectors already containing oligohistidine tail sequences along with adjacent protease cleavage sites to allow removal of the tail after purification. Vectors have been developed for expression of histidine fusion proteins by bacteria, yeast, baculovirus, vaccinia, and various eukaryotic promoters. The protocols that follow are designed for use with a Novagen pET vector (Fig. 10.11B.2) but may easily be modified for other vectors.

NATIVE MCAC FOR PURIFICATION OF SOLUBLE HISTIDINE-TAIL FUSION PROTEINS

**BASIC
PROTOCOL**

This protocol describes expression and purification of histidine fusion proteins in *E. coli*. A desired protein is induced in bacteria, which are then harvested and lysed. The lysate is loaded directly onto a column containing Ni²⁺-NTA resin and the protein is eluted in nearly pure form with MCAC buffer containing imidazole.

Materials

- M9ZB medium (see recipe) containing 50 µg/ml ampicillin and 25 µg/ml chloramphenicol
- E. coli* BL21(DE3)pLysS or other suitable strain (Novagen) containing a pET vector (Fig. 10.11B.2) expressing a histidine-tail fusion protein
- 0.1 M IPTG (Table 1.4.2), filter sterilized
- NTA resin slurry: 50% (w/v) suspension in 20% (v/v) ethanol (Qiagen)
- 100 mM NiSO₄·6H₂O
- MCAC-0, MCAC-20, MCAC-40, MCAC-60, MCAC-80, MCAC-100, MCAC-200, and MCAC-1000 buffers (see recipe)
- 150× protease inhibitor cocktail (see recipe)
- 10% (v/v) Triton X-100
- 1 M MgCl₂
- MCAC-EDTA buffer (see recipe)
- DNase I solution (see recipe)
- Centrifuge with Beckman JA-20 rotor or equivalent
- 1 × 10-cm glass or polypropylene column
- Additional reagents and equipment for growth of bacteria in liquid medium (*UNIT 1.2*) and analysis and processing of purified proteins by SDS-PAGE (see Support Protocol 1)

Express the protein

1. Inoculate 10 ml M9ZB/ampicillin/chloramphenicol with *E. coli* BL21(DE3)pLysS containing a pET vector expressing a desired histidine-tail fusion protein. Grow overnight with shaking at 37°C (*UNIT 1.2*).
2. Inoculate 100 ml M9ZB/ampicillin/chloramphenicol with 1 ml of the overnight culture and grow with shaking at 37°C to OD₆₀₀ = 0.7 to 1.0.

**Analysis of
Proteins**

10.11.11

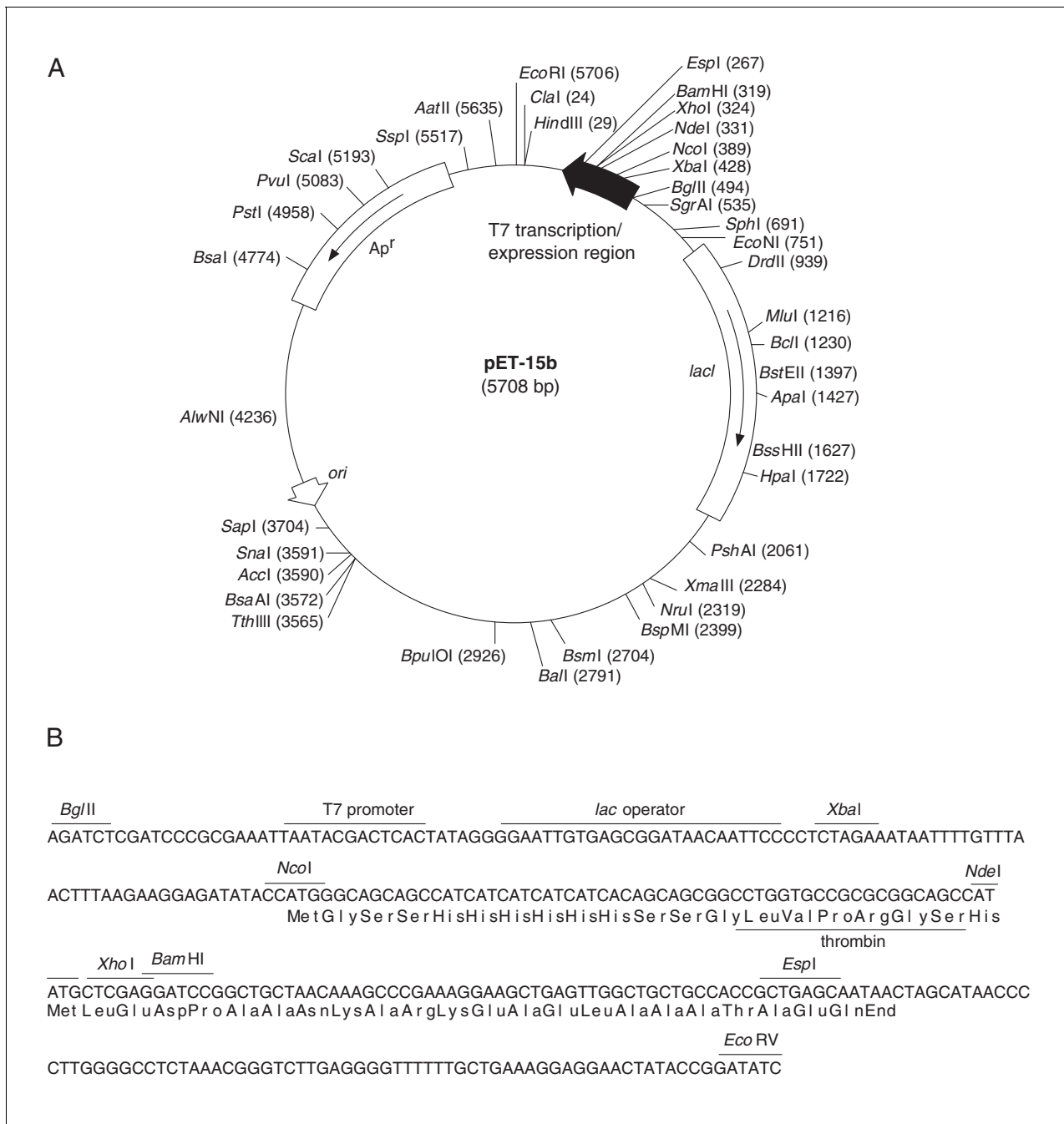


Figure 10.11B.2 pET-15b, one of a series of pET vectors designed for cloning, expression, and purification of recombinant proteins. Reprinted with the permission of Novagen. **(A)** pET-15b vector; **(B)** sequence of pET-15b cloning/expression region. The target gene is cloned into the pET plasmid such that its expression is under the control of bacteriophage T7 transcription and translation signals (see UNIT 16.2). pET-15b encodes an amino-terminal His-Tag leader that allows purification of the resulting recombinant protein over Ni²⁺-NTA resin. Following purification, the His-Tag can be removed by thrombin cleavage (see UNIT 16.7). The pET series are derivatives of vectors originally described by Studier et al. (1990) and are available from Novagen. Figure reproduced by permission of Novagen, Inc.

3. Add 1 ml of 0.1 M IPTG (to 1 mM final) and continue shaking incubation 1 to 3 hr at 37°C.

Incubation time will depend on the solubility of the expressed protein. A 3-hr incubation will allow maximal expression but may result in mostly insoluble protein. A 1-hr incubation will produce less protein but more of it may be soluble.

4. Centrifuge 10 min at 4400 × g (5000 rpm in a Beckman JA-20 rotor), 4°C. Discard supernatant. Freeze pellet at -70°C.

It is not necessary to wash the cell pellet. The wet weight of the cell pellet will be ~0.5 g. The pellet can be stored indefinitely at -70°C before proceeding. Alternatively, extract preparation (steps 9 to 13) can be carried out immediately and the column prepared during the centrifugation at step 13.

Prepare the affinity column

5. Add 0.2 ml NTA resin slurry to a 1 × 10-cm column and allow liquid to drain.

During this and subsequent column washes, liquid should be allowed to drain to the top of the packed resin bed that forms as resin settles (packed bed volume should be 0.1 ml) and resin should not be allowed to dry.

6. Wash column with 1 ml (5 bed volumes) deionized water.
7. Charge column by washing with 1 ml (5 bed volumes) of 100 mM NiSO₄·6H₂O.
8. Wash column with 2 ml MCAC-0 buffer.

NTA resin is nitrilotriacetic acid covalently coupled to Sepharose CL-6B. When charged with nickel ions, it has a light blue-green color, and when stripped of nickel, it is white. See supplier's instructions for resin preparation.

Charged resin can be stored at 4°C. If column is to be stored >1 day, wash with 10 bed volumes of 20% ethanol and add 1 bed volume of 20% ethanol prior to storage. Keep column sealed to prevent evaporation.

Prepare the extract

IMPORTANT NOTE: Beginning with this step, all procedures should be performed on ice or in a cold room unless otherwise indicated.

9. Thaw cell pellet (from step 4) on ice. Add 5 ml MCAC-0 buffer and 33 µl of 150× protease inhibitor cocktail and resuspend by pipetting, sonication, or homogenization.

All MCAC buffers contain phenylmethylsulfonyl fluoride (PMSF) as a protease inhibitor; to reduce expense, protease inhibitor cocktail is added only to the crude extract.

10. Add 0.05 ml of 10% (v/v) Triton X-100 (0.1% final). Mix thoroughly and subject the sample to 3 cycles of freezing at -70°C and thawing on ice.

Ionic detergents may interfere with binding of the protein to the resin and should not be used.

Cell lysis is evidenced by a visible increase in viscosity. The pLysS plasmid in these cells encodes an endogenous lysozyme that eliminates the need for exogenous lysozyme treatment to disrupt the bacterial cell wall.

11. Add 0.05 ml of 1 M MgCl₂ (final concentration 10 mM) and 0.05 ml of DNase I solution (final concentration 10 µg/ml DNase I). Mix gently and incubate 10 min at room temperature.

The DNase I treatment reduces the viscosity of the lysate.

12. Centrifuge 15 min at $27,000 \times g$ (15,000 rpm in JA-20 rotor), 4°C. Decant the supernatant into a clean container on ice and discard the pellet. Set aside and freeze a 10- μ l aliquot at -70°C for later analysis by SDS-PAGE (see Support Protocol 1).

The supernatant can be frozen at -70°C indefinitely before continuing with the procedure.

Purify protein

13. If extract is frozen, thaw on ice. Load onto Ni^{2+} -NTA column and allow to flow through at a rate of 10 to 15 ml/hr. Collect column flowthrough and save for SDS-PAGE (see Support Protocol 1).

Charged NTA resin has a capacity of 5 to 10 mg histidine-tagged protein per milliliter of packed resin. The amount of extract that can be loaded on the column will depend on the amount of soluble histidine-tagged protein in the extract.

14. Wash column with 5 ml MCAC-0 buffer at a flow rate of 20 to 30 ml/hr. Discard flowthrough.
15. Wash column in stepwise fashion with 5 ml each of MCAC-20, MCAC-40, MCAC-60, MCAC-80, MCAC-100, MCAC-200, and MCAC-1000 buffers at a flow rate of 10 to 15 ml/hr. Collect 0.5-ml fractions and save on ice for SDS-PAGE (see Support Protocol 1).

Alternatively, the column can be eluted with a 5-ml linear gradient (UNIT 10.10) of 0 to 400 mM imidazole in MCAC buffer.

The second and third fractions of each wash will contain most of the eluted proteins. Most proteins with hexahistidine tails will remain bound in 60 mM imidazole (MCAC-60) and elute with 100 to 200 mM imidazole (MCAC-100 or -200); therefore, the purified protein will elute in MCAC-100 or -200. Proteins with longer histidine tails (e.g., 10 residues) bind to Ni^{2+} -NTA with greater affinity and require higher imidazole concentrations for elution. However, optimum washing and elution conditions must be determined for each protein.

Once optimum washing and elution conditions are established, it is possible to prepare the crude extract in a buffer that contains the highest imidazole concentration in which the histidine tail remains bound to the Ni^{2+} -NTA (e.g., MCAC-40 or -60 buffer). This decreases nonspecific binding of proteins to resin and permits use of a single buffer for extract preparation, column loading, and column washing.

16. Elute column with 1 ml MCAC-EDTA buffer at a flow rate of 10 to 15 ml/hr, collecting 0.5-ml fractions.

The blue-green color of the column will disappear as nickel is removed by EDTA. More tightly bound proteins may be found in these fractions. The resin can now be recharged (repeat steps 6 to 8) and the column reused.

If protein is eluted adequately with imidazole (as determined by overall yield for subsequent preparations), EDTA washing can be omitted. The column can be reequilibrated with MCAC-0 buffer and the purification repeated. The same column can be used three to five times before EDTA stripping and nickel recharging are necessary. Only one protein should be purified on any given column.

17. Analyze fractions for the presence of eluted protein.

An ultraviolet (280-nm) absorbance flow monitor is helpful for following column elution but is not necessary. An alternative is to measure the OD_{280} of individual fractions to identify protein-containing fractions. However, imidazole will also absorb at 280 nm.

A quick and easy method to determine which fractions contain eluted protein is to place 2 μ l undiluted Protein Assay Dye Reagent Concentrate (Bio-Rad) on a piece of

Parafilm, add 8 μ l from fraction to be tested, and mix by pipetting up and down. Immediate appearance of blue color indicates that the fraction contains protein. This does not work in the presence of Triton X-100 because the detergent itself produces an intense blue color; for this reason, Triton X-100 is excluded from the washing and elution buffers.

18. Combine the fractions containing eluted protein and remove a 10- μ l aliquot for SDS-PAGE (see Support Protocol 1). Freeze the remainder in smaller aliquots at -70°C or in liquid nitrogen.

If a different buffer for the protein is desired (e.g., for proteolytic removal of the histidine tail), the protein should be dialyzed against the buffer of choice to remove the MCAC buffer prior to storage at -70°C or in liquid nitrogen.

19. If time permits, proceed immediately to analysis of fractions by SDS-PAGE and processing of protein (see Support Protocol 1). Otherwise, freeze all samples at -70°C until ready for analysis and processing.

DENATURING MCAC FOR PURIFICATION OF INSOLUBLE HISTIDINE-TAIL FUSION PROTEINS

**ALTERNATE
PROTOCOL 1**

High-level expression of foreign proteins in bacteria and other cells frequently results in poor solubility of the expressed protein (see *UNIT 16.4*). These insoluble proteins form inclusion bodies in bacteria, and strong chaotropic agents such as guanidine, urea, or SDS are usually required to solubilize them. These agents denature the protein and destroy the secondary structure that is essential to other affinity purification methods (e.g., maltose-binding protein or glutathione-S-transferase fusion proteins; *UNITS 16.6 & 16.7*). A significant advantage of metal-chelate affinity chromatography is that the oligohistidine tail will bind to the Ni^{2+} -NTA resin even when the protein is denatured. In denaturing MCAC, the protein extract is solubilized with 6 M guanidine and the entire affinity purification procedure is carried out in guanidine. The purified, denatured protein is renatured during dialysis.

Additional Materials (also see *Basic Protocol*)

GuMCAC-0, GuMCAC-20, GuMCAC-40, GuMCAC-60, GuMCAC-100, and GuMCAC-500 buffers (see recipe)
GuMCAC-EDTA buffer (see recipe)
Appropriate final buffer for protein (e.g., for proteolytic cleavage or long-term storage)
Guanidine-HCl

Additional reagents and equipment for analysis and processing of purified proteins (see Support Protocol 1) and dialysis (*APPENDIX 3C*)

Express the protein

1. Prepare the pellet of *E. coli* expressing a histidine-tail fusion protein (see Basic Protocol, steps 1 to 4).

The pellet can be stored indefinitely at -70°C before proceeding. Alternatively, extract preparation (steps 5 and 6) can be carried out immediately and the column prepared during the centrifugation at step 6.

Prepare the affinity column

2. Prepare column (see Basic Protocol, steps 5 to 7).
3. Wash column with 2 ml GuMCAC-0 buffer.

During this and subsequent column washes, liquid should be allowed to drain to top of packed resin bed and resin should not be allowed to dry.

**Analysis of
Proteins**

10.11.15

Prepare cell extract

4. Thaw cell pellet (from step 1) on ice. Resuspend in 5 ml GuMCAC-0 buffer by pipetting, sonication, or homogenization.
5. Freeze 10 min at -70°C and thaw at room temperature.

Protease inhibitors are omitted because proteases are inactivated by guanidine. Triton X-100 is not needed at this step. Freezing is not necessary but is included because it ensures complete lysis of cells.

Subsequent steps can be performed at room temperature. However, if solid-phase renaturation is used (see Alternate Protocol 2), it is better to maintain lower temperatures throughout the process.

6. Gently mix samples for 30 min using a rocker, rotating mixer, or magnetic stirrer. Centrifuge 15 min at $27,000 \times g$ (15,000 rpm in Beckman JA-20 rotor), 4°C . Decant supernatant into a clean container and discard pellet. Set aside a 10- μl aliquot for analysis by SDS-PAGE (see Support Protocol 1).

The supernatant can be frozen at -70°C indefinitely before continuing with the procedure.

Purify protein

7. If extract from step 5 is frozen, thaw at room temperature. Load onto Ni^{2+} -NTA column and allow to flow through at a rate of 10 to 15 ml/hr. Collect flowthrough and save a 10- μl aliquot for SDS-PAGE (see Support Protocol 1).
8. Wash column with 5 ml GuMCAC-0 buffer at a rate of 20 to 30 ml/hr. Discard the flowthrough.
9. Wash column in stepwise fashion with 5 ml GuMCAC-20, -40, -60, -100, and -500 buffers at a rate of 10 to 15 ml/hr. Collect 0.5-ml fractions and save for SDS-PAGE (see Support Protocol 1).

The second and third fractions from each wash will contain most of the unbound protein. The histidine tail binds slightly less avidly under denaturing conditions. Lower imidazole concentrations are therefore required for washing and elution than in the Basic Protocol.

10. Elute with 1 ml GuMCAC-EDTA buffer at a rate of 10 to 15 ml/hr, collecting 0.5-ml fractions.
11. Identify fractions containing the protein, pool together, transfer to dialysis tubing, and seal.

Alternatively, fractions can be frozen at -70°C indefinitely before continuing with the procedure.

Guanidine precipitates in the presence of SDS and must be removed by dialysis before SDS-PAGE. An alternative technique employs buffers that switch from 6 M guanidine to 8 M urea during affinity column washing (Stüber et al., 1990). This permits samples to be taken directly from urea fractions without dialysis and analyzed by SDS-PAGE or injected into animals for antibody production.

Renature purified protein by dialysis

12. Prepare appropriate final buffer for protein (e.g., for proteolytic cleavage or long-term storage) and add sufficient guanidine to bring final concentration to 4 M.
13. Dialyze purified protein from step 11 for ≥ 2 hr at 4°C against 500 ml buffer/4 M guanidine (see APPENDIX 3C).

The MWCO of the dialysis membrane should be chosen to be smaller than the MW of the purified protein. In most cases an MWCO of 12 to 14 kDa is sufficient.

14. Remove 250 ml buffer/guanidine and add 250 ml buffer without guanidine. Continue dialysis ≥ 2 hr. Repeat.

With some proteins, renaturation by dialysis may require longer dialysis periods and more gradual decrements in the guanidine concentration of the buffer. Conditions for each protein must be determined empirically.

15. Remove dialysis bag to a container containing 500 ml of fresh buffer without guanidine at 4°C. Continue dialysis 2 hr to overnight.

16. Remove sample from dialysis bag, divide into aliquots, and freeze at -70°C or in liquid nitrogen.

If protein precipitates during dialysis, solid-phase renaturation (see Alternate Protocol 2), in which protein bound to the column is renatured before elution, should be employed.

17. Analyze fractions and process protein (see Support Protocol 1).

SOLID-PHASE RENATURATION OF MCAC-PURIFIED PROTEINS

ALTERNATE PROTOCOL 2

In Alternate Protocol 1, removal of denaturants by dialysis will occasionally lead to precipitation of protein, possibly due in part to entanglement or aggregation of separate protein molecules as they refold. To avoid this problem, solid-phase renaturation may be attempted. In this procedure, protein extract is prepared and bound to the column under denaturing conditions. A series of washes removes the denaturing agent before the target protein is eluted and the resulting renatured protein is eluted from the column under native conditions.

Additional Materials (also see *Basic Protocol*)

- 1:1 (v/v) MCAC-20/GuMCAC-20 buffer (see recipes)
- 3:1 (v/v) MCAC-20/GuMCAC-20 buffer (see recipes)
- 7:1 (v/v) MCAC-20/GuMCAC-20 buffer (see recipes)

1. Prepare protein extract, bind to column, and wash with GuMCAC buffers (see Alternate Protocol 1, steps 1 to 9).

2. Wash column with 5 ml of 1:1 (v/v) MCAC-20/GuMCAC-20 buffer.

During this and subsequent washes, liquid should be allowed to drain just to top of packed resin bed and resin should not be allowed to dry.

3. Wash column with 5 ml of 3:1 (v/v) MCAC-20/GuMCAC-20 buffer.

4. Wash column with 5 ml of 7:1 (v/v) MCAC-20/GuMCAC-20 buffer.

5. Wash column with MCAC buffers, elute proteins, and analyze (see Basic Protocol, steps 15 to 19).

Slow elution (between 1 and 2 hr) with a 5-ml linear gradient from 100% GuMCAC-20 buffer to 100% MCAC-20 buffer may also yield efficient renaturation.

ANALYSIS AND PROCESSING OF PURIFIED PROTEINS

The success of the purification scheme (particularly during a small pilot study) should be monitored at each stage by SDS-PAGE. If the fusion protein contains a specific protease cleavage site, the histidine tail can be removed using an appropriate proteolytic procedure, if desired.

Materials

Fractions from MCAC column purification (crude extract, flowthroughs, and purified protein; see Basic Protocol or Alternate Protocols 1 or 2)

2× SDS sample buffer (UNIT 10.2)

MCAC-0 buffer (see recipe)

Additional reagents and equipment for one-dimensional SDS-PAGE (UNIT 10.2), cleavage of proteins with factor Xa or thrombin (UNIT 16.4B; optional), and dialysis (APPENDIX 3C)

1. Thaw aliquots of fractions to be analyzed on ice. Mix 5 µl from crude extract and crude flowthrough fractions and 10 µl from the second and third fractions from each washing step with an equal volume of 2× SDS sample buffer.
2. Load samples onto a standard SDS-PAGE gel. Run gel and visualize to identify the fractions containing purified protein (UNIT 10.2).

Guanidine must be removed by dialysis prior to addition of SDS sample buffer.

3. Thaw the remaining aliquots of fractions containing purified protein, dialyze against the appropriate proteolysis buffer, and carry out cleavage procedure if desired. If necessary, after cleavage dialyze the protein against an appropriate storage buffer and freeze in aliquots.

The size of the cleaved histidine tail will generally be <3 kDa, depending on the design of the fusion protein. This fragment will usually be removed by the dialysis. It may also be removed by ultrafiltration (which will also concentrate the protein), size-exclusion chromatography, or a second MCAC (see Basic Protocol, steps 13 to 19). The advantage of a second MCAC is that the histidine tail and any uncleaved protein will bind to the column and only the cleaved protein will be found in the flowthrough from step 12. Aliquots of each fraction should be analyzed by SDS-PAGE to verify recovery of pure protein.

If an expression vector other than pET has been used, and the fusion protein does not contain an appropriate cleavage site for factor Xa or thrombin, or if the histidine tail will not interfere with subsequent studies, omit the cleavage procedure. Presence of the histidine tail will not generally affect the protein's biological functions (see Background Information).

NTA RESIN REGENERATION

NTA resin can be repeatedly charged with Ni²⁺ and stripped. Over time, however, some protein residue will accumulate and decrease the efficiency of the resin, resulting in slow flow rates or lack of blue-green color after charging with Ni²⁺. Periodic regeneration (e.g., after 5 to 10 cycles of charging and stripping) permits recycling of resin for long-term use.

Materials

2.5 ml used NTA resin (packed volume)

Stripping solution: 0.2 M acetic acid/6 M guanidine-HCl

2% (w/v) SDS

20%, 25%, 50%, 75%, and 100% (v/v) ethanol

0.1 M EDTA, pH 8.0

1. Wash resin with 5 ml stripping solution.
In this and all subsequent washes, liquid added to the column should be allowed to drain to top of packed resin bed and resin should not be allowed to dry.
2. Wash with 5 ml water.
3. Wash with 7.5 ml of 2% (w/v) SDS.
4. Wash with 2.5 ml of 25% (v/v) ethanol.
5. Wash with 2.5 ml of 50% (v/v) ethanol.
6. Wash with 2.5 ml of 75% (v/v) ethanol.
7. Wash with 12.5 ml of 100% (v/v) ethanol.
8. Wash with 2.5 ml of 75% (v/v) ethanol.
9. Wash with 2.5 ml of 50% (v/v) ethanol.
10. Wash with 2.5 ml of 25% (v/v) ethanol.
11. Wash with 2.5 ml water.
12. Wash with 12.5 ml of 0.1 M EDTA, pH 8.0.
13. Wash with 7.5 ml water. Proceed either to nickel charging (see Basic Protocol, step 7) or to step 14 below for long-term storage of the resin.
14. Add 2.5 ml of 20% (v/v) ethanol to resin and store at 4°C.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

DNase I solution

Dissolve lyophilized DNase I (U.S. Biochemical) at a final concentration of 1 mg/ml in 50% glycerol/1 mM CaCl₂. Store ≤1 year at -20°C.

GuMCAC buffers (store ≤6 months at 4°C)

GuMCAC-0 buffer:

20 mM Tris·Cl, pH 7.9
0.5 M NaCl
10% glycerol
6 M guanidine·HCl

GuMCAC-500 buffer:

20 mM Tris·Cl, pH 7.9
0.5 M NaCl
10% glycerol
6 M guanidine·HCl
0.5 M imidazole

GuMCAC-20, GuMCAC-40, GuMCAC-60, and GuMCAC-100 buffers:

These buffers (containing different concentrations of imidazole) are made by mixing GuMCAC-0 buffer and GuMCAC-500 buffer in the appropriate ratios: e.g., for GuMCAC-20 buffer, use 96:4 (v/v) GuMCAC-0/GuMCAC-500.

GuMCAC-EDTA buffer

20 mM Tris·Cl, pH 7.9
0.5 M NaCl
10% glycerol
6 M guanidine·HCl
0.1 M EDTA, pH 8.0
Store ≤6 months at 4°C

M9ZB medium

Dissolve 10 g N-Z-Amine A (Sigma) and 5 g NaCl in 889 ml water. Autoclave, cool, and add 100 ml of 10× M9 medium (see UNIT 1.1), 1 ml of 1 M sterile MgSO₄, and 10 ml of 40% (w/v) glucose (filter sterilized). Store ≤1 year at room temperature.

MCAC buffers (store ≤6 months at 4°C)*MCAC-0 buffer:*

20 mM Tris·Cl, pH 7.9
0.5 M NaCl
10% glycerol
1 mM PMSF (phenylmethylsulfonyl fluoride)

MCAC-1000 buffer:

20 mM Tris·Cl, pH 7.9
0.5 M NaCl
10% glycerol
1 M imidazole
1 mM PMSF

Add PMSF immediately before use from a 0.2 M stock in 100% ethanol stored at room temperature. All buffers used with Ni²⁺-NTA resin contain high salt concentrations to reduce nonspecific electrostatic interactions between proteins and resin. Lower salt concentrations can be used but may lead to nonspecific binding of unwanted proteins to the resin.

*MCAC-20, MCAC-40, MCAC-60, MCAC-80, MCAC-100, and**MCAC-200 buffers:*

These buffers (containing different concentrations of imidazole) are made by mixing MCAC-0 buffer and MCAC-1000 buffer in the appropriate ratios: e.g., for MCAC-60 buffer, 94:6 (v/v) MCAC-0/MCAC-1000.

MCAC-EDTA buffer

20 mM Tris·Cl, pH 7.9
0.5 M NaCl
10% glycerol
0.1 M EDTA, pH 8.0
1 mM PMSF
Store ≤6 months at 4°C

Protease inhibitor cocktail, 150×*Stock solutions:*

2 mg/ml aprotinin in H₂O
1 mg/ml leupeptin in H₂O
1 mg/ml pepstatin A in methanol

Cocktail: Combine 1.5 vol of each stock solution with 5.5 vol sterile water.

Protease inhibitors can be obtained from Sigma or U.S. Biochemical. The stock solutions and cocktail are stable ≥6 months at -20°C.

COMMENTARY

Background Information

Metal-chelate affinity chromatography (also called immobilized metal affinity chromatography or IMAC) for protein purification was first described in 1975 (Porath et al., 1975). It is based on the ability of certain amino acids acting as electron donors on the surface of proteins (histidine, tryptophan, tyrosine, or phenylalanine) to bind reversibly to transition-metal ions that have been immobilized by a chelating group covalently bound to a solid support. Of these amino acids, histidine is quantitatively the most important in mediating the binding of most proteins to immobilized metal ions. Histidine binds selectively to immobilized metal ions even in the presence of excess free metal ions in solution (Hutchens and Yip, 1990b). Copper and nickel ions have the greatest affinity for histidine (Yip et al., 1989).

Proteins and peptides containing phospho-amino acids (phosphoserine, phosphotyrosine, or phosphothreonine) have also been purified by MCAC via selective interaction of the terminal phosphate group with Fe^{3+} (Muszynska et al., 1992). Phosphoproteins also bind to less common metals such as lutetium, scandium, and thorium (Andersson and Porath, 1986).

The chelating group that has been used most extensively for MCAC is iminodiacetic acid (IDA). The tridentate IDA group binds to three sites within the coordination sphere of divalent metal ions such as copper, nickel, zinc, and cobalt. When copper ions (coordination number of 4) are bound to IDA, only one site remains available for interaction with proteins (Hochuli et al., 1987). For nickel ions (coordination number of 6) bound to IDA, three sites are available for binding to proteins. Thus Cu^{2+} -IDA complexes are stable on the column but have lower capacity for protein binding. Conversely, Ni^{2+} -IDA complexes bind proteins more avidly, but the Ni^{2+} -protein complexes are more likely to dissociate from the solid support.

A pentadentate chelating group—*N,N,N'*-tris(carboxymethyl)ethylenediamine (TED)—has also been used for MCAC. This chelator binds five coordination sites of the metal ions, providing highly stable metal ion–TED groups. However, for metal ions with a coordination number of 6, this leaves only one site available for protein interaction.

The development of a new metal-chelating adsorbent, nitrilotriacetic acid (NTA), has pro-

vided a convenient and inexpensive tool for purification of proteins containing histidine residues (Hochuli et al., 1987). The quadridentate NTA moiety covalently coupled via a spacer arm to Sepharose CL-6B forms an extremely stable complex with metal ions possessing a coordination number of 6 (e.g., Ni^{2+}). After binding to the NTA, the Ni^{2+} ion has two sites within its octahedral coordination sphere available for binding to electron donor groups (i.e., histidine) on the surface of proteins. Thus the advantage of NTA over IDA is that the Ni^{2+} ion is bound by four rather than three of its coordination sites. This minimizes leaching of the metal from the solid support and allows for more stringent purification conditions. The NTA also binds Cu^{2+} ions with high affinity, but this occupies all of the coordination sites, rendering the resulting complex ineffective for MCAC.

The affinity of histidine residues for immobilized Ni^{2+} ions allows selective purification of proteins containing a high proportion of histidine residues on the surface. Immobilized copper or nickel ions bind native proteins with a K_d of $1\text{--}17 \times 10^{-5}$ M (Hutchens and Yip, 1990a). This affinity is enhanced significantly by designing proteins to contain short stretches (6 to 10 residues) of histidines in regions likely to be surface exposed (amino or carboxyl termini) using recombinant DNA technology. Addition of a histidine tail results in a protein that binds to the Ni^{2+} -NTA complex with a K_d of 10^{-13} M at pH 8.0 even in the presence of detergent, ethanol, 2 M KCl (Hoffmann and Roeder, 1991), 6 M guanidine (Hochuli et al., 1988), or 8 M urea (Stüber et al., 1990).

The histidine tail binds to the Ni^{2+} -NTA resin via the imidazole side chains of the histidine residues. At $\text{pH} \geq 7.0$, the imidazole side chain is deprotonated, with a net negative charge; at $\text{pH} 5.97$ (the pK of the imidazole side chain of histidine), 50% of the histidines are protonated; and at $\text{pH} \leq 4.5$, almost all of the histidines are protonated and do not interact with Ni^{2+} -NTA. Thus, there are two methods of dissociating histidine tails from the Ni^{2+} -NTA resin. The first method, presented in this unit, uses increasing concentrations of imidazole at constant pH to displace the histidine tail from the Ni^{2+} -NTA. This technique provides great versatility because it can be used for both native and denaturing MCAC and buffer preparation is simplified. The second method uses buffers

of decreasing pH to elute the histidine tail and is quite efficient for producing pure protein (Hochuli et al., 1988). Disadvantages are that the pH must be maintained accurately at all temperatures and that some proteins may not be able to withstand the extreme pH change.

Creation of a small histidine tail has advantages over other fusion protein systems. Addition of six histidines to the protein adds only 0.84 kDa to the mass of the protein, whereas other fusion protein systems utilize much larger affinity groups that often must be removed to allow normal protein function (e.g., glutathione-*S*-transferase, protein A, maltose-binding protein, and lacZ have masses of 26, 30, 40, and 116 kDa, respectively). The small histidine tail is not immunogenic and therefore need not be removed before the purified protein is injected into animals for antibody production. Histidine tail fusion proteins often retain their normal biologic functions—e.g., dihydrofolate reductase and adenylylcyclase retain their enzymatic activities (Hochuli, 1990; Tausig et al., 1993) and the TATA box binding protein retains its transcriptional regulatory function (Parvin et al., 1992). The aforementioned studies also demonstrate the technique's suitability for purification of membrane, cytoplasmic, and nuclear proteins. However, although it often is not necessary to remove the histidine tail, it is possible to include this step and still generate sufficient quantities of protein pure enough for detailed X-ray crystallographic studies (Nikolov et al., 1992).

The histidine tail can also be used to detect the recombinant protein. Ni²⁺-NTA covalently coupled to alkaline phosphate (available from Qiagen; see *SUPPLIERS APPENDIX*) can bind directly to histidine tailed proteins, and the bound alkaline phosphatase can be quantitated by standard techniques (*UNIT 10.8*). This allows for direct quantitation of histidine-tailed proteins blotted to membranes without the need for primary and secondary antibodies (Botting and Randall, 1995).

The protocols presented in this unit utilize protein produced in bacteria; an example protocol for the purification of HIV-1 integrase expressed in *E. coli* is presented in *UNIT 6.5*. Histidine-tail fusion proteins have also been expressed and purified from HeLa cells using transient transfection techniques (Janknecht and Nordheim, 1992; Papavassiliou et al., 1992) or a vaccinia expression system (Janknecht et al., 1991) and from baculovirus-infected insect cells (Gearing et al., 1993; Tausig

et al., 1993). Therefore, MCAC can be used with any protein expression system currently available.

This technique can also be used to prepare immunoaffinity columns (*UNIT 10.11A*). Histidine tails have been covalently linked to oligosaccharide moieties of monoclonal antibodies and then subsequently bound to Ni²⁺-NTA columns (Loetscher et al., 1992). The advantage of this method over more conventional techniques of antibody immobilization is that the immobilized antibody will retain a much greater ability to bind antigen and is recoverable from the column.

Critical Parameters

NTA can be synthesized according to the protocol provided by Hochuli (1990); more commonly, it is purchased from Qiagen, which provides detailed protocols for its use.

Binding and elution of proteins from the Ni²⁺-NTA resin depend upon the histidine content of the protein. Each protein will have a slightly different histidine content and optimal elution profile on the Ni²⁺-NTA resin. Therefore, it is important to establish the buffer imidazole concentrations that will give the best purification of a particular protein. As outlined in the Basic Protocol, gradient or stepwise elution of small-scale preparations should be done before large-scale preparations are attempted. It is also possible to perform batch purification by adding charged resin to a tube containing the extract, mixing for 1 hr, loading into a column, washing, and eluting.

Troubleshooting

The most common problem is low yield of purified protein in the expected fraction. This can be caused by insufficient protein loaded on the column, protein not binding to the column, or protein not eluting from the column.

Insufficient protein

It is important to verify that the desired protein is adequately expressed in the chosen system. Check the crude extract for the presence of the expressed protein by SDS-PAGE or other methods before proceeding to the affinity purification steps.

Protein in the crude lysate may undergo significant proteolytic degradation in some cell types. Always use PMSF, minimize freezing and thawing, maintain low temperatures during purification, and use additional protease inhibitors whenever practical.

Lack of binding

If the protein does not bind to the column, either the protein lacks a histidine tail, the column is inadequately charged with nickel, or the histidine tail is buried in the protein and is inaccessible to the resin. Before expression and purification are attempted, verify that the DNA sequence of the histidine tail and adjacent protein-coding sequence are in frame with each other and that initiation and termination codons are present. If the histidine tail is at the amino terminus, it may be helpful to delete the endogenous codon for initiator methionine; if this is present, translation may initiate downstream of it, producing a full-length protein without a tail. Placing the histidine tail at the carboxyl terminus ensures that only the full-length protein will bind to the Ni²⁺-NTA resin (if protease activity is controlled). However, a histidine tail at the carboxyl terminus is more likely to be inaccessible to the resin.

Chelating agents (e.g., EDTA and EGTA) must be excluded from all solutions because they will strip the Ni²⁺ ions from the affinity column. Strong reducing agents such as dithiothreitol should also be avoided because they reduce the Ni²⁺ ions to metallic nickel, forming a brown precipitate. Some proteins will require the presence of reducing agents to maintain structure. It is possible to use up to 10 mM 2-mercaptoethanol with the Ni²⁺-NTA resin, although the lowest possible concentration should be used.

The Ni²⁺-NTA should retain a distinct blue-green color throughout the purification process until EDTA is applied. If the resin is white or if the blue-green color is faint after charging the column with Ni²⁺, check all buffers and make sure no chelating and reducing agents are present. Recharge the column with Ni²⁺ and if the color is still faint or absent, regenerate the NTA resin (see Support Protocol 2) and repeat Ni²⁺ charging (see Basic Protocol, steps 6 to 8). If the resin is still white, discard it and start over with a new batch.

If the protein has or should have a histidine tail but is not binding to the resin, perform the purification under denaturing conditions (see Alternate Protocol 1 or 2) to unmask the inaccessible histidine tail. Placing the histidine tail at the opposite terminus of the protein may also be helpful.

Failure to elute

If the protein appears to be binding to the column but not eluting, make a final elution with EDTA-containing buffer. Any protein

bound to the Ni²⁺-NTA will be removed with the EDTA. Occasionally the protein bound to the resin will precipitate and not elute; if this problem is suspected, purify the protein under denaturing conditions and analyze all eluted fractions by SDS-PAGE for the presence of the expressed protein.

Anticipated Results

The maximum binding capacity of Ni²⁺-NTA resin is 5 to 10 mg of protein per milliliter of packed resin. Under ideal conditions, as much as 80% of the bound protein can be recovered. The limiting factor in most instances will probably be the amount of protein loaded on the resin, which depends on the protein expression system.

Time Considerations

Expression of protein and preparation of crude extract require 1 day for the system described here. Column preparation requires 30 to 60 min. Loading, washing, and eluting the column require 4 to 6 hr.

Literature Cited

- Andersson, L. and Porath, J. 1986. Isolation of phosphoproteins by immobilized metal (Fe³⁺) affinity chromatography. *Anal. Biochem.* 154:250-254.
- Botting, C.H. and Randall, R.E. 1995. Reporter enzyme-nitrilotriacetic acid-nickel conjugates: Reagents for detecting histidine-tagged proteins. *Biotechniques* 19:362-363.
- Gearing, K., Göttlicher, M., Teboul, M., Widmark, E., and Gustafsson, J. 1993. Interaction of the peroxisome-proliferator-activated receptor and retinoid X receptor. *Proc. Natl. Acad. Sci. U.S.A.* 90:1440-1444.
- Hochuli, E. 1990. Purification of recombinant proteins with metal chelate adsorbent. *In Genetic Engineering, Principles and Practice*, Vol. 12 (J. Setlow, ed.) pp. 87-98. Plenum, New York.
- Hochuli, E., Dobeli, H., and Schacher, A. 1987. New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J. Chromatogr.* 411:177-184.
- Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R., and Stüber, D. 1988. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *BioTechnology* 6:1321-1325.
- Hoffmann, A. and Roeder, R. 1991. Purification of His-tagged proteins in nondenaturing conditions suggests a convenient method for protein interaction studies. *Nucl. Acids Res.* 19:6337-6338.
- Hutchens, T.W. and Yip, T.-T. 1990a. Protein interactions with immobilized transition metal ions: Quantitative evaluations of variations in affinity

- and binding capacity. *Anal. Biochem.* 191:160-168.
- Hutchens, T.W. and Yip, T.-T. 1990b. Differential interaction of peptides and protein surface structures with free metal ions and surface-immobilized metal ions. *J. Chromatogr.* 500:531-542.
- Janknecht, R., De Martynoff, G., Lou, J., Hipskind, R., Nordheim, A., and Stunnenberg, H. 1991. Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus. *Proc. Natl. Acad. Sci. U.S.A.* 88:8972-8976.
- Janknecht, R. and Nordheim, A. 1992. Affinity purification of histidine-tagged proteins transiently produced in HeLa cells. *Gene* 121:321-324.
- Loetscher, P., Mottlau, L., and Hochuli, E. 1992. Immobilization of monoclonal antibodies for affinity chromatography using a chelating peptide. *J. Chromatogr.* 595:113-119.
- Muszynska, G., Dobrowolska, G., Medin, A., Ekman, P., and Porath, J.O. 1992. Model studies on iron(III) affinity chromatography. II. Interaction of immobilized iron(III) ions with phosphorylated amino acids, peptides and proteins. *J. Chromatogr.* 604:19-28.
- Nikolov, D., Hu, S., Lin, J., Gasch, A., Hoffman, A., Horikoshi, M., Chua, N., Roeder, R., and Burley, S. 1992. Crystal structure of TFIID TATA-box binding protein. *Nature* 360:40-46.
- Papavassiliou, A., Bohmann, K., and Bohmann, D. 1992. Determining the effect of inducible protein phosphorylation on the DNA-binding activity of transcription factors. *Anal. Biochem.* 203:302-309.
- Parvin, J., Timmers, H., and Sharp, P. 1992. Promoter specificity of basal transcription factors. *Cell* 68:1135-1144.
- Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. 1975. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* 258:598-599.
- Stüber, D., Matile, H., and Garotta, G. 1990. System for high-level production in *Escherichia coli* and rapid purification of recombinant proteins: Application to epitope mapping, preparation of antibodies, and structure-function analysis. *Immunol. Methods* 4:121-152.
- Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. 1990. Use of T7 polymerase to direct expression of cloned genes. *Methods Enzymol.* 185:60-89.
- Taussig, R., Quarmby, L., and Gilman, A. 1993. Regulation of purified type I and type II adenylcyclases by G protein β - γ subunits. *J. Biol. Chem.* 268:9-12.
- Yip, T.-T., Nakagawa, Y., and Porath, J. 1989. Evaluation of the interaction of peptides with Cu(II), Ni(II), and Zn(II) by high-performance immobilized metal ion affinity chromatography. *Anal. Biochem.* 183:159-171.

Key Reference

Hochuli, 1990. See above.

Describes basic principles of MCAC with detailed protocols.

Contributed by Kevin J. Petty
University of Texas Southwestern
Medical Center
Dallas, Texas

PURIFICATION OF PROTEINS BY HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY

HPLC of Peptides and Proteins: Preparation and System Set-Up

High-performance liquid chromatography (HPLC) is an essential tool for the purification and characterization of biomacromolecules. The choice of the chromatographic method and the type of high-performance equipment are determined by the molecular nature of the investigated molecules and the aim of the research. There are eight basic modes of HPLC currently in use for peptide and protein analysis and purification, namely size-exclusion chromatography (HP-SEC), ion exchange chromatography (HP-IEX), normal phase chromatography (HP-NPC), hydrophobic interaction chromatography (HP-HIC), reversed-phase chromatography (RP-HPLC), hydrophilic interaction chromatography (HP-HILIC), immobilized metal ion affinity chromatography (HP-IMAC), and biospecific/biomimetic affinity chromatography (HP-BAC), and a number of subsets of these chromatographic modes, e.g., mixed mode chromatography (HP-MMC), charge transfer chromatography (HP-CTC), or ligand-exchange chromatography (HP-LEC).

In terms of usage, versatility, and flexibility, RP-HPLC techniques dominate the application world with peptides and proteins at the analytical- and laboratory-scale preparative levels (see *UNITS 10.13 & 10.14*). All of these various chromatographic modes can be operated under isocratic or gradient elution conditions, and in analytical or preparative situations. They all have common start-up procedures, which are outlined in this unit, and specific standard conditions that are detailed for the major modes and which represent starting points for further method development, as presented in *UNIT 10.13*.

THE PROPERTIES OF PEPTIDES AND PROTEINS AND THEIR IMPLICATIONS FOR HPLC METHOD DEVELOPMENT

Peptides and proteins are a class of molecules containing amino acids as the fundamental units. The chemical organization (i.e., the primary structure or amino acid sequence) and

the folded structure (i.e., the secondary, tertiary and quaternary structure) are the essential features of a polypeptide or protein, around which a chromatographic separation can be designed. Two sets of factors must be considered. The first relates to the structural properties of the amino acid entities themselves; the second relates to the chemical and physical attributes of the separation system per se.

Biophysical Properties of Peptides and Proteins

The 20 naturally occurring L- α -amino acids found in peptides and proteins vary dramatically in terms of the properties of the side chain or R-groups. Table 10.12.1 lists some of the fundamental properties of the common L- α -amino acids found in peptides and proteins. This chemical diversity becomes even greater in circumstances where some of these side chains have been post-translationally modified with carbohydrates or lipid moieties. The side-chains are generally classified according to their polarity (e.g., non-polar or hydrophobic versus polar or hydrophilic). The polar side chains are divided into three groups: uncharged polar, positively charged or basic, and negatively charged or acidic side chains. Peptides and proteins generally contain several ionizable basic and acidic functionalities. They therefore typically exhibit characteristic isoelectric points with the overall net charge and polarity in aqueous solution varying with pH, solvent composition and temperature. Cyclic peptides without ionizable side chains will have zero net charge, and they represent an exceptional subgroup.

The number and distribution of charged groups will influence the polarizability and ionization status of a peptide or protein, as well as the microscopic and global hydrophobicity. These important factors ultimately determine the selection of the optimal separation conditions for the resolution of peptide and protein mixtures. Table 10.12.1 can be used to evaluate

Table 10.12.1 Properties of the Common L- α -Amino Acid Residues

3-Letter code	1-Letter code	Mass (Da)	Partial spec. vol. (\AA^3) ^a	Accessible surface area (\AA^2) ^b	pK _a of side-chain ^c	Rel. hydrophobicity ^d
Ala	A	71.08	88.6	115	—	0.06
Arg	R	156.19	173.4	225	12.48	-0.85
Asn	N	114.10	117.7	160	—	0.25
Asp	D	115.09	111.1	150	3.9	-0.20
Cys	C	103.14	108.5	135	8.37	0.49
Gln	Q	128.13	143.9	180	—	0.31
Glu	E	129.12	138.4	190	4.07	-0.10
Gly	G	57.05	60.1	75	—	0.21
His	H	137.14	153.2	195	6.04	-2.24
Ile	I	113.16	166.7	175	—	3.48
Leu	L	113.16	166.7	170	—	3.50
Lys	K	128.17	168.6	200	10.54	-1.62
Met	M	131.20	162.9	185	—	0.21
Phe	F	147.18	189.9	210	—	4.8
Pro	P	97.12	122.7	145	—	0.71
Ser	S	87.08	89	115	—	-0.62
Thr	T	101.11	116.1	140	—	0.65
Trp	W	186.21	227.8	255	—	2.29
Tyr	Y	163.18	193.6	230	10.46	1.89
Val	V	99.13	140	155	—	1.59
α -amino ^e	—	—	—	—	7.7-9.2 ^g	—
α -carboxyl ^f	—	—	—	—	2.75-3.2 ^g	—

^aFrom Zamyatin, 1972.^bFrom Chothia, 1974.^cFrom Dawson et al., 1986.^dFrom Wilce et al., 1995.^e α -amino group present on all primary amino acids.^f α -carboxyl group present on all primary amino acids.^gFrom Rickard et al., 1991.

the impact of amino acid composition on retention behavior. For example, this information can be used to direct the choice of eluent composition or the gradient range in RP-HPLC; to assess the impact on retention of amino acid substitution or deletion with small peptides; or alternatively to guide the identification of peptide fragments derived from tryptic digestion of proteins for further sequence analysis.

Parameters of the Mobile Phase/Stationary Phase

These parameters directly impact on the molecular properties of the polypeptide or protein during liquid chromatographic separations, and are listed in Table 10.12.2. In solution, a polypeptide or protein can, in principle, explore a relatively large array of conforma-

tional space. For small peptides (up to ~15 amino acid residues) a defined secondary structure (α -helical, β -sheet or β -turn motif) is generally absent. With increasing polypeptide chain length, depending on the nature of the amino acid sequence, specific regions/domains of a polypeptide or protein can adopt preferred secondary, tertiary or quaternary structures. In aqueous solutions this folding, which internalizes the hydrophobic residues and thus stabilizes the polypeptide structure, becomes a significant feature of peptides and proteins for chromatographic separations. A critical factor in the selection of an HPLC procedure is that the choice of experimental conditions will inevitably cause perturbations of the conformational status of these biomacromolecules. Although polypeptide and protein conformational

Table 10.12.2 Chemical and Physical Factors of the Chromatographic System that Contribute to the Variation in the Resolution and Recovery of Polypeptides, Proteins and Other Biomacromolecules in HPLC Systems (Hearn, 2000)

Mobile phase contributions	Stationary phase contributions
Organic solvents	Ligand composition
pH	Ligand density
Metal ions	Surface heterogeneity
Chaotropic reagents	Surface area
Oxidizing or reducing reagents	Pore diameter
Temperature	Pore diameter distribution
Buffer composition	Particle size
Ionic strength	Particle size distribution
Loading concentration and volume	Particle compressibility

Table 10.12.3 Relevant Absorption Bands and Extinction Coefficients in Proteins (Campbell and Dwek, 1984)

Group	Wavelength (nm)	Log ϵ_{\max}
Peptide bond	190-210	2.0-3.8
His	211	3.8
Cys	250	2.5
Trp	280	3.7
Tyr	274	3.1
Phe	257	2.3

stability can be manipulated in a number of ways (e.g., mobile and stationary phase composition, temperature) in HPLC, in most cases an integrated biophysical experimental strategy—including ^1H 2-dimensional NMR, Fourier transformed infrared (FTIR), ESI-MS (UNIT 10.21), or circular dichroism–optical rotatory dispersion (CD-ORD) spectroscopy—is required in order to determine the secondary and higher-order structure of a polypeptide or protein in solution or in the presence of specific ligands. Availability of such instrumentation is not mandatory, but the quality of the interpretation of the experimental results will become more substantial when additional results are independently obtained with such spectrometric procedures to confirm the participation of conformational or self-self aggregation effects with peptides or proteins under HPLC conditions.

DETECTION OF PEPTIDES AND PROTEINS IN HPLC

The peptide bond absorbs strongly in the far-ultraviolet (UV) region of the spectrum ($\sim\lambda = 205$ to 215 nm). Hence UV detection is the most widely used method for detection of peptides and proteins in HPLC (Table 10.12.3). Besides absorbing in the far-UV range, the aromatic amino acid residues (and to some extent cysteine) also absorb light above 250 nm. Knowledge of the UV spectra, in particular the extinction coefficients of the non-overlapping absorption minima of these amino acids, allows, in conjunction with UV-diode array detection (DAD) and second derivative or difference UV-spectroscopy, verification of peak purity and determination of the aromatic amino acid content of peptides and proteins. Moreover, the knowledge of the relative UV/VIS absorbency of a peptide or protein is therefore crucial, since the choice of detection wavelength of peptides and proteins in RP-HPLC (and in the other HPLC modes) depends on the

Table 10.12.4 UV Cutoff Values of Different Organic Solvents in RP-HPLC

Eluent	UV cutoff (nm) ^a
Methanol	205
Acetonitrile	188
Isopropanol	205

^aWavelength at which the absorbance of a 1-cm-long cell filled with the solvent was 1.0, measured against water as reference.

different UV cutoffs of the eluents used (Table 10.12.4). The common use of $\lambda = 215$ nm as the preferred detection wavelength for most analytical reversed-phase applications (and for those of other HPLC modes) with peptides and proteins is a good compromise between detection sensitivity and potential detection interference due to buffer absorption. However, wavelengths between 230 and 280 nm are frequently employed in preparative applications, where the use of more sensitive detection wavelengths could result in overloading of the detector response (usually above an absorbance value of 2.0 to 2.5 AU).

START-UP PROCEDURES

Correct selection of these first important steps may take more time than the ultimate experimental procedure if a high-performance separation of high resolution, robustness, and reproducibility is to be achieved. They require good planning and thorough work. The following details are representative of the types of equipment, materials, chemicals, and experimental protocols that can be routinely required for isocratic or gradient elution HPLC.

Sample

Peptide or protein sample (kept at 4°C if not used)

Apparatus

Pump module
Mixing chamber
Spectrophotometer with analytical or preparative flow cell
Injection valve
Analytical (10 to 100 μ l) or preparative (500 to 1000 μ l) sample loop
Column oven or thermostated column coolant-jacket coupled to recirculating cooler
Autosampler (optional)

Computer, printer, and software, e.g., Beckman, System Gold; Hewlett-Packard HP-1090A liquid chromatograph, or Waters 600/486 HPLC system with attendant data management systems and system automation controllers

Chemicals

Acetonitrile (HPLC grade)
Methanol (HPLC grade)
Acetone (HPLC grade)
Thiourea or sodium nitrate
Milli-Q water

Glassware

Two 1-liter eluent bottles
Two 1-liter measuring cylinders
10-ml measuring cylinder
Waste bottle
All glassware coming into contact with sample before and during analysis should be rinsed three times with Milli-Q water

Mobile phase filtration facility

Vacuum pump
1-liter reservoir
Support base with glass frit and integral vacuum connection
Funnel
Clamp
47-mm membrane filter (0.2 μ m PTFE)

Gases

Helium
Nitrogen (for autosampler)

Columns

See respective sections below and in *UNIT 10.13*

HPLC peptide standards

See respective sections below and in *UNIT 10.13*

Tools

Screwdrivers:

1/8-in. and 1/4-in. flat-head screwdrivers
Phillips screwdriver no. 2

Wrenches:

12-in. adjustable wrench (for compressed gas tank regulator)
Three open-end wrenches (two 1/4-in. × 1/16-in.; one 1/2-in. × 1/16-in.) for fittings columns and valves
Two long-jaw needle-nose pliers
Two bastard files
Two hex key (Allen) wrench sets (metric and nonmetric)
Tweezers
Pump seal insertion tool (if required)
Inner reamer
Teflon tape
Flashlight
Magnetic pick-up tool

Logbook

All steps must be documented to facilitate troubleshooting and reproduction

Spare parts

Zero dead volume union
Ferrules (steel, rheodyne) 1/16-in. short and long (for rheodyne valves)
Bushing nuts
One-piece PEEK fitting
Tubing (steel, PEEK)
Column frits
In-line filter
Inlet filter
Fuses

Miscellaneous

Graduated 25- μ l Hamilton glass syringe
1-ml syringe with truncated needle
Conical vials for autosampler
Laboratory coat
Gloves
Safety glasses
Stopwatch

PREPARATION OF THE SAMPLE

The following considerations are relevant to the preparation of samples containing peptides and proteins.

1. Dissolve the sample with half the target volume of eluent A (weak mobile phase). If the sample is not soluble, a small amount of eluent B (strong mobile phase) may be added (typically <25% of total volume). Small amounts of strong mobile phase may cause pre-elution of the sample depending on the sample loop size,

column dimensions, and starting mobile phase composition of the gradient.

2. Inspect the sample for clearness and filter through a 0.2 μ m PTFE filter if insoluble, opalescent, or solid particles are present. Alternatively centrifuge the sample using the supernatant for injection. Samples that are not fully dissolved should not be injected, since they can block injector and column.

3. Store sample if not in use at 4°C or -20°C depending on the planned storage time and usage. Peptides and protein samples can degrade at room temperature. Avoid repeated freeze-thawing of the sample. Rather, prepare small aliquots of the peptide or protein sample, which are kept at -20°C and are used for one task.

PREPARATION OF THE MOBILE PHASE

The following considerations are relevant to the preparation of mobile phases to be used with samples containing peptides and proteins.

1. Prepare 1 liter each of eluent A (weak mobile phase) and B (strong mobile phase).
2. Mix each eluent either by stirring with a magnetic “flea” or by shaking a stoppered cylinder (time depends on volume).
3. Filter the eluents, first A, then B, through a 0.2 μ m PTFE filter. Filtering of eluents increases the column lifetime and contributes to degassing.
4. Close unused eluent bottles with a stopper to avoid evaporation of the organic solvent.

SETTING UP THE HPLC INSTRUMENT

The following steps can be done either with or without a column connected to the HPLC instrumentation.

1. Switch on gas supply, nitrogen for the operation of the autosampler (if available), and helium for the degassing of the eluents.
2. Degas the eluents for 10 min at a rate of 100 ml/min, then reduce to a rate of 20 ml/min, which may be maintained throughout the experiments. Avoid excessive sparging since this will change the eluent composition.
3. Switch the detector lamp on in order to warm it up.

PRIMING OF THE PUMPS AND LOW PRESSURE LINES WITH ELUENTS

The following steps can be done either with or without a column connected to the HPLC instrumentation.

1. Open purge valve.
2. Set the pump at 100% buffer A and a flow rate of 1 ml/min. Under these settings, the eluent will bypass the column and travel directly to the waste fluid container (all wastes are collected for appropriate disposal).
3. Prime all tubing lines by switching the selector to the "prime line" position or by opening an additional valve which has an outlet to attach a syringe (see manufacturers' handbook for specific details).
4. Draw (via the opened valve) 20 ml out of the eluent bottle with the syringe. Close the valve and discharge eluent from syringe in waste.
5. Prime all tubing lines again by drawing 10 ml in syringe, but keep the eluent in the syringe.
6. Switch selector to the "prime pump" position and push eluent gently through the pump.
7. Repeat the last four steps with 100% buffer B.
8. Put the selector in the "operate" position, close the valve, and remove the syringe from instrument.
9. Close the purge valve (now the eluent will travel through the column). This procedure should expel air bubbles from the pump heads and replace the previous eluent (~20 ml) in the solvent lines.

PREPARATION OF THE HPLC SYSTEM

The considerations below are relevant to the preparation of the HPLC system for use with samples containing peptides and proteins (see *UNIT 10.13*). The following steps can be done either with or without a column connected to the HPLC instrumentation.

1. Test the pump delivery system with eluents comprising 100% A and 100% B each for 5 min at a flow rate of 1 ml/min, collecting the eluent into a 10 ml cylinder. Some HPLC systems allow on-line pump diagnostics (e.g., Beckman System Gold). This procedure tests the reliability of the inlet and outlet valve, which may be blocked or not closing properly, e.g., due to salt from previous eluents. Some instrument systems allow the removal of the valve, which then can be sonicated in methanol for 15 min (care must be taken not to mix up the inlet and outlet valves, which often cannot be visually distinguished).
2. Flush the HPLC system with the eluent (e.g., 50% A and 50% B) and monitor the detector baseline. If spikes occur after 15 min of flushing in the absence of the column, the

pressure in the detector cell is too low, leading to an outgassing of air and cycling bubble formation. Use the back-pressure restrictor (alternatively a restriction capillary) on the detector outlet to slightly, and very carefully, enhance the back-pressure. Every piece of capillary must be checked for blockage beforehand by connecting to a pump, thus bypassing the column and detector, since the detector cell is very sensitive to high pressure (consult manufacturer's handbook for details).

3. Flush the needle port with the eluent B using a truncated 1-ml syringe in systems with manual injector. This procedure removes sample residues from previous injections.
4. Flush the Hamilton glass syringe (for manual injection) with methanol, then water to rinse it.
5. Flush the sample loop (in the load position) with three times the sample loop volume with eluent B using a Hamilton glass syringe.

THE GRADIENT DELAY (DWELL VOLUME) OF THE HPLC SYSTEM

The following steps must be done without a column connected to the HPLC instrumentation.

1. Connect the injector directly to the detector with union piece (zero length column).
2. Prepare a special eluent A and B, 200 ml each:

Eluent A: acetonitrile

Eluent B: acetonitrile/0.2% acetone.

3. Run a gradient of 10% to 90% B in 10 min at a flow rate of 2.0 ml/min. Detection is carried out at 254 nm. The measured value of the dwell volume can be influenced by the injection technique. If after the injection the valve remains in the inject position, the dwell volume will include the volume of the sample loop; if the valve is put back in load position, the dwell volume will not. This effect can produce errors with sample loops >100 μ l. The same consideration is valid if the sample loop is exchanged from an analytical separation (e.g., a sample loop of 50 μ l) to a semipreparative separation (e.g., a sample loop of \geq 500 μ l) on the same column.

4. Determine the gradient delay and present the results graphically in a format similar to that shown as Figure 10.12.1.

The dwell volume, V_D , is the volume of eluent from the pump heads to the column inlet (including the mixing chamber volume). The dwell volume values range from 2 to 7 ml; autosamplers in particular make a large contribution to the delay volume. It should be deter-

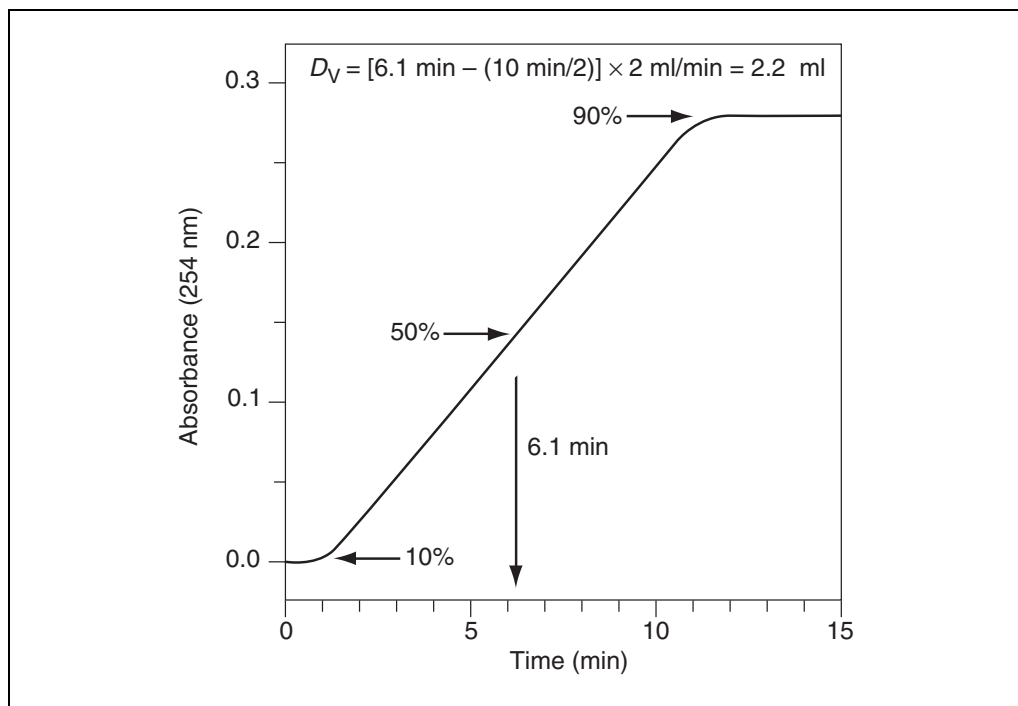


Figure 10.12.1 Graphical illustration of the approach employed to determine the gradient delay volume, D_V . In this figure the gradient profile is illustrated for a defined flow rate (2 ml/min), with the gradient profile recorded from 10% to 90% buffer B at a specified wavelength such as 254 nm, as described in the text with acetonitrile-water-acetone mixtures.

mined with an accuracy of ± 0.5 ml. The profile can be used for diagnostic purposes, since the volume accuracy of the pump delivery is also monitored.

Knowledge of the gradient delay is essential for method development, since it allows the accurate calculation of the S and k_0 values (Ghrst and Snyder, 1988; Hearn, 1991). Its determination is particularly important when establishing segmented gradients (since various errors can accumulate here), and when an established HPLC method is transferred from one instrument to another instrument.

CONNECTING THE COLUMN

The following considerations are relevant to the preparation of the HPLC column for use with samples containing peptides and proteins.

1. Flush the column with eluent B with the inlet connected to the injector and the outlet facing the waste collector (5 min at 1 ml/min for analytical columns). This procedure removes air that may have been trapped and replaces the storage buffer.

2. Connect the column outlet to the detector. Start the flow with 0.5 ml/min and slowly increase the flow rate to 1 ml/min.

3. Equilibrate the column first with eluent B until a stable baseline is reached or alterna-

tively with 10 column volumes (~ 15 min for analytical column at 1 ml/min) and then with eluent A again with 10 column volumes. The pressure should be monitored and documented for each eluent since it can be used for diagnostic purposes.

PROGRAMMING THE HPLC INSTRUMENT

The following considerations are relevant to the programming of the HPLC instrument for use with samples containing peptides and proteins.

1. Program, according to the manufacturer's handbook, the pump, the detector, the integration module, and the autosampler. Test the method with a test run before leaving the instrument alone.

2. Program a shutdown method for overnight runs, which will switch off the lamp and pump. This approach prolongs lamp life and saves eluents.

INJECTING THE SAMPLE

The following considerations are relevant to the injection of the samples containing peptides and proteins onto the HPLC column.

1. Switch the sample loop to the load position and rinse with eluent A. This procedure

removes eluent B (which may be present from rinsing the loop or previous runs). Failure to do so can cause pre-elution of the peptide or protein sample, particularly in conjunction with partially filled sample loops.

2. Load the sample slowly into the loop avoiding air bubbles. Do not squirt the sample into the loop too fast as it will end up in the waste.

3. Inject the sample by switching the valve swiftly into the inject position. If the switching is done too slowly, the pumps might shut down because the pressure limit is exceeded, as the valve is blocked in the intermediate switching position.

TESTING THE FUNCTIONAL HPLC SYSTEM

The following considerations are relevant to the testing of the HPLC system for use with samples containing peptides and proteins.

1. Produce a blank run (inject eluent A) and run a gradient from 100% eluent A to 100% eluent B under the same conditions as intended for the peptide or protein sample. Repeat if peaks occur. This procedure cleans the column of peptides and proteins from previous separations, which have not been removed by the flushing process.

2. Measure the dead volume of the column with thiourea or sodium nitrate (or any other noninteractive solute).

3. Test the column performance with a gradient run and an appropriate test mixture (see specific sections in *UNIT 10.13* for details). First, this test allows the evaluation of the column bed integrity (low integrity will be associated with split, fronting or tailing peaks) and column performance (in terms of plate numbers). Second, this test allows, if repeated at regular intervals, the monitoring of the performance during the lifetime of a column, and the assessment of batch-to-batch differences of column fillings.

LOGBOOKS

Record keeping is essential for liquid chromatography system maintenance and confirmation/substantiation of the experimental results. Three types of logbooks are recommended, the system logbook, the column logbook, and the experiment/assay logbook (Dolan and Snyder, 1989).

System Logbook

This logbook should contain information on:

1. The module identification (brand, model, serial number, purchase data, warranty information) for the entire liquid chromatography (LC) system: injector, autosampler, pump(s), detector, software, column.

2. Module replacements, maintenance records and upgrades.

3. Reference chromatograms and operating parameters.

4. Maintenance: activity, time, operator, and name/reliability of service engineer.

5. Column replacement (cross-reference to column logbook).

Column Logbook

This logbook should contain information including a summary of use, column life in months and number and type of samples, cause of failure, and suggestions for extending life. For each column this would be tabulated as follows.

1. Date column first used.
2. Specification: mobile phase flow rate, plate number, peak shape, dead volume over the lifetime of the column.
3. Performance of new column (validated with a test mixture).
4. Record of use (instrument, operator, number and type of samples).
5. Storage information.
6. Maintenance performed (type of back-flush protocols, frit replacement, etc.).
7. Reevaluation of column performance.
8. Cause of failure.

Experiment/Assay Logbook

This logbook should contain information on:

1. Equipment configuration.
2. Operating condition(s).
3. Mobile phase recipe(s) (literature references if available).
4. Sample pre-treatment method(s) (literature references if available).
5. Assay procedure(s) (literature references if available).
6. Data analysis procedure(s).

UNIT 10.13 presents illustrative examples of standard operating protocols for the major modes of HPLC, based on the instrumental system validation procedures described above.

LITERATURE CITED

- Campbell, I.D. and Dwek, R.A. 1984. Biological Spectroscopy pp. 255-277. Benjamin/Cummings, Menlo Park, N.J.
- Chothia, C. 1974. Hydrophobic bonding and accessible surface area in proteins. *Nature* 248:338-339.
- Dawson, R.M.C., Elliot, D.C., Elliot, W.H., and Jones, K.M. 1986. Data for Biomedical Research. Clarendon Press, Oxford.
- Dolan, J.W. and Snyder, L.R. 1989. Troubleshooting LC Systems pp. 1-515. Humana Press, Clifton, N.J.
- Ghrist, B.F.D. and Snyder, L.R. 1988. Design of optimized high-performance liquid chromatographic gradients for the separation of either small or large molecules. II. Background and theory. *J. Chromatogr.* 459:25-41.
- Hearn, M.T.W. 1991. High-performance liquid chromatography of peptides and proteins: General principles and basic theory. In *High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis and Conformation* (C.T. Mant and R.S. Hodges, eds.) pp. 95-104. CRC Press, Boca Raton, Fla.
- Hearn, M.T.W. 2000. Physico-chemical factors in polypeptide and protein purification and analysis by high-performance liquid chromatographic techniques: Current status and future challenges. In *Handbook of Bioseparation* (S. Ahuja, ed.) pp. 72-235. Academic Press, New York.
- Rickard, E.C., Strohl, M.M., and Nielsen, R.G. 1991. Correlation of electrophoretic mobilities from capillary electrophoresis with physico-chemical properties of proteins and peptides. *Anal. Biochem.* 197:197-207.
- Wilce, M.C.J., Aguilar, M.I., and Hearn, M.T.W. 1995. Physicochemical basis of amino acid hydrophobicity scales: Evaluation of four new

scales of amino acid hydrophobicity coefficients derived from RP-HPLC of peptides. *Anal. Chem.* 67:1210-1219.

Zamyatin, A.A. 1972. Protein volume in solution. *Progress Biophys. Mol. Biol.* 24:107-123.

KEY REFERENCES

- Dolan and Snyder, 1989. See above.
An excellent practical manual.
- Hearn, M.T.W. 1991. HPLC of Peptides, Proteins and Polynucleotides, Fundamental Principles and Contemporary Applications. pp. 1-776. VCH Publishers, Deerfield Beach, Fla.
A comprehensive text encompassing most applications.
- Mant, C.T. and Hodges, R.S. 1991. High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis and Conformation. pp. 1-938. CRC Press, Boca Raton, Fla.
A comprehensive survey of applications.
- McMaster, M.C. 1994. HPLC: A Practical User's Guide. pp. 1-211. VCH Publishers, Inc., New York.
A general guide to HPLC.
- Vijayalakshmi, M.A. 2000. Theory and Practice of Biochromatography. pp. 1-721. Harwood Academic Publishers.
A useful guide to modern chromatography.

Contributed by Reinhard I. Boysen
and Milton T.W. Hearn
Monash University
Victoria, Australia

HPLC of Peptides and Proteins: Standard Operating Conditions

This unit builds upon the general discussion of HPLC preparation and system setup in UNIT 10.12 by presenting standard operating conditions for HP-SEC, HP-NPC, HP-HIC, HP-IEX, HP-HILIC, HP-IMAC, HP-BAC, and RP-HPLC. As an example for appropriate method development, the RP-HPLC mode is described in greater detail and discussed according to four possible intended purposes, i.e.: (1) the purification of one component out of a natural or synthesized sample; (2) the simultaneous purification of several components; (3) the desalting of proteins or polypeptides obtained from other purification procedures; and (4) the characterization of the physicochemical properties of peptides or proteins. They all require a good understanding of the underlying common principles of polypeptide-ligand interaction. The basics of these principles are touched upon with references to further reading. Finally, a short section of this unit is dedicated to troubleshooting; however, many of the “check-back” confirmatory procedures implicit to sound operational practices and the identification of suitable alternatives for the separation strategy are included in the section on method development.

STANDARD OPERATING CONDITIONS FOR HP-SEC

The separation of peptides and proteins by high-performance size-exclusion chromatography (HP-SEC) is based on the concept that molecules of different sizes (hydrodynamic volume, Stoke's radius) permeate to different extents into porous SEC separation media and thus exhibit different permeation coefficients according to differences in their molecular weights (Regnier, 1983). However, many SEC materials are slightly hydrophobic or can weakly act as ion exchangers. These properties lead to nonideal behavior (specifically electrostatic or hydrophobic interactions between the peptide or protein and the matrix). This feature is not necessarily a disadvantage, since mixed-mode selectivities can be achieved, but can be suppressed by the addition of a salt at a reasonably high ionic strength, i.e., ≥ 100 mM, to the mobile phase (Mant et al., 1987).

Chromatographic Conditions

Column: e.g., TSK-250 (10 μ m, 300 Å, 300-mm length \times 7.5-mm i.d.)

Sample size: <2 mg peptide/protein

Sample loop size: 20 to 200 μ l

Isocratic elution

Eluent A: 50 mM KH_2PO_4 , pH 6.5, 0.1 M KCl

Flow rate: 0.5 ml/min

Detection: 214 nm

Temperature: room temperature

Peptide standards for column testing as described, e.g., in Mant and Hodges (1991a)

Method development in the HP-SEC of peptides and proteins can be performed via the following steps:

1. Select sorbent of the most appropriate average pore size, packed into a column of suitable length.
2. Check for “ideal” and “non-ideal” retention effects.
3. Optimize plate number (adjust the flow rate or change to a column of different length).

An example of HP-SEC for the separation of peptides and proteins is illustrated in Figure 10.13.1. In this example, the resolution of the 50S ribosomal proteins from *Thermus aquaticus* was achieved on a tandem TSK-250 column (75-mm length \times 7.5-mm i.d. and 300-mm length \times 7.5-mm i.d.) at a flow rate of 0.5 ml/min with a 50 mM $(\text{NH}_4)_2\text{SO}_4$ /20 mM NaH_2PO_4 buffer system, pH 5.0.

STANDARD OPERATING CONDITIONS FOR HP-NPC

Chromatographic systems in which the stationary phase is more polar than the mobile phase were developed at the beginning of the modern era of liquid chromatography and were known under the acronym NPC (“normal phase” liquid chromatography). In contrast to RPC with immobilized *n*-alkyl ligands, where the interaction of solute and stationary phase is based on solvophobic phenomena, the interaction in NP-LC is based on adsorption. The retention behavior of peptides and proteins in NP-LC is often described in terms of the classical concepts of multisite displacement and site occupancy theory (Snyder, 1970). Today, NP-HPLC is mainly used for the separation of polyaromatic hydrocarbons (PAHs), heteroaromatic compounds, nucleotides, nucleosides, etc., and much less frequently for protected

synthetic peptides, deprotected small peptides in the “flash chromatographic mode,” and protected amino acid derivatives used in peptide synthesis (Ballschmiter and Wossner, 1998). Originally, normal phase chromatography was limited to unmodified silica columns. Recent work has, however, utilized polar bonded phases such as amino (-NH₂), cyano (-CN), or diol (-COHCOH-) coated sorbents. Chromatography on such modified normal phase packing materials is also known as polar bonded phase chromatography (PBPC), which is used for the separation of peptides (Yoshida, 1998) and proteins (Buchholz et al., 1982). Today, one of the main applications of modified normal phase silica materials is in HPLC-integrated solid phase extraction procedures (SPE; Papadoyannis et al., 1995). These types of sorbents, particularly when used as precolumn packing materials in LC-LC column switching settings in conjunction with restricted access sorbents materials (RAM), allow multiple injections of untreated complex biological samples such as hemolyzed blood, plasma serum, fermentation broth, cell tissue homogenates, etc., for the isolation of bioactive peptides. Typically, with RAM materials, hydrophilic, electroneutral diol groups are immobilized onto the outer surface of spherical particles.

This layer prevents nonspecific interactions between the support matrix and protein(s) or other high-molecular-weight biomolecules, which are thus excluded from the interior regions of the particle and elute as nonretained components. The inner surfaces of the porous RAM particles are, however, chemically modified with *n*-alkyl ligands, which are only freely accessible for low molecular analytes, such as peptides. As a consequence, significant enrichment or partial resolution of peptide analytes can be achieved.

Chromatographic Conditions

Column: e.g., diol-phase, aminopropyl-phase, cyano-phase column, 250-mm length × 4.6-mm i.d.

Sample size: <2 mg peptide/protein

Sample loop size: 20 to 200 μl

Linear A→B gradient

Eluent A: 0.1% TFA in water

Eluent B: 0.1% TFA in 20% acetonitrile/80% water (v/v)

Gradient range and time: 0% to 100% eluent B in 60 min

Flow rate: 1 ml/min

Detection: 214 nm

Temperature: room temperature

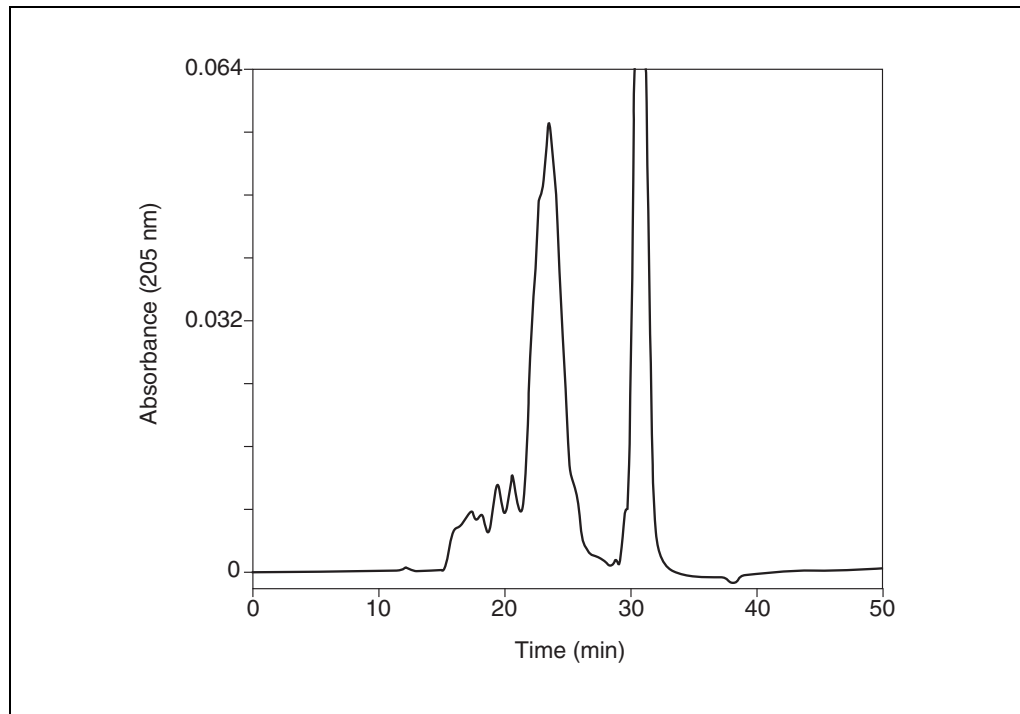


Figure 10.13.1 Illustrative example of the use of HP-SEC in peptide and protein separation and analysis. In this example the resolution of 50S ribosomal proteins from *Thermus aquaticus* in 2% acetic acid (Molnar et al., 1989b) was achieved with a tandem TSK-250 column (75-mm length × 7.5-mm i.d. and 300-mm length × 7.5-mm i.d.) using a flow rate of 0.5 ml/min and a 50 mM (NH₄)₂SO₄, 20 mM NaH₂PO₄, pH 5.0, with detection at 205 nm.

An example of the use of HP-NPC for the separation of peptides and proteins is illustrated in Figure 10.13.2. In this example, the resolution of ferritin, bovine serum albumin, chymotrypsinogen, cytochrome *c*, and alanine was achieved on a LiChrosorb Diol column (250-mm length \times 16-mm i.d.) at flow rate of 2.5 ml/min and a 0.1 M phosphate buffer system, pH 5.0 (Buchholz et al., 1982).

STANDARD OPERATING CONDITIONS FOR HP-HIC

As in reversed-phase chromatography, the hydrophobic interaction between the peptide or protein and the nonpolar ligands immobilized onto the surface of the sorbent represents the dominant effect in hydrophobic interaction chromatography (Fausnaugh et al., 1984; Fausnaugh and Regnier, 1986; Gooding et al., 1986; Wu et al., 1986a,b; Melander et al., 1989; Antia et al., 1995; Hearn, 2001a). In both techniques, peptides and proteins are eluted by lowering the surface tension of the mobile phase. However, in HP-HIC, this is achieved with a decreasing

salt concentration, i.e., by increasing the water content of the eluent, in contrast to RP-HPLC where the decrease in surface tension of the eluent is achieved through an increase in the organic solvent content of the mobile phase. As such, the minimum surface tension reached in the HP-HIC of polypeptides or proteins with binary water-salt systems corresponds to the surface tension of pure water, i.e., 78 dynes/cm. These differences between RP-HPLC and HP-HIC have fundamental effects on the recovery of proteins in bioactive state, as well as on the selectivity of the system. Typically, kosmotropic (anti-chaotropic) salts, i.e., ammonium sulfate, sodium sulfate, or magnesium chloride of high molal surface tension increment, are to be preferred for HP-HIC applications with polypeptides and proteins. In HP-HIC, non-polar ligands with lower hydrophobicity and lower ligand density ($\sim 1/10$ th that of RP-HPLC sorbents) are employed. HP-HIC sorbents should be selected on the basis of the critical hydrophobicity concept (Hearn, 2000, 2001a). In combination with nondenaturing mobile

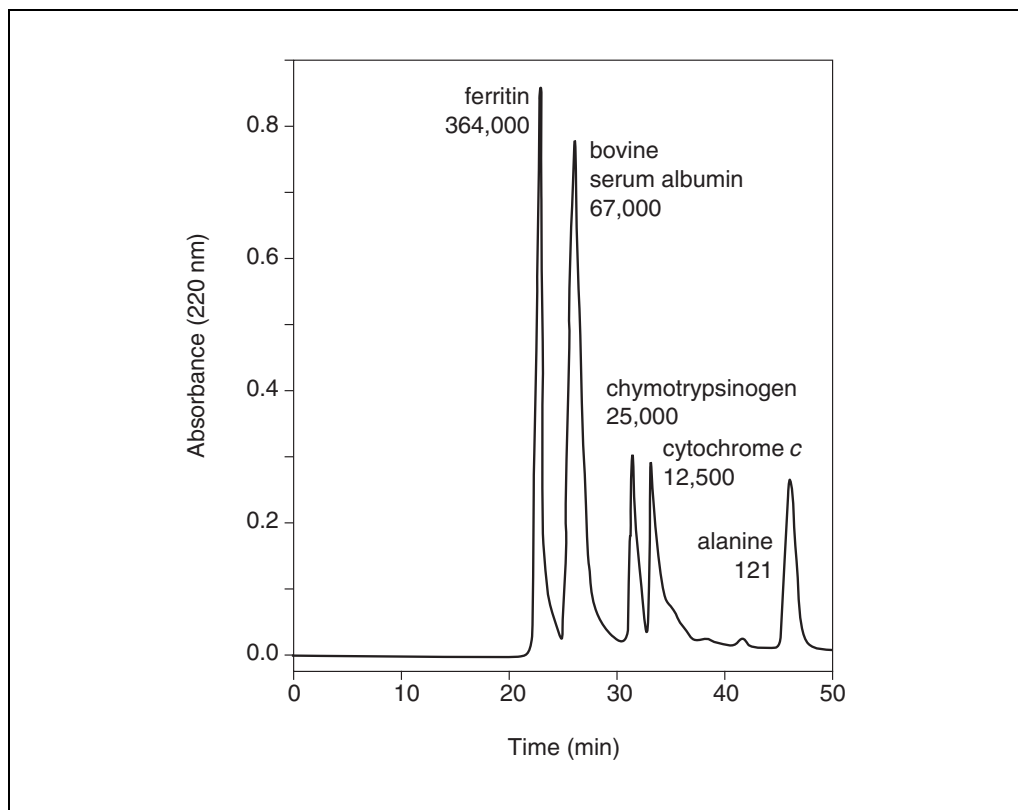


Figure 10.13.2 Illustrative example of the use of HP-NPC in peptide and protein separation and analysis. In this example the resolution of ferritin (M_r , 364,600), bovine serum albumin (M_r , 67,000), chymotrypsinogen (M_r , 25,000), cytochrome *c* (M_r , 12,500), and alanine (M_r , 121; Buchholz et al., 1982) was achieved on a LiChrospher Diol column (250-mm length \times 16-mm i.d.) at a flow rate of 2.5 ml/min and a 0.1 M phosphate buffer system, pH 5.0. In this specific case, size exclusion effects dominated the global resolution, with polar phase effects contributing to the individual selectivities of these proteins, as assessed from the “nonideality” of the log M_r versus V_e plots.

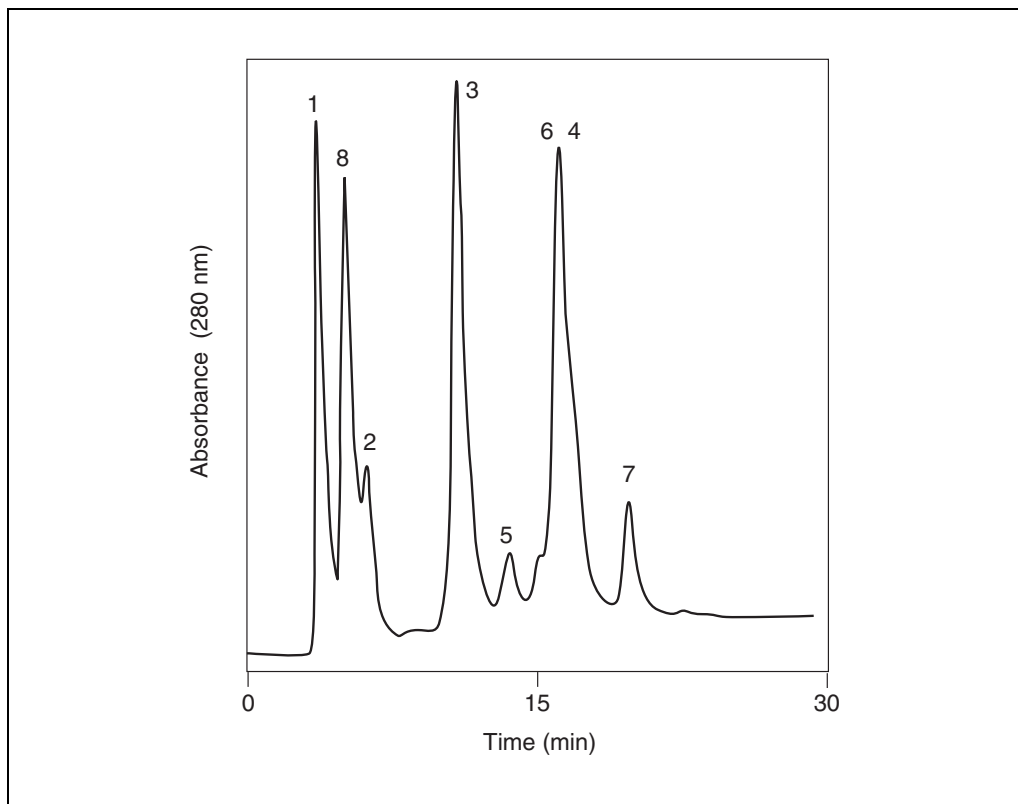


Figure 10.13.3 Illustrative example of the use of HP-HIC in peptide and protein separation and analysis. In this example the resolution of: 1, cytochrome *c*; 2, ribonuclease; 3, lysozyme; 4, bovine serum albumin; 5, ovalbumin; 6, α -chymotrypsin; 7, α -chymotrypsinogen; and 8, myoglobin was achieved (Kato et al., 1983) on a Butyl-G3000 SW column (150-mm length \times 6-mm i.d., 10 μ m) with a 60-min linear gradient from 0% to 100% eluent B (eluent A: 1.5 M ammonium sulfate, 0.1 M phosphate buffer, pH 6.0; eluent B: 0.1 M phosphate buffer, pH 6.0).

phases, proteins, in particular, can potentially be eluted in their native conformation from HP-HIC sorbents.

Chromatographic Conditions

Column: e.g., TSK Phenyl column, 75-mm length \times 7.5-mm i.d., 10 μ m

Sample size: <2 mg peptide/protein

Sample loop size: 20 to 200 μ l

Linear A \rightarrow B gradient is the preferred method with “hold” options

Eluent A: 0.1 M NaH₂PO₄, 2.0 M (NH₄)₂SO₄, pH 7.0

Eluent B: 0.1 M NaH₂PO₄, pH 7.0

Gradient rate: 5% eluent B/min

Gradient range and time: 0 to 100% B in 20 min

Flow rate: 1 ml/min

Detection: 214 nm

Temperature: room temperature

The method development in HP-HIC can be performed in the following steps:

1. Select type of salt (antichaotropic) and concentration range, taking into account the concentration of the salt that is required to reach

saturation, and then keep the maximum concentration below this value by \sim 20%.

2. Optimize the gradient conditions (gradient run time, starting and final mobile phase composition).

3. Optimize the band spacing (pH, organic solvent).

4. Optimize the column conditions (flow, column length).

Examples of the use of HP-HIC with proteins are described in Fausnaugh et al. (1984); Fausnaugh and Regnier (1986); Gooding et al. (1986); Melander et al. (1989); Wu et al. (1986a,b); Antia et al. (1995); and Hearn (2000, 2001a). Figure 10.13.3 illustrates the resolution of cytochrome *c* (1), ribonuclease (2), lysozyme (3), bovine serum albumin (4), ovalbumin (5), α -chymotrypsin (7), and myoglobin (8) on a Butyl-G3000 SW column (150-mm length \times 6-mm i.d., 10 μ m) with a 60-min linear gradient from 0% to 100% B using eluent A (1.5 M ammonium sulfate, 0.1 M phosphate buffer, pH 6.0) and eluent B (0.1 M phosphate buffer, pH 6.0).

STANDARD OPERATING CONDITIONS FOR HP-IEX

Peptides and proteins can be eluted in ion exchange chromatography by either isocratic or gradient elution (Chang et al., 1976; Kopaciewicz and Regnier, 1983a, 1986; Regnier, 1984; Kopaciewicz et al., 1985; Hearn et al., 1988; Heinitz et al., 1988; Hodder et al., 1990). Gradient elution is usually performed with a linear A→B gradient of a salt such as sodium or potassium chloride in phosphate buffer. The retention of peptides and proteins on ion-exchange sorbents arises from electrostatic interactions between the ionized surface of the solute and the charged surface of the sorbent. For peptide and protein separations, the use of a strong cation exchanger has a considerable advantage over other varieties of ion-exchanger (Mant and Hodges, 1985), since the column can retain its negatively charged character over the whole range from acidic to neutral pH. Both weak and strong cation exchangers, e.g., based on carboxymethyl or sulfonopropyl ligands, as well as weak and strong anion exchangers, e.g., dimethylamino or quaternary ammonium ligands, are available and can be applied to the HP-IEX of peptides and proteins.

With peptides and proteins, at neutral pH, the side-chain carboxyl groups of the acidic amino acid residues—glutamic acid and aspartic acid—are completely ionized. Below pH 3.0 they are almost completely protonated. A change of pH therefore allows the retention of peptides and proteins to be varied according to the net charge of these biosolute(s). Even when the most dominant effect in IEX is electrostatic, in many cases the participation of a mixed mode separation, whereby hydrophobic interaction contributes, cannot be excluded (Burke et al., 1989). This is not necessarily a disadvantage, since it permits selectivity modulation with complex mixtures of peptides and proteins. Hydrophobic interactions can be suppressed by adding a nonpolar solvent, such as acetonitrile, or a nonionic detergent, such as Brij-25, to the mobile phase.

Chromatographic Conditions

Column: e.g., strong cation exchanger with sulfonate functionality (5 μm , 300 \AA , 75-mm length \times 7.5-mm i.d.)

Sample size: <2 mg peptide/protein

Sample loop size: 20 to 200 μl

Linear A→B gradient

Eluent A: 5 mM KH_2PO_4 , pH between 3.0 and 7.0

Eluent B: 5 mM KH_2PO_4 /0.5 M KCl, pH between 3.0 and 7.0

Gradient rate: 3.3% eluent B/min

Gradient range and time: 0% to 100% eluent B in 30 min

Flow rate: 1 ml/min

Detection: 214 nm

Temperature: room temperature

Peptide standards for column testing as described, e.g., in Mant and Hodges (1991a)

The method development in HP-IEX can be performed in the following steps:

1. Select type of anion or cation exchanger.
2. Optimize gradient conditions (gradient run time, starting and final mobile phase composition).
3. Optimize band spacing (pH, salt type).
4. Optimize column conditions (flow, column length).

Examples of the use of HP-IEX with proteins are described in Chang et al. (1976); Regnier (1984); Kopaciewicz et al. (1985); Kopaciewicz and Regnier (1986); Hearn et al. (1988); Heinitz et al. (1988); Burke et al. (1989); Hodder et al. (1990); Mant and Hodges (1985, 1991b); and Hearn (2000). Figure 10.13.4 illustrates the use of HP-IEC for the resolution under gradient elution conditions from 0% to 100% B (eluent A: 0.1 M Tris-Cl, pH 7.5; Eluent B: 0.1 M Tris-Cl/0.2 M NaCl, pH 7.5) of ovalbumin isoforms on a TSK-GEL IEX-545 DEAE SIL column (150-mm length \times 6-mm i.d.) using a 90-min linear gradient at a flow rate of 1.0 ml/min.

STANDARD OPERATING CONDITIONS FOR HP-HILIC

Peptides can be separated on strong hydrophilic materials, i.e., poly(2-hydroxyethyl-aspartamide) silica (Alpert, 1990), whereby their selectivity changes with the concentration of the organic modifier. At high organic modifier concentrations (i.e., >55% *n*-propanol), the solute is retained on the column. The solute is subsequently eluted with a decreasing gradient of organic modifier, whereby with low organic modifier concentrations the separation of the solutes is governed by molecular sieving effects. HP-HILIC allows specifically the evaluation of highly polar compounds, which cannot be retained on traditional reversed-phase stationary phases. To date a variety of HP-HILIC sorbent packings, including amide-, poly(2-hydroxyethyl)-aspartamide-, cyclodextrin-, or teicoplanin-derivatized stationary phases are available (Strege, 1998; Risley and Strege, 2000). HP-HILIC can be applied in a variety of

challenging separation tasks, e.g., the separation of glycopeptides (Zhang and Wang, 1998) or the desalting of electroeluted proteins from SDS-PAGE systems (Jeno et al., 1993). In this mixed-mode hydrophilic interaction/cation exchange chromatographic mode (Mant et al., 1998a,b; Litowski et al., 1999), peptides are usually subjected to linear, decreasing salt gradients in the presence of high levels of organic modifier. The following example is representative of the conditions that can be employed in the HP-HILIC of peptides and small proteins.

Chromatographic Conditions

Column: e.g., HILIC PolySulfoethyl A sorbent with a poly(2-sulfonethylaspartimide) functionality (5 μm , 300 \AA , 200-mm length \times 4.6-mm i.d.)

Sample size: <2 mg peptide/protein

Sample loop size: 20 to 200 μl

Linear A \rightarrow B gradient

Eluent A: 20 mM triethylammonium phosphate/80% acetonitrile, pH 3.0

Eluent B: eluent A with 400 mM NaClO_4 , pH 3.0.

Gradient rate: 2.5% B/min

Gradient range and time: 0% to 100% B in 90 min

Flow rate: 1 ml/min

Detection: 214 nm

Temperature: 30°C

Method development in the HP-HILIC of peptides and proteins can be performed via the following steps:

1. Select sorbent of the most appropriate average pore size, packed into a column of suitable length.

2. Check for “ideal” and “non-ideal” retention effects.

3. Optimize selectivity through the changes in the composition or concentration of the organic solvent and kosmotropic/chaotropic salt additives.

4. Optimize plate number (by adjusting the flow rate, or changing to a column of different length).

Examples of the use of HP-HILIC with peptide standards for column testing are described in Mant and Hodges (1991b). In Figure 10.13.5 is shown the separation of the tryptic glycopeptides of Asn^{97} for recombinant human interferon γ (Zhang and Wang, 1998) on a polyhydroxyethylaspartamide column (150-mm length \times 1-mm i.d.) using a flow rate of 50 $\mu\text{l}/\text{min}$. In this example, a 60-min gradient from 100% A (85% acetonitrile, 15% water, 10 mM triethylamine) to 100% B (20% acetonitrile, 15% water, 10 mM triethylamine, 25 mM sodium perchlorate), pH 6.0 was employed.

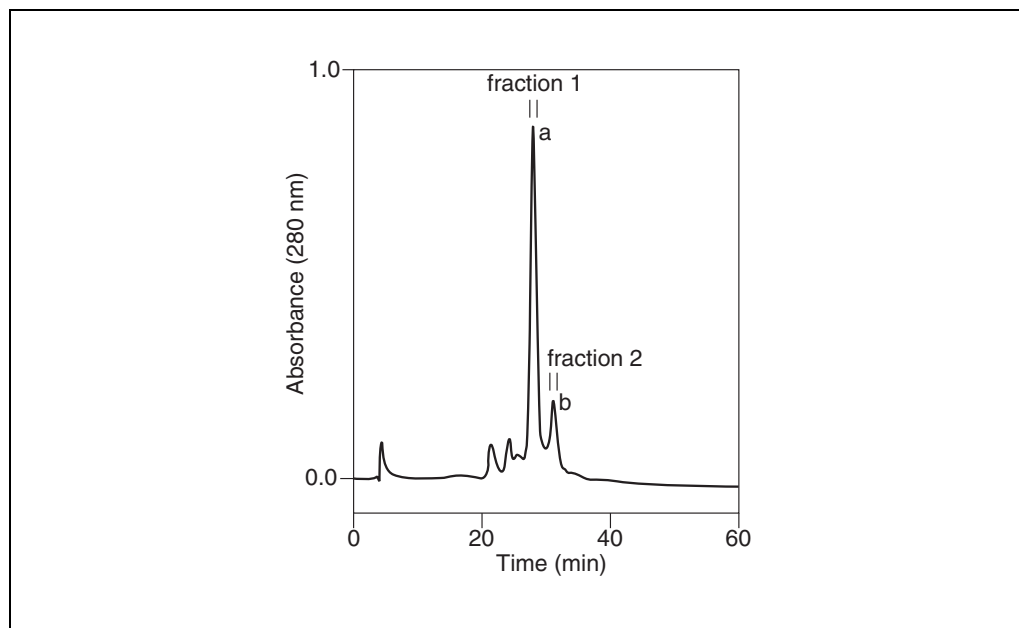


Figure 10.13.4 Illustrative example of the use of HP-IEX in peptide and protein separation and analysis. In this example the resolution of ovalbumin isoforms (Kato et al., 1982), Frla and FrIb, was achieved on a TSK-GEL IEX-545 DEAE SIL column (150-mm length \times 6-mm i.d.) using a 90 min linear gradient at a flow rate of 1.0 ml/min from 0-100% B (eluent A: 0.1 M Tris-Cl, pH 7.5; eluent B: to 0.1 M Tris-Cl, 0.2 M NaCl, pH 7.5).

STANDARD OPERATING CONDITIONS FOR HP-IMAC

Immobilized metal-chelate affinity chromatography (IMAC) exploits the affinities of the side chain moieties of specific surface accessible amino acids in peptides and proteins for the coordination sites of immobilized transition metal ions (Porath et al., 1975; Zachariou et al., 1993). The majority of investigations employed tri- or tetradentate ligands, such as iminodiacetic acid (IDA), nitrilotriacetic acid (NTA), tris-(carboxymethyl)ethylenediamine (TED), *O*-phosphoserine (OPS) or carboxymethylaspartic acid (CMA) (Zachariou et al., 1993). Novel immobilized chelate systems, such as 2,6-diaminomethylpyridine or *cis* and *trans* carboxymethylproline, have shown binding properties different from the IMAC behavior of other chelating ligands (Hearn et al., 1997; Chaouk and Hearn, 1999). These techniques, in conjunction with soft gel matrices, have been predominantly applied for the preparative purification of globular proteins. Procedures to immobilize the IMAC ligand at the surface of silica supports, based on classical ligand exchange principles, have permitted useful guidelines to be developed for the design

of HP-IMAC applications for peptides and proteins at the analytical level (Wirth and Hearn, 1993). However, for the majority of the possible HP-IMAC applications with peptides or proteins no standard chromatographic conditions exist. Rather, different elution regimes have tended to be employed with HP-IMAC systems, based on the selection of either empirical step- or gradient-elution protocols involving changes in pH, buffer composition, or the concentration of salts or another competitive binding reagent such as imidazole or malonic acid. Various applications of different IMAC systems for the separation of peptides and proteins have recently been reviewed elsewhere (Porath et al., 1975; Jenó et al., 1993; Mant et al., 1998a,b; Litowski et al., 1999; Hearn, 2000). The following example is representative of the conditions that can be employed in the HP-IMAC of peptides and small globular proteins.

Chromatographic Conditions

Column: e.g., HP-IMAC sorbent with a Cu^{2+} ion chelated to immobilized iminodiacetic acid functionality, i.e., IDA-Cu(II) TSK gel chelate-5PW column (5 μm , 300 \AA , 100-mm length \times 4.6-mm i.d. or 75-mm length \times 8-mm i.d.)

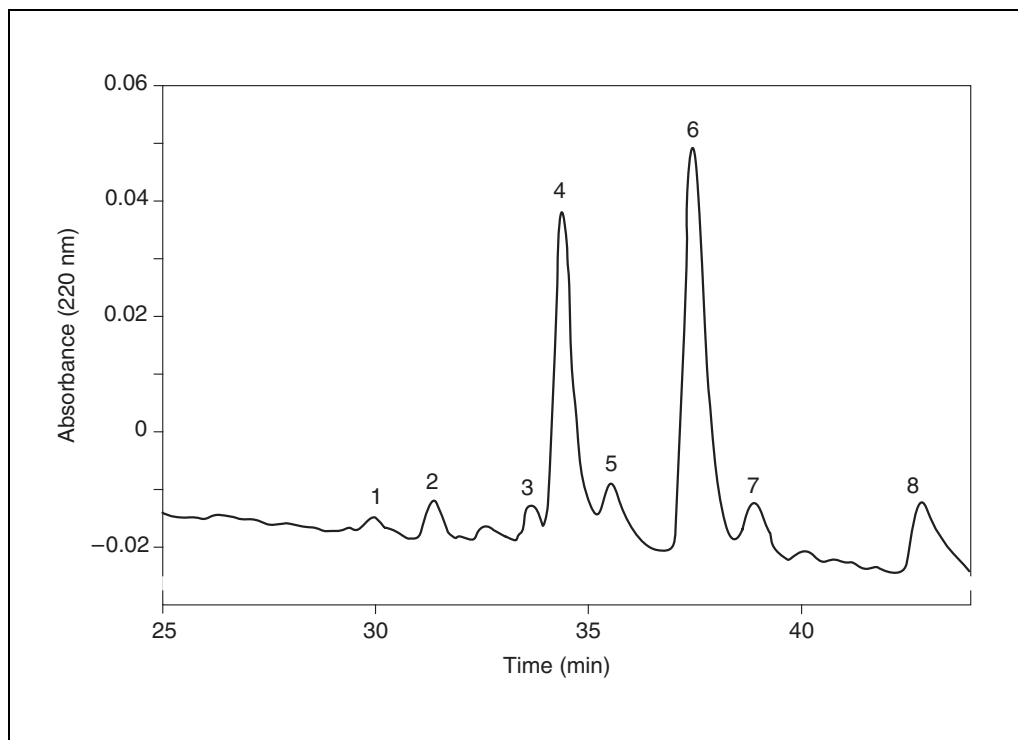


Figure 10.13.5 Illustrative example of the use of HP-HILIC in peptide and protein separation and analysis. In this example the resolution of the tryptic Asn⁹⁷ glycopeptides (1-8) of recombinant human interferon γ (Zhang and Wang, 1998) was achieved on a polyhydroxyethyl aspartamide column (150-mm length \times 1-mm i.d.) using a flow rate of 50 $\mu\text{l}/\text{min}$ and a 60-min gradient from 100% A (85% acetonitrile, 15% water, 10 mM triethylamine) to 100% B (20% acetonitrile, 15% water, 10 mM triethylamine, 25 mM sodium perchlorate), pH 6.

Sample size: <1 mg peptide/protein
Sample loop size: 20 to 200 μ l
Linear A \rightarrow B gradient
Eluent A: 50 mM acetic acid/50 mM MES/50 mM HEPES/80 mM Na₂SO₄/2 \times 10⁻⁶ M Cu²⁺, pH 8.0; or 1 mM imidazole/20 mM sodium phosphate buffer containing 0.5 M NaCl, pH 7.0
Eluent B: 50 mM acetic acid/50 mM MES/50 mM HEPES/50 mM ammonium acetate/2 \times 10⁻⁶ M Cu²⁺, pH 5.5; or 20 mM sodium phosphate buffer containing 0.5 M NaCl, pH 7.0 with 20 mM imidazole gradient
Gradient rate: 5% B/min
Gradient range and time: 0% to 100% eluent B in 20 min
Flow rate: 1 ml/min
Detection: 280 nm
Temperature: 25°C

Method development in the HP-IMAC of peptides and proteins can be performed via the following steps:

1. Select sorbent of the most appropriate average pore-size, packed into a column of suitable length.
2. Check for "ideal" and "non-ideal" retention effects.
3. Optimize selectivity through the changes in the type of chelating ligand and metal ion, i.e., whether borderline, hard or soft metal ion, the pH, the type and concentration of the buffer used at the loading and washing stages, and finally at the elution stage.
4. Optimize plate number (by adjusting the flow rate or changing to a column of different length).

Examples of the use of HP-IMAC with proteins are described in Porath et al. (1975); Wirth and Hearn (1993); Zachariou et al. (1993); Jiang et al. (1998); Porath (1988); and Hearn (2000). Figure 10.13.6 illustrates the use of HP-IMAC for the resolution of human gastrin (24), human GIP (29), Trp(for)- human GIP (30), human big gastrin (35), porcine GIP (38), and bovine GIP on an IDA-Cu(II) TSK gel chelate-5PW column (75-mm length \times 8-mm i.d.) at a flow rate of 1 ml/min equilibrated with a 1 mM imidazole/20 mM sodium phosphate buffer containing 0.5 M NaCl, pH 7.0 and eluted with a 1 to 20 mM imidazole gradient (dotted line).

STANDARD OPERATING CONDITIONS FOR HP-BAC

During the last 25 years, biospecific affinity chromatography (BAC) had evolved as a well established method for analytical separation

(Porath, 1988), as well as the preparative scale purification of biopolymers of complex biological origin. Additionally, HP-BAC is a useful technique to investigate peptide-protein and protein-protein interactions. In HP-BAC, the separation is based on the propensity for biorecognition between the protein or peptide and its naturally occurring ligand (or suitable mimetic), in particular, on its relative binding affinity for a specific ligand. The ligand should have a high specificity for the protein or peptide, should bind to the protein or peptide reversibly with rapid on-and-off kinetics, and should be chemically stable under elution conditions. The introduction of HP-BAC dramatically increased the speed of affinity chromatographic separation of proteins, including sample loading and column regeneration (Ohlsson et al., 1978). However, since the specificity and the selectivity of the HP-BAC both are very high, the range of applications is, paradoxically, much more limited than the other modes of HPLC. Every separation ideally demands its own ligand type, ligand density, column configuration, and particle type in order to enable the optimal purification of a particular component. Consequently, there are no "standard" HP-BAC conditions due to the broad range of ligand types. Some ligands are specific for a particular class of protein; however, other ligands can be multifunctional and bind to a number of related molecules. Currently, commercially available HP-BAC columns are supplied as activated supports (epoxy-, N-hydroxysuccinimido-, tresyl-, hydroxylpropyl-) ready for coupling appropriate ligands (via accessible SH, OH, or NH₂ groups) or alternatively as preformed matrices with commonly used ligands already attached (e.g., protein A). The following example is representative of the conditions that can be employed in the HP-BAC with peptides and small globular proteins.

Chromatographic Conditions

Column: custom designed, i.e., glycidopropyl-activated silica such as epoxy-activated LiChrosorb containing a suitable immobilized ligand functionality (10 μ m, 10000 Å , 100-mm length \times 4.6-mm i.d.)

Sample size: <2 mg peptide/protein

Sample loop size: 20 to 200 μ l

Linear gradient or step elution

Eluent for equilibration: nondenaturing conditions with adequate buffer capacity

Eluent for desorption: non-denaturing conditions of different pH to the equilibration conditions, with adequate buffer capacity and ap-

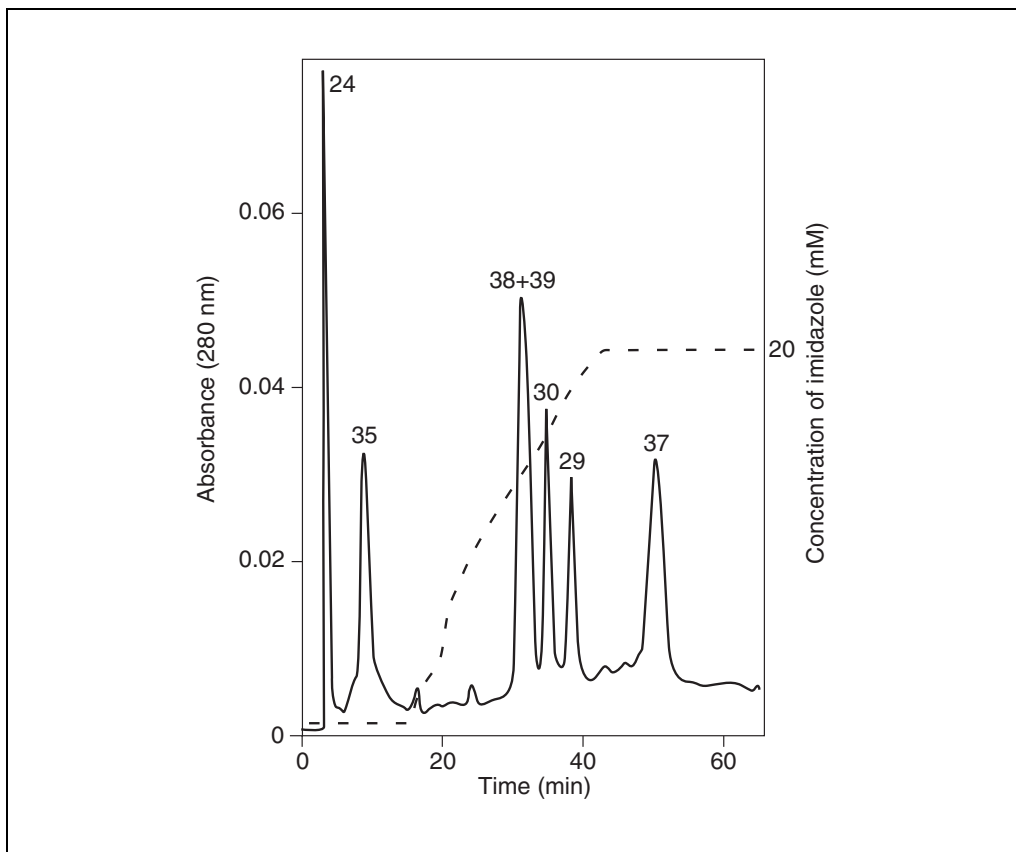


Figure 10.13.6 Illustrative example of the use of HP-IMAC in peptide and protein separation and analysis. In this example the resolution of human gastrin-I (24), human GIP(21-42) (29), Trp(for)-human GIP(21-42) (30), human big gastrin (35), human GIP (37), porcine GIP (38), and bovine GIP(39) (Yip et al., 1989) was achieved on an IDA-Cu(II) TSK gel chelate-5PW column (75-mm length \times 8-mm i.d.) with a flow rate of 1 ml/min equilibrated with a 1 mM imidazole, 20 mM sodium phosphate buffer containing 0.5 M NaCl, pH 7.0, and eluted with an 1-20 mM imidazole gradient (dotted line).

appropriate competitive species to affect efficient dissociation of the affinant-affinate cognate interaction

Gradient rate: according to affinity of interaction, HP-BAC sorbent characteristics, and flow rate

Flow rate: 0.5 to 1 ml/min

Detection: 214 to 280 nm

Temperature: 4°C

In Figure 10.13.7 an example is shown of the use of HP-BAC for the resolution of rabbit muscle lactate dehydrogenase (LDH) on a Procion Red MX5B dye immobilized onto nonporous silica column (40-mm length \times 6-mm i.d.) using a flow rate of 1 ml/min and step elution with eluent A: 10 mM phosphate buffer, pH 8.0, and eluent B: 0.5 M sodium chloride in 10 mM phosphate buffer, pH 8.0.

STANDARD OPERATING CONDITIONS FOR RP-HPLC

As noted above, RP-HPLC procedures currently represent the majority of applications for peptide analysis and purification, and over 80% of all analytical studies with proteins. The dominant effect in reversed-phase chromatography is a hydrophobic interaction between the nonpolar amino acid residues of peptides or proteins and the nonpolar ligands, typically immobilized onto the surface of a spherical, porous silica particle (Hearn, 2001b), although nonpolar polymeric sorbents derived, e.g., from cross-linked polystyrene-divinylbenzene, can also be employed. In this technique, isocratic elution, step elution, or gradient elution modes can be utilized to purify peptides and proteins. Besides the requirement for an organic solvent to be used as a surface tension modifier, ion-pair reagents (Mant and Hodges, 1991c) are utilized at low pH (e.g., pH 2.1) to

suppress silanophilic interactions between free silanol groups on the silica surface and basic amino acid residues. Silica-based packing materials of 3 to 10 μm average particle diameter and ≥ 300 Å pore size, with *n*-butyl, *n*-octyl, or *n*-octadecyl ligands, are widely used.

Chromatographic Conditions

Chemicals

HPLC-grade acetonitrile, 2-propanol, methanol, or other suitable organic solvents with UV transparency down to 210 nm

Trifluoroacetic acid (TFA)

NaH_2PO_4 or other suitable salts

H_3PO_4

Milli-Q water or equivalent

TFA employed for protein sequencing may not be suitable because it can contain antioxidants.

Preparing the mobile phase

1. Prepare eluent A and B, e.g.: eluent A: 0.1% TFA in water and eluent B: 0.09% TFA in 60% acetonitrile/40% water (v/v).

The volumes of organic solvent and water are measured in two different cylinders and then combined, because of the volume contraction upon mixing, which may be up to 30 ml per liter of prepared solvent (depending on the nature of the solvent). Failure to do so can lead to substantial errors in mobile phase composition.

To compensate for the baseline shift in gradient elution (because organic components absorb more light at low wavelengths) when working with water/organic eluents, the amount of ion pair reagent (TFA, H_3PO_4) in eluent B is usually decreased by 10% to 15% in comparison with eluent A, yielding a flat baseline (Dolan and Snyder, 1989; Dolan, 1991).

It is imperative that eluents be prepared in the fume hood. TFA is extremely corrosive; laboratory coat, gloves, and protective glasses must be worn.

2. Mix the solvent either by stirring with a Teflon-coated magnetic flea or by shaking a stoppered cylinder.

Parafilm should not be used, under any circumstances, to cover the eluents, since the or-

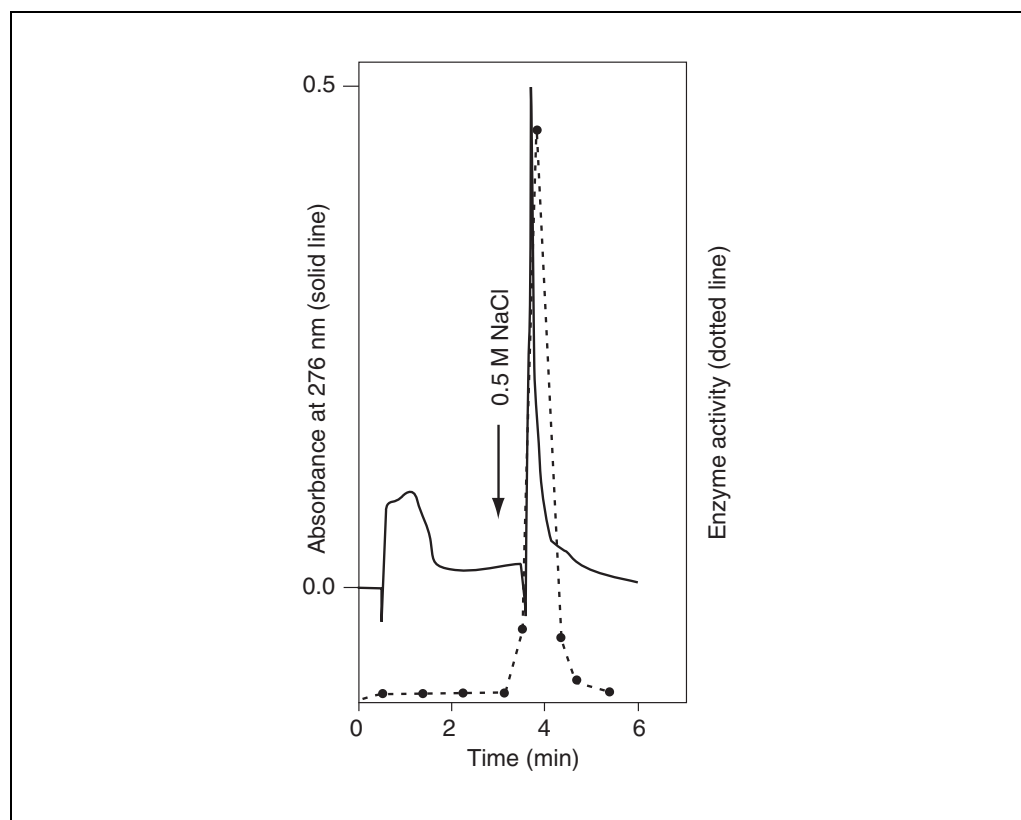


Figure 10.13.7 Illustrative example of the use of HP-BAC in peptide and protein separation and analysis. In this example, the enzyme activity is shown for rabbit muscle lactate dehydrogenase (LDH) (Anspach et al., 1988) separated on a Procion Red MX5B immobilized non-porous silica column (40-mm length \times 6-mm i.d.) using a flow rate of 1 ml/min and step elution (eluent A: 10 mM phosphate buffer, pH 8.0; eluent B: 0.5 M sodium chloride in 10 mM phosphate buffer, pH 8.0).

ganic solvents in combination with acidic ion-pair reagent components in the eluents will dissolve components in the Parafilm, yielding extra peaks in the chromatogram.

Testing the RP-HPLC stationary phase

The following considerations are relevant to the evaluation of HPLC stationary phases for use with samples containing peptides and proteins.

1. Test the column performance with a gradient run and a test mixture: RP-HPLC test mixture example: 0.15% (w/v) dimethylphthalate, 0.15% (w/v) diethylphthalate, 0.01% (w/v) diphenyl, 0.03% (w/v) *O*-terphenyl, 0.32% (w/v) dioctylphthalate in methanol.

2. Peptide standards for column testing are described, for example, in Mant and Hodges (1991b).

Peptide purification by RP-HPLC

In addition to the reagents and procedures described in the preceding section the following additional materials are required:

Chemicals: acetonitrile, trifluoroacetic acid (TFA)

Columns: analytical C-18 or other appropriate *n*-alkylsilica sorbent material (5 μ m, 300 \AA , 150-mm length \times 4.6-mm i.d.)

Preparative C-18 or other appropriate *n*-alkylsilica sorbent material (10 μ m, 300 \AA , 300-mm length \times 21.5-mm i.d.)

Apparatus: in addition to an analytical HPLC instrument with autosampler: preparative HPLC pump(s), manual injector, detector with preparative flow cell, and fraction collector are required

Purification of peptides derived from Solid-Phase Peptide Synthesis (SPPS) can be performed in the following steps.

1. Separation of the crude peptide (~100 μ g) with an analytical RP-HPLC procedure allows the assessment of the sample in terms of purity, peak profile and elution conditions. Appropriate equipment/conditions are as follows: Column: e.g., C-4, C-8, C-18, etc. (5 μ m, 300 \AA , 150-mm length \times 4.6-mm i.d.)

Sample size: <2 mg peptide/protein

Sample loop size: 20 to 200 μ l

Linear A \rightarrow B gradient

Eluent A: 0.9% aqueous TFA

Eluent B: 0.1% TFA in acetonitrile/water

Gradient rate: 1% B/min

For example, for 60% ACN/water: gradient range and time: 0% to 100% eluent B in 60 min

Flow rate: 1.0 ml/min

Detection: 214 nm

Temperature: room temperature

An example of the analytical use of RP-HPLC in peptide analysis is shown in Figure 10.13.8. In this case, resolution of the crude *N*-acetyl lipocortin-1[2-26] product from solid phase peptide synthesis (panel A), and the purified product (panel B) was achieved with a TSK-ODS-120 T column (150-mm length \times 4.6-mm i.d., 120 \AA , 5 μ m, end-capped) at a flow rate of 1 ml/min and a 60 min gradient from 100% A (0.1% TFA in water) to 100% B (90% acetonitrile 0.09% TFA), pH 2.1.

2. Separation of the crude peptide (~25 to 100 mg) with preparative RP-HPLC procedure. Appropriate equipment/conditions are as follows:

Column: e.g., C-4, C-8, C-18, etc. (10 μ m, 300 \AA , 300-mm length \times 21.5-mm i.d.)

Sample size: <150 mg peptide/protein

Sample loop size: 1 ml, multiple injection

Linear A \rightarrow B gradient

Eluent A: 0.9% aqueous TFA

Eluent B: 0.1% TFA in acetonitrile/water

Gradient rate: 0.66% eluent B/min

For example for 60% ACN/water: gradient range and time: 0% to 100% eluent B in 90 min

Flow rate: 7.5 ml/min

Detection: 254 nm

Temperature: room temperature

The preparative use of RP-HPLC in peptide purification is shown in Figure 10.13.9. In this example, the isolation of the synthetic *N*-acetyl lipocortin-1[2-26] peptide product obtained from solid-phase peptide synthesis was achieved on a TSK-ODS-120 T column (300-mm length \times 21.5-mm i.d., 120 \AA , 10 μ m, end-capped) using a flow rate of 7.5 ml/min and a 135 min gradient from 25% B to 100% B (A: 0.1% TFA in water, B: 60% acetonitrile 0.09% TFA), pH 2.1.

In order to avoid detector response overload effects, usually wavelengths from 230 to 280 nm are chosen. However, small amounts of chemical scavengers (used during the SPPS procedures) present in the crude peptide solution can absorb very strongly in this wavelength range. It may be worthwhile to sacrifice up to 1 mg of sample and to perform a preparative separation with detection at 214 nm in order to unambiguously determine the retention time of the main peptide product.

3. Collection of HPLC-fractions (3 to 7.5 ml).

4. Analysis of aliquots (30 to 50 μ l) of the collected fraction with an analytical RP-HPLC. A blank, the crude peptide solution, the fraction

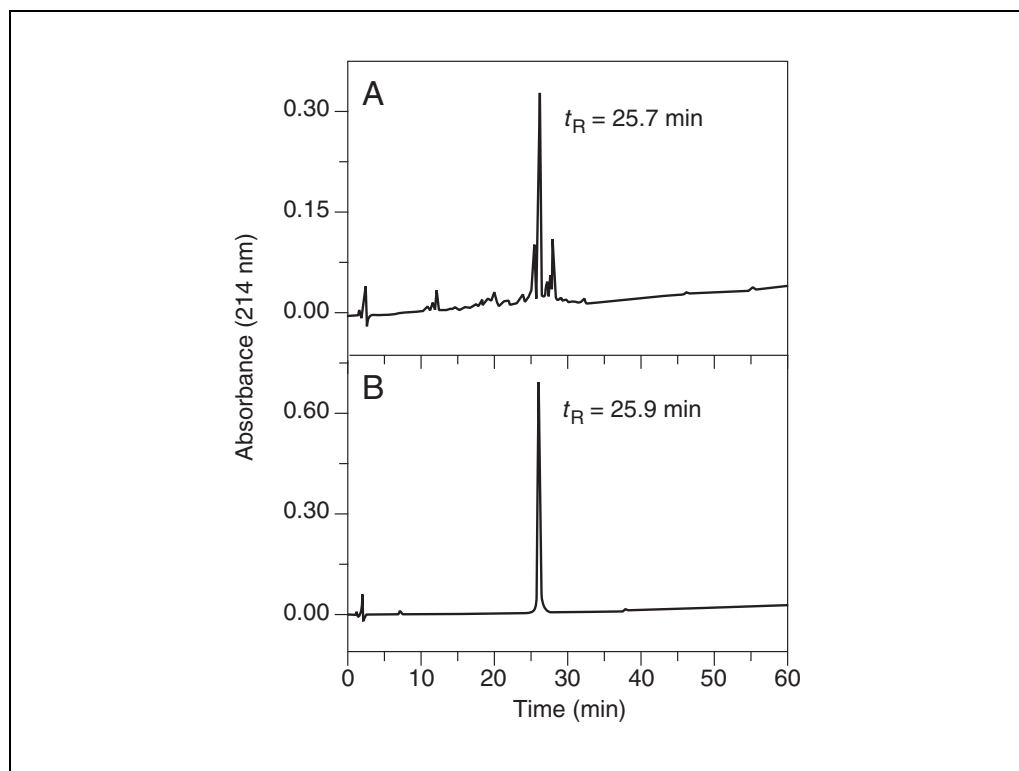


Figure 10.13.8 Illustrative example of the use of RP-HPLC in peptide separation and analysis. Resolution of *N*-acetyl lipocortin-1 [2-26] peptide product [Ac-AMVSEFLKQAW-FIENEEQEYVQTVK-CONH₂] obtained by solid phase peptide synthesis on a TSK-ODS-120 T column (150-mm length \times 4.6-mm i.d., 120 Å, 5 μ m, end-capped) using a flow rate of 1 ml/min and a 60 min gradient from 100% A (0.1% TFA in water) to 100% B (90% acetonitrile 0.09% TFA), pH 2.1. (A) Crude peptide; (B) purified peptide.

of interest and the 2 fractions before and 2 fractions after are typically analyzed.

5. Freeze drying of selected fraction.
6. Carry out off-line or on-line the ES-MS of purest fraction.

DESALTING OF PEPTIDE AND PROTEIN MIXTURES BY RP-HPLC TECHNIQUES

RP-HPLC can be utilized to desalt peptide or protein samples derived from extraction procedures or from previous HP-HIC, HP-IEC, HP-IMAC, HP-HILIC, or HP-BAC separations. Peptide or protein solutions are injected onto a small RP-HPLC column. An aqueous buffer is used to elute the salts, while the peptides or proteins are concentrated on the top of the column. After elution of the salts, monitored by UV detection, the peptides or proteins are eluted with water-acetonitrile or water 2-propanol mobile phases. The loading capacity of an analytical column (100- to 300-mm length \times 4-mm i.d.) is typically \sim 8 mg, while the

loading capacity for a semi-preparative column (30-mm length \times 16-mm i.d.) is \sim 34 mg (Pohl and Kamp, 1987).

In addition to the reagents and procedures described in the preceding sections the following materials, reagents and conditions are required:

Chromatographic Conditions

Chemicals: acetonitrile, 2-propanol, trifluoroacetic acid (TFA)

Column: e.g., C-4, C-8, C-18, etc. (10 μ m, 300 Å, 300-mm length \times 21.5-mm i.d.)

Sample size: 8 mg peptide or protein sample

Sample loop size: 1 ml

Step elution

Eluent A: 0.1% aqueous TFA

Eluent B: 0.1% TFA in acetonitrile or 2-propanol

Elution conditions: 100% eluent A for 3 min, then 100% eluent B for 3 min

Flow rate: 2.5 ml/min

Detection: 230 nm

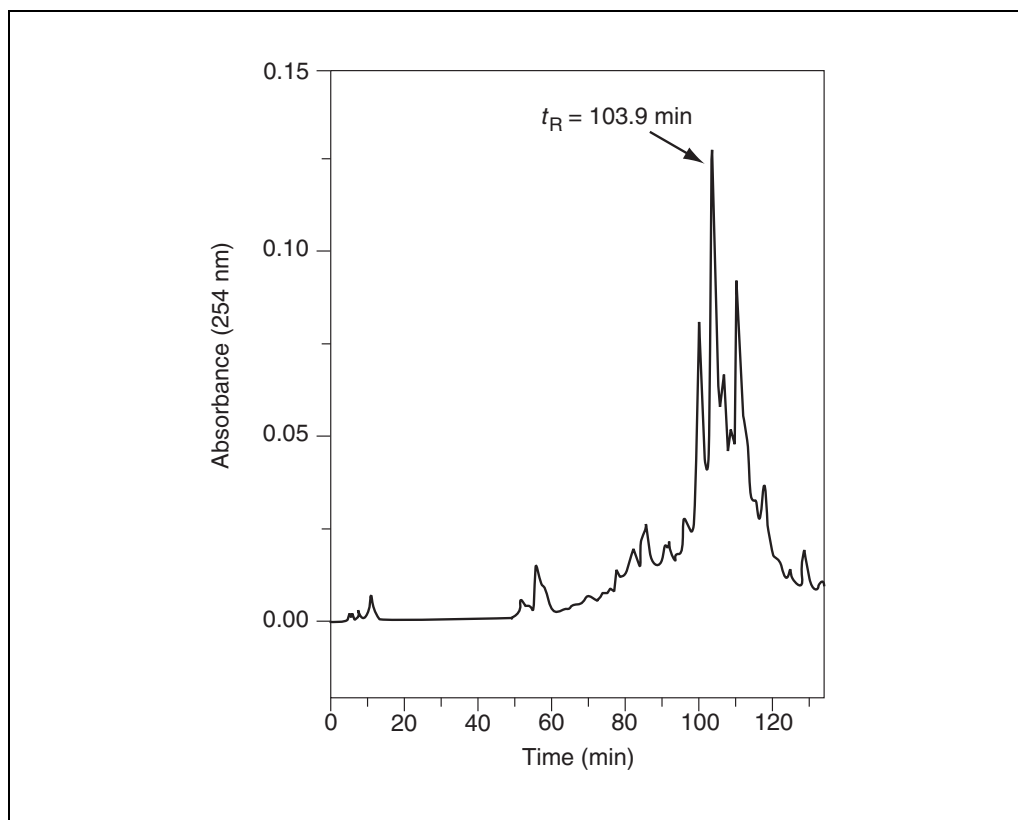


Figure 10.13.9 Illustrative example of the use of RP-HPLC in the preparative isolation of synthetic polypeptides. Here, the isolation of the peptide, *N*-acetyl lipocortin-1[2-26]—Ac-AMVSEFLKQAW-FIENEEQEYVQTVK-CONH₂—(70 mg) from the crude mixture obtained from solid phase peptide synthesis using a TSK-ODS-120 T column (300-mm length × 21.5-mm i.d., 120 Å, 10 μm, end-capped) using a flow rate of 7.5 ml/min and a 135-min gradient from 25% B to 100% B (A: 0.1% TFA in water, B: 60% acetonitrile 0.09% TFA), pH 2.1.

Temperature: room temperature

Illustrative of the use of RP-HPLC in the desalting of peptide and protein samples isolated by other techniques is the example shown in Figure 10.13.10 for the step elution of a 50S ribosomal protein sample derived from *Thermus aquaticus* preparation on a C-4 column (40-mm length × 4-mm i.d., 300 Å, 5 μm) using a flow rate of 2.5 ml/min.

MULTIMODAL HPLC: COLUMN SWITCHING

Electronically controlled multimodal HPLC is a valuable technique (Kopaciewicz and Regnier, 1983b; Hulpke and Werthmann, 1986), which allows automated fractionation and sample cleanup. In addition to the reagents and procedures described in the preceding sections, the following materials, reagents, and conditions are required for such experiments: (1) an autosampler with two rheodyne 7010-080 valves, third pump (optional), and (2) columns (e.g., for HP-SEC and RP-HPLC).

Automated fractionation, for example, with a HP-SEC-RP-HPLC coupling, can be performed with the setup shown in Figure 10.13.11. The valves V₂ and V₃ allow the on-line or off-line position of the HP-SEC and RP-HPLC column to be independently controlled. To establish an electronically controlled automated fractionation method, the following steps are needed:

1. Establish a suitable method for HP-SEC-RP-HPLC, whereby the HP-SEC is performed in the isocratic mode and the RP-HPLC is performed in the gradient elution mode, and the HP-SEC eluent and the eluent A of the RP-HPLC are identical.
2. Perform a HP-SEC run (the RP-HPLC column is off-line).
3. Perform a RP-HPLC run (the HP-SEC column is off-line).
4. Determine the valve switching times.
5. Equilibrate both columns with eluent A (with pump P₁), with both columns connected.
6. Switch the RP-HPLC column off-line.

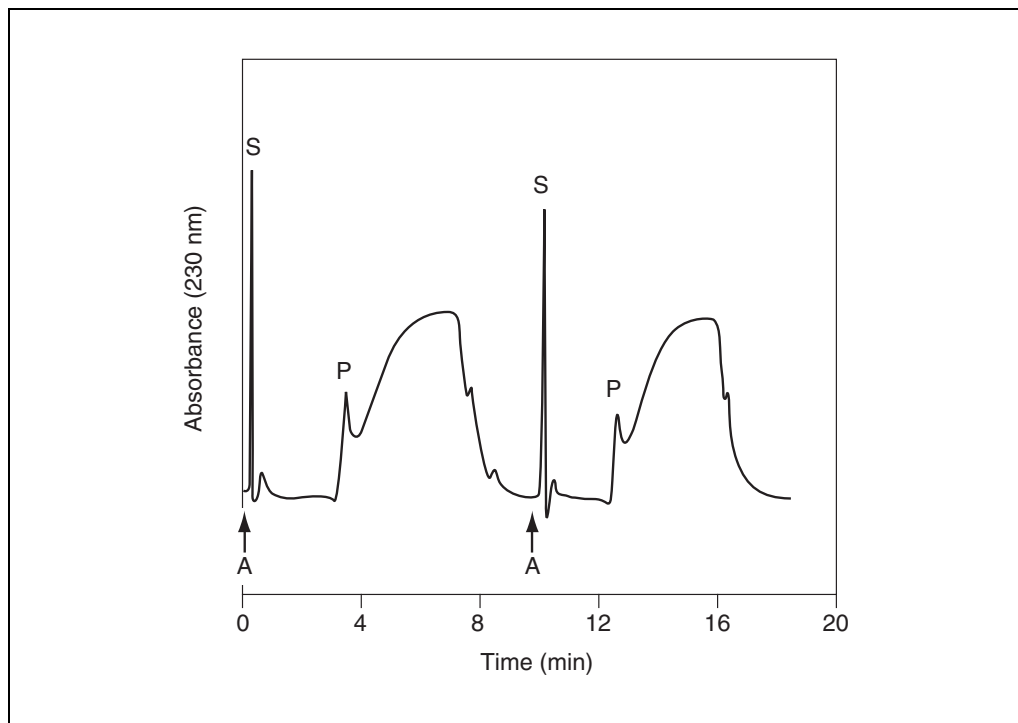


Figure 10.13.10 Illustrative example of the use of RP-HPLC in the desalting of peptide and protein samples isolated by other techniques. In this example, the 50S ribosomal protein from *Thermus aquaticus* on a C-4 column (40-mm length \times 4-mm i.d., 300 Å, 5 μ m) were desalted at a flow rate of 2.5 ml/min using a step elution protocol with 3 min elution with eluent A (0.1% TFA in water) and 3 min elution with eluent B (0.1% TFA in 2-propanol), pH 2.1. Abbreviations: S, salt; P, protein; A, time of injection.

7. Perform a HP-SEC run according to step 1 (see Table 10.13.1).

8. Switch to LC-LC coupling according to step 2 (see Table 10.13.1). The fraction is collected at the inlet of the RP-HPLC column.

9. Switch back to the HP-SEC only mode according to step 3 (see Table 10.13.1). The fraction collection on the RP-HPLC column is stopped. The sample stays on the column (no flow) until the HP-SEC column run is completed.

10. After completion of the HP-SEC run, the HP-SEC column is switched off-line, but flushed with buffer by pump 3, while a RP-HPLC gradient elution run is performed according to step 4 (see Table 10.13.1).

Illustrative of the use of multimodal HPLC techniques for the separation and analysis of peptide and protein samples is the example shown in Figure 10.13.12, panels A to C, for the separation of closely related 50S ribosomal proteins from *Thermus aquaticus* under RP-HPLC, HP-SEC, and again RP-HPLC protocols under different elution conditions to achieve optimal resolution of specific proteins in high purity for further investigations.

METHOD DEVELOPMENT IN RP-HPLC

For the separation of the diverse components of a sample containing peptides or proteins of unknown composition, usually a model is first developed which aims to create separation conditions that result in different retention times of the various components. Currently no algorithm can predict with absolute fidelity, on the basis of the amino acid sequence, the separation behavior of peptides or proteins. In many cases, the nature of the components is not known anyway. Moreover, changes in retention and peak shape will always occur when sample overload or volume overload conditions prevail. There are, however, empirical concepts that describe the retention behavior of peptides and proteins with a ligand in the presence of different solvent combinations. The most commonly adapted concepts are based on the solvophobic theory (Horvath et al., 1976, 1977) and the linear solvent strength theory (Snyder, 1980). These concepts allow the development of fast, robust, and cost-effective separation methods (Boysen et al., 1998).

Various aspects of these theoretical approaches can be used to reach different aims

with a specific separation (Table 10.13.2). These aims may be:

- a. the purification of one component, or alternatively several components simultaneously;
- b. the desalting of a peptide or protein sample;
- c. the characterization of physico-chemical properties of peptides and proteins in hydrophobic environments;
- d. examination of the unfolding behavior of proteins under different sorbent or mobile phase conditions;
- e. determination of linear free energy dependencies between different members of a peptide analog family; selection and optimization of different elution protocols;
- f. examination of the effects of different hydrogen bonding solvents on the retention behavior of peptides or proteins;
- g. the generation of a large variety of empirical data that permits different sorbent types to be validated or batch-to-batch variations characterized. The requirements, in terms of preliminary measurements, for these optimization procedures and data analyses are outlined in Table 10.13.2.

SYSTEMATIC APPROACH TO METHOD DEVELOPMENT

The quality of a separation is determined by the resolution of individual peak zones (Hearn, 1991a). Hence, method development is always an optimization of the resolution, with the trade-off either being the speed or the capacity (sample size) of the separation. The resolution of adjacent peak zones can be defined as:

$$R_s = \frac{(t_1 - t_2)}{(1/2)(\omega_1 + \omega_2)}$$

Equation 10.13.1

Here, t_1 and t_2 are the retention times, while ω_1 and ω_2 are the peak widths of two adjacent peaks. The baseline separation (an overlap of the peaks less than 1%) is given by definition as $R_s = 1.5$. The resolution of two peak zones depends on a large number of factors, including effects that influence the peak symmetry but also on the ratio of the peak areas. R_s values are easily estimated using tables and graphs (Snyder et al., 1988). In a chromatogram, every peak pair has a different R_s value. In developing a very high-resolution analytical separation of a complex mixture of peptide or protein components, method development always focuses on

the least well resolved peak pair. If conditions can be developed to ensure that this “critical peak pair” is well resolved, then all other peaks are also well resolved. However, the peaks that constitute this critical pair can change as a consequence of changes in the experimental conditions. For preparative separations, method development always focuses on the peak of interest and the two adjacent contaminant peaks. In this case, all other peaks can be viewed as superfluous, and directed to the waste. Optimization of the resolution of the peak of interest from the adjacent peaks has to take into account the sample size and the relative abundances of the three components that form the basis of the separation task.

In gradient elution, the resolution also depends on the plate number N , the selectivity α , and the capacity factor \bar{k} , all of which can be experimentally influenced through systematic changes in individual chromatographic parameters. In this case, resolution is determined from:

$$R_s = (1/4)N^{1/2}(\alpha - 1) \left[\frac{\bar{k}}{1 + \bar{k}} \right]$$

Equation 10.13.2

The plate number, N , is the band broadening of the peak zone caused by the column and is a measure of the column performance. The selectivity α describes the selectivity of a chromatographic system for a defined peak pair and is the ratio between the \bar{k} value of the second peak zone and the \bar{k} value of the first peak zone. The capacity factor \bar{k} is a dimensionless parameter of the retention in gradient elution. Its calculation from gradient elution data will be described further below.

In contrast to the isocratic elution, in a gradient elution system N , α , and \bar{k} are the median values for these variables (Hearn, 1991b), since they change during the separation as the shape or duration of the gradient changes. The plate number N has no influence on the selectivity or the retention (except for conditions of temperature change). The selectivity α and the capacity factor \bar{k} have only a minor influence on N . While N and α change only slightly during the solute migration through the column, the \bar{k} value can change by a factor of 10 or more. The best chromatographic separation is achieved within a \bar{k} value range between 1 and 20. Although the resolution is mainly influenced by the mobile phase variables α and \bar{k} , and nearly independent of N for a given column, an optimization strategy should start, from a logistical

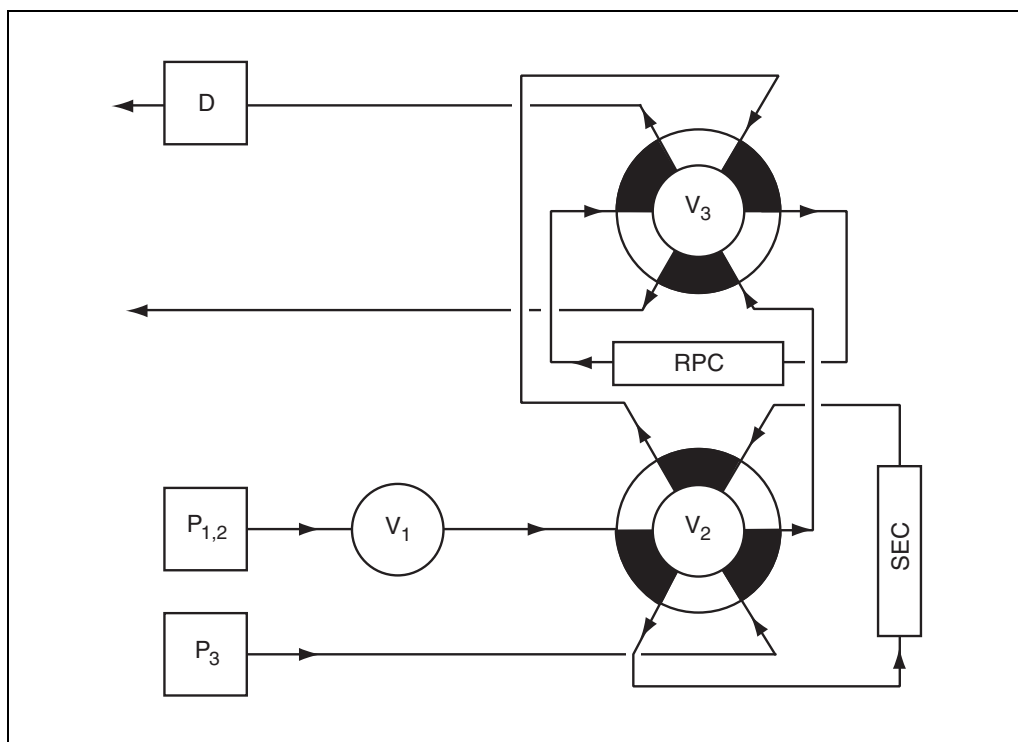


Figure 10.13.11 Diagram of column switching unit employed in multi-dimensional HPLC. Abbreviations: D, detector; P_{1,2}, two-pump module for gradient elution mode; P₃, pump equilibrates offline SEC column, while RPC gradient is performed; V₁, injection valve; V₂ and V₃, switching valves. The black field in the switching valve position symbolizes the joined valve positions. In the depicted position, the SEC and the RPC column are connected.

Table 10.13.1 Valve Positions and Pump Activities During the Four Steps in LC-LC Coupling

Step	1	2	3	4
Event	HP-SEC	Start LC-LC	End LC-LC	RP-HPLC
Valve 2: HP-SEC	On	On	On	Off
Valve 3: RP-HPLC	Off	On	Off	On
Mode	Isocratic	Isocratic	Isocratic	Gradient
Pump 1	On	On	On	On
Pump 2	Off	Off	Off	On
Mode	—	—	—	Isocratic
Pump 3	Off	Off	Off	On

Figure 10.13.12 (At right) Illustrative example of the use of multimodal HPLC techniques for the separation and analysis of peptide and protein samples. **(A)** RP-HPLC resolution of the 50S ribosomal proteins from *Thermus aquaticus* in 2% acetic acid (Molnar et al., 1989b) on a C8 column (250-mm length × 4.6-mm i.d.) using a flow rate of 1.5 ml/min and a 30 min gradient from 10% to 100% eluent B [eluent A: 125 mM NaH₂PO₄, 125 mM H₃PO₄; eluent B: acetonitrile: 250 mM NaH₂PO₄, 250 mM H₃PO₄ (50:50), buffer system, pH 2.1]. **(B)** HP-SEC separation of the 50S ribosomal proteins from *Thermus aquaticus* in 2% acetic acid on a tandem TSK-250 column (75-mm length × 7.5-mm i.d. and 300-mm length × 7.5-mm i.d.) using a flow rate of 1.5 ml/min with a isocratic separation with 10% B (eluent as described above) with a fraction cut out and directed on to the reversed-phase column. **(C)** RP-HPLC separation of the fraction of the 50S ribosomal proteins from *Thermus aquaticus* in 2% acetic acid obtained on-line from the size-exclusion chromatography on a C8 column (250-mm length × 4.6-mm i.d.) using a flow rate of 1.5 ml/min and a 30 min gradient from 10% to 100% B (eluent as described above). Proteins 6, 9, and 10, which were coeluted with other proteins in chromatogram A, could now be obtained in relatively pure form for further investigations.

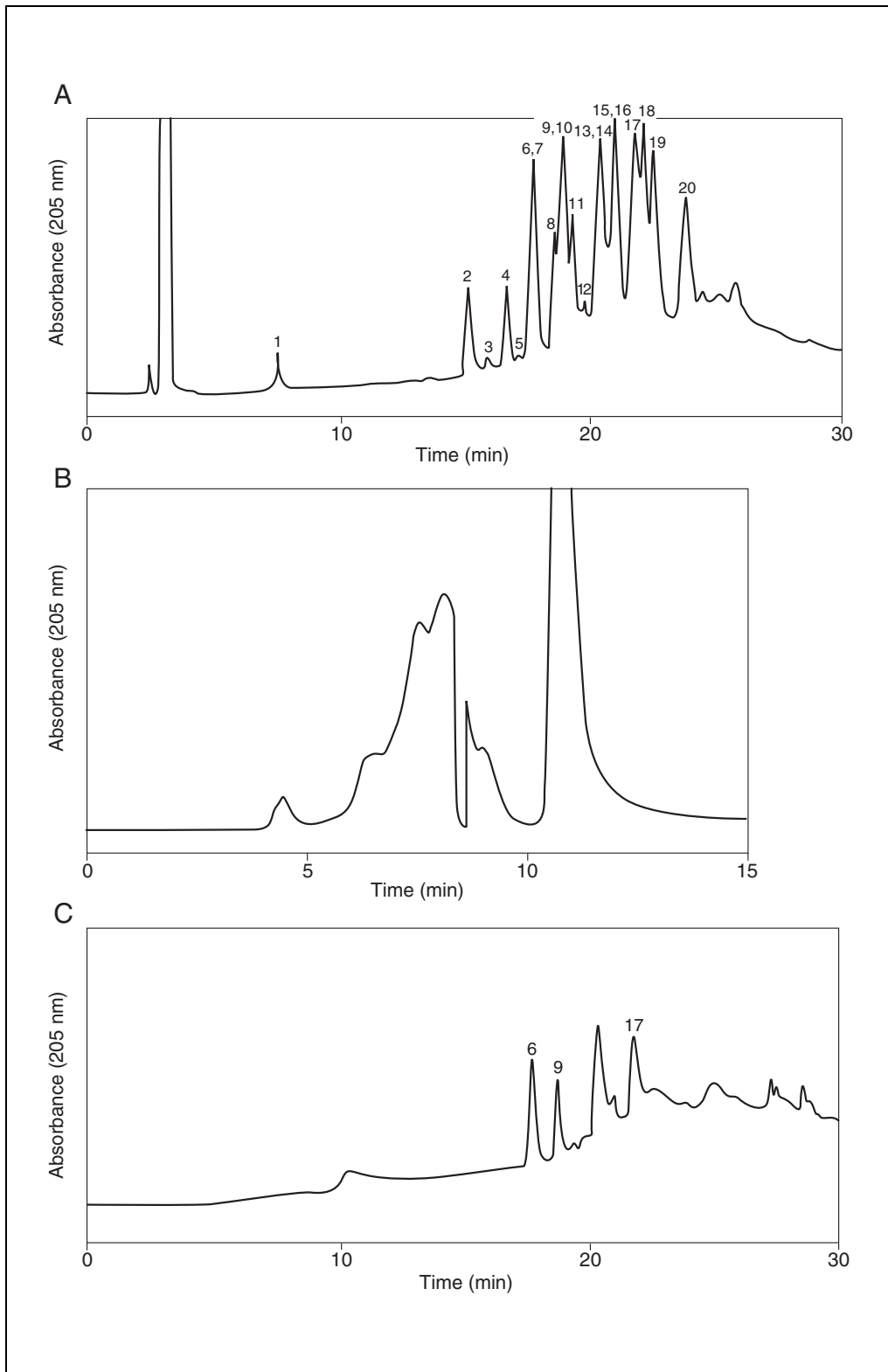


Figure 10.13.12 (See legend on facing page.)

Table 10.13.2 Methodological Requirements of the Different Tasks in RP-HPLC

Task	Purification of one component	Purification of several components simultaneously	Desalting of peptide or protein sample	Characterization of physico-chemical properties
Example	Peptides from solid phase peptide synthesis, enzymatic digestion of a protein	Peptide from tryptic digests, ribosomal proteins, proteins from biological extracts	Fractions from IEX or HIC	Thermodynamic parameters of solute/ligand interaction
Mode	Gradient elution	Gradient elution	Step elution	Isocratic elution
Analytical scale	Yes	Yes	Yes	Yes
Preparative scale	Yes	Yes	Yes	Yes
Measuring of t_0	Yes	Yes	Yes	Yes
Measuring of t_D	Yes	Yes	Yes	—
Column testing	Yes	Yes	Yes	Yes
A value	Yes	Yes	—	Yes
X value	—	Yes	—	—
Y value	—	Yes	—	—
Extra-column band broadening (σ)	—	Yes	—	Yes
Regression analysis	Yes	Yes	—	Yes
Optimization procedures	Yes	Yes	Yes	Yes
Computer simulation	—	Yes	Yes	Yes

point of view, with the optimization of the stationary phase, because many ab initio choices (like column dimensions, choice of ligand, etc.) are determined by the overall strategy (e.g., analytical versus preparative separations). It should be noted, from a theoretical point of view, that optimization of the stationary phase is often performed last (as in some computer simulation programs) when an established analytical method is up-scaled to a preparative method. Typically, a particular RP-HPLC sorbent will be selected empirically as the starting point for the separation. Rarely do most investigators go to the trouble of optimizing the surface chemistries of their own sorbents, although a number of notable exceptions exist. Based on these considerations and in view of the implication of Equation 10.13.2, the optimization can be performed in three modules in the following order: (1) the optimization of the plate number N , then (2) optimization of the selectivity α , and finally, (3) optimization of the \bar{k} values.

Optimization Module 1: The Stationary Phase

The optimization of the peak efficiency, expressed as the theoretical plate number, N , requires an independent optimization of each of the contributing factors that influence the band broadening of the peak zones due to column and the extra-column effects. With a particular sorbent (ligand type, particle size, and pore size) and column configuration, this can be achieved through optimization of the linear velocity (flow rate), the temperature, the detector time constant, and the column packing characteristics, as well as by minimizing extra-column effects, e.g., by using zero-dead volume tubing and connectors. The temperature of the column and the eluents should be constant ($\pm 0.1^\circ\text{C}$) using a thermostatically controlled system in order to facilitate the reproducible determination of the various column parameters and to ensure reproducibility in the resolution.

The theoretical plate numbers measured on the same column are smaller for proteins than for small peptides. This difference in peak width behavior between peptides and proteins can be explained directly from the Knox equation (Knox and Scott, 1983; Freebairn and Knox, 1984), if the N values are expressed as reduced plate heights, i.e., $h = L/Nd_p$, where L is the column length, and d_p is the average particle diameter of the sorbent. According to the Knox equation, h is related to the reduced mobile phase velocity, v , through the relationship $h = Av^{0.33} + Bv^{-1} + Cv$, where $v = ud_p/D_m$, u is the linear velocity, and A , B , and C are constants that describe the packing and solute transport characteristics of a particular sorbent in a column. Observations of reduced plate heights (h values) in the range $3 < h < 15$ for small peptides with reversed-phase sorbents of average particle size of 3 to 10 μm , and for proteins and other high molecular weight macromolecules where h values are >50 , have frequently been made, confirming that the Knox equation can be approximated (Snyder et al., 1983) for small peptides as $h = Av^{0.33}$ and for proteins as $h = Cv$, because, in the latter case, only the intraparticle pore transport and surface interaction (term C), contributes significantly to the band broadening.

The following parameters can be used to assess column and extra column effects:

The A value of the column (a term that reflects the quality of the packing of the sorbent in the column) can be determined according to the procedures described in Stadalius et al. (1987). The A values for well packed columns fall within the limits of $0.5 \leq A \leq 1$. An A value equal to one corresponds to a column containing a theoretically ideal bed of particles in a dodecahedral close-packed arrangement (Stout et al., 1983).

The X value (the column volume of mobile phase outside the pores of the particles) can be calculated according to the procedures described in Stadalius et al. (1987).

The average Y value (an estimate of the diffusional restriction of proteins within the pores of the sorbent) can be calculated according to the procedures described in Stadalius et al. (1987). When no restricted diffusion occurs, the value of Y , which corresponds to the ratio (D_p/D_m) is equal to one. Here D_p is the diffusion constant of the protein within the pores of the particle and D_m is the diffusion coefficient of the peptide or protein in the bulk mobile phase. For proteins with molecular weights within the range 10,000 to 80,000, separated using RP-

HPLC sorbents of large pore diameters (e.g., ≥ 30 nm or larger), the Y values typically fall within the narrow range of 0.8 ± 0.1 , indicating that limited restricted diffusion can occur but that the D_p/D_m ratio is largely independent of the molecular size and weight of the examined proteins. However, with particles of smaller pore size, such as ≤ 10 nm, a larger range of Y values (e.g., $0.05 \leq Y \leq 0.80$) will occur with a significant dependency of the Y value on the molecular characteristics of the test proteins.

The measurement of the extra-column band broadening σ (which describes the decrease in column performance due to instrumental design) can be determined with a "zero-length column" that connects the injector with a sample loop of 1 μl directly to the detector (Freebairn and Knox, 1984; Hupe et al., 1984). Its value can be calculated based on the calculated peak asymmetry values according to the procedures described in Stadalius et al. (1987).

The extra-column band broadening or volume dispersion effects caused by the injector, tubing, and detector influence the performance of the column and should be less than 10% of the plate number of the column. The extra-column band broadening expressed as peak standard deviation should therefore be less than a third of the band broadening due to the characteristics of the packed column, e.g., as given by:

$$\sigma_{\text{COL}} = \frac{V_0}{\sqrt{N}} \text{ with } \sigma_{\text{DI}} \leq (1/3)\sigma_{\text{COL}}$$

Equation 10.13.3

Hence, the LC system should be designed so as to have minimal extra-column band broadening. Care should be taken to attach fittings in such a way as to avoid unnecessary void volumes and to choose tubing with the smallest possible inner diameter, as well as to keep the volume of the detector cell low.

For a standard analytical column, the maximum extra-column band broadening can be therefore be calculated as:

$$\sigma_{\text{COL}} = \frac{1900\mu\text{l}}{\sqrt{10000}} = 19\mu\text{l} \text{ with } \sigma_{\text{DI}} \leq 6.3\mu\text{l}$$

Equation 10.13.4

In order to obtain this value, the detector cell volume must be 5 μl or less.

The flow rate to achieve the minimum plate height, H , for a column can be taken from the literature or experimentally determined accord-

ing to procedures described in Stadalius et al. (1987).

Optimization Module 2: The Mobile Phase Composition

Change in selectivity of the separation with peptides and proteins is the most effective way to influence resolution. This is mainly achieved by changing the chemical nature or concentration of the organic solvent modifier (acetonitrile, methanol, isopropanol, etc.) or the choice of an ion pair reagent (Hancock et al., 1978).

Choice of the organic solvent modifier

A good starting point is the solvent selectivity triangle approach. Here, solvents are classified according to their relative dipole moment, basicity, or acidity in a triangle. Blends of three different solvents, plus water to provide an appropriate \bar{k} range, are selected to differ as much as possible in their polar interactions. This selection permits the solvent combinations to mimic the selectivity that is possible for any given solvent, and confines the boundaries of the triangle (Snyder, 1974, 1978). At the same time, these solvents must be totally miscible with each other and with water. Three solvents that best meet these requirements are methanol, acetonitrile, and tetrahydrofuran. Four-solvent mobile phase optimization using three organic solvents and water provides a significant control over α values in reversed-phase HPLC. If different organic solvents are used, the different eluotropic strengths (Schoenmakers et al., 1979; Patel and Jefferies, 1987) must be considered in order to elute the sample in the appropriate \bar{k} range.

Choice of ion pair reagent

The retention of peptides can be influenced by the presence of ion pair reagents (Horvath et al., 1977; Hearn et al., 1979; Goldberg et al., 1984; Mant and Hodges, 1991c). The ion pair reagents interact with the ionized groups of the peptides: anionic counterions (e.g., hexanesulfonic acid, orthophosphoric acid, TFA, HFBA) interact with basic residues (arginine, lysine, histidine) of a peptide as well as with the protonated *N* terminus; cationic counterions (e.g., triethylammonium, tetrabutyl-ammonium) interact with ionized acidic residues (glutamic and aspartic acid, tyrosine and cysteic acid) as well as ionized free C-terminal carboxylic groups. The actual effect on retention depends strongly on the hydrophobicity of the ion-pair reagent and the number of oppositely charged groups on the peptide. Typically, the retention

of peptides and proteins in the presence of ion pairing reagents follow asymptotic dependencies on the concentration of the different ion pairing reagent.

Once the selectivity parameter is fixed due to the *initial* choices of the mobile and stationary phase, the further optimization should concentrate on resolution optimization via achieving the most appropriate \bar{k} value for the different peptides or proteins in the mixture.

Optimization Module 3: The Gradient Conditions

For resolution optimization, the strategy takes advantage of the relationship between the gradient retention time of a protein (expressed as the median capacity factor, \bar{k}) and the median volume fraction of the organic solvent modifier $\bar{\phi}$, in regular RP-HPLC systems based on the concepts of the linear-solvent-strength theory (Horvath et al., 1976; Snyder, 1980; Hearn and Grego, 1983), such that:

$$\ln \bar{k} = \ln k_0 - S\bar{\phi}$$

Equation 10.13.5

where k_0 is the capacity factor of the solute in the absence of the organic solvent modifier, and S is the slope of the plot of $\ln \bar{k}$ versus $\bar{\phi}$. The values of $\ln k_0$ and S can be calculated by linear regression analysis. The underlying principles of an intuitively performed optimization and manually achieved optimization (using Excel spreadsheets, for example, to calculate the $\ln k_0$ and S values), or, alternatively, optimization via computer simulation software (e.g., Simplex methods, multivariate factor analysis programs, DryLab *G/plus*, etc.) are essentially the same. However, the outcomes result in different levels of precision. Two representative approaches are collectively outlined below.

Resolution R_S of peak zones is optimized through adjustment of \bar{k} by successive change of the parameters t_G and $\Delta\phi$ in the gradient elution mode according to the following steps:

- Initial experiments;
- Peak tracking and assignment of the peaks;
- Calculation of $\ln \bar{k}$ and S values from initial chromatograms;
- Optimization of gradient run time t_G over the whole gradient range;
- Determination of new gradient range;
- Calculation of new gradient retention times t_g ;
- Change of gradient shape (optional);
- Verification of results.

Initial experiments

In initial experiments, the peptide or protein sample is separated under two linear gradient conditions differing by a factor of 3 in their gradient run times t_G (all other chromatographic parameters being held unchanged; Dolan et al., 1989) to obtain the RP-HPLC retention times of each peptide or protein. Irrespective of what optimization strategy will be then used, it is advisable to separate any sample with at least two different gradient run times, in order to identify overlapping peaks. For optimization of the gradient shape and to achieve maximum resolution between adjacent peak zones, the ability to determine the retention times of the peptides or proteins and to classify the parameters that reflect contributions from the mobile phase composition and column dimensions (Table 10.13.3) is essential (Ghrist and Snyder, 1988a,b; Ghrist et al., 1988). The determination of the volume, V_{mix} , is useful (Stadalius et al., 1984); determination of dead volume and gradient delay are essential (Snyder, 1980). With these compiled input parameters, various algorithms including DryLab *G/plus* can generate the relative resolution map (RRM), based on calculation of the corresponding S values and k_0 values for each component. If no DryLab *G/plus* is available, the resolution information can be extracted directly from the vertical distances of the individual R_S versus t_G plots, whereby a cross-over of lines reflects a peak overlap as shown in Figure 10.13.13A,B,C. Over a small range of $\bar{\phi}$ values, the S values and k_0 values for an individual peptide or protein will remain essentially constant. It is assumed that no significant conformational change of the individual peptides or proteins occurs as a consequence of the differences in column residency time and gradient time (Purcell et al., 1993). The typical retention behavior of the peptides and proteins is depicted in Figure 10.13.13 as plots of $\ln \bar{k}$ versus $\bar{\phi}$. These plots become nonparallel, and are frequently intersecting, as the mobile phase composition is varied.

Choice of starting conditions

The S values of peptides and proteins can be empirically correlated with their molecular weights (Snyder et al., 1983), based on a correlation of $S = 0.48(MW)^{0.44}$. Thus S values, estimated with this equation or, alternatively, values of similar peptides or proteins taken from the literature, can be used to calculate a reasonable starting point for the initial experiments according to the following equation:

$$t_G = \frac{V_m \times \Delta\phi \times S \times \bar{k}}{0.87 \times F}$$

Equation 10.13.6

The required gradient run time, t_G , for a separation of peptides and proteins with expected S values of around 20 can be calculated for a gradient of 0% to 100% (60% ACN) at a flow rate of 1 ml/min and the ideal \bar{k} value of 5 as follows:

$$t_G = \frac{1.9 \text{ ml} \times 0.6 \times 20 \times 5}{0.87 \times 1 \text{ ml/min}} = 131 \text{ min}$$

Equation 10.13.7

Based on this calculation, the initial gradients from 0% to 100% eluent B for gradient times of 1 and 3 hr duration can be selected.

On the other hand, availability of the S values, derived from two gradient elution RP-HPLC experiments, can be used as an analytical criterion early in the separation of the target peptide or protein from other components in soluble extracts of biological sources, to distinguish low- from high-molecular-weight molecules and, e.g., to exclude peptide fragments participating in the optimization process, without having to resort to SDS-PAGE experiments. The RP-HPLC behavior of small peptides with molecular weights from 300 to 1000 are approximately correlated (Hearn and Grego, 1981; Snyder, 1990; Wilce et al., 1995) to S values between 3 and 10. For medium-molecular-weight globular proteins, S values above 20 are expected (Aguilar et al., 1985). As an example, cytochrome *c*, with a molecular weight of ~12,000, has an S value of 28.8 on a Nucleosil C-18 column (Stadalius et al., 1984).

Calculation of $\ln k_0$ and S values

The retention times t_{g1} and t_{g2} for a peptide or protein solute separated under conditions of two different gradient run times (t_{G1} and t_{G2} , whereby $t_{G1} < t_{G2}$) can be given by the following equations (Quarry et al., 1986; Ghrist et al., 1988):

$$t_{g1} = \left(\frac{t_0}{b_1} \right) \log(2.3k_0b_1) + t_0 + t_D$$

Equation 10.13.8

$$t_{g2} = \left(\frac{t_0}{b_2} \right) \log(2.3k_0b_2) + t_0 + t_D$$

Equation 10.13.9

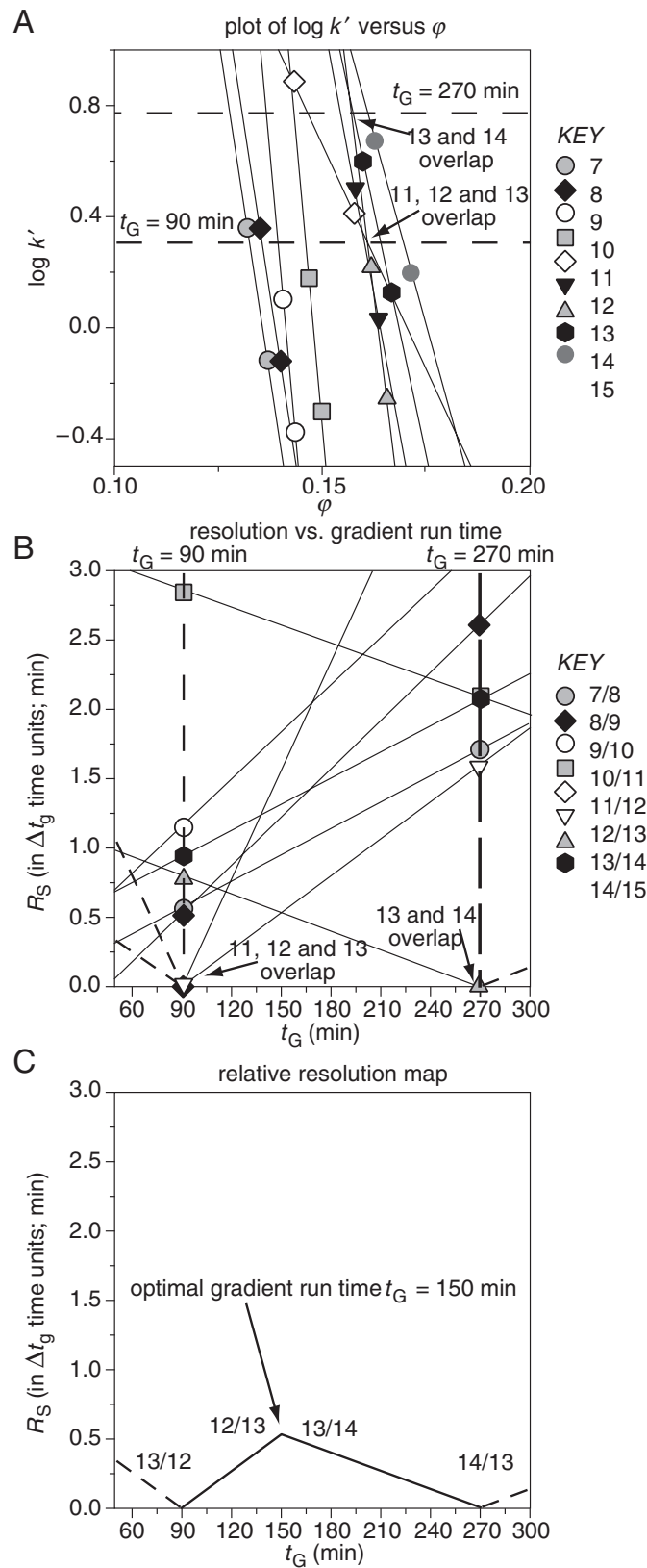


Figure 10.13.13 (A) Plots of logarithmic capacity factor $\ln k'$ versus the volume fraction of organic solvent, ϕ , for 9 different proteins, denoted 7 to 15, obtained from *Thermus aquaticus* on a C18 RP-HPLC column. As evident from this plot, the intersection of two plots represents complete overlap of the peak zones at the specific gradient run times t_G 90 and 270 min. (Legend continues on next page.)

where:

$$\frac{b_1}{b_2} = \frac{t_{G2}}{t_{G1}} = \beta$$

Equation 10.13.10

Here, t_{G1} , t_{G2} are the gradient run time values of t_G for two different gradient runs, resulting in different values of b (b_1 , b_2) and t_g (t_{g1} , t_{g1}) for a single solute; t_{g1} , t_{g1} are the gradient retention times for a single solute in two different gradient runs; b_1 , b_2 are the gradient steepness parameters for a single solute and two gradient runs differing only in their gradient times. Steep gradients correspond to large b values and small \bar{k} values; k_0 is the solute capacity factor at the initial mobile phase composition; β is the ratio of t_{G2} and t_{G1} which is equivalent to the ratio of b_1 and b_2 ; t_0 is the column dead time; and t_D is the gradient delay time.

For peptides and proteins there is an explicit solution (Stadalius et al., 1984; Quarry et al., 1986) for b and k_0 , namely:

$$b_1 = \frac{t_0 \log \beta}{\left[t_{g1} - \left(\frac{t_{g2}}{\beta} \right) + (t_0 + t_D) \left(\frac{1 - \beta}{\beta} \right) \right]}$$

Equation 10.13.11

$$\log k_0 = \left(\frac{b_1}{t_0} \right) (t_{g1} - t_0 - t_D) - \log(2.3b_1)$$

Equation 10.13.12

From the knowledge of b and k_0 the values of \bar{k} and $\bar{\varphi}$ can be calculated (Snyder, 1980; Hearn, 1991b):

$$\bar{k} = \frac{1}{1.15b_1}$$

Equation 10.13.13

$$\bar{\varphi} = \frac{\left[t_{g1} - t_0 - t_D - \left(\frac{t_0}{b_1} \right) \log 2 \right]}{t_{G1}^0}$$

Equation 10.13.14

Table 10.13.3 Parameters Required for the Calculation of $\ln k_0$ and S Values and for the Manual Optimization and Optimization with DryLabG/plus

Parameter	S and $\log k$	RRM
	Manual	DryLab
Gradient delay volume V_D (ml)	Yes	Yes
Dead volume t_0 (ml)	Yes	—
Gradient range (%B)	Yes	Yes
Gradient run time t_{G1} , t_{G2} (min)	Yes	Yes
Column length L (mm)	—	Yes
Column diameter d_C (mm)	—	Yes
Flow rate F (ml/min)	—	Yes
Column plate number N	—	Yes
Retention times t_{g1} for all peaks	Yes	Yes
Retention times t_{g2} for all peaks	Yes	Yes

Figure 10.13.13 (Continued) **(B)** Plots of the resolution R_S (resolution taken as the difference in retention times, Δt_g , for two adjacent peaks) versus the gradient run time t_G of these 9 different proteins from *Thermus aquaticus*. These data represent the blueprint for generation of the relative resolution map (RRM). Since the plotted values of the resolution are absolute values, negative resolution values of peaks that change their elution order are depicted as broken lines. **(C)** RRM is shown for the nine different proteins from *Thermus aquaticus*, depicting only the critical peak pairs including their change in retention order. The resolution optimum for this specific set of chromatographic conditions occurs at $t_G = 150$ min.

Here, \bar{k} is the value of k' (capacity factor) for a solute when it reaches the column midpoint during elution; ϕ is the volume fraction of solvent in the mobile phase; $\Delta\phi$ is the change in ϕ for the mobile phase during the gradient elution ($\Delta\phi = 1$ for a 0% to 100% gradient); $\bar{\phi}$ is the effective value of ϕ during gradient elution and the value of ϕ at band center when the band is at the midpoint of column, and t_G^0 is the normalized gradient time with $t_{G1}^0 = t_{G1}/\Delta\phi$.

By linear regression analysis, using \bar{k} and $\bar{\phi}$, the S value (empirically related to the hydrophobic contact area between solute and ligand) can be derived from the slope of the $\log\bar{k}$ versus $\bar{\phi}$ plots, and $\ln k_0$ (empirically related to the affinity of the solute towards the ligand) as the y-intercept (Horvath et al., 1976):

$$S = \frac{(\ln k_0 - \ln \bar{k})}{\bar{\phi}}$$

Equation 10.13.15

Peak tracking and assignment of the peaks

Complex chromatograms that result from reversed-phase gradient elution can often exhibit changes in peak order when the gradient steepness is changed. Before $\ln k_0$ and S values are calculated, or computer simulation is used, the peaks from the two initial runs need to be correctly assigned. Several approaches to peak tracking have been described, using algorithms based on relative retention and peak areas (Glajch et al., 1986; Lankmayr et al., 1989; Molnar et al., 1989a), or alternatively, based on diode-array detection (Berridge, 1986; Strasters et al., 1989; Round et al., 1994).

The assignment of peaks can be done in the following steps:

1. Integrate the chromatograms (using the integration software of the HPLC system or alternatively an integrator) of the initial runs and correct integration due to baseline drift or other instrumental causes where necessary.
2. Print out both chromatograms including the peak area percent reports.
3. Number all relevant peaks in the chromatogram with the better resolution. Relevant peaks have, e.g., an area >0.5% of the overall peak area.
4. Assign the peaks according to their peak areas allowing a reasonable elution window.

The total peak areas, A_{T1} and A_{T2} for the initial runs 1 and 2 (whereby $t_{G1} < t_{G2}$) are determined, as:

$$A_T = \sum_{i=1}^n A_i$$

Equation 10.13.16

and their ratio is calculated:

$$R_T = \frac{A_{T2}}{A_{T1}}$$

Equation 10.13.17

Consequently the ratio of the peak areas, A_{i1} and A_{i2} of a baseline separated single component in the initial runs 1 and 2, respectively, is:

$$R_i = \frac{A_{i2}}{A_{i1}}$$

Equation 10.13.18

If the peaks are correctly assigned, $R_T = R_i$. Difficulties arise when peaks partially or completely overlap. If $R_T \ll R_i$ and the peak in run 2 is composed of a single component, the peaks are wrongly matched. If $R_T \gg R_i$ then the peaks might be correctly matched, but the peak area in run 1 could be enlarged by a hidden additional peak. In order to resolve these difficulties, the deviation ($\%R_T$) of R_i from R_T can be calculated and taken as a measure of the likelihood that a peak assignment hypothesis is correct:

$$\%R_T = \frac{100R_i}{R_T} - 100$$

Equation 10.13.19

Optimization of the gradient time, t_G , over the whole gradient range

The capacity factor \bar{k} is a linear function of the gradient run time t_G if $\Delta\phi$ is kept constant. Hence:

$$\frac{\bar{k}}{t_G} = \frac{0.87F}{V_m \times \Delta\phi \times S} = \text{constant} = C$$

Equation 10.13.20

The optimized gradient run time t_{GRRM} can be obtained from the RRM or alternatively, from the plot of R_S versus t_G (see Fig. 10.13.13, panels A to C), and yields for each peptide or protein the new values of \bar{k}_{new} by t_{GRRM} being multiplied by C :

$$C t_{GRRM} = \bar{k}_{new}$$

Equation 10.13.21

Determination of the new gradient range

If the gradient run time t_{GRRM} is changed in relation to $\Delta\phi$ with $t_{\text{G1}}^0 = \text{constant}$, the \bar{k} values do not change, as can be seen from the following equation:

$$t_{\text{G1}}^0 = \frac{t_{\text{GRRM}}}{\Delta\phi} = \frac{V_m \times S \times \bar{k}}{0.87 \times F}$$

Equation 10.13.22

where:

$$\Delta\phi_{\text{opt}} = \frac{t_{\text{Gopt}}}{t_{\text{G1}}^0}$$

Equation 10.13.23

and the retention time t_g of the first peak is $>(t_0 + t_D)$ and the retention time t_g of the last peak is $<\Delta t_{\text{Gopt}}$.

Calculation of the new gradient retention times t_g

Based on the knowledge of the S and the $\ln k_0$ values, new gradient retention times can then be calculated.

Change of gradient shape (optional). A multisegmented gradient should only be performed when the gradient delay has been measured. With multisegmented gradients, an error in the gradient delay will reoccur at the beginning and at the end of each gradient step. In addition, the effect of V_{mix} (which can be determined according to the procedures described in Ghrist et al., 1988), which modifies the composition of the gradient at the start and end (rounding of the gradient shape), can lead to deviation of the experimentally determined from the predicted “ideal” retention times in DryLab *G/plus* simulations.

Verification of the results. After completion of the optimization process, the simulated chromatographic separation can now be verified experimentally using the predicted chromatographic conditions.

DETERMINATION OF THERMODYNAMIC PARAMETERS ASSOCIATED WITH PEPTIDE OR PROTEIN INTERACTIONS WITH IMMOBILIZED LIGANDS

Isocratic elution can be utilized to determine thermodynamic parameters associated with the interaction of a peptide or protein with an immobilized ligand in all of the various HPLC modes. Thus, the enthalpy, $\Delta H^0_{\text{assoc}}$, the entropy, $\Delta S^0_{\text{assoc}}$, and the heat capacity, ΔC^0_p , of

the association of a peptide or protein interacting with immobilized nonpolar ligand in RP-HPLC can be evaluated from the dependency of $\ln k'$ on T , i.e., from the van't Hoff plots (Melander et al., 1984; Haidacher et al., 1996; Vailaya and Horvath, 1996; Boysen et al., 1999; Hearn and Zhao, 1999; Hearn, 2001b), and from the Boltzmann-Helmholtz expression:

$$\ln k' = \frac{-\Delta H^0_{\text{assoc}}}{RT} + \frac{\Delta S^0_{\text{assoc}}}{R} + \ln \Phi$$

Equation 10.13.24

In order to derive this fundamental information, in addition to the procedures and reagents described in the sections above, the following materials and conditions are required:

Two 5-liter flasks

Peptide or protein (purity >95%) sample at a concentration of 1 mg/ml in a suitable buffer

1. Measure the dead volume t_0 of the system as described above.

2. Determine the isocratic conditions at which the peptides or proteins elute in a k' range of 2 to 10. Typically, these mobile phase conditions will encompass only a narrow range of values, e.g., 22% to 27% acetonitrile/water (v/v)/0.1% TFA for a RP-HPLC separation, and must be determined by running the sample under different eluent conditions.

3. Prepare stock solutions: e.g., 5 liters each of eluents of appropriate compositions that are ~2% above and 2% below the determined elution range (per the example above 20% and 30% v/v acetonitrile/water/0.1% TFA).

4. Manually mix final buffer in 1% steps from the two stock solutions, according to mixing tables (e.g., Tables for the Laboratory, Merck). Anticipate different eluent consumption for different mobile phase compositions.

5. Inject the sample in order to determine the retention times of the peptide or protein in the temperature range of 5° to 65°C (or higher depending on column specifications) at different isocratic eluent compositions in 5°C increments, allowing no more than 0.5°C column temperature fluctuation, measuring each data point at least twice.

6. Establish the van't Hoff plots and fit the data to a polynomial equation using linear regression analysis as described in Boysen et al. (1999).

7. Use the derived coefficients to calculate the thermodynamic parameters, the enthalpy $\Delta H^0_{\text{assoc}}$ of the association, the entropy $\Delta S^0_{\text{assoc}}$ of the association, and the heat capacity ΔC^0_p exemplified for a second-order approximation

of the solute-ligand interaction by the following equations (Boysen et al., 1999):

$$\ln k' = b_{(0)} + \frac{b_{(1)}}{T} + \frac{b_{(2)}}{T^2} + \ln \Phi$$

Equation 10.13.25

$$\Delta H_{\text{assoc}}^{\circ} = -R \left[b_{(1)} + \frac{2b_{(2)}}{T} \right]$$

Equation 10.13.26

$$\Delta S_{\text{assoc}}^{\circ} = R \left[b_{(0)} - \frac{b_{(2)}}{T^2} \right]$$

Equation 10.13.27

$$\Delta C_p^{\circ} = R \left[\frac{2b_{(2)}}{T^2} \right]$$

Equation 10.13.28

The physical basis and applications of this approach to characterize peptide- or protein-ligand interactions have been extensively described in a variety of papers and reviews (Hearn, 2000, 2001a,b).

TROUBLESHOOTING

There are three approaches to troubleshooting, which are usually practiced in combination: (1) preventive maintenance; (2) anticipation of problems during use by small signs of malfunction; and (3) complacency until complete breakdown.

The choice of the troubleshooting approach ultimately affects the spare part strategy and general instrument management and maintenance approach (see Table 10.13.4). As a consequence, the concept of preventative maintenance

addresses both the hardware and software, as well as the selection of sorbent characteristics, mobile phase composition, and sample preparation.

Preventive Maintenance

By practicing preventive maintenance, the downtime of HPLC instrumentation used in peptide and protein separation can be substantially minimized.

As a general principle, use clean buffer and reagents, use proper sample preparation techniques, keep air out of the LC system, check the system for leaks, remove all unwanted substances (sample, buffer components) from the system at the end of each day's work, and be aware of how the system works under normal conditions.

Mobile phase

Filter all mobile phases except pure HPLC-grade solvents through 0.2- μm filters. During filtration, avoid contamination that might result from the use, e.g., of dirty glassware or rubber stoppers. Degas all solvents thoroughly. Ensure that all fittings on the low-pressure side of the HPLC pump are tight, so that air is not drawn in. Add sodium azide (10,000 ppm) to the mobile phase in order to stop microbial growth, particularly in phosphate and acetate buffers, if they are stored at room temperature for longer than one day.

Use a solvent inlet line filter (usually 10 μm porosity) in order to prevent particulate matter entering the pump and to act as a weight to keep the inlet line at the bottom of the solvent reservoir. Cap reservoir loosely to prevent dust from entering the mobile phase and to allow air influx. Avoid cross-contamination of buffer, discard stale mobile phases, and clean the eluent reservoir regularly.

Table 10.13.4 Troubleshooting Strategy and Resulting Spare Part Policy

Troubleshooting strategy	Spare part strategy	Parts replaced
Preventive maintenance	Have always available and change frequently	Precolumn Guard column Column frits In-line filter Inlet filter
Anticipation of problems	Have available as backup parts	Column Fitting and tubing valves
Wait until complete breakdown	Purchase as needed	Lamps Circuit boards Pump heads Fuses

Stationary phase

Ensure compatibility of the sample solvent with the mobile phase. Incompatibility of the sample buffer with the mobile phase can cause a precipitation of sample in the pores of the column packing.

Ensure compatibility of the mobile phase components with each other. Mobile-phase strength effects can cause precipitation of eluent components, such as salts, inside the column, resulting in column blockage or a packing void.

Ensure compatibility of the column with the mobile phase. The set of guidelines supplied with the column reveal the pressure and temperature limit at which pH range the column can be used (the wrong pH can cause loss of bonded phase with many silica-based sorbents) and which solvents are compatible with the stationary phase (an incompatible solvent can shrink or swell column particles with the polymeric-type of sorbents).

Use in-line filter or guard columns. For SEC and other chromatographic modes with impure samples, the use of an 0.5- μm in-line filter is recommended in order to prevent blockage of the column inlet frit. A guard column, preferably packed with the same material as the separation sorbent, prevents a blockage of the column inlet frit and deterioration of the column performance. The integrity of the guard column must be checked regularly, and it must be exchanged when necessary, since a poor guard column can adversely affect column performance.

Avoid pressure shocks. Pressure shocks can be minimized by using a pulse damper and by avoiding drastic flow rate changes.

Avoid exceeding the pressure limit by using the upper limit switches of the pump in the method program. Do not use default values. Select the upper limit of pressure according to the back-pressure of a new column with an associated method (depending on the viscosity of the mobile phase, it could be the mid-gradient pressure when using water-organic solvents, e.g., 1500 psi) and add a margin of 1000 psi in order to avoid stopping of overnight runs, etc. (yielding a pressure limit of, e.g., 2500 psi).

Wash the column after use. All columns should be flushed with 10 column volumes of either an organic solvent/water mixture or water with 10,000 ppm sodium azide to suppress microbial growth.

Store the column, when not in use, in the appropriate solvent (see manufacturers' instructions)—either in >10% organic solvent or

with sodium azide when organic solvents have to be avoided. The mobile phase in which the column is stored should not contain any salts, acids, or bases.

Fittings and tubing

Assemble the fittings so that the tubing contacts the bottom of the fitting body. Tighten the fitting sufficiently to prevent leaks, but not so much as to distort the ferrule. In order to minimize the extra column band broadening, all tubing on the high-pressure side from the injector to the detector (excluding detector outlet) should be 0.01-in. and as short as convenient. Larger i.d. tubing (e.g., 0.020-in.) is suitable for connecting the pump with the injector. The extra column band broadening can be measured with a zero length column.

Manual injector

Filter all samples. Avoid pressure values above 5000 psi. Do not use syringes with sharp needles (e.g., from gas chromatography). Maintain adequate gas pressure for air- or nitrogen-operated valves. Flush injector at the end of each day's work. If a leaking valve requires the replacement of a damaged seal, rebuild the valve using the kit and instructions supplied by valve vendor. Finally, store the injector in a nonbuffered mobile phase when not in use.

Pump(s)

Avoid air bubbles entering pump. If an air bubble arises in the pump see discussion of Priming the Pumps and Low-Pressure Lines with Eluents in *UNIT 10.12*.

Always flush the HPLC after use to prevent salt deposits at the pump valves. If blocked, see discussion on Preparing the HPLC System in *UNIT 10.12*.

Flush behind the pump seal at the end of each day's LC operation, first with 5 to 10 ml water, then with 5 to 10 ml methanol or isopropanol in order to remove the water. Many LC systems have a flushing port. By design, the pump seal does not seal completely around the piston. Crystalline buffer residues may act as an abrasive on the pump seal when restarted, damaging the seal with possible scratching of the piston. Store pumps in nonbuffered mobile phase or pure organic solvent.

Detector(s)

Keep the detector cell clean by washing it at the end of each day's work with the strong mobile phase. Do not perform acid washing on

a routine basis. Switch the detector lamp off when not in use for the next 2 hr or longer. Avoid frequent "on and off" switching.

Recorder, printer and data systems

Replace ink when recorder ink begins to fade. Lift the pen from the paper and cap it, when recorder is not in use. Check paper supply before HPLC instrument is started and before leaving the HPLC instrument unattended. Ensure that adequate disk storage space exists within the computer for unattended runs.

Anticipation of Problems

With this strategy parts are replaced when minor signs of malfunction occur.

Troubleshooting after HPLC Breakdown

The first aim of troubleshooting is the reduction of HPLC instrument down time. Equally important is the prevention of further damage to the HPLC system, which includes the consequences of troubleshooting. It is therefore crucial to identify the point at which help is needed or whether an event is a case for the manufacturer's service (as most electronic problems are). Before calling the service engineer, the following sources of help can be utilized if a problem occurs:

Operation manual

Internal laboratory records, including the log-books

Troubleshooting guides

Other people in the laboratory

Manufacturer's technical support line.

If technical support from the manufacturer is needed, agree on precise appointment times, be there in order to learn from the support personnel, or alternatively, to challenge them. Brief them thoroughly, and have all of the activities recorded for further reference. Check on warranty times of parts and on proper invoicing.

With analytical thinking, and the following good practices, most of the problems can be solved:

Describe the problem in as unprejudiced a way as possible.

Observation of a problem should be confirmed twice before taking further action.

Operate the instrument under reference conditions, if possible.

Take notes of every change you make while trying to solve the problem.

Label all removed parts.

Make only one change at a time.

Only use replacement parts that have been proven not to be faulty or worn out.

Change a part back, if it did not resolve the problem. This does not apply if a risk exists to cause further damage to parts when they are replaced, if parts are inexpensive, if reinstallation creates the risk of damaging a module, or if parts are scheduled for replacement anyway. Document problem solution.

TERMINOLOGY

In this chapter, the nomenclature follows that proposed for the chromatographic sciences by the IUPAC recommendations (Ettre, 1994).

SUMMARY

In this overview, a variety of protocols and instrumental considerations have been introduced, and various factors involved in the correct selection of experimental procedures have been described. This information represents a starting point for good laboratory practices for the resolution of complex mixtures of peptides and proteins by HPLC methods—not the ultimate solution. As a consequence, the various sections have been prepared with the novice practitioners in mind, rather than the expert chromatographic scientist or peptide/protein chemist with advanced chromatographic skills. These introductory sections should thus facilitate the successful development by biologists and other life scientists with little formal training in the modern separation sciences, but who nevertheless face numerous experimental challenges requiring the high-resolution separation of complex mixtures of peptides and proteins by HPLC techniques.

LITERATURE CITED

- Aguilar, M.I., Hodder, A.N., and Hearn, M.T.W. 1985. Studies on the optimization of the reversed-phase gradient elution of polypeptides: Evaluation of retention relationships with β -endorphin related polypeptides. *J. Chromatogr.* 327:115-138.
- Alpert, A.J. 1990. Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. *J. Chromatogr.* 499:177-196.
- Anspach, B., Unger, K.K., Davies, J., and Hearn, M.T.W. 1988. Affinity chromatography with triazine dyes immobilized onto activated non-porous monodisperse silicas. *J. Chromatogr.* 457:195-204.
- Antia, F.D., Fellegvari, I., and Horvath, C. 1995. Displacement of proteins in hydrophobic interaction chromatography. *Ind. Engin. Chem. Res.* 34:2796-2804.

- Ballschmiter, K. and Wossner, M. 1998. Recent developments in adsorption liquid chromatography (NP-HPLC): A review. *Fresenius J. Anal. Chem.* 361:743-755.
- Berridge, J.C. 1986. Techniques for the automated optimization of HPLC separations. Wiley Interscience, Chichester.
- Boysen, R.I., Erdmann, V.A., and Hearn, M.T.W. 1998. Systematic, computer-assisted optimization of the isolation of *Thermus thermophilus* 50S ribosomal proteins by reversed-phase high-performance liquid chromatography. *J. Biochem. Biophys. Methods* 37:69-89.
- Boysen, R.I., Wang, Y., Keah, H.H., and Hearn, M.T.W. 1999. Observations on the origin of the non-linear van't Hoff behaviour of polypeptides in hydrophobic environments. *J. Biophys. Chem.* 77:79-97.
- Buchholz, K., Goedelmann, B., and Molnar, I. 1982. High-performance liquid chromatography of proteins: Analytical applications. *J. Chromatogr.* 238:193-202.
- Burke, T.W., Mant, C.T., Black, J.A., and Hodges, R.S. 1989. Strong cation-exchange high-performance liquid chromatography of peptides. Effect of non-specific hydrophobic interactions and linearization of peptide retention behaviour. *J. Chromatogr.* 476:377-389.
- Chang, S., Noel, R., and Regnier, F.E. 1976. High speed ion exchange chromatography of proteins. *Anal. Chem.* 48:1839-1845.
- Chaouk, H. and Hearn, M.T.W. 1999. New ligand, *N*-(2-pyridylmethyl)aminoacetate, for use in the immobilised metal ion affinity chromatographic separation of proteins. *J. Chromatogr. A* 852:105-115.
- Dolan, J.W. 1991. Preventive maintenance and troubleshooting LC instrumentation. In *High Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis and Conformation* (C.T. Mant and R.S. Hodges, eds.) pp. 23-29. CRC Press, Boca Raton, Fla.
- Dolan, J.W. and Snyder, L.R. 1989. Troubleshooting LC Systems pp. 1-515. Humana Press, Clifton, N.J.
- Dolan, J.W., Lommen, D.C., and Snyder, L.R. 1989. DryLab computer simulation for high-performance liquid chromatographic method development. II. Gradient elution. *J. Chromatogr.* 485:91-112.
- Ettre, L.S. 1994. New, unified nomenclature for chromatography. *Chromatographia* 38:521-526.
- Fausnaugh, J.L. and Regnier, F.E. 1986. Solute and mobile phase contributions to retention in hydrophobic interaction chromatography of proteins. *J. Chromatogr.* 359:131-146.
- Fausnaugh, J.L., Pfannkoch, E., Gupta, S., and Regnier, F.E. 1984. High-performance hydrophobic interaction chromatography of proteins. *Anal. Biochem.* 137:464-472.
- Freebairn, K.W. and Knox, J.H. 1984. Dispersion measurements on conventional and miniaturized HPLC systems. *J. Chromatogr.* 19:37-47.
- Ghrist, B.F.D. and Snyder, L.R. 1988a. Design of optimized high-performance liquid chromatographic gradients for the separation of either small or large molecules. II. Background and theory. *J. Chromatogr.* 459:25-41.
- Ghrist, B.F.D. and Snyder, L.R. 1988b. Design of optimized high-performance liquid chromatographic gradients for the separation of either small or large molecules. III. An overall strategy and its application to several examples. *J. Chromatogr.* 459:43-63.
- Ghrist, B.F.D., Coopermann, B.S., and Snyder, L.R. 1988. Design of optimized high-performance liquid chromatographic gradients for the separation of either small or large molecules. Minimising errors in computer simulations. *J. Chromatogr.* 459:1-23.
- Glajch, J.L., Quarry, M.A., Vasta, J.F., and Snyder, L.R. 1986. Separation of peptide mixtures by reversed-phase gradient elution. Use of flow rate changes for controlling band spacing and improving resolution. *Anal. Chem.* 58:280-285.
- Goldberg, A.P., Nowakowska, A.E., Antle, P.E., and Snyder, L.R. 1984. Retention-optimization strategy for the high-performance liquid chromatographic ion-pair separation of samples containing basic compounds. *J. Chromatogr.* 316:241-260.
- Gooding, D.L., Schmuck, M.N., Nowlan, M.P., and Gooding, K.M. 1986. Optimization of preparative hydrophobic interaction chromatographic purification methods. *J. Chromatogr.* 359:331-337.
- Haidacher, D., Vailaya, A., and Horvath, C. 1996. Temperature effects in hydrophobic interaction chromatography. *Proc. Nat. Acad. Sci. U.S.A.* 93:2290-2295.
- Hancock, W.S., Bishop, C.A., Prestidge, R.L., Harding, D.R., and Hearn, M.T.W. 1978. Reversed-phase, high-pressure liquid chromatography of peptides and proteins with ion-pairing reagents. *Science* 200:1168-1170.
- Hearn, M.T.W. 1991a. High-performance liquid chromatography of peptides and proteins: General principles and basic theory. In *High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis and Conformation* (C.T. Mant and R.S. Hodges, eds.) pp. 95-104. CRC Press, Boca Raton, Fla.
- Hearn, M.T.W. 1991b. High-performance liquid chromatography of peptides and proteins: Quantitative relationships for isocratic and gradient elution of peptides and proteins. In *High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis and Conformation* (C.T. Mant and R.S. Hodges, eds.) pp. 105-122. CRC Press, Boca Raton, Fla.
- Hearn, M.T.W. 2000. Physico-chemical factors in polypeptide and protein purification and analysis by high-performance liquid chromatographic techniques: Current status and future challenges.

- In Handbook of Bioseparation* (S. Ahuja, ed.) pp. 72-235. Academic Press, New York.
- Hearn, M.T.W. 2001a. Reversed-phase and hydrophobic interaction chromatography of peptides and proteins. *In HPLC of Biopolymers* (K. Gooding and F.E. Regnier, eds.). Marcel Dekker, New York.
- Hearn, M.T.W. 2001b. Conformational behaviour of polypeptides and proteins in reversed-phase environments. *In Theory and Practice of Biochromatography* (M.A. Vijayalakshmi, ed.) pp. 72-235. Harwood Academic Publishers, Chur, Switzerland.
- Hearn, M.T.W. and Grego, B. 1981. Organic solvent modifier effects in the separation of unprotected peptides by reversed-phase liquid chromatography. *J. Chromatogr.* 218:497-507.
- Hearn, M.T.W. and Grego, B. 1983. Further studies on the role of the organic modifier in reversed-phase high-performance liquid chromatography of polypeptides. Implications for gradient optimization. *J. Chromatogr.* 255:125-136.
- Hearn, M.T.W. and Zhao, G.L. 1999. Investigations into the thermodynamics of polypeptide interaction with nonpolar ligands. *Anal. Chem.* 71:4874-4885.
- Hearn, M.T., Grego, B., and Hancock, W.S. 1979. Investigation of the effects of pH and ion-pair formation on the retention of peptides on chemically-bonded hydrocarbonaceous stationary phases. *J. Chromatogr.* 185:429-444.
- Hearn, M.T.W., Hodder, A.N., and Aguilar, M.I. 1988. Comparison of retention and bandwidth properties of proteins eluted by gradient and isocratic anion-exchange chromatography. *J. Chromatogr.* 458:27-44.
- Hearn, M.T.W., Middleton, S., Jackson, R.W., and Chaouk, W. 1997. Constrained immobilised metal ion affinity absorbents for protein purification. *Int. J. Biochem.* 2:153-190.
- Heinitz, M.L., Kennedy, L., Kopaciewicz, W., and Regnier, F.E. 1988. Chromatography of proteins on hydrophobic interaction and ion-exchange chromatographic matrices: Mobile phase contributions to selectivity. *J. Chromatogr.* 443:173-182.
- Hodder, A.N., Aguilar, M.I., and Hearn, M.T.W. 1990. The influence of the gradient elution mode and displacer salt type on the retention properties of closely related protein variants separated by high-performance anion-exchange chromatography. *J. Chromatogr.* 506:17-34.
- Horvath, C., Melander, W., and Molnar, I. 1976. Solvophobic interactions in liquid chromatography with nonpolar stationary phases. *J. Chromatogr.* 125:129-156.
- Horvath, C., Melander, W., and Molnar, I. 1977. Liquid chromatography of ionogenic substances with nonpolar stationary phases. *Anal. Chem.* 49:142-154.
- Hulpke, H. and Werthmann, U. 1986. Column switching. *In Practice of High Performance Liquid Chromatography* (H. Engelhardt, ed.) pp. 147. Springer Verlag, New York.
- Hupe, K.P., Jonker, R.J., and Roing, G. 1984. Determination of band-spreading effects in high-performance liquid chromatographic instruments. *J. Chromatogr.* 285:253-265.
- Jeno, P., Scherer, P.E., Manninkrieg, U., and Horst, M. 1993. Desalting electroeluted proteins with hydrophilic interaction chromatography. *Anal. Biochem.* 215:292-298.
- Jiang, W., Graham, B., Spiccia, L., and Hearn, M.T.W. 1998. Protein selectivity with immobilized metal ion tacn sorbents: Chromatographic studies with human serum proteins and several other globular proteins. *Anal. Biochem.* 255:47-58.
- Kato, Y., Komiya, K., and Hishimoto, T. 1982. Study of experimental conditions in high-performance ion-exchange chromatography. *J. Chromatogr.* 46:13-22.
- Kato, Y., Kitamura, T., and Hashimoto, T. 1983. High-performance hydrophobic interaction chromatography of proteins. *J. Chromatogr.* 266:49-54.
- Knox, J.H. and Scott, H.P. 1983. B and C terms in the van Deemter equation for liquid chromatography. *J. Chromatogr.* 282:297-313.
- Kopaciewicz, W. and Regnier, F.E. 1983a. Mobile phase selection for the high-performance ion-exchange chromatography of proteins. *Anal. Biochem.* 133:251-259.
- Kopaciewicz, W. and Regnier, F.E. 1983b. A system for coupled multiple-column separation of proteins. *Anal. Biochem.* 129:472-482.
- Kopaciewicz, W. and Regnier, F.E. 1986. Synthesis of cation-exchange stationary phases using an adsorbed polymeric coating. *J. Chromatogr.* 358:107-117.
- Kopaciewicz, W., Rounds, M.A., and Regnier, F.E. 1985. Stationary phase contributions to retention in high-performance anion-exchange protein chromatography: Ligand density and mixed mode effects. *J. Chromatogr.* 318:157-172.
- Lankmayr, E.P., Wegscheider, W., and Budna, K.W. 1989. Global optimization of HPLC separations. *J. Liq. Chromatogr.* 12:35-58.
- Litowski, J.R., Semchuk, P.D., Mant, C.T., and Hodges, R.S. 1999. Hydrophilic interaction/cation-exchange chromatography for the purification of synthetic peptides from closely related impurities: Serine side-chain acetylated peptides. *J. Peptide Res.* 54:1-11.
- Mant, C.T. and Hodges, R.S. 1985. Separation of peptides by strong cation-exchange high-performance liquid chromatography. *J. Chromatogr.* 327:147-155.
- Mant, C.T. and Hodges, R.S. 1991a. Requirement for peptide standards to monitor ideal and non-ideal size exclusion chromatography. *In High Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis and Conformation* (C.T. Mant and R.S. Hodges, eds.) pp. 125-134. CRC Press, Boca Raton, Fla.

- Mant, C.T. and Hodges, R.S. 1991b. The use of peptide standards for monitoring ideal and non-ideal behaviour in cation-exchange chromatography. *In High Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis and Conformation* (C.T. Mant and R.S. Hodges, eds.) pp. 171-185. CRC Press, Boca Raton, Fla.
- Mant, C.T. and Hodges, R.S. 1991c. The effects of anionic ion-pairing reagents on the peptide retention in reversed-phase chromatography. *In High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis and Conformation*, Vol. 327-341 (C.T. Mant and R.S. Hodges, eds.). CRC Press, Boca Raton, Fla.
- Mant, C.T., Parker, J.M., and Hodges, R.S. 1987. Size-exclusion high-performance liquid chromatography of peptides. Requirement for peptide standards to monitor column performance and non-ideal behaviour. *J. Chromatogr.* 397:99-112.
- Mant, C.T., Litowski, J.R., and Hodges, R.S. 1998a. Hydrophilic interaction/cation-exchange chromatography for separation of amphipathic alpha-helical peptides. *J. Chromatogr.* 816:65-78.
- Mant, C.T., Kondejewski, L.H., and Hodges, R.S. 1998b. Hydrophilic interaction/cation-exchange chromatography for separation of cyclic peptides. *J. Chromatogr.* 816:79-88.
- Melander, W.R., Corradini, D., and Horvath, C. 1984. Salt-mediated retention of proteins in hydrophobic-interaction chromatography. Application of solvophobic theory. *J. Chromatogr.* 317:67-85.
- Melander, W.R., el Rassi, Z., and Horvath, C. 1989. Interplay of hydrophobic and electrostatic interactions in biopolymer chromatography. Effect of salts on the retention of proteins. *J. Chromatogr.* 469:3-27.
- Molnar, I., Boysen, R.I., and Jekow, P. 1989a. Peak tracking in high-performance liquid chromatography based on normalized band areas. A ribosomal protein sample as an example. *J. Chromatogr.* 485:569-579.
- Molnar, I., Boysen, R.I., and Erdmann, V.A. 1989b. Reversed-phase high-performance liquid chromatography of *Thermus aquaticus* 50S and 30S ribosomal proteins. *Chromatographia* 28:39-44.
- Ohlson, S., Hansson, L., Larsson, P.O., and Mosbach, K. 1978. High-performance liquid affinity chromatography (HPLAC) and its application to the separation of enzymes and antigens. *FEBS Letts.* 93:5-9.
- Papadoyannis, I.N., Zotou, A.C., and Samanidou, V.F. 1995. Solid-phase extraction study and RP-HPLC analysis of lamotrigine in human biological fluids and in antiepileptic tablet formulations. *J. Liq. Chromatogr.* 18:2593-2609.
- Patel, H.B. and Jefferies, T.M. 1987. Eluotropic strength of solvents. Prediction and use in reversed-phase high-performance liquid chromatography. *J. Chromatogr.* 389:21-32.
- Pohl, T. and Kamp, R.M. 1987. Desalting and concentration of proteins in dilute solution using reversed-phase high-performance-liquid-chromatography. *Anal. Biochem.* 160:388-391.
- Porath, J. 1988. High-performance immobilized-metal-ion affinity chromatography of peptides and proteins. *J. Chromatogr.* 443:3-11.
- Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. 1975. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* 258:598-599.
- Purcell, A.W., Aguilar, M.I., and Hearn, M.T.W. 1993. Dynamics of peptides in reversed-phase high-performance liquid chromatography. *Anal. Chem.* 65:3038-3047.
- Quarry, M.A., Grob, R.L., and Snyder, L.R. 1986. Prediction of precise isocratic retention data from two or more gradient elution runs. Analysis of some associated errors. *Anal. Chem.* 58:907-917.
- Regnier, F.E. 1983. High-performance liquid chromatography of biopolymers. *Science* 222:245-252.
- Regnier, F.E. 1984. High-performance ion-exchange chromatography. *Methods Enzymol.* 104:170-189.
- Risley, D.S. and Stregge, M.A. 2000. Chiral separations of polar compounds by hydrophilic interaction chromatography with evaporative light scattering detection. *Anal. Chem.* 72:1736-1739.
- Round, A.J., Aguilar, M.I., and Hearn, M.T.W. 1994. Peak tracking of peptides in reversed-phase chromatography. *J. Chromatogr.* 661:61-75.
- Schoenmakers, P.J., Biliot, H.A.H., and Galan, L.D. 1979. Influence of organic modifiers on the retention behaviour in reversed-phase liquid chromatography and its consequences for gradient elution. *J. Chromatogr.* 185:179-195.
- Snyder, L.R. 1970. Adsorption chromatography: Scope, technique and equipment. *Methods Med. Res.* 12:11-36.
- Snyder, L.R. 1974. Classification of the solvent properties of common liquids. *J. Chromatogr., Chromatogr.* 92:223-230.
- Snyder, L.R. 1978. Classification of the solvent properties of common liquid solvents. *J. Chromatogr. Sci.* 16:223-234.
- Snyder, L.R. 1980. Gradient elution. *In HPLC: Advances and Perspectives*, Vol. 1 (C. Horvath, ed.) pp. 208-316. Academic Press, New York.
- Snyder, L.R. 1990. Gradient elution separation of large biomolecules. *In HPLC of Biological Macromolecules. Methods and Application* (K.M. Gooding and F.E. Regnier, eds.) pp. 231-259. Marcel Dekker, New York.
- Snyder, L.R., Stadalius, M.A., and Quarry, M.A. 1983. Gradient elution in reversed-phase HPLC separation of macromolecules. *Anal. Chem.* 55:1413-1430.
- Snyder, L.R., Glajch, J.L., and Kirkland, J.J. 1988. Practical HPLC Method Development pp. 1-260. Wiley Interscience, New York.

- Stadalius, M.A., Gold, H.S., and Snyder, L.R. 1984. Optimization model for the gradient elution separation of peptide mixtures by reversed-phase high-performance liquid chromatography. Verification of retention relationships. *J. Chromatogr.* 296:31-59.
- Stadalius, M.A., Ghrist, B.F., and Snyder, L.R. 1987. Predicting bandwidth in the high-performance liquid chromatographic separation of large biomolecules. II. A general model for the four common high-performance liquid chromatography methods. *J. Chromatogr.* 387:21-40.
- Stout, R.W., DeStefano, J.J., and Snyder, L.R. 1983. High-performance liquid chromatographic column efficiency as a function of particle composition and geometry and capacity factor. *J. Chromatogr.* 282:263-286.
- Strasters, J.K., Billiet, H.A.H., de Galan, L., Vandeginste, B.G.M., and Kateman, G. 1989. Automated peak recognition from photodiode array spectra in liquid chromatography. *J. Liq. Chromatogr.* 12:3-22.
- Strege, M.A. 1998. Hydrophilic interaction chromatography electrospray mass spectrometry analysis of polar compounds for natural product drug discovery. *Anal. Chem.* 70:2439-2445.
- Vailaya, A. and Horvath, C. 1996. Retention thermodynamics in hydrophobic interaction chromatography. *Ind. Engin. Chem. Res.* 35:2964-2981.
- Wilce, M.C.J., Aguilar, M.I., and Hearn, M.T.W. 1995. Physicochemical basis of amino acid hydrophobicity scales: Evaluation of four new scales of amino acid hydrophobicity coefficients derived from RP-HPLC of peptides. *Anal. Chem.* 67:1210-1219.
- Wirth, H.J. and Hearn, M.T.W. 1993. High-performance liquid chromatography of amino acids, peptides and proteins. CXXX. Modified porous zirconia as sorbents in affinity chromatography. *J. Chromatogr.* 646:143-151.
- Wu, S.L., Benedek, K., and Karger, B.L. 1986a. Thermal behavior of proteins in high-performance hydrophobic-interaction chromatography. On-line spectroscopic and chromatographic characterization. *J. Chromatogr.* 359:3-17.
- Wu, S.L., Figueroa, A., and Karger, B.L. 1986b. Protein conformational effects in hydrophobic interaction chromatography. Retention characterization and the role of mobile phase additives and stationary phase hydrophobicity. *J. Chromatogr.* 371:3-27.
- Yip, T.T., Nakagawa, Y., and Porath, J. 1989. Evaluation of the interaction of peptides with Cu(II), Ni(II), and Zn(II) by high-performance immobilized metal ion affinity chromatography. *Anal. Biochem.* 183:159-171.
- Yoshida, T. 1998. Calculation of peptide retention coefficients in normal-phase liquid chromatography. *J. Chromatogr.* 808:105-112.
- Zachariou, M., Traverso, I., and Hearn, M.T.W. 1993. *O*-phosphoserine as a new chelating ligand for use with hard Lewis metal ions in the immobilized-metal affinity chromatography of proteins. *J. Chromatogr.* 646:107-120.
- Zhang, J.F. and Wang, D.I.C. 1998. Quantitative analysis and process monitoring of site-specific glycosylation microheterogeneity in recombinant human interferon-gamma from Chinese hamster ovary cell culture by hydrophilic interaction chromatography. *J. Chromatogr. B* 712:73-82.

KEY REFERENCES

Dolan and Snyder, 1989. See above.

An excellent practical manual.

Glajch, J.L. and Snyder, L.R. 1990. Computer Assisted Method Development for High-Performance Liquid Chromatography. pp. 1-682. Elsevier Science Publishers, Amsterdam.

An advanced text for experienced scientists.

Hearn, M.T.W. 1991. HPLC of Peptides, Proteins and Polynucleotides, Fundamental Principles and Contemporary Applications. pp. 1-776. VCH Publ, Deerfield Beach, Fla.

A comprehensive text encompassing most applications.

Hearn, M.T.W. 2000. Physicochemical factors in polypeptide and protein purification by high-performance liquid chromatography: Current status and future challenges. *In Handbook of Bioseparation* (S. Ahuja ed.) pp. 1-652. Academic Press, New York.

Hearn, M.T.W. 2001. RPC and HIC of peptides and proteins. *In HPLC of Biological Macromolecules. Methods and Applications* (K.M. Gooding and F.E. Regnier, eds.). Marcel Dekker, New York.

A detailed survey of application and theory.

Mant, C.T. and Hodges, R.S. 1991. High-Performance Liquid Chromatography of Peptides and Proteins: Separation, Analysis and Conformation. pp. 1-938. CRC Press, Boca Raton, Fla.

A comprehensive survey of applications.

McMaster, M.C. 1994. HPLC: A Practical User's Guide. pp. 1-211. VCH Publishers, Inc., New York.

A general guide to HPLC.

Snyder et al., 1988. See above.

An excellent advanced text.

Vijayalakshmi, M.A. 2001. Theory and Practice of Biochromatography. pp. 1-721. Harwood Academic Publishers.

A useful guide to modern chromatography.

Contributed by Reinhard I. Boysen
and Milton T.W. Hearn
Monash University
Victoria, Australia

Reversed-Phase Isolation of Peptides

UNIT 10.14

Reversed-phase high-performance liquid chromatography (HPLC) is a fundamental tool for the isolation and analysis of peptides. Peptides are separated on a hydrophobic stationary phase and eluted with a gradient of increasing organic solvent concentration. Peptides in 5- to 500-pmol quantities are separated on a narrow-bore (2-mm-i.d.) or microbore (1-mm-i.d.) column (see Basic Protocol 1). Separation of peptides in quantities <5 pmol can be accomplished using capillary HPLC columns (see Basic Protocol 2). Capillary HPLC columns, however, require a gradient flow rate of 3 to 5 $\mu\text{l}/\text{min}$, which most current HPLC pumps cannot attain without modifications. A procedure is therefore provided for constructing a capillary HPLC system using readily available components (see Support Protocol).

HPLC peaks that appear to be symmetrical may actually contain coeluting peptides. Matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (see Basic Protocol 3 and Alternate Protocol; also see *UNITS 10.21 & 10.22*) and capillary electrophoresis (see Basic Protocol 4; also see *UNIT 10.20*) can be used to analyze a small portion of an HPLC fraction and to determine the number of components present in a small sample. These methods can be utilized to screen fractions prior to automated sequencing.

REVERSED-PHASE PEPTIDE SEPARATION AT THE 5- TO 500-PMOL LEVEL

**BASIC
PROTOCOL 1**

A 2-mm narrow-bore column requires a flow rate of 100 to 200 $\mu\text{l}/\text{min}$. Many commercial HPLC instruments are capable of providing flow rates in this range, although some have tubing with a large internal volume that may add excessive delay time to the separation. Some HPLC instruments that are suitable without modification for 1- and 2-mm columns are listed in the materials section below. Many older HPLC instruments can be utilized for 2-mm columns, but most cannot reproducibly deliver the 50- to 100- $\mu\text{l}/\text{min}$ flow rates required by 1-mm columns.

Materials

Mobile-phase solvent A: 0.1% (v/v) trifluoroacetic acid (TFA; Pierce) in Milli-Q water (TFA sold for protein sequencing may not be suitable because some manufacturers add antioxidants that can generate artifacts)

Mobile-phase solvent B: 0.07% to 0.1% (v/v) TFA (Pierce) in acetonitrile or 1- or 2-propanol (Burdick & Jackson or Baker; HPLC-grade)

Solvent modifier: TFA (Pierce)

Milli-Q grade water or equivalent (distilled water is not suitable)

Peptide sample

HPLC peptide standards, commercial (e.g., PE Biosystems) *or* tryptic digest (see Troubleshooting), for testing column

HPLC system (e.g., Hewlett-Packard HP-1090 liquid chromatograph; PE Biosystems model 170A; Michrom BioResources Ultrafast Microprotein Analyzer; Beckman System Gold; Waters Alliance System)

Detector flow cell (see recipe)

2-channel strip-chart recorder (Kipp & Zonen or equivalent)

C18, C8, or C4 reversed-phase columns, 300 Å, 1- or 2-mm i.d. (e.g., Micra Scientific or Vydac; many other columns from numerous manufacturers can be used)

**Analysis of
Proteins**

10.14.1

1. Prepare and degas the mobile phase and set up an appropriate gradient of mobile-phase solvents A and B. Select a gradient based on number of peaks expected and resolution desired.

Normal gradient: 0% to 70% solvent B in 70 min

Fast gradient: 0% to 70% solvent B in 35 min

Slow gradient: 0% to 70% solvent B in 90 to 140 min.

A normal, fast, or slow gradient should be used when the expected number of peaks is ~30 to 60, <30, or >60, respectively.

2. Program flow rate to 200 $\mu\text{l}/\text{min}$ or 100 $\mu\text{l}/\text{min}$ for 2-mm- or 1-mm-i.d. columns, respectively. Set detector to 0.1 and 0.02 AUFS and the wavelengths on chart recorder to 214 and 280 nm for channels 1 and 2, respectively. Equilibrate column with initial conditions and establish a flat baseline.

The 280-nm pen on the chart recorder should be set to a scale five times more sensitive than the 214-nm pen. If only one wavelength is available, use 214 nm.

3. Fill injector loop with starting solvent and run a blank gradient consisting of the same gradient as will be utilized for separation of sample.

This will wash the column, removing any peptides that may be present from previous injections.

4. Prepare sample for injection as follows (or prepare peptide standards, when using a new column or when troubleshooting):

- a. If the sample contains organic solvent, lower the organic solvent concentration to <10% (v/v) by evaporating it in a Speedvac evaporator or by diluting it with water or aqueous buffer.
- b. Check that the sample volume is appropriate for the volume of the injector sample loop; if not, either reduce the sample volume by evaporation or change to a larger sample loop.
- c. Spin the sample in a microcentrifuge to remove particulates prior to injection.

The presence of organic solvent in the sample may prevent peptides from adsorbing to the stationary phase.

The size of the sample loop will not have a significant effect on the separation. However, if a larger loop is used, an appropriate delay should be added to the gradient program to allow the entire sample volume to load on the column and to allow salts that may be in the sample to wash off the column before the gradient begins.

5. Change injector to load position and rinse injector with starting solvent (mobile-phase solvent A).

If the sample loop was switched from inject to load position during the gradient run, the sample loop will contain organic solvent. If the sample only partially fills the sample loop, the presence of organic solvent in the loop might prevent the sample from binding to the column; rinsing with starting solvent averts this.

6. Inject sample, being careful to avoid injecting air into sample loop.

Air that is injected onto the column can produce air bubbles that may lodge in the detector.

7. Prepare a rack with 1.5-ml microcentrifuge tubes for collecting peak fractions. Number tubes with felt-tipped pen.

8. Calculate the delay for peak collection by dividing the volume of tubing exiting the flow cell by the flow rate.

Table 10.14.1 Volumes of Tubing Exiting Flow Cell

Tubing inner diameter	Volume ($\mu\text{l}/\text{cm}$)
0.075-mm (75- μm) FSC	0.045
0.005-in. (127- μm) PEEK	0.127
0.007-in. (178- μm) PEEK	0.249

This is particularly important to prevent miscollection of peaks with low flow rates (<200 $\mu\text{l}/\text{min}$). Table 10.14.1 provides a list of volumes for different diameters of PEEK and FSC tubing.

Powder-free gloves should be worn during peak collection to avoid sample contamination by free amino acids.

9. Collect peaks by monitoring absorbance change on a strip-chart recorder and allowing appropriate delay time (calculated in step 8) before changing tubes.

A stopwatch can be used to accurately measure the delay time when collecting poorly resolved peaks. Data systems on most HPLC systems have signal processing delays that can make manual peak collection difficult. Monitoring a strip-chart recorder provides a real-time measurement of absorbance and eliminates this problem.

REVERSED-PHASE PEPTIDE SEPARATION AT ≤ 5 PMOL

Separation of peptides that are present in quantities ≤ 5 pmol requires the use of a capillary column. These columns have internal diameters <1 mm and provide high-resolution peptide mapping of peptides collected from in gel or in situ digests of proteins electroblotted from one- and two-dimensional gels. A 0.32-mm-i.d. capillary column is capable of separating peptides at the 500-fmol level, but requires a flow rate of 3 to 5 $\mu\text{l}/\text{min}$. Several commercial HPLC systems are currently available that are capable of delivering gradient flow rates in this range; other HPLC pumps can be adapted to deliver capillary gradient flow rates using a flow splitter (see Support Protocol).

Materials

Solvent A: 0.1% (v/v) trifluoroacetic acid (TFA) in Milli-Q water

Solvent B: 0.05% to 0.1% (v/v) TFA in acetonitrile

Peptide sample

Detector flow cell (LC Packings)

Strip-chart recorder (Kipp & Zonen or equivalent)

Capillary HPLC system (see Support Protocol)

Graduated 10- μl Hamilton glass syringe connected to outlet of flow cell with 0.25-mm-i.d. Teflon tubing (LC Packings)

Two stopwatches

Additional reagents and equipment for reversed-phase peptide separation at 5 to 500 pmol (see Basic Protocol 1) and capillary HPLC system assembly (see Support Protocol)

1. Prepare and degas the mobile phase (see Basic Protocol 1, step 1).
2. Set detector to 0.1 AUFS and the wavelength on the chart recorder to 195 nm.
3. Equilibrate capillary column in solvent A until a flat baseline is established. Connect a graduated 10- μl Hamilton glass syringe to the outlet tubing of capillary flow cell

BASIC PROTOCOL 2

with 0.25-mm-i.d. Teflon tubing. Measure the time it takes for liquid to reach a given volume and calculate flow rate as follows: flow rate (in $\mu\text{l}/\text{min}$) = volume/time.

4. Accurately measure the length of tubing from the flow cell to its outlet end and calculate the delay according to the following formula: delay (min) = volume of outlet tubing (μl)/flow rate ($\mu\text{l}/\text{min}$).

See Table 10.14.1 for a list of tubing internal volumes. A typical delay time is 35 sec.

5. If necessary, adjust the flow rate to between 3 and 5 $\mu\text{l}/\text{min}$ (see Support Protocol, step 1).
6. Run blank gradient, prepare sample, and inject (see Basic Protocol 1, steps 3 to 6).
7. Collect peaks in 0.5-ml microcentrifuge tubes using two stopwatches. Mark the start of the peak with one stopwatch and time the end of the peak with the second watch.

Peak collection is the most difficult aspect of preparative capillary HPLC and requires a little practice. However, the degree of peak purity will be surprisingly high for closely eluting peaks if the delay time is calculated accurately. With Z-shaped flow cells, the flow cell is a length of fused silica capillary tubing. Very minor resolution loss occurs as the peaks travel from the end of the column to the outlet of the detector tubing.

Only one stopwatch will be needed if the capillary flow cell is modified as described for capillary HPLC assembly (see Support Protocol, step 3b).

SUPPORT PROTOCOL

CAPILLARY HPLC SYSTEM ASSEMBLY

Peptide separations at the low picomole level require columns with an i.d. of <1 mm, which are called capillary columns. A 0.32-mm-i.d. capillary column is capable of separating peptides at the 500-fmol level. However, these columns require a flow rate of 3 to 5 $\mu\text{l}/\text{min}$, and at present only a few commercial HPLC systems can deliver gradient flow rates in this range without modifications. Using the material listed below, it is quite easy to construct or modify an existing HPLC using readily available components. PE Biosystems, Waters Associates, MicroTech Scientific, and Michrom Bioresources have capillary HPLC pumps.

Materials

Capillary flow cell (LC Packings)

Piston- or syringe-pump splitter (LC Packings) *or* laboratory-constructed splitter:

1/16 in. tee (Valco); fingertight fittings; 0.012-in.-i.d. sleeve for capillary tubing; fingertight fitting for 0.012-in. sleeve (Upchurch Scientific); and 150 cm of 0.05-mm-i.d., 300-mm-o.d. FSC tubing (Polymicro Technologies) *or* 0.250-mm-i.d. Teflon tube inserted into a PEEK fingertight nut and ferrule

Capillary pump (PE Biosystems 140-D) *or* HPLC piston pump (see Basic Protocol 1) for use with an LC Packings splitter, *or* HPLC syringe pump (PE Biosystems models 120A, 140A, or 170A) for use with a laboratory-constructed splitter

Injector equipped with a 20- μl loop (Rheodyne)

In-line precolumn filter (Upchurch Scientific)

0.32-mm \times 15-cm C18 column (LC Packings, Keystone Scientific, Metachem Technologies, Micro-Tech Scientific, or Michrom Bioresources)

- 1a. *If using a capillary pump that can provide a flow rate of 3 to 5 $\mu\text{l}/\text{min}$:* Proceed to step 3a or 3b.
- 1b. *If assembling a splitter:* Connect a 150-cm length of 0.05-mm-i.d. FSC tubing to Valco tee with a 0.012-in.-i.d. sleeve *or* with a 0.250-mm-i.d. Teflon tube inserted

into a PEEK fingertight nut and ferrule. Connect the inlet side of splitter to HPLC pump with 0.005-in.-i.d. PEEK tubing.

A schematic diagram of the capillary fittings and the splitter assembly is shown in Figure 10.14.1.

A splitter must be connected to the HPLC pump to allow the flow rate to be reduced from 50 to 200 $\mu\text{l}/\text{min}$ to the 3 to 5 $\mu\text{l}/\text{min}$ required for a 0.32-mm-i.d. column. A splitter can be purchased commercially or assembled as described. The split ratio is controlled by varying

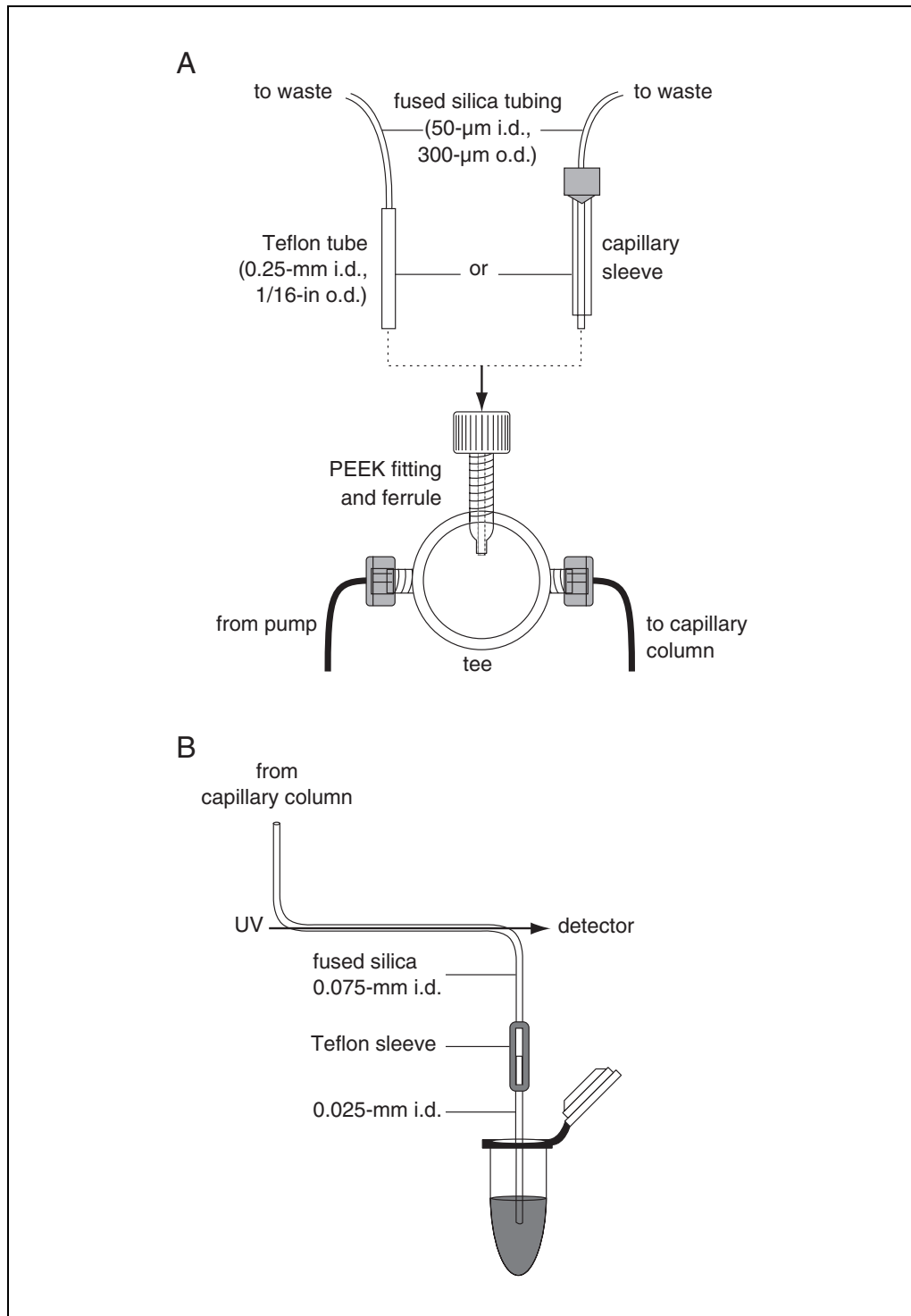


Figure 10.14.1 Modifications for capillary HPLC system. **(A)** Assembly of capillary fittings and splitter. **(B)** Modification of U or Z flow cell.

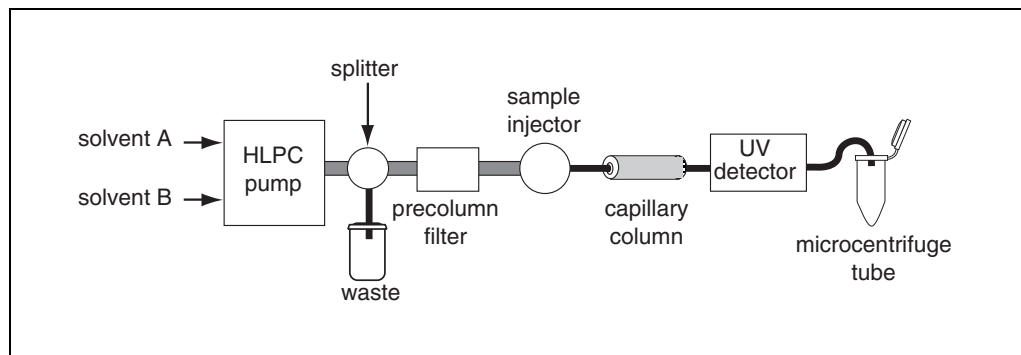


Figure 10.14.2 Capillary HPLC system.

the flow rate from the pump or by adjusting the length of the 0.05-mm-i.d. capillary tubing connected to the Valco tee. The pump flow rate should be set to the lowest flow that the pump can reliably deliver. A homemade splitter requires a pulse-free pump that can reliably deliver a flow rate of <math> < 200 \mu\text{l}/\text{min}</math>. PE Biosystems syringe pumps can accurately deliver a pulse-free flow rate of $70 \mu\text{l}/\text{min}</math>. These pumps have a dynamic mixer and a static mixer, which are removed to reduce the gradient delay caused by their large internal volumes. The high-pressure outlet of the mixing tee is connected directly to the splitter. Figure 10.14.2 shows a schematic diagram of the capillary HPLC system.$

2. Connect outlet of splitter to injector and connect the injector to the in-line precolumn filter with 0.005-in.-i.d. PEEK tubing.

The precolumn filter protects the capillary tubing and the column from clogging. The injector is fitted with a 20- μl sample loop. Any size sample loop can be used; however, a large loop will require a long wait for the sample to load on a 0.32-mm-i.d. column at a flow rate of 3 to 5 $\mu\text{l}/\text{min}</math>.$

- 3a. Install capillary flow cell.

The small peak volumes of capillary HPLC prevents the use of conventional flow cells. Instead, a length of capillary tubing with a short section of its polyamide coating removed is used as a flow cell. LC Packings has developed a capillary tube that is bent in the shape of a U or Z. The UV beam is aligned with the short axis of the U, increasing the path length of the flow cell. Although these flow cells have shorter path lengths than many conventional ones, the short path length and thin walls of the capillary tubing allow peptide detection at 195 nm.

- 3b. *Optional:* Modify the capillary flow cell (Fig. 10.14.1B) to reduce the delay volume by carefully cutting the 0.075-mm-i.d. outlet tubing of the flow cell, leaving ~1 to 2 cm adjacent to the Z cell (the fused silica tubing can be easily cut with a ceramic tubing cutter). Connect a 30-cm length of 0.025-mm-i.d., 0.280-mm-o.d. fused silica tubing with a 0.25-mm-i.d. Teflon sleeve to the remaining short outlet of the Z cell. The connection should be made inside the detector housing and the lengths of glass capillary tubing must be touching each other to avoid introducing any dead volume.

This modification eliminates the need for two stopwatches for peak collection. The flow cell modification reduces the delay volume to 0.45 μl , corresponding to a delay of 6 sec for a flow rate of 3.5 $\mu\text{l}/\text{min}</math>. The short delay greatly facilitates hand collection of HPLC fractions and the small-i.d. outlet tubing provides back-pressure that prevents bubble formation in the flow cell.$

4. Connect the 0.32-mm-i.d. C18 capillary column to the sample prefilter with the $\frac{1}{16}$ -in. PEEK fitting supplied with the column. Connect outlet tubing (75- μm -i.d.) of capillary column to capillary flow cell with a short length of 0.25-mm-i.d. Teflon tubing, also supplied with the column.

The 0.32-mm × 15-cm C18 column from LC Packings is the best general-purpose column for tryptic peptides. Keystone Scientific, Metachem Technologies, and Micro-Tech Scientific also sell capillary columns.

The Teflon tubing slips over the ends of the capillary tubing, creating a zero-dead-volume union between the column outlet and the detector union. Resolution will be significantly diminished if a gap is present between the adjacent pieces of capillary tubing.

5. Prepare the capillary HPLC system for sample run (see Basic Protocol 2).

PEPTIDE MAPPING BY MATRIX-ASSISTED LASER DESORPTION/IONIZATION (MALDI) MASS SPECTROMETRY

**BASIC
PROTOCOL 3**

Protein digests are mass analyzed either directly as an unfractionated mixture or as individual fractions from reversed-phase HPLC. The peptide-containing solution is mixed with a UV-absorbing matrix and applied to the sample stage of the mass spectrometer. Crystals of the peptide/matrix mixture form on the stage. This stage is inserted into the time-of-flight mass spectrometer vacuum system and irradiated with a UV laser to obtain a spectrum (UNITS 10.21 & 10.22). Greater mass accuracy is achieved by addition of known peptide(s) as internal mass standards. Alternate protocols are found in UNIT 16.3.

Materials

Peptide sample

30% (v/v) acetonitrile/0.1% (v/v) trifluoroacetic acid (TFA)

MALDI matrix solution (see recipe)

Mass standard solution (see recipe)

Time-of-flight mass spectrometer (e.g., PE Biosystems, Amersham Pharmacia Biotech, Micromass, Bruker, or Kratos)

1. Dilute samples in 30% acetonitrile/0.1% TFA to a final concentration of 0.05 to 5 pmol/μl.

Salt and buffer concentrations should be <50 mM. Most detergents, especially SDS, are not tolerated. If necessary, reduce the ionic strength by dilution, HPLC (UNITS 10.12 & 10.13), or size-exclusion chromatography (UNITS 10.9 or UNIT 10.13). If volatile buffers have been used, they may be removed by lyophilization and the dried sample redissolved in 2% TFA/50% acetonitrile.

2. Mix 2 μl of sample with 2 μl MALDI matrix solution in a 0.5-ml microcentrifuge tube.
3. Apply 1 to 2 μl of the mixture to the sample target, depending on the target size. Allow sample to dry completely and crystals to form without heating or applying vacuum.

Alternatively, 0.5 μl of sample solution and 0.5 μl of matrix solution may be mixed directly on the target. The solutions should be mixed well by repeatedly drawing the liquid into a pipettor.

The α-cyano-4-hydroxycinnamic acid (4HCCA) matrix solution works well for most peptides. An alternative is 2,5-dihydroxybenzoic acid (DHB). Another alternative is 3,5-dimethoxy-4-hydroxy-trans-cinnamic acid (sinapinic acid), but this produces a large background below 2000 Da and is not recommended for smaller peptides. Sinapinic acid or 4HCCA work best for larger peptides (>3000 Da). Due to preferential ionization of particular peptides in mixtures, it is advisable to obtain spectra for peptide mixtures with more than one matrix.

4. Insert sample into vacuum chamber of the mass spectrometer and follow the mass measurement procedure recommended by the instrument's manufacturer.

**Analysis of
Proteins**

10.14.7

Laser intensity (fluence) should be set just above the threshold for appearance of ion intensity. The observed signal (m/z) corresponds to $M+H^+$. Larger peptides may yield some $(M+2H)^{2+}$, which will appear at half the peptide mass ($M/2$). At high laser fluence, dimers may be formed as artifacts of the ionization process. Peptide mixtures may display preferential ionization for some peptides, and the relative peptide levels may vary from crystal to crystal. For this reason, spectra should be acquired from several different crystals on the target to obtain a representative display of peaks. The relative peak abundances in some cases (but not always!) give a semiquantitative image of the relative concentrations of peptides in the mixture, but they should not be relied upon for quantitation. Additional spectra acquired from the peptide mixture in a different matrix sometimes provide signals for peptides with suppressed ionization.

5. Calibrate the mass spectrometer with a mixture of peptides of known mass. Obtain calibrant peptide mass spectra as described in steps 2 to 4. Calibrate the instrument according to the manufacturer's recommendation.

An equimolar solution (~1 pmol/ μ l of each component) of des-Arg-bradykinin (e.g., Sigma) and ACTH-CLIP (18-39) provides two peaks that cover the masses of most peptides. The centroid of the peak corresponds to the isotopically averaged mass and should be used for calibration and mass assignment. The average masses ($M+H^+$) for des-Arg-bradykinin (905.1 Da) and ACTH-CLIP (18-39) (2466.7 Da) are generally reliable as calibrants. When sufficient resolution is available, the monoisotopic peaks should be used for calibration: des-Arg-bradykinin 904.468, ACTH-CLIP (18-39) 2465.199. Mass accuracy is best when the calibration peaks are close in mass and abundance to the unknown sample being analyzed. The laser fluence should be constant for the calibrant and unknown sample spectra.

- 6a. *If >500 fmol of unknown peptide sample is available:* Mix an equimolar concentration of des-Arg-bradykinin/ACTH-CLIP (18-39) standard solution with the unknown sample. Repeat mass analysis of sample with internal mass standards as described in steps 2 to 4.
- 6b. *If no additional unknown peptide sample is available:* Redissolve sample and matrix remaining on target after analysis in step 4 with 1 μ l of 2% TFA/50% acetonitrile. Add an equimolar concentration of mass standard solution. Allow mixed sample to dry on sample target and repeat mass analysis as described in step 4.

Internal mass calibration provides greater mass accuracy than external calibration. The amount of internal standard may need to be varied in several samples in order to obtain peak intensities that are approximately the same for the standard and the unknown sample. Avoid adding an excess of internal standard, which can suppress ionization of the unknown sample.

ALTERNATE PROTOCOL

PEPTIDE MAPPING BY MALDI MASS SPECTROMETRY USING THE MATRIX FAST-EVAPORATION METHOD

Preparation of a thin layer of matrix onto which the sample is applied offers advantages for small quantities of peptides and for contaminated samples. Matrix is applied to the sample stage in a highly volatile solvent that leaves a thin layer of matrix after evaporation. The uniform layer yields more uniform, intense signal across the sample. The sample is easily washed for efficient removal of contaminating salts and other buffer components that may interfere with ionization. Nitrocellulose is added to the matrix to provide greater peptide binding during the rinse step and to give greater structural support to the matrix, especially in the presence of higher concentrations of organic solvent that readily dissolve the matrix. This method is applicable to any type of MALDI instrument and gives superior results for a broad range of peptide and protein samples.

Additional Materials (also see Basic Protocol 3)

Fast-evaporation MALDI matrix solution (see recipe)
10% and 0.1% formic acid

1. Apply 0.5 to 2.0 μl fast-evaporation MALDI matrix solution to a clean sample target.

The solution must be applied quickly due to the high volatility of the acetone. A uniform, white surface 3 to 5 mm in diameter will form in 3 to 5 sec.

2. Apply 0.3 to 2.0 μl peptide sample to target surface. Allow sample to dry completely.

The matrix dissolves easily in basic solution or in the presence of a high concentration of organic solvent (e.g., >30% acetonitrile). If the peptide sample is not already acidic, apply 1.0 μl of 10% formic acid to the target before applying sample. For samples that contain high concentrations of organic solvent, apply 1.0 μl of water to the target before applying the sample.

3. Rinse sample with 4.0 μl cold 0.1% formic acid. After 10 sec, carefully remove the solution from the target surface with a pipet, or blow solution off target with a stream of nitrogen. Repeat rinse.

Rinse step may be eliminated if the peptide sample is free of salts. If the salt content of the sample is unknown, try to obtain signal without rinsing. If the results are unsatisfactory, remove the sample from the mass spectrometer and perform the rinse steps.

4. Insert sample into vacuum chamber of the mass spectrometer and follow the mass measurement procedure recommended by the instrument's manufacturer (see Basic Protocol 3).

Best results are obtained when samples are analyzed immediately. The authors have observed a decrease in signal intensity when samples are stored >24 hr. They have not observed any decrease in signal when samples are prepared according to Basic Protocol 3 using α -cyano-4-hydroxycinnamic acid (4HCCA).

5. Calibrate the mass spectrometer with a mixture of peptides of known mass (see Basic Protocol 3).

Because the peptide sample often covers only a part of the matrix spot, a small volume (0.1 to 0.3 μl) of the mass standard solution may be applied to the side of the matrix spot (after the peptide sample is dried on the other side of the matrix spot and rinsed) to provide a pseudo-internal standard. The mass spectrum of this small calibration sample may be taken immediately before or after the peptide spectrum to provide an external calibration with high accuracy. Alternatively, the signal from this sample may be included in the sample spectrum to provide internal calibration without the risk of ion suppression by the calibration compounds.

CAPILLARY ELECTROPHORESIS ANALYSIS

Capillary electrophoresis (also see *UNIT 10.20*) separates peptides by electrophoretic migration and electroosmotic flow using a high-voltage electric field in a fused silica capillary. The high voltage (20 to 25 kV) and the small internal volume of the fused silica capillary provide resolution that often exceeds 100,000 plates/meter. This high resolution combined with a different separation mechanism will often provide baseline resolution of peaks that coeluted during a reversed-phase HPLC separation. Very small volumes are required (10 to 200 nl), allowing the majority of the sample to be used for other methods of analysis or, if necessary, for repurification.

Materials

- 0.1 N NaOH
- Peptide sample
- 20 mM sodium phosphate, pH 2.5
- Capillary electrophoresis (CE) instrument
- 75- μ m-i.d. uncoated capillary column with separating length of 50 cm (from manufacturer of CE instrument or Polymicro Technologies)

1. Set detector wavelength to 200 nm.
If excessive baseline noise occurs at 200 nm, increase wavelength to 214 nm.
2. Wash capillary with 0.1 N NaOH for 1 min, followed by Milli-Q water for 1 to 2 min.
3. Equilibrate capillary with 20 mM sodium phosphate buffer, pH 2.5, for 5 min or until a stable baseline is established.

4. Inject sample using a pressure or vacuum injection of 5 to 30 sec.

*Some instruments inject the sample by creating a vacuum, whereas others apply a gas pressure to the sample reservoir. Both methods are capable of producing reproducible injections. Other instruments employ electrokinetic injection, in which an electric field is used to move the sample into the separating capillary. A drawback to this procedure is that the velocity of sample migration is dependent on the charge of the sample. This can result in a disproportionate injection of the sample components that have a higher charge. Also see *UNIT 10.20*.*

The peptide sample does not generally require any special preparation before injection.

5. Separate the components of the sample at 20 to 25 kV.
6. Repeat steps 2 to 5 for each additional sample. Replace capillary inlet and outlet buffer reservoir with fresh 20 mM sodium phosphate after every five runs.

Buffer ion depletion can affect the reproducibility of the separation.

REAGENTS AND SOLUTIONS

*Use Milli-Q-purified water or equivalent in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.*

Detector flow cell

A flow cell with an internal volume <12 μ l is required for 2-mm columns and with internal volume <5 μ l for 1-mm columns. Many of the newer HPLC systems designed for 2-mm columns are supplied with flow cells that have appropriately small (0.005-in.-i.d.) outlet tubing that allows peaks to be collected without loss of resolution caused by mixing. Older instruments designed only for 4.6-mm-i.d. columns, such as the Hewlett-Packard HP-1090, may need to be modified. This can easily be accomplished for most instruments by connecting 75- μ m-i.d. FSC tubing

continued

(Polymicro Technologies) or 0.005-in.-i.d. PEEK tubing (Upchurch Scientific) to the outlet of the flow cell. This tubing should be kept as short as possible to prevent loss of resolution from mixing and to minimize back-pressure on the flow cell (excessive back-pressure can result in a broken flow cell).

Fast-evaporation MALDI matrix solution

5 mg nitrocellulose

20 mg α -cyano-4-hydroxycinnamic acid (4HCCA)

Dissolve nitrocellulose and 4HCCA in 500 μ l acetone in a 1.5-ml microcentrifuge tube. Add 500 μ l isopropanol and mix completely. Store at -20°C 1 to 2 months. Keep the cap tightly sealed at all times—evaporation of the solvent is the main problem encountered in repeated use.

Before adding isopropanol, it may be necessary to vortex the material for 1 to 2 min to dissolve the nitrocellulose completely.

Use only pure nitrocellulose (e.g., from Schleicher & Schuell). Some nitrocellulose membranes contain other material supports.

MALDI matrix solution

Prepare in 30% (v/v) acetonitrile/0.1% (v/v) trifluoroacetic acid either:

1. 20 mg/ml α -cyano-4-hydroxycinnamic acid (use only the supernatant as matrix solution); *or*
2. 50 mg/ml 2,5-dihydroxybenzoic acid (DHB); *or*
3. Saturated solution of 3,5-dimethoxy-4-hydroxy-*trans*-cinnamic acid (sinapinic acid; use only the supernatant as matrix solution).

Trifluoroacetic acid from Pierce is recommended because some manufacturers add antioxidants that can generate artifacts.

The 4HCCA may yield better signal if it is further purified by recrystallization.

Mass standard solution

Dissolve 1 pmol/ml des-Arg-bradykinin (e.g., Sigma) and 2 pmol/ml ACTH-CLIP (18-39) in 30% (v/v) acetonitrile/0.1% (v/v) trifluoroacetic acid (Pierce).

COMMENTARY

Background Information

Reversed-phase HPLC has been widely utilized for analysis and preparative fractionation of peptides. Although many different buffers have been described in the literature, trifluoroacetic acid (TFA) has become the most widely used ion-pairing agent for reversed-phase chromatography of peptides. The volatility of TFA enables it to be compatible with many analytical methods including protein sequencing, on-line electrospray mass spectrometry, amino acid analysis, and capillary electrophoresis.

The development of columns with small internal diameters (2 and 1 mm) has greatly facilitated the ease of handling of small amounts (<200 pmol) of proteolytic digests of proteins (see Basic Protocol 1; Schlabach and Wilson, 1987). The use of 2-mm-i.d. columns generally results in resolution equivalent or superior to that obtained on a 4.6-mm-i.d. col-

umn, with significant cost savings in solvent purchase and disposal. Another advantage of smaller-i.d. columns is the increase in sensitivity they provide. The increase in detector signal is a function of the inverse square of the column radius. Reducing the column diameter from 4.6 to 1 mm results in a signal that is 20 times larger.

Microbore 1-mm-i.d. columns have been more problematic than 2-mm-i.d. columns. The inability of many HPLC pumps to deliver samples at a rate of 50 to 100 μ l/min reliably has prevented widespread use of these smaller columns. Many commercially packed 1-mm-i.d. columns have had low theoretical plate counts compared to 2-mm-i.d. columns. This is primarily due to difficulty in packing these columns in small stainless steel tubes.

The problems associated with packing 1-mm stainless steel columns are generally not observed with columns of diameters <1 mm.

These columns are packed into fused silica capillaries that have extremely smooth walls compared to stainless steel tubing. Capillary columns have a high theoretical plate count that often exceeds 100,000 plates per meter, resulting in higher resolution than can usually be achieved in a 1- or 2-mm-i.d. column. The high resolution and small internal diameters of capillary columns are ideal for high-sensitivity peptide mapping at the 5-pmol level and below (see Basic Protocol 2; Cobb and Novotny, 1989; Davis and Lee, 1992). Capillary columns can be fabricated in the laboratory using a slurry packing technique with an HPLC pump (Moritz and Simpson, 1992). Many different types and sizes of capillary columns are commercially available. Capillary columns of 0.8-mm i.d. can be used on some HPLC systems without using a splitter. These HPLC systems must be capable of delivering a flow rate of 25 to 50 $\mu\text{l}/\text{min}$ and be equipped with a micro flow cell ($<1 \mu\text{l}$) to prevent a loss of resolution caused by mixing in the flow cell. The short path-length of these flow cells can often outweigh any gain in sensitivity.

Replacing a micro flow cell with a 200- μm -i.d. U-shaped capillary cell can increase detector sensitivity. This type of capillary cell is available for a variety of UV detectors from LC Packings. An 0.32-mm-i.d. capillary column can provide high sensitivity and good recovery for analytical and micropreparative collection of peptides from in situ digests of proteins electroblotted on one- and two-dimensional gels (UNITS 10.2 & 10.3; Wong et al., 1993). A 0.32-mm-i.d. capillary column requires a capillary flow cell to prevent loss of resolution from mixing.

Although capillary flow cells typically have short path lengths, the thin walls of the capillary and the short path length of the detector cell allow transmission of UV light at 195 nm. The shorter wavelength allows higher-sensitivity peptide detection than can be obtained at 214 nm.

Peak collection of capillary HPLC fractions can be tedious. However, a significant advantage to this method is the high concentration present in the collection tube as a result of the small peak volumes (0.5 to 3 μl). The high concentration in the sample tube minimizes sample loss by adsorption to the tube walls. When a 5-pmol sample is diluted to 100 to 200 μl —the typical peak volume obtained from a separation on a 1- to 2-mm column—adsorption to the tube walls may occur.

Despite the small internal volume of capillary columns, the capacity can be relatively

large ($>1 \text{ nmol}$). Large-volume injections ($>1 \text{ ml}$) may require a wait of several hours for the sample to load on the column. The increased volume will not have a significant effect on resolution for peptide separations, however, because of the absorption/desorption mechanism of peptide retention. Another advantage of capillary HPLC is its compatibility with on-line electrospray ionization mass spectrometry (UNIT 10.21), a result of the low flow rates (3 to 5 $\mu\text{l}/\text{min}$) employed with these columns. On-line electrospray ionization mass spectrometry allows mass measurement of each component that elutes from the column and in some cases provides sequence information (Griffin et al., 1991; Huang and Henion, 1991; Hess et al., 1993).

Small-i.d. columns can provide high-resolution separations, but some peaks may still contain multiple components. MALDI mass spectrometry (see Basic Protocol 3) and capillary electrophoresis (see Basic Protocol 4) are rapid methods for analyzing peak purity. MALDI mass spectrometry (Hillenkamp et al., 1991; Beavis and Chait, 1996; UNIT 10.22) is capable of analyzing low-femtomole quantities of peptide mixtures or isolated peptide fractions (Billeci and Stults, 1993). For example, an in situ tryptic digest of 2 to 5 pmol of a PVDF-electroblotted protein may yield 1- to 2-pmol peptide fractions after enzymatic digestion and separation on a capillary HPLC column. Only a small aliquot (1% to 10%) of a 1-pmol peptide fraction is usually required to determine both purity and molecular weight by MALDI mass spectrometry (Henzel et al., 1993). This information can be useful in deciding which peptide fraction to sequence as well as deciding on the number of cycles with which to program the sequencer. The peptide mass is indispensable for corroborating sequence data and identifying post-translational modifications (Geromanos et al., 1994). Peptide masses may also be used for protein identification by peptide mass mapping (Henzel et al., 1993; Patterson and Aebersold, 1995; Gevaert and Vandekerckhove, 2000).

Capillary electrophoresis (CE; UNIT 10.20) can also provide an indication of peak purity. The electropherogram (chromatogram) produced by the CE instrument can be used to approximate the degree of purity of peptide fractions. This is in contrast to MALDI mass spectrometry, which is not quantitative and in which two components with the same concentration may result in significantly different signals. CE generally requires a peptide concentration of 10

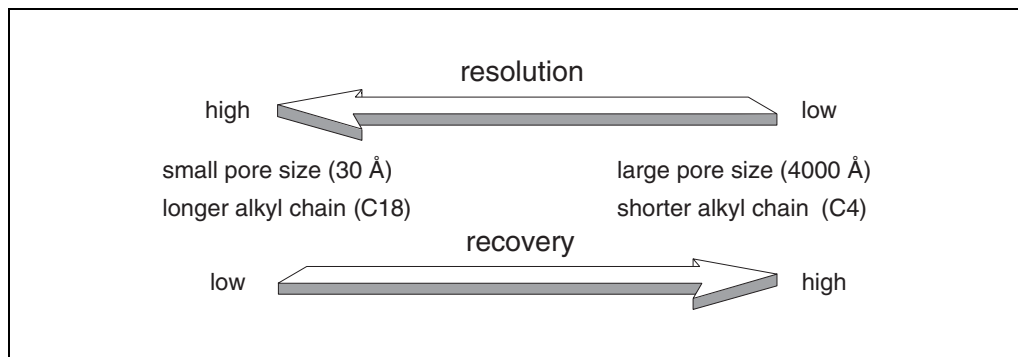


Figure 10.14.3 Effects of pore size and alkyl chain length on the resolution and recovery of peptides.

pmol/ μ l, which is 2 to 3 orders of magnitude larger than that required by MALDI mass spectrometry. Although some progress has been made in injecting larger volumes by concentrating the sample within the CE column (Aebersold and Morrison, 1990; Chien and Burgi, 1992; Figeys et al., 1996; Tomlinson et al., 1996), CE injection volumes are generally limited to <50 nl.

Critical Parameters

Selection of stationary phase

A large variety of reversed-phase stationary phases have been developed. The most commonly employed packings for peptide separation are based on porous silica derivatized with C4, C8, or C18. These columns are available with pore sizes ranging from 100 to 4000 Å. Columns of small particles containing small pores generally provide higher resolution than columns made of materials with larger pores, but peptide recovery is often lower. Peptides obtained from trypsin, chymotrypsin, or pepsin digests, which are often short (<15 residues), are preferably separated on a C18 column. Peptides from Lys-C and cyanogen bromide cleavages, which are often >15 residues long, are usually separated on a C4 column with 300- or 1000-Å pores.

Selection of mobile-phase solvent

Another important variable is the choice of the mobile phase. Mixtures of trifluoroacetic acid (TFA) in water with either acetonitrile or propanol as the organic solvent are the most widely utilized types of mobile phase for peptide separation. The choice of acetonitrile or propanol can affect peptide recovery and resolution. Acetonitrile will generally result in greater resolution than propanol; however, propanol usually provides higher peptide re-

covery. Complex peptide separations resulting from digestion of large proteins (>100 kDa) demand high-resolution separations and hence are best done using acetonitrile.

Increasing the length of the gradient can significantly increase resolution. A tryptic digest of a 100-kDa protein may require a 2- to 3-hr gradient. Peptides that fail to resolve in the course of an extended gradient can often be separated by repeating the chromatography using a different column (C4 or C8 in place of C18). The selectivity of the separation also can be affected by substituting a different organic solvent (i.e., propanol in place of acetonitrile) or by changing the pH of the separation. Rechromatography on the same column using a different organic modifier will often resolve coeluting peaks. Rechromatography can be repeated as many times as necessary but at the expense of some losses with each step. The sample must be diluted with water or concentrated in a Speedvac evaporator to lower the organic solvent concentration prior to rechromatography in order to allow the peptides to adsorb to the reversed-phase column. Figure 10.14.3 shows the relationship of pore size and alkyl chain length to resolution and peptide recovery.

Sample preparation and storage

Peptide fractions should be stored at 5°C in polypropylene tubes, not in polystyrene or glass. Prolonged storage can result in some evaporation of organic solvent from the sample, which may lead to sample adsorption to the microcentrifuge tube. Adding neat TFA to the sample tube (33% final concentration) can resolubilize peptides that may have precipitated or adsorbed to the tube walls (Erdjument-Bromage et al., 1993).

Solubility of peptides for MALDI can also be enhanced by addition of up to 50% acetonitrile.

trile and 2% TFA. Alternatively, up to 1% octyl- β -glucoside may assist in solubilization and is compatible with MALDI (Gharahdaghi et al., 1996). As fractions isolated by RP-HPLC with a volatile mobile phase are normally quite clean, no further cleanup of the sample is usually needed before analysis by MALDI. Preparation of samples for MALDI is easily done by taking a 0.3- to 0.6- μ l aliquot of an HPLC fraction straight from the fraction collection tube and mixing it with matrix directly on the sample target.

Mass measurement of unfractionated proteolytic digests is useful for rapidly confirming the sequence of known proteins (Billeci and Stults, 1993), for identifying proteins by searching a sequence database with peptide masses (Henzel et al., 1993; *UNIT 19.2*), and for assessing completeness of digestion. The digestion is preferably performed in a volatile buffer, such as ammonium bicarbonate, ammonium acetate, or *N*-ethylmorpholine. The buffer should be removed by vacuum centrifugation and the peptides solubilized in 50% acetonitrile/2% TFA. If nonvolatile buffer components are used, they must be diluted below 50 mM. Samples can be washed after application when the fast-evaporation matrix is used (Shevchenko et al., 1996; see Alternate Protocol). Analysis of subpicomole amounts of peptide may only be successful at substantially lower buffer concentrations. Buffer components may also be removed by pipet tip purification methods (Erdjument-Bromage et al., 1988; Gobom et al., 1999). Similar pipet tips are available as ZipTips from Millipore. Additional details on sample preparation are found in *UNITS 10.21 & 10.22*.

Special consideration for MALDI and CE

Mass accuracy for peptide measurement (1- to 3-kDa peptides) with a linear time-of-flight instrument is typically ± 1 to 2 Da with external calibration of the mass axis. Higher mass accuracy is achieved by measuring the sample with an internal mass standard, usually a pair of peptides of known mass with signal intensity approximately the same as the unknown sample. Repeated experiments may be necessary to achieve the proper ratio of standard to sample. With proper internal mass standards and an adequate signal-to-noise ratio, the mass accuracy can be $\pm 0.01\%$ (± 0.2 Da at m/z 2000). Higher mass accuracy and resolution may be obtained with reflector instruments (*UNIT 10.21*) or delayed extraction (Vestal et al., 1995).

The main variable in capillary electrophoresis of peptides is the choice of the separation buffer. Peptides that comigrate can often be separated by changing the pH of the buffer (Grossman et al., 1988). Changing the pH from 2.5 to 7.0 will often have a major effect on the separation. The ionic strength of the buffer can be increased for peptides that have slow migration rates. A different buffer can be tried when pH and ionic strength fail to improve the separation. Buffer additives can also be utilized to further enhance selectivity (Cobb and Novotny, 1992).

Troubleshooting

One of the main problems associated with reversed-phase HPLC is the occurrence of a drifting baseline. This is usually the result of contaminants in the mobile phase or inadequate mobile phase degassing. It is essential to use high-purity HPLC reagents and Milli-Q-quality water and to wear powder-free gloves for all manipulations. A baseline rise that increases with increasing organic modifier (solvent B) can be corrected by adjusting the amount of TFA added to the organic solvent. Titrating the amount of TFA in the solvents is especially important for capillary HPLC. In general a concentration of 0.05% to 0.08% TFA in solvent B will usually result in an acceptable baseline.

Baseline artifacts, which may appear as negative fluctuations in the baseline, can be caused by a number of HPLC pump malfunctions, including leaks (Dolan, 1991). It is essential to keep HPLC pumps in optimum condition for high-sensitivity HPLC. It is sometimes difficult to detect minor leaks, which can cause a major baseline disturbance during use of capillary HPLC columns. These leaks can usually be detected by placing a dry Kimwipe on the suspected fitting.

Occasionally the reversed-phase column can become contaminated by the sample, causing ghost peaks to appear. These are peaks that appear in blank runs when a gradient is run without injecting a sample. Sometimes several blank runs may be necessary before the ghost peaks completely disappear. Extended washing with 80% to 100% organic modifier can also be tried. Another method of cleaning columns, tubing, and flow cell is the use of methanol as a solvent. The HPLC is run in isocratic mode with 100% methanol in solvents A and B for 30 to 60 min. If all of the above methods have been tried without success, the column should be replaced.

Increased back-pressure (>500 p.s.i.) above the normal back-pressure of the column may indicate a clogged frit or a contaminated column. Use of a precolumn filter can prevent the column frit from clogging. The precolumn filter is easily replaced and relatively inexpensive. Column inlet frits can be replaced, but care must be exercised to prevent loss of column packing material. Capillary columns are particularly prone to clogging and should always be used in conjunction with a precolumn filter.

Peptide standards have been developed to monitor HPLC performance. Commercial standards (e.g., PE Biosystems) should be run when a new column is first utilized and when resolution appears to deteriorate (Mant and Hodges, 1990). Alternatively, a tryptic digest of a protein that results in well-resolved peptides on reversed-phase chromatography can be utilized to monitor column resolution. The digest should be solubilized in 0.1% TFA and stored frozen at -70°C between uses.

Reversed-phase columns should be stored in organic solvent that does not contain TFA, as TFA can slowly dissolve silica. Capillary columns will last significantly longer if this practice is followed.

Unsuccessful mass measurement by MALDI may be due to insufficient sample (<10 to 50 fmol). Sample loss is often a result of poor solubility of the peptide or the presence of contaminants that suppress ionization or crystal formation. Peptide solubility is aided by addition of acid, organic solvent, or octyl- β -glucoside. Alternative solvents such as formic acid or methanol may also assist in solubility and ionization (Cohen and Chait, 1996). In the case of an unfractionated digest or a nonvolatile buffer system, incompatible buffer components may be accommodated simply by dilution of the sample, if the peptide concentration permits, or by pipet tip purification. Alternatively, salts and buffers are often excluded from the less soluble peptide/matrix crystal, so a brief (~5-sec) wash of the sample target with cold water may remove many of the interfering salts.

CE instruments are relatively simple and are less problematic than HPLC equipment. The two major problems in CE are clogging of the capillary tubing and artifacts resulting from reagent or sample contaminants. Complete absence of current flow usually indicates a blocked capillary. The clog is often found on the sample inlet side and can be usually corrected by removing a short section (2 to 5 mm) of capillary from the column inlet end. The high sensitivity of CE requires high-purity reagents.

A number of chemical distributors now sell CE-purity reagents.

Anticipated Results

Reversed-phase HPLC should result in resolution and good recovery of the majority of the peptides (<30 residues) present in a typical proteolytic digest of a protein (<30 kDa). The use of capillary HPLC for tryptic peptides should allow recovery of peptides at the 1- to 2-pmol level. CE should be capable of separating peptides that coelute as a single peak on HPLC; however, the concentration must be at least 10 pmol/ μl . Peptides in the concentration range of 10 fmol/ μl to 50 pmol/ μl can be analyzed by MALDI mass spectrometry.

Time Considerations

Reversed-phase separations usually require 30 min to 1 hr per analysis. When additional blank gradients are run, several hours may be required. CE analysis of a purified peptide fraction may require from 15 to 60 min, depending on the charge and mass of the sample and the buffer utilized. Mass analysis by MALDI requires 1 to 2 min for preparation of each sample target, 5 to 10 min drying time, and 2 to 5 min per sample for data acquisition. Data processing may add an additional 5 to 10 min.

Literature Cited

- Aebersold, R. and Morrison, H.D. 1990. Analysis of dilute peptide samples by capillary zone electrophoresis. *J. Chromatogr.* 516:79-88.
- Beavis, R.C. and Chait, B.T. 1996. Matrix-assisted laser desorption ionization mass spectrometry of proteins. *Methods Enzymol.* 270:519-551.
- Billeci, T.M. and Stults, J.T. 1993. Tryptic mapping of recombinant proteins by matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.* 65:1709-1716.
- Chien, R.L. and Burgi, D.S. 1992. On-column sample concentration using field amplification in CZE. *Anal. Chem.* 64:489A-496A.
- Cobb, K.A. and Novotny, M. 1989. High-sensitivity peptide mapping by capillary zone electrophoresis and microcolumn liquid chromatography using immobilized trypsin for protein digestion. *Anal. Chem.* 61:2226-2231.
- Cobb, K.A. and Novotny, M. 1992. Peptide mapping of complex proteins at the low-picomole level with capillary electrophoretic separations. *Anal. Chem.* 64:879-886.
- Cohen, S.L. and Chait, B.T. 1996. Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. *Anal. Chem.* 68:31-37.

- Davis, M.T. and Lee, T.D. 1992. Analysis of peptide mixtures by capillary high performance liquid chromatography: A practical guide to small-scale separations. *Protein Sci.* 1:935-944.
- Dolan, J.W. 1991. Preventive maintenance and troubleshooting LC instruments. In *HPLC of Peptides and Proteins: Separation, Analysis, and Conformation* (C.T. Mant and R.S. Hodges, eds.) pp. 23-30. CRC Press, Boca Raton, Fla.
- Erdjument-Bromage, H., Geromanos, S., Chodera, A., and Tempst, P. 1993. Successful peptide sequencing with femtomole level PTH-analysis: A commentary. In *Techniques in Protein Chemistry IV* (R.H. Angeletti, ed.) pp. 419-426. Academic Press, San Diego.
- Erdjument-Bromage, H., Lui, M., Lacomis, L., Grewal, A., Annan, R.S., McNulty, D.E., Carr, S.A., and Tempst, P. 1998. Examination of micro-tip reversed-phase liquid chromatographic extraction of peptide pools for mass spectrometric analysis. *J. Chromatogr. A.* 826:167-181.
- Figeys, D., Ducret, A., Yates, J.R., III, and Aebersold, R. 1996. Protein identification by solid phase microextraction-capillary zone electrophoresis-microelectrospray-tandem mass spectrometry. *BioTechnology* 14:1579-1583.
- Geromanos, S., Casteels, P., Elicone, C., Powell, M., and Tempst, P. 1994. Combined Edman-chemical and laser-desorption mass spectrometric approaches to micro peptide sequencing: Optimization and applications. In *Techniques in Protein Chemistry V* (J.W. Crabb, ed.) pp. 143-150. Academic Press, San Diego.
- Gevaert, K. and Vandekerckhove, J. 2000. Protein identification methods in proteomics. *Electrophoresis* 21:1145-54.
- Gharahdaghi, F., Kirchner, M., Fernandez, J., and Mische, S.M. 1996. Peptide-mass profiles of polyvinylidene difluoride-bound proteins by matrix-assisted laser desorption ionization time-of-flight mass spectrometry in the presence of nonionic detergents. *Anal. Biochem.* 233:94-99.
- Gobom, J., Nordhoff, E., Mirgorodskaya, E., Ekman, R., and Roepstorff, P. 1999. Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* 34:105-16.
- Griffin, P.R., Coffman, J.A., Hood, L.E., and Yates, J.R., III. 1991. Structural analysis of proteins by capillary HPLC electrospray tandem mass spectrometry. *Int. J. Mass Spectrom. Ion Process.* 111:131-149.
- Grossman, P.D., Wilson, K.J., Petrie, G., and Laura, H.H. 1988. Effect of buffer pH and peptide composition on the selectivity of peptide separations by capillary zone electrophoresis. *Anal. Biochem.* 173:265-270.
- Henzel, W.J., Billeci, T.M., Stults, J.T., Wong, S.C., Grimley, C., and Watanabe, C. 1993. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc. Natl. Acad. Sci. U.S.A.* 90:5011-5015.
- Hess, D., Covey, T.C., Winz, R., Brownsey, R.W., and Aebersold, R. 1993. Analytical and micro-preparative peptide mapping by high performance liquid chromatography/electrospray mass spectrometry of proteins purified by gel electrophoresis. *Protein Sci.* 2:1342-1351.
- Hillenkamp, F., Karas, M., Beavis, R.C., and Chait, R.C. 1991. Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal. Chem.* 63:1193A-1203A.
- Huang, E.C. and Henion, J.D. 1991. Packed-capillary liquid chromatography/ion-spray tandem mass spectrometry determination of biomolecules. *Anal. Chem.* 63:732-739.
- Mant, C.T. and Hodges, R.S. 1990. HPLC of peptides. In *HPLC of Biological Macromolecules* (K.M. Gooding and F.E. Regnier, eds.) pp. 301-332. Marcel Dekker, New York.
- Moritz, R.L. and Simpson, R.J. 1992. Application of capillary reversed-phase high-performance liquid chromatography to high-sensitivity protein sequence analysis. *J. Chromatogr.* 599:119-130.
- Patterson, S.D. and Aebersold, R. 1995. Mass spectrometric approaches for the identification of gel-separated proteins. *Electrophoresis* 16:1791-814.
- Schlabach, T.D. and Wilson, K.J. 1987. Microbore flow-rates and protein chromatography. *J. Chromatogr.* 385:65-74.
- Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. 1996. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal. Chem.* 68:850-858.
- Tomlinson, A.J., Benson, L.M., Guzman, N.A., and Naylor, S. 1996. Preconcentration and micro-reaction technology on-line with capillary electrophoresis. *J. Chromatogr.* 744:3-15.
- Vestal, M.L., Juhasz, P., and Martin, S.A. 1995. Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 9:1044-1050.
- Wong, S.C., Grimley, C., Padua, A., Bourell, J., and Henzel, W.J. 1993. Peptide mapping of 2-D gel proteins by capillary HPLC. In *Techniques in Protein Chemistry IV* (R. H. Angeletti, ed.) pp. 371-378. Academic Press, San Diego.

Key References

Davis and Lee, 1992. See above.

Provides a guide to setting up and using a capillary HPLC.

Grossman, P.D. and Colburn, J.C. (eds.) 1992. *Capillary Electrophoresis—Theory and Practice.* Academic Press, San Diego.

Provides an introduction to capillary electrophoresis.

Mant, C.T. and Hodges, R.S. (eds.) 1991. HPLC of Peptides and Proteins: Separation, Analysis, and Conformation. CRC Press, Boca Raton, Fla.

A comprehensive guide to HPLC of peptides and proteins

Wilm, M. 2000. Mass spectrometric analysis of proteins. *Adv. Protein Chem.* 54:1-30.

A general description of mass spectrometers and their utility for solving problems in protein chemistry.

Contributed by William J. Henzel and
John T. Stults
Genentech, Inc.
South San Francisco, California

Purification of Recombinant Proteins and Study of Protein Interaction by Epitope Tagging

UNIT 10.15

A protein molecule can be engineered to include a short stretch of residues corresponding to an epitope to facilitate its subsequent biochemical and immunological analysis; a technique often referred to as “epitope tagging.” This is achieved by appending the sequence of the epitope tag to the protein-coding sequence in an appropriate expression vector. Several short epitope tags are currently available (Table 10.15.1) and have been widely used. Epitope-tagged proteins can be affinity purified using highly specific antibodies raised against various tags, which are also readily available. Basic Protocol 1 provides a general protocol for small-scale immunoprecipitation of epitope-tagged recombinant proteins expressed in transiently transfected mammalian cells. The immunoprecipitant can then be analyzed by SDS-PAGE (UNIT 10.2). Basic Protocol 2 provides an immunoprecipitation protocol that has been optimized for use with a baculovirus over-expression system.

In many cases, formation of a protein complex is required to execute, and thus to study, a particular biological function. Assembly of a functional complex *in vitro* offers a powerful tool for revealing the activities of the complex. Epitope tagging provides an efficient means of assembling such protein complexes (see Alternate Protocol). Briefly, one of the proteins is chosen as the “core” subunit or foundation for building the complex. This core subunit is bound to the affinity resin, and only the core subunit possesses the epitope tag that is recognized by the antibody-linked beads. One by one, each successive subunit is allowed to bind to the immobilized core subunit or to intermediate complexes. This process results in the assembly of a high-purity complex, even when the starting materials are relatively crude.

Table 10.15.1 Commonly Used Epitopes

Epitope	Peptide sequence	Antibody sources ^a	Affinity matrix source
FLAG	DYKDDDDK	Eastman Kodak ^b and BAbCO (M); Santa Cruz Biotechnology (P)	BAbCO
HA	YPYDVPDYA	BAbCO and Boehringer Mannheim (M); Santa Cruz Biotechnology (P)	BAbCO
c-Myc	EQKLISEEDL	BAbCO, Invitrogen, Boehringer Mannheim, and Sigma (M); Santa Cruz Biotechnology (P, M)	Santa Cruz Biotechnology
6-His	HHHHHH	BAbCO, Invitrogen ^c , and Sigma (M); Santa Cruz Biotechnology (P)	Qiagen ^d
AU1	DTYRYI	BAbCO (M)	BAbCO

^aM, monoclonal; P, polyclonal.

^bThree types of anti-FLAG monoclonal antibodies (M1, M2, M5) are available. Their specificities are dependent on the location of FLAG in the fusion protein.

^cThis antibody recognizes only the oligohistidine amino acid sequence at the carboxy-terminus of the fusion protein.

^dMetal-chelate affinity resin is used for purification of recombinant protein containing a polyhistidine tag (UNIT 10.11B).

Contributed by Ning Zhang and Jin-Long Chen

Current Protocols in Molecular Biology (1998) 10.15.1-10.15.9

Copyright © 1998 by John Wiley & Sons, Inc.

**Analysis of
Proteins**

10.15.1

Supplement 41

IMMUNOPRECIPITATION OF EPITOPE-TAGGED RECOMBINANT PROTEINS

This protocol describes immunoprecipitation of the epitope-tagged recombinant protein of interest. The protein is complexed with anti-epitope antibodies. It is a general procedure that can be used with a variety of transfection systems. The calcium phosphate–DNA precipitation method (UNIT 9.1) is recommended for transient transfection of mammalian cells with the recombinant plasmid vector, although other transfection methods may also be used. For detailed protocols addressing transfection of mammalian cells, see Chapter 9.

Materials

- Transfected cells (see Chapter 9), adherent (on a 100-mm plate) or nonadherent, expressing epitope-tagged protein of interest
- PBS (APPENDIX 2), ice cold, with and without 0.02% NaN_3
- Lysis buffer (see recipe)
- Protein A- or protein G-conjugated agarose beads (Pharmacia Biotech)
- Normal mouse IgG
- Epitope-specific antibody (see Table 10.15.1)
- 1.5× SDS sample buffer (prepare as a dilution of 2× SDS sample buffer; see UNIT 10.2)
- 10- and 15-ml centrifuge tubes
- Additional reagents and equipment for SDS-polyacrylamide gel electrophoresis (SDS-PAGE; UNIT 10.2)

Prepare lysate from transfected cells

For adherent cells

- 1a. At 36 to 48 hr posttransfection, aspirate the media from a 100-mm plate containing transfected cells that have been grown to 80% to 90% confluency ($0.5\text{--}2 \times 10^7$ cells, depending on the cell type), and wash the cells twice with 10 ml ice-cold PBS.
- 2a. Remove PBS and add 1 ml lysis buffer. Scratch the cells to one side of the dish with a rubber policeman.
- 3a. Transfer the lysate into a 1.5-ml microcentrifuge tube and incubate 10 min on ice.

For nonadherent cells

- 1b. At 48 hr posttransfection, transfer $\sim 2 \times 10^7$ cells into a 10-ml centrifuge tube and centrifuge 5 min at $2000 \times g$, 4°C .
- 2b. Remove the supernatant by aspiration, and resuspend the cell pellet in 1 ml ice-cold PBS. Transfer cells to a 1.5-ml microcentrifuge tube prechilled on ice and centrifuge 2 min at $2000 \times g$, 4°C .
- 3b. Aspirate the supernatant and resuspend the cell pellet in 1 ml lysis buffer by pipetting up and down several times. Incubate the sample 10 min on ice.
4. Microcentrifuge cells 10 min at $14,000 \times g$, 4°C .
5. Transfer the supernatant to a new microcentrifuge tube.

At this stage, the sample can be used immediately (step 9) or stored up to 1 year at -70°C . Thaw frozen samples on ice before use.

Protein expression should be monitored directly through immunoblot analysis (UNIT 10.8) using epitope-specific antibodies. To express two or more proteins with different epitope tags through cotransfection, each of the epitope-tagged recombinant proteins needs to be examined for its level of expression using epitope-specific antibodies.

Table 10.15.2 Relative Affinities of Protein A/G for Antibodies

Species and subclasses	Antibody	Affinity for protein A	Affinity for protein G
Rabbit	polyclonal	++++	+++
Mouse IgG1	monoclonal		++++
Mouse IgG2a	monoclonal	++++	++++
Mouse IgG2b	monoclonal	+++	+++
Mouse IgG3	monoclonal	+	+++

Prepare protein A- or protein G-agarose beads

- Transfer 0.5 ml (bed volume) of protein A- or protein G-agarose beads to a 15-ml centrifuge tube.
- Add 10 ml PBS and centrifuge 5 min at $2000 \times g$, 4°C . Discard the supernatant. Repeat twice (total three washes).
- Suspend beads in 0.5 ml PBS/0.02% NaN_3 .

At this stage, the 50% (v/v) slurry can be used immediately (step 9) or stored up to 6 months at 4°C .

Protein A and protein G may have different affinities for antibodies depending on their sources and subclasses (see Table 10.15.2). In addition to protein A/G-linked matrices, other matrices covalently linked directly to epitope-specific monoclonal antibodies are also available (see Table 10.15.1) and can be used for affinity purification of epitope-tagged proteins (see Basic Protocol 2).

Perform selection with epitope-specific antibody

- Preclarify 1 ml sample (step 5) by adding 50 μl of 50% (v/v) protein A- or protein G-agarose slurry (step 8) plus 5 μg normal mouse IgG. Shake 40 min at 4°C on an orbital shaker.

Preclarifying removes nonspecifically bound proteins from the extract. Antibodies raised against an irrelevant protein can also be used for this purpose.

- Centrifuge 2 min at $2000 \times g$, 4°C .
- Transfer the supernatant to a clean tube (discarding the pellet) and add 1 μg epitope-specific antibody.
- Shake the mixture 3 to 4 hr at 4°C on an orbital shaker.
- Centrifuge 2 min at $2000 \times g$, 4°C , to collect the immunoprecipitate. Discard the supernatant.
- Wash pellet with 1 ml lysis buffer and centrifuge 2 min at $2000 \times g$, 4°C . Discard supernatant and repeat twice (total three washes). Discard as much supernatant as possible after the last wash.
- Add 40 μl of 1.5 \times SDS sample buffer to the pellet and boil 5 min to denature the protein.
- Briefly centrifuge to pellet the debris. Electrophorese 20 μl of the supernatant by SDS-PAGE.

When performing immunoprecipitation with extracts prepared from metabolically labeled cells that have been transfected with an expression vector, the sample separated by

SDS-PAGE should be checked by autoradiography (APPENDIX 3A) to determine the presence of any coimmunoprecipitated proteins. If the aim of the study is to confirm whether two proteins interact in vivo, immunoprecipitants prepared from cells cotransfected with different expression constructs can be analyzed by immunoblotting (UNIT 10.8). The expressed recombinant proteins can be tagged with different epitopes, and their interaction can be tested by immunoprecipitation with an antibody against one epitope followed by immunoblotting with an antibody against the other.

IMMUNOPRECIPITATION OF EPITOPE-TAGGED RECOMBINANT PROTEINS FROM A BACULOVIRUS OVEREXPRESSION SYSTEM

This affinity purification protocol has been optimized for use with proteins that have been overexpressed using baculovirus-infected cells (UNIT 16.11). Additionally, the immunoprecipitation is performed using protein-A–Sepharose beads to which the anti-epitope antibody is already conjugated. This method can be modified to allow stepwise assembly of protein complexes (see Alternate Protocol).

Materials

- Baculovirus-infected Sf9 cells (UNIT 16.11), expressing epitope-tagged protein of interest
- PBS (APPENDIX 2), ice cold
- Binding buffer (see recipe), prepared as indicated *and* with an additional 60 mM KCl
- 50% (v/v) slurry of anti-epitope beads
- Normal mouse IgG
- Epitope elution buffer (see recipe)

Prepare lysate from transfected cells

1. Centrifuge $\sim 2 \times 10^7$ baculovirus-infected Sf9 cells 10 min at $2000 \times g$, 4°C.
2. Gently suspend the cell pellet in 10 ml ice-cold PBS. Centrifuge 10 min at $2000 \times g$, 4°C.
3. Aspirate the supernatant and repeat step 2.
4. Aspirate the supernatant. Add 1 ml binding buffer to the pellet and sonicate the sample twice for 15 seconds, as described (UNIT 16.5). Incubate 30 min on ice.
5. Centrifuge 10 min at $14,000 \times g$, 4°C.
6. Transfer the supernatant to a new tube. Preclear the sample (see Basic Protocol 1, steps 9 and 10) and transfer the supernatant to a new tube.

If desired, the sample can be stored up to 1 year at -70°C .

Precipitation caused by freezing and thawing can be removed by centrifugation as above. Ultracentrifugation (e.g., $100,000 \times g$) is recommended if nonspecific background proteins are detected in the resulting complex.

Bind sample to epitope-specific antibody beads

7. Add 50 μl of a 50% (v/v) slurry of anti-epitope beads to a 1.5-ml microcentrifuge tube. Wash the beads three times with PBS (see Basic Protocol 1, step 7).

Epitope-specific antibodies are covalently conjugated to protein A–Sepharose beads as described by Harlow and Lane (1988). For some epitopes, antibody affinity matrices are also commercially available (see Table 10.15.1).

8. Add 1 ml binding buffer and gently resuspend the beads.

9. Centrifuge 1 min at $2000 \times g$, 4°C , and discard the supernatant.
10. Add 1 ml cell lysate (step 6) to the beads.
11. Suspend the beads gently and incubate 2 hr at 4°C on an orbital shaker.
12. Centrifuge 1 min at $2000 \times g$, 4°C to recover the beads.
13. Wash the beads twice as in steps 8 and 9, using 1 ml binding buffer with a final of 160 mM KCl.

Elute protein of interest

14. Add 100 μl epitope elution buffer and resuspend the pellet. Incubate 2 hr at 4°C on an orbital shaker.

Epitope peptides are usually used for competitive elution of recombinant proteins bound to the affinity resin. The concentration of the peptide in the elution buffers is usually between 0.1 and 1.0 mg/ml. The complex can be eluted at room temperature if the elution efficiency at 4°C is low.

15. Centrifuge 1 min at $2000 \times g$, 4°C , and collect the supernatant.

Any residual beads in the supernatant can be removed by passing the supernatant through a suitable spin column (e.g., 1-ml Bio-spin, Bio-Rad).

16. Divide the eluted complex into 10- to 20- μl aliquots, quickly freeze the tubes in liquid nitrogen, and store the samples up to several months at -70°C .

STEPWISE ASSEMBLY OF PROTEIN COMPLEXES

Multisubunit complexes can be assembled by starting with a core protein (protein I) affixed to beads via an epitope tag (see Basic Protocol 2), and adding the other members of the complex in a stepwise manner. Protein II and subsequent members of the complex may be purified by conventional methods. It is not necessary to purify each of the participating proteins to homogeneity (see Critical Parameters and Troubleshooting); crude cell extracts may be also used. The necessity and/or degree of purification must be determined empirically. If protein II and/or successive proteins are affinity purified (see Basic Protocol 2), a different epitope from that used to affix the multisubunit complex to the beads must be employed.

Materials (also see Basic Protocol 2)

Baculovirus-infected Sf9 cell lines (UNIT 16.11), each expressing an additional epitope-tagged protein of interest (protein II and subsequent complex proteins)

1. Prepare cell lysates for all proteins to be complexed (see Basic Protocol 2, steps 1 to 6).

Alternatively, for protein II and any successive proteins, prepare purified or partially purified protein by other conventional chromatography methods (UNITS 10.9-10.11), as desired. Dialyze the sample against binding buffer prior to the assembly process.

2. Perform binding of protein I to anti-epitope beads (see Basic Protocol 2, steps 7 to 13).
3. Add 0.1 to 1 ml protein II cell lysate/extract to the tube and suspend the beads gently. Repeat the binding and washing procedure (see Basic Protocol 2, steps 11 to 13).

To assemble a protein complex that involves more than two proteins, repeat the above assembly process, incorporating each of the protein components into the complex in a stepwise fashion.

**ALTERNATE
PROTOCOL**

4. Wash the beads two more times.

These additional washes are required only after the final protein has been complexed.

5. Elute the formed complex (see Basic Protocol 2, steps 14 to 16), using epitope elution buffer that contains the peptide used to tag protein I.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Binding buffer

25 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES), pH 7.6
0.1 mM EDTA
100 mM KCl
12.5 mM MgCl₂
10% glycerol
0.1% Nonidet P-40 (NP-40)
Store up to 6 months at 4°C
Add DTT and proteinase inhibitors at 1:1000 from the following stock solutions immediately before use
1.0 M DTT (store up to several months at –20°C)
0.2 M 4-(2-aminoethyl)-benzenesulfonyl fluoride (AEBSF) hydrochloride (Calbiochem-Novabiochem; store up to 6 months at 4°C)
1 M sodium metabisulfite (store up to 6 months at –20°C)
2.7 mg/ml pepstatin in methanol (store up to 3 months at –20°C)
2 mg/ml leupeptin (store up to 3 months at –20°C)

Epitope elution buffer

Add purified epitope peptide to binding buffer (see recipe) immediately before use.

Use the same peptide that is used as the epitope tag. Use FLAG peptide at 0.2 mg/ml; use HA peptide at 1 mg/ml. The concentrations of other peptides in the elution buffer are usually between 0.1 and 1.0 mg/ml. The optimal concentration should be determined empirically. See Table 10.15.1 for peptide sequences.

Lysis buffer

50 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES), pH 7.5
1% Nonidet P-40 (NP-40)
0.5% sodium deoxycholate
150 mM NaCl
1 mM EDTA
0.1 mM Na₃VO₄
Store up to 6 months at 4°C
Add proteinase inhibitors at 1:1000 from the stock solution immediately before use

COMMENTARY

Background Information

Advantages of epitope tagging methodologies

Expression of proteins from newly cloned cDNAs in mammalian cell lines is an essential step towards a functional analysis of the proteins (Pati, 1992; Witzgall et al., 1994). In many instances, however, this is hindered by the difficulty in purifying the expressed protein and/or the lack of specific antibodies against the protein. To circumvent these problems, a technique known as epitope tagging has been developed. A recombinant fusion protein is generated, in which an epitope is appended at either the N- or C-terminus of the coding region of the protein of interest. The proteins expressed can be easily identified in whole cell extracts and can be purified by immunoprecipitation using antibodies against the epitope (Sells and Chernoff, 1995; Chubet and Brizzard, 1996).

Epitope tagging has been widely adopted as a technique for the detection, characterization, and purification of recombinant proteins for a number of reasons (Brown et al., 1995; Murray et al., 1995; Dear et al., 1997). First, it obviates the need for conventional methods of protein purification and animal immunization to generate antibodies, which are often more tedious and time-consuming. Second, the small size of the epitope presents little risk of disrupting the native conformation of the recombinant protein, which generally ensures the recovery of a biologically active protein. Third, highly specific antibodies raised against various tags are readily available. When a protein encoded by a gene of interest is epitope-tagged, it can be affinity purified before specific antibodies have been raised against it. This is particularly useful for studying proteins encoded by newly isolated genes. Fourth, expression and purification of epitope-tagged proteins has met with success in cell types as diverse as *E. coli*, yeast, insect cells, and mammalian cells.

Applications

Epitope-specific antibodies, when used in combination with secondary polyclonal antibodies and protein A or protein G conjugated with horseradish peroxidase (HRP), alkaline phosphatase (AP), or fluorescein isothiocyanate (FITC), have been applied in a variety of assays including immunoblotting (UNIT 10.8), immunoprecipitation (UNIT 10.16) and immunoprecipitation, enzyme-linked immunosor-

bant assays (ELISAs), fluorescence microscopy, gel retardation, mobility shift assays (UNIT 12.2), and fluorescence-activated cell sorting (FACS). Some specific examples of biological applications for epitope tagging methodology include: determining the size, intracellular localization, and abundance of proteins (Molloy et al., 1994; Canfield et al., 1996); monitoring posttranslational modification; analyzing the function of individual protein domains; studying receptor binding and internalization of exogenous proteins (Brown et al., 1995); and establishing the identity of a protein within a protein complex (Zhou et al., 1992).

For example, immunoprecipitants from cells transfected with the recombinant expression vector can be compared by SDS-PAGE with those from control cells transfected with the vector alone. This allows the detection of proteins that interact *in vivo* with a protein of interest in the cell (Murray et al., 1995; Thibault et al., 1997). Because these experiments can be carried out using transient transfection, they can be performed quickly and lead to the purification of the interacting protein(s) for further characterization. By cotransfecting multiple expression constructs, the *in vivo* interaction between proteins encoded by different genes can also be confirmed.

Proteins that have been epitope tagged can be purified following expression in any overexpression system (see Chapter 16). An appropriate lysis procedure must be used to solubilize the epitope-tagged protein, which can then be purified using the procedures described in Basic Protocol 1 or 2. In addition, epitope-tagged subunits can be used to generate purified multiprotein complexes in an intact form (see Alternate Protocol). In this case, one protein of a complex is tagged and expressed in a cell type that normally expresses that protein. The complex is then purified by techniques similar to those described in Basic Protocol 2. For this to be successful, the epitope on the tagged protein must be exposed to solution, and hence available to the antibody, when the protein is incorporated in the complex. This technology has been successful in purifying multiprotein complexes such as TFIID from human cells following expression of a tagged subunit by retrovirus-mediated expression (Zhou et al., 1992; also see Chapter 9) or from *Drosophila* embryos following expression of a tagged subunit by P element-mediated expression (Burke and Kadonaga, 1996).

Critical Parameters and Troubleshooting

Transfection and protein expression. The calcium phosphate–DNA precipitation method is recommended for transient transfection of recombinant expression vectors into most mammalian tissue culture cell lines. The level of protein expression may vary dramatically among different proteins being expressed. When the level of expression of a tagged protein is too low to be analyzed by immunoblotting, tandem repeats of epitope tags can be used to improve the sensitivity of detection (Nakajima and Yaoita, 1997). Alternatively, recombinant protein with multiple tags can be produced and purified in a sequential manner using epitope-specific antibodies against each tag (Pathak and Imperiali, 1997). Proteolysis should be minimized following cell lysis. This can be achieved by adding EDTA and proteinase inhibitors to the lysis buffer and by keeping the samples at 4°C. Proteinase inhibitors should be added fresh each time.

Cross-reactivity. One common problem in studies using epitope-specific antibodies is the presence of cross-reactivity. The reagents used to detect antibodies on the membrane often cross-react to the precipitating antibodies. This becomes worse when the secondary antibodies react to the antibodies used for immunoprecipitation. Therefore, it is necessary to carry out pilot experiments to optimize the quantity of protein extract and concentration of epitope-specific antibodies in order to minimize any adverse effects caused by this problem. If both monoclonal and polyclonal epitope-specific antibodies are available, cross-reactivity can be reduced by performing immunoprecipitation with a monoclonal antibody and analyzing the resulting products with the polyclonal antibodies.

Protein complexes. When assembling a protein complex, having a sufficient level of protein expression from each individual protein is likely to be the most critical factor that determines success. Expression in baculovirus-infected (see Chapter 16) or *E. coli* systems is thus recommended. Optimal conditions for the expression of each protein need to be determined empirically. The concentration and quality of protein subunits in the starting material are usually more important than their absolute purity. Thus, in some cases, pure complexes can be obtained using proteins derived from crude cell extracts. However, if the assembled protein complexes have a high level of impurities or the efficiency of generating the final assembled

products is very low, each of the protein components should be further purified prior to the assembly process.

Anticipated Results

Immunoprecipitated products are ready for electrophoresis by SDS-PAGE (UNIT 10.2). Analysis by immunoblotting (UNIT 10.8) should detect the transiently expressed recombinant protein using an epitope-specific antibody. Silver staining (UNIT 10.6) of the immunoprecipitants will reveal the protein profiles of the final product. Since nonspecific background is a common problem, the results should be interpreted with caution and proper control experiments should also be performed.

To determine the stoichiometry and purity of the assembled complexes, 5 to 10 μ l of eluted material can be analyzed by SDS-PAGE, followed by silver staining or immunoblotting. The integrity of the eluted complexes can be verified by reimmunoprecipitation using antibodies directed against different subunits. A relatively large amount (>60 μ l) of the eluted complexes is required for one reimmunoprecipitation.

Time Considerations

Preparation of the cell lysate takes 1 hr; an additional 4 to 5 hr are required to immunoprecipitate the recombinant protein from the cell lysate. The sample can be fractionated by SDS-PAGE on the same day and analyzed by immunoblotting on the following day. Preparation of epitope antibody–conjugated beads takes about 4 hr. The protein complex assembly process takes one day.

Literature Cited

- Brown, P.M., Tagari, P., Rowan, K.R., Yu, V.L., O'Neill, G.P., Middaugh, C.R., Sanyal, G., Ford-Hutchinson, A.W., and Nicholson, D.W. 1995. Epitope-labeled soluble human interleukin 5 (IL-5) receptors. *J. Biol. Chem.* 270:29236-29243.
- Burke, T.W. and Kadonaga, J.T. 1996. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes & Dev.* 10:711-724.
- Canfield, V.A., Norbeck, L., and Levenson, R. 1996. Localization of cytoplasmic and extracellular domains of Na,K-ATPase by epitope tag insertion. *Biochemistry* 35:14165-14172.
- Chubet, R.G. and Brizzard, B.L. 1996. Vectors for expression and secretion of FLAG epitope-tagged proteins in mammalian cells. *BioTechniques* 20:136-141.

- Dear, T.N., Hainzl, T., Follo, M., Nehls, M., Wilmore, H., Matena, K., and Boehm, T. 1997. Identification of interaction partners for the basic-helix-loop-helix protein E47. *Oncogene* 14:891-898.
- Harlow, E.D. and Lane, D. 1988. Immunoaffinity purification. In *Antibodies: A Laboratory Manual*. pp. 522-523. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Molloy, S.S., Thomas, L., Van Slyke, J.K., Stenberg, P.E., and Thomas, G. 1994. Intracellular trafficking and activation of the furin proprotein convertase: Localization of the TGN and recycling from the cell surface. *EMBO J.* 13:18-33.
- Murray, P.J., Watowich, S.S., Lodish, H.F., Young, R.A., and Hilton, D.J. 1995. Epitope tagging of the human endoplasmic reticulum HSP70 protein, BiP, to facilitate analysis of BiP-substrate interactions. *Anal. Biochem.* 229:170-179.
- Nakajima, K. and Yaoita, Y. 1997. Construction of multiple-epitope tag sequence by PCR for sensitive Western blot analysis. *Nucl. Acids Res.* 25:2231-2232.
- Pathak, R. and Imperiali, B. 1997. A dual affinity tag on the 64-kDa Nlt1p subunit allows the rapid characterization of mutant yeast oligosaccharyl transferase complexes. *Arch. Biochem. Biophys.* 338:1-6.
- Pati, U.K. 1992. Novel vectors for expression of cDNA encoding epitope-tagged proteins in mammalian cells. *Gene* 114:285-288.
- Sells, M.A. and Chernoff, J. 1995. Epitope-tag vectors for eukaryotic protein production. *Gene* 152:187-189.
- Thibault, C., Sganga, M.W., and Miles, M.F. 1997. Interaction of phosphatidylinositol-3-OH kinase-like protein with G protein $\beta\gamma$ subunits. *J. Biol. Chem.* 272:12253-12256.
- Witzgall, R., O'Leary, E., and Bonventre, J.V. 1994. A mammalian expression vector for the expression of GAL4 fusion proteins with an epitope tag and histidine tail. *Anal. Biochem.* 223:291-298.
- Zhou, Q., Lieberman, P.M., Boyer, T.G., and Berk, A.J. 1992. Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Genes & Dev.* 6:1964-1974.

Key References

- Chen, J.-L. and Tjian, R. 1996. Reconstitution of TATA-binding protein-associated factor/TATA-binding protein complexes for in vitro transcription. *Methods Enzymol.* 273:208-217.
- Chiang, C.-M. and Roeder, R.G. 1993. Expression and purification of general transcription factors by FLAG epitope tagging and peptide elution. *Pept. Res.* 6:62-64.

Contributed by Ning Zhang and
Jin-Long Chen
Tularik, Inc.
South San Francisco, California

Immunoprecipitation

UNIT 10.16

Immunoprecipitation is a technique in which an antigen is isolated by binding to a specific antibody attached to a sedimentable matrix. The source of antigen for immunoprecipitation can be unlabeled cells or tissues, metabolically or extrinsically labeled cells (UNIT 10.18), subcellular fractions from either unlabeled or labeled cells (see Chapter 3), or *in vitro*-translated proteins (UNIT 10.17). Immunoprecipitation is also used to analyze protein fractions separated by other biochemical techniques such as gel filtration (UNIT 10.9) or sedimentation on density gradients. Either polyclonal or monoclonal antibodies from various animal species can be used in immunoprecipitation protocols. Antibodies can be bound noncovalently to immunoabsorbents such as protein A- or protein G-agarose, or can be coupled covalently to a solid-phase matrix.

Immunoprecipitation protocols consist of several stages (Fig. 10.16.1; see Basic Protocol 1). In stage 1, the antigen is solubilized by one of several techniques for lysing cells. Soluble and membrane-associated antigens can be released from cells grown either in suspension culture (see Basic Protocol 1) or as a monolayer on tissue culture dishes (see Alternate Protocol 1) with nondenaturing detergents. Alternatively, cells can be lysed under denaturing conditions (see Alternate Protocol 2). Soluble antigens can also be extracted by mechanical disruption of cells in the absence of detergents (see Alternate Protocol 3). All of these procedures are suitable for extracting antigens from animal cells. By contrast, yeast cells require disruption of their cell wall in order to allow extraction of the antigens (see Alternate Protocol 4).

In stage 2, a specific antibody is attached, either noncovalently or covalently, to a sedimentable, solid-phase matrix to allow separation by low-speed centrifugation. In this unit, two methods for achieving this are described: the noncovalent attachment of antibody to protein A- or protein G-agarose beads (see Basic Protocol 1) and covalent coupling to Sepharose (see Alternate Protocol 5 and Support Protocol).

Stage 3 is the actual immunoprecipitation, which can be achieved by incubating the solubilized antigen from stage 1 with the immobilized antibody from stage 2, followed by extensive washing to remove unbound proteins (see Basic Protocol 1). Another method is to precipitate the immune complexes using antibodies contained in an anti-immunoglobulin (anti-Ig) serum (see Alternate Protocol 6). Immunoprecipitated antigens can be dissociated from antibodies and reprecipitated by a protocol referred to as "immunoprecipitation-recapture" (see Basic Protocol 2). This procedure can be used with the same antibody for further purification of the antigen, or with a second antibody to identify components of multisubunit complexes or to study protein-protein interactions (Fig. 10.16.3). Immunoprecipitated antigens can be analyzed by one-dimensional electrophoresis (UNIT 10.2), two-dimensional electrophoresis (UNITS 10.3 & 10.4), or immunoblotting (UNIT 10.8). In some cases, immunoprecipitates can be used for structural or functional analyses of the isolated antigens. Immunoprecipitates can also be used as sources of immunogens for production of monoclonal or polyclonal antibodies.

IMMUNOPRECIPITATION USING CELLS IN SUSPENSION LYSED WITH A NONDENATURING DETERGENT SOLUTION

In this protocol, cells in suspension (labeled or unlabeled) are extracted by incubation in nondenaturing lysis buffer containing the nonionic detergent Triton X-100 (steps 1 to 7). This procedure results in the release of both soluble and membrane proteins; however, many cytoskeletal and nuclear proteins, as well as a fraction of membrane proteins, are

**BASIC
PROTOCOL 1**

**Analysis of
Proteins**

10.16.1

Contributed by Juan S. Bonifacino, Esteban C. Dell'Angelica, and Timothy A. Springer

Current Protocols in Molecular Biology (1999) 10.16.1-10.16.29

Copyright © 1999 by John Wiley & Sons, Inc.

Supplement 48

not efficiently extracted under these conditions. The procedure allows immunoprecipitation with antibodies to epitopes that are exposed in native proteins.

For immunoprecipitation, a specific antibody is immobilized on a sedimentable, solid-phase matrix (steps 8 to 14). Although there are many ways to attach antibodies to matrices (see Commentary), the most commonly used methods rely on the property of immunoglobulins to bind *Staphylococcus aureus* protein A, or protein G from group G *Streptococcus* (Table 10.16.1). The best results are obtained by binding antibodies to protein A or protein G that is covalently coupled to agarose beads. In this protocol, Sepharose beads are used (Sepharose is a more stable, cross-linked form of agarose). Immunoprecipitation is most often carried out using rabbit polyclonal or mouse monoclonal antibodies, which,

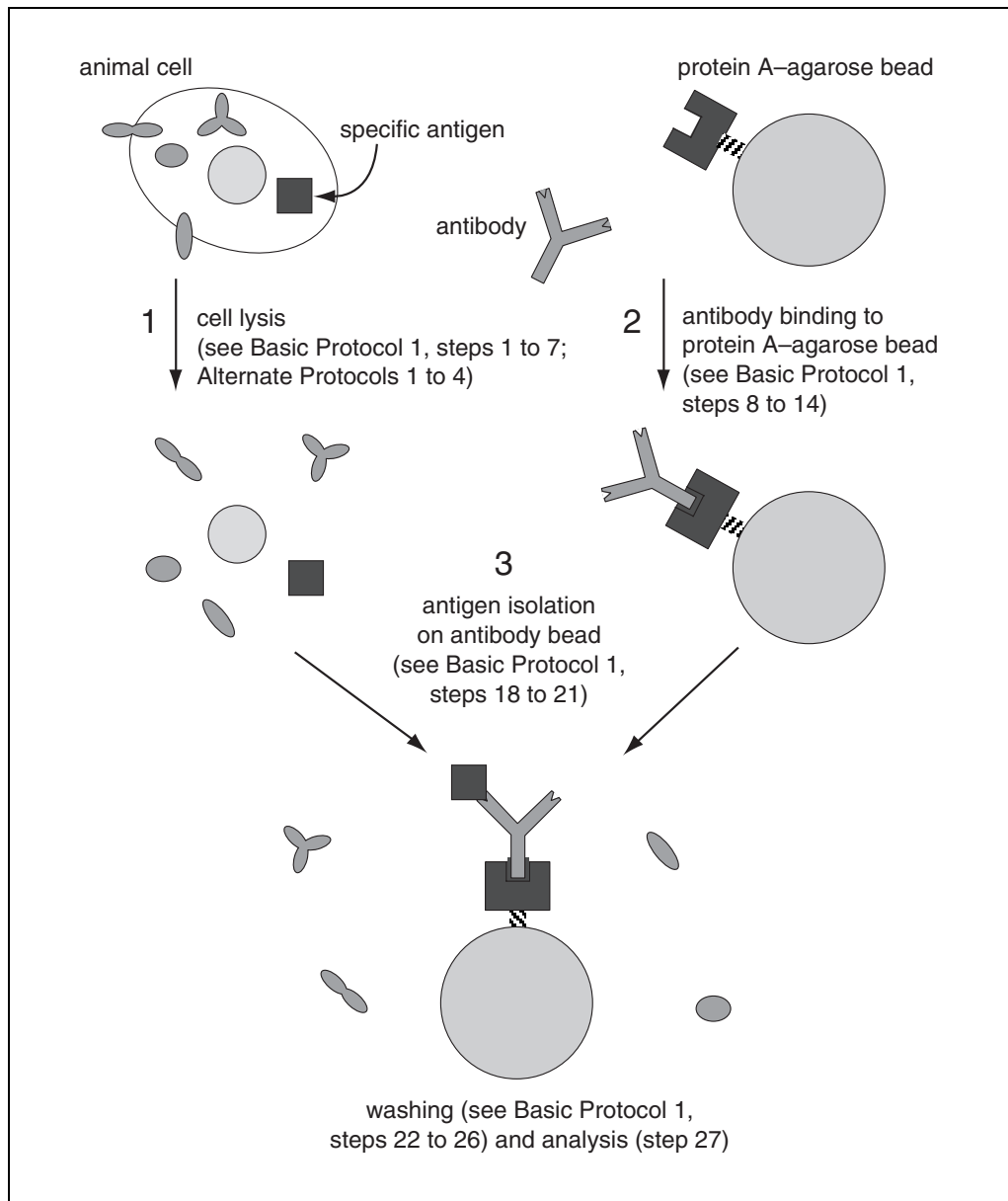


Figure 10.16.1 Schematic representation of the stages of the immunoprecipitation protocol presented in Basic Protocol 1. (1) Cell lysis: antigens are solubilized by extraction of the cells in the presence or absence of detergents. To increase specificity, the cell lysate can be precleared with protein A-agarose beads (steps 15 to 17, not shown). (2) Antibody immobilization: a specific antibody is bound to protein A-agarose beads. (3) Antigen capture: the solubilized antigen is isolated on antibody-conjugated beads.

with some exceptions (e.g., mouse IgG1), bind well to protein A (Table 10.16.1). Antibodies that do not bind to protein A–agarose can be adsorbed to protein G–agarose (Table 10.16.1) using exactly the same protocol. For optimal time management, incubation of antibodies with protein A–agarose can be carried out either before or during lysis of the cells.

The final stage in immunoprecipitation is combining the cell lysate with the antibody-conjugated beads and isolating the antigen (steps 18 to 26). This can be preceded by an optional preclearing step in which the lysate is absorbed with either “empty” protein A–agarose beads or with an irrelevant antibody bound to protein A–agarose (steps 15 to 17). The need for preclearing depends on the specific experimental system being studied and the quality of the antibody reagents. The protocol described below incorporates a preclearing step using protein A–agarose. Protein fractions separated by techniques such

Table 10.16.1 Binding of Antibodies to Protein A and Protein G^{a,b,c}

Antibody	Protein A binding	Protein G binding ^d
<i>Monoclonal antibodies^e</i>		
Human IgG1	++	++
Human IgG2	++	++
Human IgG3	–	++
Human IgG4	++	++
Mouse IgG1	+	++
Mouse IgG2a	++	++
Mouse IgG2b	++	++
Mouse IgG3	++	++
Rat IgG1	+	+
Rat IgG2a	–	++
Rat IgG2b	–	++
Rat IgG2c	++	++
<i>Polyclonal antibodies</i>		
Chicken	–	–
Donkey	–	++
Goat	+	++
Guinea pig	++	+
Hamster	+	++
Human	++	++
Monkey	++	++
Mouse	++	++
Rabbit	++	++
Rat	+	+
Sheep	+	++

^a++, moderate to strong binding; +, weak binding; –, no binding.

^bA hybrid protein A/G molecule that combines the features of protein A and protein G, coupled to a solid-phase matrix, is available from Pierce.

^cInformation from Harlow and Lane (1999), and from Amersham Pharmacia Biotech, Pierce, and Jackson ImmunoResearch.

^dNative protein G binds albumin from several animal species. Recombinant variants of protein G have been engineered for better binding to rat, mouse, and guinea pig IgG, as well as for avoiding binding to serum albumin.

^eProtein A binds some IgM, IgA, and IgE antibodies in addition to IgG, whereas protein G binds only IgG.

as gel filtration (UNIT 10.9) or sedimentation on sucrose gradients can be used in place of the cell lysate at this stage. After binding the antigen to the antibody-conjugated beads, the unbound proteins are removed by successive washing and sedimentation steps.

Materials

Unlabeled or labeled cells in suspension
PBS (APPENDIX 2), ice cold
Nondenaturing lysis buffer (see recipe), ice cold
50% (v/v) protein A–Sepharose bead (Sigma, Amersham Pharmacia Biotech)
slurry in PBS containing 0.1% (w/v) BSA and 0.01% (w/v) sodium azide
(NaN₃)
Specific polyclonal antibody (antiserum or affinity-purified immunoglobulin) or
monoclonal antibody (ascites, culture supernatant, or purified immunoglobulin)
Control antibody of same type as specific antibody (e.g., preimmune serum or
purified irrelevant immunoglobulin for specific polyclonal antibody; irrelevant
ascites, hybridoma culture supernatant, or purified immunoglobulin for specific
monoclonal antibody; see Critical Parameters)
10% (w/v) BSA
Wash buffer (see recipe), ice cold

Microcentrifuge with fixed-angle rotor (Eppendorf 5415C or equivalent)
Tube rotator (capable of end-over-end inversion)
Pasteur pipet attached to a vacuum trap

CAUTION: When working with radioactivity, take appropriate precautions to avoid contamination of the experimenter and the surroundings. Carry out the experiment and dispose of wastes in an appropriately designated area, following the guidelines provided by the local radiation safety officer (also see APPENDIX 1F).

NOTE: All solutions should be ice cold and procedures should be carried out at 4°C or on ice.

Prepare cell lysate

1. Collect cells in suspension by centrifuging 5 min at 400 × g, 4°C, in a 15- or 50-ml capped conical tube. Place tube on ice.

Approximately 0.5–2 × 10⁷ cells are required to yield 1 ml lysate, which is the amount generally used for each immunoprecipitation.

Labeled cells are likely to have been pelleted earlier as part of the labeling procedure. If the cells are frozen, they should be thawed on ice before solubilization.

2. Aspirate supernatant with a Pasteur pipet attached to a vacuum trap.

CAUTION: Dispose of radioactive materials following applicable safety regulations (APPENDIX 1F).

3. Resuspend cells gently by tapping the bottom of the tube. Rinse cells twice with ice-cold PBS as in steps 1 and 2, using the same volume of PBS as in the initial culture.
4. Add 1 ml ice-cold nondenaturing lysis buffer per ~0.5–2 × 10⁷ cells and resuspend pellet by gentle agitation for 3 sec with a vortex mixer set at medium speed.

Do not shake vigorously, as this could result in loss of material or protein denaturation due to foaming.

5. Keep suspension on ice 15 to 30 min and transfer to a 1.5-ml conical microcentrifuge tube.

Tubes can have flip-top or screw caps. Screw-capped tubes are preferred because they are less likely to open accidentally during subsequent procedures. They are also recommended for work with radioactivity.

6. Clear the lysate by microcentrifuging 15 min at $16,000 \times g$ (maximum speed), 4°C .

Centrifugation can be carried out in a microcentrifuge placed in a cold room or in a refrigerated microcentrifuge. Take precautions to ensure that the 4°C temperature is maintained during the spin (e.g., use a fixed-angle rotor with a lid, as the aerodynamics of this type of rotor reduces generation of heat by friction). If it is necessary to reduce background, the lysate can be spun for 1 hr at $100,000 \times g$ in an ultracentrifuge.

7. Transfer the supernatant to a fresh microcentrifuge tube using an adjustable pipet fitted with a disposable tip. Do not disturb the pellet, and leave the last 20 to 40 μl of supernatant in the centrifuge tube. Keep the cleared lysate on ice until preclearing (step 15) or addition of antibody beads (step 18).

NOTE: Resuspension of even a small amount of sedimented material will result in high nonspecific background due to carryover into the immunoprecipitation steps. A cloudy layer of lipids floating on top of the supernatant will not adversely affect the results of the immunoprecipitation.

When the lysate is highly radioactive—as is the case for metabolically labeled cells—the use of tips with aerosol barriers is recommended to reduce the risk of contaminating internal components of the pipet.

Cell extracts can be frozen at -70°C until used for immunoprecipitation. However, it is preferable to lyse the cells immediately before immunoprecipitation in order to avoid protein degradation or dissociation of protein complexes. If possible, freeze the cell pellet from step 3 rather than the supernatant from step 7.

Prepare antibody-conjugated beads

8. In a 1.5-ml conical microcentrifuge tube, combine 30 μl of 50% protein A–Sepharose bead slurry, 0.5 ml ice-cold PBS, and the following quantity of specific antibody (select one):

- 1 to 5 μl polyclonal antiserum
- 1 μg affinity-purified polyclonal antibody
- 0.2 to 1 μl ascitic fluid containing monoclonal antibody
- 1 μg purified monoclonal antibody
- 20 to 100 μl culture supernatant containing monoclonal antibody.

The quantities of antibody suggested are rough estimates based on the expected amount of specific antibodies in each preparation. Quantities can be increased or decreased depending on the quality of the antibody preparation (see Commentary).

Substitute protein G for protein A if antibodies are of a species or subclass that does not bind to protein A (see Table 10.16.1).

If the same antibody will be used to immunoprecipitate multiple samples (e.g., samples from a pulse-chase experiment; UNIT 10.18), the quantities indicated above can be increased proportionally to the number of samples and incubated in a 15-ml capped conical tube. In this case, the beads should be divided into aliquots just prior to the addition of the cleared cell lysate (step 18).

Antibody-conjugated beads can be prepared prior to preparation of the cell lysate (steps 1 to 7), in order to minimize the time that the cell extract is kept on ice.

9. Set up a nonspecific immunoprecipitation control in a 1.5-ml conical microcentrifuge tube by incubating 30 μl of 50% protein A–Sepharose bead slurry, 0.5 ml ice-cold PBS, and the appropriate control antibody (select one):

- 1 to 5 μl preimmune serum as a control for a polyclonal antiserum
- 1 μg purified irrelevant polyclonal antibody (an antibody to an epitope that is not present in the cell lysate) as a control for a purified polyclonal antibody
- 0.2 to 1 μl ascitic fluid containing irrelevant monoclonal antibody (an antibody to an epitope that is not present in the cell lysate and of the same species and immunoglobulin subclass as the specific antibody) as a control for an ascitic fluid containing specific monoclonal antibody
- 1 μg purified irrelevant monoclonal antibody as a control for a purified monoclonal antibody
- 20 to 100 μl hybridoma culture supernatant containing irrelevant monoclonal antibody as a control for a hybridoma culture supernatant containing specific monoclonal antibody.

The amount of irrelevant antibody should match that of the specific antibody and the antibody should be from the same species as the specific antibody.

10. Mix suspensions thoroughly. Tumble incubation mixtures end over end ≥ 1 hr at 4°C in a tube rotator.

Addition of 0.01% (w/v) Triton X-100 may facilitate mixing of the suspension during tumbling. Incubations can be carried out for as long as 24 hr. This allows preparation of the antibody-conjugated beads prior to immunoprecipitation.

11. Microcentrifuge 2 sec at $16,000 \times g$ (maximum speed), 4°C .
12. Aspirate the supernatant (containing unbound antibodies) using a fine-tipped Pasteur pipet connected to a vacuum aspirator.
13. Add 1 ml nondenaturing lysis buffer and resuspend the beads by inverting the tube three or four times.

For lysates prepared with detergents (this protocol and see Alternate Protocols 1 and 2), use 1 ml nondenaturing lysis buffer; for lysates prepared by mechanical disruption (see Alternate Protocol 3), use detergent-free lysis buffer (see recipe).

Use of a repeat pipettor is recommended when processing multiple samples.

14. Wash by repeating steps 11 to 13, and then steps 11 and 12 once more.

At this point the beads have been washed twice with lysis buffer and are ready to be used for immunoprecipitation. Antibody-bound beads can be stored up to 6 hr at 4°C until used.

Preclear lysate (optional)

15. In a microcentrifuge tube, combine 1 ml cell lysate (from step 7) and 30 μl of 50% protein A–Sepharose bead slurry.

The purpose of this step is to remove from the lysate proteins that bind to protein A–Sepharose, as well as pieces of insoluble material that may have been carried over from previous steps. If the lysate was prepared from cells expressing immunoglobulins—such as spleen cells or cultured B cells—the preclearing step should be repeated at least three times to ensure complete removal of endogenous immunoglobulins.

If cell lysates were frozen and thawed, they should be microcentrifuged 15 min at $16,000 \times g$ (maximum speed), 4°C , before the preclearing step.

16. Tumble end over end 30 min at 4°C in a tube rotator.
17. Microcentrifuge 5 min at $16,000 \times g$ (maximum speed), 4°C .

Immunoprecipitate

18. Add 10 μl of 10% BSA to the tube containing specific antibody bound to protein A–Sepharose beads (step 14), and transfer to this tube the entire volume of cleared lysate (from step 7 or 17). If a nonspecific immunoprecipitation control is performed, divide lysate in two $\sim 0.4\text{-ml}$ aliquots, one for the specific antibody and the other for the nonspecific control.

In order to avoid carryover of beads with precleared material, leave 20 to 40 μl of supernatant on top of the pellets in the preclearing tubes. Discard beads and remaining supernatant. The BSA quenches nonspecific binding to the antibody-conjugated beads during incubation with the cell lysate.

19. Incubate 1 to 2 hr at 4°C while mixing end over end in a tube rotator.

Samples can be incubated overnight, although there is an increased risk of protein degradation, dissociation of multiprotein complexes, or formation of protein aggregates.

20. Microcentrifuge 5 sec at 16,000 $\times g$ (maximum speed), 4°C.

21. Aspirate the supernatant (containing unbound proteins) using a fine-tipped Pasteur pipet connected to a vacuum aspirator.

The supernatant can be kept up to 8 hr at 4°C or up to 1 month at -70°C for sequential immunoprecipitation of other antigens or for analysis of total proteins. To reuse lysate, remove the supernatant carefully with an adjustable pipet fitted with a disposable tip. Before reprecipitation, preabsorb the lysate with protein A–Sepharose (as in steps 15 to 17) to remove antibodies that may have dissociated during the first immunoprecipitation.

CAUTION: *Dispose of radioactive materials following applicable safety regulations.*

22. Add 1 ml ice-cold wash buffer, cap the tubes, and resuspend the beads by inverting the tube 3 or 4 times.

Use of a repeat pipettor is recommended when processing multiple samples.

23. Microcentrifuge 2 sec at 16,000 $\times g$ (maximum speed), 4°C.

24. Aspirate the supernatant, leaving $\sim 20\ \mu\text{l}$ supernatant on top of the beads.

25. Wash beads three more times (steps 22 to 24).

Total wash time (steps 22 to 26) should be ~ 30 min, keeping the samples on ice for 3 to 5 min between washes if necessary (see Critical Parameters).

26. Wash beads once more using 1 ml ice-cold PBS and aspirate supernatant completely with a drawn-out Pasteur pipet or an adjustable pipet fitted with a disposable tip.

The final product should be 15 μl of settled beads containing bound antigen.

Immunoprecipitates can either be processed immediately or frozen at -20°C for later analysis. For subsequent analysis of the isolated proteins prior to electrophoresis (e.g., comparison of the electrophoretic mobility of the antigen with or without treatment with glycosidases), samples can be divided into two or more aliquots after addition of PBS. Transfer aliquots of the bead suspension to fresh tubes, centrifuge and aspirate as in the previous steps.

27. Analyze immunoprecipitates by one-dimensional electrophoresis (UNIT 10.2B), two-dimensional electrophoresis (UNIT 10.3 & 10.4), or immunoblotting (UNIT 10.8).

IMMUNOPRECIPITATION USING ADHERENT CELLS LYSED WITH A NONDENATURING DETERGENT SOLUTION

Immunoprecipitation using adherent cells can be performed in the same manner as with nonadherent cells (see Basic Protocol 1). This protocol is essentially similar to steps 1 to 5 of Basic Protocol 1, but describes modifications necessary for using the same nondenaturing detergent solution to lyse cells attached to tissue culture plates. It is preferable to use cells grown on plates rather than in flasks, because the cell monolayer is more easily accessible.

Additional Materials (also see Basic Protocol 1)

- Unlabeled or labeled cells grown as a monolayer on a tissue culture plate (*UNIT 10.18*)
- Rubber policeman

NOTE: All solutions should be ice cold and procedures should be carried out at 4°C or on ice.

1. Rinse cells attached to a tissue culture plate twice with ice-cold PBS. Remove the PBS by aspiration with a Pasteur pipet attached to a vacuum trap.

CAUTION: Dispose of radioactive materials following applicable safety regulations.

2. Place the tissue culture plate on ice.
3. Add ice-cold nondenaturing lysis buffer to the tissue culture plate.

Use 1 ml lysis buffer for an 80% to 90% confluent 100-mm-diameter tissue culture plate. Depending on the cell type, a confluent 100-mm dish will contain $0.5\text{--}2 \times 10^7$ cells. For other plate sizes, adjust volume of lysis buffer according to the surface area of the plate.

4. Scrape the cells off the plate with a rubber policeman, and transfer the suspension to a 1.5-ml conical microcentrifuge tube using an adjustable pipettor fitted with a disposable tip. Vortex gently for 3 sec and keep tubes on ice for 15 to 30 min.

Tubes can have flip-top or screw caps. Screw-capped tubes are preferred because they are less likely to open accidentally during subsequent procedures. They are also recommended for work with radioactivity.

5. Clear the lysate and perform immunoprecipitation (see Basic Protocol 1, steps 6 to 27).

IMMUNOPRECIPITATION USING CELLS LYSED WITH DETERGENT UNDER DENATURING CONDITIONS

If epitopes of native proteins are not accessible to antibodies, or if the antigen cannot be extracted from the cell with nonionic detergents, cells should be solubilized under denaturing conditions. This protocol is based on that for nondenaturing conditions (see Basic Protocol 1, steps 1 to 7), with the following modifications. Denaturation is achieved by heating the cells in a denaturing lysis buffer that contains an ionic detergent such as SDS or Sarkosyl (*N*-lauroylsarcosine). The denaturing lysis buffer also contains DNase I to digest DNA released from the nucleus. Prior to immunoprecipitation, the denatured protein extract is diluted 10-fold with nondenaturing lysis buffer, which contains Triton X-100; this step protects the antigen-antibody interaction from interference by the ionic detergent. Immunoprecipitation is performed as described (see Basic Protocol 1).

The following protocol is described for cells in suspension culture, although it can be adapted for adherent cells (see Alternate Protocol 1). Only antibodies that react with denatured proteins can be used to immunoprecipitate proteins solubilized by this protocol.

Additional Materials (also see Basic Protocol 1)

Denaturing lysis buffer (see recipe)
Heating block set at 95°C (Eppendorf Thermomixer 5436 or equivalent)
25-G needle attached to 1-ml syringe

1. Collect cells in suspension culture (see Basic Protocol 1, steps 1 to 3). Place tubes on ice.
2. Add 100 µl denaturing lysis buffer per $\sim 0.5\text{--}2 \times 10^7$ cells in the pellet.
3. Resuspend the cells by vortexing vigorously 2 to 3 sec at maximum speed. Transfer suspension to a 1.5-ml conical microcentrifuge tube.

The suspension may be very viscous due to release of nuclear DNA.

Tubes can have flip-top or screw caps. Screw-capped tubes are preferred because they are less likely to open accidentally during subsequent procedures. They are also recommended for work with radioactivity.

4. Heat samples 5 min at 95°C in a heating block.
5. Dilute the suspension with 0.9 ml nondenaturing lysis buffer. Mix gently.

The excess 1% Triton X-100 in the nondenaturing lysis buffer sequesters SDS into Triton X-100 micelles.

6. Shear DNA by passing the suspension five to ten times through a 25-G needle attached to a 1-ml syringe.

If the DNA is not digested by DNase I in the denaturing lysis buffer or thoroughly sheared mechanically, it will interfere with the separation of pellet and supernatant after centrifugation. Repeat mechanical disruption until the viscosity is reduced to manageable levels.

7. Incubate 5 min on ice.
8. Clear the lysate and perform immunoprecipitation (see Basic Protocol 1, steps 6 to 27).

IMMUNOPRECIPITATION USING CELLS LYSED WITHOUT DETERGENT

Immunoprecipitation of proteins that are already soluble within cells (e.g., cytosolic or luminal organellar proteins) may not require the use of detergents. Instead, cells can be mechanically disrupted by repeated passage through a needle, and soluble proteins can be separated from insoluble material by centrifugation. The following protocol describes lysis of cells in a PBS-based detergent-free lysis buffer. Other buffer formulations may be used for specific proteins.

Additional Materials (also see Basic Protocol 1)

Detergent-free lysis buffer (see recipe)
25-G needle attached to 3-ml syringe

NOTE: All solutions should be ice-cold and procedures should be carried out at 4°C or on ice.

1. Collect and wash cells in suspension (see Basic Protocol 1, steps 1 to 3).
2. Add 1 ml of ice-cold detergent-free lysis buffer per $\sim 0.5\text{--}2 \times 10^7$ cells in a pellet.
3. Resuspend the cells by gentle agitation for 3 sec with a vortex mixer set at medium speed.

**ALTERNATE
PROTOCOL 3**

**Analysis of
Proteins**

10.16.9

4. Break cells by passing the suspension 15 to 20 times through a 25-G needle attached to a 3-ml syringe.

Extrusion of the cell suspension from the syringe should be rapid, although care should be exercised to prevent splashing and excessive foaming. Cell breakage can be checked under a bright-field or phase-contrast microscope. Repeat procedure until >90% cells are broken.

It is helpful to check ahead of time whether the cells can be broken in this way. If the cells are particularly resistant to mechanical breakage, they can be swollen for 10 min at 4°C with a hypotonic solution containing 10 mM Tris·Cl, pH 7.4 (APPENDIX 2) before mechanical disruption.

5. Clear the lysate and perform immunoprecipitation (Basic Protocol 1, steps 6 to 27).

ALTERNATE PROTOCOL 4

IMMUNOPRECIPITATION USING YEAST CELLS DISRUPTED WITH GLASS BEADS

Unlike animal cells, yeast cells have an extremely resistant, detergent-insoluble cell wall. To allow extraction of cellular antigens, the cell wall needs to be broken by mechanical, enzymatic, or chemical means. The most commonly used procedure consists of vigorous vortexing of the yeast suspension with glass beads. The breakage can be done in the presence or absence of detergent, as previously described for animal cells (see Basic Protocol 1, Alternate Protocol 2, and Alternate Protocol 3). The protocol described below is suitable for mechanical disruption of most yeast species, including *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. A protocol for metabolic labeling for yeast has been described by Franzusoff et al. (1991).

Additional Materials (also see Basic Protocol 1)

Unlabeled or radiolabeled yeast cells

Lysis buffer, ice cold: nondenaturing, denaturing, or detergent-free lysis buffer (see recipes)

Glass beads (acid-washed, 425- to 600- μ m diameter; Sigma)

NOTE: All solutions should be ice-cold and procedures should be carried out at 4°C or on ice.

1. Collect 10 ml of yeast culture at 1 OD₆₀₀ per immunoprecipitation sample, and centrifuge 5 min at 4000 \times g, 4°C. Place tube on ice.
2. Remove supernatant by aspiration with a Pasteur pipet attached to a vacuum trap.

CAUTION: *Dispose of radioactive materials following applicable safety regulations.*

3. Loosen pellet by vortexing vigorously for 10 sec. Rinse cells twice with ice-cold distilled water as in steps 1 and 2.

Radiolabeled yeast cells are likely to have been pelleted earlier as part of the labeling procedure. If the pellets are frozen, they should be thawed on ice prior to cell disruption.

4. Add 3 vol ice-cold lysis buffer and 3 vol glass beads per volume of pelleted yeast cells.

Use nondenaturing lysis buffer or detergent-free lysis buffer as required for the antigen under study. If the experiment requires denaturation of the antigen, the procedure can be adapted to include this (see Alternate Protocol 2 for higher eukaryotic cells); however, the yeast cells must be broken with glass beads before heating the sample at 95°C.

5. Shake cells by vortexing vigorously at maximum speed for four 30-sec periods, keeping the cells on ice for 30 sec between the periods.

Check cell breakage under a bright-field or phase-contrast microscope. It is helpful to check ahead of time if the cells can be broken in this way.

6. Remove the yeast cell lysate from the beads using a pipettor with a disposable tip. Transfer to a fresh tube.
7. Add 4 vol (see step 4) lysis buffer to the glass beads, vortex for 2 sec, and combine this supernatant with the lysate from step 6.
8. Clear the lysate and perform immunoprecipitation (see Basic Protocol 1, steps 6 to 27).

IMMUNOPRECIPITATION WITH ANTIBODY-SEPHAROSE

This protocol, which follows the steps presented in Figure 10.16.2, relies on the formation of an insoluble immune complex between a protein antigen and an antigen-specific monoclonal (or polyclonal) antibody covalently bound to Sepharose.

Materials

Unlabeled cells, surface-labeled cells (e.g., with ^{125}I or biotin; *UNIT 3.18*) or biosynthetically ^{35}S -, ^3H -, or ^{14}C -labeled cells (*UNIT 10.18*)
Triton X-100 lysis buffer (see recipe)
Dilution buffer (see recipe)
Antibody (Ab)-Sepharose (see Support Protocol)
Activated, quenched (control) Sepharose, prepared as for Ab-Sepharose (see Support Protocol) but eliminating Ab or substituting irrelevant Ab during coupling
Tris/saline/azide (TSA) solution (see recipe)
0.05 M Tris-Cl, pH 6.8 (*APPENDIX 2*)
2× SDS sample buffer (*UNIT 10.2A*)

NOTE: Carry out all procedures in a 4°C cold room or on ice.

Lyse cells and preclear the lysate

1. Incubate cells in Triton X-100 lysis buffer (5×10^7 cells/ml) for 1 hr at 4°C.
2. Centrifuge the lysate 10 min at $3000 \times g$ to remove nuclei and save the supernatant.
3. Centrifuge the supernatant 1 hr at $100,000 \times g$ and save the supernatant.

Supernatants may also be prepared by microcentrifugation ($10,000 \times g$) for 30 min.

The supernatant must be used within several days or stored at -70°C . The length of storage is limited by autoradiolysis and the half-life of the isotope. ^3H - and ^{14}C -labeled samples can often be stored frozen for years. Storage of ^{125}I -labeled samples is usually limited to 1 to 2 months because of autoradiolysis, while the usefulness of ^{35}S -labeled samples is usually limited to 6 months because of half-life. Repeated freezing and thawing may disrupt antigenic determinants and dissociate some protein complexes, especially those that are noncovalently associated.

4. Preclear supernatant to be used in one batch by adding 10 μl activated, quenched (control) Sepharose per 200 μl supernatant. Shake on an orbital shaker 2 hr at room temperature or overnight at 4°C. Centrifuge 1 min at $200 \times g$ and save supernatant.

Preclearing removes nonspecifically absorbing material. Control Sepharose can be prepared without antibody or coupled with irrelevant (nonspecific) antibody. Irrelevant antibody is an antibody directed against an unrelated protein, and could also be whole IgG; it must not cross-react with the protein being immunoprecipitated.

ALTERNATE PROTOCOL 5

Analysis of Proteins

10.16.11

Immunoprecipitate the antigen

5. Precoat 1.5-ml microcentrifuge tubes by filling with Triton X-100 lysis buffer 10 min at room temperature. Remove the solution by aspiration.

Precoating minimizes antigen absorption to the tube.

6. Add 10^5 to 10^6 cpm of radiolabeled (^{125}I or ^{35}S) supernatant containing antigen (from step 4) to a precoated microcentrifuge tube and bring the volume to 200 μl with dilution buffer.

The recommended amount of radioactivity is appropriate for eukaryotic cells with >1000 molecules of antigen/cell. It is assumed that detection on slab gels of ^{125}I -labeled proteins will be carried out with enhancing screens and ^{35}S -labeled proteins with fluorography.

For nonradiolabeled samples, use 0.2 to 1 ml of precleared lysate.

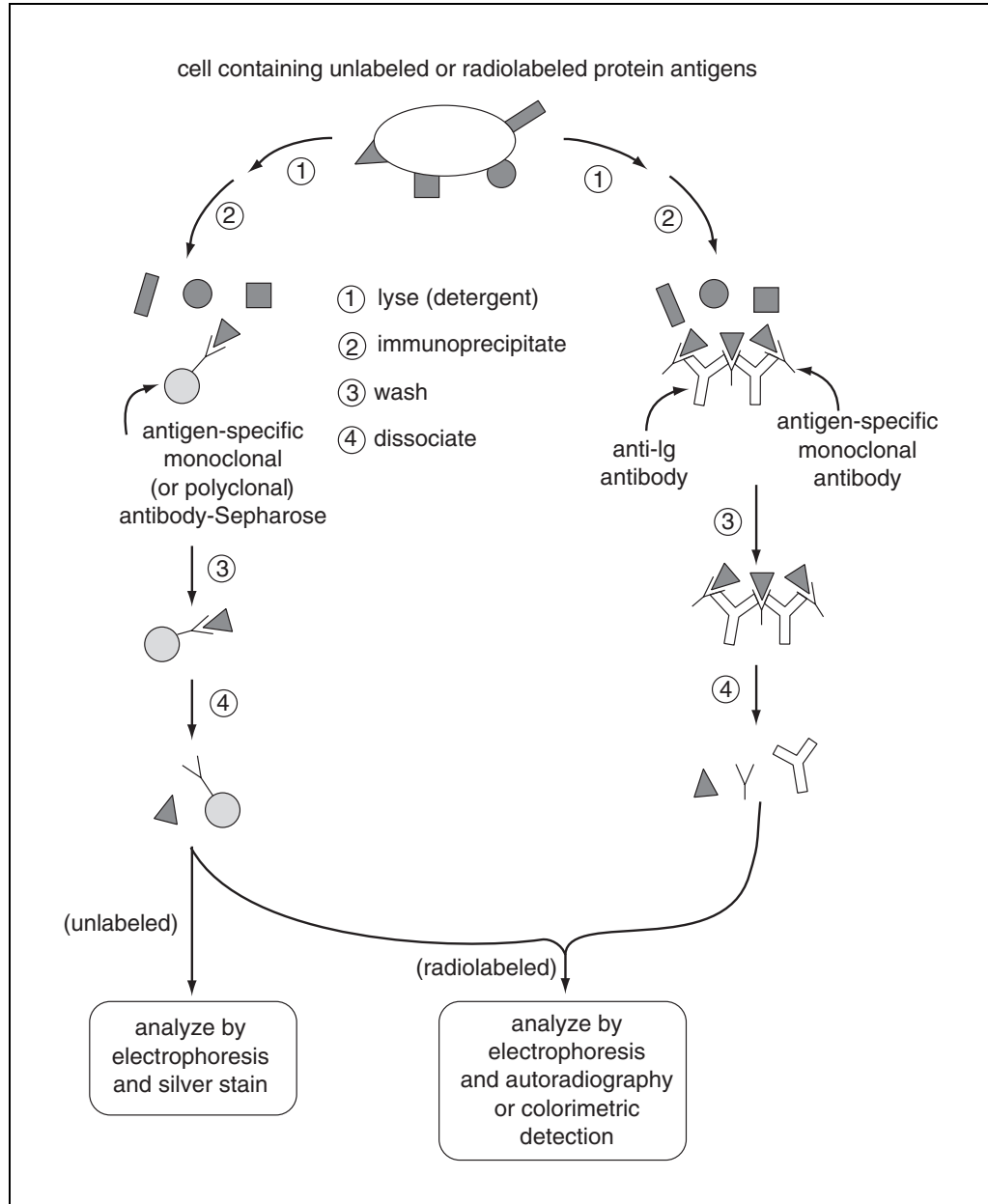


Figure 10.16.2 Schematic representation of the stages of the immunoprecipitation protocols using either antibody-Sepharose (left, see Alternate Protocol 5) or anti-Ig serum (right, see Alternate Protocol 6). (1) Cell lysis. (2) Immunoprecipitation using specific antibodies coupled covalently to Sepharose beads (left) or specific antibodies combined with anti-Ig serum (right). (3) Washing. (4) Dissociation of the antigen-antibody complex in sample buffer for electrophoresis.

7. Add ~10 μ l of a 1:1 slurry of Ab-Sepharose/dilution buffer and shake 1.5 hr at 4°C on an orbital shaker.

The antibody coupled to Sepharose is antigen specific. As described in the following support protocol, 5 mg/ml antibody per milliliter Sepharose is coupled, and the amount actually coupled can be estimated as described in step 10 of the support protocol. Shaking must be vigorous enough to suspend the Sepharose. Shaking may be extended to 3 hr; longer periods may increase background.

Wash, dissociate, and analyze the immunoprecipitate

8. Wash the Ab-Sepharose with 1 ml of the buffers listed below. After each wash, centrifuge 1 min at 200 \times g or microcentrifuge 5 sec. Then, carefully aspirate the supernatant with a fine-tipped Pasteur pipet and leave 10 μ l of fluid above the pellet. After the fourth wash, centrifuge again to bring down any residual drops on the side of the tube, aspirate, and leave 10 μ l over the pellet.

First wash: dilution buffer

Second wash: dilution buffer

Third wash: TSA solution

Fourth wash: 0.05 M Tris-Cl, pH 6.8.

Prepare a fine-tipped Pasteur pipet by pulling the pipet in a flame, scoring with a diamond pen, and breaking at the score.

9. Add 20 to 50 μ l of 2 \times SDS sample buffer. Because the sample buffer has a higher density than the wash solution, it will sink into the Sepharose; do not vortex, because Sepharose may stick to side of tube above buffer level. Cap the tube securely and incubate 5 min at 100°C.
10. Vortex and centrifuge 1 min at 200 \times g or microcentrifuge 5 sec. Load the supernatant, carefully avoiding the Sepharose, into a gel lane and analyze by SDS-PAGE (UNIT 10.2).
11. Detect labeled proteins by autoradiography (APPENDIX 3A) with an enhancing screen (125 I), by fluorography (35 S, 14 C, or 3 H), or by colorimetric or chemiluminescent detection (biotinylated proteins; UNIT 3.18).

PREPARATION OF ANTIBODY-SEPHAROSE

This protocol details the procedure for covalently linking an antibody to Sepharose (an insoluble, large-pore-size chromatographic matrix) using the cyanogen bromide activation method. It is necessary to first prepare the antibody and Sepharose separately. Next, the Sepharose is activated with cyanogen bromide (alternatively, CNBr-activated Sepharose can be purchased from Amersham Pharmacia Biotech and used according to the manufacturer's instructions). Finally, the CNBr-activated Sepharose is coupled to the antibody.

Materials

- 1 to 30 mg/ml antigen-specific monoclonal or polyclonal antibody
- 0.1 M NaHCO₃/0.5 M NaCl
- Sepharose CL-4B (or Sepharose CL-2B for high-molecular-weight antigens; Amersham Pharmacia Biotech)
- 0.2 M Na₂CO₃
- Cyanogen bromide (CNBr)/acetonitrile (see recipe)
- 1 mM and 0.1 mM HCl, ice-cold
- 0.05 M glycine (or ethanolamine), pH 8.0
- Tris/saline/azide (TSA) solution (see recipe)

SUPPORT PROTOCOL

Analysis of Proteins

10.16.13

Dialysis tubing (molecular weight cutoff >10,000)
Whatman no. 1 filter paper
Buchner funnel
Erlenmeyer filtration flask
Water aspirator

Prepare the antibody

1. Dialyze 1 to 30 mg/ml antibody against 0.1 M NaHCO₃/0.5 M NaCl at 4°C with three buffer changes during 24 hr. Use a volume of dialysis solution that is 500 times the volume of antibody solution.

Dialysis is performed to remove all small molecules containing free amino or sulfhydryl groups (see APPENDIX 3C).

2. Centrifuge 1 hr at 100,000 × g, 4°C, to remove aggregates. Save the supernatant.

Removal of aggregates is important. Because only some of the antibody molecules in an aggregate will be directly coupled to the Sepharose, the noncoupled antibody molecules may leach out during elution.

3. Measure the A₂₈₀ of an aliquot of the solution and determine the concentration of the antibody (mg/ml IgG = A₂₈₀/1.44). Dilute with 0.1 M NaHCO₃/0.5 M NaCl to 5 mg/ml (or to the same concentration as desired for Ab-Sepharose) and keep at 4°C. Measure the A₂₈₀ of this solution for later use in step 11.

Prepare the Sepharose

4. Allow the Sepharose slurry to settle in a beaker and decant and discard the supernatant. Weigh out the desired quantity of Sepharose (assume density = 1.0).

5. Set up a filter apparatus using Whatman no. 1 filter paper in a Buchner funnel and an Erlenmeyer filtration flask attached to a water aspirator. Wash the Sepharose on the filter apparatus with 10 vol water.

Sintered-glass funnels are traditionally recommended but rapidly become clogged unless coarse-porosity funnels are used.

Activate Sepharose with cyanogen bromide

6. Transfer Sepharose to 50-ml beaker and add an equal volume of 0.2 M Na₂CO₃.
7. Activate Sepharose at room temperature using 3.2 ml CNBr/acetonitrile per 100 ml Sepharose. Add CNBr/acetonitrile dropwise with a Pasteur pipet over 1 min, while slowly stirring the slurry with a magnetic stirrer. Continue stirring slowly for 5 min.

Excessive and vigorous stirring may fracture the Sepharose beads. The protocol uses 2 g CNBr/100 ml Sepharose. Two to four grams of CNBr/100 ml Sepharose can be used to couple 1 to 20 mg of antibody/ml Sepharose.

CAUTION: *Activation should be carried out in a fume hood.*

8. Rapidly filter the CNBr-activated Sepharose as in step 5. Aspirate to semidryness (i.e., until the Sepharose cake cracks and loses its sheen).
9. Wash with 10 vol ice-cold 1 mM HCl, then with 2 vol of ice-cold 0.1 mM HCl. Hydrate the cake with enough ice-cold 0.1 mM HCl so the cake regains its sheen, but so there is no excess liquid above the cake.

Washing is most efficient if the wash solution is added evenly over the surface of the cake at about the same rate as the solution is removed by filtration. CNBr-activated Sepharose is very unstable at the alkaline pH necessary for activation; it is much more stable in dilute HCl. CNBr-activated Sepharose can be purchased premade from Amersham Pharmacia Biotech, but the coupling capacity will be lower.

Couple antibody to CNBr-activated Sepharose

10. Immediately transfer a weighed amount of Sepharose (assume density = 1.0) to a beaker. Add an equal volume of a solution of antibody dissolved in 0.1 M NaHCO₃/0.5 M NaCl (from step 2). Stir gently with a magnetic stirrer or rotate end over end 2 hr at room temperature or overnight at 4°C.
11. Add 0.05 M glycine (or ethanolamine), pH 8.0, to saturate the remaining reactive groups on the Sepharose and allow the slurry to settle. Remove an aliquot of the supernatant, centrifuge to remove any residual Sepharose, and measure A_{280} . Compare absorbance to that of the A_{280} of the antibody solution from step 2 to determine the percentage coupling.
12. Store the Ab-Sepharose in TSA solution.

IMMUNOPRECIPITATION OF RADIOLABELED ANTIGEN WITH ANTI-Ig SERUM

This protocol relies on the formation of soluble immune complexes between a protein and an antigen-specific antibody, followed by immunoprecipitation of the immune complexes by antibodies contained in anti-immunoglobulin (Ig) serum. This procedure is usually only used with radiolabeled or biotinylated antigen, as the unlabeled antibody remains in the precipitate and greatly complicates the use of any other detection method.

Additional Materials (also see Alternate Protocol 5)

- Normal serum
- Anti-Ig serum (Zymed Laboratories)
- Antigen-specific antiserum *or* antigen-specific purified monoclonal antibody *or* antigen-specific hybridoma culture supernatant

Follow the procedures in Alternate Protocol 5, with the following modifications at the indicated steps:

- 4a. Preclarify by adding normal serum at a concentration of 2 µl/ml radiolabeled antigen. Add the proper amount of anti-Ig serum and let stand 12 to 18 hr at 4°C. Centrifuge 10 min at 1000 × g and reserve supernatant.

Normal serum is the source of carrier Ig. The proper amount of anti-Ig serum must be determined by titration with radiolabeled antigen or Ig. For high-titered anti-Ig serum, this amount would be 20× to 40× the volume of antigen-specific antiserum, 2 to 4 µl/µg purified MAb, or one-third the volume of hybridoma culture supernatant.

- 7a. Add 1 µl antigen-specific antiserum, 3 µg antigen-specific purified MAb, or antigen-specific hybridoma culture supernatant (30 µl cloned line or 100 µl uncloned line). Vortex and allow to stand 2 hr at 4°C. Then add the proper amount of anti-Ig serum, vortex, and allow to stand 12 to 18 hr at 4°C.
- 8a. Wash the immunoprecipitate (see Alternate Protocol 5, step 8), except centrifuge 7 min at 1000 × g.
- 9a. Add 20 to 50 µl of 2× SDS sample buffer. Do not vortex, as immunoprecipitates may stick to side of tube above buffer level. Cap the tube securely. For immunoprecipitates, incubate first 1 hr at 56°C and then 5 min at 100°C.

The initial 56°C incubation enhances the dissolution of the immunoprecipitates by reducing irreversible aggregation which occurs when precipitated protein is rapidly heated to 100°C. Proteolytic degradation has never been noted, probably because of the high IgG protein concentration.

ALTERNATE PROTOCOL 6

Analysis of Proteins

10.16.15

IMMUNOPRECIPITATION-RECAPTURE

Once an antigen has been isolated by immunoprecipitation, it can be dissociated from the beads and reimmunoprecipitated (“recaptured”) either with the same antibody used in the first immunoprecipitation or with a different antibody (Fig. 10.16.3). Immunoprecipitation-recapture with the same antibody allows identification of a specific antigen in cases where the first immunoprecipitation contains too many bands to allow unambiguous identification. By using a different antibody in the second immunoprecipitation, immunoprecipitation-recapture can be used to analyze the subunit composition of multi-protein complexes (Fig. 10.16.4). The feasibility of this approach depends on the ability of the second antibody to recognize denatured antigens.

Dissociation of the antigen from the beads is achieved by denaturation of antigen-antibody-bead complexes at high temperature in the presence of SDS and DTT. Prior to recapture, the SDS is diluted in a solution containing Triton X-100, and the DTT is neutralized with excess iodoacetamide. Recapture is then performed as in the first immunoprecipitation (see Basic Protocol 1, steps 6 to 26).

Materials

- Elution buffer (see recipe)
- Beads containing bound antigen (see Basic Protocol 1, step 26)
- 10% (w/v) BSA
- Nondenaturing lysis buffer (see recipe)
- Heating block set at 95°C (Eppendorf Thermomixer 5436 or equivalent)

1. Add 50 μ l elution buffer to 15 μ l beads containing bound antigen. Mix by vortexing.
The DTT in the elution buffer reduces disulfide bonds in the antigen and the antibody, and the SDS contributes to the unfolding of polypeptide chains.
2. Incubate 5 min at room temperature and 5 min at 95°C in a heating block. Cool tubes to room temperature.
3. Add 10 μ l of 10% BSA. Mix by gentle vortexing.
BSA is added to prevent adsorption of antigen to the tube, and to quench nonspecific binding to antibody-conjugated beads.
4. Add 1 ml nondenaturing lysis buffer.
The iodoacetamide in the nondenaturing lysis buffer reacts with the DTT and prevents it from reducing the antibody used in the recapture steps. The presence of PMSF and leupeptin in the buffer is not necessary at this step.
5. Incubate 10 min at room temperature.
6. Clear the lysate and perform second immunoprecipitation (see Basic Protocol 1, steps 6 to 26).

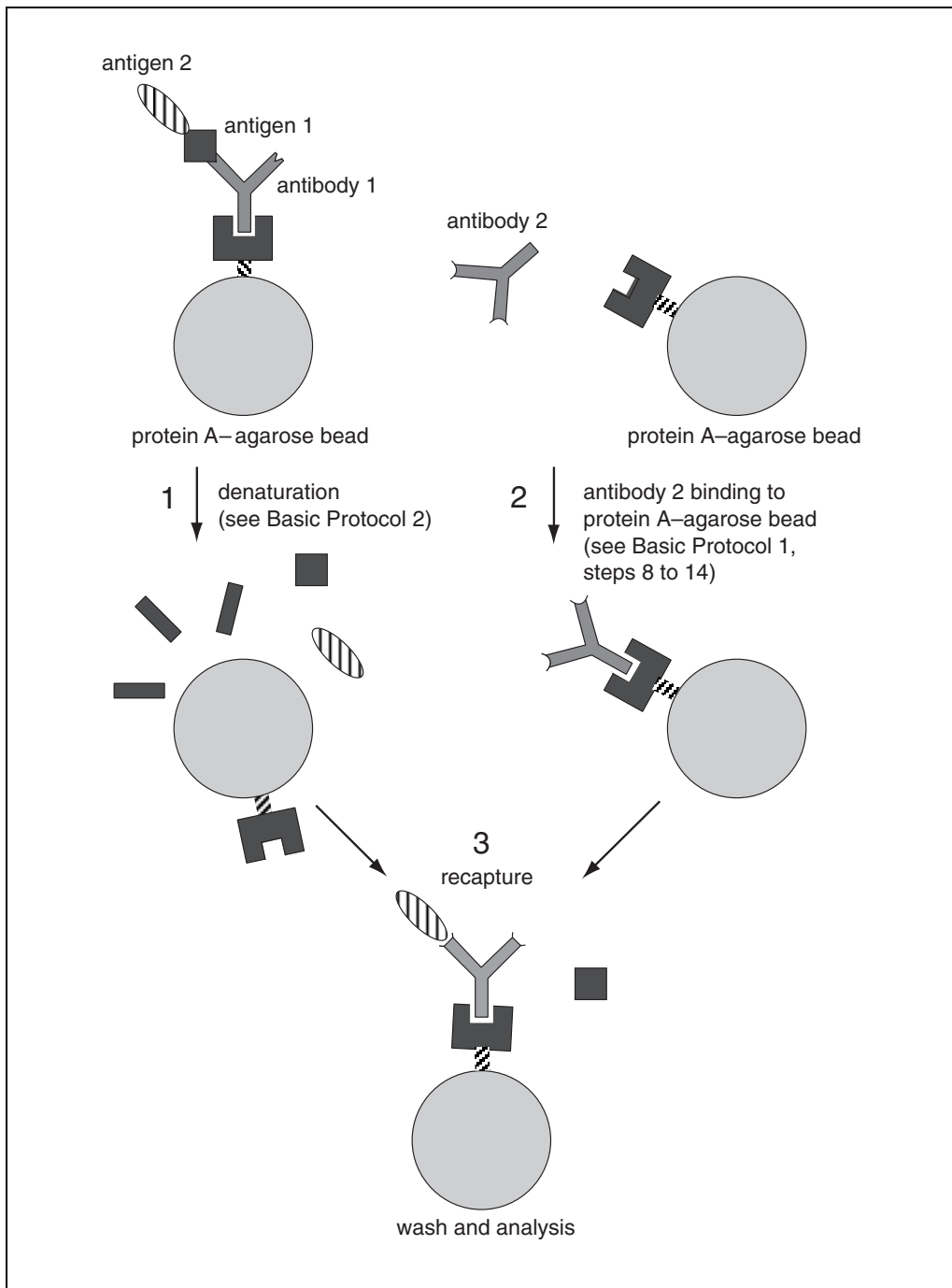


Figure 10.16.3 Scheme showing the stages of immunoprecipitation-recapture. (1) Dissociation and denaturation of the antigen: an antigen immunoprecipitated with antibody 1 bound to protein A-agarose beads is dissociated and denatured by heating in the presence of SDS and DTT. (2) Immobilization of the second antibody: antibody 2 is bound to protein A-agarose beads. (3) Recapture: the denatured antigen 2 (striped oval) is recaptured on antibody 2 bound to protein A-agarose beads. Alternatively, antibody 1 can be used again for further purification of the original antigen (square).

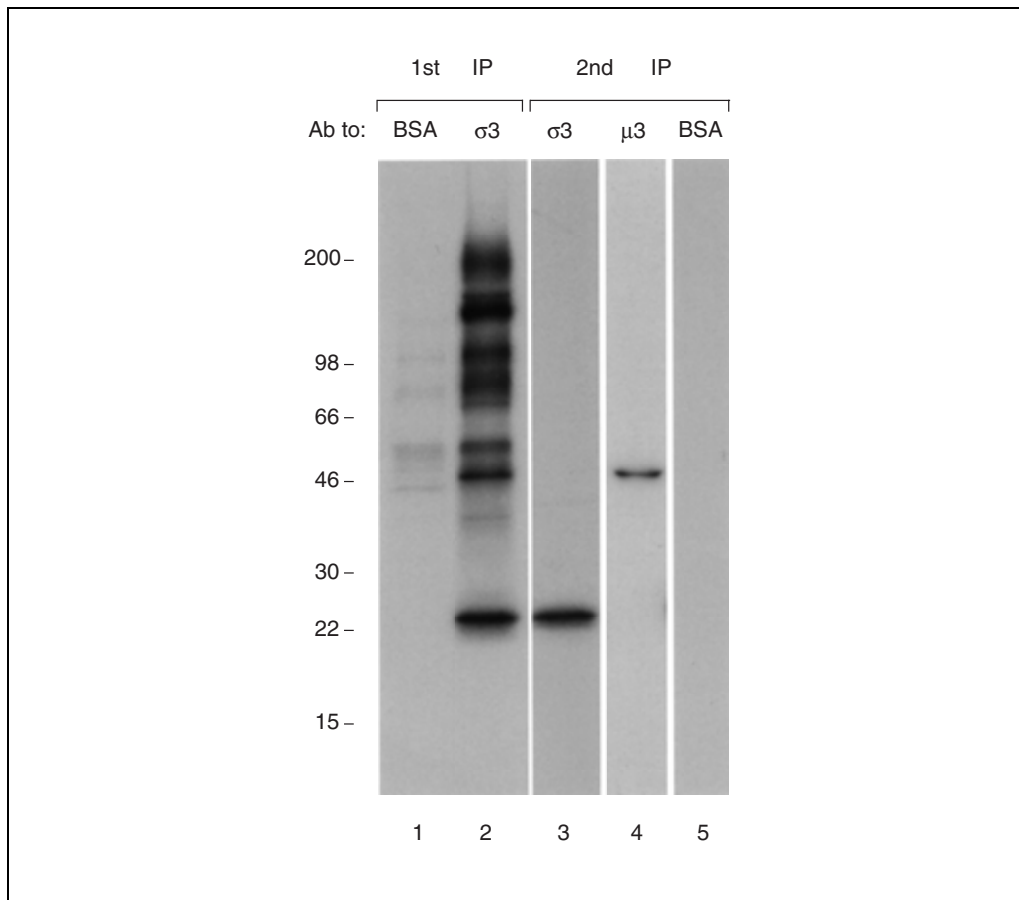


Figure 10.16.4 Example of an immunoprecipitation-recapture experiment. Human M1 fibroblasts were labeled overnight with [³⁵S]methionine (*UNIT 10.18*) and extracted with nondenaturing lysis buffer (see Basic Protocol 1). The cell extract was then subjected to immunoprecipitation with antibodies to BSA (irrelevant antibody control; lane 1) and to the AP-3 adaptor ($\sigma 3$; lane 2), a protein complex involved in protein sorting. Notice the presence of several specific bands in lane 2. The AP-3 immunoprecipitate was denatured as described in Basic Protocol 2 and individual components of the AP-3 complex were recaptured with antibodies to two of its subunits: $\sigma 3$ ($M_r \sim 22,000$; lane 3) and $\mu 3$ ($M_r \sim 47,000$; lane 4). An immunoprecipitation with an antibody to BSA was also performed as a nonspecific control (lane 5). The amount of immunoprecipitate loaded on lanes 1 and 2 is $\sim 1/10$ the amount loaded on lanes 3 to 5. Notice the presence of single bands in lanes 3 and 4. The positions of M_r standards (expressed as $10^{-3} \times M_r$) are shown at left. IP, immunoprecipitation.

REAGENTS AND SOLUTIONS

Use deionized or distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

CNBr/acetoneitrile

To 25 g of cyanogen bromide (*CNBr should be white, not yellow, crystals*), add 50 ml acetoneitrile to make a 62.5% (w/v) solution. This may be stored indefinitely at -20°C in a desiccator over silica. Allow to warm before opening.

CAUTION: *CNBr is a highly toxic lachrymator; handle in fume hood.*

Denaturing lysis buffer

1% (w/v) SDS

50 mM Tris·Cl, pH 7.4 (*APPENDIX 2*)

5 mM EDTA (*APPENDIX 2*)

Store up to 1 week at room temperature (SDS precipitates at 4°C)

Add the following fresh before use:

10 mM dithiothreitol (DTT; from powder)

1 mM phenylmethylsulfonyl fluoride (PMSF; store 100 mM stock in 100% ethanol up to 6 months at –20°C)

2 µg/ml leupeptin (store 10 mg/ml stock in H₂O up to 6 months at –20°C)

15 U/ml DNase I (store 15,000 U/ml stock solution up to 2 years at –20°C)

1 mM 4-(2-aminoethyl)benzenesulfonyl fluoride (AEBSF), added fresh from a 0.1 M stock solution in H₂O, can be used in place of PMSF. AEBSF stock can be stored up to 1 year at –20°C.

Detergent-free lysis buffer

PBS (*APPENDIX 2*) containing:

5 mM EDTA (*APPENDIX 2*)

0.02% (w/v) sodium azide

Store up to 6 months at 4°C

Immediately before use add:

10 mM iodoacetamide (from powder)

1 mM PMSF (store 100 mM stock in 100% ethanol up to 6 months at –20°C)

2 µg/ml leupeptin (store 10 mg/ml stock in H₂O up to 6 months at –20°C)

1 mM AEBSF, added fresh from a 0.1 M stock solution in H₂O, can be used in place of PMSF. AEBSF stock can be stored up to 1 year at –20°C.

Dilution buffer

TSA solution (see recipe) containing:

0.1% Triton X-100 (store at room temperature in dark)

0.1% bovine hemoglobin (store frozen)

Elution buffer

1% (w/v) SDS

100 mM Tris·Cl, pH 7.4 (*APPENDIX 2*)

Store up to 1 week at room temperature

10 mM DTT (add fresh from powder before use)

Nondenaturing lysis buffer

1% (w/v) Triton X-100 (store at room temperature in dark)

50 mM Tris·Cl, pH 7.4 (*APPENDIX 2*)

300 mM NaCl

5 mM EDTA (*APPENDIX 2*)

0.02% (w/v) sodium azide

Store up to 6 months at 4°C

Immediately before use add:

10 mM iodoacetamide (from powder)

1 mM PMSF (store 100 mM stock in 100% ethanol up to 6 months at –20°C)

2 µg/ml leupeptin (store 10 mg/ml stock in H₂O up to 6 months at –20°C)

1 mM AEBSF, added fresh from a 0.1 M stock solution in H₂O, can be used in place of PMSF. AEBSF stock can be stored up to 1 year at –20°C.

Tris/saline/azide (TSA) solution

10 mM Tris·Cl, pH 8.0 (APPENDIX 2)

140 mM NaCl

0.025% NaN₃

CAUTION: Sodium azide (NaN₃) is poisonous; wear gloves.

Triton X-100 lysis buffer

TSA solution (see recipe) containing:

1% Triton X-100 (store at room temperature in dark)

1% bovine hemoglobin (store frozen)

1 mM iodoacetamide (from powder)

Aprotinin (0.2 trypsin inhibitor U/ml)

1 mM PMSF (store 100 mM stock in 100% ethanol up to 6 months at -20°C)

Prepare fresh

1 mM AEBSF, added fresh from a 0.1 M stock solution in H₂O, can be used in place of PMSF.

AEBSF stock can be stored up to 1 year at -20°C.

Wash buffer

0.1% (w/v) Triton X-100 (store at room temperature in dark)

50 mM Tris·Cl, pH 7.4 (APPENDIX 2)

300 mM NaCl

5 mM EDTA (APPENDIX 2)

0.02% (w/v) sodium azide

Store up to 6 months at 4°C

COMMENTARY

Background Information

The use of antibodies for immunoprecipitation has its origin in the *precipitin* reaction (Nisonoff, 1984, and references therein). The term precipitin refers to the spontaneous precipitation of antigen-antibody complexes formed by interaction of certain polyclonal antibodies with their antigens. The precipitation arises from formation of large networks of antigen-antibody complexes, due to the bivalent or polyvalent nature of immunoglobulins and to the presence of two or more epitopes in some antigens. This phenomenon was quickly exploited to isolate antigens from protein mixtures; however, its use remained limited to antibodies and antigens that were capable of multivalent interaction. In addition, the efficiency of precipitate formation was highly dependent on the concentrations of antibody and antigen. Thus, the precipitin reaction was not generally applicable as a method for immunoprecipitation. A significant improvement was the use of secondary anti-immunoglobulin reagents (generally anti-immunoglobulin serum) to cross-link the primary antibodies, thus promoting the formation of a precipitating network.

In the 1970s, immunoprecipitation became widely applicable to the study of cellular antigens as a result of several technological advances. A critical development was the introduction of methods for the production of monoclonal antibodies (Köhler and Milstein, 1975). The ability to produce unlimited amounts of antibodies with specificity against virtually any cellular antigen had a profound impact in many areas of biology and medicine. The fact that preparation of monoclonal antibodies did not require prior purification of the antigens accelerated the characterization of cellular proteins and organelles, a process in which immunoprecipitation protocols played a major role. To this day, monoclonal antibodies produced in mice or rats continue to be among the most useful tools in biology.

Another important development was the discovery of bacterial Fc receptors, proteins found on the surface of bacteria that have the property of binding a wide range of immunoglobulins. Two of the most widely used bacterial Fc receptors are protein A from *Staphylococcus aureus* (Kessler, 1975) and protein G from group G streptococci. Protein A and protein G bind both polyclonal and monoclonal antibodies.

ies belonging to different subclasses and deriving from different animal species (Table 10.16.1). Protein A was initially used to adsorb immunoglobulins as part of fixed, killed *S. aureus* particles. Both protein A and protein G are now produced in large quantities by recombinant DNA procedures and are available coupled to solid-phase matrices such as agarose. In most cases, the binding of polyclonal or monoclonal antibodies to immobilized protein A (or G) avoids the need to use a secondary antibody to precipitate antigen-antibody complexes. Because of their broad specificity and ease of use, protein A-agarose and protein G-agarose (and related products) are the state-of-the-art reagents for the isolation of soluble antigen-antibody complexes in immunoprecipitation protocols.

Recent progress in the field of antibody engineering (reviewed by Rapley, 1995; Irving et al., 1996) promises to make antibody production a less time-consuming and haphazard process. Antibody fragments with high affinity for specific antigens can now be selected from phage display antibody libraries. Selected recombinant antibodies can then be produced in large quantities in *Escherichia coli*. Techniques have been developed for producing antibodies in soluble, secreted form. Affinity tags are added to the recombinant antibody molecules to facilitate purification, detection, and use in procedures such as immunoprecipitation. While attractive in principle, the production of recombinant antibodies has been plagued by technical difficulties that so far have limited their widespread use in biology. However, as technical problems are overcome, recombinant techniques will progressively replace immunization of animals as a way of producing antibodies for immunoprecipitation and for other applications.

Critical Parameters

Extraction of antigens

Isolation of cellular antigens by immunoprecipitation requires extraction of the cells so that the antigens are available for binding to specific antibodies, and are in a physical form that allows separation from other cellular components. Extraction with nondenaturing detergents such as Triton X-100 (see Basic Protocol 1, Alternate Protocol 1, and Alternate Protocol 5) or in the absence of detergent (see Alternate Protocol 3) allows immunoprecipitation with

antibodies to epitopes that are exposed on native proteins. Other nondenaturing detergents such as Nonidet P-40, CHAPS, digitonin, or octyl glucoside are also appropriate for extraction of native proteins (APPENDIX II). Some of these detergents (e.g., digitonin) preserve weak protein-protein interactions better than Triton X-100. If the antigen is part of a complex that is insoluble in nondenaturing detergents (e.g., cytoskeletal structures, chromatin, membrane “rafts”) or if the epitope is hidden within the folded structure of the protein, extraction under denaturing conditions is indicated (see Alternate Protocol 2). Alternate Protocol 2 may also be indicated for in vitro-translated products, which often tend to form aggregates (Anderson and Blobel, 1983).

The number of cells necessary to detect an immunoprecipitated antigen depends on the cellular abundance of the antigen and on the efficiency of radiolabeling. The protocols for radiolabeling (UNIT 10.18) and immunoprecipitation described in this book are appropriate for detection of antigens that are present at low to moderate levels (10,000 to 100,000 copies per cell), as is the case for most endogenous integral membrane proteins, signal transduction proteins, and transcription factors. For more abundant antigens, such as cytoskeletal and secretory proteins or proteins that are expressed by viral infection or transfection, the quantity of radiolabeled cells used in the immunoprecipitation can be reduced accordingly.

Production of antibodies

Immunoprecipitation can be carried out using either polyclonal or monoclonal antibodies (see discussion of selection below). Polyclonal antibodies are most often prepared by immunizing rabbits, although polyclonal antibodies produced in mice, guinea pigs, goats, sheep, and other animals are also suitable for immunoprecipitation. Antigens used for polyclonal antibody production can be whole proteins purified from cells or tissues, or whole or partial proteins produced in bacteria or insect cells by recombinant DNA procedures. Another useful procedure is to immunize animals with peptides conjugated to a carrier protein. Production of polyclonal antibodies to recombinant proteins and peptides has become the most commonly used approach to obtain specific probes for immunoprecipitation and other immunochemical techniques, because it does not require purification of protein antigens from their

native sources. The only requirement for making these antibodies is knowledge of the sequence of a protein, which is now relatively easy to obtain as a result of cDNA library production and genomic DNA sequencing projects. Polyclonal antibodies can be used for immunoprecipitation in the form of whole serum, ammonium sulfate-precipitated immunoglobulin fractions, or affinity-purified immunoglobulins. Although all of these are suitable for immunoprecipitation, affinity-purified antibodies often give lower backgrounds and are more specific.

Most monoclonal antibodies are produced in mice or rats. The sources of antigen for monoclonal antibody production are the same as those for production of polyclonal antibodies, namely proteins isolated from cells or tissues, recombinant proteins or protein fragments, and peptides. A significant advantage of using monoclonal antibodies is that antigens do not need to be purified to serve as immunogens, as long as the screening method is specific for the antigen. Another advantage is the unlimited supply of monoclonal antibodies afforded by the ability to grow hybridomas in culture or in ascitic fluid. Many monoclonal antibodies can now be produced from hybridomas deposited in cell banks or are directly available commercially. Ascitic fluid, cell culture supernatant, and purified antibodies are all suitable sources of monoclonal antibodies for immunoprecipitation. Ascitic fluid and purified antibodies should be used when a high antibody titer is important. Cell culture supernatants have lower antibody titers, but tend to give cleaner immunoprecipitations than ascitic fluids due to the lack of contaminating antibodies.

Selection of antibodies: Polyclonal versus monoclonal

What type of antibody is best for immunoprecipitation? There is no simple answer to this question, as the outcome of both polyclonal and monoclonal antibody production protocols is still difficult to predict. Polyclonal antibodies to whole proteins (native or recombinant) have the advantage that they frequently recognize multiple epitopes on the target antigen, enabling them to generate large, multivalent immune complexes. Formation of these antigen-antibody networks enhances the avidity of the interactions and increases the efficiency of immunoprecipitation. Because these antibodies recognize several epitopes, there is a better

chance that at least one epitope will be exposed on the surface of a solubilized protein and thus be available for interaction with antibodies. Thus, the likelihood of success is higher. These properties can be a disadvantage, though, as some polyvalent antibodies can cross-react with epitopes on other proteins, resulting in higher backgrounds and possible misidentification of antigens. Because they are directed to a short peptide sequence, anti-peptide polyclonal antibodies are less likely to cross-react with other proteins. However, their usefulness is dependent on whether the chosen sequence turns out to be a good immunogen in practice, as well as on whether this particular epitope is available for interaction with the antibody under the conditions used for immunoprecipitation.

Unfractionated antisera are often suitable for immunoprecipitation. However, there is a risk that serum proteins other than the antibody will bind nonspecifically to the immunoabsorbent, and in turn bind proteins in the lysate that are unrelated to the antigen. For instance, transferrin can bind nonspecifically to immunoabsorbents, potentially leading to the isolation of the transferrin receptor as a contaminant (Harford, 1984). Polyclonal antisera can also contain antibodies to other antigens (e.g., viruses, bacteria) to which the animal may have been exposed, and these antibodies can also cross-react with cellular proteins during immunoprecipitation. Affinity-purified antibodies are a better alternative when antisera do not yield clean immunoprecipitations. Affinity-purification can lead to loss of high-affinity or low-affinity antibodies; however, the higher specificity of affinity-purified antibodies generally makes them “cleaner” reagents for immunoprecipitation.

The specificity, high titer, and limitless supply of the best immunoprecipitating monoclonal antibodies are unmatched by those of polyclonal antibodies. However, not all monoclonal antibodies are useful for immunoprecipitation. Low-affinity monoclonal antibodies can perform acceptably in immunofluorescence microscopy protocols (UNIT 14.10) but may not be capable of holding on to the antigen during the repeated washes required in immunoprecipitation protocols. The use of ascitic fluid has the same potential pitfalls as the use of polyclonal antisera, as ascites may also contain endogenous antibodies to other antigens

and proteins such as transferrin that can bind to other proteins in the lysate.

In conclusion, an informed empirical approach is recommended in order to select the best antibody for immunoprecipitation. In general, it is advisable to generate and/or test several antibodies to a particular antigen in order to find at least one that will perform well in immunoprecipitation protocols.

Antibody titer

The importance of using the right amount of antibody for immunoprecipitation cannot be overemphasized. This is especially the case for quantitative immunoprecipitation studies, in which the antibody should be in excess of the specific antigen. For instance, in pulse-chase analyses of protein degradation or secretion (UNIT 10.18), it is critical to use sufficient antibody to deplete the antigen from the cell lysate. This is particularly important for antigens that are expressed at high levels, a common occurrence with the growing use of high-yield protein expression systems such as vaccinia virus (UNITS 16.15 - 16.19) or replicating plasmids in COS cells (UNIT 16.12). Consider, for example, a protein that is expressed at high levels inside the cell, and of which only a small fraction is secreted into the medium. If limiting amounts of antibody are used in a pulse-chase analysis of this protein, the proportion of protein secreted into the medium will be grossly overestimated, because the limiting antibody will bind only a small proportion of the cell-associated protein and a much higher proportion of the secreted protein. The same considerations apply to degradation studies. Thus, it is extremely important in quantitative studies to ensure that the antibody is in excess of the antigen in the cell samples. This can be ascertained by performing sequential immunoprecipitations of the samples (see Basic Protocol 1, annotation to step 21). If the second immunoprecipitation yields only a small amount of the antigen relative to that isolated in the first immunoprecipitation (<10%), then the antibody titer is appropriate. If, on the other hand, the amount of antigen isolated in the second immunoprecipitation is >10%, either more antibody or less antigen should be used.

Too much antibody can also be a problem, as nonspecific immunoprecipitation tends to increase with increasing amounts of immunoglobulins bound to the beads. Thus, titration of

the antibody used for immunoprecipitation is strongly advised.

Immunoabsorbent

If cost is not an overriding issue, the use of protein A- or protein G-agarose is recommended for routine immunoprecipitation (see Basic Protocol 1). Protein A- or protein G-agarose beads (or equivalent products) have a very high capacity for antibody binding (up to 10 to 20 mg of antibody per milliliter of gel). Both protein A and protein G bind a wide range of immunoglobulins (Table 10.16.1). Backgrounds from nonspecifically bound proteins are generally low. Protein A- and protein G-agarose beads are also stable and easy to sediment by low-speed centrifugation. A potential disadvantage, in addition to their cost, is that some polyclonal or monoclonal antibodies bind weakly or not at all to protein A or protein G (Table 10.16.1). This problem can be solved by using an intermediate rabbit antibody to the immunoglobulin of interest. For example, a goat polyclonal antibody can be indirectly bound to protein A-agarose by first incubating the protein A-agarose beads with a rabbit anti-goat immunoglobulin, and then incubating the beads with the goat polyclonal antibody. Anti-immunoglobulin antibodies (e.g., rabbit anti-goat immunoglobulins) coupled covalently to agarose can also be used for indirect immunoprecipitation in place of protein A- or protein G-agarose.

A less expensive alternative to protein A- or protein G-agarose is the use of anti-Ig serum to crosslink the primary antibody (see Alternate Protocol 6). This procedure can result in very low backgrounds, although it requires proper titration of the anti-Ig serum. Protein A-agarose can also be substituted by fixed *Staphylococcus aureus* particles (Pansorbin). They have a lower capacity, can give higher backgrounds, and take longer to sediment. However, they work quite well in many cases. In order to establish if they are appropriate for a particular experimental setup, conduct a preliminary comparison of the efficiency of protein A-agarose with *Staphylococcus aureus* particles as immunoabsorbent.

Specific antibodies coupled covalently to various affinity matrices can also be used for direct immunoprecipitation of antigens (see Alternate Protocol 5). After binding to protein A-agarose, antibodies can be cross-linked with dimethylpimelimidate (Gersten and Marchalo-

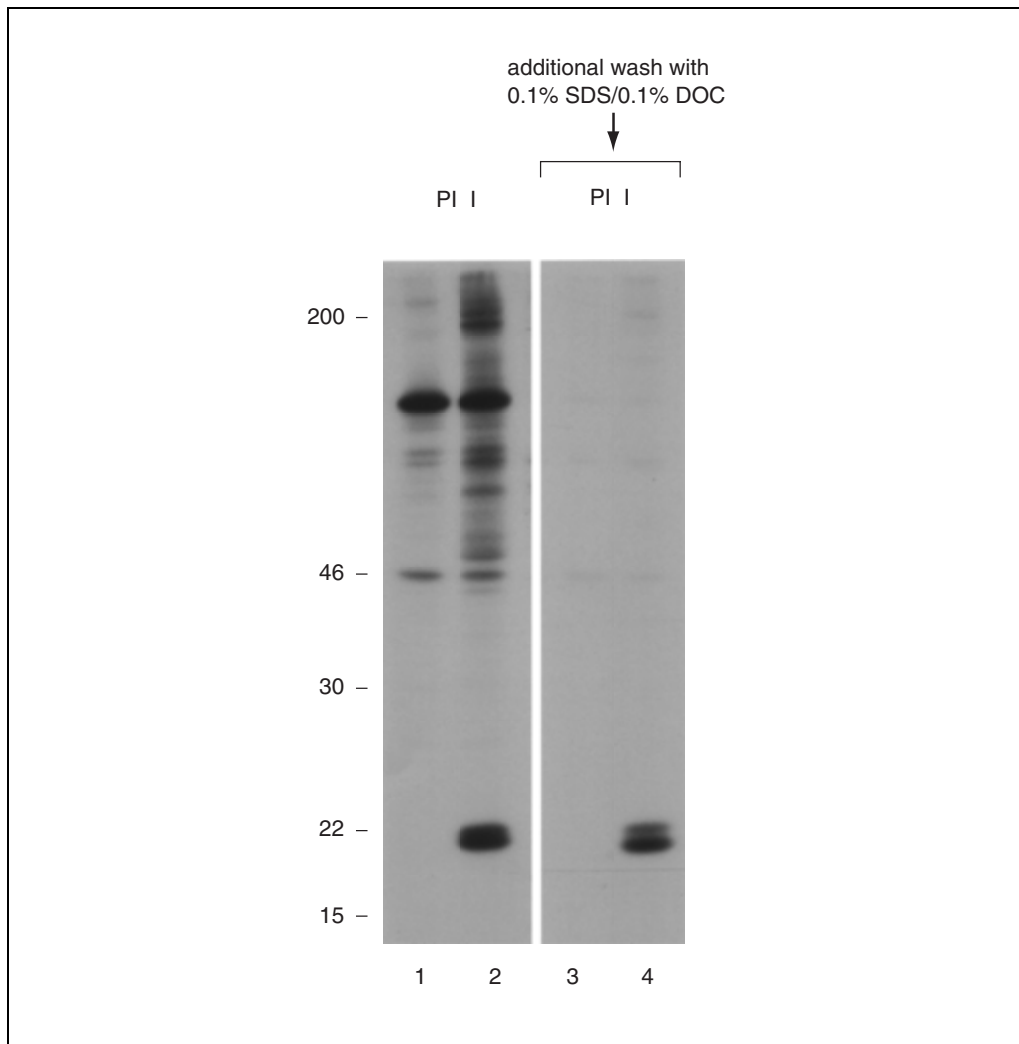


Figure 10.16.5 Lowering background by washing with SDS and sodium deoxycholate (DOC). In this experiment, BW5147 cells (mouse thymoma) labeled with [³⁵S]methionine for 1 hr were extracted with nondenaturing lysis buffer (see Basic Protocol 1). The extracts were subjected to immunoprecipitation with protein A–agarose beads incubated with either preimmune (PI) or immune (I) serum from a rabbit immunized with the ribosomal protein L17 (doublet at $M_r \sim 22,000$). Lanes 1 and 2 correspond to immunoprecipitates obtained using the protocols described in this unit. Notice the presence of nonspecific bands and/or associated proteins in lane 2. Lanes 3 and 4 correspond to beads that were washed an additional time with 0.1% (w/v) SDS and 0.1% (w/v) DOC. Notice the disappearance of most of the nonspecific bands and/or associated proteins. The positions of M_r standards (expressed as $10^{-3} \times M_r$) are shown at left.

nis, 1978). Purified antibodies can also be coupled directly to derivatized matrices such as CNBr-activated Sepharose (see Support Protocol). This latter approach avoids having to bind the antibody to protein A–agarose. Covalently bound antibodies should be used when elution of immunoglobulins from the beads complicates further analyses of the complexes. This is the case when proteins in immunoprecipitates

are analyzed by one- or two-dimensional gel electrophoresis (UNITS 10.2-10.4) followed by Coomassie blue or silver staining, or are used for microsequencing. Also, the released immunoglobulins could interfere with detection of some antigens by immunoblotting (UNIT 10.8) following immunoprecipitation.

The best results in terms of complete precipitation and lower backgrounds are obtained

when an antigen-specific antibody is coupled directly to Sepharose. Sepharose CL-4B appears to give a lower background than CL-2B, perhaps because aggregates are better excluded.

The support protocol for coupling protein antigens to CNBr-activated Sepharose is a modification of the methods of Cuatrecasas (1970) and March et al. (1974). As originally described, the washing was done at alkaline pH. Because activated Sepharose is very unstable at this pH, it was originally recommended that washing, adding the protein ligand, and mixing be done in <90 sec (Cuatrecasas, 1970). However, using an acid wash buffer (Gelb, 1973) as described in the support protocol causes activated groups to remain stable for ≥ 30 min. The 0.1 M NaHCO₃ buffer in which antibody is dissolved provides an optimal pH of ~ 8.4 , after mixing with activated-Sepharose. The low amount of CNBr recommended is sufficient to obtain a coupling yield of 80% to >99%. Higher amounts of CNBr may result in multipoint attachment of IgG molecules to the matrix, thereby reducing accessibility to antigen. Most investigators purchase CNBr-activated Sepharose, while others, to achieve a higher coupling efficacy or to avoid the expense of the commercial product, prefer to prepare it themselves.

Quantities and ratios recommended in these protocols have been found to work well with several hundred monoclonal antibodies and more than 40 different antigens. However, titration of Ab-Sepharose or sandwich reagents versus the protein antigen may further optimize a given immunoprecipitation.

Nonspecific controls

For correct interpretation of immunoprecipitation results, it is critical to include appropriate nonspecific controls along with the specific samples. One type of control consists of setting up an incubation with an irrelevant antibody in the same biochemical form as the experimental antibody (e.g., serum, ascites, affinity-purified immunoglobulin, antibody bound to protein A-agarose or directly conjugated to agarose), and belonging to the same species and immunoglobulin subclass as the experimental antibody (e.g., rabbit antiserum, mouse IgG2a). For an antiserum, the best control is preimmune serum (serum from the same animal obtained before immunization). Nonimmune serum from the same species is an acceptable substitute for preimmune serum in some

cases. “No-antibody” controls are not appropriate because they do not account for nonspecific binding of proteins to immunoglobulins. In immunoprecipitation-recapture experiments, control immunoprecipitations with irrelevant antibodies should be performed for both the first and second immunoprecipitation steps (Fig. 10.16.4). Another type of control is to perform an immunoprecipitation from cells that do not express a specific antigen in parallel with immunoprecipitation of the antigen-expressing cells. For instance, untransfected cells are a perfect control for transfected cells. In yeast cells, null mutants that do not express a specific antigen are an ideal control for wild-type cells.

Order of stages

In the immunoprecipitation procedure described in Basic Protocol 1, the antibody is prebound to protein A-agarose before addition to the cell lysate containing the antigen. This differs from other methods in which the free antibody is first added to the lysate and the antigen-antibody complexes are then collected by addition of the immunoadsorbent. Although both procedures can give good results, the authors prefer the protocol described here because this method allows better control of the amount of antibody bound to the immunoadsorbent. Prebinding antibodies to the immunoadsorbent beads allows removal of unbound antibodies. The presence of unbound antibodies in the incubation mixture could otherwise result in decreased recovery of the antigen on the immunoadsorbent beads. Another advantage of the prebinding procedure is that most proteins other than the immunoglobulin in the antibody sample (e.g., serum proteins) are removed from the beads and do not come in contact with the cell lysate. This eliminates potential adverse effects of these proteins on isolation of the antigen.

Washing

The five washes described in Basic Protocol 1 (four with wash buffer and one with PBS) are sufficient for maximal removal of unbound proteins; additional washes are unlikely to decrease the background any further. The last wash with PBS removes the Triton X-100 that can lead to decreased resolution on SDS-PAGE. It also removes other components of the wash buffer that could interfere with enzymatic treatment of immunoprecipitates. It is not advisable

Table 10.16.2 Troubleshooting Guide for Immunoprecipitation

Problem	Cause	Solution
<i>No specific radiolabeled antigen band</i>		
Gel is completely blank after prolonged autoradiographic exposure	Poorly labeled cells: too little radiolabeled precursor, too few cells labeled, lysis/loss of cells during labeling, too much cold amino acid in labeling mix, wrong labeling temperature	Check incorporation of label by TCA precipitation (<i>UNIT 10.18</i>); troubleshoot the labeling procedure
Only nonspecific bands present	Antigen does not contain the amino acid used for labeling	Label cells with another radiolabeled amino acid or, for glycoproteins, with tritiated sugar
	Antigen expressed at very low levels	Substitute cells known to express higher levels of antigens as detected by other methods; transfect cells for higher expression
	Protein has high turnover rate and is not well labeled by long-term labeling	Use pulse labeling
	Protein has a low turnover rate and is not well labeled by short-term labeling	Use long-term labeling
	Protein is not extracted by lysis buffer used to solubilize cells	Solublize with a different nondenaturing detergent or under denaturing conditions
	Antibody is nonprecipitating	Identify and use antibody that precipitates antigen
	Epitope is not exposed in native antigen	Extract cells under denaturing conditions
	Antibody does not recognize denatured antigen	Extract cells under nondenaturing conditions
	Antibody does not bind to immunoadsorbent	Use a different immunoadsorbent (<i>Table 10.16.2</i>); use intermediate antibody
	Antigen is degraded during immunoprecipitation	Ensure that fresh protease inhibitors are present
<i>High background of nonspecific bands</i>		
Isolated lanes on gel with high background	Random carryover of detergent-insoluble proteins	Remove supernatant immediately after centrifugation, leaving a small amount with pellet; if resuspension occurs, recentrifuge
High background in all lanes	Incomplete washing	Cap tubes and invert several times during washes
	Poorly radiolabeled protein	Optimize duration of labeling to maximize signal-to-noise ratio
	Incomplete removal of detergent-insoluble proteins	Centrifuge lysate 1 hr at $100,000 \times g$
	Insufficient unlabeled protein to quench nonspecific binding	Increase concentration of BSA

continued

Table 10.16.2 Troubleshooting Guide for Immunoprecipitation, continued

Problem	Cause	Solution
	Antibody contains aggregates	Microcentrifuge antibody 15 min at maximum speed before binding to beads
	Antibody solution contains nonspecific antibodies	Use affinity-purified antibodies; absorb antibody with acetone extract of cultured cells that do not express antigen; for yeast cells, absorb antibody with null mutant cells
	Too much antibody	Use less antibody
	Incomplete preclearing	Preclear with irrelevant antibody of same species of origin and immunoglobulin subclass bound to immunoadsorbent
	Nonspecifically immunoprecipitated proteins	Fractionate cell lysate (e.g., ammonium sulfate precipitation, lectin absorption, or gel filtration) prior to immunoprecipitation; after washes in wash buffer, wash beads once with 0.1% SDS in wash buffer or 0.1% SDS/0.1% sodium deoxycholate
<i>Immunoprecipitating antibody detected in immunoblots</i>		
Complete immunoglobulin or heavy and/or light chains visible in immunoblot	Protein A conjugate or secondary antibody recognizes immunoprecipitating antibody	Use antibody coupled covalently to solid-phase matrix for immunoprecipitation; probe blots with primary antibody from a different species and the appropriate secondary antibody specific for immunoblotting primary antibody

to complete all the washes quickly (e.g., in 5 min), because this may not allow enough time for included proteins to diffuse out of the gel matrix. Instead, beads should be washed over ~30 min, which may require keeping the samples on ice for periods of 3 to 5 min between washes. In order to reduce nonspecific bands, samples can be subjected to an additional wash with wash buffer containing 0.1% (w/v) SDS, or with a mixture of 0.1% (w/v) SDS and 0.1% (w/v) sodium deoxycholate (Fig. 10.16.5). This wash should be done before the last wash with PBS (in Basic Protocol 1) or before the wash with 0.05 M Tris-Cl, pH 6.8 (in Alternate Protocol 5).

Troubleshooting

Two of the most common problems encountered in immunoprecipitation of metabolically labeled proteins are failure to detect specific antigens in the immunoprecipitates, and high background of nonspecifically bound proteins for antigens that were radiolabeled *in vivo* and analyzed by SDS-PAGE (UNIT 10.2) followed by autoradiography or fluorography (APPENDIX 3A). When immunoprecipitates are analyzed by immunoblotting (UNIT 10.8), an additional problem may be the detection of immunoprecipitating antibody bands in the blots (Table 10.16.2).

Anticipated Results

For antigens that are present at >10,000 copies per cell, the radiolabeling and immunoprecipitation protocols described in this book

can be expected to result in the detection of one or more bands corresponding to the specific antigen and associated proteins in the electrophoretograms. Specific bands should not be present in control immunoprecipitations done with irrelevant antibodies. If antigens are labeled with [³⁵S]methionine (UNIT 10.18), specific bands should be visible within 2 hr to 2 months of exposure. Due to the relatively low yield of the immunoprecipitation-recapture procedure (<10% of that of a single immunoprecipitation step), detection of specific bands is likely to require longer exposure times. This may turn out to be problematic due to the radioactive decay of ³⁵S (half-life = 88 days). In immunoprecipitation-recapture experiments in which the antibodies used for the first and second immunoprecipitation steps recognize different antigens (e.g., for the study of protein-protein interactions; Fig. 10.16.3), it is advisable to include in the second immunoprecipitation step a positive control containing either the antibody used for the first immunoprecipitation (if the antibody recognizes both the native and denatured forms of the antigen) or a different antibody with specificity for the same antigen (Fig. 10.16.4, lane 3).

Time Considerations

Preparation of cell extracts (Basic Protocol 1 and Alternate Protocols 1 to 4) takes 1 to 3 hr to complete, and isolation of the antigen on antibody-conjugated beads takes 2 to 3 hr. Binding antibodies to immunoadsorbent beads can be done prior to or simultaneously with preparation of the cell extracts and also takes 1 to 3 hr. Therefore, the whole immunoprecipitation procedure can be completed in 1 day. Similar time requirements apply for Alternate Protocols 5 and 6. Reaction between antibody and antigen can be extended overnight, but this may increase the background when using Ab-Sepharose. Antigen-antibody complexes should not be left overnight in the midst of washing, as significant dissociation may occur.

Immunoprecipitation-recapture experiments (Basic Protocol 2) require an additional 1 to 2 hr to denature and prepare the antigen for immunoprecipitation, and 2 to 3 hr to isolate the antigen. Completion of an entire immunoprecipitation-recapture experiment requires a very long workday. Alternatively, samples can be frozen after the first immunoprecipitation, and the elution and recapture can be carried out another day. Immunoprecipitates can be ana-

lyzed immediately (e.g., resolved by SDS-PAGE) or frozen and analyzed another day.

Literature Cited

- Anderson, D.J. and Blobel, G. 1983. Immunoprecipitation of proteins from cell-free translations. *Methods Enzymol.* 96:111-120.
- Cuatrecasas, P. 1970. Protein purification by affinity chromatography. *J. Biol. Chem.* 245:3059.
- Franzusoff, A., Rothblatt, J., and Schekman, R. 1991. Analysis of polypeptide transit through yeast secretory pathway. *Methods Enzymol.* 194:662-674.
- Gelb, W.G. 1973. Affinity chromatography: For separation of biological materials. *Am. Lab.* 81:61-67.
- Gersten, D.M. and Marchalonis, J.J. 1978. A rapid, novel method for the solid-phase derivatization of IgG antibodies for immune-affinity chromatography. *J. Immunol. Methods* 24:305-309.
- Harford, J. 1984. An artefact explains the apparent association of the transferrin receptor with a ras gene product. *Nature* 311:673-675.
- Harlow, E. and Lane, D. 1999. *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Irving, R.A., Hudson, P.J., and Goding, J.W. 1996. Construction, screening and expression of recombinant antibodies. In *Monoclonal Antibodies: Principles and Practice* (J.W. Goding, ed.). Academic Press, London.
- Kessler, S.W. 1975. Rapid isolation of antigens from cells with a staphylococcal protein A-antibody adsorbent: Parameters of the interaction of antibody-antigen complexes with protein A. *J. Immunol.* 115:1617.
- Köhler, G. and Milstein, C. 1975. Continuous cultures of fused cells secreting antibody of predetermined specificity. *Nature* 256:495-497.
- March, S.C., Parikh, I., and Cuatrecasas, P. 1974. A simplified method for cyanogen bromide activation of agarose for affinity chromatography. *Anal. Biochem.* 60:149-152.
- Nisonoff, A. 1984. *Introduction to Molecular Immunology*. Sinauer Associates, Sunderland, Mass.
- Rapley, R. 1995. The biotechnology and applications of antibody engineering. *Mol. Biotech.* 3:139-154.

Key References

- Harlow and Lane, 1988. See above.
- Helenius, A., McCaslin, D.R., Fries, E., and Tanford, C. 1979. Properties of detergents. *Methods Enzymol.* 56:734-749.
- Hjelmeland, J.M. and Chrambach, A. 1984. Solubilization of functional membrane proteins. *Methods Enzymol.* 104:305-318.

The above references describe various detergents and solubilization conditions.

Contributed by Juan S. Bonifacino and
Esteban C. Dell'Angelica
National Institute of Child Health and
Human Development
Bethesda, Maryland

Timothy A. Springer (precipitation with
antibody-Sepharose and with anti-Ig)
Center for Blood Research
Harvard Medical School
Boston, Massachusetts

Synthesizing Proteins In Vitro by Transcription and Translation of Cloned Genes

BASIC PROTOCOL

The availability of a cloned gene makes it possible to synthesize the encoded protein by in vitro transcription and translation. Such in vitro synthesized proteins are extremely useful for a wide variety of purposes including the analysis of DNA-protein interactions (UNIT 12.9). The basis of the method is to clone the protein-coding sequences into a vector containing a promoter for SP6 or T7 RNA polymerase (UNIT 1.5), to produce messenger RNA by transcribing the DNA template with this enzyme (UNIT 3.8), and to synthesize the desired protein (often as a ³⁵S-labeled species) by translation of this mRNA in vitro. A major advantage of this method is that any desired mutant protein can be generated simply by altering the DNA template.

Materials

Plasmid DNA containing SP6 or T7 promoter
 DNA containing cloned gene *or* cDNA encoding protein of interest
 Appropriate restriction endonucleases (UNIT 3.1)
 TE buffer (APPENDIX 2)
 5× ribonucleoside triphosphate mix
 10× SP6/T7 RNA polymerase buffer (UNIT 3.4)
 10 mM spermidine (for SP6 RNA polymerase only)
 Pancreatic ribonuclease inhibitor (e.g., RNasin from Promega Biotec)
 SP6 or T7 RNA polymerase (UNIT 3.8)
 Buffered phenol (UNIT 2.1)
 Isobutanol
 10 M ammonium acetate
 100% ethanol
 In vitro translation kit (wheat germ extract or reticulocyte lysate)
³⁵S-labeled methionine (1400 Ci/mmol)
 0.1 M NaOH
 10% trichloroacetic acid (TCA)
 EN³HANCE (Du Pont NEN)

Additional materials for subcloning DNA fragments (UNIT 3.16), preparing highly purified plasmid DNA (UNIT 1.7), digesting DNA with restriction endonucleases (UNIT 3.1), phenol extraction and ethanol precipitation (UNIT 2.1), agarose and nondenaturing polyacrylamide gel electrophoresis (UNITS 2.5, 2.7, & 12.2), quantitating radioactivity by acid precipitation on glass-fiber filters (UNIT 3.4), one-dimensional SDS-PAGE (UNIT 10.2), and autoradiography (APPENDIX 3).

Prepare the DNA template

1. Subclone protein-coding DNA sequences of interest into a plasmid vector that contains a promoter for SP6 or T7 RNA polymerase (e.g., pSP64) at a site downstream of the promoter.

The protein-coding sequences must be contiguous—i.e., uninterrupted by introns. It is also crucial that the protein-coding sequences be cloned in the correct orientation downstream of the bacteriophage promoter such that the correct initiation codon is the first AUG in the RNA to be synthesized. It is best if the AUG initiation codon is relatively close (25 to 100 bases) to the 5'-end of the RNA; this may require additional DNA manipulations of the protein-coding sequences such as nested deletions (UNIT 7.2) or oligonucleotide-directed mutagenesis (UNIT 8.1). See critical parameters.

2. Prepare plasmid DNA containing protein-coding sequences downstream of the bacteriophage promoter (step 1) by CsCl/ethidium bromide centrifugation or PEG precipitation (UNIT 1.7).

It is important that the DNA be of high quality; DNA produced by standard miniprep procedures (UNIT 1.6) is generally not adequate. It is important that the DNA preparations be devoid of ribonuclease activity.

3. Cleave 10 μg DNA with a restriction endonuclease that cuts just downstream of termination codon (ideally 50 to 200 bp) and does not cut within the protein-coding region. Remove a small aliquot, and check the efficiency of cleavage by agarose gel electrophoresis.

Although not essential, this step will ultimately result in the synthesis of higher molar amounts of the desired mRNA because it minimizes the length of the transcript. However, mRNAs containing less than 50 bases beyond the termination codon may be translated less efficiently.

4. Purify the DNA by phenol extraction and ethanol precipitation. Resuspend in 50 μl TE buffer.

This amount of DNA is enough for ten separate in vitro transcription and translation reactions.

Prepare the mRNA by in vitro transcription

5. Set up the following 25- μl reaction mixture at room temperature (to avoid precipitation of the DNA template by spermidine):

8 μl H₂O
5 μl DNA (total 1 μg)
5 μl 5 \times ribonucleoside triphosphate mix
2.5 μl 10 \times SP6/T7 RNA polymerase buffer
2.5 μl 10 mM spermidine (for SP6 RNA polymerase only)
1 μl RNasin (30 to 60 U)
1 μl SP6 or T7 RNA polymerase (5 to 20 U).

Incubate 60 min at 40°C. For reactions with T7 RNA polymerase, omit the spermidine and add an additional 2.5 μl water.

If desired, the transcription reaction can be monitored by electrophoresis in formaldehyde gels (UNIT 4.9) or by including a small amount of labeled ribonucleotides in the reaction mixture.

6. Add 25 μl buffered phenol, vortex, and extract immediately. Transfer the aqueous phase to a new microcentrifuge tube, extract twice with isobutanol, and add 6 μl of 10 M ammonium acetate and 70 μl ethanol. Ethanol precipitate with RNA and wash once with ethanol.
7. Resuspend the RNA in 24 μl TE buffer, add 6 μl of 10 M ammonium acetate and 70 μl ethanol, reprecipitate, and wash once with ethanol. Resuspend RNA in 10 μl TE buffer.

The RNA should be translated immediately or quick frozen and stored at -70°C.

Prepare the protein by in vitro translation

8. Add 1 to 10 μl RNA to the appropriate reagents in an in vitro translation kit and follow the directions of the manufacturer. Add 15 μCi of [³⁵S]methionine (1400 Ci/mmol) to radioactively label the protein. Reactions are typically performed in 30- μl volumes at room temperature for 30 to 60 min. After the reaction is complete, store at 0° to 4°C.

The amount of RNA to be added depends on the efficiency of the in vitro transcription reaction. In general, 1 to 2 μ l should be enough, but if necessary, up to 10 μ l can be used.

Translation reactions are generally carried out with commercially available kits that use either wheat germ extracts or reticulocyte lysates (see commentary). Although commercial kits are relatively expensive, they are highly recommended as active translation extracts can be difficult and tedious to prepare. For each new preparation of translation extract, it is useful to perform a control reaction that lacks added RNA.

For some applications, unlabeled methionine is used to synthesize proteins that are unlabeled or of lower specific activity. Proteins should be used as soon as possible, although they are generally stable at 0° to 4°C for a week.

9. Remove 1 μ l of the translation products and add to 50 μ l of 0.1 M NaOH; incubate 15 min at 37°C. Add 1 ml of 10% TCA and incubate an additional 15 min on ice.
10. Collect the precipitated protein on glass fiber filters and quantitate the incorporated [³⁵S]methionine by scintillation counting (UNIT 3.4).

By comparing the amount of incorporated [³⁵S]methionine in the sample to the control lacking added RNA, it is possible to estimate the amount of protein synthesized. Wheat germ extracts lack any endogenous methionine, hence making the calculation straightforward; reticulocyte lysates contain variable amounts of endogenous methionine.

11. Remove 3 μ l of the translation reaction and analyze the products by one-dimensional SDS–polyacrylamide gel electrophoresis, including lanes for molecular weight standards. Visualize the [³⁵S]proteins by fluorography with EN³HANCE and autoradiography.

Ideally, the synthesis should yield a single species of predicted molecular weight. The band should be visible in 1 to 4 hr.

REAGENTS AND SOLUTIONS

5× ribonucleoside triphosphate mix

5 mM each ATP, UTP, CTP

5 mM diguanosine triphosphate (G-5'ppp5'-G)TP

0.5 mM GTP

COMMENTARY

Background Information

The procedure for synthesizing proteins in vitro from cloned genes can be divided into several steps. First, protein-coding sequences are cloned into a vector containing a promoter for SP6 or T7 RNA polymerase (UNIT 1.5). The protein-coding sequence can be obtained from a cDNA clone, or for genes lacking introns, it can be generated from cloned genomic DNA. In some cases, it is necessary to modify the initial cDNA or genomic clone to optimize the lengths of the 5' or 3' untranslated sequences adjacent to the protein-coding region (see critical parameters).

Second, messenger RNA encoding the pro-

tein is generated by transcribing the DNA template with the appropriate bacteriophage RNA polymerase. This step is carried out essentially as described in UNIT 3.8, except that 90% of the GTP in the reaction mixture is replaced by diguanosine GTP. In this way, 90% of the RNA synthesized contains a 5'-capped structure that is normally found in eukaryotic mRNAs; this is important for efficient translation in vitro (Melton et al., 1984). Given the very high specificity of SP6 and T7 RNA polymerases for transcriptional initiation from their cognate promoters, the RNA synthesized is a single species that encodes the desired protein.

Third, the essentially pure mRNA is translated *in vitro* using wheat germ extracts or reticulocyte lysates. By using [³⁵S]methionine during the translation reaction, the protein is synthesized as a radiolabeled species. Moreover, because the translation reaction is programmed with a pure mRNA species, the protein generated should be radiochemically pure; i.e., analysis by SDS-PAGE should yield a single band of expected molecular weight assessed by autoradiography. Of course, the protein synthesized *in vitro* will not be biochemically pure, because of the vast excess of other proteins present in the translation extract. For most applications, *in vitro* synthesized proteins can be used directly without any further purification.

A significant advantage of this approach is that any desired mutant protein can be created simply by altering the DNA template, and then tested for the property of interest. For example, by creating a set of N- or C-terminal deletions of the protein, a DNA-binding domain can be localized (Hope and Struhl, 1986; *UNIT 12.9*). The availability of truncated but functional proteins can be useful for determining the subunit structure of a protein (Hope and Struhl, 1987). [³⁵S]proteins are also useful for structural studies involving proteases (Hope et al., 1988) and for examining protein-protein interactions. In addition, they can be used as substrates to assay for factors that modify the structure and/or activity of the protein.

Critical Parameters

In order to synthesize the desired protein, the uninterrupted protein-coding sequence must be cloned in the correct orientation downstream of the bacteriophage promoter. For efficient translation, it is essential that the correct initiation codon is the 5'-proximal AUG in the mRNA to be synthesized; initiation from internal AUG codons is generally very inefficient. It is best if the AUG initiation codon is relatively close (25 to 100 bases) to the 5'-end of the RNA; this may require additional DNA manipulations of the protein-coding sequences. Occasionally, synthesis of a desired protein may be inefficient because the AUG initiation codon resides in a poor sequence context or is inaccessible due to secondary structure of the mRNA. To circumvent either of these problems, the best approach is use the appropriate mutagenesis procedure(s) to replace the normal 5'-untranslated sequences (and possibly the first few translated codons) with the equivalent region from an efficiently translated protein.

For synthesizing mRNA, the most important consideration is the quality of the DNA preparation. In particular, RNA synthesis reactions can be inhibited by high concentrations of RNA, and the transcription products are easily destroyed by contaminating ribonucleases. DNA prepared by standard miniprep procedures (*UNIT 1.6*) is generally not suitable; preparation by CsCl gradient centrifugation or PEG precipitation (*UNIT 1.7*) is highly recommended. With high-quality DNA, the transcription of full-length mRNA is rarely a problem. However, if necessary, the amount and quality of the RNA synthesized can be examined by electrophoresis in formaldehyde gels (*UNIT 4.9*).

Although many mRNAs can be translated efficiently in wheat germ extracts or in reticulocyte lysates, synthesis of individual proteins may be more efficient in one of these systems. In general, it is thought that wheat germ extracts initiate translation somewhat better than reticulocyte lysates but are more prone to premature termination (or degradation of the mRNA template). Thus, wheat germ extracts may be preferred for shorter proteins, whereas reticulocyte lysates may be the best choice for longer proteins. However, many other factors can influence the translation efficiency for any particular protein. Wheat germ extracts contain no exogenous methionine, and hence generate proteins with somewhat higher specific activity than those produced in reticulocyte lysates, which contain variable amounts of exogenous methionine. For all experiments involving *in vitro* synthesized proteins, it is essential to determine the quality of the translation product by SDS-PAGE (*UNIT 10.2*).

Finally, a general consideration is that the protein synthesized *in vitro* may not be active either because its structure is different from the native protein or because a critical cofactor is missing. For such reasons, it is possible that proteins synthesized in wheat germ extracts or reticulocyte lysates might have different properties.

Anticipated Results

The *in vitro* transcription and translation reactions should yield a single band on an SDS-polyacrylamide gel whose molecular weight corresponds to that expected for the protein of interest. For some proteins, the mobility of the *in vitro* synthesized product is much slower than expected and the apparent molecular weight can be strongly influenced by the gel conditions (Hope and Struhl, 1985). Further confirmation that the correct protein is

being synthesized can be obtained by cleaving the template DNA at various positions within or just beyond the protein-coding sequences, and examining the proteins synthesized. Even in successful syntheses, minor bands are often observed. The most common reasons for bands corresponding to lower-molecular-weight products are premature termination of translation, aberrant initiation at internal AUG codons, and proteolysis. Bands corresponding to higher-molecular-weight species are less common and may be due to failure to terminate at the correct stop codon and post-translational modification. Protein modification may account for why certain in vitro synthesis reactions yield a set of distinct products of similar molecular weight.

A standard reaction should generate 1 to 10 ng of radiochemically pure protein at a specific activity between 10^8 and 10^9 cpm/ μ g. Because the materials necessary for in vitro protein synthesis are essentially unlimited, it should be possible to make at least 1 μ g of protein.

Time Considerations

The time necessary to generate a DNA template suitable for efficient synthesis of protein in vitro is highly variable; it can be as short as a few days. Assuming a high-quality preparation of such a DNA template, it is possible in a single day to cleave the DNA, transcribe the template into mRNA, synthesize the protein, and analyze the products by SDS-PAGE and autoradiography. More typically, these steps

are carried out over 2 days. The procedure can be stopped at essentially any point; however, mRNA should be stored at -70°C . Once synthesized, protein can be stored at 0° to 4°C without loss of stability for ~ 1 week. If a given protein is to be synthesized on multiple occasions, it is useful to prepare excess amounts of cleaved DNA template and/or mRNA.

Literature Cited

- Hope, I.A. and Struhl, K. 1985. GCN4 protein synthesized in vitro, binds HIS3 regulatory sequences: Implications for general control of amino acid biosynthetic genes in yeast. *Cell* 43:177-188.
- Hope, I.A. and Struhl, K. 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* 46:885-894.
- Hope, I.A. and Struhl, K. 1987. GCN4, a eukaryotic transcriptional activator protein, binds DNA as a dimer. *EMBO J.* 6:2781-2784.
- Hope, I.A., Mahadevan, S., and Struhl, K. 1988. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. *Nature* 333:635-640.
- Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K., and Green, M.R. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucl. Acids Res.* 12:7057-7070.

Contributed by Kevin Struhl
Harvard Medical School
Boston, Massachusetts

Metabolic Labeling with Amino Acids

Metabolic labeling techniques are used to study biosynthesis, processing, intracellular transport, secretion, degradation, and physical-chemical properties of proteins. In this unit, protocols are described for metabolically labeling mammalian cells with radiolabeled amino acids (Table 10.18.1). Cells labeled using these procedures are suitable for analysis by immunoprecipitation (UNIT 10.16), characterization of cellular proteins, analysis of protein trafficking, and one- and two-dimensional gel electrophoresis (UNITS 10.2-10.4). The first protocols describe pulse-labeling (10 to 30 min) of mammalian cells in suspension with [³⁵S]methionine (see Basic Protocol), and necessary modifications for adherent mammalian cells (see Alternate Protocol 1). Alternate protocols are also presented for pulse-chase analysis (see Alternate Protocol 2) and long-term labeling (“steady state,” 6 to 32 hr; see Alternate Protocol 3). This is followed by conditions for metabolic labeling of cells with radiolabeled amino acids other than [³⁵S]methionine (see Alternate Protocol 4). The degree of label incorporation can be determined by precipitation with trichloroacetic acid (TCA; see Support Protocol).

NOTE: All solutions and equipment coming into contact with living cells must be sterile, and aseptic technique should be used accordingly.

NOTE: All culture incubations should be performed in a humidified, 37°C, 5% CO₂ incubator unless otherwise specified. Some media (e.g., DMEM) may require altered levels of CO₂ to maintain pH 7.4.

Table 10.18.1 Radiolabeled Amino Acids Used in Metabolic Labeling of Proteins

Amino acid ^a	Frequency (%) ^b	Radioisotope	Specific activity (Ci/mmol)	Comments ^c
Leucine	10.4	³ H	5-190	E
Serine	8.1	³ H	5-40	T, I
Glutamic acid	7.3	³ H	15-80	T, I
Lysine	7.0	³ H	40-110	E
Alanine	7.0	³ H	10-85	T
Valine	6.2	³ H	10-65	E
Glycine	5.7	³ H	10-60	T, I
Threonine	5.6	³ H	5-25	E
Arginine	5.0	³ H	30-70	E
Aspartic acid	4.9	³ H	10-50	T, I
Proline	4.9	³ H	15-130	—
Glutamine	4.5	³ H	20-60	T, I
Phenylalanine	4.5	³ H	15-140	E
Tyrosine	3.6	³ H	15-60	—
Asparagine ^d	3.5	—	—	—
Cysteine	3.4	³⁵ S	>800	—
Isoleucine	2.9	³ H	30-140	E
Histidine	2.5	³ H	30-70	E
Methionine	1.8	³⁵ S	>800	E
Tryptophan	1.3	³ H	20-30	E

^aAll amino acids in this table are in the L configuration.

^bFrequency of amino acid residues in proteins, taken from Lathe (1985).

^cE, essential amino acids; T, amino acids that are modified by transamination (Coligan et al., 1983); I, amino acids that are converted by cells to other amino acids.

^dAsparagine is difficult to label (Coligan et al., 1983).

Contributed by Juan S. Bonifacino

Current Protocols in Molecular Biology (1998) 10.18.1-10.18.10

Copyright © 1998 by John Wiley & Sons, Inc.

Analysis of
Proteins

10.18.1

Supplement 44

SAFETY PRECAUTIONS FOR WORKING WITH ³⁵S-LABELED COMPOUNDS

When working with radioactive materials, take appropriate precautions to avoid contamination of the experimenter and the surroundings. Carry out the experiment and dispose of wastes in an appropriately designated area, following guidelines provided by the local radiation safety officer (also see *APPENDIX 1F*).

Solutions containing ³⁵S-labeled compounds have been found to release volatile radioactive substances (Meisenhelder and Hunter, 1988). In addition to the usual safety practices followed when handling radioactive materials (see *APPENDIX 1F*), some extra precautions should be taken when using ³⁵S-labeled amino acids:

1. Vials containing ³⁵S-labeled compounds should always be handled in a designated fume hood equipped with an activated charcoal filter. This includes thawing the solution, opening the vial, and adding the radiolabeled amino acid to the labeling medium. Before opening, vials should be vented with a needle attached to a syringe packed with activated charcoal. Avoid using tissue culture hoods for this purpose as they are likely to become contaminated.
2. Minimize exposure of ³⁵S-containing solutions to the air.
3. Use baths, incubators, and centrifuges designated for work with radioactive materials. Place a tray containing a layer of activated charcoal in the CO₂ incubator to reduce the amount of ³⁵S-labeled compounds released to the air during cell labeling. Alternatively, filters impregnated with activated carbon (β-Safe, Schleicher & Schuell) can be attached to the covers of tissue culture dishes.
4. Monitor areas used during labeling by conducting wipe tests.
5. Dispose of solid and liquid ³⁵S waste quickly and with appropriate precautions.

BASIC PROTOCOL

PULSE-LABELING OF CELLS IN SUSPENSION WITH [³⁵S]METHIONINE

Pulse-labeling of proteins is performed by incubating cells for short periods (≤30 min) in culture medium containing a radiolabeled amino acid. The conditions described below are optimized for labeling times of 10 to 30 min. The same protocol is used for labeling cells for up to 6 hr, although a smaller number of cells and/or a larger amount of labeling medium should be used for optimal results. Very short pulses (<10 min) may require special conditions that are discussed elsewhere (see Critical Parameters). For labeling times >6 hr, see Alternate Protocol 3. [³⁵S]methionine is the radiolabeled amino acid of choice for metabolic labeling of proteins because of its high specific activity (>800 Ci/mmol) and ease of detection. A potential disadvantage of [³⁵S]methionine is its low abundance in proteins (~1.8% of the average amino acid composition; see Table 10.18.1). For proteins that contain little (e.g., only one) or no methionine, other radiolabeled amino acids should be used (see Alternate Protocol 4).

Materials

[³⁵S]L-Methionine (>800 Ci/mmol) or [³⁵S]-labeled protein hydrolyzate (>1000 Ci/mmol)

Pulse-labeling medium (see recipe), warmed to 37°C

Cell suspension (e.g., Jurkat, RBL, K562, BW5147, T and B cell hybridomas), grown in a humidified, 37°C, 5% CO₂ incubator or prepared from tissues (e.g., lymphocytes)

PBS (*APPENDIX 2*), ice cold

Vacuum aspirator with trap for liquid radioactive waste

Additional reagents and equipment for TCA precipitation (optional; see Support Protocol)

Metabolic Labeling with Amino Acids

10.18.2

1. Thaw [³⁵S]methionine at room temperature and prepare a 0.1 to 0.2 mCi/ml working solution in prewarmed (37°C) pulse-labeling medium.

CAUTION: Volatile ³⁵S-containing compounds can be released during the labeling procedure (refer to Safety Precautions and APPENDIX 1F). Keep [³⁵S]methionine-containing medium in a tightly capped tube in a 37°C water bath until use. Do not let it sit >60 min at 37°C.

2. Harvest $0.5\text{--}2 \times 10^7$ cells in suspension by centrifuging 5 min at $300 \times g$, room temperature.
3. Wash cells with ~10 ml prewarmed pulse-labeling medium. Centrifuge 5 min at $300 \times g$, room temperature, and aspirate supernatant. Resuspend cells by gently tapping the bottom of the tubes and repeat wash.
4. Resuspend cells at 5×10^6 cells/ml in prewarmed pulse-labeling medium and incubate 15 min in a 37°C water bath to deplete intracellular pools of methionine. Invert tubes periodically to resuspend cells.
5. Centrifuge cells 5 min at $300 \times g$, room temperature, and aspirate supernatant.
6. Resuspend cells in a fresh 15-ml centrifuge tube, using 2 ml [³⁵S]methionine working solution (step 1). Cap tubes tightly. Incubate cells 10 to 30 min (see Critical Parameters) in a 37°C water bath, resuspending frequently by gentle inversion of the tubes.

Alternatively, use a rotator placed in a 37°C incubator.

7. Centrifuge cells 5 min at $300 \times g$, 4°C, and aspirate supernatant. Resuspend with gentle swirling in 10 ml ice-cold PBS and repeat centrifugation.

CAUTION: The medium and wash are radioactive—follow applicable safety regulations for disposal.

Because of the short times and high concentrations of radiolabeled amino acids employed, pulse-labeling may not result in complete depletion of label from the medium. If this is the case, the labeling mixture can be reused. Collect the labeling medium containing [³⁵S]methionine or ³⁵S-labeled protein hydrolyzate and filter carefully through a 0.45- μ m filter unit. Estimate the percentage of unincorporated radioactivity by scintillation counting aliquots of the labeling mixture before and after labeling the cells. Store the labeling mixture up to 2 months frozen at -20°C .

8. *Optional:* Determine amount of label incorporation by trichloroacetic acid (TCA) precipitation (see Support Protocol 1).

Labeled cells are now ready for the desired processing and analysis. If cell pellets cannot be processed immediately, they can be kept on ice for a few hours or frozen at -80°C for several days. Thaw frozen cell pellets on ice before analysis. In most cases, freezing cells will not affect the biochemical properties of the proteins, but freezing and thawing cell extracts after solubilization with detergents can cause dissociation of multisubunit complexes or proteolysis of labeled proteins.

PULSE-LABELING OF ADHERENT CELLS WITH [³⁵S]METHIONINE

Labeling adherent cells is essentially the same as described for cells in suspension except that cells are labeled while attached to a dish. As with suspended cells, adherent cells are pulse-labeled 10 to 30 min (see Basic Protocol introduction). Alternatively, some adherent cells can be detached from plates by incubation for 10 min at 37°C with 10 mM EDTA in PBS (*APPENDIX 2*) and labeled as a suspension (see Basic Protocol). This is particularly advantageous when labeling cells for pulse-chase experiments, because it simplifies handling of multiple samples and allows more uniform labeling of the cells.

Additional Materials (also see Basic Protocol)

Adherent cells (e.g., HeLa, NRK, M1, COS-1, CV-1, or fibroblasts or endothelial cells in primary culture)
100-mm tissue culture dishes

1. Grow adherent cells to 80% to 90% confluency in 100-mm tissue culture dishes.

Depending on the cell type, a confluent 100-mm dish will contain $0.5\text{--}2 \times 10^7$ cells.

2. Prepare [³⁵S]methionine working solution as described (see Basic Protocol, step 1).
3. Aspirate culture medium from the dishes and wash twice by gently swirling with 10 ml prewarmed (37°C) pulse-labeling medium, aspirating medium after each wash.
4. Add 5 ml prewarmed pulse-labeling medium and incubate 15 min in a humidified, 37°C, 5% CO₂ incubator to deplete intracellular pools of methionine.
5. Remove medium from cells, add 2 ml [³⁵S]methionine working solution (from step 2), and incubate 10 to 30 min in a CO₂ incubator.

If necessary, as little as 1 ml of labeling medium per plate can be used. It is critical, however, that the plates sit on a perfectly horizontal surface during incubation to avoid drying of the cell monolayer.

6. Remove medium from cells. Wash once with 10 ml ice-cold PBS and remove PBS.

CAUTION: *The medium and wash are radioactive—follow applicable safety regulations for disposal.*

If significant detachment of cells occurs during labeling, it may be necessary to scrape the cells carefully in the [³⁵S]methionine-containing medium and transfer the suspension to a 15-ml centrifuge tube before centrifuging and washing with PBS.

7. Add 10 ml ice-cold PBS and scrape cells carefully with either a disposable plastic scraper or a rubber policeman.
8. Transfer the suspension to a 15-ml centrifuge tube, centrifuge 5 min at $300 \times g$, 4°C, and discard supernatant.
9. *Optional:* Determine amount of label incorporation by trichloroacetic acid (TCA) precipitation (see Support Protocol 1).

Labeled cells are now ready for the desired processing and analysis (see Basic Protocol, step 8 annotation, if cells are not to be used immediately).

PULSE-CHASE LABELING OF CELLS WITH [³⁵S]METHIONINE

Pulse-chase protocols are used to analyze time-dependent processes, such as posttranslational modification, transport, secretion, or degradation of newly synthesized proteins. Cells in suspension or adherent cells are pulse-labeled with [³⁵S]methionine (see Basic Protocol and Alternate Protocol 1), after which they are incubated (chased) in complete medium containing excess unlabeled methionine, as described below.

Additional Materials (also see *Basic Protocol* and *Alternate Protocol 1*)

Chase medium (see recipe), 37°C

1. Prepare $0.5\text{--}2 \times 10^7$ cells per sample per time point, and pulse-label 10 to 30 min with 2 ml of 0.1 to 0.2 mCi/ml [³⁵S]methionine per $0.5\text{--}2 \times 10^7$ cells (see Basic Protocol, steps 1 to 6, or see Alternate Protocol 1, steps 1 to 5).
2. Remove the [³⁵S]methionine working solution, wash once with 10 ml 37°C chase medium, and add 10 ml 37°C chase medium.

Rapid termination of the labeling reaction can be achieved by adding two times the volume of chase medium containing excess unlabeled methionine (15 mg/liter) directly to the labeling mixture.

Final concentration of cells in suspension should be 2×10^6 cells/ml for chases ≤ 2 hr, or 0.5×10^6 cells/ml for chases > 2 hr. For adherent cells, add 10 ml/100-mm dish.

3. Incubate for the desired time at 37°C. Incubate cell suspensions with rotation in tightly capped tubes. Incubate adherent cells in a CO₂ incubator.
- 4a. *For cells in suspension:* Collect cells by centrifuging 5 min at $300 \times g$, 4°C.

The supernatant can be discarded or can be collected for analysis of proteins that are secreted or shed into the medium.

- 4b. *For adherent cells:* Scrape off adherent cells and transfer to 15-ml centrifuge tubes. Centrifuge 5 min at $300 \times g$, 4°C, and either save or discard the supernatant as in step 4a.

5. *Optional:* Determine amount of label incorporation by trichloroacetic acid (TCA) precipitation (see Support Protocol 1).

Labeled cells are now ready for the desired processing and analysis (see Basic Protocol, step 8 annotation, if cells are not to be used immediately).

LONG-TERM LABELING OF CELLS WITH [³⁵S]METHIONINE

Long-term labeling refers to continuous metabolic labeling of cells for periods of 6 to 32 hr. Long-term labeling is particularly advantageous when studying proteins that are synthesized at low rates. It is also used to accumulate mature labeled proteins, rather than biosynthetic precursors, for characterization of their properties. Steady-state labeling is a form of long-term labeling in which cells are incubated with the radiolabeled amino acid until the rates of synthesis and degradation of the radiolabeled proteins are equal. Steady-state labeling allows calculation of the stoichiometry of subunits within a protein complex, provided that the primary structure of the subunits is known. In these procedures, unlabeled methionine is added to the medium to maintain cell viability and to sustain incorporation of label for the duration of the experiment. The amount of unlabeled methionine added depends on factors such as the length of the labeling period and the cell density. Media containing between 5% and 20% the normal amount of unlabeled methionine are generally used. The conditions described below are suitable for overnight (~16 hr) labeling of cells in suspension or adherent cells.

**ALTERNATE
PROTOCOL 2**

**ALTERNATE
PROTOCOL 3**

**Analysis of
Proteins**

10.18.5

Additional Materials (also see *Basic Protocol and Alternate Protocol 1*)

Long-term labeling medium (see recipe), warmed to 37°C
75-cm² tissue culture flask

1. Prepare a 0.02 to 0.1 mCi/ml [³⁵S]methionine working solution in prewarmed (37°C) long-term labeling medium (see *Basic Protocol*, step 1).

For cell suspensions

- 2a. Prepare and wash cells once with prewarmed long-term labeling medium (see *Basic Protocol*, steps 2 and 3).
- 3a. Resuspend 0.5–2 × 10⁷ cells in 25 ml of 0.02 to 0.1 mCi/ml [³⁵S]methionine working solution and transfer to a 75-cm² tissue culture flask.

For adherent cells

- 2b. Grow 0.5–2 × 10⁷ adherent cells in a 75-cm² tissue culture flask (80% to 90% confluency) in a CO₂ incubator. Wash once with prewarmed long-term labeling medium (see *Alternate Protocol 1*, step 3).
- 3b. Add 25 ml of 0.02 to 0.1 mCi/ml [³⁵S]methionine working solution to each 75-cm² flask.
4. Tighten caps to prevent release of volatile ³⁵S-labeled compounds. Incubate 16 hr in a CO₂ incubator.
5. Wash cells and determine incorporation (see *Basic Protocol*, steps 7 and 8, or *Alternate Protocol 1*, steps 6 to 9).

*Labeled cells are now ready for the desired processing and analysis (see *Basic Protocol*, step 8 annotation, if cells are not to be used immediately).*

**ALTERNATE
PROTOCOL 4**

METABOLIC LABELING WITH OTHER RADIOLABELED AMINO ACIDS

In some instances, it may be necessary to label proteins with radiolabeled amino acids other than [³⁵S]methionine—e.g., when the proteins have a low methionine content or have no methionine residues at all. In these cases, the best choices are [³⁵S]cysteine or [³H]leucine.

Cysteine residues are more abundant in proteins than are methionine residues (3.4% versus 1.8%; Table 10.18.1), although cysteine is less stable. Formulations of [³⁵S]cysteine that have high specific activity (>800 Ci/mmol) can be obtained from several companies. Leucine has the advantage of being the most frequent amino acid in proteins (10.4%; Table 10.18.1) and specific activities of [³H]leucine are the highest among ³H-labeled amino acids (up to 190 Ci/mmol). If proteins are known to be rich in a particular amino acid or if special methods are to be performed (e.g., radiochemical sequencing or multiple labeling), ³H-labeled amino acids other than [³H]leucine can be used.

Label cells with these amino acids as described for [³⁵S]methionine in the previous protocols. Substitute [³⁵S]cysteine or ³H-labeled amino acids in the labeling medium. Use labeling media lacking cysteine or any other respective amino acid.

CAUTION: Solutions containing [³⁵S]cysteine release volatile ³⁵S-labeled compounds (refer to Safety Precautions).

TCA PRECIPITATION TO DETERMINE LABEL INCORPORATION

In metabolic labeling experiments, it is often useful to monitor the incorporation of radioactivity into total cellular proteins. This can easily be achieved by precipitation with trichloroacetic acid (TCA) using BSA as a carrier protein. This protocol can be used at the end of the labeling procedures in each protocol.

Materials

Labeled cell suspension (see Basic Protocol or Alternate Protocols 1 to 4)
BSA/ NaN_3 : 1 mg/ml BSA containing 0.02% (w/v) sodium azide (NaN_3)
10% (w/v) TCA solution (see recipe), ice cold
Ethanol

Filtration apparatus attached to a vacuum line
2.5-cm glass microfiber filter disks (Whatman GF/C)

CAUTION: TCA is extremely caustic. Protect eyes and avoid contact with skin when preparing and handling TCA solutions.

1. Add 10 to 20 μl of a labeled cell suspension to 0.1 ml BSA/ NaN_3 . Place on ice.
2. Add 1 ml ice-cold 10% (w/v) TCA solution. Vortex vigorously and incubate 30 min on ice.
3. Filter the suspension onto 2.5-cm glass microfiber filter disks in a filtration apparatus under vacuum.
4. Wash the disks twice with 5 ml ice-cold 10% (w/v) TCA solution and twice with ethanol. Air dry 30 min.

CAUTION: *The wash fluids should be handled as mixed chemical/radioactive waste—follow applicable safety regulations for disposal (see APPENDIX 1F).*

5. Spot the same volume of the radiolabeled cell suspension used in step 1 (10 to 20 μl) on a glass microfiber disk. Air dry.
This disk will be used to measure the total amount of radiolabeled amino acid in the cell suspension.
6. Transfer disks from steps 4 and 5 to 20-ml scintillation vials, add 5 ml scintillation fluid, and measure the radioactivity in a scintillation counter.
7. Calculate the ratio of TCA-precipitable label to total radioactivity (i.e., the ratio of sample radioactivity in step 4 to the total radioactivity in step 5).

REAGENTS AND SOLUTIONS

Use deionized or distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

NOTE: Use sterile tissue culture technique to prepare these reagents.

Chase medium

Pulse-labeling medium (see recipe) containing 15 mg/liter unlabeled methionine or equivalent excess amount of other amino acid used for radiolabeling. Store up to 2 weeks at 4°C.

Long-term labeling medium (90% methionine-free medium)

Mix 9 vol pulse-labeling medium (see recipe), lacking methionine or other appropriate amino acid, with 1 vol chase medium (see recipe) containing the same amino acid. Store up to 2 weeks at 4°C.

Pulse-labeling medium

Use supplemented Dulbecco's modified Eagle medium (DMEM; see APPENDIX 3F but omit nonessential amino acids) or RPMI, each lacking methionine or other specific amino acid, but containing 10% (v/v) FBS (dialyzed overnight against saline solution to remove unlabeled amino acids). Add 25 mM HEPES (with pH adjusted to 7.4 with NaOH). Store up to 2 weeks at 4°C.

The amino acid used to label cells is omitted from this medium (e.g., if [³⁵S]methionine is employed, methionine-free DMEM or RPMI 1640 must be used). Specific amino acid-free media can be obtained from several tissue culture media suppliers, or can be prepared (see APPENDIX 3F) from amino acid-free medium by adding individual amino acid components, omitting the one that will be used to label. A kit for the preparation of such media is available (Select-Amine from Life Technologies).

Trichloroacetic acid (TCA) solution, 10% (w/v)

Prepare a 100% (w/v) TCA stock solution by dissolving the entire contents of a newly opened TCA bottle in water (e.g., dissolve the contents of a 500-g bottle of TCA in sufficient water to yield a final volume of 500 ml). Store up to 1 year at 4°C. Prepare 10% (w/v) TCA by dilution and store up to 3 months at 4°C.

CAUTION: TCA is extremely caustic. Protect eyes and avoid contact with skin when preparing and handling TCA solutions.

COMMENTARY

Background Information

Metabolic labeling of cellular proteins is achieved by placing cells in a nutritional medium containing all components necessary for growth of cells in culture, except for one amino acid that is substituted by its radiolabeled form. The radiolabeled amino acids are transported across the plasma membrane by carrier-mediated systems and, once in the cytosol, are loaded onto tRNA molecules before being incorporated into newly synthesized proteins.

Because metabolic labeling techniques use the metabolic machinery of the cell to incorporate radiolabeled amino acids, there are limitations on the type of radiolabeled amino acids that can be employed. The list of potential precursors is restricted to L-amino acids normally found in proteins, in which one or more atoms are substituted by a radioisotope (Table 10.18.1). Sulfuric amino acids, such as methionine and cysteine, are conveniently labeled with ³⁵S. Most other amino acids can be obtained labeled with ³H.

Critical Parameters

The protocols in this unit should be considered as models from which more specific methods can be designed. The conditions described here have been optimized for radiolabeling proteins that are expressed at low to moderate levels (10⁴ to 10⁵ copies per cell) and assume that the labeled cells will be used for immunoprecipitation (UNIT 10.16). This level of expres-

sion is characteristic of most endogenous membrane proteins, luminal organellar proteins, signal transduction proteins, and transcription factors. More abundant proteins, such as cytoskeletal proteins or proteins expressed by infection or transfection, can be labeled with less radiolabeled amino acid and/or fewer cells. The choice of a particular labeling protocol and its modification will be aided by careful consideration of a number of parameters that influence the incorporation of radiolabeled amino acids into proteins, as discussed below.

Selection of amino acid label

Purified [³⁵S]methionine of high specific activity (>800 Ci/mmol) can be obtained from several suppliers of radiolabeled amino acids. Protein hydrolyzates of *Escherichia coli* grown in the presence of [³⁵SO₄]²⁻ (e.g., Tran³⁵S-label from ICN Biomedicals or Expre³⁵S³⁵S from NEN Life Sciences) can be used as substitutes for [³⁵S]methionine in metabolic labeling techniques. These preparations contain ³⁵S distributed among several different compounds. In a typical batch, ~70% of the radioactivity will be present as [³⁵S]methionine and ~15% as [³⁵S]cysteine; the remainder are other ³⁵S-labeled compounds. Labeling with these radioactive protein hydrolyzates in methionine-free medium will result in incorporation of label only in methionine residues; use of methionine- and cysteine-free medium will result in labeling of both methionine and cysteine residues.

Purified [³⁵S]methionine should be used whenever certainty of labeling with only [³⁵S]methionine is required. Another reason for using purified [³⁵S]methionine is that it tends to be more stable and emit less volatile decomposition products than ³⁵S-labeled protein hydrolyzates. In most cases, however, the relatively inexpensive radioactive protein hydrolyzates (~1/3 the cost of [³⁵S]methionine) can be used instead of purified [³⁵S]methionine. [³⁵S]methionine preparations should be stored frozen at -80°C. Under these conditions, they are stable for at least 1 month; the half-life of ³⁵S is 88 days.

For proteins with one or no methionine residues but several cysteine residues, labeling with [³⁵S]cysteine (available with specific activities >800 Ci/mmol) is a good option. ³⁵S-labeled protein hydrolyzates are not good sources of [³⁵S]cysteine for radiolabeling because only ~15% of the ³⁵S-labeled compounds are [³⁵S]cysteine.

Purified [³H]leucine (available with specific activities of up to 190 Ci/mmol) is a good alternative to ³⁵S-labeled amino acids. The half-life of ³H is ~12 years. The specific activities of other ³H-labeled amino acids range between 5 and 140 Ci/mmol (Table 10.18.1). Several problems can arise when using certain ³H-labeled amino acids due to their participation in metabolic pathways. The following problems must be considered when using tritiated amino acids for metabolic labeling of proteins.

Nonessential versus essential amino acids. Cells are able to synthesize nonessential amino acids from other compounds. If the radiolabeled amino acids used are nonessential, their specific activity will be reduced by dilution with the endogenously synthesized amino acids. Therefore, essential amino acids (E in Table 10.18.1) should be favored for metabolic labeling.

Transamination. The α -amino groups of many amino acids can be removed in reactions catalyzed by transaminases (T in Table 10.18.1). Deamination of the radiolabeled amino acids can be prevented by addition of the transaminase inhibitor (aminoxy)acetic acid during starvation and labeling of the cells (Coligan et al., 1983).

Interconversion. Certain radiolabeled amino acids can be converted by cells into other amino acids (I in Table 10.18.1). This problem is of particular importance in methods used to determine the amino acid composition or sequence of radiolabeled proteins.

Labeling time

The experimental purpose, the turnover rate of the protein of interest, and the viability of the cells should all be considered when determining the length of time for labeling. If the purpose of the experiment is to identify or characterize a protein precursor, pulse-labeling (10 to 30 min) must be used (see Basic Protocol and Alternate Protocol 1). If the protein of interest has a low turnover rate or if it is necessary to accumulate a labeled mature product, a protocol for long-term labeling (6 to 32 hr) is more appropriate (see Alternate Protocol 3). If the biosynthesis, posttranslational modification, intracellular transport, or fate of newly synthesized proteins is being analyzed, a pulse-chase protocol should be used (see Alternate Protocol 2).

The turnover rate of the protein of interest is an important parameter to be considered when determining the labeling time. Proteins with a high turnover rate are optimally labeled for short times. Longer times will result in increased labeling of other cellular proteins, causing a decrease in the relative abundance of the labeled protein of interest in the cell lysate. This could result in increased detection of non-specific proteins in immunoprecipitates (see UNIT 10.16). Conversely, proteins that turn over slowly should be labeled for longer times.

Incorporation of radiolabeled amino acids into proteins is directly proportional to their length of labeling time for a certain period, after which it tends to plateau. When all the limiting amino acid is consumed, protein synthesis ceases. The length of the initial phase of linear incorporation will vary with the concentration of the labeling amino acid and the density and metabolic activity of the cells.

Concentration and specific activity of the radiolabeled amino acid

Because radiolabeled amino acids are used in limiting amounts, their incorporation into proteins is directly proportional to their concentration in the labeling medium. If very short pulses are required, the concentration of labeled amino acid can be increased to compensate for the shorter labeling time. The incorporation of radiolabeled amino acids is also directly proportional to their specific activity. Addition of unlabeled amino acids, as is required in long-term labeling protocols, will result in reduced incorporation over short periods.

Cell density

At low cell densities, the amount of labeled protein synthesized is directly proportional to the cell concentration. At high densities, however, incorporation will increase nonlinearly. Very concentrated cell samples ($>5 \times 10^7$ cells/ml) can only be labeled for short periods (≤ 5 min) as rapid loss of metabolic activity and cell viability will occur due to acidification of the medium and accumulation of toxic metabolites.

Conditions for very short pulses

Very short pulses (2 to 10 min) are required to study posttranslational modifications that occur shortly after synthesis. This includes folding, disulfide bond formation, and early carbohydrate modifications of newly synthesized proteins. Proteins that are rapidly degraded after synthesis are also best studied with very short pulses. The protocol for very short pulses is similar to the Basic Protocol for pulse-labeling, with the following modifications: (1) the pulse-labeling medium contains higher concentration of radiolabeled amino acid (up to 1 mCi/ml), (2) labeling is stopped by addition of ice-cold PBS containing 20 mM freshly prepared *N*-ethylmaleimide (NEM) to prevent oxidation of free sulfhydryls, and (3) if a chase is necessary, 1 mM cycloheximide is added to the chase medium to quickly stop the elongation of nascent polypeptide chains. For additional details on very short pulses, see Braakman et al. (1991).

Temperature and pH

Unless otherwise required for special conditions (e.g., heat shock, temperature-sensitive mutants), metabolic labeling should be performed at 37°C. For this reason, it is important that the labeling medium be warmed to 37°C before adding it to the cells. A dramatic reduction in incorporation occurs at room temperature. It is also important that the pH of the labeling medium be ~ 7.4 .

Anticipated Results

A typical incorporation of labeled amino acid precursor after a 30-min pulse under the conditions described in this unit (see Basic Protocol and Alternate Protocols 1 and 2) is 5% to 20%. In the case of long-term labeling, the incorporation typically reaches 30% to 60%. If the labeled cells are used for immunoprecipitation (UNIT 10.16), followed by SDS-PAGE (UNIT 10.2) and autoradiography (APPENDIX 3A), specific bands should be visible within two hours to two months of exposure.

Time Considerations

It takes 1 to 2 hr to prepare cells and materials for labeling. The actual labeling time depends upon the protocol chosen—pulse-labeling takes 10 to 30 min and long-term labeling takes 6 to 32 hr. Washing and processing cells may take an additional 1.5 hr.

Literature Cited

- Braakman, I., Hoover-Litty, H., Wagner, K.R., and Helenius, A. 1991. Folding of influenza hemagglutinin in the endoplasmic reticulum. *J. Cell Biol.* 114:401-411.
- Coligan, J.E., Gates, F.T. III, Kimball, E.S., and Maloy, W.L. 1983. Radiochemical sequence analysis of metabolically labeled proteins. *Methods Enzymol.* 91:413-434.
- Lathe, R. 1985. Synthetic oligonucleotide probes deduced from amino acid sequence data. Theoretical and practical considerations. *J. Mol. Biol.* 183:1-12.
- Meisenhelder, J. and Hunter, T. 1988. Radioactive protein labelling techniques. *Nature* 335:120.

Key Reference

Coligan et al., 1983. See above.

Contains a detailed description of conditions used to metabolically label proteins with different radiolabeled amino acids.

Contributed by Juan S. Bonifacino
National Institute of Child Health
and Human Development
Bethesda, Maryland

Isolation of Proteins for Microsequence Analysis

UNIT 10.19

The first basic protocol can be used to determine the amino-terminal sequence of polypeptides or proteins. It is particularly appropriate for large fragments of insoluble or hydrophobic proteins or proteins that cannot be purified to >90% molar purity without electrophoresis. Although the efficacy of this technique varies with the protein, it is possible to obtain useful sequence information starting with ≤ 50 pmol of the protein of interest.

If the protein is blocked at the amino terminus, chemical cleavage or partial enzymatic digestion must be performed prior to electrophoresis. Upon isolation, the internal amino acid sequence is analyzed as described in the second basic protocol. This method requires ~ 200 pmol of protein for analysis. It is understood that both of the analyses will be done in association with an expert operator of an automated protein sequencer. In addition, the second basic protocol requires expertise with reversed-phase HPLC to separate peptides.

DETERMINATION OF AMINO ACID SEQUENCE BY SDS-PAGE AND TRANSFER TO PVDF MEMBRANES

**BASIC
PROTOCOL**

A minigel is prepared and preelectrophoresed. A sample containing the desired protein is loaded onto the minigel and fractionated by SDS-PAGE at neutral pH. The proteins are then electrophoretically transferred to a sheet of polyvinylidene difluoride (PVDF) and stained with Coomassie blue. The separated bands are excised, then analyzed in an automated protein sequencer.

Materials

- Separating and stacking gel solutions (Table 10.19.1)
- 4 \times gel buffer
- Glutathione, reduced powder (Sigma #G4251)
- 10 \times lower reservoir buffer
- 10 \times upper reservoir buffer
- Mercaptoacetic acid, sodium salt
- Protein sample in sample buffer (see support protocol)
- Methanol
- Transfer buffer
- 0.1% Coomassie blue in 50% methanol (v/v)
- 10% acetic acid in 50% methanol (v/v)
- Vertical minigel unit (e.g., Bio-Rad Mini-Protean II; Hoefer Mighty Small SE 250/280 is not recommended for this procedure)
- Power supply (constant voltage and constant current)
- Microvolume syringe or gel-loading pipet tip
- Powder-free plastic gloves
- PVDF membranes (e.g., Immobilon-P or -PSQ, Millipore; ProBlott, Applied Biosystems)
- Small-format transfer apparatus (Midget MultiBlot, Hoefer or Pharmacia LKB; Mini Trans-Blot, Bio-Rad)
- Automated protein sequencer (Applied Biosystems)
- Additional reagents and equipment for minigel preparation (UNIT 10.2)

**Analysis of
Proteins**

10.19.1

Contributed by Malcolm Moos, Jr.

Current Protocols in Molecular Biology (2000) 10.19.1-10.19.12

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 21

Table 10.19.1 Recipes for Polyacrylamide Separating and Stacking Gels

SEPARATING GEL											
Stock solutions	Final acrylamide concentration in the separating gels (%)										
	5	6	7	8	9	10	11	12	13	14	15
30% acrylamide monomer ^a	3.33	4.00	4.67	5.33	6.00	6.67	7.33	8.00	8.67	9.33	10.00
H ₂ O	11.49	10.83	10.16	9.50	8.84	8.17	7.51	6.84	6.18	5.51	4.85
TEMED ^b	0.040	0.033	0.029	0.025	0.025	0.020	0.018	0.017	0.015	0.014	0.013

Mix the above ingredients (listed in milliliters of stock solution) with 5 ml of 4× gel buffer and 0.14 ml of 5% potassium persulfate or 70 μl of 10% ammonium persulfate.

STACKING GEL

0.666 ml 30% acrylamide monomer^a
 3.033 ml H₂O
 0.025 ml 10% ammonium persulfate *or* 0.05 ml 5% potassium persulfate
 0.025 ml TEMED^b

^aGas-stabilized monomer solution (containing 37.5:1 acrylamide/*N,N*-methylene-bisacrylamide) from which acrylic acid and carbonyl-containing compounds have been removed (Protogel, National Diagnostics; or PAGE1 protein gel mix, Boehringer Mannheim)

^bTEMED may have to be altered to facilitate proper polymerization. Values given are reasonable approximations.

Pour and preelectrophorese the minigels

1. Pour denaturing minigels as in UNIT 10.2, substituting the separating and stacking gel solutions listed in Table 10.19.1.

Deaerate the gel polymerization mixtures prior to adding persulfate and TEMED and be careful not to draw in air through the polymerization mixtures when transferring or pipetting the solutions.

After removal of combs, the wells should be rectangular and firm. If they are not, prepare a fresh gel; poorly polymerized stacking gels are the most common cause of low sequencing yields.

2. Assemble vertical minigel unit.
3. Dilute 80 ml of 4× gel buffer to 320 ml (to 1× gel buffer). Pour 200 ml of 1× gel buffer into lower buffer reservoir. Add reduced glutathione (powder) to remaining 1× gel buffer to 1.0 mM final. Pour this into upper buffer reservoir. Attach gel to a constant voltage/constant current power supply (see UNIT 10.2 introduction for a discussion of electricity and electrophoresis).

Glutathione acts as a scavenger to eliminate the by-products of acrylamide polymerization.

4. Preelectrophorese by applying 10 mA per minigel for 45 min.

Voltage applied for >45 min may impair resolution.

5. Turn off power supply. Allow gel to stand overnight.

Gels may be stored this way for several days.

Load and electrophorese the sample

6. Pour off gel buffer and blot wells with tissue or filter paper.
7. Dilute 10× lower and upper reservoir buffers to 1×. Pour 200 ml of 1× lower reservoir buffer into lower reservoir. Add 0.1 g mercaptoacetic acid (sodium salt) to 150 ml of 1× upper reservoir buffer and pour into upper reservoir.

Mercaptoacetic acid is used as a substitute for glutathione at this step because it is less likely to interfere with sequence analysis.

8. Load protein sample dissolved in sample buffer with a microvolume syringe or gel-loading pipet tip.

When using 5-well combs and 0.75-mm gels, results are best with sample volumes <30 μ l/well.

It is advisable to run a known protein such as β -lactoglobulin (Applied Biosystems or Sigma) as a control for the procedure.

9. Electrophorese sample by applying 10 mA until pink tracking dye (pyronin Y; see support protocol) reaches bottom of gel. Turn off power supply.

The tracking dye should focus to sharp zones within 5 to 10 min. Pyronin Y is used instead of bromphenol blue because it binds tightly to transfer membranes, thereby identifying the bottom of the blot. Electrophoresis should be complete in 60 to 80 min.

Transfer the minigel samples to PVDF membranes

NOTE: Wear powder-free plastic gloves for all subsequent steps; one fingerprint contains many times more amino acids than the sample.

10. Wet two PVDF membranes with methanol and immerse them in transfer buffer.
11. Disassemble minigel unit and gently separate glass plates.
12. Assemble blotting sandwich supplied with the small-format transfer apparatus in the following sequence: plastic frame, sponge, filter paper, gel, two PVDF membranes, filter paper, sponge, plastic frame (see UNIT 10.8, Fig. 10.8.1; substitute PVDF membrane for nitrocellulose).

As with other blotting procedures, it is crucial to exclude air bubbles from the sandwich. The easiest way to do this is to assemble the sandwich in a tray filled with the transfer buffer. PVDF membranes may be placed on both sides of the gel to guard against inadvertent reversal of the electrode connections if desired.

13. Insert sandwich into transfer apparatus, placing membranes closest to the anode (red or positive electrode). Fill apparatus with transfer buffer.
14. Apply ~6 V/cm (e.g., 50 V in the LKB Midget MultiBlot apparatus) across electrodes of transfer apparatus for a period of time appropriate to protein of interest and gel concentration. Turn off power supply.

The efficiency with which a particular protein will transfer to the PVDF membrane depends on several variables. Transfer shows a rough inverse correlation with molecular weight, but there are many marked exceptions. Isoelectric point, hydrophobicity, and tertiary structure following partial renaturation of the protein are all important. Many smaller, highly charged proteins will travel through conventional PVDF membranes before others have migrated out of the gel, so a particular transfer should be optimized for the protein of interest. This means that some trial and error may be necessary. Duplicate PVDF sheets allow material migrating through the first membrane to be recovered and, in conjunction with staining of the gel following transfer, help optimize subsequent experiments.

PVDF-based membranes have been developed (e.g., ProBlott from Applied Biosystems and Immobilon-PSQ from Millipore) which adsorb blotted proteins much more avidly than others. These membranes make the transfer less critical for most samples and also permits analysis of low-molecular-weight peptides. With these membranes, one can err on the side of prolonged transfer without adverse effects.

At the specified voltage, transfer takes 30 to 40 min for 20- to 50-kDa proteins in a 15% gel, 60 to 70 min for 50- to 100-kDa proteins in a 10% to 12% gel, and 90 min for 150- to

200-kDa proteins in a 7% gel. For proteins >60 kDa, reduce the amount of methanol in transfer buffer to 1%.

Stain the blots and excise the bands

15. Disassemble transfer apparatus. Immerse PVDF membrane blots in 0.1% Coomassie blue in 50% methanol and agitate 5 min.
16. Destain blots in 10% acetic acid prepared in 50% methanol by agitating until bands become clearly visible (5 to 10 min).
17. Transfer to water and photograph (optional; *UNIT 10.6*).
18. Excise band of interest with a razor blade (use a new razor blade for each band). Place each band in a microcentrifuge tube and allow to air dry at room temperature (do not heat). Store excised bands at -20°C .

Sequence the proteins

19. Insert excised band into sequencer reaction cartridge, protein side facing the solvent delivery (see manufacturer's instructions).

All pieces of PVDF must fit in a single layer in the reaction cartridge; they may be cut or trimmed to fit as required. When using conventional chemistry ("NORMAL" program on Applied Biosystems instruments), two alternate configurations have proven optimal, depending on the sequencer. For many gas-phase instruments, placing the PVDF pieces between the upper-cartridge block and the TFA-etched glass-fiber filter gives the best results. For some gas-phase and most pulsed-liquid instruments, placement between the glass-fiber disk and the cartridge seal is better. The results of the cartridge leak test should be normal. The Blott cartridge should be used on Applied Biosystems instruments, if available.

20. Sequence the sample on an automated protein sequencer.

The "Blott-2 reaction cycle" should be used if a Blott reaction cartridge is installed. The conversion cycle described by Tempst and Riviere (1989) is recommended.

SUPPORT PROTOCOL

PREPARATION OF PROTEIN SAMPLES FOR SDS-PAGE

Samples are concentrated and freed of contaminants that interfere with electrophoresis by precipitation. They are then solubilized in a sample buffer containing SDS.

Additional Materials

Protein samples
1 M NaHCO_3 (optional)
100% ethanol, ice-cold (containing no denaturants; USP grade)
Sample buffer
0.1% (w/v) pyronin Y
Ultrafiltration concentrator (Amicon) or Speedvac evaporator (Savant)
Drawn-out Pasteur pipet *or* gel-loading pipet tip
Boiling water bath

1. Adjust the salt concentration of the protein sample to >100 mmol/liter with 1 M NaHCO_3 , if necessary.
2. Concentrate samples to 50 to 100 μl by ultrafiltration (follow manufacturer's instructions explicitly) or vacuum centrifugation (it is crucial to avoid introduction of airborne debris when vacuum is released). Transfer to 1.5-ml microcentrifuge tubes.

3. Add 9 vol ice-cold 100% ethanol to the samples. Incubate 1 hr on dry ice or overnight at -20°C .

Most samples may be kept indefinitely at this stage.

4. Microcentrifuge 15 min at maximum speed. Aspirate the supernatant with a drawn-out Pasteur pipet or a gel-loading pipet tip and save the pellet.

A drum rotor, which holds the microcentrifuge tubes at a 90° angle, is more likely to give compact, easily visualized pellets than the commonly used fixed-angle rotor.

5. Dissolve the pellet in 10 μl sample buffer by drawing the sample buffer up and down with a pipettor. Boil 3 min in a boiling water bath.
6. Add 1 μl of 0.1% pyronin Y (tracking dye) to the sample. Load on the minigel in the basic protocol.

DETERMINATION OF INTERNAL AMINO ACID SEQUENCE FROM ELECTROPHORETICALLY-SEPARATED PROTEINS

BASIC PROTOCOL

As many as 50% of all eukaryotic proteins are blocked at the amino terminus, making sequence determination of the intact protein impossible. The method presented here allows multiple stretches of internal amino acid sequence to be obtained from most proteins that can be isolated by any one- or two-dimensional gel electrophoresis and transfer method. About 200 pmol of protein is an adequate sample size; smaller amounts have sufficed in several cases.

The protein(s) of interest is separated by electrophoresis and transferred to nitrocellulose. Protein bands or spots are visualized with Ponceau S and excised. After destaining, the nitrocellulose is treated with polyvinylpyrrolidone (PVP) to prevent adsorption and inactivation of proteolytic enzyme, which is added to cleave the protein. The resulting peptides, which elute from the membrane, are separated by reversed-phase HPLC. Each peptide is analyzed individually in an automated microsequencer.

Materials

- Protein sample
- 0.1% (w/v) Ponceau S (Sigma) in 1% (v/v) acetic acid
- 1% acetic acid (v/v)
- 0.2 mM NaOH
- 0.5% (w/v) polyvinylpyrrolidone (Sigma # PVP-40) in 0.1 M acetic acid
- Digestion buffer
- 1 mg/ml sequencing-grade trypsin (Promega #V511A)
- Chromatography solvent A: 5% (v/v) acetonitrile in 0.1% (v/v) trifluoroacetic acid (TFA)
- Chromatography solvent B: 70% (v/v) acetonitrile in 0.085% (v/v) TFA
- 0.22- μm nitrocellulose membrane (e.g., Schleicher & Schuell #BA83)
- Acid-washed glass plate or petri dish
- Powder-free gloves
- Fine-tipped forceps
- 0.5-ml microcentrifuge tube
- Bath sonicator (e.g., Branson 12)
- Centrifugal filter device, 0.22- μm membrane, low-protein-binding (e.g., Millipore #UFC3-OGV-00)
- Reversed-phase HPLC column (e.g., Vydac #214TP52), UV column monitor, and chart recorder
- Column oven (optional)

Analysis of Proteins

10.19.5

Electrophoresis and transfer the proteins

1. Resolve protein(s) of interest by electrophoresis and transfer to nitrocellulose membrane as described in *UNIT 10.8*.

It is important to end up with as much protein as possible bound to the smallest possible area of nitrocellulose. Apart from this constraint, there are few limitations on the design of the separation, as chemical modification of the amino terminus will not interfere with subsequent analyses.

Although PVDF membranes may be used, nitrocellulose seems to give better yields (see commentary).

The membrane must never be allowed to dry from this point on.

Stain and excise the separated proteins

2. Place nitrocellulose membrane in an acid-washed glass petri dish (or similar vessel) containing 50 ml (for an 8 × 10-cm minigel) of 0.1% Ponceau S prepared in 1% aqueous acetic acid. Agitate gently 1 min.

For larger blots, increase the volume of stain in proportion to the area.

3. Transfer to 1% acetic acid for 1 min, changing the solution as necessary to allow easy visualization of band(s).

Use enough 1% acetic acid to cover the membrane generously.

4. Using a new razor blade and wearing powder-free gloves, cut out the band(s) of interest. Place in 1.5-ml microcentrifuge tube(s).

It is important to remove all areas of blank nitrocellulose to minimize subsequent adsorption of the protease used for digestion. Meticulous technique is crucial at this step to eliminate contamination of the sample by adventitious proteins. A clean work area, scrupulously clean glassware and instruments, and powder-free gloves are minimum precautions; tryptic fragments of human keratin have been identified in some samples!

To distinguish UV-absorbing peaks corresponding to the protein from those derived from the proteolytic enzyme or other reaction constituents, cut a piece of nitrocellulose from a blank area of the blot to use as a control.

The bands may be transferred to microcentrifuge tubes containing 0.5 ml water and stored at -20°C at this step.

Destain the membrane pieces

5. Transfer membrane pieces to 1 ml of 0.2 mM NaOH and vortex 1 min. Aspirate the NaOH and immediately proceed to step 6.

Proteins are more apt to be lost from the nitrocellulose under alkaline than acidic conditions. Because release of the peptides is desired at a later stage, the length of time for destaining should be minimized. Residual stain on the blot will not adversely affect the final results.

Block the membrane

6. Add 1 ml of 0.5% PVP-40 in 0.1 M acetic acid and agitate tube gently 30 min at room temperature.
7. Aspirate the PVP-40 and wash membrane five times with 1 ml water. Be sure to remove any liquid droplets caught under tube cap.

- Using clean fine-tipped forceps, transfer excised band to a clean glass surface (e.g., acid-washed glass plate or petri dish) and cut it into 1- to 2-mm pieces. Use forceps to collect the pieces and squeeze out excess liquid. Immediately proceed to step 9.

Digest the protein with protease

- Transfer pieces to a 0.5-ml microcentrifuge containing 25 μ l digestion buffer. Add 1 μ l 1 mg/ml trypsin. Mix so that membrane pieces are evenly coated with solution.
- Incubate overnight at room temperature.

Elute the peptides

- Microcentrifuge sample 1 sec at high speed, room temperature, to recover liquid that may have condensed on tube walls and cap.
- Sonicate sample 5 min at room temperature.
- Microcentrifuge 1 min at top speed, room temperature.
- Transfer supernatant to a centrifugal filter device. Rinse membrane pieces with 100 μ l digestion buffer and add to supernatant. Microcentrifuge sample 20 to 30 sec at top speed.

Filtration removes nitrocellulose particles that might clog the HPLC column. Samples should be stored frozen unless they will be fractionated immediately.

Chromatograph and process the samples

- Equilibrate a 2.1-mm-i.d. \times 250-mm reversed-phase HPLC column with 95% chromatography solvent A/5% solvent B at a flow rate of 0.15 ml/min. For optimal resolution, perform the separation at 60°C if a column oven is available.
- Inject the sample. Wash column 10 to 15 min with 95% solvent A/5% solvent B.
- Elute the peptides with a gradient between chromatography solvents A and B as follows: 5% to 40% solvent B over 1 hr; 40% to 75% solvent B over 30 min; and 75% to 100% solvent B over 15 min. Monitor elution of the peptides at 215 nm.

Elution gradient guidelines are described in Stone and Williams (1986).

The TFA concentration should be adjusted to equalize the UV absorbance (215 nm) of the eluants. Alternatives to TFA, such as phosphoric acid or hydrochloric acid, are compatible with the procedure. Buffers containing ammonia or UV-containing impurities should not be used.

Trypsin elutes at 60% solvent B and can be used as an internal standard. Characteristically, it elutes in a broader peak than the majority of peptides.

- Monitor the appearance of peptide peaks with a chart recorder adjusted so that 0.05 to 0.1 AUFS corresponds to a full-scale deflection.

Minimize the length and capacity of the capillary tubing between the column and the point of collection as much as possible. Polyethylether ketone (PEEK) tubing (0.005-in. i.d.) is helpful for this purpose.

- When chart-recorder pen begins a deflection indicative of a peak, wipe tip of capillary tubing against a clean surface (e.g., a Kimwipe or the previous collection tube) and collect this fraction in a microcentrifuge tube.

Clean gloves should be worn at this point. Commercial microcentrifuge tubes are sufficiently clean as supplied by the manufacturer if they are stored in the original container and handled with clean gloves.

20. Immediately cap tube and place it on dry ice.

Rapid freezing prevents adsorption of separated peptides to the walls of the microcentrifuge tube. The samples can be stored at -80°C indefinitely.

21. Sequence samples using glass support disks precycled with polybrene as prescribed by Applied Biosystems.

Improved reaction (Speicher, 1989) and conversion (Tempst and Riviere, 1989) cycles will reduce background and improve sensitivity.

It is very useful to compare chromatograms obtained from the sample and from the blank nitrocellulose to identify peaks corresponding to peptides derived from the protein of interest. UV-absorbing impurities may originate from several sources, including residual PVP-40, contaminated glassware, and the proteolytic enzyme (autolysis of the protease). These can be eliminated easily by inspecting the control chromatogram. Peaks as small as 0.002 AUFS have often been sequenced successfully.

REAGENTS AND SOLUTIONS

Digestion buffer

0.1 M Tris-Cl, pH 8
1 mM CaCl_2
10% (v/v) acetonitrile
Store frozen

4× gel buffer

41.24 g *bis*-Tris (0.493 M)
7.08 ml 37.5% HCl (reagent grade)
 H_2O to 400 ml
Final pH should be 6.61; do not adjust the pH with acid or base. This buffer may be stored indefinitely at -20°C or ≤ 2 months at 4°C .

10× lower reservoir buffer

52.4 g *bis*-Tris (0.626M)
12.2 ml 37.5% HCl
 H_2O to 400 ml
Final pH should be 5.90; do not adjust the pH with acid or base. This buffer may be stored indefinitely at -20°C or ≤ 2 months at 4°C .

10× upper reservoir buffer

40.28 g *N*-tris-(hydroxymethyl)-methyl-2-aminoethanesulfonic acid (TES; 0.439 M)
94.64 g *bis*-Tris (1.131 M)
4.0 g sodium dodecyl sulfate (SDS)
 H_2O to 400 ml
Final pH should be 7.25; do not adjust the pH with acid or base. This buffer may be stored indefinitely at -20°C or ≤ 2 months at 4°C .

Sample buffer

50 mM Tris-Cl, pH 6.8
5% (v/v) 2-mercaptoethanol
10% (v/v) glycerol
1% (w/v) SDS
Store frozen in small aliquots for ≤ 2 months

(continued)

Transfer buffer

2.21 g cyclohexylaminopropane sulfonic acid (CAPS), free acid
0.5 g dithiothreitol
150 ml methanol
H₂O to 1 liter
Adjust pH to 10.5 with NaOH and chill to 4°C
Prepare just before use

For proteins >60 kDa, reduce amount of methanol to 1%.

COMMENTARY

Background Information

The most difficult part of any protein sequencing project is to obtain a sample that is sufficiently free of both protein and nonprotein contaminants. SDS-PAGE is used because of its resolving power and its ability to maintain solubility of almost any protein. Many proteins of interest are identified only by migration on an analytical gel or by immunoreactivity in immunoblot experiments. The sequencing procedures outlined here allow the direct comparison of primary structures.

Following the demonstration that blotting procedures could be used to prepare samples for microsequencing (Aebersold et al., 1986; Vandekerckhove et al., 1986; Matsudaira, 1987), acceptance of these techniques was slow due to the inability of many investigators to obtain adequate sequencing signals. Evaluation of several parameters has indicated that preelectrophoresis with a scavenger such as glutathione or mercaptoacetic acid to eliminate reactive by-products of acrylamide polymerization, and lowering the separation pH to reduce the reactivity of the N-terminal amine, can dramatically improve the signals (Moos et al., 1988).

Recently, Tris-Tricine gels have become popular, especially for low-molecular-weight proteins and peptides (Schägger and von Jagow, 1987). These gels address the concerns of pH-dependent N-terminal blockage and give good results in sequencing experiments. Their commercial availability in precast format makes them an attractive choice for the novice.

The ability to sequence proteins blocked at the amino terminus following electrophoretic separation on gels (second basic protocol; Aebersold et al., 1987; Tempst et al., 1990) is an important improvement in protein analysis. Although meticulous cleanliness and exacting technique are required, the protocol is straightforward and highly reliable. Its particular strength is that it can be used following any type of gel electrophoresis or electrophoretic trans-

fer. It is even possible to combine small spots obtained from several gels (e.g., minor components detected by two-dimensional analysis). Its principal limitation is the relatively low yield of material for sequencing from a given amount of blotted protein; a dark band of Ponceau S-staining material corresponding to 100 to 250 pmol (or more for some proteins) is often required. This problem is obviated somewhat by the extremely high quality of sequence data that can be obtained from minute (e.g., 2 mAUFs) HPLC peaks.

Modifications to enhance the sensitivity of the microsequencer 5- to 8-fold have been developed (Tempst and Riviere, 1989); these have proven extremely useful for all microsequencing procedures. A new automated chemistry cycle specifically optimized for PVDF significantly improves results (Speicher, 1989). Applied Biosystems has developed a reaction cartridge designed for PVDF blots that simplifies sample placement and substantially improves sequencing results (Sheer et al., 1990). Finally, a procedure for covalently entrapping blotted proteins in a polymer network may prove beneficial (information available from Milligen/Biosearch; see APPENDIX 4).

Critical Parameters and Troubleshooting

Determination of amino acid sequence by SDS-PAGE and transfer to PVDF membranes

In practice, the electrophoresis conditions described in the first basic protocol improve the initial sequencing yield by 2- to 3-fold over those obtained with earlier procedures. Because many proteins are available in very small quantities, this consideration is often crucial. If the sample is *not* in short supply, it may be reasonable to attempt sequencing with electrophoresis conditions commonly in use (UNIT 10.2; Matsudaira, 1987), especially if the reagents are available in the laboratory. Samples already

transferred to PVDF by other procedures may also be tried; these should be rewet with methanol and washed with water (3 times, 5 min each) before sequencing to remove dust and materials such as glycine that may have been used in the transfer.

Efficient acrylamide polymerization is crucial to obtain good sequencing signals. Even the most efficiently polymerized gels contain 50 to 60 mM acrylamide monomer and possibly other compounds that react in a pH-dependent manner with the N terminus during electrophoresis. Therefore, fresh, high-quality reagents, ultrapure water, and thorough deaeration of the polymerization mixtures are necessary. TEMED is the most likely reagent to deteriorate and a fresh bottle should be opened each month.

The buffer system used in this protocol operates at a lower ionic strength than others in common use and therefore is more sensitive to the presence of ionic impurities in the sample. The most efficient way to remove these is by precipitation with organic solvent. Ethanol was chosen for this application (support protocol) because it is least likely to contain impurities that could react with amino termini of proteins (Yuan et al., 1987). Precipitation is most efficient with concentrated samples, so centrifugal ultrafiltration can be useful. Too much SDS in the sample will impair stacking. This is why only 1% SDS, which is sufficient to saturate almost all proteins, is used in the sample buffer. Pyronin Y is used as a tracking dye because it binds to PVDF tightly and unambiguously locates the dye front.

The gel buffers and polymerization mixtures should be prepared exactly as specified. No adjustments in pH should be necessary. SDS is not included in the gel polymerization mixtures as this could impair stacking. The same buffer is used in both the stacking and resolving gel. If efficient stacking and good resolution are not observed, the buffers are incorrect, the sample contains too much salt (see support protocol), or the gel is overloaded (10 to 20 $\mu\text{g}/\text{band}$ is about right).

If poor sequencing yields are obtained with a standard protein known to be unblocked, there are problems with gel polymerization, transfer, or placement of the blots in the reaction cartridge of the sequencer. Optimization of transfer is straightforward, and samples that have been applied directly to PVDF pieces without prior electrophoresis can be used to assess the transfer. It is strongly recommended that β -lac-

toglobulin, which has worked well in many laboratories and is widely available, be used on a test run before using valuable experimental samples. In addition, it is a good idea to use β -lactoglobulin as a positive control in every experiment. This will distinguish between existing N-terminal blockage and blockage that may occur during the procedure. Exposure of the blots to high pH and high temperature should be minimized.

When preparing the sample for electrophoresis, the supernatant should be aspirated very carefully—a magnifying lens may be helpful to avoid aspirating the pellet. A 10- μg pellet is usually visible to the naked eye.

Determination of internal amino acid sequence from electrophoretically-separated proteins

Once a sufficiently pure protein has been obtained, the most troublesome aspect of sequencing is chemical modification of the N terminus of the protein, either *in vivo* or *in vitro*. Approximately 50% of eukaryotic cellular proteins are thought to have naturally occurring N-terminal blockage. For these proteins, enzymatic or chemical cleavage procedures are required in order to obtain sequence information by automated Edman degradation (Hewick et al., 1981; Tempst and Riviere, 1989). In addition, many reagents commonly encountered in protein isolation procedures can irreversibly modify unblocked N-terminal amino groups and prevent Edman degradation. Because of this, a standard protein, such as β -lactoglobulin, should be analyzed as a control. Failure to obtain any sequence or a low-yield sequence for the standard protein suggests problems with the reagents.

The success of this method depends on a number of factors. (1) A reasonably homogeneous band or spot must be obtained. If two or more proteins are overlapping, it will be difficult to resolve individual peptides and impossible to assign sequence data to specific proteins. (2) It is important to maintain meticulous cleanliness throughout the procedure. Keratins from the skin are particularly problematic. (3) UV-absorbing contaminants in polyvinylpyrrolidone (PVP) can obscure the peptide peaks, which are often very small. Therefore, careful washing after the PVP blocking step is necessary. (4) The digestion should be performed in a minimum volume, as proteases are less active when dilute. The high concentrations used (compared to traditional

solution digestion procedures) are not a problem because proteases are generally separated from most of the peptides by HPLC. If an autolysis product should be isolated, this will be apparent when its sequence is compared with databases. (5) Careful manual collection of the HPLC fractions can improve the results; a bit of practice, and paying special attention to changes in the *rate of change* of the recorder pen's movement will be helpful. In general, quite closely spaced peaks will give unique amino acid sequences.

In the author's experience, trypsin is the most reliable enzyme for this procedure. Other enzymes—including Lys-C, Asp-N, and Glu-C—may also be tried to provide overlapping sequences or in case the tryptic pattern does not provide the desired information. In this event, the digestion buffer recommended by the supplier of the protease, supplemented as above with 10% acetonitrile, should be substituted.

The most common problems are spurious peaks, no peaks, and only a trypsin peak in the HPLC analysis. Spurious peaks arise from UV-absorbing contaminants, as discussed above. If a trypsin peak is absent, a control reaction containing the buffer and protease but no protein or nitrocellulose should be injected. If there is still nothing, the trypsin concentration should be checked. If a trypsin peak is present in the control samples, blocking was inadequate or the amount of nitrocellulose without adsorbed protein was too high. If only a trypsin peak is detected, either not enough protein was present or the trypsin was inactive. The latter possibility can be assessed by digesting some standard protein with the same lot of protease.

Anticipated Results

In the first basic protocol, if 100 pmol β -lactoglobulin are loaded onto a 12% gel prepared as described, the initial sequencing signal should indicate a yield of 50 to 80 pmol with gas-phase instruments (e.g., Applied Biosystems 470A). In some cases, usable sequence may be obtained from as little as 10 pmol. In general, if a band is easy to see with Coomassie blue staining, there is enough to sequence, but if it is so faint that excision is difficult, it is doubtful that there is enough protein. However, very faint bands, particularly of low-molecular-weight proteins, have yielded sequence.

For N-terminally blocked proteins, if enough material is present in a band or spot, usable sequence from 2 to 3 peptides can be obtained for most proteins. In several cases, two dozen or more HPLC peaks from a single band

have yielded high-quality sequence data, allowing determination of as many as two hundred residues from a single experiment.

Though the efficiency of the precipitation procedure (support protocol) varies with the sample, it is usual to recover 70% to 100% of samples >1 μg from a volume of 50 μl .

Time Considerations

Time required for the SDS-PAGE and transfer procedure: casting the gels—20 min hands-on time and ~2 hr total, including deaeration and polymerization; preelectrophoresis—just a few minutes hands-on time and 45 min total; electrophoreses—60 to 80 min; setting up the transfer—15 min; transfer—30 to 90 min; staining—15 min; band excision—5 to 30 min; sample preparation—15 min hands-on time and 1 hr for incubation on dry ice. The entire procedure can be completed in a single day, though it is preferable to cast the gels a day in advance.

Time required for determination of internal amino acid sequences following electrophoretic separation and transfer of proteins: destaining and PVP blocking—30 min; washing—5 min/sample; mincing the nitrocellulose—2 to 3 min/sample; digestions—overnight 12 to 24 hr; sample recovery and filtration—10 min; reversed-phase HPLC separations—2 hr/sample. A set of samples may be separated on a gel, transferred, and processed for digestion in ~5 hr. If several proteins from a given gel are of interest, a few days of HPLC will be required; at 2 hr per run, only 3 samples per day can be analyzed.

Literature Cited

- Aebersold, R.H., Teplov, D.B., Hood, L.E., and Kent, S.B.H. 1986. Electroblothing onto activated glass. *J. Biol. Chem.* 261:4229-4238.
- Aebersold, R.H., Leavitt, J., Saavedra, R.A., Hood, L.E., and Kent, S.B.H. 1987. Internal amino acid sequence analysis of proteins separated by one- or two-dimensional electrophoresis after in situ protease digestion on nitrocellulose. *Proc. Natl. Acad. Sci. U.S.A.* 84:6970-6974.
- Hewick, R.M., Hunkapiller, M.W., Hood, L.E., and Dreyer, W.J. 1981. A gas-liquid solid phase peptide and protein sequenator. *J. Biol. Chem.* 256:7990-7997.
- Matsudaira, P. 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.* 262:10035-10038.
- Moos, M., Nguyen, N.Y., and Liu, T-Y. 1988. Reproducible, high-yield sequencing of proteins electrophoretically separated and transferred to an inert support. *J. Biol. Chem.* 263:6005-6008.

- Schägger, H. and von Jagow, G. 1987. Tricine-sodium dodecylsulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 KDa. *Anal. Biochem.* 166:368-379.
- Sheer, D.G., Yuen, S.W., Yamone, D.K., Mattaliano, R.J., and Yuan, P-M. 1990. A modified reaction cartridge for sequencing samples on polymeric sample matrices. Abstract T139, Fourth Symposium of the Protein Society, San Diego.
- Speicher, D.W. 1989. Microsequencing with PVDF membranes: Efficient electroblotting, direct protein adsorption and sequencer program modifications. *In* Techniques in Protein Chemistry (T.E. Hugli, ed.) pp. 24-35. Academic Press, San Diego.
- Stone, K.L. and Williams, K.R. 1986. High-performance liquid chromatographic peptide mapping and amino acid analysis in the subnanomole range. *J. Chromatogr.* 359:203-212.
- Tempst, P. and Riviere, L. 1989. Examination of automated polypeptide sequencing using standard phenylisothiocyanate reagent and subpicomole high-performance liquid chromatographic analysis. *Anal. Biochem.* 183:290-300.
- Tempst, P., Link, A.J., Riviere, L.R., Fleming, M., and Elicone, C. 1990. Internal sequence analysis of proteins separated on polyacrylamide gels at the picomole level: Improved methods, applications and gene cloning strategies. *Electrophoresis* 11:537-553.
- Vandekerckhove, J., Bauw, G., Puype, M., van Damme, J., and van Montagu, M. 1986. Protein blotting from polyacrylamide gels onto glass microfiber filters. *In* Advanced Methods in Protein Microsequence Analysis (B. Wittmann-Liebold, J. Salnikow, and V.A. Erdman, eds.) pp. 179-193. Springer-Verlag, Berlin.
- Yuan, P., Hawke, D., Blacher, R., Hunkapiller, M., and Wilson, K. 1987. Ethanol precipitation of electroeluted-electrodialyzed samples. Applied Biosystems User Bulletin #27.

Key References

Moos et al., 1988. See above.

Tempst, et al. See above.

Address common problems encountered in procedures of this type.

Contributed by Malcolm Moos, Jr.
Center for Biologics Evaluation & Research
Food and Drug Administration
Bethesda, Maryland

Capillary Electrophoresis of Proteins and Peptides

Capillary electrophoresis (CE) is a high-resolution technique for the separation of a wide variety of molecules of biological interest such as metabolites, drugs, amino acids, nucleic acids, and carbohydrates. This unit focuses on the use of CE to separate proteins and peptides. As with polyacrylamide gel electrophoresis (PAGE; UNIT 10.2), CE separations of proteins and peptides are based on charge-to-mass ratios. While PAGE separations are restricted to polyacrylamide matrices and a relatively small number of buffer systems, CE separations can be achieved in a variety of different matrices using a wide range of electrophoresis buffers. Consequently, there is a much greater flexibility in the design of optimal separation protocols.

CE employs a fused-silica capillary column, which may or may not be derivatized, either in free solution or in the presence of a fluid matrix. In contrast to PAGE, the resulting protein or peptide bands cannot be fixed or stained. The CE separation is, in essence, a dynamic one and is more analogous to high-performance liquid chromatography (HPLC) than PAGE. However, CE separation is based on electrophoretic parameters, so it can be viewed as orthogonal to HPLC, which is based predominantly on solvent partitioning parameters. The combination of these two techniques is very powerful for analyzing and characterizing proteins and peptides. The initial expectation that CE would play a major role in protein separations has not been realized. However, CE is particularly useful for determining the purity of a sample and for rapid, efficient, and quantitative evaluation of protein purification.

For best separation results, it is necessary to optimize the procedure. If some properties of the protein are known—e.g., its isoelectric point (pI)—the selection of a suitable running buffer is relatively straightforward. In fact, CE itself can be used to determine the isoelectric point of a protein, either in purified form or in a mixture, by focusing the sample in a pH gradient that is generated within the capillary during electrophoresis (Basic Protocol 1). After refocusing, the sample is mobilized either by changing the anode buffer, or by applying a hydrodynamic force to the column to move the separated components past the detector. However, there is little information available that will predict the potential of a protein to interact with the capillary wall. Underivatized silica capillaries have a strong tendency to adsorb proteins, thereby immobilizing them and making protein separations impractical. If such interactions occur, then an additive or a coated capillary column is required. A certain amount of trial and error is involved, and it may be necessary to perform several CE runs to optimize the separation conditions for a given protein. One approach to optimization is described in Basic Protocol 2.

CE is most useful for separations of peptides (Basic Protocol 3), because it offers great flexibility in separation parameters. It can be used to monitor proteolytic digestions and optimize digestion conditions for the production of a representative peptide fingerprint of a protein. This profile can subsequently be used to provide structural information about the protein, especially when used in conjunction with reversed-phase HPLC (RP-HPLC; UNIT 10.14). It can be used to screen peptide fractions that are obtained from a preparative RP-HPLC separation of a protease digestion. Fractions that contain single or major components are suitable candidates for protein sequence analysis. Similarly, CE can be used to assess the purity of synthetic peptides. In the presence of an internal standard, it can provide quantitative information about the various components that are present in a peptide mixture. CE can also be used as a micropreparative technique—with either multiple separations that are pooled (Basic Protocol 4) or a single, larger-scale separation

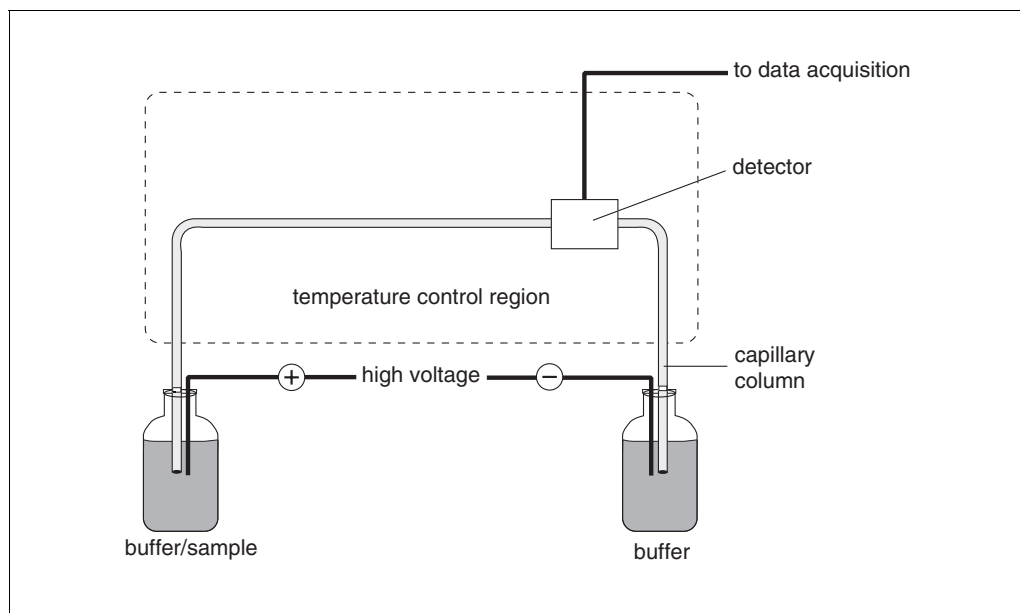


Figure 10.20.1 Schematic of a CE instrument.

(Alternate Protocol)—for the isolation of peptides from a protease digestion (in much the same way that RP-HPLC is currently used). In most of these examples the same capillary column can be used for all the separations. Only changes in buffer composition, ionic strength, and the presence or absence of additives are required for each specific application.

INSTRUMENTATION

CE separation in its simplest form can be achieved by passing a high voltage between two buffer reservoirs that are joined by a liquid-filled fused-silica capillary (Fig. 10.20.1). This results in the generation of electroosmotic flow (EOF; see Separation Theory), which allows the molecules of interest to be carried from one end of the capillary to the other. The capillaries are generally 30 to 50 cm long with 50 to 75 μm i.d. The net total volume of these capillaries is in the low microliter range. For comparison, the volume of a slab gel lane is ~ 1000 μl . The capillaries are thin-walled; this allows rapid and efficient exchange of the Joule heating that results from the high voltages (20 to 25 kV) that are necessary for electrophoretic separations. This heat dissipation minimizes the negative convective effects that could result in band-broadening during electrophoresis. The fused-silica capillary is coated on the outside with a polyimide layer that confers excellent tensile strength to the otherwise fragile capillary. The polyimide sheathing is carefully burned from a small portion of the capillary to expose a section of the silica. This clear section of the capillary is inserted into the light path of a UV detector and becomes the flow cell. As the protein and peptide molecules are swept through the capillary by EOF, they pass through the detector light path and are registered on the UV monitor. In effect the capillary becomes a very-low-volume flow cell.

IMPORTANT NOTE: The removal of the polyimide coating makes the capillaries susceptible to breakage. Capillaries that are not provided in cartridges by the manufacturer should be handled with care to avoid breakage.

The combination of high field strength and large surface-to-volume ratios results in rapid and very efficient separations of both proteins and peptides. Sample loading volumes are routinely in the nanoliter range, with starting sample concentrations of ~ 0.1 $\mu\text{g}/\mu\text{l}$ for UV detection. On-capillary preconcentration protocols are available when starting concentra-

Table 10.20.1 Commercially Available CE Instruments

Manufacturer	Model	Detection method	Reversal of polarity	Capillary heating	Cooling of sample/buffer	Capillary format	Software available
Applied Biosystems	270-HT	UV: 190-700 nm	Controlled by software	FAC	Sample	Free-hanging	DA
Beckman	P/ACE 5000 P/ACE 5510 P/ACE 5510	UV: filter; Diode-array: 190-600 nm; Fluorescence: argon laser; MS interface	Manual	Liquid (Peltier)	Sample/buffer	Cartridge	C, DA
Bio-Rad	Biofocus 3000	UV: 190-800 nm	Controlled by software	Liquid	Sample/buffer	Cartridge	C, DA
Hewlett-Packard	HP 3D CE	Diode-array: 190-600 nm	Controlled by software	FAC (Peltier)	Sample/buffer	Cartridge	C, DA
Waters	Quantra 4000E	UV: filter	Manual	FAC (Peltier)	Sample/buffer	Free-hanging	DA

^aAll instruments have hydrostatic/electrokinetic sample loading and fraction collecting.

^bAbbreviations: C, CE system control software; DA, data acquisition; FAC, forced-air convection; MS, mass spectrometry.

tions are below these levels. Sample loading can be achieved either hydrodynamically or electrokinetically. Hydrodynamic loading can be performed by pressure injection, vacuum injection, or gravity injection, e.g., pressure loading at 0.5 lb/in² for 5 sec. Any of these methods will deliver a small aliquot of sample into the capillary column. Hydrodynamic sample loading has limitations where the capillary contains a viscous matrix or the sample itself has a high viscosity. Electrokinetic loading is performed by applying a low voltage across the capillary column for a fixed time, e.g., 7.5 kV for 4 sec. Sample buffer conductivity, run buffer conductivity, and EOF all play a part in determining the amount of material loaded into the capillary column. In general, samples in high salt buffers should be avoided because buffer ions will carry the charge rather than the sample ions. There are several methods for concentrating the sample onto the capillary column (Burgi and Chien, 1992). All instruments listed in Table 10.20.1 have both sample loading capabilities. Clearly, with respect to sensitivity, speed, and versatility, CE can offer some significant advantages over gel electrophoresis for the separation of proteins and peptides.

As shown in Figure 10.20.1, CE can require minimal instrumentation. However, in reality, due to the high voltages that are utilized, safety issues require that the capillary column be incorporated into a dedicated CE instrument. These instruments can provide efficient capillary cooling and online detection of analytes. Table 10.20.1 lists a number of commercially available instruments and some of their characteristics. For the purpose of this discussion, it is assumed that a generic CE system possessing the following capabilities is used: an active capillary cooling system, a UV detector capable of monitoring at 200 nm, a thermostatted autosampler, and a chromatographic data package. Given this basic unit a variety of separations can be performed, depending upon the nature of the proteins or peptides that are being separated.

SEPARATION THEORY

CE is part of the family of electrophoretic techniques that separate species based upon their sizes and ionic properties. An ion (*i*) placed into an electric field will move in the direction parallel to the field with a velocity (v_i) as follows:

$$v_i = \mu_i E$$

Here μ_i is the electrophoretic mobility of species i , and E is the electric field, as defined by the following equation, where V is the voltage and L is the total length from electrode to electrode:

$$E = \frac{V}{L}$$

The electrophoretic mobility of an ion (μ_i) is:

$$\mu_i = \frac{q}{6 \pi \eta a_i}$$

where q is the charge on the ion, η is the viscosity of the solution, and a_i is the radius of the ion.

As seen from these three equations, the movement of the molecule in the column is dependent on the applied voltage, the length of the column, the charge on the molecule, and the size of the molecule.

In CE, buffer flow is generated inside the column when the electric field is applied. This flow is from the cathode electrode to the anode electrode in a fused-silica capillary column (from left to right in Fig. 10.20.1); this movement of the buffer is called the electroosmotic flow (EOF). In addition, normal electrophoretic separation occurs whereby a positively charged molecule moves in the same direction as the EOF, a negatively charged molecule moves in a direction opposite to the flow, and a neutral molecule is carried along by the flow. Thus, the total velocity of the molecule (μ_t) is given by the following equation, where μ_{eo} is the EOF in the column.

$$\mu_t = (\mu_{eo} + \mu_i) E$$

The magnitude of the EOF will vary as a function of the pH of the carrier buffer. In uncoated fused-silica columns, a low pH generates a slow EOF, whereas a high pH generates a fast flow. The flow can be suppressed or completely reversed depending on the coating on the column or organic modifier added to the carrier buffer (Landers et al., 1992; Tsuiji and Little, 1992). Thus, EOF is another parameter that can be optimized to aid in difficult separations. Neutral molecules, because they are all carried along by the flow, can be difficult to separate. Addition of a micelle to the carrier buffer results in the partitioning of molecules between the micelle and carrier buffer and can effectively resolve neutral molecules (Khaledi, 1994).

STRATEGIC PLANNING

Proteins can be separated on coated or uncoated capillary columns, and the choice of separation protocol depends on the specific properties of the target protein. The most important property is the pI, which can be determined by either conventional gel electrophoresis or CE. The use of a buffer ~2 pH units above the pI is optimal for CE separations. Table 10.20.2 lists a number of buffers and their pH ranges. The presence of high salt concentrations in the sample can interfere with the separation process, and dialysis or dilution may be required. Protein concentrations of microgram per microliter are optimal for CE separations; however, on-capillary concentration protocols exist that allow the separation of proteins with concentrations 1 to 2 orders of magnitude lower.

Peptides can be effectively separated in open-tube fused-silica capillary columns. The EOF generated within the capillary causes separation of both charged and neutral peptides. The respective migration times are dependent upon both the pH of the electrophoretic separation and the charge-to-mass ratios of the peptides. For example, at low pH, peptides with net positive charge migrate towards the anode faster than neutral or negatively charged peptides. Even negatively charged peptides are swept towards the anode by EOF.

Table 10.20.2 Useful Electrolytes for CE of Peptides

Electrolyte	Useful pH range	Minimum useful wavelength (nm)
Phosphate	1.14-3.14	195
Citrate	3.06-5.40	260
Acetate	3.76-5.76	220
MES ^a	5.15-7.15	230
Phosphate	6.20-8.20	195
Tris	7.30-9.30	220
Borate	8.14-10.14	180

^aAbbreviation: MES, 2-(*N*-morpholino)ethanesulfonic acid.

In practice this migration order is further modified by the mass of the peptides, with small positively charged peptides having faster migration rates.

The choice of buffer for CE is of primary importance. The use of buffer systems that separate at high pH significantly alters the migration positions of peptides relative to those of low-pH separations. The selection of a specific buffer is dictated both by its buffering capacity at a selected pH and by its minimum usable UV absorbance wavelength. Characteristics of some useful buffer systems are shown in Table 10.20.2. Small differences in the pH or composition of the buffer can have a significant impact on the absolute mobilities of the peptides, so it is advisable to adjust the pH of buffers accurately and reproducibly. Alternatively, a suitable internal standard can be incorporated into the separation in order to obtain relative mobilities. The protein can be solubilized in water or 40 mM buffer in the presence or absence of urea. The sample buffer does not need to be the same composition as the running buffer, however, use of ionic detergents is not recommended. For the separation of peptides from a tryptic digest of a protein, where almost every peptide carries at least two positive charges, a phosphate buffer of pH 2.0 to 3.0 is commonly used. Likewise, for V8 protease digestions (which cleaves C-terminal to glutamic acid residues), the preferred separation conditions use a borate buffer with a pH range of 8.0 to 9.0.

SEPARATION OF PROTEINS BY ISOELECTRIC FOCUSING

Isoelectric focusing (IEF) separations can be performed readily by CE; this can be a useful first step in selecting subsequent buffer systems for protein separations. Proteins and peptides are amphoteric in nature, so their charge is dictated by the surrounding carrier buffer. When a protein or peptide is placed in an electric field, it moves to a region where the surrounding pH equals its isoelectric point (pI); at this point the molecule stops moving. A pH gradient is generated in a coated capillary column by filling the column with a sample solution that contains ampholytes. A high-pH solution (sodium hydroxide) is placed in the cathode reservoir and a low-pH solution (phosphoric acid) is placed in the anode reservoir. An electric field is then applied across the column, and the ampholytes and the sample move to a location in the column corresponding to their respective pIs. If the molecule drifts out of the isoelectric region, a charge is induced by the surrounding pH solution, and the molecule moves back to the region of zero charge. Narrow regions are formed with pI resolutions on the order of 0.01 pH units. To determine the pI of a particular protein in the column, markers of known pIs are added to the sample mixture, and extrapolation between markers yields the pI of the protein of interest.

After separation of the sample in the absence of EOF, mobilization of the proteins past the detector is necessary. A simple replacement of the anode reservoir with the buffer in

BASIC PROTOCOL 1

Analysis of Proteins

10.20.5

the cathode reservoir establishes EOF and causes the sample zone to migrate past the detector. The separation discussed here assumes there is no EOF in the column during focusing. However, separations may also be done in the presence of EOF (Mazzeo and Krull, 1992).

Materials

Ampholyte mixture, pH 3 to 10 (Bio-Rad)
Sample containing 0.5 to 1.0 mg protein/ml in water
IEF markers (Bio-Rad; optional)
20 mM sodium hydroxide (NaOH; store at 4°C)
10 mM phosphoric acid (store at 4°C)
50- μ m-i.d. coated silica capillary column
CE instrument (see Table 10.20.1)

1. Add ampholyte mixture to 0.5 ml protein sample to give a final concentration of 2.5% (v/v) ampholytes.

If desired, add IEF markers to a final concentration of 0.1 mg/ml to calibrate the column.

2. Add the solution to the sample reservoir. Fill the capillary by pressurizing the reservoir (0.5 lb/in²).
3. Place 10 mM phosphoric acid in the anode reservoir (negative end). Place 20 mM NaOH in the cathode reservoir (positive end).
4. Focus the sample 4 to 6 min at 8 to 10 kV constant voltage. Monitor the current until it reaches steady state.

The instrument should be operated in accordance with the manufacturer's instructions.

5. Mobilize the sample by placing 20 mM NaOH in the anode reservoir and setting the voltage to 10 kV. Monitor the mobilization of the proteins past the detector 15 to 20 min.

Another method for mobilizing the proteins past the detector is to apply a hydrodynamic force to one end of the column and push the zones past the detector. Commercially available CE instruments have this capacity as an option. To minimize distortion of the pI regions, a constant electric field (10 kV) must be applied at the same time. The pressure applied is small (0.5 lb/in²) and is maintained for the entire time of mobilization.

6. Wash the column after each run with 10 mM phosphoric acid for 1 min at 0.5 lb/in². Store the column at room temperature in running buffer (short-term) or water (long-term).

BASIC PROTOCOL 2

SEPARATION OF PROTEINS

To use CE for protein analyses, the more that is known about the particular protein, the higher the success rate for developing a separation method. Knowing the pI of a protein permits selection of the appropriate run buffer to maximize differences in the charge-to-size ratio. IEF separation performed using CE (Basic Protocol 1) can be a useful first step in selecting subsequent buffer systems for protein separations on an underivatized capillary column. If no protein is detected, it has probably adsorbed to the silica surface of the column. Repeating the separation in the presence of an ionic detergent such as SDS will coat the protein with negative charge and prevent adsorption. However, this means that proteins will separate on the basis of size alone. Alternatively, if the pI of the protein is unknown, the separation can often be achieved by using a coated capillary column in the presence of high-pH, high-ionic-strength buffer.

Materials

- Sample containing 10 mg protein/ml in water
- 50 mM and 500 mM sodium borate, pH 8.0 to 9.5 (for separation of proteins with unknown pI)
- 5 mM and 50 mM buffer at pH >pI (for separation of proteins with known pI; see Table 10.20.2)
- 50- μ m-i.d. coated (unknown pI) or uncoated (known pI) fused-silica capillary column
- CE instrument (see Table 10.20.1)

For separations of proteins with unknown pI

- 1a. Dilute protein sample 1/10 (v/v) with 50 mM sodium borate buffer to give a final concentration of 1.0 mg/ml.

Alternatively, dialyze sample (at a concentration of 1 mg/ml) in 50 mM sodium borate buffer.

- 2a. Fill a coated column with 500 mM sodium borate buffer.
- 3a. Inject the sample hydrodynamically, e.g., 3 sec at 1 lb/in².

Consult the manufacturer's instruction manual for the proper procedure.

- 4a. Separate using the following conditions:

- Detector wavelength: 200 nm
- Temperature: 25°C
- Run voltage: 10 kV
- Run time: 30 min.

For separations of proteins with known pI

- 1b. Dilute protein sample 1/10 (v/v) with a 5 mM buffer that has a pH above the pI of the protein to induce a charge on the protein (final protein concentration = 1.0 mg/ml).

Alternatively, dialyze sample (at a concentration of 1 mg/ml) in 5 mM buffer.

- 2b. Fill an uncoated capillary column with the same buffer at 50 mM.
- 3b. Inject the sample hydrodynamically, e.g., 3 sec at 1 lb/in².

Consult the manufacturer's instruction manual for the proper procedure.

- 4b. Separate using the following conditions:

- Detector wavelength: 200 nm
- Temperature: 25°C
- Run voltage: 25 kV
- Run time: 30 min.

If the protein of interest is not detected, add 10 mM SDS to the run buffer to induce a charge on the molecule. If the peaks are too broad, add 0.1% (w/v) methylcellulose to coat the column and reduce adsorption of the molecule onto the wall of the column. The addition of SDS and/or methylcellulose will increase the run time of the separation.

ANALYTICAL PEPTIDE SEPARATIONS

CE is particularly useful for separating specific peptides from complex mixtures such as the products of proteolytic digestion. In this procedure, the peptide mixture to be analyzed is placed in a sample vial at the cathode end of the capillary column. The anode reservoir contains running buffer. The sample is loaded into the capillary either electrokinetically or hydrodynamically by pressure or vacuum, then the sample vial is replaced by a vial containing running buffer and electrophoresis begins. The following protocol describes the use of a P/ACE 5000 CE instrument (Beckman) using low-pressure sample injection. However, other instruments (see Table 10.20.1) are capable of similar separations when operated according to the manufacturer's instructions.

Materials

Peptide mixture: e.g., tryptic digest of β -lactoglobulin (Applied Biosystems)
0.05 M and 0.25 M sodium phosphate buffer, pH 2.30 (store at 4°C)
0.1 M sodium hydroxide (NaOH)
75- μ m-i.d. fused-silica capillary column (Beckman)
CE instrument (e.g., P/ACE 5000, Beckman, or equivalent; see Table 10.20.1)

1. Precondition the capillary by flushing with the following solutions:

10 column volumes of 0.1 M NaOH at low pressure (0.5 lb/in²)
10 column volumes water
4 column volumes of 0.25 M sodium phosphate buffer, pH 2.30.

Store the column in 0.25 M sodium phosphate buffer, pH 2.30, at 25°C.

When changing to a different separation buffer, equilibrate the column 4 hr with the new buffer (Strickland and Strickland, 1990).

If a new capillary is being used, it is essential that this preconditioning step be performed prior to attempting separations.

2. Prepare the peptide mixture by dissolving 10 nmol (~300 μ g) in 10 ml of 0.05 M sodium phosphate buffer, pH 2.30. Freeze unused mixture in 100- μ l aliquots.
3. Apply 10 to 20 nl sample using low-pressure injection for 10 sec at 0.5 lb/in².

Consult the manufacturer's manual for the proper procedure.

4. Separate the peptide mixture using the following conditions:

Electrolyte: 0.05 M sodium phosphate buffer, pH 2.3
Detector wavelength: 200 nm
Temperature: 25°C
Voltage: 25 kV.

Consult the manufacturer's manual for proper operating conditions.

5. Rinse the column with the following solutions:

0.5 min with water
1.0 min with 0.1 N NaOH
1.5 min with water
1 min with 0.25 M sodium phosphate buffer, pH 2.30.

Store the column in running buffer or water at room temperature.

An example of a separation protocol for a Beckman P/ACE 5000 instrument with a 37-min run time is described in Table 10.20.3.

Table 10.20.3 Operating Program for Peptide Analysis^a

Step	Process	Duration	Inlet ^b	Outlet ^b	Control summary ^c
1	Set Temp	—	—	—	Temp 25°C
2	Set Detector	—	—	—	UV 200nm
3	Rinse	1.0 min	33	7	Forward: HP
4	Rinse	1.5 min	29	7	Forward: HP
5	Rinse	0.5 min	29	9	Forward: HP
6	Inject	10.0 sec	11	9	Low pressure
7	Separate	30.0 min	29	9	Voltage 25.00 kV
8	Rinse	0.5 min	33	7	Forward: HP
9	Rinse	1.0 min	32	7	Forward: HP
10	Rinse	1.5 min	33	7	Forward: HP
11	Rinse	1.0 min	34	9	Forward: HP

^aFor analysis of proteolytic digests, HPLC fractions, and synthetic peptides.

^bVial contents: 7, waste vial (water level just to contact capillary effluent); 9, electrolyte (50 mM sodium phosphate buffer, pH 2.30); 11, sample; 29, electrolyte (50 mM sodium phosphate buffer, pH 2.30); 32, regeneration solution (0.1 N NaOH); 33, water; and 34, 0.25 M phosphate buffer, pH 2.30.

^cAbbreviation: HP, high pressure.

MICROPREPARATIVE CAPILLARY ELECTROPHORESIS: MULTIPLE SEPARATIONS

BASIC PROTOCOL 4

Although CE has been most commonly used for analytical separations, considerable interest has developed in using its high resolving power in micropreparative applications. Two basic approaches have evolved, one utilizing multiple separations and collections from an analytical capillary (Bergman and Jörnval, 1992) and the other utilizing a single separation and collection on a much larger-diameter (150- μm -i.d.) capillary (see Alternate Protocol). For the multiple collection approach to be effective, the elution times of the peptides must be reproducible. The following method uses the P/ACE 5000 capillary electrophoresis instrument to separate a mixture of peptides.

Materials

Peptides: ACTH 4-10, angiotensin I, and angiotensin II (Sigma)
 0.05 mM and 0.25 mM sodium phosphate buffer, pH 2.30 (store at 4°C)
 0.1 M sodium hydroxide (NaOH)
 75- μm -i.d. fused-silica capillary column (Beckman)
 CE instrument (e.g., P/ACE 5000, Beckman, or equivalent; see Table 10.20.1)
 Conical microvials (Beckman)

1. Precondition the capillary column by flushing with the following solutions at low pressure (0.5 lb/in²):

10 column volumes of 0.1 M NaOH
 10 column volumes of water
 4 column volumes of 0.25 M sodium phosphate buffer, pH 2.30.

Store in 0.25 M sodium phosphate, pH 2.30.

When changing to a different separation buffer, equilibrate the column 4 hr with the new buffer (Strickland and Strickland, 1990).

2. Prepare a peptide mixture by dissolving 0.2 mg each of ACTH 4-10, angiotensin I, and angiotensin II in 1 ml of 0.05 M sodium phosphate buffer, pH 2.30 (final concentration = 0.6 mg peptide/ml). Store 100- μl aliquots at -20°C.
3. Load 10 to 20 nl peptide mixture by low-pressure injection for 10 sec at 0.5 lb/in².

Analysis of Proteins

10.20.9

Consult the manufacturer's manual for the proper procedure.

4. Separate the mixture using the following conditions:

Electrolyte: 0.05 M sodium phosphate, pH 2.30

Detector wavelength: 200 nm

Temperature: 25°C

Run voltage: 25 kV

Fraction size: 3-min per collection vial.

Consult the manufacturer's manual for proper operating conditions.

5. Replace the standard outlet reservoir (anode) with a series of conical microvials, each containing 10 μ l of 0.05 M sodium phosphate buffer, pH 2.30. Collect 3-min fractions into each of the vials for the length of the separation.
6. Repeat the injection and separation (steps 3 to 5) four times and combine the contents from the corresponding fraction numbers.

It will take 2 to 3 hr to complete the fractionation.

7. Screen fractions for peptide content using the analytical separation previously described (see Basic Protocol 3).

ALTERNATE PROTOCOL

MICROPREPARATIVE CAPILLARY ELECTROPHORESIS: SINGLE SEPARATION

In this micropreparative protocol, a larger quantity of the peptide mixture is loaded onto a 150- μ m-i.d. capillary column and the fractions are collected from a single separation (Kenny et al., 1993). For the single-collection approach the CE instrument must be able to effectively control the capillary temperature at the elevated power levels that are required for electrophoresis. Forced air cooling is generally inadequate for this purpose.

Additional Materials (also see Basic Protocol 4)

4:1 (v/v) 0.5 M sodium phosphate buffer (pH 2.50)/ethylene glycol

150- μ m-i.d. fused-silica capillary column (Polymicro)

1. Prepare peptide mixture and precondition the column (see Basic Protocol 4, steps 1 and 2).
2. Load 0.1 μ l peptide mixture by low-pressure injection for 10 sec at 0.5 lb/in².

Consult the manufacturer's manual for the proper procedure.

3. Separate the peptide mixture using the following conditions:

Electrolyte: 0.05 M sodium phosphate buffer, pH 2.30

Detector wavelength: 200 nm

Temperature: 25°C

Run voltage: 7.5 kV

Fraction size: 3-min per collection vial.

Consult the manufacturer's manual for proper operating conditions.

4. Replace the standard outlet reservoir (anode) by a series of conical microvials each containing 10 μ l of 4:1 (v/v) 0.5 M sodium phosphate buffer/ethylene glycol. Collect 3-min fractions into each vial for the length of the separation.

It will take ~2 hr to fractionate the sample.

5. Screen fractions for peptide content using the analytical separation protocol (see Basic Protocol 3).

COMMENTARY

Background Information

Protein separation is one of the more difficult types of separation to perform by capillary electrophoresis (CE; Novotny et al., 1990). Each protein has its own set of optimal separation conditions, and one optimized protocol will not necessarily transfer well to the separation of a different protein. The more that is known about the properties of the molecule, the easier it will be to optimize the separation conditions. The most important property is probably the pI, because it will dictate the selection of pH and ionic strength conditions for the initial separation attempt. However, a variety of other properties must also be taken into account when optimizing a separation protocol: shape (globular or fibrous), aggregation tendencies, solubility in dilute solutions, salt requirements, thermal stability, and hydrophobic nature. Several parameters can be varied within CE separations in order to utilize this knowledge—e.g., temperature, buffer/salt selection, ionic and nonionic detergents, and matrix additives (polyacrylamide, methylcellulose). Several protocols have been developed that can serve as useful starting points for most protein separations. Table 10.20.4 lists references for some groups of proteins. Furthermore, many of the manufacturers listed in Table 10.20.1 have developed additional protocols for protein separations.

Ultimately, the selection of an appropriate separation protocol for a protein depends on the specific properties of that protein. However, there are separation approaches that utilize specific properties of proteins. IEF-CE separates on the basis of charge alone; CE in the presence of ionic detergents such as SDS separates on the basis of size alone; and CE using underivatized capillaries at low pH or derivatized capillaries at high salt and high pH separates on the basis of both size and charge. Successful CE separations also depend on the behavior of

other contaminants that are present in the mixture. It may also be possible to take advantage of the knowledge gained from previous purification steps, e.g., ion-exchange or size-exclusion chromatography, to select CE separation conditions.

Analytical-scale separation of peptides by CE has found application in a variety of areas. Proteolytic digestions can be monitored by placing the digestion mixture in a sample vial on the sample table that is equilibrated to 37°C. The instrument can then be programmed to analyze an aliquot at desired intervals, such as every 4 hr. Digestion is complete when a stable CE profile is obtained. Also, fractions collected from an HPLC separation of a protease digest can readily be screened orthogonally by CE to determine the homogeneity of any given peak. Fractions with a single major component can then be subjected to protein sequence analysis.

The sensitivity of UV detection at 200 nm is such that digest levels of ≥ 100 pmol can be effectively screened by this method. Below these load levels the fractions have to be reduced in volume or concentrated on the capillary to achieve a sufficiently high concentration for visualization with UV detection (Dolnik et al., 1990; Burgi and Chien, 1992).

Selection of an appropriate micropreparative protocol will depend on the type of instrumentation that is available. The single-separation micropreparative protocol will only work with a CE instrument that has adequate cooling capabilities. The multiple-separation micropreparative protocol will work on any CE instrument that has reproducible and stable electrophoretic mobilities.

Critical Parameters

For protein analysis, the charge on the protein is the most important parameter for separation. However, the charge-to-shape ratio also plays a role in effective separation of proteins.

Table 10.20.4 References for Separations of Different Protein Types

Protein type	Reference
Basic proteins	Wiktorowicz and Colburn (1990); Bullock and Yuan (1991)
Glycoproteins	Tran et al. (1991); Tsuiji and Little (1992)
Antibody-antigen complexes	Nielsen et al. (1991)
Milk proteins	Chen et al. (1992)

The charge on the protein is controlled by the buffer and the additives used for the run. The shape is determined by the denaturing conditions of the solutions. The surface charge on a protein may ultimately dictate the extent of wall interactions in addition to electrophoretic mobility, whereas shape has its greatest effect upon mobility alone.

The recovery of peptides from underivatized silica capillary columns appears to be very good—even at subpicomole levels. However, proteins are frequently irreversibly bound to these same capillaries. Clearly, at some point a “peptide” must become a “protein” in the context of CE. There is little published data available to clarify this issue, although good recoveries of peptides with 30–40 amino acid residues in length have been obtained from underivatized capillary columns. If the recoveries of higher molecular weight peptides are poor, then the separation should be repeated using a coated capillary column. During peptide synthesis, single-amino-acid-deletion products are frequently produced. The properties of these deletion products are very similar to those of the expected species, so the deletion products may not be readily resolved by reversed-phase HPLC. In this case, CE has proven to be a very useful orthogonal technique because of its high resolution power and separation based on charge-to-mass ratios. However, it must be emphasized that the effective separation and detection of proteins and peptides requires relatively high solute starting concentrations.

When using uncoated columns, a sodium hydroxide rinse is required to maintain the capillary surface. However, this rinse is never used with coated capillaries because it will remove the coating. Commercially available coated columns are satisfactory, but linear polymers (e.g., methylcellulose) have been used with more success (Palmieri and Nolan, 1994). If a protocol can be worked out for the molecule of interest, rapid CE analysis with little loss of starting material can augment the tools used now for protein analysis (e.g., HPLC).

Troubleshooting

Mechanical and electrical problems that might be encountered during instrument use are addressed in the troubleshooting sections of the manufacturers’ manuals. One of the common difficulties is the loss of electrical contact during a run. This loss of current can be caused if a small bubble forms in the column during

injection or if the power generated by the run is so great that the solution boils or outgasses. Purging the column after a failed run removes any bubbles in the column. The sample vial must contain enough sample to cover the end of the column during injection to prevent injection of air into the column. The sample vial can dry out during a long run, so sample temperature control or the use of the proper cap on the vial is required to slow evaporation. Reducing the buffer concentration or the run voltage can eliminate bubble formation during a run. Degassing the run buffer is also useful if the outgassing problems continue.

The instrument sometimes arcs during a run; this is caused by salt deposits or moisture on the high-voltage electrode stations. Particular attention should be given to keeping the electrode stations salt free. Moisture buildup can be minimized by reducing the separation voltage by 5 kV. Frequent visual inspection is also advised.

The column can plug up because of salt deposition if the column dries out. Washing the column with water dissolves the plug; storing the column in water, if it is not going to be used for a long period of time, prevents plugging. Proteins can precipitate during a run and plug the column. Using lower protein concentrations prevents this type of plugging.

Anticipated Results

Electrophoretic profiles for both protein and peptide separations should be highly reproducible. However, absolute migration times may vary from run to run, especially if the separation temperature is not adequately controlled. For this reason, the use of an appropriate internal standard is highly recommended (see the manufacturer’s instructions).

At load levels <50 pmol of digest, the recovery of peptides from a narrow-bore (1-mm) HPLC separation is poor. A convenient upper separation range for micropreparative CE is ~50 pmol of digest. The use of CE as a preparative separation technique therefore complements conventional HPLC capabilities. Recoveries from micropreparative CE separations are essentially quantitative down to the low picomole level, although postseparation handling losses can be significant. To use the multiple-separation micropreparative approach (see Basic Protocol 4), the CE instrument must be capable of maintaining very stable migration times. Not all instruments have this capability, so it is advisable that the method be used for simple mixtures of peptides. The single-sepa-

ration micropreparative approach is recommended for more complex mixtures. Because the fractions are collected "blind," they should be screened by matrix-assisted laser desorption/ionization (MALDI; Henzel and Stults, 1995) mass spectrometry prior to sequencing.

Time Considerations

Complete analysis times for IEF are ~30 min. Analysis times for protein separations are ~30 min for each step in method development. Peptide separations are normally completed in 40 min and frequently in as little as 20 min. Micropreparative peptide separations using either the multiple or single-separation approach are normally achievable within 90 min, but may take 2 hr or longer. It should be noted that these times are approximate, and sufficient time should be allowed to electrophorese all the components completely from the capillary.

Literature Cited

- Bergman, T. and Jörnvall, H. 1992. Capillary electrophoresis for preparation of peptides and direct determination of amino acids. *In Techniques in Protein Chemistry III* (R.A. Hogue Angeletti, ed.) pp. 129-134. Academic Press, San Diego.
- Bullock, J.A. and Yuan, L.-C. 1991. FSCE of basic proteins in uncoated fused-silica capillary tubing. *J. Microcol. Sep.* 3:241-248.
- Burgi, D.S. and Chien, R.-L. 1992. On-column sample concentration using field-amplification in CZE. *Anal. Chem.* 64:489A-495A.
- Chen, F.A., Kelly, L., Palmieri, R., Biehler, R., and Schwartz, H. 1992. Use of high ionic strength buffers for the separation of proteins and peptides with CE. *J. Liq. Chromatogr.* 15:1143-1161.
- Dolnik, V., Cobb, K.A., and Novotny, M. 1990. Capillary zone electrophoresis of dilute samples with isotachophoretic preconcentration. *J. Microcol. Sep.* 2:127-140.
- Henzel, W.J. and Stults, J.T. 1995. Reversed-phase isolation of peptides. *In Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 11.6.1-11.6.14. John Wiley & Sons, New York.
- Kenny, J.W., Ohms, J.I., and Smith, A.J. 1993. Micropreparative capillary electrophoresis (MPCE) and micropreparative HPLC of protein digests. *In Techniques in Protein Chemistry, IV* (R.A. Hogue Angeletti, ed.) pp. 363-370. Academic Press, San Diego.
- Khaledi, M.G. 1994. Micellar electrokinetic capillary chromatography. *In Handbook of Capillary Electrophoresis* (J.P. Landers, ed.) pp. 43-94. CRC Press, Boca Raton, Fla.

- Landers, J.P., Madden, B.J., Oda, R.P., and Spelsberg, T.C. 1992. The use of modifiers of endosmotic flow for analysis of microheterogeneity. *Anal. Biochem.* 205:115-124.
- Mazzeo, J.R. and Krull, I.S. 1992. Improvements in the methods developed for performing isoelectric focusing in uncoated capillaries. *J. Chromatogr.* 606:291-296.
- Nielsen, R.G., Richard, E.C., Santa, P.F., Sharknas, D.A., and Sittampalam, G.S. 1991. Separation of antibody-antigen complexes by CZE, IEF, and HPSEC. *J. Chromatogr.* 539:177-185.
- Novotny, M.V., Cobb, K.A., and Liu, J. 1990. Recent advances in CE of proteins, peptides, and amino acids. *Electrophoresis* 11:735-749.
- Palmieri, R. and Nolan, J.A. 1994. Protein capillary electrophoresis: Theoretical and experimental considerations for methods development. *In Handbook of Capillary Electrophoresis* (J.P. Landers, ed.) pp. 325-368. CRC Press, Boca Raton, Fla.
- Strickland, M. and Strickland, N. 1990. FSCE using phosphate buffer and acidic pH. *Am. Lab.* 22:60-65.
- Tran, A.D., Parks, S., Lisi, P.J., Huynh, O.T., Ryall, R.R., and Lane, P.A. 1991. Separations of carbohydrate-mediated microheterogeneity of human erythropoietin by free solution CE. *J. Chromatogr.* 542:459-471.
- Tsujii, K. and Little, R.J. 1992. Charge-reversed, polymer-coated capillary column for the analysis of a recombinant chimeric glycoprotein. *J. Chromatogr.* 594:317-324.
- Wiktorowicz, J.E. and Colburn, J.C. 1990. Separation of cationic proteins via charge reversal in CE. *Electrophoresis* 11:769-773.

Key References

- McCormick, R.M. 1994. Capillary zone electrophoresis of peptides. *In Handbook of Capillary Electrophoresis* (J.P. Landers, ed.) pp. 287-324. CRC Press, Boca Raton, Fla.

Excellent review of both theory and practice for separation of peptides by CE.

Palmieri and Nolan, 1994. See above.

Good review of methods development for protein separations.

Contributed by Dean Burgi
Genomyx
Foster City, California

Alan J. Smith
Stanford University Medical Center
Stanford, California

Overview of Peptide and Protein Analysis by Mass Spectrometry

UNIT 10.21

WHY IS MS AN ESSENTIAL TOOL IN PROTEIN STRUCTURE ANALYSIS?

Mass spectrometry (MS) is now regarded as an indispensable tool for peptide and protein primary structure analysis. The increasingly widespread use of this technique is due in large part to its ability to solve structural problems not easily handled by conventional protein chemistry techniques. MS can provide accurate molecular weights for peptides and proteins with masses up to 500,000 Da using only a few picomoles (tens of femtomoles in favorable cases) of material. The accuracy achieved by MS is frequently better than 0.01% of the calculated mass (see Table 10.21.1). This means, for example, that the experimentally determined M_r of a 25-kDa protein will be ± 2.5 Da of the actual mass. In contrast, molecular weight estimates obtained by SDS-PAGE have accuracies of 5% to 10%; moreover, the mobility of the protein in the gel can be grossly affected (up to 50%) by the presence of covalent modifications such as lipids and carbohydrates. The accuracy of protein molecular weight estimates obtained by other methods, such as sedimentation velocity and gel permeation, can also be affected by the presence of lipid or carbohydrate. For correction and interpretation of this data, a molecular weight derived by mass spectrometry is needed (Hensley et al., 1994).

Unlike these other techniques, mass spectrometry provides molecular mass information independent of any modifications to the structure (such as glycosylation and phosphorylation). A mass difference between the observed and the expected mass of the protein or a peptide suggests the presence of post-translational modification or proteolytic processing, and the mass difference will be the starting point for defining the structure of the modification. A difference of +42 Da, for example, together with the lack of a free NH_2 -terminus, is indicative of the presence of an α -*N*-acetyl group. Similarly, a difference of +80 Da suggests the presence of a phosphate or sulfate moiety, and a cluster of molecular ion peaks separated by 162, 203, 291, or 146 Da indicates glycoforms differing by the presence of hexose, *N*-acetylhexosamine, *N*-acetylneuraminic acid, or deoxyhexose (see Table A.1C.4 in APPENDIX 1C). Thus molecular weight information provides a

useful measure of the integrity, purity, and overall state of modification of a peptide or protein. When evaluated in the context of other available structural data, the molecular mass information alone can be used to solve a variety of peptide and protein structural problems.

The ability of MS to determine the molecular weights of peptides in mixtures (such as result from a proteolytic digest of a protein) with better than 0.01% accuracy forms the basis of the mass spectrometric peptide mapping strategy, a rapid and highly efficient method to verify the fidelity of translation of protein sequences deduced from DNA or cDNA sequencing. This can be done either by direct analysis of unfractionated mixtures (Billeci and Stults, 1993) or on-line with liquid chromatography (Hess et al., 1993; Carr et al., 1993; Kassel et al., 1994) or capillary electrophoresis (Wahl et al., 1992; Foret et al., 1994; Smith et al., 1996).

Mass spectrometry also has the ability to provide amino acid sequence information on peptides using methodology generically referred to as tandem mass spectrometry (MS/MS). Using MS/MS, complete or partial sequence information may be obtained at the femtomole to picomole level for peptides containing up to 25 amino acid residues. In favorable cases some internal sequence data can often be obtained for peptides up to 30 or 40 residues. Unlike Edman sequencing, MS can provide this type of sequence information even if the peptides are present in complex mixtures, and, more importantly, even if they are modified. Methodologies such as Edman degradation and amino acid analysis have serious limitations with regard to characterizing structurally modified proteins and peptides, because many commonly occurring modifications are lost or destroyed by the harsh cleavage and derivatization conditions employed (Uy and Wold, 1977; Wold, 1981; Wold and Moldave, 1984; Carr et al., 1991). Furthermore, the sensitivity and speed of MS-based sequencing now exceeds that of Edman-based sequence analysis (Hunt et al., 1992; Wilm et al., 1996).

Vast amounts of cDNA sequence data, obtained by grand-scale sequencing of the human and other genomes, are presently flooding into databases. However, little is known about the functions of most of the proteins the various genes encode, nor is much known about the

Contributed by Steven A. Carr and Roland S. Annan

Current Protocols in Molecular Biology (1997) 10.21.1-10.21.27

Copyright © 1997 by John Wiley & Sons, Inc.

Analysis of
Proteins

10.21.1

Supplement 38

Table 10.21.1 Key Ionization Methods for Biomolecule Analysis

Ionization method	Tolerance for buffers, salts, and detergents	MW range	Mass assignment accuracy	Compatible with on-line LC/CE-MS
Electrospray ionization (ESI)	Low tolerance (≤ 20 mM) for nonvolatile buffers, alkali metal salts, detergents; on-line LC-MS used for contaminant removal	≤ 150 kDA	$\pm 0.005\%$ – 0.01% to 50 kDA $\pm 0.02\%$ – 0.03% to 150 kDA	Yes
Matrix-assisted laser desorption/ionization (MALDI)	High tolerance (≥ 50 mM) for alkali metal salts, phosphate, urea, etc.; tolerant of some nonionic detergents; intolerant of SDS	≥ 350 kDA	$\pm 0.01\%$ – 0.05% to 25 kDA $\pm 0.05\%$ – 0.3% to 300 kDA	No

processed states of the expressed, mature forms of the proteins. In contrast, the protein content of a cell, referred to as the proteome (Kahn, 1995; Wilkins et al., 1995), can be studied by protein differential display using two-dimensional gel electrophoresis (UNIT 10.4), immunoprecipitation, and related procedures, allowing physiologically relevant proteins to be recognized by changes in their expression levels in response to stimuli. Because of its speed and sensitivity, mass spectrometry, through the use of strategies that employ molecular weight and/or partial sequence information to search protein databases, has emerged as the primary technology for identifying these proteins (Henzel et al., 1993; Mann et al., 1993; Pappin et al., 1993; Wilm and Mann, 1994; Eng et al., 1994).

This introductory overview has been designed with several goals in mind. The first and perhaps the most important one is to try and dispel the still widely held, but now mistaken, belief that mass spectrometry is beyond the technical resources of the typical laboratory or facility engaged in the structural characterization or synthesis of peptides and proteins. Over the last five years, mass spectrometers and associated data systems have become available that can be operated by anyone capable of running a modern amino acid analyzer or Edman sequencer, yet are powerful enough to handle the most demanding analyses that the peptide or protein investigator might require. As a result, it is no longer necessary to be a “card-carrying” mass spectroscopist to use the instruments and techniques of MS effectively and to apply it successfully to one’s own research program. The remaining goals of this article are to familiarize peptide and protein chemists with the types of mass spectrometers that are appropriate for the majority of their

analytical needs, to describe the kinds of experiments that can be performed with these instruments on a routine basis, and to discuss the kinds of information that these experiments provide. The emphasis here is on established tools and techniques that can realistically be used for problem solving. As a result, many useful instruments and techniques employed by the trained mass spectroscopist are not discussed here, because they do not satisfy these criteria or are not yet widely available. A few of these are briefly mentioned at the end of this introductory overview. The units that follow this overview will further discuss many of the principles introduced here and provide detailed protocols and applications of the methods. The overview concludes with a tutorial discussion on the meaning and practical importance of a number of fundamental experimental and performance-related issues that have important implications for interpretation and use of mass spectral data, such as mass accuracy and resolution.

More comprehensive discussion of mass spectrometry of biological molecules may be found in Chapter 16 of *Current Protocols in Protein Science* (Coligan et al., 1997). Additional information is available in several recently published books (Burlingame and McCloskey, 1990; McCloskey, 1990; Burlingame and Carr, 1996) and review articles (Carr et al., 1991; Biemann, 1992; Chait and Kent, 1992; Aebersold, 1993; Wang and Chait, 1994; Mann and Wilm, 1995; Williams and Carr, 1995; Mann and Talbo, 1996). The biannual review of mass spectrometry in *Analytical Chemistry* by Burlingame and colleagues provides a particularly good annotated bibliography of the recent literature on techniques and application of MS in the biological sciences

(Burlingame et al., 1994). Mass spectrometry information is also rapidly proliferating on the World Wide Web.

WHAT IS MS?

Mass spectrometry is a powerful analytical technique for forming gas-phase ions from intact, neutral molecules and subsequently determining their molecular masses. All mass spectrometers have three essential components: an ion source, a mass analyzer, and a detector. Ions are produced from the sample in the ion source using a specific ionization method (see discussion of Key Ionization Methods and Related Ancillary Techniques). The ions are separated in the mass analyzer based on their mass-to-charge (m/z) ratios and then detected, usually by an electron multiplier. The data system produces a mass spectrum, which is a plot of ion abundance versus m/z . Except in the case of electrospray mass spectrometry (where the ion source is at atmospheric pressure), the ion source, mass analyzer, and detector are usually situated inside a high-vacuum chamber (pressure between 10^{-8} and 10^{-4} Torr).

KEY IONIZATION METHODS AND RELATED ANCILLARY TECHNIQUES

The decade of the 1980s marked a revolutionary period in the development of "soft ionization" techniques for the analysis by mass spectrometry of large, polar, nonvolatile molecules (Carr et al., 1991). These ionization techniques produced primarily intact protonated molecular ions of peptides and proteins. In the early 1980s, fast atom bombardment (FAB) and, to a lesser extent, plasma desorption (PD) mass spectrometry were used extensively for structural characterization of peptides and small proteins. In 1988, electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) mass spectrometry were first shown to be useful for the analysis of peptides, proteins, carbohydrates, and oligonucleotides. Within a few years of their introduction, ESI and MALDI techniques became the MS methods of choice for biopolymers analysis, largely supplanting FAB-MS and PD-MS. For this reason these techniques will be the focus of this overview. The capabilities and sample requirements of these two MS techniques are compared in Table 10.21.1. The performance specifications in the table are those that can be routinely achieved by trained operators and do not represent the ultimate performance of each technique.

Several factors account for the widespread acceptance of ESI-MS and MALDI-MS. Both ionization methods (see detailed descriptions below) are optimally combined with relatively low-cost, simple-to-operate mass analyzers. A quadrupole mass filter is typically used for ESI-MS, whereas a time-of-flight (TOF) analyzer is commonly employed for MALDI-MS. The sensitivity of both methods is in the low-to subpicomole range, and both are capable of analyzing molecules with a wide range of molecular weights (Table 10.21.1). The methods have many analytical capabilities in common, but do differ in a number of respects that are of practical significance. Both techniques work best with clean (salt- and detergent-free) samples, but MALDI is capable of providing information directly on relatively dirty samples. Both methods can provide molecular weight information for large proteins, but the information is invariably easier to obtain and more sensitive with MALDI-MS. In the molecular weight range from 5 to 50 kDa, however, analysis by ESI-MS usually provides much more accurate molecular weight measurements. ESI-MS in this molecular weight range also typically exhibits better mass resolution than MALDI, which permits detection of protein variants that differ only slightly in mass. These practical differences between the techniques, and their impact on the information that can be obtained about a protein structure, are discussed below.

In addition to the development of the ionization approaches described above, a number of key enabling technologies that have significantly improved the performance and sensitivity of ESI-MS and MALDI-MS have been introduced in the last few years. These technologies have moved MS onto the center stage of protein discovery and characterization. The first such advance has been the development and refinement of microscale desalting procedures such as efficient liquid-liquid extraction for peptides produced by in-gel digestion (Shevchenko et al., 1996), micropacked glass capillaries for cleanup of 1 to 5 μ l of sample (Houthaeve, et al., 1995), and microscale high-performance liquid chromatography (HPLC) employing columns with internal diameters ≤ 1 mm (Zhang et al., 1995). Second has been the development of ultra-low-flow (< 50 nl/min) electrospray sources that enable 1 to 2 μ l of a dilute sample containing femtomole per microliter concentrations of peptides to be analyzed for periods of > 60 min (Wilm and Mann, 1996). The long duration of analysis permits

many experiments, including conventional MS experiments and MS/MS, to be carried out on a single sample loading. A third development has been the introduction of novel scanning methods whereby MS can be used to selectively detect specific protein modifications (Carr et al., 1993; Huddleston et al., 1993a,b). Fourth is the development of the so-called post-source decay (PSD) experiment (Kaufmann et al., 1994), which provides the ability to mass select particular peptide molecular ions and to analyze the metastable fragment ions from this precursor that are formed during the MALDI process. Like the ESI tandem MS experiment, PSD can provide a substantial amount of peptide sequence information (see section entitled What Is Tandem MS?). Finally, a variety of computer-based searching strategies have been developed whereby known proteins can be identified from protein and DNA databases by searching MS-derived molecular weights with and without partial sequence information (Henzel et al., 1993; Mann et al., 1993; Pappin et al., 1993; Eng et al., 1994; Wilm and Mann, 1994; Hall et al., 1996).

WHAT IS TANDEM MS?

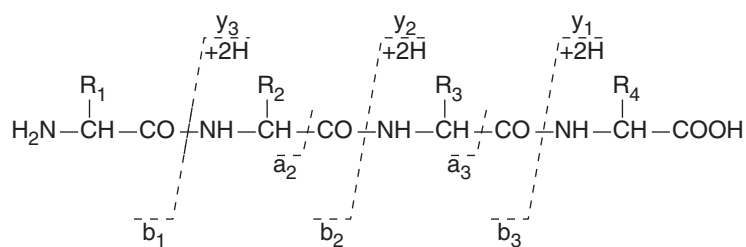
In tandem mass spectrometry (commonly referred to as MS/MS), two consecutive stages of mass analysis are used to detect secondary fragment ions that are formed from a particular precursor ion. The first stage serves to isolate a peptide precursor ion of interest based on its m/z , and the second stage to mass analyze the product ions formed by spontaneous or induced fragmentation of the selected precursor ion. Interpretation of the product-ion spectrum provides sequence information for the peptide selected. The manner in which product ions are produced and detected is different in MALDI and ESI mass spectrometry, and therefore is described in the following sections that are devoted to these specific techniques.

MS/MS provides information that is highly complementary to that provided by Edman degradation; in addition, MS/MS has several advantages relative to Edman degradation for determining peptide sequence. First, MS/MS can produce sequence information for blocked or otherwise modified peptides that are either difficult or impossible to sequence by Edman degradation (for example, see the MS/MS spectrum of the phosphopeptide in Fig. 10.21.6). Second, the ability to individually select and fragment peptides in mixtures eliminates the need for extensive purification of peptides prior to sequence analysis. Third, MS-based peptide

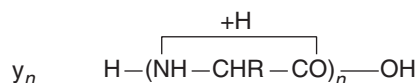
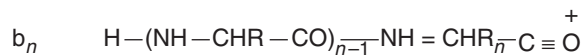
sequencing is faster than Edman sequencing because it is not a stepwise process. All of the structural information, in the form of fragment ions, is presented in a single mass spectrum acquired in a single experiment. Finally, MS-based approaches are now at least as sensitive as Edman degradation (e.g., 1 to 5 pmol of peptide or protein is required), and state-of-the-art MS-based sequencing by either ESI-MS or MALDI-MS is capable of providing sequence information at the 50- to 250-fmol level (Hunt et al., 1992; Wilm et al., 1996).

Peptide fragment ions are produced in the mass spectrometer primarily by cleavage of the amide bonds (—CO—NH—) that join pairs of amino acid residues. The most commonly observed fragment ions and their nomenclature are shown in Figure 10.21.1, and a list of commonly observed low-mass fragment ions indicative of the presence of particular amino acid residues is presented in Table 10.21.2. The fragmentation of peptides in MS has been well described (Hunt et al., 1986; Biemann, 1988, 1990; Falick et al., 1993). The interested reader is referred to these studies and references therein for more detailed discussions of fragmentation and its possible mechanisms. The difference in mass between adjacent sequence ions of the same type defines an amino acid. Of the twenty commonly occurring amino acids, sixteen have unique masses, Leu and Ile have exactly the same mass, and Lys and Gln have nearly the same mass (Table A.1C.3 in APPENDIX 1C). A mass difference between two adjacent sequence ions of the same fragment type that does not correspond to the mass of an amino acid residue may represent the presence of a post-translational modification. A list of common post-translational modifications and their masses is given in Table A.1C.4 in APPENDIX 1C.

Two caveats to the above discussion on MS/MS should be noted. First, although it is generally very easy to obtain the internal sequence of a peptide by this procedure, it is often difficult to determine the order of the first two and sometimes the last few residues in the peptide without additional experiments. Several methods involving reanalysis by MS/MS after chemical derivatization have been developed to help sort out the sequence at the ends of the peptide when required (Sherman et al., 1995; Gulden et al., 1996). Alternatively, there are several sequence ladder approaches, using either chemical derivatives or peptidases, that can be useful for identifying the N- and C-terminal residues (Chait et al., 1993; Bartlett-Jones et al., 1994; Thiede et al., 1995; Patterson et al.,



sequence-specific ions



non-sequence-specific ions

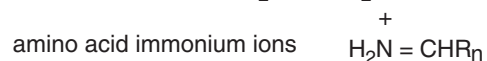


Figure 10.21.1 Nomenclature for peptide fragment ions. R represents the variable amino acid side chain (see APPENDIX 1C). The non-sequence-specific ions are designated by the single-letter amino acid codes (Table A.1C.1) for the R groups they include.

Table 10.21.2 Common Low-Mass Ions Characteristic of Amino Acids^a

m/z	Amino acid present
<u>70</u> , 87, 100, <u>112</u> , 129	Arginine
<u>70</u> (w/o 87, 100, and 112)	Proline
<u>72</u>	Valine
84 (w/o 101)	Lysine
<u>86</u>	Leucine or isoleucine
101 (w/o 84)	Glutamine
102	Glutamic acid
104	Methionine
106	Pyridylethyl cysteine
<u>110</u>	Histidine
120	Phenylalanine
133	Carbamidomethylated cysteine
134	Carboxymethylated cysteine
<u>136</u>	Tyrosine
<u>147</u>	Acrylamide-modified cysteine
159	Tryptophan
<u>216</u>	Phosphotyrosine

^aUnderlining indicates more abundant ions; w/o, without.

1995). The second caveat is that, unlike Edman degradation, MS cannot at present be used reliably to determine the NH₂-terminal sequence of proteins directly, although several successful attempts have been made (Speir et al., 1995).

MALDI-MS

Matrix-assisted laser desorption/ionization (MALDI)-MS as commonly practiced today was first described by Karas and Hillenkamp (1988). They demonstrated that by cocrystallizing analyte molecules with a large molar excess of a small, UV-absorbing organic acid, large proteins such as serum albumin could be desorbed and ionized by short, intense pulses from a UV laser and detected using a TOF mass spectrometer. Refinement of the technique was rapid. New laser wavelengths and correspondingly optimized matrices were introduced (Beavis et al., 1992). In less than 2 years, proteins >100 kDa were being routinely analyzed with subpicomole sensitivity (Hillenkamp et al., 1991) and molecular weights could be assigned with an accuracy of ~0.1%. Beavis and Chait (1990) demonstrated that for small proteins (<30 kDa) mass accuracies on the order of 0.01% to 0.02% could be achieved by coadding peptides or proteins of known mass to the sample to serve as internal mass reference standards (Beavis and Chait, 1990). This is 2 to 3 orders of magnitude more accurate than SDS-PAGE.

MALDI-MS has also been shown to be extremely useful for the analysis of peptides (Billeci and Stults, 1993). Unlike ESI, MALDI produces predominately singly charged molecular ions from peptides and proteins (although larger proteins can produce multiply charged ions as well; see Fig. 10.21.8), making the analysis of mixtures very much more straightforward. The ease of spectral interpretation combined with femtomole sensitivity and a tolerance for many of the common biological buffers makes MALDI an obvious choice for the analysis of unfractionated enzyme digests. Recent developments in the use of MALDI-MS for the analysis of peptides and proteins have been reviewed (Wang and Chait, 1994; Mann and Talbo, 1996).

The MALDI process can result in extensive fragmentation of the sample. Decomposition that occurs in the ion source, upstream from the field-free region of the flight tube, is referred to as "prompt fragmentation" and can be observed in a linear spectrum. Ions formed by prompt fragmentation are usually of low abun-

dance, and are formed when very high laser irradiance is used. For example, disulfide bonds have been demonstrated to undergo prompt fragmentation to yield the free peptides (Zhou et al., 1993; Patterson and Katta, 1994). Fragmentation that occurs after the source, in the flight tube, is referred to as metastable decay or post-source decay (Kaufmann et al., 1994). These fragment ions, which are often abundant, cannot be observed in a linear TOF instrument but can be observed with the use of a reflectron. As described below, they can be very useful for peptide sequencing.

The mechanism by which MALDI operates is still a matter of some debate (Dreisewerd et al., 1995). The one point upon which everyone agrees, however, is that the matrix is critical to the success of the experiment. In its simplest form, sample preparation for MALDI involves mixing the sample with a large (10⁴) molar excess of matrix and applying 0.5 to 2.0 μl of this solution to the target, where it air dries. Once in the mass spectrometer, the sample is bombarded with photons of UV light. The matrices, all compounds that are strongly UV-absorbing at the designated wavelength, serve three major functions in the experiment. First, the large excess of matrix separates the analyte molecules from one another, thereby reducing intramolecular interactions. Second, the matrix rapidly absorbs large amounts of energy from the incoming photons, resulting in an explosive breakdown of the matrix-analyte lattice. This explosion sends both matrix and analyte molecules into the gas phase. Third, the matrix is necessary for ionization of the analyte molecules. Transfer of protons from the matrix to the peptide and protein molecules is believed to occur via gas-phase reactions in the dense cloud that forms above the target during laser irradiance. The choice of matrix depends on the irradiance wavelength and the type of sample to be analyzed. A N₂ laser operating in the UV at 337 nm is the most typical laser in use on commercial MALDI instruments. Matrices that have been shown to be routinely useful at 337 nm are 3,5-dimethoxy-4-hydroxy-*trans*-cinnamic acid (sinapinic acid), 2,5-dihydroxybenzoic acid, and α-cyano-4-hydroxy-*trans*-cinnamic acid (Hillenkamp et al., 1991; Beavis et al., 1992). Lasers that operate in the infrared at 2.94 μm have also been shown to be useful in MALDI applications, particularly for the analysis of proteins directly from gels or blots (Strupat et al., 1994). MALDI has been demonstrated on magnetic-sector (Annan et al., 1992; Medzihradzky et al., 1996), ion-trap

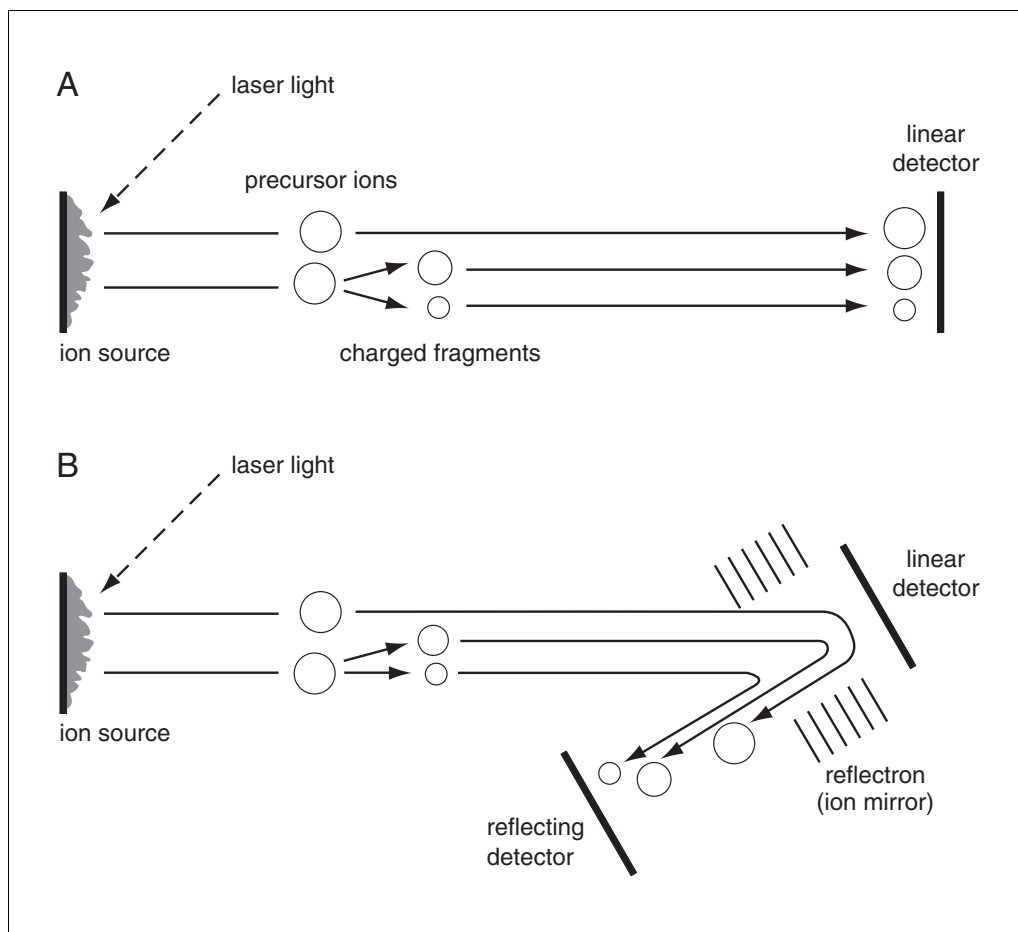


Figure 10.21.2 Schematic diagrams of linear (**A**) and reflectron (**B**) MALDI-TOF mass spectrometers. The flight paths of precursor and fragment ions in each instrument are shown.

(Qin and Chait, 1995), and Fourier-transform ion-cyclotron resonance (Wu et al., 1995) mass spectrometers, but by far the most common application of MALDI has been on TOF instruments.

Typical TOF mass spectrometers can be classified into two broad categories, linear and reflector-type instruments. A simple linear instrument (Fig. 10.21.2, panel A) consists of an ion source, where the MALDI process takes place; a flight tube, where ions of different masses are separated from one another; and a detector. All of these components, with the exception of the laser, are housed in a high vacuum maintained at 10^{-5} to 10^{-8} Torr. After ions are formed by the MALDI process, they are accelerated out of the source under the influence of a strong electric field. All of the ions have essentially the same final kinetic energy upon entering the flight tube; thus, the arrival time of a given ion at the linear detector will vary depending on its mass. These relationships are mathematically expressed by classical equation for kinetic energy (here expressed in

two forms), where E is the kinetic energy of an ion, m is its mass, and v is its velocity:

$$E = \frac{1}{2} mv^2$$

$$v = (2E/m)^{0.5}$$

These equations tell us that because all ions of the same charge have been accelerated to the same energy, their velocities (and therefore, their flight times) will be inversely proportional to the square root of their molecular masses. Thus ions of greater mass will travel slower than lighter ones and will therefore reach the detector later.

Simple linear MALDI-TOF instruments are rather low-resolution mass spectrometers, operating at a resolving power of ~450 to 600 for peptides and 50 to 400 for proteins. For peptides, this means that one cannot measure the distribution of the various naturally occurring isotopes for each peptide, or separate peptides which differ in mass by only a few daltons. It also means that given a resolution (R) of 400, it will be difficult to resolve two proteins >30 kDa that differ only by an initiator Met residue

(e.g., see Fig. 10.21.8). These issues are considered in more detail in the discussion of Fundamentals of Mass Measurement Accuracy and Mass Resolution.

Several important considerations relating to the production of ions in the MALDI source limit the resolving power of linear TOF instruments. The most important of these is the initial kinetic energy spread of individual ion populations: as a result of this spread, ions of any particular mass will, after reaching final accelerating voltage, exhibit a spread in velocities. This means that various members of the same ion population will arrive at the detector at slightly different times, and that in turn gives rise to a broader peak for each ion than would be observed if there were no initial kinetic energy spread. Because the initial energy spread is mass dependent, peaks for higher-mass ions will be disproportionately broader than those for lower-mass ions. Resolution in a linear TOF instrument can be improved by increasing the length of the flight tube. This enhances the time dispersion of ions of different m/z ; unfortunately, it also increases the spread of arrival times for ions of the same m/z due to the initial kinetic energy distribution. This energy spread is typically reduced in magnitude by increasing the final accelerating voltage.

A more effective way to correct for energy distributions is through the use of an ion mirror, or reflectron. A reflectron TOF instrument (Fig. 10.21.2, panel B) corrects for the initial energy spread by acting as an energy-focusing device. The reflector works by slowing an ion down until it stops (the back of the reflector is at a voltage slightly higher than the source-accelerating voltage), turning it around, and then reaccelerating it back out to a second detector. Ions with an initial kinetic energy (and velocity) slightly lower than full accelerating potential will not penetrate the reflector as deeply, and will therefore turn around sooner, catching up to those ions with full kinetic energy. Ions with slightly greater energies (and thus higher velocities) will penetrate more deeply into the reflector, be turned around later, and have their flight times retarded, allowing the other ions to catch up. All of this causes ions of a given mass-to-charge ratio to be spatially focused into packets having flight times that are closer together, thus improving the resolution. This improvement in resolution is illustrated in Figure 10.21.3, which shows the molecular ion region of the linear (panel A) and reflector (panel B) spectra of a synthetic tyrosine phosphorylated peptide TRDIYETDpYYRK. The

isotopically resolved ion cluster in the reflector spectrum shows the contributions of all of the isotopes of C, H, N, O, and P that make up the elemental composition of this peptide. The improvement in resolution is most noticeable at masses of ~ 3000 and lower. For larger molecules, where the resolution is no longer sufficient to provide separation of the isotopes, the natural width of the unresolved isotopic envelope has a large influence on the apparent resolution of the instrument, as explained in the discussion of Fundamentals of Mass Measurement Accuracy and Mass Resolution. In addition to resolution, mass accuracy is also improved when spectra are recorded in the reflector mode (Vorm and Mann, 1994).

Tremendous improvement in the resolution achievable on MALDI-TOF instruments has recently been demonstrated using a focusing technique developed many years ago by Wiley and McLaren (1953). This method, referred to as time-lag focusing or “delayed extraction,” has been shown to provide a resolution of ~ 2000 to 4000 (full-width, half-maximum; or FWHM) in linear mode for peptides and resolution of ~ 3000 to 6000 FWHM in reflecting mode, on a 2-meter-long instrument (Brown and Lennon, 1995; King et al., 1995; Vestal et al., 1995; Whittal and Li, 1995). In a delayed extraction source, ions are created in a field-free region and allowed to spread out before an extraction voltage is applied and the ions are accelerated into the drift tube. The delayed extraction also probably limits the number of collisions the ions undergo on their way out of the source, thereby reducing additional peak broadening caused by metastable decomposition. Mass accuracy with the use of an internal reference standard may also be improved with the use of the delayed extraction source.

Assigning an accurate m/z value to an ion whose flight time has been measured requires that the mass spectrometer be properly calibrated. This is usually accomplished by measuring the flight times of two compounds whose molecular weights are accurately known. In practice, it is usually best to select two compounds with molecular weights that bracket the mass range over which samples are to be examined. For molecules $>10,000$ Da, the $[M+H]^+$ and the $[M+2H]^{2+}$ ions of a well-characterized protein standard work very well for calibration. Application of an internal calibration, where the calibrant peaks are present in the same spectrum as the sample, can provide accuracies of 0.01% to 0.05% (e.g., 0.1 to 0.5 Da at $m/z = 1000$). Internal calibration is sometimes experi-

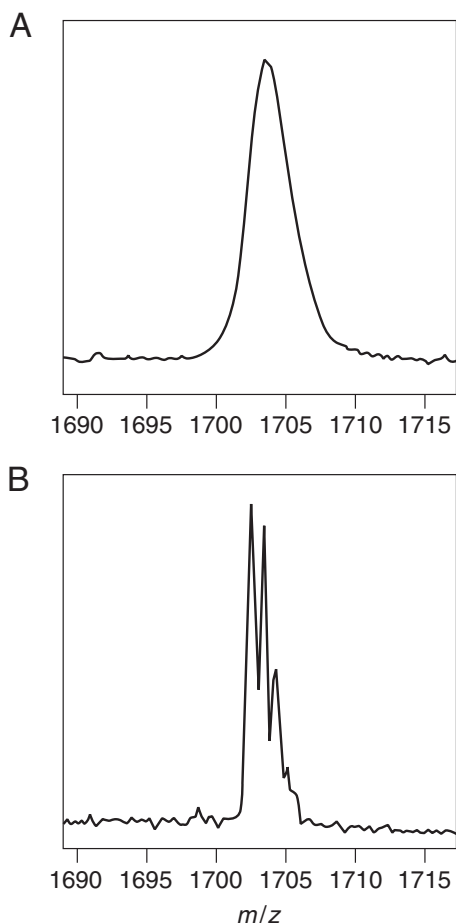


Figure 10.21.3 Molecular ion region from MALDI spectra of the peptide TRDIYETDpYYRK (monoisotopic $M_r = 1701.8$) recorded in the (+) ion mode. Spectra were recorded in (A) the linear mode, where the average mass is measured at the centroid of the unresolved isotopic cluster, and (B) the reflector mode, where the monoisotopic mass is measured at the centroid of the first peak in the cluster. This peak contains the lightest isotopes of the elements present (see discussion of Fundamentals of Mass Measurement Accuracy and Mass Resolution).

mentally difficult owing to the need to adjust the amount of calibrant compounds added so that the calibrant peaks and the analyte peaks are simultaneously observed with reasonable intensity.

MALDI mass spectra may also be calibrated externally; this approach is experimentally simpler. In this method calibrant compounds are measured using the same matrix and laser irradiance as were used for sample analysis. A mass accuracy of ~ 0.1 to 0.2% (e.g., ± 1 to 2 Da at $m/z = 1000$) can be achieved in linear mode using external calibration (Karas et al., 1989). In reflector mode, one can routinely expect mass accuracies on the order of 0.01% to 0.05% using external calibrations. In reflector mode it is also possible to use either the $[M+H]^+$ ion or the $[2M+H]^+$ ion (the protonated gas-phase dimer) of the matrix as one of the two calibrant peaks, which means that only one calibrant compound needs to be added.

MALDI Post-Source Decay MS for Peptide Sequencing

A peptide molecular ion that is sufficiently stable to be transported out of the ion source

but insufficiently stable to survive the flight to the detector will decompose in the flight tube, giving rise to a series of fragment ions that are characteristic of the molecule's original structure. This process is referred to as metastable decay or post-source decay. It is believed that molecular ions acquire excess internal energy necessary for fragmentation via multiple collisions with matrix ions in the source (Kaufmann et al., 1994). When a precursor ion decomposes in the flight tube, all of the fragment ions will have the same velocity as the precursor, but they will have only a fraction of its kinetic energy. This means that the precursor and metastable fragments will all strike the linear detector at the same time and will be detected at the same apparent mass. Thus metastable or post-source-decay fragment ions are never observed in a linear spectrum.

In contrast, a reflector TOF is capable of discriminating fragment ions by flight-time dispersion. Although all ions will have essentially the same velocity, the kinetic energy of a post-source-decay fragment ion will be determined by the ratio of its mass to the mass of the precursor. Thus the fragments, which have less

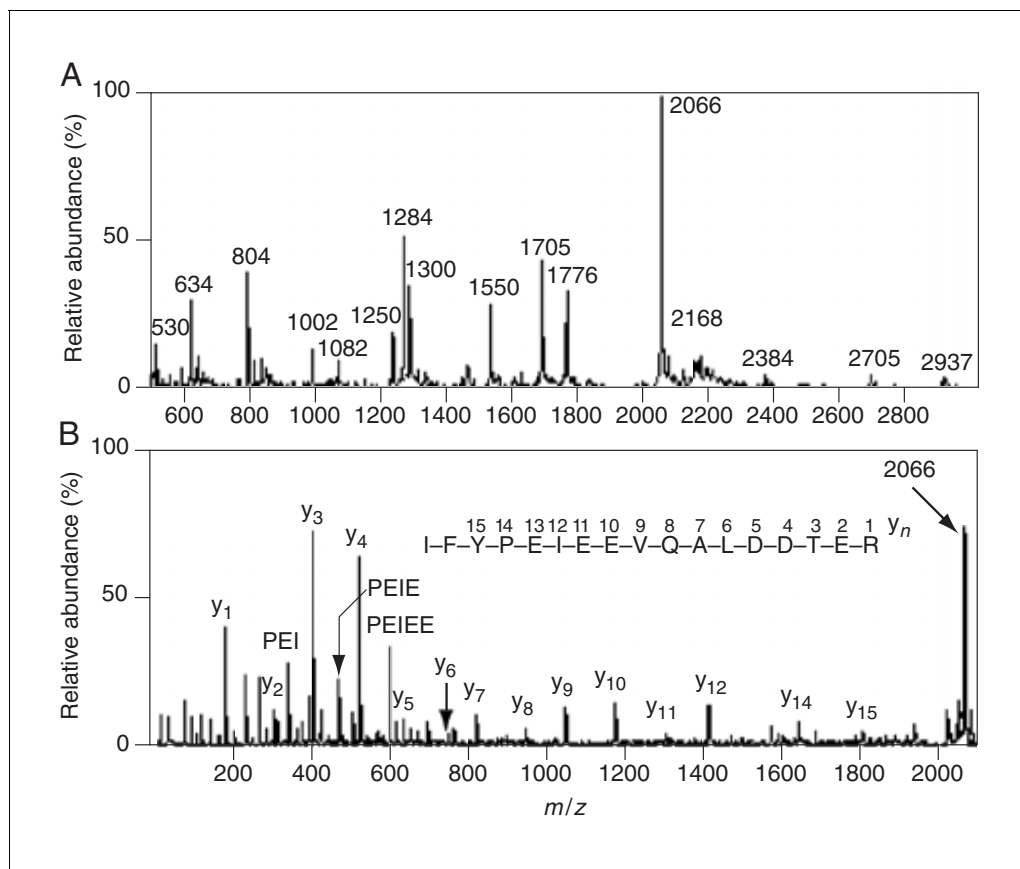


Figure 10.21.4 MALDI-MS analysis of a tryptic digest of an 18-kDa protein (Ladner et al., 1996). A total of 30 pmol of protein was separated on an SDS-PAGE gel, the band of interest excised from the gel, and the protein digested in situ with trypsin; then $\frac{1}{20}$ of the sample was analyzed by MALDI-MS. **(A)** MALDI mass spectrum of the tryptic peptides recorded in the linear mode; **(B)** MALDI-PSD spectrum of the $[M+H]^+$ ion at m/z 2066. The peak was selected using an ion gate ($R = 100$). A nearly complete series of y_n ions (see Fig. 10.21.1) makes it possible to deduce the sequence of residues 3 to 17.

kinetic energy than the precursor, will not penetrate as deeply into the ion mirror, and so will have their flight times decreased relative to the precursor. The fragments are therefore detected at a lower mass. By successively lowering the reflector voltage, thereby bringing lower- and lower-mass fragment ions into focus, a complete post-source decay fragment ion spectrum can be acquired in segments. The individual segments are calibrated and assembled afterward automatically by the data system.

The energy-focusing features of reflectron-type mass spectrometers make it possible to use post-source decay for peptide sequencing (Kaufmann et al., 1994). Post-source decay spectra resemble low-energy collision-induced dissociation (CID) spectra, similar to those commonly recorded on triple-quadrupole mass spectrometers (Hunt et al., 1986). The spectra contain mainly b_n and y_n ions, often with very abundant internal fragments and immonium ions (see Fig. 10.21.1 and Table 10.21.2). The

use of an ion gate makes it possible to perform a low-resolution ($R = 100$) selection of a precursor ion of interest from an unfractionated mixture, such as a protein digest. This approach is very powerful for providing internal sequence information on in situ-digested proteins that have been purified by SDS-PAGE (see Fig. 10.21.4 for an example).

ESI-MS

In electrospray ionization (ESI)-MS, ions are formed from peptides and proteins by spraying a dilute solution of these analytes at atmospheric pressure from the tip of a fine metal capillary (Fig. 10.21.5). The spray process, often assisted by pneumatic nebulization, creates a fine mist of droplets (Fenn et al., 1990; Kebarle and Tang, 1993). The droplets are formed in a very high electric field created by applying a high voltage (~ 4 kV) to either the spray tip (Fig. 10.21.5) or the counter electrode; in this process, the droplets become highly

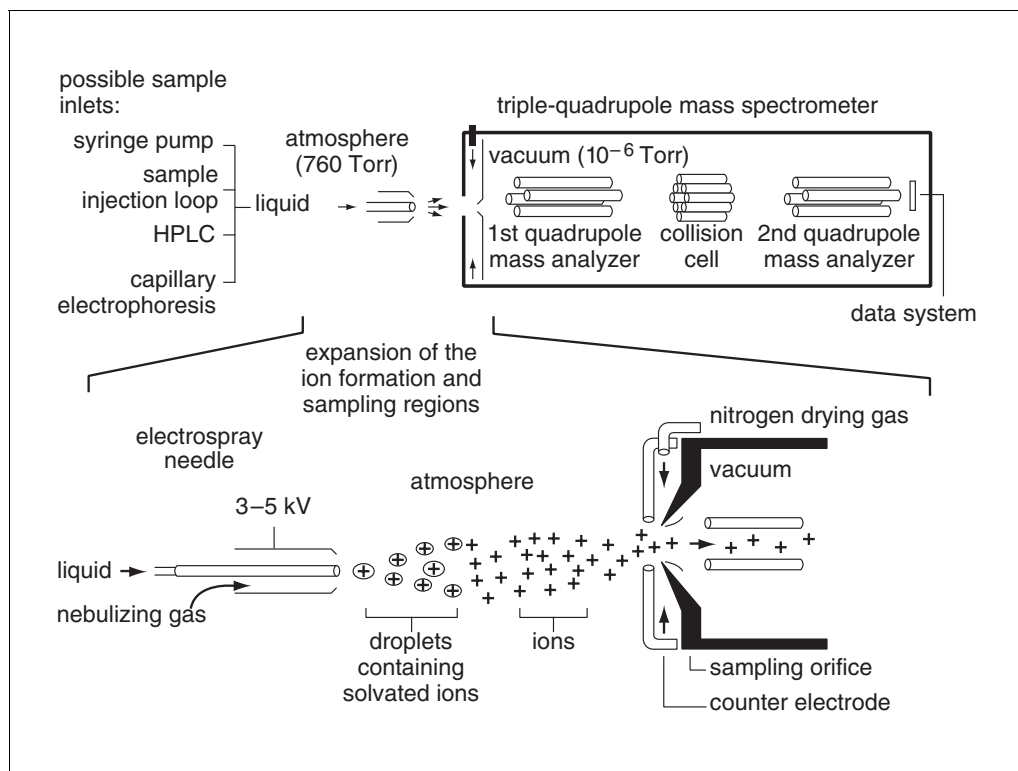


Figure 10.21.5 Schematic diagram of a triple-quadrupole mass spectrometer (top) and the electrospray interface (bottom).

charged. Solvent evaporation is rapid from these small droplets. As the droplets evaporate, the peptide and protein molecules in the droplets pick up one, two, or more protons from the solvent to form singly or, more frequently, multiply charged ions (e.g., $[M+H]^+$, $[M+2H]^{2+}$, etc.). The number of charges acquired by a molecule is roughly equivalent to the number of possible sites of proton attachment. As the droplets continue to shrink, the charge density on the surface of each one increases to the point where charge repulsion overcomes the forces holding together the droplet and the solvated ions contained within it. Ions are then “emitted” or “evaporated” from the droplet surface. The ions are sampled into the high-vacuum region of the mass spectrometer for mass analysis and detection, most often using a quadrupole (or triple-quadrupole) mass analyzer (Fig. 10.21.5). Usually little or no fragmentation is observed in the normal ESI mass spectra of peptides. The typical solvent for peptides and proteins is a mixture of water, an organic modifier such as CH_3CN , and up to a few percent by volume of acetic, trifluoroacetic, or another volatile acid (the latter included to enhance ionization of sample constituents). Because the ions are produced at atmospheric pressure from flowing liquid

streams, ESI is ideally suited for on-line coupling to high-performance liquid chromatography, making it possible to analyze mixtures of peptides and proteins rapidly (see below).

As mentioned above, a key feature of the electrospray process is the formation of multiply charged molecular species from analytes that contain more than one possible site of proton attachment (Fenn et al., 1990; Smith et al., 1991). Proteins usually exhibit a characteristic series of multiply charged ions (e.g., see Fig. 10.21.8). As one proton (i.e., one charge) is typically attached for each 1000 Da of molecular mass, the ion series for a protein, even a large protein, will fall in the m/z range of 800 to 3000. This makes it possible for simple instruments like quadrupole mass spectrometers to be used for mass analysis of ions produced by ESI. The molecular mass of the protein can easily be calculated using the observed masses of any two adjacent ions in the series (Covey et al., 1988; Mann et al., 1989; Reinhold and Reinhold, 1992). Normally these calculations are performed automatically by the data system. Because each multiply charged peak provides an independent measure of the molecular mass of the protein, an ion series from a single experiment provides a measure of the mass measurement precision. Most data sys-

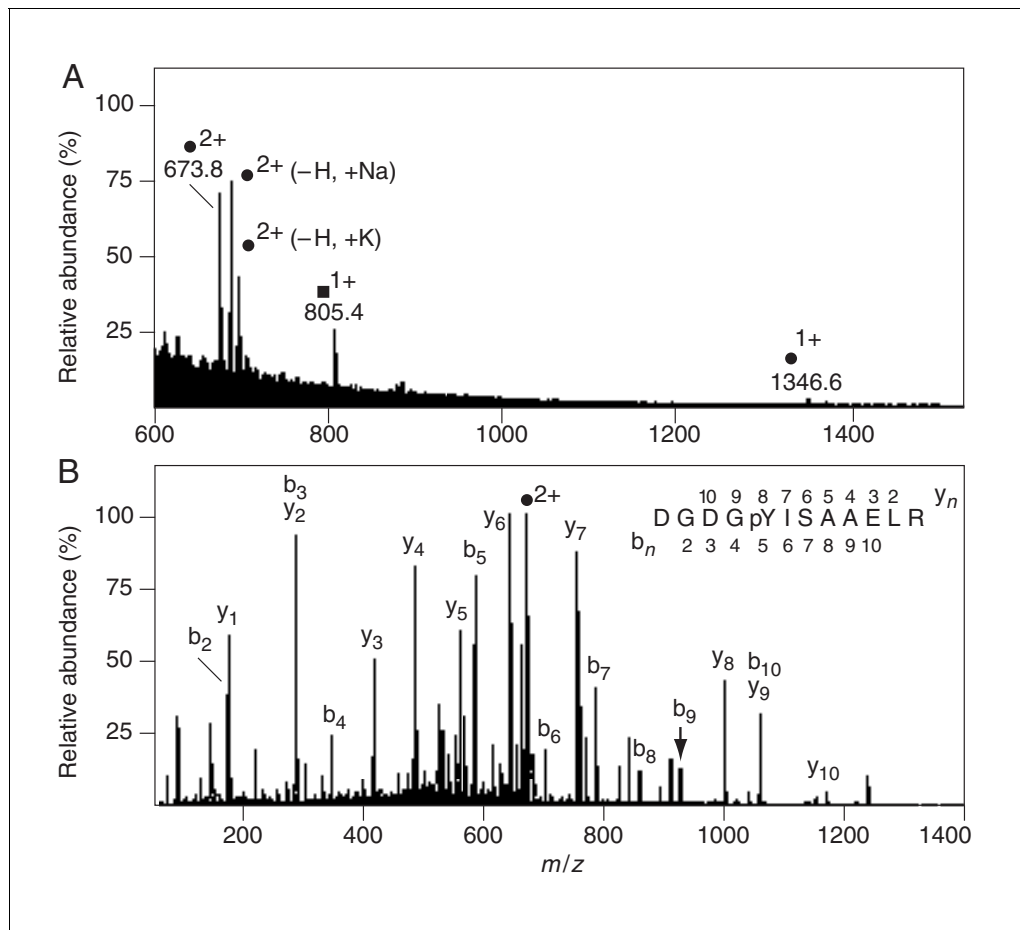


Figure 10.21.6 Electrospray analysis of an HPLC fraction from the tryptic digest of a 17-kDa phosphoprotein. The protein was purified by SDS-PAGE, electroblotted onto PVDF membrane, eluted, digested with trypsin, and the peptides analyzed by on-line LC-ESI-MS using a phosphopeptide-selective scanning method (Huddleston et al., 1993b). The column flow was split after UV detection, with 10% going to the mass spectrometer and 90% to a fraction collector. **(A)** Nanoelectrospray mass spectrum of a putative phosphopeptide-containing fraction. Peaks labeled (+Na) and (+K) correspond to $[M + H + Na]^{2+}$ and $[M + H + K]^{2+}$, respectively. **(B)** Collision-induced decomposition (CID) spectrum of the doubly charged ion ($[M+2H]^{2+}$) at $m/z = 673.8$. This spectrum confirms the sequence of the peptide and identifies the site of phosphorylation as Tyr⁵.

tems also provide algorithms for transforming the mass-to-charge axis into a molecular mass axis for simplifying interpretation and presentation of the data. Proteins with molecular masses of up to 150,000 Da have been successfully analyzed by ESI using commercially available quadrupole instrumentation. Mass assignment accuracies of 0.01% are routine (see Table 10.21.1 and discussion of Fundamentals of Mass Measurement Accuracy and Mass Resolution). ESI has also been successfully interfaced with ion-trap mass spectrometers (Korner et al., 1996), Fourier-transform ion-cyclotron resonance mass spectrometers (Winger et al., 1993; O'Conner et al., 1995; Wu et al., 1995; Smith et al., 1996), and TOF mass spectrometers (Mirgorodskaya et al., 1994; Verentchikov et al., 1994).

In the case of peptides, the maximum number of charges observed often correlates reasonably well with the number of amino acid side chains that can readily accept a proton at the low pH of the analyte stream (i.e., Arg, Lys, and His, plus the free amino terminus). In the case of tryptic peptides, at least two charge sites are normally present, and tryptic peptides with molecular weights of 800 to 3000 typically exhibit a major, if not dominant, $[M+2H]^{2+}$ ion. For example, the ESI mass spectrum of an HPLC fraction from a tryptic digest of a phosphoprotein shows a simple mixture of peptides, in which doubly charged forms of the major peptide predominate (Fig. 10.21.6). In this spectrum, in addition to proton attachment, doubly charged ions due to Na and K adduction are also abundant. In general, the smaller the pep-

tion, the more likely it is that the singly charged $[M+H]^+$ ion will also be observed. As the size of the peptide increases, triply and higher-charged states of the peptide will increase in abundance in the spectra while the abundance of ions with lower charge states will decrease (Fig. 10.21.7).

Electrospray mass spectra are keen reminders that mass spectrometers measure the ratio of mass to charge (m/z) and not molecular mass directly. The peak at $m/z = 1346.6$ in Figure 10.21.6 (panel A) is the monoisotopic $[M+H]^+$ of a peptide of molecular weight 1345.6 (see section on Fundamentals of Mass Measurement Accuracy and Mass Resolution for a discussion of monoisotopic mass). Evidence that this ion carries a single charge (1^+) as opposed to a higher number of charges is obtained from the spacing of the isotope peaks, which are one m/z unit apart. The doubly charged form of this same peptide is observed on the m/z scale at $[1345.6 + 2]/2 = 673.8$. The isotope peaks in the $[M+2H]^{2+}$ isotope cluster are spaced 0.5 m/z units apart, indicating that these ions are doubly charged. Quadrupole mass analyzers can routinely achieve resolution sufficient to distinguish singly from doubly charged ions up to $m/z = 3000$ (if sufficient sample is available and the instrument is scanned slowly). This capability is particularly relevant to tryptic digests of proteins where, as noted above, the doubly charged ion may be the only ion observed for a given peptide. In the absence of the peak spacing information, the analyst could have difficulty determining if such an ion is singly or doubly charged. For ions with three or more charges, a charge series (e.g., 2^+ and 3^+) is usually present that permits the analyst to determine the charge state and mass of any given ion in the spectrum without the need to determine the isotopic peak spacing.

LC-ESI-MS

Electrospray ionization has made on-line liquid chromatography (LC)-MS a practical and robust method for sample introduction because gas-phase ions, as opposed to liquid droplets, are what is mainly being sampled by the instrument. The ideal milieu for presenting a sample to any mass spectrometer, regardless of the ionization mechanism, is in mixtures of highly pure water and volatile organic solvents containing low levels of volatile acid or base. Fortunately, such mixtures (e.g., water, acetonitrile, and trifluoroacetic acid at 0.05% to 0.1% by volume) are also those that provide optimal separation of peptides and proteins by reversed-

phase liquid chromatography. From a practical standpoint, LC-MS reduces or eliminates the need to preparatively fractionate complex mixtures prior to MS, thereby saving valuable instrument time and preventing possible sample losses. LC-MS data may also be searched retrospectively for components present at very minor levels that may not have been considered worthy of collection based on a weak UV response.

In an LC-MS experiment, mass spectra are recorded continuously as the components elute from the HPLC column. If the mass spectrometer is made to scan once every 4 sec, the LC-MS data file will contain ~900 mass spectra. To identify regions of data that are likely to contain useful spectra, the mass spectrometer's data system produces a plot of the total number of ions detected during each mass spectrum scan versus time (the scan number). This plot, which is called a mass chromatogram or a total ion current (TIC) trace, looks very much like the UV chromatogram (Fig. 10.21.7). If it is of interest to know whether a particular peptide is present in the digest and where it elutes, the analyst simply requests the data system to display a selected ion-current trace for the specific masses of the most likely charge states (e.g., 1^+ to 4^+). Figure 10.21.7 illustrates the type of data typically acquired during the LC-ESI-MS analysis of a protein digest. In this example, the protein is heterogeneously glycosylated at a single site, and the multiplicity of peaks evident is due to differing carbohydrates attached to a single tryptic peptide (Fig. 10.21.7, panel C).

Because an ESI mass spectrometer behaves as a concentration-dependent detector, much like a UV detector, only a few percent of the total column effluent (~1 to 5 $\mu\text{l}/\text{min}$) needs to be directed into the mass spectrometer. The remainder can be diverted, via a flow splitter, through an in-line UV detector to a fraction collector. This arrangement makes it possible to perform straightforward correlation of the UV trace with the total ion current trace and then analyze individual fractions further by MS/MS (see discussion of ESI MS/MS for Peptide Sequencing) or some other method. The UV response is also useful for estimating the amount of a peptide that has been injected on the column. This is generally not possible to do by MS without appropriate internal standards (see section entitled Is MS Data Quantitative?). LC-ESI-MS has become a key analytical method for the analysis of natural and recombinant proteins, and data from these experiments is now often included in the regulatory

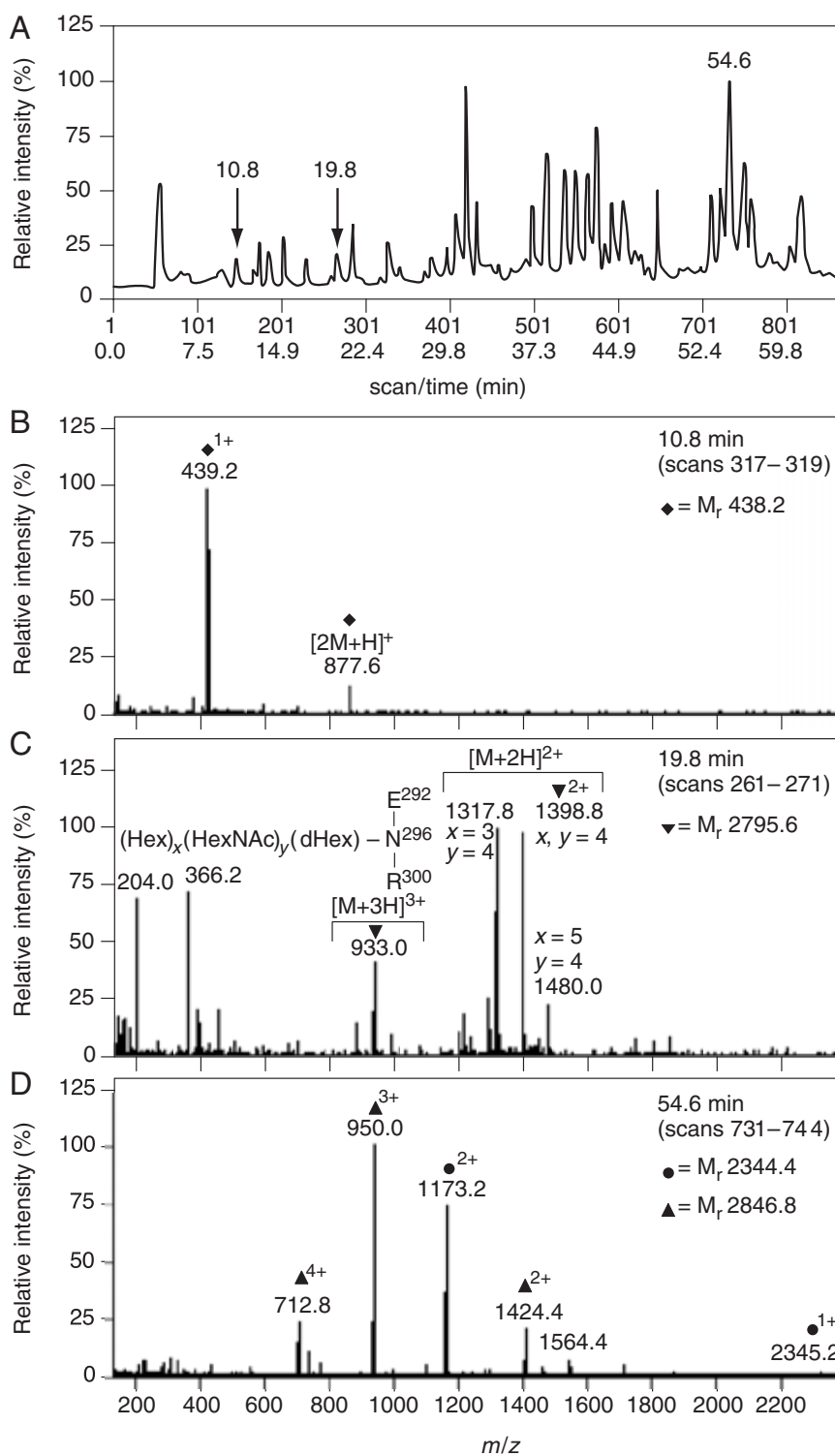


Figure 10.21.7 On-line LC-ESI-MS analysis of a tryptic digest of reduced and alkylated heavy chain (51 kDa) from a recombinant monoclonal antibody. **(A)** Total ion current trace showing the distribution of tryptic peptides. **(B)** ESI spectrum of the peak at 10.8 min showing a small singly charged peptide and its gas-phase dimer. **(C)** ESI spectrum of the peak at 19.8 min corresponding to a tryptic glycopeptide with heterogenous carbohydrate. At least five different carbohydrate structures are indicated. Abundant low- m/z ions at $m/z = 204$ and $m/z = 366$ are characteristic fragment ions produced from the carbohydrate (Roberts et al., 1995). **(D)** ESI spectrum of the peak at 54.6 min showing two large tryptic peptides, each represented by several different charge states.

filings for new biopharmaceutical products (Carr et al., 1989; Palczewski et al., 1994; Stults et al., 1994; Watts et al., 1994; Resing et al., 1995; Roberts et al., 1995; Rush et al., 1995; Schindler et al., 1995; Hemling et al., 1996).

ESI MS/MS for Peptide Sequencing

The principles of MS/MS using triple-quadrupole mass spectrometers have been reviewed (Yost and Boyd, 1990). Briefly, the first quadrupole (Fig. 10.21.5) acts as a mass filter to select specifically the $[M+nH]^{n+}$ ion (where n is typically 1, 2, 3, or 4) of a peptide of interest and isolate it from other ions produced in the source. The selected precursor ion is then fragmented with a neutral gas such as nitrogen or argon in a quadrupole collision cell. This process is known as collision-induced decomposition. The fragment or product ions produced are then transmitted into the quadrupole second mass analyzer, which is being used as a scanning mass analyzer. The product ions are separated based on m/z and detected. The result is a product-ion mass spectrum of the selected precursor that is free of interferences from other components present in the mixture. An example of this process is shown in Figure 10.21.6. The fragmentation of this peptide reveals its sequence and the location of the phosphorylated residue.

Fragmentation can also be induced in the ESI source by increasing the energetics of the ion sampling process. This is done by adjusting the lens potentials in the high-pressure region upstream from the first quadrupole mass filter. Ions with increased kinetic energy undergo decomposition via collisions with residual gas molecules (Katta et al., 1991). Structurally useful fragmentation can thus be obtained on a single-quadrupole instrument, with the caveat that because this method has no ability to preferentially select precursor ions, it requires a highly purified sample. Recently strategies have been developed that take advantage of this ability of the ESI interface to generate fragment ions prior to mass analysis. Selective detection of glycosylated peptides (Carr et al., 1993; Huddleston et al., 1993a), phosphorylated peptides (Huddleston et al., 1993b; Till et al., 1994), and sulfated peptides (Bean et al., 1995) during on-line LC-ESI-MS analysis of protein digests can be accomplished by monitoring the total ion current trace for highly diagnostic low- m/z fragment ions produced in the ESI interface. These procedures are equally applicable to the detection of these modifications in unseparated mixtures.

IS MS DATA QUANTITATIVE?

A question that frequently arises in the analysis of peptide or protein mixtures is whether the relative peak heights (or areas) are quantitatively representative of the relative solution concentrations of the sample components. Unfortunately, the answer to this question is no, unless appropriate internal reference standards are used. Due to differing ionization and desorption efficiencies, the absolute yields of different ions from one peptide in a mass spectrometer can vary by perhaps as much as 2 orders of magnitude, depending on the sequence, composition, and size of the peptides (Dunayevskiy et al., 1995; Jespersion et al., 1995; Cohen and Chait, 1996). Furthermore, strongly ionizing components can frequently suppress the ionization of more weakly ionizing components present in the same sample or chromatographic fraction. In MALDI, the matrix itself can strongly influence the detectability of specific sample constituents. Quantitation by mass spectrometry requires that the response of a specific sample component be determined relative to that of a reference compound added to the sample and analyzed under identical conditions. In certain cases, semi-quantitative estimates of relative amounts of components may be made if the components are of nearly identical structure and their chemical differences do not significantly affect the overall charge or hydrophobic/hydrophilic character of the molecules. For example, the relative ratios of the variants of β -casein observed in the ESI mass spectrum (Fig. 10.21.8) are probably a good reflection of the relative solution concentrations of these variants, as they differ from one another by only one or two amino acid substitutions out of more than 200 amino acids. The smaller the molecule, the more likely it is that such estimates will be in error.

SAMPLE PREPARATION

The ideal method for preparing samples for any type of MS analysis is reversed-phase high-performance liquid chromatography (RP-HPLC). If purification by RP-HPLC is not an option because of known problems with sample loss, then the cleanup protocol employed should produce a sample composition that resembles as closely as possible that of a sample which has been purified by RP-HPLC. That is, the sample should contain the least amount possible of buffers, salts, detergents, and anything else besides water, organic modifier, and volatile acid or base.

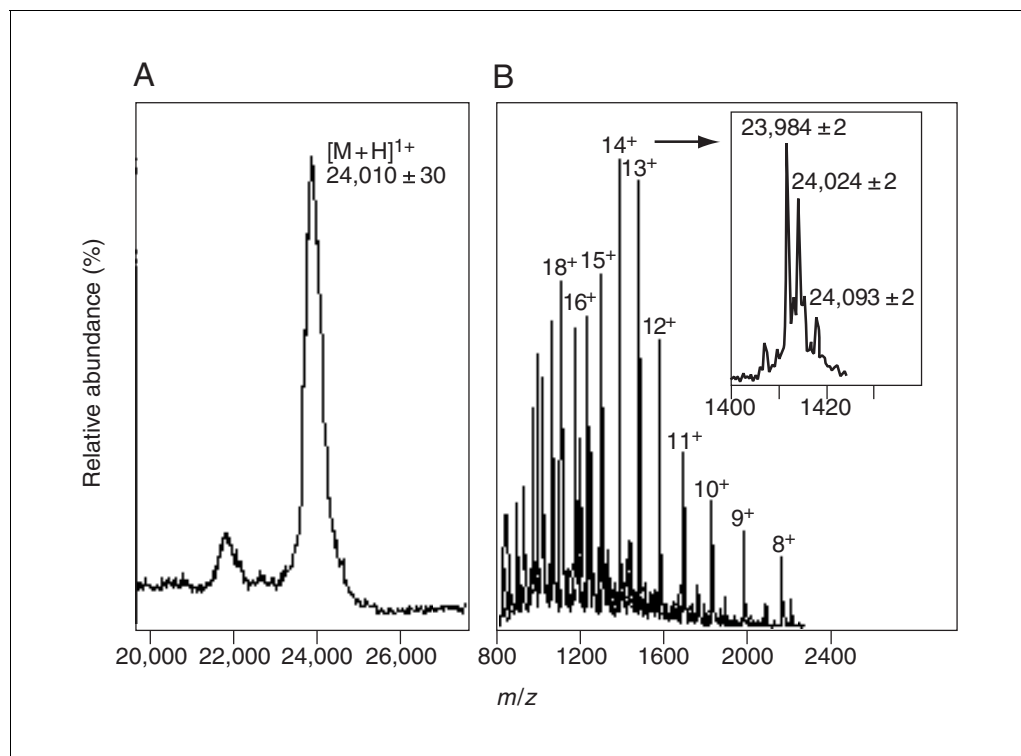


Figure 10.21.8 Molecular weight determination of β -casein, a 24-kDa phosphoprotein with known sequence heterogeneity. **(A)** Linear MALDI-TOF spectrum. Molecular weight was determined from the centroid of the smoothed peak using an external calibration. **(B)** ESI spectrum. The inset, an expansion of the region around the 14^+ charge state, reveals the presence of at least three variants. The molecular weights shown are the averages of determinations made from ten individual charge states.

The quality of the data obtained using ESI-MS is very dependent on the type and concentration of excipients present in the sample. If buffers are necessary for protein solubility or enzymatic digestion, it is best to use a volatile buffer such as ammonium bicarbonate or ammonium acetate at a concentration of ≤ 30 mM. In general it is best to avoid using any ionic detergents (surfactants), even though it is possible to obtain spectra for some proteins when these are included as long as the surfactant concentration is kept $< 0.01\%$. ESI tolerates nonionic detergents better. Analytically useful data may be obtained for sample solutions containing certain nonionic detergents at concentrations between 0.01% and 0.1% (Ogorzalek Loo et al., 1994).

Samples containing unacceptable amounts of buffers, salts, and detergents need to be purified prior to ESI-MS analysis. If the sample can be eluted from an RP column, some type of on-line procedure is both rapid and efficient. For water-soluble salts and buffers, a short, small-i.d. RP cartridge installed in place of the sample loop on a HPLC injector serves to both concentrate and purify samples (Stoney and

Nugent, 1995). By placing the trapping cartridge in the injector loop, it is possible to wash the buffers and salts into waste rather than into the mass spectrometer. After loading the aqueous sample onto the trap (sample volume is no longer important except for very hydrophilic sample components), the trap is washed with 1% to 5% organic solvent and the salts and buffers are diverted to the waste line. The trap column is then put in-line with the solvent delivery pump by switching the valve, and the sample is eluted into the mass spectrometer with the appropriate solvent. This is also a convenient way to analyze dilute solutions, as the trap serves to preconcentrate the sample.

The same approach can be used to purify samples containing anionic and nonionic detergents (Stoney and Nugent, 1995). However, on-line use of detergent columns require that this step be done in tandem with an RP column, because either the sample will have to be eluted with a buffer and thus will need to be desalted prior to introduction into the mass spectrometer, or it will not be retained on the column and will need to be trapped on an RP column. An alternative approach for removing SDS from

samples is to load the sample, dissolved in a high concentration of organic solvent, onto a polyhydroxyethyl aspartamide column (Jeno et al., 1993). Under these conditions, peptides and proteins are retained while the SDS passes through with the solvent front. The sample is then eluted using a gradient of decreasing organic strength.

All of the purification strategies described above can, of course, also be performed off-line. In either case, the use of small-bore HPLC columns and a low-flow HPLC system will provide better results with sample amounts in the picomole range or below. Because the ESI mass spectrometer acts as a concentration-dependent detector, the highest sensitivity will be achieved using the smallest-i.d. columns. For off-line cleanup, the use of the lower flow rates required for small-bore columns means that samples are collected in smaller volumes. This can be important for small amounts of sample, where losses associated with reducing solvent volumes prior to the next step of analysis often are great. The authors have found that 0.5- to 1-mm-i.d. columns provide a practical compromise between on-line sensitivity and easy collection of fractions using a post-column flow splitter. A number of groups have developed sample purification and fraction collection methods that use packed fused silica capillary columns and solvent flow rates of 0.1 to 5.0 $\mu\text{l}/\text{min}$ (Zhang et al., 1995; Gulden et al., 1996) or micropipets filled with packing material (Houthaeve, et al., 1995).

Other approaches to protein purification for MS include ultrafiltration (APPENDIX 3C), TCA precipitation, size exclusion chromatography (UNIT 10.14), and microdialysis (APPENDIX 3C). Among these alternatives, the authors find the easiest and most reliable methods for desalting samples or effecting a buffer change to be ultrafiltration using membranes of varying molecular weight cutoff (MWCO), and small-column (1-mm-i.d.) size-exclusion chromatography. Final volumes of 25 μl can be achieved using the ultrafiltration devices.

If the protein is being purified for subsequent proteolytic digestion with the ultimate goal being mass spectrometry-based peptide mapping, a suitable approach is SDS-PAGE followed by *in situ* enzymatic digestion (Stone and Williams, 1995; Williams and Stone, 1995; Shevchenko et al., 1996; Wilm et al., 1996). The broad applicability and high resolving power of SDS-PAGE applied to protein mixtures makes this approach very useful for achieving difficult separations or for working

with proteins that have poor solubility characteristics. Electrophoresing as little as a few micrograms of individual protein on an analytical gel will usually provide an informative peptide map. Following SDS-PAGE, the protein can also be blotted onto a membrane and enzymatically digested (Hess et al., 1993; Fernandez and Mische, 1995; Patterson, 1995). Recovery of peptides from individual proteins appears to be similar regardless of whether they were produced from an *in-gel* digest or digestion off a membrane. If the peptide map is to be recorded using MALDI, the peptides can be sampled either directly from the digest or from the peptide extract. If the map is to be recorded using ES, then on-line LC-MS is usually the best approach.

Although MALDI is very much more tolerant of the presence of salts, buffers, and some detergents, it too generally benefits from having the sample in the purest form possible, especially when dealing with subpicomole amounts of proteins or low femtomole amounts of peptides. All of the methods discussed above work equally well for preparing samples for MALDI-MS analysis.

FUNDAMENTALS OF MASS MEASUREMENT ACCURACY AND MASS RESOLUTION

Monoisotopic Mass and Average Mass

Most of the elements that comprise typical biological molecules have more than one naturally occurring isotope, each with a unique mass and natural abundance. For example, carbon has two principle isotopes, ^{12}C and ^{13}C (see Table 10.21.3). The lighter, more abundant isotope of carbon has, by definition, a mass of 12.000000 and a natural abundance of 98.9%, while the heavier isotope has a mass of 13.003355 and a natural abundance of 1.1%. In a mass spectrometer one measures a large, statistical sampling of the molecules present and, therefore, observes all of the isotopes for each element present. The **monoisotopic mass** of a molecule is the sum of the accurate masses (including the decimal component, referred to as the mass defect) for the most abundant isotope of each element present (Yergey et al., 1983). As the number of atoms of any given element increases, the percentage of the population of molecules having one or more atoms of a heavier isotope of this element also increases. The most significant contributor to the isotopic peak pattern for biomolecules is the ^{13}C isotope of carbon. For peptides with mo-

Table 10.21.3 Isotopic Mass and Abundance Values

Isotope	Mass ^a	Natural abundance (%) ^b
¹ H	1.007 825 035	99.985
² H	2.014 101 779	0.015
¹² C	12.000 000 000	98.90
¹³ C	13.003 354 826	1.10
¹⁴ N	14.003 074 002	99.634
¹⁵ N	15.000 108 97	0.366
¹⁶ O	15.994 914 63	99.762
¹⁷ O	16.999 131 2	0.038
¹⁸ O	17.999 160 3	0.200
¹⁹ F	18.998 403 22	100
²³ Na	22.989 767 7	100
²⁸ Si	27.976 927 1	92.23
²⁹ Si	28.976 494 9	4.67
³⁰ Si	29.973 770 1	3.10
³¹ P	30.973 762 0	100
³² S	31.972 070 698	95.02
³³ S	32.971 458 428	0.75
³⁴ S	33.967 866 650	4.21
³⁶ S	35.967 080 620	0.02
³⁵ Cl	34.968 852 728	75.77
³⁷ Cl	36.965 902 619	24.23
³⁹ K	38.963 707 4	93.2581
⁴⁰ K	39.963 999 2	0.0117
⁴¹ K	40.961 825 4	6.7302
⁷⁹ Br	78.918 336 1	50.69
⁸¹ Br	80.916 289	49.31
¹²⁷ I	126.904 473	100

^aData from Wapstra and Audi (1985). Standard errors are given in this article but are omitted from the present table.

^bLide (1995).

lecular masses <2000 Da, the first peak in the resolved isotopic cluster corresponding to the all-¹²C species will be the most abundant (Fig. 10.21.9, panel A). However, for peptides with molecular masses ≥2500 Da, the first peak in the isotopic cluster will no longer be the most abundant (Fig. 10.21.9, panel B). In this example, the species containing one ¹³C atom is now the most abundant.

For proteins, the heavy isotopes of nitrogen, sulfur, and oxygen (in addition to carbon) make significant contributions to the mass, and, as a result, the monoisotopic peak in the molecular ion cluster appears to vanish (Fig. 10.21.9, panels C and D). In this case a different definition for the mass of the molecule must be used. Whenever it is not possible to observe the all-¹²C peak for a molecular ion, the average

mass should be used. The **average mass** of a molecule is the sum of the chemical average masses of the elements present (Fig. 10.21.9, peaks labeled A). The chemical average mass of an element is the sum of the abundance-weighted masses of all of its stable isotopes (Yergey et al., 1983). For example, the isotope-weighted average mass of carbon is 12.0111, of hydrogen is 1.0079, of nitrogen is 14.0067, of oxygen is 15.9994, of sulfur is 32.0666, and of phosphorus is 30.9738. These masses are known with less precision than the monoisotopic masses of the elements because of variations in the natural abundances of the isotopes. The theoretical appearance of a molecule's isotope envelope may be precisely modeled by solving a polynomial expression that calculates the abundance-weighted sum of the isotopes of

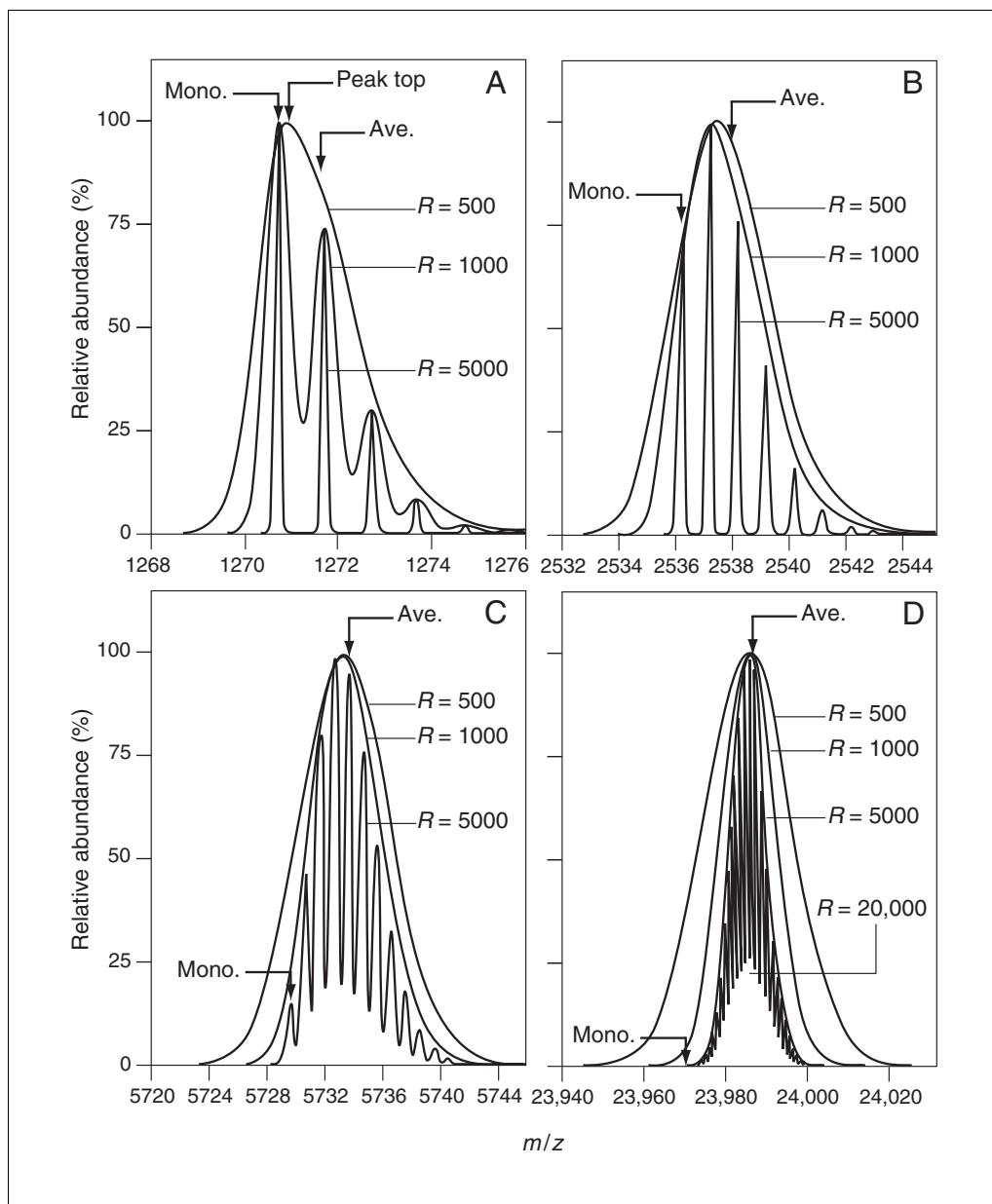


Figure 10.21.9 Effect of resolution (10% valley definition) on the appearance and relative peak intensity of the isotopic envelopes for representative peptides and proteins in four different mass ranges. Abbreviations: *R*, resolution; *M*, monoisotopic mass (Mono.); *A*, average mass (Ave.); *P*, peak top. **(A)** $C_{60}H_{86}N_{16}O_{15}$: $M = 1270.7$, $A = 1271.5$, $P = 1270.8$. **(B)** $C_{115}H_{161}N_{31}N_{35}$: $M = 2536.2$, $A = 2537.7$. **(C)** $C_{254}H_{377}N_{65}O_{75}S_6$: $M = 5729.6$, $A = 5733.6$. **(D)** $C_{1080}H_{1697}N_{268}O_{325}P_5S_6$: $M = 23968.2$, $A = 23983.3$.

each element to the molecular ion cluster (Yergey, 1983). Nearly all modern mass spectrometer data systems come with programs that perform these calculations and can simulate the appearance of the resulting isotope pattern at different instrumental resolving powers.

Resolution

Whether or not one observes the individual isotopes that make up the molecular ion cluster for a given compound depends on the mass

resolution of the instrument being used. The resolution of the mass analyzer is a measure of its ability to separate signals of adjacent molecular mass. Mass resolution is often expressed as the ratio $m/\Delta m$, where m and $m + \Delta m$ are the masses (in atomic mass units or Da) of two adjacent peaks of approximately equal intensity in the mass spectrum (Fig. 10.21.10). There are presently two definitions of resolution in widespread use. In the “10% valley” definition, the two overlapping, adjacent peaks

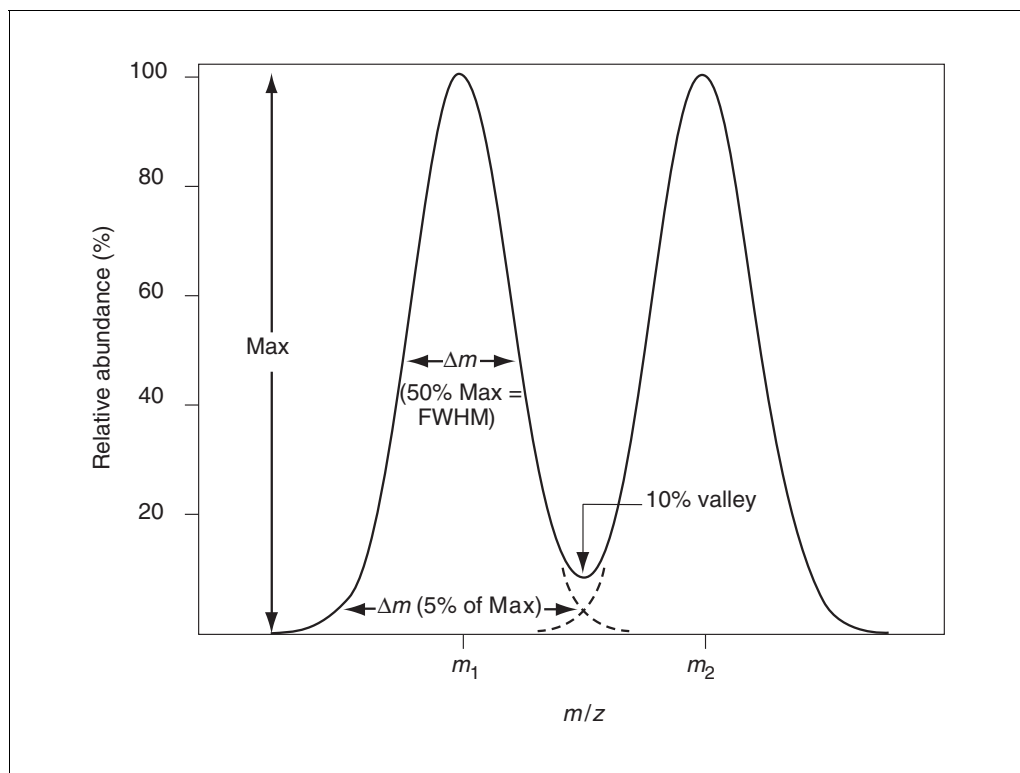


Figure 10.21.10 Schematic representation of the two common definitions of resolution used in mass spectrometry—10% valley and full-width, half-maximum (FWHM).

each contribute 5% to the resulting 10% valley between them (Fig. 10.21.10). The resolution of the mass analyzer may also be determined using a single peak by measuring the peak width (in m/z) at the 5% level and dividing this number into the m/z of the peak. The 10% valley definition is most frequently used for molecules with masses <5000 Da, where it is possible on certain analyzers to resolve the isotope peaks. For *unresolved* isotopic clusters the “full-width, half-maximum” (FWHM) definition is used. By this definition, the resolution is equal to the mass of the peak (in m/z) divided by the width (in m/z) of the isotopic envelope measured at the half-height of the peak (Fig. 10.21.10). An unfortunate consequence of the use of this definition is that, given an FWHM resolving value, it is more difficult to visualize what the separation between two peaks of a given masses will be. For example, a resolution of 1000 by the FWHM definition means that a peak at mass 1001 will be unresolved from the peak at mass 1000. Creation of a detectable separation between these two peaks would require a FWHM resolution >1200. A useful rule of thumb is that the value for the resolution determined using the FWHM definition is approximately twice that obtained using the 10% valley definition. For example, a resolution of

1000 using the 10% valley definition is approximately equivalent to a resolution of 2000 using the FWHM definition (i.e., the measured degree of separation of the peaks in Da/ z is identical).

Mass Accuracy

Getting as accurate a molecular weight as possible is, of course, a central goal of any MS experiment. Here mass accuracy is taken to mean the difference in mass between the experimentally measured value and the theoretical value. By definition, the mass measurement accuracy can be directly determined only for compounds of known elemental composition and, therefore, known molecular mass. For unknown samples, the accuracy of the mass determination is approximated by measuring the mass of a peptide or protein of known elemental composition and similar mass to the unknown. It is important that these measurements be made under experimental conditions as close as possible to those used for analysis of the unknown. How useful this estimate of the mass measurement accuracy is depends on how well the instrument has been calibrated and on the reproducibility of the measurement. The reproducibility, or precision, of the measurement is best evaluated by making replicate measure-

ments of the sample. However, in cases where the amount of sample is limited or the data is too noisy, a reasonable approach is to estimate the precision from replicate measurements of a known peak of mass and intensity similar to the unknown. It cannot be overemphasized that the mass spectrometer's performance with respect to mass measurement accuracy and precision must be regularly evaluated using well-characterized standards covering the mass range of interest in order to provide confidence in the measurements of unknowns being made. It is desirable to achieve a mass measurement precision of ± 0.5 Da or less for peptides and ± 1 Da for proteins. In practice, it is relatively easy to achieve the stated value for peptides but more difficult for proteins (Table 10.21.1). Whether the monoisotopic mass or the average mass should be used when measuring and reporting molecular weights will depend on the mass of the substance and the resolving power of the mass spectrometer (see discussion of Interplay Between Resolution and Molecular Weight).

The accuracy of the mass measurement may be stated in several different ways: as either an absolute mass error or a percentage of the measured mass (e.g., molecular mass = 1000 ± 0.1 Da or $\pm 0.01\%$), or as parts per million (e.g., molecular mass = 1000 ± 100 ppm). As the mass being considered increases, the absolute mass error corresponding to the percent or ppm error will also increase proportionally (e.g., 0.01% or 100 ppm = 0.1 Da at $m/z = 1000$, 0.5 Da at $m/z = 5000$, or 5 Da at $m/z = 50,000$). It is important not to confuse mass accuracy with the term "accurate mass"; the latter is reserved for mass measurements accurate to ≤ 10 ppm.

Interplay Between Resolution and Molecular Weight

In the mass range up to ~ 3000 Da it is recommended that, if possible, a resolution sufficient to partially resolve the isotope cluster be used to facilitate mass assignment of the monoisotopic peak (Ingendoh et al., 1994). Resolution sufficient to separate single-mass-unit differences is also important in order to detect the presence of modifications such as deamidation of Asn to Asp and to make it possible to directly determine the charge state of an ESI ion by measuring the isotope peak spacing. For poorly resolved or unresolved isotope clusters, the mass accuracy can be dramatically affected by how the data is treated by the instrument's data system. It is important to be aware of whether the data system is assigning

masses to the centroid of the peak or to the peak top. In general, the centroid and the peak top mass will not be the same in this mass range because the isotopic distribution comprising the unresolved envelope of peaks is asymmetric in shape. Below mass ~ 3000 it is sometimes difficult to accurately determine the centroid, and therefore the average mass, of an unresolved isotope cluster. The reason for this difficulty is that the bottom of the cluster, which must be included to get an accurate determination of the centroid for peaks in this mass range, is often distorted or poorly defined due to sample background (or, in MALDI, matrix) noise and/or poor ion statistics. These factors can easily contribute errors of ≥ 1 Da to the mass determination for compounds in the mass range of 1000 to 3000 Da. Furthermore, it should be noted that the peak top mass, which is the most easily determined value, corresponds to neither the monoisotopic mass or the average mass in this mass range. At ~ 1200 and below, the mass at the peak top of the unresolved envelope will be closer to the monoisotopic mass, whereas at ~ 3000 Da and above, the peak top will be closer to the average mass. For these reasons the use of resolved isotopes in this mass range is clearly preferable.

Determining the monoisotopic mass of a protein is difficult or impossible due to the insignificant contribution that the monoisotopic peak makes to the isotopic envelope (Fig. 10.21.10, panel D). Furthermore, quadrupole and TOF mass analyzers cannot provide the resolving power necessary to separate the isotope peaks for molecules with masses in excess of ~ 8000 Da. Fortunately, at masses above $\sim 10,000$, the unresolved peak profiles are quite symmetrical and the mass at the peak top is nearly identical to the chemical average mass of the protein. For proteins, errors in measuring the centroid can be caused by the presence of unresolved adducts along with the protonated molecular ion (with, for example, Na^+ , K^+ , or photochemical adducts of MALDI matrices) or of other proteins of similar but not identical mass that are not resolved. The relative abundance of any of these adducts is unpredictable, and will shift the apparent centroid of the unresolved cluster to higher mass in an unpredictable fashion. In these cases, it will be impossible to obtain an accurate mass assignment unless the instrument resolving power can be increased or the contributing adduct ions can be removed by more stringent cleanup of the sample (Ingendoh et al., 1994; Zubarev et al., 1995).

Thus it is still important even at higher mass to have a reasonable amount of resolving power. Resolution should be sufficient to enable detection of common protein variants that may be present in the sample. For example, β -casein exists as a family of isoforms differing by the substitution of one or more amino acids. The MALDI and ESI mass spectra of β -casein are compared in Figure 10.21.8. Detection of one of these variants, which differs by 40 Da in 24,000 Da, requires a resolution of ~ 1200 FWHM ($24,000/40_{10\% \text{ valley def.}} \times 2$), which is easily achievable by ESI-MS on quadrupole mass analyzers. The resolution in MALDI-MS with a TOF analyzer is typically <1000 at this mass, and is insufficient to detect such small modifications. The minimum desirable resolution can be estimated from the natural width of the molecular ion envelope obtained from the computer programs mentioned above. For example, the envelope of isotopes for a 24-kDa protein has a width of ~ 10 Da at the half-height of the peak (Fig. 10.21.9, panel D). This corresponds to an apparent resolution of 2400 ($24,000/10$). Having an *instrumental* resolution higher than 2400 but less than $\sim 20,000$ (see Fig. 10.21.9) will not improve the apparent resolution because the minimum width of the unresolved protein molecular ion cluster is determined by the weighted average of the natural abundance of the isotopes present. It should be clearly stated, however, that having a higher resolution than is theoretically required is not detrimental (as long as it does not result in a decrease in sensitivity), and it may have some beneficial effect on the ability to precisely define the peak top and therefore the average mass.

OUTLOOK

Mass spectrometry and its application to biological problems are clearly enjoying a second childhood. MS-based methods for protein sequencing have sensitivity in the femtomole range, exceeding that of Edman-based sequencing, and MS-based methods are uniquely capable of analyzing the ever-broadening spectrum of protein modifications. Mass spectrometry is even showing promise as a tool for studying protein folding (Englander, 1993; Smith and Zhang, 1994; Wagner and Andereg, 1994) and noncovalent interaction (Smith and Light-Wahl, 1993; Cheng et al., 1995). Further increases in the understanding of the ionization/desorption processes in ESI and MALDI are likely to translate into even higher sensitivity through higher ion-sampling efficiency. It

can also be expected that new methods for dissociating molecular ions of peptides and proteins (possibly involving surface-induced collisions or photoexcitation) will be developed that will be more efficient or more selective than collision-induced or post-source decay. One can also anticipate that computer algorithms for the interpretation or “fingerprinting” of MS/MS data will continue to improve so that soon a minimum amount of human intervention will be required for identification (for an example, see Eng et al., 1994). Although MS-based approaches have sufficient raw sensitivity to analyze low femtomole amounts of peptides and proteins, purifying and handling samples at these levels remains a significant challenge for which there is no general strategy at present. With further improvements, gel- and membrane-based methods and microchromatography-based methods for immobilizing minute amounts of sample for biochemical manipulation and cleanup may meet this challenge.

The mantra for MS now and into the next millennium is “smaller, cheaper, faster, simpler, more sensitive.” These criteria are based in large part on the need to make MS more accessible to the biological community. A variety of new types of mass spectrometers have recently been introduced that satisfy one or more of these criteria, and certain of these new instruments may provide unique benefits for molecular mass determination and sequencing of minute amounts of peptides and proteins. ESI and MALDI have shown particular promise when coupled to quadrupole ion trap analyzers (Qin and Chait, 1995; Korner et al., 1996). These devices are capable of very high resolution and very rapid scanning, and can perform MS^n experiments, in which fragment ions formed in the first MS/MS (MS_2) experiment are further fragmented to provide additional structural information. TOF mass analyzers have been coupled to ESI sources (Mirgorodskaya et al., 1994; Verentchikov et al., 1994) and offer the promise of improved resolution, very high sensitivity, and very fast sampling rates. An interesting variation on these two approaches couples an ion trap to a TOF analyzer, providing structural information as well as molecular mass data (Lee and Lubman, 1995). Additional variations on this theme of connecting two different, relatively inexpensive analyzers together to produce a more versatile instrument are under development. Although neither small or inexpensive, Fourier-transform ion-cyclotron resonance mass spectrometers equipped with MALDI and ESI are extremely versatile

instruments with the potential for simultaneously providing very high mass accuracy, high resolution (routinely $\geq 20,000$), and extremely high sensitivity for peptide and protein mass measurement (Winger et al., 1993; O'Conner et al., 1995; Wu et al., 1995; Smith et al., 1996). These devices are also capable of performing MSⁿ experiments.

LITERATURE CITED

- Aebersold, R. 1993. Mass spectrometry of proteins and peptides in biotechnology. *Curr. Opin. Biotechnol.* 4:412-419.
- Annan, R.S., Köchling, H.J., Hill, J.A., and Biemann, K. 1992. Matrix-assisted laser desorption using a fast atom bombardment for source and a magnetic sector mass spectrometer. *Rapid Commun. Mass Spectrom.* 6:298-302.
- Bartlet-Jones, M., Jeffery, W.A., Hansen, H.F., and Pappin, D.J. 1994. Peptide ladder sequencing by mass spectrometry using a novel, volatile degradation reagent. *Rapid Commun. Mass Spectrom.* 8:737-742.
- Bean, M.F., Annan, R.S., Hemling, M.E., Mentzer, M., Huddleston, M.J., and Carr, S.A. 1995. LC-MS methods for selective detection of posttranslational modifications in proteins: Glycosylation, phosphorylation, sulfation, and acylation. *In Techniques in Protein Chemistry VI* (J. Crabb, ed.) pp. 107-116. Academic Press, San Diego.
- Beavis, R.C. and Chait, B.T. 1990. High-accuracy molecular mass determination of proteins using matrix-assisted laser desorption mass spectrometry. *Anal. Chem.* 62:1836-1840.
- Beavis, R.C., Chaudhary, T., and Chait, B.T. 1992. α -Cyano-4-hydroxycinnamic acid as a matrix for matrix assisted laser desorption mass spectrometry. *Org. Mass Spectrom.* 27:156-158.
- Biemann, K. 1988. Contributions of mass spectrometry to peptide and protein structure. *Biomed. Environ. Mass Spectrom.* 16:99-111.
- Biemann, K. 1990. Sequencing of peptides by tandem mass spectrometry and high-energy collision-induced dissociation. *Methods Enzymol.* 193:455-479.
- Biemann, K. 1992. Mass spectrometry of peptides and proteins. *Annu. Rev. Biochem.* 61:977-1010.
- Billeci, T.M. and Stults, J.T. 1993. Tryptic mapping of recombinant proteins by matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.* 65:1709-1716.
- Brown, R.S. and Lennon, J.J. 1995. Mass resolution improvement by incorporation of pulsed ion extraction/ionization linear time-of-flight mass spectrometry. *Anal. Chem.* 67:1998-2003.
- Burlingame, A.L. and Carr, S.A. 1996. Mass Spectrometry in the Biological Sciences. Humana Press, Totowa, N.J.
- Burlingame, A.L. and McCloskey, J.A. (eds.) 1990. Biological Mass Spectrometry. Elsevier, New York.
- Burlingame, A.L., Boyd, R.K., and Gaskell, S.J. 1994. Mass spectrometry. *Anal. Chem.* 66:634R-683R.
- Carr, S.A., Hemling, M.E., Folena-Wasserman, G., Sweet, R.W., Anumula, K., Barr, J.R., Huddleston, M.J., and Taylor, P. 1989. Protein and carbohydrate structural analysis of a recombinant soluble CD4 receptor by mass spectrometry. *J. Biol. Chem.* 35:21286-21295.
- Carr, S.A., Hemling, M.E., Bean, M.F., and Roberts, G.D. 1991. Integration of mass spectrometry in analytical biotechnology. *Anal. Chem.* 63:2802-2824.
- Carr, S.A., Huddleston, M.J., and Bean, M.F. 1993. Selective identification and differentiation of N- and O-linked oligosaccharides in glycoproteins by liquid chromatography-mass spectrometry. *Protein. Sci.* 2:183-196.
- Chait, B.T. and Kent, S.B. 1992. Weighing naked proteins: Practical, high-accuracy mass measurement of peptides and proteins. *Science* 257:1885-1893.
- Chait, B.T., Wang, R., Beavis, R.C., and Kent, S.B. 1993. Protein ladder sequencing. *Science* 262:89-92.
- Cheng, X., Chen, R., Bruce, J.E., Schwartz, B.L., Anderson, G.A., Hofstadler, S.A., Gale, D.C., and Smith, R.D. 1995. Using electrospray ionization FTICR mass spectrometry to study competitive binding of inhibitors to carbonic anhydrase. *J. Am. Chem. Soc.* 117:8859-8860.
- Cohen, S.L. and Chait, B.T. 1996. Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. *Anal. Chem.* 68:31-37.
- Coligan, J.E., Dunn, B.M., Ploegh, H.L., Speicher, D.W., and Wingfield, P.T. (eds.) 1997. Current Protocols in Protein Science. John Wiley & Sons, New York.
- Covey, T.R., Bonner, R.F., Shushan, B.I., and Henion, J. 1988. The determination of protein, oligonucleotide and peptide molecular weights by ion-spray mass spectrometry. *Rapid Commun. Mass Spectrom.* 2:249-256.
- Dreisewerd, K., Schurenberg, M., Karas, M., and Hillenkamp, F. 1995. Influence of the laser intensity and spot size on the desorption of molecules and ions in matrix-assisted laser desorption/ionization with a uniform beam profile. *Int. J. Mass Spectrom. Ion Process.* 141:127-148.
- Dunayevskiy, Y., Vouros, P., Carell, T., Wintner, E.A., and Rebek, J., Jr. 1995. Characterization of the complexity of small-molecule libraries by electrospray ionization mass spectrometry. *Anal. Chem.* 67:2906-2915.
- Eng, J.K., McCormack, A.L., and Yates, R.J. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass. Spectrom.* 5:976-989.
- Englander, S.W. 1993. In pursuit of protein folding. *Science* 262:848-849.

- Falick, A.M., Hines, W.M., Medzihradzky, K.F., Baldwin, M.A., and Gibson, B.W. 1993. Low-mass ions produced from peptides by high-energy collision-induced dissociation in tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* 4:882-893.
- Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. 1990. Electrospray ionization—Principles and practice. *Mass Spectrom. Rev.* 9:37-70.
- Fernandez, J. and Mische, S. 1995. Enzymatic digestion of proteins on PVDF membranes. In *Current Protocols in Protein Science*. (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 11.2.1-11.2.10. John Wiley & Sons, New York.
- Foret, F., Thompson, T.J., Vouros, P., Karger, B.L., Gebauer, P., and Bocek, P. 1994. Liquid sheath effects on the separation of proteins in capillary electrophoresis/electrospray mass spectrometry. *Anal. Chem.* 66:4450-4458.
- Gulden, P.H., Hackett, M., Addona, T.A., Guo, L., Walker, C.B., Sherman, N.E., Shabanowitz, J., Hewlett, E.L., and Hunt, D.F. 1996. Mass spectrometric methods for peptide sequencing: Applications to immunology and protein acylation. In *Mass Spectrometry in the Biological Sciences* (A.L. Burlingame and S.A. Carr, eds.) pp. 281-305. Humana Press, Totowa, N.J.
- Hall, S.C., Smith, D.M., Clauser, K.R., Andrews, L.E., Walls, F.C., Webb, J.W., Tran, H.M., Epstein, L.B., and Burlingame, A.L. 1996. Mass spectrometric identification of proteins isolated by two-dimensional gel electrophoresis. In *Mass Spectrometry in the Biological Sciences* (A.L. Burlingame and S.A. Carr, eds.) pp. 171-202. Humana Press, Totowa, N.J.
- Hemling, M.E., Mentzer, M.A., Capiou, C., and Carr, S.A. 1996. A multifaceted strategy for the characterization of recombinant gD-2, a potential herpes vaccine. In *Mass Spectrometry in the Biological Sciences* (A.L. Burlingame and S.A. Carr, eds.) pp. 307-331. Humana Press, Totowa, N.J.
- Hensley, P., McDevitt, P.J., Brooks, I., Trill, J.J., Feild, J.A., McNulty, D.E., Connor, J.R., Griswold, D.E., Kumar, N.V., Koppke, K.D., Carr, S.A., Dalton, B.J., and Johanson, K. 1994. The soluble form of E-selectin is an asymmetric monomer. *J. Biol. Chem.* 269:23949-23958.
- Henzel, W.J., Billeci, T.M., Stults, J.T., Wong, S.C., Grimley, C., and Watanabe, C. 1993. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc. Natl. Acad. Sci. U.S.A.* 90:5011-5015.
- Hess, D., Covey, T.C., Winz, R., Brownsey, R.W., and Aebersold, R. 1993. Analytical and micro-preparative peptide mapping by high performance liquid chromatography/electrospray mass spectrometry of proteins purified by gel electrophoresis. *Protein Sci.* 2:1342-1351.
- Hillenkamp, F., Karas, M., Beavis, R.C., and Chait, B.T. 1991. Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal. Chem.* 60:2299-2301.
- Houthaeve, T., Gausepohl, H., Mann, M., and Ashman, K. 1995. Automation of micro-preparation and enzymatic cleavage of gel electrophoretically separated proteins. *FEBS Lett.* 376:91-94.
- Huddleston, M.J., Bean, M.F., and Carr, S.A. 1993a. Collisional fragmentation of glycopeptides by electrospray ionization LC/MS and LC/MS/MS: Methods for selective detection of glycopeptides in protein digests. *Anal. Chem.* 66:87-884.
- Huddleston, M.J., Annan, R.S., Bean, M.F., and Carr, S.A. 1993b. Selective detection of phosphopeptides in mixtures by electrospray LC-MS. *J. Am. Soc. Mass Spectrom.* 4:710-717.
- Hunt, D.F., Yates, J.R., Shabanowitz, J., Winston, S., and Hauer, C.R. 1986. Protein sequencing by tandem mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* 83:6233-6237.
- Hunt, D.F., Henderson, R.A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., Cox, A.L., Appella, E., and Engelhard, V.H. 1992. Characterization of peptides bound to class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* 255:1261-1266.
- Ingendoh, A., Karas, M., Hillenkamp, F., and Giessmann, U. 1994. Factors affecting the resolution in MALDI mass spectrometry. *Int. J. Mass Spectrom. Ion Process.* 131:345-354.
- Jeno, P., Scherer, P.E., Manning-Krieg, U., and Horst, M. 1993. Desalting electroeluted proteins with hydrophilic interaction chromatography. *Anal. Biochem.* 215:292-298.
- Jespersen, S., Niessen, W.M.A., Tjaden, U.R., and van der Greef, J. 1995. Quantitative bioanalysis using matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* 30:357-364.
- Kahn, P. 1995. From genome to proteome: Looking at a cell's proteins. *Science* 270:369-370.
- Karas, M. and Hillenkamp, F. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* 60:2299-2301.
- Karas, M., Bahr, U., and Hillenkamp, F. 1989. UV laser matrix desorption/ionization mass spectrometry of proteins in the 100,000 dalton range. *Int. J. Mass Spectrom. Ion Process.* 92:231-242.
- Kassel, D.B., Shushan, B., Sakuma, T., and Salzmann, J.-P. 1994. Evaluation of packed capillary perfusion column HPLC/MS/MS for the rapid mapping and sequencing of enzymatic digests. *Anal. Chem.* 66:236-243.
- Katta, V., Chowdhury, S.K., and Chait, B.T. 1991. Use of a single-quadrupole mass spectrometer for collision-induced dissociation studies of multiply charged peptide ions produced by electrospray ionization. *Anal. Chem.* 63:174-178.
- Kaufmann, R., Kirsch, D., and Spengler, B. 1994. Sequencing of peptides in a time-of-flight mass spectrometer: Evaluation of post-source decay following matrix-assisted laser desorption ionization.

- zation (MALDI). *Int. J. Mass Spectrom. Ion. Process.* 131:355-385.
- Kebarle, P. and Tang, L. 1993. From ions in solution to ions in the gas phase. *Anal. Chem.* 65:972A-986A.
- King, T.B., Colby, S.M., and Reilly, J.P. 1995. High resolution MALDI-TOF mass spectra of three proteins obtained using space-velocity correlation focusing. *Int. J. Mass Spectrom. Ion Process.* 145:L1-L7.
- Korner, R., Wilm, M., Morand, K., Schubert, M., and Mann, M. 1996. Nano electrospray combined with a quadrupole ion trap for the analysis of peptides and protein digests. *J. Am. Soc. Mass Spectrom.* 7:150-156.
- Ladner, R.D., McNulty, D.E., Carr, S.A., Roberts, G.D., and Caradonna, S.J. 1996. Characterization of distinct nuclear and mitochondrial forms of human dUTPase. *J. Biol. Chem.* 271:7745-7751.
- Lee, H. and Lubman, D.M. 1995. Sequence-specific fragmentation generated by matrix-assisted laser desorption/ionization in a quadrupole ion trap/reflectron time-of-flight device. *Anal. Chem.* 67:1400-1408.
- Lide, D.R. (ed.). 1995. CRC Handbook of Chemistry and Physics, 75th ed. CRC Press, Boca Raton, Fla.
- Mann, M. and Talbo, G. 1996. Developments in matrix-assisted laser desorption/ionization peptide mass spectrometry. *Curr. Opin. Biotechnol.* 7:11-19.
- Mann, M. and Wilm, M. 1995. Electrospray mass spectrometry for protein characterization. *Trends Biochem. Sci.* 20:219-224.
- Mann, M., Meng, C.K., and Fenn, J.B. 1989. Interpreting mass spectra of multiply charged ions. *Anal. Chem.* 61:1702-1708.
- Mann, M., Hojrup, P., and Roepstorff, P. 1993. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biol. Mass Spectrom.* 22:338-345.
- McCloskey, J.A. (ed.) 1990. Mass Spectrometry. *Methods Enzymol.*, Vol. 193.
- Medzihradzky, K.F., Adams, G.W., Burlingame, A.L., Bateman, R.H., and Green, M.R. 1996. Peptide sequence determination by matrix-assisted laser desorption ionization employing a tandem double focusing magnetic-orthogonal acceleration time-of-flight mass spectrometer. *J. Am. Soc. Mass Spectrom.* 7:1-10.
- Mirgorodskaya, O.A., Shevchenko, A.A., Chernushevich, I.V., Dodonov, A.F., and Miroshnikov, A.I. 1994. Electrospray ionization time-of-flight mass spectrometry in protein chemistry. *Anal. Chem.* 66:99-107.
- O'Connor, P.B., Speir, J.P., Senko, M.W., Little, D.P., and McLafferty, F.W. 1995. Tandem mass spectrometry of carbonic anhydrase (29 kDa). *J. Mass Spectrom.* 30:88-93.
- Ogorzalek Loo, R.R., Dales, N., and Andrews, P.C. 1994. Surfactant effects on protein structure examined by electrospray ionization mass spectrometry. *Protein Sci.* 3:1975-1983.
- Palczewski, K., Buczylo, J., Ohguro, H., Annan, R.S., Carr, S.A., Crabb, J.W., Kaplan, M.W., Johnson, R.S., and Walsh, K.A. 1994. Characterization of a truncated form of arrestin isolated from bovine rod outer segments. *Protein Sci.* 3:314-324.
- Pappin, D.J.C., Hojrup, P., and Bleasby, A.J. 1993. Rapid identification of proteins by peptide-mass fingerprinting. *Curr. Biol.* 3:327-332.
- Patterson, S.D. 1995. Matrix-assisted laser-desorption/ionization mass spectrometric approaches for the identification of gel-separated proteins in the 5-50 pmol range. *Electrophoresis* 16:1104-1114.
- Patterson, S.D. and Katta, V. 1994. Prompt fragmentation of disulfide-linked peptides during matrix-assisted laser desorption ionization mass spectrometry. *Anal. Chem.* 66:3727-3732.
- Patterson, D.H., Tarr, G.E., Regnier, F.E., and Martin, S.A. 1995. C-terminal ladder sequencing via matrix-assisted laser desorption mass spectrometry coupled with carboxypeptidase Y time-dependent and concentration-dependent digestions. *Anal. Chem.* 67:3971-3978.
- Qin, J. and Chait, B. 1995. Preferential fragmentation of protonated gas-phase peptide ions adjacent to acidic residues. *J. Am. Chem. Soc.* 117:5411-5412.
- Reinhold, B.B. and Reinhold, V.N. 1992. Electrospray ionization mass spectrometry: Deconvolution by an entropy-based algorithm. *J. Am. Soc. Mass Spectrom.* 3:207-215.
- Resing, K.A., Mansour, S.J., Hermann, A.S., Johnson, R.S., Candia, J.M., Fukasawa, K., Vande Woude, G.F., and Ahn, N.G. 1995. Determination of v-Mos-catalyzed phosphorylation sites and autophosphorylation sites on MAP kinase kinase by ESI-MS. *Biochemistry* 34:2610-2620.
- Roberts, G.D., Johnson, W.P., Burman, S., Anumula, K.R., and Carr, S.A. 1995. An integrated strategy for structural characterization of the protein and carbohydrate components of monoclonal antibodies: Application to anti-respiratory syncytial virus Mab. *Anal. Chem.* 67:3613-3625.
- Rush, R.S., Derby, P.L., Smith, D.M., Merry, C., Rogers, G., Rohde, M.F., and Katta, V. 1995. Microheterogeneity of erythropoietin carbohydrate structure. *Anal. Chem.* 67:1442-1452.
- Schindler, P.A., Settineri, C.A., Collet, X., Fielding, C.J., and Burlingame, A.L. 1995. Site-specific detection and structural characterization of the glycosylation of human plasma proteins lecithin:cholesterol acyltransferase and apolipoprotein D using HPLC/electrospray mass spectrometry and sequential glycosidase digestion. *Protein Sci.* 4:791-803.
- Sherman, N.E., Yates, N.A., Shabanowitz, J., Hunt, D.F., Jeffery, W., Jones, M.B., and Pappin, D.J. 1995. A novel N-terminal derivative designed to simplify peptide fragmentation. *In Proceedings*

- of the 43rd ASMS Conference on Mass Spectrometry and Allied Topics, Atlanta, Ga., p. 626.
- Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. 1996. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal. Chem.* 68:850-858.
- Smith, R.D. and Light-Wahl, K.J. 1993. The observation of noncovalent interactions in solution by electrospray ionization mass spectrometry: Promise, pitfalls and prognosis. *Biol. Mass Spectrom.* 22:493-501.
- Smith, D.L. and Zhang, Z. 1994. Probing noncovalent structural features of proteins by mass spectrometry. *Mass Spectrom. Rev.* 13:411-429.
- Smith, R.D., Loo, J.A., Ogorzalek Loo, R.R., Busman, M., and Udseth, H.R. 1991. Principles and practice of electrospray ionization-mass spectrometry for large polypeptides and proteins. *Mass Spectrom. Rev.* 10:359-451.
- Smith, R.D., Bruce, J.E., Wu, Q., Cheng, X., Hofstadler, S.A., Anderson, G.A., Chen, R., Bakhtiar, R., Van Orden, S.O., Gale, D.C., Sherman, M.G., Rockwood, A.L., and Udseth, H.R. 1996. The role of Fourier transform ion cyclotron resonance mass spectrometry in biological research—New developments and applications. *In* Mass Spectrometry in the Biological Sciences (A.L. Burlingame and S.A. Carr, eds.) pp. 25-68. Humana Press, Totowa, N.J.
- Speir, J.P., Senko, M.W., Little, D.P., Loo, J.A., and McLafferty, F.W. 1995. High-resolution tandem mass spectra of 37-67 kDa proteins. *J. Mass Spectrom.* 30:39-42.
- Stone, K.L. and Williams, K.R. 1995. Digestion of proteins in gels for sequence analysis. *In* Current Protocols in Protein Science. (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 11.3.1-11.3.13. John Wiley & Sons, New York.
- Stoney, K. and Nugent, K. 1995. Online preparation of complex biological samples prior to analysis by HPLC, LC/MS and/or protein sequencing. *In* Techniques in Protein Chemistry IV (J.W. Crabb, ed.) pp. 277-284. Academic Press, San Diego.
- Strupat, K., Karas, M., Hillenkamp, F., Eckerskorn, C., and Lottspeich, F. 1994. Matrix-assisted laser desorption ionization mass spectrometry of proteins electroblotted after polyacrylamide gel electrophoresis. *Anal. Chem.* 66:464-470.
- Stults, J.T., O'Connell, K.L., Garcia, C., Wong, S., Engel, A.M., Garbers, D.L., and Lowe, D.G. 1994. The disulfide linkages and glycosylation sites of the human natriuretic peptide receptor-C homodimer. *Biochemistry* 33:11372-11381.
- Thiede, B., Wittmann-Liebold, B., Bienert, M., and Krause, E. 1995. MALDI-MS for C-terminal sequence determination of peptides and proteins degraded by carboxypeptidase Y and P. *FEBS Lett.* 357:65-69.
- Till, J.H., Annan, R.S., Carr, S.A., and Miller, W.T. 1994. Use of synthetic peptide libraries and phosphopeptide-selective mass spectrometry to probe protein kinase substrate specificity. *J. Biol. Chem.* 269:7423-7428.
- Uy, R. and Wold, F. 1977. Posttranslational covalent modification of proteins. *Science* 198:890-896.
- Verentchikov, A.N., Ens, W., and Standing, K.G. 1994. Reflecting time-of-flight mass spectrometer with an electrospray ion source and orthogonal extraction. *Anal. Chem.* 66:126-133.
- Vestal, M.L., Juhasz, P., and Martin, S.A. 1995. Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 9:1044-1050.
- Vorm, O. and Mann, M. 1994. Improved mass accuracy in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of peptides. *J. Am. Soc. Mass Spectrom.* 5:955-958.
- Wagner, D.S. and Anderegg, R.J. 1994. Conformation of cytochrome c studied by deuterium exchange-electrospray ionization mass spectrometry. *Anal. Chem.* 66:706-711.
- Wahl, J.H., Goodlett, D.R., Udseth, H.R., and Smith, R.D. 1992. Attomole-level capillary electrophoresis/electrospray ionization mass spectrometric protein analysis using 5 μ m i.d. capillaries. *Anal. Chem.* 64:3194-3196.
- Wang, R. and Chait, B.T. 1994. High-accuracy mass measurement as a tool for studying proteins. *Curr. Opin. Biotechnol.* 5:77-84.
- Wapstra, A.H. and Audi, G. 1985. The 1983 atomic mass revolution. I. Atomic mass table. *Nucl. Phys.* A432:1-54.
- Watts, J.D., Affolter, M., Krebs, D.L., Wange, R.L., Samelson, L.E., and Aebersold, R. 1994. Identification by electrospray ionization mass spectrometry of the sites of tyrosine phosphorylation induced in activated Jurkat T cells on the protein tyrosine kinase ZAP-70. *J. Biol. Chem.* 25:29520-29529.
- Whittall, R.M. and Li, L. 1995. High resolution matrix-assisted laser desorption/ionization in a linear time-of-flight mass spectrometer. *Anal. Chem.* 67:1950-1954.
- Wiley, W.C. and McLaren, I.H. 1953. Time-of-flight mass spectrometer with improved resolution. *Rev. Sci. Instrum.* 26:1150-1157.
- Wilkins, M.R., Sanchez, J.-C., Gooley, A.A., Appel, R.D., Humphery-Smith, I., Hochstrasser, D.F., and Williams, K.L. 1995. Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. *Biotechnol. Genet. Eng. Rev.* 13:19-50.
- Williams, K.R. and Carr, S.A. 1995. Protein analysis by integrated sample preparation, chemistry, and mass spectrometry. *In* Molecular Biological Biotechnology: A comprehensive desk reference (R.A. Meyers, ed.) pp. 731-737. VCH Publishers, New York.
- Williams, K.R. and Stone, K.L. 1995. In gel digestion of SDS PAGE-separated proteins: Observations from internal sequencing of 25 proteins. *In* Techniques in Protein Chemistry VI (J. Crabb, ed.) pp. 143-152. Academic Press, San Diego.

- Wilm, M. and Mann, M. 1994. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal. Chem.* 66:4390-4399.
- Wilm, M. and Mann, M. 1996. Analytical properties of nano-electrospray ion source. *Anal. Chem.* 68:1-8.
- Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. 1996. Femtomole sequencing of proteins from acrylamide gels by nano-electrospray mass spectrometry. *Nature* 379:466-469.
- Winger, B.E., Hofstadler, S.A., Bruce, J.E., Udseth, H.R., and Smith, R.D. 1993. High-resolution accurate mass measurements of biomolecules using a new electrospray ionization ion cyclotron resonance mass spectrometer. *J. Am. Soc. Mass Spectrom.* 4:566-577.
- Wold, F. 1981. In vivo chemical modification of proteins (post-translational modification). *Annu. Rev. Biochem.* 50:783-814.
- Wold, F. and Moldave, K. (eds.) 1984. Posttranslational modifications, Parts A and B. *Methods Enzymol.*, Vols. 106 and 107.
- Wu, Q., Van Orden, S., Cheng, X., Bakhtiar, R., and Smith, R.D. 1995. Characterization of cytochrome c variants with high-resolution FTICR mass spectrometry: Correlation of fragmentation and structure. *Anal. Chem.* 67:2498-2509.
- Yergey, J.A. 1983. A general approach to calculating isotopic distributions for mass spectrometry. *Int. J. Mass Spectrom. Ion Process.* 52:337-349.
- Yergey, J.A., Heller, Hansen, D.G., Cotter, R.J., and Fenselau, C. 1983. Isotopic distributions in mass spectra of large molecules. *Anal. Chem.* 55:353-356.
- Yost, R.A. and Boyd, R.K. 1990. Tandem mass spectrometry: Quadrupole and hybrid instruments. *Methods Enzymol.* 193:154-200.
- Zhang, H., Andren, P.E., Caprioli, R.M., and Annan, R. 1995. Micro-preparation procedure for high-sensitivity matrix-assisted laser desorption ionization mass spectrometry. *J. Mass. Spectrom.* 30:1768-1771.
- Zhou, J., Ens, W., Poppe-Schriemer, N., Standing, K.G., and Westmore, J.B. 1993. Cleavage of interchain disulfide bonds following matrix-assisted laser desorption. *Int. J. Mass Spectrom. Ion Process.* 126:115-122.
- Zubarev, R.A., Demirev, P.A., Hakansson, P., and Sundqvist, B.U.R. 1995. Approaches and limits for accurate mass characterization of large biomolecules. *Anal. Chem.* 67:3793-3798.

Contributed by Steven A. Carr
and Roland S. Annan
SmithKline Beecham Pharmaceuticals
King of Prussia, Pennsylvania

Protein Identification and Characterization by Mass Spectrometry

UNIT 10.22

Many aspects of science are driven by technological advances. The pace of molecular biology research continues to increase through such advances, many of which are described in *Current Protocols in Molecular Biology*. However, in the last decade significant technological advances in many aspects of protein chemistry have been realized. Most of these advances have been driven by new ionization capabilities of mass spectrometers, making the analysis of biological materials feasible, and most importantly, by a wider range of scientists. Together with the massive increase in the amount of nucleotide sequence data that is publicly available, this has led to the development of new methods for protein identification using mass spectrometry-derived data (reviewed in Patterson and Aebersold, 1995; and Pennington et al., 1997). In addition, protein identification by mass spectrometry-based methods is being adapted for massively parallel analysis on the one hand or very rapid serial analysis on the other, thereby resembling some of the extremely powerful molecular biological approaches currently available.

As we enter a post-genome era for many organisms, including ultimately *Homo sapiens*, these technologies will be applied to high-throughput approaches as well as in more traditional settings (Patterson, 1997). Such approaches will facilitate analyses of many of the proteins expressed by a given tissue or cell, now being referred to as proteome analysis. The term proteome refers to the *protein* complement expressed by a *genome* (Wilkins et al., 1996) and is often applied to two-dimensional gel electrophoresis-separated proteins and their subsequent analysis (James, 1997; Pennington et al., 1997). With so much nucleotide sequence information at hand in the form of full-length cloned genes and expressed sequence tags (ESTs) for a few species, the need for traditional de novo amino acid sequencing for these genetically well-characterized species is rapidly diminishing. A new era of protein identification by correlation of primary structural data with these databases has begun.

This overview will describe some of the new technologies that can be employed to facilitate rapid identification and characterization of proteins, including the use of correlative approaches for protein identification, rapid post-

translational modification analysis, identification of components in complex mixtures, and direct mass analysis of gel-separated proteins. Other fields of protein chemistry that have been impacted by the advances in mass spectrometry include analysis of protein:protein interactions of both purified proteins and ligands in solution/gas phase (Chapter 20; Robinson et al., 1994; Loo, 1997), and direct screening of specific analytes in natural fluids through the use of affinity surfaces (Nelson et al., 1995). Determination of domain structure through accurate mass analysis after limited proteolysis has also been used as an adjunct to X-ray crystallographic studies, including designing molecules of appropriate size for crystallization (Cohen, 1996). However, as it is not possible within this overview to cite all of the relevant literature, only a limited number of examples of each approach will be given, and where relevant, more extensive reviews will be cited.

The mass spectrometric methods that will be referred to in this overview include matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) and electrospray ionization mass spectrometry (ESI-MS). This overview will not include a detailed description of the technical side of mass spectrometry, as this is well covered in *UNIT 10.21*, and a discussion of its use for analysis of biological molecules appears in Chapter 16 of *Current Protocols in Protein Science* (Coligan et al., 1997). However, a brief description of the two types of mass spectrometers utilized in the studies mentioned in this overview follows.

Mass spectrometers can be thought of simply as having two main components, an ionization source and a mass analyzer. Charged molecules are generated by an ionization source, e.g., MALDI or ESI, into the gas phase of a mass analyzer which is held under vacuum. MALDI sources are usually coupled with time-of-flight (TOF) mass analyzers. In MALDI-MS, the sample (analyte) is cocrystallized in a matrix that absorbs at the wavelength of the laser—either an ultraviolet (UV) laser at 337 nm, or an infrared (IR) laser at 2.94 μm . Ionization occurs in a static electric accelerating field that is generated by holding the sample at high potential (e.g., 20 to 30 kV) with respect to a closely spaced grounded electrode. The accelerated analyte ions pass through an orifice

Contributed by Scott D. Patterson

Current Protocols in Molecular Biology (1998) 10.22.1-10.22.24

Copyright © 1998 by John Wiley & Sons, Inc.

Analysis of
Proteins

10.22.1

Supplement 41

in the grounded electrode (referred to as continuous extraction) into a field-free drift region of 50 to 300 cm, and signals are recorded when they strike a detector at the end of the flight tube. All of the ions are accelerated by the same potential difference, hence their velocity is proportional to their mass over charge (m/z), i.e., smaller ions arrive at the detector sooner than larger ions, and all of the ions generated in each ionization event are measured. Analytes of known mass can be used to calibrate the instrument. Depending upon the matrix used to cocrystallize the sample, the analytes are usually singly charged, i.e., only one proton has been added to the analyte. Mass over charge ratios of up to ~500,000 have been measured with MALDI-MS. In addition to continuous extraction, another form of analyte extraction from the MALDI probe has recently been introduced onto commercial MALDI-MS instruments and is referred to as delayed extraction (time-lag focusing; see *UNIT 10.21*). This allows the extracted ions time to focus prior to an acceleration of the gas-phase ions into the field-free drift tube resulting in improved sensitivity and increased mass accuracy by reducing the spread of flight times of ions of the same mass (e.g., due to surface nonhomogeneity; Vestal et al., 1995).

In ESI-MS, the analytes are introduced into the gas phase by passing them through a fine needle at high potential at atmospheric pressure (e.g., +5 kV, for positive-ion spectra) resulting in the formation of a fine spray of highly-charged droplets. This spray is directed across the small inlet orifice (held at a low potential, e.g., +50 V) to the evacuated quadrupole mass analyzer. Between the spray needle and the orifice, the droplets are desolvated by directed flows of dry inert gas and/or heat. These dry ions can now be mass analyzed in the quadrupole device. Due to this process, most analytes entering the mass analyzer are multiply charged, i.e., more than one proton has been added to the analyte during the ionization process; peptides, for example, are often only doubly (mass of peptide + 2 protons divided by 2 = m/z value) or triply charged, but proteins can carry many protons. Quadrupole mass analyzers are scanning instruments, unlike the TOF mass analyzer, where the masses of all of the analytes are measured following each ionization event. At any one point in time, only ions of one particular m/z value are passed through the quadrupoles; all other ions are lost for further analysis. A detector counts the number of ions transmitted at each specific m/z value,

and by scanning a mass range, a spectrum of m/z versus abundance (a mass spectrum) can be plotted. The mass range of quadrupole instruments is limited compared with TOF analyzers, usually up to 2000 to 4000 m/z , but as stated above, analytes ionized by ESI are multiply charged, and hence their m/z values are measurable in these instruments even when their masses are considerably lower than 2000 u (atomic mass units, equivalent to Da). In addition, in some mass spectrometers designed for that purpose, intact molecules (peptides usually) can be caused to break apart (i.e., fragment in a sequence-specific manner), and the masses of the resulting pieces can be measured (see Protein Identification Using Mass Spectrometry-Based Correlative Approaches and *UNIT 10.21*). For instruments that are appropriately equipped, this is referred to as collisional-induced dissociation (CID), or tandem MS (MS/MS) for ESI-MS, and as post-source decay (PSD) for MALDI-MS (see *UNIT 10.21*). The fragment ions produced are predominantly formed via fragmentation through the peptide bonds. The mass difference between fragment ions of the same series (either N- or C-terminal) is the residue mass of the amino acid at that position. Thus, the fragment-ion spectrum can be correlated with the peptide amino acid sequence.

Most simply put, mass spectrometers are very accurate balances providing an absolute physical measurement of analytes including proteins, peptides, and following fragmentation in the gas phase, sequence-specific fragment-ion spectra of peptides selected in the mass spectrometer.

PROTEIN IDENTIFICATION USING MS-BASED CORRELATIVE APPROACHES

Peptide-Mass Searching

Definitive protein identification for many years has been achieved by generating an amino acid sequence through chemical degradation of the N-terminus of the isolated protein (Edman degradation), or peptides derived from it (*UNIT 10.19*). Proteins have also been identified by immunodetection with antibodies recognizing the protein of interest (*UNIT 10.8*). A few years after it was practical to accurately measure the mass of peptides generated from gel-separated proteins, several groups published variations on the following theme as a means to identify proteins already resident in a sequence database: use an algorithm to match a set of peptide

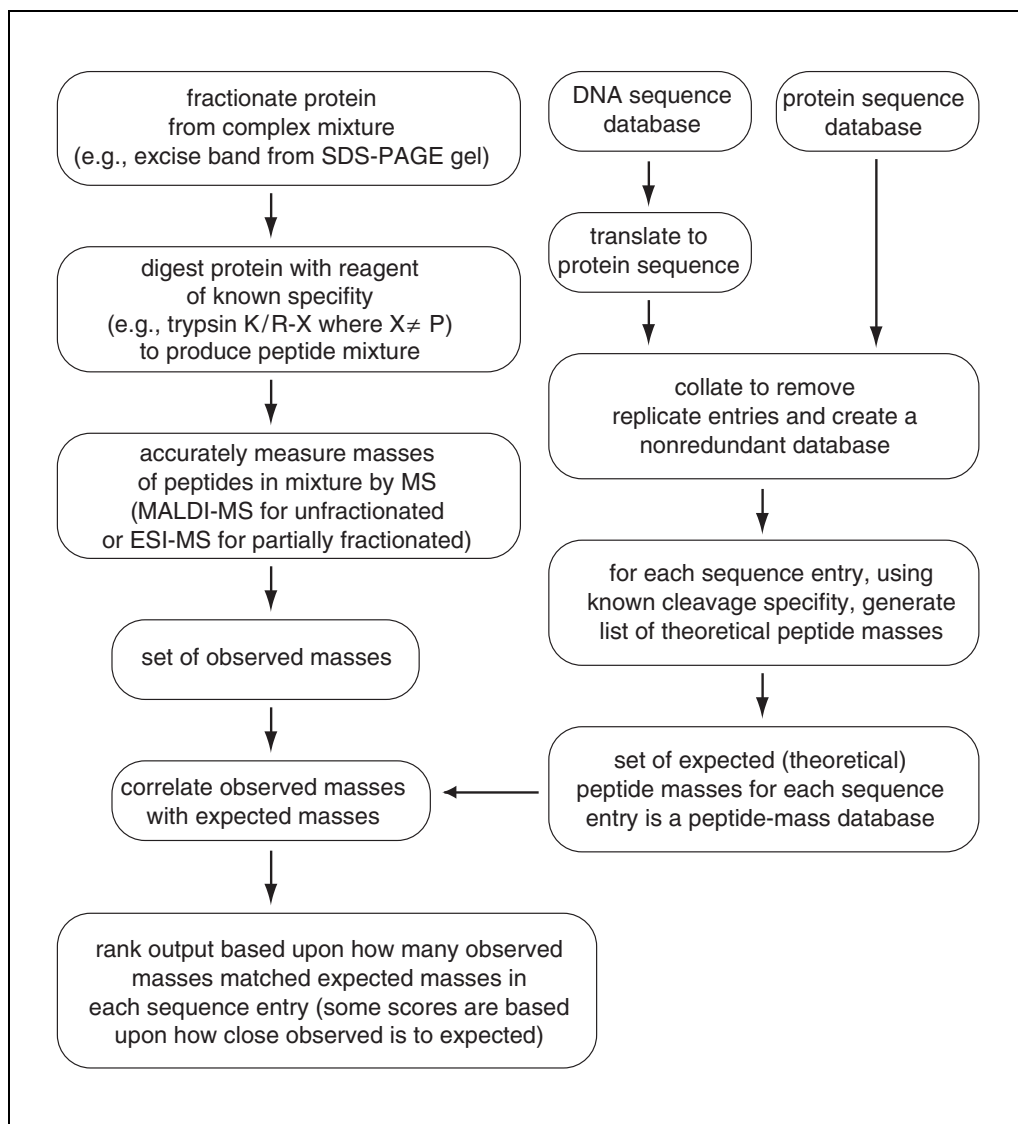


Figure 10.22.1 Protein identification using peptide-mass data generated by accurate mass measurement by either ESI-MS or MALDI-MS

masses generated in a specific manner (e.g., by enzymatic digestion with trypsin) from the protein of interest, with theoretical peptide masses calculated from each sequence entry in the database using the same cleavage specificity as the reagent in the experiment. Then derive a ranking (or score) to provide a measure of the fit between the observed and expected peptide masses (see Fig. 10.22.1; reviewed in Cottrell, 1994; Patterson, 1994; Patterson and Aebersold, 1995). Many of these original groups and others have placed their search programs on the World Wide Web for public access, and many include examples for their use (see Table 10.22.1).

The obvious caveat to such an analysis is that the protein to be identified already exists within a sequence database, either translated

nucleotide or protein sequence database. Therefore, this approach is ideal when the protein to be identified is derived from an organism whose genome is sequenced or at least whose major proteins are sequenced. Of course, such searches are not suited to translations of expressed sequence tag (EST) databases which consist of only limited coding regions and therefore, may not contain sufficient information for a match. Peptide-mass data is most rapidly obtained by MALDI-MS analysis of an aliquot of a protein digest. However, peptide-mass data can also be obtained from liquid chromatography (LC)-MS analysis of a protein digest.

It is rare that all of the input peptide masses match with the highest ranking sequence entry, and it is important to determine the origin of

Table 10.22.1 Sites Providing Online Protein Identification using Mass Spectrometry–Derived Data

Resource	WWW Uniform Resource Locator (<i>http://...</i>)	Features and comments
Rockefeller University New York, N.Y.	<i>prowl.rockefeller.edu/PROWL/prot-id-main.html</i>	Both peptide-mass and fragment-ion search programs
ETH-Zurich Zurich, Switzerland	<i>cbrg.inf.ethz.ch/subsection3_1_3.html</i>	Peptide-mass search program
University of California San Francisco, Calif.	<i>prospector.ucsf.edu</i>	Both peptide-mass (MS-Fit) and fragment-ion (MS-Tag) search programs
Max Delbrück Centre Berlin, Germany	<i>www.mdc-berlin.de/~emu/peptide_mass.html</i>	Peptide-mass search program
EMBL Heidelberg, Germany	<i>www.mann.embl-heidelberg.de/Services/PeptideSearch/PeptideSearchIntro.html</i>	Both peptide-mass and fragment-ion search programs
SEQNET Daresbury, UK	<i>www.dl.ac.uk/SEQNET/mowse.html</i>	Both peptide-mass and fragment-ion search programs

these masses (Patterson and Aebersold, 1995). The following should be considered as reasons for additional masses. (1) The correct protein was identified in the search, but the masses are the result of posttranslational or artifactual modification or posttranslational processing. Plausible modifications may be taken into consideration but should only be considered tentative unless confirmed experimentally. (2) The correct protein was identified, but some peptides were derived by nonspecific proteolysis or cleavage by a contaminating protease. This theory can be tested by determining whether the protein candidate can produce the peptide masses without any assumptions as to the cleavage specificity. (3) The correct protein was identified, but the “pure” protein was contaminated with one or more additional proteins. If there are enough additional masses, a separate peptide-mass search using these masses can be performed to attempt to identify the contaminant(s). (4) A sequence homologue, or processing variant, from either the same or a different species was identified. Some search programs allow species-specific searches to be performed, and that could eliminate this possibility; however, if confirmatory data is obtained, matches to proteins from other species can be useful, especially for scientists working with organisms whose genomes are relatively poorly characterized. (5) The match was a false positive! This potentially disastrous result is difficult to verify or disprove, particularly if the highest ranked protein did not have a high score; sometimes confidence in the highest-ranked score can be gained from the difference

between it and the second highest and subsequent scores; the protein may also not yet reside in the database and it may be truly novel.

Implicit to the theory of peptide-mass searching is the accurate measurement of peptide masses. Therefore, the greater the accuracy, the greater the confidence of the assignment. In addition, the reagent used to generate the peptides should also be one that shows a high degree of specificity. This latter point is limited to the specificity of enzymes and chemical cleavage reagents. Trypsin is the most commonly used reagent for peptide-mass searching analyses, but it can also cleave at sites other than C-terminal to Lys or Arg, if not followed by a Pro, and as with most enzymes it may not cleave a substrate to completion (i.e., missed cleavage sites). The number of missed cleavage sites is often a variable for peptide-mass search routines, but spurious cleavage sites for trypsin or other enzymes cannot be predicted. Endoproteinase LysC is another good enzyme used in such analysis, but it can also miss cleavage sites and cleave nonspecifically; this enzyme has the advantage that there are few if any autolysis products. Thus, there is little that can be done to improve the specificity of the reagents to cleave isolated proteins, but more accurate mass measurement is in large part dependent upon the type of mass spectrometer employed.

Most peptide-mass analyses are performed on MALDI-mass spectrometers due to the tolerance of the MALDI process to sample buffers and its rapid and simple analysis. This has increased further in the last couple of years

following implementation of delayed extraction which dramatically improves the mass accuracy obtainable on an ever-increasing range of commercial instruments (Vestal et al., 1995). The ability to measure masses at 5 ppm allows assignment of monoisotopic masses and the use of very tight tolerances in peptide-mass searches, thereby limiting the number of potential matches (Jensen et al., 1996a; Takach et al., 1997). Even though increasing the mass accuracy can't affect any of the situations mentioned above, there is less chance of a spurious match.

In addition to more accurately measuring the mass, a number of orthogonal methods can be performed on a separate aliquot of the peptide digest (or sometimes for MALDI-MS on the probe after the original data has been obtained) to provide information on the amino acid composition of each peptide, which can be used as an additional piece of information for the peptide-mass search algorithms. This is achieved through one of the following approaches: site-specific chemical modification [e.g., methyl esterification, which adds +14 u for each acidic residue—Asp, Glu, or C-terminus—in the peptide (Pappin et al., 1995) or iodination, which adds +126 u for each tyrosine (Craig et al., 1995)]; determination of partial amino acid composition of the peptide (hydrogen/deuterium exchange in which each amino acid can exchange a defined number of its hydrogens for the heavier deuterium, and the total mass increase reflects the peptide composition from 0 to 5 u per residue; James et al., 1994); one-step Edman degradation to determine the N-terminal amino acid of each peptide in the mixture (Jensen et al., 1996b); or an additional enzymatic digestion [subdigestion (Pappin et al., 1995), or parallel digestion (James et al., 1994)].

These approaches have proven quite successful for a number of investigators (reviewed in Patterson and Aebersold, 1995; Jungblut et al., 1996; Lamond and Mann, 1997). Most importantly, the data can be obtained very rapidly when using MALDI-MS of peptide mixtures, and some groups have implemented automated approaches so that large numbers of proteins can be analyzed in an unattended mode (see below). In one report, a peptide-mass search result provided sufficient information to be able to clone a homologous gene because the sequence of the matched peptides could be used to generate primers (Salmeron et al., 1996).

Peptide-Sequence Tags, Uninterpreted Fragment-Ion Searching, and de Novo Sequencing by MS/MS

The most discriminating criteria used to identify a protein is contiguous amino acid sequence information. The amino acid sequence information can be obtained either by chemical sequencing (automated Edman degradation) or by generating sequence-specific fragment ions from individual peptides in the gas-phase of a mass spectrometer (see *UNIT 10.21* for detailed description of fragmentation; see Figure 10.21.1 for fragmentation nomenclature). If sufficient contiguous amino acid sequence is obtained from either method, it is unnecessary to generate additional data to confirm the identification of the protein because standard string-based search protocols (e.g., TFASTA; see *UNIT 19.3*) can be employed. However, this is not always possible, and use of accurate mass measurement in conjunction with partial amino acid sequence information can be a very powerful approach.

Measurement of the masses of fractionated peptides (usually by reversed-phase-HPLC) prior to automated Edman sequencing provides the researcher with a number of options in addition to the amino acid sequence obtained. (1) Mass information can make it possible to determine more accurately the number of sequencing cycles required, thus minimizing reagent expense and sequencer time. (2) The peptide mass can confirm the determined peptide sequence and whether the sequence is complete (the sum of the masses of the identified amino acids should equal the mass of the peptide). (3) The mass analysis usually indicates the purity of the sample, thereby avoiding sequencing mixtures (although this is not always the case). (4) Given that proteins are cleaved with rather high enzyme-to-substrate ratios, knowledge of the mass of the peptide to be sequenced can avoid sequencing of an enzyme autolysis product. (5) An amino acid sequence can sometimes be completed if only a couple of amino acid residues have not been assignable, i.e., by summing the masses of the assigned residues and determining which residues could account for the mass difference between that and the mass of the intact peptide (Tempst et al., 1995). Thus, simple measurement of the mass of a peptide is an extremely useful exercise. Another scenario, often faced when sequencing low levels of peptides, is that only a few amino acid residues of a sequence are obtained, insufficient to identify the protein or to construct an oligonucleotide primer. If the

mass of the peptide has been determined, this limited amino acid sequence data may still be sufficient for further use (e.g., Patterson et al., 1996).

The partial amino acid sequence, which does not need to be contiguous, and the mass of the peptide can be used to construct a “peptide-sequence tag.” This term was coined by Mann and Wilm (1994) to reflect the power of this relatively small amount of information to identify the protein from which specific peptides were obtained, much the same way short stretches of nucleotide sequence as expressed sequence tags (ESTs) have proven so useful. By using the partial amino acid sequence, the cleavage specificity of the enzyme or other reagent used to generate the peptide, and the peptide mass, the peptide-sequence tag can be used to search amino acid sequence databases (including translations of EST databases). Mann and Wilm (1994) also described a method by which partial interpretation of sequence-specific MS/MS fragment-ion spectra (see UNIT 10.21) combined with the mass of the peptide from which the spectrum was obtained can be used to construct a peptide-sequence tag (see Fig. 10.22.2). In this case a portion of the fragment-ion spectrum is interpreted to determine whether an ion series (either *b* or *y*; see Fig. 10.21.1 in UNIT 10.21) can be established, i.e., a series of fragment ions whose mass differences correspond to known amino acid residues. The peptide-sequence tag is composed of three regions: the mass of the first ion in the series which provides the mass of the peptide fragment from an end of the peptide to the uninterpreted region; the interpreted amino acid sequence for anywhere from one to many residues; and the mass of the last ion in the series (the difference between this and the mass of the intact peptide is the mass of the uninterpreted region between the sequence and the other end of the peptide; Mann and Wilm, 1994; see Fig. 10.22.2). If the cleavage specificity of the enzyme is used, two additional pieces of information are then included in the search—the N-terminal and C-terminal specificity. These can also be left out if the search is conducted in error-tolerant mode which allows for either database sequence errors or posttranslational modifications.

Yates and colleagues have developed identification algorithms, referred to as SEQUEST, that are based upon automated interpretation of MS/MS fragment-ion spectra in a manner compatible with direct searching of sequence databases (Eng et al., 1994). In this approach, the

program first generates a list of theoretical peptide masses for each sequence entry, using either enzyme cleavage specificity rules or the sum of contiguous amino acid residues, and calculates which of these match the measured mass of the peptide ion (within a stated tolerance), thereby generating a list of candidate peptides. In the second step, the program then calculates the fragment ions expected for each of the candidate peptides if they were fragmented under the experimental conditions employed to generate a predicted spectrum; then it compares the experimentally determined MS/MS spectrum with the predicted spectra using cluster-analysis algorithms. Each comparison receives a score, and the highest-scoring peptide sequences are then reported. If a significant score is achieved, the protein is identified based on the peptide MS/MS spectrum without any explicit determination of the peptide sequence (see Fig. 10.22.2).

The advantages of such an approach include the following (Patterson and Aebersold, 1995). (1) Proteins can be effectively identified in complex mixtures because each peptide that generates an MS/MS spectrum provides data for an independent database search. (2) When applied to analysis of a digest from a single protein, the results are inherently autoconfirmatory, and with the same protein as a top-ranked candidate for a number of peptides, identification is absolutely confirmed. (3) The approach allows for easy automation when ions for fragmentation are selected in a data-dependent fashion, and the MS/MS results are automatically analyzed (Yates et al., 1995). (4) The method is able to be adapted to find peptides carrying specific posttranslational modifications (the program can be set to anticipate a modification of a specific mass on specific residues) as well as to identify the protein from which the peptide originated (Yates et al., 1995). A number of other investigators have developed algorithms for uninterpreted fragment-ion searches based upon matching of observed and expected fragment-ion masses, and a list of sites on the Internet that provide these search capabilities is provided in Table 10.22.1. One of these referred to as MS-Tag (Matsui et al., 1997), developed at the University of California, San Francisco, uses an algorithm to match the observed fragment-ions with a user-selected list of specific ion series; this is also shown in Figure 10.22.2.

Manual or computer-assisted interpretation of MS/MS spectra can be very difficult in many instances because it is impossible to determine

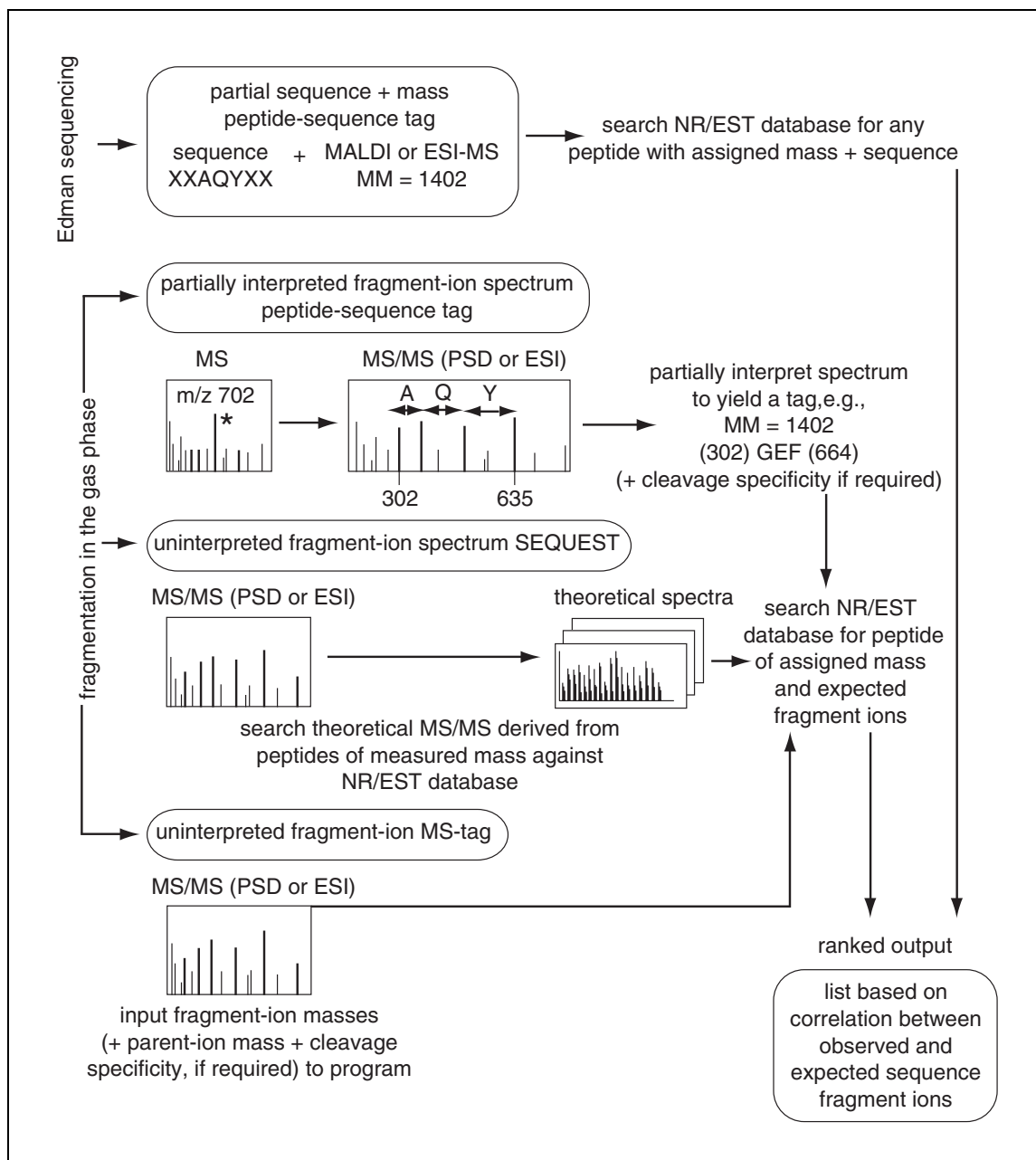


Figure 10.22.2 Protein identification using fragment-ion data, interpreted or not. For a peptide of interest the mass is accurately measured by MS and sequence information is generated by either Edman sequencing (if an incomplete sequence is obtained), fragmentation in the gas phase by ESI-MS/MS, or PSD-MALDI-MS to yield a fragment-ion spectrum.

which series-specific ions belong to N-terminal- (*b*) or C-terminal- (*y*) derived peptides; hence, the great utility of the methods described above for peptide sequence tagging and uninterpreted fragment-ion searching is for searching protein sequence and translated nucleotide sequence databases. However, for these approaches to be successful, the sequence that matches the peptide under analysis has to reside in the database. ESTs constitute only a portion of the sequence of many but not all expressed genes. Many widely used experimental organ-

isms are poorly represented in the sequence databases, and although conserved regions between proteins may be able to be matched (Cordwell et al., 1997), more often than not an accurate extended amino acid sequence is required so that oligonucleotide primers can be made to obtain the gene of interest from an appropriate library via the polymerase chain reaction (PCR; *UNIT 22.1*).

Mass spectrometry-based approaches are the methods of choice for generating sequence-based information from very low levels of pro-

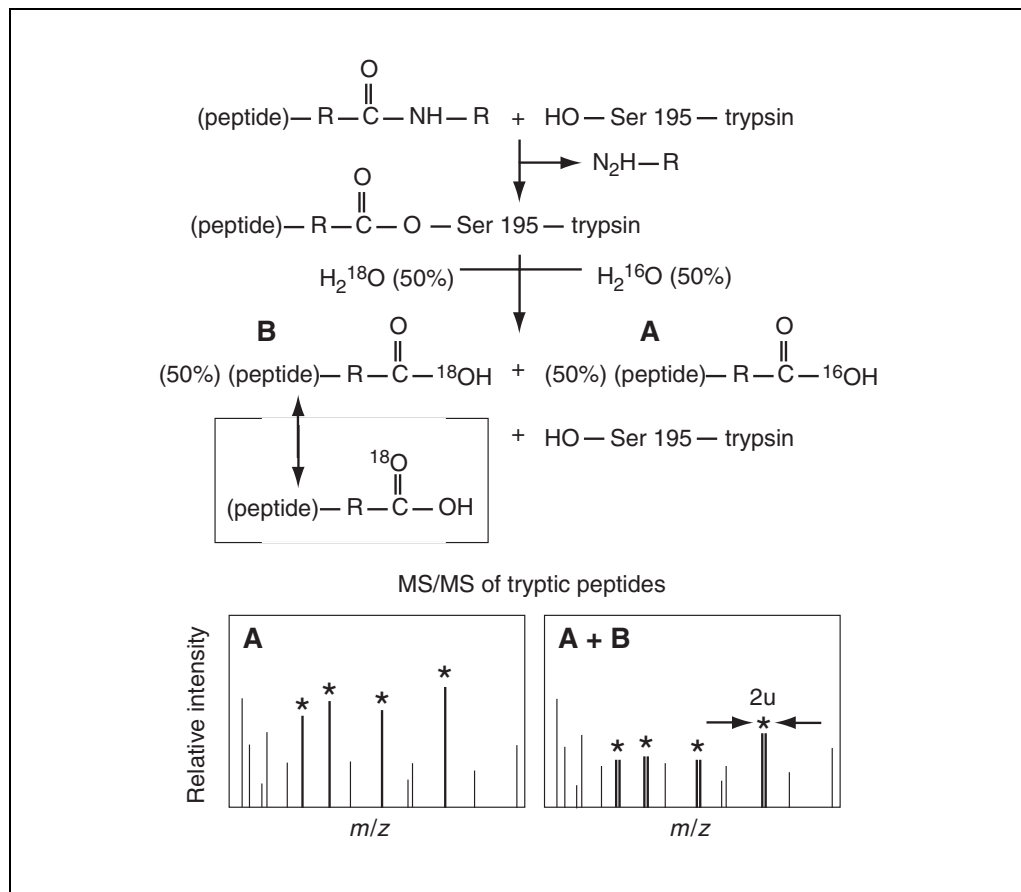


Figure 10.22.3 A scheme for de novo sequencing using H_2^{18}O . Tryptic digestion in the presence of 50% H_2^{18}O facilitates de novo sequencing by MS/MS by generating two sets of γ series ions (*) that differ in mass by 2 u (compare MS/MS of peptide A with the mixture of peptides A and B).

tein. Therefore, for proteins from species with limited sequence information, de novo sequencing from MS/MS spectra is important. This can be accomplished in two ways; (1) by generating two MS/MS spectra of the peptide with and without methyl esterification, and (2) by conducting tryptic digestion in the presence of 50% (v/v) H_2^{18}O (see Fig. 10.22.3). In the first method, a MS/MS spectrum is obtained after all of the carboxyl groups have been modified by esterification with a methyl group thereby increasing the mass of the peptide by 14 u for each carboxyl group present. This method is most useful when the peptide of interest does not contain any acidic residues, and it is of least use when either multiple acidic residues are present or the one acidic residue is located on the N-terminus. If there are no acidic residues, then comparison of the MS/MS spectra with and without esterification will reveal that all γ -series ions are 14 u heavier in the methyl-esterified spectrum and that all of the other ions are unaffected. If there are acidic residues in the peptide, the comparison becomes more complicated.

A more rapid approach for de novo sequence interpretation is afforded by conducting the original tryptic or LysC digestion in the presence of 50% (v/v) H_2^{18}O /50% (v/v) H_2^{16}O (Schnolzer et al., 1996). This results in all of the newly generated peptides but not the C-terminal peptide being present as doublets differing in mass by 2 u; it also reduces the intensity of the total signal for the intact species by half. This occurs because trypsin cleaves the peptide bond after lysine or arginine via the activated OH group of serine-195 (of trypsin), which forms a covalent ester bond with the carboxyl group of the lysine or arginine residue (the newly formed N-terminus is released). This intermediate is then cleaved by attack of an activated water molecule, leading to release of the peptide fragment, whose carboxyl group contains one oxygen from solvent water (Schnolzer et al., 1996). If the solvent water is 50% H_2^{18}O then half of the peptides will have a mass 2 u greater than those that incorporate ^{16}O from normal water. Therefore, when these peptides are fragmented, all of the C-terminal-derived peptides (γ -series ions) will exist as

doublets allowing identification of one ion series. In many cases complete interpretation of the MS/MS spectrum in a single experiment is possible, especially if high-resolution MS instruments are employed (Shevchenko et al., 1997).

DIRECT MS ANALYSIS OF GEL-SEPARATED PROTEINS AND OTHER MULTI-DIMENSIONAL APPROACHES FOR PROTEIN SEPARATION AND MASS MEASUREMENT

MS and MS/MS Analysis of Proteins and Peptides Following Gel Electrophoresis

The last separation step prior to protein identification and/or characterization often involves electrophoretic separation. As described, mass spectrometry has provided many powerful approaches to protein identification. Methods compatible with mass spectrometry analysis are most often applied to gel-separated proteins and involve generation of peptides by digestion of the protein in the gel or following transfer to a membrane (reviewed in Patterson, 1994; Patterson and Aebersold, 1995). The peptides are then extracted and analyzed as a mixture or alone following separation. One way to speed up this process and potentially avoid losses due to multiple handling steps is to perform all of the mass spectrometry analyses either directly out of the gel or following transfer to a membrane. As MALDI-MS is compatible with solid-phase methods, it is the analytical method of choice. It also provides an opportunity to measure the mass of the intact molecule prior to digestion and analysis of the resulting peptides.

Transfer of the intact proteins to a membrane has always been a preferred step because it allows many potential ionization-suppressing contaminants to be washed away and matrix to be introduced easily. Therefore, it was considered appropriate to transfer proteins to a membrane prior to any attempt to ionize them intact in a mass spectrometer. A number of different membranes have been evaluated for subsequent MALDI-MS analysis of intact proteins using either ultraviolet (UV) or infrared (IR) lasers (reviewed in Vestling and Fenselau, 1995). For proteins or peptides to be ionized directly from a membrane, they first have to be extracted into a solvent that contains matrix, so the protein can cocrystallize with matrix. Therefore, the stronger the interactions between the protein

and the membrane, the less protein will be available for ionization. The choice of solvent for the matrix is also important as solvents with high organic content can eliminate the spatial resolution of the electrophoretically separated proteins due to diffusion (Vestling and Fenselau, 1994a; Loo et al., 1996). In addition, when proteins are electroblotted onto membrane, they often penetrate the surface of the membrane and may not be accessible to the laser (especially the short-wave-length UV lasers generally employed in commercial MALDI-MS instruments). Data suggests that UV lasers are not as efficient at ion extraction as IR lasers are when the analyte is embedded in a membrane following electroblotting, although the effect is less pronounced when the protein is spotted and not electroblotted onto the membrane (Strupat et al., 1994; Schreiner et al., 1996). This was confirmed by the rather large quantities of protein required for detection on the membrane compared to that which could be analyzed by enzymatic digestion (Patterson, 1995; Schreiner et al., 1996). IR-MALDI-MS has been shown to be more efficient at extracting proteins from electroblotted membranes (Strupat et al., 1994), and it has recently been shown to be directly compatible with two-dimensional gel electrophoresis separations and blotting to PVDF (Eckerskorn et al., 1997). This latter study revealed that an area of a blotted membrane could be effectively converted into a mass contour plot that displayed significant microheterogeneity within individual two-dimensional gel electrophoresis-separated protein spots that were not resolved electrophoretically (Eckerskorn et al., 1997). Thus the power of two-dimensional gel electrophoresis combined with accurate mass measurement of intact proteins by IR-MALDI-MS will no doubt be of significant benefit to identification and characterization projects. Current limitations include the small area of the membrane that can be placed into the MALDI-MS, but more of an issue is the amount of data generated and its collation into a format for easy analysis. However, there is no reason to imagine that either of these problems is insurmountable.

Another important issue to be addressed is the further analysis of the protein after the intact mass has been determined. The intact mass is very useful for characterization purposes, but it will not lead to confident identification, which requires generation of peptides for peptide-mass searches and sequence-specific fragment ions from individual peptides for peptide-

sequence tag or uninterpreted fragment-ion searches (see Protein Identification Using Mass Spectrometry–Based Correlative Approaches). This has already been demonstrated, at relatively high picomole loads on the gel, by a number of investigators where chemical or enzymatic cleavages followed by direct UV-MALDI-MS of the resulting peptides (without elution) have been achieved, allowing peptide-mass searching to be used for protein identification (Vestling and Fenselau, 1994a,b). Fabris et al. (1995) also demonstrated that peptides generated by enzymatic cleavage on the membrane could be fragmented by post-source decay (PSD) UV-MALDI-MS. These latter two steps have been achieved for IR-MALDI-MS in a continuation of previously published work (Eckerskorn et al., 1997), according to reports at a recent meeting (F. Lottspeich, pers. comm.), where results were shown for measurement of the intact mass of the protein, peptide masses produced by enzymatic digestion, and PSD-MALDI-MS spectra of some individual peptides, all obtained from individual electrobotted proteins. Such analyses demonstrate the utility of the approach and lay the groundwork for its adaptation to larger scale “proteome” projects.

Originally it was also considered difficult for proteins and peptides to be extracted in an electric field from a gel-matrix, so membrane-based methods were used. However, the Andrews group has demonstrated in a series of proof-of-principle experiments that direct UV-MALDI-MS from certain types of gels is possible (Loo et al., 1996). By separating proteins in an isoelectric focusing (IEF) gel and subjecting this gel to delayed extraction-MALDI-MS after drying the gel and applying matrix, they were able to generate a virtual two-dimensional gel (Loo et al., 1996; see Fig. 10.22.4). The same considerations outlined above for membrane-based approaches apply to analyses of proteins out of the gel—e.g., sensitivity, visualization of the protein in a manner compatible with further analysis, matrix and solvent combinations to allow cocrystallization of the protein from the gel without loss of the spatial resolution, thin gels so that the protein concentration per gel volume ratio isn't too low, and methods compatible with peptide generation and subsequent sequence-specific fragmentation generation. This research was originally performed on a standard continuous-extraction UV-MALDI-MS, but with delayed-extraction UV-MALDI-MS significant improvements in mass accuracy were realized for the analysis of

intact ultrathin (0.35-mm) IEF gel-separated proteins, as well as cyanogen bromide digests of these same proteins (Loo et al., 1997). Sensitivities were in the low to subpicomole range, were most compatible with IEF gels, but were also successful, albeit at higher protein loads, with SDS-PAGE gels (Loo et al., 1997). It is yet to be demonstrated whether PSD-MALDI-MS can be performed on peptides generated in the gel, but as this is an ionization-dependent event, it would not be considered to be a problem.

Therefore, these new approaches for direct analysis of gel-separated proteins either from the gel or following transfer to a membrane show considerable promise for future automated protein identification and characterization protocols.

Multidimensional Separations with MS as the Final Dimension

In classical two-dimensional gel electrophoresis the second dimension provides a mass estimate based upon comparison of SDS-denatured proteins with similarly treated standards. As described in the previous section, substitution of mass spectrometric detection for gel electrophoresis (or as an addition) provides a very accurate mass for each separated component that can be ionized. By employing a liquid-phase first dimension, introduction of the protein mixtures into ESI-mass spectrometers becomes possible by facilitating integration of a series of on-line separation devices (see Fig. 10.22.5). Two examples of such approaches are described for protein analysis, one using IEF coupled to ESI-MS, the other using two orthogonal steps of chromatography on-line with ESI-MS.

Tang et al. (1997) described a capillary IEF gel system interfaced to an ESI-MS for the analysis of total cellular proteins from *E. coli*. They were able to discern just over 100 proteins by this approach, which is comparable with a minigel two-dimensional gel electrophoresis approach where an equal number of proteins were observed by silver staining. Comparison with high-resolution two-dimensional gel electrophoresis would have yielded a dramatically different result, but as with other methods described in this unit, it provides a glimpse of potential future approaches for analysis of complex protein mixtures. Opiteck and coworkers (Opiteck et al., 1997a) utilized two types of chromatography, cation-exchange and reversed-phase, coupled to two separate sets of LC pumps under computer control to separate

proteins on-line for subsequent ESI-MS analysis. Storage loops were used to collect effluent from the cation exchanger for subsequent injection onto the RP-HPLC which was connected to a UV detector as well as an ESI-MS. A limited protein mixture as well as an *E. coli* lysate were separated using this approach in a 2-hr run time (Opiteck et al., 1997a).

Both of these approaches demonstrate the ability to separate and accurately mass analyze

complex protein mixtures with some success. Before such approaches are competitive with two-dimensional gel electrophoresis, they will need to be able to detect proteins over a much larger range of concentrations and to demonstrate that the ionization intensity of any species is not affected by the concentration of a co-ionizing species, so the approach is truly as quantitative as is possible within limits for two-dimensional gel electrophoresis. Another draw-

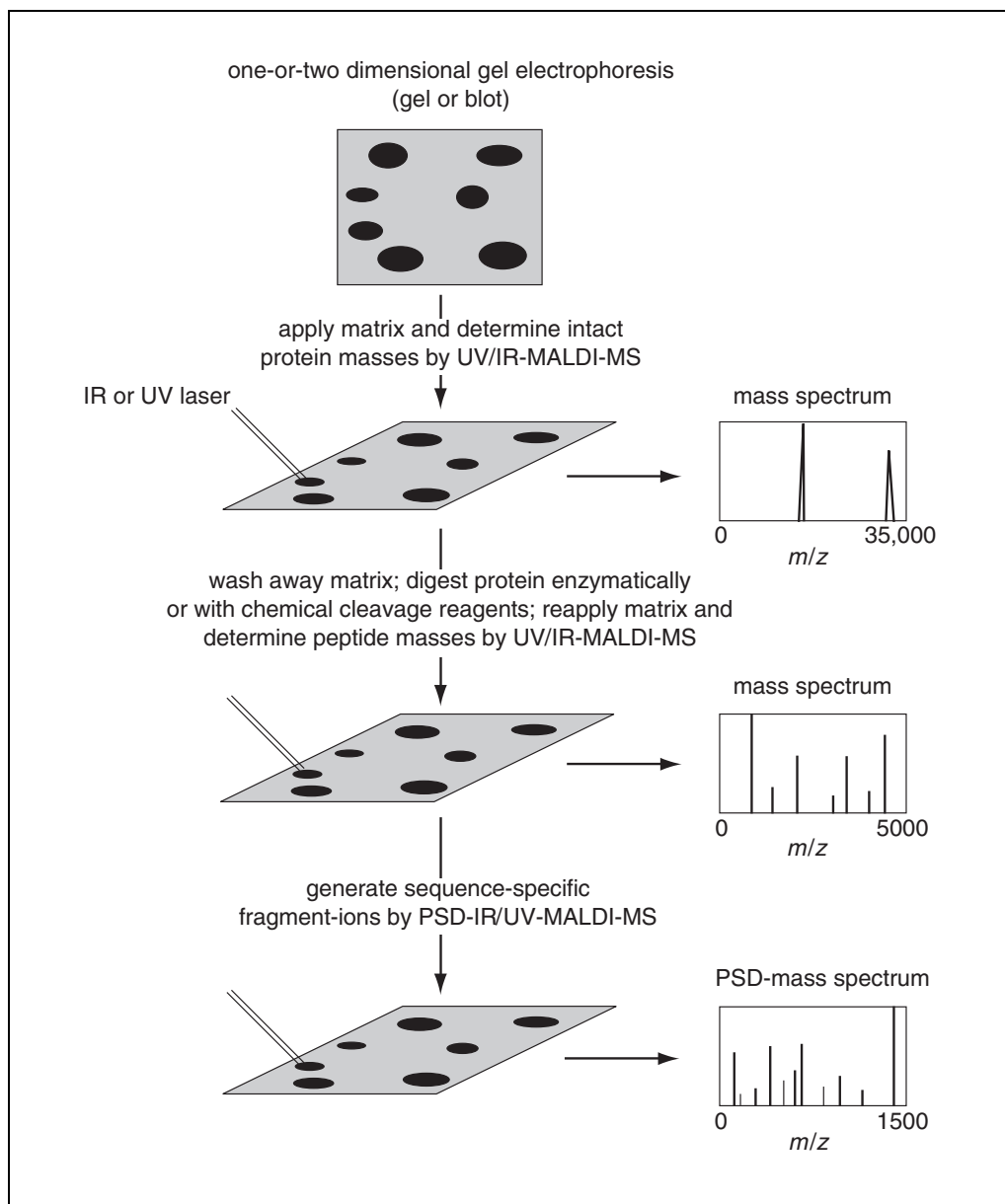


Figure 10.22.4 Direct mass measurement of gel-separated proteins. By inserting either dried thin gels or blotted membranes into a MALDI-MS with an ultraviolet (UV) or an infrared (IR) source, accurate mass measurements of the gel-separated proteins can be made. A “virtual” two-dimensional gel is the result of multiple MALDI-mass measurements of isoelectric (IEF) gel-separated proteins along the length of the IEF gel, thereby generating a mass contour plot of the IEF separation. Ultimately, sequential mass analyses (intact, following enzymatic/chemical cleavage, and sequence-specific fragmentation of individual peptides) will be able to be performed on the same protein spot/band as illustrated.

back of the approach is the inability (at this stage) to fragment the proteins from coeluting mixtures and determine the mass or sequence of the fragments, thereby allowing protein identification. However, it has been demonstrated that using a high-resolution mass spectrometer, CID fragmentation of intact proteins yielded short stretches of sequence-specific fragment ions that could be used in peptide-sequence-tag searches to identify the protein (Mortz et al.,

1996). This was applied to pure protein preparations but as specific ions from each protein were chosen, this approach may become suitable for application to the on-line protein-separation methods described above. With the significant advances that have been made in this field, it will be one to watch in the future.

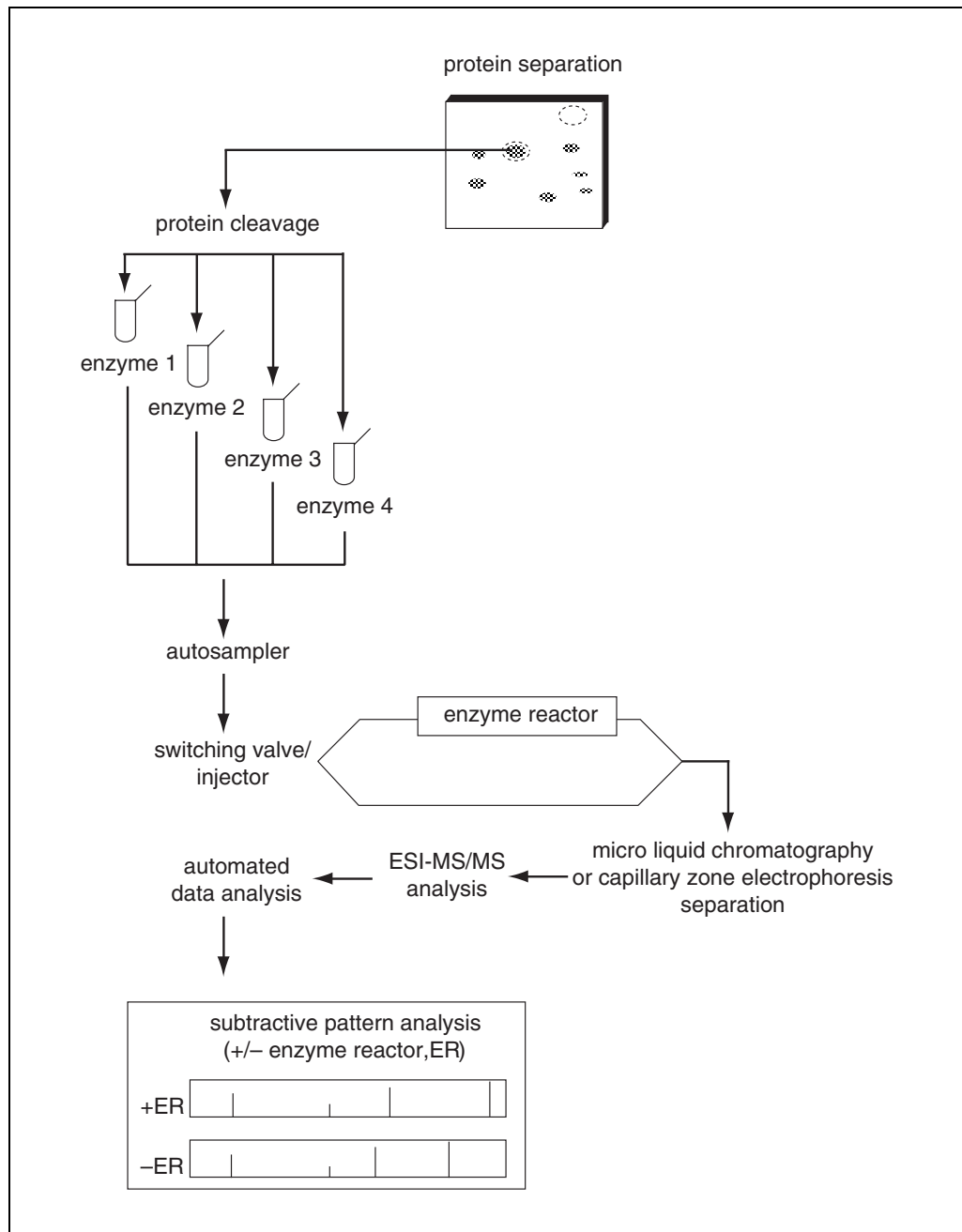


Figure 10.22.5 Schematic representation of an integrated system for determining posttranslational modifications. A range of enzymes can be employed to generate peptides covering the entire sequence of the protein of interest, and when a specific modification is being targeted, an enzyme reactor column can be placed on-line to remove the modification from an aliquot of the peptide digest (modified from figures supplied by Dr. R. Aebersold, University of Washington, Seattle, USA).

RAPID POSTTRANSLATIONAL MODIFICATION ANALYSIS

Given that mass spectrometers are accurate balances, they are ideal instruments to identify posttranslational modifications in an unbiased manner without the use of radioactive precursors, e.g., one will only observe phosphorylation if radiolabeled phosphate is measured in an experiment. An ideal situation would be to be able to identify and characterize all of the posttranslational modifications present on a protein in a single experiment. To achieve this aim requires an approach to unambiguously characterize any and all posttranslational modifications that are stable to the mass spectrometric ionization process on an identified protein. This can be accomplished by mass measurement of (1) the intact protein to estimate the type and number of modifications, (2) peptides which represent all of the protein to localize the site of modifications, and (3) fragment ions of the modified peptide(s) to identify the site of modification. Many posttranslational modifications have been identified in an unbiased manner through the use of mass spectrometry analysis, and in many cases Edman sequencing, of recombinant proteins, often as a consequence of their overexpression (Tu et al., 1995 and references therein). However, fulfilling all of the above-mentioned criteria is a daunting task even when the protein has been obtained from recombinant sources and is therefore available in relatively high quantity. Characterization of *in vivo* protein modifications is the ultimate aim, and progress is being made towards this end. The applications of new protein technologies to this field are demonstrated through examples of analysis of phosphorylation.

The most reported posttranslational modification is that of phosphorylation due to its critical role in signaling pathways, predominantly through either the addition or removal of phosphate esters from the hydroxyl groups of serine, threonine, and tyrosine residues (Hunter, 1987; Charbonneau and Tonks, 1992). Phosphoramidates can also be formed with histidine, arginine, and lysine residues, and acylphosphates with aspartic and glutamic acid residues, but these are considered less frequent and therefore, their functions are less well understood. However, due to the labile nature of some of these modifications they may well be more frequent and important than currently assumed (Huang et al., 1991).

Classical Approaches for Study of Protein Phosphorylation

Briefly, classical methods for the analysis of protein phosphorylation use two different approaches to add radiolabeled phosphate to proteins—metabolic labeling which involves equilibration of the cellular ATP pool with $^{32}\text{PO}_4$ (see *UNIT 18.2*) or an *in vitro* kinase reaction in which purified kinases (possibly immunoprecipitated) are used to phosphorylate crude or fractionated (sometimes coimmunoprecipitated) cell lysates using $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ (*UNIT 18.7*). The mixture is then separated by gel electrophoresis and the migration of the labeled protein detected by autoradiography. The type of amino acid residues modified can be identified by mild base hydrolysis of the excised proteins, and separation of the amino acids by thin-layer chromatography (TLC) or electrophoresis (*UNIT 18.3*). To determine the site of modification, the protein must be fragmented and the resulting peptides separated, usually by two-dimensional TLC with autoradiographic detection, or sometimes by RP-HPLC with Cerenkov counting of collected fractions. Again, the type of residue phosphorylated can be determined by mild base hydrolysis followed by one- or two-dimensional TLC (see *UNIT 18.3*). Alternatively, the site can be revealed by Edman sequencing of the fractionated peptide, in the rare cases where sufficient sample is available, or by comigration of synthetic phosphorylated peptides designed from potential sites of phosphorylation within the protein of interest. For both of these approaches to work the amino acid sequence of the protein has to be known. The ability to raise antisera to phosphotyrosine has provided a very simple nonradioactive means of identifying proteins carrying this modification (see *UNITS 18.5 & 18.6*). Antibodies are available from several commercial sources, providing a rapid and sensitive means of monitoring tyrosine phosphorylation on SDS-PAGE (including two-dimensional gel electrophoresis) separated proteins. Antisera directed against phosphoserine and phosphothreonine detect some but not all proteins modified in this manner.

MS Approaches for Identifying Regions and Sites of Phosphorylation

A number of mass spectrometry-based approaches have been developed to address the issue of protein phosphorylation. They fall into two main categories: searching for mass differences consistent with phosphorylation (i.e.,

+80 u or multiples thereof) and then treating with phosphatases to confirm the assignment, and searching for phosphorylated peptides through the use of high-orifice scanning in ESI-MS to find phosphate-specific signature ions (see *UNIT 10.21*). Although it is always preferable to determine the intact mass of the protein as the first step in analysis, this is often not possible; therefore, the protein is enzymatically digested and the masses of the resulting peptides are determined.

If the entire sequence of the protein under investigation is known, then putative phosphorylated peptides can be assigned based upon which peptides do not match the expected masses. Of course, this only localizes the site of modification, but doesn't confirm that the expected residue is modified with phosphate; that requires sequencing of the peptide either chemically or by MS/MS. Confirmation of the observed mass difference can be made by treating the digest with a phosphatase and measuring the peptide masses again. This can be conducted with a separate aliquot in solution, or as Craig et al. (1994) and Liao et al. (1994) have demonstrated for MALDI-MS, directly on the probe, making the analysis rapid and with minimal sample loss. If the putative assignments were correct, those previously phosphorylated peptides will now display masses 80 u, or multiples thereof, less than the first analysis. This seems relatively straightforward, but the approach may not always result in complete assignment of all phosphorylated sites. Independent of the mass spectrometer used (MALDI-MS or ESI-MS), not all of the peptides generated from the digest may be able to be mass analyzed, e.g., some peptides may only be ionized in either positive- or negative-ion mode but not both, or some peptides may become insoluble following digestion. Mass spectrometry of peptides is usually conducted in positive-ion mode (see *UNIT 10.21*), but because phosphorylation is a negatively charged modification, it sometimes changes the properties of the peptide such that it will only be ionized in negative-ion mode. For MALDI-MS, conducting experiments in both modes sequentially is not a problem because the sample is immobilized, and many analyses can be performed on the same sample. In ESI-MS, the analysis is usually performed in one mode (positive or negative) or the other.

To overcome the problem of incomplete coverage of the protein due to insolubility or other problems, a parallel digest with an enzyme of differing specificity is often used. In

an elegant study of a PAK/STE20 homolog from *Acanthamoeba castellanii*, Szczepanowska et al. (1997) used MALDI-MS on both a TOF and an ion-trap instrument to identify the site of *in vivo* modification by analysis of unfractionated tryptic and Glu-C digests. In this study, all of the possible sites of phosphorylation were observed as tryptic or Glu-C-derived peptides, and the one site of phosphorylation was identified in both the baculovirus-expressed truncated form of the protein as well as the endogenous full-length form. Both phosphatase treatment and MS/MS fragmentation were employed for confirmation that the peptide was phosphorylated and to determine the site of modification.

In the second approach, advantage was taken of the observation that under ionization conditions that leave most peptide bonds intact, phosphate ester bonds in phosphopeptides can be dissociated at the source (or in a collision cell) of an ESI-MS. When the analysis was conducted in negative-ion mode (the potential on the spraying needle is negative) with the orifice voltage at a high potential (e.g., -350 V), phosphopeptides containing phosphoserine or phosphothreonine yielded signature low-mass fragment ions at 97, 79, and 63 u, and phosphotyrosine-containing phosphopeptide low-mass ions at 79 and 63 u. In the original methods of Huddleston et al. (1993, 1994) and Ding et al. (1994), peptide digests were separated by RP-HPLC. On-line ESI-MS was employed to determine both the mass of the peptide as well as its phosphorylation status by switching the voltage on the orifice between a high potential when scanning (measuring masses) in the low-mass region, to a normal potential when scanning over the remainder of the mass range, in a cyclical manner. Of course, fractions can be collected during the run for subsequent analysis. This is a rapid method to identify which peptides in a given mixture are phosphorylated, as well as what their intact masses are (unless the mixture is too complex and many peptides coelute on the RP-HPLC). However, as with the MALDI-MS approach, this only localizes the modified residues and does not confirm the site of modification until the fractionated peptides are sequenced either chemically or by MS/MS. Another feature of the approach is that even when the RP-HPLC is conducted on a 320- μm -i.d. capillary column, the limit of detection is suggested to be 50 pmol (Huddleston et al., 1993), thereby limiting the RP-HPLC/ESI-MS approach to *in vitro* phosphorylation site analysis.

The ability to spray sample into the ESI-MS at very low flow rates (so called nanospray, due to low nanoliter per minute flow rates; Wilm and Mann, 1994), has dramatically increased the sensitivity of MS/MS sequencing. This method can also be used to advantage for post-translational modification analysis of *in vivo*-generated samples due to the high sensitivity. In addition to the method where high-orifice scanning generates signature low-mass ions for phosphorylated residues, another MS/MS-based approach can also be used in conjunction with a triple quadrupole instrument (see *UNIT 10.21*). This instrument can be run in a number of modes, two of which are described here. In normal MS/MS mode to generate fragment ions from a specific precursor ion (peptide), the first set of quadrupoles (Q1) is set to allow transmission of only the peptide ion of interest into Q2 which in turn is set to transmit all ions, not just a specific m/z into which a curtain of gas (e.g., argon) has been injected. The peptide ions undergo collision-induced decomposition (CID) into sequence-specific fragment ions, whose masses can be determined by scanning Q3. Thus, the fragment ions from a specific precursor ion can be generated and measured. In a mode referred to as precursor (or parent)-ion scanning, the third set of quadrupoles (Q3) are set to transmit only a specific m/z ; in this example for identification of phosphorylated peptides in a mixture, this is set to $m/z = 79$. The collision gas is left on in Q2, and ions are scanned over the mass range in Q1. Therefore, a signal will only be recorded when a fragment ion of 79 u is generated from a parent ion of the mass being scanned at that point in time in Q1. A spectrum is produced of any peptide that following CID produces a fragment ion of 79 u, i.e., a phosphopeptide. The advantage of this approach of precursor-ion scanning is that the phosphopeptides can be selected from a mixture for fragmentation analysis in the normal mode (i.e., select in Q1, fragment in Q2, mass analyze in Q3).

This approach has been employed successfully by Betts et al. (1997) in their identification of sites of *in vivo* phosphorylation on neurofilament proteins (NF-M and NF-L) isolated from rat brain. They separated the neurofilament-enriched fraction from rat brain by SDS-PAGE, digested the protein in the gel, desalted it, and performed nanoESI-MS/MS on the resulting mixtures. They also analyzed peptide mixtures that had been enriched in phosphopeptides by binding the *in-gel* digest eluant to an immobilized metal-affinity column

(IMAC; *UNIT 10.11B*) equilibrated with FeCl_3 followed by elution with Tris-Cl (Betts et al., 1997). They analyzed not only the phosphopeptides but all of the peptides possible to determine whether the phosphorylation was heterogeneous at the identified sites. By using three enzymes, trypsin, Glu-C, and Asp-N, they achieved 81% sequence coverage for NF-L and 64% sequence coverage for NF-M. Interestingly, no sites of phosphorylation were found on NF-L, and the four sites identified on NF-M were shown to be heterogeneous.

Automated Approaches for Identifying Phosphorylation Sites

The nanoESI-MS/MS approach described above provides high-sensitivity analytical capability for both protein identification and post-translational modification determination. The drawbacks are that setup of the device is not rapid (although a number of simpler devices are now becoming available), and the analysis is manual, i.e., an operator manually sets the instrument in all different modes and selects peptides for fragmentation. Other groups have been investigating automating posttranslational modification analysis by constructing integrated on-line systems. These consist of an RP-HPLC-ESI-MS/MS (LC-MS/MS) system for automated MS/MS of separated peptides, with an additional column on-line in front of the RP-HPLC system which provides the specificity to the system. Such columns include an IMAC column (for general capture of phosphorylated peptides), immunoaffinity and SH2-domain affinity column (phosphotyrosine specific), and an enzyme reactor (a tyrosine phosphatase). A general schema for such devices is shown in Figure 10.22.5 and a brief description of these approaches follows.

Watts et al. (1994) used an on-line IMAC column in front of a LC-MS to capture phosphopeptides in a protocol that resulted in determination of the sites of *in vivo* tyrosine phosphorylation of the lymphocyte-specific protein tyrosine kinase ZAP-70. In this approach, standard two-dimensional TLC phosphopeptide maps were produced for both *in vivo* phosphorylated ZAP-70 and recombinant ZAP-70 phosphorylated with recombinant p56lck *in vitro*. From the *in vitro* generated two-dimensional TLC plate, spots were excised, the slurry extracted and recovered peptides subjected to IMAC-LC-MS to determine the mass of the phosphopeptides and identify the sites of phosphorylation. Comparison (including comigration of excised spots) of the *in vivo*- and *in vitro*-

vitro-generated two-dimensional TLC patterns allowed identification of the sites of autophosphorylation and T cell receptor-stimulated phosphorylation. The sensitivity of the system was demonstrated to be ~250 fmol (Watts et al., 1994). This approach provides a high-sensitivity system that complements the extensively utilized two-dimensional TLC mapping procedure (Watts et al., 1996).

Specific capture of tyrosine-phosphorylated peptides has been achieved using an antiphosphotyrosine column in-line with an LC-MS (Kassel et al., 1995). The specificity of the system was demonstrated, but the sensitivity of the approach was not determined. Kassel et al. (1995) also produced a more specific affinity column by immobilizing a recombinant SH2 domain to the first column, thus allowing capture of only tyrosine phosphopeptides containing the specific sequence motif recognized by this domain. Amankwa et al. (1995) developed an “enzyme reactor” system employing a broad-specificity tyrosine phosphatase on-line prior to either capillary zone electrophoresis (CZE)-MS or LC-MS. Comparison of peptide maps with and without passage through the enzyme reactor allowed determination of sites of tyrosine phosphorylation at low levels (2.6 pmol).

Role of MS in Posttranslational Modification Analysis

More than 400 posttranslational modifications have been described to date (Krishna and Wold, 1993), and mass spectrometry has been used extensively in the characterization of many of them (for a list of posttranslational modifications and their masses see *APPENDIX 1B*, Table A.1C.4, or for an extensive on-line list see Dr. K. Mitchell-Hill’s Delta Mass site at <http://www.medstv.unimelb.edu.au/WWWD/OCS/SIVMRDocs/MassSpec/deltamassV2.html>). However, mass spectrometry has yet to be employed in a high-throughput mode for a variety of modifications, and in many cases it is not the primary method of choice in characterization of posttranslational modifications. With the increased sensitivity, availability, and ease of use of these instruments, mass spectrometry will be used more and more in routine analyses to detect and characterize posttranslational modifications. As stated previously, characterization of the posttranslational modifications present on a protein requires an initial measurement of the intact mass of the protein. For electrophoretically separated proteins, this has proven to be a difficult task at low levels

when the protein is eluted from the gel slice for subsequent ESI-MS analysis. Although a few protocols recently have been demonstrated for mass analysis of low-picomole quantities of pure protein eluted from gels (e.g., Schuhmacher et al., 1996; Cohen and Chait, 1997), the methods described previously for direct mass measurement of gel-separated proteins either off membranes or out of gels may simplify this first step in the analysis. Enzymatic digestion and analysis of the resulting fragments allows the region of modification to be determined, and generation of fragment-ion spectra from these peptides in most cases will assign the site of modification. Thus, the ability to characterize small quantities of gel-separated proteins is close to becoming reality.

An extremely promising advance in posttranslational modification analysis is the demonstration by Bean et al. (1995) of the lability of a number of posttranslational modifications to stepped high-orifice scanning in LC-MS analyses as described for phosphorylation analysis. This approach results in the generation of signature ions for specific modifications to indicate their presence on peptides eluting at that retention time, and in combination with MS analysis at normal orifice voltage, the intact mass of the modified peptides. Modifications amenable to stepped high-orifice scanning include phosphorylation, palmitoylation, sulfation, and glycosylation (Bean et al., 1995). In the case of glycosylation the presence of individual classes of carbohydrates present on individual peptides—hexose, hexosamine, and neuraminic acid—are revealed. Fragmentation of the peptide in an automated mode may reveal the site of modification. However, in some cases (e.g., glycosylation), the modifying group has to be removed for sequence-specific fragment-ion data to be obtained.

Identification of the site of modification is critical to complete analysis, but if the modification is not stable to the mass spectrometry ionization or fragmentation process, then the modifying group needs to be removed and in doing so, “tag” the site from which it was removed. In this case the peptide is then fragmented and the site of former modification identified. Gonzalez et al. (1992) used this approach to identify sites of N-glycosylation by using the enzyme peptide-*N*-4(*N*-acetyl- β -glucosaminyl)asparagine amidase F (also known as PNGase F) in the presence of 50% H₂¹⁸O. PNGase F completely removes the carbohydrate side chain from the asparagine residue and in so doing results in conversion of this

residue to aspartic acid, which is 1 u heavier than asparagine following incorporation of H₂O from the solvent (see *UNIT 17.14A*). By conducting the reaction in 50% (v/v) H₂¹⁸O, 50% of the oxygen atoms incorporated during the removal of the carbohydrate side chain and conversion of asparagine to aspartic acid will have a mass 2 u greater than normal aspartic acid (see Fig. 10.22.6). Therefore, a pair of peptides differing in mass by 2 u will be produced following this reaction. MS/MS analysis of this peptide pair will result in doublets 2 u apart for all fragment ions containing the site of modification thereby allowing identification of the site of former modification (Gonzalez et al., 1992).

MS/MS analysis of peptides containing O-linked sugar or phosphate on serine and threonine residues does not always allow unambiguous determination of the site of modification due to the lability of the O-linked modifying group; often significant neutral loss occurs during MS/MS, complicating the analysis. However, removal of either of these modifying groups by β -elimination results in dehydration of the serine and threonine residues, leaving unique residue masses for each, 18 u less than normal (see Fig. 10.22.6; see *UNIT 17.15B* for β -elimination protocols applied to glycosylated peptides). This has been applied to determination of sites of O-GlcNAc modification of intracellular proteins and to characterize sites of phosphorylation in the highly phosphorylated profilaggrin (Resing et al., 1995; Greis et al., 1996). In both cases, the altered residue mass (-18 u for each site of modification) was used to localize the peptides upon which the modification occurred, and the site was identified by MS/MS. Greis et al. (1996) demonstrated the ability to determine the intact mass of glycopeptides by LC-MS analysis of the native sample and, following β -elimination, the site of glycosylation on the peptides. Resing et al. (1995) noted that MS/MS fragmentation was more pronounced at the peptide bonds either side of the dehydrated residue thus facilitating analysis of the fragment-ion spectrum.

Tagging sites of modification by changing the mass of the amino acid carrying the modifying group provides a powerful tool in determining the site of modification. Conducting MS and MS/MS analysis of the sample prior to removal of the group also allows some characterization of the modifying group. Interestingly it has recently been demonstrated that mixtures of N-linked oligosaccharide structures can be determined from gel-separated proteins at the

low microgram level by in-gel PNGase F digestion followed by MALDI-MS analysis, after which the remaining protein can be digested for subsequent peptide-mass mapping experiments (Kuster et al., 1997).

HIGH-THROUGHPUT PROTEIN IDENTIFICATION

The highest-resolution protein separation method is still two-dimensional gel electrophoresis. Two-dimensional gel electrophoresis provides the ability to display literally thousands of proteins in one experiment, revealing their relative levels of expression and, depending on how the sample was prepared, their subcellular localization and predetermined posttranslational modifications (i.e., by radiolabeling with specific metabolic precursors). Since the method of two-dimensional gel electrophoresis was first introduced in 1975 (Klose, 1975; O'Farrell, 1975; Scheele, 1975; also see *UNIT 10.3*), a frustrating aspect has been the inability to identify individual proteins and extensively characterize their modifications. This unit and others in this volume (and *Current Protocols in Protein Science*) describe the increasing ability to identify minute quantities of gel-separated protein, but large-scale identification of many proteins in a two-dimensional gel electrophoresis pattern has only recently begun to be addressed successfully.

One semiautomated approach uses existing methods for parallel enzymatic digestion of excised spots and automated loading of a fraction of the digest onto a MALDI-MS plate for subsequent unattended data generation and capture followed by peptide-mass searching (Shevchenko et al., 1996). This approach was used in a yeast proteome project where proteins isolated by two-dimensional gel electrophoresis were identified using a two-part strategy: peptide-mass searching employing an automated 32-port parallel enzyme digestion robot (Houthaeve et al., 1995), followed by automated MALDI-MS with delayed-extraction of a small aliquot of the peptide digest and subsequent automated peptide-mass searching. If this didn't conclusively identify the protein, nano-ESI-MS/MS to generate a sequence tag was undertaken using half of the remaining sample after it was reduced in volume and cleaned up using a microdesalting step. Peptide-sequence tags generated from a number of peptides in the mixture were then used in searches to identify the protein. If this was not successful, then the remainder of the digest could be derivatized by methyl esterification

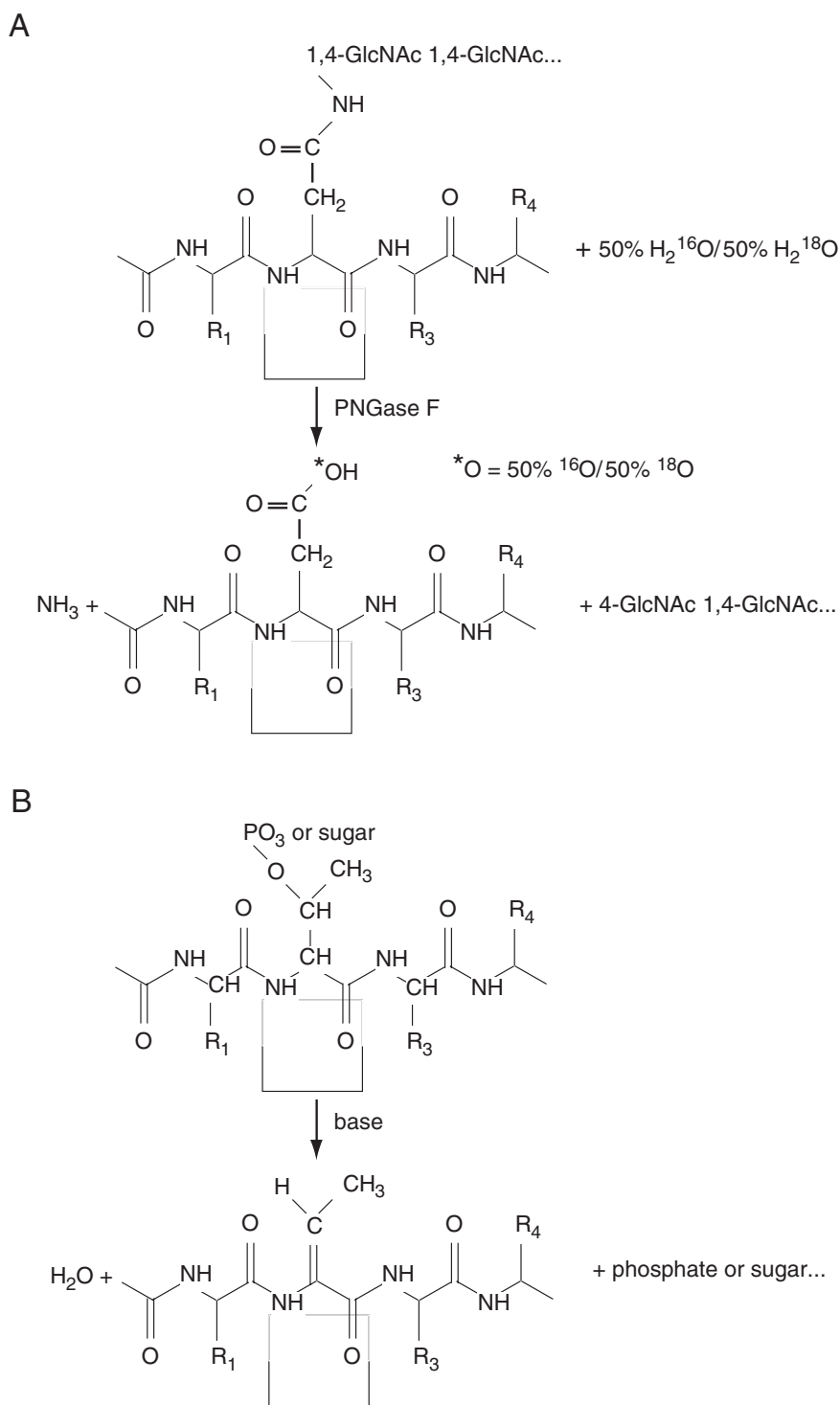


Figure 10.22.6 “Tagging” sites of posttranslational modification for mass spectrometric analysis. **(A)** Tagging sites of N-linked glycosylation. PNGase F is used to remove N-linked carbohydrate side chains in the presence of 50% H₂¹⁸O (50% normal water) changing the asparagine residue (residue mass 114 u) to an aspartic acid residue (which exists as a doublet of residue masses 115 u and 117 u). **(B)** Tagging sites of O-linked phosphorylation or glycosylation. β -elimination (base hydrolysis) will remove O-linked phosphate or sugar from threonine (residue mass 103 u, shown) or serine (residue mass 87 u) residues and in doing so dehydrate the side chain resulting in a 18 u decrease in mass to yield dehydroamino-2-butyric acid (residue mass 83 u) and dehydroalanine (residue mass 69 u) respectively. The amino acid residues of interest are bracketed.

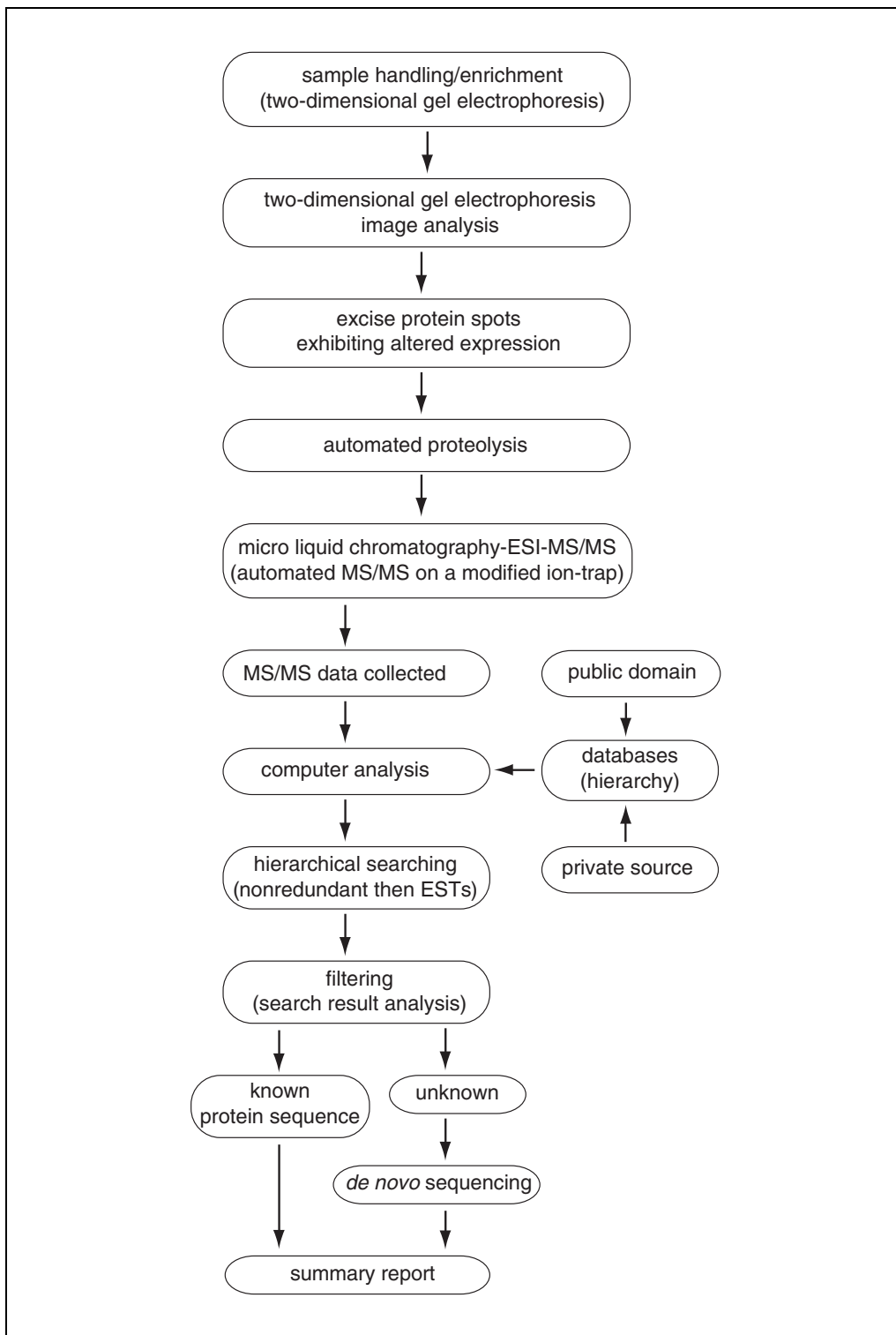


Figure 10.22.7 A scheme for automated identification of proteins exhibiting altered expression when assayed by two-dimensional gel electrophoresis (adapted from figures supplied by Dr. R. B. Parekh, Oxford GlycoSciences, Abingdon, UK). The samples to be analyzed are fractionated as appropriate prior to two-dimensional gel electrophoresis separation using an automated system. The two-dimensional gel electrophoresis–resolved proteins are imaged and differences between samples highlighted by a computer system, which then excises protein spots from the gel and digests them—all under robotic control. The digest is then separated by micro-LC-ESI-MS/MS in an automated mode, and the MS/MS spectra are transferred to a database from which automated searches are conducted against protein and translated nucleotide (including ESTs) databases. If the sequence resides in a database, the match is determined; if not, the MS/MS pattern is used to interpret the sequence *de novo*.

and subjected to an additional nano-ESI-MS/MS analysis to interpret the sequence de novo from the fragment-ion spectra (Shevchenko et al., 1996). This approach was able to identify 150 yeast proteins separated by two-dimensional gel electrophoresis in a high-throughput mode.

A fully automated system has been developed by Oxford GlycoSciences (Dr. R. B. Parekh, pers. comm.), where two-dimensional gels are run under computer control followed by automated image analysis of fluorescently labeled proteins. This allows the user to select spots of interest on a computer screen for automated excision, digestion, LC-MS/MS, and SEQUEST software analysis (see Fig. 10.22.7). For large-scale two-dimensional gel electrophoresis projects this type of approach can be advantageous, and it fulfills the requirements of a proteomics approach to two-dimensional gel electrophoresis-separated proteins. However, these approaches are most likely to be adapted only in core facilities or even more specialized laboratories.

When the mixtures are less complex, a chromatography-based approach has been demonstrated to provide rapid automated protein identification without employing any steps of electrophoresis. All of the approaches described so far analyze peptides derived from one or a few proteins and therefore, a number of fragment-ion spectra are obtained from each protein, and these allow redundant identifications. Other automated approaches rely upon only one peptide fragment-ion spectrum for identification as the protein mixture is digested prior to peptide separation. This provides a simpler approach for automation, but at least one peptide from each protein in the mixture must be analyzed for a completely successful analysis to be performed.

Enzymatic digestion of relatively complex protein mixtures followed by data-dependent LC-MS/MS (i.e., the mass spectrometer selects the peptides on which to perform MS/MS on-the-fly) and automated uninterpreted fragment-ion searching of sequence databases using the SEQUEST program, has been demonstrated to be capable of identifying many components within a mixture (McCormack et al., 1997). The key to the success of such an approach is the ability to isolate mixtures of limited complexity that are of biological interest. Yates and colleagues (McCormack et al., 1997) demonstrated identification of components of a co-immunoprecipitation (see UNIT 10.16), a micro-

tubule complex, as well as a set of protein standards, to demonstrate the ability to detect a component at lower concentration in a mixture where the other components were at a much higher level. This analysis revealed that fragment-ion spectra from only a few peptides were usually obtained, but of course, in using this approach only one MS/MS spectrum needs to be obtained per protein if the spectral quality is good enough. The main advantage of this approach is the speed with which the analysis can take place compared even with automated two-dimensional gel electrophoresis approaches and the relative simplicity of the setup. The described system uses capillary LC which allows low amounts of proteins to be analyzed. The disadvantages of the approach include: (1) difficulty in comparing relative protein levels, aside from presence or absence; (2) generation of MS/MS data from contaminants when they could have been excluded from analysis if a comparative (e.g., electrophoretic) analysis had been performed; and (3) the limited number of components that can be analyzed in one experiment. However, given these caveats there are still many applications that could benefit from an analytical approach capable of rapidly identifying a mixture of protein components through analysis of their derived peptides. For example, in addition to the examples presented above, identification of the major components present during a specific purification step may allow the purification to be directed to remove unwanted components. An increased number of peptide components could be analyzed if a two-dimensional chromatographic system were employed.

Opiteck et al. (1997b) described a two-dimensional chromatography system where peptides were separated by size-exclusion chromatography (SEC; UNIT 10.14) and loaded alternately onto one of two HPLC columns for sequential, rapid RP-HPLC separation followed by ESI-MS analysis. Four-minute SEC fractions were loaded onto the RP-HPLC column, and while this was being separated over 3 min (with 1 min reequilibration), the other HPLC column was being loaded. The two HPLC columns were run alternately over 160 min, each running 20 gradients in that time (40 consecutive gradients were run by the one set of HPLC pumps). This approach provided a much greater resolving power than either method alone and in an automated format (Opiteck et al. 1997b). Of course, such a system could be used in conjunction with an ESI-MS

system capable of data-dependent fragmentation, yielding fragment-ion spectra for each peptide component as well.

In an extension of the LC-MS/MS method, additional on-line steps have been added at the beginning to automate the entire analysis by allowing (1) immunoaffinity chromatography purification, (2) desalting and buffer exchange on a mixed-bed strong ion-exchange absorbent, (3) digestion on an immobilized trypsin column, (4) capture of the tryptic peptides on a short perfusion capillary reversed-phase column, and (5) separation on an analytical reversed-phase column with the column effluent being analyzed by ESI-MS(/MS) (Hsieh et al., 1996). The entire computer-controlled setup requires an autosampler, five columns, and three 10-port switching valves to allow all of the steps to be performed on-line without any manual transfers of material (Hsieh et al., 1996). The method was described only for analysis of protein purified by the system, but it may be able to be extended to the analysis of many components of a complex protein mixture.

CONCLUSION

This unit has provided the reader with a glimpse into some new technologies that have been developed for protein analysis. Not all aspects of protein analysis have been covered, with an emphasis being placed on protein identification and characterization. However, it is hoped that the reader will have a better understanding of how mass spectrometry techniques can be used to answer questions of biological importance both rapidly and accurately. Proteome analysis in its broadest definition should provide evidence of the steady-state level of expression for all individual gene products in a cell and their modifications at a defined point in the life of the cell. In its simplest form, it can be used just to confirm open reading frames in DNA sequences. In addition, unless the cell population is truly synchronous, such an analysis needs to be undertaken on a very small number of cells. Unless radiolabelling of newly synthesized proteins is performed, only the steady-state levels of proteins can be determined. In addition to expression levels, the subcellular localization of each protein needs to be identified, and this is an extremely complex process requiring many parallel experiments. Such an analysis is still far from being realized without using a large number of experimental protocols. However, with the tools currently in hand, the prospect of attaining such

a goal in a reasonable time frame looks a little less daunting.

LITERATURE CITED

- Amankwa, L.N., Harder, K., Jirik, F., and Aebersold, R. 1995. High-sensitivity determination of tyrosine-phosphorylated peptides by on-line enzyme reactor and electrospray-ionization mass spectrometry. *Protein Sci.* 4:113-125.
- Bean, M.F., Annan, R.S., Hemling, M.E., Mentzer, M., Huddleston, M.J., and Carr, S.A. 1995. LC-MS methods for selective detection of posttranslational modifications in proteins: Glycosylation, phosphorylation, sulfation, and acylation. *In Techniques in Protein Chemistry VI.* (J.W. Crabb, ed.) pp. 107-116. Academic Press, San Diego.
- Betts, J.C., Blackstock, W.P., Ward, M.A., and Anderton, B.H. 1997. Identification of phosphorylation sites on neurofilament proteins by nanoelectrospray mass spectrometry. *J. Biol. Chem.* 272:12922-12927.
- Charbonneau, H. and Tonks, N.K. 1992. 1002 protein phosphatases? *Annu. Rev. Cell Biol.* 8:463-493.
- Cohen, S.L. 1996. Domain elucidation by mass spectrometry. *Structure* 4:1013-1016.
- Cohen, S.L. and Chait, B.T. 1997. Mass spectrometry of whole proteins eluted from sodium dodecyl sulfate-polyacrylamide gel electrophoresis gels. *Anal. Biochem.* 247:257-267.
- Coligan, J.E., Dunn, B.M., Ploegh, H.L., Speicher, D.W., and Wingfield, P.T. (eds.). 1997. *Current Protocols in Protein Science.* John Wiley & Sons, New York.
- Cordwell, S.J., Wasinger, V.C., Cerpa-Poljak, A., Duncan, M.W., and Humphery-Smith, I. 1997. Conserved motifs as the basis for recognition of homologous proteins across species boundaries using peptide-mass fingerprinting. *J. Mass Spectrom.* 32:370-378.
- Cottrell, J.S. 1994. Protein identification by peptide mass fingerprinting. *Pept. Res.* 7:115-124.
- Craig, A.G., Hoeger, C.A., Miller, C.L., Goedken, T., Rivier, J.E., and Fischer, W.H. 1994. Monitoring protein kinase and phosphatase reactions with matrix-assisted laser desorption/ionization mass spectrometry and capillary zone electrophoresis: Comparison of the detection efficiency of peptide-phosphopeptide mixtures. *Biol. Mass Spectrom.* 23:519-528.
- Craig, A.G., Fischer, W.H., Rivier, J.E., McIntosh, J.M., and Gray, W.R. 1995. MS-based scanning methodologies applied to *Conus* venom. *In Techniques in Protein Chemistry VI.* (J.W. Crabb, ed.) pp. 31-38. Academic Press, San Diego.
- Ding, J., Burkhart, W., and Kassel, D.B. 1994. Identification of phosphorylated peptides from complex mixtures using negative-ion orifice-potential stepping and capillary liquid chromatography/electrospray ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* 8:94-98.

- Eckerskorn, C., Strupat, K., Schleuder, D., Hochstrasser, D., Sanchez, J.-C., Lottspeich, F., and Hillenkamp, F. 1997. Analysis of proteins by direct-scanning infrared-MALDI mass spectrometry after 2D-PAGE separation and electroblotting. *Anal. Chem.* 69:2888-2892.
- Eng, J.K., McCormack, A.L., and Yates, J.R. III. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* 5:976-989.
- Fabris, D., Vestling, M.M., Cordero, M.M., Doroshenko, V.M., Cotter, R.J., and Fenselau, C. 1995. Sequencing electroblotted proteins by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* 9:1051-1055.
- Gonzalez, J., Takao, T., Hori, H., Besada, V., Rodriguez, R., Padron, G., and Shimonishi, Y. 1992. A method for determination of N-glycosylation sites in glycoproteins by collision-induced dissociation analysis in fast atom bombardment mass spectrometry: Identification of the positions of carbohydrate-linked asparagine in recombinant alpha-amylase by treatment with peptide-N-glycosidase F in ¹⁸O-labeled water. *Anal. Biochem.* 205:151-158.
- Greis, K.D., Hayes, B.K., Comer, F.I., Kirk, M., Barnes, S., Lowary, T.L., and Hart, G.W. 1996. Selective detection and site-analysis of O-GlcNAc-modified glycopeptides by beta-elimination and tandem electrospray mass spectrometry. *Anal. Biochem.* 234:38-49.
- Houthaeve, T., Gausepohl, H., Mann, M., and Ashman, K. 1995. Automation of micro-preparation and enzymatic cleavage of gel electrophoretically separated proteins. *FEBS Lett.* 376:91-94.
- Hsieh, Y.L., Wang, H., Elicone, C., Mark, J., Martin, S.A., and Regnier, F. 1996. Automated analytical system for the examination of protein primary structure. *Anal. Chem.* 68:455-462.
- Huang, J.M., Wei, Y.F., Kim, Y.H., Osterberg, L., and Matthews, H.R. 1991. Purification of a protein histidine kinase from the yeast *Saccharomyces cerevisiae*. The first member of this class of protein kinases. *J. Biol. Chem.* 266:9023-9031.
- Huddleston, M.J., Bean, M.F., and Carr, S.A. 1993. Collisional fragmentation of glycopeptides by electrospray ionization LC/MS and LC/MS/MS: Methods for selective detection of glycopeptides in protein digests. *Anal. Chem.* 65:877-884.
- Huddleston, M.J., Annan, R.S., Bean, M.F., and Carr, S.A. 1994. Selective detection of Thr-, Ser-, and Tyr-phosphopeptides in complex digests by electrospray LC-MS. In *Techniques in Protein Chemistry V*. (J.W. Crabb, ed.) pp. 123-130. Academic Press, San Diego.
- Hunter, T. 1987. 1001 protein kinases. *Cell* 50:823-829.
- James, P. 1997. Of genomes and proteomes. *Biochem. Biophys. Res. Commun.* 231:1-6.
- James, P., Quadroni, M., Carafoli, E., and Gonnet, G. 1994. Protein identification in DNA databases by peptide mass fingerprinting. *Protein Sci.* 3:1347-1350.
- Jensen, O.N., Podtelejnikov, A., and Mann, M. 1996a. Delayed extraction improves specificity in database searches by matrix-assisted laser desorption/ionization peptide maps. *Rapid Commun. Mass Spectrom.* 10:1371-1378.
- Jensen, O.N., Vorm, O., and Mann, M. 1996b. Sequence patterns produced by incomplete enzymatic digestion or one-step Edman degradation of peptide mixtures as probes for protein database searches. *Electrophoresis* 17:938-944.
- Jungblut, P., Thiede, B., Zimnyardt, U., Muller, E.C., Scheler, C., Wittmann-Liebold, B., and Otto, A. 1996. Resolution power of 2-dimensional electrophoresis and identification of proteins from gels. *Electrophoresis* 17:839-847.
- Kassel, D.B., Consler, T.G., Shalaby, M., Sekhri, P., Gordon, N., and Nadler, T. 1995. Direct coupling of an automated 2-dimensional microcolumn affinity chromatography-capillary HPLC system with mass spectrometry for biomolecule analysis. In *Techniques in Protein Chemistry VI*. (J.W. Crabb, ed.) pp. 39-46. Academic Press, San Diego.
- Klose, J. 1975. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues: A novel approach to testing for induced point mutations in mammals. *Humangenetik* 26:231-243.
- Krishna, R.G. and Wold, F. 1993. Post-translational modification of proteins. *Adv. Enzymol. Relat. Areas Mol. Biol.* 67:265-298.
- Kuster, B., Wheeler, S.F., Hunter, A.P., Dwek, R.A., and Harvey, D.J. 1997. Sequencing of N-linked oligosaccharides directly from protein gels: In-gel deglycosylation followed by matrix-assisted laser desorption/ionization mass spectrometry and normal-phase high-performance liquid chromatography. *Anal. Biochem.* 250:82-101.
- Lamond, A.I. and Mann, M. 1997. Cell biology and the genome projects—a concerted strategy for characterizing multiprotein complexes by using mass spectrometry. *Trends Cell Biol.* 7:139-142.
- Liao, P.C., Leykam, J., Andrews, P.C., Gage, D.A., and Allison, J. 1994. An approach to locate phosphorylation sites in a phosphoprotein: Mass mapping by combining specific enzymatic degradation with matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Biochem.* 219:9-20.
- Loo, J.A. 1997. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spec. Rev.* 16:1-23.
- Loo, R.R.O., Stevenson, T.I., Mitchell, C., Loo, J.A., and Andrews, P.C. 1996. Mass-spectrometry of proteins directly from polyacrylamide gels. *Anal. Chem.* 68:1910-1917.
- Loo, R.R.O., Mitchell, C., Stevenson, T.I., Martin, S.A., Hines, W., Juhasz, P., Patterson, D., Peltier, J., Loo, J.A., and Andrews, P.C. 1997. Sensitivity and mass accuracy for proteins analyzed directly

- from polyacrylamide gels: Implications for proteome mapping. *Electrophoresis* 18:382-390.
- Mann, M. and Wilm, M. 1994. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal. Chem.* 66:4390-4399.
- Matsui, N.M., Smith, D.M., Clauser, K.R., Fichman, J., Andrews, L.E., Sullivan, C.M., Burlingame, A.L., and Epstein, L.B. 1997. Immobilized pH gradient two-dimensional gel electrophoresis and mass spectrometric identification of cytokine-regulated proteins in ME-180 cervical carcinoma cells. *Electrophoresis* 18:409-417.
- McCormack, A.L., Schieltz, D.M., Goode, B., Yang, S., Barnes, G., Drubin, D., and Yates, J.R. III. 1997. Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femtomole level. *Anal. Chem.* 69:767-776.
- Mortz, E., O'Connor, P.B., Roepstorff, P., Kelleher, N.L., Wood, T.D., McLafferty, F.W., and Mann, M. 1996. Sequence tag identification of intact proteins by matching tandem mass spectral data against sequence data bases. *Proc. Natl. Acad. Sci. U.S.A.* 93:8264-8267.
- Nelson, R.W., Krone, J.R., Bieber, A.L., and Williams, P. 1995. Mass spectrometric immunoassay. *Anal. Chem.* 67:1153-1158.
- O'Farrell, P.H. 1975. High-resolution two-dimensional gel electrophoresis of proteins. *J. Biol. Chem.* 250:4007-4021.
- Opiteck, G.J., Lewis, K.C., Jorgenson, J.W., and Anderegg, R.J. 1997a. Comprehensive on-line LC/LC/MS of proteins. *Anal. Chem.* 69:1518-1524.
- Opiteck, G.J., Jorgenson, J.W., and Anderegg, R.J. 1997b. Two-dimensional SEC/RPLC coupled to mass spectrometry for the analysis of peptides. *Anal. Chem.* 69:2283-2291.
- Pappin, D.J.C., Rahman, D., Hansen, H.F., Bartlett-Jones, M., Jeffery, W., and Bleasby, A.J. 1995. Chemistry, mass spectrometry and peptide-mass databases: Evolution of methods for the rapid identification and mapping of cellular proteins. *In Mass Spectrometry in the Biological Sciences.* (A.L. Burlingame, and S.A. Carr, eds.) pp. 135-150. Humana Press, Totowa, N.J.
- Patterson, S.D. 1994. From electrophoretically separated protein to identification: Strategies for sequence and mass analysis. *Anal. Biochem.* 221:1-15.
- Patterson, S.D. 1995. Matrix-assisted laser-desorption/ionization mass spectrometric approaches for the identification of gel-separated proteins in the 5-50 pmol range. *Electrophoresis* 16:1104-1114.
- Patterson, S.D. 1997. Identification of low- to sub-picomolar quantities of electrophoretically separated proteins: Towards protein chemistry in the post-genome era. *Biochem. Soc. Trans.* 25:255-262.
- Patterson, S.D. and Aebersold, R. 1995. Mass spectrometric approaches for the identification of gel-separated proteins. *Electrophoresis* 16:1791-1814.
- Patterson, S.D., Thomas, D., and Bradshaw, R.A. 1996. Application of combined mass spectrometry and partial amino acid sequence to the identification of gel-separated proteins. *Electrophoresis* 17:877-891.
- Pennington, S.R., Wilkins, M.R., Hochstrasser, D.F., and Dunn, M.J. 1997. Proteome analysis: From protein characterization to biological function. *Trends Cell Biol.* 7:168-173.
- Resing, K.A., Johnson, R.S., and Walsh, K.A. 1995. Mass spectrometric analysis of 21 phosphorylation sites in the internal repeat of rat profilaggrin, precursor of an intermediate filament associated protein. *Biochemistry* 34:9477-9487.
- Robinson, C.V., Gross, M., Eyles, S.J., Ewbank, J.J., Mayhew, M., Hartl, F.U., Dobson, C.M., and Radford, S.E. 1994. Conformation of GroEL-bound alpha-lactalbumin probed by mass spectrometry. *Nature* 372:646-651.
- Salmeron, A., Ahmad, T.B., Carlile, G.W., Pappin, D., Narsimhan, R.P., and Ley, S.C. 1996. Activation of MEK-1 and SEK-1 by Tpl-2 proto-oncogene, a novel MAP kinase kinase kinase. *EMBO J.* 15:817-826.
- Scheele, G.A. 1975. Two-dimensional gel analysis of soluble proteins. Characterization of guinea pig exocrine pancreatic proteins. *J. Biol. Chem.* 250:5375-5385.
- Schnolzer, M., Jedrzejewski, P., and Lehmann, W.D. 1996. Protease-catalyzed incorporation of O-18 into peptide-fragments and its application for protein sequencing by electrospray and matrix-assisted laser desorption/ionization mass spectrometry. *Electrophoresis* 17:945-953.
- Schreiner, M., Strupat, K., Lottspeich, F., and Eckerskorn, C. 1996. Ultraviolet matrix-assisted laser-desorption ionization-mass spectrometry of electroblotted proteins. *Electrophoresis* 17:954-961.
- Schuhmacher, M., Glocker, M.O., Wunderlin, M., and Przybylski, M. 1996. Direct isolation of proteins from sodium dodecyl sulfate-polyacrylamide gel-electrophoresis and analysis by electrospray-ionization mass-spectrometry. *Electrophoresis* 17:848-854.
- Shevchenko, A., Jensen, O.N., Podtelejnikov, A.V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko, A., Boucherie, H., and Mann, M. 1996. Linking genome and proteome by mass spectrometry: Large-scale identification of yeast proteins from two-dimensional gels. *Proc. Natl. Acad. Sci. U.S.A.* 93:14440-14445.
- Shevchenko, A., Chernushevich, I., Ens, W., Standin, K.G., Thomson, B., Wilm, M., and Mann, M. 1997. Rapid "de novo" peptide sequencing by a combination of nanoelectrospray, isotopic labeling and a quadrupole/time-of-flight mass spectrometer. *Rapid Commun. Mass Spectrom.* 11:1015-1024.

- Strupat, K., Karas, M., Hillenkamp, F., Eckerskorn, C., and Lottspeich, F. 1994. Matrix-assisted laser desorption ionization mass spectrometry of proteins electrophoresed after polyacrylamide gel electrophoresis. *Anal. Chem.* 66:464-470.
- Szczepanowska, J., Zhang, X., Herring, C.J., Qin, J., Korn, E.D., and Brzeska, H. 1997. Identification by mass spectrometry of the phosphorylated residue responsible for activation of the catalytic domain of myosin I heavy chain kinase, a member of the PAK/STE20 family. *Proc. Natl. Acad. Sci. U.S.A.* 94:8503-8508.
- Takach, E.J., Hines, W.M., Patterson, D.H., Juhasz, P., Falick, A.M., Vestal, M.L., and Martin, S.A. 1997. Accurate mass measurements using MALDI-TOF with delayed extraction. *J. Prot. Chem.* 16:363-369.
- Tang, Q., Harrata, A.K., and Lee, C.S. 1997. Two-dimensional analysis of recombinant *E. coli* proteins using capillary isoelectric focusing electrospray ionization mass spectrometry. *Anal. Chem.* 69:3177-3182.
- Tempst, P., Erdjument-Bromage, H., Casteels, P., Geromanos, S., Lui, M., Powell, M., and Nelson, R.W. 1995. MALDI-TOF mass spectrometry in the protein biochemistry lab: From characterization of cell-cycle regulators to the quest for novel antibiotics. In *Mass Spectrometry in the Biological Sciences*. (A.L. Burlingame, and S.A. Carr, eds.) pp. 105-133. Humana Press, Totowa, N.J.
- Tu, G.-F., Reid, G.E., Zhang, J.-G., Moritz, R.L., and Simpson, R.J. 1995. C-terminal extension of truncated recombinant proteins in *Escherichia coli* with a 10S RNA decapeptide. *J. Biol. Chem.* 270:9322-9326.
- Vestal, M.L., Juhasz, P., and Martin, S.A. 1995. Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 9:1044-1050.
- Vestling, M.M. and Fenselau, C. 1994a. Poly(vinylidene difluoride) membranes as the interface between laser desorption mass spectrometry, gel electrophoresis, and in situ proteolysis. *Anal. Chem.* 66:471-477.
- Vestling, M.M. and Fenselau, C.C. 1994b. Protease digestions on PVDF membranes for matrix-assisted laser desorption mass spectrometry. In *Techniques in Protein Chemistry V*. (J.W. Crabb, ed.) pp. 59-67. Academic Press, San Diego.
- Vestling, M.M. and Fenselau, C. 1995. Surfaces for interfacing protein gel electrophoresis directly with mass spectrometry. *Mass Spectrom. Rev.* 14:169-178.
- Watts, J.D., Affolter, M., Krebs, D.L., Wange, R.L., Samelson, L.E., and Aebersold, R. 1994. Identification by electrospray ionization mass spectrometry of the sites of tyrosine phosphorylation induced in activated Jurkat T cells on the protein tyrosine kinase ZAP-70. *J. Biol. Chem.* 269:29520-29529.
- Watts, J.D., Krebs, D.L., Wange, R.L., Samelson, L.E., and Aebersold, R. 1996. Electrospray ionization mass spectrometric investigation of signal transduction pathways: Determination of sites of inducible protein phosphorylation in activated T-cells. In *Biochemical and Biotechnological Applications of Electrospray Ionization Mass Spectrometry* (A.P. Snyder, ed.) pp. 381-407. ACS Symposium Series, Washington, D.C.
- Wilkins, M.R., Pasquali, C., Appel, R.D., Ou, K., Golaz, O., Sanchez, J.C., Yan, J.X., Gooley, A.A., Hughes, G., Humphery-Smith, I., Williams, K.L., and Hochstrasser, D.F. 1996. From proteins to proteomes—Large-scale protein identification by 2-dimensional electrophoresis and amino acid analysis. *Bio/Technology* 14:61-65.
- Wilm, M.S. and Mann, M. 1994. Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *Int. J. Mass Spectrom. Ion Process.* 136:167-180.
- Yates, J.R. III, Eng, J.K., McCormack, A.L., and Schieltz, D. 1995. Method to correlate tandem mass-spectra of modified peptides to amino-acid-sequences in the protein database. *Anal. Chem.* 67:1426-1436.

Contributed by Scott D. Patterson
Amgen, Inc.
Thousand Oaks, California

Difference Gel Electrophoresis (DIGE) Using CyDye DIGE Fluor Minimal Dyes

UNIT 10.23

DIGE (difference gel electrophoresis) is a technique used for labeling protein samples, prior to electrophoresis, with different fluorescent dyes that have distinct spectral characteristics, to analyze quantitative differences between multiple samples on the same gel. In this technique for multiplex analysis of related but different protein samples, the labeled samples are mixed together and then subjected to gel electrophoresis. The reagents and technical advice to perform this technique using CyDyes are available from Amersham Biosciences (now part of GE Healthcare; http://www4.amershambiosciences.com/aptrix/upp01077.nsf/content/na_homepage). Three different CyDye DIGE fluor minimal dyes are available for DIGE analysis: CyDye DIGE fluor Cy2, CyDye DIGE fluor Cy3, and CyDye DIGE fluor Cy5. Two-dimensional (2-D) gel electrophoresis (UNITS 10.3 & 10.4) is usually carried out for multiplex analysis of CyDye-labeled protein samples. However, prior to carrying out the 2-D separation, it is important to check the efficiency of labeling of protein samples by one-dimensional (1-D) sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (SDS-PAGE; UNIT 10.2A). Other potential applications of 1-D PAGE of CyDye-labeled proteins are discussed in Background Information.

LABELING OF PROTEINS WITH CyDye DIGE FLUOR MINIMAL DYES

**BASIC
PROTOCOL**

There are three different steps that need to be performed for labeling of proteins with CyDye DIGE fluor minimal dyes: (1) preparation of the CyDye DIGE fluor minimal dye stock solution, (2) dilution of the CyDye DIGE fluor minimal dye stock solution, and (3) labeling of protein samples with CyDye DIGE fluor minimal dyes. The same protocol is applicable to CyDye DIGE fluor Cy2, CyDye DIGE fluor Cy3, and CyDye DIGE fluor Cy5. The entire labeling procedure can be carried out on the same day or on different days.

The protocol described here is according to the manufacturer's instructions for the Ettan DIGE system (Amersham Biosciences).

Materials

CyDye DIGE fluor minimal dyes (store at -20°C in dark)
Dimethyl formamide (DMF)
Protein sample (see Support Protocol)
10 mM lysine
2 \times gel loading buffer (see recipe)
5% to 15% (w/v) acrylamide gradient gel (UNIT 10.2A)
Low-fluorescence glass plates (Amersham Biosciences)
Typhoon variable mode imager 9400 series (Amersham Biosciences) *or*
charge-coupled device (CCD)-based imaging system (AC1 Autochemisystem
from UVP) with UV transilluminator (UVP)
ImageQuant software (Amersham Biosciences) or similar software
Additional reagents and equipment for 1-D SDS-PAGE (UNIT 10.2A)

Prepare CyDye DIGE fluor minimal dye stock solution

1. Remove CyDye DIGE fluor minimal dye tubes from -20°C and allow tubes to reach room temperature (in the dark) before opening.

Absorption of moisture will lead to the hydrolysis of the dyes.

Each tube contains a known amount of a specific CyDye DIGE fluor minimal dye.

**Analysis of
Proteins**

10.23.1

2. Add required volume of DMF to each tube of stock CyDye DIGE fluor minimal dye to reconstitute the dye to a final concentration of 1 nmol/ μ l.
3. Close the cap of microcentrifuge tube containing dye, vortex vigorously for 30 sec to dissolve the dye, and centrifuge 30 sec at $12,000 \times g$ in a tabletop microcentrifuge. Store the stock solution of the dyes at -20°C . Wrap tubes with aluminum foil.

Dilute stock solution

4. Dilute the stock solution of each dye to the required working dye concentration before using in the labeling reaction. Briefly microcentrifuge the dye stock solution (30 sec) in a microcentrifuge at room temperature.

This is to prepare 5 μ l of working dye solution (400 pmol/ μ l). Follow the same ratio for preparing larger volumes of working dye solution.

5. Add 3 μ l DMF to a microcentrifuge tube. Add 2 μ l of dye stock solution to the tube. Ensure that all dye is removed from the pipet tip by pipetting up and down several times into the working dye solution.

The commonly recommended dye/protein ratio is 400 pmol of dye for labeling 50 μ g of protein. Although a different dye/protein ratio can be used, it must be tested by 1-D PAGE for optimum labeling. Since the labeling efficiencies of the CyDyes depend on the dye/protein ratio, it is recommended that for the same experiment, the same dye/protein ratio should be used for all samples.

The working dye solution is only stable for 2 weeks at -20°C .

Label protein samples

6. Add the required volume of protein sample (see Support Protocol) containing 50 μ g of protein to a microcentrifuge tube. Add 1 μ l of the working dye solution to the microcentrifuge tube containing the protein sample.

For labeling with a CyDye DIGE fluor minimal dye, the recommended concentration of the protein sample is 5 to 10 mg/ml. However, protein samples of 1mg/ml have been labeled successfully. If necessary, the process can be scaled up for labeling a larger amount of samples. All new protein lysates or samples containing new chemical components should always be checked for successful labeling.

For this example, 50 μ g of protein is labeled with 400 pmol of dye.

7. Vortex to mix the dye and protein solution thoroughly and microcentrifuge for 30 sec.
8. Wrap tubes with aluminum foil to carry out the labeling reaction in the dark. Leave on ice for 10 min to complete the labeling reaction.
9. Stop the reaction by adding 1 μ l of 10 mM lysine. Mix and briefly microcentrifuge.
10. Immediately analyze the sample by 1-D PAGE (steps 11 to 15).

If necessary, the aluminum foil-wrapped tubes may be stored for up to 3 months at -70°C (in dark) before performing gel electrophoresis.

Analyze samples by 1-D PAGE

11. Transfer the labeled sample from step 9 to a clean microcentrifuge tube. Add an equal volume of $2\times$ gel loading buffer to the CyDye-labeled samples.
12. Heat the samples 5 min at 95°C to carry out the reduction of the protein samples.
13. Pour a 5% to 15% acrylamide gradient gel using low-fluorescence glass plates.
14. Load the serially diluted protein samples (5 to 25 μ g) in successive lanes on the gel. Carry out the electrophoresis until the bromophenol blue (the dye front) almost reaches the bottom of the gel.

Table 10.23.1 Detection of Fluorescence of CyDye DIGE Fluor Minimal Dyes: Laser Excitation Source and Emission Filters^a

Fluorescent dye	Laser excitation source (nm)	Emission filter (nm) ^b
CyDye DIGE fluor Cy2 minimal dye	Blue2 (488)	520 BP 40
CyDye DIGE fluor Cy3 minimal dye	Green (532)	580 BP 30
CyDyeDIGE fluor Cy5 minimal dye	Red (633)	670 BP 30

^aGade et al. (2003).

^bBP, bandpass.

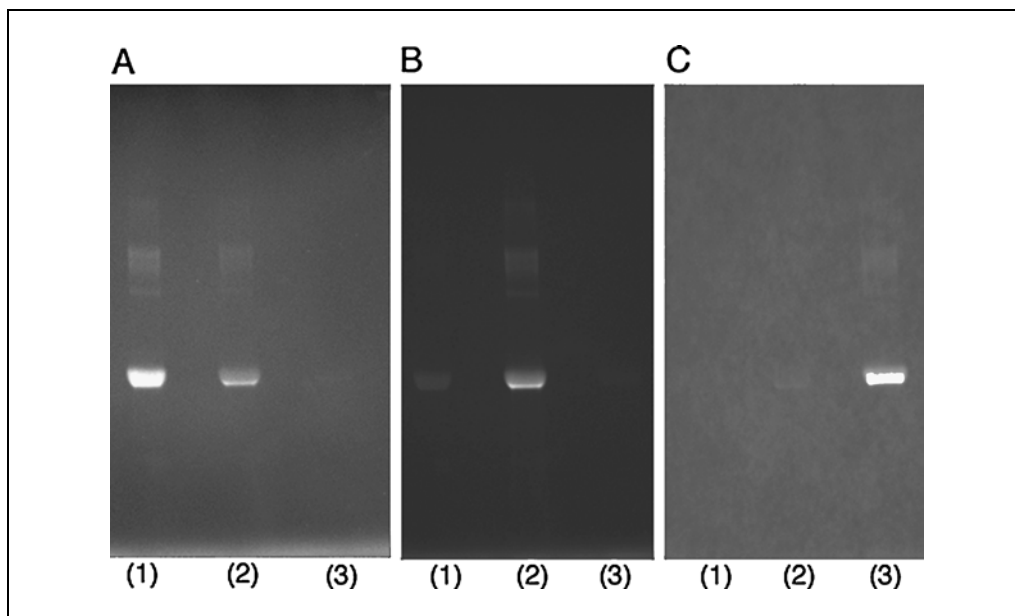


Figure 10.23.1 This photograph of gels depicts 50 μg of BSA labeled with 400 pmol of Cy2, Cy3, or Cy5 using CyDye DIGE fluor minimal dyes. Lanes 1, 2, and 3 were loaded with 5 μg BSA labeled with Cy2, Cy3, and Cy5, respectively, on a polyacrylamide gel (5% to 12.5% w/v acrylamide) and subjected to 1-D PAGE. Images were captured using a UV-transilluminating light source and CCD camera (AC1 Autochemisystem, UVP). The following emission filters were used for the detection of the CyDye-labeled proteins: green filter (515 to 570 nm) for Cy2-labeled BSA (**A**); red filter (570 to 640 nm) for Cy3-labeled BSA (**B**); and blue filter (650 to 690 nm) for Cy 5-labeled BSA (**C**).

15. After electrophoresis, take the gel cassette out of the apparatus and thoroughly clean the outside of the glass plates with deionized water. If using plastic backing, take that off before taking images of the gel.
16. Take the images of the gel at the appropriate wavelength using a laser-based scanner such as a Typhoon variable mode imager.

The excitation and emission wavelengths used for all three CyDyes are listed in Table 10.23.1. A 100- μm pixel resolution is recommended to obtain accurate image information.

Alternatively, images can be captured using a UV transilluminator, appropriate emission filters, and a CCD camera (Fig. 10.23.1).

17. Quantify the labeling of each protein sample using ImageQuant software or similar software available from other vendors.

PREPARATION OF SAMPLE FOR LABELING

Prior to carrying out the labeling reaction, the samples have to be prepared in a way compatible with the labeling reaction. Since most of the biological samples to be analyzed by DIGE are usually of cellular origin, this protocol focuses on such samples.

Materials

Cells of interest: 20 to 50 ml of microbial culture medium (e.g., *Escherichia coli*, $\sim 1 \times 10^8$ cells/ml), 20 to 50 ml of tissue culture medium for mammalian cells ($\sim 5 \times 10^5$ cells/ml), or ~ 200 mg of tissue

Cell wash buffer (see recipe)

Lysis buffer (see recipe)

1. Wash cells with cell wash buffer.

This wash allows the removal of the cell culture medium or tissue fluids, depending on how the cells were collected. It will not lyse the cells.

2. Lyse cells in an appropriate lysis buffer.

The pH of the lysis buffer is important since maximum labeling efficiency is obtained at a pH between 8.0 and 9.0. Since the pH of the Tris-Cl buffer is temperature dependent and the labeling reaction is carried out in the cold, it is important to check and adjust the pH when the solution is cold.

3. Collect cells by centrifuging 10 min at $12,000 \times g$, 4°C . Discard any liquid present.

4. Resuspend the cell pellet in 1 ml of cell wash buffer. Centrifuge 10 min at $12,000 \times g$, 4°C , and repeat this wash step two additional times.

5. Resuspend the final cell pellet in the lysis buffer, leave on ice for 10 min, and then sonicate intermittently to lyse the cells.

6. Centrifuge cell lysate 10 min at $12,000 \times g$, 4°C , and collect the supernatant. Discard the pellet.

The pH of the supernatant should be checked by spotting a few microliters on a pH indicator strip. If the pH is below 8.0, adjust, e.g., by careful addition of 50 mM sodium hydroxide.

7. Store the cell lysate up to several months at -70°C until further use.

Before performing the labeling reaction, the protein concentration of the lysate should be determined with a protein assay kit that is detergent-, urea-, and thiourea-compatible.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

Cell wash buffer

10 mM Tris-Cl, pH 8.0

5 mM magnesium acetate

Store for several months at 4°C or at room temperature as long as there is no visible microbial growth.

Gel loading buffer, $2\times$

120 mM Tris-Cl, pH 6.8

20% (v/v) glycerol

4% (w/v) SDS

200 mM dithiothreitol

Store for several months at room temperature.

Lysis buffer

7 M urea
2 M thiourea
30 mM Tris·Cl, pH 8.5
4% (w/v) CHAPS
Alternatively, use:
8 M urea
30 mM Tris·Cl, pH 8.5
4% (w/v) CHAPS

Store in small aliquots for several months at 4°C.

COMMENTARY

Background Information

In 1997, Ünlü and co-workers first described the DIGE technique to profile proteins (Ünlü et al., 1997). It involves the pre-electrophoretic labeling of samples with one of three spectrally distinct fluors, cyanine-2 (Cy2), cyanine-3 (Cy3), or cyanine-5 (Cy5). More than one set of samples is prepared, each of which is first labeled with a different CyDye, and then combined. The sample containing the mixture is then subjected to gel electrophoresis and viewed individually by scanning the gel at different wavelengths, thus circumventing problems with spot matching between gels. Image analysis programs are then used to generate volume ratios for each spot. Volume ratios essentially describe the intensity of a

particular spot in each test sample, and thus enable expression differences to be identified and quantified. This methodology has since been commercialized by Amersham Biosciences and all three CyDye DIGE fluors are available.

CyDye DIGE fluors have a reactive *N*-hydroxysuccinimide ester group that forms a covalent bond with the ϵ -amino group of lysine side chains (Fig. 10.23.2). The content of lysine is generally very high in most proteins such that an excessive amount of dye is required to direct the reaction to completion. When the proteins undergo such extensive chemical modification, they may become insoluble (Ünlü et al., 1997; Patton, 2002) and cannot be analyzed using 1-D PAGE or 2-D gel electrophoresis. However, with the

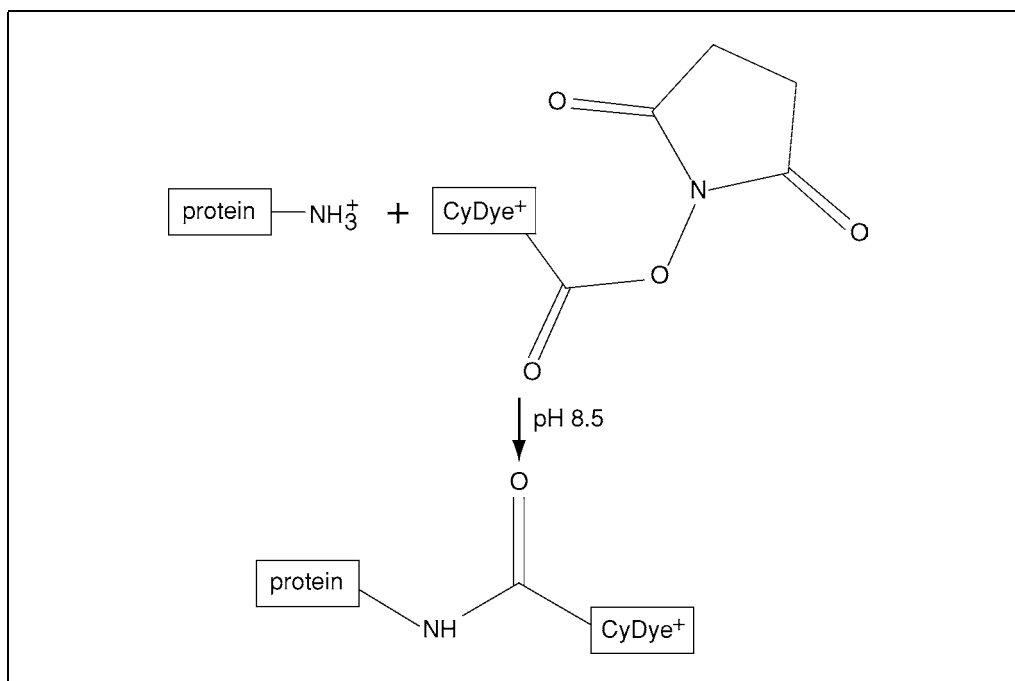


Figure 10.23.2 Schematic presentation of the reaction of CyDye DIGE fluor minimal dyes containing *N*-hydroxy succinimidyl (NHS) ester group with the ϵ -amino group of a lysine residue of a protein. A covalent amide bond is formed between the NHS ester active group of the CyDyes and the ϵ -amino group of the lysine residue of a protein.

Table 10.23.2 Absorption and Emission Maxima of CyDye DIGE Fluor Minimal Dyes

Fluorescent dye	Absorption maxima (nm)	Emission maxima (nm)
CyDye DIGE fluor Cy2 minimal dye	482	501
CyDye DIGE fluor Cy3 minimal dye	552	570
CyDye DIGE fluor Cy5 minimal dye	640	660

Table 10.23.3 Molecular Weight Added to Proteins Following Labeling with CyDye DIGE Fluor Minimal Dyes

Fluorescent dye	Molecular weight added to protein (Da)
CyDye DIGE fluor Cy2 minimal dye	435.52
CyDye DIGE fluor Cy3 minimal dye	467.69
CyDye DIGE fluor Cy5 minimal dye	465.67

minimal fluor technology, labeling reactions are designed such that the stoichiometry of protein to fluor results in only 1% to 2% of the total number of lysine residues being labeled.

All three CyDye DIGE fluor minimal dyes have different absorption and emission spectra (Table 10.23.2). Because of the different absorption and emission spectra, the CyDyes can be used for multiplexing of up to three different protein samples on the same 1-D or 2-D gel. Due to the high degree of sensitivity of CyDye labeling, a single protein as low as 125 pg can be detected following polyacrylamide gel electrophoresis. A linear response can be obtained with protein concentrations up to five orders of magnitude. Labeling with CyDyes allows a greater level of detection of protein spots than Coomassie Blue and a comparable level of detection of protein spots as with silver stain or SYPRO Ruby (Tonge et al., 2001; Gharbi et al., 2002; Gade et al., 2003).

The CyDye DIGE fluor minimal dyes are matched for charge and mass. As a result, if the same protein is labeled with different CyDyes, they will migrate to the same position following 1-D PAGE (resolved according to molecular mass) or 2-D gel electrophoresis (resolved according to isoelectric point in the first dimension and molecular mass in the second dimension). At a neutral or acidic pH, the amino acid lysine carries a net +1 charge. CyDye DIGE fluor minimal dyes also carry a +1 charge. During its reaction with the lysine residues of proteins, the intrinsic +1 charge in the dye molecule compensates for the loss of the +1 charge of the lysine side chain. Hence, the isoelectric points of the proteins do not alter significantly. Also, due to the addition of each CyDye DIGE fluor minimal dye, the mass of

the protein is altered ~500 Da (Table 10.23.3). Other than the low-molecular-weight proteins, this mass shift is not significant on 1-D PAGE or 2-D gel electrophoresis.

The advantage of multiplex analysis using CyDyes is the ability to use the same internal standard (a pool of all the samples within an experiment) on every 1-D PAGE or 2-D gel electrophoresis. A major drawback of 2-D gel electrophoresis, where each sample is loaded on a separate gel, is the gel-to-gel variation observed with the separation profile, even between the same samples. This can interfere with the interpretation of the data. Inclusion of an internal standard and running different samples on the same gel is particularly useful to overcome the above problem.

Although DIGE is mainly used for 2-D gel electrophoresis, there are some definitive applications of 1-D PAGE using DIGE technology. For example, before carrying out a 2-D gel electrophoresis using DIGE technology, it is necessary to optimize the protein/dye ratio for the labeling reaction and test the efficiency of the labeling reaction using 1-D PAGE. This is particularly important for labeling new samples or samples containing any new chemical component.

One-dimensional PAGE using DIGE technology is also good for comparing expression of recombinant proteins under different sets of experimental parameters, such as variation of growth conditions and use of different microbial strains. It can also be used for comparative analysis of relative protein profiles in fractions obtained from chromatographic separation of a mixture of cellular extracts.

The technique is also useful for comparing differential expression of proteins in

isolated fractions of cellular extracts that are less complex in terms of number of proteins present or are difficult to separate by 2-D gel electrophoresis. Membrane fractions, such as bacterial outer-membrane protein fractions from different strains (e.g., pathogenic and non-pathogenic strains of the same bacterial species), seem to be well suited for 1-D PAGE using DIGE technology. Identification of these classes of proteins is very important for drug and vaccine discovery. These proteins are often underrepresented in 2-D gel electrophoresis due to their insolubility under conditions used in the isoelectric focusing step of 2-D gel electrophoresis. On the contrary, use of SDS in 1-D PAGE aids in the solubilization of membrane proteins. Due to compatibility of CyDye minimal labeling technique with mass spectrometry, the proteins of interest can be subsequently characterized by mass spectrometry following trypsin digestion of the gel bands of interest. In this connection, 1-D PAGE, rather than 2-D gel electrophoresis, followed by mass spectrometric analysis was used to identify a large number of proteins in the membrane fractions of *Streptococcus pneumoniae* as putative vaccine candidates (Chakravarti et al., 2002).

The images obtained following 1-D PAGE using DIGE technology can be analyzed by appropriate 1-D image analysis software such as ImageQuant (Amersham Biosciences) and expression of protein(s) can be compared. Subsequently, the bands of interest can be excised, subjected to trypsin digestion and identified by mass spectrometry.

Critical Parameters

General considerations

Since CyDyes are transported as dry powders in microcentrifuge tubes, the powder may spread around the inside surface of the tube during transportation. Therefore, in order to achieve maximum recovery, the reconstituted dye solution should be pipetted around the tube (and the lid) to resuspend all of the dye powder.

Once reconstituted in DMF, the dye is stable for 2 months at -20°C or the date of expiry, whichever is earlier. The dye should be stored in the absence of light. After its use, it should be returned to -20°C as soon as possible to minimize its exposure to the room temperature. The tubes should be wrapped with aluminum foil.

High quality anhydrous DMF ($\text{H}_2\text{O} \leq 0.005\%$, purity $\geq 99.8\%$) should be used. Care should be taken so that there is no water contamination. When the bottle containing DMF

is opened, it starts to generate amine compounds, which react with the CyDyes. This will decrease the concentration of dye available for protein labeling. A new bottle of DMF should be opened every 3 months. The same quality of DMF should also be used for dilution of the stock solution.

Wear powder-free laboratory gloves. The powder present in the gloves may fluoresce and interfere with the quality of the image.

Although different protein assay kits are available, it is important to use kits appropriate for samples containing detergents (as well as urea and thiourea if the lysis solution contains these components).

Labeling of the proteins

Care should be taken about the following: use the recommended protein concentration of 5 to 10 mg/ml; maintain the recommended pH of 8 to 9. A buffer such as Tris or bicarbonate is useful for maintaining the pH range. The buffer concentration should be ~ 30 mM (a high buffer concentration may affect the isoelectric focusing step in 2-D gel electrophoresis). If the appropriate buffer is not used, it may interfere with the labeling efficiency. Also, presence of any primary amine in the solution before labeling should be avoided since these will compete with the proteins for the dye. The presence of the following compounds is not recommended: reducing agents such as >2 mg/ml dithiothreitol, >1 mM (Tris-[2-carboxyethyl]phosphine) TCEP, β -mercaptoethanol at any concentration; detergents are compatible with the labeling reactions at the following concentrations: up to 0.25% for SDS, up to 1% for NP-40, and $<1\%$ for Triton X-100.

Precautions following SDS-PAGE analysis of labeled proteins with CyDye DIGE fluor minimal dyes

Since the proteins in the gel diffuse after electrophoresis is complete, scan the gels and capture the images of the gels immediately following SDS-PAGE analysis and keep the gels between the glass plates. Alternatively, the gels should be stored in SDS-PAGE electrophoresis running buffer in the dark at ambient temperature. However, even those gels should be scanned on the same day. If stored for a longer period of time, a storage temperature of 4°C is recommended and the gels should be kept moist to avoid tearing. Before scanning or taking the image, the gels should attain room temperature. However, the possibility of diffusion of the proteins will still remain. Although the problem of diffusion can be avoided by fixing

the proteins, it is not recommended to do the fixing before scanning because it can interfere with subsequent image analysis using appropriate software such as DeCyder Differential Analysis Software.

Most plastic materials will have intense fluorescence at the wavelength used for fluorescent dyes. Hence, it is recommended that gels with any kind of plastic backing should not be used. For spot picking, the 2-D gel may be placed on a low-fluorescence glass plate. If excision of the spots (2-D gel electrophoresis) or bands (1-D PAGE) is not necessary, the gel can be taken off the plastic backing and images can be captured.

Troubleshooting

If optimal labeling of protein(s) is not observed, the following points should be taken into consideration.

The DMF used should be of high quality as mentioned in Critical Parameters. The presence of moisture in the DMF may lead to the hydrolysis of the CyDye fluor minimal dyes, thereby reducing their effective concentrations and, hence, the labeling efficiencies. Similarly, exposure to light will cause photodegradation of the dyes, therefore, reducing their labeling efficiencies.

An appropriate protein assay reagent compatible with detergent (as well as urea and thiourea if used in the lysis solution) should be used for accurate estimation of protein concentrations to obtain efficient labeling. See UNIT 10.1A for additional details.

The pH of the solution should be checked before labeling and should be ≥ 8.0 .

The samples should not contain any primary amines or thiol reagents at a concentration greater than that mentioned earlier.

Anticipated Results

If one follows the protocol as described above, following 1-D PAGE, the protein(s) labeled with Cy2, Cy3, and Cy5 using CyDye DIGE fluor minimal dyes can be visualized when the gels are scanned using a suitable scanner such as the Typhoon variable mode imager 9400 series (Amersham Biosciences) or the AC1 Autochemi System (UVP; see Fig. 10.23.1).

Time Considerations

Making the CyDye solution should take 15 min. Preparing the protein samples (see Support Protocol) may take >1 hr. The labeling reaction requires 45 min, and the separation

of the labeled proteins by gel electrophoresis can take 45 min to 5 hr, depending upon the gel format used. Scanning and image capture analysis will take 15 min to 1 hr or longer depending upon the type of scanner used to capture the images.

Acknowledgments

The authors would like to acknowledge the generous support provided by the Keck Graduate Institute of Applied Life Sciences, Claremont, CA and UVP, Inc., Upland, CA. The authors would also like to acknowledge the support from the Arnold and Mabel Beckman Foundation, the Ralph M. Parsons Foundation, and the Henry L. Guenther Foundation. In addition, the authors would like to acknowledge Melissa Louie for performing the experiment described in Figure 10.23.1.

Literature Cited

- Chakravarti, D.N., Russell, D.P., Wooters, J.L., Greene, B.A., Masi, A.W., and Zagursky, R.J. Novel *Streptococcus pneumoniae* Open Reading Frames Encoding Polypeptide Antigens and Uses Thereof. International patent application number WO 02/083855, published October 24, 2002.
- Gade, D., Thiermann, J., Markowsky, D., and Rabus, R. 2003. Evaluation of two-dimensional difference gel electrophoresis for protein profiling. *J. Mol. Microbiol. Biotechnol.* 5:240-251.
- Gharbi, S., Gaffney, P., Yang, A., Zvelebil, M.J., Cramer, R., Waterfield, M.D., and Timms, J.F. 2002. Evaluation of two-dimensional differential gel electrophoresis for proteomic expression analysis of a model breast cancer cell system. *Mol. Cell. Proteomics* 1:91-98.
- Patton, W.F. 2002. Detection technologies in proteome analysis. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* 771:3-31.
- Tonge, R., Shaw, J., Middleton, B., Rowlinson, R., Rayner, S., Young, J., Pognan, F., Hawkins, E., Currie, I., and Davison, M. 2001. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. *Proteomics* 1:377-396.
- Ünlü, M., Morgan, M.E., and Minden, J.S. 1997. Difference gel electrophoresis: A single gel method for detecting changes in protein extracts. *Electrophoresis* 18:2071-2077.

Contributed by Bulbul Chakravarti and
Deb N. Chakravarti
Keck Graduate Institute of Applied Life
Sciences
Claremont, California

Sean R. Gallagher
UVP, Inc.
Upland, California

Solution Radioimmunoassay of Proteins and Peptides

UNIT 10.24

**BASIC
PROTOCOL**

Radioimmunoassays (RIAs) allow the identification and quantitation of molecules with endocrinological, immunological, neurological, and toxicological functions. In particular, they are useful for assaying specific proteins (or peptides) in a complex mixture (e.g., plasma, serum, cerebrospinal fluid, urine, tissue culture media, cell lysates, or tissue homogenates). There are two formats for RIAs: solution (homologous) and solid-phase (UNIT 11.17). Solid-phase assays do not work well with small peptides, since the antigen must have two antigenic sites (one for each antibody used in the assay). This unit presents a solution RIA using a double-antibody/polyethylene glycol (PEG) method, in which a secondary antibody and PEG are used to precipitate an antigen bound to a primary antibody. The amount of protein or peptide antigen in an unknown sample is determined based on its ability to compete with a radiolabeled protein or peptide tracer for binding to the primary anti-protein or anti-peptide antibody, as illustrated in Figure 10.24.1. The protocol described here is based on a method developed by LINCO Research (St. Charles, Missouri; see Internet Resources) for measuring primate leptin in plasma, serum, and tissue culture media.

Prior to developing an RIA, a sufficient quantity of the protein (or peptide) being assayed must be isolated and purified. It will be used in the RIA as the protein tracer, and protein standards. Some of the protein is radiolabeled to yield the protein tracer. Unlabeled protein is used as the protein standards. The exact protein concentration in each of the standards must be known, since the cpm values derived from the RIA of the standards are used to construct a standard curve. This standard curve allows the determination of the concentration of the protein from any unknown sample, as long as the concentration is within the operating range of the RIA. Quality controls are used to determine inter-assay variability.

Under ideal conditions, an RIA is capable of measuring picograms (10^{-12} g) of a protein (or peptide) antigen, but only if high-specific-activity radiolabeling of the protein tracer can be achieved. Radioiodine (either ^{125}I or ^{131}I) is the label of choice because of the significant increase in radioactive disintegrations per unit time in comparison to ^3H or ^{14}C . Radioiodination is most commonly done by one of five methods. Four are chemical modifications, employing one of the following reagents: chloramine-T, IODO-GEN (1,3,4,6-tetrachloro-3 α ,6 α -diphenylglycouril; Pierce), PIB (*N*-succinimidyl 4-iodobenzoate), and Bolton-Hunter reagent (*N*-hydroxysuccinimide ester of 3-(*p*-hydroxyphenyl)propionic acid). The other method is enzymic, employing lactoperoxidase. UNIT 11.17 provides protocols for radiolabeling a standard protein (or peptide) with either chloramine-T or IODO-GEN.

Enzyme-linked immunosorbent assays (ELISAs; UNIT 11.2) are alternative methods for those wishing to avoid the use of radioactivity and in situations where a large number of tests will be performed. The cost of assay development is higher for ELISAs than for RIAs, but, for large numbers of assays, the cost of radioisotope and its disposal become significant factors.

Materials

- Assay buffer (see recipe)
- Protein standards (see recipe)
- Protein quality controls 1 and 2 (see recipe)

Contributed by John A. Smith

Current Protocols in Molecular Biology (2005) 10.24.1-10.24.12
Copyright © 2005 by John Wiley & Sons, Inc.

**Analysis of
Proteins**

10.24.1

Supplement 72

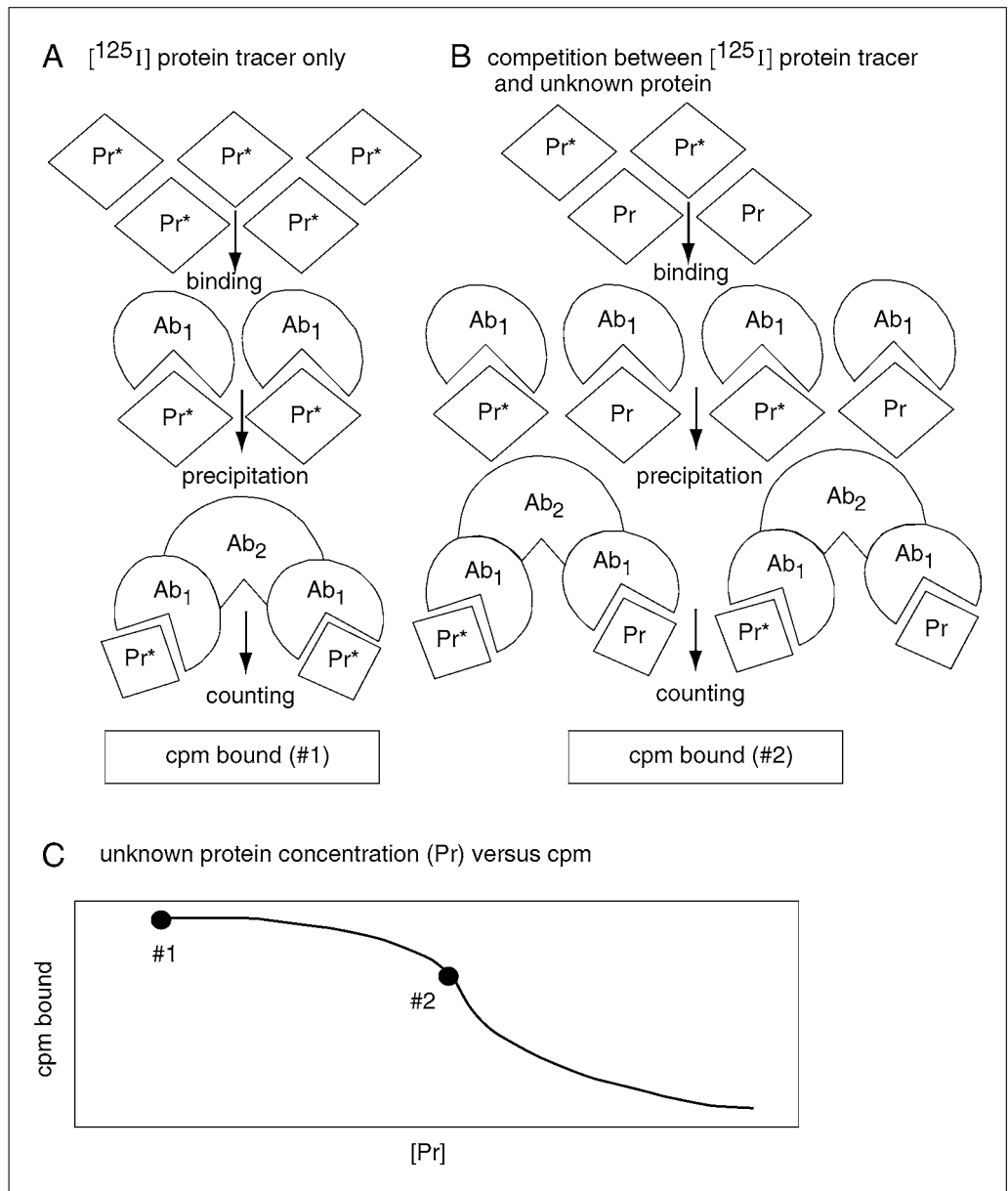


Figure 10.24.1 Solution radioimmunoassay for a protein of interest (Pr) using the double-antibody/PEG method. The protein of interest is isolated, purified, and radiolabeled to generate a protein tracer (Pr^*) that binds specifically to a primary antibody (Ab_1). The soluble $\text{Pr}^*\text{-Ab}_1$ complex is precipitated by adding a specific secondary antibody (Ab_2). The addition of a low concentration of polyethylene glycol (PEG) increases the formation of the insoluble $\text{Pr}^*\text{-Ab}_1\text{-Ab}_2$ complex. **(A)** Measurement the precipitate in the presence of Pr^* only yields the B_0 value (Table 10.24.2), which corresponds to the maximum Pr^* cpm that will bind to the amount of Ab_1 used in the assay. **(B)** In the presence of unlabeled protein (Pr) from standards or unknown samples, Pr and Pr^* compete equally for Ab_1 binding sites, and the soluble $\text{Pr}^*\text{-Ab}_1$ and $\text{Pr}\text{-Ab}_1$ complexes are precipitated equally by Ab_2 and PEG. **(C)** A plot of B values (Table 10.24.2) versus Pr concentration illustrates the principle of the RIA. As the concentration of Pr increases, the amount of Pr^* that binds to Ab_1 decreases. Such a curvilinear plot can be made linear by constructing a log-log plot of various protein concentrations (derived from a series of protein standards) versus % B/B_0 (Fig. 10.24.2). By comparing the % B/B_0 value of an unknown sample to this standard curve, the concentration of the protein in the unknown sample may be determined.

Table 10.24.1 Setup for a Sample Solution RIA^a

	Steps 2-4	Steps 5-8	Step 9	Step 10	Step 12
Tube #	Add assay buffer	Add standards/QCs/unknowns	Add [¹²⁵ I]tracer	Add primary antibody	Add precipitating secondary antibody/PEG
<i>Total protein tracer</i>					
1, 2	—	—	100 µl	—	—
<i>Nonspecific binding</i>					
3, 4	200 µl	—	100 µl	—	1.0 ml
<i>B₀</i>					
5, 6	200 µl	—	100 µl	100 µl	1.0 ml
<i>Standards</i>					
7, 8	100 µl	100 µl of 0.5 ng/ml	100 µl	100 µl	1.0 ml
9, 10	100 µl	100 µl of 1 ng/ml	100 µl	100 µl	1.0 ml
11, 12	100 µl	100 µl of 2 ng/ml	100 µl	100 µl	1.0 ml
13, 14	100 µl	100 µl of 5 ng/ml	100 µl	100 µl	1.0 ml
15, 16	100 µl	100 µl of 10 ng/ml	100 µl	100 µl	1.0 ml
17, 18	100 µl	100 µl of 20 ng/ml	100 µl	100 µl	1.0 ml
19, 20	100 µl	100 µl of 50 ng/ml	100 µl	100 µl	1.0 ml
21, 22	100 µl	100 µl of 100 ng/ml	100 µl	100 µl	1.0 ml
<i>Quality controls</i>					
23, 24	100 µl	100 µl of QC #1	100 µl	100 µl	1.0 ml
25, 26	100 µl	100 µl of QC #2	100 µl	100 µl	1.0 ml
<i>Unknowns</i>					
27, 28	100 µl	100 µl of unknown #1	100 µl	100 µl	1.0 ml
29, 30	100 µl	100 µl of unknown #2	100 µl	100 µl	1.0 ml
Others	100 µl	100 µl of unknown <i>N</i>	100 µl	100 µl	1.0 ml

^aAdapted from data derived from a primate leptin RIA kit available from LINCO Research (St. Charles, Mo.). Abbreviations: B₀, bound protein tracer in absence of competing unlabeled protein; QC, protein quality control.

Unknown samples (e.g., plasma, serum, tissue culture medium)

Protein tracer, iodinated with ¹²⁵I (see recipe)

Protein tracer buffer (see recipe)

Anti-protein antiserum (containing primary antibody; see recipe)

Precipitating secondary antibody/PEG solution (see recipe)

Assay tubes: 12 × 75–cm glass test tubes (polypropylene or polyethylene tubes may be used provided that the pellets remain compacted following the decanting step)

Refrigerated centrifuge with swinging-bucket rotor

γ counter

Log/log graph paper

Set up the assay

1. Label a series of 12 × 75–mm glass tubes as described in Table 10.24.1.
2. Pipet 200 µl assay buffer into each of the duplicate nonspecific binding tubes (labeled #3 and #4).

- Pipet 200 μl assay buffer into each of the duplicate B_0 tubes (labeled #5 and #6).

The B_0 value is the amount of radiolabeled antigen ($[^{125}\text{I}]$ protein tracer) that is bound to the amount of primary antibody used in the assay. When the unknown samples are added, the $[^{125}\text{I}]$ tracer will be displaced by antigen contributed by the unknown samples. The lower the concentration of antigen in the unknown samples, the more $[^{125}\text{I}]$ tracer that will remain bound to the primary antibody.

- Pipet 100 μl assay buffer into the duplicate standard tubes, quality control tubes, and the unknown sample tubes (i.e., tube #7 through the final tube of the assay).
- Pipet 100 μl of the 0.5 ng/ml protein standard into the duplicate tubes #7 and #8, 100 μl of the 1 ng/ml standard into tubes #9 and #10, and so forth until all standards have been dispensed.

The lowest level of protein that can be detected by the assay should be used as the lowest concentration of standard to construct the standard curve (i.e., 0.5 ng/ml in this example). The limit of the linear range of the assay determines the highest concentration, which in this assay is 100 ng/ml.

- Pipet 100 μl of quality control 1 into the duplicate tubes #23 and #24.

The quality control samples (also see Reagents and Solutions) are other complex samples with known quantities of the protein of interest, and the values obtained in this assay should agree with the previously determined quantities. Such controls are useful for inter-assay comparisons.

- Pipet 100 μl of quality control 2 into the duplicate tubes #25 and #26.
- Pipet 100 μl of each unknown sample into the appropriately numbered duplicate tubes.

If the antigen is expected to be present in high concentration in the unknown sample, or if the sample quantity is limited, smaller aliquots of sample may be used. If so, add assay buffer to bring the sample volume to 100 μl . The sample may be plasma, serum, tissue culture medium, or some other complex biological mixture.

- Pipet 100 μl of $[^{125}\text{I}]$ protein tracer in protein tracer buffer ($\sim 5,000$ to 20,000 cpm per tube) into all tubes.

The assay results illustrated in Table 10.24.2 indicate that $\sim 11,500$ cpm per tube were used.

- Pipet 100 μl primary antibody into all tubes except #1, #2, #3, and #4.

Perform the radioimmunoassay

- Vortex tubes, place in a rack, and cover with Parafilm. Incubate 24 hr at 4°C.

Covering the tubes with Parafilm prevents evaporation and contamination during the incubation.

- Pipet 1 ml cold (4°C) precipitating secondary antibody/PEG to all tubes except tubes #1 and #2.
- Vortex, then incubate 30 min at 4°C.
- Centrifuge 20 min at $2500 \times g$, 4°C, in a centrifuge with a swinging-bucket rotor.
- For all tubes except #1 and #2, invert the tubes to decant the supernatant and gently place the tubes upside-down in a wire test tube rack that has paper towel on top of the rack base. Allow the tubes to drain onto the paper towel for 1 min.

Invert the tubes only once, since the precipitant pellets are fragile. With the paper towel on top of the rack base (not on the benchtop), the excess liquid from the open ends of the tubes will wick into the paper towel.

- Count all the tubes in a γ counter for 1 min.

Required counting times will vary depending on the cpm used in the assay.

Many γ counters are capable of processing the data from assays such as the one described in this protocol. If not, the raw data can be processed using commercially available software for RIAs (e.g., Hewlett-Packard 41, Beckman EIARIA). Alternatively, the data can be manually processed as described below.

Process data and calculate results

- Average the duplicate counts for the total counts added (tubes #1 and #2), nonspecific binding (#3 and #4), total tracer bound to primary antibody (B_0 ; tubes #5 and #6), all standard dilutions (tubes #7 to #22), quality controls (tubes #23 to #26), and all unknown samples (tubes #27 and #28, #29 and #30, and so on).

If the difference between the duplicate results is >10%, repeat the assay for that sample.

- Subtract the average nonspecific binding counts (average of tubes #3 and #4) from each average count of all tubes except the total counts (tubes #1 and #2).

In this assay, nonspecific binding refers to tracer protein that remains in the tube without addition of primary or secondary antibody.

- Calculate the percent of tracer bound to primary antibody according to the equation:

$$\% \text{ tracer bound} = \{(\text{total tracer bound})/(\text{total tracer added})\} \times 100$$

See Table 10.24.2 for sample assay results. Total tracer specifically bound to primary antibody (average of tubes #5 and #6 minus average of tubes #3 and #4) equals $\sim 2,200$ cpm. The total counts of tracer added to each tube (average of tubes #1 and #2) equals $\sim 11,500$ cpm/tube. Ideally, the amount of tracer bound would equal 15% to 50%. In this case, it is $\sim 19\%$. If the amount of tracer bound does not fall within this ideal range, a primary antibody with higher affinity for the protein or peptide of interest must be prepared or obtained.

- Calculate the $\% B/B_0$ for each standard and unknown sample, where B is the net average counts (the average counts minus the average nonspecific binding, as determined in step 18) and B_0 is the total tracer bound (determined in step 19).

See sample calculations in Table 10.24.2.

- Using log/log graph paper, plot $\% B/B_0$ for each standard on the y axis and the known concentration of each standard on the x axis (see Fig. 10.24.2).

- Construct a standard curve by joining the points in a smooth curve.

A standard curve should be generated for each protein assay.

- Calculate the concentration (ng/ml) of antigen in the unknown samples by comparison with the standard curve.

If the unknown sample volumes used are less than 100 μl and a dilution was performed (see step 8), it will be necessary to account for the dilution in the calculations (e.g., if 25 μl was used, the calculated quantity of protein must be multiplied by 4).

The $\% B/B_0$ values for unknown samples #1 and #2 were 39.4 and 61.3, respectively, which correspond to 7.6 and 3.0 ng/ml (see Table 10.24.2 and Fig. 10.24.2).

- Calculate the concentration (ng/ml) of antigen in the quality control samples (tubes #23 and #24 and tubes #25 and #26) and compare to the known values.

The quality control samples (also see Reagents and Solutions) are other complex samples with known quantities of the protein of interest, and the values obtained in this assay should agree with the previously determined quantities. Such controls are useful for inter-assay comparisons.

Table 10.24.2 Calculations of B/B_0 and Protein Concentration for a Sample RIA^a

Tube #	Identification	cpm	Avg. cpm	Avg. net cpm	% B/B_0	Protein conc. (ng/ml) ^b
1	Total tracer counts	11574				
2	Total tracer counts	11449	11512			
3	Nonspecific binding	745				
4	Nonspecific binding	796	771			
5	B_0 counts	2854				
6	B_0 counts	3022	2938	2167 (B_0)		
<i>Standards</i>						
7	0.5 ng/ml	2902				
8	0.5 ng/ml	2821	2862	2091 (B)	96.5	
9	1 ng/ml	2576				
10	1 ng/ml	2654	2615	1844 (B)	85.1	
11	2 ng/ml	2273				
12	2 ng/ml	2293	2283	1512 (B)	69.8	
13	5 ng/ml	1859				
14	5 ng/ml	1727	1793	1022 (B)	47.2	
15	10 ng/ml	1536				
16	10 ng/ml	1518	1527	756 (B)	34.9	
17	20 ng/ml	1238				
18	20 ng/ml	1283	1261	490 (B)	22.6	
19	50 ng/ml	1112				
20	50 ng/ml	1082	1097	326 (B)	15.1	
21	100 ng/ml	1001				
22	100 ng/ml	1014	1008	237 (B)	10.9	
<i>Controls</i>						
23	QC #1	2147				
24	QC #1	2141	2144	1373 (B)	63.4	2.8
25	QC #2	1314				
26	QC #2	1296	1305	534 (B)	24.7	18.5
<i>Unknowns</i>						
27	Unknown #1	1600				
28	Unknown #1	1650	1625	854 (B)	39.4	7.6
29	Unknown #2	2000				
30	Unknown #2	2200	2100	1329 (B)	61.3	3.0

^aAdapted from data derived from a primate leptin RIA kit available from LINCO Research. Abbreviations: B , cpm of bound protein tracer in presence of varying amounts of unlabeled protein; B_0 , cpm of bound protein tracer in absence of competing unlabeled protein; QC, protein quality control.

^bCalculated from standard curve illustrated in Figure 10.24.2.

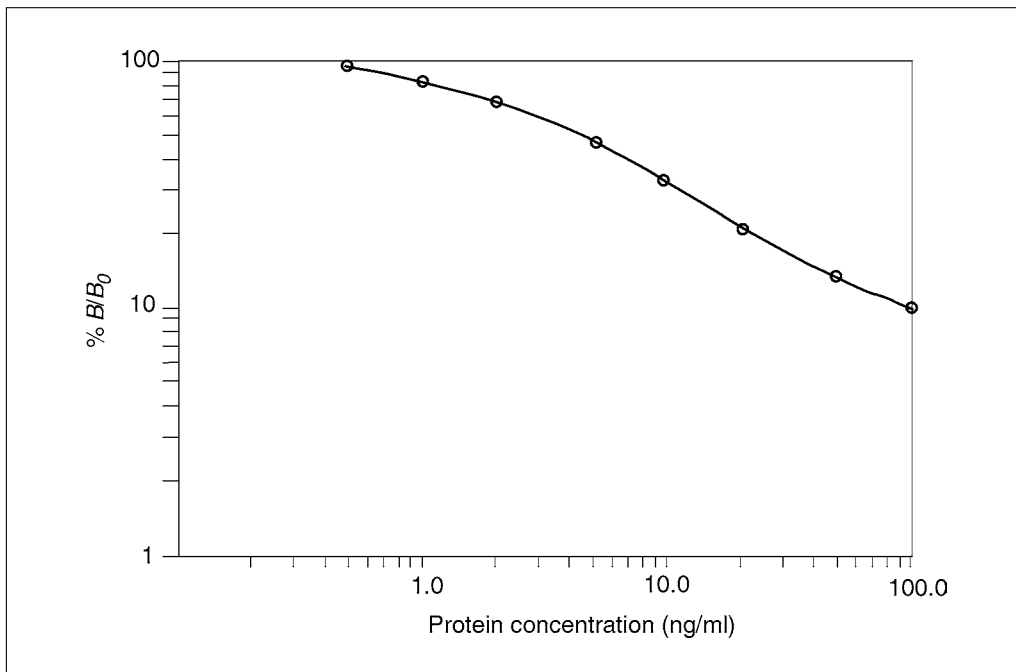


Figure 10.24.2 Log-log plot of protein concentration versus % B/B_0 . This graph is derived from the data in Table 10.24.2. The protein concentration range is between 0.5 ng/ml and 100 ng/ml. This range was determined by using protein standards at the concentrations of 0.5, 1, 2, 5, 10, 20, 50, and 100 ng/ml. The B/B_0 values are determined by dividing the B values (cpm of bound tracer in the presence of variable amounts of unlabeled protein standard) by B_0 (cpm of bound tracer without the presence of competing, unlabeled protein standard).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Anti-protein antiserum (primary antibody solution)

This solution contains the primary (first) antibody used to complex the [¹²⁵I]tracer or the unlabeled antigen in the standards, quality controls, and unknown samples. If the unknown samples are of human origin, an example of an appropriate primary antibody might be a rabbit monoclonal or polyclonal antiserum against the human protein to be quantitated (see Chapter 11 for production of monoclonal and polyclonal antisera). Dilute the antiserum 1:20 to 1:200 in assay buffer (see recipe), depending on the titer of specific immunoglobulin in the antiserum. Prepare dilution on day of use and keep at 2° to 8°C.

The primary antibody is either a polyclonal antibody (derived from serum; therefore referred to as an antiserum) or a monoclonal antibody (derived from ascites fluid or hybridoma supernatant). The serum, ascites fluid, or hybridoma supernatants may be used without purifying the primary antibody from these sources. In each case, a dilution with the diluting buffer is required. A dilution of 1:20 to 1:200 is specified above; however, the titer of a primary antibody directed at a protein (or peptide) of interest will differ depending on the immunogenicity of the antigen. Hence, the dilution of the anti-protein (or anti-peptide) antiserum, ascites fluid, or hybridoma supernatant may be increased (e.g., to 1:10,000) if the titer is high. The amount of primary antibody must be sufficient to bind 15% to 50% of the radiolabeled protein (or peptide) tracer used in the assay.

Assay buffer

Phosphate-buffered saline (PBS), pH 7.4 (see recipe)

0.025 M EDTA (APPENDIX 2)

0.1% (v/v) Triton X-100

0.08% (w/v) sodium azide

1% (w/v) BSA (RIA grade; Sigma-Aldrich)

Store short term (up to 24 hr) at 2° to 8°C or long term at –20°C

CAUTION: Sodium azide may react with lead or copper plumbing to form explosive metal azides; dispose of azide solutions according to local regulations and institutional guidelines; also see APPENDIX 1H.

Phosphate-buffered saline (PBS)

1.15 g NaH₂PO₄ (anhydrous; 9.5 mM)

5.75 g Na₂HPO₄ (anhydrous; 40.5 mM)

9.00 g NaCl (154 mM)

Add H₂O to 900 ml

Adjust to pH 7.4 using 1 M NaOH or 1 M HCl

Add H₂O to 1 liter

Precipitating secondary antibody/PEG solution

This solution contains an antibody against the immunoglobulin of the species used to raise the primary antibody. For example, if the primary antibody is a rabbit anti-human immunoglobulin, an example of an appropriate secondary antibody might be a goat anti-rabbit IgG. The composition of the solution is as follows:

Phosphate-buffered saline (PBS), pH 7.4 (see recipe) containing:

1% to 2% (v/v) antiserum

3% (w/v) polyethylene glycol (PEG; mol. wt. 8000)

0.1% (v/v) Triton X-100

0.025 M EDTA (APPENDIX 2)

0.08% (w/v) sodium azide

Store short term (up to 24 hr) at 2° to 8°C or long term at –20°C

CAUTION: Sodium azide may react with lead or copper plumbing to form explosive metal azides; dispose of azide solutions according to local regulations and institutional guidelines; also see APPENDIX 1H.

Protein quality controls

Select complex samples (e.g., sera, tissue culture supernatants) from the same species as the unknown samples (e.g., human), but in which the concentration of the desired antigen is already known. Before the assay, if needed, dilute in assay buffer (see recipe) to bring the final antigen concentration within the assay range (e.g., 3 and 18 ng/ml, as in Table 10.24.2). Store short term (up to 24 hr) at 2° to 8°C or long term at –20°C.

Protein standards

Dilute the purified protein of interest (antigen) from the same species as the unknown samples (e.g., human) in assay buffer (see recipe) at 0.5, 1, 2, 5, 10, 20, 50, and 100 ng/ml. Prepare dilutions on the day of use and keep at 2° to 8°C.

Protein tracer buffer

Prepare assay buffer (see recipe) containing 2% (w/v) serum from the species of origin of the unknown sample (e.g., human) as carrier. Store short term (up to 24 hr) at 2° to 8°C or long term at –20°C.

This buffer is used to dissolve iodinated protein.

Protein tracer, iodinated with ¹²⁵I

Radiolabel the purified protein of interest (antigen) from the same species as the unknown samples (e.g., human) with ¹²⁵I (see, e.g., UNIT 11.17) to a specific activity of 130 to 140 μCi/μg, and purify by HPLC (UNITS 10.12 & 10.13). To maximize stability of the batch of iodinated protein, lyophilize in aliquots that are sufficient to complete a whole RIA in which each assay tube contains 5,000 to 20,000 cpm. Store lyophilized tracer up to 1 month at –20°C. Before use, dilute in protein tracer buffer (see recipe) such that each 100 μl contains 5,000 to 20,000 cpm, and keep at 2° to 8°C.

COMMENTARY

Background Information

The first radioimmunoassay was developed to measure insulin in human plasma (Yalow and Berson, 1960). RIAs are most commonly used in clinical laboratories. They are used to measure, e.g., peptides and proteins with endocrine function, antibiotics, cocaine, steroids, antigens associated with autoimmunity, bacteria, viruses, parasites, neuropeptides, and expressed proteins.

Radioimmunoassay is a simple, rapid and precise method for measuring picomolar amounts of an antigen if the antigen is radiolabeled to high specific activity. An RIA, as described by Re and Haber (1980), requires: (1) a substance that binds with high affinity to the protein (or peptide) being measured; (2) a radiolabeled protein (or peptide)—a so-called “tracer”—that competes specifically for binding with the protein (or peptide) being measured; (3) a simple method for separating free and bound forms of the radiolabeled tracer; and (4) the availability of sufficient and known amounts of the protein (or peptide) being measured—so-called calibrators or standards—free from any impurities that could also displace the radiolabeled tracer protein (or peptide) from its binding site. The quantification of a specific antigen (protein or peptide) in a complex biological sample results from allowing differing amounts of the antigen of interest contained in a complex mixture (e.g., cell lysates, serum, saliva, urine, cerebrospinal fluid) to combine with a fixed quantity of an antigen-specific antibody in the presence of a fixed quantity of radiolabeled tracer antigen. The radiolabeled

tracer and the antigen being measured compete for binding to the antigen-specific antibody. Labeled and unlabeled antigens bind with equal affinity and in indirect proportion to each other. The smaller the amount of sample antigen, the larger the amount of labeled antigen that will bind. Hence, the higher the number of counts detected on a γ counter, the lower the amount of sample antigen present. A plot (a so-called standard curve) of the percent of bound tracer in the presence of differing amounts of unlabeled standard antigen—referred to as % B/B_0 —versus the concentration of standard antigen can be made, and the concentration of antigen in the unknown samples can be readily determined by comparing the amount of bound radiolabeled tracer in each unknown sample with the standard curve.

RIA reagents and supplies are available from numerous commercial sources (e.g., Sigma Chemical Co., ICN Biochemicals, Perkin-Elmer, Pierce Chemicals, LINCO Research). These are supplied as reagents or kits, which, in the case of double-antibody assays (used in this unit), include labeled antigen, antigen-specific primary antibody, and a secondary antibody (a so-called precipitating antibody) recognizing the primary antibody, which is a component of the bound antigen-antibody complex. Under some circumstances, it is also possible to purchase unlabeled antigen and to iodinate with one of the five available methods: chloramine-T (Hunter and Greenwood, 1962; UNIT 11.17), IODOGEN (Salacinski et al., 1981; UNIT 11.17), PIB (Wilbur et al., 1989), Bolton-Hunter

(Bolton and Hunter, 1973), or lactoperoxidase (Marchalonis, 1969).

As described by Hunter (1973), radioiodine (either ^{125}I or ^{131}I) is the label of choice because of the significant increase in radioactive disintegrations per unit time (in comparison to ^3H or ^{14}C). Unless the radiolabeled antigen is unstable (due to radiochemical damage), ^{125}I , not ^{131}I , is the radiolabel of choice, because of its longer half-life (60 days versus 8 days), the higher isotope abundance available in the carrier-free, commercially available ^{125}I label, its higher counting efficiency (80% versus 40%), and the higher specific activity of labeled protein (or peptide) with one molecule of ^{125}I ($\sim 100 \mu\text{Ci}/\mu\text{g}$ at a molecular mass of 20,000 and $\sim 1,000 \mu\text{Ci}/\mu\text{g}$ at a molecular mass of 2,000) in comparison to ^{131}I .

Chloramine-T (the sodium salt of the *N*-monochloro derivative of *p*-toluene sulfonate) in the presence of NaI (labeled with ^{125}I or ^{131}I) enables proteins (or peptides) containing tyrosyl residues to be specifically labeled in aqueous solutions under mildly alkaline conditions. The method was developed by Hunter and Greenwood (1962). The pH optimum is 7.5. At pH > 8.5, iodination of other amino acids (histidine or tryptophan) is possible. The efficiency of the reaction is dependent on the concentration of the protein (the lower the amount of protein, the less efficient the labeling). However, if only trace labeling is needed, only very low concentrations of chloramine-T are required.

iodo-GEN (1,3,4,6-tetrachloro-3 α ,6 α -diphenylglycouril; a trademark of Pierce developed by Salacinski, 1981), is a more gentle iodination method than chloramine-T, because the reaction occurs on a solid-phase surface of the oxidant, thereby minimizing the exposure of the protein or peptide to oxidative damage. The specific activity of the labeled proteins (or peptides) is less than that achievable by chloramine-T.

The PIB (*N*-succinimidyl 4-iodobenzoate) method, developed by Wilbur et al. (1989), iodinate tyrosine residues. Because these iodinated proteins (or peptides) do not contain an *ortho*-phenol, they are not susceptible to delabeling by natural deiodinases selective for *ortho*-iodophenols. This leads to increased stability of the iodinated proteins (or peptides).

The Bolton-Hunter (1973) method is a two-step labeling approach for introducing an iodinated *p*-hydroxyphenyl moiety (i.e., a tyrosine mimic) into proteins (or peptides) lacking tyrosine residues or under circum-

stances where an active tyrosine is involved in a binding or an antigenic site and therefore should not be iodinated. The first step is the iodination by chloramine-T or IODO-GEN of the *N*-hydroxysuccinimide ester of 3-(*p*-hydroxyphenyl)propionic acid. The second step is the active ester coupling of the iodinated 3-(*p*-hydroxyphenyl)propionic acid to lysyl residues in the protein (or peptide). Addition of a protein solution to the dried reagent residue results in a covalent attachment of the radiolabeled moiety to lysyl residues in the protein. The advantage of this acylating method is that the protein is not exposed to the oxidizing conditions associated with the chloramine-T or IODO-GEN. In addition, the method allows the introduction of a radiolabeled iodine into proteins lacking a tyrosyl residue but containing a lysyl residue. When the antibodies used in the RIA are directed at an antigenic site containing a tyrosyl residue, the Bolton-Hunter method does not affect the antigenic site unless a modified lysyl residue is also in the antigenic site.

The lactoperoxidase method is an enzymic reaction method developed by Marchalonis (1969). Lactoperoxidase catalyzes the oxidation of iodine in the presence of hydrogen peroxide. In turn, selective iodination of tyrosyl residues in proteins occurs. The specific activity of the labeled proteins is less than that achievable by chloramine-T.

Critical Parameters and Troubleshooting

In order to develop an RIA for a given protein (or peptide), the following questions should be asked and answered (http://www.cyfc.psu.edu/funding/L2_dr2.html).

1. Is it feasible to develop a radioimmunoassay for a specific protein (or peptide)? Begin by searching the literature to determine the following:

a. Information about the protein's availability: i.e., has it been isolated and, if so, how easy will it be to obtain the amount required for developing and running subsequent RIAs?

b. What is the molecular mass (and amino acid composition) of the protein?

c. What is its half-life?

d. What is its intracellular, organ, or body distribution? Has a previous assay been reported (i.e., are sensitivity, sample volumes, and assay protocols defined) for measuring the protein of interest? If so, was the assay validated? Such validation requires determination of:

i. Minimal detection limit, defined as the minimum concentration of the protein that may be distinguished from zero at a 95% confidence limit, by extrapolating the mean minus two standard deviations for multiple replicates (e.g., 5 to 10 assays) of the zero calibrator (i.e., not containing standard protein).

ii. Reproducibility of the standard curve using replicate runs (e.g., 5 to 10 assays) by comparing averages, standard deviations, and confidence values for the slope and correlation coefficient.

iii. Precision, by measuring the intra-assay variation (confidence value) for replicate samples at various concentrations (i.e., low, medium, and high).

iv. Accuracy, by adding known amounts of unlabeled protein (or peptide) to samples containing various endogenous concentrations.

v. Drift, due to delays in adding reagents due to distribution of samples in a given assay configuration (i.e., arrangement of tubes or placement of samples with presumed low, medium, or high concentrations).

vi. Sensitivity sufficient to capture the full range of developmental, individual, gender, and diurnal variations.

2. Does an immunodiagnostic vendor provide a source of polyclonal (or monoclonal) antibodies specific to the protein (or peptide) of interest, chemically or enzymatically labeled standard protein (or peptide), and calibrators (i.e., standard protein or peptide at various defined concentrations required to construct a standard curve)? If so, is the laboratory prepared to develop an RIA?

3. Is an RIA commercially available for the protein (or peptide) of interest?

If the answer to any of the above questions is yes, one is ready to develop or purchase a solution RIA. If not, a large amount of work may be required to prepare the required assay reagents, which may involve expressing and/or purifying a sufficient quantity of a standard protein or synthesizing and/or purifying a standard peptide (Chapter 10), preparing polyclonal or monoclonal antibodies (Chapter 11), and radiolabeling the standard protein (or peptide).

A simple method to determine the minimal concentration of iodinating reagent required to label a protein (or peptide) is to carry out the labeling procedure and then measure, in a small aliquot, the amount of labeled protein (or peptide) that can be precipitated using 10% trichloroacetic acid (TCA). Addition of a carrier protein (e.g., 2% to 3% serum) will be required to precipitate most peptides,

since they are not readily precipitated by 10% TCA.

The most common diluent is sodium phosphate-buffered saline (e.g., 0.01 to 0.05 M, pH 7.5). However, for membrane proteins, more complex buffer mixtures (containing nonionic or ionic detergents) are required to solubilize the proteins for use in a soluble RIA format. Furthermore, RIAs require the addition of a carrier protein at very large molar excess (e.g., 2% to 3% serum) over antigen and labeled tracer when working at low concentrations, to minimize losses due to adsorption to glass or plastic surfaces. In addition, the following should be checked: (1) the carrier protein concentration should prevent adsorption of radiolabeled antigen at the lowest concentration being assayed; (2) the carrier protein should be free of the antigen being assayed or any cross-reacting antigens; (3) the stability of the radiolabeled antigen in the presence of the carrier protein may be reduced, especially under prolonged incubation times at room temperature; and (4) a bacteriostatic agent (e.g., sodium azide at 0.1 g/liter or thimerosal at 0.01%, w/v) increases the stability of the assay by reducing inadvertent degradation of antigen or carrier protein by contaminating proteases.

Separation of bound from free tracer can be achieved physicochemically (via electrophoresis, chromatography, or gel diffusion; see Chapter 10) or by immunoprecipitation, as in (the so-called double-antibody method in which the antibody-tracer complex is immunoprecipitated with a secondary anti-immunoglobulin raised in another species; *UNIT 10.16*). In the double-antibody/PEG method described here, PEG is added to enhance the precipitation reaction. A perfect double-antibody separation system would show no tracer in the bound fraction in the absence of capture antibody and 100% of the tracer in the bound fraction after incubation with excess capture antibody.

As described by Hunter (1973), any separation system may affect the system in two discrete ways. First, antigen-antibody interactions are reversible, even if the experimental system is at equilibrium. Hence, some bound antigen may dissociate if the separation removes free antigen from the system. Such dissociation would only occur if the association and dissociation constants for the antigen-antibody interaction were equal. With antibodies of high avidity, this would not occur. Secondly, the separation system may disrupt the antigen-antibody complex. This could occur when nonpolar compounds are being assayed,

since the organic solvents used to extract unbound label could also extract free and bound compounds equally. This is not a problem with the RIAs of proteins (or peptides) using the double-antibody/PEG method, since organic solvents are not used.

Anticipated Results

As described by Hunter (1973), radioiodinated proteins (or peptides) are usually stable at 2° to 8°C for 30 days. High-specific-activity labeling of proteins (or peptides) can lead to a decrease in immunoreactivity due to radiochemical damage. Such damage can be differentiated from chemical damage by the fact that it is proportional to the amount of radioactivity reacted and should not vary with different batches of iodine. Chemical damage results from iodine's propensity to oxidize sulfhydryl groups and/or from denaturation by the organic solvents, if used to extract the unreacted iodine. Iodination may also alter the immunoreactivity of proteins (or peptides) if the iodinated tyrosine is contained within the antigenic site recognized by the antibody. Hence, to minimize this possibility, trace labeling should not exceed one atom of iodine per mole of protein (or peptide). Interestingly, such a decrease in the apparent immunoreactivity may be specific for a single antiserum and not others, so it may be beneficial to use several antisera during the development of an RIA.

Time Considerations

The solution RIA can be completed in 2 days or less. The incubation time for the assay using the double-antibody method is complete by 16 to 24 hr at 4° to 25°C. The separation of immunoprecipitated antigen requires different lengths of time depending on the number of samples being assayed, due to the time required to centrifuge the precipitate and to decant the supernatants (usually < 1 to 3 hr). The amount of time needed to count the samples with a γ counter depends on the specific activity of the labeled tracer, but is usually <5 min per tube.

Literature Cited

- Bolton, A.E. and Hunter, W.M. 1973. The labeling of proteins to highly specific radioactivities by conjugation to a ^{125}I -containing acylating agent. *Biochem. J.* 133:529-539.
- Hunter, W.M. 1973. Radioimmunoassay. *In* Handbook of Experimental Immunology, Immunochimistry (D.M. Weir, ed.), Vol. 1, pp.17.1-17.36. Blackwell Scientific Publications, Oxford.

- Hunter, W.M. and Greenwood, F.C. 1962. Preparation of iodine-131 labeled human growth hormone of high specific activity. *Nature* 194:495-496.
- Marchalonis, J.J. 1969. An enzymic method for the trace iodination of immunoglobulins and other proteins. *Biochem. J.* 113:299-305.
- Re, R.N. and Haber, E. 1980. Radioimmunoassay. *In* Methods in Immunodiagnosis (N.R. Rose and P.E. Bigazzi, eds.) pp.147-160. John Wiley & Sons, New York.
- Salacinski, P.R.P., McLean, C., Sykes, J., Clement-Jones, V., and Lowry, P. 1981. Iodination of proteins, glycoproteins, and peptides using a solid phase oxidation reagent, 1,3,4,6-tetrachloro-3 α ,6 α -diphenylglycouril (Iodogen). *Anal. Biochem.* 117:136-146.
- Wilbur, D.S., Hadley, S.W., Hylarides, M.D., Abrams, P.G., Beaumier, P.A., Morgan, A.C., Reno, J.M., and Fritzberg, A.R. 1989. Development of a stable radioiodinating reagent to label monoclonal antibodies for radiotherapy of cancer. *J. Nucl. Med.* 30:216-226.
- Yalow, R.S. and Berson, S.A. 1960. Immunoassay of endogenous plasma insulin in man. *J. Clin. Invest.* 39:1157-1175.

Key References

Hunter, 1973. See above.

Excellent review of all aspects of radioimmunoassay development.

Yalow and Berson, 1960. See above.

This paper describes the development of the first radioimmunoassay and its use to measure insulin in human plasma. Dr. Rosalind Yalow was awarded the 1977 Nobel Prize in Physiology or Medicine for the invention of this method.

Internet Resources

http://www.cyfc.psu.edu/funding/L2_dr2.html

This URL presents an excellent assay development plan for measuring biomarkers by immunological assays (e.g., RIAs and EIAs). It discusses feasibility considerations and internal and external validation, as well as aspects of sample selection, collection, handling, and storage.

<http://www.lincoresearch.com/products/pl-84k.html>

This URL provides access to a detailed protocol for a commercially available primate leptin RIA, developed by LINCO Research. This protocol is an excellent example of the use of a double-antibody/PEG method for doing RIAs. This protocol was used as the basis for the protocol in this unit.

Contributed by John A. Smith
University of Alabama at Birmingham
Birmingham, Alabama

CHAPTER 11

Immunology

INTRODUCTION

Certain technological advances in the field of molecular biology were made possible in part by earlier progress in the field of immunology. A review of the earlier chapters in this book documents the importance of immunological methods to the purification of proteins as well as to the identification of specific cDNA clones. Specific antibodies have greatly facilitated the purification of proteins by immunoaffinity chromatography (*UNIT 10.11*) and immunoprecipitation (*UNIT 10.16*). One limitation of immunoaffinity chromatography has been that the harsh dissociation conditions required to elute bound antigens from high-affinity antibodies sometimes denature the eluted antigens. *UNIT 11.18* presents a method to circumvent this problem by utilizing polyol-responsive antibodies that release their bound antigens under gentle dissociation conditions, employing a combination of various low molecular weight polyhydroxylated compounds (e.g., ethylene glycol) and nonchaotropic salts (e.g., ammonium sulfate). These polyol-responsive antibodies can be readily identified and isolated from typical fusions, prepared by standard hybridoma procedures. When pure protein has been unavailable for deducing the complementary oligonucleotide sequence, specific antibodies have been utilized to screen recombinant DNA libraries for the desired cDNA clones (*UNIT 6.7*) and selected mRNA for the translation of desired protein (*UNIT 6.8*). Specific antibodies have also been utilized to identify antigen by western blotting (*UNIT 10.8*).

Just as immunology has facilitated the advances made in the field of molecular biology, the latter in turn has contributed to a better understanding of the basis for antibody diversity. The clonal selection theory proposed by Sir Macfarlane Burnet in 1959 is now an accepted concept: each B cell differentiates into a plasma cell committed to the production of antibodies specific for one antigen—i.e., the antibodies are monoclonal in nature. “Clonal selection” refers to the fact that when an antigen binds to one of these antibodies on the membrane of the B cell, the cell is stimulated to proliferate (at which point some variation may be introduced in the “monoclonal” cell line). Generally, many clones respond to a single antigen, as most proteins carry multiple antigenic sites (called epitopes). The overall immune response is polyclonal, with specific recognition of multiple, discrete epitopes.

An understanding of the genetic mechanisms responsible for antibody (or immunoglobulin) diversity requires some knowledge of antibody structure. Man has five major immunoglobulin classes: IgG, IgA, IgD, IgE, and IgM, which share the same type of combining site for antigen. The immunoglobulin molecule is similar for the first four classes; it consists of four polypeptides—two heavy chains and two light chains—arranged in the shape of the letter “Y,” with a molecular weight of ~150,000. The IgM class, with a molecular weight of ~800,000, consists of five Y-shaped molecules arranged in a cyclic pentamer, with the antigen-binding sites facing outward. Although the different immunoglobulin classes can share the same κ or λ light chains, they are each distinguished by their unique heavy chains, designated γ (IgG), α (IgA), δ (IgD), ϵ (IgE), and μ (IgM). The heavy and light chains are each composed of constant and variable regions. The antigen-binding site, a cleft of about $15 \text{ \AA} \times 20 \text{ \AA} \times 10 \text{ \AA}$ deep formed by interactions of hypervariable regions of the heavy- and light-chain variable regions, is unique for each antibody.

Contributed by John A. Smith

Current Protocols in Molecular Biology (2001) 11.0.1-11.0.3

Copyright © 2001 by John Wiley & Sons, Inc.

Immunology

11.0.1

Supplement 53

For many years it was assumed that the mammalian germ line must include a separate gene for every polypeptide that ultimately appears in an antibody; this model presupposes a vast number of immunoglobulin genes. In the past decade, however, recombinant DNA technology has shown that diversity in antigen-binding sites arises through genetic recombination in somatic cells—i.e., while B lymphocytes are maturing and differentiating in the bone marrow. Located on different chromosomes are approximately 50 genes coding for the “constant” **C** regions, the “variable” **V** regions, the “joining” **J** segments (which combine with the **C** and **V** regions to make up the antibody’s light chain) and the “diversity” **D** segments (which combine with **C**, **J**, and **V** regions to comprise the antibody’s heavy chain). Mouse germ cells have a few hundred **V** segments, approximately 20 **D** segments, and 4 **J** segments, which can be assembled in >10,000 combinations. Subsequent assemblage of heavy and light chains could yield >10 million specific antigen-binding sites. (For an excellent review of the molecular biology of the immune system, see Tonegawa, 1985.)

This chapter presents the methodologies for the preparation of both monoclonal and polyclonal antibodies. Section I describes the enzyme-linked immunosorbent assay (ELISA), a highly sensitive, versatile, and quantitative technique that requires little equipment and for which critical reagents are readily available. The preparation of enzyme-antibody conjugates, which forms the basis of this assay, is described in *UNIT 11.1*. The versatility of ELISAs is demonstrated by the six distinct ELISA protocols presented in *UNIT 11.2*. These provide general methods for the detection of specific antibodies, soluble antigens, or cell-surface antigens. Protocols for determining the isotype (i.e., serological class) of antibodies are described in *UNIT 11.3*.

The pioneering studies of Kohler and Milstein (1975) enable investigators to obtain milligram quantities of specific monoclonal antibodies after immunizing mice with relatively impure antigen. The spleen is removed from a previously immunized mouse that has a sufficient antibody titer. After separation into individual cells, B cells from the spleen are fused with myeloma cells of B cell origin to produce immortal antibody-secreting hybridoma cells of predetermined specificity. Each hybridoma cell is capable of producing an unlimited supply of a single, antigen-specific monoclonal antibody. Section II describes the preparation of these antigen-specific monoclonal antibodies in separate protocols that cover immunization of mice (*UNIT 11.4*), cell preparation and cell fusion for generating hybridoma cell lines (*UNITS 11.5-11.7*), cloning by limiting dilution to ensure the production of truly monoclonal antibodies derived from a single antibody-secreting cell (*UNIT 11.8*), freezing and recovery of hybridoma cell lines (*UNIT 11.9*), production of cell culture supernatants of monoclonal antibodies in ascites fluid (*UNIT 11.10*), and purification of these monoclonal antibodies by affinity chromatography (*UNIT 11.11*). Detection of antibody in serum, hybridoma supernatants (micrograms per milliliter), and ascites fluid (milligrams per milliliter) by ELISA is described in *UNIT 11.2*.

Although monoclonal antibodies can be made available in unlimited quantities and without the need to purify the antigen to homogeneity, the reliance upon only monoclonal antibodies for detection and identification of antigen and cDNA clones can produce equivocal results. Because monoclonal antibodies may be specific for short peptide sequences, there is a possibility of obtaining false positives, since unrelated proteins can share small regions of homology. One way in which this uncertainty can be minimized is to utilize several different monoclonal antibodies specific for different sites on the antigen. Another disadvantage of using a monoclonal antibody is that it may have a relatively low affinity for a given antigenic site.

These problems caused by the use of monoclonal antibodies may be circumvented by generating polyclonal antibodies, which consist essentially of numerous monoclonal

antibodies with different epitope specificities (Section III). When a purified antigen is available in sufficient amount for immunization, it is possible to obtain specific polyclonal antibodies with high affinity after repeated immunizations (*UNIT 11.12*; Klinman and Press, 1975). Choice of animal is determined by the amount of antiserum required for subsequent experiments. Although animals such as goats, sheep, or horses can provide larger volumes of antiserum, few institutions have adequate facilities for their care and maintenance. Mice, rats, and guinea pigs, on the other hand, may not yield sufficient volumes of antiserum. For these reasons, rabbits have become the animal of choice for the generation of polyclonal antibodies. *UNIT 11.12* describes the proper preparation of antigen as well as various routes of immunization in rabbits to optimize the antibody response. Although a schedule for immunization and boosting is provided, this procedure is only a recommendation of what has worked for the author; optimal conditions should be determined empirically. *UNIT 11.13* describes systems for in vitro antibody production, and subsequent measurement of secreted antibodies. *UNIT 11.14* discusses the purification from serum, ascites fluid, or hybridoma supernatant of the immunoglobulin G fraction, which becomes the predominant antibody class after the booster injection.

If purified antigen is in limited supply, polyclonal (as well as monoclonal) antibodies can still be raised by immunization with synthetic peptides whose sequences are based on that of the protein, which it is designed to mimic (Section IV). In this case, the selection of an immunogenic peptide is vital for obtaining a good antibody response. *UNIT 11.15* discusses the necessary parameters to consider in the selection of a particular peptide sequence that will elicit an antibody that recognizes the native form of the protein. To enhance the immunogenicity of the peptide, it can be chemically cross-linked to a carrier molecule (*UNIT 11.16*). Such cross-linking of the peptide has been demonstrated to be helpful in generating an antibody response to peptides that might not otherwise elicit antibody production.

The quantitation of specific antibody (as well as its isotypes) in polyclonal antisera, ascites fluid, or hybridoma supernatant by solid-phase radioimmunoassay (RIA) is described in *UNIT 11.17*. Although this method is more laborious than the nonradioactive ELISA described in *UNIT 11.2*, the disadvantages are offset by greater sensitivity and better reproducibility from assay to assay.

Literature Cited

- Burnet, F.M. 1959. *The Clonal Selection Theory of Acquired Immunity*. Vanderbilt University Press, Nashville.
- Klinman, N.R. and Press, J. 1975. The B Cell specificity repertoire: Its relationship to definable subpopulations. *Transplant. Rev.* 24:41-83.
- Kohler, G. and Milstein, C. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature (Lond.)* 256:495-497.
- Tonegawa, S. 1985. The molecules of the immune system. *Sci. Am.* 253:122-131.

John A. Smith

Antigen can be detected and quantitated using an enzyme-linked immunosorbent assay (ELISA). The direct and sandwich ELISAs discussed in *UNIT 11.2* are particularly useful for determining the presence or amount of antigen in samples ranging from crude bacterial lysates to highly purified protein antigen preparations. These assays require the preparation of enzyme-antibody conjugates. How these conjugates are prepared and what enzymes should be linked to the antigen-specific antibody in the conjugate are presented in *UNIT 11.1*.

UNIT 11.1

Conjugation of Enzymes to Antibodies

Conjugation of enzymes to antibodies involves the formation of a stable, covalent linkage between an enzyme [e.g., horseradish peroxidase (HRPO), urease, or alkaline phosphatase] and an antigen-specific monoclonal or polyclonal antibody in which neither the antigen-combining site of the antibody nor the active site of the enzyme is functionally altered. The chemistry of cross-linking HRPO or urease to immunoaffinity-purified monoclonal or polyclonal antibodies (IgG) is presented in Figures 11.1.1 and

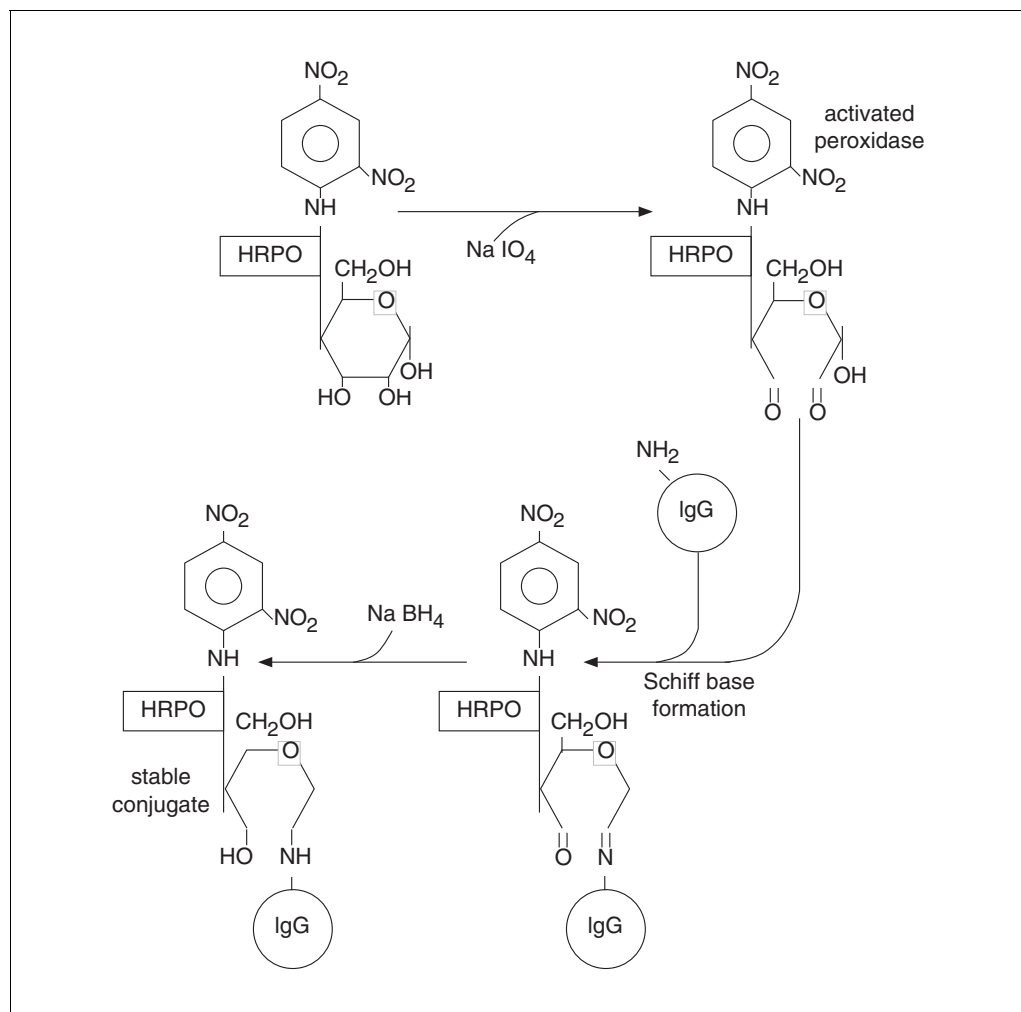


Figure 11.1.1 Conjugation of horseradish peroxidase (HRPO) to antibody (IgG) using the periodate oxidation method. The method involves three chemical steps: (1) sodium periodate (NaIO_4) oxidation of the carbohydrate side chains of HRPO, (2) Schiff base formation between activated peroxidase and amino groups of the antibody, and (3) sodium borohydride (NaBH_4) reduction of the Schiff base to form a stable conjugate.

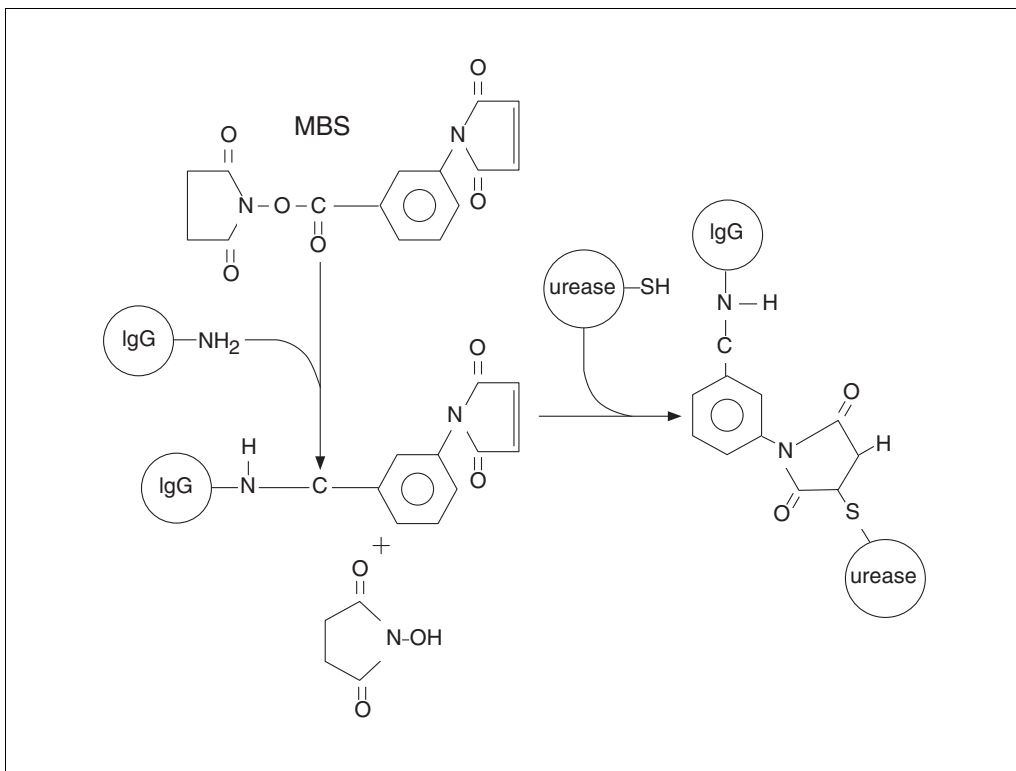


Figure 11.1.2 Conjugation of urease to antibody (IgG) with *m*-maleimidobenzoyl *N*-hydroxysuccinimide ester (MBS). The first step involves benzoylation of the amino groups of the antibody (IgG) oxidate oxidation method. The second step involves the thiolation of the maleimide moiety by the sulfhydryl groups of the urease enzyme.

11.1.2, respectively. The chemistry of cross-linking alkaline phosphatase to antibodies is presented in Figure 11.16.2.

CONJUGATION OF HORSERADISH PEROXIDASE TO ANTIBODIES

Horseradish peroxidase–antibody conjugates (Tijssen and Kurstak, 1984) can be used in ELISA (enzyme-linked immunosorbent assay; *UNIT 11.2*) and western blotting (*UNIT 10.8*).

Materials

1 mg/ml antibody solution (affinity-purified polyclonal or monoclonal antibodies; *UNIT 11.11*)

0.1 M phosphate buffer, pH 6.8

Horseradish peroxidase (HRPO; Sigma Type VI #P8375)

0.1 M carbonate buffer, pH 9.2

Sodium periodate (NaIO₄) solution, freshly prepared

Sodium borohydride (NaBH₄) solution, freshly prepared

Saturated ammonium sulfate [(NH₄)₂SO₄] solution

Tris/EDTA/NaCl (TEN) buffer, pH 7.2

Bovine serum albumin (BSA)

Glycerol

Dialysis membrane (see reagents and solutions and *APPENDIX 3*)

Pasteur pipet fitted with glass wool

Sephadex G-25, medium (size of gel matrix)

1. Dialyze 1 mg/ml antibody solution against 2 liters of 0.1 M phosphate buffer, pH 6.8, overnight at 4°C, stirring gently.

**BASIC
PROTOCOL**

Immunology

11.1.2

Antibody may be polyclonal or monoclonal, purified as described in UNIT 11.11.

A₂₈₀/1.44 = mg IgG/ml. Concentration of antibody should be at least 1 mg/ml.

Soak dialysis membrane 1 hr in 50% (v/v) aqueous ethanol, 1 hr in 10 mM NaHCO₃, and 1 hr in 1 mM EDTA. Then rinse twice in distilled water and store at 4°C in phosphate buffer containing 0.01% (w/v) sodium azide.

2. Dissolve 10 mg HRPO in 1 ml 0.1 M carbonate buffer, pH 9.2.
3. Mix 0.25 ml freshly prepared NaIO₄ solution with 0.25 ml of 10 mg/ml HRPO/carbonate mixture from step 2, cap tightly, and incubate at room temperature for 2 hr in the dark (NaIO₄ is light sensitive).
4. Into a Pasteur pipet fitted with glass wool and blocked at the tip with Parafilm, add 1 ml of the dialyzed 1 mg/ml antibody solution from step 1 to 0.5 ml of the 10 mg/ml HRPO solution from step 3. Add 0.25 g Sephadex G-25 to the antibody/HRPO mixture.

Addition of Sephadex increases the concentration of antibody and HRPO by absorbing water. This enhances the conjugation of enzyme to antibody.

5. Incubate 3 hr at room temperature in the dark.
6. Wash column with 0.75 ml carbonate buffer to elute conjugate.
7. Add 38 µl freshly prepared NaBH₄ solution to the eluate and incubate 30 min at room temperature in the dark.
8. Add 112 µl freshly prepared NaBH₄ and incubate 60 min in the dark.
9. Add 0.9 ml of saturated (NH₄)₂SO₄ solution and stir gently for 30 min at 4°C. Centrifuge 15 min at 10,000 × g, 4°C.
10. Decant, discard supernatant, and resuspend pellet in 0.75 ml TEN buffer.
11. Dialyze resuspended pellet overnight at 4°C against 2 liters TEN buffer. Change TEN solution in the morning and continue dialyzing for 4 hr.
12. Remove conjugate from dialysis membrane and add sufficient BSA to bring the conjugate solution to a final concentration of 20 mg BSA/ml.
13. Add an equal volume of glycerol and store at -20°C.

**ALTERNATE
PROTOCOL**

CONJUGATION OF UREASE TO ANTIBODIES

Urease conjugates can be used in ELISA (UNIT 11.2) but not western blotting (UNIT 10.8).

Additional Materials

- 20 mg/ml urease (Sigma Type VII #U0376; source is important) in 0.1 M phosphate buffer
- m*-Maleimidobenzoyl *N*-hydroxysuccinimide ester in dimethylformamide (MBS/DMF solution)
- 0.143 M 2-mercaptoethanol (prepare from 14.3 M stock)
- Phosphate-buffered saline (PBS; APPENDIX 2)
- 12 × 75-mm glass tubes
- 1.5 × 5-cm PD-10 column (Pharmacia)
- Nitrogen tank

1. Repeat step 1 of the basic protocol.
2. Dialyze 0.25 ml of 20 mg/ml urease in 0.1 M phosphate buffer against 2 liters 0.1 M phosphate buffer overnight at 4°C while stirring gently.
3. In the morning replace the phosphate buffer and continue to dialyze for 2 hr.
4. Remove antibody and urease solutions from dialysis membranes and place in separate glass tubes. Read A_{280} of antibody solution and dilute with phosphate buffer to 0.5 mg/ml.
5. Add 0.075 ml MBS/DMF solution to 1.5 ml of 0.5 mg/ml dialyzed antibody solution in a glass tube (MBS/antibody molar ratio, 120:1). Place a magnetic stir bar in the tube and stir gently at room temperature for 30 min.
6. Load on a PD-10 column (see support protocol, *UNIT 10.9*) preequilibrated with 100 ml phosphate buffer. Run column with phosphate buffer and collect 0.6-ml fractions. Read A_{280} of fractions and collect first peak that elutes.

The first peak contains activated antibody and the second peak contains free MBS.

7. Pool first peak (generally about 2.5 to 3.0 ml) and add 0.15 ml of 20 mg/ml urease in 0.1 M phosphatase buffer (urease/antibody weight ratio, 4:1).
8. Stir at room temperature under N_2 for 1.5 hr or until solution appears cloudy.
9. Add 0.143 M 2-mercaptoethanol to a final concentration of 2 mM (0.014 ml mercaptoethanol solution/ml urease-antibody conjugate) and stir at room temperature for 30 min.
10. Dialyze overnight at 4°C against 2 liters PBS. In the morning, replace PBS and dialyze 4 hr.
11. Add an equal volume of glycerol, divide into small aliquots, and store at -20°C (stable for 1 year).

CONJUGATION OF ALKALINE PHOSPHATASE TO ANTIBODIES

Alkaline phosphatase conjugates can be used in ELISA (*UNIT 11.2*) and western blotting (*UNIT 10.8*).

Additional Materials

- 5 mg/ml antibody solution (affinity-purified polyclonal or monoclonal antibodies, *UNIT 11.11*)
- 10 mg/ml alkaline phosphatase (enzyme immunoassay grade; Boehringer Mannheim; source is important)
- 25% glutaraldehyde in H_2O
- Tris/ovalbumin solution
- Sodium azide

1. Dialyze 5 mg/ml antibody solution as in step 1 of the basic protocol, except dialyze against PBS.
2. Remove antibody solution from dialysis membrane and place in a tube. Read A_{280} and dilute with PBS to 3 mg/ml.
3. Add 100 μ l dialyzed antibody solution to 90 μ l of 10 mg/ml alkaline phosphatase in a 1.5-ml microcentrifuge tube.

***ALTERNATE
PROTOCOL***

Immunology

11.1.4

4. Add 5 ml 25% glutaraldehyde and mix gently. Let stand at room temperature.
5. Remove 25- μ l samples at time 0, 5, 10, 15, 30, 60, and 120 min and place in separate 1.5-ml microcentrifuge tubes. Add 125 μ l PBS to each sample, then add 1.1 ml Tris/ovalbumin solution. Store each sample on ice until the time course is completed.
6. Dialyze the samples against PBS as described in step 1 of this protocol. Test each sample for alkaline phosphatase activity using a direct ELISA assay (*UNIT 11.2*) to determine which conjugation time yields the most active enzyme conjugate.
7. Repeat steps 1 to 4, but in step 4 allow the reaction to proceed for the optimal conjugation time, as determined in step 6.
8. Add sodium azide to 0.1% and store the conjugate protected from light at 4°C for up to 1 year. Alternatively, add an equal volume of glycerol and store the conjugates at -20°C for 1 year.

REAGENTS AND SOLUTIONS

0.1 M carbonate buffer, pH 9.2

1.36 g sodium carbonate
 7.35 g sodium bicarbonate
 950 ml H₂O
 Adjust pH to 9.2 with 1 M HCl or 1 M NaOH, if necessary
 Add H₂O to 1 liter

m-Maleimidobenzoyl N-hydroxysuccinimide ester in dimethylformamide (MBS/DMF solution)

Prepare 0.25% (w/v) solution by adding 2.3 mg MBS (Pierce #22310) to 0.92 ml DMF. Use within 1 hr of preparation.

MBS will deteriorate upon prolonged storage if repeatedly thawed, opened, and refrozen. Dispense into 10-mg aliquots and store desiccated at -20°C.

0.1 M phosphate buffer, pH 6.8

Stock A, 0.2 M: 31.2 g NaH₂PO₄ in 1 liter H₂O
 Stock B, 0.2 M: 28.39 g Na₂HPO₄ in 1 liter H₂O
 Mix 51 ml Stock A, 49 ml Stock B, and 100 ml H₂O

Saturated ammonium sulfate [(NH₄)₂SO₄] solution

Prepare 0.01 M Tris solution by adding 1.21 g Tris base to 990 ml water, adjust to pH 7.0, and bring to a final volume of 1 liter. Weigh 767 g (NH₄)₂SO₄ and dissolve in 1 liter 0.01 M Tris by stirring and gently warming. Adjust the pH to 7.0 and store at 4°C. (NH₄)₂SO₄ crystals should be seen at the bottom of the solution at 4°C.

Sodium borohydride (NaBH₄) solution

5 mg NaBH₄/ml 0.1 mM NaOH (use immediately)

Sodium periodate (NaIO₄) solution

1.71 mg NaIO₄/ml H₂O (use immediately)

Tris/EDTA/NaCl (TEN) buffer, pH 7.2

To 930 ml H₂O, add:
 6.06 g Tris base
 0.37 g Na₂EDTA
 8.77 g NaCl
 Adjust pH to 7.2 with HCl
 Add H₂O to 1 liter

Tris/ovalbumin buffer

0.05 M Tris·Cl, pH 8.0

5% ovalbumin

5 mM MgCl₂

0.5% NaN₃

0.5% mertiolate

COMMENTARY

Background Information

Direct conjugation of enzymes to antibodies has greatly simplified the development and performance of many different types of immunoassays. The conjugation of HRPO (Nakane and Kawaoi, 1974) to antibody is dependent on the generation of aldehyde groups by periodate oxidation of the carbohydrate moieties on HRPO. Combination of these active aldehydes with amino groups on the antibody forms Schiff bases that, upon reduction by sodium borohydride, become stable. Horseradish peroxidase (HRPO) conjugates are useful in all types of immunological assays, but are generally less stable than urease conjugates. In addition, endogenous peroxidases may cause false positive reactions. For urease conjugation (Healey et al., 1983), cross-linking the enzyme and antibody with MBS is achieved through benzoylation of free amino groups on antibody. This is followed by thiolation of the maleimide moiety of MBS by the cysteine sulfhydryl groups of urease. The advantages of urease conjugates are their stability in solution at normal working dilutions, the rapid turnover rate of the enzyme, the easily discernible color change when substrate is added, and the fact that urease is not found in most mammalian or bacterial systems. The disadvantage is that since no precipitable substrate is available, urease conjugates cannot be used for immunohistology or western blotting.

Alkaline phosphatase conjugates are useful for all types of immunological assays depending on the alkaline phosphatase substrate used (i.e., *p*-nitrophenyl phosphate in diethanolamine is the preferred substrate for ELISA with colorimetric detection, 4-methylumbelliferyl phosphate is useful for ELISA with fluorimetric detection, and nitroblue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate is the preferred substrate for western blotting). Alkaline phosphatase conjugates are as stable as urease conjugates and more stable than HRPO conjugates. Endogenous phosphatases can cause false positive reactions. However, levamisole will inhibit alkaline phosphatase in many mammalian tissues but not the alkaline phosphatase (i.e., bovine intestinal) used in the conjugates,

and for this reason levamisole may be added to the substrate solution.

The one-step glutaraldehyde method (Voller et al., 1976) is the most simple available procedure for preparing alkaline phosphatase-antibody conjugates. Various alternative procedures for preparing alkaline phosphatase conjugates have been compared (Jeanson et al., 1988).

The sensitivity that can be achieved with either HRPO, urease, or alkaline phosphatase conjugates is comparable and between 1 ng/ml and 10 ng/ml of antigen can be detected.

Critical Parameters

The most critical parameters of both conjugation methods are the quality of enzyme and the cross-linking reagents. Several lots of these reagents should be tested as described in the protocol before conjugating to larger quantities of antibodies. It is imperative that the *m*-maleimidobenzoyl *N*-hydroxysuccinimide ester (MBS), sodium periodate (NaIO₄), and sodium borohydride (NaBH₄) be stored in a desiccator and that solutions containing these chemicals be prepared immediately prior to use. The method described is applicable to most antibodies and should produce conjugates that are useful for developing an ELISA for detecting sensitively and specifically a given antigen. However, not all antibodies conjugate in an identical manner. It may be necessary to vary the ratio of MBS/antibody or urease/antibody for the urease conjugation and the NaIO₄/HRPO and HRPO/antibody ratios for a given HRPO conjugation.

The quality and grade of alkaline phosphatase is crucial to the generation of effective conjugates. Immunoassay grade material is recommended over lower grades, and the enzyme should not be conjugated beyond its expiration date. In the case of polyclonal antisera, the specificity and titer of the antiserum will be reflected in the conjugate and any purification procedures that increase these values, such as immunoaffinity chromatography (UNIT 10.11) will enhance conjugate performance.

The selection of an optimal conjugation time

for preparing alkaline phosphatase–antibody conjugates varies for different antibodies, in particular when monoclonal antibodies are used. In contrast, polyclonal antibodies may be reliably conjugated in 120 min.

Troubleshooting

There are several factors that may contribute to the production of poor enzyme-antibody conjugates. It is important to determine first whether a poor conjugate is the result of inactivation of either the antibody or the enzyme (or both) or the result of insufficient or excessive cross-linking. The affinity of the antibody for substrate can be measured by determining the presence of bound antibody with another immunoassay employing anti-antibody conjugated to a different enzyme. Enzyme activity can be measured by cleavage of substrate at different enzyme concentrations. Precipitation of material in the conjugate solution or opaque solutions are indicative of excessive cross-linking. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (UNIT 10.2) is useful for monitoring the extent of cross-linking by determining the M_r of the cross-linked species.

Insufficient cross-linking usually results from the use of inactive or poor-quality cross-linking agents. Try using fresh reagents or different lots of reagent. Excessive cross-linking and inactivation of antibody or enzyme can be eliminated by either reducing the concentration of antibody and enzyme or by reducing the time of reaction.

It may not be possible to generate effective alkaline phosphatase conjugates with all antibodies using the one-step glutaraldehyde method. An alternative is to try a different conjugation technique (see Jeanson et al., 1988). Another alternative is to use an anti-species antibody–alkaline phosphatase conjugate to detect the antibody in question. These reagents may be purchased or prepared using the above technique.

Anticipated Results

Most monoclonal antibodies will couple to alkaline phosphatase very quickly, and a 5-min conjugation time will often be optimal. In general, polyclonal antibodies will take longer to conjugate, usually between 1 and 2 hr. The yield and titer of the resultant conjugate will depend on the original antibody's properties and spe-

cific application.

It is difficult to estimate the yield or working dilution of the conjugates, as it is dependent on numerous factors such as antibody affinity, type of ELISA, and quality of antigen. In general, the working dilutions range from 1:100 to 1:10,000.

Time Considerations

The total time for conjugation is 1 to 3 days, with working times of several hours per conjugation. The majority of time is spent dialyzing or stirring. Once started, a protocol should be completed as described because the reactant products and solutions are unstable.

Literature Cited

- Healey, K., Chandler, H.M., Cox, J.C., and Hurrell, J.G.R. 1983. A rapid semi-quantitative capillary enzyme immunoassay for digoxin. *Clin. Chim. Acta* 134:51-58.
- Jeanson, A., Cloes, J.-M., Bouchet, M., and Rentier, B. 1988. Comparison of conjugation procedures for the preparation of monoclonal antibody-enzyme conjugates. *J. Immunol. Methods* 111:261-270.
- Nakane, P.K. and Kawaoi, A. 1974. Peroxidase-labeled antibody. A new method of conjugation. *J. Histochem. Cytochem.* 22:1084-1091.
- Tijssen, P. and Kurstak, E. 1984. Highly efficient and simple methods for the preparation of peroxidase and active peroxidase-antibody conjugates for enzyme immunoassay. *Anal. Biochem.* 136:451-457.
- Voller, A., Bidwell, D.E., and Barlett, A. Enzyme immunoassays in diagnostic medicine. *Bull. W.H.O.* 53:55-65.

Key Reference

- Van Vunakis, H. and Langone, J.J., eds. 1980. Immunochemical techniques. *Methods Enzymol.* 70:1-525.

An excellent collection of articles on immunoassay techniques, including several on enzyme-antibody conjugation techniques.

Contributed by Scott E. Winston,
Steven A. Fuller, and Michael J. Evelegh
ADI Diagnostics
Rexdale, Ontario

John G.R. Hurrell
Boehringer Mannheim
Indianapolis, Indiana

Enzyme-Linked Immunosorbent Assays (ELISA)

This unit describes six different ELISA systems for the detection of specific antibodies, soluble antigens, or cell-surface antigens. In all six systems, soluble reactants are removed from solution after specifically binding to solid-phase reactants. Table 11.2.1 summarizes the different ELISA protocols, which are illustrated in Figures 11.2.1-11.2.6.

In the first four protocols, solid-phase reactants are prepared by adsorbing an antigen or antibody onto plastic microtiter plates; in the next two protocols, the solid-phase reactants are cell-associated molecules. In all protocols, the solid-phase reagents are incubated with secondary or tertiary reactants covalently coupled to an enzyme.

Table 11.2.1 Summary of ELISA Protocols

ELISA protocol	Uses	Required reagents	Comments
Indirect	Antibody screening; epitope mapping	Antigen, pure or semipure; test solution containing antibody; enzyme conjugate that binds Ig of immunized species	Does not require the use of preexisting specific antibodies; requires relatively large amounts of antigen
Direct competitive	Antigen screening; detect soluble antigen	Antigen, pure or semipure; test solution containing antigen; enzyme-antibody conjugate specific for antigen	Rapid assay with only two steps; excellent for measuring antigenic cross-reactivity
Antibody-sandwich	Antigen screening; detect soluble antigen	Capture antibody (purified or semi-purified specific antibody); test solution containing antigen; enzyme-antibody conjugate specific for antigen	Most sensitive antigen assay; requires relatively large amounts of pure or semi-pure specific antibody (capture antibody)
Double antibody-sandwich	Antibody-screening; epitope mapping	Capture antibody: (specific for Ig of immunized species); test solution containing antigen; enzyme-antibody conjugate specific for antigen	Does not require purified antigen; relatively long assay with five steps
Direct cellular	Screen cells for expression of antigen; measure cellular antigen expression	Cells that express antigen of interest; enzyme-antibody conjugate specific for cellular antigen	Sensitive assay for bulk screening; insensitive to heterogeneity of expression in mixed population of cells
Indirect cellular	Screen for antibodies against cellular antigens	Cells used for immunizing; test solution containing antibodies; enzyme conjugate that binds Ig of immunized species	May not detect antibodies specific for cellular antigens expressed at a low density

Unbound conjugates are washed out and a chromogenic or fluorogenic substrate is added. As the substrate is hydrolyzed by the bound enzyme conjugate, a colored or fluorescent product is generated. Finally, the product is detected visually or with a microtiter plate reader. The amount of product generated is proportional to the amount of analyte in the test mixture. The first support protocol can be used to optimize the different ELISAs. The second support protocol provides a method for preparing lysates for use as test antigen from bacterial cultures containing expressed protein.

BASIC PROTOCOL

INDIRECT ELISA TO DETECT SPECIFIC ANTIBODIES

This assay is useful for screening antisera or hybridoma supernatants for specific antibodies when milligram quantities of purified or semipurified antigen are available (1 mg of purified antigen will permit screening of 80 to 800 microtiter plates; Fig. 11.2.1). Antibodies are detected by coating the wells of microtiter plates with antigen, incubating the coated plates with test solutions containing specific antibodies, and washing away unbound antibodies. A solution containing a developing reagent, (e.g., alkaline phosphatase conjugated to protein A, protein G, or antibodies against the test solution antibodies) is then added to the plate. After incubation, unbound conjugate is washed away and substrate solution is added. After a second incubation, the amount of substrate hydrolyzed is assessed with a spectrophotometer or spectrofluorometer. The measured amount is proportional to the amount of specific antibody in the test solution. Visual inspection can also be used to detect hydrolysis.

Materials

Developing reagent: protein A–alkaline phosphatase conjugate (Sigma #P9650), protein G–alkaline phosphatase conjugate (Calbiochem #539304), or anti-Ig-alkaline phosphatase conjugate (*UNIT 11.1*)

Antigen solution

PBS (*APPENDIX 2*) containing 0.05% NaN₃ (PBSN)

Water, deionized or distilled

Blocking buffer

Test antibody samples

4-methylumbelliferyl phosphate (MUP) or *p*-nitrophenyl phosphate (NPP) substrate solution

0.5 M NaOH (optional)

Multichannel pipet and disposable pipet tips

Immulon 2 (Dynatech #011-010-3450), Immulon 4 (Dynatech #011-010-3850), or equivalent microtiter plates

Plastic squirt bottles

Microtiter plate reader (optional)—spectrophotometer with 405-nm filter or spectrofluorometer (Dynatech #011-970-1900) with 365-nm excitation filter and 450-nm emission filter

Determine developing reagent and antigen concentrations

1. Determine the optimal concentration of the developing reagent (conjugate) by criss-cross serial dilution analysis (see first support protocol).

Good conjugates of many specificities are available commercially. Choice of developing reagent (i.e., conjugate) is determined by the goals of the assay. If it is necessary to detect all antibodies that bind to antigen, conjugates prepared with antibodies specific for Ig κ and λ light chains should be used. Alternatively, protein A– or protein G–enzyme conjugates may be preferable when screening monoclonal antibodies. Specific monoclonal antibodies that bind protein A or protein G are easy to purify and characterize.

- Determine the final concentration of antigen coating reagent by criss-cross serial dilution analysis (see first support protocol). Prepare an antigen solution in PBSN at this final concentration. The final concentration of antigen is usually 0.2 to 10.0 $\mu\text{g/ml}$. Prepare ~ 6 ml antigen solution for each plate.

Pure antigen solution concentrations are usually ≤ 2 $\mu\text{g/ml}$. Although pure antigen preparations are not essential, $>3\%$ of the protein in the antigen solution should be the antigen. The total concentration of protein in the antigen solution should be increased for semipurified antigen preparations. Do not raise the total protein concentration in the antigen solution to >10 $\mu\text{g/ml}$, since this concentration usually saturates $>85\%$ of the available sites on Immulon microtiter plates. For some antigens, coating may occur more efficiently at different pHs.

Coat plate with antigen

- Using a multichannel pipet and tips, dispense 50 μl antigen solution into each well of an Immulon microtiter plate. Tap or shake the plate to ensure that the antigen solution is evenly distributed over the bottom of each well.

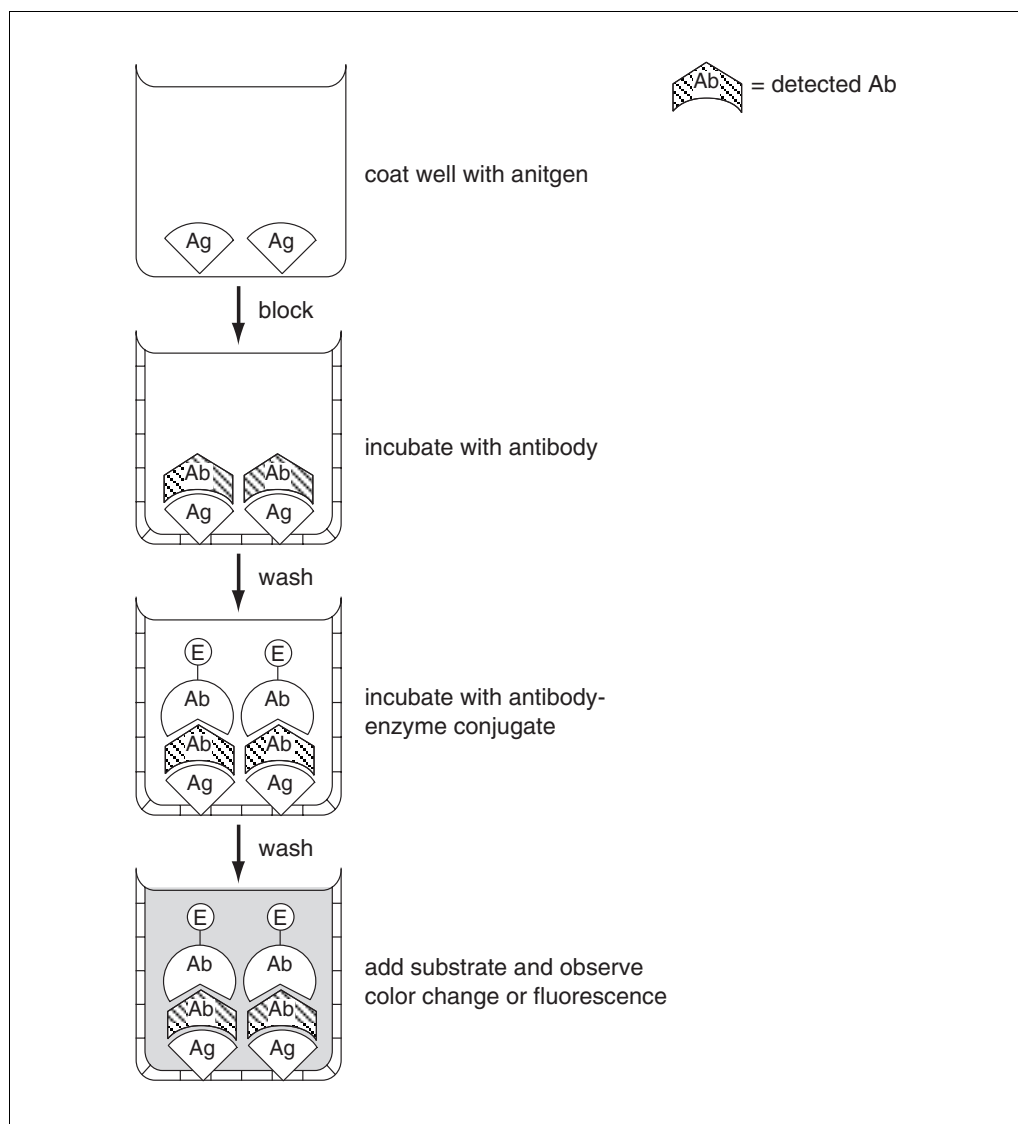


Figure 11.2.1 Indirect ELISA to detect specific antibodies. Ag = antigen; Ab = antibody; E = enzyme.

4. Wrap coated plates in plastic wrap to seal and incubate overnight at room temperature or 2 hr at 37°C.

Individual adhesive plate sealers are sold commercially but plastic wrap is easier to use and works as well. Sealed plates can be stored at 4°C with antigen solution for months.

5. Rinse coated plate over a sink by filling wells with deionized or distilled water dispensed either from a plastic squirt bottle or from the tap. Flick the water into the sink and rinse with water two more times, flicking the water into the sink after each rinse.

Block residual binding capacity of plate

6. Fill each well with blocking buffer dispensed as a stream from a squirt bottle and incubate 30 min at room temperature.

Residual binding capacity of the plate is blocked in this step. Tween 20 (0.05%) by itself is more effective at blocking than any protein tested, but because the combination of protein and Tween 20 may be more effective than Tween 20 alone in some cases, bovine serum albumin (BSA; 0.25%) is included in the blocking buffer.

7. Rinse plate three times in water as in step 5. After the last rinse, remove residual liquid by wrapping each plate in a large paper tissue and gently flicking it face down onto several paper towels laying on the benchtop.

Rinsing with water is cheaper and easier than rinsing with buffered solutions and is as effective.

Add antibody to plate

8. Add 50 µl antibody samples diluted in blocking buffer to each of the coated wells, wrap plate in plastic wrap, and incubate ≥2 hr at room temperature.

While enough antibody may be bound after 1 to 2 hr to generate a strong signal, equilibrium binding is generally achieved after 5 to 10 hr. Thus, the specific signal may be significantly increased by longer incubations.

For this and all steps involving the delivery of aliquots of many different solutions to microtiter plates with multichannel pipets, such as the primary screening of hybridoma supernatants, the same pipet tips can be reused for hundreds of separate transfers. Wash tips between transfers by expelling any liquid remaining in the tips onto an absorbent surface of paper tissues, rinsing tips five times in blocking buffer, and carefully expelling any residual liquid from tips onto the tissues. Avoid bubbles in the tips; any tip with intractable bubbles should be replaced.

Wash the plate

9. Rinse plate three times in water as in step 5.
10. Fill each well with blocking buffer, vortex, and incubate 10 min at room temperature.

Plates are vortexed to remove any reagent remaining in the corners of the wells.

11. Rinse three times in water as in step 5. After the final rinse, remove residual liquid as in step 7.

Add developing reagent to plate

12. Add 50 µl developing reagent in blocking buffer (at optimal concentration determined in step 1) to each well, wrap in plastic wrap, and incubate ≥2 hr at room temperature.

The strength of the signal may be increased by longer incubations (see annotation to step 8).

13. Wash plates as in steps 9 to 11.

After final rinsing, plates may be wrapped in plastic wrap and stored for months at 4°C prior to adding substrate.

Add substrate and measure hydrolysis

14. Add 75 µl MUP or NPP substrate solution to each well and incubate 1 hr at room temperature.
15. Monitor hydrolysis qualitatively by visual inspection or quantitatively with a microtiter plate reader (see below). Hydrolysis can be stopped by adding 25 µl of 0.5 M NaOH.
 - a. Visually, hydrolysis of NPP can be detected by the appearance of a yellow color. If using a microtiter plate reader to measure NPP hydrolysis, use a 405-nm filter.
 - b. Visually, hydrolysis of MUP can be monitored in a darkened room by illumination with a long-wavelength UV lamp. If using a microtiter plate spectrofluorometer to measure MUP hydrolysis, use a 365-nm excitation filter and a 450-nm emission filter.

The fluorogenic system using the MUP substrate is 10 to 100 times faster than the chromogenic system using NPP. Furthermore, the rate of spontaneous hydrolysis of MUP is much lower than that of NPP.

To detect bound antibodies that are present at low concentration, measure hydrolysis at a later time. To calculate when to measure hydrolysis the second time, remember that the amount of hydrolysis is almost directly proportional to the time of hydrolysis. For example, if the hydrolysis in the wells of interest reads 200 at 1 hr and a reading of 2000 is desired, incubate the plate ~10 hr before taking the second reading.

DIRECT COMPETITIVE ELISA TO DETECT SOLUBLE ANTIGENS

This assay is used to detect or quantitate soluble antigens and is most useful when both a specific antibody and milligram quantities of purified or semipurified antigen are available (Fig. 11.2.2). To detect soluble antigens, plates are coated with antigen and the binding of specific antibody-enzyme conjugates to antigen-coated plates is inhibited by test solutions containing soluble antigen. After incubation with mixtures of the conjugate and inhibitor in antigen-coated wells, unbound conjugate is washed away and substrate is added. The amount of antigen in the test solutions is proportional to the inhibition of substrate hydrolysis and can be quantitated by interpolation onto an inhibition curve generated with serial dilutions of a standard antigen solution.

The direct assay may also be adapted as an indirect assay by substituting specific antibody for specific antibody-enzyme conjugate. The amount of specific antibody bound is then detected using a species-specific or isotype-specific conjugate as a tertiary reactant.

Additional Materials

- Specific antibody–alkaline phosphatase conjugate (*UNIT 11.1*)
- Standard antigen solution
- Test antigen solutions
- Round- or cone-bottom microtiter plates

1. Determine the optimal concentration of coating reagent and antibody–alkaline phosphatase conjugate by criss-cross serial dilution analysis in which the concentrations of both the antigen (coating reagent) and the conjugate (developing reagent) are varied (see first support protocol). Prepare a 2× conjugate solution by diluting the

ALTERNATE PROTOCOL

Immunology

11.2.5

specific antibody–alkaline phosphatase conjugate in blocking buffer to twice the optimal concentration.

The final concentration is usually 25 to 500 ng antibody/ml. Prepare 3 ml antibody–alkaline phosphatase conjugate for each plate.

2. Coat and block wells of an Immulon microtiter plate with 50 μ l antigen solution as in steps 2 to 7 of the basic protocol.
3. Prepare six 1:3 serial dilutions of standard antigen solution in blocking buffer (see first support protocol for preparation of serial dilutions)—these antigen concentrations will be used in preparing a standard inhibition curve (see step 10).

Antigen concentrations should span the dynamic range of inhibition. The dynamic range of inhibition is defined as that range of inhibitor concentrations wherein changes of inhibitor concentration produce detectable changes in the amount of inhibition. The dynamic range of inhibition is empirically determined in an initial assay in which antigen concentration is typically varied from the micromolar (10^{-6} M) to the picomolar (10^{-12} M) range. For most protein antigens, initial concentration should be ~ 10 μ g/ml, followed by nine 1:4 serial dilutions in blocking buffer. These antigen dilutions are assayed for their ability to inhibit the binding of conjugate to antigen-coated plates under standard assay conditions. From this initial assay, six 1:3 antigen dilutions spanning the dynamic range of inhibition are selected for further use as standard antigen-inhibitor dilutions. Prepare ≥ 75 μ l of each dilution for each plate to be assayed.

Inhibitor curves are most sensitive in the region of the curve where small changes in inhibitor concentrations produce maximal changes in the amount of inhibition. This

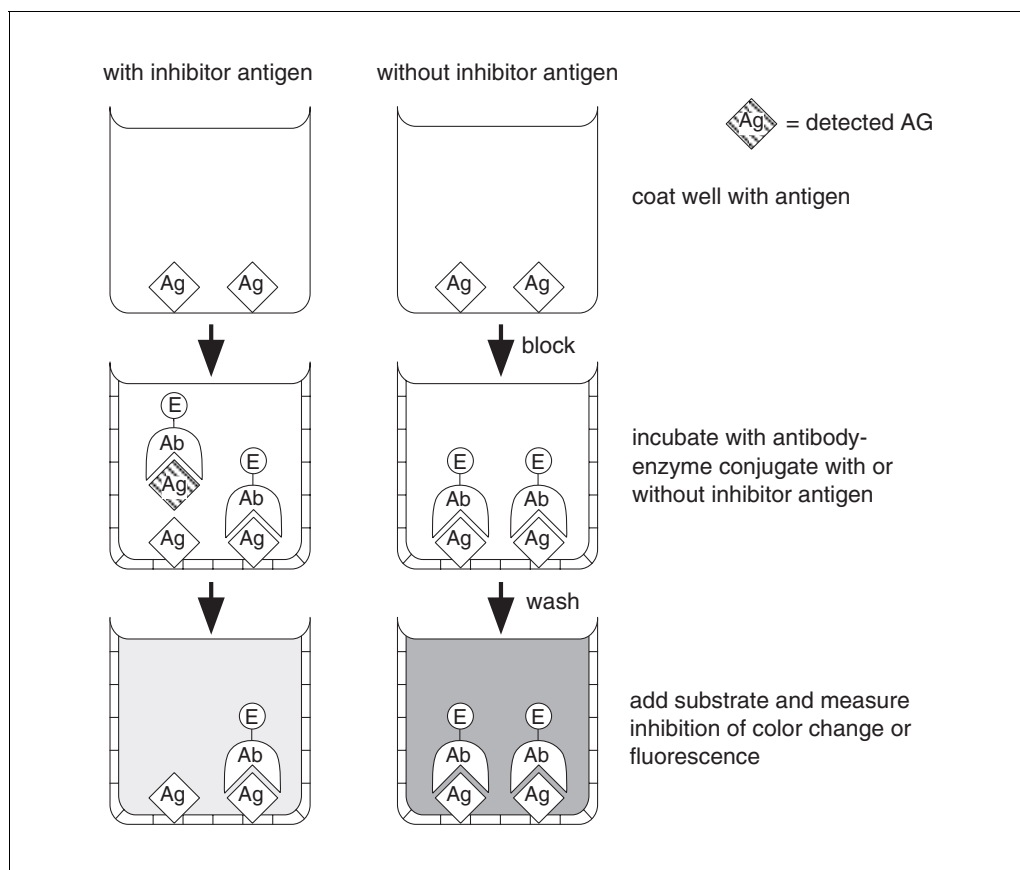


Figure 11.2.2 Direct competitive ELISA to detect soluble antigens. Ag = antigen; Ab = antibody; E = enzyme.

region of the curve normally spans 15% to 85% inhibition. In most systems, this range of inhibition is produced by concentrations of inhibitor between 1 and 250 ng/ml.

4. Mix and incubate conjugate and inhibitor by adding 75 μ l of 2 \times conjugate solution (from step 1) to each well of a round- or cone-bottom microtiter plate, followed by 75 μ l inhibitor—either test antigen solution or standard antigen solution (from step 3). Mix the conjugate and inhibitor solutions by pipetting up and down in the pipet tip three times (see annotation to step 8 in the basic protocol) and incubate \geq 30 min at room temperature.

For accurate quantitation of the amount of antigen in the test solutions, test antigen solutions should inhibit conjugate binding between 15% to 85%. It may be necessary to assay two or three different dilutions of the test solutions to produce inhibitions within this range.

5. Prepare uninhibited control samples by mixing equal volumes of 2 \times conjugate solution and blocking buffer.
6. Transfer 50 μ l of the mixture of conjugate plus inhibitor (from step 4) or conjugate plus blocking buffer (from step 5) to an antigen-coated plate (from step 2) and incubate 2 hr at room temperature.

If samples are to be assayed in duplicate, the duplicates should be in adjacent columns on the same plate. Reserve column 11 for uninhibited control samples (step 5) and column 12 for substrate alone without any conjugate. If the concentration of antigen in the test samples is to be accurately quantitated, dilutions of homologous antigen solutions (step 3) should be included on each plate.

7. Wash plate as in steps 9 to 11 of the basic protocol.
8. Add 75 μ l of MUP or NPP substrate solution to each well and incubate 1 hr at room temperature.
9. Read plates on the microtiter plate reader after \geq 1 hr, at which time enough substrate has been hydrolyzed in the uninhibited reactions to permit accurate measurement of the inhibition.
10. Prepare a standard antigen-inhibition curve constructed from the inhibitions produced by the dilutions of the standard antigen solutions from step 3. Plot antigen concentration on the x axis, which is a log scale, and fluorescence or absorbance on the y axis, which is a linear scale.
11. Interpolate the concentration of antigen in the test solutions from the standard antigen-inhibition curve.

The dynamic range of the inhibition curve may deviate from linearity if the specific antibodies are heterogeneous and possess significantly different affinities or if the standard antigen preparation contains heterogeneous forms of the antigen. Antigen concentration in test samples can be accurately interpolated from the inhibition curve as long as the test antigen is antigenically identical to the standard antigen and the concentration of test antigen falls within the dynamic range of inhibition.

ANTIBODY-SANDWICH ELISA TO DETECT SOLUBLE ANTIGENS

Antibody-sandwich ELISAs may be the most useful of the immunosorbent assays for detecting antigen because they are frequently between 2 and 5 times more sensitive than those in which antigen is directly bound to the solid phase (Fig. 11.2.3). To detect antigen, the wells of microtiter plates are coated with specific (capture) antibody followed by incubation with test solutions containing antigen. Unbound antigen is washed out and a different antigen-specific antibody conjugated to enzyme (i.e., developing reagent) is added, followed by another incubation. Unbound conjugate is washed out and substrate is added. After another incubation, the degree of substrate hydrolysis is measured. The amount of substrate hydrolyzed is proportional to the amount of antigen in the test solution.

Additional Materials

Specific antibody or immunoglobulin fraction from antiserum or ascites fluid, *or* hybridoma supernatant (UNIT 11.10), *or* bacterial lysate (second support protocol)

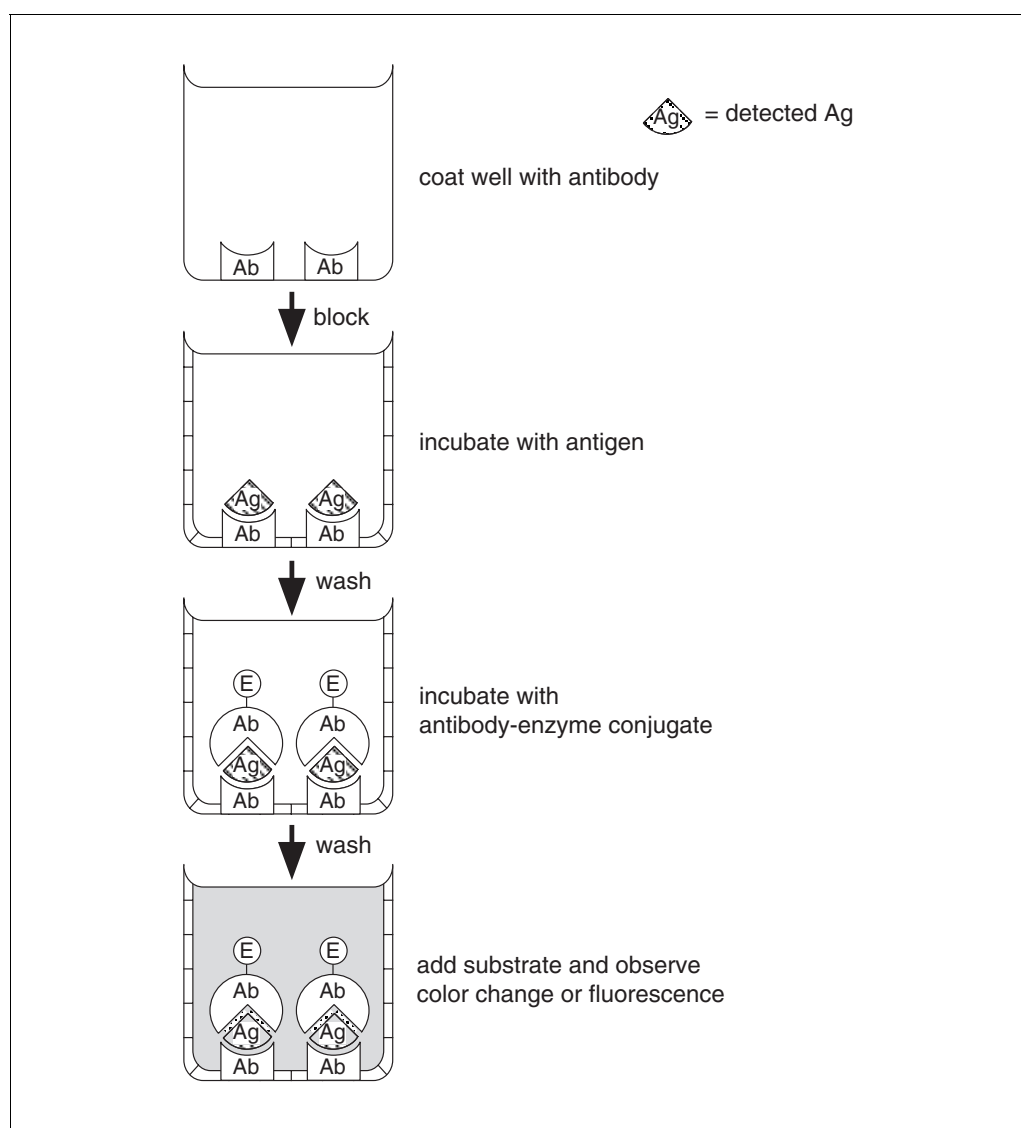


Figure 11.2.3 Antibody-sandwich ELISA to detect antigen. Ag = antigen; Ab = antibody; E = enzyme.

1. Prepare the capture antibody by diluting specific antibody or immunoglobulin fraction in PBSN to a final concentration of 0.2 to 10 $\mu\text{g/ml}$.

The capture antibodies can be monoclonal or polyclonal.

If the immunoglobulin fraction from an antiserum or ascites fluid is used, the concentration of total protein may need to be increased to compensate for the lower content of specific antibody. Little advantage is gained by increasing the total protein concentration in the capture antibody solution beyond 10 $\mu\text{g/ml}$.

2. Determine the concentration of capture antibody and conjugate necessary to detect the desired concentration of antigen by criss-cross serial dilution analysis (see first support protocol). Prepare a capture antibody solution in PBSN at this concentration.
3. Coat wells of an Immulon plate with capture-antibody solution as in steps 3 to 5 of the basic protocol.
4. Block wells as in steps 6 and 7 of the basic protocol.
5. Prepare a standard antigen-dilution series by successive 1:3 dilutions of the homologous antigen stock in blocking buffer (see first support protocol).

In order to measure the amount of antigen in a test sample, the standard antigen-dilution series needs to span most of the dynamic range of binding. This range typically spans from 0.1 to 1000 ng antigen/ml. The dynamic range of binding is defined as that range of antigen concentrations wherein small, incremental changes in antigen concentration produce detectable differences in the amount of antigen bound (see annotation to step 3, in the preceding alternate protocol). In most assay systems, the amount of antigen in a test solution is most accurately interpolated from the standard curve if it produces between 15% to 85% of maximal binding.

NOTE: While standard curves are necessary to accurately measure the amount of antigen in test samples, they are unnecessary for qualitative “yes/no” answers.

6. Prepare dilutions of test antigen solutions in blocking buffer.

It may be necessary to assay one or two serial dilutions of the initial antigen test solution to ensure that at least one of the dilutions can be accurately measured. For most assay systems, test solutions containing 1 to 100 ng/ml of antigen can be accurately measured.

7. Add 50- μl aliquots of the antigen test solutions and the standard antigen dilutions (from step 5) to the antibody-coated wells and incubate ≥ 2 hr at room temperature.

For accurate quantitation, samples should be run in duplicate or triplicate, and the standard antigen-dilution series should be included on each plate (see step 5). Pipetting should be performed rapidly to minimize differences in time of incubation between samples.

8. Wash plate as in steps 9 to 11 of the basic protocol.
9. Add 50 μl specific antibody-alkaline phosphatase conjugate and incubate 2 hr at room temperature.

The conjugate concentration is typically 25 to 400 ng specific antibody/ml.

When the capture antibody is specific for a single determinant, the conjugate must be prepared from antibodies which recognize different determinants that remain available after the antigen is bound to the plate by the capture antibody.

10. Wash plate as in steps 9 to 11 of the basic protocol.
11. Add 75 μl of MUP or NPP substrate solution to each well and incubate 1 hr at room temperature.

12. Read the plate on a microtiter plate reader.

To quantitate low-level reactions, the plate can be read again after several hours of hydrolysis.

13. Prepare a standard curve constructed from the data produced by serial dilutions of the standard antigen (step 5). Plot antigen concentration on the x axis which is a log scale, and fluorescence or absorbance on the y axis which is a linear scale.
14. Interpolate the concentration of antigen in the test solutions from the standard curve.

ALTERNATE PROTOCOL

DOUBLE ANTIBODY–SANDWICH ELISA TO DETECT SPECIFIC ANTIBODIES

This assay is especially useful when screening for specific antibodies in cases when a small amount of specific antibody is available and purified antigen is unavailable (Fig. 11.2.4). Additionally, this method can be used for epitope mapping of different monoclonal antibodies that are directed against the same antigen. Plates are coated with capture antibodies specific for immunoglobulin from the immunized species. The test antibody solution is incubated on the plates coated with the capture antibodies. Plates are then washed, incubated with antigen, washed again, and incubated with specific antibody conjugated to an enzyme. After incubation, unbound conjugate is washed out and substrate is added. Wells that are positive for hydrolysis may contain antibodies specific for the antigen.

Additional Materials

Capture antibodies specific for immunoglobulin from the immunized species
Specific antibody–alkaline phosphatase conjugate

1. Coat wells of an Immulon microtiter plate with 50 μ l of 2 to 10 μ g/ml capture antibodies as in steps 2 to 5 of the basic protocol.

NOTE: Capture antibodies must not bind the antigen or conjugate antibodies. When analyzing hybridoma supernatants or ascites fluid, coat plates with 2 μ g/ml capture antibody. When analyzing antisera, coat plates with 10 μ g/ml capture antibody.

2. Block wells as in steps 6 and 7 of the basic protocol.
3. Prepare dilutions of test antibody solutions in blocking buffer. Add 50 μ l to coated wells and incubate ≥ 2 hr at room temperature.

Hybridoma supernatants, antisera, or ascites fluid can be used as the test samples. Dilute hybridoma supernatants 1:5 and antisera or ascites fluid 1:200.

4. Wash plate as in steps 9 to 11 of the basic protocol.
5. Prepare an antigen solution in blocking buffer containing 20 to 200 ng/ml antigen.

Although purified antigen preparations are not essential, the limit of detectability for most protein antigens in this type of system is 2 to 20 ng/ml. A concentration of 20 to 200 ng antigen/ml is recommended.

6. Add 50- μ l aliquots of the antigen solution to antibody-coated wells and incubate ≥ 2 hr at room temperature.
7. Wash plate as in steps 9 to 11 of the basic protocol.
8. Add 50 μ l specific antibody–alkaline phosphatase conjugate to the wells and incubate 2 hr at room temperature.

The conjugate antibodies must not react with the capture antibody or the test antibody. The conjugate concentration is typically between 25 to 500 ng specific antibody/ml, and should be high enough to result in ~0.50 absorbance units/hr at 405 nm when using NPP as a substrate or a signal of 1000 to 1500 fluorescence units/hr when using MUP as a substrate. If no specific antibodies from the appropriate species are available to serve as a positive control, then a positive control system should be constructed out of available reagents. Such reagents can be found in Linscott's Directory of Immunological and Biological Reagents.

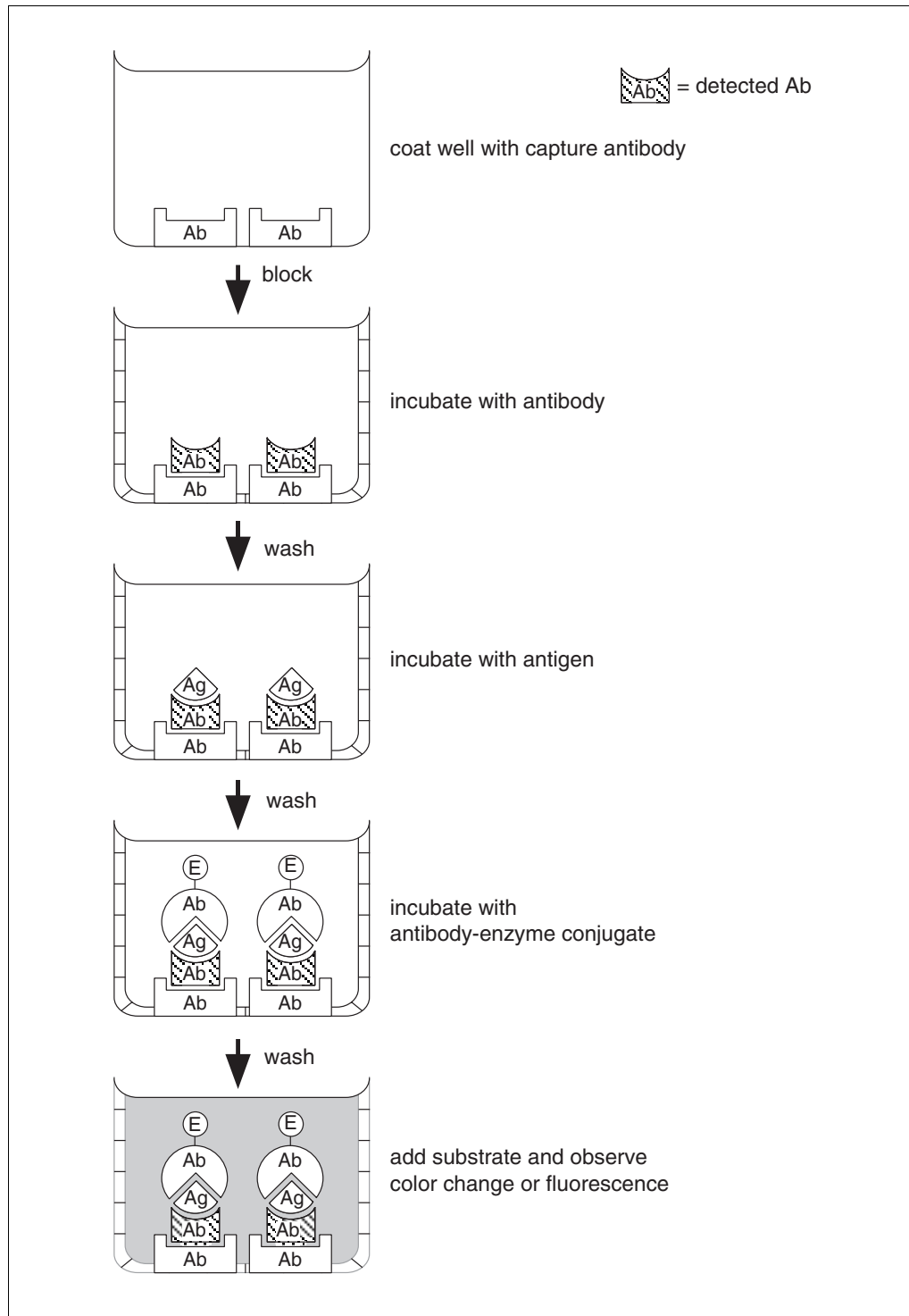


Figure 11.2.4 Double antibody-sandwich ELISA to detect specific antibodies. Ag = antigen; Ab = antibody; E = enzyme.

9. Wash plate as in steps 9 to 11 of the basic protocol.
10. Add 75 μ l of MUP or NPP substrate solution to each well and incubate 1 hr at room temperature. After 1 hr, examine hydrolysis visually or spectrophotometrically (see step 15 of the basic protocol).

In order to detect low-level reactions, the plate can be read again after several hours or days of hydrolysis.

11. Check for false positives by rescreening samples that test positive for antigen-specific antibody. For each positive sample, coat four wells with capture antibody and arm the capture antibody with test antibody (steps 1 to 4). Incubate two of the wells with antigen (steps 5 to 7) and two of the wells with blocking buffer. Add conjugate and substrate to all four wells (steps 8 to 10) and measure hydrolysis after 1 hr.

This procedure will eliminate false positives resulting from test antibodies that react with the enzyme-antibody complex.

ALTERNATE PROTOCOL

DIRECT CELLULAR ELISA TO DETECT CELL-SURFACE ANTIGENS

The expression of cell-surface antigens or receptors is measured using existing antibodies or other ligands specific for cell-surface molecules (Fig. 11.2.5). Cells are incubated with enzyme conjugated to antibodies that are specific for a cell-surface molecule. Unbound conjugate is washed away and substrate is added. The level of antigen expression is proportional to the amount of substrate hydrolysis. This procedure can be as sensitive as flow cytometry analysis in quantitating the level of antigen expression on a population of cells (Coligan et al., 1991). Unlike the flow cytometry analysis, however, this method is not sensitive for mixed populations. This assay can be converted to an indirect assay by substituting biotinylated antibody for the enzyme-antibody conjugate, followed by a second incubation with avidin-alkaline phosphatase.

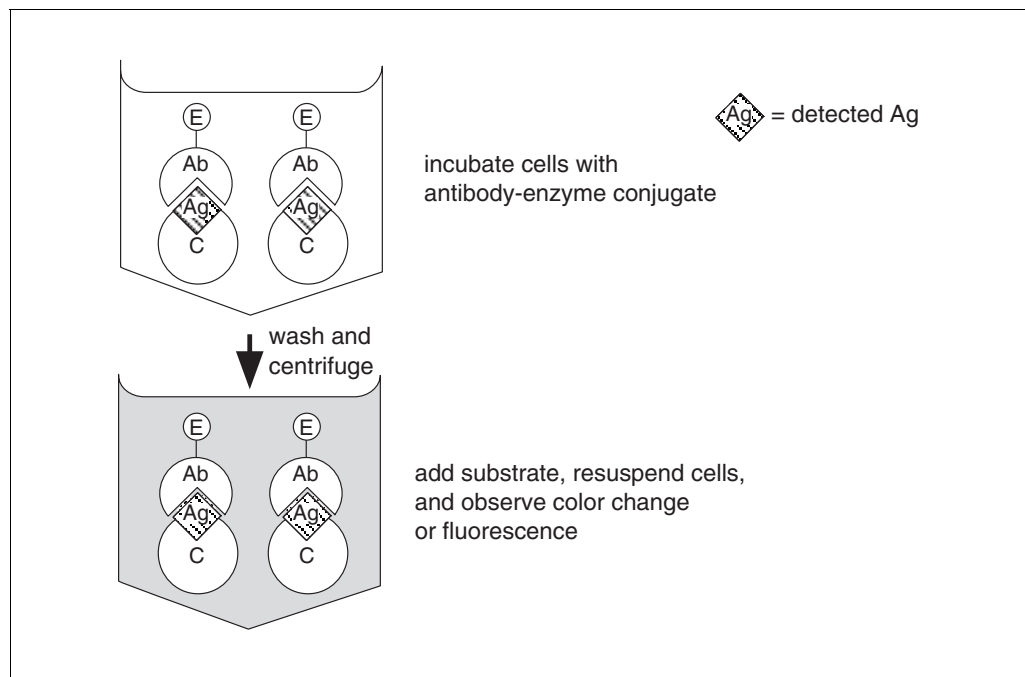


Figure 11.2.5 Direct cellular ELISA to detect cell-surface antigens. Ab = antibody; E = enzyme; C = cell.

Additional Materials

Cell samples
Specific antibody–alkaline phosphatase conjugate (see second support protocol)
Wash buffer, ice-cold

Cone- or round-bottom microtiter plates
Sorvall H-1000B rotor (or equivalent)

1. Determine the optimal number of cells per well and the antibody-conjugate concentration by criss-cross serial dilution analysis (see first support protocol) using variable numbers of positive- and negative-control cell samples and varying concentrations of antibody-biotin conjugate.

Titrate cells initially at $1-5 \times 10^5$ /well and conjugate at 0.5 to 10 $\mu\text{g/ml}$. For preparation and handling of cells, consult steps 2 to 5.

Because eukaryotic cells express variable amounts of alkaline phosphatase, test cells must be assayed in a preliminary experiment for alkaline phosphatase by incubation with substrate alone. If the test cells express unacceptable levels of alkaline phosphatase, another enzyme conjugate such as β -galactosidase should be used. Both chromogenic and fluorogenic substrates are available for β -galactosidase.

2. Centrifuge cell samples in a table-top centrifuge 5 min in Sorvall H-1000B rotor at 1500 rpm ($450 \times g$), 4°C , in a 15- to 50-ml centrifuge tube. Count cells (APPENDIX 3) and resuspend in ice-cold wash buffer at $1-5 \times 10^6$ cells/ml.

If the surface antigen retains its antigenicity after fixation, cells may be fixed at the beginning of the experiment—but do not fix cells unless it can be demonstrated that the antigenicity is retained after fixation. Fix cells by suspending in glutaraldehyde (0.5% final; from a 25% stock, EM grade Sigma #G5882), and incubating 30 min at room temperature. Pellet cells, resuspend in PBSLE (see second support protocol), and incubate for 30 min at 37°C . Wash twice in PBSLE and resuspend in wash buffer. Cells can be kept for months at 4°C after fixation.

3. Dispense 100 μl of cell suspension ($1-5 \times 10^5$ cells) into wells of cone- or round-bottom microtiter plates, and centrifuge 1 min at $450 \times g$, 4°C . Remove supernatant by vacuum aspiration, and disrupt pellet by briefly shaking microtiter plate on a vortex mixer or microtiter plate shaker.
4. Resuspend pellet in 100 μl of conjugate in ice-cold wash buffer at the optimal concentration (see step 1). Incubate 1.5 hr at 4°C , resuspending cells by gently shaking at 15-min intervals.

Be careful not to splash cell suspensions out of wells.

5. Centrifuge cells 1 min at $450 \times g$, 4°C , remove supernatant by vacuum aspiration, briefly vortex pellet, and resuspend in 200 μl ice-cold wash buffer. Repeat three times.
6. Add 100 μl MUP or NPP substrate solution. Incubate 1 hr at room temperature, resuspending cells by gently shaking at 15-min intervals during hydrolysis.
7. Determine extent of hydrolysis by visual inspection or using a microtiter plate reader.

INDIRECT CELLULAR ELISA TO DETECT ANTIBODIES SPECIFIC FOR SURFACE ANTIGENS

This assay is designed to screen for antibodies specific for cell-surface antigens (Fig. 11.2.6). Antibodies against surface antigens are detected by incubating whole cells with a test solution containing the primary antibody. The unbound antibody is washed away and the cells are then incubated with an enzyme conjugated to antibodies specific for the primary antibody. Unbound enzyme conjugate is washed away and substrate solution added. The level of bound primary antibody is proportional to the amount of substrate hydrolysis.

Additional Materials

Positive-control antibodies (i.e., those that react with the experimental cells and are from the immunized species)

Negative-control antibodies (i.e., those that do not react with the experimental cells)

Test antibody solution

Antibody or $F(ab')_2$ (against immunoglobulin from the immunized species) conjugated to alkaline phosphatase

Cone- or round-bottom microtiter plates

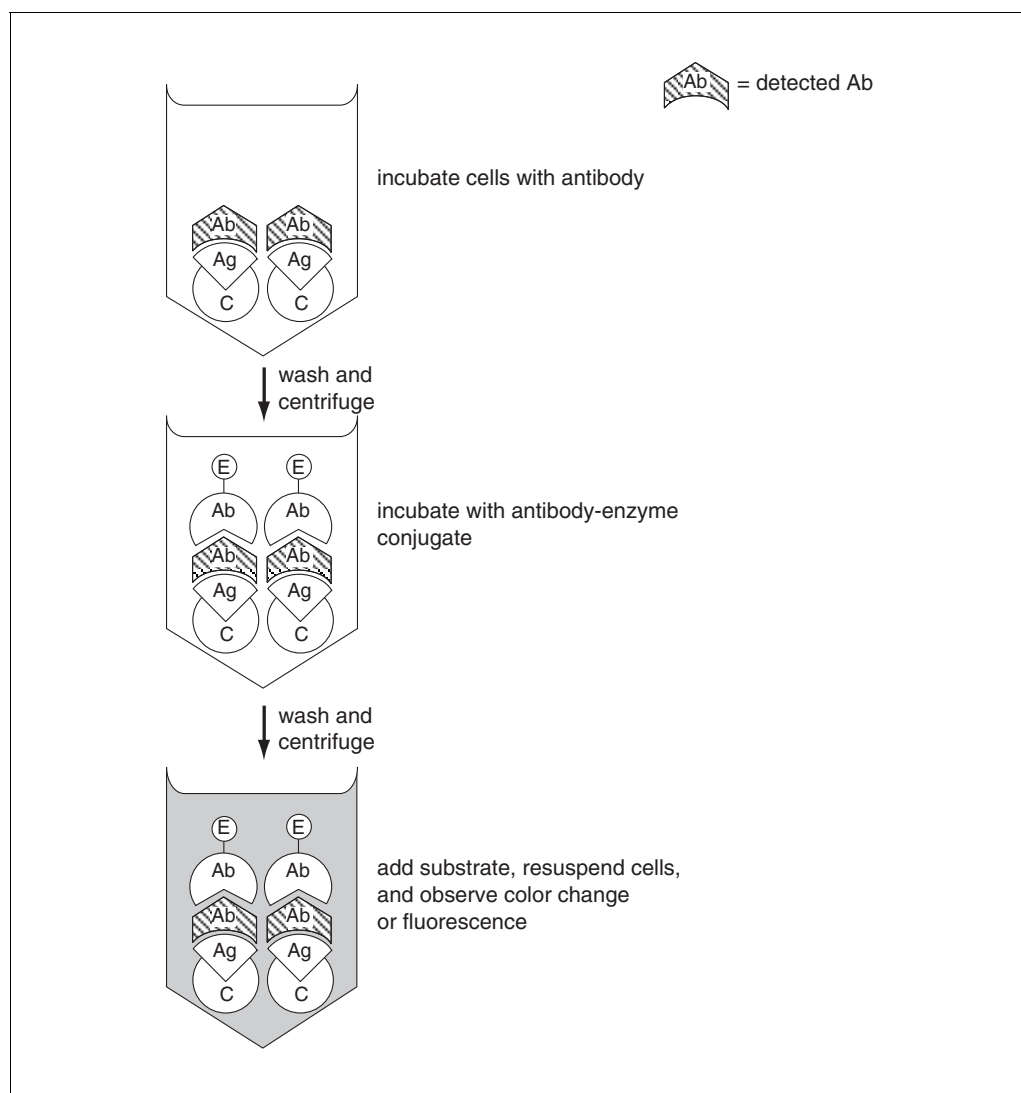


Figure 11.2.6 Indirect cellular ELISA to detect antibodies specific for surface antigens. Ab = antibody; E = enzyme; C = cell.

1. Centrifuge and resuspend cell samples as in step 2 of the previous alternate protocol at $1-5 \times 10^6$ cells/ml.

Because this technique detects antibodies against uncharacterized epitopes, fixation prior to analysis is not recommended. Fixation may destroy the antigenicity of the epitope. All steps must be performed at 4°C in physiological buffers containing NaN₃.

Because eukaryotic cells express variable amounts of alkaline phosphatase, test cells must be assayed for alkaline phosphatase activity. If the endogenous alkaline phosphatase level is too high, another enzyme should be substituted for alkaline phosphatase in the antibody-enzyme conjugate (see annotation to step 1 of the previous alternate protocol).

2. In preliminary assays, determine the optimal number of cells per well and conjugate concentration by criss-cross serial dilution analysis using positive- and negative-control antibodies instead of test antibodies (see first support protocol). In adapting the criss-cross serial dilution analysis, whole cells replace the solid-phase coating reagent; see techniques for handling cells are outlined in steps 3 to 8. Set up titrations by varying the number of cells between 1×10^5 and 5×10^5 /well, the concentration of positive- and negative-control antibodies between 0.1 and 10 µg/ml, and the concentration of antibody-enzyme conjugate between 0.1 and 10 µg/ml.
3. Dispense 100 µl of cell suspension ($1-5 \times 10^5$ cells) into wells of round- or cone-bottom microtiter plates. Centrifuge 1 min at 1500 rpm, 4°C, remove supernatant by vacuum aspiration, and disrupt pellet by briefly shaking microtiter plate on the vortex mixer.
4. Resuspend cells in 100 µl solutions containing 1 to 10 µg/ml test antibody or control antibodies in ice-cold wash buffer. Incubate 1.5 hr at 4°C, resuspending cells by gently shaking at 15-min intervals.

Be careful not to splash cell suspensions out of wells.

5. Centrifuge cells 1 min at 1500 rpm, 4°C, remove supernatant by vacuum aspiration, briefly vortex pellet, and resuspend in 200 µl ice-cold wash buffer. Repeat twice.
6. Resuspend pellet in 100 µl enzyme-antibody conjugate or F(ab')₂-enzyme conjugate diluted in ice-cold wash buffer. The optimal concentration of antibody, determined in step 2, is usually 100 to 500 ng/ml. Incubate 1.5 hr at 4°C, resuspending cells by gently shaking at 15-min intervals.

When working with cells that may express Fc receptors, it is best to use enzyme conjugated to F(ab')₂ fragments. F(ab')₂ fragments have had the Fc portion of the antibody enzymatically removed and no longer bind to Fc receptors.

7. Wash cells as in step 5. Repeat three times.
8. Add 100 µl MUP or NPP substrate solution. Allow hydrolysis to proceed until the signal has reached the desired levels; resuspend cells by gently shaking at 15 min intervals during hydrolysis. If desired, stop hydrolysis by adding 25 µl of 0.5 M NaOH.
9. Determine extent of hydrolysis by visual inspection or spectrophotometrically using a microtiter plate reader.

**CRISS-CROSS SERIAL DILUTION ANALYSIS TO DETERMINE
OPTIMAL REAGENT CONCENTRATIONS**

Serial dilution titration analyses are performed to determine optimal concentrations of reagents to be used in ELISAs. In this protocol, all three reactants in a three-step ELISA—a primary solid-phase coating reagent, a secondary reagent that binds the primary reagent, and an enzyme-conjugated tertiary developing reagent that binds to the secondary reagent—are serially diluted and analyzed by a criss-cross matrix analysis (Fig 11.2.7). Once the optimal concentrations of reagents to be used under particular assay conditions are determined, these variables are kept constant from experiment to experiment. The coating (primary), secondary, and developing (tertiary) reagents will vary depending upon which of the previous protocols needs to be optimized.

Additional Materials

- Coating reagent
- Secondary reagent
- Developing reagent
- 17 × 100–mm and 12 × 74–mm test tubes

		Secondary reactant												Rows
		homologous (antigen)						heterologous (antigen)						
		200	50	12.5	3.12	0.78	0	200	50	12.5	3.12	0.78	0	
Tertiary reactant (antibody-alkaline phosphatase) (ng/ml)	500	over	over	over	3200	1000	0	500	120	40	20	10		A
	250	over	over	over	2060	560	0	300	80	20	0	0		B
	125	over	over	3650	1370	360	0	195	40	10	10	0		C
	62.5	3600	4000	2270	790	240	0	120	30	10	10	10		D
	31.25	2700	2100	1200	410	120	0	60	10	10	10	0		E
	0	0	0	0	0	0	0	0	0	0	0	0		F
														G
														H
	1	2	3	4	5	6	7	8	9	10	11	12		
	Columns													

Figure 11.2.7 Results of a criss-cross serial dilution analysis (for optimization of secondary and tertiary reactant concentrations) of an antibody-sandwich ELISA to detect antigen. The numbers in columns 1 to 11 and rows B to G represent relative fluorescence units observed for each well on a 96-well microtiter plate.

Plates were coated overnight with the capture antibody at 2 μg/ml. The secondary reactants, 4-fold serial dilutions of the homologous antigen and a non-cross-reactive heterologous antigen, were incubated on the plate 2 hr. The tertiary reactant, 2-fold serial dilutions of specific antibody-alkaline phosphatase conjugates, were incubated on the plate 2 hr. After 1 hr of incubation with the substrate MUP, the fluorescence was read in a microtiter plate spectrofluorometer.

Reagent concentrations depend upon individual assay variables that are set by the investigator. If the time of hydrolysis is set at 1 hr, the relative fluorescence at ~1000 relative fluorescence units, and the sensitivity at 780 pg/ml of homologous antigen, then 500 ng/ml of enzyme-antibody conjugate must be used in the ELISA. If, however, the assay has to detect only 3.12 ng/ml of homologous antigen, then the concentration of conjugate can be reduced to 125 ng/ml. It should be noted by comparing the homologous with the heterologous reactions (wells B5 versus B11 and D4 versus D10) that both the specificity and the signal-to-noise ratio for this assay are excellent.

Prepare coating-reagent dilutions

1. Place four 17 × 100–mm test tubes in a rack and add 6 ml PBSN to the last three tubes. In tube 1, prepare a 12-ml solution of coating reagent at 10 µg/ml in PBSN. Transfer 6 ml of tube 1 solution to tube 2. Mix by pipetting up and down five times. Repeat this transfer and mix for tubes 3 and 4; the tubes now contain the coating reagent at 10, 5, 2.5, and 1.25 µg/ml.
2. Using a multichannel pipet, dispense 50 µl of the coating reagent solutions into wells of four Immulon microtiter plates (i.e., each plate is filled with one of the four dilutions). Incubate overnight at room temperature or 2 hr at 37°C.
3. Rinse and block plates with blocking buffer as in steps 5 to 7 of the basic protocol.

Prepare secondary-reagent dilutions

4. Place five 12 × 75–mm test tubes in a rack and add 3 ml blocking buffer to the last four tubes. In tube 1, prepare a 4-ml solution of secondary reagent at 200 ng/ml in PBSN. Transfer 1 ml of tube 1 solution to tube 2. Pipet up and down five times. Repeat this transfer and mix for tubes 3 to 5; the tubes now contain the secondary reactant at 200, 50, 12.5, 3.125, and 0.78 ng/ml. If possible, prepare and test serial dilutions of a nonreactive heterologous form of the secondary reactant in parallel (Fig. 11.2.7).

If the assay is especially insensitive, it may be necessary to increase the secondary reactant concentrations so the tube-1 solution is 1000 ng/ml.

5. Dispense 50 µl of the secondary reagent solutions into the first five columns of all four coated plates. The most dilute solution is dispensed into column 5, while solutions of increasing concentration are added successively into columns 4, 3, 2, and 1. Thus, the fifth column contains 0.78 ng/ml and the first column 200 ng/ml. Incubate 2 hr at room temperature.
6. Wash plates as in steps 9 to 11 of the basic protocol.

Prepare developing-reagent dilutions

7. Place five 17 × 100–mm test tubes in a rack and add 3 ml blocking buffer to the last four tubes. In tube 1, prepare a 6-ml solution of developing reagent at 500 ng/ml in blocking buffer. Transfer 3 ml of tube 1 solution into tube 2 and mix. Repeat this transfer and mixing for tubes 3 and 4—the tubes now contain the developing reagent at 500, 250, 125, 62.5, and 31.25 ng/ml.
8. Dispense 50 µl of the developing reagent solutions into the wells of rows 2 to 6 of each plate, dispensing the most dilute solution into row 6 and solutions of increasing concentration successively into rows 5, 4, 3, and 2. Incubate 2 hr at room temperature.
9. Wash plates as in steps 9 to 11 of the basic protocol.

Measure hydrolysis

10. Add 75 µl MUP or NPP substrate solution to each well, incubate 1 hr at room temperature, and measure the degree of hydrolysis visually or with a microtiter plate reader. An appropriate assay configuration results in 0.50 absorbance units/hr at 405 nm when using NPP as a substrate or 1000 to 1500 fluorescence units/hr when using MUP as a substrate.

These results can be used to adjust optimal concentrations in the basic and alternate protocols.

PREPARATION OF BACTERIAL CELL LYSATE ANTIGENS

A culture of *E. coli* containing proteins expressed from cloned genes is lysed for use as test antigen in any of the first three protocols of this unit. For more extensive discussion on protein expression for antigen production, see *UNITS 16.4-16.7* (expression by fusion protein vectors).

Materials

Escherichia coli culture in broth or agar (*UNITS 1.2 & 1.3*)

Cell resuspension buffer

Lysozyme solution

Tris/EDTA/NaCl (TEN) buffer (*UNIT 11.1*)

10% sodium dodecyl sulfate (SDS)

8 M urea (optional)

Nylon-tipped applicator (Falcon #2069, Becton Dickinson)

1. For liquid culture, centrifuge 5 ml of cells at 2500 rpm in a tabletop centrifuge for 10 min. Decant supernatant and resuspend pellet in 5 ml cell resuspension buffer by vortexing gently. For agar culture, remove about 10 colonies from the plate using a nylon-tipped applicator and resuspend in 2 ml cell resuspension buffer. Press swab against side of tube to remove as much liquid as possible.

Yield of expressed protein may vary with growth phase. Samples should be taken for analysis at various periods of growth (e.g., mid-log and stationary phases). If samples are taken from agar plates, the culture should be grown overnight at 37°C.

2. Place 1 ml of resuspended cells in a microcentrifuge tube on ice.
3. Add 0.2 ml lysozyme solution to the tube and leave 5 min on ice.
4. Microcentrifuge 5 min. Decant supernatant and save. Resuspend pellet in 1.2 ml TEN buffer.

Since many expressed proteins are insoluble, it is worthwhile to assay both the pellet and supernatant for activity.

5. Add 0.065 ml of 10% SDS solution to each sample. Incubate 10 min at 37°C. Samples are ready for ELISA at this point. Store frozen if not used within several hours.

Alternatively, add urea to a final concentration of 8 M (4.8 g to a final volume of 10 ml) to denature and solubilize proteins.

REAGENTS AND SOLUTIONS

Borate-buffered saline (BBS)

0.017 M Na₂B₄O₇·10H₂O

0.12 M NaCl

Adjust to pH 8.5 with NaOH

Blocking buffer

BBS (see above) containing:

0.05% Tween 20

1 mM EDTA

0.25% bovine serum albumin (BSA)

0.05% NaN₃

Store at 4°C

Gelatin may be substituted for BSA; 5% instant milk has been successfully used but may interfere nonspecifically with antibody binding.

Cell resuspension buffer (10 mM HEPES)

2.38 g HEPES

Add H₂O to 1 liter

Lysozyme solution

5 mg chicken egg white lysozyme (Sigma Grade VI #L2879)

1 ml TEN buffer (UNIT 11.1)

Make fresh immediately before use

MUP substrate solution

0.2 mM 4-methylumbelliferyl phosphate (MUP; Sigma #M8883)

0.05 M Na₂CO₃

0.05 mM MgCl₂

Store at room temperature

NPP substrate solution

3 mM *p*-nitrophenyl phosphate (NPP; Sigma #104-0)

0.05 M Na₂CO₃

0.05 mM MgCl₂

Store at 4°C

Test antibody solution

Hybridoma supernatants (UNIT 11.10) can usually be diluted 1:5 and ascites fluid and antisera (UNIT 11.12) diluted 1:500 in blocking buffer and still generate a strong positive signal. Dilutions of nonimmune ascites or sera should be assayed as a negative control. Prepare antibody dilutions in cone- or round-bottom microtiter plates before adding them to antigen-coated plates.

Sources of appropriate antibodies and conjugates can be found in Linscott's Directory of Immunological and Biological Reagents.

Test antigen solution

0.2 to 10 µg/ml antigen, purified or partially purified in PBSN; store at 4°C

Wash buffer

Hanks balanced salt solution (HBSS; APPENDIX 2)

1% fetal calf serum (FCS; heat-inactivated 60 min, 56°C)

0.05% NaN₃

Store at 4°C

COMMENTARY

Background Information

Since their first description in 1971 (Engvall and Perlman), ELISAs have become the system of choice when assaying soluble antigens and antibodies. Factors that have contributed to their success include their sensitivity, the long shelf-life of the reagents (alkaline phosphatase conjugates typically lose only 5% to 10% of their activity per year), the lack of radiation hazards, the ease of preparation of the reagents, the speed and reproducibility of the assays, and the variety of ELISA formats that can be generated with a few well-chosen reagents. Additionally, no sophisticated equipment is necessary for many ELISA applications, including

screening hybridoma supernatants for specific antibodies and screening biological fluids for antigen content.

The ELISAs described here combine the special properties of antigen-antibody interactions with simple phase separations to produce powerful assays for detecting biological molecules. The multivalency of antibodies can result in the formation of long-lived antigen-antibody complexes, thus allowing long periods of time during which such complexes can be measured. By designing an assay so that a capture reagent initiates the binding of antigen-antibody complexes and enzyme conjugates onto a solid phase, the unbound reagents can be easily and

rapidly separated from the solid phase. The solid phase is washed and the amount of bound conjugate is visualized by incubating the solid phase with a substrate that forms a detectable product when hydrolyzed by the bound enzyme. ELISAs are similar in principle to radioimmunoassays, except that the radioactive label is replaced by an enzyme conjugate.

A number of different enzymes have been successfully used in ELISAs, including alkaline phosphatase, horseradish peroxidase, β -galactosidase, glucoamylase, and urease. Alkaline phosphatase—perhaps the most widely used conjugated enzyme—is recommended because of its rapid catalytic rate, excellent intrinsic stability, availability, ease of conjugation, and resistance to inactivation by common laboratory reagents. Additionally, the substrates of alkaline phosphatase are nontoxic and relatively stable. Solutions of *p*-nitrophenyl phosphate (NPP) are stable for months at 4°C, while solutions of 4-methylumbelliferyl phosphate (MUP) can be kept for months at room temperature without any significant spontaneous hydrolysis. The biggest disadvantage of alkaline phosphatase is that if NPP is used as a substrate, the yellow color of the nitrophenyl product is relatively difficult to detect visually. Using the substrate MUP instead of NPP can greatly enhance the sensitivity of the assay. The fluorogenic system using MUP is 10 to 100 times faster than the chromogenic system using NPP, and appears to be as sensitive as an enhanced chromogenic assay in which alkaline phosphatase generates NAD⁺ from NADP (Macy et al., 1988). The disadvantage of using fluorogenic substrates is that they require a microplate fluorometer costing twice as much as a high-quality microtiter plate spectrophotometer.

Cellular ELISAs have been shown to be as sensitive as flow cytometry analysis in detecting some cell-surface antigens (Bartlett and Noelle, 1987) and are potentially of great value in rapidly screening hybridoma supernatants for antibodies against surface molecules (Feit et al., 1983). Using ELISAs for screening large numbers of hybridoma supernatants has been hindered by the large number of cells required and high background signal. The increased sensitivity of the fluorogenic system should reduce the number of cells needed by a factor of 5, making the system more useful as a screening assay.

In addition to the methods described here, hundreds of other ELISA applications have been described, including the determination of

antibody affinities (Beatty et al., 1987; Schots et al., 1988), the detection of antibodies specific for hormone receptors (Quinn et al., 1988; Wang and Leung, 1985), expression cloning of lymphokine receptors (Harada et al., 1990), and homogeneous assays in which a solid phase is not needed because the antigen-antibody interaction itself modifies the enzymatic activity (Rubenstein et al., 1972). A number of books are devoted to ELISAs and can be consulted for further discussion (Maggio, 1981; Kurstak, 1986).

Critical Parameters

Sensitive ELISAs require antibodies of high affinity and high specificity. Since the sensitivity of an ELISA depends upon the affinity of the antibodies involved, antibodies with the highest affinities should be used when setting up ELISAs. Antibodies should be screened for unwanted cross-reactions. For instance, capture antibodies must not bind conjugate antibodies and vice versa. There are many commercial sources of reliable reagents. *Linscott's Directory of Immunological and Biological Reagents* is an excellent source book for locating reagents used in ELISAs. If reagents from one source are inadequate, try another.

When screening for expressed proteins in *E. coli*, it is important to utilize conjugates with antibodies that recognize nonnative and native molecules. Many foreign proteins expressed in *E. coli* will not assume their native conformation, and expression of such proteins will not be detected if antibody specific for the native form is used. It is also important to test enzyme-antibody conjugates for cross-reactivity or non-specific binding to host cell molecules. This potential problem can be eliminated by incorporating these as control antigens in the screening procedures used to select the original antibodies in the basic protocol.

When coating plates with antigen, the antigen preparation need not be pure, but should generally comprise >3% of the protein in the coating solution. In some situations, dilution of the antigen solution with BSA has greatly improved the sensitivity of the ELISA (Jitsukawa et al., 1989).

All steps after coating the microtiter plates should be carried out in solutions containing 0.05% Tween 20 and a carrier protein (0.25% BSA or gelatin).

When using ELISAs for quantitative determinations of antigen or antibody concentrations, four guidelines should be followed.

First, it is essential that all experimental conditions up to the final wash after incubation with conjugate—including incubation times, wash times, reagent concentrations, and temperature—be kept constant between experiments. This is especially important in assays using polyvalent antibodies and complex mixtures of antigens. The optimal concentrations of all reagents for each system should be determined in an initial criss-cross serial dilution experiment (see first support protocol). Second, because the efficiency of binding and other micro-environmental conditions can vary from plate to plate, a standard curve should be included on each plate. Third, all samples must be analyzed at least in duplicate. Fourth, the concentration of the reagent being quantitated must lie within the dynamic range of the standard curve.

Anticipated Results

Antibody-sandwich assays are generally the most sensitive ELISA configuration and can detect concentrations of protein antigens between 100 pg/ml and 1 ng/ml. ELISAs in which antigen is directly bound to plates are usually an order of magnitude less sensitive than sandwich techniques.

Either the direct or sandwich ELISA may be used to detect and quantitate a bacterially expressed antigen or a purified or partially purified antigen in the range of 1 ng/ml to 1 μ g/ml. A major disadvantage of the direct ELISA is that when an impure antigen preparation like a bacterial lysate is coated directly onto the surface of the microtiter well, the antigen must compete with all the other macro-molecules in the lysate for binding to the plastic and very little of the desired antigen may be bound. The sandwich ELISA bypasses this problem by relying on selective adsorption of an antigen to an antigen-specific antibody-coated surface.

Time Considerations

These assays are designed to take ~6 hr, but the incubation times may be abbreviated or expanded as needed. Since equilibrium binding between the soluble and solid phases frequently takes 5 to 10 hr, stronger specific signals can usually be obtained by longer incubations.

Fluorogenic ELISAs are generally 10 to 100 times faster than assays using chromogenic substrates.

Literature Cited

- Bartlett, W.C. and Noelle, R.J. 1987. A cell-surface ELISA to detect interleukin 4-induced class II MHC expression on murine B cells. *J. Immunol. Methods* 105:79-85.
- Beatty, J.D., Beatty, B.G., and Vlahos, W.G. 1987. Measurement of monoclonal affinity by non-competitive immunoassay. *J. Immunol. Methods* 100:173-179.
- Coligan, J.E., Kruisbeek, A.M., Margulies, D.H., Shevach, E.M., and Strober, W., eds. 1991. Current Protocols in Immunology, Chapter 5: Immunofluorescence and cell sorting. Greene Publishing and Wiley-Interscience, New York.
- Engvall, E. and Perlman, P. 1971. Enzyme-linked immunosorbent assay (ELISA): Quantitative assay of immunoglobulin G. *Immunochemistry* 8:871-879.
- Feit, C., Bartal, A.H., Tauber, G., Dymbort, G., and Hirshaut, Y. 1983. An enzyme-linked immunosorbent assay (ELISA) for the detection of monoclonal antibodies recognizing antigens expressed on viable cells. *J. Immunol. Methods* 58:301-308.
- Harada, N., Castle, B.E., Gorman, D.M., Itoh, N., Schreurs, J., Barrett, R.L., Howard, M., and Miyajima, A. 1990. Expression cloning of a cDNA encoding the murine interleukin 4 receptor based on ligand binding. *Proc. Natl. Acad. Sci. U.S.A.* 87:857-861.
- Jitsukawa, T., Nakajima, S., Sugawara, I., and Watanabe, H. 1989. Increased coating efficiency of antigens and preservation of original antigenic structure after coating in ELISA. *J. Immunol. Methods* 116:251-257.
- Kurstak, E. 1986. Enzyme Immunodiagnosis. Academic Press, San Diego.
- Linscott's Directory of Immunological and Biological Reagents, Santa Rosa, Calif.
- Macy, E., Kemeny, M., and Saxon, A. 1988. Enhanced ELISA: How to measure less than 10 picograms of a specific protein (immunoglobulin) in less than 8 hours. *FASEB J.* 2:3003-3009.
- Maggio, E.T. 1981. Enzyme Immunoassay. CRC Press, Boca Raton, Fla.
- Quinn, A., Harrison, R., Jehanli, A.M.T., Lunt, G.G., and Walsh, S.S. 1988. An ELISA for the detection of anti-acetylcholine receptor antibodies using biotinylated α -bungarotoxin. *J. Immunol. Methods* 107:197-203.
- Rubenstein, K.E., Schneider, R.S., and Ulmann, E.L. 1972. Homogeneous enzyme immunoassay: A new immunochemical technique. *Biochem. Biophys. Res. Commun.* 47:846.
- Schots, A., Van der Leede, B.J., De Jongh, E., and Egberts, E. 1988. A method for the determination of antibody affinity using a direct ELISA. *J. Immunol. Methods* 109:225-233.

Wang, K.C. and Leung, B.S. 1985. Fluorometric ELISA methods for rapid screening of anti-estrogen receptor antibody production in hybridoma cultures. *J. Immunol. Methods* 84:279.

Key Reference

Linscott's Directory. See above.

Highly recommended publication listing sources of immunological reagents, kits, and cells/organisms, including addresses and phone numbers of commercial suppliers (updated quarterly).

Contributed by Peter Hornbeck
University of Maryland
Baltimore, Maryland

Scott E. Winston (bacterial cell
lysate antigens)
Univax Biologics
Rockville, Maryland

Steven A. Fuller (bacterial cell
lysate antigens)
Allelix Inc.
Mississauga, Ontario

Isotype Determination of Antibodies

UNIT 11.3

Frequently it is necessary to know the amount and serological class of antibodies made by an immunized animal, produced by hybridomas, or present in the serum of patients with inflammatory or neoplastic conditions. The immunologist's approach to such a problem is to consider the antibody or immunoglobulin molecules themselves as antigens and to use anti-immunoglobulin antibodies as the specific and sensitive agents of detection. This unit describes two methods for measurement and classification of supernatant or serum immunoglobulins—an ELISA (first basic protocol) and a method employing electrophoresis and immunofixation (second basic protocol).

SANDWICH ELISA FOR ISOTYPE DETECTION

The speed and sensitivity of sandwich ELISAs make them the assays of choice for isotype determination. The antibody-sandwich ELISA to measure soluble antigen (UNIT 11.2; Fig. 11.2.3) is adapted in this unit for isotype detection. Microtiter wells are coated with isotype-specific capture antibodies followed by incubation with test solutions containing the antibodies to be isotyped. After test antibodies have been bound to the plate by reacting with capture antibodies, unbound test antibodies are washed out. Developing reagent is added to the wells, followed by another incubation. Unbound conjugate is washed out and substrate is added. Substrate hydrolysis indicates that the test solution contained the appropriate isotype. A lack of hydrolysis indicates that the test solution did not contain the appropriate isotype. UNIT 11.2 should be consulted for additional information.

**BASIC
PROTOCOL**

Materials

Capture anti-isotype antibodies: heavy-chain class-specific antibodies (anti- μ , - α , - γ , - δ , - ϵ), heavy-chain subclass-specific antibodies (anti- γ 1, - γ 2a, - γ 2b, - γ 3, - γ 4), or light-chain isotype-specific antibodies (anti- κ , - λ)

PBS (APPENDIX 2) containing 0.05% NaN_3 (PBSN)

Test antibodies: hybridoma supernatants, ascites fluid, or antisera

Blocking buffer (UNIT 11.2)

Standard isotype antibodies (i.e., purified antibodies of known isotypes)

Developing reagent: anti-Ig antibody (specific for all heavy-chain classes)—alkaline phosphatase conjugate (see UNIT 11.1; Southern Biotechnology or *Linscott's Directory*)

MUP or NPP substrate solution (UNIT 11.2)

Immulon 2 or 4 microtiter plates (or equivalent; UNIT 11.2)

Additional reagents and equipment for ELISA (UNIT 11.2)

1. Prepare the capture anti-isotype antibodies by diluting in PBSN to 2 $\mu\text{g}/\text{ml}$ final.

Capture antibodies can be monoclonal or polyclonal.

When screening hybridoma supernatants for IgG (anti- γ), it is advisable to initially use a capture-antibody preparation that recognizes all IgG isotypes. Subsequent analysis using subclass-specific capture antibodies can determine which IgG subclass is expressed. It is crucial that the capture antibodies do not bind to the antibodies in the developing reagent.

2. Coat each well of an Immulon microtiter plate with 50 μl capture antibody solution as in steps 3 to 5 of the basic protocol in UNIT 11.2.
3. Block wells as in steps 6 and 7 of the basic protocol in UNIT 11.2.
4. Prepare dilutions of the test antibodies in blocking buffer—typically, hybridoma supernatants are diluted 1:5 while sera are diluted 1:500.

Immunology

11.3.1

Contributed by Peter Hornbeck, Thomas A. Fleisher, and Nicholas M. Papadopoulos

Current Protocols in Molecular Biology (1992) 11.3.1-11.3.6

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 18

5. Prepare the standard isotype antibodies at ~500 ng/ml to serve as positive controls.
6. Transfer 50- μ l aliquots of test or standard isotype antibodies into the antibody-coated wells and incubate >2 hr at room temperature.

Specificity controls consisting of wells coated with capture antibodies and a series of standard antibodies of different isotypes should be included on each plate. Negative controls should be included on each plate and consist of wells coated with capture antibodies but receive no test or control antibodies.

7. Wash plate as in steps 9 to 11 of the basic protocol in *UNIT 11.2*.
8. Prepare the developing reagent so that the final conjugate solution contains ~200 ng anti-Ig/ml.

Crucial that antibodies in developing reagent do not bind to capture antibodies.

9. Add 50 μ l developing reagent to each well and incubate 2 hr at room temperature.

All test wells, positive-control wells, and negative-control wells should receive the developing reagent.

10. Wash as in steps 9 to 11 of the basic protocol in *UNIT 11.2*.

11. Add 75 μ l MUP or NPP substrate solution to each well and incubate at room temperature. Periodically check the plate for substrate hydrolysis.

Hydrolysis of NPP results in liberation of a yellow product that can be detected by visual inspection in ambient light. Hydrolysis of MUP results in the liberation of a fluorescent product that can be detected by visual inspection under a long-wavelength UV lamp in a darkened room. For more quantitative estimates of isotype concentrations, plates can be read with microtiter plate reader. Positive-control wells give strong signals by 1 hr. Weaker reactions can be detected by incubation for many hours or overnight.

BASIC PROTOCOL

DETECTING AND ISOTYPING ANTIBODIES BY ELECTROPHORESIS AND IMMUNOFIXATION

Qualitative identification and quantitative determination of serum (and other biological fluid) proteins provide useful information concerning pathologic conditions of the lymphoid system. High-resolution zone electrophoresis is a simple method to separate serum proteins based on their classification defined by five electrophoretic zones: albumin, α 1-globulin, α 2-globulin, β -globulin, and γ -globulin. Following electrophoresis, the proteins can be detected by staining with amido black, or immunofixation can be performed first to achieve more precise identification. In immunofixation procedure described below, immunoglobulins within separated protein bands are identified, and clonality is established, using antisera to the α , γ , μ , ϵ , and δ heavy chains and κ and λ light chains of immunoglobulins. Other serum proteins—including glycoproteins, transferrin, and C3—can be identified in electrophoresed sample by same technique.

The serum (or other fluid) is loaded on an agarose gel-covered microscope slide and electrophoresed. Proteins are detected after zone electrophoresis and can be identified by immunofixation.

Materials

- Serum (or other biological fluid)
- Normal saline
- 95% methanol/5% acetic acid
- 1% amido black (1 g in 100 ml of 2.5% acetic acid)
- 2.5% (v/v) acetic acid

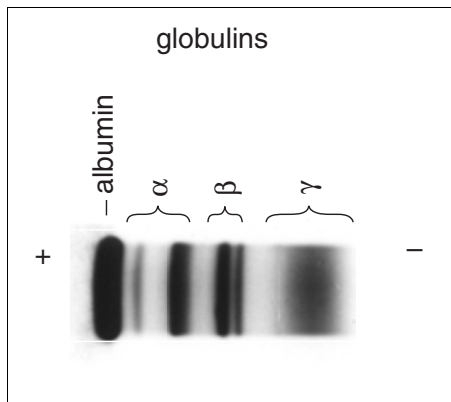


Figure 11.3.1 Normal serum protein electrophoretic pattern obtained by agarose gel zone electrophoresis demonstrating the five protein zones.

2 × 2-cm cellulose acetate strip
 Monospecific anti-Ig, heavy-chain-specific (α , γ , μ , ϵ , or δ) or
 light-chain-specific (κ or λ)
 0.85% (w/v) NaCl
 Agarose gel-covered microscope slides
 Plexiglas electrophoresis cell with two agarose bridges
 Electrophoresis power supply (e.g., Pharmacia EPS 500/400)

Electrophorese the sample

1. Cut a narrow slit with a razor blade in the middle of an agarose gel-covered microscope slide. For standard fixation and staining (steps 3 and 4), place 1.5 μ l undiluted serum in the slit. For immunofixation (steps 5 to 8), dilute serum in normal saline to generate 1 to 2 mg/ml of the specific protein being evaluated (e.g., IgG, IgA, or IgM) and place in the slit.
2. Place the slide in an electrophoresis cell with two agarose bridges having the same composition as the gel on the slide. Electrophorese 12 min at 140 V. Proceed to steps 3 and 4 for standard fixation and staining. Proceed to steps 5 to 8 for immunofixation.

Standard fixation and staining

3. Fix slides by submerging in 95% methanol/5% acetic acid 15 min. Air dry with filter paper placed on the gel.
4. Stain agarose gel by submerging in 1% amido black 10 min. Destain in 2.5% acetic acid until the background clears, then rinse in water. Dry 5 min at 60°C in a drying oven for visual inspection (Fig. 11.3.1).

Immunofixation

5. Transfer the *unstained* and *unfixed* gel (from step 2) to a petri dish containing damp filter paper. Overlay the γ -globulin zone of the gel with a cellulose acetate strip impregnated with a monospecific anti-Ig.

The γ -globulin zone is that nearest the cathode (Fig. 11.3.1).

The cellulose acetate strip is cut from any commercial membrane and dipped into the appropriate antibody solution (see critical parameters). The concentration of the antibody solution provided by supplier (usually 0.5 to 1 mg/ml) is suitable.

6. Allow the antibody to diffuse into the gel 15 min at room temperature.
7. Remove the slides and immerse in 0.85% NaCl for 2 hr to wash out unfixed proteins. Air dry with filter paper placed on the gel.
8. Stain as described in step 4 and visually inspect for immunoprecipitated protein bands (Fig. 11.3.2).

REAGENTS AND SOLUTIONS

Agarose gel-covered microscope slides

Add 0.5 g agarose (Seakem agarose HE, FMC Bioproducts) to 100 ml of 0.05 M barbital buffer (Sigma #B0500); store at room temperature and heat to 100°C. Stir until agarose has dissolved and allow to cool to 70°C. Pour 2.5 ml of the agarose solution on each microscope slide and allow to gel 2 to 3 min at room temperature. Store the prepared slides at 4°C in a humidified chamber. Set up a chamber by half-filling a petri dish with the above agarose solution. Allow to cool and place in a refrigerator. This humidified chamber can be used to store agarose slides up to 2 weeks.

COMMENTARY

Background Information

Sandwich ELISA

Antibodies are heteromeric molecules consisting of heavy and light chains, each of which contains a variable and a constant region. Heavy-chain constant regions include μ , α , $\gamma 1$, $\gamma 2a$, $\gamma 2b$, $\gamma 3$, $\gamma 4$, δ , or ϵ , depending upon the species; light-chain constant regions include κ and λ . Immunoglobulin constant regions, commonly referred to as isotypes, determine many of the biological and immunochemical properties of the antibody molecule including complement fixation, binding to Fc receptors, and binding to proteins A and G. Because the iso-

type may influence the method of purification (Andrew and Titus, 1991), it is routine to determine the isotypes of monoclonal antibodies or other specific antibody preparations as part of their initial characterization. The identification of antibody isotypes can easily be performed with an ELISA employing commercially available anti-isotype reagents. Alternatively, isotypes can be determined using electrophoresis/immunofixation (second basic protocol) or a double-immunodiffusion assay (Hornbeck, 1991).

See UNIT 11.2 for a full discussion of the ELISA technique.

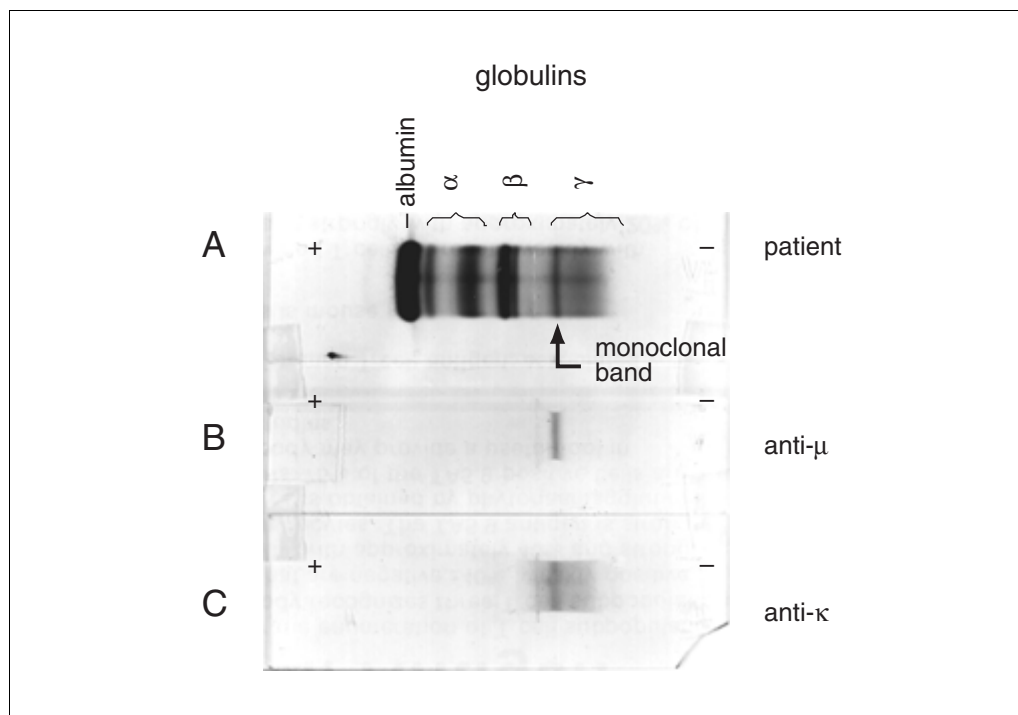


Figure 11.3.2 Agarose gel zone electrophoresis of patient serum demonstrating a monoclonal band in the γ -globulin zone (A; arrow). Immunofixation electrophoresis with anti- μ (B) and anti- κ (C) demonstrate that the band is a μ - κ monoclonal immunoglobulin. There is no reactivity with the antisera to the other heavy (α , γ , δ , ϵ) and light (λ) chains (data not shown).

Electrophoresis and immunofixation

Zone electrophoresis is based on the principle that charged particles migrate at different rates in an electric field based on the net charge of the particle. The application of this method to the evaluation of serum proteins was first described in a seminal paper by Tiselius (1937). In this paper it was shown that serum proteins were separable into defined zones (albumin, α -globulins, β -globulins, and γ -globulins). While protein separation by zone electrophoresis is excellent, protein quantitation using this method is poor. To overcome these shortcomings, various support media have been employed, including cellulose acetate, agar gel, and, more recently, agarose gel. Agarose gel provides improved stability and clarity, as well as greater sensitivity than other materials (Papadopoulos et al., 1982).

Immunofixation electrophoresis combines the high resolution of agarose gel zone electrophoresis and the unique specificity of an antigen-antibody reaction (Johnson, 1982). After electrophoretic separation, the antigen of interest is reacted with an overlaid, monospecific antibody to form an immunoprecipitate, which can be easily detected. The location of the immunoprecipitate depends on the electrophoretic migration of the specific protein antigen. The unreacted proteins and antibody reagents are washed out of the agarose gel and the precipitin band is stained for visualization. Using this method, polyclonality, oligoclonality, or monoclonality can be ascertained (Fig. 11.3.2). Commercial sources for complete immunofixation electrophoresis setups are available. Immunoelectrophoresis, a classic method in which diffusion of antibody into the gel is combined with electrophoresis, is an alternative method for evaluation of protein clonality. However, this approach is less sensitive and more difficult to interpret as compared with immunofixation (Johnson, 1986).

Critical Parameters and Troubleshooting

Sandwich ELISA

Sources of isotype-specific antibodies can be found in *Linscott's Directory of Immunological and Biological Reagents* (see key references, UNIT 11.2). Isotype-specific antibodies should have no detectable cross-reactivity against other isotypes and should not cross-react with other antibodies that might be used in the assay. Check isotype-specific reagents against standard isotype proteins to confirm

their specificity. Sources of purified isotypes from the species of interest, to be used as experimental standards, can also be found in *Linscott's Directory*. Alternatively, standard isotype proteins can be prepared from myeloma and hybridoma lines of known isotypes using standard techniques (UNIT 11.8). Many myelomas and hybridomas of defined isotypes and specificities are available from ATCC and other sources; see full listing in Knapp et al. (1991).

The concentration of the developing reagent should be adjusted so that the positive control gives a strong signal by 1 hr. Since the hydrolysis of MUP is at least 10-fold easier to detect than the hydrolysis of NPP, assays using MUP can be significantly faster than those using NPP. Test solutions should be scored as positive only when they give 3-fold higher signals than the negative controls.

While this assay is designed to qualitatively determine the presence or absence of a given isotype in the test solution, it can be easily modified to quantitate the concentration of isotype by including serial dilutions of standard isotype proteins (for details of quantitation using a standard curve, see UNIT 11.2 protocol for antibody-sandwich ELISA to measure soluble antigen).

Electrophoresis and immunofixation

Antigen and antibody interaction at or near the point of equivalence results in the formation of immune complexes that produce an insoluble precipitate. In the case of a monoclonal (homogeneous) protein, this precipitate is found in a very narrow band, while a polyclonal protein will generate a broad band. Immunoprecipitation is optimal at antigen/antibody equivalence. It is often useful to quantitate total immunoglobulin levels in the sample to allow for the dilution of the serum (see step 1 of basic protocol). Antigen excess will result in clear spots (lack of detectable immunoprecipitate) in the location where the band(s) are anticipated. If this occurs, dilute the sample and repeat immunofixation electrophoresis.

Failure to see a clearly discernible electrophoretic protein pattern after step 4 suggests a technical problem at some point during steps 1 through 4. Immunofixation should not be performed and the initial electrophoresis procedure should be repeated.

Absence of a detectable immunoprecipitate after step 8 (in the presence of a gamma globulin band after step 4) suggests antigen excess. Quantitate immunoglobulins, dilute serum accordingly, and repeat entire experiment.

Anticipated Results

Sandwich ELISAs are typically sensitive to 0.5 to 2.0 ng/ml of antibody isotype.

During electrophoresis and immunofixation, the specific bands observed in immunofixation electrophoresis are dependent on both the presence of particular proteins and the appropriate antisera being used in the detection process. In the case of normal immunoglobulins, all three major isotypes should be observed. Monoclonal or oligoclonal immunoglobulins produce single or multiple darker (clonal) bands (Fig. 11.3.2). The detection limit for a specific protein using this immunofixation technique is 5 to 10 $\mu\text{g/ml}$.

Time Considerations

The sandwich ELISA requires 6 to 8 hr. However, the times allotted for the various incubation steps can usually be reduced by half, so results can be obtained in 3 to 4 hr.

For electrophoresis and immunofixation, running and developing the gel takes 2 to 3 hr. Additional time (10 to 15 min) is needed to prepare and titrate the serum sample if immunoglobulin levels are increased.

Literature Cited

- Andrew, S.M. and Titus, J.A. 1991. Purification of immunoglobulin G. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 2.7.1-2.7.12. Greene Publishing and Wiley-Interscience, New York.
- Hornbeck, P. 1991. Double-immunodiffusion assay for detecting specific antibodies. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 2.3.1-2.3.4. Greene Publishing and Wiley-Interscience, New York.

Johnson, M.A. 1982. Immunofixation electrophoresis. *Clin. Chem.* 28:1797-1800.

Johnson, M.A. 1986. Immunoprecipitation in gels. *In* Manual of Clinical Laboratory Immunology. (N.R. Rose, H. Friedman, and J.L. Fahey, eds.) pp. 14-24. Am. Soc. Microbiol., Washington, D.C.

Knapp, W., Stockinger, H., Majdic, O., and Shevach, E.M. 1991. The CD system of leukocyte surface molecules. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. A.4.1-A.4.28. Greene Publishing and Wiley-Interscience, New York.

Papadopoulos, N.M., Elin, R.J., and Wilson, D.M. 1982. Incidence of γ banding in a healthy population by high-resolution immunofixation electrophoresis. *Clin. Chem.* 28:707-708.

Tiselius, A. 1937. A new apparatus for electrophoretic analysis of colloidal mixtures. *Trans. Faraday Soc.* 33:524-526.

Key References

Johnson, 1986. See above.

A concise discussion of the principles of immunoprecipitation with specific reference to immunofixation electrophoresis.

Maggio, E.T. 1981. Enzyme Immunoassay. CRC Press, Boca Raton, Fla.

A valuable reference describing parameters of ELISA technology.

Contributed by Peter Hornbeck
(sandwich ELISA)

University of Maryland
Baltimore, Maryland

Thomas A. Fleisher and Nicholas M.
Papadopoulos (electrophoresis
and immunofixation)

Warren Grant Magnuson Clinical Center
Bethesda, Maryland

PREPARATION OF MONOCLONAL ANTIBODIES

The preparation of monoclonal antibodies should be undertaken carefully, since the production of monoclonal antibodies is expensive and time-consuming. Figure 11.4.1 summarizes the experimental procedures that must be carried out to prepare monoclonal antibodies. The various procedures are presented in individual protocols and described in sufficient detail to allow an individual with no prior experience to carry out a cell fusion, to produce monoclonal antibodies in ascites fluid, and to purify antigen-specific monoclonal antibodies.

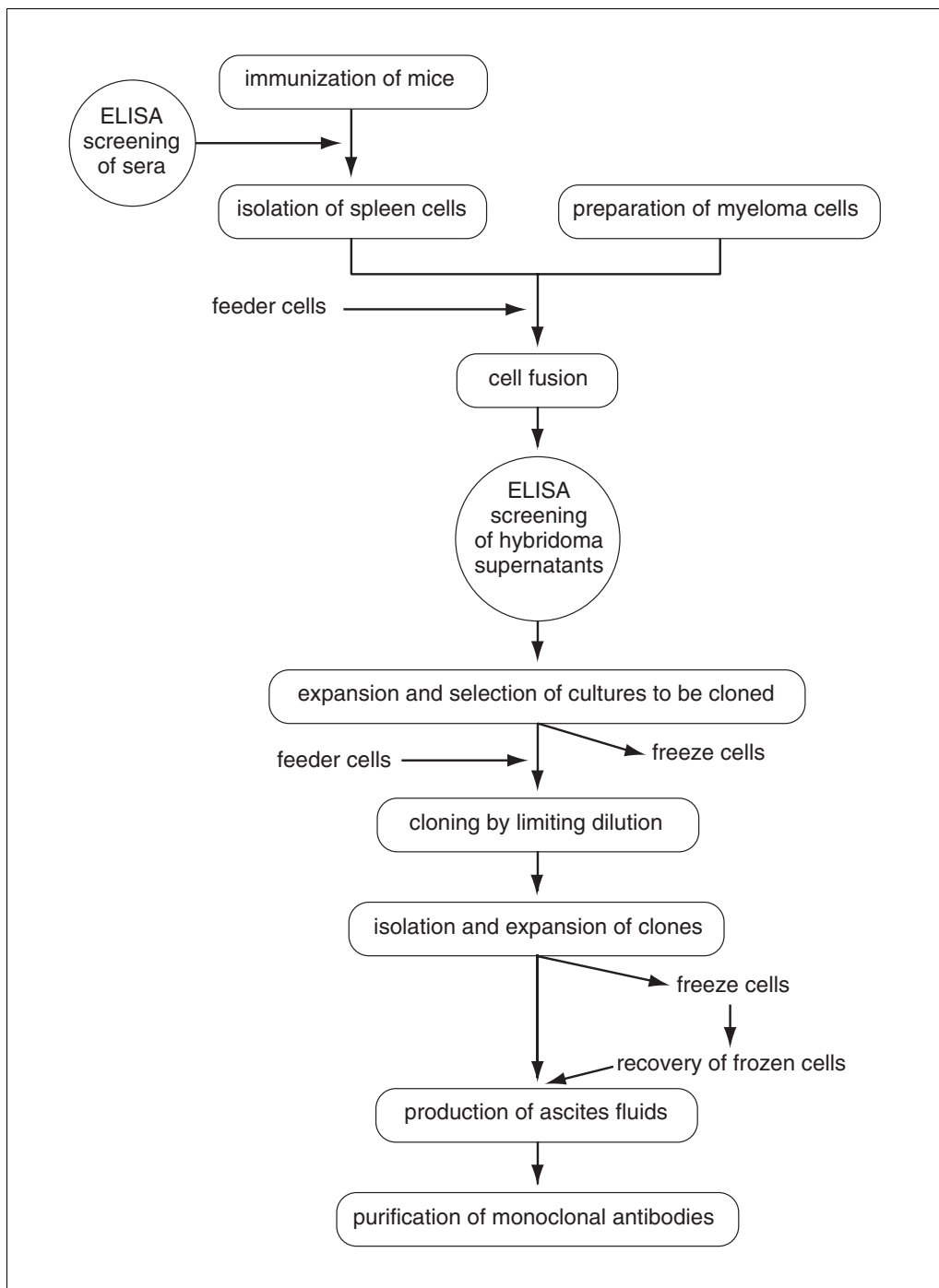


Figure 11.4.1 Flow chart for preparation of monoclonal antibodies.

Contributed by Steven A. Fuller, Miyoko Takahashi, and John G.R. Hurrell
Current Protocols in Molecular Biology (1992) 11.4.1-11.4.6
Copyright © 2000 by John Wiley & Sons, Inc.

Immunization of Mice

Antigen is prepared for injection either by emulsifying an antigen solution with Freund's adjuvant or by homogenizing a polyacrylamide gel slice containing the protein antigen. Mice are immunized at 2- to 3-week intervals. Test bleeds are collected 7 days after each booster immunization to monitor serum antibody levels. Mice are chosen for hybridoma fusions when a sufficient antibody titer is reached.

BASIC PROTOCOL

PRODUCTION OF IMMUNE SPLEEN CELLS: IMMUNIZATION WITH SOLUBLE ANTIGEN

Materials

Phosphate-buffered saline (PBS; *APPENDIX 2*)

Antigen

Complete Freund's adjuvant

Any strain mice, 6 to 8 weeks old

Incomplete Freund's adjuvant

22-G needles

3-ml syringes with locking hubs (Luer-Lok, Becton Dickinson)

Double-ended locking hub connector (Luer-Lok, Becton Dickinson)

Sterile sharp scissors

Sterile razor blades or scalpel blades

Wooden applicator sticks

200- μ l pipettor

Additional reagents and equipment for ELISA (*UNITS 11.2 & 11.3*) and western blotting (optional; *UNIT 10.8*)

1. Prepare an emulsion (200 to 400 μ l/mouse) of equal volumes PBS containing 25 to 100 μ g antigen and complete Freund's adjuvant. Using a 22-G needle, inject mice intraperitoneally. For each antigen, 3 to 5 mice are immunized.

Complete Freund's adjuvant contains mycobacteria—incomplete Freund's does not.

An emulsion is most readily prepared by linking two locking syringes, one loaded with antigen and the other loaded with adjuvant, using a double-ended locking connector (see Fig. 11.4.2). Press syringe barrels back and forth, transferring contents from one syringe to the other, for 5 to 10 min until a stable emulsion is produced. For best antibody production, inject antigen in as small an emulsion volume as practicable.

A stable emulsion is an oil-in-water emulsion which will not disperse when dropped into water. This is a useful check for the emulsification endpoint. Further, at the endpoint the emulsion will thicken noticeably.

Mice may be restrained for immunization in the following manner: Place mouse on gridded cage top. Lift mouse by the tail (generally, when mice are lifted by the tail they will grab the bars of the cage top with their front feet, thus stabilizing themselves for restraint). Immobilize the mouse's head by pinching together the skin at the base of the skull between thumb and forefinger. Turn hand over so that mouse is lying with its back against the palm. Wrap fourth finger around tail and stretch mouse over arched palm for intraperitoneal injection.

CAUTION: *Handle Freund's adjuvant carefully, since self-injection can cause a positive TB test and lead to a granulomatous reaction.*

2. Boost mice 3 weeks later by intraperitoneally injecting an emulsion (200 to 400 μ l) of equal volumes PBS containing 10 to 50 μ g antigen and incomplete Freund's adjuvant. The emulsion is prepared and injected as in step 1.

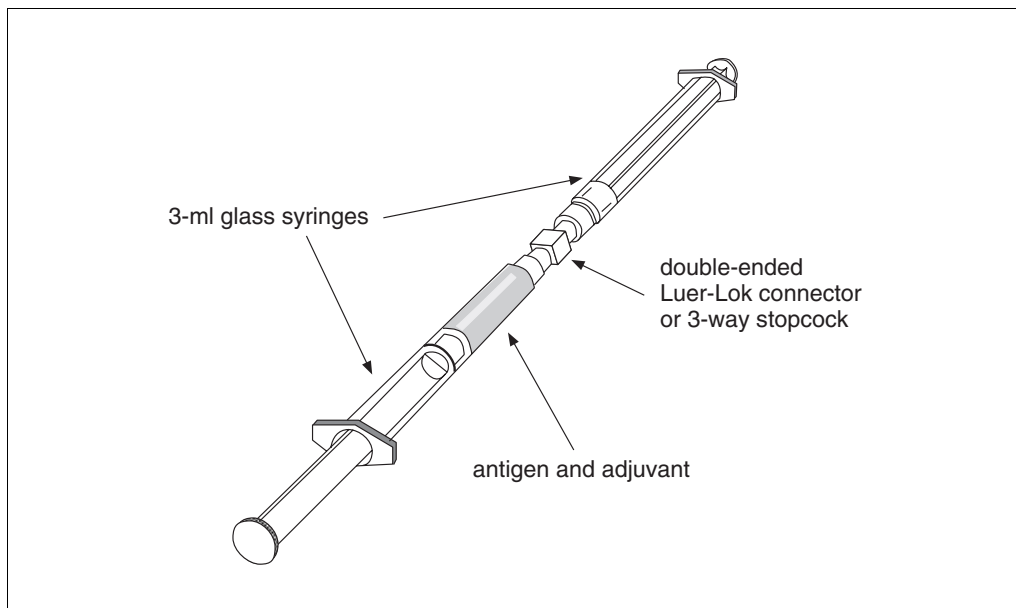


Figure 11.4.2 Double-syringe device for preparation of antigen-adjuvant emulsions.

- Bleed mice 7 days after second immunization by cutting off 0.5 cm of the tail with sterile sharp scissors or a razor blade. Collect 100 to 200 μl blood into a 1.5-ml microcentrifuge tube. After clot formation, rim the clot with a wooden applicator stick to dislodge the clot from the surface of the tube, but do not break up the clot. After clot retraction, transfer the serum into another microcentrifuge tube with a 200- μl pipettor. If test bleeds are collected more than three times, it will be necessary to cut the tail vein to obtain further samples rather than cutting off additional lengths of the tail itself. This is done by nicking one of the lateral tail veins with a razor blade.

The collection of blood may be facilitated by using a heat lamp to warm the mouse for 30 sec to 1 min prior to cutting of the tail. Additionally, if blood flow from the cut tail is slow, the tail may be "milked" from base to the cut tip with thumb and forefinger.

- Determine the antibody titer in the serum by ELISA (UNITS 11.2 & 11.3). If desired, further characterize the antibody specificity by western blotting (UNIT 10.8).

Antibody titer is operationally defined as that dilution of serum that results in 0.2 absorbance units above background in the ELISA procedure.

- If the antibody titer is considered too low ($\geq 1/1000$) for cell fusion, mice can be boosted every 2 weeks until an adequate response is achieved. Bleed the mice and test the serum with an ELISA.
- When the antibody titer is sufficient ($> 1/1000$), boost mice by injecting 10 to 50 μg antigen in PBS intraperitoneally (200 to 400 μl), or intravenously (50 to 100 μl) via the tail veins, 3 days before fusion but >2 weeks after previous immunization.

In general, the higher the serum antibody titer, the more antigen-specific antibody-producing hybridomas are obtained per fusion.

If an antibody against a nonimmunodominant epitope is desired, the cell fusion may be done at an earlier or later time, since the percentage of antibody-producing cells in the spleen directed at these less immunogenic regions of the antigen may vary with time in an unpredictable fashion.

- Perform cell fusion (UNIT 11.7) 3 days after the immunization (step 6).

**ALTERNATE
PROTOCOL**

**IMMUNIZATION WITH COMPLEX ANTIGENS (MEMBRANES, WHOLE
CELLS, AND MICROORGANISMS)**

1. Prime the mice and boost intraperitoneally with adjuvant (i.e., complete Freund's for priming and incomplete Freund's for booster immunizations) as described for soluble antigen (see basic protocol, steps 1 and 2) or suspend antigen in PBS and inject. Use 1 to 2×10^7 cells for mammalian species or 10^8 to 10^9 bacterial or yeast cells.
2. Bleed the mice and determine the antibody titer of the serum as described for soluble antigen (see basic protocol, steps 3 to 6).
3. Perform cell fusion (UNIT 11.7) 3 days after final immunization.

**ALTERNATE
PROTOCOL**

IMMUNIZATION WITH ANTIGEN ISOLATED BY ELECTROPHORESIS

In some instances the antigen under investigation can be purified most conveniently by gel electrophoresis (UNIT 10.2). Mice can be immunized with protein antigens still contained in a polyacrylamide gel slice, as described in this protocol.

Additional Materials

0.1 M KCl, cold

Tissue grinder

Additional reagents and equipment for denaturing (SDS) discontinuous gel electrophoresis (UNIT 10.2)

1. Apply a protein mixture containing 10 to 50 μg of the desired protein antigen to an appropriate denaturing (SDS) discontinuous gel electrophoresis system (e.g., the Laemmli gel system) and complete the electrophoresis as described in UNIT 10.2.
2. Soak gel 5 to 15 min in cold 0.1 M KCl. Protein bands will appear as white precipitates against a clear gel background.
3. Cut out the appropriate bands from the gel with a razor blade or scalpel blade.
4. Prepare gel suspension by homogenizing the gel slice in a minimum volume of PBS using a tissue grinder. Minimum volume is defined by adding successive 100- μl volumes of PBS until the homogenized gel is liquid.

Alternatively, the gel may be air dried for 1 to 2 hr, smashed with a glass rod, and suspended in a minimum volume of PBS.

5. Immunize each mouse with 200 to 400 μl gel suspension containing 10 to 50 μg antigen via an intraperitoneal injection.

Amount of antigen is estimated from prior observation of the proportion of desired protein antigen to other antigens in the sample as determined by the relative intensity of stained bands on the polyacrylamide gel (see UNIT 10.6 for staining procedures).

6. Boost mice after 3 weeks with 200 to 400 μl gel suspension containing 10 to 25 μg antigen.
7. Bleed the mice and determine the antibody titer of the serum as described for soluble antigen (see basic protocol).

Mice immunized repeatedly with polyacrylamide tend to form adhesions that can make aseptic removal of the spleen difficult.

8. Perform cell fusion (UNIT 11.7) 3 days after final immunization.

COMMENTARY

Background Information

The stimulation of an effective humoral immune response in mice is critical to the production of monoclonal antibodies directed at a particular antigen. The variety and quality of the monoclonal antibodies prepared is generally directly proportional to the serum antibody titer in the particular mouse used for cell fusion. Any means of antigen preparation, antigen delivery, or immunization schedule that increases antibody titer in the serum of the immunized mouse will potentiate the isolation of hybridomas secreting monoclonal antibodies of interest. We have described two methods of antigen preparation: (1) antigen emulsified in Freund's adjuvants (probably the most common technique used) and (2) antigen isolated in a polyacrylamide gel slice and homogenized. Other preparation methods (e.g., adsorption of antigen to supports such as aluminum hydroxide or aluminum phosphate, polystyrene beads, or nitrocellulose paper, and alternate sites of injection such as footpads) are discussed in the key references.

Critical Parameters

It is desirable to use antigen of the highest available purity for immunizations, particularly for primary immunizations. Contaminants may be more immunogenic than the antigen of interest and as such may result in a low specificity antibody. Mice given primary immunizations of highly pure antigen may be boosted with less pure material (containing as little as one-third specific antigen in a complex protein mixture).

Troubleshooting

Poor success in raising an adequate antibody titer to an antigen of interest can be attributed to several factors. Improperly prepared emulsion when using Freund's adjuvant (i.e., the aqueous and oil phases separate upon standing) is ineffective in stimulation of an immune response. Contaminants in an antigen preparation may be more immunogenic, necessitating a more homogeneous preparation of the desired antigen. Other parameters that can be varied in an effort to produce a higher antibody titer and increased specificity include presentation of antigen (Freund's adjuvant emulsion versus

polyacrylamide gel slice), site of immunization (intraperitoneal versus footpad or tail vein), antigen dose, and frequency of immunization. Alternate immunization protocols are presented in the key references below.

Anticipated Results

Isolation of high-quality monoclonal antibodies correlates with high-serum antibody titers. A serum ELISA titer of $1/1000$ is the minimum level before attempting a cell fusion. Titers for most antigens (particularly from animals injected with highly purified antigen) will range from $1/1000$ to $1/100,000$ after 3 to 4 immunizations. Occasional serum samples will titer at greater than 10^6 . The proportion of monoclonal antibodies of IgG class rather than IgM class generally increases proportionally to the duration of the immunization schedule, although this can vary dramatically among different antigens. [In general, IgG class antibodies are more suitable for immunoassays, western blotting (UNIT 10.8), immunoaffinity chromatography (UNIT 10.11), and immunoprecipitation (UNIT 10.16)].

Time Considerations

A primary immunization followed by two booster immunizations and test bleeds will occupy 6 weeks. For many antigens, however, an adequate antibody response in the mice is achieved only after several months and multiple immunizations.

Key References

- Hurrell, J.G.R., ed. 1982. *Monoclonal Hybridoma Antibodies: Techniques and Applications*. CRC Press, Boca Raton, Fla.
- Langone, J.J. and Van Vunakis, H., eds. 1986. Immunological techniques, Part I: Hybridoma technology and monoclonal antibodies. *Methods Enzymol.* 121:1-947.

Contributed by Steven A. Fuller and
Miyoko Takahashi
ADI Diagnostics
Rexdale, Ontario

John G.R. Hurrell
Boehringer Mannheim Diagnostics
Indianapolis, Indiana

Preparation of Myeloma Cells

BASIC PROTOCOL

Myeloma cells are cultured with 8-azaguanine to ensure their sensitivity to the HAT selection medium (see UNIT 11.6) used after cell fusion (UNIT 11.7). One week prior to cell fusion, myeloma cells are grown in medium without 8-azaguanine. Cell culture conditions are adjusted such that the Sp2/0 cells are in the log phase of growth and exhibit high viability at the time of collection for fusion (UNIT 11.7).

Materials

Sp2/0 murine myeloma cell line (ATCC #CRL 1581)
Complete culture medium
20 µg/ml 8-azaguanine

Tissue culture flasks, 25 cm² or 75 cm²
8% CO₂-in-air gas mixture
Humidified 37°C, 8% CO₂ incubator
Inverted microscope

1. Recover frozen cells from liquid N₂ storage, as described in UNIT 11.9.
2. Grow Sp2/0 cells overnight in complete medium in tissue culture flasks at 37°C in a CO₂ incubator in 8% CO₂-in-air atmosphere with 98% relative humidity.
3. Determine that the cells are growing by examining the cell cultures in the flasks with an inverted microscope and return culture flask to CO₂ incubator for continuation of cell growth.
4. To ensure that the Sp2/0 cells remain aminopterin sensitive for the selection process following fusion, supplement the complete culture medium with 8-azaguanine at 20 µg/ml during maintenance. One week prior to fusion, culture cells in medium *without* 8-azaguanine.

A seeding cell density of 2.5 to 5 × 10⁴ cells/ml works well with Sp2/0 cells.

Sp2/0 cells will grow to a maximum density of 6 to 9 × 10⁵ cells/ml, with a doubling time of 10 to 15 hr. When this density is reached, there is a rapid decline in cell viability. The Sp2/0 cultures are split every 2 to 3 days either by discarding an appropriate volume from the old flask and replacing with fresh medium or by transferring an appropriate volume of cells to a new flask and adding fresh medium. A 1-in-10 or 1-in-20 split is recommended.

5. A total of 1 × 10⁷ Sp2/0 cells (i.e., 1:10 ratio to immune spleen cells) is used for fusion. Cell viability at the time of collection should be greater than 95%. To ensure that cells are collected in log phase of growth, adjust the cell density to 2 × 10⁵ cells/ml the day before the fusion by adding fresh medium. Determine cell viability using the trypan blue exclusion method (see support protocol, below) on cells suspended in serum-free medium or PBS.

SUPPORT PROTOCOL

CELL VIABILITY TEST BY TRYPAN BLUE EXCLUSION

This procedure is used to determine the number of viable cells present in the cell culture. A non-viable cell will have a blue cytoplasm; a viable cell will have a clear cytoplasm.

Additional Materials

Phosphate-buffered saline (PBS; APPENDIX 2) or serum-free
complete culture medium
0.4% trypan blue solution

Binocular microscope
Hemocytometer

1. Centrifuge 1 ml cell suspension at $100 \times g$ for 5 min.
2. Resuspend the cell pellet in 1 ml PBS or serum-free complete culture medium.

Serum proteins stain with trypan blue and can produce misleading results. Determinations must be made in serum-free solution.

3. Mix 1 part of trypan blue solution and 1 part cell suspension ($\frac{1}{2}$ dilution).
4. Using a binocular microscope, count the unstained (viable) and stained (dead) cells separately in a hemacytometer. Each of the four corner squares (composed themselves of 16 smaller squares) have 1 mm sides and are 0.1 mm deep (0.1 mm^3). Count all cells within each of the four corner squares, including those that lie on the bottom and left-hand perimeters, but not those that lie on the top and right-hand perimeters. Count any clumps of cells as one cell. Calculate the mean number of cells per 0.1-mm^3 volume. Multiply by 10^4 to obtain the number of cells/ml (i.e., cells/ cm^3). Apply dilution factor for trypan blue ($2\times$) to obtain the number of cells per milliliter of culture.
5. Calculate the percentage of viable cells as follows:

$$\text{Viable cells (\%)} = \frac{\text{Number of viable cells}}{\text{Total number of cells (dead and viable)}} \times 100$$

REAGENTS AND SOLUTIONS

Complete culture medium

Dulbecco modified Eagle medium (DMEM), high-glucose formula (4.5 g glucose/liter; GIBCO/BRL #430-2100) supplemented to the indicated concentrations with the following additives:

- 2.8 g/liter sodium bicarbonate (33.3 mM)
- 4.8 g/liter HEPES (20 mM)
- 10% fetal calf serum (v/v)
- 10 ml/liter L-glutamine (2 mM)
- 10 ml/liter sodium pyruvate (1 mM)
- 10 ml/liter penicillin (50 IU/ml) and streptomycin (50 $\mu\text{g/ml}$)

The last four additives are available as $100\times$ solutions from GIBCO/BRL and other major suppliers of cell culture media. Penicillin and streptomycin are combined in one solution.

Samples of fetal calf serum lots should be tested for ability to support efficient cell growth and cloning before a large purchase because there is much variability between lots of a given supplier. The fetal calf serum must be mycoplasma free. If low volume usage of fetal calf serum precludes testing of serum lots, purchase of mycoplasma-free, virus-free, low endotoxin sera from suppliers such as GIBCO/BRL, Flow Laboratories, or Sigma will generally provide satisfactory results. Horse or bovine serum is not an adequate substitute!

COMMENTARY

Background Information

The Sp2/0 cell line was chosen as the fusion partner for immune spleen cells because of its good rate of growth, the efficiency with which hybridomas are obtained after fusion, and, most importantly, because it does not synthesize or secrete any immunoglobulin heavy or light chains itself. The Sp2/0 myeloma cell line was developed by Schulman et al. (1978). Other commonly used cell lines are P3X63-Ag8.653 (Kearney et al., 1979), which does not secrete immunoglobulins, and NS-1 (Kohler and Milstein, 1976), which produces only κ light chains.

Critical Parameters

Optimal growth of myeloma cells is density dependent. Cultures should be split at regular intervals to maintain >95% viability. Do not culture Sp2/0 cells longer than 1 month to avoid genetic drift and development of antibiotic-resistant contaminants. Maintain several aliquots of Sp2/0 cells in liquid nitrogen storage.

Anticipated Results

Proper care yields a healthy log phase

myeloma cell culture able to sustain good production of hybridomas upon fusion.

Time Considerations

Depending on culture conditions, 10^5 cells can be expanded to the 10^7 cells required for fusion in 4 to 6 days.

Literature Cited

- Kearney, J.F., Radbruch, A., Liesegang, B., and Rajewsky, K. 1979. A new mouse myeloma cell line that has lost immunoglobulin expression but permits the construction of antibody-secreting hybrid cell lines. *J. Immunol.* 123:1548-1550.
- Kohler, G. and Milstein, C. 1976. Fusion between immunoglobulin-secreting and nonsecreting myeloma cell lines. *Eur. J. Immunol.* 6:511-519.
- Schulman, M., Wilde, C.D., and Kohler, G. 1978. A better cell line for making hybridomas secreting specific antibodies. *Nature* 276:269-270.

Contributed by Steven A. Fuller, Miyoko Takahashi, and John G.R. Hurrell
Allelix Inc.
Mississauga, Ontario

Preparation of Mouse Feeder Cells for Fusion and Cloning

UNIT 11.6

**BASIC
PROTOCOL**

Chilled sucrose solution is injected intraperitoneally into mice. When withdrawn, the solution contains feeder cells (macrophages and other cells) that are placed in the wells of microtiter plates 1 day prior to seeding of hybridomas from cell fusion (UNIT 11.7) or cloning (UNIT 11.8) procedures.

Materials

- 0.34 M sucrose solution, sterile and chilled
- Mice (any strain)
- 70% ethanol
- HAT medium, chilled
- Sterile phosphate-buffered saline (PBS; APPENDIX 2)
- 10-ml syringe, sterile
- 18-G needle, sterile
- 50-ml conical centrifuge tube, sterile
- Dissecting board
- Forceps, sterile
- Scissors, sterile
- 96-well microtiter plates
- 8% CO₂-in-air gas mixture
- Humidified CO₂ incubator
- Additional reagents and equipment for estimating cell viability by trypan blue exclusion (support protocol, UNIT 11.5)

1. Just prior to sacrificing a mouse, fill a 10-ml syringe with 8 ml chilled sucrose solution and attach 18-G needle.

To avoid macrophages adhering to plastic surfaces, it is important to use chilled solutions to optimize cell harvest.

2. Chill the 50-ml conical centrifuge tube in ice.
3. Kill mouse by cervical dislocation.

This is accomplished by firmly holding a thick pencil or similar rod-shaped object to the neck of the mouse just behind the skull and quickly and firmly pulling the tail.

4. Immerse the mouse in a 100-ml beaker containing 70% ethanol.
5. Lay out mouse on dissecting board.
6. Snip skin at diaphragm level and pull skin back, exposing the lower part of the rib cage and abdomen.

With forceps pull skin from underlying tissue at the diaphragm level and snip with a scissors. With forceps or sterile gloved hands, pull skin back at both sides of the incision to expose the lower part of the rib cage and abdomen.

Care must be taken not to tear or cut the peritoneal membrane.

7. Insert the needle into the peritoneal cavity at the base of the sternum and rest the tip of the needle over the liver. Inject sucrose solution. Gently squeeze the abdomen two or three times.
8. Harvest the peritoneal feeder cells by withdrawing as much solution as possible into the syringe.

Care must be taken not to puncture the digestive organs, which may lead to fecal contamination of the feeder cells.

Enough peritoneal feeder cells can usually be isolated from one mouse to seed ~100 to 300 wells. However, some mice do not yield effective feeder cells. Depending on the total number of wells that must be seeded with mouse feeder cells, an appropriate number of mice must be killed. Peritoneal exudate feeder cells can be prepared up to 3 days prior to use.

9. Transfer the feeder cell-containing sucrose solution into the 50-ml centrifuge tube.
10. In a sterile fume hood, add 20 ml chilled HAT medium.
11. Centrifuge at $100 \times g$ for 5 min at room temperature.
12. Resuspend the pellet in 1 ml chilled HAT medium and perform cell viability test by trypan blue exclusion as described in the support protocol, *UNIT 11.5*.
13. Suspend the cell pellet in chilled HAT medium at 1×10^5 cells/ml.
14. Add 100 μ l cell suspension to each of the 60 inner wells of the 96-well plates. The peripheral 36 wells are filled with sterile PBS.

Plates having 24 wells may be used. If this is the case, add 1 ml cell suspension/well.
15. Incubate plates overnight at 37°C in a CO₂ incubator in 8% CO₂-in-air with 98% relative humidity.

REAGENTS AND SOLUTIONS

The following solutions are sterilized by filtration through a 0.22- μ m membrane. A suitable sterilization system is a disposable filter unit (e.g., Nalgene #120-0020). Glass-distilled water should be used for all preparations.

0.34 M sucrose solution

58.2 g sucrose

H₂O to 500 ml

Filter sterilize and store at 4°C in 100-ml aliquots

HAT (hypoxanthine/aminopterin/thymidine) medium

Complete culture medium (see reagents and solutions, *UNIT 11.5*) supplemented to the indicated concentrations with the following additives:

20% (v/v) fetal calf serum

0.1 mM nonessential amino acids

100 μ M hypoxanthine

0.4 μ M aminopterin

16 μ M thymidine

These additives may be purchased in concentrated and sterile solutions from the major suppliers of cell culture media and reagents. Concentrated solutions of hypoxanthine and thymidine (HT) and aminopterin may also be prepared in the laboratory (see following recipes).

100 \times HT solution

Weigh 340.3 mg hypoxanthine and 96.9 mg thymidine; add water to 250 ml. Heat to 70°C to dissolve. Filter sterilize and store in 20-ml aliquots at -20°C. Thaw at 70°C for 10 to 15 min.

1000 \times aminopterin solution

Weigh 17.6 mg aminopterin. Add 60 ml water and dissolve by adding 0.1 M NaOH dropwise. Titrate with HCl to pH ~8.5. Adjust volume to 100 ml and filter sterilize. Make 100 \times working solution by diluting stock in complete culture medium. Store in 5-ml aliquots at -20°C.

Aminopterin precipitates at low pH and is light sensitive.

COMMENTARY

Background Information

To maximize the yield of hybrids from the fusion and cloning procedures, feeder cells are required to be cocultured with the hybrids, while hybrid cell density is low. Mouse peritoneal cells, most of which are macrophages, have been found to be convenient and effective feeder cells which are a source of soluble growth factors for hybridoma cells.

Critical Parameters

Feeder cells such as peritoneal cells provide best support of hybridoma growth when used 1 to 3 days after harvest. Use of chilled solutions is necessary for optimum cell harvest, to prevent macrophages adhering to plastic surfaces.

Anticipated Results

From 1 to 3×10^6 peritoneal feeder cells are harvested from one mouse. The number of feeder cells will be enough to seed 100 to 300 wells.

Time Considerations

Peritoneal feeder cells from one mouse can be processed in 1 hr or less.

Contributed by Steven A. Fuller,
Miyoko Takahashi, and John G.R. Hurrell
Allelix Inc.
Mississauga, Ontario

Fusion of Myeloma Cells with Immune Spleen Cells

BASIC PROTOCOL

Freshly harvested spleen cells and myeloma cells are copelleted by centrifugation and fused by addition of polyethylene glycol solution to the pellet. Cells are centrifuged again and the PEG solution diluted by slow addition of medium. Fused cells are centrifuged, resuspended in selection medium, and aliquoted into 96-well microtiter plates. Hybridomas are grown to 10 to 50% confluence and then assayed for production of antigen-specific antibody.

Materials

- Any strain immunized mouse (*UNIT 11.3*)
- Sp2/0 murine myeloma cells in active log phase (Am. Type Culture Collection #CRL 1581; *UNIT 11.5*)
- Diethyl ether
- 70% ethanol
- Dulbecco modified Eagle medium (DMEM) with supplements
- Sterile polyethylene glycol (PEG) solution
- HAT medium (*UNIT 11.6*)
- HT medium
- Phosphate-buffered saline (PBS; *APPENDIX 2*)
- 15- and 50-ml centrifuge tubes
- Glass desiccator or metal can with lid
- Dissecting board
- 10.5-cm scissors (Irex #IR-105), sterile
- 10.5-cm forceps (Irex #IR-1393), sterile
- 60- and 100-mm petri dishes
- Stainless-steel strainer (Collector; GIBCO #1985-8500), sterile
- 3-cc glass syringes with 26-G needle
- 5-ml serological pipets
- 37° C water bath
- Stopwatch
- 8% CO₂-in-air gas mixture
- Humidified CO₂ incubator
- Polyvinyl or polystyrene 96-well microtiter plates
- Inverted microscope
- Additional reagents and equipment for estimating cell viability by trypan blue exclusion (*UNIT 11.5*) and for detection of antibodies (*UNIT 11.4*)

Preparation of myeloma and spleen cells

1. Just prior to sacrificing the mouse, transfer 1×10^7 Sp2/0 murine myeloma cells (prepared as described in *UNIT 11.5*) to a 50-ml centrifuge tube. Check the percentage of viable cells using the trypan blue exclusion method (support protocol, *UNIT 11.5*).
2. Sacrifice the mouse by anesthetizing with diethyl ether in a closed container (e.g., a glass desiccator or a metal can with a lid).

At this point, a blood sample may be collected from the mouse by severing the blood vessels of one forelimb. Collect the blood with a Pasteur pipet and place the blood in a microcentrifuge tube.
3. Immerse the mouse in a beaker containing 70% ethanol and lay out on a dissecting board.

4. Using sterile forceps, lift skin over the thorax area and snip with sterile scissors. Peel skin over both sides to expose left side of the rib cage.
5. Using another set of sterile forceps and scissors, remove the spleen from the left upper abdomen of the mouse. The spleen is a small dark red organ.
6. Place the spleen in a 60-mm petri dish containing 3 ml supplemented DMEM.
7. Take spleen in petri dish to sterile hood and carefully dissect away surface fat and other adhering tissue by using a sterile forceps and scissors.
8. Transfer spleen to sterilized, stainless-steel strainer in 100-mm petri dish with 10 ml DMEM.
9. Fill the 3-cc syringe with 2 ml supplemented DMEM. Using a 26-G needle, fill the spleen with DMEM by injecting at several sites.
10. With sterile scissors, cut the supplemented DMEM-filled spleen in 3 or 4 places.
11. Using circular movements, press the spleen against the screen of the stainless-steel strainer with the glass syringe plunger of the 3-cc syringe until only fibrous tissue remains on top of the strainer screen. The tissue that is forced through the strainer is collected in a sterile petri dish underneath.
12. Rinse the screen with 2 ml supplemented DMEM.
13. Transfer the suspension of spleen cells to a 15-ml centrifuge tube. Using a 5-ml serological pipet, disperse the clumps by drawing up and expelling several times.
14. Let suspension stand for 3 min at room temperature.
15. Transfer the top 95% of the cell suspension to a 15-ml centrifuge tube.
16. Perform a viable cell count using the trypan blue exclusion procedure (support protocol, *UNIT 11.5*); record the percentage of viable cells.
Ignore red blood cells that are substantially smaller than the nucleated cells.
17. Transfer 1×10^8 viable spleen cells into a 15-ml centrifuge tube.
18. Wash myeloma cells (from step 1) twice with supplemented DMEM followed by centrifugation at $200 \times g$ for 5 min and resuspend the cells in 5 ml supplemented DMEM.
19. Add spleen cells to myeloma cells in the 50-ml tube and fill the tube with DMEM.
20. Centrifuge the suspension at $200 \times g$ for 5 min at room temperature.
21. Resuspend the cell pellet in 50 ml supplemented DMEM and centrifuge as above.
22. Warm the cell pellet by placing the tube in a 37°C water bath in a beaker for 2 min.
23. Loosen the pellet by flicking the tip of the tube gently.

Cell fusion

24. Fuse spleen and Sp2/0 cells with sterile PEG solution:

Over the first 1 min—Add 1 ml PEG solution at 37°C. Mix gently.

Over the next 2 min—Spin at $100 \times g$ (2 min, total time).

Over the next 3 min—Add 4.5 ml supplemented DMEM.

Over the next 2 min—Add 5 ml supplemented DMEM.

Fill the tube with supplemented DMEM.

Timing is critical and should be monitored with a stopwatch.

25. Centrifuge at $100 \times g$ for 5 min at room temperature.
26. Aspirate supernatant.
27. Resuspend the cell pellet in 35 ml HAT medium.
Do not force the dispersion of small cell clumps.
28. Incubate cell suspension at 37°C in a CO_2 incubator in 8% CO_2 -in-air with 98% relative humidity for a minimum of 30 min.

Plating and culture of fused cells

29. Add 100 μl cell suspension to each of the 60 inner wells of six 96-well plates. Peripheral 36 wells are filled with sterile PBS.

Twenty-four (24) hr prior to use, the 60 inner wells are conditioned with 1×10^4 mouse peritoneal macrophages per well in 100 μl HAT medium.

30. Incubate plates at 37°C in CO_2 incubator in 8% CO_2 -in-air with 98% relative humidity. (This is day 1 of the culture).
31. On day 5 of the culture, add 100 μl HAT medium to each well.
32. On day 7 of the culture, remove 100 μl from each well and add 100 μl fresh HAT medium.
33. Repeat step 32 every other day until hybrid cell growth covers 10% to 50% of the surface area of the wells. This is monitored by examining the bottom of the wells with an inverted microscope. At this time, the wells should be screened for antibody (UNIT 11.4).
34. Grow the hybrids in HAT medium for 2 weeks after fusion.
35. After 2 weeks, change the medium to HT medium.
36. The hybrids are grown in the HT medium until the completion of two cloning procedures (as described in the protocol for cloning of hybridoma cells by limiting dilution, UNIT 11.8).

REAGENTS AND SOLUTIONS

Dulbecco modified Eagle medium (DMEM), high-glucose formula (GIBCO/BRL #430-2100), supplemented to the indicated concentrations with the following additives:

- 2.8 g/liter sodium bicarbonate (33.3 mM)
- 4.8 g/liter HEPES (20 mM)

HT medium

Prepare as described in UNIT 11.6 for HAT medium but without aminopterin solution.

PEG (polyethylene glycol) solution

Weigh 10 g PEG 4000 (Merck) into a 100-ml glass bottle. Autoclave PEG for 10 min at 121°C . Cool the molten PEG to 50°C . Mix with 10 ml prewarmed supplemented DMEM (50°C) that contains 5% (v/v) dimethylsulfoxide (DMSO). Aliquot the mixture into small glass bottles (3 ml/bottle). Store at 4°C in the dark.

PEG 4000, Merck's gas chromatography grade, appears to be the best PEG for cell fusions using Sp2/0 myeloma cells regardless of lot number. Remove PEG from the autoclave as soon as the pressure is down in order to avoid prolonged heating of PEG. Incubate PEG solution at 37°C for 24 hr prior to use to test its sterility.

Alternatively, if the temperature control of the autoclave is uncertain, melt the PEG at 65°C on the day of fusion. Mix the PEG with supplemented DMEM to make a 45% PEG solution and sterilize by filtration through a 0.22- μ m membrane filter.

COMMENTARY

Background Information

Murine spleen cells, some of which are involved in production of the desired antibodies, are fused with a murine myeloma cell line to form a stable antibody-producing hybridoma cell line. The myeloma cells (Sp2/0) are hypoxanthine–guanine phosphoribosyltransferase deficient (HGPRT⁻) and therefore are unable to use the purine salvage pathway when de novo purine synthesis is blocked by aminopterin, which is included in the HAT selection medium. See discussion of selectable markers in *UNIT 9.5*. Hypoxanthine and thymidine in the HAT selection medium allow HGPRT⁺ cells, the spleen cell–myeloma hybrids, to survive and grow. Unfused spleen cells eventually die.

Literature Review

Kohler and Milstein (1975) first demonstrated that somatic cell fusion could be used to generate a hybridoma cell line producing a monoclonal antibody of predetermined specificity. Cell fusion was initially accomplished by addition of Sendai virus. The fusion procedure used in this protocol is a modification of the method of Gefer et al. (1977), which uses polyethylene glycol (PEG) as the fusogen. A similar procedure is presented by Oi and Herzenberg (1980). Each protocol encompasses careful timing of the PEG addition to the cell pellet and its subsequent dilution after fusion.

Critical Parameters

The choice and use of PEG in the fusion protocol is the most critical factor. PEG can vary dramatically in efficiency between manufacturers and among lots of a particular manufacturer. The suggested source (Merck) has provided the most consistent lots of PEG. Care must be taken to autoclave the PEG at 121°C for only 10 min in order to minimize the production of toxic aldehydes. In the fusion procedure itself, it is important to adhere carefully to the time schedule established. Extended incubation of the cells with PEG results in decreased cell viability. Dilution of the PEG by medium must be done carefully to avoid lysis of the cells.

Troubleshooting

When poor fusion results (i.e., poor hybridoma growth) are obtained despite close adherence to the protocol, the likely causes are the PEG, the myeloma cells used as fusion partner, or the culture conditions. Most critical is the PEG (see critical parameters). The quality of the fusion partner (i.e., the myeloma culture; see *UNIT 11.5* for details) is also very important. The CO₂ incubator must provide a stable temperature, pH, and humidity for optimal hybridoma growth.

Anticipated Results

Hybridoma growth should be observed with the aid of an inverted microscope in nearly all wells after a few days in culture. For wells with viable cells, generally 10 to 50% will contain antigen-specific antibody. Results can vary widely, of course, and it has been observed that 0 to 99% of the supernatants in the wells will contain specific antibody.

Time Considerations

The fusion procedure requires 3 to 4 hr to complete. For best results, it should be accomplished without interruption. Wells with hybridoma growth can be assayed for specific antibody 7 to 12 days after fusion.

Literature Cited

- Gefer, M.L., Margulies, D.H., and Scharff, M.D. 1977. A simple method for polyethylene glycol-promoted hybridization of mouse myeloma cells. *Somat. Cell Genet.* 3:231-236.
- Kohler, G. and Milstein, C. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256:495-497.
- Oi, V.T. and Herzenberg, L.A. 1980. Immunoglobulin-producing hybrid cell lines. In *Selected Methods in Cellular Immunology* (B.B. Mishell and S.M. Shiigi, eds.) pp. 351-372. W.H. Freeman, San Francisco.

Contributed by Steven A. Fuller, Miyoko Takahashi, and John G.R. Hurrell
Allelix Inc.
Mississauga, Ontario

Cloning of Hybridoma Cell Lines by Limiting Dilution

BASIC PROTOCOL

Hybridomas to be cloned are diluted to 0.8 cells/well. This dilution provides 36% of wells with 1 cell/well by Poisson statistics. When cultures are 10 to 50% confluent, antibody is assayed by ELISA. Two or more cloning procedures are carried out until >90% of the wells containing single clones are positive for antibody production.

Materials

HAT medium (UNIT 11.6)
HT medium (UNIT 11.7)
Complete culture medium (UNIT 11.5)

Polyvinyl or polystyrene 96-well microtiter plates
6- and 24-well culture plates
8% CO₂-in-air gas mixture
Humidified CO₂ incubator
Cryotubes (Nunc #3-63401)
2-, 5-, and 10-ml serological pipets
Multichannel pipet and tips
Inverted microscope

Additional reagents and equipment for preparing mouse feeder cells for fusion and cloning (UNIT 11.6), for ELISA screening (UNIT 11.4), and for estimating cell viability by trypan blue exclusion (UNIT 11.5)

1. *The day before cloning*, isolate mouse feeder cells and prepare 96-well plates with feeder cells in appropriate medium, as described in UNIT 11.6.

Choice of medium depends on the current stage of the cloning process. The first cloning uses HAT medium, and the second cloning uses HT medium. Any other clonings use complete culture medium.

2. Transfer all of the cells from each well containing antigen-specific antibody in its hybridoma supernatant (as determined by an ELISA, UNIT 11.4) into a separate well of a 24-well plate that has been preincubated with 0.5 ml of an appropriate culture medium (see note in step 1) and culture overnight at 37°C, 8% CO₂-in-air, and 98% humidity in a CO₂ incubator.

If there are greater than 40 to 50 positive wells, it is difficult to manage conveniently the cloning procedure. The ELISA assays (UNIT 11.4) or western blotting (UNIT 10.8) can be performed using the 24-well culture supernatants in order to select the most promising samples to be cloned. The remaining positive samples can be transferred to cryotubes and frozen, as described in UNIT 11.9.

Cloning efficiency is always improved when hybrid cells are grown to log phase in 24-well plates. The efficiency is still better if the 24 wells are preconditioned with feeder cells prior to transferring the hybrids.

3. Perform cell viability count using trypan blue exclusion method (support protocol, UNIT 11.5) on the overnight cultures in 24-well plates.
4. Using a 6-well plate, make dilutions of cells from overnight cultures in HT or complete medium. In the first well, make a 1:100 dilution in a total of 3 ml; in the second well, dilute an aliquot of the first dilution to 80 cells/ml in 5 ml; in the third well, prepare 8 cells/ml in 10 ml (i.e., a 1:10 dilution from the second well).

Choice of medium is discussed in step 1. Using 6-well plates is far easier for preparing cell dilutions than using tubes. However, using 6-well plates is expensive.

5. With a multichannel pipet, fill the upper 50 wells of the inner 60 wells of the 96-well plate from step 1 with 100 µl of 8 cells/ml dilution (i.e., 0.8 cells/well)

and the 10 wells of the bottom row with 100 μ l of 80 cells/ml dilution (i.e., 8 cells/well).

6. Incubate at 37°C in a CO₂ incubator in 8% CO₂-in-air with 98% relative humidity (day 1).

As a precaution, the hybridoma cells remaining in the 24-well plate can be transferred to cryotubes and frozen, as described in UNIT 11.9.

7. On day 6, feed the culture with the addition of 100 μ l/well of fresh medium, using a multichannel pipet. Thereafter, if necessary, refeed the culture every other day by removing 100 μ l media from each well and adding 100 μ l fresh media.
8. When cell growth in the bottom of the wells is 10 to 50% confluent (as monitored using an inverted microscope), assay for specific antibody in the hybridoma supernatants using an ELISA (UNIT 11.4).
9. Transfer 2 to 3 selected positive subclones from each plate into a 24-well plate (as for step 2) and incubate overnight.

Expand the subclones and freeze one aliquot for each subclone in a cryotube. This is done as a precaution in case one fails to recover positive clones.

10. Repeat cloning procedure from the beginning until a stable and single hybridoma cell line is established.

Hybridomas that yield >90% antibody-positive cultures upon recloning are considered to be stable. Those that yield <90% positive cultures are subjected to further cloning. When clones become stable, reduce the 20% fetal calf serum level in the HT medium to 10%, gradually reduce HT level, and finally remove HT from the medium entirely. Some clones are more sensitive to this HT weaning process than others.

11. Once established as stable cell lines, hybridomas are maintained in complete culture medium in a similar manner to the myeloma cell line (UNIT 11.5). Cells are then propagated for liquid N₂ storage (UNIT 11.9) and for antibody production in ascites fluid (UNIT 11.10).

Recloning of established hybridoma lines may become necessary when they are cultured for longer than 30 days. Somatic mutation or chromosome loss may occur during an extended culturing, which could lead to a loss of antibody production.

COMMENTARY

Background Information

Cloning by limiting dilution is a method based on the Poisson distribution. Dilution of cells to an appropriate number per well can maximize the proportion of wells that contain one single clone. The cloning protocol presented here is a modification of the method of Galfre and Milstein (1981).

Critical Parameters

Accurate cell counting is necessary to obtain the proper dilution of 0.8 cells/well at step 5. Clone as soon as antibody-positive hybridomas are identified in order to avoid overgrowth by nonsecreting hybrids.

Anticipated Results

Depending on accuracy of cell counts and growth characteristics of each hybridoma, 20 to 50% of wells seeded at 0.8 cells/well can be expected to

exhibit growth. Normally, all wells at 8 cells/well will exhibit growth. Upon ELISA screening of the first cloning, some hybridomas will have lost production of specific antibody. The proportion of hybridomas that lose antibody production generally decreases at each successive cloning.

Time Considerations

Two to three clonings require 4 to 6 weeks.

Literature Cited

Galfre, G. and Milstein, C. 1981. Preparation of monoclonal antibodies: Strategies and procedures. *Meth. Enzymol.* 73(B):3-46.

Contributed by Steven A. Fuller, Miyoko Takahashi, and John G.R. Hurrell
Allelix Inc.
Mississauga, Ontario

FREEZING OF CELL LINES

Hybridoma cells are suspended in dimethyl sulfoxide/fetal calf serum and frozen rapidly in a dry ice–ethanol and glycerol bath followed by transfer to liquid nitrogen storage. Cells are recovered by thawing rapidly at 37°C, with immediate replacement of freezing medium by culture medium.

Materials

Freezing medium

95% ethanol

10% glycerol

Cryotubes (Nunc #3-63401)

15-ml plastic centrifuge tubes

Liquid N₂ freezer

Additional reagents and equipment for estimating cell viability by trypan blue exclusion (UNIT 11.5)

Ideally, hybridoma cells to be frozen should be in the log phase of growth (as described for the Sp2/0 myeloma cells, UNIT 11.5, step 5).

1. Perform a cell viability count using the trypan blue exclusion procedure (support protocol, UNIT 11.5).
2. Using a pencil, label cryotubes with identification and date.
3. Centrifuge cell suspension at 100 × g for 5 min at room temperature.
4. Aspirate the supernatant.
5. Resuspend cell pellet in freezing medium to give a cell density of 1 × 10⁷ viable cells/ml.
6. Aliquot 0.5 ml/cryotube (i.e., 5 × 10⁶ cells/tube).
7. Freeze in dry ice/ethanol and glycerol bath for 60 min (see Fig. 11.9.1).

For optimal cell viability, it is important that the time interval between steps 5 and 7 is not less than 15 min or greater than 30 min.

Alternatively, place the tubes in a styrofoam box, then place the box in a –70°C freezer overnight. The styrofoam box slows the rate of freezing of the medium in the tubes.

8. Transfer cryotubes to liquid nitrogen freezer.

RECOVERY OF FROZEN CELL LINES**Materials**

Complete culture medium (as described for Sp2/0 cells, UNIT 11.5)

37°C water bath

Alcohol swabs

15-ml centrifuge tubes

25-cm² tissue culture flask (Costar or Falcon)

1. Thaw cryotubes completely in 37°C water bath.

Thawing should be completed within 1 min. Immerse tubes only to level of contents.

2. Wipe top of cryotube with alcohol swab, transfer cells to 15-ml centrifuge tube and add 5 ml of complete culture medium warmed in a 37°C water bath.

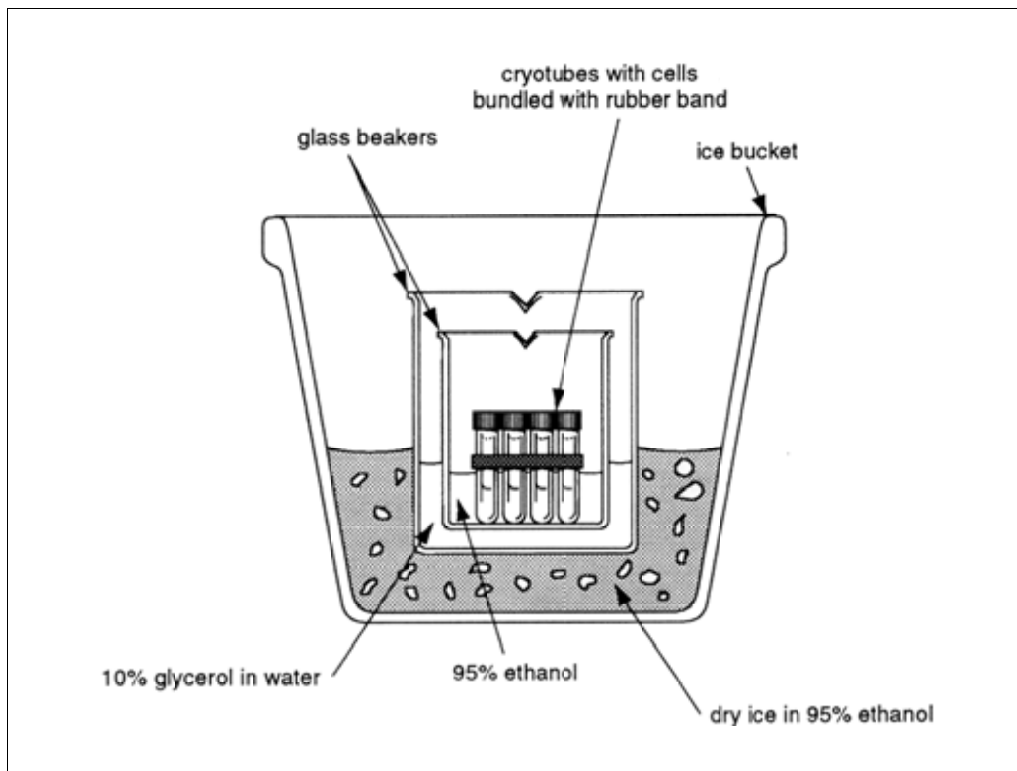


Figure 11.9.1 Apparatus for freezing of hybridoma cell lines in dry ice-ethanol and glycerol bath.

3. Centrifuge at $100 \times g$ for 5 min.
4. Aspirate supernatant.
5. Resuspend cell pellet in 5 ml warm complete culture medium (37°C).
6. Transfer suspended cells to a 25-cm^2 tissue culture flask.
7. Incubate overnight at 37°C in a CO_2 incubator in $8\% \text{CO}_2$ -in-air with 98% humidity. Keep flask upright.
8. Add 5 ml warm complete culture medium (37°C). Lay flask flat.
9. Propagate cells in complete culture media as described in *UNIT 11.10*, step 3.

REAGENTS AND SOLUTIONS

Freezing medium

10% (v/v) dimethyl sulfoxide, analytical grade

90% (v/v) fetal calf serum

Medium is prepared on day of use and chilled to 4°C before use.

COMMENTARY

Background Information

Liquid nitrogen storage is the method of choice for long-term safekeeping of hybridoma cell lines. Frozen aliquots of originally isolated hybridomas provide insurance against loss of antibody production and vigor during culture. There are many variations of cell freezing methods in use. However, for freezing hybridomas and lymphoid cells in general, this protocol is simple and has been successful.

Critical Parameters

Cell viability in the freezing and recovery protocols is drastically affected by the length of time spent in liquid freezing medium prior to and after frozen storage. Adhere carefully to time limits of the protocols.

Anticipated Results

Viability upon recovery of frozen aliquots ranges from 50% to 95%. Any recovery less than 50% viable is considered substandard.

Time Considerations

Freezing of cells can be accomplished in 1.5 to 2 hr. Recovery of hybridoma cell lines should require less than 10 min.

Contributed by Steven A. Fuller,
Miyoko Takahashi, and John G.R. Hurrell
Allelix Inc.
Mississauga, Ontario, Canada

Production of Monoclonal Antibody Supernatant and Ascites Fluid

UNIT 11.10

A major advantage of using monoclonal antibodies over polyclonal antisera is the potential availability of large quantities of the specific monoclonal antibody. In general, preparations containing the monoclonal antibody include a hybridoma supernatant, ascites fluid from a mouse inoculated with the hybridoma, and purified monoclonal antibody. Hybridoma supernatants are easy to produce, especially for large numbers of different monoclonal antibodies, but are relatively low in monoclonal antibody concentration. Ascites fluid contains a high concentration of the monoclonal antibody but the fluid is not a pure monoclonal antibody preparation. To obtain a purified preparation of the monoclonal antibody, affinity chromatography (UNITS 10.9 & 10.10) of culture supernatants or ascites fluid can be performed; however, this obviously requires more effort (see commentary).

Procedures detailing the production of monoclonal antibody supernatants (first basic protocol), including the production of larger amounts (liters; first alternate protocol) are presented here. The second alternate protocol for large-scale production of hybridomas or cells (e.g., for isolation of cellular proteins) involves a similar procedure. A method for producing and obtaining ascites fluid containing the monoclonal antibody is presented in the second basic protocol.

PRODUCTION OF A MONOCLONAL ANTIBODY SUPERNATANT

**BASIC
PROTOCOL**

There are a variety of methods for producing monoclonal antibody supernatants. In the easy version presented here, the hybridoma is grown and split 1:10. The cells are then overgrown until cell death occurs. The supernatant is harvested and the titer determined. If the titer is high, the hybridoma can be used for large-scale production in anticipation of purification or ascites production.

Materials

- Hybridoma of interest (UNIT 11.7)
- Complete DMEM-10 medium (APPENDIX 3F)
- 175-cm² tissue culture flasks
- 50-ml conical centrifuge tubes, sterile
- Beckman TH-4 rotor (or equivalent)

NOTE: All incubations are performed in a humidified 37°C, 5% CO₂ incubator.

1. Place hybridoma in a 175-cm² tissue culture flask in complete DMEM-10 medium. Incubate until vigorously growing and ready to split.

Most cell lines need to be split into new medium or new flasks when cell density reaches 1-2 × 10⁶ cells/ml. Tissue culture flasks can be inspected with an inverted microscope and cell viability and density determined. In addition, the culture can be monitored for contamination. Cells should never be allowed to become so crowded that cell death occurs because this crisis phase increases the likelihood of phenotypic change. With experience, most investigators will be able to determine whether or not to split the cells by their macroscopic appearance and the color of the medium—i.e., cells may need to be split when the medium in the flask becomes turbid and more yellow than medium in the DMEM bottle (macroscopic turbidity in the absence of a color change is often a sign of bacterial contamination). The density can then be confirmed by microscopic examination.

2. Split cells 1:10 in a new 175-cm² flask. Fill the flask with complete DMEM-10 to

Contributed by Wayne M. Yokoyama

Current Protocols in Molecular Biology (1992) 11.10.1-11.10.8

Copyright © 2000 by John Wiley & Sons, Inc.

Immunology

11.10.1

Supplement 18

100 ml total and incubate until cells are overgrown, medium becomes acidic (yellow), and cells die (~5 days).

Alternatively, add hybridoma cells at high density ($1-2 \times 10^6$ cells/ml) in fresh medium to a tissue culture flask. Incubate 2 to 3 days, at which time the cells will die and the supernatant can be collected (step 3).

3. Transfer flask contents to sterile 50-ml conical centrifuge tubes. Centrifuge 10 min in a TH-4 rotor at 2700 rpm ($1500 \times g$), room temperature. Collect the supernatant and discard the pellet.
4. Assay titer of MAb supernatant by appropriate method (see commentary).
5. Store the supernatant under sterile conditions; it is generally stable at 4°C for weeks to months, at -20°C for months to years, and indefinitely at -70°C. Minimize thawing and refreezing by storing several aliquots.

ALTERNATE PROTOCOL

LARGE-SCALE PRODUCTION OF MONOCLONAL ANTIBODY SUPERNATANT

The first step in monoclonal antibody purification by affinity chromatography (*UNIT 11.11*) is the production of large amounts of the culture supernatant. The procedure described below utilizes readily available equipment and supplies and a special tissue culture roller apparatus. Cells are first grown in smaller flasks, then gradually expanded into large-volume roller flasks. The supernatant is harvested and stored until needed.

Additional Materials

Complete DMEM-10 medium (*APPENDIX 3F*) with 5 to 10 mM HEPES, pH 7.2 to 7.4

70% ethanol

850-cm² roller flask

Roller apparatus in 37°C room or incubator

250-ml conical centrifuge tubes, sterile

Beckman JS-5.2 rotor (or equivalent)

1. Repeat step 1 of the first basic protocol and split 1:10 in complete DMEM-10/HEPES to 100 ml total.

Each 175-cm² flask will ultimately seed 2.35 to 2.5 liters of culture medium. Scale up the experiment as desired. If the supernatant will be used for MAb purification by affinity chromatography, the yield will be 1 to 10 mg MAb/liter. It is usually not necessary to adapt and grow cells in serum-free media (which frequently decreases yield); however, if the supernatant will be used for MAb purification by protein A-affinity chromatography, test the culture medium with FCS alone for contaminants (i.e., other proteins) that may co-purify with the MAb. Bovine newborn serum frequently contains significant amounts of Ig that will bind to protein A and should not be used as a medium supplement.

2. When cells are ready to split, transfer contents of the 175-cm² flask (100 ml) to an 850-cm² roller flask. Add an additional 150 ml complete DMEM-10/HEPES (250 ml total volume). Cap tightly, place on roller apparatus in 37°C room or incubator and grow 1 to 2 days.

Timing of the split will depend on the cell line (see step 1 of the first basic protocol).

3. Wipe cap and neck of roller flask with a sterile gauze sponge soaked in 70% ethanol.
4. Open roller flask and add 250 ml complete DMEM-10/HEPES (500 ml total volume). Cap tightly and incubate on roller apparatus 1 to 2 days at 37°C.

- Repeat wiping as in step 3.
- Open roller flask and add ~2 liters complete DMEM-10/HEPES until flask is almost full (~2.5 liters total volume depending on capacity of flask). Cap tightly and incubate on roller apparatus at 37°C until the medium turns yellow (~ 5 days).

Avoid foaming by first pouring in growth medium without FCS followed by FCS to 10% final.

Avoid prolonged rolling, as cell fragmentation will occur and the debris will be difficult to pellet with the large centrifuge tubes.

- Pour the culture into sterile 250-ml conical centrifuge tubes. Harvest the supernatant by centrifuging 20 min in JS-5.2 rotor at 1000 rpm (250 × g), room temperature. Collect the supernatant and discard the pellet. Freeze the supernatant in aliquots.

If the supernatant will not be used in a bioassay, add 10% sodium azide to 0.02% final. If the supernatant will undergo affinity chromatography or salt fractionation, sterile filtration through a 0.45- μ m filter is recommended to eliminate debris.

LARGE-SCALE PRODUCTION OF HYBRIDOMAS OR CELL LINES

The following procedure is used to produce large amounts of cells which can be used to isolate cellular components such as membrane proteins. Individual small flasks are grown, then each is used to inoculate a larger roller flask. The cells are gradually expanded by addition of fresh medium, and are harvested when the cells are near saturation densities.

- Follow steps 1 to 6 for large-scale production of MAb supernatants (first alternate protocol) but harvest when the density is appropriate or if the cell growth plateaus.

Estimate the amount of cells needed. This procedure will yield ~10⁶ cells/ml. More cells can be obtained with faster-growing cell lines or lines that can tolerate higher densities, but we usually grow 80 liters (33 flasks) and obtain 10¹⁰ to 10¹¹ cells.

Seed (introduce cells into) the number of 175-cm² flasks necessary to produce the amount of cells required (one 175-cm² flask for every 2.4 liters of medium). Each flask should be treated as an independent culture. Thus, if there is contamination at any time after the 175-cm² flasks are initially seeded, the contamination should spread vertically (i.e., stay in the flask and roller bottle seeded by the contaminated flask) and not horizontally (i.e., not involve any other flasks or roller bottles).

Estimate the time of harvesting by macroscopic inspection for medium color and turbidity and by taking daily cell counts and checking viability on several flasks. When the cell concentration reaches a plateau, harvesting is indicated. Do not allow cells to overgrow or cell viability will drop precipitously.

- Pour cells into sterile 250-ml conical centrifuge tubes, centrifuge 15 min in JS-5.2 rotor at 1000 rpm (250 × g), 4°C, and discard supernatant.

Each tube can be used for two spins. Harvesting 80 liters of cells requires at least three centrifuges to spin four to six tubes (1 to 1.5 liters) each, two or three people, and nearly one day. Conical centrifuge tubes are recommended because pellets in the flat-bottom bottles are harder to work with.

- Place cell pellets on ice. Pool 10 cell pellets into one tube by resuspending cells in 250 ml of 4°C PBS. Centrifuge at 250 × g, 4°C, and discard supernatant.
- Repeat and further consolidate tubes into one tube. After three washes, the cells are ready for further processing (i.e., cell lysis, radiolabeling, or other procedures).

Alternatively, lyse smaller cell pellets as they become ready after three washes.

ALTERNATE PROTOCOL

PRODUCTION OF ASCITES FLUID CONTAINING MONOCLONAL ANTIBODY

High-titer monoclonal antibody preparations can be obtained from the ascites fluid of mice inoculated intraperitoneally with monoclonal antibody-producing hybridoma cells. The fluid is collected several times after injection of the cells. It is heat-inactivated, titered, and stored.

Materials

Nude mice, 6 to 8 weeks old and specific-pathogen free, *or* syngeneic host if mouse-mouse hybridomas are injected

Pristane (2,6,10,14-tetramethylpentadecane; Aldrich)

Hybridoma of interest (UNIT 11.7)

Complete DMEM-10 medium (APPENDIX 3F) with 10 mM HEPES and 1 mM sodium pyruvate

PBS (APPENDIX 2) *or* HBSS, sterile and without FCS

20- or 22-G needle and 18-G needle

175-cm² tissue culture flask

Beckman TH-4 rotor (or equivalent)

50- and 15-ml polypropylene conical centrifuge tubes, sterile

56°C water bath

Additional reagents and equipment for ELISA (UNIT 11.2) and counting cells (UNIT 11.5)

1. Using a 20- or 22-G needle, inject mice intraperitoneally with 0.5 to 1 ml Pristane per mouse 1 week prior to inoculation with cells (Donovan and Brown, 1991).

Alternatively, the injections can be done at the same time, but the week interval is recommended to avoid leakage through two needle sites. Mice should be maintained in specific-pathogen free (spf) facility (see commentary).

2. Grow hybridoma cells in a 175-cm² tissue culture flask in complete DMEM-10/HEPES/pyruvate under conditions that promote log-phase growth.

Before injecting mice (step 7), test the supernatants for MAb activity by ELISA or appropriate assay, preferably before the cells are expanded.

To minimize the risk of introducing a pathogen into the rodent colony, screen cells for pathogens by antibody-production assay (Donovan and Brown, 1991).

3. Transfer the culture to 50-ml conical centrifuge tubes. Centrifuge 5 min in TH-4 rotor at 1500 rpm (500 × g), room temperature, and discard supernatant.

4. Wash cells by resuspending in 50 ml sterile PBS or HBSS without FCS, then centrifuging 5 min at 500 × g, room temperature, and discarding supernatant. Repeat twice and resuspend cells in 5 ml PBS or HBSS.

Avoid washing in FCS-containing medium because the mouse will produce antibodies to the FCS.

5. Count cells and determine viability by trypan blue exclusion.

Cells should be nearly 100% viable.

6. Adjust cell concentration to 2.5×10^6 cells/ml with PBS or HBSS without FCS.

7. Draw up cells in 10-ml sterile syringe. Using a 22-G needle, inject nude mouse intraperitoneally with 2 ml cells. Wait for ascites to form (1 to 2 weeks).

Typically, three mice are injected at one time. In most cases, at least one and frequently all of the mice will develop ascites.

8. Harvest ascites by grasping and immobilizing the mouse in one hand in such a way as to stretch the abdominal skin taut. With the other hand, insert an 18-G needle 1 to 2 cm into the abdominal cavity. Enter either the left or right lower quadrants to avoid the vital organs in the upper quadrants and the major vessels in the midline. Allow the ascites to drip into a sterile 15-ml polypropylene conical centrifuge.

If the mouse has a large amount of ascites and the fluid stops dripping from the 18-G needle, it may be necessary to reposition the needle tip by withdrawing it slowly and reinserting it in a different plane. If no ascites fluid accumulates, the mouse may be reinjected (see commentary).

Occasionally, the ascites is under such high pressure that a large amount squirts out as soon as the needle is inserted. For this reason, be sure that the hub of the needle is pointed into a tube before inserting the needle into the peritoneal cavity.

Rather than tapping the mouse as soon as the ascites is apparent, allow the fluid to build up (3 to 7 days) to obtain the highest yield. Frequently, 5 to 10 ml (sometimes >40 ml) of ascites can be collected from each mouse.

9. Centrifuge the ascites 10 min in TH-4 rotor at 2700 rpm (1500 × g), room temperature. Harvest supernatants and discard pellet. Store ascites fluid at 4°C until all collection is completed (<1 week).

If the fluid clots, “rim” the clot by passing with a wooden applicator stick around its edge (between clot and tube) before centrifugation. The clot may adhere to the applicator stick and thus may be discarded or it will remain in the tube and become part of the cell pellet after centrifugation.

10. Allow the mouse to reaccumulate ascites (2 to 3 days) before reharvesting as in step 8. Process the ascites as in step 9. Repeat this process until no further ascites accumulate, the fluid cannot be collected, or the mouse becomes ill. The mouse should be euthanized at this point (Donovan and Brown, 1991).

11. Pool ascites fluid collected on different days and heat-inactivate 45 min in a 56°C water bath. If a clot reforms, remove it by rimming and centrifuge as in step 9.

12. Assay the titer of MAb-containing ascites by the appropriate method (see commentary).

Saturating concentration (maximal activity) of the MAb should be apparent at 0.5% or higher dilutions. Lower titers usually are the result of unstable hybridomas that stop producing MAb or too many (>2) in vivo serial passes of the hybridomas.

13. Dilute >1:10 and filter sterilize through a 0.45-μm filter. Aliquot and freeze at -70°C, avoiding repeated freezing and thawing (Yokoyama, 1991a). Shelf life should be several years.

Add sodium azide to 0.02% final if the ascites will not be used for bioassay.

It is possible that ascites fluid will not form. If the mice die without any ascites forming, particularly within 2 weeks of inoculation, try fewer cells. If the mice do not form ascites after 2 weeks and they appear healthy, inject those mice—as well as naive, Pristane-primed mice—with more cells. If solid tumors form, tease cells into suspension and inject tumor cells into another Pristane-primed mouse. Even if a little ascites forms, the fluid can be transferred to another mouse (~0.5 ml/mouse), and large amounts of ascites should accumulate. Once the ascites is formed, the mouse-adapted cells can be frozen and used to reinoculate mice in the future.

COMMENTARY

Background Information

Production of MAb supernatant

There are three basic preparations that contain monoclonal antibodies: supernatant from a MAb-producing hybridoma, ascites from a mouse inoculated with the hybridoma, or purified MAb. In the first protocol, a MAb-containing supernatant is produced. The hybridoma continues to secrete MAb into the culture fluid until cell death occurs. Because the MAb is not metabolized, it accumulates in the culture supernatant. Hybridoma supernatants are advantageous because small amounts (<100 ml) can be easily obtained in a few days (<1 week), and from multiple hybridomas with relatively little effort. Such small quantities of the MAb can be used in preliminary experiments to determine if the MAb has the desired property before expending the effort required to produce ascites or purified MAb.

The first step in MAb purification, usually by affinity chromatography, is the production of large amounts of the MAb. For the noncommercial laboratory, several liters of culture supernatant are adequate for most MAb-purification protocols. The use of bioreactors or large spinner flasks would require expensive, somewhat fragile equipment. The alternate protocol described here involves the use of sealed roller bottles that can be rotated on a roller apparatus in any 37°C room, as the pH of the medium is maintained by the HEPES instead of a bicarbonate-based buffer system and a CO₂ atmosphere.

Production of ascites containing MAb

A convenient source of large concentrations of the desired MAb is the ascites fluid of mice inoculated with the appropriate hybridoma cells. These fluids often have a titer >100-fold more than culture supernatants and thus can be diluted significantly. Frequently, high-titer ascites preparations will have saturating MAb concentrations of 10⁻³ to 10⁻⁴ by flow cytometry analysis of cell-surface-antigen binding (Holmes and Fowlkes, 1991). This high titer minimizes the nonspecific functional effects (e.g., in proliferation assays) of equivalent concentrations of MAb in spent culture supernatants. However, ascites used at concentrations >0.5% may also have significant nonspecific effects. In addition to the appropriate

isotype control MAb ascites, ascites harvested from mice inoculated with the nonsecretory partner cell line (i.e., SP2/0) also provides a useful nonspecific control.

The general procedure involves the elicitation of nonspecific inflammation in the peritoneal cavity of an appropriate host mouse, usually with Pristane, and injection of the hybridoma cells. The tumor cells then grow as an ascites tumor and should continue to secrete the MAb. Eventually, the mouse develops a monoclonal gammopathy similar to the human disease, multiple myeloma, in that a monoclonal protein reaches high titer in the serum. Since the combination of the ascites tumor and Pristane results in an inflammatory exudate in the peritoneal cavity, the ascites should contain concentrations of the MAb similar to that in serum.

Critical Parameters

Monoclonal antibody production by hybridomas is an unstable phenotype. Hybridoma cells always should be grown under log-phase growth conditions. Prolonged *in vitro* culture and *in vivo* passage should be avoided. Thus, the most critical parameter is whether the hybridoma of interest is secreting a high titer of the MAb. This should be checked before a major effort is made to grow large amounts of supernatant or to produce ascites fluid. The MAb titer can be determined by serial dilution of the culture supernatant in the assay appropriate for that MAb, such as ELISA (UNIT 11.2) or flow cytometry (Holmes and Fowlkes, 1991; Yokoyama, 1991b). Titers of ≥1:10 should be saturating if spent culture supernatants are examined. If necessary, the hybridoma can be recloned by limiting dilution (UNIT 11.8) to find high-producing clones. If cells are known to produce MAb at high titers, aliquots frozen immediately after cloning (Yokoyama, 1991a) will retain this phenotype.

The most critical parameter in the large-scale production of cell lines and hybridomas is the adaptation of the cells to roller flasks. Most hybridomas and other nonadherent cells that grow in suspension can be easily adapted. If the cells (particularly adherent cells) cannot be adapted, other methods should be tried. For example, large-scale production of cells for use in isolating cellular components can be performed using multiple 175-cm² flasks instead of roller flasks. Unfortunately, the

relative surface area is small, and therefore the number of flasks required can become prohibitive. Adherent cell lines are less easily adaptable to growth in roller flasks. The surface area for growing adherent cells can be increased by the use of dextran beads (e.g., Cytodex beads from Pharmacia). These beads can increase the surface area of a culture flask severalfold.

If there is any suspicion that the cells may be mycoplasma contaminated, diagnosis and treatment are indicated (Fitch et al., 1991). Mycoplasma-contaminated lines will produce a much poorer yield of final cell numbers because they do not grow to as high a cell density as normal cells.

Supernatants frequently contain 1 to 10 $\mu\text{g/ml}$ of MAb, but the concentration is cell-line dependent. The supernatants could be concentrated by salt precipitation, but this is not generally recommended because large volumes of culture supernatants have to be concentrated to derive the amount of MAb in small amounts of ascites. Moreover, the FCS in the supernatant will also be concentrated. While the hybridoma could be adapted to culture in serum-free medium, this requires additional testing and yields may decrease. Affinity purification of the culture supernatants would take a similar amount of effort and produce purified MAb at high concentrations. Thus, instead of concentrating supernatants (if high concentrations of MAb are desired), it is recommended to produce purified MAb by first growing hybridomas at a larger scale (liters) or to produce ascites.

For ascites production, it is important to consider the appropriate host for the hybridoma since an injection of allogeneic or xenogeneic cells may result in rejection. For most mouse-mouse hybridomas, an F1 hybrid—between the BALB/c strain (origin of the commonly used SP2/0 fusion partner) and the strain from which the normal cells were obtained—could be used. For xenogeneic hybridomas, nude mice or low-dose irradiated normal mice are potential hosts. Outbred nude mice are somewhat more expensive than normal mice but do not require irradiation. It is not necessary to use the prohibitively expensive inbred nude mouse strains.

Because the level of normal immunoglobulin in mouse serum is in the same range as ascites fluid (mg/ml), ascites fluid can be only partially purified by salt fractionation or anti-mouse-Ig- or protein A–affinity chromatography (UNIT 2.7). However, it is a convenient source of raw material from which to affinity purify rat MAb with a mouse anti-rat κ MAb (e.g., MAR 18.5) column (Andrew and Titus, 1991).

Troubleshooting

It is possible that ascites fluid will not form. The reasons for this are unclear but are probably related to a property of the individual hybridoma. If the mice die without any ascites forming, particularly within 2 weeks of inoculation, try fewer cells. If the mice do not form detectable ascites after 2 weeks and they appear healthy, inject those mice—as well as naive, Pristane-primed mice—with more cells. If solid tumors form, tease cells into suspension and inject the tumor cells into another Pristane-primed mouse. Even if a little ascites forms, the fluid can be transferred to another mouse (~ 0.5 ml/mouse), and large amounts of ascites should accumulate. Once the ascites is formed, the mouse-adapted cells can be frozen and used to reinoculate mice in the future.

Anticipated Results

Most culture supernatants will have saturating MAb titers of $\geq 1:10$ when tested at 100 μl for 10^6 cells. If the spent culture supernatant is used for MAb purification by affinity chromatography (UNITS 10.9 & 10.10; Andrew and Titus, 1991), 1 to 10 mg of purified MAb/liter can be anticipated. If a much lower titer or yield is achieved, recloning of the hybridoma line may be indicated. Hamster-mouse hybridomas are particularly notorious for instability.

Most hybridomas can be grown as ascites tumors. The saturating concentration of the MAb in such fluids should be detected at dilutions of 1:500. If MAb titers are significantly lower, the hybridoma may be a poor producer. If ascites do not form, see troubleshooting above.

Most tumor cells that grow in suspension should be amenable to growth in roller flasks and densities of $>10^6$ cells/ml should be attained. Careful work should result in no contamination.

Time Considerations

For high-titer and large-scale production of MAb supernatants, a flask split 1:10 will be overgrown, with cell viability definitely decreasing by day 5 to 6, at which time the supernatants can be harvested. If several liters of supernatant are required, ~ 10 days are required to expand a 25-cm² flask (10 ml) to 2.4 liters.

For production of ascites fluid containing MAb, 4 to 6 weeks are necessary for growth of the cells for inoculation, ascites accumulation, tapping the fluid, centrifugation, and determination of the MAb titer.

Once the 175-cm² flasks are seeded for large-scale production of hybridomas and cell lines, <2 weeks are required to reach saturating cell densities in the roller bottles.

Literature Cited

Andrew, S.M. and Titus, J.A. 1991. Purification of immunoglobulin G. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 2.7.1-2.7.12. Greene Publishing and Wiley-Interscience, New York.

Donovan, J. and Brown, P. 1991. Care and handling of laboratory animals. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 1.1.1-1.1.4, 1.6.1-1.6.9, 1.8.1-1.8.4. Greene Publishing and Wiley-Interscience, New York.

Fitch, F.W., Gajewski, T.F., and Yokoyama, W.M. 1991. Diagnosis and treatment of mycoplasma-contaminated cell cultures. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. A.3.9-A.3.12. Greene Publishing and Wiley-Interscience, New York.

Holmes, K. and Fowlkes, B.J. 1991. Preparation of cells and reagents for flow cytometry. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 5.3.1-5.3.11. Greene Publishing and Wiley-Interscience, New York.

Yokoyama, W.M. 1991a. Cryopreservation of cells. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. A.3.15-A.3.17. Greene Publishing and Wiley-Interscience, New York.

Yokoyama, W.M. 1991b. Flow cytometry analysis using the Becton Dickinson FACScan. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 5.4.1-5.4.15. Greene Publishing and Wiley-Interscience, New York.

Contributed by Wayne M. Yokoyama
University of California School of Medicine
San Francisco, California

Purification of Monoclonal Antibodies

UNIT 11.11

PURIFICATION USING PROTEIN A–SEPHAROSE

BASIC
PROTOCOL

Ascites fluid is diluted in high pH buffer and loaded onto protein A–Sepharose. The column is washed and bound antibody is eluted with a succession of buffers of decreasing pH. The procedure for other affinity columns is similar, except that only one eluting buffer is used. Antibodies are concentrated by ultrafiltration.

Materials

Protein A–Sepharose CL-4B (Pharmacia)

Tris buffer, pH 8.6

Ascites fluid (UNIT 11.10)

Citrate buffer, pH 5.5

Acetate buffer, pH 4.3

Glycine-Cl buffer, pH 2.3

Neutralizing buffer, pH 7.7

2.5-cm (inner diameter) glass chromatography column

UV flowthrough detector or UV spectrophotometer and cuvette

Ultrafiltration cells and XM50 membranes (Amicon)

Additional reagents and equipment for ELISA screening (UNIT 11.4)

1. Fully swell protein A–Sepharose in Tris buffer (50-fold excess, vol/vol) in a beaker. Pour the protein A–Sepharose into a 2.5-cm glass chromatography column and equilibrate protein A–Sepharose with Tris buffer at room temperature (see UNIT 10.11A; Figure 10.11A.1).

The size of the protein A–Sepharose column and the volume of ascites used are subject to the requirements of the researcher. A column of 10-ml bed volume (~3 g dry weight protein A–Sepharose) is sufficient to purify antibody from 10 to 20 ml of ascites.

2. Dilute ascites fluid in 3 vol Tris buffer and apply to column at a flow rate of 1 to 5 ml/min.

Ascites can be applied directly to the top of the column by a Pasteur pipet or a flow adapter.

3. Wash column with Tris buffer until entire unbound proteins are eluted. This can be monitored by measuring the column effluent at A_{280} using a UV flowthrough detector or a UV spectrophotometer and cuvette.
4. Collect unbound protein and successively eluted protein peaks by fraction collector or by pooling an entire buffer elution volume.
5. Elute bound proteins successively with 2 to 3 column volumes of citrate, acetate, and glycine-Cl buffers directly into test tubes in a fraction collector containing neutralizing buffer in an amount equal to one-quarter of the collected volume.

Alternatively, bound material may be eluted in one step using glycine-Cl buffer only.

Collecting individual buffer fractions provides a better separation of monoclonal antibody from nonspecific antibody.

6. Eluted fractions or pools may be assayed for antigen-specific monoclonal antibody by ELISA (UNIT 11.4).

Immunology

11.11.1

Contributed by Steven A. Fuller, Miyoko Takahashi, and John G.R. Hurrell

Current Protocols in Molecular Biology (1997) 11.11.1-11.11.5

Copyright © 1997 by John Wiley & Sons, Inc.

Supplement 37

7. Concentrate appropriate eluates containing monoclonal antibodies to 1 to 5 mg/ml using an ultrafiltration cell with an XM50 ultrafiltration membrane. Store at -20°C .

Antibody concentration may be estimated by measuring absorbance at 280 nm. A mouse IgG solution at 1 mg/ml = 1.44 absorbance units.

ALTERNATE PROTOCOL

ALTERNATIVE BUFFER SYSTEM FOR PROTEIN A-SEPHAROSE

Pharmacia recently suggested another buffer system that results in higher binding capacity for certain monoclonal antibodies. The protocol was as described above; however, the buffers used were different. These alternative buffers are as listed below.

1. Glycine-OH buffer. Replaces Tris buffer in the basic protocol.
2. Elution buffers (0.04 M sodium citrate/0.02 M NaCl at pH 6.0, 5.0, 4.0, and 3.2). These buffers replace those in the basic protocol as follows:

Elution buffers, pH 6.0 and pH 5.0. Replace citrate buffer.

Elution buffer, pH 4.0. Replaces acetate buffer.

Elution buffer, pH 3.2. Replaces glycine·Cl buffer.

ALTERNATE PROTOCOL

PURIFICATION BY ANTIGEN-SEPHAROSE AND ANTI-MOUSE IMMUNOGLOBULIN-SEPHAROSE

This protocol describes the preparation of other affinity columns for purification of monoclonal antibodies if protein A-Sepharose will not bind the antibody (e.g., IgM or some IgG1 antibodies). It is useful for coupling protein antigens or anti-mouse immunoglobulin antibodies to commercially available CNBr-activated Sepharose 4B (see support protocol, *UNIT 10.16* for alternate procedure). Columns to which antigen or anti-mouse immunoglobulin antibody are attached may be used as immunoadsorbents to isolate specific antibody (antigen-Sepharose) or all mouse immunoglobulins (anti-mouse immunoglobulin-Sepharose) from ascites fluid.

Additional Materials

CNBr-activated Sepharose 4B (Pharmacia)
1 mM HCl
Coupling buffer
Protein antigen (previously purified; see Chapter 10) or anti-mouse immunoglobulin antibody (commercially available)
1 M ethanamine, pH 8.0
Phosphate-buffered saline (PBS; *APPENDIX 2*)
Washing buffer
60- or 150-ml sintered glass funnel, medium porosity
50-ml conical plastic centrifuge tube

1. Allow required amount of CNBr-activated Sepharose 4B to swell in 1 mM HCl.
One gram of dry gel swells to about 3 ml. It is convenient to use 5 to 10 mg of ligand per milliliter of gel. The remainder of the protocol assumes that 10 ml of gel is being prepared.
2. Transfer to sintered glass funnel and wash successively with (a) 1 mM HCl, 200 to 500 ml, and (b) coupling buffer, 200 to 500 ml (wash over a period of 0.5 hr).
3. Transfer swollen gel to a 50-ml conical plastic centrifuge tube. Add 50 to 100 mg of ligand at a concentration of 2 to 5 mg/ml.
4. Mix gel and ligand on rocker platform or end-over-end mixer overnight at 4°C .

5. Centrifuge at $250 \times g$ for 10 min.
6. Fill the tube with coupling buffer. Centrifuge as in step 5.
7. Incubate with 20 to 40 ml of 1 M ethanolamine, pH 8.0, on rocker platform for 4 to 5 hr at 4°C .
8. Centrifuge at $250 \times g$ for 10 min. Wash gel successively by filling the tube with PBS and washing buffer followed by centrifugation.
9. Transfer gel to a 2.5-cm glass column and equilibrate with 100 ml PBS.
10. Ascites fluid is diluted with 3 vol PBS and applied to the column at a flow rate of 1 to 5 ml/min at room temperature.
11. Wash unbound protein from the column with 30 ml PBS.
12. Elute bound protein using glycine-Cl buffer from the basic protocol (p. 11.11.1, step 5).
13. Concentrate eluate containing monoclonal antibodies to 1 to 5 mg/ml using ultrafiltration cell and XM50 ultrafiltration membrane. Store at -20°C .

Antibody concentration may be estimated by measuring absorbance at 280 nm. A mouse IgG solution at 1 mg/ml = 1.44 absorbance units.

REAGENTS AND SOLUTIONS

Acetate buffer, pH 4.3

To 900 ml H_2O , add:

6.80 g sodium acetate (trihydrate) (0.05 M)

8.77 g NaCl (0.15 M)

Titrate with acetic acid to pH 4.3

Add H_2O to 1 liter

Citrate buffer, pH 5.5

2.45 g citric acid (anhydrous) } (0.05 M citrate)

10.96 g trisodium citrate dihydrate }

8.77 g NaCl (0.15 M)

Add H_2O to 1 liter

Coupling buffer, pH 8.3 (0.125 M phosphate)

Solution A: 17.75 g Na_2HPO_4 (anhydrous)

Add H_2O to 1 liter

Solution B: 1.95 g $\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$

Add H_2O to 100 ml

Titrate solution A with solution B to pH 8.3

Elution buffers (0.04 M sodium citrate/0.02 M NaCl at pH 6.0, 5.0, 4.0, and 3.2)

To 900 ml H_2O , add:

11.76 g trisodium citrate dihydrate

1.17 g NaCl

Titrate with HCl to pH 6.0, 5.0, 4.0, or 3.2

Add H_2O to 1 liter

1 M ethanolamine, pH 8.0

61.1 ml ethanolamine

Titrate with HCl to pH 8.0

Add H_2O to 1 liter

Glycine-Cl buffer, pH 2.3

To 900 ml H₂O, add:
3.75 g glycine (0.05 M)
8.77 g NaCl (0.15 M)
Titrate with HCl to pH 2.3
Add H₂O to 1 liter

Glycine-OH buffer

To 700 ml H₂O, add:
108.9 g glycine (1.45 M)
175.3 g NaCl (3 M)
Titrate with NaOH to pH 8.9
Add H₂O to 1 liter

Neutralizing buffer, pH 7.7 (0.5 M phosphate)

Solution A: 70.98 g Na₂HPO₄ (anhydrous)
Add H₂O to 1 liter
Solution B: 7.80 g NaH₂PO₄·2H₂O
Add H₂O to 100 ml
Titrate solution A with solution B to pH 7.7

Tris buffer, pH 8.6

To 900 ml H₂O, add:
6.06 g Tris-Cl (0.05 M)
8.77 g NaCl (0.15 M)
0.2 g NaN₃ (0.02%)
Titrate with HCl to pH 8.6
Add H₂O to 1 liter

Washing buffer

To 900 ml H₂O, add:
13.6 g sodium acetate (trihydrate) (0.1 M)
29.2 g NaCl (0.5 M)
Titrate with acetic acid to pH 4.0
Add H₂O to 1 liter

COMMENTARY

Background Information

The uses of monoclonal antibodies as enzyme conjugates and as immunoaffinity reagents require their purification from the crude ascites fluid. Protein A–Sepharose chromatography and affinity chromatography are superior to ammonium sulfate precipitation, gel filtration, or ion-exchange chromatography for the preparation of contaminant-free antibody.

Literature Review

Staphylococcal protein A is known to bind immunoglobulins of several mammalian species. Ey et al. (1978) were able to use protein A–Sepharose to fractionate murine IgG antibodies using a stepwise pH elution. This protocol uses the buffer systems described by Oi and Herzenberg (1980), which allow efficient

purification of most IgGs and occasional IgM monoclonal antibodies at room temperature. The alternate protocol for preparation of specific affinity columns is a modification of the procedure suggested by Pharmacia.

Critical Parameters

Low-pH buffer elutions should be collected into vessels containing neutralizing buffer to avoid undue loss of antibody activity because of denaturation.

Anticipated Results

For most antibodies bound by the affinity columns, yield will be quantitative. Protein A–Sepharose will bind some IgM, most IgG₁, and nearly all IgG_{2a}, IgG_{2b}, and IgG₃ monoclonal antibodies. Yield of other affinity tech-

niques depends on the antibody-ligand interaction.

Time Considerations

The time involved in affinity chromatography depends largely on sample size and column dimensions. Research scale procedures (less than 20 ml of ascites on a 1 × 10 cm column) can be accomplished in 2 to 4 hr.

Literature Cited

Ey, P.L., Prowse, S.J., and Jenkin, C.R. 1978. Isolation of pure IgG1, IgG2a and IgG2b immunoglobulins from mouse serum using protein A-Sepharose. *Immunochemistry* 15:429-436.

Oi, V.T. and Herzenberg, L.A. 1980. Immunoglobulin-producing hybrid cell lines. *In* Selected Methods in Cellular Immunology (B.B. Mishell and S.M. Shiigi, eds.) pp. 351-372. W.H. Freeman, San Francisco.

Key References

Hurrell, J.G.R. ed. 1982. Monoclonal Hybridoma Antibodies: Techniques and Applications. CRC Press, Boca Raton, FL.

Langone, J.J. and Van Vunakis, H. eds., 1986. Immunological techniques, Part I: Hybridoma technology and monoclonal antibodies. *Meth. Enzymol.* 121: 1-947.

These volumes present a large body of material on basic hybridoma methodology and detail the use of monoclonal antibodies in the study of hormones, structural proteins, viruses, parasites, and mammalian cell types.

Galfre, G. and Milstein, C. 1981. Preparation of monoclonal antibodies: Strategies and procedures. *Meth. Enzymol.* 73(B):3-46.

Details the strategies and procedures for the preparation of monoclonal antibodies.

Contributed by Steven A. Fuller, Miyoko Takahashi, and John G.R. Hurrell
Allelix Inc.
Mississauga, Ontario

Production of Polyclonal Antisera

UNIT 11.12

Antibodies are serum immunoglobulins with binding specificity for particular antigens. Although antibodies can be identified in the serum of individuals or patients that have been exposed to particular pathogens, the usual methods for eliciting antibodies involve immunization with purified or partially purified antigen preparations. Antigens used are most commonly proteins or peptides, but carbohydrates, nucleic acids, small organic molecules (haptens) conjugated to appropriate protein carriers, cells, and cell and tissue extracts can also be employed.

The first consideration is usually whether polyclonal or monoclonal antibodies are needed. Polyclonal antibodies are particularly valuable for immunoprecipitation (*UNIT 10.16*) and immunoblotting (*UNIT 10.8*), whereas monoclonal antibodies can have exquisite specificity and can be derived for almost any purpose. Choice of the species of animal to be used for immunization is based in part on whether antibodies of great specificity are required—in which case genetically defined strains can be very helpful—or antibodies of wide cross-reactivity are needed.

The amount of antibody needed must also be evaluated. Clones of hybridomas (somatic cell hybrids of B cells from an immunized animal's spleen and myeloma tumors permissive for the production of monoclonal immunoglobulins) provide an essentially limitless supply of a constant reagent. Nevertheless, the initial investment in producing a monoclonal antibody is quite large, whereas relatively large amounts of a polyclonal antiserum can be obtained from a single rabbit or from several genetically identical rats or mice.

In this unit, the Basic and Alternate Protocols describe the production of polyclonal antisera specific for protein antigens in rabbits, rats, mice, and hamsters. The Support Protocol presents a method for preparing serum from blood. Polyclonal anti-peptide antisera can be produced by substituting carrier-conjugated peptides (*UNIT 11.16*) for the purified protein antigens.

STRATEGIC PLANNING

Production of good antisera depends in large part upon the quality, purity, and amount of available antigen as well as on the specificity and sensitivity of the assay. For protein antigens, if possible, the material should be biochemically homogeneous and, depending on the intended use, should be in either a native or denatured conformation. Be aware that minor contaminants are often (unfortunately) more antigenic than the immunogen of interest, and antisera resulting from immunization may have more activity against the contaminants than against the protein of interest. Antisera to be used for screening bacterial expression cDNA libraries or for immunoblots are best made against denatured protein, whereas those to be used for screening cDNAs expressed in eukaryotic transfection systems or for immunoprecipitation of native-cell-synthesized structures might best be made against native protein.

Although the advances offered by the development of monoclonal antibody techniques have revolutionized the specificity, uniformity, and quantity of antibodies, there remain many circumstances in which polyclonal antibodies are more desirable than monoclonal antibodies. Production of polyclonal antisera takes less time and effort than production of monoclonal antibodies, requires relatively simple and readily available equipment, and

produces reagents that can be used for immunoprecipitation, immunoblotting, and enzyme-linked immunosorbent assays (ELISAs).

Choice of animal for the production of antibodies depends upon the amount of antiserum desired, the evolutionary distance between the species from which the protein of interest has been derived and the species of the animal to be immunized, and prior experience with the immunogens. Rabbits are the usual animal of choice because they are genetically divergent from the human and mouse sources of the proteins most often studied. Rabbits provide as much as 25 ml of serum from each bleed without significant harmful effects. For smaller-scale experiments, or for those that rely on precisely defined antibody specificities, inbred mouse strains may be the system of choice. Because mice are smaller, the volume of antigen suspension used for immunization is significantly less and the amount of serum that can be obtained from a single bleed does not exceed 0.5 ml. Rats and hamsters may be used when larger amounts of serum are needed, or when the greater evolutionary distance is advantageous. With repeated bleeding, as much as 5 ml of serum can be obtained from these species. Additional discussion of the choice of species for the production of monoclonal antibodies is given in *UNIT 11.10*.

The choice of adjuvant for in vivo animal use has become problematic in recent years. Freund's adjuvant has been reliably and widely used for over fifty years (Freund et al., 1937). However, there is a degree of distress and discomfort to the animal associated with its use which requires that the responsible investigator explore alternatives. In most studies that have compared available alternatives to Freund's adjuvant, the Basic Protocol given here, which uses a CFA/IFA immunization regimen, gives far superior antibody titers to the commonly used commercially available adjuvants (Johnston et al., 1991; Deeb et al., 1992; Smith et al., 1992). However, in some instances TiterMax (Bennett et al., 1992) and Ribi Adjuvant Systems (Mallon et al., 1991) have performed as well as Freund's adjuvant. The use of these adjuvant systems as alternatives to Freund's adjuvant is described in the Alternate Protocol.

**BASIC
PROTOCOL**

**IMMUNIZATION TO PRODUCE POLYCLONAL ANTIBODIES USING
FREUND'S ADJUVANT**

In the presence of adjuvant, the protein antigen is injected intramuscularly, intradermally, or subcutaneously into an animal of the chosen species. Booster immunizations are started 4 to 8 weeks after the priming immunization and continued at 2- to 3-week intervals. Prior to the priming immunization and following the primary and each booster immunization, the animal is bled and serum prepared from whole blood (see Support Protocol).

Instructions on the different strategies for immunization (intramuscular, intradermal, or subcutaneous) can be found in Donovan and Brown (1995a) and for bleeding (from marginal vein or artery of ear for rabbit; various other sites for mouse, rat, or hamster) in Donovan and Brown (1995b). Factors important in preparing specific high-titer antisera, procedures for modifying protein antigens to enhance their immunogenicity, and choice of host animal are discussed in the Commentary. Protocols for immunization prior to production of monoclonal antibodies (*UNIT 11.11*) should be reviewed for these purposes.

Materials

- Rabbit, rat, mouse, or hamster of appropriate strain
- Complete Freund's adjuvant (CFA; Sigma)
- 1 to 2 mg/ml purified protein antigen in PBS (*APPENDIX 2*)
- Incomplete Freund's adjuvant (IFA; Sigma)
- 50-ml disposable polypropylene centrifuge tubes
- 3-ml glass syringes with 19-, 21-, and 22-G needles

**Production of
Polyclonal
Antisera**

11.12.2

Double-ended locking hub connector (Luer-Lok, Becton Dickinson) or plastic 3-way stopcock

CAUTION: CFA is an extremely potent inflammatory agent, particularly if introduced intradermally or into the eyes and may cause profound sloughing of skin or loss of sight. Self-injection can cause a positive TB test and lead to a granulomatous reaction. Use gloves and protective eyewear when handling CFA.

1. Bleed the animal prior to immunization and collect blood sample in a 50-ml centrifuge tube. Prepare serum from blood, then assay and store (see Support Protocol).

This preimmune bleed is critical as a control to ensure that the antibody activity detected in later bleeds is due to the immunization.

2. Shake CFA to disperse insoluble heat-killed *Mycobacterium tuberculosis* bacilli. Add 2 ml CFA to 2 ml of 0.25 to 0.5 mg/ml purified protein antigen in PBS at 4°C.

These volumes produce immunogen sufficient to immunize 4 rabbits or up to 80 mice. Do not use Tris-based buffers for generating the emulsion.

An effective and simple method for preparing purified protein antigen is by preparative SDS-PAGE (UNIT 10.2). If a standard-size 1.5-mm slab gel is used with a large-toothed comb, as much as 2 mg of a homogeneous protein can be loaded across the entire gel. Following electrophoresis, an edge can be sliced off with a razor blade, fixed and stained, and used to identify the region containing the protein band (UNIT 10.6). The gel slice containing the protein may then be directly added to several milliliters of PBS (APPENDIX 2) and emulsified as described below with an equal volume of CFA. The acrylamide serves as an additional component for the protein depot provided by the adjuvant.

3. Draw up the CFA/antigen mixture into a 3-ml glass syringe with a 19-G needle. Remove needle, expel as much air as possible, and attach syringe to the double-ended locking hub connector or the plastic 3-way stopcock (see Fig. 11.12.1). Attach an empty 3-ml glass syringe at the other end and force the mixture back and forth from one syringe to the other repeatedly. When the mixture is homogeneous and white, disconnect the connector or stopcock, attach a 21-G needle, and test whether the

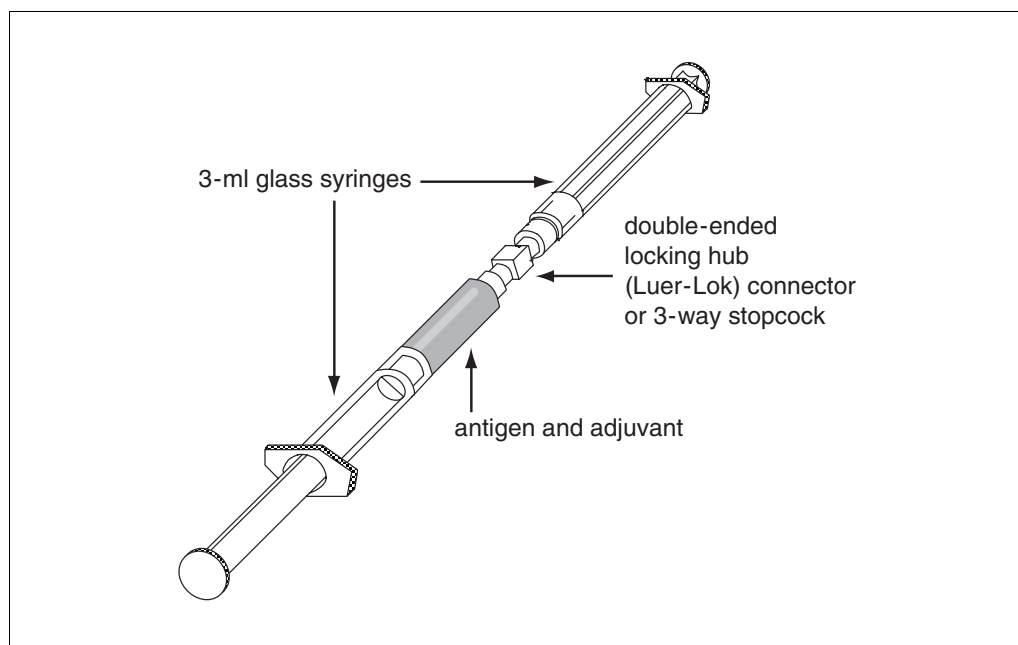


Figure 11.12.1 Double-syringe device for preparation of antigen-adjutant emulsions.

emulsion is stable by extruding a small drop onto the surface of 50 ml cold water in a 100-ml beaker. A good oil-in-water emulsion should hold together as a droplet on the surface of the water. If the drop disperses, mix the antigen using the hub-connected syringes until it forms an emulsion.

Heat will be generated by this procedure. Chill on a bed of ice from time to time to keep the mixture as close to 4°C as possible.

4. Transfer all of the adjuvant-antigen emulsion to one syringe and remove the connector or stopcock. Attach a 22-G needle to the syringe and remove air bubbles.
5. Restrain the animal and inject the adjuvant/antigen emulsion into multiple intramuscular (i.m.), intradermal (i.d.), or subcutaneous (s.c.) sites.

Discard the unused immunogen. For extremely valuable antigens, the emulsion may be stored at 4°C for several weeks and re-emulsified before use. However, denaturation of protein antigens may take place under these conditions. For immunization of small rodents (e.g., mice), it is often better to carry out injections intraperitoneally (i.p.).

6. Bleed the animal 10 to 14 days following the priming immunization and collect blood sample. Prepare serum from blood (see Support Protocol).
7. Prepare antigen for booster immunizations, following steps 2 to 4. When CFA is the primary adjuvant, use IFA as the adjuvant for all subsequent immunizations.
8. Administer the first booster immunization 4 to 8 weeks after the priming immunization, bleed the animal 7 to 14 days later, and collect blood sample. Prepare serum from blood (see Support Protocol).

Some investigators will administer the first booster immunization as early as 2 weeks after the primary immunization.

9. Administer further booster immunizations at 2- to 3-week intervals. Bleed animal 10 to 14 days after each boost and collect blood sample. Prepare serum from blood (see Support Protocol).

Repeated intradermal immunization should be avoided as it can cause skin ulceration. Following primary intradermal or subcutaneous immunization, it is preferable to boost with intramuscular injections for the rabbit. Some investigators prefer primary intramuscular injections with boosters at other sites.

ALTERNATE PROTOCOL

IMMUNIZATION TO PRODUCE POLYCLONAL ANTISERUM USING OTHER ADJUVANTS

For highly immunogenic antigens the use of Freund's adjuvant can certainly be avoided. For other immunogens it may be necessary to test a number of adjuvant systems. The use of two commercially available adjuvants is described in this protocol.

Additional Materials (also see Basic Protocol)

TiterMax #R-1 (CytRx; store <24 months at 4°C) *or* Ribi Adjuvant System (RAS; Ribi ImmunoChem; store at 2 to 8°C and do not freeze)
1-ml plastic syringes

- 1a. *Using TiterMax:* Emulsify aqueous antigen with TiterMax adjuvant (see Basic Protocol; follow steps 1 through 5, except use 0.5 ml antigen and 0.5-ml vial TiterMax in step 2 and plastic syringe in step 3).

TiterMax contains microparticulate silica coated with block copolymer CRL-8941, sorbitan mono-oleate, and squalene.

Although glass syringes are recommended for Freund's adjuvant emulsions, all-plastic syringes should be used with TiterMax.

Each reconstituted 0.5-ml vial will immunize 20 mice or 10 rabbits. Unused antigen/adjuvant emulsion can be stored at 4°C, -20°C, or -70°C for as long as the antigen is stable. It may be necessary to re-emulsify before using.

- 1b. *Using Ribi Adjuvant System:* Warm the vial of RAS for 5 to 10 min at 40° to 45°C. Add 2 ml antigen in PBS directly through the rubber stopper using a syringe with a 21-G needle. Vigorously vortex the vial 3 min at room temperature with the cap seal in place.

The final volume of adjuvant/antigen is 2 ml containing 2% squalene oil.

Each vial of RAS contains 0.5 mg each of monophosphoryl Lipid A (MPL), synthetic trehalose dicorynomycolate (TDM), and cell wall skeleton (CWS) in 44 µl squalene and Tween 80.

Each reconstituted vial will immunize 10 mice or 2 rabbits. Unused adjuvant/antigen emulsion can be stored several months at 4°C. However, if the entire vial will not be used initially, it is better to reconstitute to 1.0 ml with saline alone, store at 4°C, and mix aliquots 1:1 with antigen in saline as needed.

2. Transfer the antigen emulsion to a 1-ml syringe, attach a 22-G needle to the syringe, and remove air bubbles.
3. Restrain the animal and inject the adjuvant/antigen emulsion.

Rabbits should receive 40 µl TiterMax/antigen emulsion i.m. in each thigh. High antibody titers have been obtained with 30 to 50 µg of antigen per rabbit. Rabbits should be immunized with 1.0 ml RAS containing 50 to 250 µg of antigen in multiple sites: 0.05 ml i.d. at six sites, 0.3 ml i.m. in each thigh and 0.1 ml s.c. in the neck.

4. Bleed the animal and prepare antigen for booster immunization (see Basic Protocol, steps 6 and 7).

Serum antibody responses have been reported to be slower for both RAS and TiterMax than for Freund's adjuvant (Smith et al., 1992).

5. Administer booster immunizations at 4, 8, and 12 weeks. Bleed the animal 10 to 14 days after each booster immunization. Prepare serum from blood (see Support Protocol) and cease immunization when high antigen-specific titers have been achieved.

Boosting with TiterMax may not be necessary for all antigens. If a second immunization is necessary, use soluble antigen in place of antigen/adjuvant at 4 weeks. If titers are still low after 10 to 14 days, a booster dose of antigen/TiterMax adjuvant can be given immediately. Increasing the amount of antigen may also help.

Ribi ImmunoChem strongly recommends that booster injections of RAS adjuvant/antigen be repeated no more frequently than every four weeks.

PREPARATION OF SERUM FROM BLOOD

Each blood sample is allowed to stand 4 hr at room temperature and overnight at 4°C until a clot forms. After removal of the clot and debris, the serum is assayed and stored at -20°C.

Additional Materials (also see Basic Protocol)

Blood samples (see Basic Protocol)
Sorvall H-1000B rotor or equivalent

Additional reagents and equipment for immunoblotting (UNIT 10.8),
immunoprecipitation (UNIT 10.16), and ELISA (UNIT 11.2)

1. Allow blood to stand in the 50-ml centrifuge tube 4 hr at room temperature to allow clot to form, then place overnight at 4°C to allow clot to retract.

SUPPORT PROTOCOL

Immunology

11.12.5

2. Gently loosen the clot from the sides of the tube with a wooden applicator stick (do not break up the clot), then remove the clot from tube with the applicator.

If a clot has not formed, initiate clotting by placing a wooden applicator stick into the tube containing the collected blood, then begin again at step 1.

3. Transfer serum to a 50-ml centrifuge tube. Pellet any remaining blood cells and debris by centrifuging 10 min at $2700 \times g$ (4000 rpm in H-1000B rotor), 4°C , and save supernatant.
4. Assay antibody titer by the appropriate method: immunoprecipitation, immunoblotting, ELISA, or double-immunodiffusion assay in agar.
5. Store serum in aliquots in screw-top tubes at -20°C .

Some sera lose activity on repeated freezing/thawing; others aren't stable at 4°C .

COMMENTARY

Background Information

The kinetics of development of a specific antibody response upon immunization of a rabbit with antigen are illustrated in Figure 11.12.2. After the primary immunization, naive B cells are stimulated to differentiate into antibody-secreting plasma cells. For most soluble protein antigens, specific antibody begins to appear in the serum 5 to 7 days after the animal is injected. The antibody concentration (titer) continues to rise and peaks around day 12, after which it decreases. Similar kinetics are observed with mice, rats, hamsters, and rabbits.

In addition to differentiating into antibody-forming cells, the antigen-stimulated B cells proliferate to form a large population of memory B cells, which quickly become activated after the booster injection is administered. Thus, the lag period before the appearance of the specific antibody is much shorter after a booster injection than that observed for the initial immunization. In addition, a significantly higher titer of specific antibody is

achieved and sustained for a longer period of time. The peak of antibody production occurs 7 to 14 days after boosting. As a consequence of the existence of the memory B cells, less antigen is required to stimulate a strong secondary response. Memory B cells are long-lived; therefore, a specific antibody response can be elicited as much as 6 months to a year after the last booster. Finally, the average affinity and degree of specificity of the antibody population for the antigen increase with repeated immunizations (Klinman and Press, 1975).

Adjuvants greatly enhance the specific antibody titer, as they allow the antigen to be released slowly, thus ensuring the continual presence of antigen to stimulate the immune system. Freund's adjuvant has been used extensively in the preparation of antigen because it induces a high, long-lasting antibody titer that is often still measurable 25 weeks or more after boosting. The presence of killed mycobacteria in complete Freund's adjuvant (CFA) activates the T cell population, providing necessary lymphokines for B cell stimulation and matu-

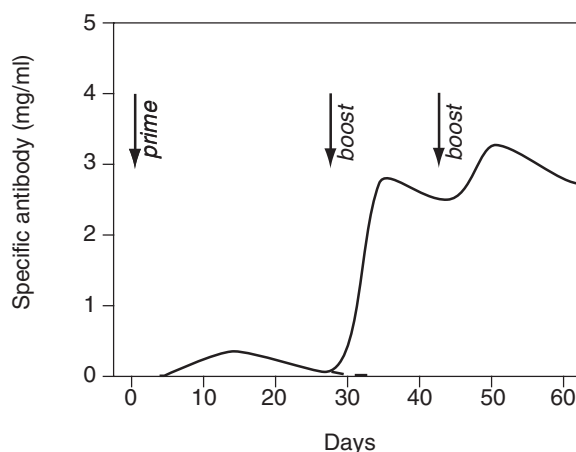


Figure 11.12.2 Kinetics of development of the specific antibody response. Arrows indicate when priming and boosting immunizations were administered. Actual amounts of specific antibody produced will vary considerably depending on immunogenicity of the protein.

ration. CFA may cause granuloma and subsequent necrotic abscesses, so it should be used only for primary immunization. Incomplete Freund's adjuvant (IFA) is adequate for booster injections. In the past decade, adjuvant research has concentrated on the production of effective adjuvants that minimize animal trauma. Wherever possible, to reduce animal discomfort, less noxious adjuvants should be used as described in the Alternate Protocol. In addition some countries are restricting the use of CFA in laboratory animals. Published comparisons of commercial adjuvants with the basic CFA/IFA protocol vary widely in their conclusions. In most cases, however, CFA/IFA produces higher titers of higher-affinity antibodies in a shorter time period (Johnston et al., 1991; Deeb et al., 1992; Smith et al., 1992). Conflicting results (Mallon et al., 1991; Bennett et al., 1992) may reflect differences in the immunogenicity of the antigens used.

Critical Parameters

New Zealand red or white rabbits are generally the best source of specific antisera because 30 to 50 ml of whole blood can be obtained at each bleed. The life span of a rabbit is 5 to 6 years, so a continual source of specific antiserum can be provided over a period of time by one rabbit after booster injections. In this regard, the recommended times between booster injections are not critical; the animal may be rested for several months between subsequent boosters, after the primary and secondary booster injections. Blood collection, however, must take place 7 to 14 days after each booster to ensure a high titer.

Preimmune serum from the same animal is the preferred negative control. If additional control serum is required, either immune serum from animals immunized with totally unrelated antigens or pooled serum from naive animals will be adequate. Occasionally, spurious antibody activities from nonimmunized animals may mimic the activity of the immune serum.

Antibody specificity may vary widely between individual animals with respect to the dominant antigenic epitopes recognized on a given protein antigen. Therefore, antiserum from a single animal should be used throughout a study. If more than one animal must be used for particular antisera, the antisera should be pooled. Large-scale production of antisera can be carried out in goats, sheep, and horses with appropriate veterinary guidance. If serum is taken from inbred animal strains, the variability

in antibody specificity, as observed in outbred rabbits, is less of a problem.

The most important factor in producing a highly specific polyclonal antiserum is the purity of the antigen preparation used for immunization. The immune system is very sensitive to the presence of foreign proteins. Any contaminating proteins in the antigen preparation can potentially induce a strong immune response when injected in the presence of adjuvant. When antisera are employed in sensitive techniques such as immunoblotting or the screening of cDNA or genomic libraries, significant antibody titers to protein contaminants can be a major problem. Thus, the antigen preparation should contain no significant contaminating proteins. Ideally, there should be no visible contaminating bands when 10 to 20 μg are analyzed on an SDS-polyacrylamide gel stained with Coomassie brilliant blue (*UNITS 10.2 & 10.6*).

If the antiserum is to be used in functional assays, extra care must be taken to ensure that the immunizing antigen is in its native form, because antibodies directed to denatured forms of the protein antigen will interact weakly, if at all, with the antigen in its native conformation. On the other hand, antibodies used in immunoblots, immunoprecipitation of primary in vitro translation products, and immunoscreening of $\lambda\text{gt}11$ expression libraries may be most effective if generated against a denatured protein with reduced and carboxymethylated disulfide bonds.

Troubleshooting

Inability to attain high-titer antiserum after several booster injections may be due to a variety of factors as described below.

Use of inappropriate adjuvant. Some experimentation may be necessary to optimize the antigen/adjuvant ratio for different antigens. If the alternate protocol still fails to produce a good antibody titer after three immunizations, switch to the basic protocol.

Inadequate antigen emulsification. If the emulsion fails the drop-on-water test described in the Basic Protocol (step 3), repeat the emulsification process. Be sure to use phosphate-buffered saline. Avoid plastic syringes and Tris-based buffers with CFA and IFA.

The antigen is a poor immunogen. In general, the immunogenicity of a protein is related to the degree to which it differs from "self" proteins (Benjamin et al., 1984). Large bacterial or viral proteins such as hemagglutinin or bacterial-coat proteins are highly immuno-

genic, whereas proteins from mammalian sources, such as polypeptide hormones or cell-surface receptors, may be poorly immunogenic due to tolerance. Protein antigens can be made more immunogenic in two ways. First, they can be chemically linked to a carrier protein that is known to be a good immunogen. Common carrier proteins include keyhole limpet hemocyanin (KLH), fowl immunoglobulin, and bovine serum albumin (BSA). Coupling peptides to carrier proteins is described in *UNIT 11.16*; the same protocols can be used to couple the protein antigen of interest to the desired carrier. Second, the immunogenicity of an antigen may be enhanced by its polymerization into large aggregates via a cross-linking agent such as glutaraldehyde. The protocol in *UNIT 11.16* for the coupling of peptide antigens to a carrier protein with glutaraldehyde can also be used to polymerize any protein antigen. With both the coupling and polymerization procedures, any insoluble antigen complexes formed should be removed prior to immunization by centrifuging 10 min at $15,000 \times g$, 4°C .

Host animal's immune system may be compromised by bacterial or viral infection. Refer to Donovan and Brown (1995c,d) for discussion of the consequences of poor animal husbandry. Utilize animals from reliable, pathogen-free sources and maintain them in appropriate infection-free facilities.

Only a few animals have been immunized. Because of the vagaries of immune-response genes in outbred animals such as rabbits, some antigens may not induce a good antibody response in a significant proportion of randomly selected animals. Thus, it is best to immunize several different animals and to screen the sera for the best responder. Obviously, this is less of a problem in homozygous inbred strains, but with a new antigen it is wise to test several strains for their antibody response.

An insufficient amount of antigen was used. Although recommended concentrations of antigen for rabbits are 0.25 to 0.5 mg/ml injected into multiple sites, for a total of 1 to 2 ml in the same animal, good results can be obtained with $1/10$ to $1/20$ of the concentration in the same volume. It is always tempting to use less of a precious antigen, but often too low a dose leads to too low a response.

Anticipated Results

For large or nonevolutionarily related proteins, a titer of 5 to 10 mg/ml of serum can be expected after repeated boosts (hyperimmunization). When immunizing with small or highly

conserved protein species, a titer of 1 to 2 mg/ml of specific antibodies is more likely. Antibody titers and affinity for the antigen will be low after primary immunization and the first booster immunization, but both titer and affinity will increase with subsequent immunizations.

Time Considerations

Preparation of the immunogen and immunization will take ~3 hr on each occasion. Collection of antisera will take 1 to 2 hr, depending on the number and species of animals.

Collection of antiserum after the primary immunization will be at 10 to 14 days. This will be a low-titer, low-affinity serum. The first booster normally is given 4 to 8 weeks after the primary immunization but can be given as early as 2 weeks after the primary if time is critical. Ideally, there should be at least 19 days between the primary and the secondary bleeds. A second booster is given at 6 weeks with a bleed on day 52 to 59. This will usually be the first high-titer bleed. If a titer of <1 mg/ml of specific antibody is obtained, subsequent boosting immunization will be necessary.

Literature Cited

- Benjamin, D.C., Berzofsky, J.A., East, I.J., Gurd, F.R.N., Hannum, C., Leach, S.J., Margoliash, E., Michael, J.G., Miller, A., Prager, E.M., Reichlin, M., Sercarz, E.E., Smith-Gill, S.J., Todd, P.E., and Wilson, A.C. 1984. The antigenic structure of proteins: A reappraisal. *Annu. Rev. Immunol.* 2:67-101.
- Bennett, B., Check, I.J., Olsen, M.R., and Hunter, R.L. 1992. A comparison of commercially available adjuvants for use in research. *J. Immunol. Methods* 153:31-40.
- Deeb, B.J., DiGiacomo, R.F., Kunz, L.L., and Stewart, J.L. 1992. Comparison of Freund's and Ribi adjuvants for inducing antibodies to the synthetic antigen (TG)-AL in rabbits. *J. Immunol. Methods* 152:105-113.
- Donovan, J. and Brown, P. 1995a. Parenteral injections. *In Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 1.6.1-1.6.10. John Wiley & Sons, New York.
- Donovan, J. and Brown, P. 1995b. Blood collection. *In Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 1.7.1-1.7.8. John Wiley & Sons, New York.
- Donovan, J. and Brown, P. 1995c. Animal health assurance. *In Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 1.1.1-1.1.4. John Wiley & Sons, New York.

- Donovan, J. and Brown, P. 1995d. Managing immunocompromised animals. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, eds.) pp. 1.2.1-1.2.6. John Wiley & Sons, New York.
- Freund, J., Casals, J., and Hismar, E.P. 1937. Sensitization and antibody formation after injection of tubercle bacilli and paraffin oil. *Proc. Soc. Exp. Biol. Med.* 37:509.
- Johnston, B.A., Eisen, H., and Fry, D. 1991. An evaluation of several adjuvant emulsion regimens for the production of polyclonal antisera in rabbits. *Lab Anim. Sci.* 41:15-21.
- Klinman, N.R. and Press, J. 1975. The B cell specificity repertoire: Its relationship to definable subpopulations. *Transplant. Rev.* 24:41-83.
- Mallon, F.M., Graichen, M.E., Conway, B.R., Landi, M.S., Leonard, T.B., and Hughes, H.C. 1991. A comparison of antibody response using a synthetic adjuvant system and Freund's complete adjuvant in rabbits. *Am. J. Vet. Res.* 52:1503-1506.
- McWilliam, A. and Niemi, S.M. 1988. Freund's adjuvant. Canadian Council on Animal Care Resource. 12:1.
- Smith, D.E., O'Brien, M.E., Palmer, V.J., and Sadowski, J.A. 1992. The selection of an adjuvant emulsion for polyclonal antibody production using a low-molecular-weight antigen in rabbits. *Lab. Anim. Sci.* 42:599-601.

Key Reference

Harlow, E. and Lane, D. 1999. Using Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

A comprehensive methods book with many modern techniques.

Contributed by Helen M. Cooper
Ludwig Institute for Cancer Research
Melbourne, Australia

Yvonne Paterson
University of Pennsylvania
Philadelphia, Pennsylvania

This unit describes the antigenic stimulation of in vitro antibody production by B cells and the subsequent measurement of secreted antibodies. Such in vitro systems have been extremely useful for studying the regulation of interacting cell types involved in antibody production—namely B cells, T cells, and macrophages—as well as their antigenic stimulants. In addition, these studies have revealed important information about cellular requirements for growth and differentiation factors during antibody production.

A generalized system for inducing in vitro antibody production, and which can accommodate various types of antigens under study, is first presented (see Basic Protocol 1). Secreted antibodies can then be measured with an enzyme-linked immunosorbent assay (ELISA; UNIT 11.2) or other soluble-antibody detection systems. Alternatively, the number of antibody-producing cells can be quantified by plaque-forming cell (PFC) assays: the Cunningham-Szenberg technique (see Basic Protocol 2) and the Jerne-Nordin technique (see Basic Protocol 3). Both methods employ specially prepared slide chambers (see Support Protocols 1 and 2) in which the antibody-producing B cells are mixed with complement and indicator sheep red blood cells (SRBC), or with trinitrophenol-modified SRBC (TNP-SRBC; see Support Protocol 3), with subsequent lysis and counting of plaques. Because IgM antibodies fix complement efficiently, whereas IgG and IgA antibodies do not, unmodified PFC assays measure only IgM antibodies. The assay can be modified, however, to measure all classes of antibodies or to enumerate total immunoglobulin-secreting B cells (see Alternate Protocol). Yet another method of measuring the number of antibody-producing B cells (in a class-specific fashion) is to use the ELISPOT technique (Lycke and Coico, 1996). The resting B cells used in these procedures are prepared as described in Support Protocol 5 for Percoll gradient centrifugation.

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique must be used accordingly.

STRATEGIC PLANNING

Choice of Antigen

In setting up an in vitro system for antibody production and measurement, the antigen to be used is a critical factor that will determine the types of cells and lymphokines necessary for an adequate response.

A first consideration is whether the antigen to be used is one that contains immunogenic epitopes that the animal from which cells are to be obtained has “seen” in vivo. Generally speaking, if the antigen is new and a primary response is being elicited in vitro, the response will be small, perhaps below the limits of measurement, because the cells utilized do not contain the number of specific precursors necessary for a measurable in vitro response. Responses to most protein antigens fall into this category, so that in vitro induction with such antigens require the use of cells from in vivo primed animals. An apparent exception to this rule is the particulate antigen SRBC, which provides good responses even with unprimed cells. This, however, can be attributed to the fact that SRBC contain immunogenic epitopes that cross-react with those present in environmental antigens, so that the cell donor has, in fact, been exposed to this antigen in vivo. It should be noted in this regard that SRBC induce only IgM antibody responses in in vitro cells obtained from unprimed animals, probably reflecting the fact that cells from unprimed animals do not contain enough SRBC-specific precursors of isotype-switched B cells.

Thus, to obtain IgG responses with SRBC antigen, one must use cells from SRBC-primed animals.

A second consideration is whether the antigen is a T cell-dependent or -independent antigen. T cell-dependent antigens (most protein antigens) require the presence of antigen-primed T cells which provide the necessary factors (lymphokines) for B cell proliferation and differentiation. In addition, for these antigens there is increasing evidence that T cells are required for promoting cell-cell interactions necessary for antibody production. Perhaps the one exception to this rule is that certain T cell-dependent antigens such as SRBC can stimulate primed B cells in the absence of T cells, provided that T cell lymphokines are present.

T cell-independent antigens, as defined originally, can stimulate B cells in vitro in the absence of T cells. Type I antigens, such as trinitrophenylated lipopolysaccharide (LPS), have intrinsic polyclonal activating properties, whereas type II antigens, such as trinitrophenylated Ficoll and dextran, do not possess this characteristic. While both stimulate in vivo responses in nude mice and in vitro responses in T cell-depleted spleen cells, type I antigens can stimulate responses in immature neonatal B cells and in cells from *xid* immune-defective mice, but type II antigens stimulate responses only in immunologically mature B cells. The T cell-independent designation for these antigens is now outdated, as it has become clear that T cell-derived lymphokines are necessary for the response when very pure resting B cells are used (Mond et al., 1983; Thompson et al., 1984). Consequently, T cell-derived lymphokines are now usually added to systems that use T cell-independent antigens as a stimulant. It is therefore most correct to refer to this class of antigen as "T cell-regulated" or as "T cell-independent type I or type II."

Decisions concerning the type of antigen used (and the cell/lymphokine requirements) must be made in order that the response to be studied is precisely defined. In this regard, B cells responding to complex, multiepitopic antigens of any kind are a clonally diverse mixture of cells; the antibody response will therefore be heterogeneous and will comprise antibodies with specificity for the various immunodominant determinants on the antigen. However, if a modified antigen is utilized—i.e., an antigen conjugated to a chemically defined low-molecular-weight antigenic epitope—the response to this epitope can then be measured. In this case the known epitope is called the hapten and the molecule to which it is attached is called the carrier; since B cell responses are being quantified, the B cells in these systems recognize the hapten and T cells recognize other epitopes on the carrier (Mitchison, 1971). TNP is frequently employed as the hapten; this chemical group can be placed on various carriers, including soluble proteins (ovalbumin), to measure highly defined T cell-dependent B cell responses, as well as on other T cell-independent carriers such as *Brucella abortus* (BA) and LPS to measure type I responses, or on Ficoll and dextran to measure type II responses.

Table 11.13.1 summarizes lymphokine requirements of cellular responses to some antigens. This table is incomplete in that the type of lymphokine that is added may have considerable effect both on the quality of the responses (such as the isotype profile) and on the magnitude of the response. For example, in most systems IL-4 and IFN- γ are required lymphokines for stimulation of IgE and IgG2a responses, respectively.

A final consideration relating to choice of antigen concerns the need for macrophages, either as antigen-presenting cells (APC) or as sources of cytokines in the culture system. Whereas B cells can present those antigens for which they express specific surface Ig receptors, they are not generally sufficient in the usual in vitro culture systems because such antigen-specific cells are present in low numbers. Macrophages, on the other hand, take up and present antigen in an antigen-nonspecific manner and are thus the primary

choice of APC in in vitro systems. Thus, for responses to T cell–dependent antigens, where T cell activation is required for inducing a response, the need for macrophages is absolute. T cell–independent responses, however, are relatively independent of the requirement for macrophages. It is important to note that the presence of too many macrophages may be suppressive for some in vitro responses, because of the release of inhibitory substances (e.g., prostaglandins and certain cytokines).

Preparation of Cells

Prior to setting up cultures to measure in vitro antibody production, considerable thought has to be given to the methods for preparing the cell populations to be studied. Unprimed, unseparated cell populations may be obtained from a variety of lymphoid organs. More frequently, however, purified T and B cell populations are used (Hathcock, 1991a,b; Mage, 1993), or purified populations of resting B cells alone can be used. Priming of T cells is described in Kruisbeek and Shevach (1991); priming of B cells is accomplished by immunizing mice with hapten-carrier conjugate emulsified in complete Freund's adjuvant 3 to 5 weeks prior to cell harvest (using a different carrier than that used for T cell priming). It is best to size-separate B cells on Percoll gradients (see Support Protocol 5) after enrichment as in Hathcock (1991b) when studying T cell–independent responses, since responses of resting and partially activated B cells vary (Thompson et al., 1984). This cannot be done when studying T cell–dependent responses, as antigen-primed (activated) B cells are required.

INDUCTION OF ANTIGEN-SPECIFIC AND POLYCLONAL ANTIBODY PRODUCTION

Antibody production is induced by culturing B cells, desired antigens (either sheep red blood cells or trinitrophenylated carriers), appropriate lymphokines, and T cells for 4 to 5 days (see Strategic Planning, above). Secreted antibodies can be measured in the supernatant with ELISA assays (UNIT 11.2). Alternatively, the number of antibody-producing cells can be measured as in Basic Protocol 2 or Basic Protocol 3.

Materials

- Antigens (see recipe): SRBC, TNP-BA, TNP-Ficoll, TNP-dextran, TNP-ovalbumin (TNP-OVA), or TNP-LPS
- Recombinant lymphokines (for use with purified B cells): rIL-2 (Cetus or Schering) and/or rIL-1 (Hoffman La Roche)
- Complete RPMI-10 medium (see recipe) or complete DMEM-10 medium (see recipe)

Table 11.13.1 Cells and Lymphokines Required for Stimulation of In Vitro Antibody Production^a

Antigen	T Cells	B Cells	Lymphokines	PFC/culture
TNP-Ficoll	–	+	IL-1, IL-2, or IL-5	400 (IgM)
TNP-BA	–	+	IL-1, IL-2, or IL-5	2000 (IgM)
SRBC (T-independent)	–	+	IL-1, IL-2, or IL-5	1000 (IgM)
SRBC (T-dependent)	T'	+	–	3000 (IgM)
				4000 (IgG)
TNP-OVA	T'	+	–	500 (IgM)
				300 (IgG)

^aAbbreviations: BA, *Brucella abortus*; OVA, ovalbumin; PFC, plaque-forming cells; SRBC, sheep red blood cells; T', antigen-primed T cells; TNP, trinitrophenylated.

B lymphocytes purified from mice (Support Protocol 5) either nonimmunized or immunized 3 to 5 weeks previously with specific hapten-carrier conjugate (UNIT 11.12 and Table 11.13.1)

T cells primed to antigen or carrier (Kruisbeek and Shevach, 1991; Table 11.13.1)

96-well flat-bottom microtiter plates (Costar)

Refrigerated low-speed centrifuge (e.g., IEC 7R with 216 rotor)

Sterile gauze

1. Add antigens and recombinant lymphokines diluted to the appropriate concentrations in supplemented complete medium in a total volume of 0.15 ml to the wells of a 96-well microtiter plate. Set up in triplicate to permit calculation of standard error.

The amount of antigen to be used must be determined separately in each system. A wide range of concentrations should be tested in pilot experiments and one or two concentrations in later (more definitive) studies.

Approximate final concentrations in the wells for representative antigens:

2 × 10⁶ cells/well SRBC

10 µg/ml TNP-OVA

1:100 dilution of stock TNP-BA

10 ng/ml TNP-Ficoll

5 to 50 µg/ml TNP-LPS or unconjugated LPS

Lymphokines are added when responses to T cell-independent antigens are studied. If TNP-Ficoll or SRBC are used as the stimulating antigen, rIL-1 and rIL-2 (or alternatively, rIL-5) are recommended. When TNP-BA is used as antigen, only IL-2 or IL-5 is necessary. rIL-1 is used at 10 U/ml and rIL-2 at 100 U/ml. When T cells are added, lymphokines are not required.

2. Add 10⁵ to 10⁶ B cells in 50 µl to each well with multichannel pipettor to give a final volume of 0.2 ml.

Appropriate cell number must be determined empirically. Use lower cell concentrations for responses elicited by polyclonal stimuli, and higher cell concentrations for responses driven by specific antigens, but never use more than 10⁶ cells per well.

3. Incubate plates for 4 to 5 days at 37°C, 5% CO₂. If an assay of secreted antibody is desired, go to step 4. If a plaque assay is to be performed, go to step 7.

During the culture period, preparation can be made for analysis of culture supernatants by an enzyme-linked immunosorbent assay (UNIT 11.2).

4. Centrifuge plates 10 min at 1000 rpm (170 × g), 4°C. In a tissue culture hood, discard supernatant by inverting and flicking plate, then blot dry with sterile gauze.

5. Add 0.2 ml fresh supplemented complete medium to each well; repeat step 4 twice.

6. Add 0.2 ml fresh supplemented complete medium to each well and incubate plates 24 to 48 hr in a humidified 37°C, 5% CO₂ incubator. Remove and save 150 µl supernatant from each well with a multichannel pipettor if secreted antibody is to be measured using ELISA or other suitable technique.

7. Centrifuge plates 10 min at 170 × g, 4°C. Invert plates in one quick motion and flick off supernatant. Without reinverting plates, remove excess fluid by blotting plates on absorbent paper.

8. Add 0.2 ml supplemented complete medium to each well and repeat step 7.

9. Add 0.2 ml supplemented complete medium to each well. Tap plates gently or place on vibrating platform to ensure complete resuspension of cells.

Repeated washing of the wells is required to remove free antibody before the subsequent assay of antibody production. Antibody-producing B cells are now ready to be measured as described for plaque-forming cell assays (see Basic Protocols 2 and 3).

PLAQUE-FORMING CELL ASSAYS

B cells prepared in Basic Protocol 1 can be assayed for antibody production using a plaque-forming cell (PFC) assay. In this approach, the B cells are plated in (or on) a layer of hapten- or antigen-conjugated SRBC under conditions in which the secreted antibody will cause local SRBC lysis, resulting in pin-point areas of clearing in the lawn of SRBC (plaques) that can be counted. The two widely used PFC assays presented below are alike in principle but different in detail. Antibody-producing B cells are mixed with a source of complement and indicator RBC if SRBC are used as stimulating antigen, or TNP-RBC if TNP-conjugated antigens are used as the stimulating antigen. The B cell/SRBC mixture is then processed in Cunningham-Szenberg chambers or with Jerne-Nordin slides. This technique measures only cells secreting IgM since only antibody of the latter isotype fixes complement under the conditions of the assay. It can, however, be modified to measure cells secreting antibody of any class by the addition of isotype-specific rabbit anti-Ig antibody that binds to secreted immunoglobulin and forms complexes that fix complement. The PFC assays can also be modified to enumerate total immunoglobulin-secreting B cells stimulated by a polyclonal B cell activator, e.g., LPS (see Alternate Protocol).

There are advantages and disadvantages of both methods. The principle advantage of the Cunningham-Szenberg assay is that six to eight plates can be processed in one day. Its disadvantage is that extra time and effort are required for preparing the slide chambers, which cannot be reused, and for preabsorbing the complement. The Jerne-Nordin assay, on the other hand, accommodates only three 96-well plates in an 8-hr period. However, while this technique requires construction of a tray for holding the slides, the tray can be reused indefinitely; in addition, complement needn't be preabsorbed. Thus, for IgG or protein A plaquing, Jerne-Nordin is the assay of choice.

Cunningham-Szenberg Assay (Cunningham and Szenberg, 1968)

Materials

Supplemented complete RPMI-10 (see recipe) *or* DMEM-10 medium (see recipe)
7.5% (v/v) SRBC or 15% (v/v) TNP-SRBC (see Support Protocol 3) in
supplemented complete medium

Source of complement: 50% (v/v) guinea pig serum (Life Technologies; also see
recipe) in supplemented complete medium

Wax (tissue preparation grade; Fisher)

96-well round-bottom microtiter plate (Linbro # 36-311-05)

Cunningham-Szenberg chambers (see Fig. 11.13.1 and Support Protocol 1)

120-cm² glass petri dish

Dissecting microscope

NOTE: All reagents should be room temperature before addition to chambers, because cold reagents result in bubble formation during incubation.

**BASIC
PROTOCOL 2**

Immunology

11.13.5

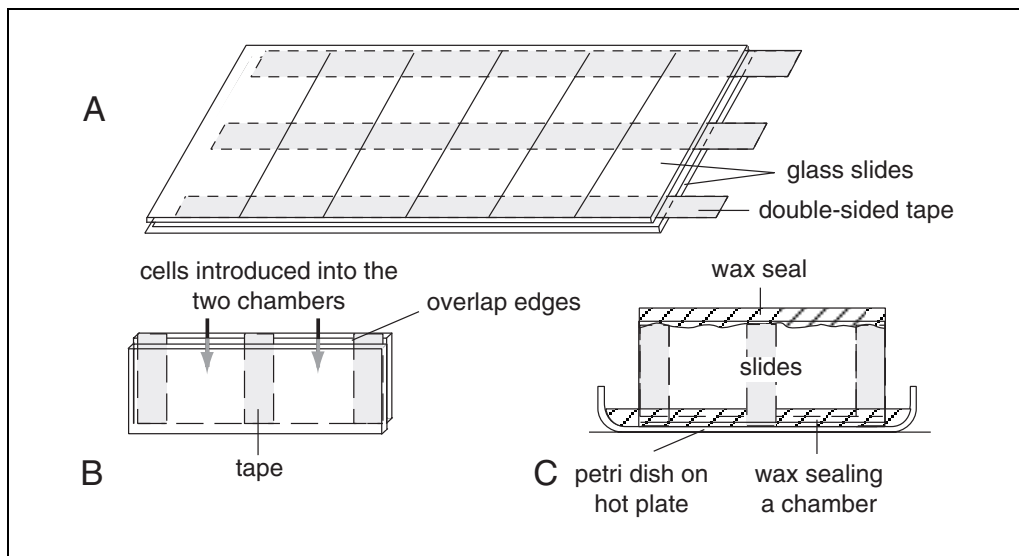


Figure 11.13.1 The Cunningham-Szenberg chamber. **(A)** The chambers as they are made in strips. **(B)** The whole chamber. **(C)** The chambers being dipped into a petri dish and sealed with hot wax.

1. Prepare molten wax by heating in a 120-cm² glass petri dish on a hot plate; maintain at low heat.
2. To each well of an empty 96-well microtiter plate, add the following:
 - 50 μ l supplemented complete medium
 - 50 μ l 7.5% SRBC or 15% TNP-SRBC
 - 50 μ l suspended B cells (from step 9 of Basic Protocol 1)
 - 50 μ l diluted guinea pig serum

Mix the contents of each well with pipettor and fill Cunningham-Szenberg chambers by slowly expelling the contents of the pipet with the tip placed at an angle at the overlap edges (see Fig 11.13.1B).

Add guinea pig serum to only 24 wells at a time to prevent inactivation of the complement during the period the slides are kept at room temperature. The time needed to fill and seal 24 slides is at the upper limit necessary to prevent the chambers from drying out.

3. After filling 24 slide chambers, seal each side with molten wax (see Fig. 11.13.1). Incubate chambers 45 to 60 min at 37°C.

Metal trays from standard 37°C incubators are excellent for carrying slides; each tray holds ~72 slides.

Seal with wax within 20 min to prevent the chamber from drying out.

4. Remove chambers from incubator and count plaques under a dissecting microscope at 10 \times magnification.

Care must be taken to differentiate between air bubbles and real plaques. Air bubbles have clearly defined borders and are highly reflective. Plaques have hazy borders and are nonreflective (see Fig. 11.13.2).

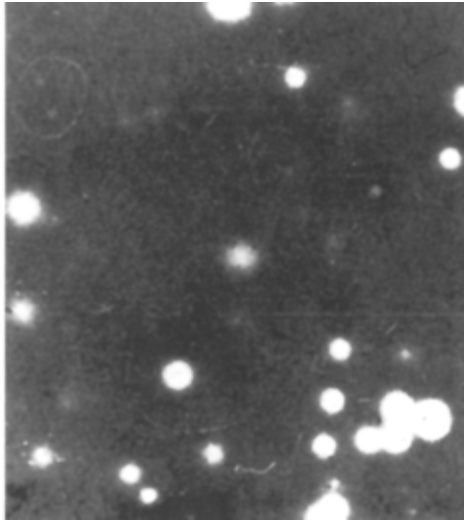


Figure 11.13.2 Typical plaques from PFC assay. The plaques show pinpoint clearing in a lawn of RBC. This is due to lysis by complement and antibody being secreted by the B cell (center of plaque).

Jerne-Nordin Assay (Jerne and Nordin, 1963)

Materials

- 1% agarose (SeaPlaque; FMC Bioproducts)
- 2× basal Eagle medium (Life Technologies)
- 10% (v/v) SRBC or 20% (v/v) TNP-SRBC (see Support Protocol 3) in supplemented complete medium
- Phosphate-buffered saline (PBS; APPENDIX 2), 4°C
- Source of lyophilized guinea pig serum (Life Technologies)
- 5-ml pipets and 5-ml glass test tubes, prewarmed
- 42°C water bath
- Precoated agarose slides (see Support Protocol 2)
- Tray for holding slides (Fig. 11.13.3; not commercially available)

1. Mix equal amounts of 1% agarose with 2× basal Eagle medium in a 50-ml centrifuge tube. Using a prewarmed pipet, add 0.4 ml of this plaquing medium to glass test tubes (maintained in a rack in a 42°C water bath).

Maintain this mixture at 42°C until the B cells are added and placed onto the slides. This is critical, because at lower temperatures agarose slowly and imperceptibly gels and gives rise to false plaques. The cooled agarose can no longer be used.

2. To each test tube, first add 1 drop of 10% SRBC or 20% TNP-SRBC, then add (6 tubes at a time) 0.1 ml B cell suspension (from step 9 of Basic Protocol 1) after carefully resuspending cells.
3. Vortex tubes and immediately pour and spread suspensions onto precoated agarose slides. Allow slides to harden and turn upside down on a tray (see Fig. 11.13.3).

Trays are designed such that eight slides (or less) can be placed upside-down.

4. Cover tray with a damp cloth and place 60 to 90 min in an incubator at 37°C.

Even if a humidified incubator is used, do not omit the damp cloth.

5. Prepare complement by adding 5 ml of cold PBS to 5 ml lyophilized guinea pig serum. When serum is dissolved, add it to 95 ml of cold PBS.

**BASIC
PROTOCOL 3**

Immunology

11.13.7

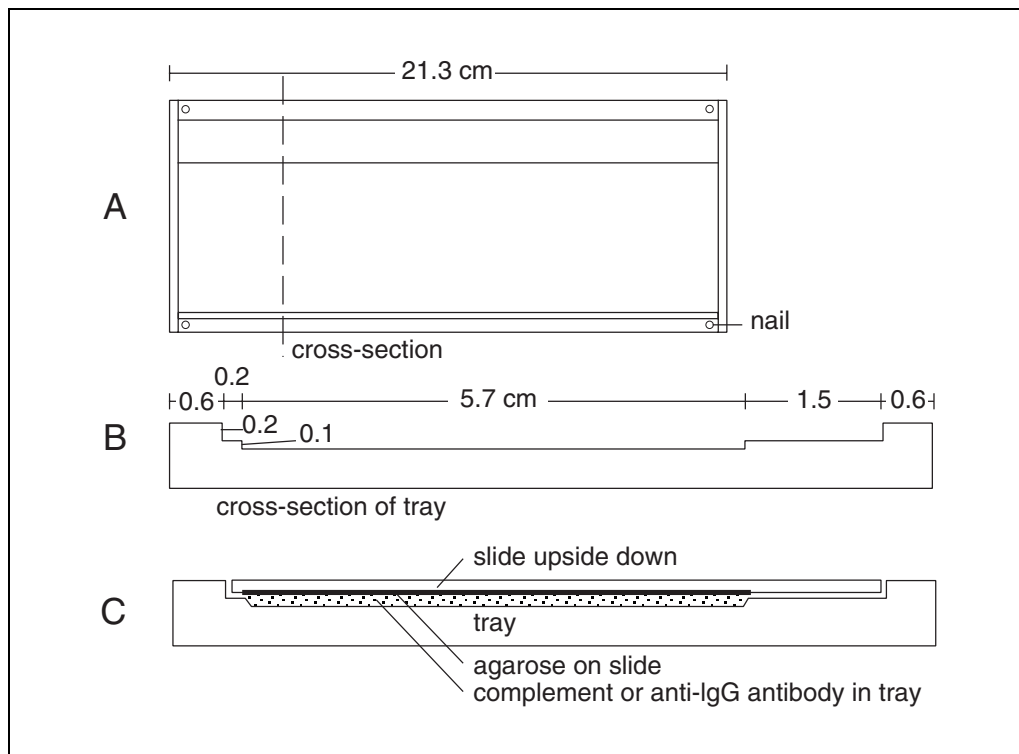


Figure 11.13.3 Jerne-Nordin slide assembly. **(A)** Top view of tray sufficient for eight slides. In the corners, nails can be worked through to facilitate stacking of multiple trays. **(B)** Amplified cross-section of the tray with detailed dimensions (in cm). **(C)** Cross-section of tray with agarose-coated slide in place for incubation.

6. Add 2 ml of freshly diluted guinea pig serum per slide to the shallow well in the tray holding the slides (i.e., 16 ml/tray; see Fig. 11.13.3). Place slides in 37°C incubator for 90 min.

There is no further need for a damp cloth, as the complement maintains the moisture.

7. Remove slides from incubator and count plaques using a dissecting microscope at 10× magnification (Fig. 11.13.2).

ALTERNATE PROTOCOL

MEASUREMENT OF ISOTYPE-SPECIFIC ANTIBODY AND POLYCLONAL ANTIBODY RESPONSES

Isotype-Specific Antibody Response

The Jerne-Nordin PFC assay (see Basic Protocol 3) can be modified for measurement of IgG- and IgA-producing B cells as follows. At the end of step 1, add 50 μ l goat anti-IgM antibody to each tube containing plaquing medium to inhibit IgM (direct) plaques. Then, after the incubation in step 4 (before preparing complement in step 5), add developing anti-IgG (or anti-IgA) to the slides to develop IgG (or IgA)-producing B cells. This is done by filling the shallow well in the tray holding the slides with 2 ml of antibody. Incubate 1 hr at 37°C, then draw off the developing antibody with a Pasteur pipet connected to a vacuum suction flask. Proceed with the preparation and addition of complement as in steps 5 and 6.

Appropriate dilution of rabbit anti-IgM, -IgG, and -IgA antibodies (Jackson Labs) in supplemented complete medium must be predetermined for each batch of antibody, but is usually in the range of 1:100. Addition of developing antibodies is critical, since IgG and IgA antibodies fix complement poorly in their absence.

Polyclonal Antibody Response

The Jerne-Nordin PFC assay (see Basic Protocol 3) can also be modified to measure total Ig-secreting B cells stimulated by a polyclonal B cell activator—e.g., LPS. In this case, Ig of unknown antigenic specificity is measured; thus, an indicator cell must be used which binds Ig of all isotypes. The indicator cells are SRBC conjugated to protein A (see Support Protocol 4), a substance that binds both mouse and rabbit IgG with high avidity. Protein A–SRBC thus binds a developing rabbit anti-mouse Ig antibody that has been added to the B cell preparation and subsequently binds secreted mouse Ig. The resulting complex is then able to activate complement and effect SRBC lysis.

To assay polyclonal antibody response, prepare protein A–SRBC indicator cells according to Support Protocol 3. Proceed with step 1 of the Jerne-Nordin protocol (see Basic Protocol 3), then add 1 drop of 15% protein A–SRBC (see Support Protocol 4) suspended in supplemented complete medium (instead of unmodified SRBC or TNP-SRBC) in step 2. Proceed with steps 3 and 4, then add developing antibody (rabbit anti-IgM and anti-IgG) to the shallow well in the tray holding the slides. Incubate slides 1 hr at 37°C, then draw off the developing antibodies with a Pasteur pipet connected to a vacuum suction flask. Proceed with the preparation and addition of complement as in steps 5 and 6.

It is best (but not essential) to use affinity-purified antibodies (Jackson Immunoresearch) that have been pretitrated to determine optimum dilution in supplemented complete medium.

PREPARATION OF CUNNINGHAM-SZENBERG CHAMBERS

Materials

70% ethanol
Microscope slides
¼-in. double-sided tape (3M)

1. Place 15 to 30 microscope slides side-by-side on a bench in a row.
2. Stick double-sided ¼-in. tape on top and bottom edges. Place a third strip of tape between the other two, dividing slide into two sections (see Fig. 11.13.1).
3. Wipe slides with a paper towel soaked with 70% ethanol and allow to air dry.
4. Remove backing of double-sided tape and wipe a second slide with 70% ethanol, air dry, and place on first slide in the row. Overlap the slides on the side edges as in Figure 11.13.1 to facilitate filling the chamber. Repeat this process so the entire row of slides on the bench are covered with a second slide.
5. Apply slight pressure with a roller to ensure that the slides are stuck firmly together.

Each slide chamber can hold just under 0.1 ml of fluid on either side of the tape for a total of 0.2 ml per slide.

SUPPORT PROTOCOL 1

**PREPARATION OF GLASS SLIDES FOR THE
JERNE-NORDIN ASSAY**

Materials

1% agarose (SeaPlaque; FMC Bioproducts)
Microscope slides
2 × 2-in. sterile gauze pad (Johnson & Johnson)

1. Prepare microscope slides as for Cunningham-Szenberg chambers (see Support Protocol 1), but omit the double-sided tape (i.e., follow steps 1 and 3).
2. Boil 1% agarose and dilute 1:10 in water (0.1% final).
3. Apply a thin film of 0.1% agarose to prepared slides (frosted end up) with a sterile gauze pad.
4. Allow slides to dry.

Precoated slides can be stored for 2 months at room temperature in a dust-free environment.

PREPARATION OF MODIFIED INDICATOR SRBC FOR PFC ASSAYS

TNP-Modified SRBC (Rittenberg and Pratt, 1969)

Additional Materials

Modified barbital buffer (MBB; see recipe)
2,4,6-trinitrobenzenesulfonic acid, sodium salt (TNBS; Eastman Kodak)
Cacodylate buffer (see recipe)
120 mg glycylglycine (gly-gly; Sigma #10022) in 15 ml MBB (see recipe for MBB)
15-ml graduated conical tube (Costar)

1. Remove SRBC from stock bottle and centrifuge 15 min in IEC 216 rotor at 3000 rpm (1525 × g). Aspirate supernatant and transfer 3 ml of packed SRBC to a tube containing 50 ml MBB.
2. Wash SRBC with 50 ml MBB by centrifuging 10 min at 1525 × g. Repeat the washing procedure two more times. Resuspend SRBC in 50 ml MBB.
3. Add 60 mg TNBS to 21 ml cacodylate buffer in a 50-ml foil-covered tube (TNBS is light sensitive). Mix and allow the tube to stand 10 min at room temperature.
4. *Slowly* add 5 ml SRBC/MBB suspension to TNBS/cacodylate solution, while shaking. Cap tube and invert occasionally over the next 10 min.
5. Fill tube with MBB. Centrifuge 10 min in IEC 216 rotor at 2000 rpm (675 × g), 4°C, and discard supernatant.
6. Add glycylglycine solution to tube and resuspend SRBC. Place tube on ice 10 min.

Formation of glycylglycine solution is such that 20 mg of glycylglycine is used for every 10 mg of TNBS.

7. Centrifuge 10 min at 675 × g and discard supernatant.

The supernatant should be yellow. If it is not, the conjugation step was not successful.

8. Wash SRBC twice in 50 ml MBB and at least once in supplemented complete medium, using the conditions in step 2. Continue washing until there is no evidence of SRBC lysis (clear supernatant).

9. Spin cells at $675 \times g$ in a graduated 15-ml conical tube to determine final packed cell volume. Discard supernatant and resuspend SRBC in supplemented complete medium at a final concentration of 15% or 20% (v/v) for Cunningham-Szenberg or Jerne-Nordin assays respectively.

TNP-SRBC can be stored ~1 week but should be washed prior to use and should not be used until lysis is absent. (i.e., until supernatant stays clear after washing). Fresh SRBC make better targets and are more likely to survive the conjugation procedure.

Protein A–Modified SRBC (Gronowicz et al., 1976)

Materials

Saline (0.15 M NaCl; Life Technologies)
0.5 mg/ml protein A (Pharmacia Biotech) in saline (store frozen in saline at 1 mg/ml)
1× chromic chloride solution (see recipe)
15-ml conical test tube (Costar)

1. Wash SRBC in 50 ml saline by centrifuging 10 min in IEC-216 rotor at 2000 rpm ($675 \times g$), 4°C. Repeat the washing procedure two more times.
2. Pipet 0.5 ml packed cells into a 15-ml conical test tube.
3. Add 0.5 ml freshly prepared 0.5 mg/ml protein A and 5 ml of 1× chromic chloride solution to SRBC. Incubate 1 hr at room temperature. Invert tube gently every 10 min.
4. Wash twice in 50 ml saline and once in 50 ml supplemented complete medium as described in step 1.
5. Resuspend SRBC at a final concentration of 15% (v/v) in supplemented complete medium.

Protein A–SRBC are best used on the same day, but will keep for 3 days at 4°C.

ISOLATION OF B CELLS BY PERCOLL GRADIENT CENTRIFUGATION

Resting B cells are isolated from a suspension of spleen cells by elimination of T cells with an anti-Thy-1 MAb (Hathcock, 1991b), followed by enrichment using flotation on Percoll gradients (also see Kruisbeek and Shevach, 1991). However, instead of using the two layers of 55% and 70% Percoll described in Kruisbeek and Shevach (1991) (for enrichment of accessory cells), B cell enrichment requires four layers of 55%, 60%, 65%, and 70% Percoll. All other conditions are identical to those described in Kruisbeek and Shevach (1991). Resting B cells are collected from the 65% to 70% interface after centrifugation.

Additional Materials

Hank's balanced salt solution (HBSS; APPENDIX 2), ice cold
Percoll (store at 4°C; Pharmacia Biotech #17-0891-01)
Percoll mix solution (see recipe)

15- and 50-ml polypropylene centrifuge tubes
Sorvall H-1000B rotor (or equivalent)

Additional reagents and equipment for preparation of depleted B cell suspensions (Hathcock, 1991b; Mage, 1993), making cell suspensions (Kruisbeek, 1993), and counting cells (APPENDIX 3F)

**SUPPORT
PROTOCOL 4**

**SUPPORT
PROTOCOL 5**

Immunology

11.13.11

1. Prepare T cell–depleted B cell suspension from spleen. Resuspend cells in HBSS using 1 ml of HBSS for each spleen.

Be sure to include anti-Thy-1, anti-CD4, and anti-CD8 antibodies to deplete all T cells. If desired, deplete RBC present in spleen cell suspension using ACK lysing buffer, pH 7.4 (see recipe); however, small RBC contamination usually does not interfere with proliferation studies.

2. Prepare 70% Percoll by mixing 290 ml Percoll with 170 ml Percoll mix solution. Keep on ice at all times.
3. Prepare 50%, 60% and 66% Percoll as follows:

	<u>50%</u>	<u>60%</u>	<u>65%</u>
70% Percoll, ice-cold (ml)	21.42	25.71	27.86
HBSS, ice-cold (ml)	8.58	4.29	2.14

4. Prepare 50%/60%/65%/70% Percoll gradients in 15-ml conical centrifuge tubes:

- a. Add 2.5 ml 70% Percoll to 15-ml tube.
- b. Gently pipet 2.5 ml of 65% Percoll over 70% Percoll layer.
- c. Gently pipet 2.5 ml of 60% Percoll over 65% Percoll layer.
- d. Gently pipet 2.5 ml of 50% Percoll over 60% Percoll layer.

Careful layering of Percoll is essential to create the sharp gradient interfaces necessary for good cell separation.

Use polypropylene tubes to withstand high-speed centrifugation. Make ≥ 4 gradients for a suspension of cells from 10 spleens and ≥ 6 gradients for a suspension of cells from 20 spleens.

5. Allow prepared gradients to sit on ice for ≥ 15 min before adding cells and centrifuging.

Good cell separations depend on use of cold gradients.

6. Gently layer 2.5 ml of cell suspension (from step 1) onto the surface of each gradient. Add cell suspension slowly and down side of tube so as not to break gradient surface.
7. Centrifuge gradients 13 min in Sorvall H-1000B rotor at 3000 rpm ($\sim 1900 \times g$), 4°C. The brake must be off.
8. Aspirate HBSS and 50% Percoll layer; then, with a pipet, separately collect cells at each interface between remaining layers and place into 50-ml centrifuge tubes.

Cells at the 65%/70% interface are the small, resting B cells, whereas those at the 50%/60% interface are the large, activated B cells. Cells at the intermediate 60%/65% interface are mainly small, resting B cells; these cells can be used as resting cells if absolutely necessary.

9. Wash cells 3 times with HBSS, centrifuging at 1200 to 1500 rpm ($300 \times g$) each time.
10. Count cells after final wash.

REAGENTS AND SOLUTIONS

ACK lysing buffer

8.29 g NH₄Cl (0.15 M)
1 g KHCO₃ (10.0 mM)
37.2 mg Na₂EDTA (0.1 mM)
Add 800 ml H₂O
Adjust pH to 7.2 to 7.4 with 1 N HCl
Add H₂O to 1 liter
Filter sterilize through a 0.2- μ m filter
Store at room temperature

Antigens

The following antigens can be used for in vitro antibody production (Basic Protocol 1). Note that when used in conjunction with nonimmunized, purified B cells, these antigens must be used with lymphokines.

Sheep red blood cells (SRBC): T cell–dependent, particulate antigen which serves as a relatively potent in vitro immunogen; does not require the use of cells from in vivo primed animals (for IgM response). Several batches may need to be screened to obtain cells that stimulate reproducible responses; once found, use that source exclusively. This antigen results in nonspecific B cell response to various epitopes. SRBC can be purchased from Becton Dickinson (#11943).

Ovalbumin (OVA): Soluble protein that is a relatively weak immunogen and requires the use of cells from in vivo primed animals. Elicits nonspecific B cell response; however, trinitrophenylation of the protein (TNP-OVA) results in production of antibodies to the specific epitope. However, in this case, T cells must be primed to the carrier molecule (OVA), and B cells must be primed to the hapten molecule (TNP). TNP-OVA can be prepared by the method of Eisen et al. (1953).

Trinitrophenylated Ficoll and dextran (TNP-Ficoll and TNP-dextran): “Type 2,” high-molecular-weight multiepitope antigens that have no intrinsic polyclonal activating properties and stimulate responses only in immunologically mature B cells (CBA/N mice do not respond, but nude mice do). Use of these antigens requires the addition of T cell-derived lymphokines (IL-1 and IL-2; alternatively, IL-5) or the use of B cells partially activated in vivo (see commentary discussion of T cell-regulated antigens). TNP-Ficoll can be purchased from Biosearch or prepared by the method of Inman (1975).

Trinitrophenylated Brucella abortus and lipopolysaccharide (TNP-BA and LPS): “Type 1” antigens that have intrinsic polyclonal activating properties and can stimulate responses in cell from all mouse strains (immature neonatal B cells as well as B cells from *xid* immune-defective mice). Use of these T cell–regulated antigens requires addition of lymphokines (IL-2 only) or use of B cells partially activated in vivo (see commentary). Killed *B. abortus* is available from the U.S. Department of Agriculture (National Veterinary Services Labs, P.O. Box 844, Ames, IA 50010) and is trinitrophenylated as described by Mond et al. (1978). LPS is available from Difco as *E. coli* 0111:B4.

Barbital buffer, modified (MBB), 5 \times stock solution

2.875 g barbituric acid
1.875 g sodium barbital
0.1103 g CaCl₂·2H₂O
0.238 g MgCl₂·6H₂O
42.5 g NaCl
H₂O to 1 liter
Dilute 1:5 for working solution

Cacodylate buffer (0.28 M)

Dissolve 24 g cacodylic acid in deionized water (~300 ml). Adjust to pH 6.9 with concentrated HCl. Dilute to 400 ml. Store at room temperature ≤6 months.

Chromic chloride solution (2.5×10^{-4} M)

For 200× stock, dissolve 134 mg chromic chloride in 10 ml saline (not buffered). Mix well. Dilute to 1× with saline just before use.

DMEM-10 medium, supplemented, complete

Prepare in *modified* Dulbeccos minimal essential medium (DMEM):

10% fetal bovine serum (FBS; APPENDIX 3F)

1 mM sodium pyruvate

2 mM L-glutamine

50 µg/ml gentamycin

0.1 mM nonessential amino acids

0.3% sodium bicarbonate

25 mM HEPES

50 µM 2-mercaptoethanol (2-ME)

Modified DMEM contains L-glutamine, 1000 mg glucose/liter, and phenol red but does not contain sodium bicarbonate. L-glutamine is labile and is therefore added fresh. The medium can be stored up to 1 month at 4°C.

Fetal bovine serum (FBS) is a critical element in obtaining good in vitro responses for any protocols in this unit. It is not unusual to have to screen 10 to 20 different batches of FBS to obtain a batch that supports high specific responses but low background responses. Reliable sources of FBS are Hyclone and Armour.

Guinea pig serum

To use as a source of complement, preabsorb with TNP-SRBC (see Support Protocol 3) 30 min on ice using 1 ml packed cells and 5 ml serum. Complement may be refrozen and stored at -70°C. Commercial sources such as Life Technologies have worked well in the PFC assays.

Complement absorbed with TNP-SRBC can also be used for unmodified SRBC. Complement does not have to be absorbed for the Jerne-Nordin assay (Basic Protocol 3), as the agarose absorbs cytotoxic activity during the assay.

Percoll mix solution

45 ml 10× phosphate-buffered saline (PBS; APPENDIX 2)

3 ml 0.6 N HCl (1:20 dilution of 12 N HCl; J.T. Baker)

132 ml H₂O

Check that pH is 7.0 to 7.2

Filter sterilize using 0.22-µm filter; store at 4°C

RPMI-10 medium, supplemented, complete

Prepare in complete RPMI/10% FBS (APPENDIX 3F):

25 mM HEPES

1 mM sodium pyruvate

50 µg/ml gentamycin

0.25 µg/ml Fungizone (this supplementation includes 100 penicillin and 100 µg/ml streptomycin; Whittaker #17-745A)

See note regarding testing of FBS in recipe above for supplemented DMEM-10.

COMMENTARY

Background Information

In vitro antibody production by B cells following stimulation with T cell–dependent antigens (i.e., antigens that activate B cells only if antigen-specific T cells are present to provide essential help) involves a number of discrete cellular events. Initially, antigen is taken up by the B cell as a result of binding to and cross-linkage of specific Ig receptors on the cell surface; this is followed by expression of the processed antigen on the cell surface in the context of MHC class II antigens. Next, there is B cell presentation of the processed antigen to the T cell, which leads to interactions between the T cell receptor and the antigen–MHC class II complex as well as between adhesion antigens on the T cell and B cell. These cell-surface events lead to stimulation of both interacting cells—the T cells producing growth and differentiation factors (lymphokines) and the B cells differentiating into antibody-producing plasma cells. Macrophages also play a role in this process as antigen-presenting cells that interact with T cells early in the process to bring about T cell activation and lymphokine production; such interaction may be critical since B cells require T cell–derived cytokines for optimum activation and differentiation. It should be noted that a given clone of B cells recognizes one determinant (usually called the haptenic determinant) on the antigen via the Ig receptor, whereas the T cell usually recognizes another determinant (usually called the carrier determinant) expressed on the processed antigen.

B cells may also be activated and induced to secrete immunoglobulin in the absence of antigen-specific T cells by cross-linkage of the Ig receptor by antigens containing multiple copies of the same determinant (multivalent antigens). These so-called T cell–independent antigens (such as TNP-Ficoll or TNP-dextran) were originally felt to be capable of activating B cells in the complete absence of T cells; however, it is now known that even these antigens require the help of T cell-derived lymphokines to induce antibody production by resting B cells. It is therefore more appropriate to refer to this group of antigens as T cell–regulated antigens. These antigens can, however, induce B cells that have been partially activated in vivo to respond even in the absence of cytokine-mediated help.

The culture system for measuring in vitro antibody production described in this unit is essentially a miniaturized version of the system

originally developed by Mishell and Dutton (1966, 1967). Much early information about T cell and B cell function was derived using this system in conjunction with mouse spleen cells as a source of T and B cells and heterologous red blood cells as a source of antigen. One advantage of using SRBC to study in vitro T and B cell responses is that this antigen induces responses in cells obtained from unprimed animals, probably because animals are naturally primed in vivo with antigens that cross-react with those associated with SRBC. However, under the usual in vitro conditions, the response of unprimed cells consists mainly of IgM antibody production; thus, either priming of the T cell/B cell source or addition of lymphokines to the culture system is necessary if IgG responses are desired. One disadvantage of using SRBC (and other complex antigens) is that a large number of B cell clones with poorly defined anti-RBC specificities are activated during the response. To study more restricted responses (of individual cell clones), it is necessary to use purified peptide antigens as the immunogen; however, in this case the response is poor with regard to levels of antibody produced unless T cells from primed animals are used or lymphokines are added to the cultures.

Critical Parameters

Stimulation of antibody production in vitro can lead to inconsistent and, at times, unreliable results. It is recommended that a laboratory starting this technology obtain reagents and cells from an experienced laboratory group. The novice should also consult the original literature cited in this unit to gain a fuller sense of the many parameters affecting in vitro response, as well as review the strategic planning section at the beginning of this unit.

Responder cells. B cells from pathogen-free mice should be used. B cells obtained from mice that are chronically infected with endemic viruses show poor in vitro responses. Healthy, uninfected mice are available from Jackson Laboratory (*APPENDIX 4*) and the National Cancer Institute (see Silverman and La Via, 1976). A purer and more homogeneous population of B lymphocytes permits a more reliable interpretation of the results. However, as mentioned in strategic planning, many culture systems require the presence of macrophages. Finally, defined T cell populations should be utilized. In this regard, the use of T cell populations defined by various surface antigens, organs or

origin—as well as other parameters or clonal/hybridoma T cell populations—should be considered.

Media components. Fetal calf serum of different sources and batches vary in their ability to support *in vitro* B cell responses (Mishell and Dutton, 1967; Shiigi and Mishell, 1975). It is best to test as many different types of FCS as possible to find one that gives a low background response and high specific response. When different sources of FCS are to be tested, be sure that the supplier has available >20 to 30 liters of a particular lot, as it is not worth the time and effort to test a batch containing only a few liters. 2-mercaptoethanol is essential for most *in vitro* murine B cell responses. The 2-ME stock should be replaced every 2 months.

RBC stimulator cells. For use as an antigen in *in vitro* B cell systems and in PFC assays, the SRBC should be as fresh as possible. For direct anti-SRBC responses, they should be <2 months old, and different batches should be tested for efficacy of response (McCarthy and Dutton, 1975a,b). If anti-TNP or α -protein A responses are being examined, the cells should be <1 week old before conjugation, because SRBC become increasingly fragile with age. When older SRBC are conjugated, they lyse very easily, making it difficult to enumerate plaques.

Incubator. The quality of the incubator is very important. It must be capable of maintaining temperature, humidity, and percent CO₂ extremely accurately.

Cunningham-Szenberg assay. The guinea pig serum must be absorbed properly as described in reagents and solutions. In addition, the complement should be used within 20 min of thawing (during the assay). All reagents should be at room temperature before they are added to the chambers, because if they are cold, air bubbles form in the chambers during incubation, making plaques difficult to count. Chambers should be sealed with wax within 20 min of filling to prevent the chamber from drying out.

Jerne-Nordin assay. It is essential that the agarose does not cool below 42°C before it is poured onto the slides. After the slides are poured, the agarose must be solidified by allowing to cool to room temperature before the slides are inverted and put on the racks.

Antigen removal. When assaying supernatants for specific antibody responses, it is important to wash the cells to remove the antigen prior to collecting supernatants for the ELISA. The timing of antigen removal and

collection of supernatants will depend on the system being employed. In general, antigen can be removed 3 to 6 days after the onset of culture, and supernatant can be collected after 4 to 7 days. In studying polyclonal immunoglobulin secretion induced by lipopolysaccharide or lymphokine, there is no need to remove these stimuli, as they do not interfere with the ELISA.

Manipulations of cells. Most importantly, when working with cells *in vitro* one should always be aware that any manipulation, however minor, may influence the responsiveness of the cell and may lead to variability in the results. To minimize these occurrences, attention should be paid to treating cells with care—e.g., not centrifuging at speeds greater than 1000 rpm, not vortexing for prolonged periods of time, and maintaining cells on ice in medium that is appropriately buffered and supplemented with FBS.

Anticipated Results

For antibody secretion, the number of PFC generated will depend on the nature of the stimulus and/or the presence of added lymphokines. However, typical responses for 10⁶ purified B cells cultured with stimulatory lymphokines are presented in Table 11.13.1.

Time Considerations

Preparing purified resting B cells by Percoll gradient centrifugation (Support Protocol 5) takes ~4 hr, while preparing antigen and distributing the cells into the plates normally takes ~2 hr. The exact length of time will depend on the size and complexity of the experiment and the skill of the investigator.

Supernatants for ELISA—washing the plates to remove excess antigen takes 0.5 hr; harvesting supernatants for assay takes 0.5 hr (the ELISA takes ~7 hr).

Preparation of modified SRBC—TNP-SRBC and Protein A-SRBC both take ~3 hr.

Jerne-Nordin assay—precoating the slides takes ~15 min; the maximum size of an experiment is three 96-well plates, which would take an experienced worker 8 hr.

Cunningham-Szenberg assay—preparation of slides takes ~1 hr per 120 slides; absorption of complement takes ~2 hr; an experienced worker can assay six 96-well plates in 8 hr.

Literature Cited

Cunningham, A.J. and Szenberg, A. 1968. Further improvements in the plaque technique for detecting single antibody forming cells. *Immunology* 14:599-600.

- Eisen, H., Belman, S., and Carsten, M. 1953. The reaction of 2,4-dinitrobenzene sulfonic acid with free amino groups of proteins. *J. Am. Chem. Soc.* 75:4583-4591.
- Gronowicz, E., Coutinho, A., and Melchers, F. 1976. A plaque assay for all cells secreting Ig of a given type or class. *Eur. J. Immunol.* 6:588-590.
- Hathcock, K. 1991a. T cell enrichment by cytotoxic elimination of B cells and accessory T cells. In *Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.), pp. 3.3.1-3.3.5. John Wiley & Sons, New York.
- Hathcock, K. 1991b. T cell depletion by cytotoxic elimination. In *Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.), pp. 3.4.1-3.4.3. John Wiley & Sons, New York.
- Inman, J. K., 1975. Thymus-independent antigens: The preparation of covalent, hapten-Ficoll conjugates. *J. Immunol.* 114:704-709.
- Jerne, N.K. and Nordin, A.A. 1963. Plaque formation in agar by single antibody producing cells. *Science* 140:405-408.
- Kruisbeek, A.M. and Shevach, E. 1991. Proliferative assays for T cell function. In *Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.), pp. 3.12.1-3.12.14. John Wiley & Sons, New York.
- Lycke, N.Y. and Coico, R. 1996. Measurement of immunoglobulin synthesis using the ELISPOT assay. In *Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.), pp. 7.14.1-7.14.9. John Wiley & Sons, New York.
- Magge, M.G. 1993. Fractionation of T cells and B cells. In *Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.), pp. 3.5.1-3.5.6. John Wiley & Sons, New York.
- McCarthy, M.M. and Dutton, R.W. 1975a. The humoral response of mouse spleen cells to two types of sheep erythrocytes. I. Genetic control of the response to H and L SRBC. *J. Immunol.* 115:1316-1321.
- McCarthy, M.M. and Dutton, R.W. 1975b. The humoral response of mouse spleen cells to two types of sheep erythrocytes. II. Evidence for gene expression in the B lymphocyte. *J. Immunol.* 115: 1322-1329.
- Mishell, R.I. and Dutton, R.W. 1966. Immunization of normal mouse spleen cell preparations in vitro. *Science* 153:1004-1006.
- Mishell, R.I. and Dutton, R.W. 1967. Immunization of dissociated spleen cell cultures from mice. *J. Exp. Med.* 126:423-442.
- Mitchison, N.A. 1971. The carrier effect in the secondary response to hapten-protein conjugates. II. Cellular cooperation. *Eur. J. Immunol.* 1:18-27.
- Mond, J.J., Scher, I., Mozier, D.E., Blacsy, M., and Paul, W.E. 1978. T independent responses in B cell defective CBA/N mice to *Brucella abortus* and to trinitrophenyl (TNP) conjugates of *Brucella abortus*. *Eur. J. Immunol.* 8:459-463.
- Mond, J.J., Farrar, W.E., Paul W.E., Fuller-Farrar, J., Schaeffer, M., and Howard, M., 1983. T cell dependence and factor reconstitution of in vitro antibody responses to TNP-*B. abortus* and restoration of depleted responses with chromatographed fractions of a T cell derived factor. *J. Immunol.* 131:633-637.
- Rittenberg, M.B. and Pratt, C. 1969. Anti-trinitrophenyl (TNP) plaque assay. Primary response of BALB/c mice to soluble and particulate immunogen. *Proc. Soc. Exp. Biol. Med.* 132:575.
- Shiigi, S.M. and Mishell, R.I. 1975. Sera and the in vitro induction of immune responses. I. Bacterial contamination and the generation of good fetal bovine sera. *J. Immunol.* 115:741-744.
- Silverman, M.S. and LaVia, M.F. 1976. Letters to the editor, *J. Immunol.* 117: 2270-2279.
- Thompson, C-B., Scher, I., Schaeffer, M., Lindsten, T., Finkelman F.D. and Mond, J.J. 1984. Size dependent B lymphocyte subpopulations: Relationship of cell volume to surface phenotype, cell cycle, proliferative response and requirement for antibody production to TNP-Ficoll and TNP-BA. *J. Immunol.* 133:2333-2342.

Contributed by James J. Mond
and Mark Brunswick
Uniformed Services University of
the Health Sciences
Bethesda, Maryland

Purification of Immunoglobulin G Fraction from Antiserum, Ascites Fluid, or Hybridoma Supernatant

UNIT 11.14

This unit describes the isolation of the immunoglobulin G (IgG) fraction (containing antibodies of all specificities) from a complex protein mixture such as antiserum, ascites fluid, or hybridoma supernatant. The Basic Protocol utilizes saturated ammonium sulfate solution to precipitate the IgG fraction, while the Alternate Protocol describes fractionation of IgG by chromatography on DEAE–Affi-Gel Blue resin. The purification of IgG by affinity chromatography utilizing staphylococcal protein A or antigen-Sepharose is described in *UNIT 11.11*.

PRECIPITATION OF IgG WITH SATURATED AMMONIUM SULFATE

**BASIC
PROTOCOL**

Antiserum or ascites fluid is adjusted to 33% with respect to the concentration of saturated ammonium sulfate, resulting in the precipitation of the IgG. At this concentration of saturated ammonium sulfate, a large percentage—but not all—of the contaminating protein species present in the antiserum or ascites remains in solution.

Materials

- Saturated ammonium sulfate (SAS) solution (see recipe)
- 33% SAS solution (see recipe)
- Immunoglobulin-containing antiserum, ascites, or tissue culture supernatant
- Additional reagents and equipment for dialysis (*APPENDIX 3C*)

1. Add 1 vol SAS solution dropwise with a Pasteur pipet to 2 vol antiserum, ascites, or tissue culture supernatant with constant mixing at 4°C.
2. Allow precipitate to form over a period of 2 to 4 hr at 4°C with constant mixing. Pellet precipitate by centrifugation for 20 min at 12,000 × *g*, 4°C.
3. Wash pellet by resuspending (vortexing) it in a volume of cold 33% SAS solution equivalent to the original volume of antiserum, ascites, or tissue culture supernatant. Repeat wash step once.
4. Dissolve pellet in appropriate cold buffer by gentle vortexing.

A convenient volume for solubilizing the IgG fraction is 5% to 10% of the original antiserum, ascites, or hybridoma supernatant volume.

5. Dialyze IgG solution over 48 hr at 4°C against three changes of the desired buffer (4 liters per change) to fully remove the ammonium sulfate.

See APPENDIX 3C for preparing dialysis membrane (molecular weight cutoff 12,000 to 14,000; Spectrapor 2 or equivalent) and for large-volume dialysis.

Immunology

11.14.1

FRACTIONATION OF IgG BY CHROMATOGRAPHY ON DEAE-AFFI-GEL BLUE

This alternate procedure can be used to isolate the IgG fraction from either ascites or antiserum or to further purify the IgG fraction obtained after SAS precipitation. Under low salt concentrations, many protein species present in ascites or antiserum, including IgG, are bound to DEAE-Affi-Gel Blue resin with varying affinities. When increasing the salt concentration in small increments, the IgG fraction is selectively eluted, leaving the majority of the contaminating protein species still bound to the resin. This protocol describes the fractionation of murine IgG.

Additional Materials (also see Basic Protocol)

Loading buffer (see recipe)
Elution buffer (see recipe)
NaN₃, crystalline form
DEAE-Affi-Gel Blue (Bio-Rad)

Additional reagents and equipment for gel-filtration chromatography (UNIT 10.9)
and SDS-polyacrylamide gel electrophoresis (UNIT 10.2)

1. Prepare a column of DEAE-Affi-Gel Blue according to manufacturer's instructions (see Fig. 10.9.1) and equilibrate with five bed volumes of loading buffer.

It is recommended that a total bed volume of 7 ml/ml antiserum or ascites be used.

This and subsequent steps should be carried out at 4°C.

2. Dialyze IgG-containing sample against two changes of loading buffer (4 liters per change) for ~40 hr at 4°C.
3. Apply sample to the column at a flow rate of 10 ml/hr. Elute unbound protein with three bed volumes of loading buffer at the same flow rate.

Alternatively, elution of the unbound material may be carried out overnight; however, at least three bed volumes of loading buffer must be run down the column before the specific elution of the IgG fraction.

4. Elute bound IgG fraction with elution buffer at a flow rate of 10 ml/hr. Collect 10-ml fractions and store at 4°C until they are pooled.

Transferrin (M_r76,000) usually coelutes with the IgG fraction.

5. Identify fractions containing IgG by analyzing 30 to 50 µl of each fraction on a 10% SDS-polyacrylamide gel (UNIT 10.2), under reducing conditions (see UNIT 10.2).

Under reducing conditions, the IgG heavy and light chains will run separately with molecular weights of 50,000 and 25,000, respectively.

The amount of transferrin contaminating the IgG will be determined at this step. The amount present will vary depending on the source of the antibody. It can be removed by gel-filtration chromatography (UNIT 10.9).

6. Pool fractions containing IgG and dialyze over 40 hr at 4°C against two changes of the desired buffer (4 liters per change).
7. Re-equilibrate DEAE-Affi-Gel Blue column with five bed volumes of loading buffer, add a few crystals of NaN₃ to retard bacterial growth, and store at 4°C.

REAGENTS AND SOLUTIONS

Elution buffer (20 mM Tris·Cl/50 mM NaCl, pH 8.0)

2.4 g Tris base
2.92 g NaCl
900 ml H₂O
Adjust to pH 8.0 with concentrated HCl
H₂O to 1 liter

Loading buffer (20 mM Tris·Cl/30 mM NaCl, pH 8.0)

2.4 g Tris base
1.75 g NaCl
900 ml H₂O
Adjust to pH 8.0 with concentrated HCl
H₂O to 1 liter

Saturated and 33% saturated ammonium sulfate (SAS) solution

450 g ammonium sulfate
H₂O to 500 ml

Heat solution on a heated stirring plate until ammonium sulfate dissolves completely. Filter solution while it is still warm and then allow it to cool. Upon cooling, crystals will form and should not be removed. Check pH of cooled solution; adjust to pH 7.5 with ammonium hydroxide. For 33% SAS solution, mix 33 ml SAS solution with 67 ml phosphate-buffered saline (PBS; APPENDIX 2).

A quantity of 760 g of ammonium sulfate is required per 1000 ml for a 100% saturated solution. A total of 450 g/500 ml is used to ensure sufficient crystal formation to maintain saturation. A slightly different recipe for saturated ammonium sulfate solution in Tris buffer is contained in UNIT 11.1 and may be substituted in this protocol.

COMMENTARY

Background Information

Affinity chromatography, in which the relevant protein antigen or peptide antigen is coupled to an affinity resin (see UNIT 11.11), is the method of choice for the purification of specific antibody. However, under certain circumstances, such purification procedures are either unnecessary or not applicable (e.g., when the antigen is not available in adequate quantities). In this unit two procedures are described for the isolation of the IgG fraction from ascites or antiserum. In both cases, however, the total IgG component is isolated.

The precipitation of protein using ammonium sulfate is achieved by dehydration in the microenvironment of the protein molecule. In solution, a large number of water molecules are bound to the sulfate ion (SO₄⁼), significantly reducing the amount of water available to interact with the protein molecules. At a particular concentration of ammonium sulfate an insufficient quantity of unbound water will remain to keep a given protein species in solution, resulting in the precipitation of that protein. The exact concentration of ammonium sulfate re-

quired to precipitate a protein will be determined by the characteristics of the amino acid side chains exposed to the solvent. Precipitation of the IgG fraction of antiserum or ascites fluid at an ammonium sulfate concentration of 33% efficiently isolates the IgG fraction from major contaminating proteins such as albumin and hemoglobin, yielding an IgG preparation that is relatively pure. However, a significant number of other proteins will coprecipitate. Therefore, this fractionation protocol serves as an excellent initial step in the purification of IgG. The SAS-precipitated IgG fraction can then be further purified by immunoaffinity chromatography or on a DEAE–Affi-Gel Blue column (see Alternate Protocol). In addition, the IgG fraction can be stored for a long period of time (1 to 2 years) as a suspension of precipitate in SAS solution without loss of antibody-binding activity. Fresh antiserum or ascites can be brought to the desired ammonium sulfate concentration and safely stored until further purification of the IgG fraction is required.

Affinity chromatography using DEAE–Affi-Gel Blue is a combination of dye-interac-

tion and ion-exchange (IEX) chromatography. The DEAE–Affi-Gel Blue is a bifunctional resin bearing both diethylaminoethyl groups and Cibacron F3GA molecules. The dye-interaction chromatographic function is based on the ability of the dye, Cibacron blue F3GA, to bind a large variety of enzymes and other proteins with a range of affinities. A comprehensive discussion on the nature and use of this dye, when attached to an affinity chromatography matrix, is given in the technical literature found on the Bio-Rad website (<http://www.bio-rad.com>). Blue Sepharose 6 Fast Flow is the equivalent matrix produced by Amersham Pharmacia Biotech (<http://www.apbiotech.com>). In addition, the IEX moieties on the resin interact with the remaining acidic proteins in the antiserum or ascites. Overall, a large percentage (70% to 80%) of contaminating proteins are separated from the IgG fraction under the elution conditions described. The chromatographic conditions given in the Basic Protocol for the selective elution of the immunoglobulin fraction from the DEAE–Affi-Gel Blue column have been optimized for murine IgG. The manufacturers recommend that 0.02 M K_2HPO_4 (adjusted to pH 8.0 with KOH) buffer be used when fractionating the IgG component from human antiserum. For isolating the IgG fraction from rabbit antiserum, it is recommended that 0.02 M Tris-Cl, pH 8.0/0.03 M NaCl (i.e., the loading buffer used in the Alternate Protocol) be used. It is stated that the human and rabbit IgG fractions do not bind to the column when using these respective buffers.

Pure preparations of immunoglobulins of the G and M isotypes can be achieved by a variety of other chromatographic procedures, including gel filtration and IEX. These protocols are described in this manual (*UNITS 10.9 & 10.10*). The application of electrophoresis to the isolation of pure immunoglobulin species is also described in the former volume.

Critical Parameters

The addition of SAS solution to the antiserum or ascites must be carried out slowly. High local concentrations of ammonium sulfate will result in the coprecipitation of other contaminating protein species that usually precipitate at SAS concentrations >33%. The precipitation should be carried out between 4° and 25°C. Outside this temperature range, the degree of saturation of ammonium sulfate deviates by >3%. Also, to maintain total saturation of the SAS solution, ensure that a significant quantity

of ammonium sulfate crystals are always present at the bottom of the saturated solution.

The optimal elution conditions isolating murine IgG using DEAE–Affi-Gel Blue have been described. As stated in background information, the optimal conditions for fractionating IgG from other species will vary. It is also possible that individual monoclonal antibodies may elute under slightly different salt concentrations than those given in the Basic Protocol. Therefore, the buffer system should be tested by running a small volume of the antibody containing sample down a mini-column of DEAE–Affi-Gel Blue resin before committing a larger quantity of sample to the column.

Troubleshooting

In general, immunoglobulins are a very homogeneous family of proteins and will therefore precipitate when the SAS solution is at a concentration of 33%. In some instances, however, precipitation may be only partial or may not occur at all under these conditions. This problem is more commonly observed when attempting to precipitate an individual monoclonal antibody from ascites fluid. Under such circumstances, it is necessary to increase the ammonium sulfate saturation to 40% or 50%. It must be noted that the number of contaminating protein species will increase with increasing ammonium sulfate concentration. If the IgG fraction is not selectively eluted from the DEAE–Affi-Gel Blue column under the conditions described in the basic protocol and in background information, it may be necessary to adjust the NaCl concentration in the elution buffer. However, do not increase the NaCl concentration above 75 mM, as contaminating protein species begin to elute above this level of salt concentration.

Anticipated Results

An efficient precipitation or ion-exchange purification of the IgG fraction will be achieved using these procedures. However, in many cases, transferrin will coprecipitate or copurify and may be removed by gel-filtration chromatography (*UNIT 10.9*; IgG: MW 150,000; transferrin: MW 76,000). The preferred procedure for purifying antibodies from serum or ascites fluid is by immunoaffinity chromatography on a ligand-bearing affinity matrix or with protein A–Sepharose (*UNIT 11.11*).

Time Considerations

Precipitation of the IgG fraction using SAS can be completed within 4 to 6 hr. After allow-

ing 24 hr for dialysis, affinity chromatography on DEAE–Affi-Gel Blue can be completed over a 24-hr period. It is usually convenient to wash the unbound material through the column overnight, using the loading buffer.

Contributed by Helen M. Cooper
Walter and Eliza Hall Institute
Melbourne, Australia

Yvonne Paterson
University of Pennsylvania
Philadelphia, Pennsylvania

PREPARATION OF ANTIPEPTIDE ANTIBODIES

Introduction to Peptide Synthesis

DEVELOPMENT OF SOLID-PHASE PEPTIDE-SYNTHESIS METHODOLOGY

A number of synthetic peptides are significant commercial or pharmaceutical products, ranging from the dipeptide sugar substitute aspartame to clinically used hormones such as oxytocin, adrenocorticotrophic hormone, and calcitonin. Rapid, efficient, and reliable methodology for the chemical synthesis of these molecules is of utmost interest. The stepwise assembly of peptides from amino acid precursors has been described for nearly a century. The concept is a straightforward one, whereby peptide elongation proceeds via a coupling reaction between amino acids, followed by removal of a reversible protecting group. The first peptide synthesis, as well as the creation of the term “peptide,” was reported by Fischer and Fourneau (1901). Bergmann and Zervas (1932) created the first reversible N^α -protecting group for peptide synthesis, the carbobenzoxy (Cbz) group. DuVigneaud successfully applied early “classical” strategies to construct a peptide with oxytocin-like activity (duVigneaud et al., 1953). Classical, or solution-phase methods for peptide synthesis have an elegant history and have been well chronicled. Solution synthesis continues to be especially valuable for large-scale manufacturing and for specialized laboratory applications.

Peptide synthesis became a more practical part of present-day scientific research following the advent of solid-phase techniques. The concept of solid-phase peptide synthesis (SPPS) is to retain chemistry that has been proven in solution but to add a covalent attachment step that links the nascent peptide chain to an insoluble polymeric support (resin). Subsequently, the anchored peptide is extended by a series of addition cycles (Fig. 11.15.1). It is the essence of the solid-phase approach that reactions are driven to completion by the use of excess soluble reagents, which can be removed by simple filtration and washing without manipulative losses. Once chain elongation has been completed, the crude peptide is released from the support.

In the early 1960s, Merrifield proposed the use of a polystyrene-based solid support for peptide synthesis. Peptides could be assembled stepwise from the C to N terminus using N^α -protected amino acids. SPPS of a tetrapeptide was achieved by using Cbz as an α -amino-protecting group, coupling with N,N' -dicyclohexylcarbodiimide (DCC), and liberating the peptide from the support by saponification or by use of HBr (Merrifield, 1963). SPPS was later modified to use the *t*-butyloxycarbonyl (Boc) group for N^α protection (Merrifield, 1967) and hydrogen fluoride (HF) as the reagent for removal of the peptide from the resin (Sakakibara et al., 1967). SPPS was thus based on “relative acidolysis,” where the N^α -protecting group (Boc) was labile in the presence of moderate acid (trifluoroacetic acid; TFA), while side-chain-protecting benzyl (Bzl)-based groups and the peptide/resin linkage were stable in the presence of moderate acid and labile in the presence of strong acid (HF). The first instrument for automated synthesis of peptides, based on Boc SPPS, was built by Merrifield, Stewart, and Jernberg (Merrifield et al., 1966). From the 1960s through the 1980s, Boc-based SPPS was fine-tuned (Merrifield, 1986). This strategy has been utilized for synthesis of proteins such as interleukin-3 and active enzymes including ribonuclease A and all-L and all-D forms of HIV-1 aspartyl protease.

In 1972, Carpino introduced the 9-fluorenylmethoxycarbonyl (Fmoc) group for N^α protection (Carpino and Han, 1972). The Fmoc group requires moderate base for removal, and thus offered a chemically mild alternative to the acid-labile Boc group. In the late 1970s, the Fmoc group was adopted for solid-phase applications. Fmoc-based strategies utilized *t*-butyl (*t*Bu)-based side-chain protection and hydroxymethylphenoxy-based linkers for peptide attachment to the resin. This was thus an “orthogonal” scheme requiring base for removal of the N^α -protecting group and acid for removal of the side-chain protecting groups and liberation of the peptide from the resin. The milder conditions of Fmoc chemistry as compared to Boc chemistry—which include elimination of repetitive moderate acidolysis steps and the

Contributed by Gregg B. Fields

Current Protocols in Molecular Biology (2002) 11.15.1-11.15.9

Copyright © 2003 by John Wiley & Sons, Inc.

final strong acidolysis step—were envisioned as being more compatible with the synthesis of peptides that are susceptible to acid-catalyzed side reactions. In particular, the modification of the indole ring of Trp was viewed as a particular problem during Boc-based peptide synthesis (Barany and Merrifield, 1979), which could be alleviated using Fmoc chemistry. One example of the potential advantage of Fmoc chemistry for the synthesis of multiple-Trp-containing peptides was in the synthesis of gramicidin A. Gramicidin A, a pentadecapeptide containing four Trp residues, had been synthesized previously in low yields (5% to 24%) using Boc chemistry. The mild conditions of Fmoc chem-

istry dramatically improved the yields of gramicidin A, in some cases up to 87% (Fields et al., 1989, 1990). A second multiple-Trp-containing peptide, indolicidin, was successfully assembled in high yield by Fmoc chemistry (King et al., 1990). Thus, the mild conditions of Fmoc chemistry appeared to be advantageous for certain peptides, as compared with Boc chemistry.

One of the subsequent challenges for practitioners of Fmoc chemistry was to refine the technique to allow for construction of proteins, in similar fashion to that which had been achieved with Boc chemistry. Fmoc chemistry had its own set of unique problems, including suboptimum solvation of the peptide/resin,

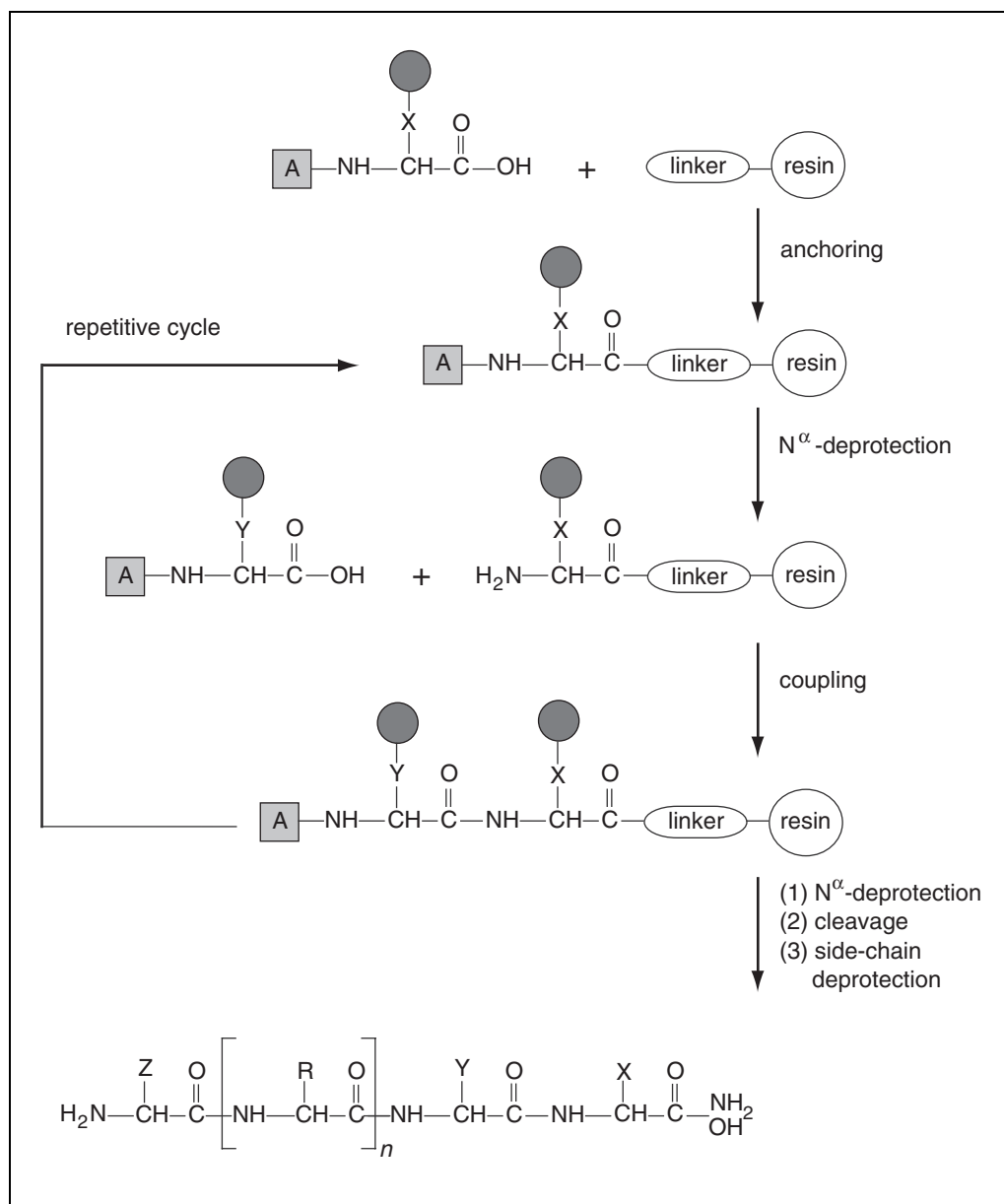


Figure 11.15.1 Generalized approach to solid-phase peptide synthesis. Symbols: A, N^α -protecting group; circle, side-chain protecting groups; R, X, Y, and Z, side-chain functionalities.

slow coupling kinetics, and base-catalyzed side reactions. Improvements in these areas of Fmoc chemistry (Atherton and Sheppard, 1987; Fields and Noble, 1990; Fields et al., 2001) allowed for the synthesis of proteins such as bovine pancreatic trypsin inhibitor analogs, ubiquitin, yeast actin-binding protein 539-588, human β -chorionic gonadotropin 1-74, mini-collagens, HIV-1 Tat protein, HIV-1 nucleocapsid protein Ncp7, and active HIV-1 protease.

The milder conditions of Fmoc chemistry, along with improvements in the basic chemistry, have led to a shift in the chemistry employed by peptide laboratories. This trend is best exemplified by a series of studies (Angeletti et al., 1997) carried out by the Peptide Synthesis Research Committee (PSRC) of the Association of Biomolecular Resource Facilities (ABRF). The PSRC was formed to evaluate the quality of the synthetic methods utilized in its member laboratories for peptide synthesis. The PSRC designed a series of studies from 1991 to 1996 to examine synthetic methods and analytical techniques. A strong shift in the chemistry utilized in core facilities was observed during this time period—i.e., the more senior Boc methodology was replaced by Fmoc chemistry. For example, in 1991 50% of the participating laboratories used Fmoc chemistry, while 50% used Boc-based methods. By 1994, 98% of participating laboratories were using Fmoc chemistry. This percentage remained constant in 1995 and 1996. In addition, the overall quality of the peptides synthesized improved greatly from 1991 to 1994. Possible reasons for the improved results were any combination of the following (Angeletti et al., 1997):

1. The greater percentage of peptides synthesized by Fmoc chemistry, where cleavage conditions are less harsh.
2. The use of different side-chain protecting group strategies that help reduce side reactions during cleavage.
3. The use of cleavage protocols designed to minimize side reactions.
4. More rigor and care in laboratory techniques.

The present level of refinement of solid-phase methodology has led to numerous, commercially available instruments for peptide synthesis (Table 11.15.1).

The next step in the development of solid-phase techniques includes applications for peptides containing non-native amino acids, post-translationally modified amino acids, and pseudoamino acids, as well as for organic molecules in general. Several areas of solid-phase

synthesis need to be refined to allow for the successful construction of this next generation of biomolecules. The solid support must be versatile so that a great variety of solvents can be used, particularly for organic-molecule applications. Coupling reagents must be sufficiently rapid so that sterically hindered amino acids can be incorporated. Construction of peptides that contain amino acids bearing post-translational modifications should take advantage of the solid-phase approach. Finally, appropriate analytical techniques are needed to assure the proper composition of products.

THE SOLID SUPPORT

Effective solvation of the peptide/resin is perhaps the most crucial condition for efficient chain assembly during solid-phase synthesis. Swollen resin beads may be reacted and washed batch-wise with agitation, then filtered either with suction or under positive nitrogen pressure. Alternatively, they may be packed in columns and utilized in a continuous-flow mode by pumping reagents and solvents through the resin. ^1H , ^2H , ^{13}C , and ^{19}F nuclear magnetic resonance (NMR) experiments have shown that, under proper solvation conditions, the linear polystyrene chains of copoly(styrene-1%-divinylbenzene) resin (PS) are nearly as accessible to reagents as if free in solution. ^{13}C and ^{19}F NMR studies of Pepsyn (copolymerized dimethylacrylamide, *N,N'*-bisacryloylethylenediamine, and acryloylsarcosine methyl ester) have shown similar mobilities at resin-reactive sites as PS. Additional supports created by grafting polyethylene glycol (polyoxyethylene) onto PS—either by controlled anionic polymerization of ethylene oxide on tetraethylene glycol-PS (POE-PS) or by coupling N^{Boc} - or Fmoc-polyethylene glycol acid or -polyethylene glycol diacid to amino-functionalized PS (PEG-PS)—combine the advantages of liquid-phase synthesis (i.e., a homogeneous reaction environment) and solid-phase synthesis (an insoluble support). ^{13}C NMR measurements of POE-PS showed the polyoxyethylene chains to be more mobile than the PS matrix, with the highest T_1 spin-lattice relaxation times observed with POE of molecular weight 2000 to 3000. Other supports that have been developed that show improved solvation properties and/or are applicable to organic synthesis include polyethylene glycol polyacrylamide (PEGA), cross-linked acrylate ethoxylate resin (CLEAR), and augmented surface polyethylene prepared by chemical transformation (ASPECT). As the solid-phase

Table 11.15.1 Instruments for Solid-Phase Synthesis

Supplier ^a	Instrument model	Fmoc	Boc	Batch	Flow	Monitoring	Scale (mmol)	No. of peptides
<i>Peptide synthesis systems</i>								
Advanced ChemTech	90	Yes	Yes	Yes	No	No	0.1–12	1–2
	Apex 396	Yes	Yes	Yes	No	No	0.005–1	96
	348	Yes	Yes	Yes	No	No	0.005–0.15	48
	357	Yes	Yes	Yes	No	No	0.005–0.25	42
Bachem Bioscience	SP4000-LAB	Yes	Yes	Yes	No	No	0.25–5	1
	SP4000-PRO	Yes	No	Yes	No	No	5–50	1
Gilson/Abimed	AMS422	Yes	No	Yes	No	Yes	0.005–1	1
Perkin-Elmer	ABI433A	Yes	Yes	Yes	No	Yes	0.05–1	1
	Pioneer-MPS	Yes	No	No	Yes	Yes	0.005–0.1	16
Rainin	PS3	Yes	Yes	Yes	No	No	0.1–0.25	3
	Sonata/Pilot	Yes	Yes	Yes	No	No	0.1–50	1
	Symphony/Multiplex	Yes	No	Yes	No	No	0.005–0.35	12
CS Bio	CS100	Yes	Yes	Yes	No	No	0.05–1	1
	CS336	Yes	Yes	Yes	No	No	0.05–0.25	3
	036	Yes	Yes	Yes	No	No	0.1–2.5	1
	136	Yes	Yes	Yes	No	No	0.1–2.5	1
	CS536	Yes	Yes	Yes	No	No	0.2–25	1
	CS936S	Yes	Yes	Yes	No	No	≤600	1
	CS936	Yes	Yes	Yes	No	No	≤12,500	1
Intavis AG	AutoSpot	Yes	No	Yes	No	No	3–4 nmol/mm ²	384
<i>Multiple organic synthesis units</i>								
Advanced ChemTech	384	Yes	Yes	Yes	No	No	—	4 × 96
	Vantage	Yes	Yes	Yes	No	No	—	96

^aFor contact information, see APPENDIX 4.

method has expanded to include organic-molecule and library syntheses, the diversity of supports will enhance the efficiency of these new applications.

Successful syntheses of problematic sequences can be achieved by manipulation of the solid support. In general, the longer the synthesis, the more polar the peptide/resin will become (Sarin et al., 1980). One can alter the solvent environment and enhance coupling efficiencies by adding polar solvents and/or chaotropic agents (Fields and Fields, 1994). Also, using a lower substitution level of resin to avoid interchain crowding can improve the synthesis (Tam and Lu, 1995). During difficult syntheses, deprotection of the Fmoc group can proceed slowly. By spectrophotometrically monitoring deprotection as the synthesis proceeds, one can detect problems and extend base-deprotection times and/or alter solvation conditions as necessary.

COUPLING REAGENTS

The classical examples of in situ coupling reagents are *N,N'*-dicyclohexylcarbodiimide (DCC) and the related *N,N'*-diisopropylcarbodiimide (Rich and Singh, 1979). The generality of carbodiimide-mediated couplings is extended significantly by the use of either 1-hydroxybenzotriazole (HOBT) or 1-hydroxy-7-azabenzotriazole (HOAt) as an additive, either of which accelerates carbodiimide-mediated couplings, suppresses racemization, and inhibits dehydration of the carboxamide side chains of Asn and Gln to the corresponding nitriles. Protocols involving benzotriazol-1-yl-oxy-tris(dimethylamino)phosphonium hexafluorophosphate (BOP), benzotriazol-1-yl-oxy-tris(pyrrolidino)phosphonium hexafluorophosphate (PyBOP), 7-azabenzotriazol-1-yl-oxytris(pyrrolidino)phosphonium hexafluorophosphate (PyAOP), *O*-benzotriazol-1-yl-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HBTU), *O*-(7-azaben-

Table 11.15.2 Coupling Reagents and Additives Used in Solid-Phase Peptide Synthesis and Suppliers

Reagent	Abbreviation	Supplier(s) ^a
<i>N,N'</i> -dicyclohexylcarbodiimide	DCC	A, ACT, AO, CI, CN, F, PI, PL, Q, S
<i>N,N'</i> -diisopropylcarbodiimide	DIPCDI	A, ACT, AO, CI, F, PE, Q, S
<i>O</i> -benzotriazol-1-yl- <i>N,N,N',N'</i> -tetramethyluronium hexafluorophosphate	HBTU	A, ACT, AS, CI, CN, F, NS, PI, PL, Q, S
<i>O</i> -benzotriazol-1-yl- <i>N,N,N',N'</i> -tetramethyluronium tetrafluoroborate	TBTU	A, ACT, B, CI, CN, F, NS, PE, PI, PL, Q, S
<i>O</i> -(7-azabenzotriazol-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium hexafluorophosphate	HATU	PE
Benzotriazol-1-yl-oxy-tris(dimethylamino) phosphonium hexafluorophosphate	BOP	A, ACT, AO, B, CI, CN, F, NS, PL, PI, Q, S
Benzotriazol-1-yl-oxy-tris(pyrrolidino)-phosphonium hexafluorophosphate	PyBOP	A, ACT, AO, CI, CN, F, S
7-azabenzotriazole-1-yl-oxy-tris(pyrrolidino)-phosphonium hexafluorophosphate	PyAOP	PE
Tetramethylfluoroformamidinium hexafluorophosphate	TFFH	ACT, PE
1-hydroxybenzotriazole	HOBt	A, ACT, AO, AS, CI, CN, NS, PE, PI, Q, S
1-hydroxy-7-azabenzotriazole	HOAt	PE
<i>N,N</i> -diisopropylethylamine	DIEA	A, ACT, AO, CI, F, PE, Q, S
<i>N</i> -methylmorpholine	NMM	A, AO, CI, F, S

^aAbbreviations: A, Aldrich; ACT, Advanced ChemTech; AO, Acros Organics; AS, AnaSpec; B, Bachem; CI, Chem-Impex; CN, Calbiochem-Novabiochem; F, Fluka; NS, Neosystem/SNPE; PE, Perkin-Elmer; PI, Peptides International; PL, Peninsula Laboratories; Q, Quantum Biotechnologies; S, Sigma. For contact information, see APPENDIX 4.

zotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HATU), and *O*-benzotriazol-1-yl-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU) result in coupling kinetics even more rapid than that obtained with carbodiimides. Amino acid halides have also been applied to solid-phase peptide synthesis (SPPS). *N*^α-protected amino acid chlorides have a long history of use in solution synthesis. Fmoc-amino acid chlorides and fluorides react rapidly under SPPS conditions in the presence of HOBt/*N,N*-diisopropylethylamine (DIEA) and DIEA, respectively, with very low levels of racemization. For convenience, tetramethylfluoroformamidinium hexafluorophosphate (TFFH) can be used for automated preparation of Fmoc-amino acid fluorides. Amino acid fluorides have been found to be especially useful for the preparation of peptides containing sterically hindered amino acids, such as peptaibols. All of the coupling reagents and additives discussed here are commercially available (see Table 11.15.2).

SYNTHESIS OF MODIFIED RESIDUES AND STRUCTURES

Peptides of biological interest often include structural elements beyond the 20 genetically encoded amino acids. Particular emphasis has been placed on peptides containing phosphorylated or glycosylated residues or disulfide bridges. Incorporation of side-chain-phosphorylated Ser and Thr by solid-phase peptide synthesis (SPPS) is especially challenging, as the phosphate group is decomposed by strong acid and lost with base in a β-elimination process. Boc-Ser(PO₃phenyl₂) and Boc-Thr(PO₃phenyl₂) have been found to be useful derivatives, where hydrogen fluoride (HF) or hydrogenolysis cleaves the peptide/resin and hydrogenolysis removes the phenyl groups. Fmoc-Ser(PO₃Bzl,H) and Fmoc-Thr(PO₃Bzl,H) can be used in conjunction with Fmoc chemistry with some care. Alternatively, peptide/resins that were built up by Fmoc chemistry to include unprotected Ser or Thr side chains may be subject to “global” or post-assembly phosphorylation. Side-chain-phosphorylated Tyr is less susceptible to strong-acid decom-

position and is not at all base-labile. Thus, SPPS has been used to incorporate directly Fmoc-Tyr(PO₃methyl₂), Fmoc-Tyr(PO₃*t*Bu₂), Fmoc-Tyr(PO₃H₂), and Boc-Tyr(PO₃H₂). Phosphorylation may also be accomplished on-line, directly after incorporation of the Tyr, Ser, or Thr residue but prior to assembly of the whole peptide.

Methodology for site-specific incorporation of carbohydrates during chemical synthesis of peptides has developed rapidly. The mild conditions of Fmoc chemistry are more suited for glycopeptide syntheses than Boc chemistry, as repetitive acid treatments can be detrimental to sugar linkages. Fmoc-Ser, -Thr, -5-hydroxylysine (-Hyl), -4-hydroxyproline (-Hyp), and -Asn have all been incorporated successfully with glycosylated side chains. The side-chain glycosyl is usually hydroxyl-protected by either benzoyl or acetyl groups, although some SPPSs have been successful with no protection of glycosyl hydroxyl groups. Deacetylation and debenzoylation are performed with hydrazine/methanol prior to glycopeptide/resin cleavage or in solution with catalytic methoxide in methanol.

Disulfide-bond formation has been achieved on the solid-phase by air, K₃Fe(CN)₆, dithiobis(2-nitrobenzoic acid), or diiodoethane oxidation of free sulfhydryls, by direct deprotection/oxidation of Cys(acetamidomethyl) residues using thallium trifluoroacetate or I₂, by direct conversion of Cys(9-fluorenylmethyl) residues using piperidine, and by nucleophilic attack by a free sulfhydryl on either Cys(3-nitro-2-pyridinesulfonyl) or Cys(*S*-carboxymethylsulfonyl). The most generally applicable and efficient of these methods is direct conversion of Cys(acetamidomethyl) residues by thallium trifluoroacetate.

Intra-chain lactams are formed between the side-chains of Lys or Orn and Asp or Glu to conformationally restrain synthetic peptides, with the goal of increasing biological potency and/or specificity. Lactams can also be formed via side-chain-to-head, side-chain-to-tail, or head-to-tail cyclization (Kates et al., 1994). The residues used to form intra-chain lactams must be selectively side-chain deprotected, while all side-chain protecting groups of other residues remain intact. Selective deprotection is best achieved by using orthogonal side-chain protection, such as allyloxycarbonyl or 1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl protection for Lys and allyl or *N*-[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl]aminobenzyl protection for Asp/Glu in combination with an Fmoc/*t*Bu strategy. Cycli-

zation is carried out most efficiently with BOP in the presence of DIEA while the peptide is still attached to the resin.

The three-dimensional orthogonal protection scheme of Fmoc/*t*Bu/allyl protecting groups is the strategy of choice for head-to-tail cyclizations. An amide linker is used for side-chain attachment of a C-terminal Asp/Glu (which are converted to Asn/Gln) and the α -carboxyl group is protected as an allyl ester. For side-chain-to-head cyclizations, the N-terminal amino acid (head) can simply be introduced as an N ^{α} -Fmoc derivative while the peptide-resin linkage and the other side-chain protecting groups are stable to dilute acid or carry a third dimension of orthogonality.

PROTEIN SYNTHESIS

There are three general chemical approaches for constructing proteins. First is stepwise synthesis, in which the entire protein is synthesized one amino acid at a time. Second is "fragment assembly," in which individual peptide strands are initially constructed stepwise, purified, and finally covalently linked to create the desired protein. Fragment assembly can be divided into two distinct approaches: (1) convergent synthesis of fully protected fragments, and (2) chemoselective ligation of unprotected fragments. Third is "directed assembly," in which individual peptide strands are constructed stepwise, purified, and then noncovalently driven to associate into protein-like structures. Combinations of the three general chemical approaches may also be employed for protein construction.

Convergent synthesis utilizes protected peptide fragments for protein construction (Albericio et al., 1997). The advantage of convergent protein synthesis is that fragments of the desired protein are first synthesized, purified, and characterized, ensuring that each fragment is of high integrity; these fragments are then assembled into the complete protein. Thus, cumulative effects of stepwise synthetic errors are minimized. Convergent synthesis requires ready access to pure, partially protected peptide segments, which are needed as building blocks. The application of solid-phase synthesis to prepare the requisite intermediates depends on several levels of selectively cleavable protecting groups and linkers. Methods for subsequent solubilization and purification of the protected segments are nontrivial. Individual rates for coupling segments are substantially lower than for activated amino acid species by stepwise synthesis, and there is always a risk of racemization at the C-terminus of each segment. Care-

ful attention to synthetic design and execution may minimize these problems.

As an alternative to the segment condensation approach, methods have been developed by which unprotected peptide fragments may be linked. "Native chemical ligation" results in an amide bond being generated between peptide fragments (Muir et al., 1997). A peptide bearing a C-terminal thioacid is converted to a 5-thio-2-nitrobenzoic acid ester and then reacted with a peptide bearing an N-terminal Cys residue (Dawson et al., 1994). The initial thioester ligation product undergoes spontaneous rearrangement, leading to an amide bond and regeneration of the free sulfhydryl on Cys. The method was later refined so that a relatively unreactive thioester can be used in the ligation reaction (Dawson et al., 1997; Ayers et al., 1999). Safety-catch linkers are used in conjunction with Fmoc chemistry to produce the necessary peptide thioester (Shin et al., 1999).

SIDE-REACTIONS

The free N^α-amino group of an anchored dipeptide is poised for a base-catalyzed intramolecular attack of the C-terminal carbonyl. Base deprotection of the Fmoc group can thus release a cyclic diketopiperazine while a hydroxymethyl-handle leaving group remains on the resin. With residues that can form *cis* peptide bonds, e.g., Gly, Pro, *N*-methylamino acids, or D-amino acids, in either the first or second position of the (*C* → *N*) synthesis, diketopiperazine formation can be substantial. The steric hindrance of the 2-chlorotrityl linker may minimize diketopiperazine formation of susceptible sequences during Fmoc chemistry.

The conversion of side-chain protected Asp residues to aspartimide residues can occur by repetitive base treatments. The cyclic aspartimide can then react with piperidine to form the α- or β-piperidide or α- or β-peptide. Aspartimide formation can be rapid, and is dependent upon the Asp side-chain protecting group. Sequence dependence studies of Asp(O*t*Bu)-X peptides revealed that piperidine could induce aspartimide formation when X = Arg(2,2,5,7,8-pentamethylchroman-6-sulfonyl; Pmc), Asn(triphenylmethyl; Trt), Asp(O*t*Bu), Cys(Acm), Gly, Ser, Thr, and Thr(*t*Bu) (Lauer et al., 1995). Aspartimide formation can also be conformation-dependent. This side-reaction can be minimized by including 0.1 M HOBt in the piperidine solution (Lauer et al., 1995), or by using an amide backbone protecting group (i.e., 2-hydroxy-4-methoxybenzyl) for the resi-

due in the X position of an Asp-X sequence (Quibell et al., 1994).

Cys residues are racemized by repeated piperidine deprotection treatments during Fmoc SPPS. Racemization of esterified (C-terminal) Cys can be reduced by using 1% 1,8-diazabicyclo[5.4.0]undec-7-ene in *N,N*-dimethylformamide (DMF). Additionally, the steric hindrance of the 2-chlorotrityl linker minimizes racemization of C-terminal Cys residues. When applying protocols for Cys internal (not C-terminal) incorporation which include phosphonium and aminium salts as coupling agents, as well as preactivation in the presence of suitable additives and tertiary amine bases, significant racemization is observed. Racemization is generally reduced by avoiding preactivation, using a weaker base (such as collidine), and switching to the solvent mixture DMF-dichloromethane (DCM) (1:1). Alternatively, the pentafluorophenyl ester of a suitable Fmoc-Cys derivative can be used.

The combination of side-chain protecting groups and anchoring linkages commonly used in Fmoc chemistry are simultaneously deprotected and cleaved by TFA. Cleavage of these groups and linkers results in liberation of reactive species that can modify susceptible residues, such as Trp, Tyr, and Met. Modifications can be minimized during TFA cleavage by utilizing effective scavengers. Three efficient cleavage "cocktails" quenching reactive species and preserving amino acid integrity, are TFA-phenol-thioanisole-1,2-ethanedithiol-H₂O (82.5:5:5:2.5:5) (reagent K) (King et al., 1990), TFA-thioanisole-1,2-ethanedithiol-anisole (90:5:3:2) (reagent R) (Albericio et al., 1990), and TFA-phenol-H₂O-triisopropylsilane (88:5:5:2) (reagent B) (Solé and Barany, 1992). The use of Boc side-chain protection of Trp also significantly reduces alkylation by Pmc or 2,2,4,6,7-pentamethyldihydro-benzofuran-5-sulfonyl (Pbf) groups.

PURIFICATION AND ANALYSIS OF SYNTHETIC PEPTIDES

Each synthetic procedure has limitations, and even in the hands of highly experienced workers, certain sequences defy facile preparation. The maturation of high-performance liquid chromatography (HPLC) has been a major boon to modern peptide synthesis, because the resolving power of this technique facilitates removal of many of the systematic low-level by-products that accrue during chain assembly and upon cleavage. Peptide purification is most commonly achieved by reversed-phase HPLC

(RP-HPLC; UNIT 10.14). Either alternatively to or in tandem with RP-HPLC, ion-exchange HPLC (UNIT 10.10) and gel-filtration HPLC (UNIT 10.9) can be used for isolation of desired peptide products. The progress of peptide purification can be monitored rapidly by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) or ion-trap electrospray MS (UNITS 10.21 & 10.22).

The homogeneity of synthetic materials should be checked by at least two chromatographic or electrophoretic techniques—e.g., RP-HPLC (UNIT 10.14), ion-exchange HPLC (UNIT 10.10), and capillary zone electrophoresis (UNIT 10.20). Also, determination of a molecular ion by MS (UNIT 10.21) using a mild ionization method is important for proof of structure. Synthetic peptides must be checked routinely for the proper amino acid composition, and in some cases sequencing data are helpful. The PSRC studies (see discussion of Development of Solid-Phase Peptide Synthesis Methodology) have allowed for a side-by-side comparison of a variety of analytical techniques. Efficient characterization of synthetic peptides best been obtained by a combination of RP-HPLC and MS, with sequencing by either Edman degradation sequence analysis or tandem MS (UNIT 10.21) being used to identify the positions of modifications and deletions. Proper peptide characterization by multiple techniques is essential.

LITERATURE CITED

- Albericio, F., Kneib-Cordonier, N., Biancalana, S., Gera, L., Masada, R.I., Hudson, D., and Barany, G. 1990. Preparation and application of the 5-(4-(9-fluorenylmethoxycarbonyl)aminomethyl-3,5-dimethoxyphenoxy)valeric acid (PAL) handle for the solid-phase synthesis of C-terminal peptide amides under mild conditions. *J. Org. Chem.* 55:3730-3743.
- Albericio, F., Lloyd-Williams, P., and Giralt, E. 1997. Convergent solid-phase peptide synthesis. *Methods Enzymol.* 289:313-336.
- Angeletti, R.H., Bonewald, L.F., and Fields, G.B. 1997. Six year study of peptide synthesis. *Methods Enzymol.* 289:697-717.
- Atherton, E. and Sheppard, R.C. 1987. The fluorenylmethoxycarbonyl amino protecting group. *In The Peptides*, Vol. 9 (S. Udenfriend and J. Meienhofer, eds.) pp. 1-38. Academic Press, New York.
- Ayers, B., Blaschke, U.K., Camarero, J.A., Cotton, G.J., Holford, M., and Muir, T.W. 1999. Introduction of unnatural amino acids into proteins using expressed protein ligation. *Biopolymers (Peptide Sci.)* 51:343-354.
- Barany, G. and Merrifield, R.B. 1979. Solid-phase peptide synthesis. *In The Peptides*, Vol. 2 (E. Gross and J. Meienhofer, eds.) pp. 1-284. Academic Press, New York.
- Bergmann, M. and Zervas, L. 1932. Über ein allgemeines Verfahren der Peptidsynthese. *Ber. Dtsch. Chem. Ges.* 65:1192-1201.
- Carpino, L.A. and Han, G.Y. 1972. The 9-fluorenylmethoxycarbonyl amino-protecting group. *J. Org. Chem.* 37:3404-3409.
- Dawson, P.E., Muir, T.W., Clark-Lewis, I., and Kent, S.B.H. 1994. Synthesis of proteins by native chemical ligation. *Science* 266:776-779.
- Dawson, P.E., Churchill, M.J., Ghadiri, M.R., and Kent, S.B.H. 1997. Modulation of reactivity in native chemical ligation through the use of thiol additives. *J. Am. Chem. Soc.* 119:4325-4329.
- duVigneaud, V., Ressler, C., Swan, J.M., Roberts, C.W., Katsoyannis, P.G., and Gordon, S. 1953. The synthesis of an octapeptide amide with the hormonal activity of oxytocin. *J. Am. Chem. Soc.* 75:4879-4880.
- Fields, C.G. and Fields, G.B. 1994. Solvents for solid-phase peptide synthesis. *In Methods in Molecular Biology*, Vol. 35: Peptide Synthesis Protocols (M.W. Pennington and B.M. Dunn, eds.) pp. 29-40. Humana Press, Totowa, N.J.
- Fields, G.B. and Noble, R.L. 1990. Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int. J. Peptide Protein Res.* 35:161-214.
- Fields, C.G., Fields, G.B., Noble, R.L., and Cross, T.A. 1989. Solid phase peptide synthesis of ¹⁵N-gramicidins A, B, and C and high performance liquid chromatographic purification. *Int. J. Peptide Protein Res.* 33:298-303.
- Fields, G.B., Otteson, K.M., Fields, C.G., and Noble, R.L. 1990. The versatility of solid phase peptide synthesis. *In Innovation and Perspectives in Solid Phase Synthesis: Peptides, Polypeptides and Oligonucleotides, Macro-organic Reagents and Catalysts* (R. Epton, ed.) pp. 241-260. Solid Phase Conference Coordination, Ltd., Birmingham, U.K.
- Fields, G.B., Lauer-Fields, J.L., Liu, R.-q., and Barany, G. 2001. Principles and practice of solid-phase peptide synthesis. *In Synthetic Peptides: A User's Guide*, 2nd ed. (G.A. Grant, ed.) in press. W.H. Freeman, New York.
- Fischer, E. and Fourneau, E. 1901. Über einige Derivate des Glykocoils. *Ber. Dtsch. Chem. Ges.* 34:2868-2877.
- Kates, S.A., Solé, N.A., Albericio, F., and Barany, G. 1994. Solid-phase synthesis of cyclic peptides. *In Peptides: Design, Synthesis and Biological Activity* (C. Basava and G.M. Anantharamaiah, eds.) pp. 39-57. Birkhaeuser, Boston.
- King, D.S., Fields, C.G., and Fields, G.B. 1990. A cleavage method which minimizes side reactions following Fmoc solid phase peptide synthesis. *Int. J. Peptide Protein Res.* 36:255-266.

- Lauer, J.L., Fields, C.G., and Fields, G.B. 1995. Sequence dependence of aspartimide formation during 9-fluorenylmethoxycarbonyl solid-phase peptide synthesis. *Lett. Peptide Sci.* 1:197-205.
- Merrifield, R.B. 1963. Solid phase peptide synthesis I: The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* 85:2149-2154.
- Merrifield, R.B. 1967. New approaches to the chemical synthesis of peptides. *Recent Prog. Hormone Res.* 23:451-482.
- Merrifield, R.B. 1986. Solid phase synthesis. *Science* 232:341-347.
- Merrifield, R.B., Stewart, J.M., and Jernberg, N. 1966. Instrument for automated synthesis of peptides. *Anal. Chem.* 38:1905-1914.
- Muir, T.W., Dawson, P.E., and Kent, S.B.H. 1997. Protein synthesis by chemical ligation of unprotected peptides in aqueous solution. *Methods Enzymol.* 289:266-298.
- Quibell, M., Owen, D., Packman, L.C., and Johnson, T. 1994. Suppression of piperidine-mediated side product formation for Asp(OBu^t)-containing peptides by the use of *N*-(2-hydroxy-4-methoxybenzyl) (Hmb) backbone amide protection. *J. Chem. Soc. Chem. Commun.* 2343-2344.
- Rich, D.H. and Singh, J. 1979. The carbodiimide method. *In* The Peptides, Vol. 1 (E. Gross and J. Meienhofer, eds.) pp. 241-314. Academic Press, New York.
- Sakakibara, S., Shimonishi, Y., Kishida, Y., Okada, M., and Sugihara, H. 1967. Use of anhydrous HF in peptide synthesis I: Behavior of various protective groups in anhydrous HF. *Bull. Chem. Soc. Jpn.* 40:2164-2167.
- Sarin, V.K., Kent, S.B.H., and Merrifield, R.B. 1980. Properties of swollen polymer networks: Swelling and swelling of peptide-containing resins in solid-phase peptide synthesis. *J. Am. Chem. Soc.* 102:5463-5470.
- Shin, Y., Winans, K.A., Backes, B.J., Kent, S.B.H., Ellman, J.A., and Bertozzi, C.R. 1999. Fmoc-based synthesis of peptide-^othioesters: Application to the total chemical synthesis of a glycoprotein by native chemical ligation. *J. Am. Chem. Soc.* 121:11684-11689.
- Solé, N.A. and Barany, G. 1992. Optimization of solid-phase synthesis of [Ala⁸]-dynorphin A. *J. Org. Chem.* 57:5399-5403.
- Tam, J.P. and Lu, Y.-A. 1995. Coupling difficulty associated with interchain clustering and phase transition in solid phase peptide synthesis. *J. Am. Chem. Soc.* 117:12058-12063.

KEY REFERENCES

Atherton, E. and Sheppard, R.C. 1989. Solid Phase Peptide Synthesis: A Practical Approach. IRL Press, Oxford.

An extensive collection of Fmoc-based synthetic methods and techniques.

Barany and Merrifield, 1979. See above.

The definitive, comprehensive overview of the solid-phase method.

Fields, G.B. 1997. Solid-phase peptide synthesis. *Methods Enzymol.* Vol. 289.

A contemporary collection of SPPS techniques and applications.

Contributed by Gregg B. Fields
Florida Atlantic University
Boca Raton, Florida

Synthetic Peptides for Production of Antibodies that Recognize Intact Proteins

Antibodies that recognize intact proteins can be produced through the use of synthetic peptides based on short stretches of the protein sequence, without first having to isolate the protein. The procedure for selecting stretches of protein sequence likely to be antigenic is relatively straightforward. However, no procedure will identify a single sequence guaranteed to be effective, nor will it usually identify the best single sequence to use. Rather, several sequences will be identified that have a higher-than-average probability of producing an effective antigen.

The steps to produce an effective antibody include: (1) designing the peptide sequence based on the sequence of the protein; (2) synthesizing the peptide; (3) preparing the immunogen either by coupling the synthetic peptide to a carrier protein or through the use of a multiple antigenic peptide (MAP); (4) immunizing the host animal; (5) assaying antibody titer in the host animal's serum; and (6) obtaining the antiserum and/or isolating the antibody. This unit covers steps 1 and 3; step 2 requires a laboratory with expertise in peptide synthesis. Peptide synthesis services are widely available both academically and commercially.

The best method to select potentially effective sequences is via a computer-assisted strategy (see Basic Protocol 1). An alternative manual method is also described (see Alternate Protocol 1) but is not recommended to replace the use of algorithms if there is a choice. A small synthetic peptide is usually insufficiently immunogenic on its own, and two methods have been developed to solve this problem. The first (see Basic Protocol 2) involves chemically coupling the synthetic peptide to a carrier protein to boost the immune response. The second method (see Alternate Protocol 2) entails direct synthesis of a MAP covalent multimer of the simple peptide sequence. Both methods have proven effective and it is a matter of personal preference which to use. Coupling to a carrier protein requires additional chemical manipulations after synthesis of the peptide, while the MAP is complete and ready for immunization at the conclusion of the synthetic protocol. Disadvantages of MAPs are that they are more difficult to produce homogeneously and to analyze postsynthetically. They also may be more prone to insolubility problems.

A carrier protein is a relatively large molecule capable of stimulating an immune response independently. A synthetic peptide coupled to a carrier protein acts as a hapten and produces antibodies specific for the hapten (antibodies against the carrier protein are also produced). The most commonly used carrier proteins are keyhole limpet hemocyanin (KLH) and bovine or rabbit serum albumin (BSA or RSA). KLH is usually preferred, because it tends to elicit a stronger immune response and is evolutionarily more remote from mammalian proteins. A common problem with KLH, however, has been its solubility. Pierce Chemical Company sells a preparation of KLH purported to have better solubility properties (see below).

Alternatively, peptides can be coupled to carrier proteins through either their amino (see Alternate Protocol 3) or carboxyl groups (see Alternate Protocol 4). These two alternate protocols are not recommended as a first choice for coupling, but are included because they have been used successfully and may be advantageous for certain special applications discussed in the Commentary. Also presented are methods for assaying free sulfhydryl content and for reducing disulfide bonds in synthetic peptides (see Support Protocols 1 and 2).

Contributed by Gregory A. Grant

Current Protocols in Molecular Biology (2002) 11.16.1-11.16.19

Copyright © 2002 by John Wiley & Sons, Inc.

Once the coupling procedure has been performed, it is possible to determine the approximate degree of coupling by amino acid analysis (see Support Protocol 3). However, in most instances this is unnecessary and the product can be used directly.

COMPUTER-ASSISTED SELECTION OF APPROPRIATE ANTIGENIC PEPTIDE SEQUENCES

An antibody produced in response to a simple linear peptide will most likely recognize a linear epitope in a protein. Furthermore, that epitope must be solvent-exposed to be accessible to the antibody. The general features of protein structure that correspond to these criteria are turns or loop structures, which are generally found on the protein surface connecting other elements of secondary structure, and areas of high hydrophilicity, especially those containing charged residues. As a consequence, computer algorithms that predict protein hydrophilicity and tendency to form turns are very useful. Several analytic programs or algorithms that attempt to do this have been developed. Although the choice of method may rely on availability or personal preference, there tends to be a high level of agreement among them. As stated earlier, none of the methods will identify the one single sequence guaranteed to produce an effective antibody against any given protein. Rather, the methods will offer several good candidates, one or several of which can be used.

Many of these algorithms may already be available on a local computer system. They are included in many commercial software packages such as GCG (Genetics Computer Group; see *APPENDIX 4*). The ExPASy Web site of the University of Geneva offers free access to a variety of different programs over the Internet at <http://expasy.org/tools>.

The following protocol utilizes the hydropathy index developed by Kyte and Doolittle (1982) and the secondary structure prediction method for β turns developed by Chou and Fasman (1974) found in the tool "ProtScale" at the ExPASy Internet address.

1. Using the selected algorithms, compute the hydropathy index and the tendency for β -turns of the protein sequence. Use a window size of 7 or 9 and give equal weight to each amino acid. Record the results in either graphical or numerical form, or both.

As an example, the graphical representation of these results for the protein sequence shown in Figure 11.16.1 is presented in Figure 11.16.2.

A window size determines the number of amino acids to be used in computing a value for the amino acid at the center of the window. For example, a window size of 9 includes 4

10	20	30	40	50	60
MAKVSLEKDK	IKFLLVEGVH	QKALESRAA	GYTNI EFHKG	ALDDEQLKES	IRDAHFIGLR
SRTHLTEDVI	NAAEKLVAIG	CFCIGTNQVD	LDAAAKRGIP	VFNAPFSNTR	SVAELVIGEL 120
LLLLRGVPEA	NAKAHRGVWN	KLAAGSFEAR	GKKLGIIGYG	HIGTQLGILA	ESLGMVVFY 180
DIENKLPLGN	ATQVQHLSDL	LNMSDVVSLH	VPENPSTKNM	MGAKEISLMK	PGSLLINASR 240
GTVVDIPALC	DALASKHLAG	AAIDVFPTEP	ATNSDPFTSP	LCEFDNVLLT	PHIGGSTQEA 300
QENIGLEVAG	KLIKYSDNGS	TLSAVNFPEV	SLPLHGGRRL	MHIHENRPGV	LTALNKIFAE 360
QGVNIAAQYL	QTSAQMGYVV	IDIEADEDVA	EKALQAMKAI	PGTIRARLLY	

Figure 11.16.1 The amino acid sequence of a 410-residue protein analyzed by the method presented in Basic Protocol 1. The results are shown in Figure 11.16.2.

amino acids on each side of the central amino acid. The value computed for the central amino acid is the simple average of the values for each amino acid in the window.

2. Compare the results of the two analyses and look for areas of sequence that are high in turn tendency and high in hydrophilicity (low in hydrophobicity).

In Figure 11.16.2, these areas correspond to positive peaks in the Chou-Fasman analysis and negative peaks in the Kyte-Doolittle analysis. The three best areas in terms of amplitude and correlation are shaded. These correspond to the sequences underlined in Figure 11.16.1. (Note the alignment of these peak optima as compared to the peaks around residue 300.)

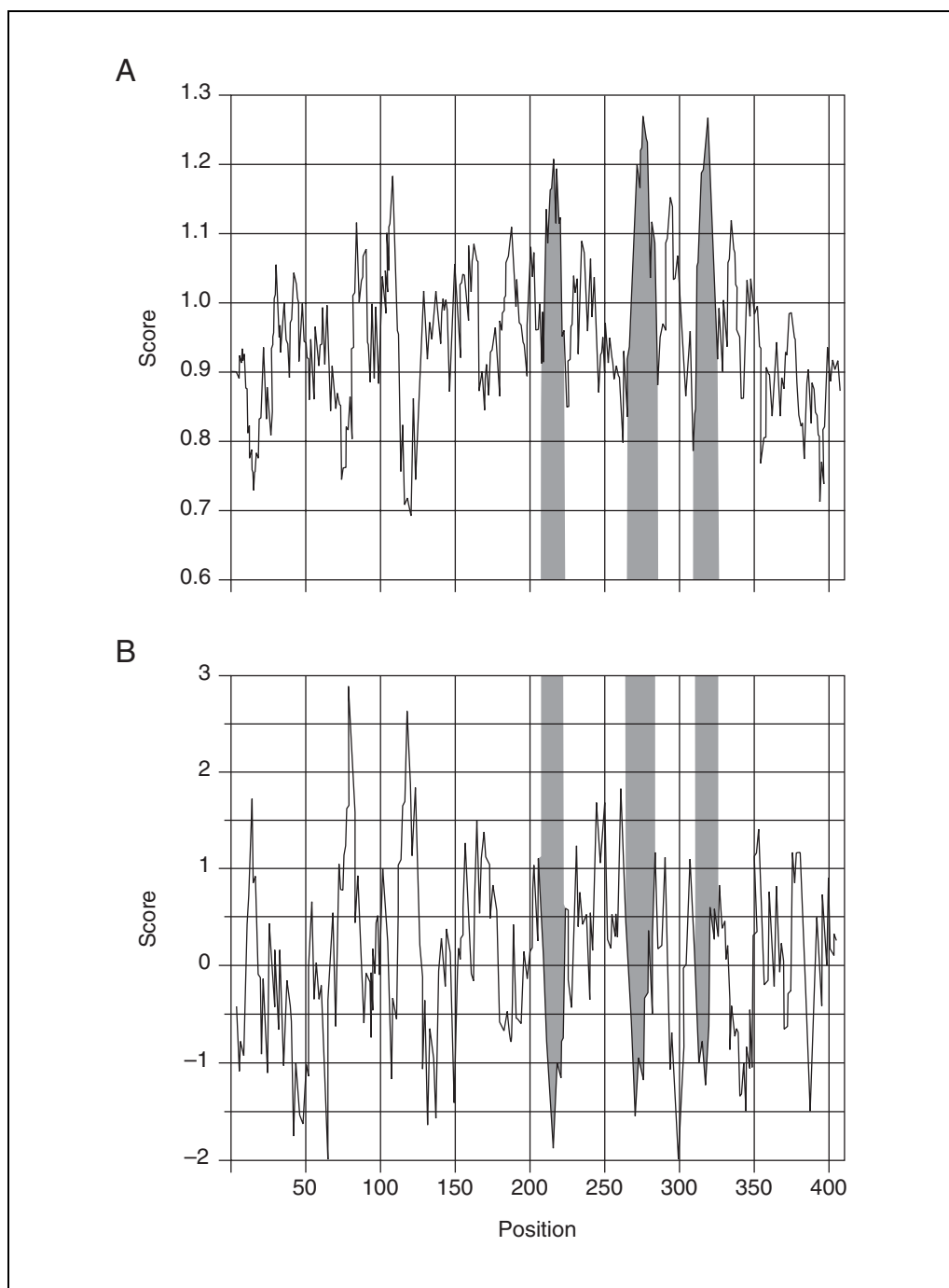


Figure 11.16.2 Graphical representation of the results generated by a computer algorithm for the sequence in Figure 11.16.1, analyzed by the method presented in Basic Protocol 1. The shaded areas represent three regions in the sequence meeting criteria for selection as potential immunogens. **(A)** Analysis for β turns (Chou and Fasman, 1974). **(B)** Analysis for hydrophobicity (Kyte and Doolittle, 1982).

3. Examine the sequences for glycosylation site motifs and discard any sequences that contain them unless it is known that the protein is not glycosylated.

Amino acids in glycosylated regions may be shielded from presentation to an antibody by masking carbohydrates.

Amino-linked carbohydrate chains can occur at Asn-X-Ser or Asn-X-Thr sequences. Hydroxyl-linked carbohydrate chains do not appear to have a set motif. A program to assist in the prediction of mucin-type GalNAc O-glycosylation sites in mammalian lipoproteins is found in the tool "NetOGlc" at the ExPASy site (<http://expasy.org/tools>). However, before using read the documentation carefully and keep in mind that such prediction methods cannot always be successful.

4. Select the best sequences resulting from this analysis to use as antigenic peptides. These are sequences where the largest positive values (peaks with positive deflection) for turn propensity correspond in position to the largest negative values (peaks with negative deflection) for hydrophobicity. The values obtained in these analyses are relative and dependent on the individual protein's composition, so it is not possible to set an arbitrary minimum value as a cutoff for rejecting a particular peak. Rather, always select the peaks of greatest magnitude in any given sequence. In addition, the immediate amino-terminal and carboxyl-terminal regions of proteins are often exposed to solvent. If these areas appear to be hydrophilic in nature, they are also acceptable candidates. Thus each analysis may provide several potential sequences. How many peptides to make (see Anticipated Results) is a matter of individual choice.

ALTERNATE PROTOCOL 1

MANUAL INSPECTION TO SELECT APPROPRIATE PEPTIDE SEQUENCES

If computer algorithms are not available, it is possible to select potential sequences by manual inspection. Although there is no evidence that a manual method is any less effective than the use of computer algorithms, there is a greater probability of overlooking potentially important areas of sequence. It is therefore recommended that computer analysis be used whenever it is available. Although it can be done, it would be very time consuming and labor intensive to manually calculate values for every overlapping peptide offset by a single amino acid in the same way that the algorithms do. For this reason, areas rich in polar residues are selected for manual calculation of hydrophilicity and turn propensity.

1. Visually inspect the protein sequence and select areas that contain at least two to three charged residues (Lys, Arg, His, Asp, Glu) within a 10- to 15-residue span.

If this criterion cannot be met, select sequences with the greatest number of charged residues.

2. From the sequences identified in step 1, select a subset of sequences that are the highest in Ser, Thr, Asn, Gln, Pro, and Tyr content.
3. Calculate average hydrophilicity and turn propensity for each amino acid in the selected sequences using the values given in Table 11.16.1 and a window of 9 residues (see Basic Protocol 1, step 1).

Be sure to include the residues flanking the selected sequence for calculation of values for the residues at the ends of the selected sequence. In other words, do not use different size windows.

4. Plot the values for each amino acid of a chosen sequence.

Sequences whose optimal values for hydrophilicity and turn propensity correspond (as in Fig. 11.16.2) are considered good candidates.

5. Inspect sequences for glycosylation motifs and discard these candidates (see Basic Protocol 1, step 3).

Table 11.16.1 Hydrophobic and β -Turn Indices of Amino Acids

Amino acid	Symbols	Hydrophobicity value ^a	β -turn propensity ^b
Arginine	Arg (R)	-4.5	0.95
Lysine	Lys (K)	-3.9	1.01
Aspartic acid	Asp (D)	-3.5	1.46
Glutamic acid	Glu (E)	-3.5	0.74
Asparagine	Asn (N)	-3.5	1.56
Glutamine	Gln (Q)	-3.5	0.98
Histidine	His (H)	-3.2	0.95
Proline	Pro (P)	-1.6	1.52
Tyrosine	Tyr (Y)	-1.3	1.14
Tryptophan	Trp (W)	-0.9	0.96
Serine	Ser (S)	-0.8	1.43
Threonine	Thr (T)	-0.7	0.96
Glycine	Gly (G)	-0.4	1.56
Alanine	Ala (A)	1.8	0.66
Methionine	Met (M)	1.9	0.60
Cysteine	Cys (C)	2.5	1.19
Phenylalanine	Phe (F)	2.8	0.60
Leucine	Leu (L)	3.8	0.59
Valine	Val (V)	4.2	0.50
Isoleucine	Ile (I)	4.5	0.47

^aKyte and Doolittle (1982).^bChou and Fasman (1974).

Amino acids of glycosylated regions may be masked in native proteins, so an antibody raised against them would be ineffective.

6. Select the best sequences (see Basic Protocol 1, step 4 for criteria), choosing a high turn-propensity-to-hydrophobicity ratio.

DESIGNING A SYNTHETIC PEPTIDE FOR COUPLING TO A CARRIER PROTEIN

Although there is no direct evidence to show that the state of the termini of the peptide affects its ability to produce antibodies that will react with the protein, most procedures suggest that the termini of the peptide should mimic their native state. Thus, sequences whose terminal residues normally are in peptide linkage in the protein can have their amino-terminal and carboxyl-terminal groups modified by acetylation and amidation, respectively, during synthesis.

Modification of the amino or carboxyl termini will decrease the polarity of the peptide in solution and could have a significant effect on the peptide's solubility. If the peptide lacks sufficient protonatable side chains, modification of the termini can be omitted. A general rule to predict solubility is that the total number of charges at a given pH should be at least 20% of the number of residues in the peptide.

1. Choose a sequence of 10 to 15 amino acid residues for the synthetic peptide.

Longer peptides are more difficult and expensive to make, and they are usually unnecessary.

BASIC PROTOCOL 2

Immunology

11.16.5

Try to choose a stretch of sequence that contains some charged residues such as Arg, Lys, His, Glu, and Asp. In addition to the high likelihood of these amino acids being located on the surface of the protein, they aid handling of the synthetic product by promoting solubility.

- 2a. *If the selected sequence does not contain an internal cysteine:* Place a cysteine on either the amino or carboxyl terminus for use in coupling to a carrier protein with a heterobifunctional cross-linking reagent such as MBS, *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester (see Basic Protocol 3).

*Cross-linking with heterobifunctional reagents is the recommended procedure for most peptides (see Basic Protocol 3). As an alternative to using a chemical cross-linking reagent, any peptide, regardless of amino acid content, can also be photochemically linked to a carrier protein if *p*-benzoyl benzoic acid is added to the peptide during synthesis (see Alternate Protocol 5). Although photochemical cross-linking is effective, it is not widely used.*

If the sequence includes the immediate amino or carboxyl terminal sequence of the protein, the cysteine should be placed on the end that would normally be engaged in the internal peptide bond. For sequences internal to the protein, the cysteine may be placed at either end according to the preference of the synthetic chemist. However, if amino-terminal capping (acetylation) is used after the coupling of each amino acid during synthesis, it is preferable to place the cysteine on the amino-terminal end of the peptide since then only the full-length peptide will contain the cysteine residue. In this way, if synthetic difficulties are encountered, only the full-length peptide will couple to the carrier. If placed on the carboxyl terminal end, the cysteine residue tends to racemize during synthesis unless a chlorotriyl resin is used. However, this should not have an effect on the rest of the peptide or the generation of antibodies.

- 2b. *If the sequence contains an internal cysteine residue:* Do not add a terminal cysteine for MBS cross-linking. Rather, use an alternative coupling procedure (see Alternate Protocols 3, 4, or 5) or synthesize a multiple antigenic peptide (MAP; see Alternate Protocol 2).

Internal cysteine sulfhydryl groups will also cross-link to the carrier protein, and multiple cysteines will result in a peptide attached at multiple points. If the sulfhydryl will eventually be important for antibody recognition of the protein, the immunization may not produce effective antibodies. Furthermore, the additional constraint produced by the existence of multiple points of coupling may affect the ultimate ability of the antibody to recognize the protein.

- 3a. *For sequences whose terminal residues are in peptide linkage within the protein:* If the peptide is coupled using a heterobifunctional reagent such as MBS (see Basic Protocol 3), modify the amino and carboxyl ends by acetylation and amidation, respectively, during the synthetic procedure. If coupling is performed with a homobifunctional reagent that reacts with amino groups such as glutaraldehyde (see Alternate Protocol 3) or by the photochemical method (see Alternate Protocol 5), only amidate the carboxyl terminus. If coupling is performed with EDC (see Alternate Protocol 4), only acetylate the amino terminus.

- 3b. *For sequences that are amino or carboxyl terminal to the protein:*

- i. *If the sequence is the immediate amino or carboxyl terminal sequence of the protein and the peptide will be coupled with a heterobifunctional reagent such as MBS (see Basic Protocol 3), leave the end that is not in peptide linkage (and the end that does not contain the additional cysteine residue for MBS coupling) as the free amino or carboxyl group unless it is known that they are normally blocked.*
- ii. *If the sequence is the immediate amino terminal sequence of the protein and the peptide will be coupled with a homobifunctional reagent that reacts with amino*

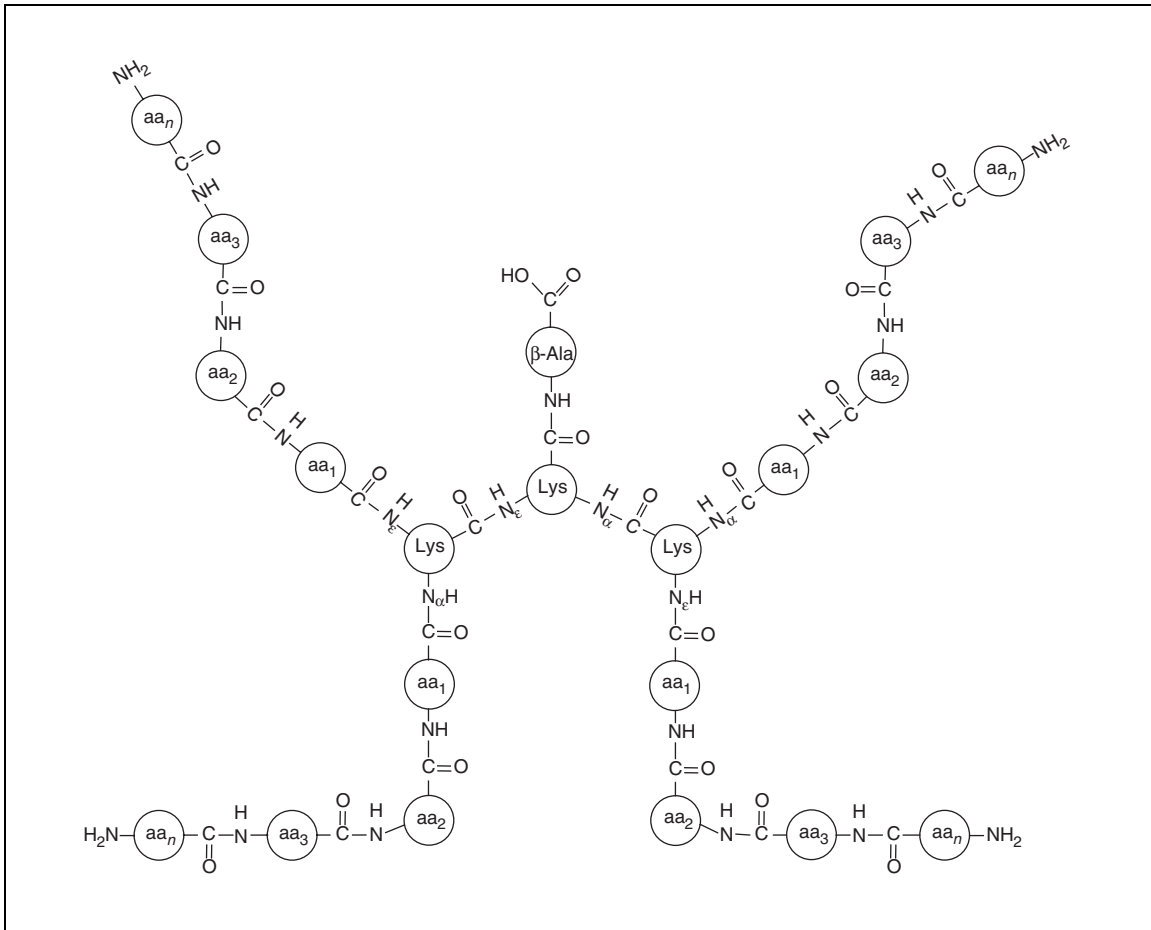


Figure 11.16.3 Representation of a four-branched multiple antigenic peptide (MAP). The MAP core is usually synthesized with a β -alanine residue attached to the solid phase support followed by a scaffold of lysine residues. The first lysine residue, attached through its α -carboxyl group to β -alanine, provides two primary amino groups (α and ϵ) for chain elongation. The next level of the lysine scaffold consists of two lysine residues which in turn provide four primary amino groups for chain elongation. MAPs with eight branches are formed by providing one additional layer of lysine to that depicted here.

groups such as glutaraldehyde (see Alternate Protocol 3) or photochemically (see Alternate Protocol 5), amidate the C-terminus.

- iii. If the sequence is the immediate carboxyl terminal sequence of the protein and the peptide will be coupled with glutaraldehyde (see Alternate Protocol 3) or photochemically (see Alternate Protocol 5), leave both termini free.
4. If the peptide will be coupled with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC; see Alternate Protocol 4), block the amino terminus by acetylation.

This should be done even if the sequence is at the immediate amino terminus of the protein, since treatment with EDC can result in the production of covalent multimers of the peptide through reaction with the N-terminal amino group. Some procedures recommend using citraconylation to temporarily protect the amino group. However, this also adds a free carboxyl group that could result in multiple site attachment of the peptide to the carrier.

DESIGNING A SYNTHETIC MULTIPLE ANTIGENIC PEPTIDE

A multiple antigenic peptide (MAP; Posnett et al., 1988; Tam, 1988) is an effective alternative to coupling a simple linear peptide to a carrier protein. MAPs are covalent constructs consisting of a simple peptide sequence synthesized on a branched core (Fig. 11.16.3) with one copy of the peptide sequence on each of four or eight branches. One advantage of a MAP is that it is suitable for use as an immunogen at the conclusion of the synthetic process. On occasion, MAPs have produced effective protein antibodies when the conventional peptide coupled to carrier protein has not. Thus, MAPs represent an effective alternative approach for antiserum production.

1. Select a sequence between 10 and 15 residues in length.

Longer sequences are unnecessary and increase the probability of synthetic problems. The presence of internal cysteine residues are not a concern with MAPs, but if present, take precautions to keep them reduced.

2. Synthesize the MAPs utilizing a four-branch core.

Synthesize the MAP core de novo, or purchase resins for solid-phase peptide synthesis with four- or eight-branched cores (available commercially from Advanced ChemTech, Novabiochem, Applied Biosystems, AnaSpec, and Bachem Bioscience). Eight-branched cores are suitable if the peptide is no more than 12 to 14 residues and has a high degree of hydrophilicity. There are more synthetic problems with eight-branched MAPs, presumably due to the higher density of structure during synthesis: they are more difficult to characterize and probably raise a diverse antibody population against some synthetic artifacts (see Mints et al., 1997).

3. *Optional.* If the selected sequence was not the amino terminus of the protein, acetylate the new amino terminus.

In the case of a MAP, the carboxyl terminus will remain in covalent linkage to the branched core.

4. Use the MAP directly as an immunogen.

Coupling to a carrier protein as described in Basic Protocol 3 is usually not necessary.

COUPLING SYNTHETIC PEPTIDES TO A CARRIER PROTEIN USING A HETEROBIFUNCTIONAL REAGENT

If the synthetic peptide was designed with a cysteine residue at one terminus (see Basic Protocol 2, step 2a), the following procedure should be followed for coupling to keyhole limpet hemocyanin (KLH) or other carrier proteins. Care must be taken to assure that the cysteine sulfhydryl group has remained reduced. Under normal synthetic conditions, if the peptide was lyophilized and stored dry immediately after synthesis, the sulfhydryl usually remains in the reduced state. The presence of free sulfhydryl groups in the peptide can be determined with Ellman's reagent (see Support Protocol 1) just prior to use; alternatively, high-resolution mass spectrometry can be used. If reduction is needed, follow the cysteine reduction procedure (see Support Protocol 2) before starting the coupling process.

The reagent most commonly used for this purpose is *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester (MBS). However, several related reagents (all available from Pierce) offer some additional features. Sulfo-MBS (Pierce) is a water-soluble alternative to MBS. Succinimidyl 4-(*N*-maleimidomethyl)-cyclohexane-1-carboxylate (SMCC) and its sulfonated analog sulfo-SMCC provide the same chemistry with a more pH-stable maleimide (see step 1). The MBS and SMCC reagents can be used interchangeably in this protocol (the sulfo reagents can be dissolved in aqueous solution, while the others must be dissolved in an organic solvent; the concentrations listed are appropriate for all four).

Materials

- Keyhole limpet hemocyanin (KLH; Pierce, Sigma, Calbiochem, or Boehringer Mannheim)
- 0.01 M sodium phosphate buffer, pH 7.5 (*APPENDIX 2*)
- 10 mg/ml MBS in fresh *N,N*-dimethylformamide (DMF)
- 0.05 M and 0.1 M sodium phosphate buffer, pH 7.0 (*APPENDIX 2*)
- Synthetic peptide with a reduced cysteine residue at either the N- or C-terminus
- 6 M guanidine-HCl (see recipe)
- Small glass vial with flat bottom
- ~0.9 × 15-cm gel filtration column with Sephadex G-25 or G-50 (Pharmacia Biotech) or Bio-Gel P2 or P4 (Bio-Rad) resin; *or* prepacked PD-10 column (Pharmacia Biotech)
- Additional reagents and equipment for gel filtration chromatography (*UNIT 10.9*)

NOTE: Do not use Tris or other buffers with primary amino groups in this procedure.

CAUTION: MBS is a moisture-sensitive irritant. Read the Material Safety Data Sheet before use.

1. Dissolve 5 mg KLH in ~0.5 ml of 0.01 M sodium phosphate buffer, pH 7.5, in a small, flat-bottomed vial.
A pH range of 7.0 to 7.5 offsets competing reactions. Although the unprotonated form of the amine reacts with the N-hydroxysuccinimide ester and would be optimal at pH >8.0, hydrolysis of the ester bond and reaction of the maleimide group with amines is enhanced at higher pH.
2. Add 100 µl of 10 mg/ml MBS/DMF solution and stir gently with a micro stir-bar 30 min at room temperature.
A small amount of precipitate may form during this procedure and is acceptable. However, if the precipitate is large, perform the procedure again with fresh components.
As an alternative to performing this coupling procedure from scratch, it is possible to purchase MBS-activated KLH (Pierce; Boehringer Mannheim) or kits containing MBS-activated KLH and an alternate MBS-activated protein for use in ELISA assays (Pierce).
3. Separate MBS-activated KLH from free MBS on a ~0.9 × 15-cm gel filtration column, equilibrating and eluting the column with 0.05 M sodium phosphate buffer, pH 7.0. Collect 0.5-ml fractions and read their absorbance at 280 nm (*UNIT 10.9*).
The first peak to elute is the KLH-MBS conjugate. These fractions may appear cloudy. The second peak is uncoupled MBS.
4. Pool the KLH-MBS conjugate fractions in a separate tube.
5. Dissolve 5 mg of the synthetic peptide in 0.01 M sodium phosphate buffer, pH 7.0, immediately prior to use. If the peptide is poorly soluble, use 6 M guanidine-HCl.
Maleimide groups react specifically with sulfhydryls at slightly acid to neutral pH.
6. Add the peptide solution to the KLH-MBS conjugate. Stir gently with a micro stir-bar 3 hr at room temperature.
The coupling may be continued overnight.
7. Dialyze against 4 liters distilled water overnight at 4°C. Use for immunizations within 24 hr.

ASSAY OF FREE SULFHYDRYLS WITH ELLMAN'S REAGENT

Free sulfhydryl content of a peptide can be quantitatively determined with Ellman's reagent, 5,5'-dithio-bis(2-nitrobenzoic acid). The molar extinction coefficient at 412 nm of thionitrobenzoate, the colored species generated when the reagent reacts with a free thiol, is 14,150 in 0.1 M sodium phosphate buffer (see Fig. 11.16.4). The sensitivity of the reaction is in the low nmol/ml range for sulfhydryl groups, making it well suited for synthetic peptides. By dry weight most synthetic peptides are only ~60% to 75% peptide, the remainder consisting of counterions and water of hydration. Amino acid analysis (see UNIT 11.9) is needed to establish the actual peptide content unambiguously, but such precise measurement is not usually necessary for qualitative evaluation of the free sulfhydryl content of a peptide sample.

Materials

Cysteine standard stock solution (see recipe)
0.1 M sodium phosphate, pH 8.0 (APPENDIX 2)
Peptide to be assayed
Ellman's reagent solution (see recipe)
13 × 100-mm glass test tubes

1. Prepare a cysteine standard curve by adding 25 μ l, 50 μ l, 100 μ l, 150 μ l, 200 μ l, and 250 μ l of cysteine standard stock solution to separate 13 × 100-mm tubes. Add \leq 250 μ l of each peptide to be tested to separate tubes. Bring the volume in each tube to 250 μ l with 0.1 M sodium phosphate, pH 8.0. Add 250 μ l of 0.1 M sodium phosphate, pH 8.0, to a blank tube.

The cysteine content of the peptide to be assayed should fall within the range of the standard curve (37.5 to 375 nmol).

2. Add 50 μ l Ellman's reagent solution and 2.5 ml of 0.1 M sodium phosphate, pH 8.0, to each tube. Mix and incubate 15 min at room temperature.
3. Measure absorbance at 412 nm (A_{412}).
4. Plot the A_{412} values of the standards after subtracting the value for the blank to produce a standard curve. Use this curve to determine the free sulfhydryl content of the peptides.

REDUCING CYSTEINE GROUPS IN PEPTIDES

When a peptide is synthesized with a terminal cysteine residue to be used for coupling with MBS (see Basic Protocol 3), the cysteine must be in the reduced state (present as free -SH rather than as a disulfide) in order to participate in the reaction with the coupling reagent. If peptides are lyophilized immediately after extraction from the resin cleavage cocktail or reversed-phase HPLC and used immediately after reconstitution, oxidation of cysteine side chains is usually not a problem. However, if oxidation to disulfides has occurred, the peptide can be reduced prior to use with the protocol presented here.

Dithiothreitol (DTT) is preferred to 2-mercaptoethanol (2-ME) as a reducing agent because its lower redox potential allows it to be effective at lower concentrations, and the reaction goes to completion because formation of the six-membered ring containing an internal disulfide is energetically favorable (see Fig. 11.16.4).

To determine if reduction is necessary, quantitate the level of free sulfhydryl groups with Ellman's reagent (see Support Protocol 1).

Additional methods for reducing disulfides include using sodium borohydride (Gailit, 1993) and Tris(2-carboxyethyl)phosphine (TCEP; Getz et al., 1999).

Materials

Synthetic peptide
0.1 M sodium phosphate, pH 8.0 (*APPENDIX 2*)
1 M aqueous dithiothreitol (DTT)
1 N HCl

100- or 250- μ l polypropylene tubes
Nitrogen gas source

Additional reagents and equipment for reversed-phase HPLC of peptides (see *UNIT 10.14*)

1. Dissolve 5 to 10 mg of peptide in 0.1 M sodium phosphate, pH 8.0.
2. Add 100 μ l of 1 M DTT.
3. Flush nitrogen over the surface of the liquid, seal the tube, and incubate 1 hr at 37°C.
4. Acidify with 1 N HCl and desalt by reversed-phase HPLC (*UNIT 10.14*).
5. Pool peptide fractions and lyophilize. Store lyophilized at 4°C until ready to use (up to several days).

The oxidation state of the peptide can usually be followed by analytical monitoring of its elution position on reversed-phase HPLC. Disulfide-linked dimers of peptides generally elute later than the monomeric peptide.

ALTERNATE PROTOCOL 3

COUPLING SYNTHETIC PEPTIDES TO A CARRIER PROTEIN USING A HOMOBIFUNCTIONAL REAGENT

The available homobifunctional reagents couple compounds through primary amino groups. Therefore, peptides with internal lysine residues should not be used in this procedure. The reagent most commonly used for this procedure is glutaraldehyde, but it should not be used with peptides containing internal Cys, Tyr, or His residues. Other homobifunctional cross-linking reagents that can be used in the same way as glutaraldehyde, but do not cross react with Cys, Tyr, or His residues, are also available: disuccinimidyl suberate (DSS), disuccinimidyl glutarate (DSG), and bis(sulfosuccinimidyl) suberate (BS3; all available from Pierce). However, these reagents are not widely employed for coupling peptides to proteins and are not considered the method of choice, and methods for their use have not been formalized. This is probably because coupling of the synthetic peptide to itself and aggregation of the carrier protein can occur with homobifunctional reagents such as these. Glutaraldehyde, on the other hand, although also subject to this limitation, has generally been used successfully.

Additional Materials (*also see Basic Protocol 3*)

50 mM sodium borate buffer, pH 8.0: adjust pH with HCl
Glutaraldehyde solution (see recipe)
1 M glycine in 50 mM sodium borate buffer, pH 8.0

NOTE: Do not use Tris or other buffers with primary amino groups in this procedure.

1. Dissolve 5 mg KLH in ~1.0 ml of 50 mM sodium borate buffer, pH 8.0.
2. Add 5 mg synthetic peptide.
3. Slowly add 1 ml fresh glutaraldehyde solution with gentle mixing at room temperature. Allow to react for an additional 2 hr with gentle mixing.

Formation of a yellowish color or milkiness is normal and does not affect the sample.

4. Add 0.25 ml of 1 M glycine to bind unreacted glutaraldehyde.

A darker yellow to brown color may develop.

5. Dialyze the reaction mixture overnight at 4°C against 4 liters of 50 mM sodium borate buffer, pH 8.0, and then overnight against water. Use immediately.

COUPLING SYNTHETIC PEPTIDES TO A CARRIER PROTEIN USING A CARBODIIMIDE

**ALTERNATE
PROTOCOL 4**

This procedure couples amino groups to carboxyl groups by way of activation of the carboxyl group with a water-soluble carbodiimide. Since the procedure is most easily performed in one step, peptides containing internal Asp, Glu, Lys, Tyr, or Cys residues should not be used. Also, in order to avoid making polymers of the peptide, the amino terminus should be blocked by acetylation during synthesis (see Basic Protocol 2). While this method has been used successfully, it is not considered to be the method of choice except in special situations.

Additional Materials (also see *Basic Protocol 3*)

- 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC; Pierce), used fresh or stored desiccated and frozen
- 0.1 N HCl

NOTE: Buffers containing amino or carboxyl groups should not be used in this procedure. According to some reports, buffers containing phosphate groups should also be avoided. Water is the safest choice as a solvent.

1. Dissolve 5 mg synthetic peptide in 1 ml water.
2. Add 25 mg EDC and carefully adjust pH to 4.0 to 5.0 by adding small amounts of 0.1 N HCl. Allow to react for 5 to 10 min at room temperature with gentle mixing.
pH paper suffices to monitor this adjustment.
3. Dissolve 5 mg KLH in 0.5 ml water and add to solution from step 2. React 2 hr at room temperature with gentle mixing.
4. Dialyze against 4 liters of water overnight at 4°C. Use immediately.

COUPLING SYNTHETIC PEPTIDES TO A CARRIER PROTEIN PHOTOCHEMICALLY

**ALTERNATE
PROTOCOL 5**

Most synthetic peptides today are made with 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry (see Fig. 18.1.1), and therefore the following protocol links the photoactive group to the free amino terminus of the peptide. An alternative approach if *t*-butyloxycarbonyl (Boc) chemistry is used is to link it to an ϵ -amino group of a terminal lysine (Gorka et al., 1989).

Additional Materials (see also *Basic Protocol 3*)

- 4-benzoyl benzoic acid (Sigma or Aldrich)
- Quartz spectrophotometry cuvettes

1. Have the synthetic chemist who is making the peptide attach a benzoyl benzoic acid group to the amino terminus simply by treating the reagent as the terminal residue during normal synthesis.

This blocks the amino terminus.

Immunology

11.16.13

**SUPPORT
PROTOCOL 3**

2. Dissolve 5 mg KLH in ~0.5 ml of 0.01 M sodium phosphate buffer, pH 7.5.
3. Add 2 mg of synthetic peptide containing the benzoyl benzoate adduct.
4. Place in a 1-cm quartz cuvette and irradiate with 366-nm light for 3 hr at a distance of 0.5 cm. Use immediately.

The peptide can be used directly for immunization.

CALCULATION OF THE MOLAR RATIO OF PEPTIDE TO CARRIER PROTEIN

The molar ratio of peptide to carrier protein coupling efficiency can be calculated to determine the level of substitution achieved by the coupling procedure. This information can be obtained using the results of amino acid compositional analysis. By performing the calculations presented in this protocol, the molecules of peptide in the conjugate per molecule of carrier protein in the conjugate can be determined.

1. Obtain the amino acid composition of the carrier protein, the peptide, and the peptide/carrier conjugate. Amino acid compositional analysis (of these hydrolysates) is usually available at sources that provide automated peptide synthesis (see *UNIT 11.15*). Be sure that the conjugate is free of unconjugated peptide (i.e., it should be well dialyzed).
2. Determine a scaling factor (SF) that relates the moles of protein in the unconjugated carrier protein to the moles of protein in the peptide/carrier conjugate. This is done by comparing the molar ratio of ≥ 3 amino acids present in the carrier protein and peptide/carrier conjugate but not present in the peptide. For example, if the peptide TGLRDSC (Table 11.16.2) is coupled to a carrier protein, choose A, K, and I. The calculation is done as:

Table 11.16.2 Sample Calculation of the Extent of Coupling of the Peptide TGLRDSC to Carrier Protein^a

Amino acid	Composition of carrier protein	Composition of peptide/carrier conjugate	Amount of carrier protein amino acids in conjugate	Amount of peptide amino acids in conjugate
D	80	185	160	25
E	110	222	220	—
G	95	215	190	25
S	65	150	130	20
T	70	163	140	23
H	10	19	20	—
P	25	51	50	—
A	103	206	206	—
M	5	11	10	—
V	60	118	120	—
F	22	45	44	—
L	55	133	110	23
I	65	135	130	—
C	7	22	14	8
Y	13	25	26	—
K	65	125	130	—
R	<u>75</u>	<u>177</u>	<u>150</u>	<u>27</u>
Total pmol amino acid	925	2002	1850	151

$$SF = [(pmol\ A\ in\ conjugate/pmol\ A\ in\ carrier) + (pmol\ K\ in\ conjugate/pmol\ K\ in\ carrier) + (pmol\ I\ in\ conjugate/pmol\ I\ in\ carrier)]/3.$$

For these amino acids, the carrier protein yields are as follows: A = 103 pmol, K = 65 pmol, and I = 65 pmol. For these same amino acids, the peptide/carrier conjugate yields: A = 206 pmol, K = 125 pmol, and I = 135 pmol.

From these values, the relative amount of carrier protein in the conjugate versus the unconjugated carrier protein (SF) can be calculated as follows: $(206/103 + 125/65 + 135/65)/3 = 2.0$, indicating that there is twice as much carrier protein in the peptide/carrier conjugate hydrolysate as in the carrier-protein hydrolysate.

3. Calculate the moles of peptide present in the conjugate by subtracting the moles of amino acid present in the carrier from the moles of amino acid present in the conjugate. Choose ≥ 3 amino acids present in the peptide. The relative amount (SF) of protein present in the carrier protein versus the amount in the conjugate as calculated in step 2, must also be considered as follows:

$$pmol\ peptide\ in\ conjugate = \{[pmol\ G\ in\ conjugate - (SF \times pmol\ G\ in\ carrier)] + [pmol\ L\ in\ conjugate - (SF \times pmol\ L\ in\ carrier)] + [pmol\ R\ in\ conjugate - (SF \times pmol\ R\ in\ carrier)]\}/3$$

Therefore, the amount of peptide in the conjugate hydrolysate for the example shown in Table 9.4.1, calculated using the amino acids G, L, and R, is $\{[215 - (2 \times 95)] + [133 - (2 \times 55)] + [177 - (2 \times 75)]\}/3 = 25$ pmol.

4. Calculate the number of moles of protein in the conjugate hydrolysate as follows:

$$pmol\ carrier\ protein\ in\ conjugate = \frac{\text{total pmol carrier protein amino acids}}{\text{molecular weight of carrier protein}} \times 110$$

where total pmol carrier protein amino acids = SF \times (total amino acid composition of carrier in pmol) and 110 is the average molecular weight of an amino acid.

In this example, there are 1850 pmol of carrier protein amino acids in the conjugate; therefore, $1850\ pmol \times (110/100,000) = 2.04$ pmol carrier protein in conjugate.

5. Determine the ratio of peptide to carrier protein as follows:

$$\frac{\text{molecules peptide in conjugate}}{\text{pmol peptide in conjugate}} / \frac{\text{molecules carrier protein in conjugate}}{\text{pmol carrier protein in conjugate}} =$$

Using the values calculated in steps 3 and 4, the result is: $25\ pmol\ peptide\ in\ conjugate / 2.04\ pmol\ carrier\ protein\ in\ conjugate = 12.2$ molecules peptide in conjugate per molecule carrier protein in conjugate.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Cysteine standard stock solution

Dissolve 26.3 mg cysteine hydrochloride monohydrate in 100 ml of 0.1 M sodium phosphate, pH 8.0 (APPENDIX 2). Prepare immediately before use.

Ellman's reagent solution

Dissolve 4 mg Ellman's reagent, 5,5'-dithio-bis-(2-nitrobenzoic acid) (Pierce), in 1 ml of 0.1 M sodium phosphate, pH 8.0 (APPENDIX 2). Prepare immediately before use.

Glutaraldehyde solution, 0.15%

Add 30 μ l of 25% aqueous glutaraldehyde solution to 5 ml of 50 mM sodium borate buffer, pH 8.0 (pH adjusted with HCl). Prepare fresh and use immediately. If the glutaraldehyde precipitates, check the pH. It should not be above 8.0; a slightly lower pH can be used (pH 7 to 8).

CAUTION: *Glutaraldehyde is a sensitizing agent that should be handled in a hood and only according to the recommendations in the Material Safety Data Sheet. When mixing solutions or performing reactions, keep the container covered to prevent vapors from escaping into the atmosphere.*

Guanidine-HCl, 6 M

Dissolve 1 g guanidine-HCl in 1 ml of 0.05 M sodium phosphate, pH 7.0 (APPENDIX 2). Store up to several weeks at room temperature.

The resulting 1.8-ml solution should be ~0.025 M phosphate/6 M guanidine-HCl at pH 7.0.

COMMENTARY

Background Information

Synthetic peptides are linear arrays of amino acids that in most instances possess a random structure in solution. While it is not difficult to produce antipeptide antibodies, it does not necessarily follow that the antibodies will recognize a protein containing the same stretch of sequence found in the peptide. In order for this to occur, the amino acids in the protein must be oriented to the antibody in a way similar to that of the synthetic peptide. This generally requires three basic features of the protein: (1) that the stretch of sequence be exposed to solvent; (2) that the sequence be a continuous stretch of amino acids; and (3) that it not possess a higher-order structure that renders it unrecognizable by the antibody population.

The large number of model protein structures now available indicate that almost all of the ionized groups in water-soluble proteins are on the protein surface. Asp, Glu, Lys, and Arg residues, on the average, comprise 27% of the protein surface and only ~4% of the protein interior. The fraction of residues that are at least 95% buried range from 0.36 to 0.60 for nonpolar residues and 0.01 to 0.23 for polar residues. Only 1% of Arg and 3% of Lys residues fall

into the 95% buried range (Creighton, 1993). Therefore, it is reasonable to expect solvent-exposed areas of proteins to display relatively high levels of polar and charged residues, particularly Arg and Lys.

Proteins display three kinds of secondary structure: α -helices, β -sheets, and turns or loops. Turns or loops generally connect elements of α -helices and β -sheets, and can either fit one of several rather strict motifs with recognizable hydrogen bonding patterns or be of a more extended, random nature. These turn or loop structures appear to be most useful for antibody production because they tend to be found on the surface of proteins connecting larger arrays of helices and sheets, and they consist of continuous stretches of amino acids. Although many amino acid residues in helices and sheets are also exposed at the surface, the regular geometry of amino acids contained within them makes them less suitable for this purpose. For instance, in β -sheet structures the side chain of each successive amino acid in the β -sheet strand points in the opposite direction to the ones immediately preceding and following it. Thus, even if the amino acid side chains are not predominantly buried in the interior of

the protein, only every other side chain is exposed on the same surface of the sheet. This can hinder recognition by an antibody produced with a linear peptide capable of assuming a more random structure. A similar situation exists for α -helices. Although the change in direction of the side chains of successive amino acids is perhaps not as abrupt as in β -sheets, only approximately every third or fourth side chain is found on the same surface of the helix. Epitopes in proteins have been identified in amphipathic helices, but unless the synthetic peptide assumes a similar helical structure in solution, recognition by the antibody may be problematic.

These considerations have led to more useful methods for predicting sequences that will produce antibodies recognizing intact proteins. A variety of different indices that predict hydrophilicity or hydrophobicity and secondary structure are available. In addition, predictive methods based on segmental mobility, side chain accessibility, and sequence variability (see Van Regenmortel et al., 1988) have also been proposed. All of these methods generally tend to yield similar results, but it must be noted that these procedures were developed for (and work best with) water-soluble proteins composed of a single globular structure. Additional complications can arise with multisubunit proteins, where normally exposed structures may be shielded by subunit interactions, or membrane proteins with large sections shielded from the solvent.

The method presented in this unit utilizes the correlation between the hydrophilic character of a peptide sequence (Kyte and Doolittle, 1982) and its propensity to form β -turn structures (Chou and Fasman, 1974). Free access to these and many other algorithms is provided at the ExpASY Web site of the University of Geneva at <http://expasy.org.tools>.

After selection of the peptide sequence, an effective immunogen is generally produced by coupling the peptide to a carrier protein or by synthesizing a multiple antigenic peptide (MAP), with four or eight identical peptides assembled simultaneously on the α and ϵ amines of the terminal lysines of a branched core (see Fig. 11.16.3).

Critical Parameters

Analyzing protein sequences with algorithms or tables of assigned values for amino acids is a well-established procedure, but evaluating these results and selecting the candidate

sequences requires some consideration. To take full advantage of the results, choose areas of sequence that give the maximum values for the properties being evaluated and that also show the highest degree of residue-by-residue correlation. In other words, choose areas of maximum amplitude where the centers of the peaks correspond to the same sequence with a divergence of no more than two to three residues. Examples of this are given in Figure 11.16.2, which shows results from the method presented in Basic Protocol 1 for the sequence shown in Figure 11.16.1. The top panel in Figure 11.16.2 predicts β -turns as calculated by the method of Chou and Fasman (1974). The bottom panel is a prediction of hydrophobicity using the parameters of Kyte and Doolittle (1982). The data are analyzed by looking for areas of high turn propensity (maximum positive deflection in the top panel) and high hydrophilicity (maximum negative deflection in the bottom panel). The shaded areas in Figure 11.16.2 designate three segments that meet these criteria. Note that the maximum and minimum values of these three stretches of protein sequence correlate very well. Additional areas of high hydrophilicity (bottom panel) are found near residues 64, 132, 137, 149, and 345, although the β -turn values of these secondary candidates are not as high as those of the three shaded areas. Two equally hydrophilic areas at residues 49 and 299 correspond to downward deflections in the β -turn profile and are thus not good candidates based on this analysis.

Many different chemistries are available for coupling synthetic peptides to carrier proteins to produce effective immunogens (Van Regenmortel et al., 1988). In many cases, however, side reactions or incompatibilities in chemistry between the coupling agent and the residues present in the peptide can be problematic. In order to simplify the process and present the greatest probability of success in most cases, only a few coupling methods are presented in this unit. In this regard, the recommended coupling procedure is cross-linking of the peptide via cysteine residues to keyhole limpet hemocyanin (KLH) with the heterobifunctional reagent *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester (MBS; see Basic Protocol 3). This effective method has enjoyed great success and can be used for virtually any peptide. The one caveat is that it is not recommended for peptides with internal cysteine residues, since they will also link to the carrier. It is also critical, when coupling with MBS through an added terminal

cysteine residue, that the sulfhydryl group of the peptide be present in the free or reduced form (see Support Protocols 1 and 2).

In addition to MBS coupling, other procedures commonly used (see Alternate Protocols 3 and 4) are included as alternatives for use in special situations, but these are not recommended as a general alternative to MBS because they are more restrictive and have the potential for undesirable side reactions. Glutaraldehyde coupling (see Alternate Protocol 3) should not be used with peptides containing internal Lys, Cys, Tyr, or His residues and, since it is a homobifunctional reagent, cross-linking of the peptide to itself and the carrier to itself can occur. The latter lowers antigenicity and can result in extensive aggregation and precipitation of the carrier. 1-ethyl-3(3-dimethylaminopropyl) carbodiimide (EDC; see Alternate Protocol 4) is a water-soluble carbodiimide and should not be used with peptides containing internal Lys, Glu, Asp, Tyr, or Cys residues. Alternate Protocol 5 describes a simple photochemical coupling strategy (Gorka et al., 1989).

Another good alternative for most peptides is the production of a multiple antigenic peptide (MAP; see Alternate Protocol 2). With this method the composition of the peptide is not a concern beyond its potential solubility properties. In most cases, since hydrophilic sequences are selected, this also is not a major problem. Both four- and eight-branched MAPs have been found to be effective. However, four-branched MAPs are recommended because they are less prone to synthesis problems and are easier to characterize.

As with any synthetic peptide, the product must be well characterized before use. If the peptide is not what it was intended to be, this decreases the probability of generating antibodies that will recognize the protein. At the very least, check synthetic peptides for homogeneity by analytical HPLC and correct mass by mass spectrometry (see *UNITS 10.21 & 10.22*). Characterization of MAP can be more problematic due to their multibranching nature (Mints et al., 1997): HPLC and mass spectrometric analysis can be compromised by the presence of four to eight peptide chains per molecule, each of which may have only a small percentage of modification at any particular residue but which in the aggregate contribute to broad spectra. However, this feature of MAPs usually does not tend to compromise their ability to form antigens of the proper peptide since the correct sequence is usually present in high enough concentration that a significant amount of specific antibody is pro-

duced among the polyclonal population. Amino acid analysis (*UNIT 10.1B*), which is less sensitive to multiple small differences, tends to give a reasonable assessment of the MAP integrity.

Anticipated Results

The methods outlined in this unit produce an effective polyclonal antiserum against an intact protein from a single peptide sequence ~50% to 70% of the time. Therefore, it is advisable to prepare two or three different peptides from a given protein to increase the probability of at least one of them being effective.

Time Considerations

Computer-assisted analysis of a protein sequence and inspection of the data to select several candidate sequences takes from 5 to 30 min. Manual analysis of a protein sequence can take several hours but can certainly be accomplished in <1 day. Selection of peptide design and manner of synthesis as well as selection of a coupling method will take <1 hr. Actual preparation of the peptide can be accomplished in 3 to 4 days, but this may vary depending on the turnaround time of the synthetic laboratory. Coupling a synthetic peptide to a carrier protein takes from 1 to 2 days. Although not covered in this unit, production of the antisera will vary with the animal and protocol used, but generally requires 2 to 3 months. It is therefore advisable to inject several animals with different peptides at one time.

Literature Cited

- Chou, P.Y. and Fasman, G.D. 1974. Prediction of protein conformation. *Biochemistry* 13:222-245.
- Creighton, T.E. 1993. *Proteins: Structure and Molecular Properties*, 2nd ed. W.H. Freeman, New York.
- Gailit, J. 1993. Restoring free sulfhydryl groups in synthetic peptides. *Anal. Biochem.* 214:334-335.
- Getz, E.B., Xiao, M. Chakrabarty, T., Cooke, R., and Selvin, P.R. 1999. A comparison between the sulfhydryl reductants Tris(2-carboxyethyl) phosphine and dithiothreitol for use in protein biochemistry. *Anal. Biochem.* 272:73-80.
- Gorka, J., McCourt, D.W., and Schwartz, B.D. 1989. Automated synthesis of a C-terminal photoprobe using combined Fmoc and *t*-Boc synthesis strategies on a single automated peptide synthesizer. *Peptide Res.* 2:376-380.
- Kyte, J. and Doolittle, R.F. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105-132.
- Mints, L., Hogue Angeletti, R., and Nieves, E. 1997. Analysis of MAPS peptides. *ABRF News* 8:22-26.

- Posnett, D.N., McGrath, H., and Tam, J.P. 1988. A novel method for producing anti-peptide antibodies. Production of site-specific antibodies to the T-cell antigen beta-chain. *J. Biol. Chem.* 263:1719-1725.
- Tam, J.P. 1988. Synthetic peptide vaccine design: Synthesis and properties of a high density multiple antigenic peptide system. *Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413.
- Van Regenmortel, M.H.V., Briand, J.P., Muller, S., and Plaué, S. 1988. Synthetic polypeptides as antigens. *In* Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 19 (R.H. Burdon and P.H. van Knippenberg, eds.). Elsevier/North-Holland, Amsterdam.

Key References

Van Regenmortel et al., 1988. See above.

Comprehensive treatment of theory and method.

Tam, J.P. 1988. High density multiple antigen peptide system for preparation of anti-peptide antibodies. *Methods Enzymol.* 168:7-15.

Original methods article for MAPs.

Internet Resource

<http://expasy.org/tools.html>

Web site for programs to analyze protein sequences.

Contributed by Gregory A. Grant
Washington University School of Medicine
St. Louis, Missouri

DETERMINATION OF SPECIFIC ANTIBODY TITER AND ISOTYPE

Determination of the Specific Antibody Titer

The amount of specific antibody present in polyclonal antiserum, ascites fluid, or hybridoma supernatant can be quantitated by either solid-phase radioimmunoassay (RIA) or by direct enzyme-linked immunosorbent assay (ELISA; UNIT 11.2). In the solid-phase assay described here (see Basic Protocol), serially diluted antiserum is incubated in microtiter wells previously coated with the relevant antigen. Bound antibody is detected by employing ^{125}I -labeled anti-immunoglobulin antibodies. The amount of specific antibody in the antiserum is then determined from a standard curve generated with a specific antibody of known concentration. The unknown antiserum and the standard antibody are assayed in parallel. The support protocols describe the chloramine T (see Support Protocol 1) and IODO-GEN (see Support Protocol 2) procedures for radioiodination of the anti-immunoglobulin reagent. The use of the solid-phase RIA procedure to determine the light-chain (κ and λ) and heavy-chain (γ , μ , α) isotypes present in polyclonal antisera and fluids containing monoclonal antibodies is also described as Support Protocol 3.

SOLID-PHASE RADIOIMMUNOASSAY (RIA) FOR DETERMINATION OF ANTIBODY TITER

In this assay, a specific antigen is used to coat the wells of the microtiter plate.

Materials

- Antigen
- Coating buffer (see recipe)
- Control antigen (non-cross-reactive protein)
- Phosphate-buffered saline (PBS; APPENDIX 2)
- ^{125}I anti-immunoglobulin reagent (see recipe; Ab must be specific for the species in which the test antibody was raised—e.g., ^{125}I goat anti-mouse for mouse hybridoma or ^{125}I goat anti-rabbit for rabbit sera)
- Wash buffer (UNIT 11.2)
- 1% bovine serum albumin in phosphate-buffered saline (BSA/PBS)
- Standardized antibody solution (5 mg/ml in diluting buffer)
- Diluting buffer (see recipe)
- 96-well microtiter plates compatible with γ counter such as Wallac ScintiStrips
- Repeater pipet (e.g., Eppendorf with disposable Combi-tips)
- Multichannel pipet
- Automated γ counter that counts 96-well plates such as a Wallac MicroBeta TriLux

1. Prepare antigen and control antigen solutions (0.005 to 0.01 mg/ml) in coating buffer.

The optimal antigen concentration will depend on the nature of the antigen employed. To determine the optimal coating concentration, incubate a relatively concentrated dilution of antiserum (1:100 dilution) in microtiter wells previously coated with serial dilutions of antigen. Bound antibody is then detected as described below in this protocol. Counts per minute (cpm) are plotted against coating antigen concentration. The optimal coating concentration is that at which coating antigen is saturating (i.e., at an antigen concentration corresponding to a point on the plateau of the curve).

The antigen need not be pure, although for accurate and reproducible antibody titers, it is desirable to use a preparation containing a consistent antigen concentration. Bovine serum albumin (BSA) is a frequently used control antigen; however, the best control antigen is an immunologically non-cross-reactive protein of similar molecular weight to the antigen of interest.

If the antigen is stable, the antigen-coated plate can be prepared in advance and stored in PBS for up to 1 week at 4°C.

2. Pipet 50 μ l of appropriately diluted antigen into columns 1, 2, and 3, and columns 7, 8, and 9 of rows A to H of a 96-well microtiter plate, as shown in Figure 11.17.1. Add control antigen to wells 4, 5, and 6, and 10, 11, and 12 of rows A to H. Cover plate to avoid evaporation and incubate 2 hr at room temperature or overnight at 4°C.

It is desirable to assay each antiserum dilution in triplicate. In general, eight serial dilutions of antiserum are sufficient.

The volumes of antigen, antibody, and [¹²⁵I]anti-immunoglobulin reagent may be halved if antigen and/or antibody are precious.

3. Remove coating antigens by shaking into sink. Add 100 μ l wash buffer and incubate 5 min at room temperature. Remove wash buffer and repeat wash cycle twice.

If the antigen is difficult to obtain, the coating antigen solution may be removed with a repeater pipet, stored at -20°C, and reused. The length of time it may be stored and the number of times it may be reused depend on the antigen, but it should not be used >5 times

4. Pipet 100 μ l BSA/PBS into each coated well, cover plate, and incubate 1 hr at room temperature or overnight at 4°C. Shake out blocking buffer and wash three times with wash buffer as described above.

An alternative to BSA/PBS is 5% preimmune horse serum prepared in PBS.

5. Make serial dilutions of polyclonal antiserum (or ascites fluid or hybridoma supernatant) and standardized antibody solution (5 mg/ml) in diluting buffer. Dilutions can be made in small glass tubes or plastic 1-ml microcentrifuge tubes. Initially, a 1:5 dilution of antiserum is made, followed by seven five-fold dilutions as follows:

- a. Label tubes from 1 to 8.
- b. Pipet 400 μ l diluting buffer into all tubes.
- c. Pipet 100 μ l undiluted antiserum into tube 1 and mix well by pipetting up and down without producing foam.
- d. Transfer 100 μ l diluted antiserum from tube 1 to tube 2 and mix well. Continue to transfer 100 μ l diluted antiserum from tubes 2 to 8 as described. Finally, remove and discard 100 μ l from tube 8.

6. Follow step 5 for serial dilutions of the standard antibody solution using tubes 9 to 16.

The standardized antibody solution (5 mg/ml) is ideally a purified solution of antibody prepared by affinity chromatography (UNIT 11.11), with the same antigen specificity as the antiserum of interest and raised in the same host animal species. The concentration of the antibody is determined from the A_{280} and adjusted to 5 mg/ml. (A solution of 1 mg/ml immunoglobulin has an A_{280} of 1.4).

7. Pipet 50 μ l diluted antiserum from tubes 1 to 8 into the first six columns of rows A to H, respectively, as indicated in Figure 11.17.1. In a similar fashion, pipet 50 μ l of each dilution of standard antibody from tubes 9 to 16 into wells 7 to 12 of rows A to H, respectively. Cover plate and incubate 2 hr at room temperature.

Dilutions may be made while coating antigen is incubating on the plate and then stored at 4°C until required.

		Specific antigen			Control antigen			Specific antigen			Control antigen			
Antibody dilution factor	Tube	1	2	3	4	5	6	7	8	9	10	11	12	
5	1	A												
25	2	B												
125	3	C												
625	4	D												
3.1×10^3	5	E												
1.6×10^4	6	F												
7.8×10^4	7	G												
3.9×10^5	8	H												
			Antiserum						Standard antibody					

Figure 11.17.1 The microtiter plate setup for the determination of specific antibody titer. Wells are initially coated with the specific antigen or the control antigen. Serial dilutions of specific antiserum, ascites fluid or hybridoma supernatant are then incubated in the coated wells. Bound antibody is detected using a [125 I]anti-immunoglobulin reagent.

8. Shake diluted antiserum and standard antibody from the plate and wash three times with wash buffer as described in step 3.
9. Pipet 50 μ l of appropriately diluted [125 I]anti-immunoglobulin reagent into each well. Cover plate and incubate ≥ 4 hr at room temperature, or overnight at 4°C.

Steps 9 to 12 should be carried out in a well-ventilated fume hood designated for radioactive work.

A repeater pipet is convenient for delivering the 50- μ l aliquots of [125 I]anti-immunoglobulin reagent.
10. Remove radioactive supernatant from wells using a multichannel pipet and discard waste into appropriate radioactive waste container. Add 100 μ l wash buffer to each well and gently pipet up and down three times with the multichannel pipet. Remove wash buffer and discard it into radioactive waste container. Repeat this procedure twice. Wash wells five more times with 100 μ l wash buffer per well.

The final five washes may be poured into the sink, as the radioactivity will be negligible; however, first check radiation safety regulations.
11. Allow plates to dry for 4 hr at room temperature or under a heat lamp for 30 min.
12. Count each plate in gamma counter for one minute.
13. Determine amount of specific antibody bound from antiserum of interest and from standardized antibody solution as follows:
 - a. Calculate average cpm (cpm_{av}) for triplicate values obtained for each antibody dilution.
 - b. For each dilution of antiserum, subtract cpm_{av} obtained when a given dilution is incubated in wells coated with control antigen (background) from cpm_{av} measured on wells coated with specific antigen ($\text{cpm}_{\text{bound}}$).
 - c. Plot $\text{cpm}_{\text{bound}}$ versus \log_{10} (antiserum dilution; see Fig. 11.17.2A).
 - d. Plot $\text{cpm}_{\text{bound}}$ versus \log_{10} (standard antibody in mg/ml; see Fig. 11.17.2B).
 - e. Select a value of $\text{cpm}_{\text{bound}}$ from linear section of binding curve (y in Fig. 11.17.2A) constructed for the antiserum of interest. Translate this value onto the standard

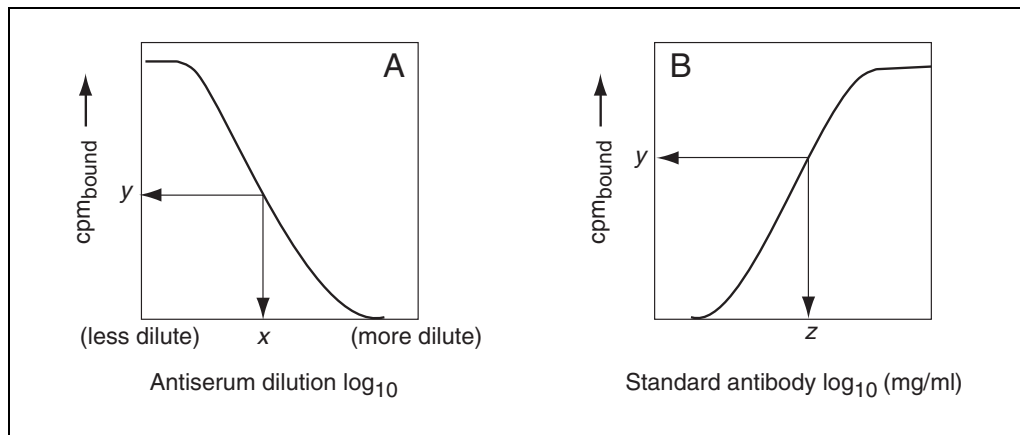


Figure 11.17.2 Theoretical binding curves for the determination of specific antibody titer. **(A)** The experimental curve. Serial dilutions of specific antiserum, ascites fluid, or hybridoma supernatant are incubated in antigen-coated wells. **(B)** The standard curve. Serial dilutions of the standard antibody are incubated in the antigen-coated wells. The specific antibody titer is calculated as shown: Specific antibody titer = $x \times z$, where x is the antibody dilution that yields y $\text{cpm}_{\text{bound}}$ and z is the standard antibody concentration that yields y $\text{cpm}_{\text{bound}}$.

curve and determine standard antibody concentration corresponding to this $\text{cpm}_{\text{bound}}$ (z in Fig. 11.17.2B). This concentration is equal to that of specific antibody present in polyclonal antiserum at the dilution giving the selected $\text{cpm}_{\text{bound}}$.

- f. Calculate initial specific antibody concentration (i.e., titer) in undiluted serum by multiplying standard antibody concentration (z) derived from selected $\text{cpm}_{\text{bound}}$ (y) by corresponding dilution factor of antiserum (x in Fig. 11.17.2A).

SUPPORT PROTOCOL 1

IODINATION OF ANTI-IMMUNOGLOBULIN REAGENT USING CHLORAMINE T

This protocol is a clean and efficient method for covalently linking ^{125}I to tyrosine residues in the protein of interest. Unbound ^{125}I is quenched with a saturated tyrosine solution and separated from the protein-bound fraction by passage down a small desalting column.

CAUTION: Under no circumstances should the iodination reaction be carried out on an open bench. Use a well-ventilated fume hood (confirmed by radiation safety department); do not use tissue culture hood. During the reaction, significant quantities of free $^{125}\text{I}_2$ are liberated which can concentrate in the thyroid gland. Therefore, as ^{125}I is a strong gamma emitter, it cannot be stressed strongly enough that extreme care must be taken to contain the radioactive reaction products within a well-ventilated fume hood so that they are isolated from the environment.

Ensure that lead bricks line the front of the fume hood to protect against gamma emission. Additionally, a lab coat and two layers of protective gloves should be worn throughout the iodination procedure.

It is recommended that the experimenter have his or her thyroid tested by the institutional radioactive safety department before and within 1 week after carrying out this reaction. If all safety precautions are followed correctly, contamination will not occur.

Determination of the Specific Antibody Titer

11.17.4

Materials

Sephadex G-25 (Amersham Pharmacia Biotech), washed and equilibrated in PBS
(see APPENDIX 2 for PBS)
DEAE-Sephadex (Amersham Pharmacia Biotech), washed and equilibrated in PBS
Mixed bed resin (Bio-Rad #AG 501-X8), washed and equilibrated in PBS
1% bovine serum albumin in phosphate-buffered saline (BSA/PBS)
Protein to be iodinated (1 mg/ml, preferably in iodinating buffer)
Iodinating buffer (see recipe)
Saturated tyrosine solution (see recipe)
Na[¹²⁵I] (1 mCi in 10 to 5 µl of 0.1 M NaOH, specific activity 17 Ci/mg; NEN
Life Sciences)
Chloramine T solution (see recipe)
Plastic tubing (3 mm i.d., 10 cm long) with clamps
Silanized glass wool (UNIT 5.6)
Small column stand
1-ml syringes equipped with 22-G needles
Plastic collection tubes with caps
Lead pig

1. Prepare desalting column by pushing plastic tubing securely onto tip of a Pasteur pipet and attaching a small clamp to tubing. Place a small wad of siliconized glass wool into neck of pipet. Clamp pipet to a small column stand.

The glass wool is the bed on which the chromatography resins are packed. Therefore, it is important that the wool is packed at the neck of the pipet tightly enough to stop the resin beads from passing through, yet loosely enough to allow the column to flow freely.

2. Layer 1.5 to 2.0 cm Sephadex G-25 onto glass wool. After the Sephadex G-25 has settled, gently layer 3 to 4 cm DEAE-Sephadex on top.

The DEAE-Sephadex removes aggregated protein.

3. Allow DEAE-Sephadex to settle and then layer 1.5 to 2.0 cm mixed bed resin on top.

The mixed bed resin adsorbs any free ¹²⁵I⁻ remaining in the solution.

4. Pour 50 ml BSA/PBS down the column.

The 1% BSA in the buffer saturates the resin bed and inhibits the nonspecific adsorption of iodinated protein during the desalting procedure.

5. Draw up 50 µl saturated tyrosine solution into a 22-G needle connected to a 1-ml syringe and 250 µl saturated tyrosine solution into a second 22-G needle connected to a 1-ml syringe.

Do not remove air bubble from inside syringe barrel. The air bubble should separate the base of the plunger from the solution so that the full contents of the syringe can be expelled when required.

6. Take all reagents and necessary equipment to a well-ventilated fume hood that has been designated previously for radioactive use (preferably one used specifically for iodination of proteins using ¹²⁵I; see cautionary statements preceding materials list). Place the following equipment in the hood ready for use:
 - a. The desalting column and collection tubes. Remove BSA/PBS from top of desalting column. Remove caps from collection tubes and place them close by the column.
 - b. A number of Pasteur pipets and pipet bulbs for column elution.

- c. The two syringes containing saturated tyrosine solution (from step 5).
 - d. One 22-G needle.
 - e. The lead pig containing a vial of Na[¹²⁵I].
 - f. A box of protective gloves beside the hood, within easy reach.
7. Into a third 1-ml syringe (connected to a 22-G needle), take up the following: 50 μl protein solution (do not remove air bubble between base of plunger and protein solution), 40 μl air space, and 30 μl chloramine T solution.

This step can be carried out outside the fume hood, but it must be done immediately prior to commencing the reaction. Because chloramine T is a strong oxidant, the protein must be in contact with it for as short a time as possible.

8. Remove vial of Na[¹²⁵I] from lead pig and remove cap from inner container, leaving the small glass vial containing the radioactivity inside the inner container.

The inner container will provide shielding from the radioactivity. Under no circumstances should the open vial be removed from the fume hood.

9. Insert 22-G needle into rubber lid of vial containing Na[¹²⁵I].

The 22-G needle acts as a vent when the protein/chloramine T solution is injected into the vial. Ensure that the needle tip lies just below the lid, well above the liquid in the vial.

10. Inject protein/chloramine T solution into vial and allow reaction to proceed one minute. Stop the reaction by injecting 50 μl saturated tyrosine solution into vial.
11. Remove contents of vial using the syringe that contained 50 μl saturated tyrosine solution and load contents carefully onto the desalting column.
12. Rinse vial by injecting 250 μl saturated tyrosine solution (from step 5). Remove this wash solution and load it onto the column.
13. Allow sample to penetrate into column bed and elute with BSA/PBS. Collect twenty 0.5-ml fractions.
14. Count 1-μl aliquots from each fraction on a gamma counter to determine where the iodinated protein peak has eluted. Pool fractions containing the radiolabeled protein and store them at -20°C in a lead pig.

The ¹²⁵I-labeled protein is usually eluted within the first 2 to 3 ml.

The sample may be stored in aliquots (e.g., 100 μl), depending on the volume required to make the appropriate dilutions for the RIA.

CAUTION: Dispose of all contaminated solutions and equipment, including the desalting column, in appropriate radioactive waste receptacles.

SUPPORT PROTOCOL 2

IODINATION OF ANTI-IMMUNOGLOBULIN REAGENT USING IODO-GEN

IODO-GEN (1,3,4,6-tetrachloro-3α,6α-diphenylglycouril) can be substituted for chloramine T as the iodination agent. This reagent is reported to be gentler than chloramine T. Because IODO-GEN has a longer reaction time (10 to 15 min) than chloramine T (1 min), the investigator can vary this parameter for labeling unstable proteins. The chloramine T support protocol (see Support Protocol 1) is modified according to the steps below when using IODO-GEN.

Materials

iodo-GEN (Pierce #28600T)
Methylene chloride
iodo-GEN-coated glass test tubes (Pierce #28601T or see recipe)
Drierite

Additional reagents and equipment for iodination of anti-immunoglobulin reagent using chloramine T (see Support Protocol 1)

1. Proceed with steps 1 to 6 of the chloramine T protocol (see Support Protocol 1).
2. Rinse iodo-GEN-coated glass tube with 200 μ l iodinating buffer to remove any loose microscopic flakes of iodo-GEN. If flakes are visible, reagent has not been properly plated and another tube should be tested. Add 100 μ l protein solution.
3. Proceed with step 8 of the chloramine T protocol (see Support Protocol 1). Insert third syringe into vial containing Na[¹²⁵I] and transfer contents into the iodo-GEN-coated vial containing protein to be labeled. Incubate mixture for 10 to 15 min at room temperature.
4. Proceed as described in steps 11 to 14 of the chloramine T protocol (see Support Protocol 1).

DETERMINATION OF ANTIBODY ISOTYPES

The optimal dilutions of the isotyping reagents (anti- μ , γ , α , λ , and κ) are determined using a set of standard monoclonal antibodies of the μ , γ , α , λ , or κ isotype, respectively. This protocol describes the determination of the optimal dilution of an anti- μ isotyping reagent, but the same protocol can be used for optimization of isotyping reagents of all specificities. Optimization assays for all isotyping reagents must be carried out in parallel.

After establishing the optimal dilutions of isotyping reagents, their use in determining the isotypes of mouse antiserum is described for the μ , γ_1 , γ_2 , and γ_3 heavy chains and κ and λ light chains. Initially, the specific antibody of interest is bound to the wells of a 96-well microtiter plate coated previously with the antigen. The appropriately diluted isotyping reagents are then incubated within the wells. The extent of specific binding is determined using a second [¹²⁵I]anti-immunoglobulin antibody specific for the animal species from which the isotyping reagents were derived.

Additional Materials (also see Basic Protocol)

Isotype standards (0.1 μ g/ml)
Isotyping reagents (see recipe)

Determine the optimal dilutions for isotyping reagents

1. Pipet 25 μ l of μ , γ , α , λ , and κ isotype standards (0.1 μ g/ml) and control antigen (0.1 μ g/ml) into all wells of columns 1 and 2, 3 and 4, 5 and 6, 7 and 8, 9 and 10, and 11 and 12, respectively, of a microtiter plate (see Fig. 11.17.3 for diagram). Cover plate to avoid evaporation and incubate 2 hr at room temperature or overnight at 4°C (see Fig. 11.17.4).
2. Remove isotype standards from plate by flicking contents into the sink. Add 100 μ l wash buffer to each well, cover plate, and incubate 5 min at room temperature. Remove wash buffer and repeat the wash step twice.

SUPPORT PROTOCOL 3

												Control antigen	
		μ	γ	α	λ						κ		
Isotyping reagent dilution	Tube	1	2	3	4	5	6	7	8	9	10	11	12
10^2	1	A											
5×10^2	2	B											
2.5×10^3	3	C											
1.3×10^4	4	D											
6.3×10^4	5	E											
3.1×10^5	6	F											
1.6×10^6	7	G											
7.8×10^6	8	H											

Figure 11.17.3 Microtiter plate setup for the determination of the optimal dilutions of the isotyping reagents. The wells are coated with the appropriate isotype standards. Serial dilutions of the isotyping reagent are incubated in the coated wells. Bound isotyping reagent is detected using a [125 I]anti-immunoglobulin reagent specific for the animal species from which the isotyping reagent is derived.

- Pipet 100 μ l blocking buffer onto each coated well, cover plate, and incubate 1 hr at room temperature or overnight at 4°C. Shake out blocking buffer and wash three times as in step 2.
 - Label a set of 1.5-ml microcentrifuge tubes 1 to 8.
 - Add 990 μ l diluting buffer to tube 1 and 800 μ l diluting buffer to tubes 2 to 8.
 - Pipet 10 μ l anti- μ isotyping reagent into tube 1 and mix well by pipetting up and down.
 - Remove 200 μ l of 1:100 dilution anti- μ isotyping reagent from tube 1, add it to tube 2, and mix well. Continue to transfer 200 μ l diluted anti- μ reagent from tube 2 to 8 as described. Finally, remove and discard 200 μ l from tube 8.
 - Pipet 50 μ l of each dilution of the anti- μ reagent from tubes 1 to 8 into wells of rows A to H, respectively, and incubate plate 2 hr at room temperature.
- Each dilution of the isotyping reagent is assayed in duplicate.*
- Remove isotyping reagent from plate and wash plate as described in step 2.
 - Carry out the remainder of the assay (see Basic Protocol, steps 9 to 12).
 - Average the cpm of the duplicate values (cpm_{av}) for each dilution of isotyping reagent bound to each isotype standard.
 - Subtract the cpm_{av} value obtained for a given dilution of anti- μ reagent incubated in wells coated with control antigen from the cpm_{av} value calculated for that dilution incubated in wells coated with each of the isotype standards, respectively ($\text{cpm}_{\text{bound}}$).
 - Choose the dilution of anti- μ isotyping reagent at which minimal, if any, cross-reactivity with the nonspecific isotype standards is observed.

The dilutions chosen for all isotyping reagents must yield equivalent cpm_{av} values to ensure comparable sensitivity of detection for each isotype within a given antibody population.

		Isotyping reagents											
		anti- μ		anti- γ_1		anti- γ_2		anti- γ_3		anti- κ		anti- λ	
		1	2	3	4	5	6	7	8	9	10	11	12
specific antigen	A	2580	2612	347	302	106	80	91	76	1386	1235	83	67
	B	80	123	2464	2398	44	98	105	125	1068	1039	55	69
control antigen	C	106	54	81	93	24	51	118	91	59	78	96	148
	D	103	51	108	102	81	68	62	80	75	37	97	66

Figure 11.17.4 Microtiter plate setup for the determination of the isotype/s of the specific antibody population. The specific antibody (10 $\mu\text{g/ml}$) is incubated in the antigen (specific or control) coated wells. The appropriately diluted isotyping reagent is then added to the wells. Bound isotyping reagent is detected using a [^{125}I]anti-immunoglobulin reagent specific for the animal species from which the isotyping reagent is derived. Typical (data as cpm [^{125}I]anti-immunoglobulin antibody bound) is shown for a mouse $\mu\kappa$ antibody in rows A and C and a mouse g1k antibody in rows B and D.

Determine the isotypes

- Pipet 50 μl coating antigen (i.e., specific antigen) into all wells of rows A and B of the microtiter plate and 50 μl control antigen into all wells of rows C and D. Cover plate to avoid evaporation and incubate 2 hr at room temperature, or overnight at 4°C.

Each isotyping reagent will be assayed in duplicate. The volumes of antigen, antibody, and [^{125}I]anti-immunoglobulin reagent may be reduced if antigens and/or antibody are in limited supply.

- Remove coating antigens and block wells (see Basic Protocol, steps 3 and 4).

In this assay only a single specific antibody concentration (10 $\mu\text{g/ml}$) is required.

- Add 50 μl antiserum (diluted to give a specific antibody concentration of 10 $\mu\text{g/ml}$) to each well. Incubate plate 2 hr at room temperature. After incubation, shake out antibody, add 100 μl wash buffer to each well, and incubate plate 5 min at room temperature. Remove wash buffer and repeat wash step twice.
- Pipet 50 μl of each of the appropriately diluted isotyping reagents (in duplicate) into the designated wells: anti- μ into rows 1 and 2, anti- γ_1 into rows 3 and 4, anti- γ_2 into rows 5 and 6, anti- γ_3 into rows 7 and 8, anti- κ into rows 9 and 10, and anti- λ into rows 11 and 12. Cover plate and incubate 2 hr at room temperature.

When using anti- γ isotyping reagents, nonspecific binding to the μ heavy chain can be significantly reduced by adding 5 μl of 0.1 M 2-mercaptoethanol to the 50 μl of antigen-specific antibody in each well prior to the 2-hr incubation period.

- Wash wells as described in step 16 above. Complete remainder of assay (see Basic Protocol, steps 9 to 12).
- Determine amount of each isotyping reagent specifically bound in the wells as follows:
 - Average the cpm for each set of duplicates (cpm_{av}).
 - For each isotype specificity, subtract the cpm_{av} observed in wells coated with control antigen from the cpm_{av} observed in wells coated with specific antigen ($\text{cpm}_{\text{bound}}$).

20. Compare cpm_{av} values obtained for each heavy- and light-chain isotyping reagent to determine the immunoglobulin isotypes present. See Figure 11.17.4 for typical results for a mouse $\gamma_1\kappa$ antibody and a mouse $\mu\kappa$ antibody.

REAGENTS AND SOLUTIONS

Chloramine T solution

Prepare an 0.56 mg/ml solution by adding 5.6 mg chloramine T to 10 ml iodinating buffer (recipe below). Use immediately.

Coating buffer (0.1 M carbonate buffer, pH 9.6)

Stock A, 0.2 M: 21.2 g Na_2CO_3 (anhydrous) per liter H_2O

Stock B, 0.2 M: 16.8 g NaHCO_3 per liter H_2O

For pH 9.6, add 80 ml A + 170 ml B + 250 ml H_2O

Diluting buffer

0.5 g Tween 20

2.5 g bovine serum albumin (BSA; 0.25%)

1.0 g sodium azide (0.1%)

Add phosphate-buffered saline (PBS; APPENDIX 2) to 1 liter

2.5 g BSA may be replaced with 10 ml preimmune horse serum (1% final concentration).

NOTE: Do not add sodium azide if the HRPO-antibody conjugate is being used.

[^{125}I]anti-immunoglobulin reagent

Choose anti-immunoglobulin reagent specific for the animal species from which the antibody population being assayed is derived and iodinate according to Support Protocol 1 or 2. Dilute the stock solution of [^{125}I]anti-immunoglobulin reagent in diluting buffer to give 20,000 cpm/50 μl .

5 ml of diluted [^{125}I]anti-immunoglobulin reagent are required per 96-well plate.

Exercise extreme caution at all times when using radioactive materials, no matter how dilute the sample. Follow all institutional radiation safety regulations with respect to pertinent biological hazards and radioactive containment procedures.

Iodinating buffer (0.5 M phosphate buffer, pH 7.5)

Stock A: 137.03 g $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ per liter H_2O

Stock B: 68.99 g NaH_2PO_4 (anhydrous) per liter H_2O

Add 400 ml stock A to 100 ml stock B and check pH

iodo-GEN-coated glass test tubes

Pipet 100 μl methylene chloride solution containing 10 μg IODO-GEN into a 12 \times 75 mm glass tube (suitable for 100 μl protein at 1 mg/ml) and leave overnight to evaporate in a well-ventilated fume hood under a gentle stream of nitrogen.

IODO-GEN is insoluble in aqueous media and will remain immobilized on the reaction vessel wall after the iodinated sample has been removed. It is stable and can be stored coated on the glass up to 2 months at -20°C over Drierite in a desiccator.

Isotyping reagents

Anti-rabbit isotyping reagents are available from ICN Immunobiologicals. Anti-mouse isotyping reagents are available from a number of immunochemical suppliers.

Saturated tyrosine solution

Mix 50 g tyrosine and 50 ml iodinating buffer in a test tube. Cap the tube and incubate 30 min in a boiling water bath. Allow to cool before use.

COMMENTARY

Background Information

The solid-phase RIA described here provides a very sensitive method for the determination of the specific antibody titer within a given antiserum. The protein antigen is immobilized on the walls of the microtiter wells, which results in the denaturation of a significant proportion of the antigen. Thus, the solid-phase RIA allows the detection of antibodies that recognize both the native and denatured forms of the antigen. The immobilization of the protein antigen greatly enhances the antigen-antibody interaction, making the assay relatively affinity independent. The specific antibody titer determined using this assay includes both conformationally dependent and independent antibody specificities, as well as antibodies whose affinity for the antigen is extremely low ($K_a = 10^1$ to 10^2 M^{-1}). In addition, the solid-phase RIA is extremely useful for studying antibody specificity when quantities of protein antigen are limited.

In the Basic Protocol, [^{125}I]anti-immunoglobulin reagent is in the RIA as an alternative to the ELISA (UNIT 11.4), which can also be used to detect bound antibody. A wide variety of species-specific anti-immunoglobulin reagents can now be easily obtained from many biotechnology companies and can be labeled with ^{125}I using either of the procedures described in the support protocols. In some cases, it may be more convenient to use ^{125}I -labeled protein A in place of the anti-immunoglobulin reagent. However, protein A does not bind all species of immunoglobulin. A list of immunoglobulin species to which protein A binds is shown in Table 11.17.1. A more extensive list is provided by Goding (1978).

Support Protocol 1 describes the covalent coupling of ^{125}I to protein molecules using chloramine T. Although this procedure yields high specific activities (12 to 16 $\mu Ci/\mu g$ of immunoglobulin), given a sufficient number of surface-exposed tyrosine residues, chloramine T is a strong oxidant and can cause significant

losses in antibody activity after limited exposure (i.e., >1 min). In addition, some proteins may be much more susceptible to denaturation through oxidation than immunoglobulins. In such cases, a less harsh method of radio-iodination should be used: the IODO-GEN procedure described in Support Protocol 2 or the lactoperoxidase procedure (see Hudson and Hay, 1980; Mishell and Shiigi, 1980). Both of these procedures also link ^{125}I to proteins through tyrosines residues but yield lower specific activities. Finally, the addition of a large atom, such as the iodine atom, to the protein surface can potentially lead to some degree of conformational change, regardless of the radio-iodination procedure employed. Therefore, it is important that no more than one atom of ^{125}I is incorporated per protein molecule. The above protocols yield a substitution level well below this level.

Monoclonal antibodies are used in a variety of techniques such as immunoprecipitation (UNIT 10.16), western blotting (UNIT 10.8), ELISAs (UNIT 11.2), radioimmunoassays, and immunohistochemistry. In such techniques a second antibody reagent, tagged with a readily detectable molecule (e.g., ^{125}I , alkaline phosphatase, or horseradish peroxidase) is often required to enhance the sensitivity and/or selectability of the technique. As a consequence, it is necessary to determine the isotype of the monoclonal antibody to be used. Since polyclonal antisera are also frequently employed in these techniques, it is desirable to know the distribution of isotypes within the specific antibody population. As a result of its monoclonality, the isotype of a given monoclonal antibody is singular; however, a polyclonal antiserum will usually contain specific antibodies bearing a variety of isotypes. The heavy-chain isotypes found in rabbits are μ , γ_1 , γ_2 , α_1 , α_2 , and ϵ . Similar isotypes are found in mice and humans, although these species have more subtypes in the μ , γ , α , and ϵ families. The light-chain isotypes in all mammals are either λ or κ but in

Table 11.17.1 Binding of Immunoglobulin Classes and Subclasses to Protein A

Species	Classes which bind well	Classes which bind weakly
Rabbit	IgG	IgM
Mouse	IgG _{2a} IgG _{2b} IgG ₃	IgG ₁ IgM
Human	IgG ₁ IgG ₂ IgG ₄ IgA ₂	IgM
Rat	IgG	IgM

varying ratios. Thus, although these assays are described for polyclonal rabbit antisera, they may also be used to isotype sera or monoclonal antibodies from other species. The RIA will give only a qualitative estimation of the spectrum of specific isotypes present in an antiserum. A more quantitative estimation of the specific isotypes present may be obtained by affinity chromatography, as described in *UNIT 11.11*, using anti-isotype reagents conjugated to affinity resins.

Critical Parameters

For the accurate determination of specific antibody titer using the solid-phase RIA, it is critical that the standardized antibody solution be of similar composition to that of the antiserum being titered. First, both standard and unknown antibody populations must be derived from the same animal species and be specific for the same antigen. Second, because the different immunoglobulin isotypes will vary in their affinity for antigen and in their interaction with the anti-immunoglobulin reagent, the isotype composition within the standard antibody population and antiserum of interest should be equivalent. Third, it is important that the standard antibody retains maximum binding activity. A significant population of denatured or otherwise inactive standard antibody will result in low binding activity per milligram of protein (standard antibody) and therefore an underestimation of specific antibody titer.

All anti-immunoglobulin reagents cross-react to a certain extent with different isotypes. Therefore, preliminary assays must be carried out on each isotyping reagent to determine the optimal dilution that will allow minimal nonspecific cross-reactivity while yielding a sensitivity equivalent to that attained for the other isotyping reagents.

To achieve unambiguous and reproducible results, the optimal working dilutions of the isotyping reagents must be determined accurately. It is important that a working dilution be chosen at which little, if any, cross-reactivity with other nonspecific isotypes occurs. At the same time, the dilution of each isotyping reagent should be chosen such that the maximum numbers of cpm specifically bound by each reagent, at the chosen dilution, are equivalent. This ensures that the RIA will be able to detect the individual immunoglobulin isotypes, via the specific recognition of the respective isotyping reagents, with equivalent sensitivity.

Troubleshooting

The solid-phase RIA is very reliable as long as all reagents are prepared carefully. However, two potential problems may arise. First, a high level of nonspecific binding may occur when assaying the antiserum of interest. For the higher dilutions of antiserum (i.e., 1:125 to 1:390,000), the nonspecific binding should be <5% of the maximum specific binding. If higher values are obtained, it is likely that other serum components are adhering to the walls of the plastic wells or to the antigen and subsequently interacting nonspecifically with the RIA reagents. Protein aggregates can be removed from the antiserum by ultracentrifugation for 1 hr at $145,000 \times g$. Initially diluting the antiserum 1:10 or 1:25, instead of 1:5, also may help to reduce the background. If these procedures do not alleviate the problem, it may be necessary to remove the nonspecifically bound serum proteins using DEAE-Affi-Gel Blue chromatography (see *UNIT 11.14*, Alternate Protocol).

A second potential problem is the denaturation of the anti-immunoglobulin reagent during the iodination procedure. As stated in Support Protocol 1, chloramine T is a strong oxidant which should not be in contact with protein for any longer than specified. If no radioactivity is detected when the wells are counted in the gamma-counter, the ^{125}I -labeled anti-immunoglobulin reagent has most likely lost its binding activity. Repeat the iodination procedure ensuring that the antiimmunoglobulin reagent is in contact with the chloramine T for as short a time as possible. Alternatively, the IODO-GEN procedure (see Support Protocol 2) can be used.

In addition, the following should be considered: although the working dilutions of the isotyping reagents are optimized so that cross-reactivity with nonspecific isotype standards are minimal, significant cross-reactivity among the heavy-chain isotypes often occurs when assaying polyclonal antisera.

Anticipated Results

The titer of specific antibody in a given antiserum is dependent on the immunogenicity of the antigen (see introduction to this chapter). A titer of 5 to 10 mg/ml of specific antibody can be expected for a plant or bacterial antigen. A mammalian protein antigen may produce a titer significantly less than this, 0.1 to 2 mg/ml depending on the evolutionary relationship between the protein antigen and the analogous protein in the host.

Time Considerations

The solid-phase RIA can be completed in 1 or 2 days, depending on the length of the coating and blocking incubations. Labeling the anti-immunoglobulin reagent with ^{125}I can be completed within 60 to 90 min. The actual reaction time is 1 to 5 min.

Literature Cited

- Goding, J.W. 1978. Use of staphylococcal protein A as an immunological reagent. *J. Immunol. Meth.* 20:241-253.
- Hudson, L. and Hay, F.C. 1980. *Practical Immunology*, 2nd ed., Blackwell Scientific Publishers, Oxford.

Key Reference

- Mishell, B.B. and Shiigi, S.M. (eds.) 1980. Solid-phase radioimmune assays. *In Selected Methods in Cellular Immunology*, pp. 373-397. W.H. Freeman, N.Y.

Contributed by Helen M. Cooper
Walter and Eliza Hall Institute
Melbourne, Australia

Yvonne Paterson
University of Pennsylvania
Philadelphia, Pennsylvania

PREPARATION AND USE OF SPECIALIZED ANTIBODIES

SECTION VI

Identification of Polyol-Responsive Monoclonal Antibodies for Use in Immunoaffinity Chromatography

UNIT 11.18

The extreme specificity of an antigen-antibody reaction can be used as a powerful step in protein purification processes. For most applications, the antibody is covalently coupled to a supporting matrix (such as cross-linked agarose). A crude or partially purified preparation of material containing the protein antigen is passed over the antibody-matrix, resulting in the selective capture of the target protein. However, one of the limitations of this technique, known as immunoaffinity chromatography (UNIT 10.11A), has been that high-affinity antigen-antibody complexes are difficult to dissociate, often leading to inactivation of the protein product during elution from the immobilized antibody.

Some antigen-antibody complexes are dissociated in the presence of a combination of a low-molecular-weight polyhydroxylated compound (polyol) and a nonchaotropic salt. These conditions seem to be generally nondenaturing and, in some cases, even protein-stabilizing. This type of antibody is designated “polyol-responsive.” These antibodies can be easily identified and isolated as monoclonal antibodies (MAbs) from a typical fusion, using standard hybridoma procedures. They have proven to be very valuable reagents for the immunoaffinity purification of active, labile, multi-subunit protein complexes.

The basic protocol describes pertinent steps in hybridoma production (steps 1 to 6), screening of MAbs for polyol responsiveness (steps 7 to 17), optimization of eluting conditions (steps 18 to 23), preparation of the MAb and the MAb-conjugated resin (steps 24 to 40), use of the MAb-conjugated resin in immunoaffinity chromatography (steps 41 to 47), and, finally, regeneration and storage of the MAb-conjugated resin (steps 48 to 50).

**BASIC
PROTOCOL**

Screening of the MAbs for polyol responsiveness is essential to identifying the appropriate MAb early in the process and can be performed even at the master-well stage during standard hybridoma production, after the fusion but before cloning. Alternatively, an existing collection of MAbs can be screened by this method before a large amount of MAb is produced for preparation of the MAb-conjugated resin. The screening procedure is a modified enzyme-linked-immunosorbent assay (ELISA), termed “ELISA-elution assay.” The ELISA-elution assay can also be used to help optimize the elution conditions.

Materials

- Purified or partially purified immunogen
- Mice for immunization (UNIT 11.4)
- 1× phosphate buffered saline (PBS; see recipe for 10×)
- HAT medium (UNIT 11.6)
- 1% nonfat dry milk (see recipe)
- PBST (see recipe)
- 1× TE buffer, pH 7.9 (see recipe for 10×)
- 1× TE buffer, pH 7.9 (see recipe for 10×) containing salt-polyol at appropriate concentration
- Polyols (e.g., ethylene glycol, propylene glycol, 2, 3-butanediol, glycerol)

Salts (e.g., ammonium sulfate, sodium chloride, potassium glutamate)
 Anti-mouse IgG conjugated to horseradish peroxidase (HRPO)
 OPD substrate (see recipe; or use other suitable substrate for HRPO)
 1 M H₂SO₄
 Mouse sub-isotyping kit (e.g., American Qualex; also see *UNIT 11.3*)
 Protein A-agarose (Repligen) or DE-52 cellulose (Whatman)
 Cyanogen bromide-activated Sepharose 4B (Sigma, or see *UNIT 10.16*)
 1 mM HCl
 Bicarbonate coupling buffer (see recipe)
 1 M ethanolamine, pH 8.3 (see recipe)
 Acetate washing buffer (see recipe)
 2% (w/v) NaN₃
 Crude material containing protein of interest
 1× TE buffer, pH 7.9 (see recipe for 10×) containing 2 M potassium thiocyanate (KSCN)
 1× TE buffer, pH 7.9 (see recipe for 10×) containing 100 mM NaCl and 0.02% NaN₃

 96-well cell culture plates
 Microtiter plate reader capable of reading at 490 nm
 30-ml sintered glass filter and vacuum flask
 End-over-end rotator
 10-ml disposable polypropylene column (Bio-Rad)

 Additional reagents and equipment for production of monoclonal antibodies (*UNITS 11.4-11.11*), ELISA (*UNIT 11.2*), immunoblotting (*UNIT 10.8*), preparation of antibody-Sepharose (*UNIT 10.16*), dialysis (*APPENDIX 3C*), and protein assays (*UNIT 10.1*)

Prepare mouse hybridomas

1. Immunize mice with desired immunogen (*UNIT 11.4*) and monitor the response by an indirect ELISA (*UNIT 11.2*) and immunoblotting (*UNIT 10.8*). Identify mice that have a titer well over 1:1000.
Several injections, administered over 4 to 6 weeks, are usually necessary to achieve such a titer.
2. About 1 month after the last immunization dose, inject the mouse to be used for the fusion intraperitoneally or intravenously with soluble immunogen dissolved in 1× PBS. Use ~2 to 3 times the amount of immunogen that was used in the last immunization dose.
3. Three days later, harvest the spleen, prepare a single-cell suspension, and fuse with the desired plasmacytoma cell line (*UNIT 11.7*).
4. Plate cells, suspended in HAT medium, into 96-well cell culture plates (~10 to 20 plates depending upon past history of fusion frequency) as described in *UNIT 11.7*.
5. About 10 to 14 days after the fusion, screen the hybridomas for specific antibody production, using 100 μl of cell culture medium and a standard ELISA procedure (*UNIT 11.2*).
*If an impure antigen is used, a secondary screen, using a microimmunoblot assay (*UNIT 10.8*) might be necessary.*
6. Add back 100 μl of HAT medium to the positive wells. After 1 to 2 days of incubation, perform the ELISA-elution assay as described in the following steps.

Screen for polyol responsiveness by the ELISA-elution assay

7. Coat a microtiter plate with antigen (30 to 100 ng/well, contained in 50 μl of PBS) for 1 hr at room temperature. Remove the antigen solution and then block with 1%

nonfat dry milk at 200 μl /well by incubating for at least 2 hr at room temperature, or overnight at 4°C.

8. Remove the blocking solution. Add 50 μl of the cell culture medium from antigen-specific antibody positive wells to each of two wells of the microtiter plate.

One well will serve as the buffer control, the other well will serve as the polyol-responsive test well.

9. Incubate at room temperature 1 to 1.5 hr, then wash wells three times with PBST. Remove residual PBST.
10. To one well, add 100 μl of 1 \times TE buffer, pH 7.9. To the other well add 100 μl of 1 \times TE buffer, pH 7.9, containing the polyol/salt combination.

For screening, a combination of 0.75 M ammonium sulfate and 40% (v/v) propylene glycol works well. This combination seems to identify more polyol-responsive MABs than the 1 M NaCl and 50% (v/v) ethylene glycol originally reported.

11. Incubate 20 min at room temperature, tapping the plate about every 5 min to ensure mixing of the solution in the wells, then wash wells three times with PBST. Remove residual PBST.
12. Add 50 μl of the proper dilution (usually 1:1000) of HRPO-conjugated anti-mouse IgG antibody in 1% non-fat dried milk. Incubate 1 hr at room temperature.
13. Wash plate eight times with PBST. Remove residual PBST.
14. Add 100 μl of OPD substrate. Allow color to develop for 1 to 5 min.

CAUTION: OPD is a carcinogen. Use appropriate care to avoid contact. Other substrates for HRPO can be used, but OPD seems to be the most sensitive.

15. Add 50 μl of 1 M H_2SO_4 to stop the reaction. Stop each control and test reaction in parallel.

Different hybridomas will produce MABs that have different titers and affinities. Therefore, different reaction times (1 to 5 min) might be needed for the individual antibodies. Other substrates may require use of different stop solutions.

16. Read the reactions on a microtiter plate reader at 490 nm.

A MAb is considered to be polyol-responsive if the OD reading of the well treated with the polyol/salt is 50% or less than the OD reading of the control well. The above wavelength is for OPD; other substrates may require a different wavelength.

17. Clone the polyol-responsive MAb-producing hybridomas at least twice by limiting dilution (UNIT 11.8). Freeze the hybridomas for future antibody production (UNIT 11.9).

Save the cell culture supernatant. It can be used for optimizing the elution conditions and determining the isotype of the MAB.

Optimize elution conditions by ELISA-elution assay

18. Coat with antigen and block a microtiter plate (see step 7) for each of the polyol-responsive MABs.

19. Prepare the desired polyol/salt combinations in 1 \times TE buffer, pH 7.9.

These solutions can be prepared at many different concentrations. Generally a range of each polyol and salt is assayed. Convenient salt concentrations are 0 M, 0.1 M, 0.25 M, 0.5 M, and 0.75 M and convenient polyol concentrations are 0%, 10%, 20%, 30%, 40%, and 50% (v/v). It is sometimes possible to use higher salt concentrations, but some salts are not soluble at high concentration in the higher percentages of polyols. These solutions

can be made ahead and stored in the refrigerator, but should be warmed up to room temperature before use.

20. Add 50 μ l of the cell culture supernatant to each well, and incubate 1 to 1.5 hr at room temperature. Wash the wells three times with PBST.
21. Set up a grid on the microtiter plate so that all each polyol/salt combination is tested in duplicate. Run several different combinations of polyol/salt on one plate.
22. Add 100 μ l of each polyol/salt solution to each of two wells. Incubate 20 min at room temperature, tapping the plate about every 5 min to ensure mixing of the solution in the wells. Wash the plate three times with PBST. Remove any residual PBST. Continue with the ELISA-elution assay (steps 12 to 16).
23. Average the OD readings for the duplicate samples and plot the value as a function of polyol/salt concentrations.

A typical set of data is shown in Figure 11.18.1.

Prepare the MAb and MAb-conjugated resin

24. Isotype each of the MAbs using a commercially available isotyping kit (or see UNIT 11.3).

Knowledge of the isotype is helpful for designing the purification of the MAb. The cell culture supernatant from freezing the hybridomas can be used for this assay.

25. Prepare ascites fluid from the hybridoma (UNIT 11.10).

Because some research institutes have placed restrictions on the production of ascites fluid in mice, alternative methods of large-scale production of MAbs should be considered (UNIT

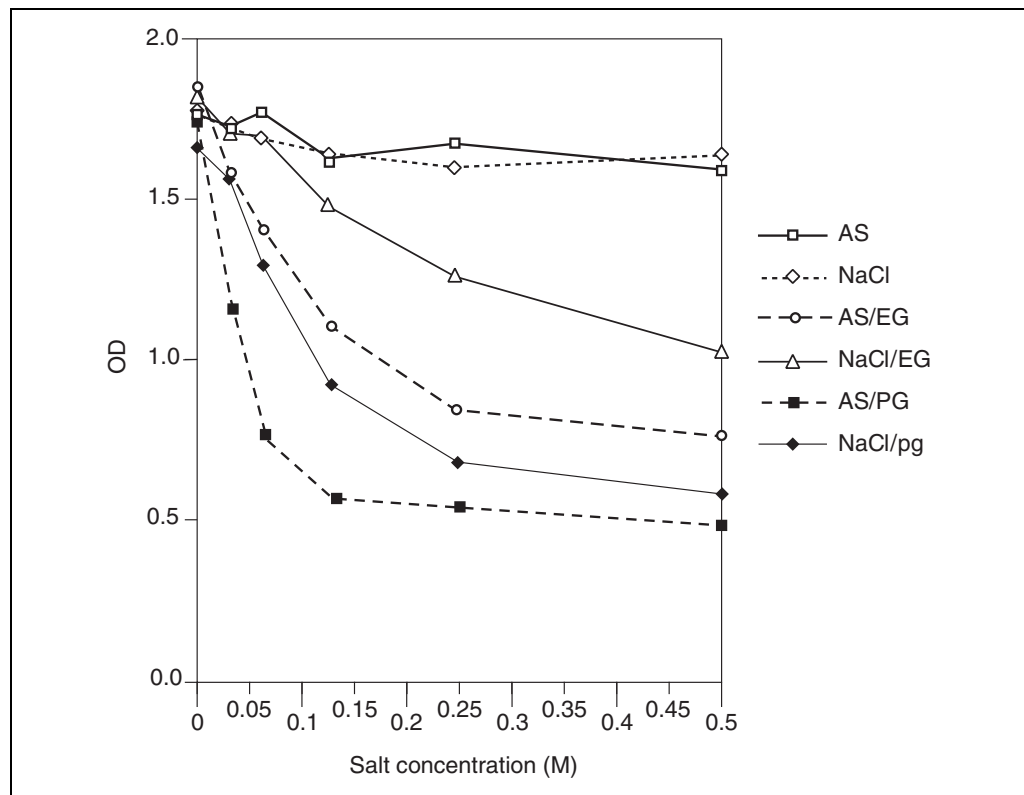


Figure 11.18.1 Optimization of eluting conditions using the ELISA-elution assay. Two different salts at several concentrations were tested with two different polyols. Sodium chloride (NaCl) and ammonium sulfate (AS) were used at the indicated concentrations. Ethylene glycol (EG) and propylene glycol (PG) were used at 40% (v/v).

11.10). The scale of the MAb production should be such that ~5 to 10 mg of purified MAb can be produced.

26. Purify mouse IgG_{2a} and IgG_{2b} on staphylococcal protein A-agarose (UNIT 11.11).

Mouse IgG₁ does not bind well to protein A, but can be purified to ~80% purity by chromatography on DE-52 cellulose. Add saturated (at 4°C) ammonium sulfate solution to achieve a 45% ammonium sulfate cut of the ascites fluid. Stir on ice bath ~30 min. Centrifuge to collect the precipitate, dissolve the precipitate in 1/2 volume of 50 mM Tris-Cl, pH 7.0, and 25 mM NaCl, and dialyze overnight against this buffer. Centrifuge to clarify. Equilibrate a 5-ml column of DE-52 in the same Tris buffer. Apply the material to the DE-52 column and collect the flowthrough fractions. Most mouse IgG₁ MAbs (and some of the IgG_{2a} and IgG_{2b} MAbs) flow through DE-52 under these conditions.

27. Determine the amount of MAb (in mg) by taking an absorbance reading at 280 nm and applying the following formula:

Absorbance × dilution factor/1.38 × volume of MAb solution (ml)

28. Calculate the amount of MAb needed for the conjugation, using an estimate of 2.5 mg MAb per ml of swollen resin.

Slightly lower amounts can be used, but increasing the amount of antibody will not necessarily increase the capacity of the resin. Hybridomas vary in the amount of MAb that is produced in the ascites fluid. This is usually in the range of 1 to 10 mg/ml with most hybridomas producing between 1 to 3 mg/ml. Thus ~5 to 10 ml of ascites fluid will generally suffice for preparation of a useful amount of MAb-conjugated resin.

29. Calculate the amount of cyanogen bromide-activated Sepharose needed to prepare the desired amount of resin.

Other matrices with other coupling chemistries can be used. However, because cyanogen bromide-activated Sepharose performs well and is relatively inexpensive, the protocol is described for this resin. 1 g of dried Sepharose swells to ~3.5 ml of resin. Thus, 1 g would require 8.75 mg of MAb. The volumes described below are useful for 3.5 ml of Sepharose.

30. Dialyze (APPENDIX 3C) the appropriate amount of purified MAb against bicarbonate coupling buffer overnight at 4°C and centrifuge 15 min at 6000 × g, 4°C, to clarify.
31. Swell the appropriate amount of cyanogen bromide-activated Sepharose in 1 mM HCl for ~15 min at room temperature.
32. Wash the Sepharose on a 30-ml sintered glass filter with ~200 ml of 1 mM HCl. Quickly wash the resin with ~20 ml of bicarbonate coupling buffer and transfer to the solution (~10 ml) containing the MAb.
33. Mix the solution, end-over-end, on a laboratory rotator for 2 hr at room temperature.
34. Collect the resin on the sintered glass filter, reserving the flowthrough to determine the amount of MAb coupled. Take the absorbance of the flowthrough at 280 nm and estimate the percentage of coupling, again assuming that 1 mg MAb/ml will give an absorbance at 280 nm of 1.38.
35. Transfer the resin to a solution of 1 M ethanolamine, pH 8.3. For 3.5 ml of resin, use ~10 ml of ethanolamine solution.
36. Mix the solution, end-over-end, on a laboratory rotator for 2 hr at room temperature.
37. Collect the resin on the sintered glass filter and wash with ~100 ml bicarbonate coupling buffer.
38. Wash the resin on the sintered glass filter with ~100 ml acetate washing buffer.

39. Repeat steps 37 and 38 twice. Wash and resuspend in bicarbonate coupling buffer (~10 ml).
40. Add 1/100 vol of 2% NaN₃ and store at 4°C.

Purify target protein on MAb-conjugated resin

41. Prepare the crude material from which the protein is to be purified by an appropriate method to maintain stability of the protein.

Immunoaffinity chromatography (UNIT 10.11A) is generally one step in a protein purification protocol. Frequently, it is the only chromatography step necessary. The initial steps in the purification procedure will vary with the nature of the material. Usually, some initial steps must be taken to remove nucleic acids to lower viscosity. In addition, generally some fractionation step might be used to reduce the volume of the material. See Chapter 10 introduction for guidelines on how to structure a protein purification procedure. The references contain protocols for preparing material from yeast (Edwards et al., 1990), wheat germ and calf thymus (Thompson et al., 1990), and bacteria (Thompson et al., 1992; Thompson and Burgess, 1994).

42. Pack a 10-ml disposable polypropylene column with the antibody-Sepharose (UNIT 10.11A). Apply the crude material to the MAb-conjugated resin.

Alternatively, this step can be performed in batch mode.

43. Wash the resin with the appropriate buffer containing ~0.5 M salt but no polyol.
44. Move the resin to room temperature. If the material had been applied to the MAb-conjugated resin in a batch mode, pour the resin into a small column.
45. Elute the column with 1× TE buffer, pH 7.9, containing the salt/polyol.

Care should be taken to apply the polyol/salt solution gently to a well-packed column to avoid causing the resin to float due to the high density of the polyol/salt solution. For reasons that are not clear, the polyol/salt elution is more effective at room temperature than at 4°C. However, if lability of the protein is a concern, all steps up to the elution step can be performed at 4°C. Fractions containing the protein can also be put in an ice bath as they come off the column. Therefore, the protein is usually at room temperature for less than 30 min.

46. Perform a quick protein assay such as a Bradford assay (UNIT 10.1) and pool the peak fractions.
47. Dialyze the protein (APPENDIX 3C) against two changes of an appropriate buffer, for 2 hr each at 4°C, to remove salt and polyol.

It is necessary to remove the salt and polyol to avoid interference in subsequent biological or biochemical assays. A buffer should be chosen that will maintain the stability and activity of the particular protein.

Regenerate and store MAb-conjugated resins

48. Wash the resin with ~10 ml of 1× TE buffer containing 2 M KSCN.

This can be performed in either batch or column mode, but the MAb-conjugated resin should not be exposed to KSCN for longer than ~20 min.

49. Wash the resin with 1× TE buffer containing 100 mM NaCl and 0.02% NaN₃.
50. Store the MAb-resin at 4°C.

MAb-conjugated Sepharose can be stored up to ~6 months at 4°C. After ~6 months, some leaching of MAb from the resin can be detected. Most MAb-conjugated Sepharose preparations can be used ~10 times. After ~10 uses, the column usually still works but more contaminants seem to be recovered in the product.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Acetate washing buffer

To 900 ml H₂O, add:
13.6 g sodium acetate trihydrate
29.22 g NaCl
Adjust to pH 4.0 at 23°C with glacial acetic acid
Add H₂O to 1000 ml
Store up to 6 months at room temperature

Bicarbonate coupling buffer

To 900 ml H₂O, add:
8.4 g NaHCO₃
29.22 g NaCl
Adjust to pH 8.3 at 23°C with 1 N NaOH
Add H₂O to 1000 ml
Prepare fresh

Citrate buffer, 0.1 M

Solution A: 29.4 g sodium citrate, dihydrate in 1000 ml H₂O
Solution B: 21.0 g citric acid, monohydrate in 1000 ml H₂O
Add solution B to ~500 ml of solution A until pH 5.0 is achieved. Filter sterilize and store up to 1 year at 4°C.

EDTA, 0.2 M, pH 7.9

74.4 g disodium EDTA dihydrate
700 ml H₂O
Adjust to pH 7.9 at 23°C with 10 M NaOH
Add H₂O to 1000 ml

Ethanolamine, 1 M, pH 8.3

To 80 ml of bicarbonate coupling buffer (see recipe), add 6 ml ethanolamine. Adjust pH to 8.3 at 23°C with 6 M HCl. Add coupling buffer to 100 ml. Prepare fresh.

Nonfat dry milk, 1%

5 g nonfat dried milk
50 ml 10× PBS (see recipe)
450 ml H₂O
Store up to 1 week at 4°C

OPD substrate

10 mg *o*-phenylenediamine (OPD)
10 ml H₂O
10 ml 0.1 M citrate buffer (see recipe)
20 µl 30% H₂O₂
Prepare fresh

PBST

100 ml 10× PBS (see recipe)
900 ml H₂O
1 ml Tween 20
Store up to 2 months at room temperature

Phosphate-buffered saline (PBS), pH 7.4, 10×

To 900 ml H₂O, add:
80.0 g NaCl
2.0 g KCl
21.6 g Na₂HPO₄·7H₂O
2.0 g KH₂PO₄
Add H₂O to 1000 ml
Store up to 6 months at room temperature

TE buffer, pH 7.9, 10×

25 ml 2 M Tris·Cl, pH 7.9 (see recipe)
0.5 ml 0.2 M EDTA, pH 7.9 (see recipe)
Add H₂O to 100 ml with H₂O
Store up to 1 year at room temperature

Tris·Cl, 2 M, pH 7.9

242.2 g Tris base
500 ml H₂O
Adjust to pH 7.9 at 23°C with 6 M HCl
Add H₂O to 1000 ml

COMMENTARY

Background Information

The development of practical methods for generating MABs has renewed interest in adapting immunoaffinity chromatography techniques for protein purification. Most early studies focused on low-affinity MABs. However, in order for a MAB to effectively capture an antigen from a dilute solution it must be a high-affinity antibody. Most high-affinity antigen-antibody interactions are difficult to disrupt, requiring conditions that are generally denaturing to proteins, such as extremes of pH or 6 M urea. About 5% to 10% of mouse MABs are polyol-responsive antibodies (Thompson et al., 1992) which release the antigen in the presence of a combination of low-molecular-weight polyol and a nonchaotropic salt. In addition to being gentle and protein-stabilizing, these reagents are very inexpensive and can be used economically on a large scale.

This procedure can be used to purify labile, multi-subunit enzymes, such as eukaryotic RNA polymerase II (Edwards et al., 1990; Thompson et al., 1990) and bacterial RNA polymerase (Thompson et al., 1992; Marshak et al., 1996). In fact, the immunoaffinity chromatography step was the critical purification step for the eventual crystallization of the yeast RNA polymerase II (Cramer et al., 2000). It has also been used for purification of eukaryotic transcription factors expressed in bacteria (Thompson and Burgess, 1994, 1999).

In some cases, it is desirable to follow the immunoaffinity step with a high-resolution ion-exchange step to remove any minor contaminants. Sepharose has some nonspecific binding capacity, and some batches of Sepharose seem to have higher nonspecific binding than others.

Using the ELISA-elution assay, it is possible to screen for polyol responsiveness at the master-well stage of hybridoma production. Thus, if the goal is to isolate this type of MAB, valuable time can be saved by identifying the MAB early in the process. In addition, the ELISA-elution assay can also be used to quickly screen different polyol/salt combinations and to help optimize eluting conditions. Some polyol-responsive MABs respond to a variety of polyol/salt combinations (Thompson et al., 1990, 1992). However, some MABs are more limited in their response (Thompson and Burgess, 1999). It is worth noting that polyol-responsive MABs are not limited to a particular sub-class of mouse IgG.

Critical Parameters and Troubleshooting

One of the most critical parameters of this procedure is that the epitope for the MAB must be accessible in solution. The ELISA-elution assay can generate some false positives. That is, a MAB might appear to be polyol-responsive in this assay, but it does not capture the antigen from solution. This is probably due to distortion

of the antigen on the polystyrene surface, resulting in exposure of epitopes that are not accessible in solution. It is wise to examine the ability of the MAb to remove the antigen from solution early in the identification procedure. This can be done with a simple immunodepletion assay using cell culture supernatant (Thompson and Burgess, 1994).

When a MAb is screened from ascites fluid or other concentrated source of antibody, care must be taken to ensure that the MAb concentration is in the linear range. Otherwise a false negative might result due to overloading of the MAb. Thus, a 50% reduction might not be noticed, resulting in a false negative.

Another critical parameter is that the elution of the antigen from the immobilized MAb is much more efficient at room temperature than at 4°C. However, with careful planning it is possible to limit the time that the protein is exposed to the higher temperature.

High levels of reducing agents in buffers can result in reduction of the disulfides in the MAb. When the antigen is eluted with the polyol/salt, some immunoglobulin chains can be detected in the protein product. To avoid this problem, the material should be applied in buffer that does not contain reductant. In most cases, reductant can be added to each fraction as it comes off the column. In cases where reducing agent is required for enzyme stability, the 0.05 to 0.1 mM DTT can be used in the column buffers, but this might result in decreasing the life of the column and recovery of some MAb chains in the product.

Anticipated Results

Approximately 5% to 10% of MAbs are polyol-responsive; therefore, the probability of isolating one from a standard fusion is quite high. In most cases, the immunoaffinity chromatography step is the only chromatography step necessary for purification of the protein, resulting in a more rapid purification procedure. Because the protein is eluted in salt and a polyol, the product is in a solution that is generally protein-stabilizing.

Time Considerations

The ELISA-elution assay takes ~0.5 days to perform. However, when screening hybridoma

at the master-well stage, the cells need to incubate ~1 day between the initial screen for specific antibody and the ELISA-elution assay. Thus, this adds ~2 days to the hybridoma production.

Production of large amount of MAb by ascites fluid production usually takes 2 to 3 weeks.

The purification and conjugation of the MAb takes 2 days. Because the MAb can start to leach off the resin upon extended storage, the resin should be used within 6 months of conjugation.

Literature Cited

- Cramer, P., Bushnell, D.A., Fu, J., Gnat, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R., and Kornberg, R.D. 2000. Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 288:640-649.
- Edwards, A.M., Darst, S.A., Feaver, W.J., Thompson, N.E., Burgess, R.R., and Kornberg, R.D. 1990. Purification and lipid-layer crystallization of yeast RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.* 87:2122-2126.
- Marshak, D.R., Kadonaga, J.T., Burgess, R.R., Knuth, M.W., Brennan, W.A. Jr., and Lin, S.-H. 1996. Strategies for Protein Purification and Characterization. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Thompson, N.E. and Burgess, R.R. 1994. Purification of recombinant human transcription factor IIB by immunoaffinity chromatography. *Protein Express. Purif.* 5:468-475.
- Thompson, N.E. and Burgess, R.R. 1999. Immunoaffinity purification of the RAP30 subunit of human transcription factor IIF. *Protein Express. Purif.* 17:260-266.
- Thompson, N.E., Aronson, D.B., and Burgess, R.R. 1990. Purification of eukaryotic RNA polymerase II by immunoaffinity chromatography. *J. Biol. Chem.* 265:7069-7077.
- Thompson, N.E., Hager, D.A., and Burgess, R.R. 1992. Isolation and characterization of a polyol-responsive monoclonal antibody useful for gentle purification of *Escherichia coli* RNA polymerase. *Biochemistry* 31:7003-7008.

Contributed by Nancy E. Thompson and
Richard R. Burgess
University of Wisconsin
Madison, Wisconsin

CHAPTER 12

DNA-Protein Interactions

INTRODUCTION

For several decades, DNA-binding proteins have been studied because of their involvement in cellular processes such as replication, recombination, viral integration, and transcription. In recent years, the number of workers interested in the study of these proteins has greatly increased as the advent of recombinant DNA technology has led to the isolation of numerous biologically important genes. Many investigators are interested in how transcription of these genes is controlled in response to environmental or developmental signals, and have started to characterize the DNA sequences responsible for this regulation. This analysis has naturally led to the detection, isolation, and characterization of the proteins that bind to these regulatory sequence elements. This chapter summarizes the techniques currently used to characterize DNA-protein interactions and to isolate DNA-binding proteins.

The preparation of cell extracts from either the nucleus or cytoplasm is often the first step in studying these proteins and their interactions (*UNIT 12.1*). The resulting preparations can be used directly in a variety of functional studies, or as the starting point for the purification of proteins involved in gene regulation. In the past two decades many advances have been made in the technologies used for detecting DNA-binding proteins and for purifying those proteins. Perhaps the most widely applied technology has been the mobility shift assay (*UNIT 12.2*), whereby proteins are detected by their ability to retard the mobility of a labeled DNA fragment through a nondenaturing gel. This technique—originally developed to characterize the interaction of purified prokaryotic proteins with DNA—has been refined to allow detection of numerous DNA-binding proteins in crude extracts from a wide variety of cells. Because the mobility shift assay is simple and rapid, it is typically the method of choice when purifying DNA-binding proteins.

After detection of a DNA-binding protein, it is often necessary to determine its binding specificity. Guesses about the binding site can be verified by using mutated DNA fragments as either cold competitor or as labeled probe in the mobility shift assay. Alternatively, one can chemically modify the DNA template, and ask how these alterations affect binding of a specific protein. The most widely used of these modification, or “interference,” techniques is methylation interference (*UNIT 12.3*). This technique allows direct detection of nucleotides that are in close contact with the binding protein. Here, guanine and adenine residues are methylated at an average of one modification per DNA molecule. Modified DNA molecules bound to a specific protein are then separated from those molecules that are not bound via the mobility shift assay. Molecules that are methylated at nucleotides which are important for binding will be underrepresented in bound DNA.

While methylation interference tends to be more informative than most protection techniques, DNase I protection mapping (or “footprinting”) can provide another level of information (*UNIT 12.4*). In these procedures, protein is first bound to DNA and then the DNA is cleaved either by DNase or chemical agents. Footprinting typically reveals, without perturbing, the general region(s) of DNA to which a protein binds. Footprinting is rapid and sensitive once it is optimized for a particular interaction and can be used as a routine assay in purification. The footprinting assay can also be used as a quantitative

technique to determine both the binding curves for individual proteins, as well as cooperative interactions among proteins bound to multiple sites along DNA.

After detection and preliminary characterization of a DNA-binding protein, investigators frequently wish to purify the protein. This can be accomplished using standard chromatography techniques (Chapter 10) and assaying the fractions for presence of the protein by mobility shift gel electrophoresis or footprinting. Another procedure that has proved to be very powerful in these purification schemes is the use of an affinity column containing large amounts of DNA specifying the binding site for the factor (*UNIT 12.10*).

These affinity columns take advantage of the extraordinary specificity that a protein has for its cognate binding site. When such a column is used, protein is applied to it in the presence of high levels of competing, nonspecific DNA, and proteins recognizing specific sequence motifs on the affinity column partition onto those sequences as the column is loaded. In an alternative procedure, protein is first fractionated on a standard, nonspecific DNA column (e.g., DNA-cellulose), and then applied in high salt to the specific affinity column. Specific protein-DNA complexes are frequently stable to moderate salt concentrations, while many nonspecific protein-DNA complexes are disrupted in high salt. A high degree of purification can therefore be achieved using a standard DNA-cellulose column and a DNA affinity column in tandem.

Similar degrees of purification can be accomplished through the use of biotinylated DNA fragments that contain the binding site for a protein and column matrices that are coated with streptavidin (*UNIT 12.6*). In this procedure, a biotinylated DNA fragment containing a specific sequence motif is mixed with crude or partially purified protein and an excess of competitor DNA. As with the affinity column described above, the specific protein will partition onto the DNA fragment containing the binding site. The specific DNA fragment—as well as the attached protein—is then fished out of this mixture using streptavidin and a column matrix, and the specific protein is eluted with high salt. This protocol is relatively easy to optimize, as the success or failure of several of the steps can be monitored using mobility shift gels. It is also extremely flexible, as one single type of column matrix is compatible with any biotinylated DNA probe, and thus can be used to purify numerous different DNA-binding proteins.

To help ensure that the appropriate protein has been purified, it is useful to know the size of the protein that interacts with a specific DNA sequence. In addition, regulatory proteins can be modified in the cell in response to environmental or developmental signals; therefore, knowing the apparent size of a DNA-binding protein under various conditions is also important as this information can provide insight into regulatory events concerning that protein.

The size of a DNA-binding protein can be determined by covalently crosslinking the protein to its regulatory sequence using UV light and resolving the protein complexes on an SDS polyacrylamide gel (*UNIT 12.5*). This procedure works even for impure proteins in crude extracts. Crosslinking can be accomplished by irradiating a protein solution containing a specific labeled DNA probe. Alternatively, protein-DNA complexes can first be separated on a mobility shift gel and then irradiated. This latter protocol increases confidence that the identified protein is actually part of an appropriate complex. In both protocols, it is critical to verify that the crosslinked protein interacts specifically with a given sequence motif, for example by performing competition studies with unlabeled DNA fragments.

In many instances, an important goal is to clone the gene that encodes a DNA-binding protein, thus allowing detailed genetic characterization of the protein. This can be

accomplished by purifying the binding protein to homogeneity, sequencing it, and using that sequence to identify a cDNA clone (see Chapter 6); however, despite the recent advances in purification techniques, the above approach can be extremely time consuming, particularly if the DNA-binding protein is present at very low levels. An alternative approach is to use the sequence known to be recognized by the protein to directly identify a cDNA clone (*UNITS 12.7 & 12.12*). In *UNIT 12.7*, a library that expresses inserted cDNAs in *E. coli* is plated out, and proteins expressed in plaques produced by recombinant phages are transferred to nitrocellulose filters. These filters are then probed with a specific labeled sequence in order to detect clones that express a given DNA-binding protein. Success of this protocol relies on the assumption that the protein as expressed in *E. coli* will be capable of specifically binding DNA. For example, if the protein binds DNA as a heterodimer, or requires a particular covalent modification to bind to its site, it will not be detectable using this approach.

A second approach to isolate a cDNA clone for a sequence-specific DNA binding protein is to take advantage of the powerful reporter gene and expression strategies that have been developed for use in yeast. *UNIT 12.12* describes the “one hybrid” approach in which a cDNA library of proteins that are fused to a transcriptional activation domain is screened. This library is introduced into yeast strains that contain a reporter gene whose expression is regulated using the candidate DNA binding sequence. Expression of a protein that binds to this sequence will result in activation of the reporter gene, which in turn allows for isolation of the yeast colonies that express the reporter. These yeast colonies are screened further for the presence of proteins that bind to the sequence of interest. This technique offers the advantage that the DNA binding proteins are expressed *in vivo*, and thus binding conditions do not require optimization and the expressed proteins do not need to be denatured and renatured.

Using the above mentioned techniques, one can identify and purify DNA-binding regulatory proteins. Once these proteins are pure it is important to characterize their ability to interact with DNA. Information on both the sequence-specificity for binding and the “strength,” or affinity, of the binding interaction are useful. Binding affinity can be measured using DNase footprinting (*UNIT 12.4*), however in many instances it is simpler to use filter binding (*UNIT 12.8*). Double-stranded DNA flows through nitrocellulose, while protein or protein-DNA complexes are retained. Filter binding provides a simple and sensitive means of characterizing the interaction of purified protein with a known regulatory sequence. Competition experiments using filter binding allow rigorous determination of the relative affinities that several sequences have for one protein. In addition, kinetic measurements can be performed that allow a good estimate of the binding constant and off-rate for any given protein-DNA interaction.

Filter binding is also a useful procedure for determining the DNA-binding properties of pure protein with unknown functions. Frequently (e.g., with developmentally important genes or with oncogenes) a gene product that is biologically important is known to reside in the nucleus, but its precise function is obscure. Filter binding can be used to screen regions of DNA for those sequences that interact with these proteins. If such sequences are identified by filter binding, more detailed analyses—such as footprinting or methylation interference—can be used to further characterize the binding site.

In many instances, it is worthwhile to identify the optimal DNA sequence for binding by a specific protein. Strategies that use PCR have been developed to “select” DNA fragments that bind to a given protein from a large collection of DNA fragments (*UNIT 12.11*). Initially, a large group of DNA fragments is bound to the protein of interest and bound fragments are separated and amplified by PCR. The procedure is repeated several times until a pool of DNA fragments are isolated that can bind to the protein of interest.

Determining the sequences of a number of the individual DNA fragment in this pool, and then comparing those sequences, usually yields a consensus DNA binding sequence.

Once a cloned gene encoding a DNA-binding protein is identified, it is possible to synthesize radiolabeled protein by *in vitro* transcription and translation (*UNIT 10.7*). The resulting labeled protein can then be used to detect and analyze DNA-protein interaction (*UNIT 12.9*). The striking advantage of this approach is the possibility for constructing mutations in the cloned gene which can affect the DNA binding properties of the expressed gene.

The protocols described in this chapter can therefore be used to detect and characterize specific DNA-protein interactions, to purify specific DNA-binding proteins, and to clone the genes for these proteins. Using these protocols, one can start with a defined regulatory sequence motif and isolate the gene that encodes the factor responsible for regulation. The characteristics of the regulatory protein can be determined, and the clone of the regulatory gene can be dissected to define functional domains of the regulatory protein. These procedures have been widely applied in mammalian systems where there is very little that can be done to dissect gene regulation by classical genetic means, and have developed to the point where it is feasible, though certainly still difficult, to dissect complex regulatory loops at the molecular level.

Robert E. Kingston

Preparation of Nuclear and Cytoplasmic Extracts from Mammalian Cells

Extracts prepared from the isolated nuclei of cultured cells are functional in accurate *in vitro* transcription and mRNA processing. Thus, such extracts can be used directly for functional studies and as the starting material for purification of the proteins involved in these processes. In the basic protocol described below, reproducible extracts from different preparations of nuclei are generated. The conditions described are optimized to produce transcriptionally active extracts from HeLa cells. The first support protocol describes how to optimize the basic protocol to increase the yield of specific proteins or activities from different cell types. The second support protocol describes the preparation of the cytoplasmic (S-100) fraction.

BASIC PROTOCOL

PREPARING NUCLEAR EXTRACTS

To prepare nuclear extracts, tissue culture cells are collected, washed, and suspended in hypotonic buffer. The swollen cells are homogenized and nuclei are pelleted. The cytoplasmic fraction is removed and nuclei are resuspended in a low-salt buffer. Gentle dropwise addition of a high-salt buffer then releases soluble proteins from the nuclei (without lysing the nuclei). Following extraction, the nuclei are removed by centrifugation, the nuclear extract (supernatant) is dialyzed into a moderate salt solution, and any precipitated protein is removed by centrifugation.

Materials

- Mammalian (i.e. HeLa) cells from spinner cultures or monolayer cultures
- Phosphate-buffered saline (PBS; *APPENDIX 2*)
- Hypotonic buffer
- Low-salt buffer
- High-salt buffer with 1.2 M KCl
- Dialysis buffer
- Liquid nitrogen
- Beckman JS-4.2 rotor (or equivalent)
- 50-ml graduated conical polypropylene centrifuge tubes (or 15-ml tubes for smaller extract volumes)
- Glass Dounce homogenizer (type B pestle)
- Beckman JA-20 rotor (or equivalent)
- Magnetic stirrer *or* tiltboard
- Dialysis membrane tubing ($\leq 14,000$ MWCO; *APPENDIX 3*)
- Conductivity meter
- Additional reagents and equipment for trypan blue exclusion (*UNIT 11.5*) and Bradford protein assay (*UNIT 10.1*)

NOTE: Perform this procedure at 0° to 4°C, preferably in a cold room. Use precooled buffers and equipment. All centrifugations are done at 4°C with precooled rotors.

Isolate the nuclei

- 1a. *Collect cells from spinner cultures.* Centrifuge cells in 1-liter plastic bottles 20 min in a Beckman JS-4.2 rotor at 3000 rpm ($1850 \times g$). Decant the supernatants and discard. Pool cells in 50-ml graduated conical centrifuge tubes (one tube for every 2 to 3 liters of cells). Proceed to step 2.

The density of the spinner culture should be $5-10 \times 10^8$ cells/liter.

- 1b. *Collect cells from monolayer cultures.* Remove the culture medium from confluent

monolayer cultures. Wash the cells by pipetting sufficient PBS to cover them, swirling gently, and pouring off the PBS. Scrape the cells into fresh PBS and pool in a graduated conical centrifuge tube. Proceed to step 2.

2. Pellet the cells by centrifuging 10 min in a JS-4.2 rotor at 3000 rpm. Proceed to step 3a for spinner cultures or step 3b for monolayer cultures.
- 3a. *Spinner cultures.* Decant the supernatants and discard. Using the graduations on the tube, measure the packed cell volume (pcv). Resuspend the cells in a volume of PBS ~5 times the pcv. Centrifuge the cells 10 min in a JS-4.2 rotor at 3000 rpm. Proceed to step 4.
- 3b. *Monolayer cultures.* Decant the supernatants and discard. Using the graduations on the tube, measure the pcv. Proceed to step 4.
4. Rapidly resuspend the cell pellets in a volume of hypotonic buffer ~5 times the pcv. Centrifuge the cells 5 min in a JS-4.2 rotor at 3000 rpm and discard supernatant.

Do this step quickly because proteins can leak out of the cell at this point and be discarded with the supernatant. This step removes salt from the PBS solution so that efficient swelling can occur in the next step; however, some swelling will occur during this step.

5. Resuspend the packed cells in hypotonic buffer to a final volume of 3 times the original pcv (step 3) and allow to swell on ice 10 min.

For example, if an original pcv of 10 ml has swelled to 20 ml in step 4, only 10 ml of additional buffer is required at this step. The cells should swell at least 2-fold.

6. Transfer the cells to a glass Dounce homogenizer. Homogenize with ten up-and-down strokes, using a type B pestle.

Perform the homogenization slowly, especially the down strokes. After homogenization, check for cell lysis in a microscope. Lysis can be observed by addition of trypan blue to an aliquot of cells (UNIT 11.5). The dye is excluded from intact cells but stains the nuclei of lysed cells. Lysis should be >80% to 90%.

7. Transfer cells to centrifuge tubes. Collect the nuclei by centrifuging 15 min in a JS-4.2 rotor at 4000 rpm (3300 × g). Remove the supernatant and save for S-100 cytoplasmic extract preparation (second support protocol).

Extract the nuclei

8. Using the graduations on the tubes, measure the packed nuclear volume (pnv) from step 7. Resuspend the nuclei in a volume of low-salt buffer equal to ½ pnv.

Resuspension of the nuclei with a small volume of low-salt buffer allows thorough rapid mixing of the nuclei during the addition of the high-salt buffer (step 9). If necessary, clumps of nuclei can be resuspended with one or two strokes in a Dounce homogenizer using a type B pestle.

9. In a dropwise fashion, while stirring gently, add a volume of high-salt buffer equal to ½ the pnv (from step 8). Homogenize in Dounce if necessary.

The high-salt buffer must be added dropwise with frequent or continuous mixing. If it is added too quickly, local concentration of salt can become high and some nuclei will lyse. After adding high-salt buffer, the nuclei can be homogenized with two to five up-and-down strokes in a glass Dounce homogenizer (type B pestle) to prevent clumping in large volumes. It is not always necessary for very small volumes (2 to 3 ml).

The final concentration of potassium chloride should be ~300 mM.

10. Allow the nuclei to extract for 30 min with continuous gentle mixing.

Mixing can be done using very gentle stirring on a magnetic stirrer or by tilting on a tiltboard.

11. Pellet the extracted nuclei by centrifuging 30 min in a Beckman JA-20 rotor at 14,500 rpm ($25,000 \times g$). Draw off the resulting supernatant and measure its conductivity (see step 12). This is the nuclear extract.

Dialyze and store the extract

12. Place the nuclear extract in dialysis tubing. Seal at least one end of the dialysis bag with a clip; the bag can then be opened to compare the conductivity of the contents with that of the dialysis buffer to determine whether further dialysis is necessary. Dialyze against 50 vol of dialysis buffer until the conductivities of the extract and buffer are equal (i.e., when the extract reaches 100 mM KCl).

Conductivity is checked by diluting 5 to 10 μ l of the extract to 1 ml water. Read the conductivity of this dilution directly with a conductivity meter and compare it to that of an equivalent dilution of dialysis buffer.

Dialyze the extract for the minimum amount of time needed to achieve the desired salt concentration: 75 to 100 ml of nuclear extract require ~5 hr of dialysis, whereas 2 ml of extract requires <1 hr. Use the support protocol for optimizing nuclear extraction to determine this. To decrease the dialysis time, use tubing with the largest reasonable surface area for the volume. Dialyzing against dialysis buffer without KCl can also reduce dialysis time. In this instance, the conductivity of the extract should be checked frequently and the dialysis stopped when the conductivity equals that of the dialysis buffer (with 100 mM KCl). This method requires extreme care as a very rapid decrease in the salt concentration and/or dialysis to <100 mM KCl can result in increased protein precipitation.

13. Remove the extract from the dialysis bag and check the conductivity a final time to ensure that dialysis is complete. Centrifuge the extract 20 min in JA-20 rotor at 14,500 rpm. Discard the pellet.

This will remove protein and nucleic acid that precipitates when the potassium chloride concentration is lowered during dialysis.

14. Determine the protein concentration of the supernatant. Aliquot into tubes if desired and rapidly freeze by submerging in liquid nitrogen. Store the extracts at -80°C .

Protein concentration can be determined using the Bradford assay (UNIT 10.1). If all of the extract is not going to be used at one time (i.e., for protein purification), divide the extract into aliquots before freezing to avoid unnecessary freezing and thawing. In general, avoid more than five cycles of freezing/thawing; however, the stability of specific proteins may differ. Thaw frozen extracts on ice.

OPTIMIZATION OF NUCLEAR EXTRACTION

This protocol describes a simple method to optimize the salt concentration during nuclear extraction for specific applications (see critical parameters). Extracts made with higher or lower potassium chloride concentrations can then be assayed in transcription, splicing, gel-shift analysis (*UNIT 12.2*) to determine the optimum extraction conditions for specific cells and applications.

Additional Materials

High-salt buffer with 0.8, 1.0, 1.2, 1.4, and 1.6 M KCl

1. Perform steps 1 to 8 of the basic protocol.

As in the basic protocol, perform all steps at 0° to 4°C and use precooled buffers and equipment.

2. After resuspending the nuclear pellet in low-salt buffer ($\frac{1}{2}$ pnv; step 8 of basic protocol), divide the suspension into five aliquots of equal volume.

Before aliquoting, be certain that the nuclei are completely resuspended. Homogenize with one or two strokes in a Dounce homogenizer using type B pestle if necessary.

3. Add $\frac{1}{3}$ aliquot vol of high-salt buffer with frequent or continuous mixing to each aliquot, as follows: add high-salt buffer with the lowest KCl concentration (0.8 M) to the first aliquot and add high-salt buffer with increasing KCl concentrations (up to 1.6 M) to subsequent aliquots.

It is critical to add this buffer in the cold room with gentle continual mixing on a magnetic stirrer.

4. Secure the tubes containing the extracting nuclei on a tiltboard and gently mix 30 min.
5. Remove the nuclei by centrifuging 30 min in JA-20 rotor at 14,500 rpm.
6. Decant the extracts (supernatants) and dialyze against dialysis buffer (as in step 12 of the basic protocol).

Dialyze all the extracts until the one with the highest KCl concentration approaches the conductivity of the dialysis buffer (100 mM KCl).

Centrifuge the extracts 20 min in JA-20 rotor at 14,500 rpm.

7. Check the conductivity of each extract, and determine the protein concentration of the supernatants (see steps 13 and 14 of basic protocol).
8. Assay the extracts directly or aliquot into tubes. Freeze in liquid nitrogen before storing at -80°C .

PREPARATION OF THE CYTOPLASMIC (S-100) FRACTION

The basic protocol also yields a cytoplasmic extract that may contain proteins of interest (e.g., RNA polymerase III transcription factors). This protocol describes the preparation of the S-100 fraction from this cytoplasmic extract. It is practical to prepare and dialyze the S-100 fraction and the nuclear extract simultaneously.

Additional Materials

Cytoplasmic extract (step 7, basic protocol)

10× cytoplasmic extract buffer

Beckman Type 50 fixed-angle rotor (or equivalent)

1. Carefully measure the volume of the cytoplasmic extract (supernatant from step 7 of the basic protocol). Add 0.11 vol of 10× cytoplasmic extract buffer and mix thoroughly.

As in the basic protocol, perform all steps at 0° to 4°C and use precooled buffers and equipment.

2. Centrifuge 1 hr in a Beckman Type 50 rotor at 40,000 rpm (100,000 × g).
3. Decant the supernatant containing the cytoplasmic (S-100) fraction and place it in dialysis tubing. Dialyze S-100 fraction against dialysis buffer until the conductivity of the former reaches that of dialysis buffer (100 mM KCl).

As with the nuclear extract, dialyze for the minimum amount of time necessary (see step 12, basic protocol).

4. Remove the S-100 fraction from the dialysis tubing and centrifuge 20 min in a Beckman JA-20 rotor at 14,500 rpm (25,000 × g) to pellet precipitated material.
5. Decant the supernatant, check the conductivity, and determine the protein concentration (steps 12 to 15 of the basic protocol). Aliquot into tubes, freeze in liquid nitrogen, and store at −80°C.

REAGENTS AND SOLUTIONS

NOTE: All solutions can be stored at 4°C for a couple of weeks. However, Dithiothreitol (DTT) and phenylmethylsulfonyl fluoride (PMSF) must be added to buffers immediately before use. PMSF (molecular weight 174.2) is dissolved in anhydrous isopropanol (Sigma #405-7) as a 0.2 M stock. The stock solution is stable for 9 months. PMSF is not stable in aqueous solution, and should be added to the solutions indicated below drop by drop with vigorous stirring to go into solution. Additional PMSF may be added again partway through a long dialysis.

A less expensive alternative to HEPES is Tris buffer, which can be used in place of HEPES in the solutions below. However, the pH of Tris buffers changes dramatically with temperature. For example, a Tris stock solution that is pH 7.9 at 4°C will drop to approximately pH 7.3 at room temperature. When using extracts prepared in Tris for functional studies (e.g., transcription or splicing) it is necessary to increase the pH of the extract at 30°C by adding HEPES, pH 8.4 (room temperature), to a final concentration of 40 mM.

10× cytoplasmic extract buffer

0.3 M HEPES, pH 7.9 at 4°C
1.4 M KCl
0.03 M MgCl₂

Dialysis buffer

20 mM HEPES, pH 7.9 at 4°C
20% glycerol
100 mM KCl
0.2 mM EDTA
0.2 mM PMSF*
0.5 mM DTT*

**Add ingredients immediately before use (see introductory note to reagents and solutions).*

High-salt buffer

20 mM HEPES, pH 7.9 at 4°C
25% glycerol
1.5 mM MgCl₂
0.8, 1.0, 1.2, 1.4 or 1.6 M KCl
0.2 mM EDTA
0.2 mM PMSF*
0.5 mM DTT*

*Add ingredients immediately before use (see introductory note to reagents and solutions).

Hypotonic buffer

10 mM HEPES, pH 7.9 at 4°C
1.5 mM MgCl₂
10 mM KCl
0.2 mM PMSF*
0.5 mM DTT*

*Add ingredients immediately before use (see introductory note to reagents and solutions).

Low-salt buffer

Prepare high-salt buffer, substituting 0.02 M KCl for 1.2 M KCl.

COMMENTARY

Background Information

Crude extracts from cultured mammalian cells were first used to demonstrate accurate initiation and termination of transcription from exogenously added purified DNA by RNA polymerase III (Weil et al., 1979a). Subsequently, cell-free soluble systems were developed from cultured cells that could accurately initiate transcription from purified genes by RNA polymerase II. These include a high-speed supernatant from a low-salt extraction of cells that was supplemented with purified RNA polymerase II (Weil et al., 1979b) and a high-salt extract of whole cells that contained endogenous polymerase II (Manley et al., 1980).

In order to take advantage of the presumed nuclear localization of transcriptional components, Dignam et al. (1983a) developed a procedure for preparing a soluble extract from isolated mammalian nuclei. This preparation was shown to support accurate transcription initiation from several class II genes by endogenous RNA polymerase II and accessory factors. Although 5S gene transcription by RNA polymerase III required supplementation with the cytoplasmic (S-100) fraction or the gene-specific factor TF IIIA, these extracts were sufficient for accurate transcription from tRNA and adenovirus VA genes by RNA polymerase III (Dignam et al., 1983b). Subsequent analyses

demonstrated that these extracts were also competent for *in vitro* pre-mRNA splicing (Krainer et al., 1984).

This protocol represents modifications of the original procedure of Dignam et al. (1983a). It employs KCl instead of NaCl and has the advantage of a normalized salt concentration during the nuclei extraction. This is achieved by adjusting the buffer volume in the nuclear extraction step to the volume of the nuclear pellet. These modifications result in more reproducible extracts from separate preparations of nuclei.

Nuclear extracts prepared from HeLa cells using this procedure support accurate initiation of transcription by RNA polymerase II and III from purified cloned promoters. Furthermore, regulatory promoter elements that function *in vivo* often regulate transcription in nuclear extracts. The presence in these extracts of proteins that bind to such promoter elements can be determined by gel-shift, methylation protection, and footprinting assays. These extracts will also accurately splice pre-mRNA into mature mRNA products. In addition to supplying a rapid functional assay of promoter elements and sequences required for RNA processing, these extracts provide starting material for purification and mechanistic analysis of the proteins involved in transcription and splicing.

Critical Parameters

In the preparation of nuclear extracts it is essential that denaturation of proteins or proteolysis be avoided as this will cause a loss of protein activity. In order to minimize these possibilities the protocols described are performed at 0° to 4°C, preferably in a cold room. In addition, fresh phenylmethylsulfonyl fluoride (PMSF; 0.2 mM final) is added to the buffers just prior to use to further inhibit proteolysis. While these treatments appear adequate for many mammalian cell types, in some instances, additional protease inhibitors may be necessary. For example, transcriptionally active nuclear extracts from yeast cells are prepared in the presence of 1 mM PMSF, 2 μM pepstatin A, and 0.6 μM leupeptin (Lue and Kornberg, 1987).

The approach described here is designed to generate reproducible preparations of extracts; however, there will be some variation in the activity of extracts prepared from separate cultures or prepared at different times. These variations can arise from several factors, including the volume and density of cells, the packed cell and nuclear volumes, and the amount of time needed to perform individual steps and to complete the procedure. Keeping a record of these factors allows one to directly compare the preparation of separate extracts and to correlate these parameters with any variations in activity. A flow chart for this purpose is presented in Figure 12.1.1. We recommend that this page be copied, filled in during each extract preparation, and kept in laboratory notebooks.

The preparation of nuclear extracts has two common pitfalls—the relevant proteins are inefficiently extracted from the nuclei or an excessive amount of nonspecific inhibitory proteins are also extracted. The significance of these problems depends upon the application for which the extract is intended. For example, if one is trying to purify a particular DNA-binding protein that is much more efficiently extracted with higher salt concentrations, the unwanted proteins also extracted can be removed with additional chromatographic procedures. Of course, if the protein of interest is efficiently solubilized at low KCl concentrations, contamination with proteins extracted at higher salt will be avoided.

If the extracts are to be used for direct functional analysis (i.e., transcription or splicing) where multiple proteins are involved, extraction conditions that achieve the maximum activity should be determined. In the case of

transcription analysis, increasing the KCl concentration during extraction of HeLa cells increases the transcriptional potential of the extracts up to the point where the nuclei begin to lyse during extraction. Nuclear lysis results in the release of chromosomal proteins that nonspecifically bind to DNA and thus inhibit transcriptional analysis. The conditions described in the basic protocol are chosen to give high transcriptional and splicing activity while remaining safely below salt concentrations where nuclei become fragile and lyse. Thus, these conditions are chosen to give reproducible, high-activity extracts. It should be noted that even under these conditions a significant fraction of inhibitory nonspecific DNA binding proteins are contained in the extracts. These nonspecific proteins inhibit transcription at low template concentrations (Abmayr et al., 1988). Contaminating proteins can be titrated by using an excess of template or nonspecific DNA.

Anticipated Results

Using this protocol it is possible to consistently produce extracts of 8 to 12 mg protein/ml. The high protein concentration in the extracts results in a high yield of transcription factors and nonspecific DNA binding proteins.

Time Considerations

The entire procedure should be completed in one day. It will require ~3 to 5 hr to prepare the extracts prior to dialysis, depending on the volume of cells and the amount of time required to centrifuge spinner cultures or scrape monolayer cells. Approximately 4 to 6 hr are required for dialysis. It is important to stop the dialysis and complete the procedure on the same day, rather than dialyze overnight. Dialysis time can be reduced by using dialysis buffer without KCl. In this instance it is necessary to check the conductivity of the extract frequently during dialysis. The dialysis should be stopped promptly when the extract reaches the conductivity of dialysis buffer with 100 mM KCl. After the dialysis is complete, centrifugating, freezing, and storing the extract can be completed in ≤45 min.

Literature Cited

Abmayr, S.M., Workman, J.L., and Roeder, R.G. 1988. The pseudorabies immediate early protein stimulates *in vitro* transcription by facilitating TFIID: Promoter interactions. *Genes Dev.* 2:542-553.

Preparation of Nuclear and Cytoplasmic Extracts

Date _____ Cell type _____ Source _____

Cell count _____ Culture volume _____ Total cell number _____

Starting time: _____

Collect and centrifuge spinner cultures 20 min at 1850 xg in conical tubes.

Wash monolayer cultures with PBS and collect in conical tubes.

Centrifuge 10 min at 1850 xg .

Packed cell volume (pcv): _____

Wash spinner cultures with 5 pcv PBS and centrifuge 10 min at 1850 xg .

Resuspend in 5 pcv hypotonic buffer. _____ (*vol hypotonic buffer*)

Centrifuge 5 min at 1850 xg .

Resuspend in hypotonic buffer to 3 pcv. _____ (*vol hypotonic buffer*)

Swell on ice 10 min.

Homogenize swollen cells 10 strokes.

Pellet nuclei 15 min at 3300 xg ; remove cytoplasmic extract.

Nuclear Extract

Packed nuclei volume (pnv): _____

Resuspend nuclei in 1/2 pnv low-salt buffer. _____ (*vol low-salt buffer*)

Add 1/2 pnv high-salt buffer. _____ (*vol high-salt buffer*)

Homogenization? _____ (*no. strokes*)

Extract 30 min: start: _____ stop: _____

Centrifuge 30 min at 25,000 xg .

Conductivity of supernatant: _____ = _____ M

Dialysis: start: _____ finish: _____

Conductivity of nuclear extract: _____ = _____ M

Centrifuge 20 min at 25,000 xg . _____ (*vol nuclear extract*)

Aliquots: no. _____ vol. _____

Freeze in liquid nitrogen; store at -80°C .

Ending time: _____

Protein concentration: _____

Cytoplasmic (S-100) Extract

Cytoplasmic extract volume: _____

Add 0.11 vol 10x cytoplasmic extract buffer _____ (*vol buffer*)

Centrifuge 1 hour at 100,000 xg .

Conductivity of supernatant: _____ = _____ M

Dialysis: start: _____ finish: _____

Centrifuge 20 min at 25,000 xg . _____ (*vol S-100 fraction*)

Conductivity of S-100 fraction: _____ = _____ M

Aliquots: no. _____ vol. _____

Freeze in liquid nitrogen; store at -80°C .

Ending time: _____

Protein concentration: _____

Figure 12.1.1 Flow sheet for recording data when preparing nuclear and cytoplasmic extracts.

- Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. 1983a. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucl. Acids Res.* 11:1475-1489.
- Dignam, J.D., Martin, P.L., Shastry, B.S., and Roeder, R.G. 1983b. Eukaryotic gene transcription with purified components. *Methods Enzymol.* 101:582-598.
- Krainer, A.R., Maniatis, T., Ruskin, B., and Green, M.R. 1984. Normal and mutant human β -globin pre-mRNAs are faithfully and efficiently spliced in vitro. *Cell* 36:993-1005.
- Lue, N.F. and Kornberg, R.D. 1987. Accurate initiation at RNA polymerase II promoters in extracts from *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 84:8839-8843.
- Manley, J.L., Fire, A., Cano, A., Sharp, P.A., and Geyer, M.L. 1980. DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. *Proc. Natl. Acad. Sci. U.S.A.* 77:3855-3859.
- Weil, P.A., Segall, J., Harris, B., Ng, S.Y., and Roeder, R.G. 1979a. Faithful transcription of eukaryotic genes by RNA polymerase III in systems reconstituted with purified DNA templates. *J. Biol. Chem.* 254:6163-6173.
- Weil, P.A., Luse, D.S., Segall, J., and Roeder, R.G. 1979b. Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* 18:469-484.

Key Reference

Dignam et al., 1983a. See above.

This is the original paper describing the preparation of nuclear extracts from HeLa cells. It also optimizes conditions for transcription from multiple class II promoters in these extracts.

Contributed by Susan M. Abmayr and
Jerry L. Workman
Pennsylvania State University
University Park, Pennsylvania

Mobility Shift DNA-Binding Assay Using Gel Electrophoresis

The DNA-binding assay using nondenaturing polyacrylamide gel electrophoresis (PAGE) provides a simple, rapid, and extremely sensitive method for detecting sequence-specific DNA-binding proteins. Proteins that bind specifically to an end-labeled DNA fragment retard the mobility of the fragment during electrophoresis, resulting in discrete bands corresponding to the individual protein-DNA complexes. The assay can be used to test binding of purified proteins or of uncharacterized factors found in crude extracts (see Basic Protocol). This assay also permits quantitative determination of the affinity, abundance, association rate constants, dissociation rate constants, and binding specificity of DNA-binding proteins. Three additional protocols describe a competition assay using unlabeled competitor DNA (see Alternate Protocol 1), an antibody supershift assay (see Alternate Protocol 2), and multicomponent gel shift assays (see Alternate Protocol 3).

STRATEGIC PLANNING

The utility of this technique is underscored by the many proteins that have been characterized using this assay. It has become clear that there is no single protocol that works best for all proteins. Rather, several variables can be changed to optimize binding. There are several options available for the design of the DNA probe, the binding reaction conditions, and the gel running conditions.

Protein-DNA binding is affected by a number of factors, including ion concentration and pH. Many of the salt, buffer, and glycerol components needed in the binding reaction may be contributed by the buffer contained in the protein extract. If only a small volume of protein extract is used, these components can be supplied by including the appropriate amount of buffer. Typical binding reaction conditions are 12% glycerol, 20 mM HEPES (pH 7.9) or 20 mM Tris·Cl (pH 7.9), 60 to 150 mM KCl, 1 mM EDTA, and 1 mM DTT. However, the DNA binding assay is compatible with a wide variety of binding conditions. Each parameter must be titrated for the specific protein-DNA interaction to be studied (see Critical Parameters).

Inclusion of BSA or other carrier protein in the binding reaction seems to improve the stability of some complexes during electrophoresis. Other reagents that have been reported to affect formation or stability of some protein-DNA complexes are nonionic detergents, polycations such as spermine and/or spermidine, and ATP. However, every protein-DNA complex is likely to have an idiosyncratic response to different binding conditions. Also, the optimal binding conditions for a purified protein may be different from the optimal binding conditions for the same protein in a crude extract.

As with the binding reaction, there are several variables for the nondenaturing gel—acrylamide concentration, buffer, and other components. Typical gels for this assay are in the range of 4% to 5% acrylamide. Early protocols used a high ratio of acrylamide to bisacrylamide (80:1), resulting in very soft gels. However, it is often preferable to run gels with a 40:1 acrylamide/bisacrylamide ratio, which gives a more cohesive gel. There are even published mobility shift experiments that resolve very large complexes on agarose gels (Lieberman and Berk, 1994).

Many gel buffers have been successfully used for mobility shift assays. The stability of protein-DNA complexes can be strongly affected by the choice of buffer, so it is worth trying different systems for the protein of interest. The most common buffers used are

**BASIC
PROTOCOL**

Tris-acetate (TAE), Tris-glycine, and Tris-borate (TBE) electrophoresis buffers at 0.25×, 0.5×, 1.0×, or 2.0× concentration. The gel running buffer should match the buffer in the gel. If low-ionic-strength buffer ($\leq 0.25\times$) is used, the buffer should be recirculated while running the gel. Although not required for stable electrophoresis of all protein-DNA complexes, glycerol in the gel at 2.5% (v/v) final concentration may help stabilize some interactions. At least one protein (the eukaryotic TATA-binding protein) also requires 3 to 5 mM Mg^{2+} in the gel and running buffer to allow it to bind stably to DNA during migration through the gel.

MOBILITY SHIFT ASSAY

This protocol can be divided into four stages: (1) preparation of a radioactively labeled DNA probe containing a particular protein binding site; (2) preparation of a nondenaturing gel; (3) a binding reaction in which a protein mixture is bound to the DNA probe; and (4) electrophoresis of protein-DNA complexes through the gel, which is then dried and autoradiographed.

Materials

10× electrophoresis buffer, e.g., TAE or TBE electrophoresis buffer (*APPENDIX 2*) or Tris-glycine electrophoresis buffer (see recipe)

30% (w/v) ammonium persulfate, prepared fresh
N,N,N',N'-tetramethylethylenediamine (TEMED)

Nondenaturing gel mix (see recipe)

Bulk carrier DNA, e.g., poly(dI-dC)·poly(dI-dC)

BSA

Protein preparation containing DNA-binding protein (crude extract or purified fraction)

10× loading buffer with dyes (*UNIT 2.5*)

Constant-temperature water bath

Two-head peristaltic pump

10- μ l glass capillary pipet (optional)

Clay-Adams screw-top loader (optional)

Whatman 3MM filter paper (or equivalent)

Additional reagents and equipment for digesting DNA with restriction endonucleases (*UNIT 3.1*), labeling DNA fragments with Klenow fragment (*UNIT 3.5*) or polynucleotide kinase (*UNIT 3.10*), agarose and nondenaturing polyacrylamide gel electrophoresis (*UNIT 2.5A* or *UNIT 2.7*), recovery of DNA from gels (*UNIT 2.6* & *UNIT 12.4*), oligonucleotide synthesis (*UNIT 2.11*), PCR (*UNIT 15.1*), ethanol precipitation (*UNIT 2.1*), ethidium bromide dot quantitation (*UNIT 2.6*), and autoradiography (*APPENDIX 3A*)

Prepare the DNA probe

DNA fragments from 20 to 300 bp long may be used as probes. However, longer fragments are likely to contain binding sites for multiple proteins, which may make interpretation of the gel difficult. The DNA probe can be prepared in one of several ways.

- 1a. *For restriction endonuclease fragments:* Isolate a small DNA fragment containing the binding site of interest from a plasmid using a standard restriction endonuclease digestion (*UNIT 3.1*). Label the fragment by filling in a 5' overhang with the Klenow fragment of *Escherichia coli* DNA polymerase and ^{32}P -labeled nucleotide (*UNIT 3.5*) or by end labeling using polynucleotide kinase (*UNIT 3.10*). Separate the fragment from the plasmid by gel electrophoresis (*UNIT 2.5A* or *UNIT 2.7*).

Kinased probes should be avoided in experiments using crude protein preparations that might contain phosphatase activity.

Agarose gels are good for resolving fragments as small as 50 bp. Fragments can be recovered using low melting temperature agarose or DEAE paper (UNIT 2.7 & UNIT 12.4). For smaller DNA fragments, nondenaturing polyacrylamide gels can be used (UNIT 2.7).

- 1b. *For synthetic oligonucleotides:* Synthesize and anneal (UNIT 2.11) complementary oligonucleotides to generate a double-stranded DNA fragment containing the binding site of interest. Label the probe using polynucleotide kinase (UNIT 3.10).

The kinase reaction can be performed before annealing to label only one strand, or after annealing to label both strands. Alternatively, the oligonucleotides can be designed to leave an overhang that can be filled in with labeled nucleotide and Klenow fragment of E. coli DNA polymerase I (UNIT 3.5).

- 1c. *For PCR fragments:* Generate a DNA fragment containing the binding site of interest by polymerase chain reaction (PCR; UNIT 15.1). End-label one of the primers with polynucleotide kinase (UNIT 3.10) before the PCR reaction, or label the double-stranded PCR product after purification.

2. Following isolation of the probe, determine its concentration by ethidium bromide dot quantitation (UNIT 2.6).

Concentrations in the range of 2 to 50 ng/μl are convenient.

3. Count 1 μl for Cerenkov counts in a scintillation counter to determine specific activity (cpm/μl).

A typical binding reaction will contain about 5,000 to 20,000 cpm and about 10 to 100 fmol probe (10 fmol DNA in a final reaction volume of 10 μl gives a total DNA concentration of 1 nM). If desired, probes can be stored up to 4 to 6 weeks at 4°C before proceeding with the experiment.

Prepare the nondenaturing gel

4. Dilute 10× electrophoresis buffer to prepare enough 1× electrophoresis buffer to fill the tank.

5. Assemble washed glass plates and 1.5-mm spacers for casting the gel (UNIT 2.7).

All traces of detergent must be removed because detergent will disrupt protein-DNA interactions.

6. Add 150 μl of 30% ammonium persulfate and 70 μl TEMED to 60 ml nondenaturing gel mix prepared using the same buffer as that used for electrophoresis. Swirl gently to mix.

The amounts of ammonium persulfate, TEMED, and nondenaturing gel mix can be scaled up or down as necessary depending on the size and number of gels.

7. Pour the gel mix between the plates and insert a comb. Allow the gel to completely polymerize for 20 min.

For optimal results, use a comb with teeth that are ≥7 mm wide.

8. Remove the comb and bottom spacer and attach the plates to the electrophoresis tank after filling the lower reservoir with 1× electrophoresis buffer. Fill the upper reservoir of the tank with 1× electrophoresis buffer. With a bent-needle syringe, remove any air bubbles trapped beneath the gel and flush out the wells (UNIT 2.7). Prerun the gel 30 to 60 min at 100 V.

For low-ionic-strength buffers (≤0.5×), use a pump with two heads and a flow rate of 5 to 30 ml per min to exchange buffer between the upper and lower reservoirs. Recirculation of the buffer is essential to prevent polarization due to the low buffering capacity of the buffer.

Prepare the binding reactions

9. While the gel is prerunning, assemble the binding reaction by combining the following in a 0.5-ml or 1.5-ml microcentrifuge tube:
 - 5,000 to 20,000 cpm radiolabeled probe DNA (0.1 to 0.5 ng, ≥ 10 fmol)
 - 0.1 to 2 μ g nonspecific carrier DNA
 - 300 μ g/ml BSA
 - $\geq 10\%$ (v/v) glycerol
 - Appropriate buffer and salt
 - DNA-binding protein (~15 μ g crude extract or ~5 to 25 ng purified protein)
 - Adjust the final reaction volume to 10 to 15 μ l with water or buffer.

The protein should be added last.

10. Mix gently by tapping the bottom of the tube with a finger.
 - Avoid introducing bubbles in the mix.*
11. Incubate the binding reaction mix 15 to 30 minutes in a constant-temperature water bath.
 - Optimal incubation temperatures for different proteins can vary from room temperature to 37°C.*

Run the gel

12. Load each binding reaction into the appropriate well of the prerun gel using either a 10- μ l glass capillary pipet and Clay-Adams screw-top loader, or a pipettor. Load a small volume of 10 \times loading buffer with dyes into a separate well.

The dyes are used to monitor the progress of electrophoresis.

There is no stacking gel in this system, so precise loading with little mixing with the gel buffer is necessary to obtain sharp bands on the gel. Allow the sample to fall along one side of the well to prevent dilution and avoid bubbles in the well.

13. Electrophorese at ~30 to 35 mA for the minimum time required to give good separation of free probe and the protein-DNA complexes. Stop the gel before the bromphenol blue approaches the bottom of the gel (~1.5 to 2 hr for a 15- to 20-cm gel).

Longer run times may cause a weaker signal due to partial dissociation of complexes during electrophoresis.

Bromphenol blue migrates at approximately the same position as a 70-bp DNA probe. For probes <70 bp, do not run the bromphenol blue to the bottom of the gel.

If electrophoresis is performed at room temperature, the glass plates should be allowed to become only slightly warm. Decrease the voltage if the plates become any hotter. To run the gel faster, put the apparatus in a cold room. Higher voltages may then be used without heating the glass plates.

Analyze the gel

14. Remove the glass plates from the gel box and carefully remove the side spacers.
15. Using a spatula, slowly pry the glass plates apart, allowing air to enter between the gel and the glass plate.

The gel should remain attached to only one of the plates. Prying the plates apart too quickly may tear the gel or cause it to stick to both plates. If this occurs or if the gel has become distorted, squirt a stream of water underneath it. This will reduce the stickiness of the gel. Be careful not to let the gel slide off the plate.

16. Lay the glass plate with the gel attached on the bench with the gel facing up. Place three sheets of Whatman 3MM filter paper cut to size on top of the gel.

17. Support both sides of the gel sandwich and carefully flip it over so that the filter paper is on the bottom and the glass plate is on the top.
18. Carefully lift up one end of the glass plate. Peel the filter paper with the gel attached to it from the plate.
19. Cover the gel with plastic wrap and dry under vacuum with heat.
20. Autoradiograph the dried filter overnight without an intensifying screen or 2 to 3 hr with an intensifying screen (*APPENDIX 3A*).

COMPETITION MOBILITY SHIFT ASSAY

One of the most powerful aspects of the mobility shift DNA-binding assay is the ease of assessing the sequence specificity of protein-DNA interactions using a competition binding assay (Carthew et al., 1985; Singh et al., 1986). This is necessary because most protein preparations will contain both specific and nonspecific DNA binding proteins. For a specific competitor, the same DNA fragment (unlabeled) as the probe can be used. The nonspecific competitor can be essentially any fragment with an unrelated sequence, but it is useful to roughly match the probe and specific competitor for size and configuration of the ends. For example, some proteins bind blunt DNA ends nonspecifically. These would not be competed by circular plasmid or a fragment with overhangs, leading to the false conclusion that the protein-DNA complex represented specific binding. Perhaps the best control competitor is a DNA fragment that is identical to the probe fragment except for a mutation(s) in the binding site that is known to disrupt function (and presumably binding).

Additional Materials (also see Basic Protocol)

Unlabeled specific and nonspecific competitor DNA fragments

For this protocol, follow steps 1 to 8 of the Basic Protocol to prepare the probe and gel, modify the binding reaction as indicated in step 9, and proceed with steps 10 through 20 of the Basic Protocol.

- 9a. Assemble binding reactions (see Basic Protocol, step 9). In addition to the labeled DNA probe, add increasing amounts of unlabeled specific and nonspecific competitor DNAs. Add DNA-binding protein last.

Assuming there is not a huge excess of binding protein to probe, typical amounts of competitor are 5×, 10×, and 50× molar excess relative to the labeled probe.

Sequence-specific complexes will show a clear differential response to the specific and nonspecific competitors.

ANTIBODY SUPERSHIFT ASSAY

Another useful variation of the mobility shift DNA-binding assay is to use antibodies to identify proteins present in the protein-DNA complex (Kristie and Roizman, 1986).

Addition of a specific antibody to a binding reaction can have one of several effects. If the protein recognized by the antibody is not involved in complex formation, addition of the antibody should have no effect. If the protein that forms the complex is recognized by the antibody, the antibody can either block complex formation, or it can form an antibody-protein-DNA ternary complex and thereby specifically result in a further reduction in the mobility of the protein-DNA complex (supershift). Results may be different depending upon whether the antibody is added before or after the protein binds DNA (particularly if there are epitopes on the DNA-binding surface of the protein).

**ALTERNATE
PROTOCOL 1**

**ALTERNATE
PROTOCOL 2**

**DNA-Protein
Interactions**

12.2.5

Additional Materials (also see Basic Protocol)

Antibody specific for DNA-binding protein
Nonspecific control antibody

For this protocol, follow steps 1 to 8 of the Basic Protocol to prepare the probe and gel, modify the binding reaction as indicated in step 9, and proceed with steps 10 through 20 of the Basic Protocol.

9a. Assemble binding reactions (see Basic Protocol, step 9). Add a small volume of antibody ($\leq 1 \mu\text{l}$) to the binding mix. To a second tube of binding reaction, add an equivalent amount of nonspecific control antibody.

The amount of antibody used should be the minimum needed to produce an observable effect, and inclusion of a control antibody reaction is critical, because the salts and other proteins in the antibody preparation may nonspecifically affect stability or mobility of the protein-DNA complexes. Antibody supershifts have been successfully performed using crude sera, purified polyclonal antibodies, and monoclonal antibodies. In general, however, the more pure the antibody preparation, the clearer the results will be.

**ALTERNATE
PROTOCOL 3**

MULTICOMPONENT MOBILITY SHIFT ASSAYS

The mobility shift assay can be used to study multicomponent assemblies of DNA-binding complexes. Indeed, many of the transcription factors studied by mobility shift assays are actually homo- or heterodimers. There are many cases in which a sequence-specific DNA-binding protein (A) acts as a platform for the association of other proteins (B and C) which themselves do not bind specific sequences (see Fig.12.2.1). This phenomenon can be observed as a supershift of the primary DNA-protein complex into a new discrete complex that is dependent upon all of the factors.

In essence, assays for multicomponent complexes are no different from the Basic Protocol. However, some additional points should be considered:

1. The concentration of factors may need to be high. Although the dissociation constant (K_d) for a sequence-specific DNA-binding protein is often very high (10^{-9} to 10^{-12} M),

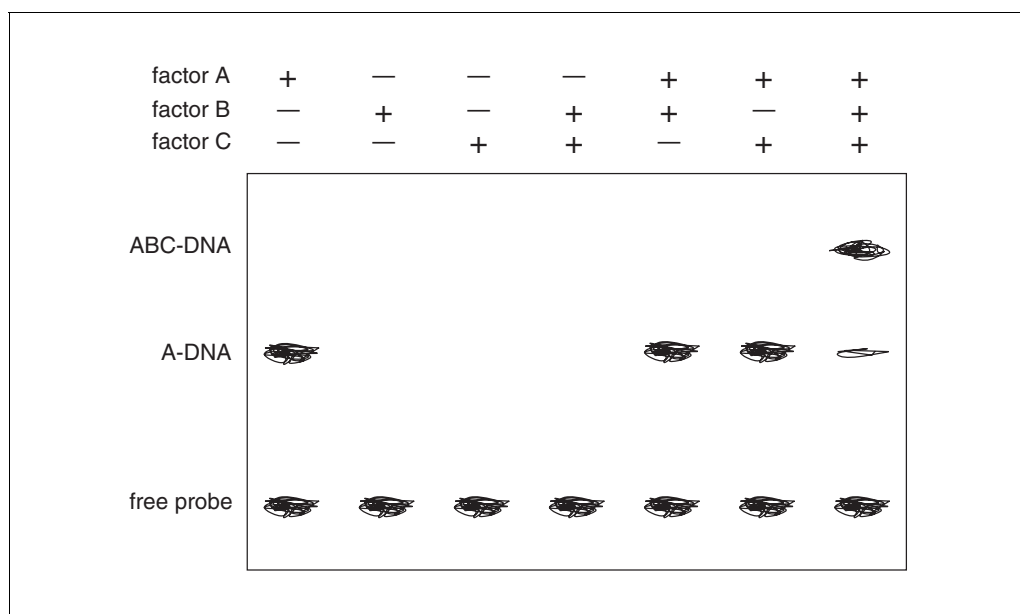


Figure 12.2.1 Hypothetical autoradiogram of a mobility shift binding experiment analyzing a multicomponent complex (see Alternate Protocol 3). In this experiment, factor A provides sequence-specific DNA binding. Factors B and C associate with the factor A-DNA complex.

the constants for subsequent protein-protein interactions may be lower. Therefore, it may be necessary to saturate the probe for binding of factor A to reach a concentration where the ABC complex can form.

2. A mobility shift is not necessarily due to association of a factor. For example, factor C may be a kinase whose function is necessary for the association of factor B with the DNA-A complex. Antibody supershift experiments or other protein-protein interaction assays may be useful for showing physical association of a factor with a particular complex.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Nondenaturing gel mix

For 60 ml:

6.0 ml 10× TAE or TBE electrophoresis buffer (APPENDIX 2) or Tris-glycine electrophoresis buffer (see recipe)

8.1 ml 30% (w/v) acrylamide

3.0 ml 2% (w/v) bisacrylamide

42.9 ml H₂O

Prepare fresh for each use

CAUTION: Acrylamide monomer is neurotoxic. Gloves should be worn when handling the solution, and the solution should not be pipetted by mouth.

Tris-glycine electrophoresis buffer, 10×

30.28 g Tris base (0.25 M final)

142.7 g glycine (1.9 M final)

3.92 g EDTA (10 mM final)

H₂O to 1 liter

Check that pH is ~8.3

Store up to several months at room temperature

COMMENTARY

Background Information

Several methods exist for detecting the sequence-specific binding of proteins to DNA, including nitrocellulose filter binding (UNIT 12.8), DNase I footprinting (UNIT 12.4), methylation protection, methylation interference (UNIT 12.3), and mobility shift gel electrophoresis, the approach presented here. The DNA-binding assay using mobility shift polyacrylamide gel electrophoresis (PAGE) is based on the observation that DNA-protein complexes migrate through polyacrylamide gels more slowly than unbound DNA fragments. This assay is generally simpler, faster, and more sensitive than other methods. Thus, it is an ideal assay for monitoring the purification of DNA-binding proteins. More importantly, the sensitivity of this assay enables femtomole quantities of DNA-binding proteins to be detected routinely. It is quite remarkable what large complexes can be resolved with acrylamide gels. For example,

the eukaryotic RNA polymerase II or RNA polymerase III transcription complexes (each of total molecular weight well over 500 kDa) can be resolved on 4% acrylamide gels.

In addition to providing quantitative information on the amount of DNA bound by the protein, the use of mobility shift PAGE provides additional information on the number and type of proteins bound. Each species of protein bound to the probe generates a complex of distinct mobility and specificity so that interactions of several proteins binding to a single DNA fragment can be observed. Even if multiple proteins recognize overlapping sites on the DNA fragment, the complexes formed by each can be resolved and characterized. If there are multiple forms of a protein that bind (for example, several phosphorylation states), these can often be distinguished using this assay.

This assay was originally developed for kinetic and equilibrium analyses of purified pro-

karyotic gene-regulatory proteins (Fried and Crothers, 1981, 1984a,b; Garner and Revzin, 1981; Hendrickson and Schleif, 1984). It was modified by the addition of bulk carrier DNA (*Escherichia coli* DNA) to the binding reaction, permitting detection of DNA-binding proteins in crude extracts (Strauss and Varshavsky, 1984). Subsequently, a synthetic alternating copolymer was used instead of heterologous DNA to facilitate the identification of sequence-specific DNA-binding proteins (Carthew et al., 1985; Singh et al., 1986). Apparently, the large mass of alternating copolymer presented a large number of nonspecific binding sites for DNA-binding proteins in the extract, but presented considerably fewer sequence-specific binding sites relative to heterologous-sequence DNA.

Interestingly, protein-DNA complexes with short half-lives (<1 min) can be easily detected despite the fact that electrophoresis takes significantly longer than this. Thus, kinetic stability is not a requirement for detection of protein-DNA complexes. The mobility shift DNA-binding assay is often more sensitive than the nitrocellulose filter-binding assay because the kinetically labile complexes dissociate during washing of the filter. For example, *lac* repressor-operator complexes formed in the presence of IPTG can be detected by the mobility shift assay but not by filter binding (Hendrickson and Schleif, 1984).

Although some dissociation of complexes during nondenaturing gel electrophoresis undoubtedly occurs, complexes within the gel are remarkably stable. At least two factors stabilize protein-DNA complexes in the gel. One of these is the low ionic strength of the buffer, which can increase binding affinity. The second factor is the "caging effect" (Fried and Crothers, 1981; Garner and Revzin, 1981). The gel matrix may provide a cage so that dissociated components cannot diffuse away and instead rapidly recombine with each other. Therefore, the separation achieved during electrophoresis is dependent upon the fraction of free and bound probe molecules as the binding reaction enters the gel.

Because the mobility shift procedure can be performed under conditions in which the DNA probe concentration is in large excess over that of the DNA-binding protein, it can be used to obtain thermodynamically rigorous physical constants of DNA-binding proteins, such as relative affinity, association rate constants, and dissociation rate constants (Riggs et al., 1970; Fried and Crothers, 1981; Chodosh et al.,

1986). The abundance of a protein in a crude extract can also be easily measured using this assay by increasing the concentration of DNA probe in the binding reaction until the amount of complex formed no longer increases (Chodosh et al., 1986). At this DNA concentration, all of the particular DNA-binding protein molecules present in the sample will be bound. By comparing the amount of probe in the protein-DNA complex to a set of standard dilutions of free probe included on the same gel, the number of moles of protein in the reaction can be calculated, assuming a 1:1 stoichiometry of binding, because the specific activity of the probe is known.

Critical Parameters and Troubleshooting

Binding reaction and gel electrophoresis conditions may affect the results obtained in this assay. Binding reaction conditions that may be varied include monovalent and divalent salt concentrations, pH, presence or absence of nonionic detergents, protein concentration, DNA concentration, and type and concentration of bulk carrier DNA.

At a fixed protein concentration, if too little carrier DNA is included in the reaction, all of the probe will be bound and will not enter the gel. Conversely, if too much carrier DNA is included, none of the probe will be bound. An appropriate amount of carrier DNA is one that is sufficient to prevent nonspecific proteins from binding to the probe, yet not so great as to prevent the specific binding of proteins that recognize sequences within the DNA probe. Most sequence-specific DNA-binding proteins have nonspecific DNA affinities several orders of magnitude below that for the recognized site. In practice, when synthetic DNA polymers are used as carrier DNA, the range of carrier DNA concentrations that will permit the detection of specific DNA-binding proteins is quite broad.

The type of carrier DNA chosen is of great importance. Optimally, the sequence should bear little resemblance to the specific binding site. Thus, if the putative binding site is TATAT-GTA, the synthetic alternating copolymer poly(dA-dT)-poly(dA-dT) would be a poor choice for a carrier DNA, whereas the synthetic polymer poly(dG-dC)-poly(dG-dC) would be a much better choice. If the binding site sequence of a putative DNA-binding protein is unknown, at least two synthetic DNA carriers of different structure should be tested. In general, heterologous-sequence DNA, such as *E. coli* or salmon sperm DNA, will reduce the

amount of specific protein-DNA complex because binding sites of varying affinities will fortuitously occur at some frequency.

The conditions under which gel electrophoresis is performed can greatly influence the results obtained. The mobility of a protein-DNA complex through a nondenaturing polyacrylamide gel is determined primarily by the size and charge of the protein bound to the DNA fragment and by the conformation of the protein-DNA complex. The mobility of the protein-DNA complex is only slightly affected by changes in the size of the DNA fragment used as probe. As a result, changing gel conditions such as acrylamide percentage, acrylamide/bisacrylamide ratio, and pH can significantly alter the mobility of a given protein-DNA complex. In some cases this alteration may be sufficient to permit the resolution of two protein-DNA complexes, allowing the DNA-binding specificity of each to be independently and unambiguously ascertained. Use of high-ionic-strength buffer systems can also enhance the resolution of protein-DNA complexes. As a result of these features, novel protein-DNA complexes may be obtained by varying the pH of the buffer system (Staudt et al., 1986). Interestingly, the mobility of a protein-DNA complex can be affected by the position of the binding site relative to the ends of the DNA fragment (Zinkel and Crothers, 1987). This is presumed to result from conformational effects such as protein-induced bending of the DNA fragment.

As previously noted, many buffer systems may be used with the mobility shift DNA-binding assay. The Tris-glycine system, in particular, has several advantages. The high ionic strength makes recirculation of buffer unnecessary. Because of the higher ionic strength, a smaller amount of bulk carrier DNA (approximately 2- to 10-fold lower) is required to abolish nonspecific binding of proteins to the probe than with low-ionic-strength systems. Moreover, the resulting bands tend to be sharper than those of the low-ionic-strength systems, allowing better resolution of protein-DNA complexes. Furthermore, the higher pH of the Tris-glycine system alters the mobility of some protein-DNA complexes, permitting detection of novel protein-DNA complexes (Staudt et al., 1986). However, although many protein-DNA interactions may survive the high ionic strength of this buffer system, some do not. Therefore, it is helpful to try both low- and high-ionic-strength buffer systems for each putative protein-DNA interaction.

For further discussion of important binding reaction conditions, see *UNIT 12.8*, Critical Parameters. A guide to troubleshooting these procedures is provided in Table 12.2.1.

Anticipated Results

Figure 12.2.2 shows a typical autoradiogram from a mobility shift assay. Lane 1 contains only the DNA probe and poly(dI-dC)-poly(dI-dC) with no added protein. A single band representing the unbound probe is seen. In lanes 2 to 5, an increasing amount of protein from a crude extract is added to a constant amount of probe and poly(dI-dC)-poly(dI-dC). In lanes 2 and 3, two discrete complexes (B1 and B2) migrate through the gel more slowly than the band corresponding to the free probe. As more protein is added to the reaction, the intensities of the B1 and B2 bands increase and the amount of free probe decreases (compare lanes 2 and 3). As even more protein is added (lanes 4 and 5), there is little or no free probe remaining. Moreover, complexes B1 and B2 are no longer observed. Because many different nonspecific DNA-binding proteins are bound, most of the probe migrates very slowly and diffusely or does not enter the gel.

The reaction shown in lane 6 is identical to that shown in lane 5, except that a larger excess of poly(dI-dC)-poly(dI-dC) has been added. Because more protein was added in lane 6 relative to that in lane 3, the intensities of the bands representing complexes B1 and B2 are greater in lane 6 than in lane 3. Note that if a more purified preparation of B1 were used in the binding reaction, all of the probe in the reaction could be specifically bound and would migrate in complex B1, rather than at the top of the gel.

In lanes 7 to 9, a competition binding experiment (see Alternate Protocol 1) is used to determine whether complexes B1 and B2 represent specific protein-DNA complexes. Lane 7 is a standard binding reaction that contains probe, protein, and poly(dI-dC)-poly(dI-dC). Lane 8 is the same as lane 7, except that a 50-fold molar excess of an unlabeled DNA fragment identical to the probe (i.e., a specific competitor) is added prior to the addition of protein. Lane 9 is the same as lane 8, except that an unlabeled DNA competitor fragment containing sequences different from the probe (i.e., a nonspecific competitor) was added.

Formation of complex B1 is competed by the specific unlabeled DNA fragment (lane 8), but not by the nonspecific DNA fragment (lane 9). Therefore, complex B1 results from the

Table 12.2.1 Troubleshooting Guide for Mobility Shift Assays

Problem	Possible cause	Possible remedies
No specific complexes observed	Not enough protein or probe (amounts $<K_d$ or below level of detection)	Titrate up the amount of protein and/or probe.
	Protein binds nonspecific competitor DNA	Try a different DNA—e.g., poly(dG-dC) instead of poly(dI-dC).
	Gel buffer system incompatible with binding	Try a different gel buffer system.
Probe smeared or stuck in the well; no free probe observed	Binding conditions not appropriate	Try different salt and buffer conditions in binding reaction.
	Protein/probe ratio too high	Titrate down the amount of protein or increase nonspecific competitor DNA. Small amounts of detergent may also help.
Complexes seen but a smear is observed leading down to the free probe	Complex too big for gel system	Try lower percentage of acrylamide or lower acrylamide/bisacrylamide ratio.
	Complexes dissociating during electrophoresis	Run gel slower or at lower temperature; try different gel buffer system.
Complexes indistinct, bands fuzzy	Poor loading technique	Increase glycerol in binding reaction; load sample as sharp band; minimize time sample sits in well before current is turned on.
	Gel run too fast or too long	Reduce voltage or running time.
	Multiple protein species due to modifications	Multiple species may be due to phosphorylation or other modifications; test for alterations upon phosphatase treatment. Multiple species may be better resolved by running gel longer or under different conditions.
	Protein or probe degraded	Try protease inhibitors in protein preparation; check protein for nucleases.

specific binding of a molecule in the extract to the DNA probe. Formation of complex B2 is not competed by either the specific competitor or the nonspecific competitor. Complex B2 results from the nonspecific binding of a molecule in the extract to the DNA probe. This is typical behavior for a nonspecific DNA-protein interaction.

Time Considerations

Following preparation of the DNA probe, the DNA-binding assay itself is extremely rapid. It may take 4 to 5 hr to complete the experiment from the time the gel is poured to the time it is dried onto filter paper. Only a small fraction of this time is labor intensive. Thus, several DNA-binding gels can be run simultaneously while other experiments are being per-

formed. Furthermore, the results are often available within 2 to 3 hr after completion of the experiment.

Literature Cited

- Carthew, R.W., Chodosh, L.A., and Sharp, P.A. 1985. An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. *Cell* 43:439-448.
- Chodosh, L.A., Carthew, R.W., and Sharp, P.A. 1986. A single polypeptide possesses the binding and activities of the adenovirus major late transcription factor. *Mol. Cell. Biol.* 6:4723-4733.
- Fried, M. and Crothers, D.M. 1981. Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucl. Acids Res.* 9:6505-6525.
- Fried, M. and Crothers, D.M. 1984a. Kinetics and mechanism in the reaction of gene regulatory proteins with DNA. *J. Mol. Biol.* 172:241-262.

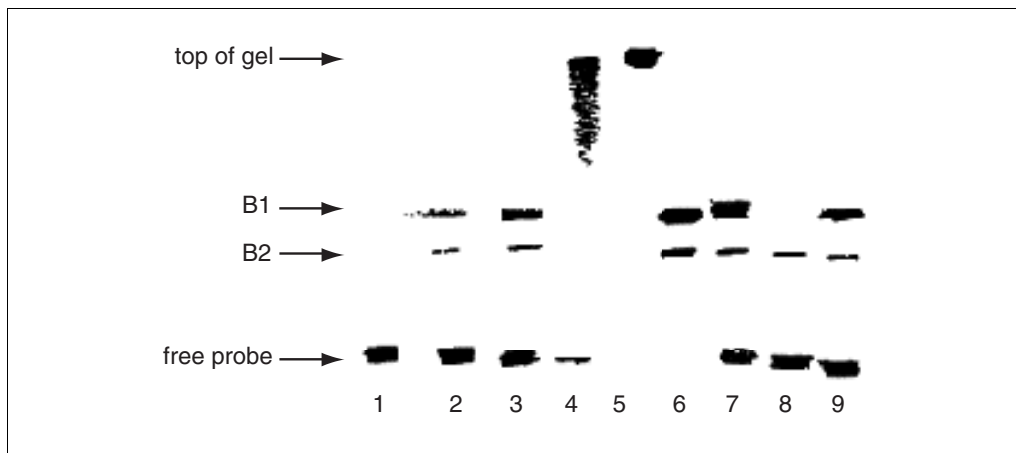


Figure 12.2.2 Hypothetical autoradiogram of a typical mobility shift DNA-binding experiment. Radioactive DNA probe and poly(dI-dC)-poly(dI-dC) were incubated with varying amounts of protein and electrophoresed from top to bottom on a low-ionic-strength polyacrylamide gel. The positions of the free probe and bound protein-DNA complexes are indicated. Lane 1, DNA probe + poly(dI-dC)-poly(dI-dC); lanes 2 to 5, DNA probe + poly(dI-dC)-poly(dI-dC) + increasing amounts of protein from crude extract; lane 6, the same as lane 5 except with a large excess of poly(dI-dC)-poly(dI-dC); lane 7, standard binding reaction; lane 8, standard binding reaction + 50-fold molar excess of unlabeled probe DNA (specific competitor); lane 9, standard binding reaction + 50-fold molar excess unlabeled DNA with a sequence unrelated to the probe (nonspecific competitor; also see Anticipated Results).

Fried, M. and Crothers, D.M. 1984b. Equilibrium studies of the cyclic AMP receptor protein-DNA interaction. *J. Mol. Biol.* 172:263-282.

Garner, M.M. and Revzin, A. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: Application to components of the *Escherichia coli* lactose operon regulatory system. *Nucl. Acids Res.* 9:3047-3060.

Hendrickson, W. and Schleif, R.F. 1984. Regulation of the *Escherichia coli* L-arabinose operon studied by gel electrophoresis DNA binding assay. *J. Mol. Biol.* 174:611-628.

Kristie, T.M. and Roizman, B. 1986. a4, the major regulatory protein of herpes simplex virus type 1, is stably and specifically associated with promoter-regulatory domains of a genes and/or selected viral genes. *Proc. Natl. Acad. Sci. U.S.A.* 83:3218-3222.

Lieberman, P.M. and Berk, A.J. 1994. A mechanism for TAFs in transcriptional activation: Activation domain enhancement of TFIID-TFIIA-promoter DNA complex formation. *Genes & Dev.* 8:995-1006.

Riggs, A.D., Suzuki, H., and Bourgeois, S. 1970. *Lac* repressor-operator interactions: I. Equilibrium studies. *J. Mol. Biol.* 48:67-83.

Singh, H., Sen, R., Baltimore, D., and Sharp, P.A. 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. *Nature* 319:154-158.

Staudt, L.M., Singh, H., Sen, R., Wirth, T., Sharp, P.A., and Baltimore, D. 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. *Nature* 323:640-643.

Strauss, F. and Varshavsky, A. 1984. A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. *Cell* 37:889-901.

Zinkel, S.S. and Crothers, D.M. 1987. DNA bend direction by phase-sensitive detection. *Nature* 328:178-181.

Key References

Carthew et al., 1985. See above.

Describes a variation of the mobility shift DNA-binding assay that is useful in detecting low-abundance molecules in crude extracts.

Chodosh et al. 1986. See above.

Presents a detailed description for measuring kinetic and thermodynamic properties of protein-DNA interactions using the mobility shift procedure.

Fried and Crothers, 1981. See above.

Garner and Revzin, 1981. See above.

Seminal articles on the mobility shift DNA-binding assay with excellent coverage of many key features of the procedure.

Contributed by Stephen Buratowski
Harvard Medical School
Boston, Massachusetts

Lewis A. Chodosh
University of Pennsylvania
Philadelphia, Pennsylvania

Methylation and Uracil Interference Assays for Analysis of Protein-DNA Interactions

UNIT 12.3

Interference assays identify specific residues in the DNA binding site that, when modified, interfere with binding of the protein. The protocols use end-labeled DNA probes that are modified at an average of one site per molecule of probe. These probes are incubated with the protein of interest, and protein-DNA complexes are separated from free probe by the mobility shift assay (UNIT 12.2). A DNA probe that is modified at a position that interferes with binding will not be retarded in this assay; thus, the specific protein-DNA complex is depleted for DNA that contains modifications on bases important for binding. After gel purification, the bound and unbound DNA are specifically cleaved at the modified residues and the resulting products analyzed by electrophoresis on polyacrylamide sequencing gels and autoradiography. In methylation interference (first basic protocol), probes are generated by methylating guanines (at the N-7 position) and adenines (at the N-3 position) with DMS; these methylated bases are cleaved specifically by piperidine. In uracil interference (second basic protocol), probes are generated by PCR amplification (UNIT 15.1) in the presence of a mixture of TTP and dUTP, thereby producing products in which thymine residues are replaced by deoxyuracil residues (which contains hydrogen in place of the thymine 5-methyl group). Uracil bases are specifically cleaved by uracil-*N*-glycosylase to generate apyrimidinic sites that are susceptible to piperidine. These procedures provide complementary information about the nucleotides involved in protein-DNA interactions.

METHYLATION INTERFERENCE ASSAY

Methylation interference identifies guanines and adenines in the DNA binding site that, when methylated, interfere with binding of the protein. The protocol described below uses a single end-labeled DNA probe that is methylated at an average of one site per molecule of probe. The labeled probe is a substrate for a protein-binding reaction. DNA-protein complexes are separated from the free probe by the mobility shift DNA-binding assay (UNIT 12.2). A DNA probe that is methylated at a position that interferes with binding will not be retarded in this assay. Therefore, the specific DNA-protein complex is depleted for DNA that contains methyl groups on purines important for binding. After gel purification, DNA is cleaved with piperidine. Finally, these fragments are electrophoresed on polyacrylamide sequencing gels and autoradiographed. Guanines and adenines that interfere with binding are revealed by their absence in the retarded complex relative to a lane containing piperidine-cleaved free probe. This procedure offers a rapid and highly analytical means of characterizing DNA-protein interactions.

Materials

- DNA containing protein-binding site
- TE buffer, pH 7.5 to 8.0 (APPENDIX 2)
- Dimethyl sulfate (DMS)
- DMS reaction buffer
- DMS stop buffer
- 10 mg/ml tRNA solution
- 0.3 M sodium acetate/1 mM EDTA, pH 5.2
- 1 M piperidine (dilute from 10 M piperidine stock)
- Stop/loading dye (UNIT 7.4)
- 90° to 95°C water bath

Additional reagents and materials for mobility shift DNA-binding assay (UNIT 12.2), phenol extraction and ethanol precipitation (UNIT 2.1), nondenaturing

**BASIC
PROTOCOL**

**DNA-Protein
Interactions**

12.3.1

Contributed by Albert S. Baldwin, Jr., Marjorie Oettinger, and Kevin Struhl

Current Protocols in Molecular Biology (1996) 12.3.1-12.3.7

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 36

polyacrylamide gel electrophoresis and electroelution of DNA fragments onto DEAE membranes (UNIT 2.7), autoradiography (APPENDIX 3), and denaturing polyacrylamide gel electrophoresis for sequencing (UNIT 7.6)

Prepare methylated DNA probe

1. Prepare a DNA probe labeled at one end (as for mobility-shift assay).
2. Suspend $\sim 10^6$ cpm of probe in 5 to 10 μl TE buffer. Add 200 μl DMS reaction buffer and 1 μl DMS. Mix well by vortexing. Incubate 5 min at room temperature.

CAUTION: DMS is a powerful poison and must be used in a fume hood with careful handling. Liquids containing DMS should be disposed of in a designated DMS waste bottle and pipet tips that come into contact with DMS should be placed in a separate DMS solid-waste bottle for disposal by institutional safety officials.

3. Add the following to the probe mixture:

40 μl DMS stop buffer
1 μl 10 mg/ml tRNA solution
600 μl 100% ethanol.

Mix and incubate ~ 10 min in a dry ice/ethanol bath. Microcentrifuge 10 min at top speed, 4°C. Carefully remove supernatant with a drawn-out Pasteur pipet and dispose in liquid DMS waste bottle.

4. Resuspend pellet in 250 μl of 0.3 M sodium acetate/1 mM EDTA. Keep the tube on ice. Add 750 μl of 100% ethanol, mix, and ethanol precipitate as in step 3.
5. Repeat ethanol precipitation once exactly as in step 4. Wash pellet once in 70% ethanol and microcentrifuge 10 min again. Carefully remove supernatant, invert tube on tissue, and air dry 10 min.
6. Measure the pellet for Cerenkov counts in a scintillation counter to determine cpm. Resuspend pellet in TE buffer at $\sim 20,000$ cpm/ μl .

The DNA can be difficult to resuspend, particularly with excessive drying of the pellet. Heating, vortexing, and pipetting up and down aid resuspension.

Bind methylated probe to protein and isolate DNA-protein complex and free probe

7. Set up a DNA-binding reaction as described in step 9 of the UNIT 12.2 basic protocol, but scaled up ~ 5 -fold (use 10^5 cpm of probe in 50 μl).

The degree of scale-up is dependent on the relative proportion of protein-DNA complex obtained. A 5-fold scale-up is sufficient for most DNA-binding activities.

8. Load binding reaction on three lanes of a native (nondenaturing) polyacrylamide gel. Electrophorese the binding reaction using the procedure for the mobility-shift assay.
9. Autoradiograph gel, cut out bands corresponding to the protein-DNA complex and free probe, and purify DNA from gel by electroelution onto a DEAE membrane.

Cleave DNA-protein complex with piperidine

10. Resuspend pellet in 100 μl of 1 M piperidine. Place in a 90° to 95°C water bath for 30 min. Put a glass plate over tubes to keep tops from popping open. Carefully remove from water bath and place on dry ice.
11. Make holes in the tops of the tubes with a large needle and lyophilize samples in a vacuum evaporator (e.g., Speedvac) for ~ 1 hr or until dry. Add 100 μl of distilled water. Freeze and lyophilize again. Repeat addition of water, freezing, and lyophilizing. Measure sample for Cerenkov counts to determine cpm.

Analyze fragments on DNA sequencing gels

12. Add sufficient stop/loading dye to pellet so that 1 to 2 μl will contain the sample to be loaded. Heat 5 min at 95°C. Quickly chill on ice.

For an overnight exposure with intensifying screen, 3000 cpm is sufficient. It is critical to equalize the number of counts applied from the DNA-protein complex and from the free probe to allow accurate comparison between samples.

13. Load the samples from the free probe and bound complex on a 6% or 8% polyacrylamide/urea sequencing gel. Electrophorese samples as for a sequencing gel and expose gel for autoradiography.

URACIL INTERFERENCE ASSAY

Uracil interference identifies thymines in a DNA binding site that, when modified, interfere with binding of the protein. Probes generated by PCR amplification in the presence of TTP and dUTP incorporate deoxyuracil in place of thymine residues. PCR products are incubated with the binding protein and resulting complexes are separated from unbound DNA. The DNA recovered from the protein-DNA complex is treated with uracil-*N*-glycosylase and piperidine, and the products are then electrophoresed on a denaturing polyacrylamide gel (Fig. 12.3.1).

Materials

DNA containing protein-binding site
Oligonucleotide primers specific for sequences flanking the binding site on the two complementary DNA strands
2 mM 4dNTP mix (UNIT 3.4)
0.5 mM dUTP
Taq polymerase buffer (UNIT 3.4) and *Taq* polymerase (UNIT 3.5)
TE buffer, pH 7.5 to 8.0 (APPENDIX 2)
Uracil-*N*-glycosylase (Perkin Elmer-Cetus)
1 M piperidine (diluted from 10 M piperidine stock)
Stop/loading dye (UNIT 7.4)
90° to 95°C water bath

Additional reagents and equipment for end-labeling DNA with T4 polynucleotide kinase (UNIT 3.9), polymerase chain reaction (UNIT 15.1), phenol extraction and ethanol precipitation (UNIT 2.1), nondenaturing polyacrylamide gel electrophoresis and electroelution of DNA onto DEAE membranes (UNIT 2.7), denaturing polyacrylamide gel electrophoresis (UNIT 7.6), mobility shift DNA-binding assay (UNIT 12.2), and autoradiography (APPENDIX 3)

Prepare deoxyuracil-substituted DNA probe

1. ³²P-label the 5' ends of oligonucleotide primers with T4 polynucleotide kinase.

Primers should be designed for PCR amplification (UNIT 15.1) and should contain sequences that flank the protein-binding site and are on opposite strands (Fig. 12.3.1).

2. Set up two parallel 50- μl PCR reactions containing the following ingredients:

5 μl DNA fragment containing protein-binding site (0.2 pmol)
5 μl of one ³²P-labeled oligonucleotide primer (20 pmol)
5 μl of other oligonucleotide primer (unlabeled; 20 pmol)
5 μl 2 mM 4dNTP mix
5 μl 0.5 mM dUTP
5 μl 10 \times *Taq* polymerase buffer
19 μl H₂O
1 μl *Taq* polymerase (5 U).

BASIC PROTOCOL

DNA-Protein Interactions

12.3.3

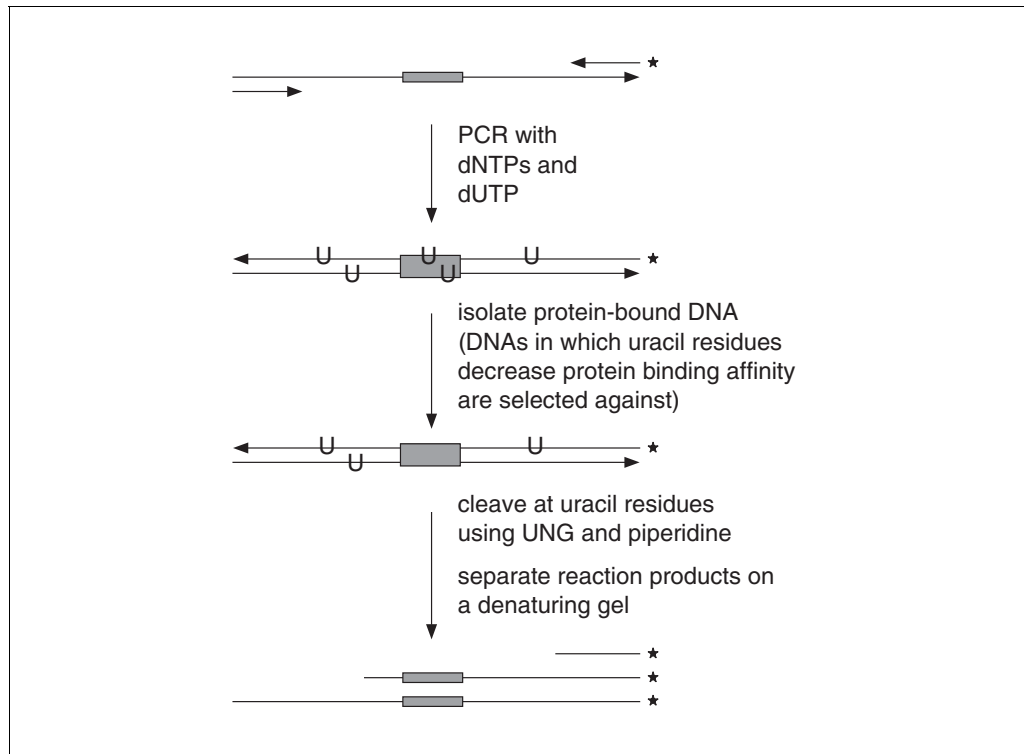


Figure 12.3.1 Uracil interference. An oligonucleotide or restriction fragment (long arrow) containing a protein-binding site (hatched box) is amplified by PCR using one unlabeled primer (short arrow) and one 5'-labeled primer (asterisk) in the presence of dGTP, dATP, dCTP, dTTP, and dUTP, producing reaction products in which deoxyuracil is randomly substituted for thymine on both strands. This collection of DNA molecules is incubated with the protein of interest and DNA molecules containing deoxyuracil substitutions that do not interfere with protein binding are selected by purifying the DNA-protein complex away from unbound DNA. The resulting DNA is cleaved at uracil residues using uracil-*N*-glycosylase followed by piperidine, and the reaction products are separated on a denaturing polyacrylamide gel. (Reprinted from Pu and Struhl, 1992, by permission of Oxford University Press.)

Carry out 8 cycles of PCR amplification.

The two PCR reactions, which differ with respect to which oligonucleotide primer is ³²P-labeled, result in binding-site probes that are specific for the individual DNA strands. The precise conditions for carrying out PCR amplification will depend on the length and base-composition of the primers (UNIT 15.1). Additional PCR cycles can be carried out if sufficient product is not generated.

3. Purify PCR products by electrophoresis on a native (nondenaturing) polyacrylamide gel followed by phenol extraction and ethanol precipitation.
4. Measure pellet for Cerenkov counts in a scintillation counter to determine cpm. Resuspend in TE buffer at ~20,000 cpm/μl.

Isolate protein-DNA complex and free probe

5. Carry out a DNA-binding reaction using 10⁵ cpm of probe in 50 μl. Electrophorese the binding reaction on a native (nondenaturing) polyacrylamide gel.

If the amount of the protein-DNA complex is not sufficient, the binding reaction can be scaled up further and multiple gel lanes can be used.

6. Autoradiograph gel, cut out bands corresponding to the protein-DNA complex and free probe, and purify DNA from gel slices. Resuspend each DNA sample in 25 μ l TE buffer.

The exposure time depends on the quantity of protein-DNA complex. An exposure time of 1 hr is usually sufficient, although longer exposures (12 hr) may be necessary in some cases.

Cleave DNAs at uracil residues and analyze on sequencing gels

7. Set up the following uracil-*N*-glycosylase reaction (50 μ l final):

19 μ l H₂O
5 μ l 10 \times *Taq* polymerase buffer
25 μ l DNA (from step 6)
1 μ l uracil-*N*-glycosylase (1 U).

Incubate 60 min at 37°C. Stop the reaction by ethanol precipitation.

8. Resuspend DNA pellets in 100 μ l of 1 M piperidine. Incubate 30 min in a 90° to 95°C water bath. Place a glass plate over tubes to keep tops from popping open. After piperidine cleavage, place tubes on dry ice.
9. Make holes in tops of tubes with a large needle and lyophilize samples in a vacuum evaporator (e.g., Speedvac) for 1 hr or until dry. Add 100 μ l water, freeze, and lyophilize again. Repeat addition of water, freezing, and lyophilizing. Measure the samples for Cerenkov counts.
10. Add stop/loading dye to the pellet. Heat 5 min at 95°C and quickly chill on ice. Electrophorese samples from free probe and bound complex on a 6% or 8% sequencing gel and autoradiograph.

To easily compare the samples, it is critical to equalize the number of counts prior to electrophoresis. For an overnight exposure with an intensifying screen, 3000 cpm is sufficient.

REAGENTS AND SOLUTIONS

DMS reaction buffer

50 mM sodium cacodylate, pH 8.0
1 mM EDTA, pH 8.0
Store at 4°C

DMS stop buffer

1.5 M sodium acetate, pH 7.0
1 M 2-mercaptoethanol
Store at 4°C

COMMENTARY

Background Information

Methylation interference was used initially by Siebenlist and Gilbert (1980) to study the interaction of RNA polymerase and the T7 promoter. This technique utilized the chemistry developed for the G reaction in DNA sequencing (Maxam and Gilbert, 1980). The principle of uracil interference was established by Goeddel et al. (1978), who synthesized oligonucleotides in which individual thymine residues

were replaced by uracil (for a review, see Levarie, 1987). To simultaneously analyze all the thymines in a given DNA fragment, randomly distributed deoxyuracil substitutions are introduced by PCR in the presence of dUTP, and uracil-specific cleavage is accomplished by treatment with uracil-*N*-glycosylase and piperidine (Pu and Struhl, 1992). Both interference methods have been coupled with the mobility shift DNA-binding assay (*UNIT 12.2*), re-

sulting in relatively rapid analyses of the contact sites for DNA-binding proteins.

Interference assays and DNase I protection assays (*UNIT 12.4*) are the two procedures typically used to map regions of DNA-protein interaction. Interference assays have several advantages over DNase I protection. First, interference analyses are not hindered by incomplete binding, because all of the probe in a retarded complex is bound by protein. In contrast, DNase I protection experiments require protein titrations to fully saturate probe DNA. Another disadvantage of DNase I digestions is that protein exchange from the binding site during the digestion can lead to cutting within the binding site. Finally, interference assays allow the determination of specific nucleotides that are in relatively close contact with the DNA-binding protein. As a result, much more information is obtained about the binding site than from the comparatively large protected region produced by a DNase I footprint. Methylation interference has determined single nucleotide differences in binding site specificities that could not be determined by DNase I protection (Baldwin and Sharp, 1988).

Although interference does not measure direct protein contacts with nucleotides, it permits detection of nucleotides that are closely apposed to the proteins. DMS methylates guanine residues at the N-7 position that protrudes into the major groove, and also at the N-3 position of adenines that protrudes into the minor groove. Adenines can be detected in this assay but these reactions are much weaker than the G reactions (Maxam and Gilbert, 1980). Uracil substitutions result in substitution of hydrogen for thymine 5-methyl groups, which lie in the major groove and are involved in short-range hydrophobic interactions.

In another interference procedure (Brunelle and Schleif, 1987), depurinated and depyrimidinated DNA probes are substrates in the binding reaction. Bound and free probes are separated by native (nondenaturing) gel electrophoresis and cleaved by piperidine, allowing detection of proteins closely apposed to T and C, as well as G and A, residues.

Critical Parameters

Maximizing the amount of specific complex is critical for DNA-binding activities of lower abundance or affinity. Before beginning the methylation interference, it is necessary to optimize conditions for the mobility shift DNA-binding assay. The degree of scale-up is important in generating enough cpm for analysis.

It is important to analyze protein-DNA interactions on both strands. In some cases no contacts are detected on one strand of the probe.

Troubleshooting

In methylation interference, a very large DNA pellet is sometimes observed after the first ethanol precipitation. If this occurs, fresh DMS stop buffer should be made or less 2-mercaptoethanol should be used in the buffer. It is also important to solubilize this methylated DNA pellet completely in TE buffer. Heating at 65°C, vortexing, and repeated pipetting may be required as well.

To obtain an interpretable gel following piperidine cleavage, the DNA probe must be free of acrylamide impurities and be labeled at one end only. The DNA probe should be modified at an average of only one guanine per molecule. If the DMS-reacted DNA has no uncleaved probe at the top of the gel and a nonhomogeneous ladder of bands, the time of the DMS reaction should be reduced. For uracil interference, it is crucial that the PCR reaction and the uracil-*N*-glycosylase cleavage proceed efficiently. The DNA probe should be modified at an average of only one uracil per molecule. To achieve this, it may be necessary to alter the dTTP:dUTP ratio in the PCR reaction.

Anticipated Results

For methylation interference, it is assumed that guanines and adenines are located in the binding site and are positioned close to the DNA-binding activity. The bound lane in the sequencing gel has some missing bands relative to the free probe lane. Adenine contacts may be detected upon careful examination (Fig. 12.3.2). Uracil interference experiments should look similar except that missing bands will correspond to thymine residues in the original DNA fingerprint.

Time Considerations

Preparation of a probe labeled on a single end requires a full day's work. This includes enzymatic treatments and electrophoresis required for probe purification. Ethanol precipitations are convenient stopping points. Methylation requires ~1 to 2 hr. The mobility shift binding assay and exposure require 3 to 5 hr, but may take 12 hr if longer exposure times are necessary. DNA isolation requires ~2 hr. Piperidine cleavage and lyophilization require 3 to 5 hr. Electrophoresis requires 2 to 3 hr followed by a 12- to 24-hr exposure. Thus, a total time of ~3 days is required to perform the procedure.

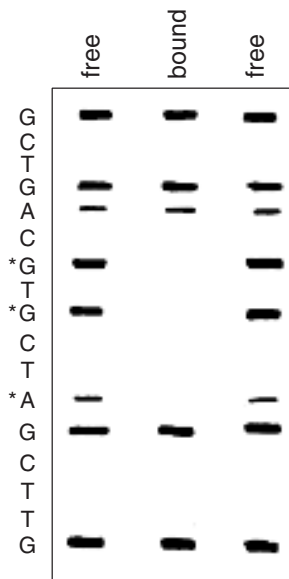


Figure 12.3.2 Hypothetical autoradiogram of a methylation interference experiment, as described in Anticipated Results. Guanines and adenines that interfere are indicated by asterisks.

Literature Cited

- Baldwin, A. and Sharp, P. 1988. Two transcription factors, H2TF1 and NF- κ B, interact with a single regulatory sequence in the class I MHC promoter. *Proc. Natl. Acad. Sci. U.S.A.* 85:723-727.
- Brunelle, A. and Schleif, R. 1987. Missing contact probing of DNA-protein interactions. *Proc. Natl. Acad. Sci. U.S.A.* 84:6673-6676.
- Goeddel, D.V., Yansura, D.G., and Caruthers, M.H. 1978. How lac repressor recognizes lac operator. *Proc. Natl. Acad. Sci. U.S.A.* 75:3579-3582.
- Ivarie, R. 1987. Thymine methyls and DNA-protein interactions. *Nucl. Acids Res.* 15:9975-9983.
- Maxam, A. and Gilbert, W. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499-560.
- Pu, W.T. and Struhl, K. 1992. Uracil interference, a rapid and general method for defining protein-DNA interactions involving the 5-methyl group of thymines: The GCN4-DNA complex. *Nucl. Acids Res.* 20:771-775.
- Siebenlist, U. and Gilbert, W. 1980. Contacts between *E. coli* RNA polymerase and an early promoter of phage T7. *Proc. Natl. Acad. Sci. U.S.A.* 77:122-126.

Key References

- Hendrickson, W. and Schleif, R. 1985. A dimer of AraC protein contacts three adjacent major groove regions at the Ara I DNA site. *Proc. Natl. Acad. Sci. U.S.A.* 82:3129-3133.
- Provides good descriptions of methylation and ethylation interference procedures as well as interpretations of major groove contacts.*
- Maxam and Gilbert, 1980. See above.
- Describes DNA labeling, modification by DMS, and polyacrylamide gel electrophoresis.*
- Pu and Struhl, 1992. See above.
- Describes original method for uracil interference.*

Contributed by Albert S. Baldwin, Jr.
University of North Carolina
Chapel Hill, North Carolina

Marjorie Oettinger (uracil interference)
Massachusetts General Hospital
Boston, Massachusetts

Kevin Struhl (uracil interference)
Harvard Medical School
Boston, Massachusetts

DNase I Footprint Analysis of Protein-DNA Binding

UNIT 12.4

Deoxyribonuclease I (DNase I) protection mapping, or footprinting, is a valuable technique for locating the specific binding sites of proteins on DNA. The basis of this assay is that bound protein protects the phosphodiester backbone of DNA from DNase I-catalyzed hydrolysis. Binding sites are visualized by autoradiography of the DNA fragments that result from hydrolysis, following separation by electrophoresis on denaturing DNA sequencing gels. Footprinting has been developed further as a quantitative technique to determine separate binding curves for each individual protein-binding site on the DNA. For each binding site, the total energy of binding is determined directly from that site's binding curve. For sites that interact cooperatively, simultaneous numerical analysis of all the binding curves can be used to resolve both the intrinsic binding and cooperative components of these energies.

The unit is divided into three parts. The basic protocol covers DNase I footprint titration and involves (1) preparation of a singly end-labeled DNA restriction fragment, (2) equilibration of the protein with DNA, (3) exposure of the equilibrium mixture to DNase I, and (4) electrophoretic separation on gels of the denatured hydrolysis products, followed by autoradiography. The support protocol describes (1) densitometric analysis of the autoradiograms to obtain binding data and (2) numerical analysis of the binding data to yield binding curves and equilibrium constants for the interactions at each of the separate sites. An alternate protocol describes the qualitative use of footprinting to identify DNA-binding proteins in crude extracts.

DNASE I FOOTPRINT TITRATION

Materials

- Plasmid DNA containing protein-binding sites (UNIT 1.6)
- Appropriate restriction endonucleases (UNIT 3.1)
- 100% and ice-cold 70% ethanol
- TE buffer (APPENDIX 2)
- Aqueous [α - 32 P]dNTP (3000 to 6000 Ci/mmol; UNIT 3.4)
- 10 \times Klenow fragment buffer (UNIT 3.4)
- Klenow fragment of *E. coli* DNA polymerase I (UNIT 3.5)
- 5 mM 4dNTP mix (UNIT 3.4)
- Deoxyribonuclease I (DNase I; EC 3.1.4.5)
- Assay buffer A or B
- Dry ice
- DNase I stop solution
- DNase I storage buffer
- Formamide loading buffer (UNIT 7.4)
- 10-ml plastic disposable tubes
- Silanized 1.5-ml microcentrifuge tubes (APPENDIX 3)
- Regulated water bath ($\pm 0.1^\circ\text{C}$)
- 12 \times 7.5-in. glass or stainless-steel tray
- Plastic microcentrifuge tube rack with open sides
- Gel comb for DNA sequencing gel with 6-mm lanes spaced on 12-mm centers
- Hamilton syringes with blunt-tip needles (29 G for 0.4-mm gels)
- Additional reagents and equipment for CsCl gradient centrifugation (UNIT 1.7), restriction enzyme digestion (UNIT 3.1), ethanol precipitation (UNIT 2.1), spin col-

**BASIC
PROTOCOL**

**DNA-Protein
Interactions**

12.4.1

Contributed by Michael Brenowitz, Donald F. Seneor, and Robert E. Kingston

Current Protocols in Molecular Biology (1989) 12.4.1-12.4.16

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 7

umn procedure (UNIT 3.4), agarose gel electrophoresis (UNIT 2.5A), electroelution (UNIT 2.6), autoradiography (APPENDIX 3), denaturing polyacrylamide gel electrophoresis (UNIT 2.12), and reversed-phase or ion-exchange chromatography (Elutip-d or DEAE-cellulose, respectively; UNIT 2.6)

Prepare singly ^{32}P end-labeled DNA

Plasmid DNA is prepared by the alkaline lysis method (UNIT 1.6), followed by two sequential CsCl gradient centrifugations (UNIT 1.7). Exposure of the plasmid to shortwave UV illumination should be minimized to prevent nicking of the DNA.

1. Digest ~5 pmol plasmid with a restriction enzyme that generates a 3'-recessed end, 25 to 100 bp from the first protein-binding site (Fig. 12.4.1).
2. Ethanol precipitate the DNA, wash once with 1 ml of cold 70% ethanol, and dry pellet in a Speedvac evaporator. Dissolve pellet in 5 μl TE buffer.

Standard precautions for handling aqueous ^{32}P -labeled nucleotides should be followed throughout the remaining steps (UNIT 3.4).

3. Add 50 μCi each of the appropriate aqueous [^{32}P]dNTPs, 5 μl of 10 \times Klenow fragment buffer, and water to 49 μl . Add 1 μl Klenow fragment (5 to 10 Kunitz units), mix gently, microcentrifuge, and incubate 25 min at room temperature (UNIT 3.5).

To generate end-labeled DNA with high specific radioactivity, perform the Klenow "fill in" reaction of the 3'-recessed end using only [α - ^{32}P]dNTPs.

4. Add 2 μl of 5 mM 4dNTP mix, mix, and incubate 5 min.

The cold chase ensures that all the labeled fragments have the same number of base pairs.

5. Remove unincorporated nucleotides with spin column procedure (UNIT 3.4). Ethanol precipitate DNA as in step 2.

6. Digest with a second enzyme to generate a restriction fragment labeled only at one end and on one strand (Fig. 12.4.1).

The second restriction site should be >150 bp beyond the protein-binding site most distal to labeled end of fragment.

7. Purify the labeled binding-site-containing DNA using agarose gel electrophoresis, followed by electroelution and reversed-phase chromatography.

Because the DNA fragment is radiolabeled, staining the gel with ethidium bromide (as described in UNIT 2.6) is unnecessary. Instead, wrap the gel in a single layer of plastic wrap. Autoradiograph the gel for 5 to 15 min and use the developed film as a template for excising the band containing the binding site from the gel.

Reversed-phase (using Elutip-d) or ion-exchange (using DEAE-cellulose) chromatography of the labeled DNA is essential for removing impurities and obtaining high-quality footprints (see UNIT 2.6).

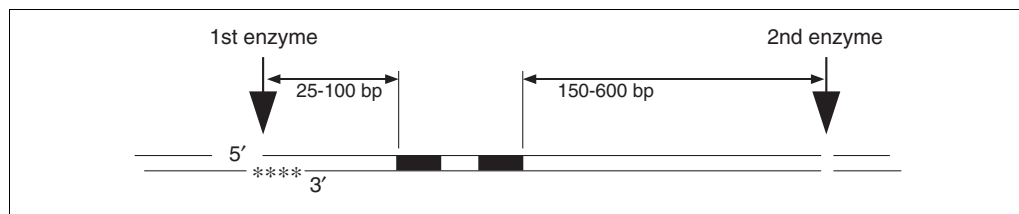


Figure 12.4.1 Correct positioning of protein-binding sites from the restriction cuts used to generate a linear, singly end-labeled fragment. Black boxes represent protein-binding sites. Asterisks (*) represent [^{32}P]dNTP incorporated in the Klenow labeling reaction.

- Ethanol precipitate DNA as in step 2. Dissolve DNA in 100 μl TE buffer, pH 8.0. Store at 4°C. Do not freeze.

Recovery of the small quantity of labeled DNA is improved by extending the -70°C precipitation for several hours.

- Determine radioactivity of the DNA by dissolving a 0.5- μl aliquot in an aqueous scintillation cocktail and counting.

Up to 80% of theoretical label incorporation has been achieved using this procedure. This figure, along with the specific radioactivity of the [^{32}P]dNTPs and the number of bases filled in, can be used to estimate the molar concentration of binding-site DNA. Storage of labeled DNA for longer than 2 weeks results in unacceptable background levels due to radiochemical nicking of the DNA.

Equilibrate protein with DNA

- Prepare $(n + 2) \times 180 \mu\text{l}$ assay buffer in a 10-ml disposable plastic tube, where n is the number of binding reactions in the experiment. Determine the volume of ^{32}P -labeled DNA to yield 10,000 to 15,000 cpm/lane. Add the ^{32}P -labeled DNA to assay buffer. Vortex gently.

The necessary number of binding-reaction mixtures (i.e., data points) for an experiment depends on the relative affinities of the binding sites if multiple sites are present, and whether the shapes of the binding curves need to be well defined. We routinely assay 18 to 24 binding-reaction mixtures. An absolute upper limit is dictated by the number of sequencing gel lanes (see commentary).

*The assay buffer used (e.g., A or B) depends on the nature of the protein, the protein-DNA system, and the questions being addressed. In all cases, millimolar concentrations of Mg^{++} and Ca^{++} are required for DNase I activity. Also, although the specific binding affinity of most DNA-binding proteins decreases with increasing monovalent cation concentration, the specificity (ratio of specific site to nonspecific affinities) often increases. Assay buffers A and B are examples of buffers used in DNA-binding studies of the *E. coli* Gal and λ cI repressors, respectively.*

- Aliquot 180 μl of assay buffer containing ^{32}P -labeled DNA into each of $n + 1$ silanized 1.5-ml microcentrifuge tubes. The extra tube is a control to which no DNase I is added.
- Prepare a series of serial protein dilutions that cover the range of concentrations to be analyzed.

The ligand concentrations should span a range from 0% to $\geq 99\%$ saturation of all of the protein-binding sites. This requires a concentration range of four orders of magnitude for even a single DNA-binding site. Site heterogeneity increases the required range. The final ligand concentrations in the binding-reaction mixtures should define an evenly spaced, logarithmic series with at least several points to define each asymptote of the titration curves (plotted versus log ligand concentration; Fig. 12.4.2).

- Pipet the desired amount of protein dilution (2 to 20 μl) into each tube. Add assay buffer (without ^{32}P -labeled DNA) to a final total volume of 200 μl per tube.

To have a sufficient amount of ^{32}P -labeled DNA for fragment detection and still keep the concentration of the DNA low, use 200- μl volumes (see commentary).

Propagation of pipetting errors can be a problem in steps 12 and 13. Accuracy is critical: wipe the outside surface of the pipet tip so that no liquid adheres to the surface.

- Gently mix each tube and microcentrifuge briefly. Equilibrate samples in a regulated water bath 30 to 45 min at the desired temperature.

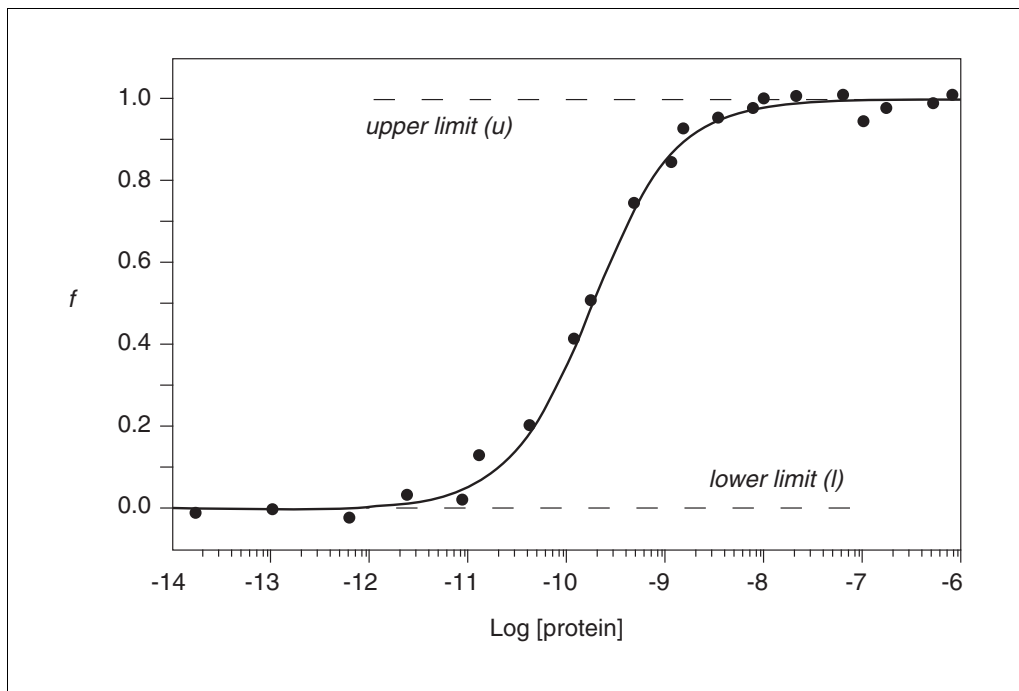


Figure 12.4.2 Footprint titration curve showing the proper range and spacing of protein concentrations. The upper and lower limits of the titration curve are indicated by dashed lines. See Equation 1 for a definition of f .

This equilibration time is a general guideline. Equilibration of the protein-DNA mixture should be verified by conducting assays using different equilibration times.

Perform DNase I footprint titration

15. Prepare an ethanol-dry ice bath in a glass or stainless-steel tray. Place a plastic microcentrifuge tube rack in the bath. The liquid level should cover tubes to just below their tops. Prepare an excess of DNase I stop solution (700 μ l is required per tube) and place in the bath to cool.
16. Prepare ≥ 500 μ l dilute DNase I solution in assay buffer *without* BSA and calf thymus DNA. Place solution in regulated water bath with samples and allow it to equilibrate.

The necessary DNase I concentration not only depends on the supplier but also on the storage and assay conditions. Routinely, 2 mg/ml enzyme is stored in DNase storage buffer at -70°C with no loss of activity. Repeated freezing and thawing drastically reduces activity (requiring storage in small aliquots); therefore, DNase I concentration must be determined empirically (see critical parameters).

Expose equilibrium mixture to DNase I

17. To start DNase I exposure, pipet exactly 5 μ l dilute DNase I into the first tube, vortex gently but quickly to mix, set timer to 2 min, and immediately return tube to the regulated bath.

To obtain titrations of high quality, maintain uniform DNase I exposure. To minimize differences, carry out all manipulations in the same order and at the same pace for each tube.

18. After precisely 2 min, rapidly add 700 μ l DNase I stop solution to the tube. Vortex vigorously to mix the viscous stop solution. Place tube in ethanol-dry ice bath.

The exposure time and DNase I concentration can be varied (between, but never within, experiments) as long as the product of the two (and hence, the number of DNA nicks produced) remains constant. The thermodynamic validity of the method depends on avoiding significant perturbation of the binding equilibria; thus, overall exposure time is not an issue.

19. Repeat steps 17 and 18 for each tube.

20. Add stop solution to the extra tube prepared in step 11 (without DNase I).

This tube provides a control for the presence of excessive background nicking of the DNA. Such nicking can occur upon preparation or over-long storage of labeled DNA.

21. Precipitate DNA in ethanol–dry ice bath for ≥ 15 min. Start timing when last tube is placed in the bath.

The minimum time required for quantitative precipitation under these conditions is 15 min. Longer times are not deleterious.

22. Microcentrifuge tubes 15 min. Carefully remove supernatant with a pipet. Add 1 ml cold 70% ethanol. Microcentrifuge again 5 min. Remove supernatant very carefully with a pipet; the pellet is frequently very loosely attached to the tube at this point. Repeat wash step. Dry pellets in a Speedvac evaporator (~10 to 15 min).

23. Resuspend DNA in 5 μ l formamide loading buffer. Pipet buffer onto the upper inside surface of tube and tap the tube to allow buffer to settle. Vortex vigorously and microcentrifuge several seconds. Repeat twice.

Quantitative resuspension of the pellet is mandatory but can be very difficult. Incomplete resuspension results in substantial variations in the amount of DNA loaded onto each lane. The difficulty appears to be in getting the DNA pellet and loading buffer in physical contact. High microcentrifuge speeds and commercially available silanized tubes facilitate good pellet formation. Vigorous vortexing also helps and does not appear to damage the DNA at this step of the procedure.

24. If the sequencing gel is not to be run immediately, store samples overnight below -70°C .

Isolate and autoradiograph denatured hydrolysis products

25. Prepare and pre-electrophorese a polyacrylamide DNA sequencing gel (UNIT 2.12).

The percentage of acrylamide depends on the size of the fragments to be separated, i.e., on the distance between the first binding site and the labeled end, and on the size of the binding-site region. In most cases, 6% to 8% is optimal. For systems with widely separated binding sites, 0.4- to 1.2-mm-thick gels are desirable.

For quantitative analysis the spacing between lanes of commercial combs (usually 3 mm) is insufficient to prevent bleeding of density across lanes when using ^{32}P . This can produce significant errors in analysis and is easily avoided by using a sample comb with 6-mm lanes and 6-mm spacing. Such combs can be cut by most university machine shops or custom ordered from commercial suppliers.

26. While pre-electrophoresing the gel, heat samples at 90°C in a dry-block heater 5 to 10 min, followed by immediate quenching in wet ice. When the temperature of the gel reaches 50° to 55°C , disconnect the power supply. Load samples carefully, being sure to take up all of the loading buffer in each tube. Either Hamilton syringes with blunt-tip needles (29 G for 0.4-mm gels) or thin micropipet tips are suitable. Care should be taken to prevent cross contamination of the wells.

It is essential that the temperature of the dry-block heater not exceed 90°C . Higher temperatures cause large portions of the DNA to remain in the sample wells and also cause streaking.

27. Electrophorese until the protein-binding sites migrate to, or just below, the middle of the gel. Fix (gradient gels only) and dry gel using standard procedures (*UNIT 2.12*).
28. Autoradiograph the dried gel with preflashed Kodak X-Omat AR film and a single calcium tungstate intensifying screen at -70° to -85°C as described in *APPENDIX 3*. Set the preflashed side of the film against the gel.

Preflashing the X-ray film is essential for overcoming threshold behavior of the film and achieving linear response to radioactivity (APPENDIX 3). With 10,000 to 20,000 cpm loaded per lane and 25% to 50% of the DNA nicked at least once, 24- to 48-hr exposures are required. The optical density (OD) range analyzable is highly dependent on the type of densitometer used (see below). An absolute limit to the darkest area to be analyzed is dictated by the significant deviation of the response of Kodak X-Omat AR film from linearity, >1.6 OD. Make two or three autoradiograms of each gel (at different exposure levels) to ensure a correct exposure.

29. Develop films under consistent and controlled conditions.

It is suggested that films be developed by hand in a standard temperature-regulated 5-gallon tank with Kodak GBX developer and fixer for the times recommended by the manufacturer (i.e., 4 min at 70°F). Automatic X-ray film processors or other developers or stabilizers have not been tested with regard to either image or background uniformity or optical density linearity.

30. Store films carefully with interleaved paper. Scratches, fingerprints, and dirt will appear as optical signals, indistinguishable from ^{32}P . The autoradiograms are now ready for quantitation.

SUPPORT PROTOCOL

QUANTITATION OF PROTEIN-BINDING EQUILIBRIA BY DENSITOMETRIC AND NUMERICAL ANALYSES

This protocol outlines the resolution of binding curves from autoradiograms and the numerical analyses of those curves to obtain equilibria binding constants. The steps involved are (1) digitization of the autoradiogram, (2) integration of optical density over the bands corresponding to binding site(s), (3) correction of integrated optical density for film background optical density, (4) standardization of corrected, integrated optical density to the total amount of DNA loaded onto each lane, and (5) calculation of fractional protection of the sites, proportional to fractional saturation.

Accurate densitometric analysis is critical to obtaining thermodynamically valid individual site-binding curves. The common practice of visually estimating titration half-points from autoradiograms is unacceptable for a number of reasons. First, the eye responds to light intensity, whereas optical density—the quantity proportional to the concentration of labeled DNA in a given band—is the log of the intensity. Thus, visual estimates of titration half-points are usually systematically wrong. Second, autoradiograms that look acceptable show appreciable saturation of dark bands. Films that are within the linear densitometric range appear light or washed out, making the intensity differences difficult to estimate. Third, visual inspection allows no estimate of the precision of the experimental data or derived binding constants. Fourth, visual estimation is not objective. There is a tendency to group estimated half-points into classes (i.e., relative dissociation constants of 1, 2, and 25) which may obscure real differences in the data. Finally, in cooperative systems, the half-point of a titration corresponds to no single equilibrium binding constant (Ackers et al., 1983).

Required hardware includes a two-dimensional optical scanning device and a micro- or minicomputer with a high-resolution graphics display. Although some investigators rely on one-dimensional analog traces of each of the separate film lanes, we consider

two-dimensional scanning to be essential for accurate quantitation. Two-dimensional scanning avoids the systematic error that is made when either the maximum peak height or the area of a one-dimensional slice through a band is equated with the total material in the band. Two-dimensional scanning also minimizes the effect of electrophoretic abnormalities, e.g., band skewing and distortion.

The variety of appropriate scanning devices currently in use include flat bed scanners (e.g., some laser scanners), flying spot detectors (e.g., rotating drum scanners), charge coupled devices (CCD) or photodiode array-imaging cameras, and video cameras. The digital image consists of an array of values corresponding to the diffuse optical density of picture elements (pixels). The pixels must be small enough to resolve closely spaced bands clearly. Spatial resolution of 250 to 200 μm is optimal. The scanner must resolve at least 256 levels of gray. This optical resolution is adequate to quantitate 0 to 1.6 OD when the levels represent equal optical density increments, such as those produced by flying spot or laser scanning devices. Camera devices, however, measure transmitted intensity (I) rather than optical density. The transformation to optical density ($\text{OD} = \log I/I_0$) must be computed. I_0 , the incident intensity, is determined by scanning the light source with no film present. The log transformation reduces the range that can be scanned to 0 to 1.0 OD (100% to 10% transmission) but benefits from the greater sensitivity of the scanner in that range.

For analysis, the film image should be displayed on a graphic display monitor with a minimum of 16 gray or color levels. Display devices with 256 levels are preferred. Spatial resolution $\geq 1024 \times 1024$ pixels is preferred, although lower resolution display devices can be used.

Computer software to expedite both film scanning and analysis is available for PC/DOS and VAX/VMS computers. Information concerning these programs can be obtained by contacting the authors. Also, some commercially available image-analysis software packages have functions that can be used for the analysis of titration autoradiograms. The protocol described can be implemented using such systems.

1. Use a film-scanning device to construct a two-dimensional digital image of the film.
2. Integrate the optical density in each lane (i.e., at each ligand concentration) for the protected DNA bands in each binding site (Fig. 12.4.3). The integrated optical density is the sum of the optical density values of the individual pixels within a contour defining the band.

Protected bands are easily identified by visual inspection. In principle, any protected bands can be chosen to follow the titration. It is best to consider all protected bands by analyzing contiguous groups of bands as single blocks (Fig. 12.4.3). This simplification is appropriate whenever all the bands included titrate equivalently.

Choose the ends of blocks by drawing boundaries at the minimum density level between bands of similar maximum density that can be visualized in all lanes (i.e., even at high protection). It is crucial that each block include precisely the same bands in every lane of the film. If there are multiple, closely spaced binding sites, ensure that the protection of bands near the edge of one site is not influenced by binding of the neighboring site.

3. Correct the integrated optical density value of each block in each lane for the local film background optical density. This can be calculated using the image-analysis software as follows:
 - a. Define small rectangles in the center of the interlane space and at the position of each block (Fig. 12.4.3).

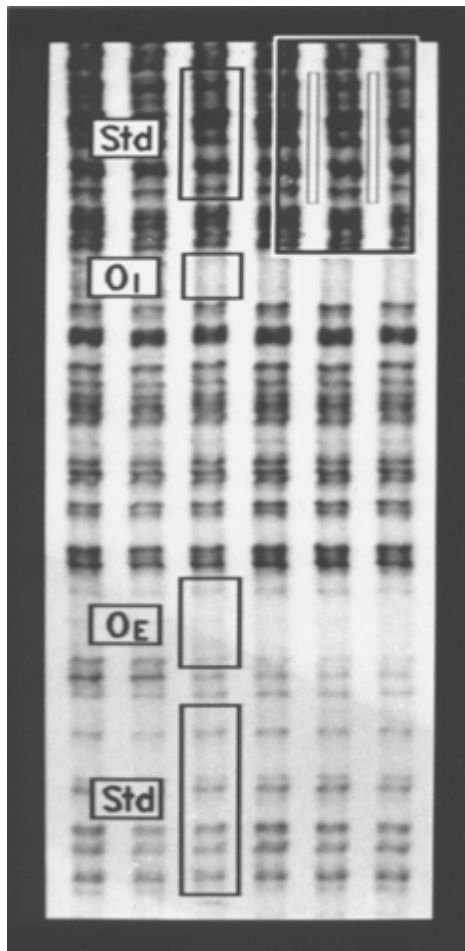


Figure 12.4.3 Portion of a footprint titration autoradiogram showing the different blocks that need to be defined for analysis. "Std." denotes a standard block used to correct for variations in the DNA loaded onto each lane. O₁ and O₂ denote protein-binding sites 1 and 2, respectively. The insert shows regions for a standard block for which background determinations will be made.

- b. Calculate a histogram plot of the pixel optical density values in each rectangle and define the local background as the most probable pixel optical density.
- c. For each block and lane, average the background values on either side of the lane.
- d. Multiply this average value by the number of pixels in the block.
- e. Subtract the product from the integrated optical density to give the corrected integrated optical density for that block.

The variability in the background density of the film is large enough to warrant local determination of background and correction of the integrated optical densities (a to e above).

4. Standardize corrected, integrated optical density values for binding-site blocks to total DNA in the lane.
 - a. Choose one or more blocks, excluding titrating sites or very hypersensitive bands, to represent the total DNA concentration in a lane.
 - b. Follow steps 2 and 3 to determine the corrected, integrated optical density for these standard blocks.
 - c. Calculate the optical density ratio ($D_{\text{site}}/D_{\text{std}}$) for each binding-site block. If more than one standard block is chosen, take the sum of their values as D_{std} .

It is practical and sufficient to choose two large standard blocks (Fig. 12.4.3), one nearer the origin of electrophoresis than the binding site(s) and one farther away. Carefully check each standard for systematic, ligand concentration-dependent

variations that invalidate its use. Variations might indicate nonspecific protein binding (decreasing density with increasing ligand), redistribution of DNase I on the DNA fragment (increasing density with increasing ligand) as binding sites are saturated and protected (Dabrowiak and Goodisman, 1989), or poor experimental technique.

5. Convert OD ratios to fractional protection (f), according to

$$f = 1 - \left\{ \frac{(D_{n,site}/D_{n,std})}{(D_{r,site}/D_{r,std})} \right\} \quad (1)$$

where n refers to any lane with finite protein ligand concentration and r refers to a reference lane (that must be included in every experiment) in which no protein ligand has been added to the reaction mixture. Plot data, f versus [ligand], for each binding-site block to define the binding curves (Fig. 12.4.2).

If outlined protocols are followed and if control experiments (see Brenowitz et al., 1986a) demonstrate no effect of exposure to DNase I on the protein–DNA binding equilibria, then f yields the fractional saturation (Y). Observed values of f frequently do not span the entire range, 0 to 1. One reason for this is the failure to account for nicking of the DNA prior to DNase I exposure (i.e., $D_{n,site} > 0$, even at saturating ligand concentration). Therefore, in analyzing data, it is best to treat f as defining a transition curve, proportional to the binding curve, and to include both curve endpoints (u and l for upper and lower endpoints, respectively) as adjustable parameters (Fig. 12.4.2).

For a single binding site or multiple sites that do not interact cooperatively, the binding curve is the familiar Langmuir isotherm given by

$$\bar{Y} = \frac{k[P]}{1 + k[P]} \quad (2)$$

where $[P]$ is the free protein concentration and k is the microscopic equilibrium association constant (the more complex case of multiple, interacting binding sites is discussed in the commentary). Thus, the equation used to analyze the data is

$$f = m \times \left\{ \frac{k[P]}{1 + k[P]} \right\} + b \quad (3)$$

with $m = 1/(u - l)$ and $b = 1/(l - u)$ and other units as defined above. Methods of nonlinear least-squares parameter estimation (Johnson and Frasier, 1985) should be used to estimate k , l , and u , along with their confidence limits and the variance of the fit. Pay particular attention to the pattern of residuals (i.e., difference between fitted and observed f) in evaluating the goodness of fit of the nonlinear least-squares parameter estimation.

DNASE FOOTPRINTING IN CRUDE FRACTIONS

DNase footprinting is frequently used to locate proteins in crude fractions, thus providing an assay for use during purification. The purpose of this type of footprinting is usually to locate a particular binding activity, not to characterize the strength of the interaction. Thus one first establishes conditions under which a complete footprint can be observed, and then assays fractions of a crude extract for that activity. For this purpose, it is necessary to develop conditions that allow binding of the protein of interest to a specific site, but that do not allow substantial nonspecific binding. This is accomplished by varying conditions such that nonspecific DNA-binding proteins are inhibited from binding to the labeled DNA containing the site of interest—i.e., by including substantial amounts of competitor DNA in the reaction, or by increasing the salt concentration of the reaction.

ALTERNATE PROTOCOL

DNA-Protein Interactions

12.4.9

1. Prepare singly ^{32}P end-labeled DNA as in steps 1 to 9 of the basic protocol.
2. Characterize dilutions of stock DNase I for a concentration of DNase I that will produce ~50% non-nicked DNA. Set up a 200- μl reaction containing probe and assay buffer, but no added protein, as described in step 10 of the basic protocol. Add 5 μl of diluted DNase I, mix gently, and microcentrifuge briefly.

The amount of DNase I required will vary with lot and storage conditions, but is frequently ~0.1 mg/ml in the dilute solution added to the reaction. Assay composition can dramatically affect DNase I activity; see critical parameters.

3. Incubate 2 min and stop as described in step 18 of the basic protocol. Ethanol precipitate and analyze products on a sequencing gel.
4. Perform a footprinting reaction as described in the basic protocol. Instead of using varying amounts of protein, however, use a constant amount of a crude fraction that is known to contain the activity of interest. Perform a set of assay conditions as described in Table 12.4.1, in order to find a condition that produces a footprint.

The goal of this experiment is to find conditions for the specific binding of the protein of interest without interference from nonspecific DNA-binding proteins. In the experiment described in the table, poly(dI-dC) concentration is varied in an attempt to find an optimal concentration that will “soak up” contaminating nonspecific proteins. A similar protocol can be used to vary salt concentration (e.g., try 200 mM, 300 mM, 500 mM, and 800 mM) or calf thymus DNA concentration. It may also be prudent to vary extract concentrations, as the level of any DNA-binding protein can vary dramatically.

It is important to vary DNase I concentration above that needed for appropriate digestion of labeled DNA when no crude fraction is present. This is because DNase digestion is inhibited in many crude fractions.

Table 12.4.1 Assay Conditions for Footprinting in Crude Fractions

Reaction no. ^a	10 \times assay buffer (μl)	Probe (cpm)	Crude fraction (μl)	Poly(dI-dC) (μg)	DNase (μl) ^b
1	20	10-20k	0	0.4	5 (1 \times)
2	20	10-20k	12	0.1	5 (1 \times)
3	20	10-20k	12	0.4	5 (1 \times)
4	20	10-20k	12	2.0	5 (1 \times)
5	20	10-20k	12	10.0	5 (1 \times)
6	20	10-20k	0	0.4	5 (4 \times)
7	20	10-20k	12	0.1	5 (4 \times)
8	20	10-20k	12	0.4	5 (4 \times)
9	20	10-20k	12	2.0	5 (4 \times)
10	20	10-20k	12	10.0	5 (4 \times)
11	20	10-20k	0	0.4	5 (16 \times)
12	20	10-20k	12	0.1	5 (16 \times)
13	20	10-20k	12	0.4	5 (16 \times)
14	20	10-20k	12	2.0	5 (16 \times)
15	20	10-20k	12	10.0	5 (16 \times)

^aReaction volume is 200 μl .

^bDilutions are shown in parentheses. 1 \times refers to the dilution that leaves 50% unnicked DNA (step 2 of alternate protocol). 4 \times and 16 \times refer to higher concentrations (as indicated) of DNase I.

REAGENTS AND SOLUTIONS

Assay buffers A and B

Assay buffer A:

10 mM Tris-Cl
5 mM MgCl₂
1 mM CaCl₂
2 mM dithiothreitol (DTT)
50 µg/ml bovine serum albumin (BSA)
2 µg/ml calf thymus DNA
100 mM KCl
Titrate to pH 8.0

Assay buffer B:

10 mM bis-Tris-Cl
2.5 mM MgCl₂
1 mM CaCl₂
0.1 mM EDTA
200 mM KCl
100 µg/ml BSA
2 µg/ml calf thymus DNA
Titrate to pH 7.0

DNase I stop solution

For each 200-µl binding-reaction mixture:

645 µl 100% ethanol
5 µl tRNA stock solution (1 mg/ml)
50 µl saturated ammonium acetate
Store at -20°C for up to 1 to 2 weeks

DNase I storage buffer

50 mM Tris-Cl, pH 7.2
10 mM MgSO₄
1 mM DTT
50% glycerol
Store DNase I in small aliquots at -70°C

COMMENTARY

Background Information

In both prokaryotes and eukaryotes the binding of proteins to specific DNA sequences is critical to the regulation of many cellular processes. In addition, cooperative interactions among bound proteins (perhaps also with non-DNA-binding proteins) are known to be critical to the regulation of transcription, replication, and recombination. A key to understanding the molecular mechanism by which these proteins regulate cellular processes is the measurement of the free energy of binding of these proteins to DNA and the free energies that describe the cooperative interactions occurring among them.

The DNase I footprint titration method, with subsequent quantitative analysis, has been shown to yield thermodynamically valid individual site-binding curves for site-specific protein-DNA interactions. In systems containing multiple sites, these curves can be used to resolve intrinsic binding constants, describing the binding of a protein to a site in the absence of other sites, and constants that describe the cooperative interactions among proteins bound

to multiple sites. The power of the method is that it is, in principle, not limited by the number of sites or the number of proteins.

The method is based on the fact that DNase I probes, *without perturbing*, the equilibrium distribution of binding protein with DNA. Control experiments show that DNase I does not perturb the equilibrium between *cI*- and Gal-repressors and their respective binding sites under the solution conditions we have studied. This conclusion may not be valid, however, for all systems under all experimental conditions. Thus, control experiments must be conducted for new systems and drastically new conditions. Some general considerations are discussed in critical parameters.

DNase I protection mapping, or footprinting, was developed as a qualitative technique to locate the specific protein-binding sites on DNA (Galas and Schmitz, 1978). Subsequently, it was used to compare the relative affinities of *cI* repressor to multiple binding sites (Johnson et al., 1979). Using the data of Johnson et al., Ackers et al. (1982) resolved the intrinsic binding and cooperative free energies

for the binding of *cI* repressor to the O_R operator of the bacteriophage λ . A general theory for individual site-binding curves has been developed (Ackers et al., 1983). The fact that quantitative DNase I footprint titrations yield thermodynamically valid individual site-binding curves for site-specific protein-DNA interactions was demonstrated by Brenowitz et al. (1986a,b). A discussion of the use of the footprint titration method (and other popular binding methods) to determine the intrinsic binding and cooperativity constants for systems of multiple, interacting sites has been presented (Senear et al., 1986). Quantitative footprint titration methods for studying drug-DNA interactions have been developed in parallel with the work on proteins presented in this protocol (Dabrowiak and Goodisman, 1989).

Other probes. As a probe of the occupancy of protein-binding sites, DNase I has several advantages over high-resolution probes such as Fe-EDTA (Tullius et al., 1987) or methidium-propyl-EDTA (MPE) (Hertzberg and Dervan, 1982). First, the uniform footprint produced by DNase I yields large density changes on the autoradiogram that can be accurately quantitated. This facilitates the precise determination of binding curves. Second, divalent cations are known to bind tightly to DNA, are frequently involved in protein structure and self-assembly, and are physiologically important. Unlike MPE and Fe-EDTA, DNase I can be used conveniently with (and requires) divalent cation-containing buffers. Third, the fact that DNase I is an enzyme specific for DNA precludes the possibility of degradation of the binding protein by promiscuous free radicals. Thus, DNase I may be particularly valuable for the analysis of binding proteins sensitive to degradation.

The single base-pair resolution of Fe-EDTA and MPE make these reagents important tools if detailed structural information is desired. There is no a priori reason that they cannot be used to monitor protein or drug titrations with single basepair resolution.

Other binding methods. A virtue of footprinting for quantitative analysis is that protein binding is determined at equilibrium. Unlike other methods, such as the mobility shift assay (UNIT 12.2), the manipulations required to separate and visualize liganded and unliganded DNA are performed after the binding and DNase I-catalyzed reactions are quenched. This feature of the method provides for its application over a wide range of precisely controlled experimental conditions.

In contrast, the use of the mobility shift assay

to obtain titration curves that separately represent each protein-DNA complex with exactly n bound proteins requires that the protein-DNA complexes do not dissociate in the gel during electrophoresis and that their distribution remains unchanged. It has been noted that quantitation of the unbound DNA band is less affected by dissociation of protein-DNA complexes in the gel. If only the unbound DNA is quantitated, however, all of the information concerning potential cooperative free energies is lost. For a theoretical discussion of the determination of cooperativity constants from the mobility shift assay, see Senear et al. (1986).

Critical Parameters

Protein purity. Meaningful determination of absolute equilibrium binding constants requires that the protein and DNA concentrations be known. Thus, the binding protein must be chemically pure, and both its DNA-binding form and specific DNA-binding activity must be known. Only relative binding affinities can be obtained from the analysis of crude extracts or partially purified protein (see alternate protocol).

Concentration of DNA-binding site(s). Accurate determination of equilibrium binding constants requires that the free protein concentration in each equilibrium mixture be known. The relationship of free to total protein is

$$[P]_f = [P]_t - [P-DNA] \quad (4)$$

where $[P]_f$ is free protein, $[P]_t$ is total protein, and $[P-DNA]$ represents DNA-bound protein. The simplest way to deal with this problem is to maintain low concentration of DNA-binding sites in the reactions, so that $[P-DNA]$ is negligible and Equation 4 reduces to

$$[P]_f = [P]_t \quad (5)$$

Because the concentration of specific binding-site-containing DNA in the assay is dictated by the radioactivity needed for good autoradiography, it is important to achieve high specific activity in labeling. Choosing restriction sites such that the Klenow fill-in reaction adds four radiolabeled deoxyribonucleotides to the DNA is useful. By using high-specific-activity nucleotides (3000 or 6000 Ci/mmol), 15,000 to 20,000 dpm/200 μ l assay mixture is achieved with 1 to 10 pM in binding-site DNA. For this simplest case of a protein ligand binding to a single site on the DNA, this concentration can be safely ignored if the dissociation equilibrium constant is 50 to 500

pM or greater. In some cases, DNA-binding proteins self-polymerize before binding to DNA. It is then necessary to calculate the free polymer concentration as the free ligand, $[P]_f$. However, the polymer concentration is buffered by the monomer with which it is in equilibrium, allowing higher concentrations of binding-site-containing DNA to be used or higher affinity binding to be analyzed.

It is also possible to make the correction to $[P]_f$ from $[P]_t$, if the total concentration of binding-site-containing DNA ($[DNA]_t$) in the reactions is known or can be estimated accurately, because

$$[P-DNA] = \bar{Y} \times [DNA]_t \quad (6)$$

where \bar{Y} (being the dependent variable in the experiment) represents the fractional saturation of the binding site(s). For systems with multiple binding sites, \bar{Y} is the sum of the Y_s for each of the individual sites. The procedure is to solve for $[P-DNA]$, hence, $[P]_f$, iteratively. At each iteration of the nonlinear least-squares procedure used to fit the Y values observed (see below), the fitted $Y(s)$ is used to correct each $[P]_t$ to $[P]_f$ for the next iteration. The limitations to this approach are the accuracy with which $[DNA]_t$ is known, the sophistication of the least-squares program used, and the available computing power.

DNase I concentration. The general requirement for frequency of single-stranded nicks introduced into the DNA is an average of one per DNA fragment. Based on a random distribution of nicks and fragments, this corresponds to ~50% of the fragments containing nicks and ~50% left intact (Table 12.4.2). This relaxes the more widely held (and stringent) interpretation of “single-hit kinetics,” in which no molecule is nicked more than once. This

interpretation leads to a very small fraction of DNA molecules being nicked during the assay, so that a significant proportion of the density on the autoradiogram represents nicks introduced during preparation and handling of the DNA more than by DNase I activity. More nicking greatly improves the experimental signal-to-noise ratio. Although this results in molecules being nicked more than once, our experimental results (Brenowitz et al., 1986a,b) indicate no effect on the binding equilibria, as long as the (intact) specific-site-containing DNA molecule is at least an order of magnitude longer than the binding sites. The fraction of uncut DNA following DNase I exposure is readily determined by comparison of the uncut DNA (top band following electrophoresis) from DNA exposed, and not exposed, to DNase I. Quantitation of the amount of DNA in this band can be accomplished by autoradiography (very short exposures) and densitometry, or the bands can be physically excised, dissolved in scintillant, and counted. The latter might be more accurate, but hydrolyzable analogs of bisacrylamide must be used to cast the gels, and acrylamide quenching of fluorescence can be a problem. The distribution of number and average number of nicks per fragment are calculable, based on Poisson statistics (Table 12.4.2).

Nonspecific carrier DNA. Nonspecific carrier calf thymus DNA is included in the assay buffer to maintain the DNase I substrate concentration constant. In the absence of carrier, binding of ligand to the DNA leads to increased nicking of the fragment outside of the binding sites (Dabrowiak and Goodisman, 1989). In our hands, carrier DNA eliminates this effect and hence, the need to apply empirical corrections. It is crucial that the carrier concentration be sufficiently low to avoid nonspecific binding of

Table 12.4.2 Calculated Distribution of DNase I Nicks Among Fragments^a

Uncut ^b (%)	Singly nicked ^c (%)	Multiply nicked ^c (%)	Ave. no. nicks/ fragment ^d
10	23.1	67.0	2.3
20	32.2	47.8	1.6
30	36.1	34.0	1.2
50	34.7	15.3	0.7
75	21.6	3.4	0.3
90	9.5	0.5	0.1

^aCalculated for a 682-bp fragment, before DNase I exposure.

^bExperimentally measured percentage of fragments left after exposure to DNase I.

^cPercentages were calculated assuming a Poisson distribution of nicks.

^dAverage number of nicks per labeled strand calculated assuming a Poisson distribution of nicks.

the protein ligand. Although we have used only calf thymus DNA in our experiments, other carriers including synthetic polynucleotides, e.g., poly(dI-dC), which might exhibit lower affinities for site-specific DNA-binding proteins, should also be effective. Binding to the carrier DNA can be monitored by conducting parallel nitrocellulose filter-binding assays (UNIT 12.8).

Assay buffer conditions. An advantage of DNase I footprinting is its applicability over a broad range of binding reaction conditions. Limitations are of two types: the binding affinity of the protein ligand, and the enzymatic activity of DNase I. The limits of high- and low-binding affinity that can be studied by DNase I footprinting have been outlined above. The extremes of temperature, pH, and salt concentration corresponding to these limits are easily determined for any particular protein-DNA system by conducting preliminary binding experiments, e.g., using filter binding (UNIT 12.8) or mobility shift (UNIT 12.2) methods. In the presence of millimolar amounts of Ca⁺⁺ and Mg⁺⁺, the specific activity of DNase I varies by only severalfold from pH 5 to 9 (most active near pH 7), from 4° to 37°C and from 50 to 200 mM KCl. Simultaneous extremes have not been tested. Changes in DNase I activity of only severalfold are easily compensated by adjusting the DNase I concentration and/or the exposure time. Although DNase I is active in the presence of Mg⁺⁺ alone or Ca⁺⁺ (and some other divalent cations), its activity varies dramatically with solution conditions. Under some conditions, the variations in activity (decreases) may be too great to be compensated without unacceptably high DNase I concentrations.

Analysis of binding to multiple, interacting sites. The footprint titration method is best suited to analyze binding to multiple, cooperatively interacting sites on DNA. No other

method provides the information necessary to separately resolve the microscopic equilibrium constants for intrinsic binding and cooperativity. The key feature of cooperativity is that the binding of ligand to one site depends on the fractional occupancy of all other interacting sites. The individual site-binding curves for this situation and the equations that describe them reflect this fact.

To formulate the binding expressions, first consider all possible configurations of liganded and nonliganded sites (Table 12.4.3). For each configuration, the microscopic equilibrium (association) constants for each of the macromolecular interactions are defined. In Table 12.4.3, the Gibbs free-energy changes associated with the interactions are substituted for equilibrium constants. These are related by the well-known thermodynamic relation, $\Delta G = -RT \ln(k)$. The ΔG s are of two types: binding of ligand to one site in the absence of binding to others (ΔG_i) and the excess (or cooperative) energy for binding to two (or more) site(s) simultaneously (ΔG_{ij}). The fractional probability of any one of the configurations is related to its relative Gibbs energy according to

$$f_s = \frac{\exp(-\Delta G_s / RT) \times [P]^j}{\sum_{s,j} \exp(-\Delta G_s / RT) \times [P]^j} \quad (7)$$

where ΔG_s is the sum of ΔG_i 's and ΔG_{ij} 's for state (configuration) s , R is the gas constant, T is the absolute temperature, and j is the number of ligands bound to s . The denominator is the summation of such terms over all states, s . Y_i (fractional saturation of site i) is the sum of probabilities of all configurations with ligand bound to i , e.g., for two sites (Table 12.4.2):

$$\bar{Y}_1 = f_1 + f_3 \quad (8)$$

$$\bar{Y}_2 = f_2 + f_3 \quad (9)$$

Table 12.4.3 Binding-Site Configurations and Associated Energy States of Protein Bound to Two Sites

Species	Binding sites		Free energy contributions ^a	Total free energy ^a
	1	2		
1	—	—	(Ref. state)	ΔG_{s1}
2	P_2	—	ΔG_1	ΔG_{s2}
3	—	P_2	ΔG_2	ΔG_{s3}
4	P_2 —	P_2	$\Delta G_1 + \Delta G_2 + \Delta G_{12}$	ΔG_{s4}

^aFree energies (kcal/mol), where $\Delta G_{12} = \Delta G_{total} - (\Delta G_1 + \Delta G_2)$; i.e., the cooperative free energy ΔG_{12} is the difference between binding to both sites simultaneously and to each site independently.

Table 12.4.4 Analysis of Gal Repressor with Wild-Type Operator^a

ΔG_1^b	ΔG_2^b	ΔG_{12}^b	s^c
-12.9 ± 0.5	-13.2 ± 0.4	-0.03 ± 0.7	0.065

^aFree energies in kcal/mol.

^b ΔG_1 and ΔG_2 refer to the intrinsic free energies to sites O_1 and O_E , respectively. The monomer-to-dimer association free energy of Gal repressor is -11.6 kcal/mol.

^cSquare root of the variance.

Substituting the expressions developed using Equation 7 into Equations 8 and 9 yields

$$\bar{Y}_1 = \frac{k_1 [P] + k_1 k_2 k_{12} [P]^2}{1 + (k_1 + k_2)[P] + k_1 k_2 k_{12} [P]^2} \quad (10)$$

and

$$\bar{Y}_2 = \frac{k_2 [P] + k_1 k_2 k_{12} [P]^2}{1 + (k_1 + k_2)[P] + k_1 k_2 k_{12} [P]^2} \quad (11)$$

Systems with more than two sites are treated in an analogous manner. These larger systems introduce the complication of more than one possible pattern of cooperative interaction. Consider, e.g., three interacting sites. With all sites occupied, cooperativity might reflect any of several combinations of the strictly pairwise interactions (ΔG_{12} , ΔG_{13} , ΔG_{23}), or a unique interaction unrelated to these (ΔG_{123}). In general, it will not be possible to infer a unique pattern from binding data alone; the model will rest on extra-thermodynamic information or on a priori assumptions.

If there are multiple specific-binding sites, simultaneous numerical analysis of the data for all sites, according to the appropriate set of binding expressions, is necessary (be sure to include all transition endpoints). In principle, it should be possible to resolve all parameters in this manner; however, infinite numerical correlation between parameters usually precludes such resolution. It can be a challenge to even conclusively demonstrate cooperativity. To resolve the binding and cooperativity terms, it is usually also necessary to analyze binding data for reduced valency mutants (i.e., DNA fragments in which mutations in one or more sites eliminate site-specific binding to those sites). It is then necessary to demonstrate that alterations in the mutated sites have no intrinsic effect on the remaining sites. A complete description of the procedures necessary to analyze the interactions of protein ligands with multiple specific sites on DNA is beyond the scope of this article. Refer to Senear et al. (1986) for a complete discussion of the issues outlined above.

Anticipated Results

The individual site-binding curves resolved from *E. coli* Gal-repressor binding to its two-site wild-type operator—converted to fractional saturation by numerical determination of the endpoints—are shown in Figure 12.4.5. Note that the protein concentration range covers six orders of magnitude and includes a significant number of points at the upper and lower ends of the binding curve. Numerical analysis of this binding data with Equations 10 and 11 yields the *intrinsic* binding and cooperative free energies shown in Table 12.4.3. As discussed above, improved resolution of the cooperative free energies can be obtained by the simultaneous analysis of binding curves of protein and wild-type DNA (such as those shown in Fig. 12.4.4) and binding curves determined for protein binding to reduced valency DNA.

Time Considerations

Following preparation of the end-labeled DNA, the footprint titration assay will take ~6 to 8 hr to complete. The footprint assay and preparation of the samples for electrophoresis is labor intensive and can be performed in 3 to 4 hr. Electrophoresis usually takes ~2.5 hr, with fixing and drying the gels requiring the remaining time. It is convenient to cast the sequencing gels the day before conducting the assay.

Literature Cited

- Ackers, G.K., Johnson, A.D., and Shea, M.A. 1982. Quantitative model for gene regulation by lambda phage repressor. *Proc. Natl. Acad. Sci. U.S.A.* 79:1129-1133.
- Ackers, G.K., Shea, M.A., and Smith, F.R. 1983. Free energy coupling within macromolecules: The chemical work of ligand binding at the individual sites in cooperative systems. *J. Mol. Biol.* 170:223-242.
- Brenowitz, M., Senear, D.F., Shea, M.A., and Ackers, G.K. 1986a. Footprint titrations yield valid thermodynamic isotherms. *Proc. Natl. Acad. Sci. U.S.A.* 83:8462-8466.

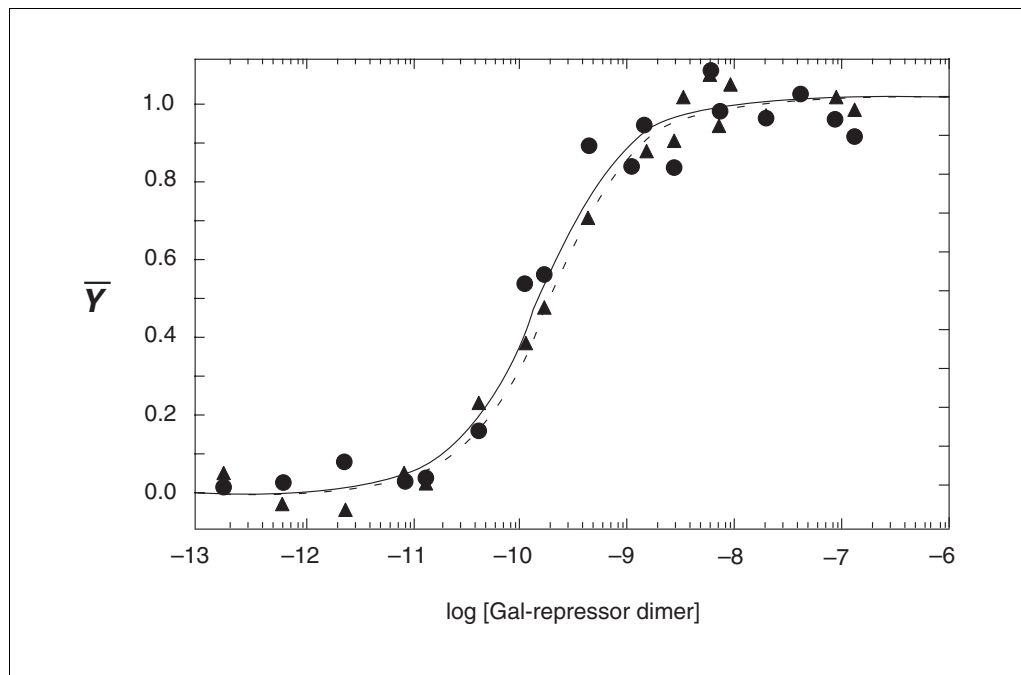


Figure 12.4.4 Individual site-binding curves determined for a footprint autoradiogram. See Equation 2 for a definition of \bar{Y} .

Brenowitz, M., Senear, D.F., Shea, M.A., and Ackers, G.K. 1986b. Quantitative DNase I footprint titration: A method for studying protein-DNA interactions. *Meth. Enzymol.* 130:132-181.

Dabrowiak, J.C. and Goodisman, J. 1989. Quantitative footprinting analysis of drug-DNA interactions. *In Chemistry and Physics of DNA-Ligand Interactions* (N.R. Kallenback, ed.). Adenine Press.

Galas, D. and Schmitz, A. 1978. DNase footprinting: A simple method for the detection of protein-DNA binding specificity. *Nucl. Acids Res.* 5:3157-3170.

Hertzberg, R.P. and Dervan, P.B. 1982. Cleavage of double helical DNA by (methidiumpropyl-EDTA) iron(II). *J. Am. Chem. Soc.* 104:313-315.

Johnson, A.D., Meyer, B.J., and Ptashne, M. 1979. Interactions between DNA-bound repressors govern regulation by the lambda phage repressor. *Proc. Natl. Acad. Sci. U.S.A.* 76:5061-5065.

Johnson, M.L. and Frasier, S.G. 1985. Nonlinear least-squares analysis. *Meth. Enzymol.* 117:301-342.

Senear, D.F., Brenowitz, M., and Ackers, G.K. 1986. Energetics of cooperative protein-DNA interactions: Comparison between quantitative DNase I footprint titration and filter binding. *Biochemistry* 25:7344-7354.

Tullius, T.D., Dombroski, B.A., Churchill, M.E.A., and Kam, L. 1987. Hydroxyl radical footprinting: A high resolution method for mapping protein-DNA contacts. *Meth. Enzymol.* 155:537-558.

Contributed by Michael Brenowitz
Albert Einstein College of Medicine
Bronx, New York

Donald F. Senear
University of California
Irvine, California

Robert E. Kingston (crude fractions)
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

UV Crosslinking of Proteins to Nucleic Acids

UNIT 12.5

Irradiation of protein-nucleic acid complexes with ultraviolet light causes covalent bonds to form between the nucleic acid and proteins that are in close contact with the nucleic acid. Thus, UV crosslinking may be used to selectively label DNA-binding proteins based on their specific interaction with a DNA recognition site. As a consequence of label transfer, the molecular weight of a DNA-binding protein in a crude mixture can be rapidly and reliably determined.

The procedure can be divided into 3 stages: (1) Extract containing the protein of interest is incubated with a radioactive, uniformly labeled DNA fragment that contains a high-affinity binding site for the protein; (2) protein-DNA complexes are crosslinked with UV irradiation and digested with nuclease, leaving only those labeled DNA fragments that are crosslinked and in close contact with the DNA-binding protein; and (3) the molecular weights of the crosslinked proteins are determined by SDS-polyacrylamide gel electrophoresis followed by autoradiography. One of the advantages of this technique is that proteins bound specifically to the DNA probe can be easily distinguished from those bound nonspecifically.

UV CROSSLINKING USING A BROMODEOXYURIDINE-SUBSTITUTED PROBE

**BASIC
PROTOCOL**

DNA molecules containing halogenated analogs of thymidine, such as bromodeoxyuridine (BrdU), are considerably more sensitive to UV-induced crosslinking compared to unsubstituted DNA. Although use of BrdU-substituted probes is not essential for detecting protein-DNA crosslinking by UV light, in many cases it is helpful. Of course, BrdU cannot be incorporated into RNA. For crosslinking to RNA, it is most convenient to generate labeled SP6 transcripts (UNIT 4.7) and crosslink as described in the alternate protocol of this unit.

Materials

- Single-stranded M13 vector with desired binding site (UNIT 1.15)
- 17-bp M13 universal primer (Pharmacia)
- 1× and 10× restriction endonuclease buffer (50 mM NaCl and 500 mM NaCl, respectively; UNIT 3.1)
- 3000 Ci/mmol [α -³²P]dCTP (UNIT 3.4)
- 50× dNTP/BrdU solution
- 0.1 M dithiothreitol (DTT)
- 25 U Klenow fragment of *E. coli* DNA polymerase I (UNIT 3.5)
- Ammonium acetate
- 100% ethanol
- TE buffer (APPENDIX 2)
- Buffered extract containing DNA-binding protein
- Bulk carrier DNA, e.g., poly(dI-dC)·poly(dI-dC)
- 0.5 M CaCl₂
- DNase I (Worthington)
- 1 U micrococcal nuclease (Worthington)
- 2× SDS/sample buffer (UNIT 10.2)
- Fluor (Du Pont NEN Enhance or equivalent)
- ¹⁴C-labeled protein markers (UNIT 10.2)

- DEAE membrane (Schleicher & Schuell NA45)
- 1.5-ml round-bottom screw-cap vial (Nunc or equivalent)

**DNA-Protein
Interactions**

12.5.1

Contributed by Lewis A. Chodosh

Current Protocols in Molecular Biology (1996) 12.5.1-12.5.8

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 36

16°, 37°, 68°, 90°, and 100°C water baths
UV transilluminator (305 nm, 7000 μW/cm²; Fotodyne)

Additional reagents and equipment for digesting with restriction nucleases (UNIT 3.1), labeling with Klenow fragment (UNIT 3.5), agarose gel electrophoresis (UNIT 2.5), mobility shift DNA-binding assay (UNIT 12.2), ethidium bromide dot quantitation (UNIT 2.6), SDS-polyacrylamide gel electrophoresis (UNITS 10.2 and 10.3), and autoradiography (APPENDIX 3)

Prepare the BrdU-substituted probe

1. Add 10 μg of single-stranded M13 vector containing a high-affinity binding site for the protein of interest, and an equimolar amount of the 17-bp M13 universal primer. Adjust final volume to 100 μl in 1× restriction buffer (50 mM NaCl). Heat 5 min at 90°C. Cool overnight at room temperature.

The M13 plasmid must contain a high-affinity protein-binding site. If a binding site has not been cloned into M13, an alternative is to anneal a 50-bp synthetic oligonucleotide (containing a high-affinity binding site for the protein of interest) to a complementary 15-bp synthetic oligonucleotide. The complementary region must be at the 3' end of the 50-bp oligonucleotide. This hybridized complex may be used as substrate for Klenow fragment. BrdU increases the efficiency of photochemical crosslinking. It also permits the use of longer UV wavelengths that are less damaging to protein. Moreover, in certain instances DNA-binding proteins recognize sites substituted with BrdU with a higher affinity than unsubstituted DNA.

2. Add the following to the hybridized mixture:

50 μl [α -³²P]dCTP (3000 Ci/mmol)
3.5 μl 50× dNTP/BrdU solution
1.75 μl 0.1 M DTT
7.5 μl 10× restriction buffer (50 mM NaCl, final concentration)
7 μl H₂O
5 μl (25 U) Klenow fragment.

Incubate 90 min at 16°C.

The radioactive dNTP should be chosen on the basis of the DNA sequence at the protein-binding site. Approximately 10 to 20 bp will remain in the immediate vicinity of the binding site following crosslinking and DNase digestion. The nucleotide appearing most frequently in the binding site (except for thymidine) should be chosen as the radioactive nucleotide. Similarly, choice of which strand to label should be based upon the nucleotide sequence of the binding site. If one strand of a binding site is highly rich in thymidine and cytidine residues, that strand should be labeled with BrdU and [α -³²P]dCTP. Of course, the unlabeled dNTP solution must be appropriately modified.

3. Heat-inactivate the Klenow fragment 10 min at 68°C.
4. Add 40 U restriction endonuclease(s) to generate a DNA fragment between 20 and 600 bp. Digest under appropriate conditions.

If protein-DNA complexes are not digested with nuclease following crosslinking, a small DNA probe (50 bp) should be used to minimize the effect of the crosslinked DNA on the mobility of the protein during SDS-polyacrylamide gel electrophoresis.

5. Add ammonium acetate to 0.3 M and precipitate with 2 vol of 100% ethanol. Resuspend in TE buffer.
6. Load onto an agarose gel containing 0.5 μg/ml ethidium bromide. Electrophorese and isolate the desired fragment using DEAE membrane as described in UNIT 2.7.
7. Determine the specific activity of the BrdU-substituted fragment using a scintillation counter and estimate the DNA concentration by ethidium bromide dot quantitation (UNIT 2.6).

Specific activities of 10^8 cpm/ μ g can easily be generated. Since high-specific-activity probes decay relatively rapidly, the probe should be used within 3 to 6 days. This probe can be used in the mobility shift DNA-binding assay (UNIT 12.2). Thus, the integrity and function of the probe can be tested.

Bind the protein to DNA

8. Set up the following binding reaction:

- 10⁵ cpm of uniformly labeled probe
- Buffered extract containing the binding protein
- 10 to 20 μ g of DNA carrier such as poly (dI-dC) \cdot poly(dI-dC).

Adjust volume to 50 μ l in a 1.5-ml round-bottom vial. Seal with plastic wrap held in place with a strip of Parafilm.

Determine conditions required to observe specific protein-DNA binding using the mobility shift assay (UNIT 12.2).

Crosslink the binding reaction with UV light and digest the unprotected DNA

9. Set the vial in a test-tube rack. Irradiate from 5 cm directly above by inverting a UV transilluminator of 305 nm and intensity 7000 μ W/cm² (see Fig. 12.5.1). Irradiate 5 to 60 min.

If a lower intensity UV source is used or if the distance is greater than 5 cm, time of exposure should be increased. Plastic wrap prevents evaporation of the sample during irradiation (which increases the temperature of the sample), but is UV transparent. The effect of UV dose on the protein extract and the probe alone can be monitored by the mobility shift DNA-binding assay (UNIT 12.2).

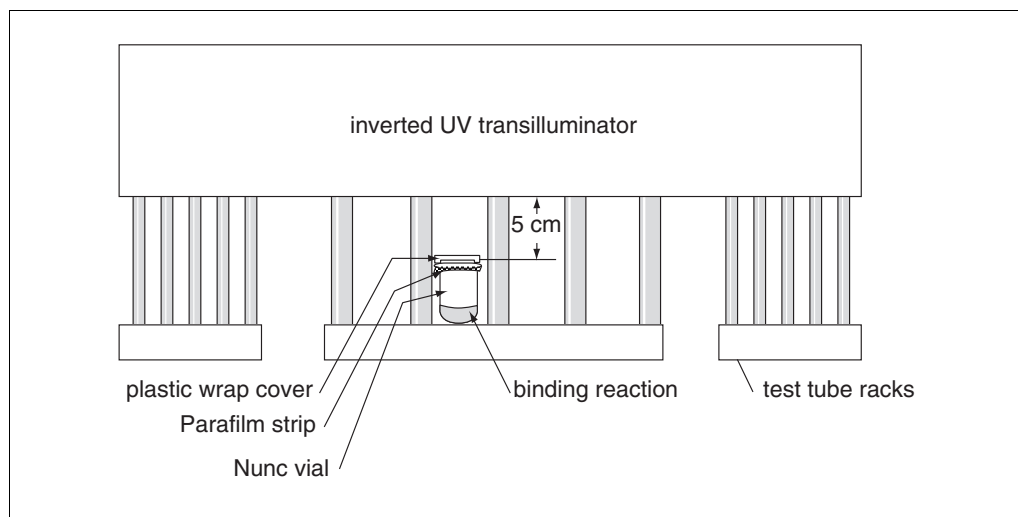


Figure 12.5.1 Experimental setup for UV crosslinking DNA-binding proteins to uniformly labeled DNA.

10. Add the following to each binding reaction:

- 1 μ l 0.5 M CaCl₂
- 4 μ g DNase I
- 1 U micrococcal nuclease.

Digest 30 min at 37°C.

Nuclease digestion generally spares the 10 to 20 bp protected by the crosslinked protein.

Electrophoresis and autoradiograph the crosslinked protein

11. Add an equal volume of 2× SDS/sample buffer to the binding reaction. Boil 5 min at 100°C.
12. Electrophorese sample through a discontinuous SDS-polyacrylamide gel of appropriate percentage (UNIT 10.2). Include a lane of radioactive protein markers such as ¹⁴C-labeled markers. To efficiently visualize ¹⁴C-labeled markers, the protein gel must be impregnated with a fluor such as Enhance before drying and autoradiography.
13. After running the gel, cut away the region that migrates with the dye front.

The bulk of digested nucleic acid products runs with the dye front and is intensely radioactive. If these products are not cut away, they can obscure signals from crosslinked proteins migrating toward the bottom of the gel.
14. Dry the gel and autoradiograph with an intensifying screen. Expose 1 to 3 days to visualize crosslinked proteins.

ALTERNATE PROTOCOL

UV CROSSLINKING USING A NON-BROMODEOXYURIDINE-SUBSTITUTED PROBE

Incorporation of BrdU into probes may have no effect upon the efficiency of UV crosslinking of proteins to sequences that do not contain thymidine residues. In addition, BrdU cannot be used to UV-crosslink proteins to RNA. In these instances, crosslinking may be performed with probes that do not contain BrdU. The procedure is similar to the basic protocol except for the steps noted below.

Additional Materials

50× dNTP/TTP solution
UV transilluminator (254 nm)

1. To prepare the uniformly labeled probe (step 2 of basic protocol), replace 50× dNTP/BrdU solution with 50× dNTP/TTP solution.

If crosslinking to RNA is desired, construct the template using SP6 RNA polymerase (UNIT 4.7).

2. The wavelength for crosslinking should be 254 nm instead of 305 nm (step 9 of basic protocol).
3. Irradiation time for crosslinking the protein may be varied from 5 min to 3 hr.

ALTERNATE PROTOCOL

UV CROSSLINKING IN SITU

A protocol has recently been described (Wu et al., 1987) in which UV crosslinking is performed as the final step in a mobility shift DNA-binding assay (UNIT 12.2). The most obvious advantage is that when multiple proteins bind to the probe, proteins present in individual protein–nucleic acid complexes can be visualized. This protocol is similar to the basic protocol except that after the binding reaction the mixture is loaded onto a low gelling/melting temperature agarose gel, electrophoresed, and irradiated. Complexes are visualized by autoradiography, excised, and melted in SDS gel solution. Samples are then resolved by SDS-PAGE as in the basic protocol.

Additional Materials

Low gelling/melting temperature agarose (UNIT 2.6)
1× TBE buffer (APPENDIX 2)
SDS gel solution

UV Crosslinking of Proteins to Nucleic Acids

12.5.4

1. Since DNA is not digested following crosslinking, short DNA probes (50 bp) should be prepared. Use the M13 system with appropriate restriction endonucleases or a small oligonucleotide primer that is complementary to the 3' end of a larger synthetic oligonucleotide. Follow steps 1 to 7 of the basic protocol.
2. Set up a standard 50- μ l binding reaction (step 8 of the basic protocol or as in the mobility shift assay, *UNIT 12.2*, scaled up 5 \times).

*Unlabeled competing DNA fragments can be added to the binding reaction to help define specific complexes and to determine which UV-crosslinked protein bands are unique to any complex (see *UNIT 12.2* and commentary to this unit).*

3. Load the binding reaction onto a 1% low gelling/melting temperature agarose gel cast in a cold room using 1 \times TBE buffer. Electrophorese at 4 V/cm for 2 to 3 hr.

Steps 3 and 4 are performed in a cold room to prevent the gel from melting due to the heat from electrophoresis and from the UV transilluminator.

4. Place the gel on plastic wrap on the surface of a 305-nm UV transilluminator. Irradiate in the cold room 5 to 30 min.
5. Autoradiograph the gel for 1 to 3 hr in a refrigerator. Excise the regions corresponding to specific protein-DNA complexes.
6. Add 10 μ l of SDS gel solution to each 50 μ l of gel slice. Boil the samples for 2 min.
7. Load the warm liquid samples onto a discontinuous SDS-polyacrylamide gel. Electrophorese and autoradiograph as in the basic protocol.

REAGENTS AND SOLUTIONS

50 \times dNTP/BrdU solution

2.5 mM dATP
2.5 mM dGTP
250 μ M dCTP
2.5 mM BrdU (5-bromo-2'-deoxyuridine triphosphate; Sigma)

50 \times dNTP/TTP solution

2.5 mM dATP
2.5 mM dGTP
250 μ M dCTP
2.5 mM TTP

SDS gel solution

0.3 M Tris-Cl, pH 6.8
6% sodium dodecyl sulfate
15% glycerol
70 mM dithiothreitol
Store at -20°C

COMMENTARY

Background Information

Crosslinking proteins to nucleic acids with UV light is a simple method for rapidly and accurately determining the molecular weight of a DNA-binding protein in a crude extract. Moreover, the specificity of the photoadduct can be rigorously determined by measuring the ability of an excess of unlabeled competitor DNA to compete for binding sites on the protein.

The goal of the UV crosslinking method is to specifically transfer a radioactive label from a DNA-binding site to the binding protein. Irradiation of DNA with UV light produces purine and pyrimidine free radicals. If a protein molecule is in close proximity to the free radical, a covalent bond can be formed, crosslinking the protein to the DNA.

In solution, pyrimidines are approximately 10-fold more sensitive to photochemical alteration than purines. Several amino acids are known to form photoadducts with pyrimidine bases, including cysteine, serine, methionine, lysine, arginine, histidine, tryptophan, phenylalanine, and tyrosine. Cells that have incorporated halogenated analogs of thymine—such as bromodeoxyuridine (BrdU)—into their DNA are several times more sensitive to UV-induced crosslinking with protein than normal cells. This is because replacement of the thymidine methyl group with the bromine atom creates a molecule more susceptible to free radical formation in the presence of UV light. Because the bromo group is approximately the same van der Waals radius as a methyl group, several cellular enzymes will use thymidine and BrdU interchangeably. Thus, it is quite simple to generate BrdU-substituted DNA probes.

Another important reason for using BrdU is that the longer wavelength of UV light used to crosslink these probes is less damaging to proteins than a shorter wavelength. In addition, substitution of BrdU into a binding site sometimes increases the affinity of the protein-DNA interaction being studied.

Literature Review

UV crosslinking has been used to study the interactions of several DNA-binding proteins with their DNA recognition sites. In 1972, Markowitz showed that UV irradiation induced the formation of DNA polymerase–DNA complexes that were resistant to high salt, phenol, heat, and 0.1 M NaOH. He concluded that a covalent bond between DNA and protein had

been formed. Lin and Riggs (1974) showed that *lac* repressor could be specifically crosslinked by UV irradiation to BrdU-substituted *lac* operator DNA. Hillel and Wu (1978) used photochemical crosslinking to demonstrate that in nonspecific T7 DNA–*E. coli* RNA polymerase complexes, subunits σ , β , and β' were crosslinked to DNA. However, in specific T7 DNA–RNA polymerase complexes only subunits σ and β were crosslinked to DNA. Simpson (1979) used UV irradiation to map the molecular interaction between *E. coli* RNA polymerase and the *lacUV5* promoter substituted with BrdU. More recently, UV crosslinking has been used to identify eukaryotic sequence-specific DNA-binding proteins present in crude extracts. Chodosh et al. (1986) specifically crosslinked the adenovirus major late transcription factor to its binding site, confirming the molecular weight of the molecule. Wu et al. (1987) used UV light to specifically label the *Drosophila* heat-shock activator protein.

Critical Parameters

Many critical parameters of this assay can be easily optimized using the mobility shift DNA-binding procedure (UNIT 12.2). If specific binding to the DNA probe can be detected in the mobility shift assay it should be possible to crosslink the protein to the probe. Particular parameters which may be varied include time of irradiation, BrdU substitution, binding conditions (especially ionic strength), and choice of radioactive dNTP. Some protein-DNA complexes are particularly sensitive to DNase, and thus protocols using only a very limited DNase digestion or no digestion may be tried.

Anticipated Results

Figure 12.5.2 represents the autoradiogram of a dried gel containing the results of a typical UV crosslinking experiment. The figure legend describes the components of each lane.

When probe DNA was irradiated in the absence of protein (lane 2) or when UV irradiation was omitted (lane 3), no nuclease-resistant labeled species were generated. However, when complete binding-reaction mixtures were irradiated and analyzed, at least two labeled species were generated, migrating at ~45 kDa and 26 kDa. Increasing the time of irradiation from 15 to 60 min increased the amount of label in these species (lanes 5 to 7).

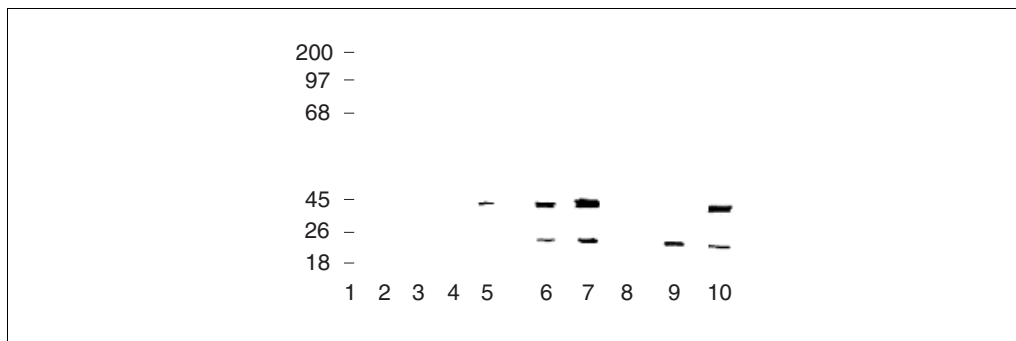


Figure 12.5.2 Representation of autoradiogram of a typical UV crosslinking experiment. As described in commentary, procedure involves UV crosslinking of protein to DNA probe, digestion of free DNA with DNase, and electrophoresis by SDS-PAGE of the treated binding reactions. Molecular weights of ^{14}C -labeled protein markers in Lane 1 are indicated to the left. Lanes 2 through 10 each contained a ^{14}C -labeled DNA probe with a site for a binding protein X. Samples applied to the lanes differed as follows (see also anticipated results): **Lane 2**—no protein; **Lane 3**—no UV irradiation; **Lane 4**—5 min irradiation of complete binding reaction; **Lane 5**—15 min irradiation of complete binding reaction; **Lane 6**—30 min irradiation of complete binding reaction; **Lane 7**—60 min irradiation of complete binding reaction; **Lane 8**—60 min irradiation of complete binding reaction then proteinase K treatment; **Lane 9**—60 min irradiation of complete binding reaction plus 100-fold molar excess of unlabeled probe containing a binding site for protein X; **Lane 10**—60 min irradiation of complete binding reaction plus 100-fold molar excess of unlabeled mutant probe lacking a binding site for protein X.

When UV-irradiated binding-reaction mixtures were treated with proteinase K, no labeled species were observed following electrophoresis—suggesting that the labeled species were protein-DNA photoadducts (lane 8). Control reactions demonstrated that probe DNA irradiated for 60 min was still able to bind protein X (data not shown). Similarly, UV irradiation of protein X alone did not significantly reduce its ability to bind probe DNA (data not shown).

To distinguish whether polypeptides were bound specifically or nonspecifically, UV crosslinking was performed with an excess of unlabeled competitor DNA. This DNA contained (lane 9) or lacked (lane 10) a high-affinity binding site for protein X. If a labeled polypeptide results from the covalent crosslinking of protein X to its recognition sequence, competitor fragments containing this sequence should abolish formation of the corresponding labeled DNA-protein adduct. Conversely, competitor fragments that lack this sequence should not affect formation of the labeled DNA-protein adduct. Because a competitor fragment with the protein X binding site abolished formation of the 45-kDa labeled polypeptide but did not affect formation of the 26-kDa labeled polypeptide (lane 9), the 45-kDa labeled polypeptide represents the covalent crosslinking of protein X to its recognition sequence. This is confirmed in lane 10 where a competitor fragment lacking the protein X

binding site did not affect formation of either labeled polypeptide.

Treatment of irradiated binding mixtures with either greater amounts of nuclease or the same amount of nuclease for longer times (data not shown) did not alter the mobility of the 45-kDa species, suggesting that this was a limit digest. Since the covalent attachment of short oligonucleotides to proteins has only a minor effect on the mobility of proteins in gels, these experiments indicate that a 45-kDa protein binds specifically to the protein X binding site.

An important point to consider is that the efficiency of UV crosslinking is usually on the order of ~0.1% to 10%. Therefore, it is exceedingly rare to see more than one crosslinking event in a single complex. If a protein binds to its target site as a dimer, UV crosslinking will typically only transfer label to one of the two bound molecules. Thus, the observed molecular weight is likely to be that of the monomer.

Time Considerations

The basic protocol can be performed in ~7 to 10 hr. However, most of this time is not labor intensive. Autoradiography takes between 12 hr and 3 days depending on the specific activity of the probe, the efficiency of crosslinking, and the abundance and binding affinity of the protein.

Literature Cited

- Chodosh, L.A., Carthew, R.W., and Sharp, P.A. 1986. A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. *Mol. Cell. Biol.* 6:4723-4733.
- Hillel, Z. and Wu, C.-W. 1978. Photochemical crosslinking studies on the interaction of *Escherichia coli* RNA polymerase with T7 DNA. *Biochemistry* 17:2954-2961.
- Lin, S.-Y. and Riggs, A.D. 1974. Photochemical attachment of *lac* repressor to bromodeoxyuridine-substituted *lac* operator by ultraviolet radiation. *Proc. Natl. Acad. Sci. U.S.A.* 71:947-951.
- Markowitz, A. 1972. Ultraviolet light-induced stable complexes of DNA and DNA polymerase. *Biochim. Biophys. Acta* 281:522-534.
- Simpson, R.B. 1979. The molecular topography of RNA polymerase-promoter interaction. *Cell* 18:277-285.
- Wu, C., Wilson, S., Walker, B., Dawid, I., Paisley, T., Zimarino, V., and Ueda, H. 1987. Purification and properties of *Drosophila* heat shock activator protein. *Science* 238:1247-1253.

Key References

- Chodosh et al., 1986. See above.
Describes UV crosslinking in crude mammalian extracts and demonstrates specificity of binding of the crosslinked proteins. Compares the technique with other methods for determining the molecular weight of DNA-binding proteins.
- Hillel and Wu, 1978. See above.
An excellent study of the differences in polymerase-DNA contacts in specific and nonspecific complexes.
- Lin and Riggs, 1974. See above.
A seminal paper describing the use of BrdU in photochemical crosslinking.

Contributed by Lewis A. Chodosh
Massachusetts Institute of Technology
and Harvard Medical School
Cambridge and Boston, Massachusetts

Purification of DNA-Binding Proteins Using Biotin/Streptavidin Affinity Systems

UNIT 12.6

BASIC
PROTOCOL

Short fragments of DNA—either natural or formed from oligonucleotides—containing a high-affinity site for a DNA-binding protein provide a powerful tool for purification. The biotin/streptavidin purification system is based on the tight and essentially irreversible complex that biotin forms with streptavidin. The experimental design of this system is illustrated in Figure 12.6.1. First, a DNA fragment is prepared that contains a high-affinity binding site for the protein of interest. A molecule of biotinylated nucleotide is incorporated into one of the ends of the DNA fragment. The protein of interest is allowed to bind to the high-affinity recognition site present in the biotinylated fragment. The tetrameric protein streptavidin is then bound to the biotinylated end of the DNA fragment. Next, the protein/biotinylated fragment/streptavidin ternary complex is efficiently removed by adsorption onto a biotin-containing resin. Since streptavidin is multivalent, it is able to serve as a bridge between the biotinylated DNA fragment and the biotin-containing resin. Proteins remaining in the supernatant are washed away under conditions that maximize the stability of the DNA-protein complex. Finally, the protein of interest is eluted from the resin with a high-salt buffer.

Materials

Plasmid DNA with desired binding site
Appropriate restriction endonucleases (UNIT 3.1)
Biotin-11-dUTP (BRL)

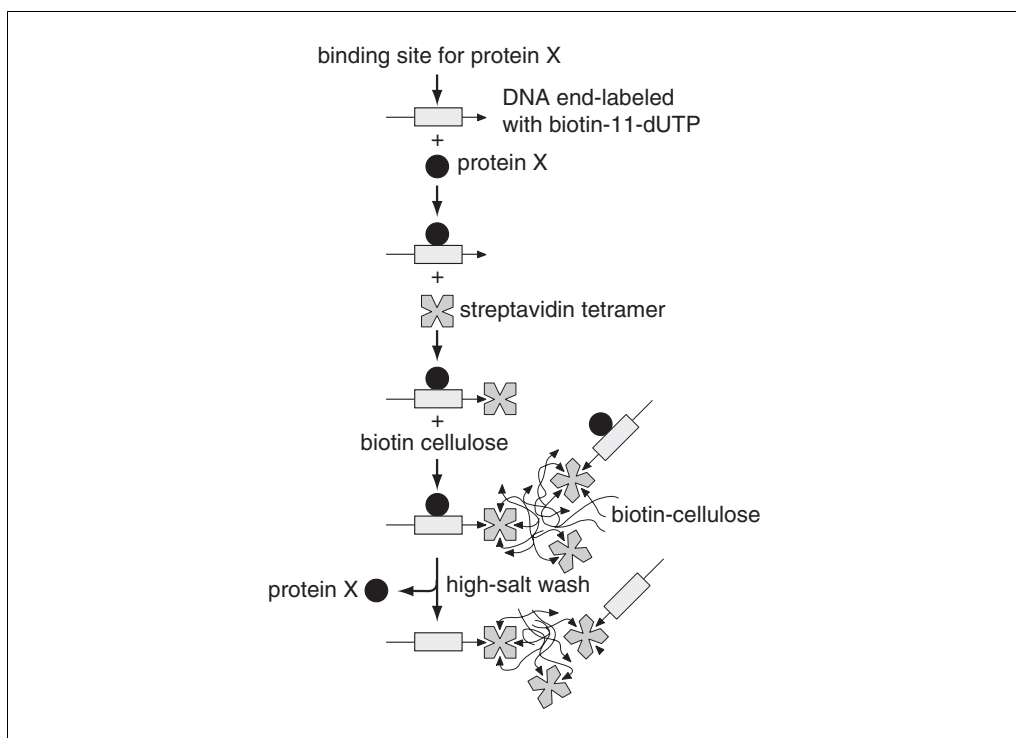


Figure 12.6.1 Purification of DNA-binding proteins using the biotin/streptavidin affinity technique. The protocol involves the following steps: (1) a biotinylated, labeled DNA fragment is prepared containing a binding site for the protein to be purified; (2) the biotin-cellulose resin is prepared; (3) a binding reaction containing a crude protein fraction and the biotinylated probe is set up; (4) free streptavidin is added to the binding reaction; (5) the protein/biotinylated DNA fragment/streptavidin complex is bound to the biotin-cellulose resin; (6) unbound protein is removed by extensive resin washing; and (7) the protein is eluted from the resin with high-ionic-strength buffer. Each of these steps can be monitored and optimized in solution, using the mobility shift DNA-binding assay (see UNIT 12.2).

Contributed by Lewis A. Chodosh

Current Protocols in Molecular Biology (1996) 12.6.1-12.6.9

Copyright © 2000 by John Wiley & Sons, Inc.

DNA-Protein
Interactions

12.6.1

Supplement 36

Labeled and unlabeled dNTPs (UNIT 3.4)
Klenow fragment of *E. coli* DNA polymerase I (UNIT 3.5)
TE buffer (APPENDIX 2)
Biotin-cellulose (Pierce)
Biotin-cellulose binding buffer
Bovine serum albumin (BSA)
Bulk carrier DNA [e.g., poly(dI-dC)·poly(dI-dC), salmon sperm DNA,
or *E. coli* DNA]
Biotin-cellulose elution buffer
Protein solution
Streptavidin (Celltech Ltd.; may be stored as 5 mg/ml stock for at least 2 months)
DEAE membrane (Schleicher & Schuell NA45)
0.025- μ m filter discs (Millipore VS)
Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5),
mobility-shift DNA-binding assay (UNIT 12.2), and ethidium bromide dot
quantitation (UNIT 2.6)

Prepare biotinylated, labeled DNA fragment

1. Digest 50 μ g of plasmid DNA in 100 μ l. The plasmid DNA must contain a binding site for the protein of interest and should be digested with one or more restriction endonucleases.

The same DNA probe used in the mobility shift DNA-binding assay (UNIT 12.2) can also be used to purify the protein.

2. Add the following to the probe mixture:

Biotin-11-dUTP to a final concentration of 20 μ M
Radioactive dNTP for incorporation into the 5' overhang
100-fold molar excess of corresponding unlabeled dNTP
Remaining two unlabeled dNTPs to a final concentration of 200 μ M
5 U Klenow fragment.

This reaction is identical to that used in the DNA-binding assay except biotinylated dUTP is incorporated into one end of the DNA fragment in place of TTP and the fragment is radiolabeled to low specific activity instead of to high specific activity.

3. Precipitate the biotinylated probe. Isolate the probe by agarose gel electrophoresis using DEAE membrane (UNIT 2.7).

Gel purification of the biotinylated fragment removes unreacted biotin-11-dUTP. This is essential because any free biotin-dUTP binds to streptavidin and then to the biotin-cellulose column, reducing the apparent capacity of both materials to react with protein.

4. Resuspend the probe in TE buffer and measure an aliquot for Cerenkov counts. Estimate the DNA concentration by ethidium bromide dot quantitation (UNIT 2.6).
5. Test the biotinylated probe to be certain it will efficiently bind to the protein of interest. Use a standard binding assay (UNIT 12.2) with the biotinylated fragment as probe.

Prepare biotin-cellulose resin

6. Place 200 μ l biotin-cellulose in a 1.5-ml microcentrifuge tube. Spin the resin for 30 sec in a microcentrifuge and remove the supernatant. Add to the pellet:

1.0 ml biotin-cellulose binding buffer
500 μ g/ml BSA
200 μ g carrier DNA.

Gently mix the tube 5 min on a rotating wheel.

This step blocks nonspecific protein and nucleic acid binding sites present on the biotin-cellulose resin.

7. Spin the resin, remove the supernatant, and resuspend pellet in 1.0 ml of biotin-cellulose elution buffer. Gently mix 5 min on a rotating wheel. Repeat this wash.

Washing removes molecules on the biotin-cellulose resin that might later be eluted from the resin by the biotin-cellulose elution buffer.

8. Spin the resin, remove the supernatant, and resuspend pellet in 1.0 ml of biotin-cellulose binding buffer. Repeat this wash.

This 1:6 dilution of pretreated biotin-cellulose is ready for use and can be stored for several months.

Set up binding reaction

9. Determine the molar concentration of the protein to be purified using the mobility-shift assay (UNIT 12.2).
10. Set up a standard binding reaction containing the protein to be purified, carrier DNA, and a 10-fold molar excess of biotinylated fragment relative to the protein to be purified. Allow the reaction to go to completion for ~15 min.

Use reaction conditions that optimize protein binding to its recognition site. The composition of the biotin-cellulose binding buffer should be the same as that optimized for protein-DNA binding.

11. Add a 5-fold molar excess of streptavidin relative to the biotinylated fragment. Continue the binding reaction for an additional 5 min at 30°C.

Bind protein/DNA/streptavidin complex to biotin-cellulose resin

12. In a separate tube, place 2 μ l pretreated biotin-cellulose (12 μ l of the 1:6 dilution) for each picomole of biotinylated DNA fragment in the binding reaction. Spin the resin and remove the supernatant.

One or two microliters of biotin-cellulose can easily be seen at the bottom of the microcentrifuge tube.

13. Transfer the binding reaction mix into the tube with the biotin-cellulose resin using a pipettor. Gently resuspend the resin and incubate on a rotating wheel for 30 min.

This incubation can be done either at 4°C or at room temperature, depending on the stability of the protein.

14. Spin the resin and remove the supernatant.

Using the mobility shift assay, the supernatant should be measured to determine what percentage of the biotinylated fragment and the protein has been removed from the supernatant. It is also useful to assay the supernatant from the binding reaction for a control DNA-binding protein to determine whether the protein of interest has been specifically depleted or whether multiple DNA-binding proteins have been depleted nonspecifically. The latter observation would suggest that the matrix is acting as a nonspecific DNA-affinity column. For some applications, a protein fraction specifically depleted for a particular DNA-binding protein is a valuable reagent.

Wash the resin

15. Resuspend the biotin-cellulose pellet in 500 μ l biotin-cellulose binding buffer. Mix by gently inverting the tube 1 to 2 min. Spin the resin and remove the supernatant. Repeat this procedure twice. The second time, transfer to a clean microcentrifuge tube.

Transferring the reaction avoids elution of proteins that were bound nonspecifically to the walls of the tube in the first binding incubation.

Elute the protein

16. Resuspend the biotin-cellulose pellet in at least an equal volume of biotin-cellulose elution buffer. Mix gently on a rotating wheel for 20 min.

The salt concentration in the biotin-cellulose elution buffer must be determined empirically. Successively higher salt concentrations may be tested until the concentration of eluted protein is maximal.

17. Spin the resin. Save the supernatant and assay for binding activity.

Small volumes of protein solutions can be dialyzed effectively on 0.025- μ m filter discs.

ALTERNATE PROTOCOL

PURIFICATION USING A MICROCOLUMN

Although the batch method in the basic protocol is rapid and well-suited for analytical-scale purification, larger volumes of biotin-cellulose resin can be better handled in a microcolumn. This method is also used to elute the protein in as small a volume (i.e., as high a concentration) as possible.

Additional Materials

Siliconized glass wool (APPENDIX 3)
1.0-ml pipet tip
Ring stand

1. Prepare the biotinylated DNA fragment and biotin-cellulose resin, and set up binding reaction, as described in steps 1 to 11 of basic protocol.
2. Place a small plug of siliconized glass wool in the bottom of a 1.0-ml pipet tip. Firmly attach the pipet-tip microcolumn to a ring stand.

Prewet the glass wool in biotin-cellulose binding buffer before insertion into the pipet tip to avoid trapping air bubbles which might denature proteins.

3. Add 500 μ l binding buffer to the microcolumn. Maintain a steady flow through the glass wool plug.

If the column does not flow smoothly, a pipettor can be gently inserted into the top of the microcolumn. Slightly depressing the plunger will start the column or increase the flow.

4. Add at least 40 μ l of 1:1 biotin-cellulose slurry to the microcolumn. Allow the buffer to run down to the surface of the resin.
5. Equilibrate the resin with 3 column-volumes of biotin-cellulose binding buffer if the resin has already been pretreated (steps 6 to 8 of basic protocol). If the resin has not been pretreated, wash sequentially with 3 column-volumes each of biotin-cellulose binding buffer, biotin-cellulose binding buffer with 500 μ g/ml BSA and 200 μ g/ml poly(dI-dC)-poly(dI-dC), and biotin-cellulose elution buffer. Finally, equilibrate with biotin-cellulose binding buffer.

The biotin-cellulose microcolumn can be washed very rapidly. Each wash takes 2 or 3 min. The drop size from the pipet tip is ~25 μ l but this will change with alterations in the ionic strength and protein concentration of the eluate.

6. Load the binding reaction mix (step 11 of basic protocol) onto the microcolumn and collect the flowthrough.

The column can be run as fast as 6 to 10 column-volumes/hr without affecting the amount of biotinylated fragment bound by the resin. If the flow rate is too slow, use a pipettor to apply pressure to the column. If the flow rate is too fast, plug the tip of the microcolumn with Parafilm in between drops.

7. Wash with 4 column-volumes of biotin-cellulose binding buffer. Discard the flow-through.
8. Wash with 3 column-volumes of biotin-cellulose elution buffer. Collect 2-drop fractions and assay.

Fractions as small as 5 μ l may be rapidly and effectively dialyzed on 0.025- μ m filter discs with minimal loss of volume and activity. Float a filter (shiny side up) in a petri dish on top of 20 ml dialysis buffer. Allow the filter 10 min to wet. Place the sample (5 to 100 μ l) to be dialyzed onto the surface of the filter. Surface tension will keep the sample confined in a drop unless there is a detergent in the sample. After 1 hr remove the sample. Once the protein has been eluted from the resin, the resin is effectively the same as a sequence-specific DNA-affinity microcolumn.

PURIFICATION USING STREPTAVIDIN-AGAROSE

When high-quality free streptavidin is not available or cellulose is an inappropriate resin, a simple variation on the basic protocol may be employed. In this protocol, the same biotinylated DNA fragment is used but is removed from solution directly by streptavidin-agarose (see Fig. 12.6.2).

Additional Materials

Streptavidin-agarose

1. Prepare biotinylated DNA fragment and resin (substituting streptavidin-agarose for biotin-cellulose), and set up binding reaction, as described in steps 1 to 11 of basic protocol.

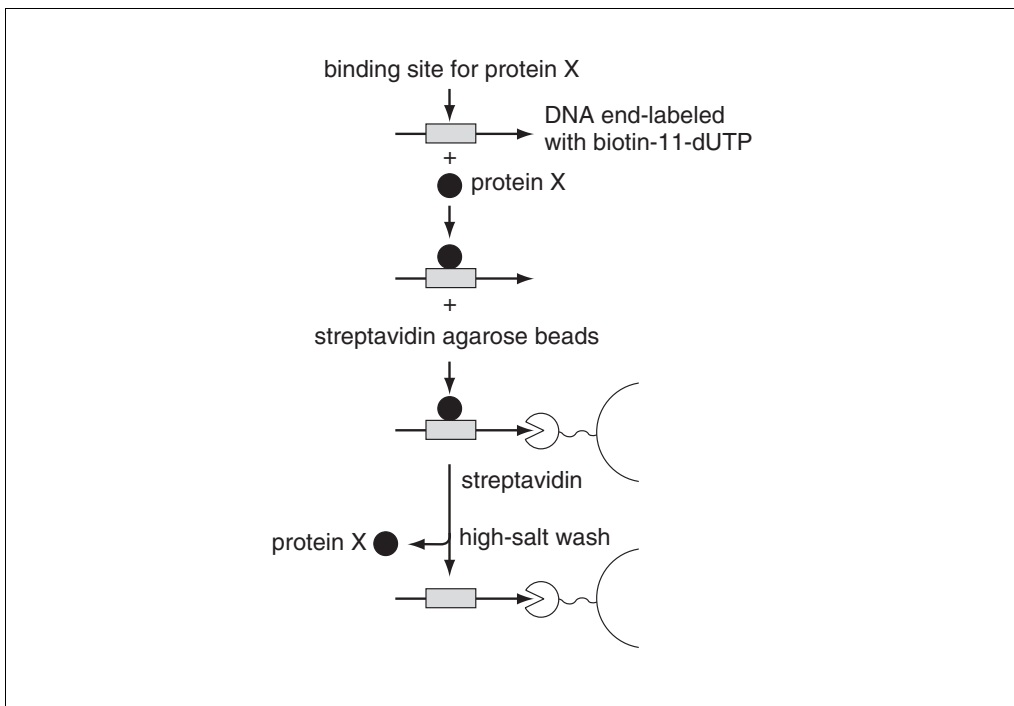


Figure 12.6.2 Purification of DNA-binding proteins using streptavidin-agarose.

**ALTERNATE
PROTOCOL**

**DNA-Protein
Interactions**

12.6.5

2. In a separate tube, add 50 μl pretreated streptavidin-agarose (300 μl of the 1:6 dilution) for each picomole of biotinylated DNA fragment in the binding reaction. Spin the resin and remove the supernatant.
3. Transfer the binding-reaction mix into the tube with the streptavidin-agarose using a pipettor. Gently resuspend the resin and incubate on a rotating wheel for 30 min to 2 hr.
4. Wash and elute the protein as described in steps 14 to 17 of the basic protocol.

Like biotin-cellulose, streptavidin-agarose may also be used in a microcolumn. Follow the microcolumn alternate protocol.

REAGENTS AND SOLUTIONS

Biotin-cellulose binding buffer

12% glycerol
 12 mM HEPES-NaOH, pH 7.9
 4 mM Tris-Cl, pH 7.9
 60 mM KCl
 1 mM EDTA
 1 mM dithiothreitol (DTT)
 Store at -20°C upon addition of DTT

This is a typical binding buffer. The composition of binding buffer—especially with respect to pH, ionic strength, and the presence or absence of MgCl_2 —should be determined by those conditions which optimize the binding of the protein of interest to its recognition site.

Biotin-cellulose elution buffer

12% glycerol
 20 mM Tris-Cl, pH 6.8
 1 M KCl
 5 mM MgCl_2
 1 mM EDTA
 1 mM DTT
 200 $\mu\text{g}/\text{ml}$ bovine serum albumin (BSA)
 Store at -20°C upon addition of DTT

This is a typical elution buffer. The composition of elution buffer—especially with respect to pH, ionic strength, and the presence or absence of MgCl_2 —should be determined by those conditions that maximize the dissociation rate of the protein from its recognition site. Another carrier protein, such as insulin or hemoglobin, may be substituted for BSA. The small size of insulin can be useful if protein gels will be run to determine the size of the regulatory factor.

COMMENTARY

Background Information and Literature Review

The biotin/streptavidin purification method works well because the interaction between biotin and avidin (or avidin-like proteins) is one of the strongest known noncovalent interactions. The dissociation constant for the streptavidin-biotin complex is $\sim 10^{-15}$ M. Avidin, from egg white, and streptavidin, from *Streptomyces avidinii*, are tetrameric proteins

containing four high-affinity binding sites for the vitamin biotin. Since streptavidin is multivalent, it is able to serve as a bridge between the biotinylated DNA fragment and the biotin-containing resin. The strong interaction is extremely useful for purification of DNA-binding proteins, because DNA-affinity columns with streptavidin/biotin bridges can be washed under a wide variety of conditions (i.e., 2 M KCl and 1% SDS) without remov-

ing either the streptavidin or the biotinylated DNA fragment from the matrix.

Two properties of streptavidin make it more suitable than avidin for use in DNA-affinity purification. The first is that streptavidin, unlike avidin, is not a glycoprotein, and the second is that streptavidin is slightly acidic whereas avidin is basic. Therefore, streptavidin is less likely to bind nonspecifically to cellular glycoproteins and to acidically charged cell components such as nucleic acids.

Several factors make this method simple, rapid, and effective. The same binding conditions and DNA fragment used in the mobility-shift DNA-binding assay to identify a protein (UNIT 12.2) can be used to effect its purification. In addition, the binding of the protein to its DNA recognition site in solution is more efficient than protein-DNA interactions that take place on a column matrix (discussed below). Most importantly, binding in solution allows each reaction parameter to be optimized on an analytical scale by using the gel binding assay.

As an analytical technique, biotin/streptavidin DNA-affinity purification permits the direct identification of a wide variety of sequence-specific DNA-binding proteins. It has already been successfully used to identify hormone receptors, components in mRNA splicing complexes, and RNA polymerase II and III transcription factors (Haeuptle et al., 1983; Grabowski and Sharp, 1986; Chodosh et al., 1986; Kasher et al., 1986).

Another method is commonly used for purifying DNA-binding proteins, whereby DNA-protein binding interactions occur on a column matrix. In this technique, catenated DNA-binding sites are covalently coupled with cyanogen bromide to Sepharose CL-2B (Kadonaga and Tjian, 1986; Rosenfeld and Kelly, 1986). A partially purified protein fraction is combined with competitor DNA and passed through a DNA-Sepharose resin, binding the protein to the surface of the matrix. Generally, multiple passes through the column are used to effect purification, with 25- to 50-fold purification achieved with each column passage. The biotin/streptavidin technique, on the other hand, typically yields 50- to 250-fold purification.

Critical Parameters

Only one end of the DNA fragment should be labeled with biotin. If both ends are labeled, the biotinylated fragment may bind to the resin in such a fashion as to sterically displace the protein bound to the DNA fragment. Thus, one

of the restriction enzymes used must be chosen to permit the incorporation of biotin-11-dUTP in place of TTP. Furthermore, since only one end of the fragment is biotin-labeled, the restriction enzyme generating the nonbiotinylated end should leave no overhang, a 3' overhang, or a 5' overhang that will not incorporate biotin-dUTP. For binding sites cloned into the pUC polylinker, an *EcoRI* and *HincII* fragment fulfills these requirements. Biotin-11-dUTP may be incorporated into the *EcoRI* overhang along with a low-specific-activity dATP radioactive label. The blunt *HincII* end of the fragment remains unlabeled.

Alternately, if both ends have 5' overhangs which permit incorporation of biotin-dUTP, the plasmid may first be digested with one of the enzymes and end-labeled with biotin-dUTP and a radioactive dNTP using the Klenow fragment. After heat-inactivating the Klenow fragment, the plasmid may then be digested with the second restriction endonuclease.

The low-specific-activity label incorporated into the biotinylated DNA fragment is used to help monitor both the binding of the protein to the biotinylated fragment and the binding of the biotinylated fragment to the biotin-cellulose resin via a streptavidin bridge. The amounts of the radioactive dNTP and the nonradioactive dNTP added to the Klenow reaction should be in a ratio such that 1% to 2% of the biotinylated fragments are labeled. For example, 100 μCi of 6000 Ci/mmol dATP (0.2 μM) may be added to a reaction containing 20 μM unlabeled dATP.

If the radioactive dNTP and the biotin-dUTP residue are incorporated on the same end of the DNA fragment, as in an *EcoRI* site, decay of the radioactive nuclide will result in formation of an unlabeled, nonbiotinylated probe that binds to the protein but is incapable of binding to streptavidin. This molecule is a competitive inhibitor during protein purification and can be kept to a minimum by labeling to a lower specific activity.

Over time, streptavidin tetramers dissociate into monomers. The integrity of the tetramer may be monitored by performing native (non-denaturing) gel electrophoresis. Since streptavidin monomers will bind to the biotinylated DNA fragment, but will be unable to bind to the biotin-cellulose resin, the monomeric form of streptavidin will inhibit purification.

Most proteins can be eluted from their binding site with 1 M KCl. Since the added elution buffer is diluted by the binding buffer in the biotin-cellulose pellet, the actual salt concentration in the resuspended pellet will be lower

than that of the elution buffer. As a result, it is useful to remove a small aliquot, dilute it, and measure the actual salt concentration with a conductivity meter. Otherwise, the time, temperature, and composition of the buffer used in the elution step should be chosen to maximize the protein's dissociation rate from its binding site. Thus, if a protein's dissociation rate is 15 sec in 1 M KCl at room temperature, and 3 hr in 1 M KCl at 4°C, a 5-min incubation in 1 M KCl at room temperature should be sufficient to elute the protein. Dissociation rates can easily be measured using the mobility shift DNA-binding assay described in this chapter (UNIT 12.2).

Troubleshooting

A major advantage of the biotin/streptavidin method is the ability to monitor the purification on an analytical scale using the mobility-shift assay (see UNIT 12.2). Because different parameters can be independently examined, it is possible to pinpoint the particular step that is not working. Several parameters can be examined routinely. Using the biotinylated radiolabeled fragment, the specific binding of the DNA fragment to its protein can be directly determined. In addition, the binding conditions can be optimized in a binding reaction containing the DNA probe and protein. The probe can be a biotinylated or a nonbiotinylated radiolabeled DNA fragment. The complex is then resolved by native (nondenaturing) gel electrophoresis.

Streptavidin binding to the probe can be monitored and titrated by varying the amount of streptavidin in a binding reaction. The resulting streptavidin-DNA complex is resolved from free DNA by native gel electrophoresis.

Binding of this complex to the biotin-cellulose matrix can be directly monitored. First, biotin-cellulose is incubated with the complex. Next the resin is spun and aliquots of the supernatant are assayed for the presence of the radiolabeled fragment on a polyacrylamide gel. This basic procedure allows several relevant parameters to be examined, including the amount of streptavidin required for complete binding of the fragment to the resin, the amount of biotin-cellulose required to bind a given amount of fragment, the binding time, and the optimal binding conditions. These three final steps can also be monitored for the specific adsorption of the protein to the column. By assaying the amount of binding activity in the supernatant before and after addition of biotin

cellulose, the adsorption efficiency can be determined.

Nonspecific resin binding can be examined by assaying the binding activity in a supernatant containing the protein, the DNA carrier poly(dI-dC)-poly(dI-dC), streptavidin, and biotin-cellulose, but lacking the biotinylated DNA fragment with the protein's binding site. Alternatively, a nonspecific biotinylated DNA fragment can be added. Interestingly, nonspecific protein binding to the resin is minimized with a lower poly(dI-dC)-poly(dI-dC) concentration than is required to generate a discrete band in the mobility shift assay.

The conditions for protein elution from the column can also be directly examined and optimized. After protein adsorption onto the matrix, elution parameters such as time, temperature, pH, carrier protein, and ionic strength can be varied. The amount of protein released into the supernatant is then assayed by the mobility shift assay. In optimizing this step, it is very helpful to know as much as possible about the variables affecting the dissociation rate of the protein from its DNA binding site. These can be studied quickly and rigorously with the mobility shift assay.

Anticipated Results

The strong interaction between biotin and avidin can result in 25 nmol of biotinylated DNA fragment binding to 1 ml of biotin-cellulose resin. This corresponds to 1.5 mg of a 60 kDa polypeptide. After parameter optimization, ~50- to 250-fold purification of the DNA-protein complex can be obtained. Once the protein has been adsorbed to the column >99% of the remaining protein can be removed with repeated column washing using the binding buffer. Moreover, ~30% to 80% of the binding activity can be recovered from the column. The recovery efficiency depends on how effectively nonspecific binding sites have been blocked. This purification scheme may be used either with crude extracts or with partially purified protein fractions.

Time Considerations

Starting with a biotinylated DNA fragment, the batch (basic) protocol can be performed in ~1.5 hr, and the microcolumn procedure in ~2.5 hr.

Literature Cited

- Chodosh, L.A., Carthew, R.W., and Sharp, P.A. 1986. A single polypeptide possesses the binding and transcription activities of the adenovirus major late promoter. *Mol. Cell. Biol.* 6:4723-4733.
- Grabowski, P.J. and Sharp, P.A. 1986. Affinity chromatography of splicing complexes: U2, U5, and U4 + U6 small nuclear ribonucleoprotein particles in the spliceosome. *Science* 233:1294-1299.
- Hauptle, M.-T., Aubert, M.L., Kjiene, J., and Kraehenbuhl, J.-P. 1983. Binding sites for lactogenic and somatogenic hormones from rabbit mammary gland and liver. *J. Biol. Chem.* 258:305-314.
- Kadonaga, J.T. and Tjian, R. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* 83:5889-5893.
- Kasher, M.S., Pintel, D., and Ward, D.C. 1986. Rapid enrichment of HeLa transcription factors IIIB and IIIC by using affinity chromatography

based on avidin-biotin interactions. *Mol. Cell. Biol.* 6:3117-3127.

- Rosenfeld, P.J. and Kelly, T.J. 1986. Purification of nuclear factor I by DNA recognition site affinity chromatography. *J. Biol. Chem.* 261:1398-1408.

Key References

Chodosh et al., 1986. See above.

Describes the biotin/streptavidin-agarose microcolumn procedure from which the second alternate protocol of this unit was drawn.

Kasher et al., 1986. See above.

Describes a biotin-cellulose/streptavidin procedure similar to the basic protocol of this unit.

Lewis A. Chodosh
Massachusetts Institute of Technology
and Harvard Medical School
Cambridge and Boston, Massachusetts

Detection, Purification, and Characterization of cDNA Clones Encoding DNA-Binding Proteins

In this unit, an appropriate recombinant clone is detected in an expression library with a DNA recognition-site probe, purified, and shown to encode a DNA-binding domain of defined sequence specificity. The strategy described below obviates purification of a sequence-specific DNA-binding protein for the purpose of isolating its gene; it simply requires a suitable cDNA library constructed in the expression vector λ gt11 and a DNA recognition-site probe. The basic protocol enables the detection and purification of clones encoding sequence-specific DNA-binding proteins. The alternate protocol describes a denaturation/renaturation procedure that can increase detection of certain clones. The support protocol provides a rapid means of characterizing the DNA-binding activities of the proteins encoded by the cloned cDNAs.

BASIC PROTOCOL

SCREENING A λ gt11 EXPRESSION LIBRARY WITH RECOGNITION-SITE DNA

A clone encoding a sequence-specific protein is detected in a λ gt11 library because its recombinant protein binds specifically to a radiolabeled recognition-site DNA. Bacteriophage from a cDNA library constructed in the vector λ gt11 are plated under lytic growth conditions. After plaques appear, expression of the β -galactosidase fusion proteins encoded by the recombinant phage is induced by placing nitrocellulose filters impregnated with IPTG onto the plates. Phage growth is continued and is accompanied by the immobilization of proteins, from lysed cells, onto the nitrocellulose filters. The filters are lifted after this incubation, blocked with protein, then reacted with a radiolabeled recognition-site DNA (containing one or more binding sites for the relevant sequence-specific protein) in the presence of an excess of nonspecific competitor DNA. After the binding reaction, the filters are washed to remove nonspecifically bound probe and processed for autoradiography. Potentially positive clones detected in the primary screen are rescreened after a round of plaque purification. Recombinants which screen positively after enrichment and whose detection specifically requires the recognition-site probe (not detected with control probes lacking the recognition site for the relevant protein) are then isolated by further rounds of plaque purification.

Materials

pUC recombinant plasmid DNA containing multiple tandem copies of recognition-site DNA and lacking the recognition site (or containing mutant versions; *UNITS 1.5 & 3.16*; see commentary)

*Eco*RI and *Hind*III restriction endonucleases (*UNIT 3.1*)

1 M Tris·Cl, pH 7.5 (*APPENDIX 2*)

5 mM each dCTP, dGTP, dTTP, and dATP (*UNIT 3.4*)

5000 Ci/mmol [α -³²P]dATP

Klenow fragment of *E. coli* DNA polymerase I (*UNIT 3.5*)

500 mM EDTA, pH 8.0

25:24:1 phenol/chloroform/isoamyl alcohol (*UNIT 2.1*)

24:1 chloroform/isoamyl alcohol

3 M sodium acetate, pH 5.2 (*APPENDIX 2*)

100% ethanol

Elution buffer

E. coli Y1090 (Table 1.4.5)

λ gt11 cDNA library (*UNIT 5.8*; see commentary)

LB medium containing 0.2% maltose and 50 μ g/ml ampicillin (*UNIT 1.1* and Table 1.4.1)

0.7% top agarose, melted and equilibrated to 47°C (UNIT 1.1)
100- and 150-mm LB plates containing 50 µg/ml ampicillin, dry and prewarmed to 37°C (UNIT 1.1 and Table 1.4.1)
10 mM IPTG (prepare 1 M stock with sterile H₂O and store at -20°C; Table 1.4.2)

BLOTTO

Binding buffer

1 mg/ml sonicated (to ~1 Kb) calf thymus DNA

Chloroform

Elutip-d columns (Schleicher & Schuell)

37° and 42°C incubators

132- and 82-mm nitrocellulose membrane filters (Schleicher & Schuell)

Additional reagents and equipment for restriction digest (UNIT 3.1), mobility shift DNA-binding assay (UNIT 12.2), phenol extraction and ethanol precipitation (UNIT 2.1), nondenaturing polyacrylamide gel electrophoresis (UNIT 2.7), Elutip-d chromatography (UNIT 2.6), titrating and plating bacteriophage (UNITS 1.11 & 6.1), autoradiography (APPENDIX 3), and purifying and storing bacteriophage clones (UNITS 6.5 & 1.12)

Prepare and analyze radiolabeled recognition-site DNA

1. Digest 20 µg of an appropriate pUC recombinant plasmid DNA (containing multiple tandem copies of the recognition site in the polylinker) with *Eco*RI and *Hind*III in a final volume of 100 µl.

It is preferable to use a probe containing multiple (2 to 10) protein-binding sites with an overall length <200 bp. Longer probes yield higher nonspecific signals. A multisite probe is constructed by cloning tandem copies of a synthetic binding-site oligomer in a pUC vector. The synthetic probe should be tested for recognition by the desired protein before use in screens. This is most easily done using the mobility shift DNA-binding assay.

2. Add the following components to the restriction digest:

1 M Tris·Cl, pH 7.5, to a final concentration of 50 mM
5 mM each dCTP, dGTP, dTTP to 100 µM each final concentration
200 µCi [α -³²P]dATP (5000 Ci/mmol)
10 U Klenow fragment.

Incubate mixture 30 min at room temperature. Add cold 5 mM dATP to 100 µM. Continue the incubation for another 30 min.

This reaction labels the ends of the probe.

3. Stop the end-labeling reaction by adding 500 mM EDTA to 20 mM. Extract the DNA sequentially with 25:24:1 phenol/chloroform/isoamyl alcohol and 24:1 chloroform/isoamyl alcohol, adjust the salt concentration with 3 M sodium acetate, pH 5.2, to 0.3 M final, and precipitate with 100% ethanol.
4. Resuspend the pellet in 200 µl water, add 22 µl of 3 M sodium acetate, pH 5.2, and reprecipitate with 100% ethanol.
5. Separate the labeled fragments by nondenaturing polyacrylamide gel electrophoresis using a 0.75-mm-wide, 5% gel.
6. Visualize the labeled fragments by autoradiography. Cut out a gel slice containing the binding-site fragment. Elute the binding-site fragment by incubating the gel slice in 1.5 ml elution buffer with shaking at 4°C overnight.
7. Remove supernatant with a pipettor and purify the labeled fragment by Elutip-d chromatography. Determine the activity of the ³²P-labeled DNA binding-site probe

by scintillation counting (cpm) and store at 4°C.

These reaction conditions yield DNA probes with specific activities of $2-4 \times 10^7$ cpm/pmol. The probe yield should be 10^8 to 2×10^8 cpm. This amount of probe is sufficient to screen 20 large filters representing $\sim 10^6$ plaques (see screening conditions below). Multisite probes can also be prepared for screening simply by catenation of a binding-site oligonucleotide with DNA ligase followed by nick translation (Vinson et al., 1988).

Prepare nitrocellulose filter replicas

8. Grow *E. coli* Y1090 to saturation (usually overnight) in LB/maltose/ampicillin medium at 37°C.
9. For each plating, mix a 0.5-ml aliquot of the overnight Y1090 culture with λ gt11 library (containing $3-5 \times 10^4$ pfu) in a tube.
10. Incubate the tubes 15 min at 37°C to adsorb phage to cells.
11. Add 9 ml top agarose to each tube. Mix by inverting a few times and plate each suspension onto a 150-mm LB/ampicillin plate.
12. Incubate the plates ~ 3 hr at 42°C until tiny plaques are visible.
13. Move plates to a 37°C incubator. Do not allow plates to cool below 37°C.
14. Prepare 132-mm nitrocellulose filters for overlays (handle filters with blunt-ended forceps) by immersing in 10 mM IPTG for 30 min. Air dry filters at room temperature for 60 min.
15. Overlay each plate with a dry nitrocellulose filter impregnated with IPTG and continue incubation 6 hr at 37°C.
16. Cool plates 10 min at 4°C. Mark the position of each filter with needle holes. Lift filters sequentially and immerse in a deep dish containing BLOTTO. Incubate 60 min at room temperature with gentle swirling on an orbital platform shaker.

Plates are cooled so as to avoid lifting the top agar with the filter. Hybridization chambers (Altec Plastics) are very useful for processing multiple filters in this and subsequent steps. Avoid trapping air bubbles while immersing the filters. The master plates are sealed with Parafilm and stored at 4°C.

17. Transfer filters to a dish containing enough binding buffer to immerse all filters (500 ml/10 filters) and incubate 5 min at room temperature with shaking. Repeat this wash twice with fresh binding buffer.

Filters can be stored in binding buffer for up to 24 hr at 4°C prior to screening. The DNA-binding activities of various recombinant proteins are quite stable under these conditions.

Screen nitrocellulose filter replicas

18. Heat denature 1 mg/ml sonicated calf thymus DNA 10 min at 100°C, quench on ice, and add to binding buffer to 5 μ g/ml final (final volume should be 25 ml/10 filters). To this solution, add 32 P-labeled DNA binding-site probe (from step 7) to $1-2 \times 10^6$ cpm/ml ($\sim 10^{-10}$ M). Add filters into the probe solution one by one in a dish and shake gently 60 min at room temperature.

Filters should be sequentially immersed in probe solution without trapping air bubbles.

19. Wash filters in batches, 4 times (7.5 min per wash, 30 min total) at room temperature with 500 ml binding buffer/10 filters.

Filters should shake freely in large dish without sticking to one another.

20. Blot filters dry, cover with plastic wrap, and expose to X-ray film with an intensifying screen at -70°C .

Positive signals should be detectable after 12 to 24 hr of autoradiography. Longer exposures result in high background.

Identify and purify sequence-specific clones

21. Align phage plates with autoradiographs, isolate agarose plugs corresponding to positive signals, and generate secondary phage stocks.

To reduce the number of irrelevant clones (false positives), it is helpful to generate autoradiography exposures of varying times with the primary filters. Such exposures help to distinguish spots with intense centers (likely to be artifacts) from those with a diffuse halo-like appearance (likely to represent true positives). Multiple autoradiographs also eliminate from consideration positive spots that do not reproduce on each exposure.

22. Plate secondary phage stocks (~ 5000 pfu/100-mm plate) as described in steps 8 to 11. For each plating, use a 0.2-ml aliquot of a Y1090 culture and 3 ml top agarose.
23. Prepare and screen secondary nitrocellulose filter replicas as described in steps 12 to 20, using 82-mm filters.

Recombinants that screen positively after enrichment represent true positives but may encode non-sequence-specific DNA-binding proteins.

24. Label a pVC recombinant plasmid DNA lacking the recognition site (or containing mutant versions) as in steps 1 to 7. Use this labeled probe to screen secondary stocks of true positives.

Phage detected with both test and control probes likely encode non-sequence-specific DNA-binding proteins.

25. Plaque purify and store phage (whose detection specifically requires the wild-type recognition-site probe over) over a drop of chloroform at 4°C .

These phage are likely to encode a β -galactosidase fusion protein that specifically binds the recognition site (see support protocol).

DENATURATION/RENATURATION CYCLING OF DRIED REPLICA FILTERS USING GUANIDINE-HCl

**ALTERNATE
PROTOCOL**

Protein replica filters suitable for screening with DNA recognition-site probes are most easily prepared as described in steps 8 to 17 of the basic protocol. Recently, Vinson et al. (1988) have shown that processing dried nitrocellulose replica filters through a denaturation/renaturation cycle using 6M guanidine-HCl significantly enhances the signal from a λ gt11 recombinant encoding C/EBP. However, it is not possible from this report to directly compare the sensitivity of the two protocols in detecting the C/EBP phage since in the basic protocol the replica filters are not dried. In this alternate protocol, the filters are first washed in HEPES binding buffer with 6 M guanidine-HCl, followed by several buffer washes containing two-fold dilutions of guanidine-HCl. They are then processed as described in the basic protocol.

Additional Materials

HEPES binding buffers

1. Follow steps 1 to 15 of the basic protocol.
2. Cool plates 10 min at 4°C . Mark the position of each filter with needle holes. Lift filters carefully and air dry 15 min at room temperature.

**DNA-Protein
Interactions**

12.7.4

3. Immerse filters in HEPES binding buffer/6 M guanidine-HCl and incubate 10 min at 4°C with gentle shaking. Repeat this once with fresh HEPES binding buffer/6 M guanidine-HCl.
4. Prepare HEPES binding buffer/3 M guanidine-HCl and immerse filters (50 ml/10 filters) 5 min at 4°C. Repeat this wash four times, each time using HEPES binding buffer containing a two-fold dilution of guanidine-HCl from the previous wash.
5. Incubate filters in HEPES binding buffer (50 ml/10 filters) 5 min at 4°C. Repeat this once with fresh HEPES binding buffer.
6. Block filters by incubating in BLOTTO as described in step 16 of the basic protocol.
7. Process filters for screening as described in steps 17 to 25 of the basic protocol.

**SUPPORT
PROTOCOL**

**PREPARATION OF A CRUDE EXTRACT FROM λ gt11
RECOMBINANT LYSOGEN TO CHARACTERIZE DNA-BINDING
ACTIVITY OF THE FUSION PROTEIN**

The specific detection of a recombinant phage with a wild-type recognition-site probe suggests that the phage expresses a β -galactosidase fusion protein that binds specifically to the recognition site. Direct evidence is obtained by analyzing a crude extract from a lysogen harboring the recombinant phage for appropriate DNA-binding activity. *E. coli* Y1089 cells are infected with the recombinant phage and lysogenized clones are isolated. Lysogenized cells growing in liquid culture are induced for the expression of the β -galactosidase fusion protein, harvested, and then lysed. A crude extract is obtained and used to characterize the DNA-binding activity of the fusion protein (see accompanying units in this chapter). Such an analysis not only confirms the sequence specificity of the fusion protein but allows a better definition of its recognition properties.

Additional Materials

- E. coli* Y1089 *hflA*150 (Table 1.4.5)
- LB medium with and without 10 mM MgCl₂ (UNIT 1.1)
- 10¹⁰ pfu/ml λ gt11 recombinant phage stock (see basic protocol)
- 1 M IPTG (Table 1.4.2)
- Extract buffer
- 5 mg/ml lysozyme in extract buffer (store at -20°C)
- 4 M NaCl
- 32°C incubator
- Sterile toothpicks
- 0.025- μ m filter disks (Millipore VS)

Isolate recombinant phage lysogen

1. Grow *E. coli* Y1089 *hflA*150 to saturation in LB/maltose/ampicillin medium at 37°C.
2. Dilute saturated cell suspension 100-fold in LB/MgCl₂ medium. Infect 100 μ l of diluted cell suspension with 5 μ l of 10¹⁰ pfu/ml λ gt11 recombinant phage stock and incubate 20 min at 32°C.

The phage and host cell mixture represents a multiplicity of infection of 5 to 10.

3. Dilute infected cell suspension 1000-fold in LB medium. Spread 100- μ l aliquots on LB/ampicillin plates and incubate overnight at 32°C.

At 32°C the temperature-sensitive phage repressor is functional and thus permits lysogeny. Approximately 100 colonies should be obtained on each plate after an overnight incubation.

4. Test single colonies for temperature-sensitive growth. Spot single colonies using sterile toothpicks onto two LB/ampicillin plates. Incubate the first plate at 42°C and the second at 32°C.

Clones that grow at 32°C but not at 42°C represent lysogens. Lysogens should arise at a frequency of 10% to 80%.

Prepare a crude extract from a recombinant phage lysogen

5. Grow overnight cultures of recombinant phage lysogens in LB/ampicillin medium at 32°C.
6. Inoculate 2-ml aliquots of LB/ampicillin medium with 20- μ l aliquots of the overnight cultures. Incubate at 32°C with good aeration.

A small air-shaker incubator should be used for this purpose. It allows for rapid temperature equilibration required in subsequent steps.

7. Grow cultures to OD₆₀₀ = 0.5 (~3 hr). Shift the temperature of the incubator to 44°C. Incubate 20 min at the elevated temperature.

The temperature shift should be accomplished as rapidly as possible while the cells are in log phase. At the elevated temperature the cI repressor is inactivated, inducing replication of the recombinant phage.

8. Add 1 M IPTG to 10 mM final and reduce the temperature of the incubator to 37°C. Continue the incubation for 1 hr.

Expression of the β -galactosidase fusion protein is induced with IPTG.

9. Microcentrifuge 1-ml aliquots of induced cultures ~45 sec at room temperature.
10. Resuspend each pellet in 100 μ l extract buffer and quick freeze with dry ice/ethanol. If desired, store cell suspensions at -70°C.
11. Rapidly thaw frozen cell suspension. Add 5 mg/ml lysozyme to 0.5 mg/ml final. Incubate 15 min on ice.

At the end of this incubation, the cell suspension should be very viscous due to efficient lysis.

12. Add 4 M NaCl to 1 M final and mix thoroughly. Incubate on rotator for 15 min at 4°C.
13. Microcentrifuge lysates 30 min at 4°C and remove supernatants carefully.
14. Float 0.025- μ m filter disks on 100 ml extract buffer in a petri dish and apply samples on the filters. Dialyze 60 min at 4°C then quick freeze aliquots with dry ice/ethanol and store at -70°C.

Crude extracts are unstable at 4°C because they contain significant proteolytic activity. Frozen aliquots are used subsequently to characterize the DNA-binding activity of the fusion protein (see accompanying units in this chapter).

REAGENTS AND SOLUTIONS

Binding buffer

- 10 mM Tris-Cl, pH 7.5
- 50 mM NaCl
- 1 mM EDTA
- 1 mM dithiothreitol (DTT)

BLOTTO

5% Carnation nonfat milk powder
50 mM Tris-Cl, pH 7.5
50 mM NaCl
1 mM EDTA
1 mM DTT
Store 1 to 2 weeks at 4°C

The milk powder should be thoroughly dissolved in sterile water containing the other components.

Elution buffer

20 mM Tris-Cl, pH 7.5
200 mM NaCl
1 mM EDTA
Store at room temperature

Extract buffer

50 mM Tris-Cl, pH 7.5
1 mM EDTA
1 mM DTT
1 mM Phenylmethylsulfonyl fluoride (PMS)
Store at -20°C

HEPES binding buffers

25 mM HEPES, pH 7.9
25 mM NaCl
5 mM MgCl₂
0.5 mM DTT

Prepare separate solutions without guanidine-HCl and with 6 M guanidine-HCl. Prepare solutions fresh for each use and keep briefly at 4°C.

COMMENTARY

Background Information

Sequence-specific DNA-binding proteins function in diverse capacities including DNA replication, recombination, and transcription. In higher eukaryotic cells, biochemical approaches have been used to identify such proteins (see accompanying units in this chapter). Typically, these proteins bind specifically and with high affinity to distinct sequence motifs. Prior to the application of the cloning strategy described by Singh et al. (1988), genes encoding these proteins could only be isolated through purification of substantial amounts of the relevant activities (Weinberger et al., 1985; Kadonaga et al., 1987). The screening strategy discussed below obviates purification of a DNA-binding protein for the purpose of isolating its gene. Therefore, it is ideally suited for facilitating the analysis of rare regulatory molecules.

The strategy is designed to directly detect recombinant clones that encode sequence-specific DNA-binding proteins. It depends on the

functional expression of the DNA-binding domain of a sequence-specific protein in *E. coli* at high levels and a strong interaction between this domain and its recognition site. If these conditions are fulfilled, a recombinant clone encoding a sequence-specific DNA-binding protein can be detected by probing nitrocellulose filter replicas of a lambda expression library with the corresponding recognition-site DNA (Singh et al., 1988; Staudt et al., 1988). This strategy is quite similar to that previously developed for the isolation of genes by screening recombinant expression libraries with antibodies specific for given proteins (Young and Davis, 1983; see *UNIT 6.7*). In fact, the phage expression vector λ gt11—designed for immunological screening—is also employed in the detection of DNA-binding protein clones. The simple procedure described in the basic protocol has permitted the isolation of many clones encoding different transcription factors including MBP-1 (Singh et al., 1988), Oct-2 (Staudt et al., 1988), E12 (Murre

et al., 1989), XBP (Hsiou-Chi et al., 1988), IRF-1 (Miyamoto et al., 1988), and CREB (Hoeffler et al., 1988). The denaturation/re-naturation cycle described in the alternate protocol may in some cases increase the detection signal by facilitating the correct folding of a larger fraction of the *E. coli* expressed protein. Alternatively, it may help to dissociate insoluble aggregates that form as a consequence of overexpression. This modified procedure has been successfully used to isolate clones encoding Oct-1 (Sturm et al., 1988), Oct-2 (Müller et al., 1988), Pit-1 (Ingraham et al., 1988), and RF-X (Reith et al., 1989).

Prospects for the isolation of cDNAs encoding sequence-specific DNA-binding proteins by this strategy can be assessed by examining the three assumptions on which it is based: (1) functional expression of the DNA-binding domain of the sequence-specific protein in *E. coli*; (2) a strong and specific interaction of the binding domain with its recognition site; and (3) high-level expression of the DNA-binding domain. A number of eukaryotic sequence-specific DNA-binding proteins have been functionally expressed in *E. coli* including the proteins GAL4 (Keegan et al., 1986) and GCN4 (Arndt and Fink, 1986) of yeast, engrailed of *Drosophila* (Desplan et al., 1985), E2 of the bovine papillomavirus (Androphy et al., 1987), and EBNA-1 of the Epstein-Barr virus (Rawlins et al., 1985). Thus it is reasonable to expect the functional expression of the DNA-binding domains of many eukaryotic sequence-specific proteins in *E. coli*. Interestingly, in most cases the functional DNA-binding domain is contained within a small tract of 100 to 300 amino acids. Therefore, successful screening is not dependent on full-length cDNA clones. It simply requires that a given recombinant expression library contain partial cDNA clones spanning the DNA binding domain of the sequence-specific protein. However, if this DNA-binding domain is encoded in the 5' region of a long mRNA, the frequency of functional cDNA clones in a library made by oligo(dT) priming is expected to be low. It is therefore preferable for the purpose of this screening strategy to construct a library by randomly priming cDNA synthesis (Staudt et al., 1988). Finally, it should be noted that some eukaryotic sequence-specific proteins have been shown to recognize target DNA only as heterodimers (Chodosh et al., 1988). Obviously, genes encoding the subunits of these proteins cannot be cloned by this strategy.

Successful screening may be restricted to

proteins with relatively high binding constants since only these are likely to form complexes with half-lives long enough to withstand the wash protocol. For example, if the DNA-binding protein has an association constant of 10^{10} M^{-1} , then under the screening conditions (the DNA probe in excess at a concentration of $\sim 10^{-10} \text{ M}$), approximately half of the active molecules on the filter will have DNA bound. Since the filters are subsequently washed for 30 min, the fraction of protein-DNA complexes that remain will be determined by their dissociation rate constant. Assuming a diffusion-limited association rate constant of $10^7 \text{ M}^{-1} \text{ S}^{-1}$, the dissociation rate constant in solution will be 10^{-3} S^{-1} . This rate constant translates into a half-life of ~ 10 min. Thus one-eighth of the protein-DNA complexes should survive the 30-min wash. For a binding constant of 10^9 M^{-1} , about one-tenth of the active protein molecules will have DNA bound, but virtually all of this signal should be lost since the half-life of these complexes in solution is ~ 1 min.

However, it is unclear whether the equilibrium and kinetic constants of a protein-DNA interaction in solution accurately describe the reversible binding of a DNA probe to a matrix of protein immobilized on a filter. Thus, it may be possible to isolate recombinants encoding proteins with binding constants of $\leq 10^9 \text{ M}^{-1}$. These considerations nevertheless suggest that the sensitivity of this methodology for low-affinity proteins might be significantly enhanced by using DNA probes containing multiple binding sites that are spaced such that the probe can simultaneously bind two or more immobilized protein molecules. Enhanced sensitivity with a multisite probe has been impressively demonstrated in the molecular cloning of a mammalian regulatory protein (Staudt et al., 1988). Finally, since the binding constants of sequence-specific proteins are dependent on ionic strength, temperature, and pH, manipulation of these factors might also enhance detection.

The DNA-binding domains of sequence-specific proteins need to be overexpressed in *E. coli* to permit detection with radiolabeled recognition-site probes. These proteins, when expressed to a level of $\sim 1\%$ of the total cellular protein, can be readily detected (Singh et al., 1988; Staudt et al., 1988). This level of recombinant protein expression is typical of λ gt11.

Critical Parameters

The success of the screening strategy is critically dependent on the quality of the cDNA

library and the recognition-site probe used. Recombinant cDNA libraries made by random priming and multisite recognition probes are the preferred reagents. The synthetic probe should be tested for recognition by the desired protein before use in screens (UNIT 12.2). If possible, the highest affinity site among a set of related sites should be chosen for the construction of the multisite probe.

Plating mixtures for 100- and 150-mm plates should contain a maximum of 2×10^4 pfu and 5×10^4 pfu, respectively, of λ phage. Higher plating densities result in diminished signals. Plating mixtures in top agarose should be spread uniformly by swirling, avoiding air bubbles. This ensures efficient contact of the nitrocellulose filters with the agarose overlay. Plates should be cooled at 4°C before lifting the nitrocellulose filters. This step minimizes peeling off of the agarose overlay. Filters should be submerged in BLOTTO without trapping air bubbles to permit uniform blocking. Filters should be transferred from one solution to the next as rapidly as possible to avoid drying.

The DNA probe preparation used for screening ($\sim 10^8$ cpm/ μ g) should not contain significant levels of either radiolabeled nucleotide or radiolabeled single-stranded DNA because these greatly increase the background. The quality of a probe preparation can be checked by screening control λ gt11 plaques (there should be no signal). The protein-DNA complexes formed on the nitrocellulose filters appear to be labile. Therefore, filters should be shaken gently during the bind and wash steps.

Sonicated and denatured calf thymus DNA is a preferred nonspecific competitor. Earlier screens using poly(dI-dC)-poly(dI-dC) as the nonspecific competitor yielded recombinant phage encoding proteins that bound nonspecifically to either double-stranded or single-stranded DNA (Singh et al., 1988). The signal from these phage, but not phage encoding sequence-specific DNA-binding proteins, could be efficiently competed with denatured calf thymus DNA (R. Clerc, personal communication; Staudt et al., 1988).

A recombinant phage specifically detected with a given recognition-site probe may not necessarily encode a protein that binds to the desired site. It is therefore necessary to characterize the recognition properties of the phage-encoded fusion protein thoroughly with in vitro binding assays. The relation between a protein cloned by this strategy and a previously char-

acterized DNA-binding activity can be explored with antibodies generated against the cloned molecule.

Anticipated Results

If 10^6 recombinant clones are screened, 1 to 10 putative clones can be expected.

Time Considerations

The screening procedure is well suited for a 3-day cycle time. The phage are plated in the morning of the first day. Nitrocellulose filters are placed on plates 3 hr after the infection. After a further 6-hr incubation, the filters are lifted and blocked for 1 hr. Filters are rinsed and stored immersed in buffer overnight at 4°C. Filters are screened on the second day with a DNA recognition-site probe and processed for autoradiography. Positive signals are detected after 12 to 24 hr of autoradiography. If the filters are to be put through a denaturation/renaturation cycle, it is convenient to stop on the first day after they are dried.

Literature Cited

- Androphy, E.J., Lowy, D.R., and Schiller, J.T. 1987. Bovine papillomavirus E2 trans-activating gene product binds to specific sites in papillomavirus DNA. *Nature (Lond.)* 325:70-73.
- Arndt, K. and Fink, G. 1986. GCN4 protein, a positive transcription factor in yeast, binds general control promoters at 5'TGACTC3' sequences. *Proc. Natl. Acad. Sci. U.S.A.* 83:8516-8520.
- Chodosh, L.A., Olesen, J., Hahn, S., Baldwin, A.S., Guarente, L., and Sharp, P.A. 1988. A yeast and a human CCAAT-binding protein have heterologous subunits that are functionally interchangeable. *Cell* 53:25-35.
- Desplan, C., Theis, J., and O'Farrell, P.H. 1985. The *Drosophila* developmental gene, engrailed, encodes a sequence-specific DNA binding activity. *Nature (Lond.)* 318:630-635.
- Hoeffler, J.P., Meyer, T., Yun, Y., Jameson, J.L., and Habener, J.F. 1988. Cyclic AMP-responsive DNA-binding protein: Structure based on a cloned placental cDNA. *Science* 242:1430-1433.
- Hsiou-Chi, L., Boothby, M.R., and Glimcher, L.H.. 1988. Distinct cloned class II MHC DNA binding proteins recognize the X box transcription element. *Science* 242:69-71.
- Ingraham, H.A., Chen, R., Mangalam, H.J., Eisholtz, H.P., Flynn, S.E., Lin, C.R., Simmons, D.M., Swanson, L., and Rosenfeld, M.G. 1988. A tissue-specific transcription factor containing a homeo domain specifies a pituitary phenotype. *Cell* 55:519-529.

- Kadonaga, J.T., Carner, K.R., Masiarz, F.R., and Tjian, R. 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* 51:1079-1090.
- Keegan, L., Gill, G., and Ptashne, M. 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* 231:699-704.
- Miyamoto, M., Fujita, T., Kimura, Y., Maruyama, M., Harada, H., Sudo, Y., Miyata, T., and Taniguchi, T. 1988. Regulated expression of a gene encoding a nuclear factor, IRF-1, that specifically binds to IFN- β gene regulatory elements. *Cell* 54:903-913.
- Murre, C., Schonleber-McCaw, P., and Baltimore, D. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD and myc proteins. *Cell* 56:777-783.
- Müller, M.M., Ruppert, S., Schaffner, W., and Matthias, P. 1988. A cloned octamer transcription factor stimulates transcription from lymphoid specific promoters in non-B cells. *Nature (Lond.)* 336:544-551.
- Rawlins, D.R., Milman, G., Hayward, S.D., and Hayward, G.S. 1985. Sequence-specific DNA binding of the Epstein-Barr viral nuclear antigen (EBNA-1) to clustered sites in the plasmid maintenance region. *Cell* 42:859-868.
- Reith, W., Barras, E., Satola, S., Kobr, M., Reinhart, D., Herrero Sanchez, C., and Mach, B. 1989. Cloning of the major histocompatibility complex class II promoter affected in a hereditary defect in class II gene regulation. *Proc. Natl. Acad. Sci. U.S.A.* 86:4200-4204.
- Singh, H., LeBowitz, J.H., Baldwin, A.S., and Sharp, P.A. 1988. Molecular cloning of an enhancer binding protein: Isolation by screening of an expression library with a recognition site DNA. *Cell* 52:415-423.
- Staudt, L.M., Clerc, R.G., Singh, H., LeBowitz, J.H., Sharp, P.A., and Baltimore, D. 1988. Molecular cloning of a lymphoid-specific cDNA encoding a protein that binds to the regulatory octamer DNA motif. *Science* 241:577-580.
- Sturm, R.A., Das, G., and Herr, W. 1988. The ubiquitous octamer protein Oct-1 contains a Pou domain with a homeo subdomain. *Genes & Dev.* 2:1582-1599.
- Vinson, C.R., LaMarco, K.L., Johnson, P.F., Landschulz, W.H., and McKnight, S.L. 1988. *In situ* detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. *Genes & Dev.* 2:801-806.
- Weinberger, C., Hollenberg, S.M., Ong, E.S., Harmon, J.M., Brower, S.T., Cidlowski, J., Thompson, E.B., Rosenfeld, M.G., and Evans, R.M. 1985. Identification of human glucocorticoid receptor complementary DNA clones by epitope selection. *Science* 228:740-742.
- Young, R.A. and Davis, R.W. 1983. Yeast RNA polymerase II genes: Isolation with antibody probes. *Science* 222:778-782.

Key Reference

- Huynh, T.V., Young, R.A., and Davis, R.W. 1985. Construction and screening cDNA libraries in λ gt10 and λ gt11. *In* DNA Cloning, Vol. 1: A Practical Approach (D.M. Glover, ed.) pp. 49-78. IRL Press, Oxford.

Provides an excellent description of the λ gt11 expression screening system.

Contributed by Harinder Singh
University of Chicago
Chicago, Illinois

Rapid Separation of Protein-Bound DNA from Free DNA Using Nitrocellulose Filters

BASIC PROTOCOL

This protocol relies on the ability of nitrocellulose to bind proteins but not double-stranded DNA. Use of radioactively labeled double-stranded DNA fragments allows quantitation of DNA bound to the protein at various times and under various conditions, permitting kinetic and equilibrium studies of DNA-binding interactions. Purified protein is mixed with double-stranded DNA in an appropriate buffer to allow interaction. After incubation, the mixture is suction filtered through nitrocellulose, allowing unbound DNA to pass through the filter while the protein (and any DNA interacting with it) is retained.

Materials

- 2× DNA-binding buffer
- Double-stranded, ³²P-labeled DNA
- Protein sample
- Bovine serum albumin (BSA; optional)
- 0.45- μ m nitrocellulose filter discs (13-, 24-, or 25-mm diameter to fit filtration apparatus being used)
- Flat, smooth-tip forceps for handling nitrocellulose
- Filtration apparatus and vacuum source

Prepare materials

1. Prepare 30 ml of fresh 2× DNA-binding buffer by adding DTT and PMSF (see reagents and solutions). Mix well and set aside 0.5 ml for use in step 3. Dilute remaining 29.5 ml to 1× with distilled water.

This will be enough 2× and 1× DNA-binding buffer to set up and filter ten reactions. This is a good general buffer but the optimal buffer conditions for binding a given protein must be individually determined. See critical parameters.

CAUTION: PMSF is toxic. Handle carefully!

2. Place the filters in a small tray containing ~30 ml of 1× DNA-binding buffer (diluted in step 1). Soak \geq 30 min. Retain at least 10 ml of 1× buffer for the filtration step.

Always handle the filters with forceps because direct handling can increase background. Any filters that do not wet uniformly should be discarded. If desired, the filters can be carefully labeled at the edge with a pen prior to wetting. Special pens are available, although ballpoint pens work well.

Bind DNA to protein

3. For each sample, set up a microcentrifuge tube as follows. Add \geq 10,000 cpm double-stranded, ³²P-labeled DNA (2 to 10 ng) to 25 μ l of 2× DNA-binding buffer. Add sufficient water to make a final volume of 50 μ l, accounting for the volume of protein solution added in step 4. Vortex and microcentrifuge 1 to 2 sec.

Typically DNA fragments of a few hundred base pairs or more—which are believed to contain a specific binding site—are used. DNA fragments can be labeled by filling in the ends with Klenow fragment (UNIT 3.5) or by kinasing (UNIT 3.10). Fragments are isolated by any of the standard techniques described in Chapter 2 or in UNIT 12.2.

4. Add protein solution to the reaction tubes. GENTLY invert or flick the tubes up to 4 times to mix in the protein solution. Microcentrifuge 1 to 2 sec to bring all the liquid to the bottom of the tubes. Incubate 25 min at 30°C.

Time and temperature parameters will vary for different proteins. In general, 20 min is sufficient to allow the protein-DNA interaction to reach equilibrium. Temperatures above 37°C may affect the stability of the protein.

The final concentration of protein needed to detect specific binding is best determined by titrating (see background information) and can range from 10^{-8} to 10^{-14} M. Low concentrations of protein may be stabilized by including in the reactions a non-DNA-binding carrier protein (e.g., BSA at ~100 mg/ml) with no detrimental effect on the binding assay.

Filter the DNA-protein complex and determine cpm

5. During the incubation, number and label scintillation vials. Attach a vacuum source to the apparatus as indicated in Figure 12.8.1 but do not apply vacuum.

The most common filtration apparatus holds a single filter and can be set up using supplies found in most labs (Fig. 12.8.1A). A multiwell apparatus is more convenient, allowing filtration of several samples simultaneously. We recommend the Hoefer 10-well apparatus, which has stopcocks to control suction for individual wells and allows filtrate to be collected from individual wells (Fig. 12.8.1B). Although a standard suction-filtration setup is described here, disposable units for filtration by centrifugation are available from Schleicher & Schuell (Centrex).

If the Hoefer 10-well apparatus is being used, number enough vials to count the filtrates and the filters and place the vials in the collection rack in the apparatus. The flowthrough and all washes can be collected into one vial for each reaction. The filter is counted separately.

6. Place filters in the apparatus with weights on top, as shown in Figure 12.8.1 (do this no more than 5 min prior to start of filtration to avoid drying out the filters). At the end of the incubation, turn on the vacuum source. Apply the contents of each tube gently but quickly to separate filters. Filtering should proceed at ~1 ml/min.
7. When the liquid is gone from each filter well, add 0.5 ml of 1× DNA-binding buffer (this comprises the first wash). Allow to filter through. Apply the second 0.5-ml 1× buffer wash. When washing is complete, increase the suction and observe the liquid dripping into the collection vessels. Turn off the vacuum when dripping has ceased.
8. Remove the weights and transfer filters to scintillation vials for Cerenkov counting. Total cpm is determined by adding the counts in the filtrate and those on the filters (where the filtrates were collected), or by estimating input cpm. Percent cpm retained = $\text{cpm on filter} / \text{total cpm} \times 100$; this reflects the percentage of input DNA bound by protein at the time of filtration (see background information).

For further characterization, see support protocol describing recovery of DNA from filters.

DETECTION OF SPECIFICITY IN DNA BINDING

When the binding site of a protein is unknown, the pure protein can be added to a mixture of fragments to select those fragments of DNA for which it has the greatest affinity. Specificity of binding can be influenced by the buffer conditions and filtering regimen. This alternate protocol creates conditions that disrupt weaker, presumably nonspecific binding interactions, while retaining the stronger binding interactions. Differences from the basic protocol include using more labeled DNA and more extensive washing. The goal is to recover enough of a single input fragment to visualize by subsequent autoradiography. This protocol uses the same binding buffer as the basic protocol, however increasing the KCl to 150 mM (final concentration in the 1× binding buffer) could further improve detection of specific binding in some cases.

ALTERNATE PROTOCOL

DNA-Protein Interactions

12.8.2

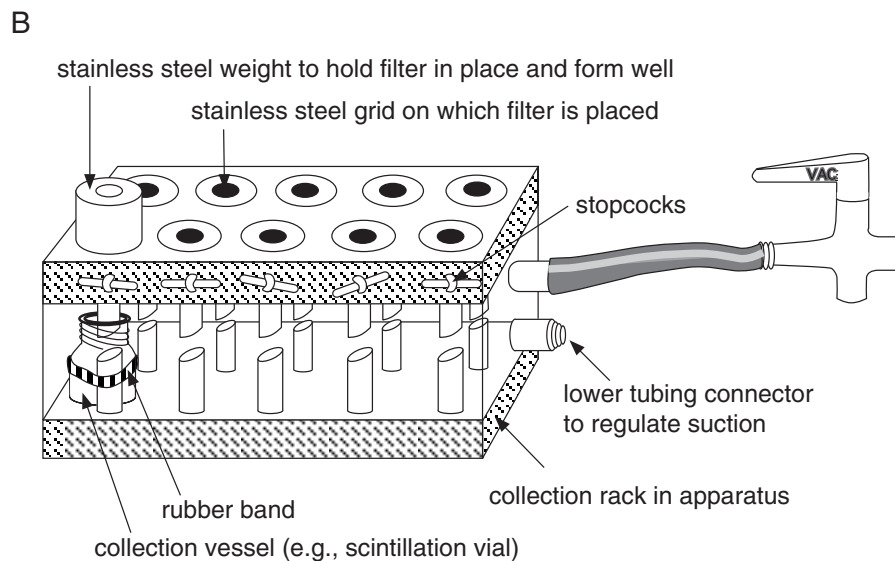
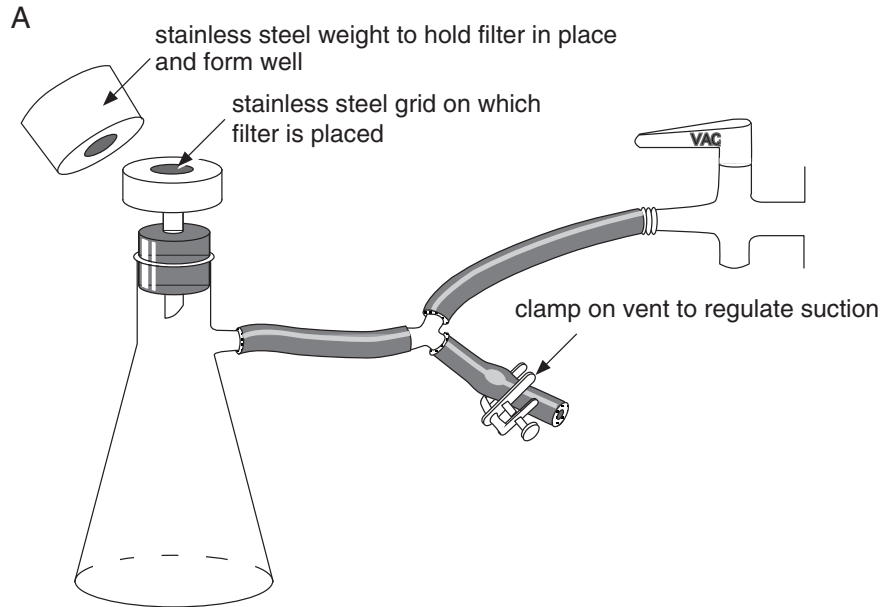


Figure 12.8.1 (A) A single filtration apparatus can be set up with a sidearm flask connected to a vacuum source. The tubing arrangement need not be precisely as depicted but should include an adjustable clamp on a vent to allow precise control of the vacuum. (B) The 10-well filtration apparatus from Hoefer contains a removable rack for collection vessels. Many types of tubes or vials can be held in place by a rubber band as shown. Stopcocks allow vacuum to be shut off to any subset of the 10 wells, so that any number of samples from 1 to 10 can be filtered at one time. The lower tubing connector can be set up with tubing and a clamp to regulate the vacuum as with the vent on the sidearm flask unit. In both setups the interlocking design and weight of the steel are sufficient to create a seal around the filter.

1. Prepare 40 ml of fresh 2× DNA-binding buffer (see step 1 of basic protocol). Set aside 3 ml and dilute remaining 37 ml to 1× with distilled water.
2. Soak filters in ~10 ml of 1× binding buffer. Retain 60 ml of 1× binding buffer for the filtration step.
3. Add ~100,000 cpm of double-stranded, ³²P-labeled DNA (20 to 100 ng of a mixture of labeled fragments) to 250 μl of 2× DNA binding buffer. Add sufficient water to make a final volume of 500 μl, accounting for the volume of protein solution added in step 4. Vortex and microcentrifuge 1 to 2 sec.

As in the basic protocol, the method of labeling is not critical. However, there may be some differences in retention of fragments based on size alone.

4. Add the protein solution to the reaction tubes, mix, microcentrifuge, and incubate, as described in step 4 of the basic protocol.

As in the basic protocol, time and temperature parameters will vary for different proteins. The amount of protein should be titrated to determine the ratio of protein to DNA that gives the optimal specific signal. High concentrations of protein result in all DNA fragments in the mixture being retained on the filter.

5. During the incubation, number and label scintillation vials for counting of filters, and of flowthrough and washes if using Hoefer apparatus. Set up filtration apparatus as described in step 6 of the basic protocol.

If using the Hoefer apparatus, the flowthrough and all washes can be collected into one vial for each reaction. However, it is often reassuring to collect the last wash separately for comparison to filter-bound cpm. It is expected that effective washing will leave fewer cpm in the last wash than retained on the filter. The importance of washing is further discussed in critical parameters.

6. Place filters in the apparatus with weights on top (Fig. 12.8.1)—again, ≤5 min prior to the start of filtration to avoid drying the filters. At the end of the incubation, apply each sample to a separate filter with the vacuum off. Apply vacuum to filter at ~1 ml/min.
7. When all liquid is gone from the filter, increase the vacuum until dripping into the collection vessel slows. Release the vacuum. Add 2 ml of 1× buffer for the first wash to the wells. If the next wash is to be collected in separate vials, the vials can be changed quickly at this time while the 2 ml of buffer keeps the filter from drying. Apply vacuum and filter the first wash as above.

Repeat this washing procedure a total of three times. On the third wash increased suction is used until dripping ceases.

ELUTION OF BOUND DNA

In some cases the quantitation (by scintillation counting) of DNA retained is not sufficient information. This support protocol describes how the DNA can be recovered from the filters for further analysis by gel electrophoresis or amplification and cloning.

Additional Materials

Filter elution buffer
Carrier nucleic acid
Ethanol

1. Remove filters and place face down in scintillation vials containing 450 μl filter elution buffer.

SUPPORT PROTOCOL

DNA-Protein Interactions

12.8.4

2. Place vials in 30°C water bath, weighted down, for 2 hr. Periodically tilt vials gently to swirl the buffer.

Use lead pig lids or plate glass to weigh down the vials. It is most efficient to keep as much of the filter as possible in contact with the buffer. Any other arrangement that combines this temperature and gentle agitation should serve just as well. Yield with this method is usually 50% to 85% of the cpm originally on the filter.

3. Remove the liquid and transfer it to microcentrifuge tubes. Add carrier nucleic acid and mix. Add 1 ml ethanol, mix, and allow to precipitate overnight at -20°C. Run on a gel (Chapter 2) to determine which subset of the input fragments was retained.

Precipitation should go overnight because of the small amount of DNA in the large volume. However, sufficient carrier (~1 µg per tube) can decrease the time needed for precipitation. Conventional carriers such as tRNA will do; however, for analysis by agarose gel electrophoresis, it is helpful to use as carrier cold DNA fragments that are identical to the labeled ones. Directly after running the gel, a picture can be taken of the gel stained with ethidium bromide. After drying and autoradiography, the radioactive band(s) can be identified by comparing the autoradiogram to the picture.

REAGENTS AND SOLUTIONS

2× DNA-binding buffer

40 mM Tris-Cl, pH 7.5
20% glycerol
100 mM KCl
0.2 mM dithiothreitol (DTT)
0.2 mM PMSF (phenylmethylsulfonyl fluoride) in isopropanol

DTT and PMSF should be added to binding buffer immediately before use, as described in step 1 of basic protocol.

CAUTION: PMSF is extremely toxic. Handle carefully!

Filter elution buffer

20 mM Tris-Cl, pH 7.8
0.2% sodium dodecyl sulfate
0.3 M sodium acetate

COMMENTARY

Background Information

Filter binding is used predominantly to characterize interactions between purified DNA-binding proteins and their specific sites as well as various nonspecific DNAs. This assay is ideal for kinetic and equilibrium studies because bound DNA can be rapidly separated from free DNA. However, slower methods such as the mobility shift (UNIT 12.2) or immunoprecipitation (UNIT 10.16) assays, have the advantage that they can be used with cruder protein preparations.

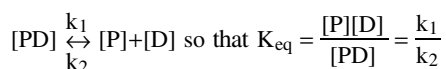
DNA retained on the filters can be recovered for gel electrophoresis. This has allowed selection of fragments containing a specific binding

site from a pool of several fragments in the initial reaction mix (Strauss et al., 1981). Filter binding has also been used to separate bound DNA from free DNA for methylation interference experiments (UNIT 12.3), although the mobility shift assay is more popular for this purpose.

The use of the filter-binding technique for kinetic and equilibrium analyses is described briefly below. For more detailed discussion of the validity of these approaches, as well as alternative approaches, consult Riggs et al. (1970) and Hinkle and Chamberlin (1972b).

The protein-DNA binding interaction can be looked at as an equilibrium represented

by the equation:



where [P] = concentration of free protein

[D] = concentration of free DNA

[PD] = concentration of protein-DNA complex

k_1 = dissociation rate constant

k_2 = association rate constant

K_{eq} = equilibrium dissociation constant (often refer to as the dissociation constant, K_D)

Equilibrium studies. The K_{eq} of a given protein-DNA interaction is a measure of the affinity of the protein for that particular piece of DNA. The apparent K_{eq} can be determined by an experiment in which protein is titrated against a known amount of DNA containing the binding site being tested. The half-maximal point on the resulting curve (see Fig. 12.8.2) is equal to the K_{eq} , since at 50% saturation of the DNA, $[PD] = [D]$. Therefore, the equation

$$K_{eq} = \frac{[P][D]}{[PD]}$$

is reduced to $K_{eq} = [P]$. This simplification is true only if $[D] \ll K_{eq}$. Therefore, the amount of DNA used in these experiments is usually very small.

This experimental method of determining K_{eq} also relies on the assumption that [P] equals

the concentration of active protein molecules with respect to DNA binding. If this is not the case (for example, if the protein is only 30% active after purification), kinetic experiments can give an estimate of the K_{eq} .

Kinetic studies. This kind of estimate of the K_{eq} is based on the relation $K_{eq} = k_1/k_2$. For many DNA binding proteins the “on rate” (rate of association) is diffusion limited. By making this assumption, and experimentally determining the “off rate” (rate of dissociation of formed complexes), an estimate for the K_{eq} can be obtained. The off rate is experimentally measured as follows: (1) protein and labeled DNA (containing the specific binding site) are mixed and allowed to reach equilibrium; (2) the reaction volume is diluted 10-fold in the presence of an excess of cold competitor DNA—e.g., 50-fold excess of unlabeled DNA containing the specific binding site; and (3) aliquots are filtered at various time points. As proteins dissociate they have little chance of reassociating with labeled DNA due to dilution and competitor DNA factors.

A dissociation curve is generated where the filter-bound cpm at $t = 0$ are defined as 100% binding (see Fig. 12.8.3). This data can be applied to the following rearranged form of the equation for first-order rates of decay:

$$\ln(\% \text{ bound}) = \ln(\% \text{ bound at start} = 100) - k_1 t.$$

From the experimentally determined half-time of dissociation (t at which % bound = 50), this equation can be solved for k_1 , the apparent dissociation rate constant. An estimate for the

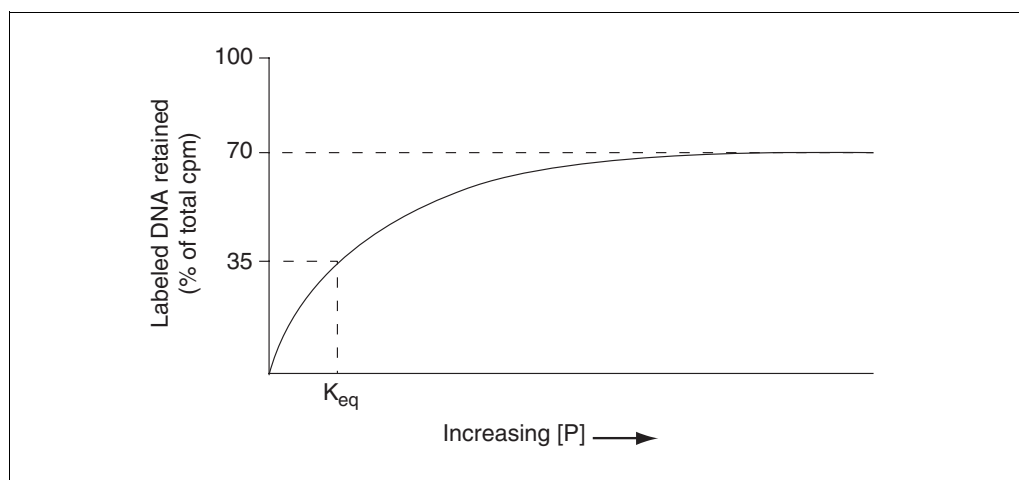


Figure 12.8.2 Plot of increasing concentration of protein vs. amount of DNA retained (expressed as % of total input labeled DNA in cpm). In this example, plateau occurs when 70% of labeled DNA is bound. Therefore, the half-saturation point is when half of this possible maximum is reached—i.e., when 35% of labeled DNA is bound. The concentration of protein necessary to achieve this half point is an experimental estimate of the K_{eq} (see discussion in background information).

K_{eq} is readily made using this value for k_1 and a diffusion limited value for k_2 (Hinkle and Chamberlin, 1972b).

The K_{eq} of a protein for DNAs other than those containing its specific site can be calculated from filter-binding data obtained by equilibrium or rate competition experiments as described by Lin and Riggs (1972). Equilibrium competition experiments involve the mixing of various amounts of cold competitor with a fixed amount of labeled specific DNA during the binding incubation. The amount of competitor needed to reduce labeled DNA retained measures the affinity of the protein for that competitor. Such experiments are particularly useful when the competitor is a single-stranded nucleic acid that could bind directly to the filter, precluding a direct measurement, or when the competitor is difficult to label (Riggs et al., 1970).

Literature Review

Filter binding was first used to study ribosome-RNA interactions (Nirenberg and Leder, 1964) and was applied to DNA-protein interactions by Jones and Berg (1966) to study RNA polymerase. Earlier studies of protein-nucleic acid interactions had relied on glycerol or sucrose gradients for separation of molecules and molecule complexes.

The first detailed discussion of the filter binding method and its general applicability was by Riggs et al. (1970). Beginning in the early 1970s this technique was used exten-

sively—e.g., by Lin and Riggs (1972, 1975) to characterize *lac* repressor binding, and by Hinkle and Chamberlin (1972a,b,c) and later Strauss et al. (1980a,b; 1981) to characterize *E. coli* RNA polymerase binding. Since then, the general applicability of this technique has been borne out as it has been used to analyze many other DNA-binding proteins.

Critical Parameters

The optimal filter-binding conditions are usually different for every protein. In addition, those conditions which permit the strongest binding to a known specific site are often not the conditions desired when one is looking for such a site (as in the alternate protocol). This is because maximizing specific binding usually increases nonspecific binding.

The following is a brief discussion of some parameters to consider when designing an experiment. It is best to titrate individual components initially to establish the optima for the protein of interest.

Incubation conditions

Monovalent salt (e.g., NaCl or KCl). Lower salt favors tighter binding (slower dissociation) whereas higher salt favors quicker dissociation. Most kinetic and equilibrium experiments are done at 50 mM salt. When attempting to detect specific binding over nonspecific binding, 150 mM or higher salt may be appropriate. At high salt concentrations all binding is lost.

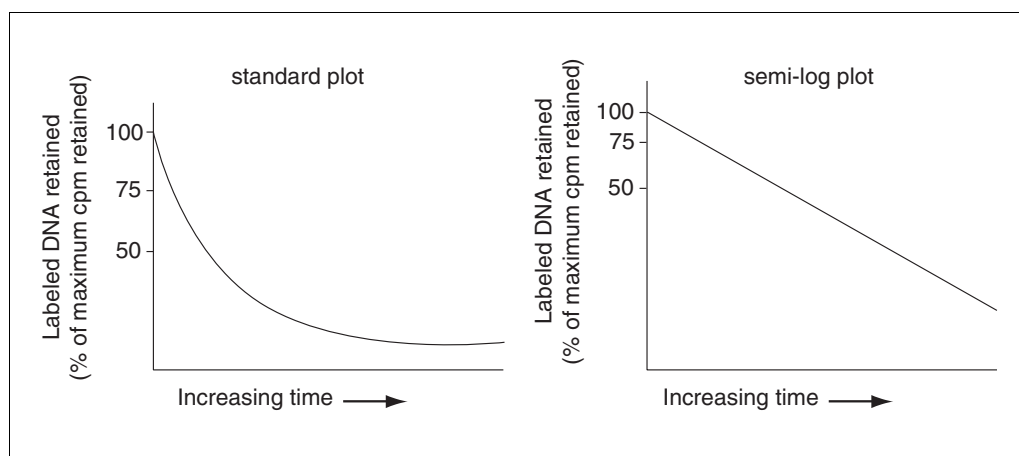


Figure 12.8.3 Two alternative ways to plot the results of a dissociation curve. Here the initial value of 100% equals the maximum amount of labeled DNA (measured in cpm) retained at the start of the experiment. Even though the maximum possible may only be 70% of the input labeled DNA (as shown in Fig. 12.6.2), if an amount of protein is used close to this plateau then this 70% value is the starting point for any dissociation and is therefore defined as 100%. The semi-log plot uses a log scale for the cpm bound to represent the data as a straight line whose slope is $-k_1$, the dissociation constant (see background information and Barkley et al., 1975).

pH. Lower pH favors tighter binding whereas higher pH favors dissociation. Most binding studies are done in the physiological range of 7.2 to 7.5. To accentuate specificity, it may be helpful to use a slightly higher pH. However, as this tends to weaken all binding, increasing either salt concentration or the number of wash steps may be preferable.

Divalent cations (e.g., Mg²⁺). The effect of divalent cations varies with each protein. For some proteins they increase binding but for others they decrease binding (thus, none were included in the basic protocol). It is best to titrate each protein against this parameter; however, binding can frequently be detected in the absence of divalent cations.

Temperature. Temperature effects in this assay are minimal. Most binding protocols consistently use one temperature in the range of 25° to 30°C. Some protocols call for dilution into ice-cold buffer prior to filtering or transfer of tubes to ice. This does not appear to be critical in most cases.

Time. Incubation time is usually between 10 min and 1 hr. Longer times favor equilibrium, but stability of the protein must be considered. Once the reaction has been filtered, complexes continue to dissociate during wash steps but presumably at a slower rate. Nonetheless, it is desirable to conclude the washing as quickly as possible.

Volume. Volumes smaller than 50 µl are not usually used due to the surface area of the filters. Larger volumes may provide more even filtering but necessitate using more protein to attain the desired concentrations. For rate experiments, it is reasonable to make up one large reaction and filter aliquots at various time points.

Manipulations

Mixing. Physical mixing has been reported to disrupt DNA-protein interactions permanently in some cases (Fried and Crothers, 1981). The minimum force possible should be used to mix the protein into the reaction. Avoid vortexing.

Filtration. Successful rates of filtration vary from less than 1 ml/min to 15 ml/min. The more rapid filtration rates have been reported to affect results adversely in some cases by allowing insufficient time for binding to nitrocellulose or by drying out the filter itself, leading to less reproducible results (Riggs et al., 1970).

Washing. The number and volume of washes depends on the application. For kinetic and equilibrium experiments it is not necessary to

wash extensively. One or two small washes (0.5 ml) can improve the reproducibility of the results by decreasing the trapping of reaction volume in the apparatus and within the filter itself. This is particularly helpful if the reaction volume is small. For detection of specificity in binding, more extensive washing is recommended to decrease the background (protein binding to nonspecific DNA). Three large (2-ml) washes as described in the alternate protocol should be sufficient for this purpose, and more may decrease the specific signal.

Troubleshooting

Background retention of labeled DNA on filters in the absence of protein can frequently be as high as 15% of input counts. Pretreating the filters in the following manner (Lin and Riggs, 1972) can reduce the background to between 0.1% and 1.5% of input counts. (Note: Because the treated filters become somewhat yellowed and more fragile, care must be taken in manipulating them.) Filters are wet and soaked in 0.4 M KOH for 40 min at room temperature (~25°C). Next, they are rinsed thoroughly with water and soaked as usual for a minimum of 30 min in 1× binding buffer. Batches of filters can be prepared this way and stored in binding buffer for days.

Anticipated Results

Graphs of some expected results are presented in Figures 12.8.2 and 12.8.3. When precise numbers are to be derived it is best to run triplicate samples. In general, 100% binding is not achieved when titrating protein against a constant amount of DNA. Plateau can be as low as 30% or as high as 70% (the higher value is depicted in Fig. 12.8.3). However, it is the shape of the curve that is most informative. In some cases, one protein molecule is sufficient for DNA binding and retention on nitrocellulose. Any indication of sigmoidicity at low concentrations of protein may be an indication that cooperativity exists in binding or that multiple protein molecules must bind to achieve retention of a given fragment (see Ptashne, 1987).

Time Considerations

Filter-binding experiments are quick. In one day, it is easily possible to set up and filter two experiments of 10 samples each, elute the bound DNA, and load the gel to run overnight. If only scintillation counts are required, approximately four or more experiments can be run in one day.

Literature Cited

- Barkley, M.D., Riggs, A.D., Jobe, A., and Bourgeois, S. 1975. Interaction of effecting ligands with *lac* repressor and repressor-operator complex. *Biochemistry* 14:1700-1712.
- Fried, M. and Crothers, D.M. 1981. Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucl. Acids Res.* 9:6505-6525.
- Hinkle, D.C. and Chamberlin, M.J. 1972a. Studies of the binding of *Escherichia coli* RNA polymerase to DNA I. The role of sigma subunit in site selection. *J. Mol. Biol.* 70:157-185.
- Hinkle, D.C. and Chamberlin, M.J. 1972b. Studies of the binding of *Escherichia coli* RNA polymerase to DNA II. The kinetics of the binding reaction. *J. Mol. Biol.* 70:187-195.
- Hinkle, D.C. and Chamberlin, M.J. 1972c. Studies of the binding of *Escherichia coli* RNA polymerase to DNA III. Tight binding of RNA polymerase holoenzyme to single-strand breaks in T7 DNA. *J. Mol. Biol.* 70:197-207.
- Jones, O.W. and Berg, P. 1966. Studies on the binding of RNA polymerase to polynucleotides. *J. Mol. Biol.* 22:199-209.
- Lin, S.-Y. and Riggs, A.D. 1972. *Lac* repressor binding to non-operator DNA: Detailed studies and a comparison of equilibrium and rate competition methods. *J. Mol. Biol.* 72:671-690.
- Lin, S.-Y. and Riggs, A.D. 1975. The general affinity of *lac* repressor for *E. coli* DNA: Implications for gene regulation in procaryotes and eucaryotes. *Cell* 4:107-111.
- Nirenberg, M. and Leder, P. 1964. RNA codewords and protein synthesis: The effect of trinucleotides upon the binding of sRNA to ribosomes. *Science* 145:1399-1407.
- Ptashne, M. 1987. A Genetic Switch: Gene Control and Phage λ , pp. 80-83 and 109-118. Cell Press, Cambridge, MA and Blackwell Scientific, Boston, MA.
- Riggs, A.D., Suzuki, H., and Bourgeois, S. 1970. *Lac* repressor-operator interactions: I. Equilibrium studies. *J. Mol. Biol.* 48:67-83.
- Strauss, H.S., Burgess, R.R., and Record, M.T. Jr. 1980a. Binding of *Escherichia coli* ribonucleic acid polymerase holoenzyme to a bacteriophage T7 promoter-containing fragment: Selectivity exists over a wide range of solution conditions. *Biochemistry* 19:3496-3504.
- Strauss, H.S., Burgess, R.R., and Record, M.T. Jr. 1980b. Binding of *Escherichia coli* ribonucleic acid polymerase holoenzyme to a bacteriophage T7 promoter-containing fragment: Evaluation of promoter binding constants as a function of solution conditions. *Biochemistry* 19:3504-3515.
- Strauss, H.S., Boston, R.S., Record, M.T. Jr., and Burgess, R.R. 1981. Variables affecting the selectivity and efficiency of retention of DNA fragments by *E. coli* RNA polymerase in the nitrocellulose-filter-binding assay. *Gene* 13:75-87.

Key References

Ptashne, 1987. See above.

A straightforward introduction to DNA-binding interactions and equations and a good place to start for theory.

Riggs et al., 1970. See above.

An early paper covering the underlying theory and limitations of the filter-binding technique. In addition to actual data it provides detail on most applications described in this protocol, and is an excellent source of practical information.

Contributed by Ophelia Papoulas
Harvard Medical School
Boston, Massachusetts

Analysis of DNA-Protein Interactions Using Proteins Synthesized In Vitro from Cloned Genes

UNIT 12.9

BASIC
PROTOCOL

The availability of a cloned gene makes it possible to synthesize the encoded protein by in vitro transcription and translation. As described in UNIT 10.17, protein-coding sequences are cloned into a vector containing a promoter for SP6 or T7 RNA polymerase (UNIT 1.5), messenger RNA is produced by transcribing the DNA template (UNIT 3.8), and the desired protein is synthesized as a ³⁵S-labeled species by in vitro translation. Such in vitro synthesized proteins are extremely useful for determining whether a cloned gene encodes a specific DNA-binding protein and for analyzing DNA-protein interactions. To detect DNA binding activity, the labeled protein is incubated with specific DNA fragments, and protein-DNA complexes are separated from free protein by electrophoresis in native acrylamide gels (UNIT 12.2). Unlike the more conventional mobility shift assay which utilizes ³²P-labeled DNA and unlabeled protein, the assay described here generally utilizes ³⁵S-labeled protein and unlabeled DNA. Major advantages of this method are that any desired mutant protein can be tested for its DNA-binding properties simply by altering the DNA template, and the subunit structure (e.g., dimer, tetramer) can be determined.

Materials

Plasmid DNA containing desired binding sites (UNIT 1.6)
³⁵S-labeled protein (UNITS 10.17 & 10.18)
5× binding buffer
10 mg/ml poly(dI-dC)·poly(dI-dC) or other bulk carrier DNA
Loading buffer
45% methanol/10% acetic acid
EN³HANCE (Du Pont NEN)

Additional materials for digesting DNA with restriction endonucleases (UNIT 3.1), phenol extraction and ethanol precipitation (UNIT 2.1), agarose and nondenaturing polyacrylamide gel electrophoresis (UNITS 2.5, 2.7, & 12.2), and autoradiography (APPENDIX 3). For some applications, materials for end-labeling DNA with T4 polynucleotide kinase (UNIT 3.9) or Klenow fragment (UNIT 3.5) will be necessary.

1. Cleave plasmid DNA containing the binding site with appropriate restriction endonucleases, generating a fragment whose length is between 150 and 700 bp and whose molecular weight differs from other fragments generated by restriction cleavage. Purify by phenol extraction and ethanol precipitation. If desired, purify the fragment with the binding site as well as control DNA fragments followed by agarose or acrylamide gel electrophoresis.

In considering how much plasmid DNA to cleave, each DNA-binding reaction requires the equivalent of ~0.5 μg DNA assuming a plasmid of 5 kb. ³²P-labeled DNA containing the binding site can also be prepared as described in the more conventional mobility shift assay (UNIT 12.2).

2. Set up the following 15-μl binding reaction:

5 μl H₂O
3 μl 5× binding buffer
5 μl DNA (DNA fragments each at 9 nM)
1 μl 10 mg/ml poly(dI-dC)·poly(dI-dC)
1 μl [³⁵S]protein (UNIT 10.17).

Incubate 20 min at room temperature.

DNA-Protein
Interactions

12.9.1

Contributed by Kevin Struhl

Current Protocols in Molecular Biology (1993) 12.9.1-12.9.4

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 24

It is essential to perform control reactions that contain no DNA and that contain nonspecific DNA (i.e., lacking a binding site). The binding conditions—especially the amount of carrier DNA, ionic strength, divalent cations, and temperature—can be varied if necessary. It is also possible to perform binding reactions with unlabeled protein and ³²P-labeled DNA fragments as described for more conventional mobility shift assays (UNIT 12.2).

3. Add 5 µl loading buffer and immediately load on a 5% nondenaturing polyacrylamide gel. Carry out electrophoresis until the bromphenol blue is near the bottom of the gel (1 to 4 hr depending on the buffer conditions); do not heat the gel above room temperature.

The polyacrylamide gel can either be standard (30:0.8 acrylamide/bisacrylamide in 90 mM TBE buffer, pH 8.3; UNIT 2.7) or low-percentage and low-ionic-strength (UNIT 12.2). The composition of the gel and the ionic strength of the buffer can be varied and may be important for detecting a given protein-DNA interaction.

4. After electrophoresis, fix the gel in 45% methanol/10% acetic acid for 1 hr at room temperature. Treat the gel with EN³HANCE for 1 hr, dry, and analyze by autoradiography.

If the DNA-binding reaction was performed with unlabeled protein and ³²P-labeled DNA, simply dry the gel and autoradiograph as described for the conventional mobility shift assay (UNIT 12.2).

REAGENTS AND SOLUTIONS

5× binding buffer

100 mM Tris·Cl, pH 7.4
250 mM KCl
15 mM MgCl₂
5 mM EDTA
500 µg/ml gelatin

Loading buffer

1× binding buffer (see above)
20% glycerol
1 mg/ml bromphenol blue
1 mg/ml Xylene Cyanol

COMMENTARY

Background Information

The procedure described here for analyzing DNA-protein interactions was first utilized in studies on the yeast GCN4 transcriptional activator protein (Hope and Struhl, 1985). It differs from conventional biochemical approaches in that the protein of interest is not obtained from cells, but rather is synthesized by in vitro transcription and translation of a cloned gene. However, once the protein is synthesized, many standard procedures for studying specific DNA-protein interactions (e.g., UNITS 12.2, 12.3, & 12.4) can be performed with only minor modifications. In addition, the method is extremely useful for analyzing the properties of mutant proteins and for determining subunit structure, issues that are much more difficult to investi-

gate by more classical biochemical techniques.

The protocol for synthesizing [³⁵S]protein by in vitro transcription and translation is detailed in UNIT 10.17. DNA-binding activity is detected by incubating the labeled protein with appropriate DNA fragments, and separating the protein-DNA complexes from free protein by electrophoresis in native acrylamide gels. This DNA-binding assay is essentially the reverse of the standard mobility shift assay (UNIT 12.2) in which unlabeled proteins are examined for their ability to retard the mobility of a ³²P-labeled DNA fragment. By incubating the protein with a variety of DNA fragments, it is possible to examine its specific and nonspecific DNA binding properties

and to precisely localize the DNA-binding sequences. This “reverse mobility shift assay” is very convenient, and it has the advantage that the fate of the ^{35}S -labeled protein is followed directly. However, *in vitro* synthesized proteins can be tested for their DNA-binding properties by the conventional mobility shift assay using ^{32}P -labeled DNA (in which case the protein does not have to be radiolabeled) or by immunoprecipitation of protein-DNA complexes (Johnson and Herskowitz, 1985).

In interpreting the results of such reverse mobility shift assays it is necessary to consider the parameters governing gel mobilities. The mobility of a given free protein in nondenaturing gels depends upon its charge:mass ratio, a property that varies greatly among proteins and is strongly affected by pH. This means that the band corresponding to free protein can appear anywhere on the gel and its location can be strongly affected by the precise gel conditions. Thus, in order to distinguish between bands corresponding to free protein and those corresponding to protein-DNA complexes, it is crucial to perform parallel control reactions that lack DNA (since the DNA is unlabeled, its mobility in the absence of protein is irrelevant). With respect to the specific protein-DNA complex, the mobility is affected by the amount of nonspecific bulk DNA in the reaction. In the absence of carrier DNA, the complex migrates very slowly because it contains nonspecific DNA-binding proteins from the translation extract in addition to the ^{35}S -labeled protein. As the amount of bulk DNA is increased, these nonspecific DNA-binding proteins are competed off the target DNA; hence the protein-DNA complex migrates further in the gel. In addition, it is important to note that because the protein is radiolabeled, nonspecific DNA-binding complexes can be observed; for this reason, it is inadvisable to use very high concentrations of bulk carrier DNA. Finally, the mobility of a protein-DNA complex is strongly influenced by the molecular weight of the protein component; complexes with larger proteins migrate more slowly.

Perhaps the most significant advantage of this approach is that any desired mutant protein can be created simply by altering the DNA template, and then tested for its DNA-binding properties. For example, by creating a set of N- or C-terminal deletions of the protein, the DNA-binding domain can be localized (Hope and Struhl, 1986). The availability of truncated but functional proteins can be useful for determining the subunit structure of a protein (Hope

and Struhl, 1987). Specifically, the target DNA is incubated with a cosynthesized mixture of two proteins of different size obtained by carrying out the protocol on an equimolar mixture of DNA templates. As mentioned above, protein-DNA complexes involving each of the individual proteins will have a different electrophoretic mobility. If, for example, the protein binds as a dimer, the cosynthesized mixture will generate three protein-DNA complexes in a 1:2:1 molar ratio, with the complex of intermediate mobility representing a heterodimer containing the two different protein species. Proteins binding as monomers will yield two equimolar complexes, whereas proteins binding as tetramers will yield five complexes in a 1:4:6:4:1 molar ratio. The principle of using different sized proteins to determine stoichiometry is not constrained to the specific DNA-binding assay. Other methods, such as glutaraldehyde crosslinking followed by SDS-PAGE can be used to examine the protein species.

Critical Parameters

The ability to demonstrate specific DNA-binding activity using the “reverse mobility shift assay” depends on a number of factors. Of course, it is crucial that a reasonable amount of the ^{35}S -labeled protein is synthesized by *in vitro* transcription and translation (see *UNIT 10.17* for a detailed discussion). In general, it is not necessary to purify the radiolabeled protein away from the unlabeled proteins in the translation extract. As is the case with the more conventional mobility shift assay, the conditions for carrying out the binding reaction and electrophoresis can have an enormous impact (for a detailed discussion, see *UNIT 12.2*). The relative amounts of radiolabeled protein, specific DNA fragment, nonspecific DNA fragments, and bulk carrier DNA are particularly important because the presence of labeled protein means that all protein-DNA complexes are observable. If multiple DNA fragments are included in the binding reaction, their molecular weights should be sufficiently different such that the individual complexes can be distinguished. It is also important to perform electrophoresis under conditions that clearly separate the protein from the protein-DNA complex; this is generally very easy. Finally, a general consideration is that the protein synthesized *in vitro* may not be active either because its structure is different from the native protein or because a critical cofactor is missing. For such reasons, it is possible that proteins synthesized in wheat

germ extracts or reticulocyte lysates might have different properties.

Anticipated Results

The ideal protein-DNA complex gel should appear as follows. In the absence of DNA, many radiolabeled proteins will not yield a band because they migrate in the wrong direction due to their positive charge at the gel pH; other "free proteins" will generate a band. In the presence of DNA containing a high-affinity binding site, a new band indicative of a protein-DNA complex will be observed. In the presence of DNA fragments containing mutant binding sites or unrelated DNA sequences, protein-DNA complex bands may be observed, but their intensities should be significantly lower than observed for the specific binding site. Such low-intensity bands reflect nonspecific, low-affinity interactions with DNA that are characteristic of specific DNA-binding proteins. The distinction between specific and nonspecific interactions can often be enhanced by competition binding experiments in which the radiolabeled protein is incubated with a collection of DNA fragments, only one of which contains the correct target sequence. In this case, there should be one intense band corresponding to the nonspecific complexes. For experiments involving different sized derivatives of a given protein, complexes formed with truncated proteins will migrate more quickly than those with full-length proteins.

Time Considerations

It is useful to prepare the DNA fragments to

be tested for binding concurrently with the preparation of protein. Once radiolabeled protein has been synthesized and the target DNA fragments prepared, it takes about an hour to set up and perform the DNA-binding reactions, a few hours to run the non-denaturing gel, and a few more hours to carry out the steps for autoradiography. An overnight exposure is usually sufficient to visualize the results.

A typical preparation of protein can be used for 20 to 50 DNA-binding reactions (including control reactions). If a given protein is to be synthesized on multiple occasions, it is useful to prepare excess amounts of cleaved DNA template and/or mRNA.

Literature Cited

- Hope, I.A. and Struhl, K. 1985. GCN4 protein, synthesized in vitro, binds *HIS3* regulatory sequences: Implications for general control of amino acid biosynthetic genes in yeast. *Cell* 43:177-188.
- Hope, I.A. and Struhl, K. 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* 46:885-894.
- Hope, I.A. and Struhl, K. 1987. GCN4, a eukaryotic transcriptional activator protein, binds DNA as a dimer. *EMBO J.* 6:2781-2784.
- Johnson, A.D. and Herskowitz, I. 1985. A repressor (*MAT α 2* product) and its operator control a set of cell type specific genes in yeast. *Cell* 42:237-247.

Contributed by Kevin Struhl
Harvard Medical School
Boston, Massachusetts

Purification of Sequence-Specific DNA-Binding Proteins by Affinity Chromatography

UNIT 12.10

Many biological processes, such as recombination, replication, and transcription, involve the action of sequence-specific DNA-binding proteins. Analysis and purification of these proteins by conventional chromatographic methods is often difficult because DNA-binding proteins typically make up <0.01% of the total cellular protein. Various methods have been developed for purifying such proteins. Affinity chromatography is a very effective means of purifying a protein based on its sequence-specific DNA-binding properties and is relatively straightforward if proper care is taken. The affinity chromatography procedure described in this unit uses DNA containing specific recognition sites for the desired protein that has been covalently linked to a solid support. Researchers in the past have purified proteins based on the proteins' ability to bind DNA, only to be disappointed to find that the DNA-binding activity of their highly purified products was not sequence-specific. This situation can be avoided by carefully performing all of the preliminary steps and experimental controls.

The first basic protocol describes preparation of a DNA affinity resin, including cyanogen bromide (CNBr) activation of the agarose support. The alternate protocol provides a method to couple DNA to commercially available CNBr-activated Sepharose. The first support protocol describes how to purify crude synthetic oligonucleotides by gel electrophoresis prior to preparation of the affinity resin. The second basic protocol outlines the affinity chromatography procedure. The second support protocol describes determination of the appropriate type and quantity of nonspecific competitor DNA that should be used in the procedure and its preparation. Parameters essential to the success of an affinity chromatography experiment are discussed in detail in the Commentary. Figure 12.10.1 provides a summary diagram of the entire chromatography procedure.

PREPARATION OF DNA AFFINITY RESIN

Correct choice of oligonucleotide sequence (discussed in detail in the Commentary) and preparation of the affinity resin are probably the most important parts of the affinity chromatography procedure. Preparation of affinity resin can be broken down into four steps: (1) preparing oligonucleotides; (2) activating Sepharose; (3) coupling DNA to resin; and (4) blocking unreacted CNBr. The first step requires highly purified oligonucleotides that contain the recognition sequence for the desired protein. Once purified, the complementary oligonucleotides are annealed, phosphorylated with T4 polynucleotide kinase, and ligated into long, multimeric chains (averaging ≥ 10 -mers) with T4 DNA ligase. Sepharose CL-2B is activated with CNBr, the ligated DNA added, and the coupling reaction carried out overnight. Because CNBr is very toxic, researchers may prefer to employ the alternate protocol, which uses commercially available CNBr-activated Sepharose and therefore avoids direct handling of CNBr. After the coupling reaction, the remaining reactive groups are blocked with ethanolamine. This protocol is designed to prepare 10 ml of affinity resin.

NOTE: Glass-distilled or other high-quality water should be used throughout these procedures.

**BASIC
PROTOCOL**

**DNA-Protein
Interactions**

12.10.1

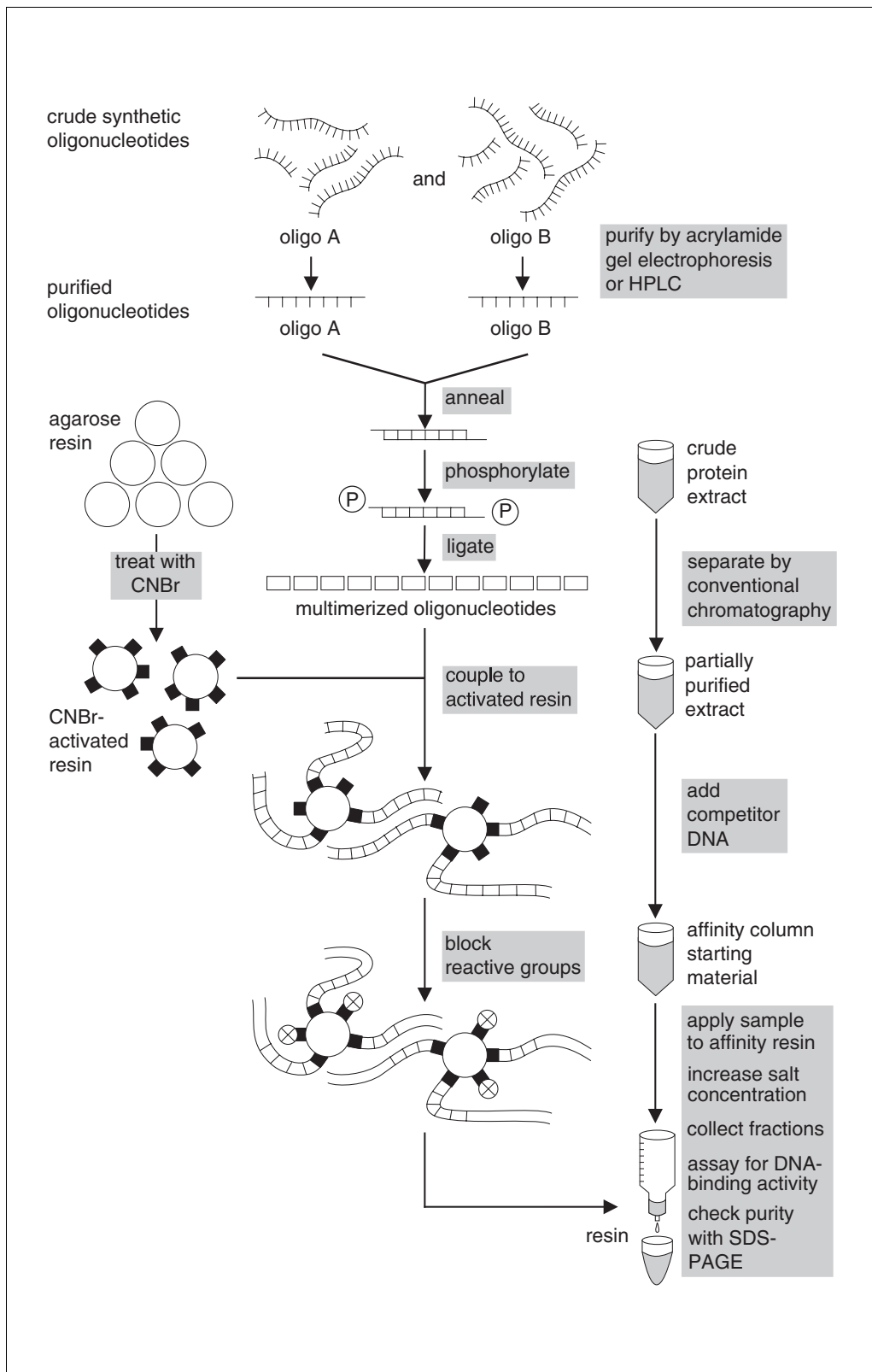


Figure 12.10.1 Purification of sequence-specific DNA-binding proteins with DNA affinity chromatography. Shown are the steps required to perform an affinity chromatography experiment using the methods described in this unit.

Materials

440 µg each of two synthetic oligonucleotides with desired binding site
(support protocol or commercial HPLC-purified)
TE buffer, pH 7.8 (*APPENDIX 2*)
10× T4 polynucleotide kinase buffer
20 mM ATP (Na⁺ salt), pH 7.0
150 mCi/ml [γ -³²P]ATP (6000 Ci/mmol)
10 U/µl T4 polynucleotide kinase (New England Biolabs; *UNIT 3.10*)
10 M ammonium acetate (*APPENDIX 2*)
25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol (*UNIT 2.1*)
24:1 (v/v) chloroform/isoamyl alcohol
3 M sodium acetate (*APPENDIX 2*)
100% and 75% ethanol
10× linker/kinase buffer
6000 U/ml T4 DNA ligase (measured in Weiss units;
New England Biolabs; *UNIT 3.14*)
Buffered phenol (*UNIT 2.1*)
Isopropanol (2-propanol)
Sephacryl CL-2B (Pharmacia Biotech)
Cyanogen bromide (CNBr; Aldrich)
N,N-dimethylformamide
5 N NaOH (*APPENDIX 2*)
10 mM and 1 M potassium phosphate buffer, pH 8.0
1 M ethanolamine hydrochloride, pH 8.0
NaOH, solid
Glycine
1 M KCl (*APPENDIX 2*)
Column storage buffer

15-ml screw-cap polypropylene tubes
Heating blocks or water baths, 15°C, 37°C, 65°C, and 88°C
60-ml coarse-sintered glass funnel
Rotating wheel

Additional reagents and equipment for agarose gel electrophoresis of
DNA (*UNIT 2.5A*) and DNA purification (*UNIT 2.1*)

5'-phosphorylate the oligonucleotides

1. In a 1.5-ml microcentrifuge tube, prepare a mixture containing 440 µg of each oligonucleotide in TE buffer in a total volume of 130 µl. Add 20 µl of 10× T4 polynucleotide kinase buffer. Incubate 2 min at 88°C, 10 min at 65°C, 10 min at 37°C, and 5 min at room temperature.

A 1-µmol synthesis of oligonucleotide should yield enough purified DNA for ~20 ml of affinity resin.

2. Divide mixture in half in separate microcentrifuge tubes. To each 75-µl aliquot, add 15 µl of 20 mM ATP (pH 7.0), ~5 µCi [γ -³²P]ATP, and 10 µl of 10 U/µl T4 polynucleotide kinase (100 U total). Incubate 2 hr at 37°C.

Dividing the reaction in half at this point facilitates the steps that follow. To add the labeled ATP, do not thaw, but simply touch (do not jab) the top of the frozen [γ -³²P]ATP with a yellow pipet tip and transfer the resulting tiny amount to the reaction tube.

3. Inactivate the kinase by adding 50 µl of 10 M ammonium acetate and 100 µl water to each tube and heating 15 min at 65°C. Allow to cool to room temperature.

Purify phosphorylated oligonucleotides

4. Add 750 μ l of 100% ethanol and mix by inversion. Microcentrifuge at high speed 15 min at room temperature to pellet DNA. Discard supernatant.
5. Resuspend each pellet in 225 μ l TE buffer.
6. Add 250 μ l of 25:24:1 phenol/chloroform/isoamyl alcohol to each tube. Vortex 1 min. Microcentrifuge at high speed 5 min to separate phases. Transfer the aqueous phase (upper layer) to a new tube.
7. Add 250 μ l of 24:1 chloroform/isoamyl alcohol to aqueous phase. Vortex 1 min. Microcentrifuge at high speed 5 min to separate phases. Transfer the aqueous phase (upper layer) to a new tube.
8. Add 25 μ l of 3 M sodium acetate to aqueous phase and mix by vortexing. Then add 750 μ l of 100% ethanol and mix by inversion. Microcentrifuge at high speed 15 min to pellet DNA. Discard supernatant.
9. Wash pellet with 800 μ l of 75% ethanol. Mix by vortexing. Microcentrifuge at high speed 5 min. Discard supernatant.
10. Dry pellet in vacuum evaporator (e.g., Speedvac).

Ligate oligonucleotides

11. Add 65 μ l water and 10 μ l of 10 \times linker/kinase buffer to each pellet. Dissolve DNA by vortexing. Add 20 μ l of 20 mM ATP (pH 7.0) and 5 μ l of 6000 U/ml T4 DNA ligase (30 Weiss units). Incubate \geq 2 hr at room temperature or overnight at 15°C.

Depending upon the oligonucleotides used, the optimal temperature for ligation will vary from 4° to 30°C. Short oligonucleotides (\leq 15-mers) tend to ligate better at lower temperatures (4° to 15°C), whereas oligonucleotides that have a moderate degree of palindromic symmetry tend to self-anneal and therefore ligate better at higher temperatures (15° to 30°C).

12. Monitor the ligation reaction by agarose gel electrophoresis, using 0.5 μ l of ligation reaction per gel lane and including lanes containing size markers. Visualize DNA by ethidium bromide staining and UV photography.

5'-phosphorylated oligonucleotides often do not ligate on the first attempt. If ligation has not occurred, extract the DNA once with 25:24:1 phenol/chloroform/isoamyl alcohol and once with 24:1 chloroform/isoamyl alcohol, then ethanol precipitate using sodium acetate as the precipitating salt. Dissolve the DNA in 225 μ l TE, add 25 μ l of 3 M sodium acetate, and reprecipitate with 750 μ l of 100% ethanol. Wash with 75% ethanol, dry in Speedvac, and repeat ligation.

The average length of the ligated oligonucleotides should be \geq 10-mers. Ligating the oligonucleotides increases the binding capacity of the affinity resin and helps avoid potential problems due to steric hindrance between the factor and the Sepharose support. However, oligonucleotide multimers $<$ 10-mers will probably also work.

Purify oligonucleotide multimers

13. Add 100 μ l buffered phenol to the 100- μ l ligation reactions. Vortex 1 min. Microcentrifuge at high speed 5 min at room temperature. Transfer aqueous phase (upper layer) to a new tube.
14. Add 100 μ l of 24:1 chloroform/isoamyl alcohol to aqueous phase. Vortex 1 min. Microcentrifuge at high speed 5 min, room temperature. Transfer aqueous phase (upper layer) to a new tube.
15. Add 33 μ l of 10 M ammonium acetate to the aqueous phase. Mix by vortexing.

16. Add 133 μ l isopropanol. Mix by inversion. Incubate 20 min at -20°C . Microcentrifuge at high speed 15 min to pellet DNA. Discard supernatant.

The ammonium acetate/isopropanol precipitation removes residual ATP, which would otherwise interfere with coupling of the ligated DNA to the CNBr-activated Sepharose.

17. Add 225 μ l TE buffer. Vortex to dissolve pellet. Add 25 μ l of 3 M sodium acetate. Mix by vortexing. Add 750 μ l of 100% ethanol. Mix by inversion. Microcentrifuge at high speed 15 min to pellet DNA. Discard supernatant.
18. Wash DNA twice with 75% ethanol. Dry pellet in vacuum evaporator.
19. Dissolve DNA in 50 μ l water. Store at -20°C .

Do not dissolve the DNA in TE buffer, as the Tris buffer in TE will interfere with the coupling reaction.

Prepare CNBr-activated Sepharose

It is best to assemble all equipment and reagents required for the activation and coupling reactions *before* proceeding with the following steps.

20. Place 10 to 15 ml (settled bed volume) of Sepharose CL-2B in a 60-ml coarse-sintered glass funnel and wash extensively with 500 ml water.

To wash, add water to resin in funnel, stir gently with a glass rod, and remove water by vacuum suction, making sure to release vacuum before the resin is suctioned into a dry cake. Repeat until all 500 ml water is used.

21. Transfer moist Sepharose resin to a 25-ml graduated cylinder, estimating 10 ml of resin. Add water to 20 ml final volume. Transfer the resulting slurry to a 150-ml glass beaker containing a magnetic stir bar. Place beaker in a water bath equilibrated to 15°C and set up over a magnetic stirrer in a fume hood. Turn on stirrer to slow medium speed.

Keep the water bath at 15°C by periodically adding small chunks of ice as needed. Occasional variation of a degree or two is not critical.

22. In the fume hood, measure 1.1 g CNBr into a 25-ml Erlenmeyer flask, keeping the mouth of the flask covered with Parafilm or a ground glass stopper as much as possible (it is better to have slightly more than 1.1 g than slightly less). Add 2 ml *N,N*-dimethylformamide (the CNBr will dissolve instantly). Over the course of 1 min, add the resulting CNBr solution dropwise to the stirring Sepharose slurry.

CAUTION: *CNBr is highly toxic and volatile. Use only in a fume hood with extreme caution. Observe appropriate decontamination and disposal procedures.*

23. Immediately add 5 N NaOH as follows: add 30 μ l to the stirring mixture every 10 sec for 10 min until 1.8 ml NaOH has been added.

It is convenient to measure 1.8 ml NaOH into a small tube before addition to the reaction to avoid having to change pipet tips during the 10-sec additions.

24. Immediately add 100 ml ice-cold water to the beaker and pour the mixture into a 60-ml coarse-sintered glass funnel.

IMPORTANT: *At this point, it is very important to avoid suction-filtering the resin into a dry cake. If the resin is accidentally dried into a cake, do not use; instead, repeat steps 20 to 24 with fresh resin.*

25. Still working in the fume hood, wash the resin in the funnel with four 100-ml washes of ice-cold ($\leq 4^{\circ}\text{C}$) water followed by two 100-ml washes of ice-cold 10 mM potassium phosphate, pH 8.0.
26. Immediately transfer the resin to a 15-ml polypropylene screw-cap tube and add ~4 ml of 10 mM potassium phosphate (pH 8.0) until the resin has the consistency of a thick slurry.

Couple oligonucleotide multimers to CNBr-Sepharose

27. Immediately add the two 50- μl aliquots of DNA from step 19. Incubate on a rotating wheel overnight (≥ 8 hr) at room temperature.
28. In the fume hood, transfer the resin to a 60-ml coarse-sintered glass funnel and wash with two 100-ml washes of water and one 100-ml wash of 1 M ethanolamine hydrochloride, pH 8.0.

Using a Geiger counter, compare the level of radioactivity in the first few milliliters of filtrate with the level of radioactivity in the washed resin to estimate the efficiency of incorporation of DNA to the resin. Usually, all detectable radioactivity is present in the resin.

29. In the fume hood, transfer the resin to a 15-ml polypropylene screw-cap tube and add 1 M ethanolamine hydrochloride (pH 8.0) until the mixture is a smooth slurry. Incubate the tube on a rotating wheel 2 to 4 hr at room temperature.

This step inactivates unreacted CNBr-activated Sepharose.

It is important to clean up all CNBr waste carefully. In the fume hood, add solid NaOH and glycine (~10 to 20 mg/ml) to inactivate the CNBr. Soak contaminated instruments in a similar solution. Let sit overnight in the fume hood, then discard.

30. Wash the resin in a 60-ml coarse-sintered glass funnel with 100 ml of 10 mM potassium phosphate (pH 8.0), 100 ml of 1 M potassium phosphate (pH 8.0), 100 ml of 1 M KCl, 100 ml water, and 100 ml column storage buffer.
31. Store the resin at 4°C (stable at least one year; do not freeze).

**ALTERNATE
PROTOCOL**

**COUPLING THE DNA TO COMMERCIALY AVAILABLE
CNBr-ACTIVATED SEPHAROSE**

The major advantage of this alternate procedure is that it begins with commercially available CNBr-activated chromatography resin, avoiding the need for preparation of CNBr-activated resin (first basic protocol). However, commercial CNBr-Sepharose is considerably more expensive than the homemade variety; moreover, the resulting column tends to run more slowly than one prepared as described above. Both resins are effective, leaving it up to the researcher's discretion which to use.

Additional Materials

1 mM HCl (*APPENDIX 2*), prepared fresh before use
CNBr-activated Sepharose 4B (Pharmacia Biotech)

1. Carry out steps 1 to 19 of the first basic protocol.
2. Weigh out 3 g CNBr-activated Sepharose 4B (1 g freeze-dried resin gives ~3.5 ml final gel volume).
3. Place the dry resin in a 15-ml conical polypropylene tube. Hydrate resin with 10 ml of 1 mM HCl and mix gently by flicking and inverting the tube. After 1 min, transfer

slurry to a 60-ml coarse-sintered glass funnel. Wash and swell the beads by gradually pouring 500 ml of 1 mM HCl through the funnel (this will take ~15 min).

4. Wash the resin with 100 ml water and then with 100 ml of 10 mM potassium phosphate, pH 8.0.
5. Proceed with steps 26 through 31 of the first basic protocol.

PURIFICATION OF OLIGONUCLEOTIDES BY PREPARATIVE GEL ELECTROPHORESIS

**SUPPORT
PROTOCOL**

Preparation of affinity resin requires a large amount (1- μ mol synthesis) of purified synthetic oligonucleotides. The oligonucleotides used to prepare the affinity resin must be of high purity because contaminating, incompletely synthesized oligonucleotides can interfere with the ligation reaction. HPLC-purified oligonucleotides are of sufficient purity for preparation of DNA affinity resins; however, it is normally expensive to obtain such oligonucleotides. An alternative is to purify crude oligonucleotides by the following method (a variation of the approach described in *UNIT 2.12*).

Additional Materials

40% (w/v) 19:1 acrylamide/bisacrylamide

10 \times TBE buffer (*APPENDIX 2*)

Urea

10% (w/v) ammonium persulfate (*UNIT 10.3*)

TEMED

Formamide loading buffer

sec-butanol (2-butanol)

Diethyl ether

1 M MgCl₂ (*APPENDIX 2*)

Saran wrap or other UV-transparent plastic wrap

Intensifying screen (e.g., Lightning Plus, DuPont NEN)

Hand-held short-wavelength UV light source

Silanized glass wool (*APPENDIX 3*)

Dry ice/ethanol bath (-78°C)

Additional reagents and equipment for purification of oligonucleotides by denaturing polyacrylamide gel electrophoresis (*UNIT 2.12*), agarose gel electrophoresis of DNA and gel photography (*UNIT 2.5A*), and butanol concentration and ether extraction of DNA (*UNIT 2.1*)

Purify oligonucleotides by gel electrophoresis

1. Prepare a 20 cm \times 40 cm \times 1.5 mm denaturing gel with four wells 3 cm in width, using 16% polyacrylamide/urea for separating oligonucleotides ~10 to ~45 bases long (or 8% or 6% polyacrylamide/urea for longer oligonucleotides). Let gel polymerize for \geq 30 min, then prerun at 30 W for \geq 1 hr.
2. Dissolve each oligonucleotide in formamide loading buffer to 200 μ l final. Heat 15 min at 65°C to remove any secondary structure in the DNA. Load 50 μ l of samples into separate wells, and run the gel at 30 W for ~4 hr.

The amount of oligonucleotide that is prepared in a 1- μ mol synthesis (~1 to 2 mg) can be loaded on one gel (0.25 μ mol per 3-cm well). This is the maximum amount that can be applied to the gel without overloading it.

It takes ~4 hr for the bromphenol blue to migrate three-quarters of the way down the gel, which is far enough to purify oligonucleotides of 15 to 30 bases. In a 16% gel, bromphenol blue comigrates with ~10-base oligonucleotides and xylene cyanol com-

**DNA-Protein
Interactions**

12.10.7

igrates with ~30-base oligonucleotides. If the oligonucleotide is 25 to 35 bases long, it is recommended that the formamide loading buffer be made without xylene cyanol.

Visualize gel by UV shadowing

3. Remove one of the glass gel plates and cover the gel with Saran wrap. Flip gel over and remove the other plate so that the gel is lying on the Saran wrap. Cover the other side of the gel with Saran wrap.
4. In a darkroom, lay the gel on an intensifying screen and hold a hand-held short-wavelength UV light source directly over it to visualize the DNA. Identify the major oligonucleotide band and mark its position directly on the Saran wrap using a marker, making sure that the light is directly over the band.

The major band (usually the largest oligonucleotide but sometimes the next-to-largest) should be visible as a thick, dark band in the gel.

Purify oligonucleotides from gel

5. Carefully cut out the band with a razor blade, trying to avoid shredding the gel material or pulverizing the gel slice into small pieces.
6. Soak the gel piece in 5 ml TE buffer in a 15-ml polypropylene tube overnight at 37°C with shaking.
7. Place a silanized glass wool plug in a Pasteur pipet and prerinse with ~5 ml water. Filter the supernatant containing the DNA through the glass wool.
8. Concentrate the DNA to $\leq 180 \mu\text{l}$ by repeated extractions with *sec*-butanol.
9. Extract the DNA once with diethyl ether and place in a vacuum evaporator (e.g., Speedvac) until all traces of ether are removed.
10. Adjust the volume of the liquid to 180 μl with TE buffer. Add 20 μl of 3 M sodium acetate and 2 μl of 1 M MgCl_2 . Mix by vortexing.
11. Add 600 μl of 100% ethanol. Mix by inversion. Chill 10 min in dry ice/ethanol. Let stand 5 min at room temperature. Microcentrifuge at high speed 15 min, room temperature, to pellet DNA. Discard supernatant.
12. Add 180 μl TE buffer. Dissolve pellet by vortexing. Add 20 μl of 3 M sodium acetate and 2 μl of 1 M MgCl_2 . Mix by vortexing.
13. Add 600 μl of 100% ethanol. Mix by inversion. Chill 10 min in dry ice/ethanol. Let stand 5 min, room temperature. Microcentrifuge at high speed 15 min, room temperature. Discard supernatant.
14. Add 800 μl of 75% ethanol. Mix by vortexing. Microcentrifuge at high speed 5 min, room temperature. Discard supernatant.
15. Carefully dry pellet in vacuum evaporator (sometimes static electricity will cause the pellet to jump out of the tube).
16. Dissolve DNA in 200 μl TE buffer and measure A_{260} and A_{280} . Store at -20°C indefinitely.

For sequence-specific DNA affinity resins, assume that 1 A_{260} unit = 40 $\mu\text{g/ml}$ DNA.

DNA AFFINITY CHROMATOGRAPHY

Affinity chromatography is performed by combining a partially purified protein sample with appropriate competitor DNA, pelleting the insoluble protein-DNA complexes by centrifugation, and loading the resulting soluble material by gravity flow onto the affinity resin. Nonspecific DNA-binding proteins flow through the column while the specific protein is retained by the column. The protein is then eluted by gradually increasing the salt concentration of the buffer and individual fractions are tested for DNA-binding activity and purity. Fractions containing appropriate DNA-binding activity can be reapplied to the affinity resin if further purification is desired. By using two sequential affinity chromatography steps, a typical protein can be purified 500- to 1000-fold with ~30% yield. The method described below is performed with buffer Z, but many other buffers will work as well. Buffer choice is addressed in the Commentary.

Materials

- Prepared DNA affinity resin (first basic or alternate protocol)
- Buffer Z or other column buffer (e.g., buffers Z^c or TM) made with varying KCl concentrations (buffers Z/0.1 M KCl through Z/1 M KCl)
- Partially purified protein fraction dialyzed against buffer Z/0.1 M KCl
- Nonspecific competitor DNA (support protocol)
- Column regeneration buffer
- Column storage buffer
- Disposable chromatography column (Poly-Prep, Bio-Rad)
- Sorvall SS-34 rotor or equivalent
- Liquid nitrogen
- Narrow glass rod, silanized (*APPENDIX 3*)
- Additional reagents and equipment for DNA-binding assays (*UNITS 12.2-12.4*), SDS-PAGE (*UNIT 10.2*), and silver staining (*UNIT 10.6*)

1. Equilibrate 1 ml settled bed volume of the DNA affinity resin in a disposable chromatography column with two 10-ml washes of buffer Z/0.1 M KCl.
2. Combine the partially purified protein fraction in buffer Z/0.1 M KCl with nonspecific competitor DNA as determined by DNA-binding studies (see Commentary and support protocols).
3. Incubate mixture 10 min on ice.
4. Centrifuge mixture 10 min at 12,000 × g (10,000 rpm in Sorvall SS-34 rotor), 4°C, to pellet insoluble protein-DNA complexes.
5. Load the supernatant onto the column at gravity flow (e.g., 15 ml/hr per column for Sepharose CL-2B).

A single 1-ml column is sufficient for a standard nuclear extract (e.g., 12 liters of HeLa cells or 150 g of Drosophila embryos). When purifying a larger quantity of material, it is preferable to use multiple 1-ml columns. It is common practice to apply as much as 50 ml of protein sample onto a single 1-ml column.

Typically, 1 ml of affinity resin contains ~80 to 90 μg DNA, which corresponds to a protein-binding capacity of 7 nmol/ml of resin, assuming one recognition site per 20 bp.

6. After loading the starting material, wash the column four times with 2-ml aliquots of buffer Z/0.1 M KCl, rinsing the sides of the column each time.

It is very important to wash the affinity column thoroughly at this step by rinsing down the sides of the column with the 2-ml wash buffer aliquots; a single 8-ml wash is not effective.

DNA affinity columns will often yield >100-fold purification of the desired factor. When a protein is purified 100-fold, however, a 1% contamination of the starting material due to inefficient washing of the column will lead to major contamination of the affinity-purified factor.

7. Elute the protein from the column by successive addition of 1-ml portions of buffer Z/0.2 M KCl, buffer Z/0.3 M KCl, buffer Z/0.4 M KCl, buffer Z/0.5 M KCl, buffer Z/0.6 M KCl, buffer Z/0.7 M KCl, buffer Z/0.8 M KCl, and buffer Z/0.9 M KCl, followed by three 1-ml aliquots of buffer Z/1 M KCl. Collect 1-ml fractions that correspond to the addition of the 1-ml portions of buffer. Quick-freeze the protein samples in liquid nitrogen and store at -80°C . Samples can generally be stored for at least 2 years.

It is convenient to save separate aliquots of each fraction for DNA-binding assays (20 μl each) and SDS-PAGE analysis (50 μl each).

8. Assay the protein fractions for the sequence-specific DNA-binding activity using a DNA-binding assay. Estimate the purity of the protein fractions by SDS-PAGE followed by silver staining to visualize the protein.

If further purification is desired, combine the fractions that contain the activity and, depending on the KCl concentration, either dilute (using buffer Z without KCl) or dialyze (against buffer Z/0.1 M KCl) to 0.1 M KCl. Then combine the protein fraction with nonspecific competitor DNA (the amount and type of which must be determined experimentally as described in the support protocol) and reapply to either fresh or regenerated DNA affinity resin.

In some instances, proteins might still be bound to the affinity resin after the 1 M salt elution step (step 7). Thus, if the desired protein is not detected in either the column fractions or flowthrough, it is possible that the protein is still on the column, and a wash at a higher salt concentration (e.g., 2 M KCl) may be needed to elute the protein from the resin.

9. Regenerate the affinity resin as follows: At room temperature, stop the column flow and add 5 ml column regeneration buffer to the column. Stir the resin with a silanized narrow glass rod to mix the resin with the regeneration buffer. Let the buffer flow out of the column. Repeat step.
10. To store column, add 10 ml column storage buffer and allow to flow through. Repeat this wash, then close the bottom of the column and add another 5 ml buffer. Cover top of column and store at 4°C .

The column may be stored for ≤ 1 year. Alternatively, the affinity resin may be removed from the column after the two washes and stored in a clean polypropylene tube with the 5 ml buffer.

SELECTION AND PREPARATION OF NONSPECIFIC COMPETITOR DNA

Proper use of competitor DNA is essential for successful purification of DNA-binding proteins by affinity chromatography. Common competitor DNAs include poly(dI-dC), poly(dA-dT), poly(dG-dC), and calf thymus DNA. The amount and type of competitor to use in a given experiment must be determined experimentally. A typical method is to mix a protein sample (the exact fraction that will be applied to the affinity resin) with varying amounts of different competitor DNAs and evaluate DNA binding using assays such as DNase I footprinting (*UNIT 12.4*) or gel mobility shifts (*UNIT 12.2*; see Commentary). The competitor DNA that inhibits DNA binding the least should be used for DNA affinity chromatography. It is perhaps easiest to think of these binding experiments as scaled-down affinity columns. First, determine the highest amount of competitor DNA that can be added to a DNA-binding reaction that does not interfere with the binding of the sequence-specific factor. To move up to a full-scale experiment, use one-fifth of the amount that would be required if the binding reaction were directly scaled up.

The following example with a hypothetical factor M illustrates how to determine the amount of competitor to use in an affinity chromatography experiment. Using DNase I footprinting, a strong footprint was observed with 5 μ l of a 0.4 M heparin fraction. Testing various competitors reveals that calf thymus DNA, poly(dI-dC), and poly(dG-dC) all strongly inhibit binding of factor M. However, no detectable inhibition of factor M binding is observed with 2 μ g poly(dA-dT), weak but detectable inhibition is observed with 3 μ g poly(dA-dT), and strong inhibition is observed with 4 μ g poly(dA-dT). In this example, with 5 μ l of the 0.4 M heparin fraction, the highest amount of poly(dA-dT) that does not inhibit factor M binding is 2 μ g. If 5 ml of the 0.4 M heparin fraction is available, the direct scale-up would be 1000-fold, so $1000 \times 2 \mu\text{g} = 2000 \mu\text{g}$ poly(dA-dT) would be needed. Because in an affinity chromatography experiment one-fifth of the direct scale-up amount is used, the appropriate amount of poly(dA-dT) to add to 5 ml of the 0.4 M heparin fraction is $2000 \mu\text{g} \times 1/5 = 400 \mu\text{g}$.

The optimal amount of competitor DNA will vary with the purity of the partially purified protein sample. Fractions that contain more nonspecific DNA-binding proteins will normally require more competitor DNA and thus it is necessary to determine experimentally the optimal amount of DNA to use with each protein fraction. In addition, the amount of competitor DNA is estimated as the mass of DNA (in micrograms) to add per volume (in milliliters) of protein fraction; this is because it is likely that competitor DNA acts by forming a complex with high-affinity, nonspecific DNA-binding proteins in the crude extract. Different competitor DNAs can be used in a single experiment.

To prepare poly(dI-dC), poly(dG-dC), and poly(dA-dT), dissolve the desired amount of competitor DNA to a final concentration of 10 A_{260} units in TE buffer per 100 mM NaCl. Heat the sample to 90°C and slowly cool to room temperature over 30 to 60 min. If the average length of the DNA is >1 kb, degrade it by sonication. Estimate the length of the DNA by agarose gel electrophoresis (*UNIT 2.5*).

REAGENTS AND SOLUTIONS

Buffer TM

50 mM Tris-Cl, pH 7.9
0 M *or* 1 M KCl
12.5 mM MgCl₂
1 mM dithiothreitol (DTT; add fresh just before use)
20% (v/v) glycerol
0.1% (v/v) Nonidet P-40 (NP-40)

Do not make a 10× buffer. To generate aliquots of buffer with the range of KCl concentrations described in the protocol, make two 500-ml batches of 1× buffer containing no KCl and 1 M KCl respectively and mix together appropriate quantities. Store at 4°C. Just prior to use, place 1.5-ml aliquots of each concentration in separate 1.5-ml microcentrifuge tubes, and add DTT.

Buffer Z

25 mM HEPES (K⁺ salt), pH 7.6
0 M *or* 1 M KCl
12.5 mM MgCl₂
1 mM DTT (add fresh just before use)
20% (v/v) glycerol
0.1% (v/v) NP-40
Adjust pH to 7.6 with KOH

Do not make a 10× buffer. To generate aliquots of buffer with the range of KCl concentrations described in the protocol, make two 500-ml batches of 1× buffer containing no KCl and 1 M KCl respectively and mix together appropriate quantities. Store at 4°C. Just prior to use, place 1.5-ml aliquots of each concentration in separate 1.5-ml microcentrifuge tubes, and add DTT.

Buffer Z^e

25 mM HEPES (K⁺ salt), pH 7.6
0 M *or* 1 M KCl
1 mM DTT (add fresh just before use)
20% (v/v) glycerol
0.1% (v/v) NP-40
Adjust the pH to 7.6 with KOH

Do not make a 10× buffer. To generate aliquots of buffer with the range of KCl concentrations described in the protocol, make two 500-ml batches of 1× buffer containing no KCl and 1 M KCl respectively and mix together appropriate quantities. Store at 4°C. Just prior to use, place 1.5-ml aliquots of each concentration in separate 1.5-ml microcentrifuge tubes, and add DTT.

Column regeneration buffer

10 mM Tris-Cl, pH 7.8
1 mM EDTA, pH 8.0
2.5 M NaCl
1% (v/v) NP-40
Store at room temperature

The solution will be cloudy and separate into two phases (NP-40 and aqueous) upon storage. Mix by swirling and shaking just before use.

Column storage buffer

10 mM Tris·Cl, pH 7.8

1 mM EDTA, pH 8.0

0.3 M NaCl

0.04% (w/v) sodium azide

Store at room temperature without sodium azide. Make a 4% (w/v) sodium azide stock solution and add just before use.

1 M ethanolamine hydrochloride, pH 8.0

1 M ethanolamine

Adjust pH to 8.0 with HCl

Filter sterilize

Store at room temperature

Formamide loading buffer

90 ml deionized formamide (*UNIT 14.3*)

10 ml 10× TBE (*APPENDIX 2*)

40 mg xylene cyanol

40 mg bromphenol blue

Store at -20°C

Linker-kinase buffer, 10×

660 mM Tris·Cl, pH 7.6

100 mM MgCl_2

100 mM DTT

10 mM spermidine

Store at -20°C

Add an extra 10 mM DTT just before use

16% polyacrylamide-urea gel

50 ml 40% (w/v) 19:1 acrylamide/bisacrylamide

12.5 ml 10× TBE (*APPENDIX 2*)

62.5 g urea

17 ml H_2O

Mix, filter, briefly degas, and then add:

750 μl 10% (w/v) ammonium persulfate

20 μl TEMED

Pour immediately into prepared gel plates.

1 M potassium phosphate buffer, pH 8.0

A: 1 M K_2HPO_4

B: 1 M KH_2PO_4

Add A to B until pH = 8.0

T4 polynucleotide kinase buffer, 10×

500 mM Tris·Cl, pH 7.6

100 mM MgCl_2

50 mM DTT

1 mM spermidine

1 mM EDTA, pH 8.0

Store at -20°C

Add an extra 50 mM DTT just before use

COMMENTARY

Background Information

Purification of sequence-specific DNA-binding proteins has historically been a difficult task, mainly because such proteins are a small fraction of the total cellular protein. However, it is now possible to purify these factors quickly, simply, and effectively by using multimerized synthetic oligonucleotides that contain the recognition sequence for a particular DNA-binding protein. Early efforts with DNA-affinity-chromatography involved adsorption or coupling of nonspecific DNA (such as calf thymus DNA) to either cellulose (Alberts and Herrick, 1971) or agarose (Arndt-Jovin et al., 1975) supports. These methods paved the way for the development of a variety of sequence-specific DNA affinity chromatography techniques, including the procedures described in this unit.

Other methods have been described to purify sequence-specific DNA-binding proteins, including chromatography using biotinylated DNA fragments attached to various supports by biotin-avidin or biotin-streptavidin coupling (UNIT 10.6; Chodosh et al., 1986; Kasher et al., 1986; Leblond-Francillard et al., 1987), oligonucleotides synthesized onto Teflon-based beads (Duncan and Cavalier, 1988), or synthetic oligonucleotide monomers attached to agarose supports (Wu et al., 1987; Blanks and McLaughlin, 1988; Hoey et al., 1993); and preparative gel mobility shifts (Gander et al., 1988). Although more than 50 sequence-specific DNA-binding proteins have been purified by the method described in this unit (Kadonaga, 1991, and references therein), it is likely that many of the techniques listed above are also effective for purifying sequence-specific DNA-binding proteins.

There are a variety of reasons that CNBr activation is commonly used in the preparation of affinity resins: it is simple, works well with agarose matrices, and is mild enough to bind ligands such as DNA. Briefly, the chemistry of the CNBr activation reaction is as follows. At high pH, the hydroxyl groups on the agarose resin react with CNBr. The majority of the CNBr added to the reaction reacts with water to yield inert cyanate ions, which is part of the reason such a large amount of CNBr is required. Additionally, the majority of the cyanate esters that are formed on the agarose either are hydrolyzed to form inert carbamate or react with the matrix hydroxyls to form imidocarbonates. The imidocarbonates that form can act effectively as chemical cross-links, thus stabilizing the

matrix (which is in most cases beneficial, particularly if the agarose resin chosen is not covalently cross-linked). The remaining active cyanate esters are coupled to the amino-containing ligands (in this case, oligonucleotides) at physiological pH. Finally, the unreacted cyanate esters are blocked with an excess of a suitable reagent, such as ethanolamine, to prevent coupling of the protein sample to the matrix (Janson and Rydén, 1989).

Critical Parameters and Troubleshooting

Basic strategy

The general approach to affinity purification of DNA-binding proteins is as follows. First, estimate the binding site of the desired protein by a technique such as DNase I footprinting (UNIT 12.4). If possible, it is best to survey a variety of promoters and enhancers. Next, determine the optimal conditions for protein binding to the DNA, considering factors such as temperature, ionic strength, pH, and Mg^{2+} concentration. Using conventional chromatography, partially purify the protein to remove contaminants such as nucleases that may degrade the affinity resin. Test a variety of nonspecific competitor DNAs, including poly(dI-dC), poly(dG-dC), poly(dA-dT), and calf thymus DNA to determine their effect on protein-DNA interactions. For best results, prepare two or more different DNA affinity resins with naturally occurring, high-affinity binding sites, preferably containing different flanking DNA sequences. The desired protein should bind with high affinity to both resins; if it does not, then it is likely that the protein that has been purified is not specific to the desired binding site. Finally, it is important to prepare a control resin that does not contain the recognition sequence for the desired protein. A control resin will make it possible to identify proteins that bind nonspecifically to DNA-Sepharose. Using this approach, it is possible to obtain a preparation containing a highly purified, sequence-specific DNA-binding protein.

Starting material and conventional chromatography

Extensive purification of the DNA-binding protein is not required for effective affinity chromatography. However, partial purification of the protein by conventional chromatography

is recommended prior to affinity chromatography to remove proteases and nucleases that might degrade either the protein or the affinity resin. Keep in mind that sequence-specific and nonspecific DNA-binding proteins will often co-purify in conventional chromatography, possibly because they may both interact with negatively charged polymers that resemble DNA. It is important to note that the most persistent contaminants in sequence-specific DNA-binding protein preparations are nonspecific, high-affinity DNA-binding proteins. Various conventional chromatographic methods may be employed prior to affinity purification of a DNA-binding protein; methods that have been used successfully in the past are discussed below.

Ion-exchange chromatography. One obvious first step might be cation-exchange chromatography with resins such as S-Sepharose Fast Flow/Mono S (Pharmacia Biotech), CM-52 (Whatman), CM Sepharose Fast Flow (Pharmacia Biotech), P11 phosphocellulose (Whatman), or Bio-Rex 70 (Bio-Rad). DNA-binding proteins (both specific and nonspecific) will normally bind to cation-exchange resins in buffers containing from 50 mM to 100 mM NaCl and can be eluted with higher salt concentrations.

Affinity chromatography. Nonspecific DNA-cellulose (Alberts and Herrick, 1971) or DNA-agarose (Arndt-Jovin et al., 1975) resins can be used for a preliminary purification step. These resins are typically prepared with salmon sperm or calf thymus DNA. Also, sequence-specific DNA affinity resins can be prepared with oligonucleotides that do not contain binding sites for the desired factor. A strategy for chromatography with either nonspecific DNA affinity resin (e.g., see Rosenfeld and Kelly, 1986) or sequence-specific DNA affinity resins that lack the binding site for the desired factor (e.g., see Kaufman et al., 1989) is as follows. Once the protein sample is applied to a nonspecific DNA affinity resin, the desired factor will either (1) flow through the resin or (2) elute from the resin at a low-salt concentration. High-affinity, nonspecific DNA-binding proteins should remain bound to the resin and will thus be separated from the desired protein. If this strategy is successful, it will not be necessary to add the competitor DNA prior to the subsequent sequence-specific affinity chromatography step because the high-affinity nonspecific DNA-binding proteins will already have been separated from the sample.

A resin that works well if the desired protein contains O-linked *N*-acetylglucosamine monosaccharide residues is wheat germ agglutinin-agarose. This method has been used successfully in the past with transcription factors such as Sp1 (Jackson and Tjian, 1989) and HNF1 (Lichtsteiner and Schibler, 1989).

Heparin-agarose chromatography. Heparin-agarose resins are also excellent for the purification of DNA-binding proteins because they possess properties similar to those of both ion-exchange and affinity resins. It is important to be aware that variability exists between different heparin-agarose preparations. This variability may be due to differences in the methods of coupling heparin to agarose, because alternate methods for coupling and blocking the agarose may produce different functional groups on the resin. If a particular batch of heparin-agarose is used successfully, it is wise to continue to use that batch to ensure reproducibility.

Gel-filtration chromatography. Finally, gel-filtration chromatography may be used effectively for partial purification of a DNA-binding protein because separation is based on size and shape, not DNA-binding properties. For example, gel filtration may be useful for separating a desired protein from contaminating nucleases and other nonspecific DNA-binding proteins, in contrast to chromatography using ion-exchange, heparin-agarose, and nonspecific DNA affinity resins that may actually enrich for DNA-binding proteins. In some cases gel filtration may be undesirable because significant sample dilution occurs; however, further concentration can be obtained by subsequent affinity chromatography.

DNA binding studies

Binding conditions. In order to use affinity chromatography to purify a DNA-binding protein, it is necessary to optimize binding conditions. It is important to determine these conditions experimentally, as ionic strength, temperature, pH, and presence or absence of Mg²⁺ can affect protein-DNA interactions. In addition, although HEPES and PIPES are often considered to be superior to Tris buffers, some commercially available preparations of HEPES and PIPES buffers contain contaminants that inhibit binding of proteins to DNA. Therefore, in some respects, it may be safer to use Tris buffer. It is also important to note that some factors, such as Sp1, bind with higher affinity to DNA in the absence of Mg²⁺ than in its

presence. Thus, in some cases it may be useful to omit Mg^{2+} in the chromatography buffer. In addition, although the procedures in this unit use buffers containing KCl, NaCl could probably be substituted. Although it is usually preferable to handle proteins at $4^{\circ}C$, proteins have been described that will not bind DNA at $4^{\circ}C$, yet bind perfectly well at room temperature; in such cases the temperature at which the affinity chromatography is performed should be adjusted accordingly. Finally, to minimize non-specific adsorption of affinity-purified proteins to plastic and glass, chromatography and storage buffers should contain a nonionic detergent such as NP-40 at a concentration of $\sim 0.01\%$ to 0.1% (v/v).

Binding assays. There are many methods for identifying proteins that are bound to DNA, including DNase I footprinting (Galas and Schmitz, 1978; *UNIT 12.4*), Fe-EDTA footprinting (Tullius et al., 1987), methidiumpropyl-EDTA- Fe^{2+} (MPE; Hertzberg and Dervan, 1982), methylation interference (*UNIT 12.3*), and gel mobility shift assays (*UNIT 12.2*). Although binding information may be obtained from all these techniques, it is most straightforward to use DNase I footprinting to identify a protein binding site. Additionally, DNase I is active over a broad range of buffer conditions, which allows for considerable variation in binding conditions. Visualization of the protected region by footprinting makes designing oligonucleotides for affinity chromatography relatively simple (as discussed later in this section). Both DNase I footprinting and gel mobility shift assays are useful for assaying affinity-column fractions.

Design of oligonucleotides

Many different parameters are important in designing oligonucleotides for an affinity resin. The first step, as described above, is to determine the protein-binding site on the DNA. The DNA affinity resin should be prepared from naturally occurring, high-affinity binding sites. Using a consensus sequence may work in some instances, but in most situations, a "real" binding site is best. Given that the length of the oligonucleotides can be reasonably flexible (14-mers to 61-mers have been used successfully), use of oligonucleotides that include the DNase I footprint as well as some flanking DNA (at least a few bases beyond the borders of the footprint) is recommended. In the early stages, it is better to copurify factors that bind to the flanking region than to fail to purify the desired factor because the binding site is too

short. It may be wise to avoid using oligonucleotides that are 21 or 42 bases in length because the DNA, if bent, might have a greater tendency to circularize during the ligation step. The oligonucleotides should also be designed with a single-stranded overhang, for two reasons: first, the ligation reaction will be more efficient (as compared to a blunt-end ligation), and second, it is likely that the DNA couples to the resin via the primary amine groups of the single-stranded overhang of the ligated multimers. The sequence GATC-XXX-BINDING SITE-XXX (where XXX represents the DNA flanking the binding site) works well, though it may also be convenient to use the sequences flanking the binding site as the overhang.

Inadvertent purification of nonspecific DNA-binding proteins

Throughout this unit, an attempt has been made to emphasize how to avoid inadvertent purification of nonspecific DNA-binding proteins. Unfortunately, this problem is quite common, particularly when the gel mobility shift assay is used to monitor DNA binding. Proper use of the control resin (to identify proteins that bind nonspecifically to DNA-Sepharose), two different affinity resins (to identify proteins that might be interacting with the linker DNA), and competitor DNA (to deplete nonspecific DNA-binding proteins from the extract) should virtually eliminate purification of nonspecific DNA-binding proteins. In addition, DNase I footprinting is recommended for assaying DNA binding because it allows visualization of the actual protein binding site and to confirm sequence specificity. To correlate DNA binding with a particular polypeptide, it would be informative to purify the polypeptide from a polyacrylamide gel by using the denaturation/renaturation procedure described by Hager and Burgess (1980). Two common nonspecific DNA-binding proteins that have been purified from HeLa cells by affinity chromatography are poly(ADP-ribose) polymerase, which has an M_r of 116,100 (Ueda and Hayaishi, 1985; Slattery et al., 1983), and the Ku antigen, which consists of two polypeptides of M_r 70,000 and 80,000 (Mimori et al., 1986).

Anticipated Results

After two rounds of affinity chromatography, it is possible to achieve a 500- to 1000-fold purification with a 30% overall yield. The concentration of affinity-purified proteins is typically 5 to 50 $\mu g/ml$. About 0.2 to 1 μl of protein (2 to 10 ng) can be expected to be sufficient for

a “blanked-out” footprint (with ~10 fmol of DNA probe). Affinity resin is stable for >1 year if stored at 4°C. Affinity-purified proteins are typically stable for >2 years if stored at –80°C in an appropriate buffer and handled properly. Proteins should always be quick-frozen in liquid nitrogen and quick-thawed in either cold or room-temperature water (although if room-temperature water is used to accelerate thawing, the sample must not be allowed to warm up above 4°C). Although many DNA-binding proteins remain active after several freeze-thaw cycles, it is preferable to divide protein preparations into small aliquots rather than to freeze and thaw the entire sample many times.

Time Considerations

The time required to optimize DNA-binding conditions of a particular protein is variable and depends on assay type, probe preparation time, and competitor choices. The time to prepare extracts and to perform conventional chromatography is also variable depending on factors such as starting material and type of chromatography. Beginning with purified oligonucleotides, preparation of the DNA affinity resin takes 2 days (with either commercially available or homemade CNBr-activated resin). If the oligonucleotides are to be gel-purified, an additional day and a half should be added. Affinity chromatography takes ~4 to 8 hr, dependent mainly on the flow rate of the resin: Sepharose CL-2B runs at ~15 ml/hr, while Sepharose-4B runs somewhat slower. Therefore, it is possible to estimate chromatography time by calculating the volume of sample that is to be applied to the column plus 8 ml for the wash steps and 11 ml for the elution steps. If the volume of starting material is not too large, it is possible to assay the affinity column on the day it is run (either by DNA-binding assays, SDS-PAGE, or both). It is important to freeze the affinity fractions in liquid nitrogen and store at –80°C as soon as the column is complete.

Literature Cited

- Alberts, B. and Herrick, G. 1971. DNA-cellulose chromatography. *Methods Enzymol.* 21:198-217.
- Arndt-Jovin, D.J., Jovin, T.M., Bähr, W., Frischauf, A.-M., and Marquardt, M. 1975. Covalent attachment of DNA to agarose: Improved synthesis and use in affinity chromatography. *Eur. J. Biochem.* 54:411-418.
- Blanks, R. and McLaughlin, L.W. 1988. An oligodeoxynucleotide affinity column for the isolation of sequence specific DNA binding proteins. *Nucl. Acids Res.* 16:10283-10299.
- Chodosh, L.A., Carthew, R.W., and Sharp, P.A. 1986. A single polypeptide possesses the binding and transcription activities of the *Adenovirus* major late transcription factor. *Mol. Cell. Biol.* 6:4723-4733.
- Duncan, C.H. and Cavalier, S.L. 1988. Affinity chromatography of a sequence-specific DNA binding protein using Teflon-linked oligonucleotides. *Anal. Biochem.* 169:104-108.
- Galas, D. and Schmitz, A. 1978. DNase footprinting: A simple method for the detection of protein-DNA binding specificity. *Nucl. Acids Res.* 5:3157-3170.
- Gander, I., Foeckler, R., Rogge, L., Meisterernst, M., Schneider, R., Mertz, R., Lottspeich, F., and Winnacker, E.L. 1988. Purification methods for the sequence-specific DNA-binding protein nuclear factor I (NFI)—generation of protein sequence information. *Biochim. Biophys. Acta* 951:411-418.
- Hager, D.A. and Burgess, R.R. 1980. Elution of proteins from sodium dodecyl sulfate-polyacrylamide gels, removal of sodium dodecyl sulfate, and renaturation of enzymatic activity: Results with sigma subunit of *Escherichia coli* RNA polymerase, wheat germ DNA topoisomerase, and other enzymes. *Anal. Biochem.* 109:76-86.
- Hertzberg, R.P. and Dervan, P.B. 1982. Cleavage of double helical DNA by (methidiumpropyl-EDTA)iron(II). *J. Am. Chem. Soc.* 104:313-315.
- Hoey, T., Welnzierl, R.O., Gill, G., Chen, J.L., Dynlacht, B.D., and Tjian, R. 1993. Molecular cloning and functional analysis of *Drosophila* TAF110 reveal properties expected of coactivators. *Cell* 72:247-260.
- Jackson, S.P. and Tjian, R. 1989. Purification and analysis of RNA polymerase II transcription factors using wheat germ agglutinin affinity chromatography. *Proc. Natl. Acad. Sci. U.S.A.* 86:1781-1785.
- Janson, J.-C. and Rydén, L. 1989. Protein Purification. Principles, High Resolution Methods, and Applications. VCH Publishers, New York.
- Kadonaga, J.T. 1991. Purification of sequence-specific DNA binding proteins by DNA affinity chromatography. *Methods Enzymol.* 208:10-23.
- Kasher, M.S., Pintel, D., and Ward, D.C. 1986. Rapid enrichment of HeLa transcription factors IIIB and IIIC by using affinity chromatography based on avidin-biotin interactions. *Mol. Cell. Biol.* 6:3117-3127.
- Kaufman, P.D., Doll, R.F., and Rio, D.S. 1989. *Drosophila* P element transposase recognizes internal P element DNA sequences. *Cell* 59:359-371.
- Leblond-Francillard, M., Dreyfus, M., and Rougeon, F. 1987. Isolation of DNA-protein complexes based on streptavidin and biotin interaction. *Eur. J. Biochem.* 166:351-355.
- Lichtsteiner, S. and Schibler, U. 1989. A glycosylated liver-specific transcription factor stimulates transcription of the albumin gene. *Cell* 57:1179-1187.

- Mimori, T., Hardin, J.A., and Steitz, J.A. 1986. Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. *J. Biol. Chem.* 261:2274-2278.
- Rosenfeld, P.J. and Kelly, T.J. 1986. Purification of nuclear factor I by DNA recognition site affinity chromatography. *J. Biol. Chem.* 261:1398-1408.
- Slattery, E., Dignam, J.D., Matsui, T., and Roeder, R.G. 1983. Purification and analysis of a factor which suppresses nick-induced transcription by RNA polymerase II and its identity with poly(ADP-ribose) polymerase. *J. Biol. Chem.* 258:5955-5959.
- Tullius, T.D., Dombroski, B.A., Churchill, M.E.A., and Kam, L. 1987. Hydroxyl radical footprinting: A high resolution method for mapping protein-DNA contacts. *Methods Enzymol.* 155:537-558.
- Ueda, K. and Hayaishi, O. 1985. ADP-ribosylation. *Annu. Rev. Biochem.* 54:73-100.
- Wu, C., Wilson, S., Walker, B., Dawid, I., Paisley, T., Zimarino, V., and Ueda, H. 1987. Purification and properties of *Drosophila* heat shock activator protein. *Science (Wash. DC)* 238:1247-1253.

Key References

Kadonaga, J.T. 1991. See above.

Techniques paper, though less descriptive than this unit, containing a table that lists (with references) >50 sequence-specific proteins that have been purified using the affinity chromatography method described herein.

Kadonaga, J.T. and Tjian, R. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* 83:5889-5893.

First paper to describe affinity chromatography with multimerized oligonucleotides; details purification of transcription factor Sp1.

Contributed by Leslie A. Kerrigan
and James T. Kadonaga
University of California San Diego
La Jolla, California

Determination of Protein-DNA Sequence Specificity by PCR-Assisted Binding-Site Selection

Binding-site selection is used to determine the target specificity of a sequence-specific DNA-binding protein. The technique has a number of applications, ranging from identifying DNA target sequences for proteins with unknown DNA-binding specificities to providing additional information on the protein-DNA interactions of previously characterized DNA binding domains.

The procedure is outlined in Figure 12.11.1. As indicated in the Basic Protocol, a pool of random-sequence oligonucleotides is used as the source of potential binding sites (see UNIT 2.11 for a discussion of the synthesis of randomized oligonucleotides). The oligonucleotide pool is incubated in a binding reaction with extract containing the DNA-binding protein of interest. Protein-DNA complexes are isolated by immunoprecipitation with an antibody specific for the protein under investigation, and unbound oligonucleotides are removed by gentle washing. Bound oligonucleotides are recovered, amplified by the polymerase chain reaction (PCR), and used as input DNA for a further round of binding,

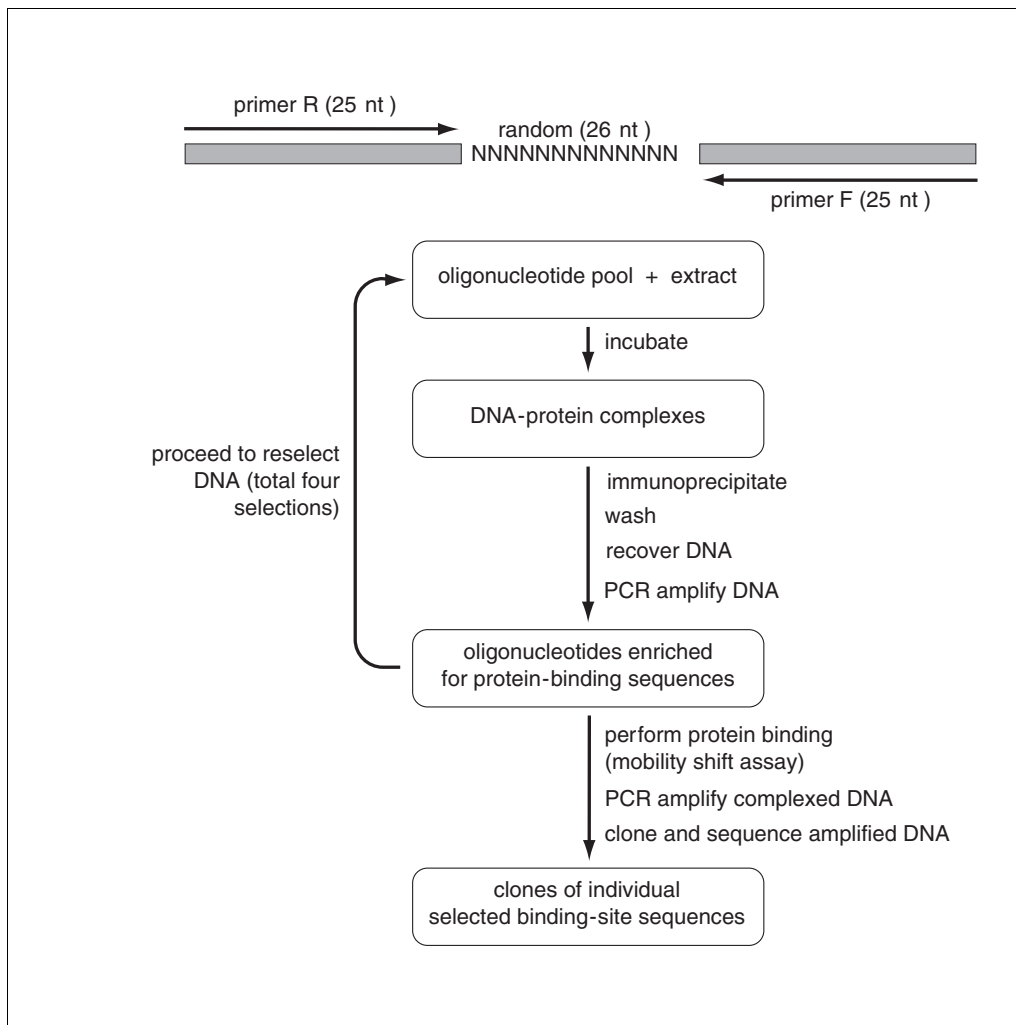


Figure 12.11.1 Flow chart for PCR-assisted DNA binding-site selection.

recovery, and amplification. After four rounds of selection, progress of the procedure is monitored by mobility shift analysis of the selected oligonucleotide pools. In the Support Protocol, individual binding sites are isolated from the appropriate complex on a mobility shift gel, cloned into plasmids, and examined by sequencing.

Materials

Random-sequence oligonucleotide R76:

5'-CAGGTCAGTTCAGCGGATCCTGTCTG(A/G/C/T)₂₆GAGGCGAATTCAGTG
CAACTGCAGC-3' (see UNIT 2.11)

Primer F: 5'-GCTGCAGTTGCACTGAATTCGCCTC-3'

Primer R: 5'-CAGGTCAGTTCAGCGGATCCTGTCTG-3'

10× *Taq* DNA polymerase buffer (UNIT 3.4)

0.5 mM 3dNTP mix (minus dCTP; UNIT 3.4)

40 μM and 0.5 mM dCTP (UNIT 3.4)

10 mCi/ml [α -³²P]dCTP (800 Ci/mmol)

5 U/μl *Taq* DNA polymerase (UNIT 3.5)

Elution buffer (see recipe)

Glycogen carrier (e.g., Boehringer Mannheim)

TE buffer, pH 7.5 to 8.0 (APPENDIX 2)

Protein A–Sepharose CL-4B (Pharmacia Biotech)

Wash buffer (see recipe), with and without 50 μg/ml BSA

Binding buffer (see recipe), with and without 50 μg/ml BSA

Carrier DNA: e.g., 100 ng/μl poly(dI-dC)·poly(dI-dC)

Protein sample: reticulocyte lysate or nuclear extract containing DNA-binding protein of interest, *or* purified DNA-binding protein (if available)

Antiserum to protein of interest, *or* (if protein is epitope-tagged and monoclonal antibody is available) appropriate ascites fluid

Recovery buffer (see recipe)

2.5 M sodium acetate

5 M ammonium acetate

Scintillation vials and counter

17 × 100–mm polystyrene centrifuge tube with snap-cap

Tumbler or rotating wheel

Whatman 3MM paper

Additional reagents and equipment for purification of oligonucleotides using denaturing (UNIT 2.12) and nondenaturing PAGE (UNIT 2.7), autoradiography (APPENDIX 3A), phenol/chloroform extraction and ethanol precipitation (UNIT 2.1), and mobility shift DNA-binding assay (UNIT 12.2)

Prepare labeled double-stranded random-sequence oligonucleotide

1. Purify oligonucleotides R76, primer F, and primer R by denaturing PAGE, then extract from gel (UNIT 2.12). Dilute R76 to 50 ng/μl and primers F and R to 80 ng/μl in water.
2. Set up the following reaction in a 0.5-ml microcentrifuge tube (20 μl total):
 - 2 μl 50 ng/μl oligonucleotide R76
 - 2 μl 10× *Taq* DNA polymerase buffer
 - 2 μl 0.5 mM 3dNTP mix (minus C)
 - 2 μl 40 μM dCTP
 - 2 μl 80 ng/μl primer F
 - 2 μl 10 Ci/ml (800 Ci/mmol) [α -³²P]dCTP
 - 7 μl H₂O
 - 1 μl 5 U/μl *Taq* DNA polymerase.

3. Carry out one PCR cycle using the following cycling parameters:

1 min 94°C (denaturation)
3 min 62°C (annealing)
9 min 72°C (extension)

Chase extension by adding 2 μ l of 0.5 mM dCTP, then heating 9 min at 72°C.

This reaction will generate double-stranded R76 oligonucleotide (dsR76) labeled to a specific activity of 3200 Ci/mmol, which is appropriate for use in a gel mobility shift assay (step 29).

Probes with higher specific activity can be generated by increasing the ratio of labeled to unlabeled dCTP in the reaction. The reaction volume can be cut to 10 μ l to economize on label.

*R76 may also be rendered double-stranded by annealing primer F and extending with Klenow fragment of *E. coli* DNA polymerase I (UNIT 3.5).*

4. Purify dsR76 on an 8% nondenaturing polyacrylamide gel (UNIT 2.7), then visualize by autoradiography (APPENDIX 3A).

A clearly visible dsR76 band can be obtained after a 60- to 90-sec exposure.

5. Excise gel slice containing labeled dsR76 using a clean scalpel and place in a 1.5-ml microcentrifuge tube containing 250 μ l elution buffer. Incubate overnight at 37°C.
6. Remove elution buffer to a fresh 1.5-ml microcentrifuge tube, add 1 μ g of glycogen carrier, and ethanol precipitate.
7. Resuspend pellet in 10 μ l TE buffer and measure 1 μ l into a scintillation vial. Measure Cerenkov counts in a scintillation counter to determine cpm.

Because the molecular weight and specific activity of the probe oligonucleotide are known, the amount of dsR76 can be quantitated by assuming that 10^6 cpm is approximately equivalent to 1 μ Ci. The double-stranding reaction in step 2 and the PCR amplification reaction in step 22 contain [α - 32 P]dCTP and unlabeled dCTP at concentrations such that four labeled C nucleotides are incorporated into every dsR76 oligonucleotide. The specific activity of dsR76 is therefore four times the specific activity of the [α - 32 P]dCTP itself (4×800 Ci/mmol = 3200 Ci/mmol).

8. Prepare a 0.2 ng/ μ l dilution of dsR76 in TE buffer for use in binding reaction (step 12).

Prepare protein A–Sepharose bead slurry

9. Rehydrate and wash freeze-dried Protein A–Sepharose CL-4B beads in water according to manufacturer's instructions.

Each binding reaction requires only 10 μ l packed volume of beads. There should be no need to prepare more than 500 μ l at a time (equivalent to ~0.125 g dried beads).

10. Wash swollen beads twice in 50 vol wash buffer without BSA.
11. Prepare a 50% (v/v) slurry in wash buffer containing 50 μ g/ml BSA and allow beads to equilibrate 2 to 3 hr at 4°C.

The 50% protein A–Sepharose slurry can be stored for several weeks at 4°C.

Set up binding reaction

12. In a 1.5-ml microcentrifuge tube on ice, prepare the following reaction:

20 μ l binding buffer containing 50 μ g/ml BSA
2 μ l 100 ng/ μ l poly(dI-dC)-poly(dI-dC)
1 to 2 μ l protein sample
2 μ l 0.2 ng/ μ l dsR76 probe oligonucleotide (from step 8)
1 μ l antiserum.

For the first round of selection use 0.4 ng probe per reaction; for subsequent rounds use 0.2 ng (dilute probe solution to 0.1 ng/ μ l in TE buffer before adding to reaction).

Antiserum should be specific for the DNA-binding protein of interest, and should immunoprecipitate protein-DNA complexes without disrupting protein-DNA interactions (see Critical Parameters). The exact quantity to be used must be optimized empirically. For proteins tagged with an epitope to which a monoclonal antibody is available, use 0.1 μ l ascites fluid (~0.1 to 1.0 μ g antibody) diluted $1/10$ with wash buffer.

If available, purified DNA-binding proteins can be used instead of extracts, although the concentration in the binding reaction should be kept low (see Critical Parameters).

13. Allow protein-DNA complexes to form on ice for 20 to 30 min.

Immunoprecipitate and wash

14. Transfer 20 μ l of 50% protein A–Sepharose slurry (step 11) to a 1.5-ml microcentrifuge tube containing 250 μ l 4°C wash buffer without BSA. Pellet protein A–Sepharose by microcentrifuging 15 sec at maximum speed, room temperature, and aspirate off wash buffer.

Loss of the protein A–Sepharose beads themselves is minimized by performing the aspiration through a 200- μ l flat tip.

15. Pipet binding reaction (step 13) into the 1.5-ml microcentrifuge tube containing 10 μ l packed volume protein A–Sepharose. Mix beads and binding reaction by pipetting up and down gently.
16. Insert the microcentrifuge tube into the mouth of a 17 \times 100–mm polystyrene tube and snap on cap. Place on a tumbler or rotating wheel and tumble or rotate overnight in a cold room.

Placing the microcentrifuge tube in the mouth of the larger tube immobilizes it and allows for convenient mixing.

IMPORTANT NOTE: *End-over-end mixing should be avoided, as protein A–Sepharose beads will smear along the side of the microcentrifuge tube.*

17. Add 250 μ l 4°C binding buffer without BSA to the binding reaction and mix rapidly by brief vortexing followed by inversion of the microcentrifuge tube. Repeat vortexing and inversion twice, then microcentrifuge 15 sec at maximum speed, room temperature, and aspirate off supernatant. Repeat this wash twice more.

These washes should be performed as fast as is reasonably possible, because the time spent performing them can be an important factor in determining the stringency of the selections (see Critical Parameters).

Recover DNA

18. Elute bound DNA from the protein A–Sepharose bead pellet by resuspending in 200 μ l recovery buffer and incubating 1 hr at 45°C.
19. Perform a phenol extraction followed by a chloroform extraction (*UNIT 2.1*).

20. Add 1 μg glycogen carrier and 12 μl of 2.5 M sodium acetate, then recover DNA by ethanol precipitation (UNIT 2.1).
21. Measure pellet for Cerenkov counts in a scintillation counter to determine cpm. Measure 0.2 ng of input probe.

Determination of the proportion of the input probe recovered allows quantitation of selected DNA (i.e., determination of the proportion of input probe recovered as selected DNA). In addition, when background recoveries (i.e., recoveries obtained with negative control extracts) are low, the proportion of input DNA recovered after each round of selection often gives a good indication of the progress of the site selection.

Amplify selected DNA

22. Prepare the following reaction mixture (20 μl total):

2 μl 10 \times *Taq* DNA polymerase buffer
3.2 μl 0.5 mM 3dNTP mix (minus C)
2 μl 40 μM dCTP
2 μl 80 ng/ μl primer F
2 μl 80 ng/ μl primer R
1 μl 10 mCi/ml (800 Ci/mmol) [α - ^{32}P]dCTP
7.3 μl H₂O
0.5 μl 5 U/ μl *Taq* DNA polymerase.

Add this mixture to 1 pg of selected DNA in a 0.5-ml microcentrifuge tube.

23. Carry out 15 of the following PCR amplification cycles:

1 min 94°C (denaturation)
1 min 62°C (annealing)
1 min 72°C (extension)

This step is optimized for amplification of oligonucleotide R76 with primers F and R. It is important to carefully calibrate the amplification reaction with respect to the amount of input DNA and number of PCR cycles performed (see Critical Parameters), particularly when using different random oligonucleotide/primer combinations.

24. Dilute PCR reaction to 150 μl with TE buffer and phenol extract (UNIT 2.1).
25. Add 1 μg of glycogen carrier and $\frac{1}{4}$ vol of 5 M ammonium acetate (to 1 M). Ethanol precipitate (UNIT 2.1).
26. Purify PCR product by electrophoresis on a nondenaturing polyacrylamide gel as described in steps 4 to 6.

Because maximum recovery is not absolutely essential, the elution step can be cut to 2 hr at 45°C.

27. Following ethanol precipitation, resuspend pellet in 10 μl TE buffer and measure 1 μl for Cerenkov counts. Quantitate selected, amplified dsR76, then resuspend at 0.1 ng/ μl .

This procedure typically yields 10 to 20 ng labeled amplified selected dsR76 oligonucleotide.

28. Use the selected and amplified oligonucleotide pool in another round of binding-site selection (steps 12 to 27). After four selection cycles, proceed to step 29 (mobility shift assay) to monitor the success of the binding-site selection procedure.

If successful, four rounds of selection should yield oligonucleotide pools that are capable of forming visible protein-DNA complexes on a mobility shift gel (see Critical Parameters).

Analyze selected oligonucleotide pools by mobility shift assay

29. Set up a series of binding reactions in which the selecting extract is incubated with probes representing oligonucleotide pools from each round of selection. The binding reactions should be identical to that described in step 12 except with the antibody omitted.

Controls are critical, and should include binding reactions in which oligonucleotide pools selected by the test protein are incubated with extracts that do not contain the protein under investigation (see Troubleshooting). They should also include a series of binding reactions representing parallel selections performed with a positive control extract.

30. Incubate 20 to 30 min at room temperature, then load reactions onto a mobility shift gel (UNIT 12.2).

For each selecting extract, binding reactions with probes derived from consecutive rounds of selection should be loaded adjacent to each other on the mobility shift gel. This allows the gradual appearance of specific complexes to be easily visualized.

31. Run gel according to the procedure described for mobility shift gels (UNIT 12.2).

32. Dry unfixed gel onto Whatman 3MM paper, then perform autoradiography overnight (APPENDIX 3A).

If the selections have been successful, mobility shift complexes should begin to appear in lanes representing later rounds of selection.

SUPPORT PROTOCOL

ISOLATION AND ANALYSIS OF BOUND OLIGONUCLEOTIDES FROM MOBILITY SHIFT GELS

Once the success of the site selection procedure has been confirmed by mobility shift analysis, bound oligonucleotides are isolated from appropriate mobility shift complexes by direct amplification from dried gel slices. Bound oligonucleotides are then digested with the appropriate restriction enzymes and subcloned for sequence analysis of individual selected sites.

Additional Materials (also see Basic Protocol)

EcoRI and BamHI restriction endonucleases and appropriate buffer (UNIT 3.1)

1. Cut the region containing the protein-DNA complex from a dried mobility shift gel using a clean scalpel. Divide into four squares.
2. For two of the squares, use a scalpel to separate the dried polyacrylamide gel from the Whatman 3MM paper. Place both squares in a 0.5-ml microcentrifuge tube.

Save the other two squares in case amplification needs to be repeated (see step 4 annotation).

3. Add the following mix to the dried acrylamide (50 μ l total):

5 μ l 10 \times *Taq* DNA polymerase buffer
8 μ l 0.5 mM 3dNTP mix (minus C)
5 μ l 40 μ M dCTP
5 μ l 80 ng/ μ l primer F
5 μ l 80 ng/ μ l primer R
2.5 μ l [α -³²P]dCTP
18 μ l H₂O
1.5 μ l 5 U/ μ l *Taq* DNA polymerase.

The dried acrylamide will rehydrate in the reaction mix as the PCR progresses.

4. Carry out 17 of the following PCR amplification cycles:

- 1 min 94°C (denaturation)
- 1 min 62°C (annealing)
- 1 min 72°C (extension)

The optimal number of PCR cycles will depend upon variables that are difficult to control, such as the amount of DNA in the dried acrylamide. For this reason it is important to save some of the dried complex so that the PCR reaction can be repeated with additional cycles if sufficient product is not generated.

5. Purify PCR product (see Basic Protocol, steps 4 to 6).
6. Digest bound oligonucleotide pool with an excess of *Eco*RI and *Bam*HI restriction endonucleases (UNIT 3.1).
7. Phenol extract digested oligonucleotides, recover by ethanol precipitation, and resuspend at a dilution appropriate for cloning (UNIT 2.1).
8. Ligate oligonucleotides into an appropriate vector and transform into *E. coli* (UNIT 3.16). Perform sequence analysis of individual selected sites (UNITS 7.4-7.6).

CAUTION: Ligations will be moderately radioactive. Take care to dispose of bacterial plates appropriately.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Binding buffer

Wash buffer (see recipe) containing:

- 0.1% Nonidet P-40 (NP-40)
- 1 mM benzamidine
- 0.5 mM phenylmethylsulfonyl fluoride (PMSF)
- 0.5 mM DTT

Add benzamidine, PMSF, and DTT fresh before use. Buffer without those components can be stored several weeks at 4°C.

Elution buffer

- 0.5 M ammonium acetate
- 1 mM EDTA
- 0.1% SDS
- Store several months at room temperature

Recovery buffer

- 50 mM Tris·Cl, pH 8
- 100 mM sodium acetate
- 5 mM EDTA
- 0.5% SDS
- Store several months at room temperature

Wash buffer

- 20 mM HEPES, pH 7.9
- 100 mM KCl
- 0.2 mM EDTA
- 0.2 mM EGTA
- 20% (v/v) glycerol
- Store several weeks at 4°C

COMMENTARY

Background Information

Determination of the DNA binding specificity of a transcription factor is a vital step in the characterization of its interactions with DNA and the identification of its regulatory targets. Before the advent of site selection methods, this was a major problem in the study of DNA-binding proteins identified on the basis of properties other than sequence-specific DNA binding.

A significant advance came with the demonstration by Oliphant et al. (1989) that specific binding sites for a DNA-binding protein could be affinity selected from random sequence oligonucleotides. Another important development was the combination of selection with *in vitro* amplification using the polymerase chain reaction (PCR). Kinzler and Vogelstein (1989) showed that iterative rounds of selection and amplification could be used to isolate specific binding sites from a complex population of small genomic DNA fragments.

Simpler and more efficient site selection methods were soon described in which iterative rounds of selection and amplification were combined with the use of random-sequence oligonucleotides as a source of binding sites. Random-sequence oligonucleotides are used in preference to genomic fragments for three reasons. First, their uniform size allows progress of site selection to be easily monitored on mobility shift gels. Second, a consensus can be derived simply by comparing recovered sequences without the need for binding-site mapping techniques. Third, small oligonucleotides are less likely to contain binding sites for other nuclear proteins, which may increase background recoveries.

Site selection methods differ mainly in the way in which DNA-protein complexes are separated from unbound oligonucleotides. A variety of separation techniques have been successfully used, including affinity columns (Ecker et al., 1991), filter binding (Thiesen and Bach, 1990), precipitation with glutathione-*S*-transferase beads (Chittenden et al., 1991), and mobility shift gels (Blackwell and Weintraub, 1990). The protocol described here uses immunoprecipitation as a means of isolating specific DNA-protein complexes (Pollock and Treisman, 1990).

Immunopurification gives a high degree of enrichment for target protein and can be used to isolate DNA-protein complexes from crude mixtures such as cell extracts. Immunoselection is also quite sensitive, and has been used

to determine the DNA-binding specificities of proteins present at low concentrations in cell extracts. The ability to use crude extracts has several advantages. First, there is no requirement for purified or overexpressed target protein. Second, the technique can be applied to proteins that require postranslational modifications in order to bind DNA. Third, sites can be selected for proteins that bind DNA only as components of heteromeric complexes. A disadvantage of immunoprecipitation is the requirement for an antibody specific for the test protein. In the absence of a specific antibody, an epitope-tagging approach can be used to recover protein-DNA complexes (Pollock and Treisman, 1990).

Site selection techniques can be used to determine consensus binding sites for proteins whose target sequences are unknown. Examples would include proteins isolated on the basis of properties other than sequence-specific DNA binding but suspected to bind DNA by virtue of their affinity for nonspecific DNA, nuclear localization, or homology to known DNA-binding proteins. The same separation techniques can be adapted to different selection protocols.

Site selection can also provide additional information about the protein-DNA interactions of previously characterized DNA-binding domains. For example, selection experiments can help to determine the full range of target sequences for a given DNA-binding domain, provide information on the relative importance of bases within a consensus, and reveal subtle differences in sequence specificity between closely related members of a transcription factor family.

Site selection can also be used in conjunction with mutational studies of DNA-binding domains to examine the contribution of individual amino acids or motifs to sequence recognition.

More recently the use of immunoselection of protein-DNA complexes from nuclear extracts has been extended to identify potential interactions between transcription factors within multicomponent transcription complexes (Funk and Wright, 1992).

Critical Parameters

Probe oligonucleotides

R76 comprises a random 26-base sequence flanked by invariant 25-nucleotide primer-

binding sequences containing *EcoRI* and *BamHI* restriction sites to facilitate cloning of binding oligonucleotides after selection (see Basic Protocol, Materials). Alternative primers may have to be used if the protein under investigation has specificity that overlaps with the forward (F) and reverse (R) primers used here.

The randomized core of R76 is long enough to encompass the consensus binding sites for most if not all known DNA-binding proteins. Longer or shorter random cores can also be used. Use of an oligonucleotide with a random core that is longer than the transcription-factor binding site increases the number of potential binding sites present in a given mass of DNA even though the number of oligonucleotides is decreased. For example, 0.4 ng dsR76 represents $\sim 5 \times 10^9$ oligonucleotides but the 26-bp random core contains 8×10^{10} ($5 \times 10^9 \times 16$) 10-mers, whereas 0.4 ng of an oligonucleotide with a 10-bp random core represents only 6.3×10^9 10-mers. Random cores longer than 26 bp can and have been used, although this may increase the difficulty of perceiving complex unknown motifs.

Binding, immunoprecipitation, and wash

The binding conditions described here are of moderate ionic strength and are designed to support stable protein-DNA complexes. They are included only as a guide and may not be optimal for every DNA-binding protein. If suitable binding conditions for the protein under investigation are already known then these should be used. Binding reactions contain relatively low concentrations of DNA-binding protein and binding probe to prevent the recovery of weak binding sites. The overnight incubation is to ensure efficient adsorption of DNA-protein complexes onto the protein A-Sepharose beads. The time required for this to occur will vary with different antibody/target protein combinations. It is likely that in many cases the incubation time can be cut significantly from that described here. A pilot immunoprecipitation experiment can be performed to determine a minimum incubation period (see Antibodies, below).

The wash step is designed to remove unbound oligonucleotides without disrupting DNA-protein complexes. The ionic strength of the wash buffer is an important factor in determining the stringency of the selections. The time spent performing the washes can also determine the affinity of selected sites, particularly for DNA-binding proteins with high dis-

sociation rates. In the protocol described here washes are performed rapidly with binding buffer under nondenaturing conditions.

DNA amplification

The PCR amplification conditions described here are designed to directly radiolabel the oligonucleotide pool to a known specific activity. This allows the recovered DNA to be quantitated after each round of selection and enables amplified selected DNA to be used directly as a probe in subsequent binding reactions.

There are two potential problems to be aware of when using PCR to amplify oligonucleotide pools. First, as the PCR reaction progresses, the concentration of the oligonucleotide pool can increase to a point at which self-priming of product competes with the re-annealing of the F and R primers. This may lead to artifactual PCR products if too many PCR cycles are performed. Second, overamplification will result in primer or dNTP depletion before thermal cycling has ceased. The resulting population of oligonucleotides is unable to fully renature and cannot be used in further rounds of selection. For these reasons, the PCR reaction should be carefully calibrated with respect to input DNA and number of cycles performed. To do this, set up a series of PCR reactions containing 1 pg input DNA. Subject each reaction to an increasing number of thermal cycles (between 14 and 20) and analyze the radiolabeled products by nondenaturing polyacrylamide gel electrophoresis. Use the maximum number of thermal cycles that gives rise to a single correctly sized product in subsequent site selection PCR reactions.

Even after the correct number of thermal cycles has been determined by calibration, it is recommended that the gel purification step be retained during this protocol to guard against overamplification artifacts.

Number of selection cycles

The exact number of selection cycles required may vary according to the source, amount, affinity, and specificity of the protein under investigation. In practice, visible complexes on a mobility shift gel are usually generated after three to four rounds of selection.

During early rounds of selection, the concentration of DNA-binding protein will be in excess over specific binding sites present in oligonucleotide pools. Under these conditions sites will be selected as long as they are above

a certain affinity determined by the stringency of the binding and wash conditions, and the concentration of DNA-binding protein. As selections are repeated, the proportion of binding sites in the selected pools will increase until the DNA-binding protein is no longer in excess. Once this point is reached selections will begin to discriminate in favor of higher-affinity sites. Because biologically relevant target sites are not necessarily those with the highest affinity, it is advisable to analyze sequences from earlier rounds of selection to ensure that the full spectrum of sequences bound by the protein under the assay conditions is represented.

Antibodies

Antibodies should be specific for the DNA-binding protein of interest and should allow efficient immunoprecipitation of the test protein. The ability of an antibody to immunoprecipitate the test protein can be confirmed in pilot experiments with the extracts and binding conditions to be used in site selections.

It is important to ensure that the antibody to the test protein does not disrupt DNA-protein complexes. For transcription factors with known binding sites this can be directly tested by performing a mobility shift assay in which antibody is included in the binding reaction. Inclusion of antibody should not abolish the protein-DNA complex; rather, a slower migrating or “supershifted” ternary complex containing protein, DNA, and antibody should result.

In cases where the effect of an antibody upon DNA binding cannot be directly assessed, tagging the test protein at the amino or carboxyl terminus with a short epitope to which a monoclonal antibody exists minimizes the risk of interfering with protein-DNA interactions.

Troubleshooting

Troubleshooting is greatly facilitated by performing site selection experiments in parallel using the appropriate control extracts. An extract containing a DNA-binding protein with a known sequence specificity should be included as a positive control. If possible, the positive control protein should be immunoprecipitated with the same antibody as the test proteins, although this may only be feasible if an epitope tagging approach is used. It is also important to confirm that selections performed with an extract that does not contain the target protein do not result in selection of specific binding sites. Such extracts include unprogrammed *in vitro* translation extract, extract

from cells that do not express the protein under investigation, and extract containing test protein without an epitope tag. Finally, when monitoring success of the selections by mobility shift assay, binding reactions should be included in which oligonucleotide pools selected with the test protein are incubated with extracts that do not contain the protein under investigation. This allows identification of appropriate mobility shift complexes from which to isolate DNA (see below).

If selection fails to produce visible complexes on a mobility shift gel, the following possibilities should be considered.

1. Binding conditions may not be optimal for the protein under investigation. Selections can be repeated with a number of different binding conditions tested. The most important parameters to vary are the salt concentration and the nature and amount of nonspecific carrier DNA (it may be possible to leave carrier DNA out of the binding reaction altogether).

2. The protein under investigation binds DNA only as part of a heteromeric complex.

This applies if selections were performed with purified proteins or proteins produced by *in vitro* translation. In this case, selections can be repeated with nuclear extracts prepared from cells transfected with a cDNA encoding the test protein.

3. The test protein may not form stable DNA-protein complexes in mobility shift gels.

This possibility should be seriously considered if probe recovery increases in later cycles, indicating that binding-site enrichment has occurred. In this case, it may be worth running mobility shift gels under different buffer conditions or bypassing the mobility shift step by sequencing sites cloned from the entire selected oligonucleotide pool.

4. The protein under investigation may be sensitive to repeated cycles of freezing and thawing.

It is always advisable to use a fresh aliquot of extract for each round of selection.

Another potential pitfall is the appearance of multiple complexes on mobility shift gels following selection. These may be due to selection by proteins in the extract that cross-react with the selecting antibody or to recognition of a subset of sites selected by the test protein by other DNA-binding proteins in the extract. It is therefore important to identify the appropriate complex. Bound oligonucleotides should be isolated only from complexes that do not appear with negative control extracts.

Anticipated Results

Recovered sequences should contain binding sites for the DNA protein under investigation and give rise to a consensus recognition sequence when appropriately aligned. The frequency with which a base occurs in a particular position will be proportional to its relative importance for binding. The derived consensus and the predicted relative importance of individual bases within the consensus can be tested by performing binding assays with oligonucleotides containing appropriate wild-type and mutant binding sites.

Contamination or PCR errors may occasionally give rise to sequences that do not conform to the derived consensus.

Time Considerations

It is possible to recover DNA from an overnight binding reaction in the morning and have amplified, purified DNA ready to add to another binding reaction in the evening. Four rounds of selection can therefore be performed in 4 days, although this becomes much harder to achieve if more than six selections are performed in parallel.

The mobility shift gel can be run on day 5 and DNA isolated from complexes on day 6. It is therefore possible to have templates ready for sequencing by day 8.

Literature Cited

- Blackwell, T.K. and Weintraub, H. 1990. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* 250:1104-1110.
- Chittenden, T., Livingston, D.M., and Kaelin, W.J. 1991. The T/E1A-binding domain of the retinoblastoma product can interact selectively with a sequence-specific DNA-binding protein. *Cell* 65:1073-1082.
- Ekker, S.C., Young, K.E., von, K.D., and Beachy, P.A. 1991. Optimal DNA sequence recognition by the Ultrabithorax homeodomain of *Drosophila*. *EMBO J.* 10:1179-1186.
- Funk, W.D. and Wright, W.E. 1992. Cyclic amplification and selection of targets for multicomponent

complexes: Myogenin interacts with factors recognizing binding sites for basic helix-loop-helix, nuclear factor 1, myocyte-specific enhancer-binding factor 2, and COMP1 factor. *Proc. Natl. Acad. Sci. U.S.A.* 89:9484-9488.

Kinzler, K.W. and Vogelstein, B. 1989. Whole genome PCR: Application to the identification of sequences bound by gene regulatory proteins. *Nucl. Acids Res.* 17:3645-3653.

Oliphant, A.R., Brandl, C.J., and Struhl, K. 1989. Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: Analysis of yeast GCN4 protein. *Mol. Cell Biol.* 9:2944-2949.

Pollock, R. and Treisman, R. 1990. A sensitive method for the determination of protein-DNA-binding specificities. *Nucl. Acids Res.* 18:6197-6204.

Thiesen, H.J. and Bach, C. 1990. Target Detection Assay (TDA): A versatile procedure to determine DNA binding sites as demonstrated on SP1 protein. *Nucl. Acids Res.* 18:3203-3209.

Key References

Pollock, R. and Treisman, R. 1990. A sensitive method for the determination of protein-DNA-binding specificities. *Nucl. Acids Res.* 18:6197-6204.

Describes the use of the method upon which this protocol is based to select binding sites for SRF and FOS protein present in nuclear or in vitro translation extracts. Also demonstrates the use of epitope tagging to recover protein-DNA complexes.

Szostak, J.W. 1992. In vitro genetics. *Trends Biochem. Sci.* 17:89-93.

Contains a useful review of site selection techniques.

Wright, W.E. and Funk, W.D. 1993. CASTing for multicomponent DNA-binding complexes. *Trends Biochem. Sci.* 18:77-80.

A review of immunoprecipitation of protein-DNA complexes from nuclear extracts and its application in selecting sites for interacting transcription factors within multicomponent complexes.

Contributed by Roy M. Pollock
Ariad Pharmaceuticals, Inc.
Cambridge, Massachusetts

Yeast One-Hybrid Screening for DNA-Protein Interactions

One-hybrid screening in yeast is a powerful method to rapidly identify heterologous transcription factors that can interact with a specific regulatory DNA sequence of interest (the bait sequence). Essentially, one-hybrid screens are derived from the two-hybrid concept. In this technique, the interaction between two proteins (bait and prey) is detected via *in vivo* reconstitution of a transcriptional activator that turns on expression of a reporter gene. In the one-hybrid system, detection is based on the interaction of a transcription factor (prey) with a bait DNA sequence upstream of a reporter gene. To ensure that DNA binding results in reporter-gene activation, cDNA expression libraries are used to produce hybrids between the prey and a strong *trans*-activating domain. The advantage of cloning transcription factors or other DNA-binding proteins via one-hybrid screenings, compared to biochemical techniques, is that the procedure does not require specific optimization of *in vitro* conditions. In fact, yeast is being used as a “eukaryotic test tube” to demonstrate interactions. This may circumvent possible difficulties, e.g., incorrect protein folding or lack of post-transcriptional modifications, that may occur in screenings for transcription factors via a prokaryotic expression system such as that used in the *E. coli*-based southwestern procedure (Singh et al., 1988).

In the one-hybrid screening procedure described in this unit, cDNA libraries are screened using specific yeast reporter strains. The detection of the interaction between transcription factor and bait sequence is via activation of the positive growth selection marker *HIS3*. The bait sequence can be a well defined *cis*-acting element, but larger regulatory sequences can also be used. For example, complete promoters can be used in a random search for regulators. The precise recognition sequences of transcription factors isolated in such a screening can subsequently be mapped, and this strategy can lead back to identification of the *cis*-acting elements. The yeast reporter strains are constructed using the efficient pHIS3/pINT1 vector system. This vector system consists of two components. First, the bait sequence is cloned in front of the *HIS3* reporter gene in either pHIS3NB or pHIS3NX (Table 12.12.1 and Fig. 12.12.3). Secondly, the *HIS3* reporter cassettes with upstream bait sequence are transferred to the integrative vector, pINT1. The complete *HIS3* reporter is then transformed to yeast and integrated via double cross-over in the genome using homologous sequences from pINT1. Apart from the *HIS3* gene itself, the system does not use any other auxotrophic markers, which are thus kept free for other purposes, including the selection for library plasmids and, for example, the construction of dual-reporter strains as discussed in the Commentary. Another important feature of the pHIS3/pINT1 system is that selection for integration of the reporter is via the dominant selection gene, *APT1*. In some other described methods, integration is dependent on the low level of so-called leaky expression that *HIS3* reporter genes often show in the absence of a *trans*-acting factor binding the bait sequence. However, the fact that not all bait sequences confer this leaky expression is a severe limitation of these other procedures. For construction and verification of a pINT1-*HIS3* reporter strain, see Basic Protocol 1.

Also described in detail is a procedure for performing one-hybrid screenings (see Basic Protocol 2). For this procedure, a cDNA expression library is required, usually made in an *E. coli*-yeast shuttle vector that contains the sequences required for replication in both organisms. In a typical screen (Fig. 12.12.1), pools of library plasmid DNA are transformed into the pINT1-*HIS3* reporter strain. Selection for presence of the library plasmid is via the auxotrophic marker on the vector (usually *LEU2*). If an expressed transcription factor recognizes the bait sequence, the *HIS3* reporter will be activated. This complements

the auxotrophic *his3* mutation, allowing colonies to grow on selective medium lacking leucine and histidine. Library vectors used in the one-hybrid system are designed to express cDNAs as translational fusions with a strong constitutive *trans*-activation domain from a transcription factor (e.g., Gal4p, VP16). Therefore, the method is not limited to isolation of transcription factors that contain an intrinsic activation function, but also allows identification of repressors or other DNA-binding proteins. The one-hybrid method described in this unit is fully compatible with the many existing Gal4p- or *LexA*-based hybrid libraries that have been constructed for two-hybrid applications (e.g., UNIT 20.1). Most commercially available libraries were produced by direct cloning in plasmid vectors. Alternatively, expression libraries can more efficiently be made in lambda phage-based vectors. The Support Protocol refers to available library construction vectors and the procedure for conversion of lambda phage into plasmid libraries via *in vivo* mass excision.

Following the actual screening, the discrimination between true and false positives is an important step. The unit describes how to analyze the positive colonies with simple molecular methods and how to develop strategies to determine whether or not an isolated prey indeed represents a transcription factor that specifically interacts with the bait sequence (see Basic Protocol 3).

All steps for making the necessary preparations to perform a one-hybrid screening and to analyze the outcome are summarized in a flowchart (Fig. 12.12.2). Furthermore, a schematic representation of the procedure is shown in Figure 12.12.1. Many foreseeable complications or possible unexpected results are extensively discussed in the Comm-

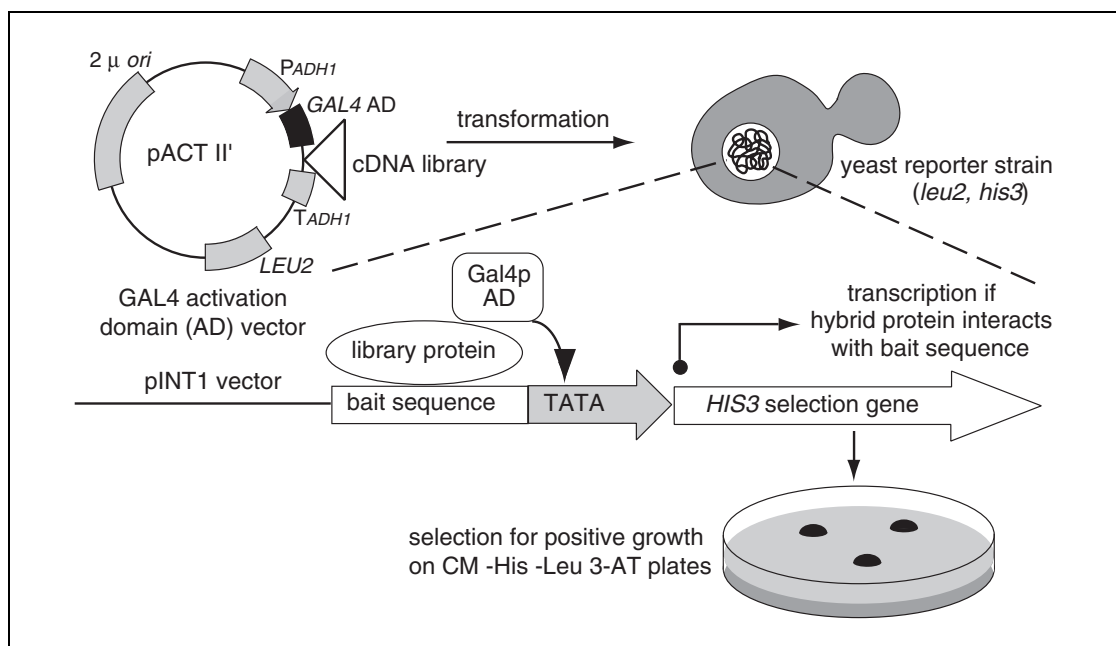


Figure 12.12.1 Schematic overview of the yeast one-hybrid screening procedure to clone transcription factors or other DNA-binding proteins. A cDNA expression library in an *E. coli*-yeast shuttle vector is required. In the example shown here, the cDNA is expressed from the yeast *ADH1* promoter and is fused to the activation domain (AD) sequence of the yeast transcription factor Gal4p. With a unidirectional cDNA library, one-third of the clones should produce in-frame fusions with the AD coding sequence (hybrid proteins). The detection of the clones is via activation of the *HIS3* growth marker. Recognition and binding of a transcription factor to the bait sequence cloned upstream of the *HIS3* gene will result in *HIS3* expression. This complements for a chromosomal *his3* mutation, thereby allowing colony formation on histidine-deficient medium. The *LEU2* marker on the cDNA library vector complements for a chromosomal mutation in the *leu2* gene.

12.12.2

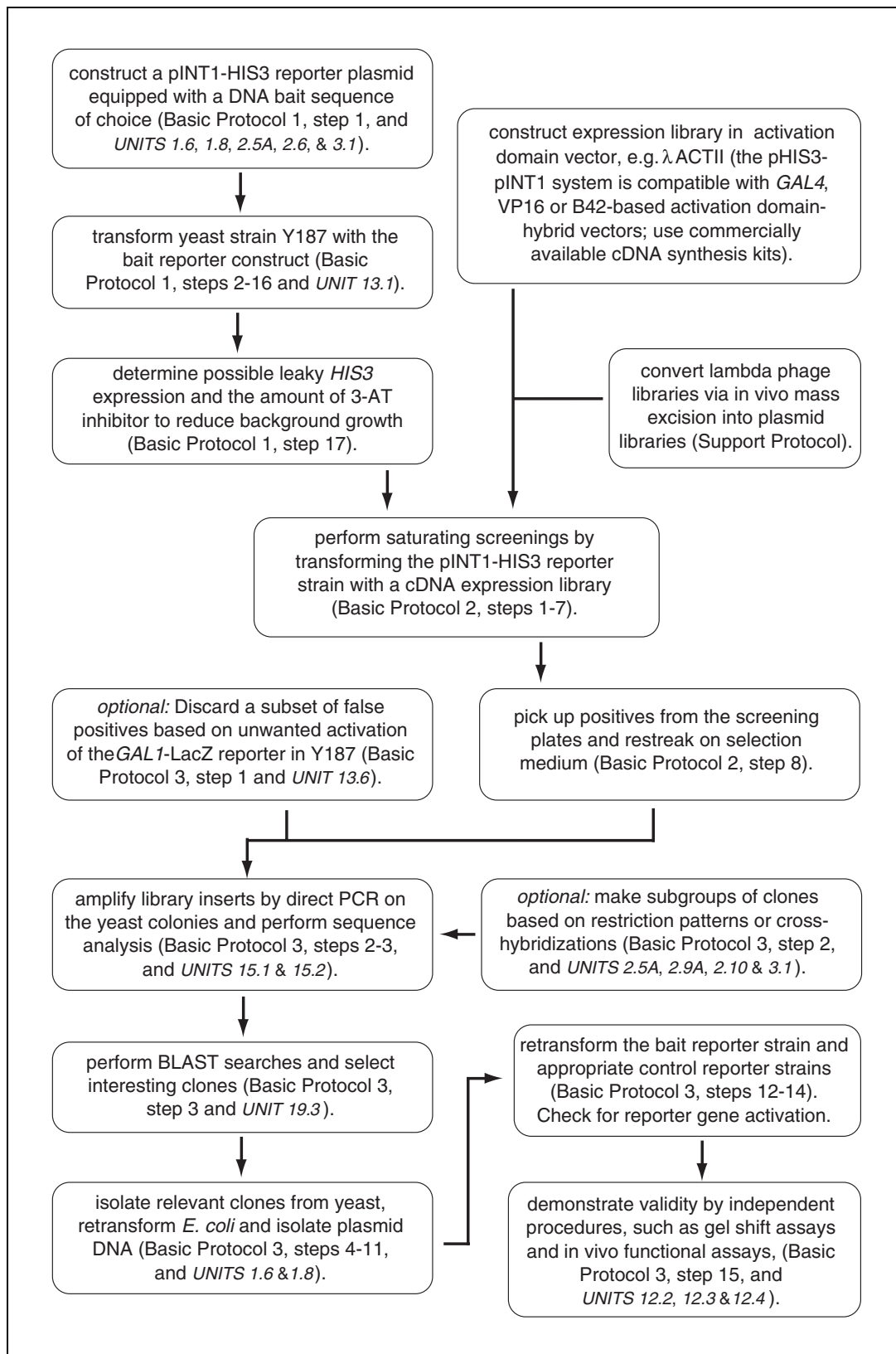


Figure 12.12.2 Flow chart for performing yeast one-hybrid screenings.

tary. The information in this unit is based on results from various screenings performed with different bait sequences and cDNA libraries from both animals and plants.

CONSTRUCTION AND CHARACTERIZATION OF YEAST REPORTER STRAINS

The first step in a one-hybrid screening is to construct a yeast strain that is equipped with a *HIS3* reporter gene construct. In this construct, the *HIS3* gene is preceded by a DNA bait sequence, which can be a defined *cis*-acting sequence, a heterologous promoter, or regulatory sequences from exon, intron, or 5'- or 3'-noncoding regions. Also, it may be desired to screen for proteins interacting with centromere or telomere regions. The *HIS3* reporter construct is integrated at the nonessential *PDC6* locus using the integrative vector, pINT1. Once this strain is obtained, its growth characteristics on His-lacking medium have to be determined in a series of control experiments. Some sequences confer a certain amount of *trans*-activation on the *HIS3* gene that is sufficient for growth on His-lacking medium. Such possible background growth due to leaky expression can be reduced by the addition of 3-amino-1,2,4-triazole (3-AT), a competitive inhibitor of the His3p enzyme. Once the required 3-AT concentration is established, the yeast strain can be used to screen cDNA expression libraries. Using the same integration procedure as described here for the bait construct, control reporter strains should be constructed that will be useful after the library screening for investigation of the specificity of putative positives. Basic Protocol 3 and the Commentary section give directions for control reporter design and application.

Materials

- DNA fragment or oligonucleotides, containing the bait sequence
- Plasmids pHIS3NB or pHIS3NX and pINT1 (Table 12.12.1, Fig. 12.12.3; available from P.B.F. Ouwkerk or A.H. Meijer; ouwkerk@rulbim.leidenuniv.nl and meijer@rulbim.leidenuniv.nl)
- Yeast strain Y187 (Table 12.12.2)
- Liquid YAPD medium: YPD medium (UNIT 13.1) supplemented with 20 mg/liter adenine hemisulfate
- 10× TE buffer, pH 7.5 (see recipe)
- 10× lithium acetate stock (see recipe)
- 50% (w/v) polyethylene glycol (PEG) stock (see recipe)
- 10 mg/ml YEASTMAKER carrier DNA (Clontech; also see UNIT 13.7 or UNIT 20.1)

Table 12.12.1 Overview of Vectors Used for Yeast Reporter Strain Construction

Vector ^a	<i>E. coli</i> selection	Yeast markers	Genbank accession no. ^b
pHIS3NB	Cb ^r	<i>HIS3</i>	AF275029
pHIS3NX	Cb ^r	<i>HIS3</i>	AF275030
pINT1	Cb ^r , Km ^r	<i>APT1</i> ^c	AF289993

^aPlasmids pHIS3NB, pHIS3NX, and pINT1 are described by Meijer et al. (1998) and can be obtained from P.B.F. Ouwkerk or A.H. Meijer. All plasmids replicate with high copy in *E. coli*.

^bAnnotated sequence information is deposited in the NCBI database.

^cThe *APT1* marker confers resistance to a relatively low concentration (25 mg/liter) of kanamycin in *E. coli*, and to G418 (150 mg/liter in YPAD) in yeast.

Table 12.12.2 Yeast Strain Genotypes

Strain	Genotype	Reference
Y187 ^a	<i>MATα ura3-52 his3-Δ200 ade2-101 trp1-901 leu2-3,112 me⁻ gal4Δ gal80Δ URA3::GAL1_{UAS}-GAL1_{TATA}-lacZ MEL1</i>	Harper et al., 1993; Clontech
YPH500 ^a	<i>MATα ura3-52 lys2-801 ade2-101 trp1-Δ63 his3-Δ200 leu2-Δ1</i>	Sikorski and Hieter, 1989
YJO ^a	<i>MATα gal4Δ gal80Δ ura3-52 his3 ade2-101 trp1 leu2-3,112 MEL1</i>	Leuther and Johnston, 1992 ^b
YM954 ^a	<i>MATα ura3-52 his3-Δ200 ade2-101 lys2-801 trp1-901 leu2-3, 112 gal4-Δ512 gal80-Δ538 tyr1-501</i>	Wilson et al., 1991
YM4271 ^a	<i>MATα ura3-52 his3-Δ200 ade2-101 lys2-801 trp1-901 leu2-3, 112 gal4-Δ512 gal80-Δ538 tyr1-501 ade5::hisG</i>	Liu et al., 1993
YPO101 ^c	Y187 <i>PDC6::pINT1-HIS3</i>	P.B.F. Ouwkerk (unpub. observ.)

^aThese strains are suitable hosts for one-hybrid screenings and are mentioned in the text. Use of Y187 is recommended for the procedure described in this unit.

^bThis reference refers to YJO-Z which is a derivative of YJO, containing a *GAL1-lacZ* fusion chromosomally integrated.

^cIn strain YPO101 a *HIS3* reporter lacking an upstream bait sequence was integrated via the pHIS3/pINT1 system. This strain can serve as a control strain for the analysis of library clones.

YAPD-G418 plates; YPD plates (*UNIT 13.1*) supplemented with 20 mg/liter adenine hemisulfate and G418 (150 mg/liter; add from 1000-fold stock solution) 87% (v/v) glycerol (sterilize by autoclaving for 20 min; store indefinitely at room temperature)

Complete minimal medium (CM) plates (*UNIT 13.1*) prepared with Difco Bacto-Agar or Duchefa Micro-Agar

CM –His plates (*UNIT 13.1*) with a concentration series of 3-amino-1,2,4-triazole (3-AT) (see recipe for 3-AT)

30°C incubator, with and without shaker

50-ml centrifuge tubes

Low-speed tabletop centrifuge (preferably swing-out) for 50-ml tubes

1.5-ml microcentrifuge tubes (e.g., Eppendorf or equivalent)

42°C water bath

10- to 15-ml tubes or flasks

Additional reagents and equipment for miniprep DNA isolation procedures (*UNIT 1.6*), transformation of *E. coli* (*UNIT 1.8*), gel electrophoresis (*UNIT 2.5A*), gel extraction procedures (*UNIT 2.6*), and restriction analyses (*UNIT 3.1*)

Clone the DNA bait sequence in the pHIS3-pINT1 vector system

1. Design a cloning strategy and clone the bait sequence in the polylinker of either pHIS3NB or pHIS3NX (Table 12.12.1, Fig. 12.12.3) to make a transcriptional fusion. For the cloning steps, use the standard techniques described in *UNITS 1.6, 1.8, 2.5A, 2.6 & 3.1*. Next, clone a fragment containing the *HIS3* reporter gene including the bait sequence as *NotI-BamHI* or *NotI-XbaI* fragment in the corresponding sites of the integrative vector pINT1.

It is not necessary to include other regulatory sequences such as the TATA box, transcription start site, and leader sequence. In the pHIS3 vectors, the HIS3 reporter gene is preceded by a minimal promoter (100-bp length) that already contains these features. When using the BclI site in pINT1, the vector needs to be isolated from a dam, dcm strain (GM48 or any of its derivatives, e.g., JM110). BclI cuts at 50°C and it is strongly recommended that

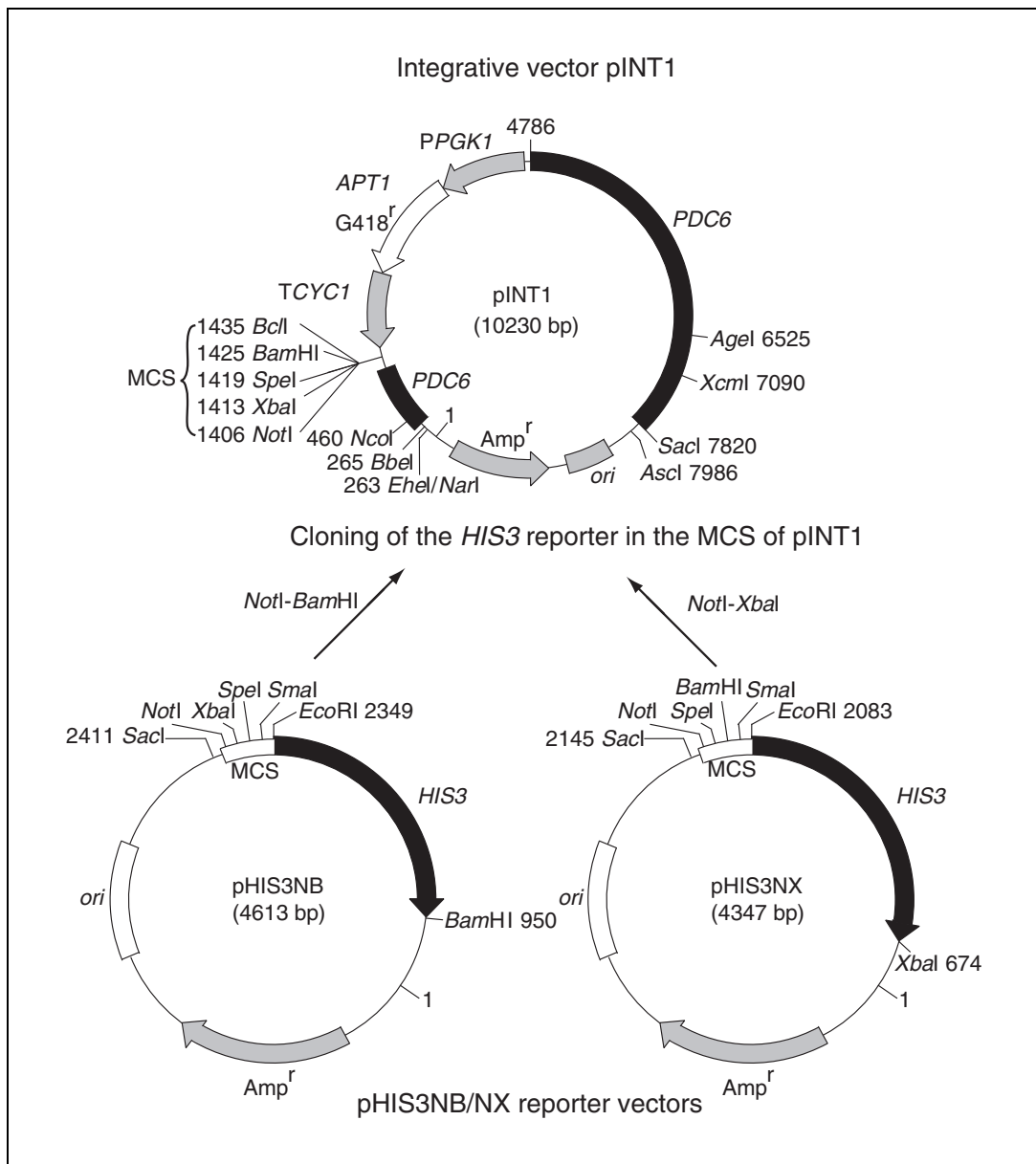


Figure 12.12.3 The pHIS3/pINT1 vector system for construction of reporter yeast strains for one-hybrid screening. The *HIS3* reporter plasmids pHIS3NB and pHIS3NX are based on SK II Bluescript (Stratagene). The *HIS3* gene is preceded by its own minimal promoter and a multiple cloning site with several unique sites for insertion of the bait sequence. Inserts can be sequenced using the M13 reverse primer. pINT1 is pUC29-based and contains the sequences necessary for integration in the yeast *PDC6* locus via double crossover. Selection for integration is via the dominant marker *APT1* that confers resistance in yeast to the antibiotic G418 (also termed gentamycin G418 or Geneticin), or kanamycin in *E. coli*. The *APT1* gene is controlled by the yeast *PGK1* promoter and the *CYC1* terminator. *APT1* was originally derived from the bacterial transposon Tn *g03* and codes for an aminoglycoside phosphotransferase (Hadfield et al., 1990). Several unique restriction sites are present downstream of the *APT1* expression cassette, which can be used for the unidirectional cloning of a bait-*HIS3* construct that is derived from either pHIS3NB or pHIS3NX. For integration of pINT1 in the *PDC6* locus, a fragment lacking the pUC29 sequences is isolated by restriction with *NcoI* (alternatively *BbeI*, *EheI* or *NarI*) and *SacI* (alternatively *Ascl*, *XcmI*, or *AgeI*). The choice for the pHIS3NB or -NX variant and for the restriction enzymes to linearize pINT1 is dependent on the sequence of the bait.

CsCl- or column-purified DNA (UNITS 1.7 or 2.1B) be used. The *APT1* marker in *pINT1* is also expressed in *E. coli* where it confers resistance to a low concentration of kanamycin (25 mg/liter). It can therefore be used in cloning strategies as an extra marker gene in addition to the carbenicillin resistance gene of *pINT1*.

2. Isolate 100 to 500 ng of the fragment from *pINT1* containing the bait sequence hooked up to the *HIS3* gene, the *APT1* marker gene, and the flanking *PDC6* sequences for homologous recombination (UNIT 2.6). Use the restriction sites indicated in Figure 12.12.3.

If necessary, it is possible to use only one restriction site for linearizing. However, this can lead to integration via single cross-over, which may affect the stable maintenance of the construct. Therefore, the use of two restriction sites is recommended. In most cases, the authors have used the combination of NcoI and SacI to linearize the pINT1 vector prior to yeast transformation, and have obtained integrations by double cross-over.

Grow and prepare yeast cells for transformation

3. Transform yeast with a lithium acetate transformation procedure as described in the following steps. First, grow a 50-ml YAPD culture of a yeast host strain of choice (e.g., Y187) in YPD medium supplemented with 20 mg/liter adenine hemisulfate overnight at 30°C.

The authors recommend the use of Y187 (Table 12.12.2) as host strain for efficient one-hybrid screenings. Y187 is a gal4gal80 strain, which is desirable since most hybrid libraries contain the Gal4p activation domain (AD). GAL4GAL80 strains such as YPH500 (Table 12.12.2) can be used, but in that case galactose should be the carbon source to avoid inhibition of Gal4p AD by Gal80p. Disadvantages of using galactose are that growth is slower and it is quite expensive. A useful aspect of Y187 is that it gives higher transformation efficiencies compared to other gal4gal80 strains such as YM954, YM4271, or YJO (Table 12.12.2). Some strains, including YPH500 and YJO, tend to become pink on YAPD medium due to the ade2-101 mutation, but Y187 remains white, although it also has this mutation.

4. Dilute the overnight culture to an OD_{600} of ~0.25 in liquid YAPD medium prewarmed to 30°C and grow the culture(s) for an additional 3 hr with shaking at 30°C. Grow one culture for every 10 planned transformation reactions.

*For optimal transformation efficiency, it is important that the cells be harvested during the exponential phase (OD_{600} 0.4 to 0.8). Therefore, it may be necessary to adjust the incubation time when using other strains. Usually step 4 is not necessary for *pINT1*-*HIS3* transformations, but the authors certainly advise that step 4, above, be included during the library transformations (see Basic Protocol 2).*

5. Harvest the yeast cells of the overnight culture in 50-ml centrifuge tubes by centrifuging for 1 min at $2400 \times g$ in a swing-out tabletop centrifuge (e.g., Heraeus Megafuge 1.0R, 2500 rpm) at room temperature.
6. Discard the supernatant fluid, resuspend the cells in 50 ml sterile water by vigorous shaking, and repeat centrifugation as in step 5.
7. Discard the supernatant fluid and resuspend the yeast cells in 1 ml of $1 \times TE/1 \times$ lithium acetate prepared from the $10 \times$ stock solutions.
8. Transfer the yeast suspension to a 1.5-ml microcentrifuge tube, microcentrifuge for 30 sec at maximum speed, and resuspend in 250 μ l of $1 \times TE/1 \times$ lithium acetate.

At this stage the cells are ready to be transformed. Please note that none of the steps should be carried out on ice. Lithium acetate-competent cells can be stored up to several weeks at 4°C, although their competency will gradually decrease. This transformation procedure is performed, with some modifications, according to Gietz et al. (1992).

Transform the pHIS3-pINT1 reporter construct to yeast

9. Directly prior to the transformation reactions, denature the required amount of YEASTMAKER carrier DNA (25 µg per reaction) by boiling for 10 min and, then place the DNA on ice. Also prepare a solution of 40% (w/v) PEG/1× TE buffer/1× lithium acetate by mixing 8 parts 50% (w/v) PEG, 1 part 10× TE buffer, and 1 part 10× lithium acetate solution in quantity sufficient for use in step 10.
10. Mix 100 to 500 ng of the isolated fragment from the pINT1-HIS3 vector (step 2) with 25 µg of 10 mg/ml carrier DNA in a maximum total volume of 10 µl in a 1.5-ml microcentrifuge tube. Add 50 µl of the yeast suspension and 300 µl of freshly prepared 40% PEG/1× TE/1× lithium acetate solution, and vortex to mix.

The authors have found the YEASTMAKER carrier DNA from Clontech to be a reliable (and not too expensive) source of high-quality carrier DNA for yeast transformations. Alternatively, a protocol to prepare carrier DNA can be found in UNIT 13.7 or UNIT 20.1. However, depending on the source of the starting material, the quality of the prepared batches can vary considerably.

11. Incubate for 30 min at 30°C with shaking.
12. Transfer and incubate the tubes for 15 min in a 42°C water bath.
13. Harvest the cells by microcentrifuging for 30 sec.
14. Allow the yeast cells to recuperate by resuspending in 1 ml YAPD medium, and transfer them to 10- to 15-ml tubes or flasks. Incubate for 3 to 6 hr at 30°C with shaking.

The recuperation step is necessary to allow the yeast cells to express the APT1 marker carried on the pINT1-HIS3 fragment, in order to gain resistance towards G418. The recuperation step in YAPD cannot be carried out in microcentrifuge tubes because pressure is built up due to CO₂ formation.

15. After the recuperation step, transfer the cells to a 1.5-ml microcentrifuge tube and harvest by microcentrifuging for 30 sec. Decant supernatant and resuspend cells in 100 µl of 1× TE buffer, then plate on YAPD-G418 plates. Incubate for 2 to 3 days at 30°C.
16. Pick and restreak G418-resistant colonies on YAPD-G418 plates and incubate for 2 to 3 days at 30°C. Keep the strains thus obtained at 4°C, and restreak every 2 to 3 months. For long-term storage, prepare glycerol stocks (final concentration of 30%) by adding 500 µl of 87% glycerol to 1-ml YAPD cultures grown overnight at 30°C. Snap-freeze the cultures in liquid nitrogen and keep at –80°C.

Be careful not to restreak very small transformants (petites), as these sometimes carry a spontaneous mitochondrial mutation. Due to their considerably slower growth, they are not useful in the screening. The obtained strains can be checked for correct integration of the reporter construct by PCR (UNIT 15.1) or Southern blot analysis (UNIT 2.9A). From this stage on, it is no longer necessary to grow the yeast strains on G418 since the insert is maintained stably. Apart from that, far more G418 would be required in CM than in YAPD, which makes its use very expensive.

Determine the concentration of 3-AT required to reduce possible leaky expression of the HIS3 gene

17. Streak G418-resistant colonies on CM and on CM –His plates containing a concentration series of 3-amino-1,2,4-triazole (3-AT)(0, 5, 10, 25 to 50 mM). Grow the plates for a week at 30°C. Determine the concentration that is required to reduce growth.

Note that the optimal 3-AT concentration required under screening conditions is also somewhat dependent on the plating density. Therefore, it can be useful to fine tune the

amount of 3-AT in test transformations with an empty library vector or in pilot library screenings under the conditions described in Basic Protocol 2. It is not strictly necessary to eliminate the growth completely.

LIBRARY SCREENING

The screening involves the large-scale transformation of the cDNA expression library to the pINT1-HIS3 reporter strain made in Basic Protocol 1. With the recommended Y187 host strain, ~20 to 30 individual transformation reactions will be required. Based on a transformation efficiency of $\geq 10^5$ clones per reaction, this will result in ≥ 2 to 3 million clones. This should be sufficient for saturating screenings of cDNA libraries, usually consisting of ~1 million independent recombinants. The yeast cells are plated on selective CM dropout plates deficient for leucine and histidine. The leucine deficiency is complemented via the *LEU2* gene on the library plasmid, but the histidine deficiency can only be complemented if the *HIS3* reporter gene is switched on, due to recognition of the bait sequence by a DNA-binding hybrid protein from the transformed library. If the *HIS3* reporter gene shows leaky expression, this should be inhibited by addition of 3-AT as outlined in Basic Protocol 1. Within 3 to 10 days after the transformations, His⁺ colonies will be obtained. After that, experiments have to be carried out to assess whether these represent true DNA-binding proteins or are false positives (see Basic Protocol 3). The required cDNA libraries can be made by direct cloning of cDNA into a plasmid vector or, more efficiently, by using a lambda vector containing an *E. coli*-yeast shuttle vector between the phage arms. Such lambda phage libraries are subsequently converted via in vivo mass excision into a plasmid library. More details concerning these types of library vectors and guidelines for handling can be found in the Support Protocol. Most of the useful cDNA library vectors (e.g., pACTII', pACT2, pGAD10, pGAD424, pAD-GAL4; see Table 12.12.3) contain the *LEU2* marker, but the choice for another type of vector requires only the adaptation of the marker usage in this protocol.

Materials

pINT1-HIS3 reporter yeast strain(s) (see Basic Protocol 1 for construction)
 Liquid YAPD medium: YPD medium (*UNIT 13.1*) supplemented with 20 mg/liter adenine hemisulfate
 cDNA library plasmid

Table 12.12.3 A Selection of Yeast Vectors for Expression of Hybrid Proteins Carrying a *trans*-activation Domain

Vectors	Yeast marker	Activation domain	Reference
λ ACTII/pACTII'	<i>LEU2</i>	Gal4p	Memelink, 1997
λ YGE15-GAD/pYGE15-GAD'	<i>URA3</i>	Gal4p	Memelink, 1997
HybriZAP-(2.1)/pAD-GAL4-(2.1) ^a	<i>LEU2</i>	Gal4p	Stratagene
λ ACT(2)/pACT(2) ^a	<i>LEU2</i>	Gal4p	Clontech
pGAD10 ^a	<i>LEU2</i>	Gal4p	Clontech
pGAD424 ^a	<i>LEU2</i>	Gal4p	Clontech
pGAD-GH	<i>LEU2</i>	Gal4p	Clontech
pGAD-GL	<i>LEU2</i>	Gal4p	Clontech
pB42AD ^a	<i>TRP1</i>	B42	Clontech

^aComplete sequences are deposited in the NCBI database.

10× TE buffer, pH 7.5 (see recipe)
Complete minimal medium (CM) standard-size plates, minus leucine (CM –Leu plates; *UNIT 13.1*) prepared with Difco Bacto-Agar or Duchefa Micro-Agar
CM –Leu –His standard and large (15-cm wide) plates (*UNIT 13.1*), with optimal concentration of 3-AT (see Basic Protocol 1, step 17; see recipe for 3-AT)
30°C incubator, with and without shaker

Transform a cDNA expression library to the pINT1-HIS3 reporter strain

1. Inoculate the pINT1-HIS3 reporter strain in YAPD medium and grow overnight at 30°C with shaking.
2. Follow Basic Protocol 1, steps 4 to 13, to prepare lithium acetate-competent yeast cells and to perform the desired number of transformation reactions with 1 µg of cDNA library plasmid each.

The use of exponentially growing cells instead of an overnight culture increases the transformation efficiency. It is not recommended that the transformation reactions be scaled up, because the use of larger volumes will require adjustment of incubation times, e.g., at the heat-shock step.

3. Resuspend the transformed yeast cells in 200 µl of 1× TE buffer, pH 7.5, per transformation.

A recuperation step is not necessary for expression of auxotrophic marker genes.

4. Randomly choose 2 to 3 of the identical transformation reactions and plate 100 µl of 100- and 1000-fold dilutions on CM –Leu plates to determine the transformation efficiency.
5. Plate each transformation reaction on a large (15-cm) CM –Leu, –His plate with the appropriate amount of 3-AT as determined in Basic Protocol 1, step 17.
6. After 3 days of incubation at 30°C, count the number of colonies on the transformation efficiency plates from step 4. Calculate the total number of transformants to determine if the screening was saturating.

When using Y187, the transformation efficiency should be at least 10⁵ per µg library plasmid. The efficiency can be dependent on the type of agar used in the CM plates. Good choices of agar include Bacto Agar (Difco) and Micro Agar (Duchefa Biochemie).

7. Grow the library screening plates from step 5 for up to 2 weeks in a 30°C incubator.

During this period, positives will appear as faster-growing colonies among the background.

Depending on the amount of leaky expression of the HIS3 reporter and concentration of 3-AT used in the CM plates, nonpositive yeast cells can still grow at a slow rate and form very small colonies. True positives should be clearly distinguished by their larger colony size; otherwise the 3-AT concentration should be raised.

8. Carefully pick up His+ positive colonies and restreak them on fresh plates with the same medium as used in step 5 and grow for 3 to 10 days in a 30°C incubator.

Usually, restreaked positives give clearly visible colonies within a few days, but the growth of yeast can be much reduced if the expression of the particular cDNA clone is somewhat toxic. Some colonies on the library plates may appear like positives because the local plating density was too high. Restreaking on fresh plates will show the potential validity.

ASSESSMENT OF HIS⁺ COLONIES: IDENTIFICATION OF TRUE POSITIVES

The number of His⁺ colonies that is obtained from a saturating screening may range from a few to hundreds, depending on the particular bait construct. The next challenge is to assess whether or not these contain a library clone encoding a transcription factor of interest. The strategy of how to discriminate most efficiently between true and false positives is somewhat dependent on the actual number of positives obtained, but basic guidelines are given in the procedure below. The nature of some of the known false positives is discussed in the commentary section of this unit. One optional step to identify a subset of common false positives is to assay the activity of the *GAL1_{UAS}-GAL1_{TATA}-lacZ* (β -galactosidase) reporter that host strain Y187 contains as a chromosomal integration. False positives encoding nonspecific DNA-binding proteins probably activate both the *lacZ* reporter and the bait-*HIS3* reporter. After this, usually the most straightforward way to continue is to amplify the library cDNA inserts by direct PCR (UNIT 15.1) on the yeast colonies, sequence the products (UNIT 15.2), and search for homology by BLAST analysis (UNIT 19.3). Promising clones should subsequently be rescued from yeast and retransformed to the bait reporter strain to validate the interaction in yeast and to rule out that a contaminant was amplified. Furthermore, the specificity for the bait sequence should be determined using other *HIS3* reporter constructs. If the binding site preferences of the transcription factor of interest are known, it is recommended that control reporters with mutant binding sites be included. Finally, gel shifts (UNIT 12.2), or other in vitro assays such as methylation or uracil interference (UNIT 12.3), or DNase footprinting (UNIT 12.4) have to be performed to confirm, by an independent method, that the library protein is a DNA-binding protein that can interact specifically with the bait sequence.

Materials

- His⁺ positive yeast colonies (see Basic Protocol 2)
- Liquid complete minimal medium (CM; UNIT 13.1) minus leucine (CM –Leu)
- 0.9 M sorbitol/50 mM EDTA, pH 8.0; sterilize by autoclaving
- 750 U/mg lyticase from *Arthrobacter luteus* (750 U/mg from Sigma, L-4025)
- 1 \times TE buffer, pH 7.5 (see recipe)
- pINT1-HIS3 reporter yeast strain(s) (see Basic Protocol 1)
- Empty cDNA library vector
- CM –Leu plates (UNIT 13.1) prepared with Difco Bacto-Agar or Duchefa Micro-Agar
- CM –Leu –His, plates with optimal concentration of 3-AT (see Basic Protocol 1, step 17; see recipe for 3-AT)
- 50-ml centrifuge tubes
- 30°C incubator, with and without shaking
- Low-speed tabletop centrifuge for 50-ml tubes (preferably swing-out)
- 1.5-ml microcentrifuge (e.g., Eppendorf tubes or equivalent)
- Additional reagents and equipment for β -galactosidase filter lift assay (UNIT 13.6), miniprep DNA isolation procedures (UNIT 1.6), transformation of *E. coli* (UNIT 1.8), restriction analyses (UNIT 3.1), gel electrophoresis (UNIT 2.5A), PCR amplification (UNIT 15.1), Southern blotting (UNIT 2.9A), hybridizations (UNIT 2.10), sequence analysis (UNIT 15.2), and in vitro binding assays (UNITS 12.2, 12.3 & 12.4)

Select initial putative positives for further analysis

1. When using Y187 as host strain, perform a colony filter lift assay for β -galactosidase activity (UNIT 13.6) and discard colonies that turn blue as false positives.

Not all false positives will turn blue, only those representing nonspecific DNA-binding proteins. Take care that a sequence similar or identical to the bait sequence is not, by

accident, also present in the GAL1 promoter of the lacZ construct, in which case the β -galactosidase assay would not be useful to determine specificity.

2. Perform PCR reactions directly on the His⁺ colonies (UNIT 15.1). Use a pipet tip to pick yeast cells from a colony and resuspend these in the PCR reaction mixture.

In the case where a large number of positives is obtained, try to reduce the number of clones to be analyzed. Organize the PCR products in smaller subgroups using Southern blots (UNITS 2.5A & 2.9A) in combination with restriction analyses (UNIT 3.1) and cross-hybridizations (UNIT 2.10). Select a representative of each group for further analysis.

Do not use wooden toothpicks to inoculate yeast cells for PCR amplification, since this interferes with the reaction.

3. Sequence the PCR products (UNIT 15.2) and run BLAST searches in GenBank or other databases in order to identify the clones (UNIT 19.3). Continue only with promising cDNAs.

See Anticipated Results for identities of commonly found false positives.

Isolate plasmid from yeast for further analysis of putative positives

To continue the analysis of the putative positives, it is necessary to isolate the plasmids from yeast. Since the amount of isolated DNA is too low to conduct further experiments, the plasmids are first introduced into calcium- or electrocompetent *E. coli* cells for further amplification. The quality of the DNA isolate from yeast is of major concern. It is not uncommon that only a few *E. coli* colonies result. Apart from the relatively low amount of DNA isolate, the difficulty is that the isolates may contain a toxic contaminant that negatively interferes with *E. coli* transformations. One useful plasmid isolation procedure is described in UNIT 13.11. An efficient commercially available alternative is the yeast plasmid isolation Y-DER Kit (Pierce). In the steps below, the authors present another simple alternative, in which spheroplasts are produced first and subsequently used in a standard *E. coli* miniprep procedure.

4. Grow 10-ml overnight cultures of the His⁺ colonies in CM –Leu medium in 50-ml tubes in a 30°C incubator with shaking.

Do not grow His⁺ colonies in YAPD medium. Usually, the cDNA libraries are based on 2- μ vectors that easily segregate without selection pressure. Only transformants carrying stable ARS-CEN-based plasmids can be grown in rich medium.

5. Harvest the cells by centrifuging for 2 min at 2400 \times g in a low-speed swing-out tabletop centrifuge.
6. Resuspend the cells in 200 μ l of 0.9 M sorbitol/50 mM EDTA, pH 8.0 containing 4 mg/ml lyticase (freshly added from 750 U/mg stock) and transfer to a 1.5-ml microcentrifuge tube.
7. Incubate 1 hr at 30°C in order to produce spheroplasts.

A quick test for proper spheroplast formation is described in UNIT 13.7.

8. Carefully microcentrifuge the spheroplasts for 5 min at 3000 rpm.
9. Discard the supernatant and use the pellet to perform a standard alkaline lysis miniprep procedure as for *E. coli* (UNIT 1.6).
10. Dissolve the DNA in 20 μ l 1 \times TE buffer, pH 7.5, and transform 1 to 2 μ l DNA to *E. coli* cells using the methods described in UNIT 1.8.
11. Isolate plasmid DNA from the resulting *E. coli* colonies (UNIT 1.6) and analyze the plasmids using restriction digestions (UNIT 3.1) and gel electrophoresis (UNIT 2.5A).

Deletions in plasmids from yeast can occur; therefore it can be worthwhile to analyze at least two E. coli colonies from each transformation with a yeast DNA isolate.

Verify positive interactions and analyze specificity

12. Reintroduce the library plasmids into the pINT1-HIS3 reporter strain and verify that they activate reporter gene expression. Perform a transformation with empty library vector as a control.

To determine the specificity for the bait sequence, also transform the library plasmids to other reporter strains to check for activation of constructs lacking the upstream bait sequence, containing a mutant bait sequence, or containing an unrelated sequence.

13. For transformation, use CM, –Leu plates and restreak the colonies on CM –Leu –His medium with the same amount of 3-AT as used in the screening or on a 3-AT titration series.
14. Compare library plasmids versus the empty library vector control for their ability to confer growth on 3-AT.
15. Proceed with other methods to validate the interaction between the identified (putative) transcription factor and the bait sequence, such as in vitro DNA binding (e.g., *UNITS 12.2, 12.3 & 12.4*) and in vivo functional assays.

AUTOMATIC SUBCLONING OF λ PHAGES INTO PLASMID LIBRARIES

The most effective way to produce and handle cDNA expression libraries is via lambda (λ) phage vectors. Cloning of cDNA in λ phages is more efficient than in plasmid vectors and the subsequent transformation and amplification rounds of the phage in *E. coli* are facilitated. Several modern λ phage vectors allow the automatic subcloning of plasmids in specific *E. coli* strains. This process can be scaled up to create a plasmid library from the original phage library via in vivo mass excision in *E. coli*. The authors standardly use the λ ACTII/pACTII' vector system (Memelink, 1997) to produce cDNA libraries for one- and two-hybrid screenings. Automatic subcloning of λ ACTII libraries is via the Cre-lox system. Vector sequences flanked by lox sites are recombined by Cre protein expressed in *E. coli*, thereby releasing a plasmid circle. Since this step is critical, a procedure for the conversion to a plasmid library is given below. Other Cre-lox-based λ phages for production of one- or two-hybrid cDNA libraries are listed in Table 12.12.3. Alternatively, the HybriZAP/pAD-GAL4 vector system can be used with automatic subcloning via the ZAP system according to the manufacturer's instructions (Stratagene). However, an advantage of the λ ACTII/pACTII' system is that it confers a higher expression level of the one-hybrid protein in yeast.

Materials

- cDNA synthesis kit (e.g., ZAP-cDNA Synthesis Kit, Stratagene)
- λ ACTII-cDNA or equivalent expression library
- Cre protein-expressing *E. coli* (e.g., BNN132)
- Liquid LB medium (*UNIT 1.1*), supplemented with 0.2% maltose, 10 mM MgSO₄, and 50 mg/liter kanamycin (add from 1000-fold kanamycin stock)
- LB-Cb plates (see recipe)
- Liquid LB medium (*UNIT 1.1*) containing 200 mg/liter carbenicillin
- Commercial maxiprep kits (optional; e.g., Qiagen or Promega)
- 30°C and 37°C incubators, with and without shaker
- 1.5-ml microcentrifuge tubes (e.g., Eppendorf or equivalent)
- Sterile cell scrapers

SUPPORT PROTOCOL

DNA-Protein Interactions

12.12.13

Centrifuge with large buckets (e.g., Sorvall GSA or GS-3 or Beckman JA-10 rotor, or equivalent)

Additional reagents and equipment for handling λ phages (UNIT 1.11), miniprep (UNIT 1.6) and maxiprep (UNIT 1.7), DNA isolation procedures, gel electrophoresis (UNIT 2.5A), restriction analyses (UNIT 3.1), and quantitation of DNA via UV spectrophotometry (APPENDIX 3D)

Prepare unidirectional cDNA expression library

1. Prepare a unidirectional cDNA expression library using λ ACTII or an equivalent vector (e.g., *EcoRI-XhoI*-digested) using a cDNA synthesis kit according to the manufacturer's instructions.
2. Determine the phage titer of the cDNA library. Carefully follow the instructions described in UNIT 1.11.

It is recommended that glassware which has been sterilized for 4 hr at 140°C be used. This procedure completely removes detergent residues, to which phages are very susceptible.

Grow Cre protein-expressing cells and determine excision frequency

3. Grow a 10-ml culture of Cre protein-expressing *E. coli* BNN132 overnight at 37°C in LB medium supplemented with 0.2% maltose, 10 mM MgSO₄, and 50 mg/liter kanamycin.

*E. coli strain BNN132 (Elledge et al., 1991) originates from JM107 and was made lysogenic with λ KC from which the Cre protein and the kanamycin resistance marker are expressed. Other λ KC-containing *E. coli* strains can also be used if compatible with the phage vector for infection.*

4. Mix 10⁹ phages with 2 ml *E. coli* BNN132 culture and incubate for 30 min at 30°C without shaking.
5. Add 4 ml LB medium and further incubate for 1 hr at 30°C with shaking.
6. Determine the excision frequency by making serial dilutions (10²-, 10⁴-, 10⁶-fold) in LB medium and plating on LB-Cb plates. For this, only take a small aliquot and store the rest at 4°C. Incubate the plates overnight at 37°C.

The excision frequency usually reaches up to 50%, but lower frequencies (>10%, resulting in 10⁸ clones) are still acceptable. Usually, the original phage library will consist of in the order of 10⁶ independent recombinants. A 100-fold excess of clones in the excised plasmid pool will ensure that the complexity of the original library is maintained.

7. Analyze the plasmid content of a number (20 to 40) of random colonies using a miniprep procedure (UNIT 1.6), digestions with restriction enzymes (UNIT 3.1), and gel electrophoresis (UNIT 2.5A). Determine the percentage of clones with insertions and the cDNA size distribution.

Process phage-infected *E. coli* and plate library

8. If the excision frequency was sufficient, divide the previously stored phage-infected *E. coli* BNN132 culture (step 6) over eight 1.5-ml microcentrifuge tubes and microcentrifuge for 1 min to pellet the bacteria.
9. Decant the supernatant, resuspend each bacterial pellet in 1 ml liquid LB medium, and spread the content of each tube on a 24 × 24-cm LB-Cb plate. Grow overnight at 37°C.

A dense field of small colonies will result. The initial growth of the plasmid library on plates avoids underrepresentation of toxic cDNA clones. If the culture is directly grown in liquid medium, bacteria containing such clones could be competed out.

10. Scrape the colonies from the plates using sterile cell scrapers. Resuspend in 1 liter prewarmed LB medium (with 200 mg/liter carbenicillin) and grow for another 4 hr, shaking at 37°C.
11. Harvest the bacteria by centrifuging in a centrifuge with large buckets (e.g., 250- or 500-ml buckets) and isolate the DNA using maxiprep procedures (*UNIT 1.7*) or commercially available maxiprep kits. Determine the DNA concentration by UV spectrometry (*APPENDIX 3D*).
12. Proceed with the library transformations, following the instructions for library screening (see Basic Protocol 2).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

3-Amino-1,2,4-triazole (3-AT), 1 M

Prepare a 1 M stock solution of 3-AT (e.g., Sigma). Do not autoclave, but filter sterilize the solution. Store at 4°C. Add the 3-AT to the CM medium after the latter has cooled down to ~60°C. Store up to 1 year at 4°C.

LB-Cb plates

Solidify LB medium (*UNIT 1.1*) with 1.5% agar and supplement with 200 mg/liter carbenicillin (add from 1000-fold stock). Sterilize LB medium with 1.5% agar by autoclaving 20 min. Add carbenicillin after LB medium has cooled to 60°C. Pour into 24 × 24-cm square dishes (Nunc) and standard (9-cm) petri dishes. Store up to 3 months at 4°C.

Lithium acetate stock, 10×

Prepare 1 M lithium acetate, pH 7.5, adjust with acetic acid; sterilize by autoclaving 20 min. Store indefinitely at room temperature.

PEG stock, 50% (w/v)

Prepare 50% (w/v) polyethylene glycol (PEG 3,350; e.g., Sigma); sterilize by autoclaving 20 min

Store indefinitely at room temperature, in small portions in well-closed bottles, because transformation efficiencies are critically dependent on the correct PEG concentration.

TE buffer stock, 10×

100 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)

10 mM EDTA

Sterilize by autoclaving 20 min

Store indefinitely at room temperature

COMMENTARY

Background Information

To gain insight into the regulation of a gene, it is necessary to identify the transcription factors binding to regulatory sequence elements. Such *cis*-acting elements mostly have been identified in the 5' promoter regions of genes, but can also be located in introns, exons, or 3' untranslated regions. Binding of factors can result in activation or silencing of gene repression, depending on the developmental stage of

the cell or on internal or external signals. Using techniques like gel-shift assays, DNA methylation interference, or DNase footprinting, binding sites for transcription factors can be mapped. Often, these are short sequences of only 5 to 10 bp. Almost all promoters contain a TATA box close to the transcription start site. The TATA box is recognized by a complex of factors, the basal transcription machinery. In turn, this complex interacts with the transcrip-

tion factors that specifically bind to the *cis*-acting elements.

Traditionally, the cloning of transcription factors relied on biochemical strategies such as affinity chromatography. Furthermore, molecular biological methods have been developed that involve the screening of cDNA expression libraries using labeled oligonucleotides. In the past decade, newly developed yeast genetic selection methods have provided particularly useful alternatives for cloning of transcription factors. One of the possibilities is to complement yeast transcription factor mutants. An illustrative example is the cDNA cloning of a human CCAAT-binding protein via complementation of the yeast *hap2* mutant (Becker et al., 1991). With the availability of the complete yeast genome sequence, it has become relatively easy to construct yeast transcription factor mutants and to assess the feasibility of a complementation strategy to clone a heterologous factor. Obviously, the possibilities are limited to cloning of functionally conserved transcription factor genes. More generally applicable is the use of a yeast genetic selection strategy that has become known as the one-hybrid system. In fact, the one-hybrid system is a simplification of the two-hybrid approach for detection of protein-protein interactions (Fields and Song, 1989; *UNIT 20.1*). The one-hybrid concept (Fig. 12.12.1), relies on the detection of the interaction between a protein expressed from a cDNA library (prey) with a DNA sequence of interest (bait) fused to a reporter gene. Activation of this reporter gene is dependent on the interaction and can be selected for. The required cDNA libraries are designed such that the expressed proteins carry a strong transcriptional activation domain from a known factor (hence the term "one-hybrid"). Therefore, an interaction with the bait sequence will usually guarantee reporter gene activation. This makes the selection method suitable to clone all types of DNA-binding proteins, independent of their normal functional properties. It is also possible to apply a "non-hybrid" selection strategy (Grueneberg et al., 1992; Chan et al., 1993), but this limits cloning possibilities to transcription factors with an intrinsic activation domain that functions properly in yeast.

One-hybrid libraries are normally constructed with shuttle vectors containing the replication sequences and marker genes necessary for maintenance in both yeast and *E. coli*. The activation domain (AD) in the expressed hybrid proteins is usually from the

yeast Gal4p transcription factor, but there are also vectors where the *Herpes simplex* VP16 AD or the acidic B42 domain (*UNIT 20.1*; Clontech) is used. To facilitate nuclear import of the hybrid proteins, the ADs used in all vectors are equipped with an SV40 nuclear localization signal. As a frequent additional feature, an epitope domain, such as HA (hemagglutinin), is included to facilitate detection on western blots or other applications. Constitutive promoters, such as the *ADHI* promoter, are often used to drive cDNA expression, but alternatives are conditional expression systems, e.g., based on the galactose-inducible *GALI* promoter. All AD vectors designed for two-hybrid screenings are exchangeable with one-hybrid systems. Some AD vectors are available as lambda vector systems (Clontech, Stratagene), with which cDNA cloning is more efficient and it is easier to achieve and maintain high complexity. Described lambda phage vectors for AD-library construction can easily be converted into *E. coli*-yeast shuttle vectors by automatic subcloning. A selection of useful vectors is shown in Table 12.12.3. The authors and others have had good experiences with use of the λ ACTII vector (Memelink, 1997) for cloning of a wide variety of transcription factors from plants (Menke et al., 1999; Meijer et al., 2000; van der Fits et al., 2000) and human cell lines (P.B.F. Ouwkerk, unpub. observ.).

The most frequently used reporter for genetic selection procedures is the *HIS3* gene, operative in the biosynthesis of histidine and encoding the enzyme imidazoleglycerol-phosphate dehydratase. Detection of its activation is through complementation of the auxotrophic *his3* mutation. Usually, *HIS3* reporters contain the minimal *HIS3* promoter, including the TATA box and transcription start site (Grueneberg et al., 1992). There are several examples of successful application of the *HIS3* reporter in one-hybrid screenings (e.g., Wilson et al., 1991; Grueneberg et al., 1992; Wang and Reed, 1993). Alternative strategies make use of the *lacZ* (Li and Herskowitz, 1993) or green fluorescence (GFP) reporter genes (Display Systems Biotech; *UNIT 9.7C*). Furthermore, the use of *MEL1* as an alternative color marker was recently described (Melcher et al., 2000). Although detection of these reporters is easy, the use of *HIS3* as a reporter has important advantages in a genetic selection procedure. Because *HIS3* is a growth selection marker, detection of positives is very simple since these appear as clear colonies in a field of nongrowing or slowly growing cells. Screening of *lacZ*, *MEL1*,

or *GFP* activity requires that all the transformed cells grow into colonies. Therefore, the plating density must be far lower than when *HIS3* is used as reporter, which complicates the setup of saturating library screens. Another advantage is that the choice of *HIS3* offers a solution to possible problems of leaky reporter expression, also termed background activity. Leaky expression can occur when bait sequences are recognized by endogenous yeast transcription factors. Since the activity of the His3p enzyme can be competitively inhibited by 3-amino-1,2,4-triazole (3-AT), leaky *HIS3* expression can often be eliminated by a 3-AT titration, so that it remains possible to detect increased *HIS3* expression resulting from the specific binding of a library protein during the screening. With *lacZ*, *MEL1*, or *GFP* reporters, the distinction between specific and background expression is difficult, and other auxotrophic markers that could be used as reporters (*LEU2*, *ADE2*) are not titratable. A good alternative for the *HIS3* reporter is a dominant antibiotic marker such as G418, hygromycin, or chloramphenicol. Here, leaky expression can be eliminated by increasing the concentration of the selective compound. An example is the cloning of the transcription factor Nrf1 using a neomycin-resistance gene in combination with G418 selection (Chan et al., 1993). With some bait sequences, the occurrence of excessively high leaky expression, which needs to be eliminated, may be observed, which is the most important constraint on the application of the one-hybrid screening system.

Initially, one-hybrid screenings involved the use of two replicating vector systems, one for the AD library and the other for the reporter (e.g., Wilson et al., 1991; Li and Herskowitz, 1993; Wang and Reed, 1993). Afterwards, a switch was made to chromosomally integrated reporters, which has the advantage that the reporter gene copy number is constant and that DNA-protein interactions are studied within the context of the nucleosome (e.g., Clontech). The pHIS3/pINT1 vectors described by the authors (Meijer et al., 1998) provide a system for the integration of reporter genes in the yeast genome. In this system, a bait sequence is first cloned upstream of the *HIS3* gene in one of the reporter vectors pHIS3NB or pHIS3NX (Table 12.12.1 and Fig. 12.12.3). In the second step, the reporter construct is cloned in the integrative vector pINT1. Finally, the pINT1-*HIS3* reporter construct is integrated via double cross-over in the nonessential *PDC6* gene (YGR087c, Chr. VII) using selection for the

dominant *APT1* antibiotic marker. Advantages of the pHIS3/pINT1 system are discussed in the introduction to this unit, and examples of its application have been published (Menke et al., 1999; Meijer et al., 2000; van der Fits et al., 2000).

A further refinement of one-hybrid screenings lies in the use of dual or triple reporter strains to facilitate discrimination between true and false positives. The Y187 strain that the authors recommend be used in conjunction with the pHIS3/pINT1 reporter system provides the opportunity to use a dual reporter system. It contains an additional *GALI_{UAS}-GALI_{TATA}-lacZ* construct, thus, a reporter gene with an upstream sequence that is unrelated to the bait sequence. Based on the unwanted activation of this second reporter, it is possible to discard a subset of false positives, probably representing nonspecific DNA-binding proteins. The *lacZ* reporter in Y187 was integrated with the *URA3* marker and can be counterselected with 0.1% (v/v) 5-fluoro-orotic acid. If desired, it can therefore be replaced to make alternative dual-reporter strains, or another construct can be added to make a triple reporter. One possibility is addition of a construct containing the bait sequence upstream of another reporter gene that also has a different minimal promoter than that of the *HIS3* reporter (see Table 12.12.4). Such a reporter should be activated by true positives, and a negative interaction would select against false positives that activated the bait-*HIS3* reporter through sequences outside the bait itself. Another elegant alternative is to add a reporter with a mutant bait sequence, but this is only applicable if it is already known which nucleotides in the bait sequence are critical for its *cis*-acting function. A good control is also to retransform isolated library clones to a strain containing a mutant bait-*HIS3* reporter integrated via the pHIS3/pINT1 system, so that it is present in an exactly identical chromosomal context as the bait reporter. Furthermore, control transformations can be performed to strains containing pHIS3/pINT1-based reporters lacking the upstream bait sequence or containing an unrelated upstream sequence. The nature of false positives is diverse (see Anticipated Results) and it should be taken into account that even with various different control reporters, not necessarily all the false positives can be eliminated. Some may follow specificity criteria in yeast exactly as expected for true positives (Schouten et al., 2000). Therefore, it remains essential that

Table 12.12.4 A Selection of Useful Cloning, Expression and Reporter Vectors for Modified Screening Applications or Procedures

Vectors ^a	Features	Reference
pRS300 series ^{b,c}	Yeast markers <i>URA3</i> , <i>TRP1</i> , <i>LEU2</i> , <i>HIS3</i> , and <i>LYS2</i> . Replication via <i>ARS-CEN</i>	Sikorski and Hieter, 1989; Sikorski and Boeke, 1991
pRS400 series ^{b,c}	Yeast markers <i>URA3</i> , <i>TRP1</i> , <i>LEU2</i> , <i>HIS3</i> , <i>MET15</i> , <i>ADE2</i> , and <i>LYS2</i> . Integrating, <i>ARS-CEN</i> , and 2- μ m versions	Sikorski and Hieter, 1989; Christianson et al., 1992; Brachmann et al., 1998
pYC series ^d	Yeast markers <i>URA3</i> , <i>MET2-CA</i> , <i>G418</i> . <i>ARS-CEN</i> and 2- μ m versions. Counterselectable marker <i>PKA3</i> regulatable by conditional <i>CHAI</i> or <i>MET25</i> promoters	Olesen et al., 2000
YcpIF series ^c	Yeast markers <i>URA3</i> , <i>TRP1</i> , <i>LEU2</i> , <i>HIS3</i> . Replication via <i>ARS-CEN</i> . Expression cassettes for translational fusions, equipped with inducible <i>GALI</i> promoter	Foreman and Davis, 1994
p4XX series ^c	Yeast markers <i>URA3</i> , <i>TRP1</i> , <i>LEU2</i> , <i>HIS3</i> . <i>ARS-CEN</i> and 2- μ m versions. Expression cassettes for transcriptional fusions, equipped with the <i>CYC1</i> terminator and <i>CYC1</i> , <i>ADH1</i> , <i>GPD</i> , <i>TEF</i> , <i>MET25</i> , or <i>GALI</i> promoters	Mumberg et al., 1994, 1995
pMEL α / β 2, YIpMEL α / β 2 ^e	<i>MEL1</i> or <i>lacZ</i> reporter vectors. Integrating versions and <i>ARS-CEN</i> -replicating versions with <i>URA3</i> marker	Melcher et al., 2000
pSLF187K	<i>lacZ</i> reporter vector, <i>URA3</i> marker, replication via 2- μ m	Forsburg and Guarente, 1988
LacZi ^b	<i>lacZ</i> reporter, <i>URA3</i> marker. Integrative vector	Clontech

^aAll vectors replicate high-copy in *E. coli* and confer carbenicillin resistance.

^bComplete sequences are deposited in the NCBI database.

^cAvailable from the American Type Culture Collection (ATCC).

^dVectors and sequences available from <http://www.crc.dk/phys/pyc/index.htm>.

^eVectors and sequences are available from http://www.rz.uni-frankfurt.de/FB/fb16/mikro/euroscarf/data/km_rep.html.

a yeast result be followed up by an independent confirmation of the specific interaction.

Finally, several interesting variations of the one-hybrid concept have been developed. One system, which the authors call “bridge one-hybrid,” can, for example, be applied for cloning coactivators of transcription factors (Sieweke et al., 1996; Yu et al., 1997). Here, a transcription factor, specifically interacting with a target sequence upstream of the reporter gene, is used to attract library proteins that increase the reporter gene expression. This type of “interactor hunt” differs from the conventional two-hybrid approach in that there is a specific selection for proteins that interact when the transcription factor is bound to its own target sequence. Therefore, this would more likely reflect authentic interactions. Another system is a reverse one-hybrid approach that can detect library proteins capable of dissociating the interaction between a known factor and its target sequence (Vidal et al., 1996). Both the one-hybrid and reverse one-hybrid systems can also be modified to mutation screens for identifica-

tion of critical amino acid residues in functional domains of the transcription factor (Bush et al., 1996; Vidal et al., 1996). A last and particularly challenging variation is to express an AD hybrid of a known transcription factor and screen for its target sites. This requires a reporter gene library with genomic DNA fragments or random oligonucleotides upstream of the reporter gene, and a procedure to discriminate between true positive target sequences and false-positive sequences that are activated by endogenous yeast factors (Wilson et al., 1991; Liu et al., 1993; Meijer et al., 1998).

Critical Parameters and Troubleshooting

A critical step in setting up yeast one-hybrid screenings is how to deal with the leaky *HIS3* expression that can occur due to recognition of the bait sequence by endogenous yeast transcription factors. Whether or not a particular reporter construct will be activated by a yeast transcription factor is difficult to predict in advance, but it is known that several *cis*-acting

sequences of plant or animal genes can also be recognized by yeast factors. Nevertheless, it may still be possible to screen in yeast for the corresponding plant or animal factors. One-hybrid screenings have even been used to clone factors from yeast itself (Li and Herskowitz, 1993; Kunoh et al., 2000). In recent years, a wide variety of transcription factors from plants (e.g., Kim et al., 1997; Menke et al., 1999; Meijer et al., 2000; van der Fits et al., 2000) or animals (e.g., Wang and Reed, 1993; Wei et al., 1999) have been cloned with a range of promoters and binding sites. Therefore, it does not seem that the one-hybrid approach is limited to cloning of particular transcription factor classes.

With the lack of knowledge to predict in advance whether a certain promoter will be activated in yeast, the best and fastest strategy is still trial and error. The authors have used promoters over 1 kb in length that needed to be titrated with only 5 mM 3-AT. On the other hand the authors have experienced the situation whereby occasionally even a short sequence could confer so much *HIS3* expression that this could not be inhibited with 3-AT up to 50 mM or higher, making its use in a screening impossible. Obviously, with longer promoters, the risk that such a strong enhancer is present becomes increased. Therefore, it is advised that several lengths or subfragments be tested. Screening conditions will be optimal if concentrations of 3-AT up to 25 mM suffice to reduce the *HIS3* expression to an acceptable minimum. Higher concentrations of 50 up to 100 mM can still be tried, but the growth of the positives will also be slowed considerably, and it will be more difficult to distinguish these from the background. If the background is unworkably high, it may be useful to test the reporter in other yeast strains (Table 12.12.2), because the level of leaky expression can depend considerably on the genotype. Furthermore, the use of a yeast TF mutant as host strain could be considered, if clues exist about the identity of the endogenous yeast factor that might cause the background problem. If there are no leaky expression problems, it can be useful to make multimers of the bait sequence. This can increase the sensitivity of the detection of positive interactions in the library screening. The authors have designed a stepwise unidirectional multimerization procedure (Ouwerkerk and Memelink, 1997) to tetramerize sequences ranging from 9 to 250 bp, and have used these successfully in one-hybrid screenings (e.g., Meijer et al., 2000).

It is important to keep track of the number of clones that have been screened to be able to decide at what stage the library screening has been saturating. This is dependent on the complexity of the primary cDNA library. Furthermore, discrimination between true and false positives is an important factor. Isolated library clones should be retransformed to the bait-*HIS3* strain, and preferably also to different strains with negative control reporters (see Background Information). The nature of common false-positive cDNA clones is discussed under Anticipated Results.

Finally, it is important to be aware of possible bacterial contamination in a yeast screening. Occasionally, bacterial contaminants can be very misleading because they can appear in close association with the yeast cells. Therefore, seemingly normal yeast colonies can occur on selective medium, while their capacity to grow actually results from bacterial cross-feeding rather than from expression of the auxotrophic marker gene. The presence of bacteria in suspected yeast colonies can easily be spotted with phase-contrast microscopy. Elimination is possible by streaking the colonies on CM –Leu –His plates supplemented with carbenicillin (50 to 200 mg/liter) and kanamycin (100 mg/liter)—antibiotics which do not affect the growth of yeast.

Anticipated Results

It is difficult to predict how many positive clones are to be expected from a saturating library screening. This is very much dependent on the bait sequence as well as on the cDNA library source. First of all, the frequency of false positives is extremely variable. With some bait sequences, these can represent the majority of the clones obtained, whereas nearly all positives in other screenings may be specific transcription factors. Furthermore, the frequency of isolation of a certain clone is no indication for its validity. This frequency is determined by factors such as the abundance in the mRNA population from which the library was derived and the stability in *E. coli* and yeast. If no positives are obtained at all, it is useful to change the cDNA library. However, it is also possible that a library protein that could potentially interact with the bait sequence was present, but is degraded in yeast or sequestered due to interaction with a certain yeast protein. Another cause for an unsuccessful result may be that the bait sequence is masked or repressed due to binding of a yeast transcription factor. For this reason, it may be worthwhile to try

different lengths or subfragments of the bait sequence. In the case where a very large number of positives are isolated, it is useful to attempt to sort them into classes, e.g., based on restriction or cross-hybridization patterns.

Several groups of possible false positives can be discriminated. The first group includes, among others, nucleoside transporters, ribosomal proteins, RNA binding proteins, thymine glycosylases, amino-acyltransferases, and histones (P.B.F. Ouwwerkerk, unpub. observ.; van der Fits, 2000). These factors all have in common the fact that their normal function is related to interactions with nucleic acids. As AD hybrids, such proteins are likely to turn into artifactual transcriptional activators.

A second group of artifacts are components of signal transduction pathways such as phosphatases, kinases (Schouten, 1999), and GTPases (van der Fits, 2000), or of proteolysis pathways. Strikingly, their possible occurrence is very much dependent on the bait sequence used. Therefore, it is most likely that such proteins interfere with signal transduction pathways in yeast in such a way that specific activation occurs of endogenous yeast transcription factors that recognize the bait sequence. An exemplifying case is the search for plant factors regulating the *T-cyt* gene of the tumor-inducing soil bacterium *Agrobacterium tumefaciens* (Schouten et al., 2000). Screenings in yeast resulted in several library clones that were able to activate *HIS3* expression via a wild-type bait sequence but not via a mutant sequence, suggesting that the interaction was specific. However, instead of DNA-binding proteins, the discovered clones turned out to represent plant proteins homologous with the yeast Skp1 protein that is involved in ubiquitin-mediated protein degradation. The following mechanism was proposed for the interference of the Skp1 homologs with reporter gene activation. The Skp1-like proteins would interact with the yeast F-box protein, Grr1p. This contact would result in the destabilization of Grr1p, which in turn would inactivate the yeast repressor protein Mig1p. The Mig1p repressor interacted with the bait sequence and consequently its inactivation could make this sequence accessible to another yeast factor, resulting in activation of the reporter gene. Because both the Mig1p repressor and the yet unidentified yeast activator appeared to have the same binding specificity as the searched plant *T-cyt* binding factor, the identification of this plant factor via yeast screening was impossible.

The third group of false positives are His3p homologs, which can complement the mutant *his3* locus in the yeast genome. These exist in plants and fungi, but not in animals. In the authors' experience, their frequency of isolation in one-hybrid screenings is low.

The fourth and last identified category of artifacts is miscellaneous and includes proteins involved in metabolic pathways, such as L-ascorbate peroxidase, pectinesterase, and others (van der Fits, 2000). Their occurrence is enigmatic and it is not very likely that these proteins directly activate reporter gene expression. The fact is that somehow they are beneficial for the yeast cells to survive under the selection conditions. It might be that they are involved in sequestering or degrading the His3p inhibitor 3-AT, due to which titration of leaky *HIS3* reporter expression would fail and a false positive colony would grow.

The protein-coding sequences of isolated cDNAs can be either partial or full-length. Preceding leader sequences can be translated to form a spacer between the AD and the library protein. Unexpectedly, sometimes positive cDNA clones are found for which the encoded protein sequence is not in frame with the AD sequence. Such clones can still represent true positives, as demonstrated by the following example. In a one-hybrid screening in yeast strains that contained reporters with upstream recognition sites for homeodomain proteins, the authors identified cDNAs from rice encoding homeodomain-leucine zipper (HD-Zip) proteins of two different families, I and II (Meijer et al., 2000). Strikingly all the family II members were in frame with the AD, but all the family I members were not. The authors concluded that the reporters must have been activated due to the fact that truncated HD-Zip I proteins were expressed, which possessed an intrinsic activation function. Subsequently, in plant experiments, it was indeed confirmed that HD-Zip I proteins could act as activators, whereas the HD-Zip II proteins turned out to be repressors. Since none of the isolated HD-ZIP I clones were in-frame fusions, in this case AD-hybrids might have been toxic, or the combination of Gal4p AD and intrinsic AD might have been incompatible.

Time Considerations

When the required pHIS3NB/NX and pINT1 reporter constructs have been cloned, it will take ~1 week to transform yeast and to select for G418-resistant pINT1-HIS3 reporter strains. The next week, the strains should be

tested on a range of 3-AT concentrations in order to determine the selective conditions for the actual library screening. Subsequently, it is useful to invest 1 to 2 weeks to fine tune the 3-AT concentration during pilot library transformations. Unless the transformation efficiency is suboptimal, all the transformations necessary to perform a saturating library screening, including the media preparation, can be finished within 1 week. The screening plates are incubated for 1 to 2 weeks during which positives can already be restreaked. As described in the protocol, from this point onward several strategies can be followed to discriminate between true and false positives, and to confirm the specificity of putative positives. Dependent on the numbers of positives and their nature, these analyses can take several weeks to many months.

Literature Cited

- Becker, D.M., Fikes, J.D., and Guarente, L. 1991. A cDNA encoding a human CCAAT-binding protein by functional complementation in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 88:1968-1972.
- Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P., and Boeke, J.D. 1998. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* 14:115-132.
- Bush, S.M., Folta, S., and Lannigan, D.A. 1996. Use of the yeast one-hybrid system to screen for mutations in the ligand-binding domain of the estrogen receptor. *Steroids* 61:102-109.
- Chan, J.Y., Han, X.-L., and Kan, Y.W. 1993. Cloning of Nrf1, an NF-E2-related transcription factor, by genetic selection in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 90:11371-11375.
- Christianson, T.W., Sikorski, R.S., Dante, M., Shero, J.H., and Hieter, P. 1992. Multifunctional yeast high-copy-number shuttle vectors. *Gene* 110:119-122.
- Elledge, S.J., Mulligan, J.T., Ramer, S.W., Spottswo, M., and Davis, R.W. 1991. λ YES: A multifunctional cDNA expression vector for the isolation of genes by complementation of yeast and *Escherichia coli* mutations. *Proc. Natl. Acad. Sci. U.S.A.* 88:1731-1735.
- Fields, S. and Song, O. 1989. A novel genetic selection system to detect protein-protein interaction. *Nature* 340:245-246.
- Foreman, P.K. and Davis, R.W. 1994. Cloning vectors for the synthesis of epitope-tagged, truncated and chimeric proteins in *Saccharomyces cerevisiae*. *Gene* 144:63-68.
- Forsburg, S.L. and Guarente, L. 1988. Mutational analysis of upstream activation sequence 2 of the *CYCL1* gene of *Saccharomyces cerevisiae*: A *HAP2-HAP3*-responsive site. *Mol. Cell. Biol.* 8:647-654.
- Gietz, D., St. Jean, A., Woods, R.A., and Schiestl, R.H. 1992. Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.* 20:1425.
- Grueneberg, D.A., Natesan, S., Alexandre, C., and Gilman, M.Z. 1992. Human and *Drosophila* homeodomains proteins that enhance the DNA-binding activity of serum response factor. *Science* 257:1089-1095.
- Hadfield, C., Fordan, B.E., Mount, R.C., Pretorius, G.H.J., and Burak, E. 1990. G418 resistance as a dominant marker and reporter for gene expression in *Saccharomyces cerevisiae*. *Curr. Genet.* 18:303-313.
- Harper, J.W., Adami, G.R., Wei, N., Keyomarsi, K., and Elledge, S.J. 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* 75:805-16.
- Kim, S.Y., Chung, H.-W., and Thomas, T. 1997. Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the *Dc3* promoter using a modified yeast one-hybrid system. *Plant J.* 11:1237-1251.
- Kunoh, T., Kaneko, Y., and Harashima, S. 2000. YHP1 encodes a new homeoprotein that binds to the IME1 promoter in *Saccharomyces cerevisiae*. *Yeast* 16:439-449.
- Leuther, K.K. and Johnston, S.A. 1992. Nondissociation of GAL4 and GAL80 in vivo after galactose induction. *Science* 256:1333-1335.
- Li, J.J. and Herskowitz, I. 1993. Isolation of *ORC6*, a component of the yeast origin recognition complex by a one-hybrid system. *Science* 262:1870-1874.
- Liu, J., Wilson, T.E., Milbrandt, J., and Johnston, M. 1993. Identifying DNA-binding sites and analyzing DNA-binding domains using a yeast selection system. In *Methods: A Companion to Methods in Enzymology*. Vol. 5 pp. 125-137. Academic Press, New York.
- Meijer, A.H., Ouwkerk, P.B.F., and Hoge, J.H.C. 1998. Vectors for transcription factor cloning and target site identification by means of genetic selection in yeast. *Yeast* 14:1407-1415.
- Meijer, A.H., de Kam, R.J., d'Erfurth I., Shen W., and Hoge, J.H.C. 2000. HD-Zip proteins of families I and II from rice: Interactions and functional properties. *Mol. Gen. Genet.* 263:12-21.
- Melcher, K., Sharma B., Ding W.V., and Nolden, M. 2000. Zero background yeast reporter plasmids. *Gene* 247:53-61.
- Memelink, J. 1997. Two yeast/*Escherichia coli* lambda/plasmid vectors designed for yeast one- and two-hybrid screens that allow directional cDNA cloning. Technical Tips Online TT01111 (<http://research.bmn.com/>). *Trends Genet.* 13:376.
- Menke, F.L., Champion, A., Kijne, J.W., and Memelink, J. 1999. A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene *Str* interacts with a jasmonate- and elicitor-inducible AP2-domain

- transcription factor, ORCA2. *EMBO J.* 18:4455-4463.
- Mumberg, D., Müller, R., and Funk, M. 1994. Regulatable promoters of *Saccharomyces cerevisiae*: Comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Res.* 14:7861-7871.
- Mumberg, D., Müller, R., and Funk, M. 1995. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 156:119-122.
- Olesen, K., Johannesen, P.F., Hoffmann, L., Sørensen, S.B., Gjermansen, C., and Hansen, J. 2000. The pYC plasmids, a series of cassette-based yeast plasmid vectors providing means of counter-selection. *Yeast* 16:1035-1043.
- Ouwerkerk, P.B.F. and Memelink, J. 1997. A simple method for directional multimerization of DNA sequences. Technical Tips Online T01051(<http://research.bmn.com/>)*Trends Genet.* 13:207.
- Schouten, J. 1999. Identification of *Arabidopsis thaliana* cDNAs coding for proteins that can activate gene expression in yeast via the *cyt-1* element of the *Agrobacterium tumefaciens* T-*cyt* gene promoter. Ph.D. Thesis, Leiden University, The Netherlands.
- Schouten, J.S., de Kam, R.J., Fetter, K., and Hoge, J.H.C. 2000. Overexpression of *Arabidopsis thaliana* SKP1 homologues in yeast inactivates the Mig1 repressor by destabilising the F-box protein Grr1. *Mol. Gen. Genet.* 263:309-319.
- Sieweke, M.H., Tekotte, H., Frampton, J., and Graf, T. 1996. MafB is an interaction partner and repressor of Ets-1 that inhibits erythroid differentiation. *Cell* 85:49-60.
- Sikorski, R.S. and Boeke, J.D. 1991. In vitro mutagenesis and plasmid shuffling: From cloned gene to mutant yeast. In *Methods of Enzymology*, Vol. 194: Guide to Yeast Genetics and Molecular Biology (C. Guthrie and G.R. Fink, eds.) pp. 302-318. Academic Press, San Diego.
- Sikorski, R.S. and Hieter, P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* 122:19-27.
- Singh, H., LeBowitz, J.H., Baldwin, A.S., Jr., and Sharp, P.A. 1988. Molecular cloning of an enhancer binding protein: Isolation by screening of an expression library with a recognition site DNA. *Cell* 52:415-423.
- van der Fits, L. 2000. Transcriptional regulation of stress-induced plant secondary metabolism. Ph.D. Thesis, Leiden University, The Netherlands.
- van der Fits, L., Zhang, H., Menke, F.H.M., Deneka, M., and Memelink, J. 2000. A *Catharanthus roseus* BPF-1 homologue interacts with an elicitor-responsive region of the secondary metabolite biosynthetic gene *Str* and is induced by elicitor via a JA-independent signal transduction pathway. *Plant Mol. Biol.* 44:675-685.
- Vidal, M., Brachmann, R.K., Fattaey, A., Harlow, E., and Boeke, J.D. 1996. Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. *Proc. Natl. Acad. Sci. U.S.A.* 93:10315-10320.
- Wang, M. and Reed, R.R. 1993. Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. *Nature* 364:121-126.
- Wei, Z., Angerer, R.C., and Angerer, L.M. 1999. Identification of a new sea urchin ets protein, Spets4, by yeast one-hybrid screening with the hatching enzyme promoter. *Mol. Cell. Biol.* 19:1271-1278.
- Wilson, T.E., Fahrner, T.J., Johnston, M., and Milbrandt, J. 1991. Identification of the DNA binding site for NGFI-B by genetic selection in yeast. *Science* 252:1296-1300.
- Yu, Y., Li, W., Su, K., Yussa, M., Han, W., Perrimon, N., and Pick, L. 1997. The nuclear hormone receptor Ftz-F1 is a cofactor for the *Drosophila* homeodomain protein Ftz. *Nature* 385:552-555.

Key References

Meijer et al., 1998. See above.

First description of the pHIS3/pINT1 vector system for yeast reporter strain construction.

Internet Resources

<http://www.atcc.org/>

American Type Culture Collection(ATCC): source for yeast vectors and strains.

<http://www.ncbi.nlm.nih.gov/>

National Center for Biotechnology Information (NCBI, Genbank): public database of genes and vectors sequences, computational tools for sequence analysis

<http://www.clontech.com>

<http://www.stratagene.com/>

Commercial sources for vectors, AD-hybrid libraries, yeast strains, and supplies.

Contributed by Pieter B.F. Ouwerkerk and
Annemarie H. Meijer
Leiden University
Leiden, The Netherlands

CHAPTER 13

Yeast

INTRODUCTION

Saccharomyces cerevisiae (baker's yeast) and *Schizosaccharomyces pombe* (fission yeast) are often considered to be model eukaryotic organisms, in a manner analogous to *Escherichia coli* as a model prokaryotic organism. Both yeasts have been extensively characterized and their genomes completely sequenced. They are as easy to grow as other microorganisms, and they have a haploid nuclear DNA content only 3.5 times that of *E. coli*. However, despite the small genome sizes, these yeasts display most of the features of higher eukaryotes. The fact that many cellular processes are conserved among different eukaryotic species—combined with the powerful genetic and molecular tools that are available—has made these yeasts important experimental organisms for a variety of basic problems in eukaryotic molecular biology.

Primarily for historical reasons, most studies on yeast have involved *Saccharomyces cerevisiae* (hereafter termed yeast). Culturing yeast is simple, economical, and rapid, characterized by a doubling time of ~90 min on rich medium. In addition, yeast has been well adapted to both aerobic and anaerobic large-scale culture. Cells divide mitotically by forming a bud, which pinches off to form a daughter cell. The progression through the cell cycle can be monitored by the size of the bud; this has been used to isolate a large collection of mutants (called *cdc* mutants) that are blocked at various stages of the cell cycle. Since yeast can be grown on a completely defined medium (see UNIT 13.1), many nutritional auxotrophs have been isolated. This has not only permitted the analysis of complex metabolic pathways but has also provided a large number of mutations useful for genetic analysis.

Yeast can exist stably in either haploid or diploid states. A haploid cell can be either of two mating types, called **a** and α . Diploid **a**/ α cells—formed by fusion of an α cell and an **a** cell (UNIT 13.2)—can grow mitotically indefinitely, but under conditions of carbon and nitrogen starvation will undergo meiosis. The meiotic products, called *spores*, are contained in a structure called an *ascus*. After gentle enzymatic digestion of the thick cell wall of the ascus, the haploid spore products can be individually isolated and analyzed (UNIT 13.2). This ability to recover all four products of meiosis has allowed detailed genetic studies of recombination and gene conversion that are not possible in most other eukaryotic organisms. The existence of stable haploid and diploid states also facilitates classical mutational analysis, such as complementation tests and identification of both dominant and recessive mutations.

The haploid yeast cell has a genome size of about 15 megabases and contains 16 linear chromosomes, ranging in size from 200 to 2200 kb. Thus, the largest yeast chromosome is still 100 times smaller than the average mammalian chromosome. This small chromosome size, combined with the advent of techniques for cloning yeast genes as well as manipulating yeast chromosomes, has allowed detailed studies of chromosome structure. Three types of structural elements required for yeast chromosome function have been identified and cloned: origins of replication (*ARS elements*), centromeres (*CEN elements*), and *telomeres*. The cloning of these elements has led to the construction of artificial chromosomes that can be used to study various aspects of chromosome behavior, such as how chromosomes pair and segregate from each other during mitosis and meiosis. In addition, systems using artificial chromosomes have been designed that allow cloning of

larger contiguous segments of DNA (up to 400 kb) than are obtainable in other cloning systems. These structural elements, as well as cloned selectable yeast genes, have permitted the construction of yeast/*E. coli* shuttle vectors that can be maintained in yeast as well as in *E. coli* (UNITS 13.4 & 13.6).

Procedures for high-efficiency transformation of yeast (UNIT 13.7) have been available for nearly two decades, allowing cloning of genes by genetic complementation (UNITS 13.8 & 13.9). Because yeast has a highly efficient recombination system, DNAs with alterations in cloned genes can be reintroduced into the chromosome at the corresponding homologous sites (UNIT 13.10). This has permitted the rapid identification of the phenotypic consequences of a mutation in any cloned gene, a technique generally unavailable in higher eukaryotes. In addition, homologous recombination permits a wide variety of genetic techniques that have greatly facilitated the analysis of biological processes.

Despite its small genome size, yeast is a characteristic eukaryote, containing all the major membrane-bound subcellular organelles found in higher eukaryotes, as well as a cytoskeleton. Yeast DNA is found within a nucleus and nucleosome organization of chromosomal DNA is similar to that of higher eukaryotes, although no histone H1 is present. Three different RNA polymerases transcribe yeast DNA, and yeast mRNAs (transcribed by polymerase II) show characteristic modifications of eukaryotic mRNAs [such as a 5' methyl-G cap and a 3' poly(A) tail], although only a few *S. cerevisiae* genes contain introns. Transcriptional regulation has been extensively studied and at least one yeast transcriptional activator has been shown to function in higher eukaryotes as well. High-molecular-weight yeast DNA and RNA can be prepared fairly quickly (UNITS 13.11 & 13.12). Another characteristic of eukaryotes is the proteolytic processing of precursor proteins to yield functional products, which is often coupled to secretion. Yeast has several well-studied examples of secreted proteins and pheromones, and the large number of genes that have been identified as involved in protease processing and secretion suggests a highly complex pathway. Yeast protein extracts can be prepared using three different protocols (UNIT 13.13); the best choice will depend on the particular application. The ease and power of genetic manipulation in yeast facilitate the use of this organism to detect novel interacting proteins using the two-hybrid system or interaction trap (UNIT 20.1).

Although *Saccharomyces cerevisiae* is the most commonly studied yeast, *S. pombe* is also an important experimental organism (UNIT 13.14). Although both yeasts are unicellular microorganisms that grow in similar medium, they are evolutionarily quite distant. It has become increasingly clear that, in terms of molecular mechanisms, *S. pombe* is more similar to higher eukaryotic organisms than *S. cerevisiae*. Experimental manipulations in *S. pombe* are broadly similar to those in *S. cerevisiae*, although the technical details often differ. The chapter includes units on *S. pombe* relating to strain maintenance and media (UNIT 13.15), growth and genetic manipulation (UNIT 13.16), and introduction of DNA into cells (UNIT 13.17).

This chapter is written for the molecular biologist who has not previously worked with yeast. The glossary below introduces the terms of yeast molecular biology.

aerobic growth growth in the presence of oxygen, utilizing the Krebs cycle.

α and a factor mating type-specific polypeptides secreted by either α or a haploid cells, respectively, which interact with haploid cells of the opposite mating type to stimulate mating.

anaerobic growth growth in the absence of oxygen, utilizing fermentation (via glycolysis).

ARS elements DNA sequences present throughout the yeast genome that confer autonomous replication on plasmids in yeast; most of these sequences also function as chromosomal origins of replication as well.

ascus thick-walled sac containing the four haploid products, called spores, resulting from meiosis.

cdc mutants strains of yeast that exhibit stage-specific blocks of the cell cycle; these mutations define genes important in DNA replication, meiosis, and sporulation.

CEN element DNA sequences present at the centromere that ensure proper segregation of chromosomes during mitosis and meiosis, presumably by promoting interaction with the mitotic spindle.

cir⁺ strains of yeast that contain the naturally occurring 2 μ m plasmid.

cir^o yeast strains that have lost this endogenous 2 μ m plasmid.

δ element ~330-bp sequence, present as direct repeats at the ends of the transposable element Ty1, and also found dispersed throughout the genome.

gene disruption a mutation constructed in vitro in a cloned gene which, upon reintroduction into the genome at the homologous chromosomal site, results in inactivation of the gene function.

glusulase a digestive enzyme isolated from snails that breaks down thick cell walls of either an ascus to allow isolation of spore products, or a yeast cell to produce spheroplasts.

heterothallic common laboratory strains of yeast which—due to a mutation in the *HO* gene—stably maintain a given allele at the *MAT* locus.

homothallic strains of yeast (typically found in the wild) that, in a haploid state, rapidly interconvert the *MAT* locus, resulting in rapid switching between the α and **a** mating types; cultures of such strains rapidly diploidize.

killer strains strains of yeast that harbor a double-stranded RNA virus; such strains kill sensitive yeast strains via secretion of a protein toxin, to which killer strains are immune.

MAT the mating-type locus which is expressed and therefore determines the mating type of a haploid cell; this locus has two alleles—the *MAT_a* allele confers the **a** mating type, while *MAT α* specifies the α mating type.

meiosis the process by which the number of chromosomes present in a diploid cell is halved to yield haploid products.

mitosis vegetative cell division (of either haploid or diploid cells) in which the chromosome

number stays the same.

petites mutants of yeast (either nuclear or mitochondrial) with impaired mitochondria function; they grow as small colonies on fermentable carbon sources and are unable to grow on nonfermentable carbon sources.

schmoo a distinctive shape of a haploid cell (pear-shaped), induced by exposure to mating pheromone.

sporulation the end product of meiosis, induced by carbon and nitrogen starvation of a diploid cell, which results in four haploid progeny contained as spores within an ascus; this complicated developmental process requires over 200 genes.

telomeres DNA sequences found at the end of linear chromosomes that are essential for chromosome stability and complete replication; in *S. cerevisiae*, telomeres consist of tandem repeats of the sequence 5'dG₁₋₃dT3'.

Ty1 elements the primary transposable element found in yeast, which functions as a retrotransposon (via reverse transcription of its RNA and subsequent reinsertion into the genome).

UAS element upstream activating sequences in yeast promoters, to which regulatory proteins bind in order to enhance the rate of transcription.

zygote a morphologically distinct cellular structure formed by the fusion of two haploid cells of opposite mating type, which results in formation of a diploid.

Zymolyase β -glucanase, isolated from *Arthro-bacter luteus*, that hydrolyzes the yeast cell wall and is used to prepare spheroplasts for a variety of purposes.

Key References

- Watson, J.D., Hopkins, N.H., Roberts, J.W., Steitz, J.A., and Weiner, A.M. 1987. Yeasts as the *E. coli* of eukaryotic cells. *In* Molecular Biology of the Gene, Vol. 1, pp. 550-594. Benjamin/Cummings, Menlo Park, Calif.
- Strathern, J.N., Jones, E.W., and Broach, J.R. (eds.) 1981. The Molecular Biology of the Yeast *Saccharomyces*: Life Cycle and Inheritance. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Strathern, J.N., Jones, E.W. and Broach, J.R. (eds.) 1982. The Molecular Biology of the Yeast *Saccharomyces*: Metabolism and Gene Expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Victoria Lundblad and Kevin Struhl

BASIC TECHNIQUES OF YEAST GENETICS

SECTION I

Like *Escherichia coli*, yeast can be grown in either liquid media or on the surface of (or embedded in) solid agar plates. Yeast cells grow well on a minimal medium containing dextrose (glucose) as a carbon source and salts which supply nitrogen, phosphorus, and trace metals. Yeast cells grow much more rapidly, however, in the presence of protein and yeast cell extract hydrolysates, which provide amino acids, nucleotide precursors, vitamins, and other metabolites which the cells would normally synthesize *de novo*. During exponential or log-phase growth, yeast cells divide every 90 min when grown in such media. Log phase can be divided into three stages based on the rate of cell division (or the proportion of budded cells within a culture), which is in turn a function of the cell density of the culture. As cell density increases, nutrient supplies drop and the rate of cell division slows (the measurement of cell density, as well as techniques for the propagation and genetic manipulation of yeast, are described in *UNIT 13.2*). Early-log phase is the period when cell densities are $<10^7$ cells/ml. Mid-log phase cultures have densities between 1 and 5×10^7 cells/ml. Late-log phase occurs when cell densities are between 5×10^7 and 2×10^8 cells/ml. At a density of 2×10^8 cells/ml yeast cultures are said to be saturated and the cells enter stationary, or G_0 phase.

Recipes for media that are commonly encountered when working with yeast are provided in *UNIT 13.1*. The rich medium, YPD (yeast extract, peptone, dextrose; also called YEPD media), is most commonly used for growing *S. cerevisiae*. Additional recipes are provided for minimal and complete minimal dropout media, which are routinely used for testing mating type, selecting diploids, determining auxotrophic requirements (*UNIT 13.2*), and selecting for transformants (*UNIT 13.7*). Finally, recipes for media required in the more advanced techniques described in *UNITS 13.6* and *13.10* are presented.

Preparation of Yeast Media

UNIT 13.1

Preparation of sterile media of consistently high quality is essential for the genetic manipulation of yeast. Recipes for media needed in the protocols in this chapter are provided below. It is recommended that ingredients always be purchased from the same manufacturer. The following sources are recommended for specific ingredients (complete addresses and phone numbers are provided in *APPENDIX 4*):

J.T. Baker

dextrose
ammonium sulfate
potassium acetate
glycerol

Difco

agar (Bacto-agar)
peptone (Bacto-peptone)
yeast extract (Bacto-yeast extract)
yeast nitrogen base (YNB, *without*
amino acids or ammonium
sulfate)

Sigma

amino acids
nucleotide bases
canavanine
cycloheximide
L- α -amino adipic acid
galactose (with 0.01%
contaminating glucose)
potato starch

PCR, Inc

5-fluoroorotic acid

Autoclaving is usually carried out for 15 min at 15 lb/in², but times should be increased when large amounts of media are being prepared (20 min for 4 to 6 liters and 25 min for 6 to 12 liters).

*Saccharomyces
cerevisiae*

13.1.1

LIQUID MEDIA

Ingredients for liquid media are dissolved in water to 1 liter, mixed until completely dissolved, and autoclaved in 100- or 500-ml media bottles. Alternatively, liquid media can be filter sterilized, resulting in faster preparation, less caramelization (of dextrose), and faster growth of cells. Recipes for “premixes” are provided to minimize the number of materials that must be weighed each time media is prepared. When preparing premixes, break up any large chunks of dextrose before mixing with other components, then shake the container vigorously until contents are homogenized. It is convenient to make the premix in the empty plastic containers in which 2.5 kg of dextrose is packaged. Throughout this chapter, YNB –AA/AS refers to yeast nitrogen base without amino acids or ammonium sulfate. (See listing of recommended suppliers of ingredients in unit introduction above.)

Rich Medium

YPD medium

<i>Per liter:</i>	<i>Premix:</i>	<i>Final concentration:</i>
10 g yeast extract	250 g yeast extract	1% yeast extract
20 g peptone	500 g peptone	2% peptone
20 g dextrose	500 g dextrose	2% dextrose
	Use 50 g/liter	

This rich, complex medium—also known as YEPD medium—is widely used for the growth of yeast when special conditions are not required.

It is preferable to use a 20% (10×) solution of dextrose that has been filter sterilized or autoclaved separately (and added to the other ingredients after autoclaving) to prevent darkening of the media and to promote optimal growth.

Minimal Media

Minimal medium

<i>Per liter:</i>	<i>Premix:</i>	<i>Final concentration:</i>
1.7 g YNB –AA/AS	68 g YNB –AA/AS	0.17% YNB –AA/AS
5 g (NH ₄) ₂ SO ₄	200 g (NH ₄) ₂ SO ₄	0.5% (NH ₄) ₂ SO ₄
20 g dextrose	800 g dextrose	2% dextrose
	Use 27 g/liter	

This minimal medium—also known as synthetic dextrose (SD) medium—can support the growth of yeast which have no nutritional requirements. However, it is used most often as a basal medium to which other supplements are added (see CM dropout medium below).

Complete minimal (CM) dropout medium, per liter:

- 1.3 g dropout powder (Table 13.1.1)
 - 1.7 g YNB –AA/AS
 - 5 g (NH₄)₂SO₄
 - 20 g dextrose
- (Alternatively, replace last three ingredients with 27 g minimal medium premix)

CM dropout powder, also known as minus or omission powder, lacks a single nutrient but contains the other nutrients listed in Table 13.1.1. Complete minimal (CM) dropout medium is used to test for genes involved in biosynthetic pathways and to select for gene function in transformation experiments. To test for a gene involved in histidine biosynthesis one would determine if the yeast strain in question can grow on CM minus histidine (–His) or “histidine dropout” plates. It is convenient to make several dropout powders, each lacking a single nutrient, to avoid weighing each component separately for all the different dropout plates required in the laboratory.

Table 13.1.1 Nutrient Concentrations for Dropout Powders^a

Nutrient ^b	Amount in dropout powder(g) ^c	Final conc. in prepared media (µg/ml)	Liquid stock conc. (mg/100 ml) ^d
Adenine (hemisulfate salt)	2.5	40	500
L-arginine (HCl)	1.2	20	240
L-aspartic acid ^e	6.0	100	1200
L-glutamic acid (monosodium salt)	6.0	100	1200
L-histidine	1.2	20	240
L-leucine ^f	3.6	60	720
L-lysine (mono-HCl)	1.8	30	360
L-methionine	1.2	20	240
L-phenylalanine	3.0	50	600
L-serine	22.5	375	4500
L-threonine ^e	12.0	200	2400
L-tryptophan	2.4	40	480
L-tyrosine	1.8	30	180 ^f
L-valine	9.0	150	1800
Uracil	1.2	20	240

^aCM dropout powder lacks a single nutrient but contains the other nutrients listed in this table. Nomenclature in this manual refers to, e.g., a preparation that omits histidine as histidine dropout powder. Conditions from Sherman et al., 1979.

^bAmino acids not listed here can be added to a final concentration of 40 µg/ml (40 mg/liter).

^cGrind powders into a homogeneous mixture with a clean, dry mortar and pestle. Store in a clean, dry bottle or a covered flask.

^dUse 8.3 ml/liter of each stock for special nutritional requirements. Store adenine, aspartic acid, glutamic acid, leucine, phenylalanine, tyrosine, and uracil solutions at room temperature. All others should be stored at 4°C.

^eWhile these amino acids can be used reliably when included in autoclaved media, they supplement growth better when added after autoclaving.

^fUse 16.6 ml/liter for L-tyrosine nutritional requirement.

It may be preferable to use a 10× solution of dropout powder (i.e., 13 g of dropout powder in 100 ml water) that has been “sterilized” separately (and added to the other ingredients after autoclaving) to improve the growth rate in this medium.

Sporulation medium, per liter

10 g potassium acetate (1% final)

1 g yeast extract (0.1% final)

0.5 g dextrose (0.05% final)

This nitrogen-deficient “starvation” medium contains acetate as a carbon source to promote high levels of respiration, which induces diploid yeast strains to sporulate. Sporulation can be carried out in liquid media or on plates (see below and UNIT 13.2). If nutrients are required, add them at the concentrations listed in Table 13.1.1.

Alternative Carbon Sources

Wild-type yeast can use a variety of carbon sources other than glucose to support growth. These include galactose, maltose, fructose, and raffinose. In particular, galactose is often used to induce transcription of sequences fused to the *GAL10* promoter (UNIT 13.6). All are used at a concentration of 2% w/v (20 g/liter) and should be made by replacing dextrose in the recipe for minimal or complete minimal (CM) dropout media. For a nonfermentable carbon source—which will not support the growth of *petites* (cells lacking functional mitochondria; see glossary)—2% potassium acetate (w/v), 3% glycerol, 3% ethanol, or 2% glycerol and 2% ethanol (v/v each) can be used. YPA medium (2% potassium acetate,

Saccharomyces cerevisiae

13.1.3

2% peptone, and 1% yeast extract) is excellent for inducing high levels of respiration in cells prior to sporulation (UNIT 13.2).

SOLID MEDIA

Making solid media for yeast is—for the most part—no different from preparing plates for bacteriological work (see UNIT 1.1). For all plates, agar is added at a concentration of 2% (20 g/liter). A pellet of sodium hydroxide (~0.1 g) should be added per liter to raise the pH enough to prevent agar breakdown during autoclaving. In addition, add a stir bar to facilitate mixing after autoclaving.

After autoclaving, flasks are left for 45 to 60 min at room temperature until cooled to 50° to 60°C. (Drugs and other nutrients are added after 30 min at room temperature.) Just prior to pouring, put the flask on a stir plate at medium to high speed and mix until contents are homogeneous (~5 min). After pouring, a few bubbles can be removed from the agar surface by passing the flame of a Bunsen burner lightly over the surface of the molten agar (“flaming” the plates). One liter of media will yield 30 to 35 plates.

While a specific brand of petri plate is not required, we recommend Fisher plates (100 × 15 mm, #8-757-12), which have ridges around the tops of the covers to allow easy stacking, making plate pouring less cumbersome.

When preparing the plate recipes below, follow the general guidelines for mixing and autoclaving in the introduction to liquid media (p. 13.1.2). Most plates can be stored at room temperature for ≤4 months. Plates containing drugs (cycloheximide, 5-fluoroorotic, canavanine, and L- α -amino adipic acid) or Xgal are stable for 2 to 3 months when stored at 4°C.

Minimal Plates and Rich Plates

YPD, minimal, and CM dropout plates

Per liter: Follow recipes for liquid media above, adding 20 g agar and a pellet of NaOH.

Premixes: Follow recipes for liquid media premixes above, adding 500 g agar for YPD premix and 800 g agar for minimal premix. To prepare plates, add one NaOH pellet and the following amounts of premix (per liter):

YPD plates—70 g YPD plate premix

Minimal plates—47 g minimal plate premix

CM dropout plates—47 g minimal plate premix + 1.3 g dropout powder

Specialty Plates

The recipes for α -amino adipate, canavanine, and cycloheximide plates are included even though no specific use for them is described in this chapter. They are commonly used in negative selection experiments in the same way that 5-FOA plates are used (see below). One can select against the wild-type *LYS2*, *CAN1*, and *CYH2* genes by growth on plates that select for cycloheximide, α -amino adipate, or canavanine resistance, respectively (for review, see Brown and Szostak, 1983).

5-fluoroorotic acid (5-FOA) plates

To a 2-liter flask (containing a stir bar), add:

1 g 5-FOA powder

500 ml H₂O

5 ml 2.4 mg/ml uracil solution

Stir with low heat ~1 hr until completely dissolved; filter sterilize

To a separate 2-liter flask, add:

1.7 g YNB –AA/AS

5 g (NH₄)₂SO₄

20 g dextrose

20 g agar

1.3 g uracil dropout premix (Table 13.1.1)

H₂O to 500 ml and autoclave

(Alternatively, replace first four ingredients with 47 g minimal plate premix)

When the molten agar cools to ~65°C, gently add the sterile 5-FOA/uracil solution to the uracil dropout medium by pouring it down the inside wall of the flask containing the latter. Swirl gently to mix and pour the plates.

URA3⁺ strains are unable to grow on media containing the pyrimidine analog 5-fluoroorotic acid (Boeke et al., 1984). This observation has led to methods that use 5-FOA to select against the functional URA3 gene (see UNIT 13.10). This type of selection—termed “negative selection” (Brown and Szostak, 1983)—can also be used to select against the wild-type LYS2, CAN1, and CYH2 genes.

5-FOA is quite expensive, and plates should be used sparingly. The material is prepared in bulk and is substantially discounted for members of the Genetics Society of America (Bethesda, Md.).

Xgal plates, per liter

1.7 g YNB –AA/AS

5 g (NH₄)₂SO₄

20 g dextrose

20 g agar

0.8 g dropout powder (omitting appropriate amino acids; see Table 13.1.1)

NaOH pellet

(Alternatively, replace first four ingredients with 47 g minimal plate premix)

Add H₂O to 900 ml and autoclave. Add 100 ml of 0.7 M potassium phosphate, pH 7.0, and 2 ml of 20 mg/ml Xgal prepared in 100% *N,N*-dimethylformamide (stored as frozen stock; see Table 1.4.2).

Dissection plates

Follow the recipe for YPD plates, keeping in mind that dissection plates used for tetrad analysis should be of uniform thickness and free of imperfections. After the plates have been poured they should be flamed. Stack the plates on a level surface in piles of six and move gently in a circular motion to “even out” the agar. Certain batches of agar produce plates that have microscopic precipitates embedded in the agar, often looking much like yeast spores. If this occurs, an agar of higher purity can be used, such as Noble agar (Difco) or agarose.

α-aminoadipate plates

1.7 g YNB –AA/AS

20 g dextrose

20 g agar

H₂O to 1 liter

Add ingredients to a 2-liter flask and autoclave. When the molten agar cools to ~65°C, add 34 ml of a solution of 6% L-α-aminoadipic acid (prepared by dissolving α-aminoadipate in 100 ml water and adjusting the pH to ~6.0 with 1 M KOH). The final concentration of α-aminoadipic acid in the plates should be 0.2%. Swirl gently to mix and pour the plates.

Lys2⁻ yeast use α-aminoadipic acid as an alternate nitrogen source. Since yeast can use certain amino acids as nitrogen sources, only those amino acids which are required by the particular strain being used should be added to these plates. These can be added from liquid

Saccharomyces cerevisiae

13.1.5

stocks prior to autoclaving. When using this media for isolating *lys2⁻* yeast, lysine must be added at 30 µg/ml.

Canavanine plates

Follow recipe for complete minimal (CM) dropout plates, omitting the nutrient arginine. When the 1-liter autoclaved solution has cooled to ~65°C, add 10 ml of 6 mg/ml filter-sterilized canavanine sulfate solution. The final concentration of canavanine should be 60 µg/ml.

The sterile canavanine sulfate solution can be stored frozen.

Cycloheximide plates

Follow the recipe for YPD plates. When the agar cools to ~65°C, add 1 ml of 10 mg/ml filter-sterilized cycloheximide solution. The final concentration of cycloheximide in the plates should be 10 µg/ml.

The sterile cycloheximide stock solution can be stored frozen.

STRAIN STORAGE AND REVIVAL

Yeast strains can be stored at -70°C in 15% glycerol (viable for >5 years), or at 4°C on slants consisting of rich medium supplemented with potato starch (viable for 1 to 2 years). Both methods are described below.

Preparation and Inoculation of Frozen Stocks

Make a solution of 30% (w/v) glycerol. Pipet 1 ml into 15 × 45-mm, 4-ml screwcap vials. Loosely cap the vials and autoclave 15 min.

To inoculate vials for storage, add 1 ml of a late-log or early-stationary phase culture, mix, and set on dry ice. Store at -70°C. Revive by scraping some of the cells off the frozen surface and streak onto plates. Do not thaw the entire vial. Cells can also be stored in the same way by adding 80 µl dimethyl sulfoxide (DMSO) to 1 ml cells (8% v/v) and storing at -70°C.

Preparation and Inoculation of Slants

1. For 250 slants, add the following ingredients to a 1-liter flask:

70 g YPD plate premix
20 mg adenine (hemisulfate salt)
20 g potato starch
H₂O to 500 ml

*Excess adenine prevents *ade⁻* mutations from being lost.*

2. Set the flask in a 4-liter beaker filled with 1 liter of water. Place the beaker on a heat-controlled, magnetic stirring apparatus and stir with the heat setting on high.

With smooth and continuous stirring, the contents should not burn and should be molten after ~1 hr.

3. With the entire setup in place, begin pipetting 2-ml aliquots into 15 × 45-mm, 4-ml screwcap vials. When all vials have been filled, put the caps on loosely, pack in the original boxes, and autoclave 15 min.
4. Lean the boxes against a support at an angle of ~70°. Allow the slants to dry 2 days and screw the caps on tightly.

Slants can be stored at room temperature for at least 6 months.

**BASIC
PROTOCOL**

**ALTERNATE
PROTOCOL**

**Preparation of
Yeast Media**

13.1.6

5. To inoculate a slant, smear cells from the flat end of a sterile toothpick onto the agar surface of the slant. Cap loosely and incubate 1 or 2 days at 30°C. After growth, screw the cap on as tightly as possible and store at 4°C.

Slants are a convenient way to store and mail strains. They can be mailed immediately after inoculating since sufficient growth will occur in transit. See another method for mailing strains below.

Mailing and Reviving Strains

Yeast strains can be conveniently mailed as slants. Alternatively, transfer cells to a piece of sterile Whatman 3MM paper by pressing the paper onto the desired yeast colony using forceps that have been dipped in ethanol and flamed. Wrap the paper in sterile aluminum foil and mail to recipient.

Revive the strain by placing the paper (yeast side down) on the surface of an agar plate. Incubate the plate at 30°C. A thick patch of yeast should be visible after lifting the paper.

LITERATURE CITED

- Boeke, J.D., LaCrute, F., and Fink, G.R. 1984. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoroorotic acid resistance. *Mol. Gen. Genet.* 197:345-346.
- Brown, P.A. and Szostak, J.W. 1983. Yeast vectors with negative selection. *Meth. Enzymol.* 101:278-290.
- Sherman, F., Fink, G.R., and Lawrence, C.W. 1979. *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Contributed by Douglas A. Treco
Massachusetts General Hospital
Boston, Massachusetts

Victoria Lundblad
Baylor College of Medicine
Houston, Texas

**BASIC
PROTOCOL**

*Saccharomyces
cerevisiae*

13.1.7

Growth and Manipulation of Yeast

Aside from different media requirements, yeast cells are physically manipulated essentially as described for bacterial cells—i.e., they are grown in liquid culture (in tubes or flasks) or on the surface of agar plates and are manipulated using the basic equipment described in *UNITS 1.1-1.3*. In addition, a well-equipped yeast laboratory requires static and shaking incubators dedicated to 30°C and a microscope with magnification up to 400×. A second microscope adapted for dissecting yeast tetrads is extremely valuable for the genetic analyses and strain constructions described in this unit. A small electric clothes dryer is indispensable when replica plating is done frequently and large numbers of velvets are regularly used.

This unit first presents the necessary details for growing yeast cells. This is followed by a description of replica plating methods for assessing the nutritional requirements and mating types of strains. Yeast genetic experiments often require the construction of strains with specific genotypes, as well as an analysis of the meiotic segregation patterns of newly introduced mutations (see *UNIT 13.8*). These genetic manipulations are carried out using the protocols presented in the final sections of this unit, which describe the construction and selection of diploids, sporulation, and tetrad analysis.

BASIC PROTOCOL

GROWTH IN LIQUID MEDIA

Wild-type *S. cerevisiae* grows well at 30°C with good aeration and with glucose as a carbon source. When using culture tubes, vortex the contents briefly after inoculation to disperse the cells. Erlenmeyer flasks work well for growing larger liquid cultures, and baffled-bottom flasks to increase aeration are especially good. It is important that all glassware be detergent-free.

For good aeration, the medium should constitute no more than one-fifth of the total flask volume, and growth should be carried out in a shaking incubator at 300 rpm. For small-scale preparations of DNA and RNA, yeast can be grown in glass or plastic culture tubes filled one-third full with medium and shaken at 350 rpm in a rack firmly attached to a shaking incubator platform.

BASIC PROTOCOL

GROWTH ON SOLID MEDIA

Yeast cells can be streaked or spread on plates as shown for bacteria in the sketches in *UNIT 1.3*. When a dilute suspension of wild-type haploid yeast cells is spread over the surface of a YPD plate and incubated at 30°C, single colonies may be seen after ~24 hr but require ≥48 hr before they can be picked or replica plated (see below). Growth on dropout media (*UNIT 13.1*) is about 50% slower.

BASIC PROTOCOL

DETERMINATION OF CELL DENSITY

The density of cells in a culture can be determined spectrophotometrically by measuring its optical density (OD) at 600 nm. For reliable measurements, cultures should be diluted such that the OD₆₀₀ is <1. In this range, each 0.1 OD₆₀₀ unit corresponds to ~3 × 10⁶ cells/ml. Thus, an OD₆₀₀ of 1 is equal to ~3 × 10⁷ cells/ml. It is advisable to calibrate the spectrophotometer by graphing the OD₆₀₀ as a function of the cell density that has been determined by some other means, such as direct counting in a hemacytometer chamber (*UNIT 1.2*) or titering for viable colonies (*UNIT 1.3*).

Growth and Manipulation of Yeast

13.2.1

Contributed by Douglas A. Treco and Fred Winston
Current Protocols in Molecular Biology (1992) 13.2.1-13.2.12
Copyright © 2000 by John Wiley & Sons, Inc.

DETERMINATION OF PHENOTYPE BY REPLICA PLATING

Cells from yeast colonies grown on any medium can be tested for their nutritional requirements by replica plating (UNIT 1.3). An inexpensive replica plating block can be constructed by gluing a circular plexiglass disk (8-cm diameter, 1 cm thick) onto the end of a hollow plexiglass tube (8 cm long with an 8-cm outer diameter). Sterile velveteen squares (velvets) are held in place by a large adjustable tube clamp (available in any automotive supply outlet) set to fit snugly around the outside of the tube.

A master plate containing the strain or strains of interest is first printed onto a velvet. A copy of this impression is transferred to plates made with all the relevant selective media, which may include various dropout and drug media, as well as alternative carbon sources (UNIT 13.1). For analysis of temperature-sensitive mutations (UNIT 13.8), a copy of the master plate is made on a plate that will provide all the nutritional requirements of the strain. This plate is then incubated at 37°C.

DETERMINATION OF MATING TYPE

Genetic analysis of yeast often requires a knowledge of the mating type. This protocol is based on the ability of a strain with a single auxotrophic requirement (the tester strain) to complement any and all nutritional requirements of strains of the opposite mating type, *as long as the genetic deficiency in the tester strain is not present in any of the uncharacterized strains*. The genetic deficiency found in the tester strain prevents it from growing on minimal plates. This deficiency is complemented by the wild-type gene in the uncharacterized strains, which themselves usually cannot grow on minimal plates due to one or more auxotrophic mutations (see commentary following this protocol for determining the mating type of strains *without* auxotrophic requirements). When strains of opposite mating type mate, the resulting diploid can grow on minimal plates. The following protocol is useful for determining the mating types of the numerous spore colonies produced during strain construction.

Materials

YPD medium and plates (UNIT 13.1)

S. cerevisiae: *MAT α thr4⁻* (tester), *MAT α thr4⁻* (tester), and uncharacterized strains

Minimal plates (UNIT 13.1)

Replica plating block (UNIT 1.3)

Sterile velvets (UNIT 1.3)

1. Grow 1-ml overnight cultures of each tester strain.
2. Spread 200 μ l of each tester strain on a YPD plate. To get a completely even distribution of cells, replica plate this freshly spread YPD plate onto a sterile velvet, lift and rotate 90°, and print again. Lift, rotate, and print once more. Discard velvet and incubate plate at 30°C.
3. On the next day replica plate the strain to be tested onto two fresh YPD plates. Discard velvet. Onto one of these two YPD plates, replica plate one of the two tester strains prepared in step 2. Using a new velvet, repeat with the other YPD plate and the other tester strain.

If many uncharacterized strains are picked into a grid pattern on a single plate (UNIT 1.3), their mating types can be determined simultaneously.

4. Incubate plates \geq 4 hr at 30°C.
5. Replica plate each plate onto a minimal medium plate. Incubate plates overnight at 30°C.

**BASIC
PROTOCOL**

**BASIC
PROTOCOL**

***Saccharomyces
cerevisiae***

13.2.2

6. Score for mating type. Growth on the minimal plate printed with the *MAT α thr 4^-* tester strain indicates that the uncharacterized strain is *MAT α* . Growth on the minimal plate printed with the *MAT α thr 4^-* tester strain indicates that the uncharacterized strain is *MAT α* .

The thr 4^- mutation makes tester strains auxotrophic for threonine (i.e., they can't grow on threonine dropout medium). However, a mutation in any gene that leads to nutritional auxotrophy can be used for the tester strains, and strains with defects in other genes required for threonine biosynthesis can be tested using MAT α or MAT α thr 4^- testers. One or a few strains can be tested simply by patch-mating the uncharacterized cells to each tester strain (see protocol for diploid construction in this unit) for 4 hr and streaking each pair onto minimal plates.

Strains without auxotrophic requirements cannot be tested using the above basic protocol for mating-type determination. Instead, mating type can be determined by observing zygote formation microscopically. Patch-mate the strains of unknown mating type separately to MAT α and MAT α strains. After 4 hr at 30°C, examine the cells microscopically for zygote formation (zygotes will appear as relatively large 2- and 3-lobed structures in which the lobes are connected by wide, smooth "necks"). If the strain with unknown mating type is haploid, zygote formation will only be obvious in one of the two mixtures. The unknown mating type is opposite that of the strain with which it forms zygotes.

Mating type can be determined in ~24 hr if fresh tester plates are available. A stock of tester plates can be prepared as described in step 1 and stored at 4°C for 1 to 2 months. Allow mating (step 4) to proceed overnight when using tester plates that have been stored for more than a few days.

STRAIN CONSTRUCTION AND TETRAD ANALYSIS

Because meiosis and sporulation are parts of the life cycle of *S. cerevisiae*, it is relatively straightforward to create strains with different genotypes. Genes on different chromosomes assort independently, and linked genes can be separated by recombination. Diploids are constructed from parents that will each contribute some of the markers desired in the haploid products. The protocol for constructing of a new yeast strain is presented in several distinct stages: (1) diploid construction, where two haploid strains are mated; (2) sporulation, where diploid cells are induced to form spores; (3) tetrad preparation, where the ascus wall is removed from the tetrad; and (4) tetrad dissection, where each of the four haploid spores from a single tetrad is specifically positioned on a plate and grown for subsequent studies. Not all of these steps are necessary in an individual experiment. Diploids can be constructed and stored indefinitely. Spores can be stored at 4°C for 1 to 2 weeks without a significant decrease in spore viability. Finally, plates that have been streaked with glucylase-treated spores can be stored at 4°C for several days prior to dissection.

Diploid Construction

Diploids are constructed by mating strains of opposite mating types on the surface of agar plates (patch mating). Mix cells from freshly grown colonies of each haploid parent with a toothpick in a circle ~0.5 cm in diameter on an agar plate (the plate should allow growth of both haploid strains). Allow mating to proceed ≥ 4 hr at 30°C, then streak the mating mixture onto a plate that will select for the diploid genotype.

When there is no selection specific for the diploid genotype (the case when one of the haploid parents has the same nutritional requirements as the diploid), isolate diploids by physically "pulling zygotes" out of the mating mixture using a dissecting microscope. After mating for 4 hr, streak the mix in several parallel lines on an agar plate that will

support the growth of the diploid. Using the dissecting microscope, identify zygotes by their characteristic shape (described above), pick them up with the dissecting needle, move them away from the streak of cells, and set them down. To ensure that the selected cells are actually diploids, patch them onto sporulation plates (UNIT 13.1) and examine microscopically for tetrad formation after appropriate incubation (see following protocols). Alternatively, attempt to mate selected cells with a pair of mating type tester strains and examine microscopically for zygote formation with each tester. Zygotes should not form if a diploid was correctly selected.

Sporulation of Diploid Cells

Starvation of diploid yeast cells for nitrogen and carbon sources induces meiosis and spore formation, during which chromosomes replicate and proceed through two divisions to produce haploid nuclei. These nuclei (along with surrounding cytoplasm) are individually packaged into spores, and the four spore products (tetrad) of a single meiosis are held together in a thick-walled sac (ascus).

The sporulation process can be induced in cells growing either on solid or in liquid medium. Because some strains do not sporulate well on plates, and other strains do not sporulate well in liquid, both methods are presented; one of the two methods should result in reasonably good spore formation for any given diploid.

Materials

Yeast cells

Sporulation plates *or* sporulation medium, with appropriate nutrients (UNIT 13.1)

YPD medium (UNIT 13.1)

Sporulation on plates

1. Patch cells that have been grown on YPD or selective plates onto a sporulation plate.

If no selective conditions are required, grow cells several days on YPD plates prior to transfer to sporulation medium. Allow single colonies to grow for 3 to 4 days on YPD; for patches of cells, allow 2 days growth on YPD. While this pregrowth is not essential, it results in much more efficient sporulation. A small dab of cells should be smeared over a relatively large area (~1 cm²) of the sporulation plate, such that no thick patches of the inoculum are visible.

2. Incubate 4 days at 25°C.

Sporulation is generally less efficient at higher temperatures. While sporulation can occur in the absence of amino acids or other nutrients that are required by the strain for mitotic growth, sporulation is much more complete when those (and only those) nutrients that the particular strain requires are added (see Table 13.1.1).

Several rounds of mitotic growth will cause the cell number to visibly increase over the sporulation incubation period.

3. Visualize tetrads by suspending a small dab of cells in a drop of water on a microscope slide and examining at a magnification of 250× to 400×.

Tetrads will appear as clusters of four small spheres (the spores), all held within a tight-fitting sac. The four spores can be in either a diamond or tetrahedral configuration.

Not all asci will contain four spores. Some cells do not package all four spore products. The proportion of cells that undergo sporulation as well as the fraction of four-spored asci that result varies from strain to strain.

Sporulation in liquid media

For unknown reasons, some strains do not sporulate well on plates. Even for strains that do, the efficiency can often be increased by sporulation in liquid. In fact, sporulation in liquid is so efficient that higher incubation temperatures (30°C) do not appear to inhibit this process. Under these conditions, sporulation is effectively complete within 48 hr. In addition, the yield of asci with four spores is usually higher using this protocol.

1. In any suitable culture vessel, grow the diploid to be sporulated to an OD₆₀₀ of 2.5 to 3.0 ($\sim 8 \times 10^7$ cells/ml) in YPD medium.
2. Transfer 1 ml culture to a sterile, disposable 15-ml polypropylene tube and centrifuge 5 min at $1200 \times g$ (3000 rpm).
3. Pour off supernatant and resuspend cells in 5 ml sterile water. Vortex to resuspend cells and spin as in step 2.
4. Pour off supernatant and resuspend cells in 1 ml liquid sporulation medium supplemented with nutritional requirements of the particular diploid.
5. Shake for 2 to 3 days at ≥ 350 rpm, 30°C, and examine the culture microscopically for spore formation.

If a strain appears refractory to induction of sporulation, try pregrowing cells in YPA medium (UNIT 13.1). The use of acetate as a carbon source requires respiration, which is a requirement for sporulation. This requirement for respiratory competence is reflected by the fact that cells which have lost mitochondrial function (petites; see glossary in introduction to this chapter) are unable to sporulate.

BASIC PROTOCOL

Preparation and Dissection of Tetrads

The analysis of yeast tetrads requires a standard light microscope with a stage that is movable along both the *x* and *y* axes in precisely measurable intervals, but that does not move up and down (focusing is accomplished by moving the objectives). The microscope must be modified with an assembly for mounting an inverted petri dish on its stage and a micromanipulator for holding and moving a fine glass needle (see support protocol that follows for preparation of dissecting needles). Plans for the construction of micromanipulators for tetrad analysis have been published (Sherman, 1973). An assembled instrument is available from Rainin Instruments.

Two methods of preparing yeast tetrads for dissection are shown below. The first employs glusulase (steps 1a-4a) and the second relies on Zymolyase-100T (steps 1b to 4b) to break down the ascus cell wall. Glusulase is relatively inexpensive and easy to work with, although the more expensive Zymolyase-100T generally gives more uniform and reproducible digestion.

Dissection of yeast tetrads requires a steady hand and a great deal of patience. Don't be discouraged. It is not uncommon for a skilled yeast geneticist to dissect up to 60 tetrads per hour. However, dissection of 20 to 30 tetrads per hour is a commendable achievement.

Materials

Glusulase (Du Pont NEN), diluted 1:10 in H₂O *or* Zymolyase-100T solution:
0.5 mg/ml Zymolyase-100T (ICN Immunobiologicals) in 1 M sorbitol

Dissecting microscope (Rainin Instruments)
Dissecting needle (support protocol)

Prepare the tetrads

For glusulase treatment:

- 1a. If spores to be examined are from a plate, wash by suspending a toothpick-full in 300 μ l sterile water in a 1.5-ml microcentrifuge tube. Vortex vigorously. If spores to be examined are from liquid cultures, centrifuge in a tabletop centrifuge 5 min at 1200 \times g (3000 rpm). Aspirate supernatant and resuspend pellet in 5 ml water. Repeat twice and resuspend final pellet in 5 ml water. Dilute 30 μ l into 270 μ l water in a 1.5-ml microcentrifuge tube. Vortex vigorously.

Subsequent glusulase treatment (for ascus removal) will not be effective unless the asci are washed extensively.

- 2a. Examine cells under a light microscope to detect intact asci.
- 3a. Add 3 μ l diluted glusulase to spore preparation. Mix by gently vortexing.

Commercial preparations of glusulase are highly concentrated and should be diluted 10-fold in water for controlled treatment of asci.

- 4a. Incubate 2 min at room temperature, then examine the cells microscopically. Continue incubation at room temperature until glusulase treatment is complete. Proceed to step 5.

The glusulase treatment is complete when approximately half of the tetrads show disruption of their ascus walls. The four spores should remain associated, but more loosely than when initially examined. The wall of the ascus should be expanded and loosely associated with the spores. Two to six minutes incubation is sufficient for most strains.

For Zymolyase-100T treatment:

- 1b. If spores to be examined are from a plate, add a small toothpick-full of tetrads to 50 μ l Zymolyase-100T solution and gently resuspend. If spores to be examined are from liquid cultures, microcentrifuge 1 ml cells for 10 sec. Pour off supernatant and resuspend pellet in 50 μ l Zymolyase-100T solution.

- 2b. Examine cells under a light microscope to detect intact asci.
- 3b. Incubate 10 min at 30°C.
- 4b. Gently add 0.8 ml sterile water by slowly running it down side of tube. Proceed with step 5.

The Zymolyase treatment is complete when approximately half of the tetrads show disruption of their ascus walls. The four spores should remain associated, but more loosely than when initially examined. The wall of the ascus should be expanded and loosely associated with the spores.

5. Set tubes on ice and leave them there. Streak treated spores (using an inoculating loop) in two parallel lines across the surface of a YPD dissection plate (*UNIT 13.1*), as shown in Figure 13.2.1.
6. Examine plate, inverted with the lid off, on the dissecting microscope. Individual tetrads, grouped into tetrahedral- or diamond-shaped clusters of spores, should be visible.

Dissect the tetrads

The following procedure is specifically adapted for the Rainin Instruments dissecting microscope (Fig. 13.2.1).

7. Position the plate so that streaks are parallel to the *x* axis of the stage. Focus on spores and position a good tetrad in center of field.

Saccharomyces cerevisiae

13.2.6

At 100× magnification, attempt to dissect a tetrad that is not near any other tetrads, so that multiple spores from unrelated tetrads are not picked up by the needle.

Be careful not to graze the tip of the dissecting needle with the petri dish as it is being placed on or removed from the stage. Lower the needle and rotate the microscope objectives away from the stage before changing plates.

8. Move dissecting needle upward using coarse adjustment to a position such that its tip is touching the plate surface when the joystick is depressed halfway.

Moving the joystick down results in the needle moving upward. When the joystick is fully up, the needle will be visible as a large black sphere looming below the plate surface. It will appear to be much larger than the tetrad itself.

9. Using a downward motion on joystick, gently touch needle to surface of plate, immediately adjacent to the tetrad.

The tetrad should disappear under the needle. When attempting to pick up individual spores or tetrads, touch the needle to the plate very close to, but not directly on top of, the spore or tetrad.

10. Gently lift joystick straight up. The tetrad should be gone.

At this point the tetrad should be on the tip of the needle. If it is still on the plate, repeat steps 9 and 10 until the spores stick to the needle. Pick up tetrads or individual spores as gently as possible, with the needle barely touching the agar surface. If glucylase or Zymolyase treatment was effective, all four spores should move together. Tetrads will fall apart at this stage if the treatment was excessive.

11. Using stage movement controls, move the plate so that the needle is at rightmost position on the plate (using x-axis control) and at a position (position *a*) ~1 cm away from streak of treated spores (using y-axis control).

*The four spores will be set in a line perpendicular to the original streak at positions *a*, *b*, *c*, and *d*. These positions should be 0.5 cm apart and are easily found by aligning an arbitrary marking on the movable stage with a fixed y-axis scale that is even with the right side of the stage (Fig. 13.2.1).*

12. Use the joystick to gently touch the needle to the agar surface. Then use the joystick to move the needle down and forward, so that the point where the needle touched the agar is visible. The tetrad should be on the agar surface. The following variations are possible:

- a. No spores were deposited. In this case repeat step 12, using more force when touching the needle to the agar.
- b. Only a single spore is on the plate (the best result). Move the y-axis adjustment to position *b* and repeat step 12.
- c. If two or three spores are on the plate, move the y-axis adjustment to position *b* and attempt to set down the remaining one or two spores on the needle. If two spores were on the needle and only one spore is deposited, move the needle to position and set down the remaining spore. If two spores were deposited at position *b*, the needle must be used to break them apart. This is best accomplished by moving the joystick in circular motion, so that the tip of the needle hits the plate and drags over the spores. When the pair of spores is broken apart, pick one up and move it to position *c*. Don't forget that there are multiple spores at position *a* that still need to be separated.
- d. If all four spores are present, use the needle to break them apart. Once this is done, pick up one or more of the spores (leaving one at position *a*) and attempt to lay down a single spore at each of the three remaining positions.

Be careful not to move the x-axis controls during the dissection of an individual tetrad. If the four spores cannot be separated after repeatedly dragging with the needle, the glusulase treatment was probably not sufficient. Let the treated culture (which should have been on ice) set at room temperature for a few minutes and then streak onto a new YPD dissection plate.

13. Use x- and y-axis controls to move the plate back to a position where the needle is directly below the streak of glusulase-treated spores. Repeat steps 7 through 12, with spores of each successive tetrad set down in a line 0.5 cm to right of and parallel to that of the preceding tetrad. Placement of these lines is determined by aligning an arbitrary marking on the movable stage (upper right-hand corner of the plate in Fig. 13.2.1) with a fixed x-axis scale that is even with back edge of stage. Alternatively, a strip of paper marked in 0.5-cm intervals can be taped along front edge of stage. In this case the lower right-hand corner of the plate holder is used as the arbitrary alignment point.
14. Continue steps 7 through 13 until all positions along the x axis are occupied by dissected tetrads. The plate is then rotated 180° for dissection of additional tetrads using the streak on the opposite side of plate as a source. Thirteen tetrads can be dissected on each side of the plate.

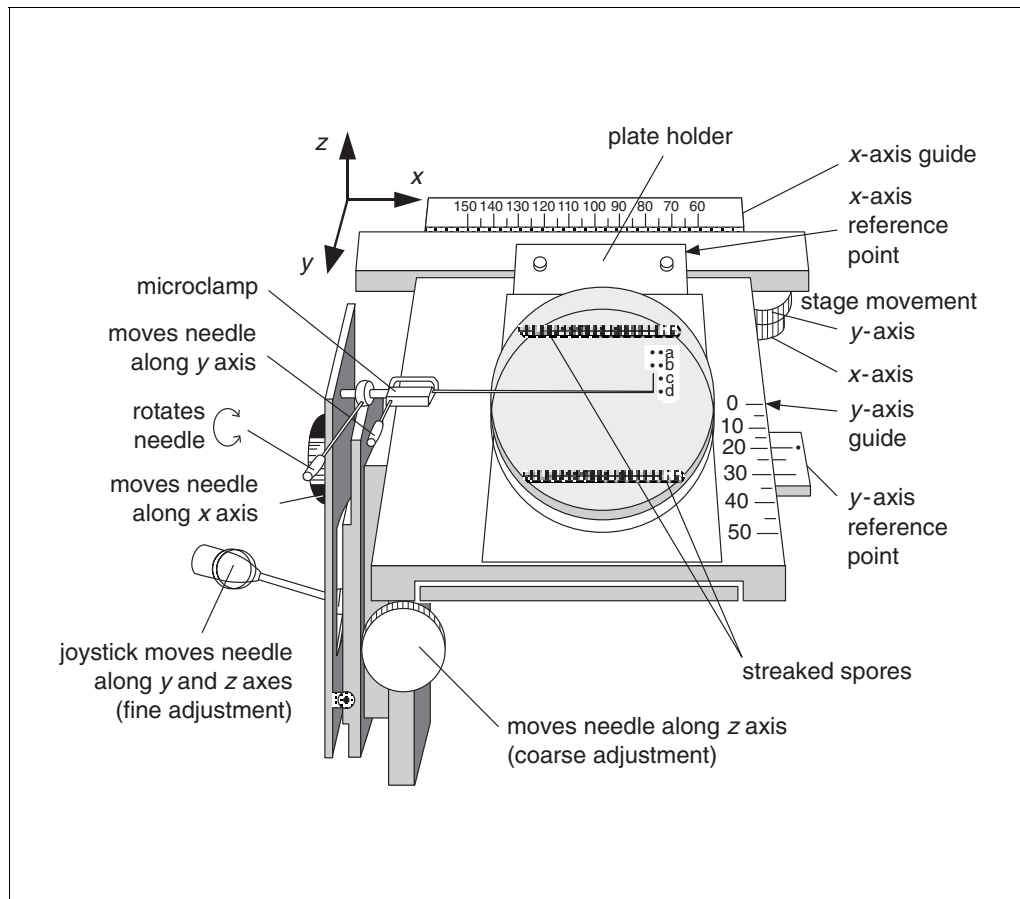


Figure 13.2.1 Essential features of the dissection microscope. Not all parts are shown, including the light source and condenser (below) and the objective and eyepiece (above). The micromanipulator is attached to the base of the microscope. The mark with the large dot on the y-axis guide is aligned with demarcations on the right side of the stage to identify positions a, b, c, and d (corresponding to positions 10, 15, 20, and 25 respectively, in this figure). The upper right-hand corner of the plate holder is aligned with demarcation along the rear of the stage to identify positions 1, 2...13 (corresponding to positions 60, 65...120, respectively in this figure).

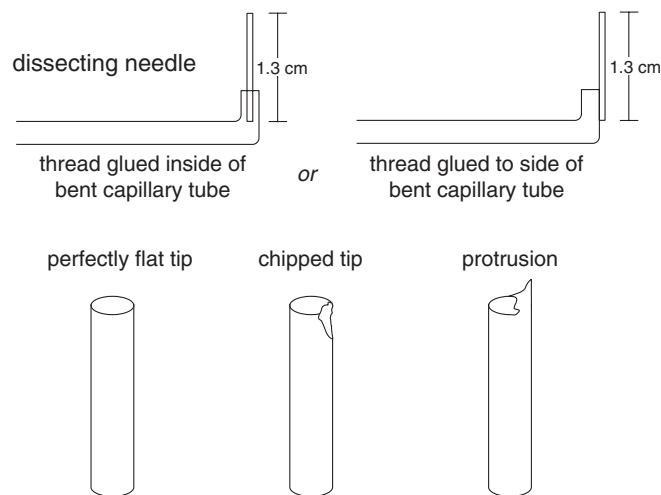
Saccharomyces cerevisiae

13.2.8

Preparation of Dissecting Needles

One of the most difficult steps involved in tetrad analysis is the preparation of the dissecting needle. Like tetrad dissection, making needles requires patience and practice. The general strategy is to first produce a long fine thread of glass, which is broken into multiple short segments. Individual segments are then glued to the ends of capillary tubes. The finished product is L-shaped, with a short arm of ~1.3 cm, culminating in a needle “tip” that is 40 to 150 μm in diameter (see sketch 13.2A).

1. A thin glass thread is created by heating a piece of glass tubing until it is flexible and, in one motion, removing it from the heat source and pulling the ends apart very quickly. The goal here is to draw out the molten glass into a very fine thread before it (nearly instantly) hardens. Timing is everything here so a few attempts may be necessary before a thread that is suitably thin can be generated. Either capillary tubes or pyrex glass pipets can be used as sources of glass tubing. Suitable threads should be between 40 and 150 μm in diameter, depending upon application and personal taste. Human hairs, which range from ~40 to 100 μm in diameter can be used as a reference.
2. Once a thread of appropriate diameter is obtained, it is broken into segments ~1.3 cm long. To do this, simply begin at one end, pulling the segments straight away from the rest of the thread using your fingers. Wear latex gloves for a better grip. Alternatively, 1.3-cm segments can be cut using a fresh razor blade.
3. Examine these segments microscopically. Look for chips and protrusions in the glass as shown in the sketch. The best needles (those that easily pick up and release cells and do not cut the surface of the agar) have an absolutely flat surface that is perpendicular to the shaft (see sketch 13.2A). Save any with an acceptable end, noting which end is which, and discard those with chipped or uneven ends.
4. A dissecting needle is created by attaching a segment with an acceptable tip to the end of a capillary tube. The finished needle is L-shaped with a short arm ~1.3 cm long culminating in the flat working surface. Introduce a right-angle bend in the capillary tube 0.5 to 1.0 cm from the end by briefly heating the tube in a Bunsen burner flame and bending the tip using forceps. Glass threads can be glued in place inside, or against the outside edge, of the short arm of the L-shaped capillary tube, using any fast-acting glue (see sketch 13.2A).



Sketch 13.2A

5. Examine the finished product after mounting the needle on a dissection microscope. Place a plate on the microscope and press the needle firmly against the surface of the agar to make sure the needle leaves an even, circular impression.

Even with all these precautions, needles will perform differently when it is time to dissect, so several candidates should be made at the same time and tested to determine which function best.

Random Spore Analysis

As an alternative to separating spores by tetrad dissection, meiotic products can be released from their asci, dispersed by sonication, and plated directly onto agar plates. The spore colonies can then be screened for the desired genotypes by replica plating.

Materials

Zymolyase-20T solution: 1 mg/ml Zymolyase-20T (ICN Immunobiologicals)
in H₂O and filter sterilized

2-mercaptoethanol (2-ME)

1.5% (v/v) Nonidet P-40 (NP-40)

Ethanol

Sonicator and probe

Additional reagents and equipment for cell counting (*UNIT 1.2*) and
replica plating (*UNIT 1.3*)

1. Prepare cells for tetrad dissection as described earlier in this unit. If sporulation was done on plates, resuspend several toothpicks-full in 5 ml water in a 50-ml flask. If sporulation was done in liquid, resuspend all the cells from a 1-ml sporulation culture in 5 ml water.
2. Add 0.5 ml Zymolyase-20T solution and 10 μ l of 2-ME.
3. Incubate overnight at 30°C with gentle shaking.

Treatment of the sporulated culture with Zymolyase-20T in a hypotonic solution results in lysis of unsporulated diploid cells. The preparation should be examined microscopically after the Zymolyase treatment to evaluate its effectiveness. A higher concentration of the enzyme or the more concentrated preparation of Zymolyase (Zymolyase-100T) can be used.

4. Add 5 ml of 1.5% Nonidet P-40 (NP-40). Transfer the suspension to a 15-ml disposable tube and set 15 min on ice.
5. Hold the tube in one hand and insert the sonicator probe as far into the liquid as possible, but without touching the bottom or the sides of the tube. Before sonicating, clean the sonicator probe with water followed by a wipe-down with ethanol.
6. Sonicate 30 sec at 50% to 75% full power, then set on ice 2 min. Repeat twice.

Sonication produces heat that will warm up the spore suspension significantly. The tube should be cooled between the sonication steps.

7. Centrifuge spores 10 min at 1200 \times g (3000 rpm). Aspirate or pour off supernatant and resuspend in 5 ml of 1.5% NP-40. Vortex vigorously. Repeat twice.
8. Sonicate as in step 6 (with repeats).

The spores should be examined after the last sonication step to ensure that no spores remain stuck together. More sonication steps at higher power settings will release the more tenacious spores. If spores remain stuck to each other, add 2 ml glass beads (Type I, Sigma) and shake 30 min at 300 rpm in an Erlenmeyer flask at 30°C. Let the beads settle and remove the supernatant containing the spores.

**ALTERNATE
PROTOCOL**

***Saccharomyces
cerevisiae***

13.2.10

9. Centrifuge spores 10 min at $1200 \times g$ (3000 rpm). Aspirate or pour off supernatant and resuspend in 5 ml water. Vortex vigorously. Repeat.
10. Count a 10-fold dilution of the treated spores using a hemacytometer.
11. Dilute the spores with water to get 10^3 spores/ml. Plate 100 μ l on several YPD plates and incubate 3 days at 30°C.
12. Screen spore colonies for markers of interest by replica plating (see fourth basic protocol and *UNIT 1.3*).

COMMENTARY

Background Information

Diploids are constructed from parents that each contribute some desired markers in the haploid products. In meiosis, homologous chromosomes assort independently, resulting in haploid cells with new combinations of wild-type and mutant genes. By definition, independent assortment holds for genes located on different chromosomes (different linkage groups). However, the high frequency of crossing-over that occurs during meiosis allows even closely linked genes to be separable. If a large number of spores are scored (generated through either tetrad analysis or a random spore protocol), a strain with almost any combination of markers can be isolated. Large numbers of spores can be analyzed by random spore analysis (alternate protocol).

The frequency with which linked markers are separated by recombination is the basis for genetic mapping in all organisms and is mainly a function of the physical distance separating them. This frequency is equal to twice the genetic distance in centimorgans (cM). Thus, two genes separated by a genetic distance of 5 cM will be recombined in 10% of all tetrads. In yeast, a genetic distance of 1 cM corresponds to a physical distance of ~ 2.5 kb.

Critical Parameters

Diploids are usually selected by streaking mating mixtures on minimal plates that are supplemented (by spreading liquid stocks of the nutrients on the surface of agar plates at the concentrations listed in Table 13.1.1; see also *UNIT 1.3*) with the known nutritional requirements predicted for the diploid. If diploids cannot be obtained by this method, make sure that the haploid parents were of opposite mating type. If they were, and the selective plates are correct, the haploids probably carry at least one additional nutrient requirement that is common to both. In this case, select on CM dropout plates that select for two or more auxotrophic

requirements of the diploid (at least one auxotrophy for each haploid parent). It is important that the haploid parents be from cultures that have been freshly grown, but cells from cultures stored 1 or 2 days at 4°C will also mate well. If cells from plates that have been stored longer are used, allow mating to proceed overnight.

In most laboratory strains of *S. cerevisiae*, many diploid cells in a culture do not undergo meiosis, resulting in a population of spores contaminated to varying degrees with the original diploid. In addition, spores produced by a single meiotic event are often notoriously difficult to separate. The random spore protocol employs an extended incubation with Zymolyase to destroy contaminating diploids, followed by sonication and detergent treatment to disperse spores. The procedure should yield a spore population with <1 diploid cell per 10^4 spores. Spore colonies should be tested for mating type (described earlier in this unit) to determine the frequency of diploid contamination (identified by their nonmating phenotype). Contamination by diploids is usually not a problem when tetrads are dissected, although sometimes three spores and an unsporulated diploid are accidentally dissected as a single tetrad.

While most strains do not sporulate synchronously, pregrowth in YPA medium (*UNIT 13.1*) can result in a degree of synchrony that is useful for monitoring gene expression as meiosis progresses. For more precise synchrony, *S. cerevisiae* strain SKI can be used (Esposito and Klapholz, 1981). This strain undergoes a highly synchronous meiosis which it completes in 8 hr. For this reason, SKI and its derivatives have been extremely useful for studying the molecular and cellular events that occur in meiosis (Wang et al., 1987).

The most critical step in dissecting tetrads is the glucylase or Zymolyase treatment. Too little digestion results in spores that cannot be separated, while too much digestion makes it impossible to pick up intact tetrads with the

needle. The optimum time for glucosylase treatment varies among genetic backgrounds, but once the conditions for proper treatment are established they will give reproducible results for strains from the same genetic background.

When spores are allowed to grow up, it is not uncommon to find tetrads where only 1, 2, or 3 of the spores have grown into colonies. This should be relatively rare (<5% of the tetrads) when closely related haploids are used, but may occur at very high frequencies when the haploid parents are less closely related (as is often the case when strains from different laboratories are mated). If one is simply trying to construct a useful genotype, this is not a major problem. However, it is a problem when viability of spore products is being analyzed to determine if a mutation is lethal. If a marker is needed from a foreign strain where spore viability is poor in crosses with commonly used laboratory strains, spore products carrying that marker must be repeatedly crossed back to a laboratory standard strain—the goal being to introduce only the desired mutation into the genetic background of the laboratory strains. This procedure, called backcrossing, is carried out until spore viability reaches acceptable levels (usually between 4 and 7 backcrosses). At each meiosis, spores are identified that carry the foreign marker in conjunction with several markers found in the laboratory standard strain that is used as the backcross parent. If the foreign marker has been cloned, it is often more efficient to introduce a defined mutation into laboratory strains by transformation, since each backcross requires ~10 days.

Time Considerations

Diploid selection will require ~3 days. Sporulation will take 2 to 4 days. Tetrad preparation will require ~1 hr, and one should

allot several hours to learning how to dissect tetrads. The dissected spores need ~3 days of growth before they form colonies that can be replica plated for genotyping, which requires an additional overnight incubation on the appropriate selective plates. Thus, between 9 and 11 days is required to construct a new haploid strain when starting from haploid parents. This time does not include pregrowing diploids on YPD medium prior to sporulation, which increases sporulation efficiency.

Literature Cited

- Esposito, R.E. and Klapholz, S. 1981. Meiosis and ascospore development. *In* The Molecular Biology of Yeast *Saccharomyces*: Life Cycle and Inheritance (J.N. Strathern, E.W. Jones, and J.R. Broach, eds.) pp. 211-287. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Sherman, F. 1987. Micromanipulators for yeast genetic studies. *Applied Microbiol.* 26:829.
- Wang, H-T., Frackman, S., Kowalisyn, J. Esposito, R.E., and Elder, R. 1987. Developmental regulation of SPO13, a gene required for separation of homologous chromosomes at meiosis I. *Mol. Cell. Biol.* 7:1425-1435.

Key References

- Sherman, F., Fink, G.R., and Lawrence, C.W. 1979. *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Provides a number of detailed procedures for genetic experiments that may be of interest to more advanced students.

Contributed by Douglas A. Treco
Massachusetts General Hospital
Boston, Massachusetts

Fred Winston
Harvard Medical School
Boston, Massachusetts

Genome-Wide Transposon Mutagenesis in Yeast

Transposon-based insertion screens provide a convenient means to identify gene function. Transposon mutagenesis allows one to rapidly generate a large set of mutations that can be screened for mutant phenotypes and specific patterns of gene expression. Moreover, transposon-insertion alleles are much easier to detect than those generated by agents that predominantly produce single-base changes. Insertional mutagenesis is most effectively performed in yeast by means of shuttle mutagenesis. In this procedure, a library of yeast genomic DNA is mutagenized with a bacterial transposon *in vivo* in *E. coli*; mutant alleles are subsequently transferred into yeast for functional analysis. In yeast, shuttle mutagenesis is preferred over transposon mutagenesis *in vivo*, as prokaryotic transposons insert more randomly than do endogenous yeast transposable elements.

This unit provides comprehensive protocols for the use of insertional libraries generated by shuttle mutagenesis. From the Basic Protocol presented here, a small aliquot of insertional library DNA may be used to mutagenize yeast, producing strains containing a single transposon insertion within a transcribed and translated region of the genome. This transposon-mutagenized bank of yeast strains may be screened for any desired mutant phenotype. Alternatively, since the transposon contains a reporter gene lacking its start codon and promoter, transposon-tagged strains may also be screened for specific patterns of gene expression. Strains of interest may be characterized by vectorette PCR (see Support Protocol 1) in order to locate the precise genomic site of transposon insertion within each mutant. A method by which *Cre/lox* recombination may be used to reduce the transposon in yeast to a small insertion element encoding an epitope tag is described below (see Support Protocol 2). This tag serves as a tool by which transposon-mutagenized gene products may be analyzed further (e.g., localized to a discrete subcellular site). For additional considerations regarding these transposon-mutagenized libraries (see Strategic Planning).

NOTE: All solutions and equipment coming into contact with living cells must be sterile, and aseptic technique should be used accordingly.

STRATEGIC PLANNING

At present, the authors have generated three different transposon-mutagenized libraries of yeast genomic DNA. Each library presents its own unique advantage as an insertional mutagen; therefore, care should be taken in choosing an appropriate library when planning a genome-wide screen for gene function.

Specifically, two mutagenized libraries contain transposon insertions bearing the reporter gene, *lacZ*, lacking both its promoter and initiator ATG codon. Our first *lacZ*-insertional library (Burns et al., 1994) was derived by transposon mutagenesis of a genomic library constructed in the Tn3-free vector pHSS6 (Seifert et al., 1986); each library plasmid contains a 3 to 4 kb fragment from a *Sau3A* partial digest of yeast genomic DNA. This library (containing 18 genome equivalents of DNA) was mutagenized *in vivo* in *E. coli* with mTn-*lacZ/LEU2*, a Tn3-derived minitransposon (mTn) carrying the yeast selectable marker *LEU2* and a bacterial marker encoding resistance to ampicillin (Fig. 13.3.1). The resulting insertional library exhibits a random distribution of transposon-tagged genes, but does contain 2 μ m plasmid DNA.

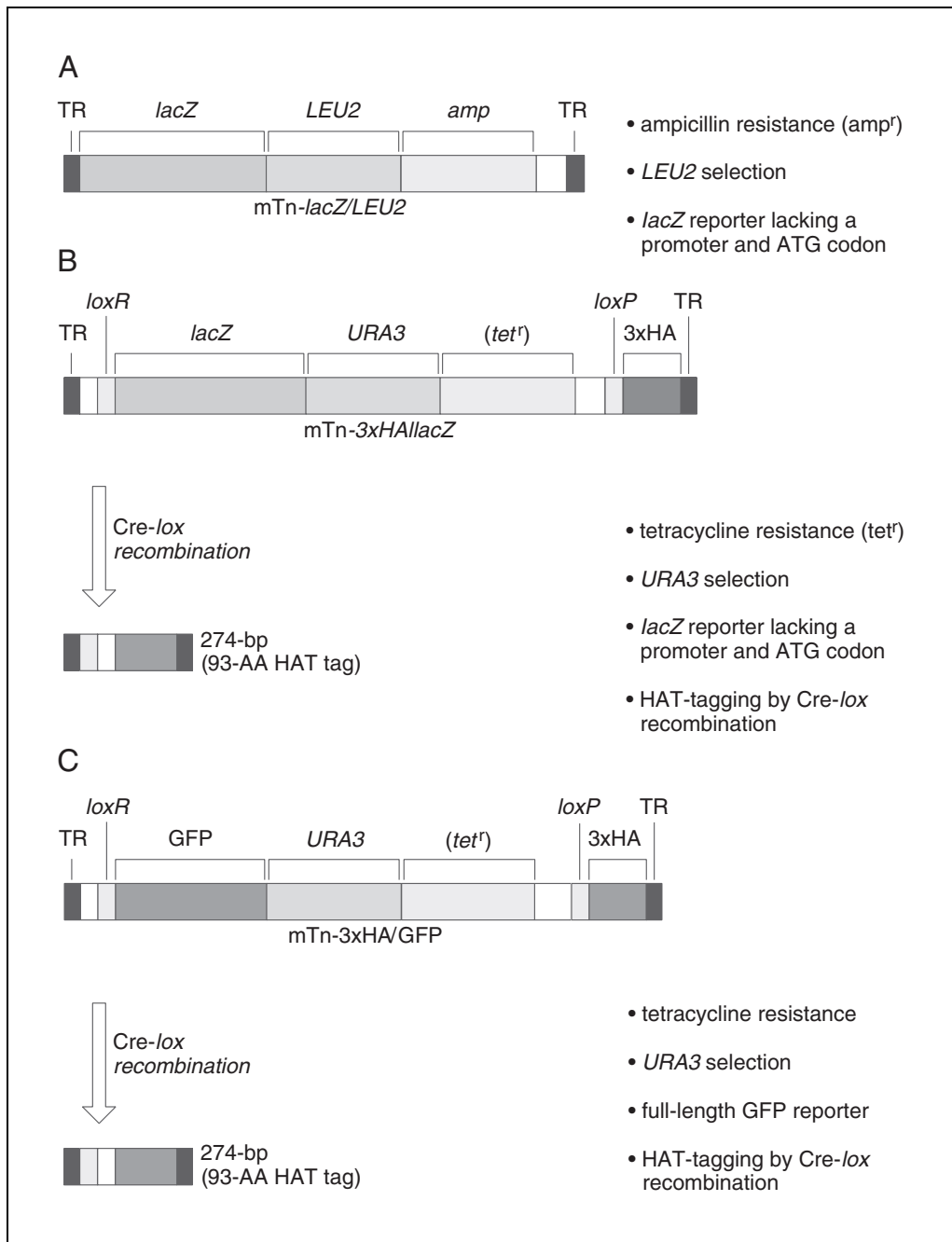


Figure 13.3.1 Transposons for genome-wide random mutagenesis in yeast. Each indicated transposon has been used to mutagenize a plasmid-based library of yeast genomic DNA in *E. coli*. Relevant features of each transposon are indicated below the corresponding schematic. **(A)** mTn-*lacZ/LEU2* Tn3-derived minitransposon capable of generating disruption alleles and reporter gene fusions. Proteins tagged by mTn-*lacZ/LEU2* insertion may be immunolocalized using antibodies directed against β -gal. **(B)** mTn-3xHA/*lacZ*. Cre/*lox* recombination may be used to reduce mTn-3xHA/*lacZ* in yeast to a 274-bp epitope-insertion element. Due to a 5-bp duplication in target-site sequence associated with Tn3 transposition, this insertion element encodes a 93-amino acid tag containing three copies of the HA epitope. mTn-3xHA/*lacZ* may be used to generate *lacZ*-fusions, disruption alleles, conditional alleles, and HA-tagged alleles. **(C)** mTn-3xHA/GFP, a minitransposon carrying full-length GFP reporter gene, in other respects, identical to mTn-3xHA/*lacZ*. Note both mTn-3xHA/GFP and mTn-3xHA/*lacZ* encode identical HA tags upon Cre/*lox* recombination. Corresponding insertion libraries may be requested at the authors' Web site (see Internet Resources).

A second *lacZ*-insertional library was generated by mutagenesis with a *URA3*-marked mTn3 transposon (mTn-3xHA/*lacZ*; Ross-Macdonald et al., 1997) modified to incorporate *lox* elements and sequence encoding three copies of an epitope from the influenza virus hemagglutinin protein (the HA epitope; Fig. 13.3.1). Cre-mediated recombination between *lox* sites (see Support Protocol 2) may be used to reduce the 6 kb mTn-3xHA/*lacZ* insertion to a smaller element encoding an in-frame epitope tag; this allele may be used to determine the subcellular localization of transposon-tagged gene products. This small 93-amino acid epitope tag is also an effective means of generating conditional alleles and hypomorphic mutants for further study. Using mTn-3xHA/*lacZ*, we have mutagenized a plasmid library (containing 50 genome equivalents) generated by *Sau3A* partial digestion of genomic DNA isolated from a *cir^ρrho⁰* yeast strain. The resulting mTn-3xHA/*lacZ* insertion library, therefore, lacks 2μ plasmid sequences and mitochondrial-encoded DNA. This mTn-3xHA/*lacZ* insertional library has been used in the bulk of our genomic studies with excellent results (Ross-Macdonald et al., 1999); however, it does possess a less random distribution of transposon insertions as compared to that found in the mTn-*lacZ*/*LEU2* library.

A third insertion library was generated by mutagenesis with a mini-Tn3 transposon (mTn-3xHA/GFP; Ross-Macdonald et al., 1997) encoding full-length green fluorescent protein (GFP) in place of β-galactosidase (Fig. 13.3.1); in all other respects, this transposon is identical to mTn-3xHA/*lacZ*. GFP is a versatile reporter facilitating a variety of biological studies in living cells (e.g., fluorescence microscopy, fluorescence-flow cytometry, and fluorescent tagging as a means of determining protein localization). We, however, have found *lacZ* to be a more sensitive reporter of gene expression and, therefore, more effective in analyzing genes expressed at a low level.

Diagrams of each transposon, with accompanying summaries, are presented in Figure 13.3.1.

Several considerations should be addressed in choosing a yeast host strain. If using an insertional library derived from mutagenesis with mTn-*lacZ*/*LEU2*, choose a *Leu2⁻* yeast strain. Use of a diploid strain will allow recovery of transposon insertions within essential genes. Also, the use of a *cir^ρ* strain will prevent recovery of insertions within the yeast 2μ plasmid (applicable to the mTn-*lacZ*/*LEU2*- and mTn-3xHA/GFP-mutagenized libraries, which were not generated from a *cir^ρ* strain). For the eventual analysis of HAT-tagged proteins (Support Protocol 2), choose a *Ura3⁻*, *Leu2⁻* strain of yeast; preferably transform this strain with pGAL-cre (pB227; see Support Protocol 2) prior to use in this protocol. Irrespective of the host strain chosen, yeast cultures should be grown to mid-log phase in order to ensure maximal transformation efficiency.

**GENERATING YEAST MUTANTS FROM mTn-MUTAGENIZED
LIBRARY DNA**

This protocol describes methods by which a transposon-mutagenized library of yeast genomic DNA may be used to produce a bank of yeast strains, each strain carrying a transposon insertion within protein-coding sequence. Specifically, the following protocol assumes use of the mTn-3xHA/*lacZ*-mutagenized library (see Strategic Planning), although an identical approach could be applied to the manipulation of any *lacZ*-based insertional library.

The mTn-3xHA/*lacZ*/*URA3*-derived insertional library (as well as any other library described above) may be requested using order forms linked to the authors' Web page (<http://ygac.med.yale.edu>; see Internet Resources). Upon receipt of this library, mutagenized DNA may be excised from pHSS6 for transformation into an appropriate Ura3⁻ yeast strain according to protocols modified from Chen et al. (1992). By homologous recombination, the mTn-mutagenized genomic DNA fragment will replace its chromosomal locus, generating a yeast strain carrying a chromosomal mTn insertion. Transformants carrying an in-frame fusion of mTn-encoded *lacZ* to yeast protein coding sequence may be identified using a color assay for β-galactosidase (β-gal) activity. The resulting bank of yeast strains containing productive *lacZ* fusions may then be screened for any desired mutant phenotype.

Materials

Transposon-mutagenized genomic library plasmid DNA (available upon request; see Internet Resources)

10× TE buffer, pH 8.0 (APPENDIX 2), sterile

E. coli tet^s, kan^s (e.g., DH5α)

14-cm LB plates and medium supplemented with 3 μg/ml tetracycline and 40 μg/ml kanamycin (UNIT 1.1)

LB medium (UNIT 1.1)

Glycerol, sterile

NotI restriction endonuclease and buffer (UNIT 3.1)

Ura3⁻ yeast culture (Chapter 13)

One-step buffer (see recipe)

10 mg/ml denatured salmon sperm DNA (UNIT 14.7)

CM dropout plates and medium without uracil (-Ura; UNIT 13.1)

YPAD plates: YPD plates (UNIT 13.1) supplemented with 80 mg/liter adenine

Chloroform

Xgal plates (UNIT 13.1)

Clinical tabletop centrifuge

45°C water bath or incubator

Sterile toothpicks

3MM filter paper (Whatman)

30°C incubator

9-cm and 15-cm glass petri dishes

Additional reagents and equipment for culturing *E. coli* (UNITS 1.1, 1.2 & 1.3), preparing and transforming competent *E. coli* (UNIT 1.8), isolating plasmid DNA by miniprep (UNIT 1.6) or large-scale preparation (UNIT 1.7), restriction endonuclease digestion (UNIT 3.1), yeast culture (UNIT 13.2), agarose gel electrophoresis (UNIT 2.5), culturing yeast (UNITS 13.1 & 13.2), transforming yeast (UNIT 13.7), and assaying for β-galactosidase activity (UNIT 13.6)

Prepare library DNA

1. Distribute plasmid DNA from individual pools of the transposon-mutagenized genomic library (~1 µg DNA per pool) as dried-down solutions. Briefly centrifuge each dry sample, and resuspend DNA from each pool in an appropriate volume of TE buffer, pH 8.0 (e.g., 10 µl per pool). Each pool of library DNA is mutagenized independently; therefore, it is advisable to process DNA from each pool separately to ensure independent insertion alleles are obtained.
2. Introduce a suitable amount of DNA from each pool into any tetracycline⁻ (tet^s) and kanamycin⁻ sensitive (kan^s) *E. coli* strain by standard transformation procedures (UNIT 1.8).

Standard E. coli strains (e.g., DH5α) are suitable hosts for this transformation. Electroporation (UNIT 9.3) is not required, although resulting transformant yields may be increased 10- to 100-fold. If electroporation is used, decrease DNA quantity appropriately.

3. Select transformants on LB plates (14 cm in diameter) supplemented with 3 µg/ml tetracycline and 40 µg/ml kanamycin.

Approximately 10,000 transformants should be obtained per pool following overnight growth at 37°C.

4. Elute transformant colonies by placing 6 ml of LB medium onto the surface of each plate and scraping cells into a homogenous suspension.

An aliquot of this suspension should be stored at -70°C in 15% glycerol.

5. Dilute a 1-ml aliquot of this eluate into 100 ml LB medium supplemented with 3 µg/ml tetracycline and 40 µg/ml kanamycin to yield a culture of nearly saturated cell density. Incubate at 37°C with aeration for 2 to 3 hr.

The remaining eluate may be stored at 4°C until completion of subsequent steps.

6. Isolate plasmid DNA by standard miniprep (UNIT 1.6) or large-scale preparation (UNIT 1.7).

Transform yeast with mTn-mutagenized library DNA

7. Digest a small aliquot (typically, at least 1 µg) of plasmid DNA from each library pool (obtained in step 6) with *NotI* restriction endonuclease (UNIT 3.1). Subsequently, analyze a portion of the reaction mixture by agarose gel electrophoresis (UNIT 2.5) to ensure release of mTn-mutagenized yeast DNA from the pHSS6 vector. Store the remaining reaction mixture at 4°C for later use (step 10).

Upon electrophoresis, a distinct 2.1-kb band (pHSS6) and broad 8-kb band (mTn-mutagenized yeast genomic DNA) should be visible.

*The broad 8-kb band generated by *NotI* digestion consists of 2- to 4-kb inserts of yeast genomic DNA carrying the 6-kb transposon construct. As this library was constructed from size-fractionated yeast genomic DNA fragments, insert sizes are relatively homogeneous.*

8. Grow a 10-ml culture of any desired Ura3⁻ yeast strain to mid-log phase (a density of 10⁷ cells/ml or OD₆₀₀ of ~1) maintaining appropriate selection, as applicable.

Several considerations should be addressed in choosing a yeast host strain (see Strategic Planning).

9. Centrifuge cells in a clinical tabletop centrifuge for 5 min at 1100 × g at room temperature. Wash pellet once with 5 vol of one-step buffer.

This initial wash in one-step buffer is particularly important if culture volumes are increased.

Saccharomyces cerevisiae

13.3.5

10. Resuspend cells in 1 ml of one-step buffer supplemented with 1 mg of denatured salmon sperm DNA. Add 100- μ l aliquots from this suspension to microcentrifuge tubes containing 0.1 to 1 μ g of *NotI*-digested plasmid DNA from step 7. Vortex tubes to mix contents thoroughly.

To minimize generation of transformants containing more than one insertion, use just enough transforming DNA to yield a reasonable number of transformants (see Anticipated Results). For best results, perform this pilot experiment after determining optimal conditions; scale up as appropriate.

11. Incubate this mixture at 45°C for 30 min.
12. Microcentrifuge cells for 5 sec at maximum speed, room temperature, and subsequently resuspend pellet in 400 μ l of CM (-Ura) dropout medium. Spread 200- μ l aliquots onto CM dropout plates (-Ura). Incubate at 30°C for 3 to 4 days.

*Up to 1×10^3 transformants may be recovered per 1 μ g of transforming DNA. If the *mTn-lacZ/LEU2*-mutagenized library and *leu2* host strain are used, recover transformants on appropriate CM without leucine (-Leu) dropout medium.*

Screen transformants for β -galactosidase activity

13. To maximize detection of *lacZ*-fusions expressed at a low level, transfer transformant colonies onto YPAD plates with sterile toothpicks (*UNIT 13.2*) at a density of up to 100 colonies per plate.

*If an *Ade2*⁻ host strain is used, the addition of adenine to growth medium will prevent any accumulation of red pigment; otherwise, accumulated red pigment may obscure the blue color produced in subsequent assays for β -galactosidase (β -gal) activity.*

14. Place a sterile disc of 3MM filter paper onto a CM dropout plate -Ura; repeat for as many plates as desired. Replicate transformant cells onto filter-covered plates, to allow easy identification of corresponding colonies on YDAP plates (step 13), and incubate overnight at 30°C.

*As some transposon insertions may impair growth even in the heterozygous state, selection for uracil prototrophy should be maintained during this period of vegetative growth. Similarly, selection for leucine prototrophy should be maintained if using a *Leu2*⁻ host. Other growth conditions (e.g., growth on sporulation medium) may be substituted above as desired.*

15. Following overnight growth, lift filters from plates and place in the lid of a 9-cm glass petri dish. Place this lid inside a closed 15-cm glass petri dish containing chloroform. Incubate for 10 to 30 min.

The minimum exposure time required to lyse a particular yeast strain must be determined empirically.

16. Place filters colony-side-up onto Xgal plates. Incubate inverted at 30°C for up to 2 days.

To minimize cost, very thin plates containing Xgal may be used with no loss in signal quality. The signal intensity observed using an Xgal plate is superior to that obtained by soaking the filters in a buffered Xgal solution.

17. Recover transformants producing β -gal (indicated by blue staining on Xgal plates) from the regrown YPAD plates generated in step 13. Maintain appropriate selection during subsequent growth and manipulation of strains.

If desired, strains of interest may be stored long-term in 15% glycerol at -70°C following overnight growth at 30°C in appropriate CM dropout medium.

VECTORETTE POLYMERASE CHAIN REACTION

Banks of transposon-mutagenized yeast strains (see Basic Protocol) represent powerful tools by which genetic screens can be performed rapidly on a genome-wide scale. Large numbers of yeast mutants exhibiting a given disruption phenotype may be easily identified in these genome-wide screens; however, in order for this analysis to be informative, the precise genomic site of mTn insertion must be located within each of these transformants. Accordingly, the following support protocol provides an adaptation of the vectorette polymerase chain reaction (PCR) method of Riley et al. (1990) by which genomic DNA from any site of mTn insertion may be easily recovered for subsequent DNA sequence analysis.

In vectorette PCR, genomic DNA is initially digested with any suitable blunt-end restriction endonuclease possessing a 4- to 6-base-pair recognition sequence. Following digestion, DNA fragments are ligated to a pair of annealed primers containing a non-homologous central region; these primer pairs form “anchor bubbles” flanking each genomic fragment. PCR is then performed using a primer complementary to mTn sequence and a primer identical to sequence within the anchor bubble. During the initial round of amplification, only the mTn primer can bind its template; however, during subsequent cycles, the anchor bubble primer can anneal to the elongated mTn primer, resulting in selective amplification of DNA sequence adjacent to the point of mTn insertion. Vectorette PCR is summarized in Figure 13.3.2.

Materials

Anchor bubble primers 1 and 2 (see Fig. 13.3.2 for sequences; see *UNITS 2.11 & 2.12* for synthesis techniques)

1 M MgCl₂ (*APPENDIX 2*)

Universal vectorette (UV) and mTn primers (see Fig. 13.3.2 for sequences; see *UNITS 2.11 & 2.12* for synthesis techniques)

Transposon-mutagenized yeast strain (see Basic Protocol)

Appropriate restriction endonucleases (e.g., *AluI* or *DraI*) and buffers (*UNIT 3.1*)

10× T4 DNA ligase buffer (*UNIT 3.4*)

5 mM ATP (*UNIT 3.4*)

400 U/μl T4 DNA ligase (measured in cohesive end ligation units)

5 U/μl *Taq* DNA polymerase and 10× buffer (*UNITS 3.5 & 15.1*)

2.5 mM 4dNTP mix (*UNIT 15.1*)

Heat block

Automated thermal cycler

Additional reagents and equipment for preparing yeast DNA (*UNIT 13.11*), restriction endonuclease digestion (*UNIT 3.1*), ligation of DNA fragments (*UNIT 3.16*), PCR amplification (*UNIT 15.1*), purification of DNA from agarose gels (*UNIT 2.6*), and DNA sequence analysis (*UNIT 7.4*)

1. In preparation for vectorette PCR, synthesize anchor bubble primers (suggested DNA sequence indicated in Fig. 13.3.2). Form anchor bubble as follows.
 - a. Prepare an aqueous solution that contains 2 to 4 mM each of anchor bubble primers 1 and 2.
 - b. Denature by incubating 5 min at 95°C in a heat block.
 - c. Add 1 M MgCl₂ to a final concentration of 2 mM.
 - d. Anneal by removing the block from the base unit and placing it on the bench top until it slowly cools to room temperature.

The annealed anchor bubble may be stored at –20°C until step 5.

Saccharomyces cerevisiae

13.3.7

2. Synthesize universal vectorette (UV) primer (suggested DNA sequence indicated in Fig. 13.3.2). Synthesize primer complementary to the mTn used to generate the insertion library (mTn primer).

If using a lacZ-insertional library, the following sequence is recommended for the mTn primer: 5'-CGC CAG GGT TTT CCC AGT CAC GAC-3'.

The UV primer is identical in sequence (not complementary) to a portion of anchor bubble primer 2 (Fig. 13.3.2). The mTn primer indicated above corresponds to sequence from the 5'-end of lacZ on the antisense strand of mTn-3xHA/lacZ. If analyzing a mTn-3xHA/GFP-mutagenized strain, the following GFP primer sequence is recommended: 5'-CAT CAC CTT CAC CCT CTC CAC TGA C-3'.

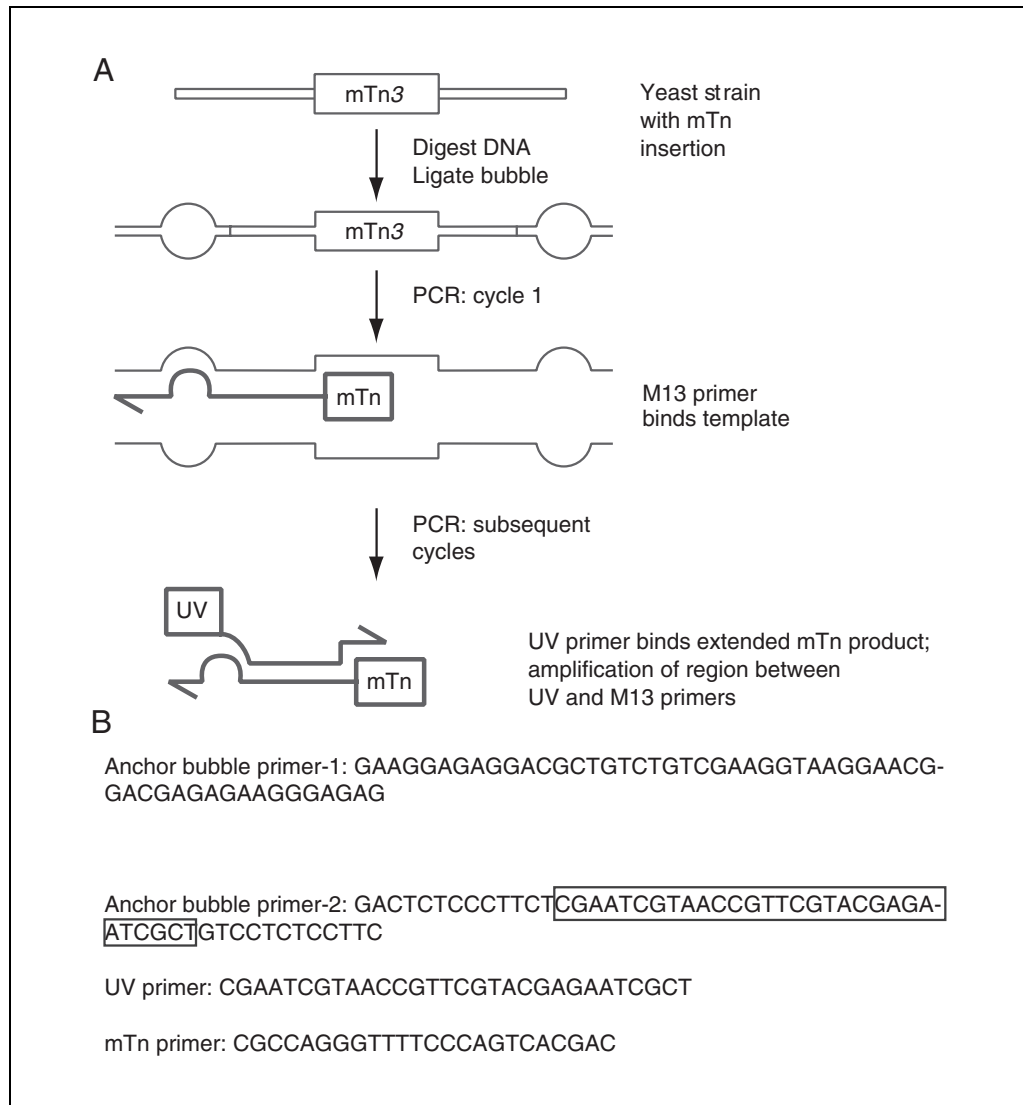


Figure 13.3.2 Vectorette PCR. **(A)** In preparation for vectorette PCR, small, blunt-ended fragments of mTn-mutagenized genomic DNA are generated by restriction enzyme digestion; DNA fragments are subsequently ligated to preannealed anchor bubble primers. Amplification is carried out using mTn and Universal vectorette (UV) primers (sequence indicated). During the first cycle of PCR, only the mTn primer can bind to template, as the UV primer is identical (not complementary) to anchor bubble sequence. During subsequent PCR cycles, the UV primer can bind to elongated product from the mTn primer, resulting in the selective amplification of genomic DNA adjacent to the mTn insertion site. Primers and major PCR products are shown in bold. **(B)** Suggested anchor bubble and PCR primers are listed below: sequence is indicated 5' to 3'. The Universal vectorette (UV) primer is identical to the boxed sequence within Anchor Bubble primer-2.

3. Prepare genomic DNA from a transposon-mutagenized yeast strain of interest (UNIT 13.11).

Care should be taken to obtain high-quality DNA; this is critical for successful PCR amplification.

4. Digest 1 to 3 μg of genomic DNA overnight at 37°C with 10 U of either *AluI* or *DraI* in a total reaction volume of 20 μl (UNIT 3.1). Following digestion, incubate at 65°C for 20 min to inactivate the enzyme.

Ideally, genomic DNA fragments should be small and blunt-ended in preparation for subsequent anchor bubble ligation; accordingly, any restriction enzyme used in step 4 should cut frequently and leave blunt-ends after cleavage.

IMPORTANT NOTE: *The enzyme must not cut between the transposon end and the mTn primer binding site.*

5. To the reaction mix from step 4, add the following:

5 μl 10 \times T4 DNA ligase buffer
22.5 μl sterile water
1 μl annealed anchor bubble (generated in step 1)
0.5 μl 5 mM ATP
1 μl (400 U) T4 DNA ligase.

Incubate 9 to 24 hr at 16°C.

To accommodate blunt-end ligation, incubation times have been extended up to 24 hr as indicated. The addition of polyethylene glycol to a final concentration of 15% may enhance blunt-end ligation.

6. Prepare PCR mix (100 μl total) as follows.

- a. Withdraw 5 μl from the ligation reaction from step 4.

- b. To this 5- μl aliquot, add:

10 μl 10 \times *Taq* PCR buffer
71 μl sterile water
8 μl 2.5 mM dNTP mix
2.5 μl 20 μM mTn primer (step 2)
2.5 μl 20 μM UV primer (step 2)
1 μl (5 U) *Taq* DNA polymerase.

If desired, the mixture above may be modified for “hot start” PCR (UNIT 15.1) using wax beads commercially available from a variety of sources; please follow manufacturer-suggested guidelines.

7. Transfer to an automated thermal cycler and denature at 92°C for 2 min. Carry out PCR (UNIT 15.1) under the following conditions:

35 cycles:	20 sec	92°C	(denaturation)
	30 sec	67°C	(annealing)
	45 sec	72°C	(extension)
1 cycle:	90 sec	72°C	(extension)

In our experience, short cycling times tend to improve product yield; however, optimum conditions for PCR amplification must be determined empirically.

- Analyze 80 μ l of the reaction mix from step 7 by nondenaturing agarose gel electrophoresis (UNIT 2.7); a single band containing 200 to 400 ng of DNA should be visible. Excise the PCR product and recover DNA in TE buffer (~12 μ l) as described in UNIT 2.6.

If multiple bands are present, excise all bands and recover DNA individually from each. DNA may be recovered with Qiaex resin (available from Qiagen) if desired.

- Analyze ~4 to 6 μ l of recovered product by DNA sequencing (UNIT 7.4) to identify the exact site of mTn insertion.

SUPPORT PROTOCOL 2

EPITOPE TAGGING OF mTn-MUTAGENIZED GENE PRODUCTS

Transposon-mutagenized yeast strains may be analyzed further by utilizing the epitope-tagging feature incorporated into mTn-3xHA/*lacZ* and mTn-3xHA/GFP. As shown in Figure 13.3.1, these minitransposons each contain a pair of *lox* elements located internal to the sequence, encoding three copies of the HA epitope. These *lox* sequences are target sites for the Cre recombinase, which catalyzes site-specific recombination between the *lox* sites. Therefore, expression of Cre results in excision of the central body of the transposon, leaving behind a 274-bp sequence containing the HA-epitope coding region. Due to a 5-bp target site duplication associated with Tn3 transposition, this reduced construct corresponds to a small element encoding 93 amino acids (the HA-epitope tag or HAT tag) inserted within the protein. This HAT tag insertion element is an effective means of generating conditional alleles, hypomorphic mutants, and epitope-tagged strains for immunodetection.

The following support protocol describes a method by which yeast strains bearing an in-frame mTn insertion may be used to derive corresponding HAT-tagged strains by Cre/*loxP* recombination in vivo. The phage P1 Cre recombinase is expressed exogenously from plasmid pGAL-*cre* (available from authors' Web site; see Internet Resources). On this plasmid, *cre* is under transcriptional control of the *GAL* promoter; therefore, induction by galactose can be used to drive *cre* expression. Following galactose induction, cells that have undergone Cre-mediated recombination (and loss of the *URA3* marker) are selected on medium containing 5-fluoroorotic acid (5-FOA). In the authors' experience, galactose induction has resulted in Cre-mediated excision of the mTn-encoded *URA3* marker in >90% of cells analyzed.

Materials

- mTn-mutagenized yeast strain (see Basic Protocol)
 - pGAL-*cre* (pB227; available upon request from authors, see Internet Resources)
 - Raff/CM dropout plates and medium –Leu, –Ura with 2% (w/v) raffinose (UNIT 13.1)
 - Gal/CM dropout medium –Leu with 2% (w/v) galactose (UNIT 13.1)
 - Glc/CM dropout medium –Leu with 2% (w/v) glucose (UNIT 13.1)
 - 5-fluoroorotic acid plates (5-FOA; UNIT 13.1)
 - Glycerol, sterile
 - 30°C shaker and incubator
- Additional reagents and equipment for culture of yeast (UNIT 13.2) and transformation of yeast cells (UNIT 13.7)

1. Transform mTn-mutagenized yeast strain with pGAL-cre (pB227) using standard procedure (see Basic Protocol or UNIT 13.7). Select transformants on CM –Leu, –Ura dropout plates.

Alternatively, pGAL-cre may be transformed into a desired yeast strain prior to transformation with mTn-mutagenized genomic DNA (see Basic Protocol, steps 8 to 11).

This plasmid contains the following elements: amp, ori, CEN, and LEU2.

2. Inoculate transformants into 2 ml Raff/CM dropout medium –Leu, –Ura with 2% raffinose as the carbon source. Incubate at 30°C with aeration until the culture has grown to saturation.

Growth in raffinose derepresses the GAL promoter.

3. Dilute cultures 100-fold into 2 ml Gal/CM –Leu dropout medium with 2% galactose as the carbon source. As a control, dilute an aliquot of this same culture 100-fold into 2 ml Glc/CM –Leu, with 2% glucose as the carbon source. Grow cultures for 2 days at 30°C with aeration.

Some strains grow very poorly in galactose; however, galactose induction is sometimes effective even without visible signs of growth.

4. Test strains for loss of *URA3* marker as follows.
 - a. If visible growth is apparent in cultures grown on 2% galactose, dilute cultures 100-fold in sterile water and withdraw a 10- μ l aliquot.
 - b. If no growth is apparent in cultures grown on 2% galactose, withdraw a 10- μ l aliquot from the undiluted culture.
 - c. Dilute cultures grown in 2% glucose 100-fold in sterile water and withdraw a 10- μ l aliquot.

Spot the aliquot onto a 5-FOA plate; isolate single colonies by streaking the droplet. Incubate 5-FOA plates at 30°C until growth is visible on those plates inoculated with strains grown in galactose.

*Loss of the transposon-encoded *URA3* gene by galactose induction of the Cre recombinase should be reflected in colony growth on 5-FOA plates containing strains grown in galactose. 5-FOA plates carrying strains grown in the presence of glucose should display little or no growth, as expression of the Cre recombinase is repressed by glucose.*

Alternatively, plate diluted cultures obtained in this step onto CM medium and replicate on CM dropout medium –Ura. Incubate at 30°C ~2 days. Cultures grown in galactose should yield roughly 100-fold more Ura⁻ cells than identical cultures grown in glucose.

5. Save single colonies from strains that have lost the *URA3* (marker exclusively following galactose induction) as a stock in 15% (w/v) glycerol at –70°C.

If desired, PCR analysis may be used to confirm the position of the HAT tag; the complete 274-bp DNA sequence encoding this tag may be viewed on the authors' Web site (see Internet Resources).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

One-step buffer

0.2 M lithium acetate
40% (w/v) PEG 4000
100 mM 2-mercaptoethanol
Store up to 1 week at 4°C in a dark bottle

COMMENTARY

Background Information

Transposable elements are powerful insertional mutagens capable of efficiently generating large numbers of informative mutant alleles for subsequent functional analysis. In particular, transposon-based mutagenesis systems offer several advantages over conventional mutagenesis techniques discussed in UNIT 13.3 (i.e., treatment with ethylmethane sulfonate and UV irradiation). In contrast to these approaches, transposon mutagenesis may be used to generate a diverse variety of insertion alleles within target DNA; transposons may be modified to incorporate reporter genes, regulatory elements, and epitope-tagging elements generating null alleles, reporter fusions, mis-expression alleles, conditional alleles, and epitope-tagged alleles. Transposon insertions are additionally advantageous in that they tag mutant loci with an easily detectable DNA element, thereby facilitating subsequent identification and analysis of mutated genes.

Several transposon systems are applicable to genome-wide mutagenesis in yeast. Endogenous yeast Ty-based transposable elements may be used for *in vivo* as well as *in vitro* transposon mutagenesis (Garfinkel and Strathern, 1991; Devine and Boeke, 1994); however, the utility of these systems may be limited by strong target bias associated with Ty transposition *in vivo* (Ji et al., 1993) and poor transposition efficiency *in vitro*. Prokaryotic transposons, such as Tn3 and Tn7, generally exhibit less bias in target-site selection and offer high mutation frequencies. Additionally, negative target regulation of bacterial transposition renders DNA molecules already containing a single transposon immune to further transposon insertion. Collectively, these properties of bacterial transposons enable them to generate large numbers of single-hit mutants—a suitable population for further functional analysis.

Using shuttle mutagenesis, yeast DNA may be mutagenized in *E. coli* with bacterial transposons; transposon-mutagenized DNA may then be returned to yeast by transformation and homologous recombination (Hoekstra et al., 1991). This unit describes the application of Tn3-based shuttle mutagenesis as a means of generating a bank of transposon-mutagenized yeast strains for functional analysis in yeast. The Basic Protocol provides methods by which a Tn3-mutagenized genomic DNA library may be used to transform yeast, generating a bank of marked genomic mutations. This mutant bank may be subsequently screened for desired disruption phenotypes and specific patterns of gene expression. Transposon insertion sites may be identified in strains of interest either by vectorette PCR (Support Protocol 1) or by plasmid rescue (Burns et al., 1994). Plasmid rescue protocols may be found at the authors' web site (<http://ygac.med.yale.edu>; see Internet Resources). Cre/lox recombination may be induced within yeast strains mutagenized by mTn-3xHA/lacZ or mTn-3xHA/GFP (Support Protocol 2); resulting epitope-tagged gene products may be immunolocalized with antibodies directed against the transposon-encoded epitope. Epitope-tagging is also an effective means of generating hypomorphic mutants exhibiting partial gene function.

These transposon insertion libraries constitute a valuable laboratory reagent facilitating genetic screening in yeast. Traditionally, genetic studies have followed a paradigm in which labor-intensive procedures such as complementation analysis or genetic mapping have been necessary as a means of identifying the affected gene within a mutant of interest. Transposon-mutagenized strain collections, as described in this unit, provide a means of altering this paradigm. Researchers may now identify affected genes within clones of interest

either by directly screening a bank of defined mutants or by simply identifying the mutated genomic loci; in either case, the mutated gene can be identified rapidly with a minimal investment in time and effort. Such labor-saving methodologies are essential in order to expedite studies of gene function.

Critical Parameters and Troubleshooting

Transposon-insertion libraries may be used as insertional mutagens in yeast through a series of simple and straightforward steps. Insertion library DNA may be amplified in *E. coli* using any standard method of DNA transformation (e.g., “heat-shock” treatment, electroporation). Transformants should be recovered on growth medium supplemented with antibiotics as indicated; note that tetracycline should be used at a final concentration of 3 µg/ml. Mutagenized DNA from the library is excised from its vector and subsequently transformed into an appropriate yeast strain; electroporation is not normally required, although its use can increase transformant yields 10- to 100-fold. Transform yeast with a small quantity of DNA in order to minimize recovery of double integrants. Diploid strains containing multiple transposon insertions may be identified by examining segregation of the transposon-encoded selectable marker (i.e., *URA3* or *LEU2*) upon tetrad dissection. Alternatively, Southern analysis (*UNIT2.9*) may be used to identify strains containing two different transposon insertions. For purposes of this analysis, *lacZ*-bearing transposons may be probed with a 817-bp *BamHI*-*ClaI* fragment from the 5′ end of *lacZ*. Care should be taken to ensure that mutant phenotypes of interest are linked to the transposon insertion, as multiple insertion events or unrelated spontaneous mutations may have occurred during yeast transformation. Within strains of interest, genetic analysis should be used to confirm segregation of the desired phenotype with the transposon-encoded marker.

The vectorette PCR (see Support Protocol 1) described in this unit allows efficient amplification of genomic DNA at the site of mTn insertion. Typical of any PCR reaction, product yield is dependent upon the quality of DNA used as template. High-quality genomic DNA may be obtained by any standard protocol; CsCl purification of prepared DNA is not required. To function as template for vectorette PCR, genomic DNA fragments must be modi-

fied by anchor bubble ligation. To facilitate blunt-ended ligation, polyethylene glycol may be added to the ligation mixture at a final concentration of 15% (w/v). Longer ligation times (in excess of 16 hr) may be beneficial. Insufficient product yield may also be addressed by trial-and-error modification of indicated cycling conditions. Per cycle, primer extension at 72°C may be increased from 45 sec to 3 min, if no PCR product is apparent; however, the authors typically have obtained better results using shorter extension times. Alternatively, annealing conditions (suggested temperature of 67°C per cycle) may be made more or less stringent as needed. If multiple PCR products are present after amplification, each product may be individually analyzed by DNA sequencing.

Anticipated Results

Amplification of library DNA in *E. coli* should generate roughly 10,000 colonies per pool. Subsequent transformation of library DNA into yeast should yield at least 2×10^5 transformants. Roughly 180,000 yeast transformants should be screened for reporter activity in order to ensure 95% coverage of the genome; the authors typically observe β-gal activity in 12% to 16% of transformants. For purposes of insertional mutagenesis without regard to in-frame reporter activity, 30,000 to 50,000 colonies should be screened.

Within yeast transformants of interest, transposon insertion sites may be identified by vectorette PCR (Support Protocol 1). Vectorette PCR can be expected to yield 200 to 400 ng of product. This quantity of DNA should constitute sufficient template for 1 or 2 sequencing reactions.

Transposon-mutagenized gene products may be HAT-tagged through galactose-induced *Cre/lox* recombination. The authors have generally observed poor growth of yeast strains on medium containing galactose as its carbon source; however, after growth for a few generations in galactose-containing medium, greater than 90% of yeast cells exhibit loss of the mTn *URA3* marker. In contrast, less than 1% of cells grown in the presence of glucose undergo recombination between mTn-encoded *lox* sites. From a pilot study of epitope-tagged proteins, the authors estimate that 40% to 75% of HAT-tagged proteins should be fully functional, while ~90% should localize properly within the cell (Ross-Macdonald et al., 1997).

Saccharomyces cerevisiae

13.3.13

Time Considerations

The protocols presented herein are neither time-consuming nor labor-intensive. Large quantities of library DNA may be prepared as described within a period of 2 days. Library DNA may be subsequently digested and transformed into yeast in one day; allow 3 to 4 days for growth of transformants. Yeast mutants may be incubated under desired growth conditions for a length of time appropriate to the given study (e.g., overnight). Subsequently, strains may be assayed for β -gal activity in <1 hr; blue staining may develop over a period of up to 2 days.

Transposon-mutagenized gene products may be epitope-tagged through Cre/lox recombination over a period of ~5 days; this protocol, however, requires very little “hands-on” time. Typically, strains should be grown 1 to 2 days in raffinose, followed by growth for 2 days in galactose or glucose. An additional 1 to 2 days may be necessary to allow growth of “pop-out” strains on 5-FOA plates.

Vectorette PCR and subsequent DNA sequence analysis may be performed within a similar time frame. In preparation for PCR, a total of 2 to 3 days may be required to synthesize appropriate primers and isolate genomic DNA. Anchor bubbles may be ligated to genomic DNA in 1 day. PCR itself requires only a few hours. PCR products may be analyzed and recovered in ~6 hr. DNA sequencing reactions can be completed in 1 day, and results may be viewed the next morning. In total, an initial attempt at vectorette PCR may require ~5 days; however, subsequent vectorette PCR may be accomplished in only 3 days. Potential stopping points for both procedures have been indicated at appropriate steps in the respective protocols.

Literature Cited

Burns, N., Grimwade, B., Ross-Macdonald, P.B., Choi, E.-Y., Finberg, K., Roeder, G.S., and Snyder, M. 1994. Large-scale analysis of gene expression, protein localization and gene disruption in *Saccharomyces cerevisiae*. *Genes Dev.* 8:1087-1105.

Chen, D.C., Yang, B.C., and Kuo, T.T. 1992. One-step transformation of yeast in stationary phase. *Curr. Genet.* 21:83-84.

Devine, S.E. and Boeke, J.D. 1994. Efficient integration of artificial transposons into plasmid targets in vitro: a useful tool for DNA mapping, sequencing and genetic analysis. *Nucleic Acids Res.* 22:3765-3772.

Garfinkel, D.J. and Strathern, J.N. 1991. Ty mutagenesis in *Saccharomyces cerevisiae*. *Methods Enzymol.* 194:342-361.

Hoekstra, M.F., Seifert, H.S., Nickoloff, J., and Heffron, F. 1991. Shuttle mutagenesis: bacterial transposons for genetic manipulation in yeast. *Methods Enzymol.* 194:329-342.

Ji, H., Moore, D.P., Blomberg, M.A., Braiterman, L.T., Voytas, D.F., Natsoulis, G., and Boeke, J.D. 1993. Hotspots for unselected Ty1 transposition events on yeast chromosome III are near tRNA genes and LTR sequences. *Cell* 73:1007-1018.

Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J.C., and Markham, A.F. 1990. A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res.* 18:2887-2890.

Ross-Macdonald, P., Sheehan, A., Roeder, G.S., and Snyder, M. 1997. A multipurpose transposon system for analyzing protein production, localization, and function in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 94:190-195.

Ross-Macdonald, P., Coelho, P., Roemer, T., Agarwal, S., Kumar, A., Cheung, K.-H., Jansen, R., Sheehan, A., Symoniatis, D., Umansky, L., Nelson, K., Iwasaki, H., Hager, K., Gerstein, M., Miller, P., Roeder, G.S., and Snyder, M. 1999. Large-scale analysis of the yeast genome by transposon tagging and gene disruption. *Nature* 402:413-418.

Seifert, H.S., Chen, E.Y., So, M., and Heffron, F. 1986. Shuttle mutagenesis: A method of transposon mutagenesis for *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 83:735-739.

Key References

Kumar, A., Cheung, K.-H., Ross-MacDonald, P., Coelho, P.S.R., Miller, P., and Snyder, M. 2000. TRIPLES: A database of gene function in *S. cerevisiae*. *Nucleic Acids Res.* 28:81-84.

Provides a helpful explanation of resources freely available from the authors' web site (see Internet Resources).

Seifert et al., 1986. See above.

An early application of Tn3-based shuttle mutagenesis to Saccharomyces cerevisiae.

Ross-Macdonald et al., 1997. See above.

Provides an in-depth description of multifunctional Tn3-minitransposons used in this unit.

Ross-Macdonald et al., 1999. See above.

Presents an extensive application of Tn3-mediated shuttle-mutagenesis towards functional genomics in yeast, with protocols for the genome-wide analysis of disruption phenotypes, gene expression, and protein localization.

Internet Resources

<http://ygac.med.yale.edu>

Many strains and reagents used in this protocol (including all transposon-insertion libraries) may be requested from this, the authors' web site.

Contributed by Anuj Kumar and
Michael Snyder
Yale University
New Haven, Connecticut

*Saccharomyces
cerevisiae*

13.3.15

Yeast Cloning Vectors and Genes

This unit describes some of the most commonly used yeast vectors, as well as the cloned yeast genes that form the basis for these plasmids. Yeast vectors can be grouped into five general classes, based on their mode of replication in yeast: YIp, YRp, YCp, YEp, and YLp plasmids. With the exception of the YLp plasmids (yeast linear plasmids), all of these plasmids can be maintained in *E. coli* as well as in *S. cerevisiae* and thus are referred to as shuttle vectors.

Table 13.4.1 summarizes the general features of a number of these vectors, including the phenotypes that allow selection in either *E. coli* or yeast (or both). These plasmids contain two types of selectable genes, both of which can confer a dominant phenotype: plasmid-encoded drug-resistance genes (UNIT 1.5) and cloned yeast genes. The drug resistance marker is dominant because the recipient (bacterial) cell does not encode a gene for drug

resistance. In contrast, the cloned yeast gene present on the plasmid is a copy of a gene that is present in the yeast genome as well. Thus, this gene functions as a dominant selectable marker only when the recipient yeast cell has a recessive mutation in the corresponding chromosomal copy of the cloned gene. Many of these cloned yeast genes encode functions involved in biosynthetic pathways of yeast and are capable of complementing certain mutations in similar biosynthetic pathways of *E. coli* (e.g., the cloned *URA3* gene of yeast can complement *ura3⁻* mutations of yeast as well as mutations in the *pyrF* gene of *E. coli*). The availability of bacterial mutations that can be complemented by these yeast genes can greatly simplify plasmid constructions, allowing genetic screening for the expected recombinant plasmid.

Table 13.4.2 presents characteristics of a number of cloned yeast genes, along with

Table 13.4.1 Yeast Plasmids and their Selectable Markers

Plasmid	Size	<i>E. coli</i> replicon	Yeast replicon	Phenotypes selectable in <i>E. coli</i>	Phenotypes selectable in yeast	Reference
YIp5	5541 bp	pMB1	none	Amp ^r , Tet ^r , PyrF ⁺	Ura ⁺	Struhl et al., 1979
YRp7	5816 bp	pMB1	<i>ARS1</i>	Amp ^r , Tet ^r , TrpC ⁺	Trp ⁺	Struhl et al., 1979
YRp17	7002 bp	pMB1	<i>ARS1</i>	Amp ^r , Tet ^r , PyrF ⁺ , TrpC ⁺	Ura ⁺ , Trp ⁺	Stinchcomb et al., 1982
YEp13	10.7 kb	pMB1	2 μ m	Amp ^r , Tet ^r , LeuB ⁺	Leu ⁺	Broach et al., 1979
YEp24	7769 bp	pMB1	25 μ m	Amp ^r ; Tet ^r in some constructions; PyrF ⁺	Ura ⁺	Botstein et al., 1979
YCp19	10.1 kb	pMB1	<i>ARS1</i>	Amp ^r , PyrF ⁺ , TrpC ⁺	Ura ⁺ , Trp ⁺	Stinchcomb et al., 1982
YCp50	7.95 kb	pMB1	<i>ARS1</i>	Amp ^r , Tet ^r , PyrF ⁺	Ura ⁺	Rose et al., 1987
YLp21	55 kb	none	<i>ARS1</i>	n.a.	Trp ⁺ , His ⁺	Murray and Szostak, 1983
pYAC3	11.4 kb	pMB1	<i>ARS1</i>	Amp ^r , Tet ^r , TrpC ⁺ , PyrF ⁺ , HisB ⁺	Trp ⁺ , Ura ⁺ , His ⁺	Burke et al., 1987
2 μ m	6318 bp	none	2 μ m	n.a.	none	Hartley and Donelson, 1980

Table 13.4.2 Cloned Yeast Genes

Yeast gene	Length of sequenced fragment (bp)	Map position in yeast genome	Selectable phenotype in <i>E. coli</i>	<i>E. coli</i> strain ^a	Selectable phenotype of wild type in yeast	Common nonreverting mutant alleles	Selectable phenotype of mutant alleles in yeast ^b	GenBank file name
<i>ARG4</i>	2296	8R	<i>ArgH</i> ⁺	JA209	Arg ⁺	none	none	YSCARG4
<i>ARS1</i>	1453 ^c	4R	none	n.a.	none	n.a.	none	YSCTRP1
<i>CAN1</i> ^d	n.s.	5L	none	n.a.	Can ^s	<i>can1-100</i> <i>can1-11</i>	canavanine sulfate ^r	YSCCAN1
<i>CEN3</i>	627	3	none	n.a.	none	n.a.	none	YSCCEN3
<i>CEN4</i>	2095	4	none	n.a.	none	n.a.	none	YSCCEN4
<i>CYH2</i> ^e	1393	7L	none	n.a.	viability ^f	n.a.	cycloheximide ^r	YSCRPL29
<i>GAL1-GAL10</i> regulatory region	907	2R	none	n.a.	none	n.a.	n.a.	YSCGAL
<i>HIS3</i>	1822	15R	<i>HisB</i> ⁺	BA1	His ⁺	<i>his3-Δ1</i> <i>his3-200</i>	none	SCHIS3Y
<i>LEU2</i>	2230	3L	<i>LeuB</i> ⁺	JA300	Leu ⁺	<i>leu2-3,112</i>	none	YSCLEU2
<i>LYS2</i>	n.s.	2R	none	n.a.	Lys ⁺	<i>lys2-Δ1</i>	α-aminoadipate ^r	n.a.
<i>TCM1</i> ^g	1529	15R	none	n.a.	viability ^f	n.a.	tricodestermin ^r	YSCR13
<i>TRP1</i>	1453 ^c	4R	<i>TrpC</i> ⁺	JA300	Trp ⁺	<i>trp1-289</i> <i>trp1-Δ901</i>	none	YSCTRP1
<i>URA3</i>	1170	5L	<i>PyrF</i> ⁺	DB6656	Ura ⁺	<i>ura3-52</i> <i>ura3-Δ1</i>	5-fluoroorotic acid ^r	YSCODCD

n.s. = not sequenced; n.a. = not applicable.

^aThe genotypes and references for these strains can be found in the legends to Figures 13.4.1, 13.4.2, and 13.4.6, except for JA209 (*argH1 metE xyl trpA36 recA56 str^r*; Clarke and Carbon, 1978).

^bSelection for α-aminoadipate resistance produces both *lys2*⁻ and *lys5*⁻ mutants. Selection for 5-FOA resistance generates both *ura3*⁻ and *ura5*⁻ mutants.

^cThis 1453-bp fragment contains both the *TRP1* and the *ARS1* genes.

^dThe wild-type *CAN1* gene encodes dominant sensitivity to the arginine analog canavanine sulfate.

^eThe sensitive (wild-type) and resistant alleles of the *CYH2* gene are codominant.

^fThe *CYH2* and *TCM1* genes encode ribosomal proteins, which are required for viability, but which can be mutated to confer resistance to cycloheximide or tricodestermin, respectively.

^gThe resistant allele of *TCM1* is dominant to the wild-type sensitive allele.

mutants of bacteria and yeast in which the wild-type cloned gene can be selected. For two yeast genes, a positive selection for mutant alleles also exists: *ura3*⁻ cells can be selected on plates containing the drug 5-fluoroorotic acid (5-FOA), whereas *lys2*⁻ cells can be selected on α-aminoadipate plates (see UNIT 13.2 for recipes for these two types of plates).

The nomenclature of different classes of yeast vectors, as well as details about their mode of replication in yeast are described below.

PLASMID NOMENCLATURE

YIp plasmids (yeast integrating plasmids) contain selectable yeast genes but lack sequences that allow autonomous replication of the plasmid in yeast. Instead, transformation of yeast occurs by integration of the YIp plasmid

into the yeast genome by recombination between yeast sequences carried on the plasmid and homologous sequences in the yeast genome. This recombination event results in a tandem duplication of the yeast sequences that bracket the rest of the plasmid DNA. If the YIp plasmid contains an incomplete portion of a cloned gene, this technique can be used to create a gene disruption (see UNIT 13.10). The reversal of the integration process occurs at a low frequency (about 0.1% to 1% per generation), with excision of the integrated plasmid occurring by recombination between duplicated yeast sequences. The frequency of transformation of YIp plasmids is only 1 to 10 transformants/μg DNA, but transformation frequency can be increased 10- to 1000-fold by linearizing the plasmid within yeast sequences

Saccharomyces cerevisiae

13.4.2

that are homologous to the intended site of integration on the yeast chromosome. Linearization also directs the integration event to the site of the cleavage, which is useful when several different homologous yeast sequences are present on the plasmid.

Three classes of yeast vectors are circular plasmids capable of extrachromosomal replication in yeast. YRp plasmids (yeast replicating plasmids) contain sequences from the yeast genome which confer the ability to replicate autonomously. These autonomous replication sequences (*ARS*) have been shown to be chromosomal origins of replication. YRp plasmids have high frequencies of transformation (10^3 to 10^4 transformants/ μg DNA), but transformants are very unstable both mitotically and meiotically. Despite the fact that *ARS*-containing plasmids replicate only once during the cell cycle, YRp plasmids can be present in high copy number (up to 100 copies per plasmid-bearing cell, although the average copy number per cell is 1 to 10). During mitosis, both instability and high copy number are due to a strong bias to segregate to the mother cell, to the extent that as few as 5% to 10% of cells grown selectively still retain the plasmid.

Incorporation of DNA segments from yeast centromeres (*CEN* elements) into YRp plasmids, to generate vectors called YCp plasmids (yeast centromeric plasmids), greatly increases plasmid stability during mitosis and meiosis. Such plasmids—present in 1 to 2 copies per cell—have a loss rate of approximately 1% per generation and show virtually no segregation bias. During meiosis, *CEN* plasmids behave like natural chromosomes, generally segregating in a $2^+ : 2^-$ ratio.

The last class of circular replicating plasmids, YE_p vectors (yeast episomal plasmids), contain sequences from a naturally occurring yeast plasmid called the $2\mu\text{m}$ circle. These $2\mu\text{m}$ sequences allow extrachromosomal replication and confer high transformation frequencies ($\sim 10^4$ to 10^5 transformants/ μg DNA). These plasmids are commonly used for high-level gene expression in yeast, due to their ability to be propagated relatively stably through mitosis and meiosis in high copy number. YE_p vectors vary in the portion of $2\mu\text{m}$ DNA that they carry, although most carry only the sequences essential for autonomous replication. If the $2\mu\text{m}$ -encoded *REP1* and *REP2* functions are present (either on the YE_p plasmid or due to the presence of endogenous $2\mu\text{m}$ circles), transformants are relatively stable and present in high

copy number (20 to 50 copies). In the absence of *REP* functions, transformants are much more unstable, with a segregation bias and copy number similar to those observed with YRp plasmids. In the $2\mu\text{m}$ circle, a highly efficient site-specific recombination event occurs between two perfect 599-bp inverted repeats, mediated by the $2\mu\text{m}$ -encoded *FLP* gene. Most YE_p plasmids carry at least one copy of this repeat; thus, FLP-mediated recombination between YE_p vectors and other plasmids carrying one or more of these repeats (either the endogenous $2\mu\text{m}$ plasmid or other shuttle vectors) can result in a variety of recombinant plasmid multimers.

YL_p plasmids (yeast linear plasmids) contain certain G-rich repeated sequences at their termini which function as telomeres and allow the plasmid to replicate as a linear molecule. In yeast, the telomeric sequence consists of tandem repeats of the sequence $5'(\text{dG}_{1-3}\text{dT})_3'$. Very short *CEN*-containing YL_p plasmids (10 to 15 kb) are unstable and present in high copy, due to random segregation during mitosis. Increasing the size to 50 to 100 kb produces YL_p vectors that disjoin from each other in a manner similar to that of natural chromosomes, resulting in a copy number of about one per cell. However, these artificial chromosomes—which are lost at a rate of 10^{-2} to 10^{-3} per cell division—are still ~ 100 -fold less stable than a natural yeast chromosome.

MAPS OF SELECTED PLASMIDS AND GENES

Restriction maps and a brief description of selected plasmid vectors from several of the five general classes are presented in Figures 13.4.1, 13.4.2, 13.4.3, 13.4.4, 13.4.5, 13.4.6, and 13.4.7. These plasmids were chosen because they are used by many investigators and are generally applicable for a wide variety of purposes. However, where different selectable markers or unique restriction sites may be required, the reader is referred to two reviews on vector systems used in yeast (Pouwels et al., 1985; Parent et al., 1985). In addition, a method for constructing new plasmids *in vivo* in yeast has been described and employed to construct an extended series of new YRp, YCp, and YE_p plasmids (Ma et al., 1987). Yeast shuttle vectors have also been constructed that are derived from either pUC18 or the Bluescript plasmids, providing a greater variety of unique cloning sites and allowing both identification of recombinant plasmids by screening for alpha-

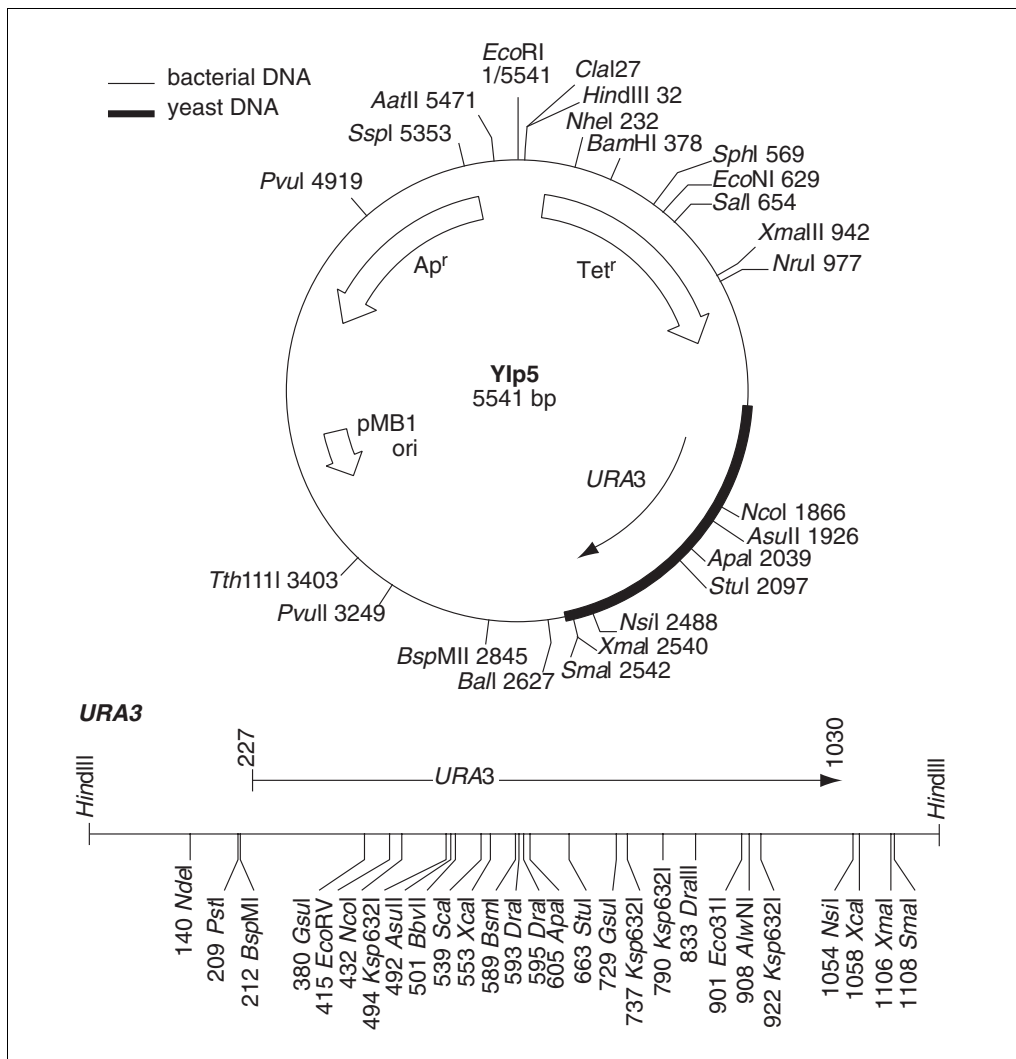


Figure 13.4.1 Yip5. Yip5 contains the 1.1-kb *HindIII* *URA3* gene cloned into the *AvaI* site of pBR322 via the addition of poly(dG-dC) tails (Struhl et al., 1979). Since this plasmid does not contain a yeast origin of replication, transformants occur by integration into the yeast genome at the *URA3* locus; the frequency of transformation can be increased by linearization of the plasmid within the *URA3* insert. The complete nucleotide sequence is available from the Vecbase database (file name: Vecbase.Yip5) and a detailed restriction map can be found in the New England Biolabs catalog.

The *URA3* gene encodes orotidine-5'-phosphate (OMP) decarboxylase, a 267-amino-acid protein required for uracil biosynthesis. The map shown is the gene from strain +D4, which can be expressed in *E. coli* without an external bacterial promoter (Rose et al., 1984). Loss of *URA3*⁺ function can be directly selected using 5-FOA: *ura3*⁻ cells are resistant to 5-FOA, whereas 5-FOA is toxic to cells synthesizing the *URA3* gene product (Boeke et al., 1984). This negative selection has been exploited in a variety of gene replacement schemes, discussed in *UNIT 13.10*. The *URA3* gene can also complement mutations in the *pyrF* gene in *E. coli* using strain DB6656 (*pyrF*::Mu, *lacZ*_{am} *trp*_{am} *hsr*_k⁻ *hsm*_k⁺; Bach et al., 1979).

complementation of the *lacZAM15* mutation of *E. coli* (see *UNIT 1.4*) and the ability to produce single-stranded DNA (Hill et al., 1986; Elledge and Davis, 1988).

Finally, a vector system has been designed using yeast artificial chromosome (pYAC) plasmids, allowing direct cloning into yeast of contiguous stretches of DNA up to 400 kb

(Burke et al., 1987). The circular pYAC plasmids (without inserts) can replicate in *E. coli*. In vitro digestion of the pYAC vector, ligation to exogenous DNA, and direct transformation of the subsequent linear molecules (with telomeric sequences at each termini) into yeast generate a library that can then be screened by standard techniques.

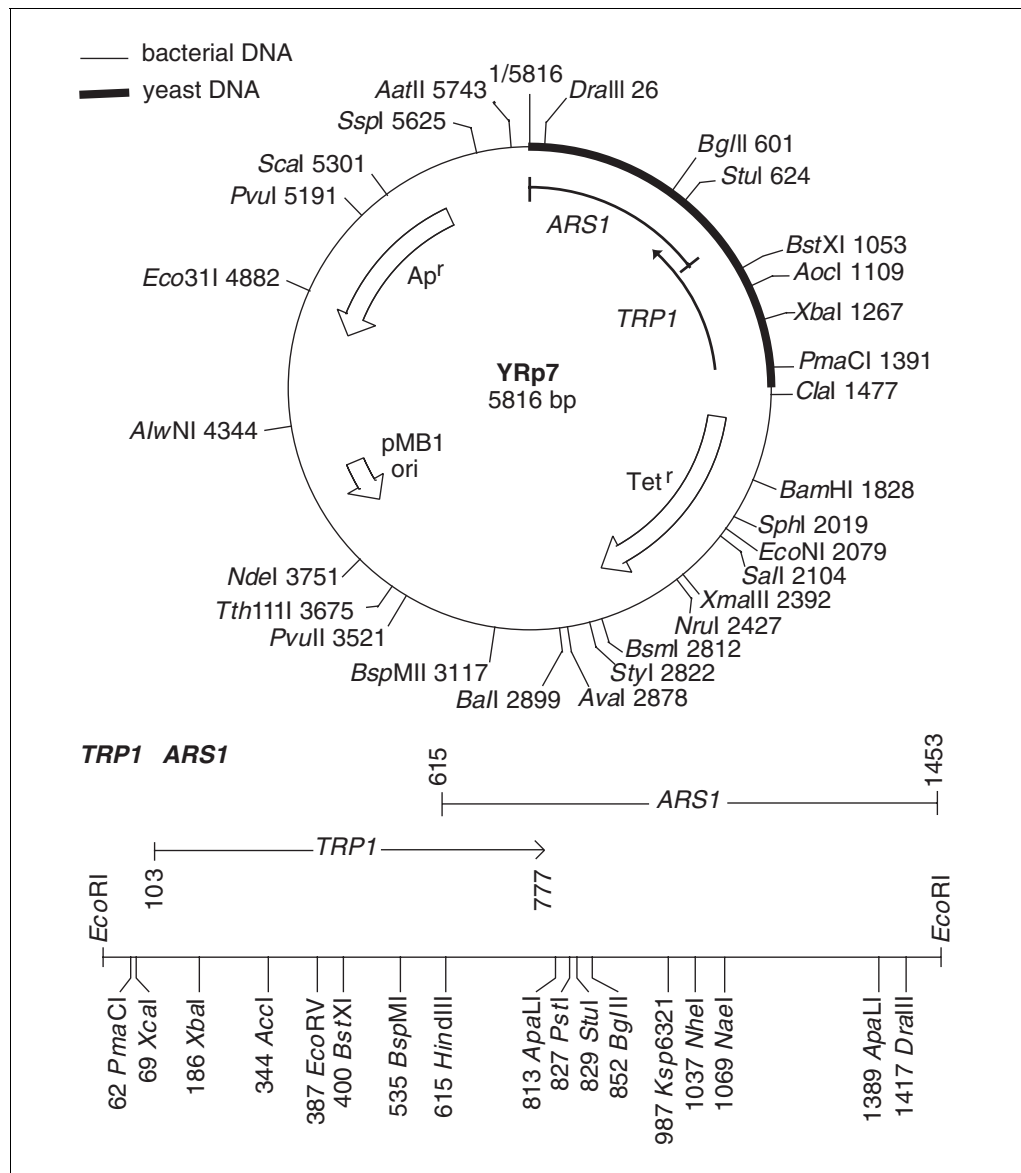


Figure 13.4.2 YRp7. This plasmid contains the 1453-bp *EcoRI* *TRP1* *ARS1* fragment from *S. cerevisiae* inserted into the *EcoRI* site of pBR322 (Struhl et al., 1979).

The *TRP1* RI circle is a derivative of YRp7 containing only the 1453-bp *TRP1* *ARS1* *EcoRI* fragment. This plasmid is mitotically and meiotically unstable, but is present in 100 to 200 copies per plasmid-bearing cell in both *cir*⁺ and *cir*⁰ strains.

A genomic plasmid bank has been constructed by inserting size-selected *Sau3A* partial fragments into the *Bam*HI site of YRp7 (Nasmyth and Reed, 1980).

The *TRP1* *ARS1* 1453-bp *EcoRI* fragment contains both the *TRP1* gene, encoding N-(5'-phosphoribosyl)-anthranilate isomerase, and the autonomous replication sequence *ARS1*. The intact *TRP1* gene can complement mutations in the *trpC* gene of *E. coli*, using *E. coli* JA300 (*thr1 leuB6 thi1 thyA trpC1117 hsr_k⁻ hsm_k⁻ str'; Tschumper and Carbon, 1982). The chromosomal replicator *ARS1* lies between positions 615 and 1453 (on a *Hind*III-*EcoRI* fragment) and is composed of three domains. Domain A contains an 11-bp core sequence (position 857 to 867) consisting of a consensus sequence found in many other *ARS* elements and which is essential for *ARS1* function. Domains B and C flank Domain A and are relatively AT-rich regions that contribute to, but are not essential, for *ARS* function.*

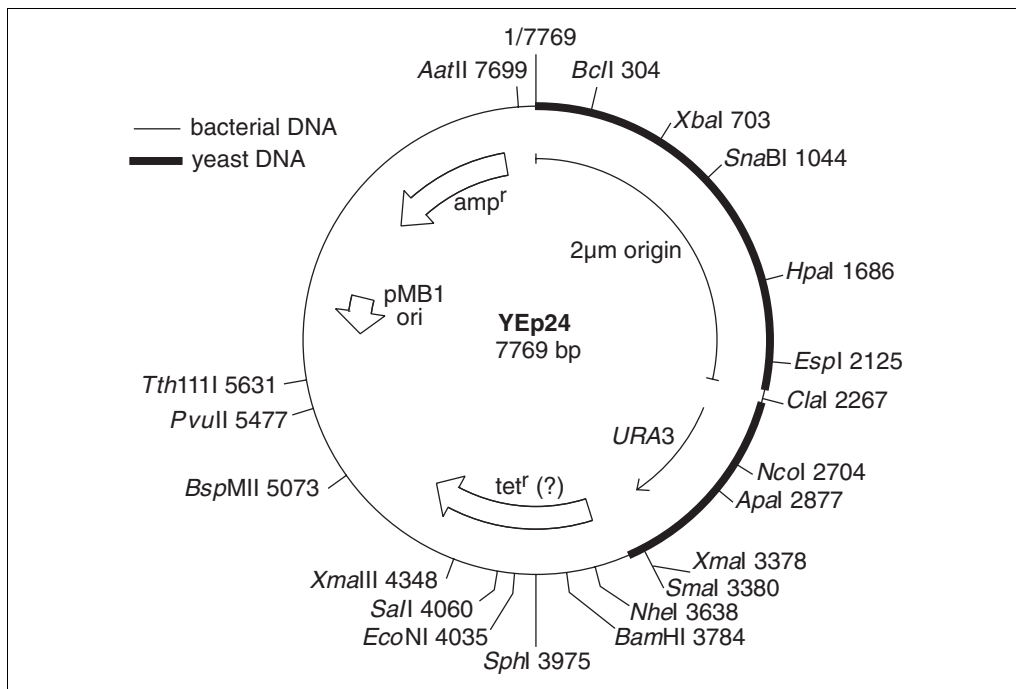


Figure 13.4.3 YEp24. YEp24 has the 2.2-kb *EcoRI* fragment of the B form of the 2 μ m plasmid and the 1.1-kb *HindIII* *URA3* gene inserted into the *EcoRI* and *HindIII* sites, respectively, of pBR322 (Botstein et al., 1979). The expression of the *tet^r* gene is variable among different isolates of this plasmid. YEp24 is mitotically stable in *cir⁺* strains at a copy number of about 20 but is unstable in *cir⁰* strains. The complete sequence of YEp24 is available from the Vecbase database (file name: Vecbase.Yep24) and a detailed restriction map can be found in the New England Biolabs catalog.

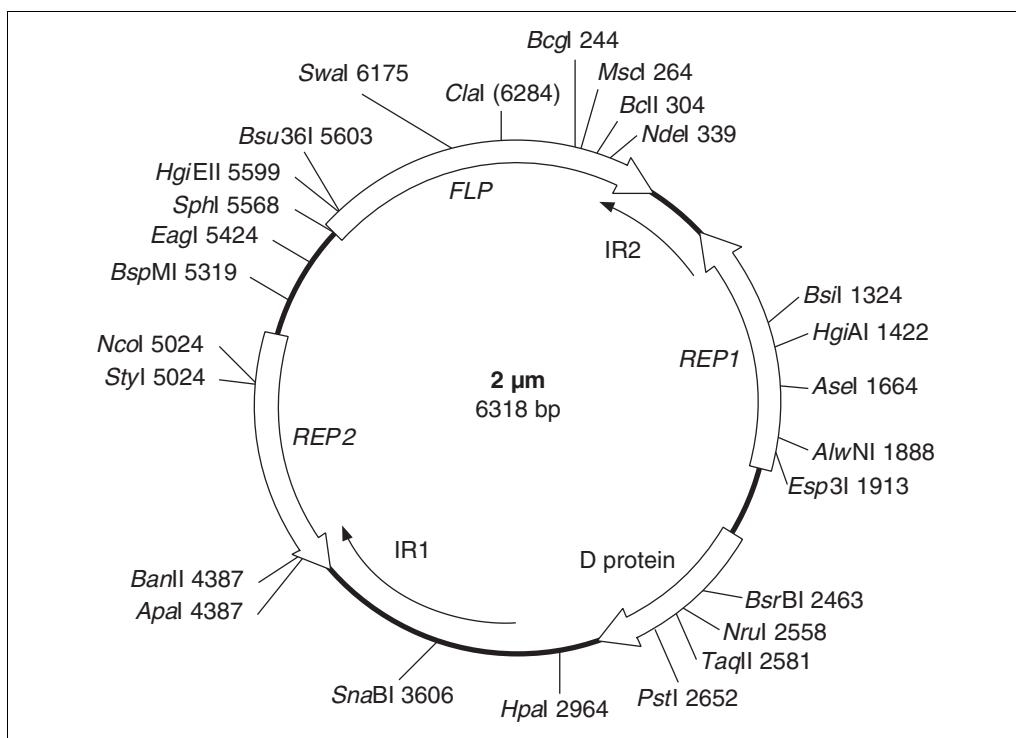


Figure 13.4.4 2 μ m plasmid. The 2 μ m circle is a naturally occurring DNA plasmid found in almost all strains of *S. cerevisiae*, with a copy number of ~20 to 80. The plasmid exists in two different forms, A and B (the former is shown above), due to intra-molecular recombination between two perfect 599-bp inverted repeats. Strains that carry this plasmid are called *cir⁺*; strains missing the plasmid *cir⁰* have been identified or isolated (see UNIT 13.9). It is extremely stable mitotically, with a spontaneous loss rate in haploid cells of 10^{-4} per generation; during meiosis the plasmid is transmitted to all four spore products. The plasmid has been completely sequenced (Hartley and Donelson, 1980) and the sequence is available from GenBank (Plant: yscplasm).

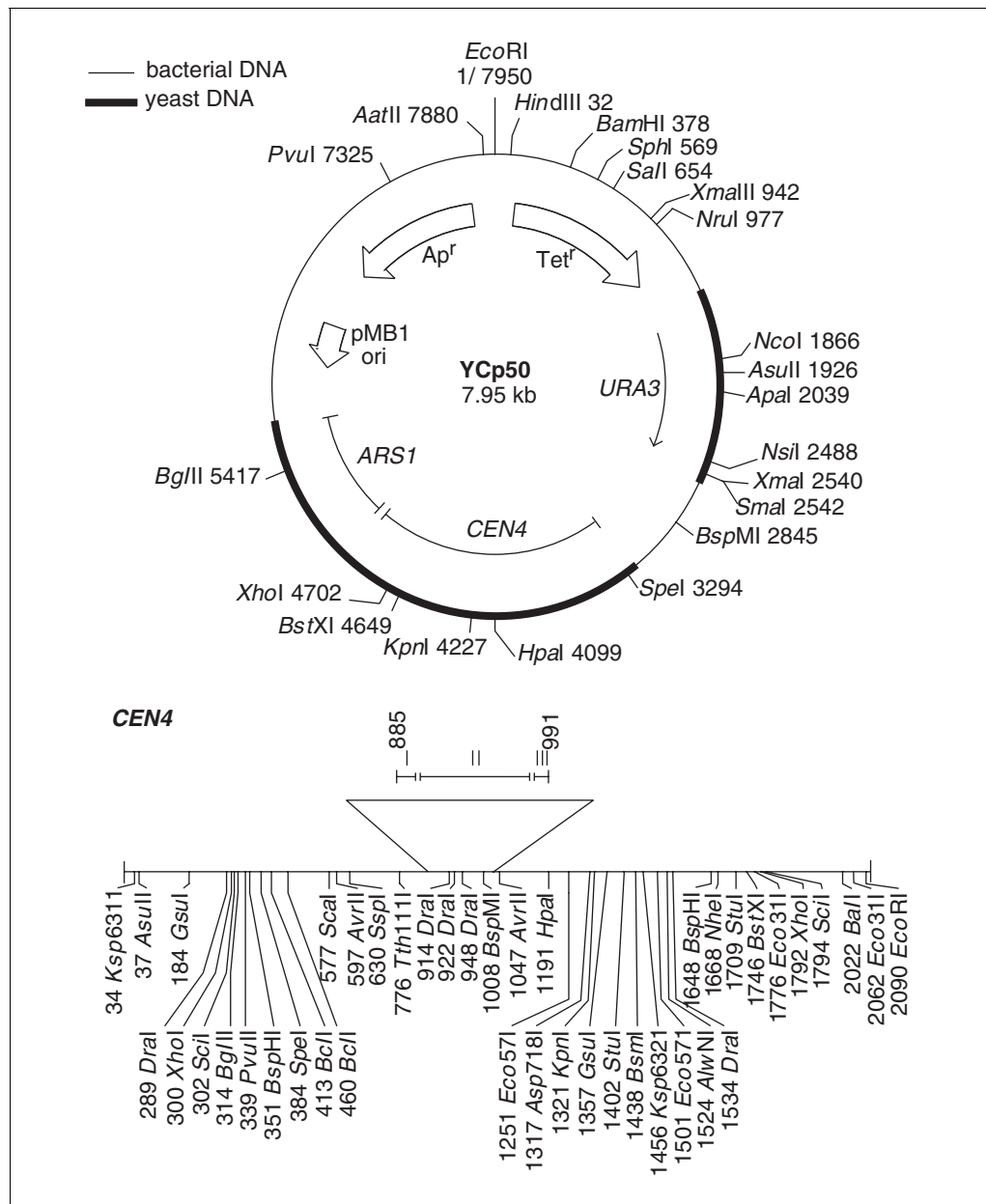


Figure 13.4.5 YCp50. This vector is a derivative of Ylp5 and YCp19. The *EcoRI* site of YCp19 was removed (producing an unsequenced deletion of about 190 bp) and a *PvuII-HindIII* fragment (containing *CEN4* and *ARS1*) from this derivative was cloned into the *PvuII* site of Ylp5, with loss of the *PvuII* site (Rose et al., 1987). Due to the presence of the *CEN* element, this plasmid exists in low copy in yeast (1 to 2 copies/cell) and is mitotically stable (<1% loss per cell per generation). This plasmid has not been completely sequenced; a more complete restriction map is available in Rose et al., 1987.

A set of genomic plasmid banks using YCp50 and size-selected DNA fragments has been constructed (Rose et al., 1987). These plasmid banks provide an alternative to genomic libraries constructed in high-copy-number vectors, useful when isolating genes that would be lethal in yeast when present in high copy.

Both the *CEN3* (Fig. 13.4.7) and *CEN4* (above) sequences were identified based on their ability to confer mitotic stability and proper meiotic segregation to autonomously replicating plasmids (Fitzgerald-Hayes et al., 1982; Mann and Davis, 1986). Nucleotide sequence comparison combined with functional analysis has shown that centromeres contain three conserved structural elements. Elements I and III show the highest degree of sequence conservation between different centromeres, and are separated by an extremely AT-rich region of about 90 bp, designated Element II. Full *CEN4* activity is contained within the 850-bp *PvuII-HpaI* fragment (which contains Elements I, II, and III), although the adjacent 905-bp *HpaI-EcoRI* fragment also confers some mitotic stability to unstable *ARS*-containing plasmids (Mann and Davis, 1986).

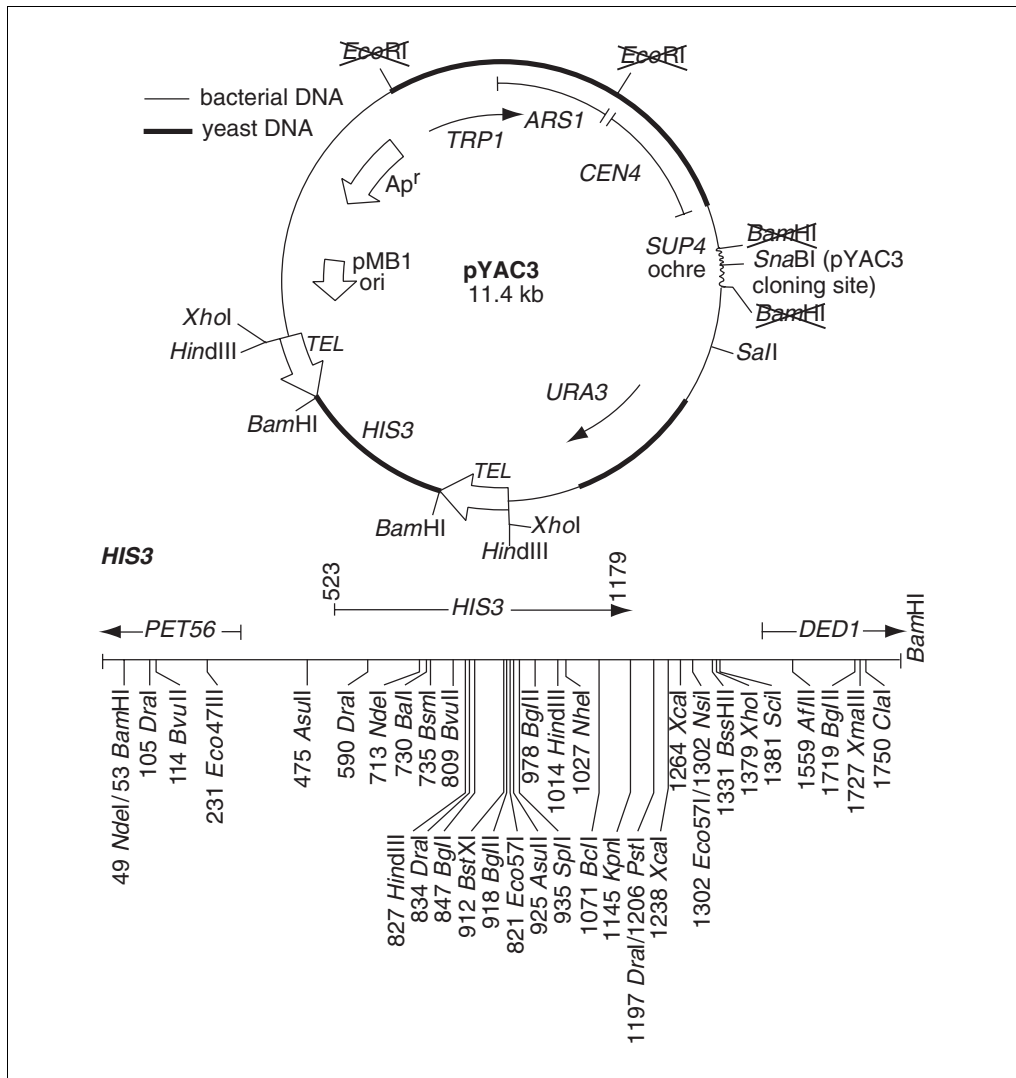


Figure 13.4.6 pYAC3. pYAC vectors are used to clone very large fragments of exogenous DNA onto artificial linear chromosomes, which can be stably maintained in yeast. This vector, which can be propagated as a circular plasmid in *E. coli*, contains a unique cloning site in the *SUP4* gene (an ochre-suppressing allele of a tyrosine tRNA), as well as *ARS1* and *CEN4* elements, required for stable single-copy propagation of the artificial chromosome. The TEL sequences are derived from *Tetrahymena* telomeres and have been shown to function as telomeres in yeast. To clone an insert, pYAC3 is digested with *Bam*HI (which cuts adjacent to the telomere sequences) and *Sna*BI; the resulting vector arms (containing either *TRP1*, *ARS1*, and *CEN4* or *URA3*) are ligated to insert fragments with *Sna*BI-compatible ends. The resulting ligation products are transformed into a *ura3*⁻ *trp1*⁻ *ade2-1* yeast strain, using the spheroplast protocol, selecting for Ura⁺ and subsequently screening for Trp⁺ (to insure that both vector arms are present). Transformants can be further screened for the presence of inserts in the middle of the *SUP4* gene by using a color assay: colonies in which the *ade2-1* ochre mutation is suppressed by *SUP4* are white, whereas inactivation of the suppressor results in red colonies.

The pYAC vector shown above is one of a collection of three plasmids, each with a different cloning site inserted into the *SUP4* gene: pYAC4 and pYAC5 contain *Eco*RI and *Not*I sites, respectively, in place of the *Sna*BI site found in pYAC3. Selected restriction sites (not necessarily unique) are shown for pYAC3, as well as sites that have been destroyed in the process of plasmid construction. The *SUP4* gene is shown as a wavy line. For more detailed discussion of the cloning protocol, as well as details of the construction of this vector, see Burke et al., 1987.

The *HIS3* gene encodes imidazoleglycerolphosphate (IGP) dehydratase, which catalyzes a step in the histidine biosynthetic pathway. This 1822-bp fragment also contains a portion of two other genes: *pet56*, required for mitochondrial function, and *ded1*, required for cell viability (Struhl, 1985). Mutations in the *hisB* gene of *E. coli* can be complemented by the cloned *HIS3* yeast gene, using *E. coli* BA1 (*thr1 leuB6 trpC1117 hisB463 Tn10::near hisB thi1 thyA hsr_k⁻ hsm_k⁻ str^r*; Murray et al., 1986).

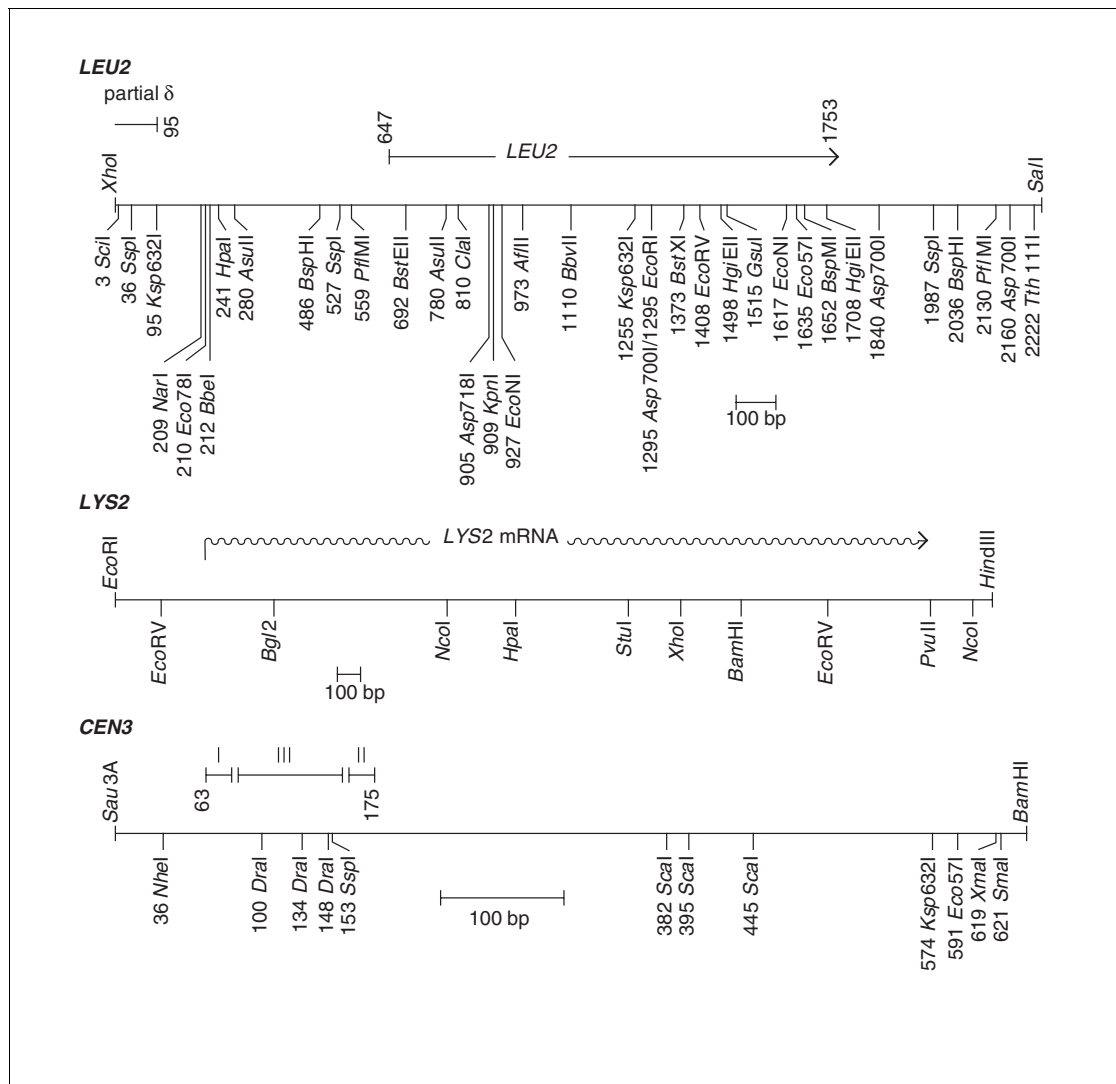


Figure 13.4.7 The *LEU2* gene encodes β -isopropylmalate (β -IPM) dehydrogenase, which catalyzes the third step in leucine biosynthesis (Andreadis et al., 1982). Unlike several other yeast genes involved in amino acid biosynthesis, *LEU2* (and *LEU1*, which is coordinately regulated with *LEU2*) is under specific amino acid control: gene expression is repressed by elevated concentrations of leucine. The *leu2-d* allele is a deletion of the 5'-flanking region of the *LEU2* message which leaves only 29 bp preceding the *LEU2* initiation codon; this derivative of the *LEU2* gene, when present on a YE_p plasmid, requires a very high plasmid copy number to give a Leu⁺ phenotype and has been used to cure cir⁺ strains of the endogenous 2 μ m plasmid (see *UNIT 13.9*). Also contained in this 2230-bp *XhoI-SalI* fragment are 95 bases of the 330-nucleotide δ element. Although this δ element diverges in sequence from other δ elements, when the entire 2230-bp fragment is used as a probe of genomic yeast DNA δ elements present elsewhere in the genome will be detected at a low level. The cloned *LEU2* gene can complement mutations in the *leuB6* gene of *E. coli* using the strain JA300 (*thr1 leuB6 thi1 thyA trpC1117 hsr_k⁻ hsm_k⁻ str^r*; Tschumper and Carbon, 1982).

The *LYS2* gene is the structural gene for α -amino adipate reductase, which catalyzes an essential step in lysine biosynthesis. The gene, which has not yet been sequenced, is present on a 4.6-kb *EcoRI-HindIII* genomic fragment, and gives rise to a 4.2-kb *LYS2* transcript, which is under general amino acid control (Eibel and Philippsen, 1983; Barnes and Thorner, 1986). Much larger genomic fragments (up to 15.7 kb) containing the *LYS2* gene have been isolated, providing a large variety of restriction sites flanking the gene for cloning purposes. As with *URA3*, a positive selection for *lys2⁻* mutants exists: such mutants can be selected on medium containing α -amino adipic acid and lysine, with a spontaneous frequency of 10⁻⁵ to 10⁻⁶. Because the pathways for lysine biosynthesis in bacteria and fungi are not the same, no *E. coli* mutations can be complemented by the cloned *LYS2* gene.

See the legend to Figure 13.4.5 for a discussion of *CEN3*.

LITERATURE CITED

- Andreadis, A., Hsu, Y.-P., Kohlhaw, G.B., and Schimmel, P. 1982. Nucleotide sequence of yeast *LEU2* shows 5' noncoding region has sequences cognate to leucine. *Cell* 31:319-325.
- Bach, M.L., LaCrute, F., and Botstein, D. 1979. Evidence for transcriptional regulation of orotidine-5'-phosphate decarboxylase in yeast by hybridization of mRNA to the yeast structural gene cloned in *E. coli*. *Proc. Natl. Acad. Sci. U.S.A.* 76:386-390.
- Barnes, D.A. and Thorner, J. 1986. Genetic manipulation of *Saccharomyces cerevisiae* by use of the *LYS2* gene. *Molec. Cell. Biol.* 6:2828-2838.
- Boeke, J., LaCrute, F., and Fink, G.R. 1984. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoroorotic acid resistance. *Mol. Gen. Genet.* 197:345-346.
- Botstein, D., Falco, S.C., Stewart, S.E., Brennan, M., Scherer, S., Stinchcomb, D.T., Struhl, K., and Davis, R.W. 1979. Sterile host yeasts (SHY): A eukaryotic system of biological containment for recombinant DNA experiments. *Gene* 8:17-24.
- Burke, D.T., Carle, G.F., and Olson, M.V. 1987. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 236:806-812.
- Clarke, L. and Carbon, J. 1978. Functional expression of cloned yeast DNA in *Escherichia coli*: Specific complementation of argininosuccinate lyase (*argH*) mutations. *J. Mol. Biol.* 120:517-532.
- Eibel, H. and Philippsen, P. 1983. Identification of the cloned *S. cerevisiae* *LYS2* gene by an integrative transformation approach. *Mol. Gen. Genet.* 191:66-73.
- Elledge, S.J. and Davis, R.W. 1988. A family of versatile centromeric vectors designed for use in the sectoring-shuffle mutagenesis assay in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* In press.
- Fitzgerald-Hayes, M., Clarke, L., and Carbon, J. 1982. Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. *Cell* 29:235-244.
- Hartley, J.L. and Donelson, J.E. 1980. Nucleotide sequence of the yeast plasmid. *Nature* 286:860-865.
- Hill, J.E., Myers, A.M., Koerner, T.J., and Tzagoloff, A. 1986. Yeast/*E. coli* shuttle vectors with multiple unique restriction sites. *Yeast* 2:163-167.
- Ma, H., Kunes, S., Schatz, P.J., and Botstein, D. 1987. Plasmid construction by homologous recombination in yeast. *Gene* 58:201-216.
- Mann, C. and Davis, R.W. 1986. Structure and sequence of the centromeric DNA of chromosome 4 in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 6:241-245.
- Murray, A.W., Schultes, N.P., and Szostak, J.W. 1986. Chromosome length controls mitotic chromosome segregation in yeast. *Cell* 45:529-536.
- Nasmyth, K.A. and Reed, S.I. 1980. Isolation of genes by complementation in yeast: Molecular cloning of a cell cycle gene. *Proc. Natl. Acad. Sci. U.S.A.* 77:2119-2123.
- Parent, S.A., Fenimore, C.M., and Bostian, K.A. 1985. Vector systems for the expression, analysis and cloning of DNA sequences in *S. cerevisiae*. *Yeast* 1:83-138.
- Pouwels, P.H., Enger-Valk, B.E., and Brammar, W.J. 1985. Cloning Vectors: A Laboratory Manual. Elsevier Science Publishing, Amsterdam.
- Rose, M., Grisaffi, P., and Botstein, D. 1984. Structure and function of the yeast *URA3* gene: expression in *Escherichia coli*. *Gene* 29:133-124.
- Rose, M.D., Novick, P., Thomas, J.H., Botstein, D., and Fink, G.R. 1987. A *Saccharomyces cerevisiae* genomic plasmid bank based on a centromere-containing shuttle vector. *Gene* 60:237-243.
- Struhl, K. 1985. Nucleotide sequence and transcriptional mapping of the yeast *pet56-his3-ded1* gene region. *Nucl. Acids Res.* 13:8587-8601.
- Struhl, K., Stinchcomb, D.T., Scherer, S., and Davis, R.W. 1979. High-frequency transformation of yeast: Autonomous replication of hybrid DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.* 76:1035-1039.
- Tschumper, G. and Carbon, J. 1980. Sequence of a yeast DNA fragment containing a chromosomal replicator and the *TRP1* gene. *Gene* 10:157-166.

Contributed by Victoria Lundblad
University of California, Berkeley
Berkeley, California

Yeast Vectors and Assays for Expression of Cloned Genes

UNIT 13.6

CONSTRUCTION OF *LACZ* FUSION VECTORS FOR STUDYING YEAST GENE REGULATION

BASIC
PROTOCOL 1

The *lacZ* gene of *Escherichia coli* encodes the enzyme β -galactosidase, which hydrolyzes a variety of β -D-galactosides including chromogenic substrates (colorless compounds that yield a colored product when hydrolyzed). Because of the ease and sensitivity of the assays (see Basic Protocol 2 and see Alternate Protocol), yeast genes are often “tagged” with a functional portion of the *lacZ* gene in order to monitor the regulation of expression of the yeast gene in question. These fusions are constructed such that the promoter region of the yeast gene—plus several amino acids from the N terminus of the protein encoded by this gene—is fused to the carboxy-terminal region of the *lacZ* gene, which encodes a protein fragment that still retains β -galactosidase activity. The liquid β -galactosidase assay (see Basic Protocol 2) is sensitive and quantitative; it is used to accurately monitor gene expression. The filter assay (see Alternate Protocol) is less sensitive and provides only a qualitative assessment of β -galactosidase activity. However, the filter assay is particularly useful for simultaneously and rapidly analyzing a large number of yeast colonies.

When constructing *lacZ* fusions, it is crucial that the translational reading frame across the fusion junction is maintained. If the sequence of the yeast gene is known, an in-frame fusion can be engineered by choosing an appropriate fragment containing the promoter region of the yeast gene as well as the N terminus of the encoded protein. Although the same unique site preceding the *lacZ* fragment may not be present in the N terminus of the yeast gene (and if present, may not generate an in-frame fusion), ligation of flush ends (using another N-terminal restriction site) or the use of oligonucleotide linkers should result in an in-frame fusion. If the sequence of the yeast gene is not known, in-frame fusions can be identified empirically by assaying potential fusions for β -galactosidase activity in either *E. coli* or yeast. A more detailed discussion of construction of in-frame fusions can be found in Guarente (1983).

Figure 13.6.1 shows a vector, pLG670-Z, that can be used for constructing *lacZ* fusions (Guarente, 1983). This plasmid contains the 2 μ m origin of replication and *URA3* as a selectable marker in yeast, with a unique *Bam*HI site at the 5' end of the *lacZ* fragment (this fragment is actually itself a translational fusion of the *lacI* and *lacZ* genes, but for simplicity will be referred to as the *lacZ* fragment). Preceding the *Bam*HI site are unique *Xho*I, *Sa*II, and *Sma*I sites, which can aid in the insertion of various yeast fragments. The *Xho*I and *Sa*II sites are in a region of yeast DNA derived from the promoter region of the *CYC1* gene; however, a deletion of a portion of this promoter has inactivated it. The translational reading frame of the *lacZ* gene immediately following the *Bam*HI site is also shown in Figure 13.6.1.

Two other plasmids (pLG200 and pLG400)—containing different translational reading frames and/or different unique restriction sites at the 5' end of the *lacZ* fragment—have also been constructed (Guarente et al., 1980). These two plasmids do not contain yeast selectable genes or yeast replication origins but can be used to first construct an in-frame fusion, which is then transferred onto a yeast shuttle vector (UNIT 13.4).

Saccharomyces cerevisiae

13.6.1

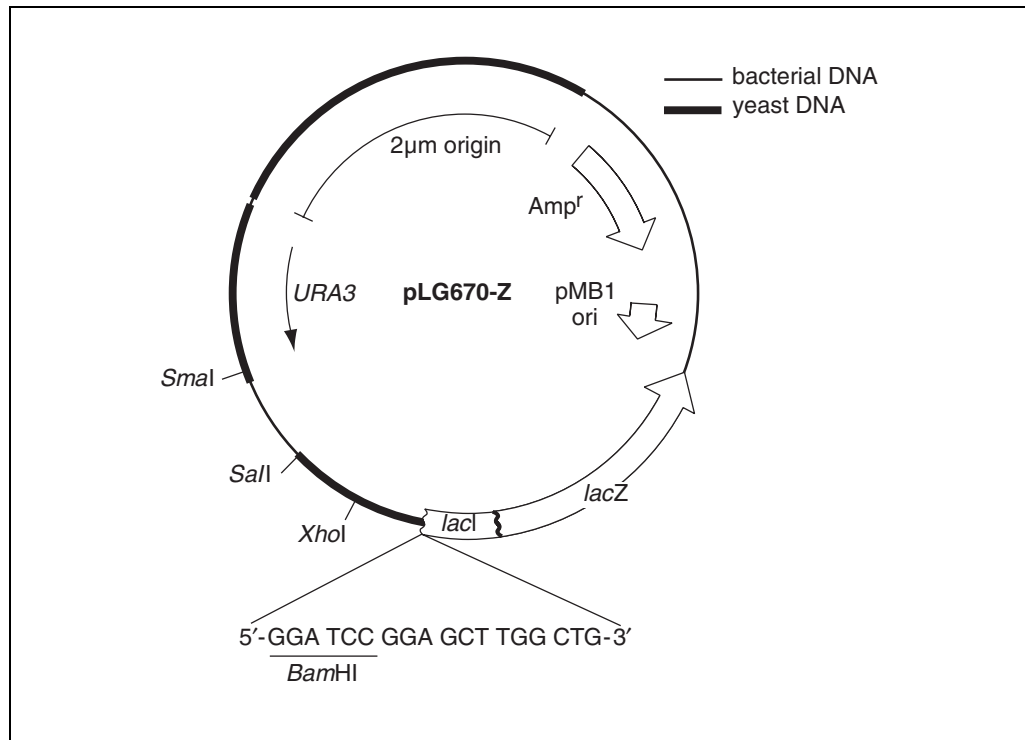


Figure 13.6.1 pLG670-Z (derived from YEp24) contains the 2µm origin of replication, *URA3* as a selectable yeast gene, and a 3' fragment of the *lacZ* gene with a unique *Bam*HI site at the 5' end of the *lacZ* fragment. The sequence just after the *Bam*HI, as well as the translational reading frame, is shown.

**BASIC
PROTOCOL 2**

ASSAY FOR β-GALACTOSIDASE IN LIQUID CULTURES

This protocol describes a rapid, quantitative assay of β-galactosidase activity in liquid cultures of yeast. Yeast cells are either permeabilized or broken open, and the chromogenic substrate *o*-nitrophenyl-β-D-galactoside (ONPG) is added in excess. After incubation at 30°C, the reaction is stopped by raising the pH to 11, inactivating β-galactosidase. Product formation is determined spectrophotometrically.

Materials

- YPD or other appropriate medium (*UNIT 13.1*)
- Yeast strain containing *lacZ* fusion gene (see Basic Protocol 1)
- Appropriate inducing agent (optional)
- Z buffer (see recipe)
- 0.1% sodium dodecyl sulfate (SDS)
- Chloroform
- 4 mg/ml ONPG (Table 1.4.2) in 0.1 M potassium phosphate, pH 7.0 (*APPENDIX 2*;
filter sterilized and stored frozen)
- 1 M Na₂CO₃
- 30°C water bath

Grow and prepare the cells

1. Inoculate 5 ml YPD (or appropriate) medium with a single yeast colony. Grow two to three independent single-colony cultures of a yeast strain containing a *lacZ* fusion gene overnight at 30°C.

Assaying in duplicate or triplicate will help increase the accuracy. If the fusion is present on a plasmid, cells should be grown in medium that selects for the plasmid.

- Inoculate 5 ml YPD medium (or appropriate selective medium with or without inducing agent) with 20 to 50 μl of each overnight culture. Grow to mid- or late-log phase: $0.5\text{--}1 \times 10^8$ cells/ml ($\text{OD}_{600} = 2.0$) for rich medium or $2\text{--}5 \times 10^7$ cells/ml ($\text{OD}_{600} = 0.5$ to 1.0) for minimal medium.

If the goal is to investigate the activity of the yeast promoter under conditions that induce its expression, cells should be grown in parallel under inducing and noninducing conditions.

- Centrifuge cells 5 min at $1100 \times g$ (2500 rpm in a tabletop centrifuge). Resuspend in an equal volume of Z buffer and place on ice.

If the anticipated level of β -galactosidase activity is low, it may be necessary to concentrate the cells. For cells demonstrating 100 to 1000 U of activity within 30 min to 4 hr, concentration is not necessary.

- Determine OD_{600} for each sample.

If cells are in mid-log phase, no dilution is necessary to obtain an accurate OD reading; however, readings above 0.7 are inaccurate.

- Set up the following two reaction tubes for each sample (1 ml each), with mixing:
(a) 100 μl cells with 900 μl Z buffer and (b) 50 μl cells with 950 μl Z buffer.

- Add 1 drop of 0.1% SDS and 2 drops chloroform to each sample using a Pasteur pipet. Vortex 10 to 15 sec and equilibrate 15 min in a 30°C water bath.

The addition of SDS and chloroform permeabilizes the cells. Alternatively, break cells open using 0.5 g of acid-washed glass beads per sample (UNIT 13.12). Pellet the cell debris and measure β -galactosidase activity and total protein concentration of the lysate. Enzyme activity can then be normalized to protein concentration for calculating specific activity.

Assay for β -galactosidase

- Add 0.2 ml of 4 mg/ml ONPG and vortex 5 sec. Place in a 30°C water bath and begin timing.

- When a medium-yellow color has developed, stop the reaction by adding 0.5 ml of 1 M Na_2CO_3 and note the time.

For accuracy, the OD_{420} should be 0.3 to 0.7. With practice, the amount of color correlating with this OD reading can be recognized visually.

- Centrifuge cells 5 min at $1100 \times g$. Determine OD_{420} and OD_{550} of the supernatant.

If the cell debris has been well-pelleted, the OD_{550} —which measures light scattering by cell debris—is usually zero and therefore is not necessary to read.

- Calculate units with the following equation:

$$U = \frac{1000 \times [(\text{OD}_{420}) - (1.75 \times \text{OD}_{550})]}{(t) \times (v) \times (\text{OD}_{600})}$$

where t = time of reaction (min)

v = volume of culture used in assay (ml)

OD_{600} = cell density at the start of the assay

OD_{420} = combination of absorbance by *o*-nitrophenol
and light scattering by cell debris

OD_{550} = light scattering by cell debris.

Saccharomyces cerevisiae

13.6.3

SCREENING FOR β -GALACTOSIDASE-EXPRESSING YEAST COLONIES USING A FILTER LIFT ASSAY

This protocol (based on Durfee et al., 1993, and Staudinger et al., 1993) describes a relatively simple and rapid method for simultaneously analyzing a large number of yeast colonies for their ability to express β -galactosidase. This technique, based on a blue/white color assay of colonies immobilized on filters, is much less laborious than the conventional liquid assay for β -galactosidase (see Basic Protocol 2). This assay can be used to qualitatively measure the level of β -galactosidase expressed by individual yeast colonies based on the intensity of the blue color, but it is less sensitive and quantitative than the liquid β -galactosidase assay. The protocol is conveniently used as a secondary screen on potential two-hybrid positive colonies (UNIT 20.1) that have already come through a previous screening procedure or in genetic screens for mutations affecting the expression of a specific reporter construct.

Materials

Yeast strain containing *lacZ* fusion gene (see Basic Protocol 1)
Selective medium plates (UNIT 13.1)
Liquid nitrogen
Z buffer (see recipe)
20 mg/ml Xgal (Table 1.4.2) in dimethylformamide
30°C incubator
Circular nitrocellulose membrane filters
Whatman 3MM paper

1. Spot or streak yeast colonies onto a plate containing appropriate selective medium and allow to grow at 30°C until they reach optimal size (generally ~48 hr).

Assaying colonies in duplicate will increase the reliability of the results.

2. Place a circular nitrocellulose membrane filter on the plate and press it gently on the surface so that all the colonies are transferred to the membrane.

It is advisable to use reinforced nitrocellulose membranes for this procedure and treat them very gently, as they become extremely fragile when placed in liquid nitrogen.

3. Carefully peel off the membrane and place in liquid nitrogen.

To avoid shattering of the membrane, use a shallow dish containing just enough liquid nitrogen to cover the membrane.

4. Carefully remove the membrane from the liquid nitrogen container using an instrument such as a broad kitchen spatula. Place the membrane on a piece of dry Whatman paper and allow to thaw (<5 min) at room temperature.

5. Place a circular piece of Whatman 3MM paper (cut to fit) into a petri plate. Add 3 to 5 ml of Z buffer containing 1 mg/ml Xgal—enough liquid to soak the paper without flooding it. Place the thawed nitrocellulose membrane onto the Whatman paper and allow the buffer to absorb slowly into the membrane.

Flooding the Whatman 3MM paper will allow diffusion of the colonies and make the screening difficult to interpret.

6. Place the petri dish at 30°C and incubate a few minutes to overnight.

Strong positives should give a blue color after a few minutes of incubation time.

REAGENTS AND SOLUTIONS

Z buffer

- 16.1 g Na₂HPO₄·7H₂O (60 mM final)
 - 5.5 g NaH₂PO₄·H₂O (40 mM final)
 - 0.75 g KCl (10 mM final)
 - 0.246 g MgSO₄·7H₂O (1 mM final)
 - 2.7 ml 2-mercaptoethanol (50 mM final)
- Adjust to pH 7.0 and bring to 1 liter with H₂O. Do not autoclave.

COMMENTARY

Background Information

Cleavage of ONPG by β -galactosidase yields two products, galactose and *o*-nitrophenol. The *o*-nitrophenol product is yellow and can be detected by its absorption at 420 nm. Because ONPG is present in excess in this assay, the amount of *o*-nitrophenol produced is proportional to the amount of enzyme present. This assay is essentially that same as that used for *E. coli* (Miller, 1972).

Critical Parameters

Liquid assay. Always use exponentially growing cells and always normalize results by measuring *lacZ* activity from a control strain containing a previously characterized *lacZ* fusion protein that has been grown in parallel. Reproducibility between measurements taken on different days can be ensured by using the same volume of inoculum in step 2 of Basic Protocol 2 and growing cells to the same OD₆₀₀ each time. The length of the reaction should be from 30 min to 4 hr. Reaction times shorter or longer than this will affect the accuracy of the reaction. Adjust the reaction time by varying the amount of Z buffer used to resuspend the cells. If the fusion is present on a plasmid that is lost at a high frequency ($\geq 10\%$), the percentage of plasmid-bearing cells in the culture should be determined (in step 2) and a correction made in calculating units of β -galactosidase activity.

LacZ filter lift assay. Always use both positive and negative control strains in this assay. The positive control strain should give a strong blue color after a few minutes of incubation time. The negative control is especially important as this may turn a faint shade of blue depending on the strain background and the length of time of development.

Anticipated Results

Liquid assay. For a highly inducible fusion (such as a *GAL1* fusion) 5000 U can be observed, whereas other fusions (such as a *HIS3-lacZ* fusion) have only 1 to 3 U. In some cases, highly inducible fusions on high-copy-number plasmids (such as 2 μ m) may give values lower than expected, presumably due to precipitation of excess β -galactosidase present in the cell; to increase assay accuracy, it may be necessary to clone such fusions onto low-copy YCp vectors or into the yeast genome.

LacZ filter lift assay. A strong two-hybrid interaction strain should give a deep blue color after just a few minutes of development time. Weaker two-hybrid interaction strains may require up to several hours of incubation, and in this instance they should be carefully compared to the negative control to ensure that the color change is not due to strain background.

Time Considerations

For the liquid assay, once mid-log cultures are available, preparation of six samples will take ~45 min. The reaction time can vary between 30 min and 4 hr. Determining OD₄₂₀ and OD₅₅₀ and calculating units require 15 to 30 min. Once the yeast plates are grown, the filter lift assay takes 20 to 30 min followed by anywhere from 5 min to overnight for development, depending on the signal.

Literature Cited

- Durfee, T., Becherer, K., Chen, P.-L., Yeh, S.-H., Yang, Y., Kilburn, A.-E., Lee, W.-H., and Elledge, S.J. 1993. The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes & Dev.* 7:555-569.
- Guarente, L. 1983. Yeast promoters and *lacZ* fusions designed to study expression of cloned genes in yeast. *Methods Enzymol.* 101:181-191.

- Guarente, L., Lauer, G., Roberts, T.M., and Ptashne, M. 1980. Improved methods for maximizing expression of a cloned gene: A bacterium that synthesizes rabbit β -globin. *Cell* 20:543-553.
- Miller, J.H. 1972. *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Staudinger, J., Perry, M., Elledge, S.J., and Olson, E.N. 1993. Interaction among vertebrate helix-loop-helix proteins in yeast two-hybrid system. *J. Biol. Chem.* 268:4608-4611.

Contributed by Ann Reynolds
University of Washington
Seattle, Washington

Victoria Lundblad
Baylor College of Medicine
Houston, Texas

David Dorris and Marie Keaveney
(filter lift assay)
Harvard Medical School
Boston, Massachusetts

MANIPULATION OF YEAST GENES

Introduction of DNA into Yeast Cells

The most commonly used yeast transformation protocol is the lithium acetate procedure (basic protocol). It is reasonably fast and provides a transformation efficiency of 10^5 to 10^6 transformants/ μg . This efficiency rivals that achieved for most, but not all, strains with the more difficult and time-consuming spheroplast procedure (first alternate protocol). However, the fastest and easiest of the transformation methods is electroporation (second alternate protocol). For a number of strains, electroporation offers the highest transformation efficiency, and may prove especially useful with limiting quantities of transforming DNA. Unlike the lithium acetate procedure, however, electroporation saturates at low DNA levels, restricting its general utility.

NOTE: All solutions and glassware coming into contact with yeast cells must be sterile. Traces of soap on glassware may decrease the transformation efficiency. In addition, the water used for washes and in solution preparation must be of the highest quality; see reagents and solutions for guidelines.

TRANSFORMATION USING LITHIUM ACETATE

**BASIC
PROTOCOL**

The lithium acetate method is based on the fact that alkali cations make yeast competent to take up DNA. After yeast is briefly incubated in buffered lithium acetate, transforming DNA is introduced with carrier DNA. Addition of polyethylene glycol (PEG) and a heat shock trigger DNA uptake. The yeast are then plated on selective media.

Materials

- YPD medium (UNIT 13.1)
- Yeast strain to be transformed
- YPAD medium: YPD medium supplemented with 30 mg/liter adenine hemisulfate
- Highest-quality sterile H_2O
- 10 \times TE buffer, pH 7.5 (modify 1 \times recipe in APPENDIX 2), sterile
- 10 \times lithium acetate stock solution: 1 M lithium acetate, pH 7.5 (adjust pH with dilute acetic acid), filter sterilized
- DNA: high-molecular-weight, single-stranded carrier DNA (see support protocol) and transforming DNA
- 50% (w/v) PEG 4000 or 3350 (do not use PEG 8000), filter sterilized
- CM dropout plates (UNIT 13.1) prepared with Difco agar
- 30 $^\circ\text{C}$ incubator with shaker
- Sorvall GSA and SS-34 rotors (or equivalents)
- 42 $^\circ\text{C}$ water bath

Grow and prepare the yeast cells

1. Two days before the experiment, inoculate 5 ml YPD medium with a single yeast colony of the strain to be transformed. Grow overnight to saturation at 30 $^\circ\text{C}$.
The saturated overnight culture may, if desired, be prepared up to several weeks in advance of the transformation and stored at 4 $^\circ\text{C}$.
2. The night before transformation, inoculate a 1-liter sterile flask containing 300 ml YPAD medium with an appropriate amount of the saturated culture and grow overnight at 30 $^\circ\text{C}$ to 1×10^7 cells/ml ($\text{OD}_{600} \cong 0.3$ to 0.5, depending on strain). For 2- to 3-fold higher efficiency, dilute at this point to 2×10^6 cells/ml in fresh YPAD medium and grow for another 1 to 2 generations (2 to 4 hr).

*Saccharomyces
cerevisiae*

The presence of adenine in the medium produces a slightly higher transformation efficiency, especially with ade^- strains.

It is often difficult to know in advance how large an inoculum from the overnight culture will produce a density of 1×10^7 cells/ml at a reasonable time the next day. Because growth phase and cell density are important in achieving the highest transformation efficiency, the easiest approach is to inoculate three independent flasks with varying amounts of the saturated culture (try 1, 5, and 25 μ l for a 12-hr growth period), then check the OD_{600} of each flask the next day.

3. Harvest cells by centrifuging 5 min at $4000 \times g$ (5000 rpm in Sorvall GSA rotor), room temperature. Resuspend in 10 ml highest-quality sterile water.
4. Transfer to a smaller centrifuge tube and pellet cells by centrifuging 5 min at 5000 to $6000 \times g$ (7000 rpm in SS-34 rotor), room temperature.
5. Resuspend in 1.5 ml buffered lithium solution, freshly prepared as follows:
 - 1 vol 10 \times TE buffer, pH 7.5
 - 1 vol 10 \times lithium acetate stock solution
 - 8 vol sterile water.

Cells may be incubated ≤ 1 hr at 30°C before adding the transforming DNA (step 6), although this incubation is not required. In addition, the yeast may be stored ≤ 2 weeks at 4°C in this solution before proceeding with transformation. The transformation efficiency will be much lower, but suffices for routine introduction of plasmids.

Transform yeast cells

6. For each transformation, mix 200 μ g carrier DNA with ≤ 5 μ g transforming DNA in a sterile 1.5-ml microcentrifuge tube. Keep total volume of DNA ≤ 20 μ l.

This protocol will yield a linear increase in the number of yeast transformants with increasing input of transforming DNA over a large range (1 ng to 5 μ g).

Maximal transformation efficiency is achieved by repeating the denaturation cycle (boiling and chilling) of carrier DNA immediately prior to use.

7. Add 200 μ l yeast suspension to each microcentrifuge tube.

If the volume of DNA exceeds 20 μ l, add appropriate volumes of 10 \times lithium acetate stock solution and 10 \times TE buffer to prevent further ($>10\%$) dilution of the lithium acetate and TE concentrations.

8. Add 1.2 ml PEG solution, freshly prepared as follows:

- 8 vol 50% PEG
- 1 vol 10 \times TE buffer, pH 7.5
- 1 vol 10 \times lithium acetate stock solution.

Shake 30 min at 30°C .

9. Heat shock exactly 15 min at 42°C . Microcentrifuge 5 sec at room temperature.

Longer or shorter duration of the heat shock will decrease transformation efficiency.

10. Resuspend yeast in 200 μ l to 1 ml of 1 \times TE buffer (freshly prepared from 10 \times stock) and spread up to 200 μ l onto CM dropout plates made with Difco agar.

Difco agar gives 3-fold higher transformation efficiency than GIBCO/BRL agar.

11. Incubate at 30°C until transformants appear.

Transformants will be visible on the surface of the agar 2 to 5 days after transformation, depending on the strain, the plasmid, and the selection.

SPHEROPLAST TRANSFORMATION

This method is more time-consuming than the lithium procedure, but can result in a higher efficiency of transformation per input DNA. Thus, it is the preferred method when the DNA to be transformed is in limiting quantities or when a high efficiency of transformation is desirable—e.g., when screening a genomic plasmid bank.

Yeast cells are prepared for transformation by enzymatic digestion of the cell wall, followed by several gentle isotonic washes (to remove the digestive enzyme). DNA is mixed with the resulting spheroplasts and treated with a high-molecular-weight polymer, polyethylene glycol (PEG), to promote DNA uptake. After removal of the PEG, spheroplasts and DNA are resuspended in osmotically stable regeneration agar and are plated on appropriate selective plates.

Materials

YPD medium and plates (*UNIT 13.1*)
Yeast strain to be transformed
1 M sorbitol
2-mercaptoethanol (2-ME)
Glusulase (Du Pont NEN)
CaCl₂ solution
Sorbitol/CaCl₂ solution
DNA: transforming DNA and 5 mg/ml carrier DNA (sheared calf thymus or salmon sperm DNA; *UNIT 6.3*)
PEG/CaCl₂ solution
Selective regeneration agar
CM drop-out plates (*UNIT 13.1*)

30°C incubator with rotating platform
55°C water bath

NOTE: Before starting, melt selective regeneration agar (microwave at low setting), aliquot 10-ml samples into sterile glass test tubes, and place in a 55°C water bath.

Grow the yeast strain

1. Two days before the experiment, inoculate 5 ml YPD medium with a single yeast colony of the strain to be transformed. Grow overnight to saturation at 30°C (or lower temperature, if strain is temperature sensitive).
2. The night before transformation, inoculate a 250-ml sterile flask containing 50 ml YPD with an appropriate amount of the 5-ml culture and grow overnight to $1-2 \times 10^7$ cells/ml ($OD_{600} \cong 0.5$ to 1.0, depending on strain).

The frequency of transformation drops with cell densities higher than this. If the exact inoculum for a given strain is not known, it is best to inoculate three independent flasks with varying amounts of the saturated 5-ml overnight as described in step 2 of the basic protocol.

Prepare the spheroplasts

3. Pellet cells in a tabletop centrifuge 5 min at $1100 \times g$ (2500 rpm) and resuspend in 10 ml of 1 M sorbitol. Repeat.

These and subsequent spins/resuspensions can be performed at room temperature.

4. Pellet cells once more and resuspend in 5 ml of 1 M sorbitol. Make 10^{-5} dilution in sterile water and plate 0.1 ml onto a YPD plate. Incubate 2 days at 30°C.
5. Transfer resuspended cells to a 50-ml sterile flask and add 5 μ l 2-ME and 150 μ l glusulase. Incubate ~30 to 60 min at 30°C with very gentle shaking. Plate 0.1 ml of

ALTERNATE PROTOCOL

*Saccharomyces
cerevisiae*

13.7.3

a 10^{-3} dilution (made in sterile water) onto a YPD plate and incubate 2 days at 30°C.

The number of colonies that grow up on this plate can be compared to the number that grew on the plate from step 4 to calculate the percentage of spheroplast formation. Each new lot of glucanase and each new strain should be assayed for the percentage of spheroplast formation.

Spheroplast formation can also be monitored during the 30- to 60-min incubation with glucanase by observing the degree of lysis in water under a microscope (spheroplasts will lyse in a nonisotonic medium such as water, whereas intact yeast cells will not). At 15-min intervals, remove 10 μ l of the cells, place on a microscope slide, and cover with a cover slip. After focusing (at 240 \times to 400 \times magnification), gently touch a drop of water to the edge of the cover slip. As the water leaks under the cover slip, observe how many cells lyse. Spheroplasts should swell slightly before lysis; upon lysis, a “ghost” (membrane) should still be visible.

6. Transfer spheroplasts to a sterile 50-ml, round-bottom centrifuge tube and pellet in a tabletop centrifuge 4 min at 400 \times g (1500 rpm). Gently decant supernatant without dislodging pellet. Add 2 ml of 1 M sorbitol; resuspend pellet by gently swirling liquid across surface of pellet (which should easily come off the side of the tube).

DO NOT vortex or otherwise vigorously agitate spheroplasts. If pellet does not easily resuspend, reduce either the time or the rpm of subsequent spins.

7. Add 8 ml of 1 M sorbitol and pellet at 400 \times g. Repeat steps 6 and 7.
8. Repeat step 6, add 7 ml of 1 M sorbitol and 1 ml CaCl_2 solution, mix gently by swirling and centrifuge 4 min at 400 \times g.
9. Resuspend in 1 ml sorbitol/ CaCl_2 solution.

It has been reported that spheroplasts can be prepared and stored at -70°C in 1 M sorbitol/15% DMSO and subsequently rethawed for use, with about a 5- to 10-fold reduction in transformation efficiency (Orr-Weaver et al., 1983).

Transform the spheroplasts

10. Mix 150 to 200 μ l cells with the DNA to be transformed (up to 10 μ g) plus 10 μ l carrier DNA. Incubate 10 min at room temperature.

The total volume of added DNA should be no greater than $1/10$ the volume of the cells. For each transformation, it is advisable to mix an aliquot of cells with only carrier DNA, to monitor the frequency of reversion of the relevant mutation.

The amount of transforming DNA to be added depends on several parameters, including the frequency with which the strain can be transformed (determined experimentally) and the state of the DNA (fragments of DNA without replication origins will transform several orders of magnitude less efficiently than intact YEp, YRp or YCp plasmids—see UNIT 13.4).

11. Add a 10-fold volume of PEG/ CaCl_2 solution and thoroughly resuspend. Let sit 10 min at room temperature.
12. Pellet in tabletop centrifuge 4 min at 400 \times g and decant the PEG-containing supernatant.

Cells can also be plated in regeneration agar without removal of PEG, as long as the volume of PEG + cells is $1/10$ to $1/20$ the volume of regeneration agar added. The presence of PEG, however, will cause a decline in transformation frequency.

13. Gently resuspend pellet in 0.5 ml of sorbitol/ CaCl_2 solution and pipet into 10 ml melted regeneration agar tempered to 55°C. Vortex briefly to mix and immediately pour onto the appropriate CM dropout (selective) plate, swirling the plate to further mix cells and agar together.

The 3% agar will quickly solidify at temperatures <55°C, so it is important to plate each 10-ml aliquot as rapidly as possible.

14. Incubate at 30°C (or other appropriate temperature) until colonies appear, both on the surface of the regeneration agar overlay and embedded in the agar. Pick individual colonies and streak for single colonies on the same selective plates. If the purpose of this transformation was to disrupt genomic sequences (UNIT 13.10), make DNA from several transformants in order to analyze the relevant chromosomal region by Southern blot analysis (UNIT 13.11).

TRANSFORMATION BY ELECTROPORATION

Yeast cells are concentrated 1000-fold from a log-phase culture using multiple washes with sterile water to remove all extracellular ions. The concentrated cells are suspended in 1 M sorbitol for osmotic stabilization. DNA is introduced without carrier, and the cells are transformed using an exponential electric pulse delivered by an electroporation device. The transformed cells are spread on a selective agar plate containing 1 M sorbitol. For maximal efficiency, yeast may be incubated before concentration with DTT and lithium acetate to render the cell wall/membrane more permeable to DNA.

Additional Materials

- 1 M dithiothreitol (DTT; filter sterilize and store at –20°C)
- 1 M sorbitol
- Sorbitol selection plates
- Gene Pulser with Pulse Controller (Bio-Rad) *or* Cell-Porator (GIBCO/BRL)
- 0.2-cm-gap disposable electroporation cuvettes (Bio-Rad) *or* 0.15-cm-gap microelectroporation chambers (GIBCO/BRL); ice-cold

Grow and harvest the yeast cells

1. Two days before the experiment, inoculate 5 ml of YPD medium with a single yeast colony of the strain to be transformed. Grow overnight to saturation at 30°C.

The saturated overnight culture may be prepared up to several weeks in advance of the transformation and stored at 4°C.

2. The night before transformation, inoculate a 2-liter sterile flask containing 500 ml YPD with an appropriate amount of the saturated culture and grow overnight with vigorous shaking at 30°C to 1×10^8 cells/ml ($OD_{600} \cong 1.3$ to 1.5, depending on strain).

This cell density is achieved in mid- to late-log phase. It is often difficult to know how much inoculum from the overnight culture will produce a cell density of 1×10^8 cells/ml at a reasonable time the next day. If the exact inoculum for a given strain is not known, inoculate three independent flasks with varying amounts of the saturated culture as described in step 2 of the basic protocol.

3. Harvest culture by centrifuging at $4000 \times g$ (5000 rpm) 4°C, and resuspend vigorously in 80 ml sterile H₂O. To increase electrocompetence of the cells, proceed to step 4. If this treatment is not required, proceed to step 7.

The rotor, the exact speed, and the duration of centrifuge spins are not critical; the principal consideration is to centrifuge sufficiently hard (usually $\sim 4000 \times g$) to pellet all of the yeast, but not so hard as to make resuspension difficult. For large volumes, 5000 rpm in a Sorvall GSA rotor will suffice; for smaller volumes, 7000 rpm in a Sorvall SS-34 rotor is recommended. This and subsequent centrifugation steps should be conducted at 4°C.

ALTERNATE PROTOCOL

*Saccharomyces
cerevisiae*

13.7.5

Treat with lithium acetate and DTT to increase electrocompetence

This treatment is optional. It increases the handling time in preparing the yeast and should not be performed if the cells are to be frozen for subsequent use.

4. Add 10 ml of 10× TE buffer, pH 7.5. Swirl to mix.
5. Add 10 ml of 10× lithium acetate stock solution (see first basic protocol). Swirl to mix. Shake gently 45 min at 30°C.
6. Add 2.5 ml of 1 M DTT while swirling. Shake gently 15 min at 30°C.

This treatment increases the transformation efficiency >5-fold. Both lithium acetate and DTT, if used individually, will increase transformation efficiency. Used together, the improvement is approximately additive.

Concentrate and wash the yeast cells

7. Dilute yeast suspension to 500 ml with water.
8. Wash and concentrate the cells three times by centrifuging at 4000 to 6000 × g, resuspending the successive pellets as follows:
 - First pellet—250 ml ice-cold water
 - Second pellet—20 to 30 ml ice-cold 1 M sorbitol
 - Third pellet—0.5 ml ice-cold 1 M sorbitol.

Resuspension should be vigorous enough to completely dissociate each pellet. The final volume of resuspended yeast should be 1.0 to 1.5 ml and the final OD₆₀₀ should be ~200.

Yeast can be stored at -70°C for subsequent use by addition of glycerol to 15% (v/v) followed by freezing in a dry ice/ethanol bath. Frozen aliquots should be thawed slowly, pelleted in a sterile microcentrifuge tube, and resuspended to the same volume in 1 M sorbitol. The wash is necessary to remove ions liberated by lysis of some yeast cells. Transformation efficiency will drop ≥10-fold with freezing.

Electroporate the yeast cells

Using the Bio-Rad Gene Pulser:

- 9a. In a sterile, ice-cold 1.5-ml microcentrifuge tube, mix 40 µl concentrated yeast cells with ≤100 ng transforming DNA contained in ≤5 µl.

Transforming DNA should be in a low-ionic-strength buffer such as TE or in sterile high-quality water (see reagents and solutions). There is no required length of incubation; this time can be varied to convenience. Do not include carrier DNA in this procedure; it drastically reduces transformation efficiency. Maximal efficiency (transformants/µg) will be obtained with <10 ng of transforming DNA, while the largest number of transformants will be obtained with 100 ng. Using >100 ng will cause a precipitous drop in both transformation efficiency and the total number of transformants, with few colonies evident in transformation with 1 µg DNA.

- 10a. Transfer to an ice-cold 0.2-cm-gap disposable electroporation cuvette.
- 11a. Pulse at 1.5 kV, 25 µF, 200 Ω. It is crucial that the Bio-Rad Pulse Controller be included in the circuit; failure to do so will result in damage to the Gene Pulser.

The time constant reported by the Gene Pulser will vary from 4.2 to 4.9 msec. Times <4 msec or the presence of a current arc (evidenced by a spark and smoke) indicate that the conductance of the yeast/DNA mixture is too high.

- 12a. Add 1 ml ice-cold 1 M sorbitol to the cuvette and recover the yeast, with gentle mixing, using a sterile 9-inch Pasteur pipet.

- 13a. Spread aliquots of the yeast suspension directly on sorbitol selection plates. Incubate at 30°C until colonies appear.

Although there is a conflicting report (Lorow-Murray and Jessee, 1991), we find that the addition of sorbitol as an osmotic protectant to the agar increases survival of transformed yeast, yielding 10-fold higher transformation efficiency (Becker and Guarente, 1991).

Transformants will be visible on the agar surface 3 to 6 days after transformation, depending on the strain, the plasmid, and the selection. This is usually 1 day longer than with the lithium acetate procedure.

Using the BRL Cell-Porator:

- 9b. In a sterile, ice-cold 1.5-ml microcentrifuge tube, mix 20 μ l concentrated yeast with \leq 100 ng transforming DNA contained in \leq 5 μ l.

Incubation time can be varied for convenience (step 9a).

- 10b. Transfer to an ice-cold, 0.15-cm-gap micro-electroporation chamber.
- 11b. Pulse at 400V, 10 μ F, low resistance.
- 12b. Remove 10 μ l of electroporated mixture to a sterile 1.5-ml microcentrifuge tube containing 0.5 ml ice-cold 1 M sorbitol.
- 13b. Spread aliquots of the yeast suspension directly on sorbitol selection plates. Incubate at 30°C until colonies appear.

PREPARATION OF SINGLE-STRANDED HIGH-MOLECULAR-WEIGHT CARRIER DNA

SUPPORT PROTOCOL

Addition of denatured high-molecular-weight carrier DNA is critical to achieving high transformation efficiency using the lithium acetate protocol. Both parameters are crucial: the DNA must be single-stranded, and the larger the fragments, the better the transformation efficiency. Carrier DNA prepared according to this support protocol may also be used for spheroplast transformations, although single-stranded DNA has been reported to give a slightly lower transformation frequency than the more easily prepared double-stranded calf thymus or salmon sperm carrier DNA. (*NOTE:* Addition of this or any other carrier DNA during electroporation drastically reduces the efficiency of transformation.)

Materials

DNA (type III sodium salt from salmon testes; Sigma #D1626)

1 \times TE buffer, pH 8.0 (*APPENDIX 2*)

Buffered phenol (*UNIT 2.1*)

1:1 (v/v) phenol/chloroform

Chloroform

3 M sodium acetate, pH 5.2 (*APPENDIX 2*)

100% ethanol, ice-cold

Probe sonicator

Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5*) and phenol extraction and ethanol precipitation of DNA (*UNIT 2.1*)

1. Dissolve DNA in 1 \times TE buffer, pH 8.0, to a final concentration of 10 mg/ml. Stir overnight at 4°C.

The solution of DNA will be extremely viscous.

Because the correct preparation of carrier DNA is crucial in achieving high transformation efficiency by the lithium acetate protocol, and because each individual transformation

*Saccharomyces
cerevisiae*

13.7.7

requires 200 μg , prepare a large amount (e.g., 1 g) and freeze aliquots at -20°C .

2. Sonicate using a large probe until viscosity appears to decrease slightly. Run 1 μg on a 0.8% agarose gel to determine size distribution of sonicated fragments. Repeat sonication as necessary to achieve the appropriate size distribution.

Each sonication should be brief— ~30 sec at 75% power (see UNIT 6.3 for guidelines on sonication of herring sperm DNA). The larger the fragments, the better the transformation efficiency will be, but the more viscous and unwieldy the solution. The optimum distribution of fragment size is that in which fragments range from 2 kb to 15 kb, with a mean size of ~7 kb.

3. Extract once with buffered phenol, once with phenol/chloroform, and once with chloroform.
4. Precipitate DNA with $\frac{1}{10}$ vol of 3 M sodium acetate, pH 5.2, and 2.5 vol ice-cold 100% ethanol.

With DNA of this size and at this concentration, the DNA will form flocculent masses immediately upon addition of the ethanol.

5. Resuspend pellet with $1\times$ TE buffer to 10 mg/ml final concentration.

Resuspension may require stirring overnight at 4°C .

6. Transfer carrier DNA to a Pyrex flask. Microwave to a rolling boil and continue to boil 2 to 3 min.
7. Chill flask rapidly in ice water. Aliquot and freeze DNA in sterile tubes at -20°C .

REAGENTS AND SOLUTIONS

CaCl₂ solution

0.1 M Tris-Cl, pH 7.4 (APPENDIX 2)

0.1 M CaCl₂

Autoclave or filter sterilize and store at room temperature

PEG/CaCl₂ solution

45% (w/v) PEG 3350

10 mM Tris-Cl, pH 7.4 (APPENDIX 2)

10 mM CaCl₂

Filter sterilize and store at room temperature

Selective regeneration agar

5 g ammonium sulfate

1.7 g yeast nitrogen base (without amino acids or ammonium sulfate)

20 g dextrose

30 g agar

1 pellet NaOH

1.3 g appropriate amino acid dropout powder (Table 13.1.1)

To above ingredients, add 500 ml of 2 M sorbitol, 20 ml YPD medium (UNIT 13.1), and 480 ml water. Mix well and autoclave 15 min in 2-liter flask. Aliquot 250-ml portions into 500-ml sterile bottles, using sterile technique. Store at room temperature or 4°C .

Sorbitol selection plates

Supplement CM dropout plates (UNIT 13.1) with sorbitol to 1 M final concentration. Sorbitol can be added as a powder prior to autoclaving. *NOTE:* Addition of sorbitol makes the agar more viscous and more prone to boiling over upon removal from the autoclave. Store at room temperature or 4°C .

Sorbitol/CaCl₂ solution

1 M sorbitol

10 mM Tris·Cl, pH 7.4 (*APPENDIX 2*)

10 mM CaCl₂

Filter sterilize or autoclave and store at room temperature

Water, high-quality and sterile

Water used in these protocols for washes and in solution preparation must be of the highest-possible quality (e.g., Millipore Milli-Q); the resistance must be at least as high as 10 MΩ/cm. Sterilize by autoclaving and store at room temperature.

COMMENTARY

Background Information

High-frequency transformation in yeast was first demonstrated using β-glucanase to remove the cell wall to generate spheroplasts (Beggs, 1978; Hinnen et al., 1978). The method has several major disadvantages. Of the procedures for transforming yeast, it is the most tedious and difficult. Yeast transformed by spheroplasting must be embedded in top agar to allow regeneration of the cell wall, which makes subsequent manipulations, such as replica plating or colony hybridization (*UNIT 13.2*), extremely difficult. In addition, the PEG treatment of spheroplasts often generates diploids and triploids by cell fusion. Nonetheless, spheroplast transformation has long remained a preferred approach for yeast transformation because it offered, until recently, the highest transformation efficiencies for most yeast strains.

More recent transformation protocols, such as that employing lithium acetate (Ito et al., 1983), do not require digestion of the cell wall and thus do not require regeneration in top agar—rather, transformants are plated on the agar surface. With lithium acetate transformation, competence for DNA uptake at the level of the cell wall is induced by treatment with lithium ions plus PEG. This procedure is easier and more rapid than spheroplasting. The finding that single-stranded, high-molecular-weight carrier DNA could increase transformation efficiency by one to two orders of magnitude (Schiestl and Gietz, 1989) has made lithium acetate transformation at least as efficient as spheroplast transformation. This modified lithium acetate procedure will likely replace spheroplast transformation for most strains.

The final and simplest approach is electroporation (Becker and Guarente, 1991). Yeast are harvested during log-phase growth, concentrated, and suspended in an osmotically stabi-

lizing medium with transforming DNA. A high-voltage electrical discharge through this suspension drives the DNA uptake. Electroporation can be extremely efficient, especially with small quantities of DNA (1 pg to 10 ng). When DNA quantities are limited, this may be the procedure of choice. Electroporation has several drawbacks, however. First, the expense of the electroporator apparatus may be prohibitive for some labs. Second, the efficiency of transformation is extremely strain-dependent. Third, the process saturates; increasing the input above 100 ng decreases the yield of transformants. Electroporation is inappropriate, therefore, for experiments that require maximizing the total number of transformants, such as screening genomic or cDNA expression libraries in yeast.

Critical Parameters

Transformation efficiency in all of these procedures depends on the yeast strain background. Lithium acetate transformation is least affected by strain variation, while electroporation is most affected. In the spheroplast method, strain-dependent variations in frequency may reflect differing sensitivities to glucanase. In electroporation, strain-dependent variations may reflect differing degrees of survival after the electric pulse. Since the response of a strain cannot be predicted a priori, strains known to transform at high efficiency should be used whenever possible. When strain selection is dictated by the particular experiment, all the protocols might have to be tested to determine which works best.

Maximal transformation efficiency in each of the procedures also depends on finding the optimal cell density. Each protocol lists the preferred starting density; strain differences may require empiric variation. The condition of the DNA has an effect as well; miniprep DNA prepared by either alkaline lysis or boil-

ing (UNIT 1.6) will transform adequately, but phenol extraction (UNIT 2.1) or purification by CsCl/ethidium bromide density gradients (UNIT 1.7) will further increase the efficiency.

In the lithium acetate protocol, the single most critical parameter is the carrier DNA. It must be of high-molecular-weight and completely denatured (Schiestl and Gietz, 1989). Denaturation can be assured by repeating the cycle of boiling and chilling immediately before use. Achieving the proper size distribution may be more difficult. If low efficiencies are obtained with several strains, another batch of carrier DNA should be prepared. In addition, using Difco agar to prepare the CM dropout plates increases the transformation efficiency.

In the spheroplast procedure, the most critical parameter is the handling of spheroplasts—vortexing, osmotic shock, or plating in regeneration agar at temperatures $>55^{\circ}\text{C}$ will reduce transformation frequency. For both the lithium and spheroplast protocols, transformation efficiency also depends on the batch of PEG used. If poor frequencies are obtained, several lots of PEG must be tested. Be certain to use PEG 3350, not PEG 8000.

In electroporation, the most important parameter after strain selection is the amount of transforming DNA used. Maximal efficiency (transformants/ μg) will be obtained with <10 ng of transforming DNA; the largest number of transformants will be obtained with 100 ng. Using more than 100 ng of plasmid or including carrier DNA will cause a precipitous drop in both transformation efficiency and total number of transformants. The electrical parameters are also important. With either of the exponential decay devices described, strain differences may require empiric variation of the initial voltage around the suggested value. For square-wave devices, see Meilhoc et al. (1990).

For all protocols, it is important that sterile technique be used throughout. Traces of soap on glassware may decrease the transformation efficiency.

Anticipated Results

Under optimal conditions, a circular plasmid (e.g., YCp or YEp; UNIT 13.4) should give 10^5 to 10^6 transformants/ μg DNA with the lithium acetate protocol, $\sim 10^4$ to 10^5 by spheroplasting, and $>10^5$ with electroporation ($>10^4$ transformants at 100 ng). With either linear molecules or circular molecules that do not

contain a yeast replication origin, stable transformation requires integration into the genome; transformation frequencies are correspondingly lower, with $\sim 10^2$ to 10^3 transformants/ μg DNA.

Time Considerations

Once the cells have reached the desired density, the lithium acetate protocol requires ~ 3 hr to complete, spheroplasting requires 3 to 4 hr, and electroporation can be completed in ~ 90 min. Because spheroplasts are metabolically active for up to 24 hr, the time spent washing the spheroplasts can be increased for convenience (e.g., each resuspension step can be left for an hour or more on a gently rocking platform).

Literature Cited

- Becker, D.M. and Guarente, L. 1991. High-efficiency transformation of yeast by electroporation. *Methods Enzymol.* 194:182-187.
- Beggs, J.D. 1978. Transformation of yeast by a replicating hybrid plasmid. *Nature (Lond.)* 275:104-109.
- Hinnen, A., Hicks, J.B., and Fink, G.R. 1978. Transformation of yeast. *Proc. Natl. Acad. Sci. U.S.A.* 75:1929-1933.
- Ito, H., Fukuda, Y., Murata, K., and Kimura, A. 1983. Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* 153:163-168.
- Lorow-Murray, D. and Jessee, J. 1991. High-efficiency transformation of *Saccharomyces cerevisiae* by electroporation. *Focus* 13:65-67.
- Meilhoc, E., Masson, J.-M., and Teissie, J. 1990. High efficiency transformation of intact yeast cells by electric field pulses. *BioTechnology* 8:223-227.
- Orr-Weaver, T.L., Szostak, J.W., and Rothstein, R.J. 1983. Genetic applications of yeast transformation with linear and gapped plasmids. *Methods Enzymol.* 101:228-244.
- Schiestl, R.H. and Gietz, R.D. 1989. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr. Genet.* 16:339-346.

Contributed by Daniel M. Becker
Stanford Law School
Stanford, California

Victoria Lundblad
Baylor College of Medicine
Houston, Texas

Cloning Yeast Genes by Complementation

UNIT 13.8

**BASIC
PROTOCOL**

Yeast genes have been cloned by a variety of techniques, including use of purified RNA as hybridization probes, differential hybridization of regulated RNA transcripts, antibody screening, transposon mutagenesis, cross suppression of mutant phenotypes, cross hybridization with heterologous cDNA or oligonucleotide probes, as well as by complementation in *E. coli* (for reviews of these methods see Rothstein, 1985, or Rose, 1987). This unit presents a generalized protocol and describes the principles involved in cloning yeast genes by complementation in yeast.

The protocol is presented using a hypothetical mutation of yeast, the *cdc101-1* mutation. This mutation was isolated as a cell cycle mutant and is both recessive and temperature-sensitive for growth: it can grow relatively normally at 30°C but is unable to make a colony at 37°C. A genomic DNA clone that complements this mutation will be isolated by transforming the *cdc101-1* strain with a yeast genomic library and subsequently screening for temperature-resistant colonies. Once isolated, two steps are necessary to prove that the insert present on the plasmid contains the wild-type *CDC101* gene. First, segregation of the complementing plasmid must result in co-loss of both the plasmid-borne selectable marker and the complementing phenotype, demonstrating that the observed complementation is plasmid-specific and is not due to reversion of the *cdc101-1* mutation. Second, it must be ruled out whether the cloned gene encodes a phenotypic suppressor of the mutation, rather than the wild-type gene. This is done via a complementation test, which demonstrates whether or not a disruption of the cloned gene that is integrated into the genome can complement the original mutation.

Cloning the Gene

Transform a mutant yeast strain

1. Transform a *leu2⁻ cdc101-1* yeast strain with a yeast genomic DNA library containing *LEU2* as a selectable marker (using either protocol in UNIT 13.7) and selecting, in this case, for *Leu⁺*. Depending on the insert size, between 2,000 and 20,000 transformants must be screened (~4 to 8 genome equivalents).

If the mutation reverts at a very low frequency (<10⁻⁸), transformants that complement the mutant phenotype (in this case, temperature sensitivity) can be selected directly after transformation. However, if the mutation reverts at higher frequencies (or if the complementing phenotype cannot be directly selected), the mutation must be present in a strain that also contains a low or nonreverting mutation in the selectable gene present on the library vector.

Whenever possible, it is also preferable to have the mutation present in a strain that transforms well. If the mutation has been isolated in a strain that transforms poorly, alternate transformation protocols can be tested (see UNIT 13.7). Transformation proficiency can also be improved by backcrosses with a known strain that is more competent for transformation.

Screen for complementation of the mutant phenotype

2. Replica plate transformants onto prewarmed selective plates (leucine dropout plates; UNIT 13.1) and incubate at 37°C, to screen for complementation of the temperature-sensitive phenotype. With overnight incubation, colonies containing plasmids with inserts that complement the *cdc101-1* mutation, as well as potential *CDC101⁺* revertants that contain random noncomplementing plasmids, will grow up. Restreak each single colony that appears on this plate and save for further analysis.

*Saccharomyces
cerevisiae*

13.8.1

Contributed by Victoria Lundblad

Current Protocols in Molecular Biology (1989) 13.8.1-13.8.4

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 5

Many mutant phenotypes, including *ts* phenotypes, cannot be distinguished by replica plating (presumably due to leakiness of the mutation). If this is the case, transformants must be recovered from the original transformation plate and replated directly on a second plate that screens for the mutant phenotype. If the lithium method of transformation is used, recover transformants by pipetting about 0.5 ml sterile water onto each plate and resuspending the individual colonies in the liquid by mashing about the plate with a sterile spreader (UNIT 1.3). Pipet the 0.5 ml back into a sterile tube and replate for single colonies onto the appropriate plate. If the spheroplast protocol is used, chop up the agar from each transformant plate (using a sterile spatula) and resuspend in 5 to 10 ml sterile water. Vortex extensively to further break up the agar (which releases the transformants into the liquid) and replate for single colonies.

If the mutation displays more than one phenotype, complementation of the additional phenotypes should also be tested. Complementation of some but not all of the mutant phenotypes indicates that the wild-type gene has not been cloned. However, such a complementation pattern suggests that a gene which performs a related function has been isolated.

Proof That the Correct Gene Has Been Cloned

Determine whether segregation of the complementing plasmid results in co-loss of both the plasmid-borne selectable marker and the complementing phenotype

3. Using the protocol described in UNIT 13.9, isolate Leu⁻ segregants for each candidate and test their ability to grow at 37°C. Save transformants that are only able to grow at 37°C in the presence of the plasmid.

If plasmid segregants can be isolated that still display “complementation,” this suggests that the complementation pattern is due to either reversion of the original mutation or acquisition of a suppressor mutation. Alternatively, if colonies are observed that still retain plasmid but are no longer complemented for the mutant phenotype, this indicates that the original transformant contained a mixed population of plasmids.

4. Isolate plasmid DNA from each transformant identified in step 3 by transforming total yeast DNA into *E. coli* (UNIT 13.11). Retransform each plasmid into the mutant yeast strain and confirm that the correct plasmid has been isolated (in our example, demonstrate that retransformation into a *cdc101-1* strain results in a temperature-resistant phenotype).
5. Analyze the plasmids by restriction mapping (UNIT 3.2) to determine whether genomic DNA inserts present in different plasmids have overlapping segments, which can help define the boundaries of the complementing region. Construction of various deletion and insertion mutations, and tests of their ability to complement the mutation in yeast, will provide more precise information about the location of the gene within the cloned insert.

Determine whether a disruption of the cloned gene that is integrated into the genome can complement the original mutation (complementation test)

The following steps are designed to test whether the cloned gene encodes a phenotypic suppressor of the mutation, rather than the wild-type gene.

6. Introduce a yeast selectable marker (UNIT 13.4) into a site in the middle of the complementing region (based on the information gained in step 5).
7. Transform this disrupted plasmid back into the original mutant strain and screen for the mutant phenotype to test whether complementation has been abolished as follows. Using one of several techniques discussed in UNIT 13.10, construct a diploid strain containing one wild-type copy of the gene and one disrupted copy of the gene. After sporulation and dissection (UNIT 13.2), cross a haploid spore product

containing the disruption (usually identified by the presence of the selectable marker associated with the disruption) is crossed to a strain carrying the original mutation (UNIT 13.2). If this diploid has the same mutant phenotype as the strain with the original mutation, this demonstrates lack of complementation and indicates that the wild-type gene corresponding to this mutation has been cloned. As a control, cross the strain with the gene disruption (which may itself display a non-wild-type phenotype) to a wild-type strain, in order to demonstrate that the phenotype of this diploid is wild-type (showing that the disruption is a recessive mutation).

Often, a genomic disruption of the cloned insert will have a mutant phenotype which is either similar to that of the original mutation or is a phenotype predicted for a null mutation in this gene. Although this is suggestive that the correct gene has been cloned, it does not constitute proof: a related gene may not only compensate for the original mutation when present on a plasmid, but may also give a similar phenotype when disrupted. Complementation, however, will demonstrate whether two different genes have been identified.

In the case of the hypothetical CDC101 gene, the cdc101-1 mutation is a ts lethal, indicating that this is an essential gene and that a haploid strain containing a disruption of the CDC101 gene would be inviable (techniques for determining whether this hypothesis is correct are presented in UNIT 13.10). For such an essential gene, the same complementation test described above can be performed, with one technical modification. Construct a plasmid-borne disruption of the insert region that complements cdc101-1, using URA3 as the selectable marker. Transform a ura3⁻/ura3⁻, leu2⁻/leu2⁻, diploid strain with this disruption (using techniques described in UNIT 13.10), selecting for Ura⁺. Now transform this heterozygous diploid with a LEU2 plasmid bearing an intact copy of the complementing region. After sporulation, identify a spore product that contains both the URA3 disruption and the LEU2 complementing plasmid. This strain can now be crossed to both a cdc101-1 strain and a control CDC101⁺ strain, the Leu⁺ plasmid segregated away (UNIT 13.9), and complementation (or lack of it) observed.

COMMENTARY

Cloning a gene by complementation in yeast requires not only a low-reverting recessive mutation in the gene of interest, but the appropriate yeast genomic DNA library. A number of yeast genomic libraries have been published and may be available (e.g., see Rose et al., 1987; Meeks-Wagner et al., 1986; Kuo and Campbell, 1983). The parameters to consider when selecting a preexisting library are the auxotrophic mutation(s) present in the strain of interest (which determines the selectable marker that must be present on the vector), whether the strain is cir⁺ or cir⁰ (libraries constructed in YEp plasmids are much less stable in cir⁰ strains; see UNIT 13.4), and the copy number of the vector (some genes, when present in high copy, are toxic to yeast). In some cases, it may be necessary to construct a yeast plasmid library *de novo*. This is the case when the only mutation available in yeast is dominant to the wild-type allele. Construction of a library from DNA isolated from the mutant strain will allow isolation of the mutant gene (by screening transformants of a wild-type strain for the mutant phenotype), from which the wild-type gene can be obtained by standard cloning techniques (UNITS 6.1 to 6.3). In addition,

some yeast genes have been demonstrated to be toxic when expressed in *E. coli*. Thus, preexisting yeast libraries which have been amplified through *E. coli* will be under-represented for plasmids containing these genes. This can be solved by constructing a library and transforming directly into yeast. A more detailed discussion of the issues involved in choosing a library, as well as a protocol for constructing a yeast genomic library, are presented in Rose, 1987 (see also Chapter 5).

Cloning yeast genes by complementation when the only available mutation is dominant presents special problems. As mentioned above, a genomic library must be constructed from the mutant strain. In addition, once a candidate gene has been isolated, if either the original mutation or a genomic disruption in this candidate gene are dominant, a complementation test cannot be used to prove that the correct gene has been cloned. In this case, another genetic test can be used to demonstrate tight genetic linkage between a selectable marker introduced at the chromosomal site of the cloned gene and the original mutation. In this test, a plasmid carrying the cloned gene as

well as a selectable marker are integrated into the genome of a wild-type haploid strain, using the technique of integrative transformation (UNIT 13.10). This haploid strain is crossed with a strain carrying the (dominant) mutation, and the subsequent diploid is sporulated and analyzed for the pattern of segregation of the mutation and the selectable marker associated with the integrated gene. If the cloned gene has integrated at the genetic locus corresponding to the original mutation, the selectable marker will always cosegregate with wild-type—i.e., spore products will never be recovered that display both the mutation and the selectable marker. If the cloned gene has integrated at another unlinked genetic locus, the selectable marker will segregate independently of the mutant/wild-type phenotypes, with approximately one-fourth of the spores containing both the mutation and the selectable marker (e.g., however, <25% of the spores will contain both markers if the cloned DNA has integrated at a site that is somewhat linked to the mutation). If tight genetic linkage can be demonstrated in this test, this provides evidence for the identity of the cloned piece of DNA. However, it is possible that a suppressor gene has been cloned that is tightly linked to the original mutant allele; such a possibility would not be ruled out by the above test.

Cloning of heterologous genes by complementation in yeast follows the same general principles presented in this section but requires a library of genomic DNA or cDNA from the relevant organism cloned into a vector that can be maintained in yeast. Although in some cases the heterologous gene can be expressed in yeast by its own promoter, in most cases the gene must be fused to a yeast promoter. Several heterologous libraries containing DNA from *S. pombe* (Beach et al., 1982) and *Drosophila* have already been constructed. Vectors for ex-

pression of genes in yeast that can be used for construction of heterologous libraries are discussed in UNIT 13.4. One principle that cannot be applied to cloned heterologous genes is the use of complementation (or linkage analysis, as described in the previous paragraph) to prove that the correct gene has been cloned. In these situations, evidence that the correct gene has been cloned may require comparison of nucleic acid and protein sequence between the two genes, as well as functional analysis (e.g., demonstrating that the two genes encode the same enzymatic function).

Literature Cited

- Beach, D., Piper, M., and Nurse, P. 1982. Construction of a *Schizosaccharomyces pombe* gene bank in a yeast bacterial shuttle vector and its use to isolate genes by complementation. *Mol. Gen. Genet.* 187:326-329.
- Kuo, C.L. and Campbell, J.L. 1983. Cloning of *Saccharomyces cerevisiae* DNA replication genes: Isolation of the *CDC8* gene and two genes that compensate for the *cdc8-1* mutation. *Mol. Cell. Biol.* 3:1730-1737.
- Meeks-Wagner, D., Wood, J.S., Garvik, B., and Hartwell, L.H. 1986. Isolation of two genes that affect mitotic chromosome transmission in *S. cerevisiae*. *Cell* 44:53-63.
- Rose, M.D. 1987. Isolation of genes by complementation in yeast. *Meth. Enzymol.* 152:481-504.
- Rose, M.D., Novick, P., Thomas, J.H., Botstein, D., and Fink, G.R. 1987. A *Saccharomyces cerevisiae* genomic plasmid bank based on a centromere-containing shuttle vector. *Gene* 60:237-243.
- Rothstein, R. 1985. Cloning in yeast. In *DNA Cloning*, Vol. 2: A Practical Approach (D.M. Glover, ed.) 4th ed., pp. 45-67. IRL Press, Oxford.

Contributed by Victoria Lundblad
University of California, Berkeley
Berkeley, California

Manipulation of Plasmids from Yeast Cells

UNIT 13.9

This unit describes several procedures for manipulating plasmids in yeast cells. The first (see Basic Protocol 1) is a general method to segregate autonomously replicating plasmids from cells: plasmid-containing yeast cells are grown in nonselective medium, and colonies lacking the plasmid are identified by replica plating. The second, plasmid shuffling (see Basic Protocol 2), represents a specialized version of plasmid segregation that is useful for analyzing the function of essential genes and for identifying conditional lethal mutations in essential genes. This method involves introduction of a plasmid containing mutated versions of essential genes into a strain carrying the wild-type gene on a *URA3* plasmid followed by a genetic selection to remove the *URA3* plasmid. The genetic selection for plasmid loss permits the rapid analysis of many thousands of colonies. The third approach, plasmid gap repair (see Basic Protocol 3), is based on the efficient homologous recombination characteristics of yeast cells. Plasmid gap repair can be used as a method to incorporate mutagenized DNA fragments into a yeast plasmid, rescue genomic mutations onto plasmids, or map alleles of a given gene.

SEGREGATION OF PLASMIDS FROM YEAST CELLS

Most *E. coli* yeast shuttle vectors are lost at a frequency of ~1% when grown nonselectively. In this simple protocol, a plasmid-bearing yeast strain is grown nonselectively (to allow segregation of the plasmid) and plated for single colonies. Replica plating onto selective plates then permits identification of colonies that no longer contain the plasmid.

Materials

Plasmid-bearing yeast strain

YPD or other nonselective liquid medium and plates (UNIT 13.1)

CM dropout plates (UNIT 13.1)

Sterile velvets and replica block (UNITS 13.2 & 1.3)

Additional reagents and equipment for growth and manipulation of yeast (UNIT 13.2)

1. Inoculate several single colonies of a plasmid-bearing yeast strain into individual 10-ml aliquots of nonselective medium (UNIT 13.2) and grow overnight at 30°C.

If no other selection is necessary, YPD medium can be used. If, however, other unstable genetic markers or plasmids are to be retained, use defined medium supplemented with the nutrient corresponding to the selectable marker to be lost (e.g., when segregating a LEU2 plasmid, grow in medium supplemented with leucine).

2. Plate for single colonies (UNIT 13.2) on the corresponding nonselective plates and incubate 2 days at 30°C.

In most cases, ~200 to 300 single colonies should be examined (~100/plate).

3. Replica plate onto selective plates that will identify colonies which have lost the plasmid-borne selectable marker. Incubate both the selective plates and the nonselective master plate overnight at 30°C.

Those colonies that are present on the master plate but fail to grow on the selective plate have lost the plasmid.

**BASIC
PROTOCOL 1**

*Saccharomyces
cerevisiae*

PLASMID SHUFFLING

This technique (based on Boeke et al., 1987) provides a method of analyzing the function of essential genes as well as a rapid means for identifying conditional lethal mutations in an essential cloned gene carried on a plasmid (Fig. 13.9.1). In this system, the chromosomal copy of the essential gene is deleted or otherwise disrupted. Viability is maintained by the presence of a YEp plasmid carrying an intact copy of this essential gene, as well as the *URA3* gene. Introduction of a second plasmid carrying a temperature-sensitive copy of this gene can relieve selective pressure on the YEp plasmid at the permissive temperature, generating *Ura*⁻ derivatives due to loss of the YEp plasmid. However, at the nonpermissive temperature, the YEp plasmid cannot be lost and no *Ura*⁻ segregants are generated. Loss of this plasmid is assayed by replica plating single colonies onto 5-FOA plates grown at the permissive and nonpermissive temperatures. On these plates, *Ura*⁻ segregants appear as 5-FOA-resistant papillae. This replica plating technique allows a large number of mutagenized clones to be screened rapidly.

Materials

- YCp vector (UNIT 13.4) bearing a selectable marker other than *URA3*
- Dropout plates +Ura (UNIT 13.1)
- Plates with medium selective for YCp vector and containing 5-FOA (UNIT 13.1)
- YPD plates (UNIT 13.1)
- YIp5 vector (UNIT 13.4)

Additional reagents and equipment for cloning yeast genes by complementation (UNIT 13.8), subcloning (UNIT 3.16), mutagenesis (see Chapter 8), lithium acetate transformation (UNIT 13.7), replica plating (UNIT 13.2), yeast plasmid DNA preparation (UNIT 13.11), *E. coli* transformation (UNIT 1.8), and gene replacement by transplacement (UNIT 13.10)

1. Construct a haploid strain that contains a chromosomal disruption of the essential gene and a YEp or YCp plasmid carrying both the intact essential gene and *URA3* as a selectable marker (see UNIT 13.8 for a discussion of how to construct such a strain).

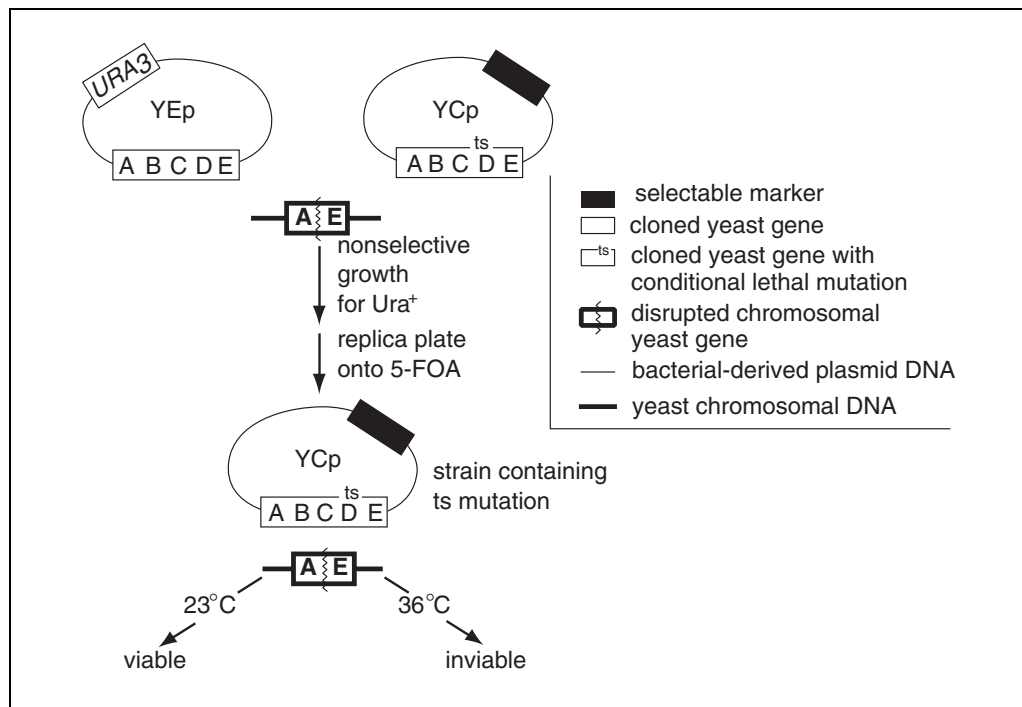


Figure 13.9.1 Plasmid shuffling is a rapid means of identifying conditional lethal mutations in an essential cloned gene carried on a plasmid (see accompanying text).

2. For mutagenesis, subclone (*UNIT 3.16*) the essential gene onto a YCp vector (bearing a selectable marker other than *URA3*) and mutagenize ~10 to 20 μg using hydroxylamine or other techniques (see Chapter 8). Alternatively, clone mutated derivatives of an essential gene into the YCp vector.
3. Transform into the strain from step 1 using lithium acetate (*UNIT 13.7*), selecting for the YCp plasmid on dropout plates that are supplemented with uracil. Incubate plates at the permissive temperature (usually 25°C).

This period of growth in the presence of uracil relieves selection for the YEp plasmid, such that in each transformant colony, a portion of the cells will have lost the plasmid (which carries the unmutagenized copy of the essential gene). However, note that any transformants carrying a null mutation in this gene on the YCp vector will be unable to lose the YEp vector.

4. Replica plate (*UNIT 13.2*) each transformation plate onto two plates that contain 5-FOA and that still maintain selection for the YCp vector. Incubate one plate at the permissive temperature and one plate at the nonpermissive temperature (usually 36°C). As a control, replica plate each transformation plate onto two YPD plates and incubate at the same two temperatures.

Replica plating distinguishes between different categories of transformants by determining whether a portion of the colony has lost the Ura^+ YEp plasmid: such Ura^- cells appear as 5-FOA⁺ papillae growing out of the background of the mostly Ura^+ replica. Transformants that carry an unmutagenized copy of the essential gene on the YCp vector will give Ura^- papillae at both temperatures, whereas transformants that contain a temperature-sensitive mutation in the YCp-borne essential gene will only produce Ura^- papillae at the permissive temperature. Those transformants containing a null mutation in this gene will not generate papillae at either temperature.

Yeast transformation is often somewhat mutagenic, introducing unlinked chromosomal mutations in the recipient. Discard all candidates that display a temperature-sensitive phenotype in the YPD control.

5. Recover from the transformation plate those colonies that gave Ura^- papillae at the permissive temperature but not at the nonpermissive temperature (and which grew normally at both temperatures on the YPD plates) and streak for single colonies.
6. Retest about six single colonies from each candidate for Ura^- papillation at the two temperatures.
7. For each candidate that retests as a temperature-sensitive mutation in step 6, recover Ura^- papillae from the permissive temperature 5-FOA plate (by streaking out papillae on a separate plate that maintains selection for the YCp vector). Recover plasmid from these strains (which now contain only the YCp vector) by isolation of the plasmid (*UNIT 13.11*) followed by transformation into *E. coli* (*UNIT 1.8*).
8. Subclone (*UNIT 3.16*) the gene (now carrying a temperature-sensitive mutation) onto a YIp5 vector, and introduce this mutagenized gene into its chromosomal site using the transplacement technique described in *UNIT 13.10*.

**PLASMID GAP REPAIR FOR LOCALIZED MUTAGENESIS AND
ALLELE REPAIR**

This gap repair technique protocol is useful for several purposes. First, it can be used as a method of localized mutagenesis in which a mutagenized DNA fragment (typically obtained by PCR; *UNIT 15.1*) can be incorporated into a desired DNA molecule in yeast cells (Muhlrad et al., 1992). Second, it can be used to introduce a mutation at a specific position using a mutagenic PCR primer (Brenner et al., 1994). Third, it can be used to rescue genomic mutations onto a plasmid-borne copy of the same gene for subsequent analysis. Fourth, it can be used to map multiple alleles of a given gene. In all of these cases, a gene conversion event initiated by the gapped molecule rescues the introduced DNA fragment or the chromosomal locus covered by the gapped region; this results in yeast strains with autonomously replicating plasmids containing the mutagenized DNA fragment or chromosomal locus of interest (Fig. 13.9.2). The resulting yeast strains can be directly analyzed for a mutant phenotype, and the plasmids can be rescued and analyzed.

Materials

- YRp or YCp plasmid (*UNIT 13.4*) with selectable marker
- Appropriate restriction enzyme(s) (*UNIT 3.1*)
- Yeast strain with mutation corresponding to selectable marker
- Plates with medium selective for the plasmid marker

Additional reagents and equipment for subcloning (*UNIT 3.16*), PCR (*UNIT 15.1*), yeast transformation (*UNIT 13.7*), growth and manipulation of yeast (*UNIT 13.2*), and plasmid segregation (*UNIT 13.9*)

1. Subclone (*UNIT 3.16*) the gene of interest into a YRp or YCp plasmid with an appropriate selectable marker.
2. Gap the plasmid by cutting at two restriction enzyme sites. Gel purify the resulting plasmid fragment.

An overlap of ~200 bp on either side of the gap is often used, although as few as 40 bp will be sufficient.

For introducing a mutation at a specific position, the plasmid may also simply be linearized rather than gapped; however, contamination with uncut plasmid must be avoided. The gap or cut must be as close as possible to the mutated site to ensure a high frequency of recovery of the mutation.

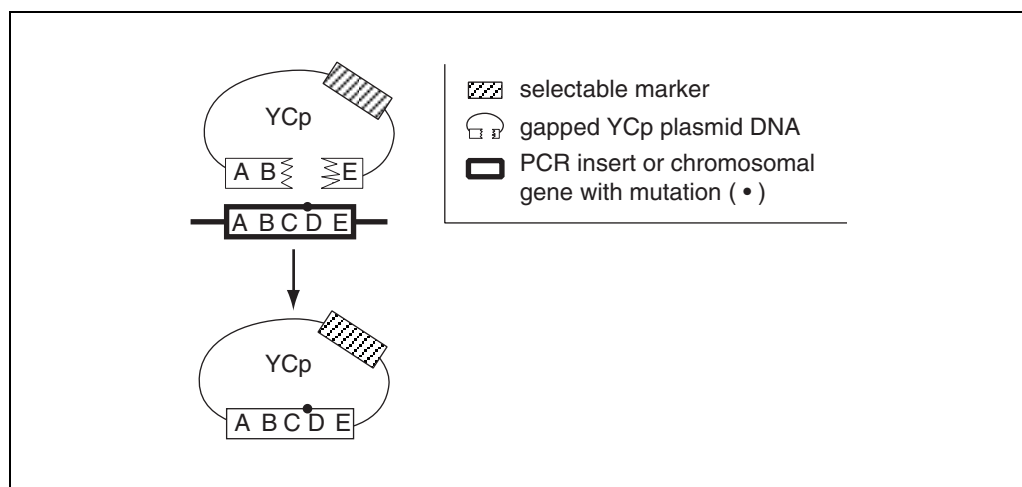


Figure 13.9.2 Plasmid gap repair rescues genomic DNA mutations onto a plasmid-borne copy of the gene (see accompanying text).

3. For mutagenesis applications, prepare the desired mutagenized fragment by PCR (UNIT 15.1). For recovering or mapping chromosomal alleles, proceed to step 4.

A mutagenic PCR reaction using Taq DNA polymerase will introduce mutations at random positions; mutations at a specific position can be made using a mutagenic primer.

4. Co-transform (UNIT 13.7) 1 µg of gapped plasmid and the PCR-generated fragment into a yeast strain with appropriate markers. If using gap repair to rescue the genomic allele, transform with gapped plasmid alone.

Although it is not essential, a 5-fold molar excess of insert is often used.

To ensure that recombination occurs only between the gapped plasmid and the PCR insert, it is preferable to use a yeast strain deleted for the genomic locus. However, the method will work even if the genomic locus is present, although a background of undesired recombination events will arise.

5. Identify successful repair events by plating transformants onto medium selective for the plasmid marker (UNIT 13.2). For recovering or mapping chromosomal alleles, identify transformants that are unstable for the selectable marker using the plasmid segregation technique (UNIT 13.9).

When a mutation is located within the gap, it will be carried on the repaired plasmid. However if it is next to the gap, the outcome will depend on where the cross-over occurs. The farther the mutation is from the gapped region, the lower will be the proportion of plasmids containing the mutation.

6. If using this as a scheme for PCR-based mutagenesis, screen these transformants for the desired mutant phenotype.

*The plasmid can be isolated from yeast (UNIT 13.11) and transformed into *E. coli* (UNIT 1.8) for subsequent analysis.*

COMMENTARY

Background Information

Techniques for segregating plasmids from yeast are useful in several experimental situations. For example, after a plasmid containing a DNA sequence that complements a yeast mutation has been isolated, it is necessary to determine whether complementation is due to sequences present on the plasmid (UNIT 13.8). This is achieved by demonstrating that the complementation phenotype is lost following segregation of the plasmid.

Plasmid shuffling is a specialized version of plasmid segregation which is particularly useful for the analysis or mutagenesis of genes that are essential for cell growth and viability. The manipulation of such essential genes is complicated by the obvious fact that cells containing mutated versions of such genes cannot be propagated. Plasmid shuffling circumvents this problem by generating a yeast strain in which the sole copy of an essential gene is present on a *URA3*-marked plasmid—i.e., the chromosomal copy of the essential gene is deleted. This strain is transformed by a second plasmid (which carries a different marker) containing a

mutated version of the essential yeast gene (carried on a plasmid with a different marker), whereupon the *URA3* plasmid carrying the wild-type genes is eliminated (shuffled out) by replica plating the cells on 5-FOA. This procedure efficiently generates yeast cells that contain only the plasmid with the mutated copy of the essential gene. If the mutated gene is sufficiently functional to support cell growth, the colonies can be examined under a variety of conditions (e.g., low or high temperature). Alternatively, if the mutated gene is nonfunctional, colonies will not arise on the plates containing 5-FOA.

Although yeast plasmids replicate autonomously from chromosomal DNA, the plasmid and chromosomal DNAs are similar in virtually all respects. For this reason, efficient homologous recombination occurs between plasmids and chromosomes, thereby permitting a number of useful genetic manipulations. In particular, the ends of linearized plasmids can recombine efficiently with introduced DNA fragments or with chromosomal DNA in a process called plasmid gap repair. This process permits

Saccharomyces cerevisiae

13.9.5

one to incorporate DNA fragments (either wild-type or mutagenized) or genomic alleles into autonomously replicating plasmids without using standard cloning methods. It is particularly useful for localized mutagenesis, rescuing genomic mutations, and genetic mapping (see Basic Protocol 2).

Literature Cited

Boeke, J.D., Trueheart, J., Natsoulis, G., and Fink, G.R. 1987. 5-fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods Enzymol.* 154:164-175.

Brenner, C., Bevan, A., and Fuller, R.S. 1994. One-step site-directed mutagenesis. *Methods Enzymol.* 244:163-165.

Muhlrad, D., Hunter, R., and Parker, R. 1992. A rapid method for localized mutagenesis of yeast genes. *Yeast* 8:79-92.

Contributed by Victoria Lundblad
Baylor College of Medicine
Houston, Texas

Heng Zhou
Harvard Medical School
Boston, Massachusetts

Manipulation of Cloned Yeast DNA

A major advantage of working with yeast is the ability to replace the wild-type chromosomal copy of a gene with a mutant derivative that is constructed *in vitro* using a cloned copy of the gene. This technique—unavailable in most other eukaryotes—allows the phenotype of the mutation to be studied under accurate *in vivo* conditions, with the mutation present in single copy at its normal chromosomal location.

Once a cloned gene has been obtained (see UNIT 13.8), there are a variety of techniques available for introducing a mutant derivative into its corresponding chromosomal site. All of these methods rely on homologous recombination between the transforming DNA and the yeast genomic sequences. If the gene has been cloned by complementation, several of these techniques can be used to demonstrate that the cloned gene corresponds to the gene defined by the mutation (UNIT 13.8). Gene replacement techniques also allow analysis of partial or complete gene deletions, which can be introduced into the genome to determine the null phenotype. In addition, once a cloned gene is available, previously identified mutant alleles can be recovered via the technique of plasmid gap repair (UNIT 13.9). Gene replacement techniques also allow the creation of modified genes that are regulated by a desirable (usually inducible) promoter. A specialized version of these methods permits a general approach for generating conditional alleles by a copper-inducible double shutoff.

INTEGRATIVE TRANSFORMATION

In the steps presented below (based on the procedure described by Hinnen et al., 1978), a YIp plasmid (UNIT 13.4) harboring both a selectable marker and a cloned gene of interest is integrated at the chromosomal location of the cloned gene via homologous recombination. The resulting integrant contains the entire plasmid, bracketed by intact copies of the gene (see Fig. 13.10.1). In the case of cloned genes for which no mutations have been identified (or where identified mutations have a phenotype that is difficult to score), the integrated plasmid can be used as a genetic marker—by virtue of the presence of the selectable gene—for mapping and other genetic studies. Integrative transformation of a cloned gene can also be used to demonstrate genetic linkage to the mutation used to clone the gene, providing evidence that the cloned DNA is the corresponding wild-type gene (discussed in the commentary in UNIT 13.8). It is important to note that this method introduces a selectable marker at the genomic site of the cloned gene but does not disrupt the copy of the duplicated gene. Often, multiple tandem integrations can occur, which can

BASIC PROTOCOL 1

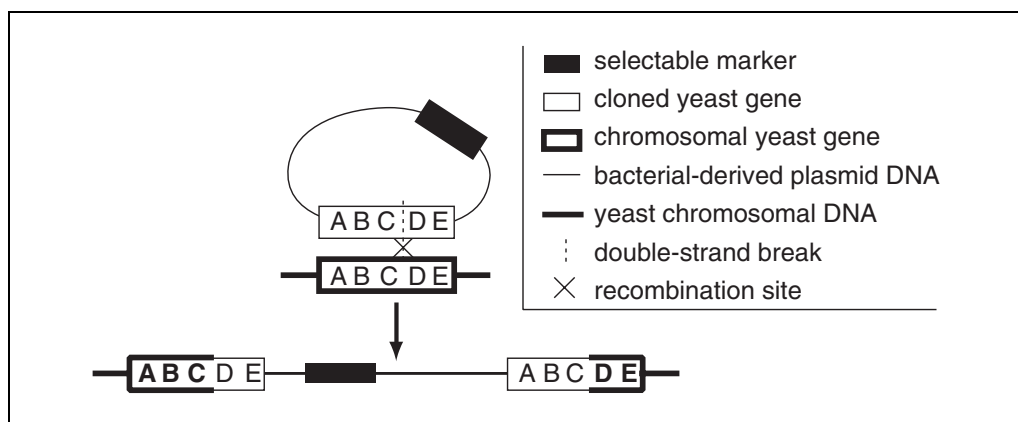


Figure 13.10.1 Integrative transformation introduces a selectable marker at the chromosomal site of a cloned gene, via integration of the entire plasmid (see accompanying text).

Saccharomyces cerevisiae

be identified by isolating genomic DNA and performing Southern blot analysis (UNITS 13.11, 2.9 & 2.10). In addition, these integrants are unstable and are lost at a frequency of ~1% per generation when grown nonselectively.

Although the frequency of integrative transformation is low, it can be increased by linearizing the plasmid at a restriction site within the cloned gene that is homologous to the intended site of integration. Linearization also directs, or targets, the recombination event to the chromosomal homologue of the cloned gene: without a double-strand break in the cloned gene, the plasmid will integrate at the chromosomal site of either the selectable gene or the cloned gene, at about equal frequencies. In Figure 13.10.1, the dashed line indicates the region in which a double-strand break should be introduced.

Materials

YIp shuttle vector (UNIT 13.4)

Appropriate restriction enzyme (UNIT 3.1)

Purified DNA from gene of interest

Carrier DNA

Selective medium (UNIT 13.1)

Additional reagents and equipment for subcloning (UNIT 3.16), yeast transformation (UNIT 13.7), yeast genomic DNA isolation (UNIT 13.11), and Southern blotting (UNITS 2.9 & 2.10) or PCR (UNIT 15.1)

1. Subclone (UNIT 3.16) the gene to be studied into a YIp vector.
2. Linearize the plasmid with a restriction enzyme that cuts within the cloned gene.

As long as no other DNA is being cotransformed, it is not necessary to inactivate the restriction enzyme prior to transformation. If a unique restriction site is not available, digestion with one or more enzymes to produce a gap in this region is also acceptable, as long as sufficient homology is present on either side of the gap (>250 bp).

3. Transform an appropriately marked strain with 1 to 10 µg DNA plus carrier DNA (UNIT 13.7), selecting for the marker present on the plasmid.
4. Purify several transformants on selective medium and isolate genomic DNA (UNIT 13.11). Confirm by Southern hybridization (UNITS 2.9 & 2.10) or PCR (UNIT 15.1) that the integration has occurred at the desired genomic site and determine whether multiple integrations have occurred.

GENE REPLACEMENT TECHNIQUES

The four methods described below provide a means of constructing a mutation in vitro in a cloned gene and reintroducing this mutation at the correct chromosomal site. This allows assessment of the genetic consequences of a mutation, and is often used to determine whether or not a gene is essential (by determining if a complete gene deletion is viable). Two of these techniques—integrative disruption and one-step gene disruption—generate either insertion or deletion mutations. The third technique—transplacement—is more generally applicable: it can be used to introduce insertion or deletion mutations containing a selectable marker, but it can also be used to introduce nonselectable mutations, such as conditional lethal mutations in an essential gene.

In each of these procedures, if the goal is to examine whether or not a null mutation is viable, a diploid strain should be transformed with the appropriate DNA construct. This allows a potentially lethal mutation to be complemented by the wild-type allele on the other chromosome. After transformation, the diploid is examined by Southern blot analysis to confirm that there is one wild-type and one disrupted copy of the gene.

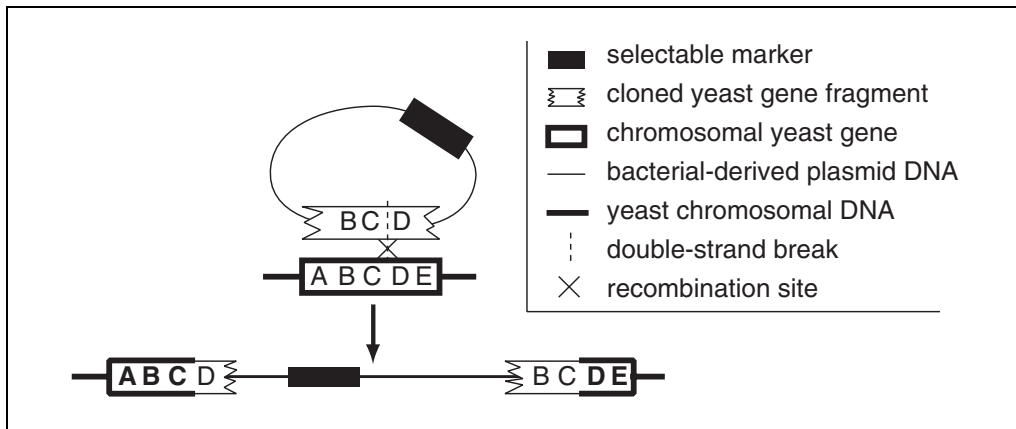


Figure 13.10.2 Integrative disruption introduces a selectable marker at the chromosomal site of a cloned gene and simultaneously generates either a 3' or 5' deletion in the two copies of the gene that bracket the plasmid (see accompanying text).

Subsequent sporulation and tetrad dissection (*UNIT 13.2*) will reveal whether the disruption is viable. Disruption of an essential gene should result in a ratio of $2^+ : 2^-$ (2 live spores:2 dead spores) in the tetrad analysis; in addition, if a selectable marker (e.g., *URA3*) is associated with the disruption, all viable spores should be auxotrophic for this marker (e.g., *Ura*⁻).

Integrative Disruption

This technique (first described by Shortle et al., 1982) generates a deletion in the chromosomal copy of a cloned gene. An internal fragment of a cloned gene is introduced into the chromosome on an integrating plasmid. As shown in Figure 13.10.2, this generates a gene duplication (which brackets the integrated plasmid), but neither copy of the gene consists of an intact copy: one copy is missing the 3' end of the gene and one copy is missing the 5' end. The procedure is the same as for integrative transformation, with two exceptions: (1) the starting YIp plasmid, instead of containing an intact gene, contains a completely internal fragment of the cloned gene, and (2) the strain to be transformed should be diploid, as explained above.

Two limitations of this technique are that a knowledge of the 5' and 3' boundaries of the coding region of the gene is required and the size of the subcloned fragment must be at least 250 bp, to promote efficient recombination. However, the disruption can be constructed in one step, without requiring an insertion or insertion/deletion mutation in the cloned gene, and a selectable phenotype (from the selectable marker on the integrated YIp plasmid) is associated with the disruption. Like integrative transformation, the integrants are unstable when grown nonselectively, and multiple integration events can occur. In addition, linearization of the plasmid increases the frequency of transformation, as well as targets the integration to the desired site.

Materials

- YIp shuttle vector (*UNIT 13.4*)
- Purified DNA from gene of interest
- Appropriate restriction enzyme (*UNIT 3.1*)
- Additional reagents and equipment for subcloning (*UNIT 3.16*) and integrative transformation (see Basic Protocol 1)

1. Subclone (*UNIT 3.16*) an internal fragment of the gene into a YIp vector.
2. Linearize the plasmid within this internal fragment by restriction enzyme digestion.

BASIC PROTOCOL 2

Saccharomyces cerevisiae

13.10.3

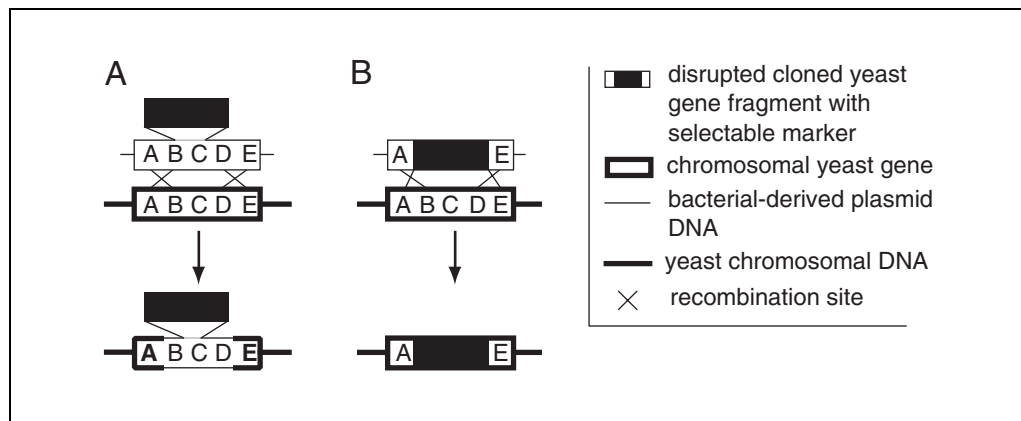


Figure 13.10.3 One-step gene disruption recombines a fragment of DNA containing a cloned gene—disrupted by the insertion of a selectable marker—into the chromosomal site of the same gene (see accompanying text). **(A)** No portion of the cloned gene fragment is deleted. **(B)** Sequences BCD are removed from the cloned gene fragment.

3. Continue with transformation and analysis of transformants as for integrative transformation (see Basic Protocol 1, steps 3 and 4).

BASIC PROTOCOL 3

One-Step Gene Disruption

Like integrative disruption, this method (from Rothstein, 1983) generates a gene disruption in one step via transformation, using a fragment of DNA containing a cloned gene that is disrupted by a selectable genetic marker. Homologous recombination between the free DNA ends, which are highly recombinogenic, and homologous sequences in the yeast genome results in replacement of the wild-type gene by the disrupted copy (see Fig. 13.10.3). The disrupted gene can contain either a simple insertion (of the selectable marker) or a deletion/insertion mutation. Introduction of these disruptions into the genome can be achieved in a single step, resulting in stable, nonreverting mutations.

Materials

- Purified DNA from gene of interest
- Purified DNA from a selectable gene
- Appropriate restriction enzyme (*UNIT 3.1*)
- Appropriate yeast strain (e.g., Table 13.10.2)

Additional reagents and equipment for subcloning (*UNIT 3.16*), gel purification of DNA (*UNIT 2.6*), yeast transformation (*UNIT 13.7*), yeast plasmid DNA isolation (*UNIT 13.11*), Southern blotting (*UNITS 2.9 & 2.10*) or PCR (*UNIT 15.1*), and tetrad analysis (*UNIT 13.2*; optional)

1. Subclone (*UNIT 3.16*) a suitable selectable gene into the gene of interest, creating in the process of subcloning a deletion as well, if desired.

Insertion mutations into the cloned gene can also be generated in vivo in E. coli by transposon mutagenesis. Several transposons have been designed with selectable markers in E. coli and yeast (Seifert et al., 1986; Huisman et al., 1987).

2. Using appropriate restriction sites, excise a linear fragment that contains the disrupted gene from the plasmid constructed in step 1 and gel-purify (*UNIT 2.6*).

Small amounts of vector sequences (≤ 200 bp) can be retained on this fragment without deleterious effects. Ideally, at least 250 bp of the cloned gene should bracket either side of the inserted selectable gene, to promote recombination at the chromosomal locus of the cloned gene, rather than at the site of the selectable marker.

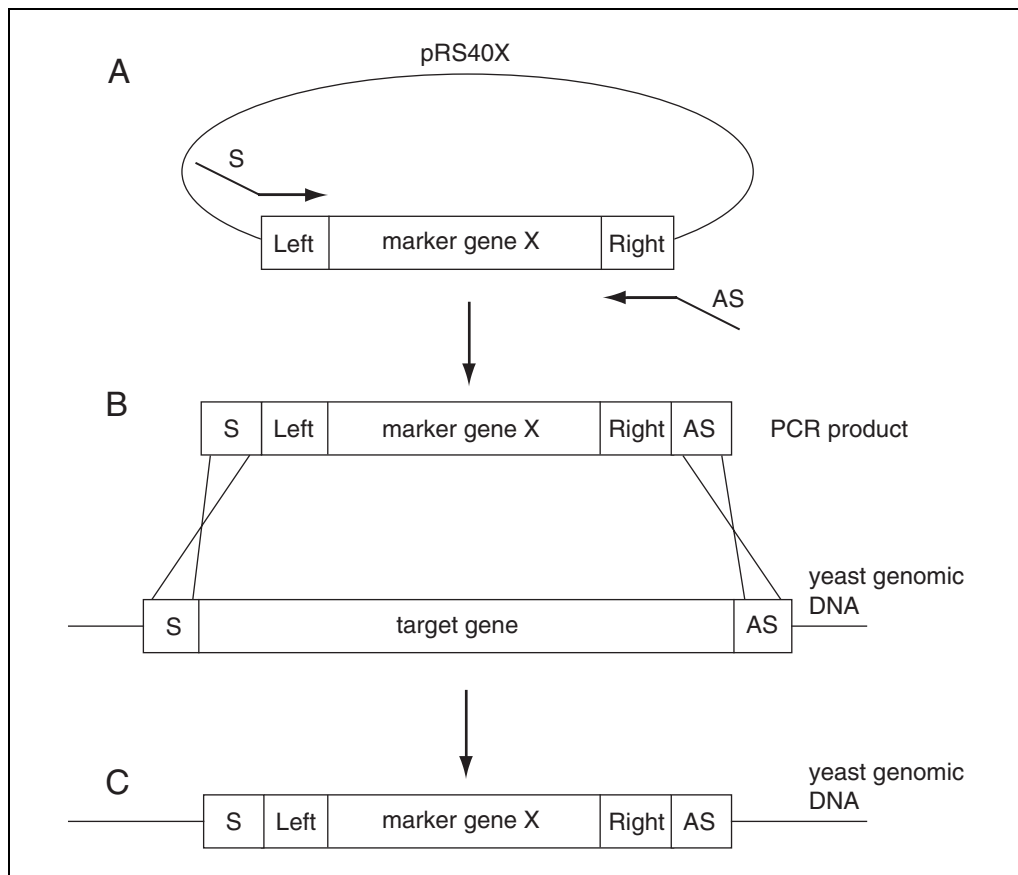


Figure 13.10.4 PCR mediated one-step gene disruption. **(A)** The one-step gene-disruption cassette is generated by PCR amplification of marker gene X using primers whose 3' ends are derived from the sequences Left and Right found in all pRS40X series vectors. The 5' 40 nucleotides of these primers are derived from sequences S (sense strand) and AS (anti-sense strand), which flank the open reading frame of the target gene. **(B)** When transformed into yeast, the gene-disruption cassette can replace the target gene by recombination between the homologous sequences S and AS. **(C)** Recombinants can be identified by selection for expression of gene X and confirmed by subsequent PCR or Southern blot analysis of the target gene locus in the transformants.

3. Transform yeast with 1 to 10 μg of the gel-purified fragment (*UNIT 13.7*), selecting for the inserted marker.
4. Confirm the structure of the disruption isolation of genomic DNA (*UNIT 13.11*) followed by Southern hybridization (*UNITS 2.9 & 2.10*) or PCR (*UNIT 15.1*). If a diploid was transformed, sporulate and dissect (*UNIT 13.2*) to obtain haploid spore products with the disruption (or to observe inviability).

PCR-Mediated One-Step Gene Disruption

One-step gene disruption (first described by Baudin, 1993) is a simple and powerful technique for removal of specific DNA sequences from the yeast genome. A gene disruption cassette is constructed that consists of a selectable marker flanked by sequences derived from the 5' and 3' ends of the target gene to be detected. When this linear DNA fragment is transformed into yeast, homologous recombination between the identical sequences flanking the marker and target genes results in replacement of the target gene by the marker gene. In the past, this method suffered from the drawback that a plasmid carrying the gene disruption cassette had to be constructed before it could be carried out. Often, suitable restriction sites near the 5' and 3' ends of the ORF of the target gene were not available, making complete deletion of the ORF difficult or impractical. This draw-

**ALTERNATE
PROTOCOL 1**

*Saccharomyces
cerevisiae*

13.10.5

Table 13.10.1 Yeast Markers and Plasmids for PCR-Mediated One-Step Gene Disruption^a

Gene	Null allele	Template plasmid
<i>URA3</i>	<i>ura3Δ0</i>	pRS406
<i>HIS3</i>	<i>his3Δ200</i>	pRS403
<i>LYS2</i>	<i>lys2Δ0</i>	pRS317
<i>LEU2</i>	<i>leu2Δ0</i>	pRS405
<i>TRP1</i>	<i>trp1Δ63</i>	pRS404
<i>MET15</i>	<i>met15Δ0</i>	pRS401

^aAuxotrophic markers and plasmids above are described in Brachman et al. (1997) and Sikorski and Hieter (1989).

Table 13.10.2 Suggested Yeast Strains for Gene Replacement^a

Strain	ATCC number	Genotype
BY4700	200866	<i>MATa ura3Δ</i>
BY4736	200898	<i>MATa ade2Δ0::hisG his3Δ200 met15Δ0 trp1Δ63 ura3Δ0</i>
BY4727	200889	<i>MATα his3Δ200 leu2Δ0 lys2Δ0 met15Δ0 trp1Δ63 ura3Δ0</i>

^aMany other suitable strains are available from the ATCC (strains 200866 to 200902). A recipient diploid can be made by mating BY4736 and BY4727 and selecting for Ade⁺ Leu⁺ Lys⁺ colonies.

back can be overcome by using PCR to create the gene disruption cassette (see Fig. 13.10.4). Because only very small regions of homology between DNA sequences are required to direct homologous recombination in yeast, PCR primers that will amplify the desired auxotrophic marker flanked by 30 to 40 nucleotides of homology to the target gene can be used to create the gene disruption cassette. The efficiency of this technique depends upon two factors. First, the homology of the PCR product to the target gene should be as long as possible, typically 40 nucleotides. Second, this technique works best if the sequences encoding the auxotrophic marker gene have been completely deleted from the yeast genome, thus minimizing the chances for homologous recombination elsewhere in the genome. A set of convenient auxotrophic markers and plasmids for this technique are shown in Table 13.10.1. Yeast strains carrying different combinations of these markers (see Brachman et al., 1997) are available from the ATCC (strains 200866 to 200902); some suggested strains are described in Table 13.10.2. Using the pRS40X series of vectors as the template for PCR, one set of PCR primers is sufficient to make a gene knockout marked by any one of the six different auxotrophic markers listed in Table 13.10.1.

Materials

- 10× PCR buffer (e.g., as supplied with *Taq* DNA polymerase from Boehringer Mannheim)
- 100 mM Tris·Cl/15 mM MgCl₂/500 mM KCl, pH 8.3 (20°C)
- 4dNTP mixture (UNIT 3.4)
- pRS40X-series vector (e.g., Table 13.10.1)
- pRS Left and Right primers (see step 1)
- Taq* DNA polymerase
- Appropriate yeast strain (e.g., Table 13.10.2)
- Selective medium (UNIT 13.1)

Additional reagents and equipment for PCR (UNIT 15.1), yeast transformation (UNIT 13.7), yeast genomic DNA isolation (UNIT 13.11) and Southern blotting (UNITS 2.9 & 2.10; optional)

1. Design PCR primers for the disruption cassette as follows (see Fig. 13.10.4). Pick 40 nucleotides of sequence from the sense strand at the 5' end of the target gene ORF and add to the 5' end of one of the two sequences given below. Similarly, pick 40 nucleotides of sequence from the anti-sense strand at the 3' end of the target-gene ORF to be deleted and add to the 5' end of the other sequence listed below (found in all pRS-series vectors). Obtain primer from oligonucleotide supplier.

Left primer: 5'-CTGTGCGGTATTTACACCG-3'

Right primer: 5'-AGATTGTACTGAGAGTGAC-3'

This results in two 60-nucleotide primers sufficient to amplify any of the auxotrophic markers in the pRS-series vectors. The overall primer length is set at 60 nucleotides because this is the longest primer length that is convenient and economical to obtain commercially, and because primers of this length typically give satisfactory results.

2. Set up several 100- μ l PCR reactions containing:

10 μ l 10 \times PCR buffer
200 μ M 4dNTP mixture
100 pmol of each Left and Right primer
10 ng pRS40X
2.5 U *Taq* DNA polymerase.

pRS40X series vectors are used as the template because they lack sequences that allow their stable maintenance in yeast, thus obviating need for purification of the PCR product away from the template DNA. Because the number of transformants obtained from a yeast transformation is proportional to the amount of input DNA, it is worthwhile to perform several 100- μ l reactions, combine and concentrate them by ethanol precipitation (UNIT 2.1), and then transform all of the DNA together.

3. Perform PCR amplification (UNIT 15.1) using the following cycling conditions:

1 cycle:	2 min	94°C	(denaturation)
30 cycles:	30 sec	94°C	(denaturation)
	1 min	55°C	(annealing)
	2 min	72°C	(extension)
1 cycle:	10 min	72°C	(extension).

4. Transform (UNIT 13.7) the PCR product into an appropriate yeast strain and select for expression of the auxotrophic marker.

Because the transformation efficiency of these PCR products is low, it is critical that a high-efficiency transformation procedure—such as lithium acetate transformation with early log-phase cell (see UNIT 13.7)—be used. Also, if the gene being deleted is essential for life, the deletion mutation must be created in a diploid strain.

5. Isolate genomic DNA (UNIT 13.11) and identify strains carrying the gene disruption by PCR (UNIT 15.1) with primers that flank the expected mutation, or by Southern blot analysis (UNITS 2.9 & 2.10).

As noted above, the efficiency of this procedure is dependent upon the complete absence of homology to the auxotrophic marker in the recipient strain. In cases where substantial homology to the auxotrophic marker remains, the frequency of transformants with the expected deletion may drop to as low as a few percent. In such cases, the deletions may still be found by screening many clones by PCR or, if possible, by first identifying potential deletion strains among the transformants by some expected mutant phenotype.

Saccharomyces cerevisiae

13.10.7

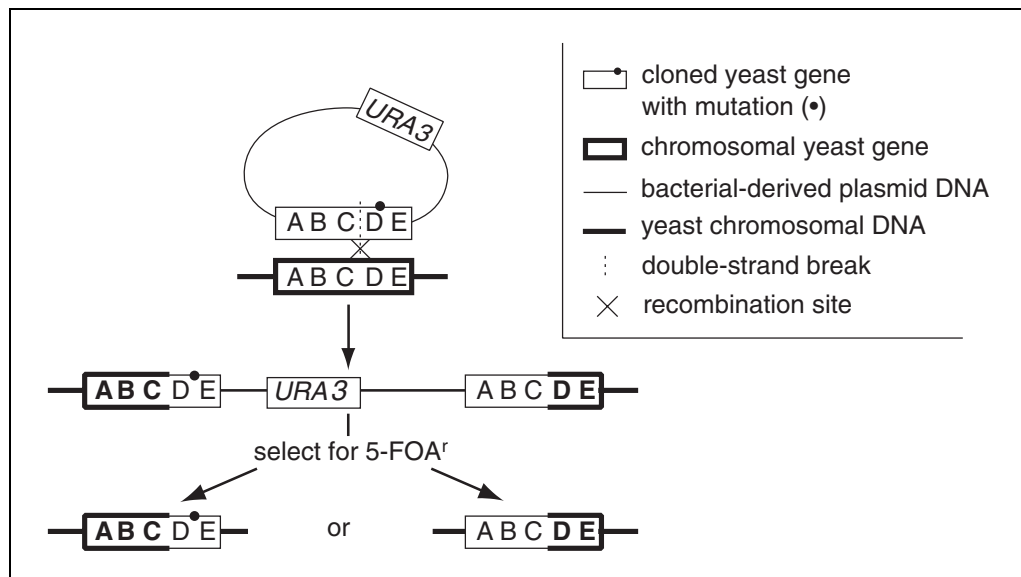


Figure 13.10.5 Transplacement allows the introduction of any mutation in a cloned gene that does not have a selectable phenotype (see accompanying text).

**BASIC
PROTOCOL 4**

Transplacement

This method (based on Scherer and Davis, 1979), also called allele replacement, provides a general means of introducing any type of mutation constructed *in vitro* that does not have a selectable phenotype into its corresponding chromosomal location. It can be used either to introduce a single defined mutation or to screen a mutagenized collection of plasmids. The mutated gene is introduced into yeast on a YIp plasmid, which usually contains the *URA3* gene as a selectable marker. After transformation, selection for Ura⁺ results in an integration event, such that mutant and wild-type alleles bracket the *URA3*-containing plasmid sequences (Fig. 13.10.5). Subsequent eviction of the plasmid (containing one copy of the gene) is monitored by screening for colonies that are URA⁻ and therefore resistant to the drug 5-fluoroorotic acid (5-FOA). The 5-FOA^r colonies are then screened for the desired mutant phenotype; the proportion of 5-FOA^r colonies containing the mutant allele will depend on the position of the mutation relative to the length of the flanking homologous DNA.

Materials

- YIp shuttle vector (UNIT 13.4)
- Purified DNA from gene of interest
- Nonselective liquid medium (YPD or uracil-containing minimal medium; UNIT 13.1)
- 5-FOA plates (UNIT 13.1)
- Additional reagents and equipment for subcloning (UNIT 3.16), integrative transformation (see Basic Protocol 1), yeast genomic DNA isolation (UNIT 13.11), and Southern blotting (UNITS 2.9 & 2.10) or PCR (UNITS 15.1.)

1. Subclone (UNIT 3.16) the gene of interest into YIp5.
2. Mutagenize the plasmid as desired, either via the techniques presented in Chapter 8, or by introducing defined deletion or insertion mutations.
3. Proceed with linearization, transformation, and analysis of transformants as for integrative transformation (see Basic Protocol 1, steps 2 to 4).
4. Grow transformants overnight in liquid medium without selection. Plate for single colonies on nonselective medium (~100 single colonies) and grow 2 days at 30°C.

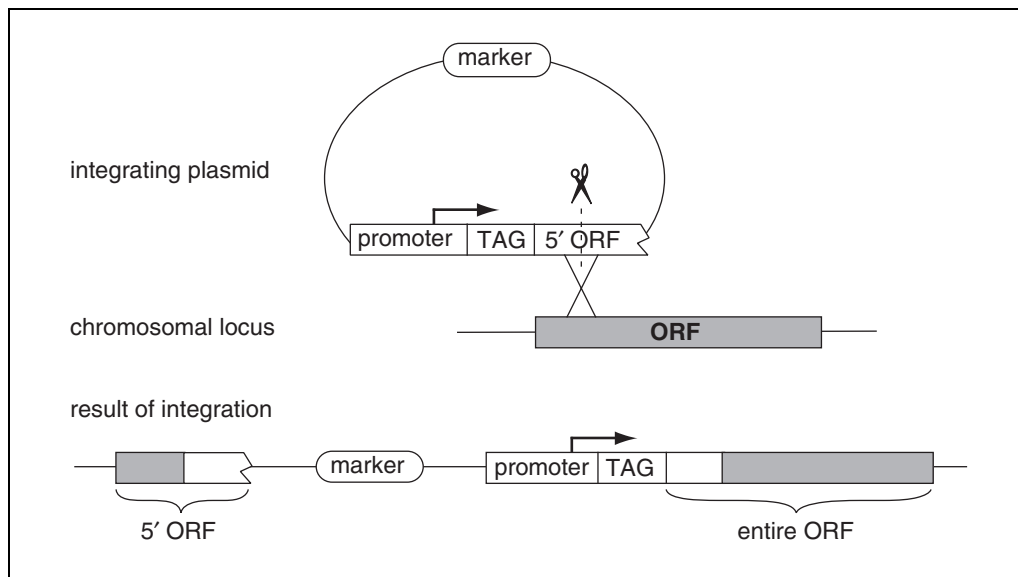


Figure 13.10.6 A single integrative transformation step introduces the desired allele at the normal genome locus and truncates the wild-type allele.

Alternatively, cells can be plated or streaked directly on 5-FOA plates, in which case proceed to step 6.

If the goal is to identify conditional lethal mutants, a large number of transformants should be examined (since some proportion of transformants will have received unmutagenized plasmid and since conditional lethal mutations can be rare events). In addition, this nonselective period of growth should be at the permissive temperature (usually 23°C).

5. Replica plate onto 5-FOA plates and grow overnight to identify Ura⁻ segregants.

Colonies that have evicted the plasmid during the period of growth in liquid medium will give completely resistant replicas on the 5-FOA plates; these colonies should be recovered from the nonselective plate (rather than from the 5-FOA plate) and used in step 6, below. Other colonies that have evicted the plasmid during the period of growth of the colony will appear “patchy” on the 5-FOA replica (often referred to as papillation).

6. Screen 5-FOA^r colonies for the presence of a mutant phenotype. If the mutation is expected to alter the restriction pattern, confirm the genotype of mutant colonies in comparison to wild-type by isolation of genomic DNA (UNIT 13.11) followed by Southern hybridization (UNITS 2.9 & 2.10) or PCR (UNIT 15.1).

CREATING MODIFIED GENES BY ONE-STEP INTEGRATIVE REPLACEMENT

This method may be used to create a yeast strain in which a gene of interest is controlled by any desired (usually inducible) promoter. The same method may be used to create an N-terminal fusion to any chosen epitope tag or other cassette. A single integration event introduces the new allele of the gene at its normal locus and simultaneously truncates the normal endogenous copy of the gene, rendering it nonfunctional (Fig. 13.10.6).

Materials

- Purified DNA from gene of interest
- Yeast integrating shuttle vector with selectable marker
- Appropriate restriction enzyme (UNIT 3.1)
- Appropriate yeast strain
- Selective medium (UNIT 13.1)

BASIC PROTOCOL 5

Saccharomyces cerevisiae

13.10.9

Additional reagents and equipment for subcloning (UNIT 3.16), yeast transformation (UNIT 13.7), yeast genomic DNA isolation (UNIT 13.11), and Southern blotting (UNITS 2.9 & 2.10) or PCR (UNIT 15.1)

1. Select a short 5' fragment of the open reading frame (ORF) of the gene of interest. This fragment should:

- a. begin at the first codon of the gene;
- b. be too short to encode a functional product; in addition, its short product should not have the potential to interfere with the normal activity of the full-length gene product;
- c. be long enough (preferably ≥ 300 bp) to be efficiently targeted by homology to the corresponding chromosomal region; and
- d. contain a restriction site (to be used for targeting the final integrating plasmid to the correct locus) that will remain unique in the final construct and that leaves reasonable stretches of homology to the target gene on both sides of the double strand break.

In order to minimize the chances of undesired integration at the promoter locus, it is preferable for the restriction site to be located slightly closer to the 3' end of the ORF fragment than to its 5' end.

2. Subclone (UNIT 3.16) this 5' fragment downstream of the desired promoter in an integrating yeast shuttle plasmid. If desired, this construct may also include an epitope tag (or other fusion cassette) appended in frame to the N terminus of the ORF fragment.

This subcloning is usually most easily accomplished by PCR. Be certain to situate the initial ATG in the appropriate promoter context. Occasionally, an N-terminal epitope tag may interfere with proper function of a gene. In these situations it is advisable to try subcloning the ORF fragment downstream of the promoter with no N-terminal fusion. Alternatively, a modification of this procedure (using a short, nonfunctional 3' piece of the ORF fused upstream of the desired tag) may be used to create a C-terminal fusion to an epitope tag.

3. Linearize the construct at a unique restriction site within the ORF fragment.

If no restriction site in the ORF fragment is unique, a partial digest may be used to create a mixed population of linearized molecules for transformation. A correspondingly larger number of transformants should then be screened in step 6 for correctly targeted integration. On the other hand, if no restriction sites are present in the ORF fragment, it may be necessary to introduce one by silent PCR mutagenesis (UNIT 8.5).

4. Transform the linearized molecule into an appropriate yeast strain (UNIT 13.7).
5. Plate the transformations onto medium selecting for the plasmid marker (UNIT 13.2).

Remember that if placing an essential gene under the control of an inducible promoter, it is important to plate the cells onto inducing medium.

6. Isolate genomic DNA (UNIT 13.11) and identify clones containing the modified genes by Southern blotting (UNITS 2.9 & 2.10), genomic PCR (UNIT 15.1), or phenotypic assessment.

When working with strains created by this method, be sure to maintain selection for the plasmid marker in order to prevent looping out of the integrated plasmid.

CREATING MODIFIED GENES BY TRANSPLACEMENT

For some genes, it may be difficult or impossible to select a 5' ORF fragment that is long enough for efficient integration yet too short to retain function. In these situations, inducible alleles may be introduced by complete allele replacement (see Basic Protocol 4). Such replacements are stable, leave no short 5' ORF fragment upstream, and avoid the need to maintain selection for the plasmid marker. The allele replacement may be accomplished as follows:

1. Create a molecule as for one-step integrative replacement (see Basic Protocol 5, steps 1 and 2), being sure to use an integrating vector with *URA3*.
2. Just upstream of the promoter, subclone (*UNIT 3.16*) a fragment (usually ≥ 500 bp) of the 5' flanking region of the target gene.

When selecting this fragment, take care to preserve the uniqueness of the targeting restriction site in the ORF fragment.

3. Linearize the molecule and integrate into yeast (see Basic Protocol 5, steps 3 to 5).
4. Streak the *URA*⁺ transformants onto 5-FOA plates (see Basic Protocol 4, step 5).
5. Screen the 5-FOA^r clones obtained for the correctly replaced allele (see Basic Protocol 4, step 6).

CREATION OF CONDITIONAL ALLELES BY COPPER-INDUCIBLE DOUBLE-SHUTOFF PROCEDURE

This method (first described in Moqtaderi et al., 1996), which is useful in the functional analysis of essential genes, allows the creation of conditional expression strains in which the addition of copper causes expression of a gene of interest to be shut off at both the RNA and the protein levels. The RNA shutoff is accomplished by harnessing a transcriptional repressor, and the protein degradation is achieved by exploitation of the N-end rule, whereby a protein's rate of turnover is dictated by its N-terminal amino acid (Varshavsky, 1992). Traditional depletion systems generally rely on the removal of positive agents or on a complete change in medium (e.g., galactose to glucose shift), requiring the cells to be washed at the time of the shift to nonpermissive conditions. In contrast, this method works indirectly, by the induction of negative agents (the simple addition of copper to stimulate the production of a transcriptional repressor and a protein degradation factor), which obviates the need to wash the cells.

Specifically, the parent strain used must harbor copper-inducible alleles of *ROX1*, a transcriptional repressor, and of *UBR1*, which encodes the N-end recognition component of the ubiquitin protein degradation pathway. In this background, the gene of interest is replaced with an allele fused to an N-end recognition sequence for rapid degradation and driven by the *ANB1* promoter, a natural target of Rox1. In the absence of copper, neither Rox1 nor Ubr1 is produced, and the target gene is expressed. The addition of copper results in rapid activation of both *ROX1* and *UBR1*, leading to the simultaneous repression of the target gene's transcription and rapid degradation of its existing protein product (Figs. 13.10.7 and 13.10.8).

Materials

Plasmid ZM168, containing the *ANB1* promoter driving a ubiquitin-arginine-*lacI*-HA (URLF) cassette immediately followed by a polylinker

Yeast integrating shuttle vector with selectable marker

Yeast strain ZMY60 (*MATa*, *ade2-101*, *HIS*⁺, *LEU*⁺, *trp1Δ1*, *ura3-52*)

**ALTERNATE
PROTOCOL 2**

**BASIC
PROTOCOL 6**

*Saccharomyces
cerevisiae*

13.10.11

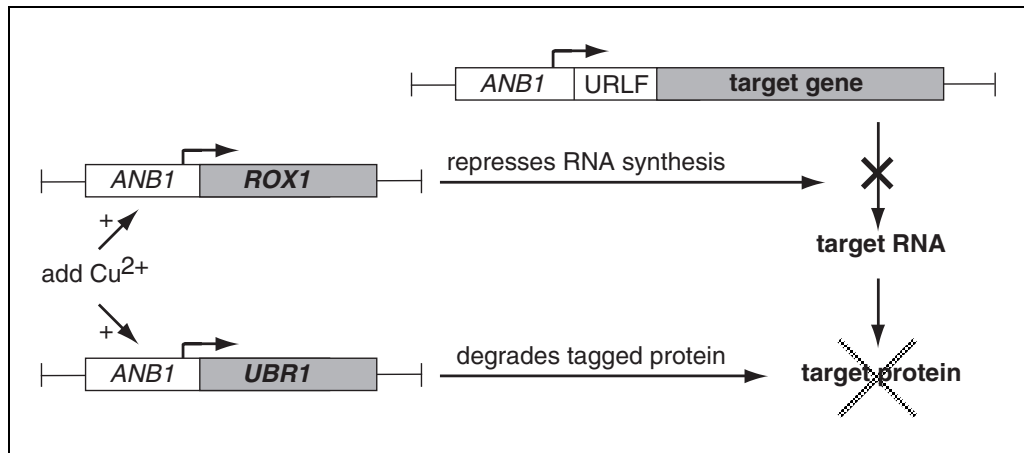


Figure 13.10.7 A two-pronged conditional knock-out strategy shuts off the target gene's expression at both the RNA and the protein levels.

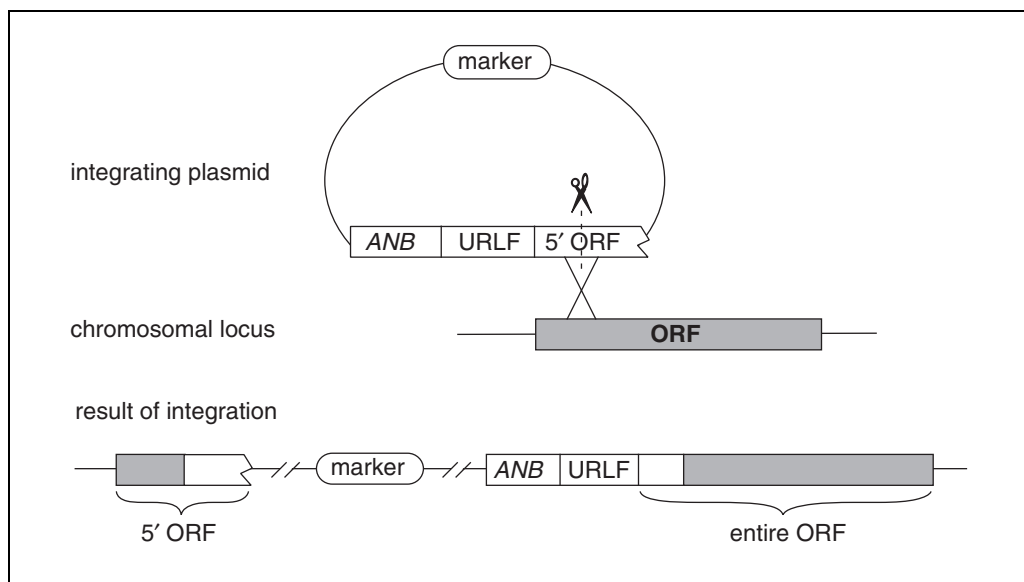


Figure 13.10.8 Introduction of URLF-tagged, *ANB1*-driven allele of the desired target gene into the genome.

Centromeric plasmid bearing the same marker as the integrating vector
 Synthetic medium plates selecting for the integrated plasmid marker (*UNIT 13.1*),
 with and without 500 μ M cupric sulfate (CuSO_4)

1. Following the guidelines provided for one-step replacement (see Basic Protocol 5, steps 1 and 2), select a short, nonfunctional 5' fragment of the intended target gene, and subclone it (*UNIT 3.16*) into the polylinker of the plasmid ZM168 in frame with the URLF cassette.
2. Subclone the entire *ANB1*-URLF-ORF fragment into a yeast integrating shuttle vector bearing the preferred selectable marker.

Take care to maintain the uniqueness of the targeting restriction site in the ORF.

3. Linearize the final construct at the unique restriction site within the ORF fragment and transform it (*UNIT 13.7*) into yeast strain ZMY60. In parallel, transform ZMY60

with a centromeric plasmid bearing the same yeast marker as the final integrating construct. Plate onto synthetic medium selecting for the plasmid marker.

A different strain background may be desired. If this is the case, it is possible to generate another strain containing stably integrated, copper-inducible alleles of ROX1 and UBR1 by using the two-step gene replacement constructs ZM195 and ZM197. (See Basic Protocol 4 for details of how to construct stable gene replacements.)

4. Streak the transformants onto (a) synthetic medium selecting for the integrated plasmid marker with no added copper and (b) synthetic medium selecting for the integrated plasmid marker and containing 500 μM CuSO_4 . Grow plates 2 days at 30°C.

When preparing solid medium for this purpose, it is generally advisable to add all necessary amino acids and the copper when pouring the plates. Attempts to spread liquid amino acid supplements or copper on top of a previously poured plate may result in an uneven copper concentration, making strain growth patterns more difficult to interpret.

5. Analyze transformant growth to assess integration. If the target gene is essential, correct integrants should fail to grow on copper-containing medium. Transformants carrying the centromeric plasmid should grow normally on both media.

If the depletion is inefficient, the copper concentration may be raised to 1 mM, but beyond this level cell growth may begin to be impaired due to copper toxicity. If the target gene is not essential, test for an expected phenotype in the presence of copper or analyze the transformants by isolating genomic DNA (UNIT 13.11) and Southern blotting (UNITS 2.9 & 2.10).

6. For time-course experiments, grow both the conditional strain and the centromeric plasmid control strain in liquid synthetic medium without copper but selective for the plasmid marker. Add 500 μM CuSO_4 to initiate depletion and harvest cells at varying time-points for analysis.

A standard procedure is to assay the cellular process under investigation at time-points over the next several hours after the addition of copper (usually up to 8 hr) to test the effects of increasing levels of depletion. The centromeric plasmid control strain is included to represent a strain expressing a wild-type level of the gene product of interest. The conditional strain, even under permissive conditions, harbors an allele of the target gene driven by a heterologous promoter, and it therefore may not yield a truly wild-type level of expression. It is important to maintain selection for the marker at all times to avoid loss of the integrated plasmid. Even after the addition of copper, it is normal for the cells to continue to grow for several hours. The need to maintain the cells in exponential growth throughout most experiments thus makes it inadvisable to add copper at too high a cell density; a typical OD_{600} at the time of shift is ~ 0.25 .

LITERATURE CITED

- Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F., and Cullin, C. 1993. A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucl. Acids Res.* 21:3329-3330.
- Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P., and Boeke, J. 1998. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* 14:115-132.
- Hinnen, A., Hicks, J.B., and Fink, G.R. 1978. Transformation of yeast. *Proc. Natl. Acad. Sci. U.S.A.* 75:1929-1933.
- Huisman, O., Raymond, W., Foelich, K., Errada, P., Kleckner, N., Botstein, D., and Hoyt, M.A. 1987. A *Tn10-lacZ-kan^R-URA3* gene fusion transposon for insertion mutagenesis and fusion analysis of yeast and bacterial genes. *Genetics* 1116:191-199.
- Moqtaderi, Z., Bai, Y., Poon, D., Weil, P.A., and Struhl, K. 1996. TBP-associated factors are not generally required for transcriptional activation in yeast. *Nature* 383:188-191.
- Rothstein, R.J. 1983. One-step gene disruption in yeast. *Methods Enzymol.* 101:202-210.

- Scherer, S. and Davis, R.W. 1979. Replacement of chromosome segments with altered DNA sequences constructed in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 76:4951-4955.
- Seifert, H.S., Chen, E.Y., So, M., and Heffron, F. 1986. Shuttle mutagenesis: A method of transposon mutagenesis for *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 83:735-739.
- Shortle, D., Haber, J.E., and Botstein, D. 1982. Lethal disruption of the yeast actin gene by integrative DNA transformation. *Science* 217:371-373.
- Sikorski, R.S. and Hieter, P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces Cerevisiae*. *Genetics* 122:19-27.
- Varshavsky, A. 1992. The N-end rule. *Cell* 69:725-735.

Contributed by Victoria Lundblad
Baylor College of Medicine
Houston, Texas

Grant Hartzog (PCR-mediated one-step gene disruption)
Harvard Medical School
Boston, Massachusetts

Zarmik Moqtaderi (creation of modified genes and conditional alleles)
Harvard Medical School
Boston, Massachusetts

PREPARATION OF YEAST DNA, RNA, AND PROTEINS

Preparation of Yeast DNA

Molecular studies in yeast often require the isolation of both plasmid and chromosomal yeast DNA. Plasmid DNA is used in the transformation of *E. coli*, whereas chromosomal DNA is used for Southern hybridization analysis, in vitro amplification by the polymerase chain reaction (PCR), or cloning of integrated plasmids. This unit presents two variations of the “smash and grab” protocol (Hoffman and Winston, 1987) that produce suitable DNA for all these applications. These protocols work for both *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, although for *S. pombe* a suitable liquid medium must be used to grow the cells.

RAPID ISOLATION OF PLASMID DNA FROM YEAST

Plasmid DNA is released from a yeast transformant along with chromosomal DNA in a rapid, 10-min protocol by vortexing with glass beads in the presence of detergents, phenol, chloroform, and isoamyl alcohol. After centrifugation to pellet cell debris, no further purification of the DNA is required for transformation of competent *E. coli* cells.

Materials

- YPD or appropriate selective medium (UNIT 13.1)
- Yeast colony containing the plasmid of interest
- Breaking buffer (see recipe)
- 25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol (with buffered phenol; UNIT 2.1)
- 0.45- to 0.52-mm acid-washed glass beads (UNIT 13.12; Thomas Scientific)
- Competent *E. coli* cells HB101 or MH1 (UNIT 1.8)
- LB plates (UNIT 1.1) containing appropriate antibiotic (UNIT 1.4)
- 13 × 100-mm glass tubes, sterile
- 30°C incubator with shaker or roller drum
- Additional reagents and equipment for growth of *E. coli* on solid medium (UNIT 1.3) and transformation of *E. coli* (UNIT 1.8)

1. Inoculate 2 ml medium in a 13 × 100-mm sterile glass tube with a single yeast colony containing the plasmid of interest. Grow overnight to stationary phase at 30°C in either a roller drum or a shaking incubator.

Transformants carrying stable plasmids (UNIT 13.4) should be grown in YPD to get a dense culture. Transformants carrying unstable plasmids should be grown in selective medium to increase the percentage of plasmid-containing cells.

2. Transfer 1.5 ml of the overnight culture to a microcentrifuge tube and spin 5 sec at high speed, room temperature. Pour off supernatant and disrupt pellet by vortexing briefly.
3. Resuspend cells in 200 µl breaking buffer. Add 0.3 g glass beads (~200 µl volume) and 200 µl phenol/chloroform/isoamyl alcohol. Vortex 2 min at highest speed.

The amount of vortexing required can vary depending upon the vortex used. Determine by microscopic examination the minimum vortexing required for a particular machine to break 80% to 90% of the cells.

For simultaneous preparation of DNA from several yeast strains, a multi-tube vortexer or a multi-tube head for a standard vortexer can be used (VWR Scientific) and vortexing time should be increased to 3 min.

4. Microcentrifuge 5 min at high speed, room temperature.
5. Transform competent *E. coli* HB101 or MH1 with 1 to 2 μ l of the aqueous layer. Plate on LB plates containing the appropriate antibiotic to select for the drug-resistance marker on the plasmid.

See Critical Parameters and Troubleshooting for guidelines on transformation.

Save 50 μ l of the aqueous layer and store at -20°C in case additional transformations are required.

ALTERNATE PROTOCOL

RAPID ISOLATION OF YEAST CHROMOSOMAL DNA

The DNA preparation described in the Basic Protocol can be easily scaled up to prepare chromosomal DNA for use in Southern hybridization analysis (UNIT 2.9A), in vitro DNA amplification by PCR (UNIT 15.1), or restriction digestion and ligation (UNITS 3.1 & 3.16) to clone integrated plasmids. This procedure is significantly faster than other protocols used to isolate high-molecular-weight DNA. Although the DNA is subject to shearing in this protocol, the resulting DNA is of sufficiently high molecular weight to enable detection of 19-kb restriction digestion products by Southern hybridization analysis.

Additional Materials (also see Basic Protocol)

- TE buffer (APPENDIX 2)
- 1 mg/ml DNase-free RNase A (UNIT 5.5)
- 4 M ammonium acetate solution (APPENDIX 2)
- 100% ethanol
- 18 \times 150-mm glass culture tubes *or* 17 \times 100-mm disposable polypropylene tubes, sterile
- Tabletop centrifuge

Prepare yeast cells

1. Grow a 10-ml culture of yeast in YPD overnight to stationary phase in either 18 \times 150-mm sterile glass culture tubes or in 17 \times 100-mm sterile, polypropylene tubes (see Basic Protocol, step 1).
2. Spin culture 5 min in a tabletop centrifuge at $1200 \times g$ (3000 rpm), room temperature. Aspirate or pour off supernatant, and resuspend cells in 0.5 ml water.
3. Transfer the resuspended cells to a microcentrifuge tube and spin 5 sec at room temperature. Pour off supernatant and disrupt pellet by vortexing briefly.

This wash step in water removes any remaining medium.

Break open the cells

4. Resuspend cells in 200 μ l breaking buffer. Add 0.3 g glass beads (\sim 200 μ l volume) and 200 μ l phenol/chloroform/isoamyl alcohol and vortex at highest speed for 3 min.

The amount of vortexing required can vary depending upon the vortex used. Determine by microscopic examination the minimum vortexing required for a particular machine to break 80% to 90% of the cells.

If using a multi-tube vortex rather than a single-tube vortex, the time of vortexing may need to be increased up to 4 or 5 min to get more efficient breakage of cells. However, longer vortexing can result in shearing of DNA molecules.

5. Add 200 μ l TE buffer and vortex briefly.
6. Microcentrifuge 5 min at high speed, room temperature, and transfer aqueous layer to a clean microcentrifuge tube. Add 1 ml of 100% ethanol and mix by inversion.

7. Microcentrifuge 3 min at high speed, room temperature. Remove supernatant and resuspend pellet in 0.4 ml TE buffer.

Degrade RNA contaminants and recover DNA

8. Add 30 μ l of 1 mg/ml DNase-free RNase A, mix, and incubate 5 min at 37°C.
9. Add 10 μ l of 4 M ammonium acetate and 1 ml of 100% ethanol. Mix by inversion.
10. Microcentrifuge 3 min at high speed, room temperature. Discard supernatant and dry pellet. Resuspend DNA in 100 μ l TE buffer.

Yields of ~20 μ g of chromosomal DNA should be obtained. This DNA is ready to use for restriction digestion (UNIT 3.1), in vitro PCR amplification (UNIT 15.1), or Southern blot analysis (UNIT 2.9). For Southern blots, best results are obtained when 5 μ l DNA (~1 μ g) is digested in a total volume of 20 μ l. To amplify by PCR, 2 μ l of DNA should be used in a 50- μ l reaction.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Breaking buffer

- 2% (v/v) Triton X-100
 - 1% (v/v) sodium dodecyl sulfate (SDS)
 - 100 mM NaCl
 - 10 mM Tris-Cl, pH 8.0
 - 1 mM EDTA, pH 8.0
- Store \leq 1 year at room temperature

COMMENTARY

Background Information

These DNA isolation protocols simultaneously break yeast cells by vortexing with glass beads in a detergent solution and separate nucleic acids from protein by phenol/chloroform extraction. They differ in the amount of additional manipulation required to obtain DNA of the quality needed for a particular procedure. Other protocols also describe the isolation of high-molecular-weight DNA (Cryer et al., 1975; Winston et al., 1983; Davis et al., 1980). However, even the protocols designed to be rapid DNA preparations require >3 hr time. In general, the protocols described here produce DNA of sufficient molecular weight for any desired application, except for the production of insert DNA for the construction of genomic libraries.

Many situations require the isolation of plasmid DNA from a yeast transformant to transform *E. coli*. Examples include cloning of (1) the wild-type copy of a gene by complementation of a recessive mutant (UNIT 13.8), (2) a dominant mutant allele of a gene by transformation of a wild-type strain, or (3) a mutant allele of a gene by plasmid gap repair (UNIT 13.10). Plasmid DNA is also needed to identify a de-

sired mutant allele of a gene by the plasmid shuffle technique (UNIT 13.10). Most yeast plasmids used in molecular biology are shuttle vectors carrying a selectable marker and an origin of replication (*ori*) for transformation and maintenance in *E. coli* (UNIT 13.4). Because significantly more plasmid DNA can be isolated from an *E. coli* transformant than from a yeast transformant, *E. coli* is used to amplify the DNA for restriction mapping and other in vitro manipulations of the DNA as well as for retransformation of yeast strains.

Other situations require the isolation of chromosomal DNA. The most common use of this DNA is for Southern hybridization analyses. Such analyses are needed to confirm that integration of a YIp plasmid or of a marked gene disruption has occurred by homologous recombination. Chromosomal DNA can also be used as a template for in vitro amplification by PCR in place of Southern analyses to determine whether DNA is integrated by homologous recombination. PCR can also be used in place of gap repair to clone mutant alleles of a given gene. Finally, chromosomal DNA is needed to clone DNA adjacent to an integrated plasmid. This is done by digesting the chromosomal

Saccharomyces cerevisiae

13.11.3

DNA with a restriction enzyme that creates a restriction fragment containing the plasmid origin of replication and selectable marker, as well as the flanking DNA of interest (Southern analyses are required to determine the appropriate restriction enzyme for this construction). Ligation of this yeast DNA, followed by transformation of *E. coli*, results in the isolation of the cloned flanking DNA, as only the appropriate piece of DNA will confer drug resistance upon the *E. coli* strain.

These protocols do not produce intact chromosomal DNA. Yeast chromosomes can be isolated by methods involving gentle lysis of cells in agarose blocks (UNIT 2.5B; Carle and Olson, 1987). Intact chromosomes can be used to map a cloned gene to a given chromosome by Southern hybridization analysis or as a size standard for field-inversion or orthogonal-field-alteration gel electrophoresis (UNIT 2.5B).

Critical Parameters and Troubleshooting

There is a significant difference among *E. coli* strains regarding their transformation efficiency with plasmid DNA isolated by this protocol. One commonly used strain, HB101 (Table 1.4.5), works well with either CaCl₂ transformation or electroporation protocols (UNIT 1.8). A less commonly used strain, MH1 (Hall et al., 1984), is even better suited for these transformations when made competent by the RbCl method (Hanahan, 1983). Because MH1 is a derivative of MC1061, it is possible that this strain would also transform well using the Hanahan protocol. Many other strains do not work as well, even when they appear to be more competent as judged by transformation with a control plasmid used to standardize transformation efficiency. Therefore, problems with transformations may be due to the choice of the *E. coli* host strain.

It is important to note that only 1 or 2 μ l of the aqueous layer should be used per transformation, as larger amounts are increasingly toxic to *E. coli*. Attempts to purify the DNA with ethanol precipitations are generally unsuccessful, possibly due to co-precipitation of material that inhibits transformation. A 1:1 isopropanol precipitation can be effective in concentrating the DNA when the original attempt fails to produce transformants. In general, however, this additional manipulation is unwarranted.

When preparing DNA for use in Southern hybridization analysis, it is important to minimize the vortexing time needed to break the cells in order to reduce shearing of the DNA.

Three minutes is generally sufficient, but this time can vary from vortex to vortex. Too much vortexing can result in a weak signal for restriction fragments <10 kb. The other consideration for optimizing Southern blot analysis is the amount of DNA used per digest. This protocol generates a very consistent yield of readily digested DNA. Increasing the amount of DNA per digest can result in incomplete cutting and is not recommended.

Anticipated Results

The basic protocol for isolating plasmid DNA from yeast transformants gives sufficient DNA for transformation of *E. coli*. The number of transformants obtained is a function of copy number of plasmid and transformation efficiency of *E. coli* host strain. The protocol for isolating chromosomal DNA yields ~20 μ g DNA from a 10-ml stationary-phase culture.

Time Considerations

It should take ~10 min to prepare plasmid DNA for transformation of *E. coli*. The genomic DNA preparation takes from 45 min to 1 hr for twelve cultures. The vortexing step can become much more time-consuming if only a single-tube vortex is available, as only three microcentrifuge tubes can be vortexed simultaneously.

Literature Cited

- Carle, G.F. and Olson, M.V. 1987. Orthogonal-field-alteration gel electrophoresis. *Methods Enzymol.* 155:468-482.
- Cryer, D.R., Eccleshal, R., and Marmur, J. 1975. Isolation of yeast DNA. *Methods Cell Biol.* 12:39-44.
- Davis, R.W., Thomas, M., Cameron, J., St. John, T.P., Scherer, S., and Padgett, R.A. 1980. Rapid DNA isolations for enzymatic and hybridization analysis. *Methods Enzymol.* 65:404-411.
- Hall, M.N., Hereford, L., and Herskowitz, I. 1984. Targeting of *E. coli* β -galactosidase to the nucleus of yeast. *Cell* 36:1057-1065.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166:557-580.
- Hoffman, C.S. and Winston, F. 1987. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene* 57:267-272.
- Winston, F., Chumley, F., and Fink, G.R. 1983. Eviction and transplacement of mutant genes in yeast. *Methods Enzymol.* 101:211-228.

Contributed by Charles S. Hoffman
Boston College
Chestnut Hill, Massachusetts

Preparation of Yeast RNA

UNIT 13.12

This unit provides two protocols for extraction of RNA from yeast that differ primarily in the method for lysing the yeast cells. The first (basic) protocol isolates RNA directly from intact yeast cells by extraction with hot acidic phenol. This procedure yields RNA that is relatively free of contaminating DNA, is convenient to perform with multiple samples, and gives little or no sample-to-sample variation. In contrast, the first alternate protocol relies upon disruption of cells by vigorous mixing with glass beads and denaturing agents. Although this procedure results in efficient breaking of the cells, the product is associated with residual DNA, and the procedure itself is troublesome when one is working with multiple samples. A second alternate protocol describes the scaling up of the first two procedures to isolate enough total RNA for poly (A)⁺ RNA preparation.

NOTE: Take precautions to avoid contamination by RNases. See UNIT 4.1, reagents and solutions, for instructions.

PREPARATION OF YEAST RNA BY EXTRACTION WITH HOT ACIDIC PHENOL

**BASIC
PROTOCOL**

Yeast RNA can be isolated efficiently and directly from intact cells by extraction with acidic phenol (pH 5) and SDS at 65°C. Because this procedure does not require vortexing individual samples with glass beads (alternate protocol), which is tedious and a source of variability, it is well-suited for obtaining reproducible quantities of RNA from multiple samples. In addition, RNA preparations are largely devoid of contaminating DNA which partitions into the interface during the extraction step.

Materials

Yeast cells and desired medium (UNITS 13.1 & 13.2)

TES solution

Acid phenol

Chloroform

3 M sodium acetate, pH 5.3

100% and 70% ethanol, ice-cold

50-ml centrifuge tube (Falcon)

Centrifuge: tabletop *or* Sorvall equipped with an SS-34 rotor

Additional reagents and equipment for ethanol precipitation (UNIT 2.1) and spectrophotometric quantitation of cells and RNA (APPENDIX 3)

1. Grow yeast cells in 10 ml of desired medium to mid-exponential phase ($OD_{600} = 1.0$).

It is not advisable to prepare RNA from cells that have reached a higher density because as the stationary phase is approached, the results are less consistent and RNA yields will vary.

2. Transfer culture to 50-ml centrifuge tube and centrifuge cells 3 min at $1500 \times g$ (7000 rpm in a tabletop centrifuge or SS-34 rotor), 4°C.

The time and speed of the centrifugation are not critical.

3. Discard supernatant, resuspend pellet in 1 ml ice-cold water. Transfer to a clean 1.5-ml microcentrifuge tube. Microcentrifuge 10 sec at 4°C, and remove supernatant. Proceed to step 4 or if desired immediately freeze pellet by placing tube in dry ice.

Liquid nitrogen may also be used to freeze the pellets. Although not essential, freezing is particularly useful when RNA is to be prepared from multiple cultures or multiple time points from a given culture; this permits simultaneous processing of the samples. The frozen

Saccharomyces cerevisiae

Contributed by Martine A. Collart and Salvatore Oliviero

Current Protocols in Molecular Biology (1993) 13.12.1-13.12.5

Copyright © 2000 by John Wiley & Sons, Inc.

13.12.1

Supplement 23

cell pellets can be stored for months at -70°C . Thaw on ice just before continuing the procedure.

4. Resuspend cell pellet in 400 μl TES solution. Add 400 μl acid phenol and vortex vigorously 10 sec. Incubate 30 to 60 min at 65°C with occasional, brief vortexing.

It is crucial to incubate for ≥ 30 min (with occasional vortexing) to obtain quantitative recovery of both large and small RNA species.

5. Place on ice 5 min. Microcentrifuge 5 min at top speed, 4°C .
6. Transfer aqueous (top) phase to a clean 1.5-ml microcentrifuge tube, add 400 μl acid phenol, and vortex vigorously. Repeat step 5.
7. Transfer aqueous phase to a clean 1.5-ml microcentrifuge tube and add 400 μl chloroform. Vortex vigorously and microcentrifuge 5 min at top speed, 4°C .
8. Transfer aqueous phase to a new tube, add 40 μl of 3 M sodium acetate, pH 5.3, and 1 ml of ice-cold 100% ethanol and precipitate. Microcentrifuge 5 min at top speed, 4°C . Wash RNA pellet by vortexing briefly in ice-cold 70% ethanol. Microcentrifuge as before to pellet RNA.
9. Resuspend pellet in 50 μl H_2O . Determine the concentration spectrophotometrically by measuring the A_{260} and A_{280} (UNIT 4.1). Store at -70°C , or at -20°C if it is to be used within 1 year.

Make sure that the RNA is well dissolved; if necessary, heat the resuspended pellet at 65°C for 10 to 20 min and/or dilute further with more water. The yield from 10 ml of cells grown in YPD medium is ~ 300 μg . Cells grown in less optimal medium will yield less RNA per ml culture.

ALTERNATE PROTOCOL

PREPARATION OF RNA USING GLASS BEADS

Yeast RNA is also efficiently released by disrupting the cells using high-speed mixing in the presence of glass beads and denaturing agents. Proteins are removed by extraction with organic solvents and the RNA is recovered by ethanol precipitation and quantitated by measuring its absorbance at 260 nm. This preparation is suitable for S1, northern hybridization, or primer extension analyses (UNITS 4.6, 4.9, and 4.8, respectively) and can be prepared quickly and easily from a relatively small quantity of yeast cells. Although the RNA isolated by this procedure is contaminated with DNA, the DNA component does not interfere with most analytical studies.

Additional Materials

- RNA buffer
- 25:24:1 phenol/chloroform/isoamyl alcohol (equilibrated with RNA buffer; see support protocol, UNIT 2.1)
- 0.45- to 0.55-mm, chilled, acid-washed glass beads (Sigma)

Prepare the cells

1. Grow and process yeast cells and freeze cell pellet as in steps 1 to 3 of the basic protocol.

If necessary, the samples can now be quick-frozen on dry ice and stored at -70°C . Thaw on ice just before processing.

2. Resuspend pellet in 300 μl RNA buffer.

Disrupt the cells

3. Add a volume of chilled acid-washed glass beads equivalent to ~ 200 μl water.

Prepare acid-washed glass beads by soaking in concentrated nitric acid for 1 hr, washing extensively with deionized water, and drying in a baking oven.

Before use the glass beads should be chilled on ice. Manipulate the beads with a stainless steel spatula that has been dried in a baking oven. Do this carefully, using a spatula or a funnel fashioned out of weighing paper. If beads stick to the lip of the microcentrifuge tube, they will prevent the tube from closing securely.

4. Add 300 μl of 25:24:1 phenol/chloroform/isoamyl alcohol equilibrated with RNA buffer.

A phenol mix previously equilibrated with TE buffer (UNIT 2.1) can be reequilibrated with RNA buffer. Remove the aqueous layer (TE buffer), add a volume of RNA buffer equal to the volume of the organic layer, shake vigorously, and allow the phases to separate. Remove the aqueous layer (RNA buffer), add fresh RNA buffer, mix well, and allow the phases to separate.

5. Close the cap, then invert and shake up and down to ensure that the beads are suspended. Vortex vigorously for 2 min at highest speed.

Hold 2 to 4 tubes on the head of a vortexer for 1 min, place these tubes on ice, and vortex another set. After 1 min with the second set, place on ice, and vortex the first set for another minute. This limits the processing to 8 samples at one time. If a high-speed horizontal shaker is available many samples can be vortexed simultaneously. Use the highest speed setting for 3 min in a cold room.

6. Microcentrifuge 1 min at room temperature. Transfer aqueous (top) layer to a clean microcentrifuge tube.

Avoid the interface by taking only the uppermost 200 to 250 μl from each sample. Taking a fixed amount from each sample results in a more uniform yield of RNA from each sample.

7. Add an equal volume of 25:24:1 phenol/chloroform/isoamyl alcohol (200 to 250 μl). Vortex vigorously 10 sec.

8. Repeat step 6.

Precipitate the RNA

9. Add 3 vol (~600 μl) of ice-cold 100% ethanol. Mix well and place at -20°C for ≥ 30 min or on dry ice for 5 min.
10. Microcentrifuge 2 min at 4°C . Aspirate or pour off the supernatant and wash pellet with ice-cold 70% ethanol.
11. Microcentrifuge 1 min at 4°C . Aspirate or pour off supernatant and dry pellet.
12. Resuspend pellet in 50 μl H_2O . Determine the concentration spectrophotometrically by measuring the A_{260} and A_{280} (UNIT 4.1). Store the RNA at -70°C .

If 2×10^8 cells are used, the RNA concentration of the final solution will be ~2 mg/ml.

PREPARATION OF POLY(A)⁺ RNA

The RNA isolated in the basic and first alternate protocols is also a suitable source of poly(A)⁺ RNA for use in constructing cDNA libraries. For this purpose, larger quantities of RNA can be isolated by simply scaling-up the procedures. For 10¹⁰ *S. cerevisiae* cells, use the following guidelines:

For the hot acidic phenol protocol: Increase volumes of TES and acid phenol solutions to 4 ml each; use 50-ml polypropylene centrifuge tubes for all manipulations. The yield will be ~10 mg of RNA.

For the glass beads protocol: Increase volumes to 15 ml RNA buffer, 10-ml volume glass beads, 15 ml phenol/chloroform/isoamyl alcohol, and 30 ml of 100% ethanol. Perform the procedure in a 50-ml disposable polypropylene tube and centrifuge 10 min, 3000 rpm (1200 × g), 4°C in a fixed-angle or swinging-bucket type rotor at each phenol extraction step. Good recovery of the precipitated nucleic acid can be accomplished by centrifugation under these same conditions. The yield will be ~5 mg total RNA.

Prepare poly(A)⁺ RNA from either method using the protocol presented in *UNIT 4.5*.

REAGENTS AND SOLUTIONS

Acid phenol

Add sufficient water to a bottle of solid phenol such that phenol is water-saturated; pH will be ~5.0. Do not buffer phenol. Store at 4°C, protected from light.

RNA buffer

0.5 M NaCl
200 mM Tris·Cl, pH 7.5
10 mM EDTA
Store indefinitely at room temperature

TES solution

10 mM Tris·Cl, pH 7.5
10 mM EDTA
0.5% SDS
Store indefinitely at room temperature

COMMENTARY

Background Information

Aside from the harsh conditions used to break open yeast cells, the RNA isolation procedures presented here are similar to the phenol/SDS method for isolating RNA from plant cells (*UNIT 4.3*). Both methods take advantage of the fact that phenol extraction is an effective means of inactivating and removing RNases. The hot acidic phenol method is preferable when working with multiple samples; because the number of manipulations and time required for each are reduced, very little sample-to-sample variation is encountered.

The glass bead disruption procedure causes the release of both RNA and DNA, and the absorbance at 260 nm will measure total nucleic acid. For precise determination of RNA con-

centrations, the DNA component can be removed as described in *UNITS 4.1, 4.3, and 4.5*. This is not necessary for routine northern blot, primer extension, and S1 analyses. Refer to *UNITS 4.1-4.5* for further information on the purification and properties of RNA.

Critical Parameters

As with all RNA manipulations, precautions must be taken to prevent RNase contamination. See *UNIT 4.1* for details. In the hot acidic phenol method, the two phenol and one chloroform extractions are critical to obtain clean RNA for analysis.

These protocols are designed to process multiple samples in tandem—for example, when monitoring gene induction over time or

when examining RNA synthesis driven by different promoter constructions. In these cases (where quantitation is critical), detection using probes in northern hybridization, S1 mapping, or primer extension studies should be carried out in parallel (preferably in the same tube or hybridization bag) with a probe for a gene in which transcription does not change over time or in response to inducing conditions. The ratios of signal intensities for the two probes in each sample are then compared.

Anticipated Results

The yield of RNA isolated using the hot acidic phenol protocol is ~300 µg from 2×10^8 . The yield resulting from the glass bead disruption protocol starting with 2×10^8 cells will be ~100 µg of total RNA. Scale-up of either method will produce ~5 to 10 mg total RNA. Ten to twenty micrograms of total RNA isolated by either procedure is sufficient for most northern blot, S1 mapping, or primer extension applications.

Time Considerations

Depending on the organization of the researcher, 12 to 24 samples can be processed conveniently in about 1 hr using the hot acidic phenol method. For the glass bead method, eight samples can be easily processed without using a high-speed shaking apparatus in ~1.5 hr. A second set of eight samples can be processed while the first set is precipitating at -20°C . With a high-speed shaking apparatus 16 samples can be processed in about 1 hr. Both phenol extraction steps can be done on the shaking apparatus.

If RNA synthesis is being monitored over time, samples from individual time points can be frozen at -70°C (see step 3, basic protocol) so that all of the samples can be processed simultaneously.

Contributed by Martine A. Collart and
Salvatore Oliviero
Harvard Medical School
Boston, Massachusetts

Preparation of Protein Extracts from Yeast

Three protocols are presented for preparing protein extracts; they differ primarily in the way the cells are broken. In the basic protocol, cells are enzymatically converted to spheroplasts, which are then lysed by a combination of osmotic shock and Dounce homogenization. A support protocol for isolating intact nuclei by differential centrifugation is also presented. An alternate protocol describes mechanical breakage of cells by vortexing in the presence of glass beads. In a second alternate protocol, growing cells are frozen immediately in liquid nitrogen and then lysed by grinding in an industrial-strength blender in the presence of liquid nitrogen. These methods each have advantages and disadvantages; the best choice will depend on the particular application. For all of these procedures, it is advantageous to use protease-deficient strains such as BJ926 or EJ101.

BASIC PROTOCOL

SPHEROPLAST PREPARATION AND LYSIS

Materials

Protease-deficient yeast cells (BJ926, EJ101, or equivalent)

YPD medium (UNIT 13.1)

Zymolyase buffer, room temperature and ice-cold

Zymolyase 100T (ICN Immunobiologicals)

1 M sorbitol (optional)

Lysis buffer

Extraction buffer

Storage buffer

Liquid nitrogen

Sorvall GS-3 or GSA rotor (or equivalent)

Sorvall SS-34 or SA-600 rotor (or equivalent)

Beckman Type 45Ti rotor (or equivalent; large capacity)

30°C shaker platform

Rubber policeman

Dounce homogenizer

Rotating wheel or rocker

Additional reagents and equipment for growing yeast cells (UNITS 13.1 & 13.2), large-volume dialysis (APPENDIX 3), and determining conductivity (UNIT 10.10)

Grow the cells

1. Grow cells to mid-log phase in YPD medium with vigorous shaking or forced aeration.

The OD_{600} can vary with aeration conditions, but should be between 1 and 5. The procedure is appropriate for between 100-ml and 20-liter cultures, although it can be scaled up if necessary.

2. Harvest cells by centrifugation 5 min at $1500 \times g$ (GS-3 or GSA rotor at ~ 3000 rpm), 4°C, in preweighed centrifuge bottles.

Prepare the spheroplasts

3. Determine the wet weight (in grams) of yeast cells in the pellet by the weight increase over that of the preweighed bottle. This is approximately equal to the packed cell volume (in milliliters), and for all subsequent steps will be considered 1 vol.

One liter of BJ926 (a diploid strain) at $OD_{600} = 1.0$ yields a packed cell volume of 2 to 3 ml.

4. Resuspend cells in 2 to 4 vol ice-cold water and immediately centrifuge 5 min at $1500 \times g$ (SS-34 or SA-600 rotor at 3500 rpm), 4°C . Discard the supernatant.
5. Resuspend the cells by adding 1 vol zymolyase buffer containing 30 mM DTT (see note in reagents and solutions), and incubate 15 min at room temperature.

This step facilitates subsequent zymolyase treatment and spheroplast lysis by breaking disulfide bonds.

6. Centrifuge 5 min at $1500 \times g$, 4°C , and resuspend in 3 vol zymolyase buffer. Add 2 mg (200 U) Zymolyase 100T per ml of original packed cell volume to the resuspended cells. Incubate 40 min at 30°C on a shaker platform at ~ 50 rpm.
7. Determine if conversion to spheroplasts has been completed by the lysis in water technique (UNIT 13.7). If spheroplasting is incomplete, continue incubation until complete.

Perform all procedures from this point on at 4°C .

8. Centrifuge spheroplasts 5 min at $1500 \times g$. Decant the supernatant carefully—the spheroplast pellet will not be as tight as the previous cell pellets.

For some procedures (e.g., extracts for in vitro transcription or translation), spheroplasts are resuspended in YPD medium containing 1 M sorbitol and incubated for 30 to 60 min at 30°C to allow metabolic recovery.

9. Wash the spheroplasts by gently resuspending the pellet in 2 vol ice-cold zymolyase buffer and centrifuging 5 min at $1500 \times g$. Repeat this step two more times.

Spheroplasts are sticky and difficult to resuspend. To facilitate resuspension, first resuspend the spheroplasts in a small volume with the aid of a rubber policeman and then add more buffer to achieve the correct final volume.

The washing step is important for removing proteases, phosphatases, and nucleases present in the zymolyase preparation. For some purposes, it may be necessary to carry out additional washes.

Lyse the spheroplasts

10. Gently resuspend the pellet in 2 vol lysis buffer. Do not try to achieve a homogeneous suspension; simply dislodge the pellet from the side of the centrifuge tube and gently swirl 10 to 20 times. Centrifuge spheroplasts 10 min at $1500 \times g$.

Extensive manipulation of the pellet may result in premature osmotic lysis.

11. Thoroughly resuspend spheroplast pellet with 1 vol lysis buffer using a glass rod.

At this point, tubes containing the resuspended spheroplasts can be quick-frozen in liquid nitrogen and stored at -80°C . Thaw frozen spheroplasts overnight on ice before proceeding.

12. Lyse spheroplasts with 15 to 20 strokes of a tight-fitting pestle (clearance 1 to 3 μm) in a Dounce homogenizer.

Extract proteins from lysate

13. Half-fill ultracentrifuge tubes with lysate. Add an equal volume of extraction buffer and seal the tubes. Gently invert tubes on rotating wheel or rocker for 15 to 30 min at 4°C .

The lysate and extraction buffer are not premixed because the resulting solution will immediately become quite viscous and hence difficult to pour if the ionic strength is above ~ 0.5 M.

14. Centrifuge 90 min at $100,000 \times g$ (Type 45Ti rotor at 33,000 rpm), 4°C .
15. Collect supernatant and dialyze 2 to 4 hr against 100 vol storage buffer. Transfer dialysis bag to 100 vol fresh storage buffer; dialyze an additional 2 to 4 hr.

A flocculent precipitate may form during dialysis. These precipitates usually contain negligible amounts of most protein factors and can be discarded.
16. Remove a few microliters of the dialysate, dilute 1:1000 with water, and determine the conductivity (UNIT 10.10). If it is equal to that of similarly diluted storage buffer or below some acceptable value (usually 100 to 250 mM NaCl), proceed to step 17. If not, continue dialysis.
17. Centrifuge dialysate 10 min at $10,000 \times g$ (SS-34 rotor at 9200 rpm or SA-600 rotor at 8500 rpm), 4°C . Collect the supernatant, freeze in small aliquots in liquid nitrogen, and store at -80°C .

This crude extract contains most DNA-binding proteins as well as transcription and replication factors. The pellet contains proteins that can be "salted in" by resuspending in storage buffer containing 0.5 to 1.0 M KCl, if desired.

SUPPORT PROTOCOL

NUCLEI PREPARATION BY DIFFERENTIAL CENTRIFUGATION

Nuclei suitable for chromatin studies and/or nuclear protein extracts are prepared by osmotically lysing spheroplasts in the presence of Ficoll, which preserves nuclear structure and prevents proteins from leaking out of the nucleus, followed by differential centrifugation.

Additional Materials

Ficoll buffer, ice-cold
Teflon pestle tissue homogenizer, motor-driven (optional; Thomas)

1. Perform steps 1 to 9 of basic protocol. Resuspend cells in 0.5 vol zymolyase buffer.
2. Pipet cells drop by drop into a beaker containing 15 to 25 vol ice-cold Ficoll buffer with continuous stirring in an ice bath or cold room.

Alternatively, resuspend cells in several volumes Ficoll buffer and homogenize with 5 to 10 strokes using a motor-driven tissue homogenizer and moderately tight Teflon pestle (clearance 0.15 to 0.23 mm) at medium speed.

3. Transfer the suspension to centrifuge tubes and centrifuge 5 min at $3000 \times g$ (SS-34 rotor at 5000 rpm), 4°C , to pellet cell debris and unlysed spheroplasts.

This step can be skipped if separation of unlysed cells is not critical, or it can be repeated multiple times if it is important to completely remove debris.

4. Transfer supernatant to new centrifuge tubes and centrifuge 20 min at $20,000 \times g$ (SS-34 rotor at 13,000 rpm), 4°C . Decant the supernatant. The pellet contains the nuclei.

The pellet can be difficult to resuspend. If necessary, resuspend as described in step 9 of basic protocol.

For chromatin studies, the nuclei can be suspended in an appropriate buffer and used directly for micrococcal nuclease or DNase I digestion. In some instances, it is useful to wash the nuclei in a buffer equivalent to Ficoll buffer but containing 10% glycerol instead of Ficoll. Nuclei can be stored in this buffer several months at -80°C if the tubes are frozen quickly on dry ice or liquid nitrogen.

5. To obtain nuclear extracts, resuspend the nuclei in 1 vol lysis buffer and perform

steps 12 to 17 of basic protocol.

Alternatively, the nuclei can be washed in buffers containing various types and molarities of salt. The supernatants will contain proteins that have been extracted by the salt.

**ALTERNATE
PROTOCOL**

CELL DISRUPTION USING GLASS BEADS

These procedures are for small-scale preparations using a vortexer, and for large-scale preparations using a Bead Beater.

Additional Materials

Glass bead disruption buffer
Chilled, acid-washed glass beads (0.45- to 0.55-mm; see reagents and solutions)
Bead Beater and vessel (Biospec; optional)

1. Grow and harvest yeast cells as described in steps 1 to 4 of the basic protocol for spheroplast lysis. Determine the packed cell volume.

All subsequent steps should be carried out at 4°C.

2. Resuspend cells in 1 vol glass bead disruption buffer.

The cells can be frozen by slowly pouring this suspension into a plastic beaker filled with liquid nitrogen. Use enough liquid nitrogen to submerge the frozen paste. This frozen "popcorn" can be stored at -80°C. Thaw overnight on ice before proceeding.

3. Mix cell paste with 2 vol glass bead disruption buffer.
4. Add 4 vol of chilled, acid-washed glass beads.

For packed cell volumes of <10 ml:

- 5a. Transfer the cell suspension to an appropriately sized screw-cap centrifuge tube. The suspension should occupy no more than 60% to 70% of the capacity of the tube.
- 6a. Vortex the suspension at maximum speed for 30 to 60 sec at 4°C; place tube on ice for 1 to 2 min. Repeat 3 to 5 more times. Check the amount of cell breakage by visual inspection under a microscope. Proceed to step 7.

Although larger volumes may be processed by using multiple aliquots, it is better to use a Bead Beater as described below.

For packed cell volumes of >10 to 20 ml:

- 5b. Transfer the cell suspension to an appropriately sized Bead Beater vessel (stainless steel is recommended for better heat transfer) and add glass bead disruption buffer to fill the vessel almost to the brim. The volume of buffer required to fill the vessel should not exceed 1 cell-suspension volume. Attach the blade and cap assembly, ensuring that all air is excluded from the vessel.

It is important to exclude air in order to prevent foaming and potential protein denaturation.

- 6b. Grind at high speed for 60 sec, then let sit 1 to 2 min on ice. Repeat 3 to 5 more times.
7. Allow the glass beads to settle out and decant the supernatant.
8. Add 2 to 4 vol glass bead disruption buffer to the glass beads and invert the tube 5 to 10 times. Allow the beads to resettle and decant the supernatant. Pool the supernatants.
9. Centrifuge the pooled supernatants 60 min at 12,000 × g (SS-34 rotor at 10,000 rpm),

***Saccharomyces
cerevisiae***

13.13.4

4°C. Collect the supernatant which represents the crude extract. For long-term storage, aliquot into small tubes, quick-freeze the tubes in liquid nitrogen, and store at -80°C.

**ALTERNATE
PROTOCOL**

CELL DISRUPTION USING LIQUID NITROGEN

This protocol is designed for processing 200 ml to ~20 liters of cells. It can be easily scaled up for processing larger amounts (e.g., from a fermentor), in which case a larger blender and cup must be used.

Additional Materials

Yeast cakes (optional; Red Star)

Liquid nitrogen

60-ml syringe

1-liter plastic beaker (Nalgene or equivalent)

1-liter stainless steel blender cup and blender (Waring or equivalent)

1. Grow cells to mid-log phase in YPD (or selective) medium with vigorous shaking or forced aeration. Centrifuge the culture to harvest cells. Discard supernatant.

Many liters can be processed at one time with this method. Alternatively, commercial yeast cakes can be used in place of the growing cells.

2. Vortex cell pellet to create a thick cell paste. If necessary, add a minimal amount of ice-cold water to allow the paste to be poured or spooned.

NOTE: Keep cell paste on ice in subsequent steps.

If using a yeast cake, mix equal volumes of cells and water and blend into a thick paste.

3. Remove the plunger from a 60-ml syringe. Seal the bottom with a plug or tightly wrapped parafilm. Pour or spoon in 50 ml of the cell paste.

4. Place 400 ml liquid nitrogen into a 1-liter Nalgene beaker.

Avoid contact with liquid nitrogen due to the danger of frostbite. Use insulated gloves when handling containers. Never use glass containers to hold liquid nitrogen as they may break.

5. Hold the syringe over the liquid nitrogen; remove plug or parafilm from bottom, insert plunger, and squeeze the cell paste into the liquid nitrogen. Repeat steps 3 and 5 no more than three times. A maximum of ~200 ml cell paste can be processed at one time when using a 1-liter blender cup.

Long spaghetti-like aggregations of frozen cells should form with a thick cell paste; spherical popcorn-like clumps form with a thinner paste. The frozen cells can be stored indefinitely at -80°C before disruption.

6. Carefully pour the liquid nitrogen with frozen cells into a blender cup that is already attached to the blender and has been thoroughly dried prior to use.

It is preferable to perform the grinding in a cold room with a precooled blender to decrease the rate of evaporation of the liquid nitrogen, but the procedure can be done at room temperature if necessary.

Only industrial-strength blenders and stainless steel blender cups can be used for this protocol (a Waring commercial blender is recommended). Care should be taken that all moving parts of both the blender and blender cup are completely free of moisture in order to avoid freezing when adding the liquid nitrogen.

7. Place the lid on the blender cup, making sure the lid vents are open so that pressure does not build up in cup as liquid nitrogen evaporates. Keeping the lid held down tightly, grind at high speed in three successive 2-min bursts. Between bursts, mix frozen powder and, if necessary, add liquid nitrogen to just cover powder.

Begin grinding as soon as possible after pouring in the liquid nitrogen to avoid freezing the moving parts. Some material will spurt out of the lid vents during grinding; take care not to get in its way. It is advisable to cover the working area with plastic sheeting prior to grinding to ease the clean-up. Pulsed grinding may help avoid the spurting problem, but keep in mind that speed is critical here—the frozen powder should not be allowed to thaw.

8. After grinding, pour the fine frozen yeast powder into a beaker containing twice the original cell-paste volume of ice-cold storage buffer. Mix and then pour the suspension into a centrifuge bottle. Keep bottle on ice.
9. Repeat steps 3 through 8 until all the cell paste has been processed.
10. Centrifuge 15 min at $5000 \times g$ (GSA or GS-3 rotor at 5500 rpm; or SS-34 rotor at 6500 rpm), 4°C . The supernatant contains a crude extract whose concentration should be between 10 and 20 mg protein/ml. It can be used immediately for assays or as a starting point for protein purification. For long-term storage, aliquot the supernatant into plastic tubes, freeze quickly on dry ice, and store at -80°C .

Alternatively, the frozen yeast powder can be resuspended in 1 vol lysis buffer and treated as described in steps 13 to 17 of the basic protocol for spheroplast lysis.

REAGENTS AND SOLUTIONS

Extraction buffer

Lysis buffer (see below)
0.8 M ammonium sulfate
20% glycerol

Ficoll buffer

18% (w/v) Ficoll-400 (Pharmacia)
10 mM Tris-Cl, pH 7.5
20 mM KCl
5 mM MgCl_2
3 mM DTT
1 mM EDTA
1 \times protease inhibitor mix (see below)
1 mM PMSF

Glass bead disruption buffer

20 mM Tris-Cl, pH 7.9
10 mM MgCl_2
1 mM EDTA
5% glycerol
1 mM DTT
0.3 M ammonium sulfate
1 \times protease inhibitor mix (see below)
1 mM PMSF

The concentration of ammonium sulfate in the buffer can be varied between 0.1 and 1.0 M. Final concentrations above 0.25 M strip specific DNA-binding proteins and histones off chromatin and are therefore useful in obtaining factors that interact with nucleic acids. KCl and NaCl can also be added to final concentrations between 0.1 and 2.0 M.

Saccharomyces cerevisiae

13.13.6

Glass beads, chilled and acid-washed, 0.45- to 0.55-mm

Wash the beads by soaking 1 hr in concentrated nitric acid. Rinse thoroughly with water. Dry the beads in a baking oven, cool to room temperature, and store at 4°C until needed.

Lysis buffer

50 mM Tris·Cl, pH 7.5
10 mM MgSO₄
1 mM EDTA
10 mM potassium acetate
1 mM DTT
1× protease inhibitor mix (see below)
1 mM phenylmethylsulfonyl fluoride (PMSF)

100× protease inhibitor mix

Listed below are representative protease inhibitors; different combinations may be more appropriate for individual applications.

10 µg/ml chymostatin
200 µg/ml aprotinin
100 µg/ml pepstatin A
110 µg/ml phosphoramidon
720 µg/ml E-64
50 µg/ml leupeptin
250 µg/ml antipain
10 mM benzamidine
10 mM sodium metabisulfite

Storage buffer

20 mM Tris·Cl, pH 7.5
0.1 mM EDTA
10% glycerol
100 mM KCl
1 mM DTT
1× protease inhibitor mix (see above)
1 mM PMSF

Zymolyase buffer

50 mM Tris·Cl, pH 7.5
10 mM MgCl₂
1 M sorbitol
1 mM *or* 30 mM DTT (see annotation below)

For step 5 of basic protocol use 30 mM DTT in this buffer. For all other applications, use 1 mM DTT.

COMMENTARY

Background Information

The three protocols for preparing protein extracts from yeast differ primarily in the way the cells are broken. Each of these methods has advantages and disadvantages that should be considered when choosing one for a particular application. Most importantly, the extracts made by these different methods may contain different proteins and different levels of protein activity. Thus, it may be

necessary to try different methods to obtain extracts capable of carrying out the *in vitro* reactions of interest.

Spheroplast procedure. Cells are converted to spheroplasts by zymolyase treatment, and then lysed by a combination of osmotic shock and Dounce homogenization. Because this method is the most gentle way to break yeast cells, it is most suitable for preparing extracts that can carry out complex enzymatic

functions (e.g., translation, transcription, DNA replication) and in which the integrity of macromolecular structures (e.g., ribosomes, spliceosomes) has been maintained. It is also useful for isolating intact nuclei that can be used for chromatin studies (Bloom and Carbon, 1982) or for nuclear protein extracts (Lue and Kornberg, 1987). The support protocol described here for isolating nuclei derives primarily from Nelson and Fangman (1979) and is based on differential centrifugation in the presence of high concentrations of Ficoll (a polymer that inhibits leakage of nuclear proteins into the cytoplasm). Another procedure for purifying the nuclei (after release by spheroplast lysis) employs a Percoll gradient (Amati and Gasser, 1988). Although these nuclei are slightly more purified, they seem to be of varying or lower quality than those produced by the differential centrifugation (based on nucleosome ladder integrity). The major disadvantages of the spheroplast lysis procedure are that it is relatively tedious and expensive, especially for large-scale preparations (>10 liters), and the long incubation periods can lead to proteolysis or protein modification.

Liquid nitrogen procedure. Initially described by Sorger and Pelham (1987), cells are frozen immediately in liquid nitrogen and then lysed by grinding in a Waring blender in the presence of liquid nitrogen. The protocol is quick and easy (albeit a bit messy and potentially dangerous to the careless investigator). It can accommodate varying amounts of yeast cells including very large cultures. Its main advantage is that cells are taken immediately from the actively growing state into liquid nitrogen (-135°C), decreasing degradative enzyme activities such as proteases and nucleases as well as activities that modify proteins (e.g., phosphatases and kinases). It is particularly suited for making whole-cell extracts from a single yeast culture for large-scale protein purification. A detailed description of the technique as applied to very large-scale cultures (100 to 400 liters) is given by Sorger et al. (1989). A drawback to the liquid nitrogen protocol is that small samples (i.e., 10- to 100-ml yeast cultures) are not easily processed because there is not enough mass of frozen cell clumps to fracture effectively in the blender. Stainless steel blender cups with very small capacities (10 to 100 ml) can be obtained and may be useful for processing such small amounts. Multiple samples pose another problem for the liquid nitrogen technique because, usually, only one blender is available, making it time-

consuming to process individual samples and to clean the equipment between uses.

Glass bead procedure. This protocol involves mechanically breaking cells by vortexing in the presence of glass beads. It is very flexible because it can be carried out easily on multiple small (10-ml) cultures as well as on very large cultures with appropriate equipment such as a Bead Beater (Klekamp and Weil, 1982). It is particularly useful when making extracts from many different small yeast cultures for assaying purposes rather than for protein purification. Unfortunately, during the glass bead procedure, proteins are treated harshly causing extensive foaming. The amount of cell breakage varies, and proteolysis and modification of the proteins may result from heating the extract above 4°C during the mechanical breakage.

Critical Parameters

For all these procedures, it is advantageous to use protease-deficient strains such as BJ926 or EJ101, but laboratory strains of a desired genotype can be used. In addition, commercial yeast cakes can be used for some applications and have the advantage that limitless quantities of material can be obtained for little cost and no work. The liquid nitrogen method is probably best for processing yeast cakes, while the spheroplast method is unsuitable for this application.

Another crucial factor, even for extracts from protease-deficient strains, is the inclusion of a wide collection of protease inhibitors at all stages during and after lysis. The protease inhibitor mix given in this procedure, along with PMSF and EDTA (to inhibit metalloproteases) should inactivate almost all types of proteases, although other combinations may be more appropriate for particular applications. A final general consideration is to be sure that all solutions coming in contact with the extracted proteins are kept ice cold.

Spheroplast procedure. For this protocol, it is important to handle the spheroplasts gently. The amount of time and zymolyase needed to achieve complete spheroplasting can vary quite a bit between strains, and some adjustments of the amounts given in the protocol may be necessary. Complete spheroplasting is essential for efficient lysis. Only the highest grade Zymolyase 100T should be used. However, because even this grade of zymolyase contains considerable proteolytic activity, it is important to wash the spheroplasts several times prior to lysis.

Liquid nitrogen procedure. For optimal yields of protein, all (or nearly all) the cells should be broken. This should be tested by suspending a small amount of the ground powder in buffer and visually inspecting under the microscope. (Remember to add more liquid nitrogen to the remaining powder to keep it frozen.) If many intact (phase-refractile) cells are seen, further grinding in liquid nitrogen should be performed. Care should be taken to ensure that the liquid nitrogen covering the cells does not evaporate. The frozen yeast powder should always be poured into cold extraction buffer as quickly as possible.

Glass bead procedure. It is crucial that the beads be the correct size. Cell breakage will be inefficient if the wrong size beads are used. As is the case for the liquid nitrogen procedure, the degree of breakage should be examined before proceeding. Finally, it is very important that the temperature of the solution during the breakage remains between 0° and 4°C. Because mechanical disruption generates heat, it is usually necessary to vortex in short bursts so that the solution cools down between intervals. It is also important to minimize foaming during vortexing as this is indicative of protein denaturation which may result in considerable loss of activity. Because the Bead Beater generates much less foaming than vortexing, it is preferred for large-scale preparations.

Anticipated Results

Cell extracts produced by any of these procedures contain ~10 to 30 mg/ml of the total cellular protein. These extracts can be used directly for some purposes, such as mobility shift DNA binding assays (Chapter 12), or can serve as the first step in protein purification (Chapter 10). The extracts will also contain large amounts of nucleic acid, especially RNA. Nuclear extracts will contain significant concentrations of chromosomal DNA.

Time Considerations

After growing the cells, the spheroplast procedure should take between 3 to 12 hr depend-

ing on the amount of yeast cells being processed. The most time-consuming and tedious steps are the resuspensions of the spheroplasts. Isolation of nuclei will take an additional 2 to 4 hr depending on the size of the preparation. For the liquid nitrogen or glass bead procedures, several liters of a single culture sample can be processed in an hour or less; larger cultures require more time.

Literature Cited

- Amati, B.B. and Gasser, S.M. 1988. Chromosomal ARS and CEN elements bind specifically to the yeast nuclear scaffold. *Cell* 54:967-978.
- Bloom, K.S. and Carbon, J. 1982. Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. *Cell* 29:305-317.
- Klekamp, M.S. and Weil, P.A. 1982. Specific transcription of homologous class III genes in yeast *Saccharomyces cerevisiae* soluble cell-free extracts. *J. Biol. Chem.* 257:8432-8441.
- Lue, N.F. and Kornberg, R.D. 1987. Accurate initiation at RNA polymerase II promoters in extracts from *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 84:8839-8843.
- Nelson, R.G. and Fangman, W.L. 1979. Nucleosome organization of the yeast 2 μ DNA plasmid: A eukaryotic minichromosome. *Proc. Natl. Acad. Sci. U.S.A.* 76:6515-6519.
- Sorger, P.K. and Pelham, H.R.B. 1987. Purification and characterization of a heat-shock element binding protein from yeast. *EMBO J.* 6:3035-3041.
- Sorger, P.K., Ammerer, G., and Shore, D. 1989. Identification and purification of sequence-specific DNA-binding proteins. *In* Protein Function: A Practical Approach (T.E. Creighton, ed.) pp. 199-223. IRL Press. Oxford.

Contributed by Barbara Dunn
Memorial Sloan-Kettering Cancer Center
New York, New York

C. Richard Wobbe
Harvard Medical School
Boston, Massachusetts

Overview of *Schizosaccharomyces pombe*

UNIT 13.14

The fission yeast *S. pombe* is a popular model eukaryote that offers complementary strengths to the more familiar budding yeast *S. cerevisiae*. The organisms are evolutionarily distant with an estimated 1000 million years separation. Fission yeast has been dubbed an archaeoscomycete, making it quite separate from common ascomycetes, including *S. cerevisiae* and the filamentous fungi (Hedges, 2002); however, like budding yeast, fission yeast is genetically tractable, and lends itself to easy molecular manipulation. Its increasing popularity in recent years, especially in labs studying mammalian cell biology, has led to the nickname “micromammal.”

Fission yeast was discovered in 1893, but wasn't commonly taken up as a model organism until the 1960s (e.g., see Mitchison, 1990; Egel, 2000; Yanagida, 2002), largely for studies of growth control and differentiation. In the 1970s, Nobel laureate Paul Nurse and his colleagues isolated cell cycle mutants that either blocked the cell division cycle, or accelerated it, making *S. pombe* an important model for cell cycle control (Nurse, 2000). *S. pombe* continues as an important system for studying cell cycle regulation, sexual differentiation, and more recently, chromosome dynamics and polarized cell growth. Other areas were less developed but are now growing as new investigators continue to move into the system. Given its wide divergence from *S. cerevisiae*, *S. pombe* provides a complementary model system, and the compare and contrast approach to problems using both species has proven remarkably fruitful in providing models for biology of larger cell types (e.g., Forsburg, 1999). With a haploid genome size of just 13.8 Mb, distributed amongst 4,824 open reading frames (compared to over 5500 for *S. cerevisiae*), *S. pombe* is the smallest sequenced eukaryote (Wood et al., 2002). Unlike *S. cerevisiae*, there is no evidence of genome-wide duplications in *S. pombe*.

There are ~145 genes in *S. pombe* that have metazoan homologs but are not found in *S. cerevisiae*, and a similar number vice versa (Wood et al., 2002). These genes tend to represent groups of proteins involved in particular cellular functions. For example, *S. pombe* has

three relatively large chromosomes of 5.7, 4.6, and 3.5 Mb. These have features typical of higher eukaryotes, including large diffuse replication origins (~1 kb) and large heterochromatic centromeres (~40 to 100 kb). These elements do not contain short consensus sequences, as in the more numerous small chromosomes of budding yeast, but are defined by activity typical of the same elements in higher cells. Many of the proteins common to *S. pombe* and metazoans that are missing in *S. cerevisiae* are implicated in chromosome functions, including heterochromatin factors such as Swi6/HP1, chromatin modifiers such as the histone methyltransferase Clr4/SuVar2-9, telomere proteins such as Taz1/TRF, centromere proteins such as CENP-B, and components of RNAi apparatus such as Argonaut and Dicer. Not surprisingly, *S. pombe* is proving to be a particularly appropriate model system for studies of mammalian chromosome biology. Similarly, given its distinct rod-shaped morphology and division by medial fission, it is also a good model for polarity and cytokinesis. Other problems are also easily approached.

The choice of system is in part driven by the problem under study, but also by the availability of local expertise. Relatively few U.S. investigators chose fission yeast as a system prior to the explosion of the cell cycle studies in the 1980s, in contrast to its strong popularity in Japan and Europe; however, *S. pombe* is now well represented throughout North America, as well as in its traditional strongholds. In addition, a number of groups with experience in mammalian cell biology have picked up *S. pombe* as their alternative model system, leading to a substantial group of fission yeast researchers who work in labs that do not consider themselves yeast groups. The fission yeast community overall is relatively small by comparison to that of budding yeast but is friendly and international in outlook. It is held together by an international meeting in alternate years, and extensive online resources.

There is a persistent myth suggesting that *S. pombe* cannot be manipulated to the same extent as budding yeast. In practice, working in *S. pombe* is not strikingly different from working in *S. cerevisiae*. Most tools available in *S.*

Contributed by Susan L. Forsburg

Current Protocols in Molecular Biology (2003) 13.14.1-13.14.3

Copyright © 2003 by John Wiley & Sons, Inc.

Yeast

13.14.1

Supplement 64

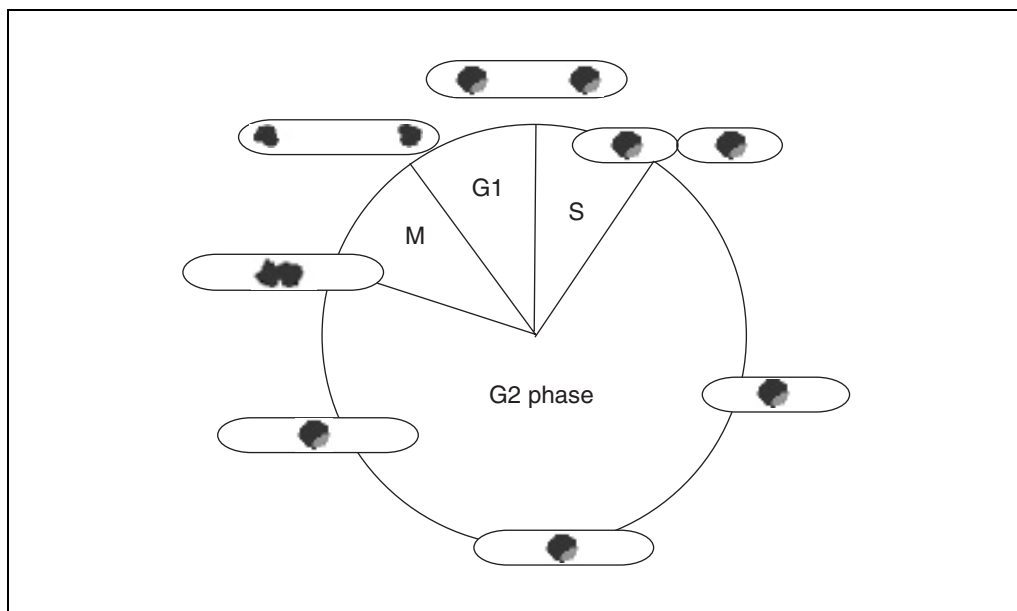


Figure 13.14.1 The fission yeast cell cycle. Rod-shaped cells grow at the ends, maintaining constant diameter. In exponentially growing cells, cytokinesis occurs after daughter nuclei have already entered their next S phase.

cerevisiae are available in *S. pombe* versions that accommodate the distinct biology of the fission yeast, and similar genetic strategies are available for both systems (e.g., Forsburg, 2001). Fission yeast still lacks well developed genomics resources, presumably because its genome sequence was only completed in 2002. Commercial microarrays or oligonucleotide libraries are not widely available, and there is as yet no standard collection of disruptions of all open reading frames. However, other community-wide resources have been developed, including libraries of temperature-sensitive and insertion mutants at FYSSION at the University of Sussex, UK (<http://pombe.biols.susx.ac.uk>), which may be freely screened. A large expression project is underway at the Sanger Centre and the data are freely searchable (http://www.sanger.ac.uk/PostGenomics/S_pombe).

S. pombe cells are rod shaped and grow by increasing length while maintaining a constant diameter. The length of a cell is a sensitive indicator of its position within the cell cycle. They divide by medial fission, producing two essentially identical daughter cells. The nuclear cell cycle is divided into distinct G1 (10%), S (10%), G2 (70%), and M (10%) phases. Following mitosis, the newly replicated nuclei enter the next cell cycle and undergo G1 and S phase prior to completion of the previous cycle's cytokinesis. This quirk means that a single cell particle almost always has a 2C DNA content, either because of the extended G2 phase, or the binucleate G1 or S phase cells (see

Fig. 13.14.1). Fission yeast is generally haploid even in the wild. Following conjugation, newly formed zygotes immediately enter meiosis and sporulation to produce four spores in a linear tetrad ascus. Diploids can be recovered in the laboratory by selection for complementing markers, but are unstable and prone to sporulate even on minimal media. One of the major differences in *S. pombe* and *S. cerevisiae* is in handling mating and diploids (see UNIT 13.16).

Approximately 43% of genes have at least one intron and multiple introns are common; however, the introns are usually short, and most are <100 bp in length (Wood et al., 2002). They do not have the rigid splicing consensus observed in *S. cerevisiae*. TATA boxes and other consensus sequences for transcription are more proximal to the ATG than observed in budding yeast (Russell, 1983). For this reason, it cannot be assumed that fission yeast genes will express appropriately in budding yeast, or vice versa. Thus, for cross-complementation experiments it is best to use a cDNA and a promoter native to the host species. Plasmids and markers for *S. pombe* are conceptually similar to those in *S. cerevisiae*, with the exception of single-copy centromere plasmids. Since the fission yeast centromeres are so large, they cannot be accommodated on a typical episome. While some *S. cerevisiae* plasmids and markers can be maintained in *S. pombe*, and were widely used in the early days of fission yeast molecular biology, they are not optimal and are prone to plasmid loss or rearrangement. Plasmids containing *S.*

pombe-specific replication origins and markers are strongly recommended and a large variety now exist.

Most protocols described in the sections on *S. cerevisiae* are broadly applicable to fission yeast with only minor changes. The author's purpose here is to identify those common *S. pombe*-specific methods that are required by a new investigator but are substantially different from similar protocols in budding yeast, including specific media (UNIT 13.15), crosses and diploids, tetrad analysis, and cell cycle synchronization (UNIT 13.16), and some types of transformation (UNIT 13.17). General handling of the organism, microbial culture, and overall purification of nucleic acid and protein are similar to budding yeast. Generally, an investigator who has handled *E. coli* in the course of performing standard molecular biology manipulations will have no trouble working with yeast of either species. Additional resources for working with *S. pombe* can be found in (Moreno et al., 1991; Alfa et al., 1993; Gould, 2003).

LITERATURE CITED

- Alfa, C., Fantes, P., Hyams, J., McLeod, M., and Warbrick, E. 1993. Experiments with Fission Yeast. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Egel, R. 2000. Fission yeast on the brink of meiosis. *BioEssays* 22.9:854-860.
- Forsburg, S.L. 1999. The best yeast. *Trends Genet.* 15:340-344.
- Forsburg, S.L. 2001. The art and design of genetic screens: Yeast. *Nat. Rev. Genet.* 2:659-668.
- Gould, K.L. (ed.) 2003. Methods for *Schizosaccharomyces pombe*. *Methods* Vol In Press.
- Hedges, S.B. 2002. The origin and evolution of model organisms. *Nat. Rev. Genet.* 3: 838-849.

- Mitchison, J.M. 1990. The fission yeast, *Schizosaccharomyces pombe*. *BioEssays* 12:189-191.
- Moreno, S., Klar, A., and Nurse, P. 1991. Molecular genetic analysis of the fission yeast *Schizosaccharomyces pombe*. *Methods Enzymol.* 194:795-823.
- Nurse, P. 2000. A Long Twentieth Century of the Cell Cycle and Beyond. *Cell* 100:71-78.
- Russell, P.R. 1983. Evolutionary divergence of the messenger-RNA transcription initiation mechanism in yeast. *Nature* 301:167-169.
- Wood, V., Gwilliam, R., Rajandream, M., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., et al. 2002. The genome sequence of the eukaryote fission yeast *Schizosaccharomyces pombe*. *Nature* 415:871-880.
- Yanagida, M. 2002. The model unicellular eukaryote, *Schizosaccharomyces pombe*. *Genome Biol.* 3:2003.1-2003.4.

INTERNET RESOURCES

- http://www.sanger.ac.uk/Projects/S_pombe
The Sanger Center S. pombe project.
- <http://www.genedb.org/genedb/pombe/index.jsp>
The Sanger Center S. pombe GeneDB project.
- http://www.sanger.ac.uk/PostGenomics/S_pombe
The Sanger Center expression analysis page.
- <http://www.pombe.net>
The Forsburg Laboratory S. pombe pages.
- <http://pombe.biols.susx.ac.uk>
The FYSSION homepage.

Contributed by Susan L. Forsburg
The Salk Institute for Biological Studies
La Jolla, California

STRAIN MAINTENANCE AND GENERAL GUIDELINES

Growth conditions

The basic methods for yeast culture are the same for both *S. pombe* and *S. cerevisiae* (see UNIT 13.2). The generation time of *S. pombe* is somewhat longer than that of budding yeast, ranging from 2 to 5 hr depending upon the medium, strain, and temperature. The preferred permissive temperature is 32°C, although cells grow well at 30°C. The maximum permissive temperature is 36°C and the minimum is 17°C. Most temperature-sensitive mutants are viable at 25°C.

Determining cell count

As an approximate guide, an OD₆₀₀ of 1.0 equals $\sim 1.5 \times 10^7$ cells/ml, subject to the same caveats described in UNIT 13.2; however, bear in mind that the OD is a measure of overall cell mass, not cell number. A strain arrested by mutation in a cell cycle gene will continue to increase in OD because the cells continue to elongate, even though the cell number is not increasing. Under these conditions, a hemacytometer or Coulter counter should be used to measure cell number.

Diploids

Diploid strains should not be maintained for prolonged periods, but constructed fresh before use (see UNIT 13.16 on mating and mating type testing). The fission yeast diploid is partially induced for meiosis, and prolonged growth leads to selection for mutants that block the meiotic pathway. The exception is diploids that contain an otherwise lethal disruption. It is common for diploids awakened from the freezer to have a high frequency of sporulation-minus segregants; this can be examined by iodine staining, as described below. Similarly, *h*⁺ strains which can revert to *h*⁹⁰ should be tested to verify the mating type. Yeast nomenclature is detailed in Table 13.15.1.

Table 13.15.1 Yeast Nomenclature

Species	Wild type gene name	Recessive mutant	Protein	Disruption
<i>S. pombe</i>	<i>yfg1</i> ⁺	<i>yfg1-1</i>	Yfg1, Yfg1p	$\Delta yfg1::ura4^+$; $\Delta yfg1-D1$; <i>yfg1</i> Δ or $\Delta yfg1$
<i>S. cerevisiae</i>	<i>YFG1</i>	<i>yfg1-1</i>	Yfg1, Yfg1p	$\Delta yfg1::URA3$; <i>yfg1</i> Δ or $\Delta yfg1$

PRESERVING STRAINS BY FREEZING

Strains are typically stored at –70°C in glycerol. Stocks should be cultured to stationary phase (OD₆₀₀ = 1.5) in YES medium (see recipe), then mixed 1:1 either with sterile yellow freezing mix (50% v/v glycerol in YES) or 50% (v/v) sterile glycerol and immediately frozen. For strains that must be maintained under selection (e.g., to maintain a plasmid), the culture should be grown in EMM (see recipe) and frozen by addition of an equal volume of 50% (v/v) glycerol. Phenotypes of strains reawakened from the freezer should be verified by replica plating.

NOTE: Rather than sterile velvets, the authors recommend using two circles of Whatman #1 filter paper, which are sterile straight out of the box.

BASIC PROTOCOL

Contributed by Susan L. Forsburg

Current Protocols in Molecular Biology (2003) 13.15.1-13.15.5

Copyright © 2003 by John Wiley & Sons, Inc.

Yeast

13.15.1

Supplement 64

MEDIA

Fission yeast can be grown on synthetic minimal or rich media (see recipes), depending upon the needs of the experiment. Ideally, investigators should use media optimized for the organism, although *S. pombe* can be maintained on budding yeast media if necessary. Rich media made from yeast extract (YE) is usually supplemented to promote maximum growth and prevent adenine limitation. There are several variations of minimal media, although EMM (see recipe) is the most efficient at suppressing mating or sporulation. There are multiple choices for mating media, including malt extract or SPAS (see recipe). Most fission yeast investigators do not use “drop-out mix,” but add the individual supplements required for their strains; however, this is a result of habit rather than a requirement. Commercially available media mixes are also available from Qbiogene.

The same general rules apply for fission yeast media as for budding yeast media (*UNIT 13.1*), including sources for compounds, autoclaving, and general procedures. Media should not be autoclaved for more than 20 min to prevent caramelization of the glucose.

Rich Medium

YES

<i>Per liter:</i>	<i>Final concentration:</i>
5 g yeast extract	0.5% yeast extract
30 g glucose	3% glucose
225 mg adenine-HCl	1.31 mM adenine
225 mg L-histidine	1.45 mM L-histidine
225 mg L-leucine	1.71 mM L-leucine
225 mg uracil	2.01 mM uracil
225 mg L-lysine-HCl	1.23 mM L-lysine
20 g Difco Bacto Agar	2% agar (solid medium only)

*YES, or YE + supplements, is preferred for general growth when selection is not required; however, when selecting for the kan-MX marker using G418/geneticin resistance (Sigma G-5013), YES is the preferred media. The supplements are added to ensure maximum growth. If the supplements are not added, normal YE is functionally a low-adenine media. Note that YE or YES contains thiamine, so it represses expression from the commonly used *nmt1* (no message in thiamine) promoter.*

Minimal Medium

EMM

<i>Per liter:</i>	<i>Final concentration:</i>
3 g potassium hydrogen phthalate	14.7 mM $C_8H_5KO_4$
2.2 g dibasic sodium phosphate	15.5 mM Na_2HPO_4
5 g ammonium chloride	93.5 mM NH_4Cl
20 g glucose	2% glucose
20 ml 50× salt stock (see recipe)	1× salt
1 ml 1000× vitamin stock (see recipe)	1× vitamins
0.1 ml 10,000× mineral stock (see recipe)	1× minerals

When using Edinburgh minimal medium, required supplements for auxotrophies (e.g., adenine, uracil) are added immediately before use to a final concentration of 225 mg/liter as required. These can be maintained as sterile stock solutions at 7.5 mg/ml in water (3.75 mg/ml for uracil). Low adenine media, which allows the development of a red color in Ade^- strains, reduces the amount of adenine to 7.5 mg/liter. Some protocols use 10 or even 30 mg/liter adenine, and still see satisfactory red color develop.

The *nmt1* promoter, which is the most commonly used promoter for heterologous gene expression (Maundrell, 1990, 1993), is repressed by the addition of 15 μM thiamine (15 μM or 5 $\mu\text{g/ml}$; from filter sterilized stock at 10 mg/ml) immediately prior to use. The levels of expression can also be titrated using intermediate amounts of thiamine (0.05 μM or 0.016 $\mu\text{g/ml}$; Javerzat et al., 1996). SD media, used for budding yeast, contains thiamine and therefore cannot be used for fission yeast when *nmt1* expression is desired.

Variations of EMM media are used for particular purposes. For example, EMM lacking nitrogen (i.e., which leaves out the ammonium chloride; EMM -N) is used to arrest cells in the G1 phase of the cell cycle at 25°C. EMM in which the ammonium chloride is replaced by 1 g/liter L-glutamate is a low nitrogen medium that promotes sexual differentiation in lieu of mating media discussed below. Under conditions where slow growth is desired to allow full expression of a particular phenotype, ammonium chloride is often replaced with a poor nitrogen source, for example 3.8 g/liter glutamate (Ekwall et al., 1996) or 1.2 g/liter proline (Rhind and Russell, 1998). EMM with 0.5% glucose instead of 2% reportedly improves some transformation procedures (Okazaki et al., 1990). EMM that lacks both nitrogen and glucose (EMM -N -G) can be used to arrest cells in the G2 phase of the cell cycle and is also a suitable buffer to maintain spores following random spore analysis (UNIT 13.16). EMM + sorbitol plates are used for plating protoplasts. Add sorbitol to a final concentration of 1.2 M in standard EMM prior to autoclaving.

Stock Solutions

Prepare all stock solutions in water. Filter sterilize and store indefinitely at 4°C.

50× salt stock

<i>Per liter:</i>	<i>Final concentration:</i>
52.5 g magnesium chloride hexahydrate	0.26 M $\text{MgCl}_2 \cdot 6 \text{H}_2\text{O}$
0.735 g calcium chloride dihydrate	5.0 mM $\text{CaCl}_2 \cdot 2 \text{H}_2\text{O}$
50 g potassium chloride	0.67 M KCl
2 g disodium sulfate	4.1 mM Na_2SO_4

1000× vitamin stock

<i>Per liter:</i>	<i>Final concentration:</i>
1 g pantothenic acid	81.2 mM pantothenic acid
10 g nicotinic acid	81.2 mM nicotinic acid
10 g inositol	4.20 mM inositol
10 mg biotin	40.9 μM biotin

10,000× mineral stock

<i>Per liter:</i>	<i>Final concentration:</i>
5 g boric acid	80.9 mM boric acid
4 g magnesium sulfate	33.2 mM MnSO_4
4 g zinc sulfate heptahydrate	13.9 mM $\text{ZnSO}_4 \cdot 7 \text{H}_2\text{O}$
2 g ferric chloride hexahydrate	7.40 mM $\text{FeCl}_2 \cdot 6 \text{H}_2\text{O}$
0.4 g molybdic acid	0.32 mM molybdic acid
1 g potassium iodide	6.02 mM KI
0.4 g cupric sulfate pentahydrate	1.60 mM $\text{CuSO}_4 \cdot 5 \text{H}_2\text{O}$
10 g citric acid	47.6 mM citric acid

Mating Media

ME

<i>Per liter:</i>	<i>Final concentration:</i>
30 g Bacto-malt extract (ME)	3% (w/v) malt extract
30 g glucose	3% glucose
225 mg adenine-HCl	1.31 mM adenine
225 mg L-histidine	1.45 mM L-histidine
225 mg L-leucine	1.71 mM L-leucine
225 mg uracil	2.01 mM uracil
20 g Difco Bacto agar	2% agar (for solid medium only)
Adjust to pH 5.5 with NaOH	

ME is the most common mating media, but may show batch-to-batch variation. More defined media can be used instead, such as SPAS, EMM -N, or EMM with 1 g/liter L-glutamate instead of NH_4Cl .

SPAS

<i>Per liter:</i>	<i>Final concentration:</i>
10 g glucose	1% (w/v) glucose
1 g potassium phosphate monobasic	7.3 mM KH_2PO_4
1 ml 1000 \times vitamin stock	1 \times vitamins
45 mg adenine-HCl	0.26 mM adenine (1/5 normal)
45 mg histidine	0.29 mM histidine (1/5 normal)
45 mg leucine	0.34 mM leucine (1/5 normal)
45 mg uracil	0.40 mM uracil (1/5 normal)
45 mg lysine-HCl	0.25 mM lysine (1/5 normal)
30 g Difco Bacto agar	3% agar (for solid medium only)

Phloxin B

This vital stain is commonly used to determine ploidy and temperature sensitivity of *S. pombe*. It colors solid media vivid pink and the colonies will range in color according to the health of the cells. However, because it is mildly toxic, it should only be used for testing and not for long-term storage. Cells on phloxin B media die rapidly if the plates are refrigerated. Wild-type haploids form pale pink colonies on phloxin B. Diploid colonies are slightly darker. Dying cells form darker colonies or patches, and lethal mutants are often dark red. The color is not apparent at the single cell level (any dead cells will stain vivid pink), but relies on macroscopic examination of colony color.

Phloxin B (Magdala Red; Sigma) is prepared at a final concentration of 5 mg/liter from a stock of 10 g/liter in water. It is then filter sterilized and kept indefinitely at room temperature. Because it is light sensitive, it should be stored in the dark (e.g., a foil-covered tube). Plates should be stored at room temperature, in the dark, for no more than a few weeks. If the plates start to dull or look brownish, they are past their usefulness.

LITERATURE CITED

- Ekwall, K., Nimmo, E.R., Javerzat, J.P., Borgstrom, B., Egel, R., Cranston, G., and Allshire, R. 1996. Mutations in the fission yeast silencing factors *clr4⁺* and *rik1⁺* disrupt the localisation of the chromo domain protein Swi6p and impair centromere function. *J. Cell. Sci.* 109:2637-2648.
- Javerzat, J.P., Cranston, G., and Allshire, R.C. 1996. Fission yeast genes which disrupt mitotic chromosome segregation when overexpressed. *Nucl. Acids Res.* 24:4676-4683.
- Maundrell, K. 1990. *nmt1* of fission yeast: A highly transcribed gene completely repressed by thiamine. *J Biol. Chem.* 265:10857-10864.

- Maundrell, K. 1993. Thiamine-repressible expression vectors prep and prip for fission yeast. *Gene* 123:127-130.
- Okazaki, K., Okazaki, N., Kume, K., Jinno, S., Tanaka, K., and Okayama, H. 1990. High-frequency transformation method and library transducing vectors for cloning mammalian cDNAs by trans-complementation of *Schizosaccharomyces pombe*. *Nucl. Acids Res.* 18:6485-6489.
- Rhind, N. and Russell, P. 1998. The *Schizosaccharomyces pombe* S-phase checkpoint differentiates between different types of DNA damage. *Genetics* 149:1729-1737.

Contributed by Susan L. Forsburg
Molecular and Cell Biology Laboratory
The Salk Institute for Biological Studies
La Jolla, California

Much of the difference in working with *S. pombe* compared to *S. cerevisiae* comes from two issues. First, *S. pombe* has a G2-oriented cell cycle, spending 70% of its lifecycle in G2—almost completely opposite to the case of *S. cerevisiae*. Second, a wild-type cell is usually haploid, because *S. pombe* cells mate and sporulate only in response to nutrient limitation in a continuous pathway from conjugation through meiosis. This contrasts with wild-type *S. cerevisiae* which will mate whenever an appropriate partner is present, but only sporulates under nutrient limitation. Thus, working with fission yeast diploids in the laboratory requires specific genetic tricks and planning ahead. Moreover, mating type pheromones cannot be used to arrest the cell cycle easily in *S. pombe*, so other methods must be employed. This unit addresses these distinct areas of working with crosses, tetrads, and diploids, as well as analyzing the cell cycle and the products of genetic analysis.

CROSSES, TETRADS, AND DIPLOIDS

Fission yeast has two mating types, h^+ and h^- , which are alleles of the *mat1* locus (*mat1-M* for h^- , *mat1-P* for h^+). Fully wild-type cells switch between h^+ and h^- mating types; these are called h^{90} (because 90% of cells can typically switch). In a culture of h^{90} cells, approximately half of the cells will be functionally h^+ and half h^- because of switching. Laboratory investigators typically use strains with rearrangements of the silent mating loci to provide a stable mating type and prevent switching. These heterothallic strains will only mate if a partner of the opposite mating type is provided to them. Since the silent loci are closely linked to the expressed *mat1* locus, it is convenient (but not quite accurate) to think of mating type in *S. pombe* as having three alleles: h^{90} , h^+ , and h^- . The common h^- is stable, but most laboratory h^+ strains revert to h^{90} at a low frequency ($\sim 10^{-3}$; for review, see Klar, 1992).

The *S. pombe* lifestyle is predominantly haploid, and in contrast to budding yeast, fission yeast haploids must be starved of nitrogen in order to undergo mating. Once they mate, they normally proceed immediately through sporulation, generating a curved, zygotic ascus (Fig. 13.16.1A). Diploids can be recovered in the laboratory by interrupting this process, and can be maintained vegetatively; however, they are prone to sporulate, and when they do, they form a linear, azygotic ascus that is easily distinguished from the zygotic ascus (Fig. 13.16.1A).

In this unit, mating type testing is described (see Basic Protocol 1), crossing strains for tetrad or random spore analysis is covered (see Basic Protocol 2), and instructions for working with diploids are given (see Basic Protocol 3).

TETRAD DISSECTION AND RANDOM SPORE ANALYSIS

Both tetrad dissection and random spore analysis (RSA) are used to construct new strains and for analysis of meiotic products. Because mating and meiosis are linked in *S. pombe*, it is generally not necessary to isolate a diploid prior to characterizing offspring. Instead, the haploid parents are crossed on mating plates and allowed to proceed all the way through sporulation. The only exceptions are generally when very efficient sporulation is required, or when one parent is homothallic (h^{90}), and would otherwise cross with itself.

The logic of tetrad dissection is exactly the same for *S. pombe* as *S. cerevisiae*, and the manipulations are similar. RSA in fission yeast is much simpler than for budding yeast

Contributed by Susan L. Forsburg

Current Protocols in Molecular Biology (2003) 13.16.1-13.16.17

Copyright © 2003 by John Wiley & Sons, Inc.

Yeast

13.16.1

Supplement 64

because vegetative cells are easily killed by the enzyme glusulase (snail gut enzyme), and the spores do not stick to one another, so they are readily dispersed. RSA provides certain advantages because it allows the analysis of a large population of offspring with relatively little labor, and is very commonly used. On the other hand, tetrad dissection uniquely allows the analysis of the products of an individual meiosis, and the unambiguous identification of double mutants. *S. pombe* tetrads are generally somewhat linear in shape, and therefore ordered, but it is seldom possible to dissect the spores in such a way as to maintain the order.

How does one choose which is most useful? For general strain construction in which the desired double mutant can be unambiguously identified, RSA is the method of choice. It is particularly useful if the two mutations are linked (e.g., *leu1* and *his7*), because a large number of offspring can be analyzed to identify a rare event. RSA can be used for linkage analysis in genetic screens, since large numbers of crosses can be processed easily, and it is commonly used to determine recombination frequency in large mapping studies. The only caveats are that all the genotypes must be equally viable and the desired recombinant must be unambiguously identified. In cases where there is some doubt about the genotype of a putative double mutant, the candidates can be backcrossed to each parent to verify linkage. Most common strain construction requirements are satisfied by RSA.

There are two special cases where RSA is particularly helpful. The first is in recovery of spores containing plasmids. Because fission yeast plasmids cannot encompass the very large fission yeast centromere (*UNIT 13.14*), plasmids are meiotically unstable and usually only 10% of spores will contain a plasmid marker. However, using RSA, even this low fraction is sufficient to isolate a plasmid-containing spore or to draw conclusions about plasmid complementation of deletion alleles in cases where plasmid shuffle is not possible.

The second special case is the analysis of lethal mutations in the population during a spore germination experiment. Diploid strains heterozygous for a marked allele are sporulated in batch culture of liquid mating medium to generate large populations of spores. Following digestion with the enzyme, the spores can be inoculated into selective medium so that only cells carrying a desired marker are able to germinate. The rest of the spores are inert so that the phenotype of the germinated mutants may be unambiguously determined.

Tetrad analysis provides an enormous amount of information even in ten tetrads. Linkage trends are usually visible in just a few tetrads (but meaningful numbers for weak linkage may take many more). Most importantly, tetrad analysis allows the unambiguous identification of double mutants and is essential if the phenotype of a double mutant is unknown or potentially lethal. Consider a cross between two viable mutants, $\Delta yfg1::ura4^+$ and $\Delta yfg1::ura4^+$. Since both are marked with *ura4*⁺ and are otherwise indistinguishable, the investigator relies upon isolation of the nonparental ditype tetrad (2 *ura4*⁺:2 *ura4*⁻ spores) to identify unambiguously the double mutant. In this tetrad, the two *ura4*⁺ spores must both be double mutants and their phenotypes and viability can be determined.

Wild-type fission yeast are efficient at germinating on YES plates. Germination is modestly reduced on EMM plates, so unless selection for specific markers is required, YES plates are preferred both for RSA and tetrad dissection.

In this unit, tetrad analysis is described (see Basic Protocol 4). Additional protocols detail RSA (see Alternate Protocol 1), nonsporulating diploids (see Alternate Protocol 2), and endoreduplication (see Alternate Protocol 3).

Cell Cycle Synchronization

In budding yeast, the mating pheromone α factor is sufficient to arrest cells in G1 phase. In wild-type fission yeast, mating pheromone has much weaker effects, and the primary G1 arrest signal is provided by nitrogen starvation. Thus, cells can be arrested in G1 by starving for nitrogen and then releasing by refeeding with complete medium. This leads to a reasonably synchronous cell cycle following a several hour delay (lag phase). Alternative methods for cell synchronization fall into two classes: those that separate small G2 cells on the basis of size by some form of centrifugation and those that rely on cell cycle block and release by starvation, drugs, or mutants. The latter are generally simpler to perform but, since they rely on a perturbation of the cell cycle, they may lead to confusing results. Any meaningful cell cycle variation in phenotype, protein, or RNA levels should be verified using cycling cells unperturbed by arrest.

In these protocols the authors describe methods used for physiological analysis, such as cell cycle dynamics, using small cultures. They can be adjusted for preparative studies of RNA or protein by increasing the culture volume and harvesting larger aliquots at each time point.

To determine cell cycle synchronization during nitrogen starvation, a time course is monitored by sampling cells at each time point and fixing the aliquot for microscopy (see Basic Protocol 5 and Support Protocol). Similar monitoring is used for other synchrony methods, including cell cycle block and release (see Alternate Protocol 4). Size selection of cells via lactose gradients is also described (see Basic Protocol 6). Use of centrifugal elutriation requires a specialized centrifuge and rotor that is not found in most labs; however, should such an instrument be available, a good description of the procedure is found in Alfa et al. (1993).

MATING TYPE TESTING

Unlike *S. cerevisiae*, mating type testing in *S. pombe* relies upon completion of sporulation rather than isolation of complementing diploids. This reflects the relative instability of diploids in the laboratory. Thus, the only markers that are relevant in the strains to be tested are their mating types. When fission yeast cells complete meiosis and sporulation, they produce starch that can be stained easily by exposure to iodine vapor; however, iodine vapor kills the cells, so this method cannot be used if further analysis of the meiotic products is required.

Materials

- Known h^+ and h^- tester strains
- Rich medium plates (e.g., YES plates; *UNIT 13.15*)
- Characterized (positive control) and uncharacterized strains
- Mating plates (ME, SPAS, or EMM -N; *UNIT 13.15*)
- Iodine crystals (Sigma)
- 25°C incubator
- Empty petri dish
- Additional equipment and reagents for replica plating (*UNITS 1.3 & 13.2*)

CAUTION: Iodine is toxic and the vapors from the crystals or the exposed tester plate should not be inhaled.

1. Patch known h^+ and h^- tester strains vertically on rich medium plates, and the uncharacterized strain on a separate plate horizontally (Fig. 13.16.1). Include a characterized strain as a positive control. Incubate at 32° (wild type) or 25°C (ts strains) until both patches are grown up (1 to 2 days).

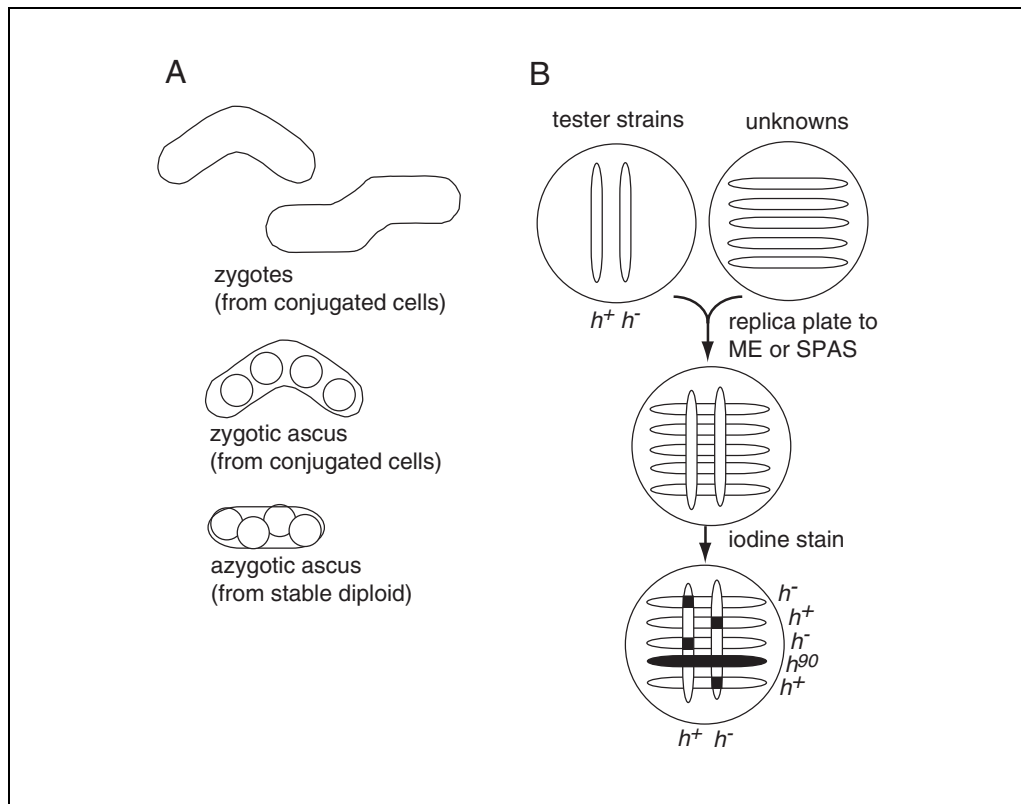


Figure 13.16.1 Fission yeast mating. (A) Zygotes and asci as seen under the microscope. (B) Mating type testing on plates.

2. Replica plate onto mating plates so that they intersect (UNITS 1.3 & 13.2).

Cells mate most efficiently when they are replica plated from freshly grown streaks on rich medium.

3. Incubate 3 to 4 days at 25° to 29°C.

Mating and sporulation are inherently temperature sensitive and should be carried out at 25°C (maximum of 29°C).

4. Prepare an iodine plate by scattering iodine crystals on the bottom or lid of an empty petri dish in a fume hood. Remove the lid of the tester plate and invert over the plate of crystals. Incubate this way until the positive control becomes dark brown where the streak intersects the tester of the opposite mating type (5 to 10 min). Do not allow the crystals to touch the agar or cells.

The plate with crystals can be stored in a sealed plastic container in the fume hood. The crystals will continue to sublime and will need to be refreshed regularly. If the tester plate is left too long, the color can be reduced by removing the plate from the crystals and letting it fade in the hood.

5. Determine the mating type by identifying the junctions in which a strain of known mating type generated a dark patch with the strain being tested (see Fig. 13.16.1B). h^{90} strains will mate with all testers and themselves, generating a continuously dark streak. The dark color does not last, so the results should be recorded immediately.

CROSSING STRAINS FOR TETRAD OR RANDOM SPORE ANALYSIS

Construction of new strains or analysis of meiotic products requires that the parent strains cross efficiently under conditions where the spores are viable. The previous protocol (see Basic Protocol 1) is adequate for assigning mating type, but mating via replica plates is relatively inefficient, and the iodine staining will kill the cells. This protocol provides more efficient mating by improving the mixing of the parent cells, and does not employ iodine. Successful mating can be determined microscopically by determining the formation of zygotes and asci (Fig. 13.16.1A).

Materials

- Strains to be crossed (i.e., strains 1 and 2) growing robustly on YES plates (EMM if selection is required; *UNIT 13.15*)
- Mating plates (ME or SPAS plates; *UNIT 13.15*)
- H₂O, sterile
- Sterile toothpicks
- 25° to 29°C incubator

For single matings

- 1a. Using a sterile toothpick, take a visible glob of cells from strain 1 and make a small patch on a mating plate (~5 mm in diameter). Using a second sterile toothpick, take an equivalent amount of strain 2 and add to the previous patch.
- 2a. Add 5 µl sterile water and mix strains gently on the surface of the agar with a sterile toothpick.

For multiple matings using the same strains

- 1b. Pick a generous amount into 100 µl sterile water.
- 2b. Pipet 5 µl of each onto a patch and use a sterile toothpick to mix each cross gently.
 - Strains growing in liquid medium can be mated, but they must be washed free of the medium or the nitrogen in the medium will suppress mating.*
3. Incubate at 25° to 29°C for 2 to 4 days (2 to 3 for tetrads, 3 to 4 for other assays). Monitor formation of asci microscopically.

Note that iodine staining (see Basic Protocol 1) is lethal and cannot be used if viable asci are required for tetrad dissection, RSA, or spore germination.

Mating and sporulation are inherently temperature sensitive and should be carried out at 25°C (maximum of 29°C). Sterile strains can be forcibly mated using protoplast fusion, a variation of the protoplast protocol used for transformation (see section on transformation in UNIT 13.17). Mating and sporulation can also occur in liquid culture. Equal numbers of both mating types, or previously isolated diploid cells, are washed free of normal growth medium and then inoculated into liquid mating medium. Cultures are incubated with gentle shaking at 25°C and monitored microscopically for the formation of zygotes and/or spores. Generally this is accompanied by substantial flocculation.

WORKING WITH DIPLOIDS

Although fission yeast diploids are unstable, they can be recovered by interrupting the mating process. Diploid strains must be selected by complementation of nutritional markers; however, after they are identified, they should be maintained on YES to suppress sporulation. Diploids are important for many assays and take only a little advance planning to be useful in the laboratory.

A particularly convenient pair of markers for diploid construction is provided by two complementing *ade6* alleles, *ade6-M210* and *ade6-M216*. When present in the same cell, these alleles complement intragenically, leading to an Ade⁺ phenotype. This also provides a color screen: the *ade6*⁺ gene is the ortholog of budding yeast *ADE2*, and mutants cause the accumulation of a pink or red color in the cell in the presence of low adenine (see *UNIT 13.15*). Ade⁺ strains are white. The two *ade6* alleles can be distinguished by their degree of color: *ade6-M210* is usually darker pink than *ade6-M216*. Neither is as dark as a null allele; however, their color can be affected by the strain background, so the marker is frequently misassigned. Assignment can be verified by attempting to make a diploid against a strain with a known *ade6* allele.

If the appropriate *ade6* markers are unavailable, a diploid can be constructed by complementation of any two auxotrophic markers. In this case, the investigator must use care to verify that the strain recovered is diploid, and not a recombinant haploid spore.

Materials

Strains to be crossed (complementing mating types and *ade6* markers) growing robustly on YES plates (EMM if required for selection; *UNIT 13.15*)

EMM plates lacking adenine (*UNIT 13.15*)

YES plates containing phloxin B (*UNIT 13.15*)

YES plates (*UNIT 13.15*)

Mating plates (ME or SPAS; *UNIT 13.15*)

Sterile toothpicks

25° or 32°C incubator

Additional reagents and equipment for mating and testing the mating type of *S. pombe* (see Basic Protocol 1)

1. Follow the steps for mating *S. pombe* (see Basic Protocol 1, steps 1 to 3), using strains with complementing mating types and *ade6* markers growing robustly on YES (or EMM) plates.
2. At 6 hr after mating and again at 12 hr, take a sterile toothpick and pick a swatch of cells. Streak to EMM medium lacking adenine (but containing any other required supplements).
3. Incubate at 32°C (25°C for temperature-sensitive diploids).
Incubation at 36°C, if the diploids are not temperature sensitive, will suppress sporulation.
4. Once clearly white colonies are visible (1 to 2 days), pick and streak several on YES plates containing phloxin B. Incubate 2 to 4 days at 32°C (25°C for temperature-sensitive mutants).

The cells in the middle of a colony on EMM may have sporulated; these will form pale pink haploid colonies when streaked on YES plates containing phloxin B. The darker pink colonies are still diploids. It is useful to streak a known haploid on YES containing phloxin B plates to verify its color. In addition, diploids are longer and wider than haploids when viewed under the microscope.

5. Pick several independent colonies and streak to YES plates. Patch the same colony in an ~5- to 10-mm spot on mating plates (ME or SPAS) and incubate 1 to 2 days at 25°C. Expose the mating plates to iodine vapor as described in mating type testing (see Basic Protocol 1) to verify clones are sporulation competent.

Diploids are stable for 1 to 2 weeks at 4°C if maintained on YES.

Prolonged maintenance selects for sporulation deficient mutants; therefore, diploids should be constructed fresh whenever possible.

TETRAD ANALYSIS

This method is almost identical to that in budding yeast and the apparatus is the same (see *UNIT 13.2*). The significant difference is that fission yeast asci do not require prior digestion with glusulase: they fall apart by themselves. Thus, the investigator must identify a mating plate with ripe asci before they have started to disintegrate, and manipulate the intact asci out to the position as shown in Figure 13.2.1. The plate is incubated for a few hours at room temperature (or overnight at 17°C), during which time the ascus breaks down, and the investigator returns to the plate and completes the dissection of the now freed tetrads. This must be completed before the freed spores enter their first cell division. The disintegration of asci is retarded at low temperatures and accelerated at high temperatures.

In addition, fission yeast spores are not particularly sticky. Thus, a very high quality needle with a completely flat plane is essential, since the spores will not stick to the needle without surface tension. In addition to pulling needles oneself, precut fiber optic needles that fit standard tetrad microscopes are also available (e.g., Singer Instruments; <http://www.singerinstruments.co.uk>).

Materials

Mating/sporulation plate grown for 2 to 3 (haploid) or 1 to 2 (diploid) days
YES plates (*UNIT 13.15*)

Sterile toothpicks
Tetrad dissecting microscope
36°C incubator (optional)
17°C incubator (optional)
25° or 32°C incubator

Additional reagents and equipment for preparation and dissection of tetrads (*UNIT 13.2*)

NOTE: Azygotic asci from previously isolated diploids form more quickly than zygotoc asci from haploids, which must mate first.

1. Under a microscope, examine a mating/sporulation plate which has been grown for 2 to 3 (zygotoc) or 1 to 2 (azygotoc) days at 25° to 29°C to identify ripe asci, which are intact but have clearly distinguished spores. If necessary, spread cells in the mating patch along the agar with a sterile toothpick to make them more apparent.
The mating/sporulation plate can be left in the refrigerator for a day or two to retard ascus breakdown if the investigator cannot immediately dissect the tetrads.
2. Transfer and lay cells across the YES plate upon which dissection will be performed using a sterile toothpick.
3. Upon identification of a ripe intact ascus, use a dissecting needle to manipulate the ascus to a fixed position (a of Fig. 13.2.1) on the plate (see protocol for preparation and dissection of tetrads, *UNIT 13.2*). Place each additional ascus 0.5 cm to one side of the preceding tetrad.
- 4a. *For tetrads identified in the morning:* Incubate the plate a few hours at warmer temperatures (i.e., 32° to 36°C) to stimulate ascus breakdown. If the strain is not temperature sensitive, an hour or two at 36°C can noticeably accelerate the process.
- 4b. *For tetrads identified at the end of the day:* Incubate the plate overnight at room temperature or 17°C.

5. Upon returning the plate to the micromanipulator, identify tetrads that have popped by the four large, round, free spores and sometimes the faint skin of an ascus. Move them down the plate to the b, c, and d positions (Fig. 13.2.1). Repeat steps 3 and 4 throughout the day as needed.
6. Once the plate is completed, incubate it at the appropriate temperature (32°C for wild type, 25°C for temperature-sensitive mutants).

Germination occurs within a few hours and it generally takes 2 to 5 days to observe colonies, depending upon the temperature of incubation.

RANDOM SPORE ANALYSIS

Fission yeast spores do not adhere to one another outside of the ascus; rather, the ascus falls apart over time and the remaining vegetative diploids in a culture are readily killed by glusulase. Thus, random spore analysis is significantly easier in *S. pombe* than in budding yeast. It is the method of choice for many strain constructions, as long as the recombinant strains are viable and easily distinguished from the parents. When used in bulk liquid culture, it allows spore germination experiments to analyze large populations of marked, lethal alleles.

Materials

Mating/sporulation plate

H₂O, sterile

5% glusulase (DuPont NEN) in sterile water

PBS (APPENDIX 2; optional), or EMM –N –G (UNIT 13.15; optional)

YES plates (UNIT 13.15)

Sterile toothpick

Additional reagents and equipment for counting cells using a hemacytometer
(APPENDIX 3F)

1. Using a sterile toothpick, pick cells from a mating/sporulation plate into 100 µl sterile water. Add 10 µl of 5% glusulase and incubate overnight at room temperature.
2. Using a hemacytometer (APPENDIX 3F), count the number of spores in 10 µl of a 1:10 dilution in sterile water. Verify that no complete asci remain.

Typically all the spores in the central 5 × 5 grid area of an improved Neubauer hemacytometer are counted, and the number multiplied by 100 to give the number of spores per milliliter.

3. Dilute in sterile water, PBS, or EMM –N –G and plate 500 spores per plate (assuming all spores are viable). If selecting directly for spores containing particular markers, adjust this number accordingly. If selecting directly for spores containing a plasmid, plate 1000 to 10,000 spores per plate.

Do not use medium containing glucose for the dilutions, because it will induce germination. The dilutions can be stored at 4°C for several days in case replating is required (e.g., if the plates were plated too densely). Alternatively, spores in the original glusulase suspension can be washed free of the enzyme in EMM –N –G and stored at 4°C for many days in case further plating is required.

Colonies will appear in 2 to 5 days, depending on the temperature, and plates can be replica plated to assign markers.

The same procedure can be employed at a larger scale on cultures of sporulated cells in liquid medium, for a spore germination experiment. Following digestion (generally overnight or longer), the spores are washed several times before inoculation into selective medium. Only the fraction of the population containing the selected marker(s) will germinate. Any phenotype associated with the marker(s) can thus unambiguously be distinguished from the inert background of wild-type spores.

NONSPORULATING DIPLOIDS

Given the instability of normal h^+/h^- diploids, it can be useful to work with nonsporulating diploids. Typically, these are either homozygous at the mating locus, or contain a mutation in the h^+ information that prevents sporulation (Egel and Egel-Mitani, 1974; Willer et al., 1995). These can sporulate by providing the missing mating information or by inducing meiosis independent of mating type signaling by inactivating *pat1* kinase. They can also be induced to haploidize using drugs promoting chromosome loss (e.g., Bodi et al., 1991). These strains have specialized uses including mapping genetic loci to chromosomes, assessing the effect of ploidy on mutant phenotypes, or testing rates of chromosome loss.

ALTERNATE
PROTOCOL 2

ENDOREDUPPLICATION

At a low frequency, wild-type fission yeast cells will occasionally skip mitosis and produce a homozygous diploid, a process called endoreduplication. Because such diploids are homozygous at the mating locus, they are incapable of sporulation; however, they are still larger and darker on phloxin B. They can be isolated by screening cells plated on YES containing phloxin B for darker pink colonies, which arise at a frequency of $\sim 10^{-3}$.

ALTERNATE
PROTOCOL 3

Materials

YES plates containing phloxin B (UNIT 13.15)
Liquid culture of strain of interest

1. Plate 500 to 1000 cells per plate on YES plates containing phloxin B.
2. Incubate at 25° or 32°C until colonies form and color develops (generally 2 to 5 days).
3. Pick darker pink colonies and restreak on YES plates containing phloxin B plates. Verify diploid morphology by microscopy and ploidy by flow cytometry.

CELL-CYCLE SYNCHRONIZATION BY NITROGEN STARVATION

At low temperatures (25°C), *S. pombe* cells placed in EMM without nitrogen or supplements will arrest predominantly in G1. This is a temperature-sensitive effect, because at higher temperatures, a larger fraction of the cells will arrest in G2 (Costello et al., 1986). Auxotrophic markers such as *ade6* may retard this arrest slightly, which can be accommodated by arresting the cells in EMM without nitrogen supplemented with 45 mg/liter adenine (1/5 normal). Cells released from nitrogen by refeeding with regular EMM will enter the cell cycle after ~ 2 hr delay (the so-called “lag phase”). Nitrogen starvation cannot be used to synchronize h^{90} or diploid cells, because absence of nitrogen triggers the sexual differentiation process.

BASIC
PROTOCOL 5

Materials

Cells from the strain of interest
EMM and EMM –N (UNIT 13.15)
Complete medium (e.g., EMM with supplements, YES; UNIT 13.15)
25°C incubator
Incubator set at growth temperature (i.e., 32° to 36°C)

1. Grow cells from the strain of interest to mid-late exponential phase (OD = 0.6 to 1.0) in EMM with nitrogen.

The amount of culture will depend upon the purpose of the experiment—i.e., whether it is physiological analysis, in which case 50 ml will do, or synchrony to recover material for protein or RNA analysis, in which case 300 to 500 ml are more appropriate.

Yeast

13.16.9

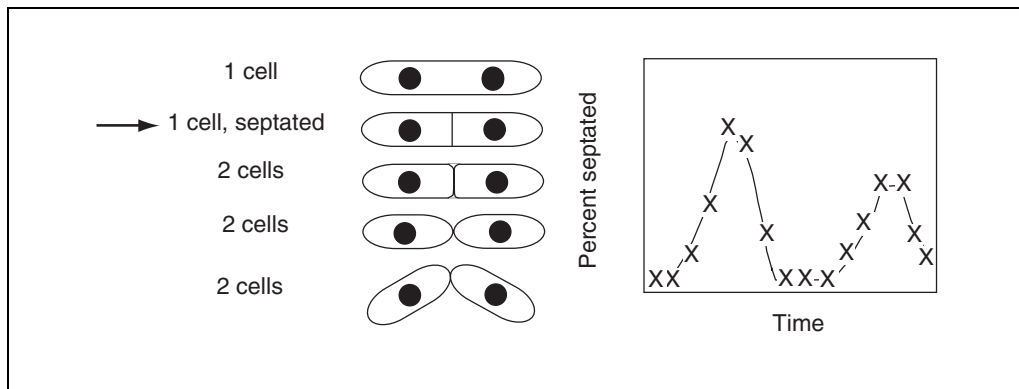


Figure 13.16.2 Determining septation index for synchronous cultures.

2. Harvest by centrifuging 5 to 10 min at $4000 \times g$, 20°C . Wash twice with an equal volume starvation medium.
3. Resuspend in an equal volume starvation medium (i.e., EMM –N) and incubate 12 hr at 25°C .
4. Add an equal volume prewarmed complete medium, preferably EMM containing any required supplements. Incubate ~ 2 hr at the appropriate growth temperature.

The temperature of the re-fed culture depends upon the purpose of the experiment. For a normal wild-type strain, the cells should be returned to 32°C . For a temperature-sensitive strain under restrictive conditions, the cells should be incubated at 36°C . It will take ~ 2 hr for cells released from nitrogen starvation in G1 to enter S phase.

YES can also be used to re-feed but, because it is a less defined medium subject to batch variations, this is not preferred.

5. Every 20 to 30 min, take an aliquot of cells ($\sim 5 \mu\text{l}$) and place on a slide with a cover slip. Examine under a phase microscope (dark-field illumination may help) and count the percent septated cells (septation index), which corresponds to G1/S phase cells (Fig. 13.16.2). Plot as a function of time.

*The degree of synchrony is determined by the maximum peak of septation; generally 30% for this method and up to 80% for the following method. This is ideally performed on live cells to monitor the time course in real time. Because of the structure of the *S. pombe* cell cycle (Fig. 13.14.1), the peak of septation in wild-type strains is roughly the same as cells in S phase.*

Aliquots may also be fixed in ethanol for FACS or microscopy and the septation index counted from fixed cells if this is more convenient (see Support Protocol). The key to using the septation index is to clearly distinguish cells that have begun division from those that have not (Fig. 13.16.2). Consistency is essential so that in a given experiment, it is best if a single investigator counts all the cultures. Preparative analysis requires larger aliquots; cells can be harvested by centrifugation, the pellets frozen on dry ice, and stored for weeks or longer at -70°C for subsequent analysis.

SUPPORT PROTOCOL

ETHANOL FIXATION

Samples fixed in 70% ethanol may be stored indefinitely at 4°C . This is particularly useful if FACS analysis or DAPI staining will also be performed, as the fixation is identical.

Materials

- Exponentially growing culture
- 70% ethanol, 4°C
- H_2O , sterile
- Mounting medium with stain (see recipe; optional)
- Positively charged microscope slides

1. Harvest 10^7 cells from an exponentially growing culture by centrifuging 5 min at $4000 \times g$, 20°C . Pour off the supernatant.
2. Vortex the tube while adding 1.0 ml cold 70% ethanol. Store at 4°C .
3. To rehydrate the cells, take 0.1 ml and add to 1 ml sterile water in a microcentrifuge tube. Mix and microcentrifuge 5 min. Decant the supernatant and resuspend in 0.1 ml sterile water. If desired, process for septation index as described (see Basic Protocol 5).
4. *Optional:* To visualize the nucleus and/or septa, pipet 5 μl cell suspension onto a cleaned microscope slide. Allow to air dry or place on a heating plate at low temperature until the liquid is just evaporated.

To ensure a monolayer of cells, pretreat the slides with poly-L-lysine (Sigma) or use a positively charged slide.

5. *Optional:* Add 5 μl mounting medium with stain and immediately cover with a cover slip. Examine using epifluorescence.

Samples can be stained with DAPI to visualize the nucleus, with Calcofluor to visualize the septa, or processed for FACS analysis.

CELL CYCLE BLOCK AND RELEASE

Alternative methods of cell cycle synchronization rely upon reversible temperature-sensitive cell cycle mutants, with the *cdc25-22* mutant being the most commonly used. It arrests cells at the G2 phase of the cell cycle at 35.5°C (Russell and Nurse, 1986). Following return to permissive temperature, *cdc25* strains rapidly enter mitosis and proceed through a synchronous cell cycle. The second cell cycle is significantly less synchronous.

The key to temperature arrest is rapid increase and decrease of temperature. All water baths should be prewarmed to the correct temperature. The volume of culture should be low for the size of the flask, to ensure most efficient temperature exchange.

Materials

Temperature-sensitive cell cycle mutants (e.g., *cdc25-22*; ATCC# 90337)
 25° and 36°C water baths
 Thermometer cleaned with ethanol

1. Grow temperature-sensitive cell cycle mutants to early/mid-log phase ($\text{OD}_{600} = 0.2$ to 0.4) at permissive temperature. If necessary, remove an aliquot for an exponentially growing cell time point, which can be processed for the same morphological or molecular tests as the synchronized cells.

The volume taken will depend upon the assay: <1 ml for cell morphology or FACS, >10 ml for protein or RNA preparation.

2. Incubate 4 to 6 hr at 36°C .

A rapid temperature shift is important. Water baths promote the most efficient thermal transfer. If a 36°C water bath is not available, swirl the flask in a 50°C bath or sink until the culture approaches 36°C as described in step 3, then transfer to an air shaker.

3. Take an aliquot for a $t = 0$ time point. Release cells to 25°C by rapidly swirling the flask in ice water and monitoring with a thermometer that has been cleaned with ethanol. Place in 25°C water bath.

ALTERNATE PROTOCOL 4

Yeast

13.16.11

4. To follow synchrony, sample time points every 20 to 30 min for septation index (correlates with S phase). This should approach 60% to 80% for an effective temperature-driven block and release.

Samples can be fixed for simple microscopy, or pellets can be harvested and frozen for RNA or protein analysis. If complex manipulations are required, it is often easier to have two people taking time points.

Because of their pronounced cell elongation, cdc25 mutants often give ambiguous FACS profiles following release to permissive temperature.

LACTOSE GRADIENT CENTRIFUGATION

Because fission yeast cells are very regular in size, size selection to isolate newborn G2 cells is an effective means to synchronize a population. Centrifugal elutriation is a frequent method in fission yeast, but requires an expensive and highly specialized centrifuge rotor, and working with more than one strain at a time is difficult. Lactose gradient centrifugation (e.g., Edwards and Carr, 1997) offers a convenient alternative that can be performed in most laboratories on multiple strains. Both these methods rely on separating cells based on cell size, so the strains must not show elongation or variation in cell morphology. Mutants that have heterogeneous phenotypes or cause elongation are much more difficult to synchronize using size selection methods.

Materials

Cells
70% ethanol
YES (*UNIT 13.15*) with 0%, 10%, and 40% lactose, 32°C
Gradient maker
20- and 50-ml conical tubes

Prepare culture

1. Grow a 125-ml culture overnight so that it is in mid- to late-exponential phase the next day ($OD_{600} = 0.5$ to 0.8).

Prepare gradient

2. Sterilize a gradient maker by passing 70% ethanol through the entire apparatus, being sure to loosen the screw so that both compartments empty into the tube. Place a tiny stir bar in the compartment closest to the exit tube (Fig. 13.16.3), put the gradient maker on a stir plate, and let the solutions drain down into a collection/waste beaker. After the ethanol has run through the gradient maker, rinse with YES containing no lactose.
3. Close the valve between the compartments. Load YES with 10% lactose in the chamber away from the exit tube. Open the central valve briefly to clear any bubbles. Load YES with 40% lactose in the compartment with the stir bar, proximal to the exit tube. Again, clear bubbles. Turn on the stir plate so that the stir-bar in the 40% compartment is mixing (Fig. 13.16.3).
4. Pour 5 ml YES with 40% lactose into a sterile 50-ml conical tube. Open both valves of the gradient maker, allowing the gradient to flow down the sides of the conical tube. Fill to 45 ml.

Once poured, the gradient is stable for an hour. Additional gradients can be poured if required to allow processing of multiple strains.

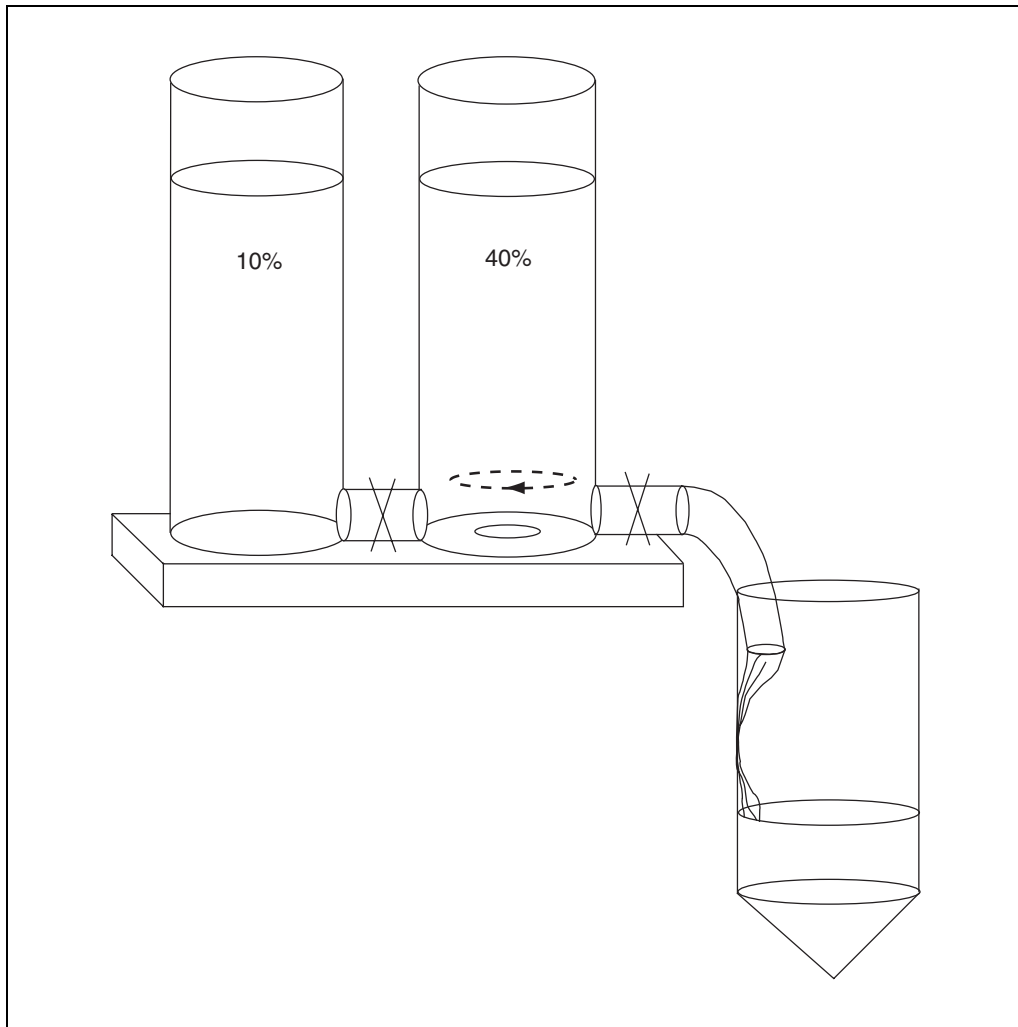


Figure 13.16.3 Assembling a lactose gradient.

Load culture onto gradient

5. Harvest cells by centrifuging at $3000 \times g$ (~5000 rpm), at 20°C until pelleted (5 to 10 min).
6. Resuspend cells in 0.5 ml YES with no lactose and pipet along the inside of the conical tube just above the gradient solution.
7. Centrifuge the conical tube in a swinging bucket centrifuge 5 min at $228 \times g$ (1000 rpm), room temperature.
8. Collect 5 ml from the top of the gradient in 1-ml aliquots using a pipettor. Examine the cells in each aliquot using the microscope. Pool all aliquots.

When removing the aliquots, the tip of the pipet should barely touch the top of the gradient.

The goal is a uniform population of small G2 cells; no longer cells nor any septated cells should be seen.

9. Spin down the aliquot of cells, wash in YES medium with 0% lactose, and resuspend in 10 ml YES medium. Follow cell progression by septation and FACS.

A maximum septation index of ~40% is expected.
10. *Optional:* If the above separation does not provide sufficient synchrony, or sufficient cell number, repeat the procedure using a larger starting culture (500 ml).

11. *Optional*: Further purify cells as follows:

- a. Add 0.5 ml of 40% lactose in YES to the bottom of a 20-ml conical tube.
- b. Add 10 ml YES/lactose from the gradient maker (step 3) on top of this initial addition.
- c. Layer the 5 ml harvested from the first gradient on the second, and centrifuge 3 min at $228 \times g$ (1000 rpm), room temperature.
- d. Recover six 0.5-ml aliquots, wash into fresh medium, and follow as above.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Calcofluor white, 1 mg/ml

Prepare Calcofluor white (fluorescent brightener 28; Sigma) at a final concentration of 1 mg/ml in 50 mM sodium citrate/100 mM sodium phosphate, pH 6.0 (APPENDIX 2). Store for months in the dark at 4°C.

DAPI stock

Dissolve 1 mg DAPI per milliliter DMSO. Store weeks to months in the dark at -20°C. Discard if the solution turns dark brown.

Mounting medium stock

Dissolve the following in 5 ml PBS (APPENDIX 2):

50 mg *n*-propyl gallate (Sigma)

50 mg *p*-phenylenediamine (PPD; Sigma)

Add glycerol to 50 ml

Aliquot and store for months at -20°C in the dark

Discard if solution darkens

Mounting medium with stain

Staining can be performed with DAPI, Calcofluor white, or both, using the following solutions.

DAPI alone

Dilute 1 μ l of DAPI stock (i.e., 1 mg DAPI/1 ml DMSO; see recipe) per 1.5 ml mounting medium stock (see recipe). Store 1 to 2 days at -20°C in the dark. Discard if the solution turns brown.

Calcofluor white alone

Dilute 1 mg/ml Calcofluor white to 50 μ g/ml in mounting medium stock (see recipe) prior to use. Prepare fresh.

Calcofluor white and DAPI

Prepare a solution of mounting medium stock (see recipe) containing both stains as described for the individual components. Prepare fresh. Note that the concentration of Calcofluor white may need adjustment to prevent overcoming the DAPI signal.

COMMENTARY

Background information

Several important biological differences between *S. pombe* and *S. cerevisiae* are particularly worth noting. First, the fission yeast is generally haploid, and can only be maintained

as a diploid with complementing markers or other genetic tricks. Thus, linkage analysis is often more efficient than complementation to assess whether two mutations affect the same gene. Second, because mating and meiosis are

coupled, it is not necessary to isolate a stable diploid prior to sporulating cells and analyzing meiotic products: haploid strains are crossed and allowed to proceed through meiosis in one process. Third, the ease of random spore analysis in fission yeast makes isolation of even rare meiotic products straightforward and, moreover, facilitates bulk sporulation and germination in liquid culture. The methods described here are sufficient for the basic genetic manipulation of fission yeast in any laboratory.

Individual labs may use slightly different protocols or medium, but in practical terms this makes little difference in assessing mating type, isolating diploids, or analyzing meiotic products by tetrads, RSA, or spore germination. For additional background on the detailed biology of mating type and meiosis, see Egel (2000), Klar (1992), and Yamamoto (1996).

There is considerably more discussion about the relative uses of different methods of cell cycle synchronization (for general reviews of cell cycle control, see MacNeill and Nurse, 1997; Moser and Russell, 2000). Methods that arrest and then release the cell cycle, such as nitrogen starvation (see Basic Protocol 5), cell cycle block and release (see Alternate Protocol 4), or drug treatments (such as hydroxyurea, which arrests early in S phase) allow very high levels of synchronization—i.e., a septation index approaching 65% or more. However, in all cases, there is substantial perturbation of the cell cycle: just because some events are blocked during the arrest, does not mean that all events are blocked. For example, *cdc25* mutants arrested in G2 phase continue to accumulate cyclin to higher than normal levels (Moreno et al., 1989). This is an important caveat that may influence downstream interpretation of the results. Generally, the use of two independent methods of synchronization is recommended to verify that the results reflect normal cell cycle variations, rather than an effect of arrest.

Not all mutants release as well as *cdc25*, but well-behaved alternatives include *cdc10-M17* (temperature sensitive, blocks at G1; Verkade and O'Connell, 1998) and *nda3-KM311* (cold sensitive; blocks in mitosis; Umesono et al., 1983). These may be employed similarly to *cdc25*, although the temperatures are reversed for *nda3*. For conditions where simple cell cycle arrest is sufficient without release, *cdc10*, *cdc25*, and *nda3* mutants may be compared with temperature-sensitive mutants affecting other parts of the cell cycle, including *cdc22* (ribonucleotide reductase; blocks in early S phase; Fernandez Sarabia et al., 1993), *cdc17*

(DNA ligase, blocks late in S phase; Nasmyth, 1977), or *nuc2* (APC component, arrests in mitosis/G1; Hirano et al., 1988).

Size-selection methods of synchrony such as lactose gradients (see Basic Protocol 6) or elutriation (Alfa et al., 1993) do not overtly perturb the cell cycle, but are less synchronous than the arrest-release approaches. The maximum septation index is ~35% to 40%, and the peaks of septation are much broader. As described in *UNIT 13.14*, newly separated fission yeast cells are in G2; therefore, methods that preferentially isolate small cells will enrich the G2 population, the basis for these procedures. The cells may suffer some modest stress from the low-speed centrifugation, but under conditions where temperature-sensitive mutants cannot be employed, or where normal cell cycle responses are particularly important, these are the methods of choice.

The reader is referred elsewhere for fundamentals of molecular analysis (Moreno et al., 1991) and other methods in fission yeast (Alfa et al., 1993; Gould, 2003).

Critical Parameters and Troubleshooting

Mating

The mating type testing protocol (see Basic Protocol 1) is often inefficient because it relies upon relatively low numbers of cells transferred by replica plating. Strains that grow slowly, that have impaired starvation response, or that have defects in mating may not generate sufficient numbers of asci to be observed by iodine staining. Under these conditions, crossing strains for tetrad or random spore analysis (see Basic Protocol 2) can be used and the plates subjected to iodine staining, or monitored for zygote formation by microscopy. If the latter, a sterile toothpick is used to pull the cells out from the mating patch and spread them out on the adjacent area of the plate, or apply them to a microscope slide in 5 μ l sterile water. Zygotes and asci as in Figure 13.16.1A can be observed under these conditions. It is generally a good idea to set up crosses for mating type testing, tetrads, or diploid isolation with a known pair of wild-type strains to confirm that the medium and conditions are suitable for analysis.

Tetrads and RSA

It is a good idea to practice both tetrads and RSA on well-behaved wild-type strains before performing them on mutants of interest. In addition, prior to replica plating, the micro-

scope should be employed, especially on tetrad plates where there are “missing” colonies, to verify that there were actually spores manipulated in the gaps. This also provides evidence of the phenotype prior to replica plating, when single spores or dead cells can be lost.

Diploids

The most common reason for failure to recover a diploid is misassignment of the mating type marker, or the *ade6* marker. These can be distinguished by observing the original mating plate to see whether zygotes and asci formed. If they do, then the likely culprit is the *ade6* allele, which can be tested by setting up crosses with a control strain with the opposite marker.

Stable diploids

Isolation of endoreduplicated strains to create stable homozygous diploids requires a method to verify ploidy in addition to phloxin B staining, because there is frequently a background of more darkly staining colonies that are still haploid in the population. FACS is unambiguous if a known haploid and diploid are used as controls. Isolation of a haploid from the stable homozygous diploid relies upon chromosome loss. If fission yeast diploids lose one chromosome, they will rapidly haploidize and this can be accelerated with certain drugs (e.g., Bodi et al., 1991).

An alternative method to create stable diploids relies upon crossing two strains with complementing markers which are able to mate but not sporulate. The *mat1-P* mating type locus produces two transcripts, one required for mating, and the second for meiosis (Egel and Egel-Mitani, 1974; Willer et al., 1995). A homothallic strain containing the *mat2-102* allele lacks the sporulation-specific P transcript. By itself on mating medium it makes zygotes but no spores. If mated as an h^+ to a wild-type h^- and selected for complementing markers, the resulting diploid is unable to enter meiosis. In contrast, the *mat2-102* strain functions as a normal h^- , so if mated to a wild-type h^+ , the resulting diploid is sporulation-competent. In this way, the *mat2-102* allele can be moved into different backgrounds.

Synchronization

The success of synchronization procedures is heavily influenced by the strain background. Not all mutants will arrest appropriately in G1 phase when starved for nitrogen. Similarly, many mutants cause a heterogeneous size phenotype that precludes use of size-selection

methods. These difficulties are only established by trial and error. The experiment should be performed on small-scale cultures, monitoring septation index and FACS if appropriate, to see if conditions are suitable before scaling up to preparative levels. Often it helps to divide the work between two people (one to harvest and fix cells, one to count septa).

Anticipated Results

Crosses, diploids, and tetrads are the fundamental components of classical genetic analysis. The methods are straightforward and allow determination of synthetic phenotypes or construction of useful strains. Synchronization methods are important because of *S. pombe*'s continuing importance as a model system for cell cycle analysis.

Time Considerations

For cell growth experiments, the limiting factor is growth rate of the individual strain. This varies substantially depending upon the genotype, medium, and temperature. Generation of recombinant strains takes about a week from crossing the haploid parents, to replica plating the germinated spores. Isolation and purification of a diploid strain generally takes about a week. Synchronous cultures generally take all day to perform, not including time for preparative analysis.

Literature Cited

- Alfa, C., Fantes, P., Hyams, J., McLeod, M., and Warbrick, E. 1993. Experiments With Fission Yeast. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Bodi, Z., Gysler-Junker, A., and Kohli, J. 1991. A quantitative assay to measure chromosome stability in *Schizosaccharomyces pombe*. *Mol. Gen. Genet.* 229:77-80.
- Costello, G., Rodgers, L., and Beach, D. 1986. Fission yeast enters the stationary phase G₀ state from either mitotic G₁ or G₂. *Curr. Genet.* 11:119-125.
- Edwards, R.J. and Carr, A.M. 1997. Analysis of radiation-sensitive mutants of fission yeast. *Methods Enzymol.* 283:471-493.
- Egel, R. 2000. Fission yeast on the brink of meiosis. *BioEssays* 22.9:854-860.
- Egel, R. and Egel-Mitani, M. 1974. Premeiotic DNA synthesis in fission yeast. *Exp. Cell. Res.* 88:127-134.
- Fernandez Sarabia, M.-J., McInerney, C., Harris, P., Gordon, C., and Fantes, P. 1993. The cell cycle genes *cdc22⁺* and *suc22⁺* of the fission yeast *Schizosaccharomyces pombe* encode the large and small subunits of ribonucleotide reductase. *Mol. Gen. Genet.* 238:241-251.

- Gould, K.L. (ed.) 2003. Methods for *Schizosaccharomyces pombe*. *Methods* Vol In Press.
- Hirano, T., Hiraoka, Y., and Yanagida, M. 1988. A temperature-sensitive mutation of the *Schizosaccharomyces pombe* gene *nuc2⁺* that encodes a nuclear scaffold-like protein blocks spindle elongation in mitotic anaphase. *J. Cell Biol.* 106:1171-1183.
- Klar, A.J.S. 1992. Molecular genetics of fission yeast cell type: Mating type and mating type interconversion. In *The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene Expression* (E. Jones, J. Pringle, and J. Broach, eds.) pp. 745-777. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- MacNeill, S.A. and Nurse, P. 1997. Cell cycle control in fission yeast. In *The Molecular and Cellular Biology of the Yeast Saccharomyces: Cell Cycle and Cell Biology* (J. Pringle, J. Broach, and E. W. Jones, eds.) pp. 697-763. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Moreno, S., Hayes, J., and Nurse, P. 1989. Regulation of p34^{*cdc2*} protein kinase during mitosis. *Cell* 58:361-372.
- Moreno, S., Klar, A., and Nurse, P. 1991. Molecular genetic analysis of the fission yeast *Schizosaccharomyces pombe*. *Meth. Enzymol.* 194:795-823.
- Moser, B.A. and Russell, P. 2000. Cell cycle regulation in *Schizosaccharomyces pombe*. *Curr. Opin. in Microbiol.* 3:631-636.
- Nasmyth, K. 1977. Temperature-sensitive lethal mutants in the structural gene for DNA ligase in the yeast *Schizosaccharomyces pombe*. *Cell* 12:1109-1120.
- Russell, P. and Nurse, P. 1986. *cdc25⁺* functions as an inducer in the mitotic control of fission yeast. *Cell* 45:145-153.
- Umesono, K., Toda, T., Hayashi, S., and Yanagida, M. 1983. Two cell division cycle genes NDA2 and NDA3 of the fission yeast *Schizosaccharomyces pombe* control microtubular organization and sensitivity to anti-mitotic benzimidazole compounds. *J. Molec. Biol.* 168:271-284.
- Verkade, H.M. and O'Connell, M.J. 1998. Cut5 is a component of the UV-responsive DNA damage checkpoint in fission yeast. *Mol. Gen. Genet.* 260:426-433.
- Willer, M., Hoffmann, L., Styrkarsdottir, U., Egel, R., Davey, J., and Nielsen, O. 1995. Two-step activation of meiosis by the *mat1* locus in *Schizosaccharomyces pombe*. *Mol Cell Biol* 15:4964-4670.
- Yamamoto, M. 1996. The molecular control mechanisms of meiosis in fission yeast. *Trends Biochem. Sci.* 21:18-22.

Contributed by Susan L. Forsburg
The Salk Institute for Biological Studies
La Jolla, California

Introduction of DNA into *S. pombe* Cells

UNIT 13.17

Transformation is the essential method by which DNA is introduced into a cell, whether it is a plasmid episome or an integration construct. Fission yeast is readily transformed using a variety of methods, but efficiencies, expressed as transformed colonies per microgram plasmid, are generally lower than *S. cerevisiae* (albeit still more than sufficient for plasmid recovery, screening of plasmid libraries, repair of gapped plasmids, and other standard yeast procedures). There are several possible explanations. First, the protocols have not been as carefully optimized. Second, maintenance of episomes may be affected by performance of the replication origin, and the absence of a centromere or other segregation method. Third, mutants affecting DNA dynamics may be harder to transform.

S. pombe is efficient at homologous integration of fragments of DNA, although it is also adept at nonhomologous recombination. For maximum efficiency of integration and minimum background, tracts of homology >300 base pairs are recommended; targeted integration occurs with smaller homology tracts (e.g., 50 base pairs), but generally have substantial background of nonhomologous events which must be screened out (Bähler et al., 1998; Grallert et al., 1993; Keeney and Boeke, 1994). Most plasmids are selected using nutritional markers on EMM plates. However, G418 resistance requires selection on YES plates.

S. pombe can be transformed by lithium acetate (see Basic Protocol and Alternate Protocol 1), electroporation (see Alternate Protocol 2), or protoplast treatment (see Alternate Protocol 3). In addition, protoplasts can be induced to fuse and undergo karyogamy, which provides a means of mating sterile strains (see Support Protocol).

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

TRANSFORMATION USING LITHIUM ACETATE

The principles of this method (Okazaki et al., 1990) are similar to *S. cerevisiae*, although maximum efficiency requires that the lithium acetate have a lower pH. Reported transformation frequency is up to 10^5 transformants per microgram DNA. This is a laborious but high-efficiency method; a more rapid version with somewhat reduced efficiency follows (see Alternate Protocol 1).

Materials

- YES medium (UNIT 13.15)
- S. pombe*
- MB medium (see recipe)
- 0.1 M lithium acetate (adjust to pH 4.9 with acetic acid)
- Plasmid DNA
- TE, pH 7.5 (APPENDIX 2)
- 50% (w/v) PEG 4000, 25° or 30°C
- 50% YE medium (see recipe for YES in UNIT 13.15)
- Selective EMM plates (UNITS 13.15)

- 25° or 32°C incubators
- 1-liter flask
- 43°C water bath

- Additional reagents and equipment for culturing yeast (UNIT 13.7)

NOTE: If necessary, a 30° rather than a 32°C incubator can be used.

**BASIC
PROTOCOL**

Contributed by Susan L. Forsburg

Current Protocols in Molecular Biology (2003) 13.17.1-13.17.8

Copyright © 2003 by John Wiley & Sons, Inc.

Yeast

13.17.1

Supplement 64

Prepare cells for transformation

1. Two days before the experiment, inoculate 5 ml YES medium with a single *S. pombe* colony of the strain to be transformed. Grow overnight to saturation at 32°C.

The saturated overnight culture may, if desired, be prepared up to several days in advance of the transformation and stored at 4°C. EMM (UNIT 13.15) should be used if selection is required.

2. The night before transformation, inoculate a 1-liter sterile flask containing 150 ml MB medium with an appropriate amount of the saturated culture and grow overnight at 32°C to an OD₆₀₀ of 0.5. Calculate the volume of starter culture from which to generate a larger overnight culture which is grown up with shaking.

$$\text{vol. inoculated (ml)} = \frac{(Y \times \text{OD}_f) / 2^n}{\text{OD}}$$

where Y is the volume of overnight culture required, n is the expected number of generations in this media during the expected grown period, OD_f refers to the desired OD for the exponential culture (generally 0.5), and OD refers to the OD₆₀₀ of the starter culture.

The expected number of generations when inoculating from a saturated starter culture must take into account a 1-generation lag phase to allow the cells to exit stationary phase and re-enter the cell cycle. For example, for a 12-hr growth period of a strain with a 3-hr doubling time, the total number of possible growth generations is $12/3 = 4$. If diluting from an already growing exponential culture, then $n = 4$. However, if inoculating from a stationary starter culture, there is one lag generation, so the actual growth generations = $4 - 1$, or 3.

The amount of culture to be inoculated depends upon the growth rate of the strain which is influenced by growth temperature, genotype, age of the starter culture, and growth media. A healthy wild-type strain in complete medium (YES or EMM) at 32°C has a generation time of between 2.5 and 3 hr during exponential growth. MB is a stringent minimal medium that has severely reduced growth rate compared to normal EMM; generation time will be ≥ 5 hr. It is not clear why this medium improves efficiency of transformation, but it may affect cell wall permeability.

Strains with particular auxotrophic markers or other mutations may be unable to grow in MB sufficiently well to generate a large working culture. In this case, cells may be grown to mid-exponential phase in EMM ($\text{OD}_{600} = 0.5$) and diluted into MB for two additional generations to gain some of the benefits of growth in MB. Low-glucose EMM can be used in place of MB, but transformation efficiency may be reduced. How much depends upon the individual strain.

3. Harvest the cells 5 min at $228 \times g$ (3000 rpm), room temperature. Wash cells in 40 ml sterile water and centrifuge again.

Transform with lithium acetate

4. Resuspend the cells at 1×10^9 cells/ml (0.5 to 1.5 ml) in 0.1 M lithium acetate, pH 4.9, and dispense 100- μ l aliquots into microcentrifuge tubes. Incubate 60 min at 32°C, or 25°C for temperature-sensitive (ts) mutants.

Cells will sediment at this stage. Cells can be left for up to 120 min without harm. Each tube is sufficient for a single transformation.

5. Add 1 μ g plasmid DNA in 15 μ l TE, pH 7.5 to each tube. Mix by gentle vortexing, completely resuspending cells that sedimented during the incubation. Do not allow the tubes to cool down.

Duplicate transformations using the same plasmid for each tube maximizes total transformants (e.g., screening a library).

6. Add 290 μ l of 50% (w/v) PEG 4000, 32°C (25°C for ts mutants) to each tube. Mix by gentle vortexing and incubate 60 min at 32°C (25°C for ts mutants).
7. Heat shock 15 min in a 43°C water bath or heat block. Cool the tubes and incubate 10 min at room temperature.
8. Microcentrifuge 2 min at 5000 rpm. Carefully remove the supernatant by aspiration.
9. Resuspend each tube in 1 ml of 50% YE medium by pipetting up and down with a 1000- μ l micropipettor.
10. Transfer each tube suspension to a 50-ml flask and dilute with 9 ml of 50% YE. Incubate with shaking at least 60 min at 32°C (25°C for ts mutants).
11. Plate <0.3-ml aliquots onto EMM plates selective for the marker of interest at 32°C (25°C for ts strains). If necessary, centrifuge the cells from a single transformation and resuspend in 1 ml EMM medium to spread more cells on a plate.

Colonies will generally form in 3 days at 32°C for wild-type cells

RAPID LITHIUM ACETATE PROCEDURE

This rapid method (Kanter-Smoler et al., 1994) reports efficiencies of up to 2×10^4 per microgram plasmid DNA. It uses small cultures and does not include a grow-out period for phenotypic expression.

Additional Materials (also see *Basic Protocol*)

Low-glucose EMM with appropriate supplements (*UNIT 13.15*)
 100 mM lithium acetate/1mM EDTA (pH 4.9)
 40% (w/v) PEG 3350/100 mM lithium acetate/1mM EDTA, pH 4.9
 42°C water bath

Additional reagents and equipment for growth of yeast (*UNIT 13.7*)

1. Grow a 10-ml culture of *S. pombe* in low-glucose EMM low to a density of 0.5–1 $\times 10^7$ cells/ml ($OD_{600} = 0.2$ to 0.5) as described above.
2. Wash once with 10 ml sterile water and resuspend in 1 ml sterile water. Transfer to a sterile 1.5-ml microcentrifuge tube and pellet cells by microcentrifuging briefly at 20°C.
3. Wash once with 0.2 ml of 100 mM lithium acetate/1 mM EDTA, pH 4.9, and resuspend in 50 μ l lithium acetate/EDTA.
4. Add 1 μ g DNA in up to 30 μ l water or TE, pH 7.5, and add 300 μ l of 40% (w/v) PEG 3350/100 mM lithium acetate/1mM EDTA, pH 4.9.
5. Incubate 30 min with agitation at 32°C.
6. Heat shock 15 min in a 43°C water bath without agitation.
7. Pellet by microcentrifuging briefly at 20°C. Resuspend in 1 ml TE, pH 7.5, and plate 200- μ l aliquots on EMM plates selective for the plasmid marker.

Colonies will generally form in 3 days at 32°C for wild-type cells.

**ALTERNATE
 PROTOCOL 1**

Yeast

13.17.3

ELECTROPORATION OF *S. POMBE* CELLS

The electroporation protocol for *S. pombe* is similar to that for *S. cerevisiae* (see UNIT 13.7); however, it is not necessary to plate the *S. pombe* cells on sorbitol-containing selection plates and in fact, doing so will actually retard growth. The sorbitol in which cells are suspended following treatment is sufficient for osmotic protection.

PROTOPLAST PROCEDURE

This is a time-consuming protocol that is seldom used any more, although it can be useful for creating diploids between nonmating strains (see Support Protocol for protoplast fusion). Transformation frequency is $\sim 1 \times 10^4$ to 5×10^4 transformants per microgram DNA.

Additional Materials (also see Basic Protocol)

- EMM medium (UNIT 13.15)
- Cit/phos/EDTA (see recipe)
- Cit/phos/sorbitol (see recipe)
- NovoZym 234 (BiosPacific)
- 10 mM Tris·Cl, pH 7.6 (APPENDIX 2)/1.2 M sorbitol
- 10 mM Tris·Cl, pH 7.6 (APPENDIX 2)/10 mM CaCl₂/1.2 M sorbitol
- 10 mM Tris·Cl, pH 7.6 (APPENDIX 2)/10 mM CaCl₂/20% (w/v) PEG 4000
- CaCl₂/sorb/YE (see recipe)
- Selective EMM sorbitol plates (UNIT 13.15)

- 50-ml plastic centrifuge tube
- 37°C incubator
- 29° to 32°C incubator

Additional reagents and equipment for counting cells using a hemacytometer
(APPENDIX 3F)

Prepare cultures

1. Grow a 200-ml culture of *S. pombe* to an OD₆₀₀ of 0.2 to 0.5 (up to 1×10^7 cells/ml) in EMM medium.
2. Harvest cells by centrifuging 5 to 10 min at $4000 \times g$ (5000 rpm), room temperature. Decant the supernatant and resuspend the pellet in 10 ml cit/phos/EDTA. Transfer to 50-ml plastic centrifuge tube.
3. Harvest cells and resuspend each tube in 5 ml cit/phos/sorbitol.

Form protoplasts

4. Add 25 mg NovoZym 234. Immediately incubate 15 to 30 min at 37°C until spheroplasts have formed (check that cells have rounded up under a microscope).

Using different enzymes (e.g., Sigma's lysing enzyme) changes the shape of the cells when protoplasting. This is because different enzymes break different sugar linkages in the cell wall. Depending upon which linkages are broken, sometimes the cells get round, sometimes they stay rodlike, and sometimes they ooze out of their jackets like bright little refractile balloons, leaving ghosts behind. It is possible to check protoplasting efficiency for any morphology by taking a small aliquot (a few microliters) and dropping them on a slide with a large drop of 0.5% to 1% SDS. Protoplasted pombe will lyse in these conditions.

From this point, the protoplasts are extremely fragile, so handle gently.

5. Add 35 ml of 10 mM Tris·Cl, pH 7.6/1.2 M sorbitol and divide between 2 to 4 tubes (no more than 3×10^8 spheroplasts/tube). Centrifuge 5 min at $1000 \times g$ (2000 rpm), 20°C.

It is better to centrifuge at slow speed and repeat than to pellet too vigorously.

6. Resuspend gently in 1 ml Tris·Cl/sorbitol, then add an additional 20 ml and invert several times to wash. Harvest and repeat. After the last resuspension in 1 ml, take a sample and count the number of protoplasts using a hemocytometer (APPENDIX 3F).
7. Resuspend at 2 to 5×10^8 protoplasts/ml in 10 mM Tris·Cl, pH 7.6/10 mM CaCl_2 /1.2 M sorbitol and pool the contents of the tubes.

Transform protoplasts

8. Add 1 to 10 μg plasmid in no more than 10 μl water to 100 μl protoplasts in a microcentrifuge tube. Incubate at room temperature for 15 min.
9. Add 1 ml of 10 mM Tris·Cl, pH 7.6/10 mM CaCl_2 /20% (w/v) PEG 4000 and incubate 15 min at room temperature.
10. Microcentrifuge briefly at 20°C. Drain well and resuspend the protoplasts in 0.2 to 0.5 ml CaCl_2 /sorb/YE. Incubate 30 to 60 min at 30°C.
11. Plate 0.2-ml aliquots onto well-dried, sterile, EMM sorbitol plates. Spread cells very gently with minimum force (e.g., using a spreader derived from a Pasteur pipet). Incubate 2 to 5 days at 29° to 32°C (i.e., until transformants appear).

Protoplasts can be aliquotted and stored in 10 mM Tris·Cl (pH 7.6)/10 mM CaCl_2 /1.2 M sorbitol (see step 7) up to 2 months at -70°C . The frequency of transformation is 1×10^3 transformants per microgram DNA for protoplasts stored in this way.

PROTOPLAST FUSION

This protocol allows creation of a diploid strain using nonmating strains. Cells lacking cell walls stick together in PEG and may fuse to generate a diploid. Recovery of a diploid requires selection for complementing markers in the two starting haploids (e.g., *ade6-M210* and *ade6-M216*, or *ura4⁺* in one parent and *leu1⁺* in the other).

Additional Materials (also see Basic Protocol and Alternate Protocol 3)

Strains with complementary markers

1. Follow protoplast transformation (see Alternate Protocol 3, steps 1 through 6) for each of two strains with complementing markers.

Complementing mating types are not required.

It is not necessary to count in a hemacytometer.

2. Combine 50 μl of each strain in a sterile tube. Incubate 15 min at room temperature.
3. Add 2 ml of 10 mM Tris·Cl (pH 7.6)/10 mM CaCl_2 /20% (w/v) PEG 4000. Incubate 25 min at room temperature.
4. Plate on EMM sorbitol plates selecting for the complementing markers at the appropriate temperature (32°C for wild type). As the cells are extremely fragile, spread gently (e.g., using a spreader from a thin Pasteur pipet).

Growth is reduced on sorbitol plates so it may take a few extra days to see colonies.

SUPPORT PROTOCOL

Yeast

13.17.5

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

CaCl₂/sorb/YE

10 mM CaCl₂
1.2 M sorbitol
0.5 mg/ml yeast extract
5 μg/ml each leu, ura, ade, and his
Sterilize by autoclaving
Store indefinitely at room temperature

Cit/phos/EDTA

2.82 g/liter Na₂HPO₄
4.2 g/liter citric acid
40 mM EDTA pH 8.0
Sterilize by autoclaving
Store indefinitely at room temperature

Cit/phos/sorbitol

7.1 g/liter Na₂HPO₄
11.5 g/liter citric acid
1.2 M sorbitol
Adjust to pH 5.6 with 5 M NaOH
Sterilize by autoclaving
Store indefinitely at room temperature

MB medium

0.5 g KH₂PO₄
0.36 g potassium acetate
0.5 g MgSO₄·7 H₂O
0.1 g NaCl
0.1 g CaCl₂·2 H₂O
5 g (NH₄)₂SO₄
500 μg H₃BO₄
40 μg CuSO₄·5 H₂O
100 μg KI
200 μg FeCl₃·6 H₂O
400 μg MnSO₄·H₂O
200 μg Na₂MoO₄·2 H₂O
400 μg ZnSO₄·7 H₂O
5 g glucose
10 μg biotin
1 mg calcium pantothenate
10 mg nicotinic acid
10 mg myo-inositol
150 mg uracil for *ura4⁻* strains or leucine for *leu1⁻* strains
Adjust volume to 1 liter with water
Sterilize by autoclaving
Store indefinitely at room temperature

COMMENTARY

Background Information

Early attempts to transform fission yeast relied upon *S. cerevisiae* plasmids and methods (Beach and Nurse, 1981). Because most *S. cerevisiae* plasmids and markers are poorly maintained in *S. pombe*, cloning of *S. pombe* specific origins and auxotrophic markers significantly improved efficiencies (for review, Russell, 1989, and Siam et al., 2003). While standard *S. cerevisiae* protocols, most notably electroporation, work to some extent in *S. pombe*, better success with the protoplast and lithium acetate methods is generally observed using *S. pombe*-specific methods.

The choice of method is determined by the equipment and time available. For example, electroporation (see Alternate Protocol 2) is extremely fast and simple to perform, but requires access to an electroporator. Recent studies report improving efficiency following treatment with DTT (Suga and Hatakeyama, 2001) and also describe storage of frozen electrocompetent cells (Suga and Hatakeyama, 2003).

The classic lithium acetate method (see Basic Protocol) is the most efficient technique, but it is also quite time consuming and relies on growth in a stringent minimal medium, as well as a grow-out period for phenotypic expression. A faster method should be suitable for most routine transformations (see Alternate Protocol 1). In addition, lithium acetate has been reported to be more efficient than protoplast treatment (see Alternate Protocol 2) for recovery of homologous integrants (Grallert et al., 1993). Methods have also been reported for preservation of frozen competent cells using this method (Broker, 1993).

Protoplast preparation was the first transformation technique reported for fission yeast (Beach and Nurse, 1981), but has fallen out of favor with the development of other methods. Homologous integration is less effective using this method than lithium acetate, with events skewed towards nonhomologous integration (Grallert et al., 1993), yet protoplast fusion mediated by PEG allows recovery of diploids from strains that are otherwise unable to mate (see Alternate Protocol 3 and Support Protocol). Competent protoplasts can also be preserved in the freezer (Jimenez, 1991).

Further information about transformation methods may be found in Moreno et al. (1991), Alfa et al. (1993), and Gould (2003).

Critical Parameters

The most important variable of any transformation procedure is the growth of the target strain. For all methods, cells in midexponential phase are preferred. Growth in low glucose medium has been reported to enhance transformation efficiency for lithium acetate (Kanter-Smoler et al., 1994; Okazaki et al., 1990), but has not been tested for electroporation. Transformation of mutant strains, especially those defective in genes affecting DNA replication or repair, or temperature-sensitive strains, may have reduced efficiency.

The plasmid or other DNA being used should be of high quality, as nonspecific inhibitors of transformation may be present in crude preparations. The number of transformants obtained per microgram generally decreases with larger amounts of DNA, although total transformants will still increase. Especially for electroporation, DNA should be in a small volume without any excess salt.

Targeted integration of linear plasmids or fragments varies enormously in efficiency, depending upon the actual locus and the amount of homology used to target the event (Grallert et al., 1993; Keeney and Boeke, 1994). A high background of nonhomologous integrants must often be screened to find the desired clone. Increasing the homology generally results in much more efficient targeting. Protoplast transformation is not recommended for homologous integration (Grallert et al., 1993).

Troubleshooting

Transformation is highly variable depending upon the strain, plasmid, and conditions of growth. Mutants defective in DNA replication are particularly difficult to transform, and may be slow to form colonies; however, simply transforming a known plasmid into a strain does not require high efficiency and can be performed without much concern or effort. In contrast, transformation of a plasmid library relies upon efficient recovery of a large number of colonies. In this case, a control plasmid containing the marker of interest should be used to optimize conditions, including temperature of growth before and after transformation, growth stage of the culture (mid- versus late exponential), and method employed.

A special consideration is transformation with a library to complement a temperature-sensitive mutant. Failure to recover any trans-

formants at the restrictive temperature may reflect absence of a complementing clone in the population, or may reflect growth conditions required to establish the plasmid. If direct plating of the transformation to the restrictive temperature doesn't work, it often helps to allow phenotypic expression of the clone at permissive temperature for a day or two before transferring the plates to the restrictive temperature. In some cases, allowing the transformed colonies to develop at permissive temperature and then replica plating to restrictive temperature (e.g., using phloxin B plates to identify the rescuing clones; see UNIT 13.15) may be the best way to proceed.

Anticipated Results

Transformation efficiencies may exceed 10^5 colonies per microgram DNA when the protocols are performed exactly with well-behaved wild-type strains, but are more likely to be in the 10^3 to 10^4 range with some mutants if frozen competent cells are used or if short cuts are taken in the procedures.

Time Considerations

Rapid lithium acetate transformation (see Alternate Protocol 1) and electroporation (see Alternate Protocol 2) require minimal manipulation of the target culture: a few washes, treatment with the DNA, and rapid plating. These can be accomplished within an hour or two. The classic lithium acetate method (see Basic Protocol) and protoplast preparation (see Alternate Protocol 3) are significantly more laborious, including prolonged incubation steps. These will occupy several hours.

Literature Cited

- Alfa, C., Fantes, P., Hyams, J., McLeod, M., and Warbrick, E. 1993. Experiments With Fission Yeast. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Bähler, J., Wu, J.Q., Longtine, M.S., Shah, N.G., McKenzie, A., Steever, A.B., Wach, A., Philippsen, P., and Pringle, J.R. 1998. Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast* 14:943-951.
- Beach, D. and Nurse, P. 1981. High frequency transformation of the fission yeast *S. pombe*. *Nature* 200:140-142.
- Broker, M. 1993. Rapid transformation of cryopreserved competent *Schizosaccharomyces pombe* cells. *Biotechniques* 15:598.
- Gould, K. L., ed. 2003. Methods. Academic Press, San Diego.
- Grallert, B., Nurse, P., and Patterson, T.E. 1993. A study of integrative transformation in *Schizosaccharomyces pombe*. *Mol. Gen. Genet.* 238:26-32.
- Jimenez, J. 1991. Cryopreservation of competent *Schizosaccharomyces pombe* protoplasts. *Trends Genet.* 7:40.
- Kanter-smoler, G., Dahlkvist, A., and Sunnerhagen, P. 1994. Improved method for rapid transformation of intact *Schizosaccharomyces pombe* cells. *Biotechniques* 16:798.
- Keeney, J.B. and Boeke, J.D. 1994. Efficient targeted integration at *leu1-32* and *ura4-294* in *Schizosaccharomyces pombe*. *Genetics* 136:849-856.
- Moreno, S., Klar, A., and Nurse, P. 1991. Molecular genetic analysis of the fission yeast *Schizosaccharomyces pombe*. *Meth. Enzymol.* 194:795-823.
- Okazaki, K., Okazaki, N., Kume, K., Jinno, S., Tanaka, K., and Okayama, H. 1990. High-frequency transformation method and library transducing vectors for cloning mammalian cDNAs by trans-complementation of *Schizosaccharomyces pombe*. *Nucl. Acids Res.* 18:6485-6489.
- Russell, P. 1989. Gene cloning and expression in fission yeast. In *Molecular Biology of the Fission Yeast* (A. Nasim, P.G. Young, and B.F. Johnson, eds.) pp. 243-271. Academic Press, San Diego.
- Siam, R., Dolan, W.P., and Forsburg, S.L. 2003. Choosing and using *S. pombe* plasmids. *Methods* In Press.
- Suga, M. and Hatakeyama, T. 2001. High efficiency transformation of *Schizosaccharomyces pombe* pretreated with thiol compounds by electroporation. *Yeast* 18:1015-1021.
- Suga, M. and Hatakeyama, T. 2003. High-efficiency electroporation by freezing intact yeast cells with addition of calcium. *Curr. Genet.* 43:206-211.

Contributed by Susan L. Forsburg
The Salk Institute for Biological Studies
La Jolla, California

CHAPTER 14

In Situ Hybridization and Immunohistochemistry

INTRODUCTION

This chapter presents some commonly used cytological techniques for determining the temporal and spatial expression patterns of both mRNA (in situ hybridization of tissue sections) and protein (immunohistochemistry) in individual cells, tissues, and even whole embryos. The study of gene products using biochemical and molecular techniques often requires tissue samples containing a considerable amount of the target molecule. This presents a difficulty because many developmentally interesting genes are expressed either in a minority of cells in complex tissues or for only brief periods of time during the differentiation of an organism or tissue. This chapter presents some commonly used cytological techniques for determining the temporal and spatial expression patterns of both mRNA (in situ hybridization of tissue sections) and protein (immunohistochemistry) in individual cells. Light microscopy coupled with sensitive detection methods allows the detection of local RNA or protein expression and adds an invaluable tool to the gene product detection repertoire.

The in situ hybridization techniques described here rely upon the hybridization of a specifically labeled nucleic acid probe to the cellular RNA in individual cells or tissue sections. Three areas of technical expertise are required. First, preparation of a suitable nucleic acid probe demands an understanding of the principles of molecular biology (e.g., subcloning, plasmid preparation, and radiolabeling). Successful tissue preparation requires practical experience in the art of histology. Finally, as with all morphological techniques, the correct interpretation of the experimental results requires familiarity with cell biology, anatomy, or embryology.

Although the extent of expertise outlined here may seem overwhelming to some, the reader should realize that the protocols are described in great detail, and that the commentary sections on critical parameters and troubleshooting can be very useful in avoiding or analyzing common mistakes. Of course, it is always reasonable to consult an experienced colleague for advice.

Immunohistochemical localization of cellular molecules exploits the ability of antibodies to bind specific antigens (usually proteins) with high affinity. The technique may be used to localize antigens to subcellular compartments or individual cells within tissues. It is especially powerful when used together with in situ hybridization to localize both the mRNA and protein products of a particular gene.

The chapter begins with a discussion of fixation and sectioning of mouse tissues (although most procedures should be applicable to those of other species). Fixation and sectioning (*UNITS 14.1 & 14.2*), which critically affect the success of both in situ hybridization and immunohistochemistry experiments, must frequently be optimized for each experimental application. Although these parameters are discussed at length, the protocols presented are not complete treatises on these methods and a good histology text and the advice of a trained histologist are highly recommended.

Contributed by J. G. Seidman

Current Protocols in Molecular Biology (2003) 14.0.1-14.0.2
Copyright © 2003 by John Wiley & Sons, Inc.

In Situ
Hybridization
and Immuno-
histochemistry

14.0.1

Supplement 64

Pretreatment, in situ hybridization, and washing of cells and tissue sections are discussed in *UNIT 14.3*. In situ hybridization can be performed on either paraffin (*UNIT 14.1*) or frozen sections (cryosections; *UNIT 14.2*). Alternate protocols are given for each case, although with the appropriate modifications either protocol can be used successfully with each type of sectioned material. The advantages of each method are detailed in the commentary of *UNIT 14.3*, but by far the most important reason to choose one method over the other is equipment availability. If a microtome is available there is generally no reason to buy a cryostat and vice versa.

Detection of hybridized probe and the counterstaining and mounting of treated cytological preparations using isotopic (radioactive; *UNITS 14.3-14.5*), non-isotopic (*UNITS 14.7 & 14.8*), and fluorescent (*UNITS 14.10 & 14.11*) probes are described. Radioisotope-labeled hybridization probes have a very high sensitivity and resolution. The hybridization signal can be detected by either film or emulsion autoradiography (*UNIT 14.4*). A number of alternate staining and mounting procedures are presented in *UNIT 14.5*. A method for quantifying radioactive signal produced by in situ hybridization of a radioactive using a phosphorimager is presented in *UNIT 14.12*. Non-isotopic methods for detecting hybridization signals (*UNIT 14.7 & 14.8*) provide an attractive alternative for RNAs expressed at slightly higher levels because they provide increased sensitivity while avoiding the hazards normally associated with the use of radioisotopes. *UNIT 14.7* describes fluorescence in situ hybridization, with methods for amplifying weak signals as well as enzymatic detection methods using horseradish peroxidase or alkaline phosphatase. Two units (*14.10 & 14.11*) deal with technical aspects of fluorescence light microscopy and confocal microscopy. *UNIT 14.10* discusses the basis of fluorescence microscopy, the methodology used to detect the fluorescence emitted from hybridized fluorescent probes. *UNIT 14.11* presents confocal microscopy, which can be used to increase the sensitivity and specificity of in situ hybridization methods. The use of the polymerase chain reaction in conjunction with in situ hybridization, presented in *UNIT 14.8*, allows the specific amplification of low-abundance sequences and can be used with any probe, regardless of label.

In situ hybridization to whole embryos, rather than cells or tissue sections, is a powerful tool for examining RNA expression during early stages of vertebrate development. Whole-mount in situ hybridization, presented in *UNIT 14.9*, can be used to detect patterns of transcription in intact vertebrate embryos, including chicken, mouse, and *Xenopus*.

UNIT 14.6 surveys immunohistochemical techniques for protein localization and focuses on the use of indirect immunofluorescence as the primary detection method. Alternative methods (immunoperoxidase and immunogold labeling), suitable for bright-field microscopy, are described and critically reviewed. An extensive troubleshooting guide with a list of suitable routine controls is also included.

Both in situ hybridization and immunohistochemistry eventually produce images of cells, tissues or whole embryos overlaid with a pattern of RNA or protein. Because most laboratories capture these images digital photography, the images can be manipulated and quantitated using standard digital image analysis tools. Valuable tools for biologic image analyses are described in *UNIT 14.15*.

J. G. Seidman

Fixation, Embedding, and Sectioning of Tissues, Embryos, and Single Cells

UNIT 14.1

This unit describes selected methods for fixing and sectioning various forms of biological materials ranging from tissues to single cells. Sections prepared according to these protocols can then be used to examine cell and tissue morphology and in studies involving in situ hybridization, immunohistochemistry, and enzyme histochemistry. The basic protocol describes how tissues and embryos can be prepared for sectioning by fixing in paraformaldehyde followed by embedding in wax, while the alternate protocol describes fixation of suspended or cultured cells. Two support protocols cover preserving and fixing organs by perfusion of whole animals with paraformaldehyde and a procedure for sectioning wax blocks of fixed tissue, plus the subsequent mounting of sections onto prepared or “subbed” glass slides.

PARAFORMALDEHYDE FIXATION AND PARAFFIN WAX EMBEDDING OF TISSUES AND EMBRYOS

BASIC
PROTOCOL

Samples are fixed for the appropriate time and dehydrated through a graded ethanol series. They are then impregnated in paraffin wax and cast into blocks.

Materials

- 4% paraformaldehyde (PFA) fixative, freshly prepared at 4°C
- Paraffin wax (e.g., Paraplast)
- 50%, 70%, 95%, and 100% ethanol
- Xylenes (Baker)
- Silicone spray
- 20-ml snap-cap glass vials (silanized for small samples; APPENDIX 3)
- 60°C oven and heating block with holes to hold 20-ml glass vials
- Embedding molds and rings (e.g., VWR Scientific)
- Hot forceps or hot Pasteur pipet with end cut off
- Hot Pasteur pipet with end drawn out and sealed

Fix the samples

1. Place dissected organs or embryos in labeled 20-ml snap-cap glass vials (for very small samples use silanized glass vials). Label the caps as well as the vials in case the label is removed by organic solvents. Fill vials with freshly prepared 4% PFA fixative, 4°C. Allow fixation to proceed at 4°C for the desired time.

Optimal fixation time is that which gives good morphology as well as good signal-to-noise ratio after in situ hybridization, and must be determined empirically for each sample. Hence, series of different fixation times, ranging from 15 min to 4 hr, should be tested.

In general, organs dissected from perfused mice have given best results when fixed for 1 to 2 hr (Zeller et al., 1987; Muller et al., 1988). Fixation times for mouse embryos vary greatly depending on age and size of the embryo: isolated gestational day 7 and 8 embryos (where the day of the copulatory plug is designated gestational day 0) should be fixed 15 to 30 min, day 9 and day 10 embryos should be fixed 1 to 2 hr, and older embryos 2 hr or longer (Zeller et al., 1987).

For better and more homogeneous fixation of individual organs, animals should be perfused with fixative prior to harvesting tissues. See support protocol for perfusion.

2. Begin melting paraffin wax in 60°C oven using a sufficient amount to process all the samples (see step 8).

In situ
Hybridization
and Immuno-
histochemistry

14.1.1

Contributed by Rolf Zeller

Current Protocols in Molecular Biology (1989) 14.1.1-14.1.8

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 7

Dehydrate the samples

3. After fixation is completed, pour off the fixative (be careful not to lose the sample), and replace with 50% ethanol. Immediately change the 50% ethanol to fresh 50% ethanol. Incubate 20 min and change the 50% ethanol again. A total of three changes of ethanol, each followed by a 20-min incubation at room temperature, is done this way.
4. Continue dehydration by incubating three times in 70% ethanol, 20 min each time (room temperature).

At this stage samples can be stored in 70% ethanol for a few days, if desired.

5. Incubate in 95% ethanol for 20 min at room temperature. Repeat two more times.
6. Incubate in 100% ethanol for 20 min at room temperature. Repeat two more times.

Do not reduce these incubation periods. Incomplete dehydration will lead to unusable paraffin blocks.

7. Replace 100% ethanol with xylenes. Immediately change to fresh xylenes and incubate 10 min. Follow with two more xylenes changes and 10-min incubations after each change.

CAUTION: Xylenes are toxic organic solvents—steps 7 and 8 should be carried out in a hood.

Prepare samples for embedding in wax

8. Pour off xylenes and add 5 ml fresh xylenes to each vial (check that samples are still there). Add an equal amount of molten wax using a hot glass pipet.

Transfer of the wax must be very fast to avoid hardening.

9. Mix and leave samples at room temperature overnight. The wax/xylenes mixture will harden, but enough paraffin is dissolved in the xylenes to start impregnation of the samples.

10. Transfer samples to 60°C oven to melt the wax/xylenes mixture.

11. Prepare a 60°C heating block (the 20-ml vials must fit its holes). When the wax/xylenes mix is molten, transfer vials to heating block.

12. Pour off the wax/xylenes mixture from one vial into a waste bottle (be careful to retain the samples). Immediately add fresh molten wax to the vial with a hot glass pipet. Put the vial back into the heat block. Repeat procedure with all the vials.

Work quickly to avoid hardening of the wax. If hardening occurs, melt the wax again at 60°C before initiating the 1-hr incubation (step 13). Otherwise, impregnation of the samples will be poor and sectioning impossible.

13. Return vials to 60°C oven and incubate 1 hr at 60°C.

14. Remove vials from oven and place again into 60°C heating block. Repeat steps 12 to 13 twice for a total of 3 hr of incubation.

Embed samples in wax and prepare blocks

15. The samples are now ready for embedding. Prepare embedding molds according to manufacturer's instructions (e.g., coated with silicone spray) and fill one of the molds with molten paraffin wax using a hot glass pipet.

16. Immediately transfer sample to the wax-filled mold using hot forceps or a hot cut-off Pasteur pipet.

In general, only one to two embryos should be embedded in each block. This facilitates ribboning (see support protocol for sectioning).

17. Place an embedding ring on the mold and fill with paraffin wax. Label the embedding ring to facilitate future identification of samples. Samples can be oriented within the mold using a hot drawn-out and sealed Pasteur pipet.

The embedding ring will be used to attach the sample to the microtome holding clamp (support protocol for sectioning, step 2).

18. Leave cast blocks at room temperature to harden completely.
19. Remove cast blocks from embedding molds and store in a dry place at room temperature. Perform sectioning according to the support protocol in this unit.

Blocks stored this way are stable and can be used successfully for years. Moisture destroys the blocks; dry storage is essential.

FIXATION OF SUSPENDED AND CULTURED CELLS

Cells are fixed to poly-L-lysine-coated slides with 4% PFA, dehydrated, and stored at -70°C until use (Berger, 1986).

Materials

4% paraformaldehyde (PFA) fixative at room temperature
3 \times and 1 \times phosphate-buffered saline (PBS), pH 7.2
50%, 70%, 95%, and 100% ethanol

Poly-L-lysine-coated glass slides (see reagents and solutions)
Moist chamber
Glass staining dishes

1. Pipet 10 μl of cells (2×10^7 cells/ml density) in a drop onto a poly-L-lysine-coated glass slide.

Resuspend cells in serum-free medium if possible.

2. Place slide in a humidified chamber and let cells settle 30 min.
3. Transfer slides into a slide rack and immerse in a glass staining dish filled with 4% PFA fixative. Allow fixation to proceed 20 min at room temperature.

Alternatively, cells can be grown directly on poly-L-lysine-coated cover slips or slides, and processed with 4% PFA fixative for 20 min at room temperature.

4. Transfer slide rack to a new dish filled with 3 \times PBS and incubate 2 min.

This step is necessary to stop the PFA fixation.

5. Transfer to a new dish containing 1 \times PBS and incubate 2 min. Replace PBS with fresh 1 \times PBS and incubate another 2 min.
6. Dehydrate slides through a series of 5-min incubations in 50% ethanol, 70% ethanol, 95% ethanol, and 100% ethanol. All dehydration steps are done in dishes.

The dehydration can be performed efficiently by setting up four glass staining dishes, each having a different ethanol concentration. Slides aligned in a slide rack can be easily dipped into each solution and transferred to the next dish after 5 min.

7. Air dry the slides until they are completely dried.
8. Transfer slides to slide boxes containing desiccant and store at -70°C . Slides must be completely dry before storing in desiccant and freezing; otherwise, freezing artifacts can occur. Slides stored in this manner can be used for several weeks.

ALTERNATE PROTOCOL

In situ Hybridization and Immuno- histochemistry

14.1.3

PERFUSION OF ADULT MICE

Perfusion of animals is essential for achieving good morphology and preservation of brains, kidneys, hearts, and many other organs. For the untrained person the first perfusion might not work—it is therefore advisable to learn this procedure using some control animals first, or to seek the advice of a trained animal pathologist or colleague familiar with this procedure.

The protocol described here for mice (applicable also to most other species) involves first exchanging the animal's blood with phosphate-buffered saline and subsequently exchanging the saline with freshly prepared ice-cold (4°C) 4% paraformaldehyde. This protocol gives reproducibly good results without the necessity of using a pressure-controlled perfusion setup. These more sophisticated setups (available in major Anatomy departments) need to be used if, for example, blood vessels or kidney glomeruli need to be fixed in situ without any possible collapse of capillaries.

Materials

- 1× phosphate-buffered saline (PBS)
- 4% paraformaldehyde (PFA) fixative, freshly prepared at 4°C
- 2 syringes (20- to 30-ml) equipped with 23-G needles
- Container (bag) for mouse and CO₂ gas
- Dissection instruments (scissors, forceps, etc.)

1. Fill one syringe with 1× PBS, and another with 4% PFA fixative, 4°C. Set aside.
2. Kill mouse with CO₂ gas in a bag. Immediately after respiratory arrest, lay mouse on its back and open the thorax carefully to avoid excessive bleeding. Cut carefully through the rib cage and remove the diaphragm for access to the heart.

It is important to work fast, but be careful to get a good perfusion. If the blood clots or the main blood vessels are harmed, perfusion will not work.

3. Carefully insert syringe filled with 1× PBS into the left ventricle. Cut open the right ventricle for drainage, allowing the 1× PBS to be slowly but constantly perfused into the heart (see Fig. 14.1.1).

If the perfusion is working well, blood-rich organs such as liver, spleen, and kidneys will turn grayish-white.

4. After most of the blood has been flushed out, remove the syringe with 1× PBS and insert syringe filled with 4% PFA fixative (4°C) into the same puncture of the left ventricle. Slowly perfuse the mouse with ~20 ml fixative.

The start of the perfusion with 4% PFA is considered the starting point of the fixation process ($t = 0$).

A sign of a good perfusion is a muscle tremor best seen on the limbs and tail. By the end of this procedure the animal should be stiff. If perfusion is unsuccessful, it is not worthwhile continuing at this point (see Commentary).

5. Following perfusion, dissect out organs and tissues, transfer into well labeled glass vials (filled with 4% PFA fixative), and store on ice (4°C). Further fixation and processing is carried out as described in the basic protocol or *UNIT 14.2*.

SECTIONING SAMPLES IN WAX BLOCKS

Sectioning paraffin blocks containing samples is a process requiring experience and should be learned from an experienced individual if at all possible. Some universities and hospitals have service facilities that perform sectioning. If so, provide instructions on the desired thickness of the sections and orientation for mounting on subbed glass slides (see below). The following is an outline of the sectioning protocol. Paraffin blocks are cut into

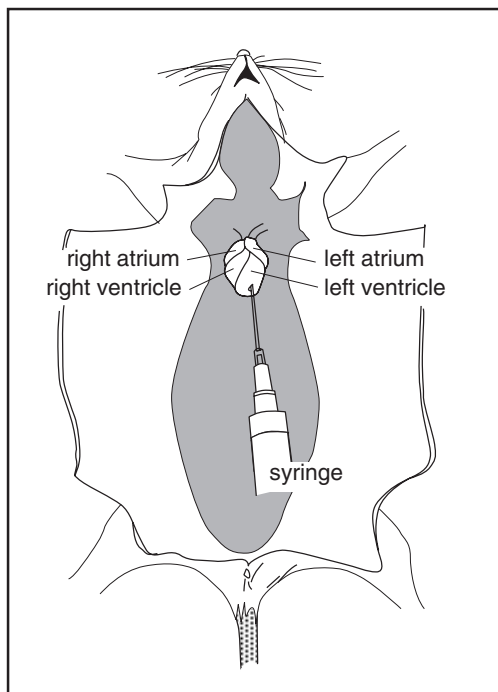


Figure 14.1.1 Perfusion of a mouse. This diagram illustrates the chambers of the heart and the correct positioning of the syringe in the left ventricle for perfusion.

thin (8- μm) tissue sections, which are mounted on subbed slides to be further processed for in situ hybridization (UNIT 14.3).

Materials

- Paraffin wax block(s) containing samples
- 0.2 \times gelatin subbing solution
- Microtome, complete with knives
- Gelatin-subbed glass slides
- Fine brushes
- Slide warmer or heating plate set between 45° and 50°C
- 42°C oven
- Desiccant (e.g., Humicaps, United Desiccants-Gates)

1. Cut a wax block containing samples into a trapezoidal shape using a razor blade. Carefully shave off the extra wax and do not cut too close to the embedded sample. Otherwise the sample might be destroyed or the paraffin wax might shatter.

The trapezoidal shape facilitates ribboning during sectioning (step 2) and allows the placement of sections close together on the slide.

2. Attach the trapezoid block (with the wide edge facing the knife) to the holding clamp of a microtome and begin sectioning (see Fig. 14.2.1c-e). Cut 8- μm sections.

Thinner sections (>1 μm) may be cut, but 8 μm is the standard thickness used. If the block is sectioning well, a ribbon will form.

CAUTION: *Microtome knives are very sharp. For the proper and safe use of the microtome, consult the manufacturer's manual.*

3. Put a drop of 0.2 \times gelatin subbing solution on a gelatin-subbed glass slide (see Fig. 14.1.2, which illustrates steps 3 to 8).
4. Transfer the ribbon of sections (up to 10 sections, depending on sample size) onto the drop on the subbed slide using fine brushes.

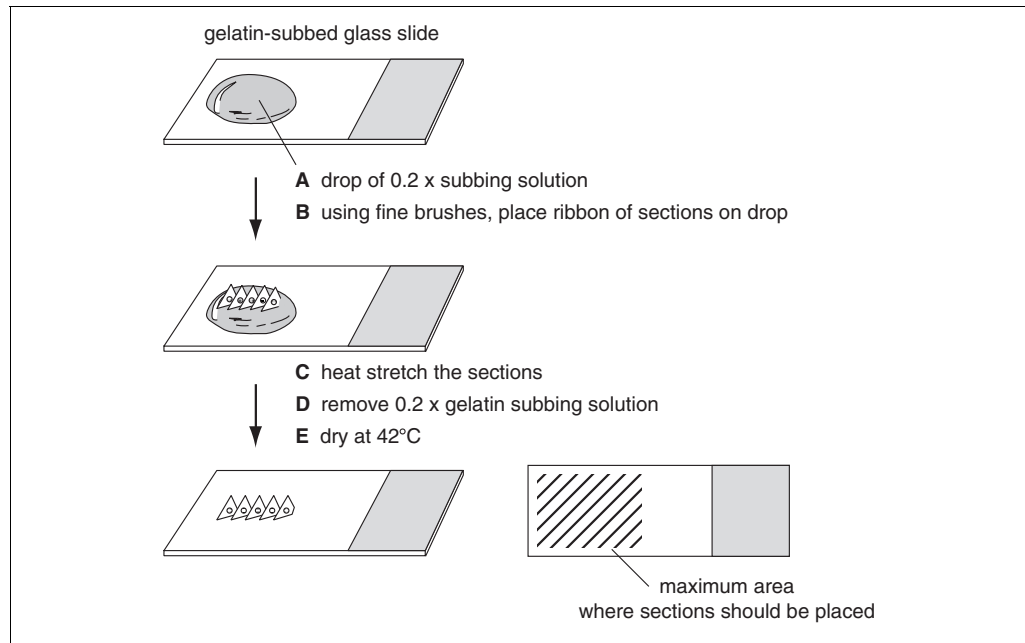


Figure 14.1.2 Placement of paraffin sections on gelatin-subbed glass slides. Steps **A** to **E** depicted here correspond to steps 3 to 8 in the support protocol.

5. Transfer the slide onto a slide warmer or heating plate set between 45° and 50°C.
Heat will stretch the sections and remove wrinkles. Do not leave slides on the slide warmer longer than necessary or the sections may be destroyed.
6. After stretching is complete, remove slide from slide warmer and carefully remove the remaining subbing solution using a Pasteur pipet.
7. Dry slide at room temperature.
8. Incubate 1 to 2 days at 42°C to firmly attach sections to subbed slides.
9. Store sections in a slide box with desiccant at –20°C for up to several weeks.
If desired, several different samples can be placed on the same slide. After steps 3 to 7 have been completed for sample 1, sample 2 is sectioned and sections are placed next to those of sample 1. Place sections as close together as possible to minimize the area on the slide that they occupy. The more space on a slide occupied by sections, the more hybridization solution will be needed for in situ hybridization (UNIT 14.3). Do not place sections too close to the edge of the slide.

REAGENTS AND SOLUTIONS

Gelatin-subbed glass slides

Use 25 × 75-mm precleaned Rite-on Microslides from Clay Adams or comparable source and prepare as follows:

1. Wash slides in water with detergent for a few minutes. Use any good soap.
2. Rinse slides in water 30 min. During this rinse, prepare 1× gelatin subbing solution (see below).
3. Place slides in a slide rack and immerse 2 min in a glass staining dish filled with 1× gelatin subbing solution (4°C). Subbing must be done carefully to avoid air bubbles.
4. Remove slide rack from subbing solution and set on its side, with the frosted sides of the slides facing downward, to let excess subbing solution drip off. Dry overnight before use. Slides subbed with gelatin are stable at room temperature for several weeks.

1× and 0.2× gelatin subbing solution

Dissolve gelatin (type Bloom 275) in water at 70°C to make a 0.5% solution. Cool the solution to room temperature, and add $\text{CrK}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$ to 0.05% (chromium potassium sulfate). Cool the solution to 4°C on ice. Use immediately, as gelling occurs after prolonged storage of the solution on ice.

Prepare 0.2× gelatin solution by diluting 1× solution 5-fold with water. It can be used for several days if prepared using sterile water and stored at room temperature. If cloudiness or bacterial growth occurs, discard the solution.

4% paraformaldehyde (PFA) fixative

Heat to 60°C a volume of water equal to slightly less than $\frac{2}{3}$ the desired final volume of fixative. Weigh out a quantity of paraformaldehyde (Baker) that will make a 4% solution and add it with a stir bar to the water. Cover. Transfer to fume hood and maintain on heating plate at 60°C with stirring. Add 1 drop 2 N NaOH with a Pasteur pipet. The solution should become almost clear fairly rapidly, but will still have some fine particles that will not go away. Be careful not to overheat the solution. Remove from heat and add $\frac{1}{3}$ vol 3× PBS. Bring pH of solution to 7.2 with HCl, add water to final volume, and filter using a Millipore or Nalgene filter. Cool to room temperature, or to 4°C on ice.

CAUTION: *Formaldehyde is a carcinogen and may cause allergic reactions.*

3× and 1× phosphate-buffered saline (PBS), pH 7.2

Prepare the following solutions: 390 mM NaCl/30 mM Na_2HPO_4 and 390 mM NaCl/30 mM NaH_2PO_4 . Mix to obtain 3× PBS, pH 7.2, and autoclave or filter sterilize. Prepare 1× PBS by diluting 3× PBS 3-fold with water.

Poly-L-lysine-coated glass slides

Clean glass slides as described for gelatin-subbed glass slides (use slide racks and glass staining dishes). Prepare a sufficient amount of 500 $\mu\text{g}/\text{ml}$ poly-L-lysine (Sigma, MW >150,000) in water. Coat the glass slides by dipping individually into this solution, air dry, and store at 4°C. Use within a week. (A useful small dipping chamber is a plastic cytology slide mailer; see *UNIT 14.4*.)

COMMENTARY

Background Information

The best fixative for in situ hybridization appears to be paraformaldehyde (see Zeller et al., 1987), although other fixatives have been used successfully. Sample fixation is one of the crucial steps of the in situ hybridization protocol (*UNIT 14.3*) and is decisive in determining subsequent success or failure. Underfixation of samples leads to bad morphology and/or loss of signal, whereas overfixation leads to fixation artifacts, loss of signal, and/or increased background (noise) signal. The 4% PFA fixation recommended here has worked very well for mouse embryos as well as on tissues, but other investigators find it imperative to use different fixatives for some species. (A list of references to alternative fixation protocols for other species is given at the end of *UNIT 14.3*.) Homogeneous fixation of large, well vascularized or-

gans, such as brains, hearts, and kidneys, is achieved only by complete perfusion of the animal. On the other hand, it is not necessary to perfuse animals if embryos of pregnant females or poorly vascularized tissues (e.g., mammary tissues) are to be harvested.

Critical Parameters

The use of freshly prepared 4% PFA fixative is essential for good sample fixation. This is in contrast to fixation of tissues to be used for histology examination where commercially available formaldehyde stock solutions are used. Perfusion and fixation are done at 4°C because formaldehyde fixation is progressive over time; lowering the temperature slows the fixation process and provides a more suitable fixation.

Troubleshooting

Fixation. Failure of paraffin blocks to section well can be due to either under- or overfixation. If the sections crumble and disintegrate very easily, or are destroyed during stretching, underfixation or poor dehydration or impregnation with wax are frequently to blame. If the cellular morphology of the section is very poor, and cavities have collapsed, underfixation is again most likely to be at fault. If the block is extremely hard and falls out in one piece during sectioning, overfixation is the reason.

These problems might be partially or completely solved by using longer or shorter fixation times. Optimal fixation, however, may not be the only factor in obtaining a strong hybridization signal and low background. Many of the parameters have to be adjusted empirically by doing a series of test experiments using different fixation times and different hybridization conditions (see *UNIT 14.3*).

Perfusion. Perfect perfusion requires working quickly and methodically. If large areas of grayish-white decolorization are not seen on the liver, spleen, and kidneys of a perfused animal, it is not worthwhile continuing with the procedure. The main causes of a failed perfusion are severing of major blood vessels during opening of the animal, improper insertion of syringe needle into the left ventricle (thereby penetrating into other heart chambers or going all the way through the heart), and slow perfusion, leading to blood clotting.

Sectioning. Problems during sectioning can be caused by blunt or scratched knives or a misadjusted microtome. It is important to sharpen the blades occasionally or arrange for a service check of your microtome. If sections are lost in large quantities from slides during subsequent treatment and hybridization (*UNIT 14.3*), insufficiently subbed slides are most likely to blame. In such cases, make up freshly subbed slides.

Last, but not least, sectioning is a routine which requires experience—if you are inexpe-

rienced, do not expect the first sections to come out perfectly. Get advice from an experienced trained histologist, or let such a person do the sectioning for you. A detailed description is also given in Luna (1968).

Time Considerations

The amount of time needed to complete this procedure varies greatly with experience and the number of samples processed. From the start of fixing tissues until sections can be processed for in situ hybridization, allow at least 4 to 5 days. Approximate times for the individual operations are as follow:

perfusion—5 to 10 min
fixation—15 min to 4 hr
dehydration—5 to 6 hr
impregnation—overnight
embedding—5 to 6 hr
sectioning—≥1 hr, depending on
number of slides and samples
processed
attachment of sections—1 to 2 days.

Literature Cited

- Berger, C.N. 1986. In situ hybridization of immunoglobulin-specific RNA in single cells of the β -lymphocyte lineage with radiolabeled DNA probes. *EMBO J.* 5:85-93.
- Luna, L.G. 1968. *Manual of Histologic Staining: Methods of the Armed Forces Institute of Pathology* (3rd ed.). McGraw-Hill, New York, N.Y.
- Muller, W.J., Sinn, E., Pattengale, P.K., Wallace, R., and Leder, P. 1988. Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated *c-neu* oncogene. *Cell* 54:105-115.
- Zeller, R., Bloch, K.D., Williams, B.S., Arceci, R.J., and Seidman, C.E. 1987. Localized expression of the atrial natriuretic factor gene during cardiac embryogenesis. *Genes Dev.* 1:693-698.

Contributed by Rolf Zeller
European Molecular Biology Laboratory
Heidelberg, Germany

Cryosectioning

This unit describes sample preparation and sectioning methods for frozen tissue. Sections of this type are used in a variety of light microscopic procedures including in situ hybridization, immunohistochemistry, and enzyme histochemistry.

SPECIMEN PREPARATION AND SECTIONING

Samples are quick-frozen and then sectioned. This protocol produces sections that are particularly suitable for immunohistochemistry and enzyme histochemistry. See commentary for critical parameters on specimen size before freezing tissue.

Materials

- Liquid N₂
- CryoKwik (Damon) or isopentane
- OCT compound (Tissue Tek II, Miles)
- Small Dewar flask or expanded polystyrene box
- Filter paper cut into 1 × 7-cm strips (e.g., Whatman 50)
- Forceps
- Metal rod

- Cryostat and microtome equipped with roll bar or plastic roll plate
- Cutting chuck (metal platform that supports specimen during sectioning)
- Heat sink or CO₂ jet freezer
- Fine brush and ¼-in. brush
- Gelatin- or poly-L-lysine-coated slides (UNIT 14.1), prelabeled with specimen details in pencil

All tools used in the cryosectioning, including the trimming razor blade, should be prechilled within the cryostat chamber. NOTE: The slides should not be chilled.

Freeze tissue specimen and prepare for sectioning

1. Chill a 50-ml Pyrex beaker by immersing in a Dewar flask (or expanded polystyrene box) filled with liquid N₂.
2. Fill beaker with CryoKwik.
3. Label one end of a filter paper strip with pencil and place specimen on the other end with forceps.

Samples prepared as a slurry are best frozen as 200- to 300-μl drops applied to the filter paper strips with a Pasteur pipet. Most tissue samples will stick to the filter paper naturally, but if they will not, it may be necessary to use a little OCT compound between the sample and filter paper.

4. Ensure the CryoKwik is liquid. If necessary, melt the CryoKwik by touching it with a metal rod.
5. Immerse specimen in cold CryoKwik ~1 min. Some of the CryoKwik may solidify on the sample, but this will evaporate in the cryostat.

At this stage specimens may be stored in either liquid N₂ or in a -70°C freezer.

6. Place filter paper strips with tissue specimens in a cryostat chamber and mount on cutting chucks with a thin layer of OCT compound (Fig. 14.2.1).

Do not use excessive OCT compound as this may lead to freezing artifacts; in particular, immersion of tissue in OCT should be avoided, unless small multiple specimens are being sectioned (see commentary).

BASIC PROTOCOL

In situ Hybridization and Immuno- histochemistry

Contributed by Simon Watkins

Current Protocols in Molecular Biology (1989) 14.2.1-14.2.8

Copyright © 2000 by John Wiley & Sons, Inc.

14.2.1

Supplement 7

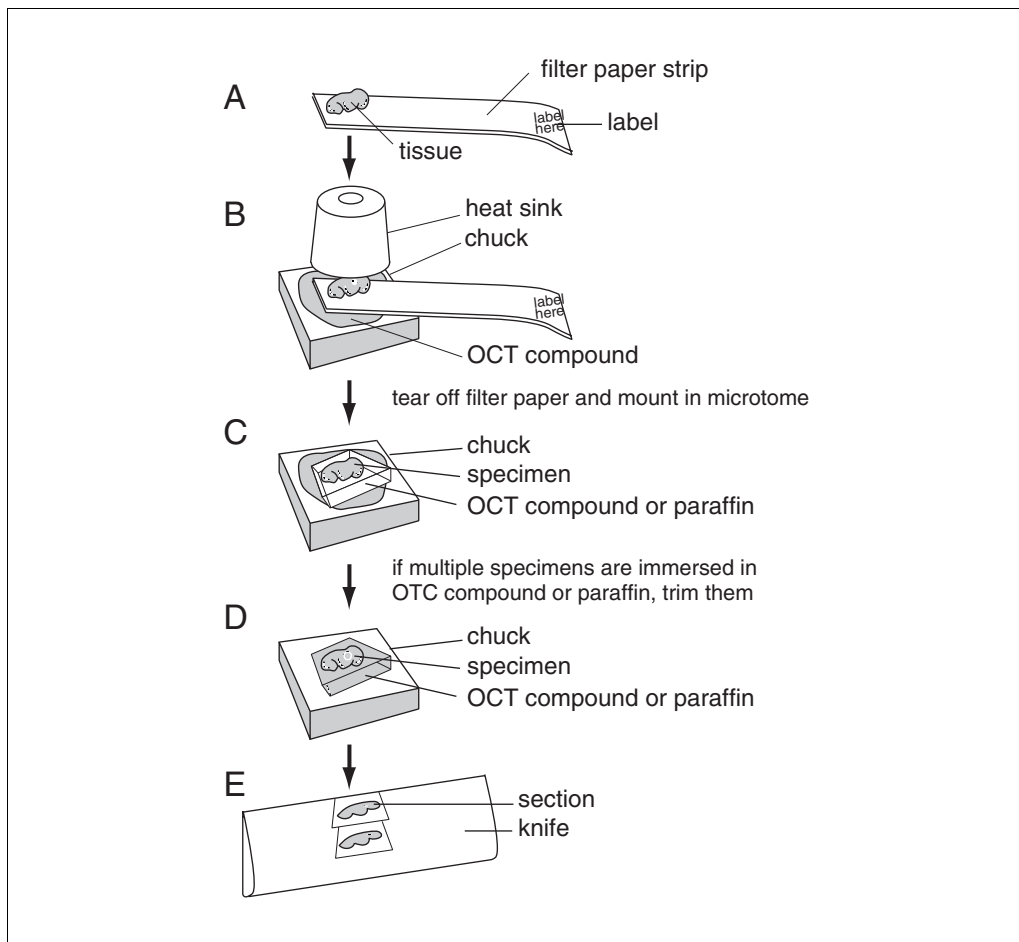


Figure 14.2.1 Sectioning procedure. **A** to **E** illustrate cryosectioning and correspond to steps 6 to 13 of basic protocol; **C** to **E** also illustrate paraffin wax sectioning and correspond to steps 1 to 4 of support protocol, *UNIT 14.1*.

- Cool the specimen/chuck mount within the cryostat with a precooled heat sink or a CO₂ jet freezer until the OCT compound solidifies.

The specimen block must remain frozen throughout the mounting procedure to avoid formation of cryoartifacts (see Table 14.2.1).

- Leave specimen within the cryostat for >10 min to reach the same temperature as the prechilled microtome.
- Trim specimen block to a trapezoid shape—if it is convenient and does not compromise the morphology—using a prechilled razor blade. Tear off filter paper.

Section specimen block

- Mount the chuck in the microtome with the parallel faces of the trapezoid in line with the knife and the wide edge toward the knife edge.
- Retract the specimen block until it easily clears the knife edge.
- Produce a smooth “face” on the block using the knife at rapid cutting speed (one section/sec).

Choice of section thickness is important as this will have considerable effect on the final resolution at the microscope. Generally, the thickest usable sections are ~10 μm and should not exceed 20 μm, since extremely thick sections will frequently

shatter, rendering the specimen useless. It is best to reduce section thickness until sectioning is only just possible ($\sim 5 \mu\text{m}$) as this will markedly improve quality.

13. Once the specimen block has been cut so that either the structure of interest can be seen within the cut face or the block dimensions are optimal (between 0.5 and 1 cm on a side), sections may be collected.

Collect cryosections

Generally one of two methods is employed to facilitate section collection: the roll-bar method (described in step 14 and Fig. 14.2.2) and the plastic roll-plate method (described in step 15 and Fig. 14.2.3).

14. The roll-bar method (this system is most suitable for large block faces, though not suitable for serial sections):
 - a. Flip the roll bar over such that it lies against the cut face of the specimen block and align it such that it is positioned ~ 1 mm from the edge of the block nearest the knife.
 - b. Cut sections slowly (~ 2 to 3 sec to traverse the cut face).
 - c. As sections are cut, the roll bar should lift onto the knife face, sandwiching the section between it and the knife, thereby preventing the section from rolling up on itself or being lost due to static electricity.
 - d. Proceed with step 16.

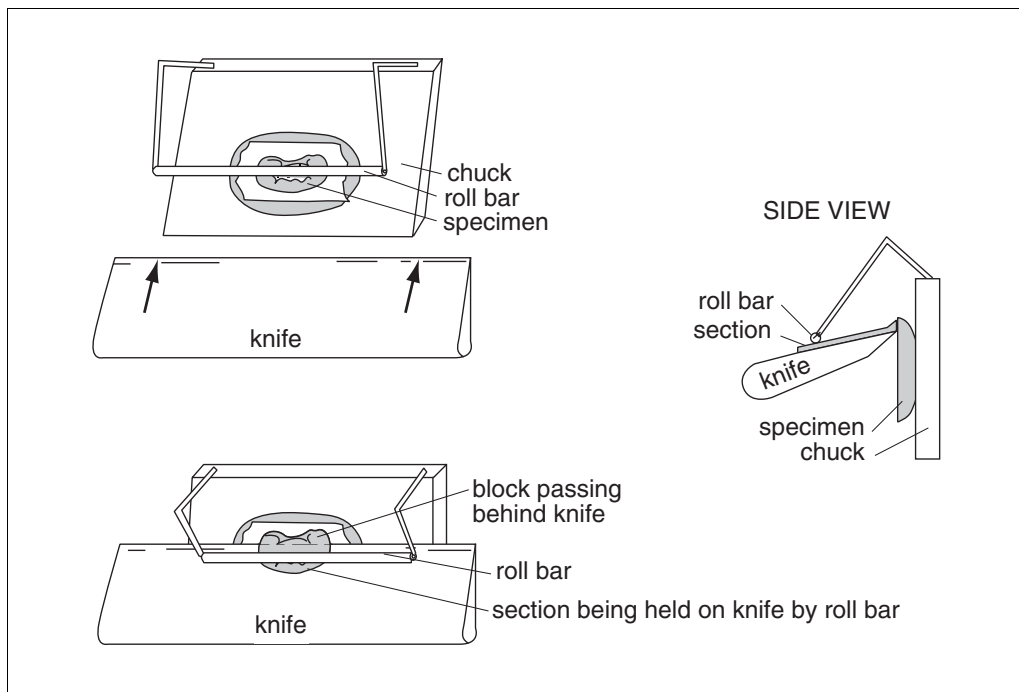


Figure 14.2.2 A roll bar (a U-shaped metal bar) is mounted on the rear of a cutting chuck onto which the specimen has been placed. As sections are cut, the roll bar will lift onto the knife face, over the section, and prevent the section from rolling up.

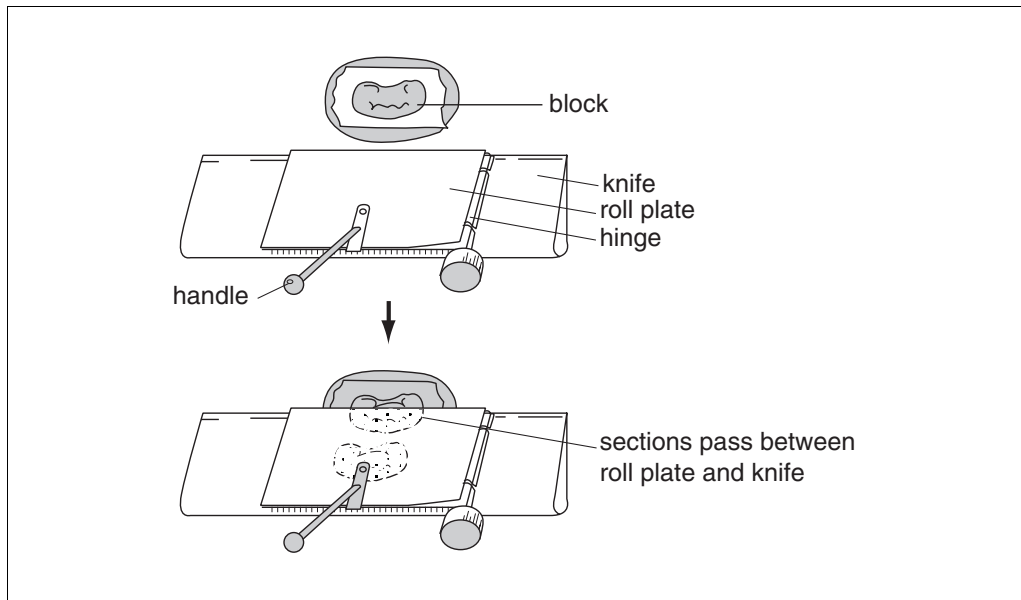


Figure 14.2.3 A plastic roll plate is hinged at the back of the microtome's knife holder and placed parallel to the plane of the knife edge. The plate touches the knife at the leading edge and is tilted up from the knife at the rear so that sections may pass between the knife and the plate.

15. The plastic roll-plate method (this system is the most commonly used as it allows easy serial sectioning, though it is less suitable for large block faces):
 - a. Prior to sectioning, flip the roll plate down such that it touches the knife and lies parallel to the knife edge, though retracted slightly behind the cutting edge of the knife.
 - b. Turn the crank and cut sections, which under these conditions will roll up at the cutting edge of the knife.
 - c. Lift the roll plate and, with a small brush, brush cut sections from the knife edge.
 - d. Drop the roll plate and advance it slightly.
 - e. Repeat steps 15a to 15d until sections pass smoothly between the roll plate and knife (generally when the roll plate leads the knife edge by ~ 0.5 mm) then proceed with step 16.

16. Lift the roll bar or plate and move the cut sections to the rear of the knife with a fine brush.

17. Collect sections by gently touching to a warm prelabeled gelatin- or poly-L-lysine-coated slide (see morphology troubleshooting, Table 14.2.1).

Mounted sections may be left in the microtome until all sectioning is complete. If they will be used for in situ hybridization, treat immediately following the support protocol for the fixation of cryosections for in situ hybridization. If they will be used for immunohistochemistry or enzyme histochemistry, the sections may be stored at -70°C for a limited time (overnight) in an airtight container (see also UNIT 14.6 on air dried sections for immunohistochemical study).

18. After lifting sections onto slide, clear condensed ice from the knife using a thick ($\frac{1}{4}$ -in.) brush.

This brushing may need to be quite vigorous and, therefore, should always be toward the knife edge; otherwise the knife will be rapidly dulled and the brush ruined.

FIXATION OF CRYOSECTIONS FOR IN SITU HYBRIDIZATION

Cryosections used for in situ hybridization must be fixed with paraformaldehyde and then dehydrated.

Materials

4% paraformaldehyde (PFA) fixative (UNIT 14.1), freshly prepared

3× and 1× phosphate-buffered saline (PBS; UNIT 14.1)

30%, 60%, 80%, 95%, and 100% ethanol

Moist chamber (Fig. 14.2.4)

Desiccant (e.g., Humicaps, United Desiccants-Gates)

1. Put slides containing sections in a moist chamber and cover the sections with 4% PFA fixative. Cover the chamber and incubate 20 min at room temperature for previously unfixed tissues or 5 min for fixed tissues.

Fixing and dehydrating the sections in a moist chamber will firmly fix most tissues to the slides throughout the remaining steps. Some tissues, however, may stick to slides better than others, and may be fixed in Coplin jars. This can be determined for the tissue of interest.

2. Aspirate off the fixative and flood the section with 3× PBS from a pipet or squirt bottle. Incubate 5 min.
3. Repeat step 2 twice with ~1× PBS.
4. Remove PBS and flood with ethanol for 2 min at each of the following concentrations: 30%, 60%, 80%, 95%, and 100%.
5. Dry the slides and store with desiccant at -70°C in an airtight box for up to several weeks. Before opening the box to use the slides, bring to room temperature.

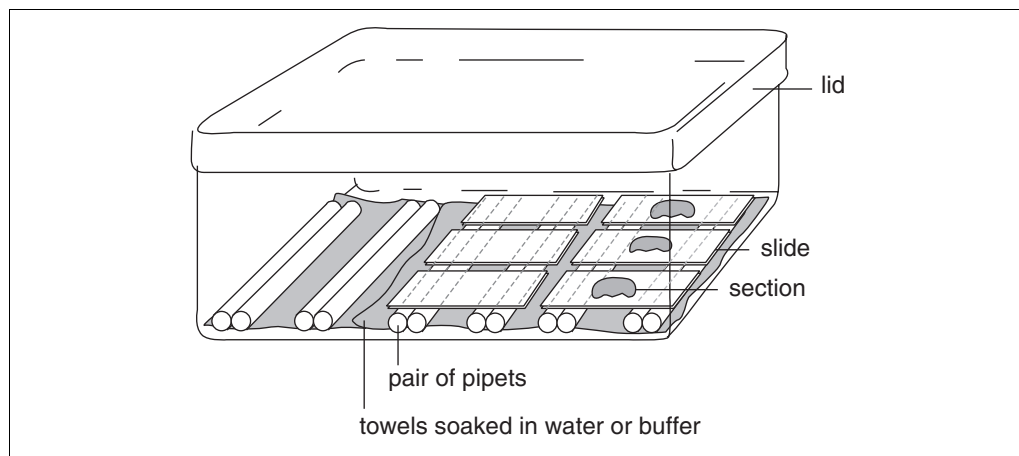


Figure 14.2.4 Moist chamber. Take a conveniently sized container with a tight lid (e.g., Tupperware or equivalent) and attach pairs of 5- to 10-ml plastic pipets to bottom so that they support slides at either end. Place pipets so that the maximum number of slides can be set on them. Alternatively, place a stainless steel cake rack in the bottom. Put absorbent paper on the bottom and drench with water or, for in situ hybridization, where incubation times are very long, use a solution of identical osmolarity and formamide concentration to your hybridization buffer (see UNIT 14.3).

TISSUE FIXATION AND SUCROSE INFUSION

Frequently for cryosectioning, morphology is much improved if the tissue is fixed and infused with sucrose prior to sectioning.

Materials

Phosphate-buffered saline (PBS; *UNIT 14.1*)

2% (for immunohistochemistry) or 4% (for in situ hybridization)
paraformaldehyde (PFA) fixative (*UNIT 14.1*; for 2%, make 1:1
dilution in PBS of 4% PFA fixative)

0.5 M sucrose in PBS

1. Fix small tissue samples or isolated embryos (e.g., day 7 or day 8 mouse embryos) in 2% (for immunohistochemistry) or 4% (for in situ hybridization) PFA fixative for 30 min at 4°C.

Larger embryos, and some tissues, may require fixation for up to 4 hr. Larger organs should be fixed in situ within the animal by perfusion (UNIT 14.1) prior to dissection.

2. Following fixation, wash tissue samples twice in PBS.
3. Infuse samples in 0.5 M sucrose in PBS until tissue sinks (~1 to 3 hr).

This infusion may be extended to overnight, but no longer.

Paraformaldehyde fixation is a continuous fixation, as the tissue will become increasingly fixed the longer it is left in the fixative. Therefore, the fixation time should be minimized, as overfixation will result in loss of reactivity for both in situ hybridization or immunohistochemistry. This is particularly important for in situ hybridization (see commentary, UNITS 14.1-14.3).

COMMENTARY

Background Information

Cryosections, which are rapidly and relatively easily prepared, provide a good system for visualizing cellular fine detail. There are currently three main sectioning methodologies available to study the morphology of cells and tissues at the light microscope level: cryosections, paraffin sections (*UNIT 14.1*), and plastic (commonly methacrylate) sections. The latter two methodologies generally present a superior morphology than cryosections and both have been used for in situ hybridization. However, the preparation procedures are more arduous, and neither method can be used reliably for immunohistochemistry or enzyme histochemistry where retention of native protein form and function may be necessary.

Critical Parameters

Fixation

As a general rule, enzyme histochemical methods are most effected by changes in the state of the protein within the sections. For these studies, fixation of any kind should be avoided

as this may impair the functional state of the protein to be studied. For sections to be used in immunohistochemical studies, in which proteins need not be functional, fixation should not proceed to the point at which antigenic activity is lost. This is rarely a problem if only formaldehyde is used as a fixative. In situ hybridization is too new a method to have any golden rules of procedure. Since proteins bound to RNAs must be removed by pronase digestion or denaturation, a mild fixation protocol should be used. Optimization of fixation conditions is critical for the success of the in situ hybridization procedure. For standard histological staining methods, where fixation is not important, the morphology may be optimized. If this type of methodology is the ultimate goal of the technique, paraffin or methacrylate sections may be the preferred methods of preparation, as the storage and handling of these types of specimens are both easy and cheap.

Size of specimens

The size of specimens to be sectioned may vary from groups of cells grown in suspen-

Table 14.2.1 Troubleshooting Guide for Morphological Problems Encountered During Cryosectioning

Problem	Possible cause	Solution
Poor morphological preservation throughout tissue	Inadequate care during dissection or excessive delay between dissection and freezing	Take great care during dissection not to damage tissue by stretching or excessive bending. Freeze tissue soon after dissection to prevent proteolysis by endogenous proteolytic enzymes. This is particularly important for delicate samples (e.g., small embryos or brain) or enzyme-rich tissues (e.g., liver). Fixation and infusion with sucrose may improve morphology considerably (see support protocol).
Holes in tissue	Ice crystal formation in tissue during freezing, or due to thawing of the block and subsequent refreezing and ice crystal formation in the cryostat	Ensure that blocks are frozen in either isopentane or CryoKwik rather than a gas such as nitrogen. The latter will boil on contact with the warm tissue giving a “shell freezing” effect, where the boiling gas insulates the outer layers of the tissue from the cold liquid nitrogen, thereby slowing freezing and allowing ice crystal formation. Ensure that when specimens are mounted in the microtome they are protected from warming by a large precooled heat sink. Ensure that you do not touch the specimens with your fingers or with warm tools.
Morphology appears fuzzy	“Pressure artifact” due to excessive pressure on the back of the slide while picking up the section	Use less pressure when picking up the section. A mixture of warmth and static charge will normally ensure that sections stick to the slide. If this is not the case, the slide is too cold and may be locally warmed by touching a finger to the back of the slide.
Morphology appears smeared	Slide was moved while lifting section	Hold slide against the rear of the knife and use this as a fulcrum while gently rocking the slide down to touch the section, thereby ensuring no lateral motion of the slide.
Sections do not stick to slide, they wash away during labeling	Slides were not adequately coated	Ensure that sections stick to slides by dip-coating them as described in <i>UNIT 14.1</i> .
Sections blow away from knife edge during sectioning	A buildup of static electricity in the cryostat	There is no reliable cure. It is possible to buy polonium brushes which may help. Antistatic guns (such as those designed for record players) have also been used successfully.
Sections have lines running up them and separate into ribbons on knife edge	Knife is dull	Sharpen knife. Generally a knife will cut ~20 to 30 blocks of soft tissue before it needs to be sharpened. If starting with new equipment, a disposable blade system (e.g., Fisher) may be cheaper over the long term.

sion to whole animals—the upper size limit of the specimen is normally defined by the capacity of the standard laboratory cryostat, ~1.5 to 2 cm³. The method of preparation prior to sectioning is important as this will ultimately affect the final image quality.

Small, multiple specimens. This category includes samples ranging in size from cells (e.g., lymphocytes grown in suspension), to small animal samples (e.g., early mouse embryos or small nematodes). With specimens of this type, orientation is generally not possible. Instead, sufficient samples are included at random orientations in each specimen such that most or all orientations are present. Prior to freezing, specimens are mixed as a slurry in an inert supporting media which is easy to section (e.g., OCT compound).

Large specimens. Larger specimens, such as organs or tissue samples, are generally sectioned individually. A factor that should be considered prior to freezing is whether the tissue sample has a homogeneous density, as this will make sectioning much easier. An example of extreme inhomogeneity is whole animal limbs, where skin, muscle, and bone may all be contained within the same sample, making sectioning very difficult; in this case, specific tissues should be dissected out prior to freezing. Small tissue samples, which are easily lost and are not available in large quantities, are best mounted in another tissue, such as in a “liver sandwich,” which frequently serves as a negative control as well as a tissue support.

Another feature which should be considered prior to sectioning is tissue orientation, as this may be indistinguishable in whole frozen tissue. This is important in tissues such as muscle or pieces of brain; here the type of information gathered from the labeled sections will depend entirely on orientation.

Sectioning

The cryostat must be thermally stable prior to sectioning. Fluctuations in temperature make sectioning very difficult as the specimen will either contract away from the knife, or worse still, expand into the knife, usually causing the sample to tear out from the chuck and be ruined. For this reason, it is necessary to allow the cryostat to stabilize for some time (30 min to 1 hr) after the chosen temperature is reached, or after altering the temperature, before proceeding with sectioning. Normal cutting temperatures are between –30° and –20°C. Usually,

softer tissues require colder cutting temperature. As a guide, softer samples (e.g., brain tissue) are cut at –30°C, while harder samples with a defined cytoskeleton (e.g., muscle) are cut at about –20°C. For very hard specimens it may be necessary to raise the temperature to –15°C, although temperatures any warmer than this may lead to excessively soft specimen blocks. It is quite easy to recognize temperature problems. Generally, if the sections appear brittle and crumble, the cryostat is too cold. If the sections wrinkle excessively, the cryostat is too warm.

Morphology

A variety of morphological problems may arise. Refer to Table 14.2.1 for a guide to troubleshooting some of the more common difficulties.

Anticipated Results

Cryosections should show good retention of morphology (Fig. 14.6.2), with no artifacts due to ice crystals and sectioning problems as described in Table 14.2.1. The sections provide an ideal substrate for immunohistochemistry (UNIT 14.6), in situ hybridization (UNIT 14.3), or enzyme histochemistry.

Time Considerations

Specimen preparation. If the specimen is to be fixed and infused with sucrose, this procedure may take several hours; otherwise, freezing the tissues takes only a few minutes.

Sectioning. With practice, this procedure becomes very rapid—<30 min for each specimen. However, it takes time to reach this degree of competence, so in the beginning, allot ≥1 hr for each specimen.

Key References

Hollands, B. 1962. Histochemistry and microtomy of fresh-frozen tissue. *In* Progress in Medical Laboratory Technique (F.J. Baker, ed.) pp. 112-135. Butterworth, London.

Zugibe, F. 1970. Diagnostic Histochemistry. C.V. Mosby, St. Louis.

These references describe the basic cryosectioning method and applications in clinical pathology.

Contributed by Simon Watkins
Harvard Medical School
and Dana-Farber Cancer Institute
Boston, Massachusetts

In Situ Hybridization to Cellular RNA

UNIT 14.3

In situ hybridization to cellular RNA is used to determine the cellular localization of specific messages within complex cell populations and tissues. Tissues may either be embedded in paraffin and sectioned on a microtome (see Basic Protocol), or frozen and sectioned in a cryostat (see Alternate Protocol). RNA contained in the specimens is hybridized to a specific radiolabeled probe (see Support Protocols 1 and 2), which is then detected using film autoradiography or emulsion autoradiography (UNIT 14.4).

HYBRIDIZATION USING PARAFFIN SECTIONS AND CELLS

Dewaxed sections and single cells are hydrated, then pretreated to denature the tissue and remove some of the proteins to make the mRNA transcripts more accessible for hybridization. On the second day, specimens are hybridized to [³⁵S]UTP-labeled riboprobe for up to 4 hr. After hybridization, specimens are washed and digested with RNase to remove excess and nonspecifically hybridized probe.

Materials

- Specimens mounted on subbed glass slides (e.g., see UNIT 14.1)
- Dewaxing/rehydration (dehydration) series consisting of 3 staining dishes of xylenes, 2 staining dishes of 100% ethanol, and 1 staining dish each of 95%, 70%, and 50% ethanol
- 0.2 N HCl
- 2× SSC (APPENDIX 2), 70°C
- 1× and 3× phosphate-buffered saline (PBS; UNIT 14.1)
- Predigested pronase solution (optional; see recipe)
- 2 mg/ml glycine in 1× PBS (optional)
- 4% paraformaldehyde (PFA) fixative (UNIT 14.1), freshly prepared at room temperature
- 10 mM dithiothreitol (DTT)/1× PBS (see recipe), freshly prepared at 45°C
- Blocking solution (see recipe), prepared immediately before use at 45°C
- Triethanolamine (TEA) buffer (see recipe), freshly prepared
- Acetic anhydride
- [³⁵S]UTP-labeled riboprobes (see Support Protocol 1)
- S-riboprobe competitor (see recipe)
- 50 mM DTT, sterile
- Hybridization mix A (see recipe)
- Moist chamber solution A (see recipe)
- Wash solutions A, B, and C (see recipes)
- RNase digestion solution (see recipe)
- 50% ethanol/0.3 M ammonium acetate
- 70% ethanol/0.3 M ammonium acetate
- 95% ethanol/0.3 M ammonium acetate
- 100% ethanol
- Two sets of slide racks (one clearly labeled for RNase use only)
- ≥10 glass staining dishes
- 45°, 55°, and 50°C water baths
- Slide box with desiccant (e.g., Humicaps, United Desiccants–Gates)
- 100°C heating block or water bath
- 45°C incubator
- Moist chambers (Fig. A.3E.4; one clearly labeled for RNase use only)
- ≥4 glass staining dishes, clearly labeled for RNase use only

**BASIC
PROTOCOL**

**In situ
Hybridization
and Immuno-
histochemistry**

14.3.1

Additional reagents and equipment for fixation and sectioning of tissues and cells (UNIT 14.1), probe preparation (see Support Protocols 1 and 2), autoradiography (UNIT 14.4), and staining and mounting (UNIT 14.5).

NOTE: All of the following steps are performed by incubating the slides containing specimens (held in a slide rack) in glass staining dishes in the indicated solutions. All solutions are made up fresh and are used only once unless indicated otherwise.

Dewax and rehydrate the specimens

1. Prepare the dewaxing/rehydration series (can be reused several times) and 0.2 N HCl while slides containing specimen sections (kept in slide boxes at -20° or -70°C ; see UNIT 14.1) are warmed to room temperature. Also, start preheating $2\times$ SSC to 70°C (for step 5).
2. Dewax slides by three changes in xylenes, 2 min each (not necessary for slides containing unembedded single cells).
3. Rehydrate through the following regimen:
 - 100% ethanol—twice, 2 min each
 - 95% ethanol—2 min
 - 70% ethanol—2 min
 - 50% ethanol—2 min.

Denature the specimens

4. Denature specimens 20 min at room temperature in 0.2 N HCl.

This step will denature proteins and nick DNA, but will also partially reverse the paraformaldehyde fixation step.
5. Heat denature 15 min at 70°C in $2\times$ SSC.

This step denatures RNA and probably also removes some of the proteins to make RNA in sections or cells more accessible to hybridization.
6. Rinse 2 min in $1\times$ PBS.

Optional step: Under certain circumstances (see Commentary), a pronase digestion step is included here. Specimens are digested 15 min at 37°C with 0.1 to 10 $\mu\text{g}/\text{ml}$ pronase digestion solution. The optimal pronase concentration has to be determined empirically for each tissue, but in general, the highest possible pronase concentration that still gives good cellular morphology is used. Pronase digestion is stopped by rinsing slides 30 sec in 2 mg/ml glycine in $1\times$ PBS.
7. Post-fix specimens 5 min at room temperature in freshly prepared 4% PFA fixative.
8. Block fixation 5 min in $3\times$ PBS.
9. Rinse twice, 30 sec each time, in $1\times$ PBS.

Block the specimens

10. Equilibrate specimens in 10 mM DTT/ $1\times$ PBS for 10 min at 45°C in a water bath.
11. Block specimens with freshly prepared blocking solution 30 min at 45°C in a water bath covered with aluminum foil (iodoacetamide is light sensitive).

CAUTION: Both iodoacetamide and *N*-ethylmaleimide are very toxic substances; handle very carefully.

This step is included to block nonspecific sulfur-binding sites on tissue specimens; it may be omitted if ^3H - or ^{32}P -labeled probes are used for in situ hybridization.

12. Rinse twice, 2 min each time, in 1× PBS at room temperature.
13. Equilibrate specimens 2 min in freshly prepared TEA buffer.
14. Transfer slide rack to fresh TEA buffer and add acetic anhydride to a concentration of 0.25%. Mix quickly and incubate slides 5 min with agitation. Add additional acetic anhydride to reach a final concentration of 0.5% and incubate another 5 min.

This step blocks polar and charged groups on tissue sections, which will cause nonspecific sticking of probes. It is important to add the acetic anhydride immediately before specimens are incubated because the half-life of acetic anhydride in aqueous solutions is very short.

15. Block specimens 5 min in 2× SSC.

Dehydrate the specimens

16. Dehydrate specimens through 50% ethanol, 70% ethanol, 95% ethanol, and 100% ethanol (twice), for 2 min each at room temperature.

For this step, use the same ethanol solutions as in step 3.

17. Air dry specimens (or dry in desiccator), making sure specimens are absolutely dry before proceeding to step 18.
18. Store specimens in a slide box with desiccant at -70°C overnight.

Prepare the probe

19. Spin down the ethanol-precipitated antisense and sense ^{35}S -labeled riboprobes (see Support Protocol 1), as well as the S-ribo probe competitor. Dry pellets. Dissolve each pellet (corresponding to one reaction) in 5 μl of sterile 50 mM DTT. Add 2.5 μl (half a reaction) of S-ribo probe competitor to both antisense and sense riboprobe.

Start probe preparation early in the morning.

20. Heat dissolved probes to 100°C for 3 min in heating block or water bath.
21. Immediately add enough hybridization mix A to obtain a 0.3 $\mu\text{g}/\text{ml}$ final probe concentration (see Support Protocol 1 for details concerning determination of the probe mass synthesized). Mix well and count 1 μl (expected counts $\geq 1 \times 10^5$ cpm/ μl). Place tubes in water bath at 45°C (hybridization temperature).

If antisense and sense probes differ extensively in their counts per μl (e.g., antisense riboprobe has five times more counts per μl than sense riboprobe, or vice versa), make sure it is not due to improper mixing, and recount the samples. If the difference is consistent, dilute the sample that has more counts to get a roughly equal count per μl for both antisense and sense riboprobes. Radiolabeled probes that are substantially less than 10^5 cpm/ μl should not be used. The poor recovery was probably caused by loss during ethanol precipitation.

Hybridize the specimens

22. Set up hybridizations by carefully spreading an appropriate amount of probe on the specimens (e.g., 20 $\mu\text{l}/20$ mm³) using the tip of a pipet (see Fig. 14.3.1).

If background is a problem, a prehybridization step can be included prior to hybridization: Take up one reaction of S-ribo probe competitor in 5 μl of 50 mM DTT, heat 3 min to 100°C , and add 500 μl of hybridization mix A. Spread ≥ 20 μl of this prehybridization mix on the specimens as described in step 22. Incubate 1 to 2 hr at 45°C in a moist chamber. To remove prehybridization mix, tip slides, causing the prehybridization mix to collect at one edge and carefully blot off buffer with Whatman 3MM paper. Then proceed with step 22.

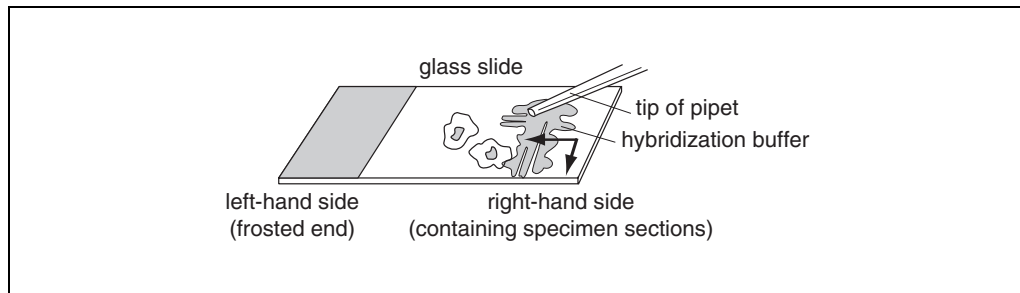


Figure 14.3.1 Application of hybridization mix to slides. Carefully spread hybridization on sections with the tip of a pipet (avoid touching sections). Arrows indicate directions of spreading of hybridization mix A (or B).

23. Place specimens in a moist chamber (see Fig. 14.2.4) containing moist chamber solution A and incubate at 45°C for the appropriate hybridization time. Perform a series of hybridizations ranging from 30 min to 4 hr.

It is crucial that slides are kept level; otherwise the hybridization mix will collect at one side, leading to uneven hybridization. Equilibrate and seal the moist chamber carefully since dried out specimens will result in high background. The osmolarity of the moist chamber solution must equal that of the hybridization mix to prevent its dilution or concentration.

24. During the last hour of hybridization, prepare and preheat wash solutions A, B, and C.

Wash the specimens

25. Start washing slides by dipping one at a time into 100 ml wash solution A at 55°C. Immediately place in a slide rack in a staining dish filled with wash solution A.

The first dip into wash solution A removes much of the radioactivity, and should be disposed of into radioactive liquid waste. For disposal of all other less radioactive liquid wash solutions, follow local guidelines.

Do not allow slides to dry during any of the hybridization and washing steps. Drying causes extremely high background.

26. Incubate twice, 15 min each time, in wash solution A at 55°C.
27. Incubate twice, 15 min each time, in wash solution B at 55°C.
28. Incubate twice, 2 min each time, in wash solution C at room temperature.

Treat specimens with RNase

29. Add 500 µl RNase digestion solution per slide, covering all specimens, and place the slides in a moist chamber (containing water) of the type shown in Figure 14.2.4 (but labeled “RNase”).

NOTE: In steps 29 to 34, use only materials clearly labeled “RNase.” Do not use these materials for any other steps in any of the protocols described in this unit. Cross-contamination with RNases can ruin the experiment.

30. Incubate 15 min at room temperature.
31. Wash slides twice, 30 min each time, in wash solution C at 50°C with gentle shaking.
32. Wash slides twice, 30 min each time, in wash solution A at 50°C with gentle shaking.
33. Wash slides twice, 5 min each time, in 2× SSC at room temperature.

34. Dehydrate through the following regimen (2 min each):

50% ethanol/0.3 M ammonium acetate
70% ethanol/0.3 M ammonium acetate
95% ethanol/0.3 M ammonium acetate
100% ethanol.

35. Air dry the slides.

36. Expose slides at least overnight against film, and then perform emulsion autoradiography (e.g., *UNIT 14.4*).

HYBRIDIZATION USING CRYOSECTIONS

Sections are briefly treated with pronase, acetylated, and hybridized to radiolabeled or digoxigenin-labeled probe. Although the protocol detailed here specifies cryosections, it may also be used successfully with dewaxed and hydrated paraffin sections (see Basic Protocol, steps 1 to 3).

Materials

Specimens mounted on subbed glass slides (*UNIT 14.2*)
Predigested pronase solution (see recipe)
50 mM Tris-Cl, pH 7.5 (*APPENDIX 2*)/5 mM EDTA
2 mg/ml glycine in 1× PBS
1× phosphate-buffered saline (PBS; *UNIT 14.1*)
Triethanolamine (TEA) buffer (see recipe), freshly prepared
Acetic anhydride
2× SSC (*APPENDIX 2*)
30%, 60%, 80%, 95%, and 100% ethanol
Labeled DNA or RNA probe (see Support Protocols 1, 2, and 3)
Hybridization mix B (see recipe)
Deionized formamide (see recipe)
50% dextran sulfate (Amersham Pharmacia Biotech)
3.3 M dithiothreitol (DTT), freshly prepared
Moist chamber solution B (see recipe)
DNA wash solution (see recipe), prewarmed to 37°C
RNA wash solutions I and II (see recipes), prewarmed to 50°C
20 µg/ml boiled ribonuclease A (*UNIT 7.3*) in 0.5 M NaCl/10 mM Tris-Cl, pH 8.0
0.6 M NaCl in 30% ethanol and in 60% ethanol

2 sets of slide racks and jars (one set will be reserved for RNase use only)
50°C water bath
Moist chamber
37° or 42°C incubator (or water bath)

Additional reagents and equipment for fixation and sectioning of tissues (*UNIT 14.2*),
probe preparation (see Support Protocols 1 and 2), autoradiography (*UNIT 14.4*),
and staining and mounting (*UNIT 14.5*).

Pretreat the sections for hybridization

1. Remove box of slides from freezer and allow to come to room temperature before opening.

***ALTERNATE
PROTOCOL***

**In situ
Hybridization
and Immuno-
histochemistry**

14.3.5

2. Place slides in rack and immerse 10 min in predigested and lyophilized pronase in 50 mM Tris·Cl, pH 7.5/5 mM EDTA.

The concentration of pronase required must be determined in a test experiment. If cryosections are used, digestion may sometimes be omitted. However, in this protocol paraffin sections require pronase digestion.

3. Rinse slides 30 sec at room temperature in PBS containing 2 mg/ml glycine, then twice in PBS for 30 sec each time.

This step stops proteolysis.

Acetylate the samples

4. Immerse slides 5 min in freshly prepared TEA buffer.
5. In a beaker, prepare enough TEA buffer to cover slides. Add acetic anhydride to a final concentration of 0.25% and pour on slides as quickly as possible. Agitate rack to mix. Incubate 10 min.
6. Wash slides twice, 5 min each time, in 2× SSC.
7. Dehydrate for 2 min each in 30%, 60%, 80%, 95%, and 100% ethanol. Dry and use immediately for hybridization.

Prepare the probes

8. Precipitate labeled DNA or RNA probes.
9. Determine the percent of label incorporated and estimate the mass of probe synthesized. Calculate the volume of final hybridization mix required to resuspend the pellet at a final concentration of 0.2 µg/ml per kilobase.
10. Resuspend the pellet first in 2 parts hybridization mix B and 2 parts deionized formamide. Subsequently, add 1 part 50% dextran sulfate. Mix thoroughly.

Probe may be prepared the previous day and stored frozen at -80°C.

Hybridize the probes

11. Boil labeled DNA probes 2 min and chill in ice bath, or place heat-labeled RNA probes 30 sec at 80°C and then hold at 50°C.
12. Add 50 mM DTT from a 3.3 M stock solution. Using the tip of a pipet, distribute sufficient probe to cover sections, e.g., 20 µl/20 mm² (see Fig. 14.3.1).

Refer to the support protocols for preparation of DNA and RNA probes.

If using single-stranded probes, several different probes may be heated at once. Double-stranded probes, however, are so concentrated that self-annealing is a problem, so denature only a few slides' worth of probe at a time.

13. Incubate 4 hr in a well-sealed moist chamber containing moist chamber solution B. Incubate DNA probes at 37°C and RNA probes at 42°C.

Wash the slides

- 14a. *DNA probes:* Wash 2 hr at 37°C with 4 to 5 changes of prewarmed DNA wash solution.

- 14b. *RNA probes:*

- a. Wash 15 min at least twice at 50°C in prewarmed RNA wash solution I.
- b. Treat slides 30 min at 37°C with 20 µg/ml boiled ribonuclease A in 0.5 M NaCl and 10 mM Tris·Cl, pH 8.0.

- c. Wash 15 min at least twice at 50°C in prewarmed RNA wash solution I.
- d. Wash 15 min twice at 50°C in prewarmed RNA wash solution II.

This wash temperature may be increased to 60°C, if desired, although morphology may suffer.

IMPORTANT NOTE: *Keep all glassware and racks used for washes and RNase treatment completely separate from glassware used for pretreatment and hybridizations.*

15. Dehydrate through the following regimen (2 min each):

0.6 M NaCl in 30% ethanol
 0.6 M NaCl in 60% ethanol
 80% ethanol
 95% ethanol
 100% ethanol.

The ethanol solutions used in step 7 may be reused here, but discard them after this step in order to avoid later RNase contamination.

16. Air dry the slides and detect hybridized probe by autoradiography (UNIT 14.4) or enzymatic detection (see Support Protocol 3; Haramis and Carrasco, 1996).

SYNTHESIS OF ³⁵S-LABELED RIBOPROBES

Materials

5× transcription buffer (see recipe)
 1 M dithiothreitol (DTT), freshly prepared
 Ribonuclease inhibitor (e.g., Amersham placental ribonuclease inhibitor or Promega Biotec RNasin)
 10 mM CTP, ATP, and GTP (UNIT 3.4)
 1 µg/µl restriction enzyme–digested plasmid (see recipe)
 1000 to 1500 Ci/mmol [³⁵S]UTP (UNIT 3.4)
 SP6 or T7 RNA polymerase (UNIT 3.8)
 10 mg/ml yeast tRNA or mouse poly(A) RNA for carrier
 1 U/µl RNase-free DNase I (e.g., Promega Biotec RQ1; UNITS 4.1 & 4.10)
 3 M sodium acetate
 7.5 M ammonium acetate
 100% and 70% ethanol, –20°C

NOTE: When ³⁵S-labeled probes are prepared, it is very important to add 10 mM DTT to all solutions containing [³⁵S]UTP, particularly after any step that may inactivate DTT (e.g. precipitation, column chromatography, and boiling). In addition, extreme care should be taken to prevent RNase contamination of reagents.

1. Prepare the following reaction mix at room temperature (20 µl total):

4.0 µl 5× transcription buffer
 0.2 µl 1 M DTT
 60 U ribonuclease inhibitor
 1.0 µl each of the three 10 mM NTPs
 1.0 µl digested DNA (1 µg/µl)
 10.0 µl [³⁵S]UTP
 16 U SP6 or T7 RNA polymerase.

2. Incubate 30 min at 37°C. Add 16 U more polymerase and continue incubation at 37°C for 40 min.

SUPPORT PROTOCOL 1

In situ Hybridization and Immuno- histochemistry

14.3.7

3. Add the following to the reaction mix:

60 U ribonuclease inhibitor
2.0 μ l 10 mg/ml carrier RNA
1.0 μ l DNase I.

Incubate 10 min at 37°C to remove template.

4. Add the following to the reaction mix:

0.8 μ l 1 M DTT
63.0 μ l sterile H₂O
10.0 μ l 3 M sodium acetate.

Remove 1 μ l and determine cpm/ μ l by scintillation counting.

5. Make reaction mix 2 M ammonium acetate by adding 36.4 μ l of 7.5 M ammonium acetate. Add 272 μ l of -20°C 100% ethanol, precipitate 10 min on dry ice, pellet, wash with -20°C 70% ethanol, and dry. Repeat if desired. Resuspend in 100 μ l of 10 mM DTT and determine cpm/ μ l.

6. Determine percent incorporation.

Under these conditions 70% to 90% of the label is incorporated, resulting in ~70 to 90 ng labeled RNA.

7. Following synthesis of labeled probe, add 50 to 100 μ g carrier tRNA per 20- to 25- μ l synthesis reaction, add sodium acetate to 0.3 M and DTT to 10 mM, and then ethanol precipitate as in step 5 (adjusting quantities proportionally).

Riboprobes should be used within 2 to 3 days of synthesis.

SUPPORT PROTOCOL 2

SYNTHESIS OF ³⁵S-LABELED DOUBLE-STRANDED DNA PROBES

Radioactive double-stranded DNA probe may be synthesized by nick translation or random oligonucleotide-primed synthesis using [³⁵S]dNTPs. Add 10 mM DTT to the standard reaction mixes (UNIT 3.5) and replace the [³²P]dNTP with two different [³⁵S]dNTPs (specific activity \geq 1000 Ci/mmol) at as high a final molarity as possible, usually 2 to 4 μ M.

Note that since the rate of [³⁵S]dNTP incorporation is slower than that of [³²P]dNTPs, the reaction should be optimized for maximum incorporation and desired DNA length. Incorporation should be as high as 5×10^8 cpm/ μ g.

SUPPORT PROTOCOL 3

SYNTHESIS OF DIGOXIGENIN-LABELED RNA PROBES

Antisense riboprobes are synthesized as run-off transcripts from linearized templates, using bacteriophage polymerases and template DNA consisting of the DNA fragment of interest cloned in a vector containing the promoter appropriate for the RNA polymerase (T3, T7, or SP6). RNA synthesis is carried out in the presence of a digoxigenin-substituted ribonucleotide. Nonradioactive probes have several advantages. They are easily synthesized in large quantities, they are stable for several months, and they can be reused up to three times. An additional advantage of RNA versus DNA probes is that they result in cleaner signals because nonspecifically bound probe is removed during ribonuclease treatment. This protocol is based on Wilkinson (1993).

Materials

Distilled water, sterile
10× transcription buffer: 400 mM Tris-Cl (pH 8.25)/60 mM MgCl₂/20 mM spermidine (Boehringer Mannheim)
0.2 M dithiothreitol (DTT; APPENDIX 2)
Nucleotide mix, pH 8: 10 mM GTP/10 mM ATP/10 mM CTP/6.5 mM UTP/3.5 mM digoxigenin-UTP (Boehringer Mannheim)
1 µg/ml linearized plasmid
100 U/ml placental ribonuclease inhibitor (RNasin, Boehringer Mannheim)
10 U/µl SP6, T3, or T7 RNA polymerase
TBE electrophoresis buffer (see recipe)
RNase-free DNase I
TE buffer, pH 8 (APPENDIX 2), DEPC-treated (UNIT 4.1)
4 M LiCl, DEPC-treated
70% ethanol in DEPC-treated water
100% ethanol
37°C water bath
Additional reagents and equipment for agarose gel electrophoresis (Voytas, 1988)

1. Mix reagents in the following order at room temperature (21 µl total):

13 µl sterile distilled water
2 µl 10× transcription buffer
1 µl 0.2 M DTT
2 µl nucleotide mix
1 µl linearized plasmid
1 µl RNasin
1 µl RNA polymerase.

Incubate 2 hr at 37°C.

Although stock solutions of sensitive reagents should be kept on ice, the reaction mixture should be mixed at room temperature; otherwise the spermidine in the transcription buffer will precipitate the DNA.

2. Remove 1 µl of the reaction and run on an agarose gel in TBE electrophoresis buffer to estimate amount of riboprobe synthesized.

The RNA band should be ~10 times more abundant than the plasmid band, indicating that ~10 µg of riboprobe have been synthesized.

3. Add 2 µl RNase-free DNase I to the reaction mix. Incubate 15 min at 37°C.
4. Add 100 µl TE buffer, 10 µl of 4 M LiCl, and 300 µl of 100% ethanol. Mix and incubate 30 min at -20°C.

LiCl precipitates high-molecular-weight RNA.

5. Microcentrifuge 10 min at maximum speed, 4°C.
6. Wash pellet with 70% ethanol.
7. Air-dry pellet (do not use a Speedvac evaporator). Redissolve pellet in TE buffer to give a concentration of ~0.1 µg/µl. Store at -20°C.

Digoxigenin-labeled riboprobes are stable at least 2 years at -20°C.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Blocking solution

Preheat 400 ml of 1× PBS (see APPENDIX 2) to 45°C. Immediately before use, add the following: 0.617 g DTT, 0.74 g iodoacetamide, and 0.5 g *N*-ethylmaleimide. Mix well and use immediately. Cover with aluminum foil.

CAUTION: *These substances are very toxic—exercise proper precautions during preparation and handle very carefully.*

Deionized formamide

Mix 50 ml of good-quality formamide (e.g., Fluka) and ~5 g of mixed-bed, ion-exchange resin (e.g., Bio-Rad AG 501-X8, 20 to 50 mesh) and stir 30 min at room temperature. Filter through Whatman filter paper, dispense into 1-ml aliquots, and store at –20°C.

DNA wash solution

0.6 M NaCl
10 mM Tris·Cl, pH 7.5 (APPENDIX 2)
1 mM EDTA (APPENDIX 2)
50% formamide (use inexpensive grade, e.g., Fluka, and do not deionize)
0.1% 2-mercaptoethanol

10 mM DTT/1× PBS

Preheat 400 ml of 1× PBS (APPENDIX 2) to 45°C; immediately before use, add 0.617 g dithiothreitol (DTT) and dissolve.

Use highest quality DTT available (Sigma).

Hybridization mix A

50% deionized formamide (see recipe above for deionization)
0.3 M NaCl, sterile
10 mM Tris·Cl, pH 8.0 (APPENDIX 2)
1 mM EDTA (APPENDIX 2)
1× Denhardt's solution (APPENDIX 2)
500 µg/ml yeast tRNA
500 µg/ml poly(A) (Amersham Pharmacia Biotech)
50 mM DTT
10% polyethylene glycol (MW 6000; EM Science)
Store in aliquots at –70°C

Hybridization mix B

1.2 M NaCl
20 mM Tris·Cl, pH 7.5 (APPENDIX 2)
4 mM EDTA (APPENDIX 2)
2× Denhardt's solution (APPENDIX 2)
1 mg/ml yeast tRNA
200 µg/ml poly(A) (Amersham Pharmacia Biotech)
Store in aliquots at –20°C

Moist chamber solution A

50% formamide (inexpensive grade)
0.3 M NaCl
10 mM Tris·Cl, pH 8.0 (APPENDIX 2)
1 mM EDTA (APPENDIX 2)

Moist chamber solution B

50% formamide (inexpensive grade)
0.6 M NaCl
10 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
2 mM EDTA (*APPENDIX 2*)

Predigested pronase

Predigest pronase by incubating a 40 mg/ml pronase solution in water for 4 hr at 37°C. Lyophilize in aliquots and store in a nondefrosting freezer at –20°C.

Determine the optimal pronase concentration by digesting a series of sections with sequential dilutions of the predigested pronase as follows: resuspend pronase in 50 mM Tris·Cl, pH 7.5/5 mM EDTA (see *APPENDIX 2* for stock solutions). For the Basic Protocol, incubate slides 15 min at 37°C with 0.1 to 10 µg/ml pronase. For the Alternate Protocol, incubate slides 10 min at room temperature with 0.125 to 1 mg/ml pronase. Pretreat experimental sections with the highest possible pronase concentration that allows the retention of adequate cellular morphology.

Restriction enzyme–digested plasmid, 1 µg/µl

Insert gene of interest into a vector containing an SP6 (e.g., pSP64 or pSP65), T3, or T7 promoter. Linearize the plasmid downstream of the coding sequence. Phenol/chloroform extract the DNA (*UNIT 2.1A*), ethanol precipitate, wash pellet in 70% ethanol, dry, and redissolve at 1 µg/µl in sterile water. Alternatively, treat with 0.1% fresh diethylpyrocarbonate (DEPC) for 10 min at room temperature, heat to 65°C for 10 min, and ethanol precipitate as above.

The DNA should be clean and free of salt after the restriction digestion in order to avoid terminating transcription.

RNase digestion solution

40 µg/ml RNase A (Sigma)
2 µg/ml RNase T1 (Sigma)
10 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
5 mM EDTA (*APPENDIX 2*)
0.3 M NaCl

RNA wash solution I

2× SSC (*APPENDIX 2*)
50% formamide (use inexpensive grade, e.g., Fluka, and do not deionize)
0.1% 2-mercaptoethanol

RNA wash solution II

0.1× SSC (*APPENDIX 2*)
1% 2-mercaptoethanol

S-riboprobe competitor (see Basic Protocol)

S-riboprobe competitor is synthesized using a vector without an insert as template. The [³⁵S]UTP in the transcription reaction (see Support Protocol 1, step 1) is replaced by 100 nM nonradioactive S-UTP (NEN Life Sciences). Half of a reaction of S-riboprobe competitor is required per ³⁵S-labeled riboprobe reaction.

TBE electrophoresis buffer, 10×

108g Tris base (890 mM)
55 g boric acid (890 mM)
960 ml H₂O
40 ml 0.5 M EDTA, pH 8.0 (20 mM final; *APPENDIX 2*)

Transcription buffer, 5×

200 mM Tris·Cl, pH 8.3 (APPENDIX 2)
30 mM MgCl₂
10 mM spermidine·Cl
0.1% Triton X-100
Store at -70°C in aliquots

Triethanolamine (TEA) buffer (0.1 M TEA)

Add 18.57 g triethanolamine·Cl to 900 ml water. Dissolve and adjust pH to 8.0 with NaOH. Adjust to 1 liter with water for 0.1 M solution. Prepare fresh daily.

Wash solution A

50% formamide (inexpensive grade, e.g., Fluka)
2× SSC (APPENDIX 2)
20 mM 2-mercaptoethanol
Preheat half to 55°C and half to 50°C

Wash solution B

50% formamide (inexpensive grade, e.g., Fluka)
2× SSC (APPENDIX 2)
20 mM 2-mercaptoethanol
0.5% Triton X-100
Preheat to 55°C

Wash solution C

2× SSC (APPENDIX 2)
20 mM 2-mercaptoethanol
Preheat to 50°C

COMMENTARY**Background Information**

In situ hybridization was first used to localize specific DNA sequences on chromosomes (Gall and Pardue, 1971). This technique has been modified for the detection and localization of RNAs within tissues from many different organisms. Alternative protocols for in situ hybridization studies using organisms other than mouse are listed below. Early studies described the localization of viral or abundant cellular messages in cultured cells or tissue sections (e.g., Brahic and Haase, 1978; Singer and Ward, 1982). Improved methods allowing detection of moderate to low abundance mRNAs were developed later in order to describe the expression and localization of genes regulating embryonic and larval development in *Drosophila*. The sensitivity of these methods, which utilized double-stranded (Hafen et al., 1983) or single-stranded (Akam, 1983) tritiated DNA probes, has been increased further by the use of single-stranded RNA probes (Cox et al., 1984) and ³⁵S-labeled probes (Awgulewitsch et al., 1986).

Nonradioactive techniques for the detection of cellular messages have also been developed.

Commonly used methods employ digoxigenin- and biotin-conjugated nucleotides to label probes which are then detected by either a fluorescent or an enzymatic system (Singer and Ward, 1982; Singer et al., 1986; Haramis and Carrasco, 1996). Digoxigenin-labeled probes are particularly useful for whole-mount in situ hybridization. Although these methods offer increased resolution and rapid detection while avoiding the safety and disposal problems of radioactive isotopes, they can be less sensitive on sections than autoradiographic techniques.

Successful in situ hybridization to cellular RNA relies on a number of factors. The most suitable cytological fixation procedure should provide adequate cellular morphology, while neither extracting nor changing the location of the RNA. Moreover, fixation must not be so thorough that it inhibits probe access. The most widely used fixative is paraformaldehyde, although many other fixatives have also been used successfully. The length of fixation time and type of fixative should be empirically determined for the tissue of interest and for any other cytological procedures (e.g., immunolo-

calization) required. Limited proteolytic digestion is commonly used to increase probe access to fixed tissues. The degree of proteolysis must be carefully balanced to permit adequate morphology while increasing signal, and often digestion may be unnecessary (Lawrence and Singer, 1985; Pardue, 1985).

The probe may be either DNA or RNA. Randomly sheared double-stranded DNA probes labeled by either nick translation or oligonucleotide-primed labeling may form networks on the cytological hybrid, thereby increasing the hybridization signal. Self-annealing of probe also occurs, however, thus decreasing the effective probe concentration. This problem has been circumvented by synthesizing single-stranded DNA probes from M13 templates (Akam, 1983), but the probes most commonly used today are single-stranded RNA probes transcribed by phage polymerases from plasmids containing SP6 or T7 promoters. These highly specific radioactive probes are easily synthesized in large quantities. RNA probes generally give lower backgrounds than equivalent DNA probes because nonspecifically bound probe is removed by ribonuclease treatment. Several studies have shown that probe length affects hybridization efficiency with probes of about 30 to 300 nucleotides being optimal. DNA probe length may be decreased by DNase I digestion, while RNA probes may be shortened by limited alkaline hydrolysis (Cox et al., 1984). Length of RNA probes does not appear to be as critical as that of DNA probes, and the protocols described here do not involve hydrolysis at all.

Acetylation of tissue preparations is frequently used to decrease background binding of radioactive probe when performing in situ hybridization to chromosomal DNA (Hayashi et al., 1978). This treatment is equally important when hybridizing to cellular mRNAs and is recommended in all cases. In the Basic Protocol an additional step for blocking nonspecific sulfur-binding sites on tissue sections is included (Zeller et al., 1987).

Critical Parameters

Many particularly critical points are noted throughout the hybridization protocol of which the following are noteworthy. The most sensitive parameter, tissue fixation, requires careful optimization in order to maximize the signal-to-noise ratio. Overfixation can entirely obscure weak signals.

It is essential that freshly prepared solutions are used where indicated since many reagents,

blocking agents in particular, are extremely labile. Moreover, be sure to purchase only as much chemical as can be used in a reasonable period of time.

[³⁵S]NTPs, the radionucleotides most commonly used for labeling in situ hybridization probes, are treated in essentially the same way as [³²P]NTPs, with a few important exceptions. In general, any factor that makes a [³²P]NTP reaction slightly substandard will almost abolish [³⁵S]NTP incorporation, presumably because thionucleotides are less easily utilized by the enzymes. Consequently, great care should be taken to use clean restriction endonuclease-digested plasmid DNA, fresh [³⁵S]NTPs (<2 weeks old), and fresh aliquots of cold NTPs. Most importantly, make sure that all solutions contain ≥10 mM DTT and that RNase contamination is eliminated.

The successful in situ hybridization experiment requires thoughtful design and suitable controls. For the initial experiments the following guidelines are helpful: (1) If possible, include a tissue in your experiments which expresses the gene of interest at high levels in a RNase protection or northern assay (UNIT 4.9). (2) Use at least two different fixation time points which have been found to yield high-quality sections (see UNITS 14.1 & 14.2). (3) To optimize the signal-to-noise ratio, use different hybridization times (30 min, 1 hr, 2 hr, 4 hr). (4) Process at least duplicate slides for each fixation and hybridization condition to get two different exposure times by emulsion autoradiography (UNIT 14.4). (5) Initially, the experimenter should be conservative in estimating how many slides can be processed.

The following controls should be done to confirm the spatial distribution pattern: First, and most importantly, hybridize one set of slides to an antisense single-stranded riboprobe (see Support Protocol 1) to detect the specific hybridization to the complementary mRNA, as well as a second set under absolutely identical conditions to a sense riboprobe (synthesized from the same template as the antisense riboprobe in opposite orientation) as a control to detect nonspecific background. It is advisable to determine that there is no transcription of the opposite strand using an RNase protection assay.

As a second control, treat sections with RNase prior to hybridization to the antisense probe in order to reveal signal resulting from binding to non-RNA cell components. Finally, use another portion of the gene, expecting to get the same result (this control is important

Table 14.3.1 Time Parameters for In Situ Hybridization

Step	Paraffin	Cryosections
Fixation and embedding	Day 1—Fix or perfuse tissues in 4% PFA fixative. Dehydrate.	Day 1—Fix or perfuse tissues in 4% PFA fixative. Infuse in sucrose. Orient and embed in OCT compound. Quick freeze and store at -70°C .
Sectioning	Days 3 to 4—Cut sections on microtome and place on slides. Dry on slides for 1 to 2 days at 42°C . Store desiccated ≤ 2 months at -20°C .	Day 2—Make cryostat sections and place on microscope slides. Fix with 4% PFA fixative, dehydrate, and store desiccated ≤ 1 month at -70°C .
Pretreatment	Day 4—Dewax. HCl and heat treat. Treat with blocking agents and acetic anhydride. Dehydrate and store desiccated overnight at -70°C .	Day 3—Pronase digest. Treat with acetic anhydride. Dehydrate and use immediately.
Hybridization	Day 5—Overlay sections with probe and hybridize in moist chamber 1 to 4 hr.	Day 3—Overlay sections with probe and hybridize in moist chamber 1 to 4 hr.
Washes	Day 5—Wash, RNase-treat, and dehydrate.	Day 3—Wash, RNase-treat, and dehydrate.
Detection	Days 5-?—Expose to X-ray film, 1 to 3 days. Estimate exposure time and dip multiple slides in autoradiographic emulsion. Develop and analyze.	Days 3-?—Expose to X-ray film, 1 to 3 days. Estimate exposure time and dip multiple slides in autoradiographic emulsion. Develop and analyze. Alternatively, detect digoxigen-labeled probe enzymatically 1 to 2 days

when the gene is a member of a multigene family), and compare the expression pattern to that of a completely unrelated gene.

Troubleshooting

Sections get lost during procedure

- Slides were subbed or coated improperly.
- Tissue requires more gentle fixation; use a moist chamber throughout fixation, and the first dehydration (cryosections only).

Low signal, high background

- Tissue is overfixed; use shorter fixation time.
- Blocking agents are old; buy small quantities, which can be used up rapidly.
- $[^{35}\text{S}]\text{UTP}$ is oxidized; use a fresh batch, store it properly at -70°C , and check that enough DTT (or 2-mercaptoethanol) is present in all solutions. Keep freezing/thawing to a minimum.

d. Overdigestion with pronase (cryosections only).

e. Certain sequences of some genes are associated with high background; use a different portion of the gene as probe.

Localized high levels of background

Certain cell types are very sticky for ^{35}S -riboprobes. If these cells are a minor component of the tissue of interest, use sense-strand control riboprobe to document that the particular signal is nonspecific. Alternatively, digest sections with higher concentrations of pronase.

Comparison of paraffin and cryosections

Advantages of paraffin wax (Paraplast). Easier to get satisfactory sections and good morphology; routine histology, antigen localization using enzymatic detection, and in situ hybridizations can be performed on serial sections; blocks can be stored indefinitely at room temperature; has been shown to work with

tissues from many different organisms; paraffin sectioning is routine in clinical settings.

Advantages of cryosections. Quicker; coimmunolocalization on the same section is possible; and immunolocalization using fluorescent tags is possible. Finer morphological studies are possible (e.g., electron microscopy) because the tissues are treated more gently.

Anticipated Results

A good in situ hybridization experiment will clearly identify the location of a message within specific cells or tissues with minimal background throughout the remainder of the tissue. Numerous examples of these experiments can be found in the listed references.

Time Considerations

Table 14.3.1 details how long it takes to complete an in situ hybridization experiment, starting with fixing the tissue and finishing with exposure of the hybridized slide. It is advisable that probes be synthesized the day before hybridization is performed.

Literature Cited

- Akam, M.E. 1983. The location of Ultrabithorax transcripts in *Drosophila* tissue sections. *EMBO J.* 2:2075-2084.
- Awgulewitsch, A., Utset, M.F., Hart, C.P., McGinnis, W., and Ruddle, F.H. 1986. Spatial restriction in expression of a mouse homeobox locus within the central nervous system. *Nature* 320:328-335.
- Brahic, M. and Haase, A.T. 1978. Detection of viral sequences of low reiteration frequency by in situ hybridization. *Proc. Natl. Acad. Sci. U.S.A.* 75:6125-6129.
- Cox, K.H., DeLeon, D.V., Angerer, L.M., Angerer, R.C. 1984. Detection of mRNAs in sea urchin embryos by in situ hybridization using asymmetric RNA probes. *Dev. Biol.* 101:485-502.
- Gall, J.G. and Pardue, M.L. 1971. Nucleic acid hybridization in cytological preparations. *Meth. Enzymol.* 38:470-480.
- Hafen, E., Levine, M., Garber, R.L., Gehring, W.J. 1983. An improved in situ hybridization method for the detection of cellular RNA in *Drosophila* tissue sections and its application for localizing transcripts of the homeotic Antennapedia complex. *EMBO J.* 2:617-623.
- Haramis, A.G. and Carrasco, A.E. 1996. Whole-mount in situ hybridization and detection of RNAs in vertebrate embryos and isolated organs. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 14.9.1-14.9.18. John Wiley & Sons, New York.
- Hayashi, S., Gilam, I.C., Delaney, A.D., Tener, G.M. 1978. Acetylation of chromosome squashes of

Drosophila melanogaster decreases the background in autoradiographs from hybridizations with ¹²⁵I-labeled RNA. *J. Histochem. Cytochem.* 26:677-679.

- Lawrence, J.B. and Singer, R.H. 1985. Quantitative analysis of in situ hybridization methods for the detection of actin gene expression. *Nucl. Acids Res.* 13:1777-1799.
- Pardue, M.L. 1985. In situ Hybridization. *In* Nucleic Acid Hybridization: A Practical Approach (B.D. Hames and S.J. Higgins, eds.) pp. 179-202. IRL Press, Oxford.
- Singer, R.H., Lawrence, J.B., Villnave, C. 1986. Optimization of in situ hybridization using isotopic and nonisotopic methods. *BioTechniques* 4:230-250.
- Singer, R.H. and Ward, D.C. 1982. Actin gene expression visualized in a chicken muscle tissue culture by using in situ hybridization with a biotinylated nucleotide analog. *Proc. Natl. Acad. Sci. U.S.A.* 79:7331-7335.
- Wilkinson, D.G. 1993. In situ hybridization. *In* Essential Development Biology. A Practical Approach (C.D. Stern and P.W. H. Hollan, eds.) pp. 257-276. IRL Press, Oxford.

Voytas, D. 1988. Agarose gel electrophoresis. *In* Current Protocols in Molecular Biology (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds.) pp. 2.5.1-2.5.9. John Wiley & Sons, New York.

Zeller, R., Bloch, K.D., William, B.S., Arceci, R.J., and Seidman, C.E. 1987. Localized expression of the atrial natriuretic factor gene during cardiac embryogenesis. *Genes Dev.* 1:693-698.

Additional References

Chicken

Hayashi, M., Ninomiya, Y., Parsons, J., Hayashi, K., Olsen, B.R., and Trelstad, R.L. 1986. Differential localization of mRNAs of collagen types I and II in chick fibroblasts, chondrocytes, and corneal cells by in situ hybridization using cDNA probes. *J. Cell Biol.* 102:2302-2309.

Wedden, S.E., Pang, K., and Eichele, G. 1989. Expression pattern of homeobox-containing genes during chick embryogenesis. *Development* 105:639-650.

Drosophila

Hafen, E. and Levine, M. 1986. The localization of RNAs in *Drosophila* tissue sections by in situ hybridization. *In* *Drosophila: A Practical Approach* (D.B. Roberts, ed.) pp. 139-157. IRL Press, Oxford.

Human tissues and the use of ¹²⁵I isotope

Pfeifer-Ohlsson, S., Goustin, A.S., Rydner, J., Bjersing, L., Wahlstrom, T., Stehelin, D., Ohlsson, R. 1984. Spatial and temporal pattern of cellular *myc* oncogene expression in developing human placenta: Implications for embryonic cell proliferation. *Cell* 38:585-596.

Nematodes

Costa, M., Weir, M., Coulson, A., Sulston, J., and Kenyon, C. 1988. Posterior pattern formation in *C. elegans* involves position-specific expression of a gene containing a homeobox. *Cell* 55:747-756.

Xenopus laevis

Carrasco, A.E. and Malacinski, G.M. 1987. Localization of *Xenopus* homeobox gene transcripts during embryogenesis and in the adult nervous system. *Dev. Biol.* 121:69-73.

Kintner, C.R., and Melton, D.A. 1987. Expression of *Xenopus* N-CAM RNA in ectoderm in early response to neural induction. *Development* 199:311-325.

Key References

Hogan, B., Constantini, F., Lacy, E. 1986. Manipulating the Mouse Embryo. Cold Spring Harbor, N.Y.

Pardue, 1985. See above.

Zeller et al., 1987. See above.

The Basic Protocol for Paraplast wax-embedded samples is entirely based on Zeller et al., 1987.

Contributed by Rolf Zeller
University of Utrecht
The Netherlands

Melissa Rogers
University of South Florida
Tampa, Florida

Anna G. Haramis (digoxigenin labeling)
University of Utrecht
The Netherlands

Andrés E. Carrasceo (digoxigenin labeling)
Buenos Aires Medical School
Buenos Aires, Argentina

Detection of Hybridized Probe

UNIT 14.4

Autoradiography is used to detect and quantitate radioactive probe hybridized to cytological preparations. Autoradiographic film is used to detect ^{32}P - or ^{35}S -labeled probe, and can be useful in experiments dealing with large organs or tissues. Emulsion autoradiography is required to obtain resolution at the level of a single cell. For further information on film autoradiography, see *APPENDIX 3*.

FILM AUTORADIOGRAPHY

Tape slides to a backing such as cardboard or an old piece of film. Expose slides to Du Pont Cronex Video Imaging Film (MRF 34 Clear) at 4°C under light pressure. Kodak XRP-1 may also be used.

When using Du Pont Cronex Video Imaging Film make sure that the emulsion side of the film is facing the specimens.

**BASIC
PROTOCOL**

EMULSION AUTORADIOGRAPHY

Single-cell resolution and ^3H detection can be obtained by coating slides with autoradiographic emulsion.

**BASIC
PROTOCOL**

Materials

- Diluted Kodak emulsion (support protocol)
- Kodak D19 developer
- Kodak fixer
- Plastic cytology slide mailer (Curtin Matheson), for use as dipping chamber
- Nonsparking fan or slide dryer (Oncor; optional)
- Kodak safelight filter #2 (optional)
- Black, light-tight slide boxes
- Desiccant (e.g., Humicaps, United Desiccants-Gates)
- 42° to 45°C water bath
- Slide racks and jars, for use in developing and fixing
- 15° to 20°C water bath (a styrofoam box works well)
- Slide rack or wire test-tube rack for drying dipped slides

Steps 1 to 7 must be performed in complete darkness, or ~4 feet from safelight.

1. Melt aliquot of diluted emulsion (prepared according to support protocol) in 42° to 45°C water bath for 10 min. Pour or pipet emulsion into clean slide mailer (dipping chamber).
2. Dip slides slowly and smoothly into slide mailer or dipping chamber. Withdraw slowly and place vertically in a test-tube rack to dry 2 hr.
Alternatively, use a nonsparking fan to decrease drying time, or place in a slide dryer and dry according to manufacturer's instructions. With tissue sections artifacts caused by running of emulsion can be a problem. Prevent this by placing the dipped slides on a cold glass plate for 10 min before drying the slides.
3. Place thoroughly dry slides in a light-tight slide box with desiccant. Seal box with electrician's black tape, cover with foil, and expose at 4°C .

If slides are not completely dry sensitivity is decreased, as humidity causes fading of the latent images in the autoradiographic emulsion.

Do not expose in a refrigerator that is used for storing ^{32}P because even very small quantities of ^{32}P will increase background. Organic solvents will also increase background.

**In situ
Hybridization
and Immuno-
histochemistry**

Contributed by Melissa Rogers

Current Protocols in Molecular Biology (1989) 14.4.1-14.4.3

Copyright © 2000 by John Wiley & Sons, Inc.

14.4.1

Supplement 7

- Put developer, water, and fixer in slide jars and bring to 15° to 20°C in water bath.

The exact temperature is not critical. Lowering the temperature, however, reduces the silver grain size; 15°C is recommended.

- Take slides out of the refrigerator and allow to warm to same temperature as developing solutions.

Slides and developing solutions should be at the same temperature since drastic temperature changes will cause the emulsion to wrinkle.

- Transfer the slides to slide racks and develop in a light-tight darkroom as follows (monitor time carefully with a timer):

2.5 min in developer
30 sec in water
3 min in fixer.

Kodak does not recommend using an acid stop bath because it may cause formation of microscopic bubbles in the emulsion. Use of rapid (ammonium) fixers may result in the loss of developed silver grains. Do not agitate slides vigorously.

- Rinse slides 10 to 15 min under gently running cool tap water in the light, then once in cool distilled water. While slides are still wet, scrape off emulsion on the back of the slide with a razor, then dry them in a dust-free location.

- Counterstain the slides as desired and mount under clean cover slips (UNIT 14.5).

Developed silver grains are best observed by dark-field microscopy, under which they appear as white dots on a black background.

SUPPORT PROTOCOL

PREPARATION OF DILUTED EMULSION FOR AUTORADIOGRAPHY

This protocol outlines the preparation of a 1:1 (v/v) dilution of Kodak NTB-2 autoradiographic emulsion for use in emulsion autoradiography. *NOTE:* All steps must be performed in complete darkness or ~4 feet from safelight.

Additional Materials

Kodak NTB-2 autoradiographic emulsion

- Heat a 4-oz bottle of Kodak NTB-2 autoradiographic emulsion in a 42° to 45°C water bath 30 min in the dark. At the same time warm an equal volume of water in a 500-ml flask to the same temperature.
- After emulsion has melted, slowly pour it down the side of the flask (which is held at an angle) and mix. Swirling the flask gently once or twice will sufficiently mix the emulsion and water. Avoid creating any bubbles.
- Aliquot diluted emulsion into nylon scintillation vials or plastic slide mailers.

If the scintillation vial caps have cork inserts, remove them because they may release organic compounds that can increase background. If slide mailers are used, the aliquots must be stored vertically in order to prevent leakage before the emulsion solidifies; however, they are convenient because they may be used for dipping directly.

- Wrap each aliquot in aluminum foil or store in a light-tight container in a refrigerator that is never used for storing ³²P or organic chemicals.

Even very small quantities of ³²P will cause an unacceptable background.

- Test diluted emulsion before use by dipping and developing two clean, unused slides (as described in basic protocol for emulsion autoradiography) and visualizing under a 100× lens. If >100 grains per microscope field are observed, the emulsion should be returned to Kodak.

COMMENTARY

Background Information

The final factor affecting hybridization sensitivity is autoradiographic detection. Two forms of autoradiography are used: film (*APPENDIX 3*) and emulsion.

Emulsion autoradiography is performed by dipping previously hybridized slides in liquid nuclear track emulsion, thus producing a thin film of autoradiographic emulsion that adheres closely to the tissue section. The thickness of this film and the type of isotope used directly affect the resulting sensitivity, resolution, and background of this procedure. Very strong β -particle emitters such as ^{32}P (1710 KeV) produce a strong signal that may be detected by ordinary X-ray film, but the long path length of each particle makes fine tissue and cell resolution impossible. Emulsion autoradiography is rarely used to detect this isotope because of this lack of resolution and because many particles escape from the emulsion before decaying. Tritium, a weak β emitter (18.5 KeV), can only be localized by emulsion autoradiography and is ideal for single-cell resolution. Because the approximate path length of each particle is <1 micron, most particles decay within the emulsion above the cytological preparation. The weak signal, however, makes extremely long exposure times necessary. ^{35}S -labeled probes (167 KeV) are a suitable compromise because they afford increased sensitivity combined with adequate cellular resolution.

Film autoradiography may be used to preview quickly the results of hybridization using a ^{35}S label, thus allowing one to estimate subsequent emulsion exposure time and to determine whether or not an experiment worked as expected. In some cases, where large organs or embryos are labeled, this exposure may provide the required information; however, single-cell resolution requires the use of emulsion autoradiography. As an approximate rule of thumb, the exposure time required for an equivalent signal in emulsion is 3- to 5-fold that of film, although multiple slides must always be used in order to obtain the best exposure.

Critical Parameters

Emulsion must be prepared and used either in complete darkness or at a reasonable distance from a safelight (~4 feet).

After exposing slides to emulsion, be sure to dry slides completely before putting them away to expose. Both the slides and the emul-

sion should never be stored in a cold room or refrigerator that is used for storing ^{32}P or organic solvents. Either material can increase background and ruin an experiment.

In preparing and using emulsion, it is important to avoid forming any bubbles, either in the solution itself or along the slides. Any bubbles in the surface will prevent the slides from being smoothly coated with emulsion and interfere with autoradiography.

Anticipated Results

Film autoradiography provides a rapid but low resolution image of hybridized ^{35}S probe. Fine resolution and detection of tritiated probes are obtained with emulsion autoradiography.

Time Considerations

Film autoradiography requires only a few minutes to place film on the specimens, a 1- or 2-day exposure, and ≤ 10 min developing time. Thus film autoradiography is frequently used to preview the results of an experiment before embarking on the time-consuming process of emulsion autoradiography.

Although the time required to emulsion coat and develop slides is brief, lengthy exposure times may be required. Moreover, it is usually necessary to prepare replicate slides for different exposure times. Preparation of emulsion requires 45 to 60 min. Dipping of slides requires 1 to 3 hr, depending on the availability of a fan or slide dryer. Slides hybridized to ^{35}S probes should be exposed to film for 1 to 2 days, then emulsion exposure time is estimated based on the signal observed; emulsion is about 3- to 5-fold slower. Usual exposure times range from days to weeks, while slides hybridized to tritiated probes may require exposure times of up to several months. Developing takes ~1 hr.

Key Reference

Pardue, M.L. 1985. In situ hybridization. *In* Nucleic Acid Hybridization: A Practical Approach (B.D. Hames and S.J. Higgins, eds.) pp. 179-202. IRL Press, Oxford.

Provides a detailed description of emulsion autoradiography and several photographs of hybridized and exposed specimens.

Contributed by Melissa Rogers
Harvard Medical School and
Dana-Farber Cancer Institute
Boston, Massachusetts

Counterstaining and Mounting of Autoradiographed In situ Hybridization Slides

UNIT 14.5

The morphology of specimen sections and the identity of specific areas are defined by lightly counterstaining sections. After staining, slides are dehydrated, mounted with coverslips, hardened, and cleaned for examination under the microscope. In the protocols provided below, Giemsa stains predominantly the nuclei, hematoxylin/eosin stain differentiates both the nuclei and cytoplasm, toluidine blue staining is a simpler procedure that lightly stains both nuclei and cytoplasm, and Hoechst staining of nuclei provides a fast, easy, and effective way to simultaneously view the entire tissue and the regions of hybridization.

NOTE: Carry out all steps at room temperature except where specified otherwise.

GIEMSA STAINING

**BASIC
PROTOCOL**

Materials

Hydrated, developed in situ hybridization slides (UNITS 14.3 & 14.4)

Giemsa stain (Fisher)

10 mM NaPO₄ buffer, pH 6.8 (1:50 dilution of 500 mM stock)

50%, 70%, 95%, and 100% ethanol

Xylenes

Mounting medium: Permout (Fisher) or Gelvatol (now called Airvol, from Air Products and Chemicals)

13 glass staining dishes

Blunted forceps

Whatman 3MM paper chips

42°C incubator

CAUTION: Xylenes are toxic organic solvents. All steps using xylenes must be carried out under a fume hood.

1. Immerse developed slides in a slide rack in a staining dish filled with water.
2. Prepare the following regimen of staining dishes filled with solutions:
 - 1 dish—25-fold dilution of Giemsa stain in 10 mM NaPO₄ buffer, pH 6.8
 - 3 dishes—H₂O
 - Dehydration series (can be reused, or use the series prepared for dewaxing in UNIT 14.3)—3 dishes each filled with 50%, 70%, and 95% ethanol, respectively, and 2 dishes filled with 100% ethanol
 - 3 dishes—xylenes.
3. Stain slides 20 sec in a 25-fold dilution of Giemsa stain.

Depending on staining time, a weak or intense stain can be achieved. Optimal staining times vary from tissue to tissue and should also be adjusted to the strength of the in situ hybridization signal. Perform a test series the first time.

4. Rinse slides in water three times, 2 min each time.
5. Dehydrate slides through the ethanol series, 2 min in each dish. Transfer to xylenes and do three changes, 2 min each time.

It is important to perform the dehydration and equilibration in xylenes carefully because Permout will not mix with water or ethanol.

**In situ
Hybridization
and Immuno-
histochemistry**

Contributed by Rolf Zeller and Melissa Rogers

Current Protocols in Molecular Biology (1993) 14.5.1-14.5.5

Copyright © 2000 by John Wiley & Sons, Inc.

14.5.1

Supplement 24

- In the hood, mount slides as follows: With blunted forceps (in left hand), remove one slide from the xylenes and hold it flat at its frosted end. Add ~4 drops Permout to the other end (where sections are located). With the right hand, grab a clean coverslip and place it very slowly on the slide as shown in Figure 14.5.1.

Gelvatol is used instead of Permout when the cytological preparation should not be dehydrated, e.g., in the case of immunohistochemistry. Mount slides with Gelvatol as follows: Place a small drop of gelvatol on a coverslip. Carefully set the slide containing the specimen on the drop, avoiding formation of bubbles. Avoid applying pressure as this may squash the specimen. Allow to harden 2 to 4 hr at room temperature. See reagents and solutions for a homemade recipe for Gelvatol.

It is important that the slide does not start to dry before the mounting medium is added and the coverslip is placed on. Otherwise, microbubbles can cause artifacts.

- Remove extra mounting medium/xylenes by blotting carefully around the edges with 3MM paper and by wiping the back of the slide (but not the coverslip).
At this stage, slides cannot be stored upright in slide boxes because the mounting medium has not yet hardened.
- Place slides flat on a cardboard tray and allow to harden ~2 days in a 42°C incubator.
- Remove excess emulsion, stain, and mounting medium from the backs of slides by carefully scraping with a razor blade; remove dust with lens paper or plain tissue.
- Transfer cleaned slides to slide boxes. If necessary, relabel slides properly.
- Slides may now be examined microscopically. For weak signals, dark-field illumination is recommended. For strong signals, transillumination is sufficient.

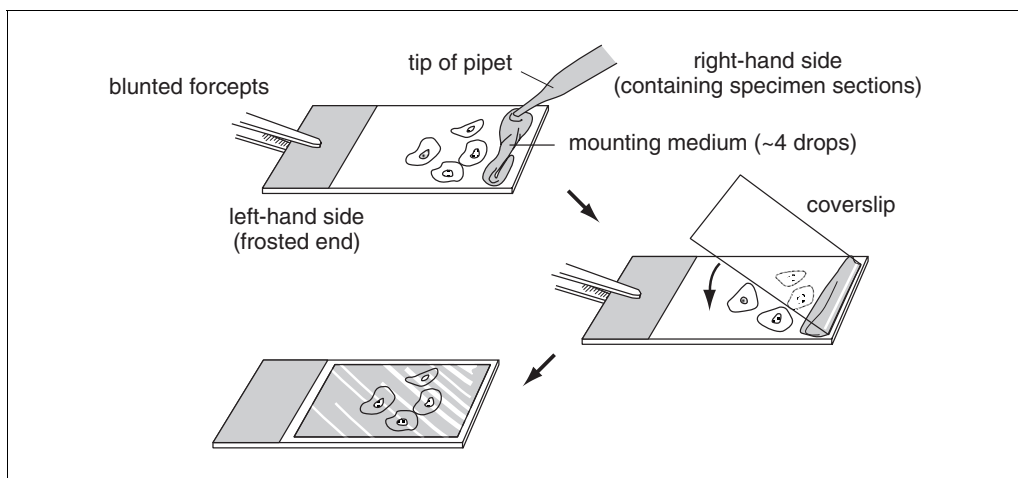


Figure 14.5.1 Mounting of slides with mounting medium.

ALTERNATE PROTOCOL

HEMATOXYLIN/EOSIN STAINING

Additional Materials

- Hematoxylin stain (Harris modified hematoxylin with acetic acid, mercury-free; Fisher)
- 0.1% ammonium hydroxide
- Eosin stain

- Place developed slides (in slide rack) in a staining dish filled with water.
- Prepare the following regimen of solutions in staining dishes:
 - 1 dish—hematoxylin stain
 - 2 dishes—H₂O

- 1 dish—0.1% NH₄OH
- 2 dishes—H₂O
- 1 dish—eosin stain
- 1 dish—95% ethanol
- 2 dishes—100% ethanol
- 3 dishes—xylenes.

3. Stain slides 20 to 30 sec in hematoxylin.

Again, optimal staining time varies a bit from tissue to tissue, and also among batches of hematoxylin. Do not overstain.

4. Rinse two times in water, 2 min each time.
5. Quickly dip into 0.1% ammonium hydroxide, then rinse in water.

If the dip in 0.1% ammonium hydroxide is too long, autoradiographic emulsion may start to come off the slides.

6. Wash 5 min in water, then eosin stain 20 to 30 sec.

If oversteining with eosin occurs, destaining can be achieved by dipping the slides in 50% ethanol for a short time. Watch the stain diffusing out. If the stain is too weak, restain again.

7. Dip slides eight times in the same 95% ethanol solution, then dip them again eight times in the same 100% ethanol solution.
8. Fully dehydrate 2 min in 100% ethanol.
9. Equilibrate three times, 2 min each, in xylenes.
10. Mount and process as described in steps 6 to 11 of the basic protocol.

TOLUIDINE BLUE STAINING

Additional Materials

Toluidine blue stain

1. Dilute Toluidine blue stain with water.

The dilution should be determined empirically, according to the degree of staining desired. Start at ~1:100.

2. Dip slides briefly in diluted stain solution. Rinse several times with water and mount as desired.

HOECHST STAINING

Additional Materials

1 mg/ml Hoechst stain in dimethyl sulfoxide (Hoechst 33258 dye, bisbenzimidazole; store at -20°C)

Mounting medium: 0.5 g/ml Canada balsam or Gelvatol

Fluorescence microscope with Hoechst or DAPI filter and dark-field optics

CAUTION: Hoechst dye is a carcinogen.

1. Dilute 1 mg/ml Hoechst stain 1:500 with water to 2 µg/ml final.
2. Cover sections with diluted stain or dip slides in diluted stain. Incubate 2 min at room temperature.

**ALTERNATE
PROTOCOL**

**ALTERNATE
PROTOCOL**

**In situ
Hybridization
and Immuno-
histochemistry**

14.5.3

3. Wash slides 2 min in water at room temperature.
4. Air dry slides at room temperature.
5. Bake slides 30 to 60 min in a 42° to 55°C incubator. Remove and bring to room temperature.
6. Mount as in Figure 14.5.1 using 0.5 g/ml Canada balsam and air dry the slides 2 days at room temperature (after mounting).

Do not use Permount because autofluorescence occurs when nuclei are viewed. Canada balsam clears paraffin sections well and does not fluoresce under UV irradiation. Gelvatol may be used for cryosections but should not be used for paraffin sections since it does not clear the tissues.

7. View nuclei using Hoechst epifluorescence optics. Cover the light source with a red filter and view silver grains simultaneously with dark-field illumination.

The silver grains will appear red and contrast well with the fluorescent blue nuclei. To photograph, make a double exposure of the nuclei and the silver grains, optimizing the exposure time required for both nucleus and silver grain visualization.

REAGENTS AND SOLUTIONS

0.5 g/ml Canada balsam

5 g Canada balsam
 10 ml methyl salicylate
 Stir until Canada balsam is dissolved
 Store at room temperature

Eosin stain

12 g Eosin Y (Fisher)
 3 g Phloxine B (Fisher)
 500 ml 70% ethanol

Gelvatol (now called Airvol)

Gelvatol (Air Products and Chemicals) is a water-soluble mounting medium made from polyvinyl alcohol (PVA) and glycerol. The following recipe is a simplification of the original method: Add 5 g PVA 2000 to 100 ml 1× PBS (APPENDIX 2) every hour for 4 hr (total 20 g) while stirring constantly. Keep covered and stir overnight at 4°C. Add 3 more grams PVA 2000 and stir until dissolved. Add a single sodium azide crystal and 50 ml glycerol. Mix thoroughly, aliquot, and store at 4°C in sealed storage vials.

500 mM NaPO₄ buffer, pH 6.8 (stock solution)

35.5 g Na₂HPO₄
 34.5 g NaH₂PO₄
 H₂O to 1 liter
 Adjust pH if necessary

Toluidine blue stain

1.0 g sodium borate
 0.5 g Toluidine blue
 H₂O to 100 ml
 Filter before use

COMMENTARY

Background Information

The histological stains described here are used to differentiate cellular morphology after completing in situ hybridization (UNITS 14.1-14.4) or immunolocalization (UNIT 14.6) procedures. Counterstaining sectioned material is necessary to positively identify the specific structures or cell types in an embryo or tissue expressing the gene of interest. Choice of histological stain depends on several factors. In general, a histological stain that is commonly used in morphological analysis of the particular tissue or cell type is suggested. For adult tissues, hematoxylin/eosin staining is preferred, whereas Giemsa and Toluidine blue stains are mostly used to stain embryos. Hoechst staining of nuclei, combined with dark-field illumination, permits simultaneous visualization of exposed silver grains and nuclei and allows the viewer to determine exactly which cells have hybridized to the probe (Edgar and O'Farrell, 1989).

For more detailed information and the use of other histological stains the reader is referred to Luna (1968), which shows expected results with color plates.

Critical Parameters

For any of the four staining procedures, the optimal staining time will depend upon the tissue as well as the type of stain used, and should be adjusted to the intensity of the hybridization signal.

Thorough exposure to xylenes after dehydration is important because Permount mounting medium will not mix with water or ethanol. When the specimen should not be dehydrated, Gelvatol can be used instead.

Canada balsam or Gelvatol mounting media must be used with Hoechst-stained tissues since Permount autofluoresces. Paraffin sections should be mounted with Canada balsam in order to clear the tissues adequately.

Troubleshooting

If the stains are too weak, sections can be restained by removing Permounted coverslips in xylenes, washing the slides extensively in xylenes, and then rehydrating them as described in UNIT 14.3. Gelvatol-mounted coverslips may be removed by soaking in water.

Overstaining for hematoxylin is irreparable, whereas both Giemsa and eosin stains can be destained in low ethanol concentrations. To avoid under- or overstaining, it is advisable to go through the procedure the first time with very few slides and then to adjust the staining times according to the results obtained.

Anticipated Results

Ideally, histological staining reveals the details of cellular morphology thus allowing identification of the cell types expressing the gene of interest but not interfering with visualization of the grains of the autoradiographic exposure (UNIT 14.4). Certain stains (e.g., hematoxylin/eosin) have the tendency to nonspecifically stain the autoradiographic emulsion, which can obscure the hybridization signal. In general, shorter times are used for staining sections after in situ hybridization or immunolocalization than for conventional histological analysis.

Time Considerations

These staining procedures are very fast and should not take more than 10 to 15 min to complete. Mounting time depends on experience and number of slides processed, although 20 slides should easily be mounted in 30 min. After mounting, several hours (Gelvatol) to 2 days (Permount and Canada balsam) are required for hardening.

Literature Cited

Edgar, B.A. and O'Farrell, P.H. 1989. Genetic control of cell division patterns in the *Drosophila* embryo. *Cell* 57:177-187.

Luna, L.G. 1968. Manual of Histologic Staining: Methods of the Armed Forces Institute of Pathology (3rd ed.). McGraw-Hill, New York.

Provides a detailed description of many histological stains.

Contributed by Rolf Zeller
European Molecular Biology Laboratories
Heidelberg, Germany

Melissa Rogers
Harvard Medical School
and Dana-Farber Cancer Institute
Boston, Massachusetts

Immunohistochemistry

This unit surveys immunohistochemical techniques for protein localization. The first four protocols focus on indirect immunofluorescence which is the easiest and most widely used method of optical detection. In the following protocols alternate methods suitable for bright-field microscopy, such as immunoperoxidase and immunogold labeling, are described and critically reviewed. The commentary contains an extensive troubleshooting guide with a list of suitable routine controls.

The chosen method of tissue or cellular preparation must be suitable to answer the fundamental objective of the protocol (i.e., where the protein is localized within the system). Thus, if cultured cells are being studied, and it is suspected that the protein exists at the cell's surface, it would be pointless to employ a protocol designed to look at cytoplasmic contents and which includes an intrinsic step that may disrupt the cell membrane. Furthermore, the choice of fluorochrome and the method of labeling are also important to optimize the sensitivity of antigen localization. The basic protocol is therefore divided into two sections: (1) cellular preparations (monolayers of adherent cells and cells grown in suspension) and (2) tissue section labeling.

The commentary discusses suitable routine control experiments that should be performed for each experiment. Advantages and disadvantages of monoclonal and polyclonal primary antibodies and secondary labeling methods, including secondary antibody-fluorochrome conjugates, protein A-fluorochrome conjugates, and antibody-biotin/streptavidin-fluorochrome conjugates, are also discussed. It is recommended that the commentary be read before starting any of the labeling protocols.

BASIC PROTOCOL

IMMUNOFLUORESCENT LABELING OF CELLS GROWN AS MONOLAYERS

Immunocytochemical studies are easily performed on adherent cultured cells. Cells are generally grown in culture dishes, on glass cover slips, or in sterile tissue culture chambers on glass slides (see Table 14.6.1 in commentary). Immobilization of the antigen varies depending on whether the antigen is considered to be associated with the cell surface membrane or contained within the cytoplasm. Low concentrations of paraformaldehyde are used to fix surface antigens. This will generally not destroy antigenicity, and will preserve cell morphology. If fixation does remove antigenic activity, one may use the same protocol and eliminate the fixation step, provided that all incubations are performed at 4°C.

When cytoplasmic antigens are to be studied, the membrane must be permeabilized to allow antibodies into the cell. Two methods are provided, in which the lipid bilayer is disrupted and the cytoplasmic protein is fixed, with generally good retention of antigenicity—the dehydration method (using methanol) and the cross-linking/ permeabilization method (using a mixture of paraformaldehyde and Triton X-100).

Materials

- Cells of choice generally grown in a 3- to 5-cm dish (a smaller dish uses less antibody; however, the dish must fit under the microscope objective)
- 4°C phosphate-buffered saline (PBS; APPENDIX 2)
- 2% paraformaldehyde (PFA) fixative (1:1 dilution with PBS of 4% PFA fixative described in UNIT 14.1), 4°C—for cell surface antigens
- 100% methanol, -10° to -20°C (solvent cooled in the ice box of a refrigerator is ideal) or 2% PFA fixative containing 0.1% Triton X-100, 4°C—for cytoplasmic antigens
- Primary antibody, ~5 to 10 µg/ml (see commentary)

Secondary antibody–fluorochrome conjugate specific to the source species of primary antibody (see commentary)

Motorized pump or water pump

1. Cool the cells on ice.

The cells should be confluent (or as close to confluent as possible).

2. Pipet off culture media and wash in 4°C PBS.

All pipetting steps are done with Pasteur pipets attached to a pump.

3. Pipet off PBS. If cell surface antigens are being studied, fix for 30 min in 2% PFA fixative on ice. If cytoplasmic antigens are being studied, fix either 30 min in 2% PFA fixative/0.1% Triton X-100 on ice, or 15 min in 100% methanol in the freezer compartment of a standard refrigerator (–10° to –20°C).

Both methods of cytoplasmic antigen fixation work well; however, the paraformaldehyde/Triton X-100 fix frequently gives a better final morphology.

If methanol fixation is being used, be sure that the cells are washed thoroughly with methanol before putting into freezer, or they will freeze.

4. Pipet off fixative and wash twice in 4°C PBS (5 min/wash).
5. Microcentrifuge diluted primary antibody 2 min at 13,500 × g, 4°C.
6. Layer primary antibody into dish such that cells are just covered and incubate 1 hr at 4°C.
7. Wash four times in 4°C PBS (5 min/wash).
8. Microcentrifuge diluted secondary antibody 2 min at 13,500 × g, 4°C.
9. Layer secondary antibody into dish and incubate 1 hr at 4°C.
10. Wash four times in 4°C PBS. Store cells in PBS. Unless cells are to be observed immediately, cover dishes, wrap in aluminum foil, and refrigerate. It is important to examine preparations of this type within 24 hr as the fluorescence rapidly fades and/or dissociates from the cells.

Do not store cells in water as osmotically competent membrane systems remaining in the cells will be disrupted.

IMMUNOFLUORESCENT LABELING OF SUSPENSION CELLS

Cells may be labeled in suspension, concentrated by centrifugation, and layered onto poly-L-lysine–coated slides prior to observation. The procedure is essentially the same as for monolayer cells, except that the washing steps are replaced by gentle centrifugation, replacement of supernatant, and resuspension of cells. Also, to preserve cellular morphology, the number of centrifugation steps is minimized. Use of PFA/Triton X-100 is the preferred fixation method for cytoplasmic labeling of cells grown in suspension, as methanol fixation frequently renders cells too fragile to be centrifuged reliably.

Additional Materials

5 × 10⁶–10⁷ cells in media in 15-ml Falcon tube

Poly-L-lysine–coated slides (UNIT 14.1)

1. Cool the cells on ice.
2. Centrifuge 5 min at 800 × g, 4°C, in tabletop centrifuge (for lymphocytes).

Speed is adjusted for cell type but must be slow enough to prevent damage. Unless otherwise stated, all centrifuge steps are done under these conditions.

**ALTERNATE
PROTOCOL**

**In situ
Hybridization
and Immuno-
histochemistry**

14.6.2

3. Pipet off culture media and resuspend cells in cold PBS.
All pipetting steps are performed with pipets attached to a pump.
4. Centrifuge to a pellet, pipet off PBS, and fix by resuspending cells (30 min on ice) in 1 to 2 ml of either 2% PFA fixative or 2% PFA fixative/0.1% Triton X-100.
5. Centrifuge to a pellet, pipet off fixative, and resuspend cells in 15 ml of 4°C PBS. Allow to wash 5 min and repeat with a second PBS wash.
A large volume of buffer allows a minimal number of washes.
6. Microcentrifuge primary antibody 2 min at $13,500 \times g$, 4°C. An antibody volume of 250 μ l is adequate for 5×10^6 cells.
7. Centrifuge cells to a pellet, pipet off PBS, resuspend in primary antibody, and incubate 1 hr at 4°C.
8. Dilute cells and antibody to 15 ml with 4°C PBS.
This dilution step serves as an initial wash.
9. Centrifuge to a pellet, pipet off PBS, and resuspend cells in 4°C PBS. Repeat once.
10. Microcentrifuge secondary antibody 2 min at $13,500 \times g$, 4°C. An antibody volume of 250 μ l is adequate for 5×10^6 cells.
11. Pipet off PBS, resuspend cells in secondary antibody, and incubate 1 hr at 4°C.
12. Repeat steps 8 and 9.
13. Unless cells are to be observed immediately, wrap tube in aluminum foil and refrigerate—*prior* to final concentration step below.
It is important to examine preparations of this type within 24 hr as the fluorescence rapidly fades and/or dissociates from the cells.
14. Centrifuge cells to a pellet and resuspend in a small volume of PBS. Pipet onto poly-L-lysine-coated slides, cover with cover slips, and observe immediately if possible.
Do not store in water as osmotically competent membrane systems remaining in the cells will be disrupted.

**BASIC
PROTOCOL**

IMMUNOFLUORESCENT LABELING OF TISSUE SECTIONS

This protocol describes a procedure for fluorescent labeling of frozen sectioned material. Material is prepared and sectioned as detailed in *UNIT 14.2*. Sections are reacted with primary antibody and secondary antibody–fluorochrome conjugate.

Materials

- Specimen tissue cryosections on glass slides (*UNIT 14.2*)
- Phosphate-buffered saline (PBS; *APPENDIX 2*)
- Primary antibody, ~5 to 10 μ g/ml (see commentary)
- Secondary antibody–fluorochrome conjugate specific to the source species of primary antibody (see commentary)
- Mounting medium (e.g., Gelvatol; *UNIT 14.5*)
- Plastic slide box or moist chamber (*UNIT 14.2*)
- Motorized pump or water pump

1. Layer wet paper towels in base of slide box to make a moist chamber, or make a chamber as in *UNIT 14.2*.

2. Remove slides with sections from cryostat or freezer, place across slide box (~6 per side), or place in moist chamber. The slides should not touch one another.
3. Once slides are at room temperature, and before they air dry, layer PBS over sections. Do not flood the slides.
4. Microcentrifuge diluted primary antibody 2 min at $13,500 \times g$, 4°C (40 to 50 μl antibody should cover the sections on each slide).
5. Remove PBS from slides by aspirating at one end of the sections with a Pasteur pipet connected to a pump and introduce antibody at the other end.
6. Close box and incubate 1 hr at room temperature.
7. Wash slides three times in PBS (5 min/wash).

For all washes, introduce new buffer at one end and aspirate off old buffer from opposite end of sections.
8. Microcentrifuge diluted secondary antibody 2 min at $13,500 \times g$, 4°C . Allow 40 to 50 μl antibody per slide.
9. Layer secondary antibody over sections and incubate in moist chamber 1 hr at room temperature (see commentary on choice of antibody).
10. Wash slides three times in PBS (5 min/wash).
11. Lay cover slips on paper towels and place a drop of Gelvatol in the middle of the cover slip.

If Gelvatol is not available, glycerol or PBS may be used for temporary mounting.
12. Invert slides on cover slip.

Do not apply pressure, as this will damage the section; instead let the Gelvatol spread naturally.
13. Leave slides 30 min on bench under aluminum foil to keep out light to allow Gelvatol to harden.
14. Observe under microscope or store at 4°C in closed slide box.

IMMUNOFLUORESCENT LABELING OF TISSUE SECTIONS USING COPLIN JARS

**ALTERNATE
PROTOCOL**

The preceding basic protocol for immunofluorescent staining of tissue sections should work for all antigens. However, the following modifications are appropriate for those antigens which do not lose their antigenicity upon air drying. It employs Coplin jars for the washing steps, and is simpler and quicker than the basic protocol.

Tissue sections are air dried and may be stored in slide boxes at room temperature prior to labeling. Air-dried sections rarely detach from glass slides and therefore may be treated more rigorously than wet sections.

1. Layer wet paper towels in base of slide box to make a moist chamber, or make a chamber as in *UNIT 14.2*.
2. Rehydrate air-dried slides by dipping in a Coplin jar containing enough PBS to cover the sections and dry the area around each section with paper towel.

Drying the area around the sections reduces the amount of antibody used as it stops the antibody from spreading all over slide.

**In situ
Hybridization
and Immuno-
histochemistry**

14.6.4

3. Place slides across slide box (~6 per side) or in moist chamber. The slides should not touch one another.
4. Microcentrifuge diluted primary antibody 2 min at $13,500 \times g$, 4°C . Pipet primary antibody over sections; 40 to 50 μl antibody should cover the sections on each slide. Close box and incubate 1 hr at room temperature.
5. Put slides in Coplin jar containing PBS and wash by changing PBS three times (5 min/wash).

PBS is changed by simply tipping it down the sink and refilling the Coplin jar with fresh buffer.
6. Microcentrifuge diluted secondary antibody 2 min at $13,500 \times g$, 4°C . Allow 40 to 50 μl antibody per slide.
7. Dry slides around sections, layer secondary antibody over sections, and incubate in moist chamber 1 hr at room temperature (see commentary on choice of antibody).
8. Wash slides three times in PBS as in step 5.
9. Follow basic protocol for sectioned material, starting at step 11.

**ALTERNATE
PROTOCOL**

IMMUNOFLUORESCENT LABELING USING STREPTAVIDIN-BIOTIN CONJUGATES

This protocol describes a triple-step reaction technique that increases the sensitivity of the immunohistochemical reaction. This is achieved using a streptavidin–secondary antibody conjugate followed by reaction with a biotin–fluorochrome conjugate (see Fig. 14.6.1d and commentary on choice of secondary antibody).

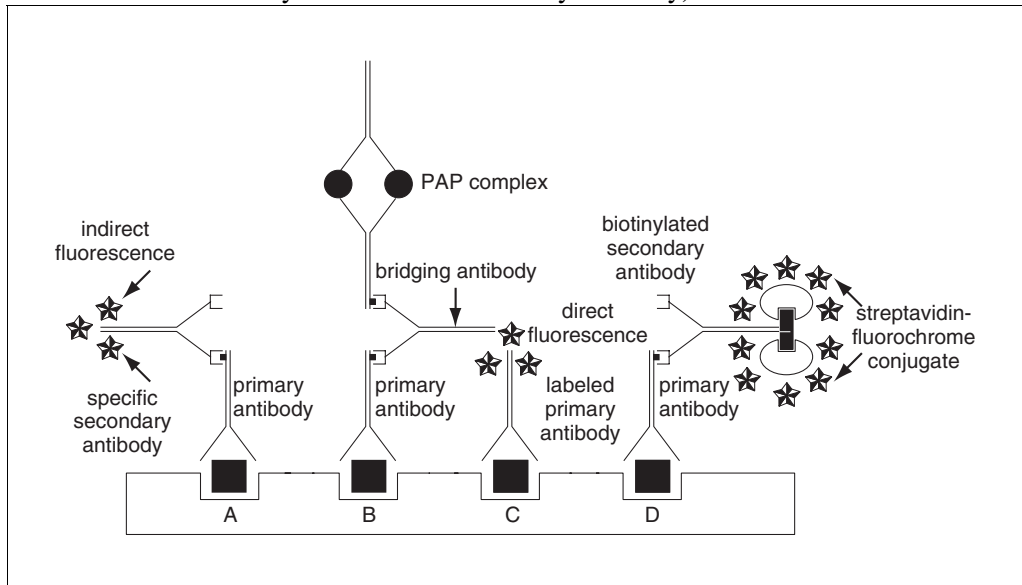


Figure 14.6.1 Various methods of immunohistochemical labeling. **(A)** The primary antibody bound to a specific epitope has been revealed by a secondary antibody conjugated to a fluorochrome. This may then be viewed by fluorescence microscopy. **(B)** The PAP–immunoperoxidase method. The primary antibody is bound to its epitope and has been coupled to PAP complex by a bridging antibody. Following development of the DAB substrate solution, a brown coloration at the epitopic site will be seen by bright-field microscopy. **(C)** The original method of labeling tissue or cells for immunofluorescence employed a primary antibody directly conjugated to a fluorochrome that could then be viewed by fluorescence microscopy. **(D)** The use of streptavidin–biotin conjugates allows considerable amplification of the section so that low-level binding may be clearly seen. The primary antibody has bound specifically to the substrate which is then labeled with biotinylated secondary antibody. This complex is then revealed with a streptavidin fluorochrome.

Additional Materials

Biotinylated secondary antibody (Vector Laboratories)
Fluorochrome-streptavidin conjugate (Vector Laboratories)

Follow steps 1 to 7 of basic protocol on immunofluorescent labeling of tissue sections.

8. Microcentrifuge diluted biotinylated secondary antibody for 2 min at $13,500 \times g$, 4°C (~40 to 50 μl antibody per slide).
9. Lay secondary antibody over sections and incubate in a moist chamber 1 hr at room temperature.
10. Wash slides three times in PBS (15 min/wash). Layer fluorochrome-streptavidin (~40 to 50 μl antibody per slide) over sections and incubate in a moist chamber for 1 hr at room temperature. Wash slides three times in PBS as in step 7 of basic protocol.
11. Follow basic protocol for sectioned material, starting at step 11.

IMMUNOGOLD LABELING OF TISSUE SECTIONS

This parallels the basic protocol on immunofluorescent labeling of tissue sections, except for steps 9 and 14 below. See commentary on silver enhancement method in nonfluorescent marker section.

Additional Materials

Immunogold conjugate (Janssen)

9. Incubate 2 hr in immunogold-conjugated secondary antibody.
14. Observe under microscope or store at room temperature in a closed slide box.

IMMUNOPEROXIDASE LABELING OF TISSUE SECTIONS

This is a very sensitive method. Currently, the most popular immunoperoxidase method is that using horseradish peroxidase–anti-peroxidase (PAP) (Sternberger et al., 1970). Here, a tertiary antibody is used that is directed at, and comes conjugated to, horseradish peroxidase (HRPO). This form of noncovalent linkage assures that the enzyme retains maximum reactivity for the diaminobenzidine (DAB) substrate. The PAP complex is cross-linked to the primary antibody via a “bridging antibody” (Fig. 14.6.1b). The antibody constituting the PAP complex must be of the same type as the primary antibody to be recognized by the bridging antibody.

Additional Materials

0.25% hydrogen peroxide in PBS
Secondary antibody (a bridging antibody recognizing both primary antibody and PAP complex)
Horseradish peroxidase–anti-peroxidase (PAP) complex (from the same source species as the primary antibody)
Diaminobenzidine (DAB) substrate solution

CAUTION: DAB is a carcinogen; handle with great care.

Follow steps 1 to 3 of the basic protocol for immunofluorescent labeling of tissue sections. All pipetting steps are performed with Pasteur pipets attached to a pump.

4. Incubate slides 30 min in 0.25% hydrogen peroxide in PBS at room temperature. Wash three times in PBS.

***ALTERNATE
PROTOCOL***

***ALTERNATE
PROTOCOL***

**In situ
Hybridization
and Immuno-
histochemistry**

14.6.6

5. Follow steps 4 to 7 of basic protocol (immunofluorescent labeling of tissue sections).
6. Incubate slides 1 hr at room temperature in specific secondary bridging antibody. Wash three times in PBS.
7. Incubate slides 1 hr at room temperature in PAP complex.
8. Wash three times in PBS.
9. Develop 2 to 5 min in DAB substrate solution at room temperature.

Developing time is empirical and must be assessed for each experiment.

10. Follow steps 10 to 14 of basic protocol.

Keep slides in the dark in a slide box. It is not necessary to refrigerate them.

**ALTERNATE
PROTOCOL**

**IMMUNOFLUORESCENT DOUBLE-LABELING
OF TISSUE SECTIONS**

Indirect immunofluorescence microscopy allows two or more antigens to be revealed on the same section at any one time. This is done by using fluorochromes which are excited by, and emit light of, different wavelengths. Generally this is limited to two fluorochromes, the most popular combination being rhodamine (excited in green range and emitting red) and fluorescein (excited in blue and emitting green), though more exotic combinations using fluorochromes such as phycoerythrin are possible (see commentary on choice of antibody). However, the following should be observed, regardless of the fluorochrome used:

1. When performing double-labeling experiments, the most important criterion is that the primary antibodies be of different types, such that secondary antibodies can recognize them independently. It is preferable that they originate in different animals; however, if monoclonals are being used, this may not be possible. When monoclonals are used, antibodies of different types such as IgG and IgM may be differentiated reliably, though antibodies of different classes (IgG₁ and IgG₂, for example) cannot generally be differentiated by secondary antibodies.
2. The protocol is the same as the basic protocol for immunofluorescent labeling of tissue sections except that the primary and secondary antibody steps are replaced by *mixtures* of two primary and two secondary antibodies.
3. Prior to attempting a double-labeling experiment, single-labeling experiments should be performed to find optimal dilutions for the different primary and secondary antibodies.

REAGENTS AND SOLUTIONS

Diaminobenzidine (DAB) substrate solution

Dissolve 60 mg 3,3' diaminobenzidine (DAB) in 200 ml phosphate-buffered saline (PBS; APPENDIX 2). Store frozen in 5- to 10-ml aliquots until use. Defrost DAB aliquot and add hydrogen peroxide to a final concentration of 0.1%. (NOTE: concentrated hydrogen peroxide is generally sold as a 30% solution.) Filter through syringe (Millipore) and use immediately.

CAUTION: *DAB is a carcinogen and must be used with great care. Exposure to dry carcinogenic DAB may be minimized by preparing large batches of DAB solution in the fume hood which are stored as frozen aliquots.*

COMMENTARY

Background Information

Nearly 50 years ago, fluorescent markers were employed in the first successful use of immunocytochemical methods to study cellular protein distributions (Coons, 1941). Early studies used direct labeling (Fig. 14.6.1c), where the primary antibody is conjugated directly to the fluorochrome. This methodology was cumbersome as every primary antibody had to be conjugated separately to the fluorescent marker. Indirect methods are more sensitive, as signal amplification occurs when multiple secondary antibodies bind to a single primary antibody (particularly when streptavidin-biotin conjugates are used; see commentary below on choice of antibody). Furthermore, they are much easier to employ, because a single secondary antibody conjugate may be used to label a variety of different primary antibodies from the same source species. Today, commercially available secondary markers have very high specificity, purity, and signal strength, and are available directed against primary antibodies from most animals. In addition, an ever-increasing number of primary antibodies is available (a fairly complete list is compiled annually in *Linscott's Directory*).

A variety of methodologies for immunohistochemical labeling of cells and sections are described in this section. Each method has advantages and disadvantages. For routine labeling of sections, we suggest using the basic protocol for immunofluorescent labeling initially. If this is too laborious, try the alternate protocol, which describes immunofluorescent labeling of sections using Coplin jars.

If the level of labeling seen on the section is inadequate using the basic protocol, it is possible to amplify the perceived signal using the alternate protocol employing streptavidin-biotin conjugates, or to use a bright-field technique such as immunoperoxidase labeling of tissue sections. Of the two bright-field methods discussed—immunogold labeling and immunoperoxidase labeling—the latter is generally considered to be more sensitive than immunofluorescence methods. However, it has some specific disadvantages. These are (1) the reaction product may diffuse away from the site of reaction and hence the antibody locus, (2) it is a lengthy triple-step procedure, (3) the specificity of the reaction may be confused by reactive endogenous peroxidase, (4) the reaction product is light sensitive, and (5) DAB is con-

sidered a carcinogen. These problems may be solved to a degree in that endogenous peroxidase may be blocked by preincubation with hydrogen peroxide and labeled slides may be kept in light-tight boxes. Other modifications are to use alternative substrate reagents (see Polak and Van Noorden, 1983) or to use alkaline phosphatase instead of peroxidase. These latter methods, however, generally give inferior results.

The immunogold labeling methodology, still gaining popularity, is best used in conjunction with silver enhancement of the gold probe (as discussed below).

It is strongly recommended that the critical parameters sections be read carefully prior to attempting any of the methods discussed here.

Critical Parameters

Dilution recommendations in the protocols—both in terms of true protein concentrations and empirical dilutions—are rough guidelines and are presented to serve only as starting points. Whenever a study uses an unknown primary antibody, a dilution series should be performed to assay the optimal dilution for that antibody. Frequently, clear differences in labeling may be seen with quite subtle changes in dilution. For example, an initial dilution series would be 1:10, 1:100, and 1:1000. Once the appropriate range is found, another series should be done, e.g., 1:100, 1:200, 1:400. While this may seem excessively time consuming, it will ensure optimal results and frequently a more economic use of antibody.

Choice of antibody

Primary antibodies. Antibodies are produced in two main forms, polyclonal antibodies and monoclonal antibodies (see Chapter 11). Their critical feature with regard to immunohistochemistry is that polyclonal antisera may contain antibodies to many different epitopes on the antigen, whereas monoclonal antibodies recognize a single epitope. The ramifications of this difference are considerable, as a multivalent polyclonal antibody will recognize many epitopic sites, the resultant signal seen using immunocytochemistry may be greater than that found using a monoclonal antibody. On the other hand, monoclonal antibodies are generally more specific than a polyclonal antiserum. Nonspecific cross-reactivity is generally avoided in polyclonal antibodies by affinity

purification of the antibody before use. Generally, best results are obtained with affinity-purified polyclonal antibodies followed by monoclonal antibodies, and finally antisera.

Monoclonal antibodies (UNITS 11.3-11.11). Monoclonal antibodies are available in three main forms: hybridoma culture supernatant, ascites and purified IgG. The effective concentration of the antibody in monoclonal culture supernatants is generally very low. This antibody should be used either concentrated or at dilutions in the 1:5 to 1:10 range.

Ascites fluid contains a large amount of cellular debris as well as antibody, although as long as it is spun before use (2 min, 13,500 × *g* in microcentrifuge) it will work very well. Both ascites and purified monoclonal immunoglobulins are generally used very dilute (1:1000 to 1:10,000).

It should be noted that monoclonal antibodies are commonly produced as IgM's. As this form of antibody effectively appears as clusters of immunoglobulin, the labeling may appear punctate due to clustering of the secondary antibody.

Polyclonal antibodies (UNITS 11.12 & 11.13) are present in serum and may be used in this form, or after purification from serum or affinity purification. If the antibodies are specific on western blots as antiserum, they will generally work adequately for immunolabeling at the light-microscope level and further purification is unnecessary. As a general rule, polyclonal antisera work well at 1:50 to 1:100 dilution. After affinity purification when the exact protein concentration is known, the working dilution should be 5 to 10 µg/ml.

Secondary markers

A variety of different fluorochromes are available commercially (see below) which are used in conjunction with one of three main second labeling methods as follows.

Secondary antibodies. This is the most common secondary marker and is the recommended choice in this unit. The source species should not compete with other primary antibodies in a double-labeling experiment. They are generally sold in lyophilized form. Following rehydration, the antibody should be spun 5 min at 4°C in a microcentrifuge, and the supernatant quick frozen and stored in 10-µl aliquots at -70°C until use. Antibodies sold as a liquid should also be centrifuged when received, the supernatant quick frozen, and stored at -70°C

in aliquots unless specifically indicated by the manufacturer.

Protein A, protein G. Protein A is generally purified from the cell wall of *Staphylococcus aureus* and binds very strongly, though variably, to the Fc fragment of antibodies from a variety of source animals—in particular, mouse, human, and rabbit antibodies. Protein G is derived from the cell wall of G group streptococci and also binds to the Fc fragment of many antibodies, frequently with greater avidity than protein A. For example, it binds to rat and goat IgG more strongly than protein A. The disadvantage of this method of labeling is that it may bind to other primary or secondary markers (in the case of a double-labeling experiment), or to a blocking antibody, if used, thereby causing increased background or apparent colocalization with another antibody.

Streptavidin-biotin. This method is used primarily as an amplification system (Fig. 14.6.1d). Biotin binds strongly and irreversibly to avidin. Biotinylated secondary antibodies are incubated with a streptavidin-fluorochrome conjugate. Several streptavidin fluorochrome conjugates will bind to each biotinylated antibody.

Fluorochrome

Choice of fluorochrome for most work is simple, although for double- or triple-labeling experiments the combinations may become more exotic. The following is a discussion of each with its potential advantages and disadvantages. See also Table 14.6.1 for troubleshooting the labeling procedure.

Fluorescein isothiocyanate (FITC) is the most commonly used fluorochrome and most fluorescent microscopes will be equipped with the correct filters. It is excited by blue light and emits a bright green fluorescence (ideal filters are Excitor 450 to 490 nm and Barrier 520 to 560 nm). The intensity of label is high and is frequently detected where other fluorochrome such as rhodamine may give a threshold signal. The major disadvantage is that it extinguishes rapidly and irreversibly, so photography must be performed as soon as a positive area is detected. This dye is available conjugated to all the markers described above from a variety of vendors (Cappel, Calbiochem, ICN).

Tetramethyl rhodamine isothiocyanate (TRITC) is a very commonly used fluorochrome, second in popularity to FITC, and again, most microscopes will be equipped

with correct filters. It is excited by green light and emits a red fluorescence (ideal filters are Excitor 510 to 560 nm and Barrier 590 nm lower cutoff). The intensity of label is less than with FITC though the persistence of the signal is longer, thus more careful examination of the label is possible before extinction below usable level occurs. This dye is available conjugated to most of the markers described above from a variety of vendors (Cappel, Calbiochem, ICN), and is commonly used in double-labeling experiments in conjunction with FITC-conjugated markers.

Texas Red is becoming increasingly popular inasmuch as it offers the high emission of FITC in conjunction with the better persistence of TRITC. It is excited and emits in the same range as TRITC, though the emitted signal has a slightly more orange characteristic. Unfortunately, it cannot be reliably distinguished from TRITC and thus cannot be used with this in double labeling work, though it is ideal for use with FITC. We prefer to use this dye as a matter of course, and its increasing popularity means that the range of available immunoconjugates will continue to increase.

Phycoerythrin has been available for some time but is only now becoming popular as a

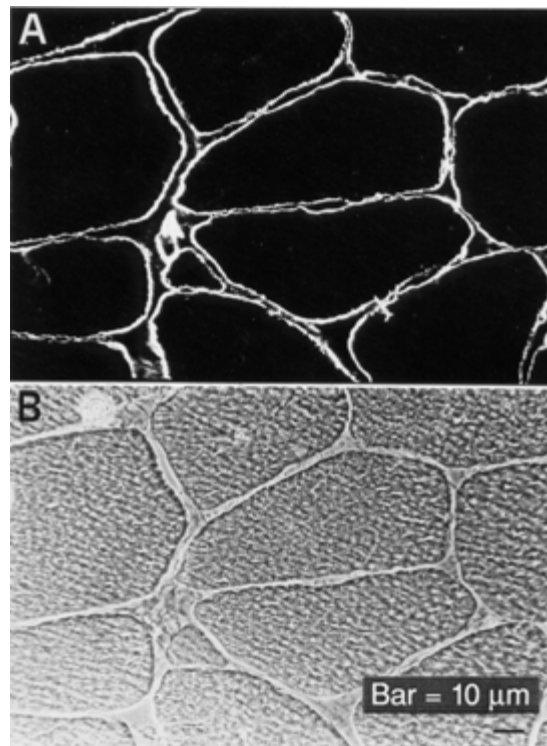
potential marker. It is excited in the same range as FITC though it emits a more yellow light, which makes it easily distinguishable from FITC. Furthermore, it generally has a longer persistence than FITC. It is discussed here as it is the most plausible third label for triple label studies. The only disadvantage is that the photography must be in color so that FITC and phycoerythrin may be distinguished (marketed by TAGO).

Nonfluorescent markers

The advantages of these methods are that labeled sections have an indefinite shelf life, and morphology and labeling may be examined using the same type of illumination. Further, the immunoperoxidase method is highly sensitive and may be used on paraffin sections if antigenicity remains where fluorescent markers cannot be used since autofluorescence is generally too great in this type of sample.

Immunogold conjugates present as a pink coloration under the light microscope. They have a high mass, and hence, high steric hindrance when compared with fluorescent conjugates, which may decrease labeling intensity. It is possible to increase the optical density of label using silver enhancement, which enlarges

Figure 14.6.2 Immunofluorescent labeling of the protein dystrophin in a cross-section of normal skeletal muscle. The protein is clearly localized at the periphery of the muscle fibers (**A**). No background is apparent within the muscle fibers themselves. The identity of the cells as muscle fibers is confirmed in the phase micrograph of the same field (**B**).
Bar = 10 μ m.



the gold particle by depositing metallic silver on it (Danscher, 1981). This method is difficult inasmuch as the reagents used are generally light sensitive and the reaction is hard to control. Thus, it is strongly recommended that a kit (such as the IntenSE kit marketed by Janssen) be used.

Controls

Because of the variety of problems of background and nonspecific labeling inherent in any immunohistochemical analysis of cells and tissues, the following controls are suggested (see also Table 14.6.1): (1) use of preimmune serum from the antibody-producing animal used at an equivalent dilution, (2) use of secondary antibody alone, (3) use of no antibodies at all, (4) use of an antibody, the antigen for which is not found in the tissue or cells to be studied, (5) use of a known positive antibody (e.g., actin), and

(6) use of a limited dilution series for the primary antibody for each experiment.

Troubleshooting

Refer to Table 14.6.1 for an itemization of the problems commonly encountered when using the procedures in this unit. Possible causes are described and solutions are recommended.

Anticipated Results

Good immunolabeling at the light microscope is characterized by clear definition of a marker within the tissue or cell being studied, with minimal background or nonspecific labeling throughout the remainder of the tissue (see Fig. 14.6.2).

Time Considerations

One advantage of the immunohistochemi-

continued on p. 14.6.13

Table 14.6.1 Troubleshooting Guide for Immunofluorescent Labeling of Specimen Sections

Problem	Possible cause	Solution
Slides appear "faded"	Slides were not kept in the dark and the cold prior to observation	It is extremely important that slides are stored in the cold in the dark
Morphology appears grainy; labeling is grainy or weak	Slide dried out when removed from the cryostat or lyophilized in the cryostat or -80°C freezer	Deal with slides promptly; <i>do not allow slides to dry out at any time</i> , unless the antigen is extremely robust and abundant
Fluorescence is seen even in unlabeled material	Autofluorescence, frequently caused by the use of aldehyde fixatives	Block most autofluorescence by preincubating sections after the first PBS wash for 10 min in 0.5% sodium borohydride (NaBH_4) in water (prepare immediately before use). Frequently, bubbles appear on the section. This is normal, but treatment is rigorous and may cause section loss.
	With monolayer labeling, autofluorescence may be caused by the culture dish; test by examining a new culture dish under the fluorescence microscope	Grow cells on glass cover slips or in sterile culture chambers in glass slides
Fluorescent labeling appears over section—even in the absence of a primary antibody	Secondary antibody is too concentrated	Dilute antibody further. Normally, commercial immunoconjugates work at a 1:100 dilution, although this is a very general starting concentration. For best results, do a dilution series with each batch of secondary antibody.
Bright spots of fluorescence occur nonspecifically across section and slide	Inadequate washing during immunolabeling or a failure to spin unconjugated fluorochrome out of the secondary antibody	Make sure secondary antibodies are spun 5 min in microcentrifuge ($13,500 \times g$) as soon as they are received. This is particularly important for antibodies shipped as lyophilized product. In most cases, a 10- to 20- μl pellet will result, which should be discarded.

continued next page

Table 14.6.1 Troubleshooting Guide for Immunofluorescent Labeling of Specimen Sections (continued)

Problem	Possible cause	Solution
Labeling appears specific, but background is excessive	Nonspecific binding of primary or secondary antibody	Preincubate sections 1 hr in 1% serum (eg., normal goat serum) prior to labeling with primary antibody. Be sure that the blocking antibody is not recognized by the secondary antibody (as this will only compound labeling problems) by blocking with serum from same animal species as that used to make the secondary antibody. It is also possible to dilute all antibodies in 1% to 10% serum, thereby avoiding the blocking step.
Labeling is specific, but not where expected, or mutually exclusive labeling patterns are seen on the same section	Antiserum contains tissue-specific antibodies other than those to the target antigen	Confirm problem by labeling sections with preimmune serum, i.e., serum taken from the antibody-producing animal prior to injection of antigen. Serum should always be taken prior to immunization and should be part of the antibody production process. Avoid nonspecific labeling of this type by affinity purification of the antibody.
No labeling is seen on sections even though clear label is seen on western blots of the same material	The tissue section used does not contain the antigen	Be sure to cut sufficient sections so that all possible structural features of the tissue are sampled
	The antigen is not recognized by the antibody in its native form. This may occur if the antibody was raised against denatured rather than native antigen and happens more frequently with monoclonal than polyclonal antibodies, because the latter include antibodies to a variety of epitopes.	Make new antibody using native antigen
	Tissue processing prior to labeling has rendered the antigen nonreactive	While many antigens may be air dried in sections and still retain strong antigenic reactions (see alternate protocol for immunofluorescent staining using Coplin jars), if labeling is not seen using this protocol, use basic protocol.
Labeling appears blurry and streaked across slide	Overfixation leading to conformational changes in the antigen prior to freezing (<i>UNIT 14.2</i>)	Reduce the fixation time of the sample
	On rare occasions, very soluble antigens may be found which dissolve out of the section during labeling. A more probable cause is smearing of the tissue during sectioning (see sectioning troubleshooting, Table 14.2.1).	Fixation of tissue prior to freezing or fixation of section itself

cal methods described here is their relative speed. It is quite possible to go from the whole animal to labeled sections in one day.

Literature Cited

- Coons, A.H., Creech, H.J., and Jones, R.N. 1941. Immunological properties of an antibody containing a fluorescent group. *Proc. Soc. Exp. Biol. Med.* 47:200-202.
- Danscher, G. 1981. Localization of gold in biological tissue: A photochemical method for light and electron microscopy. *Histochemistry* 71:81-93.
- Polak, J.M. and Van Noorden, S., eds. 1983. Immunocytochemistry: Practical Applications in Pathology and Biology. John Wright PSG, Littleton, Mass.
- Sternberger, L.A., Hardy, P.H. Jr., Cuculis, J.J., and Meyer, H.G. 1970. The unlabeled antibody-enzyme method of immunohistochemistry. Preparation of soluble antigen-antibody complex (horseradish peroxidase-anti-horseradish peroxidase) and its use in the identification of spirochetes. *J. Histochem. Cytochem.* 18:315-333.

Key References

- Linscott's Directory of Immunological and Biological Reagents. Mill Valley, Calif.
- Highly recommended publication listing sources of immunological reagents, kits, and cells/organisms, including addresses and phone numbers of commercial suppliers (updated quarterly).*
- Polak and Van Noorden, 1983. See above.
- Sternberger, L.A. 1986. Immunocytochemistry (3rd ed.). Wiley, N.Y.
- Both these texts discuss the various methods of immunohistochemistry in extensive detail.*

Contributed by Simon Watkins
Harvard Medical School
and Dana-Farber Cancer Institute
Boston, Massachusetts

In situ Hybridization and Detection Using Nonisotopic Probes

Nonisotopic in situ hybridization can be used to determine the cellular location and the relative levels of expression of specific transcripts within cells and tissues. RNA in prepared specimens is hybridized with a probe labeled nonisotopically with biotin or digoxigenin. Nonisotopic probes are generally detected by fluorescence or enzymatic methods (Figs. 14.7.1 & 14.7.2). Hybridization and detection of ^{35}S - and ^{32}P -labeled probes are described in UNITS 14.3 & 14.4, respectively.

Fluorescence in situ hybridization (FISH), probably the most widely used method, is described in Basic Protocol 1. Support Protocols 1 and 2 describe techniques for amplification of weak fluorescent signals obtained in FISH. Nonisotopic probes can also be detected by enzymatic reactions using horseradish peroxidase or alkaline phosphatase, as described in Alternate Protocols 1 and 2. Nonisotopic labeling of DNA probes by nick translation or by random oligonucleotide primed synthesis is described in UNIT 3.18.

NOTE: Water should be treated with DEPC and all solutions prepared with DEPC-treated water to inhibit RNase activity. See UNIT 4.1, Reagents and Solutions, for instructions.

CAUTION: DEPC is a suspected carcinogen and should be handled carefully.

FLUORESCENCE IN SITU HYBRIDIZATION

Fluorescence in situ hybridization (FISH) involves hybridization of biotin- or digoxigenin-labeled DNA probes to prepared specimens (cells or tissue sections) and visualization of those probes with fluorochrome-conjugated reagents. The accompanying Support Protocols describe amplification of weak fluorescent signals.

Materials

- Glass slide containing specimen (UNIT 14.1)
- 20 to 150 ng nonisotopically labeled DNA probe (UNIT 3.18)
- Deionized formamide (American Bioanalytical *OR* UNIT 14.3)
- 10 mg/ml sonicated salmon sperm DNA (see recipe)
- Master hybridization mix (see recipe)
- 50% (v/v) formamide (not deionized)/2× SSC (APPENDIX 2)

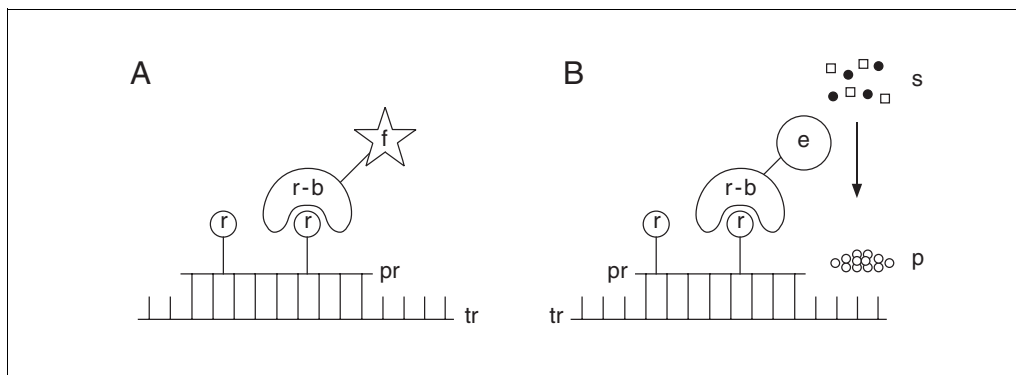


Figure 14.7.1 Detection of hybridized probe. **(A)** Fluorescence in situ hybridization (FISH). **(B)** Enzymatic detection. Abbreviations: e, enzyme (e.g., alkaline phosphatase or horseradish peroxidase); f, fluorochrome (e.g., fluorescein, rhodamine, or Texas red); p, colored precipitate product; pr, probe labeled with reporter molecule; r, reporter molecule (e.g., biotin, digoxigenin); r-b, reporter-binding molecule (e.g., avidin, streptavidin, or digoxigenin antibody); s, soluble substrate; tr, transcript. Arrow indicates reaction catalyzed by enzyme.

BASIC PROTOCOL

Contributed by Joan H.M. Knoll and Peter Lichter
Current Protocols in Molecular Biology (1995) 14.7.1-14.7.14
Copyright © 2000 by John Wiley & Sons, Inc.

In situ Hybridation and Immuno- histochemistry

14.7.1

1×, 2×, and 4× SSC, pH 7.0 (*APPENDIX 2*)
 Biotin detection solution *or* digoxigenin detection solution *or* biotin/digoxigenin detection solution (see recipes)
 0.1% (v/v) Triton X-100/4× SSC
 DAPI *or* propidium iodide staining solution (see recipes)
 Appropriate antifade mounting medium (see recipe)
 39°, 42°, and 72°C water baths
 Phase-contrast microscope
 22-mm² coverslips
 Rubber cement
 Moist chamber (Fig. 14.2.4)
 Slide box with desiccant (Baxter Scientific)
 Nail polish
 Fluorescence microscope with epillumination and filter set(s) appropriate for fluorochrome(s) used including dual-band pass filter (fluorescein/Texas red) or triple-band pass filter (fluorescein/Texas red/DAPI; Omega Optical or Chroma Technology)
 Ektar-1000 *or* Ektachrome-400 color film *or* Technical Pan 2415 black-and-white film (Kodak)

Additional reagents and equipment for preparing paraffin sections or cells for hybridization (*UNIT 14.3*, Basic Protocol, steps 1 to 16) or for preparing cryosections for hybridization (*UNIT 14.3*, Alternate Protocol, steps 1 to 7) and ethanol precipitation of DNA (*UNIT 2.1*)

CAUTION: Formamide, DAPI, and propidium iodide are hazardous; see manufacturer's information for guidelines on handling, storage, and disposal.

Prepare the specimens for hybridization

1a. *For hybridization to paraffin sections or cells:* Prepare slides by dewaxing, rehydrating, blocking, and dehydrating as in *UNIT 14.3*, Basic Protocol steps 1 to 7, finishing with air drying the specimens (or drying in dessicator), followed by storage in a slide box with dessicant overnight at −70°C.

It is essential that the specimens be absolutely dry before they are stored.

1b. *For hybridization to cryosections:* Pretreat sections and acetylate as described in *UNIT 14.3*, Alternate Protocol steps 1 to 7, finishing with dehydration of slides.

It is essential that the specimens be absolutely dry. Slides are now ready for hybridization (step 3), which should be done immediately.

Prepare the probe

2. For each hybridization precipitate 10 to 15 ng of probe in ethanol and dissolve in 10 μl deionized formamide. Add 5 μg (0.5 μl) sonicated salmon sperm DNA. Heat denature the probe 10 min at 70 to 80°C.

*Probe can be labeled with biotin or digoxigenin (*UNIT 3.18*) or with fluorochrome-tagged deoxynucleotides (see Background Information). Quantitation and detection of nonisotopic probes is discussed in the Support Protocol of *UNIT 3.18*.*

Hybridize probe to sample

3. Add 10 μl master hybridization mix to denatured probe (0.1 to 0.5 μg/ml final probe concentration). Mix well, spin briefly at high speed in a microcentrifuge, and transfer onto preparation on slide. Cover with 22-mm² coverslip, remove any large air bubbles with gentle pressure, and incubate 2 to 4 hr in a moist chamber at 37°C.

Hybridizations can also be performed overnight. In this case, seal the coverslip using rubber cement.

When hybridizing with labeled oligonucleotide probes, the optimal hybridization temperature is a function of the melting point of the probe. This must be calculated for each probe (see UNIT 6.4, Basic Protocol 1).

4. During last 30 min of hybridization, warm 50 ml of 50% formamide/2× SSC wash solution and 50 ml of 2× SSC in Coplin jars in a 37°C water bath.
5. Remove slide from moist chamber. Peel off rubber cement (if present) and carefully remove coverslip. Wash hybridized slide 15 min in 37°C 50% formamide/2× SSC, 15 min in 37°C 2× SSC, and 15 min in room temperature 1× SSC.
6. Allow slide to equilibrate 5 min in 4× SSC at room temperature. Remove slide and drain excess buffer. Do not allow slide to dry at any point during the procedure.
7. Add 50 µl biotin detection solution, digoxigenin detection solution, or biotin/digoxigenin detection solution to hybridized preparation on slide and cover with 22-mm² square of Parafilm. Incubate 45 min in an aluminum foil-wrapped moist chamber at 37°C.

Detection reagents and DAPI (step 10) are light sensitive; keep exposure to light to a minimum by wrapping containers in aluminum foil.

8. Soak slide sequentially in aluminum foil-wrapped Coplin jars containing room temperature 4× SSC, 0.1% Triton X-100/4× SSC, and 4× SSC—10 min in each solution.

Counterstain DNA

9. Add 50 µl DAPI or propidium iodide staining solution to preparation on slide, cover with 22-mm² square of Parafilm, and allow to stain 5 min at room temperature. Rinse briefly in 1× SSC in Coplin jar to remove excess stain. Blot slide, but do not dry.

DAPI staining of nuclei can be viewed simultaneously with fluorescein and rhodamine or Texas red. Propidium iodide and fluorescein can be viewed simultaneously, but propidium iodide and rhodamine or Texas red cannot because their emission spectra overlap.

10. Add 7 µl of appropriate antifade mounting medium to stained slide and add a coverslip. Gently squeeze out excess antifade medium, taking care not to damage tissue, and seal with nail polish. Store at -20°C in slide box with desiccant.

The antifade mounting medium chosen depends on the fluorochrome used (see Critical Parameters).

Slide is ready to examine immediately after it is mounted, but it may be stored several weeks at 4° or -20°C with minimal loss of fluorescence.

View and photograph slides

11. Examine slide using a fluorescence microscope with epillumination and filter set appropriate for the fluorochrome used.

A DAPI filter set is used for nuclear identification and a dual-band-pass or a triple-band-pass filter set is used for simultaneous viewing of multiple fluorochromes.

12. Photograph using either Ektar-1000 (for prints) or Ektachrome-400 (for slides) color film.

Exposure times vary with brightness of hybridization signal, but exposures for DAPI are ~2 sec and exposures through dual- or triple-band-pass filter sets are 30 to 90 sec and 3 to 8 sec, respectively. Probes with bright hybridization signals can be photographed with black-and-white film, e.g., Kodak Technical Pan 2415, exposed at ASA 200.

AMPLIFICATION OF HYBRIDIZATION SIGNALS

Signals of both biotin- and digoxigenin-labeled probes can be amplified if necessary (see Fig. 14.7.2). Amplification can be performed at any point after post-hybridization washes, including after viewing, and it may be repeated.

SUPPORT PROTOCOL 1

Amplification of Biotinylated Signals

Slides hybridized with biotinylated probes and exposed to avidin are incubated with a biotin-labeled antibody to avidin. Every avidin molecule contains four biotin-binding sites, so biotin-labeled antibody to avidin can potentially bind the avidin molecule via a biotin-avidin interaction or an antibody interaction. A final application of fluorescein-conjugated avidin results in significant signal amplification. Because the signal-to-noise ratio decreases with successive amplifications, amplification is recommended only when the probe yields weak fluorescence with very little background.

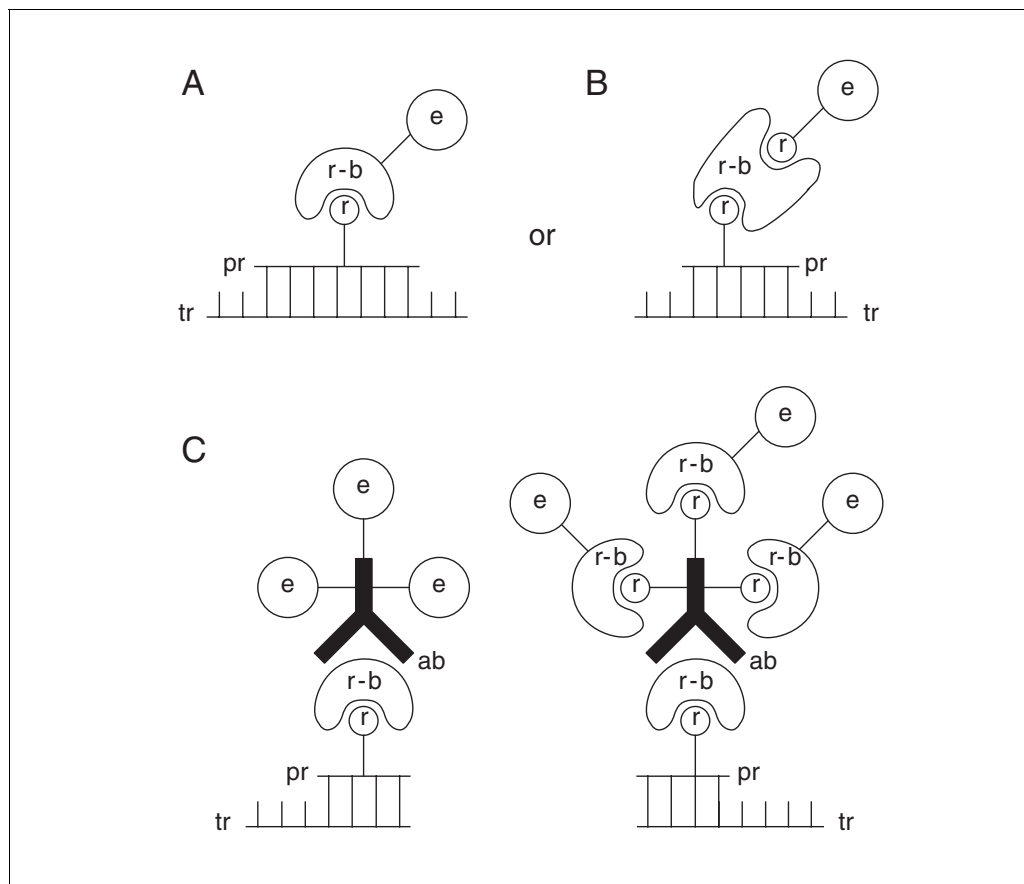


Figure 14.7.2 Enzyme-mediated detection of reporter molecules. **(A)** For direct detection, enzyme is conjugated to reporter-binding molecule. **(B)** For a two-step procedure, reporter-binding molecule is applied first, followed by an incubation with reporter-conjugated enzyme. **(C)** For signal amplification, incubation with reporter-binding molecule is followed by incubation with an antibody to reporter-binding molecule. The antibody may be conjugated with enzyme or reporter molecule. In the latter case an incubation with enzyme-conjugated reporter-binding molecule follows. The final step is addition of substrate (not shown). Fluorochrome can be substituted for enzyme molecule. Abbreviations: ab, antibody to reporter-binding molecule conjugated with enzyme or reporter molecule; e, enzyme (e.g., alkaline phosphatase or horseradish peroxidase); pr, probe labeled with reporter molecule; r, reporter molecule (e.g., biotin or digoxigenin); r-b, reporter-binding molecule (e.g., avidin, streptavidin, or anti-digoxigenin); tr, transcript.

Additional Materials (also see Basic Protocol)

- 1 to 3 µg/ml biotinylated anti-avidin antibodies (Vector Laboratories) in 4× SSC/1% (w/v) BSA (fraction V)
Biotin amplification solution: 2 to 5 µg/ml fluorescein-avidin DCS (Vector Laboratories) in 4× SSC/1% (w/v) BSA (fraction V)

NOTE: This procedure should be performed with minimum exposure to light.

1. Hybridize slide with biotinylated probe, wash, and perform first round of signal detection (see Basic Protocol).

Amplification of the signal from biotinylated probes can be carried out at any time after step 8 of the Basic Protocol.

2. If slide has been mounted and sealed, remove sealed coverslip by breaking the nail polish seal with a needle or scalpel, lifting off the nail polish, and soaking slide 15 min in 0.1% Triton X-100/4× SSC in an aluminum foil-wrapped Coplin jar with gentle agitation. If coverslip is not loose, carefully lift it off. Repeat wash twice.
3. Drain excess wash solution from slide and add 50 µl of 1 to 3 µg/ml biotinylated anti-avidin antibody to slide. Cover with 22-mm² square of Parafilm, place in moist chamber, wrap moist chamber in aluminum foil, and incubate 30 min at 37°C.
4. Remove Parafilm and wash 15 min in 0.1% Triton X-100/4× SSC with gentle agitation.
5. Drain excess solution from slide and add 50 µl biotin amplification solution. Cover with Parafilm, then incubate 30 min at 37°C in a foil-wrapped moist chamber.
6. Remove Parafilm and wash 15 min in 0.1% Triton X-100/4× SSC, then 15 min in 4× SSC, both with gentle agitation.

Multiple layers of avidin-fluorescein and biotinylated anti-avidin can be applied, but background will be significantly increased.

7. Counterstain, mount, and examine slide (see Basic Protocol, steps 9 to 12).

Amplification of Signals from Digoxigenin-Labeled Probes

Improved fluorescence signal for digoxigenin-labeled probes is achieved by “sandwiching” a nonfluorescent Fab fragment of sheep anti-digoxigenin between the digoxigenin-labeled probe and a fluorochrome-labeled rabbit anti-sheep IgG. Because more than one Fab fragment can attach to a single digoxigenin molecule and each of these Fab fragments is recognized by the fluorochrome-conjugated second antibody, this procedure leads to signal amplification.

Additional Materials (also see Basic Protocol)

- 10 µg/ml Fab fragment of sheep anti-digoxigenin (Boehringer Mannheim) in 4× SSC/1% (w/v) BSA (fraction V)
Digoxigenin amplification solution: 3.5 to 7.0 µg/ml fluorescein-conjugated rabbit anti-sheep IgG (Sigma) in 4× SSC/1% (w/v) BSA (fraction V)

NOTE: This procedure should be performed with minimum exposure to light.

1. Hybridize slide with digoxigenin-labeled probe and wash (see Basic Protocol, steps 1 through 6).
2. Add 50 µl of 10 µg/ml Fab fragment of sheep anti-digoxigenin to slide and cover with 22-mm² square of Parafilm. Incubate 30 min at 37°C.

**SUPPORT
PROTOCOL 2**

**In situ
Hybridation
and Immuno-
histochemistry**

14.7.5

3. Remove Parafilm. Wash in 4× SSC, 0.1% Triton X-100/4× SSC, then 4× SSC, 15 min each.
4. Drain excess 4× SSC and add 50 μl digoxigenin amplification solution to slide. Cover with 22-mm² square of Parafilm, place in moist chamber, wrap chamber in aluminum foil, and incubate 30 min at 37°C.
5. Repeat washes as in step 3.
6. Counterstain, mount, and examine slide (see Basic Protocol, steps 9 to 12).

ENZYMATIC DETECTION OF NONISOTOPICALLY LABELED PROBES

Hybridized probes can also be detected by enzymatic reactions that produce a colored precipitate at the site of hybridization (Fig. 14.7.1). The most commonly used enzymes for this application are alkaline phosphatase (AP) and horseradish peroxidase (HRPO). Although these enzymes can be conjugated directly to nucleic acid probes, such enzyme-coupled probes are often inappropriate for in situ hybridization to cellular preparations because probe penetration is hampered by the presence of the conjugated enzyme. Therefore, indirect methods are preferred. As illustrated in Figure 14.7.1, probes labeled by reporter molecules such as biotin or digoxigenin are detected via AP or HRPO conjugated to reporter-binding molecules such as avidin (for biotin) or antibodies (for digoxigenin). To visualize the site of probe hybridization, a slide is incubated with the appropriate enzyme substrate. The substrate is converted by the enzymatic activity to a stable colored precipitate, thus providing a permanent in situ hybridization signal that can be analyzed by phase-contrast microscopy. Because of their stability, such signals are often preferred for diagnostic analysis in routine pathology laboratories.

Table 14.7.1 summarizes the most commonly used combinations of enzyme and substrate for detection of in situ hybridization probes.

Table 14.7.1 Commonly Used Enzyme/Substrate Combinations for Detection of In Situ Hybridized Probes

Enzyme	Substrate ^a	Color ^b
Alkaline phosphatase	5-bromo-4-chloro-3-indolyl phosphate (BCIP) and nitroblue tetrazolium (NBT)	Bluish-purple precipitate
	Naphtol-AS-MX-phosphate and fast red TR	Red precipitate and red fluorescence
	Naphtol-AS-MX-phosphate and fast blue BN (Boehringer)	Blue precipitate
	Naphtol-AS-MX-phosphate and fast green BN (Boehringer)	Green precipitate
	Vector Red	Red precipitate
	Vector Black	Black precipitate
	Vector Blue	Blue precipitate
Horseradish peroxidase	3,3-diaminobenzidine tetrahydrochloride (DAB)	Brown precipitate ^c

^aThe main suppliers for these reagents are Boehringer Mannheim, Life Technologies, Promega, Sigma, and Vector Laboratories, but other suppliers deliver equivalent reagents.

^bThe detection systems listed result in hybridization signals that are generally analyzed by conventional bright-field microscopy. A fluorescence microscope is needed when utilizing the fluorescence signals produced by fast red. Optimization of the fluorescence detection procedure has been reported (Speel et al., 1992).

^cCan be intensified by silver deposition.

Enzymatic Detection Using Horseradish Peroxidase

**ALTERNATE
PROTOCOL 1**

Additional Materials (also see *Basic Protocol*)

Blocking solution: 1% (w/v) BSA in PBS (*APPENDIX 2*)
Streptavidin solution (see recipe)
0.1% (v/v) Tween 20/PBS (*APPENDIX 2*), 42°C
Biotinylated horseradish peroxidase (HRPO) solution (see recipe)
DAB substrate solution: 500 µg/ml 3,3'-diaminobenzidine tetrahydrochloride
(DAB) in PBS (*APPENDIX 2*), prepared fresh
3% H₂O₂
PBS (*APPENDIX 2*)
90% (v/v) glycerol *or* appropriate antifade mounting medium (see recipe)
24 × 60-mm coverslips
42°C shaking water bath

CAUTION: DAB is hazardous; see manufacturer's information for guidelines on handling, storage, and disposal.

NOTE: This procedure must be performed in the dark with minimum exposure to ambient light.

Apply detecting reagents

1. Hybridize slide preparation with biotinylated probe and wash slide (see *Basic Protocol*, steps 1 to 6).
2. Remove slide from 1× SSC. Drain as much buffer as possible, but do not allow slide to dry. Add 200 µl blocking solution to slide, and place 24 × 60-mm coverslip on top of applied solution. Place slide in an aluminum foil-wrapped moist chamber and incubate 30 min at 37°C.
3. Remove slide from chamber, tilt it to let coverslip slide off, and drain off as much blocking solution as possible without allowing slide to dry. Add 200 µl streptavidin solution to slide and place a 24 × 60-mm coverslip on top of applied solution. Place slide in an aluminum foil-wrapped moist chamber and incubate 30 min at 37°C.
4. Remove slide from chamber, tilt it to let the coverslip slide off, and place slide in Coplin jar containing 42°C 0.1% Tween 20/PBS. Place Coplin jar in a 42°C shaking water bath and agitate 5 min. Repeat wash twice with 42°C 0.1% Tween 20/PBS.
5. Remove slide and drain thoroughly without allowing to dry. Add 200 µl biotinylated HRPO solution to slide and place a 24 × 60-mm coverslip on top of applied solution. Incubate slide in an aluminum foil-wrapped moist chamber 30 min at 37°C.

Alternatively, steps 3 to 5 can be accomplished in a single step by incubating the slide with 3 µg/ml streptavidin-conjugated HRPO.

HRPO-conjugated anti-digoxigenin antibody should be used to detect digoxigenin-labeled probes.

6. Repeat wash as in step 4.

Enzymatically detect bound probe

7. Remove slide from wash solution and drain thoroughly without allowing to dry. Add 0.015% H₂O₂ to DAB substrate solution. Immediately add 200 µl DAB substrate solution to slide and place a 24 × 60-mm coverslip on top of applied solution. Incubate slide 10 to 20 min in the dark at room temperature.

**In situ
Hybridization
and Immuno-
histochemistry**

14.7.7

**ALTERNATE
PROTOCOL 2**

8. When the colored precipitate becomes visible to the eye, wash slide 5 min in room temperature PBS to stop the reaction.
9. If desired, apply a fluorescent counterstain to identify nuclei (see Basic Protocol, step 9).
10. Mount slide in 90% glycerol *or* appropriate antifade mounting medium (see Critical Parameters). View and photograph with a phase-contrast microscope.

Enzymatic Detection Using Alkaline Phosphatase

Additional Materials (also see Basic Protocol)

- Blocking solution: 1% (w/v) BSA/PBS (*APPENDIX 2*)
- Streptavidin solution (see recipe)
- 0.1% (v/v) Tween 20/PBS (*APPENDIX 2*), 42°C
- Biotinylated alkaline phosphatase (AP) solution (see recipe)
- Alkaline phosphatase buffer, pH 9.5 (see recipe), 42°C
- NBT/BCIP substrate solution (see recipe)
- PBS (*APPENDIX 2*)
- 90% glycerol *or* appropriate antifade mounting medium (see recipe)
- 24 × 60-mm coverslips
- 42°C shaking water bath

NOTE: This procedure must be performed in the dark with minimum exposure to ambient light.

1. Hybridize slide with biotinylated probe, wash, block, and incubate in streptavidin solution (see Alternate Protocol 1, steps 1 to 4).
2. Remove slide from 0.1% Tween 20/PBS and drain without allowing it to dry. Add 200 µl biotinylated AP solution to slide and place a 24 × 60-mm coverslip on top of applied solution. Incubate slide in an aluminum foil-wrapped moist chamber 30 min at 37°C.

Alternatively, this procedure can be accomplished in a single step by incubating the slide with 3 µg/ml streptavidin-conjugated AP (Life Technologies).

3. Remove slide from chamber, tilt it to let the coverslip slide off, and place slide in Coplin jar containing 42°C 0.1% Tween 20/PBS. Agitate 5 min at 42°C in shaking water bath. Repeat wash twice with 42°C 0.1% Tween 20/PBS.
4. Transfer slide to Coplin jar containing 42°C alkaline phosphatase buffer, pH 9.5. Agitate 5 min at 42°C. Replace buffer and agitate again 5 min at 42°C.
5. Place slide in aluminum foil-wrapped Coplin jar containing 50 ml freshly prepared NBT/BCIP substrate solution. Incubate in dark at 37°C or at room temperature (to slow down reaction) until color development is suitable.

Usually, 15 to 60 min incubation is sufficient. Longer incubation often results in higher background.

6. Wash slide 5 min in PBS at room temperature to stop reaction.
7. If desired, Apply a fluorescent counterstain to identify nuclei (see Basic Protocol, step 9).
8. Mount in 90% glycerol or appropriate antifade mounting medium (see Critical Parameters). View and photograph with a phase-contrast microscope.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Alkaline phosphatase buffer, pH 9.5

0.1 M Tris-Cl, pH 9.5

0.1 M NaCl

50 mM MgCl₂ (add immediately before use)

Store up to 1 year (without MgCl₂) at room temperature

Antifade mounting media

DABCO mounting medium: Dissolve 0.233 g 1,4-diazobicyclo-[2.2.2]octane (Sigma; 0.21 M final) in 800 µl H₂O. Add 200 µl 1 M Tris-Cl, pH 8.0 (0.02 M final) and 9 ml glycerol (90% final). Mix. Store 100-µl aliquots, wrapped in foil, at -20°C. Thaw and use once.

CAUTION: *DABCO is hazardous; see manufacturer's information for guidelines on handling, storage, and disposal.*

Phenylenediamine dihydrochloride: Dissolve 50 mg *p*-phenylenediamine dihydrochloride in 5 ml PBS (9 mM final). Adjust to pH 8 with 0.5 M carbonate/bicarbonate buffer, pH 9.0 (see recipe). Add to 45 ml glycerol (90% final), mix, and filter through 0.22 µm filter. Store in small aliquots in dark at -20°C. Thaw and use at once.

Vectashield: Purchase from Vector Laboratories; follow manufacturer's instructions.

Biotin detection solution

Dilute fluorescein-avidin DCS *or* rhodamine-avidin D (Vector Laboratories) to 2 µg/ml in 4× SSC/1% (w/v) BSA (fraction V). Prepare fresh daily.

Biotin/digoxigenin detection solution

Dilute fluorescein-avidin DCS (Vector Laboratories) and rhodamine-conjugated Fab fragment of sheep anti-digoxigenin (Boehringer Mannheim) to 2 µg/ml each in 4× SSC/1% (w/v) BSA (fraction V). Prepare fresh daily.

Biotinylated alkaline phosphatase solution

PBS (APPENDIX 2) containing:

1% (w/v) BSA

0.1% (v/v) Tween 20

2.5 µg/ml biotinylated alkaline phosphatase

Prepare fresh

Biotinylated horseradish peroxidase solution

PBS (APPENDIX 2) containing:

1% (w/v) BSA

0.1% (v/v) Tween 20

3 µg/ml biotinylated horseradish peroxidase

Prepare fresh

Carbonate/bicarbonate buffer (pH 9.0), 0.5 M

0.42 g NaHCO₃

10 ml H₂O

Adjust pH to 9 with NaOH

Store up to 1 year at room temperature

DAPI staining solution

Stock solution: Dissolve 1 mg 4',6-diamidino-2-phenylindole (DAPI; 0.3 mM final) in 10 ml H₂O. (Add a few drops of methanol before adding H₂O to help dissolve DAPI). Aliquot into aluminum foil–wrapped tubes and store a year or more at –20°C.

Working solution: Dilute stock solution 1/1000 in PBS (APPENDIX 2). Store in aluminum foil–wrapped tubes several weeks at 4°C.

CAUTION: *DAPI is hazardous; see manufacturer's information for guidelines on handling, storage, and disposal.*

Digoxigenin detection solution

2 µg/ml fluorescein *or* rhodamine-conjugated Fab fragment of sheep anti-digoxigenin (Boehringer Mannheim) in 4× SSC/1% (w/v) BSA (fraction V). Prepare fresh daily.

Master hybridization mix

1 ml 20× SSC (APPENDIX 2; 4× final)

0.5 ml 20 mg/ml nuclease-free BSA (2 mg/ml final)

1.5 ml sterile H₂O

2 ml 50% (w/v) dextran sulfate (Pharmacia Biotech, mol. wt. 500,000; autoclaved; 20% final)

Store at 4°C and use up to 6 weeks

NBT/BCIP substrate solution

Add 220 µl nitroblue tetrazolium (NBT) solution (75 mg/ml in dimethylformamide; 330 µg/ml final) to 50 ml alkaline phosphatase buffer, pH 9.5 (see recipe) and mix gently (do not vortex). Add 170 µl 5-bromo-4-chloro-3-indolyl phosphate (BCIP) solution (50 mg/ml in dimethylformamide; 170 µg/ml final) and mix gently again. Prepare fresh each time.

CAUTION: *Dimethylformamide is hazardous; see manufacturer's information for guidelines on handling, storage, and disposal.*

Propidium iodide staining solution

Stock solution: Dissolve 1 mg propidium iodide in 10 ml H₂O (0.15 mM final). Aliquot into aluminum foil–wrapped tubes and store up to 1 year at –20°C.

Working solution: Dilute stock solution 1/1000 in PBS (APPENDIX 2) and store in aluminum foil–wrapped tube up to 6 months at 4°C.

CAUTION: *Propidium iodide is hazardous; see manufacturer's information for guidelines on handling, storage, and disposal.*

Sonicated salmon sperm DNA, 10 mg/ml

Dissolve 10 mg salmon sperm DNA (Worthington) in 1 ml sterile water in a polycarbonate tube. Sonicate five times, 30 sec each time, at maximum power, chilling tube on ice between bursts. Check molecular size of DNA by gel electrophoresis (UNIT 2.5A); it should be 200 to 400 bp. Store in 50-µl aliquots up to 1 year at –20°C.

Streptavidin solution

PBS (APPENDIX 2) containing:

1% (w/v) BSA

0.1% (v/v) Tween 20

3 µg/ml streptavidin

Prepare fresh

COMMENTARY

Background Information

In situ hybridization (ISH) is a powerful technology for examining the expression, processing, and location of specific nucleic acid sequences in single cells or in tissue sections. Historically, ISH has been performed using isotopically labeled probe sequences that are detected by autoradiography (see *UNITS 14.3 & 14.4*; see also Knoll and Lichter, 1995 for a more extensive version of this unit optimized for hybridization to chromosomes and nuclei). This method has been used to identify and localize both repetitive and unique sequences of DNA in chromosomes (Morton et al., 1984; Lawrence et al., 1988) as well as RNA transcripts (Zhang et al., 1994). In recent years, nonisotopic ISH has gained recognition because of technical advances that make its sensitivity comparable to that of isotopic ISH (Lichter et al., 1988, 1990). Nonisotopic techniques can be used in combination with radioactive hybridization and immunohistochemistry as a powerful approach to detect more than one nucleic acid target and antigen in the same section (Trembleau et al., 1993).

The method described in this unit is a modification of methodology previously described for fluorescent detection of single-copy nuclear DNA (Lawrence et al., 1988; Knoll and Lichter, 1995). The approach described here has been adapted to allow direct visualization of specific transcripts. The method has been used most recently to study the sites at which RNA transcription and processing occur—in particular, to localize splicing events and the factors associated with processing nascent transcripts in mammalian nuclei (Zhang et al., 1994).

Results are obtained more rapidly with nonisotopic ISH because there is no need for lengthy autoradiographic exposures. Nonisotopically labeled probes do not require special handling and disposal. With fluorochrome detection using multi-band-pass filters, two or more sequences can be examined simultaneously and distinguished from one another. Isotopic ISH (IISH; *UNITS 14.3 & 14.4*), however, is useful for mapping small sequences that cannot be detected easily by nonisotopic methods.

FISH

The advantages and availability of commercially prepared probes have contributed to widespread application of nonisotopic ISH, especially fluorescence in situ hybridization (FISH), in clinical and research laboratories

(Knoll and Lichter, 1995). Probes are generally labeled by nick translation, but other methods such as random oligonucleotide-primed labeling are also effective. Detailed protocols for both labeling procedures are provided in *UNIT 3.18*. Biotin-labeled and digoxigenin-labeled deoxynucleotides are most frequently used for labeling, but other modified nucleotides, such as direct-labeled fluorescent nucleotides (e.g., fluorescein-12-dUTP; Boehringer Mannheim) are also used (Wiegant et al., 1991). Biotin- and digoxigenin-labeled probes are detected by incubating the hybridized preparation with a fluorochrome- or enzyme-conjugated reporter-binding molecule (Fig. 14.7.1). Digoxigenin detection yields lower background fluorescence than biotin detection because digoxigenin is not a natural component of mammalian cells. Hybridized probe is detected by fluorescent (Lawrence et al., 1989) or enzymatic methods (Lawrence and Singer, 1988) and viewed by transmission, reflection, or epifluorescence microscopy. Analysis may be facilitated by using cameras such as charge-coupled device (CCD) cameras, or for three-dimensional analysis, confocal microscopy.

Enzymatic detection

In this unit, two protocols for enzymatic detection of biotinylated and digoxigenin-labeled probes are described; one uses horseradish peroxidase (HRPO) with diaminobenzidine (DAB) as substrate; the other uses alkaline phosphatase with 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and nitroblue tetrazolium (NBT) as substrates. These procedures were developed in several different laboratories and the protocols described here are based on published work (e.g., Lichter et al., 1988; Emmerich et al., 1989).

Enzyme-mediated detection procedures result in permanent colored precipitates (Figs. 14.7.1 & 14.7.2); therefore, these techniques are often preferred for preparing routine specimens that are to be archived. However, until recently it has been very difficult to simultaneously label different targets in the same specimen, a goal that is achievable using fluorescence methods, because the precipitates produced by HRPO- and AP-mediated reactions were often hard to distinguish in cellular preparations. With the advent of new substrates (see Table 14.7.1) that yield differently colored precipitates, this limitation has been overcome. For a typical dual-color experiment, the bio-

tinylated first probe is detected using AP and a substrate that produces a red precipitate, and the second digoxigenin-labeled probe is detected using HRPO and a substrate that produces a brown precipitate.

Critical Parameters and Troubleshooting

For a detailed discussion of special considerations for in situ hybridization to RNA see *UNIT 14.3*, Commentary.

FISH

To maximize probe hybridization efficiency and signal-to-noise ratio, it is important to have high-quality cytological preparations. For best hybridization results with hypotonically treated and methanol/acetic acid-fixed cells, unstained nuclei should appear dark gray when examined with a phase-contrast microscope. They should not appear black and refractile or light gray and almost invisible. In addition, there should be minimal extracellular debris in the preparations. Avidin nonspecifically binds to extracellular debris and increases background. Pretreatment—e.g., with pepsin or pronase—is recommended when there is too much debris (*UNIT 14.3*).

Probe hybridization efficiency and background are also influenced by the size of the labeled probe. Optimal probe size is 100 to 500 bp; background increases significantly when this size is exceeded. If the probe extends beyond 800 bp, additional nick translation is suggested. If the DNA is difficult to nick to smaller sizes (i.e., it contains inhibiting contaminants), fresh probe DNA should be prepared. An additional phenol/chloroform extraction and ethanol precipitation of the unlabeled probe DNA is often sufficient to prepare clean DNA for subsequent nick translations. Biotin- and digoxigenin-labeled DNA should not be subjected to phenol/chloroform extraction because it will partition to the interface.

If the hybridization signal is not visible or is difficult to see, it is possible that the target is too small, the level of nucleotide incorporation is inadequate, the concentration of probe utilized is too low, or denaturation is not optimal. The level of incorporation of labeled nucleotide should be checked and the concentration of probe utilized for hybridization should be increased, if necessary (see *UNIT 3.18*). Amplification of the signal may improve fluorescence intensity. Additionally, the optimal antifade mounting medium for the fluorochrome de-

tected should be used. DABCO (1,4-diazobicyclo-[2.2.2]octane) permits optimal green fluorescence of fluorescein (Johnson et al., 1982); *p*-phenylenediamine dihydrochloride (Johnson and de C. Nogueira Araujo, 1981) and Vectashield (Vector) permit optimal red fluorescence of rhodamine and Texas red. Fluorescein fluorescence is less bright with *p*-phenylenediamine dihydrochloride and Vectashield but is still very good. Vectashield does not readily discolor and therefore does not obscure a less intense signal, as may occur with *p*-phenylenediamine. Digital imaging systems (e.g., CCD cameras and confocal microscopes) are also useful to detect small or weak hybridization signals (see Lichter et al., 1991, for an overview of digital imaging systems). These systems, however, are expensive and not accessible to many laboratories.

Image shifting as filter sets are changed is a potential problem in orienting sequences that are physically close, or in localizing a hybridization signal with precision. This problem is solved for practical applications by simultaneous viewing of multiple-color fluorescence through a dual- or triple-band-pass filter set, instead of viewing each color through a separate filter set.

Enzymatic detection

In general, enzymatic detection can be carried out in two ways, as illustrated in Figure 14.7.2. The reporter-binding molecule can be coupled directly to the enzyme, as in Figure 14.7.2A, requiring only one incubation with detection reagents prior to the enzyme reaction. However, these conjugates are often large and do not always penetrate into cytological preparations as well as the individual reporter-binding and enzyme molecules. Therefore, a two-step application of detection reagents, as outlined in Figure 14.7.2B and described in the protocols, may be preferred. Weak signals can be amplified as indicated in Figure 14.7.2C.

An important parameter of AP-mediated detection is the accurate adjustment of pH in the alkaline phosphatase buffer to pH 9.5. Another extremely important critical parameter is the time period of the color reaction described in the protocol. The reaction can continue for hours. Although longer reaction times may result in an increase of detection sensitivity, in general they also tend to result in a considerable increase in background. In contrast, HRPO-mediated detection using DAB/H₂O₂ is substrate-inactivated. Therefore, longer incuba-

tion times do not improve signal intensity, but can still increase background.

Silver intensification increases the optical density of the colored precipitate, facilitating the use of enzyme-mediated procedures for electron microscopy. Applications of silver-intensified DAB products have been reported (Manuelidis and Ward, 1984).

Anticipated Results

Detection of RNAs either in the nucleus or cytoplasm is dependent upon the size of the RNA target, its abundance, and its distribution. Generally, molecules smaller than a few kilobases are not easily visible without the use of amplification procedures, which may tend to enhance background. Examples of results of these experiments are described and illustrated in many of the references cited here (e.g., Lawrence et al., 1988; Zhang et al., 1994).

This FISH protocol has allowed detection of cDNA probes as small as 1.8 kb without use of signal amplification or digital imaging.

The sensitivity of enzymatic detection is less than that of FISH so this method is not suitable for probes of cosmid size or smaller. However, many diagnostic probes are highly complex and the colorimetric signals obtained with enzymatic detection are permanent, so enzymatic detection protocols are often preferred in clinical laboratories.

Time Considerations

Denaturation and hybridization of probe and cellular DNA require ≥ 4 hr with a repetitive probe, and hybridization can be performed in a few hours or overnight. Post-hybridization washing, signal detection, and counterstaining require a total of 3 to 4 hr. An additional 2 to 3 hr are required if signal amplification is performed. After mounting in antifade medium, preparations from cDNA can be stored at -20°C for ~ 2 weeks.

The time required for enzyme-mediated detection is not significantly longer than that required for fluorescence detection.

Literature Cited

Emmerich, P., Loos, P., Jauch, A., Hopman, A.H.N., Wiegant, J., Higgins, M., White, B.N., van der Ploeg, M., Cremer, C., and Cremer, T. 1989. Double in situ hybridization in combination with digitized image analysis: A new approach to study interphase chromosome topography. *Exp. Cell Res.* 181:126-140.

Johnson, G.D. and de C. Nogueira Araujo, G.M. 1981. A simple method of reducing the fading of

immunofluorescence during microscopy. *J. Immunol. Methods* 43:349-350.

- Johnson, G.D., Davidson, R.S., McNamee, K.C., Russell, G., Goodwin, D., and Holborrow, E.L. 1982. Fading of immunofluorescence during microscopy: A study of its phenomenon and its remedy. *J. Immunol. Methods* 55:231-242.
- Knoll, J.H.M. and Lichter, P. 1995. In situ hybridization to metaphase chromosomes and interphase nuclei. In *Current Protocols in Human Genetics* (N.C. Dracopoli, J.L. Haines, B.R. Korf, D.T. Moir, C.C. Morton, C.E. Seidman, J.G. Seidman, and D.R. Smith, eds.) pp. 4.3.1-4.3.29. John Wiley & Sons, New York.
- Lawrence, J.B. and Singer, R.H. 1988. Quantitative analysis of in situ hybridization methods for the detection of actin gene expression. *Nucl. Acids Res.* 13:1777-1799.
- Lawrence, J.B., Villnave, C.A., and Singer, R.H. 1988. Sensitive, high-resolution chromatin and chromosome mapping in situ: Presence and orientation of two closely integrated copies of EBV in a lymphoma cell line. *Cell* 52:51-61.
- Lawrence, J.B., Singer, R.H., and Marselle, L.M. 1989. Highly localized tracks of specific transcripts within interphase nuclei visualized by in situ hybridization. *Cell* 57:493-502.
- Lichter, P., Cremer, T., Tang, C.C., Watkins, P.C., Manuelidis, L., and Ward, D.C. 1988. Rapid detection of human chromosome 21 aberrations by in situ hybridization. *Proc. Natl. Acad. Sci. U.S.A.* 85:9664-9668.
- Lichter, P., Tang, C.C., Call, K., Hermanson, G., Evans, G., Housman, D., and Ward, D.C. 1990. High-resolution mapping of human chromosome 11 by in situ hybridization with cosmid clones. *Science* 247:64-69.
- Lichter, P., Boyle, A.L., Cremer, T., and Ward, D.C. 1991. Analysis of genes and chromosomes by nonisotopic in situ hybridization. *Genet. Anal. Techn. Appl.* 8:24-35.
- Manuelidis, L. and Ward, D.C. 1984. Chromosomal and nuclear distribution of the *HindIII* 1.9 kb human DNA repeat segment. *Chromosoma* 91:28-38.
- Morton, C.C., Kirsch, I.R., Taub, R.A., Orkin, S.H., and Brown, J.A. 1984. Localization of the beta-globin gene by chromosomal in situ hybridization. *Am. J. Hum. Genet.* 36:576-585.
- Speel, E.J.M., Schutte, B., Wiegant, J., Ramaekers, F.C., and Hopman, A.H.N. 1992. A novel fluorescence detection method for in situ hybridization, based on the alkaline phosphatase-fast red reaction. *J. Histochem. Cytochem.* 40:1299-1308.
- Trembleau, A., Roche, D., and Calas, A. 1993. Combination of non-radioactive and radioactive in situ hybridization with immunohistochemistry: A new method allowing the simultaneous detections of two mRNAs and one antigen in the same brain tissue section. *J. Histochem. Cytochem.* 41:489-498.

Wiegant, J., Tied, T., Nederlof, P.M., van der Ploeg, M., Tanke, H.J., and Raap, A.K. 1991. In situ hybridization with fluoresceinated DNA. *Nucl. Acids Res.* 19:3237-3241.

Zhang, G. Taneja, K., Singer, R.H., and Green, M.R. 1994. Localization of pre-mRNA splicing in mammalian nuclei. *Nature* 372:809-812.

Contributed by Joan H.M. Knoll
Harvard Medical School
Boston, Massachusetts

Peter Lichter (enzymatic detection)
Deutsches Krebsforschungszentrum
Heidelberg, Germany

In situ Polymerase Chain Reaction and Hybridization to Detect Low-Abundance Nucleic Acid Targets

This unit presents a novel approach for detecting low-abundance nucleic acid targets in nuclear and cytoplasmic regions by amplification of specific target sequences using an in situ polymerase chain reaction (ISPCR). If the target sequence is RNA, ISPCR is preceded by in situ reverse transcription. Following ISPCR, in situ hybridization is performed. The Strategic Planning section discusses specific considerations for the overall approach, including selection and design of primers, determination of annealing temperatures for reverse transcription and amplification, design of controls, and validation of experiments. Basic Protocol 1 describes ISPCR amplification of DNA or RNA targets. An Alternate Protocol describes a variant in situ method for simultaneously reverse transcribing and amplifying RNA transcripts using recombinant *Thermos thermophilus* (*rTth*) polymerase. Basic Protocol 2 describes in situ hybridization and detection of amplified targets.

Two specific variations on the standard in situ hybridization protocols (in *UNITS 14.1 & 14.3*) are recommended for this particular application. The first is the use of slides subbed with 3-aminopropyltriethoxysilane (AES) in place of gelatin-subbed slides (see recipe in *UNIT 14.1*). Support Protocol 1 describes preparation of AES-subbed slides. Support Protocol 2 outlines methods for attaching tissue sections and cells to AEC-subbed slides in preparation for ISPCR. The second variation is the use of ³³P-labeled oligonucleotide probes instead of ³⁵S-labeled riboprobes for in situ hybridization. Support Protocol 3 describes radiolabeling of oligonucleotide probes using ³³P.

STRATEGIC PLANNING

Figure 14.8.1 outlines the major steps for in situ amplification and hybridization. It is essential that appropriate planning and consideration be given to each aspect of the project—including design of primers, amplification parameters, detection and fixation conditions, and validation and controls—before proceeding with the actual protocols. These considerations and the corresponding procedural variations, which are discussed below, must be defined and optimized for each system. For further details on optimization of various aspects of the protocols, see Critical Parameters.

Design of Primers to Reverse Transcribe and Amplify Target

The design of primers for amplification reactions in general requires careful planning (see *UNIT 15.1*); for in situ amplification, the design of the primer set is even more critical. In the authors' laboratory, antisense downstream primers for the gene of interest are used, as the authors already know the sequence of most genes that they study. However, it is equally possible to use oligo(dT) primers to convert all mRNA populations into cDNA, and then perform the in situ amplification for a specific cDNA. This technique may be useful when several different gene transcripts are being amplified at the same time in a single cell. For example, if attempting to detect expression of various cytokines in a cell or tissue section, it is possible to use an oligo(dT) primer to reverse transcribe all of the mRNA copies in the specimen. More than one type of cytokine cDNA can then be amplified, and the various types can be detected with differentially labeled probes (see Commentary for discussion of detecting multiple signals in individual cells).

There are two possible choices in designing a specific method to detect expression of a particular RNA. The more elegant method is to use primer pairs that are complementary to the sequence spanning spliced sequences of mRNA, as these particular sequences will

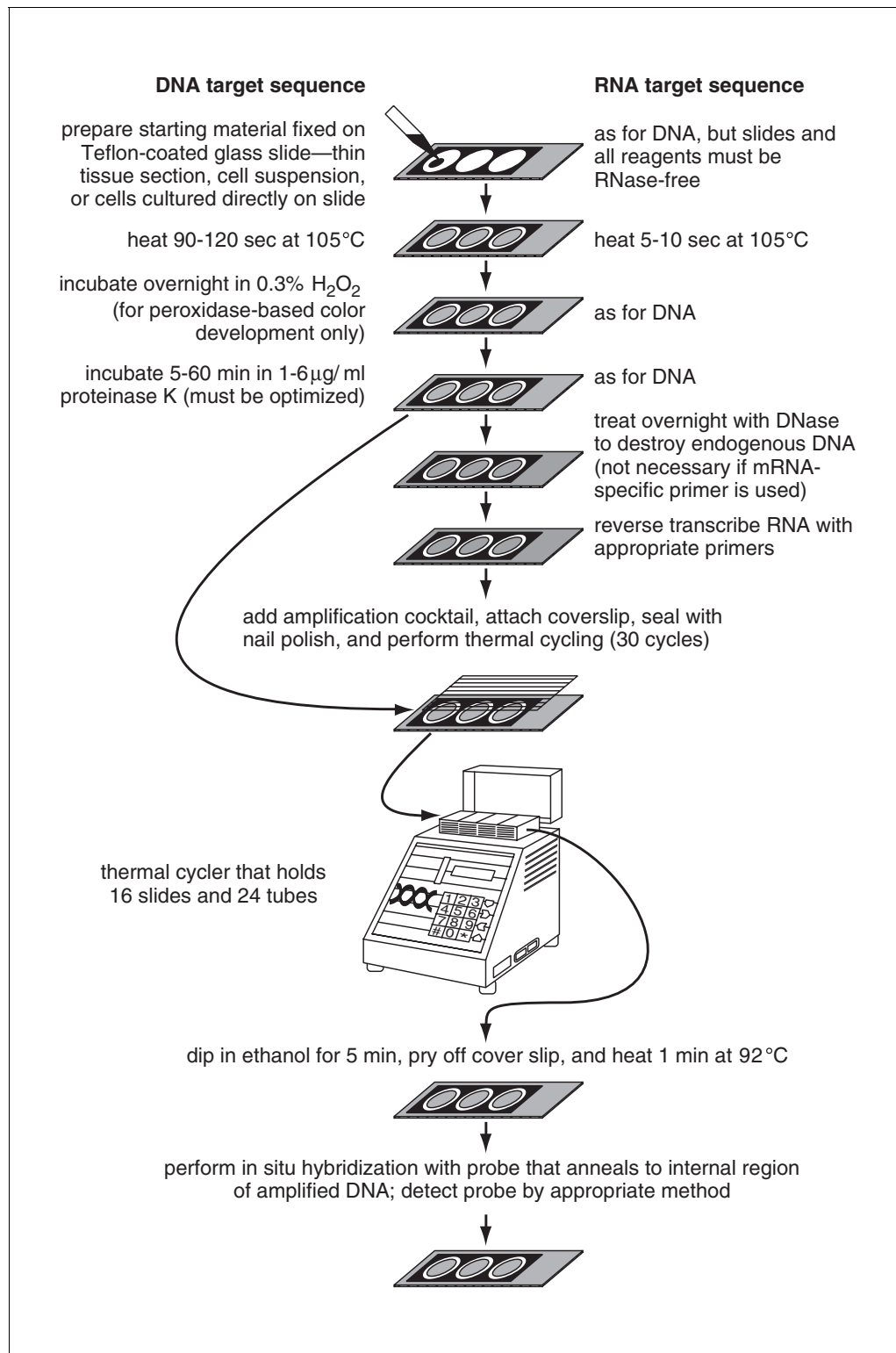


Figure 14.8.1 Flow chart for in situ amplification/hybridization.

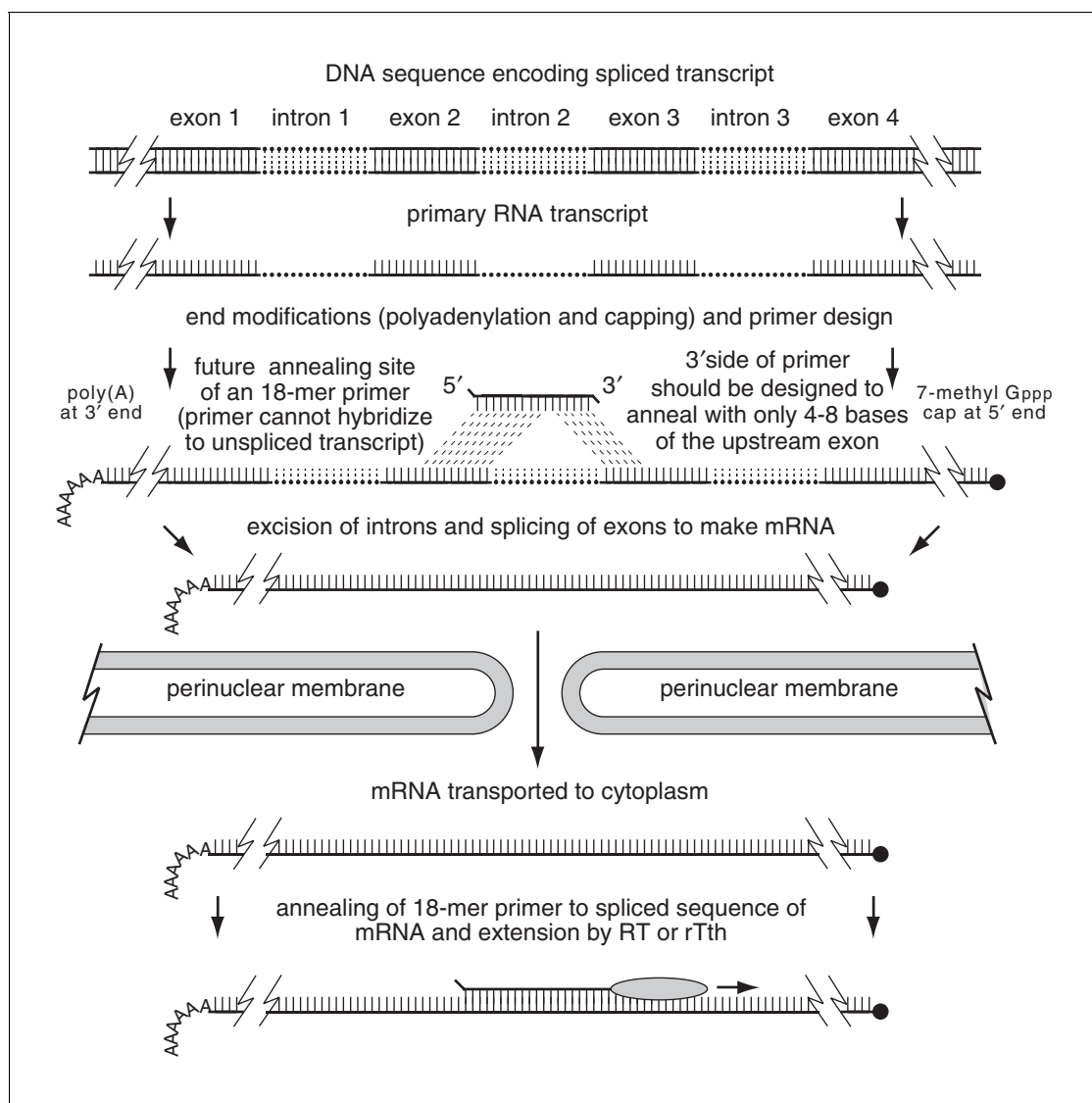


Figure 14.8.2 Design of primers to amplify mRNA without interference from DNA. Primers are designed so that their sequences span the splice junction where two exon sequences are joined. The resulting sequence thus exists only in mRNA, not in the encoding DNA. Use of such primers allows DNase treatment to be omitted and facilitates simultaneous amplification of RNA and DNA signals.

be found only in spliced mRNA and will not fully hybridize to the encoding sequence in DNA (see Fig. 14.8.2). Thus, by using these primers, it is possible to omit the DNase treatment (to destroy endogenous DNA) and proceed directly to reverse transcription. The more brute-force, yet often necessary, approach is to treat the cells or tissue with DNase subsequent to the proteinase K digestion. This step destroys all of the endogenous DNA in the cells so that only RNA survives for amplification and subsequent detection.

In all reverse transcription reactions, it is advantageous to reverse transcribe only relatively small fragments of mRNA (<1500 bp). Larger fragments may not be completely reverse transcribed due, for example, to the presence of secondary structures. Furthermore, the reverse transcriptase enzymes—avian myeloblastosis (AMV) reverse transcriptase and Moloney murine leukemia virus (MoMuLV) reverse transcriptase, at least—are not very efficient in transcribing long mRNAs. However, this size restriction does not apply to DNA amplification; DNA target lengths up to 5000 bp are routinely amplified in situ in the authors' laboratory. Several new recombinant enzymes and buffer systems

have been described that allow efficient amplifications of much larger fragments (up to 10 kb) of DNA or cDNA (Lundberg et al., 1991; Barnes, 1994). An alternative approach is to create a multiple primer set, choosing primers positioned 200 to 300 nucleotides apart along the sequence to be amplified (see Haase et al., 1990).

Several additional points should be kept in mind when designing primers for ISPCR and reverse transcription: (1) the length for both sense and antisense primers should be 18 to 22 bp; (2) at the 3' ends, primers should contain at least one GC-type dinucleotide (i.e., GG, CC, GC, or CG) to facilitate complementary-strand formation (as two GC-type base pairs will provide six hydrogen bonds as compared to four with two AT-type base pairs); (3) the preferred GC content of the primers is 45% to 55%; (4) primers should be designed so that they do not form intra- or interstrand base pairs; (5) the 3' ends of the primers should not be complementary to each other, or they will form primer-dimers; and (6) reverse transcription primers should be designed so that they do not contain secondary structures (see Pallansch et al., 1990)

Determining Optimal Annealing Temperature

The optimal primer annealing temperature for reverse transcription and DNA amplification is usually 2°C above the T_m of the primer, which is calculated according to the formula:

$$T_m \text{ of the primers} = 81.5^\circ\text{C} + 16.6 (\text{Log } M) + 0.41 (\%GC) - 500/n$$

where T_m is the melting temperature of the primer, n is the length of the primer, and M is the molarity of the salt in the buffer (usually 0.047 M for DNA reactions and 0.070 M for reverse transcription reactions)

If using AMV reverse transcriptase, the value of the melting temperature will be lower according to the following formula:

$$T_m \text{ of the primers} = 62.3^\circ\text{C} + 0.41 (\%GC) - 500/n$$

These formulas provide only an approximate temperature for annealing, because base stacking, near-neighbor effect, and buffering capacity may play a significant role in determining the annealing temperature of a particular primer.

The logic of determining the correct annealing temperature for ISPCR is that during amplification, spurious products often appear in addition to those desired. Even if the cells do not contain DNA homologous to the primer sequences, many artifactual bands may appear as a result of false priming, which will occur if the determined melting temperature (T_m) between primer and template is not accurate. Annealing at a temperature too far above the T_m will yield no products, and annealing at a temperature too far below the T_m will often give unwanted products resulting from false priming.

Recently, MJ Research has devised a thermal cycler (the PTC-100) that has the capacity to perform in situ gene amplification on slides and solution-based amplification in tubes simultaneously in the same block. This kind of thermal cycler can be very useful in determining optimal amplification conditions for a gene of interest. Many methods for eliminating high background resulting from false priming in PCR have appeared in the literature, including hot-start PCR (Nuovo et al., 1992, 1994a) as well as use of DMSO, formamide, and anti-Taq DNA polymerase antibodies (e.g., TaqStart from Clontech; Nuovo, 1994a). These variations and modifications are discussed in *UNIT 15.1*.

The annealing temperature should be optimized with solution-based reactions before the corresponding in situ reactions are attempted, as in situ reactions are simply not as robust as solution-based ones (O'Leary et al., 1994). One hypothesis is that this is due to primers

not having easy access to DNA templates inside cells and tissues; the numerous membranes, folds, and other small structures can prevent primers from binding complementary sites as readily as they do in solution. In solution-based PCR, determining the appropriate optimal annealing temperature may not be as critical as in ISPCR, because primers in solution may anneal at the correct temperatures while temperatures are changing during thermal cycling. For example, if the optimal annealing temperature is empirically calculated to be 48°C but in reality is 52°C, freely moving primers may anneal at the proper temperature as the temperature cycles on the downward slope from the denaturation to the calculated annealing temperature of 48°C. For ISPCR, determining the optimal annealing temperature is more critical because primers do not move as freely as a result of partial hindrance by intracellular organelles, multiple layers of membranes, and numerous subcompartments. Two additional ways to determine the real annealing temperature in situ are to use “touchdown PCR” (Don et al., 1991), or to utilize a thermal cycler designed for determining actual annealing temperature (e.g., Robocycler from Stratagene).

Validation and Controls

The validity of in situ amplification/hybridization should be examined in every run. Attention is especially necessary in laboratories that are using the technique for the first time, because occasional technical pitfalls lie on the path to mastery. Even in an experienced laboratory, it is necessary to continuously validate the procedure and to confirm the efficiency of amplification. In the authors’ laboratory, two or three sets of experiments are run simultaneously in multiwell slides. This is done not only to validate the amplification, but also to confirm the subsequent hybridization/detection steps.

The authors frequently use ISPCR in working with HIV-1. A common validation procedure used in the authors’ laboratory is to mix HIV-1-infected cells and uninfected cells in a known proportion (e.g. 1:10 or 1:100), then to confirm that the results are appropriately proportional. To examine the efficiency of amplification, the authors use a cell line that carries a single copy or two copies of cloned HIV-1 virus, then check that proper amplification and hybridization have occurred. In all amplification procedures, the authors use one slide in which the amplified cells are hybridized with an unrelated probe as a control for nonspecific binding. The authors also use HLA-DQ α probes and primers with human peripheral blood mononuclear cells (PBMCs) as positive controls to check various parameters of the system. If tissue sections are used, a cell suspension lacking the gene of interest can be used as a control. These cells can be added on top of the tissue section, then retrieved after the amplification procedure and analyzed with the specific probe to see if the signal from the tissue leaked out and entered the cells floating above.

It is suggested that researchers carefully design and employ appropriate positive and negative controls for their specific experiments. In the case of reverse transcription/in situ amplification, RNAs for β -actin or HLA-DQ α , or other abundant endogenous RNAs may be used as the positive markers. Of course, a reverse-transcription-negative control should always be used for reverse transcription/in situ amplification, as well as DNase-treated and non-DNase-treated controls. Controls without DNA polymerase should always be included, with and without primers.

**IN SITU PCR (ISPCR) AMPLIFICATION OF DNA AND RNA TARGETS
WITH IN SITU REVERSE TRANSCRIPTION FOR RNA**

The first steps of the protocol comprise the basic preparatory work that must be done before performing any in situ amplification/hybridization procedures; subsequent steps describe separate methods for amplifying DNA or RNA targets. For optimal results, use of slides subbed with AES (see Support Protocol 1) is recommended. Slides with fixed tissue specimens are pretreated with proteinase K to optimize resolution and reduce background during subsequent reactions. For RNA, the target transcript is reverse transcribed to make DNA copies suitable for amplification. The last several steps of the protocol give the details of the ISPCR amplification process.

Materials

Slides containing fixed specimens (see Support Protocol 2)
0.3% H₂O₂ in PBS (prepare fresh)
PBS (APPENDIX 2)
1 mg/ml proteinase K (Sigma; store in aliquots at -20°C)
RNase-free DNase solution (see recipe)
Rinse buffer: RNase-free DNase solution (see recipe) without DNase
DEPC-treated H₂O (UNIT 4.1)
10× AMV/MoMuLV reaction buffer (see recipe)
10 mM 4dNTP mix: 10 mM each dNTP in TE buffer, pH 7.5 (see APPENDIX 2 for TE buffer; store 4dNTP mix at -20°C)
40 U/μl RNasin (Promega)
20 μM downstream primer (for reverse transcription; see Strategic Planning)
20 U/μl avian myeloblastosis virus (AMV) or Moloney murine leukemia virus (MoMuLV) reverse transcriptase *or* 20 U/μl SuperScript II (with 5× reaction buffer; Life Technologies)
0.1 M DTT
25 μM forward and reverse primers (for PCR; see Strategic Planning and UNIT 15.1)
1 M Tris·Cl, pH 8.3 (APPENDIX 2)
1 M KCl (APPENDIX 2)
100 mM MgCl₂
0.01% (w/v) gelatin
5 U/μl *Taq* DNA polymerase
100% ethanol
2× SSC (APPENDIX 2)
55°C (optional), 92°C, 95°C, and 105°C heating blocks accommodating glass slides
Moist chamber (UNIT 14.2)
42°C incubator (optional)
20 × 60-mm glass coverslips
Clear nail polish or varnish (see Critical Parameters)
Thermal cycler accommodating glass slides (see Critical Parameters)

NOTE: Where the target sequence is RNA, all reagents should be prepared with DEPC-treated water (UNIT 4.1). In addition, the AES-subbed glass slides and all glassware should be RNase free; in the authors' laboratory this is ensured by baking overnight in an oven usually set at 250° to 300°C.

Pretreat samples with peroxide and proteinase K

1. Incubate slide with fixed specimens 5 to 120 sec on a 105°C heating block.

This step serves to stabilize the cells or tissue on the glass surface of the slide. The time is absolutely critical, and it may be necessary to experiment with different time

periods in order to optimize for specific tissues. The authors' laboratory routinely uses 90 sec for DNA target sequences, and 5 to 10 sec for RNA sequences. The shorter incubation is recommended for RNA targets because certain mRNAs may be unstable at high temperature.

2. Incubate slide overnight at 37°C or room temperature in 0.3% H₂O₂ in PBS, then wash once with PBS.

This treatment will inactivate any endogenous peroxidase activity that may interfere with subsequent assays; it is required only where peroxidase-based color development (see Basic Protocol 2) will be used for hybridization/detection.

3. Dilute 1 ml of 1 mg/ml proteinase K in 150 ml PBS (6 µg/ml final). Immerse slide in this solution and incubate at room temperature. After 5 min, examine cells under microscope at 400×; if the majority of cells of interest exhibit small round “bubbles,” “blobs,” or “peppery dots” (as in Fig. 14.8.4), proceed immediately to step 4. Otherwise, continue incubation up to 60 min, examining the slide at 400× at 5 min intervals and proceeding immediately to step 4 at the point where cell surface bubbles appear.

Performing this treatment properly is absolutely critical. The time and temperature of incubation should be optimized carefully for each cell line or tissue-section type (see Critical Parameters). If development of “peppery dots” takes >2 hr at room temperature, digestion may be performed at 48° or 55°C to speed up the process. Examination for dots should be done with an ordinary bright-field microscope with light adjusted for optimal visualization. Too much digestion will either cause membranes to deteriorate during repeated denaturation or (in the worst case) cause signals to leak out of the cells. In the former case, cells will not contain the signal and high background will result. In the latter case, many cells will show pericytoplasmic staining, representing the leaked signals contaminating other cells (see Basic Protocol 2). Attention to detail at this step can often mean the difference between success and failure, and this procedure should be practiced rigorously with extra samples before proceeding to the amplification steps.

Figure 14.8.3 shows typical cytoplasmic membranes before proteinase K treatment. Figure 14.8.4 shows membranes exhibiting “bubbles” after treatment.

4. Heat slide 2 min in a 95°C heating block to inactivate the proteinase K, then rinse 10 sec in PBS and 10 sec in water. Allow slide to air dry.

To reverse transcribe RNA targets, proceed with step 5, except if primer pairs are to be used that are complementary to the sequence joining spliced sequences of mRNA (to specifically amplify mRNA; see Strategic Planning, discussion of design of primers to reverse transcribe and amplify target), in which case proceed to step 7. To amplify DNA targets, proceed to step 10.

Treat slides with DNase (RNA samples only)

5. Add 10 µl RNase-free DNase solution to each well of slide. Incubate overnight at 37°C in a moist chamber.

For liver tissue, this incubation should be extended an additional 18 to 24 hr.

6. Rinse slide once with rinse solution, then twice with DEPC-treated water and allow to air dry.

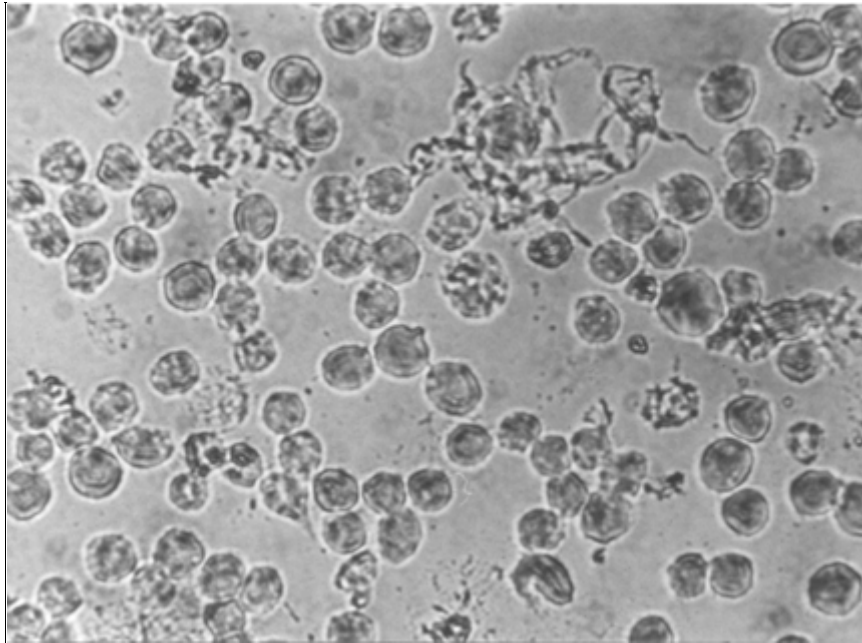


Figure 14.8.3 Lymphocytes before proteinase K treatment. Cells have been subjected to heat treatment and fixation. Note the smooth cytoplasmic membranes without “bubbles”, “peppery dots,” or “salt and pepper dots” on the cell surface.

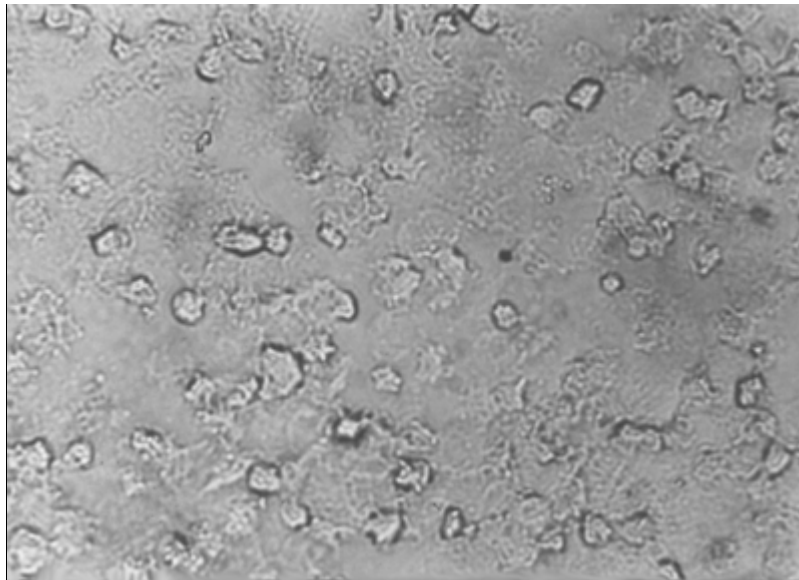


Figure 14.8.4 Lymphocytes after proteinase K treatment. Note the minute cell surface “bubbles,” which are presumably the result of partial digestion of transmembrane proteins. The “bubbles” should appear uniform and be fairly evenly distributed.

Reverse transcribe RNA

If using AMV or MoMuLV reverse transcriptase:

7a. Make up the following reverse transcription cocktail (20 μ l total volume):

- 2 μ l 10 \times AMV/MoMuLV RT buffer
- 2 μ l 10 mM 4dNTP mix
- 0.5 μ l 40 U/ μ l RNasin
- 1.0 μ l 20 μ M downstream primer
- 0.5 μ l 20 U/ μ l AMV or MoMuLV reverse transcriptase
- 8 μ l DEPC-treated H₂O.

Avian myeloblastosis virus (AMV) reverse transcriptase and Moloney murine leukemia virus (MoMuLV) reverse transcriptase give comparable results in the authors' laboratory. Other reverse transcriptase enzymes will probably work also. However, it is important to read the manufacturer's descriptions of the reverse transcriptase enzyme and to make certain that the proper buffer is used. An alternative reverse transcriptase enzyme, SuperScript II (Life Technologies), has significantly reduced RNase H activity and is therefore suitable for reverse transcription of long mRNAs. It is also suitable for routine reverse transcription, and in our laboratory has proven to be more efficient than the two enzymes described above.

If using SuperScript II reverse transcriptase:

7b. Make up the following reverse transcription cocktail (20 μ l total volume):

- 4 μ l 5 \times reaction buffer (supplied with enzyme)
- 2 μ l 10 mM 4dNTP mix
- 0.5 μ l 4 U/ μ l RNasin
- 1.0 μ l 20 μ M downstream primer
- 0.5 μ l 20 U/ μ l SuperScript II
- 1.2 μ l 0.1 M dTT
- 4.8 μ l DEPC-treated H₂O.

RNasin inhibits ribosomal RNases and should be included for optimal yields.

8. Add 10 μ l reverse transcription cocktail to each well of slide and carefully cover each well with a 20 \times 60-mm glass coverslip. Incubate 1 hr at 42°C or 37°C in a moist chamber.
9. Incubate slides 2 min at on a 92°C heating block. Remove coverslips and wash slides twice with water.

Perform ISPCR amplification

10. Make up the following amplification cocktail (100 μ l total volume):

- 5 μ l 25 μ M forward primer (1.25 μ M final)
- 5 μ l 25 μ M reverse primer (1.25 μ M final)
- 2.5 μ l 10 mM 4dNTP mix (200 μ M each dNTP final)
- 1.0 μ l 1.0 M Tris-Cl, pH 8.3 (10 mM final)
- 5.0 μ l 1.0 M KCl (50 mM final)
- 2.5 μ l 100 mM MgCl₂ (2.5 mM final)
- 10 μ l 0.01% gelatin (0.001% final)
- 2 μ l 5 U/ μ l Taq DNA polymerase (0.1 U/ μ l final)
- 66 μ l H₂O.

Appropriate primers for HIV-1 are SK 38 (forward) and SK 39 (reverse). Thermostable DNA polymerases other than Taq DNA polymerase have also been used quite successfully.

11. Layer 8 μl (if using 3-well slide) or 12 to 20 μl (if using single-well slide) ISPCR amplification cocktail onto each well using a 20 μl micropipettor, so that the whole surface of the well is covered with the solution.

Be careful not to touch the surface of the slide with the tip of the micropipettor.

12. Place a 20 \times 60-mm glass coverslip over each slide and carefully seal the edge of the coverslip to the slide with clear nail polish or varnish.

If using tissue sections, use a second slide instead of a cover slip. Be certain to carefully paint the polish around the entire periphery of the coverslip or the edges of the dual slide, as the polish must completely seal the coverslip/slide assembly in order to form a small reaction chamber that can contain the water vapor during thermal cycling (see Critical Parameters).

13. Incubate slide 90 sec on a 92°C heating block, then transfer to thermal cycler. See Critical Parameters for discussion of thermal cyclers.

14. Carry out PCR using the following amplification cycles:

30 cycles:	30 sec	94°C (denaturation)
	1 min	~45°C (annealing)
	1 min	72°C (extension)
Final step:	indefinitely	4°C (hold).

These times and temperatures will likely require optimization for the specific materials and thermal cycler being used. Furthermore, the annealing temperature should be optimized (see Strategic Planning discussion of optimization of annealing temperature). These particular incubation parameters work well with SK38 and SK39 primers for the HIV-1 gag sequence, amplified in an MJ Research PTC-100-60 or PTC-100-16MS thermal cycler.

Mount slides

15. Remove slide from thermal cycler and soak in 100% ethanol \geq 5 min to dissolve nail polish. Pry off coverslip using a razor blade or other fine blade. Scratch off any remaining nail polish so that fresh coverslips can be placed evenly in the hybridization/detection steps (Basic Protocol 2).

16. Incubate slide 1 min on a 92°C heating block, then soak 5 min in 2 \times SSC at room temperature.

The 92°C incubation helps immobilize the intracellular signals. The amplification protocol is now complete and the slide is ready for hybridization (see Basic Protocol 2). If not to be used immediately, slides can be stored 2 to 3 weeks at 4°C.

ALTERNATE PROTOCOL

ONE-STEP REVERSE TRANSCRIPTION AND AMPLIFICATION

This method uses a single, recombinant enzyme—*rTth* DNA polymerase—which performs both reverse transcription and DNA amplification at once in a single reaction, thus avoiding the need for two different buffer systems as required in Basic Protocol 1.

Additional Materials (also see Basic Protocol 1)

3 mM 4dNTP mix: 3 mM each dNTP in TE buffer, pH 7.5 (see APPENDIX 2 for TE buffer; store 4dNTP mix at -20°C)

10 mM MnCl₂

25 mM MgCl₂

10 \times *rTth* transcription buffer: 100 mM Tris·Cl, pH 8.3 (APPENDIX 2)/900 mM KCl

10 \times chelating buffer (see recipe)

1.7 mg/ml BSA

2.5 U/ μl *rTth* DNA polymerase (Perkin-Elmer or Life Technologies)

1. Pretreat slides with peroxidase, proteinase K, and DNase (see Basic Protocol 1, steps 1 to 6).
2. Make up the following single-step reaction cocktail (100 μ l total volume):
 - 0.5 μ l 100 μ M forward primer
 - 0.5 μ l 100 μ M reverse primer
 - 6 μ l 3 mM 4dNTP mix
 - 2 μ l 10 mM MnCl₂
 - 10 μ l 25 mM MgCl₂
 - 2 μ l 10 \times r*Tth* transcription buffer
 - 8 μ l 10 \times chelating buffer
 - 10 μ l 1.7 mg/ml BSA
 - 2 μ l 2.5 U/ml r*Tth*
 - 59 μ l DEPC-treated H₂O.
3. Carry out steps 11 to 13 of Basic Protocol 1, using the single-step reaction cocktail from step 2 in place of the ISPCR reaction cocktail.
4. Carry out the following thermal cycling program for reverse transcription:

1 cycle:	15 min	70°C
	3 min	92°C
	15 min	70°C
	3 min	92°C
	15 min	70°C
5. Carry out the following thermal cycling program for ISPCR:

29 cycles:	1 min	93°C	(denaturation)
	1 min	53°C	(annealing)
	1 min	72°C	(extension).
Final step:	indefinitely	4°C	(hold).
6. Carry out steps 15 and 16 of Basic Protocol 1.

HYBRIDIZATION AND DETECTION OF ISPCR-AMPLIFIED TARGET MATERIAL

Amplified target nucleic acid from the Basic Protocol 1 or the Alternate Protocol is detected using an adaptation of standard in situ hybridization procedures (also see *UNIT 14.3*, *UNIT 14.4* & *UNIT 14.7*). Hybridized specimens are detected using a method suitable for the particular labeled probe employed in the hybridization reaction and slides are subsequently counterstained and mounted for visualization and analysis. Alternative steps are provided for detection of ³³P-labeled probes (by emulsion autoradiography) and biotin- or digoxigenin-labeled probes (by peroxidase- or alkaline phosphatase-based methods). More detail concerning in situ hybridization and detection can be found in *UNITS 14.3* & *14.7*.

Materials

- 200 pM probe, ³³P-labeled (see Support Protocol 3) or biotin- or digoxigenin-labeled (*UNIT 3.18*)
- Deionized formamide
- 20 \times and 2 \times SSC (*APPENDIX 2*)
- 100 \times Denhardt solution (*APPENDIX 2*)
- 10 mg/ml sonicated salmon sperm DNA (*UNIT 2.10*; denature 10 min at 94°C before adding to hybridization mix)

**BASIC
PROTOCOL 2**

**In situ
Hybridization
and Immuno-
histochemistry**

14.8.11

10% (w/v) SDS
Slide containing ISPCR-amplified nucleic acids (see Basic Protocol 1 or see Alternate Protocol)
Diluted Kodak emulsion (UNIT 14.4)
Kodak D19 developer
Kodak Unifix fixer
2% Gills hematoxylin (Sigma)
PBS (APPENDIX 2)
Streptavidin-peroxidase conjugate working solution (see recipe)
AEC working solution (see recipe)
Blocking solution (see recipe)
Streptavidin-alkaline phosphatase conjugate working solution (see recipe)
100 mM Tris·Cl, pH 7.5 (APPENDIX 2)/150 mM NaCl
Alkaline phosphatase substrate buffer (see recipe)
75 mg/ml nitroblue tetrazolium (NBT) in 70% (v/v) dimethylformamide
50 mg/ml 5-bromo-4-chloro-3-indolyl phosphate (BCIP) in 100% dimethylformamide
Nuclear fast red stain (Sigma)
50%, 70%, 90%, and 100% (v/v) ethanol
50% (v/v) glycerol in PBS
Mounting medium, water-based (e.g., CrystalMount; Stephens Scientific or GelMount; Biomed) or organic-based (e.g., Permount; Fisher)
20 × 60-mm glass coverslips
95°C heating block accommodating glass slides
48°C incubator
Moist chamber (UNIT 14.2)
Light-tight slide box
Dessicant
Coplin jars or glass staining dishes

NOTE: All washes and incubations are performed at room temperature unless otherwise noted.

1. Prepare the following hybridization mix:
 - 2 μ l 200 pM probe (20 pM final)
 - 50 μ l deionized formamide (50% final)
 - 10 μ l 20 \times SSC (2 \times final)
 - 10 μ l 100 \times Denhardt solution (10 \times final)
 - 10 μ l 10 mg/ml sonicated salmon sperm DNA (1 mg/ml final)
 - 1 μ l 10% SDS (1% final)
 - 7 H₂O.
2. Add 10 μ l hybridization mix to each well of slide containing ISPCR-amplified nucleic acids. Cover wells with coverslips and heat slide 5 min on a 95°C heating block.
3. Incubate slides 2 to 4 hr at 48°C in a moist chamber.

The optimal hybridization temperature is a function of the T_m (melting temperature) of the probe, which must be calculated for each probe. See Strategic Planning, discussion of optimization of annealing temperature. However, the hybridization temperatures used should not be too high. If it is necessary to use a very high hybridization temperature, 40% formamide should be substituted for 50% formamide in the hybridization mix in step 1 (see UNIT 14.3, Commentary).

For ³³P-labeled probe

- 4a. Remove coverslips and wash slide in 2× SSC for 5 min.
- 5a. Dip slides in diluted Kodak emulsion.
- 6a. Air dry slides, then incubate 3 to 10 days in light-tight slide box with a dessicant.
- 7a. Develop slides in darkroom as follows:
 - 3 min in Kodak D19 developer
 - 30 sec rinse in H₂O
 - 3 min in Kodak Unifix fixer.
- 8a. Counterstain slide by incubating 2 to 3 min in 2% Gills hematoxylin in a Coplin jar at room temperature.

The slide is now ready for viewing; omit steps 9 to 11.

For peroxidase-based detection of biotin- or digoxigenin-labeled probe:

- 4b. Remove coverslips and wash slide twice in PBS, immersing for 5 min each time.
- 5b. Add 10 μl 100 μg/ml streptavidin-peroxidase to each well of slide. Gently cover wells with new coverslips and incubate 1 hr at 37°C.
- 6b. Remove coverslip and wash slides twice in PBS, immersing 5 min each time .
- 7b. In the dark, add 100 μl AEC working solution to each well of slide. Incubate 10 min at 37°C, then observe slides under microscope; if color is not strong, develop another 10 min.
- 8b. Rinse slide with tap water, then allow to dry.

Positive cells will appear brownish-red under the microscope.

For alkaline phosphatase-based detection of biotin- or digoxigenin-labeled probe

- 4c. Remove coverslips and wash slides twice in 2× SSC, immersing 15 min each time at room temperature, then cover surface of each well with 100 μl blocking solution. Place slide flat in moist chamber and incubate 15 min at room temperature.
- 5c. For each well to be developed, mix 10 μl of 40 μg/ml streptavidin-alkaline phosphatase conjugate with 90 μl conjugate dilution buffer.
- 6c. Remove blocking solution by touching absorbent paper to edge of slide. Cover surface of each well with 100 μl of the diluted conjugate solution prepared in step 5c. Place slide flat in moist chamber and incubate in 15 min at room temperature.

Do not allow the tissue sample to dry after adding the conjugate.

- 7c. Wash slides twice in 100 mM Tris-Cl, pH 7.5/150 mM NaCl, 15 min each time at room temperature, then wash once more in alkaline phosphatase substrate buffer, 5 min at room temperature.
- 8c. Prewarm 50 ml alkaline phosphatase substrate buffer to 37°C in a Coplin jar. Add 200 μl 75 mg/ml NBT and 166 μl of 50 mg/ml BCIP. Mix well, then incubate slides at 37°C in this solution at 37°C until desired level of signal is achieved (usually 10 min to 2 hr; check color development periodically by removing slide from solution and examining under 10× objective without allowing slide to dry), then stop reaction by rinsing slides in several changes of deionized water.

Be careful not to allow specimen to dry when removing slides to check color development.

Counterstain and mount slide (for nonisotopic probes only)

9. Stain slide for 5 min at room temperature in either 0.2% Gills hematoxylin (if peroxidase-based color development was used) or in 1% nuclear fast red stain (if alkaline phosphatase-based color development was used). Rinse slide in several changes of tap water.
10. Dehydrate slide by incubating 1 min at room temperature successively in 50%, 70%, 90%, and 100% ethanol. Air dry at room temperature.
11. Mount slide by applying one drop mounting medium per well and covering with a coverslip. View immediately, taking care not to disrupt the coverslip, or allow mounting medium to dry overnight at room temperature.

**SUPPORT
PROTOCOL 1**

PREPARATION OF AES-SUBBED SLIDES

Teflon-coated glass slides with wells are treated with 3-aminopropyltriethoxysilane (AES), rinsed extensively, and dried before use. The Teflon coating serves to form distinct wells, each of which serves as a small reaction chamber for in situ PCR (Basic Protocol 1) when covered with a coverslip and sealed with nail polish. The Teflon coating also helps to keep the nail polish from entering the reaction chamber. Use of slides with multiple wells allows for both a positive and negative control on the same slide. See Critical Parameters for further considerations regarding choice of slides.

Materials

Teflon-coated glass slides with three 10-, 12-, or 14-mm wells for cell suspensions or single oval well for tissue sections (Cel-Line Associates and Erie Scientific)
2% (v/v) 3-aminopropyltriethoxysilane (AES; Sigma) in acetone (prepare in Coplin jar or glass staining dish immediately before use)
DEPC-treated H₂O (UNIT 4.1)

Coplin jars or glass staining dishes
Vessel accommodating 1000 ml of liquid

1. Dip Teflon-coated glass slides in 2% AES for 5 min, then allow to dry 10 to 15 min at room temperature.
2. Dip slides for five min in a separate vessel containing 1000 ml of DEPC-treated water.
3. Repeat step 2 three times using three changes of DEPC-treated water, then air dry in laminar-flow hood overnight.

Slides should be used within 15 days of subbing.

**SUPPORT
PROTOCOL 2**

PREPARATION OF SPECIMENS ON SLIDES FOR ISPCR

Paraffin-Fixed Tissue

Routinely prepared paraffin-fixed tissue sections (UNIT 14.1) can be amplified quite successfully, permitting evaluation of individual cells in the tissue for the presence of a specific RNA or DNA sequence (Greer et al., 1991a,b). Specially designed slides that have single wells and that have been subbed with 3-aminopropyltriethoxysilane are used for this purpose (see Support Protocol 1). The authors routinely use sections of placental, CNS, cardiac, and similar tissues that have been sliced to a 3- to 5- μ m thickness. Other laboratories prefer to use sections up to 10- μ m thickness, but in the authors' experience, amplification is often less successful with thicker sections as multiple cell layers can often lead to difficult interpretation resulting from superposition of cells. However, if tissues that contain particularly large cells are being studied—e.g., ovarian follicles—thicker

sections may be appropriate and even desirable.

Plastic Sections

In situ amplification can be performed successfully on plastic sections (see Pereira et al., 1995 for plastic section technique). First, plastic-sectioned tissue must be deplasticized to remove methyl methacrylate by incubating in four successive baths of fresh methyl celuloacetate ether (MCA; Sigma) for 15 min each time, then in three successive baths of fresh acetone for 10 min each time, and then in multiple baths of xylene up to 4 hr each time. After this, the tissue is fixed in 4% paraformaldehyde (*UNIT 14.1*), and the in situ procedure is performed as with other tissue preparations (R. Boyce, pers. comm.).

Frozen Sections

It is possible to use frozen sections for in situ amplification, but the morphology of the tissue following the amplification process is generally not as good as with paraffin sections (see above). It seems that the freezing of the tissue, combined with the lack of paraffin substrate during slicing, compromises the integrity of the tissue. Usually, thicker slices must be made, and the tissue “chatters” in the microtome. As any clinical pathologist will relate, definitive diagnoses are made from paraffin sections, and this rule of thumb seems to apply to ISPCR as well. The exception to the rule is when immunohistochemical techniques are to be used for detecting additional signals in the cells (see Background Information) as certain immunohistochemical techniques require frozen sections.

To use frozen sections, the first step is to freeze the tissue properly (*UNIT 14.2*). For this purpose cut a 1 × 1–mm piece of styrofoam from a sheet (or a cup if a sheet is not available). Cut a piece of tissue about the same size (1 × 1–mm; 0.2- to 0.3-mm thickness), attach it to the styrofoam with ~2 to 4 ml of OCT solution (*UNIT 14.2*) and immerse the whole specimen attached to the styrofoam in liquid nitrogen in an insulated vessel. The tissue will freeze in few seconds. The frozen tissue can then be mounted in a cryostat and cut with a microtome (*UNIT 14.2*) or frozen at –70°C for future use. This method prevents the formation of icicles in the tissues and preserves the morphology much better than simply storing the tissue in a deep freeze. If liquid nitrogen is not available, the tissue (with styrofoam and OCT) may be wrapped in aluminum foil and placed on dry ice for 10 to 15 min before being stored in the deep freeze. As thin a slice as possible (3 to 4 μm) should be cut during sectioning. This is then applied to an AEC-subbed slide (see Support Protocol 1), dehydrated for 10 min in 100% methanol, and air dried in a laminar-flow hood.

Cell Cultures

ISPCR has been successfully performed by the authors on cells cultured on either 4 or 8-well Nunc slides (VWR Scientific) or AES-subbed Teflon ISPCR slides (see Support Protocol 1). If Nunc slides are used, 4- or 8-well slides with a glass base should be selected because plastic-base slides will melt during thermal cycling. This type of slide has a rubber gasket that should be left on, but the coverslip will not fit unless some protruding portions of the gasket are shaved off. Also, in the author’s laboratory, the stability of these gaskets is reinforced by applying a layer of nail polish at the junction of the gasket and the glass base. Aside from these modifications, the Nunc slides can be processed like the Teflon slides (they need not be AES-subbed, but can be subjected to thermal cycling in the PTC-100 thermal cycler). It should be noted that occasional leakage occurs in these slides. AES-subbed Teflon slides may also be used for tissue culture. They should be sterilized after silanization by incubating 30 min in 70% ethanol. Cells can be then placed on these slides and cultured overnight in a sterile humidified box. After this, the cells may be fixed as in *UNIT 14.1* and subjected to ISPCR.

LABELING OLIGONUCLEOTIDE PROBES USING ³³P

³³P-labeled probes are preferable for isotopic in situ hybridization (ISPCR) to avoid possible contamination of the thermal cycler with the more hazardous ³²P. Oligonucleotides are easily labeled with ³³P using polynucleotide kinase.

Materials

- 2 μM oligonucleotide probe (see Strategic Planning)
- 10× T4 polynucleotide kinase buffer (*UNIT 3.4*)
- 10 μCi/μl [γ -³³P]ATP (10 Ci/mmol; Amersham)
- 10 U/μl T4 polynucleotide kinase
- 0.8-ml Sephadex G-50 column (e.g., QuickSpin, Boehringer Mannheim)
- TE buffer, pH 7.4 (*APPENDIX 2*)

Additional reagents and equipment for labeling with T4 polynucleotide kinase (*UNITS 3.4 & 3.10*)

1. Make up the following reaction mix in a microcentrifuge tube (20 μl total volume):
 - 1.0 μl 2 μM oligonucleotide probe
 - 2 μl 10× polynucleotide kinase buffer
 - 1.0 μl 10 μCi/μl [γ -³³P]ATP
 - 15 μl H₂O
 - 1 μl 10 U/μl T4 polynucleotide kinase.

Incubate 30 min at 37°C.

2. Apply reaction mix to an 0.8-ml Sephadex G-50 column. Elute with TE buffer, and collect fractions at the following volumes:

Fraction 1	300 μl
Fraction 2	100 μl
Fraction 3	100 μl
Fraction 4	100 μl
Fraction 5	100 μl
Fraction 6	100 μl.

3. Count radioactivity in 1.0 μl of each fraction and identify those containing the labeled probe.

The labeled probe should be contained in fractions 2 through 4.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Acetate buffer, 50 mM (pH 5.0)

Add 74 ml of 0.2 M acetic acid (11.55 ml glacial acetic acid/liter) and 176 ml of 0.2 M sodium acetate (27.2 g sodium acetate trihydrate in 1 liter) to 1 liter of deionized water and mix.

AEC solution

Stock solution: Dissolve one tablet (20 mg) of 3-amino-9-ethylcarbazole (AEC; Sigma) in 2.5 ml *N,N* dimethylformamide. Store at 4°C in the dark up to 6 months.

Working solution: Add 250 μl stock solution to 5 ml of 50 mM acetate buffer, pH 5 (see recipe), then add 25 μl 30% H₂O₂. Make fresh before each use; keep solution in the dark.

Alkaline phosphatase substrate buffer

100 mM Tris·Cl, pH 9.5 (APPENDIX 2)
150 mM NaCl
50 mM MgCl₂
Store at 4°C up to 3 months

AMV/MoMuLV reaction buffer, 10×

100 mM Tris·Cl, pH 8.3 (APPENDIX 2)
500 mM KCl
15 mM MgCl₂
Prepare fresh

Blocking solution

100 mM Tris·Cl, pH 7.8 (APPENDIX 2)
150 mM NaCl
50 mg/ml BSA
0.2 mg/ml sodium azide
Prepare fresh

Chelating buffer, 10×

100 mM Tris·Cl, pH 8.3
1 M KCl
7.5 mM EGTA
0.5% Tween 20
50% (v/v) glycerol

Conjugate dilution buffer

100 mM Tris·Cl, pH 7.3 (APPENDIX 2)
150 mM MgCl₂
10 mg/ml BSA
0.2 mg/ml sodium azide
Prepare fresh

RNase-free DNase solution

40 mM Tris·Cl, pH 7.4 (APPENDIX 2)
6 mM MgCl₂
2 mM CaCl₂
1 U/μl RNase-free DNase (e.g., RQ1, Promega)
Prepare fresh

For use with cells that are particularly rich in RNase, add 1000 U/ml placental ribonuclease inhibitor (e.g., RNasin, Promega) and 1 mM DTT to the above solution.

Streptavidin–alkaline phosphatase conjugate solution

Stock solution: 40 μg/ml streptavidin–alkaline phosphatase conjugate (e.g., Life Technologies). Store at 4°C up to 8 months.

Working solution: For each well to be developed, dilute 10 μl stock solution with 90 μl conjugate dilution buffer (see recipe) immediately before use.

Streptavidin–peroxidase conjugate solution

Stock solution: 1 mg/ml streptavidin–horseradish peroxidase conjugate (e.g., Life Technologies) in PBS, pH 7.2 (APPENDIX 2). Store at 4°C.

Working solution: When stock solution is fresh (i.e., stored <1 month), dilute 1:100 in PBS, pH 7.2, to prepare working solution. Because the activity of the solution declines with prolonged storage, dilute 1:80 after 1 month of storage, 1:50 after 2 months of storage, and 1:30 after 3 months of storage. Discard after 4 months of storage.

COMMENTARY

Background Information

In situ hybridization, described in *UNIT 14.3* & *UNIT 14.7*, provides a critical and powerful means of determining the level of gene expression in individual cells and tissues, thereby allowing the investigator to assess and compare varying levels of gene expression within a population. A serious limitation of in situ hybridization, however, has been its inability to detect and quantitate target nucleic acid sequences present at levels below which the signal can be clearly distinguished from background. Use of the polymerase chain reaction (PCR) in conjunction with in situ hybridization (in situ PCR or ISPCR) has allowed specific amplification of previously undetectable sequences (Haase et al., 1990; Nuovo et al., 1991; Nuovo, 1994b; Bagasra et al., 1992, 1993a; Embretson et al., 1993a). Over the past several years, ISPCR has been used successfully by a number of research groups to identify the presence of targets that fall below the threshold of detection of traditional in situ hybridization. Thus, in situ amplification technology has been exploited to confirm the presence in specific cells of various subfamilies of retroviruses, including lentivirus (Haase et al., 1990) and HIV-1 (Bagasra et al., 1992; 1993b, Patterson et al., 1993, Nuovo et al., 1994). ISPCR has also been used to study elements involved in shifting an integrated virus from a latent to an infectious state (Chieu et al., 1992, Embretson et al., 1993a,b; Winslow et al., 1993; Bagasra and Pomerantz, 1993; Guatelli et al., 1994). Other applications of ISPCR include the detection of transcripts to facilitate the study of rearrangement of immunoglobulin heavy and light chain V-genes (Embleton et al., 1992), the variation in expression of cellular EGF receptor (Heniford et al., 1993; Patel et al., 1994), the expression of surfactant protein A and estrogen-receptor mRNA (Bibbo et al., in press), the expression of platelet factor XI mRNA (Hsu et al., 1994), and the localization of human and viral cDNAs (Nuovo et al., 1992; Mankowski et al., 1994).

Application of ISPCR in conjunction with electron microscopy (EM)

The authors' group at Thomas Jefferson University and another group at Uppsala University in Sweden have been developing techniques that make it possible to observe the results of ISPCR under an electron microscope. The procedure used by the authors is a simple modification of the immunogold EM

technique (Hacker et al., 1994). First, in situ amplification is carried out in solution rather than with the cells or tissue attached to a glass slide. The cells or tissue that have been subjected to ISPCR are then fixed in 4% paraformaldehyde for 4 hr (see *UNIT 14.1*), washed in PBS, and treated with proteinase K (see Basic Protocol 1). As in Basic Protocol 1, the cells are observed under an optical microscope for the development of "bubbles" or "peppery dots" (Fig. 14.8.3). Amplification of DNA or RNA is then carried out as described in Basic Protocol 1. To check the results, a small aliquot of the cells is sampled and placed on a slide, then in situ hybridization is performed with the specific biotinylated probe and an unrelated biotinylated probe as a negative control. Color is developed with streptavidin-peroxidase or streptavidin-alkaline phosphatase. If color develops with the specific probe but not the non-specific control, immunolabeling can be performed with the remaining cells. To do this, cells are first hybridized with biotinylated probe(s) and then labeled with streptavidin-immunogold conjugate. After 1 hr of incubation at 37°C, the cells are washed extensively to remove unbound conjugate, then pelleted by centrifugation. The cell pellet is incubated with freshly prepared 2.5% glutaraldehyde and processed for EM work. EM grids can also be used as the substrate for the cells or tissues. If used, the grid is placed between two slides and in situ amplification is performed as described above. Labeling with biotin/streptavidin-immunogold is then carried out. In the authors' laboratory, it has been found that the harsh treatment inflicted by repeated denaturation tends to destroy the internal organelles of cells. However, clear signals can be detected in the perinuclear areas.

Application of ISPCR in conjunction with immunohistochemistry

Immunohistochemistry to detect proteins (e.g., surface antigens) and in situ amplification to detect corresponding DNA sequences and RNA transcripts can be performed simultaneously in a single cell. In the authors' laboratory, cells or frozen sections of tissue, already attached to slides, are fixed with 100% methanol for 10 min and washed in PBS. Labeling of surface antigen(s) can then be carried out by standard immunohistochemical methods—i.e., the cells are incubated with unconjugated antibody for 1 hr at 37°C, washed, and fixed in

4% paraformaldehyde for 2 hr (see UNIT 14.6). To further facilitate these analyses, immunohistochemistry panels have been compiled tabulating specific surface antigens that will still bind specific FITC-conjugated monoclonal antibodies even after treatment with 10% formalin and similar reagents used in histopathology laboratories. It is possible, therefore, to use routinely prepared paraffin sections for the detection of such cellular antigens. In addition to the immunohistochemical detection of the antigens, in situ amplification is carried out as described in Basic Protocol 1. The amplified product is hybridized and detected as in Basic Protocol 2. At this point, secondary antibody conjugated with FITC, rhodamine, or another appropriate label may be applied. After a 1 hr incubation, the slides are washed and the cells or tissue are viewed under UV and visible light in an alternate fashion, to detect both the immunohistochemical and ISPCR/hybridization signals in a single cell.

Detection of multiple signals and multiple labels in individual cells

DNA, mRNA, and protein can all be detected simultaneously in individual cells. As described above, it is possible to label proteins with FITC-labeled antibodies, then perform in situ amplification of both RNA and DNA in the cells. If primers are used for spliced mRNA, which will not bind any sequences in DNA (see Strategic Planning), then both DNA amplification by ISPCR and RNA amplification by in situ reverse transcription/amplification can be carried out simultaneously. Subsequently, the products can be detected by hybridizing with differentially labeled probes that result in different-color signals (Komminoth et al., 1992). For example, proteins can show a rhodamine signal, mRNA can show a fluorescent FITC (20 colors of rhodamine-conjugated probe are available), and DNA can be labeled with a biotinylated probe and detected with peroxidase or phosphatase as in Basic Protocol 2. Each label will show a different signal within an individual cell.

Direct incorporation of nonisotopically labeled nucleotides into amplification products

Several nonisotopically labeled nucleotides are available from various sources—e.g., dCTP-biotin and digoxigenin-11-dUTP. By incorporating them into the ISPCR cocktail (see Basic Protocol 1) these nucleotides can be used to directly label amplification products—as an

alternative to the in situ hybridization steps (see Basic Protocol 2). The proper secondary reagents and chromogens can then be used to detect the directly labeled in situ amplification products. However, in the opinion of the authors—as well as in the opinion of several other investigators—the greatest specificity is achieved only by conducting amplification followed by subsequent in situ hybridization with the specific probe or probes. In the protocols used for labeling with nonisotopically labeled nucleotides, nonspecific incorporation can be significant, and even if this incorporation is minor, it still leads to false-positive signals similar to nonspecific bands that appear in gel electrophoresis following solution-based DNA amplification or RNA reverse transcription/amplification. Therefore, direct incorporation of labeled nucleotides as part of an in situ amplification protocol is strongly discouraged.

The only exception to this recommendation is when a large number of primer pairs is being screened for optimization of a specific assay. In that case, direct incorporation of labeled nucleotides may be useful. To perform such screenings, 4.3 μM labeled nucleotide—either biotin-14-dCTP, biotin-14-dATP, or digoxigenin-11-dUTP—is added along with the non-labeled 4dNTPs to achieve a 0.14 mM total nucleotide concentration.

Critical Parameters and Troubleshooting

Choice and treatment of slides

Choice and treatment of slides is critical to the success of the protocol (also see Support Protocol 1). Glass slides that are partially covered by a Teflon coating should always be used. Not only does glass withstand the stress of thermal cycling, but it also presents the right chemical surface—silicon oxide—needed for proper silanization. Slides with special Teflon coatings that form individual wells are useful because vapor-tight reaction chambers can be formed on the surface of the slides when coverslips are attached with coatings of nail polish around the periphery. These reaction chambers are necessary because within them, proper tonicity and ion concentrations can be maintained in aqueous solution during thermal cycling—conditions that are vital for proper amplification. The Teflon coating serves a dual purpose in this regard. First, the Teflon helps keep the glass surfaces of the slide and coverslip slightly separated, allowing reaction chambers ~20- μm in height to form in between. Second,

the Teflon border helps keep the nail polish from entering the reaction chambers when being applied. This is important because any leakage of nail polish into a reaction chamber can compromise results.

Normally, glass slides are silanized to make the slide surface hydrophobic. However, AES treatment (see Support Protocol 1) has a very different surface effect. This silicon compound imparts a strong, persistent, positive electrical charge to the slide by forming an aminopropyl derivative of glass at the surface of the slide. The resulting positive charge—and electrostatic attraction—causes the cells or tissues to adhere with great tenacity throughout the amplification/hybridization procedure. Experiments have been conducted using slides coated with alternative adhesives including white glue, albumin, chrome-gelatin, and poly-L-lysine. For slides prepared using alternative subbing techniques, however, the authors have found that slides treated with AES result in superior tissue adhesion and lower background with all tissue types tested, including mixed-cell suspensions, paraffin sections, and frozen sections. The authors hypothesize that this adhesion occurs because the negatively charged molecules on membrane surfaces bond ionically to the positively charged surface of the glass when the cells or tissues are initially placed on the slide. However, the limited positive charge is probably lost rather quickly, so later in the procedure, when solutions containing additional negatively-charged molecules are added, the new molecules float freely without becoming bound to the slide surface.

Optimizing proteinase K digestion

In the authors' laboratory, proper proteinase K digestion parameters (also see Basic Protocol 1) vary considerably with tissue type. Typically, lymphocytes require 5 to 10 min at 25°C or room temperature, CNS tissue requires 12 to 18 min at room temperature, and paraffin-fixed tissue requires between 15 to 30 min at room temperature. However, these time periods can vary widely and the appearance of the "salt and pepper dots" (see Fig. 14.8.3 and Fig. 14.8.4) is the important factor. Unfortunately, the appearance of the "salt and pepper dots" is less prominent in paraffin sections. Representative examples of "salt and pepper dots" on lymphocytes are illustrated in Figure 14.8.4. The critical importance of these dots should not be underestimated, as an extra 2 to 3 min of treatment after the appearance of dots will result in

leakage of signals. An alternative to the observation of "dots" is to select a constant time and treat slides with varying amounts of proteinase K—15 min using a range of from 1 to 6 µg/ml proteinase K.

The authors have also experimented with the use of other proteinases—e.g., amylase, trypsin, pronase, pepsinogen, and pepsin—instead of proteinase K. This has proven successful in many circumstances—e.g., incubation of slides for with 1 mg/ml pepsinogen for 10 min at pH 2.0. However, proteinase K has almost always been found to give better overall results.

Considerations for amplification

Attaching coverslip (or top slide). To ensure that a small, completely closed reaction chamber is formed to contain water vapor during thermal cycling, nail polish must be evenly distributed around the periphery of the coverslip or edges of the slide. For effective sealing, do not use colored nail polish or polish that is especially "runny"—the authors prefer to use "Wet & Wild" clear nail polish (Procter & Gamble). Proper sealing is very important in keeping reaction concentrations consistent through the thermal cycling procedure, which is critical to proper amplification. However, it is necessary to apply the nail polish very carefully so that none gets into the actual chamber where the cells or tissues reside. If any nail polish does enter the chamber, that slide should be discarded as the results will be questionable. Sealing the coverslip with nail polish is truly a learned skill; therefore, it is strongly recommended that researchers practice this procedure several times with mock slides before attempting an experiment. Where tissue sections are used for ISPCR, it is best to use another, identical blank slide to cover the specimen instead of a coverslip. This is done by applying the amplification cocktail to the appropriate well of the blank slide (a modification to Basic Protocol 1, step 11), placing the inverted tissue-containing slide (prepared as in Basic Protocol 1, steps 1 to 9) atop the blank slide, and sealing the edges with nail polish as described in Basic Protocol 1, step 12. The slide is then inverted once again so that the tissue-containing slide is on the bottom. A videotape illustrating this procedure is available from TAO Biomedical.

Hot-start PCR. There is much debate as to whether a "hot start" (see UNIT 15.1) helps to improve the specificity and sensitivity of ISPCR amplification reactions. In the authors'

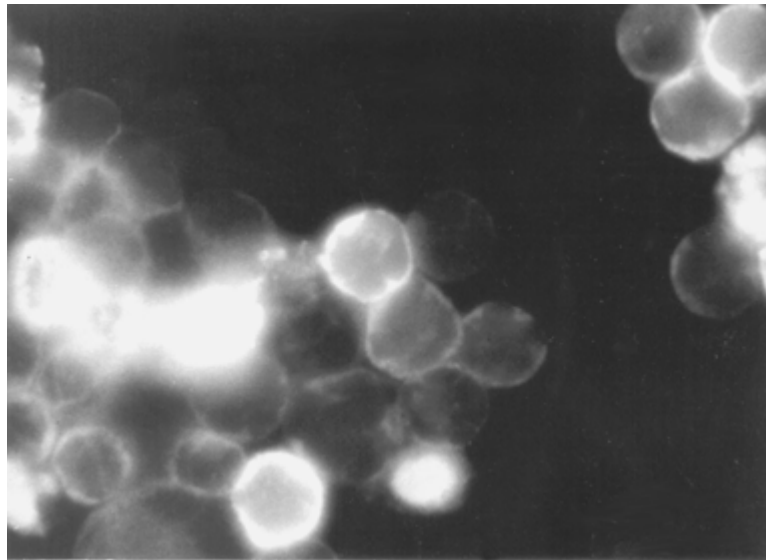


Figure 14.8.5 In situ reverse transcription/polymerase chain reaction using FITC-labeled (*tat-rev*) probes for HIV-1 expression in lymphocyte cell line chronically infected with HIV-1. Note cytoplasmic staining showing HIV-1 mRNA expression.

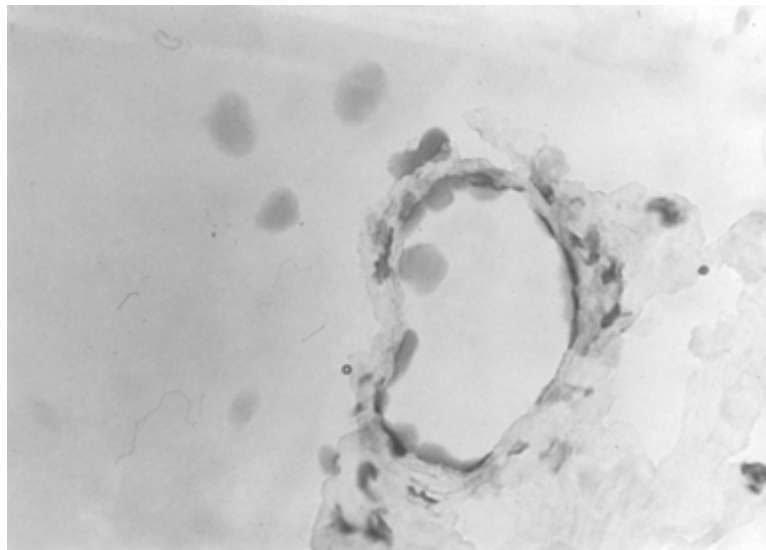


Figure 14.8.6 DNA-ISPCR of HIV-1-infected microvascular endothelial cell line. HIV-1 provirus gag sequence is amplified with an SK-38/39 primer pair and hybridized with a biotinylated SK-19 probe. Color was developed with AEC (see Basic Protocol 2) resulting in a red color.

laboratory this variation has been found to add no advantage; rather, it adds only technical difficulty to the practice of ISPCR. Recently, however, a variation of the hot-start method has been reported (Nuovo, 1994b) that essentially serves the same function without its difficulties. In this procedure, anti-*Taq* DNA polymerase antibody is used in the PCR cocktail to block *Taq* DNA polymerase activity until the first cycle of 94°C, at which temperature the antibody becomes denatured and full *Taq* DNA

polymerase activity is restored.

Thermal cyclers. Various types of thermal cyclers will work in this application; however, some work much better than others. The authors use two types: a standard, block-type thermal cycler that normally holds 60 0.5-ml tubes, but which can be adapted to hold 4 to 6 slides by use of aluminum foil, paper towels, and a weight. The authors also use dedicated thermocyclers that are specifically designed to hold 16 slides. Other laboratories have used

stirred-air, oven-type thermal cyclers quite successfully; however, problems with the cracking of glass slides during cycling have been reported. Thermal cyclers dedicated to glass slides are now available from several vendors, including Barnstead Thermolyne of Iowa, Coy Corporation of Minnesota, Hybaid of England, and MJ Research of Massachusetts. The authors' laboratory has used an MJ Research PTC-100-16MS quite successfully.

Several manufacturers of thermal cyclers, as well as suppliers of labeled nucleotides, have reported contamination problems associated with a chemical breakdown product from ^{35}S -labeled nucleotides. It is hypothesized that when these particular compounds are subjected to high temperatures, a gaseous, radioactive breakdown product forms—possibly $\text{H}_2[^{35}\text{S}]$ —which leaks from the sample vessel. This can lead to contamination of the incubating device and possibly the air in the laboratory. Therefore, for labeling probes, use of nucleotides labeled with the newer phosphorus isotope, ^{33}P , rather than ^{35}S is recommended. Biotin or digoxigenin labeling is another alternative.

It is suggested that the manufacturer's instructions on the use of a particular thermal cycler be followed, and that several additional points be kept in mind. First, glass does not easily make good thermal contact with the surface on which it rests. Therefore, a weight to press down the slides and/or a thin layer of mineral oil to fill in the space between the bottom of the slide and the heating surface will help thermal conduction. If mineral oil is used, the oil should be well-smearred over the bottom surface of the slide so that it is not merely floating on air bubbles beneath it. Second, the top surfaces of slides lose heat quite rapidly through radiation and convection; therefore, a thermal cycler should be selected that envelops the slide in an enclosed chamber (as in some dedicated instruments), or else the tops of the slides should be insulated in some manner. Insulation is particularly critical when using a weight on top of the slides because the weight can serve as an unwanted heat sink if it is in direct contact with the slides. Third, good thermal uniformity is imperative for good results—poor uniformity or irregular thermal change can result in cracked slides, uneven amplification, or completely failed reactions. If adapting a thermocycler that normally holds plastic tubes, a layer of aluminum foil should be used to spread out the heat.

In the authors' laboratory it has been ob-

served that most dedicated thermal cyclers do not transfer heat in an even manner to all parts of the thermal cycler block. The authors therefore recommend use of either the PTC-100-16MS or the newly designed "DNA Engine" (PTC-200 with slide block), both from MJ Research.

Anticipated Results

Amplification products can be detected and identified in the cytoplasm and in nuclear regions using fluorescent, colorimetric, or isotopic probes. Figure 14.8.5 shows results of *in situ* reverse transcription and amplification using FITC-labeled probes to detect HIV-1 expression in lymphocytes. Figure 14.8.6 confirms the presence of HIV-1 sequences in microvascular epithelial cells (MVEC) as determined by DNA ISPCR.

Time Considerations

The time required for preparing slides of tissue samples is as described in *UNITS 14.1 & 14.2*. Basic Protocol 1 requires ~8 hr of laboratory work in addition to the time required for thermal cycling, which is performed overnight. After amplification and hybridization, detection and visualization of the products requires ~1.5 hr of laboratory work. More time, up to 12 hr, is required for detection of a radiolabeled probe.

Hybridization (Basic Protocol 2) takes 4 to 18 hr. AES-subbing of slides and preparation of probes (Support Protocols 1 and 3) should take 1 hr each.

Literature Cited

- Bagasra, O. and Pomerantz, R.J. 1993. HIV-1 provirus is demonstrated in peripheral blood monocytes *in vivo*: A study utilizing an *in situ* PCR. *AIDS Res. Hum. Retroviruses* 9:69-76.
- Bagasra, O., Hauptman, S.P., Lischner, H.W., Sachs, M., and Pomerantz, R.J. 1992. Detection of HIV-1 provirus in mononuclear cells by *in situ* PCR. *New Engl. J. Med.* 326:1385-1391.
- Bagasra, O., Seshamma, T., and Pomerantz, R.J. 1993a. Polymerase chain reaction *in situ*: Intracellular amplification and detection of HIV-1 proviral DNA and other specific genes. *J. Immunol. Methods* 158:131-145.
- Bagasra, O., Seshamma, T., Oakes, J., and Pomerantz, R.J. 1993b. High Percentages of CD4-positive lymphocytes harbor the HIV-1 provirus in the blood of certain infected individuals. *AIDS* 7:1419-1425.
- Barnes, W. 1994. PCR Amplification of up to 35-kb DNA with high fidelity and high yield from [λ] bacteriophage templates. *Proc. Natl. Acad. Sci. U.S.A.* 91:2216-2220.

- Bibbo, M., Pestanev, J.P., Scavo, L.M., Bobroski, L., Seshamma, T., and Bagasra, O. 1995. Case report: Surfactant protein A mRNA expression utilizing the reverse transcription *in situ* polymerase chain reaction for metastatic adenocarcinoma. *Cell Vision*. In press.
- Chieu, K.-P., Cohen, S.H., Morris, D.W., and Jordan, G.W. 1992. Intracellular amplification of proviral DNA in tissue sections using the polymerase chain reaction. *J. Histochem. Cytochem.* 40:333-341.
- Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K., and Mattick, J.S. 1991. Touchdown PCR to circumvent spurious priming during gene amplification. *Nucl. Acids Res.* 19:4008-10.
- Embleton, M.J., Gorochov, G., Jones, P.T., and Winter, G. 1992. In-cell PCR from mRNA amplifying and linking the rearranged immunoglobulin heavy and light chain V-genes within single cells. *Nucl. Acids Res.* 20:3831-3837.
- Embretson, J., Zupanic, M., Beneke, T., Till, M., Wolinsky, S., Ribas, J.L., Burke, A., and Haase, A.T. 1993a. Analysis of human immunodeficiency virus-infected tissues by amplification and *in situ* hybridization reveals latent and permissive infections at single-cell resolution. *Proc. Natl. Acad. Sci. U.S.A.* 90:357-361.
- Embretson, J., Zupancic, M., Ribas, J.L., Burke, A., Racz, P., Tenner-Racz, K., and Haase, A.T. 1993b. Massive covert infection of helper T lymphocytes and macrophages by HIV during the incubation period of AIDS. *Nature* 62:359-362.
- Greer, C.E., Lund, J.K., and Manos, M.M. 1991a. PCR amplification from paraffin-embedded tissues: Recommendations on fixatives for long-term storage and prospective studies. *PCR Meth. Appl.* 95:117-124.
- Greer, C.E., Peterson, S.L., Kiviat, N.B., and Manos, M.M. 1991b. PCR amplification from paraffin-embedded tissues: Effects of fixative and fixative times. *Am. J. Clin. Pathol.* 95:117-124.
- Guatelli, J.C., Whitfield, K.M., Kwoh, D.Y., Barringer, K.J., Richman, D.D., and Gingeras, T.R. 1994. Isothermal, *in vitro* amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. *Proc. Natl. Acad. Sci. U.S.A.* ~7:1874-1878.
- Haase, A.T., Retzel, E.F., and Staskus, K.A. 1990. Amplification and detection of lentiviral DNA inside cells. *Proc. Natl. Acad. Sci. U.S.A.* 37:4971-4975.
- Hacker, G.W., Zehbe, L., Hauser-Kronberger, C., Gu, L., Graf, A.-H., and Dietze, O. 1994. *In situ* detection of DNA and mRNA sequences by immunogold-silver staining (IGSS). *Cell Vision* 1:30-37.
- Heniford, B.W., Shum-Siu, A., Leonberger, M., and Hendler, F.J. 1993. Variation in cellular EGF receptor mRNA expression demonstrated by *in situ* reverse transcription polymerase chain reaction. *Nucl. Acids Res.* 21:3159-66.
- Hsu, T.C., Bagasra, O., Seshamma, T., and Walsh, P.N. 1994. Platelet factor XI mRNA amplified from human platelets by reverse transcriptase polymerase chain reaction and detected by *in situ* amplification and hybridization. *FASEB J.* 8:1375.
- Komminoth, P., Merk, F.B., Leav, L., Wolfe, H.J., and Roth, J. 1992. Comparison of 32S and digoxigenin-labeled RNA and oligonucleotide probes for *in situ* hybridization expression of mRNA of the seminal vesicle secretion protein 11 and androgen receptor genes in the rat prostate. *Histochemistry* 93:217-228.
- Lundberg, K.S., Shoemaker, D.D., Adams, M.W., Short, J.M., Sorge, J.A., and Mathur, E.J. 1991. High fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* 1:81-6.
- Mankowski, J.L., Spelman, J.P., Reesetar, H.G., Strandberg, J.D., Latterra, J., Carter, D.L., Clements, J.E., and Zink, M.C. 1994. Neurovirulent simian immunodeficiency virus replicates productively in endothelial cells of the central nervous system *in vivo* and *in vitro*. *J. Virol.* 68:8202-8208.
- Nuovo, G.J. 1994a. PCR in Situ Hybridization Protocols and Applications, 2nd ed. Raven Press, New York.
- Nuovo, G.J. 1994b. Questioning *in situ* PCR. *In situ* cDNA polymerase chain reaction: A novel technique for detecting mRNA expression. *Am. J. Path.* 145:741.
- Nuovo, G.J., Gallery, F., MacConnell, P., Becker, J., and Bloch, W. 1991. An improved technique for the *in situ* detection of DNA after polymerase chain reaction amplification. *Am. J. Path.* 139:1239-1244.
- Nuovo, G.J., Gorgone, G.A., MacConnell, P., Margiotta, M., and Gorevic, P.D. 1992. *In situ* localization of PCR-amplified human and viral cDNA. *PCR Meth. Appl.* 2:117-223.
- Nuovo, G.J., Gallery, F., MacConnell, P., and Braun, A. 1994. *In situ* detection of polymerase chain reaction-amplified HIV-1 nucleic acids and tumor necrosis factor- α RNA in the central nervous system. *Am. J. Path.* 144:659-666.
- O'Leary, J.J., Browne, G., Landers, R.J., Crowley, M., Healy, I.B., Street, J.T., Pollock, A.M., Murphy, J., Johnson, M.I., and Lewis, F.A. et al. 1994. The importance of fixation procedures on DNA template and its suitability for solution-phase polymerase chain reaction and PCR *in situ* hybridization. *Histochem. J.* 26:337-346.
- Pallansch, L., Beswick, H., Talian, J., Zelenka, P. 1990. Use of an RNA folding algorithm to choose regions for amplification by the polymerase chain reaction. *Anal. Biochem.* 185:57-62.
- Patel, V.G., Shum-Siu, A., Heniford, B.W., Wieman, T.J., and Hendler, F.J. 1994. Detection of epidermal growth factor receptor mRNA in tissue sections from biopsy specimens using *in situ* polymerase chain reaction. *Am. J. Pathol.* 144:7-14.

Patterson, B.K., Till, M., Otto, P., and Goolsby, C. et al. 1993. Detection of HIV-I DNA and messenger RNA in individual cells by PCR-driven in situ hybridization and flow cytometry. *Science* 260:976-979.

Pereira, R.F., Halford, K.W., O'Hara, D.B., Leeper, B.P., Sokolov, M.D., Bagasra, O., and Prockop, D.J. 1995. Cultured stromal cells from marrow serve as stem cells for bone, lung, and cartilage in irradiated mice. *Proc. Natl. Acad. Sci. U.S.A.* In press.

Winslow, B.J., Pomerantz, R.J., Bagasra, O., and Trono, D. 1993. HIV-1 Latency due to the site of proviral integration. *Virology* 196:849-854.

Key References

Bagasra et al., 1995. See above.

Haase et al., 1990. See above.

Nuovo et al., 1991. See above.

Describe the basic in situ amplification methods and discuss its applications.

Contributed by Omar Bagasra,
Thikkavarapu Seshamma, and
Roger Pomerantz
Thomas Jefferson University
Philadelphia, Pennsylvania

John Hanson
MJ Research
Watertown, Massachusetts

Whole-Mount In Situ Hybridization and Detection of RNAs in Vertebrate Embryos and Isolated Organs

The spatial and temporal distribution of transcripts (e.g., during embryogenesis) provides important clues to possible functions of the encoded gene products and/or possible interactions with other genes.

The whole-mount in situ hybridization method, first introduced by Tautz and Pfeifle (1989) and applied to *Drosophila* embryos, has been extended to vertebrate embryos: *Xenopus* (Harland, 1991; Franco et al., 1999), zebrafish (Schulte-Merker, 1993), and mouse and chicken (Wilkinson, 1993). Given the complex anatomy of embryos, the method simplifies detection and interpretation of gene expression patterns, especially for genes with graded or dynamic distribution. This method circumvents the tedious detection of RNA hybrids by emulsion autoradiography (UNIT 14.4) and subsequent reconstruction of their three-dimensional distribution.

Whole-mount in situ hybridization using digoxigenin-labeled riboprobes is probably the most commonly used method to detect RNA transcripts in intact embryos or isolated organs (see Basic Protocol 1 and Alternate Protocol 1). The hybrids are subsequently detected with anti-digoxigenin antisera conjugated to alkaline phosphatase and immune complexes are visualized by an enzymatic reaction of alkaline phosphatase with a chromogenic substrate (see Basic Protocols 2 and 4 and Figs. 14.9.1, 14.9.2, and 14.9.3).

The methods described in this unit are based on procedures by Riddle et al. (1993) and Wilkinson (1993). They have been used successfully for the detection of mRNAs in mouse and chicken embryos. The methods are also applicable to experimentally manipulated embryos—e.g., for detection of transcripts in mouse tissue grafted to chicken embryos (Haramis et al., 1995).

Whole-mount in situ hybridization of *Xenopus* embryos requires different pretreatment, hybridization, and detection conditions because these embryos are yolk-rich and prone to high background (see Basic Protocols 3 and 4). Both hybridizations use digoxigenin-labeled antisense RNA probes (riboprobes; see Support Protocol 2), and hybridizations with mouse and chicken embryos use anti-digoxigenin antibody that will be preabsorbed with embryonic powder to reduce nonspecific binding of the detecting antibody to embryonic tissue (see Support Protocol 1).

CAUTION: Paraformaldehyde, formaldehyde, methanol, hydrogen peroxide, glutaraldehyde, formamide, NBT, BCIP, and DEPC are hazardous. Please consult manufacturer's guidelines for their proper handling and disposal.

NOTE: All solutions should be treated with diethylpyrocarbonate (DEPC, see UNIT 4.1) and autoclaved to inhibit RNase activity.

NOTE: The samples should always be kept covered with liquid to avoid drying, which causes high levels of background staining. Therefore, when exchanging solutions, always leave the sample covered by a small quantity of liquid. This also prevents physical damage or loss of samples.

WHOLE-MOUNT IN SITU HYBRIDIZATION WITH MOUSE OR CHICKEN EMBRYOS AND ORGANS

Mouse or chicken embryos or organs are isolated and fixed to preserve their morphology, then pretreated to make the target transcripts accessible to the riboprobe. The pretreatments and prehybridization block nonspecific binding sites, thus reducing nonspecific hybridization of riboprobes and background staining during enzymatic detection of RNA hybrids. Finally, the sample is hybridized to digoxigenin-labeled riboprobe (see Support Protocol 1) and visualized using an enzyme-conjugated antibody to digoxigenin.

Materials

Mouse or chicken embryos or organs
PBS (APPENDIX 2), ice-cold
4% (w/v) paraformaldehyde in PBS (4% PFA; UNIT 14.1), 4°C
PBT (also see recipe): 0.1% or 1% (v/v) Tween 20 in PBS, 4°C and room temperature
25%, 50%, and 75% methanol in PBT (methanol/PBT)
100% methanol
6% (v/v) hydrogen peroxide (H₂O₂; Aldrich) in PBT (optional)
10 µg/ml proteinase K in PBT, not predigested
2 mg/ml glycine (Merck) in PBT, freshly prepared
0.2% (v/v) glutaraldehyde (EM grade, Sigma)/4% (w/v) paraformaldehyde in PBT, freshly prepared
Hybridization solution (see recipe)
Hybridization solution containing 1 µg/ml digoxigenin-labeled riboprobe previously denatured (see Support Protocol 1)
Dissecting tools (e.g., scissors and forceps), wiped with 70% ethanol
Dissecting microscope
6-well culture plate with polycarbonate membrane inserts
1.5- to 2-ml microcentrifuge tubes
Heating block with adapter to hold 2-ml microcentrifuge tubes
Rocker platform

NOTE: All washes are carried out in 10 ml (6-well culture plates with polycarbonate membrane insert) or 1.5 to 2 ml (2-ml microcentrifuge tubes) of solution at room temperature with gentle agitation, unless stated otherwise. Solutions should be warmed to the indicated temperature before use.

Fix sample

1. Dissect mouse or chicken embryo in ice-cold PBS using appropriate dissecting tools and a dissecting microscope, if necessary. Completely remove the extraembryonic membranes and the amnion. Open cavities.

Riboprobes and antibodies that are trapped in embryonic cavities are one of the major causes of nonspecific signals. Therefore, it is important to open (i.e., pierce) cavities such as heart chambers and brain ventricles before or after fixation; this is particularly necessary for mouse embryos older than embryonic day 10 or chicken embryos older than embryonic stage 20. For advanced embryos (i.e., mouse embryos from embryonic day 14 or later) or postnatal stages, it is necessary to isolate the organs or tissues of interest to achieve best results.

Extraembryonic membranes can be stored up to 6 months at -80°C and used for preparation of genomic DNA (UNIT 2.2). DNA is used to genotype embryos by PCR (UNIT 15.1) or Southern blot hybridization (UNITS 2.9 & 2.10).

2. Transfer the sample into a 6-well culture plate containing ~10 ml of cold 4% PFA. Fix 4 hr to overnight at 4°C.

If desired, it is possible to fix several samples in one well of a 6-well culture plate with an excess of fixative. In general, fixation time (from 4 hr to overnight) is not critical and does not affect signal strength.

The 4% PFA can be prepared (UNIT 14.1) and stored frozen in aliquots at -20°C. Each aliquot should be thawed and used once.

3. Wash sample two times in PBT at 4°C.

For mouse embryos, it is recommended that all solutions contain 1% Tween 20 rather than 0.1%. Tween 20 is routinely used at 0.1% for the chick protocol.

It is advisable to perform the dehydration, rehydration, and bleaching steps (steps 4 to 8) even if the embryos are to be used immediately and not stored. The authors have found that these steps improve permeabilization of the tissues and increase the sensitivity of the method. Alternatively, the embryos can be stored in 1% PFA at 4°C for several weeks, and processed directly from step 8 onwards.

4. Dehydrate sample by successive washes in 25%, 50%, and 75% methanol/PBT at room temperature for 5 min each. Finally, wash two times in 100% methanol.

The samples can be stored in 100% methanol up to 6 months at -20°C. Alternatively, the samples can be processed to step 14 and then stored at -20°C.

Pretreat sample

5. Rehydrate sample using a reverse order of the methanol/PBT series from step 4 and wash two times in PBT at room temperature for 5 min.

6. Bleach sample with 6% hydrogen peroxide in PBT for 1 hr at room temperature.

See Critical Parameters and Troubleshooting for further discussion.

7. Wash sample three times in PBT for 5 min each.

Treat with proteinase K

Proteinase K digestion removes proteins, unmasks the target RNA, and helps to permeabilize the tissues, allowing the riboprobe to access its target mRNAs during hybridization (step 15). Two parameters can be manipulated: the concentration of proteinase K and/or the length of treatment. Use the following information as a guideline.

- 8a. *For chick embryos:* Digestion conditions indicate stage, concentration of proteinase K, and incubation time:

<10	1 to 3 µg/ml	15 min
10 to 18	10 µg/ml	15 min
18 to 24	10 µg/ml	20 to 25 min
26 to 29	10 µg/ml	up to 40 min

- 8b. *For embryos at stage 26 and older:* Increase the concentration of proteinase K while keeping the incubation time constant at 15 min:

26 to 27	20 µg/ml
28 to 29	30 µg/ml
30 to 31	40 µg/ml
32 and older	50 µg/ml

These conditions are an alternative treatment to step 8a. They may be used for late-stage embryos that are very difficult to permeabilize. A higher proteinase K concentration reduces the incubation time, which at these stages, should be >1 hr.

8c. *For mouse embryos:* Incubate in 10 µg/ml proteinase K. Modify the incubation time as embryo age in day post-coitum (dpc) increases, as follows:

6.5 dpc	4 min
7.5 dpc	4 to 5 min
8.5 dpc	6 min
9.5 dpc	10 min
10.5 dpc	15 min
11.5 to 12.5 dpc	20 min
13.5 to 14.5 dpc	30 min

9. Stop digestion by washing 10 min in freshly prepared 2 mg/ml glycine in PBT.

10. Wash two times for 5 min each in PBT.

Postfix and hybridize sample

11. Postfix sample with freshly prepared 0.2% glutaraldehyde/4% paraformaldehyde in PBT for 20 min at room temperature.

It is important that the 4% PFA used for this step is freshly prepared. Inclusion of 0.2% glutaraldehyde in the fixative is optional but seems to enhance the signal-to-noise ratio.

12. Wash three times in PBT for 5 min each.

For mouse embryos, a slow equilibration in hybridization solution is suggested. This approach has not been routinely used for chick embryos, but it could probably be used successfully.

13. Transfer the sample to a 1.5- to 2-ml microcentrifuge tube containing 1.5 ml of a 1:1 mixture of hybridization solution/PBT. Wash for 10 min at room temperature.

A small spoon or spatula can be used to transfer embryos.

Depending on sample size, more than one sample can be hybridized in a tube provided that the samples are completely submerged in solution. This facilitates direct comparative analysis of different samples.

14. Wash 10 min in hybridization solution.

The samples can be stored up to 6 months in hybridization solution at -20°C either before or after completing this step.

15. Incubate at 70°C in prewarmed hybridization solution for ≥1 to 3 hr.

16. Replace the hybridization solution with fresh hybridization solution containing 1 µg/ml digoxigenin-labeled riboprobe (previously denatured). Hybridize overnight at 70°C.

See Critical Parameters and Troubleshooting for suggestions on recommended hybridization temperatures.

Denatured probe in hybridization solution can be stored at -20°C and reused up to three times. Because the probe in the hybridization solution can be reused, the authors recommend that the samples be completely submerged in the microcentrifuge tubes to avoid any drying of the samples during the hybridization step.

BASIC PROTOCOL 2

ISH and Detection of RNAs in Vertebrate Embryos

14.9.4

ENZYMATIC DETECTION OF RNA HYBRIDS IN MOUSE AND CHICKEN EMBRYOS AND ORGANS

The antibody detection reaction can take 30 min for abundant messages, 3 to 4 hr for less abundant messages, or overnight to several days for rare messages. The reaction can be carried out at room temperature, in a 37°C warm room, or at 4°C. Each temperature will speed up (warm) or slow down (cold) the reaction as desired. To avoid background, do not overstain embryos. This protocol is preferred over Alternate Protocol 2 because it

contains fewer steps, which decreases the probability of damaging the samples. This protocol has been successfully used for most of the probes. Basic Protocol 2 may be used with Alternate Protocol 1.

Materials

Hybridized samples (see Basic Protocol 1)
Solution I (see recipe), 70°C
Solution II (see recipe), 65°C
TBST (also see recipe): 0.1% (chick) or 1% (mouse) (v/v) Tween 20 in TBS (see recipe for TBS)
Blocking solution for chick: 10% heat-inactivated sheep serum (see recipe) in TBST
Blocking solution for mouse: 10% heat-inactivated sheep serum (see recipe), 0.1% blocking reagent (see recipe) in TBST
Alkaline phosphatase–conjugated Fab fragments of anti-digoxigenin antibodies (Boehringer Mannheim), 4°C
Alkaline phosphatase buffer, pH 9.5 (NTMT; see recipe), freshly prepared
Reaction mix: 125 µg/ml BCIP and 250 µg/ml NBT in NTMT, pH 9.5
PBT buffer: 0.1% (chick) or 1% (mouse) Tween 20 in PBS
4% (w/v) paraformaldehyde in PBS (4% PFA; UNIT 14.1) with 0.1% glutaraldehyde
PBS (APPENDIX 2)
20%, 50%, and 80% glycerol/PBS solution
Rocker platform
Aluminum foil
6-well culture plate with polycarbonate membrane inserts

Perform post-hybridization washes, blocking, and antibody incubation

1. Hybridize samples (see Basic Protocol 1). Remove and keep (recycle) the hybridization solution containing the riboprobe. Prewarm solution I to 70°C and solution II to 65°C.

This riboprobe solution can be used up to three times and is stable for up to 2 years if kept at –20°C.

2. Wash embryos three times for 30 min each at 70°C with prewarmed solution I.
3. Wash embryos three times for 30 min each at 65°C with prewarmed solution II.
4. Wash three times with fresh TBST for 5 min each at room temperature.
5. For chick embryos, block with 10% heat-inactivated sheep serum in TBST for at least 1 hr at room temperature. For mouse embryos, block embryos by incubating at room temperature for 60 to 90 min in 10% heat-inactivated sheep serum and 0.1% blocking reagent in TBST.

Blocking solution for mouse takes a long time to solubilize. Prepare it early in the morning and heat at 37°C to help the dissolution.

6. Remove blocking solution from samples and add diluted (1:2000 to 1:10,000) alkaline phosphatase–conjugated Fab fragments of anti-digoxigenin antibody in 1% sheep serum/TBST solution. Incubate overnight at 4°C.
7. Remove and keep (recycle) the TBST solution containing the antibody.

Keep the antibody/TBST solution at 4°C, this antibody solution can be used up to three times.

8. Perform post-antibody washes by washing three times, 10 min each, in TBST at room temperature, then wash five times, 1 to 1.5 hr each, in TBST at room temperature. Finally, wash overnight in TBST at 4°C with gentle rocking on a platform rocker.

The purpose of the washes is to remove excess and nonspecifically bound antibody from the tissues to prevent nonspecific reaction with chromogenic substrate. For this step the samples can be transferred into a 6-well culture plate with polycarbonate membrane inserts. This will increase the volume of solution used for the washes and allow for better results.

The second group of five washes can be done over 2 days at 4°C—e.g., one wash in the morning and one wash at night each day with ice-cold TBST. The authors have found that longer washes result in less background during development.

Detect antibody

9. Wash hybridized samples three times for at least 10 min each (up to 1 hr) in 10 ml NTMT, pH 9.5.
10. Remove NTMT and add reaction mix. Allow the reaction to develop at the chosen temperature with gentle rocking. Throughout the development reaction, keep samples covered with aluminum foil to protect from light and periodically check to monitor progress of the reaction.

Do not hesitate to change the staining solution when it becomes purple.

11. When the reaction is judged to be complete, wash at least three times, 10 min each, in PBT buffer.

The reaction is judged to be complete when the appearance of a specific pattern of staining is observed.

An optional additional wash with TBST can be performed overnight (protected from light) at 4°C, if necessary, to reduce background.

12. Post-fix the samples in 10 ml of 4% PFA in PBS with 0.1% glutaraldehyde solution for at least 1 hr at room temperature (the reaction can go overnight at 4°C, if necessary).
13. Wash briefly 5 min in 10 ml PBS at room temperature.
14. (Optional) To clarify samples to clearly distinguish the specific staining from the background, incubate in increasing concentrations of glycerol solutions at room temperature as follows:

20% glycerol/PBS solution for at least 15 min

50% glycerol/PBS solution for at least 15 min

80% glycerol/PBS solution for at least 15 min

All the unstained tissue will become clear, allowing easier detection of the specific staining. The samples can now be analyzed prior to storage at 4°C.

ALTERNATE PROTOCOL 1

WHOLE-MOUNT IN SITU HYBRIDIZATION WITH MOUSE OR CHICKEN EMBRYOS AND ORGANS

This protocol is used for very high background probes and for the detection of very rare transcripts by modification of the hybridization and washing conditions.

Materials

Mouse or chicken embryos or organs

PBS (APPENDIX 2), ice cold

4% (w/v) paraformaldehyde in PBS (4% PFA; UNIT 14.1), 4°C

PBT (also see recipe): 0.1% (v/v) Tween 20 in PBS, 4°C and room temperature

25%, 50%, and 75% methanol in PBT (methanol/PBT)

100% methanol
6% (v/v) hydrogen peroxide (H₂O₂; Aldrich) in PBT (optional)
10 µg/ml proteinase K in PBT, not predigested
2 mg/ml glycine (Merck) in PBT, freshly prepared
0.2% (v/v) glutaraldehyde (EM grade, Sigma)/4% (w/v) paraformaldehyde in PBT, freshly prepared
Prehybridization solution A (see recipe), 65°C
Hybridization solution A: prehybridization solution A (see recipe) containing 1 µg/ml digoxigenin-labeled riboprobe (see Support Protocol 1)

Dissecting tools (e.g., scissors and forceps), wiped with 70% ethanol
Dissecting microscope
20-ml snap-cap glass vials
Heating block with adapter to hold 2-ml microcentrifuge tubes
Platform rocker

NOTE: All washes are carried out in 5 to 10 ml (20-ml glass vials) or 1.5 to 2 ml (2-ml microcentrifuge tubes) of solution for 5 min at room temperature with gentle agitation, unless stated otherwise. Solutions should be warmed to the indicated temperature before use.

Fix sample

1. Dissect mouse or chicken embryo in ice-cold PBS using appropriate dissecting tools and a dissecting microscope, if necessary. Completely remove the extraembryonic membranes and the amnion and open cavities.

See Basic Protocol 1, step 1 annotation, for information about the dissection.

2. Transfer sample into a 20-ml snap-cap glass vial containing ~10 ml of cold 4% PFA. Fix 4 hr to overnight at 4°C.

If desired, it is possible to fix several samples in one glass vial with an excess of fixative.

3. Wash sample two times in PBT at 4°C.
4. Dehydrate sample by successive washes, 5 min each wash, in 25%, 50%, and 75% methanol/PBT at room temperature. Finally, wash two times in 100% methanol.
5. Rehydrate sample using the methanol/PBT series from step 4 in reverse order and wash two times in PBT at room temperature.

Pretreat sample

6. (Optional) Bleach sample with 6% hydrogen peroxide in PBT for 15 min at room temperature.

See Critical Parameters and Troubleshooting for further discussion.

7. Wash the sample three times in PBT at room temperature.
8. Digest sample with 10 µg/ml proteinase K in PBT for 15 min at room temperature.

The samples are very fragile during this step.

The length of treatment and/or concentration of proteinase K needs to be optimized. For example, if the distribution of RNAs in embryonic ectodermal structures is studied, proteinase K digestion must be reduced to 4 µg/ml for 5 min or 10 µg/ml for 2 min to keep the ectoderm intact (Crossley and Martin, 1995). Alternatively, detergent permeabilization (Rosen and Beddington, 1993) can be substituted for proteinase K digestion.

9. Stop digestion by washing in freshly prepared 2 mg/ml glycine in PBT.
10. Wash two times in PBT.

Postfix and hybridize sample

11. Postfix sample with freshly prepared 0.2% glutaraldehyde/4% paraformaldehyde in PBT for 20 min at room temperature.
12. Wash three times in PBT.
13. Add 1 ml prewarmed (65°C) prehybridization solution A to the glass vial and transfer the sample to a 2-ml microcentrifuge tube.

A small spoon or spatula can be used to transfer embryos. Alternatively, the sample with prehybridization solution A can be carefully poured into the microcentrifuge tube.

Depending on sample size, more than one sample can be hybridized in a tube provided that the samples are completely submerged in solution. This facilitates direct comparative analysis of different samples.

14. Remove the solution and add 2 ml fresh prewarmed prehybridization solution A. Close the microcentrifuge tube, place it in a 65°C heating block, and secure it. Place the heating block sideways on a platform rocker. Prehybridize the sample 3 hr at 65°C.

This setup enables optimal and gentle mixing of the prehybridization solution, and is used for hybridization and all subsequent washes not carried out at room temperature. Alternatively, a shaking water bath may be used.

The samples can be stored up to 6 months in prehybridization solution A at -20°C either before or after completing step 14.

15. Replace prehybridization solution with 1 ml hybridization solution A containing 1 µg/ml digoxigenin-labeled riboprobe and hybridize overnight at 70°C.

See Critical Parameters and Troubleshooting for suggestions on recommended hybridization temperatures. Prehybridization and hybridization solutions can be stored at -20°C and reused up to three times.

One milliliter of hybridization solution is sufficient for three day-10.5 mouse embryos, two day-11.5 mouse embryos, or five stage-17 chicken embryos.

16. Proceed with enzymatic detection of RNA hybrids (see Alternate Protocol 2).

ALTERNATE PROTOCOL 2

ENZYMATIC DETECTION OF RNA HYBRIDS IN MOUSE AND CHICKEN EMBRYOS AND ORGANS

After hybridization (see Alternate Protocol 1) nonspecifically bound riboprobes are removed by extensive washes and RNase A digestion. Subsequently, digoxigenin-labeled RNA hybrids are detected immunologically using preabsorbed alkaline phosphatase-conjugated Fab fragments of anti-digoxigenin antibodies (see Support Protocol 2). Immune complexes are detected by an alkaline phosphatase color reaction.

This protocol is preferred to Basic Protocol 2 when there is high background. The steps with RNase, maleic acid buffer, and the preabsorption of the antibody may help to decrease unspecific staining.

Materials

Hybridized mouse or chicken embryo or organ sample in hybridization solution A (see Alternate Protocol 1)

Prehybridization solution A (see recipe), 70°C

2× SSC, pH 4.5 (see recipe for 20× SSC), 70°C

0.1% (v/v) CHAPS (3-[(3-cholamidopropyl)dimethylammonio]-1-propane-sulfonate)/2× SSC, 70°C

20 µg/ml RNase A (Boehringer Mannheim) in 0.1% CHAPS/2× SSC, 37°C
Maleic acid buffer (MAB; see recipe), 70°C and room temperature
PBS (APPENDIX 2)
PBT (also see recipe): 0.1% (v/v) Tween 20 in PBS
Blocking solution (see recipe)
Alkaline phosphatase–conjugated Fab fragments of anti-digoxigenin antibodies
(Boehringer Mannheim), preabsorbed (see Support Protocol 2), 4°C
0.1% (w/v) BSA in PBT
Alkaline phosphatase buffer, pH 9.5 (NTMT; see recipe), freshly prepared
2 mM levamisole (Sigma) in NTMT (optional)
NBT/BCIP substrate solution (UNIT 14.7) or BM purple AP substrate (Boehringer
Mannheim)

Plastic Pasteur pipets
70°C heating block with adapter for 2-ml microcentrifuge tubes
Rocker platform
37° and 70°C water bath (optional)
20-ml snap-cap glass vials
Aluminum foil
Dissecting microscope equipped with a camera

CAUTION: Formamide and DEPC are hazardous; see manufacturer's guidelines for handling, storage, and disposal.

NOTE: All incubations and washes are carried out at the temperatures indicated and with gentle rocking (see Basic Protocol 1, step 14). Solutions should be warmed to the indicated temperature before use.

Wash hybridized sample

1. Remove hybridization solution from hybridized mouse or chicken embryo or organ sample and add 800 µl prewarmed prehybridization solution A. Wash 5 min at 70°C.

The riboprobe in the hybridization solution is stable up to 2 years at –20°C and can be reused up to three times.

2. Add 400 µl of 2× SSC, pH 4.5, to the tube without removing prehybridization solution. Wash 5 min at 70°C. Repeat the addition of 2× SSC and wash two additional times.

Successive additions of 2× SSC to the prehybridization solution produce a stepwise dilution to 25% prehybridization solution/75% 2× SSC.

At high hybridization temperatures (e.g., 70°C), low pH (4.5) stabilizes the negative charges of the riboprobe and facilitates hybridization.

3. Remove the mix and wash two times, 30 min each wash, in 0.1% CHAPS/2× SSC at 70°C.
4. Digest sample with 20 µg/ml RNase A in 0.1% CHAPS/2× SSC for 1 hr at 37°C.

The conditions of RNase A digestion may need to be adjusted (see Critical Parameters and Troubleshooting).

5. Wash two times, 10 min each wash, in MAB at room temperature. Wash two times, 30 min each wash, in MAB at 70°C.

These washes significantly decrease nonspecific hybridization signals by blocking nonspecific antibody binding sites.

6. Wash two times, 10 min each wash, in PBS at room temperature.
7. Wash 5 min in PBT at room temperature.

Incubate with antibody

8. Incubate samples 2 to 3 hr in blocking solution at room temperature.
This step blocks nonspecific antibody-binding sites. The Fab fragments can be conveniently preabsorbed during step 8 (see Support Protocol 2).
9. Replace the blocking solution with 2 ml preabsorbed alkaline phosphatase-conjugated Fab fragments of anti-digoxigenin antibodies. Incubate overnight at 4°C with gentle agitation.
The Fab fragment solution can be reused up to three times within ~6 weeks. Store the solution at 4°C and check for possible contaminations before use.
The purpose of the subsequent washes is to remove excess and nonspecifically bound antibody from the tissues to prevent nonspecific reaction with chromogenic substrate.
10. Replace the Fab fragment solution with 0.1% BSA in PBT and transfer the sample to a 20-ml snap-cap glass vial.
For all subsequent steps, ~10 ml solution is used and all washes are carried out with gentle rocking at room temperature.
11. Wash five times, 45 min each wash, in 0.1% BSA in PBT.
12. Wash two times, 30 min each wash, in PBT.
13. Wash three times, 10 min each wash, in NTMT buffer.

Levamisole (2 mM final) may be included in the NTMT buffer to inhibit endogenous alkaline phosphatases.

If the Boehringer Mannheim BM Purple AP substrate is used, wash two times in NTMT buffer.

Visualize hybrids

14. Incubate in 3 ml NBT/BCIP substrate solution or BM Purple substrate solution. Cover glass vial with aluminium foil and gently rock the vial for the first 20 min of the development.
The sample must be completely submerged in substrate solution.
The NBT/BCIP substrate solution is light sensitive, and it is therefore important to keep the samples in the dark. Monitor progression of staining with minimal exposure to light.
BM Purple substrate solution is stabilized and presents superior signal-to-noise ratios. It produces clean signals with minimum background, even after overnight incubations. It is recommended for use when long developing times (overnight to 3 to 4 days) are required.
15. Monitor staining by eye or using a dissecting microscope. After the signal has developed to the desired extent, stop the staining reaction by washing the samples at least six times with PBT.

The stained samples can be stored ~2 to 3 months in PBT containing 100 mM EDTA at 4°C without significant effects on signal strength.

Signal strength depends on abundance of the target RNA, hybridization conditions, and additional factors (see Critical Parameters and Troubleshooting). Development times of up to 48 hr are required for some RNAs, whereas other probes yield strong signals after only 20 min. However, for most probes, incubation times of ~2 hr are required to obtain convincing results. The substrate solution should be replaced for overnight incubations.

It is possible to restain samples by repeating the procedure from steps 13 to 15.

16. (Optional) Fix sample overnight in 4% PFA in PBT at 4°C. Wash several times and store in PBT at 4°C.

This step fixes the signal, but may affect overall morphology. Furthermore, sample cannot be restained after fixation.

17. Photograph samples submerged in PBT in a petri dish.

Results should be documented as soon as possible. Samples are photographed submerged to avoid reflections and drying artifacts. It is best to illuminate the sample from the top or side using glass-fiber optics, and to use tungsten slide film or black-and-white negative film.

PREABSORPTION OF Fab FRAGMENTS WITH EMBRYONIC POWDER

To reduce nonspecific binding of the detecting antibody to embryonic tissue, it is important to preabsorb the antibody against an acetone powder prepared from the embryo or organ under study. The embryonic acetone powder (prepared as described in Wilkinson, 1993) is heat-inactivated and incubated with the alkaline phosphatase-conjugated Fab fragments of anti-digoxigenin antibody. The preabsorbed antibody is then ready for use in enzymatic detection of RNA hybrids in Basic Protocol 2 and Alternate Protocol 2. However, recent studies suggest that the preabsorption may be optional (A. Pizard, pers. comm.). The authors have obtained good results using Basic Protocols 1 and 2 without preabsorption when the antibody is used at a dilution of 1:2000 to 1:10,000.

Materials

Embryos or organs under study (e.g., embryonic day-10 to -12 mouse embryos or stage-24 to -26 chicken embryos)

Liquid nitrogen

Acetone, ice cold

PBT (also see recipe): 0.1% Tween 20 in PBS (*APPENDIX 2* for PBS)

BSA

Sheep serum (see recipe), heat-inactivated

Alkaline phosphatase-conjugated Fab fragment of anti-digoxigenin antibody (Boehringer Mannheim)

Blocking solution (see recipe)

Mortar and pestle

50-ml centrifuge tubes

Beckman TJ-6 centrifuge or equivalent, 4°C

Filter paper

Air-tight microcentrifuge tubes

Rocker platform

Prepare embryonic powder

1. Collect mouse embryos of embryonic day 10 to 12. Remove the extraembryonic tissues. Freeze the embryos in liquid nitrogen and store at -70°C until needed.

The frozen embryos may be stored up to 12 months at -70°C.

To produce 1 ml of embryo powder, use ~20 mouse embryos of embryonic day 11. For a high yield of chicken embryonic powder, use ~10 chicken embryos of stage 24 to 26.

2. Using a mortar and pestle, homogenize the embryos in liquid nitrogen to make a fine powder. Transfer the powder to a 50-ml centrifuge tube.
3. Evaporate the liquid nitrogen. Add 4 vol ice-cold acetone and mix well. Incubate 30 min on ice.

4. Centrifuge 10 min at $2500 \times g$ (4000 rpm in TJ-6 centrifuge), 4°C . Remove supernatant.
5. Wash pellet with ice-cold acetone and centrifuge again 10 min at $2500 \times g$, 4°C .
6. Spread the pellet on a sheet of filter paper and grind it to a fine powder. Allow to air dry and store in an air-tight microcentrifuge tube at -20°C .

The powder is stable for several months at -20°C .

7. Weigh out 3 mg embryo powder. Add 1 ml PBT and heat-inactivate 30 min at 70°C .

Each 3-mg embryo powder is sufficient to prepare 2 ml preabsorbed antibody at the appropriate dilution.

Preabsorb antibody

8. Cool the solution on ice. Add BSA to 1% (w/v) final concentration, heat-inactivated sheep serum to 10% (v/v) final concentration, and $1 \mu\text{l}$ alkaline phosphatase-conjugated Fab fragment of anti-digoxigenin antibody. Incubate 2 to 3 hr on a rocker platform at 4°C .
9. Centrifuge 10 min at $2500 \times g$, 4°C .
10. Dilute the supernatant to 2 ml with blocking solution to produce a final 1:2000 dilution of antibody.

Preabsorbed antibody may be stored up to 6 weeks at 4°C .

BASIC PROTOCOL 3

WHOLE-MOUNT IN SITU HYBRIDIZATION WITH *XENOPUS* EMBRYOS

Xenopus embryos are fixed to preserve morphology, then subjected to pretreatment to prepare the sample for hybridization with digoxigenin-labeled riboprobes. The authors have found that background can be substantially reduced by including a high concentration (1 mg/ml) of denatured salmon sperm DNA in the hybridization solution. Proteinase K treatment significantly increases the sensitivity of the method. This protocol is based on Harland (1991) with some modifications (Franco et al., 1999; López et al., 2003).

Materials

- Albino *Xenopus laevis* embryos
- 2% (w/v) cysteine, pH 7.8
- MEMFA buffer (see recipe)
- PBS (APPENDIX 2)
- 50% and 75% ethanol in sterile distilled water
- 100% ethanol
- 25% methanol/PBT
- PBT (also see recipe): 0.1% (v/v) Tween 20 in PBS, pH 7.8
- 2.5 $\mu\text{g/ml}$ proteinase K in PBT
- 0.1 M triethanolamine (TEA) buffer in PBT, pH 7.8 (UNIT 14.3, except use PBT), prepared fresh
- Acetic anhydride (Sigma)
- 4% formaldehyde in PBT
- Prehybridization solution B (see recipe)
- Hybridization solution B: prehybridization solution B (see recipe) containing 1 $\mu\text{g/ml}$ digoxigenin-labeled riboprobe (see Support Protocol 2)
- 2-ml screw-cap plastic conical tubes
- 37° and 60°C shaking water bath

CAUTION: All solutions should be treated with diethylpyrocarbonate (DEPC; see *UNIT 4.1*) and autoclaved to inhibit RNase activity. DEPC is a suspected carcinogen and should be handled carefully.

NOTE: The samples should always be kept covered with liquid to avoid drying, which causes high levels of background staining. Therefore, when exchanging solutions, always leave the sample covered by a small quantity of liquid. This also prevents physical damage or loss of samples.

NOTE: It is important that the embryos be kept in suspension during the early fixation steps, otherwise they will flatten.

Fix embryos

1. Incubate albino *Xenopus laevis* embryos in 2% cysteine, pH 7.8, for 3 to 5 min at room temperature to remove the jelly coat.
2. Transfer the embryos to a small petri dish and let the embryos settle. Remove most of the liquid and fill the vial with MEMFA buffer. Incubate 60 min at room temperature with gentle agitation to fix.
3. Replace the MEMFA buffer with PBS and wash three times, 10 min each wash, at room temperature.
4. Dehydrate gradually by incubating sample successively in 50% ethanol, 70% ethanol, and 100% ethanol, 10 min each, at room temperature.
5. Wash three times, 10 min each, in 100% ethanol, at room temperature. Store sample at -20°C .

Embryos can be stored for months in 100% ethanol at -20°C .

6. Rehydrate samples by incubating successively in 1 ml of 75% methanol, 50% methanol, and 25% methanol/PBT, 5 min each, at room temperature.
7. Wash three times, 5 min each, in PBT at room temperature.
8. Incubate embryos 10 min at room temperature in 2.5 $\mu\text{g/ml}$ proteinase K in PBT.

Skip this step if immunohistochemistry is to be performed after in situ hybridization, to preserve the antigenic structure.

Pretreat embryos

9. Wash two times, 5 min each, in 500 μl of 0.1 M TEA buffer in PBT, pH 7.8.
10. Add 1.25 μl acetic anhydride. Rock tubes 5 min at room temperature. Add an additional 1.25 μl acetic anhydride and rock for another 5 min.

*Tissues are acetylated to decrease background binding of probe. See *UNIT 14.3* (see *Background Information*) for a discussion of the rationale for the use of acetic anhydride and special considerations.*

11. Wash two times, 5 min each, in 1 ml PBT.
12. If proteinase K digestion was performed, refix embryos for 20 min with 4% formaldehyde in PBT, if not, skip steps 12 and 13.
13. Wash five times, 5 min each wash, in PBT.
14. Wash 10 min at room temperature in 1 ml fresh PBT and 250 μl prehybridization solution B.

15. Once embryos have settled through the solution, prehybridize them with 500 μ l fresh prehybridization solution B. Incubate 10 min in a 60°C shaking water bath.

A 500- μ l quantity of solution is sufficient for ten tadpoles.

16. Replace prehybridization solution B with a fresh 500- μ l aliquot. Prehybridize 3 hr at 60° to 65°C.

Hybridize

17. Replace prehybridization solution B with 500 μ l hybridization solution B containing 1 μ g/ml digoxigenin-labeled riboprobe. Hybridize overnight at 60° to 65°C.

18. Proceed with enzymatic detection of RNA hybrids (see Basic Protocol 4).

BASIC PROTOCOL 4

DETECTION OF RNA HYBRIDS IN *XENOPUS* EMBRYOS BY ANTI-DIGOXIGENIN ANTIBODIES

After hybridization (see Basic Protocol 3), nonspecifically bound probe is removed by a series of washes, and RNA hybrids are detected by anti-digoxigenin antibodies. For this protocol, maleic acid is included in antibody blocking and incubation steps and post-antibody washes (as recommended in the DIG system manual of Boehringer Mannheim and in Sive et al., 1994). It has been reported that several specific batches of antibody to digoxigenin have produced high backgrounds. This problem can be alleviated by replacing phosphate buffer with 2% blocking reagent (Boehringer Mannheim) in maleic acid buffer (MAB) for antibody washes and incubation (Lamb et al., 1993). Furthermore, the authors have found that extensive washing in maleic acid buffer before and after the antibody incubation result in clean signals with minimum background (Franco et al., 1999).

Materials

Hybridized *Xenopus laevis* embryos (see Basic Protocol 3)

Prehybridization solution B (see recipe), 60°C

2 \times and 0.2 \times SSC (see recipe for 20 \times), 60°C and room temperature

2 \times SSC containing 20 μ g/ml RNase A and 10 U/ml RNase T1 (see recipes)

0.1 (v/v) DEPC

Maleic acid buffer (MAB; see recipe)

2% (w/v) blocking reagent (see recipe) in MAB

Alkaline phosphatase-conjugated Fab fragments of anti-digoxigenin antibody (Boehringer Mannheim) diluted 1:2000 in 2% blocking reagent (see Support Protocol 1)

Alkaline phosphatase buffer, pH 9.5 (NTMT; see recipe) containing 5 mM levamisole (Sigma), filtered

NBT/BCIP substrate solution (UNIT 14.7) or BM purple AP substrate (Boehringer Mannheim)

Ethanol

4% (w/v) formaldehyde in PBT

37° and 60°C shaking water baths

Aluminum foil

Dissecting microscope equipped with a camera

CAUTION: Dimethylformamide and DEPC are hazardous; see manufacturer's guidelines for handling, storage, and disposal.

NOTE: All solutions should be treated with diethylpyrocarbonate (DEPC; see UNIT 4.1) and autoclaved to inhibit RNase activity.

NOTE: The samples should always be kept covered with liquid to avoid drying, which causes high levels of background staining. Therefore, when exchanging solutions, always leave the sample covered by a small quantity of liquid. This also prevents physical damage or loss of samples.

Wash embryos

1. Remove hybridization solution from hybridized *Xenopus laevis* embryos and add 500 μ l prehybridization solution B. Incubate 10 min at 60°C.
2. Wash three times, 20 min each wash, in 2 \times SSC at 60°C.
3. Remove 2 \times SSC and incubate in 2 \times SSC containing 20 μ g/ml RNaseA and 10 U/ml RNase T1. Incubate 30 min at 37°C.

RNases A and T1 have different specificities. Both are used to guarantee complete degradation of unhybridized probe and reduce background.

4. Wash 10 min in 2 \times SSC at room temperature.
5. Wash two times, 30 min each wash, in 0.2 \times SSC at 60°C. Add 0.1% fresh DEPC to each wash.
6. Wash two times, 15 min each wash, in MAB at room temperature.
7. Wash in 2% blocking reagent 15 min at room temperature.
8. Replace with fresh blocking reagent and incubate 1 hr at room temperature.

Detect RNA hybrids

9. Replace blocking reagent with 500 μ l alkaline phosphatase-conjugated Fab fragments of anti-digoxigenin antibody in 2% blocking reagent. Incubate 4 hr at room temperature with very gentle agitation.
10. Wash sample three times, 15 min each and leave overnight at 4°C in MAB with very gentle agitation.
11. Wash two times, 5 min each, in filtered NTMT buffer containing 5 mM levamisole at room temperature.
12. Replace last wash with NBT/BCIP substrate solution, and then wrap the container in aluminum foil. Incubate 5 min to 1 day at 37°C until the colored reaction product is visible.
13. After the signal has developed to the desired extent, stop the staining reaction by incubating in 100% methanol 30 min at room temperature.

This step is crucial to remove background.

14. Rehydrate embryos using 75% methanol/PBS; 50% methanol/PBS; PBS alone and then wash two times in PBS at room temperature.
15. Store embryos in 4% formaldehyde in PBS at 4°C.

SYNTHESIS OF DIGOXIGENIN-LABELED RNA PROBES

Antisense riboprobes are synthesized as run-off transcripts from linearized templates, using bacteriophage polymerases and template DNA consisting of the DNA fragment of interest cloned in a vector containing the promoter appropriate for the RNA polymerase (T3, T7, or SP6). RNA synthesis is carried out in the presence of a digoxigenin-substituted ribonucleotide. Nonradioactive probes have several advantages. They are easily synthe-

**SUPPORT
PROTOCOL 2**

**In Situ
Hybridization
and Immuno-
histochemistry**

14.9.15

sized in large quantities and are stable for several months. An additional advantage of RNA versus DNA probes is that they result in cleaner signals because nonspecifically bound probe is removed during ribonuclease treatment. See *UNIT 4.6* and the Commentary in *UNIT 14.3* for a discussion of RNA synthesis. This protocol is based on Wilkinson (1993). Also see Paganelli et al. (2001).

Materials

Distilled water, sterile
10× transcription buffer: 400 mM Tris·Cl (pH 8.25)/60 mM MgCl₂/20 mM spermidine (Boehringer Mannheim)
Nucleotide mix (pH 8): 10 mM GTP/10 mM ATP/10 mM CTP/6.5 mM UTP/ 3.5 mM digoxigenin-UTP (Boehringer Mannheim)
1 μg/ml linearized plasmid (*UNIT 4.7*)
40 U/ml placental ribonuclease inhibitor (RNasin, Boehringer Mannheim)
20 U/μl SP6, T3, or T7 RNA polymerase
RNase-free DNase I (*UNIT 4.1*)
TE buffer, pH 8 (*APPENDIX 2*) containing 0.1% SDS, DEPC-treated
Sephadex G-25 (coarse)
100% ethanol
3 M sodium acetate, pH 5.2 (*APPENDIX 2*)
70% ethanol in DEPC-treated water (see *UNIT 4.1*)
RNA loading buffer (see recipe)
37°C water bath
1-ml syringe
Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5A*)

1. Mix reagents in the following order at room temperature (20 μl total):

13 μl sterile distilled water
2 μl 10× transcription buffer
2 μl nucleotide mix
1 μl 1 μg/ml linearized plasmid
1 μl 40 U/ml RNasin
1 μl 20 U/μl RNA polymerase

Incubate 2 hr at 37°C.

Although stock solutions of sensitive reagents should be kept on ice, the reaction mixture should be mixed at room temperature; otherwise the spermidine in the transcription buffer will precipitate the DNA.

2. Add 2 μl RNase-free DNase I to the reaction mix. Incubate 15 min at 37°C.

Purify the probes

3. Add 80 μl TE buffer containing 0.1% SDS to the reaction mix.
4. Prepare the spin column filling a 1-ml syringe with Sephadex G25 (coarse) in TE buffer/0.1% SDS.
5. Eluate the sample by centrifuging 2 min at 1500 × g, room temperature.
6. Add 200 μl of 100% ethanol and 10 μl of 3 M sodium acetate, pH 5.2.
7. Mix and incubate overnight at -20°C.
8. Microcentrifuge 20 min at maximum speed, 4°C.

9. Wash pellet with 70% ethanol.
10. Air-dry pellet (do not use a Speedvac evaporator). Redissolve pellet in 20 μ l DEPC-treated water and store at -70°C .

Digoxigenin-labeled riboprobes are stable at least 2 years at -70°C .

Check the riboprobe

11. Mix 2 μ l of the probe with 23 μ l of RNA loading buffer.
12. Incubate 15 min at 65°C and then place 5 min on ice.
13. Run on 1% agarose gel in TBE buffer (*UNIT 2.5A*).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Alkaline phosphatase buffer (NTMT), pH 9.5

- 100 mM NaCl
- 100 mM Tris-Cl, pH 9.5 (*APPENDIX 2*)
- 50 mM MgCl_2
- 0.1% (v/v) Tween 20

Prepare fresh on the day of use to avoid decrease in pH during storage caused by absorption of carbon dioxide (Wilkinson, 1993).

*Levamisole (Sigma) is added to NTMT to block the activity of endogenous alkaline phosphatases. It should be added, just before use, to 5 mM for mouse and chicken embryos and 5 mM for *Xenopus* embryos.*

*It is necessary to filter the solution before use with *Xenopus* embryos.*

Blocking reagent, 2% (w/v)

Dissolve blocking powder (Boehringer Mannheim) in MAB (see recipe) by autoclaving 20 min, 1 atm, to give a 2% final concentration. Store at -20°C in aliquots.

Blocking solution

- 10% (v/v) heat-inactivated sheep serum (see recipe)
 - 1% (w/v) BSA
 - 0.1% (v/v) Tween 20
 - PBS (*APPENDIX 2*)
- Prepare fresh on the day of use

Hybridization solution

- 50% (v/v) deionized formamide (*UNIT 14.3*)
 - 5 \times SSC, pH 4.5 (see recipe for 20 \times)
 - 1% SDS
 - 50 μ g/ml yeast RNA
 - 50 μ g/ml heparin
- Store 25-ml aliquots up to 1 year at -20°C

Maleic acid buffer (MAB)

- 100 mM maleic acid (Sigma)
 - 150 mM NaCl
- Adjust pH to 7.5 with HCl
Autoclave
Prepare fresh on day on use

MEMFA buffer

0.1 M 3-(*N*-morpholino)propanesulfonic acid (MOPS), pH 7.4
2 mM EGTA
1 mM MgSO₄
3.7% (w/v) formaldehyde
Prepare fresh

PBT buffer

1× PBS (*APPENDIX 2*), DEPC-treated (*UNIT 4.1*)
0.1% or 1% (v/v) Tween 20
Store up to 6 months at 4°C

Prehybridization solution A

50% (v/v) deionized formamide (*UNIT 14.3*)
5× SSC, pH 4.5 (see recipe for 20×)
2% (w/v) blocking reagent (see recipe)
0.1% (v/v) Tween 20
0.5% (v/v) 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS)
50 µg/ml yeast RNA
5 mM EDTA
50 µg/ml heparin
Store 25-ml aliquots up to 1 year at –20°C

Add CHAPS from a 2% stock solution prepared in distilled sterile water (do not autoclave this solution). Add yeast RNA and heparin from 10 mg/ml stock solutions prepared in sterile distilled water.

Prehybridization solution B

50% (v/v) deionized formamide (*UNIT 14.3*)
5× SSC (see recipe for 20×)
1 mg/ml torula RNA (see recipe)
100 µg/ml heparin
1× Denhardt solution (*APPENDIX 2*)
0.1% (v/v) Tween 20
0.1% (v/v) 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS)
10 mM EDTA
1 mg/ml salmon sperm DNA (*UNIT 14.7*), denatured 10 min at 95°C
Adjust volume with DEPC-treated H₂O (*UNIT 4.1*)
Prepare fresh on day of use

RNA loading buffer

55% (v/v) formamide
20% (v/v) formaldehyde
22 mM 3-(*N*-morpholino) propanesulfonic acid (MOPS)
5.5 mM sodium acetate, pH 5.2 (*APPENDIX 2*)
0.5 mM EDTA, pH 8.0 (*APPENDIX 2*)
22% (v/v) glycerol solution: 30% (v/v) glycerol in water containing 0.25% (w/v) bromophenol blue
17 µg/ml ethidium bromide
Store up to 6 months at –20°C

RNase A

Dissolve RNase A (Sigma) to a concentration of 10 mg/ml in 10 mM Tris·Cl, pH 7.5/15 mM NaCl. Boil 10 min. Store in 100- μ l aliquots up to 12 months at -20°C .

RNase T1

Dissolve RNase T1 (Sigma) to a concentration of 10,000 U/ml in 0.1 M sodium acetate buffer, pH 5.5 (*APPENDIX 2*). Boil 10 min. Store 100- μ l aliquots at -20°C .

Keep a working stock of RNase at 4°C to avoid repeated freezing and thawing. It is stable for several months.

Sheep serum

Thaw sheep serum (Sigma-Aldrich) in a 37°C water bath. Heat inactivate the liquid serum 30 min at 70°C . Divide into 30-ml aliquots and centrifuge 15 min at $2500 \times g$ (4000 rpm in Beckman TJ-1 centrifuge), 4°C . Prepare 2-ml aliquots from the supernatant. Store 2-ml aliquots up to 1 year at -20°C ; do not refreeze aliquots once thawed.

Solution I

50% formamide
5 \times SSC, pH 4.5 (see recipe for 20 \times)
1% SDS
Store up to 6 months at room temperature

Solution II

50% formamide
2 \times SSC, pH 4.5 (see recipe for 20 \times)
0.1% (chick embryos older than stage 4 to stage 9) or 1% (mouse) Tween 20
Store up to 6 months at room temperature

SSC, pH 4.5, 20 \times

Prepare SSC as described in *APPENDIX 2*. Adjust pH to 4.5 with citric acid. Autoclave, then store at room temperature (stable for months).

TBS, 10 \times

1.4 M NaCl
27 mM KCl
250 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
Store up to 6 months at room temperature

TBST

1 \times TBS (see recipe for 10 \times) plus 0.1/1% Tween 20
2 mM 0.5 mg/ml Levamisole
Add Tween 20 on day of experiment

NOTE: *Most no longer add levamisole to TBST.*

Torula yeast RNA, 10 mg/ml

Dissolve torula yeast RNA (Sigma) in water at a concentration of 20 mg/ml. Heat to 60°C and sonicate until completely resuspended. Extract with 25:24 (v/v) phenol/chloroform and precipitate with 2 vol ethanol. Resuspend pellet in DEPC-treated water (*UNIT 4.1*) to a final concentration of 10 mg/ml. Store 100- μ l aliquots up to 12 months at -20°C .

Torula yeast RNA is less expensive than yeast tRNA.

COMMENTARY

Background Information

Nonisotopic in situ hybridization using intact embryos or organs is a widely used method that enables fast determination of RNA spatial distributions. Large numbers of samples can be analyzed simultaneously, which permits direct comparison of different developmental stages and/or genotypes. Furthermore, the sensitivity and reproducibility of whole-mount in situ hybridization facilitates the analysis of genes expressed in restricted patterns, i.e., in small structures easily overlooked when using serial sections (UNITS 14.1-14.5). In addition, the opportunity for inspecting histological details is not lost because the embryos can be sectioned after staining. Therefore, whole-mount in situ hybridization is often the method of choice to determine the expression patterns of novel gene products (Rosen and Beddington, 1993).

Protocols enabling simultaneous detection of two different transcripts by whole-mount in situ hybridization have been developed. Probes labeled with digoxigenin and fluorescein can be detected differentially by antibodies coupled to different enzymatic detection systems

(Hauptmann and Gerster, 1994; Jowett and Lettice, 1994). The methods described here can also be used for detecting multiple gene products using digoxigenin-labeled riboprobes, provided the target transcripts are localized in distinct structures or cell types of the embryo.

Critical Parameters and Troubleshooting

The methods described in this unit have been used to localize different types of transcripts in mouse, chicken, and *Xenopus laevis* embryos (e.g., Haramis et al., 1995; Figs. 14.9.1, 14.9.2, and 14.9.3). However, the parameters discussed in the following paragraphs may require optimization depending on particular tissues and target mRNAs. Furthermore, controls are suggested that should aid in the adaptation of these protocols to “difficult” tissues or probes. In general, it is important to assure that the target RNAs remain intact during fixation and subsequent manipulation of samples. Therefore, embryos or organs should be isolated at 4°C and fixed as fast as possible. It is also important to use RNase-free solutions to avoid possible

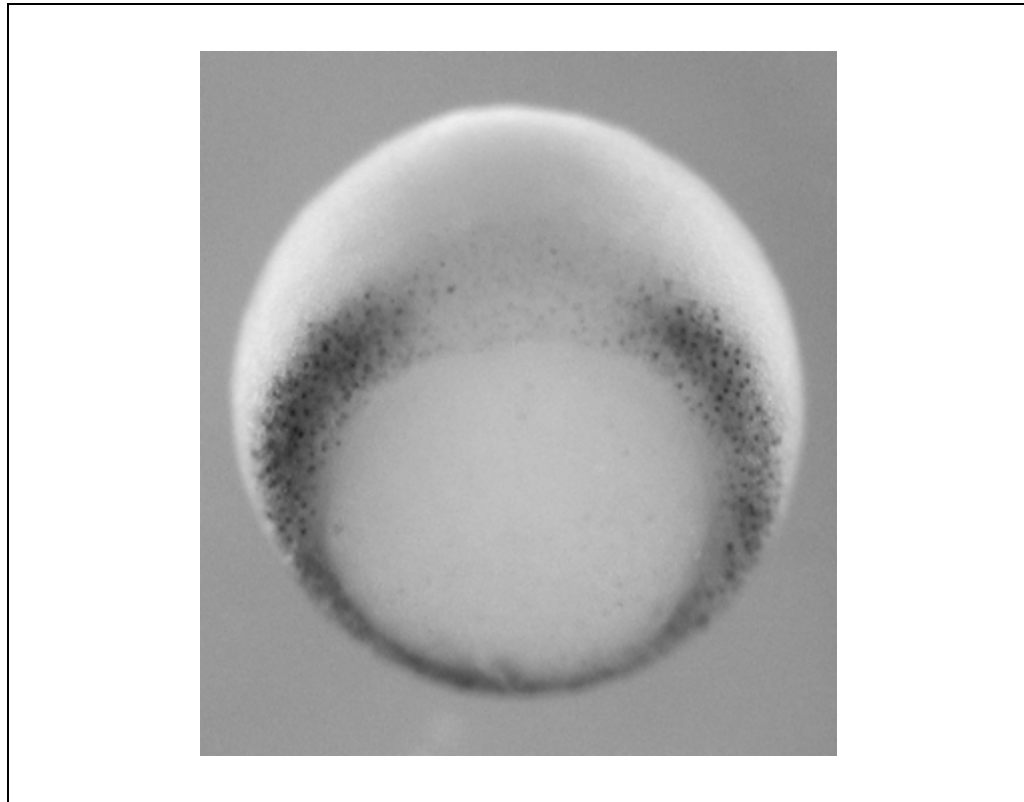


Figure 14.9.1 Whole mount ISH of *Xenopus* embryo. Detection of *X-delta-1* mRNA in *Xenopus laevis* whole-mounted gastrula (stage 11) using digoxigenin-labeled RNA antisense probe. Magnification at ~200 \times . (Contributed by Carrasco et al.)

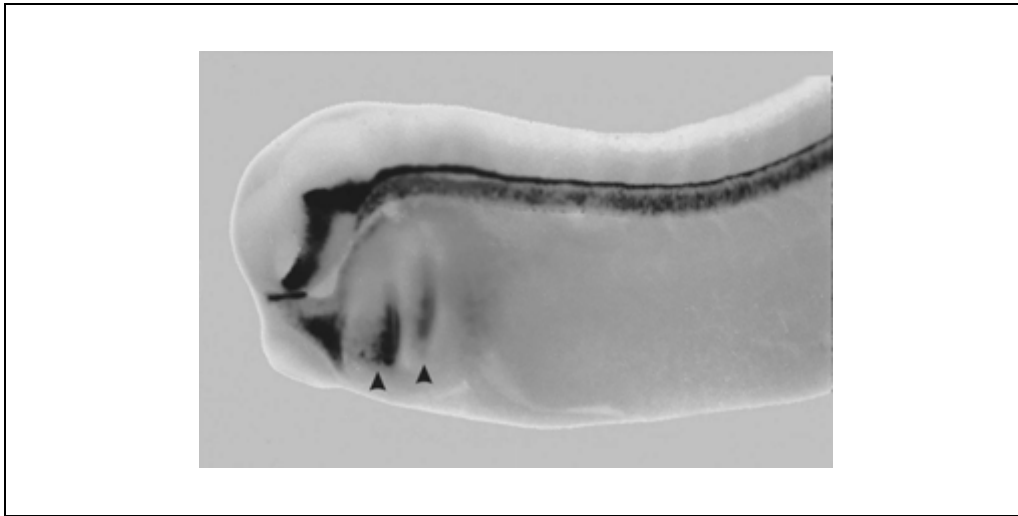


Figure 14.9.2 Whole mount ISH of *Xenopus* tadpole. Detection of sonic hedgehog (*X-shh*) mRNA in *Xenopus laevis* whole-mounted tadpole (stage 30) using digoxigenin-labeled RNA antisense probe. Magnification at ~100 \times . (Contributed by Carrasco et al.)

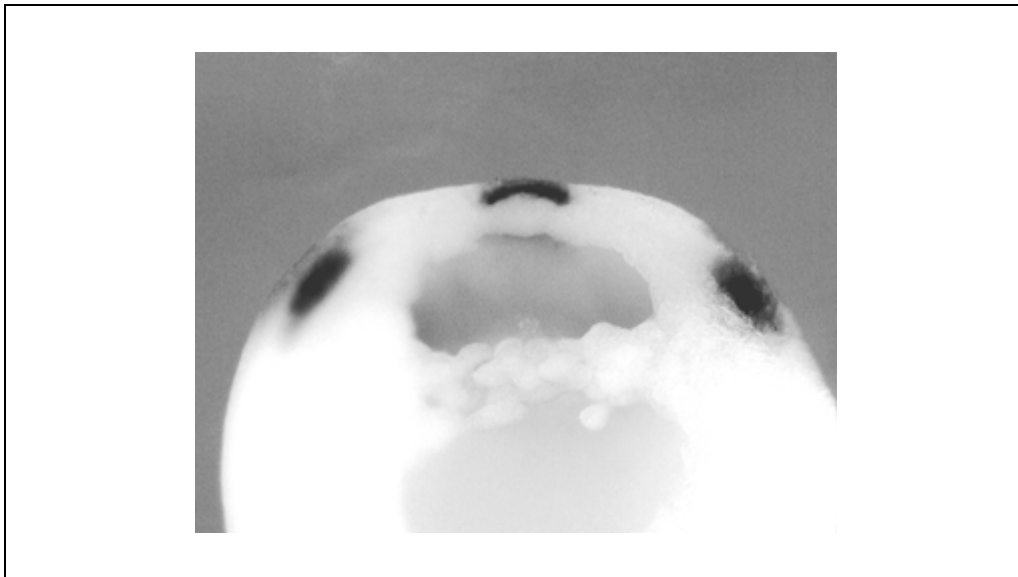


Figure 14.9.3 Whole mount ISH of *Xenopus* embryo. Detection of *hairy 2A* mRNA in *Xenopus laevis* transversal section of whole-mounted neurula (stage 14) using digoxigenin-labeled RNA antisense probe. Magnification at ~200 \times . (Contributed by Carrasco et al.)

degradation of target RNAs, so solutions should be autoclaved or treated with diethylpyrocarbonate (DEPC; see *UNIT 4.1*). During initial experiments, it is advisable to perform the procedure exactly as described and to complete it without intermediate storage. Furthermore, a probe with a known distribution in the tissue(s) of interest should be included as a positive control. Samples from which the riboprobe and/or anti-digoxigenin Fab fragments have been omitted serve as negative controls for nonspecific labeling and/or activity of endogenous alkaline phosphatases. Alternatively, a sense-strand digoxigenin-labeled riboprobe

can be synthesized and used as a negative control.

Fixation

Fixation of embryos and organs for whole-mount in situ hybridization is less critical than when using ^{35}S -labeled riboprobes on histological sections (Carrasco and Malacinski, 1987; Franco et al., 1999).

Bleaching with hydrogen peroxide

In the case of pigmented *Xenopus* embryos, extensive treatment may be required to bleach pigment granules. However, embryos of an

Table 14.9.1 Troubleshooting Guide for Whole-Mount In Situ Hybridization

Problem	Cause	Solution
Low or no hybridization signal	Low abundance or degradation of target RNA	Optimize fixation time. Optimize concentration of enzyme and duration of proteinase K digestion.
	Problems with riboprobe	Lower hybridization temperature, e.g., to 55°C. Increase concentration of riboprobe.
	Overdigestion with RNase	Use lower concentration of enzyme or shorter digestion time, e.g., 30 minutes. Perform digestion at room temperature.
High levels of nonspecific signal	Hybridization conditions not stringent enough	Increase stringency of hybridization and washes (temperature, ionic strength, or incubation time). For <i>Xenopus</i> , increase concentration of salmon sperm DNA.
	Incomplete RNase digestion	Increase RNase to 100 (µg/ml).
	Nonspecific binding of Fab fragments of antibody	Preabsorb antibody with tissue of interest. Extend preabsorption to overnight at 4°C.
	Residual endogenous alkaline phosphatase activity	Add levamisole to NTMT buffer just before use (2 mM for mouse/chicken, 5 mM for <i>Xenopus</i>).
Loss of morphology	Chromogenic substrate too old	Use fresh substrate or stabilized BM Purple Substrate Solution (Boehringer Mannheim).
	Damage to sample during handling	Handle samples carefully and check for damage at each transfer step.
	Insufficient fixation or postfixation after proteinase K digestion	Increase fixation or postfixation time.
	Overtreatment with hydrogen peroxide	Decrease treatment time.

albino mutant strain are used to circumvent the necessity of bleaching.

Proteinase K digestion

Proteinase K digestion permeabilizes the tissues, unmask target RNAs by removing proteins, and inactivates endogenous ribonucleases. It is important to establish optimal proteinase K digestion conditions for each sample. Optimal conditions might vary depending on the batch and/or supplier of the enzyme. Conditions are determined empirically in pilot experiments using varying concentrations of proteinase K. Short, nonoptimal digestion may reduce signals and increase background levels, whereas overdigestion will cause degradation of sensitive structures, especially when very young embryos are studied. The amount of enzyme, the duration of digestion, or both may be varied to optimize proteinase K digestions (see Basic Protocols 1 and 3). Postfixation of samples after proteinase K digestion is essential to retain unmasked RNAs and preserve morphology during subsequent hybridization. Proteinase K digestion must be omitted when an immunohistochemistry procedure is performed to preserve the antigenic structure.

Selection of riboprobe and hybridization conditions

Before attempting to study the distribution of RNAs of interest in a particular tissue, it is important to establish their abundance and temporal distribution by northern blot hybridization (UNIT 4.9) and/or an RNase protection assay (UNIT 4.7). RNase protection assays should be used to select possible probes for subsequent whole-mount in situ hybridization (also see UNIT 14.3, Background Information). Several parameters affect the strength of the hybridization signal and signal-to-noise ratio. Both GC- and AT-rich sequences should be avoided. The best probes are often derived from coding regions. However, when analyzing expression of members of highly conserved gene families, the probes must be derived from the least conserved regions to ensure detection of specific RNAs. Sensitivity of detection and signal-to-noise ratios depend on probe length and concentration. However, probe length seems less critical when using digoxigenin (or other haptens) than when using ³³S-labeled riboprobes (UNIT 14.3). Riboprobes up to 2.5 kb long have been used successfully without previous alkali hydrolysis (Franco et al., 1999).

RNase A digestion

RNase A digestion eliminates nonspecifically hybridized riboprobes and is essential to obtaining the best signal-to-noise ratios. Specifically hybridized RNA hybrids are stable, but RNase concentrations may have to be lowered or digestion omitted when detecting heterologous or rare mRNAs. Pilot experiments using varying RNase concentrations and/or digestion times are necessary to establish optimal conditions.

Troubleshooting

Table 14.9.1 is a guide to troubleshooting problems associated with whole-mount in situ hybridization.

Anticipated Results

A successful experiment will reveal the spatial distribution of target RNAs in the samples of interest. Good examples for the analysis of expression patterns by whole-mount in situ hybridization can be easily found in the references listed and in major scientific journals (Franco et al., 1999; Paganelli et al., 2001; López et al., 2003). Representative examples of experiments carried out using the methods described in this unit are illustrated in Figures 14.9.1, 14.9.2, and 14.9.3.

Time Considerations

Before starting whole-mount in situ hybridizations, large amounts of tissue must be collected for preparation of the embryonic acetone powder. Also, riboprobes should be tested by northern blot hybridization (UNIT 4.9) or RNase protection assays (UNIT 4.7). In general, three overnight steps are required during the procedure (fixation, hybridization, and antibody incubation). Linearization of the template DNA can be done overnight during the fixation step. Preparation of riboprobes requires 4 to 5 hr. Pretreatment of samples before hybridization requires 4 hr of hands-on work and an additional 3 hr for prehybridization. Posthybridization washes and blocking prior to antibody treatment require a total of 8 hr. Post-antibody washes require ~6 hr. Signal detection requires a minimum of 2 hr and can extend overnight. If different samples are to be collected over a period of time and analyzed at one time, they may be stored at -20°C either after fixation in 100% ethanol or after prehybridization or in prehybridization buffer.

Literature Cited

- Carrasco, A.E. and Malacinski, G. 1987. Localization of *Xenopus* homeo-box gene transcripts during embryogenesis and in the adult nervous system. *Dev. Biol.* 121:69-81.
- Crossley, P.H. and Martin, G.R. 1995. The mouse *Fgf8* gene encodes a family of polypeptides and is expressed in regions that direct outgrowth and patterning in the developing embryo. *Development* 121:439-451.
- Franco, P., Paganelli, A., López, S., and Carrasco, A.E. 1999. Functional association of retinoic acid and hedgehog signaling in *Xenopus* primary neurogenesis. *Development* 126:4257-4265.
- Haramis, A.G., Brown, J.M., and Zeller, R. 1995. The limb deformity mutation disrupts the SHH/FGF-4 feedback loop and regulation of 5' HoxD genes during limb pattern formation. *Development* 121:4237-4245.
- Harland, R.M. 1991. In situ hybridization: An improved whole-mount method for *Xenopus* embryos. *Methods Cell Biol.* 36:685-695.
- Hauptmann, G. and Gerster, T. 1994. Two-color whole-mount in situ hybridization to vertebrate and *Drosophila* embryos. *Trends Genet.* 10:266.
- Jowett, T. and Lettice, L. 1994. Whole mount in situ hybridizations on zebrafish embryos using a mixture of digoxigenin- and fluorescein- labeled probes. *Trends Genet.* 10:73-74.
- Lamb, T.M., Knecht, A.K., Smith, W.C., Stachel, S.E., Economides, A.N., Stahl, N., Yancopolous, G.D., and Harland, R.M. 1993. Neural induction by the secreted polypeptide noggin. *Science* 262:713-718.
- López, S., Paganelli, A., Rosato Siri, M., Ocaña, O., Franco, P., and Carrasco, A.E. 2003. Notch activates sonic hedgehog and both are involved in the specification of dorsal midline cell-fates in *Xenopus*. *Development* 130:2225-2238.
- Paganelli, A., Ocaña, O., Prat, M., Franco, P., López, S., Morelli, L., Adamo, A., Riccomagno, M., Matsubara, E., Shoji, M., Affranchino, J., Castaño, E., and Carrasco, A.E. 2001. The Alzheimer-related gene *presenilin-1* facilitates sonic hedgehog signaling in *Xenopus* primary neurogenesis. *Mechan. Dev.* 107:119-131.
- Riddle, R.D., Johnson R.L., Laufer, E., and Tabin, C. 1993. Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell* 75:1401-1416.
- Rosen, B. and Beddington, S.P. 1993. Whole mount in situ hybridization in the mouse embryo: Gene expression in three dimensions. *Trends Genet.* 9:162-167.
- Schulte-Merker, S. 1993. In situ hybridization. In *The Zebrafish Book*, 2nd ed. (M. Westerfield, ed.) pp. 9.16-9.21. University of Oregon Press, Eugene, Oreg.
- Sive, H.L., Grainger, R.M., and Harland, R.M. (eds.) 1994. *Early Development of Xenopus laevis*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Tautz, D. and Pfeifle, C. 1989. A nonradioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. *Chromosoma* 98:81-85.

Wilkinson, D.G. 1993. In situ hybridization. *In* Essential Development Biology. A Practical Approach (C.D. Stern and P.W.H. Holland, eds.) pp. 257-276. IRL Press, Oxford.

Contributed by Anne Pizard (ISH in amniotic embryos and organs)
Harvard Medical School
Boston, Massachusetts

Anna Haramis
EMBL
Heidelberg, Germany

Andrés E. Carrasco, Paula Franco, Silvia López, and Alejandra Paganelli
(*Xenopus* embryos)
University of Buenos Aires-CONICET
School of Medicine
Buenos Aires, Argentina

Principles and Application of Fluorescence Microscopy

Fluorescence is the luminescent emission that results from absorption of photons. Fluorescence is distinguished from its counterpart, a longer-lasting afterglow called phosphorescence, by the magnitude of the decay time. Fluorescent emission ceases abruptly when the exciting energy is shut off. The decay time, or afterglow, of the emission is on the order of 10^{-8} sec and results in a negative frequency-shifted emission. In contrast, the decay in phosphorescence takes place in milliseconds to seconds.

The fluorescent effect is used in a number of spectroscopy techniques, and it is particularly useful for fluorescence microscopy. The principal use of fluorescence microscopy is to examine specimens that have been treated with special fluorescent reagents. These reagents are able to absorb light of a certain wavelength and emit light at a longer wavelength slightly shifted toward the red end of the spectrum from the absorbed light. If, for example, blue light is absorbed, green light will be emitted. Green is shifted to yellow, yellow to red, and invisible UV light to visible blue light. This phenomenon

is termed the Stokes shift and is defined as the separation of the spectral maxima of excitation and emission (see Fig. 14.10.1).

The λ_{\max} of the spectrum is typically ~20 to 50 nm longer than that of the absorbed exciting light. The Stokes shift, however, can range from <10 to >100 nm (see Fig. 14.10.2). Each fluorochrome exhibits its own very specific absorption and emission spectra, depending on the structure of the molecules and sometimes also on their surroundings.

Fluorescence microscopy allows selective examination of a particular component of a complex biomolecular assembly. A specimen labeled with fluorescent dye(s) is illuminated with filtered light of the absorbing wavelength and is viewed using a barrier filter that is opaque to the absorbing wavelength but which transmits the longer wavelength of the emitted light. The structures marked with the fluorescent molecules will light up against the black background. Additional experimental information can be derived from the combination of optical and biochemical responses exhibited by the

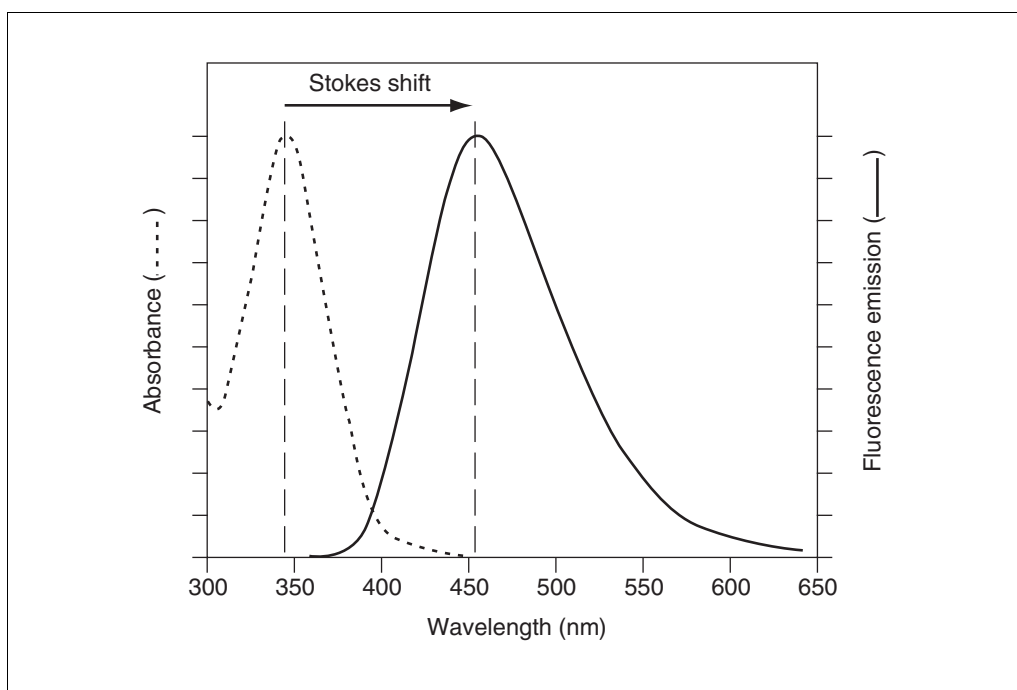


Figure 14.10.1 Absorption and emission spectra of Hoechst 33342 (Molecular Probes) bound to DNA. The dye solution is excited with a fixed-wavelength monochromatic light (typically the λ_{\max} of the absorption spectrum) and the emission measured with a spectrofluorometer. Spectra have been normalized, and relative intensities are plotted to illustrate the peak shift from absorbed to emitted light termed the “Stokes shift.” Spectra courtesy of Molecular Probes.

Contributed by Donald Coling and Bechara Kachar

Current Protocols in Molecular Biology (1998) 14.10.1-14.10.11

Copyright © 1998 by John Wiley & Sons, Inc.

fluorescent probe.

The growing importance in biology of fluorescence microscopy is due to (1) the extraordinary development of new fluorescent molecular probes and (2) the development of improved low-light-level imaging systems and confocal microscopy techniques (see *UNIT 14.11*).

FLUORESCENT MOLECULAR PROBES

Fluorescent dyes or fluorochromes can be linked to specific antibodies or to other specific molecular probes. Many such fluorescent molecular probes are commercially available (e.g., from Amersham, Calbiochem-Novabiochem, Molecular Probes, and Sigma). Fluorescent probes are identified and quantitated by their absorption and fluorescence emission wavelengths and intensities. Some examples are shown in Figure 14.10.2.

An important experimental consideration is separation of the fluorescence emission signal from the scattered excitation light. This is facilitated by choosing a fluorescent probe with a large Stokes shift. Maximization of the spectral separation can be achieved by choosing

fluorophores with narrow spectral peaks.

During excitation of a fluorescent reagent only some of the absorbed photons result in fluorescent emission. "Quantum yield" is a term that describes the ratio of emitted to absorbed photons. Good fluorescent reagents have a high quantum yield. Additional optical signals may be due to endogenous background fluorescence or to an analytically useful second emitting species.

Fluorescence detection is compromised by background signals. Many biological molecules fluoresce when illuminated with ultraviolet light. This effect is called autofluorescence. The fluorescent amino acid tryptophan, for instance, occurs in a large number of proteins. This molecule is of limited value in cell biology because of the lack of selectivity. Fortunately, autofluorescence of cells and biological tissues and fluids can be minimized by using probes that can be excited at >500 nm.

FILTERS AND FILTER SETS

Filters for fluorescence microscopy (available, e.g., from Omega Optical, Chroma Technology, and microscope manufacturers) must

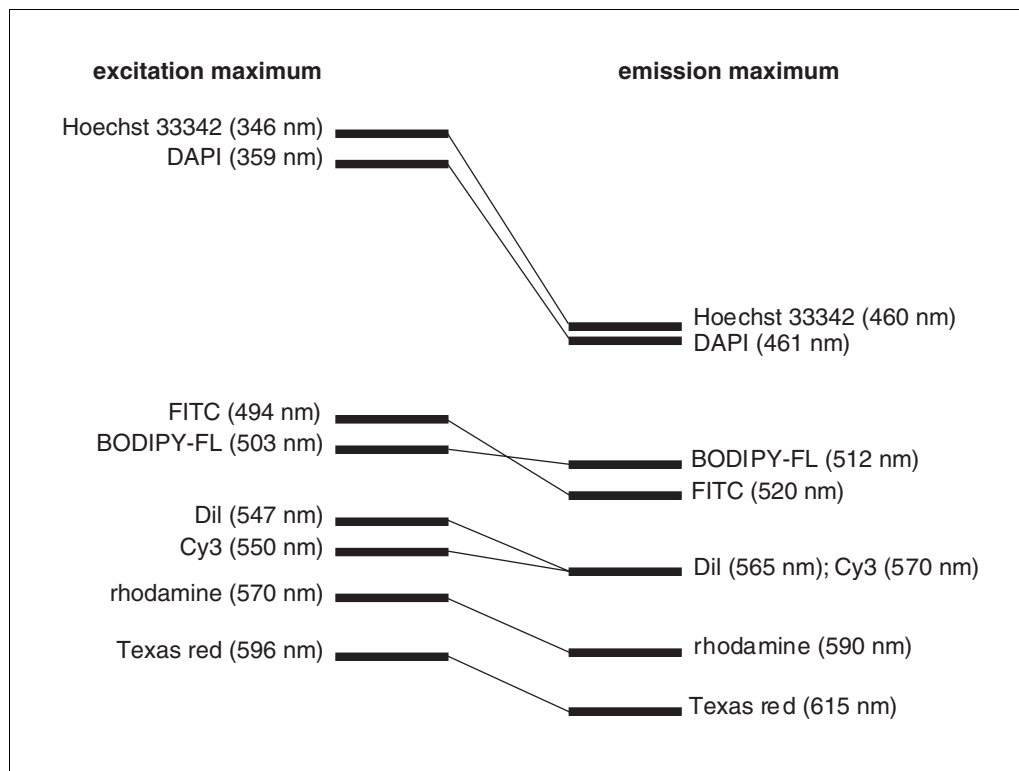


Figure 14.10.2 Excitation and emission maxima for commonly used fluorophores. DAPI, 4',6-diamidino-2-phenylindole; FITC, fluorescein isothiocyanate; Dil, 1,1'-dihexyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate. BODIPY FL and CY3 have complex chemical structures and names. More information can be obtained from the respective suppliers, Molecular Probes and Amersham Life Science.

transmit light of the required wavelengths and block other light as completely as possible. This is difficult to do, because the excitation light intensity is many times higher than that of the emission light. Ideally, light at the excitation wavelength is blocked from the microscopic image, while the emitted light arrives with minimal attenuation.

In practice, a narrow optical bandwidth (<20 nm) is usually used for fluorescence excitation, whereas the fluorescence detection bandwidth is variable, ranging from full spectrum for maximal sensitivity to narrow band (~20 nm) for maximal resolution.

Early types of filters consisted of simple colored glass or the combination of two filters to form so-called band-pass filters. They were made by combining a filter passing long wavelengths with one passing short wavelengths, giving a band-pass where the transmission curves overlap. Contemporary interference filters can achieve much sharper cutoffs with high transmission in the selected region. They consist of a sheet of glass coated with several thin layers of transparent materials of different refractive indices. Light partially reflected at each boundary between layers combines to interfere constructively for some wavelengths and destructively for others. Such interference filters can be used for reflection as well as for transmission, and they can be tailor-made to meet the spectral requirements for a particular fluorescent dye. Figure 14.10.3 shows an example of both an excitation and an emission scan of FITC along with transmission spectra of filters

optimized for FITC spectra. The resulting unfiltered signals overlap significantly, but the filtered signals are discrete.

Fluorescence microscopy is implemented using epiillumination whereby the specimen is illuminated by light passing through the objective. Since epifluorescence microscopy requires the use of a high-intensity light source, a main advantage of this arrangement over transillumination is the safety afforded to the retina. The optical paths of excitation and emission light are shown in Figure 14.10.4. The excitation filter allows only a narrow band of monochromatic light to project on the surface of the dichroic mirror (a chromatic beam splitter, sometimes referred to as a “dichroic”). The dichroic, inclined 45° to the incident light, reflects short-wavelength radiation and transmits longer wavelengths. It reflects the excitation radiation (and any back-scatter from the specimen) and transmits to the observer all of the fluorescent radiation above the cutoff wavelength. Above the beam splitter, the emission light and the remainder of the excitation light hit the emission (barrier) filter where only the emission light can pass to form the fluorescence image of the specimen against a black background (see Fig. 14.10.5).

MULTIBAND FILTERS AND MULTIDYE FLUORESCENCE

The combined use of two or three different dyes simultaneously is becoming increasingly prevalent. Different structures light up in different colors that can be viewed separately, or

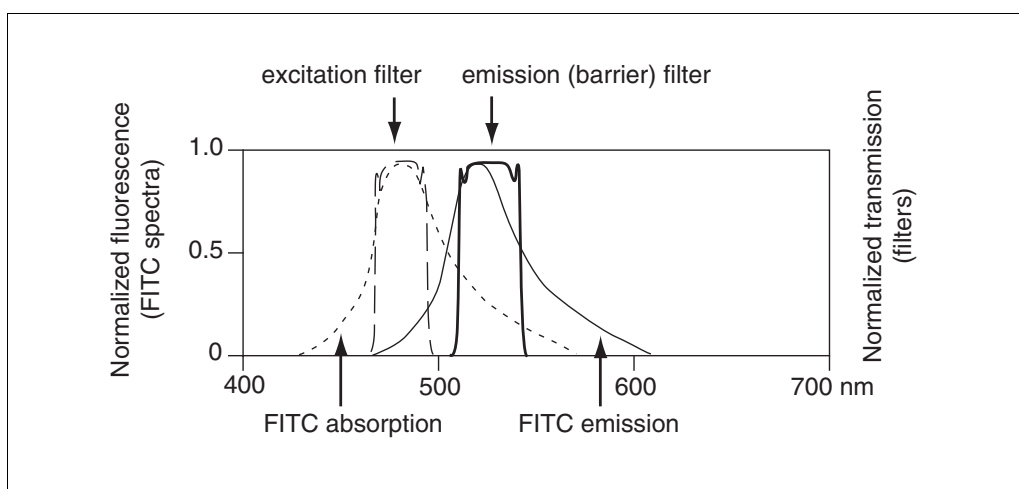


Figure 14.10.3 Transmission spectra of excitation and emission (barrier) filters. Filter characteristics are optimized to match the excitation and emission spectra of fluorescein isothiocyanate (FITC; rolling slopes). Spectra are normalized for comparison. Note the sharp drop-off of the filter spectra in relation to the FITC spectra. (For a pure sample, the absorption and excitation spectra are superimposable.) Courtesy of Omega Optical.

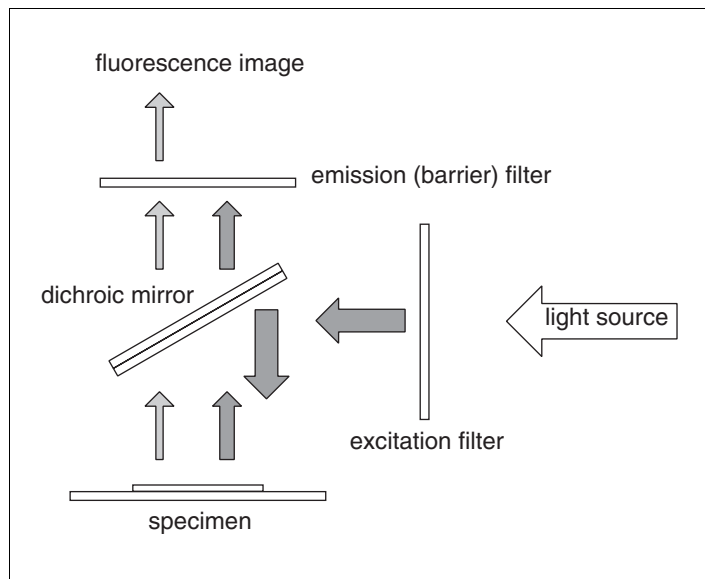


Figure 14.10.4 Physical arrangement of filters and light source in a typical epifluorescence microscope.

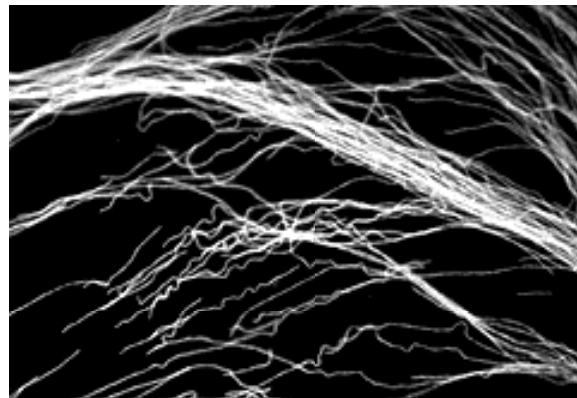


Figure 14.10.5 Single-color fluorescence microscopy of microtubules labeled with mouse anti- β -tubulin and FITC-coupled anti-mouse immunoglobulin (Amersham). This is a 35-mm micrograph of a small area in the cytoplasm of a cultured epithelial cell obtained using a 100 \times , 1.4-NA (numerical aperture) objective. Individual microtubules and bundles can be visualized within the complex structure of the cytoplasm.

simultaneously, using special filter sets. Most microscopes have reflector sliders that permit up to three or four sets of filters to be mounted for different fluorescent markers, making the changeover easy.

Since the mid 1980s, multidye fluorescence filter sets have been used more frequently. The sets have the same general components as single-color filter sets, i.e., an excitation filter, a beam splitter, and an emission filter. In each component, there are two or more regions of controlled reflection and transmission. Such an assembly allows the simultaneous imaging of at least two fluorescent labels.

Several standard dual-dye sets are available from some of the major microscope and filter manufacturers (e.g., Zeiss, Leica, Nikon, Olympus, Chroma Technology, and Omega Optical). Omega Optical specializes in producing not only standard multidye sets but also custom-made multidye filter sets including triple (see Fig. 14.10.6) and quadruple dye sets.

Figure 14.10.7 shows an example of three-color staining using cultured endothelial cells labeled with three different fluorescent markers.

The most common use of multiband filters is to investigate the relative spatial localization of two antigens. This is typically done with red and green dyes which form a yellow image from spectral summation when the antibodies colocalize. Colocalization of antigens on organelles as small as stress fibers and vesicles can be detected. Multidye filter sets with more than two-color capability are useful for illuminating intracellular organization of organelles and organization of cells within a tissue. To this end, nucleic acid dyes that stain the nucleus blue, such as DAPI and Hoechst 33342, are among the most commonly used dyes to complement red and green. Use of four or more dyes simultaneously is technically possible and has been accomplished for artistic views of cellular architecture. The use of more than three dyes at once, however, has not yet proved to have the

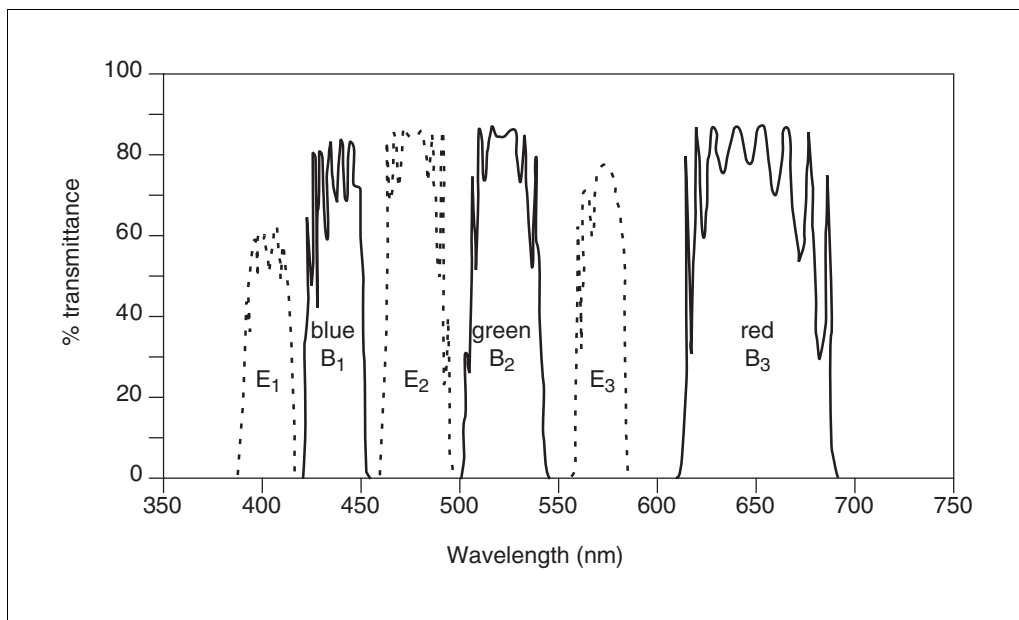


Figure 14.10.6 Spectra for a triple-dye filter set for blue, green, and red dyes. E₁ to E₃, excitation filters; B₁ to B₃, emission filters. Courtesy of Omega Optical.

analytical value of simple double labeling. The reason for this is undoubtedly linked to the difficulty of optimizing the conditions required for color balance that are encountered in sample preparation, data acquisition, and final image reproduction.

LIGHT SOURCES

Lamps illuminated by filaments are not suitable as light sources in fluorescence microscopy, because the glowing metal filaments convert most of the electrical energy into red or even invisible infrared light. The light source must provide a large amount of excitation energy in very narrow ranges of the spectrum, typically 10 to 50 nm. For this purpose, so-called line emitters, usually high-pressure mercury lamps, are used. The high-pressure mercury lamp uses the gas discharge principle and features a discrete light spectrum. Figure 14.10.8 shows the line spectrum of a mercury arc lamp with its characteristic emission peaks.

The ability to choose absorption and emission filters that take advantage of the intensity lines and the weak spectral ranges is a major benefit of this line emitter. Excitation can be made with a single line, and, because of the Stokes shift, the fluorescence can be viewed at a wavelength at which the light of the illuminator causes minimal disturbance.

Light from a mercury arc lamp radiates from ionizing gases in an arc between discrete points on the anode and cathode. A 2-hr “burning-in”

period should be allowed after lamp replacement before adjusting the optics. After that, the points at which the arc contacts the surfaces of the electrodes stabilize, and readjustment is required only occasionally to maintain uniform illumination of the field. Although an arc lamp may remain stable for 6 months to a year, it should be checked for even illumination with each use. The simplest way is to move the specimen stage and verify that the intensity of a brightly labeled cell or organelle does not vary within the field. In this way, a novice, but conscientious, user can visually detect most alignment problems that stem either from tampering, curious hands, or a spontaneously occurring change in the position of the arc on the surface of the lamp electrodes. Alignment procedures are not complex and are well documented in operators manuals. Additionally, most manufacturers have technical specialists who are able to assist users either by telephone or through courtesy visits. In making adjustments, note that fluorescence filters are sensitive to the heat emitted by the illuminator, so the heat protection filter should never be removed.

Lamps are expensive, and the time to install and align them is costly as well. Short on-off cycling times should be avoided since this necessitates more frequent realignments or lamp replacement. The life of a mercury arc lamp can be extended by planning imaging sessions for >1 hr whenever possible. Check with other users before turning off the lamp, and try to

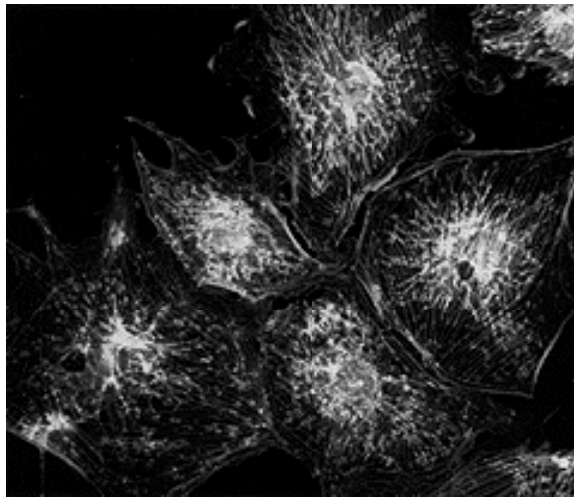


Figure 14.10.7 Three-color staining observed simultaneously with a multidy fluorescence filter set. Cultured endothelial cells were labeled with Mito Tracker Red CMXRos, BODIPY-FL phalloidin, and DAPI. The slide, which was purchased from Molecular Probes, is an excellent diagnostic and teaching tool, showing red fluorescent mitochondria, green fluorescent actin filaments, and blue fluorescent nuclei. *This black and white facsimile of the figure is intended only as a placeholder; for full-color version of figure go to <http://www.currentprotocols.com/colorfigures>*

schedule experimental sessions so that the lamp does not need to be turned off between users (<1 hr between users).

MICROSCOPE OBJECTIVES

The fluorescence emission from the excited specimen is radiated in all directions. The objective must gather as much of this radiation as possible. High-NA objectives are the most effective for this purpose. If the objective aperture is doubled, approximately four times more fluorescent light is gathered. Moreover, the use of immersion oil eliminates the loss of light caused by reflection on the surface of the objective front lens.

The best general-purpose microscope objectives offer optimum image corrections. This is achieved by using many individual lenses. However, objectives for fluorescence should provide high transmission values into the UV range. This is achieved by minimizing the number of individual lenses. Therefore, there is always a trade-off between image correction and high transmission. An additional problem is that some lenses and immersion liquids autofluoresce, producing a brightened background and low contrast.

IMAGE RESOLUTION AND THE POINT-SPREAD FUNCTION (PSF)

A fluorescence microscope maps the light emitting from structures in the microscope plane of focus to a two-dimensional image in the microscope image plane. By changing the focus of the microscope, one can map the three-dimensional structures of a specimen to a volume image that is constructed from a stack of two-dimensional images. Thus, the relationship between the image and the actual structure

of the specimen depends on how the microscope is used.

From wave optics, it follows that light emitted from a point on the specimen is not focused to an infinitely small point in the image plane. Rather, passing through the objective lens, the light waves converge and interfere around the image plane to produce a three-dimensional diffraction pattern.

Microscope resolution is limited by the diffraction effects of the finite aperture of the objective. For a well-corrected microscope objective lens, the three-dimensional diffraction pattern is symmetrical about the optical axis. When the three-dimensional diffraction pattern is sectioned in the focal plane, it is observed as the two-dimensional diffraction pattern, or Airy disk. Every point in the microscope is therefore represented by an Airy disk, rather than by a simple small point.

A point-spread function (PSF) is a mathematical representation of the image of a point source. For a diffraction-limited optical system operating in the absence of aberration, the PSF is the Airy disk (see Fig. 14.10.9).

The resolving power of an objective lens can be determined by examining the size of the Airy disk formed by that lens. The size of the Airy disk is a function of the NA of the objective lens and of the wavelength of light. Therefore, both lateral resolution and depth of field (resolution along the axis of observation) are a function of the NA of the objective lens. For a wavelength of light of 546 nm and a 0.65-NA dry lens objective, the depth of field is $\sim 1.3 \mu\text{m}$. For a 1.4-NA oil immersion objective, the depth of field is $\sim 0.29 \mu\text{m}$ (Francon, 1961; Spencer, 1982).

Careful alignment and adjustment are required for high-precision imaging with fluores-

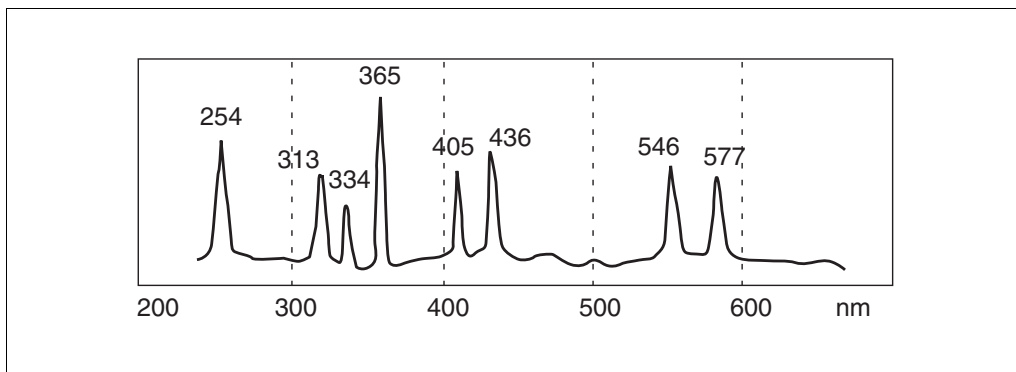


Figure 14.10.8 Line spectrum of a mercury arc lamp. Spectral peaks, the so-called spectral lines, are characteristic of the photoemission from the arcing of electrodes in a high-pressure mercury gas. Courtesy of Omega Optical.

cent probes, particularly when imaging multiple probes simultaneously or when performing three-dimensional reconstruction studies. Fluorescent microspheres, microscope point sources, and image intensity calibration standards are available from Molecular Probes. With video cameras, the shallow focal plane of the camera, the high contrast achievable at high-NA objectives, and the high magnification of the picture displayed contribute to an improved optical section resolution and critically define the focal level of a given structure in the specimen.

Digital image analysis systems, consisting of a microscope fitted with a charge-coupled device (CCD) camera and a video digitizer card installed on a personal computer, are commonly sold as packages from specialized companies (e.g., Scanalytics, Universal Imaging) or from the major microscope companies (e.g., Leica, Nikon, Olympus and Zeiss). An excellent, widely used image analysis software package, NIH Image, is available over the Internet at <ftp://zippy.nimh.nih.gov/pub/nih-image>. The reader is also directed to the Confocal Microscopy Resources on the WWW directory (http://www.pharm.arizona.edu/centers/tox_center/swehsc/exp_path/conf_www.html). The microscope stage can be controlled manually by a stepper motor or by a piezoelectric device, and its position automatically controlled by the computer. Such systems can automatically take stacks of images from parallel planes in the specimen and digitally reconstruct the three-dimensional organization of the specimen. Figure 14.10.10 shows the result of digital reconstruction of a blurry microscope image (a labeled smooth muscle cell) to a sharp image using deconvolution computer algorithms.

FLUORESCENCE MICROSCOPY OF LIVING CELLS

A variety of biologically inert fluorescent probes have been developed to be used as tracers of cell morphology or location. They have been used to investigate neuronal cell connectivity, to study dye translocation through gap junctions, and to track movements of cells in culture, tissue, or intact organisms.

Some fluorescent dyes are able to permeate living cells. Others, like fluorescein, permeate cells selectively. This property has been exploited, for example, in the study of primary sensory neurons of *Caenorhabditis elegans* (Perkins et al., 1986). Still other dyes, like the calcium indicator Fluo-3, are excluded by cells. A common method of conferring plasma membrane permeability to fluorescent dyes and other cell biology reagents is to covalently couple them to hydrophobic groups, the acetoxymethyl (AM) moiety, for example, through an ester linkage. The AM groups are cleaved by endogenous phosphodiesterases, effectively trapping the dye in the cell.

Some dyes are too polar to diffuse passively through cell membranes and require special methods for loading such as microinjection or transient permeabilization techniques. Some dyes, when injected into or applied to cells, produce subtle or sometimes quite apparent changes in cellular organization. An extensive array of literature is available and should be consulted to find procedures suitable for each experiment. A useful handbook of fluorescent reagents and applications is available from Molecular Probes (Haugland et al., 1996).

Photostability is an important concern when using fluorescent probes for imaging or for making fluorescence measurements. Under

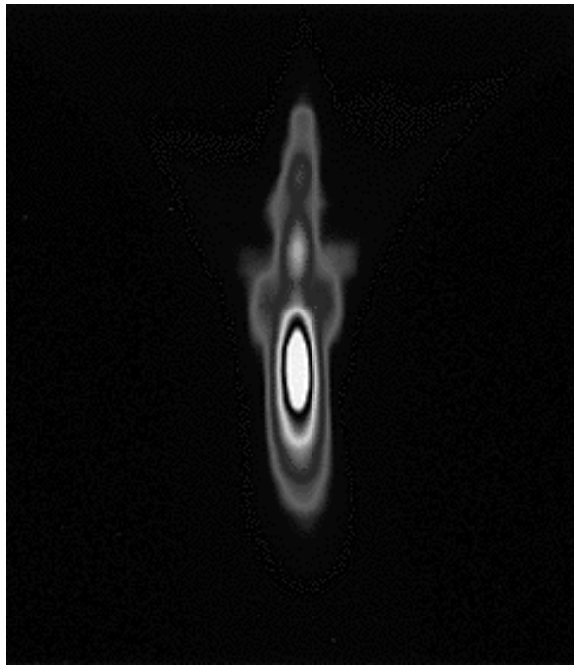


Figure 14.10.9 Orthogonal view of a point-spread function (PSF). A 0.2- μm fluorescent microsphere was embedded in a thin gel, and a stack of images in the x,y plane was collected. The microsphere and surrounding Airy disk were reconstructed in three dimensions. The figure shows the intensity pattern in the x,z plane that passes through the center of the microsphere. The PSF shows how a simple microsphere is mapped to a three-dimensional complex volume image. The PSF is strongly influenced by how the microscope is used. Significant distortions can arise from intrinsic limitations such as the finite aperture of the objective and from extraneous imperfections in the microscope optics. Courtesy of Scanalytics. *This black and white facsimile of the figure is intended only as a placeholder; for full-color version of figure go to <http://www.currentprotocols.com/colorfigures>*

some conditions, fluorescent dyes become very sensitive to photobleaching. Moreover, stability varies significantly among dyes. Exposure to light can cause photobleaching. Therefore, one should avoid illuminating the specimen when not observing or recording a signal, and one should maximize collection of the fluorescence by using high-NA objectives, low-magnification, high-quality optical filters optimal for the spectral properties of the fluorophore, and high-speed film or sensitive video cameras. Antifade reagents added to the mounting medium (see discussion of Immunolabeling, section on Mounting) can help reduce photobleaching on fixed cell preparations but are not compatible with living cells.

IMMUNOLABELING: GENERAL STEPS FOR LABELING FIXED CELLS AND TISSUES

Several excellent reference books are devoted to immunocytochemistry. The most often cited, and still quite applicable, is Sternberger (1979). Other sources include Pawley (1995) and Rost (1992). All too often, contemporary texts focus on technological aspects of imaging and image processing and not on sample preparation. Therefore, a very basic discussion is included here. This discussion also complements UNIT 14.6, which provides a thorough discussion of immunohistochemistry techniques.

Commercial Antibody Selection

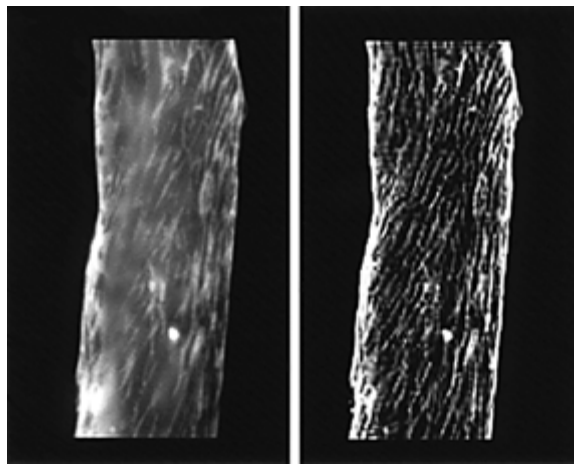
Linscott's Directory of Immunological and Biological Reagents (see Literature Cited), updated quarterly, is a useful database of commercial sources of monoclonal and polyclonal antibodies with 40,000 different products and reagents. Printed and floppy disk versions are available. This resource is an invaluable aid in the face of the growing number of large and small companies that supply antibodies.

Selection and Preparation of Starting Material

Fluorescence microscopy is a useful tool for the study of living tissues, of isolated living cells, and of fixed material, including cells, tissues, and living organisms. Preparation and fixation steps can be as diverse as nature itself. This section is focused on fluorescence microscopy of cultured animal cells, which constitutes the vast majority of applications in the literature. After fixation, subsequent steps for labeling and visualization are essentially the same as might be used for any fixed specimen.

For tissue culture applications, cells are generally grown on glass coverslips. In a sterile hood, coverslips are sterilized with 70% (v/v) ethanol and flame. Cells may be grown directly on the glass or on glass coated with extracellular matrix proteins. The most popular coating agents are collagen, fibronectin, and laminin. References and protocols for the coating of

Figure 14.10.10 Image of a smooth muscle cell labeled with antibodies against vinculin, before and after digital reconstruction. On the basis of the PSF digital image, restoration microscopy techniques use deconvolution computer algorithms to reassign locations for out-of-focus photons to produce a sharp image of the object (Carrington et al., 1995). The computer initially learns how the microscope optics transform a pointlike bead into the complex PSF, and then during normal imaging operation it reassigns all out-of-focus photons to their correct xyz points of origin. Courtesy of Scanalytics. Image acquired and processed by Dr. Edwin Moore, Biomedical Imaging Center, University of Massachusetts Medical School, Worcester.



coverslips may be readily obtained from suppliers (e.g., Life Technologies, Sigma).

Rinsing Cells (for Tissue Culture)

The rinsing step removes debris and serum components that may mask certain epitopes as a result of fixation. For many antigens, elimination of this step is possible without consequence. Most protocols use phosphate-buffered saline (PBS) at 37°C for rinsing. Dulbecco's PBS (Dulbecco and Vogt, 1954) with added calcium (Life Technologies) is a sensible choice for two reasons. First, calcium stabilizes cell-cell contact and membrane structure. Second, Dulbecco's PBS mimics extracellular fluid and various culture media more closely in the ionic composition of electrolytes than does PBS. Alternatives are fresh medium for debris removal and serum-free medium when serum proteins may interfere with antigenicity. To preserve cell morphology, it is important to follow temperature constraints, as microtubules may rapidly depolymerize at temperatures lower than cell culture conditions.

Fixation

The choice of fixation method will depend on the structure or protein being labeled. A common practice with tissue culture is to use 2% to 4% (w/v) formaldehyde in PBS. The fixative is warmed to 37°C and applied to coverslips immediately after rinsing. Usual incubation conditions are 10 min at room temperature. A stock formaldehyde solution may be prepared from the crystalline polymer paraformaldehyde as an 8% (w/v) aqueous solution

and stored at 4°C for up to 1 week. (**CAUTION:** Formaldehyde vapors are toxic. Preparation of formaldehyde stock solutions is discouraged unless large volumes are needed—e.g., for cardiac perfusion. Paraformaldehyde is insoluble in water at room temperature. Preparation requires heating to 80°C in a fume hood, tedious pH adjustment with solid sodium hydroxide pellets, and filtering.) For fixation of tissue culture cells and local perfusion of tissues, readily available commercial 16% (w/v) formaldehyde solutions are recommended. As an alternative fixation procedure, rinsed coverslips may be quickly immersed in -20° to -30°C acetone or methanol with 1% to 2% (w/v) formaldehyde. This fixative is especially useful when the question of artifactual reorganization of soluble antigen pools is being considered. Acetone and methanol fixatives are prepared by adding the appropriate volume of a 16% (w/v) aqueous formaldehyde stock to the organic solvent. Most protocols specify the use of this fixative at -20°C. In practice, it is easy to achieve -30°C by setting 10-ml glass coverslip containers on the surface of dry ice.

Wash

After fixation, further manipulations are carried out at room temperature. Three washes in PBS are normally performed to remove fixative and to rehydrate samples fixed in organic solvents.

Permeabilization

Permeabilization may be done as a separate step following fixation or by combining deter-

**In situ
Hybridization
and Immuno-
histochemistry**

14.10.9

gent with an aqueous fixation step. Omission of a detergent permeabilization step is often used to assess whether an epitope is located within a cell or on the extracellular surface. Usually 0.3% to 1% (v/v) Triton X-100 is used. For tissues, the authors suggest the use of 1% (v/v) Triton X-100 in PBS for a 30-min incubation followed by three 10-min washes in PBS. For monolayers of tissue culture cells fixed with formaldehyde in PBS, 0.3% (v/v) Triton X-100 is included with the fixative. For cells fixed with organic solvents, no further permeabilization is required because the solvent extracts lipids from the bilayer of the plasma membrane.

Blocking

Nonspecific binding can be blocked using either bovine serum albumin (BSA) or serum. The authors conservatively use both 2% (w/v) BSA and 5% (v/v) goat serum for 1 hr at room temperature or overnight at 4°C.

Primary Antibody

Primary antibodies are diluted in blocking solution and used to probe tissues and cells. The appropriate dilution depends on antibodies and antigens. If a supplier recommends a particular dilution, say, 1:200, then it is best to initiate experiments by bracketing that dilution (i.e., 1:100, 1:200, and 1:1000). To remove aggregates, solutions are centrifuged 10 min at $14,000 \times g$ just before use. For tissue culture cells, one can typically place coverslips in a moist chamber, face up, on Parafilm. The Parafilm helps to keep primary and secondary antibody solutions on the coverslip and aids in transferring coverslips to six-well dishes for washing. A 22×22 -mm coverslip can be covered with 150 μ l of antibody solution. Incubation times vary from 1 to 3 hr. Alternatively, coverslips may be placed, face down, on Parafilm on 30- to 50- μ l droplets. The surface tensions that develop in this procedure are very large. Thus, although this method results in satisfactory images for most antibodies, its use should be limited to cases of extreme shortage of antibody.

Wash

The intermediate wash step is more critical than previous ones. For high-quality images, use five 10-min washes in PBS. For routine screening of monoclonal antibodies, for example, this step can be reduced to three 5-min washes.

Secondary Antibody

Secondary antibodies coupled to fluorophores must be screened for titer with each new batch or lot. Dilutions are made in blocking serum, followed by centrifugation. Centrifugation is especially important for fluorescein isothiocyanate (FITC)-coupled antibodies to reduce particulate background. Conditions for centrifugation are the same as for primary antibody. The supernatant should be carefully transferred to a fresh tube without disturbing the pellet. Titers can usually be adjusted for a 1-hr room temperature incubation.

Final Wash

The final wash step is as critical as the previous wash. The same considerations apply.

Mounting

A number of antifade mounting solutions can be made at the bench or purchased. One that is particularly easy to prepare and use is made by mixing 10 mg *p*-phenylenediamine (Sigma) with 1 ml PBS and 9 ml glycerol. Store the solution in aliquots at -20°C in the dark. Discontinue using when a ruby red color develops. Four dots of nail polish, one at each corner of the coverslip, or four dots of High Vacuum Grease (Dow Chemical), one under each corner, will suffice to hold the coverslip in place. The silicone-based grease may be loaded into a syringe equipped with a wide-bore needle for easy application.

Many other solutions are available to delay fading of fluorescent dyes. The commercial mounting medium ProLong (Molecular Probes) is recommended. In addition to an antifade reagent, the product contains a hardening agent that circumvents the need for nail polish or grease. It works well with cells or tissue. It requires a final water wash and more careful handling. However, ProLong gives results that are noticeably superior to phenylenediamine/glycerol solutions.

Double Labeling

Two basic methods may be used for double labeling. The two primary antibodies are typically generated in mouse and rabbit. They may, however, be from any two distinct host species, so long as there is no cross-reaction of respective secondary antibodies. In the first method, the processing of the first primary antibody and its respective secondary antibody is completed, and then the second primary antibody and its respective secondary antibody are processed separately. The time for the complete operation

is effectively doubled. This method offers the advantage that interactions between the two primary antibodies are avoided. By the same logic, it is absolutely necessary to prepare controls to ensure that the pattern of labeling is not dependent on the order of binding of the two primary antibodies.

An alternative method is to mix the two primary antibodies and two secondary antibodies together. This method produces satisfactory results for many pairs of antibodies and offers significant time savings. Additionally, it provides an opportunity to test for specificity in preadsorption controls (see below). Because of the possibilities of steric hindrance or other interactions in the binding of antibodies to closely situated epitopes, any double-labeling experiments must always be accompanied by controls using each primary antibody separately. This is especially important when one of the antibodies is monoclonal, a typical situation.

Controls

Mixing antigen, in the form of protein or synthetic peptide, with primary antibody provides the best control for specificity. This so-called preadsorption control may also be accomplished in the solid phase by coupling antigen to chromatography resins or by blotting large amounts of antigen on nitrocellulose. Solid-phase adsorption controls are especially useful with antigens such as myosins that normally require high salt concentrations for solubilization. Some protocols call for overnight incubation. However, complete blocking may be obtained after a 1-hr incubation of anti-peptide antibodies with 1 mg/ml immunizing peptide, whereas the same concentration of irrelevant peptide has no effect. A good way to implement this control is to titrate the antibody with antigen and combine the fluorescence microscopy results with immunoblotting. If titration is not performed, it is still important to show that the antigen is not so concentrated that it is blocking sample labeling by nonspecific mechanisms. This may be accomplished by performing double labeling with two primary antibodies simultaneously. The requirement is

that the antigen specifically block labeling by the appropriate antibody only. Other control conditions introduced at this step are the use of preimmune serum, serum from the same host, or blocking solution alone.

LITERATURE CITED

- Carrington, W.A., Lynch, R.M., Moore, E.D.W., Isenberg, G., Fogarty, K.E., and Fay, F.S. 1995. Superresolution three-dimensional images of fluorescence in cells with minimal light exposure. *Science* 268:1483-1487.
- Dulbecco, R. and Vogt, M. 1954. Plaque formation and isolation of pure lines of poliomyelitis viruses. *J. Exp. Med.* 98:167-182.
- Francon, M. 1961. Progress in Microscopy. Row, Peterson, Evanston, Ill.
- Haugland, R.P., Spence, M.T.Z., and Johnson, I.D. 1996. Handbook of Fluorescent Probes and Research Chemicals. Molecular Probes, Eugene, Ore.
- Linscott's Directory of Immunological and Biological Reagents (updated quarterly). W.D. Linscott, Santa Rosa, Calif.
- Pawley, J.B. (ed.) 1995. Handbook of Biological Confocal Microscopy. Plenum, New York.
- Perkins, L.A., Hedgecock, E.M., Thomson, J.N., and Culotti, J.G. 1986. Mutant sensory cilia in the nematode *Caenorhabditis elegans*. *Dev. Biol.* 117:456-487.
- Rost, F.W.D. 1992. Fluorescence Microscopy. Cambridge University Press, Cambridge.
- Spencer, M. 1982. Fundamentals of Light Microscopy. Cambridge University Press, Cambridge.
- Sternberger, L. 1979. Immunocytochemistry. John Wiley & Sons, New York.

KEY REFERENCE

- Inoue, S. and Spring, K.R. 1997. Video Microscopy. Plenum, New York.
- General reference for interested readers.*

Contributed by Donald Coling
University of California Berkeley
Berkeley, California

Bechara Kachar
National Institute on Deafness and
Other Communication Disorders
Bethesda, Maryland

Confocal microscopy produces sharp images of structures within relatively thick specimens (up to several hundred microns). It is particularly useful for examining fluorescent specimens. Thick fluorescent specimens viewed with a conventional widefield fluorescent microscope appear blurry and lack contrast because fluorophores throughout the entire depth of the specimen are illuminated and fluorescence signals are collected not only from the plane of focus but also from areas above and below. Confocal microscopes selectively collect light from thin (~1- μm) optical sections representing single focal planes within the specimen. Structures within the focal plane appear more sharply defined than they would with a conventional microscope because there is essentially no flare of light from out-of-focus areas. A three-dimensional view of the specimen can be reconstructed from a series of optical sections at different depths.

The confocal microscope is the instrument of choice for examining fluorescence-stained cells in tissue slices or small, intact organisms such as *Drosophila* (Fig. 14.11.1A,B) and zebrafish embryos. It is also useful for localizing fluorescent-tagged molecules in dissociated cells (Fig. 14.11.1C,D). Its sensitivity even allows fluorescence in living specimens to be monitored, making it feasible to follow the movements in living cells of fluorescent probes such as the green fluorescent protein (GFP; Fig. 14.11.1D). In addition, some types of confocal microscopes can be configured to perform photobleach experiments (Fig. 14.11.1D) and to photoactivate “caged” molecules (molecules that are inactive until released with UV illumination).

Biologists use confocal microscopy in a number of creative ways that are beyond the scope of this article. The information presented herein is intended to provide background and practical tips needed to get started with confocal microscopy. An excellent source of theoretical and technical information is the *Handbook of Biological Confocal Microscopy* (1995; edited by J. Pawley). Also recommended are *Cell Biological Applications of Confocal Microscopy* (1993; edited by B. Matsumoto), a good source of practical information; *Confocal Microscopy* (1990; edited by T. Wilson), for theoretical background; and *Video Microscopy* (1997; Inoue and Spring) for fundamentals of microscopy.

THE BASIS OF OPTICAL SECTIONING

Confocal microscopes accomplish optical sectioning by scanning the specimen with a focused beam of light and collecting the fluorescence signal from each spot via a spatial filter (generally a pinhole aperture) that blocks signals from out-of-focus areas of the specimen. The physical basis of optical sectioning in fluorescence confocal microscopy is illustrated in Figure 14.11.2. A point light source (typically a laser) evenly illuminates the back focal plane of the objective, which focuses the light to a diffraction-limited spot in the specimen. The irradiation is most intense at the focal spot, although areas of the specimen above and below the focal spot also are illuminated. Fluorescent molecules excited by the incident light emit fluorescence in all directions. The fluorescence collected by the objective comes to focus in the image plane, which is conjugate (confocal) with the focal plane in the specimen. A pinhole aperture in the image plane allows fluorescence from the illuminated spot in the specimen to pass to the detector but blocks light from out-of-focus areas.

The diameter of the pinhole determines how much of the fluorescence emitted by the illuminated spot in the specimen is detected, and the thickness of the optical section. From wave optics we know that a point light source in the plane of focus of an objective produces a three-dimensional diffraction pattern in the image plane. The cross section at the image plane is an Airy disk (see Fig. 14.10.9), a circular diffraction pattern with a bright central region. The radius of the bright central region of the Airy disk in the reference frame of the specimen is given by

$$R_{\text{Airy}} = 0.61\lambda/\text{NA}$$

where λ is the emission wavelength and NA is the numerical aperture of the objective (see UNIT 14.10 for a discussion of NA). At the image plane (the location of the pinhole aperture), the radius of the central region is R_{Airy} multiplied by the magnification at that plane (for a more complete explanation see Wilson, 1995).

Adjustment of the pinhole to a diameter slightly less than the diameter of the central region of the Airy disk allows most of the light from the focal point to reach the detector and reduces the background from out-of-focus areas by ~1000-fold relative to widefield micros-

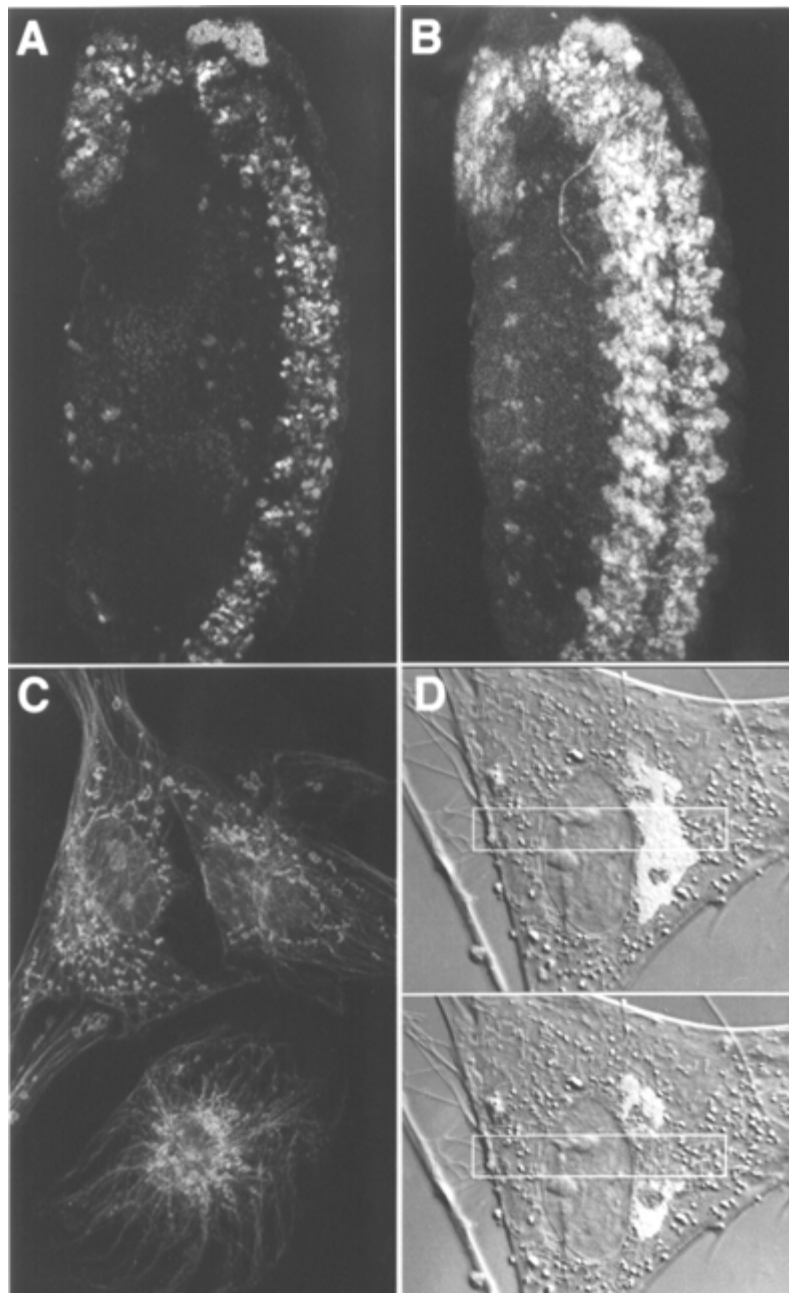


Figure 14.11.1 Applications of laser scanning microscopy. (A,B) (From W. Oldenwald; see Kamabadur et al., 1998.) 3-D analysis of thick specimens. Different neuronal populations of an $\sim 250\text{-}\mu\text{m}$ -thick *Drosophila* embryo were immunolabeled with antibodies against three transcription factors. A, $\sim 2.5\text{-}\mu\text{m}$ optical section collected with $25\times$, 0.8-NA objective using a detector pinhole diameter of ~ 1.3 Airy units. Labeled neurons in the plane of focus appear sharply defined, while those outside it are not visualized. B, projection (superimposition) of 65 optical sections collected at $2\text{-}\mu\text{m}$ intervals in the z axis. Neurons at different focal planes appear to overlap in this flattened image, but are distinct in a 3-D reconstruction. (C) Localization of intracellular structures. Dissociated rat fibroblasts were immunolabeled with anti-tubulin antibodies to visualize microtubules (green) and stained with fluorescent probes for mitochondria (Mitotracker, red) and DNA (DAPI, blue). The image is a projection of 20 optical sections collected at $0.3\text{-}\mu\text{m}$ intervals in the z axis with $100\times$, 1.4-NA objective. (D) Measuring molecular motility. In a living fibroblast expressing a Golgi membrane protein (galactosyltransferase) fused to GFP (S65T-GFP), GFP fluorescence (green) localized to the Golgi complex, shown superimposed on a DIC image of the cell. After the first image was collected, the boxed region (yellow) was scanned with full laser power; this photobleached the GFP in the boxed area as shown in the second image collected ~ 2 sec later. The rate of fluorescence recovery into the photobleached zone (not illustrated) indicated that GFP-galactosyltransferase fusion is highly mobile in Golgi membranes. *This black and white facsimile of the figure is intended only as a placeholder; for full-color version of figure go to <http://www.currentprotocols.com/colorfigures>*

copy (Sandison et al., 1995). The separation of the in-focus signal from the out-of-focus background achieved by a properly adjusted pinhole is the principle advantage of confocal microscopy for examination of thick specimens (see Fig. 14.11.1A,B).

Point illumination and the presence of a pinhole in the detection light path also produces improved lateral and axial resolution relative to conventional microscopy (Table 14.11.1). The actual extent of improvement depends on the size of the pinhole. Near-maximal axial resolution is obtained with a pinhole radius $\sim 0.7 \times R_{\text{Airy}}$ whereas optimal lateral resolution is obtained with a pinhole less than $0.3 \times R_{\text{Airy}}$ (Wilson, 1995). However, a pinhole smaller than $\sim 0.7 \times R_{\text{Airy}}$ significantly reduces the total signal, a sacrifice that may not be worth the gain in resolution, especially when imaging dim samples. In fluorescence imaging, resolution also is influenced by the emission and excitation wavelengths (Table 14.11.1).

TYPES OF CONFOCAL MICROSCOPES

Several types of confocal microscopes are available, each having unique features and advantages. The types most commonly used for examining fluorescence specimens are laser-scanning confocal microscopes. These microscopes, as their name implies, use lasers as light sources and collect images by scanning the laser beam across the specimen.

Lasers provide intense illumination within a narrow range of wavelengths. The emission wavelengths of several types of lasers, together with the excitation spectra of familiar fluorophores, are illustrated in Figure 14.11.3. Mixed krypton-argon gas lasers are popular for multi-wavelength confocal microscopy because they emit at three well-separated wavelengths (488, 568, and 647 nm) that can be used to simultaneously image two or three fluorophores (e.g., FITC, lissamine rhodamine, and Cy5). The disadvantage of krypton-argon lasers is that

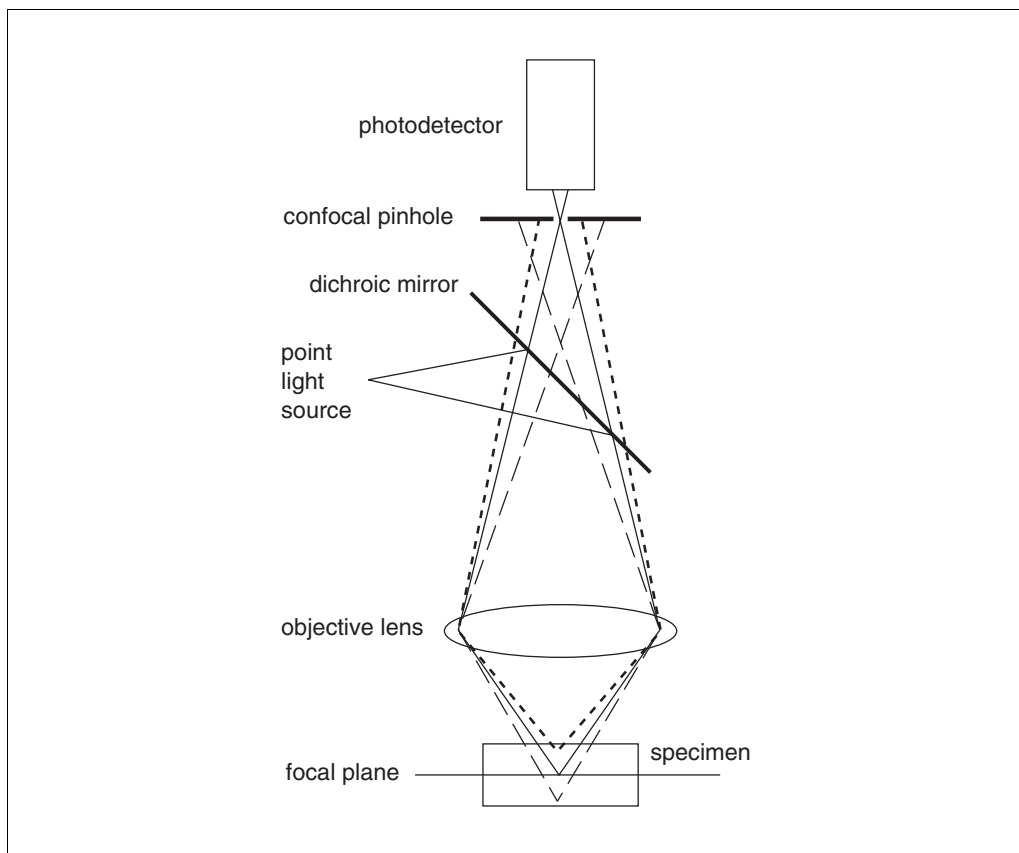


Figure 14.11.2 The basis of optical sectioning in confocal epifluorescence microscopy. Illumination from the point light source is reflected by the dichroic mirror and focused by the objective lens to a diffraction-limited spot within the specimen. Fluorophores within the focal spot as well as in the cone of light above and below it are excited, emitting fluorescence at a longer wavelength than the incident light. The fluorescence captured by the objective passes through the dichroic mirror because of its longer wavelength. The confocal pinhole allows fluorescence from the plane of focus in the specimen to reach the photodetector but blocks fluorescence from areas above and below the plane of focus. Redrawn from Shotton (1993).

Table 14.11.1 Theoretical Resolutions of Confocal and Conventional Microscopes^a

$\lambda_{\text{ex}}/\lambda_{\text{em}}$	Objective					
	10×, 0.4 NA, air		40×, 0.85 NA, air		60×, 1.4 NA, oil	
	Lat. res.	Ax. res.	Lat. res.	Ax. res.	Lat. res.	Ax. res.
<i>Confocal fluorescence microscope</i>						
488/518	0.55	4.50	0.26	0.99	0.16	0.56
568/590	0.64	5.17	0.30	1.09	0.18	0.64
647/677	0.72	5.88	0.34	1.28	0.21	0.72
<i>Conventional fluorescence microscope</i>						
518	0.79	6.48	0.37	1.43	0.24	0.93
590	0.90	7.38	0.42	1.63	0.28	1.06
680	1.04	8.50	0.49	1.88	0.32	1.22

^aData reprinted from Brelje et al. (1993) by permission of Academic Press. λ_{ex} and λ_{em} , excitation and emission wavelengths; lat. res. and ax. res., lateral and axial resolutions.

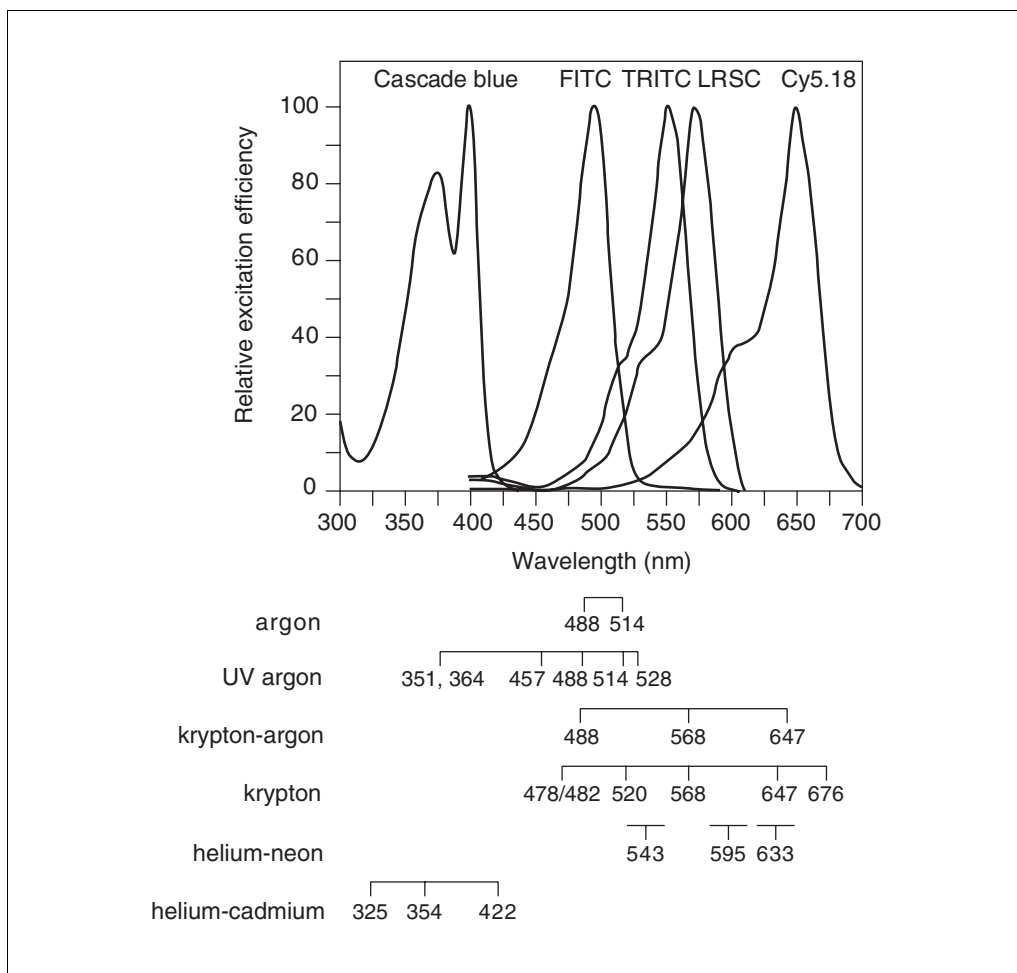


Figure 14.11.3 Comparison of the emission wavelengths of various lasers and the excitation spectra of representative fluorophores. The lasers most commonly used for laser-scanning confocal microscopy are air-cooled argon (488 and 514 nm), krypton-argon, and helium-neon lasers. UV argon lasers generally require water cooling and are more expensive. They may be configured to provide only UV wavelengths (351 nm and 364 nm) or both UV and longer wavelengths. Data for the excitation spectra of Cascade blue, fluorescein (FITC), tetramethylrhodamine (TRITC), lisamine rhodamine (LRSC), and cyanine 5.18 (Cy5.18) are from Wessendorf and Brelje (1993) and were downloaded from the web page of Aryeh Weiss, <http://optics.jct.ac.il/~aryeh/Spectra>. Modified from Brelje et al. (1993).

their life spans are short (~2000 hr). Another way to achieve multiwavelength excitation is to combine the outputs of two or more lasers.

Several methods have been devised for scanning the sample with the laser beam to illuminate different positions in the specimen. The most common method employs a pair of galvanometer mirrors to both scan the laser beam across the specimen and collect the fluorescence emitted from the specimen (Fig. 14.11.4). One galvanometer mirror scans sequential spots along the *x* axis, and the second mirror moves from line to line in the *y* axis. The fluorescence emission is separated from the illuminating beam by a dichroic beam splitter and is directed to a photomultiplier tube which collects the fluorescence produced as each spot in the specimen is illuminated. The photodetector output is converted to a digital image that can be displayed on a monitor and stored as a digital image file for later analysis. Most laser-scanning confocal microscopes have 8-bit digi-

tizers that encode 256 gray levels, although some recent models have 12- or 16-bit digitizers. Collection of a full-size image (typically 1024×1024 pixels) takes ~2 sec. Laser-scanning microscopes that employ galvanometer mirror scanners sometimes are called “slow-scan” microscopes because of their relatively slow image acquisition rates. Slow-scan microscopes are available from several sources (Bio-Rad, Zeiss, Leica, Olympus, Nikon, Molecular Dynamics, and Meridian; see APPENDIX 4).

The movements of the galvanometer mirrors in laser-scanning microscopes are under the control of a computer, providing flexibility in the scanning pattern. For example, it is possible to “zoom” a region of interest (visualize it at higher magnification) by reducing the scan area and the distances between sample points. In addition, many laser-scanning microscopes have the ability to repetitively scan a single line or to “park” the scanner to monitor fluorescence at a single spot. The latter technique is

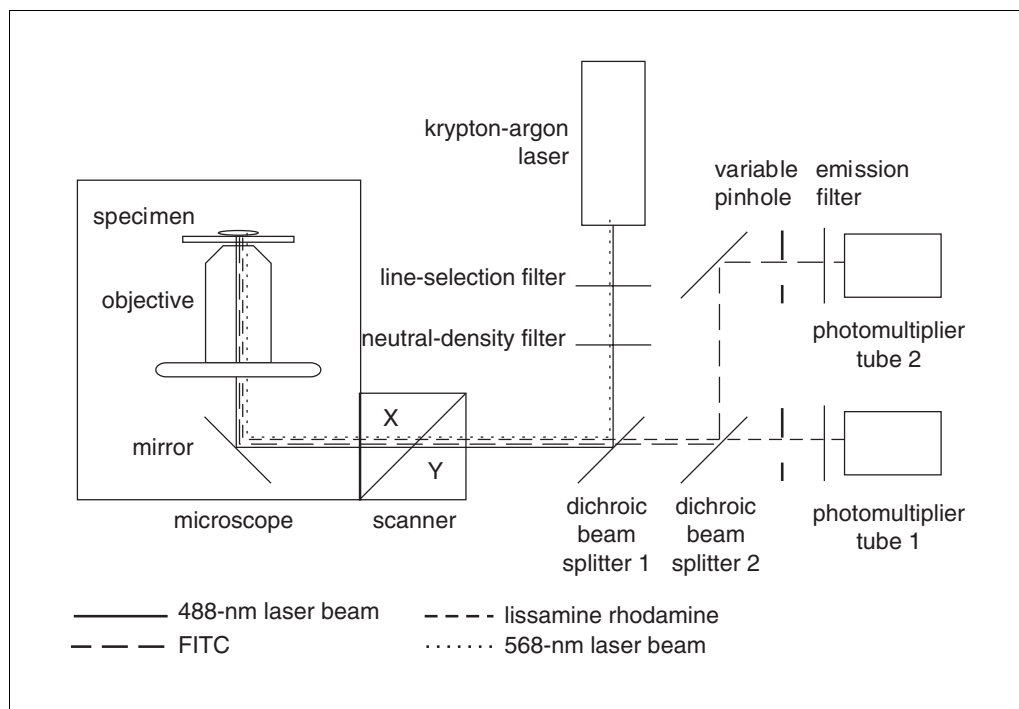


Figure 14.11.4 The light path of a laser-scanning confocal microscope set up for simultaneous imaging of FITC and lissamine rhodamine. The 488-nm and 568-nm lines of a krypton-argon laser are reflected by dichroic beam splitter 1 into the optical axis of the microscope. The scanner contains two galvanometer mirrors, which generate the *x* and *y* axis movements of the beam. The beam is reflected by a mirror into the objective which focuses the beam onto the specimen. The specimen is scanned line by line in a raster pattern. Fluorescence emitted by the specimen as each spot is illuminated travels the reverse path through the scanning system. The FITC fluorescence (peak at 520 nm) and lissamine rhodamine fluorescence (peak at 590 nm) pass through dichroic beam splitter 1 to dichroic beam splitter 2, which transmits the lissamine rhodamine fluorescence to photomultiplier tube 1 and reflects the FITC fluorescence to photomultiplier tube 2. A variable pinhole in front of each photodetector blocks light from out-of-focus areas of the specimen while allowing light from the illuminated spot to reach the detector.

particularly useful for studying rapidly changing fluorescence signals, such as those produced by a Ca²⁺ indicator in an active neuron.

Laser-scanning microscopes are available (from Norau, Life Sciences Resources, and Meridian; see APPENDIX 4) that can collect images at video rates (30 frames/sec) or faster. Several methods for achieving rapid scanning rates have been employed, such as acousto-optical deflection devices, rotating mirrors, or resonating mirrors (reviewed by Art and Goodman, 1993; Tsien and Bacsikai, 1995). The gain in imaging speed always comes at a cost, however. For example, rapid-scan confocal microscopes do not provide the degree of control over the scan pattern offered by top-of-the-line slow-scan microscopes, and some video-rate confocal microscopes are incapable of multiwavelength illumination. Video-rate microscopes that rely on slit apertures rather than pinhole apertures have slightly poorer lateral and axial resolution.

A type of rapid-scan confocal microscope that deserves mention because of its lower cost (among other reasons) uses a spinning disk with multiple pinholes (~200,000) to simultaneously illuminate and detect emission from many spots in the specimen. The light source can be a laser or a broad-spectrum lamp like that used for conventional epifluorescence microscopy. The principle advantage of this type of confocal microscope is that it is capable of collecting images very rapidly (up to 700 frames/sec at 5000 lines resolution; Kino, 1995). The images can be examined directly by eye or captured with a sensitive camera. The main disadvantage is that the disk transmits only ~1% of the available light because the holes in the spinning disk need to be widely spaced. A new type of spinning-disk confocal microscopy has recently become available that uses "microlenses" to improve optical throughput and achieve high-speed confocal imaging with better sensitivity (Ultra View; Life Sciences Resources; see APPENDIX 4).

Another form of laser-scanning microscopy that promises to be of great value uses two-photon (and three-photon) excitation to induce fluorescence emission (Denk et al., 1995). Two-photon excitation occurs when a fluorophore absorbs two photons, each having half the energy needed to raise the fluorophore to the excited state. The light intensities required for simultaneous absorption occur only at the focal point, so only fluorophores at the focal point are excited. Therefore, two-photon excitation allows optical sectioning without a spa-

tial filter in front of the detector. Moreover, since fluorophores outside the focal point are not excited, the specimen is less subject to photobleaching than in a conventional laser-scanning microscope. The wavelengths needed to excite standard visible light fluorophores by two-photon absorption are longer and penetrate tissue better than the wavelengths used for one-photon excitation, making it possible to look deeper into a specimen. In addition, UV fluorophores can be imaged without many of the problems that arise when UV wavelengths are used in conventional laser-scanning microscopes. A current drawback of two-photon confocal microscopy is the high cost of an appropriate laser (~\$100,000). Two-photon scanning microscopes are now available from commercial sources (Bio-Rad, Leica; see APPENDIX 4).

PRACTICAL GUIDELINES

Sample Preparation: Immunofluorescence in Fixed Specimens

Additional guidelines for sample preparation are discussed in UNITS 14.6 & 14.10.

Fixation

The best fixative is one that accurately preserves the three-dimensional geometry of the specimen. The standard fixative for fluorescence microscopy (2% to 4% formaldehyde in PBS) is not ideal because it can cause blebbing of the plasma membrane, vesiculation of intracellular membrane compartments, and other alterations in cellular morphology. Moreover, some commercial preparations of formaldehyde contain methanol, which shrinks cells. Techniques for optimizing formaldehyde fixation are described by Bacallao et al. (1995; also see UNITS 14.6 & 14.10). The buffer should be chosen to match the osmolality and pH of the specimen. Fixatives containing 0.125% to 0.25% glutaraldehyde in addition to formaldehyde preserve cellular morphology better than those containing formaldehyde alone. Some investigators avoid using glutaraldehyde for fluorescence microscopy because it induces autofluorescence. However, autofluorescence can be reduced by treating the sample after fixation with NaBH₄ (1 mg/ml in PBS, pH 8.0, using two treatments of 5 min each for dissociated cells, longer for thicker samples). A more serious drawback of glutaraldehyde for immunofluorescence studies is that it destroys the antibody recognition sites of some antigens. An alternative fixation technique that preserves tissue

better than chemical fixation is rapid freezing followed by freeze substitution (Bridgman and Reese, 1984).

Choices of fluorophores

Criteria to consider in selecting fluorophores for fluorescence microscopy are described in *UNITS 14.6 & 14.10*. The only additional consideration for confocal microscopy is to choose fluorophores that can be excited by the wavelengths provided by the available lasers. However, it is not essential for the excitation spectrum peak to precisely match the laser wavelength because the lasers on most microscopes are sufficiently powerful to maximally excite fluorophores at off-peak wavelengths. For experiments that depend on imaging two fluorophores, it is best to select fluorophores whose excitation and emission spectra have minimal overlap. Good choices for multiwavelength imaging with a krypton-argon laser are: FITC/Oregon green/Alexa 488 (Molecular Probes) for excitation at 488 nm; lisamine rhodamine/Cy3/Texas red/Alexa 568 (Molecular Probes) for excitation at 568 nm; and Cy5 for excitation at 647 nm. UV fluorophores also are good for multicolor imaging (with absorption at 350 to 390 nm; some of the best dyes for DNA are UV fluorophores).

Control samples

Confocal microscopes rely on electronic image enhancement techniques that can make even a dim autofluorescence signal or nonspecific background staining look bright. In order to be able to distinguish a real signal from background it is essential to prepare appropriate control samples. For immunofluorescence experiments with one primary antibody, the appropriate control samples are unstained specimens and specimens treated with the secondary antibody but no primary antibody. Experiments with two primary and secondary antibodies require additional controls to test whether the secondary antibodies cross-react with the “wrong” primary antibody. Other control experiments may be required to verify the specificity of labeling (see *UNIT 14.10*).

Mounting the specimen

The mounting medium should preserve the three-dimensional structure of the specimen. PBS (*APPENDIX 2*) or a mounting medium consisting of 50% glycerol/50% PBS preserves the shapes of cells quite well, but Mowiol and gelvatol cause a 10% decrease in height (Bacallao et al., 1995). Adding an antioxidant to

the mounting medium helps to alleviate photobleaching. One of the best antioxidants is 100 mg/ml 1,4-diazabicyclo[2,2,2]octane (DABCO; Sigma; Bacallao et al., 1995). *n*-propyl gallate (Giloh and Sedat, 1982) and *p*-phenylenediamine (PPD; Johnson et al., 1982) also are effective antibleaching agents, but the former may cause dimming of the fluorescence while the latter may damage the specimen (Bacallao et al., 1995).

The choice of mounting medium should take into account the type of microscope objective that will be used to observe the specimen. In order for an objective to perform optimally, the mounting medium should have the same refractive index as the objective immersion medium. Table 14.11.2 gives the refractive indexes of standard objective immersion media and mounting media. Mismatches in the refractive indexes produce spherical aberration leading to loss of light at the detector, as well as decreased *z* axis resolution and incorrect depth discrimination. Image deterioration caused by spherical aberration increases with depth into the specimen. Significant losses of signal intensity and axial resolution are apparent at distances of just 5 to 10 μm when an oil immersion objective is used to examine a specimen in an aqueous medium (Keller, 1995).

Most microscope objectives are designed for viewing specimens through a glass coverslip of a specific thickness (typically 0.17 μm , a no. 1½ coverslip). Correct coverslip thickness is especially critical for high-NA (>0.5) dry objectives and water immersion objectives (Keller, 1995). Use of a coverslip that differs from the intended thickness by only 5% causes significant spherical aberration. High-NA dry and water immersion objectives typically have an adjustable collar to correct for small variations in coverslip thickness.

The specimen should be mounted as close to the coverslip as possible, especially for observation with immersion objectives, which have short working distances (~100 to 250 μm , depending on the type of objective). This also helps to avoid image deterioration due to spherical aberration. Fragile specimens should be protected by supporting the coverslip; for example, using a thin layer of nail polish, strips of coverslips, or a gasket made from a sheet of silicon rubber (Reiss; see *APPENDIX 4*). Sealing the edges of the coverslip—with nail polish or silicon vacuum grease (Dow Corning; see *APPENDIX 4*)—helps to prevent specimen desiccation and movement.

Table 14.11.2 Refractive Indexes of Common Immersion and Mounting Media

Medium	Refractive index (RI)
<i>Immersion media</i>	
Air	1.00
Water	1.338
Glycerol	1.47
Immersion oil	1.518
<i>Mounting media</i>	
50% glycerol/PBS/DABCO	1.416 ^a
5% <i>n</i> -propyl gallate/0.0025% <i>p</i> -phenylene diamine (PPD) in glycerol	1.474 ^a
0.25% PPD/0.0025% DABCO/5% <i>n</i> -propyl gallate in glycerol	1.473 ^a
VectaShield (Vector Labs)	1.458 ^a
Slow Fade (Molecular Probes)	1.415 ^b
Prolong (Molecular Probes)	1.3865 ^{b,c}

^aData from Bacallao et al. (1995).

^bData from Molecular Probes.

^cRI for liquid medium (RI for solidified medium will be higher).

Living Specimens

Confocal microscopy of living preparations is challenging for several reasons. The specimen must be mounted in a chamber that keeps it healthy and immobile while at the same time providing access for the objective. For high-resolution transmitted-light imaging (e.g., by laser-scanning differential interference contrast microscopy), the chamber must be thin enough to accommodate a high-NA (oil immersion) condenser. Fluorescence signals in living specimens generally are weak and the illumination levels needed to detect them can be damaging to the specimen. Photobleaching inevitably is a problem for experiments that require collecting many images. Temperature fluctuations in specimens kept at nonambient temperatures make it difficult to maintain accurate focus.

A simple chamber for culture preparations grown on glass coverslips can be made by forming a well on a glass slide with a gasket cut from a sheet of silicon rubber or a plastic ruler. To prevent the well from leaking, it should be sealed with silicon vacuum grease, a mixture of melted paraffin and petroleum jelly, or Sylgard (Dow Corning; see APPENDIX 4). The well is filled with medium and then the coverslip with attached cells is placed, cell side down, on top of the well. The preparation can be kept warm during obser-

vation on the microscope with a heated air blower—e.g., a hair dryer with variable power source or a commercial air-stream incubator (e.g., Neutek; see APPENDIX 4)—or with infrared lamps. More elaborate chambers, some of which have built-in heaters and ports for changing solutions, are available from commercial sources (see Terasaki and Dailey, 1995, for a partial listing of manufacturers). An important factor to consider in choosing a chamber is whether it maintains the desired temperature while in contact with an immersion objective that acts as a heat sink. One solution to this problem is to heat the objective as well as the chamber. A heated chamber and objective warmer designed for microscopy with a high-NA objective and condenser are available from Bioprotechs (see APPENDIX 4).

Addition of an oxygen quencher to the medium can help to alleviate photobleaching of the fluorophores. Photobleaching not only leads to dimming of the signal but also to generation of oxygen radicals that can damage cells. Several oxygen quenchers have been reported to be effective, including oxyrase (0.3 U/ml; Oxyrase [see APPENDIX 4]; Waterman-Storer et al., 1993); ascorbic acid (0.1 to 3.0 mg/ml; Sigma; Terasaki and Dailey, 1995); a mixture of Trolox (10 μM; Aldrich) and *N*-acetylcysteine (50 μM; Sigma; M. Burack and G.

Banker, pers. comm.); and crocetin (Tsien and Waggoner, 1995).

Optimizing Imaging Parameters

Choice of objectives

High-NA objectives generally are preferable for fluorescence microscopy because they collect more light than low-NA objectives (brightness is proportional to NA^4). Most high-quality high-NA objectives have >80% transmission at visible wavelengths, but some have low transmission at UV wavelengths (Keller, 1995).

Water immersion objectives are the best choice for visualizing specimens in aqueous solutions (e.g., living specimens). Several microscope manufacturers recently have introduced high-NA water immersion objectives specifically designed for confocal microscopy of biological specimens. These objectives differ from previously available types of water immersion objectives in that they are intended for viewing specimens mounted under a coverslip. They have working distances of $\sim 250 \mu\text{m}$.

Oil immersion objectives can have higher NAs than water immersion objectives. Most have fairly short working distances ($\sim 100 \mu\text{m}$) although some recently introduced oil objectives have working distances of $\sim 200 \mu\text{m}$. A long-working-distance oil objective will be useful only if the specimen is mounted in a medium that matches the refractive index of immersion oil ($\eta = 1.518$). If an aqueous mounting medium is used, images from depths at more than $\sim 20 \mu\text{m}$ into the specimen will be noticeably degraded by spherical aberration. Also, distance measurements in the z axis will need to be corrected. The actual movement of the focal plane in the specimen (d_s) produced by a movement of the objective (d_{obj}) depends on the ratio of the refractive indexes of the specimen and immersion medium. A reasonable approximation (Majilof and Forsgren, 1993) of the relationship is given by:

$$d_s/d_{\text{obj}} = \eta_s/\eta_{\text{obj}}$$

Pinhole size

As was explained above (see Basis of Optical Sectioning), the size of the pinhole has a critical influence on image quality. A pinhole with a radius equal to the radius of the first minimum of the Airy disk—which is approximately equivalent to the diameter at half maximal intensity (Amos, 1995)—will let most of the light from the plane of focus reach the

detector, while blocking most of the out-of-focus flare. The lateral resolution will be $\sim 20\%$ better than that obtainable by conventional microscopy with the same optics (Centonze and Pawley, 1995), although not as good as can be achieved with a smaller pinhole. Lateral resolution continues to improve as pinhole radius is decreased down to a pinhole size of $\sim 0.2 \times$ Airy disk radius, but a pinhole this small excludes $\sim 95\%$ of the signal (Wilson, 1995). Axial resolution improves as pinhole size decreases, down to $\sim 0.7 \times$ Airy disk radius, then levels off. The best trade-off between signal intensity and resolution will depend on the characteristics of the sample and aims of the experiment.

Zoom factor

The zoom setting on a confocal microscope determines the size of the scan region and the apparent magnification of the image. A zoom factor of 2 will scan an area half as long and wide as a zoom factor of 1. Images are made up of the same number of samples (points along the horizontal axis, lines along the vertical axis) and are displayed on the image monitor by a fixed number of pixels regardless of the zoom factor. Therefore, the pixels in a zoom-2 image will represent areas within the specimen half as large in each dimension as the areas represented by the pixels at zoom 1. If the pixel size for an objective at zoom 1 represents $0.25 \mu\text{m} \times 0.25 \mu\text{m}$, then the pixel size at zoom 2 will be $0.125 \times 0.125 \mu\text{m}$. The pixel dimensions (referring to the specimen) are inversely related to the zoom setting.

For each objective, there is an optimal zoom setting which yields pixel dimensions small enough to take advantage of the full resolution of the objective but large enough to avoid oversampling. In order for the minimum resolvable entity to be visible on the display monitor, the pixel dimensions need to be smaller than (less than one-half) the optical resolution. However, if the pixel size is made too small by using a higher-than-optimal zoom factor, the specimen is subjected to more irradiation than necessary with an increased risk of photobleaching. The rate of photobleaching increases proportionally to the square of the zoom factor (Centonze and Pawley, 1995). A guideline for selecting an appropriate zoom factor derived from information theory (the Nyquist Sampling Theorem) states that the pixel dimensions should be equal to the optical resolution divided by 2.3 (see Webb and Dorey, 1995). However, pixel dimensions smaller than this may produce more informative images.

Z axis sectioning interval

In order to study the three-dimensional structure of a specimen, images are collected at a series of focal levels at intervals determined by the commands sent to the focus motor. The most straightforward way to ensure that the reconstructed images have correct proportions in the *x*, *y*, and *z* axes is to collect optical sections at *z* axis intervals equal to the *x*, *y* pixel dimension. However, the interfocal plane interval needed to adequately sample the specimen in the *z* axis is not as small as the *x*, *y* pixel dimension because the axial resolution is poorer than the lateral resolution (see Table 14.11.1). The optimal interfocal plane interval (according to the Nyquist Sampling Theorem) is equal to the axial resolution divided by 2.3. Collecting images at shorter intervals results in oversampling with an increased risk of photobleaching.

Illumination intensity

Fluorescence emission increases linearly with illumination intensity up to a level at which emission saturates. Optimal signal-to-background and signal-to-noise ratios are obtained with illumination levels well below saturation (Tsien and Waggoner, 1995). The illumination intensity on a laser-scanning microscope can be adjusted by inserting neutral-density filters into the light path and/or by operating the laser at submaximal power. In general, the best images are obtained with illumination levels that are as high as possible without producing unacceptable rates of photobleaching.

PMT black level and gain

The contrast and information content of confocal images are influenced by the black level and gain of the photomultiplier tube (PMT) amplifiers. To obtain maximal information, the black level and gain should be adjusted to take advantage of the full dynamic range of the PMTs. The appropriate black level setting can be found by scanning while the light path to the PMT is blocked. The image that appears on the display monitor should be just barely brighter than the background, which is black (gray level = 0). To set the gain, scan the specimen and adjust the gain so that the brightest pixel in the image is slightly below white (gray level = 255). Selecting black level and gain settings which ensure that all signals fall within the dynamic range of the PMT is important for quantitative imaging experiments. The software provided with many confocal microscopes includes a pseudocolor image display mode that facilitates selection of appropriate

black level and gain settings by highlighting pixels with intensity values near absolute black and absolute white.

Averaging

Confocal images of dimly fluorescent specimens captured at typical scan rates (1 to 2 sec/frame for a slow-scan confocal microscope) appear noisy because of the small numbers of photons collected from each spot. In some instances, it may be possible to improve the signal-to-noise ratio by scanning the specimen at slower rates. Another way to obtain a better image is by summing and averaging the signals obtained in multiple scans (frame averaging). Some confocal microscopes provide a second averaging method (line averaging), in which individual lines are repeatedly scanned and averaged. Line averaging generally produces sharper images than frame averaging (which averages full frames) because there is less risk of blurring due to movements or changes in the specimen.

Image display

Commercial confocal microscope packages provide software for some types of image enhancement and display. The display options for three-dimensional datasets typically include “*z* projections” (see Fig. 14.11.1B), which are two-dimensional displays formed by superimposition of stacks of optical sections, and stereoscopic views, which are made by combining two image stacks, one aligned in the *z* axis and the other with a displacement between successive images. Many systems also have the capability to compute cross-sections and projections of the specimen from varying angles. Computed projections for a sequence of view angles can be played as a movie in which the specimen appears to rotate around an axis. Such movies give the viewer a striking impression of the three-dimensional geometry of the specimen. Additional display options are available in various integrated software/hardware packages specifically designed for visualization and analysis of three-dimensional images.

Anticipated Results

Fluorescence in fixed specimens protected with an antifade agent is often sufficiently bright and resistant to photobleaching to make it possible to reconstruct three-dimensional images using imaging parameters that provide optimal resolution. Superb three-dimensional views may be obtained of structures as small and complex as a cell's cytoskeleton or the

terminal arbor of an axon. The maximum depth in the specimen at which adequate images can be obtained depends on a number of factors (e.g., the match in refractive indexes of the immersion and mounting media, the wavelength of light, and the extent of scattering and absorption by the specimen). Under optimal conditions, it may be possible to image structures at depths near the limit allowed by the working distance of the objective; in practice, image quality usually deteriorates at depths in the range of a few hundred micrometers or less.

Although confocal microscopy on living cells is more difficult and damage to the tissue may preclude extensive three-dimensional reconstruction, the added time dimension and confidence in the reality of the images makes it well worth the effort. In addition, it is possible to study dynamic processes lasting for hours by collecting sequences of time-lapse images. Robust fluorophores such as certain variants of GFP (S65T, EGFP) can be imaged repeatedly with minimal loss of fluorescence (see, for example, Ellenberg et al., 1997). In addition, modern laser-scanning confocal microscopes provide a versatile optical bench and sophisticated specimen positioner which permit a wide range of experiments with the controlled application of laser light to living tissues. Current examples of these approaches are photobleaching (Cole et al., 1996; Wedekind et al., 1996) and release of caged compounds (Callaway and Katz, 1993; Svoboda et al., 1996). These are only the harbingers of many future applications of light probe physiology made possible by the versatility of the confocal microscope.

Resources Available via Internet

NIH Image, a powerful image analysis program for Macintosh computers developed by W. Rasband (Research Services Branch, National Institute of Mental Health, NIH), has many useful tools for analysis of confocal images. It can be downloaded from <http://rsb.info.nih.gov/nih-image/> or obtained via FTP from zippy.nimh.nih.gov. A version of NIH Image modified for operation under Windows also is available. Much information about fluorescent probes can be obtained from the Molecular Probes Web Site at <http://www.probes.com/>.

Many topics of interest to confocal microscopists are discussed on the confocal e-mail listserver network. To subscribe to the list, send the message "subscribe confocal<your name>" to listserv@ubvm.cc.buffalo.edu.

LITERATURE CITED

- Amos, W.B. 1995. Appendix 1: Optical units. *In Handbook of Biological Confocal Microscopy*, 2nd ed. (J. Pawley, ed.) pp. 579-580. Plenum, New York.
- Art, J.J. and Goodman, M.B. 1993. Rapid scanning confocal microscopy. *In Cell Biological Applications of Confocal Microscopy* (B. Matsumoto, ed.) pp. 47-78. Academic Press, San Diego.
- Bacallao, R., Kiai, K., and Jesaitis, L. 1995. Guiding principles of specimen preservation for confocal fluorescence microscopy. *In Handbook of Biological Confocal Microscopy*, 2nd ed. (J. Pawley, ed.) pp. 311-326. Plenum, New York.
- Brelje, T.C., Wessendorf, M.W., and Sorenson, R.L. 1993. Multicolor laser scanning confocal immunofluorescence microscopy: Practical applications and limitations. *In Cell Biological Applications of Confocal Microscopy* (B. Matsumoto, ed.) pp. 98-182. Academic Press, San Diego.
- Bridgman, P.C. and Reese, T. 1984. The structure of cytoplasm in directly frozen cultured cells. 1. Filamentous meshworks and the cytoplasmic ground substance. *J. Cell Biol.* 99:1655-1668.
- Callaway, E.M. and Katz, L.C. 1993. Photostimulation using caged glutamate reveals functional circuitry in living brain slices. *Proc. Natl. Acad. Sci. U.S.A.* 90:7661-7665.
- Centonze, V. and Pawley, J. 1995. Tutorial on practical confocal microscopy and use of the confocal test specimen. *In Handbook of Biological Confocal Microscopy*, 2nd ed. (J. Pawley, ed.) pp. 549-570. Plenum, New York.
- Cole, N., Smith, C., Sciaky, N., Terasaki, M., Edidin, M., and Lippincott-Schwartz, J. 1996. Diffusional mobility of Golgi proteins in membranes of living cells. *Science* 237:797-801.
- Denk, W., Piston, D.W., and Webb, W.W. 1995. Two-photon molecular excitation in laser-scanning microscopy. *In Handbook of Biological Confocal Microscopy*, 2nd ed. (J. Pawley, ed.) pp. 445-458. Plenum, New York.
- Ellenberg, J., Siggia, E.D., Moreira, J.E., Smith, C.L., Presley, J.F., Worman, J.J., and Lippincott-Schwartz, J. 1997. Nuclear membrane dynamics and reassembly in living cells: Targeting of an inner nuclear membrane protein in interphase and mitosis. *J. Cell Biol.* 138:1193-1206.
- Giloh, H. and Sedat, J.W. 1982. Fluorescence microscopy: Reduced photobleaching of rhodamine and fluorescein protein conjugates by *n*-propyl gallate. *Science* 217:1252-1255.
- Inoue, S. and Spring, K.R. 1997. *Video Microscopy: The Fundamentals*, 2nd ed. Plenum, New York.
- Johnson, G.D., Davidson, R.S., McNamee, K.C., Russell, G., Goodwin, D., and Holborow, E.J. 1982. Fading of immunofluorescence during microscopy: A study of the phenomenon and its remedy. *J. Immunol. Methods*. 55:231-242.
- Kamabadur, R., Koizumi, K., Stivers, C., Nagle, J., Poole, S., and Oldenwald, W. 1998. Regulation of *POU* genes by *castor* and *hunchback* establishes layered compartments in the *Drosophila* CNS. *Genes & Devel.* 12:246-260.

- Keller, E. 1995. Objective lenses for confocal microscopy. *In Handbook of Biological Confocal Microscopy*, 2nd ed. (J. Pawley, ed.) pp. 111-126. Plenum, New York.
- Kino, G.S. 1995. Intermediate optics in Nipkow disk microscopes. *In Handbook of Biological Confocal Microscopy*, 2nd ed. (J. Pawley, ed.) pp. 155-166. Plenum, New York.
- Majilof, L. and Forsgren, P. 1993. Confocal microscopy: Important considerations for accurate imaging. *In Cell Biological Applications of Confocal Microscopy* (B. Matsumoto, ed.) pp. 79-97. Academic Press, San Diego.
- Matsumoto, B. (ed.) 1993. *Cell Biological Applications of Confocal Microscopy*. Academic Press, London.
- Sandison, D.R., Williams, R.M., Wells, K.S., Strickler, J., and Webb, W.W. 1995. Quantitative fluorescence confocal laser scanning microscopy. *In Handbook of Biological Confocal Microscopy*, 2nd ed. (J. Pawley, ed.) pp. 39-54. Plenum, New York.
- Shotton, D.M. 1993. Electronic acquisition of light microscope images. *In Electronic Light Microscopy* (D.M. Shotton, ed.) pp. 1-38. Wiley-Liss, New York.
- Svoboda, K., Tank, D.W., and Denk, W. 1996. Direct measurement of coupling between dendritic spines and shaft. *Science* 272:716-719.
- Terasaki, M. and Dailey, M.E. 1995. Confocal microscopy of living cells. *In Handbook of Biological Confocal Microscopy*, 2nd ed. (J. Pawley, ed.) pp. 327-346. Plenum, New York.
- Tsien, R.Y. and Bacskai, B.J. 1995. Video-rate confocal microscopy. *In Handbook of Biological Confocal Microscopy*, 2nd ed. (J. Pawley, ed.) pp. 459-478. Plenum, New York.
- Tsien, R.Y. and Waggoner, A. 1995. Fluorophores for confocal microscopy. *In Handbook of Biological Confocal Microscopy*, 2nd ed. (J. Pawley, ed.) pp. 267-280. Plenum, New York.
- Waterman-Storer, C.M., Sanger, J.W., and Sanger, J.M. 1993. Dynamics of organelles in the mitotic spindles of living cells: Membrane and microtubule interactions. *Cell Motil. Cytoskel.* 26:19-39.
- Webb, R.H. and Dorey, C.K. 1995. The pixilated image. *In Handbook of Biological Confocal Microscopy*, 2nd ed. (J. Pawley, ed.) pp. 55-68. Plenum, New York.
- Wedekind, P., Kubitscheck, U., Heinrich, O., and Peters, R. 1996. Line-scanning microphotolysis for diffraction-limited measurements of lateral diffusion. *Biophys. J.* 71:1621-1632.
- Wessendorf, M.W. and Brelje, T.C. 1993. Multicolor fluorescence microscopy using the laser-scanning confocal microscope. *Neuroprotocols* 2:121-140.
- Wilson, T. (ed.) 1990. *Confocal Microscopy*. Academic Press, London.
- Wilson, T. 1995. The role of the pinhole in confocal imaging system. *In Handbook of Biological Confocal Microscopy*, 2nd ed. (J. Pawley, ed.) pp. 167-182. Plenum, New York.

KEY REFERENCES

- Inoue and Spring, 1997. See above.
Covers basics of light microscopy, video microscopy, and much more.
- Matsumoto, B. (ed.) 1993. See above.
Good source of practical information.
- Pawley, J. (ed.) 1995. *Handbook of Biological Confocal Microscopy*, 2nd ed. Plenum, New York.
Excellent source of theoretical and technical information.
- Russ, J. 1995. *The Image Processing Handbook*. CRC Press, Boca Raton, Fla.
Excellent source of information about digital image processing.
- Shotton, D.M. (ed.) 1993. *Electronic Light Microscopy*. John Wiley and Sons, New York.
Covers many aspects of light microscopy, including confocal microscopy.
- Wilson, T. (ed.) 1990. See above.
Good source of theoretical background information.

INTERNET RESOURCES

- zippy.nimh.nih.gov
Use to obtain NIH Image via FTP.
- <http://optics.jct.ac.il/~aryeh/Spectra>
Source of excitation emission spectra for common fluorophores.
- <http://rsb.info.nih.gov/nih-image/>
Use to obtain NIH Image.
- <http://www.probes.com>
Molecular Probes web site, including product listings and much more.
- <http://www.mwrn>
Microworld Resources, a comprehensive list of microscopy products and vendors.
- listserv@ubvm.cc.buffalo.edu
Confocal e-mail listserv network.

Contributed by Carolyn L. Smith
National Institute of Neurological Disorders
and Stroke
Bethesda, Maryland

Measurement of In Situ Hybridization

UNIT 14.12

Hybridization of labeled specific molecular probes to nucleic acids in tissues allows geometric and functional location of gene expression or of foreign genome sequences. Estimates of amounts and location of target nucleic acid sequence can be made with phosphor storage imaging and molecular controls (see Basic Protocol). A reference material for the procedure may be prepared from RNA, DNA, or virus particles (see Support Protocol).

DETERMINING THE DISTRIBUTION OF RADIOLABELED HETERODUPLEXES IN IN SITU HYBRIDIZATIONS BY PHOSPHOR STORAGE IMAGING

**BASIC
PROTOCOL**

This protocol describes the application of phosphor storage imaging to determining the distribution of radiolabeled heteroduplexes in in situ hybridizations of cells and tissues on microscope slides. The slides and relevant controls and reference materials (labeled with β -emitting isotopes) are exposed on phosphor image storage plates. The resulting signals are nearly linear with respect to radioactivity; they may be digitized and the pixels can be used for quantitative measurements. Alternatively the images so produced may be analyzed by software to determine geometric distribution of label. Inclusion of suitable reference standards allows calculation of relative specific activities at a resolution of 25 μm . Digital results can be confirmed by autoradiography of the slides after imaging.

Materials

- Reference standards (clots; see Support Protocol) hybridized with the same radiolabeled probe used for in situ hybridization
- Radiation standards consisting of ^{14}C embedded in plastic (suitable for ^{35}S or ^{33}P ; may be purchased on a microscope slide from American Radiolabeled Chemicals; <http://www.arc-inc.com>)
- Hybridized slides from in situ hybridization (Chapter 14)
- Phosphor storage imager (e.g., Fuji BAS 5000) with phosphor plates
- Light box or high-intensity light for erasing phosphor plates
- Manila folders
- Double-sided adhesive tape (Scotch brand)
- Plastic wrap (e.g., Saran wrap)
- Image processing software: NIH Image (available for free from the NIH website; <http://rsb.info.nih.gov/nih-image/Default.html>) or MCID (Imaging Research)

Prepare slides for phosphor imaging

1. Assemble reference standards and dry hybridized slides. Erase the phosphor plate completely by exposure to visible light.

Manufacturers provide light boxes that will accomplish erasure, or a high-intensity light may be purchased from the Sun Box (<http://www.sunboxco.com/>).

2. Using double-sided tape, attach slides, with the sections facing up, to a manila folder. Cover this array of slides with a piece of plastic wrap, securing the plastic in place with tape (do not use plastic wrap for ^3H -labeled slides).

Select plastic wrap brands that are as thin as possible. Some brands have plasticizers that may be removed from the slides by washing them in 95% ethanol; Saran wrap, however, is recommended since it does not present this problem.

**In situ
Hybridization
and Immuno-
histochemistry**

3. Place the array of slides in the plate holder, face down on the phosphor plate (be very careful with the imaging plate). Seal the plate holder and store at room temperature in a desiccating cabinet for the requisite exposure time.

Storage plate quality deteriorates if the plate is exposed to moisture; therefore sections should be dry and the plates should be stored in a desiccating cabinet. The authors do not advise attaching slides to the tray with rubber cement, since the solvent in the cement may have an adverse effect on the plate or its properties.

4. Following exposure, working in the dark, take the manila folder with attached slides out of the holder. Take the plate, in the plate holder, to the instrument. Remove the plate from the holder, and in the case of the Fuji BAS 5000, insert it into the machine with the lights out.

The machine is instructed to "read" by the software. The Fuji BAS 5000 instrument reads at a resolution of 25 μm and produces a data file of 160 Mb. This is best handled by a computer with at least 220 Mb RAM, an external storage system with at least 1 Gb capacity, a CD read-only recorder, and a hard disk with 2 Gb storage. After reading the plate, the data is transferred to an external storage system. Up to four data files may be archived on a CD-ROM 650 MB recorder disc.

In contrast to many other instruments used for measuring radioactive decay, storage phosphor images record derivatives of radiation by detecting ionizing radiation from β and γ emitters. The energy is captured by phosphors in the detection screen. The detection screen is swept with a laser light source and visible light is emitted from the phosphorescent screen. This visible light is registered on a photomultiplier and assembled into an image by the computer (Fig. 14.12.1). The signal is linear with respect to radioactivity over several orders of magnitude.

5. Once the plate has been read and the image verified, prepare the slides for autoradiography (UNIT 14.4).

This is an important step since the image on the radiogram may be used to verify events seen on the imaging plate and to assist in the measuring process.

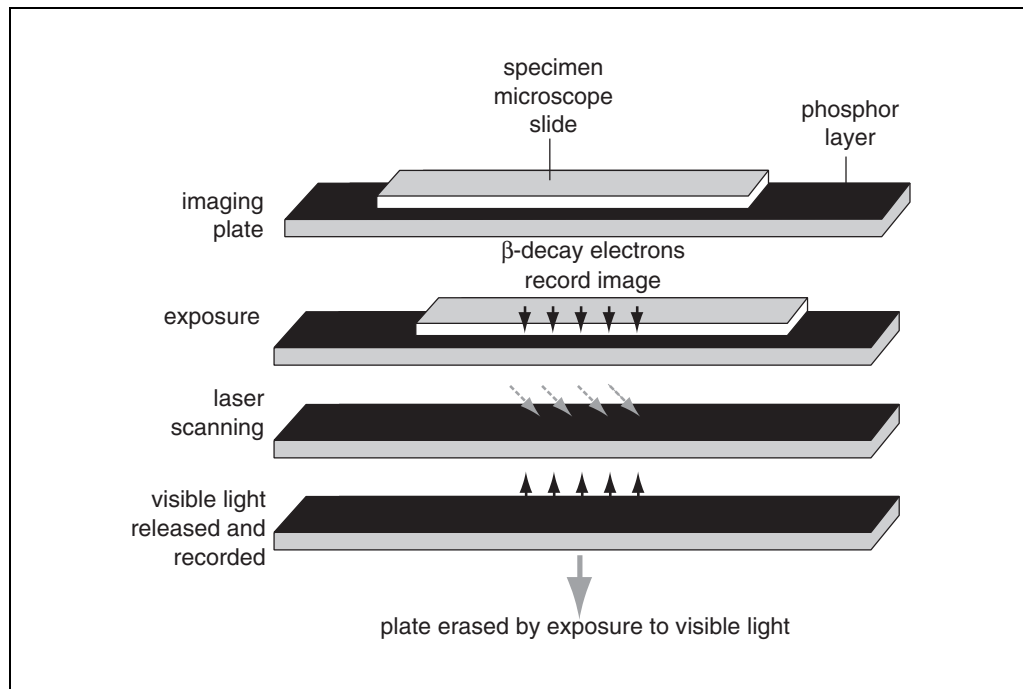


Figure 14.12.1 Phosphor storage imaging.

Evaluate results

6. Extract specific information from the image and to confirm the validity of the information. Keep the following in mind.
 - a. The data from the plate will appear in the computer memory as an image file. Details will vary among instruments and with the computer platform but such files can be very large and appropriate RAM memory should be available.
 - b. On the image of the plate there should be outlines of the slides and of the radiolabeled material. The ^{14}C reference slide should show a progressive gradation of density or, if converted to color, an increase in color from the lookup table.
 - c. The phosphor units should be linear. If they are not, there may be problems with the screens or the hardware.
 - d. A typical experiment is diagrammed in Figure 14.12.2, panels A and B. Mouse fetuses were hybridized with a probe pair and the images were captured on the same plate.
7. Analyze the images using software supplied by the phosphor image manufacturer, or export them to other image processing software such as NIH Image or MCID. When importing files into NIH image, set the width and height of the image in pixels in the custom import menu and then bring in the image as a 16-bit unsigned image on a fixed scale of 0 to 65535.

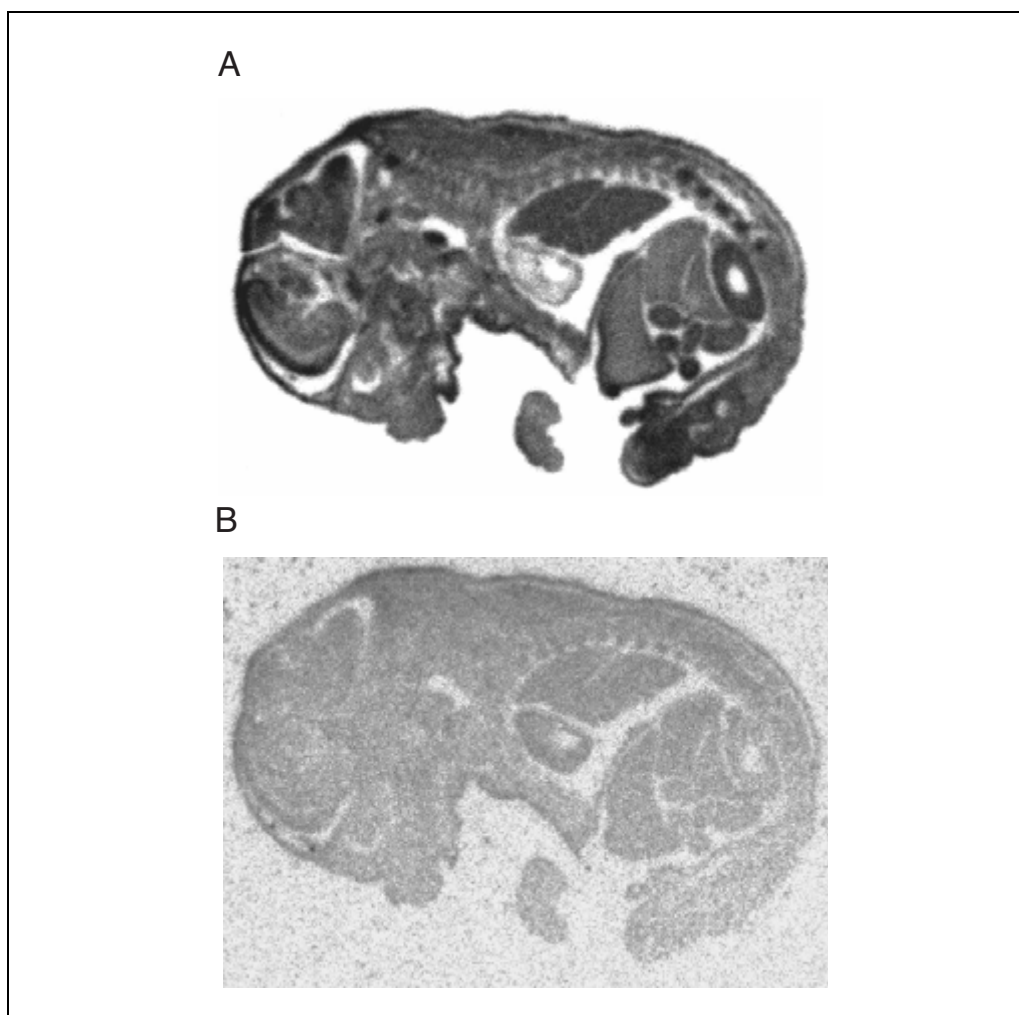


Figure 14.12.2 (A) Storage phosphor image of a mouse fetus labeled with ^{35}S antisense riboprobe. (B) ^{35}S sense riboprobe hybridized with mouse fetus.

See *Critical Parameters* for interpretation of hybridized control slides.

Although manufacturers report a detection limit as low as 0.9 dpm/mm²/hr; for ¹⁴C (and presumably ³⁵S and ³³P) the time of exposure can vary considerably. With the high-resolution plate, the exposure time may extend to 4 days without overexposure. Other plates will have a satisfactory image after only a few hours.

Estimate amounts of hybridization

- Determine the activity of the ¹⁴C plastic standards per mm² using the instrument's software.

These values can be plotted by the software or by use of a statistical program. The values should be linear or nearly linear with respect to radioactivity. Occasionally, a second-order polynomial will provide a statistically better correlation. Always include the ¹⁴C reference in each exposure.

- Determine the activity per mm² of the RNA clot reference slide. If the distribution of RNA is inhomogeneous, either prepare more reference clots (see Support Protocol) with greater attention to mixing, or, if the clot area is large enough, measure several 1-mm sites and take the mean.

At this point there should be several images on the exposure plate (Fig. 14.12.3). One is the radiation reference slide with multiple rectangles of radioactive plastic standards. Another is the image of the hybridization "standards" from the plasma clots probed with antisense probes accompanied by sections from the sense control and the nonsense control.

The clot hybridized with the sense probe may be used as a molecular control for antisense probe and in addition, a nonsense clot may be used to subtract background from the activity of the antisense probes.

Signals from nucleic acid hybridization are seldom as directly proportional as other chemical processes. There are a number of physical states for both the probe and the target nucleic acids in hybridization reactions that are highly variable. These may preclude conventional stoichiometric measurements. With that understanding, it is only possible to estimate the amounts of hybridization by constructing references such as the plasma clot containing known amounts of RNA. It is unknown what size of RNA will diffuse out of the clot and affect results. The relative amounts of probe hybridization for each experiment can be compared by reference to the permanent ¹⁴C objects, allowing a reasonable reliability of the process.

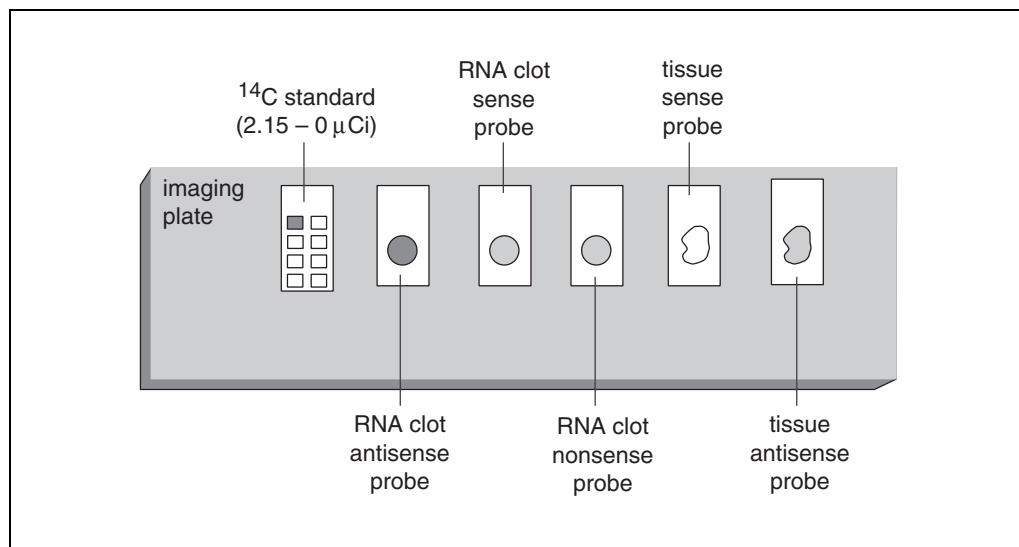


Figure 14.12.3 Estimating phosphor images.

10. Using the software provided with the imager, determine the phosphor values for the reference ^{14}C sections per mm^2 . If desired, a correction factor can be used to adjust for the differences in decay energy between the ^{14}C standard and the isotope used.
11. Determine the activity of the RNA clot, preferably at several 1-mm locations. Do the same for the sense and nonsense clot sections.

The authors prefer to use the nonsense sections for subtraction as a blank and reserve the sense section as a technique control. The test slides may be compared to a sense or nonsense control. It is assumed that the sections are infinitely thin and that self-absorption is constant. The thickness of the sections is assumed to be $6\ \mu\text{m}$, although microtome sections are subject to a number of variations.

There are several methods for estimating amounts of RNA or virion particles. One is to express values in pCi based on the radiation standards. Another is to convert the concentration of RNA in the corrected clot to phosphor units. This may be done either on the basis of molar concentrations in the clot or “viral genome equivalents” based on the amounts of RNA within the viral particles. Each investigator should develop experience with either of these possibilities.

PRODUCING A REFERENCE SYSTEM

Well characterized suspensions of a variety of viral particles are available from Advanced Biotechnologies. These preparations contain a known number of virions per ml. A quantity of RNA or DNA may also be made using a standard transcription system (*UNIT 14.8*) and assayed for concentration. The authors strongly recommend that a “nonsense” RNA be made as a control. Most suppliers of transcription reagents have a control DNA that can be transcribed to yield a nonsense probe that is irrelevant to the target gene or any other likely to be found. In situ hybridization using radiolabeled riboprobes may be conducted using established protocols (*UNIT 14.3* or Fox and Cottler-Fox, 1993). No special alterations in technique are required. The isotope used may be ^{33}P , ^{35}S , or ^3H . Tritium requires special screens for the phosphorimager. To test for linearity, a series of dilutions may be made of the calibration reference.

For a reference system, the following items are required: (1) radiation standards consisting of ^{14}C embedded in plastic are suitable for ^{35}S or ^{33}P ; these may be purchased on a microscope slide from American Radiolabeled Chemicals—and (2) an artificial tissue (Cottler-Fox and Fox, 1991), which is generated in this protocol using fibrinogen and thrombin and then hybridized with the same radiolabeled probe used in the in situ hybridization.

Materials

Radiolabeled standards: sample of RNA, DNA, or virion suspension (see recipe)
Fibrin glue (see recipe)
Thrombin, USP (see recipe; dilute to 10,000 U/ml)
1.3 M formaldehyde solution without buffers or salt (see recipe)
70% ethanol

1. Combine the fibrin glue with sample standard material in a ratio of 2 parts glue to 1 part standard (e.g., 1 ml glue to 0.5 ml). If the Tisseel system is used, add the standard suspension directly to the fibrinogen diluent used to reconstitute the lyophilized fibrinogen.

The quantities can be varied as long as the fibrinogen glue solution is not diluted >2:1.

CAUTION: *The fibrinogen is of human origin, and suitable biohazard precautions should be used.*

SUPPORT PROTOCOL

In situ Hybridization and Immuno- histochemistry

14.12.5

2. Coagulate the fibrinogen mixture by one of the following methods.
 - a. *If the Tisseel system is used:* Mix using the convenient dispenser provided with the kit, which has two paired syringes connected by a plastic valve allowing both solutions to be expelled and mixed simultaneously. Remove the hypodermic needle from the dispenser assembly for faster mixing. Expel the two solutions simultaneously and vigorously, which will cause a clot to be formed almost instantaneously.
The thrombin solution may be prepared at double strength if necessary.
 - b. *If the Tisseel system is not used:* Mix the fibrinogen solution thoroughly with the test (reference) solution (step 1) and add thrombin solution while the tube containing the mixture is rotated on a vortex mixer. Allow the clot to harden for 3 min.
3. Remove the clot from the tube with a very thin spatula or a broken applicator stick. Place the clot in 20 vol of formaldehyde solution at room temperature. Allow the clot to fix for 24 hr on a slowly (15 rpm) rotating shaker. Transfer the clot to a 70% ethanol solution and process as ordinary tissue (UNIT 14.1).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Fibrin glue

Available in some departments of transfusion medicine or may be obtained in a convenient pharmaceutical grade product kit as the Tisseel system (Baxter or Hyland-Immuno).

Fibrin glue is a concentrated fibrinogen solution made from human plasma and is used in surgery to reduce oozing of wounds.

Formaldehyde solution without buffers or salt, 1.3 M

Prepare by depolymerizing paraformaldehyde (Fluka-Sigma. See UNIT 14.1 for procedure, substituting water for saline).

IMPORTANT NOTE: *Do not use 37% commercial formaldehyde and do not add saline.*

CAUTION: *Formaldehyde is a putative carcinogen.*

Sample of standard material

RNA, DNA, or a virion suspension may be used. Concentration of the RNA and DNA varies considerably. It is a function of the length of the polymer and must be determined empirically. The concentration of the viral suspension should be $\sim 1 \times 10^{12}$ particles per ml.

Thrombin, USP

Topical bovine thrombin manufactured by GenTrac is a convenient product since it comes with diluent and is in a stable lyophilized form that may be stored at room temperature.

COMMENTARY

Background Information

In situ hybridization was first used by Gall and his students for genetic studies (Gall and Pardue, 1971). Since that time, a variety of strategies have been developed for using specific nucleic acid polymers (probes) for hybridizing with target sequences in cell and tissues (see Chapter 14 introduction). There are, however, serious problems in applying the basic principles of Southern or northern hybridization to the inhomogeneous and tortuous compartments of cellular DNA and RNA in tissues. The problem is further compounded by tissue fixation, probe accessibility, stringency of washes, and the enormous variety of biological structures that compromise routine or general protocols.

The functional principles of in situ hybridization in cells and tissues can be briefly stated as follows. The longer a sequence represented in the probe, the greater the sensitivity of target detection. Smaller probes require proportionately more target copies. If riboprobes are to be synthesized, the maximum length of insert should not exceed the capacity of the polymerase used, often 2 to 3 kb. These products should be truncated by hydrolysis to a more practical length. Since hybridization does not involve establishing covalent products and because of the vagaries of microenvironment reaction kinetics, stoichiometry is problematic. Probes based on nonradioactive chromogens are infinitely more difficult to measure than radiolabeled ones. Finally, it is important to remember that negative results in a well tuned and controlled in situ hybridization procedure are important results as well.

Critical Parameters

To estimate the amount of target molecules in tissue specimens, it is first necessary to have a functioning and well proven in situ technique and considerable experience with the plethora of problems in the procedure. The authors have chosen to adapt a protocol (Fox and Cottler-Fox, 1993) to formaldehyde-fixed tissues and test objects, since most archival tissue is available in paraffin blocks.

The expectations of the investigator must be carefully reasoned. If a simple estimate of the total amount of probe hybridized to a tissue section is all that is required, liquid scintillation counting, the use of ^{32}P for probe labeling, or estimations from γ counting of ring-labeled ^{125}I probes will suffice. For more sophisticated

studies, it may be necessary to determine the total number of copies in a particular cell, tissue, or organ, or to estimate the numbers of infected cells in viral diseases. It may also be desirable to determine the size or volume of tissues and targets by image analysis.

Slides that bear sense hybridization specimens should show marked differences from antisense hybridized slides, of at least one to two orders of magnitude. If levels of signal are the same in both sense and antisense, there are serious problems. Common difficulties include probe misidentification, incomplete washing, or, in the case of probe cocktails, the inclusion of an antisense fragment in the sense probe. A theoretical consideration is translation of the opposite strand of DNA resulting in the accumulation of antisense strands. Obviously these are technical details inherent in any in situ hybridization. Further cursory inspection of the image should show a gradient of signals from the ^{14}C control slide.

The principles of phosphor storage imaging were developed in Japan (Miyahara, 1989) and have revolutionized radiation measurement. The authors applied the phosphor imaging principle to measuring virus in tissues some time ago (Fox et al., 1994). Equipment that can be used (other than that used for in situ hybridization) is rather simple. Molds for constructing standards may be adapted from existing laboratory supplies. Imaging equipment is produced by a number of companies. We have chosen the FUJI BAS 5000 for its 25 μm resolution and dual Macintosh/PC compatibility. Other manufacturers, such as Packard Instruments, produce instruments at a lower price with only slightly less resolution and their own variety of software programs. Imaging screens also vary in resolution and sensitivity.

Troubleshooting

Some common hybridization problems are presented in Table 14.12.1.

Anticipated Results

Although manufacturers report a detection limit as low as 0.9 dpm/mm²/hr, for ^{14}C (and presumably ^{35}S and ^{33}P), the time of exposure can vary considerably. With the high-resolution plate, the exposure time may extend to four days without overexposure. Other plates will have a satisfactory image after only a few hours.

An example of how amounts of HIV in a tissue section are estimated is shown in Figure

Table 14.12.1 Troubleshooting

Problem	Possible cause	Solution
High signal in sense or nonsense specimen	Inadequate washing	Increase stringency of SSC washes, increase wash times
	Mislabeled	Repeat procedure
No signal in antisense specimen	Ribonucleases	Improve hygiene
No labeling	No labeling	Remake probe
	Low expression	Consider whether ISH is appropriate. Use higher-energy isotope.
Streaking of image	Inadequate washing	Refer to autoradiogram to see if it occurs in both images
Poor resolution	Defective screens or instrument	Purchase new screens. Service instrument.
Nonlinear measurements of clot controls	Poor mixing, or ribonucleases in fibrinogen	Recast clot
	Loss of RNA through diffusion	Add RNase inhibitors, fix clot promptly, increase fibrinogen, or use longer transcripts

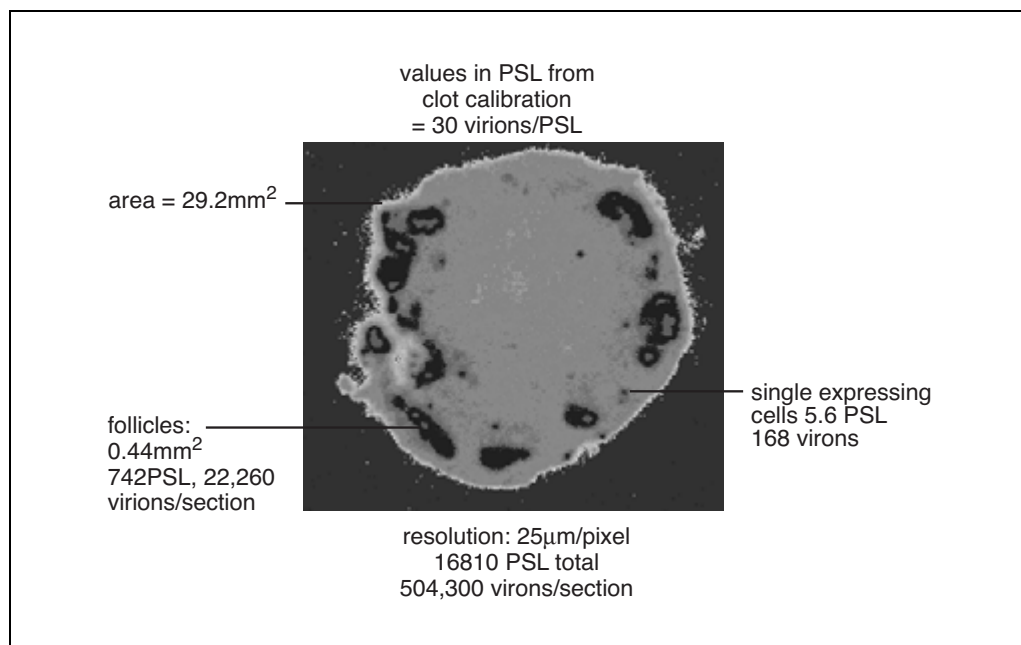


Figure 14.12.4 Quantitative anatomy of an HIV lymph node from phosphor imaging with antisense probe.

14.12.4. The measurement unit is in PSL, a term coined by Fuji for their radiation unit (Photo Stimulated Luminescence). The results are shown in PSL converted to viral genome equivalents. The image demonstrates the possibility of estimating viral expression on a single-cell basis.

Time Considerations

Time considerations are not especially critical. In general, a typical experiment runs as follows.

- a. Preparing probes: 1 day with 3 hr net time.
- b. Processing tissues and microtomy: 2 days, 3 hr net.
- c. Preparing reference: 2 days, 3 hr net.
- d. In situ hybridization of tissue and reference slides: 2 days, 8 hr net.
- e. Exposure of plate: 1 to 7 days, 3 days for the highest sensitivity/resolution.
- f. Scanning plate: 15 to 30 min.
- g. Analysis of results: 1 hr per section.

Literature Cited

- Cottler-Fox, M. and Fox, C.H. 1991. Examining cells for infectious agents: a novel approach. *J. Infect. Dis.* 164:1239-1240.
- Fox, C.H. and Cottler-Fox, M. 1993. In situ hybridization for detection of HIV RNA. *In* Current

Protocols in Immunology (J. Coligan, A. Krusbeek, D. Margulies, E. Shevach, W. Strober, eds.) pp. 12.8.1-12.8.21. John Wiley & Sons, New York.

Fox, C.H., Hoover, S., Currall, V.R., Bahre, H.J., and Cottler-Fox, M. 1994. HIV in infected lymph nodes. *Nature* 370:256.

Gall, J.G. and Pardue, M. L. 1971. Nucleic acid hybridization in cytological preparations. *Methods Enzymol.* 38:470-480.

Miyahara, J. 1989. Visualizing things never seen before: The imaging plate, a new radiation sensor. *Chemistry Today* Oct. 1989:29-36.

Contributed by Shelley B. Hoover
and Cecil H. Fox
Molecular Histology Laboratory, Inc.
Gaithersburg, Maryland

Jim Bahre
Fuji Medical Systems, USA
Stamford, Connecticut

Michele Cottler-Fox
University of Maryland Greenebaum
Cancer Center
Baltimore, Maryland

Cynthia Sung
Bioengineering and Physical Science
Program/National Institutes of Health
Bethesda, Maryland

Morphological, Biochemical, and Flow Cytometric Assays of Apoptosis

UNIT 14.13

As programmed cell death (PCD) or apoptosis has emerged as an important regulator of development and homeostasis in multicellular organisms, methods to quantify apoptosis and to distinguish it from necrosis have been developed. This unit presents a set of assays for these purposes, many of which are technically very simple. Necrosis refers to the morphology usually associated with accidental cell death, while apoptosis is seen when cell death is programmed or physiologically regulated. Loss of viability, whether the result of a necrotic or apoptotic process, is often defined experimentally as the loss of membrane integrity. The determination of whether a cell dies by apoptosis as opposed to necrosis is best made on the basis of (1) alterations in the cell membrane and cytoplasm and (2) changes in the cell's chromatin, both of which occur prior to the lysis of the membrane. The latter changes include extensive condensation of the chromatin as assessed by light or electron microscopy, and DNA fragmentation as detected by sedimentation assays, gel electrophoresis, or end labeling of the DNA fragments. Other apoptotic changes include membrane boiling or blebbing, which can only be appreciated microscopically, and the display of phosphatidylserine on the cell surface, which can be detected by flow cytometry. Loss of membrane integrity is conveniently measured by uptake of certain dyes such as trypan blue, eosin, ethidium bromide, or propidium iodide, or by the release of lactate dehydrogenase or radioactive chromium.

Basic Protocol 1 allows the qualitative and quantitative assessment of apoptosis in lymphocyte cell cultures using light or fluorescent microscopy. Using vital or DNA-binding fluorescent dyes, this method provides an easy way to determine the percentage of cells undergoing apoptosis and/or dying in a given population.

Next is a technique that quantitates apoptotic cells using flow cytometry. Basic Protocol 2 quantitatively detects DNA fragmentation in apoptotic cells based on the observation that apoptotic cells fixed with ethanol and stained with DNA binding dyes display an increased sub-G₀/G₁ peak in the flow histogram of DNA content.

Nucleotide end-labeling to detect DNA fragmentation is another technique for identifying apoptotic cells. Basic Protocols 3 and 4 describe TdT-mediated dUTP-biotin nick end-labeling (TUNEL) methods, which detect apoptotic cells by using terminal deoxynucleotidyl transferase (TdT) to directly label the ends of broken DNA strands. This approach is suitable for both qualitative and quantitative analysis. Basic Protocol 3 outlines the flow cytometric quantitation of apoptotic cells using TUNEL and Basic Protocol 4 describes TUNEL staining of tissue sections to identify apoptotic cells.

A great deal has been learned about the biochemical events and specific families of molecules that regulate the apoptotic process. However, much remains incompletely understood about the molecular pathways of programmed death, and it is probably best to perform more than one of the basic protocols to confirm an observation of apoptotic cell death.

NOTE: Unless the cells are to be used for flow cytometry or other processes that will require no further in vitro cultivation of the cells, all solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

In situ
Hybridization
and Immuno-
histochemistry

**MICROSCOPIC QUANTITATION OF APOPTOTIC INDEX AND CELL
VIABILITY USING VITAL AND FLUORESCENT DYES**

Perhaps the simplest and most accurate means to assess apoptosis is by careful microscopic inspection. By examining cells in the culture dish using a low-power inverted microscope, a qualitative diagnosis of apoptosis can often be made immediately. The tell-tale signs of apoptosis include a condensed nucleus, a blebbing cell membrane, and clear cytoplasm. With extensive cell death *in vitro*, a prominent feature is the accumulation of a large amount of debris. This is because phagocytosis, which normally efficiently removes early apoptotic cells *in vivo*, does not occur *in vitro*. This allows apoptosis to progress to full cellular disintegration and even secondary necrosis.

To easily quantify cell death, hemacytometer counting can be performed (*APPENDIX 3F*). An aliquot of cell suspension is mixed with an equal volume of 0.01% trypan blue and loaded into the trough of a standard hemacytometer. Usually three major morphologies are apparent: (1) healthy, live cells, which are clear white disks that have excluded trypan blue; (2) early apoptotic cells, which are white due to trypan blue exclusion but have an irregular shape (like a wad of crumpled paper) or a shrunken nucleus due to early apoptosis; and (3) cells in end-stage apoptosis or secondary necrosis, which are irregular, blue-stained shells or remnants of dead cells. Both live (white) and dead (apoptotic-appearing or blue) cells can be counted. However, the live cell number is more accurate, since the dead cells disintegrate and can be over- or underestimated depending on the timing of the culture and the nature of the debris that has formed. Although light microscopic examination can be time-consuming for a large number of samples, it is relatively easy and is the most direct means to evaluate apoptosis qualitatively and quantitatively for many biologists. Though this is beyond the scope of this unit, the characteristic changes in nuclear chromatin during apoptosis can also be readily visualized by electron microscopy.

For investigators with ready access to a fluorescent microscope, a more elaborate method can be used. This procedure gives slightly better distinction between apoptotic and necrotic cells.

In this protocol, a cell suspension is mixed with fluorescent DNA-binding dyes and examined by fluorescence microscopy to visualize and count cells with aberrant chromatin organization. Acridine orange is used to determine how many cells within a given population have undergone apoptosis, but it cannot differentiate between viable and nonviable cells. To do this, a mixture of acridine orange and ethidium bromide is used. The differential uptake of these two dyes allows for the identification of viable and nonviable cells (Mishell et al., 1980).

Any fluorescent DNA-binding dye capable of permeating viable cells (e.g., Hoechst 33342 or DAPI) may be used in place of acridine orange; however, these dyes require UV excitation wavelengths of <350 nm, which may not be available on standard fluorescence microscopes.

CAUTION: Acridine orange and ethidium bromide have been found by the Ames test to be highly mutagenic and should be handled with care.

Materials

Dye mix: 0.01% trypan blue (Life Technologies), 100 µg/ml acridine orange (Sigma), *or* 100 µg/ml acridine orange + 100 µg/ml ethidium bromide (Sigma), all prepared in PBS (*APPENDIX 2*)

Cell suspension at $\sim 5 \times 10^5$ to 5×10^6 cells/ml (Table 14.13.1) in complete RPMI 1640 medium (*APPENDIX 3F*)

12 × 75-mm glass tube
 Microscope slide and 22-mm² coverslip (no. 1 thickness; VWR Scientific)
 Fluorescence microscope (e.g., Zeiss or equivalent) equipped with a fluorescein filter set
 Additional reagents and equipment for counting cells (*APPENDIX 3F*)

Trypan blue staining:

- 1a. Place 10 to 20 μl trypan blue solution in the bottom of a 12 × 75-mm glass tube. Add an equal amount of a well-mixed cell suspension and gently mix by hand.

Sometimes cells loosely adhere to the bottom of the flask and may require liberation by gently scraping with a rubber policeman.

- 2a. Load 10 μl of the mixture into the trough of a standard hemacytometer. Examine with a 40× to 60× dry objective of a brightfield microscope.

- 3a. Count a minimum of 200 total cells and record the number of normal versus apoptotic or dead cells.

Cells in early apoptosis may appear white due to trypan blue exclusion, but they will have an irregular shape or shrunken nucleus.

- 4a. Calculate the percentage of apoptotic cells (apoptotic index) as follows:

$$\% \text{ apoptotic cells} = \frac{(\text{total no. of cells counted}) - (\text{total no. of live cells})}{\text{total no. of cells counted}} \times 100$$

This calculation is based on the number of live cells counted, as the dead cell count is less accurate.

Acridine orange staining:

- 1b. Place 1 μl of the dye mix in the bottom of a 12 × 75-mm glass tube. Add 25 μl of the cell suspension and mix gently by hand.

- 2b. Place 10 μl of this mixture on a clean microscope slide and cover with a 22-mm² coverslip. Examine with a 40× to 60× dry objective using epiillumination and a filter combination suitable for observing fluorescein.

Table 14.13.1 Stimulation of Apoptosis in Cultured Cells^a

Cell type	Treatment	Assay time (hr) ^b	Apoptotic index ^c	Protein synthesis required ^d
Thymocytes	100 nM dexamethasone	24	80-100	Yes
	600 rad γ-irradiation	24	80-100	Yes
D011.10 T hybridoma	Anti-CD3	24	80	Yes
P815 mastocytoma	Cytotoxic T cells	4	50-100	No
Raji B cells	Cytotoxic T cells	4	10-20	No

^aSee Background Information for a further description of the cell types and treatments employed in these studies.

^bApproximate time until the cells are assessed for apoptosis.

^cApoptotic index refers to the expected values for the percent of apoptotic/dead cells at the indicated times as quantified according to Basic Protocol 1.

^dRequirement of protein synthesis refers to whether inhibitors of RNA (e.g., actinomycin D) or protein (e.g., cycloheximide, emetine, and pactamycin) synthesis prevent apoptosis (nuclear damage as well as cell lysis).

A hemacytometer may be used for visualization (APPENDIX 3F). However, the ordinary slide/cover slip combination results in a greater flattening of the cells than can be obtained with the rigid coverslip used with hemacytometers, making apoptosis more readily apparent.

- 3b. Count a minimum of 200 total cells and record the number of normal versus apoptotic nuclei.

Acridine orange intercalates into DNA and makes it appear green. It also binds to RNA, but because it cannot intercalate, the RNA stains red-orange. Thus a stained cell will have a green nucleus and may have a red-orange cytoplasm.

It is extremely important to acquaint oneself with normal chromatin distribution in the cell line being investigated (see Commentary). Both live and dead nonapoptotic cell nuclei will fluoresce green and have "structure;" variations in fluorescent intensity reflect the distribution of euchromatin and heterochromatin. Apoptotic nuclei, in contrast, have highly of condensed crescents around the periphery of the nucleus; alternatively, the entire nucleus can be present as one or a group of featureless, bright spherical beads. In advanced apoptosis, the cell will have lost DNA or become fragmented into "apoptotic bodies" and the overall brightness will be less than that of a normal cell.

- 4b. Calculate the percentage of apoptotic cells (apoptotic index) as follows:

$$\% \text{ apoptotic cells} = \frac{\text{total no. of cells with apoptotic nuclei}}{\text{total no. cells counted}} \times 100$$

Acridine orange/ethidium bromide staining:

- 1c. Place 1 μl of the dye mix in the bottom of a 12 \times 75-mm glass tube. Add 25 μl of the cell suspension and mix gently by hand.
- 2c. Place 10 μl of this mixture on a clean microscope slide and cover with a 22-mm² coverslip. Examine with a 40 \times to 60 \times dry objective using epiillumination and a filter combination suitable for observing fluorescein.

A hemacytometer may be used for visualization (APPENDIX 3F). However, the ordinary slide/cover slip combination results in a greater flattening of the cells than can be obtained with the rigid coverslip used with hemacytometers, making apoptosis more readily apparent.

- 3c. Count a minimum of 200 total cells, recording the number of each of the following four cellular states: (1) viable cells with normal nuclei (VN; bright green chromatin with organized structure); (2) viable cells with apoptotic nuclei (VA; bright green chromatin which is highly condensed or fragmented); (3) nonviable cells with normal nuclei (NVN; bright orange chromatin with organized structure); (4) nonviable cells with apoptotic nuclei (NVA; bright orange chromatin which is highly condensed or fragmented).

As noted in step 3b, both live and dead cells take up acridine orange. It intercalates into DNA, making it appear green, and binds to RNA, staining it red. Thus a viable cell will have bright green chromatin in its nucleus and red-orange cytoplasm. Ethidium bromide is only taken up by nonviable cells. It intercalates into DNA, making it appear orange, but binds only weakly to RNA, which may appear slightly red. Thus a dead cell will have bright orange chromatin (the ethidium overwhelms the acridine) and its cytoplasm, if it has any contents remaining, will appear dark red. Cells that have undergone necrosis will have the fluorescent features of nonviable cells but will not have apoptotic nuclear morphology. A similar effect can be obtained with the membrane-impermeant dye propidium iodide, which will stain the cytoplasm and nucleus of necrotic cells orange.

Using this method, both normal or apoptotic nuclei in live cells will fluoresce bright green. In striking contrast, normal or apoptotic nuclei in dead cells will fluoresce bright orange.

- 4c. Calculate the percentage of apoptotic cells (apoptotic index) and percentage of necrotic cells.

$$\% \text{ apoptotic cells} = \frac{VA + NVA}{VN + VA + NVN + NVA} \times 100$$

$$\% \text{ necrotic cells} = \frac{NVN}{VN + VA + NVN + NVA} \times 100$$

$$\% \text{ dead cells} = \frac{NVN + NVA}{VN + VA + NVN + NVA} \times 100$$

If the cells have been undergoing programmed cell death for some time, or have fragmented into apoptotic bodies, they may have lost so much chromatin as to render accurate determination of cell viability impossible. "Viable" apoptotic (VA) cells are included in the numerator of the fraction of dead cells since "viability" refers to membrane integrity, not cell fate, and the appearance of apoptotic changes in the nucleus is always associated with cellular demise.

DETERMINATION OF APOPTOSIS USING SUB-G₀/G₁ DNA PEAK

This protocol depends on the observation that apoptotic cells have reduced or hypodiploid DNA fluorescence using conventional cell cycle profile analysis (see Telford et al., 1994). Like the other protocols using flow cytometry, this protocol allows for the quantitative analysis of intact cells.

Materials

Cells to be analyzed, in FACS buffer (see recipe; 10⁵ to 10⁶ total)
FACS buffer (see recipe)
Ice-cold 100% ethanol (70% if cells are to be antibody-labeled)
1 mg/ml RNase A
50 µg/ml propidium iodide (PI) in PBS (APPENDIX 2)

1. Pellet cells to be analyzed from culture following stimulus for apoptosis (see Table 14.13.1 for examples of stimuli). Also pellet cells from an untreated control culture. Resuspend in 3 ml cold FACS buffer, then pellet the cells by centrifuging 5 min at 300 × g. Resuspend pellets by gently flicking the tube.

It is imperative to prepare an untreated control in order to measure the sub-G₀/G₁ region.

If the cells are to be simultaneously stained for surface expression of other markers, this should be done prior to the fixation step.

- 2a. *If the cells have not been stained with other fluorochrome-labeled antibodies:* Fix the cells by resuspending the cells in FACS buffer, and gently adding 1 vol ice-cold 100% ethanol while vortexing. Incubate the cells for 30 min at room temperature.
- 2b. *If the cells have been stained with other fluorochrome-labeled antibodies:* Fix the cells by resuspending them in 500 to 1000 µl FACS buffer and adding 3 vol of 70% ice-cold ethanol while gently vortexing. Incubate cells ≤1 hr at 4°C.

The use of ethanol as a fixative is crucial. Other fixatives such as glutaraldehyde or formaldehyde do not consistently give the hypodiploid peak possibly due to chromatin cross-linking. Although the mechanism is not clear for obtaining the sub-G₀/G₁ peak, it is

BASIC PROTOCOL 2

In situ
Hybridization
and Immuno-
histochemistry

14.13.5

thought to be due to changes in the chromatin accessibility to dye (possibly from increased nuclear condensation from the ethanol) or the actual loss of DNA from the cells.

3. Pellet the cells. Resuspend in 3 ml cold FACS buffer, then pellet the cells by centrifuging 5 min at $300 \times g$. Resuspend the cells in 250 to 500 μl FACS buffer with 40 $\mu\text{g}/\text{ml}$ RNase A and 50 $\mu\text{g}/\text{ml}$ propidium iodide. Incubate ≤ 30 min in the dark at room temperature.

This step allows for the fragmented soluble DNA to leak out of the fixed cells, while simultaneously preventing nonspecific staining of RNA. Other DNA-binding dyes, including ethidium bromide, acridine orange, and anthracyclines, can be used for fixed cell analysis. This allows for greater flexibility when performing multicolor staining.

4. Analyze the cells using flow cytometry (FL2 or FL3) in a linear mode (Holmes and Fowlkes, 1991). The apoptotic fraction of cells can be measured by gating on the cells staining in the region below the large G_0/G_1 peak.

BASIC PROTOCOL 3

FLOW CYTOMETRIC QUANTITATION OF APOPTOTIC CELLS USING TUNEL

Terminal deoxynucleotidyl transferase-mediated dUTP-biotin nick end labeling (TUNEL) is a method for detecting apoptotic cells that exhibit DNA fragmentation. TUNEL (Gavrieli et al., 1992; Gorczyca et al., 1993) involves end-labeling the broken ends of the double-stranded DNA with biotin-conjugated dUTP using the enzyme terminal deoxynucleotidyl transferase (TdT). This protocol outlines a method for quantitating TUNEL-stained apoptotic cells by flow cytometric analysis. It allows apoptotic cells to be analyzed individually on a flow cytometer and is compatible with cell-surface staining, which permits the quantitation of cell death in specific subpopulations of cells that are discernible only by means of multicolor flow cytometric analysis (Kishimoto et al., 1995). If desired, this TUNEL procedure can be performed immediately following (but not before) cell-surface marker staining (Holmes and Fowlkes, 1991).

Materials

- Cells for analysis
- PBS (APPENDIX 2)
- 95% ethanol (not 100% ethanol), ice cold
- Paraformaldehyde fixative solution (see recipe)
- TdT reaction buffer (see recipe)
- TdT/biotin-dUTP mix (see recipe)
- Fluorescein isothiocyanate (FITC)-conjugated streptavidin (Jackson Immunoresearch; follow manufacturer's instructions for appropriate dilution)
- 12 \times 75-mm round-bottom centrifuge tubes
- IEC 6R6000 centrifuge with model 269 rotor (or equivalent)
- Additional reagents and equipment for immunofluorescence staining (optional; Holmes and Fowlkes, 1991) and flow cytometric analysis (Otten et al., 1995)

- 1a. *For multicolor analysis:* Perform immunofluorescence staining (Holmes and Fowlkes, 1991). Resuspend stained cells in PBS and transfer an aliquot containing 5×10^5 cells to a 12 \times 75-mm round-bottom centrifuge tube.

If this step is performed, it must be done prior to the TUNEL procedure. The final wash should be performed in PBS (instead of the staining buffer used in Holmes and Fowlkes, 1991) in 12 \times 75-mm round-bottom centrifuge tubes and the cells should not be treated with propidium iodide. It is essential that the fluorochromes used for multicolor staining not be destroyed by the fixation steps used in the TUNEL procedure. FITC, PE, and Texas

red are not affected. The duochromes (conjugates of PE and Texas red—e.g., Red 613) are not affected either, but allophycocyanin (APC) is destroyed by the fixation.

1b. *For TUNEL detection alone:* Transfer an aliquot containing 5×10^5 cells to a 12×75 -mm round-bottom centrifuge tube.

2. Add 1 to 2 ml PBS, then centrifuge cells 5 min at $300 \times g$ (1200 rpm in an IEC model 269 rotor), 4°C , and decant supernatant. Resuspend pellet in 250 μl PBS, then add 750 μl ice-cold 95% ethanol dropwise over a period of 5 to 10 sec while gently vortexing. Incubate 20 min at 4°C .

Gradually adding ethanol while gently vortexing reduces clumping during this fixation step.

3. Wash cells by adding 2 ml PBS and centrifuging 5 min at $400 \times g$ (1500 rpm in an IEC model 269 rotor), 4°C . Decant supernatant.

Higher centrifugation speed is required for fixed cells than for unfixed cells, because ethanol fixation causes cell shrinkage.

4. Flick the tube to resuspend cells in residual buffer, and add 1 ml paraformaldehyde fixative dropwise over a period of 5 to 10 sec while gently vortexing cells. Incubate 30 min at room temperature.

Paraformaldehyde fixation makes the intracellular constituents more accessible and greatly increases the sensitivity of TUNEL staining.

5. Wash cells with PBS as in step 3.

The washed cells can be stored for a few days at 4°C in the dark.

6. Wash cells as in step 3, using 0.5 ml TdT reaction buffer in place of PBS. Remove as much supernatant as possible by touching the lip of the inverted tube on absorbent paper immediately after decanting the supernatant.

7. Add 50 μl TdT/biotin-dUTP mix to cell pellet. Incubate 45 min at 37°C .

8. Wash cells with PBS as in step 3, then add 10 μl FITC-conjugated streptavidin (at the dilution recommended by the manufacturer). Incubate 30 min at room temperature, then wash again with PBS as in step 3.

The best staining is obtained with FITC-conjugated streptavidin. Other fluorochromes—e.g., PE and the “duochromes”—give much weaker staining.

9. Perform flow cytometric analysis (Otten et al., 1995).

See Anticipated Results for representative histograms obtained from flow cytometry using the TUNEL method.

IN SITU DETECTION OF APOPTOTIC CELLS IN TISSUE SECTIONS BY TUNEL

The TUNEL method of detecting cells that exhibit DNA fragmentation can also be performed on tissue sections to locate apoptotic cells in situ (Gavrieli et al., 1992). This can be done by end labeling with biotinylated dUTP and detecting with enzyme-conjugated streptavidin, but more sensitive staining is obtained by end labeling with digoxigenin-conjugated dUTP and detecting with two layers of antibodies, the last of which is conjugated to an enzyme that permits colorimetric detection (Surh and Sprent, 1994). An additional advantage of this approach is that it circumvents background staining from endogenous biotin. The TUNEL method described here, which uses digoxigenin-conjugated dUTP, is for frozen sections but can also be performed on paraffin sections. See

BASIC PROTOCOL 4

**In situ
Hybridization
and Immuno-
histochemistry**

14.13.7

UNITS 14.1 & 14.2 for relevant information on preparation and handling of frozen and paraffin sections. It should also be noted that there are now commercially available kits that eliminate one step through the use of enzymatically conjugated dUTP.

Materials

Fresh tissues for analysis
1% (w/v) paraformaldehyde in PBS (dissolve by stirring with low heat overnight and filter before use)
Tris-buffered saline (TBS; APPENDIX 2)
0.1% (v/v) H₂O₂ in TBS
TdT reaction buffer (see recipe)
TdT/digoxigenin-dUTP mix (see recipe)
2% (v/v) horse serum or FBS in TBS
Sheep anti-digoxigenin primary antibody solution (see recipe)
HRPO-conjugated anti-sheep secondary antibody solution (see recipe)
AEC substrate working solution (see recipe)
Mayer's hematoxylin (Sigma)
Crystal Mount mounting medium (Fisher)
Hydrophobic-barrier slide marker (e.g., PAP Pen; Research Products International)
Coplin jars or staining trays
Humidified container (see recipe)
Additional reagents and equipment for preparing frozen sections (as in immunoperoxidase staining; UNIT 14.2 & 14.6)

Prepare and fix sections

1. Prepare 5- to 8- μ m frozen sections (see UNIT 14.2), drying slides overnight at room temperature.

Slides can be stored 2 to 3 weeks at 4°C or a few months at -70°C. This prolonged storage is possible because DNA is fairly stable. Paraffin sections may also be used; these should be prepared, rehydrated, and air dried.

2. Draw a hydrophobic boundary on the glass around each section with a PAP Pen.

A tight boundary is critical because each section must be covered with a small volume of TdT reaction buffer (see step 7).

3. Generously cover the entire section with 1% paraformaldehyde and incubate 30 min at room temperature in a closed container.

A closed container is necessary to prevent evaporation but a humidified container is not required.

Sections can be fixed with acetone instead of paraformaldehyde for 5 min at room temperature. Acetone-fixed sections tend to stain with better resolution, but such staining is generally weaker and tends to have a higher background. If the sections are to be acetone fixed, draw the boundaries with the PAP Pen after the acetone has completely evaporated.

Carry out TUNEL reaction

4. Pour off the paraformaldehyde and wash the slide by incubating 5 min at room temperature in a Coplin jar or staining tray containing TBS, dipping the slide in and out of the solution three to four times during the course of the incubation.

This step is not required for acetone-fixed sections.

5. Cover section with 0.1% H₂O₂ in TBS. Incubate 30 min at room temperature in a closed container to quench endogenous peroxidase activity.

For acetone-fixed sections use 0.01% H₂O₂.

6. Wash with TBS as in step 4, then cover section with TdT reaction buffer to rinse out TBS.

This step is critical because the TdT reaction is highly sensitive to buffer conditions.

7. Remove as much reaction buffer as possible and add 25 µl TdT/digoxigenin-dUTP mix. Incubate 45 to 60 min at 37°C in a humidified container.

Carry out detection reaction

8. Wash slide once with TBS, then once with 2% horse serum or FBS in TBS, each time using the washing technique described in step 4.
9. Cover section with sheep anti-digoxigenin primary antibody solution and incubate 1 hr at room temperature in a closed container.
10. Wash slide as in step 8, then cover section with HRPO-conjugated anti-sheep secondary antibody solution and incubate 1 hr at room temperature in a closed container.

Develop color and mount slide

11. Wash slide as in step 8, then cover section with AEC substrate working solution and incubate 10 to 20 min at room temperature.

The intensity of color development can be visually monitored by low-power light microscopy. Develop until positive staining is strong with a minimal background.

Diaminobenzidine (DAB) may be used as a substrate in place of HRPO (see UNIT 14.6). This helps avert one critical drawback in using AEC—i.e., that the color precipitate starts to fade slowly within a few weeks. Although fading is not a problem with DAB, sections developed with this substrate have lower resolution because DAB tends to diffuse.

12. Wash slide as in step 8, then counterstain by incubating 0.5 to 1 min in Mayer's hematoxylin, then washing 5 min with tap water in a Coplin jar.

This step is optional; omitting it greatly reduces the fading problem with AEC staining. Where sections are not counterstained, the AEC-stained slide (from step 11) should be washed as in step 8; the coverslip should then be mounted as in step 13.

13. Wipe excess water from around the section and mount coverslip with Crystal Mount.

See Anticipated Results for representative stainings obtained using in situ TUNEL.

Because of the fading problem with AEC, sections developed with this reagent should be photographed within a few days after staining.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

AEC substrate solution

AEC stock solution:

Prepare 4 mg/ml 3-amino-9-ethylcarbazole (AEC; Sigma) in DMSO. Store up to 2 to 3 weeks in the dark at room temperature.

Working solution:

1 ml 0.17 M sodium acetate, pH 5.2 (APPENDIX 2)

25 µl AEC stock solution

1 µl 30% H₂O₂

Prepare immediately before use

FACS buffer

PBS (APPENDIX 2)

1% BSA (fraction IV, protease free; e.g., Sigma)

0.01% (w/v) sodium azide

Store indefinitely at 4°C if not contaminated

HRPO-conjugated anti-sheep secondary antibody solution

TBS (APPENDIX 2) containing:

2% (v/v) horse serum or FBS

5 µg/ml horseradish peroxidase (HRPO)-conjugated F(ab')₂ fragment of donkey anti-sheep IgG (Jackson Immunoresearch)

Prepare immediately before use

Humidified container (moist chamber; Fig. 14.2.4)

A moist chamber is constructed from any convenient-sized container with a tight lid (e.g., Tupperware or equivalent). Slide supports are made by attaching pairs of 5- or 10-ml pipets to the bottom, spaced so they support slides at both ends, or by using a stainless steel cake rack or equivalent. The bottom is lined with absorbent paper soaked with water or buffer. For use with light-sensitive reagents, the moist chamber is wrapped in aluminum foil, painted black, constructed from a black container, or shielded from ambient light.

Paraformaldehyde solution

Dissolve 1% (w/v) paraformaldehyde and 0.01% (v/v) Tween 20 in PBS. Stir at low heat (70°C) to dissolve paraformaldehyde. Store up to 6 months at room temperature. Filter through a 0.2-µm filter before use.

Sheep anti-digoxigenin primary antibody solution

TBS (APPENDIX 2) containing:

2% (v/v) horse serum or FBS

5 µg/ml sheep anti-digoxigenin Fab (Boehringer Mannheim)

Prepare immediately before use

TdT/biotin-dUTP mix

TdT reaction buffer (see recipe) containing:

0.5 to 1 µM biotin-dUTP (Boehringer Mannheim or Clontech)

2 to 4 U TdT (Promega or Boehringer Mannheim)

Prepare immediately before use

The TdT and biotin-conjugated dUTP should be titrated; the optimal concentrations should be near the indicated ranges. The best-quality biotin-dUTP preparation is the 1 mM biotin-16-dUTP from Boehringer Mannheim; the most economical is the 10 mM biotin-21-dUTP from Clontech (sold in 100 µl portions).

TdT/digoxigenin-dUTP mix

TdT reaction buffer (see recipe) containing:

1 to 2 µM digoxigenin-11-dUTP (Boehringer Mannheim)

2 to 4 U TdT (Promega or Boehringer Mannheim)

Prepare immediately before use

TdT and digoxigenin-conjugated dUTP should be titrated; the optimal concentrations should be near the indicated ranges. For acetone-fixed sections, a higher concentration of TdT (5 to 10 U) is required.

TdT reaction buffer

0.5 M cacodylic acid, sodium salt, pH 6.8

1 mM CoCl₂

0.5 mM DTT

0.05% (w/v) BSA

0.15 M NaCl

Store at room temperature or 4°C; stable several months

COMMENTARY

Background Information

Cell death in multicellular organisms has been classified into two broad but distinct categories—necrosis and apoptosis (Kerr et al., 1972; Duvall and Wyllie, 1986). Necrosis, also called accidental or pathological cell death, occurs when cells are exposed to conditions that vary significantly from physiological conditions. In contrast, apoptosis takes place under normal physiological conditions. Determination of whether cell death is necrotic or apoptotic is based on a number of morphological and biochemical criteria, with morphology being the first described (Kerr et al., 1972). In this brief discussion, apoptotic cell death will be delineated in terms of morphology, molecular biology, and biochemistry, and its importance in the immune system will be summarized. It is now clear that there are both positive and negative regulators of apoptosis that have important roles in immunoregulation (Vaux et al., 1988; Watanabe-Fukunaga et al., 1992). For more comprehensive discussions of apoptosis, consult Vaux and Strasser (1996), Lenardo (1996), and Chinnaiyan and Dixit (1997).

Necrotic cell death occurs in circumstances of wide departure from physiological conditions. Examples include the pathological death associated with hypoxia, inhibition of oxidative phosphorylation, glycolysis, or Krebs cycle enzymes, extreme hyperthermia, and excessive membrane damage due to toxins or complement activation. Necrosis results in an inflammatory response, oftentimes with additional tissue damage as a result of the inflammation. Apoptosis, in contrast, is a mode of cell death that occurs under normal and abnormal physiological conditions. However, unlike cells undergoing necrosis, apoptotic cells maintain the stability of important (“vital”) cellular machinery that enables it to actively participate its own demise. Importantly, apoptotic cell death does not lead to an inflammatory response, as dying cells are phagocytized prior to the release of pro-inflammatory intracellular contents.

Apoptosis has been best understood in terms of embryogenesis, metamorphosis, normal tissue turnover, and the homeostasis of lymphocytes in the immune system. Work over the past few decades has demonstrated the importance of apoptosis in the formation of the overall body plan of multicellular organisms. More recently apoptosis has been demonstrated to be involved in the precise regulation of cell number (Hengartner et al., 1992; Raff et al., 1993). In the immune system, apoptosis is involved in the removal of potentially harmful cells; self-reactive lymphocytes, tumor cells, and virally infected cells have all been shown to be removed through apoptosis. The role of apoptosis in maintaining homeostasis during immune responses is now well-established (Lenardo, 1991, 1996; Nagata and Golstein, 1995).

The morphological and biochemical changes associated with apoptosis (see below) are highly conserved in various cell types and systems. Apoptosis and the apoptotic machinery have been studied in phylogenetically distant organisms reflecting the evolutionary importance of this process. Studies in invertebrates, including genetic and biochemical studies in the nematode *C. elegans* by Horvitz et al. (Hengartner and Horvitz, 1994), have provided knowledge about families of molecules involved in regulating apoptosis. Much work in murine and human systems has provided an understanding of signaling pathways that regulate apoptosis, as well as the role of apoptosis in the immune system. There is a preponderance of evidence that an evolutionarily conserved cell death program exists in all cells, and that it can be activated under the appropriate conditions (Vaux et al., 1992). Although incompletely understood at present, it also seems likely that this common death pathway can be dysregulated leading to pathological conditions common in a variety of human diseases. Recently, a clearcut association between deficient apoptosis, lymphocyte accumulation, and autoimmunity has been elucidated in the human Autoimmune/Lymphoproliferative Syn-

drome (ALPS), which is associated with inherited mutations in Fas/CD95 (Fisher et al., 1995).

The cell death program

It is now apparent that there are several stages in the apoptotic process. In the first stage, a healthy functioning cell receives a stimulus to initiate the cell death program. This signal can come in the form of an active extracellular signal, such as that through the surface receptor APO-1/Fas/CD95 or the TNF receptor, or through a passive extracellular signal, such as the withdrawal of survival factors (e.g., IL-2 for cycling lymphocytes). Alternatively, a cell might initiate the death program via an intracellular signal such as that caused by a drug or radiation that activates p53. Intracellular signals can also affect the expression of death receptor-ligand systems (Hueber et al., 1997; Rehemtulla et al., 1997). The second stage involves the transduction of the death stimulus into the cell, which entails the activation of the effector arm. The effectors of apoptosis have been studied extensively and include the growing family of proteases known as caspases which have the unique ability to cleave directly following an aspartate residue in target proteins (for a more complete review of caspases, see Miller, 1997). The *C. elegans* genes *ced-3*, *ced-4*, and *ced-9*, and the corresponding mammalian families of genes that have been detected (see discussion below), are involved in the effector stage of the death program. During the effector stage of apoptosis, many biochemical changes occur to prepare a cell for engulfment and degradation by phagocytes, which is the final stage of apoptosis. A number of changes in the surface receptor and phospholipid composition occur, allowing for the selective removal of cells. A similar number of intracellular changes also occur, including drastic alterations in nuclear chromatin (Skalka et al., 1976), that presumably prepare a cell for degradation. Temporal aspects of the cell death program (i.e., which events occur early versus late during the apoptotic program) are poorly understood at the present time, although it is likely that many of the events occur simultaneously following activation of the protease cascade.

Morphological changes in apoptotic and necrotic cells

The initiating event in necrosis appears to involve an impairment of the cell's ability to maintain homeostasis, leading to an influx of

water and extracellular ions. Intracellular organelles, most notably the mitochondria, and eventually the entire cell, swell and rupture. The structure of the nucleus is relatively unchanged and the DNA is randomly degraded. In vivo, necrotic cell death is often associated with extensive tissue damage resulting in an intense inflammatory response.

Cells undergoing apoptosis do not swell, but show a progressive contraction of cell volume, with preservation of the integrity of cytoplasmic organelles (Wyllie, 1981). A decrease in cell volume results in the cells becoming more dense, thereby allowing apoptotic cells to be separated from normal cells by Percoll density gradient centrifugation (Wyllie and Morris, 1982). The reason for this shrinkage is not known, although it is likely that certain key structural molecules, including cytoskeletal proteins, are cleaved following the activation of proteases (see below).

While cell contraction during apoptotic cell death is impressive, widespread changes in nuclear morphology also occur (Kerr et al., 1972; Wyllie, 1981). These changes involve compaction of the nuclear chromatin into sharply circumscribed, uniformly dense masses aligned with the nuclear envelope. As described in Basic Protocol 1, condensation of chromatin is readily discerned and quantified by light microscopy, but is most dramatically observed with the electron microscope.

Late in the apoptotic process, the cells break apart into a number of membrane-bound apoptotic bodies containing one or more fragments of the nucleus. In vivo, these apoptotic bodies are rapidly phagocytosed and digested before they actually lyse. Phagocytosis presumably protects the adjoining cells and tissues from the damaging effects of the release of the intracellular contents of the dying cell which could elicit an inflammatory response. Thus, in contrast to necrosis, apoptosis is often remarkably inconspicuous histologically, making in vivo studies of the biochemical events difficult. This phenomenon is dramatically demonstrated in the thymus, where the vast majority of thymocytes undergo apoptosis during differentiation and selection, yet there are few apoptotic cells detectable.

Biochemical events in apoptosis

Early work in understanding the biochemical events that occur during programmed cell death focused on DNA degradation. DNA from apoptotic cells, when analyzed by agarose gel electrophoresis, is cleaved into a distinctive

ladder pattern consisting of multiples of an ~180-bp subunit (Wyllie, 1980). These chromatin fragments do not co-sediment with intact chromatin when the apoptotic cells are lysed and subjected to centrifugation. The temporal appearance of DNA cleavage correlates with chromatin condensation, suggesting that DNA fragmentation is responsible for the changes in nuclear morphology. The ladder pattern of DNA fragmentation is identical to that obtained when isolated nuclei are exposed to micrococcal nuclease, suggesting that DNA cleavage in apoptosis is the result of activation of an endonuclease that cleaves the DNA in the linker regions between nucleosome cores (Cohen and Duke, 1983). Very recently a protein (DFF) has been identified that is capable of triggering DNA fragmentation following its activation by the caspase family of proteases (Liu et al., 1997).

An important advance in understanding biochemical events that occur during apoptosis resulted from genetic studies in the nematode *Caenorhabditis elegans* (Hengartner et al., 1992; Hengartner and Horvitz, 1994). A large number of genes were identified that, when mutated, affected the highly regulated developmental pattern by altering the number of surviving cells. Two of these genes, *ced-3* and *ced-4* (*ced* standing for cell death defective), are required for cell death to occur, while a third gene, *ced-9*, protects cells that are not normally destined to die. Both *ced-3* and *ced-9* have mammalian homologues that have helped to define different functional families of molecules in the apoptotic pathway. Interleukin-1 converting enzyme (ICE), a previously described cysteine protease (see below) that has the unique ability to cleave following aspartic acid residues (X-D), was the original member of an expanding family of mammalian genes shown to be homologous to *ced-3*. Similarly *ced-9* is homologous in structure and function to an expanding family of mammalian genes that are named after the original member Bcl-2. The role for *ced-4* in apoptosis has been difficult to establish, and only recently has a mammalian homologue for this gene been discovered (Zou et al., 1997). Recent work suggests an adaptor/inhibitor function for *ced-4* whereby it physically interacts with both the *ced-9* and *ced-3* family members (Chinnaiyan et al., 1997; Spector et al., 1997; Wu et al., 1997).

The role of cysteine proteases in apoptosis is evident from studies in a number of systems. These proteases are all part of a specific family

of hydrolytic enzymes termed caspases (for cystinyl-aspartate-requiring proteases). *ced-3* mutants lack all somatic cell death. Transient overexpression of a number of the caspases leads to apoptosis in mammalian cells. Both viral genes and chemical substrates are capable of inhibiting caspases (Shen and Shen, 1995; Bertin et al., 1997); both can competitively inhibit by binding at the active site, and can protect cells from apoptosis in response to a number of stimuli (Ray et al, 1992; Clem and Miller, 1994). Owing to the large number of family members there is thought to be some redundancy in the function of the caspases. Targeted deletion of other cysteine proteases in mice will likely shed much light on their function and regulation.

The number of substrate molecules identified that are specifically cleaved by cysteine proteases during apoptosis has increased in the past 3 years. These intracellular substrates include a number of inactive cysteine proteases themselves, structural molecules (including nuclear lamins, actin, and non-erythroid spectrin), molecules involved in DNA repair [poly-(ADP ribose) polymerase, DNA-dependent protein kinase catalytic subunit], inactive transcription factors (steroid response element binding protein, or SREBP), and signalling molecules, such as Ly-GDI. Presently it remains to be formally demonstrated whether the cleavage of individual molecules has any causal role in the apoptotic process, whether the cleavage itself prepares the cell for further stages of programmed cell death, or whether they simply reflect downstream effects of the activation of proteases.

Initially it was thought that nuclear events were required for the apoptotic program, based on experiments showing that RNA or protein synthesis inhibitors could suppress apoptosis. However, it has become clear that the apoptotic machinery is present in the cytoplasm. This is seen most compellingly in studies involving anucleated “cytoplasts” that can assume the characteristic apoptotic morphology upon appropriate stimulation (Jacobson et al., 1994). This study makes it clear that the pre-existing machinery is present in the cytoplasm and can be triggered by apoptotic stimuli.

Detection of apoptosis using flow cytometry

Flow cytometry (also see Coligan et al., 2000) is a powerful tool that can be used to classify individual cells according to a number of different parameters including light diffraction—i.e., forward scatter (FSC) and side scat-

ter (SSC)—and uptake of fluorescent dyes—e.g., fluorescein isothiocyanate (FITC), phycoerythrin (PE), CyChrome/RED613, and propidium iodide (PI). For use in flow cytometry, FITC, PE, and CyChrome fluorochromes are conjugated to antibodies that bind epitopes on the cell surface. PI is a vital dye that is excluded by live cells, which have intact plasma membranes; however, it enters dead cells and intercalates into DNA, thereby making it possible to distinguish live cells from dead cells and to quantitate the number of live cells in a particular sample (Boehme and Lenardo, 1993a; Martin and Lenardo, 1998). Apoptosis can be distinguished from necrosis using FSC and SSC scatter profiles as cells undergo condensation and shrinkage prior to the loss of membrane integrity (Dive et al., 1992; Ormerod et al., 1995). Flow cytometry with PI staining is particularly valuable when coupled with staining by fluorochrome-conjugated antibodies (see Martin and Lenardo, 1998), because the combination provides a means of quantitating live cells that have a certain surface phenotype within a heterogeneous cell population (Boehme and Lenardo, 1993b; Critchfield et al., 1994). The PI emission spectrum, however, overlaps the FL2 channel, making dual staining with fluorochromes in the FL2 channel difficult, and there are a number of recently developed nucleic acid dyes that have a tighter emission spectra (Frey, 1995). Recently, the expression of the algal green fluorescent protein in mammalian cells has added a new dimension to fluorescence labeling of cells for flow cytometry, although it cannot replace antibody-linked fluorochromes.

One application of this combination of techniques is in experiments where activated, non-transformed T cells and T cell hybridomas undergo apoptotic cell death upon stimulation with antigen-presenting cells (APCs) and antigen (Ashwell et al., 1987; Lenardo, 1991; Wang et al., 1993; Martin and Lenardo, 1998). The activated T cells can be distinguished from the APCs by a number of parameters. Because irradiated APCs degenerate in culture, their size and FSC profile decreases and their PI-uptake profile increases in comparison to viable cells, making it possible to quantitate them according to these flow cytometry parameters. Also, the live cells can be further subdivided on the basis of antibody staining. Thus, by labeling the T cells with anti-clonotypic antibodies (e.g., anti-V β 8 or anti-V α 11) or T cell-specific antibodies (e.g., against Thy-1 or the TCR/CD3 complex, if the APCs have been depleted of T cells),

the number of live T cells of a particular phenotype can be quantitated despite the presence of APCs in the culture.

Another advantage of using the FACScan to quantitate the number of live cells is that it avoids inaccuracies inherent in determining cell loss on the basis of the number of dead cells. Cells undergoing programmed death fragment into multiple apoptotic bodies with variable size and dye uptake properties. This fragmentation makes accurate quantitation of cell loss through quantitation of dead cells extremely difficult, particularly if cell loss is measured late in the death program. Quantitation of apoptotic cells on the basis of DNA content (i.e., $<2n$) as determined by DNA-binding dyes has the same shortcoming, resulting from the fact that apoptotic cells eventually disintegrate. Furthermore, if antibody staining is to be used to determine the level of cell loss in a subpopulation of cells (see Alternate Protocol 2), only live cells can be clearly enumerated. Antibodies may bind nonspecifically to dead cells, thereby creating artifacts.

The major disadvantage of quantitating cell loss on the basis of the number of live cells is that this method does not provide strong evidence that cell death was due to apoptosis. The method outlined in Basic Protocol 2 should therefore be used in conjunction with other techniques (e.g., agarose gel analysis of DNA; *UNIT 2.5A*; see Martin and Lenardo, 1998) to demonstrate DNA fragmentation. Quantitation of cell loss based on the number of live cells also fails to determine definitively whether cell loss was due to actual cell death resulting from the apoptotic stimulus or to a cell-cycle blockade. If apoptosis has occurred, an increase in the number of dead cells as determined by flow cytometric analysis (i.e., cells that stain with PI and have a smaller FSC profile than live cells) should become apparent (see Fig. 14.13.1, upper left quadrant of panel B, for appearance of dead cells in a contour plot). Other methods that reveal dead cells (e.g., microscopy) can also be used to distinguish cell death from cell-cycle blockade. Furthermore, [3 H]TdR or BrdU uptake of cells that survive the apoptosis-inducing stimulus can be measured to determine if these cells are progressing through the S phase of the cell cycle.

Basic Protocol 2 is specifically designed to quantify cell death that occurs in activated mature T cells and T cell hybridomas upon TCR stimulation. This assay, however, can easily be applied to cell death occurring in other cell types under a wide range of conditions (see

Table 14.13.1), provided that various controls are maintained—i.e., resuspending each sample in a constant volume and collecting data for each sample over an equal length of time. Other applications of these protocols include quantitating cell death in primary B cells (Nemazee and Burki, 1989; Goodnow, 1992), in the murine B cell lymphoma WEHI-231 upon membrane Ig cross-linking (De Franco et al., 1987), and in hematopoietic cells cultured in the absence of growth factors such as IL-2, IL-3, and erythropoietin (Duke and Cohen, 1986; Koury and Bondurant, 1990). The authors have compared the quantitation of live cells by trypan blue dye exclusion and by flow cytometric analysis as detailed in this unit (Boehme and Lenardo, 1993a). These studies revealed that the relationship between these two methods is

linear (R value = 0.97); however, further comparison of the methods may be required in different cell systems.

TUNEL

By labeling the ends of DNA strand breaks, the TUNEL method detects apoptotic cells that are undergoing DNA fragmentation (Gavrieli et al., 1992; Gorczyca et al., 1993). There are several advantages to methods that employ TUNEL. First, because the end labeling is performed inside the cells, TUNEL is by far the most direct method for analyzing cells in which DNA fragmentation is occurring. Second, because DNA fragmentation is one of the earliest events in apoptosis, cells undergoing programmed cell death are detected at a very early stage in the process. Third, by using flow cy-

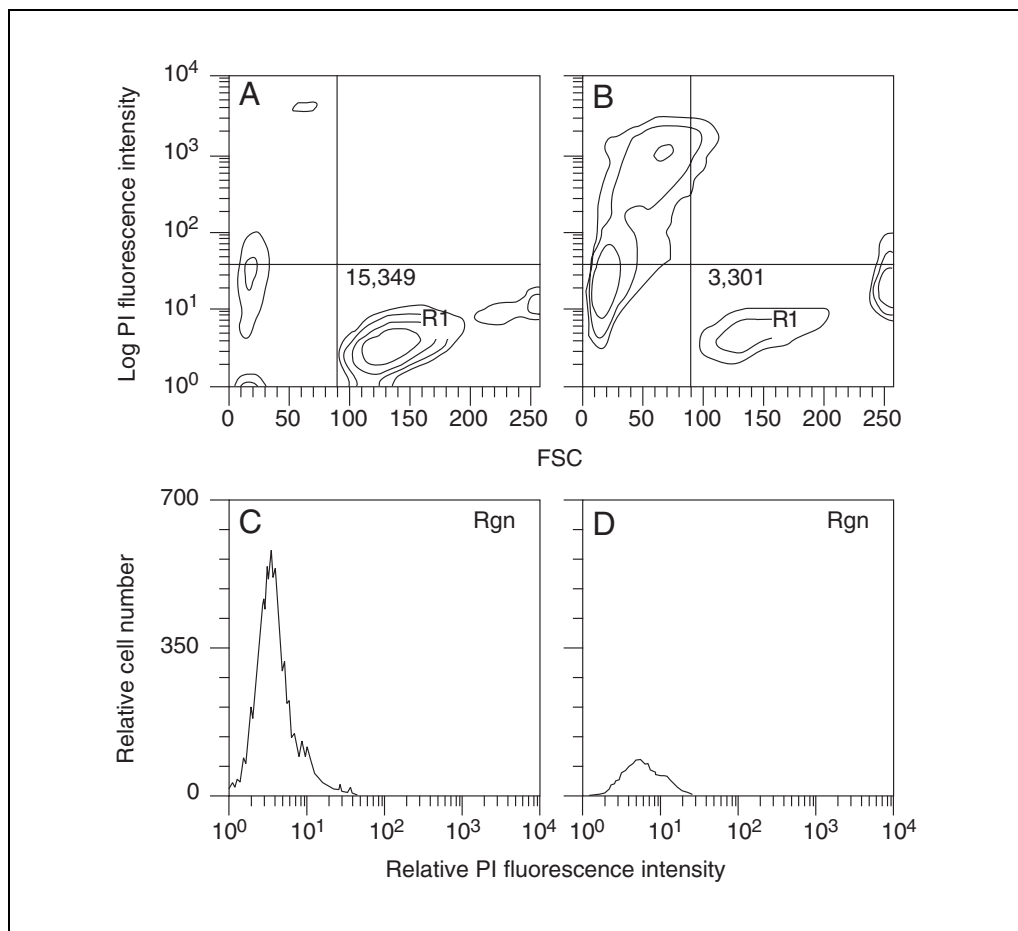


Figure 14.13.1 Dual-parameter FACS contour plots illustrating live and apoptotic cells, with corresponding histograms. **(A)** IL-2 stimulated T cell clone (A.E7) cultured 48 hr in medium alone. Note larger percentage of live cells (i.e., cells that exclude PI and have a large FSC profile) in the lower right quadrant and smaller percentage of apoptotic cells (i.e., cells that stain with PI and have a reduced FSC profile) in the upper left quadrant. **(B)** Same T cell clone cultured 48 hr with plate-bound anti-CD3 MAb. Note smaller percentage of live cells in the lower right quadrant and larger percentage of apoptotic cells in the upper left quadrant. **(C)** Histogram derived from the gated live cells in the lower right quadrant of panel A. **(D)** Histogram derived from gated live cells in the lower right quadrant of panel B. See *UNIT 5.2* for detailed discussion of contour plots and histograms.

tometry, the end-labeled cells can be individually analyzed to accurately quantitate apoptotic cells. Fourth, the TUNEL method is more sensitive than methods using vital dyes such as propidium iodide (PI); hence DNA fragmentation may be detected by TUNEL 4 to 6 hr before cell lysis (Sellins and Cohen, 1987; Kishimoto et al., 1995). Fifth, TUNEL staining is compatible with multiple-color cell-surface staining; cells can be stained for multiple surface markers with various fluorochromes prior to the TUNEL assay, thus allowing definition of the cell-surface phenotype of apoptotic cells (Kishimoto et al., 1995). As noted above, multicolor staining is also possible with vital dyes—e.g., PI or acridine orange—but because the vital dyes have a wide spectrum of emission, they cannot be used together with fluorochromes such as phycoerythrin (PE) and Red 613 unless they are used solely to “gate out” dead cells. Finally, the TUNEL method is suitable for staining apoptotic cells in tissue sections, thus allowing in situ visualization of dying cells with low-power microscopy (Gavrieli et al., 1992; Surh and Sprent, 1994).

The major disadvantage of the TUNEL method is that the reagents required for end labeling (dUTP and TdT) are expensive, probably making this the most costly means of detecting apoptotic cells. Other potential problems with the procedure apply generally to methods that detect apoptotic cells on the basis of DNA fragmentation. For example, TUNEL will not stain apoptotic cells that do not undergo typical double-stranded DNA strand breaks but instead undergo single-strand nicks or no DNA degradation at all (see above). Again, it is stressed that other methods of identifying apoptotic cells should be used in conjunction with this one to verify that cell death is due to apoptosis per se.

Critical Parameters and Troubleshooting

Fluorescence microscopy

Morphological appearance of the cells is the single most important criterion in apoptosis studies and it is important to include proper controls. These should include untreated cells and, if it is not manifestly apparent whether the cells of interest are undergoing apoptosis or necrosis, cells killed by regimens that should cause necrosis (e.g., 43°C for ≥2 hr; Harmon et al., 1990; Sellins and Cohen, 1991). It is extremely important to acquaint oneself with normal chromatin distribution in the cell line being

investigated; different cell types (e.g., normal thymocytes versus tumor cells) may present strikingly different patterns of “normal” chromatin organization. Similarly, it is a good idea to obtain experience in recognizing apoptotic nuclei in one of the systems that has been characterized (see Table 14.13.1).

Flow cytometry

Sample preparation. Each sample must be resuspended in the same volume of FACS buffer/propidium iodide (PI) and analyzed for the identical period of time for comparable results to be obtained. Therefore the FACS buffer and PI should be pipetted using a well-calibrated micropipettor for accuracy. The method also assumes that the FACScan flow rate is identical from sample to sample. This should be stringently tested by measuring the time that it takes for the FACScan to take up a certain volume from various samples with different cell concentrations.

Data acquisition. It is critical to acquire data on the basis of a constant collection time for all samples rather than a constant number of events. This is because a comparison will be made between samples on the basis of the actual number of live cells acquired over time rather than on the basis of the percentage of live cells in each sample. Using a constant flow rate, an absolute cell count per unit volume is obtained for each sample. To acquire data on the basis of collection time, the Time collection mode on the FACScan (found under the Protocol heading) is activated and the number of events collected at the upper limit (100,000; also found under the Protocol heading) is set. The length of time for collection should be chosen so that a statistically significant number of events is measured in the apoptotic cultures, but also so that the total number of events does not exceed the upper limit allowed (i.e., 100,000). Data should be acquired in the List-Mode storage setting (under the Protocol heading). This is done by activating the Auto-Save mode and giving each sample a file name. In the List Output Parameters file, the resolution for all parameters to be analyzed is adjusted to “256.” Parameters that will not be used for analyzing the samples, such as side scatter and FL1 or FL2, may be turned off to save data-storage space. Be sure not to change the flow-rate pressure setting between samples, because this will invalidate the comparison between samples. Do not gate any populations during this collection, although the FSC threshold may be raised (to a setting of 100 to 200 PMT volts) to

exclude very small particles, which will be numerous if significant cell death has occurred.

TUNEL with flow cytometry. When analyzing cells *ex vivo* immediately after preparation or after a period of incubation, the cells should be put into suspension using the gentlest method possible; this helps prevent induction of apoptosis during manipulation. For example, thymocytes should be obtained by gently teasing the thymus lobes with needles instead of by grinding them in a tissue grinder or through a steel screen. Best results are obtained when cells are fixed with ethanol followed by paraformaldehyde. Because both the ethanol and paraformaldehyde fixation steps can result in cell clumping, it is strongly recommended that the fixative be added slowly to cells as they are gently vortexed. As ethanol fixation shrinks cells, higher centrifugation speed is needed to produce a pellet at this stage than before fixation (i.e., $400 \times g$ instead of $300 \times g$). The final ethanol concentration used for fixation should

be 60% to 70%, as cells will shrink excessively with higher ethanol concentrations, making it impossible to obtain a pellet by low-speed centrifugation. When acquiring data by flow cytometry, the PMT voltage for FSC must be set higher to account for cell shrinkage, although the relative positions at which clumps and debris appear in the plot will not change. Total cell counts are also critical for the overall assessment of cell death. It is thus advisable to obtain relative cell counts prior to staining or when cells are acquired in flow cytometry. This can be done by resuspending the cells at constant volume and collecting them for a constant time period (see discussion of flow cytometry, above). Again, it is important that small particles, including debris and apoptotic bodies, be excluded to obtain accurate cell counts.

TUNEL in tissue sections. Tissues should be frozen while still as fresh as possible to reduce apoptosis induced during the handling period. Sections must be completely dried prior to

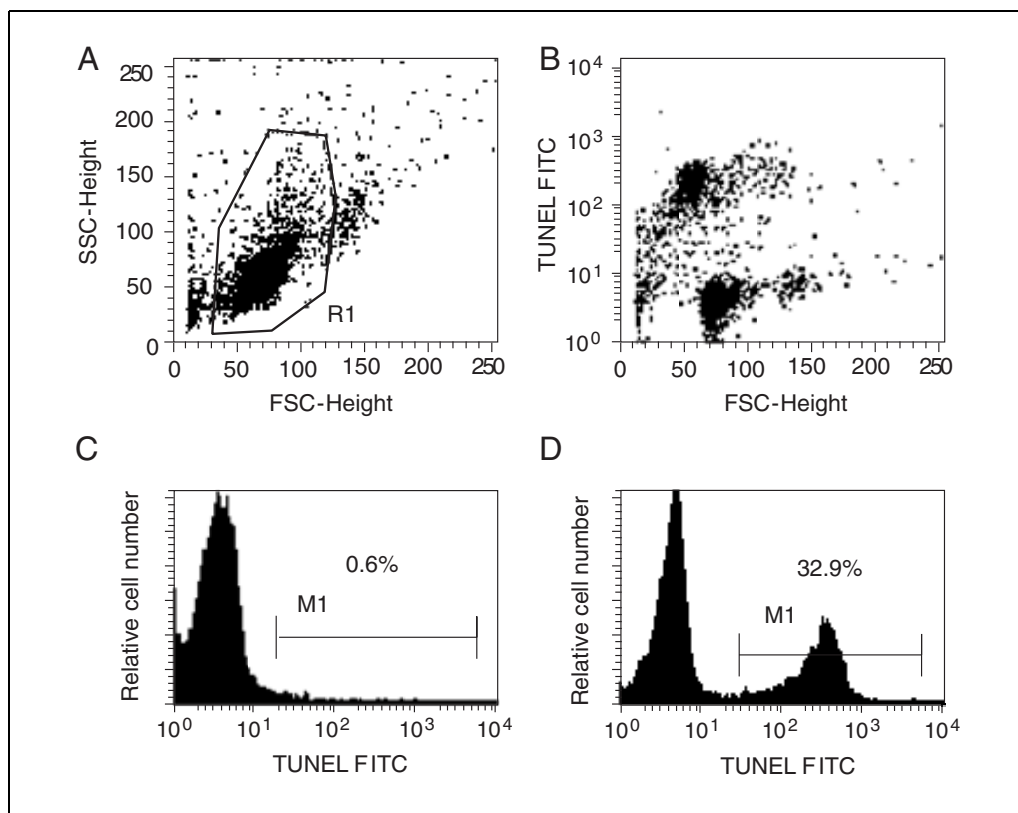


Figure 14.13.2 Detection of apoptotic cells by the TUNEL method using flow cytometric analysis. Thymocytes from 6-week-old C57BL/6 mice were stained for apoptotic cells with FITC by the TUNEL method before and after an 18-hr incubation in RPMI 1640/10% FBS at 37°C. **(A)** Dot plot depicting FSC versus SSC profile of cells after incubation. The bulk of thymocytes (in the gate R1) are easily distinguished from smaller cellular debris (low FSC) and larger cell clumps (high FSC). **(B)** Dot plot depicting profile of FITC fluorescence versus FSC for cells after incubation. Most of the TUNEL-stained apoptotic cells are slightly smaller than unstained live cells. **(C)** Histogram depicting cells before incubation in the gate R1. Very few apoptotic cells are detected. **(D)** Histogram depicting cells after incubation in the gate R1. Apoptotic cells (brightly stained by TUNEL) are detected.

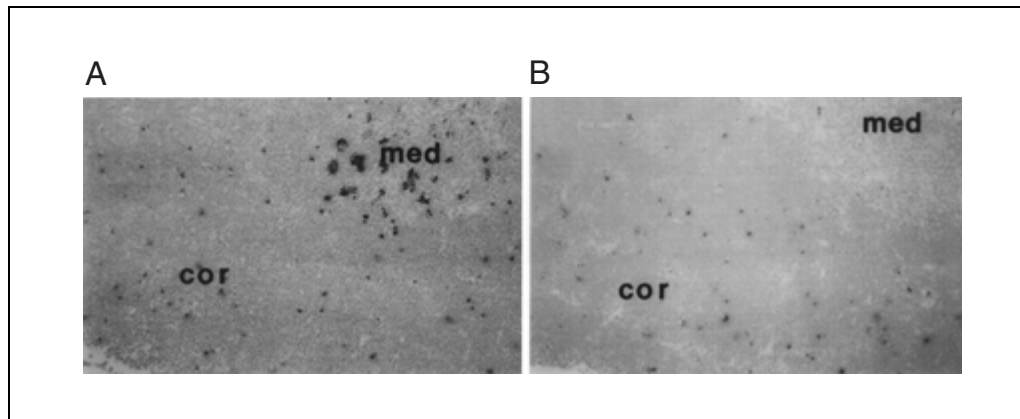


Figure 14.13.3 In situ detection of apoptotic cells in thymus sections by the TUNEL technique. Sections were counterstained with hematoxylin and photomicrographed at $200\times$ magnification (also see Anticipated Results). **(A)** Thymus from 5-week-old I-E⁺/V β 5 transgenic mouse. Small apoptotic cells are scattered in the cortex (cor) but are unusually prominent in the medulla (med) as large aggregates. **(B)** Thymus from 5-week-old I-E⁻/V β 5 transgenic mouse. This is similar to the normal mouse thymus; apoptotic cells are apparent in the cortex but rare in the medulla.

staining (i.e., at least overnight at room temperature). Sections can be stored for 2 to 3 weeks at 4°C. Paraformaldehyde fixation results in better staining than acetone fixation, probably because much higher proportions of low-molecular-weight DNA fragments are retained inside apoptotic cells with paraformaldehyde fixation. However, in sections from tissues with unusually large numbers of apoptotic cells, such as irradiated thymus, acetone fixation is adequate as apoptotic cells are easier to label.

Anticipated Results

Results of typical apoptosis studies are outlined in Table 14.13.1. There are no universal inducers of apoptosis available at the present time, so entirely new systems must be empirically studied.

As shown in Figure 14.13.1, flow cytometric analysis can distinguish live cells from apoptotic cells and quantitate live cells on the basis of FSC and PI uptake properties. Live cells should have a large FSC profile and be negative for propidium iodide uptake (as illustrated in Fig. 14.13.1, lower right quadrants of panels A and B). By contrast, apoptotic cells should have a smaller FSC profile, which results from a decrease in cell volume. They should also stain positive for propidium iodide if analyzed late in the death program when membrane viability has been lost (as illustrated in Fig. 14.13.1, upper right and left quadrants of panels A and B). The number of cells with decreased FSC profile that fail to exclude PI gives an impression of the number of dead cells; however, it

would be erroneous to calculate an exact number of dead cells on the basis of this parameter, because apoptotic cells fragment into multiple apoptotic bodies during the course of the cell-death process. Cellular debris and apoptotic bodies should have a very small FSC profile and may or may not take up propidium iodide (as illustrated in Fig. 14.13.1, upper and lower left quadrants of panel B). This analysis can be further used to measure the number of live cells of a subpopulation based on antibody binding.

Results from a typical flow cytometric analysis using TUNEL staining are shown in Figure 14.13.2. Freshly prepared normal mouse thymocytes contain very few detectable apoptotic cells (0.6%; see panel C). In contrast, incubation of the same cells in normal medium with FBS for 18 hr at 37°C causes considerable apoptosis (~33%; see panels A, B, and D). The apoptotic cells are slightly smaller than normal thymocytes but are difficult to distinguish on the basis of cell size alone (see panels A and B). Note that in panel A, which shows FSC/SSC profiles, both cell debris (consisting of particles much smaller than the bulk of thymocytes) and cell clumps (which are much larger) can be easily gated out from the analysis.

Photomicrographic results from in situ detection of spontaneously occurring apoptotic cells by the TUNEL method are shown in Figure 14.13.3. In this example, panel A depicts a section of thymus from a transgenic mouse strain produced by a cross between TCR V 5 and I-E⁺ mice. Apoptotic cells are visible in the cortex as scattered small cells, and are prominent in the medulla as large aggregates. The

apoptotic cells in the medulla reflect negative selection to endogenous mammary tumor virus antigens (Mtv 8,9) associated with I-E molecules. Panel B depicts a section of thymus from a transgenic mouse strain produced by a cross between V β 5 and I-E⁻ mice. Apoptotic cells are spread throughout the cortex but are largely absent in the medulla—a pattern of apoptosis that is also observed in normal mouse thymus (Surh and Sprent, 1994). The apoptotic cells in the cortex reflect lack of positive selection, as similar numbers of apoptotic cells are detected in the thymic cortex of MHC-deficient mice (Surh and Sprent, 1994). These data indicate that the TUNEL method is able to detect apoptosis occurring under in vivo physiological conditions.

Time Considerations

Determination of the apoptotic index using acridine orange and/or ethidium bromide staining takes ~2 to 3 min/sample. Separation of fragmented DNA from intact chromatin takes ~30 min. Isolation of DNA for agarose gel electrophoresis takes 2 to 48 hr depending on how much and how large the DNA in the sample is. Agarose gel electrophoresis requires 2 to 4 hr to perform. It takes ~1 hr to harvest 15 samples (in 45 wells including duplicates and controls; number will vary with each experiment) from a 96-well plate for flow cytometric analysis. Each sample takes ~1 to 2 min to run on the FACScan depending upon the amount of time for collection, so that an additional 1 to 1.5 hr are required for collection and data analysis.

In both the flow cytometric and in situ protocols, TUNEL staining takes 4 to 6 hr to complete once the cells or tissue sections are prepared. This estimate does not including the time required for flow cytometric analysis.

Literature Cited

Ashwell, J.D., Cunningham, R.E., Noguchi, P.D., and Hernandez, D. 1987. Cell growth cycle block of T cell hybridomas upon activation with antigen. *J. Exp. Med.* 165:173-194.

Bertin, J., Armstrong, R.C., Otilie, S., Martin, D.A., Wang, Y., Banks, S., Wang, G.H., Senkevich, T.G., Alnemri, E.S., Moss, B., Lenardo, M.J., Tomaselli, K.J., and Cohen, J.I. 1997. Death effector domain-containing herpesvirus and poxvirus proteins inhibit both Fas- and TNFR1-induced apoptosis. *Proc. Natl. Acad. Sci. U.S.A.* 94:1172-1176.

Boehme, S.A. and Lenardo, M.J. 1993a. Ligand-induced apoptosis of mature T lymphocytes (proapoptotic regulation) occurs at distinct stages of the cell cycle. *Leukemia* 7(Suppl. 2):S45-S49.

Boehme, S.A. and Lenardo, M.J. 1993b. Propriocidal apoptosis of mature T lymphocytes occurs at S phase of the cell cycle. *Eur. J. Immunol.* 23:1552-1560.

Chinnaiyan, A.M. and Dixit, V.M. 1997. Portrait of an executioner: The molecular mechanism of Fas/Apo-1-induced apoptosis. *Semin. Immunol.* 9:69-76.

Chinnaiyan, A.M., O'Rourke, K., Lane, B.R., and Dixit, V.M. 1997. Interaction of CED-4 with CED-3 and CED-9: A molecular framework for cell death. *Science* 275:1122-1126.

Clem, R.J. and Miller, L.K. 1994. Control of programmed cell death by the baculovirus genes p35 and iap. *Mol. Cell. Biol.* 14:5212-5222.

Cohen, J.J. and Duke, R.C. 1983. Glucocorticoid activation of a calcium-dependent endonuclease in thymocyte nuclei leads to cell death. *J. Immunol.* 132:38-42.

Coligan, J.E., Kruisbeek, A.M., Margulies, D.H., Shevach, E.M., and Strober, W. 2000. Current Protocols in Immunology. John Wiley & Sons, New York.

Critchfield, J.M., Racke, M.K., Zuniga-Pflucker, J.C., Cannella, B., Raine, C.S., Goverman, J., and Lenardo, M.J. 1994. T cell deletion in high antigen dose therapy of autoimmune encephalomyelitis. *Science* 263:1139-1143.

De Franco, A.L., Gold, M.R., and Jakway, J.P. 1987. B-lymphocyte signal transduction in response to anti-immunoglobulin and bacterial lipopolysaccharide. *Immunol. Rev.* 95:161-176.

Dive, C., Gregory, C.D., Phipps, D.J., Evans, D.L., Milner, A.E., and Wyllie, A.H. 1992. Analysis and Discrimination of necrosis and apoptosis (programmed cell death) by multiparameter flow cytometry. *Biochim. Biophys. Acta* 1133:275-285.

Duke, R.C. and Cohen, J.J. 1986. IL-2 addiction: Withdrawal of growth factor activates a suicide program in dependent T cells. *Lymphokine Res.* 5:289-299.

Duvall, E. and Wyllie, A.H. 1986. Death and the cell. *Immunol. Today* 7:115-119.

Fisher, G.H., Rosenberg, F.J., Straus, S.E., Dale, J.K., Middleton, L.A., Lin, A.Y., Strober, W., Lenardo, M.J., and Puck, J.M. 1995. Dominant interfering Fas gene mutations impair apoptosis in a human autoimmune lymphoproliferative syndrome. *Cell* 81:935-946.

Frey, T. 1995. Nucleic acid dyes for detection of apoptosis in live cells. *Cytometry* 21:265-274.

Gavrieli, Y., Sherman, Y., and Ben-Sasson, S.A. 1992. Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation. *J. Cell. Biol.* 119:493-501.

Goodnow, C.C. 1992. Transgenic mice and analysis of B-cell tolerance. *Ann. Rev. Immunol.* 10:489-518.

Gorczyca, W., Gong, J., and Darzynkiewicz, Z. 1993. Detection of DNA strand breaks in individual apoptotic cells by the in situ terminal

deoxynucleotidyl transferase and nick translation assays. *Cancer Res.* 53:1945-1951.

- Harmon, B.V., Corder, A.M., Collins, R.J., Gobe, G.C., Allen, J., Allan, D.J., and Kerr, J.F.R. 1990. Cell death induced in a murine mastocytoma by 42-47°C heating in vitro: Evidence that the form of death changes from apoptosis to necrosis above a critical heat load. *Int. J. Radiat. Biol.* 58:845-858.
- Hengartner, M.O. and Horvitz, H.R. 1994. Programmed cell death in *Caenorhabditis elegans*. *Curr. Opin. Genet. Devel.* 7:581-586.
- Hengartner, M.O., Ellis, R.E., and Horvitz, H.R. 1992. *Caenorhabditis elegans* gene ced-9 protects cells from programmed cell death. *Nature* 356:494-499.
- Holmes, K. and Fowlkes, B.J. 1991. Preparation of cells and reagents for flow cytometry. In *Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 5.3.1-5.3.11. John Wiley & Sons, New York.
- Hueber, A.O., Zornig, M., Lyon, D., Suda, T., Nagata, S., and Evan, G.I. 1997. Requirement for the CD95 receptor-ligand pathway in c-Myc-induced apoptosis. *Science* 278:1305-1309.
- Jacobson, M.D., Burne, J.F., and Raff, M.C. 1994. Programmed cell death and Bcl-2 protection in the absence of a nucleus. *EMBO J.* 13:1899-1910.
- Kerr, J.F.R., Wyllie, A.H., and Currie, A.R. 1972. Apoptosis: A basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br. J. Cancer* 26:239-257.
- Kishimoto, H., Surh, C.D., and Sprent, J. 1995. Upregulation of surface markers on dying thymocytes. *J. Exp. Med.* 181:649-655.
- Koury, M.J. and Bondurant, M.C. 1990. Erythropoietin retards DNA breakdown and prevents programmed cell death in erythroid progenitor cells. *Science* 248:378-381.
- Lenardo, M.J. 1991. Interleukin-2 programs mouse T lymphocytes for apoptosis. *Nature* 353:858-860.
- Lenardo, M.J. 1996. Fas and the art of lymphocyte maintenance. *J. Exp. Med.* 183:721-724.
- Liu, X., Zou, H., Slaughter, C., and Wang, X. 1997. DFF, a heterodimeric protein that functions downstream of caspase-3 to trigger DNA fragmentation during apoptosis. *Cell* 89:175-184.
- Martin, D. and Lenardo, M. 1998. Morphological, biochemical, and flow cytometric assays of apoptosis. In *Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W.S. Strober, eds.) pp. 3.17.1-3.17.39. John Wiley & Sons, New York.
- Miller, D.K. 1997. The role of the Caspase family of cysteine proteases in apoptosis. *Semin. Immunol.* 9:35-50.
- Mishell, B.B., Shiigi, S.M., Henry, C., Chan, E.L., North, J., Gallily, R., Slomich, M., Miller, K., Marbrook, J., Parks, D., and Good, A.H. 1980. Preparation of mouse cell suspensions. In *Selected Methods in Cellular Immunology* (B.B. Mishell and S.M. Shiigi, eds.) pp. 21-22. W.H. Freeman, New York.
- Nagata, S. and Golstein, P. 1995. The Fas death factor. *Science* 26:1449-1456.
- Nemazee, D.A. and Burki, K. 1989. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature* 337:562-566.
- Ormerod, M.G., Paul, F., Cheetham, M., and Sun, X.M. 1995. Discrimination of apoptotic thymocytes by forward light scatter. *Cytometry* 21:300-304.
- Otten, G., Yokoyama, W.M., and Holmes, K.L. 1995. Flow cytometry analysis using the Becton Dickinson FACScan. In *Current Protocols in Immunology* (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 5.4.1-5.4.19. John Wiley & Sons, New York.
- Raff, M.C., Barres, B.A., Burne, J.F., Coles, H.S., Ishizaki, Y., and Jacobson, M.D. 1993. Programmed cell death and the control of cell survival: lessons from the nervous system. *Science* 262:695-670.
- Ray, C.A., Black, R.A., Kronheim, S.R., Greenstreet, T.A., Sleath, P.R., Salvesen, G.S., and Pickup, D.J. 1992. Viral inhibition of inflammation: Cowpox virus encodes an inhibitor of the interleukin-1 beta converting enzyme. *Cell* 69:597-604.
- Rehmtulla, A., Hamilton, C.A., Chinnaiyan, A.M., and Dixit, V.M. 1997. Ultraviolet radiation-induced apoptosis is mediated by activation of CD-95 (Fas/APO-1). *J. Biol. Chem.* 272:25783-25786.
- Sellins, K.S. and Cohen, J.J. 1987. Gene induction by gamma-irradiation leads to DNA fragmentation in lymphocytes. *J. Immunol.* 139:3199-3206.
- Sellins, K.S. and Cohen, J.J. 1991. Hyperthermia induces apoptosis in thymocytes. *Radiat. Res.* 126:88-95.
- Shen, Y. and Shen, T.E. 1995. Viruses and apoptosis. *Curr. Opin. Genet. Devel.* 5:105-111.
- Skalka, R.E., Mattyasova, J., and Cejkova, M. 1976. DNA in chromatin of irradiated lymphoid tissues degrades in vivo into regular fragments. *FEBS Lett.* 72:271-274.
- Spector, M.S., Desnoyers, S., Hoepfner, D.J., and Hengartner, M.O. 1997. Interaction between the *C. elegans* cell-death regulators CED-9 and CED-4. *Nature* 385:653-656.
- Surh, C.D. and Sprent, J. 1994. T-cell apoptosis detected in situ during positive and negative selection in the thymus. *Nature* 372:100-103.
- Telford, W.G., King, L.E., and Fraker, P.J. 1994. Rapid quantitation of apoptosis in pure and heterogeneous cell populations using flow cytometry. *J. Immunol. Methods* 172:1-16.

- Vaux, D.L. and Strasser, A. 1996. The molecular biology of apoptosis. *Proc. Natl. Acad. Sci. U.S.A.* 93:2239-2244.
- Vaux, D.L., Cory, S., and Adams, J.M. 1988. Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells. *Nature* 335:440-442.
- Vaux, D.L., Weissman, I.L., and Kim, S.K. 1992. Prevention of programmed cell death in *Caenorhabditis elegans* by human bcl-2. *Science* 258:1955-1957.
- Wang, R., Murphy, K.M., Loh, D.Y., Weaver, C., and Russell, J.H. 1993. Differential activation of antigen-stimulated suicide and cytokine production pathways in CD4⁺ T cells is regulated by the antigen-presenting cell. *J. Immunol.* 150:3832-3842.
- Watanabe-Fukunaga, R., Brannan, C.I., Copeland, N.G., Jenkins, N.A., and Nagata, S. 1992. Lymphoproliferation disorder in mice explained by defects in Fas antigen that mediates apoptosis. *Nature* 356: 314-317.
- Wu, D., Wallen, H.D., and Nunez, G. 1997. Interaction and regulation of subcellular localization of CED-4 by CED-9. *Science* 275:1126-1129.
- Wyllie, A.H. 1980. Glucocorticoid-induced thymocyte apoptosis is associated with endogenous endonuclease activation. *Nature* 284:555-556.
- Wyllie, A.H. 1981. Cell death: A new classification separating apoptosis from necrosis. In *Cell Death in Biology and Pathology* (I.D. Bowen and R.A. Lockshin, eds.) pp. 9-34. Chapman & Hall, London.
- Wyllie, A.H. and Morris, R.G. 1982. Hormone-induced cell death. Purification and properties of thymocytes undergoing apoptosis after glucocorticoid treatment. *Am. J. Pathol.* 109:78-87.
- Zou, H., Henzel, W.J., Liu, X., Lutschg, A., and Wang, X. 1997. Apaf-1, a human protein homologous to *C. elegans* CED-4, participates in cytochrome c-dependent activation of caspase-3. *Cell* 90:405-413.

Key Reference

Kerr et al., 1972. See above.

Generalizes the distinction between programmed and pathological cell death based on differences in morphology and suggests the term apoptosis, derived from Greek and meaning "falling off;" as of leaves from trees.

Contributed by David Martin and
Michael Lenardo
National Institute of Allergy and Infectious
Diseases
Bethesda, Maryland

The authors thank Richard C. Duke, J. John Cohen, Stefen A. Boehme, Charles D. Surh, Hidehiro Kishimoto, and Jonathan Sprent, whose previous contributions to Current Protocols in Immunology (Coligan et al., 2000) formed the basis for portions of the present version of this unit.

Whole-Mount Histochemical Detection of β -Galactosidase Activity

UNIT 14.14

Seminal studies on the *lac* operon led to the isolation of β -galactosidase activity and the use of the *E. coli lacZ* gene as a fusion protein and reporter in bacteria (for review see Silhavy and Beckwith, 1985). Its first use as a reporter gene to determine the spatial and temporal expression in a more complex organism was in *Drosophila* (Lis et al., 1983). Subsequently, numerous reports documented its use in diverse experimental applications in all model organisms.

The combination of direct detection of endogenous gene expression by whole-mount in situ hybridization (UNIT 14.9) and the use of a reporter gene in regulatory sequence analyses using transgenics and chimeric studies results in a better understanding of the potential domain of function of a particular gene or cell lineage. Whole-mount in situ histochemical β -galactosidase activity staining is a relatively simple and highly sensitive method that yields a three-dimensional representation of gene expression. If the analysis is performed at different stages (e.g., during embryogenesis), the dynamic nature of endogenous gene expression, implicating possible function, and tissue- and region-specific promoters, may be visualized in a spatial and temporal manner.

The main procedure (see Basic Protocol) describes fixation and staining for β -galactosidase activity. It is based on Beddington and Lawson (1990), and has been successfully used on vertebrate embryos and tissue explants (some recent examples include: Collignon et al., 1996; Martineau et al., 1997; Wang et al., 1998; Pereira et al., 1999). To help achieve consistent results with larger samples, an Alternate Protocol details preparation of thick sections. Support Protocol 1 describes clearing of tissues for better resolution and Support Protocol 2 describes paraffin-embedding the whole embryo.

CAUTION: Glutaraldehyde, formaldehyde, sodium deoxycholate, potassium ferrocyanide and ferricyanide, dimethylformamide, Histoclear, benzyl benzoate and benzyl alcohol, and Permount are hazardous. Please consult manufacturer's guidelines for their proper handling and disposal; also see APPENDIX 1H.

WHOLE-MOUNT STAINING AND HISTOCHEMICAL DETECTION OF β -GALACTOSIDASE ACTIVITY

BASIC
PROTOCOL

Mouse and chicken embryos, tissues, or explants are isolated, partially fixed so as to not inactivate the β -galactosidase enzyme, washed extensively to remove fixative, thick sectioned if necessary (Alternate Protocol), and stained in the presence of a chromogenic substrate. Following photography of whole-mount stained tissues, samples may be embedded and sectioned to reveal cell and tissue-specific expression of the reporter (Support Protocol 2). It is strongly recommended that pilot studies be performed to assess and standardize optimal conditions for the particular experiment and tissue (see Critical Parameters and Troubleshooting).

Materials

- Mouse or chicken whole embryos, tissues, or explants
- Phosphate-buffered saline (PBS; APPENDIX 2), ice-cold
- β -galactosidase fixative (see recipe)
- β -galactosidase wash buffer (see recipe)
- β -galactosidase staining solution (see recipe)

In situ
Hybridization
and Immuno-
histochemistry

Contributed by Fred A. Pereira

Current Protocols in Molecular Biology (2000) 14.14.1-14.14.8

Copyright © 2000 by John Wiley & Sons, Inc.

14.14.1

Supplement 50

5-ml or 10-ml polystyrene or polypropylene capped tubes *or* 1.5- to 2-ml microcentrifuge tubes
Platform rocker
Sealable container for staining
Depression slides with coverslips or chamber slides
Dissecting microscope with fiber-optic lighting
Photomicrography equipment

NOTE: All steps are performed in 5-ml, 10-ml polystyrene or polypropylene capped tubes or in 1.5- to 2-ml microcentrifuge tubes in solutions (20:1 v/v reagent-to-tissue ratio) with gentle end-to-end rocking unless stated otherwise.

1. Dissect embryos or tissues in ice-cold PBS. Place in β -galactosidase fixative for appropriate length of time at room temperature.

Fixation time for tissues will depend on size, type, and whether the animal had been perfused with the β -galactosidase fixative. Whole-animal heart perfusion (see UNIT 14.1) should be performed on postnatal-to-adult animals prior to isolation of specific tissues for further fixation.

Since it is critical that tissues not be overfixed while other samples are being dissected, maintain small groups of embryos or tissues (5 to 10) in 100 mM sodium phosphate buffer, pH 8.0 (APPENDIX 2), until ready to process.

The following are some guidelines for fixation of mouse embryos and tissues: 5 to 10 min for explants and E6.5 to E7.5 embryos; 15 to 30 min for E8 to E9.5 embryos or tail cuts; 30 min to 1.5 hr for E10 to E12.5 embryos; and 2 hr for E13 to 14.5 embryos. E15.5 and larger tissues need to be dissected out or cut in half and fixed for 2 hr. Perfusion of newborn pups and older mice is required prior to dissection of particular tissues for staining.

2. Wash three times in β -galactosidase wash buffer for 15 min each time (for small embryos or tail cuts) or for 30 min each time (for E11 embryos and large tissues) at room temperature, with gentle end-to-end rocking.

It is critical to wash out the fixative sufficiently to allow the enzyme to function. Insufficient washing typically results in surface staining with minimal internal β -galactosidase staining.

3. Incubate tissues in β -galactosidase staining solution (at room temperature, 30°C, or 37°C depending on abundance of β -galactosidase expression) in a 5- or 10-ml capped tube or 1.5 to 2-ml microcentrifuge tube to prevent evaporation. Monitor staining by eye or using a dissecting microscope.

Xgal is the most popular substrate; however, Red-gal (6-chloro-3-indoyl- β -D-galactopyranoside; Research Organics) and Bluo-gal (halogenated indoyl- β -D-galactosidase; Life Technologies) have also been used successfully.

Signal strength depends on the abundance of β -galactosidase expression and appropriate fixation, washing, and buffer conditions. Development times of up to 48 hr are sometimes required, whereas abundant expression may be evident within 1 hr. If staining overnight, place at room temperature or 27° to 30°C to prevent excessive evaporation and possible crystal deposits. Crystals may form in the staining solution or on tissues if the phosphate buffer is incorrectly made. Optimal conditions will have to be determined in preliminary experiments.

*Use of pH 8.0 buffer minimizes the endogenous microsomal β -galactosidase activity, which has a pH optimum of 3 to 6 (MacGregor et al., 1987). Such activity is observed, if the pH drops, in E10 hindbrain, E14 and larger intestines, and kidneys. Use of the optimal pH of 7 to 7.4 for the *E. coli* enzyme is possible if the expected expression is not being investigated in tissues with significant microsomal activity. In addition, timely monitoring of staining*

and the use of nuclear lacZ (β -galactosidase containing a nuclear localization sequence) allows better discrimination of specific and nonspecific activity.

4. After desired signal-to-background is achieved, wash the tissues three times, each time for 30 min, in PBS.

A yellow discoloration may result if tissues are left in staining solution. Samples may be photographed at this stage. See Support Protocol 1 for long-term storage.

5. View and photograph the whole-mount, stained tissues through a dissecting microscope, using low-power objectives and fiber-optic lighting from the sides and top. For best results, place tissues in a depression slide with a coverslip, or a chamber slide, with sufficient PBS to just cover the tissue.

This eliminates reflections from overhead illumination.

PREPARING THICK SECTIONS OF LARGE TISSUES FOR β -GALACTOSIDASE STAINING

**ALTERNATE
PROTOCOL**

Use of thick vibratome sections (<1 mm) for activity staining is sometimes desirable. This step standardizes the protocol for routine and consistent analysis, especially of diffuse staining, in large embryos and tissues (such as adult mouse brain). Use of a cryostat to make frozen sections is possible (UNIT 14.2), but the ease and speed of the vibratome makes it the method of choice. Also, simply cutting the tissues into blocks with a sharp scalpel, for better penetration of solutions, is possible. This cutting would be done at the midpoint of the fixation period (see Basic Protocol, step 1). In the case of vibratome sectioning, one would immediately proceed with β -galactosidase staining (see Basic Protocol, step 3) after sectioning.

Additional Materials (also see Basic Protocol)

4% agarose (see recipe)
Cyanoacrylate glue (e.g., Superglue)
100 mM sodium phosphate buffer, pH 8 (APPENDIX 2)
Vibratome (e.g., model VT1000S; Leica)

1. Fix and wash tissues (see Basic Protocol, steps 1 and 2).
2. Place tissue in a 100-ml beaker and gently aspirate most of the wash buffer with a pipet. Add sufficient 4% intermediate-strength agarose to cover the tissue (when beaker is tilted to its side) at 37° to 40°C.

To avoid air gaps that can get trapped between the tissue and agarose, swirl the beaker to mix and coat the tissue well.

3. Place the beaker at a 45° angle at room temperature until the agarose has hardened (~10 min).
4. Trim the agarose block with a razor blade to square off the sides, leaving 2 to 3 mm on each side.
5. Using cyanoacrylate glue, affix the agarose block, in the appropriate orientation (e.g., for coronal sectioning), to the vibratome specimen-holder plate. Place the specimen-holder plate in the vibratome and fill the bath with 100 mM sodium phosphate buffer, pH 8.
6. Section at ≤ 1 mm (consult vibratome manual for specifics). Carefully remove slices from the vibratome bath and place in β -galactosidase wash buffer until ready to stain. Proceed to stain tissues (see Basic Protocol, step 3).

After staining, the thick, 1-mm sections may be resectioned in the vibratome, at 40 to 100 μ m, to reveal more subtle expression of the reporter. Alternatively, for cellular detail,

**In situ
Hybridization
and Immuno-
histochemistry**

14.14.3

samples may be secured in 3% agarose, and processed for paraffin and microtome sectioning at 10 μ m (see Support Protocol 2, step 1). View and photograph the stained thick sections (see Basic Protocol, step 5).

SUPPORT PROTOCOL 1

STORAGE AND TISSUE CLEARING

For analysis at a later time point, whole-mount stained samples (see Basic Protocol) should be fixed in formaldehyde for storage (up to 1 year). Samples may also be cleared to reveal β -galactosidase expression within internal structures. However, with thick tissues or embryos with bilateral expression, this may complicate the three-dimensional picture. In these situations, confocal microscopy would be valuable to give a better three-dimensional view.

Materials

Stained samples (see Basic and Alternate Protocol)
3.7% formaldehyde: dilute 37% formaldehyde 1:10 with PBS
Phosphate-buffered saline (PBS; APPENDIX 2)
50%, 70%, 80%, 95%, and 100% methanol
2:1 (v/v) benzyl benzoate/benzyl alcohol

1. Post-fix sample by soaking in 3.7% formaldehyde at 4°C (1 hr to overnight).
2. Wash several times in PBS and store at 4°C.

This procedure results in no appreciable loss of stain. Samples can be photographed (see Basic Protocol, step 5) prior to post-fixation, but post-fixation is recommended for long-term storage.

3. *Optional:* To clear tissue, dehydrate samples successively for 15 to 30 min each in 50%, 70%, 95%, and 100% methanol. Finally, place samples in 2:1 benzyl benzoate/benzyl alcohol in a polypropylene tube or glass vial (polystyrene will dissolve in the clearing agent). Allow the tissue to clear with gentle rocking (10 to 30 min). Proceed with photography (Basic Protocol, step 5).

Clearing can be performed any time after samples have been fixed, but the sample should first be photographed once, just in case the three-dimensional picture is too complicated.

Rehydration of the tissue (through a descending graded series of methanols to PBS) and storage in PBS or processing for sectioning is possible (see Support Protocol 2). Do not store in clearing solution since it dissolves the precipitate over time.

Alternatively, samples may be dehydrated and cleared in 1:1 glycerol/alcohol, but it is not as effective as the above solution.

SUPPORT PROTOCOL 2

PARAFFIN EMBEDDING, SECTIONING, AND COUNTER-STAINING

The cellular detail of β -galactosidase expression can be determined upon microsectioning the whole-mount stained tissues (see Basic Protocol and Alternate Protocol). This is often informative for revealing expression within internal structures such as the endothelium of blood vessels (Fig. 14.14.1, panel B) or in post-mitotic neurons of the brain (Fig. 14.14.1, panel C) which are not readily seen in whole-mount view.

Materials

Whole-mount stained embryos (see Basic Protocol)
3% agarose (see recipe)
Histoclear (National Diagnostics)
70%, 95%, and 100% ethanol

**Whole-Mount
Histochemical
Detection of
 β -Galactosidase
Activity**

14.14.4

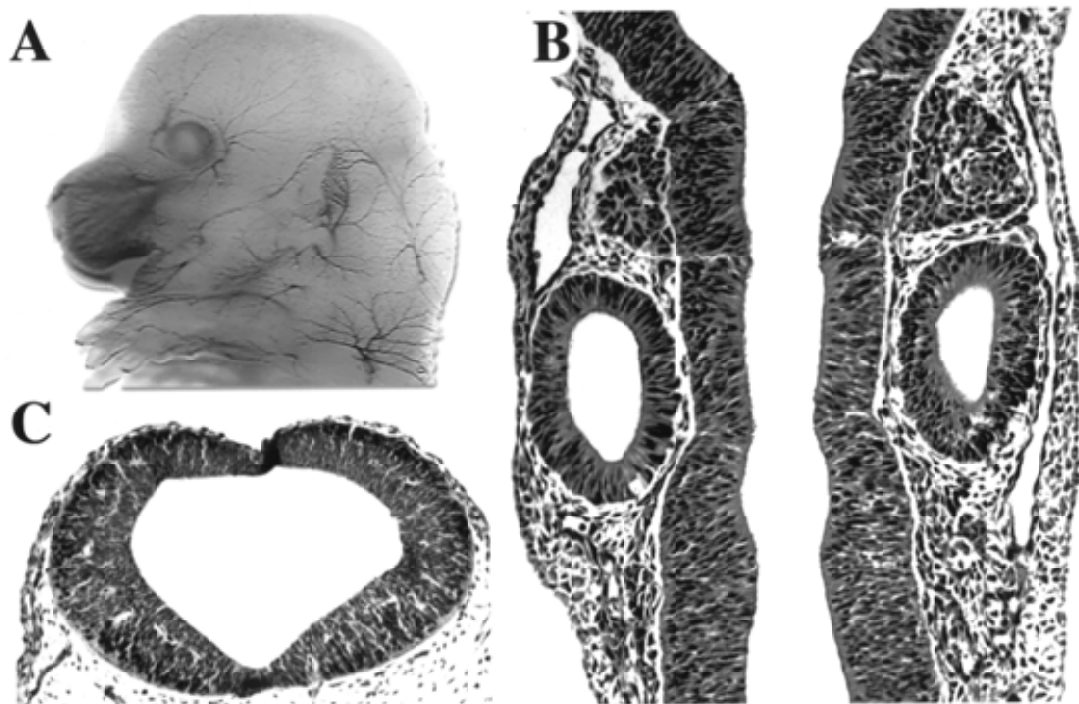


Figure 14.14.1 (A) Whole-mount view of an E13 embryo stained for β -galactosidase activity, driven by a neuronal enhancer, in peripheral nerves (Eric Nemoz-Gaillard, unpub. observ.). (B) Frontal section of *Tie1-lacZ* expression in endothelial cells of an E9.5 embryo (for detailed expression see Pereira et al., 1999, and Partanen et al., 1996). (C) Section through the midbrain of an E9.5 embryo showing β -galactosidase activity, expressed from the *BETA2/NeuroD* locus, marking post-mitotic neurons (for detailed expression see Naya et al., 1997). *This black and white facsimile of the figure is intended only as a placeholder; for full-color version of figure go to <http://www.currentprotocols.com/colorfigures>*

0.1% (w/v) nuclear fast red in 5% (w/v) aqueous aluminum sulfate: (Kernechtrot; Poly Scientific)

Permunt or aqueous mounting medium

Microtome

Subbed microscope slides (e.g., Superfrost Plus, Fisher)

42°C and 65°C ovens

Coplin jars or staining dishes

Photomicrography equipment

Additional reagents and equipment for paraffin embedding (UNIT 14.1)

Embed and section tissue

1. For tissue explants and embryos <E10.5: Embed in 3% agarose by placing tissues in a petri dish, removing excess solution, covering tissue with 3% agarose that has been cooled to 37° to 40°C, and allowing to harden. Trim the agarose for easy handling of these small samples. Carefully mark one corner with a small cut for easy orientation when embedding in paraffin.
2. To maintain tissue morphology, follow the paraffin embedding protocol (UNIT 14.1) to process and embed the stained samples, but substitute HistoClear for xylene in all appropriate steps.

HistoClear is preferred since it does not dissolve the Xgal precipitate as effectively as does xylene.

**In situ
Hybridization
and Immuno-
histochemistry**

14.14.5

Do not dehydrate samples to be cleared until ready to process and embed, since the precipitate is slightly soluble in ethanol.

3. Section tissue on a microtome at 10 μm or thicker. Place in a 42°C water bath and pick sections up onto subbed slides (e.g., Superfrost Plus; Fisher). Allow to dry.

Strong β -galactosidase expression can be readily visualized in 7- μm sections, but diffuse staining may require 10- to 15- μm or thicker sections prepared using a vibratome.

Dewax, coverslip, and examine

4. Place slides at 65°C for 30 to 60 min to completely dehydrate and fix tissue to slides.

Since sections are very thick, allow sufficient time for all steps.

5. Dewax three times, 5 min each time, in HistoClear.
6. For counterstaining, rehydrate through alcohols in Coplin jars or staining dishes as follows:

three times, 1 min each, in 100% ethanol
once, 1 min, in 95% ethanol
twice, 5 min each, in 70% ethanol
five dips in water.

If expression is weak or diffuse, no counterstaining is necessary and the tissue is simply coverslipped and viewed.

7. Counterstain 1 min in 0.1% nuclear fast red staining solution.
8. Remove excess stain with several changes of water.
9. Dehydrate slides as follows:

3 to 5 min in 70% ethanol
1 min in 95% ethanol
three times, 1 min each, in 100% ethanol.

10. Pass slides through HistoClear twice, 1 min each.
11. Coverslip with Permount or aqueous mounting media and dry ≥ 30 min before viewing.
12. Photograph and analyze the samples.

Analysis and photography of sections are optimal when using differential interference contrast or Nomarski microscopy but is satisfactory using bright-field microscopy (deep blue precipitate) or dark-field microscopy (pink precipitate).

REAGENTS AND SOLUTIONS

Agarose, 3% or 4% (w/v)

Melt intermediate-strength agarose (MetaPhor agarose; FMC Bioproducts) in 100 mM sodium phosphate buffer, pH 8 (APPENDIX 2) by heating in a microwave oven. Heat for intervals of 30 sec with gentle swirling to mix solution between intervals, until dissolved (melting temperature is $\leq 75^\circ\text{C}$). Cool to 37° to 40°C (gelling temperature is $\leq 35^\circ\text{C}$) in a water bath before use.

β-galactosidase fixative

0.2% (v/v) glutaraldehyde (from 25% stock; EM grade)
1.5% (v/v) formaldehyde (from 37% stock)
5 mM EGTA (from 100 mM stock, pH 8.0)
2 mM MgCl₂ (from 1 M stock)
100 mM sodium phosphate, pH 8.0 (*APPENDIX 2*)
Prepare fresh on day of use

β-galactosidase staining solution

Prepare the following stock solutions:

25 mg/ml 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside (Xgal) in dimethylformamide (store up to 1 year in dark at -20°C)

100 mM potassium ferrocyanide [K₄Fe(CN)₆] in 100 mM sodium phosphate buffer, pH 8.0 (see *APPENDIX 2* for buffer; store up to 1 year in dark at 4°C)

100 mM potassium ferricyanide [K₃Fe(CN)₆] in 100 mM sodium phosphate buffer, pH 8.0 (see *APPENDIX 2* for buffer; store up to 1 year in dark at 4°C)

At time of use, dilute the above stocks in β-galactosidase wash buffer (see recipe) to yield the following final concentrations:

1 mg/ml Xgal

5 mM K₄Fe(CN)₆

5 mM K₃Fe(CN)₆

Unused portions may be filtered, stored at 4°C, and reused for up to a week.

β-galactosidase wash buffer

2 mM MgCl₂ (from 1 M stock)
0.01% (w/v) sodium deoxycholate (from 1% w/v stock)
0.02% (v/v) Nonidet P-40 (from 10% stock)
100 mM phosphate buffer, pH 8.0 (*APPENDIX 2*)

COMMENTARY

Background Information

Reporter genes have been used to detect chimeric fusion proteins in transgenics and in tissue explants. The *E. coli lacZ* gene has proven to be an excellent reporter gene for many applications—from cell-lineage mapping to enhancer and promoter analyses to gene-trap analysis—in many systems (Hogan et al., 1994; Sullivan and Lo, 1997). The combination of ease of cloning (the first 26 amino acids are dispensable for activity; Muller-Hill and Kania, 1974), cytoplasmic or nuclear localization (Bonnerot et al., 1987), stability of and lack of deleterious effects of the β-galactosidase enzyme in vivo (Sanes et al., 1986), and the variety of substrates available, has resulted in high sensitivity and reproducibility of detection. The β-galactosidase expression reports, in situ, on the spatial, temporal, and tissue-specific expression. This reporter has proven indispensable in many cases for whole-animal or tissue analyses. Some recent examples are the analysis of EphB2 expression in arteries but not veins (Wang et al., 1998), asym-

metric nodal expression around the mouse node (Collignon et al., 1996), and ROSA26 cells to mark gonadal lineages in chimeric aggregates (Martineau et al., 1997).

Critical Parameters and Troubleshooting

There are a number of possible reasons for failure to observe staining. Pilot studies should be performed to standardize the procedure for the particular experiment and tissue. A standard rule of thumb is to use a 20:1 ratio of reagent to tissue, which helps to avoid minor procedural errors. If different samples are to be collected over an extended period, one should work with small batches and keep tissues in sodium phosphate buffer so as not to overfix the tissues prior to staining. If the β-galactosidase expression is weak, using pH 7 to 7.4 (optimal for bacterial β-galactosidase) will enhance the staining, but it is important to pay attention to specific signal in tissues (e.g., kidney and gut), to differentiate from background endogenous microsomal staining. Some protocols use 20 mM Tris-Cl,

pH 7.3, to stabilize the pH, which may also help to avoid crystal formation after overnight staining. Large tissues should be cut (hemi-sect or vibratome thick sections) to allow penetration of reagents. To avoid dissolution of weak staining, do not store in alcohols.

Anticipated Results

Strong expression will be visible within an hour with typical results (panel A in Figure 14.14.1) yielding a robust deep blue to magenta stain. Weak or diffuse staining appears light blue and may only be visible upon close examination using a dissecting microscope and appropriate lighting. Upon sectioning, cellular detail is evident (panels B and C in Figure 14.14.1) but some strong nuclear staining may leach in to the cytoplasm or diffuse into the surrounding interstitial space. In summary, β -galactosidase is a proven reporter enzyme that is highly specific, sensitive, and useful in many applications from lineage tracing to gene regulatory sequence analysis.

Time Considerations

Processing of tissues from dissection through fixation and washing to staining (Basic and Alternate Protocol) can be performed within 5 to 7 hr, but the staining time varies depending on the strength of the promoter driving the *lacZ* gene. Staining times of 48 hr or more may be necessary for weak promoters/enhancers. Paraffin embedding, sectioning, and counterstaining sections of the whole-mount stained tissues will take at least a couple of days. In all, a result may be obtained within 2 to 3 days from isolation of tissues to viewing the whole-mount stained tissues. Another 2 to 3 days will be required for processing and microsectioning the tissues to view cellular detail of expression.

Literature Cited

Beddington, R.S.P. and Lawson, K.A. 1990. Clonal analysis of cell lineages. *In* Postimplantation Mammalian Embryos: A Practical Approach (A.J. Copp and D.L. Cockroft, eds.) pp. 267-292. IRL Press, Oxford.

Bonnerot, C., Rocancourt, D., Briand, P., Grimber, G., and Nicolas, J.F. 1987. A β -galactosidase hybrid protein targeted to nuclei as a marker for developmental studies. *Proc. Natl. Acad. Sci. U.S.A.* 84:6795-6799.

Collignon, J., Varlet, I., and Robertson, E.J. 1996. Relationship between asymmetric nodal expression and the direction of embryonic turning. *Nature* 381:155-158.

Hogan, B., Beddington, R., Constantini, F., and Lacy, E. 1994. *Manipulating the Mouse Embryo*, 2nd Ed. Cold Spring Harbor Press, Cold Spring Harbor, New York.

Lis, J.T., Simon, J.A., and Sutton, C.A. 1983. New heat shock puffs and beta-galactosidase activity resulting from transformation of *Drosophila* with an hsp70-lacZ hybrid gene. *Cell* 35:403-410.

MacGregor, G.R., Mogg, A.E., Burke, J.F., and Caskey, C.T. 1987. Histochemical staining of clonal mammalian cell lines expressing *E. coli* beta galactosidase indicates heterogeneous expression of the bacterial gene. *Somatic Cell Mol. Genet.* 13:253-265.

Martineau, J., Nordqvist, K., Tilmann, C., Lovell-Badge, R., and Capel, B. 1997. Male-specific cell migration into the developing gonad. *Curr. Biol.* 7:958-968.

Muller-Hill, B. and Kania J. 1974. Lac repressor can be fused to beta-galactosidase. *Nature* 249:561-563.

Naya, F.J., Huang, H.P., Qiu, Y., Mutoh, H., DeMayo, F.J., Leiter, A.B., and Tsai, M.J. 1997. Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/neuroD-deficient mice. *Genes Dev.* 11:2323-2334.

Partanen, J., Puri, M.C., Schwartz, L., Fischer, K.D., Bernstein, A., and Rossant, J. 1996. Cell autonomous functions of the receptor tyrosine kinase TIE in a late phase of angiogenic capillary growth and endothelial cell survival during murine development. *Development* 122:3013-3021.

Pereira, F.A., Qiu, Y., Zhou, G., Tsai, M.J., and Tsai, S.Y. 1999. The orphan nuclear receptor COUP-TFII is required for angiogenesis and heart development. *Genes Dev.* 13:1037-1049.

Sanes, J.R., Rubenstein, J.L., and Nicolas, J.F. 1986. Use of a recombinant retrovirus to study post-implantation cell lineage in mouse embryos. *EMBO J.* 5:3133-3142.

Silhavy, T.J. and Beckwith, J.R. 1985. Uses of lac fusions for the study of biological problems. *Microbiol. Rev.* 49:398-418.

Sullivan, R. and Lo, C.W. 1997. Histochemical and fluorochrome-based detection of β -galactosidase. *In* Methods in Molecular Biology, Vol. 63: Recombinant Protein Protocols: Detection and Isolation (R. Tuan, ed.), Humana Press Inc., Totowa, New Jersey.

Wang, H.U., Chen, Z.F., and Anderson, D.J. 1998. Molecular distinction and angiogenic interaction between embryonic arteries and veins revealed by ephrin-B2 and its receptor Eph-B4. *Cell* 93:741-753.

Contributed by Fred A. Pereira
Baylor College of Medicine
Houston, Texas

Overview of Image Analysis, Image Importing, and Image Processing Using Freeware

UNIT 14.15

Quantitative image analysis turns microscopic data that might otherwise be merely descriptive into reliable mechanistic information. Such reliable structural information is a significant part of any biological investigation. "Structural information" refers to the patterns, shapes, contours, and relationships of objects made visible by their mass, as by staining, or by an electronic/computational technique applied to a biological specimen. Reliability of interpretation often depends on the statistical analysis of multiple images of the same structure. Great strides have been made in recent years in the ability to analyze multiple images quantitatively, due to the ease of acquiring images in digital formats and exporting them to various software packages. The types of analysis most often used involve: (1) elimination of background noise; (2) averaging of multiple specimens by superimposition to determine structural regions in common; (3) morphometric analysis of common measurable features obtained from many similar specimens of the same structure; (4) analysis of expression patterns of gene products over time or after an experimental treatment; (5) subcellular or tissue localization of gene products; and (6) quantitative measurements of rates of movements of cells or subcellular structures or molecules. Image analysis is also used to compare various quantitative aspects of biochemical data when such data are obtained in a visual format, such as a gel or blot. While this unit will focus on the analysis of microscopic material, some of the same tools and programs can also be useful when applied to images of biochemical data. The Commentary section of this unit will present an overview of the types of analysis that have been useful in the past as well as comments on emerging applications (see Background Information).

To maintain the value of the unit for a wide variety of investigations, the focus will be on basic techniques common to all biological applications for extracting reliable information from digital images, rather than on the equipment and software used to do this. Brief step-by-step protocols for importing data and for the first steps of data manipulation in the most common analytical applications, using two different freeware applications, NIH image (see Basic Protocol) and ImageJ (see Alternate Protocol), are provided.

IMAGE PROCESSING USING NIH IMAGE 1.62

The following is a protocol to import and process digital images for analysis or for production of figures using NIH Image 1.62, a freeware application available from the NIH at <http://rsb.info.nih.gov/nih-image/>. NIH Image is an image-processing program in the public domain designed for Macintosh computers. Digital images acquired in any format can be imported into NIH Image 1.62 if they are saved in a TIFF format. Once imported, NIH Image can be used to measure area, mean, centroid, and perimeter of user-defined regions of interest. The program includes tools for automated particle analysis and for measuring particle path lengths and angles. Density calibrations as compared to standards with user-specified units are also possible. Analyses can be printed, exported to text files, or copied to the clipboard for export to other compatible software.

NIH Image is most useful for black-and-white images. It can also be used to process such images by averaging to reduce background, by adjusting contrast and brightness, by smoothing pixelation, and by cutting and pasting to arrange a gallery of similar specimens for comparisons. NIH Image provides useful tools for editing and transforming digital

**BASIC
PROTOCOL**

**In Situ
Hybridization
and Immuno-
histochemistry**

Contributed by E.L. Bearer

Current Protocols in Molecular Biology (2003) 14.15.1-14.15.9

Copyright © 2003 by John Wiley & Sons, Inc.

14.15.1

Supplement 63

video images in black and white obtained from confocal microscopy. NIH Image can also be used to transform analog video to digital sequences for archiving. Some of these applications require additional hardware, including the appropriate computer ports to plug in the analog device.

This protocol covers the step-by-step process used to import various types of digital data into NIH Image to process and edit video sequences. NIH Image is capable of many other processes, and detailed instructions for these can be obtained from the NIH Image Manual at <http://rsb.info.nih.gov/nih-image/manual/>.

Materials

Color-capable Macintosh computer with ≥ 2 Mb of free RAM and the following programs:

MacOS v. 7 or later, or OS X capable of running programs in Classic mode
Web browser, e.g., Netscape or MS Internet Explorer
Stuffit Expander
Adobe Photoshop

Digital images and image sequences: these can be translated from a wide variety of file types, although TIFF formats are easiest to work with; importing from many other document types is handled through an Input-Output Macro that downloads with NIH Image

Download freeware

1. Go to the Web site <http://rsb.info.nih.gov/nih-image/>, download NIH Image 1.62, and decompress (“Unstuff”) the file onto the desktop.
2. Restart the computer.
3. On the desktop, highlight the icon for NIH Image application (a black microscope).
4. Under the File menu, select Get Info, then select Memory from the submenu which appears. Set Minimum Size to 75,000, and set Preferred Size to 100,000.

Smaller memory allocations will work for images that use less RAM, but the author of this unit has found that some operations can fail, incomplete loading of digital sequences can occur, or sequences can be loaded in the wrong order. For very large digital sequences, even more memory may be needed.

5. Open the application by double-clicking on the icon.

Import or open the image to process or analyze

To open Bio-Rad confocal files, electron micrographs from a scanner and some other specific formats

- 6a. Use the Input/Output Macro as follows. Choose Special from the menu bar at the top, then select Load Macros from the resulting submenu from the NIH Image folder. When the dialog box appears, open the Macros folder that downloaded automatically when NIH Image was loaded and which will be found inside the NIH Image folder. Scroll down within the Macros folder and highlight Input/Output macros.

Under Special, a list of file types that can be imported will now be seen. These include TIFF files, electron-microscopic images from Siemens, and confocal files from Bio-Rad.

- 7a. To import Bio-Rad confocal digital sequences or single frames, select Import BioRad 600 Z Series. In the dialog box, select the folder containing the Bio-Rad image sequence to be opened and scroll down in the window to find the specific file. Select Custom 512 and “8-bit” (the Bio-Rad will give a single icon for the entire image sequence).

- 8a. To import files from a different confocal, e.g., Zeiss images saved as TIFF sequences, select Open All; in this case, all the images in a sequence must be placed into a single folder.

To open TIFF files

- 6b. First, attempt to open the file by selecting Open from the File menu.
- 7b. If the file cannot be opened as in step 6b, select Import from the File menu. In the pop-up menu that appears, select TIFF and Custom 512 × 512 (or other image pixel size as necessary under the Set button).

It is possible to open all the files in a folder by selecting Open All.

- 8b. To work with Metamorph or Zeiss images, save them in the TIFF format in the Zeiss LSM program by selecting Export from the File menu in that program. Open the saved TIFF as in step 6b or 7b, above.

It is also possible to download the program Browser for free from the Zeiss Web site. This smaller program has only some of the applications included in the full Zeiss LSM program, but can help with saving images in alternative formats.

For images captured in two or more channels, it is better to save each channel separately and then work with them in ImageJ (see Alternate Protocol).

Manipulate the images

9. Upon opening the file, allow the image sequence to load onto the screen.
10. To play through the sequence, open the Stacks menu and select Animate. Click the mouse on the image to stop the playback. To toggle through the images in the forward direction, click on the “greater-than” key (>) without pressing the Shift key. To toggle through in the reverse direction, click on the “less-than” key (<) without pressing the Shift key.

The keyboard numbers can control the speed, with 1 being the slowest and 9 being the fastest.

11. To work with individual frames, open the Stacks menu and select Stacks to Windows, which will cause the images in the sequence to be posted on the screen individually.

Using this view, it is possible to delete an image or process any image in the stack.

IMPORTANT NOTE: *Be careful to put the images back in order on the screen before making them into a stack (see step 12).*

12. To put a series of individual images into a sequence, open the Stacks menu and select Windows to Stacks. Make the sequence as short as desired by deleting images. Add letters/arrows to the first image in a sequence, or to each image in the windows.

13. To crop or process an entire sequence, load the macro Stacks (Special>Load Macros) menu and select Stacks from the Macro folder in the popup window; a new set of commands appears in the scroll-down menu under Special.

Most of these are self-explanatory. The Crop command is one of the most useful, as it will make it possible to crop all images in the stack to exactly the same size.

14. To create a series of cropped images to present as still frames for publication as a figure, or for analysis of a printout, use the rectangular tool from the tool bar and draw the region of interest (ROI) on each frame of the sequence when it is in the stack mode on the screen. For example, if every sixth image is desired, draw the square on the first image, copy the ROI and paste into a new file (typically, at this stage, the author uses Adobe Photoshop to create the new file to paste in the copied image). Without moving the rectangular tool, toggle five frames, and copy the same ROI from

the sixth frame. Continue until all images in the series have been pasted into the new still-frame document.

Most digital cameras and confocals produce images with 72-bit resolution. This does not produce attractive images for publication. To improve the bit mapping in Photoshop, open the Image menu and select Image Size. A resolution of 400 is usually more than enough to produce a smooth figure. Photoshop's Gaussian Blur option (selected by opening the Filter menu and selecting Gaussian from the Blur submenu that appears) equalizes pixels at a radius that can be selected by the user. This will smooth the sharp edges of adjacent pixels where they differ by large contrast in gray scale between pixel edges.

15. Perform frame averaging by merging digital images.

Background subtraction is best performed during image capture. However, some frame averaging can be done by merging digital images. Frame averaging decreases random variations in pixel intensities.

Merging digital images after capture is time-consuming. The author uses Adobe Photoshop; in that program, open two or more frames to be averaged. Paste one image on top of the other. Open the Layers palette by selecting Show Layers from the Window menu. In the Layers palette, set opacity at 50% for each of two images, 25% for each of four images, etc. One can adjust the opacity levels to check that overlay is exact. It is best to align a stable object common to all frames. Select Merge Visible from the Layer menu at the top of the screen to compile the image.

16. Save the new sequence in NIH Image with a different name by opening the File menu and selecting Save As.

If the QuickTime radio button is selected in the Save As dialog box, encoded information may be lost when transporting back into TIFF. QuickTime format is very useful as it can be linked into Powerpoint for presentations or posted on the Web. In the Save As QuickTime pop-up menu, select Video and Color under Compressor, and High or better on the Image Quality slider bar. The optimum speed depends on how many images are in the stack and how fast the change is of the object under study. Generally, a 10- to 100-fold increase in the speed of capture is useful. For images captured every 10 sec, the author uses 6 to 10 frames/sec. Marking the video with a "keyframe" at regular intervals allows one to do manipulations of the sequence in Adobe Premiere, a movie-editing software package, or in Flash by Dreamweaver. Every 10 frames is often a useful frequency.

**ALTERNATE
PROTOCOL**

IMAGE PROCESSING USING ImageJ

ImageJ is a more sophisticated image processing system than NIH Image (see Basic Protocol) and can handle color images and superimposition of digital sequences. ImageJ can be downloaded from the NIH Web site at <http://rsb.info.nih.gov/ij/>.

ImageJ has additional capture macros that can accept data directly from a camera, such as the Hamamatsu Orca 12-bit camera, and from Scine frame grabbers. It also has TWAIN for direct download from most image scanners. It can calculate surface area and line distances, measure angles, and obtain radial profile plots. ImageJ can identify edges and measure changes in the Analyze>Measure descriptors menu. Useful macros include Combine Images to Create Mosaic, which allows digital background subtraction, and Image Calculator Plus, which can combine images of different sizes in different formats.

The ImageJ Web site lists a number of other sites where additional image processing freeware is available. An interesting site is <http://bigwww.epfl.ch/thevenaz/differentials/>. The Image Differentials plug-in available at that site allows one to perform several differentiation operations on images seen as continuous functions, even though they are stored as discrete pixels. The general principle is to construct the continuously defined function

$$f(x,y) = \sum_{k,l} c[k,l]\phi(x-k, y-l)$$

As for NIH Image (see Basic Protocol), memory has to be set at a high value (at least 100,000) if digital sequences of 100 or more frames are to be processed. A caveat is that Java allocates memory from the Systems heap, which means that too much memory allocated to ImageJ can reduce the amount of memory available for loading images, but without enough memory, one can run into error messages when processing QuickTime image series. It is therefore necessary to monitor memory under the Apple menu when running ImageJ and make alterations as necessary according to the computer and image sequence with which one is working.

Materials

Color-capable Macintosh computer or PC with 64 Mb (256 Mb recommended) of free RAM and the following programs:

MacOS v. 7.6 or higher or Windows 95 or higher with Java Runtime Environment 1.1.3

MRJ 2.1 (download from <http://www.apple.com/java/>)

QuickTime for Java (installed with Mac OS 9.04 or higher, or can be installed from QuickTime 1 "custom" Install menu)

1. Once ImageJ is loaded, open the application and open files from within it. If images have been saved as QuickTime files from the original sequences, open by selecting Import from the File menu. To open sequences saved as TIFF, first put all the images into one folder, then open from inside ImageJ by selecting Import from the File menu, then selecting Image Sequence and opening a file within the folder, which will in turn open the whole sequence.
2. Become familiar with the most commonly used image processing procedures in ImageJ (main menu items listed first, then submenus):

Image > Adjust > Brightness/Contrast

Process > Subtract Background (this requires loading of a background image that is devoid of any structures of interest)

Process > Filters > Gaussian Blur

Analyze > Measure

3. Rotate digital sequences by selecting Rotate from the Image menu for export to QuickTime, which can then be played in Powerpoint presentations or uploaded onto the Web in the orientation that makes most sense in comparison to other image sequences or to a model/diagram.

In ImageJ, if a processing routine is applied to a digital sequence, one will be stuck with the results (i.e., this is destructive editing). Therefore, one should always keep a file of the original data separate from the sequence being processed, and save often to preserve desired changes in case subsequent processing is not useful.

4. Save the new image sequences by selecting Save As from the File menu. For a TIFF series, select Image Sequence from the submenu that appears. For QuickTime, select options from the popup menu:

Compressor: Video

Color

Quality: Medium to High

Timing: Frames per Second 6-10

Keyframe every 10

Further processing and analysis can be performed in ever increasingly complex software applications.

COMMENTARY

Background Information

Image analysis to determine the structure of molecules and particles

Prior to the advent of digital image processing, various manual manipulations were required to extract quantitative information from structural data. The term “structural data” in this context refers to photographic or digital images of biological specimens, ranging from whole animals to tiny single molecules. For example, diffraction patterns can be obtained by passing laser light through film containing images of subcellular structures displaying repetitive patterns resembling crystalline arrays, such as the catalase crystal in the peroxisome (Bearer and Orci, 1986), the acetylcholine receptor channel (Unwin, 2002), and the protein coat of viruses (DeRosier and Klug, 1972). Reconstruction of the “ideal” crystalline structure is obtained by solving the diffraction pattern in reciprocal space. Another approach superimposes multiple electron microscopic images of the same structure, which “averages” the image, decreasing the intensity of nonrepeating structures, including those of unrelated background particles or of the movable parts within a fluid region of the structure. This technique has commonly been used in the analysis of electron micrographs. Multiple exposures of the same photographic paper with different images or rotations of the same image superimposed upon itself identifies repeating structures and internal symmetries within a structure. This technique was used to discover the eight-fold symmetry of the endothelial fenestral diaphragm, responsible for the passage of nutrients from blood to tissue in human organs (Bearer and Orci, 1985).

Statistical analysis of biological specimens, also known as “morphometrics,” involves compiling measurements of multiple specimens of the same structure and then mathematically deriving the average size and other statistical values such as standard deviation within the population studied. Such mathematical compilations rely on the acquisition of a large number of measurements from multiple specimens, which until very recently was performed by hand. However, only relatively small sample sizes are easily analyzed by hand, since this requires a large time investment.

Recent advances in structural determination of complex cellular components not amenable to crystallization include tomographic recon-

struction of electron-microscopic images. In this approach, images are captured from thick sections of biological specimens embedded in either plastic or vitreous ice using an intermediate-voltage electron microscope. The section is photographed at various angles of tilt, and the 3-D structure assembled digitally. The structure of the neuronal synapse is yielding to this approach (Harlow et al., 1998), which was recently cited as one of the ten top scientific advances in 2002 by the journal *Science* (News and Editorial Staff, 2002).

New technology exploiting digital image processing includes laser scanning and two-photon confocal microscopy, which allows elimination of the out-of-focus light from fluorescently labeled specimens. Confocal microscopy allows optical sectioning and thereby the acquisition of sharp images of objects deep within thick specimens (White et al., 1987). In confocal microscopy, the image is filtered before capture. Confocal microscopy relies on monochromatic laser excitation of fluorescent tags, and collects only in-focus emission using a pinhole to block out-of-focus emitted light. In two-photon confocal imaging, a recent advance, excitation and collection occurs only from the same optical section of interest, thus prolonging the length of time that a living biological specimen can be observed without laser damage (Kaeck et al., 2001; Meyer et al., 2003). New computational approaches (Meta and Multi-Tracker options in the Zeiss 510 LSM package) narrow the bandwidth of detected emitted light, thus providing the possibility of imaging multiple fluorochromes with the same excitation and even only slightly differing emission spectra in the same specimen. An alternative way to eliminate out-of-focus light is deconvolution, which involves digital processing of a series of images obtained by focusing up and down in the same microscopic field. In this approach, out-of-focus light is eliminated digitally after image capture (Hiraoka et al., 1990). Deconvolution requires a motorized stage to collect a series of images in the *z*-plane (Kam et al., 2001). These images are then processed digitally to extract in-focus structure and detail and eliminate fuzz from out-of-focus or autofluorescent emissions. The equipment is less costly than that used for confocal, but collection times are longer, and thus deconvolution is not as suitable for living specimens undergoing rapid changes. Both of these approaches allow visualization of struc-

tures by fluorescence microscopy that would not be resolved by conventional microscopy. Useful texts on microscopy include Matsumoto (1993) and Inoué and Spring (1997). Microscopy courses at the Marine Biological Laboratory in Woods Hole, Massachusetts (<http://www.mbl.edu>) are also excellent for viewing demonstrations of the latest equipment.

Video microscopy, pioneered by Robert Allen in the late 1970s using Nomarski optics (differential interference contrast or DIC), is capable of detecting structures that are theoretically below the resolution of the light microscope if lenses with high numerical apertures are used (Allen et al., 1979; Hayden and Allen, 1984; Brady et al., 1985). Such lenses have narrow focal planes, and often have short working distances as well; thus, they have not yet proven useful for imaging objects deep within thick specimens. However, like confocal microscopy, video imaging with either DIC or fluorescent tags can be performed on live specimens. Analysis of such sequence data can provide information about structural changes over time, such as the rates of changes in cellular surface contours (Zhang et al., 2002), the velocities of movement of intracellular particles such as viruses (Bearer et al., 2000), and even the dynamic behavior of purified cytoskeletal filaments during interactions with motors or severing factors (Vale et al., 1985; Bearer, 1991). In this case, microscopy becomes an assay for enzymatic activity.

An Open Microscopy Environment (OME) is being developed as an informatics solution for storage and analysis of optical microscope data (Swerdlow et al., 2003). This public resource hopes to provide image analysis automation, modeling tools, and large sets of data for mining biological material. The goal is for data to be stored as an XML-encoded standard file which can be uploaded and processed by any software. The URL for this project <http://www.openmicroscopy.org>.

Image analysis to determine clues of function from subcellular location and behavior

In addition to providing information about the exact measurements of the “ideal” for a particular structure, image analysis is also useful to discover expression patterns, location within cells, and dynamic behaviors of molecules both large (proteins and RNA) and small (ATP and calcium ions). The expression patterns and behaviors of proteins inside cells

detected by microscopy complement those detected by biochemical and molecular approaches, and readily provide functional information about a protein or gene product.

Because the genome projects have identified a large number of novel genes whose function is not easily predicted from DNA sequence, functional clues derived from structural analysis have become increasingly fashionable. Morphological data are being collected in high-throughput screening assays to identify key proteins from large data sets. Such key proteins can then be selected for analysis by more time-consuming approaches in subsequent experiments.

Strategies using localization to derive functional information about individual proteins and thereby select interesting proteins from large data sets were originally developed using antibodies to localize the protein (Bearer and Alberts, 1988; Miller et al., 1989). Antibody staining provides evidence concerning the type of cell within a tissue that expresses the protein and the subcellular location of that protein in cells. In conjunction with other stains, including antibodies, against other cellular proteins whose function is known, the contribution of the unknown protein to a particular aspect of cellular behavior can be inferred.

In addition to antibody staining, several other methods have been developed to detect gene expression. In situ hybridization (ISH) to detect mRNA expression speeds up the analysis because it uses the DNA as a probe and bypasses antibody production, which is time-consuming and not always reliable. However, results from ISH do not produce information about protein function, only about relative expression between cells in a tissue.

Another method, transfection and overexpression of fusions of the target protein with tags such as green fluorescent protein (GFP), allows observation of the dynamic behavior of a protein in the living cell. This method has most commonly been applied to individual proteins once they are determined to be of interest and is not a commonly used screen, as construction of the GFP fusion takes time. Another powerful approach is to label a protein whose function is known with GFP and then observe the effect of overexpression, deletion, or mutation of the unknown protein on the GFP-labeled structures. Finally, labeling of a protein within a larger complex of proteins allows microscopic tracking of the complex within cells, a strategy used to study the movement of viruses within cells (Bearer et al., 2000).

Questions that microscopic analyses usually address include: (1) the timing of expression changes after stimulation, (2) the types of cells within a multicellular tissue in which the gene product is expressed, and (3) the location of the gene product within the cell and its association with other cellular structures of known function. The function of an unknown protein or small molecule can be assessed using substrate cells that are transfected with some other protein whose function is already known, and monitoring the effect of the novel protein on the behavior of the known protein. This approach has proven very useful in identifying proteins whose function affects the secretory pathway, such as Golgi markers (Lippincott-Schwartz, 2001; Kreitzer et al., 2003) or the actin cytoskeleton (Westphal et al., 1997; Fischer et al., 1998).

While all of these types of structural expression analysis have in common the use of the microscope to obtain images for further analysis, the exact method and interpretation of results differ depending on the question being addressed. Microscope assays can be divided into two subsets: those acquired from living specimens in real time or by time lapse, and those acquired from fixed specimens. These assays use a variety of different microscope technologies and experimental methods. This is a fast-moving area of research, and specific applications and equipment are likely to have evolved between the writing of this unit and its appearance in print.

Critical Parameters

Image analysis is only as good as the images being analyzed. It is often a fruitless waste of time to use digital image processing to improve poor images. Therefore, before embarking on an analytical or processing project, images should be carefully scrutinized for: (1) focus; (2) good signal-to-noise ratios; (3) reproducibility; and (4) background contamination. Artifacts to be avoided are those from: (1) laser or other light damage; (2) aggregation of fluorochrome marker; (3) nonphysiologic localizations of fusion proteins; (4) movement of the stage during collection of video images; and (5) other effects of the experiment on the cell that might damage it or perturb its normal processes.

Scientists engaged in monitor-based digital processing tasks should ensure that the monitor's colors, brightness, and contrast are set appropriately. Instructions for this come with the Adobe Photoshop package, and must be

tailored to the individual monitor being used. Printing of processed images requires that the printer be coordinated with the monitor as well. This is dependent on the manufacturer's specifications for the particular printer and monitor being used. Monitor-based processing requires a low-light work area and a maximum-resolution screen.

Time Considerations

Digital image processing and analysis are usually more time-consuming than expected. The better the original image(s), the less time is required for processing. The type of data being extracted also define the length of time that must be invested. If software routines are available, then applying them can be fast once a sequence of steps is defined. It usually takes a graduate student a week to learn a routine, then a half-day to a day to make a processed QuickTime movie from a preselected, excellent series of images.

Literature Cited

- Allen, R.D., Zacharski, L.R., Widirstky, S.T., Rosenstein, R., Zaitlin, L.M., and Burgess, D.R. 1979. Transformation and motility of human platelets: Details of the shape change and release reaction observed by optical and electron microscopy. *J. Cell Biol.* 83:126-142.
- Bearer, E.L. 1991. Direct observation of actin filament severing by gelsolin and binding by gCap 39 and CapZ. *J. Cell Biol.* 115:1629-1638.
- Bearer, E.L. and Alberts, B.M. 1988. Novel actin-binding proteins associated with platelet activation. *FASEB J.* 2:A1538.
- Bearer, E.L. and Orci, L. 1985. Endothelial fenestral diaphragms: A quick-freeze, deep-etch study. *J. Cell Biol.* 100:418-428.
- Bearer, E.L. and Orci, L. 1986. A simple method for quick-freezing. *J. Electron Microsc. Tech.* 3:233-241.
- Bearer, E.L., Breakefield, X.O., Schuback, D., Reese, T.S., and LaVail, J.H. 2000. Retrograde axonal transport of Herpes Simplex Virus: Evidence for a single mechanism and a role for tegument. *Proc. Natl. Acad. Sci. U.S.A.* 97:8146-8150.
- Brady, S.T., Lasek, R.J., and Allen, R.D. 1985. Video microscopy of fast axonal transport in extruded axoplasm: A new model for study of molecular mechanisms. *Cell Motil.* 5:81-101.
- DeRosier, D.J. and Klug, A. 1972. Structure of the tubular variants of the head of bacteriophage T4 (polyheads). I. Arrangement of subunits in some classes of polyheads. *J. Mol. Biol.* 65:469-488.
- Fischer, M., Kaeck, S., Knutti, D., and Matus, A. 1998. Rapid actin-based plasticity in dendritic spines. *Neuron* 20:847-854.

- Harlow, M., Ress, D., Koster, A., Marshall, R.M., Schwarz, M., and McMahan, U.J. 1998. Dissection of active zones at the neuromuscular junction by EM tomography. *J. Physiol. Paris* 92:75-78.
- Hayden, J.H. and Allen, R.D. 1984. Detection of single microtubules in living cells: Particle transport can occur in both directions along the same microtubule. *J. Cell Biol.* 99:1785-1793.
- Hiraoka, Y., Agard, D.A., and Sedat, J.W. 1990. Temporal and spatial coordination of chromosomal movements. *J. Cell Biol.* 111:2815-2828.
- Inoué, S. and Spring, K.R. 1997. Video Microscopy. Plenum, New York.
- Kaech, S., Parmar, H., Roelandse, M., Barnmann, C., and Matus, A. 2001. Cytoskeletal microdiffraction: A mechanism of organizing morphological plasticity in dendrites. *Proc. Natl. Acad. Sci. U.S.A.* 98:7086-7092.
- Kam, Z., Hanser, B., Gustafsson, M.G., Agard, D.A., and Sedat, J.W. 2001. Computational adaptive optics for live three-dimensional biological imaging. *Proc. Natl. Acad. Sci. U.S.A.* 98:3790-3795.
- Kreitzer, G., Schmoranzler, J., Low, S.H., Li, X., Gan, Y., Weimbs, T., Simon, S.M., and Rodriguez-Boulan, E. 2003. Three-dimensional analysis of post-Golgi carrier exocytosis in epithelial cells. *Nat. Cell Biol.* 5:126-136.
- Lippincott-Schwartz, J. 2001. The secretory membrane system studied in real-time: Robert Feulgen Prize Lecture, 2001. *Histochem. Cell Biol.* 116: 97-107.
- Matsumoto, B. 1993. Cell Biological Applications of Confocal Microscopy. *In Methods in Cell Biology* vol. 38 (B. Matsumoto, ed.) Academic Press, San Diego.
- Meyer, M.P., Niell, C.M., and Smith, S.J. 2003. Brain imaging: How stable are synaptic connections? *Curr. Biol.* 13:R180-182.
- Miller, K.G., Field, C.M., and Alberts, B.M. 1989. Actin-binding proteins from *Drosophila* embryos: A complex network of interacting proteins detected by F-actin affinity chromatography. *J. Cell Biol.* 109:2963-2975.
- News and Editorial Staff. 2002. Breakthrough of the year: The runners-up. *Science* 298:2297-2303.
- Swerdlow, J., Goldberg, I., Brauner, E., and Sorger, P.K. 2003. Informatics and quantitative analysis in biological imaging. *Science* 300:100-102.
- Unwin, N. 2002. Structure of the acetylcholine-gated channel. *Novartis Found Symp.* 245:5-21.
- Vale, R.D., Schnapp, B.J., Reese, T.S., and Sheetz, M.P. 1985. Movement of organelles along filaments dissociated from the axoplasm of the squid giant axon. *Cell* 40:449-454.
- Westphal, M., Jungbluth, A., Heidecker, M., Muhlbauer, B., Heizer, C., Schwartz, J.M., Marriot, G., and Gerisch, G. 1997. Microfilament dynamics during cell movement and chemotaxis monitored using a GFP-actin fusion protein. *Curr. Biol.* 7:176-183.
- White, J.G., Amos, W.B., and Fordham, M. 1987. An evaluation of confocal versus conventional imaging of biological structure by fluorescence light microscopy. *J. Cell Biol.* 105:41-48.
- Zhang, H., Wessels, D., Fey, P., Daniels, K., Chisholm, R.L., and Soll, D.R. 2002. Phosphorylation of the myosin regulatory light chain plays a role in motility and polarity during *Dictyostelium* chemotaxis. *J. Cell Sci.* 115:1733-1747.

Contributed by E.L. Bearer
Brown University
Providence, Rhode Island

Three-Dimensional Reconstruction of Tissues

UNIT 14.16

**BASIC
PROTOCOL**

Three-dimensional (3-D) tissue reconstruction relies on the acquisition of three-dimensional data followed by the assembly of the data into a 3-D data block from which information can then be extracted. This unit provides a protocol for performing serial microscopy for acquiring the data, reconstructing the data into data blocks, and extracting cross-sectional areas of some features of the tissue. Because of the wide range of tissues that can be processed by serial microscopy and the reconstruction of the data into data blocks, and because each investigator will focus on different parameters of interest, the last section of the protocol is limited to high-contrast areas.

Three-dimensional imaging is advancing at a rapid rate. However, most 3-D histological techniques rely on expensive instrumentation, often with extensive custom assembly. As a result, this unit will focus on a protocol that requires standard histology and fluorescence microscopy equipment. The protocol outline starts with a brief guide to labeling and histological processing, followed by detailed steps for imaging and post-processing of the images to reconstruct 3-D data. The end result is a series of images that are merged from multiple images from each histological slide, aligned with each other for purposes of reconstruction. This data is useful because the user can display a much larger field of view than what the microscope camera is capable of acquiring, and because the user can follow features in three dimensions with alignment much more easily.

This protocol was written with the assumption that the user is competent with anesthesia, harvesting and processing of tissue, and euthanasia in the desired organism, as well as with fluorescence microscopy, including knowledge of the objectives, filters, cameras, and image output. For post-processing, experience with Matlab or adapting Matlab coding to another programming language is highly encouraged. Although the code provided for this protocol can run without the user knowing Matlab, the most effective use of the coding will require user interaction. Thus, at least minimal knowledge of Matlab is highly recommended for this protocol.

Labeling and histology. To make full use of the protocol and avoid multiple repetitions, it is essential to carefully plan the labels to be used ahead of time. Specifically, interference in wavelengths, the use of bright-field versus fluorescence labels, the expected number and distribution of the components being labeled, and the orientation of the tissue in the paraffin block should be carefully considered in advance of the experiment in order to make the study most fruitful. A few general guidelines that may be helpful include the following:

1. Fluorescence emission wavelengths tend to bleed into higher wavelengths, i.e., a blue dye will leak into green, but a green dye is unlikely to leak into blue. Thus, the smaller, dimmer, and more diffuse components should be labeled with the shorter (closer to blue) wavelengths and filtered appropriately.
2. Some light-based dyes will quench fluorescent labels. While both bright-field and fluorescent dyes can be used, and this can theoretically increase the amount of information acquired, interference and photobleaching must be considered. It is important to test combinations on sample tissues prior to experimentation on the real specimens.

**In Situ
Hybridization
and Immuno-
histochemistry**

3. The tissue should be oriented in the paraffin block (or other cutting block) in such a way that slices yield the information desired with minimal error. The errors in z (the direction perpendicular to each slice) are generally higher than in the plane of the slice (x - y), due both to cutting and to limitations of the optics.
4. With any 3-D reconstruction technique, it will be most difficult to identify the ends of features (cells, vessels) unambiguously in the direction with the least resolution. This will generally be the z direction.
5. The images being acquired should be the same size (the use of different binning for images to be stitched should be avoided), and should be arranged in a rectangular array (2 by 1, 5 by 3, and so on). One should avoid moving samples diagonally.

Materials

Labeled tissue, or animal to be labeled and appropriate anesthesia
 Labeling compounds (e.g., Invitrogen): 10 mg/ml Hoechst stain; 5 mg/136 μ l Texas red maleimide in DMEM; 500 μ g/ml Alexa Fluor 488 in PBS with Ca^{2+} and Mg^{2+} ; or other label of interest
 Fixative (e.g., 4% paraformaldehyde)
 70% ethanol
 Paraffin
 1-ml or smaller syringes
 Small-gauge needle
 Microtome
 Glass slides and coverslips
 Fluorescence microscope with appropriate objectives, filters, cameras, and acquisition software
 Computer (minimum requirements are Pentium 4 with 1 GHz or faster processor, 512 MB or higher RAM, and hard drive of 100-GB capacity or greater, depending on size and number of image files, running MacOS X or Windows 2000)
 Mathematical processing package (e.g., MatLab, Mathworks)
 Additional reagents and equipment for fixation, embedding and sectioning of tissues (e.g., *UNIT 14.1*)

NOTE: Using other programming languages, such as Visual C, requires having the appropriate compiler and knowledge of the programming syntax, which is not provided here. For the code presented in this protocol (available for download at <http://vascular.bwh.harvard.edu/cpmb3d/cpmb3d.html>), Matlab 7.0 or higher with Image Processing and Curve Fitting toolboxes is required.

Label tissue of interest

If the tissue is already labeled ex vivo and embedded in some blocking medium, proceed to step 8. Steps 1 to 7 describe intravital fluorescence labeling of an entire mouse.

1. Anesthetize the animal.
2. Using a 1-ml or smaller syringe, inject the appropriate amounts of the labeling compounds into the tail vein. For example:

50 μ l Hoechst stain for nuclear labeling (stock 10 mg/ml)
 20 μ l Texas red maleimide for matrix labeling (5 mg/136 μ l DMEM)
 50 ml Alexa Fluor 488 isolectin stain for endothelial labeling (500 μ g/ml PBS with Ca^{2+} and Mg^{2+}).

3. Keep the animal under anesthesia for 15 min to 1 hr to allow intravital staining.

4. Sacrifice the animal and harvest/fix the tissue by removing the organs and fixing in 4% paraformaldehyde overnight (UNIT 14.1).
5. Take the tissue from the 4% paraformaldehyde solution and place in 70% ethanol overnight (may be extended up to 1 week).
6. Embed the tissue in paraffin using standard histological protocols (UNIT 14.1).
7. *Optional:* In each block, puncture the paraffin around the tissue using a small-gauge needle (those used for insulin injections are usually adequate). Do not puncture the tissue itself, but place the punctures as close as possible to the tissue to avoid having to acquire extraneous data.

These holes can be used to help perform gross alignment once the images have been registered. This is especially useful for homogeneous tissue.

Section tissue

8. *Optional:* Notch the corner of the block through the entire height of the block, so that the notch will be visible in each slice.

This will help orient the tissue, and should be done if the tissue does not have gross features for orientation purposes.

9. Using a microtome (UNIT 14.1), slice 50 to 100 consecutive times and plate the samples on glass slides (coverslips are optional).

Multiple sections may be placed on the same slide as long as the order is preserved.

Each slice should be 5 to 20 μm thick, although this can be adjusted depending on the nature of the tissue and the features being examined. In general, the thickness of each slice should be selected to give $\sim 1/10$ the resolution of the features of interest, and the number of slices should be chosen to span roughly two to three times the size of the features. For example, if the features of interest are individual cardiac myocytes (expected length of 100 to 150 μm , expected diameter 10 to 30 μm), and the myocytes are being cut in cross-section, slice thicknesses of 10 to 15 μm should be used, and at least 30 consecutive slices should be acquired. If the myocyte is being cut longitudinally (along its long axis) then at least 60 slices of 1 to 2 μm should be acquired.

10. Examine two to three random slides to ensure that the labels are intact and that the notches and holes from steps 7 and 8, if any, are visible.

Image tissue

11. Using a fluorescence microscope, place one slide on the stage and acquire images from the first section in the region of interest. If holes were made in step 7, also acquire adjacent images that include at least two holes in the paraffin.

It is vital that all images have some overlap with each other. This can be done manually using visual landmarks, or by using a properly calibrated automated stage. The images should also form one large rectangular array to be assembled later. For ease of post-processing, the images should be acquired top-to-bottom (the y direction in plane, not the z direction through the section) and then left-to-right. After completion of imaging from one section, the acquired images should fit together to form one large image without separations. Small gaps in the images where there is no useful information are acceptable, but stand-alone islands are unacceptable. For each section, it is recommended that the same number and arrangement of images be acquired.

12. Acquire images from the other sections in sequence, using the same template of images as in the first section from step 11. That is, acquire images from all sections in a similar fashion, so that they encompass the area of interest and the holes, and so that they form the same-sized rectangular array.

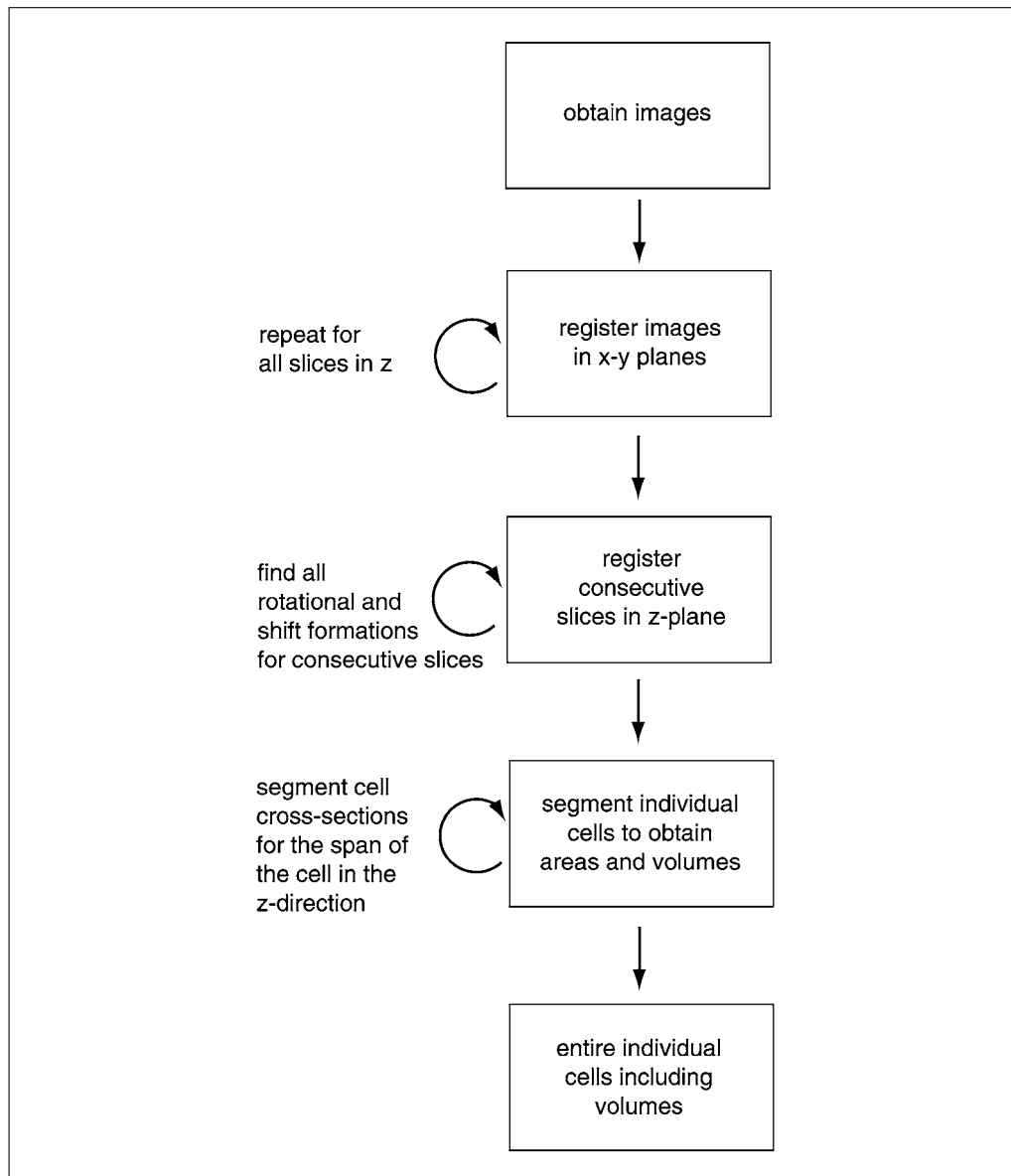


Figure 14.16.1 Flowchart of the post-processing for extraction of 3-D data from serial slices.

It is useful to have the file names follow a standard convention so that post-processing will be easier. For example, a file from a heart sample can be labeled heart_3_s6.tif to indicate that it is the third image of section 6.

For processing, the image files must be the same size and dimensions. Thus, binning or cropping should be avoided. If they are necessary, image processing should be performed using software to ensure the images are identical in size, resolution, and format before proceeding to the next step.

The flowchart in Figure 14.16.1 depicts the outline of the rest of the protocol. Images will be registered in-plane, then assembled in the z direction, and then features will be isolated.

Perform two-dimensional post-processing

13. Copy the images to the computer for image processing with the mathematical image analysis software. Make sure the images are in the Matlab root directory (for Matlab Version 7 Release 14 on a PC, this would default to C:\MATLAB701\work).

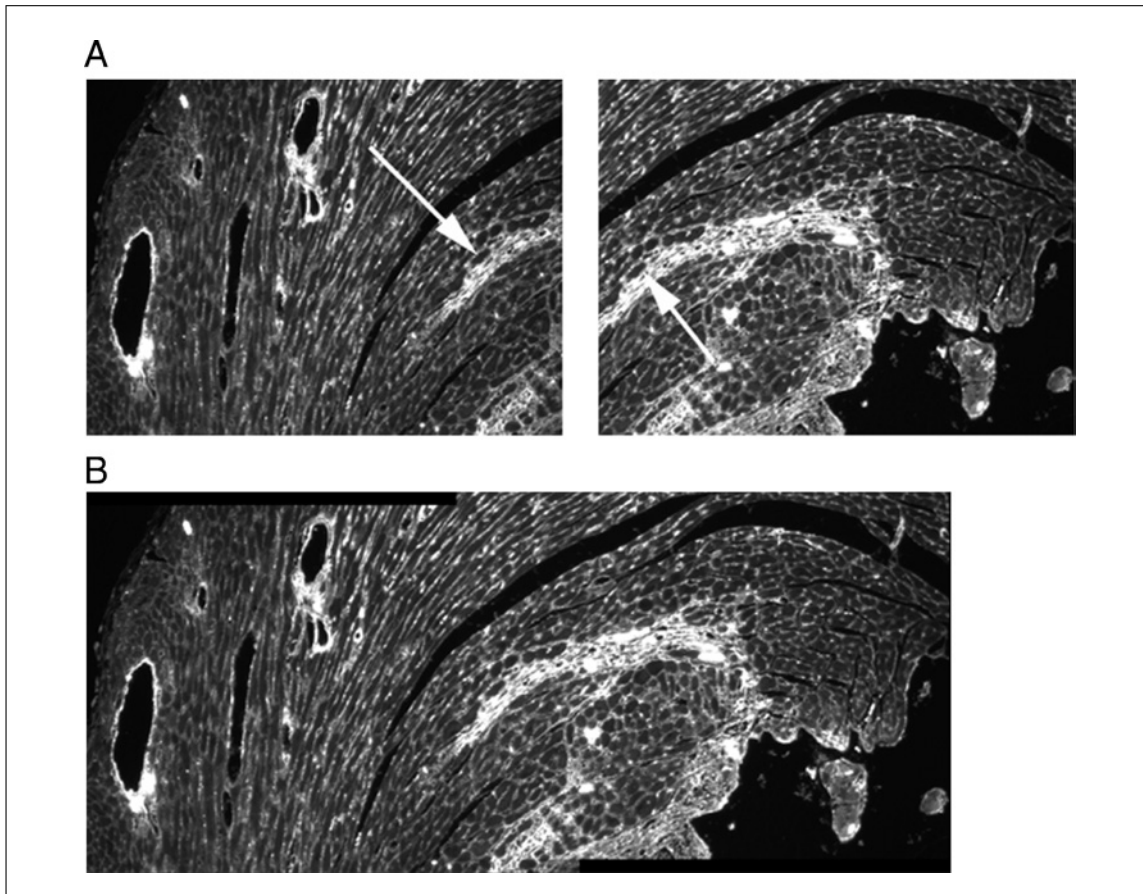


Figure 14.16.2 Two-dimensional registration of images. (A) Two images to be registered. Notice that the images overlap horizontally but are shifted slightly vertically. (B) Registered image.

14. Load images from a single plane (multiple images from a single slice in the rectangular array) into Matlab for viewing purposes. To do this, type

```
>> imshow filename.tif
```

at the Matlab command prompt, where `filename.tif` is the image file to be displayed. Examine the images to ensure the overlap is visible by looking for common features on the edges of the images. For example, in the images shown in Figure 14.16.2A, the white arrows point to a section of strong labeling that is common to both images.

After beginning this step, do not type 'clear' at the Matlab command prompt because some of the values are needed to execute subsequent commands. Any clearing that has to be done will be in the corresponding code.

15. Register the images in a single plane. Perform registration by cross-correlation using the Matlab codes named `plane_register.m`, with functions `TBreg.m` and `RLreg.m`, if downloaded (for codes, go to <http://vascular.bwh.harvard.edu/cpmb3d/cpmb3d.html>). If manually entering the code, type the contents of the code into separate notepad files and save as `plane_register.m`, `TBreg.m`, and `RLreg.m` in the Matlab directory. If the files have `.txt` appended to them, highlight the files one at a time and press F2 to rename the files without the `.txt` extension. Launch Matlab and type

```
>> plane_register
```

at the Matlab command prompt to run the script.

- a. A dialog box will open asking for the Overlay Pattern. Depending on how the images are designed to fit together, the input should consist of two numbers separated by a space. The first number is the number of rows (vertical number of images), and the second number is the number of columns (horizontal number of images). For the images in Figure 14.16.2A, one would enter 1 2 and then click OK.
- b. Enter the percent overlap of the images. For the images in Figure 14.16.2A, a good initial overlap would be 25%, although this can be adjusted if the program does not obtain a satisfactory result. For automated stages, once one overlap is figured, the rest should be nearly identical. Note that the percent should be entered as a true percentage (and not a fraction). Thus for Figure 14.16.2A, one would enter 25 and then click on OK.
- c. An open file dialog will now open for each image, up to the number of images based on the input in step a. Thus, if the input into step a is 2 3, a total of 6 files will be requested. These must be opened in the order of the image acquisition top to bottom, then left to right. For example:

```
1 3 5
2 4 6
```

would be the order of the images for a 2 by 3 array of images.

- d. The program will then run and register the images into one big image. At the end, a large image will appear of the registered image with some black space used for buffer space. It is recommended that the black space not be removed if z-registration is to be performed immediately.
16. Save the displayed registered images as a single large image. This can be done using the Matlab command `imwrite.m`. At the Matlab command prompt, type:

```
>> imwrite(final_image, 'desired_name.tif', 'tif')
```

where `'desired_name.tif'` can be custom (`'heart001.tif'` for example).

This will save the registered image as a TIFF file. Do not save by using the menu option File/Save As. Make sure the file type is of a suitable format (TIFF is recommended). It is recommended that each registered slice be assigned a slice number for z-registration (heart001.tif being the first slice, heart002.tif being the second, and so on) for easier recall later. Figure 14.16.2B depicts the images from panel A, now in register.

17. Repeat steps 15 and 16 for all the slices (e.g., 50 to 100) that were imaged.

Perform three-dimensional reconstruction

18. Close all images in the workspace to clear memory. Close the figure window or type

```
>> close all
```

at the Matlab command prompt.

19. Open two consecutive (big) images of the slice (saved in step 17) by typing at the Matlab command prompt:

```
>> cpselect('image1name.tif', 'image2name.tif')
```

where `image1name.tif` and `image2name.tif` are the filenames of the images.

A new window will open with four windows (ignore the legend). The bottom two windows display the images in total, while the upper portion displays magnified images of the areas in the white-bordered rectangle in the bottom images. Using the scroll bars in the upper

images, it is possible to move around the whole image. The rectangle in the lower image will move to show the region being magnified in the upper image.

20. Identify several (at least three) morphological features that appear in the same location in both images (e.g., tears, dark spots, a cell with particularly odd morphology, or, failing that, the holes punctured in the paraffin from step 7). Left-click on one feature in the upper left image, followed by the corresponding feature in the upper right image (lower images are for reference). Repeat for at least three pairs of points.
21. Under the menu option File, select Save control points to workspace. Use default checkboxes and click OK.
22. *Optional:* Save the coordinates to a file for future reference by typing at the Matlab command prompt:

```
>> csvwrite('filename.csv', [input_points base_points])
```

where `filename.csv` is a comma-separated value file with a filename of one's choice.

The output is four columns of x and y coordinates of the features in the left-hand image followed by the coordinates of the same features in the right-hand image. The number of rows should be the same as the number of features selected in step 20.

23. Perform least-squares fitting to determine the best orientation and translation of the second image with respect to the first. Use the code `rotate_second_image.m` from download or by typing/saving as a new M-file. Execute the code by typing at the Matlab command prompt:

```
>> rotate_second_image
```

When the open file dialog box appears, choose the name of the second image in the stack to be rotated and shifted in reference to the first, which was done in steps 19 and 20. Once the program has completed its run, save the image by entering the following at the Matlab command prompt:

```
>> imwrite(transformed, 'desired_name.tif', 'tif')
```

where `desired_name.tif` is the filename of the image to be saved.

24. Repeat steps 17 to 23 for the second and third images, then for the third and fourth images, and so on.
25. Once completed, display a data block of the corrected scan volume as individual images or as a movie if the images are stitched together.

Isolate features

This section of the protocol calculates the cross-sectional area of a feature that is dark and bordered by a bright zone. If the image in question is reversed, use an image processing program to invert the image.

26. *Optional:* Open the two consecutive (big and corrected) images aligned from the three-dimensional reconstruction (i.e., ending at step 25). Open the images so that they are the same size and take up the same position in the window when they are active.

This may help with identification of features or landmarks in successive sections. It is best to open images from the center of the stack (images 25 and 26 in a stack of 50 images or 50 and 51 in a stack of 100 images).

27. Run the script `feature_extract.m`, with the two functions `image_mask.m` and `bwcfind.m` (again, type each script/function into separate files and save, or download). Open the image at the top of the z-stack when prompted by the dialog box. Set threshold at the desired value (25 is default; lower means dimmer signal gets included, higher means brighter signals get excluded) and click OK. Drag the mouse to select a region containing the cells or features of interest (click once and hold the mouse button to start dragging, once the area has been selected, release the button). Two images will display, one with the cropped original image and one entitled 'water,' which has a watershed and edge-detected segmentation of the cropped image. Save the water image if desired, otherwise continue to step 28 with the cropped original image and ignore the water image.
28. Identify a cell or feature of interest for which the cross-sectional area is desired by clicking in the middle of the cell/feature. Repeat for as many cells/features as desired, then click the middle mouse button/scroll wheel (if using a PC) to end. If the mouse does not have a center button, hit Enter, and type at the Matlab command prompt:


```
>> Areas
```

The cross-sectional areas (in pixels) will be displayed in the order the cells/features were selected.
29. *Optional:* Save the images for future reference. Use the `imwrite` command as previously described (File/Save As is possible but not recommended). Make sure the desired file format is selected (TIFF is recommended).
30. Repeat steps 27 to 29 for subsequent images and identify the same cell or feature. This can be accomplished using several landmarks:
 - a. The cell is in the same location and overlaps significantly with the first.
 - b. The cell is fluorescently labeled and the ones near it are unlabeled or labeled differently.
 - c. The morphology of the cell is distinctive and the same distinctiveness shows up in the second image.

In general, method b is best for clear identification, but is most difficult to attain.
31. To identify the ends of the cell, use one or more of the following landmarks:
 - a. The cell is fluorescently labeled and the ones adjacent to it are not. When the fluorescence ends, so does the cell.
 - b. The location of the cell is no longer taken up by a cell but by non-cell features. Alternatively, if a cell overlaps with previous cells by a large percentage, then suddenly no cells overlap by the same amount, that is an indication that the cell has ended.
 - c. If a nuclear stain was used, and the number of nuclei in the cells is known, locations between the nuclei of a single cell where the cell visibly shrinks are likely to represent a partition.
 - d. Use the image stack in 3-D, and extract a perpendicular section. Use the perpendicular section to identify the ends of the cells.

In general, method a is best, although method d permits similar analysis without requiring additional staining.
32. Once the complete cell is identified, calculate the volume by adding the cross-sectional areas of each slice containing the cell and multiplying by the thickness of the slices. Convert from pixels to microns according to the microscope/camera setup,

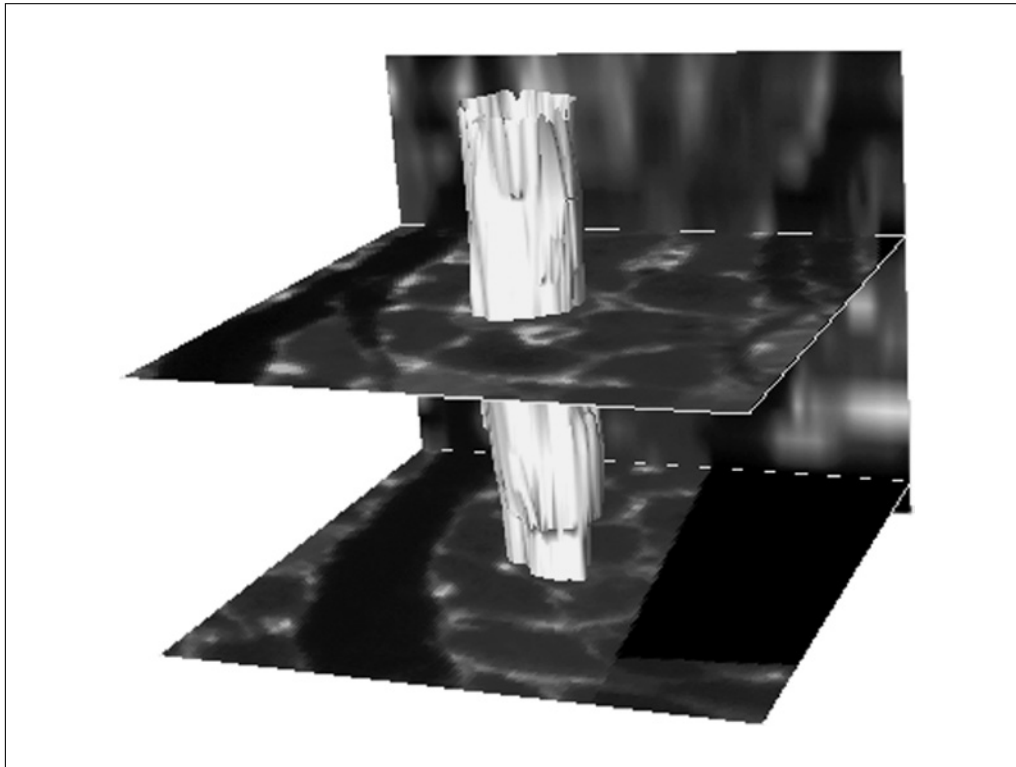


Figure 14.16.3 Example of extracted cell with various planes showing part of a cardiac myocyte. This processing was performed using Amira.

using a Ronchi ruling or similar, if desired. Calculate the surface area by adding the perimeter of each cross-sectional area multiplied by distance between slices.

The image stack can also be imported for use in other 3-D packages, such as Amira (Fig. 14.16.3).

COMMENTARY

Background Information

Three-dimensional (3-D) image processing of tissue data is a powerful tool that can provide clear visualization of the tissue properties and enhance quantization of relevant morphology. This unit refers to any reconstruction and analysis requiring all three dimensions, regardless of the thickness of the specimen. For example, a globe of the Earth is three-dimensional, despite the fact that the information on the globe has no theoretical thickness, because simple attempts to flatten the map will induce distortions or introduce discontinuities. Thus, all three dimensions are necessary.

From a visualization perspective, being able to see and rotate a solid model in space has advantages over two-dimensional (2-D) representations. For example, consider explaining the structure of coronary arteries using a 3-D surface map versus one where a movie of serial slices is shown and the audience is expected to mentally construct the map. The 3-D

map is much less likely to create confusion. Additionally, once a 3-D data block has been acquired, the tissue can be sectioned in any desired plane. Having 3-D data also obviates problems that might arise when determining the length, breadth, and width of the same cell using classic 2-D techniques.

Quantitatively, there are numerous advantages to using 3-D analysis compared to 2-D. Determination of volumes, surface areas, and spatial orientation is facilitated with 3-D data. Statistical advantages exist as well. Due to heterogeneity in some tissues, 2-D data reflect not only the inherent variability of the cells in the tissue but also the variability resulting from cutting. This weakens statistical analysis, resulting in changes of significance or increasing the number of specimens required to make a significant finding. To clarify this point, consider sampling a number of cylinders with varying cross-sectional areas. If one has the 3-D image of each cylinder, the cross-sectional

area determined by a caliper-type program would reflect the variability in the cylinders themselves (along with measurement errors). However, if one were to cut the cylinders at various angles, the variances in cross-sectional areas will generally be increased, artificially inflating the values and the variance itself.

Despite the advantages of 3-D analysis, few studies have focused on developing techniques and even fewer studies have involved 3-D analysis on tissues. Some reasons for this include the expense and difficulty in use of the necessary instrumentation, difficulty of post-processing, large amounts of data storage and processing required, and increased time required for analysis.

With the advent of commercial confocal and two-photon microscopes and improved computation, more studies are relying upon three-dimensional analysis. Confocal microscopy uses standard fluorescence illumination (laser or lamp based) and then filters out-of-plane signals using a pinhole (see, for example, Periasamy, 2001). Depending on the illumination source, the scanner/pinhole setup, and the objective, the expected z signal can be narrowed down to an $\sim 0.5\text{-}\mu\text{m}$ -thick plane. Studies using confocal microscopy show improved image quality in relatively thin planes as well as in deep scanning. Disadvantages of confocal microscopy include the cost of a laser-scanning system (typically \$200,000 USD and above), limited penetration depths (especially for shorter wavelengths), and photobleaching when taking volumetric data over an extended time. Despite these disadvantages, the signal clarity and rejection of out-of-plane noise provide sufficient advantages that several studies in three-dimensional cell and tissue structure rely on confocal microscopy and post-imaging reconstruction (Roeder et al., 2004; Sun et al., 2004; Sugawara et al., 2005; Tobita et al., 2005).

Two-photon microscopy typically uses a red-to-infrared ($>700\text{-nm}$) laser to generate an excitation volume confined to the area of maximal focus after the objective (Denk et al., 1990, 1991; Williams et al., 1994). Because the excitation relies on two photons exciting the same molecule nearly simultaneously, only the focal region has sufficient photon density to generate a fluorescent signal. The practical result is that the signal received by the camera or photomultiplier tube is emanated from a small volume, with a z -plane thickness of $\sim 1\ \mu\text{m}$ (depending on the objective used). There is subsequently no need for a pinhole/scanner arrangement, although this can be added to fur-

ther clarify the signal. Two-photon microscopy has two main advantages: (1) the limitation of excitation to a small volume, which reduces overall photobleaching and out-of-plane noise, and (2) the use of long wavelengths that have greater penetration depth into tissues, allowing scanning of optically favorable tissues of up to several hundred microns. Disadvantages of two-photon microscopy include heat damage to the specimen, expense of laser and microscope maintenance, and, in some cases, loss of resolution and difficulty in locating the feature of interest due to the scanning nature of the majority of two-photon microscopes. Two-photon microscopy is well-suited for deep scanning of tissues and for performing three-dimensional cell and tissue scanning, in part because of its penetration depth.

When using standard epifluorescence (sometimes called wide-field fluorescence), the out-of-plane noise is generally unavoidable. However, software programs can perform deconvolution on acquired data, assuming that information regarding the optical setup is known. There are several types of deconvolution, and the technique most suitable for a given application depends on the equipment, the tissue sample, and the desired outcome. In general, deconvolution can greatly improve epifluorescence images but not to the extent of confocal image acquisition. However, the combined cost of a standard fluorescence microscope and deconvolution software is much lower than the cost of a confocal system. Deconvolution has been used to clarify three-dimensional images of cells mainly, but also of some tissues (see, for example, Maierhofer et al., 2003; Mudry et al., 2003; Segretain et al., 2003; Carreno et al., 2004; Bick et al., 2005). The technique described in this unit does not rely on deconvolution. The out-of-plane noise is eliminated simply by imaging only the thin section that is desired, and then the 3-D reconstruction is performed by using cross-correlation algorithms to find a match to splice the images together; finally, a least-squares fit is applied to the spliced images to perform the alignment in z . While some commercial microscope and imaging processing packages can perform the in-plane registration automatically when coupled with an automated stage, the protocol described in this unit does not require an automated stage and can do alignment in the vertical direction based on serial slices.

Three-dimensional tissue scanning has been applied to an increasing number of tissue types and sizes. Certain new techniques

such as optical projection tomography permit the acquisition of larger and increasingly complex specimens (Sharpe et al., 2002; Sharpe, 2003). The addition of a lateral microtome to remove scanned sections on confocal or two-photon microscopes allows theoretically unlimited scanning (Young et al., 1998; Ewald et al., 2002; Ragan et al., 2003; LeGrice et al., 2004). Laser ablation (the burning away of scanned tissues using a second laser) has been used in a similar manner (Tsai et al., 2003). Certain other techniques—such as MRI, ultrasound, and, more recently, optical coherence tomography—allow fairly deep (millimeters to tens of centimeters), nondestructive scanning but are limited in resolution and feature extraction. Current challenges include increasing the speed of image acquisition, performing high-resolution non-invasive scans, labeling of specific tissue and cell components (at the organ level or above), and storage and post-processing of data.

Critical Parameters

In performing 3-D analysis using the method described here, it is vital that potential errors associated with the method be understood.

The sections to be imaged must be cut with a microtome with care so that distortion is minimized and all or almost all slices can be saved. This is not a major issue if the reconstruction is meant to be a rough estimate or if the slices are extremely thin (1 to 3 μm), but for most purposes, dropping slices leads to missing volumes that might contain the end of a cell or other feature that is important.

The programs offered in this protocol allow the user to register (in x - y) and assemble (in z) stacks of slices in order to obtain a three-dimensional tissue stack. However, it cannot be overemphasized that the user must check each step to make sure the alignments are realistic. Even with guides and cross-correlations, it is possible for the program to settle on a fit that is rotated, shifted entirely out of view, or matched to incorrect features. The old adage of GIGO (garbage in, garbage out) is especially applicable in three-dimensions, where one needs to look at successive images to see if the construct makes sense.

When locating features of interest, it is useful to check a few slices in z ahead of where the current slice is located to see whether the feature will be useful. Good features (1) persist for several slices instead of just one or two, (2) do not apparently change in shape or position much, if at all, (3) are unique in some

way, and are therefore unambiguously identifiable from slice to slice, and (4) are small or have small features that are uniform. Small size permits the coordinates of the features to have the smallest error from slice to slice.

Finally, for statistical analyses or analyses involving differential determinations, it is important that the person performing the registration and extraction be blind to the experimental groups for the specimens being analyzed. The user of the programs can unintentionally distort the size and shape of the features. Use of more advanced technologies, such as the lateral microtome or laser ablation techniques, might alleviate user bias.

Troubleshooting

Poor registration can be a result of several issues. If the image does not have sufficient resolution for feature alignment, then the imaging setup might need to be examined. In particular, the illumination, exposure time, objective, focus, and amount of overlap may need to be adjusted for each specimen.

Incorrect image registration usually results from input error. First, make sure the images are the correct ones being registered. It may be helpful when acquiring the images to tag filenames consistently with a code so that the slice and position can be determined. Overlap should be visible on opening the images. The user should attempt to input as accurate an overlap percentage as possible. Input of a 60% overlap when the images only overlap 20% will increase the chances of registering incorrectly due to mathematical noise; in contrast, attempting a 10% overlap when the images have a 20% overlap will almost never yield correct results. The overlap percentage should be slightly greater than the actual overlap, but not by too much. During imaging, overlaps of 20% to 40% are optimal for most cases.

Registering images may present the following error: Matrix index must be positive or logical. This occurs when Matlab is creating the final image and either an image below the first input image has an offset >200 pixels to the left, or an image to the right has an offset >200 pixels above of the first image. If this occurs, two options are accessible without editing the programs. The first one is to reexamine and retake the images of the section with lower (diagonal) shifts. The second option is to crop the images to a smaller size using Matlab or another image processing package, making certain that all the images in the layer are of the same dimensions.

Tissue distortion may result in a poor 3-D stack. If this is the case, the sample preparation needs to be adjusted. For example, the thickness of the cut can be adjusted (keeping in mind not to increase the thickness far beyond 1/10 the useful resolution), the tissue fixation can be adjusted (i.e., by increasing or decreasing the length of fixation), and/or the embedding medium can be adjusted from the standard paraffin (e.g., with frozen sections in specialty media, or epoxy resins).

If the ends of cells are difficult to identify (this is a common problem), then selective labeling should be used to highlight the cells of interest. Some examples include transfecting or infecting cells of the organ with a fluorescent reporter, using dyes such as Cell-Tracker (Invitrogen), labeling the nuclei (using Hoechst, for example, which will help identify the number of expected cells if the cells are known to have exactly one nucleus), sectioning in a perpendicular direction, or using thinner slices, which may help expose the ends of the cells.

If there is a lack of features for z alignment, for example, because the tissue is extremely homogeneous under a microscope, then one can introduce artificial features by puncturing the paraffin surrounding the tissue, puncturing the tissue itself, or cutting the tissue block into a rectangular piece so that the overall geometry is easily identified. Puncturing the tissue is not recommended, since this will likely cause tears in the tissue during sectioning and may introduce distortions into the tissue directly.

During segmentation, too little or too much segmentation may occur (i.e., the area selected may be much larger or much smaller than the region that should be segmented). In this case, change the threshold value parameter depending on the contrast of the image to obtain appropriate segmentation. The default value is 25; a lower value will generally yield larger areas of segmentation while a higher value will yield a smaller area. If possible, increase the contrast of the image prior to segmentation by using an image processing package or the Matlab command:

```
>> imadjust('filename.tif')
```

Occasionally, Matlab will not write an image using `imwrite` after many registrations have been performed and the images written to files using `imwrite`. This generally occurs because the computer memory is full. If this happens, type at the Matlab command prompt:

```
>> pack; dumpmemex;
```

or alternatively, exit and restart Matlab. The last program may need to be rerun.

Anticipated Results

By following this method, the user should obtain 3-D data of individual cells that can be quantitatively analyzed. Extraction of cell data and information includes cell volume, length, and surface visual representation. Location of inner features such as nuclei can be made. Depending on the quality of the slices, the entire image section may have usable data; however, due to shifts in overlap and expected distortions, not all of the cells will be usable.

As an illustration, for a section that is $\sim 1400 \times 1000$ pixels from a single image, combining six images together with 40% overlap will yield one large image that is 3000×1600 pixels. At $20\times$ magnification, with $1 \mu\text{m}$ per pixel, the field is therefore $3 \times 1.6 \text{ mm}$. Given that a cardiac myocyte is $\sim 500 \mu\text{m}^2$ in cross-section, one might expect there to be nearly 10,000 usable cells. However, not all of the area will have cells, not all of the area will have correct overlap (and portions outside the overlap region cannot be used), not all of the cells will be complete in the z direction, and not all of the cells will have identifiable ends.

Time Considerations

Registration in x - y should take 6 hr for a stack of 50 slices, each slice having six images with image dimensions of 1392×1040 with 40% overlap (using a 2.2 GHz Pentium 4, with 1 GB of RAM, running Matlab 7 Release 14 under Windows XP). Registration in z should take 1 to 2 hr for a stack of 50 images (same configuration as above). Identifying a single cell should take 6 to 10 min with practice.

A single data set (50 slices, all aligned and pre-processing files deleted) should require 4 to 10 GB of HDD space, post-processing. While this information is not strictly a time consideration, it is presented here because copying, backing up, and/or moving the files can take substantial time even at high transfer rates. For example, data transfer of a 10 GB set of files at 100 KB/sec (Ethernet speed) would take over 24 hr.

Literature Cited

Bick, R.J., Bagwell, S.H., Jones, C.E., Poindexter, B.J., Buja, L.M., Youker, K.A., Grigore, A., Clubb, F., Radovancevic, B., and Frazier, O.H. 2005. Fluorescence imaging microscopy of cellular markers in ischemic vs non-ischemic cardiomyopathy after left ventricular unloading. *J. Heart Lung Transplant.* 24:454-461.

- Carreno, S., Engqvist-Goldstein, A.E., Zhang, C.X., McDonald, K.L., and Drubin, D.G. 2004. Actin dynamics coupled to clathrin-coated vesicle formation at the trans-Golgi network. *J. Cell Biol.* 165:781-788.
- Denk, W., Strickler, J.H., and Webb, W.W. 1990. Two-photon laser scanning fluorescence microscopy. *Science* 248:73-76.
- Denk, W.J., Strickler, J.P., and Webb, W.W. 1991. Two-photon laser microscopy. United States Patent 5,034,613.
- Ewald, A.J., McBride, H., Reddington, M., Fraser, S.E., and Kerschmann, R. 2002. Surface imaging microscopy, an automated method for visualizing whole embryo samples in three dimensions at high resolution. *Dev. Dyn.* 225:369-375.
- LeGrice, I., Sands, G., Hooks, D., Gerneke, D., and Smaill, B. 2004. Microscopic imaging of extended tissue volumes. *Clin. Exp. Pharmacol. Physiol.* 31:902-905.
- Maierhofer, C., Gangnus, R., Diebold, J., and Speicher, M.R. 2003. Multicolor deconvolution microscopy of thick biological specimens. *Am. J. Pathol.* 162:373-379.
- Mudry, R.E., Perry, C.N., Richards, M., Fowler, V.M., and Gregorio, C.C. 2003. The interaction of tropomodulin with tropomyosin stabilizes thin filaments in cardiac myocytes. *J. Cell Biol.* 162:1057-1068.
- Periasamy, A. 2001. *Methods in Cellular Imaging*. Oxford University Press, Oxford.
- Ragan, T., Huang, H., and So, P.T. 2003. In vivo and ex vivo tissue applications of two-photon microscopy. *Methods Enzymol.* 361:481-504.
- Roeder, B.A., Kokini, K., Robinson, J.P., and Voytik-Harbin, S.L. 2004. Local, three-dimensional strain measurements within largely deformed extracellular matrix constructs. *J. Biomech. Eng.* 126:699-708.
- Segretain, D., Decrouy, X., Dompierre, J., Escalier, D., Rahman, N., Fiorini, C., Mograbi, B., Siffroi, J.P., Huhtaniemi, I., Fenichel, P., and Pointis, G. 2003. Sequestration of connexin43 in the early endosomes: An early event of Leydig cell tumor progression. *Mol. Carcinog.* 38:179-187.
- Sharpe, J. 2003. Optical projection tomography as a new tool for studying embryo anatomy. *J. Anat.* 202:175-181.
- Sharpe, J., Ahlgren, U., Perry, P., Hill, B., Ross, A., Hecksher-Sorensen, J., Baldock, R., and Davidson, D. 2002. Optical projection tomography as a tool for 3D microscopy and gene expression studies. *Science* 296:541-545.
- Sugawara, Y., Kamioka, H., Honjo, T., Tezuka, K., and Takano-Yamamoto, T. 2005. Three-dimensional reconstruction of chick calvarial osteocytes and their cell processes using confocal microscopy. *Bone* 36:877-883.
- Sun, Y., Rajwa, B., and Robinson, J.P. 2004. Adaptive image-processing technique and effective visualization of confocal microscopy images. *Microsc. Res. Tech.* 64:156-163.
- Tobita, K., Garrison, J.B., Liu, L.J., Tinney, J.P., and Keller, B.B. 2005. Three-dimensional myofiber architecture of the embryonic left ventricle during normal development and altered mechanical loads. *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* 283:193-201.
- Tsai, P.S., Friedman, B., Ifarraguerri, A.I., Thompson, B.D., Lev-Ram, V., Schaffer, C.B., Xiong, Q., Tsien, R.Y., Squier, J.A., and Kleinfeld, D. 2003. All-optical histology using ultrashort laser pulses. *Neuron* 39:27-41.
- Williams, R.M., Piston, D.W., and Webb, W.W. 1994. Two-photon molecular excitation provides intrinsic 3-dimensional resolution for laser-based microscopy and microphotochemistry. *FASEB J.* 8:804-813.
- Young, A.A., Legrice, I.J., Young, M.A., and Smaill, B.H. 1998. Extended confocal microscopy of myocardial laminae and collagen network. *J. Microsc.* 192:139-150.

Contributed by Francisco Cruz, Richard T. Lee, and Hayden Huang
Brigham and Women's Hospital
Boston, Massachusetts

CHAPTER 15

The Polymerase Chain Reaction

INTRODUCTION

The polymerase chain reaction (PCR) is a rapid procedure for *in vitro* enzymatic amplification of a specific segment of DNA. Like molecular cloning, PCR has spawned a multitude of experiments that were previously impossible. The number of applications of PCR seems infinite—and is still growing. They include direct cloning from genomic DNA or cDNA, *in vitro* mutagenesis and engineering of DNA, genetic fingerprinting of forensic samples, assays for the presence of infectious agents, prenatal diagnosis of genetic diseases, analysis of allelic sequence variations, analysis of RNA transcript structure, genomic footprinting, and direct nucleotide sequencing of genomic DNA and cDNA.

The theoretical basis of PCR is outlined in Figure 15.0.1. There are three nucleic acid segments: the segment of double-stranded DNA to be amplified and two single-stranded oligonucleotide primers flanking it. Additionally, there is a protein component (a DNA polymerase), appropriate deoxyribonucleoside triphosphates (dNTPs), a buffer, and salts.

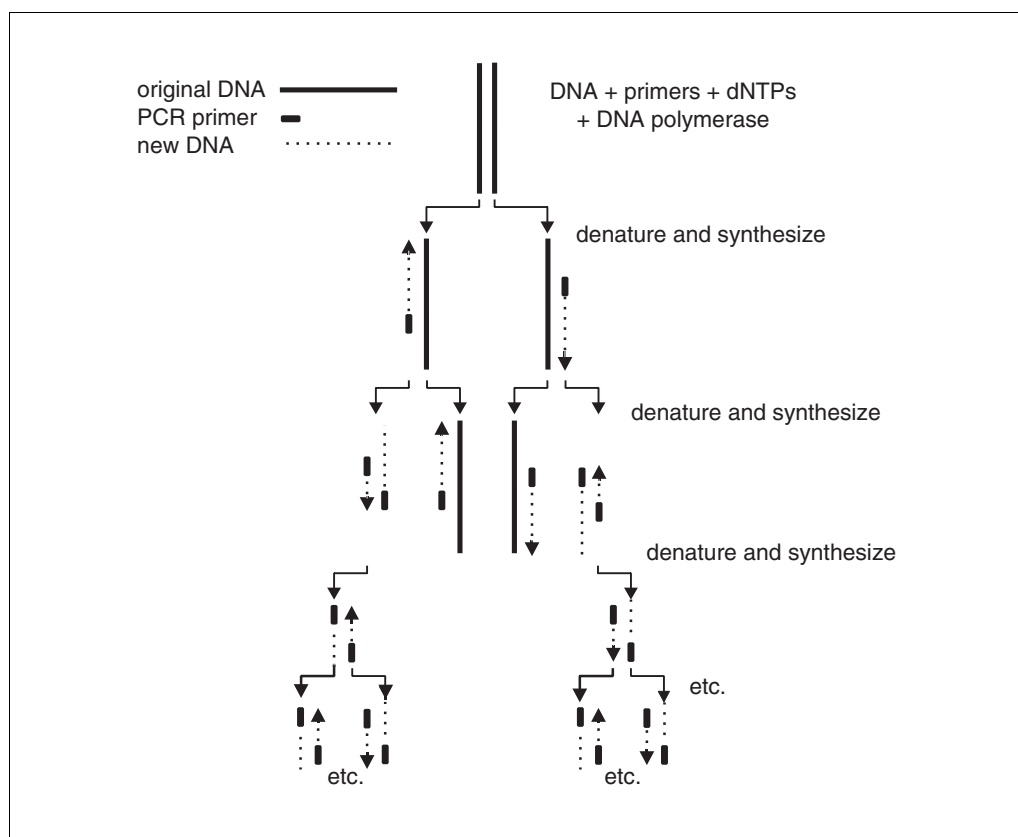


Figure 15.0.1 The polymerase chain reaction. DNA to be amplified is denatured by heating the sample. In the presence of DNA polymerase and excess dNTPs, oligonucleotides that hybridize specifically to the target sequence can prime new DNA synthesis. The first cycle is characterized by a product of indeterminate length; however, the second cycle produces the discrete “short product” which accumulates exponentially with each successive round of amplification. This can lead to the many million-fold amplification of the discrete fragment over the course of 20 to 30 cycles.

Contributed by Donald M. Coen

Current Protocols in Molecular Biology (2001) 15.0.1-15.0.3

Copyright © 2001 by John Wiley & Sons, Inc.

The primers are added in vast excess compared to the DNA to be amplified. They hybridize to opposite strands of the DNA and are oriented with their 3' ends facing each other so that synthesis by DNA polymerase (which catalyzes growth of new strands 5'→3') extends across the segment of DNA between them. One round of synthesis results in new strands of indeterminate length which, like the parental strands, can hybridize to the primers upon denaturation and annealing. These products accumulate only arithmetically with each subsequent cycle of denaturation, annealing to primers, and synthesis.

However, the second cycle of denaturation, annealing, and synthesis produces two single-stranded products that together compose a discrete double-stranded product which is exactly the length between the primer ends. Each strand of this discrete product is complementary to one of the two primers and can therefore participate as a template in subsequent cycles. The amount of this product doubles with every subsequent cycle of synthesis, denaturation, and annealing, accumulating exponentially so that 30 cycles should result in a 2^{28} -fold (270 million-fold) amplification of the discrete product.

This chapter consists of protocols that cover some of the more common applications of PCR. For many applications, the first step is simply to get PCR working with a known segment of DNA and a set of primers. Therefore, *UNIT 15.1* presents a basic PCR protocol and ways to optimize it for the sequence of interest.

PCR permits direct sequencing of nucleic acids without requiring cloning, thus avoiding cloning difficulties and artifacts. Several different protocols for preparing PCR products for sequencing using either dideoxy (Sanger) sequencing methods or chemical (Maxam-Gilbert) methods are presented in *UNIT 15.2*. This unit should permit the practitioner to choose a protocol best suited to the problem at hand and to his or her taste.

Several PCR methods have been developed that require knowledge of only a small stretch of sequence (30-40 bases) and add sequence to the ends of amplified molecules to facilitate analyses. One of these, ligation-mediated PCR (*UNIT 15.3*) has broad applications including genomic footprinting and sequencing.

PCR can be used to help clone and manipulate sequences. Various methods for generating suitable ends to facilitate the direct cloning of PCR products are detailed in *UNIT 15.4*. Other protocols for cloning and mutagenesis of DNA using PCR can be found in *UNIT 3.7* and *UNIT 8.5*.

An important application of PCR is to detect RNA transcripts, analyze their structure, and amplify their sequences to permit cloning and/or sequencing. *UNIT 15.5* presents procedures that adapt PCR to RNA templates, via production of a cDNA copy of the RNA by reverse transcriptase (RT-PCR). Anchored PCR, which, like ligation-mediated PCR, requires little knowledge of sequence and makes use of the ends of nucleic acids, is applied in *UNIT 15.6* to analysis of mRNAs.

PCR is frequently used because it is the most sensitive assay for rare sequences. A protocol that not only detects rare DNAs but quantitates them as well is presented in *UNIT 15.7*. The downside of sensitivity is contamination by infinitesimal amounts of unwanted exogenous sequences. Procedures designed to avoid contamination with undesired DNA sequences are emphasized in this unit.

In Supplement 56, we have rearranged the order of the units and have updated some of them. Further updates and new units, especially on quantitative approaches to PCR, are forthcoming. We have also moved one unit, formerly 15.8, which describes a powerful application of PCR called differential display to Chapter 25, "Discovery and analysis of differentially expressed genes in single cells and cell populations", where it is now *UNIT*

25B.3. Chapter 25 includes other PCR-based protocols as well. Of course, as PCR has become an invaluable tool for nearly every branch of molecular biology, applications of PCR can be found in many other chapters of *Current Protocols in Molecular Biology* including Chapters 3, 7, 8, 12-14, 16, 21, 22, and 24.

Donald M. Coen
Contributing Editor (Chapter 15)
Harvard Medical School

Enzymatic Amplification of DNA by PCR: Standard Procedures and Optimization

UNIT 15.1

**BASIC
PROTOCOL**

This unit describes a method for amplifying DNA enzymatically by the polymerase chain reaction (PCR), including procedures to quickly determine conditions for successful amplification of the sequence and primer sets of interest, and to optimize for specificity, sensitivity, and yield. The first step of PCR simply entails mixing template DNA, two appropriate oligonucleotide primers, *Taq* or other thermostable DNA polymerases, deoxyribonucleoside triphosphates (dNTPs), and a buffer. Once assembled, the mixture is cycled many times (usually 30) through temperatures that permit denaturation, annealing, and synthesis to exponentially amplify a product of specific size and sequence. The PCR products are then displayed on an appropriate gel and examined for yield and specificity.

Many important variables can influence the outcome of PCR. Careful titration of the MgCl_2 concentration is critical. Additives that promote polymerase stability and processivity or increase hybridization stringency, and strategies that reduce nonspecific primer-template interactions, especially prior to the critical first cycle, generally improve amplification efficiency. This protocol, using *Taq* DNA polymerase (see UNIT 3.5), is designed to optimize the reaction components and conditions in one or two stages. The first stage (steps 1 to 7) determines the optimal MgCl_2 concentration and screens several enhancing additives. Most suppliers (of which there are many) of *Taq* and other thermostable DNA polymerases provide a unique optimized MgCl_2 -free buffer with MgCl_2 in a separate vial for user titration. The second stage (steps 8 to 13) compares methods for preventing pre-PCR low-stringency primer extension, which can generate nonspecific products. This has come to be known as “hot start,” whether one omits an essential reaction component prior to the first denaturing-temperature step or adds a reversible inhibitor of polymerase. Hot-start methods can greatly improve specificity, sensitivity, and yield. Use of any one of the hot-start approaches is strongly recommended if primer-dimers or other nonspecific products are generated, or if relatively rare template DNA is contained in a complex mixture, such as viral nucleic acids in cell or tissue preparations. This protocol suggests some relatively inexpensive methods to achieve hot start, and lists several commercial hot-start options which may be more convenient, but of course more expensive.

NOTE: Use only molecular biology-grade water (i.e., DNase, RNase, and nucleic acid free) in all steps and solutions.

Materials

- 10× MgCl_2 -free PCR buffer (see recipe)
- 50 μM oligonucleotide primer 1: 50 pmol/ μl in sterile H_2O (store at -20°C)
- 50 μM oligonucleotide primer 2: 50 pmol/ μl in sterile H_2O (store at -20°C)
- Template DNA: 1 μg mammalian genomic DNA or 1.0 to 100.0 pg of plasmid DNA (UNIT 2.1-2.4)
- 25 mM 4dNTP mix (see recipe)
- 5 U/ μl *Taq* DNA polymerase (native or recombinant)
- Enhancer agents (optional; see recipe)
- 15 mM (L), 30 mM (M), and 45 mM (H) MgCl_2
- Mineral oil
- TaqStart Antibody (Clontech)
- Ficoll 400 (optional): prepare as 10× stock; store indefinitely at room temperature
- Tartrazine dye (optional): prepare as 10× stock; store indefinitely at room temperature

**The Polymerase
Chain Reaction**

15.1.1

Table 15.1.1 Master Mixes for Optimizing Reaction Components

Components	Final concentration	Per reaction	Master mix ^a (μl)			
			I	II	III	IV
10× MgCl ₂ -free PCR buffer	1×	10 μl	40.0	40.0	40.0	40.0
Primer 1	0.5 μM	1 μl	4.0	4.0	4.0	4.0
Primer 2	0.5 μM	1 μl	4.0	4.0	4.0	4.0
Template DNA ^b	Undiluted	1 vol	4 vol	4 vol	4 vol	4 vol
25 mM 4dNTP mix ^c	0.2 mM	0.8 μl	3.2	3.2	3.2	3.2
<i>Taq</i> polymerase	2.5 U	0.5 μl	2.0	2.0	2.0	2.0
DMSO ^d (20×)	5%	5 μl	—	20.0	—	—
Glycerol ^d (10×)	10%	10 μl	—	—	40.0	—
PMPE ^d (100×)	1%	1 μl	—	—	—	4.0
H ₂ O	—	To 90 μl	To 360	To 360	To 360	To 360

^aTotal volume = 360 μl (i.e., enough for $n + 1$ reactions).

^bTemplate DNA volume (“vol”) is generally 1 to 10 μl.

^cIf 2 mM 4dNTP mix is preferred, use 10 μl per reaction, or 40 μl for each master mix; adjust the volume of water accordingly.

^dSubstitute with other enhancer agents (see recipe in Reagents and Solutions) as available.

0.5 ml thin-walled PCR tubes
Automated thermal cycler

Additional reagents and equipment for DNA preparation (UNIT 2.1-2.4, agarose gel electrophoresis (UNIT 2.5A), nondenaturing PAGE (UNIT 2.7), or sieving agarose gel electrophoresis (UNIT 2.8), restriction endonuclease digestion (UNIT 3.1), and Southern blotting and hybridization (UNITS 2.9 & 6.4)

NOTE: Do not use DEPC to treat water, reagents, or glassware.

NOTE: Reagents should be prepared in sterile, disposable labware, taken directly from its packaging, or in glassware that has been soaked in 10% bleach, thoroughly rinsed in tap water followed by distilled water, and if available, exposed to UV irradiation for ~10 min. Multiple small volumes of each reagent should be stored in screw-cap tubes. This will then serve as the user’s own optimization “kit.” Thin-walled PCR tubes are recommended.

Optimize reaction components

1. Prepare four reaction master mixes according to the recipes given in Table 15.1.1.

Enhancing agents probably work by different mechanisms, such as protecting enzyme activity and decreasing nonspecific primer binding. However, their effects cannot be readily predicted—what improves amplification efficiency for one primer pair may decrease the amplification efficiency for another. Thus it is best to check a panel of enhancers during development of a new assay.

2. Aliquot 90 μl master mix I into each of three 0.5-ml thin-walled PCR tubes labeled I-L, I-M, and I-H. Similarly, aliquot mixes II through IV into appropriately labeled tubes. Add 10 μl of 15 mM MgCl₂ into one tube of each master mix (labeled L; 1.5 mM final). Similarly, aliquot 10 μl of 30 mM and 45 mM MgCl₂ to separate tubes of each master mix (labeled M and H, respectively; 3.0 and 4.5 mM final concentrations respectively).

It is helpful to set the tubes up in a three-by-four array to simplify aliquotting. Each of the three Mg²⁺ concentrations is combined with each of the four master mixes.

3. Overlay the reaction mixture with 50 to 100 μ l mineral oil (2 to 3 drops).

To include hot start in the first step, overlay reaction mixes with oil before adding the MgCl₂, heat the samples to 95°C in the thermal cycler or other heating block, and add the MgCl₂ once the elevated temperature is reached. Once the MgCl₂ has been added, do not allow the samples to cool below the optimum annealing temperature prior to performing PCR.

Alternatives to mineral oil include silicone oil and paraffin beads. Additionally, certain cyclers feature heated lids that are designed to obviate the need for an oil overlay.

Choose cycling parameters

4. Using the following guidelines, program the automated thermal cycler according to the manufacturer's instructions.

30 cycles: 30 sec	94°C	(denaturation)
30 sec	55° (GC content \leq 50%) or	
	60°C (GC content $>$ 50%)	(annealing)
~60 sec/kb		
product sequence	72°C	(extension)

Cycling parameters are dependent upon the sequence and length of the template DNA, the sequence and percent complementarity of the primers, and the ramp times of the thermal cycler used. Thoughtful primer design will reduce potential problems (see Commentary). Denaturation, annealing, and extension are each quite rapid at the optimal temperatures. The time it takes to achieve the desired temperature inside the reaction tube (i.e., the ramp time) is usually longer than either denaturation or primer annealing. Thus, ramp time is a crucial cycling parameter. Manufacturers of the various thermal cyclers on the market provide ramp time specifications for their instruments. Ramp times are lower with thin-walled reaction tubes. The optimal extension time also depends on the length of the target sequence. Allow ~1 min/kb for this step for target sequences $>$ 1 kb, and as little as a 2-sec pause for targets $<$ 100 bases in length.

The number of cycles depends on both the efficiency of the reaction and the amount of template DNA in the reaction. Starting with as little as 100 ng of mammalian genomic DNA ($\sim 10^4$ cell equivalents), after 30 cycles, 10% of the reaction should produce a band that is readily visible on an ethidium bromide-stained gel as a single predominant band. With more template, fewer cycles may suffice. With much less template, further optimization is recommended rather than increasing the cycle number. Greater cycle numbers (e.g., $>$ 40) can reduce the polymerase specific activity, increase nonspecific amplification, and deplete substrate (nucleotides). Many investigators lengthen the time for the last extension step—to 7 min, for example—to try to ensure that all the PCR products are full length.

These guidelines are appropriate for most commercially available thermal cyclers. For rapid cyclers, consult the manufacturers' protocols.

Analyze the product

5. Electrophorese 10 μ l from each reaction on an agarose (UNIT 2.5A), nondenaturing polyacrylamide (UNIT 2.7), or sieving agarose gel (UNIT 2.8) appropriate for the PCR product size expected. Stain with ethidium bromide.

For resolution of PCR products between 100 and 1000 bp, an alternative to nondenaturing polyacrylamide gels or sieving agarose is a composite 3% (w/v) NuSieve (FMC Bioproducts) agarose/1% (w/v) SeaKem (FMC Bioproducts) agarose gel. SeaKem increases the mechanical strength of the gel without decreasing resolution.

An alternative to ethidium bromide, SYBR Gold Nucleic Acid Gel Stain (Molecular Probes), is 25 to 100 times more sensitive than ethidium bromide, is more convenient to use, and permits optimization of 10- to 100-fold lower starting template copy number.

Table 15.1.2 Master Mixes for Optimizing First-Cycle Reactions

Components	Final concentration	Master mix (μl)			
		A	B	C	D
10 \times PCR buffer	1 \times	10	10	10	10
MgCl ₂ (L, M, or H)	Optimal	10	10	10	10
Primer 1	0.5 μM	1.0	1.0	1.0	1.0
Primer 2	0.5 μM	1.0	1.0	1.0	1.0
Additive	Optimal	V ^a	V ^a	V ^a	V ^a
Template DNA	— ^b	V ^a	V ^a	V ^a	V ^a
25 mM 4dNTP mix ^b	0.2 mM	0.8	0.8	0.8	0.8
<i>Taq</i> polymerase	2.5 U	0.5	0.5	—	—
<i>Taq</i> pol + <i>Taq</i> Start	2.5 U	—	—	—	1.0
H ₂ O	To 100 μl	V ^a	V ^a	V ^a	V ^a
Preparation temperature		Room temperature	Ice slurry	Room temperature	Room temperature

^aV, variable amount (total volume should be 100 μl).

^bUse undiluted or diluted template DNA based on results obtained in step 6.

- Examine the stained gel to determine which condition resulted in the greatest amount of product.

Minor, nonspecific products may be present even under optimal conditions.

- To ensure that the major product is the correct one, digest an aliquot of the reaction with a restriction endonuclease known to cut within the PCR product. Check buffer compatibility for the restriction endonuclease of choice. If necessary, add Na⁺ or precipitate in ethanol (*UNIT 2.1A*) and resuspend in the appropriate buffer. Electrophore the digestion product on a gel to verify that the resulting fragments have the expected sizes.

Alternatively, transfer the PCR products to a nitrocellulose or nylon filter and hybridize with an oligonucleotide derived from the sequence internal to the primers (UNITS 2.9 & 6.4). With appropriately stringent hybridization and washing conditions, only the correct product (and possibly some minor related products) should hybridize.

Optimize the first cycle

These optional steps optimize initial hybridization and may improve efficiency and yield. They are used when primer-dimers and other nonspecific products are detected, when there is only a very small amount of starting template, or when a rare sequence is to be amplified from a complex mixture. For an optimal reaction, polymerization during the initial denaturation and annealing steps should be prevented. *Taq* DNA polymerase activity can be inhibited by temperature (reaction B), physical separation (reaction C), or reversible antibody binding (reaction D). PCR without hot start is performed for comparison (reaction A).

- Prepare four reaction mixtures using the optimal MgCl₂ concentration and additive requirement determined in step 6. Prepare the mixes according to the recipes in Table 15.1.2. Use the following variations for addition of *Taq* polymerase.
 - Prepare reactions A and C at room temperature.
 - Chill all components of reaction B in an ice slurry before they are combined.

- c. For reaction D, combine 1.0 μl TaqStart antibody with 4.0 μl of the dilution buffer provided with the antibody, add 1.0 μl *Taq* DNA polymerase (for 1:4:1 mixture of these components), mix, and incubate 5 to 10 min at room temperature before adding to reaction mixture D (glycerol and PMPE are compatible with TaqStart antibody but DMSO will interfere with antibody binding).

To ensure that the reaction does not plateau and thereby obfuscate the results, use the smallest amount of template DNA necessary for visualization of the PCR product by ethidium bromide staining. Use the results from step 6 to decide how much template to use. If the desired product stains intensely, dilute the starting material as much as 1/100. If only a faint signal is apparent, use undiluted sample.

9. Overlay each reaction mixture with 50 to 100 μl mineral oil.
10. Heat all reactions 5 min at 94°C.

It is most convenient to use the automated thermal cycler for this step and then initiate the cycling program directly.

11. Cool the reactions to the appropriate annealing temperature as determined in step 4. Add 0.5 μl *Taq* DNA polymerase to reaction C, making sure the pipet tip is inserted through the layer of mineral oil into the reaction mix.

*Time is also an important factor in this step. If the temperature drops below the annealing temperature and is allowed to remain low, nonspecific annealing will occur. *Taq* DNA polymerase retains some activity even at room temperature.*

12. Begin amplification of all four reactions at once, using the same cycling parameters as before.
13. Analyze the PCR products on an agarose gel and evaluate the results as in steps 5 and 6.
14. Prepare a batch of the optimized reaction mixture, but omit *Taq* DNA polymerase, TaqStart antibody, PMPE, and 4dNTP mix—these ingredients should be added fresh just prior to use. If desired, add Ficoll 400 to a final concentration of 0.5% to 1% (v/v) and tartrazine to a final concentration of 1 mM.

Adding Ficoll 400 and tartrazine dye to the reaction mix precludes the need for a gel loading buffer and permits direct application of PCR products to agarose or acrylamide gels. At these concentrations, Ficoll 400 and tartrazine do not decrease PCR efficiency and do not interfere with PMPE or TaqStart antibodies. Other dyes, such as bromphenol blue and xylene cyanol, do inhibit PCR. Tartrazine is a yellow dye and is not as easily visualized as other dyes; this may make gel loading more difficult.

Ficoll 400 and tartrazine dye may be prepared as 10 \times stocks and stored indefinitely at room temperature.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Enhancer agents

For a discussion of how to select enhancer agents, see Commentary.

5 \times stocks:

25% acetamide (20 μl /reaction; 5% final)

5 M *N,N,N*-trimethylglycine (betaine; 20 μl /reaction; 1 M final)

40% polyethylene glycol (PEG) 8000 (20 μl /reaction; 8% final)

10× stocks:

Glycerol (concentrated; 10 µl/reaction; 10% final)

20× stocks:

Dimethylsulfoxide (DMSO; concentrated; 5 µl/reaction; 5% final)

Formamide (concentrated; 5 µl/reaction; 5% final)

100× stocks:

1 U/µl Perfect Match Polymerase Enhancer (Stratagene; 1 µl/reaction; 1 U final)

10 mg/ml acetylated bovine serum albumin (BSA) or gelatin (1 µl/reaction; 10 µg/ml final)

1 to 5 U/µl thermostable pyrophosphatase (PPase; Roche Diagnostics; ; 1 µl/reaction; 1 to 5 U final)

5 M tetramethylammonium chloride (TMAC; betaine hydrochloride; 1 µl/reaction; 50 mM final)

0.5 mg/ml *E. coli* single-stranded DNA-binding protein (SSB; Sigma; 1 µl/reaction; 5 µg/ml final)

0.5 mg/ml Gene 32 protein (Amersham Pharmacia Biotech; 1 µl/reaction; 5 µg/ml final)

10% Tween 20, Triton X-100, or Nonidet P-40 (1 µl/reaction; 0.1% final)

1 M (NH₄)₂SO₄ (1 µl/reaction; 10 mM final; use with thermostable DNA polymerases other than *Taq*)

MgCl₂-free PCR buffer, 10×

500 mM KCl

100 mM Tris·Cl, pH 9.0 (at 25°C; see APPENDIX 2)

0.1% Triton X-100

Store indefinitely at –20°C

This buffer can be obtained from Promega; it is supplied with Taq DNA polymerase.

4dNTP mix

For 2 mM 4dNTP mix: Prepare 2 mM each dNTP in TE buffer, pH 7.5 (APPENDIX 2). Store up to 1 year at –20°C in 1-ml aliquots.

For 25 mM 4dNTP mix: Combine equal volumes of 100 mM dNTPs (Promega). Store indefinitely at –20°C in 1-ml aliquots.

COMMENTARY

Background Information

The theoretical basis of the polymerase chain reaction (PCR; see chapter introduction) was probably first described in a paper by Kleppe et al. (1971). However, this technique did not excite general interest until the mid-1980s, when Kary Mullis and co-workers at Cetus developed PCR into a technique that could be used to generate large amounts of single-copy genes from genomic DNA (Saiki et al., 1985, 1986; Mullis et al., 1986; Embury et al., 1987).

The initial procedure entailed adding a fresh aliquot of the Klenow fragment of *E. coli* DNA polymerase I during each cycle because this enzyme was inactivated during the subsequent denaturation step. The introduction of thermostable *Taq* DNA polymerase from *Thermus*

aquaticus (Saiki et al., 1988) alleviated this tedium and facilitated automation of the thermal cycling portion of the procedure. *Taq* DNA polymerase also permitted the use of higher temperatures for annealing and extension, which improved the stringency of primer–template hybridization and thus the specificity of the products. This also served to increase the yield of the desired product.

All applications of PCR depend upon an optimized PCR. The basic protocol in this unit optimizes PCR for several variables, including MgCl₂ concentration, enhancing additives—dimethyl sulfoxide (DMSO), glycerol, or Perfect Match Polymerase Enhancer (PMPE)—and prevention of pre-PCR mispriming. These and other parameters can be extremely impor-

tant, as every element of PCR can affect the outcome (see Critical Parameters and Troubleshooting for discussion of individual parameters).

There are several PCR optimization kits and proprietary enhancers on the market (Table 15.1.3). Optimization kits generally provide a panel of buffers in which the pH, buffer, non-ionic detergents, and addition of $(\text{NH}_4)_2\text{SO}_4$ are varied, MgCl_2 may be added at several concentrations, and enhancers (e.g., DMSO, glycerol, formamide, betaine, and/or proprietary compounds) may be chosen. The protocol presented here is aimed at keeping the costs low and the options broad.

Critical Parameters and Troubleshooting

MgCl₂ concentration

Determining the optimum MgCl_2 concentration, which can vary even for different primers from the same region of a given template (Saiki, 1989), can have an enormous influence on PCR success. In this protocol three concentrations are tested—1.5 mM (L), 3.0 mM (M), and 4.5 mM (H)—against three enhancers. Enhancers tend to broaden the MgCl_2 optimal range, contributing to the success of the PCR at one of these concentrations. A 10× buffer optimized for a given enzyme and a separate vial of MgCl_2 are typically provided with the polymerase, so that the user may titrate the MgCl_2 concentration for their unique primer-template set.

Reagent purity

For applications that amplify rare templates, reagent purity is the most important parameter, and avoiding contamination at every step is critical.

To maintain purity, store multiple small volumes of each reagent in screw-cap tubes.

For many applications, simply using high-quality reagents and avoiding nuclease contamination is sufficient; however, avoid one common reagent used to inactivate nucleases, diethylpyrocarbonate (DEPC). Even tiny amounts of chemical left after treatment of water by autoclaving are enough to ruin a PCR.

Primer selection

This is the factor that is least predictable and most difficult to troubleshoot. Simply put, some primers just do not work. To maximize the probability that a given primer pair will

work, pay attention to the following parameters.

General considerations. An optimal primer set should hybridize efficiently to the sequence of interest with negligible hybridization to other sequences present in the sample. If there are reasonable amounts of template available, hybridization specificity can be tested by performing oligonucleotide hybridization as described in UNIT 6.4. The distance between the primers is rather flexible, ranging up to 10 kb. There can be, however, a considerable drop-off in synthesis efficiency with distances >3 kb (Jeffreys et al., 1988). Small distances between primers, however, lessen the ability to obtain much sequence information or to reamplify with nested internal oligonucleotides, should that be necessary.

Design primers to allow demonstration of the specificity of the PCR product. Be sure that there are diagnostic restriction endonuclease sites between the primers or that an oligonucleotide can detect the PCR product specifically by hybridization.

Several computer programs can assist in primer design (see Internet Resources at end of unit). These are most useful for avoiding primer sets with intra- and intermolecular complementarity, which can dramatically raise the effective T_m . Given the abundance of primers relative to template, this can preclude template priming. Computer primer design is not foolproof. If possible, start with a primer or primer set known to efficiently prime extensions. In addition, manufacturers' Web sites offer technical help with primer design.

Complementarity to template. For many applications, primers are designed to be exactly complementary to the template. For others, however, such as engineering of mutations or new restriction endonuclease sites, or for efforts to clone or detect gene homologs where sequence information is lacking, base-pair mismatches will be intentionally or unavoidably created. It is best to have mismatches (e.g., in a restriction endonuclease linker) at the 5' end of the primer. The closer a mismatch is to the 3' end of the primer, the more likely it is to prevent extension.

The use of degenerate oligonucleotide primers to clone genes where only protein sequence is available, or to fish out gene homologs in other species, has sometimes been successful, but it has also failed an untold (and unpublished) number of times. When the reaction works it can be extremely valuable, but it can

Table 15.1.3 PCR Optimization Products

Optimization goal	Supplier	Product
Optimization support	Perkin-Elmer	Technical information in appendix to catalog
Optimization support	Promega	PCR troubleshooting program on the Internet: http://www.promega.com/amplification/assistant
Optimization kits	Boehringer-Mannheim, Invitrogen, Stratagene, Sigma, Epicentre Technologies, Life Technologies	Several buffers, Mg ²⁺ , and enhancers which may include DMSO, glycerol, formamide, (NH ₄) ₂ SO ₄ , and other unspecified or proprietary agents
Quick startup	Amersham Pharmacia Biotech	Ready-To-Go Beads “optimized for standard PCR” and Ready-To-Go RAPD Analysis Beads (buffer, nucleotides, <i>Taq</i> DNA polymerase)
Quick startup	Fisher	EasyStart PCR Mix-in-a-Tube—tubes prepackaged with wax beads containing buffer, MgCl ₂ , nucleotides, <i>Taq</i> DNA polymerase
Quick startup	Life Technologies	PCR SuperMix—1.1× conc.—premix containing buffer, MgCl ₂ , nucleotides, <i>Taq</i> DNA polymerase
Quick startup	Marsh Biomedical	Advanced Biochemicals Red Hot DNA Polymerase—a new rival for <i>Taq</i> polymerase with convenience features
Hot-start/physical barrier	Fisher, Life Technologies	Molecular Bio-Products HotStart Storage and Reaction Tubes—preadhered wax bead in each tube; requires manual addition of one component at high temperature
Hot-start/separate MgCl ₂	Invitrogen	HotWax Mg ²⁺ beads—wax beads contain preformulated MgCl ₂ which is released at first elevated-temperature step
Hot-start/separate MgCl ₂	Stratagene	StrataSphere Magnesium Wax Beads—wax beads containing preformulated Mg ²⁺
Hot Start/separate polymerase	Promega	TaqBead Hot Start Polymerase—wax beads encapsulating <i>Taq</i> DNA polymerase which is released at first elevated-temperature step
Hot-start/reversible inactivation of polymerase by antibody binding	Clontech	TaqStart Antibody, TthStart Antibody—reversibly inactivate <i>Taq</i> and <i>Tth</i> DNA polymerases until first denaturation at 95°C
Hot-start/antibody binding	Life Technologies	PlatinumTaq—contains PlatinumTaq antibody
Hot-start/antibody binding	Sigma	JumpStart Taq—contains TaqStart antibody
Hot-start/reversible chemical modification	Perkin-Elmer	AmpliTaq Gold—activated at high temperature
Hot-start/reversible chemical modification	Qiagen	HotStarTaq DNA Polymerase—activated at high temperature
Enhancer	Boehringer Mannheim, New England Biolabs	<i>Tth</i> pyrophosphatase, thermostable
Enhancer	Clontech	GC-Melt (in Advantage-GC Kits)—proprietary
Enhancer	CPG	Taq-FORCE Amplification System and MIGHTY Buffer—proprietary
Enhancer	Fisher	Eppendorf MasterTaq Kit with TaqMaster Enhancer—proprietary
Enhancer	Life Technologies	PCRx Enhancer System—proprietary
Enhancer	Promega	<i>E.coli</i> Single Stranded Binding Protein (SSB)
Enhancer	Qiagen	Q-Solution—proprietary
Enhancer	Stratagene	Perfect Match Polymerase Enhancer—proprietary
Enhancer	Stratagene	TaqExtender PCR Additive—proprietary

15.1.8

also generate seemingly specific products that require much labor to identify and yield no useful information. The less degenerate the oligonucleotides, especially at the 3' end, the better. Caveat emptor.

Primer length. A primer should be 20 to 30 bases in length. It is unlikely that longer primers will help increase specificity significantly.

Primer sequence. Design primers with a GC content similar to that of the template. Avoid primers with unusual sequence distributions, such as stretches of polypurines or polypyrimidines, as their secondary structure can be disastrous. It is worthwhile to check for potential secondary structure using one of the appropriate computer programs that are available.

"Primer-dimers." Primer-dimers are a common artifact most frequently observed when small amounts of template are taken through many amplification cycles. They form when the 3' end of one primer anneals to the 3' end of the other primer, and polymerase then extends each primer to the end of the other. The ensuing product can compete very effectively against the PCR product of interest. Primer-dimers can best be avoided by using primers without complementarity, especially in their 3' ends. Should they occur, optimizing the MgCl₂ concentration may minimize their abundance relative to that of the product of interest.

Template

Aside from standard methods for preparing DNA (UNIT 2.1-2.4), a number of simple and rapid procedures have been developed for particular tissues (Higuchi, 1989). Even relatively degraded DNA preparations can serve as useful templates for generation of moderate-sized PCR products. The two main concerns regarding template are purity and amount.

A number of contaminants found in DNA preparations can decrease the efficiency of PCR. These include urea, the detergent SDS (whose inhibitory action can be reversed by nonionic detergents), sodium acetate, and, sometimes, components carried over in purifying DNA from agarose gels (Gelfand, 1989; Gyllenstein, 1989; K. Hicks and D. Coen, unpub. observ.). Additional organic extractions, ethanol precipitation from 2.5 M ammonium acetate, and/or gel purification on polyacrylamide rather than agarose, can all be beneficial in minimizing such contamination if the simplest method (precipitating the sample with ethanol and repeatedly washing the pellet with 70% ethanol) is not sufficient.

Clearly the amount of template must be sufficient to be able to visualize PCR products using ethidium bromide. Usually 100 ng of genomic DNA is sufficient to detect a PCR product from a single-copy mammalian gene. Using too much template is not advisable when optimizing for MgCl₂ or other parameters, as it may obscure differences in amplification efficiency. Moreover, too much template may decrease efficiency due to contaminants in the DNA preparation.

Amount of template, especially in terms of the amount of target sequence versus nonspecific sequences, can have a major effect on the yield of nonspecific products. With less target sequence, it is more likely that nonspecific products will be seen. For some applications, such as certain DNA sequencing protocols where it is important to have a single product, gel purification of the specific PCR product and reamplification are advisable.

Taq and other thermostable DNA polymerases

Among the advantages conferred by the thermostability of *Taq* DNA polymerase is its ability to withstand the repeated heating and cooling inherent in PCR and to synthesize DNA at high temperatures that melt out mismatched primers and regions of local secondary structure. The enzyme, however, is not infinitely resistant to heat, and for greatest efficiency it should not be put through unnecessary denaturation steps. Indeed, some protocols (e.g., UNIT 15.7 and the "hot start" method described here) recommend adding it after the first denaturation step.

Increasing the amount of *Taq* DNA polymerase beyond 2.5 U/reaction can sometimes increase PCR efficiency, but only up to a point. Adding more enzyme can sometimes increase the yield of nonspecific PCR products at the expense of the product of interest. Moreover, *Taq* DNA polymerase is not inexpensive.

A very important property of *Taq* DNA polymerase is its error rate, which was initially estimated at 2×10^{-4} nucleotides/cycle (Saiki et al., 1988). The purified enzyme supplied by manufacturers lacks a proofreading 3'→5' exonuclease activity, which lowers error rates of other polymerases such as the Klenow fragment of *E. coli* DNA polymerase I. For many applications, this does not present any difficulties. However, for sequencing clones derived from PCR, or when starting with very few templates, this can lead to major problems. Direct sequencing of PCR products (UNIT 15.2),

sequencing numerous PCR-generated clones, and/or the use of appropriate negative controls can help overcome these problems. Alternatively, changing reaction conditions (Eckert and Kunkel, 1990) or changing to a non-*Taq* DNA polymerase (with greater fidelity) may be useful.

Another important property of *Taq* DNA polymerase is its propensity for adding nontemplated nucleotides to the 3' ends of DNA chains. This can be especially problematic in cloning PCR products. It is frequently necessary to "polish" PCR products with enzymes such as other DNA polymerases before adding linkers or proceeding to blunt-end cloning. Conversely, addition of a nontemplated A by *Taq* DNA polymerase can be advantageous in cloning (UNIT 15.4).

Certain PCR protocols may work better with one thermostable polymerase rather than another. Table 15.1.4 lists currently available thermostable DNA polymerases by generic and trade names, the original source of native and recombinant enzymes, the supplier, the end generated (3'A addition versus blunt), and associated exonuclease activities. A 3' to 5' exonuclease activity is proofreading. Removal of the 5' to 3' exonuclease activity of *Taq* DNA polymerase (N-terminal deletion) is reported to produce a higher yield. A 5' to 3' exonuclease activity may degrade the primers somewhat. Proofreading enzymes synthesize DNA with higher fidelity and can generate longer products than *Taq*, but tend to generate low yields. Enzyme blends (Table 15.1.5) have been optimized for increased fidelity and length along with sensitivity and yield.

Hot start

What happens prior to thermal cycling is critical to the success of PCR. *Taq* DNA polymerase retains some activity even at room temperature. Therefore, under nonstringent annealing conditions, such as at room temperature, products can be generated from annealing of primers to target DNA at locations of low complementarity or having complementarity of just a few nucleotides at the 3' ends. The latter would in effect create new templates "tagged" with the primer sequences. Subsequent cycles amplify these tagged sequences in abundance, both generating nonspecific products and possibly reducing amplification efficiency of specific products by competition for substrates or polymerase. Thus conditions preventing polymerization prior to the first temperature-controlled steps are desirable. In this protocol, three

methods of inhibiting polymerization prior to the temperature-controlled step are compared. These include physical separation of an essential reaction component prior to the first denaturation step, cooling reagents to 0°C, and reversibly blocking enzymatic activity with an antibody.

Denaturation of the template before *Taq* polymerase or MgCl₂ is added to the reaction provides a dramatic improvement in specificity and sensitivity in many cases (Chou et al., 1992). The main drawback of this method is that it requires opening the reaction tubes a second time to add the essential missing component. This creates both an inconvenience and an increase in the risk of contamination, an important consideration when testing for the presence of a given sequence in experimental or clinical samples.

Cooling all components of the reaction mixture to 0°C prior to mixing is more convenient and the least expensive method but is also the least reliable. Transferring the PCR reaction tubes from the ice slurry to a 95°C preheated thermocycler block may improve the chance of success.

Reversible inhibition of *Taq* DNA polymerase by TaqStart antibody (Clontech) is the most convenient method and very effective (Kellogg et al., 1994). Complete reactions can be set up, overlaid with oil, and stored at 4°C for up to several hours prior to thermal cycling with no loss of sensitivity or specificity compared to the other hot start methods (M.F. Kramer and D.M. Coen, unpub. observ.). Cycling is initiated immediately following 5-min denaturation of the antibody at 94°C. DMSO inhibits antibody binding and should not be used with TaqStart.

Several hot-start products are now commercially available (Table 15.1.3). Success with each may depend on strict adherence to the manufacturer's protocols, even on a specific thermocycler. Wax barrier and reversible antibody binding methods are more forgiving, while chemical modifications have more stringent activation temperature requirements.

Deoxyribonucleoside triphosphates

In an effort to increase efficiency of PCR, it may be tempting to increase the concentration of dNTPs. Don't! When each dNTP is 200 mM, there is enough to synthesize 12.5 mg of DNA when half the dNTPs are incorporated. dNTPs chelate magnesium and thereby change the effective optimal magnesium concentration. Moreover, dNTP concentrations >200

Table 15.1.4 Thermostable DNA Polymerases

DNA polymerase		Biological source	Supplier	Product ends	Exonuclease activity
Generic name	Trade name				
<i>Pfu</i>	—	<i>Pyrococcus furiosus</i>	Stratagene, Promega	Blunt	3'-5' (proofreading)
<i>Pfu</i> (exo-)	—	<i>Pyrococcus furiosus</i>	Stratagene	Blunt	No
<i>Psp</i>	Deep Vent	<i>Pyrococcus</i> sp.GB-D	New England Biolabs	Blunt	3'-5' (proofreading)
<i>Psp</i> (exo-)	Deep Vent (exo-)	<i>Pyrococcus</i> sp.GB-D	New England Biolabs	Blunt	No
<i>Pwo</i>	—	<i>Pyrococcus woesei</i>	Boehringer Mannheim	Blunt	3'-5' (proofreading)
<i>Taq</i> (native and/or recombinant)	—	<i>Thermus aquaticus</i>	Ambion, Amersham Pharmacia Biotech, Boehringer Mannheim, Clontech, Fisher, Life Technologies, Marsh Biomedical, Perkin Elmer, Promega, Qiagen, Sigma, Stratagene	3'A	5'-3'
<i>Taq</i> , N-terminal deletion	Stoffel fragment Klen-Taq	<i>Thermus aquaticus</i>	Perkin-Elmer, Sigma	3'A	No
<i>Tbr</i>	DyNAzyme	<i>Thermus brocianus</i>	MJ Research	— ^a	5'-3'
<i>Tfl</i>	—	<i>Thermus flavus</i>	Promega, Epicentre Technologies	Blunt	— ^a
<i>Tli</i>	Vent	<i>Thermococcus litoralis</i>	New England Biolabs (Vent), Promega	Blunt	3'-5' (proofreading)
<i>Tli</i> (exo-)	Vent (exo-)	<i>Thermococcus litoralis</i>	New England Biolabs	Blunt	No
<i>Tma</i>	UITma	<i>Thermotoga maritima</i>	Perkin-Elmer	Blunt	3'-5' (proofreading)
<i>Tth</i>	—	<i>Thermus thermophilus</i>	Amersham Pharmacia Biotech, Boehringer Mannheim, Epicentre Technologies, Perkin Elmer, Promega	3' A	5'-3'

^aNo information at this time.

mM each increase the error rate of the polymerase. Millimolar concentrations of dNTPs actually inhibit *Taq* DNA polymerase (Gelfand, 1989).

The protocol in this unit calls for preparing 4dNTPs in 10 mM Tris·Cl/1 mM EDTA (TE buffer), pH 7.4 to 7.5. This is easier and less prone to disaster than neutralization with sodium hydroxide. However, EDTA also chelates magnesium, and this should be taken into account if stocks of dNTPs are changed. Alternatively, to lower the risk of contamination, a 4dNTP mix can be made by combining equal volumes of commercially prepared stocks.

Enhancers

Enhancers are used to increase yield and specificity and to overcome difficulties encountered with high GC content or long templates. Nonionic detergents (Triton X-100, Tween 20, or Nonidet P-40) neutralize charges of ionic detergents often used in template preparation, and should be used in the basic reaction mixture, rather than as optional enhancers. Higher yields can be achieved by stabilizing/enhancing the polymerase activity with enzyme-stabilizing proteins (BSA or gelatin), enzyme-stabilizing solutes such as betaine or betaine-HCl (TMAC), enzyme-stabilizing solvents (glycerol), solubility-enhancing sol-

Table 15.1.5 Thermostable DNA Polymerase Blends

Product (trade name)	Supplier	Thermostable DNA polymerases and other components
Expand High Fidelity, Expand Long Template, and Expand 20kb PCR Systems	Boehringer Mannheim	<i>Taq</i> + <i>Pwo</i>
KlenTaq LA Polymerase Mix	Clontech, Sigma	KlenTaq-1 (5'-exonuclease deficient <i>Taq</i>) + unspecified proofreading polymerase
Advantage-HF PCR Kit	Clontech	KlenTaq-1 + unspecified proofreading polymerase + TaqStart Antibody
Advantage-cDNA and Advantage-GC cDNA Polymerase Mixes and Kits	Clontech	KlenTaq-1 + unspecified proofreading polymerase + TaqStart Antibody; GC Kit contains GC Melt
Advantage Genomic and Advantage-GC Genomic Polymerase Mixes and Kits	Clontech	<i>Tth</i> + unspecified proofreading polymerase + TthStart Antibody; GC Kit contains GC Melt
eLONGase Enzyme Mix	Life Technologies	<i>Taq</i> + <i>Psp</i> + unspecified proofreading polymerase(s) + eLONGase Buffer
Platinum Taq DNA Polymerase	Life Technologies	<i>Taq</i> + <i>Psp</i> + Platinum <i>Taq</i> Antibody
Platinum High Fidelity DNA Polymerase	Life Technologies	<i>Taq</i> + <i>Psp</i> + <i>Taq</i> Antibody
DyNAzyme EXT Polymerase	MJ Research	<i>Tbr</i> with unspecified enhancer
GeneAmp XL PCR and XL RNA PCR Kits	Perkin-Elmer	<i>Tth</i> + <i>Tli</i>
OmniBase Sequencing Enzyme Mix	Promega	Unspecified proofreading polymerase(s) with thermostable pyrophosphatase
AccuTaq LA DNA Polymerase Mix	Sigma	<i>Taq</i> + unspecified proofreading polymerase
TaqPlus Long and TaqPlus Precision PCR Systems	Stratagene	<i>Pfu</i> + <i>Taq</i> ; TaqPlus Precision Reaction Buffer (proprietary)
Accurase Fidelity PCR Enzyme Mix; Calypso High Fidelity Single Tube RT-PCR System	Tetralink	<i>Thermus sp.</i> + <i>Thermococcus sp.</i> ; Calypso also contains AMV-RT

vents (DMSO or acetamide), molecular crowding solvents (PEG), and polymerase salt preferences [(NH₄)SO₄ is recommended for polymerases other than *Taq*]. Greater specificity can be achieved by lowering the T_M of dsDNA (using formamide), destabilizing mismatched-primer annealing (using PMPE or hot-start strategies), and stabilizing ssDNA (using *E. coli* SSB or T4 Gene 32 Protein). Amplification of high-GC-content templates can be improved by decreasing the base pair composition dependence of the T_M of dsDNA (with betaine; Rees et al., 1993). Betaine is an osmolyte widely distributed in plants and animals and is nontoxic, a feature that recommends it for convenience in handling, storage, and disposal. Betaine may be the proprietary ingredient in various commercial formulations. For long templates, a higher pH is recommended (pH 9.0). The pH of Tris buffer decreases at high temperatures, long-template PCR requires

more time at high temperatures, and increased time at lower pH may cause some depurination of the template, resulting in reduced yield of specific product. Inorganic phosphate (PPi), a product of DNA synthesis, may accumulate with amplification of long products to levels that may favor reversal of polymerization. Accumulation of PPi may be prevented by addition of thermostable PPase. When large numbers of samples are being analyzed, the convenience of adding PCR products directly to a gel represents a significant time savings. Some companies combine their thermostable polymerase with a red dye and a high density component to facilitate loading of reaction products onto gels without further addition of loading buffer.

Thermal cycling parameters

Each step in the cycle requires a minimal amount of time to be effective, while too much

time can be both wasteful and deleterious to the DNA polymerase. If the amount of time in each step can be reduced, so much the better.

Denaturation. It is critical that complete strand separation occur during the denaturation step. This is a unimolecular reaction which, in itself, is very fast. The suggested 30-sec denaturation used in the protocol ensures that the tube contents reach 94°C. If PCR is not working, it is well worth checking the temperature inside a control tube containing 100 µl water. If GC content is extremely high, higher denaturation temperatures may be necessary; however, *Taq* DNA polymerase activity falls off quickly at higher temperatures (Gelfand, 1989). To amplify a long sequence (>3 kb), minimize the denaturation time to protect the target DNA from possible effects, such as depurination, of lowered pH of the Tris buffer at elevated temperatures.

Annealing. It is critical that the primers anneal stably to the template. Primers with relatively low GC content (<50%) may require temperatures lower than 55°C for full annealing. On the other hand, this may also increase the quantity of nonspecific products. For primers with high GC content, higher annealing temperatures may be necessary. It can be worthwhile, although time-consuming, to experiment with this parameter. Some manufacturers have thermal cyclers on the market which are capable of forming a temperature gradient across the heating units, thus permitting annealing temperature optimization in one run. As with denaturation, the time for this step is based mainly on the time it takes to reach the proper temperature, because the primers are in such excess that the annealing reaction occurs very quickly.

Extension. The extension temperature of 72°C is close to the optimal temperature for *Taq* DNA polymerase (~75°C), yet prevents the primers from falling off. Indeed, primer extension begins during annealing, because *Taq* DNA polymerase is partially active at 55°C and even lower temperatures (Gelfand, 1989).

The duration of extension depends mainly on the length of the sequence to be amplified. A duration of 1 min per kb product length is usually sufficient.

Certain protocols, including others in this chapter, end the PCR with a long final extension time in an attempt to try to make products as complete as possible.

Ramp time. Ramp time refers to the time it takes to change from one temperature to another. Using water baths and moving samples

manually from temperature to temperature probably gives the shortest ramp times, which are mainly the time required for the tube's contents to change temperature. Different thermal cyclers have different ramp times; basically, the shorter the better.

The Stratagene Robocycler uses a robotic arm to move samples from one constant-temperature block to another, virtually eliminating block ramp time, but a ramp time for tube contents must be calculated (~1 sec/°C) and added to denaturation, annealing, and extension times. Rapid cyclers that utilize positive-displacement pipet tips or capillary tubes for the PCR reactions dramatically reduce the ramp times.

Generally, the more "high-performance" thermal cyclers with short ramp times are proportionally more costly. There are many new thermal cyclers on the market priced below \$5000, which perform quite well (Beck, 1998).

Anticipated Results

Starting with ≥100 ng mammalian DNA (≥10⁴ molecules), the basic protocol can be used to determine which MgCl₂ concentration, enhancing additive, and initial conditions will yield a predominant PCR product from a single-copy sequence that is readily visible on an ethidium bromide-stained gel. It is possible that other minor products will also be visible.

Time Considerations

The basic protocol can be completed in a single day. Assembly of the reaction mixtures should take ~1 hr. Cycling should take less than 3 hr. Preparing, running, and staining the gel should take another few hours. Further checks on specificity of the product such as restriction endonuclease digestion or Southern blot hybridization will take another few hours or days, respectively.

Literature Cited

- Beck, S. 1998. How low can you go? Nineteen thermal cyclers priced under \$5000. *The Scientist* 12:19-20.
- Chou, Q., Russell, M., Birch, D.E., Raymond, J., and Bloch, W. 1992. Prevention of pre-PCR mispriming and primer dimerization improves low-copy-number amplifications. *Nucl. Acids Res.* 20:1717-1723.
- Eckert, K.A. and Kunkel, T.A. 1990. High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucl. Acids Res.* 18:3739-3752.
- Embury, S.H., Scharf, S.J., Saiki, R.K., Gholson, M.A., Golbus, M., Arnheim, N., and Erlich, H.A. 1987. Rapid prenatal diagnosis of sickle cell

- anemia by a new method of DNA analysis. *N. Engl. J. Med.* 316:656-661.
- Gelfand, D.H. 1989. *Taq* DNA polymerase. In *PCR Technology: Principles and Applications for DNA Amplification* (H.A. Erlich, ed.) pp. 17-22. Stockton Press, New York.
- Gyllensten, U. 1989. Direct sequencing of in vitro amplified DNA. In *PCR Technology: Principles and Applications for DNA Amplification* (H.A. Erlich, ed.) pp. 45-60. Stockton Press, New York.
- Higuchi, R. 1989. Simple and rapid preparation of samples for PCR. In *PCR Technology: Principles and Applications for DNA Amplification* (H.A. Erlich, ed.) pp. 31-38. Stockton Press, New York.
- Jeffreys, A.J., Wilson, V., Neumann, R., and Keyte, J. 1988. Amplification of human minisatellites by the polymerase chain reaction: Towards DNA fingerprinting of single cells. *Nucl. Acids Res.* 16:10,953-10,971.
- Kellogg, D.E., Rybalkin, I., Chen, S., Mukhamedova, N., Vlasik, T., Siebert, P.D., and Chencik, A. 1994. TaqStart antibody: "Hot start" PCR facilitated by a neutralizing monoclonal antibody directed against *Taq* DNA polymerase. *BioTechniques* 16:1134-1137.
- Kleppe, K., Ohtsuka, E., Kleppe, R., Molineux, I., and Khorana, H.G. 1971. Studies on polynucleotides. XCVI. Repair replication of short synthetic DNA's as catalyzed by DNA polymerases. *J. Mol. Biol.* 56:341-361.
- Mullis, K.B., Faloona, F., Scharf, S.J., Saiki, R.K., Horn, G.T., and Erlich, H.A. 1986. Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. *Cold Spring Harbor Symp. Quant. Biol.* 51:263-273.
- Rees, W.A., Yager, T.D., Korte, J., and von Hippel, P.H. 1993. Betaine can eliminate the base pair composition dependence of DNA melting. *Biochemistry* 32:137-144.
- Saiki, R.K. 1989. The design and optimization of the PCR. In *PCR Technology: Principles and Applications for DNA Amplification* (H.A. Erlich, ed.) pp. 7-16. Stockton Press, New York.
- Saiki, R.K., Scharf, S., Faloona, F., Mullis, K., Horn, G., Erlich, H.A., and Arnheim, N. 1985. Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230:1350-1354.
- Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B., and Erlich, H.A. 1986. Analysis of enzymatically amplified β -globin and HLA-DQ α DNA with allele-specific oligonucleotide probes. *Nature* 324:163-166.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487-491.

Key Reference

Saiki et al., 1988. See above.

Demonstrates the ease and power of PCR using Taq DNA polymerase.

Internet Resources

<http://www.promega.com/amplification/amptech.html>

Offers Amplification Assistant, a PCR troubleshooting program.

<http://www.genome.wi.mit.edu/>

Provides access to www Primer Picking (Primer 3); select experimental web-based software under Genome Center Software.

<http://www.alkami.com/primers/>

Contains free primer design tools and tips.

<http://bioinformatics.weizmann.ac.il/mb/bioguide/pcr/contents.html>

Contains useful tips and links.

Contributed by Martha F. Kramer and
Donald M. Coen
Harvard Medical School
Boston, Massachusetts

Direct DNA Sequencing of PCR Products

UNIT 15.2

This unit describes two basic approaches for direct DNA sequencing of polymerase chain reaction (PCR) products. Either approach permits the rapid characterization of sequences of interest, without the need for library construction or screening.

PCR products can be sequenced using either the dideoxy (Sanger) approach or the chemical (Maxam-Gilbert) approach. Dideoxy sequencing methods are presented in Basic Protocol 1 and Alternate Protocols 1 to 4. In Basic Protocol 1 and Alternate Protocol 1, the target sequence is amplified and an excess of one strand of the target sequence (relative to its complement) is then generated by “asymmetric PCR,” where one primer is present in vast excess over the other. This single-stranded product serves as the template for conventional dideoxy sequencing methods (UNIT 7.4). The Alternate Protocol 2 describes how PCR products can be prepared so that they will be suitable templates for double-stranded dideoxy sequencing methods (UNIT 7.4). The Alternate Protocol 3 uses λ exonuclease to generate single-stranded template from double-stranded PCR products. Alternate Protocol 4 describes a one-step purification procedure that renders the PCR products suitable for direct sequencing.

Chemical sequencing methods are presented in Basic Protocol 2, where an end-labeled product is used, and in Alternate Protocol 5, where an unlabeled product is characterized using the genomic sequencing approach. The dideoxy method is somewhat simpler, and can be used for most applications, especially where short stretches of sequence are being characterized. The chemical sequencing method, although requiring more steps, is well suited for larger-scale sequencing projects where a number of separate sequences need to be amplified and determined, as in “multiplex sequencing” (Church and Kieffer-Higgins, 1988; see Commentary).

GENERATING SINGLE-STRANDED PRODUCTS FOR DIDEOXY SEQUENCING BY ASYMMETRIC PCR

BASIC
PROTOCOL 1

In this approach, an excess of one amplified strand (relative to its complement) is generated by the addition of one primer in vast excess over the other. The resulting excess of single-stranded product is then used as a template for the production of the dideoxy-terminated chains from which the sequence is derived.

Materials

- Oligonucleotide primers 1 and 2
- ^{32}P -labeled dNTPs (optional; UNIT 3.4)
- 10 M ammonium acetate (APPENDIX 2)
- 100% and 70% ethanol, room temperature
- 0.1 \times TE buffer, pH 8.0 (APPENDIX 2)
- Centricon 30 or 100 column (optional; Amicon)

Additional reagents and equipment for PCR (UNIT 15.1), electrophoresis using agarose gels (UNIT 2.5A) or nondenaturing polyacrylamide gels (UNIT 2.7), Southern blotting and hybridization (UNITS 2.9 & 6.4), ethanol precipitation (UNIT 2.1A), and dideoxy sequencing (UNIT 7.4)

1. Assemble a PCR with optimized components (UNIT 15.1) but use ~100:1 ratio of the two oligonucleotide primers.

The ideal amounts of the primers should be determined empirically, but will generally range from 0.2 to 1 pmol for the limiting primer, and from 10 to 30 pmol for the

The Polymerase
Chain Reaction

15.2.1

Contributed by Robert L. Dorit, Osamu Ohara, Charles B-C. Hwang, Jae Bum Kim and Seth Blackshaw

Current Protocols in Molecular Biology (2001) 15.2.1-15.2.13
Copyright © 2001 by John Wiley & Sons, Inc.

Supplement 56

primer present in excess. Because the primers are present in substantial excess under standard PCR conditions (~50 pmol of each primer), the limiting primer must be exhausted prior to the completion of PCR in order to produce an excess of one strand. Empirically, this means that one primer concentration must be kept below 1 pmol/100- μ l reaction.

2. Carry out PCR using optimized times and temperatures (UNIT 15.1) for 40 to 45 cycles, ending with a long extension step. Analyze an aliquot to confirm the presence of one predominant single-stranded product, proceeding to step 3 only upon verification.

The limiting primer is generally exhausted by the 25th cycle. Thereafter, amplification proceeds by the arithmetic accumulation of a single strand, initiated by the nonlimiting primer. The yield of single-stranded DNA can be verified by the addition of a small amount (0.1% of cold dNTP concentration) of [³²P]dNTPs, or of 0.001 to 0.1 pmol of [³²P] “nonlimiting” primer, during the final 3 to 5 rounds of amplification. In this way, amplified products can be directly visualized by autoradiography of an agarose or nondenaturing polyacrylamide gel. The single- and double-stranded products exhibit different electrophoretic mobilities. Include double-stranded DNA size markers to aid the analysis.

Alternatively, 2% to 5% of the unlabeled PCR products can be electrophoresed on an agarose or nondenaturing polyacrylamide gel. The agarose or polyacrylamide gel is then blotted using capillary or electroblotting methods (UNIT 2.9), and the resulting blots hybridized to a labeled oligonucleotide complementary to the 3' end of the single-stranded product.

If a single predominant product is not found, see Critical Parameters and Troubleshooting and consider performing one of the Alternate Protocols.

3. After the final long extension, and verification of single-stranded DNA product, bring the reaction slowly to room temperature and remove the mineral oil with a Pasteur pipet.
4. Precipitate by adding 10 M ammonium acetate to a final concentration of 2.5 M, followed by 1 vol of room temperature 100% ethanol. Leave the sample 5 min at room temperature. Microcentrifuge 5 min at room temperature, maximum speed, then wash pellet with room-temperature 70% ethanol and dry under vacuum. Resuspend pellet in 50 μ l water or 50 μ l of 0.1 \times TE buffer, pH 8.0.

This precipitation should result in the clean, amplified product, free of amplification buffer components and of most of the unincorporated PCR primers. Alternatively, the products can be purified by spin-dialyzing through a Centricon 30 or 100 column (according to manufacturer's instructions), depending on the size of the amplified fragment. If products are to be sequenced using Taq DNA polymerase (UNIT 7.4), this purification step may be omitted, because PCR and Taq sequencing buffers are compatible.

5. Carry out dideoxy sequencing (UNIT 7.4) using the asymmetric PCR-generated single-stranded templates.

The sequencing primer can be either the PCR-limiting oligonucleotide primer or any complementary sequence internal to the 3' end of the single-stranded template.

ALTERNATE PROTOCOL 1

GENERATING SINGLE-STRANDED TEMPLATE FOR DIDEOXY SEQUENCING BY SINGLE-PRIMER REAMPLIFICATION

The use of asymmetric primer ratios does not always result in reproducible high yields of single-stranded product. An alternate method entails isolation of double-stranded PCR products, followed by reamplification—this time in the presence of a single primer. This method of asymmetric reamplification supplies sufficient single-stranded template for a full set of dideoxy sequencing reactions.

For materials, see Basic Protocol 1.

1. Carry out 25 to 40 cycles of PCR under optimized conditions (*UNIT 15.1*), using equimolar amounts (20 to 50 pmol) of each primer, and ending with a long extension step.
2. Remove the mineral oil with a Pasteur pipet.
3. Precipitate, wash, and dry amplified products (see Basic Protocol 1, step 4). Resuspend the pellet in 30 μ l of 0.1 \times TE buffer, pH 8.0.

Because the double-stranded PCR product is to be reamplified in the presence of only a single primer, it is important to remove all excess primer remaining in this first amplification. The precipitation conditions described ensure the recovery of the double-stranded amplified product, but leave the unincorporated primers in solution.

4. Use 1 μ l of the resuspended double-stranded product as template for a new *asymmetric* PCR amplification. Add 20 to 40 pmol of *one* of the two primers, and carry out 20 cycles of PCR under the same conditions as in step 1.

Determine the yield and quality of the single-stranded product before sequencing it. Add 0.5 to 1 pmol of 32 P-labeled PCR primer (in addition to the cold primer) to the asymmetric reamplification (step 4). When PCR is completed, electrophore 5% to 10% of the product on a non-denaturing agarose or polyacrylamide gel, dry the gel, and expose to X-ray film.

Alternatively, the labeled primer can be omitted, and the unlabeled products from the asymmetric PCR can be electrophoresed on a non-denaturing agarose or polyacrylamide gel. Southern or electroblotted without denaturing the gel or filter (thus ensuring that only single-stranded products are targeted by the probe), and the filter probed with a labeled oligonucleotide complementary to the 3' end of the single-stranded product.

In both cases, a single band corresponding to the single-stranded product should be visible (a band corresponding to the double-stranded product may occasionally appear, but should be much fainter than the single-strand band). Although the single-stranded product increases only arithmetically with each cycle, 15 to 20 asymmetric PCR cycles generally produce sufficient single-stranded product for 6 to 12 sequencing reactions. If a single predominant product is not found, see Critical Parameters and Troubleshooting and consider performing one of the Alternate Protocols.

5. Precipitate amplified products (see Basic Protocol 1, step 4). Dry pellet under vacuum.

*This step may not be necessary if Taq DNA polymerase is used to sequence (*UNIT 7.4*).*

6. Carry out dideoxy sequencing using the PCR-generated single-stranded template (*UNIT 7.4*).

*The sequencing primer can be either the PCR primer (the primer **not** used in the asymmetric reamplification) or any complementary sequence internal to the 3' end of the single-stranded template.*

PREPARING DOUBLE-STRANDED PCR PRODUCTS FOR DIDEOXY SEQUENCING

Although double-stranded, closed circular DNA templates can be alkaline-denatured and sequenced using dideoxy methods (*UNITS 7.3 & 7.4*), this method of template preparation gives poor results with linear PCR products. A modified method suitable for dideoxy sequencing of PCR-generated double-stranded DNA is presented in this protocol. Amplified DNA is purified, freed of oligonucleotide primers, then heat-denatured and snap-cooled. Primers used for the PCR amplification or other internal primers can then be used in the sequencing reaction.

ALTERNATE PROTOCOL 2

The Polymerase Chain Reaction

15.2.3

Additional Materials (also see *Basic Protocol 1*)

PEG/NaCl solution (see recipe)
70% ethanol, cold
1 pmol (2 to 5 ng) sequencing primer
5× annealing buffer (see recipe)

Additional reagents and equipment for phenol/chloroform extractions (*UNIT 2.1A*)

1. Carry out PCR under optimized conditions (*UNIT 15.1*), using 50 pmol of each oligonucleotide primer, ending with a long extension step.

Use enough template DNA and perform a sufficient number of PCR cycles to yield enough products for several sequencing reactions. For mammalian genomic DNA, 1 μg and 30 cycles should be sufficient. Yield should be checked on a gel (step 2) and PCR conditions altered as necessary.

2. Analyze a 5-μl aliquot of the PCR product by electrophoresis in agarose, nondenaturing polyacrylamide, or sieving agarose gels appropriate for the PCR product size expected, and stain with ethidium bromide.

*Identification and determination of purity of the PCR product is important for deciding which method to use for the purification. If the desired PCR product is the predominant species, proceed with steps 3 and 4. If more than one PCR product is observed, the desired product can be purified on a low gelling/melting temperature agarose gel (*UNIT 2.6*) before denaturation (step 5).*

3. Remove the mineral oil from the PCR sample with a Pasteur pipet. Extract DNA with phenol/chloroform and then chloroform (*UNIT 2.1A*). Transfer the aqueous phase to a fresh tube.
4. Add 0.6 vol PEG/NaCl solution to DNA and incubate 10 min at 37°C. Microcentrifuge 10 min at top speed to collect the precipitate. Wash pellet with cold 70% ethanol and dry under vacuum. Resuspend at a concentration of ≥ 0.1 pmol/μl (60 ng/μl of 1-kbp fragment) in 0.1× TE buffer, pH 8.0.

This precipitation should remove the excess oligonucleotide primers and free nucleotides.

5. Pipet ~0.1 pmol of purified PCR product (i.e., ~1 μl of ~60 ng/μl of 1-kbp fragment) into a microcentrifuge tube. Adjust volume to 9 μl with 0.1× TE buffer, pH 8.0. Denature template by heating 2 min at 95°C, then quickly cool 1 min in a dry ice/ethanol bath. Proceed immediately to step 6.

Only ~0.1 pmol of template is required for a clear autoradiographic signal from a sequencing gel after an overnight exposure.

6. Quickly add 1 μl sequencing primer and 2 μl of 5× annealing buffer to the frozen DNA sample. Microcentrifuge tube briefly at room temperature to thaw and mix the reaction.

A 5-fold molar excess of primer over template is recommended. However, each primer should be tested for its optimal concentration.

7. Incubate template and primer 30 min at room temperature.
8. Carry out dideoxy sequencing reactions (*UNIT 7.4A & B*).

GENERATING SINGLE-STRANDED TEMPLATE FOR DIDEOXY SEQUENCING BY λ EXONUCLEASE DIGESTION OF DOUBLE-STRANDED PCR PRODUCTS

ALTERNATE
PROTOCOL 3

An alternative approach for generating single-stranded products does not require the use of unequal primer concentrations (Higuchi and Ochman, 1989). This method involves the selective digestion of one of the two strands of the amplified product by the action of λ exonuclease. This enzyme is a double-strand-specific 5'→3' exonuclease, but is only active if a phosphate is present at the 5' position (UNIT 3.11). If only one of two PCR primers is phosphorylated, only the strand flanked by that primer will be degraded.

Additional Materials (also see Basic Protocol 1)

- T4 polynucleotide kinase (UNIT 3.10) and 1× buffer (UNIT 3.4)
- 3 M sodium acetate, pH 5.2 (APPENDIX 2)
- λ exonuclease (UNIT 3.11) and 1× buffer (UNIT 3.4)

1. Phosphorylate 10 to 50 pmol of either PCR oligonucleotide primer 1 or 2 (*but not both*) with T4 polynucleotide kinase, using unlabeled ATP (UNIT 3.10).
2. Ethanol precipitate the phosphorylated primer following addition of 3 M sodium acetate to 0.3 M final concentration, using 2.5 vol cold 100% ethanol. Place 1 hr at -70°C, microcentrifuge 20 min at 4°C, and wash pellet once (carefully) with 70% ethanol. Dry pellet briefly under vacuum.

If the phosphorylation was carried out in a volume of ≤ 10 μ l, the entire reaction can be incubated 5 min at 65°C to inactivate the polynucleotide kinase, then added directly to the 100- μ l PCR mix in step 3.

3. Carry out PCR under optimized conditions (UNIT 15.1) with equimolar amounts (10 to 100 pmol) of both the phosphorylated and nonphosphorylated PCR oligonucleotide primers.
4. Phenol extract the PCR products (UNIT 2.1), then ethanol precipitate, wash, and dry (see Basic Protocol 1, step 4).
5. Resuspend the pellet in 50 to 100 μ l of 1× λ exonuclease buffer. Add 5 U λ exonuclease and incubate 15 to 30 min at 37°C.

The yield and quality of the template may be verified before sequencing, as in step 4 of the second alternate protocol.

6. Phenol extract twice, then ethanol precipitate, wash, dry, and resuspend the single-stranded product (see Basic Protocol 1, step 4).
7. Carry out dideoxy sequencing (UNIT 7.4) with the PCR- and λ exonuclease-generated single-stranded templates.

ONE-STEP ENZYMATIC PURIFICATION OF PCR PRODUCTS FOR DIRECT SEQUENCING

ALTERNATE
PROTOCOL 4

This protocol offers convenient enzymatic purification of PCR products for direct sequencing without the need for further subcloning, precipitation, or column purification. All reactions may be done in 96-well PCR plates and are compatible with large-scale sequencing. This method is designed to be convenient and rapid. The enzymatic purification and subsequent sequencing reaction may be done in the same PCR tubes or wells, eliminating the decreased yield associated with multiple sample transfers. The principal elements of a PCR reaction that interfere with conventional sequencing reactions are primers and dNTPs. Exonuclease I (Exo I), which preferentially hydrolyzes single-

The Polymerase
Chain Reaction

15.2.5

stranded DNA, is used to degrade the primers. Shrimp alkaline phosphatase (ShrAP) is used to dephosphorylate the remaining deoxynucleotides. A mixture of the above enzymes is added in one pipetting step to a small aliquot of PCR product. The samples are then placed in a thermal cycler for incubation and subsequent heat inactivation. The resulting product is then ready for direct sequencing. Commercial chromatography columns can be used as an alternative to the following protocol. However, purification columns may decrease yield and are significantly more expensive, especially when sequencing is performed on a large scale.

The following purification approach is suitable for a variety of PCR templates, including plasmid, lambda, and genomic DNA. This protocol is even compatible with PCR products from direct bacterial and yeast colony PCR.

Materials

PCR product to be sequenced (*UNIT 15.1*), 25 to 250 ng/μl
1 U/μl shrimp alkaline phosphatase (ShrAP; e.g., USB)
10 U/μl exonuclease I (Exo I; e.g., USB)
50 mM Tris-Cl, pH 8.0 (*APPENDIX 2*)

PCR tubes or 96-well PCR plate
Thermal cycler to accommodate tubes or 96-well plates

Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5A*)

1. Run a small sample of a PCR product on an agarose gel (*UNIT 2.5*). Use an appropriate DNA molecular size marker to estimate DNA concentration.

The presence of a single clear PCR product should be confirmed.

The DNA concentration should be in the range of 25 to 250 ng/μl. One of the principal requirements for high-quality sequences is the careful titration of DNA template used in the cycle sequencing reaction. The most common mistake is to use an excess of template, which causes early termination of the sequencing reactions. Keep in mind that there will be no loss of PCR template during this purification.

2. Mix together the enzyme cocktail as follows:

0.1 μl 1 U/μl ShrAP
0.1 μl 10 U/μl Exo I
1.8 μl 50 mM Tris-Cl, pH 8.0.

This mixture should be scaled up according to the number of PCR products being purified.

3. Add a 2-μl aliquot of enzyme cocktail into each PCR tube or well of a 96-well PCR plate.

When executing large-scale reactions using 96-well PCR plates, it is often helpful to use a multichannel pipettor that can dispense small volumes (1 to 10 μl).

4. Add 2 μl PCR product to the enzyme cocktail and mix by gently pipetting a few times.

Remember that there should be no loss of DNA template during this purification process. To minimize intervening pipetting steps, it is best to titrate the amount of PCR product used prior to enzymatic treatment. This would allow subsequent sequencing reactions to be done in the same PCR tubes or wells. Most sequencing reactions will require ~50 to 500 ng DNA template. Optimization of the amount of DNA template will depend upon the specific sequencing protocol and facility. Often, the direct pipetting of 2 μl PCR product will give satisfactory results. However, some contemporary automated sequencing facilities will require the PCR product to be diluted prior to treatment with the enzyme cocktail, as undiluted PCR product sometimes has too much DNA template for the sequencing reaction.

To optimize DNA concentration, a good starting point is the undiluted PCR product, with subsequent dilutions made using 50 mM Tris·Cl, pH 8.0.

There is usually enough magnesium in a typical PCR reaction to support enzymatic treatment by *ShrAP* and *Exo I*, even with a 3-fold dilution. If PCR products are diluted ≥ 4 -fold, the authors suggest using 50 mM Tris·Cl (pH 8.0)/100 mM MgCl₂ for dilutions. This buffer should provide the divalent cations that might otherwise be lost with higher dilutions of the PCR product.

5. Briefly centrifuge the product to bring down any residual fluid along the sides of the tube or well.

This step is optional, but relatively small amounts of residual primer or dNTPs are capable of interfering with the sequencing reactions. Therefore, it is important that all of the PCR product be in contact with the enzyme cocktail.

6. Place the PCR tube or plate into a thermal cycler and run using the following conditions:

15 min	37°C	(enzyme incubation)
15 min	80°C	(heat inactivation of enzymes)
Indefinitely	4°C	(hold).

The samples may be stored for several days at 4°C or frozen for several weeks at -20°C until sequencing.

7. Carry out dideoxy sequencing (UNITS 7.4A & 7.4B).

Cycle sequencing reactions may be performed in the very same PCR tubes or wells using 4 pmol sequencing primer. (The amount of sequencing primer may vary with the sequencing PCR conditions optimized for a given facility.) One of the primers used in the original PCR reaction may be used successfully as the sequencing primer. However, a nested sequencing primer may give improved results.

The purified products from this protocol are suitable for use with typical dye-terminator sequencing (e.g., UNIT 7.4B). The resulting sequencing extension products may be analyzed on a typical sequencing gel (e.g., ABI 377 DNA Sequencer; Applied Biosystems) or a polymer-based multicapillary sequencer (e.g., UNIT 7.6; ABI 3700 DNA Analyzer; Applied Biosystems).

LABELING PCR PRODUCTS FOR CHEMICAL SEQUENCING

One of the two PCR oligonucleotide primers is radioactively 5' end-labeled with polynucleotide kinase and then used in PCR. This results in the production of a double-stranded amplified product containing one labeled and one unlabeled strand. The double-stranded product is then used for chemical sequencing.

Materials

5 to 10 pmol oligonucleotide primer 1 (to be end-labeled)
T4 polynucleotide kinase (UNIT 3.10) and 1× buffer (UNIT 3.4)
[γ -³²P]ATP, specific activity 6000 Ci/mmol
20 to 40 pmol oligonucleotide primer 2
10× PCR amplification buffer (UNIT 15.1)
2 mM 4dNTP mix (UNIT 15.1)
DNA template (20 to 1000 ng for eukaryotic DNA; 10 to 100 ng for bacterial DNA; 1 to 20 ng for cloned DNA inserts)
Taq DNA polymerase (UNIT 15.1)
Mineral oil

BASIC PROTOCOL 2

The Polymerase Chain Reaction

15.2.7

Disposable columns or dialysis cartridges for DNA purification (optional; NACS prepack cartridge, Life Technologies; Centricon 30 column, Amicon; or Select 6L column, 5 Prime→3 Prime)

Additional reagents and equipment for PCR (*UNIT 15.1*), end-labeling with T4 polynucleotide kinase (Table 3.4.1 & *UNIT 3.10*), electrophoresis using agarose gels (*UNIT 2.5A*) or nondenaturing polyacrylamide gels (*UNIT 2.7*), autoradiography (*APPENDIX 3A*), electroelution from agarose gels (*UNIT 2.6*), phenol extraction and ethanol precipitation (*UNIT 2.1A*), sequencing by the chemical method (*UNIT 7.5*), and sequencing gels (*7.6*)

1. In a volume of 10 μ l, radioactively end-label 5 to 10 pmol oligonucleotide primer 1 for 30 min using T4 polynucleotide kinase. Heat-inactivate kinase by incubating 10 min at 65°C.

[γ -³²P]ATP with a specific activity of 6000 Ci/mmol is recommended. Use lower specific-activity label if longer autoradiograph exposures (>48 hr) are acceptable.

2. Prepare a microcentrifuge tube for PCR containing the following:

10 μ l 10 \times amplification buffer (optimized as in *UNIT 15.1*)
10 μ l 2 mM 4dNTP mix
20 to 40 pmol oligonucleotide primer 2 (not labeled in step 1)
DNA template
1.5 to 2.5 U *Taq* DNA polymerase
10 μ l kinase reaction (step 1)
H₂O to 100 μ l.

Overlay with 100 μ l mineral oil. Carry out PCR under optimized conditions for 5 to 10 cycles (see *UNIT 15.1*, steps 6 to 8). Complete PCR with a final long extension step of 5 to 7 min.

The final extension ensures that all PCR products are complete and blunt-ended. If the DNA template is a single band from a previous PCR reaction that has been cut out and extracted from a gel, >10 cycles of PCR may be necessary. Similarly, if only a small amount of initial template is present, 20 to 25 cycles of PCR may be necessary to generate sufficient labeled product.

3. Remove the mineral oil with a Pasteur pipet.
4. Separate the labeled PCR product from the unincorporated labeled primer and the unincorporated labeled nucleotides by electrophoresis on an agarose or nondenaturing polyacrylamide gel. Using the characteristic mobilities of the loading dyes (bromphenol blue and/or xylene cyanol) as guideposts (*UNITS 2.5A & 2.7*), run the gel until the unincorporated primers and nucleotides migrate off the gel and into the lower buffer chamber. Visualize the labeled PCR products by autoradiography for 10 to 60 min.

Be conservative in running the gel. Avoid running the PCR products off the gel.

5. Cut and isolate the desired labeled PCR product from the gel, using one of the described protocols for fragment purification (*UNITS 2.5A, 2.6, or 2.7*).

*A preparative gel confirms the presence of the desired amplified product. However, amplified DNA can be also be purified using disposable columns, which yield clean double-stranded DNA. Chemical sequencing can also be carried out without gel or column purification. In such cases, the PCR sample should be phenol extracted and ethanol precipitated in the presence of 0.3 M sodium acetate and 2.5 vol cold 100% ethanol (*UNIT 2.1A*). The resulting pellet should be washed with 70% ethanol.*

The columns and extraction procedures described above should not be used instead of a preparative gel unless it is known that only a single, homogeneous product has been amplified during PCR.

6. Resuspend the labeled PCR product in 30 to 50 μ l water.
7. Carry out chemical sequencing using 8% to 10% of the product for each sequencing reaction (UNIT 7.5).
8. Resolve sequencing products on a sequencing gel (UNIT 7.6). Wrap gel in plastic wrap and autoradiograph at -20°C for 2 to 48 hr. Read sequence as described in UNIT 7.5.

The gel may be dried, but it is generally not necessary to do so.

GENOMIC SEQUENCING OF PCR PRODUCTS

This method combines the PCR amplification method with the genomic sequencing technique (Church and Gilbert, 1984). Following PCR, the amplified DNA is chemically sequenced, transferred by electroblotting, and covalently bound by UV crosslinking onto a nylon filter that can be repeatedly probed with short, sequence-specific oligonucleotides. This method is ideally suited to situations where large amounts of sequence are sought, or where more than one region is being amplified. Several amplified fragments can be mixed, simultaneously sequenced, run out on a single set of lanes, and the sequence of the different fragments successively visualized by the use of appropriate probes.

Materials

Filter paper (Schleicher & Schuell #410), precut to gel size

Additional reagents and equipment for phenol extraction and ethanol precipitation (UNIT 2.1A); transfer by electroblotting, UV cross-linking, and hybridization (UNITS 2.9, 6.3, & 6.4); and labeling with terminal deoxynucleotidyltransferase (UNIT 3.6)

1. Carry out PCR under optimized conditions, ending with a final long extension step (UNIT 15.1).

The template concentrations that yield optimum amplifications must be empirically determined. These concentrations range from 20 to 1000 ng for eukaryotic genomic DNA, 10 to 100 ng for bacterial DNA, and 1 to 20 ng for cloned DNA inserts.

2. After the final long extension step, remove the mineral oil with a Pasteur pipet.
3. Phenol extract the amplified product (UNIT 2.1A). Precipitate with 0.1 vol of 3 M sodium acetate, pH 5.2, and 2.5 vol of 100% ethanol for 20 min in a dry ice/ethanol bath. Microcentrifuge 5 min at top speed, 4°C . Wash the pellet twice with 70% ethanol and dry under vacuum (UNIT 2.1A).
4. Resuspend the pellet in 30 to 50 μ l water. Electrophorese 5% to 10% of the PCR sample on an agarose or nondenaturing polyacrylamide gel to determine the yield and purity of the amplified product.

Because every double-stranded product generated by PCR will give rise to a sequence ladder when chemically sequenced, the homogeneity of the amplified product is critical. If amplification yields at least 2 bands of similar intensity, a gel separation prior to sequencing is necessary, unless the different products can be separately visualized by using different oligonucleotide probes. The use of internal probes (3' of the PCR primers) to visualize the sequence will prevent spurious amplification product sequences from appearing, but will not allow sequencing of the entire amplified fragment. The presence of short products ("primer-dimers") will generally not interfere with the majority of the sequence, but will obscure the bottom section of the sequencing gel.

**ALTERNATE
PROTOCOL 5**

**The Polymerase
Chain Reaction**

15.2.9

5. Carry out chemical sequencing using 8% to 10% of the amplified product for each sequencing reaction (UNIT 7.5).
6. Electrophorese the sequenced samples on a denaturing polyacrylamide sequencing gel (UNIT 7.6).

The number of resolvable base pairs can be increased by using a gradient or wedge gel, performing multiple loadings of the sequenced sample, and/or using longer gel plates (≤ 1 meter; UNIT 7.6).

7. Carefully remove one of the glass gel plates. Place the sheet of filter paper carefully on the gel, making sure not to trap air bubbles between the filter paper and gel. Lift the gel (adhered to the filter paper) onto the appropriate transfer apparatus.
8. Wet the gel with a thin layer of TBE electrophoresis buffer and place the nylon filter over the gel, again being careful not to trap air bubbles between the gel and membrane. Transfer the DNA onto the filter membrane by electroblotting (UNIT 2.9).

It is necessary to use a large electroblotting apparatus for this purpose.

9. UV cross-link the DNA onto the filter (UNIT 2.9).
10. Hybridize the filter with the appropriate probe.

The filter can now be probed, visualized, stripped, and reprobed as many as 40 times.

Best results are obtained by probing with the PCR oligonucleotide primers, ensuring that only completed PCR products are visualized. Degenerate probes or probes internal to the amplified fragment can also be used. A high-specific-activity probe can be obtained by "tailing" 3 to 6 pmol of probe using terminal deoxynucleotidyltransferase.

REAGENTS AND SOLUTIONS

Annealing buffer, 5 \times

200 mM Tris·Cl, pH 7.5
100 mM MgCl₂
250 mM NaCl
Store at -20°C

PEG/NaCl solution

20%(w/v) polyethylene glycol (PEG) 6000
2.5 M NaCl
Store at room temperature

COMMENTARY

Background Information

The methods presented in this unit permit rapid sequencing of PCR-amplified DNA. Although the cloning of amplified DNA is relatively straightforward, direct DNA sequencing of PCR products facilitates and speeds the acquisition of sequence information. As long as the PCR reaction produces a single amplified product (or a set of discrete products that can be conveniently separated), such products are amenable to direct sequencing.

In contrast to methods where the PCR product is cloned and a single clone sequenced, the

direct sequencing of "bulk" PCR product is generally unaffected by the comparatively high error rate of *Taq* DNA polymerase. Unless a certain feature of the template causes consistent misincorporation at the same point in the sequence, errors are likely to be stochastically distributed throughout the molecule—thus, the overwhelming majority of the amplified product consists of the correct sequence. The only exceptions to this rule may be those cases where only a very small number of template molecules (<5) are present at the outset of the reaction. In such cases, an error occurring in the first cycle

of amplification may be geometrically amplified.

Both the dideoxy and the chemical sequencing approaches can be used to characterize PCR products directly. The general advantages and shortcomings of the two sequencing approaches are discussed in the introduction to Chapter 7, but those specific to PCR products are emphasized here.

The dideoxy method involves fewer steps than chemical sequencing. Where *Taq* DNA polymerase is used for sequencing reactions, the same buffer will serve for the amplification and sequencing steps, and products do not have to be cleaned and isolated repeatedly.

The technology and procedures used in dideoxy sequencing are by now well standardized. A number of commercially available dideoxy “kits” yield good results. Provided that good yields of high-quality single-stranded template can be obtained, the dideoxy methods described here and elsewhere (UNIT 7.4A & B) permit rapid sequencing of amplified products. The limitations on the maximum amount of sequence obtainable by dideoxy sequencing can frequently be overcome by the use of an internal sequencing primer to generate additional sequence ladder.

The genomic sequencing method (for chemically sequenced PCR products) offers the advantage of a permanent matrix onto which the amplification product(s) can be bound. This method has proven particularly useful for sequencing large regions (>1000 bp). In such cases, the sequences of interest can be amplified simultaneously in a single PCR reaction, or separately, as a set of discrete adjacent or overlapping fragments. The fragments can then be mixed, simultaneously sequenced, electrophoresed on gels, and transferred onto nylon filters. Each fragment can be successively visualized by probing with the appropriate labeled oligonucleotide. In addition, sequences from both strands can be derived from a single filter, with a consequent increase in sequence accuracy. This “multiplex” approach, however, requires the additional blotting and hybridization steps. The direct chemical sequencing of a labeled strand is a fast alternative if only a single product (and a single strand) is being sequenced. The utility of this direct approach is thus limited to those cases when the PCR yields a single DNA species or when the product of interest can be readily gel purified.

Several excellent papers dealing with direct DNA sequencing of PCR products have been published. Detailed discussions of dideoxy se-

quencing approaches can be found in Wrischnik et al. (1987), Gyllensten and Erlich (1988), Innis et al. (1988), Kreitman and Landweber (1989), and Gyllensten (1989). Chemical sequencing of PCR products is described in DiMarzo et al. (1988) and Ohara et al. (1989).

Purification methods commonly used to purify PCR products involve extractions and precipitations, the use of commercial filtration or chromatography columns, or enzymatic clean-up. The protocols used for column purification benefit from their simplicity. However, enzymatic purification (Alternate Protocol 4) is as fast as column purification. This method, developed by Hanke and Wink (1994) and Werle et al. (1994), requires relatively little hands-on time and results in a negligible loss of template, as all steps are performed in the very same PCR tube or well. In addition, the principal advantage of enzymatic purification is the relative ease with which it can be scaled up to high-throughput capacity at a mere fraction of the cost.

Critical Parameters and Troubleshooting

The majority of problems that may hinder direct DNA sequencing of PCR products are described in the units on dideoxy and chemical sequencing (UNITS 7.4 & 7.5) and in the unit pertaining to optimization of PCR amplification (UNIT 15.1).

Regardless of the source of DNA or the sequencing methods employed, the quality of a DNA sequence depends critically on the quality of the starting material. In the case of PCR products, it is imperative (before sequencing) to optimize the reaction conditions (UNIT 15.1) in order to obtain a homogeneous amplification product. For this reason, a final long extension step must be included in any PCR reaction. In certain cases, no amount of optimization will produce a single amplified product, and individual products must be purified by gel electrophoresis. In general, gel purification is sufficient to generate a homogeneous product. On occasion, even a single isolated band contains more than one amplified product (as revealed by the sequence), and more sophisticated methods, such as denaturing gel gradient electrophoresis (Myers et al., 1988), or reamplification with a set of nested internal primers (Ohara et al., 1989), must be used.

Because PCR is normally carried out under conditions strongly favoring synthesis (high concentrations of enzyme, primers, and nucleotides), artifact sequences are easily generated.

This is a particular risk when amplifying in the presence of only one primer. In such cases, DNA synthesis is not bounded by a second primer, and a heterogeneous collection of molecules, including “hybrid” or “artificial” molecules, can easily arise. In addition, snap-back synthesis can take place, again resulting in artificial amplified molecules. Whenever possible, the identity of amplified fragments should be confirmed prior to sequencing by probing of Southern blots, preferably with specific sequences internal to the amplified region (UNIT 2.9), or by restriction endonuclease digestion analyses of amplified products to confirm the expected restriction sites (UNIT 3.1).

Conditions for obtaining single-stranded template must be optimized empirically before undertaking dideoxy sequencing. Alternatively, the alternate protocol described here for sequencing double-stranded PCR products can be employed. Many difficulties with the dideoxy method are caused by the secondary structure of the template strand. The inclusion of nucleotide analogs, coupled with the use of a thermostable polymerase (e.g., *Taq* DNA polymerase) for chain extension, can reduce or resolve many of these problems.

In sequencing double-stranded PCR products, the crucial point is to prevent reannealing of the template strands that compete with binding of the sequencing primer. This can be overcome by two critical manipulations—heat denaturation of the template and quick addition of the primer to the template.

For the one-step procedure (Alternate Protocol 4), titration of the amount of DNA template used for the cycle sequencing reaction is critical to sequence quality. Most polymerase-based capillary sequencers require ~50 to 500 ng DNA per sequencing reaction. Insufficient DNA will result in low-amplitude peaks on the sequencing chromatograms. Excessive DNA, which is more often the case, will result in the rapid depletion of the fluorophore-dNTPs in the cycle sequencing reaction. The sequencing chromatogram will appear to have extremely high-amplitude peaks that burn out relatively early, resulting in a very short region of usable sequence data. Often, the direct pipetting of undiluted PCR product will result in satisfactory sequences. However, optimal sequence lengths may require dilution of the PCR product prior to enzymatic purification.

The shrimp alkaline phosphatase (ShrAP) should be screened for contaminating endonuclease or exonuclease activity. The ShrAP from USB has undetectable or negligible levels of

the above contaminating activities. Exonuclease I (Exo I) preferentially hydrolyzes single-stranded DNA. However, Exo I does have some detectable double-stranded DNA exonuclease activity. Therefore, prolonged incubation with this enzyme may result in minor hydrolysis at the ends of the double-stranded DNA template. For this reason, using a nested primer for sequencing may give improved sequence data over the use of one of the original PCR primers. However, using one of the original PCR primers as the sequencing primer often gives satisfactory results. Further, the enzyme incubation time should be kept to a minimum (i.e., 15 min) prior to heat inactivation.

The amount of hydrolyzing enzyme used in this protocol has been titrated as low as possible without sacrificing sequencing quality. There are two reasons for using minimal enzyme quantities. First, large amounts of protein contaminants have been suspected to cause precipitation within the capillary arrays of ABI 3700 analyzers. The small amounts of enzyme used in this protocol have not caused any increased wear on capillary sequencers to date. Second, using less enzyme results in cost savings for high-throughput sequencing efforts.

Anticipated Results

These protocols capitalize on the efficiency and specificity of PCR and streamline the process of DNA sequencing. From 150 to 300 bp of sequence can routinely be obtained from overnight exposures of either dideoxy or chemical (or genomic) sequencing gels, provided oligonucleotides or strands have been labeled with ³²P. Longer exposures (24 to 72 hr) may be needed when using ³⁵S as a label. With appropriate modifications, including multiple loadings, repeated probings, or a series of sequencing primers, the yield of legible sequence can be substantially increased.

Rapid purification of PCR products can be achieved within one PCR tube or well, which can directly proceed to cycle sequencing using conventional dye-terminator reactions. The sequence quality should be comparable to or better than that achieved through column purification or ethanol precipitation. Further, this protocol can be easily scaled for high-throughput use.

Time Considerations

Each protocol described here can be completed in a single day or less, with additional time required for sequencing reactions and analysis.

Literature Cited

- Church, G.M. and Gilbert, W. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 81:1991-1995.
- Church, G.M. and Kieffer-Higgins, S. 1988. Multiplex DNA sequencing. *Science* 240:185-188.
- DiMarzo, R., Dowling, C.E., Wong, C., Maggio, A., and Kazazian, H.H. 1988. The spectrum of β -thalassaemia mutations in Sicily. *Br. J. Haematol.* 69:393-397.
- Gyllensten, U.B. 1989. Direct sequencing of in vitro amplified DNA. In *PCR Technology* (H.A. Erlich, ed.) pp. 45-60. Stockton Press, New York.
- Gyllensten, U.B. and Erlich, H.A. 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to the direct sequencing of the HLA-DQA locus. *Proc. Natl. Acad. Sci. U.S.A.* 85:7652-7656.
- Hanke, M. and Wink, M. 1994. Direct DNA sequencing of PCR-amplified vector inserts following enzymatic degradation of primer and dNTPs. *Biotechniques* 17:858-860.
- Higuchi, R.G. and Ochman, H. 1989. Production of single-stranded DNA templates by exonuclease digestion following the polymerase chain reaction. *Nucl. Acids Res.* 17:5865.
- Innis, T.M.A., Myambo, K.B., Gelfand, D.H., and Brow, M.A. 1988. DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc. Natl. Acad. Sci. U.S.A.* 85:9436-9440.
- Kreitman, M. and Landweber L. 1989. A strategy for producing single-stranded DNA in the polymerase chain reaction. *Gene Anal. Technol.* 6:84-88.
- Myers, R.M., Sheffield, V., and Cox, D.R. 1988. Detection of single-base changes in DNA: Ribonuclease cleavage and denaturing gel electrophoresis. In *Genomic Analysis: A Practical Approach* (K. Davies, ed.) pp. 95-139. IRL Press, Oxford.
- Ohara, O., Dorit, R.L., and Gilbert, W. 1989. One-sided polymerase chain reaction: The amplification of cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 86:5673-5677.

Werle, E., Schneider, C., Renner, M., Volker, M., and Fiehn, W. 1994. Convenient single-step, one tube purification of PCR products for direct sequencing. *Nucl. Acids Res.* 22:4354-4355.

Wrischnik, L.A., Higuchi, R.G., Stoneking, M., Erlich, H.A., Arnheim, N., and Wilson, A.C. 1987. Length mutations in human mitochondrial DNA: Direct sequencing of enzymatically amplified DNA. *Nucl. Acids Res.* 15:529-542.

Key References

Kusukawa, N., Uemori, T., Asada, K., and Kato, I. 1990. Rapid and reliable protocol for direct sequencing of material amplified by polymerase chain reaction. *BioTechniques* 9:66-72.

Outlines the method for double-stranded sequencing of PCR products.

Maxam, A.M. and Gilbert, W. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499-559.

Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74:5463-5467.

These two papers outline the basic principles and techniques of chemical and dideoxy sequencing.

Contributed by Robert L. Dorit
Yale University
New Haven, Connecticut

Osamu Ohara
Shionogi Research Laboratories
Osaka, Japan

Charles B-C. Hwang
(double-stranded sequencing)
Harvard Medical School
Boston, Massachusetts

Jae Bum Kim and Seth Blackshaw
(one-step purification)
Harvard Medical School
Boston, Massachusetts

Ligation-Mediated PCR for Genomic Sequencing and Footprinting

The polymerase chain reaction (PCR) can be used to exponentially amplify segments of DNA located between two specified primer hybridization sites. This unit describes a single-sided PCR method that initially requires specification of only one primer hybridization site; the second is defined by the ligation-based addition of a unique DNA linker. This linker, together with the flanking gene-specific primer, allows exponential amplification of any fragment of DNA. Because a defined, discrete-length sequence is added to every fragment, complex populations of DNA such as sequence ladders can be amplified intact with retention of single-base resolution.

Although it is also suitable for other applications, the ligation-based protocol (see Basic Protocol) was specifically designed for genomic footprinting and direct sequencing reactions, and is described in this context. The starting material is genomic DNA that has been cleaved so that it retains a 5' phosphate at the cleavage site. In the schematic of the protocol (Fig. 15.3.1), only one cleavage product is shown for clarity; in practice, there would be a population of fragments resulting from the partial cleavage of DNA during footprinting or sequencing reactions. Preparation of DNA for footprinting analysis and genomic sequencing is described in the support protocols. Support Protocol 1 details *in vivo* and *in vitro* dimethyl sulfate (DMS) treatment and isolation of DNA from monolayer cells and its subsequent piperidine cleavage. Support Protocol 2 describes *in vivo* DMS treatment and harvesting of DNA from suspension cells. Support Protocol 3 describes how control DNA generated in the first two support protocols can be prepared for genomic sequencing.

NOTE: It is essential that fresh, high-quality reagents be used throughout this unit. In addition, all solutions should be prepared with glass-distilled water.

LIGATION-MEDIATED SINGLE-SIDED PCR

In this protocol, cleaved DNA is denatured and a gene-specific primer (primer 1) is annealed to the region of interest. In the first-strand synthesis, this primer is extended with a processive polymerase (Vent DNA polymerase) to the cleavage site to create a blunt end. DNA ligase catalyzes the attachment of a unidirectional (staggered) linker to this blunt end. The 3' end of the longer strand of the linker is ligated to the 5' end of the genomic DNA. The shorter strand of the linker lacks a 5' phosphate and therefore is not ligated to the extension product of the gene-specific primer. The DNA is denatured and a second gene-specific primer (primer 2) is annealed to the genomic DNA and extended by Vent DNA polymerase through the ligated linker region. (In theory, the first gene-specific primer could be used here, but in practice substituting a second primer greatly reduces background.) The extended product is now a suitable substrate for a PCR reaction; on one end (left in Fig. 15.3.1) there is a linker sequence to which a linker primer can anneal and on the other end (right in Fig. 15.3.1) there is a genomic sequence to which a gene-specific primer can anneal. Only molecules that have both sequences will be exponentially amplified during the subsequent PCR reaction; molecules with only one of the sequences will be linearly amplified. For the last extension, a third gene-specific primer (primer 3) that overlaps the second is used to label the DNA indirectly. These end-labeled extension products are visualized on a sequencing gel.

High-quality, reproducible footprinting or sequencing reactions can be obtained starting from 6×10^5 genomes (3×10^5 diploid nuclei). This corresponds to 2 μg of mouse (mammalian) genomic DNA. For other species, the recommended minimum number of

BASIC PROTOCOL

The Polymerase Chain Reaction

15.3.1

Contributed by Paul R. Mueller, Barbara Wold, and Paul A. Garrity

Current Protocols in Molecular Biology (2001) 15.3.1-15.3.26

Copyright © 2001 by John Wiley & Sons, Inc.

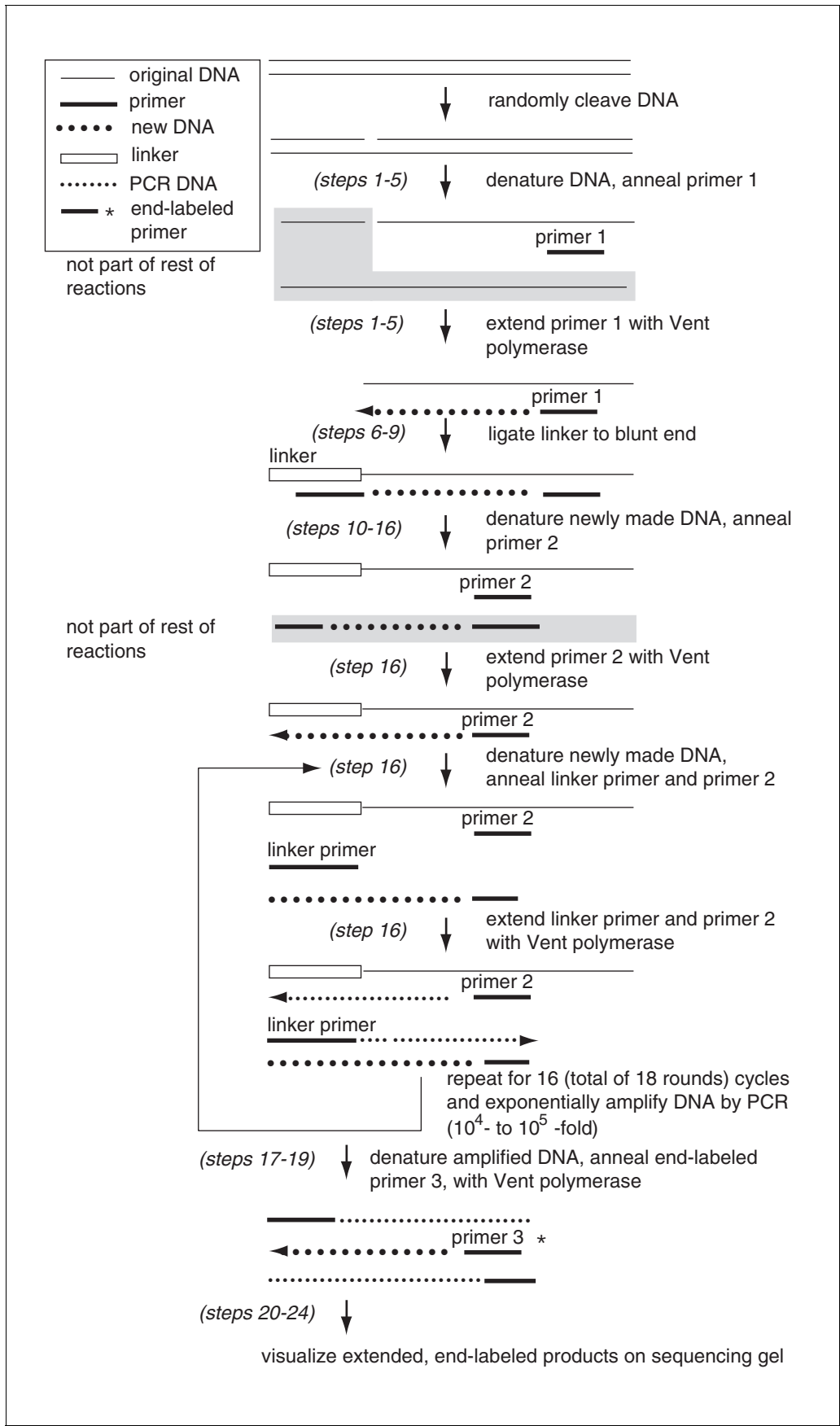


Figure 15.3.1 Flowchart of ligation-mediated PCR protocol (see text for details). The steps correspond to those listed in the Basic Protocol.

haploid genomes remains 6×10^5 , but different genome sizes lead to corresponding differences in the absolute mass of DNA.

Materials

- 0.4 $\mu\text{g}/\mu\text{l}$ cleaved genomic DNA in TE buffer, pH 7.5 (Support Protocol 1, 2, or 3)
- First-strand synthesis mix (see recipe), containing oligonucleotide primer 1 (Figs. 15.3.1 and 15.3.2; also see Critical Parameters)
- 20 μM unidirectional linker mix (see recipe)
- Ligase dilution solution (see recipe)
- Ligase mix (see recipe)
- 2000 to 3000 “Weiss” U/ml T4 DNA ligase (*UNIT 3.14*; Promega or Pharmacia Biotech)
- Precipitation salt mix (see recipe)
- 100% ethanol, ice-cold and room temperature
- 75% ethanol, room temperature
- Amplification mix (see recipe), containing linker primer and primer 2
- 2 U/ μl Vent DNA polymerase mix (see recipe; also see *UNIT 7.4*)
- Mineral oil
- End-labeling mix (see recipe), containing end-labeled primer 3 (Fig. 15.3.3)
- Vent DNA polymerase stop solution (see recipe)
- 25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol (see recipe)
- Loading buffer (see recipe)
- 1.5-ml microcentrifuge tubes, silanized (*APPENDIX 3*) and with Lid-Loks (optional; Intermountain Scientific)
- 4° and 17°C water baths
- Automated thermal cycler *or* water baths at 95°, 60°, 76°, and 60° to 70°C
- Additional reagents and equipment for PCR (*UNIT 15.1*), denaturing gel electrophoresis for DNA sequencing (*UNIT 7.6*), and autoradiography (*APPENDIX 3A*)

Carry out first-strand synthesis

1. Transfer 5 μl (2 μg) cleaved genomic DNA to a silanized 1.5-ml microcentrifuge tube and chill several minutes in an ice-water bath.
The cleaved DNA sample must be clean; contaminants left in the DNA (e.g., piperidine) will interfere with the reaction.
2. Prepare first-strand synthesis mix containing primer 1 and chill several minutes in an ice-water bath. Add 25 μl to DNA sample, gently mix with a pipettor, and return sample to the ice-water bath. Place Lid-Loks on sample to prevent tube from popping open during denaturation.

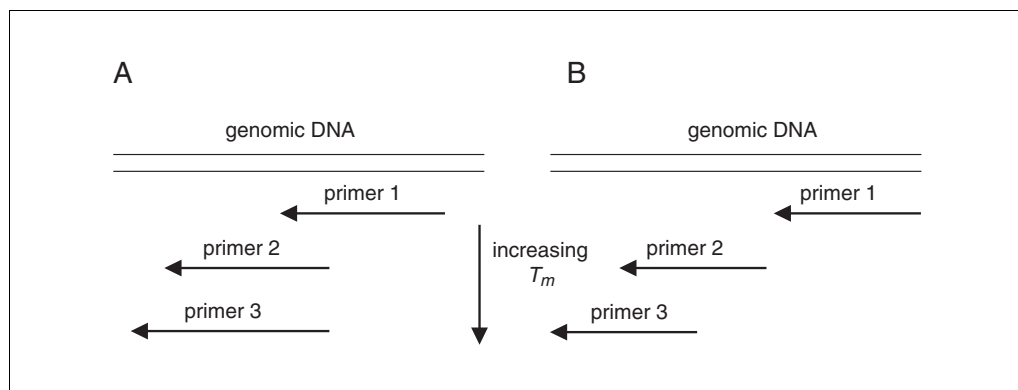


Figure 15.3.2 Two possible arrangements of gene-specific primers (see Critical Parameters).

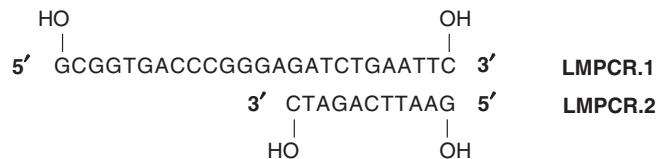


Figure 15.3.3 Staggered linker used in ligation-mediated PCR (LMPCR). This staggered linker is made by annealing the oligonucleotides LMPCR.1 and LMPCR.2 to each other. LMPCR.1 has 25 bases and a GC content of 60%, whereas LMPCR.2 has 11 bases and a GC content of 36%. Other oligonucleotides could be used in place of these as long as they meet the criteria defined in text. Note the lack of 5' phosphates on oligonucleotides (i.e., 5' OH instead) and the unconventional orientation of LMPCR.2 (3' to 5') in the figure.

When mixing with the pipettor, three or four careful strokes are usually sufficient. Try to avoid splashing the sample while mixing; if this occurs, microcentrifuge tubes briefly at 4°C to collect droplets.

- Denature DNA 5 min at 95°C, anneal primer 30 min at 60°C, and extend 10 min at 76°C. As first-strand synthesis proceeds, prepare solutions in step 4.

First-strand synthesis is most easily performed in an automated thermal cycler, but can be performed manually by transferring the tubes to different temperature water baths. It is important that the DNA be completely denatured so the primer can anneal. For this reason it is a good idea to check the calibration of the thermal cycler. The authors have found that some machines must be set to 96° or 97°C to actually obtain 95°C. Furthermore, if a thermal cycler is used, put the samples in the machine only after it is close to the denaturation temperature. This will minimize polymerase activity prior to the initial denaturation step. Most thermal cyclers have a “hold” function that allows sample addition followed by continuation of the program for the full denaturation time. If this reaction is performed manually using water baths, covering the tubes with a styrofoam block during denaturation and annealing minimizes condensation.

The extension step creates a blunt-end substrate for the subsequent ligation reaction (see Fig. 15.3.1); therefore, success at this step is absolutely crucial because it determines the molecules that will ultimately be amplified (see Critical Parameters).

- Thaw 20 μM unidirectional linker mix in ice-water bath. Prepare ligase dilution solution and partially prepare ligase mix (minus unidirectional linker and T4 DNA ligase; see reagents and solutions). Chill both in ice-water bath.

Do not add unidirectional linker mix and T4 DNA ligase to ligase mix until step 6.

- When extension in step 3 is complete, immediately transfer sample to the ice-water bath. Microcentrifuge tube briefly at 4°C to collect condensation, then return to ice-water bath.

The samples should be kept cold during steps 6, 7, and 8 to minimize Vent polymerase activity (see Critical Parameters).

Carry out ligation

- Finish preparing ice-cold ligase mix from step 4 by adding unidirectional linker mix, mixing, adding T4 DNA ligase, mixing, and keeping in ice-water bath.

This order of addition is important because the unidirectional linker is prepared in 250 mM Tris·Cl, pH 7.7, which is required by the ligase.

- Add 20 μl ice-cold ligase dilution solution from step 4 to sample, gently mix with a pipettor, and return sample to the ice-water bath. Add 25 μl ice-cold ligase mix from step 6 to sample, gently mix with a pipettor, and return to the ice-water bath.

Microcentrifuge tube briefly at 4°C and incubate overnight in a 17°C water bath (see critical parameters).

Addition of ligase dilution solution makes the first-strand synthesis mix into a buffer compatible with T4 DNA ligase (the NaCl concentration and, more importantly, the pH, are reduced). Addition of ligase mix provides a convenient way to add ligase and linker to each sample.

8. Prepare sample for precipitation by placing in the ice-water bath for several minutes. Microcentrifuge tube briefly at 4°C, then return to ice-water bath.

Placing samples in the ice-water bath is a convenient way of transferring them from the 17°C bath to the 4°C microcentrifuge without warming them.

9. Prepare and chill precipitation salt mix. Add 9.4 µl ice-cold precipitation salt mix and 220 µl of ice-cold 100% ethanol to the sample. Mix thoroughly by inversion and chill ≥2 hr at -20°C.

If desired, the experiment may be stopped at this point. Samples under ethanol are stable for weeks at -20°C.

Carry out PCR

10. Microcentrifuge precipitated ligation reaction 15 min at 4°C and discard supernatant.

11. Add 500 µl of room-temperature 75% ethanol and invert several times to wash pellet and walls of tube. Microcentrifuge sample ~5 min at room temperature and discard supernatant. Remove last traces of ethanol with a pipettor and allow any remaining ethanol to evaporate by air drying or by using a Speedvac evaporator.

The precipitated pellet will be rather large and will spread up the side of the tube.

12. Add 70 µl water and leave sample at room temperature to dissolve pellet. Vortex tube occasionally to assist dissolution, and after each vortexing, collect droplets by microcentrifuging 2 to 3 sec. When pellet is dissolved (usually ≤30 min), chill sample in ice-water bath. While pellet is dissolving, prepare and chill amplification mix containing linker primer and primer 2.

13. Add 30 µl ice-cold amplification mix to sample, gently mix with a pipettor, and return sample to the ice-water bath.

14. Add 3 µl (1 U) Vent DNA polymerase mix to sample and mix carefully with pipettor. Return sample to the ice-water bath.

The reaction is very sensitive to the amount of Vent DNA polymerase used; excess polymerase results in high background. Surprisingly, this effect has been observed when the amount of Vent polymerase is merely double the optimal level. The basis for the empirically observed threshold for this background is not known. Different lots of Vent polymerase may need to be titrated (see Critical Parameters).

15. Cover sample with 90 µl mineral oil, microcentrifuge briefly at 4°C, and return to the ice-water bath.

If the PCR reaction is performed manually using water baths, a Lid-Lock should be used to prevent the tube from popping open during denaturation.

16. Carry out 18 cycles of PCR. Perform first denaturation for 3 to 4 min at 95°C, and subsequent ones for 1 min. Anneal primers 2 min at a temperature 0° to 2°C above the calculated T_m s (if primer 2 and linker primer have different T_m s, use the lower T_m ; see critical parameters for T_m calculations). Extend 3 min at 76°C; for every cycle, add an extra 5 sec to the extension step. Allow final extension to proceed 10 min.

Transfer sample to ice-water bath, remove Lid-Lock (if applicable), and microcentrifuge briefly at 4°C to collect any condensation. Keep sample in ice-water bath.

In this procedure, 18 rounds of amplification yield highly reproducible sequence ladders that can be visualized (after the radiolabeling step) following an overnight exposure without an intensifying screen. Performing additional rounds of amplification could reduce the exposure time required, but this has not been tested exhaustively. A theoretical concern is that if too many amplification cycles are carried out, some reagents (e.g., polymerase, primers, or dNTPs) may become limiting, leading to undesirable lane-to-lane variability (18 rounds of amplification do not show variability).

An important parameter in this step is the temperature of denaturation. If it is too low, no signal or shortened sequence ladders will be seen; if it is too high, the polymerase will be destroyed. This may be a “hidden” problem because some automatic thermal cycling machines and thermometers do not accurately measure the denaturation temperature (see Critical Parameters).

Carry out end-labeling

17. Prepare end-labeling mix containing labeled primer 3 and chill several minutes in the ice-water bath. Add 5 µl to the sample. Mix aqueous phase by gently pipetting up and down, keeping the sample on ice as much as possible. Microcentrifuge briefly at 4°C and return to ice-water bath.

CAUTION: The label mix contains a significant amount of ³²P. Appropriate care should be taken in handling and disposal of mix and samples to which it is added.

18. Carry out two rounds of PCR to label the DNA. Perform first denaturation 3 to 4 min at 95°C; the second for 1 min. Anneal end-labeled primer 3 for 2 min at a temperature 0° to 2°C above its calculated T_m . Extend 10 min at 76°C. When second extension is complete, transfer sample to ice-water bath.

Although it is rare, in a few experimental situations, having two cycles of PCR during the end-labeling step has led to increased background and/or unusual primer artifacts such as primer-dimer products. It is best to try two cycles of end-labeling, cutting back to one cycle if this problem is encountered. It is also possible to label the amplification products using a filter-blotting procedure (Pfeifer et al., 1989).

If this reaction is performed manually using water baths, a Lid-Lock should be used to prevent the tube from popping open during denaturation.

19. Place sample at room temperature and immediately add 295 µl Vent DNA polymerase stop solution. Mix by vortexing. Microcentrifuge briefly to collect radioactive droplets from top and sides of tube. Add 500 µl phenol/chloroform/isoamyl alcohol and mix either by vigorous shaking or vortexing. Microcentrifuge 3 to 5 min at room temperature. Transfer upper aqueous layer (~400 µl; avoid interface if any) to a clean, silanized 1.5-ml microcentrifuge tube, and thoroughly mix. Microcentrifuge briefly to collect droplets from sides of tube.

EDTA in the Vent DNA polymerase stop solution chelates the magnesium required by Vent polymerase. The organic extraction removes proteins before precipitation and, more importantly, removes the mineral oil (there should be little or no interface). Transferring the aqueous layer to a new tube and mixing it ensures the reaction products will be equally aliquoted for multiple loadings on the sequencing gel (step 20).

20. Set up four clean, silanized 1.5-ml microcentrifuge tubes and add 235 µl of room-temperature 100% ethanol to each tube. Transfer 94 µl of the aqueous layer into each tube, thoroughly mix by vortexing, and chill ≥2 hr at -20°C. Discard any remaining aqueous layer.

This yields enough material for four samples of each reaction to be run on a sequencing gel. This accommodates multiple loadings to maximize the number of bases that can be read and provides extra samples for use in case a gel problem is encountered. If desired, this precipitation step can be a stopping point. Samples under ethanol are stable for several weeks at -20°C ; however, the specific activity of the product decreases with time.

21. Microcentrifuge precipitated samples 15 min at 4°C and discard supernatants.
22. Add 500 μl of room temperature 75% ethanol, vortex, microcentrifuge samples 5 min at room temperature, and discard supernatants. Remove last traces of ethanol by using a pipettor and allow any remaining ethanol to evaporate by air drying or by using a Speedvac evaporator.
23. Add 7 μl loading buffer to each tube and leave at room temperature while pellets are dissolving. Vortex occasionally to assist pellets in dissolving and to recover any of the DNA pellet that is on the side of the tube. After each vortexing, collect droplets by microcentrifuging 2 to 3 sec. Samples usually resuspend rapidly (within 5 min). Check resuspension by removing the sample from the tube with a pipet and using a Geiger counter to ensure that the radioactivity is in the sample and not left behind in the tube. Return sample to same tube. If $>10\%$ of the total radioactivity has remained in the tube, vortex, microcentrifuge, and repeat the resuspension check.

Do not let the samples sit longer than necessary in loading buffer, as this may result in smeared background and poor band resolution.

Run and analyze sequencing gel

24. Denature samples 5 min at 85° to 90°C . Load entire content of each tube on a 6% sequencing gel. At the completion of the run, fix and dry gel, then autoradiograph 6 to 24 hr without an intensifying screen.

The exposure time varies and is dependent upon several factors, including the gene being amplified, the number of amplification and labeling cycles, and the specific activity of primer 3. The ladder will appear 25 bases longer than the original footprinting or sequencing products due to the addition of the 25-base linker. Because a large amount of material is being applied to the gel, use of the usual 0.2-mm-thick sequencing gels may result in smeared bands; therefore, results are usually better if a relatively thick gel (0.35 to 0.56 mm) is used instead. The first readable sequence band will run at ~ 50 nucleotides, depending on the length of the labeling primer used. Therefore, to read as much sequence as possible, multiple loadings on 80-cm gels are recommended (60-cm gels also work well if an electrolyte gradient is used; UNIT 7.6), and any fragments shorter than 50 nucleotides are run into the bottom buffer tank. This means that the bottom buffer tank will contain a significant amount of radioactive ^{32}P in the form of unused primer 3; thus, disposal guidelines must be rigorously applied.

PREPARATION OF GENOMIC DNA FROM MONOLAYER CELLS FOR DMS FOOTPRINTING

This protocol describes the preparation of in vivo and in vitro dimethyl sulfate (DMS)-treated genomic DNA from monolayer cells for use in conjunction with ligation-mediated PCR. Figure 15.3.4 outlines the parallel processing of duplicate sets of cultured cells: one set is left untreated (control) and the second is treated with DMS (in vivo treatment). Both cultures are lysed and their DNA harvested. The control-cell DNA is then treated with DMS (in vitro treatment). Both in vivo and in vitro DMS-treated DNA are cleaved with piperidine before use in the ligation-mediated PCR assay (Basic Protocol). The protocol can be modified to prepare DNA from suspension cells (Support Protocol 2). The deproteinized control DNA (not treated with DMS) is of high quality and thus can also be used for genomic sequencing as described in Support Protocol 3.

SUPPORT PROTOCOL 1

The Polymerase Chain Reaction

15.3.7

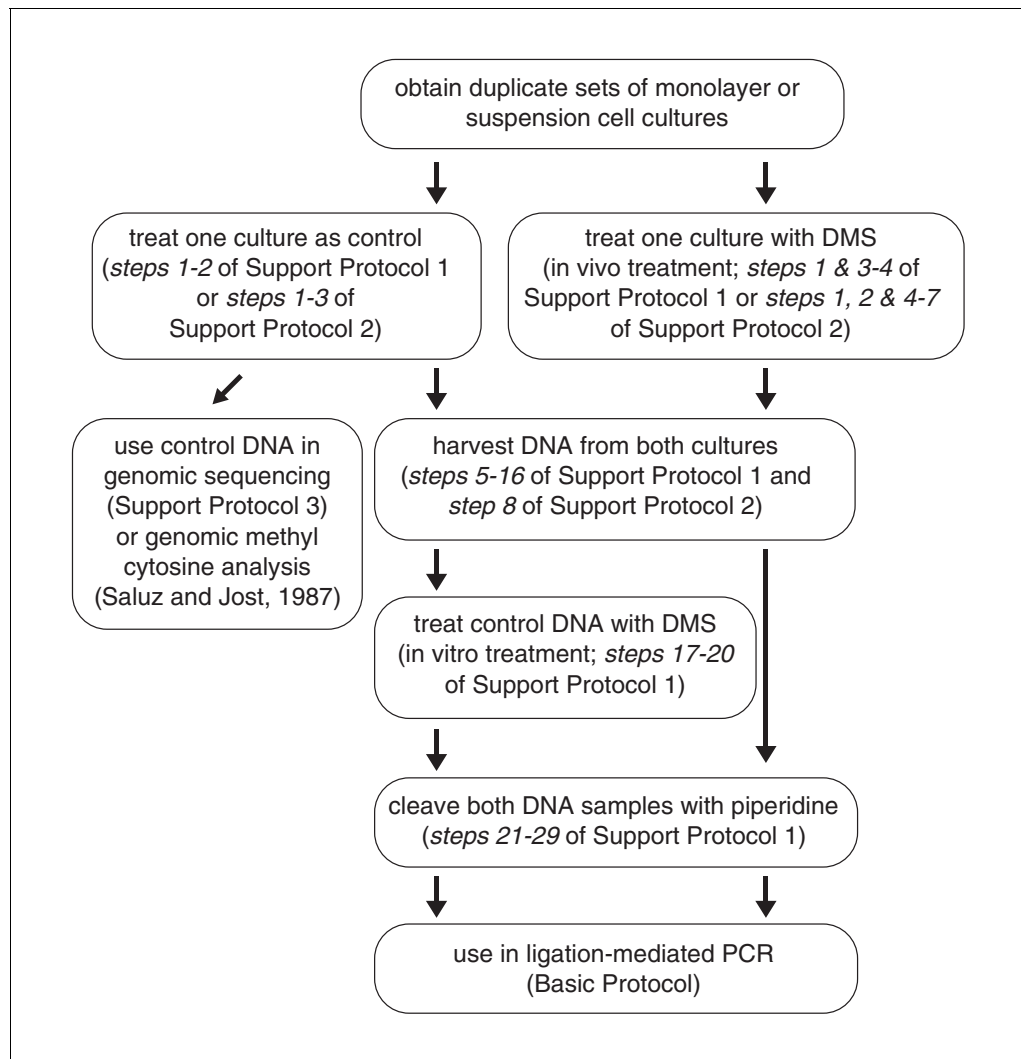


Figure 15.3.4 Flow chart of preparation of DMS-footprinted genomic DNA for use in ligation-mediated PCR.

DMS methylates guanine residues at the N7 position, rendering them susceptible to subsequent cleavage with piperidine. DMS is commonly used for *in vivo* footprinting analysis because cellular membranes are freely and rapidly permeable to it. Treatment of both control (*in vitro*-treated) and experimental (*in vivo* treated) samples with DMS is limited so that only ~1 in 150 guanines is methylated. On the other hand, piperidine treatment is quantitative so that all of the methylated guanines are cleaved. The extraction and purification of the DNA must be thorough so that all proteins and reagents are removed before the DNA is used in ligation-mediated PCR (Basic Protocol).

This procedure works well with $\sim 5 \times 10^7$ cells/set (i.e., one 15-cm plate at ~70% confluence for many fibroblasts), although more cells can be used by scaling up the extractions and $< 10^7$ cells/set will work if extra precautions are taken in precipitating and handling the DNA. The cells should be healthy and adhere well to the plate (see critical parameters).

Materials

- Phosphate-buffered saline (PBS; APPENDIX 2)
- Tissue culture medium appropriate for sample cells
- Lysis solution (see recipe)
- Duplicate 15-cm plates of cells in monolayer culture (see Critical Parameters)

100% dimethyl sulfate (DMS; Aldrich)
 Equilibrated buffered phenol (UNIT 2.1A)
 25:24:1 (v:v:v) phenol/chloroform/isoamyl alcohol (see recipe)
 24:1 (v/v) chloroform/isoamyl alcohol
 Ethyl ether
 Isopropanol
 TE buffer, pH 7.5 (APPENDIX 2)
 3 M sodium acetate, pH 7.0
 100% ethanol, room temperature and chilled on dry ice
 75% ethanol, room temperature
 DMS stop buffer (UNIT 7.5), ice-cold
 Piperidine (Aldrich)
 8 M ammonium acetate

 50- and 15-ml disposable polypropylene screw-cap tubes (e.g., Corning)
 Aspirator attached to waste flask
 Disposable cell scraper
 Tabletop centrifuge (e.g., IEC Centra-7R)
 1.5-ml microcentrifuge tubes, silanized (APPENDIX 3) and with Lid-Loks
 (Intermountain Scientific)
 Sealed Pasteur pipet or thin glass rod

 Additional reagents and equipment for quantitation of DNA (APPENDIX 3)

CAUTION: DMS and concentrated piperidine are extremely toxic. All manipulations that use DMS or piperidine should be performed in a properly functioning fume hood. Before starting this procedure, review precautions for working with and disposing of DMS and piperidine in UNIT 7.5.

Treat monolayer cells with DMS (in vivo treatment)

1. Prepare the following solutions before beginning the procedure: Prewarm PBS (~75 ml/15-cm plate) to 37°C in a water bath inside a fume hood. Aliquot 24 ml tissue culture medium into a 50-ml disposable polypropylene screw-cap tube and prewarm to 37°C as for PBS. Prepare lysis solution.

It is convenient to have the 37°C water bath in or near the fume hood.

2. *For control (untreated) cells:* Remove and discard medium from cells on one 15-cm tissue culture plate, and immediately add ~25 ml prewarmed PBS. Gently swirl plate a few times and remove and discard PBS. Immediately proceed to step 5.
3. *For in vivo DMS-treated cells:* Remove and discard medium from cells on the second 15-cm tissue culture plate. Add 24 µl of 100% DMS to aliquot of prewarmed medium (0.1% DMS final). Screw on cap, mix by inversion, and gently pour entire contents onto tissue culture plate. Leave this 0.1% DMS medium on cells exactly 2 min.

Do not add DMS to the medium until immediately before it is needed; the water in the medium quickly inactivates it. The high heat capacity of the liquid medium should keep the temperature of the reaction at ~37°C for the 2-min incubation, but this can be controlled more effectively by placing a styrofoam sheet (prewarmed to 37°C) under the plate to insulate it.

The incubation time and DMS concentration listed here are good starting conditions; it may be necessary to modify them to get an optimal DMS footprint. Typically, DMS treatment is done at the temperature at which cells are grown (e.g., 37°C for most mammalian cells); if the treatment is done at a different temperature, it is necessary to modify the incubation time or DMS concentration by empirical determination to achieve similar levels of alkylation.

4. Remove and discard the 0.1% DMS medium using an aspirator attached to a flask to hold the DMS waste. Pour ~25 ml prewarmed PBS gently on the cells. Gently swirl plate a few times and aspirate PBS. Wash plate three times with prewarmed PBS, allowing PBS to sit on cells ~30 sec each wash.

The first wash quickly removes most of the remaining DMS from the medium; subsequent washes allow the intracellular DMS to diffuse from the cell. After the last wash, the plate may be removed from the fume hood. When pouring PBS, be very careful not to dislodge the cells; they may be less adherent after DMS treatment.

5. For both untreated and treated cells: After removing as much PBS as possible, pipet 1.5 ml lysis solution onto cells. Gently tip and rock plate to spread solution evenly over cells, then let plate sit at room temperature ~5 min.

The cells should lyse within seconds of contact with the lysis solution. This will be apparent from the viscosity of the DNA/cell lysis mixture.

6. Tilt plate and scrape DNA/lysed-cell slurry to one side with a disposable cell scraper. Slowly remove as much of slurry as possible with a pipet and transfer to 15-ml polypropylene tube.

The DNA/lysed-cell slurry from up to two identically treated plates can be combined in one 15-ml polypropylene tube. If additional plates are used, the slurry may be combined in one or more 50-ml polypropylene tubes. The in vivo DMS-treated DNA must be kept separate from the untreated DNA.

Polypropylene tubes can withstand the organic solvents used in the DNA preparation. Do not use polystyrene tubes.

Harvest the DNA

7. Incubate lysed-cell slurry 3 to 5 hr at 37°C. Mix by inversion every 30 to 60 min.

During this incubation, proteinase K is digesting cellular proteins. After incubation, samples may be stored at -20°C indefinitely. They should be thawed and warmed to room temperature before proceeding with the protocol.

8. Add 1.25 vol buffered phenol, mix thoroughly by gently inverting ~30 times, and centrifuge in a tabletop centrifuge 10 min at 1300 × g (2500 rpm in an IEC Centra-7R), room temperature. Slowly remove phenol (bottom layer) from underneath by inserting a 9-in. Pasteur pipet through the aqueous layer and interface to the bottom of the tube. If the viscous DNA begins to trail along into the organic layer, it can be freed by blowing a few bubbles of air through the pipet using a pipet bulb. Do not remove the interface; much of the DNA may be trapped in it. Repeat phenol extraction once.

For this protocol, it is convenient to remove the organic layer rather than the aqueous layer. All extractions should be performed at room temperature and mixing should be by gentle inversion to avoid shearing the DNA.

Do not be alarmed if there is a large interface and a cloudy aqueous layer after the first extraction; this will clear during subsequent extractions. Both the interface and the aqueous layer will be quite viscous—this is desirable because it means the DNA has not been degraded or sheared, but it can make separating the phases a challenge. It is acceptable to leave some phenol behind, as it will be removed in subsequent extractions.

9. Add 1 vol phenol/chloroform/isoamyl alcohol, mix thoroughly by gently inverting ~30 times, and centrifuge 10 min at 500 × g (1500 rpm in an IEC Centra-7R), room temperature. Remove phenol/chloroform/isoamyl alcohol (bottom layer) from underneath with a 9-in. Pasteur pipet as in step 8. Leave the interface, if any, behind. Repeat this extraction once.

By this point the interface should be disappearing or be gone and the aqueous layer should be clear but still quite viscous. It is acceptable to leave some phenol/chloroform/isoamyl alcohol behind, as it will be removed in the next step.

10. Add 1 vol of chloroform/isoamyl alcohol, mix thoroughly by gently inverting ~30 times, and centrifuge 5 min at $200 \times g$ (1000 rpm in an IEC Centra-7R), room temperature. Remove chloroform/isoamyl alcohol (bottom layer) from underneath with a 9-in. Pasteur pipet as in step 8.

There should be little or no interface; any interface still present should be removed with the organic layer. Remove the entire organic layer at this point.

11. Add 1 vol ethyl ether, mix thoroughly by gently inverting ~30 times, and allow phases to separate by gravity (~1 min). Remove ethyl ether (top layer) with a Pasteur pipet; evaporate any remaining ether in a fume hood (~5 min).

CAUTION: Ethyl ether is highly flammable; use appropriate caution and work in a fume hood.

This extraction will remove small amounts of remaining phenol or chloroform (see UNIT 2.1A). Try to evaporate as much ether as possible; if a small amount remains after evaporation, it will be removed during the precipitations.

12. Add 1 vol isopropanol and mix thoroughly by gently inverting ~30 times.

DNA should instantly precipitate and be visible as a stringy white precipitate.

If $\leq 10^7$ cells were used as starting material, there may not be an instant precipitate. If this is the case, the DNA will need to be recovered by centrifugation (UNIT 2.1A); be careful not to allow it to dry out.

13. Spool the stringy DNA precipitate around the end of a sealed 9-in. Pasteur pipet or thin glass rod. Carefully lift the spooled DNA out of the isopropanol solution, and gently touch the DNA to the inside of the tube to remove excess liquid.

Seal the end of the Pasteur pipet in the flame of a Bunsen burner; allow the hot glass to cool before retrieving the DNA. There is no need to remove all of the liquid from the spooled DNA; never allow the DNA to dry out or it will be impossible to resuspend.

14. Unspool the DNA in 3 ml TE buffer, pH 7.5. Redissolve by gently rocking the sample several hours to overnight at room temperature.

The DNA may be released from the pipet tip by gently shaking and twisting the pipet in the TE buffer. Genomic DNA is difficult to resuspend, but the process should not be forced by vortexing as this will shear the DNA.

15. Add 330 μ l of 3 M sodium acetate and 6.7 ml of ice-cold 100% ethanol; mix thoroughly by gently inverting ~30 times. Reprecipitate and spool DNA as in steps 12 and 13.

16. Unspool DNA in 200 to 500 μ l TE buffer, pH 7.5. Redissolve DNA by storing it overnight at room temperature or for several days at 4°C, inverting tube gently a few times during the process. Quantitate DNA using a spectrophotometer and adjust concentration to 0.5 to 1 mg/ml.

The DNA is stable for years at 4°C. The solution will probably contain a significant amount of RNA—at least as great as the amount of DNA; this will be removed by the piperidine treatment.

The control DNA (not treated with DMS) is of high quality, and at this point can also be used for genomic sequencing (Support Protocol 3) and genomic methyl cytosine analysis (Garrity and Wold, 1992).

Treat control DNA with DMS (in vitro treatment)

17. Put ~75 to 175 μg control DNA in a silanized 1.5-ml microcentrifuge tube. Add TE buffer to bring the volume to 175 μl .

The DNA will be quite viscous, so care should be exercised in pipetting and aliquoting it. It is helpful to use a wider-bore pipet tip (1000- μl pipettor tip or 200- μl pipettor tip with the end cut off).

18. Make a 1% DMS solution by adding 5 μl of 100% DMS to 495 μl water. Mix thoroughly by vortexing 25 sec and microcentrifuge briefly to collect droplets.

Several concentrations of DMS may have to be tested to find conditions that will perfectly match the in vivo conditions used in steps 3 and 4 (see Critical Parameters). Once the 1% DMS solution is made it should be used promptly, as DMS is rapidly inactivated by water. It is useful to have a microcentrifuge near or in the fume hood at this point.

19. Add 25 μl of 1% DMS solution to the control DNA and mix thoroughly by vortexing gently ~25 sec. Avoid getting the DMS solution on the lid; microcentrifuge briefly if it does. Incubate exactly 2 min at room temperature, then add 50 μl ice-cold DMS stop buffer. Immediately add 750 μl of 100% ethanol prechilled on dry ice, mix by vigorously shaking, and plunge tube into powdered dry ice. Leave sample in dry ice for ~30 min.

The combination of DMS stop solution and low temperature (-70°C) ends the reaction and precipitates the DNA. Once the ethanol is added and mixed, the lid of the microcentrifuge tube will have a tendency to pop open; this can be avoided by using a finger to maintain pressure on the lid until the tube is in the dry ice.

20. Prepare in vivo-treated DNA for piperidine treatment in parallel with the in vitro-treated DMS samples. Mix 200 μl DNA (from step 16) with 50 μl ice-cold DMS stop buffer and vortex briefly (3 sec). Add 750 μl of ice-cold 100% ethanol, mix by vigorously shaking, and plunge tube into dry ice. Leave sample in dry ice ~30 min.

Cleave in vivo and in vitro DMS-treated DNA with piperidine

21. Microcentrifuge the two precipitated DNA samples (from steps 19 and 20) for 10 min at 4°C . Remove and discard supernatants, add 1 ml room-temperature 75% ethanol to the DNA pellets, and vortex briefly (~5 sec) until each pellet is dislodged. Microcentrifuge 10 min at 4°C , then remove and discard supernatants.

It is acceptable to leave some of the 75% ethanol wash (i.e., a few microliters) with the pellets. Do not allow the DNA pellets to dry out.

22. Dilute piperidine 1:10 in water (to obtain 1 M final) and add 200 μl to each DNA pellet. Resuspend DNA by incubating at room temperature with intermittent vortexing; pellets usually dissolve in ~15 min.

CAUTION: Aliquot piperidine in a fume hood.

Carefully examine the samples to ensure that the DNA is dissolved; an undissolved pellet will appear as a clear, floating lens in the 1 M piperidine.

23. When pellets are completely dissolved, microcentrifuge briefly to collect droplets and place Lid-Loks on tubes. Heat 30 min at 90°C in the fume hood.

Lid-Loks prevent the tubes from popping open during the heating step, which keeps the volatile piperidine from escaping. Heat treatment with piperidine cleaves the genomic DNA at the methylated guanines, denatures the DNA, and destroys contaminating RNA.

24. Remove Lid-Loks and briefly microcentrifuge tubes to collect condensation. Chill samples on dry ice for 10 min, then evaporate piperidine in a Speedvac evaporator for 1 to 2 hr at room temperature.

Chilling the samples on dry ice will facilitate evaporation of piperidine and will minimize its vaporization during loading into the Speedvac evaporator.

25. Resuspend pellets in 360 μ l TE buffer, pH 7.5, add 40 μ l of 3 M sodium acetate, pH 7.0, and mix by vortexing. Add 1 ml of ice-cold 100% ethanol. Shake vigorously to mix and chill at -20°C for ≥ 2 hr.

Because the DNA is cleaved and single-stranded, it will resuspend easily. The two precipitations in steps 25 to 27 remove piperidine and degraded RNA.

26. Microcentrifuge samples 15 min at 4°C and discard supernatant. Resuspend pellets in 500 μ l TE buffer, pH 7.5, and add 170 μ l of 8 M ammonium acetate. Mix by vortexing. Add 670 μ l isopropanol, shake vigorously to mix, and chill at -20°C for ≥ 2 hr.
27. Microcentrifuge samples 15 min at 4°C and discard supernatant. Add 500 μ l room-temperature 75% ethanol, vortex, and microcentrifuge samples 15 min at 4°C . Remove and discard supernatants, and remove the last traces of ethanol with a pipettor.
28. Resuspend pellets in 50 μ l water and dry ~ 1 hr in a Speedvac evaporator. Resuspend pellets in TE buffer such that the final DNA concentration is ~ 1 $\mu\text{g}/\mu\text{l}$.

This step removes any remaining traces of piperidine.

29. Microcentrifuge samples 10 min at room temperature. Transfer supernatants to a fresh silanized 1.5-ml microcentrifuge tube; discard gelatinous pellet if present. Quantitate DNA with a spectrophotometer and adjust concentration to 0.4 $\mu\text{g}/\mu\text{l}$ with TE buffer, pH 7.5.

The samples are now ready for ligation-mediated PCR.

This centrifuge step removes insoluble matter that has been carried through the procedure; not all DNA preparations will have this. At this point, most of the DNA will be single-stranded because of the piperidine treatment, so it can be assumed that an A_{260} of 1.0 = 40 $\mu\text{g}/\text{ml}$ DNA. The DNA is stable for years at 4°C .

PREPARATION OF GENOMIC DNA FROM SUSPENSION CELLS FOR DMS FOOTPRINTING

Many cell types do not adhere well to tissue culture substrates and/or grow better in suspension culture. This protocol describes *in vivo* treatment of suspension cells with DMS. It works well with $0.5\text{--}1 \times 10^8$ cells, although $<10^7$ cells will suffice if extra precautions are taken to prevent losses during later steps in DNA preparation and handling. The remaining steps of the DNA footprinting procedure (DNA harvest, *in vivo* control treatment, and piperidine treatment) required to prepare DNA for the ligation-mediated PCR assay (Basic Protocol) are detailed in Support Protocol 1.

Materials

- Cells in suspension culture
- Phosphate-buffered saline (PBS; APPENDIX 2)
- Lysis solution (see recipe)
- 100% dimethyl sulfate (DMS; Aldrich)
- 100% ethanol, room-temperature
- 50-ml polypropylene screw-cap tubes (e.g., Corning)
- Tabletop centrifuge (IEC Centra-7R or equivalent)

CAUTION: DMS is extremely toxic. All manipulations that use DMS should be performed in a properly functioning fume hood. Before starting this procedure, review precautions for working with and disposing of DMS in UNIT 7.5.

SUPPORT PROTOCOL 2

1. Prepare three 49-ml aliquots of PBS in 50-ml polypropylene screw-cap tubes and chill at least 30 min on ice prior to use. Make lysis solution and prepare three 2.7-ml aliquots in 15-ml polypropylene screw-cap tubes.
2. Transfer duplicate aliquots of medium containing $0.5\text{--}1 \times 10^8$ cells to 50-ml polypropylene tubes. Centrifuge cells 5 min in a tabletop centrifuge at $500 \times g$ (1500 rpm in IEC Centra-7R), room temperature. Aspirate and discard supernatant, leaving behind sufficient tissue culture medium to resuspend cell pellet in a final volume of 1 ml. Resuspend cells by gently flicking the bottom of the tube with a finger or by gently pipetting up and down.

Different suspension-cell cultures may require adjustments in centrifugation to optimize cell recovery and pellet resuspension without cell damage. These conditions were worked out for EL4 T lymphocytes.

3. *For control (untreated) suspension cells:* Transfer 1 ml resuspended cells to one ice-cold 49-ml PBS aliquot, mix by gentle inversion, and centrifuge 5 min at $500 \times g$, 4°C . Proceed to step 8.
4. *For DMS-treated suspension cells:* Transfer 1 ml resuspended cells to a 1.5-ml microcentrifuge tube and place in 37°C water bath inside a fume hood.
5. Make 10% DMS solution by adding 10 μl of room-temperature 100% DMS to 90 μl of 100% ethanol. Mix thoroughly by vortexing ~ 25 sec, and microcentrifuge briefly to collect droplets.

The 10% DMS solution is made in ethanol because DMS is not soluble in water at this concentration, and it should be used as soon as it is made. It is useful to have a microcentrifuge near or in the fume hood at this point.

6. Add 10 μl of 10% DMS solution to the warmed cells, mix by gentle inversion, and incubate 1 min at 37°C . Immediately transfer cells to the second ice-cold 49-ml PBS aliquot, mix by gentle inversion, and centrifuge 5 min at $500 \times g$, 4°C .

The incubation time and amount of DMS used in the treatment listed here are good starting conditions; it may be necessary to modify them to get an optimal DMS footprint. Typically, the DMS treatment is carried out at the same temperature that the cells are grown (i.e., 37°C); if the treatment is carried out at a different temperature, it will be necessary to modify the incubation time or DMS concentration to achieve similar levels of alkylation.

7. Aspirate and discard supernatant. Quickly resuspend cells with gentle pipetting in 1 to 5 ml ice-cold PBS from the third aliquot. Immediately add ice-cold PBS to fill the 50-ml tube, mix by gentle inversion, and centrifuge 5 min at $500 \times g$, 4°C .
8. *For both control and treated cells:* Aspirate and discard supernatant. Resuspend cells in 300 μl ice-cold PBS and transfer to separate 2.7-ml aliquots of lysis solution. Mix by gentle inversion. Continue with DNA harvesting, in vitro DMS treatment of the control DNA, and piperidine cleavage as in steps 7 through 29 of Support Protocol 1.

When lysis solution is added to the cells, the solution should become quite viscous.

SUPPORT PROTOCOL 3

PREPARATION OF GENOMIC DNA FOR CHEMICAL SEQUENCING

This protocol describes the preparation of genomic DNA for LMPCR-aided direct genomic sequencing (i.e., sequencing without intermediate cloning or amplification steps). For this application genomic DNA, prepared as described in the accompanying protocol, is sequenced using a modification of the chemical strategy described in detail in *UNIT 7.5* (this method also reveals in vivo cytosine methylation, as cytosines methylated at the C5 position do not participate in the C+T and C reactions).

Ligation- Mediated PCR

15.3.14

After the base-specific modification reaction, the DNA is cleaved with piperidine before use in the ligation-mediated PCR assay (Basic Protocol).

Materials

- 0.5 to 1.0 µg/µl untreated genomic DNA in TE buffer, pH 7.5 (control DNA from steps 1 to 16 in Support Protocols 1 and 2)
- TE buffer, pH 7.5 (*APPENDIX 2*)
- 3 M sodium acetate, pH 7.0 (*APPENDIX 2*)
- 100% ethanol, room temperature and chilled on dry ice
- 75% ethanol, room temperature
- 88% formic acid (Fisher Chemical)
- G+A stop solution (see recipe)
- Hydrazine, 98% anhydrous (Aldrich #21,515-5)
- C+T/C stop solution (see recipe)
- 5 M sodium chloride (*APPENDIX 2*)

CAUTION: DMS, formic acid, and hydrazine are toxic. All manipulations using these chemicals should be performed in a properly functioning fume hood. Before starting this procedure, review precautions for working with and disposing of DMS and hydrazine in *UNIT 7.5*.

Prepare G>A reaction

1. Place ~50 to 175 µg control DNA in a 1.5-ml microcentrifuge tube. Add sufficient TE buffer, pH 7.5, to bring the total volume to 175 µl. Label this tube G>A.

The DNA will be quite viscous, so pipet and aliquot it carefully.

2. Treat with DMS as described in Support Protocol 1, steps 18 and 19. Proceed to step 9.

Prepare G+A, C+T, and C reactions

3. Place 20 to 40 µg control DNA in each of three 1.5-ml microcentrifuge tubes. Add sufficient TE buffer to bring the total volume in each to 100 µl. Label one tube G+A, one tube C+T, and one tube C.

The DNA will be quite viscous, so carefully pipet and aliquot it.

4. Add 11 µl of 3 M sodium acetate and 222 µl of room-temperature 100% ethanol to each tube. Mix thoroughly by gently inverting a few times.

The DNA should precipitate immediately.

5. Microcentrifuge precipitated samples 5 min at maximum speed, room temperature, then remove and discard supernatants. Add 500 µl of room-temperature 75% ethanol to each pellet and mix by inversion until pellets are dislodged. Microcentrifuge 3 min at maximum speed, room temperature, then remove and discard supernatants.

Try to remove as much of the 75% ethanol as possible with a pipettor, but do not allow the DNA pellets to dry out. Do not use a Speedvac evaporator, as it will be impossible to resuspend the DNA once it is dried.

6. Redissolve samples in water as follows: add 18 µl to G+A tube, 40 µl to C+T tube, and 10 µl to C tube. Incubate overnight at 4°C. Warm redissolved DNA to room temperature.

Uncleaved genomic DNA redissolves slowly at these concentrations. Gentle mixing by flicking the tube with a finger a few times during the resuspension can help speed this process. Overnight incubation is usually sufficient to allow resuspension. A small amount of insoluble material is occasionally seen at this step. It does not interfere with subsequent

reactions and can be ignored; it will be discarded after the piperidine cleavage reaction in step 9.

For the G+A sample:

- 7a. Add 54 μl of 88% formic acid, mix thoroughly by vortexing 25 sec, and microcentrifuge briefly to collect droplets. Incubate 7 min at room temperature.
- 8a. Add 164 μl ice-cold G+A stop solution and vortex briefly (3 sec). Add 750 μl of 100% ethanol prechilled on dry ice, mix by vigorously shaking, and plunge tube into powdered dry ice. Leave samples in dry ice for \sim 30 min, then proceed to step 9.

The 7-min incubation time used here is a good starting condition. The extent of reaction can be changed by varying the incubation time; the average size of the final product after piperidine cleavage will decrease as incubation time is increased.

For the C+T sample:

- 7b. Add 60 μl hydrazine, mix thoroughly by vortexing 25 sec, and microcentrifuge briefly to collect droplets. Incubate 3 min at room temperature.
- 8b. Add 150 μl ice-cold C+T/C stop solution and vortex briefly (3 sec). Add 750 μl of 100% ethanol prechilled on dry ice, mix by vigorously shaking, and plunge tube into powdered dry ice. Leave samples in dry ice for \sim 30 min, then proceed to step 9.

The 3-min incubation time used here is a good starting condition. The extent of reaction can be changed by varying the incubation time, as described in step 8a.

For the C sample:

- 7c. Add 30 μl of 5 M sodium chloride, mix thoroughly by vortexing 25 sec, and microcentrifuge briefly to collect droplets. Add 60 μl hydrazine, mix thoroughly by vortexing 25 sec, and microcentrifuge briefly to collect droplets. Incubate 3 min at room temperature.
- 8c. Add 150 μl ice-cold C+T/C stop solution and vortex briefly (3 sec). Add 750 μl of 100% ethanol prechilled on dry ice, mix by vigorously shaking, and plunge tube into powdered dry ice. Leave samples in dry ice for \sim 30 min, then proceed to step 9.

The 3-min incubation time used here is a good starting condition. The extent of reaction can be changed by varying the incubation time, as described in step 8a.

Carry out cleavage and sequencing

9. Cleave DNA from each of the four reactions with piperidine as described in Support Protocol 1, steps 21 to 29. Proceed to ligation-mediated PCR (Basic Protocol).

REAGENTS AND SOLUTIONS

4dNTP mix, pH 7.0, 25mM

Prepare as in UNIT 3.4. Alternatively, purchase as individual 100-mM stock solutions from Pharmacia Biotech and combine equal amounts of dATP, dCTP, dGTP, and dTTP to obtain the 25 mM 4dNTP mix. Store at -20°C .

Amplification buffer, 5 \times

200 mM NaCl
100 mM Tris·Cl, pH 8.9 (APPENDIX 2), room temperature
25 mM MgSO₄
0.05% gelatin (from bovine skin; Sigma)
0.5% Triton X-100
Store at -20°C

Amplification mix, per reaction (30 μ l):

- 20.0 μ l 5 \times amplification buffer (see recipe)
- 1.0 μ l 10 pmol/ μ l oligonucleotide LMPCR.1 (linker primer; Fig. 15.3.3)
- 1.0 μ l 10 pmol/ μ l oligonucleotide primer 2 (Figs. 15.3.1 & 15.3.2)
- 0.8 μ l 25 mM 4dNTP mix, pH 7.0 (see recipe)
- 7.2 μ l H₂O

Prepare immediately before use and chill on ice

Addition of LMPCR.1 may not be required, but this has not been tested.

C+T/C stop solution

- 400 mM sodium acetate, pH 7.5
 - 0.14 mM EDTA, pH 8.0 (APPENDIX 2)
- Prepare fresh and chill on ice before use

End-labeling mix, per reaction (5 μ l):

- 1.0 μ l 5 \times amplification buffer (see recipe)
- 2.3 μ l 1 pmol/ μ l end-labeled primer 3 (Fig. 15.3.3)
- 0.4 μ l 25 mM 4dNTP mix, pH 7.0 (see recipe)
- 0.8 μ l H₂O
- 0.5 μ l 2 U/ μ l *Thermococcus litoralis* DNA polymerase (Vent; New England Biolabs)

Prepare immediately before use. Mix first four components and chill on ice before adding Vent DNA polymerase and chill on ice. Then add polymerase and keep on ice.

5' end-labeled primer 3 can be prepared as in UNIT 3.10 (forward reaction) with the following modifications: end-label 20 to 100 pmol of primer 3 with labeling-grade (i.e., less expensive) [γ -³²P] ATP (~6000 Ci/mmol); incubate at 37°C for 30 min instead of 60 min. Remove unincorporated ³²P by gel purification (UNIT 2.12) or by using a Nensorb-20 nucleic acid purification cartridge (Du Pont NEN)—elute primer with 50% ethanol instead of 50% methanol. Resuspend primer 3 in TE buffer, pH 7.5, so that its concentration is 1 pmol/ μ l; the specific activity should be 4–9 \times 10⁶ cpm/pmol.

First-strand buffer, 5 \times

- 200 mM NaCl
 - 50 mM Tris-Cl, pH 8.9 (APPENDIX 2) room temperature
 - 25 mM MgSO₄
 - 0.05% gelatin (from bovine skin; Sigma)
- Store at –20°C

First-strand synthesis mix, per reaction (25 μ l):

- 6.0 μ l 5 \times first-strand buffer (see recipe)
- 0.3 μ l 1 pmol/ μ l oligonucleotide primer 1
- 0.24 μ l 25 mM 4dNTP mix, pH 7.0 (see recipe)
- 18.21 μ l H₂O
- 0.25 μ l 2 U/ μ l *Thermococcus litoralis* DNA polymerase (Vent; New England Biolabs #254L)

Prepare immediately before use. Mix all components except Vent DNA polymerase and chill on ice. Then add polymerase and keep on ice.

Primer 1 should be stored in a silanized tube at a concentration \geq 10 pmol, then diluted in TE buffer, pH 7.5, immediately before use. This will minimize primer loss due to sticking to the tube walls.

G+A stop solution

360 mM sodium acetate, pH 7.5
0.14 mM EDTA, pH 8.0 (*APPENDIX 2*)
Prepare fresh and chill on ice before use

Ligase dilution solution, per reaction (20 μ l):

2.2 μ l 1 M Tris-Cl, pH 7.5, at room temperature (*APPENDIX 2*; 110 mM final)
0.35 μ l 1 M MgCl₂ (17.5 mM final)
1.0 μ l 1 M DTT (50 mM final)
0.25 μ l 10 mg/ml BSA (DNase-free; Pharmacia Biotech; 125 μ g/ml final)
16.2 μ l H₂O
Prepare immediately before use and chill on ice

Ligase mix, per reaction (25.0 μ l):

0.25 μ l 1 M MgCl₂ (*APPENDIX 2*; 10 mM final)
0.50 μ l 1 M DTT (*APPENDIX 2*; 20 mM final)
0.75 μ l 100 mM ATP, pH 7.0 (Pharmacia Biotech; 3 mM final)
0.125 μ l 10 mg/ml BSA (DNase-free; Pharmacia Biotech; 50 μ g/ml final)
17.375 μ l H₂O
5.0 μ l 20 μ M unidirectional linker mix (see recipe; 4 μ M linker and 50 mM Tris-Cl, pH 7.7, final)
1.0 μ l 3 “Weiss” U/ μ l T4 DNA ligase (*UNIT 3.14*; 3 U final)
Prepare immediately before use. First mix MgCl₂, DTT, rATP, BSA, and H₂O and chill on ice. Next add ice-cold unidirectional linker mix and T4 DNA ligase.
Tris-Cl added in the linker mix is 50 mM, pH 7.7, at final concentration.

Loading buffer

80% (v/v) formamide, deionized (see *UNIT 14.3*)
45 mM Tris base
45 mM boric acid
1 mM EDTA
0.05% (w/v) bromphenol blue
0.05% (w/v) xylene cyanol
Store in aliquots at -20°C and discard after 3 months

Lysis solution

300 mM NaCl
50 mM Tris-Cl, pH 8.0 (*APPENDIX 2*), room temperature
25 mM EDTA, pH 8.0 (*APPENDIX 2*)
0.2% (v/v) SDS (prepare fresh; add just before use)
0.2 mg/ml proteinase K (prepare fresh; add just before use)
Lysis solution without SDS and proteinase K can be stored indefinitely at room temperature. Immediately before use, add 10 μ l of 20% SDS and 10 μ l of 20 mg/ml proteinase K per milliliter lysis solution.

Phenol/chloroform/isoamyl alcohol

Mix 25 parts phenol (equilibrated in 150 mM NaCl/50 mM Tris-Cl (pH 7.5)/1 mM EDTA) with 24 parts chloroform and 1 part isoamyl alcohol. Add 8-hydroxyquinoline to 0.1%. Store in aliquots at -20°C and discard after 6 months.

Precipitation salt mix, per reaction (9.4 μ l):

8.4 μ l 3 M sodium acetate, pH 7.0 (2.7 M final)
1.0 μ l 10 mg/ml yeast tRNA (~1 mg/ml final)
Prepare immediately before use

Unidirectional linker mix, 20 μ M

20 μ M oligonucleotide LMPCR.1 (Fig. 15.3.3)

20 μ M oligonucleotide LMPCR.2 (Fig. 15.3.3)

250 mM Tris·Cl, pH 7.7 (APPENDIX 2)

Prepare this mix in advance as follows: (1) purify the oligonucleotides on denaturing polyacrylamide gels (UNIT 2.12), (2) combine the two oligonucleotides and Tris·Cl and heat 5 min at 95°C, (3) transfer to 70°C and gradually cool ~1 hr to room temperature, (4) leave ~1 hr at room temperature and then gradually cool ~1 hr to 4°C, and (5) leave ~12 hr at 4°C and store in aliquots at –20°C. Thaw linker on ice before use.

Hybridization of the linker undoubtedly takes place more rapidly than allowed for in this procedure, but this has reproducibly worked well. Because the monovalent salts (e.g., NaCl) more traditionally used to raise ionic concentration inhibit T4 DNA ligase, and Tris·Cl is required by the ligase in any case, Tris·Cl is used as the salt in this hybridization reaction. Consequently, it is difficult to calculate the kinetics of oligonucleotide annealing, and so the question of timing has not been studied in detail.

Vent DNA polymerase mix, per reaction (3.0 μ l)

Mix 0.6 μ l of 5 \times amplification buffer (see recipe) with 1.9 μ l H₂O and chill several minutes in an ice-water bath. Add 0.5 μ l (1 U) *Thermococcus litoralis* DNA polymerase (Vent; New England Biolabs), gently mix, and return to ice-water bath. Prepare just before use.

Vent DNA polymerase stop solution, per reaction (295 μ l):

25 μ l 3 M sodium acetate, pH 7.0 (260 mM final)

266 μ l TE buffer, pH 7.5 (APPENDIX 2; 10 mM Tris·Cl, pH 7.5, final)

2 μ l 0.5 M EDTA, pH 8.0 (APPENDIX 2; ~4 mM final)

2 μ l 10 mg/ml yeast tRNA (68 μ g/ml final)

Prepare immediately before use

COMMENTARY

Background Information

Ligation-mediated PCR was originally developed to study in vivo protein-DNA interactions at regions of genes important for transcriptional control of expression. Several genomic sequencing strategies have been previously applied to in vivo footprinting. These relied on the principles of Southern blotting (Church and Gilbert, 1984), solution hybridization (Jackson and Felsenfeld, 1985), or primer extension (Huibregtse and Engelke, 1986). Unfortunately, with these strategies, in vivo footprinting of a single-copy regulatory region in large genomes (e.g., mammalian) is technically challenging, requires large cell numbers, and often produces results with unacceptable signal-to-noise ratios.

The polymerase chain reaction (Saiki et al., 1988; White et al., 1989) is used to amplify specific fragments of single-copy genes (see Chapter 15 introduction and UNIT 15.1). It relies on two primers that flank the specific fragment of DNA to be amplified and uses repeated

cycles of template denaturation, primer annealing, and extension by a thermally stable DNA polymerase to exponentially amplify that fragment. PCR is not immediately applicable to sequencing or footprinting because it requires that every fragment of DNA have two defined ends (a sequence ladder is composed of a population of related DNA fragments that differ in length; they all have one end in common but the other end varies depending on the sequence). The Basic Protocol in this unit circumvents this problem by attaching a common sequence to the variable end of every fragment, thus creating substrates suitable for PCR and resulting in exponential amplification of each representative of the sequence ladder. One end of every fragment is primed with a gene-specific primer and the other is primed with the common attached sequence. Footprint and sequence information contained in the original sample of genomic DNA is amplified by an additional 10⁴ to 10⁵ copies. When visualized on a sequencing gel, each band in the ladder is

shifted to a higher position that corresponds to the length of attached sequence. In practical terms, starting with as little as 2 μg of mammalian DNA (equivalent to 3×10^5 cells), a genomic sequence ladder of a single-copy gene can be seen following overnight exposure without an intensifying screen.

Ligation-mediated PCR differs significantly from another single-sided PCR procedure that uses terminal transferase to add a homopolymeric tail of guanines or adenines to one end of the substrate DNA (Frohman et al., 1988; Loh et al., 1989; Ohara et al., 1989). This tail is added nonspecifically to all DNA and is expected to be somewhat variable in length. In contrast, ligation-mediated PCR results in the blunt-end ligation-based addition of a sequence of defined length and composition. The blunt end is generated by the extension of a gene-specific primer. The initial hybridization of this primer selects a small subset of specific gene fragments (to be modified in later steps) from the original complex DNA sample. This level of selection is unique to ligation-mediated single-sided PCR. These features, together with the use of multiple gene-specific primers, greatly reduce nonspecific background and provide the single-base resolution required for genomic footprinting, sequencing genomic DNA (Pfeifer et al., 1989; Garrity and Wold, 1992), studying *in vivo* methylation patterns of cytosine residues (Pfeifer et al., 1989, 1990; Rideout et al., 1990; Garrity and Wold, 1992), and cloning promoter elements from restriction-digested genomic DNA (Fors et al., 1990). This last adaptation uses gene-specific primers selected from known sequences located at the 5' end of an mRNA to clone the corresponding unknown promoter region.

Support Protocols 1, 2, and 3 describe methods for the preparation of genomic DNA for genomic DMS footprinting and genomic sequencing. Both of these methods are based on the Maxam-Gilbert chemistry described in *UNIT 7.5*. In addition, DNA cleaved randomly by a nuclease such as DNase I (*UNIT 12.4*; Tanguay et al., 1990; Rigaud et al., 1991), or site-specifically by a restriction endonuclease (*UNIT 3.1*) can also in principle be utilized in ligation-mediated PCR. In any case, the cleaved DNA must contain a 5' phosphate to participate in the ligation reaction.

Critical Parameters

Several important parameters are discussed in this section, but two underlying technical

points deserve emphasis at the outset. First, in order to keep background to a minimum, the hybridizations and extensions should be performed at the highest possible temperatures. Hybridizations are incubated at or 2°C above the calculated melting temperature (T_m) of the primers, and extensions are performed at 76°C with Vent DNA polymerase. (In specific cases, these temperatures may need to be adjusted, but they have worked with all combinations of primers tested to date). Second, background can be greatly reduced by using multiple primers, each primer having an end that extends 3' to the previous one. This produces a level of specificity that cannot be obtained with a single 25-base primer.

It is recommended that a test experiment be done before any footprinting or sequencing experiments are undertaken. This should include several identical samples that are processed in parallel. When visualized on a sequencing gel, these samples should appear identical. In addition, this is a convenient way to test for any problems in the operation of the automated thermal cycler or in the execution of ligation-mediated PCR. Whenever possible, prepare solutions as a reaction cocktail; this will minimize sample variation. Oligonucleotides should be gel-purified (*UNIT 2.12*) to remove contaminants and stored in TE buffer, pH 7.5, at ≥ 10 pmol/ μl and at -20°C .

Gene-specific primers

Primer position. Figure 15.3.2 shows possible arrangements for gene-specific primers. Several different primer combinations have been tested and the following operating rules were derived (P.R.M. and B.W., unpub. observ.). Primers 1 and 2 may or may not overlap but the extending end of primer 2 must be 3' to primer 1. If they do overlap, the overlap should be less than half their length. On the other hand, primer 3 can completely overlap primer 2 and extend a few extra bases 3' to it (Fig. 15.3.2A), and it must be positioned with an overlap of at least half the length of primer 2 (Fig. 15.3.2B). If primers 2 and 3 do not overlap, the labeling extension will not be successful, because it is necessary for these primers to compete for the same binding site. The basis of this phenomenon is not known. It may be caused by the premature extension product of primer 2 excluding the hybridization of primer 3. The hybridization temperatures used during the annealing phase of the labeling cycle(s) are close to the optimal temperature for Vent DNA po-

lymerase. This is done to minimize background, but probably results in the immediate extension of any primer following its binding.

Primer composition and melting temperature. Elevated temperatures minimize the formation of imperfect duplexes between genomic DNA and primers. Thus, to minimize background, hybridizations and reactions should be carried out under temperature conditions that demand as much as much specificity as possible. For hybridizations, the optimal temperature is at or even above the calculated T_m of the primer (see below for calculation). This principle is also used to favor the binding of the labeling primer (primer 3) over the amplification primer (primer 2). In general, the T_m of the primers should increase in the order primer 1 < primer 2 < primer 3. Thus, the binding of the first-strand synthesis primer (primer 1) will be less stable under amplification conditions and the binding of the amplification primer (primer 2) will be less stable under labeling conditions. In practice, this doesn't appear to be essential, but it would probably be unwise to have the relative T_m of the primers reversed.

The T_m of the primers can be adjusted by choosing sequences with differing GC content and/or length (see below). The first-strand synthesis primer (primer 1) should have a T_m of $\geq 60^\circ\text{C}$; primers of 20 (~60% GC content) to 25 bases (~50% GC content) work well. If the T_m of the primer is below 60°C , it may be necessary to reduce the hybridization temperature below the suggested 60°C (see Basic Protocol, step 3). The amplification primer (primer 2) should have a T_m equal to or higher than that of primer 1, and should closely match the T_m of the linker primer; a primer of 25 bases with a GC content of ~60% works well. The labeling primer (primer 3) should have an even higher T_m . During the labeling extension, the linker primer does not need to extend, and the labeling primer can better compete with the amplification primer for occupancy if it has a higher T_m (recall that the amplification primer and labeling primer must overlap or the labeling primer will be excluded).

Overall, the exact GC content or length of the primers is probably not very important. What is important is their relationship to each other and to the temperature of the reactions. Primers must be long enough to be specific to the gene of interest, but they should not be excessively long or have a T_m greater than $\sim 76^\circ\text{C}$ (this may lead to nonspecific binding). The following formula (Wahl et al., 1987) may be used to approximate the T_m of primers:

$$T_m = 81.5 + 16.6(\log M) + 0.41(\%GC) - (500/n)$$

where n = length of primer and M = molarity of the salt in the buffer. To determine the molarity, ignore the contribution of Mg^{2+} and multiply the Tris concentration by 0.67 (N. Davidson, pers. comm.). Therefore, for the amplification buffer, the molarity would be $0.040(\text{NaCl}) + 0.67 \times 0.020(\text{Tris}) = 0.053 \text{ M}$ salt. Hybridizations work best if done at or 2°C above this calculated T_m . As an example, the T_m calculation for a 25-base primer with a 60% GC content is as follows: $81.5 + 16.6(\log 0.053) + .41(60) - (500/25) = 65^\circ\text{C}$. Hybridization with this primer should be carried out between 65° and 67°C . This is only an approximation, and base stacking and near-neighbor effects may be significant for particular primers. Therefore, it may be necessary to determine empirically the optimal hybridization temperatures for each primer.

Unidirectional linker

The staggered design of the oligonucleotide linker is important because it assures that the ligation event is directional (see Fig. 15.3.3). The 3' hydroxyl of the long oligonucleotide (LMPCR.1) will be ligated to the 5' phosphate of the genomic DNA. For cloning purposes, LMPCR.1 contains restriction sites for *Bst*EII, *Sma*I, *Bgl*II, and *Eco*RI; however, these are not used in sequencing or footprinting experiments. This oligonucleotide will also function as the linker primer in the PCR amplification. On the other hand, the short oligonucleotide (LMPCR.2) functions only in a structural role. When it is hybridized to the long oligonucleotide, it creates a blunt-end duplex at the 3' end of LMPCR.1. This enables T4 DNA ligase to attach LMPCR.1 to the blunt-end duplex of the genomic DNA. By making the short oligonucleotide small and low in GC content, it is possible to make it unable to serve as a primer in subsequent Vent DNA polymerase reactions that are carried out far above its T_m . To prevent linker self-ligation, neither oligonucleotide has a 5' phosphate. By default, this prevents LMPCR.2 from ligating to the first-strand synthesis extension products, but this is not an essential design feature (see Fig. 15.3.1).

The exact sequence, GC content, length, and type of restriction endonuclease sites in this linker are not thought to be important for the function of the linker. The crucial properties for the linker are that (1) it contain no 5' phosphates and be staggered to eliminate self-ligation and assure unidirectionality, (2) it be a ligatable

structure, (3) the duplex between the two oligonucleotides be stable during ligation conditions but not during PCR conditions, and (4) the long oligonucleotide have a T_m in Vent DNA polymerase buffer comparable to that of primer 2.

Ligation

Most products of the first-strand synthesis reaction appear to be ligated after 45 min at 15°C. However, particular bands in some ladders require longer incubation times to reach full intensity (Mueller and Wold, 1991). For this reason, an overnight (6 to 15 hr) ligation at 15° to 17°C is suggested. Although a higher ligation temperature might decrease the time required for complete ligation, this could inadvertently permit undesirable Vent DNA polymerase activity. Prior to ligation, no attempt is made to destroy the Vent DNA polymerase used during the first-strand synthesis. A possible concern is that if the polymerase is active during the ligation, it may extend the short oligonucleotide of the staggered linker (this would compromise the unidirectionality of the linker addition). With ligation at 15° to 17°C, this does not appear to be a problem, but it may be unwise to attempt higher temperatures unless the polymerase is inactivated.

Vent DNA polymerase

Earlier versions of ligation-mediated PCR used a combination of Sequenase and *Taq* DNA polymerase for the first-strand synthesis and amplification/labeling reactions, respectively. Both of these enzymes have been replaced by *Thermococcus litoralis* DNA polymerase (Vent DNA polymerase). Vent DNA polymerase virtually eliminates the unequal representation of some members of a sequence ladder and occasional spurious bands that were previously observed with Sequenase/*Taq* polymerase-based ligation-mediated PCR. It also provides an approximate three-fold increase in signal strength to all members of the sequence ladder (Garrity and Wold, 1992).

A very important procedural point is that ligation-mediated PCR is very sensitive to the amount of Vent DNA polymerase activity. A small (e.g., two-fold) increase in polymerase can result in an unacceptable background level. The reason for this threshold response to the amount of Vent DNA polymerase activity is unknown, but it does mean that this polymerase must be used carefully. If high background is a problem (or suddenly becomes one), the polymerase should be titrated. Normally it is not necessary to titrate each lot of Vent DNA po-

lymerase, but a change in manufacturing at New England Biolabs occurred between early lots (i.e., prior to the fall of 1991), which were less active, and more recent ones. Thus, originally 1 U of Vent DNA polymerase was needed in the first-strand synthesis and 3 U were needed at the amplification step (Garrity and Wold, 1992), but now 0.5 U in the first-strand synthesis and 1.0 U at the amplification step appears to be the ideal amount (P.G. and B.W., unpub. observ.). The 1.0 U added at the labeling step does not need to be altered. The “exonuclease-minus” version of Vent DNA polymerase has not been tested in this procedure.

The buffers recommended here for use with Vent DNA polymerase in ligation-mediated PCR have been developed through empirical testing and differ from the one recommended by the manufacturer (P.G. and B.W., unpub. observ.). KCl has been replaced by NaCl, which gives better extension through certain guanine-rich regions. BSA has been replaced by gelatin, because this reduces the interface formed after the organic extraction. $(\text{NH}_4)_2\text{SO}_4$ has been eliminated, because this reduces background and the amount of excess salt deposited in the ethanol-precipitation pellets. Finally, the concentration of MgSO_4 has been increased, because this increases signal strength. These conditions may need to be adjusted in specific cases, but so far they have worked with all combinations of primers tested. The first-strand synthesis buffer differs slightly from the amplification/labeling buffer (i.e., there is less Tris and no Triton X-100). This change facilitates the conversion of first-strand buffer to ligation buffer without affecting polymerase activity for the single round of extension.

Vent DNA polymerase is added last to each sample and then only after the sample has been chilled on ice. This is to minimize the single- and double-stranded exonuclease activities reported by the manufacturer that are inhibited by cold and by the presence of dNTPs.

Efficiency and statistical limitations (founder effects)

In this procedure, exponential amplification is so effective that the minimum amount of starting DNA needed is not dictated by signal strength; instead statistical considerations become important. Conventional PCR can be successfully performed using the DNA of a single genome (White et al., 1989). Obviously, footprinting or sequencing reactions require much more material than this, but determining exactly how much more is needed is not straight-

forward (UNITS 7.5 & 12.4). It is useful to first estimate the number of starting genomes or founder molecules that would be required in an idealized case in which all steps in the procedure were 100% efficient. To generate a proper sequence ladder, all “rungs” must be represented at least once (i.e., by at least one starting chromosome). Limited cleavage is used to generate DNA fragments that are, on average, ~600 bases long (UNITS 7.5 & 12.4), and each of these can be used to generate a rung on the ladder. In the specific case of a DMS/piperidine guanine ladder, there are ~150 possible cleavage sites (guanine residues) in an average 600-nucleotide fragment. Therefore, ≥ 150 chromosomes are required to generate all of the rungs in a ladder. However, because cleavage is random, ≥ 10 representatives (on average) must be generated to ensure that each rung is derived at least once. Furthermore, in footprinting, quantitative differences in the intensity of individual bands is important, and the maximum acceptable variation in band intensity due to sampling considerations should be $< 10\%$. This level of certainty could be expected with 100 representatives for each rung. Therefore in theory, 15,000 genomes or founder molecules will be required to generate a footprint ladder (150 genomes per rung \times 100 representatives per rung).

This idealized example assumes that it is possible to cleave and represent all fragments equally. This is not the case in experimental practice. The intensity of any particular band on a sequencing gel is influenced by several parameters. For example, some bases are not as easily cleaved as others and the first-strand synthesis and blunt-end ligation steps of ligation-mediated PCR are not 100% efficient. If these variables combine to cause up to a 20-fold reduction in the frequency of the least prevalent fragments, then the weakest bands may require 20-fold more genomes than in the ideal case to be represented adequately. Therefore, by this estimation, 3.0×10^5 ($20 \times 15,000$) genomes or founder molecules will be required to generate a reliable footprint ladder. Anything that reduces the generation (cleavage conditions) or utilization (first-strand synthesis and linker ligation) of founder molecules will further increase the number of required starting genomes.

This crude calculation highlights an important point: if too small a number of genomes are used as starting material, the results may be subject to artifacts caused by founder effects. Empirically, footprinting experiments on multiple genes indicate that 6×10^5 genomes ($3 \times$

10^5 diploid cells) are sufficient to eliminate any apparent founder effects (P.R.M. and B.W., unpub. observ.). Using fewer genomes may result in missing bands or fluctuation in band intensity. Of course, the amount of starting material required will depend on the application. Founder effects must be kept to a minimum in a quantitative footprinting experiment, whereas a sequencing experiment will be more tolerant of band-intensity fluctuation as long as each band is still visible. Cloning applications of ligation-mediated PCR are much less subject to this problem and will tolerate use of substantially less starting material.

DMS footprinting

Genomic DMS footprinting is a comparison between the DMS sensitivity of DNA bound by cellular proteins inside the cell (in vivo DNA) with DMS sensitivity of deproteinized DNA outside the cell (in vitro or “naked” DNA). This requires high-quality in vitro and in vivo DNA samples that are carefully matched as to genotype, cell type, cleavage condition, and concentration. This can be achieved by using the same cell preparation for both in vitro and in vivo DNA and by processing the samples in parallel. DMS and piperidine should be of high quality and be replaced 9 to 12 months after they are opened. The final preparation of DNA must be free of all contaminants (especially piperidine), and its concentration must be determined accurately.

The DNA preparation procedure outlined in the Support Protocol 1 has been used successfully to isolate DNA from fibroblasts and a number of other cultured cell types. Cells should be healthy and adhere well to the plate. If the cell line does not adhere well, consider precoating the plate with collagen or polylysine-fibronectin or using suspension cells (Support Protocol 2). Some cell types may require more rigorous treatment of the DNA to remove impurities. In these cases, additional extractions or the use of chaotropic agents such as guanidinium in the DNA preparation may be necessary. The physiological status of the cell source for genomic DNA is important; specifically, the gene(s) to be footprinted should be uniformly expressed within the cell population. A heterogeneous background of expressing and nonexpressing cells will result in weak or confusing footprints.

Troubleshooting

In cases where comparison between multiple reactions is required (e.g., footprinting), it

is important that the samples be prepared and processed in parallel. Cocktails of reagents and enzyme solutions should be used whenever applicable. A potential problem associated with the use of an automated thermal cycler is well-to-well variability in the temperature cycling parameters (Linz, 1990). This could potentially lead to lane-to-lane variability in the footprint ladder, although this has not been a widely observed problem.

As previously mentioned, it is important that the highest-quality reagents available be used. In chemical footprinting and sequencing reactions, it is very important to use fresh DMS and piperidine, as old stocks give high levels of background. Moreover, all traces of piperidine must be removed from the DNA before any enzymatic step (Saluz and Jost, 1987). This can be accomplished through use of a Speedvac evaporator and doing at least two ethanol and isopropanol precipitations (UNIT 2.1). The precipitations will also remove the ribonucleotides and small oligoribonucleotides generated by piperidine degradation of contaminating RNA. (Unless the genomic DNA preparation was previously treated with RNase or base/heat, it will contain RNA. The heating in the presence of base performed as part of the piperidine treatment conveniently destroys this RNA, but the resulting RNA fragments must be removed to get an accurate determination of the DNA concentration.)

The primer and linker oligonucleotides should be gel purified on polyacrylamide gels and unincorporated ^{32}P should be removed from the end-labeled primer by gel purification or with a Nensorb-20 nucleic acid purification cartridge. Gel reagents (e.g., loading dye and polyacrylamide) must also be fairly fresh, as old stocks may cause high background and fuzzy bands.

If reagents and starting material are good, but results are still less than optimal, it may be necessary to modify the protocol. The first thing to check is the amount of Vent DNA polymerase being used, as even a small excess in polymerase activity will lead to high background. Alternatively, background can sometimes be reduced by increasing the annealing temperatures during the polymerase reactions. Changing primers can also improve results. As has been observed in other PCR procedures, some primers just do not work (UNIT 15.1). Unfortunately, it is not always simple to predict which primers will fail or succeed. A problem that has occurred on multiple occasions is that the published sequence of a gene does not always

match the sequence in the experimental genome, either because of polymorphisms or sequencing errors. Because high annealing and extension temperatures are used in ligation-mediated PCR, even single-base sequence mismatches can lead to failure. It is important to test the primers by conventional PCR on the experimental DNA sample.

Anticipated Results

Two points need to be considered when discussing anticipated results: the first is sensitivity (how long an exposure should be required to visualize the footprint or sequence ladder) and the second is appearance (how that ladder should look). The exposure time will depend largely on the number of input genomes and the numbers of amplification and labeling cycles performed. However, in general, starting with 6×10^5 genomes of cellular DNA (i.e., 2 μg of mammalian DNA), a sequence or footprint ladder should be visible in 6 to 24 hr using Kodak XAR-P film without an intensifying screen.

The quality of the footprint and sequence ladders is a more critical measure of the success of the experiment. The ladder should extend ≥ 500 bases and with a double loading on an 80-cm, 6% polyacrylamide sequencing gel, it should be possible to resolve footprints or sequences to ~ 300 bases. Except for the expected differences in footprint or sequence reactions (i.e., when different base-specific cleavage reactions are performed or when cytosine methylation is present), the lanes should appear identical. For example, in a genomic footprinting experiment, it is crucial that the *in vitro* and *in vivo* levels of DMS alkylation match (Mueller et al., 1988). If they don't, footprints may be incorrectly assigned. Variations in background, weak signals, and unexpected banding patterns are all signs of problems, which in most cases can be traced to the quality of the starting material.

The DNA yield varies, but $\sim 5 \times 10^7$ mammalian cells should yield ~ 300 to 600 μg nucleic acid (DNA and RNA) before piperidine treatment. During the piperidine treatment, RNA is destroyed (and removed by the precipitations) so ~ 40 to 60% of the input nucleic acid recovered will be DNA.

For additional comments on results, see Critical Parameters for discussion of efficiency and statistics.

Time Considerations

Although ligation-mediated PCR is a relatively complex, time-intensive procedure, it can

routinely yield high-quality, reproducible results. Organization of time simplifies the procedure and markedly improves the chance of success.

A good approach is to split the experiment into a 3- or 4-day procedure, depending on the availability of an automated thermal cycler. Late in the day on day 1, perform the first-strand synthesis and ligation reactions, allowing the ligation incubation to proceed overnight. On day 2, stop the ligation reactions and do the exponential amplification and labeling. Allow at least one day to load and run the sequencing gel. If the PCR reactions are carried out in an automated thermal cycler, a sequencing gel can be prepared while the amplification is performed. This gel can be loaded with the first loading late on day 2, run overnight, loaded with the second loading on the following morning, and run through day 3. Alternatively, a gel can be prepared and loaded with the first loading on the morning of day 3, run all day, loaded with the second loading late in the day, and stopped on the next morning.

For DMS footprinting, on day one, the *in vivo* DMS treatment and DNA harvest from several sets of cultured cells can be completed in ≤ 2 hr including set-up time. The lysed-cell slurry is then incubated for 3 to 5 hr. At this point the lysates can be frozen at -20°C overnight. DNA purification by multiple organic extractions followed by precipitation is carried out on day 2 and will take from 3 to 6 hr depending on sample number. The DNA is allowed to resuspend overnight. On day 3, a second precipitation can be performed in ≤ 1 hr and the DNA allowed to resuspend at least overnight. On day 4, the concentration of resuspended DNA can be determined spectrophotometrically in ~ 1 hr. The *in vitro* DMS and piperidine treatments can be carried out up to the precipitation in step 26 in ~ 7 hr. The samples can be recovered from the precipitation in step 27, dried from water in step 28 and finally resuspended in step 29 in ~ 2 to 3 hr. The DNA is allowed to redissolve overnight, and the concentration determined spectrophotometrically in ~ 1 hr on day 5. The DNA is now ready for LMPCR.

Literature Cited

Becker, P.D. and Schutz, G. 1988. Genomic footprinting. *In Genetic Engineering Principles and Methods*, Vol. 10 (J.K. Setlow and A. Hollaender, eds.) pp. 1-19. Plenum, New York.

Church, G.M. and Gilbert, W. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 81:1991-1995.

Clark, J.M. 1988. Novel nontemplated nucleotide addition reactions catalyzed by prokaryotic and eukaryotic DNA polymerases. *Nucl. Acids Res.* 16:9677-9686.

Fors, L., Saavedra, R.A., and Hood, L. 1990. Cloning of the shark Po promoter using genomic walking technique based on the polymerase chain reaction. *Nucl. Acids Res.* 18:2793-2799.

Frohman, M.A., Dush, M.K., and Martin, G.R. 1988. Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. U.S.A.* 85:8998-9002.

Garrity, P.A. and Wold, B.J. 1992. Effects of different DNA polymerases in ligation-mediated PCR: Enhanced genomic sequencing and *in vivo* footprinting. *Proc. Natl. Acad. Sci. U.S.A.* 89:1021-1025.

Huibregtse, J.M. and Engelke, D.R. 1986. Direct identification of small sequence changes in chromosomal DNA. *Gene* 44:151-158.

Jackson, P.D. and Felsenfeld, G. 1985. A method for mapping intranuclear protein-DNA interactions and its applications to a nuclease hypersensitive site. *Proc. Natl. Acad. Sci. U.S.A.* 82:2296-2300.

Linz, U. 1990. Thermocycler temperature variation invalidates PCR results. *BioTechniques* 9:286-293.

Loh, E.Y., Elliott, J.F., Cwirla, S., Lanier, L.L., and Davis, M.M. 1989. Polymerase chain reaction with single-sided specificity: Analysis of T cell receptor δ chain. *Science* 243:217-220.

Mueller, P.R. and Wold, B. 1991. Ligation-mediated PCR: Applications to genomic footprinting. *Methods* 2:20-31.

Mueller, P.R., Salser, S.J., and Wold, B. 1988. Constitutive and metal inducible protein: DNA interactions at the mouse metallothionein-I promoter examined by *in vivo* and *in vitro* footprinting. *Genes & Dev.* 2:412-427.

Mueller, P.R., and Wold, B. 1989. *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* 246:780-786.

Ohara, O., Dorit, R.L., and Gilbert, W. 1989. One-sided polymerase chain reaction: The amplification of cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 86:5673-5677.

Pfeifer, G.P., Steigerwald, S.D., Mueller, P.R., Wold, B., and Riggs, A.D. 1989. Genomic sequencing and methylation analysis by ligation mediated PCR. *Science* 246:810-813.

Pfeifer, G.P., Tanguay, R.L., Steigerwald, S.D., and Riggs, A.D. 1990. *In vivo* footprint and methylation analysis by PCR-aided genomic sequencing: Comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1. *Genes & Dev.* 4:1277-1287.

Rideout III, W.M., Coetzee, G.A., Olumi, A.F., and Jones, P.A. 1990. 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science* 249:1288-1290.

- Rigaud, G., Roux, J., Pictet, R., and Grange, T. 1991. In vivo footprinting of Rat TAT Gene: Dynamic interplay between the glucocorticoid receptor and a liver-specific factor. *Cell* 67:977-986.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487-491.
- Saluz, H.P. and Jost, J.P. 1987. A Laboratory Guide to Genomic Sequencing: The Direct Sequencing of Native Uncloned DNA. Birkhauser, Boston.
- Tanguay, R.L., Pfeifer, G.P., and Riggs, A.D. 1990. PCR-aided DNase I footprinting of single copy gene sequences in permeabilized cells. *Nucl. Acids Res.* 18:5902.
- Wahl, G.M., Berger, S.L., and Kimmel, A.R. 1987. Molecular hybridization of immobilized nucleic acids: Theoretic concepts in practical considerations. *Methods Enzymol.* 152:399-407.
- White, T.J., Arnheim, N., and Erlich, H.A. 1989. The polymerase chain reaction. *Trends Genet.* 5:185-189.
-

Contributed by Paul R. Mueller and
Barbara Wold
California Institute of Technology
Pasadena, California

Paul A. Garrity
University of California, Los Angeles
Los Angeles, California

Molecular Cloning of PCR Products

UNIT 15.4

It is often desirable to clone PCR products to establish a permanent source of cloned DNA for hybridization studies, to obtain high-quality DNA sequencing results, or to separate products when PCR amplification yields a complex mixture. The efficiency of direct cloning of PCR products can be improved by generating suitable ends on the amplified fragments. In the first protocol presented (see Basic Protocol), the A residue added to the 3' end of DNA by *Taq* DNA polymerase serves as a one-base overhang that facilitates ligation when the complementary T base is added to the cloning vector. In the first alternate protocol (see Alternate Protocol 1), three bases that are part of a restriction endonuclease site are synthesized as part of the PCR primers, and are thus added to the 5' end of the DNA fragments during amplification. The one-base A overhang left by the *Taq* DNA polymerase is removed and the fragments are ligated into long concatemers, generating full restriction sites at the ligated junctions. Subsequent digestion renders these fragments suitable for subcloning into appropriate vectors. These protocols are not exclusive—i.e., a T-A overhang can be used for one end of a fragment and a synthetic restriction site can be on the other end. An alternate method for creating restriction sites is presented in UNIT 8.5, and other variations are discussed in the Commentary of this unit. In the second alternate protocol (see Alternate Protocol 2), uracil DNA glycosylase (UDG) is used to generate a PCR product with specific 3' overhangs and to clone this product into a specially prepared UDG cloning vector. An accompanying protocol (see Support Protocol) discusses the design and use of specific dU-containing primers both to construct one's own UDG cloning vector and to prepare a PCR product suitable for cloning in the vector.

GENERATION OF T-A OVERHANGS

Taq DNA polymerase normally adds a single nontemplated nucleotide (nearly always A) to the 3' ends of all duplex DNA strands, making direct blunt-ended cloning of PCR fragments inefficient. This extra A may be removed using Klenow fragment or T4 DNA polymerase (UNIT 3.5) followed by direct cloning of the resulting blunt-ended fragment (UNIT 3.16), but it is simpler and more efficient to use the extra A as a one-base overhang. In the presence of dTTP alone, *Taq* DNA polymerase will add a single T to the “blunt ends” of a vector, generating the necessary complementary one-base overhang. This procedure is summarized in Figure 15.4.1.

Materials

- 5 μ g vector DNA—e.g., pUC19 (UNIT 1.5) or M13mp18 (UNIT 1.14)
- TE buffer, pH 8.0 (APPENDIX 2)
- 10 \times PCR buffer (with optimized Mg²⁺ concentration, see UNIT 15.1)
- 5 mM dTTP (UNIT 3.4)
- 5 U/ μ l *Taq* DNA polymerase
- Target DNA

Additional reagents and equipment for restriction endonuclease digestion (UNIT 3.1), agarose gel electrophoresis (UNIT 2.5A), purification of DNA (UNIT 2.1A), PCR (UNIT 15.1), and subcloning DNA fragments (UNIT 3.16)

1. Digest 5 μ g vector DNA with a restriction endonuclease that yields blunt ends (UNIT 3.1). Check that digestion is complete by electrophoresing 50 ng on an agarose minigel (UNIT 2.5A). Extract the DNA with phenol/chloroform and precipitate with ethanol (UNIT 2.1A). Microcentrifuge 5 min at top speed, 4°C, and resuspend pellet in 25 μ l TE buffer, pH 8.0.

In the example in Figure 15.4.1, EcoRV is used to generate the blunt ends of the vector.

**BASIC
PROTOCOL**

**The Polymerase
Chain Reaction**

15.4.1

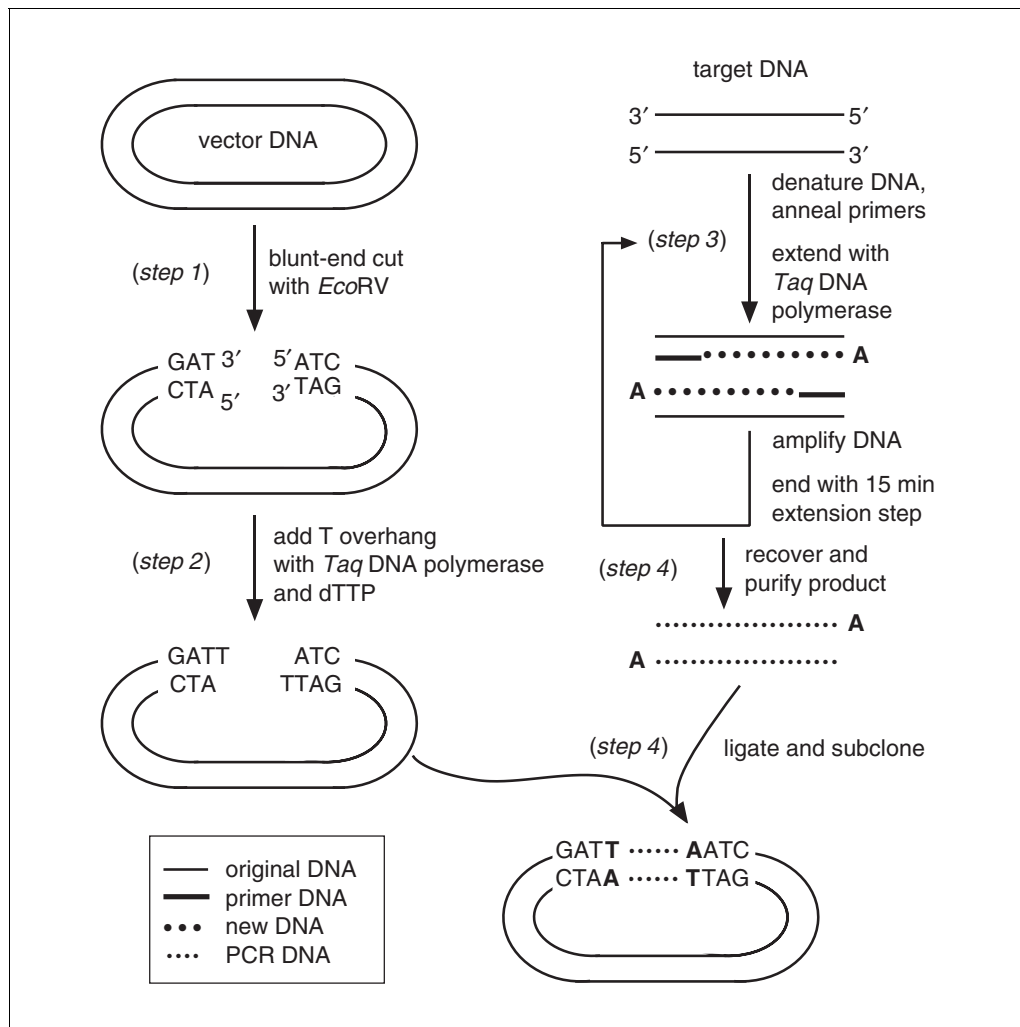


Figure 15.4.1 T-A overhang cloning. A single T is added to the 3' ends of blunt-cut vector DNA using *Taq* DNA polymerase and dTTP. A normal PCR reaction produces a fragment with a single A added to each 3' end. The As and Ts form complementary 1-base overhangs, facilitating ligation of insert to vector.

2. Set up the following T-addition reaction:

5 μ g blunt-ended vector DNA
 10 μ l 10 \times PCR amplification buffer
 20 μ l 5mM dTTP
 1 μ l (5 U) *Taq* DNA polymerase
 H₂O to 100 μ l final.

Incubate 2 hr at 75°C.

For this reaction, the Mg²⁺ should be 2 to 5 mM.

This yields sufficient T-tailed vector for many cloning procedures.

3. Carry out PCR (UNIT 15.1) to amplify the desired target DNA under optimized conditions (UNIT 15.1), ending the reaction with an extra 5- to 15-min extension at 70° to 75°C to be sure all fragments are A-tailed.

The number of amplification cycles needed can vary from 15 to 40, depending on the input DNA. Overcycling can result in amplification artifacts.

4. Recover and subclone the PCR fragments, beginning with step 6 of the Basic Protocol in *UNIT 3.16*.

The first time any preparation of T-tailed vector is used for cloning, a full set of controls should be performed as described in the Commentary.

GENERATION OF HALF-SITES

A second strategy for generating “sticky ends” on PCR products is to add to the 5′ end of each primer three bases, carefully chosen to create a full restriction site when the PCR products are concatemerized by blunt-end ligation. This procedure is summarized in Figure 15.4.2.

Additional Materials (also see Basic Protocol)

T4 DNA ligase (*UNIT 3.14*) and 10× buffer (*UNIT 3.4*)

1 mM ATP (*UNIT 3.4*)

2 mM 4dNTP mix (*UNIT 15.1*)

Klenow fragment of *E. coli* DNA polymerase I (*UNIT 3.5*)

T4 polynucleotide kinase (*UNIT 3.10*)

Additional reagents and equipment for oligonucleotide synthesis (*UNIT 2.11*) and quantitation of DNA (*APPENDIX 3D*)

1. Design and synthesize a pair of oligonucleotide primers with the three 3′ nucleotides of a palindromic six-base restriction site joined to the 5′ end of each specific primer (*UNIT 2.11*).

In the example in Figure 15.4.2, the last three bases of a BamHI site—TCC—are added to the 5′ end of the primers.

2. Anneal primers to desired DNA sequences and carry out PCR under optimized conditions (*UNIT 15.1*). Analyze a 5- μ l aliquot by agarose gel electrophoresis to verify amplification and quantitate yield (*UNIT 2.5A*). Remove the mineral oil with a pipettor. Extract DNA with phenol/chloroform and precipitate with ethanol (*UNIT 2.1A*).

Optimization of amplification conditions may be necessary because of the decreased specificity inherent in the specially designed primers (see Commentary and UNIT 15.1).

3. Resuspend the DNA in TE buffer, pH 8.0, at a concentration of 0.2 μ g/ μ l. Set up the following ligation reaction:

5 μ l (1 μ g) DNA

1 μ l 10× T4 DNA ligase buffer

1 μ l 1 mM ATP

1 μ l 4dNTP mix

1 U Klenow fragment

1 U T4 polynucleotide kinase

500 U (cohesive end) T4 DNA ligase

H₂O to 10 μ l.

Incubate \geq 2 hr at 15°C. Analyze a 1- μ l aliquot by agarose gel electrophoresis to verify ligation (*UNIT 2.5A*).

For convenience, the incubation may be extended up to 24 hr.

Klenow fragment removes the one-base overhang left by the Taq DNA polymerase, T4 polynucleotide kinase phosphorylates the ends, and T4 DNA ligase joins the DNA into long concatemers, generating restriction sites from the half-sites at the ends (see Fig. 15.4.2).

4. Subclone fragments into appropriate vectors, beginning with step 1 of the Basic Protocol in *UNIT 3.16*.

**ALTERNATE
PROTOCOL 1**

**The Polymerase
Chain Reaction**

15.4.3

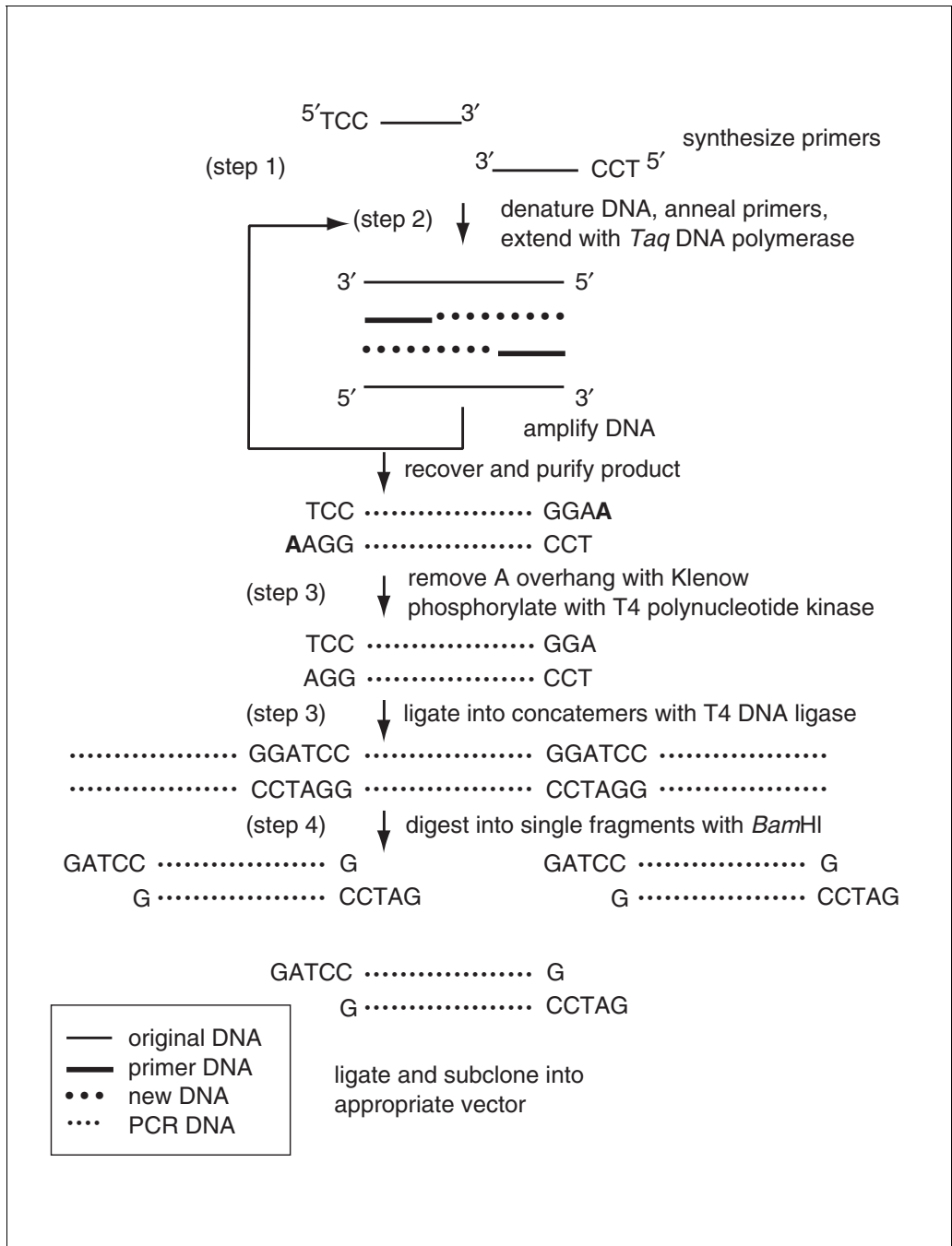


Figure 15.4.2 Restriction enzyme half-sites. Fragments are generated by a normal PCR using oligonucleotide primers containing the three 3' nucleotides of a 6-base recognition site at their 5' ends. In a single step, the nontemplated A is removed by Klenow fragment, and the fragments are phosphorylated at their 5' ends by T4 polynucleotide kinase and ligated into concatemers. The ligation generates the 6-base recognition sequence, which can then be cleaved with the appropriate restriction enzyme. The resulting fragments can be easily cloned.

CLONING PCR PRODUCTS WITH URACIL DNA GLYCOSYLASE

Cloning of PCR products with uracil DNA glycosylase (UDG) may be utilized to clone any PCR-generated DNA rapidly and efficiently. This method requires synthesis of oligonucleotide primers containing twelve specific deoxynucleotides at the 5' end. Four of the additional twelve nucleotides are dU, and these are distributed evenly throughout the 12-base sequence. (e.g., 5'-CUACUACUACUA-template specific sequence-3'). These primer sequences are incorporated into the PCR product during amplification; the four dU residues provide substrates for UDG at the 5' ends of the PCR product. Excision of dU residues by UDG results in the generation of 3' overhangs on PCR products that are complementary to sequences at the ends of UDG cloning vectors (Fig. 15.4.3). Annealing of PCR products to the vector after UDG treatment yields circular molecules that can transform competent *E. coli* cells. A number of UDG cloning vectors that allow unidirectional as well as bidirectional cloning of PCR-amplified DNA have been described and are commercially available (see Table 15.4.1; Nisson et al., 1991; Rashtchian et al., 1992a). Steps for constructing one's own UDG cloning vector and for preparing PCR products suitable for UDG cloning are described elsewhere (see Support Protocol).

Cloning with UDG is a rapid and efficient method for cloning PCR-amplified DNA. Advantages of the UDG-cloning method are: (1) ease of cloning—no restriction endonuclease digestions or ligations are required and the reaction is completed in 30 min; (2) efficiency of cloning— 1×10^7 clones/ μg DNA are routinely obtained; (3) the PCR product is used directly for cloning without the need for purification from primers and nucleotides or for buffer exchange; and (4) vector background is extremely low and virtually all transformants contain an insert.

Additional Materials (also see *Basic Protocol*)

PCR product amplified using dU-containing primers (see Support Protocol)
25 ng/ μl linearized UDG cloning vector (see Support Protocol)
5 \times PCR buffer (with optimized Mg^{2+} concentration; see *UNIT 15.1*)
1 U/ μl uracil DNA glycosylase (UDG)

Additional reagents and equipment for introduction of DNA into cells (*UNIT 1.8*), screening transformants (*UNIT 1.4*), mapping by restriction endonuclease digestion (*UNIT 3.2 & 3.3*)

1. Set up reaction to generate single-stranded 3'-overhang ends and anneal PCR product and vector by adding the following to a microcentrifuge tube:

1 to 2 μl (100 ng) PCR product amplified using dU-containing primers
2 μl 25 $\mu\text{g}/\text{ml}$ UDG cloning vector
1 μl 10 \times PCR buffer
1 μl 1 U/ μl UDG
H₂O to 10 μl final.

Mix and incubate 30 min at 37°C.

The 12-base sticky ends obviate the requirement for ligase.

2. Transform 100 μl competent *E. coli* cells with 5 μl of the annealing reaction (from step 1; *UNIT 1.8*) and plate on appropriate medium for screening and selection (e.g., LB/ampicillin plates containing Xgal and IPTG). Incubate 16 hr at 37°C.
3. Pick individual transformants and analyze for the presence of insert by screening (e.g., on Xgal; *UNIT 1.4*), mapping by restriction endonuclease digestion (*UNIT 3.2 & 3.3*), or PCR (*UNIT 15.1*).

DESIGNING PRIMER SETS FOR AMPLIFICATION AND CONSTRUCTION OF UDG CLONING VECTORS

The following Support Protocol is designed to enable the amplification of PCR products using dU-containing primers. Although UDG cloning vectors are commercially available (Life Technologies) it may be desirable to generate one's own. In theory, virtually any vector of the appropriate size can be adapted for UDG cloning with careful consideration and design of complementary primer sets for the vector and for the sequence to be cloned.

Preparation of a cloning vector for use in UDG cloning can be achieved in one of two ways: either by PCR amplification of the vector sequence using dU-containing primers, or by ligation of the dU-containing sequence to the ends of a linearized vector. In the first approach outlined in the steps below, the design of a primer set to be used to amplify a vector generally should include a 12-base sequence at the 5' end of the primer that contains at least four dU residues distributed evenly throughout the sequence (for example, see Table 15.4.1 and Figure 15.4.3), followed by another 12- to 16-base sequence complementary to the sequences flanking the site of insertion (cloning site) of the vector. In this approach, the pair of dU-containing primers is combined with vector DNA, linearized at the site of insertion, to amplify the entire vector sequence. This method has been used successfully to amplify the entire sequence of pUC19: dU-containing tails were incorporated into the primer set (5'-UAGUAGUAGUAGACCATCGTCGACCTGCAG-3' and 5'-UAGUAGUAGUAGACCATCGGATCCCCGGGT-3') to create a UDG cloning vector (for details, see Nisson et al., 1991; Table 15.4.1).

Alternatively, the UDG cloning vector can be prepared by ligating specific adaptor sequences encoding 12-base dU-containing sequences to the linearized vector DNA, using the methods described in *UNIT 3.16*. In this case, the specific adaptor oligonucleotides must be designed to accommodate the restriction endonuclease site(s) used for cloning inserts into the vector (for details, see Rashtchian et al., 1992a).

In either case, directional cloning can be facilitated by incorporating discrete or nonidentical dU-containing sequences at the 5' end of each primer. Table 15.4.1 lists a series of commercially available vectors that were constructed from the commonly used cloning vectors pUC18, pUC19, and pSPORT1 using these approaches.

Table 15.7.1 Characteristics of Various UDG Cloning Vectors

Plasmid	Cloning	Cloning sequence used	Remarks
pAMP1	Directional	Upstream Primer (sense, or 5'): 5'-CAUCAUCAUCAU...-3' Downstream Primer (antisense, or 3'): 5'-CUACUACUACUA...-3'	Derived from pSport 1, T7 and SP6 promoters, F1 <i>ori</i> , Amp ^r
pAMP10	Non-directional	Upstream Primer (sense, or 5'): 5'-CUACUACUACUA...-3' Downstream Primer (antisense, or 3'): 5'-CUACUACUACUA...-3'	Derived from pSPORT 1, T7 and SP6 promoters, F1 <i>ori</i> , Amp ^r
pAMP18	Directional	Upstream Primer (sense, or 5'): 5'-ACGCGUACUAGU...-3' Downstream Primer (antisense, or 3'): 5'-AUGGAGAUCUCU...-3'	Derived from pUC18
pAMP19	Directional	Upstream Primer (sense, or 5'): 5'-AUGGAGAUCUCU...-3' Downstream Primer (antisense, or 3'): 5'-ACGCGUACUAGU...-3'	Derived from pUC19

For UDG cloning vectors constructed using either of these approaches, the primer sequences incorporated at the ends of the vector must be used to design a complementary primer set for PCR amplification of the DNA sequence to be cloned. This secondary primer set will encode a 5' 12-base dU-containing sequence that is complementary to the corresponding primer sequence used to amplify the vector, followed by sequences complementary to the ends of the fragment to be amplified. Amplification of a DNA fragment or a clone containing the fragment of interest using the secondary primers results in a PCR product suitable for use with the appropriate vector in UDG cloning (see Alternate Protocol 2).

Materials

Vector DNA linearized at appropriate cloning or insertion site
DNA sequence to be amplified

Additional reagents and equipment for synthesis of oligonucleotides (UNIT 2.11),
PCR (UNIT 15.1), and agarose gel electrophoresis (UNIT 2.5A)

1. Design and synthesize a primary oligonucleotide primer pair (UNITS 2.11 & 15.1).

Each primer should begin with a 5' tail of 12 bases containing at least 4 dU residues distributed evenly throughout, followed by 12 to 16 bases complementary to one end of the vector sequence to be amplified (one primer for each end; see Fig. 15.4.3 and Table 15.4.1 for examples).

2. Design and synthesize a secondary oligonucleotide primer pair (UNITS 2.11 & 15.1).

The 5' 12-base sequence of each of these primers should be complementary to that of one of the primary primers synthesized in step 1, and should be followed by 12 to 16 bases complementary to one end of the DNA sequence to be amplified.

3. Anneal the primary oligonucleotide primers to linearized vector DNA and carry out PCR under optimized conditions (see UNIT 15.1).

Amplify 1 ng linearized vector DNA.

4. In a separate tube, anneal the secondary oligonucleotide primers to the desired insert DNA and carry out PCR under optimized conditions using the hot-start protocol (see UNIT 15.1).

5. Analyze the PCR products by agarose gel electrophoresis (UNIT 2.5A) to verify amplification and quantitate yield.

6. Clone amplified DNA into the prepared vector as described in Alternate Protocol 2.

COMMENTARY

Background Information

T-A cloning

The surprising finding that *Taq* DNA polymerase adds a single nontemplated nucleotide to the 3' end of double-stranded DNA (Clark, 1988) explains why direct blunt-end cloning of PCR fragments is inefficient. At the same time, it affords a new means of direct cloning. The nontemplated nucleotide is almost always A, so a complementary vector with 3' T overhangs allows the fragments to be cloned unmodified. An added advantage is that inefficient blunt-end ligations are avoided.

Several groups have constructed special vectors for T-A overhang cloning that produce single T 3' overhangs when cut with *XcmI* or *HphI* (e.g., Kovalic et al., 1991). Such vectors, which are commercially available, can be used in the Basic Protocol (starting at step 3), but they do not offer the generality that tailing any blunt-end cut vector can offer. When no dATP is present, *Taq* DNA polymerase will add any of the other bases to the end of a strand (Clark, 1988). Thus *Taq* DNA polymerase can be used to add a single T to a blunt site, allowing PCR fragments to be cloned into any blunt site in any vector (Marchuk et al., 1991). The Basic Protocol is not limited, of course, to PCR frag-

ments, and can be used to transform most blunt-end ligations into sticky-end ligations.

Generation of restriction sites

There are three ways in which restriction sites can be used for cloning PCR fragments. The first and simplest way is to employ natural sites within the target DNA. Depending on the target sequence and the intended application, such natural sites can often be found in suitable positions. The other two ways, detailed below, involve using PCR to change the DNA sequence so as to introduce a new restriction endonuclease site. Any method that changes DNA sequence involves unpaired bases in a primer oligonucleotide and therefore may decrease the specificity of the PCR reaction (see Critical Parameters). Cryptic sites require one to two unpaired bases; half-sites use three, and full sites need eight to ten.

In random DNA, a site for a particular enzyme with a 6-base recognition site is found on average every 4 kb. But a cryptic site with a $5/6$ -base match to the recognition site is found every 170 bases. Some computer DNA-analysis packages (e.g., the Map program, GCG package; *UNIT 7.7*) can find cryptic sites automatically. Because the polylinker in a typical cloning vector has ten or more 6-base sites that generate overhangs, useful cryptic sites can be found frequently. Thus, even in special cases (e.g., when the encoded amino acid sequence cannot be changed), it may well be possible to take advantage of cryptic sites by using site-directed mutagenesis of PCR products. The procedure is based on the protocols in *UNIT 8.5* for introducing point mutations, except that the PCR product is digested and subcloned directly. In designing oligonucleotide primers to change cryptic sites to real sites, three unchanged bases must be left at the 3' end, and at least three bases must be left 5' to the restriction site (see Critical Parameters). The resulting fragments can be easily cloned by following the procedures in *UNIT 3.16*.

Finally, restriction sites may be added directly to the 5' ends of oligonucleotide primers during synthesis. A technique for adding full restriction sites to the ends of fragments is discussed in *UNIT 8.5*. In the protocol for half-site generation (see Alternate Protocol 1), half of a restriction site is added to the 5' end of each primer to be used in the amplification reaction. When the reaction is complete, the products are made blunt-ended by removing the 3' A overhang with Klenow fragment, then kinased and concatemerized by ligation; all of these reac-

tions can be performed conveniently in a single buffer. Ligation generates the complete restriction site at each junction (Kaufman and Evans, 1990). The concatemers can then be cut, yielding the full-length fragment with overhangs.

Advantages of the half-site method include lower expense as well as better quality of shorter oligonucleotides and a smaller loss of specificity caused by extra bases in the primers (see Critical Parameters). The major disadvantage is loss of efficiency caused by the blunt-end ligation step. Advantages of the full-site method include the possibility of directional cloning (using a different restriction site at each end) and the flexibility to use each primer independently; a disadvantage is the inherent loss of specificity that may be prohibitive in critical applications.

Cloning PCR products with UDG

Rapid cloning of PCR products is facilitated by the use of PCR primers containing dU residues and using uracil DNA glycosylase (UDG; Nisson et al., 1991). The incorporation of dU-containing primers into the ends of PCR products and into the ends of the cloning vector results in the selective placement of dUMP residues in place of dTMP at the 5' ends of these molecules. UDG is an important part of the uracil excision pathway in many organisms; it cleaves the N-glycosidic bond between the deoxyribose moiety and uracil in single- and double-stranded DNA. The cleavage reaction results in the formation of abasic sites that destabilize base pairing in double-stranded DNA. Thus, treatment of dU-containing DNA with UDG disrupts base pairing at these sites and generates 3' overhanging ends (Fig. 15.4.3). The plasmid vector is amplified with primers whose ends are complementary to the sequences incorporated into the ends of the product to be cloned. Because the vector has identical tails (noncomplementary with respect to one another), the probability of circularization of vector without insert is extremely unlikely, as is suggested by the extremely low background of nonrecombinant transformants. Furthermore, the use of long (12-base) ends permits formation of relatively stable duplexes between these cohesive tails on insert and vector, thus eliminating the need for an *in vitro* ligation step (Aslandis and de Jong, 1990). Subsequent annealing of the 3' protruding termini of the PCR products to the complementary 3' ends of the vector results in products that can be conveniently transformed directly with high efficiency and without ligation or purification.

This method allows joining of PCR-amplified DNA fragments at virtually any point in the DNA without the need for restriction endonuclease sites. A number of UDG cloning vectors using a variety of dU-containing sequences have been developed recently (Table 15.4.1; Buchman et al., 1992, 1993).

UDG cloning has been utilized for creating novel gene constructs (Booth et al., 1994) and for site-directed mutagenesis (Rashtchian et al., 1992b; Owen et al., 1994). These applications are accomplished simply by designing two overlapping primers in which dU residues have been substituted for T residues in the overlap region. The overlap regions in PCR products amplified with such primers can specifically anneal to each other after excision of dU residues with UDG and be cloned into an appropriate UDG cloning vector. This methodology has provided a convenient and rapid approach for ordered assembly of various PCR fragments to construct recombinant genes (Rashtchian, 1995).

Critical Parameters

An extensive discussion of critical parameters for cloning DNA fragments—with examples of various cloning applications—is found in *UNIT 3.16*. Critical parameters for PCR are discussed in *UNIT 15.1*.

A point to emphasize here is that the specificity of an oligonucleotide primer depends on the number of base pairs needed to form a stable duplex between the primer and the target at the annealing temperature used. When extra bases are added to a primer, they may contribute to the necessary base pairing and thus cause the primer to hybridize in unintended places. In noncritical applications, such as amplification from cloned DNA (*UNIT 8.5*), loss of specificity is inconsequential and bases may be added with impunity. However, in stringent applications such as the use of very short or degenerate primers on genomic DNA or cDNA, extra bases should be added only with great caution. Higher annealing temperatures, up to 65° or 70°C, may help compensate for decreased specificity.

The UDG protocol in this unit requires that all reaction components except *Taq* DNA polymerase be heated 5 min at 94°C before adding *Taq* DNA polymerase; this “hot start” procedure improves the productivity of the reaction.

The secondary PCR primers illustrated in Figure 15.4.3 each contain twelve nucleotides (CAUCAUCAU) at the 5′ end that allow the PCR products to be cloned using uracil DNA glycosylase (Alternate Protocol 2). Only

portions of secondary PCR primers hybridize specifically to the template (for example, 5′-CTGAGGAGTGAATTGGTTCG-3′; and 5′-GTGAACTGCACTGTGACAAGCTGC-3′).

The primers can also be modified to allow PCR products to be cloned by other methods—e.g., by adding nucleotides at 5′ ends to install specific restriction endonuclease sites for cloning.

The significant error rate of *Taq* DNA polymerase should also be kept in mind (see *UNIT 15.1*). It is a good idea to sequence all clones derived from PCR.

Troubleshooting

The vast majority of DNA amplification failures can be attributed to oligonucleotides of incorrect sequence or poor quality. Oligonucleotide sequence can be incorrect because of errors in the target sequence (even in published sequences), in transcription, or in synthesis; the synthesis can fail, yielding no full-length product; or oligos can degrade with time. In cases of persistent failure, oligos should be checked on a gel (*UNIT 2.12*) to see if they are full length, and resynthesized if necessary. Because the percentage of full-length product decreases exponentially with the length of the oligo (*UNIT 2.11*), primers should routinely be made as short as is practical.

For T-A overhang cloning, it is essential that the nontemplated T be added to a high fraction of vector ends; otherwise, the vector will recircularize readily, giving a background of clones without inserts. This underscores the need for transfection controls as detailed in Critical Parameters of *UNIT 3.16*.

For the Basic Protocol and Alternate Protocol 1, vectors encoding the α fragment of *lacZ* are preferred, so that blue/white screening can be done if necessary (*UNIT 1.4*). Each new batch of vector should be tested by setting up parallel ligations containing: (1) gel-purified blunt-ended vector; (2) gel-purified T-tailed vector; and (3) gel-purified T-tailed vector and gel-purified PCR product. Some of each ligation mix should be transformed into an appropriate host strain and plated for blue/white screening. Ligation 1 should yield many colonies, most of them blue. Ligation 2 should yield a much smaller number of colonies, mostly white. The insert in ligation 3 should give a large boost in the number of colonies, which are usually white. A large number of blue colonies from ligation 2 indicates incomplete tailing. If this occurs, the tailing reaction should be repeated using fresh dTTP.

Although PCR products with unphosphorylated ends can be T-A cloned, yield may be improved by starting with 5' phosphorylated oligonucleotide primers (UNIT 3.10) or including 1 U of T4 polynucleotide kinase in the ligation reaction. It is unnecessary, even counterproductive, to dephosphorylate the ends of the vector—the T tails prevent self-ligation.

For UDG cloning, the success of the cloning procedure can be assessed by setting up three sets of parallel transformations containing (1) amplified, UDG-treated insert alone, (2) amplified, UDG-treated vector alone, and (3) amplified UDG-treated mixture of vector and insert. In the first two cases, the background should be extremely low or nonexistent. The third case should yield a substantial number of transformants at a frequency of 10^5 to 10^6 recombinants per microgram DNA, which occur predominantly as white colonies if the *lacZ α* -complementation screening is used.

Additional troubleshooting procedures for cloning and PCR amplification are found in UNITS 3.16 & 15.1.

Anticipated Results

A single small-scale (20 μ l) PCR reaction should yield many clones, most of which should have inserts. For example, in UDG cloning (without screening) approximately 90% to 95% of transformants will carry the appropriate insert. The efficiency of transformation depends on the method used (see UNIT 3.16, Commentary, for guidelines).

Time Considerations

Preparation of vector and insert can be completed in one day except for the half-site method which requires an extra ligation reaction often performed overnight. Ligation of vector to insert is also often performed overnight, and growth of bacterial colonies takes 10 to 12 hr. Until bacterial transformation, the procedure can be safely paused at the end of any step by freezing the sample at -20°C .

Literature Cited

Aslandis, C. and de Jong, P.J. 1990. Ligation-independent cloning of PCR products (LIC-PCR). *Nucl. Acids Res.* 18:6069-6074.

Booth, P.M., Buchman, G.W., and Rashtchian, A. 1994. Assembly and coding sequences for

neurotrophic factors directly from genomic DNA using polymerase chain reaction and uracil DNA glycosylase. *Gene* 146:303-308.

- Buchman, G.W., Schuster, D.M., and Rashtchian, A. 1992. Rapid and efficient cloning of PCR products using the CloneAmp system. *Focus* 14:41-45.
- Buchman, G.W., Booth, P., and Rashtchian, A. 1993. The CloneAmp pUC18 system: Cloning of the brain-derived neurotrophic factor gene. *Focus* 15:36-41.
- Clark, J.M. 1988. Novel nontemplated nucleotide addition reactions catalyzed by prokaryotic and eukaryotic DNA polymerases. *Nucl. Acids Res.* 16:9677-9686.
- Kaufman, D.L. and Evans, G.A. 1990. Restriction endonuclease cleavage at the termini of PCR products. *BioTechniques* 9:304-306.
- Kovalic, D., Kwak, J.-H., and Weisblum, B. 1991. General method for direct cloning of DNA fragments generated by the polymerase chain reaction. *Nucl. Acids Res.* 19:4650.
- Marchuk, D., Drumm, M., Saulino, A., and Collins, F.S. 1991. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucl. Acids Res.* 19:1154.
- Nisson, P.E., Rashtchian, A., and Watkins, P.C. 1991. Rapid and efficient cloning of Alu-PCR products using uracil DNA glycosylase. *PCR Methods Appl.* 1:120-123.
- Owen, J.L., Hay, C., Schuster, D.M., and Rashtchian, A. 1994. A highly efficient method of site-directed mutagenesis using PCR and UDG cloning. *Focus* 16:39-44.
- Rashtchian, A. 1995. Novel methods for cloning and engineering genes using the polymerase chain reaction. *Curr. Opin. Biotechnol.* 6:30-36.
- Rashtchian, A., Buchman, G., Schuster, D., and Berninger, M. 1992a. Uracil DNA glycosylase-mediated cloning of PCR-amplified DNA: Application to genomic and cDNA cloning. *Anal. Biochem.* 200:91-97.
- Rashtchian, A., Thornton, C.G., and Heidecker, G. 1992b. A novel method for site-directed mutagenesis using PCR and uracil DNA glycosylase. *PCR Meth. Appl.* 2:124-130.

Contributed by Michael Finney
MJ Research
Watertown, Massachusetts

Paul E. Nisson and Ayoub Rashtchian
(UDG cloning)
Life Technologies, Inc.
Gaithersburg, Maryland

Enzymatic Amplification of RNA by PCR (RT-PCR)

UNIT 15.5

This unit describes methods for enzymatic amplification of RNA by the polymerase chain reaction (PCR). The essential method (see Basic Protocol) is especially useful for rare RNAs because all steps (annealing, reverse transcription, and amplification) are performed under optimal conditions, thereby maximizing efficiency and recovery. These are also important considerations when amplifying heterogeneous RNA populations or large RNAs.

In the Basic Protocol, an oligonucleotide primer is coprecipitated with the RNA to maximize the efficiency of their annealing to each other. Following annealing, cDNA is synthesized using reverse transcriptase (RT). Enzymatic amplification of this cDNA is then performed by PCR. Alternate Protocols 1 and 2 provide modifications to save time and effort when the optimal conditions in the Basic Protocol are not required. The Support Protocol describes an easy method for making crude RNA preparations suitable for amplification by PCR.

Investigators may also wish to consider the use of commercially available kits for performing RT-PCR reactions. Several are available which offer advantages of convenience and speed, although at some sacrifice in terms of cost and one's ability to customize the reaction conditions for specific target RNAs.

PCR AMPLIFICATION OF RNA UNDER OPTIMAL CONDITIONS

**BASIC
PROTOCOL**

Materials

- Poly(A)⁺ (UNIT 4.5) or crude RNA (see Support Protocol)
- 25 µg/ml cDNA primer in H₂O
- 3 M sodium acetate, pH 5.5 (APPENDIX 2)
- 100% and 70% ethanol
- 400 mM Tris·Cl, pH 8.3 (APPENDIX 2)
- 400 mM KCl (APPENDIX 2)
- Reverse transcriptase buffer (see recipe)
- 32 U/µl AMV reverse transcriptase (UNIT 3.7; Boehringer-Mannheim)
- 10 mM Tris·Cl/10 mM EDTA, pH 7.5
- Phenol buffered with 10 mM Tris·Cl/10 mM EDTA, pH 7.5 (store at room temperature)
- 24:1 chloroform/isoamyl alcohol (UNIT 2.1A)
- ~150 µg/ml amplification primers in H₂O (~20 µM each)
- 5 mM 4dNTP mix (5 mM each dNTP in H₂O; UNIT 3.4)
- 10× PCR amplification buffer (UNIT 15.1)
- Taq DNA polymerase (UNIT 15.1)
- Mineral oil
- Additional reagents and equipment for PCR, (UNIT 15.1), preparation of poly (A)⁺ RNA (UNIT 4.5), total RNA (UNITS 4.2-4.4), or cytoplasmic RNA (UNIT 4.1), and for agarose or nondenaturing polyacrylamide gel electrophoresis (UNITS 2.5A & 2.7)

NOTE: Reagents and solutions used for RNA procedures should be prepared using standard methods for handling RNA (UNIT 4.1).

**The Polymerase
Chain Reaction**

15.5.1

Anneal the primer to the RNA

1. Coprecipitate RNA and cDNA primer by adding the following ingredients to a microcentrifuge tube:

2 μg poly(A)⁺ RNA
25 ng (3 pmol) cDNA primer
H₂O to 90 μl .

Mix and add 10 μl of 3 M sodium acetate, pH 5.5, and 200 μl of 100% ethanol. Mix and incubate overnight at -20°C or 15 min at -70°C .

This coprecipitation step maximizes the efficiency of annealing. The template source can also be total RNA or crude RNA (see Support Protocol). Less than 2 μg can be used, depending upon the abundance of the RNA within the sample. The cDNA primer is frequently the same as one of the amplification primers (see commentary).

2. Microcentrifuge 15 min at high speed, 4°C , and discard supernatant.
3. Add 200 μl of 70% ethanol and mix gently by inversion. Microcentrifuge 5 min at high speed, room temperature, and discard supernatant. Dry pellet briefly in a desiccator under vacuum or in a Speedvac evaporator.

Do not overdry the pellet or the RNA will not resuspend well.

4. Add the following to the RNA pellet:

12 μl H₂O
4 μl 400 mM Tris·Cl, pH 8.3
4 μl 400 mM KCl.

Heat to 90°C , then cool slowly to 67°C .

This step removes any base-paired regions in the RNA and then allows the RNA and primer to begin to anneal. The best way to heat and cool the sample is to place the microcentrifuge tube in a 400-ml beaker of 90°C water and allow it to cool to 67°C on the benchtop.

5. Microcentrifuge the sample 1 sec to collect the condensate at the bottom of the tube. Incubate 3 hr at 52°C . Microcentrifuge 1 sec to collect condensate.

Perform the centrifugation in this step quickly so the sample does not cool too much. This final annealing temperature can be adjusted according to the base composition of the primer, and increased or decreased depending on the specificity of annealing required (UNIT 6.4).

Synthesize the cDNA

6. Add 29 μl reverse transcriptase buffer and 0.5 μl (16 U) AMV reverse transcriptase. Mix and incubate 1 hr at 42°C .

The temperature can be adjusted from 37° to 55°C if desired. Lower temperatures may be helpful for primers with high A-T base compositions. Higher temperatures may lead to more full-length products by lessening RNA secondary structure.

7. Add 150 μl 10 mM Tris·Cl/10 mM EDTA, pH 7.5, and mix. Add 200 μl buffered phenol and vortex briefly. Microcentrifuge 5 min at high speed, room temperature, and save upper aqueous phase.
8. Add 200 μl of 24:1 chloroform/isoamyl alcohol and vortex briefly. Microcentrifuge 5 min at high speed, room temperature, and save upper aqueous phase.
9. Add 20 μl 3 M sodium acetate, pH 5.5, and 500 μl of 100% ethanol. Mix and precipitate overnight at -20°C or 15 min at -70°C .
10. Microcentrifuge 15 min at high speed, 4°C , and discard supernatant.
11. Dry pellet briefly and resuspend in 40 μl water.

Amplify the cDNA by PCR

12. Mix the following:

- 5 μ l cDNA from step 11
- 5 μ l each amplification primer (~150 μ g/ml or ~20 μ M each)
- 4 μ l 5 mM 4dNTP mix
- 10 μ l 10 \times amplification buffer
- 70.5 μ l H₂O.

Heat 2 min at 94°C. Microcentrifuge 1 sec to collect condensate.

Usually one of the amplification primers is the same as the cDNA primer. If a different amplification primer is used, the cDNA primer should be removed from the cDNA reaction using methods described in UNIT 15.2.

13. Add 0.5 μ l (2.5 U) *Taq* DNA polymerase, mix, and microcentrifuge 1 sec. Overlay with 100 μ l mineral oil.

14. Set up the following automated amplification cycles:

39 cycles:	2 min	55°C
	2 min	72°C
	1 min	94°C
1 cycle:	2 min	55°C
	7 min	72°C.

The number of cycles can be varied depending upon the abundance of the RNA. Forty total cycles is sufficient for rare mRNAs in 2 μ g poly(A)⁺ RNA but more cycles may be necessary if smaller amounts of template are used.

15. Analyze products by electrophoresis in agarose or nondenaturing polyacrylamide gels (UNIT 2.5A & 2.7), choosing a concentration of the gel matrix appropriate for the expected size of the amplification products.

AVOIDING LENGTHY COPRECIPIATION AND ANNEALING STEPS

This Alternate Protocol saves considerable time compared to the Basic Protocol because it eliminates the relatively lengthy coprecipitation and annealing steps. Instead, a quick heating and cooling step is used. Although this reduces the specificity and quantitative efficiency of annealing, the losses incurred are frequently insignificant and should not affect most applications.

Replace steps 1 to 5 of the Basic Protocol with the following:

1. Mix the following:

- 2 μ g poly(A)⁺ RNA
- 25 ng (3 pmol) cDNA primer
- H₂O to 21 μ l.

2. Heat 3 to 15 min at 65°C.

3. Cool on ice and add to reverse transcriptase buffer (see Basic Protocol, step 6).

ALTERNATE PROTOCOL 1

**ALTERNATE
PROTOCOL 2**

INTRODUCING cDNA DIRECTLY INTO THE AMPLIFICATION STEP

This Alternate Protocol saves time by eliminating the extraction and precipitation steps following cDNA synthesis in the Basic Protocol. The products of the cDNA reaction are diluted sufficiently to avoid interference when added to the enzymatic amplification step. This procedure may not be appropriate when the anticipated number of template molecules is small.

Replace steps 7 to 11 of the Basic Protocol with the following:

1. Add 450 μ l of 10 mM Tris·Cl/10 mM EDTA, pH 7.5, to 50 μ l of the reaction mix from step 6 of the Basic Protocol.
2. Add 1 to 5 μ l of the mixture from step 1 of this protocol to the amplification reaction in step 12 of the Basic Protocol.

**SUPPORT
PROTOCOL**

RAPID PREPARATION OF CRUDE RNA

This Support Protocol avoids the need for purifying RNA prior to enzymatic amplification. It is especially useful when many samples are to be analyzed, or when the amount of starting material is small because some of the losses associated with poly(A)⁺ RNA purification are avoided (Higuchi, 1989).

Materials

Phosphate-buffered saline (PBS; *APPENDIX 2*)
Diethylpyrocarbonate (DEPC) solution (see recipe)

CAUTION: DEPC is volatile and toxic and should be handled with precautions. DEPC is also rapidly inactivated in aqueous buffers and in the presence of Tris·Cl, so work quickly.

1. Suspend 10⁶ to 10⁷ cells (or equivalent amount of dispersed tissue) in PBS. Microcentrifuge 5 min and discard supernatant.
2. Resuspend cell pellet in 200 to 400 μ l DEPC solution and vortex briefly.
3. Microcentrifuge nuclei 10 sec and transfer supernatant to a new tube. Incubate 20 min at 37°C, then 10 min at 90°C.
4. Microcentrifuge 5 min to remove precipitate and transfer supernatant to a new tube. Use 5 to 10 μ l directly in step 1 of the Basic Protocol or ethanol precipitate and resuspend (*UNIT 2.1A*). Store frozen.

REAGENTS AND SOLUTIONS

DEPC solution

Dilute diethylpyrocarbonate (DEPC) 1:10 in 100% ethanol. *Immediately* before use, further dilute 1:10 DEPC 1:1000 in lysis buffer (see recipe).

Lysis buffer

140 mM NaCl
10 mM Tris·Cl, pH 8.0 (*APPENDIX 2*)
15 mM MgCl₂
0.5% Nonidet P-40 (NP-40)
Store at room temperature

Reverse transcriptase buffer

Per sample:

2.5 μ l 400 mM Tris·Cl, pH 8.3 (*APPENDIX 2*)

2.5 μ l 400 mM KCl

1 μ l 300 mM MgCl₂

5 μ l 100 mM DTT

5 μ l 5 mM 4dNTP mix

2 μ l 2 mg/ml actinomycin D

11 μ l H₂O

Optional: 10 U RNasin (Promega; reduce volume of H₂O accordingly)

Prepare fresh before each reverse transcription from separate stocks. DTT, dNTP, and actinomycin D stocks should be stored frozen; all other stocks can be stored at room temperature.

COMMENTARY

Background Information

Enzymatic amplification by PCR is well suited for RNA analyses because quantities available for study are frequently limited and often insufficient for more standard methods of analysis. Enzymatic amplification can be used to prepare large amounts of cDNA for subsequent steps, or as an analytical tool to map the structure of RNAs.

RNAs offer additional possibilities relative to DNAs that can readily be exploited: they are relatively small molecules, with discrete 5' and 3' ends, and in many cases the sequence of one or more of the ends is defined. The most obvious example of a defined end is the polyadenylated tail on the 3' end of most (but not all) eukaryotic messenger RNAs. These features can be incorporated in the design of oligonucleotide primers.

Although most commonly used to synthesize quantitative amounts of specific double-stranded cDNAs, amplification of RNA can be utilized as a rapid analytical tool. Both 3' and 5' ends of RNAs can be mapped using internal primers specific for a given transcript, in conjunction with primers specific for 3' ends [oligo(dT)] or 5' ends (common leaders or added primer sites; *UNIT 15.6*; Kapler et al., 1990).

Critical Parameters and Troubleshooting

As discussed in *UNIT 15.1*, several parameters affecting the efficiency and specificity of the amplification process require empirical evaluation. Among these are magnesium concentration, concentration of primers, cyclic annealing temperature, and the number of cycles. Generally, lower primer concentration, fewer cycles, and increased temperature will contribute to

increased specificity.

To confirm amplification of the expected molecules, several methods can be used, such as cleavage with restriction endonucleases or hybridization with an internal oligonucleotide or fragment (*UNIT 15.1*). Hybridization analysis will also increase the sensitivity, allowing detection of amplification products not visualized by ethidium bromide staining.

In planning RNA amplifications, the following points should be considered: (1) method of preparation of the template RNA; (2) design of the specific oligonucleotide primers; (3) synthesis of the first strand of cDNA using the appropriate primer; and (4) enzymatic amplification.

Protocols for many of these steps are discussed in detail in other units as indicated, while parameters specifically relevant to enzymatic amplification of RNA are discussed below.

Source of RNA

Depending upon available quantities, total RNA, cytoplasmic RNA, or poly(A)⁺ RNA can be used. However, because even single molecules can be enzymatically amplified efficiently, dilute crude cellular preparations containing denatured proteins and free RNA can be utilized, as in the Support Protocol.

Design of oligonucleotide primers

Specific oligonucleotide pairs are designed and utilized with many of the same considerations discussed for enzymatic amplification of DNA (*UNIT 15.1*). A point worth emphasizing here is that increased specificity can be gained by using an additional primer internal to that used for cDNA synthesis for PCR. Moreover, reamplification with nested primers can some-

times rescue an RNA amplification that seems to have failed. With RNA, also consider the primers discussed below.

In certain viruses or organisms, many mRNAs contain a common 5' leader sequence. The most extreme example of this is found in trypanosomatid protozoans in which every mRNA contains a common 39-nucleotide leader sequence. These leader sequences can be utilized for priming (Kapler et al., 1990).

Oligo(dT) and enzymatically added primer sites are discussed in *UNIT 15.6*.

Discrimination between RNA and DNA

Given the proclivity of enzymatic amplification to reveal rare molecules, one must consider whether the amplified products arise from RNA or DNA templates. In some cases it is possible to choose primers so that specific amplification products cannot arise from DNA. For example, primers from different exons will yield products of different sizes if DNA or mRNA is used as a template. In other cases, it will not be possible to select such discriminatory primers. Enzymatic treatment of samples with RNase-free DNase, followed by phenol extraction and ethanol precipitation, may be helpful (*UNIT 4.1*).

Anticipated Results

These procedures should allow specific amplification of a defined segment of RNA into double-stranded DNA. The products obtained should be suitable for hybridization analysis, DNA sequencing, or molecular cloning as desired.

Time Considerations

For the Basic Protocol, annealing of the primers requires ~4 hr, cDNA synthesis requires ~3 hr, and enzymatic amplification by

PCR requires ~4 hr. However, the PCR requires little effort from the investigator, because once the reaction mix is assembled, the amplification steps are usually performed by a machine. Alternate Protocol 1 reduces the time required for annealing to 20 min, and the second alternate protocol reduces the time required for cDNA synthesis to 1.5 hr. For all protocols, it is feasible to perform the annealing and extension in a single day, followed by automated amplification overnight.

Literature Cited

- Berchtold, M.W. 1989. A simple method for direct cloning and sequencing cDNA by the use of a single specific oligonucleotide and oligo(dT) in a polymerase chain reaction (PCR). *Nucl. Acids Res.* 17: 453.
- Frohman, M.A., Dush, M.K., and Martin, G.R. 1988. Rapid production of full length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. U.S.A.* 85: 8998-9002.
- Higuchi, R. 1989. Simple and rapid preparation of samples for PCR. In *PCR Technology: Principles and Applications for DNA Amplification* (H.A. Ehrlich, ed.) pp. 31-38. Stockton Press, New York.
- Kapler, G.M., Coburn, C.M., and Beverley, S.M. 1990. Transfection of the human parasite *Leishmania* delineates a 30 kb region sufficient for extra-chromosomal replication and expression. *Mol. Cell. Biol.* 10:1084-1094.

Key Reference

Frohman et al., 1988. See above.

Outlines several strategies for using PCR to analyze RNA structure and for production of full-length cDNAs.

Contributed by Stephen M. Beverley
Harvard Medical School
Boston, Massachusetts

cDNA Amplification Using One-Sided (Anchored) PCR

Conventional approaches to the characterization and sequencing of mRNA generally rely on the preparation and screening of a cDNA library, and on the subsequent isolation of single cDNA clones from within the library (UNITS 5.5, 5.6 & 5.8A). While this methodology has proven useful overall, success depends critically on the quality of the cDNA library. Frequently, cDNA libraries are vitiated by prematurely terminated cDNA molecules that do not reflect full-length message. In addition, rare messages may frequently be lost altogether in the preparation of cDNA libraries.

The polymerase chain reaction (PCR) offers an opportunity to directly select, amplify, and isolate a message of interest. This protocol presents a modification of PCR, called anchored PCR, that allows amplification of full-length mRNA when only a small amount of sequence information is available. Unlike the amplification of RNA or cDNA by conventional PCR (UNIT 15.4)—which requires prior knowledge of the sequences flanking the region of interest to design the PCR primers—anchored PCR can be employed when only a small region of sequence lying within the mRNA is known in advance.

In theory, only a single specific primer is required to perform anchored PCR. In practice, amplification of a single product requires a second round of PCR amplification, using a second sequence-specific primer. This scheme is depicted in Figures 15.6.1 and 15.6.2. Both the original and reamplifications use an oligo(dT) primer complementary either to the poly(A) tail of the mature mRNA [when amplifying downstream (3') to the known sequence] or to an enzymatically synthesized homopolymer tail added to the cDNA following first strand synthesis [when amplifying upstream (5') to the known sequence]. The two rounds of PCR amplification result in a single product that can be sequenced directly or cloned into an appropriate vector for further analysis.

This unit presents two basic protocols for the amplification of cDNA targets using anchored PCR. The first basic protocol can be used to amplify targets where the downstream (3') sequences are unknown. The second basic protocol permits amplification of target molecules whose upstream (5') sequences are unknown.

AMPLIFICATION OF REGIONS DOWNSTREAM (3') OF KNOWN SEQUENCE

BASIC PROTOCOL

This method can be used to characterize messages of interest from any source of poly(A)⁺ RNA prepared by conventional protocols (UNITS 4.1-4.5). Although the amount of poly(A)⁺ RNA required will depend on the genome complexity of the organism being studied and on the relative abundance of the targeted mRNA, in general 100 to 300 ng of total poly(A)⁺ RNA prepared from vertebrate tissue is sufficient for this procedure. Three PCR primers are required: a 20-mer oligo(dT) primer and two sequence-specific primers, one for the original amplification and the second for the reamplification (see Fig. 15.6.1). The internal sequence-specific primer can be immediately adjacent (3') to or can partially overlap primer 1 at the 3' end.

Materials

- Source of RNA (UNITS 4.1-4.5)
- 5× Moloney murine leukemia virus (MoMLV) reverse transcriptase buffer
- 5 µg/µl bovine serum albumin (BSA)
- 10 mM 4dNTP mix (10 mM each dNTP in TE buffer, pH 7.5, stored at -20°C)
- 500 ng/µl actinomycin D
- 200 U/µl MoMLV reverse transcriptase (Table 3.4.1 & UNIT 3.7)

The Polymerase Chain Reaction

15.6.1

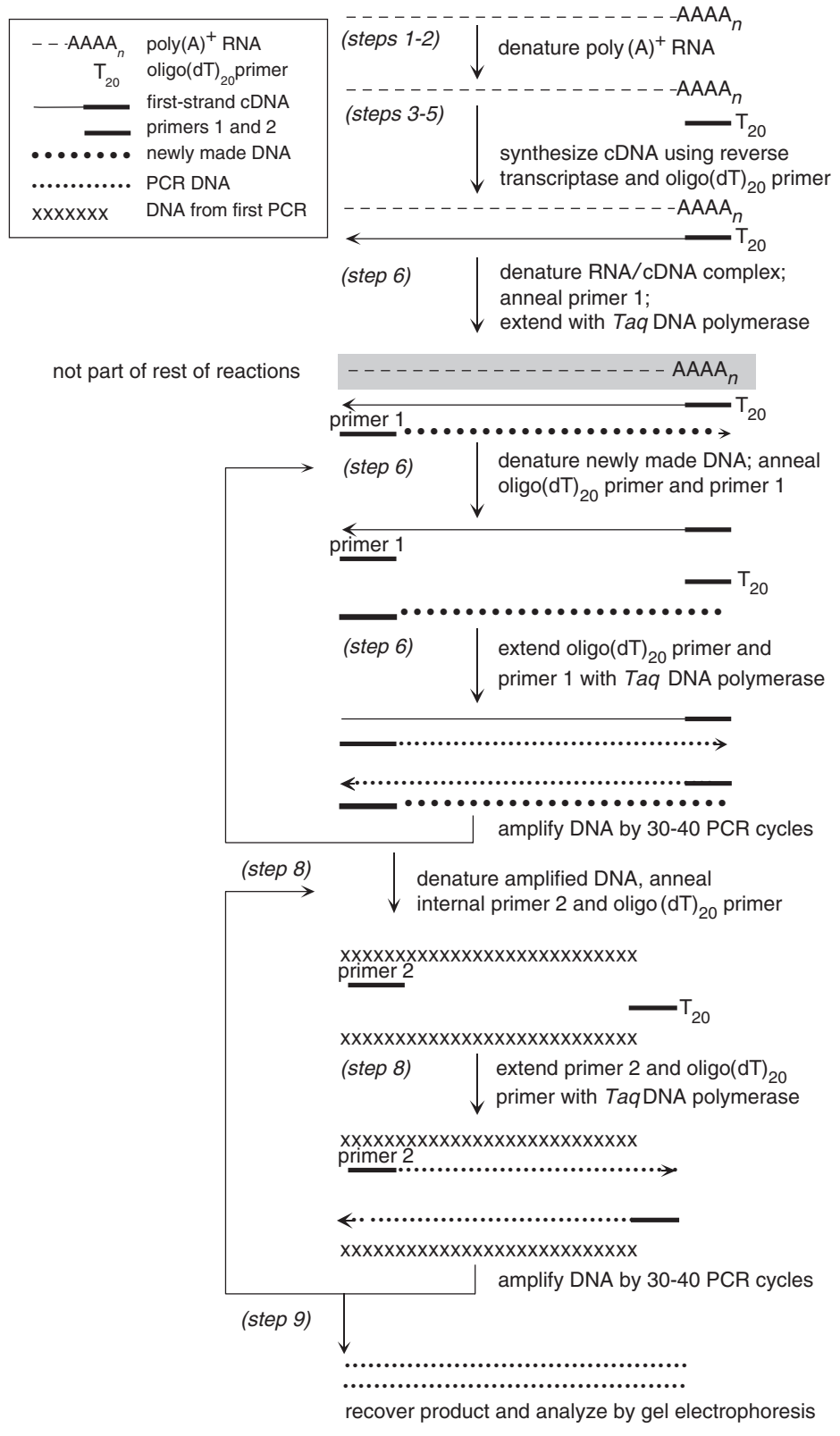


Figure 15.6.1 Flow chart of downstream (3') anchored PCR protocol (see text for details). The steps correspond to those listed in the basic protocol. Note that primer 2 can be immediately adjacent (3') to, or can partially overlap with, primer 1.

cDNA Amplification Using One-Sided (Anchored) PCR

15.6.2

15 pmol/μl (100 ng/μl) oligo(dT)₂₀ primer
100 pmol/μl each of sequence-specific primers 1 and 2 (see Fig. 15.6.1)
TE buffer (APPENDIX 2)
2.5 mM 4dNTP mix (2.5 mM each dNTP in TE buffer, pH 7.5, stored at -20°C)
10× amplification buffer
2.5 U/μl *Taq* DNA polymerase (UNITS 15.1 & 3.5)
Mineral oil

Additional reagents and equipment for preparation of poly(A)⁺ RNA (UNIT 4.1),
PCR amplification (UNIT 15.1), agarose gel electrophoresis (UNIT 2.5), and
Southern blotting and hybridization (UNITS 2.9 & 6.4)

1. Prepare 100 ng poly(A)⁺ RNA at a concentration of ≥100 ng/μl and place in a 1.5-ml microcentrifuge tube.

This high concentration is necessary because the final volume of the reverse transcriptase reaction should be kept as low as possible, preferably ≤10 μl. Larger amounts of poly(A)⁺ RNA (up to 1 μg) can be used if readily available. Amounts below 100 ng may not include the full complement of rare mRNAs.

2. Incubate poly(A)⁺ RNA 2 min at 65°C. Microcentrifuge briefly at room temperature and place immediately on ice.

This incubation melts the secondary structure of the mRNA, removing hairpins and loops that interfere with the synthesis of cDNA.

3. In a clean 1.5-ml microcentrifuge tube, prepare the following reverse transcriptase reaction mix on ice (10 μl total):

2 μl 5× MoMLV reverse transcriptase buffer
1 μl (5 μg) BSA
1 μl poly(A)⁺ RNA (from step 2)
1 μl 10 mM 4dNTP mix
1 μl (500 ng) actinomycin D
1 μl (15 pmol) oligo(dT)₂₀ primer
2 μl sterile H₂O
1 μl MoMLV reverse transcriptase.

4. Mix gently, microcentrifuge briefly at room temperature and incubate 1 hr at 37°C.

This results in synthesis of cDNA.

5. Add 40 μl TE buffer to bring the total reaction volume up to 50 μl.

6. Prepare the following PCR reaction mix on ice (100 μl final):

1 μl cDNA template (from step 5)
10 μl 10× amplification buffer
1 μl (100 pmol) oligo(dT)₂₀ primer
1 μl (100 pmol) sequence-specific primer 1
6 μl 2.5 mM 4dNTP mix
80 μl sterile H₂O
1 μl (2.5 U) *Taq* DNA polymerase.

Overlay reaction mixture with mineral oil. Carry out 30 to 40 cycles of amplification (UNIT 15.1).

Initially, amplifications should be carried out in the presence of 1.5 mM MgCl₂ (final concentration), with annealing temperatures not exceeding 42°C. These conditions can be subsequently modified and optimized for the specific template used (UNIT 15.1). Primer

concentrations can be reduced to 30 pmol of each primer without reducing the efficiency of amplification.

7. Analyze an aliquot by agarose gel electrophoresis.

The result of this first PCR is usually a smear around the expected size range, because specificity is seldom conferred by a single sequence-specific primer. The presence of the desired product can be confirmed by hybridizing a Southern blot (UNITS 2.9 & 6.4) of the gel containing the PCR products to the sequence-specific primer 2 internal to the amplified sequence.

8. Remove a 1- μ l aliquot of the amplification product from step 6 to serve as the template for reamplification. Carry out a second round of PCR as described in step 6, using 40 to 100 pmol each of the oligo(dT)₂₀ primer and the internal sequence-specific primer 2.

Although 100 pmol of each primer are usually added in conventional PCR amplifications, lower concentrations of primer do not appear to reduce amplification efficiency or yield, and may reduce the cost of this procedure.

9. Analyze an aliquot of the second amplification by agarose gel electrophoresis. The amplified product should now appear as a single band on an agarose gel.

10. If desired, characterize the amplification product by cloning into an appropriate vector (UNIT 3.16) or by direct sequencing (UNIT 15.2).

**BASIC
PROTOCOL**

AMPLIFICATION OF REGIONS UPSTREAM (5') OF KNOWN SEQUENCE

In contrast to the first protocol, this procedure uses one of the sequence-specific primers to initiate the synthesis of the cDNA strand. This cDNA is modified by the addition of a poly(A) tail. PCR amplifications—mediated by two sequence-specific primers and a 20-mer oligo(dT) primer complementary to the newly synthesized tail—yield the desired unique product (Fig. 15.6.2).

Materials

- Source of RNA (UNITS 4.1-4.5)
- 100 pmol/ μ l sequence-specific primers 3 and 4 (Fig. 15.6.2)
- 15 pmol/ μ l (~100 ng/ μ l) oligo(dT)₂₀ primer
- 1 M NaCl
- 200 mM Tris·Cl, pH 7.5 (APPENDIX 2)
- 25 mM EDTA
- 100% and 70% ethanol, ice-cold
- 5 \times Moloney murine leukemia virus (MoMLV) reverse transcriptase buffer
- 5 μ g/ μ l bovine serum albumin (BSA)
- 10 mM 4dNTP mix (10 mM each dNTP in TE buffer, pH 7.5, stored at -20°C)
- 500 ng/ μ l actinomycin D
- 200 U/ μ l MoMLV reverse transcriptase (Table 3.4.1 & UNIT 3.7)
- 3 M sodium acetate (APPENDIX 2)
- TE buffer, pH 7.5 (APPENDIX 2)
- 5 \times terminal deoxynucleotidyltransferase (TdT) buffer
- 15 mM CoCl₂
- 1 mM dATP (UNIT 3.4)
- Terminal transferase (Table 3.4.1 & UNIT 3.6)
- 40°C water bath

Additional reagents and equipment for preparation of poly(A)⁺ RNA (UNIT 4.5), phenol extraction (UNIT 2.1), and PCR amplification (UNIT 15.1)

**cDNA
Amplification
Using One-Sided
(Anchored) PCR**

15.6.4

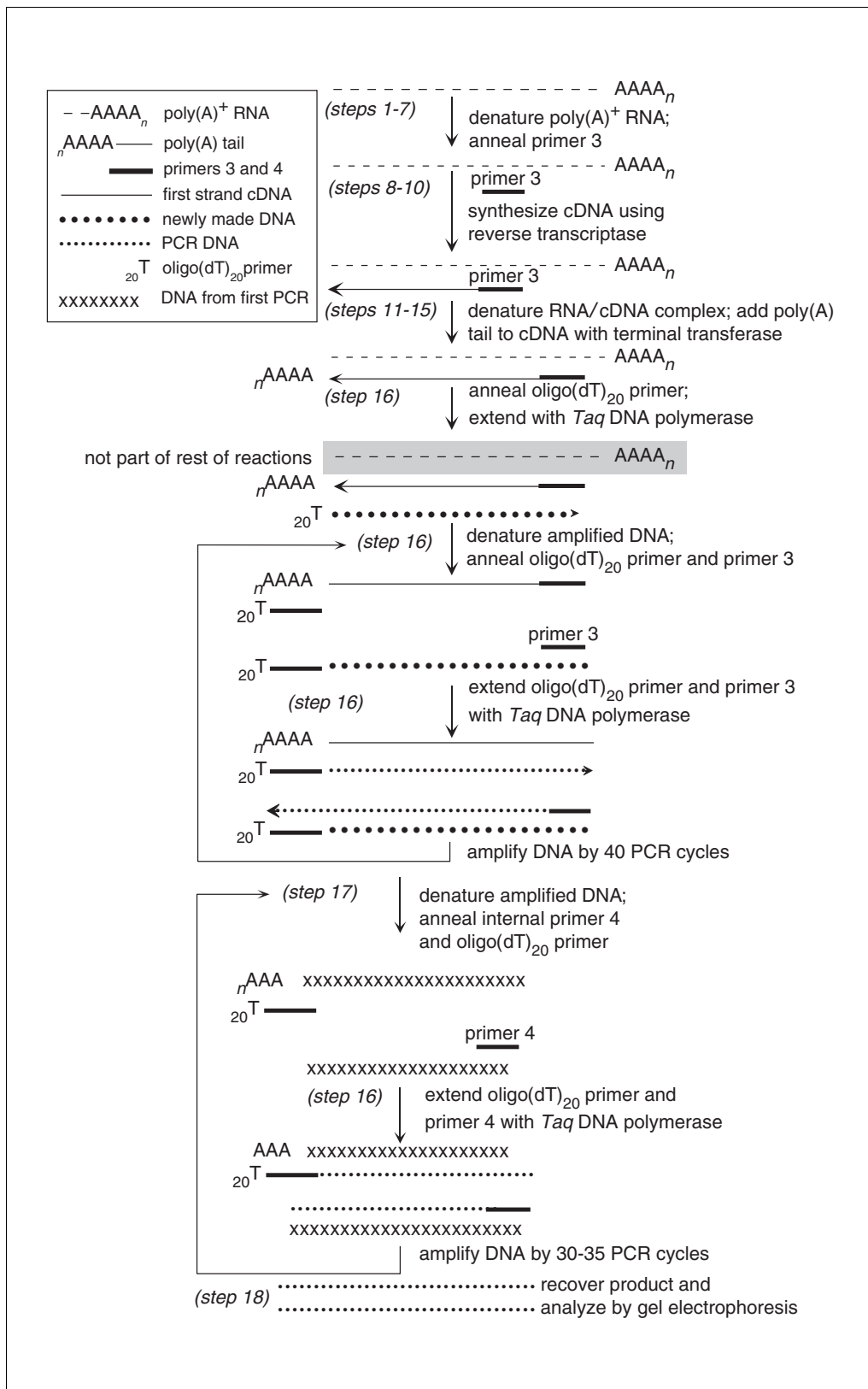


Figure 15.6.2 Flow chart of upstream (5') anchored PCR protocol (see text for details). The steps correspond to those listed in the protocol. Note that primer 4 can be immediately adjacent to, or can partially overlap with, primer 3.

Prepare RNA and anneal sequence-specific primer

1. Prepare 100 ng of poly(A)⁺ RNA at a concentration of 100 ng/μl.
2. Prepare a 5-μl annealing mix containing:
 - 1 μl (1 fmol) sequence-specific primer 3
 - 1 μl 1 M NaCl
 - 1 μl 200 mM Tris·Cl, pH 7.5
 - 1 μl 25 mM EDTA
 - 1 μl (100 ng) poly(A)⁺ RNA (from step 1).

Amount of primer to use for sequence-specific cDNA synthesis depends in part on the abundance of target mRNA. As a rule, a 5- to 10-fold molar excess of oligomer primer relative to target template yields the best results. Note that small amount of primer boosts specificity of targeted cDNA synthesis. Primer 3 should be diluted with water or 0.1× TE buffer, pH 7.5, to reach the desired concentration.

3. Incubate annealing mix 3 min at 65°C, microcentrifuge briefly at room temperature, and place immediately on ice for 2 min.

This incubation melts the secondary structure of the mRNA, removing hairpins and loops that interfere with the synthesis of cDNA. “Snap-cooling” on ice prevents the secondary structure from re-forming.

4. Incubate 3 to 4 hr at 40°C.

The long incubation time and low primer concentration increase the specificity of the primer-template annealing.

5. Add 15 μl ice-cold 100% ethanol, place in dry ice/ethanol bath for 10 min, and microcentrifuge 10 min at high speed, 4°C.
6. Decant supernatant. Add 50 μl ice-cold 70% ethanol and gently invert tube several times to rinse and desalt the pellet (do not vortex). Microcentrifuge 2 min at room temperature, remove supernatant, and dry briefly under vacuum.
7. Resuspend pellet in 10 μl water.

Synthesize first-strand cDNA

8. On ice, prepare reverse transcriptase mix containing (25 μl total):
 - 5 μl 5× MoMLV reverse transcriptase buffer
 - 2.5 μl (12.5 μg) BSA
 - 2.5 μl 10 mM 4dNTP mix
 - 2.5 μl (1.25 μg) actinomycin D
 - 10 μl annealed primer/template (from step 7)
 - 1.5 μl sterile H₂O
 - 1 μl (200 U) MoMLV reverse transcriptase.

Incubate 1 hr at 37°C.

9. Phenol extract the sample. Carefully transfer supernatant to new microcentrifuge tube. Add 2.5 μl of 3 M sodium acetate (0.3 M final) and 75 μl ice-cold 100% ethanol. Place 5 min in dry ice/ethanol bath, then microcentrifuge 20 min at high speed, 4°C. Discard supernatant.

In cases where small amounts of poly(A)⁺ RNA are being used, 1 to 5 μg carrier tRNA (UNIT 4.6) may be added at this step to facilitate the precipitation.

10. Resuspend pellet in 25 μl of TE buffer, pH 7.5. Add 2.5 μl of 3 M sodium acetate (0.3 M final) and repeat the ethanol precipitation as described in step 9, discarding supernatant after spin.

11. Add 100 μ l ice-cold 70% ethanol to pellet and rinse by gently inverting tube several times (do not vortex). Microcentrifuge 5 min at high speed, room temperature; discard supernatant, and dry pellet briefly under vacuum.
12. Resuspend pellet in 5 μ l water and boil 2 min. Microcentrifuge briefly at room temperature and place immediately on ice.

This step denatures the cDNA/RNA hybrid molecules prior to tailing the cDNA molecule with terminal transferase.

Add poly(A) tail to cDNA

13. On ice prepare terminal transferase mix in order listed below (10 μ l final):

2 μ l 5 \times TdT buffer
1 μ l 15 mM CoCl₂ (1.5 mM final)
1 μ l 1 mM dATP (100 μ mol final)
5 μ l cDNA mix (from step 12)
1 μ l (25 U) terminal transferase.

Incubate 30 min at 37°C.

14. Inactivate enzyme by heating 2 min at 65°C.
15. Add 1 μ l of 3 M sodium acetate. Ethanol precipitate, wash and dry pellet as described in step 10.

Amplify by PCR and analyze results

16. Resuspend pellet in 10 μ l water. Carry out 40 amplification cycles as in step 6 of the first basic protocol, using 40 to 100 pmol each of sequence-specific primer 3 and oligo(dT)₂₀.

Although 100 pmol of each primer are usually added in conventional PCR amplifications, lower concentrations of primer do not appear to reduce amplification efficiency or yield, and may reduce the cost of this procedure.

Initially, amplifications should be carried out in the presence of 1.5 mM MgCl₂ (final), with annealing temperatures not exceeding 42°C. These conditions can be subsequently modified and optimized for the specific template used (UNIT 15.1).

The products of this round of PCR will not be visualized by ethidium bromide staining of an agarose gel, but can be seen when a Southern blot of the gel is hybridized to sequence-specific primer 4 internal to the amplified region.

17. Remove a 1- μ l aliquot of the amplification product of step 16 to serve as template for a new round of PCR amplification. Carry out 35 cycles of PCR as in step 6 of the first basic protocol, using 40 to 100 pmol each of primer 4 and the oligo(dT) primer.

Primer 4 can be immediately adjacent to, or partially overlapping with, primer 3. Because the anchored PCR procedures are carried out independently in the 3' and 5' directions, sequence-specific primers 3 and 4 can be the complements of primers 1 and 2 used in the previous protocol (see Fig. 15.6.2).

18. Analyze a 10- μ l aliquot by agarose gel electrophoresis.

The amplified product, stretching from the target site of the sequence-specific primer 4 to the 5' end of the cDNA, should now appear as a single band on an agarose gel.

19. If desired, characterize the PCR product by cloning into an appropriate vector (UNIT 3.16) and/or by direct sequencing (UNIT 15.2).

REAGENTS AND SOLUTIONS

5× MoMLV reverse transcriptase buffer

250 mM Tris·Cl, pH 8.3
375 mM KCl
50 mM DTT
15 mM MgCl₂
Store at −20°C for ≤3 months

10× amplification buffer

500 mM KCl
100 mM Tris·Cl, pH 8.8
15 mM MgCl₂
30 mM DTT
1 mg/ml BSA
Store at −20°C for ≤3 months

5× terminal transferase (TdT) buffer

1 M potassium cacodylate
125 mM Tris·Cl, pH 7.4
1.25 μg/μl BSA
Store at −20°C for ≤6 weeks

COMMENTARY

Background Information

The advent of PCR has spawned a number of procedures that do not require full knowledge of the sequence being amplified (Frohman et al., 1988; Ochman et al., 1988; Loh et al., 1989; Mueller and Wold, 1989; Ohara et al., 1989; Tam et al., 1989), and that assume only a small region of sequence is known. In many of these methods, the PCR amplifications are anchored by a primer derived from the region of known sequence; the second primer is non-specific and is targeted to a general feature of the sequence, such as the poly(A) tail present on most mature mRNAs. Alternatively, the target sequence can be enzymatically modified to provide an annealing site for the second primer. Such modifications include the addition (via terminal transferase) of a homopolymeric tail (UNIT 3.6) or the ligation of a sequence “splint” at the end of the target molecule (UNIT 15.5).

The protocols presented here are specifically designed for the amplification of full-length cDNA when only a fragment of the cDNA sequence is known. The amplification into the unknown downstream (3′) region capitalizes on the poly(A) tail present on most mRNAs. It therefore will not succeed with poly(A)[−] mRNAs, such as those encoding histones. The region of interest is amplified by annealing the oligo(dT) primer to this poly(A) tail and the sequence-specific primer to the

appropriate known region of the cDNA.

Amplification into the unknown upstream (5′) region makes use of an artificial poly(A) tail that is enzymatically added to the 3′ end of an initial, specific cDNA product. In theory, then, upstream and downstream amplifications can succeed with any mRNA for which sufficient sequence to design a single sequence-specific oligonucleotide is known.

In practice, however, a single sequence-specific primer does not always confer sufficient selectivity in the amplification, and the primer may bind to inappropriate or partially complementary sequences, resulting in the amplification of more than one sequence. This problem is further compounded in situations where the only sequence available at the outset comes from the partial sequencing of the protein produced by the mRNA of interest, or from the identification of conserved regions in homologous mRNAs extracted from other organisms. Such cases require the use of degenerate sequence-specific primers, further reducing the specificity of the anchored PCR amplifications. The consequences of using degenerate primers will depend on the template being amplified, the degree of degeneracy, and the amplification conditions. The use of nondegenerate sequence-specific primers is strongly recommended for the protocols described here (Ohara et al., 1989).

Critical Parameters and Troubleshooting

As with any method for the analysis of mRNA, the success of these protocols depends on the quality and integrity of the mRNA being used. Appropriate precautions to ensure high yield of intact mRNA are detailed in Chapter 4 (see introduction to that chapter). Although these protocols can be performed using untreated microcentrifuge tubes and pipet tips, several steps involve single-stranded nucleic acids (mRNA, first-strand cDNA) and the use of silanized microcentrifuge tubes and pipet tips (APPENDIX 3) may be desirable.

It is important to keep in mind that the initial stages of anchored PCR will seldom result in a fully specific amplification, and the amplification products will most often appear as a smear when visualized on an agarose gel. The PCR products produced by this protocol should be monitored by Southern blot hybridization to a target-specific internal oligomer. Such monitoring confirms that the desired product is in fact being amplified in the initial phases of the protocols.

A number of steps may be taken to increase the target specificity of anchored PCR. As described in these protocols, a reamplification carried out using a second internal (nested) sequence-specific primer greatly reduces spurious amplification products. This second primer can lie immediately adjacent to the first specific primer or can partially overlap at the 3' end. Thus, 40 to 50 nucleotides of known sequence are sufficient to fully anchor this two-step procedure.

A second problem that may interfere with the production of a single, distinct PCR product stems from the length of either the naturally occurring or the enzymatically synthesized poly(A) tail. A long poly(A) tail provides a large number of potential pairing sites for the nonspecific oligo(dT) primer. The resultant PCR products may consequently be of variable length.

The difficulties that stem from the variable length of the homopolymeric tails can be partially resolved by synthesizing an oligo(dT) primer with an additional specific sequence at the 5' end, as is done in the race amplification of cDNA ends (RACE) protocol (Frohman et al., 1988). Subsequent reamplifications will then rely on a primer directed to this synthetic sequence. Alternatively, the oligo(dT) primer can be synthesized with a degenerate (A,G,C) 3'-most position, resulting in more precise an-

nealing at the poly(A) tail boundary of the mRNA.

Finally, specificity increases if the first-round amplification products are size-selected by extracting an aliquot of the main PCR band within the smear (or of the "correct" target band, previously determined by Southern hybridization with an internal specific oligomer). The selected products can serve as templates for a second round of PCR amplification. Thus, the first-round products are run out on a low gelling/melting agarose gel (UNIT 2.6), and the desired target is size selected by piercing the appropriate band on the gel with a Pasteur pipet. This small gel core sample can then be placed directly into a new PCR mix and reamplified without further purification.

The products of this second amplification may still appear as a smear when visualized by ethidium bromide staining of an agarose gel (see commentary, UNIT 15.1). This smear results from the multiple positions that the oligo(dT)₂₀ primer can occupy along the length of the poly(A) tail. In such cases, select a small aliquot of product at the upper range of the smear and reamplify with 20 to 25 PCR cycles using the oligo(dT)₂₀ and internal primers (primers 2 or 4).

We have found that certain DNA targets cannot be efficiently amplified using standard PCR conditions. In such cases, low-yield PCR products can be readily cloned into an appropriate vector (Holton and Graham, 1991; Marchuck et al., 1991) and subsequently screened by colony hybridization (UNIT 6.2) to isolate the appropriate clones. Given the relatively high error rate of *Taq* polymerase (UNITS 3.5 & 15.1), sequences obtained from cloned PCR fragments should be confirmed by isolating and sequencing multiple clones.

Time Considerations

With the mRNA already in hand, each of the basic protocols will require 1.5 to 2 days to complete. Upstream and downstream anchored PCR can be carried out synchronously.

Literature Cited

- Frohman, M.A., Dush, M.K., Martin, G.R. 1988. Rapid production of full length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Nat. Acad. Sci U.S.A.* 85:8998-9002.
- Holton, T.A. and Graham, M.W. 1991. A simple and direct method for direct cloning of PCR products using ddT-tailed vectors. *Nucl. Acids Res.* 19:1156.

- Loh, E.Y., Elliott, J.F., Cwirla, S., Lanier, L.L., and Davis, M.M. 1989. Polymerase chain reaction with single-sided specificity: Analysis of a T-cell receptor delta chain. *Science* 243:217-220.
- Marchuck, D., Drumm, M., Saulino, A., and Collins, F.S. 1991. Construction of T-vectors: A rapid and general system for direct cloning of unmodified PCR products. *Nucl. Acids Res.* 19:1154.
- Mueller, P.R. and Wold, B. 1989. In vivo footprinting of a muscle-specific enhancer by ligation-mediated PCR. *Science* 246: 780-786.
- Ochman, H., Gerber, A.S., and Hartl, D.L. 1988. Genetic applications of an inverse polymerase chain reaction. *Genetics* 120:621-623.
- Ohara, O., Dorit, R.L. and Gilbert, W. 1989. One-sided polymerase chain reaction: The amplification of cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 86:5673-5677.

- Tam, A.W., Smith, M.M., Fry, K.E., and Larrick, J.W. 1989. Construction of cDNA libraries from small numbers of cells using sequence independent primers. *Nucl. Acids Res.* 17: 1269.

Key Reference

Ohara et al., 1989. See above.

The protocols described here are used to isolate and characterize α -tropomyosin transcripts in common frog and zebra fish.

Contributed by Robert L. Dorit
Yale University
New Haven, Connecticut

Osamu Ohara
Shionogi Research Laboratories
Osaka, Japan

Quantitation of Rare DNAs by PCR

UNIT 15.7

**BASIC
PROTOCOL**

This unit presents a protocol that uses the polymerase chain reaction (PCR) to quantitate the numbers of a particular DNA sequence from 1 to 20,000 molecules per sample. In addition, it helps assess the presence of contaminating sequences, the bane of this kind of procedure.

The DNA of interest is prepared and its concentration is determined. A known amount of this DNA is then mixed with two sets of oligonucleotide primers, one set specific for the DNA of interest (e.g., a virus) and the other set specific for an internal control (e.g., a single-copy gene encoded by the host organism). The sequences between the primers are amplified, electrophoresed on a gel, transferred to a filter, and probed with oligonucleotides specific for each amplified product. The amounts of the amplified products from the DNA of interest can then be quantitated by comparison to the internal control. For simplicity, the protocol is written in terms of quantitating viral DNA molecules relative to host cellular sequences; however, it can be adapted readily for other applications.

NOTE: Use molecular biology-grade water solutions (DNase-free, RNase-free, nucleic acid free).

Materials

- Proteinase digestion buffer (see recipe)
- 20 mg/ml proteinase K (store at -20°C)
- Phenol buffered with 50 mM Tris-Cl/10 mM EDTA, pH 7.4 (store at room temperature)
- 24:1 chloroform/isoamyl alcohol (UNIT 2.1A)
- 10 M ammonium acetate (APPENDIX 2)
- Cold 100% ethanol
- 70% ethanol
- TE buffer, pH 7.5 (APPENDIX 2)
- Reaction mix cocktail (see recipe)
- Mineral oil
- 0.8 U/ μl *Taq* DNA polymerase (UNIT 15.1)
- Oligonucleotide primers for hybridization (see recipe)
- Screw-cap microcentrifuge tubes, autoclaved
- Microcapillary pipets *or* positive displacement pipets with disposable tips and plungers *or* micropipettors with barrier tips
- Additional reagents and equipment for PCR (UNIT 15.1), tissue sample preparation (UNIT 2.2), agarose and nondenaturing polyacrylamide gel electrophoresis (UNITS 2.5A & 2.7), ethidium bromide dot quantitation (UNIT 2.6), Southern blotting or electroblotting (UNIT 2.9A), labeling oligonucleotides (UNITS 3.10, 4.6, & 4.8), hybridizing blots with oligonucleotides (UNITS 2.9 & 6.4), UV cross-linking DNA to filters (UNIT 2.9), and autoradiography (APPENDIX 3A)

NOTE: Use sterile, distilled water to prepare all reagents. Do NOT use diethylpyrocarbonate (DEPC) to treat reagents. To avoid contamination with unwanted nucleic acids, prepare reagents and solutions solely for use in this protocol (see Critical Parameters and Troubleshooting). Wear disposable gloves and change them frequently.

Prepare the DNA

1. Place cells or tissue sample in a screw-cap microcentrifuge tube. Add ~ 100 μl proteinase digestion buffer per $\sim 2 \times 10^6$ cells and 20 mg/ml proteinase K to 100 $\mu\text{g}/\text{ml}$. Incubate sample overnight at 50°C .

Contributed by Donald M. Coen

Current Protocols in Molecular Biology (2001) 15.7.1-15.7.8

Copyright © 2001 by John Wiley & Sons, Inc.

**The Polymerase
Chain Reaction**

15.7.1

Supplement 56

This is a modification of the protocol in UNIT 2.2. It is best to process samples in the order of increasing likelihood of their containing the sequences of interest. Perform the extractions away from where products of the amplification reactions (PCR products) or large quantities of plasmid DNA are handled. Wear disposable gloves and change them frequently.

Always include several negative control samples that contain no viral sequences. Take care not to contaminate cells or tissue samples with unwanted DNA sequences.

2. Mix sample gently up and down with a 200- μ l microcapillary pipet. Add 100 μ l buffered phenol to digestion mixture, mix gently, add 100 μ l of 24:1 chloroform/isoamyl alcohol, and mix gently.

The initial mixing shears the DNA very slightly, allowing for more efficient extraction. The gentle extraction maintains the DNA as high-molecular-weight molecules. This aids in estimating its concentration but is less important if mere detection is desired.

Screw-cap microcentrifuge tubes and microcapillary pipets minimize aerosol contamination which commonly results from microcentrifuges and automatic pipettors. To prepare microcapillary pipets, autoclave on dry cycle or bake in a 200°C oven overnight. Dedicate a single bulb for each reagent used. Positive displacement pipets with disposable tips and plungers are a more expensive, but easier-to-use alternative to microcapillary pipets.

3. Microcentrifuge 5 min at high speed, room temperature, and transfer aqueous phase (which contains DNA) to a new microcentrifuge tube.

Even with the use of screw-cap microcentrifuge tubes, avoid using a microcentrifuge that has been used for PCR products or large amounts of plasmid DNA.

4. Back-extract organic phase with proteinase digestion buffer (\sim 1/2 original volume).
5. Microcentrifuge 5 min at high speed, room temperature, and add aqueous phase to aqueous phase from step 3.
6. Extract aqueous phase twice with an equal volume of 24:1 chloroform/isoamyl alcohol, centrifuging 5 min at high speed, room temperature, each time to separate the phases.
7. Add 10 M ammonium acetate to 2.5 M (final) in the last aqueous phase and mix gently. Add 2.5 vol cold 100% ethanol and mix gently. Place on crushed dry ice 30 min. Microcentrifuge 15 min at high speed, 4°C, to pellet DNA, and pour off supernatant.

The use of ammonium acetate during ethanol precipitation is critical for subsequent efficient amplification. DNA precipitated from sodium acetate does not amplify as well, but can be improved by reprecipitation with ethanol from 2.5 M ammonium acetate.

8. Wash DNA pellet with 1 ml of 70% ethanol by inverting the tube several times. Microcentrifuge 15 min, pour off supernatant, and dry pellet in a desiccator under vacuum or in a Speedvac evaporator.

Avoid using a desiccator or evaporator that has been used for PCR products or large amounts of plasmid DNA.

9. Resuspend pellet in TE buffer, pH 7.5 (100 μ l for a sample prepared from 2×10^6 cells). Store DNA at 4°C.

The final DNA concentration should be 5 to 100 ng/ μ l. It may take several hours or overnight before the DNA is totally dissolved.

10. Estimate DNA concentration either by running an aliquot on an agarose gel alongside known amounts of standard DNA or more rapidly by ethidium bromide dot quantitation.

This step is important because the subsequent quantitation will be less accurate if $\geq 1 \mu\text{g}$ or $\leq 10 \text{ ng}$ DNA is assayed. Estimating the amount of DNA on a gel has the advantage over ethidium bromide dot quantitation in that it assesses how intact the DNA is. The extra precautions outlined above will diminish yield somewhat.

Amplify the DNA by PCR

11. Prepare one tube containing 110 ng, one tube containing 90 ng, and several tubes each containing 100 ng DNA from uninfected cells or tissue. Use these tubes to make a set of 10-fold serial dilutions of the sequence of interest as follows: Add a known amount (e.g., 20,000 molecules) of DNA containing the sequence of interest to the tube containing 110 ng DNA and mix. Then add one-tenth of this material to a tube containing 100 ng DNA and mix. Add one-tenth of this material to a new tube containing 100 ng of DNA and mix. Repeat several more times until a tube contains ≤ 10 molecules of the sequence of interest. Then add one-tenth of the material from that tube to the tube containing 90 ng of DNA. The final total volume of each tube should be $\leq 71 \mu\text{l}$.

To compare with samples from a mouse infected with a virus, prepare the first tube so there is one copy of viral DNA per mouse cell equivalent in 100 ng of DNA from an uninfected mouse, which is $\sim 20,000$ molecules of viral DNA. Make serial dilutions to yield samples with 0.1, 0.01, 0.001 (etc.) copy per cell equivalent. As a control, prepare a sample containing 100 ng uninfected cell DNA without added viral sequences.

12. For each amplification reaction, prepare a PCR tube containing 24 μl reaction mix cocktail and enough sterile distilled water for a final volume of 100 μl (after addition of DNA and *Taq* DNA polymerase, steps 13 and 15). Mix and overlay each reaction with 2 drops ($\sim 100 \mu\text{l}$) mineral oil, to completely cover the surface of the reaction mixture.

*Overlaying with oil at this step rather than after adding *Taq* DNA polymerase helps minimize contamination because everything is assembled before adding the experimental samples. There is also less chance of contaminating the oil with DNA.*

13. Open only those tubes that will contain equivalent samples (e.g., duplicate samples or samples from animals infected with the same virus inoculum). Add 100 ng sample DNA to each appropriate tube and close the tubes. Microcentrifuge briefly to mix.

Begin with the no-virus control (e.g., 100 ng mouse DNA) followed by other negative controls (include at least two other negative controls). Proceed in order of increasing likelihood that a sample will contain viral sequences of interest, adding serial dilution tubes last. Adding samples in this order lessens the likelihood that aerosols of more concentrated samples will contaminate less concentrated samples.

14. Heat-denature the samples 1 min at 94°C, either in a water bath or in an automated cycler.
15. Open tubes (containing equivalent samples) and add 5 μl of 0.8 U/ μl *Taq* DNA polymerase to each tube. Close tubes and repeat steps 13 to 15 with the next set of equivalent samples.

*Denaturing before adding the *Taq* DNA polymerase puts the polymerase through one less cycle at high temperature.*

16. Microcentrifuge tubes briefly. Cycle tubes one time for 2 min at 55°C (reannealing) and 3 min at 72°C (extension).

The microcentrifuge spin mixes in the polymerase. This step completes the cycle started by the heat denaturation in step 14.

17. Carry out PCR (UNIT 15.1) according to the following program:

29 cycles:	1 min	94°C	(denaturation)
	1 min	55°C	(annealing)
	1 min	72°C	(extension)
1 cycle:	7 min	72°C	(extension)

Store completed reactions at 4°C.

Analyze the PCR products

- Electrophorese product aliquots on an appropriate agarose or polyacrylamide gel (UNIT 2.5A & 2.7). Include lanes with DNA molecular weight markers that will be visible both upon ethidium bromide staining and by autoradiography.

One-tenth of the reaction should be sufficient for detection of the amplified host sequences by ethidium bromide staining and for detection of the amplification products of single molecules following hybridization. Do not use markers that are so radioactive that they will obliterate other lanes during autoradiography.

- Stain and photograph the gel.

The DNA fragment corresponding to the PCR product from the host DNA should be readily visible following ethidium bromide staining and will serve as a measure of the efficiency of each amplification reaction. Don't be surprised to see a number of other bands, especially of higher molecular weights, which are nonspecific PCR products. However, if the reaction has been optimized (UNIT 15.1), the specific PCR product should predominate. Check for specificity by restriction endonuclease digestion (if the product contains a site) or by hybridization with an internal oligonucleotide as described below.

If the PCR reaction is robust and yields a single product, sometimes staining with dyes such as SYBR Gold rather than with ethidium bromide can provide sufficient sensitivity and the bands can be quantitated using fluorimetric equipment. If not, proceed to steps 20 to 22.

- Transfer gel to a nitrocellulose or nylon filter (UNIT 2.9A). If desired, UV cross-link DNA to filter.

Choice of filter and transfer method is an individual preference; however, electroblotting from acrylamide gels to nylon filters followed by UV cross-linking (UNIT 2.9) has proven to be a quantitative method for transferring and retaining small DNA fragments on filters during multiple cycles of hybridization, stripping, and reprobing.

- Prehybridize the filter and hybridize with end-labeled oligonucleotide specific for the sequence of interest. Analyze by autoradiography (APPENDIX 3A) or phosphorimager.

If each reaction is amplified with similar efficiency (which can be assessed from the ethidium bromide-stained gel), the dilution series should give a monotonically decreasing signal of one predominant band with increasing dilution. It should be possible to detect a signal from dilutions containing only one or a few molecules. There should be no signal from the negative controls. A reasonable estimate for the number of molecules of the sequence of interest in each sample can be obtained by comparison with the dilution series, assuming that each reaction amplified with similar efficiency.

- For more quantitative analysis, strip the filter of the previous probe by boiling in water 15 min (if necessary) and hybridize with a probe specific for the host single-copy sequence. Quantitate the signals by densitometric scanning (following the densitometer manufacturer's instructions) and compute a standard curve from the dilution series, normalizing to the host sequence signals. Determine the number of molecules in the experimental samples by interpolation from the standard curve.

The standard curve will ideally give a linear relationship between the log of the autoradiographic signal and the log of the amount of DNA. However, it may not necessarily be completely linear, and it is unlikely to have a slope of 1. At the high end (0.1 to 1 copy per cell equivalent), it may plateau. If this is a problem, it may be necessary to reduce the number of amplification cycles or vary other parameters (see UNIT 15.1).

REAGENTS AND SOLUTIONS

Oligonucleotide primers for amplification

One pair each for each target sequence. Prepare each primer at 50 pmol/ μ l in sterile, distilled water. Store at -20°C .

Oligonucleotide primers for hybridization

One for each target sequence at 50 pmol/ μ l in sterile, distilled water. Store at -20°C .

Proteinase digestion buffer

20 mM Tris·Cl, pH 7.4 (prepared from autoclaved, 1 M stock; *APPENDIX 2*)

20 mM EDTA, pH 8 (prepared from autoclaved, 0.5 M stock; *APPENDIX 2*)

0.5% sodium dodecyl sulfate (SDS)

Store at room temperature

Reaction mix cocktail (per amplification reaction)

10 μ l 10 \times PCR amplification buffer (*UNIT 15.1*)

10 μ l 2 mM 4dNTP mix (*UNIT 15.1*)

1 μ l each oligonucleotide primer for amplification (4 μ l total; see recipe above)

COMMENTARY

Background Information

PCR technology provides the most sensitive methods for detecting nucleic acids. These methods have proven exceedingly useful in the detection of infectious agents in experimental and clinical settings (Kwok et al., 1987) and in analyses of unusual and precious small samples of tissue, such as those from extinct animals (Paabo, 1989). They also have considerable potential for forensic applications (von Beroldingen et al., 1989). Most initial efforts to utilize the sensitivity of PCR were geared to issues of detection. For example, Saiki et al. (1988) showed that they could detect single target sequences (a b-globin gene) in 10^5 to 10^6 cells. Similarly, procedures for amplification of DNA from single cells exist (Li et al., 1988). However, these procedures were not designed to quantitate amounts of target sequence in different samples.

The Basic Protocol presented here to quantitate rare DNAs was developed to measure the amount of herpes simplex virus (HSV) DNA in the ganglia of mice infected with various HSV mutants. Initial efforts used slot-blot hybridization, but failed to detect viral sequences reliably below 0.01 to 0.1 copy of HSV DNA per mouse cell equivalent (Leib et al., 1989). For this reason, a number of parameters in the basic PCR protocol (*UNIT 15.1*) were varied to keep the replication machinery in excess to the number of templates (see Critical Parameters and Troubleshooting). Using this approach, it has been possible to quantitate HSV DNA over a 10^4 -fold range, from a few molecules per 100 ng ganglion DNA from mice infected with mu-

tants with serious growth impairments to tens of thousands of molecules in ganglion DNA from mice infected with wild-type virus (Katz et al., 1990). This approach should be pertinent to numerous other applications where small amounts of nucleic acid need to be detected and quantitated, including diagnostics and forensics. In fact, similar assays have been developed to quantitate nucleic acids encoded by human retroviruses (Arrigo et al., 1989; Pang et al., 1990) and to measure cytokine mRNA and DNA (Gilliland et al., 1990).

Critical Parameters and Troubleshooting

All parameters that are critical in the basic PCR protocol (*UNIT 15.1*) are also critical here. Several additional points particular to this protocol bear emphasis.

Optimization

The PCR must be optimized to the particular set of templates and primers and to the lot of *Taq* DNA polymerase used, following the guidelines in *UNIT 15.1*. Begin optimization with a reconstructed mixture of DNA, such as viral DNA mixed with host DNA at fairly high copy number, where both products can be visualized by ethidium bromide staining (e.g., 1 copy per cell equivalent).

DNA

The amount of added DNA is critical for achieving a quantitative assay. Too much DNA will saturate the replication machinery, not

only in terms of specific products generated, but also nonspecific products. In the protocol presented above, 100 ng of DNA are used per 100 µl reaction. Test a dilution series of the sequence of interest mixed with the endogenous sequences (e.g., various copies of viral DNA per host cell equivalent) at various amounts of total sample DNA to see how much sample DNA will allow a monotonically increasing signal with increasing amounts of the sequence of interest. In addition, impurities in DNA preparations can interfere with PCR; additional organic extractions and/or ethanol precipitations in the presence of ammonium acetate may help.

Internal control

A set of primers to yield an internal control sequence is critical for two major reasons. First, if no signal from the sequence of interest is obtained in a given sample, the internal control will verify whether this is a true or a false negative. It is not uncommon for amplifications to fail, especially since some experimental samples contain contaminants that interfere with PCR. Second, the internal control allows for quantitation, since it normalizes for several factors including variation in the amount of sample DNA, efficiency of amplification, and the amount loaded on the gel. The internal control PCR product should be different enough in size to be resolved easily from the PCR product of interest, but close enough in size so there is no concern about differences in transfer efficiency due to size differences. It should also be clearly distinguished in size from an artifactual “primer-dimer” product (see UNIT 15.1).

In the protocol described here, the internal control product is derived from endogenous cellular sequences. For some applications, this would not be appropriate or practical. An alternative is to spike the samples with a known amount of a foreign sequence and to include primers specific for this DNA in the amplification reaction. This would control for the efficiency of the reaction and certain other factors, but not for variation in the amount of sample DNA.

DNA polymerase

The amount of *Taq* DNA polymerase added per reaction is critical in achieving a quantitative assay. The protocol outlined above uses more polymerase than the Basic Protocol in UNIT 15.1 to insure that the replication machinery is in excess of the number of templates. How-

ever, because *Taq* DNA polymerase is not the cheapest reagent, and because too much of this enzyme is actually inhibitory, vary the amounts of polymerase to determine an amount that permits a quantitative assay without bankrupting the laboratory.

Contamination

It is critical in any procedure that can detect only a few molecules to be absolutely scrupulous in avoiding contamination from exogenous nucleic acids. This protocol explicitly states several precautions one can take to avoid contamination. They are summarized below in order of importance.

Never prepare DNA or PCR reagents or assemble PCRs in the same place that is used to handle large amounts of plasmid DNAs or PCR products. PCR products are the worst possible contaminants because they give rise to more of themselves faster than normal DNA. If possible, perform the two different procedures in different rooms, avoiding even the presence of PCR products in the room where DNA is prepared or PCRs are assembled. In addition, try preparing DNA and PCR reagents and assembling PCRs in a laminar flow hood equipped with a UV light to inactivate contaminating DNA.

If possible, use only sterile disposable plasticware for preparing and storing reagents. If glassware must be used, try to ensure that it has not been contaminated with plasmid or genomic DNAs.

Most pipetting devices can be readily contaminated by the aerosols that they produce. This protocol was originally developed using sterile disposable microcapillary pipets with a separate bulb for each reagent and sample type. This degree of caution may be extreme. A more expensive, but more convenient, alternative may be to use positive displacement pipetting devices with disposable tip and plunger, but even then, be sure not to use the same device for handling PCR products. The least expensive, most convenient approach is to use disposable tips with barriers and a standard micropipettor to minimize aerosol contamination.

Reagents should be sterile whenever possible and stored as small aliquots.

Wear disposable gloves and change them frequently during preparation of DNA and assembly of PCR.

Be careful of static electricity and avoid aerosols from opening tubes. As outlined in the protocol above, keep all tubes closed except those receiving sample DNA. If possible, pre-

pare only equivalent DNA samples at the same time. The use of screw-cap tubes can be helpful. It may also be useful to dedicate a microcentrifuge solely to PCR assembly.

As stated in the protocol, it is critical to include a sufficient number of negative controls (≥ 3) including a zero-copy reconstruction, which can help assess reagent contamination, as well as controls for DNA preparation, such as samples from mock-infected animals.

If contamination by plasmid DNAs or PCR products is suspected, it can be confirmed or ruled out using appropriately chosen sets of primers.

Aside from PCR products and plasmid DNAs, be suspicious of other sources of contamination, including various pieces of lab apparatus that come into contact with large amounts of target sequence. This extends to tools used to extract tissue specimens. In addition, contamination from the investigator should be avoided. Some have found that wearing a mask and hair covering is helpful. Paranoia can be a positive character trait in this situation.

Several investigators have used nucleases or UV light to inactivate contaminants in solutions. These methods can be useful but have their limitations, especially if the contaminating DNA is short and thus represents a small target.

The quantitative protocol described here can be helpful in assessing contamination because unlike nonquantitative assays, it can give an estimate of the level of contamination. Even if negative controls yield positive signals, if they are much lower than those in experimental samples it provides at least tentative assurance that the experimental signals are real.

Transfer and cross-linking

It is very important that PCR products be transferred and retained completely to ensure a quantitative assay. If the products are electrophoresed on acrylamide, the electroblotting and UV-cross-linking protocols in *UNIT 2.9* provide methods that have proven successful in this regard.

An alternative approach to transfer and hybridization is to use PCR primers prelabeled to high-specific activity with polynucleotide kinase. The products can then be quantitated by direct autoradiography of the gel (Arrigo et al., 1989; Pang et al., 1990). Potential disadvantages of this approach are that it adds steps where contamination can be introduced and detects more nonspecific products.

Anticipated Results

Results are obtained at different stages of the protocol. The DNA preparation should yield relatively intact and pure DNA, but due to the precautions described, yields may not be as high as ordinarily obtained.

After the PCR products are electrophoresed on a gel, ethidium bromide staining should reveal an easily visible band corresponding to the internal control product in every product lane. A visible band corresponding to the experimental product of interest should appear in the lane containing the 1-copy-per-cell-equivalent reconstruction and in any lanes that contain similar amounts of the sequence of interest. Because the sequences of interest are relatively nonabundant, a variety of nonspecific products may be seen as well.

After hybridization to the sequence of interest and autoradiography, the expected results are labeled bands of the appropriate size at monotonically decreasing intensities in the reconstructed dilution series, no bands in the negative controls at that size, and bands with varying intensities in the experimental samples. Depending on the probe and the stringency of the hybridization and wash conditions, nonspecific sticking of the probe to the abundant internal control product may be seen. It is also common to see minor specific PCR products of slightly greater or lower mobility than the major specific product, especially at high copy number.

Time Considerations

The time to prepare DNA is mainly occupied by the proteinase K digestion, which usually runs overnight. Therefore, it is usually possible to start the procedure one afternoon and complete the manipulations by noon the next day. It may take somewhat longer for the DNA to dissolve completely.

Assembling the amplification reactions requires only a couple hours or less, depending on the number of samples. Thirty cycles of denaturation, annealing, and extension require ~4 hr in a good automated cycler; slightly less if done by hand with three water baths. With an automated cycler, the gel can be poured during the cycling.

Running the gel only requires a few hours or an overnight run. Staining, destaining, and photographing requires 1 to 2 hr, electroblotting ≥ 3 hr, and prehybridization and hybridization a few hours to overnight. When everything is optimal, the PCR signal from one or a few molecules can be visualized in an overnight

exposure. Thus, from the point of acquiring a sample to autoradiographic signal is usually 4 normal working days.

Literature Cited

- Arrigo, S.J., Weitsman, S., Rosenblatt, J.D., and Chen, I.S.Y. 1989. Analysis of *rev* gene function on human immunodeficiency virus type 1 replication in lymphoid cells by using a quantitative polymerase chain reaction. *J. Virol.* 63:4875-4881.
- Gilliland, G., Perrin, S., Blanchard, K., and Bunn, H.F. 1990. Analysis of cytokine mRNA and DNA: Detection and quantitation by competitive polymerase chain reaction. *Proc. Natl. Acad. Sci. U.S.A.* 87:2725-2729.
- Katz, J.P., Bodin, E.T., and Coen, D.M. 1990. Quantitative polymerase chain reaction analysis of herpes simplex virus DNA in ganglia of mice infected with replication-incompetent mutants. *J. Virol.* 64:4288-4295.
- Kwok, S., Mack, D.H., Mullis, K.B., Poiesz, B., Ehrlich, G., Blair, D., Friedman-Kien, A., and Sninsky, J.J. 1987. Identification of human immunodeficiency virus sequences by using in vitro enzymatic amplification and oligomer cleavage detection. *J. Virol.* 61:1690-1694.
- Leib, D.A., Coen, D.M., Bogard, C.L., Hicks, K.A., Yager, D.R., Knipe, D.M., Tyler, K.L., and Schaffer, P.A. 1989. Immediate-early regulatory gene mutants define different stages in the establishment and reactivation of herpes simplex virus latency. *J. Virol.* 63:759-768.
- Li, H., Gyllensten, U.B., Cui, X., Saiki, R.K., Erlich, H.A., and Arnheim, N. 1988. Amplification and analysis of DNA sequences in single human sperm and diploid cells. *Nature (Lond.)* 335:414-417.
- Paabo, S. 1989. Ancient DNA: Extraction, characterization, molecular cloning and enzymatic amplification. *Proc. Natl. Acad. Sci. U.S.A.* 86:1939-1943.
- Pang, S., Koyanagi, Y., Miles, S., Wiley, C., Vinters, H.V., and Chen, I.S.Y. 1990. High levels of unintegrated HIV-1 DNA in brain tissue of AIDS dementia patients. *Nature (Lond.)* 343:85-89.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487-488.
- von Beroldingen, C.H., Blake, E.T., Higuchi, R., Sensabaugh, G.F., and Erlich, H.A. 1989. Applications of PCR to the analysis of biological evidence. In *PCR Technology: Principles and Applications for DNA Amplification* (H.A. Erlich, ed.) pp. 209-223. Stockton Press (W.H. Freeman), New York.

Key Reference

Katz et al., 1990. See above.

Uses the protocol outlined here and presents examples of data generated.

Contributed by Donald M. Coen
Harvard Medical School
Boston, Massachusetts

High-Throughput Real-Time Quantitative Reverse Transcription PCR

This unit describes the use of real-time quantitative PCR (QPCR) for high-throughput analysis of RNA expression. The topics covered include: design and validation of QPCR primers and probes for both SYBR Green- and TaqMan-based assays (see Support Protocol); the standard curve method (see Basic Protocol 1); an efficiency-corrected ΔC_t (cycle time, also called cycle threshold or crossing point) method (see Basic Protocol 2); and the comparative cycle time, or $\Delta\Delta C_t$ method (see Alternate Protocol). While the unit describes the use of the Applied Biosystems 7900HT (high-throughput, 384-well) instrument, the protocols may be utilized for any real-time PCR instrument. The high-throughput design allows analysis of the levels of transcript from a number of genes of interest (GOIs) at one time by using the appropriate primer set for each gene. (Within this unit, the term GOI will refer to the actual gene of interest as well as its RNA product or cDNA copy.)

Because of the simplicity of the mathematical application, the standard curve method (Basic Protocol 1) is the most basic and straightforward QPCR assay described in the unit. In this method, standard curves are constructed for all of the GOIs from which RNA expression is being measured, and linear regression analysis is applied to interpolate arbitrary unknown sample values. The standard curve assay may be performed even if the PCR amplification efficiencies of the primer sets (as determined by the template dilution assay in the Support Protocol) are not equal, since correction for unequal efficiencies is intrinsic to the linear regression formula. One drawback of the standard curve method is that standard curves must be run for each of the primer sets on an assay plate. This results in less space on the plate for the unknown samples, and requires the use of additional reagents. This use of resources is particularly excessive when the PCR amplification efficiencies of the primer sets have been determined to be 100% and relative fold-change is the preferred outcome of the measurements. In such cases, the $\Delta\Delta C_t$ method (see Alternate Protocol) should be employed instead. Another limitation is that unless the levels of all of the GOIs in the cDNAs used to construct the standard curves are known, the relative concentration of one GOI cannot be compared to that of another GOI. If comparison between the levels of different GOIs (without the knowledge of the relative level of the transcripts in the standards) is desired, the efficiency-corrected ΔC_t method (see Basic Protocol 2) should be applied.

The efficiency-corrected ΔC_t method builds upon the standard curve method by incorporating PCR efficiency (E) into the quantity calculations. The standard curve slopes are used to calculate PCR efficiency according to the relationship $E = 10^{(-1/\text{slope})}$. The efficiency has a maximum value of 2 for perfect doubling of the PCR template (see Basic Protocol 2 for an in-depth explanation of E). Note that in this method the standard curve is used only to determine slope and not to interpolate the RNA values of the unknown samples. The efficiency correction is then applied to determine the relative amount of RNA using the measured C_t values for the test samples, and is particularly important when comparing the levels of different RNAs whose standard-curve slopes deviate from each other by greater than ± 0.1 . This application is useful for comparing the expression profiles of many different RNAs, e.g., those belonging to gene families or related biological pathways. The sample-space and reagent-use limitations mentioned for the standard curve method also apply here.

The $\Delta\Delta C_t$ method is the method of choice when the desired output is "fold-change," because standards are not necessary, thus saving both reagents and space on the reaction plate. However, this method requires that the amplification efficiencies of the primer/probe sets be 100%. If the amplification efficiencies are suboptimal, but the primers generate a single product as determined by melting curve analysis (Support Protocol), Basic Protocol 1 should be used to determine fold-changes.

Chapter 4 describes the isolation of RNA from several sources and *UNIT 15.5* details the traditional procedure for reverse transcription of RNA into complementary DNA (cDNA). For QPCR, it is recommended that total RNA be resuspended in diethylpyrocarbonate (DEPC)-treated water or an equivalent nuclease-free buffer that does not contain EDTA. The RNA should be treated with DNase and then reverse transcribed using 0.08 $\mu\text{g}/\mu\text{l}$ (final concentration) random hexamer or nonamer primers. Purified messenger RNA—i.e., poly(A)⁺ RNA (*UNIT 4.5*) or RNA that has been reverse transcribed using oligo(dT) or gene-specific reverse primers—may also be successfully used in the assay. See Commentary for more detailed information.

Following the assay, the resulting raw data are analyzed using second-party software, usually Microsoft Excel or equivalent. The data analyses are dependent on the type of assay performed, and are outlined in detail as part of each protocol.

NOTE: General precautions for working with RNA are described in *UNIT 4.1* and other Chapter 4 units, and general precautions necessary for PCR are described in *UNIT 15.1*. In particular, the use of molecular-biology-grade water, RNase/DNase/nucleic acid-free tubes, aerosol-barrier pipet tips, and dedicated pipettors of all types (i.e., pipettors used only for RNA or PCR applications, which are kept out of areas used for plasmid or genomic DNA work) is strongly recommended for all steps in this unit. If dedicated pipettors are not available, the available pipettors should be thoroughly cleaned to remove nucleases and potential contaminants such as plasmid or genomic DNA. In addition, the use of gloves is required, since even a small amount of any contaminant can greatly impact the results of the assay.

STRATEGIC PLANNING

To begin performing a QPCR assay, design and validation of the appropriate primers and probes must first be completed (refer to Support Protocol). Second, the appropriate assay is selected based on the goal of the experiment and the desired data output (refer to Basic Protocols 1 and 2, and the Alternative Protocol, for guidelines used in making this determination). The chosen assay is then performed and the data are analyzed.

STANDARD CURVE METHOD

The standard curve method is used for determining the level of a gene of interest (GOI) relative to an endogenous reference RNA, and for calculating relative fold-changes of a GOI between experimental samples. The assay is useful for determining an "expression profile" of a single GOI within a group of samples. A dilution series of standard cDNA samples is constructed for the GOI and reference gene, and linear regression analysis is applied. The formulas resulting from the standard curves are used to interpolate the GOI and reference-gene quantities in the unknown samples. An endogenous reference gene, often a housekeeping gene, is used as a control to normalize the amount of input template for each sample (see Commentary for parameters used in choosing the appropriate reference gene). The data are expressed as "normalized RNA level" in arbitrary units. The type of nucleic acid standard chosen depends upon the nature of the unknown samples and is discussed in the Commentary.

BASIC PROTOCOL 1

High-Throughput Real-Time Quantitative Reverse Transcription PCR

15.8.2

The standard curve assay does not require that the amount of the GOI or reference RNA in the standards be known. It depends on the linear regression formula produced by plotting the C_t versus the log nanogram (log ng) of input standard total RNA. It should be noted that cDNA concentrations are not typically determined following reverse transcription of the RNA. Here, quantity refers to total RNA input prior to reverse transcription. The standard curve–plotting function is available in most instrument software. If it is not, graphing software may be used instead. Since the input ng values refer to the input amount of total RNA, and not to a known amount of target molecules, the numbers generated are simply arbitrary and may not be compared with the numbers calculated for a different GOI. If comparison of the relative RNA levels between different RNA targets is desired, refer to the efficiency-corrected ΔC_t method (Basic Protocol 2). The standard curve method should be used instead of the $\Delta\Delta C_t$ method (Alternate Protocol) to find fold-changes between samples when the amplification efficiencies of the primer sets are not 100% as determined by the template dilution assay (Support Protocol).

Materials

- 20 ng/ μ l experimental cDNA samples (concentration based on RNA input for cDNA synthesis; see *UNIT 15.5*)
- Dilution series of standard cDNAs (see recipe)
- No-template control sample (NTC; prepared at the same time as the cDNA samples using molecular-biology-grade water instead of RNA; see Commentary)
- No-reverse-transcriptase control samples (–RT; prepared at the same time as the cDNA samples using molecular-biology-grade water instead of reverse transcriptase; see Commentary)
- 2 \times SYBR Green or TaqMan mix containing ROX (Applied Biosystems, Bio-Rad, Invitrogen, Sigma, *or* see recipe for 2 \times SYBR Green mix)
- Primer mixes, 1.25 μ M each forward and reverse primer (see recipe and Support Protocol), for each reference gene and GOI to be tested
- 5 μ M TaqMan probe (for TaqMan protocol only; see recipe and Support Protocol)
- Molecular-biology-grade water (nucleic acid and nuclease free)
- 8-tube PCR tube strips (optional, but recommended; can be of low quality since they will only be used for mixing reaction components; ISC Bioexpress)
- 96-well PCR tube racks (optional, but recommended; ISC Bioexpress)
- Digital multichannel pipettor, 8- or 12-channel, 5- to 100- μ l capacity (recommended)
- Centrifuge with swinging-bucket rotor and microtiter plate carriers
- 384-well optical reaction plates (Applied Biosystems)
- Optical adhesive covers (Applied Biosystems)
- Real-time thermal cycler: e.g., Applied Biosystems 7900HT
- Microsoft Excel or spreadsheet program with equivalent statistical features

Set up plates

1. Plan the plate arrangement according to the number of samples and primer sets to be assayed (reference gene plus GOIs). For each primer set, include the standards, the NTC, and the –RT control as part of the sample group. See Figure 15.8.1 for an example of a typical plate setup.

All steps may be performed at room temperature if a vendor-supplied SYBR Green or TaqMan mix is used. Most vendor-supplied 2 \times mixes are stable at room temperature for a number of hours.

If the instrument is equipped with a plate-loading robotic arm, several plates may be prepared at once and stacked into the robotic arm queue. Alternatively, plates may be made several days in advance and stored at 4°C until ready to run, without a decline in the performance of the PCR enzyme. Advance preparation of plates is not recommended if SYBR Green or TaqMan mixes are prepared in the laboratory, since the home-made

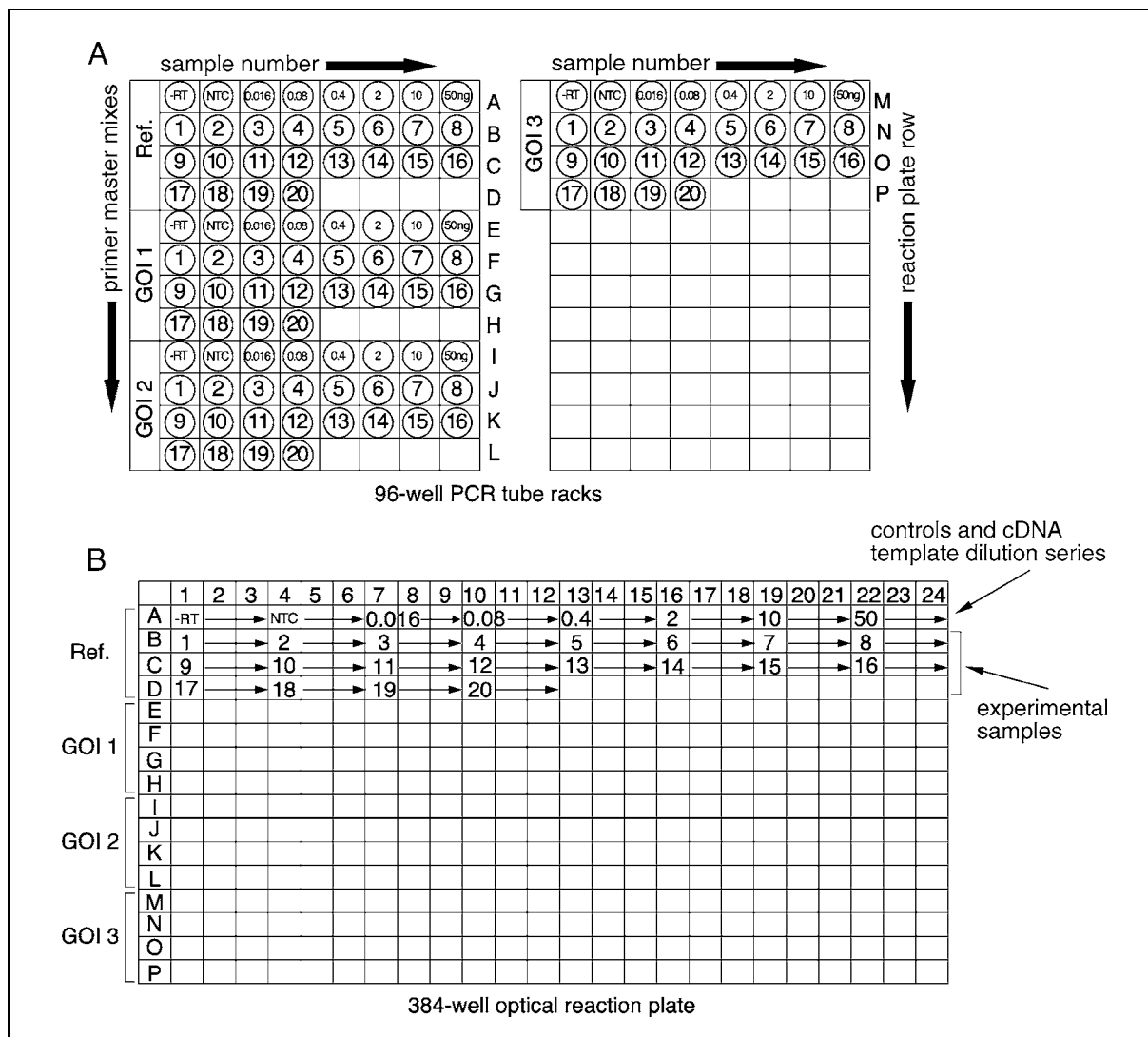


Figure 15.8.1 Typical plate setup for the standard curve and efficiency-corrected ΔC_t assays. **(A)** Organization of PCR tube strips on a 96-well PCR rack when preparing for any QPCR assay. It is advisable to premix the cDNA templates with the primer master mixes in tube strips before putting them into the reaction plate. This practice decreases the variability between replicate wells. A typical cDNA/primer mix setup is shown in relation to the final 384-well reaction plate **(B)**. The plate arrangement shown represents an experiment in which 20 samples taken from experimental animals will be assayed for levels of one endogenous control RNA and three GOIs. The samples are plated in triplicate for each of the RNAs assayed, and a standard curve is required for each of the RNAs to be measured (the row before the unknown samples contains the standard samples, whose amounts are shown in ng total RNA). Abbreviations: GOI, gene of interest; NTC, no-template control; Ref, endogenous control gene; -RT, no-reverse-transcriptase control.

mixes lack additives present in the commercial mixes that confer stability to the reaction components.

2. Prepare primer master mixes according to Table 15.8.1, but without template cDNA.

This can be done in advance, and the tubes may be put on ice or stored at 4°C for a few hours before preparing the reaction plate.

The authors always use the same mixes because this greatly increases the high-throughput nature of the assay. Keeping the mixes constant allows not only universal mix conditions but universal cycling conditions. If the initial primer set does not perform well, new primers are designed (also see Support Protocol).

Table 15.8.1 Master Mixes for the Standard Curve and Efficiency-Corrected Δ Ct Assays

Components	Final concentration	Volume per well	Volume for each cDNA/primer mix (per sample in triplicate + 1 extra)	Master mix (no. of samples + 8 standards ^a + 1 extra)
<i>SYBR Green assay</i>				
2× SYBR Green mix	1×	5 μ l	20 μ l	
1:1 primer mix (1.25 μ M each)	150 nM each primer	1.2 μ l	4.8 μ l	
Template cDNA	10-25 ng ^b	1.25 μ l	5 μ l	—
H ₂ O	N/A	to 10 μ l	to 40 μ l	
Total volume		10 μ l	40 μ l	Aliquot 35 μ l into each tube containing cDNA
<i>TaqMan assay</i>				
2× TaqMan mix	1×	5 μ l	20 μ l	
1:1 primer mix (1.25 μ M each)	300 nM each primer	2.4 μ l	9.6 μ l	
5 μ M TaqMan probe	250 nM	0.5 μ l	2.0 μ l	
Template cDNA	10-25 ng ^b	1.25 μ l	5 μ l	—
H ₂ O	N/A	to 10 μ l	to 40 μ l	
Total volume		10 μ l	40 μ l	Aliquot 35 μ l into each tube containing cDNA

^aFor this assay, the no-template control (NTC) and no-reverse-transcriptase (–RT) control are included as part of the standard sample set.

^bRecommended amount of template for detection of both high and low levels of GOIs. If necessary, significantly less template may be used (picograms). Note that the cDNA template quantity is based upon the amount of total RNA input into the reverse transcription reaction.

- Place the appropriate number of 8-tube PCR strips into a 96-well PCR tube rack (see Fig. 15.8.1A). For convenience, use a different color tube strip for each different primer master mix that has been prepared. Label the side of each strip with the letter of the row of the reaction plate into which the samples will be placed.

Alternatively, cDNA/primer mixes may be made in 0.65-ml microcentrifuge tubes or 0.2- or 0.5-ml PCR tubes.

The cDNA may be put directly into the optical reaction plates followed by the primer master mix. However, this is not recommended because it introduces a potential source of experimental error, since the assay replicates are not premixed, but pipetted individually. The end result may be lower assay precision.

- Using a multichannel pipettor, put 5 μ l of cDNA (experimental samples, standards, and controls) into the bottom of the appropriate tubes in the strips.
- Being careful not to touch the cDNA inside the tubes, use a multichannel pipettor to place a 35- μ l aliquot of the appropriate primer master mix into each tube.
- Cover the entire rack of tube strips with Parafilm and gently vortex to mix. Gently tap or briefly centrifuge the PCR tube racks (2 to 3 min at 1700 \times g, 4°C or room temperature, in a swinging-bucket rotor with microtiter plate carriers) to get contents to the bottoms of the tubes.

- Using a multichannel pipettor, dispense 10 μ l of each cDNA/primer mix into the appropriate three wells of the optical reaction plate to generate each sample in triplicate (as planned in step 1; see Fig. 15.8.1B).

If the multichannel pipettor has 8- or 12-channel dispensing capability, the triplicates can be dispensed at the same time for different rows. Note that due to the spacing between rows of a 384-well reaction plate, every other row can be added at once (i.e., A, C, E, G, I, K, M, O, and then B, D, F, H, J, L, N, P), thus greatly minimizing the pipetting time.

- Cover the plate with the optical adhesive cover and then briefly centrifuge the plate as above to get contents to the bottoms of the wells.

Perform real-time PCR

- For real-time PCR:* Transfer the plate to the real-time thermal cycler and run real-time PCR using the following program (consult the instrument manual for specific instructions):

1 cycle:	10 min	95°C (activates the hot-start <i>Taq</i> DNA polymerase)
40 cycles:	15 sec	95°C (collect data throughout)
	1 min	60°C (collect data throughout).

- For melting (dissociation) curve analysis (for use with SYBR Green only):* Add these steps following the 40 cycles of the thermal cycling program.

15 sec	95°C
15 sec	60°C (collect data)
Increase from 60° to 95°C at a 2% temperature ramping rate (collect data)	
15 sec	95°C (collect data).

See Support Protocol for a description of the use of melting curve analyses.

Analyze data

- Analyze and export raw data (see instrument manual for detailed instructions about document setup, baseline, and threshold settings).

Some instrument software applications contain a standard curve plotting feature. If this function is not available, use Excel or another graphing program to plot C_t versus the log nanograms (ng) of input total RNA for each standard, and apply a best-fit line to generate the linear regression formula $y = mx + b$, where y is C_t of the unknown sample, m is the slope, x is the quantity of the unknown sample (in log ng), and b is the y intercept for both the reference gene and each GOI. Interpolate the unknown sample quantities using the resulting formulas.

The transformation of the fluorescence signal into C_t data, as well as methods for baseline and threshold settings, vary by instrument. The specific instrument manual should be consulted.

In analyzing the raw data, it is important to adjust the cycle threshold (C_t) of the amplification plot to within the geometric phase of amplification. This is critical for proper analysis because the geometric phase represents the point of the reaction at which C_t is quantitatively related to the amount of initial PCR template. Note that a C_t decrease of 1 unit represents a two-fold increase in initial PCR template.

It is also important that the coefficient of determination, or R^2 -value, for the linear regression formula be 0.99. If the R^2 value is less than 0.99, this suggests that one or more points of the standard curve are deviating significantly from the best-fit line. In this case, the accuracy of the data obtained from the linear regression formula of this standard curve may be compromised. The R^2 value will never be above 0.99.

- Import data into Microsoft Excel or equivalent spreadsheet program with statistical features.

12. For each of the three replicates of a sample, calculate the average quantity (avg) of target cDNA interpolated from the standard curve, the standard deviation of the average (stdev), and the coefficient of variation (CV) according to the formula $CV = \text{stdev}/\text{avg}$.
13. Remove any outlier points (>17% CV). After removing the outlier point, recalculate avg, stdev, and CV.

Only one point per replicate may be removed.

A 17% CV correlates with the maximum allowable standard deviation that can distinguish a two-fold change with 99% confidence when samples are assayed in triplicate wells for both the endogenous reference and the GOI. If a 95% confidence interval is acceptable, a 21.8% CV may be used as the threshold for removing outliers. On the other hand, the Q-test (a test for rejection of discordant data) may be used to determine outlier points. Refer to Shoemaker et al. (1974) for a more in-depth description of this test.

14. For each sample, normalize the GOI quantity to that of the reference gene for the sample according to Equation 15.8.1. Use the recalculated values if outlier points were removed in step 13.

$$\text{normalized value} = \frac{\text{avg GOI quantity}}{\text{avg reference quantity}}$$

Equation 15.8.1

15. Calculate the standard deviation of the normalized value (SD) according to Equation 15.8.2.

$$SD = (\text{normalized value}) \times \sqrt{(\text{CV of reference})^2 + (\text{CV of GOI})^2}$$

Equation 15.8.2

16. Plot the resulting values as a bar graph of normalized value versus sample name or experimental treatment group, with the error bars equal to the SD.
17. If desired, calculate fold-changes between samples by choosing a calibrator sample (usually vehicle-treated or wild-type control) and dividing all of the normalized values from step 14 and the SD calculated in step 15 by the normalized value of this calibrator.

The resulting values are then expressed as fold-changes relative to the calibrator sample, which should now be equal to 1.

EFFICIENCY-CORRECTED ΔC_t METHOD

The efficiency-corrected ΔC_t method is used for determining the relative amounts of different GOIs that are normalized to an endogenous reference RNA. It may also be used to determine fold-changes of a specific RNA between samples, but may be excessive if the desired output is only fold-change (see Basic Protocol 1 or Alternate Protocol). Data obtained from the efficiency-corrected ΔC_t method are expressed as "normalized RNA level" in arbitrary units, and the calculated levels may be compared to those of other GOIs when the same threshold setting and assay chemistry are used (i.e., SYBR Green or TaqMan chemistry). It should be noted that an assumption is made that the reverse transcription efficiency is equal for all RNA transcripts in a single sample and for the same transcript between samples. In some cases, this may not be true (see Pfaffl, 2004, for further discussion); therefore, it is recommended that the RNA extraction method

**BASIC
PROTOCOL 2**

**The Polymerase
Chain Reaction**

15.8.7

remain the same for all samples, and that all samples under study be reverse transcribed at the same time with the same reaction buffer.

The basis of quantitative PCR lies in the principle that for every additional thermocycle, a two-fold increase of template-specific product occurs. Several factors affect whether a change in one cycle truly represents a two-fold growth in product, in other words, whether the reaction is 100% efficient. To assess the reaction's efficiency, linear regression analysis is applied to a standard cDNA dilution series, just as in Basic Protocol 1. The slope of the resulting standard curve is used as a measure of PCR efficiency (E) according to the equation $E = 10^{(-1/\text{slope})}$. Note that different GOIs may produce different E values.

A slope of -3.3 produces an E value of 2, indicating that a perfect doubling of the template has occurred. Calculated E values of less than 2 imply that the template has not been perfectly doubled. Template, primer, and probe quality and quantity, in addition to pipet calibration and buffer conditions like MgCl_2 , salt, additives, and deoxynucleotide concentrations, all contribute to this efficiency (Pfaffl, 2004; see *UNITS 15.1 & 15.5* for further details on these parameters). Given optimal buffer conditions and adequate primer, probe, and sample qualities, slight fluctuations in efficiency may still be observed between primer sets run on the same plate or between assay plates, even if they have been assembled and run on the same day by the same user. Efficiency correction is a means to account for inter- or intra-assay variability that is attributable to the aforementioned parameters.

The materials and setup of the assay are the same as in the standard curve method (see Basic Protocol 1), except that the linear regression formula produced by plotting the Ct versus the log ng of input standard total RNA is used only to determine PCR efficiency. This computed efficiency is then used to calculate the RNA levels (in arbitrary units) of the GOI and the endogenous control genes. The GOI RNA level in each sample is then expressed as a ratio relative to the endogenous control RNA level in that sample. Because the data are dependent upon Ct values and not an arbitrary standard curve, the resulting values may be compared to those of another RNA.

1. Set up and run assay as described in Basic Protocol 1 (steps 1 to 9). Refer to Figure 15.8.1 for an example of a typical plate setup and to Table 15.8.1 for master mix components.
2. Analyze and export raw data (see Basic Protocol 1, step 10), and then import into Microsoft Excel or equivalent spreadsheet program.

The threshold values for all RNAs measured (including the endogenous reference) must be the same. It is important to determine a suitable threshold within the geometric phase of the amplification plots for all RNA transcripts to be compared.

3. Calculate PCR efficiency, $E = 10^{(-1/\text{slope})}$, for the endogenous control RNA and each GOI from the slopes of their corresponding standard curves.
4. Calculate the quantity of the endogenous control RNA and each GOI from their Ct values according to the formula quantity = $E^{-\text{Ct}}$.

When the efficiency is 100% (i.e., slope = -3.3 and $E = 2$), the equation becomes quantity = $2^{-\text{Ct}}$. This serves as the basis for the calculation performed in $\Delta\Delta\text{Ct}$ method (Alternate Protocol).

5. For each of the three replicates of a sample, calculate the average quantity (avg), the standard deviation of the average (stdev), and the coefficient of variation (CV), where $\text{CV} = \text{stdev}/\text{avg}$.
6. Remove outliers, normalize the GOI, calculate the SD, and plot the results (see Basic Protocol 1, steps 13 to 17).

COMPARATIVE OR $\Delta\Delta$ Ct METHOD

The comparative Ct or $\Delta\Delta$ Ct method is used for measuring the fold-changes in expression of a particular RNA transcript between experimental samples. Typically, this assay is used when investigating gene-expression differences between wild-type and knockout or transgenic animals, or between vehicle-control and drug-treated samples. The results are then expressed as "fold-changes" relative to a calibrator, such as an untreated or wild-type sample. The $\Delta\Delta$ Ct method is only applicable when the primer sets for both the GOI and the endogenous reference gene have been shown to give perfect standard curve slopes (slopes = -3.3 ± 0.1 with $R^2 = 0.99$) as assayed in the Support Protocol.

1. Set up assay as described in Basic Protocol 1 (steps 1 to 9), except refer to Figure 15.8.2 for an example of a typical plate setup and to Table 15.8.2 for master mix components.

Standard cDNA samples are not needed in this method.

2. Analyze and export raw data (see Basic Protocol 1, step 10) and then import into Microsoft Excel or equivalent spreadsheet program.
3. For each of the three replicates of a sample, calculate the average (avg) cycle time (Ct) and then calculate the standard deviation (stdev).
4. Remove any outlier wells from the averaged Ct values (>0.3 stdev).

Only one point per replicate may be removed.

%CV may not be used, due to the logarithmic nature of both the Ct avg and Ct stdev. Instead, stdev must be used, where a stdev of 0.3 correlates with the maximum allowable standard deviation that can distinguish a two-fold change with 99% confidence; 0.4 stdev may be used for a 95% confidence interval.

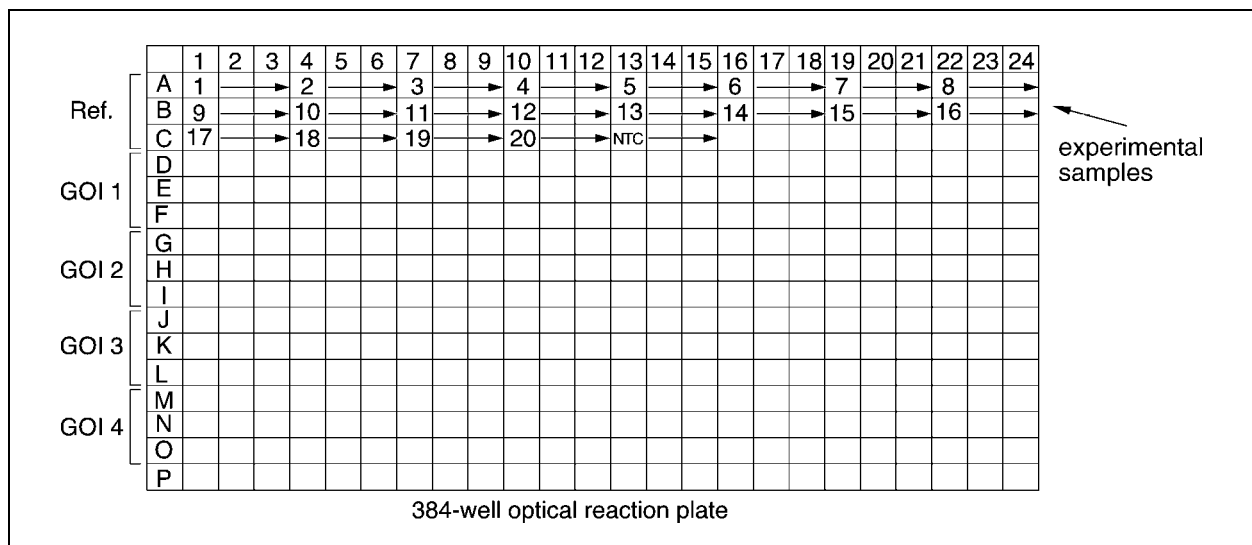


Figure 15.8.2 Typical plate setup for the $\Delta\Delta$ Ct method. The plate arrangement shown represents an experiment in which 20 samples taken from experimental animals will be assayed for one endogenous control gene and four GOIs. The samples are plated in triplicate for each of the RNAs assayed. Abbreviations: GOI, gene of interest; NTC, no-template control; Ref, endogenous control gene.

Table 15.8.2 Master Mixes for the $\Delta\Delta\text{Ct}$ Assay

Components	Final concentration	Volume per well	Volume for each cDNA/primer mix (per sample in triplicate + 1 extra)	Master mix (no. of samples + 1 NTC + 1 extra)
<i>SYBR Green assay</i>				
2× SYBR Green mix	1×	5 μl	20 μl	
1:1 primer mix (1.25 μM each)	150 nM each primer	1.2 μl	4.8 μl	
Template cDNA	10-25 ng ^a	1.25 μl	5 μl	—
H ₂ O	N/A	to 10 μl	to 40 μl	
Total volume		10 μl	40 μl	Aliquot 35 μl into each tube containing cDNA
<i>TaqMan assay</i>				
2× TaqMan mix	1×	5 μl	20 μl	
1:1 primer mix (1.25 μM each)	300 nM each primer	2.4 μl	9.6 μl	
5 μM TaqMan probe	250 nM	0.5 μl	2.0 μl	
Template cDNA	10-25 ng ^a	1.25 μl	5 μl	—
H ₂ O	N/A	to 10 μl	to 40 μl	
Total volume		10 μl	40 μl	Aliquot 35 μl into each tube containing cDNA

^aRecommended amount of template for detection of both high and low levels of GOIs. If necessary, significantly less template may be used (picograms). Note that the cDNA template quantity is based upon the amount of total RNA input into the reverse transcription reaction.

5. For each sample, normalize the GOI Ct values to those of the reference gene for the same sample according to the equation:

$$\Delta\text{Ct} = \text{avgCt}_{\text{GOI}} - \text{avgCt}_{\text{ref}}$$

Equation 15.8.3

Calculate the standard deviation of ΔCt ($\text{stdev}_{\Delta\text{Ct}}$) as:

$$\text{stdev}_{\Delta\text{Ct}} = \sqrt{(\text{stdev of reference})^2 + (\text{stdev of GOI})^2}$$

Equation 15.8.4

6. Choose a calibrator.

This will be the sample, tissue, gene, or control group to which the others will be compared.

7. Find the $\Delta\Delta\text{Ct}$, or calibrated value, for each sample, according to the equation:

$$\Delta\Delta\text{Ct} = \Delta\text{Ct}_{\text{sample}} - \Delta\text{Ct}_{\text{calibrator}}$$

Equation 15.8.5

The $\text{stdev}_{\Delta\Delta\text{Ct}}$ will be the same as $\text{stdev}_{\Delta\text{Ct}}$, since the calibrator is arbitrarily set to be a constant.

8. Find the fold-change for each sample relative to the calibrator according to the equation:

$$\text{fold-change} = 2^{(-\Delta\Delta Ct)}$$

Equation 15.8.6

For the sample that is chosen as the calibrator, the $\Delta\Delta Ct = 0$ and therefore the fold-change = $2^{(-\Delta\Delta Ct)} = 1$.

9. Plot the resulting fold-changes on a bar graph of fold-change versus sample name or experimental treatment group. Determine the measure of experimental error as:

$$SD_{\text{fold-change}} = (\ln 2)(\text{stdev}_{\Delta\Delta Ct})(2^{(-\Delta\Delta Ct)})$$

Equation 15.8.7

DESIGN AND VALIDATION OF SYBR GREEN AND TaqMan PRIMER/PROBE SETS

SUPPORT PROTOCOL

Considering that QPCR relies on the quality and the fidelity of the primers and probes that are used, very strict parameters for their design and subsequent validation are required. A common misconception in performing QPCR assays is that if the primer works for traditional end-point PCR, it is suitable for QPCR. In some cases this is true. However, the primer set must be tested in a QPCR validation assay before it can be used for RNA expression analysis. In keeping with the high-throughput capacity of QPCR, the thermocycling conditions are kept constant for all assays: 10 min at 95°C activates the hot-start *Taq* polymerase, followed by 40 cycles of the two-step 95°C melting and 60°C annealing. An extension step in the thermocycler program is not required, since all of the PCR products are 50 to 150 bp, thus making the run last only 1.5 to 2 hr. The concentrations of PCR reagents such as *Taq* DNA polymerase, MgCl₂, other salts, and dNTPs remain constant within the same chemistry (i.e., SYBR Green or TaqMan), and these so-called universal cycling conditions make primer and probe sequences the only point of flexibility in performing the assays.

The design of primer/probe sets requires the availability of reliable sequence information that may be obtained from databases like NCBI's GenBank or Ensembl, or from data produced by direct sequencing. The assay does not tolerate base mismatches between primer and template, especially in the probe sequence, a feature that allows for the detection of single-nucleotide polymorphisms (SNPs). Several primer/probe design software packages are available either for purchase or online. Otherwise, the user may design the primer sets by directly examining the sequence and choosing primers with the correct characteristics, as outlined in this protocol. The probe is labeled at the 5' end with a fluorescent reporter such as 6-FAM or VIC, and at the 3' end with a fluorescent or nonfluorescent quencher. The user should consult with the vendor that will synthesize the probe for the availability of each type of label.

The assay consists of a standard cDNA dilution series from which linear regression curves may be plotted. The slope of the resulting curve gives a measure of PCR efficiency, where -3.3 ± 0.1 with a coefficient of determination (R^2) of 0.99 indicates a reaction efficiency of 100%. Part of the initial SYBR Green validation also includes a melting (or dissociation) curve analysis. At the end of the repetitive cycles of the PCR, an additional melt-anneal-melt cycle is performed. The final melt occurs very slowly and

The Polymerase Chain Reaction

15.8.11

the changes in both temperature and fluorescent signal are monitored over time. This decrease in fluorescence correlates with the dissociation of the double-stranded PCR product releasing the bound SYBR Green I fluorophores. The instrument software uses an algorithm to transform and display the melting curve as the negative first derivative of the normalized fluorescence versus temperature (Applied Biosystems, 2001a). The presence of a single peak in the melting curve is indicative of a single PCR product, and occurs at the melting temperature of the product. Multiple peaks in this plot indicate that nonspecific products or primer dimers have been formed. Formation of a single product can be confirmed by running the PCR products on a 2% agarose gel following the QPCR run. Occasionally, when two products are observed, the second product may have been formed during the plateau phase, which would not affect quantitation. To confirm whether this has occurred, the QPCR run could be repeated and stopped during the exponential phase, and the reaction products run on an agarose gel. However, this is not feasible in practice because of the high-throughput nature of the assay. The best course of action when multiple products are observed in the dissociation curve is to redesign and validate a new primer set. Only primer sets that give a single peak in this curve should be used for experimental assays. Once a SYBR Green-based primer set has passed validation testing, the corresponding TaqMan probe is ordered and validated for PCR efficiency only. In rare cases, the SYBR Green assay conditions (e.g., primer concentration, Mg^{2+} concentration) will not be appropriate for TaqMan assays. This is observed as a decline in PCR efficiency. In this case, new primers may be designed to flank the probe sequence.

Additional Materials (also see *Basic Protocol 1*)

Primer/probe design software (Primer Express, Applied Biosystems)

Design primers

1. Retrieve the RNA sequence information from the appropriate source (e.g., Genbank or Ensembl).
2. Determine the locations of exon boundaries by aligning the mRNA sequence with its gene or by using NCBI's Entrez Gene Evidence Viewer (<http://www.ncbi.nlm.nih.gov>) or Ensembl's Genome Browser (<http://www.ensembl.org>).

Some genes do not have introns, so this step may not be applicable.

3. Copy the sequence into the design software.

Several design programs are available both commercially as stand-alone applications and as Web-based applications. Alternatively, primers and probes may be designed "by hand." Omit this step if designing by hand.

4. If using software other than Primer Express, use the following parameters:
 - a. *QPCR primers*: Should have 40% to 60% GC content and melting temperatures around 60°C. Should not contain runs of the same nucleotide, repetitive sequences, or more than two G's and/or C's on the 3' end (also called GC clamp).
 - b. *PCR product (amplicon)*: Should be 50 to 150 bases in length with an approximate melting temperature between 85° and 95°C.
 - c. *TaqMan probe (anneals to sequence between primers)*: Should have the same properties as the primers, except that the melting temperature should be around 70°C, and the sequence should not contain G's within a few bases of the 5' end because of increased reporter quenching. In addition, the sequence must have more C's than G's, which can be accomplished by using the complementary strand sequence for the probe (Applied Biosystems, 2002b).

Primer Express contains templates into which these parameters have been preloaded.

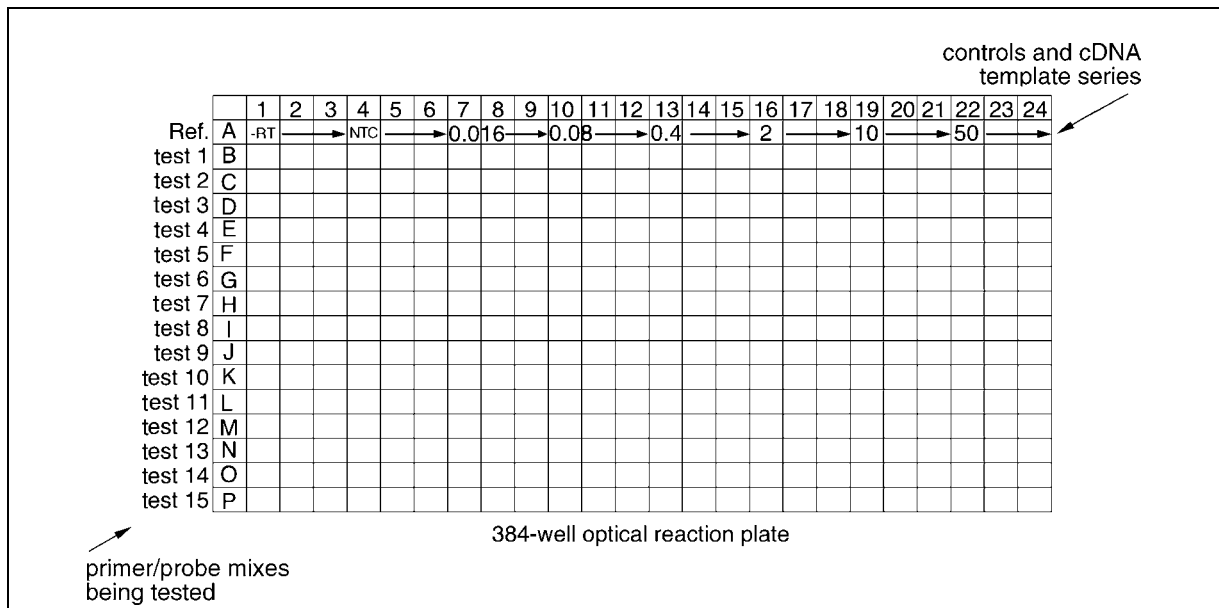


Figure 15.8.3 Typical plate setup for primer and probe validation assays. The plate arrangement shows a standard cDNA template dilution series that is being used to test 15 primer sets along with an endogenous control primer set that has already been validated. The same format is used when testing new probe sets. If different standard cDNAs are used to test primers/probes on the same plate (i.e., standards derived from different tissues), the validated endogenous reference primers/probes must also be run for that standard.

- Choose a primer/probe set for which the primers anneal in different exons, or which have less than 5-bp overhangs into the adjacent exon on the 3' end of the primer.

This step is necessary to avoid amplification of contaminating genomic DNA. Although the RNA is DNase-treated prior to reverse transcription, complete removal of genomic DNA is never achieved.

For intron-less transcripts or other primer sets that bind sequence within a single intron, the -RT controls are essential for each sample, to ensure that genomic DNA is not being amplified.

- Perform a BLAST (or equivalent) search of both primers of the set together to verify that they will fully anneal to the correct sequence and only that sequence.
- If the TaqMan probe will be used, run BLAST (*UNIT 19.3*) on the probe sequence to assess whether it binds to the correct sequence with 100% identity.
- Order a small-scale synthesis of the primers from a suitable vendor. Standard desalting of the primers is sufficient, and no additional purification (e.g., HPLC) is required.

Once a primer set has been validated, large-scale synthesis may be more cost effective, especially for frequently used primers like the reference genes.

- Validate the primer set according to the remaining steps of this protocol. If performing TaqMan-based assays, validate the primer set before ordering the probe. Once the primer set is validated, order the dual-labeled probe from an appropriate vendor, and validate according to the following steps.

Validate primer set

- Set up assay as described in Basic Protocol 1 (steps 1 to 9), except refer to Figure 15.8.3 for an example of a typical plate setup and to Table 15.8.3 for master mix components.
- Test the new primer set using SYBR Green chemistry. If valid, test the TaqMan-based chemistry.

Table 15.8.3 Master Mixes for Primer/Probe Validation Assays

Components	Final concentration	Volume per well	Volume for each cDNA/primer mix (per sample in triplicate + 1 extra)	Master mix (for 8 standards ^b + 1 extra)
<i>SYBR Green assay</i>				
2× SYBR Green mix	1×	5 μl	20 μl	180 μl
1:1 primer mix (1.25 μM each)	150 nM each primer	1.2 μl	4.8 μl	43.2 μl
Template cDNA	0.016-50 ng ^a	1.25 μl	5 μl	—
H ₂ O	N/A	to 10 μl	to 40 μl	91.8 μl
Total volume		10 μl	40 μl	Aliquot 35 μl into each tube containing cDNA
<i>TaqMan assay</i>				
2× TaqMan mix	1×	5 μl	20 μl	180 μl
1:1 primer mix (1.25 μM each)	300 nM each primer	2.4 μl	9.6 μl	86.4 μl
5 μM TaqMan probe	250 nM	0.5 μl	2.0 μl	18 μl
Template cDNA	0.016-50 ng ^a	1.25 μl	5 μl	—
H ₂ O	N/A	to 10 μl	to 40 μl	30.6 μl
Total volume		10 μl	40 μl	Aliquot 35 μl into each tube containing cDNA

^aRecommended template dilution series. The user may modify the range of cDNA concentrations based on the experimental system. Note that the cDNA template quantity is based upon the amount of total RNA input into the reverse transcription reaction.

^bFor this assay, the NTC and –RT controls are included as part of the standard sample set.

12. Following the instrument run, first check the dissociation curve. If more than one peak is present, the primer set is invalid and no other parameters are checked.
13. If a single peak is found in the dissociation curve, assess the PCR efficiency by calculating the slope of the linear regression curve as follows.
 - a. Plot Ct (or crossing point, CP) versus log ng of the standard cDNA input for each concentration of standard as an *xy* scatter plot.

This may be performed directly in the instrument software on the Applied Biosystems instrument, or can be done in Excel (or equivalent).

- b. Apply a best-fit curve and display the corresponding linear regression formula.

If the slope of the curve is -3.3 ± 0.1 with $R^2 = 0.99$, the primer set amplifies at 100% efficiency, and the set is considered valid.

Efficiency is dependent upon several factors including pipet calibration, primer quality and dilution, and even variability in the instrument run. The slope, therefore, may not be exactly -3.3 ± 0.1 . In this case, the slope of the test primer set must match that of a previously validated endogenous reference gene run for the same standard cDNA within ± 0.1 . For example, if the test primer gives a slope of -3.6 and the endogenous reference for the standard cDNA used to test the set gives a slope of -3.5 , then the test set is valid. However, if the endogenous reference primer set gives a slope of -3.3 and the test primer has a slope of -3.6 , the test set is invalid.

14. Repeat the linear regression analysis for the TaqMan probe/primer set to assess PCR efficiency.

Melting curve analysis cannot be performed for TaqMan-based assays, since cleavage of the probe releases the reporter that continuously fluoresces.

REAGENTS AND SOLUTIONS

Use molecular-biology-grade (nucleic acid- and nuclease-free) or sterile-filtered double-deionized water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Primer mixes, 1.25 μ M each forward and reverse primer

Mix small aliquots of 2.5 μ M forward and reverse primer stocks (see recipe) in equal volumes (1:1). Store up to 1 to 2 months at 4°C in screw-capped tubes to prevent evaporation. Do not freeze, as repeated freeze-thaw cycles will degrade the primers.

Primer stocks, 100 and 2.5 μ M

Purchase lyophilized oligonucleotides from any commercial source (synthesized at a 25-nM scale; standard desalting is sufficient, no additional purification is needed). Briefly centrifuge the tubes of powdered oligonucleotide to get contents to the bottom. Resuspend in molecular-biology-grade water to 100 μ M. Before preparing primer mixes, dilute each stock (forward and reverse) to 2.5 μ M. Store at -20°C. When stored properly and subjected to minimal freeze-thaw cycles, primer stocks can last >2 years.

Primer pairs produced at a 25-nM scale will yield enough reagent to test ~4000 samples using SYBR Green, or ~2000 samples using TaqMan assays.

Sterile-filtered double-deionized water may be used instead of purchased molecular-grade water. Do not use DEPC-treated water, because the slightly acidic pH may promote primer degradation. Tris buffer may be used instead of water, but should not contain EDTA, which acts as a Mg²⁺-chelating agent and can inhibit the PCR.

Standard cDNAs, dilution series

DNase-treat (UNIT 3.12) and reverse transcribe (UNIT 3.7) a suitable RNA using 0.08 μ g/ μ l random hexamer primers such that the final concentration of the standard is 40 ng/ μ l (based on RNA quantity; see UNIT 15.5). Include both a no-template control (NTC) for which water is used instead of RNA, and a no-reverse-transcriptase control (-RT). Following the reverse transcription, make a 5-fold dilution series of the 40 ng/ μ l standard to obtain working concentration standards of 40, 8, 1.6, 0.32, 0.064, and 0.0128 ng/ μ l. Store up to 1 month at 4°C or >2 years at -20°C (with minimal freeze-thaw cycles).

By using 1.25 μ l/well of each of the standards in the final reaction plate, the resulting amount of starting template will be 50, 10, 2, 0.4, 0.08, and 0.016 ng. See Commentary for more information and suggestions about control and standard RNA samples.

SYBR Green mix, 2 \times

Combine the following components as indicated in Table 15.8.4:

25 mM MgCl₂, molecular biology grade (store at -20°C)

10 \times Gold PCR buffer (supplied with PCR enzyme; Applied Biosystems; store at -20°C)

10 \times dNTP mix: equal volumes of 2 mM dATP, dTTP, dCTP, and dGTP (store at -20°C)

Dimethylsulfoxide (DMSO), molecular biology grade (store at room temperature)

continued

Table 15.8.4 Preparation of 2× SYBR Green Mix

Component	Volume (μl)				
25 mM MgCl ₂	48	288	600	720	1200
10× Gold PCR buffer	40	240	500	600	1000
10× dNTP mix	40	240	500	600	1000
DMSO	40	240	500	600	1000
1:1000 SYBR Green I	20	120	250	300	500
50× ROX	8	48	100	120	200
AmpliTaq Gold polymerase (5 U/μl)	2	12	25	30	50
H ₂ O	2	12	25	30	50
Total volume in 2× buffer	200	1200	2500	3000	5000

SYBR Green I dye (Molecular Probes; store at −20°C), diluted 1:1000 in water

50× ROX passive reference dye (Invitrogen; store at −20°C)

AmpliTaq Gold polymerase (5 U/μl; Applied Biosystems; store at −20°C)

Water, molecular biology grade

Prepare fresh and keep on ice prior to use in primer master mixes. Protect dyes and all prepared mixes from prolonged exposure to light by wrapping tubes in foil.

To maintain the high-throughput nature of the assay, all buffer conditions, including the concentrations of Mg²⁺, dNTPs, and other additives, are kept constant. It is strongly recommended that a preformulated buffer be purchased from a reliable vendor, since these are stable at room temperature and have been quality-control tested to ensure optimal performance. The authors have obtained comparable results using mixes from Applied Biosystems, Bio-Rad, Invitrogen, and Sigma.

TaqMan probe, 100 and 5 μM

Depending on the vendor, the probe may be supplied in a lyophilized form. In this case, resuspend to 100 μM in Tris·Cl, pH 8.0 (APPENDIX 2), prepared with molecular-biology-grade water and reagents. Before use, dilute a small amount of 100 μM stock to 5 μM with more Tris·Cl, pH 8.0. Store either concentration at −20°C. Avoid repetitive freeze-thaw cycles, and thaw on ice prior to use to preserve the integrity of the probe. Protect the probe from excessive exposure to light (e.g., using amber-colored screw-cap tubes) to prevent photobleaching of the fluorescent dyes and evaporation. See Support Protocol and Commentary for additional considerations regarding the TaqMan probe.

COMMENTARY

Background Information

Quantitative PCR is a rapid, robust, and highly sensitive polymerase chain reaction method used to quantify specific nucleic acid targets. Real-time quantitative PCR is different from end-point, or in-gel, analysis (see UNIT 15.7) in several ways. For real-time analysis, the increase in fluorescent signal resulting from PCR product synthesis is recorded during the course of the thermocycle. This allows the user to specify the point in the assay at which to "read" the data. Measurements are obtained from the geometric phase of the amplification

reaction. This is the phase during which all of the components required for the PCR (e.g., dNTPs, primers, polymerase) are in excess, and therefore the deficit of an essential reaction component will not quench the efficiency of product synthesis. Following geometric amplification, the fluorescence curve reaches a plateau (i.e., the saturation point) as the reaction components begin to become limited and the kinetics of the reaction become unpredictable. At this stage, an increase of one thermocycle no longer correlates with a two-fold change in product (Applied Biosystems,

2002a). In contrast, in end-point PCR, quantitation is often obtained by in-gel densitometry measurements at the end (or the saturation point) of the reaction. The unpredictable nature of the PCR at this stage makes this method semiquantitative at best.

Real-time PCR depends both on a set of universal thermocycling and buffer conditions and on primer efficiency testing and correction where necessary. As a result, the accuracy and precision of the resolution (smallest detectable fold-change) is less than 2-fold, whereas resolution for end-point in-gel measurement is limited to about 10-fold (Applied Biosystems, 2002a). Semiquantitative PCR assays (in which samples are removed at measured cycle times and electrophoresed and possibly hybridized) have better resolving power than end-point analysis. However, both in-gel methods suffer from the lengthy processing steps, compared to real-time PCR.

The two most widely used fluorescent detection methods, or chemistries, for QPCR are SYBR Green, a DNA-intercalating dye, and the fluorogenic probe. Several types of fluorogenic probes are currently available: the popular, dual-labeled hydrolysis probe (TaqMan probe) and the hybridization probes known as molecular beacons and scorpions. Both types of probes bind the sequence intervening the forward and reverse primer binding sites, and both rely on fluorescence resonance energy transfer (FRET) to silence the signal from the reporter dye while it is in proximity to the quencher dye. A hydrolysis probe is cleaved by the 5' nuclease activity of the polymerase during the primer-extension phase of the reaction, and the reporter is released and becomes free to fluoresce continuously. Hybridization probes rely on a stem-loop structure to keep the reporter and quencher in proximity. Upon hybridization to the specific sequence, the distance between the reporter and quencher becomes too large to silence the reporter, and signal is detected (Tyagi and Kramer, 1996; Whitcombe et al., 1999).

There are advantages and disadvantages to each type of chemistry (SYBR Green versus TaqMan) in terms of sensitivity and specificity. SYBR Green I binds any double-stranded DNA and does not depend on a probe-cleavage event. Therefore, SYBR Green produces earlier Ct values, resulting in an apparent enhanced sensitivity (Whittwer et al., 1997; Morrison et al., 1998). TaqMan probes, on the other hand, supply another layer of sequence specificity in addition to the forward and reverse primers.

There are two methods of quantitation that may be performed using real-time PCR: absolute and relative. Absolute quantitation allows for the assessment of the copy number of a specific nucleic acid target in a sample, while relative quantitation results in the ratio of a specific target relative to an endogenous reference gene. The absolute method requires the availability of a standard sample in which the copy number of the particular target has been carefully and accurately measured. From this standard sample, a dilution series is made and assayed for the target at the same time as the unknown samples. From the values obtained from linear regression analysis of the standard dilution series, the GOI copy number values may be interpolated. This will allow the user to assess the sensitivity (i.e., the lowest detectable copy number) of the GOI primer set. As an example, this type of analysis is used in both clinical and food science for the assessment of pathogen load and gene copy number (Pfaffl, 2004).

Relative quantitation, which is covered in detail in protocols in this unit, does not rely on the knowledge of a given transcript copy number in a standard sample. Instead, the changes in gene expression or the levels of a specified transcript may be measured and described as an arbitrary unit relative to some control sample ($\Delta\Delta\text{Ct}$ or standard curve method), or to the level of some other control transcript in the same sample (standard curve or efficiency-corrected ΔCt method). For example, relative quantitation allows for the measurement of the fold-change in expression for gene A in a treated sample versus an untreated sample, or for the assessment of the level of gene A in a sample relative to some housekeeping gene in the same sample.

Critical Parameters and Troubleshooting

Primer/probe design and validation

For RNA analysis, it is important to differentiate message from genomic DNA. In this respect, the use of amplicons that span an exon junction allows this requirement to be met. This is achieved by designing primer sets with the forward and reverse primers sitting in different exons. However, when the intervening intron is small, genomic DNA may still be amplified. This can be avoided by allowing a few base pairs on the 3' end of either primer to overhang onto the neighboring exon.

To distinguish knockout or mutant samples using QPCR, primers/probes should be

designed in the knocked-out or mutated region of the transcript. In this way, no amplification of the transcript will occur for the knockout, since the region recognized by the primers has been deleted or altered. Primers for transcripts that are rapidly degraded should be placed near the 5' end of the RNA sequence, as degradation by ribonucleases generally occurs 3' to 5' (Brown, 2002). However, RNA transcripts that have undergone linear RNA amplification (for instance, RNA isolated from laser capture microdissected cells) are around 200 to 1000 bases in length and represent the 3' ends of the transcripts. In this case, primers must be designed in a region near the poly(A)⁺ tail.

It is absolutely mandatory to validate both primers and probes on the same instrument used to perform the experimental assay, i.e., when the instrument used to validate the primer sets is a different brand than the one that will be used for the experimental assays (e.g., Roche iCycler versus ABI 7900HT). Published primer/probe sets for which the PCR product is ≤ 150 base pairs and the annealing temperature is around 60°C, which have been validated properly, should be transferable to any system, but this must be empirically determined to ensure the reliability of the data.

In rare cases, the PCR efficiency of a validated primer set changes when switched from SYBR Green to TaqMan-based chemistry. This is due to the difference in primer concentration and/or MgCl₂ concentrations between the two buffers, and may often be overcome by redesigning the primers to recognize a sequence around the already synthesized probe.

Choice of chemistry

The selection between SYBR Green I and TaqMan-based assays depends upon the RNA sequence. If the RNA of interest has no polymorphisms or other variations in the region to which the primers bind, and the primers have been correctly designed and validated, then SYBR Green-based assays are adequate for gene-expression analyses. It is up to the user to decide if the addition of the often-costly TaqMan probe is worth the additional specificity it confers. In the case of polymorphisms or variants, differentiation between different RNA species may require the specificity of the probe, which can discriminate a single base difference.

Template quality

The quality of the cDNA template depends upon the integrity of the RNA. QPCR will tol-

erate some degradation of the RNA when random hexamers (or other -mers) are used to prime the reverse transcription reaction. However, it is not good laboratory practice to use degraded RNA, and the cause of the degradation should be addressed. Transcripts decay at different rates and have variable stability, so partial degradation of a sample at any point could lead to complete absence of detection of the desired target RNA in a subsequent QPCR assay, with little to no change in the house-keeping gene used as a control.

Total RNA is used for QPCR to reduce the number of steps and potential sources of degradation during sample preparation. Purified messenger, or poly(A)⁺, RNA can also be used. However, this subtractive purification could lead to the loss of transcripts that do not have a poly(A)⁺ tail, or to the preferential enrichment of RNAs that have internal A tracts. The added processing reduces the recovery of material for subsequent use, and can cause degradation. If poly(A)⁺ RNA is used, 50- to 100-fold less (~100 pg) sample template is required in the reaction mixture because mRNA represents 3% of total cellular RNA (Alberts et al., 1994).

Reverse transcription primers

Reverse transcription (RT) of RNA into complementary DNA (cDNA; refer to *UNIT 15.5*) may be performed using several different types of oligonucleotide primers. For QPCR, the preferred primer is a random hexamer, nonamer, or dodecamer oligo with 6, 9, or 12-base stretches of random sequences, respectively. Random primers have a much higher probability of efficiently amplifying all RNA transcripts, due to their indiscriminate nature (ABI, pers. comm.). This property also enables RNA secondary structure to be overcome, since the priming occurs in random places along the length of the transcript. mRNA-specific priming by oligo(dT)s at the poly(A)⁺ tail and any internal poly(A) tract is another method of RT priming. This method will allow only polyadenylated RNAs to be converted to cDNA, thus limiting amplification of some GOIs or partially degraded samples. A mixture of a random oligo with an oligo(dT) primer may enhance detection of rare messages while still allowing for the detection of transcripts that lack polyadenylation. However, the use of this procedure may skew the measurement of relative RNA abundance towards intact, full-length mRNA over incomplete or rapidly degraded messages. Users should decide and test these parameters

in their particular experimental systems. The third method of reverse transcription priming is the gene-specific reverse primer. The reverse PCR primer is used to specifically target the GOI for conversion to cDNA. This is often performed in the same QPCR plate as the PCR by adding RT enzyme to the PCR mix and adding an incubation step prior to the first step in the PCR cycling program (known as one-step RT-PCR). While this may enhance detection of a specific RNA target, RNA secondary structure may not be overcome, depending on the priming site, and RT efficiency must be considered in addition to PCR efficiency.

An important assumption that is made when performing RT-QPCR is that RT efficiency is similar for the GOI between samples and for different GOIs. However, different tissues/sample types may contain variable levels of RT-inhibiting or -enhancing factors (Pfaffl, 2004). To control for these variables, it is recommended that all samples to be compared be prepared under the same conditions and at the same time (i.e., with regard to RNA extraction method and RT reaction).

Endogenous reference gene

Normalization of sample loading is essential in any quantitative comparative analysis to ensure that the measured differences between samples is not attributable to disproportionate amounts of starting material. For gene-expression assays, the normalizer must be an endogenous gene that is expressed at equal levels in all tissue or cell types and treatment conditions under study. Traditionally, one of the so-called housekeeping genes (e.g., GAPDH, cyclophilin, β -actin, HPRT, U36B4, 18S rRNA) is selected to serve this function. The choice is a point of controversy, since there are examples of fluctuations for most of the abovementioned genes under various treatment or physiological conditions (Schmittgen and Zakrajsek, 2000; Suzuki et al., 2000; Guo et al., 2001; Vandesompele et al., 2002; Dheda et al., 2004). The user should run a small pilot experiment to determine which endogenous reference is appropriate for the particular studies. A good method for doing this is to perform a $\Delta\Delta C_t$ assay using several potential normalizer RNAs and a GOI that is not expected to show any fold-changes within a small set of samples. The GOI is then normalized to each of the reference RNAs individually for all of the samples. Finally, the housekeeping gene for which there is no detectable fold-change of the GOI between the test samples is chosen for use in experimental assays (Applied

Biosystems, 2001b; Guo et al., 2001; Roche Applied Science, 2002).

Controls and standards

Controls for the assay that are made alongside the cDNA standards and unknowns include a no-template control (NTC), made by substituting water for RNA, and a no-reverse-transcriptase (-RT) control, made by omitting the reverse transcriptase. For primers that span an exon junction, a -RT control is not needed for every sample if, during the validation process, this control shows no amplification product. If the system under study is the result of introduction of an expression construct into cells, or if the GOI does not have introns or has a known processed pseudogene, then -RT controls should be made and assayed for every sample. An NTC should be run for every primer set of an assay to facilitate the detection of contaminants that contribute to fluorescent signal.

If amplification of the NTC occurs, primer-dimer or other nonspecific PCR products may have been formed, or contamination of a reagent or degradation of the primer mix may have taken place. If amplification occurs in the -RT control, this indicates the presence of genomic DNA if the primer/probe set does not span an exon junction, or the presence of primer-dimer or nonspecific PCR products, a contaminant, or degradation of the primer mix. Although all amplification products (including those that are nonspecific) contribute to the fluorescence signal, this fluorescence contribution can be considered negligible if the C_t values of the NTC and -RT are ≥ 7 cycles different from the experimental samples (Applied Biosystems, pers. comm.).

If the unknowns are RNA samples, the standard of choice is an RNA that is reverse transcribed in the same manner as the unknowns. A suitable standard RNA is one in which the expression of the GOI is at a moderate to high level and that has a similar composition to the unknown samples. The use of a plasmid or linear DNA standard is not advised for measurement of an endogenous tissue transcript, because these types of nucleic acids have different background compositions than RNA, are extracted differently, and are not reverse transcribed, and therefore may not have the same RT or PCR amplification efficiency as the experimental samples (Applied Biosystems, 2003; Pfaffl, 2004). Several vendors (e.g., Ambion, Clontech, Stratagene) supply total RNA preparations of many cell and tissue types collected from many different

species. Both Stratagene and Clontech make total RNA pools termed "universal reference RNA," composed of mixtures of either various cell lines or whole tissues, respectively. These RNA pools represent $\geq 90\%$ gene coverage on microarrays (Novoradovskaya et al., 2000; Clontech, 2002), and are very useful as standards both for validation of primer/probe sets and for large-scale multigene studies.

Anticipated Results

The setup of the $\Delta\Delta C_t$ assay allows 64 samples (one of which should be the NTC) to be assayed for an endogenous reference gene and one gene of interest in triplicate; the standard curve and efficiency-corrected ΔC_t assays will accommodate an endogenous reference gene and one gene of interest in triplicate for 56 unknowns, six standards, and two control samples. Assays performed on the ABI 7900HT are completed in 1.5 and 2 hr for TaqMan and SYBR Green-based assays, respectively. The results yield data that are highly reproducible and correlate well with traditional northern blotting and RNase-protection assays.

Time Considerations

In preparation for performing a QPCR assay, RNA and subsequent cDNA preparation may be carried out in advance. Prior to an actual experimental assay, primers and probes must be designed and validated. In most cases, design and validation may take several days to a few weeks. This includes the time to design, order, and synthesize the primers (~ 2 days, depending on the vendor), test the primers (a few hours), and, if desired, synthesize (~ 7 to 14 days, depending on the vendor) and test (a few hours) the corresponding probes.

Once the required primer/probe sets are validated, the experimental assays are performed. A single-plate assay may take 0.5 to 2 hr to prepare and 2 hr to run. The preliminary raw-data analyses on the instrument software may take less than half an hour. The time required for the following final analyses will depend upon the user's familiarity with both the mathematical and software applications. The typical workflow for an experimental assay is as follows:

Day 1: Prepare total RNA and determine concentration by UV or fluorescence spectroscopy (1 to 4 hr, depending on number of samples and method of preparation).

DNase-treat and reverse transcribe RNA to cDNA (~ 3 hr for setup and incubations).

Day 2: Prepare master mixes for the assay (~ 30 min to 1 hr).

Prepare QPCR plate(s) (~ 30 min to 1 hr per plate).

Run plate on instrument (1.5 to 2 hr).

Collect and analyze data (1 to 3 hr).

Literature Cited

- Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J.D. 1994. The cell nucleus. *In* Molecular Biology of the Cell, 3rd ed., p. 370. Garland Publishing, New York.
- Applied Biosystems. 2001a. ABI Prism 7900HT User Manual, rev. 4. Applied Biosystems, Foster City, Calif.
- Applied Biosystems. 2001b. TaqMan human endogenous control plate: Protocol: Revision C. <http://docs.appliedbiosystems.com/pebi/docs/04308134.pdf>. Applied Biosystems, Foster City, Calif.
- Applied Biosystems. 2002a. Real-time PCR vs. traditional PCR. <https://www.appliedbiosystems.com/support/tutorials/pdf/rtpcr-vs-tradpcr.pdf>. Applied Biosystems, Foster City, Calif.
- Applied Biosystems. 2002b. Designing TaqMan MGB probe and primer sets for gene expression using Primer Express software. http://www.appliedbiosystems.com/support/tutorials/pdf/taqman_mgb_primersprobes_for_gene_expression.pdf. Applied Biosystems, Foster City, Calif.
- Applied Biosystems. 2003. Creating standard curves with genomic DNA or plasmid DNA templates for use in quantitative PCR. https://www.appliedbiosystems.com/support/tutorials/pdf/quant_pcr.pdf. Applied Biosystems, Foster City, Calif.
- Brown, T.A. 2002. How genomes function. *In* Genomes, 2nd ed. (S. Carlson, ed.) section 10.4. John Wiley & Sons, Hoboken, N.J.
- Clontech. 2002. Control RNA for microarray experiments. *Clontechniques* XVII:6. <http://www.clontech.com/clontech/archive/APR02UPD/pdf/ControlRNA.pdf>. Clontech, Palo Alto, Calif.
- Dheda, K., Huggett, J.F., Bustin, S.A., Johnson, M.A., Rook, G., and Zumla, A. 2004. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *BioTechniques* 37:112-119.
- Guo, D., Henriksson, R., and Hedman, H. 2001. The iCycler iQ detection system for evaluating reference gene expression in normal human tissue, rev. A. *Amplification* 2804. http://www.bio-rad.com/LifeScience/pdf/Bulletin_2804.pdf. Bio-Rad, Hercules, Calif.
- Morrison, T.B., Weis, J.J., and Wittwer, C.T. 1998. Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *BioTechniques* 24:954-962.
- Novoradovskaya, N., Payette, T., Novoradovskaya, A., Braman, J., Chin, N., Pergamenschikov, A., Fero, M., and Botstein, D. 2000. Pooled, high-quality reference RNA for human microarrays. *Strategies* 13:121-122. <http://www.stratagene.com/news/newsletter.aspx?iid=6>. Stratagene, La Jolla, Calif.

Pfaffl, M.W. 2004. Quantification strategies in real-time PCR. *In* IUL Biotechnology Series, No. 5: A-Z of Quantitative PCR (S.A. Bustin, ed.) pp. 87-120. International University Line, La Jolla, Calif.

Roche Applied Science. 2002. Selection of housekeeping genes. Technical Note No. LC 15/2002. http://www.roche-applied-science.com/lightcycler-online/lc_support/pdfs/lc_15.pdf. Roche Applied Science, Indianapolis, Ind.

Schmittgen, T.D. and Zakrajsek, B.A. 2000. Effect of experimental treatment on housekeeping gene expression: Validation by real-time, quantitative RT-PCR. *J. Biochem. Biophys. Methods* 46:69-81.

Shoemaker, J.P., Garland, C.W., and Steinfeld, J.I. 1974. Experiments in Physical Chemistry, pp. 34-39. McGraw-Hill, New York.

Suzuki, T., Higgins, P.J., and Crawford, D.R. 2000. Control selection for RNA quantitation. *BioTechniques* 29:332-337.

Tyagi, S. and Kramer, F.R. 1996. Molecular beacons: Probes that fluoresce upon hybridization. *Nat. Biotechnol.* 14:303-308.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Psepe, A., and Speleman, F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3:1-12.

Whitcombe, D., Theaker, J., Guy, S.P., Brown, T., and Little, S. 1999. Detection of PCR products using self-probing amplicons and fluorescence. *Nat. Biotechnol.* 17:804-807.

Whittwer, C.T., Herrmann, M.G., Moss, A.A., and Rasmussen, R.P. 1997. Continuous fluorescence monitoring of rapid cycle DNA amplification. *BioTechniques* 22:130-138.

Key References

Ambion. 2001. The top 10 most common quantitative RT-PCR pitfalls. *Technotes Newsletter* 8:8. Ambion, Houston, Tex.

A short, but useful checklist of critical considerations for performing any type of reverse transcription PCR.

Applied Biosystems. 1997. Relative quantitation of gene expression: ABI PRISM 7700 Sequence Detection System: User Bulletin #2: Rev B. Applied Biosystems, Foster City, Calif.

This bulletin outlines both the standard curve and $\Delta\Delta C_t$ methods and shows, by comparing data obtained using both calculations, that the resulting values are very similar regardless of the assay.

Applied Biosystems. 2001a. See above.

This instrument manual contains explanations about the transformation of fluorescence signal into Ct data in addition to outlining the proper method of baseline and threshold settings for ABI machines. Users should consult their specific instrument manuals, since each type of instrumentation will require knowledge of slightly different terminology and parameters.

Livak, K.J. and Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_t}$ method. *Methods* 25:402-408.

This article presents a detailed review of the derivation of the mathematical applications described in this unit.

Internet Resources

<http://www.ncbi.nlm.nih.gov>

NCBI Web site.

<http://www.ensembl.org>

Ensembl Web site.

<http://www.gene-quantification.info/>

The Gene Quantification Web site contains a host of information concerning QPCR.

<http://pga.mgh.harvard.edu/primerbank/index.html>

The Primer Bank database, hosted by Harvard University, contains user-submitted primer sequences for several mouse and human genes.

<http://web.ncifcrf.gov/rtp/gel/primerdb/>

The Quantitative PCR Primer Database (QPPD), maintained by the National Cancer Institute, contains primer and probe sequences for mouse and human genes collected from articles cited in PubMed.

Commercial vendor-sponsored technical support Web sites are also a very good resource for tips about RNA and QPCR applications.

Contributed by Angie L. Bookout,
Carolyn L. Cummins, and
David J. Mangelsdorf
Howard Hughes Medical Institute
University of Texas Southwestern
Medical Center
Dallas, Texas

CHAPTER 16

Protein Expression

INTRODUCTION

Protein expression, as used in this chapter, refers to the directed synthesis of large amounts of desired proteins. In early applications, molecular biologists interested in obtaining great quantities of prokaryotic regulatory proteins arranged their synthesis in large amounts, a process that came to be called overproduction, expression, or overexpression. These early techniques used genetic manipulations to select *in vivo* recombination events that inserted the desired gene into bacteriophages. Later, as it was developed, recombinant DNA technology was used to create phages and plasmids *in vitro*, which directed the synthesis of large amounts of the products of cloned genes.

This chapter describes methods to express proteins. In all these methods, a gene whose product is to be expressed is introduced into a plasmid or other vector, and that vector is introduced into living cells. Typical expression vectors contain promoters that direct the synthesis of large amounts of mRNA corresponding to the gene. They may also include, for example, sequences that allow their autonomous replication within the host organism, sequences that encode genetic traits allowing cells containing the vectors to be selected, and sequences that increase the efficiency with which the mRNA is translated.

There are three basic strategies used to express proteins, each of which has advantages. *E. coli* expression techniques are probably the most popular: the organism is already used by most investigators, the techniques necessary to express usable amounts of protein are relatively simple, the amount of time necessary to generate an overexpressing strain is very short, and a familiarity with standard recombinant DNA techniques is all that is necessary to begin pilot expression experiments. *E. coli* has other advantages that have made it widely used for expression of commercially important proteins: it is cheap to grow, and the vast body of knowledge about it has made it possible to tinker intelligently with its genetics and physiology, so that strains producing 30% of their total protein as the expressed gene product can often be obtained. However, expression in *E. coli* does have disadvantages. First, eukaryotic proteins expressed in *E. coli* are not properly modified. Second, proteins expressed in large amounts in *E. coli* often precipitate into insoluble aggregates called “inclusion bodies,” from which they can only be recovered in an active form by solubilization in denaturing agents followed by careful renaturation. Third, it is relatively difficult to arrange the secretion of large amounts of expressed proteins from *E. coli*, although it has often been possible to secrete small amounts into the periplasmic space and to recover them by osmotic shock.

The baculoviral expression system has a number of advantages that have contributed to its recent popularity: proteins are almost always expressed at high levels; expressed proteins are usually expressed in the proper cellular compartment (that is, membrane proteins are usually localized to the membrane, nuclear proteins to the nucleus, and secreted proteins secreted into the medium); and the expressed protein is often properly modified. Expression using baculoviral vectors also has some drawbacks: the techniques to grow and work with the virus are more sophisticated than those used with *E. coli*, and thus the system is more prone to “crashing” due to cell growth difficulties; the expressed proteins

are not always properly modified; and generation of a recombinant baculovirus to express a given protein takes several steps.

Compared with the above systems, all mammalian expression techniques have certain advantages, particularly for the expression of higher eukaryotic proteins: expressed proteins are usually properly modified, and they almost always accumulate in the correct cellular compartment. Generally speaking, mammalian expression techniques are more difficult, time-consuming, and expensive than those used to express proteins in *E. coli*, and they are much more difficult to perform on a large scale; however, they are quite practical for small- and medium-scale work by investigators already familiar with mammalian cell culture techniques. The several procedures described here for mammalian expression can produce transient expression, stable expression, or inducible expression. Each technique has specific advantages, and considerable care should be taken in deciding which technique to pursue (see below) because of the amount of time required to establish these techniques.

GENERAL OVERVIEW

Section I (UNITS 16.1-16.8) describes techniques for expressing proteins in *E. coli*. UNIT 16.1 contains an introduction to *E. coli* expression. UNIT 16.2 describes the use of T7 vectors, in which synthesis of large amounts of foreign gene products is directed by the phage T7 gene 10 promoter, which uses T7 RNA polymerase. This polymerase transcribes the gene 10 promoter so efficiently that it uses up most of the ribonucleotide triphosphates in the cell and drastically inhibits transcription of genes by the host polymerase. UNIT 16.3 describes the use of p_L -derived vectors and their appropriate host strains. These vectors carry the powerful bacteriophage p_L promoter and take advantage of a number of other useful aspects of phage lambda biology. The next units contain techniques for expression of fusion proteins in which the expressed protein carries an additional stretch of amino acids at its N terminus to aid its expression and purification. UNIT 16.4 introduces the concept of fusions and provides methods for cleavage of fusion proteins. UNITS 16.5, 16.6, 16.7 & 16.8 describe techniques for expressing *lacZ* protein (β -galactosidase), *trpE* protein, maltose-binding protein, glutathione-S-transferase (GST), and thioredoxin fusions.

Section II (UNITS 16.9-16.11) describes the use of the baculovirus system. In this system, genes for proteins to be expressed are inserted into an insect virus in lieu of a highly expressed dispensable gene. The foreign protein is produced by growing the recombinant virus in cultured insect cells. UNIT 16.9 introduces the system. UNIT 16.10 describes how to grow the cultured insect cells and viral stocks and how to isolate recombinant baculoviruses and use them to produce the desired protein. Finally, UNIT 16.11 describes the optimization of protein expression, first on a small scale, then maximized for large-scale production, and then provides instructions for purifying the recombinant proteins using GST- and His-tag systems.

Section III (UNITS 16.12-16.24) describe techniques for expressing proteins in mammalian cells. UNIT 16.12 describes expression using COS cell vectors. In this approach, vectors containing the gene to be expressed are transiently transfected into COS cells, which constitutively produce SV40 large T antigen. COS cell vectors contain an SV40 replication origin; when they are transfected into COS cells, they replicate, and protein is expressed from mRNA synthesized by hundreds of copies of the vectors. UNITS 16.13 & 16.14 describe expression from stably transfected cell lines. UNIT 16.13 describes the fusion of a protein of interest to an epitope tag, and the production of a stable cell line that expresses the tagged protein. Cell lines of this type are widely used to purify either individual proteins, or more frequently intact complexes that contain as one component the tagged protein. This strategy does not always work, as in certain cases even slight misexpression of a

protein can impair cell growth. To get around this problem, inducible systems are used; one widely used system uses the bacterial Tet repressor protein to engineer an expression system that responds to tetracycline concentration in the medium (UNIT 16.14).

Expression of proteins using viral vectors is presented in UNITS 16.15-16.22 and UNIT 16.24. UNITS 16.15-16.19 outline how to construct recombinant vaccinia viruses and characterize their products. UNIT 16.20 describes how to carry out expression using Semliki Forest Virus (SFV) vectors. Both vaccinia systems and SFV systems are able to produce large amount of protein following infection of cells with the appropriate virus. In both cases, the normal viral machinery is used to produce a single protein at very high levels, and thus these systems are useful for overexpression of protein in mammalian cells, but not useful if it is desirable to isolate the protein of interest in complex with a group of cellular proteins, as the expressed protein will be made at much higher levels than any endogenous protein.

An important goal that cannot be achieved by any of the above methods is the stable expression of protein in nondividing cells. In both therapeutic and basic science settings it can be important to express a protein in a post-mitotic cell, a need that has led to the development of vectors based on HIV and other related lentiviruses. Use of these systems is described in UNITS 16.21 & 16.22.

Another approach is to use vectors based on Adenovirus. Earlier generations of this strategy, based on replication-competent viruses, raised some concerns related to safety and duration of protein expression. Both issues have been addressed, at least in part, by a newer generation of helper-dependent (sometimes called “gutless”) adenoviral vectors (HDAds). The only Adenoviral *cis* elements present in these vectors are those required for vector genome replication and packaging into the capsid. Preparation of these defective viruses is dependent on a helper virus that provides the various necessary factors but lacks the packaging sequences and is not incorporated into viral particles. Advantages of these vectors include the large insert size of approximately 30 kb and the ability to transduce a variety of cell types, both in vitro and in vivo.

UNIT 16.23 addresses a somewhat different problem than is addressed in the previous units. While it is relatively easy to produce milligram amounts of any protein using viral expression systems, it is considerably more challenging to make stable cell lines that will continue to produce milligram quantities of a given protein over extended periods of time. Specially constructed CHO cell lines are used to solve this problem. These lines contain integrated constructions to direct the synthesis of the mammalian protein; the constructions carry either the dihydrofolate reductase or the glutamine synthetase gene, whose products confer drug resistance. Lines that carry increased numbers of the constructs are obtained by selecting cell lines that grow in increasing drug concentrations of methotrexate. Once selected, these lines are permanent reagents, which can be stored frozen and used to produce the protein whenever desired.

STRATEGIC CONSIDERATIONS WHEN USING MAMMALIAN EXPRESSION SYSTEMS

The importance of the initial planning stage cannot be overemphasized when discussing strategies for mammalian overexpression. With the exception of COS cell transfection, each of the strategies described in Section III requires a significant amount of set-up time. While a lot of this time is not hands-on time, as it consists of waiting for cells to grow, the total time between conception and completion of a specific experimental approach might easily be several months. Thus, it is more important than usual to devise an appropriate strategy, as strategic errors might not become apparent for several months.

The first consideration in choosing a strategy is to consider the goals of the expression work. If expression is desired to demonstrate that a clone has some functional activity or to perform an initial characterization of this activity, then transfection of COS cells might prove sufficient. If large amounts of an individual protein are required, then viral systems, such as vaccinia or SFV, or stable cell lines, such as CHO, will be the method of choice. If expression of an ectopic protein at the level of endogenous proteins is the goal, then creation of stable cell lines using retroviral (*UNIT 16.13*, see also *UNITS 9.9-9.14*) or lentiviral vectors should work.

The decision can come down to choosing between a viral system or the creation of stable cell lines. Creating stable cell lines offers the advantage that the engineered cells can be frozen and then thawed for use whenever the expressing line is needed. In addition, selection techniques can be used to isolate a clonal line that has desired expression levels of the protein. This clonal line should express similar amounts of protein in every cell, an important consideration if the goal is to purify a complex of proteins, or if the function of a given protein in an entire cell population is under study.

Creating stable cell lines requires both the choice of a suitable host cell and a suitable means of introducing the gene to be expressed. Obviously, the appropriate cell type (e.g., a lymphoid cell or neuronal cell) must be used if a specific cellular process is under study. Other considerations can be important later in the process. If eventually it will be important to grow large amounts of cells, then a cell line that can be adapted easily to growth in spinner culture is needed. If cells must be infected at a later stage in the experimental plan, then a cell line that is easily infected is important. Thus, planning several steps ahead when creating stable cell lines can help create a cell line that is maximally useful. Since it can take several months to isolate a clonal cell line in reasonable amounts, this planning takes on added importance.

Stable cell lines can be made using either transfection procedures or infection procedures. Transfection can be used to introduce several genes at once; however, the integrated genes might not prove to be completely stable. Retroviral vectors can be used, and can lead to more stable expression, but are considerably more cumbersome if more than one protein is to be introduced.

If overproduction of a single protein is the goal, frequently viral systems are the best choice. It usually takes less time to create an overexpressing virus than it does to make a high production stable cell line. Viral overproduction can also offer flexibility in that coinfection can be used to express two, three, or even more proteins in the same cell at the same time. This approach is not useful when it is important to purify the expressed protein with interacting proteins within the cell (e.g., as described in *UNIT 16.13*). The overexpressed protein will be expressed at such high levels that there will be insufficient endogenous protein to copurify in anywhere near stoichiometric amounts.

This necessarily brief discussion of mammalian expression systems points out just a few of the numerous specific issues that might be important in choosing an effective strategy. A thorough plan, with all future goals for the experimental system explicitly taken into account, will lead to the most efficient strategy for creating a maximally useful mammalian expression system.

Bob Kingston and Roger Brent

EXPRESSION OF PROTEINS IN *ESCHERICHIA COLI*

Overview of Protein Expression in *E. coli*

The study of *Escherichia coli* during the 1960s and 1970s made it the best understood organism in nature (Chapter 1). Today's recombinant DNA technology is a direct extension of the genetic and biochemical analyses carried out at that time. Even before the advent of molecular cloning, genetically altered *E. coli* strains were used to produce quantities of proteins of scientific interest. When cloning techniques became available, most cloning vectors utilized *E. coli* as their host organism. Thus, it is not surprising that the first attempts to express large quantities of proteins encoded by cloned genes were carried out in *E. coli*.

E. coli has two characteristics that make it ideally suited as an expression system for many kinds of proteins: it is easy to manipulate and it grows quickly in inexpensive media. These characteristics, coupled with more than 10 years' experience with expression of foreign genes, have established *E. coli* as the leading host organism for most scientific applications of protein expression.

Despite a growing literature describing successful protein expression from cloned genes, each new gene still presents its own unique expression problems. No one, and certainly no laboratory manual, can provide a set of methods that will guarantee successful production of every protein in a useful form. Nevertheless, the vast body of accumulated knowledge has led to a general approach that often helps to solve specific expression problems. This unit introduces general considerations and strategies, while subsequent units (16.2-16.7) describe procedures that can be applied to specific expression problems.

GENERAL STRATEGY FOR GENE EXPRESSION IN *E. COLI*

The basic approach used to express all foreign genes in *E. coli* begins with insertion of the gene into an expression vector, usually a plasmid. This vector generally contains several elements: (1) sequences encoding a selectable marker that assure maintenance of the vector in the cell; (2) a controllable transcriptional promoter (e.g., *lac*, *trp*, or *tac*) which, upon induction, can produce large amounts of mRNA from

the cloned gene; (3) translational control sequences, such as an appropriately positioned ribosome-binding site and initiator ATG; and (4) a polylinker to simplify the insertion of the gene in the correct orientation within the vector. Once constructed, the expression vector containing the gene to be expressed is introduced into an appropriate *E. coli* strain by transformation (UNIT 1.8).

SPECIFIC EXPRESSION SCENARIOS

Although this general approach—insertion of the gene of interest into an expression vector followed by transformation in *E. coli*—is common to all expression systems, specific procedures differ greatly. When choosing a procedure, it is helpful to consider the final application of the expressed protein, as this often dictates which expression strategy to use (UNIT 16.4A).

Antigen Production

If the goal is to use the expressed protein as an antigen to make antibodies, several approaches are available to make protein reliably and to allow for rapid purification of the antigen. The two best approaches are synthesis of fusion proteins with specific "tag" sequences that can be retrieved by affinity chromatography (UNITS 16.5, 16.6, 16.7 & 16.8; see also UNIT 10.11B) and synthesis of the native protein, or a fragment of it, under conditions that cause it to precipitate into insoluble inclusion bodies (UNITS 16.4A & 16.6). These inclusion bodies can be purified sufficiently by differential centrifugation so preparative denaturing polyacrylamide gel electrophoresis (UNIT 10.2) will yield an isolated band that can be cut out and crushed, or electroeluted, to provide antigenic material for injection into an animal.

Biochemical or Cell Biology Studies

If the goal is to use the expressed protein as a reagent in a series of biochemical or cell biology experiments, other considerations are relevant. In this case, the authenticity of the protein's function (e.g., high-specific-activity enzyme, binding protein, or growth factor) is

Contributed by Paul F. Schendel

Current Protocols in Molecular Biology (1998) 16.1.1-16.1.3
Copyright © 1998 by John Wiley & Sons, Inc.

very important, while the ease of preparing the protein matters less. For this application, it is possible to express the protein as a fusion protein containing a specific protease-sensitive cleavage site so the N-terminal peptide tail can be removed easily, leaving only the native amino acid sequence (UNITS 16.4, 16.6, 16.7 & 16.8). Alternatively, direct expression vectors of the type described in UNITS 16.2 & 16.3 may be used to produce the authentic primary sequence. When expressed, the protein may be soluble and active, as is the case with many intracellular enzymes. If it is insoluble, as is the case for many secreted growth factors when they are made cytoplasmically in *E. coli*, it may be necessary to isolate inclusion bodies, solubilize the protein using denaturing agents, and refold the protein. Refolding is usually not too difficult when the protein is of moderate size (Marston and Hartley, 1990). Whether the protein is expressed in a soluble form or whether it requires refolding, its integrity can usually be checked by specific enzyme assays or by bioassays.

Structural Studies

If the goal is to do structural studies of the expressed protein, the greatest constraints are imposed on the expression system. Because it is nearly impossible to show that a protein of unknown structure has been precisely refolded after denaturing, the protein must generally be made in a soluble form so its purification does not require a denaturation/renaturation step. Usually, the soluble form of the protein—either intracellular or secreted—must be made in strains and by induction protocols that minimize proteolytic degradation.

Soluble expression of most eukaryotic proteins is best achieved with systems that allow induction of synthesis without changing the temperature; for example, by inducing transcription from the *trp* (Edman et al., 1981; de Boer et al., 1983) or *tac* (de Boer et al., 1983) promoters. Maximum accumulation of soluble product is best achieved by testing expression in several strains and at several temperatures, and picking the combination that works best. This is an active area of research at present (Schein, 1989); the rules are not yet understood, so little more than trial and error can be recommended.

TROUBLESHOOTING GENE EXPRESSION

Once an expression strategy has been chosen and the gene is introduced into an appropriate

expression vector, several strains of *E. coli* should be transformed with the vector and protein production should be monitored. Ideally, the protein of interest will be produced in an active form and in sufficient amounts to allow its isolation. Often, however, the protein will be made either in very small amounts or in an insoluble form, or both. If this happens, there are various approaches that may correct the problem.

If not enough protein is produced:

1. Reconstruct the 5'-end of the gene, maximizing its A+T content while preserving the protein sequence it encodes. This may reduce secondary structure within the mRNA (DeLamar et al., 1985), or it may alter an as yet undefined parameter of the reaction. Regardless of the underlying cause, this procedure usually increases translation efficiency.

2. Determine if a transcriptional terminator is present. If the vector does not have a transcriptional terminator downstream from the site at which the gene is inserted, put one in. This often aids expression, probably by increasing mRNA stability and by decreasing nucleotide drain on the cell.

3. Examine the sequence of the cloned gene for codons used infrequently in *E. coli* genes. These so-called rare codons are usually not a rate-limiting problem, but if four or more happen to occur contiguously, they can reduce expression significantly (Robinson et al., 1984), perhaps by causing ribosomes to pause. Ribosomal pausing can uncouple transcription from translation, leading to premature termination of the message. Even if transcription proceeds normally, the mRNA 3' to the stalled ribosomes can be exposed to degradation by host ribonucleases, reducing its stability. Thus, if stretches of rare codons are found, they should be altered to codons more favorable to high expression in *E. coli*.

If enough protein is produced, but it is insoluble when the application requires it to be active and soluble:

1. Vary the growth temperature. As mentioned above, many proteins are more soluble at lower than at higher temperatures (Schein and Noteborn, 1988). On the other hand, some enzymes have a higher specific activity when made at temperatures >37°C (J. McCoy and P. Schendel, unpub. observ.). *E. coli* can synthesize proteins at temperatures ranging from 10° to 43°C, so trying expression at different temperatures is often worthwhile.

2. Change fermentation conditions. Many proteins contain metals as structural and catalytic cofactors. If the protein is being made faster than metals can be transported into the cell, the apoprotein without its metal cofactor will accumulate. This apoprotein will not fold correctly and will likely be insoluble. At the very least, the average specific activity of the expressed protein will be lower than expected. Different media and metal supplements can be tested and the best combination used. Clearly, if there is information about the metal content of the protein, these supplements can be designed more rationally. If no information is available, a more random approach must be tried.

3. Alter the rate of expression by using low-copy-number plasmids. This can be done by using the pACYC family (Chang and Cohen, 1978) or using single-copy chromosomal inserts of the cloned gene into a suitable target gene (Hamilton et al., 1989). Such reductions in gene dosage often reduce the final yield of protein, but the slower kinetics of synthesis they afford can sometimes result in production of soluble proteins.

To restate the obvious, protein expression is an inexact science at present. However, most proteins can be made in *E. coli* in a form that is useful for a variety of functions. The procedures employed are relatively quick and uncomplicated, and the rewards for success are great.

Literature Cited

Chang, A.C.Y. and Cohen, S.N. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.* 134:1141-1156.

de Boer, H.A., Comstock, L.J., and Vasser, M. 1983. The *tac* promoter: A functional hybrid derived from the *trp* and *lac* promoters. *Proc. Nat. Acad. Sci. U.S.A.* 80:21-25.

DeLamararter, J.F., Mermod, J.J., Liang, C.M., Eliason, J.F., and Thatcher, D.R. 1985. Recombinant murine GM-CSF from *E. coli* has biological activity and is neutralized by a specific antiserum. *EMBO J.* 4:2575-2581.

Edman, J.C., Hallewell, R.A., Valenzuela, P., Goodman, H.M., and Rutter, W.J. 1981. Synthesis of Hepatitis B surface and core antigens in *E. coli*. *Nature* 291:503-506.

Hamilton, C.M. Aldea, M., Washburn, B.K., Babitzke, P., and Kushner, S.R. 1989. New methods for generating deletions and gene replacements in *Escherichia coli*. *J. Bacteriol.* 171:4617-4622.

Marston, F.A.O. and Hartley, D.L. 1990. Solubilization of protein aggregates. *Methods Enzymol.* 182:264-276.

Robinson, M., Lilley, R., Little, S., Emtage, J.S., Yarranton, G., Stephens, P., Millican, A., Eaton, M., and Humphreys, G. 1984. Codon usage can affect efficiency of translation of genes in *Escherichia coli*. *Nucl. Acids Res.* 12:6663-6671.

Schein, C.H. 1989. Production of soluble recombinant proteins in bacteria. *BioTechnology* 7:1141-1148.

Schein, C.H. and Noteborn, M.H.M. 1988. Formation of soluble recombinant proteins in *Escherichia coli* is favored by lower growth temperatures. *BioTechnology* 6:291-294.

Contributed by Paul F. Schendel
Genetics Institute
Cambridge, Massachusetts

Expression Using the T7 RNA Polymerase/Promoter System

This unit describes the expression of genes by placing them under the control of the bacteriophage T7 RNA polymerase. This approach has a number of advantages compared to approaches that rely on *E. coli* RNA polymerase. First, T7 RNA polymerase is a very active enzyme: it synthesizes RNA at a rate several times that of *E. coli* RNA polymerase and it terminates transcription less frequently; in fact, its transcription can circumnavigate a plasmid, resulting in RNA several times the plasmid length in size. Second, T7 RNA polymerase is highly selective for initiation at its own promoter sequences and it does not initiate transcription from any sequences on *E. coli* DNA. Finally, T7 RNA polymerase is resistant to antibiotics such as rifampicin that inhibit *E. coli* RNA polymerase, and consequently, the addition of rifampicin to cells that are producing T7 RNA polymerase results in the exclusive expression of genes under the control of a T7 RNA polymerase promoter (hereafter referred to as p_{T7}).

To use the two-plasmid p_{T7} system, it is necessary to clone the gene to be expressed into a plasmid containing a promoter recognized by the T7 RNA polymerase. The gene is then expressed by induction of T7 RNA polymerase. The gene for T7 RNA polymerase is present on a second DNA construction. This second construction can either permanently reside within the *E. coli* cell (basic protocol), or can be introduced into the cell at the time of induction by infection with a specialized phage, such as an M13 vector (mGP1-2; Tabor and Richardson, 1987) or a λ vector (CE6; Studier et al., 1990) containing the T7 RNA polymerase gene (second alternate protocol).

In the basic protocol, two plasmids are maintained within the same *E. coli* cell. One (the expression vector) contains p_{T7} upstream of the gene to be expressed. The second contains the T7 RNA polymerase gene under the control of a heat-inducible *E. coli* promoter. Upon heat induction, the T7 RNA polymerase is produced and initiates transcription on the expression vector, resulting in turn in the expression of the gene(s) under the control of p_{T7} . If desired, the gene products can be uniquely labeled by carrying out the procedure in minimal medium, adding rifampicin to inhibit the *E. coli* RNA polymerase, and then labeling the proteins with [³⁵S]methionine (first alternate protocol).

EXPRESSION USING THE TWO-PLASMID SYSTEM

The gene to be induced is subcloned into an expression vector containing p_{T7} . Two series of vectors have been developed for this purpose—the pT7 series (Fig. 16.2.1) and the pET series (Studier et al., 1990); see commentary for discussion of choice of vector. The plasmid containing the introduced gene is then used to transform an *E. coli* strain already containing the plasmid pGP1-2 (Fig. 16.2.2). pGP1-2 contains the gene for T7 RNA polymerase under the control of the λp_L promoter that is repressed by a temperature-sensitive repressor (*cI857*). pGP1-2 contains a p15A origin of replication that is compatible with the ColE1 origin of replication on the expression vector. The two plasmids are maintained in the same cell by selection with kanamycin (pGP1-2) and ampicillin (the expression vector).

Cells containing the two plasmids are grown for several hours at 30°C and then the gene for T7 RNA polymerase is induced by raising the temperature to 42°C. The production of T7 RNA polymerase in turn induces expression of the genes under the control of p_{T7} . (Rifampicin can be subsequently added to inhibit transcription by *E. coli* RNA polymerase, although this is usually not necessary since T7 RNA polymerase becomes responsible for most of the transcription even in the absence of rifampicin.) After expression

BASIC PROTOCOL

Protein Expression

16.2.1

Contributed by Stanley Tabor

Current Protocols in Molecular Biology (1990) 16.2.1-16.2.11

Copyright © 2000 by John Wiley & Sons, Inc.

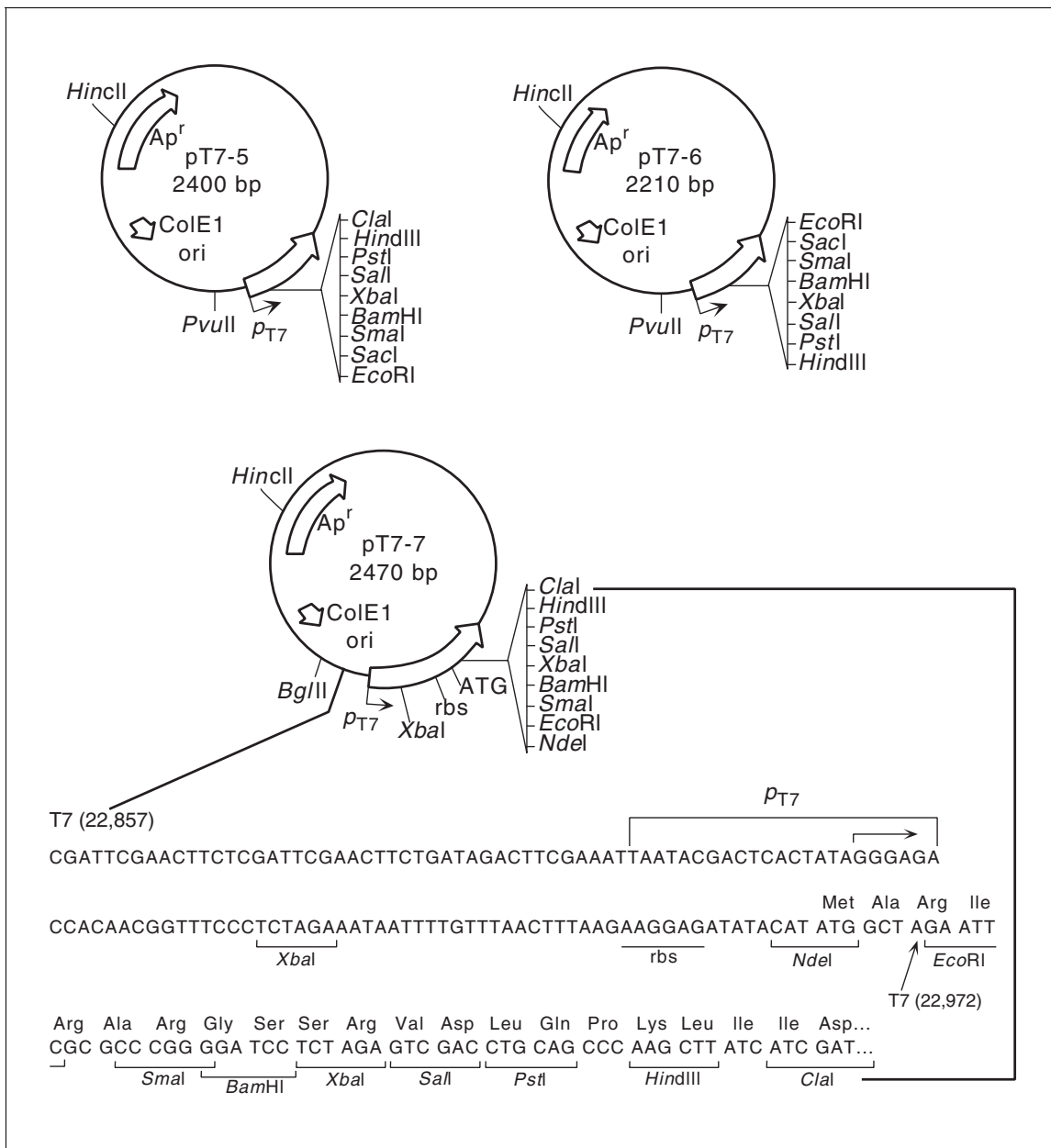


Figure 16.2.1 pT7-5, pT7-6, and pT7-7. pT7-5, pT7-6, and pT7-7 are cloning vectors that contain a T7 promoter and are used to express genes using T7 RNA polymerase. All three vectors contain a T7 RNA polymerase promoter, the gene encoding resistance to the antibiotic ampicillin and the ColE1 origin of replication. pT7-7 has a strong ribosome-binding site (rbs) and start codon (ATG) upstream of the polylinker sequence; the sequence of this region is shown below the map of pT7-7. pT7-5 and pT7-6 lack any ribosome-binding site upstream of the polylinker sequence and consequently are only useful when expressing genes that already contain the proper control sequences. pT7-5, pT7-6, and pT7-7 were constructed by S. Tabor and are derivatives of pT7-1 described in Tabor and Richardson (1985).

of the genes at 37°C, the cells are harvested and the induced proteins are analyzed. An alternative approach is to induce T7 RNA polymerase with IPTG rather than by heat induction. In this method, the expression plasmid containing p_{T7} can be placed into *E. coli* BL21 (DE3), which contains the gene for T7 RNA polymerase on the *E. coli* chromosome under the control of the *lac* promoter (Studier and Moffatt, 1986; Studier et al., 1990).

Materials

pT7-5, pT7-6, or pT7-7 vectors (available from author)
E. coli JM105, DH1, or equivalent (Table 1.4.5)
 LB plates and medium containing 60 µg/ml ampicillin (UNIT 1.1)
E. coli K38 or equivalent (Table 1.4.5)
 pGP1-2 (available from author)
 LB plates and medium containing 60 µg/ml kanamycin (UNIT 1.1)
 LB plates and medium containing 60 µg/ml ampicillin plus 60 µg/ml kanamycin (UNIT 1.1)

Cracking buffer
 Sorvall SS-34 or GS-3 rotor or equivalent

Additional reagents and equipment for subcloning DNA fragments (UNITS 1.4 & 3.16), transformation of competent *E. coli* cells (UNIT 1.8), minipreps of plasmid DNA (UNIT 1.6), restriction mapping (UNITS 3.1-3.3), and SDS-PAGE (UNIT 10.2).

1. Subclone the fragment containing the gene to be expressed into pT7-5, pT7-6, or pT7-7. Transform a standard *E. coli* strain (e.g., JM105 or DH1); this strain should *not* carry a plasmid that directs synthesis of T7 RNA polymerase (i.e., pGP1-2). Plate the transformants on LB/ampicillin plates and grow overnight at 37°C.

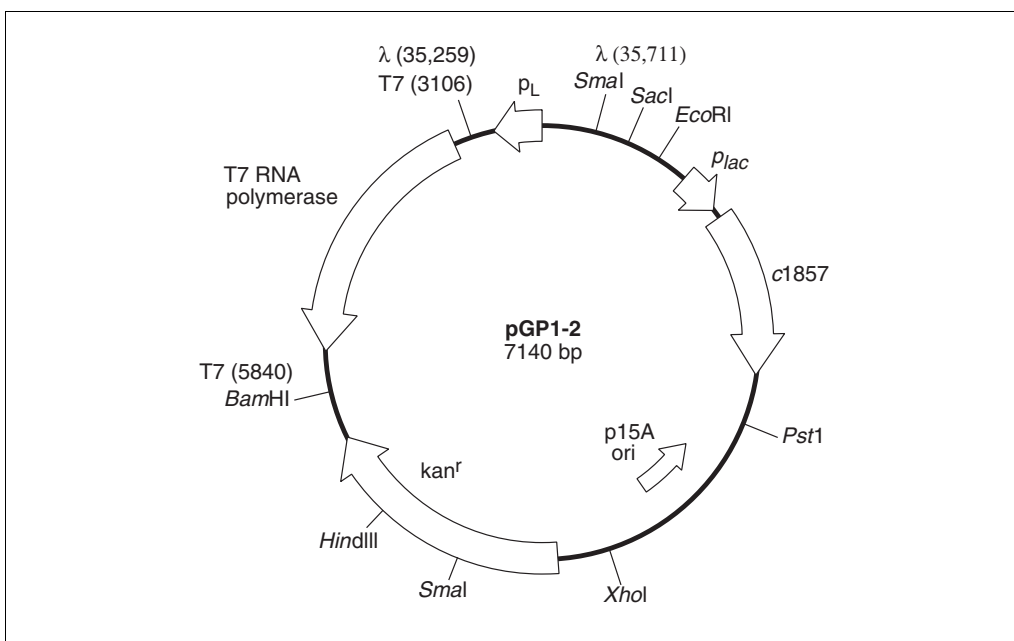


Figure 16.2.2 pGP1-2. pGP1-2 enables T7 RNA polymerase to be produced by heat induction in any *E. coli* host. pGP1-2 contains the gene for T7 RNA polymerase under the control of the λ p_L promoter. It also contains the gene for the λ repressor (*c1857*) that is expressed under the control of *E. coli* p_{lac} promoter. This repressor inhibits transcription from the λ p_L promoter at low temperature (30°C); however, at high temperature (42°C) it is inactivated, resulting in induction of the p_L promoter, that in turn results in induction of the T7 RNA polymerase. pGP1-2 also contains the gene encoding resistance to the antibiotic kanamycin, and the p15A origin of replication. pGP1-2 is described in Tabor and Richardson (1985).

It is important to first transform the plasmid into a strain that contains no T7 RNA polymerase, in case small amounts of the gene product are toxic to the cell (see critical parameters for discussion on toxic genes).

2. Grow individual transformants in LB/ampicillin medium at 37°C and obtain plasmid DNA by a miniprep procedure. Confirm that the gene has been correctly inserted by restriction mapping.
3. Transform *E. coli* K38 with pGP1-2, plate on LB/kanamycin plates, and grow overnight at 30°C. Grow an individual *E. coli* K38/pGP1-2 transformant in LB/kanamycin medium at 30°C.

*Colonies take ~24 hr to appear on plates at 30°C. *E. coli* K38/pGP1-2 can be stored in the absence of the plasmid containing p_{T7} as a glycerol stock at -80°C (see commentary).*

4. Transform the vector containing the gene to be expressed under the control of p_{T7} into *E. coli* K38/pGP1-2 grown in LB/kanamycin medium. Plate the transformants (containing both plasmids) on LB/ampicillin/kanamycin plates and grow overnight at 30°C.

Cells may be heat-shocked during transformation; the T7 RNA polymerase gene, under the control of a heat-inducible promoter, is not induced by this brief heating step.

*As a control, transform *E. coli* K38/pGP1-2 with the parent p_{T7} vector (without an insert). If the transformation efficiency of the vector containing the insert is significantly lower (by more than a factor of 50) than that of the parent vector, the gene product may be toxic to *E. coli* cells. This toxicity arises from background expression of the gene product by basal levels of T7 RNA polymerase. In this situation, the transformants that do arise invariably contain deletions or other mutations in one of the two plasmids, and the desired gene product will not be produced. If the expression of the inserted gene is toxic, it is necessary to use an alternative strategy for the repression and induction of the T7 RNA polymerase gene (see discussion on toxic genes in critical parameters).*

5. Pick a single *E. coli* colony that contains the two plasmids with a sterile toothpick or pipet. Inoculate it into 5 ml LB/ampicillin/kanamycin medium and grow overnight at 30°C.
6. Dilute the overnight culture of cells 1:40 into fresh LB/ampicillin/kanamycin medium and grow several hours at 30°C to an $OD_{590} \cong 0.4$.

The size of the culture will depend on the amount of cells needed. For an analytical preparation, use ~1 ml of cells.

7. Induce the gene for T7 RNA polymerase by raising the temperature to 42°C for 30 min, which in turn induces the genes under the control of p_{T7} .

To obtain consistent results, raise the temperature relatively quickly. If small cultures (~1 ml) are being induced, place the cultures into a 42°C water bath. For larger cultures (~500 ml), place the flask under hot tap water until the temperature of the media reaches 42°C (measured by inserting a thermometer wiped with ethanol into the flask). Once the cells reach 42°C, continue incubating at 42°C for 30 min.

*The *E. coli* RNA polymerase can be inhibited by adding rifampicin to a final concentration of 200 µg/ml; when used, it should be added after T7 RNA polymerase has been induced at 42°C for 30 min. Although rifampicin reduces the background of host proteins being expressed, in general it does not significantly increase the final accumulation of gene products, and in some cases it decreases the final yield. Thus, as a general rule, rifampicin is only added to cells when the plasmid-encoded proteins are being uniquely labeled with [³⁵S]methionine (see first alternate protocol).*

8. Reduce temperature to 37°C and grow the cells an additional 90 min with shaking.
9. Harvest the cells by centrifuging and discarding the supernatant. For 1-ml cultures, microcentrifuge 20 sec at 10,000 rpm (14,000 × *g*), room temperature. For 2-ml to 100-ml cultures, centrifuge 5 min in a Sorvall SS-34 rotor at 5000 rpm (3000 × *g*), 4°C. For >100-ml cultures, centrifuge 10 min in a Sorvall GS-3 rotor at 5000 rpm (4000 × *g*), 4°C.
10. To analyze the induced proteins by SDS-PAGE, resuspend the equivalent of 1.0 ml of cells in 0.1 ml cracking buffer. Heat at 100°C for 5 min immediately prior to loading a 20- μ l aliquot of each sample onto an SDS-polyacrylamide gel (UNIT 10.2). To analyze the cells for an induced enzymatic activity, prepare an appropriate cell extract from ~10 ml of cells.

One example of the preparation of an extract for the purification of T7 RNA polymerase is described in Tabor and Richardson (1985).

SELECTIVE LABELING OF PLASMID-ENCODED PROTEINS

**ALTERNATE
PROTOCOL**

Plasmid-encoded proteins under the control of a p_{T7} (see basic protocol) can be exclusively labeled by inducing the T7 RNA polymerase in cells growing in minimal medium, inhibiting the host *E. coli* RNA polymerase with rifampicin, and labeling the newly synthesized proteins with [³⁵S]methionine. This procedure provides an attractive alternative to maxicells or minicells for labeling of plasmid-encoded proteins (Dougan and Sherratt, 1977; Sancar et al., 1981).

Additional Materials

- M9 medium (UNIT 1.1) without and with 5% (v/v) of 18 amino acid mixture
- 20 mg/ml rifampicin in methanol (e.g., Sigma #R-3501; store in dark at 4°C for 2 weeks; Table 1.4.1)
- 10 mCi/ml [³⁵S]methionine (>800 Ci/mmol) diluted 1:10 in M9 medium
- Fluorographic enhancing agent (e.g., Enlightning from Du Pont NEN or Amplify from Amersham)

1. Repeat steps 3 to 6 of the basic protocol (using the T7-promoter expression plasmid obtained from steps 1 and 2 of the basic protocol).

*An alternative to the use of LB/ampicillin/kanamycin medium for growing cells is M9 medium containing 25 μ g/ml ampicillin and 25 μ g/ml kanamycin, and any required nutrients. The addition of one part in twenty of the 18 amino acid mixture (0.1% stock, 0.005% final concentration) stimulates the growth of cells in M9 medium without interfering with the subsequent labeling of the proteins with [³⁵S]methionine. Note that to grow in this medium, the *E. coli* strain must be Cys⁺ and Met⁺.*

2. When OD₅₉₀ \cong 0.4, remove 1 ml of cells, microcentrifuge 10 sec, and discard supernatant.
3. Wash cell pellet with 1 ml M9 medium, microcentrifuge 10 sec at room temperature, and discard supernatant.

Washing the cells after growth in LB medium is very important in order to remove the unlabeled methionine present in LB medium that otherwise dilutes the [³⁵S]methionine during labeling.

4. Resuspend cell pellet in 1 ml M9 medium containing 18 amino acid mixture. Grow cells 60 min at 30°C with shaking.

A time of 30 to 180 min is adequate for adapting cells to M9 medium. Although the OD₅₉₀ may not increase significantly during this step, induction of T7 RNA polymerase and

**Protein
Expression**

16.2.5

efficient labeling of the plasmid-encoded proteins will occur even in the absence of apparent cell growth.

5. Induce the gene for T7 RNA polymerase by placing the cells in a 42°C water bath for 20 min.
6. Add 20 mg/ml rifampicin to 200 µg/ml final. Keep cells at 42°C for an additional 10 min after adding rifampicin.

It is important to incubate the cells at 42°C for an additional 10 min after adding rifampicin, since rifampicin is more effective at inhibiting expression of host proteins at 42°C, possibly because the cells are more permeable to it at this temperature. The temperature of the cells is subsequently reduced for the labeling since in general the labeling is less efficient at 42°C than at 30° or 37°C.

7. Shift cells to a 30°C water bath for an additional 20 min. Remove 0.5 ml of cells for labeling with [³⁵S]methionine.

The other 0.5 ml can be used to label the cells at a later time point (e.g., after an additional 30 min) in order to follow the duration of protein synthesis.

8. Label newly synthesized proteins by adding 10 µl (10 µCi) diluted [³⁵S]methionine to 0.5 ml of cells and incubating for 5 min at 30°C.
9. Microcentrifuge cells 10 sec and discard supernatant. (**CAUTION:** the supernatant is radioactive; discard properly.) Resuspend cell pellet in 100 µl cracking buffer.
10. Heat samples to 100°C for 5 min. Load a 20-µl aliquot onto an SDS-polyacrylamide gel and electrophorese (*UNIT 10.2*).
11. Treat the gel with a fluorographic-enhancing agent by soaking it in the fluor for 30 min. Dry the gel under vacuum 2 hr at 65°C and autoradiograph (*APPENDIX 3*).

A 1-hr exposure should be adequate to visualize most proteins induced with this system.

*To determine whether the plasmid-encoded proteins are susceptible to proteases in the *E. coli* cell, prepare and induce the cells as described above; however, reduce the duration of the labeling step to 1 min (step 8), and follow this with a chase of nonradioactive methionine at 0.5% final concentration. Remove an aliquot for analysis both immediately prior to the chase, and after a chase reaction of 5, 15, and 60 min. After removing each aliquot, immediately pellet the cells by centrifugation, resuspend in cracking buffer, and heat the aliquot to 100°C for 5 min to inactivate the proteases. Analyze as in step 10.*

ALTERNATE PROTOCOL

EXPRESSION BY INFECTION WITH M13 PHAGE mGP1-2

Whenever the gene for T7 RNA polymerase is present in *E. coli* cells, low levels of T7 RNA polymerase are constitutively produced. This can be a problem when the gene products under the control of p_{T7} are toxic. One strategy to avoid this is to keep the gene for T7 RNA polymerase out of the cell until the time of induction. In the protocol presented here, T7 RNA polymerase is introduced into the cell by infection with the M13 phage mGP1-2. This phage contains the gene for T7 RNA polymerase under the control of the *lac* promoter (Fig. 16.2.3). Host cells for this phage must carry the F factor so that they are susceptible to M13 infection (e.g., JM101 or K38). The cells are transformed with the single plasmid that contains the gene to be expressed under the control of p_{T7} . The cells are grown at 37°C, and induction occurs by infection with a high multiplicity of mGP1-2 in the presence of IPTG. A λ vector, CE6, that contains the gene for T7 RNA polymerase has also been used to express toxic genes (Studier and Moffatt, 1986; Studier et al., 1990).

Additional Materials

M13 phage mGP1-2 (available from author)

PEG solution (UNIT 1.7)

100 mM IPTG (Table 1.4.2)

Additional reagents and equipment for preparing M13 phage (UNIT 1.15) and titering phage (UNIT 1.11)

1. Prepare a stock of M13 phage mGP1-2 and concentrate the phage by precipitation with PEG solution. (DO NOT proceed to add TE buffer or phenol.) Resuspend phage in M9 medium and titer.

If the cell proteins are to be labeled, it is important that the phage used to infect the cells are free of unlabeled methionine. In this case, precipitate the phage with PEG twice, each time resuspending the pellet in M9 medium. For long-term storage of the M13 phage mGP1-2, it is best to purify the phage through a CsCl gradient (Nakai and Richardson, 1986).

2. Transform *E. coli* cells susceptible to M13 infection (e.g., JM101 or K38) with the T7-promoter expression plasmid obtained from steps 1 and 2 of the basic protocol. Plate the transformants on LB/ampicillin plates and grow overnight at 37°C.
3. Pick a single colony and grow in LB/ampicillin medium overnight at 37°C.
4. Dilute the overnight culture of cells 1:100 in LB/ampicillin medium and grow several hours at 37°C with gentle shaking to $OD_{590} \cong 0.5$.

*It is very important that only gentle shaking is used when growing cells for M13 infection. Vigorous agitation results in shearing of the pili on the surface of the *E. coli* cells, resulting in inefficient infection.*

5. Infect cells with M13 phage mGP1-2 (from step 1) at a ratio of ~10 phage for each *E. coli* cell. Add 100 mM IPTG to 1 mM final (a 1:100 dilution) to induce production

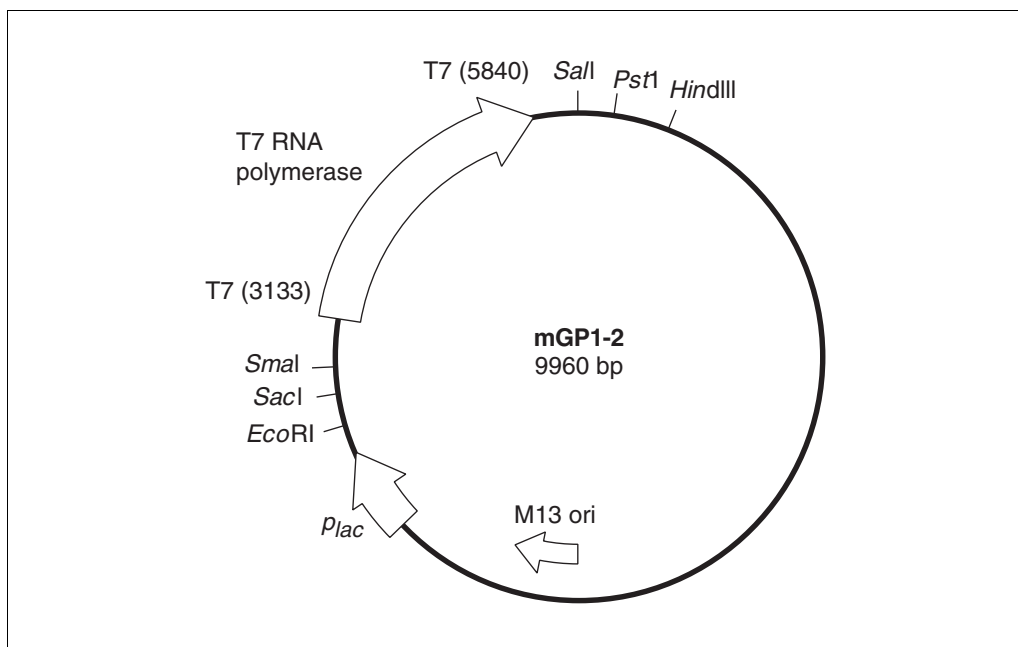


Figure 16.2.3 mGP1-2. M13 phage mGP1-2 contains the gene for T7 RNA polymerase under the control of the *E. coli* p_{lac} promoter. It is especially useful for the production of gene products that are toxic to the *E. coli* cell. When *E. coli* cells are infected with this phage, and IPTG is added to induce the p_{lac} promoter, T7 RNA polymerase is produced. As a result, any genes within the cell under the control of p_{T7} will be induced. mGP1-2 is described in Tabor and Richardson (1987).

of T7 RNA polymerase. Incubate the cells 2 hr at 37°C.

*At $OD_{590} \cong 0.5$, the density of *E. coli* cells will be $\sim 2 \times 10^8$ cells/ml. Thus, it is necessary to add M13 mGP1-2 phage at a final concentration of 2×10^9 phage/ml to obtain a multiplicity of infection of 10. Small cultures (~ 50 ml) can be incubated in a water bath without shaking. Larger cultures should be incubated at 37°C with gentle shaking.*

6. Harvest cells and analyze induced proteins as in steps 9 and 10 of the basic protocol.

REAGENTS AND SOLUTIONS

18 amino acid mixture

Prepare a solution containing 0.1% (v/v) of each amino acid except cysteine (minus cysteine) and methionine (minus methionine). Filter sterilize through a 0.2- μ m filter. Store at -20°C for several years.

Cracking buffer

60 mM Tris-Cl, pH 6.8
1% 2-mercaptoethanol
1% sodium dodecyl sulfate (SDS)
10% glycerol
0.01% bromphenol blue

COMMENTARY

Background Information

Bacteriophage T7 and T7-related phage (e.g., SP6, T3) encode their own RNA polymerase (see *UNIT 3.8*). Compared to other known RNA polymerases, this RNA polymerase is both relatively simple and highly efficient. T7 RNA polymerase is a single polypeptide of 96,000 kDa. It initiates transcription specifically at a 23-nucleotide promoter sequence, a sequence not present on the *E. coli* genome. Transcription is very processive, producing transcripts that are many thousands of nucleotides in length. Transcription is relatively rapid—five times the rate of *E. coli* RNA polymerase. All of these properties make T7 RNA polymerase and its promoter an attractive system for controlling the expression of foreign genes in *E. coli* and in other organisms. Expression systems in *E. coli* based on the controlled induction of T7 RNA polymerase have been developed by Tabor and Richardson (1985) and Studier and his colleagues (Studier and Moffatt, 1986; Rosenberg et al., 1987; Studier et al., 1990). The vectors described here are those developed by Tabor and Richardson. T7 RNA polymerase/promoter expression systems have also been successfully applied in yeast (Chen et al., 1987) and mammalian cells (Dunn et al., 1988; Fuerst et al., 1986).

Critical Parameters

Choice of vector

Questions that determine what vector to use to express a gene using T7 RNA polymerase include: Is there a ribosome-binding site upstream of the gene? What are the restriction sites available on each end of the gene? Is the gene product toxic to the *E. coli* cell? Examples of three standard vectors (pT7-5, pT7-6, and pT7-7) are shown in Figure 16.2.1. These vectors are derivatives of pBR322. The β -lactamase gene encoding amp^r is in the opposite orientation of p_{T7} ; consequently the only plasmid-encoded genes expressed by T7 RNA polymerase are those cloned into the polylinker region. pT7-5 and pT7-6 contain the polylinker region located immediately downstream of pT7 in opposite orientations. There is no ribosome-binding sequence in these two plasmids; they should thus be used either for the production of transcripts without expectation of good translation of the protein, or for the expression of genes that already have strong ribosome-binding sequences. pT7-7 differs from pT7-5 and pT7-6 in that it contains a strong ribosome-binding sequence between p_{T7} and the polylinker region; it is recommended for the expression of genes that lack a strong ribosome binding sequence or for the production of fusion proteins.

An extensive series of additional vectors containing p_{T7} , the pET series, have been described by Studier et al. (1990). These vectors are particularly useful for applications that require a greater selection of restriction endonuclease sites to insert the gene into, or that involve the expression of a gene that is toxic to the cell (see below). Some of these vectors contain other transcriptional regulatory elements (i.e., terminators, operators, RNase III cleavage sites) that could be of use for specific applications.

A large number of commercially available vectors contain a T7 RNA promoter (e.g., pIBI vectors, available from IBI; pSP6/T7-19, available from GIBCO/BRL; pBluescript II vectors, available from Stratagene; and pTZ18R and pTZ19R, available from U.S. Biochemical). These are intended to be used for producing specific transcripts in vitro using T7 RNA polymerase. In principle, they should be useful for the expression of genes using T7 RNA polymerase in vivo as well. In practice, however, the use of some of these vectors can result in some unexpected problems. (1) Most commercial vectors have extremely high copy numbers within the cell; this can accentuate the problems encountered with toxic genes. (2) In most vectors, the β -lactamase gene is oriented in the same direction as p_{T7} , complicating the analysis of radiolabeled proteins. (3) Some commercial vectors have p_{T7} oriented in a potentially deleterious direction. Derivatives of pBR322 that contain p_{T7} oriented clockwise with respect to the standard map are inviable in some *E. coli* strains that contain the gene for T7 RNA polymerase. This is due to the fact that high levels of transcription through the origin region of these plasmids in this orientation interferes with the replication of the plasmids. (4) Most commercial vectors have a *lac* operator sequence within them. This can titrate out the *lac* repressor (UNIT 1.4) and cause problems when the p_{lac} is used to control the T7 RNA polymerase gene.

Toxic genes

In some cases the gene to be expressed is toxic to the cells, even when it is not induced. This is due to a low level of constitutive expression present even under uninduced conditions. Although most genes are not toxic when expressed using the two-plasmid p_{T7} system, it is important to recognize the symptoms of toxicity to avoid selecting for mutations and to allow alternate systems for induction to be tried. The degree of toxicity varies greatly with each gene.

The symptoms encountered with toxic genes are discussed below, in order of increasing toxicity.

Some genes are mildly toxic to the cells when expressed using the two-plasmid p_{T7} system. In such cases, the cells can be stably transformed with the two plasmids and the gene product is produced at a high level. However, after the cells are several days old, they no longer induce the expected gene product even though they remain resistant to ampicillin and kanamycin. To avoid this problem, it is recommended that the *E. coli* K38/pGP1-2 be stored in the absence of the plasmid containing p_{T7} as a glycerol stock at -80°C (UNIT 1.3). The plasmid containing p_{T7} and the gene to be expressed should be stored as DNA at -20°C or -70°C (UNIT 1.6). To prepare the strain for induction, streak K38/pGP1-2 on an LB/kanamycin plate at 30°C , grow up a single colony, transform with the plasmid containing p_{T7} and the gene to be expressed, and plate the transformants on LB/ampicillin/kanamycin plates at 30°C . A single colony should then be grown at 30°C and induced as described above. This procedure is not necessary for genes that are not toxic. Strains that do not induce toxic genes can be stored in glycerol at -80°C for many months (UNIT 1.3).

A more toxic class of genes consists of those that can be successfully cloned into a plasmid under the control of p_{T7} , but that render the resulting plasmid unable to stably transform a cell that contains the gene for T7 RNA polymerase. Genes that are toxic to the cells only in the presence of pGP1-2 (which expresses the T7 RNA polymerase) are relatively common, occurring on the average ~5% of the time (S. Tabor, unpublished observation). Note that such plasmids will give transformants in *E. coli* cells containing pGP1-2, but that the frequency of transformation will be greatly reduced (>50-fold) compared to the frequency of transformation by the parent vector alone. The cells that do grow in the presence of ampicillin and kanamycin will invariably contain deletions or other mutations in one of the two plasmids, and the desired gene product will not be produced.

When genes are toxic at this level, it is necessary to use an alternative strategy that reduces the expression of the gene under uninduced conditions. One strategy is to remove the gene for T7 RNA polymerase from the cell until induction is desired, and then introduce it by a phage infection. Such an alternate protocol is described using an M13 phage harboring the gene for T7 polymerase, mGP1-2. A lambda

vector, CE6, that contains the gene for T7 RNA polymerase has also been used for this purpose (Studier and Moffatt, 1986; Studier et al., 1990). Another strategy is to retain the gene for T7 RNA polymerase in the cell but reduce the level of transcription by T7 RNA polymerase under uninduced conditions. For example, a system has been developed that expresses an inhibitor of T7 RNA polymerase—the T7 lysozyme—to reduce the activity of T7 RNA polymerase until it is induced (Studier et al., 1990). Another recent modification is the placement of p_{T7} under the control of the *lac* repressor, reducing the activity of T7 RNA polymerase until IPTG has been added (Studier et al., 1990).

Finally, some genes are difficult to clone in multicopy plasmids even in the absence of a known *E. coli* promoter. The difficulty in cloning these genes arises from the fact that their products are extremely toxic and that the residual low level of transcription by *E. coli* RNA polymerase in most plasmids is sufficient to direct the synthesis of small amounts of these proteins. One strategy that can be used to clone such toxic genes is to insert the gene near a strong *E. coli* promoter that is oriented so that transcription by the *E. coli* RNA polymerase results in the accumulation of RNA that is antisense to the toxic gene, reducing the level of its gene product. It is important to remember that the amount of a gene product synthesized is a function not only of the level of transcription but also of the efficiency at which translation is initiated. This is determined primarily by the ribosome-binding sequence located upstream of the start codon. Thus, some toxic genes with relatively weak ribosome-binding sequences can be cloned into multicopy plasmids, but not into a multicopy plasmid that also introduces a strong ribosome-binding sequence (S. Tabor, unpublished observations).

In summary, the first step in using the T7 RNA polymerase/promoter system is to clone the gene into an appropriate vector containing a p_{T7} and be certain it has an efficient ribosome-binding sequence. Once this is accomplished, the next step is determining whether the plasmid can stably transform an *E. coli* cell containing pGP1-2 at an efficiency comparable to that of the parent vector alone. If this is successful, the system is ready to be induced. If unsuccessful, it is necessary to induce the gene either by infection with M13 phage mGP1-2

(see second alternate protocol), or to use one of the more specialized vectors that further reduce the expression of T7 RNA polymerase in the cell under uninduced conditions (Studier et al., 1990).

Troubleshooting

For gene expression, one of the major advantages of the T7 RNA polymerase/promoter system over an *E. coli* RNA polymerase system is the ability to exclusively label the gene products under the control of p_{T7} . If the level of induction of the gene is estimated by inspection of a standard SDS-polyacrylamide gel, and it is difficult to see the expected induced product, then it is recommended that the induced proteins be labeled using [³⁵S]methionine as described in the first alternate protocol. This is a much more sensitive and specific assay for the specific protein production. Be sure that there is at least one methionine codon in the gene other than the one at the start of the protein (which is often removed in *E. coli*; Kirel et al., 1989); if not, then it is necessary to label with a cysteine or some other amino acid.

If it is not possible to detect the expected labeled product, there may be a problem with one of the two plasmids in the cell. One possibility is that the expressed protein is toxic to the cell, and as a result, a mutation has been selected for such that the toxic product is not synthesized. For more information on determining whether a gene is toxic, see the discussion on toxic genes in critical parameters. To determine if the cells and T7 RNA polymerase gene (e.g., pGP1-2) are inducing T7 RNA polymerase, attempt to induce a control protein that has been shown to work well in this system (e.g., the β -lactamase gene in pT7-1; Tabor and Richardson, 1985).

If the expressed protein does not accumulate significantly after induction, determine its stability in *E. coli* cells by pulse labeling with [³⁵S]methionine and chasing for various time periods with unlabeled methionine. If it is rapidly degraded, try to induce the gene in a protease-deficient strain. It should be noted that there are no known mutations that inactivate several very active *E. coli* proteases, and thus there is a strong probability that the mutant strains available (e.g., *lon*⁻) will have no effect on the stability of the gene product. In addition, such mutant strains generally grow poorly, and as a consequence the gene products are poorly produced upon induction of T7 RNA polymerase.

The most common reason for poor induc-

tion of a gene is that the translation does not initiate efficiently. Therefore, it is very important that there be an efficient ribosome-binding sequence the proper distance upstream of the gene. If a gene product does not induce well, and the problem is not the stability of the product, try a different ribosome-binding sequence—one that is known to work efficiently. The sequence and spacing between the ribosome-binding sequence and the start codon is critical. Because of this, it is recommended that the gene be inserted into a vector such as pT7-7, without altering any of the sequences between the ribosome-binding sequence and the start codon.

Anticipated Results

Under optimal conditions, the gene product expressed by the T7 RNA polymerase/promoter system can accumulate to >25% of the total cellular protein. However, in most instances the amount of gene product that accumulates is significantly less than this. There are numerous reasons for poor yields of gene product, as discussed in troubleshooting (see above).

Time Considerations

It should take ~1 week to insert the gene of interest into the pT7 vector, prepare minipreps of the DNA, and characterize the recombinants for the correct size and orientation of the insert. It should then take 3 days to transform the recombinant plasmid into the *E. coli* strain containing pGP1-2, induce the cells, and test the extracts for the production of the expected gene product.

Literature Cited

- Chen, W., Tabor, S., and Struhl, K. 1987. Distinguishing between mechanisms of eukaryotic transcriptional activation with bacteriophage T7 RNA polymerase. *Cell* 50:1047-1055.
- Dougan, G. and Sherratt, D. 1977. The transposon TnI as a probe for studying ColE1 structure and function. *Mol. Gen. Genet.* 151:151-160.
- Dunn, J.J., Krippel, B., Bernstein, K.E., Westphal, H., and Studier, F.W. 1988. Targeting bacteriophage T7 RNA polymerase to the mammalian cell nucleus. *Gene* 68:259-266.
- Fuerst, T.R., Niles, E.G., Studier, F.W., and Moss, B. 1986. Eukaryotic transient-expression system based on recombinant vaccinia virus that synthe-

sizes bacteriophage T7 RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 83:8122-8126.

- Kirel, P.-H., Schmitter, J.-M., Dessen, P., Fayat, G., and Blanquet, S. 1989. Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc. Natl. Acad. Sci. U.S.A.* 86:8247-8251.
- Nakai, H. and Richardson, C.C. 1986. Interactions of the DNA polymerase and gene 4 protein of bacteriophage T7. *J. Biol. Chem.* 261:15208-15216.
- Rosenberg, A.H., Lade, B.N., Chui, D., Lin, S., Dunn, J.J., and Studier, F.W. 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* 56:125-135.
- Russel, M. and Model, P. 1984. Replacement of the *flp* gene of *Escherichia coli* by an inactive gene cloned on a plasmid. *J. Bacteriol.* 159:1034-1039.
- Sancar, A., Wharton, R.P., Seltzer, S., Kacinsky, B.M., Clark, N.D., and Rupp, W.D. 1981. Identification of the *uvrA* gene product. *J. Mol. Biol.* 184:45-62.
- Studier, F.W. and Moffatt, B.A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189:113-130.
- Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. 1990. Use of T7 RNA polymerase to direct the expression of cloned genes. *Methods Enzymol.* 185. In press.
- Tabor, S. and Richardson, C.C. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. U.S.A.* 82:1074-1078.
- Tabor, S. and Richardson, C.C. 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 84:4767-4771.

Key References

- Studier et al., 1990. See above.
- Gives extensive list of vectors and protocols for expression using T7 RNA polymerase.*
- Tabor and Richardson, 1985. See above.
- Describes the use of the two-plasmid system for expression of genes using T7 RNA polymerase.*

Contributed by Stanley Tabor
Harvard Medical School
Boston, Massachusetts

Expression Using Vectors with Phage λ Regulatory Sequences

Many expression systems have been developed that utilize pBR322-based plasmids into which transcriptional and translational regulatory signals have been inserted. In the system described here, however, plasmids (pSKF) utilize regulatory signals—such as the powerful promoter p_L —from the bacteriophage λ . Transcription from p_L can be fully repressed and plasmids containing it are thus stabilized by the λ repressor, cI . The repressor is supplied by an *E. coli* host which contains an integrated copy of a portion of the λ genome. This so-called defective lysogen supplies the λ regulatory proteins cI and N but does not provide the lytic components that would normally lead to cell lysis. Thus, cells carrying these plasmids can be grown initially to high density without expression of the cloned gene and subsequently induced to synthesize the product upon inactivation of the repressor.

This system also ensures that p_L -directed transcription efficiently traverses any gene insert, which is accomplished by providing the phage λ antitermination function, N , to the cell and by including on the p_L transcription unit a site necessary for N utilization (Nut site). The N protein interacts with and modifies the RNA polymerase at the Nut site so as to block transcription termination at distal sites in the transcription unit.

In order to express the coding sequence, efficient ribosome-recognition and translation-initiation sites have been engineered into the p_L transcription unit. Expression occurs after temperature or chemical induction inactivates the repressor (see first and second basic protocols). Restriction endonuclease sites for insertion of the desired gene have been introduced both upstream and downstream from an ATG initiation codon. Thus, the system allows either direct expression or indirect expression (via protein fusion) of any coding sequence, thereby potentially allowing expression of any gene insert. Direct expression generates “authentic” gene products (first support protocol), while expression of heterologous genes fused to highly expressed gene partners generates chimeric proteins that differ from the native form. In the latter case, the fusion partner can be removed to obtain an unfused version of the gene product (second support protocol).

BASIC PROTOCOL

TEMPERATURE INDUCTION OF GENE EXPRESSION

Expression from p_L -containing vectors can be induced by raising the temperature. The *E. coli* lysogens used with these vectors are typically defective for phage replication and carry a temperature-sensitive mutation in the phage λ cI gene ($cI857$). After transformation and growth, induction is accomplished by raising the temperature of the culture from 32° to 42°C.

Materials

- Expression vector (e.g., pSKF series; see support protocols)
- E. coli* AR58 or equivalent (Table 1.4.5)
- LB plates containing the appropriate antibiotic (UNIT 1.1)
- LB medium containing appropriate antibiotic (room temperature and prewarmed to 65°C; UNIT 1.1)
- SDS/sample buffer (UNIT 10.2)
- Gyrotory air or water shaker, 32° and 42°C
- Additional reagents and equipment for transformation (UNIT 1.8)

1. Transform the expression vector into an *E. coli* λ lysogen (such as AR58) carrying a temperature-sensitive mutation in its repressor gene (λ $cI857$). Plate on LB/antibiotic plates and incubate transformants at 32°C.

Expression Using Vectors with Phage λ Regulatory Sequences

16.3.1

Contributed by Allan R. Shatzman, Mitchell S. Gross, and Martin Rosenberg

Current Protocols in Molecular Biology (1994) 16.3.1-16.3.11

Copyright © 2000 by John Wiley & Sons, Inc.

Heat-shock at 37° or 42°C for ≤90 sec during transformation is not a problem.

2. Grow the transformed cells overnight at 32°C in LB/antibiotic medium.
3. Dilute the overnight culture ≥1:20 into fresh LB/antibiotic medium. Grow the culture at 32°C in a gyrotory shaker at 250 to 300 rpm until OD₆₅₀ = 0.6 to 0.8.
4. Add 1/3 vol of 65°C LB/antibiotic medium with swirling in order to elevate the culture temperature rapidly to 42°C.

In our experience, a rapid increase in temperature favors production. Small shake-flask cultures (≤25 ml) are more easily induced by transfer to a 42°C gyrotory water bath without addition of prewarmed media. This generally raises the culture temperature to 42°C within 3 to 5 min.

5. Continue growing the culture 2 to 3 hr at 42°C.
6. Remove a 1-ml aliquot for analysis and harvest the remainder of cells by centrifuging 15 min in a low-speed rotor at 3000 × g, 4°C. Discard the supernatant.

Freeze cell pellet at -70°C until ready to isolate the gene product.

7. Spin the 1-ml aliquot 1 min at top speed in a microcentrifuge, then resuspend the pellet in 50 μl SDS/sample buffer. Boil 5 to 10 min and analyze gene product by SDS–polyacrylamide gel electrophoresis.

CHEMICAL INDUCTION OF GENE EXPRESSION

Expression using the pSKF system can also be induced chemically in lysogens that carry a wild-type (*ind*⁺) repressor gene (*cI857* cannot be used as it is *ind*⁻). This is accomplished by treating the bacterial host with an agent such as nalidixic acid. Nalidixic acid inhibits DNA gyrase and leads to DNA damage, which induces the SOS response. During the SOS response, wild-type repressor protein is cleaved. In this case, the wild-type repressor protein is cleaved by the RecA protease, which is induced by the SOS response. In contrast to induction by heat (product accumulates in 45 to 90 min) nalidixic acid–mediated induction of protein expression is comparatively slow (product accumulates in 5 to 6 hr).

Materials

Expression vector (e.g., pSKF series; see support protocols)

E. coli AR120 or equivalent (Table 1.4.5)

LB plates containing appropriate antibiotic (UNIT 1.1)

LB medium containing appropriate antibiotic (UNIT 1.1)

60 mg/ml nalidixic acid in 1 N NaOH (not necessary to filter sterilize; Table 1.4.1)

Additional reagents and equipment for transformation (UNIT 1.8)

1. Transform the expression vector into a replication-defective, *E. coli* *cI*⁺ lysogen (e.g., AR120). Plate on LB/antibiotic plates and incubate the transformants at 37°C.
2. Grow the transformed cells overnight at 37°C in LB/antibiotic medium.
3. Dilute the overnight culture ≥1:20 into fresh LB/antibiotic medium. Grow the culture at 37°C in a gyrotory shaker at 250 to 300 rpm until OD₆₅₀ = 0.4.
4. Add 1/1000 vol of 60 mg/ml nalidixic acid solution to give 60 μg/ml final concentration.
5. Continue growing the culture 5 to 6 hr at 37°C.
6. Harvest cells and analyze gene product (steps 6 and 7 of first basic protocol).

BASIC PROTOCOL

Protein Expression

16.3.2

AUTHENTIC GENE CLONING USING pSKF VECTORS

It is often most desirable to express a gene product in a form as similar to the native protein as possible. Such an “authentic” gene product will have the greatest chance of having a structure and activity identical to that of the native protein. Efficient translation of a coding sequence for an authentic gene product is typically accomplished by placing the inserted information immediately adjacent to a ribosome-binding site (a translational regulatory signal that interacts with the 16S rRNA of *E. coli* and contains an ATG initiation codon; Gold et al., 1981).

Strategic Planning

The translation-initiation signal utilized here is that of the phage λ *cII* gene. In order to make the translational information generally useful, the coding region of the gene has been removed from the vectors, leaving only their initiator *fMet* codon and upstream translational regulatory sequences. Additionally, these vectors have been engineered to provide restriction endonuclease sites on either side of the ATG, such that the initiation codon can be supplied by either the plasmid or the gene being inserted. Finally, restriction sites have also been engineered upstream of the translational regulatory region to permit insertion of other ribosome-binding sites. Those genes that contain restriction sites compatible with the sites on the vector may be inserted directly into the vector. As most genes do not contain appropriately positioned restriction sites, it is often necessary to adapt existing restriction cloning sites within the gene to fuse it to the translation-initiation signals provided by the vectors.

For example, pSKF101 (Fig. 16.3.1) and pSKF102 both have a *Bam*HI site adjacent to the initiation codon (ATGgatcc), while pSKF201 has an *Nco*I site (ccATGg) and pSKF301 (Fig. 16.3.2) has an *Nde*I site (catATG). The protocol presented below summarizes the steps to obtain an authentic gene clone using pSKF101 as an example.

Sample Protocol**Materials**

Appropriate restriction endonucleases and buffers (UNIT 3.1)

pSKF101 vector (available from A. Shatzman; Fig. 16.3.1)

Competent *E. coli* AS1 (Table 1.4.5; also known as MM294cI⁺)

Additional reagents and equipment for restriction digestion, (UNIT 3.1), oligonucleotide synthesis and purification (UNITS 2.11 & 2.12), nondenaturing PAGE (UNIT 2.7), isolation, recovery, and quantitation of DNA (UNIT 2.6 & APPENDIX 3), subcloning DNA fragments (UNIT 3.16), transforming, plating, and growing *E. coli* (UNITS 1.8, 1.1, & 1.3), and DNA miniprep (UNIT 1.6)

1. Identify a unique restriction endonuclease site close to the 5' end of the coding sequence of the gene to be expressed, as well as another unique site 3' to this gene's termination codon.
2. Synthesize two single-stranded DNA oligonucleotides, recreating the coding sequence immediately preceding the unique restriction endonuclease site near the 5' end of the gene to be expressed. Purify and quantitate the DNA, then anneal in order to obtain double-stranded DNA.

This synthetic DNA sequence is used to link the gene to be expressed to the initiating ATG of the pSKF expression vector. The double-stranded oligonucleotide should be designed to have ends that are complimentary to the restriction sites identified at the 5' end of the gene to be expressed as well as the chosen restriction site in the expression vector.

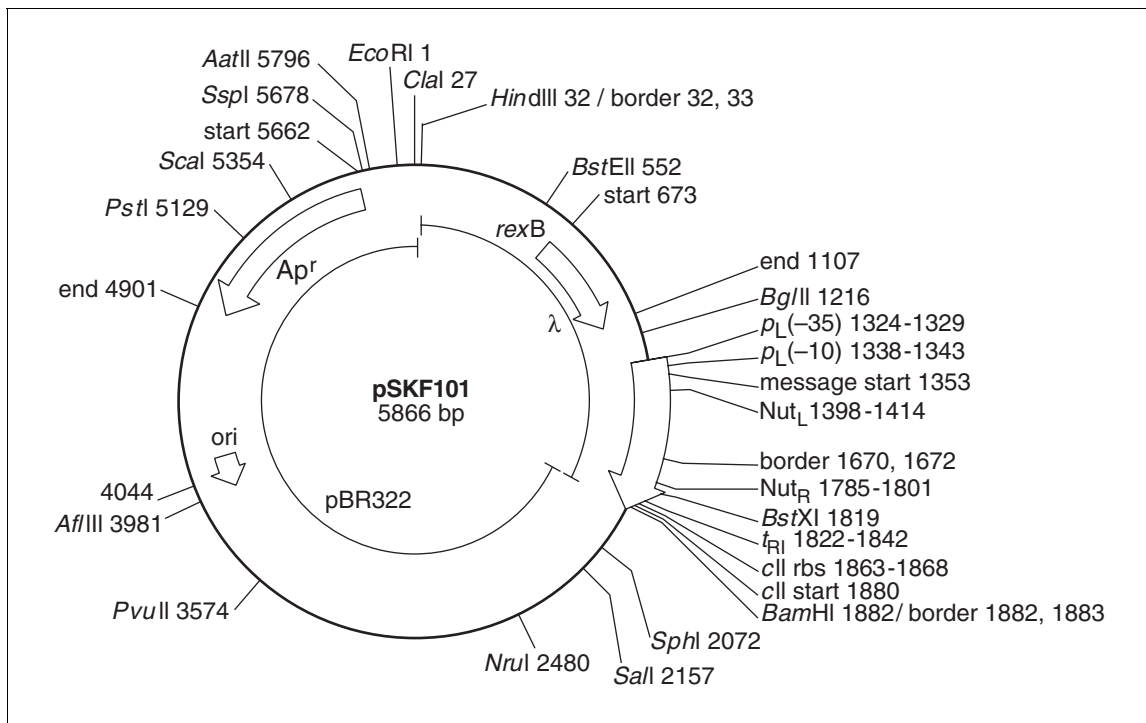


Figure 16.3.1 pSKF101. pSKF101 is a vector used for authentic gene cloning which allows direct expression of the inserted gene. It is a derivative of pBR322 (UNIT 1.5) containing sequences inserted between *Hind*III and *Bam*HI sites of pBR322. The inserted λ sequences contain the p_L promoter and *cII* ribosome-binding site (rbs); these are the transcriptional and translational regulatory sequences necessary to express heterologous genes in *E. coli*. Within this region are several unique restriction sites that permit insertion of the gene. The regions derived from pBR322 and λ are indicated. This plasmid can be maintained stably in a λ -lysogenized *E. coli* strain. The selectable marker is ampicillin, encoded by β -lactamase.

An alternate name for pSKF101 is pAS1 (Rosenberg et al., 1983). Alternative names of related vectors are as follows: pSKF102 is pOTSV (Shatzman and Rosenberg, 1987); pSKF201 is pOTS-Nco (Shatzman and Rosenberg, 1987); and pSKF301 is pMG1.

3. Digest 25 to 50 μ g plasmid DNA containing the gene to be expressed with the restriction endonucleases identified in step 1.

To ensure complete digestion, determine that the restriction endonuclease buffer is appropriate for each enzyme to be used. If the endonucleases require different buffers, then each restriction digestion must be done separately.

4. Electrophorese the doubly digested plasmid DNA on a polyacrylamide gel.

If the DNA fragment to be isolated is between 150 and 1100 bp, a 6% gel can be used. Either a borate- or acetate-buffer system can be used. If digestion was done in a large volume, ethanol precipitate the DNA (UNIT 2.1) and resuspend in 40 to 100 μ l TE buffer. Mix with loading dye and load.

5. Locate the fragment of interest by staining with an agent such as ethidium bromide and cut the DNA fragment out of the gel.

6. Recover the DNA by electroelution and quantitate the amount of DNA.

Confirm that the correct fragment has been isolated by running a small aliquot on an agarose gel. Be sure to run appropriate size markers in an adjacent lane.

7. Digest 10 μ g pSKF101 with *Bam*HI and a restriction endonuclease that generates ends compatible with the 3' end of the coding sequence (to accommodate the 3' end of the gene to be expressed).

Confirm that complete digestion of the vector has occurred by analysis of digested DNA

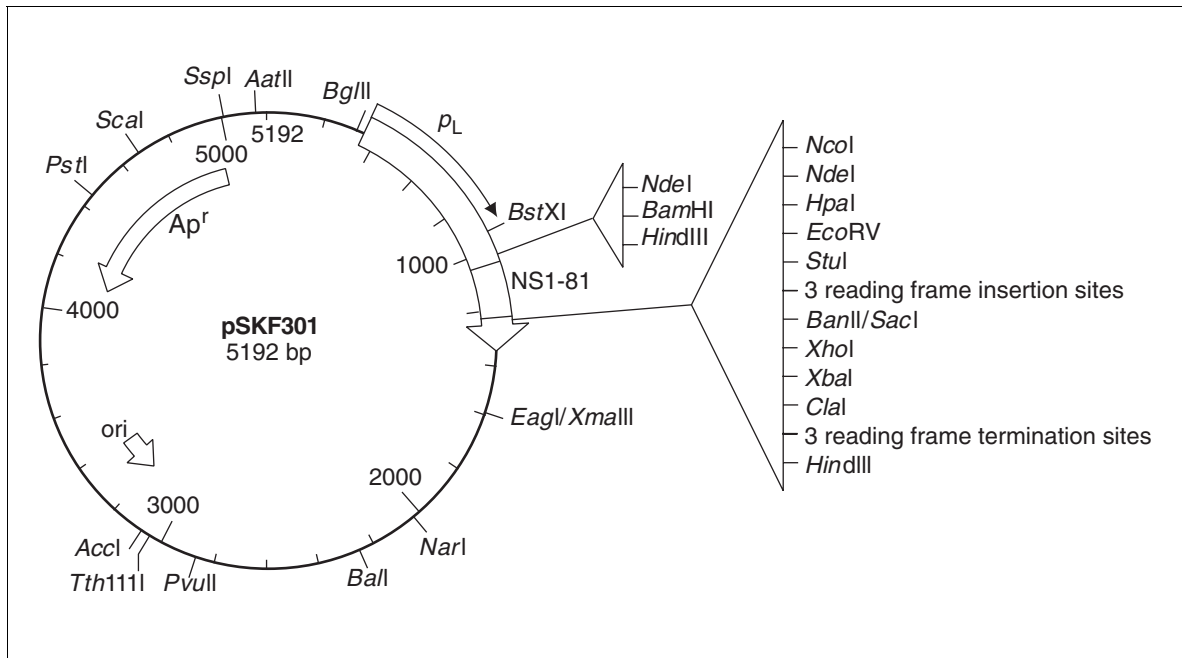


Figure 16.3.2 pSKF301. pSKF301 is a vector that can be used for both indirect and direct expression. It is similar to pSKF101 in that it contains the same transcriptional and translational regulatory sequences as well as selectable markers; it differs in that it contains a shorter segment of λ DNA than pSKF101. pSKF301 also contains the coding sequence of the first 81 amino acids of the influenza protein, NS1, shown as NS1-81. This region is adjacent to the *dl* ribosome-binding site (fbs) and contains restriction sites at the 3' end of NS1-81 that allow construction of translational fusions in any of the three reading frames. Removal of NS1-81 permits direct expression of the cloned gene. (This vector is also known as pmG1.)

on an agarose gel. Compare undigested pSKF101 with digested to make sure that pSKF101 has been completely linearized.

8. Prepare a ligation reaction (using the conditions described in UNIT 3.16) by combining the following ingredients:

- 1 ng digested pSKF101 vector DNA
- 10 ng of the gene fragment to be expressed (from step 6)
- 20 ng synthetic oligonucleotide (from step 2)
- T4 DNA ligase.

Ligate 10 to 12 hr at 4°C.

There is no need to dephosphorylate pSKF101 as long as there is at least a 5-fold molar excess of vector DNA to isolated DNA fragment and synthetic DNA.

9. Remove one-third of the ligation reaction and transform 50 to 100 μ l competent *E. coli* AS1. Plate on LB/ampicillin plates and incubate overnight at 37°C.
10. Pick 12 to 24 colonies and transfer with a sterile toothpick to 3 ml LB/ampicillin medium. Grow cells 5 to 18 hr and isolate DNA by a miniprep method.

Cells may be harvested once the broth appears turbid. For best results, allow 8 to 12 hr of growth.

11. Perform appropriate restriction endonuclease digests to determine which clones contain the desired construction of the gene to be expressed.
12. Transform an *E. coli* strain with the DNA and express the gene as in the basic protocols.

CONSTRUCTION AND DISASSEMBLY OF FUSED GENES IN pSKF301

By fusing the gene to be expressed to a coding region of another gene (the fusion partner), a chimeric gene can be constructed in an appropriate vector. Numerous vectors are available for this purpose but most share the common feature of a fusion partner that is a highly expressed gene. When expression of the chimeric gene is induced, the resulting proteins carry additional peptide information at the N terminus. Although the fusion product may have physical and/or functional properties that differ from the “authentic” protein, advantages of the approach include highly efficient expression (up to 30% of total cell protein) without complicated alterations on the gene, and the presence of a “handle” on the expressed protein which can help to identify and purify it. Such proteins are often used to develop antisera to specific proteins that have diagnostic potential, and have been used successfully to identify and define a variety of gene products (Casadaban et al., 1983; Rose and Botstein, 1983; Guarente, 1983).

Strategic Planning

Plasmid pSKF301 has been constructed to permit initial expression of a gene as a fusion product, followed by removal of the DNA encoding the fusion portion by restriction digestion. Finally, the unfused version of the gene is expressed as an authentic protein.

pSKF301 contains an *NdeI* restriction site adjacent to the ATG following the *cII* ribosome-binding site (Fig. 16.3.3). This ATG also serves as the translational start (Gold et al., 1981) of the *NS1* gene derived from the influenza nonstructural gene. This gene has been truncated to express only its first 81 amino acids. Just beyond the coding sequence for the 81st amino acid is a second *NdeI* site followed by three unique blunt-ended restriction sites, *HpaI*, *EcoRV*, and *StuI*, which allow for the insertion of genes into any of three reading frames. Immediately following the *StuI* site are sequences coding for translational stops in any of the three reading frames.

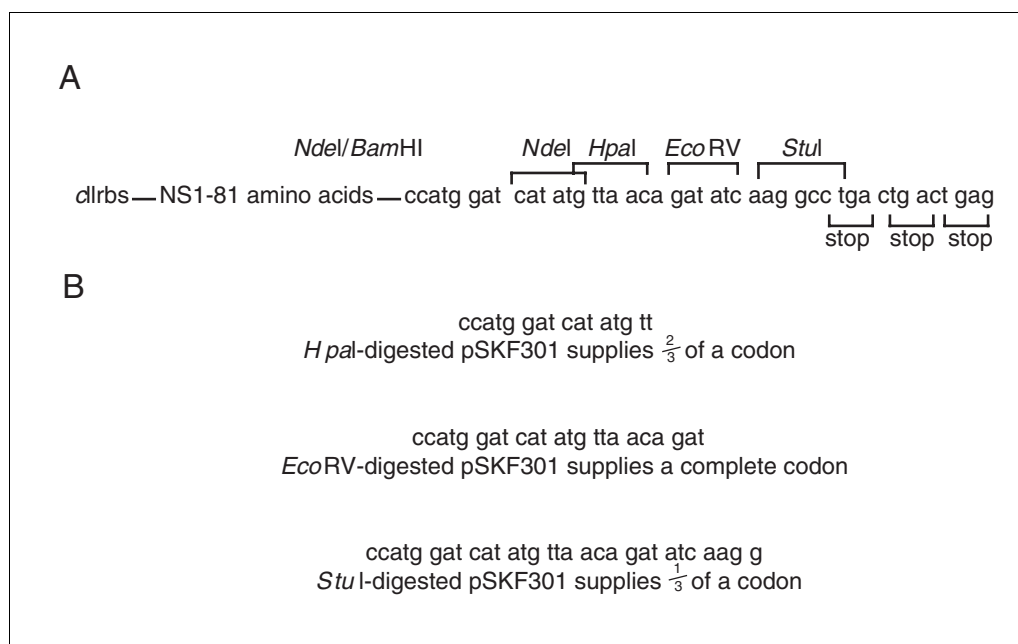


Figure 16.3.3 Sequence and restriction endonuclease sites (in the region used for cloning) of pSKF301 (**A**). Restriction endonuclease digestion shows the strategy utilized to obtain pSKF301 as a vehicle for expression in all three reading frames (**B**).

The expression of a gene of interest as a fusion protein may be achieved by utilizing any of the following restriction sites in pSKF301: *Nco*I, *Hpa*I, *Eco*RV, or *Stu*I. Choice of restriction site depends upon the reading frame necessary for the translation of a specific protein sequence. First, a unique restriction site close to the 5' end of the gene (or portion of the gene) to be expressed must be identified. Second, the appropriate restriction endonuclease is selected for digesting pSKF301 such that the gene will be expressed.

If the chosen restriction site is a blunt-end cutter, no further manipulation of that end is required. In the event the restriction site identified leaves either a 5' or 3' protruding end, further manipulation is required. "Filling in" using the Klenow fragment of *E. coli* DNA polymerase for 5' protrusions, or T4 DNA polymerase, S1, or mung bean nuclease for 3' protrusions, are methods of choice (see UNIT 3.16).

Sample Protocol

Materials

- Appropriate restriction endonucleases and buffers (UNIT 3.1)
- Klenow fragment of *E. coli* DNA polymerase I (UNIT 3.5)
- pSKF301 vector (available from A. Shatzman; Figs. 16.3.1 & 16.3.2)
- T4 DNA ligase (UNIT 3.14)
- Competent *E. coli* AS1 (Table 1.4.5; also known as MM294cI⁺)
- Additional reagents and equipment for large-scale plasmid prep (UNIT 1.7), agarose gel electrophoresis (UNIT 2.5), extraction and precipitation of DNA (UNIT 2.1), transformation of competent cells (UNIT 1.8), and restriction digestion and mapping (UNITS 3.1-3.3)

Construct a gene fusion in pSKF301

1. Assume the restriction site identified in the gene is a *Bam*HI site. Digest with *Bam*HI to obtain:

GATCC	XXX	XXX	XXX	XXX
G	YYY	YYY	YYY	YYY

2. Treat with Klenow fragment to fill in the unpaired bases to obtain:

GATCC	XXX	XXX	XXX	XXX
CTAGG	YYY	YYY	YYY	YYY

As noted above, Klenow fragment is used to fill in for 5' protrusions. For 3' protrusions, use T4 DNA polymerase (UNIT 3.5) or S1 or mung bean nuclease (UNIT 3.12).

3. Determine the proper reading frame of the gene. In this example assume XXX XXX XXX XXX is the proper reading frame; therefore, the coding sequence of the filled-in fragment should read:

GA	TCC	XXX	XXX	XXX
----	-----	-----	-----	-----

4. Determine which restriction endonuclease should be used to digest pSKF301 to allow expression of the fusion protein. For this example, *Stu*I is required to yield:

ccatg	gat	cat	atg	tta	aca	gat	atc	aag	gGA	TCC	XXX	XXX	XXX	XXX
pSKF301									fusion gene					

5. Prepare the vector and the fragment of the gene to be expressed as in the first support protocol, steps 3 to 12 (*except* no synthetic DNA is required).

Generate an authentic version of the gene

Once a gene has been expressed as a fusion protein, it may be desirable to obtain an unfused version of the gene product. If this is useful, follow steps 6 to 12.

To convert a fusion protein to an unfused protein when using pSKF301, be certain that the gene of interest does not contain an *NdeI* site. The following theoretical fusion construct will be used as an example in these steps:



6. Set up a large-scale plasmid preparation of the fusion construct to yield ~100 µg plasmid DNA.
7. Digest 10 µg of the construct with *NdeI*. Verify that all of the vector DNA has been completely digested by taking a small aliquot of the digested material and running it on an agarose gel next to lanes containing uncut plasmid and appropriate size markers. A 280-bp fragment should be observed; this contains the NS1-81 gene sequence being liberated from the construct.

Confirmation of complete digestion is extremely important.

8. Purify the digested construct by phenol/chloroform/isoamyl alcohol extraction followed by ethanol precipitation.
9. Add T4 DNA ligase to 1 µg of the *NdeI*-digested construct and incubate overnight at 4°C.
10. Transform ligated DNA into competent *E. coli* AS1 cells (or any other suitable *cI*⁺ lysogen).
11. Determine that the construct no longer contains the NS1-81 gene sequence by restriction analysis.

*Consult the restriction map of pSKF301 and the gene to be expressed to determine which endonucleases are diagnostic for identifying the construct devoid of the NS1-81 gene. If the *NdeI* digestion was complete upon ligation, reclosure is highly efficient. Expect 95% to 100% of the resulting transformants to contain the unfused construct.*

12. Transform the DNA and express the gene by temperature or chemical induction as in the basic protocols.

COMMENTARY

Background Information

Expression of a heterologous gene or gene fragment in *E. coli* requires that the coding sequence be placed under the transcriptional and translational control of regulatory elements recognized by the bacterial cell. The pSKF vectors were designed specifically to direct gene expression by providing regulatory signals from bacteriophage λ. Phage regulatory signals were chosen because of their high efficiency and ability to be tightly regulated. This system uses a promoter that can be tightly controlled, eliminating problems with “leaky”

basal expression sometimes found in other expression systems (see below). This system uses an antitermination mechanism to help assure efficient transcription across any gene insert. The different vectors used with this system offer several choices of antibiotic selection markers, contain elements that optimize plasmid stability, and carry a variety of restriction sites that permit relatively easy insertion of the gene of interest adjacent to the efficient translation regulatory information.

The pSKF system offers some advantages that differentiate it from many other expression

systems. Perhaps most important is the “tightness” of regulation of the p_L promoter. Several other strong regulatable promoters— p_{tac} (de Boer et al., 1982), p_{trp} (Edman et al., 1981), and p_{T7} (promoter of T7 gene 10; Studier and Moffatt, 1986)—are also used routinely for optimizing heterologous gene expression in *E. coli*. These promoters, along with p_L , are all of comparable strength and are sufficient to achieve very high levels of mRNA production (UNIT 16.1). In fact, these promoters are so powerful that further enhancement of promoter strength would not be expected to result in an increase of protein production; indeed, these promoters are so strong that it is very difficult to keep them fully turned off even in the “repressed” state. Because of this basal transcription under repressed conditions, use of the p_{tac} , p_{trp} , or p_{T7} (coupled with the p_{tac} -T7 polymerase) systems often leads to some expression of the cloned gene even under nonpermissive conditions. This may lead to plasmid loss or rearrangement, or possibly cell death, if small amounts of the gene product are lethal to *E. coli*. In contrast, one does not typically see expression of the cloned gene in the p_L system until cultures have been induced.

A second advantage of the p_L system over other promoter systems is the flexibility gained from having completely different induction systems (thermal and chemical). In contrast, the p_{tac} and p_{T7} systems mentioned above permit induction only by a chemical route. Different routes of induction lead to completely different cellular states (e.g., different physiology, morphology, and growth patterns) and these variations can lead to significant differences in gene product accumulation and stability (unpub. observ.). It should be pointed out that p_{trp} also permits dual modes of induction (by β -indolyl acetic acid or Trp starvation) and that a different version of the p_{T7} system has been developed (UNIT 16.2; Tabor and Richardson, 1985) in which the T7 RNA polymerase is thermally regulated via the p_L -cI857 system (however, this system is not chemically inducible as well).

The third major advantage of the pSKF system is the availability of a single vector that permits expression of either an authentic or a fusion gene product; furthermore, this vector allows the fusion gene to be converted to an authentic gene by a simple restriction digest followed by self ligation. Thus, a gene may be rapidly expressed at high levels as a protein fusion to give an initial reagent for use in activity studies and antisera preparation. Time

may then be taken to optimize the expression of the authentic (nonfusion) gene product, which will be better suited for functional and structural studies. Most other expression systems do not provide this flexibility.

Critical Parameters

Gene expression is not solely a function of message levels. The efficiency of the ribosome-binding region—including the sequences both upstream and downstream of the ATG initiation codon—also play a role in determining the extent to which a protein is made. Alterations in these sequences may affect the secondary structure of a message and the conformational presentation of the initiation signals which, as a result, can alter translational efficiency (Gold et al., 1981).

From our experience, the host strain plays a major role in determining the ultimate levels of gene expression. The reasons for the rather dramatic differences seen in product yield from different host strains are poorly understood. Product stability is, however, one determining factor that has been somewhat characterized. Host strains have been developed that are defective in certain proteases (UNIT 16.6). These specialized host strains can have a significant impact on the expression of certain gene products. However, proteases are not the only factor involved in strain-to-strain variations observed in protein expression. Other uncharacterized factors can have equally dramatic effects. It is therefore recommended that expression be tried in a number of different *E. coli* strains.

Following the induction of cultures carrying the desired expression vector, cells may be analyzed in a variety of ways to detect the presence of the cloned gene product. Most typically, the presence of the novel gene product is determined directly by observing in SDS-polyacrylamide gels a new, inducible protein band not present in lanes from control cultures.

The expression of any gene insert can also be identified and/or confirmed in several ways related to the activity or function of the protein including: (1) direct detection of a novel function or activity imparted to the living bacterial host; (2) genetic complementation of the appropriate mutant host; (3) assay of whole-cell extracts for the activity of the cloned gene product; and (4) assay after partial or complete purification of the cloned gene product.

Immunochemical methods such as immunoprecipitation (UNIT 10.16) or western blotting (UNIT 10.8) are some of the most sensitive meth-

ods available to detect expression of a gene product. These methods, of course, require that an antiserum be available which is specific for the protein to be expressed. These methods, however, are primarily quantitative and do not necessarily indicate anything about the level of expression, homogeneity, or activity of the gene product.

If a good antiserum to the protein of interest is not available, purification of sufficient amounts of a gene product allows generation of high-titer, antigen-specific mono- or polyclonal antisera (UNITS 11.3-11.13). One approach to generate an antiserum is to produce the desired heterologous gene product in bacteria as a native protein, as a fusion, or as a protein fragment. The protein may then be purified and used to produce high-titer mono- or polyclonal antisera. Such antisera have been used to (1) map natural expression of the gene product with respect to cell type, subcellular distribution, and temporal regulation; (2) determine relative levels of expression in various cell types; (3) study protein processing and stability; (4) map immuno-dominant domains; (5) purify by immunoaffinity both the native and modified forms of the protein; and (6) provide in vivo diagnostic reagents for examining tissue distribution and expression of the gene product by immunofluorescent methods.

Troubleshooting

There is never a guarantee that a gene will be expressed at high levels, but poor expression upon initial trials does not signify defeat. As mentioned earlier, transcription is rarely limiting and is, therefore, not the first parameter to be addressed in attempting to improve expression. Instead, the easiest parameter to change is the host strain being used for production. Typically, five or six different strains (which might or might not be closely related to each other) may have to be tested in pilot experiments to see which gives optimal production.

The next parameter to examine in the event of poor expression is translation. Expression may be increased by altering ribosome-binding sites to improve complementarity to the 16S rRNA, or by increasing the A-T richness of the 5'-end of the gene's coding region.

After steps have been taken to optimize translation, it is often helpful to alter the promoter and repressor system in order to change the induction system and the physiology of the cells during the production phase. For example, inducing the *cI857*-containing p_L system via a temperature shift generates a cellular heat-

shock response and protein synthesis at 42°C. Induction of this system with nalidixic acid leads to a cellular SOS response (see glossary, Chapter 1 introduction) and protein production at 37°C. Induction of the *trp* system by tryptophan starvation turns on the host stringent response (a generalized response of *E. coli* to amino acid starvation). Thus, in each case, a different host response leads to induction of a different set of host proteins as well as to greatly different physiological effects (such as changes in respiration, filamentation, and growth rate).

Finally, it may be possible to improve expression by optimizing the temperature at which the protein is made, as this parameter has often been shown to affect the proteins' solubility, stability, and activity.

Anticipated Results

Expression of most gene products as fusions with the first 81 amino acids of the NS1 protein (using pSKF301) can be achieved at levels between 5% and 30% of total cellular protein. Expression levels of nonfusion proteins (authentic) are less predictable and may vary from <1% to 30% of total cell protein. By systematically optimizing each of the parameters described in the troubleshooting section, it may be possible to increase the level of expression of a nonfusion gene product from the low end of this range to the high end. In shake flask cultures this is equivalent to hundreds of milligrams per liter and in fermenters (where greater cell densities are achieved) to grams per liter of the desired product. The expressed product can usually be visualized by running a small aliquot of cell extract on an SDS-polyacrylamide gel and staining with Coomassie Brilliant Blue. The majority of the protein produced will be insoluble, but can often be solubilized and renatured to an active state (UNIT 16.5).

Time Considerations

Cell growth, induction, and harvesting require 6 to 8 hr depending on the strain of *E. coli* and the mode of induction used. Following harvest, the cell pellets may be frozen at -70°C for long periods with no obvious loss of gene product. It is often convenient to analyze a small aliquot of the induced culture (removed prior to harvesting the remainder of the culture) by SDS-PAGE the next day, as this step will require several hours including gel preparation, gel running, staining, and destaining. Once it is clear that sufficient levels of protein have been produced to merit purification efforts, cells may be lysed and product extracted (if insoluble) by

use of detergents and/or chaotropes. This will take 1 to 2 days depending on the number of extraction steps required and the length of dialysis steps chosen between extractions.

Literature Cited

- Casadaban, M., Martinez-Arias, A., Shapira, S., and Chou, J. 1983. Translation initiation in prokaryotes. *Methods Enzymol.* 100:293-308.
- de Boer, H.A., Comstock, L.J., Yansura, D., and Heynecker, H. 1982 *In* Promoters: Structure and Function (R.L. Rodriguez and M.J. Chamberlin, eds.) pp. 462-481. Praeger, New York.
- Edman, J.C., Hallewell, R.A., Valenzuela, P., Goodman, H.M., and Rutter, W.J. 1981. Synthesis of hepatitis B surface and core antigens in *E. coli*. *Nature (London)* 291: 503-506.
- Gold, L., Pribnow, D., Schneider, T., Slinedling, S., Singer, B., and Stormo, G. 1981. β -Galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. *Annu. Rev. Microbiol.* 35:365-403.
- Guarente, L. 1983. Construction and use of gene fusions to *lac Z* (β galactosidase) that are expressed in yeast. *Methods Enzymol.* 101:167-181.
- Rose, M. and Botstein, D. 1983. Yeast promoters and *lacZ* fusions designed to study expression of cloned genes in yeast. *Methods Enzymol.* 101:181-192.
- Rosenberg, M., Ho, Y.S., and Shatzman, A.R. 1983. The use of pKC30 and its derivatives for controlled expression of genes. *Methods Enzymol.* 101:123-138.
- Shatzman, A.R. and Rosenberg, M. 1987. Expression, identification and characterization of recombinant gene products in *Escherichia coli*. *Methods. Enzymol.* 152:661-673.
- Studier, F.W. and Moffatt, B. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189:113-130.
- Tabor, S. and Richardson, C. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. U.S.A.* 82:1074-1078.

Contributed by Allan R. Shatzman,
Mitchell S. Gross, and Martin Rosenberg
SmithKline Beecham Pharmaceuticals
King of Prussia, Pennsylvania

Introduction to Expression by Fusion Protein Vectors

UNIT 16.4A

Expression—the directed synthesis of a foreign gene—is often the logical next step for researchers who have isolated a gene and want to study the protein it encodes. During the early days of recombinant DNA technology, it was thought that a strong promoter and a start codon at the beginning of the gene would be sufficient for good expression in *Escherichia coli*. Since then it has been learned that the requirements for efficient translation are a good deal more complicated. In addition to a promoter and a start codon, good expression requires that the mRNA encoding the protein to be expressed contain a ribosome-binding site that is not blocked by mRNA secondary structure. The level of expression is also affected by codon preferences, especially in the second codon of the gene (Stormo et al., 1982), and may be affected by the coding sequence in other ways that are not yet well understood (UNIT 16.1). In virtually all cases, these problems can be solved by altering the sequence preceding the start codon, and/or by making changes in the 5' end of the coding sequence that do not change the protein sequence, taking advantage of the degeneracy of the genetic code.

However, it is often quicker to solve these problems by making fusions between genes. In this approach the cloned gene is introduced into an expression vector 3' to a sequence (carrier sequence) coding for the amino terminus of a highly expressed protein (carrier protein). The carrier sequence is often from an *E. coli* gene, but it can be from any gene that is strongly expressed in *E. coli*. The carrier sequence provides the necessary signals for good expression, and the expressed fusion protein contains an N-terminal region encoded by the carrier. In such vectors, the portion of the fusion protein encoded by the carrier can be as small as one amino acid (UNIT 16.3; Amann and Brosius, 1985), although expression from such vectors can still be subject to problems caused by the coding sequence of the expressed protein. Perhaps more typical examples of short carrier sequences are those contained in the *trpE* vectors (UNIT 16.5) or the λ cII vectors (Nagai and Thøgersen, 1987).

The carrier sequence can also code for an entire functional moiety or even for an entire protein. For example, the following four units

(UNITS 16.5-16.8) describe the use of vectors that express β -galactosidase and *trpE* fusions, maltose-binding protein (MBP) fusions, glutathione-S-transferase (GST) fusions, and thioredoxin (Trx) fusions. These carrier regions often can be exploited in purifying the protein, either with antibodies or with an affinity purification specific for that carrier protein. Alternatively, unique physical properties of the carrier protein (e.g., heat stability) can be exploited to allow selective purification of the fusion protein. In addition, some carrier proteins such as MBP and Trx can be selectively released from intact cells by osmotic shock or freeze/thaw procedures, even though they reside in different cellular compartments. Often, proteins fused to these carriers can be separated from the bulk of intracellular contaminants by taking advantage of this attribute.

There are three problems often encountered when expressing fusion proteins: solubility of the expressed protein, stability of the expressed protein, and presence of the carrier protein. The first two problems are often encountered with both fusion and nonfusion expression systems (UNIT 16.1), while the third is unique to fusion systems.

SOLUBILITY OF THE EXPRESSED PROTEIN

The high-level expression of many proteins can lead to the formation of *inclusion bodies*, very dense aggregates of insoluble protein and RNA that contain most of the expressed protein (Schein, 1989). Precipitation of a protein into inclusion bodies sometimes can work to one's advantage, because inclusion bodies are insoluble and dense, and can be purified relatively easily by centrifugation (UNIT 16.5). In addition, some proteins that are degraded when expressed in the soluble fraction are quite stable as inclusion bodies. Once purified, protein in inclusion bodies can be solubilized by denaturation with guanidine·HCl or urea, and then can often be refolded by dialyzing away the denaturant. A problem, however, with denaturation/renaturation is that the yield of properly refolded protein is variable and sometimes quite low; some proteins, especially large ones, cannot be properly refolded at all (see UNIT 16.5).

If expression of a particular fusion protein produces insoluble aggregates and a soluble protein is required, there are several things to try. One important variable is temperature; for reasons not well understood, higher temperatures (37° and 42°C) promote inclusion-body formation and lower temperatures (30°C) inhibit it (Bishai et al., 1987; Schein, 1989). Another variable is the level of expression; sometimes lowering the expression level can increase the proportion of protein that is soluble. A third variable is the strain background of the cells bearing the expression vector; large differences in the proportion of a particular expressed protein that is soluble are seen among different strains (M. Southworth, S. Levitt, and F. Perler, unpub. observ.; it is not known which of the genetic differences between the strains is responsible for the differences in solubility). Finally, it is worth noting that changes in the carrier protein can affect the solubility of an expressed fusion protein (La Vallie et al., 1993).

STABILITY OF THE EXPRESSED PROTEIN

Stability problems are often encountered when foreign proteins, especially eukaryotic proteins, are expressed in *E. coli*. The carrier protein can sometimes stabilize an expression fusion protein (Lee et al., 1984). Sometimes, however, the expressed protein is degraded but the carrier protein is not. Moreover, fusion proteins are sometimes cleaved in vivo at the fusion joint between the carrier and expressed portions of the fusion, which obviously creates problems if the carrier protein is to be used as an aid in purification. These facts about fusion proteins are consistent with a model in which the carrier and the rest of the protein form independent domains. In this view, it can be imagined that there are cases where the carrier domain folds correctly and the expressed protein does not (and is degraded). There are also cases where both domains fold correctly but the joint region between them is sensitive to one or more *E. coli* proteases.

Approaches that have been used to stabilize fusion proteins are generally the same as those used to stabilize nonfusion proteins. One method is to arrange for the fusion protein to be expressed as insoluble aggregates. Another method is to use *E. coli* strains deficient in known proteases. For example, a *lon htpR* double-mutant strain—which is deficient in several cytoplasmic proteases—shows reduced degra-

ation of unstable proteins (Baker et al., 1984; I. Hall and P. Riggs, unpub. observ.). Similarly, the *degP* mutant has been shown to stabilize fusion proteins in the periplasm (Strauch and Beckwith, 1988) and *ompT* mutants have proven useful in preventing cleavage between exposed basic residues (e.g., Arg-Arg) in several nonfusion proteins during preparation of crude extracts (Grodberg and Dunn, 1988; Sugimura and Higashi, 1988). Finally, the stability of a particular fusion can vary even among different “wild-type” lab strains, perhaps due to uncharacterized differences in protease levels among the strains (I. Hall, P. Riggs, M. Southworth, S. Levitt, and F. Perler, unpub. observ.).

CLEAVAGE OF FUSION PROTEINS TO REMOVE THE CARRIER

The use of fusion proteins is growing rapidly for the many reasons described above. The various systems described in the following units have been used to produce many different kinds of proteins ranging from enzymes and growth factors to transmembrane receptors and DNA binding proteins. Often it is advantageous to remove the carrier protein moiety from the protein of interest to facilitate biochemical and functional analyses. Several methods for site-specific cleavage of fusion proteins have been developed (UNIT 16.4B). The choice of method is usually determined by the composition, sequence, and physical characteristics of the particular protein. Chemical cleavage of fusion proteins can be accomplished with reagents such as cyanogen bromide (Met↓, Itakura et al., 1977), 2-(2-nitrophenylsulphenyl)-3-methyl-3'-bromoindolenine (BNPS-skatole, Trp↓, Dykes et al., 1988), hydroxylamine (Asn↓Gly, Bornstein and Balian, 1977), or low pH (Asp↓Pro, Szoka et al., 1986). Chemical cleavage procedures tend to be inexpensive and efficient, and often can be accomplished under denaturing conditions to cleave otherwise insoluble fusion proteins (Szoka et al., 1986). However, their use is hampered by the likely occurrence of cleavage sites in the protein of interest, along with the propensity for side reactions that result in unwanted modifications to the protein. As an alternative to chemical methods, enzymatic cleavage procedures are desirable for their relatively mild reaction conditions and, most importantly, for the high degree of specificity exhibited by some proteases commonly used for this purpose. Among the useful enzymes are factor Xa (Nagai and

Thøgersen, 1984, 1987; Gardella et al., 1990), thrombin (Smith and Johnson, 1988; Gearing et al., 1989), enterokinase (Dykes et al., 1988; LaVallie et al., 1993), renin (Haffey et al., 1987), and collagenase (Germino and Bastia, 1984). All of these enzymes have extended substrate recognition sequences (up to 7 amino acids in the case of renin), which greatly reduces the likelihood of unwanted cleavages elsewhere in the protein. Of the above-mentioned proteases, factor Xa and enterokinase are most useful in this application because they cleave on the carboxy-terminal side of their respective recognition sequences, allowing the release of fusion partners containing their authentic amino-termini.

UNITS 16.5, 16.6, 16.7 & 16.8 describe five different fusion protein vector systems; of these, only three include recognition sites for interdomain cleavage. The MBP fusion system (UNIT 16.6) provides a factor Xa cleavage site. The GST fusion system (UNIT 16.7) includes vectors that contain either a thrombin cleavage site, a factor Xa cleavage site, or an Asp-Pro acid cleavage site. The Trx fusion system (UNIT 16.8) uses an enterokinase cleavage site. UNIT 16.4B describes fusion protein cleavages in detail, including specific protocols for cleaving fusion proteins produced with each of the aforementioned vector systems, along with methodologies for the site-specific cleavage of proteins using various chemical reagents.

Literature Cited

- Amann, E. and Brosius, J. 1985. "ATG vectors" for regulated high-level expression of cloned genes in *Escherichia coli*. *Gene* 40:183-190.
- Baker, T.A., Grossman, D., and Gross, C.A. 1984. A gene regulating the heat shock response in *Escherichia coli* also affects proteolysis. *Proc. Natl. Acad. Sci. U.S.A.* 81:6779-6783.
- Bishia, W.R., Rappuoli, R., and Murphy, J.R. 1987. High-level expression of a proteolytically sensitive diphtheria toxin fragment in *Escherichia coli*. *J. Bacteriol.* 169:5140-5151.
- Bornstein, P. and Balian, G. 1977. Cleavage at Asn-Gly bonds with hydroxylamine. *Methods Enzymol.* 47:132-145.
- Dykes, C.W., Bookless, A.B., Coomber, B.A., Noble, S.A., Humber, D.C., and Hobden, A.N. 1988. Expression of atrial natriuretic factor as a cleavable fusion protein with chloramphenicol acetyltransferase in *Escherichia coli*. *Eur. J. Biochem.* 174:411-416.
- Gardella, T.J., Rubin, D., Abou-Samra, A.-B., Keutmann, H.T., Potts, J.T. Jr., Kronenberg, H.M., and Nussbaum, S.R. 1990. Expression of human parathyroid hormone-(1-84) in *Escherichia coli* as a factor X cleavable fusion protein. *J. Biol. Chem.* 265:15854-15859.
- Gearing, D.P., Nicola, N.A., Metcalf, D., Foote, S., Willson, T.A., Gough, N.M., and Williams, R.L. 1989. Production of leukemia factor in *Escherichia coli* by a novel procedure and its use in maintaining embryonic stem cells in culture. *BioTechnology* 7:1157-1161.
- Germino, J. and Bastia, D. 1984. Rapid purification of a gene product by genetic fusion and site-specific proteolysis. *Proc. Natl. Acad. Sci. U.S.A.* 81:4692-4696.
- Grodberg, J. and Dunn, J.J. 1988. *ompT* encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *J. Bacteriol.* 170:1245-1253.
- Haffey, M.L., Lehman, D., and Boger, J. 1987. Site-specific cleavage of a fusion protein by renin. *DNA* 6:565-571.
- Itakura, K., Hirose, T., Crea, R.M., Riggs, A.D., Heyneker, H.L., Bolivar, F., and Boyer, H.W. 1977. Expression in *E. coli* of a chemically synthesized gene for the hormone somatostatin. *Science* 198:1053-1056.
- LaVallie, E.R., Rehemtulla, A., Racie, L.A., DiBlasio, E.A., Ferenz, C., Grant, K.L., Light, A., and McCoy, J.M. 1993. Cloning and functional expression of a cDNA encoding the catalytic subunit of bovine enterokinase. *J. Biol. Chem.* 268:23311-23317.
- Lee, N., Cozzikorto, J., Wainwright, N., and Testa, D. 1984. Cloning with tandem gene systems for high level gene expression. *Nucl. Acids Res.* 12:6797-6812.
- Nagai, K. and Thøgersen, H. C. 1984. Generation of β -globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli*. *Nature* 309:810-812.
- Nagai, K. and Thøgersen, H. C. 1987. Synthesis and sequence-specific proteolysis of hybrid proteins produced in *Escherichia coli*. *Methods Enzymol.* 153:461-481.
- Schein, C.H. 1989. Production of soluble recombinant proteins in bacteria. *BioTechnology* 7:1141-1149.
- Smith, D.B. and Johnson, K.S. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67:31-40.
- Stormo, G.D., Schneider, T.D., and Gold, L. 1982. Characterization of translation initiation sites in *E. coli*. *Nucl. Acids Res.* 10:2971-2996.

- Strauch, K.L. and Beckwith, J. 1988. An *Escherichia coli* mutation preventing degradation of abnormal periplasmic proteins. *Proc. Natl. Acad. Sci. U.S.A.* 85:1576-1580.
- Sugimura, K. and Higashi, N. 1988. A novel outer-membrane-associated protease in *Escherichia coli*. *J. Bacteriol.* 170:3650-3654.
- Szoka, P.R., Schreiber, A.B., Chan, H., and Murthy, J. 1986. A general method for retrieving the components of a genetically engineered fusion protein. *DNA* 5:11-20.

Contributed by Paul Riggs
New England Biolabs
Beverly, Massachusetts

Edward R. La Vallie
and John M. McCoy
Genetics Institute
Cambridge, Massachusetts

Enzymatic and Chemical Cleavage of Fusion Proteins

The use of gene fusion expression systems has become an increasingly popular method of producing foreign proteins in *Escherichia coli*. This popularity is due in large part to the development of fusion systems that are capable of producing large amounts of fusion protein in a soluble form. The maltose-binding protein (MBP, UNIT 16.6), glutathione-S-transferase (GST, UNIT 16.7), and thioredoxin (Trx, UNIT 16.8) fusion systems have proven singularly successful in producing properly folded and biologically active proteins. Each of these systems also provides convenient methods for specific purification of the fusion protein from cellular contaminants. As a result, proteins produced using these systems are readily amenable to the study of their biological activities and/or interactions. As a consequence of the popularity of fusion protein expression strategies, the ability to cleave the N-terminal fusion “carrier” protein from the C-terminal protein of interest has become increasingly important.

This unit provides protocols for some commonly used methods of site-specific cleavage of fusion proteins. The first three protocols describe enzymatic cleavage of proteins using proteases that display highly restricted specificities, which greatly decrease the likelihood that unwanted secondary cuts will occur. The first basic protocol describes the use of factor Xa, a mammalian serine protease that cleaves following the sequence Ile-Glu(or Asp)-Gly-Arg↓. This protocol can be applied to fusion proteins produced with either the MBP (UNIT 16.6) or the GST system (pGEX3X vector; UNIT 16.7); both systems utilize expression vectors that encode a factor Xa cleavage site. A support protocol describes conditions for denaturing proteins for factor Xa cleavage if necessary. The next two protocols (first and second alternate protocols) describe cleavage with thrombin, a site-specific protease that recognizes the sequence Leu-Val-Pro-Arg↓Gly-Ser in one of the GST expression vectors (pGEX2T) described in UNIT 16.7. The third enzymatic cleavage protocol (third alternate protocol) uses enterokinase (enteropeptidase), a mammalian intestinal protease that cleaves following the sequence Asp-Asp-Asp-Asp-Lys↓. The Trx fusion vectors pTRXFUS and hpTRXFUS (UNIT 16.8) encode an enterokinase cleavage site immediately prior to their fusion junctions.

Three additional protocols describe cleavage of fusion proteins with chemical reagents as an alternative to enzymatic cleavage. These have some advantages. Though these methods may require modification of the fusion protein so a scissile or labile bond resides at the desired point of cleavage, they may be useful for cleaving fusion proteins with solubility problems or those that are otherwise refractory to enzymatic cleavage. Chemical cleavage methods have the disadvantage of being less specific, and it is necessary to ensure that a susceptible peptide bond does not exist in the protein of interest. The first of the chemical cleavage protocols (second basic protocol) uses cyanogen bromide to cleave after methionine residues. The second method (fourth alternate protocol) uses hydroxylamine to specifically cleave between asparagine and glycine residues. The final chemical cleavage protocol (fifth alternate protocol) cleaves fusion proteins by exploiting the lability of the Asp-Pro bond at low pH.

ENZYMATIC CLEAVAGE OF FUSION PROTEINS WITH FACTOR Xa

Fusion proteins that have been produced with the MBP fusion vectors pMAL-c2, pMAL-p2, or the GST fusion vector pGEX3X contain a recognition sequence for coagulation factor Xa encoded in the DNA immediately preceding the polylinker cloning site. Fusion proteins produced in other systems must be adapted to encode this recognition sequence. It is important to note that factor Xa will not cleave if a proline residue follows the arginine of the recognition sequence. Purification of the fusion protein prior to cleavage is recommended to minimize degradation of the product by nonspecific cellular proteases during incubation with factor Xa protease. Prior purification of the fusion protein also allows subsequent isolation of the cleaved product by simply repeating whatever affinity purification step was performed to purify the fusion protein. This step now removes the fusion partner.

Factor Xa is typically added to the fusion protein substrate at a ratio of 1% to 2% (w/w). However, cleavage efficiency varies depending upon the individual fusion, and ratios ranging from 0.1% to 5% may be effective. Incubation times can be from 1 hr to several days at either room temperature or 4°C. A support protocol describes denaturing and renaturing soluble fusion proteins that do not cleave well under the standard native cleavage conditions. This procedure should be considered a last resort because of the uncertainty of regaining properly folded protein and the inevitable decrease in overall yield.

Materials

For recipes, see Reagents and Solutions in this unit (or cross-referenced unit); for common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

- 1 mg/ml fusion protein
- 200 µg/ml factor Xa (New England Biolabs) in reaction buffer (see step 1)
- 2× SDS sample buffer (UNIT 10.2)
- Boiling water bath
- Additional reagents and equipment for SDS-PAGE (UNIT 10.2)

1. Prepare two small-scale trial reactions to determine optimum incubation time as follows:

Reaction 1: 20 µl of 1 mg/ml fusion protein with 1 µl of 200 µg/ml factor Xa.

Reaction 2: 5 µl of 1 mg/ml fusion protein and no factor Xa (mock digestion).

Incubate at room temperature.

Fusion protein in column buffer from amylose-resin purification (UNIT 16.6; with 1 mM CaCl₂) or in PBS from glutathione-agarose purification (UNIT 16.7) is suitable for factor Xa digestion; otherwise the protein should be prepared in 20 mM Tris·Cl (pH 8.0)/1 mM CaCl₂/100 mM NaCl.

Although most fusion proteins could be kept at 4°C, any remaining fusion protein solution can be stored at -70°C, in 10% glycerol, until used in step 6.

2. At 2, 4, 8, and 24 hr, remove 5-µl aliquots of the factor Xa reaction, add 5 µl of 2× SDS sample buffer, and freeze at -20°C.
3. At 24 hr mix 5 µl mock digestion with 5 µl of 2× SDS sample buffer.
4. Mix 5 µl of original fusion protein solution with 5 µl of 2× SDS sample buffer (uncut control).

5. Heat all samples 10 min in a boiling water bath and load onto an SDS-polyacrylamide gel. Evaluate extent of cleavage to determine correct incubation time.

Gel composition and running conditions will be determined by the size of the fusion protein.

If only partial cleavage is evident, increase amount of enzyme and/or incubation time. If no cleavage is apparent, proceed to the next support protocol.

6. Once satisfactory cleavage conditions have been determined, scale up the trial reaction for the remainder of the fusion protein sample, saving a small amount of uncleaved fusion protein for comparison purposes. Monitor the extent of cleavage by SDS-PAGE.

The cleavage products can be separated using any of the support protocols in UNIT 16.6.

DENATURING A FUSION PROTEIN FOR FACTOR Xa CLEAVAGE

It has been observed that some fusion proteins are resistant to cleavage with factor Xa. This problem can sometimes be alleviated by denaturing the fusion protein, renaturing it, and then incubating it with protease. The following protocol has been adapted from New England Biolabs' recommendations for MBP fusion proteins, and should be applicable to any fusion protein that contains an inaccessible factor Xa cleavage site. Denaturation is accomplished by incubating the fusion protein in 6 M guanidine-HCl followed by dialysis against the reaction buffer.

Additional Materials

*For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.*

20 mM Tris·Cl (pH 7.4)/6 M guanidine-HCl

20 mM Tris·Cl (pH 8.0)/1 mM CaCl₂

1. Dialyze fusion protein for ≤4 hr against ≥10 vol of 20 mM Tris·Cl (pH 7.4)/6 M guanidine-HCl, or add guanidine-HCl to the fusion protein to give a final concentration of 6 M.
2. Dialyze the sample for 4 hr against 100 vol of 20 mM Tris·Cl (pH 8.0)/1 mM CaCl₂.
3. Repeat the second dialysis for an additional 4 hr against 100 vol fresh buffer.

This denaturation procedure is intended to allow better accessibility of the cleavage site to factor Xa before the protein can completely reassume its former protease-resistant conformation. Therefore, it is best to proceed with the cleavage reaction immediately following dialysis. However, rapid removal of denaturant sometimes results in precipitation of the protein; in these cases, gradual removal of denaturant by stepwise dialysis against 2-fold dilutions of the guanidine-HCl solution may keep the protein from precipitating. Alternatively, the fusion protein remaining in solution after rapid dialysis precipitation can be recovered and cleaved, and the insoluble material discarded.

4. Proceed with step 1 of the first basic protocol for factor Xa cleavage.

SUPPORT PROTOCOL

ENZYMATIC CLEAVAGE OF FUSION PROTEINS WITH THROMBIN

Thrombin is a mammalian serine protease that cleaves in a trypsin-like manner; that is, it cleaves after arginine and lysine residues. However, thrombin displays distinct subsite preferences, with optimum cleavage occurring at sites containing P4-P3-Pro-Arg↓P1'-P2' (where P4 and P3 are hydrophobic amino acids and P1' and P2' are nonacidic amino acids). The GST fusion system (UNIT 16.7) utilizes a vector that encodes a cleavage site with this restricted specificity (Leu-Val-Pro-Arg↓Gly-Ser). GST fusion proteins expressed with the pGEX2T vector (Fig. 16.7.1) can be cleaved with thrombin either after affinity purification on glutathione-agarose, or alternatively, while still bound to the affinity matrix. The following alternate protocols describe both methods for cleaving the fusion protein: first, cleavage of fusion proteins with thrombin in solution, an approach that is applicable to any fusion protein containing a thrombin recognition sequence; and second, thrombin cleavage of GST fusion proteins bound to glutathione-agarose. The latter technique is preferred for GST fusion proteins because it is faster and usually requires less work. However, the approach may not always work: some proteins may become insoluble when separated from the GST carrier, thus complicating their physical separation from the affinity matrix. In this case, the first alternate protocol should be used.

Additional Materials

For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

Thrombin cleavage buffer (see recipe)
Heparin, sodium salt (with ≥ 140 U/mg activity, Sigma; optional)
Thrombin (human, with ~ 3000 U/mg activity; Sigma or Boehringer Mannheim)

1. Prepare two pilot cleavage reactions to determine optimal reaction conditions as follows:

Reaction 1: 20 μ l of 1 mg/ml fusion protein solution (in appropriate buffer) and 0.2 μ g thrombin.

Reaction 2: 5 μ l of 1 mg/ml fusion protein solution only (mock digestion).

Incubate at 25°C.

GST fusion protein that has been eluted from glutathione-agarose in 50 mM Tris·Cl (pH 7.5)/5 mM reduced glutathione can be used after addition of NaCl to 150 mM and CaCl₂ to 2.5 mM and adjustment of the protein concentration to 1 mg/ml. Other fusion proteins can be resuspended or dialyzed in thrombin cleavage buffer (without glutathione) for subsequent cleavage.

Addition of 10 μ M heparin to the cleavage reaction is optional. It has been reported (Chang, 1985) that this increases the rate of some cleavages by 10% to 50%, apparently due to a direct interaction with the enzyme.

2. At 30 min, 1, 2, and 4 hr, remove 5 μ l from the thrombin reaction and mix with 5 μ l of 2 \times SDS sample buffer. Freeze at -20°C .
3. At the 4 hr time point, add 5 μ l of 2 \times SDS sample buffer to the mock digestion.
4. Mix 5 μ l of original fusion protein solution with 5 μ l of 2 \times SDS sample buffer (untreated control).
5. Boil all samples 10 min and load on an SDS-polyacrylamide gel to analyze sample stability and efficiency of cleavage.
6. Use those conditions determined empirically to be best for cleaving the fusion protein to scale up the cleavage reaction for the desired quantity of protein.

In addition to varying the time of incubation, the amount of thrombin and the temperature of incubation (up to 37°C) may also be varied to determine the optimum conditions for cleavage of a particular fusion protein.

ENZYMATIC CLEAVAGE OF MATRIX-BOUND GST FUSION PROTEINS

ALTERNATE PROTOCOL 2

In this alternate protocol, GST fusion proteins that contain a thrombin cleavage site are bound to glutathione-agarose as described in *UNIT 16.7*. Prior to elution from the matrix, thrombin is added and the protein of interest is cleaved from the GST carrier. The cleaved protein is collected in the wash buffer and the GST carrier remains bound to the beads, permitting easy and efficient physical separation of the reaction products.

Additional Materials

*For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.*

GST fusion protein bound to glutathione-agarose beads (*UNIT 16.7*)
1% (v/v) Triton X-100 in phosphate-buffered saline (PBS; *APPENDIX 2*)
GST wash buffer: 50 mM Tris·Cl (pH 7.5)/150 mM NaCl
GST elution buffer: 50 mM Tris·Cl (pH 8.0)/5 mM reduced glutathione
20- or 50- ml screw-cap tube

1. Wash GST fusion protein bound to glutathione-agarose beads with 20 vol of 1% Triton X-100 in PBS, using a 20- or 50-ml screw-cap tube. Centrifuge 10 sec in a tabletop centrifuge at $500 \times g$, room temperature, to pellet the beads. Carefully remove and discard the supernatant. Resuspend the beads in 20 vol Triton X-100 buffer and repeat wash.
2. After the second centrifugation, carefully remove and discard the supernatant. Resuspend the beads in 20 vol GST wash buffer.
3. Pellet the beads and discard the supernatant. Resuspend the beads in 20 vol thrombin cleavage buffer. Repeat the centrifugation and resuspend the beads in ≤ 1 ml thrombin cleavage buffer.

Although it is easier to wash the beads in large volumes, the amount of thrombin cleavage buffer to use in the cleavage reaction is best kept to a minimum.

4. Remove a small aliquot of resuspended beads and add an equal volume of $2\times$ SDS sample buffer. Store at -20°C until analyzed by SDS-PAGE (step 7).

This sample is used to estimate the amount of fusion protein bound to the beads.

5. Add thrombin to the remaining bead slurry at a ratio of 1% (w/w) thrombin to the estimated amount of bound fusion protein. Incubate 1 hr at 25°C .

As in solution cleavage, the amount of thrombin, time, and temperature of incubation can be adjusted to optimize the cleavage efficiency.

6. Elute the cleaved and released protein by washing the beads with 1 bed volume of GST wash buffer. Centrifuge as in step 1 to pellet beads and collect supernatant. Repeat five times, but keep each wash fraction separate. Remove 20- μl aliquots from each wash fraction for SDS-PAGE.
7. Elute bound GST by repeating step 6 with GST elution buffer instead of GST wash buffer. Remove 20- μl aliquots from each fraction and analyze by SDS-PAGE to determine extent of cleavage. Include the aliquot of beads from step 4 on this gel.

If cleavage is incomplete, the time of incubation and/or the amount of enzyme can be increased.

Protein Expression

16.4.9

Because this is an analytical-scale experiment, it is easiest to discard these beads after one use. Upon scaling up the procedure, the beads can be regenerated, as described in *Reagents and Solutions*, UNIT 16.7.

ALTERNATE PROTOCOL 3

ENZYMATIC CLEAVAGE OF FUSION PROTEINS WITH ENTEROKINASE

Enterokinase (also called enteropeptidase) is a mammalian trypsin-like serine protease that displays a high degree of specificity for the sequence (Asp)₄-Lys, cleaving on the carboxy-terminal side of the lysine residue of the recognition sequence. Although in mammals the enzyme has evolved to recognize and cleave this sequence from the amino-termini of trypsinogens, it has been shown that enterokinase is also capable of cleaving fusion proteins that are expressed in bacteria and that contain this recognition sequence inserted between the carrier protein and the carboxy-terminal fusion partner. Enterokinase is capable of cleaving fusion proteins under a wide range of reaction conditions, with pH ranging from 4.5 to 9.5 and temperatures ranging from 4° to 45°C. Enterokinase is also extremely tolerant of the nature of the amino acid residue in the P1' position (except that the peptide bond between Lys-Pro at this position is totally refractory to cleavage; E. LaVallie and L. Racie, unpub. observ.). At sufficiently low ionic strength, enterokinase can cleave fusion proteins at a weight ratio of 1:500 to 1:2000. At these ratios, typical cleavage reactions are carried out for 16 to 24 hr at 37°C, but these parameters (time, temperature, and enzyme/substrate ratio) can be adjusted as needed.

The thioredoxin fusion vector pTRXFUS (UNIT 16.8) encodes an enterokinase cleavage site immediately preceding the polylinker cloning region. Proteins produced as Trx fusions using this system can be subsequently released by incubation with enterokinase, leaving their authentic amino-terminal sequence. The protocol below describes the use of bovine enterokinase in this application.

Additional Materials

For recipes, see *Reagents and Solutions* in this unit (or cross-referenced unit); for common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

- 1 mg/ml thioredoxin fusion protein (UNIT 16.8) in 50 mM Tris·Cl (pH 8.0)/1 mM CaCl₂
- 10 µg/ml bovine enterokinase (Biozyme EK-3 grade) in 50 mM Tris·Cl (pH 8.0)/1 mM CaCl₂

NOTE: Many commercial preparations of enterokinase (bovine or porcine), with the exception of the source listed, are extremely impure and tend to be contaminated with, among other things, trypsin and chymotrypsin which can extensively degrade the fusion protein. It is recommended that only commercial enterokinase of the highest quality be used.

1. Perform a pilot experiment to monitor the efficiency of cleavage with various ratios of enterokinase to fusion protein. Prepare five reactions:

Reactions 1 to 4: 20 µl of 1 mg/ml fusion protein, 1 µl, 2 µl, 5 µl, and 10 µl of 10 µg/ml bovine enterokinase, and 50 mM Tris·Cl (pH 8.0)/1 mM CaCl₂, to a total of 30 µl.

Reaction 5: 20 µl of 1 mg/ml fusion protein and 10 µl of 50 mM Tris·Cl (pH 8.0)/1 mM CaCl₂ (mock digestion).

Incubate samples ≥ 16 hr at 37°C.

The fusion protein must be (at least) partially purified prior to digestion with enterokinase because the enzyme is inactive in crude bacterial lysates.

2. Stop the reaction by adding 30 μ l of 2 \times SDS sample buffer to each reaction. Boil 10 min.

For larger-scale applications, the reaction can be stopped by adding p-aminobenzamidinium (PABA) to 5 mM. PABA is a competitive inhibitor of most intestinal serine proteases. It should provide protection from nonspecific proteolysis of the reaction products by contaminants in the enzyme preparation before the protein of interest is purified further.

3. Load 10 μ l of each sample onto an SDS-polyacrylamide gel to analyze the extent of cleavage. Adjust enterokinase concentration and length of incubation accordingly to accomplish complete digestion.
4. Scale up the reaction components linearly to digest a larger amount of fusion protein.

Calcium ions marginally increase the efficiency of cleavage, but their presence sometimes promotes fusion protein degradation by stimulating contaminating proteolytic activities. If degradation of the cleaved fusion protein occurs, omit calcium and add 5 mM EDTA to the cleavage reaction to try to eliminate the problem.

CHEMICAL CLEAVAGE OF FUSION PROTEINS USING CYANOGEN BROMIDE

**BASIC
PROTOCOL 2**

Cyanogen bromide (CNBr) has been used to cleave proteins at methionine residues for many years. CNBr has been used industrially for the production of both somatostatin (Itakura et al., 1977) and insulin (Chance et al., 1981). The reaction is typically carried out at low pH in 70% formic acid, and cleavage occurs at the C-terminal side of methionine residues. The protein concentration is relatively unimportant, as the CNBr is in vast excess for hydrolysis. The technique is useful only if the protein of interest lacks methionine residues. Cleavage with CNBr is usually efficient, but side chain modifications and nonspecific cleavages are common upon prolonged incubation at low pH. These problems, along with the potential for reduction of intramolecular disulfide bonds during treatment with 70% formic acid, can be minimized by replacing the formic acid with 6 M guanidine·HCl/0.2 M HCl.

CAUTION: Cyanogen bromide is extremely toxic. It should only be used in a properly ventilated fume hood. Exercise appropriate caution in its use and disposal.

Materials

*For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.*

- 1 mg/ml fusion protein
- 50 mg/ml cyanogen bromide (CNBr)/70% (v/v) formic acid
- 70% (v/v) formic acid
- 1 \times SDS sample buffer (*UNIT 10.2*)

Additional reagents and equipment for SDS-PAGE (*UNIT 10.2*)

1. Perform a pilot experiment to determine minimum incubation time. Lyophilize two 50- μ l aliquots of fusion protein solution. Resuspend one aliquot in 50 μ l of 50 mg/ml CNBr/70% formic acid. Resuspend the other in 50 μ l of 70% formic acid *without* CNBr. Incubate at room temperature.
2. At 0, 8, 24, and 48 hr, remove a 5- μ l aliquot and lyophilize.

**Protein
Expression**

16.4.11

3. Resuspend all aliquots in 20 μ l of 1 \times SDS sample buffer, boil 10 min, and load onto an SDS-polyacrylamide gel.
4. Based on analysis of the gel, determine the minimum incubation time necessary to completely cleave the protein.

The protocol can be easily scaled up to accommodate larger amounts of fusion protein. Some proteins are resistant to cleavage with cyanogen bromide. In such cases, or when the fusion protein to be cleaved is insoluble, guanidine-HCl can be added to the reaction at a final concentration of 6 M.

ALTERNATE PROTOCOL 4

CHEMICAL CLEAVAGE OF FUSION PROTEINS USING HYDROXYLAMINE

Hydroxylamine cleaves proteins at Asn-Gly bonds and can be used as a reagent for chemical cleavage of fusion proteins. This cleavage site is less common than that for cyanogen bromide, and therefore the presence of a susceptible bond in the protein of interest is less likely. One disadvantage is that the released carboxy-terminal fusion partner will retain a glycine residue at its amino terminus, which is unacceptable in some applications. Also, the reaction requires incubation of the fusion protein at alkaline pH, which may cause modification of some amino acid side chains. Finally, protein digestions by this technique are usually incomplete due to the nature of the cleavage mechanism (E.L., unpub. observ.), reducing yield and possibly complicating post-cleavage purification of the desired protein product. However, the technique does have advantages: speed, economy, and the ability to perform digestions under denaturing conditions (e.g., 6 M guanidine-HCl) for otherwise insoluble fusion proteins.

CAUTION: Hydroxylamine is potentially explosive if mishandled. Be sure to follow all precautions indicated by the manufacturer.

Additional Materials

*For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.*

- 1 mg/ml fusion protein in 10 mM Tris-Cl (pH 8.0)/150 mM NaCl
- 2 \times hydroxylamine cleavage solution (see recipe)
- Guanidine-HCl (optional)
- 2 \times SDS sample buffer (*UNIT 10.2*)
- Boiling water bath

1. Perform a pilot experiment to determine minimum incubation time. Mix 50 μ l of 1 mg/ml fusion protein in 10 mM Tris-Cl (pH 8.0)/150 mM NaCl with 50 μ l of 2 \times hydroxylamine cleavage solution in a 1.5-ml microcentrifuge tube. Incubate at 45°C.

If the fusion protein is insoluble, guanidine-HCl can be added to the cleavage reaction at a final concentration of 6 M. This may also help in cases where a particular Asn-Gly bond appears to be resistant to cleavage.

2. At 0, 2, 4, 8, 16, and 24 hr, remove 10- μ l aliquots from the cleavage reaction and mix with 10 μ l of 2 \times SDS sample buffer. Freeze each tube on dry ice until all time points have been collected.
4. Heat samples 10 min in a boiling water bath. Load all samples onto an SDS-polyacrylamide gel to analyze the extent of cleavage.
5. Determine the minimum incubation time necessary for maximum cleavage.

If cleavage after 24 hr is still poor, add guanidine-HCl to 6 M final, increase hydroxylamine concentration to 3 M final, or both.

CHEMICAL CLEAVAGE OF FUSION PROTEINS BY HYDROLYSIS AT LOW pH

This method exploits the fact that the Asp-Pro bond is labile at low pH. Hydrolysis of this peptide bond occurs at elevated temperatures (37° to 40°C) under acidic conditions (pH 2.5). Nonspecific cleavages can occur upon prolonged incubation under these conditions, and it is necessary to determine empirically the minimum length of time necessary for cleavage. Like the other chemical cleavage methods described in this unit, the reaction conditions are somewhat harsh and may result in denaturation or modification of the protein. On the other hand, this treatment allows insoluble fusion proteins to be cleaved by acid hydrolysis of Asp-Pro bonds in the presence of 6 M guanidine-HCl. To use this procedure the amino acid sequence of the carboxy-terminal fusion partner should first be examined carefully to verify the absence of other Asp-Pro bonds. The released protein will retain a proline residue at its amino-terminus.

This method will potentially cleave any protein containing an Asp-Pro bond. The GST fusion vector pGEX1 (UNIT 16.7, and Fig. 16.7.1 therein) contains an Asp-Pro cleavage site encoded by the *Bam*HI site of the polylinker cloning region, so fusions at this site will result in fusion proteins that can be released from GST using this protocol.

Additional Materials

For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

Fusion protein containing an Asp-Pro bond between the component domains
70% (v/v) formic acid
13% (v/v) acetic acid
0.1 M Tris base
Guanidine-HCl

1. Perform a pilot experiment to determine optimal hydrolysis conditions. Prepare four reaction mixtures:

Reaction 1: ~20 µg fusion protein in 70% formic acid

Reaction 2: ~20 µg fusion protein in 70% formic acid/6 M guanidine-HCl

Reaction 3: ~20 µg fusion protein in 13% acetic acid

Reaction 4: ~20 µg fusion protein in 13% acetic acid/6 M guanidine-HCl.

Incubate all samples at 37°C.

2. At 0, 24, 48, and 72 hr, remove a 5-µg aliquot of each reaction mixture and lyophilize to dryness.

The 0 time point can serve as the negative control.

3. Resuspend the hydrolyzed protein in 20 µl of 1× SDS sample buffer and neutralize by gradual addition of 0.1 M Tris base until the sample turns from yellow to blue. Analyze samples on a tricine SDS-polyacrylamide gel for extent of digestion.
4. Choose the mildest condition and shortest incubation time that give the desired extent of cleavage. Scale up to larger amounts of fusion protein accordingly.

There is a great deal of variation in the susceptibility of Asp-Pro bonds to cleavage. Some Asp-Pro bonds cleave readily under mild conditions, whereas others are resistant to cleavage and require incubation in stronger acid conditions and/or strong denaturants to attain hydrolysis. Even under strong conditions, some Asp-Pro bonds remain uncleaved and others may not be cleaved to completion, i.e., they may be cleaved in only a fraction of the proteins. However, to avoid unwanted denaturation or modification of the protein of interest, it is important to determine the mildest conditions that give the desired degree of cleavage.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Hydroxylamine cleavage solution, 2×

4 M hydroxylamine
0.4 M CHES buffer
Adjust pH to 9.5 with NaOH
Prepare fresh

Thrombin cleavage buffer

50 mM Tris-Cl, pH 7.5
150 mM NaCl
2.5 mM CaCl₂
Store indefinitely at -20°C

COMMENTARY

Background Information

The use of fusion proteins in *Escherichia coli* for production of proteins from other organisms is becoming increasingly popular. The principal advantages of fusion proteins include high expression levels, ease of purification, and ease of detection with biochemical or immunological reagents. However, early protein fusion methods fused the target protein to the carboxy-terminus of *E. coli* β -galactosidase; such proteins are usually insoluble, have restricted usefulness (except as antigens for antibody production), and require refolding of the protein in an attempt to retain biological activity. The development of improved fusion protein expression systems that are often capable of producing properly folded and biologically active proteins in *E. coli*, such as the MBP system (UNIT 16.6), the GST system (UNIT 16.7), and the Trx system (UNIT 16.8), has increased the need for suitable methods of separating the protein of interest from the so-called carrier protein. Such separation is desirable only after the carrier domain has been exploited for its particular attributes—e.g., to facilitate specific purification or detection, or to enhance stability or solubility. Subsequent site-specific cleavage of the carrier domain from the correctly folded protein of interest then allows evaluation of the biological activity of the protein without potential interference from a covalently-attached fusion partner.

Over the years, many different enzymatic and chemical methods for site-specific cleavage of polypeptides have been developed (Gross, 1967; Spande et al., 1970; Maroux et al., 1971; Landon, 1977; Nagai and Thøgersen, 1984; Chang, 1985; Szoka et al., 1986; Smith and Johnson, 1988; Gearing, et al., 1989; Ham-

mond et al., 1991). The choice of cleavage reagents for those expression vectors obtained from molecular biology suppliers that are used “off the shelf” will usually be dictated by the recognition sequences that lie at their respective fusion junctions. For example, MBP vectors (UNIT 16.6; Maina et al., 1988) are designed for factor Xa cleavage, GST vectors (UNIT 16.7) offer a choice of factor Xa, thrombin, or acid cleavage, and the Trx vector (UNIT 16.8; LaVallie et al., 1993a) utilizes enterokinase for cleavage. However, it is important to remember that these vectors, or any other fusion protein expression vector, can be manipulated by the user to include junction amino acids that will allow enzymatic or chemical hydrolysis with other cleavage reagents.

Critical Parameters and Troubleshooting

The cleavage methods described in this unit have been chosen based upon their specificity, efficiency, and reagent availability. Ultimately, the choice of cleavage reagent will depend upon many factors. First of all, the primary sequence of the protein of interest must be scrutinized to identify sequences that will be susceptible to the cleavage reagents in question. Because of the highly restricted specificity in the case of the proteases described in the preceding protocols, the occurrence of such additional sites is unlikely. However, the use of chemical cleavage methods or less specific proteases such as trypsin (Lys↓ or Arg↓) requires careful consideration of the composition of the protein of interest to avoid unwanted fragmentation of the product. Secondly, the physical characteristics of the fusion protein are an important consideration in choosing an

appropriate cleavage method. Fusion proteins that are insoluble generally require a chemical cleavage method that allows incubation in the presence of protein denaturants such as guanidine-HCl or urea followed by proper refolding. Conversely, soluble fusion proteins should be cleaved under conditions where denaturation is minimized to preserve their structure and/or biological activity. In addition, some cleavage methods require incubation at pH extremes, which may result in modifications to some of the side chains or aggregation and precipitation of the protein and may ultimately be deleterious to the usefulness of the cleaved product.

When the fusion protein is produced in a soluble fashion, enzymatic cleavage protocols are preferred for many reasons. These reactions are carried out under mild conditions of neutral pH, low ionic strength, and moderate temperature (25° to 37°C). Such conditions approximate physiological environments and typically should be least harmful to the integrity of the protein. The high degree of specificity exhibited by these proteases ensures a low probability of unwanted cleavage elsewhere in the protein of interest. Finally, the extent of enzymatic cleavage typically can be altered by modulating parameters such as amount of enzyme, substrate concentration, and length of incubation.

The use of proteases sometimes causes problems, however. The most common problem is unwanted secondary proteolysis or degradation of the fusion protein by contaminating proteolytic activities in the cleavage reaction. Degradation may be caused by *E. coli* proteases that have not been purified away from the fusion protein prior to cleavage. Such contaminants can often be alleviated by additional purification prior to enzymatic digestion. Alternatively, the unwanted proteolytic activity may result from a contaminant in the enzyme preparation itself. These enzymes are serine proteases that have been purified from natural sources, so it is probable that they are contaminated with other proteases that copurify in trace amounts. For example, even highly purified enterokinase from bovine intestine contains trace amounts of tryptic and chymotryptic activity that can cause minor secondary proteolysis of thioredoxin fusion proteins (LaVallie et al., 1993a). This degradation can be reduced by omitting Ca⁺⁺ ions from the digestion and using EDTA to chelate residual Ca⁺⁺. The best solution to this problem, however, is to use a recombinant source of enzyme produced in cell cul-

ture; such enzymes are free of contaminating proteases found in intestinal preparations (LaVallie et al., 1993b and unpub. observ.).

Another potential problem that has been observed with some enzymatic cleavages is cleavage at sites other than the anticipated peptide bonds. This has been reported for factor Xa (Nagai and Thøgersen, 1987; Lauritzen et al., 1991) and thrombin (Chang, 1985); enterokinase has been observed to cleave at subsites that resemble the (Asp)₄-Lys recognition sequence when the substrate is denatured or otherwise improperly folded (Light et al., 1980, E. LaVallie and L. Racie, unpub. observ.). Almost always, this relaxed site specificity can be minimized by decreasing the enzyme/substrate ratio and/or the time of incubation. There are also more exotic strategies; for example, reversible acylation of the fusion protein has been used to eliminate nonspecific cleavage by factor Xa (Wearne, 1990).

Although it is less desirable than enzymatic digestion in most circumstances, chemical cleavage of fusion proteins is sometimes necessary and may be advantageous in certain applications. This is in spite of the significant disadvantages. First, chemical cleavage procedures almost universally employ harsh conditions, such as pH extremes or high temperatures, that can denature the fusion protein and/or modify amino acid side chains (e.g., deamidation of Asn residues and oxidation of Met residues). Second, cleavage specificity tends to be limited to single amino acids or, at most, dipeptide sequences, greatly decreasing the utility of chemical reagents in site-specific cleavage of large polypeptide substrates. Even this low degree of specificity is not absolute, and low levels of side reactions at alternate sites have been noted for hydroxylamine (Bornstein and Balian, 1970; Steinman et al., 1974) and cyanogen bromide (Langley and Smith, 1971). Even acidic cleavage of Asp-Pro bonds is sometimes accompanied by nonspecific peptide bond hydrolysis when incubation is prolonged (Landon, 1977). Third, using chemical cleavage reagents involves the danger of working with hazardous compounds such as cyanogen bromide and hydroxylamine. Extreme care must be used in the storage, use, and disposal of these reagents. Material Safety Data Sheets (MSDSs) should be obtained from the manufacturer for any of these compounds and should be read carefully.

In spite of these shortcomings, however, specific chemical hydrolysis of junction peptide bonds in fusion proteins is preferred for

some applications. For instance, fusion proteins that are produced in an insoluble form cannot be cleaved enzymatically unless they have been solubilized and refolded, but site-specific cleavage of the protein domains often can be accomplished using chemical hydrolysis in the presence of strong chaotropic agents such as guanidine-HCl (Landon, 1977; Szoka et al., 1986; Villa et al., 1989). Chemical cleavages are also useful when a soluble fusion protein is refractory to cleavage under nondenaturing conditions using either enzymatic or chemical cleavage methods. Other advantages of chemical cleavage reagents are economy, purity, and wide availability.

Anticipated Results

Enzymatic digestion of fusion proteins is sensitive to many different parameters such as temperature, pH, ionic strength, buffer composition, substrate concentration, enzyme concentration, and length of incubation. Optimum parameters must be determined empirically for each fusion protein. However, if the fusion protein is of adequate purity and the cleavage site is accessible, in most cases the standard conditions described in the protocols will give satisfactory cleavage.

Chemical cleavages are much more forgiving of small variations in reaction conditions, and cleavage should be attainable using the reaction conditions given. However, the extent of protein modifications caused by these general reaction conditions may often be significant to the ultimate usefulness of the cleaved protein. Reaction parameters may be adjusted as described to minimize any modifications or secondary cleavages that may occur.

With any of these methods, complete digestion of the fusion protein is often difficult to attain, and cleavage efficiencies of 70% to 80% should be considered satisfactory.

Time Considerations

Fusion protein cleavage can be accomplished in hours or days, depending upon the reagents used and the reaction conditions. For factor Xa, thrombin, and enterokinase, the enzyme-to-substrate ratio and substrate concentration are often chosen so the digestions are usually complete in 8 to 24 hr. By contrast, chemical cleavages are often highly variable in the amount of time necessary for total cleavage. Overnight digestion under standard conditions is common for cyanogen bromide and acid cleavages, whereas hydroxylamine cleavages are often complete in 2 to 4 hr.

Literature Cited

- Bornstein, P. and Balian, G. 1970. The specific nonenzymatic cleavage of bovine ribonuclease with hydroxylamine. *J. Biol. Chem.* 245:4854-4856.
- Bornstein, P. and Balian, G. 1977. Cleavage at Asn-Gly bonds with hydroxylamine. *Methods Enzymol.* 47:132-145.
- Chance, R.E., Hoffman, J.A., Kroeff, E.P., Johnson, M.G., Schirmer, E.W., Bromer, W.W., Ross, M.J., and Wetzel, R. 1981. The production of human insulin using recombinant DNA technology and a new chain combination procedure. *In* Peptides: Synthesis-Structure-Function (D.H. Rich and E. Gross, eds.), pp. 721-728. Pierce Chemical, Rockford, Ill.
- Chang, J.-Y. 1985. Thrombin specificity. *Eur. J. Biochem.* 151:217-224.
- Gearing, D.P., Nicola, N.A., Metcalf, D., Foote, S., Willson, T.A., Gough, N.M., and Williams, R.L. 1989. Production of leukemia factor in *Escherichia coli* by a novel procedure and its use in maintaining embryonic stem cells in culture. *BioTechnology* 7:1157-1161.
- Gross, E. 1967. The cyanogen bromide reaction. *Methods Enzymol.* 11:238-255.
- Hammond, P.M., Atkinson, T., Sherwood, R.F., and Scawen, M.D. 1991. Manufacturing new-generation proteins. Part I: The technology. *Biopharm.* 4:16-23.
- Itakura, K., Hirose, T., Crea, R., Riggs, A.D., Heyneker, H.L., Bolivar, F., and Boyer, H.W. 1977. Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science* 198:1056-1063.
- Langley, T.J. and Smith, E.L. 1971. Sequence of bovine liver glutamate dehydrogenase. IV. Cyanogen bromide peptides. *J. Biol. Chem.* 246:3789-3801.
- Landon, M. 1977. Cleavage at aspartyl-prolyl bonds. *Methods Enzymol.* 47(E):145-149.
- Lauritzen, C., Tüchsen, E., Hansen, P.E., and Skovgaard, O. 1991. BPTI and N-terminal extended analogues generated by Factor Xa cleavage and cathepsin C trimming of a fusion protein expressed in *Escherichia coli*. *Prot. Expr. Purif.* 2:372-378.
- LaVallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., and McCoy, J.M. 1993a. A thioredoxin gene fusion system that circumvents inclusion body formation in the *E. coli* cytoplasm. *BioTechnology* 11:187-193.
- LaVallie, E.R., Rehemtulla, A., Racie, L.A., DiBlasio, E.A., Ferenz, C., Grant, K.L., Light, A., and McCoy, J.M. 1993b. Cloning and functional expression of a cDNA encoding the catalytic subunit of bovine enterokinase. *J. Biol. Chem.* 268:23311-23317.
- Light, A., Savithri, H. S., and Liepnieks, J.J. 1980. Specificity of bovine enterokinase toward protein substrates. *Anal. Biochem.* 106:199-206.

- Maina, C.V., Riggs, P.D., Granda, A.G., Slatko, B.E., Moran, L.S., Tagliamonte, J.A., McReynolds, L.A., and Guan, C. 1988. An *Escherichia coli* vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein. *Gene* 74:365-373.
- Maroux, S., Baratti, J., and Desnuelle, P. 1971. Purification and specificity of porcine enterokinase. *J. Biol. Chem.* 246:5031-5039.
- Nagai, K. and Thøgersen, H. C. 1984. Generation of β -globin by sequence-specific proteolysis of a hybrid protein in *Escherichia coli*. *Nature* 309:810-812.
- Nagai, K. and Thøgersen, H. C. 1987. Synthesis and sequence-specific proteolysis of hybrid proteins produced in *Escherichia coli*. *Methods Enzymol.* 153:461-481.
- Smith, D.B. and Johnson, K.S. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67:31-40.
- Spande, T.F., Witkop, B., Degani, Y., and Patchornik, A. 1970. Selective cleavage and modification of peptides and proteins. *Adv. Prot. Chem.* 24:97-260.
- Steinman, H.M., Naik, V.R., Abernethy, J.L., and Hill, R.L. 1974. Bovine erythrocyte superoxide dismutase: Complete amino acid sequence. *J. Biol. Chem.* 249:7326-7338.
- Szoka, P.R., Schreiber, A.B., Chan, H., and Murthy, J. 1986. A general method for retrieving components of a genetically engineered fusion protein. *DNA* 5:11-20.
- Villa, S., DeFazio, G., and Canosi, U. 1989. Cyanogen bromide cleavage at methionine residues of polypeptides containing disulfide bonds. *Anal. Biochem.* 177:161-164.
- Wearne, S.J. 1990. Factor Xa cleavage of fusion proteins: Elimination of nonspecific cleavage by reversible acylation. *FEBS Lett.* 263:23-26.

Contributed by Edward R. LaVallie
and John M. McCoy
Genetics Institute
Cambridge, Massachusetts

Donald B. Smith (thrombin cleavage)
University of Edinburgh
Edinburgh, Scotland

Paul Riggs (denaturing fusion proteins)
New England Biolabs
Beverly, Massachusetts

Expression and Purification of *lacZ* and *trpE* Fusion Proteins

BASIC PROTOCOL

Fusion proteins are commonly used as a source of antigen for producing antibodies and in many cases can be useful for biochemical analyses. This unit describes how to express fusion proteins and prepare extracts for both applications.

Two widely used expression systems for producing large amounts of proteins in *E. coli* are presented. One system expresses *lacZ* fusions using the pUR series of vectors (UNIT 1.5; Rütther and Müller-Hill, 1983) and the other expresses *trpE* fusions using the pATH vectors (Koerner et al., 1990). The gene of interest is first subcloned into either a pUR or pATH vector in the correct reading frame. The correct transformant is selected, grown, and then induced with either IPTG or IAA. The method for preparing extracts—i.e., sonication of cells in the presence of protease inhibitors—is suitable for both types of fusion proteins, as well as for other types of proteins overexpressed in *E. coli*. The extracts are checked for the presence of fusion protein on an SDS-polyacrylamide gel.

Materials

- pUR (UNIT 1.5) or pATH (GenBank file name M32985) vectors
- E. coli* C600, HB101, RR1 or equivalent (Table 1.4.5)
- LB plates and medium containing 50 µg/ml ampicillin (UNIT 1.1)
- 100 mM IPTG (store at –20°C)
- M9 plates and medium containing 50 µg/ml ampicillin, 0.5% Casamino acids, 10 µg/ml thiamine, and with/without 20 µg/ml tryptophan (supplemented M9; UNIT 1.1)
- 2.5 mg/ml indoleacrylic acid (IAA) in 95% ethanol (store at –20°C)
- Phosphate-buffered saline (PBS; APPENDIX 2), ice-cold
- HEMGN buffer, ice-cold
- 50 mg/ml lysozyme in 0.25 M Tris·Cl, pH 8.0 (store at –20°C)
- HEMGN buffer/8 M guanidine·HCl (prepare 100 ml and store at 4°C)
- HEMGN buffer/1 M guanidine·HCl (prepare 500 ml and store at 4°C)
- Sorvall RC-5B centrifuge with GSA rotor (or equivalent) and 200-ml bottles
- Sorvall Omnispin clinical centrifuge (or equivalent) and 15-ml conical tubes
- Sonicator with a microtip
- Sorvall SS-34 rotor (or equivalent) and 50-ml tubes
- Ultracentrifuge with Beckman 60Ti rotor (or equivalent) and tubes
- Dialysis tubing, MWCO 12,000 to 14,000 (APPENDIX 3)
- Additional reagents and equipment for subcloning of DNA fragments (UNITS 1.4 & 3.16), transformation of competent *E. coli* cells (UNIT 1.8), quantitation of proteins by the Bradford method (UNIT 10.1), and SDS-PAGE (UNIT 10.2)

Subclone, grow, and induce cells

To express *lacZ* fusion proteins using pUR vectors:

- 1a. Subclone gene of interest into a pUR vector in the correct reading frame, transform competent *E. coli* cells, and select transformants on LB/ampicillin plates.

E. coli C600, HB101, and RR1 (Bolivar et al., 1977) have been used successfully as hosts for pUR expression vectors.

- 2a. Inoculate 2 to 5 ml LB/ampicillin medium with a single colony containing the expression vector. Grow overnight at 37°C with shaking.
- 3a. Add 1 ml of the overnight culture to 400 ml LB/ampicillin medium in a 2-liter flask.

Grow at 37°C with vigorous shaking until OD₆₀₀ reaches 0.5.

- 4a. Add 1.6 ml of 100 mM IPTG (0.4 mM final). Grow cells an additional 2 hr.

IPTG is a nonmetabolizable analog of lactose, the natural inducer of the lac promoter.

To express trpE fusion proteins using pATH vectors:

- 1b. Subclone gene of interest into a pATH vector in the correct reading frame, transform competent *E. coli* cells, and select transformants on supplemented M9/tryptophan plates.

- 2b. Inoculate 2 to 5 ml supplemented M9/tryptophan medium with a single colony containing the expression vector. Grow overnight at 37°C with shaking.

E. coli RR1 (Bolivar et al., 1977) has been used successfully for expressing trpE fusions; C600 and HB101 have also been used.

Inclusion of tryptophan in the medium represses expression of the trp operon. If fusion protein production is toxic for the host cell, repression of expression will promote maintenance of the plasmid in cells prior to induction.

- 3b. Add 1 ml of the overnight culture to 400 ml supplemented M9 medium *without* tryptophan. Grow at 37°C with vigorous shaking until OD₆₀₀ reaches 0.5.

- 4b. Add 1.6 ml of 2.5 mg/ml IAA (10 µg/ml final). Grow an additional 2 hr.

Growth of cells in medium without tryptophan induces expression of the trp operon. The addition of IAA to the medium further induces expression (Morse et al., 1969). IAA is an analog of the co-repressor tryptophan and competes with tryptophan for binding to trp repressor protein. The IAA-repressor complex is unable to bind to the trp operator and repress transcription, thereby allowing high levels of expression (Joachimiak et al., 1983).

Prepare protein extracts

Perform all steps on ice or at 4°C.

5. Split the cell culture into two 200-ml centrifuge bottles. Harvest cells by centrifugation for 10 min in a Sorvall RC-5B using a GSA rotor at 5000 rpm (4000 × g) and discard supernatant.
6. Resuspend each pellet in 5 ml PBS by pipetting up and down. Transfer each to a 15-ml conical centrifuge tube.
7. Centrifuge 10 min in a Sorvall Omnispin clinical centrifuge at 3500 rpm (3000 × g) and discard supernatant.
8. Resuspend cells in 2 ml HEMGN buffer with protease inhibitors. Add 20 µl of 50 mg/ml lysozyme (0.5 mg/ml final). Incubate 15 to 30 min on ice.
9. Disrupt the cells by sonicating two times for 15 sec each using a microtip, placing the sample on ice between rounds of sonication. Pool the lysate into one 50-ml centrifuge tube.

Hold the tube containing the sample so the tip of the sonicator probe is at the surface of the solution. Adjust the output level of the sonicator to the minimum setting required to achieve vigorous churning of the solution without creating foam.

Alternatively, a cup sonicator may be used. The advantage is that it is easier to keep the extract cold during sonication by filling the cup with ice water. The major disadvantage is that DNA is not as completely sheared, resulting in a lysate that is more viscous, thereby making it more difficult to get good separation between the soluble and insoluble fractions

during subsequent centrifugation. Sonicate two times for 30 sec each if using a cup sonicator.

10. Centrifuge cell lysate 15 min in a Sorvall SS-34 rotor at 15,000 rpm ($27,000 \times g$).
11. Pour off the supernatant and save (check for the presence of the fusion protein and its biological activity in step 17). Save the pellet, which usually contains almost all of the induced protein in an insoluble form. To prepare soluble protein, proceed to the next step.

The insoluble material from this step can be used for purification of the fusion protein by gel electrophoresis.

Solubilize the fusion protein

12. Resuspend the pellet in 2 ml HEMGN buffer with protease inhibitors.

The pellet is usually very viscous and can be difficult to resuspend. Scrape the pellet off the side of the tube using a pipettor and pipet up and down a few times. Resuspension of the pellet can be aided by further sonication.

13. Add 2 ml HEMGN buffer/8 M guanidine-HCl (4 M final guanidine). Incubate with gentle shaking 30 min at 4°C.
14. Centrifuge 30 min in a precooled ultracentrifuge using a Beckman 60Ti rotor at 35,000 rpm ($87,000 \times g$).
15. Transfer the supernatant to dialysis tubing and dialyze in three steps, each ≥ 3 hr to overnight: first, against 500 ml HEMGN buffer/1 M guanidine-HCl, and then twice against 1 liter HEMGN buffer excluding guanidine.

Only the protease inhibitors PMSF and sodium meta-bisulfite are necessary in the dialysis buffer (see reagents and solutions); the others are relatively expensive and have been safely omitted at this step. During dialysis about half of the protein in the extract will come out of solution as the guanidine is removed, resulting in the formation of a large amount of white precipitate.

16. Transfer all material in the dialysis bag to a centrifuge tube and centrifuge 5 min in an SS-34 rotor at 10,000 rpm ($12,000 \times g$) to remove insoluble material. Save the supernatant (~4 ml), which should be a clear, colorless solution with a protein concentration of ~1 mg/ml. (The fusion protein typically constitutes between 1% and 10% of the total protein.) Save the insoluble pellet, which usually contains most of the fusion protein and can be used for gel purifying the protein.

This pellet fraction usually contains a higher percentage of the fusion protein than the first insoluble fraction of the lysate (from step 11).

17. Determine the protein concentration of the supernatant and pellet from the cell lysate (steps 11 and 12) and of the final supernatant and pellet after guanidine extraction (step 16) by the Bradford method. Check the results of the induction by SDS-PAGE, loading 5 to 10 μg of protein per lane.

*For analysis of fusion protein production, it is useful to compare extracts containing the fusion protein with control extracts prepared from uninduced cells or from cells containing the expression vector with no insert. The molecular weight of the β -gal protein is 116 kDa, and the *trpE* protein is 37 kDa. The molecular weight of the fusion protein can be predicted from the length of the open reading frame ligated to the vector sequences.*

REAGENTS AND SOLUTIONS

HEMGN buffer

100 mM KCl
25 mM HEPES, pH 7.6
0.1 mM EDTA, pH 8.0
12.5 mM MgCl₂
10% glycerol
0.1% Nonidet P-40

Prepare 2 liters of the above solution. For resuspending the cells (step 8) and resuspending the pellet (step 12), take 10 ml and add 10 μ l of each of the following just before using (save the remainder of the 2 liters for dialysis):

1 M dithiothreitol (DTT; 1 mM final)
2 mg/ml aprotinin (2 μ g/ml final)
1 mg/ml leupeptin (1 μ g/ml final)
1 mg/ml pepstatin in methanol (1 μ g/ml final)
100 mM phenylmethylsulfonyl fluoride (PMSF) in ethanol (0.1 mM final)
100 mM sodium *meta*-bisulfite (0.1 mM final)

The last five ingredients are protease inhibitors. Solutions of DTT, aprotinin, leupeptin, and pepstatin should be stored at -20°C. PMSF solution should be stored at 4°C. Sodium meta-bisulfite should be made fresh and stored on ice during use.

For dialysis, only PMSF and sodium meta-bisulfite are necessary for protease inhibition (see annotation to step 15). DTT should be included in all the dialysis buffers.

COMMENTARY

Background Information

The production of fusion proteins in *E. coli* has become a widely used technique among molecular biologists. The principle advantages of this strategy for overproducing proteins are the relative ease of constructing the appropriate expression vector and the large amounts of protein that can be produced. A potential disadvantage of fusion proteins is that they are usually predominantly present in the insoluble fraction of the cell lysate. However, the insolubility of fusion proteins is actually beneficial for their most common application, producing protein for use as an antigen for generating antibodies, because the insolubility of the protein can be used as a significant purification step prior to final purification by preparative-scale SDS-PAGE (UNIT 10.8; Rio et al., 1986; Spindler et al., 1984). In addition, insolubility of fusion proteins may protect them from degradation by host cells.

In many cases, fusion proteins have also been used to study the biochemical properties of the induced protein. For example, *lacZ* fusions have been used to study the DNA binding activities of several proteins (Johnson and Herskowitz, 1985; Desplan et al., 1985; Kadonaga et al., 1987), and *trpE* fusions have been used

to characterize the properties of reverse transcriptase (Tanese et al., 1985) and a protein-tyrosine kinase (Sadowski et al., 1986).

An advantage of using fusion proteins for biochemical studies is that the β -gal or *trpE* moiety can be used as a means of identifying the protein. Antibodies directed against the β -gal or *trpE* protein can be used for affinity purification of the protein or to follow the protein during chromatographic fractionation (Rüther and Müller-Hill, 1983; Johnson and Herskowitz, 1985). Of course, a potential problem in using fusion proteins for biochemical analyses is that the properties of the native protein may be significantly altered because of inclusion of the foreign residues encoded by *trpE* or *lacZ*. Therefore, if the major goal is to purify the protein to study its enzymatic or other biochemical activities, a different type of expression vector containing little or no coding sequence downstream of the translation start site might be more suitable (UNITS 16.2 & 16.3).

Critical Parameters and Troubleshooting

Typically, fusion proteins are expressed at a very high level in the induced cell using the

methods described in this protocol. However, if expression of the expected product is low, it can sometimes be attributed to extreme toxicity caused by the induced protein or by protein instability (or both). One approach that is usually successful in overcoming the toxicity problem is to keep the cell density low prior to induction. If a small amount of fusion protein production kills the cells, the cells that are making protein productively before induction will be lost during growth of an overnight culture, and the culture will be overtaken by cells that have lost the plasmid or by cells containing mutations that prevent expression of the protein. To circumvent this problem, the amount of ampicillin in the medium can be increased (up to 200 $\mu\text{g/ml}$), and an overnight culture should not be used to inoculate the 400-ml culture. Instead, start a 2-ml culture from a fresh plate and let it grow 2 to 3 hr until it reaches mid-log phase. Pour the small culture into 400 ml fresh medium and grow to an OD_{600} of 0.5 (this usually takes 4 to 5 hr).

Proteolysis sometimes occurs in fusion proteins at or near the junction of the β -gal or *trpE* sequences and the heterologous sequences. Most commonly, proteolysis occurs within the cell after induction. This can be tested by checking the amounts of full-length protein and breakdown products at various times after induction. Whole cells can be lysed by boiling in 1 \times SDS sample buffer and then loaded onto an SDS-polyacrylamide gel (a sample prepared in this way is very viscous and difficult to load on a gel; the viscosity can be reduced by sonication). Shorter induction times may increase the ratio of full-length protein to breakdown products, although the total amount of protein will be lower.

Alternatively, proteolysis can occur during preparation of the extract. To check this, compare the relative amounts of full-length protein and truncated versions from various stages in making the extract with the state of the protein immediately after harvesting the cells. If a significant amount of proteolysis occurs during preparation of the extract, try more or different kinds of protease inhibitors (see the Boehringer Mannheim catalog for the different types of protease inhibitors). The use of protease-deficient mutants of *E. coli* might also be useful (Buell et al., 1985).

Other potential problems are complete insolubility of the fusion protein even after guanidine treatment or loss of biochemical ac-

tivity due to the guanidine denaturation. The guanidine denaturation/renaturation method described here works quite well for most proteins. However, each protein has some unique characteristics and this method cannot be expected to be optimal in every case. For both problems, one thing to try is to use urea as a denaturant rather than guanidine. Although guanidine and urea denature proteins by a similar mechanism (forming hydrogen bonds with amino acid residues, leading to disruption of the secondary and tertiary structure of the protein), some proteins may respond differently to the two denaturants. The insoluble fraction of the lysate can be resuspended in urea at concentrations up to 8 M, then dialyzed in steps to gradually reduce the urea concentration. In some cases, overexpressed proteins can be solubilized without any denaturant by extraction with high salt (Johnson and Herskowitz, 1985; Gross et al., 1976). If the protein cannot be solubilized or recovered with any activity, try expressing the protein in a different type of construct. Smaller proteins might be more stable or more soluble expressed from other constructs while retaining the domains required for biological activity (Tanese et al., 1985; Desplan et al., 1985). Other approaches that may result in improved solubility are to shorten the time of induction or to perform the induction at a lower temperature. In some cases the solubility of the expressed protein has been increased by inducing for 30 min at 30°C.

Anticipated Results

A 400-ml culture of *E. coli* will yield 15 mg to 20 mg of total protein. About two-thirds of the total protein will be in the soluble fraction of the lysate, and one-third in the insoluble fraction (step 11). The induced protein usually represents ~1% of the soluble protein and 10% to 20% of the insoluble protein. Therefore, a typical yield from a 400-ml culture is between 0.5 mg and 1 mg of induced protein. The guanidine denaturation/renaturation procedure usually solubilizes half of the protein in the pellet, and, on average, the fusion protein is ~5% to 10% of the total in the guanidine-extracted supernatant. The material that precipitates during removal of guanidine by dialysis is ~25% to 50% induced protein.

Time Considerations

It takes 2 to 3 hr to grow a 400-ml culture to an OD_{600} of 0.5 starting with an overnight

culture. The cells are usually grown an additional 2 hr after induction. Preparation of the protein extract requires ~3 hr to reach the first dialysis step which can then go 3 hr to overnight. Analysis of the protein extract takes 1 to 2 days.

Literature Cited

- Bolivar, F. Rodriguez, R.L., Greene, T.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W. 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* 2:95-113.
- Buell, G., Schulz, M.F., Selzer, G., Chollet, A., Movva, N.R., Semon, D., Escanez, S., and Kawishima, E. 1985. Optimizing the expression in *E. coli* of a synthetic gene encoding somatomedin-C (IGF-I). *Nucl. Acids Res.* 13:1923-1938.
- Desplan, C., Theis, J., and O'Farrell, P.H. 1985. The *Drosophila* developmental control gene, engrailed, encodes a sequence-specific DNA binding activity. *Nature (London)* 318:2513-2522.
- Gross, C., Engbaek, F., Flammang, T., and Burgess, R. 1976. Rapid micromethod for the purification of *Escherichia coli* ribonucleic acid polymerase and the preparation of bacterial extracts active in ribonucleic acid synthesis. *J. Bacteriol.* 128:382-389.
- Joachimiak, A., Schevitz, R.W., Kelley, R.L., Yanofsky, C., and Sigler, P.B. 1983. Functional inferences from crystals of *Escherichia coli* *trp* repressor. *J. Biol. Chem.* 258:12641-12643.
- Johnson, A.D. and Herskowitz, I. 1985. A repressor (MAT α 2 product) and its operator control expression of a set of cell type specific genes in yeast. *Cell* 42:237-247.
- Kadonaga, J.T., Carner, K.R., Masiarz, F.R., and Tjian, R. 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* 51:1079-1090.
- Koerner, T.J., Hill, J.E., Myers, A.M., and Tzagoloff, A. 1990. High-expression vectors with multiple

cloning sites for construction of *trpE*-fusion genes: pATH vectors. *Methods Enzymol.* In press.

- Morse, D.E., Mosteller, R.D., Baker, R.F., and Yanofsky, C. 1969. Direction of in vivo degradation of tryptophan messenger RNA—a correction. *Nature (London)* 223:40-43.
- Rio, D.C., Laski, F.A., and Rubin, G.M. 1986. Identification and immunological analysis of biologically active *Drosophila* P element transposase. *Cell* 44:21-32.
- Rüther, U. and Müller-Hill, B. 1983. Easy identification of cDNA clones. *EMBO J.* 2:1791-1794.
- Sadowski, I., Stone, J.C., and Pawson, T. 1986. A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami Sarcoma Virus P130^{gag-*trp*s}. *Mol. Cell. Biol.* 6:4396-4408.
- Spindler, K.R., Rossner, D.S.E., and Berk, A.J. 1984. Analysis of Adenovirus transforming proteins from early regions 1A and 1B with antisera to inducible fusion antigens produced in *Escherichia coli*. *J. Virol.* 49:132-141.
- Tanese, N., Roth, M., and Goff, S.G. 1985. Expression of enzymatically active reverse transcriptase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 82:4944-4948.

Key Reference

Koerner et al., 1990. See above.

Contains a complete description of the different trpE expression vectors as well as a thorough discussion of potential problems involved in maximizing expression.

Contributed by Timothy Hoey
University of California
Berkeley, California

Expression and Purification of Maltose-Binding Protein Fusions

The maltose-binding protein (MBP) vectors allow the expression and purification of a protein encoded by a cloned gene by fusing it to MBP, which is encoded by the *malE* gene of *Escherichia coli* (Fig. 16.6.1). This method uses the strong, inducible *tac* promoter and the *malE* translation initiation signals to give high-level expression of the cloned gene, and isolation of the fusion protein is facilitated by an affinity purification for MBP (Kellerman and Ferenci; Fig. 16.6.2).

The gene or open reading frame of interest is subcloned into one of the MBP vectors—pMAL-c2 or pMAL-p2—so the coding sequence of interest is fused in-frame to the 3' end of the *malE* gene. The vector pMAL-c2 has an exact deletion of the *malE* signal sequence which leads to cytoplasmic expression of the fusion protein. Fusion plasmids made with this vector usually give a higher yield of fusion protein than those made with pMAL-p2. The vector pMAL-p2 contains the normal signal sequence of the *malE* gene, which potentially directs the fusion protein through the cytoplasmic membrane to the periplasm. This promotes proper folding and disulfide bond formation for some proteins and allows purification from the periplasm (Hsiung et al., 1986; Lauritzen et al., 1991).

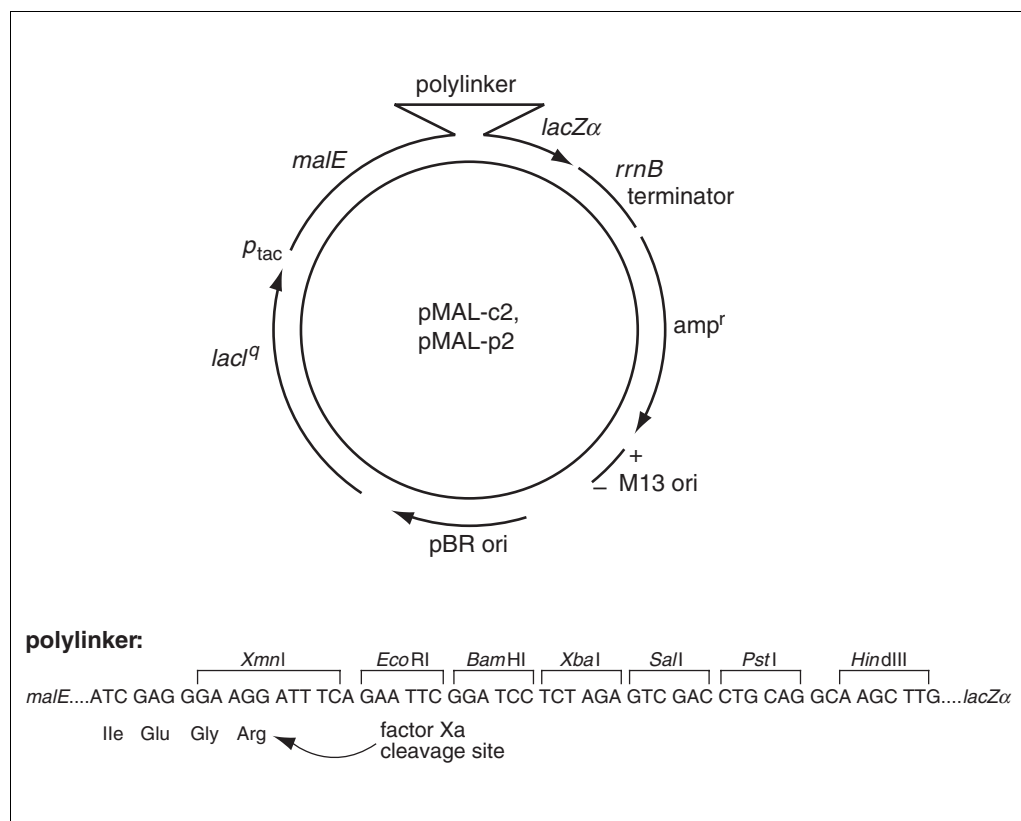


Figure 16.6.1 pMAL vectors. The vectors contain the inducible p_{tac} promoter positioned to transcribe a *malE-lacZ* gene fusion. The *lacI* gene encodes the *lac* repressor, which turns off transcription from p_{tac} until IPTG is added. The polylinker provides several restriction endonuclease sites for inserting the gene of interest so it is fused to the *malE* gene. The *rrnB* terminators prevent transcription from p_{tac} from interfering with plasmid replication or maintenance functions. pMAL-c2 (6646 base pairs) has an exact deletion of the *malE* signal sequence. pMAL-p2 (6721 base pairs) includes the *malE* signal sequence. Arrows indicate the direction of transcription. Restriction sites indicated are unique.

The MBP vectors also include a sequence that encodes the four-amino-acid recognition site for the specific protease factor Xa. The site is placed so it can be used to separate the protein of interest from MBP after affinity purification. Factor Xa cuts after arginine in the sequence Ile-Glu(or Asp)-Gly-Arg. Depending on which restriction endonuclease site is used to insert the gene of interest, this means that few or no amino acids encoded by the vector are present on the protein of interest after cleavage by factor Xa (see UNIT 16.4B).

The basic protocol of this unit outlines subcloning the sequence encoding the protein of interest into an MBP vector, and expressing and purifying the protein from the cytoplasm. The first support protocol provides a pilot experiment for analyzing the solubility, affinity for the amylose resin, and export of a particular fusion protein. The alternate protocol gives instructions for purifying a fusion protein from the periplasm for fusions that are

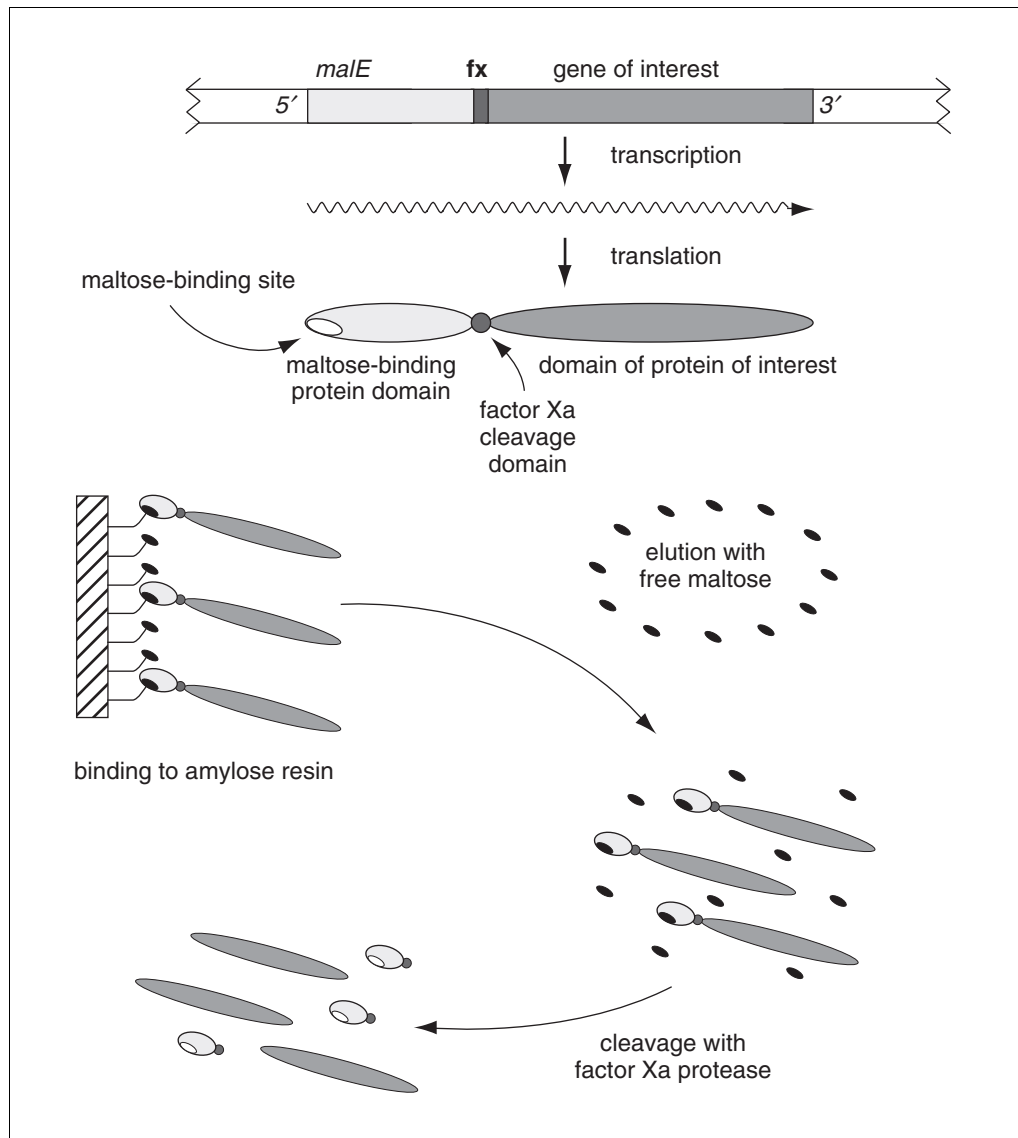


Figure 16.6.2 Schematic representation of expression and purification of a protein using the MBP vectors. “fx” denotes the sequence coding for the recognition site of factor Xa. The gene of interest is cloned 3’ to the *malE* gene and expressed, cells are lysed, and the extract is poured over a column of amylose resin. The MBP fusion protein binds to the column and the remaining proteins in the cell extract are washed through the column (Kellerman and Ferenci, 1982). The fusion protein is eluted with free maltose and then cleaved with factor Xa to separate MBP from the protein of interest. Modified from Maina et al., 1988 with permission of Elsevier Science Publishing.

made in the signal sequence vector and are exported. Methods for cleaving the purified fusion protein with factor Xa and for preparing a fusion protein for factor Xa cleavage when the cleavage site is initially blocked are provided in *UNIT 16.4B*. The second and third support protocols detail two different chromatographic methods for separating the protein of interest from MBP after factor Xa cleavage.

CONSTRUCTION, EXPRESSION, AND PURIFICATION OF MBP FUSION PROTEINS

The sequence encoding the protein of interest is subcloned into an MBP vector. Cells bearing the fusion plasmid are grown, the tac promoter is induced with isopropyl-1-thio- β -D-galactoside (IPTG), and the cells are harvested. A crude cell extract is prepared and passed over a column containing an agarose resin derivatized with amylose (a polysaccharide consisting of maltose subunits). The fusion protein binds to the column because of MBP's affinity for amylose, is eluted with free maltose, and is analyzed by SDS-PAGE (Fig. 16.6.2).

Materials

For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX2**; for suppliers, see **APPENDIX4**.

- pMAL-c2 and/or pMAL-p2 (New England Biolabs; see Fig. 16.6.1)
- Escherichia coli* TB1 (New England Biolabs) or any transformable strain of *E. coli* (see Table 1.4.5)
- LB plates and medium containing 0.2% glucose and 100 μ g/ml ampicillin (*UNIT 1.1*)
- LB plates containing 100 μ g/ml ampicillin, 0.3 mM isopropyl-1-thio- β -D-galactoside (IPTG), and 80 μ g/ml Xgal (*UNIT 1.1*)
- 1 \times and 2 \times SDS sample buffer (*UNIT 10.2*)
- 0.1 M IPTG
- Column buffer, without and with 10 mM maltose
- Coomassie brilliant blue solution (*UNIT 10.1*)
- Amylose resin (New England Biolabs)
- Beckman JS-4.2 and JA-17 rotors (or equivalents)
- Sonicator
- 2.5 \times 10-cm column
- Centricon or Centriprep concentrator *or* stirred-cell concentrator (Amicon), or equivalent
- Boiling water bath
- Additional reagents and equipment for replica plating (*UNIT 1.3*), subcloning of DNA fragments (*UNITS 1.4 & 3.16*), preparing miniprep plasmid DNA (*UNIT 1.6*), restriction mapping (*UNIT 3.2*), quantitation of protein by the Bradford method (*UNIT 10.1*), and SDS-PAGE (*UNIT 10.2*)

NOTE: A kit containing the pMAL vectors, TB1, amylose resin, factor Xa, anti-MBP serum, MBP as a gel marker, and an MBP fusion as a factor Xa control is available from New England Biolabs.

Obtain and grow the fusion plasmid

1. Subclone target gene or open reading frame into pMAL-c2 and/or pMAL-p2 so it is in the same translational reading frame as the *malE* gene of the vector (see Fig. 16.6.1). The insert should have a stop codon at the end of the coding sequence, and if the *XmnI* site is used, the first three bases of the insert should not code for a proline residue.

If an appropriate restriction site is not available at the 5' end of the gene, one must be introduced, e.g., by oligonucleotide-directed mutagenesis (UNIT 8.1), by cloning with synthetic oligonucleotides (UNIT 3.16), by PCR (UNIT 15.1), or by generating half restriction sites (UNIT 15.7).

If a stop codon is not present at the end of the open reading frame, a linker containing a stop codon can be inserted into one of the downstream polylinker sites after step 2. A proline codon immediately following the arginine codon of the factor Xa site will preclude separation of MBP from the protein of interest, as factor Xa cannot cleave the Arg-Pro bond (Nagai and Thøgersen, 1984, 1987).

2. Transform *E. coli* TB1 with the ligated vector + insert and plate on LB/glucose/ampicillin. Incubate overnight at 37°C.
3. Replica plate transformants onto an LB/glucose/ampicillin master plate and an LB/ampicillin/IPTG/Xgal indicator plate. Incubate 8 to 16 hr at 37°C. Identify transformants on the indicator plate (positive transformants are white, those with vector alone are blue). Recover clones with inserts from the corresponding spot on the master plate.

*The vectors are lethal to *E. coli* when *p_{tac}* is induced, so screening for inserts cannot be done with IPTG and Xgal in the transformation plates. For directional cloning (UNIT 3.16), the ligation efficiency is usually so high that the screen is unnecessary.*

4. Prepare miniprep DNA from candidate transformants and confirm the presence and orientation of the insert by restriction mapping.
5. Inoculate 10 ml LB/glucose/ampicillin medium with a single colony containing the fusion plasmid. Grow overnight at 37°C with shaking.
6. Inoculate 1 liter LB/glucose/ampicillin medium with 10 ml overnight culture (a 1:100 dilution). Grow at 37°C, shaking, to 2×10^8 cells/ml ($OD_{600} = 0.4$ to 0.6).
7. Remove a 1-ml sample, microcentrifuge 2 min, and discard supernatant. Resuspend cells in 50 μ l of 1 \times SDS sample buffer and set aside on ice for analysis in steps 20 and 21 (uninduced cells).

Induce the promoter and harvest the cells

8. Add 3 ml of 0.1 M IPTG (0.3 mM final) to remainder of culture. Incubate 2 hr at 37°C with good aeration.
9. Remove another 1-ml sample, microcentrifuge 2 min, and discard supernatant. Resuspend cells in 100 μ l of 1 \times SDS sample buffer, and set aside on ice for analysis in steps 20 and 21 (induced cells).
10. Centrifuge remaining cells 20 min at $4000 \times g$ (4200 rpm in a JS-4.2 rotor), 4°C, and discard supernatant.
11. Resuspend cells in 50 ml column buffer. Freeze sample at -70°C (or overnight at -20°C). Thaw in cold water.

Freezing weakens the cells so that lysis occurs more readily. Slow freezing overnight at -20°C is more effective than quick freezing at -70°C . This step can be omitted, but doing so necessitates more sonication in step 12.

12. Put cell suspension in a plastic or metal tube and place in an ice-water bath. Sonicate using short bursts to avoid heating the extract (if sonicator has a “pulse” mode, use a 50% duty cycle). Monitor release of protein by adding 10 μ l sonicate to 1.5 ml Coomassie brilliant blue to obtain a Bradford reaction. Check protein concentration

every 30 sec of sonication time. Continue sonication until released protein reaches a maximum (usually ~1 to 3 min of sonication time).

13. Centrifuge the sonicated cells 20 min at $14,000 \times g$ (10,000 rpm in a JA-17 rotor), 4°C . Discard pellet.
14. Remove a 5- μl sample of the supernatant (crude extract), add 5 μl of 2 \times SDS sample buffer, and set aside on ice for analysis in steps 20 and 21. Save the rest of the supernatant for passage over the amylose resin column in step 16.

Purify and analyze the fusion protein

15. Pour amylose resin in a 2.5×10 -cm column. Wash with 8 column volumes of column buffer.

The amount of resin needed depends on the amount of fusion protein produced. The resin binds about 3 mg/ml bed volume; a column of ~15 ml should be sufficient for a yield of ≤ 45 mg fusion protein/liter culture. A 50-ml syringe plugged with silanized glass wool can be substituted for the column, but the glass wool should cover the bottom of the syringe (not just fill the tip) to achieve an acceptable flow rate.

16. Dilute crude extract 1:5 with column buffer. Load at a flow rate of ~1 ml/min [$10 \times$ (diameter of column in cm)² ml/hr].
17. Wash with 12 column volumes of column buffer.
18. Elute fusion protein with column buffer/10 mM maltose, collecting ten to twenty 3-ml fractions. Assay fractions for protein by the Bradford method or by measuring the A_{280} .

The fractions containing the MBP fusion should have easily detectable protein. The fusion protein starts to elute after ~5 ml. Proteins that contain tyrosine and/or tryptophan residues absorb at 280 nm, so A_{280} measurements can often be used instead of the Bradford method to detect protein.

19. Pool the protein-containing fractions. If necessary, concentrate to ~1 mg/ml in a Centricon or Centriprep concentrator or stirred-cell concentrator.
20. Remove 3 μg eluted protein and mix with an equal volume of 2 \times SDS sample buffer. Heat uninduced cells, induced cells, crude extract, and eluted protein 5 min in a boiling water bath.
21. Electrophorese on an SDS-polyacrylamide gel. Load 20 μl uninduced and induced cell samples, all 10 μl crude extract sample, and ~3 μg eluted protein.

If the fusion protein is present in the crude extract, but absent from the maltose-eluted fraction, use the following pilot experiment (first support protocol) to determine if the fusion protein is soluble and has affinity for the amylose resin.

SUPPORT PROTOCOL 1

PILOT EXPERIMENT TO CHARACTERIZE THE BEHAVIOR OF AN MBP FUSION PROTEIN

This is a small-scale experiment to determine the behavior of a particular MBP fusion protein. It can be used either as a pilot experiment or to troubleshoot the basic protocol. The protocol results in five (pMAL-c2) or seven (pMAL-p2) samples: uninduced and induced cells, a crude extract of the soluble proteins, a suspension of the insoluble material, a sample of protein that binds to the amylose resin, and (if the fusion was made with pMAL-p2) a periplasmic fraction prepared by the cold osmotic shock procedure and the cells that remain after osmotic shock.

Additional Materials

For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

Cells containing fusion plasmid (basic protocol)
30 mM Tris·Cl/20% sucrose, pH 8.0
0.5 M EDTA, pH 8.0 (**APPENDIX 2**)
5 mM MgSO₄, ice-cold

Grow, induce, and harvest the cells

1. Inoculate 80 ml LB/glucose/ampicillin medium with cells containing the fusion plasmid (e.g., a 1:100 dilution from an overnight culture from step 3, basic protocol).
2. Grow at 37°C with good aeration to 2×10^8 cells/ml ($OD_{600} = 0.4$ to 0.6).
3. Remove a 1-ml sample, microcentrifuge 2 min, and discard supernatant. Resuspend cells in 50 μ l of 1 \times SDS sample buffer and place on ice for analysis in step 12 (uninduced cells).
4. Add 0.24 ml of 0.1 mM IPTG (0.3 mM final) to remaining culture. Continue incubation 2 hr at 37°C.
5. Remove a 1-ml sample, microcentrifuge 2 min, and discard supernatant. Resuspend cells in 100 μ l of 1 \times SDS sample buffer and place on ice for analysis in step 12 (induced cells).

Take samples at 1 and 3 hr to help determine optimum expression time.

6. Divide culture into two aliquots (samples A and B). Centrifuge 10 min at $7000 \times g$ (7000 rpm in a JA-17 rotor), 4°C, and discard supernatant.
7. Resuspend one pellet (sample A) in 5 ml column buffer. Resuspend the other pellet (sample B) in 10 ml of 30 mM Tris·Cl/20% sucrose, pH 8.0 (8 ml for each 0.1 g cells, wet weight).

Prepare samples A and B

For sample A (*pMAL-c2* and *pMAL-p2*):

- 8a. Freeze sample in a dry ice/ethanol bath or at -20°C and thaw in cold water. Put cell suspension in a plastic or metal tube, place in an ice-water bath, and sonicate using short bursts as in step 12 of the basic protocol.
- 9a. Centrifuge the cells 20 min at $14,000 \times g$ (10,000 rpm in a JA-17 rotor), 4°C. Decant supernatant (crude extract) and save on ice for analysis in steps 11a and 12. Resuspend pellet (insoluble matter) in 5 ml column buffer and save on ice for analysis in step 12.
- 10a. Place 200 μ l amylose resin in a 1.5-ml tube and microcentrifuge briefly. Remove supernatant by aspiration and discard. Resuspend resin in 1.5 ml column buffer, then microcentrifuge briefly and discard supernatant. Repeat this wash and resuspend the resin in 200 μ l column buffer.
- 11a. Mix 100 μ l crude extract with 50 μ l amylose resin slurry. Incubate 15 min on ice. Microcentrifuge the sample 1 min, then remove supernatant and discard. Wash pellet with 1 ml column buffer, microcentrifuge 1 min, and resuspend the resin (bound protein) in 50 μ l of 1 \times SDS sample buffer. Save for analysis in step 12.

For sample B (pMAL-p2 only):

- 8b. Add 20 μ l of 0.5 M EDTA, pH 8.0 (1 mM final) and incubate 5 to 10 min at room temperature with shaking or stirring.
- 9b. Centrifuge the cells 10 min at $7000 \times g$ (7000 rpm in a JA-17 rotor), 4°C. Remove and discard all the supernatant.
- 10b. Resuspend pellet in 10 ml ice-cold 5 mM $MgSO_4$. Shake or stir 10 min in an ice bath.
- 11b. Centrifuge the cells 10 min at $14,000 \times g$ (10,000 rpm in a JA-17 rotor), 4°C. Save the supernatant (cold osmotic shock fluid). Resuspend pellet (shocked cells) in 5 ml column buffer.

Analyze samples by SDS-PAGE

12. Add 5 μ l of 2 \times SDS sample buffer to 5 μ l crude extract, insoluble matter, and (for pMAL-p2) shocked cells. Add 10 μ l of 2 \times SDS sample buffer to 10 μ l cold osmotic shock fluid. Heat these samples, along with the uninduced and induced cell samples and the amylose resin sample, 5 min in a boiling water bath. Microcentrifuge 2 min.
13. Load the samples on a 10% SDS-polyacrylamide gel as follows: 20 μ l uninduced cells, 20 μ l induced cells, 10 μ l crude extract, 10 μ l insoluble matter, 20 μ l bound protein, and (for pMAL-p2 fusions only) 20 μ l osmotic shock fluid and 10 μ l shocked cells. Electrophorese.

If fusion protein is present in cold osmotic shock fluid, use the purification procedure in the following alternate protocol.

If desired, prepare an immunoblot (UNIT 10.8) and develop with anti-MBP serum (New England Biolabs) and, if available, serum directed against the protein of interest.

ALTERNATE PROTOCOL

PURIFICATION OF FUSION PROTEINS FROM THE PERIPLASM

This protocol is useful if the protein of interest is exported and must be purified from the periplasm. The method includes preparation of a periplasmic fraction by cold osmotic shock followed by affinity purification as in the Basic Protocol.

Additional Materials

*For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.*

Cells containing fusion plasmid (basic protocol)
30 mM Tris·Cl/20% (w/v) sucrose, pH 8.0
0.5 M EDTA, pH 8.0 (*APPENDIX 2*)
5 mM $MgSO_4$
1 M Tris·Cl, pH 7.4 (*APPENDIX 2*)
Beckman JA-14 rotor (or equivalent)

1. Inoculate 1 liter LB/glucose/ampicillin medium with cells containing the fusion plasmid (e.g., a 1:100 dilution from an overnight culture from step 3, basic protocol).
2. Grow to 2×10^8 cells/ml ($OD_{600} = 0.4$ to 0.6).
3. Add 3 ml of 0.1 M IPTG (0.3 mM final). Incubate cells 2 hr at 37°C with good aeration.
4. Centrifuge the cells 10 min at $7000 \times g$ (7000 rpm in a JA-14 rotor), 4°C, and discard supernatant.

5. Resuspend cells in 400 ml of 30 mM Tris·Cl/20% sucrose, pH 8.0 (80 ml for each gram of cells, wet weight). Add 0.8 ml of 0.5 M EDTA, pH 8.0 (1 mM final) and incubate 5 to 10 min at room temperature with shaking or stirring.
6. Centrifuge the cells 10 min at $10,000 \times g$ (8000 rpm in a JA-14 rotor), 4°C. Remove all supernatant and resuspend pellet in 400 ml ice-cold 5 mM MgSO₄. Shake or stir 10 min in an ice bath.
7. Centrifuge again as in step 6. Save the supernatant (cold osmotic shock fluid) and add to it 8 ml of 1 M Tris·Cl, pH 7.4.
8. Prepare amylose resin column as in step 15 of the basic protocol. Load cold osmotic shock fluid on column at a flow rate of ~1 ml/min [$10 \times (\text{diameter of column in cm})^2$ ml/hr].
9. Recover and analyze protein fractions as in steps 17 to 21 of the basic protocol.

PURIFYING THE CLEAVED PROTEIN BY ION EXCHANGE CHROMATOGRAPHY

**SUPPORT
PROTOCOL 2**

After factor Xa cleavage (UNIT 16.4B), the protein of interest can be separated from MBP, factor Xa, and trace contaminants by Sepharose ion-exchange chromatography. Cleaved fusion protein amounts of ≤ 25 mg can be purified by this procedure, which can also be scaled up for larger amounts.

Additional Materials

For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

Fusion protein cleaved with factor Xa (UNIT 16.4B)

20 mM Tris·Cl/25 mM NaCl, pH 8.0

DEAE-Sepharose or Q-Sepharose resin

20 mM Tris·Cl/500 mM NaCl, pH 8.0

1 \times 10–cm column

Additional reagents and equipment for purification of proteins by conventional chromatography (UNIT 10.10)

1. Dialyze fusion protein cleaved with factor Xa two or three times against ≥ 100 vol of 20 mM Tris·Cl/25 mM NaCl, pH 8.0.
2. Wash ~6 ml Sepharose resin twice in 20 ml of 20 mM Tris·Cl/25 mM NaCl, letting resin settle and pouring off supernatant between washes.
3. Pour resin into a 1 \times 10–cm column to give a bed volume of 5 ml with a 6 to 7-cm bed height. Wash with 15 ml of 20 mM Tris·Cl/25 mM NaCl.
4. Load dialyzed fusion protein onto column. Collect 2.5-ml fractions of the column flow-through.
5. Wash column with 3 to 5 column volumes of 20 mM Tris·Cl/25 mM NaCl. Continue collecting 2.5-ml fractions.
6. Start a gradient of 20 mM Tris·Cl/25 mM NaCl to 20 mM Tris·Cl/500 mM NaCl, pH 8.0. Collect 1-ml fractions.
7. Identify protein-containing fractions by the Bradford method or by measuring A_{280} .
MBP elutes as a sharp peak at 100 to 150 mM NaCl. Factor Xa elutes at ~400 mM NaCl. The target protein may flow through the column, or it may elute during the gradient.

**Protein
Expression**

16.6.8

- Analyze relevant fractions by SDS-PAGE. Pool fractions containing the target protein free of MBP and concentrate as needed.

**SUPPORT
PROTOCOL 3**

**PURIFYING THE CLEAVED PROTEIN BY AFFINITY
CHROMATOGRAPHY**

In this method, all the protein from the factor Xa cleavage reaction (*UNIT 16.4B*) is bound to hydroxylapatite and the maltose is washed away from the bound protein. This step is a substitute for dialysis, which is impractical for removing all traces of maltose from the MBP binding site (Silhavy et al., 1975). The protein mixture is then passed over a fresh amylose column and the MBP binds to the column, allowing the protein of interest to be collected in the flow-through. Although this protocol requires two steps, no dialysis is needed and both columns are step-eluted, so the procedure is fairly simple. It should be carried out at room temperature to avoid precipitation of the 0.5 M sodium phosphate buffer. This method does not remove factor Xa or any trace contaminants, and any MBP that has been denatured or otherwise damaged will not bind to the amylose column. The procedure is written for ≥ 25 mg, and can be scaled up for larger amounts.

Additional Materials

For recipes, see Reagents and Solutions in this unit (or cross-referenced unit); for common stock solutions, see APPENDIX2; for suppliers, see APPENDIX4.

Hydroxylapatite resin
Fusion protein cleaved with factor Xa (*UNIT 16.4B*)
0.5 M sodium phosphate, pH 7.2
1 or 1.5×10 -cm column

Additional reagents and equipment for hydroxylapatite chromatography (*UNIT 2.10*)

1. Swell 1 g hydroxylapatite in column buffer for 15 min. Wash resin three times in column buffer by swirling it, then letting it settle 1 min. Pour off milky supernatant.

This washes the resin and removes the "fines" that can cause the column to flow slowly.

2. Pour hydroxylapatite into a 1 or 1.5×10 -cm column.

The larger-diameter column will give a faster flow rate.

3. Load fusion-protein cleaved with factor Xa onto the column. Wash column with 80 ml column buffer.

This step removes the maltose.

4. Elute protein with 20 ml of 0.5 M sodium phosphate, pH 7.2. Collect 2-ml fractions. Assay for protein by the Bradford method or by measuring the A_{280} .

Most of the protein usually elutes in the first 8 ml.

5. Prepare 15-ml amylose column as in step 15 of the basic protocol.

6. Load hydroxylapatite-eluted protein onto the amylose column. Reapply flow-through two times. Collect final flow-through as 5-ml fractions. Wash column with 50 ml column buffer, continuing to collect 5-ml fractions. Assay for protein by the Bradford method or by measuring the A_{280} . Concentrate the protein as needed.

Most of the protein flows through in the first 35 ml. Protein should be free of MBP and consist of target protein and factor Xa.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX2; for suppliers, see APPENDIX4.

Column buffer

Per liter:

20 ml 1.0 M Tris·Cl, pH 7.4 (20 mM final)

11.7 g NaCl (0.2 M final)

2.0 ml 0.5 M EDTA (1 mM final)

0.7 ml 2-ME (optional; 10 mM final)

1 ml 1 M NaN₃ (optional; 1 mM final)

Store ~1 month at room temperature

Isopropyl-1-thio-β-D-galactoside, 0.1 M

1.41 g isopropyl-β-D-thiogalactoside (IPTG)

H₂O to 50 ml

Filter sterilize and store ~1 year at 4°C

0.5 M sodium phosphate buffer, pH 7.2

A: 69.0 g NaH₂PO₄·H₂O, to 1 liter with H₂O

B: 134.0 g Na₂HPO₄·7H₂O, to 1 liter with H₂O

Mix 117 ml A with 383 ml B

Store indefinitely at room temperature

COMMENTARY

Background Information

The maltose binding protein (MBP) vectors (Fig. 16.6.1) were constructed as a way of expressing and purifying a cloned protein or peptide by fusing it to maltose-binding protein, coded for by the *malE* gene of *Escherichia coli* (Guan et al., 1988; Bedouelle and Duplay, 1988; Maina et al., 1988). The vectors have the strong, inducible *tac* promoter (Amann and Brosius, 1985) expressing the *malE* gene (with or without its signal sequence), fused to the *lacZα* gene. When present, the signal peptide on pre-MBP directs fusion proteins to the periplasm. Restriction endonuclease sites between *malE* and *lacZα* are available for insertion of the coding sequence of interest in a way that interrupts the *malE-lacZα* fusion.

The *p_{tac}* region of these plasmids is from pTP201 (T. Poteet, unpub. observ.; the same construction as pKK207-1, Amann & Brosius, 1985), with a fragment carrying *lacI^q* (*EcoRI-BanI* partial-digest product, filled in with Klenow fragment) adjacent upstream from *p_{tac}*. The *lacI^q* gene codes for the *lac* repressor, which keeps expression from *p_{tac}* low in the absence of isopropyl-1-thio-β-D-galactoside (IPTG) induction. The *malE* region is the *HinFI* (filled in) fragment of *malE* (Duplay et al., 1984). The vector backbone is from *AvaI* (filled in) to *EcoRI* (filled in) of pKK233-2 (Amann

& Brosius, 1985). This gives the vectors a pBR322-like copy number. The vectors also include the origin of DNA replication of bacteriophage M13 from the plasmid pZ150 (Zagursky and Berman, 1984), to facilitate the production of single-stranded DNA for sequencing or mutagenesis (UNIT 1.15).

Insertion of the target sequences downstream of *malE* interrupts the *malE-lacZα* fusion. This allows use of the blue-to-white screen for inserts when transforming into an α-complementing *E. coli* host, such as TB1 or JM107, and plating on Xgal plates (Yanisch-Perron et al., 1985; UNIT 1.4). Because the vectors are lethal to *E. coli* when *p_{tac}* is induced, screening for inserts cannot be done with IPTG in the transformation plates. The screen can be accomplished by picking colonies with a sterile toothpick or by replica plating onto an LB/ampicillin master plate and an LB/ampicillin/IPTG/Xgal indicator plate.

Choice of MBP vector

If little is known about the protein of interest, the best strategy is to construct fusions in both vectors and empirically determine which is best. However, some information about the protein, if available, may help guide the choice of vector. The vector lacking the signal sequence, pMAL-c2, produces a cytoplasmic fu-

sion protein and has several advantages. Most importantly, it generally gives at least a four-fold higher yield than pMAL-p2. In addition, unstable proteins are often more stable when expressed from pMAL-c2 than when expressed from pMAL-p2, suggesting that there is more protease activity in the periplasm. Moreover, cytoplasmic expression can sometimes circumvent the instability and folding problems caused by the attempt to export cytoplasmic proteins (Bedouelle and Duplay, 1988; Gentz et al., 1988).

However, there are cases where use of the signal sequence vector, pMAL-p2, may prove advantageous. Some foreign proteins cannot fold correctly when expressed in the *E. coli* cytoplasm, and are produced in an inactive or unstable form. Secretion into the periplasm has been used to promote proper folding of such proteins (Hsiung et al., 1986; Lauritzen et al., 1991). In particular, the *E. coli* cytoplasm is a reducing environment, so disulfide bonds do not usually form until the cells are lysed to make a crude extract. This may allow abnormal disulfide bonds to form (e.g., interchain rather than intrachain). *E. coli* proteins do, however, form disulfide bonds in the periplasm, and in fact an accessory protein that catalyzes this process in *E. coli* has been described (Bardwell et al., 1991). Thus, for some foreign proteins, secretion into the periplasm can provide a protein-folding pathway that more closely mimics the native environment of the protein of interest. For these reasons, pMAL-p2 may be the vector of choice for proteins that are secreted in their native host or have disulfide bonds. In addition, in spite of the lower yield of fusion protein when expressed from pMAL-p2 relative to pMAL-c2, the number of different *E. coli* proteins in the periplasm is much lower than in the cytoplasm. This makes export to the periplasm a purification step, and the number of different proteins present in trace amounts in the affinity-purified protein is probably lower for fusion proteins purified from the periplasm.

Separation of the protein from MBP

After purification of the fusion protein, the MBP domain can be separated from the protein of interest by cleavage with factor Xa protease (UNIT 16.4B; Nagai and Thøgersen, 1984, 1987). About 50% of the time, the cleavage reaction works on the native fusion protein as it is eluted from the column. The remaining 50% of the cases break down into two categories: fusion proteins that do not cut with factor Xa and

fusion proteins that are cleaved at secondary sites within the protein of interest (see Troubleshooting). Many fusion proteins that do not cut with factor Xa after elution from the amylose column can be cleaved after denaturation and removal of the denaturant (UNIT 16.4B).

Once the fusion protein has been cleaved with factor Xa, the two domains can be separated chromatographically. Sepharose chromatography has an advantage over the affinity method, because it provides for the removal of factor Xa and any other trace contaminants. For Sepharose and most other chromatography resins, there is an equivalent HPLC method that increases the resolution of this step (UNIT 10.13).

Troubleshooting

One problem that is often encountered when expressing foreign proteins in *E. coli* (including MBP fusion proteins) is that the expressed protein is unstable. In these cases, conditions for expression and crude extract preparation can be optimized in a number of ways. The time allowed for expression can be shortened to 1 hr, which usually decreases the yield by half but can lead to a higher proportion of full-length fusion protein. In addition, the stability of a particular fusion protein can show large differences among different strain backgrounds. Apparently, different laboratory strains of *E. coli* differ a good deal in the levels and types of proteases that are produced.

If the fusion protein is being expressed cytoplasmically (in pMAL-c2), the use of *E. coli* protease mutants such as *lon* and *htrP* can reduce the degree of degradation (Baker et al., 1984; I. Hall and P. Riggs, unpub. observ.). In addition, some cytoplasmically expressed proteins are stabilized by using a host deficient in the chaperonin *DnaJ* (Straus et al., 1988; Reidhaar-Olson et al., 1990). It is possible that fusions that are exported to the periplasm might be stabilized by using a *degP* mutant (Strauch and Beckwith, 1988) or a *tsp* mutant (Silber et al., 1992). These mutations also could lead to stabilization of a cytoplasmically expressed fusion protein because such proteins are exposed to periplasmic proteases upon cell lysis. Similarly, fusion proteins being prepared from either the cytoplasm or the periplasm might be stabilized by using an *ompT* mutant, which lacks an outer membrane protease that has been shown to cleave prokaryotic and eukaryotic proteins during the preparation of crude extracts (Grodberg and Dunn, 1988; Sugimura and Higashi, 1988). Often, most of the degradation of MBP fusion proteins takes place in

vivo or during harvesting the cells and preparing the crude extract; once the crude extract is made, most proteins are fairly stable. However, if a fusion protein is unstable in the crude extract, the addition of protease inhibitors could help. One reason for the EDTA in the column buffer is to chelate calcium and inhibit the activity of calcium-dependent proteases once the cells are lysed.

A second problem sometimes seen when expressing MBP fusion proteins is that high levels of fusion protein can lead to the formation of insoluble aggregates. If a fusion protein is insoluble, lowering the temperature of expression and/or the level of induction of p_{tac} sometimes yields soluble fusion protein (e.g., induction with 10 to 100 μ M IPTG at $\leq 30^\circ\text{C}$; Tagaki et al., 1988; Schein, 1989). Similarly, switching to a different host strain can result in large differences in the proportion of a particular fusion protein that is soluble (M. Southworth, S. Levitt, and F. Perler, unpub. observ.). However, there are cases when no conditions can be found that result in soluble fusion protein expression. In these cases, the fusion protein can usually be solubilized by dissolving it in 6 M guanidine-HCl or 8 M urea, and then dialyzing away the denaturant. The fusion protein can then be applied to the amylose column for affinity purification. Alternatively, the denatured protein can be purified by conventional means (Nagai and Thøgersen, 1987). One advantage to the latter approach is that insoluble protein is usually very stable and can be purified away from proteases before it is refolded.

A third problem seen with MBP fusion proteins is failure to cleave with factor Xa. There are two possible explanations for this failure. The first involves the primary sequence of the cleavage site. Although factor Xa has a low degree of discrimination for the residue that follows the arginine in its recognition site, it will not cleave if that residue is a proline. The second explanation involves the three-dimensional conformation of the fusion protein. For some fusions, the factor Xa site can be blocked by the protein of interest. In this case, cleavage can often be obtained by denaturing the fusion protein to render the site accessible. However, some proteins remain resistant to cleavage, presumably because the protein quickly refolds into a factor Xa-resistant conformation upon removal of the denaturant. Cleavage can sometimes be facilitated by inserting a spacer between the factor Xa site and the protein of interest, but this results in extra residues (en-

coded by the spacer) attached to the amino-terminus of the protein of interest.

Another problem that sometimes arises during factor Xa cleavage is proteolysis at secondary sites within the protein of interest. This problem often (but not always) coincides with instability in *E. coli* during expression of the fusion protein (P. Riggs and I. Hall, unpub. observ.). Cleavage at arginines that are not in the factor Xa site has been observed (Nagai and Thøgersen, 1987; Lauritzen et al., 1991). Because there are arginine residues in other positions that are not cleaved, it is thought that the three-dimensional structure of the protein is important in determining these secondary cleavage sites. Cleavage at lysine residues has also been observed, and can be blocked by reversible acylation of lysine residues at the ϵ -amino group (Wearne, 1990).

Anticipated Results

Fusion protein expressed from pMAL-c2 usually constitutes 20% to 40% of the total cellular protein. Fusion protein expressed from pMAL-p2 usually constitutes 5% to 10% of the total cellular protein. In both cases, a band corresponding to the fusion protein can often be seen by running small samples of uninduced and induced cells on an SDS-polyacrylamide gel (e.g., 400 μ l of cells at $\text{OD}_{600} = 0.5$; first support protocol). The yield of fusion protein from the affinity purification ranges from ≤ 1 mg to 100 mg/liter culture at $\text{OD}_{600} = 1.0$; the remaining fusion protein is present in the column flow-through. The yield varies greatly depending upon the sequences fused to *malE*. In cases where the yield has been compared directly, pMAL-c2 (no signal sequence) gives from 4- to 15-fold more protein in the affinity purification than pMAL-p2.

Approximately 70% of the time, a fusion protein is reasonably stable and will yield ≥ 1 mg/liter culture in the affinity purification. When the fusion protein is insoluble or unstable, changing the conditions of expression has helped $\sim 5\%$ of the time. The rest of the time, affinity purification yields little or no fusion protein. In most of these cases, the fusion protein does not bind well to the column, perhaps because the target protein either blocks or distorts the binding site of MBP. In this situation, it is probably best to purify large amounts of fusion protein by conventional means (*UNITS 10.9-10.16*) or to try a different expression system.

Time Considerations

Growth of the cells, expression of the fusion protein, and harvest of the cells takes 6 to 7 hr. After harvest, the cells resuspended in column buffer can be stored at -20°C for ≥ 1 month. The thawed cell suspension should be stored for as short a time as possible, but once the crude extract is made and diluted 1:5 with column buffer, the fusion protein is usually reasonably stable at 4°C . After sonication and centrifugation of the cell debris, the crude extract (or the diluted crude extract) can again be frozen at -20°C for ≥ 1 month. Affinity chromatography takes ~ 9 hr and should be done in one day, if at all possible; if the fusion protein is left on the column too long, it starts to lose affinity for the amylose matrix. This may be due to maltose that is released from the column by trace amounts of amylase that are purified from *E. coli* along with the fusion protein. The purified protein can be stored for short periods of time (1 day) at 4°C , or at -20°C for long-term storage. Many proteins denature after freezing and thawing, so freezing in small aliquots or with the addition of 50% glycerol is recommended.

Literature Cited

- Amann, E. and Brosius, J. 1985. 'ATG vectors' for regulated high-level expression of cloned genes in *Escherichia coli*. *Gene* 40:183-190.
- Baker, T.A., Grossman, D., and Gross, C.A. 1984. A gene regulating the heat shock response in *Escherichia coli* also affects proteolysis. *Proc. Natl. Acad. Sci. U.S.A.* 81:6779-6783.
- Bardwell, J.C.A., McGovern, K., and Beckwith, J. 1991. Identification of a protein required for disulfide bond formation in vivo. *Cell* 67:581-589.
- Bedouelle, H. and Duplay, P. 1988. Production in *Escherichia coli* and one-step purification of bifunctional hybrid proteins which bind maltose. *Eur. J. Biochem.* 171:541-549.
- Duplay, P., Bedouelle, H., Fowler, A., Zabin, I., Saurin, W., and Hofnung, M. 1984. Sequences of the *malE* gene and of its product, the maltose-binding protein of *Escherichia coli* K12. *J. Biol. Chem.* 259:10606-10613.
- Gentz, R., Kuys, Y., Zwieb, C., Taatjes, D., Taatjes, H., Bannwarth, W., Stueber, D., and Ibrahim, I. 1988. Association of degradation and secretion of three chimeric polypeptides in *Escherichia coli*. *J. Bacteriol.* 170:2212-2220.
- Grodberg, J. and Dunn, J.J. 1988. *ompT* encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *J. Bacteriol.* 170:1245-1253.
- Guan, C., Li, P., Riggs, P.D., and Inouye, H. 1987. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* 67:21-30.
- Hsiung, H.M., Mayne, N.G., and Becker, G.W. 1986. High level expression, efficient secretion, and folding of human growth hormone in *Escherichia coli*. *Bio/Technology* 4:991-995.
- Kellerman, O.K. and Ferenci, T. 1982. Maltose binding protein from *E. coli*. *Methods Enzymol.* 90:459-463.
- Lauritzen, C., Tüchsen, E., Hansen, P.E., and Skovgaard, O. 1991. BPTI and N-terminal extended analogues generated by Factor Xa cleavage and cathepsin C trimming of a fusion protein expressed in *Escherichia coli*. *Prot. Expr. Purif.* 2:372-378.
- Maina, C.V., Riggs, P.D., Grandea, A.G. III, Slatko, B.E., Moran, L.S., Tagliamonte, J.A., McReynolds, L.A., and Guan, C. 1988. A vector to express and purify foreign proteins in *Escherichia coli* by fusion to, and separation from, maltose binding protein. *Gene* 74:365-373.
- Nagai, K. and Thøgersen, H.C. 1984. Generation of β -globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli*. *Nature* 309:810-812.
- Nagai, K. and Thøgersen, H.C. 1987. Synthesis and sequence-specific proteolysis of hybrid proteins produced in *Escherichia coli*. *Methods Enzymol.* 153:461-481.
- Neu, H.C. and Heppel, L.A. 1965. The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts. *J. Biol. Chem.* 240:3685-3692.
- Reidhaar-Olson, J.F., Parsell, D.A., and Sauer, R.T. 1990. An essential proline in λ repressor is required for resistance to intracellular proteolysis. *Biochemistry* 29:7563-7571.
- Schein, C.H. 1989. Production of soluble recombinant proteins in bacteria. *Bio/Technology* 7:1141-1149.
- Silber, K.R., Keiler, K.C., and Sauer, R.T. 1992. Tsp: A tail-specific protease that selectively degrades proteins with nonpolar C termini. *Proc. Natl. Acad. Sci. U.S.A.* 89:295-299.
- Silhavy, T.J., Szmelcman, S., Boos, W., and Schwartz, M. 1975. On the significance of the retention of ligand by protein. *Proc. Natl. Acad. Sci. U.S.A.* 72:2120-2124.
- Strauch, K.L. and Beckwith, J. 1988. An *Escherichia coli* mutation preventing degradation of abnormal periplasmic proteins. *Proc. Natl. Acad. Sci. U.S.A.* 85:1576-1580.
- Straus, D.B., Walter, W.A., and Gross, C.A. 1988. *Escherichia coli* heat shock gene mutants are defective in proteolysis. *Genes Dev.* 2:1851-1858.
- Sugimura, K. and Higashi, N. 1988. A novel outer-membrane-associated protease in *Escherichia coli*. *J. Bacteriol.* 170:3650-3654.

Takagi, H., Morinaga, Y., Tsuchiya, M., Ikemura, G., and Inouye, M. 1988. Control of folding of proteins secreted by a high expression secretion vector, pIN-III-ompA: 16-fold increase in production of active subtilisin E in *Escherichia coli*. *Bio/Technology* 6:948-950.

Wearne, S.J. 1990. Factor Xa cleavage of fusion proteins: Elimination of nonspecific cleavage by reversible acylation. *FEBS Lett.* 263:23-26.

Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103-119.

Zagursky, R.J. and Berman, M.L. 1984. Cloning vectors that yield high levels of single stranded DNA for rapid DNA sequencing. *Gene* 27:183-191.

Contributed by Paul Riggs
New England Biolabs
Beverly, Massachusetts

Expression and Purification of Glutathione-S-Transferase Fusion Proteins

UNIT 16.7

**BASIC
PROTOCOL**

This unit describes how pGEX vectors can be used in bacterial systems to express foreign polypeptides as fusions with glutathione-S-transferase (GST). In general, such fusion proteins are soluble and are easily purified from lysed cells under nondenaturing conditions by absorption with glutathione-agarose beads, followed by elution in the presence of free glutathione. Potential applications of the pGEX vectors include the expression and purification of individual polypeptides (including short peptides) for use as immunogens and as biochemical and biological reagents, and in the construction of cDNA expression libraries.

The screening procedure given here has the advantage (over restriction analysis of plasmid DNA, for example) of rapidly identifying transformants that express fusion proteins, and also revealing whether such fusion proteins are readily purified by the affinity absorption step. If true transformants are likely to be rare because of a high background with the vector alone, screening is better conducted by DNA hybridization to lysed colonies (UNITS 6.3 & 6.4) or by immunological screening (UNITS 6.7 & 6.8).

This protocol describes production and screening of pGEX transformants and purification of milligram quantities of fusion proteins from 1-liter cultures. UNIT 16.4B describes removal of the GST portion of a fusion protein by cleavage with site-specific proteases. Because the vagaries of heterologous protein expression in *Escherichia coli* are such that the basic protocol for purifying fusion proteins may not always suffice, Critical Parameters and Troubleshooting describe several modifications to the expression and purification protocol that may be useful in cases where fusion proteins are insoluble or unstable.

Each pGEX vector contains an open reading frame encoding GST, followed by unique restriction endonuclease sites for *Bam*HI, *Sma*I, and *Eco*RI, followed in turn by termination codons in all three frames (Fig. 16.7.1). The cloning sites are present in a different reading frame in each of the three vectors, so the vector in which the foreign polypeptide will be expressed in-frame with GST must be chosen first. pGEX2T or pGEX3X should be used if the GST carrier is eventually to be removed by site-specific proteolysis (UNIT 16.4B). Fusion proteins produced using pGEX1 can be cleaved by chemical hydrolysis at low pH (UNIT 16.4B).

Materials

For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

- pGEX vector (pGEX1 from Amrad or pGEX2T and pGEX3X from Pharmacia Biotech)
- Transformation-competent *Escherichia coli* (UNIT 1.8)
- LB plates containing 50 µg/ml ampicillin (UNIT 1.1)
- LB medium containing 10 µg/ml ampicillin (UNIT 1.1)
- 100 mM isopropyl-1-thio-β-D-galactoside (IPTG), filter sterilized
- Phosphate-buffered saline (PBS; APPENDIX 2), ice-cold
- Glutathione-agarose bead slurry (see recipe)
- 2× SDS sample buffer (UNIT 10.2)
- 10% (v/v) Triton X-100
- 50 mM Tris·Cl (pH 8.0)/5 mM reduced glutathione (freshly prepared; pH 7.5, final)
- Glycerol
- 37°C shaking incubator
- Beckman JA-10 and JA-20 rotors (or equivalents)

**Protein
Expression**

16.7.1

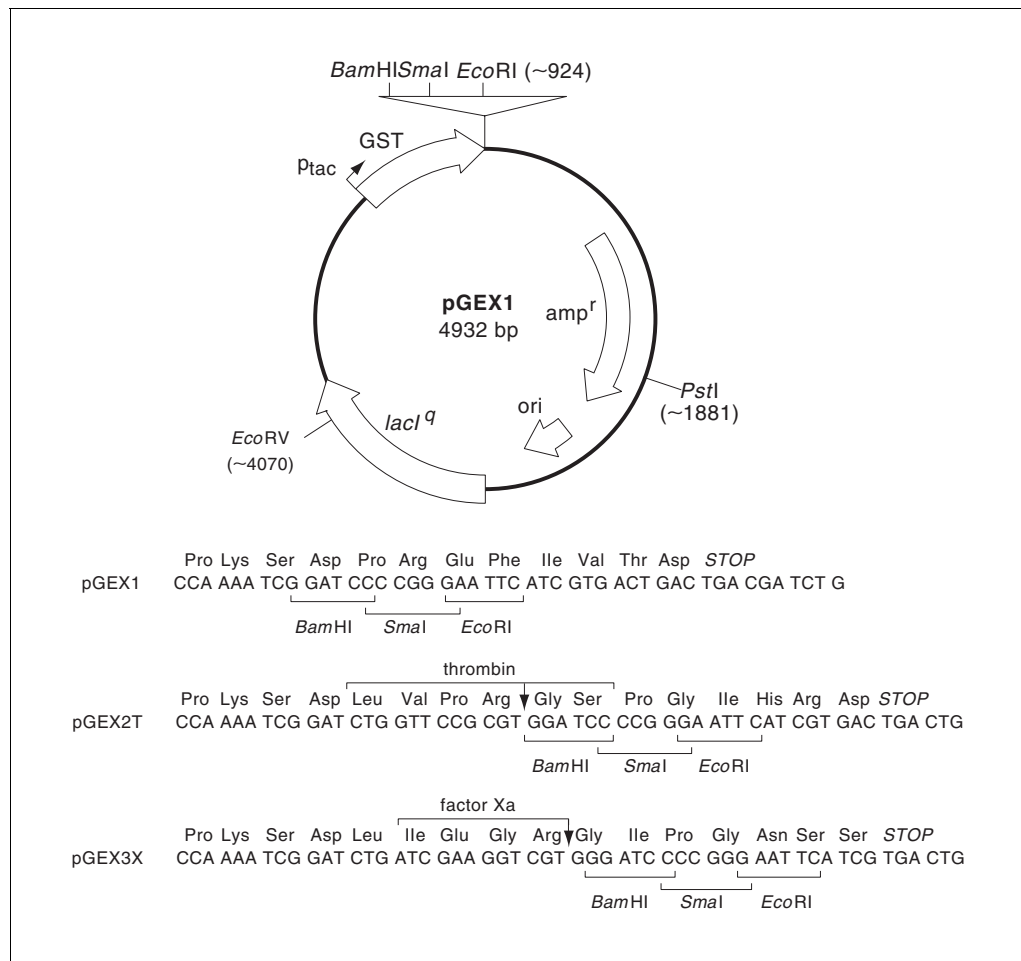


Figure 16.7.1 pGEX1. pGEX1 is a fusion-protein expression vector that expresses a cloned gene as a fusion protein to GST. The *lac* repressor (product of the *lacI* gene) binds to the p_{tac} promoter, repressing the expression of GST fusion protein. Upon induction with IPTG, derepression occurs and GST fusion protein is expressed. The gene of interest can be inserted into the polylinker located at the end of the GST gene. The polylinker sequences are shown below the map of pGEX1, where the restriction endonuclease cloning sites are bracketed. The polylinker of pGEX2T and pGEX3X contains protease cleavage sites so the cloned protein can be released from the GST moiety (or “carrier”). The recognition sequences for thrombin and factor Xa are bracketed above the polylinker sequences, with the actual cleavage site between Arg and Gly. The nucleotide sequence of pGEX1 is available under Genbank accession number M21676.

Probe sonicator (with 2- and 5-mm-diameter probes)

Additional reagents and equipment for subcloning DNA fragments (UNITS 1.4 & 3.16), SDS-PAGE (UNIT 10.2), and Coomassie blue staining (UNIT 10.6)

Produce and analyze transformants

1. Subclone the chosen DNA fragment into the appropriate pGEX vector in the correct reading frame, transform competent *E. coli* cells, and select transformants on LB/ampicillin plates. Incubate plates 12 to 15 hr at 37°C.

*Include a control of vector ligated to itself in the absence of insert DNA. Unless the insert DNA is being cloned with ends generated by digestion with two different restriction endonucleases, treat the vector with phosphatase after digestion to minimize screening. Any standard *E. coli* strain should be suitable for this transformation because the *lacI^q* repressor allele that controls expression of the fusion protein via the *tac* promoter (a *trp/lac* promoter fusion) is present on each pGEX vector.*

- Pick transformant colonies into 2 ml LB/ampicillin medium and streak out onto a master LB/ampicillin plate. Inoculate a control tube with bacteria transformed with the parental pGEX vector. Incubate the master plate 12 to 15 hr at 37°C. Grow liquid cultures with vigorous agitation in a 37°C shaking incubator until visibly turbid (3 to 5 hr).

The number of transformants to be screened can be judged from the number over background obtained with the self-ligated vector alone.

- Induce fusion protein expression by adding 100 mM IPTG to 0.1 mM. Continue incubation another 1 to 2 hr.
- Transfer liquid cultures to labeled microcentrifuge tubes, microcentrifuge 5 sec at maximum speed, room temperature, and discard supernatants. Resuspend pellets in 300 μ l ice-cold PBS. Remove 10 μ l to labeled tubes (for use in step 7).

Except where noted, keep all samples and tubes on ice.

- Lyse cells using a probe sonicator with a 2-mm-diameter probe. Microcentrifuge 5 min at maximum speed, 4°C, to remove insoluble material. Transfer supernatants to fresh tubes.

Lysis is complete when the cloudy cell suspension becomes translucent. The frequency and intensity of sonication should be adjusted so no frothing occurs and so lysis is complete in 10 sec. For a Tekmar Sonic Disruptor with a microtip probe, the output frequency should be set to the microtip limit noted on the dial. An alternative to sonication that may be useful when large numbers of samples have to be processed is lysis by repeated cycles of freezing in dry ice followed by thawing at 20°C. However, fusion proteins are more likely to suffer degradation with the freeze-thaw method.

- Add 50 μ l of 50% slurry of glutathione-agarose beads to each supernatant and mix gently \geq 2 min at room temperature. Add 1 ml PBS, vortex briefly, microcentrifuge 5 sec at maximum speed, room temperature to collect beads, and discard supernatants. Repeat the PBS wash twice.
- Add an equal volume of 1 \times SDS sample buffer to the washed beads, and 30 μ l to the 10- μ l samples of resuspended whole cells (from step 4). Heat 3 min at 100°C, vortex briefly, and load onto a 10% SDS-polyacrylamide gel. Run the gel for the appropriate time and stain with Coomassie blue solution to visualize the parental GST (made in control cells carrying a pGEX vector) and the fusion protein (Fig. 16.7.2).

Ideally, transformants expressing the desired fusion protein will be identified by the absence from total cellular proteins of the 27.5-kDa molecule specified by parental pGEX vectors, and by the presence of a novel, larger species. Where two orientations of the insert DNA are possible, the relative mobilities of the new species should distinguish between them; otherwise, conventional restriction endonuclease analysis will be required (UNIT 3.1). If the fusion protein has absorbed to the glutathione-agarose beads, proceed to large-scale purification as in steps 8 to 15. If, on the other hand, the fusion protein is absent from the purified material, it may be insoluble; see Critical Parameters and Troubleshooting for a discussion of this problem and for instances where neither the 27.5-kDa protein nor a larger fusion protein is visible (perhaps indicating toxicity or instability of the fusion protein). Interpretation is sometimes complicated when unstable fusion proteins break down and release the 27.5-kDa GST moiety. Such cases are usually recognized by the reduced level of the 27.5-kDa species, and by the series of larger, partial proteolytic fragments above it.

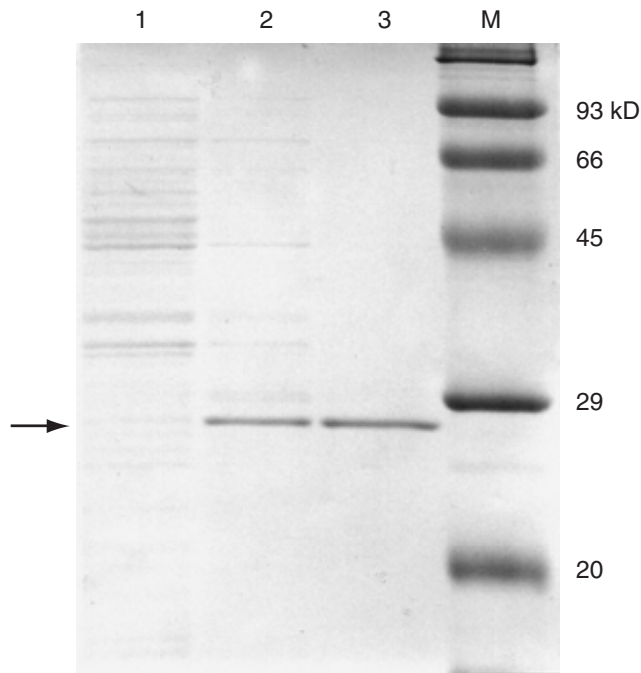


Figure 16.7.2 Stained protein gel showing expression of GST. The total cell lysate of a pGEX1 transformant is shown before IPTG induction (lane 1) and after IPTG induction (lane 2). The GST protein after elution from glutathione-agarose beads is shown at the position indicated by the arrow (lane 3). The lane marked M contains molecular-weight markers for sizes indicated.

Purify fusion proteins on a large scale

8. Inoculate a colony of the pGEX transformant into 100 ml LB/ampicillin medium and grow 12 to 15 hr at 37°C in a shaking incubator.
9. Dilute this culture 1:10 into 1 liter fresh LB/ampicillin medium, split between two 2-liter flasks, and grow 1 hr at 37°C.
10. Add 100 mM IPTG to 0.1 mM (final) and continue incubation an additional 3 to 7 hr.

By optimizing growth conditions, the yield of fusion protein may be greatly improved. Investigate the effects of delaying the addition of IPTG and altering the induction period. For example, if the fusion protein is not detrimental to the growth of the bacteria, and the protein is stable, the yield may be improved by increasing the duration of the induction period. However, if the protein is detrimental to the cells, the yield may be improved by delaying the addition of IPTG until the culture is more dense (i.e., when more cells are present) and also by decreasing the duration of induction, to decrease cell loss during this period.

11. Centrifuge 10 min at 5000 × g (~5500 rpm in a Beckman JA-10 rotor), room temperature, to collect cells. Discard supernatant and resuspend pellet in 10 to 20 ml ice-cold PBS.
12. Immerse the tube in ice and lyse cells using a probe sonicator with a 5-mm-diameter probe.

Lysis can be detected as a change in color from a rich straw brown (intact cells) to a dull gray-brown (lysed cells). Adjust the frequency and intensity of sonication so lysis occurs in ~30 sec, without frothing. Excessive sonication can result in contamination of purified fusion protein with other proteins. Lysis by freeze-thaw is a slower alternative.

13. Add 10% Triton X-100 to 1% (final) and mix. Centrifuge 5 min at 10,000 × g (~9500 rpm in a Beckman JA-20 rotor), 4°C, to remove insoluble material and intact cells. Alternatively, it may be convenient to microcentrifuge 1.5-ml aliquots for 5 min at top speed, 4°C. Collect supernatants (carefully avoiding the pellets) and pool them.

Triton X-100 is added to the lysed cells to minimize association of fusion protein with

bacterial proteins, and thus to prevent appearance of these contaminants in the final preparation.

14. Add supernatant to 1 ml of 50% slurry of glutathione-agarose beads and mix gently ≥ 2 min at room temperature. Wash by adding 50 ml ice-cold PBS, mixing, and centrifuging 10 sec in a tabletop centrifuge at $500 \times g$, room temperature. Repeat the wash two more times. Resuspend the beads in a small volume (1 to 2 ml) of ice-cold PBS and transfer to a 1.5-ml microcentrifuge tube.

The capacity of glutathione-agarose is ≥ 8 mg protein/ml swollen beads.

To remove the GST carrier by protease cleavage, see UNIT 16.4B.

15. Centrifuge 10 sec at $500 \times g$, room temperature, to collect beads; discard supernatant. Elute fusion protein by adding 1 ml of 50 mM Tris-Cl (pH 8.0)/5 mM reduced glutathione. Mix gently 2 min, centrifuge 10 sec at $500 \times g$, and collect supernatant. Repeat elution two to three times and analyze each fraction by SDS-PAGE. Store eluted protein in aliquots containing 10% glycerol at -70°C .

The majority of fusion protein should appear in the first two elutions. If a concentrated solution of the fusion protein is required, elution can be carried out while the beads are immobilized in a small column. See Critical Parameters and Troubleshooting regarding problems with contamination, instability, or yield. The concentration of glutathione in the elution buffer may be increased to 10 or 15 mM if the fusion protein is not quantitatively eluted using the lower concentration. Determine the yield of fusion protein by measuring the absorbance at 280 nm. For the GST carrier, $A_{280} = 1$ corresponds to a protein concentration of 0.5 mg/ml.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Glutathione-agarose bead slurry

Preswell S-linkage glutathione-agarose beads (Sigma or Pharmacia Biotech) 1 hr in 10 vol phosphate-buffered saline (PBS, APPENDIX 2). Wash twice with PBS and store as a 50% (v/v) slurry ≤ 1 month at 4°C .

The beads can be recycled by boiling 5 min in PBS containing 1% SDS (Frangioni and Neel, 1993), but should then only be used to purify the same fusion protein to prevent cross-contamination.

COMMENTARY

Background Information

The pGEX vectors are designed so that foreign polypeptides can be expressed in *Escherichia coli* in a form that allows them to be purified rapidly under nondenaturing conditions (Smith and Johnson, 1988). Foreign polypeptides are expressed as fusions to the C terminus of glutathione-S-transferase (GST), a common 26-kDa cytoplasmic protein of eukaryotes. The GST gene used to generate the pGEX vectors was originally cloned from the parasitic helminth, *Schistosoma japonicum* (Smith et al., 1986).

The fusion proteins typically remain soluble within the bacteria and can be purified from lysed cells because of the affinity of the GST

moiety for glutathione immobilized on agarose beads. Recovery of the fusion proteins is by elution with free reduced glutathione at neutral pH.

The main advantage of this system for expressing and recovering foreign proteins from *E. coli* is that most fusion proteins remain soluble; native proteins expressed in *E. coli* often denature and precipitate. Furthermore, denaturing conditions are not required at any stage during purification, and consequently, foreign polypeptides may retain their functional activities and antigenicity. Additional features are the efficiency and rapidity of purification, the high level of inducible expression achieved with the strong tac promoter, and the

broad range of suitable bacterial hosts (because the pGEX vectors carry the *lacI^q* repressor allele). Furthermore, the GST carrier is small compared to other carriers (e.g., β -galactosidase), and lacks the immunological complications of fusion proteins containing fragments of protein A. (Peptides containing protein A are likely to interact and/or combine with a wide variety of antibodies, but the GST moiety will not interfere with specific antibody-antigen interactions.)

The main determinant for successful purification of foreign polypeptides using the pGEX system is solubility of the fusion protein. To some extent, this can only be discovered empirically. More difficulties are encountered as the size of the desired fusion protein increases (particularly when >50 kDa), or when the protein contains regions that are strongly hydrophobic or highly charged (D.B.S. and L.M.C., unpub. observ.). Insoluble fusion proteins can sometimes be coaxed into solution (see Critical Parameters and Troubleshooting) or can otherwise be purified after solubilization in denaturing reagents.

If necessary, the GST moiety can be removed from fusion proteins by cleavage with site-specific proteases (UNIT 16.4B). Note however, that often the GST carrier does not compromise the antigenicity or functional activity of the foreign polypeptide. Modified versions of the original pGEX vectors have been produced that simplify cloning, cleavage or detection of fusion proteins (for review, see Smith, 1993).

Critical Parameters and Troubleshooting

Contamination of fusion proteins by host cell proteins is usually a sign that sonication has been too severe, perhaps because denaturation exposes regions on proteins that are more likely to cause aggregation. Some contaminating species may represent degraded fragments of the fusion protein that bind to glutathione. Such fragments are not easy to eliminate, except by increasing the stability of the fusion protein. The addition of protease inhibitors [e.g., 1% (w/v) aprotinin and 1 mM phenylmethylsulfonyl fluoride (PMSF)] may help in this regard, and inclusion of 50 to 100 mM EDTA in the lysis buffer has also been beneficial. The *E. coli* host strain can have a major and unpredictable effect on stability, and it is worth testing several different strains, including the protease-deficient *lon*⁻ strains (D.B.S. and L.M.C., unpub. observ.; UNIT 1.4). Alterna-

tively, degradation of fusion proteins can be minimized by adding isopropyl-1-thio- β -D-galactoside (IPTG) later in the course of the culture and keeping the induction period to a minimum. The overall yield of fusion protein can sometimes be improved by increasing the quantity of glutathione-agarose beads, minimizing the volume of liquid during absorption, and extending the period of absorption to 1 hr.

Insolubility of fusion proteins can be addressed by several means. In some cases, growth of cells at 30°C is sufficient to alter solubility (D.B.S., unpub. observ.), but in other cases it may be necessary to investigate the effect of mild detergent treatment after cell lysis. Examples of conditions under which binding of GST to glutathione-agarose is unaffected are 1% Triton X-100, 1% Tween-20, 1% CTAB, 10 mM DTT, and 0.03% SDS. Alternatively, fusion proteins can be solubilized in 1% to 2% Sarkosyl prior to sonication; they can be solubilized in 2% to 4% Triton X-100 prior to binding to glutathione agarose (Frangioni and Neel, 1993). Even if a fusion protein is largely insoluble, it may be possible to purify the small proportion that is still soluble, with a yield of perhaps 50 μ g/liter. Otherwise, a fusion protein that is abundant, but stubbornly insoluble, can be purified by gel filtration after solubilization under denaturing conditions or by electroelution from a gel after SDS-PAGE.

A different approach is to express the polypeptide as smaller fragments, particularly if it is possible to express parts of the protein that do not contain severely hydrophobic or highly charged regions. If the fusion protein is not of the expected size, it may be worth sequencing across the cloning site to ensure that the reading frame of the vector and the insert are matched. There is a hairpin in GST sequences near the cloning sites, so a sequencing primer complementary to either the extreme 3' end of the GST gene or to sequences within the DNA insert should be used.

Anticipated Results

Yields of fusion protein vary from more typical yields of 1 to 3 mg/liter up to 10 mg/liter, and can be as low as 50 μ g/liter if most of the fusion protein is insoluble. The single affinity chromatography step can generate fusion protein preparations that are >90% pure.

Time Considerations

A full day is required to screen colonies of transformants for expression of fusion proteins,

although continuous attention is not required. Large-scale purification and cleavage of fusion proteins require another day of intermittent work. Cells can be stored as pellets at -70°C before lysis, although with fusion proteins that are unstable, this may be undesirable. Purification is best completed in one session.

Literature Cited

- Frangioni, J.V. and Neel, B.G. 1993. Solubilization and purification of enzymatically active glutathione-S-transferase (pGEX) fusion proteins. *Anal. Biochem.* 210:179-187.
- Smith, D.B. 1993. Purification of glutathione-S-transferase fusion proteins. *Methods Mol. Cell Biol.* 4:220-229.
- Smith, D.B., Davern, K.M., Board, P.G., Tiu, W.U., Garcia, E.G., and Mitchell, G.F. 1986. Mr 26,000 antigen of *Schistosoma japonicum* recognized by resistant WEHI 129/J mice is a parasite glutathione S-transferase. *Proc. Natl. Acad. Sci. U.S.A.* 83:8703-8707.

Key Reference

- Smith et al., 1986. See above.
Original description of the pGEX system.
- Smith, 1993. See above.
Summary of modified pGEX vectors and alternative purification methods.
- Smith, D.B. and Johnson, K.S. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67:31-40.
First description of GST fusion system.

Contributed by Donald B. Smith
University of Edinburgh
Edinburgh, Scotland

Lynn M. Corcoran
Walter & Eliza Hall Institute
Victoria, Australia

Expression and Purification of Thioredoxin Fusion Proteins

This unit describes a gene fusion expression system that uses thioredoxin, the product of the *Escherichia coli trxA* gene, as the fusion partner. The system is particularly useful for high-level production of soluble fusion proteins in the *E. coli* cytoplasm; in many cases heterologous proteins produced as thioredoxin fusion proteins are correctly folded and display full biological activity. Although the thioredoxin gene fusion system is routinely used for protein production, high-level production of peptides—i.e., for use as antigens—is also possible because the prominent thioredoxin active-site loop is a very permissive site for the introduction of short amino acid sequences (10 to 30 residues in length). The inherent thermal stability of thioredoxin and its susceptibility to quantitative release from the *E. coli* cytoplasm by osmotic shock can also be exploited as useful tools for thioredoxin fusion protein purification. In addition, a more generic method for purification of any soluble thioredoxin fusion employs a modified form of thioredoxin (called “His-patch Trx”), which has been designed to bind to metal chelate resins. Protein fusions to His-patch Trx can usually be purified in a single step from cell lysates (see Strategic Planning).

The basic protocol outlines the construction of a fusion of *trxA* to any desired gene and expression of the fusion protein in an appropriate host strain at 37°C. Additional protocols describe *E. coli* cell lysis using a French pressure cell and fractionation (first support protocol), osmotic release of thioredoxin fusion proteins from the *E. coli* cytoplasm (second support protocol), and heat treatment to purify some thioredoxin fusion proteins (third support protocol).

STRATEGIC PLANNING

The thioredoxin gene fusion expression vectors pTRXFUS and hpTRXFUS, both of which carry the *E. coli trxA* gene (Fig. 16.8.1), are used for high-level production of C-terminal fusions to thioredoxin. The vector hpTRXFUS differs from pTRXFUS in that it contains a modified *E. coli trxA* gene which produces a mutant protein (“His-patch” thioredoxin) that can specifically bind to metal chelate matrices charged with nickel or cobalt, otherwise known as native metal-chelate affinity chromatography (MCAC; UNIT 10.11B). The *trxA* translation termination codon has been replaced in both vectors by DNA encoding a ten-residue peptide linker sequence that includes an enterokinase (enteropeptidase; LaVallie et al., 1993a) cleavage site. This highly specific site can be cleaved with enterokinase following purification of the fusion protein to release the protein of interest from its thioredoxin fusion partner (cleavage of the fusion protein is covered in UNIT 16.4B). Immediately downstream of the DNA encoding the enterokinase site in pTRXFUS and hpTRXFUS lies a DNA polylinker sequence containing a number of unique restriction endonuclease sites that can be used for forming in-frame translational fusions of any desired gene to *trxA*. Downstream of the DNA polylinker lies the *E. coli aspA* transcription terminator. Replication of these vectors is controlled by a modified colE1 replication origin similar to that found in pUC vectors (Norrander et al., 1983). Plasmid selection and maintenance is ensured by the presence of the β -lactamase gene on the vector. The vector pALtrxA-781 (Fig. 16.8.1) is very similar to pTRXFUS. However in this plasmid the *trxA* gene is followed by a translation termination codon, and the sequences encoding the enterokinase-site peptide linker are absent. A unique *RsrII* site, present in both pALtrxA-781 and pTRXFUS, allows for the easy insertion of short peptide-encoding DNA sequences into *trxA* within the region that encodes the active-site loop.

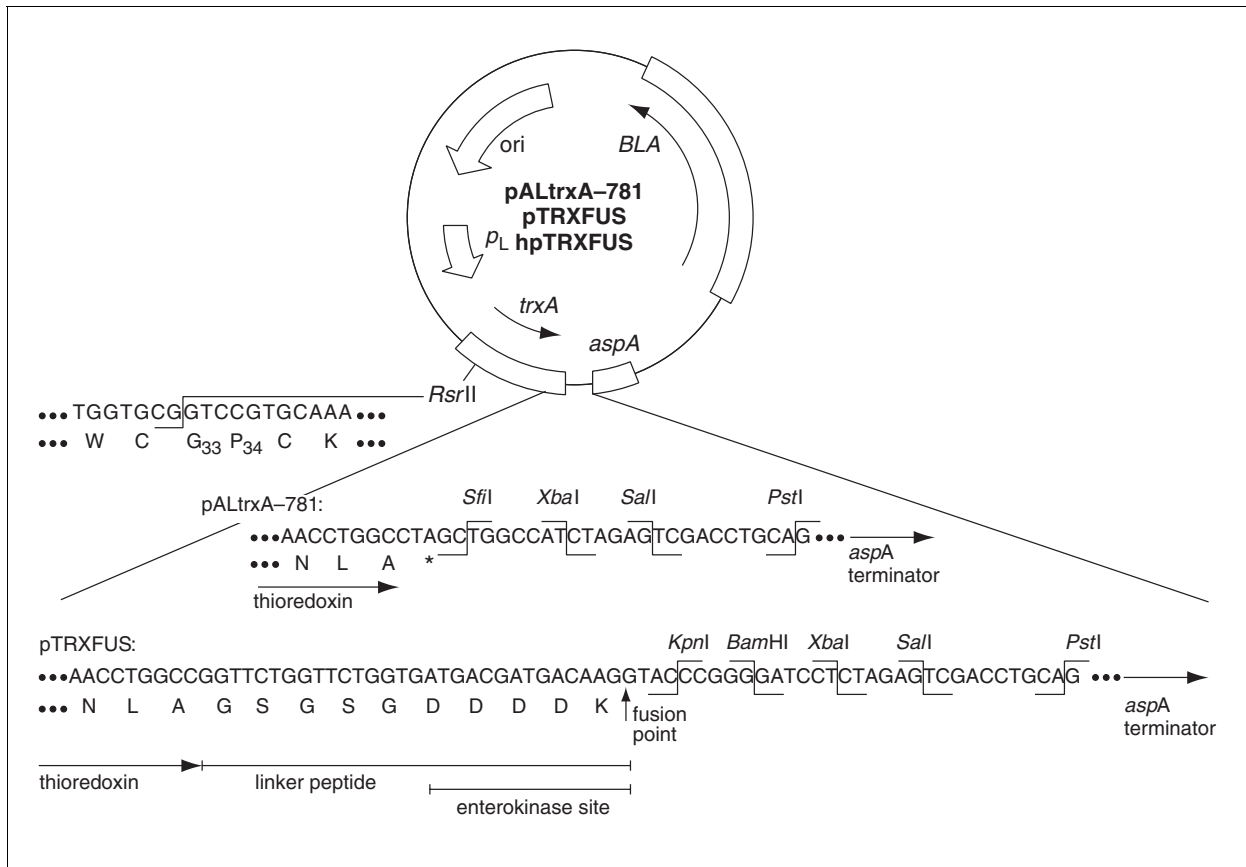


Figure 16.8.1 Thioredoxin gene fusion expression vectors pTRXFUS, hpTRXFUS, and pALtrxA-781. pALtrxA-781 contains a polylinker sequence at the 3' end of the *trxA* gene. pTRXFUS and hpTRXFUS contain a linker region encoding a peptide that includes the enterokinase cleavage site between the *trxA* gene and the polylinker. The sequence surrounding the active site loop of thioredoxin has a single *RsrII* site that can be used to insert peptide coding sequence. The asterisk indicates a translational stop codon. Abbreviations: *trxA*, *E. coli* thioredoxin gene; *BLA*, β -lactamase gene; *ori*, *colE1* replication origin; *pL*, bacteriophage λ major leftward promoter; *aspA* terminator, *E. coli* aspartate amino-transferase transcription terminator.

pTRXFUS, hpTRXFUS, and pALtrxA-781 carry the strong bacteriophage λ promoter *pL* (Shimatake and Rosenberg, 1981) positioned upstream of the *trxA* gene. Transcription initiation at the *pL* promoter is controlled by the intracellular concentration of λ repressor protein (*cI*). *UNIT 16.3* describes λ strains that carry either a temperature-sensitive form of *cI* (*cI*⁸⁵⁷) or a wild-type *cI* repressor protein. *cI*⁸⁵⁷-containing strains can be used for heat inductions of *pL* at 42°C; alternatively, in the strains carrying the wild-type repressor, *pL* can be induced by a prior induction of the *E. coli* SOS stress response. However, it is often desirable to express heterologous genes in *E. coli* at temperatures considerably lower than 42°C, or under conditions where cells are not undergoing a physiological stress. Strains GI698, GI724 and GI723 were designed to allow the growth and induction of *pL* expression vectors, including pTRXFUS, hpTRXFUS, and pALtrxA-781, under mild conditions over a wide range of temperatures (see Table 16.8.1; Mieschendahl et al., 1986). Each of these strains carries a wild-type allele of *cI* stably integrated into the *E. coli* chromosome at the nonessential *ampC* locus. A synthetic *trp* promoter integrated into *ampC* upstream of the *cI* gene in each strain directs the synthesis of *cI* repressor only when intracellular tryptophan levels are low. When tryptophan levels are high, synthesis of *cI* is switched off; therefore, the presence of tryptophan in the growth medium of GI698, GI723, or GI724 will block expression of λ repressor and thus will turn on *pL*. Because

Table 16.8.1 *E. coli* Strains for Production of Thioredoxin Fusion Proteins at Varying Temperatures

Strain	Desired production temperature (°C)	Pre-induction growth temperature (°C)	Induction period (hr)
GI698	15	25	20
GI698	20	25	18
GI698	25	25	10
GI724	30	30	6
GI724	37	30	4
GI723	37	37	5

the three strains carry ribosome-binding sequences of different strengths at the 5'-end of their respective *cI* genes, they maintain intracellular concentrations of λ repressor that increase in the order GI698 < GI724 < GI723. The choice of which strain to use for a particular application is dependent on the desired culture conditions as described below.

Although some thioredoxin fusion proteins produced at 37°C are insoluble, expression at lower temperatures can often result in the fusion protein being produced in a soluble form. Each of the three pL host strains GI698, GI723, and GI724 is suitable for the production of thioredoxin fusion proteins over a particular temperature range. Table 16.8.1 indicates the correct strain for expression of thioredoxin fusion proteins at any temperature between 15°C and 37°C. The induction protocol at any of these temperatures is the same as that described in the basic protocol for induction of GI724 at 37°C, except the preinduction growth temperature and the length of the induction period vary according to the strain used and the temperature chosen. Cultures should be grown at the indicated preinduction growth temperature until they reach a density of 0.4 to 0.6 OD₅₅₀/ml. They should then be moved to the desired induction temperature and induced by the addition of 100 µg/ml tryptophan.

Low-temperature inductions are best performed in strain GI698. However, this strain makes only enough *cI* repressor protein to maintain the vectors in an uninduced state at temperatures below 25°C. GI698 should therefore never be grown above 25°C when it carries a pL plasmid. A nonrefrigerated water bath can be maintained below room temperature by placing it in a 4°C room and setting the thermostat to the desired temperature.

It is often a good idea to collect timepoints during the course of a long induction period and to fractionate cells from these timepoints using the procedure in the first support protocol (steps 9 to 13). Although a particular fusion protein may be soluble during the early part of an induction, during the later phases of induction, it may become unstable or its concentration inside the cell may exceed a critical threshold above which it will precipitate and appear in the insoluble fraction.

**CONSTRUCTION AND EXPRESSION OF A THIOREDOXIN
FUSION PROTEIN**

This protocol describes construction and subsequent expression of a gene fusion between *trxA* (encoding thioredoxin) and a gene encoding a particular protein or peptide. After a clone carrying the correct fusion sequence is constructed, analyzed, and isolated, cultures are grown and expression is induced. The protocol is described in terms of the *E. coli* host strain GI724 with expression at 30°C; it may also be applied to strains GI698 and GI723 (also available from Genetics Institute) for expression at other temperatures by using the parameters specified in Table 16.8.1 (see Strategic Planning).

Materials

For recipes, see Reagents and Solutions in this unit (or cross-referenced unit); for common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

- DNA fragment encoding desired sequence
- Thioredoxin expression vectors (Fig. 16.8.1): pTRXFUS or pALtrxA-781 (Genetics Institute or Invitrogen) or hpTRXFUS (Genetics Institute)
- E. coli* strain GI724 (Genetics Institute or Invitrogen), grown in LB medium and made competent (UNIT 1.8)
- LB medium (UNIT 1.1)
- IMC plates containing 100 µg/ml ampicillin (see recipe)
- CAA/glycerol/ampicillin 100 medium (see recipe)
- IMC medium containing 100 µg/ml ampicillin (see recipe)
- 10 mg/ml tryptophan (see recipe)
- SDS-PAGE sample buffer (see recipe)
- 30°C convection incubator
- 18 × 50–mm culture tubes
- Roller drum (New Brunswick Scientific)
- 250-ml culture flask
- 70°C water bath
- Microcentrifuge, 4°C
- Additional reagents and equipment for subcloning of DNA fragments (UNIT 3.16), transforming competent *E. coli* cells (UNIT 1.8), preparing miniprep DNA (UNIT 1.6), restriction mapping (UNIT 3.2), direct sequencing of plasmid DNA (UNITS 7.3 & 7.4), SDS-PAGE (UNIT 10.2), and Coomassie brilliant blue staining (UNIT 10.6)

Construct the *trxA* gene fusion

1. Use DNA fragment encoding the desired sequence to construct either an in-frame fusion to the 3'-end of the *trxA* gene in pTRXFUS or hpTRXFUS, or a short peptide insertion into the unique *RsrII* site of pALtrxA-781.

A precise fusion of the desired gene to the enterokinase linker sequence in pTRXFUS or hpTRXFUS can be made by using the unique KpnI site trimmed to a blunt end with the Klenow fragment of E. coli DNA polymerase. The desired gene can usually be adapted to this blunt-end construct by using a synthetic oligonucleotide duplex ligated between it and any convenient downstream restriction site close to the 5' end of the gene. When designing the fusion junction, note that enterokinase is able to cleave —DDDDK↓X—, where X is any amino acid residue except proline. Synthetic oligonucleotides encoding short peptides for insertion into the thioredoxin active-site loop at the RsrII site will insert only in the desired orientation, because the RsrII sticky end consists of three bases.

2. Transform the ligation mixture containing the new thioredoxin fusion plasmid into competent GI724 cells. Plate transformed cells onto IMC plates containing 100 µg/ml

ampicillin to select transformants. Incubate plates in a 30°C convection incubator until colonies appear.

Strains GI698, GI723, and GI724 are all healthy prototrophs that can grow under a wide variety of growth conditions, including rich and minimal media and a broad range of growth temperatures (see Table 16.8.1). These strains can be prepared for transformation with pL-containing vectors by growing them in LB medium at 37°C. LB medium may also be used for these strains during the short period of outgrowth immediately following transformation. This growth period of 30 min to 1 hr is often used to express drug resistance phenotypes before plating out plasmid transformations onto solid medium. Subsequently, however, these strains should be grown only on minimal or tryptophan-free rich media, such as IMC medium containing 100 µg/ml ampicillin (for expression of the fusion protein) or CAA/glycerol/ampicillin 100 medium (for plasmid DNA preparations). Except during transformation, LB medium should never be used with these three strains when they carry pL plasmids because LB contains tryptophan. The pL promoter is extremely strong and should be maintained in an uninduced state until needed so that expression of the protein will not lead to selection of mutant or variant cells with lower expression due to undesirable genetic selections or rearrangements in the expression strain.

3. Grow candidate colonies in 5 ml CAA/glycerol/ampicillin 100 medium overnight at 30°C. Prepare minipreps of plasmid DNA and check for correct gene insertion into pTRXFUS by restriction mapping.
4. Sequence plasmid DNA of candidate clones to verify the junction region between thioredoxin and the gene or sequence of interest.

Induce expression

5. Streak out frozen stock culture of GI724 containing thioredoxin expression plasmid to single colonies on IMC plates containing 100 µg/ml ampicillin. Grow 20 hr at 30°C.

Occasionally there is induction of pL plasmids grown in GI698 and GI724 at 37°C, even in medium containing no tryptophan. Such induction appears to be a temperature-dependent phenomenon. If growth at 37°C prior to pL induction is essential, then GI723 should be used as the host strain because GI723 produces higher levels of cI repressor than both GI698 and to GI724. Otherwise, plasmid-containing GI698 should be grown at 25°C and plasmid-containing GI724 should be grown at 30°C prior to induction (see Table 16.8.1).

6. Pick a single fresh, well-isolated, colony from the plate and use it to inoculate 5 ml IMC medium containing 100 mg/ml ampicillin in an 18 × 150–mm culture tube. Incubate overnight at 30°C on a roller drum.
7. Add 0.5 ml overnight culture to 50 ml fresh IMC medium containing 100 µg/ml ampicillin in a 250-ml culture flask (1:100 dilution). Grow at 30°C with vigorous aeration until absorbance at 550 nm reaches 0.4 to 0.6 OD/ml (~3.5 hr).
8. Remove a 1-ml aliquot of the culture (uninduced cells). Measure the absorbance at 550 nm and harvest the cells by microcentrifuging 1 min at maximum speed, room temperature. Carefully remove all the spent medium with a pipet and store the cell pellet at –80°C.
9. Induce pL by adding 0.5 ml of 10 mg/ml tryptophan (100 µg/ml final) to remaining cells immediately.
10. Incubate 4 hr at 37°C. At hourly intervals during this incubation, remove 1-ml aliquots of the culture and harvest cells as in step 8.

11. Harvest the remaining cells from the culture 4 hr post-induction by centrifuging 10 min at 3000 rpm (e.g., in a Beckman J6 rotor), 4°C. Store the cell pellet at –80°C.

Procedures for further analysis of these cells are outlined in the support protocols.

Verify induction

12. Resuspend the pellets from the induction intervals (steps 8 and 10) in 200 µl of SDS-PAGE sample buffer/OD₅₅₀ cells. Heat 5 min at 70°C to completely lyse the cells and denature the proteins. Run the equivalent of 0.15 OD₅₅₀ cells per lane (30 µl) on an SDS-polyacrylamide gel.
13. Stain the gel 1 hr with Coomassie brilliant blue. Destain the gel and check for expression.

Most thioredoxin fusion proteins are produced at levels that vary from 5% to 20% of the total cell protein. The desired fusion protein should exhibit the following characteristics: it should run on the gel at the mobility expected for its molecular weight; it should be absent prior to induction; and it should gradually accumulate during induction, with maximum accumulation usually occurring 3 hr post-induction at 37°C.

SUPPORT PROTOCOL 1

E. COLI LYSIS USING A FRENCH PRESSURE CELL

A small 3.5-ml French pressure cell can be used as a convenient way to lyse *E. coli* cells. The whole-cell lysate can be fractionated into soluble and insoluble fractions by microcentrifugation. Other lysis procedures may be used—for example, sonication (UNITS 4.4 & 16.6) or treatment with lysozyme-EDTA (UNIT 4.4).

Additional Materials

*For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.*

Cell pellet from 4-hr post-induction culture (basic protocol)

20 mM Tris·Cl, pH 8.0 (APPENDIX 2), 4°C

Lysis buffer: 20 mM Tris·Cl (pH 8.0) with protease inhibitors (optional)—0.5 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM *p*-aminobenzamidine (PABA), and 5 mM EDTA

French press and 3.5-ml mini-cell (Fig. 16.8.2; SLM Instruments), 4°C

Lyse the cells

1. Resuspend cell pellet from 4-hr post-induction culture in 20 mM Tris·Cl, pH 8.0, to a concentration of 5 OD₅₅₀/ml.

Protease inhibitors can be included in the resuspension if desired. Cells can also be resuspended at densities of 100 OD₅₅₀/ml or greater; however, at high densities cell lysis may be less efficient.

2. Place 1.5 ml resuspended cell pellet in the French pressure cell. Hold the cell upside down with the base removed, the piston fully extended downwards, and the outlet valve handle that holds the nylon ball seal in the open position (loose).

Before filling the pressure cell, check that the nylon ball, which seals the outlet port and sits on the end of the outlet valve handle, is not deformed. If it is, replace it with a new one. Both the condition of the nylon ball and its seat in the pressure cell body are critical for the success of the procedure.

3. Bring the liquid in the pressure cell to the level of the outlet port by raising the piston slowly to expel excess air from the cell. With the outlet valve open and at the same time maintaining the piston in position, install the pressure cell base. Gently close the outlet valve.

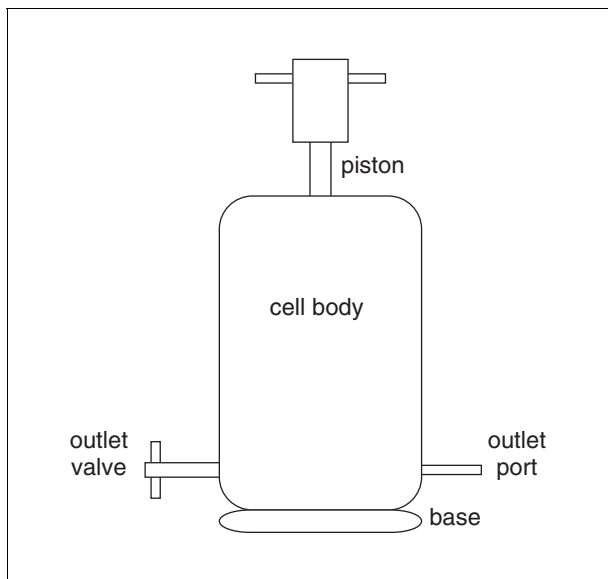


Figure 16.8.2 French pressure cell, equipped with 3.5-ml mini-cell.

CAUTION: Do not over-tighten the valve as this will deform the nylon ball and may irreparably damage its seat on the pressure cell body.

4. Turn the sealed cell right side up and place it in the hydraulic press.
5. Turn the pressure regulator on the press fully counter-clockwise to reset it to zero pressure. Set the ratio selector to medium. Turn on the press.

CAUTION: The larger (50-ml) pressure cell is usually used with the selector set on high. The small (3.5-ml) cell is only used on medium ratio.

6. Slowly turn the pressure regulator clockwise until the press just begins to move. Allow the press to compress the piston. It will stop moving after a few seconds.
7. Position a collection tube under the pressure cell outlet. Slowly increase the pressure in the cell by turning the pressure regulator clockwise. Monitor the reading on the gauge and increase the pressure to 1000 on the dial, corresponding to an internal cell pressure of 20,000 lb/in².
8. While continuously monitoring the gauge, very slowly open the outlet valve until lysate begins to trickle from the outlet.

The lysate should flow slowly and smoothly, and the cell pressure should not drop more than 100 divisions on the dial.

At 20,000 lb/in² and 5 OD₅₅₀/ml, cell lysis will be complete after one passage through the press. Lower pressures and/or higher cell densities may require a second passage.

Fractionate the lysate

9. Remove a 100- μ l aliquot of the lysate and freeze at -80°C (whole-cell lysate).
10. Fractionate the remainder of the lysate by microcentrifuging 10 min at maximum speed, 4°C .
11. Remove a 100- μ l aliquot of the supernatant and freeze at -80°C (soluble fraction). Discard the remainder of the supernatant.

Because this is a pilot experiment, it would not produce enough material to warrant saving any remaining supernatant.

12. Resuspend the pellet in an equivalent volume of lysis buffer. Remove a 100- μ l aliquot and freeze at -80°C (insoluble fraction).
13. Lyophilize the 100- μ l aliquots to dryness in a Speedvac evaporator. Solubilize in 100 μ l SDS-PAGE sample buffer. Analyze 30- μ l samples by SDS-PAGE.

This crude fractionation provides a fairly reliable indication of whether a protein has folded correctly. Usually proteins in the soluble fraction have adopted a correct conformation and proteins in the insoluble fraction have not. However, occasionally proteins found in the soluble fraction are not truly soluble; instead they form aggregates that do not pellet in the microcentrifuge. Conversely, sometimes a protein found in the insoluble fraction may be there because it has an affinity for cell wall components and cell membranes, and it may not be intrinsically insoluble. Occasionally proteins can be recovered from these insoluble fractions by extracting with agents such as mild detergents.

SUPPORT PROTOCOL 2

OSMOTIC RELEASE OF THIOREDOXIN FUSION PROTEINS

Thioredoxin and some thioredoxin fusion proteins can be released with good yield from the *E. coli* cytoplasm by a simple osmotic shock procedure.

Additional Materials

*For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.*

Cell pellet from 4-hr post-induction cultures (basic protocol)
20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA/20% (w/v) sucrose, ice-cold
20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA, ice-cold

1. Resuspend cell pellet from 4-hr post-induction cultures at a concentration of 5 $\text{OD}_{550}/\text{ml}$ in ice-cold 20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA/20% sucrose. Incubate 10 min on ice.
2. Microcentrifuge 30 sec at maximum speed, 4°C , to pellet the cells.
3. Discard the supernatant and gently resuspend the cells in an equivalent volume of ice-cold 20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA. Incubate 10 min on ice and mix occasionally by inverting the tube.

Osmotic release from the cytoplasm occurs at this stage.

4. Microcentrifuge 30 sec at maximum speed, 4°C . Save the supernatant (osmotic shockate). Resuspend the cell pellet in an equivalent volume 20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA (retentate).
5. Lyophilize 100- μ l aliquots of osmotic shockate and retentate to dryness in a Speedvac evaporator.
6. Solubilize each in 100 μ l SDS-PAGE sample buffer. Analyze 30- μ l aliquots by SDS-PAGE.

*The osmotic shock procedure provides a substantial purification step for some thioredoxin fusion proteins. This procedure will remove most of the contaminating cytoplasmic proteins as well as almost all of the nucleic acids. However the shockate will contain as contaminants about half of the cellular elongation factor-Tu (EFTu) and most of the *E. coli* periplasmic proteins.*

PURIFICATION OF THIOREDOXIN FUSION PROTEINS BY HEAT TREATMENT

Wild-type thioredoxin is resistant to prolonged incubations at 80°C. A subset of thioredoxin fusion proteins also exhibit corresponding thermal stability, and heat treatment at 80°C can sometimes be used as an initial purification step. Under these conditions the majority of contaminating *E. coli* proteins are denatured and precipitated.

Additional Materials

For recipes, see Reagents and Solutions in this unit (or cross-referenced unit); for common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Cell pellet from 4-hr post-induction cultures (basic protocol)

20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA

80°C water bath

10-ml glass-walled tube

1. Resuspend cell pellet from 4-hr post-induction cultures at a concentration of 100 OD₅₅₀/ml in 20 mM Tris·Cl (pH 8.0)/2.5 mM EDTA.

It is important to start off with a high protein concentration in the lysate to ensure efficient precipitation of denatured proteins.

2. Lyse the cells at 20,000 lb/in² in a French pressure cell as described in steps 2 to 8 of the first support protocol. Collect whole-cell lysate in a 10-ml glass-walled tube.
3. Incubate whole-cell lysate 10 min at 80°C. Remove 100-μl aliquots after 30 sec, 1 min, 2 min and 5 min and plunge immediately into ice. At 10 min plunge the remaining heated lysate into ice.

A glass-walled tube (not plastic) provides good thermal conductivity to provide a rapid rise in temperature to 80°C and then a rapid drop in temperature to 4°C. A suitable volume to use in a 10-ml glass tube is 1.5 ml lysate. For large-scale work, a glass-walled vessel should be used and the lysate should be mixed well during both heat treatment and cooling.

4. Microcentrifuge the aliquots 10 min at maximum speed, 4°C to pellet heat-denatured, precipitated proteins.
5. Remove 2-μl aliquots of the supernatants and add 28 μl SDS-PAGE sample buffer. Analyze the samples by SDS-PAGE to determine the heat stability of the fusion protein and the minimum time of heat treatment required to obtain a good purification.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

Casamino Acids (CAA), 2% (w/v)

20 g Casamino Acids (Difco certified)
H₂O to 1 liter
Autoclave or filter sterilize through a 0.45- μ m filter
Store \leq 2 months at room temperature

Do not use technical-grade Casamino Acids because it has a higher NaCl content.

CAA/glycerol/ampicillin 100 medium

800 ml 2% (w/v) Casamino Acids (see recipe; 1.6% final)
100 ml 10 \times M9 salts (see recipe; 1 \times final)
100 ml 10% (v/v) glycerol (sterile; 1% final)
1 ml 1 M MgSO₄ (sterile; 1 mM final)
0.1 ml 1 M CaCl₂ (sterile; 0.1 mM final)
1 ml 2% (w/v) vitamin B1 (sterile; 0.002% final)
10 ml 10 mg/ml ampicillin (sterile; 100 μ g/ml final)
Prepare fresh

IMC medium

200 ml 2% (w/v) Casamino Acids (see recipe; 0.4% final)
100 ml 10 \times M9 salts (see recipe; 1 \times final)
40 ml 20% (w/v) glucose (sterile; 0.5% final)
1 ml 1 M MgSO₄ (sterile; 1 mM final)
0.1 ml 1 M CaCl₂ (sterile; 0.1 mM final)
1 ml 2% (w/v) vitamin B1 (sterile; 0.002% final)
658 ml glass-distilled H₂O (sterile)
10 ml 10 mg/ml ampicillin (sterile; optional; 100 μ g/ml final)
Use fresh

IMC plates

15 g agar [Difco; 1.5% (w/v)]
4 g casamino acids [Difco-certified; 0.4% (w/v)]
858 ml glass-distilled H₂O (sterile)
Autoclave 30 min
Cool in a 50°C water bath
100 ml 10 \times M9 salts (see recipe; 1 \times final)
40 ml 20% (w/v) glucose (sterile; 0.5% final)
1 ml 1 M MgSO₄ (sterile; 1 mM final)
0.1 ml 1 M CaCl₂ (sterile; 0.1 mM final)
1 ml 2% (w/v) vitamin B1 (sterile; 0.002% final)
10 ml 10 mg/ml ampicillin (sterile; optional; 100 μ g/ml final)
Mix well and pour into Petri plates
Store \leq 1 month at 4°C

M9 salts, 10 \times

60 g Na₂HPO₄ (0.42 M)
30 g KH₂PO₄ (0.24 M)
5 g NaCl (0.09 M)
10 g NH₄Cl (0.19 M)
H₂O to 1 liter

continued

Adjust pH to 7.4 with NaOH
Autoclave or filter sterilize through a 0.45- μ m filter
Store \leq 6 months at room temperature

SDS-PAGE sample buffer

15% (v/v) glycerol
0.125 M Tris-Cl, pH 6.8 (APPENDIX 2)
5 mM Na₂EDTA
2% (w/v) SDS
0.1% (w/v) bromphenol blue
1% (v/v) 2-mercaptoethanol (2-ME; add immediately before use)
Store indefinitely at room temperature

Tryptophan, 10 mg/ml

Heat 500 ml glass-distilled H₂O to 80°C. Stir in 5 g L-tryptophan until dissolved. Filter sterilize the solution through a 0.45 μ m filter and store \leq 6 months in the dark at 4°C.

COMMENTARY

Background Information

Two significant problems plague researchers who hope to express heterologous proteins in *Escherichia coli*: inefficient initiation of translation of many eukaryotic mRNA sequences on bacterial ribosomes (Stormo et al., 1982), and proteins that often form insoluble aggregates, called inclusion bodies, that are composed of misfolded or denatured proteins (Mitraki and King, 1989). Although successful protocols for refolding eukaryotic proteins from inclusion bodies can be developed, the process is always uncertain and usually time-consuming; in most instances it is preferable to prevent inclusion-body formation in the first place. The use of *trxA* fusions provides a solution to both problems.

Inefficient initiation of translation of eukaryotic messages in *E. coli* can often be improved by modifying sequences at the 5' end of the gene. A more reliable technique that avoids the problem entirely is to use a gene fusion strategy in which the gene of interest is linked in-frame to the 3' end of a highly translated partner gene. In this case protein synthesis always initiates on the same efficiently translated fusion partner mRNA, thus high-level expression is assured. Some earlier gene fusion expression systems, for example the *trpE* and *lacZ* systems described in UNIT 16.5, offer very reliable ways of producing large quantities of any desired eukaryotic protein. However, these gene fusion systems still suffer from the pervasive inclusion-body problem. They are thus mainly useful for the production of antigens, rather than correctly folded, biologically active

proteins. More recently the maltose binding protein (MBP) and glutathione-S-transferase (GST) gene fusion expression systems (see UNITS 16.6 & 16.7) have proven more successful in producing soluble fusion proteins; these systems retain the translation advantage of the earlier fusion systems. Apart from the obvious advantages in making a correctly folded product, the synthesis of soluble fusion proteins also allows for the development of generic purification schemes based on some unique property of the fusion partner.

Why would any particular eukaryotic protein produced in the *E. coli* cytoplasm be more soluble when it is linked to a fusion partner than it would be by itself? It is likely that physical properties of the fusion partner protein are important, with efficient self-folding and high solubility being useful in this role. It is possible that some good fusion partners (proteins that fold efficiently and are highly soluble), by virtue of their desirable physical qualities, are able to keep folding intermediates of linked heterologous proteins in solution long enough for them to adopt their correct final conformations. In this respect the fusion partner may serve as a covalently joined chaperon protein, in many ways fulfilling the role of authentic chaperon proteins (McCoy, 1992), analogous to the covalent chaperon role proposed for the N-terminal pro regions of a number of protein precursors (Silen et al., 1989; Shinde et al., 1993).

Many of the known properties of *E. coli* thioredoxin (Holmgren, 1985) suggested that it would make a particularly effective fusion partner in an expression system. First, thiore-

doxin, when overproduced from plasmid vectors, can accumulate to 40% of the total cellular protein, yet even at these expression levels all of the protein remains soluble. Second, the molecule is small (11,675 M_r) and would contribute a relatively modest amount to the total mass of any fusion protein, in contrast to other systems such as the *lacZ* system. Third, the tertiary structure of thioredoxin (Katti et al., 1990) reveals that both the N- and C-termini of the molecule are accessible on the surface and in good position to link to other proteins. The structure also shows that the molecule has a very tight fold, with >90% of its primary sequence involved in strong elements of secondary structure. This provides an explanation for thioredoxin's observed high thermal stability (T_m 85°C), and suggests that the molecule might possess the robust folding characteristics that could make it a good fusion partner protein. In support of this view, complete thioredoxin domains are found in a number of naturally occurring multidomain proteins, including *E. coli* DsbA (Bardwell et al., 1991), the mammalian endoplasmic reticulum proteins ERp72 (Mazzarella et al., 1990), and protein disulfide isomerase (PDI; Edman et al., 1985). These proteins can all be considered as natural precedents for thioredoxin fusion proteins.

The synthesis of small peptides in *E. coli* is often difficult, with the products frequently being extensively degraded or insoluble. The thioredoxin tertiary structure revealed that the characteristic active site, —CGPC—, protrudes from the body of the protein as a surface loop, with few interactions with the rest of the molecule. The loop does not seem to contribute to the overall stability of thioredoxin, so the production of peptides as insertions at this site was an attractive possibility. In this location they would be protected from host-cell amino- and carboxypeptidases, and thioredoxin's high solubility should help keep them in solution. In addition, the conformation of peptides inserted at this position would be constrained, which could be an advantage for applications in which it is desirable for the peptide to adopt a particular form.

Thioredoxin has indeed proven to be an excellent partner for the production of soluble fusion proteins in the *E. coli* cytoplasm (LaVallie et al., 1993b). Figure 16.8.3 demonstrates the production of soluble fusion proteins between thioredoxin and eleven human and murine cytokines and growth factors using the *trxA* vectors. All of these mammalian proteins had been previously produced in *E. coli* only as

insoluble inclusion bodies. As thioredoxin fusions, the growth factors are not only made in a soluble form, but in most cases they are also biologically active in *in vitro* assays.

Experience gained while working with these and a number of other *trxA* fusion proteins shows that two further characteristics of thioredoxin can be exploited as purification tools. The first is the inherent thermal stability of the molecule, a property that is retained by some thioredoxin fusion proteins. This enables heat treatment to be used as an effective purification step. The second additional property relates to thioredoxin's cellular location. Although *E. coli* thioredoxin is a cytoplasmic protein, it has been shown to occupy a special position within the cell—it is primarily located on the cytoplasmic face of the adhesion zones that exist between the inner and outer membranes of the *E. coli* cell envelope (Lunn and Pigiet, 1982). From this location thioredoxin is quantitatively released to the exterior of the cell by simple osmotic shock or freeze/thaw treatments, a remarkable property that is retained by some thioredoxin fusion proteins, thus providing a simple purification step.

A more generic method for purification of any soluble thioredoxin fusion employs a modified form of thioredoxin (called “His-patch Trx”), which has been designed to bind to metal chelate resins (E.A. DiBlasio, J.M. McCoy, and E.R. LaVallie, manuscript in preparation).

Critical Parameters

Lack of protein solubility leading to inclusion-body formation in *E. coli* is a complex phenomenon with many contributing factors: simple insolubility as a result of high-level expression, insolubility of protein-folding intermediates, lack of appropriate bacterial chaperon proteins, and lack of glycosylation mechanisms in the bacterial cytoplasm. Fusion of heterologous proteins to thioredoxin or to other fusion partners can help address most of these solubility issues. However, another important factor contributing to inclusion body formation is the inability to form essential disulfide bonds in the reducing environment of the bacterial cytoplasm, which leads to incorrect folding. Thermal lability of even correctly folded heterologous proteins in the absence of these stabilizing disulfide cross-links is a significant problem, so the expression of fusion genes should be attempted over a wide range of temperatures, even as low as 15°C (the limit for *E. coli* growth is ~8°C). Thermal denaturation is

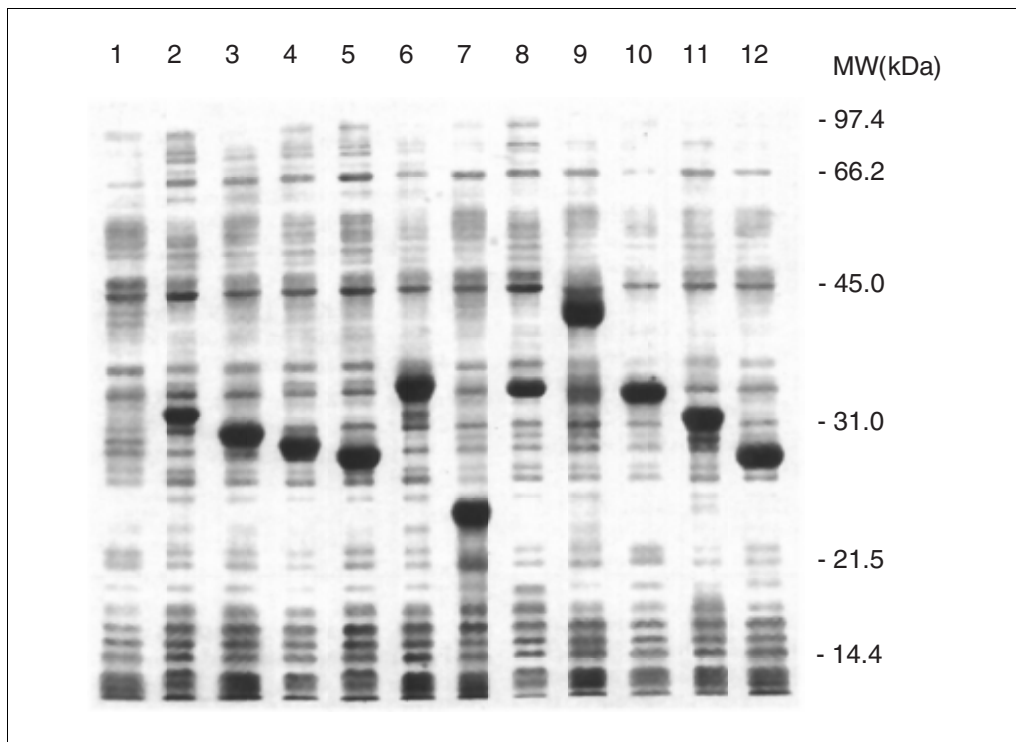


Figure 16.8.3 Expression of thioredoxin gene fusions. The gel shows proteins found in the soluble fractions derived from *E. coli* cells expressing eleven different thioredoxin gene fusions. Lane 1, host *E. coli* strain GI724 (negative control, 37°C); lane 2, murine interleukin-2 (IL-2; 15°C); lane 3, human IL-3 (15°C); lane 4, murine IL-4 (15°C); lane 5, murine IL-5 (15°C); lane 6, human IL-6 (25°C); lane 7, human MIP-1a (37°C); lane 8, human IL-11 (37°C); lane 9, human macrophage colony-stimulating factor (M-CSF; 37°C); lane 10, murine leukemia inhibitory factor (LIF; 25°C); lane 11, murine steel factor (SF; 37°C); and lane 12, human bone morphogenetic protein-2 (BMP-2; 25°C). Temperatures in parentheses are the production temperature chosen for expressing each fusion. This is a 10% SDS-polyacrylamide gel, stained with Coomassie brilliant blue.

a time-dependent process, so it is also prudent to monitor the solubility of the expressed fusion protein over the time course of induction.

A great many proteins contain distinct structural domains. For example, hormone receptor proteins usually have an extracellular ligand-binding domain, a transmembrane region, and an intracellular effector domain. Sometimes expressing these domains individually as fusion proteins can yield better results than expressing the entire protein. The exact positions chosen for boundaries of the domains to be expressed in the fusion protein are important and can be determined from a knowledge of the tertiary structure of the protein of interest, by homology comparisons with similar proteins, by limited proteolysis or other domain-mapping experiments, or empirically by generating multiple fusions that test different boundary positions.

It is important to be consistent in treating samples for loading on gels. For example, using different heating conditions from one experi-

ment to the next can result in a mobility shift for the protein of interest.

Anticipated Results

Thioredoxin fusion protein yields are usually in the range of 5% to 20% of total cell protein. At these expression levels a 1-liter induction culture in a shaker flask will yield ~3 g (wet weight) of cells, 300 mg total protein, and 15 to 60 mg of thioredoxin fusion protein. The final recovered yield will depend on factors such as solubility of the fusion protein and the efficiency of downstream purification procedures.

Time Considerations

From a single colony on a plate, the basic induction protocol requires an overnight growth to prepare a liquid inoculum and a 3.5-hr preinduction growth at 30°C the next day, followed by a 4-hr 37°C induction period. These times are significantly longer if lower induction temperatures are required (see Table

16.8.1). Lysis of a sample in the French pressure cell should require ≤ 5 min, and both the heat-treatment and osmotic-shock procedures require <1 hr each. SDS-PAGE takes 2.5 hr.

Literature Cited

- Bardwell, J.C.A., McGovern, K., and Beckwith, J. 1991. Identification of a protein required for disulfide bond formation in vivo. *Cell* 67:581-589.
- Edman, J.C., Ellis, L., Blacher, R.W., Roth, R.A., and Rutter, W.J. 1985. Sequence of protein disulfide isomerase and implications of its relationship to thioredoxin. *Nature* 317:267-270.
- Holmgren, A. 1985. Thioredoxin. *Ann. Rev. Biochem.* 54:237-271.
- Katti, S.K., LeMaster, D.M., and Eklund, H. 1990. Crystal structure of thioredoxin from *Escherichia coli* at 1.68 angstroms resolution. *J. Mol. Biol.* 212:167-184.
- LaVallie, E.R., Rehemtulla, A., Racie, L.A., DiBlasio, E.A., Ferenz, C., Grant, K.L., Light, A., and McCoy, J.M. 1993a. Cloning and functional expression of a cDNA encoding the catalytic subunit of bovine enterokinase. *J. Biol. Chem.* 268:23311-23317.
- LaVallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., and McCoy, J.M. 1993b. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *BioTechnology* 11:187-193.
- Lunn, C.A. and Pigiet, V.P. 1982. Localization of thioredoxin from *Escherichia coli* in an osmotically sensitive compartment. *J. Biol. Chem.* 257:11424-11430.
- Mazzarella, R.A., Srinivasan, M., Haugejorden, S.M., and Green, M. 1990. ERp72, an abundant luminal endoplasmic reticulum protein, contains three copies of the active site sequences of protein disulfide isomerase. *J. Biol. Chem.* 265:1094-1101.
- McCoy, J.M. 1992. Heat-shock proteins and their potential uses for pharmaceutical protein production in microorganisms. In *Stability of Protein Pharmaceuticals, Part B.* (T. Ahern and M. Manning, eds.) pp 287-316. Plenum Press, New York.
- Mieschendahl, M., Petri, T., and Hanggi, U. 1986. A novel prophage independent trp regulated lambda pL expression system. *BioTechnology* 4:802-808.
- Mitraki, A. and King, J. 1989. Protein folding intermediates and inclusion body formation. *BioTechnology* 7:690-697.
- Norrander, J., Kempe, T., and Messing, J. 1983. Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. *Gene* 26:101-106.
- Shimatake, H. and Rosenberg, M. 1981. Purified λ regulatory protein cII positively activates promoters for lysogenic development. *Nature* 292:128-132.
- Shinde, U., Chatterjee, S., and Inouye, M. 1993. Folding pathway mediated by an intramolecular chaperone. *Proc. Natl. Acad. Sci. U.S.A.* 90:6924-6928.
- Silen, J.L., Frank, D., Fujishige, A., Bone, R., and Agard, D.A. 1989. Analysis of prepro- α -lytic protease expression in *Escherichia coli* reveals that the pro region is required for activity. *J. Bacteriol.* 171:1320-1325.
- Stormo, G.D., Schneider, T.D., and Gold, L. 1982. Characterization of translation initiation sites in *E. coli*. *Nucl. Acids Res.* 10:2971-2996.

Contributed by John McCoy
and Edward LaVallie
Genetics Institute
Cambridge, Massachusetts

EXPRESSION OF PROTEINS IN INSECT CELLS USING BACULOVIRUS VECTORS

Overview of the Baculovirus Expression System

Baculoviruses have emerged as a popular system for overproducing recombinant proteins in eukaryotic cells (Miller et al., 1986; Luckow and Summers, 1988; Miller, 1988; Luckow, 1991). Several factors have contributed to this popularity. First, unlike bacterial expression systems, the baculovirus-based system is a eukaryotic expression system and thus uses many of the protein modification, processing, and transport systems present in higher eukaryotic cells. In addition, the baculovirus expression system uses a helper-independent virus that can be propagated to high titers in insect cells adapted for growth in suspension cultures, making it possible to obtain large amounts of recombinant protein with relative ease. The majority of this overproduced protein remains soluble in insect cells, in contrast to the insoluble proteins often obtained from bacteria. Furthermore, the viral genome is large (~130 kbp) and thus can accommodate large segments of foreign DNA. Finally, baculoviruses are non-infectious to vertebrates, and their promoters have been shown to be inactive in most mammalian cells (Carbonell et al., 1985), which gives them a possible advantage over other systems when expressing oncogenes or potentially toxic proteins.

This unit gives an overview of the baculovirus expression system. A comprehensive guide describing further details can be found in O'Reilly et al. (1992).

BACULOVIRUS LIFE CYCLE

Currently, the most widely used baculovirus expression system utilizes a lytic virus known as *Autographa californica* nuclear polyhedrosis virus (AcMNPV; hereafter called baculovirus). This virus is the prototype of the family *Baculoviridae*. It is a large, enveloped, double-stranded DNA virus that infects arthropods. The baculovirus expression system takes advantage of some unique features of the viral life cycle (Fig. 16.9.1). See Doerfler and Bohm (1986) for a comprehensive review.

As with mammalian DNA viruses, the bacu-

lovirus life cycle is divided temporally into immediate early, early, late, and very late phases. Viruses enter the cell by adsorptive endocytosis and move to the nucleus, where their DNA is released. DNA replication begins ~6 hr after infection and is followed by viral assembly in the nucleus of the infected cell. Two types of viral progeny are produced during the life cycle of the virus: extracellular virus particles (nonoccluded viruses) during the late phase and polyhedra-derived virus particles (occluded viruses) during the very late phase of infection. Extracellular virus is released from the cell by budding, beginning at ~12 hr postinfection, and is produced at a logarithmic rate until 20 hr postinfection, after which production drops off. Polyhedra-derived virus, on the other hand, appears in the nucleus at ~18 hr postinfection and continues to accumulate as late as 72 hr postinfection, or until the cells lyse. Occluded viral particles are embedded in proteinaceous viral occlusions called polyhedra within the nucleus of infected cells. The polyhedrin protein (29 kDa) is the major protein component of the occlusion bodies.

The polyhedrin protein serves an important function for the survival and propagation of the virus in nature. Because baculoviruses are lytic, they quickly kill their insect host after infection. The polyhedrin protein serves to sequester, and thereby protect, hundreds of virus particles from proteolytic inactivation by the decomposing host tissue. The virus is transmitted when occlusion bodies are ingested by a new host as it feeds on a contaminated food source. The polyhedrin protein dissolves in the alkaline environment of the new host's gut and the occluded virus is released. This virus infects the gut epithelial cells and virus replication takes place. Nonoccluded virus is then produced and budded from the infected gut cells. At this point, the virus spreads throughout the tissues of its new host. Although the polyhedrin protein is essential for survival of the virus in nature, it is dispensable for virus survival and propagation in tissue culture cells.

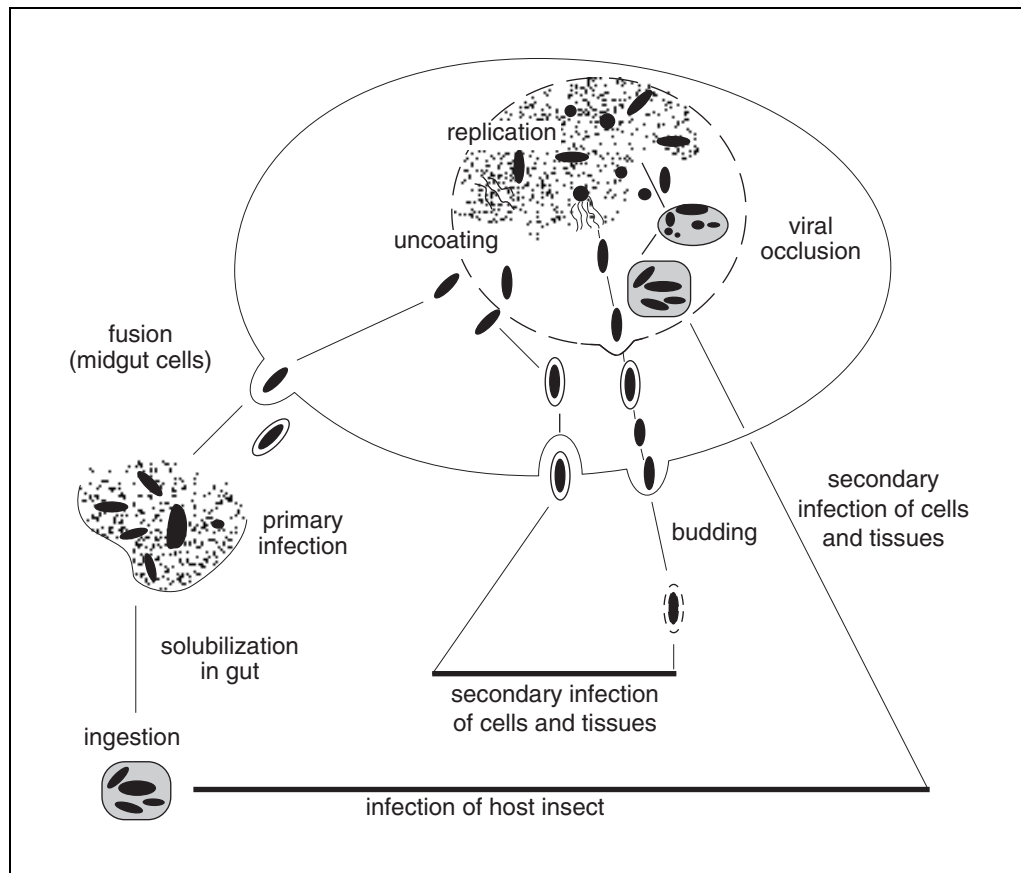


Figure 16.9.1 Baculovirus life cycle. Viruses enter cells by adsorptive endocytosis and move to the nucleus where their DNA is released. Both DNA replication and viral assembly take place in the nuclei of infected cells to generate two types of viral progeny. These include extracellular (nonoccluded) virus particles and polyhedra-derived (occluded) virus particles. Extracellular virus is released from the cell by budding, starting at ~12 hr postinfection and ending ~36 hr postinfection. Polyhedra-derived virus, on the other hand, appears later (~18 hr postinfection) and accumulates in the nuclei of infected cells ≤ 72 hr postinfection or until cellular lysis. Polyhedra-derived virus is embedded in proteinaceous viral occlusions, the major protein component of which is the viral polyhedrin protein. Secondary infection of cells and tissues occurs by two pathways. In the first, the extracellular virus, once budded from the site of primary infection, is free to infect neighboring cells by the pathway just described. Alternatively, polyhedra-derived virus is released from occlusion bodies after an infected food source is ingested by a new host. Reproduced from Summers and Smith (1987) with permission from the Texas Agricultural Experiment Station.

BACULOVIRUS EXPRESSION SYSTEM

The baculovirus expression system takes advantage of several facts about polyhedrin protein: (1) that it is expressed at very high levels in infected cells, constituting more than half of the total cellular protein late in the infectious cycle; (2) that it is nonessential for infection and replication of the virus, meaning that the recombinant virus does not require any helper function; and (3) that viruses lacking the polyhedrin gene have a plaque morphology that is distinct from that of viruses containing the gene. Recombinant baculoviruses are generated by replacing the polyhedrin gene with a foreign gene through homologous recombina-

tion. In this system, the distinctive plaque morphology provides a simple visual screen for identifying the recombinants.

To produce a recombinant virus that expresses the gene of interest, the gene is first cloned into a transfer vector (see discussion below under Choosing a Baculovirus Transfer Vector). Most baculovirus transfer vectors contain the polyhedrin promoter followed by one or more restriction enzyme recognition sites for foreign gene insertion. Once cloned into the transfer vector, the gene is flanked both 5' and 3' by viral-specific sequences. Next, the recombinant vector is transfected along with wild-type viral DNA into insect cells. In a homologous recombination event, the foreign gene is

inserted into the viral genome and the polyhedrin gene is replaced. Thus recombinant viruses lack the polyhedrin gene and in its place contain the inserted gene, whose expression is under the control of the polyhedrin promoter.

Homologous recombination between circular wild-type DNA and the recombinant plasmid DNA occurs at a low frequency (typically 0.1% to 0.5%). This limitation was recently overcome by the development of viruses having *Bsu*36I restriction sites positioned within an essential gene—*Open Reading Frame (ORF) 1629*—which is downstream of the AcMNPV polyhedrin gene—and also in the upstream *ORF 603*—in such a way that digestion releases a fragment containing a sequence necessary for virus growth (Kitts and Possee, 1993). When insect cells are cotransfected with an appropriate recombinant transfer plasmid and linearized *ORF 1629*-deleted baculovirus DNA, the necessary *ORF 1629* is provided by the transfer plasmid through homologous recombination. The vast majority of the progeny viruses, in many cases >99.9%, that are derived from these cotransfections contain the repaired virus with the target gene, thus minimizing the need to screen and plaque-purify recombinants. Several companies (Pharmingen, Invitrogen, Clontech, and Novagen; see APPENDIX 4) market linearized *ORF 1629*-deleted AcMNPV DNA. To further facilitate the identification of recombinants, several of these commercially available baculovirus DNAs contain the bacterial *lacZ* gene, which codes for β -galactosidase in lieu of the AcMNPV polyhedrin gene, thereby allowing *lacZ*-negative recombinants to be distinguished visually from any residual nonrecombinant viruses via a plaque assay. Nonrecombinant viruses form blue plaques on Xgal plates because they contain a functional *lacZ* gene, whereas recombinants form colorless, opaque plaques. Recombinant viruses can also be identified by DNA hybridization and polymerase chain reaction (PCR) amplification.

Another rapid and efficient method for generating recombinant baculoviruses uses site-specific transposition to insert foreign genes by homologous recombination into a bacmid propagated in *E. coli* rather than in insect cells. In this case, recombinant viral DNA is isolated from individual bacterial colonies and is free of any wild-type viral DNA. Upon transfection of insect cells, recombinant virus is generated free of parental nonrecombinant virus, thereby eliminating the need for multiple rounds of plaque purification. This is the basis of the Bac-to-Bac

baculovirus expression system, which is available commercially from Invitrogen.

POSTTRANSLATIONAL MODIFICATION OF PROTEINS IN INSECT CELLS

Because baculoviruses infect invertebrate cells, it is possible that the processing of proteins produced by them is different from the processing of proteins produced by vertebrate cells. Although this seems to be the case for some posttranslational modifications, it is not the case for others. For example, two of the three posttranslational modifications of the tyrosine protein kinase, pp60^{c-src}, that occur in higher eukaryotic cells (myristylation and phosphorylation of serine 17) also take place in insect cells. However, another modification of pp60^{c-src} observed in vertebrate cells, phosphorylation of tyrosine 527, is almost undetectable in insect cells (Piwnicka-Worms et al., 1990).

In addition to myristylation, palmitylation has been shown to take place in insect cells. However, it has not been determined whether all or merely a subfraction of the total recombinant protein contains these modifications. Cleavage of signal sequences, removal of hormonal prosequences, and polyprotein cleavages have also been reported, although cleavage varies in its efficiency. Internal proteolytic cleavages at arginine- or lysine-rich sequences have been reported to be highly inefficient, and alpha-amidation, although it does not occur in cell culture, has been reported in larvae and pupae (Hellers et al., 1991). In most of these cases a cell- or species-specific protease may be necessary for cleavage. Protein targeting seems conserved between insect and vertebrate cells. Thus, proteins can be secreted and localized faithfully to either the nucleus, cytoplasm, or plasma membrane. Although much remains to be learned about the nature of protein glycosylation in insect cells, proteins that are glycosylated in vertebrate cells will also generally be glycosylated in insect cells. However, with few exceptions the N-linked oligosaccharides in insect cell-derived glycoproteins are only high-mannose type and are not processed to complex-type oligosaccharides containing fucose, galactose, and sialic acid. O-linked glycosylations have been even less well characterized in *Sf9* cells, but have been shown to occur. For further information on posttranslational modifications of proteins and protein processing in insect cells, see Davidson et al. (1990), O'Reilly et al. (1992), Jarvis and Sum-

mers (1992), Grabenhorst et al. (1993), James et al. (1995), Jarvis and Finn (1995), Davis and Wood (1995), and Ogonah et al. (1996).

STEPS FOR OVERPRODUCING PROTEINS USING THE BACULOVIRUS SYSTEM

The use of the baculovirus expression system is presented in detail in *UNITS 16.10 & 16.11*. The following steps comprise a brief overview (also see Fig. 16.9.2).

1. Clone the gene of interest into the appropriate baculovirus expression vector and cotransfect with linearized baculoviral DNA (available from various vendors) or use transposition to create recombinant bacmid DNA in *E. coli* (BAC-TO-BAC system; Invitrogen). Alternatively, purify circular wild-type baculovirus DNA (*UNIT 16.10*, Alternate Protocol 1).

2. Cotransfect baculovirus DNA with the recombinant baculovirus plasmid or transfect purified bacmid DNA into *Sf9* insect cells (*UNIT 16.10*, Basic Protocol 2).

3. Collect the medium, which contains the baculoviral particles (*UNIT 16.10*, Basic Protocol 2), and plaque the virus on *Sf9* cells to separate recombinant from nonrecombinant virus (*UNIT 16.10*, Basic Protocol 4). This and subsequent rounds of plaque purification are optional when linearized virus that contains a lethal deletion is used (e.g., BaculoGold from Pharmingen), as in that case >99.9% of all amplified virus particles will be recombinant because of selection pressure. Similarly, when bacmid DNA is used, no purification is required.

4. Amplify the virus stock by infecting fresh insect cells (*UNIT 16.10*, Basic Protocol 3) and determine titer of the amplified virus stock (*UNIT 16.10*, Basic Protocol 4).

5. Express the protein of interest by infecting a new batch of insect cells with the high-titer baculovirus stock (*UNIT 16.11*, Basic Protocol 1). Determine the expression level of the recombinant protein of interest and analyze its biological activity (*UNIT 16.11*, Support Protocols 1 and 2).

CHOOSING A BACULOVIRUS TRANSFER VECTOR

The majority of available baculovirus vectors are pUC-based and confer ampicillin resistance. Most contain the polyhedrin gene promoter and insertion site(s) for cloning a foreign gene of interest, flanked by viral sequences that lie 5' to the promoter and 3' to the foreign gene insert. These flanking sequences facilitate ho-

mologous recombination between the vector and baculovirus DNA. Other baculovirus vectors contain the p10 promoter, another strong, very late promoter, or the basic protein promoter expressed late in the infection process. Some vectors are designed to express more than one heterologous gene or to express genes as fusions to N-terminal signal sequences or leader peptides, which facilitate secretion and purification of the recombinant protein. For more information on these vectors, see Table 16.9.1, O'Reilly et al. (1992), and relevant catalogs from Pharmingen, Clontech, Invitrogen, Novagen, and Stratagene.

A major consideration when choosing the appropriate baculovirus expression vector is whether to express the recombinant protein as a fusion or nonfusion protein in insect cells. Fusion proteins containing a specific tag have the advantage of easy purification and detection. For nonfusion proteins, there are several vectors available—notably pVL1392 and pVL1393 (Pharmingen and Invitrogen)—each of which contains a polylinker in the opposite orientation from the other. These vectors are derived from pAcYM1 (Matsuura et al., 1987) and pAcCL29 (Livingston and Jones, 1989) and differ only in the order of the cloning sites. Other nonfusion vectors include pBacPAK8 and pBacPAK9 (Clontech), pBac-1 (Novagen), pFAST Bac 1 (Invitrogen), and pAcSG2 (Pharmingen). To express the protein as a polyhistidine fusion protein, there are several available vectors: pAcHLT-A, -B, and -C (Pharmingen), pBlueBacHis-A, -B, and -C and pFAST BacHT -A, -B, and -C (Invitrogen), and pBac-2cp (Novagen). These vectors allow easy purification of the recombinant protein using a nickel-chelating resin (*UNIT 10.11B*). To express the protein as a glutathione-S-transferase (GST) fusion protein, the pAcGHLT-A, -B, and -C vectors are commercially available (Pharmingen). These produce an N-terminal GST fusion protein. As with all fusion protein vectors, the gene of interest has to be cloned in the proper reading frame with respect to the tag. Several other affinity tags, such as FLAG, HA, and myc, have also been used successfully for purification using commercially available affinity resins. These tags can be fused to coding sequences using PCR.

Secretion of recombinant proteins into insect cell medium simplifies purification and characterization of expressed recombinant proteins. The ease of purification is further enhanced by the use of serum-free medium. There are several vectors that contain signal se-

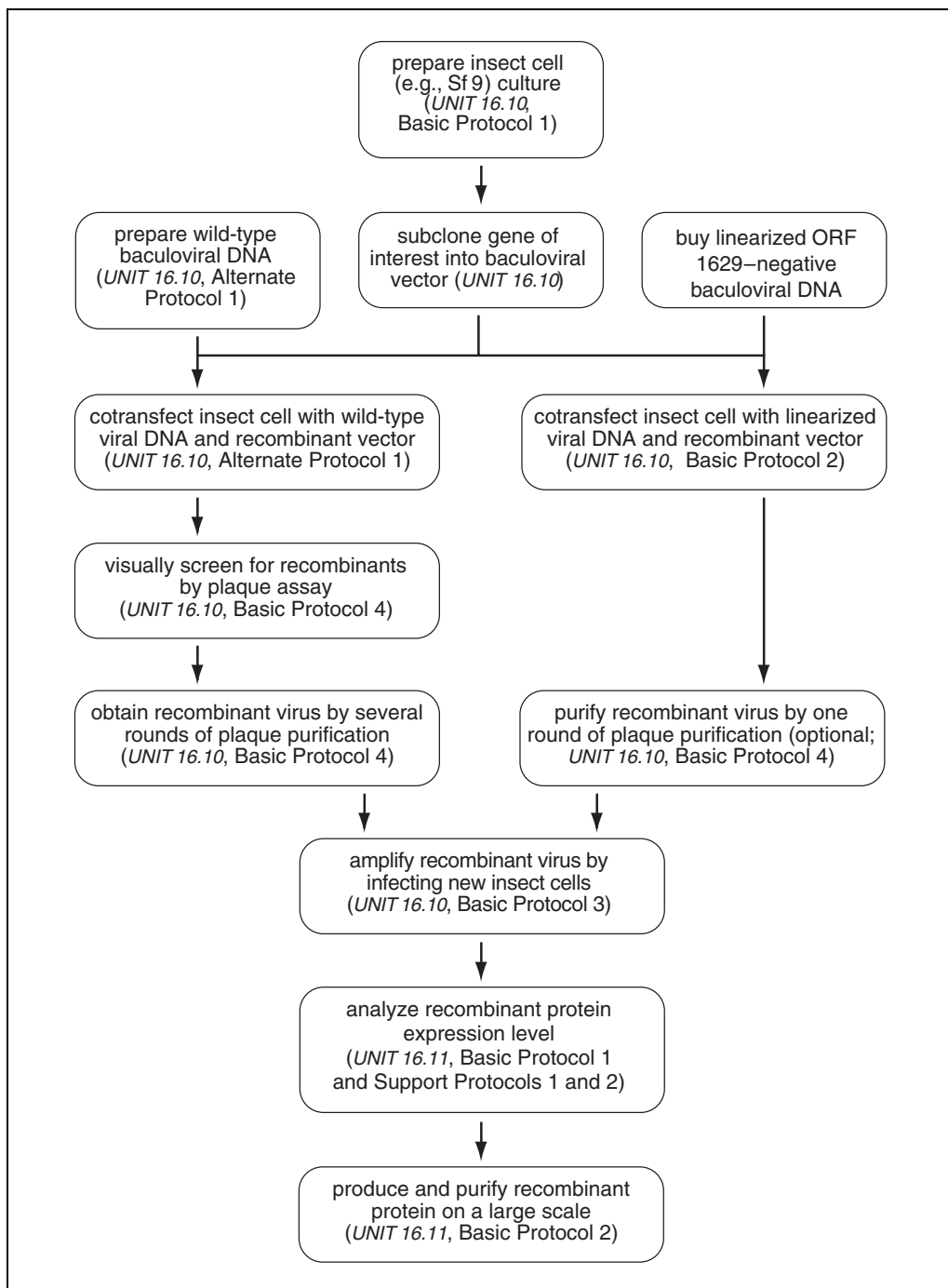


Figure 16.9.2 Flow chart for the expression of proteins in insect cells using the baculovirus system. The flow chart describes the production of recombinants using circular wild-type baculovirus DNA and the more recently developed linearized *ORF 1629*-negative variants. The use of linearized baculovirus DNA, now widely accepted, simplifies the whole protocol and produces viral-infection stocks with little or no background of nonrecombinants, which may be used directly without plaque purification.

Table 16.9.1 Baculoviral Expression Vectors^a

Vector	Promoter	Promoter type	Fusion protein	Features	Source ^b
Polyhedrin locus–based					
<i>Single-baculovirus promoter vectors</i>					
pVL1392/1393 (pair)	P	VL	No	Standard transfer vectors	PG, IN
pAcSG2	P	VL	— ^c	Recommended for large inserts	PG
pAcMP2/3 (pair)	BP	L	No	Facilitates post-translational modifications	PG
pAcUW21	P	VL	No	Polyhedrin gene, F1 origin	PG
pBacPak 8/9 (pair)	P	VL	No	Standard transfer vectors	CT
pBAC-1	P	VL	— ^c	F1 origin	NG
pBacgus-1	P	VL	— ^c	<i>gus</i> reporter gene	NG
pBlue Bac III	P	VL	No	<i>lacZ</i> gene	IN
pAcGHLT-A, B, C (set)	P	VL	Yes	GST and 6×His tags, thrombin cleavage site	PG
pAcHLT-A, B, C (set)	P	VL	Yes	6×His tag, thrombin cleavage site	PG
pBac-2cp	P	VL	Yes	6×His and S tags, F1 origin	NG
pBACgus-2cp	P	VL	Yes	6×His and S tags, F1 origin, <i>gus</i> reporter gene	NG
pBlue Bac His, A, B, C (set)	P	VL	Yes	<i>lacZ</i> gene, 6×His tag	IN
pAc 360	P	VL	Yes	Translational fusion with polyhedrin gene	IN
pAcG1	P	VL	Yes	GST tag	PG
pAcG2T	P	VL	Yes	GST tag, thrombin cleavage site	PG
pAcG3X	P	VL	Yes	GST tag, factor Xa cleavage site	PG
BioColors BV Control (set)	P	VL	Yes	Fusion with GFP or its variants	PG
BioColors His (set)	P	VL	Yes	Fusion with GFP or its variants, 6×His tag, thrombin cleavage site	PG
<i>Secretory-signal vectors</i>					
pAcGP67, A, B, C (set)	P	VL	Yes	gp67 signal sequence for secretion	PG
pAcSecG2T	P	VL	Yes	gp67 signal sequence for secretion, GST tag	PG
pPbac	P	VL	Yes	Placental AKP signal sequence for secretion	SG
pMbac	P	VL	Yes	Melittin signal sequence for secretion	SG
pBac surf-1	P	VL	Yes	gp67 signal sequence for secretion, F1 origin	NG
<i>Ligation-independent cloning vectors</i>					
pAcSG2-LIC	P	VL	No	LIC site for fast PCR cloning	PG
pAcGST-LIC-2T	P	VL	No	GST tag, thrombin cleavage, and LIC sites	PG
pAcGST1-LIC	P	VL	No	GST tag, LIC site	PG

continued

Table 16.9.1 Baculoviral Expression Vectors^a, continued

Vector	Promoter	Promoter type	Fusion protein	Features	Source ^b
<i>Ligation-independent cloning vectors, continued</i>					
pBACgus-2cp LIC	P	VL	No	<i>gus</i> reporter gene, 6×His and S tags, thrombin cleavage site, F1 origin, LIC site	NG
pBAC-2cp LIC	P	VL	No	6×His and S tags, thrombin cleavage site, F1 origin, LIC site	NG
<i>Multiple baculovirus promoter vectors</i>					
pAcUW51	P, p10	VL	No	Expression of 2 foreign genes, F1 origin	PG
p2Bac	P, p10	VL	No	Expression of 2 foreign genes	IN
pAcAB3	P, p10	VL	No	Expression of 3 foreign genes	PG
pAcAB4	P, p10	VL	No	Expression of 4 foreign genes	PG
pAcUW31	P, p10	VL	No	Expression of 2 foreign genes, M13 origin	CT
pBAC4x-1	P, p10	VL	— ^c	Expression of 4 foreign genes, F1 origin	NG
pBACgus 4x-1	P, p10	VL	— ^c	Expression of 4 foreign genes, <i>gus</i> reporter gene, F1 origin	NG
p10 locus-based					
<i>Single-baculovirus promoter plasmids</i>					
pAcUW1	p10	VL	No	Standard p10 locus vector	PG
<i>Multiple promoter vector</i>					
pAcUW32/43 (pair)	P, p10	VL	No	Expression of 2 foreign genes, F1 origin	PG
Bacmid expression vectors					
FastBac1	P	VL	No	Bacmid expression system	IN
FastBacHT A,B,C (set)	P	VL	Yes	Bacmid expression system, 6×His tag	IN
FastBacDUAL	P	VL	Yes	Bacmid expression system	IN

^aAbbreviations: 6×His, six-histidine (tag); BP, basic protein (promoter); CT, Clontech; GFP, green fluorescent protein; GST, glutathione-S-transferase (tag); *gus*, β-glucuronidase; IN, Invitrogen; L, late (promoter); LIC, ligation-independent cloning; NG, Novagen; P, polyhedrin; PG, Pharmingen; SG, Stratagene; S tag, peptide tag from Novagen; VL, very late (promoter).

^bSee APPENDIX 4 for source addresses and telephone numbers.

^cFusion protein optional.

quences under strong polyhedrin-promoter control that direct the nascent polypeptide chain toward the secretory pathway of the cell. The sequences to be expressed are inserted downstream with respect to the signal sequences to generate a fusion gene that is transcribed under strong polyhedrin-promoter control. The secretion of biologically active protein from insect cells is the final step in a complex pathway of posttranslational modifications performed in the endoplasmic reticulum (ER) and the Golgi complex (GC). Proteins destined for secretion are first cotranslationally translo-

cated into the lumen of the ER, where initial steps of carbohydrate processing occur. Later, the protein is transported to the GC, where further modifications take place. During translocation, the amino-terminal leader peptide sequence is, in most cases, proteolytically removed. The major determinants for the final form of posttranslational modification of a protein are its primary structure and the conformation presented to successive processive steps. Each secreted protein will present its own characteristics and potential problems for efficient secretion in a biologically active form. In gen-

**Protein
Expression**

16.9.7

eral, the protein produced by baculovirus-infected insect cells carries modifications that are very similar to those of the native protein. With no need for cell lysis, purification of the secreted recombinant proteins is extremely easy. For an example of a recombinant protein expressed and purified in this way, see Murphy et al. (1993). Several vectors have been developed that utilize the gp67 secretory sequence of the baculovirus envelope protein: pAcGP67-A, -B, or -C (Pharmingen) and pBac surf-1 (Novagen). Two additional vectors, pMbac and pPbac (Stratagene), have insertions, respectively, of the melittin and human placental alkaline phosphatase secretory signal sequences. Additionally, these two vectors contain a p10-promoter-driven *lacZ* gene, allowing color selection of recombinants on Xgal plates.

To improve screening of recombinant baculoviruses when using wild-type AcMNPV DNA, pBlueBacIII (Invitrogen), derived from pJVNheI (Vialard et al., 1990), was developed. pBlueBacIII has a multiple cloning site and contains two promoters—the polyhedrin promoter and the early-to-late (ETL) promoter—downstream with respect to which the *lacZ* gene has been inserted. As with the other baculovirus vectors, the gene of interest is cloned downstream of the polyhedrin promoter, which then controls synthesis of the recombinant protein. The recombinant virus is plaqued using agarose overlays containing 150 µg/ml Xgal; the plaques are visualized as described in Basic Protocol 4 of *UNIT 16.10*. Recombinant viruses generate plaques that are blue and lack occlusion bodies.

CHOOSING A BACULOVIRUS DNA

There are several methods for generating recombinant baculovirus. Initially it was required that researchers cotransfect the recombinant transfer vector with wild-type baculovirus DNA, generate a supernatant containing recombinant baculovirus, and screen out the nonrecombinant wild-type virus background through several rounds of plaque purification. This was a time-consuming process requiring technical expertise developed over a long period of time. The development of modified linearized baculovirus DNA allowed the generation of an initial viral stock containing little or no nonrecombinant virus, thus abolishing the need for plaque purification. The principle of this technique lies in the construction of a modified type of baculovirus DNA which, after linearization, contains a lethal deletion and no longer codes for any viable virus.

Cotransfection of the linearized baculoviral DNA with a complementing plasmid construct rescues the lethal deletion of the essential gene—*ORF 1629*—that lies downstream of the AcMNPV polyhedrin gene of the baculovirus DNA (see Fig. 16.9.3). Therefore, the baculovirus transfer vector must be polyhedrin-locus based to rescue this deficiency. This means that the flanking sequences of its promoter region must be derived from the polyhedrin locus of the AcMNPV wild-type virus, otherwise it will not recombine with the polyhedrin locus of linearized *ORF 1629*-deleted baculoviral DNA. Protocols for both methods are covered in *UNIT 16.10*, which also discusses a new baculovirus variant allowing the gene of interest to be cloned directly into the baculovirus genome, thus obviating the need for transfer vectors.

Another method for avoiding the time-consuming purification by plaque assay is the Bacto-Bac system (Invitrogen). The gene of interest is cloned into a donor plasmid, pFastBac1, and transformed into competent *E. coli* cells containing a helper plasmid and a baculovirus shuttle vector (bacmid). pFastBac1 contains Tn7 sites, and is transposed into the bacmid using functions supplied by the helper plasmid in *trans*. The recombinant bacmid is isolated from the competent bacteria by miniprep and transfected into insect cells using a cationic lipid reagent (Cellfectin). Screening for recombinants is done in *E. coli* and can therefore be done much more quickly than via multiple rounds of plaque purification, as is done in generation of recombinant baculovirus using wild-type virus.

REAGENTS, SOLUTIONS, AND EQUIPMENT FOR THE BACULOVIRUS SYSTEM

Reagents and solutions commonly used for expression of proteins using baculovirus vectors are summarized below. See *APPENDIX 4* for supplier contact information.

1. *Suitable insect cell lines.* Sf9 cells are derived from the ovaries of the fall armyworm (*Spodoptera frugiperda*) and are available from American Type Culture Collection, Pharmingen, or Invitrogen. A similar cell line, Sf21, is available from the same vendors. As an alternative to *S. frugiperda* cell lines, the *Trichoplusia ni* High Five line, derived from *T. ni* egg-cell homogenates, is available from Invitrogen. Several proteins have been reported to have significantly higher expression using this *T. ni* cell line. Additionally, High Five cells have a

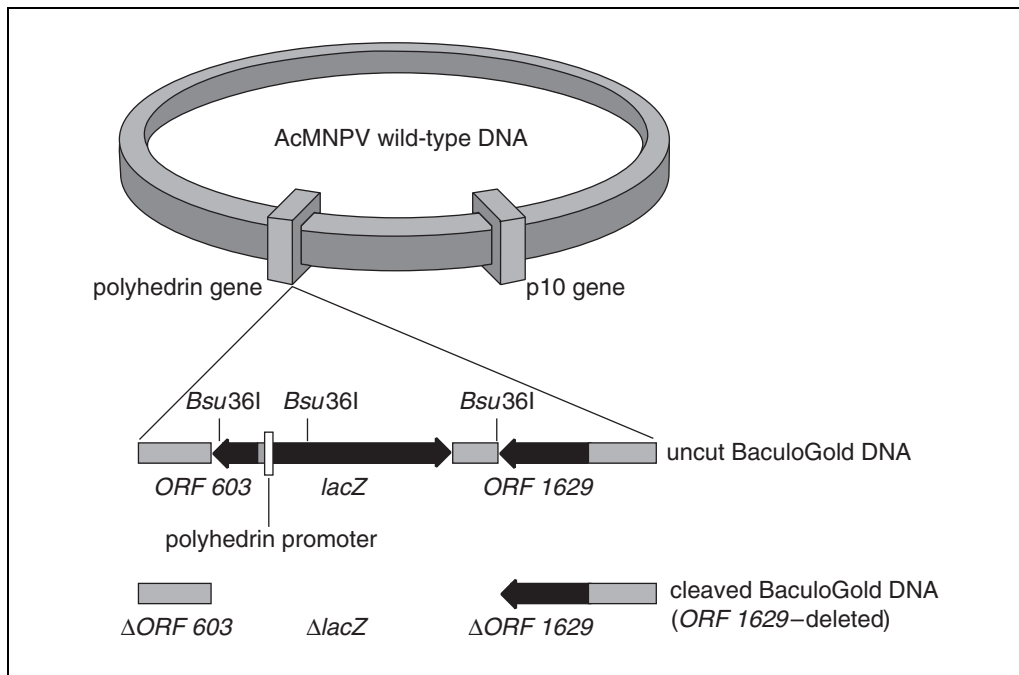


Figure 16.9.3 Generation and purification of recombinant baculovirus.

rapid doubling time as adherent cultures and adapt quickly to serum-free media.

2. *Fully prepared TNM-FH insect medium.* This will contain trace metals, lactalbumin hydrolysate, yeastolate, 10% fetal bovine serum (FBS), and gentamicin. The medium can be purchased from Pharmingen and several other vendors. It can also be prepared from Grace's insect-cell culture medium (available at 1× and 2× concentration in powdered or liquid form from Invitrogen). For instructions on preparing media from individual components, see O'Reilly et al. (1992). FBS is available from many vendors. Obtain and test different lots of serum from a number of suppliers. The lot that promotes the best growth rate and cell viability should be purchased in bulk. See APPENDIX 3F for additional discussion of FBS. Alternatively, a serum-free insect cell culture medium can be purchased from several vendors (BaculoGold medium from Pharmingen, Sf-900 II from Invitrogen, HYQ-CCM3 and SFX-Insect from Hyclone, or ExCell 401 from JRH Biosciences). These synthetic, low-protein media are recommended for secreted proteins and facilitate subsequent purification.

3. *Incubator at 27°C ± 1°C.* CO₂ is not required and humidification is optional. The Biological Oxygen Demand (B.O.D.) low-temperature incubator (VWR Scientific) or the larger Isotemp (Fisher) are good examples.

4. *Magnetic spinner flasks.* These are available in a variety of sizes from Techne or Bellco.

5. *Stir plate for multiple spinners.* This is available from Techne or Bellco.

6. *SeaKem ME agarose.* (FMC Bioproducts).

7. *60-mm, 100-mm, and 150-mm tissue culture plates.* (Falcon or Corning).

8. *Antibiotics (optional).* Gentamicin (available from numerous vendors) and amphotericin B (Fungizone from Flow Laboratories) are used.

9. *Microscope.* Either an inverted light microscope or a dissecting microscope is required.

10. *Appropriate cloning vectors.* These are available from many vendors (see Table 16.9.1). Several additional vectors, a manual of methods, and wild-type baculovirus DNA are also available upon request from Dr. Max D. Summers, Department of Entomology, Texas Agricultural Experiment Station, Texas A & M University, College Station, Texas 77843 (Phone: 409-845-9730). It is necessary to sign a licensing agreement before the material will be sent. Commercial kits are available from Pharmingen, Invitrogen, Clontech, Novagen and Stratagene.

11. *Linearized ready-to-use baculovirus DNA.* This can be purchased from many of the same vendors.

12. *Shaking incubators (90 to 150 rpm)* can also be used to culture Sf9 cells. In this case, disposable polycarbonate Erlenmeyer flasks (Corning) or Fernbach flasks with screw caps

(Bellco) should be used instead of magnetic spinner flasks.

LITERATURE CITED

- Carbonell, L.F., Klowliden, M.J., and Miller, L.K. 1985. Baculovirus-mediated expression of bacterial genes in dipteran and mammalian cells. *J. Virol.* 56:153-160.
- Davidson, D.J.C., Fraser, M.J., and Castellino, F.J. 1990. Oligosaccharide processing in the expression of human plasminogen cDNA by lepidopteran insect (*Spodoptera frugiperda*) cells. *Biochemistry* 29:5584-5590.
- Davis, T.R. and Wood, H.A. 1995. Intrinsic glycosylation potentials of insect-cell cultures and insect larvae. *In Vitro Cell. Dev. Biol. Animal* 31:659-663.
- Doerfler, W. and Bohm, P. 1986. *The Molecular Biology of Baculoviruses*. Springer-Verlag, New York.
- Grabenhorst, E., Hofer, B., Nimitz, M., Jager, V., and Contradt, H.S. 1993. Biosynthesis and secretion of human interleukin-2 glycoprotein variants from baculovirus-infected insect *Sf21* cells: Characterization of polypeptides and post-translational modifications. *Eur. J. Biochem.* 215:189-197.
- Hellers, M., Gunne, H., and Steiner, H. 1991. Expression and post-translational processing of prepro-cropin-A using a baculovirus vector. *Eur. J. Biochem.* 199:435-439.
- James, D.C., Freedman, R.B., Hoare, M., Ogonah, O.W., Rooney, B.C., and Larionov, O.A. 1995. N-glycosylation of recombinant human interferon- γ produced in different animal expression systems. *BioTechnology* 13:592-596.
- Jarvis, D.L. and Finn, E.E. 1995. Biochemical analysis of the N-glycosylation pathway in baculovirus-infected lepidopteran insect cells. *Virology* 212:500-511.
- Jarvis, D.L. and Summers, M.D. 1992. Baculovirus expression vectors. In *Recombinant DNA Vaccines: Rationale and Strategies* (R.E. Isaacson, ed.) pp. 265-291. Marcel Dekker, New York.
- Kitts, P.A. and Possee, R.D. 1993. A method for producing recombinant baculovirus expression vectors at high frequency. *BioTechniques* 14:810-817.
- Livingston, C. and Jones, I. 1989. Baculovirus expression vector with single-strand capability. *Nucl. Acids Res.* 17:2366.
- Luckow, V.A. 1991. Cloning and expression of heterologous genes in insect cells with baculovirus vectors. In *Recombinant DNA Technology and Applications* (A. Prokop, R.K. Bajpai, and C. Ho, eds.) pp. 97-152. McGraw-Hill, New York.
- Luckow, V.A. and Summers, M.D. 1988. Trends in the development of baculoviral expression vectors. *BioTechnology* 6:47-55.
- Matsuura, Y., Possee, R.D., and Bishop, D.H.L. 1987. Baculovirus expression vectors: The requirements for high level expression of proteins, including glycoproteins. *J. Gen. Virol.* 68:1233-1250.
- Miller, L.K. 1988. Baculoviruses as gene expression vectors. *Annu. Rev. Microbiol.* 42:177-199.
- Miller, D.W., Safer, P., and Miller, L.K. 1986. An insect baculovirus host vector for high-level expression of foreign genes. In *Genetic Engineering*, Vol. 8 (J.K. Setlow and A. Hollaender, eds.) pp. 277-298. Plenum, New York.
- Murphy, C.I., McIntire, J.R., Davis, D.R., Hodgdon, H., Seals, J.R., and Young, E. 1993. Enhanced expression, secretion, and large-scale purification of recombinant HIV-1 gp120 in insect cells using the baculovirus egt and p67 signal peptides. *Protein Expr. Purif.* 4:349-357.
- Ogonah, O.W., Freedman, R.B., Jenkins, N., Patel, K., and Rooney, B.C. 1996. Isolation and characterization of an insect cell line able to perform complex N-linked glycosylation on recombinant proteins. *BioTechnology* 14:197-202.
- O'Reilly, D.R., Miller, L.K., and Luckow, V.A. 1992. *Baculovirus Expression Vectors*. W.H. Freeman, New York.
- Piwnica-Worms, H., Williams, N.G., Cheng, S.H., and Roberts, T.M. 1990. Regulation of pp60 c-src and its association with polyoma virus middle T antigen in insect cells. *J. Virol.* 64:61-68.
- Summers, M.D. and Smith, G.E. 1987. *A manual of methods for baculovirus vectors and insect cell culture procedures*. Texas Agricultural Experiment Station Bulletin No.1555. College Station, Texas.
- Vialard, J., Lalumiere, M., Vernet, T., Briedis, D., Alkhatib, G., Henning, D., Levin, D., and Richardson, C. 1990. Synthesis of the membrane fusion and hemagglutinin proteins of measles virus, using a novel baculovirus vector containing the β -galactosidase gene. *J. Virol.* 64:37-50.

KEY REFERENCE

O'Reilly et al., 1992. See above.

A guide assembled to aid researchers using the baculoviral expression system, containing detailed protocols for using this system effectively.

Contributed by Cheryl Isaac Murphy
Aquila Biopharmaceuticals
Worcester, Massachusetts

Helen Piwnica-Worms
Washington University School of Medicine
St. Louis, Missouri

Stefan Grünwald
Pharming
San Diego, California

William G. Romanow
Corning Costar
Portsmouth, New Hampshire

Nicole Francis and Hua-Ying Fan
Massachusetts General Hospital
Boston, Massachusetts

Maintenance of Insect Cell Cultures and Generation of Recombinant Baculoviruses

UNIT 16.10

This unit describes the maintenance and care of insect cell cultures as well as the generation, purification, and storage of recombinant baculoviruses. Procedures are included for maintenance and subculturing of insect cells (see Basic Protocol 1) and cotransfection of insect cells with linearized baculovirus DNA and recombinant transfer plasmid containing the gene of interest (see Basic Protocol 2). Bacmid DNA may be prepared for this procedure (see Alternate Protocol 1). In the event that the linearized virus is not available, wild-type baculovirus (AcMNPV) DNA may be used to produce recombinant baculoviruses (see Alternate Protocol 2). A procedure is also included for the generation of recombinant baculoviruses using a novel method, direct cloning (see Alternate Protocol 3), which eliminates the need to first clone the gene of interest into a baculoviral transfer vector. Preparation of baculovirus infection stocks from both monolayer and suspension cultures is also described (see Basic Protocol 3). Finally, a protocol is given for a plaque assay to be used for determining the titer of baculoviral stocks as well as for selection of recombinants and plaque purification (see Basic Protocol 4).

NOTE: All reagents and equipment coming into contact with live cells must be sterile and proper sterile technique should be used accordingly.

MAINTENANCE AND CULTURE OF INSECT CELLS

This protocol describes how to maintain and subculture *Spodoptera frugiperda* (*Sf9*) cells in both monolayer and suspension (spinner) cultures, in either serum-containing or serum-free medium. A culture of insect cells is begun using frozen *Sf9* cells. Cultures are maintained by subculturing and their viability is checked periodically. Aliquots of these cultures can be frozen in a liquid nitrogen freezer for long-term storage.

Materials

- TNM-FH insect medium (see recipe) containing 10% fetal bovine serum (FBS), with and without 20% (v/v) DMSO
- Spodoptera frugiperda* (*Sf9*) cells (ATCC #CRL 1711) derived from fall armyworm ovaries (also see UNIT 16.9)
- 70% ethanol
- 0.4% trypan blue stain (Life Technologies)
- Serum-free insect cell culture medium (BaculoGold Protein-Free Insect Medium from Pharmingen; *Sf*-900 II from Life Technologies; HyQ-CCM3 and HyQSFX-Insect from Hyclone; or ExCell 401 from JRH Biosciences)
- 60-mm tissue culture plates or 25-cm² flasks
- 27°C incubator (humidification optional)
- Spinner culture flasks (for suspension cultures) and stir plate for multiple spinner flasks (all available from Techne or Bellco) *or* disposable (plastic) shaker flasks (Corning) and shaking incubator set to 27°C, 90 to 150 rpm
- Beckman GPR centrifuge with GH-3.7 horizontal rotor (or equivalent)
- Screw-top cryostat freezing vials
- Liquid nitrogen freezer
- Additional reagents and equipment for counting cells with a hemacytometer (APPENDIX 3F)

**BASIC
PROTOCOL 1**

**Protein
Expression**

Contributed by Cheryl Isaac Murphy, Helen Piwnica-Worms, Stefan Grünwald, William G. Romanow, Nicole Francis, and Hua-Ying Fan

Current Protocols in Molecular Biology (2004) 16.10.1-16.10.19

Copyright © 2004 by John Wiley & Sons, Inc.

16.10.1

Supplement 65

Begin culture of Sf9 cells

1. Place 3 ml TNM-FH medium containing 10% FBS in a 60-mm tissue culture plate or 25-cm² flask.
2. Thaw a frozen ampule of *Sf9* cells rapidly in a 37°C water bath by moving it back and forth by hand. When ampule contents are almost completely thawed, immerse ampule in 70% ethanol to sterilize the outside.
3. Break the neck of the ampule and transfer contents to 60-mm tissue culture plate or 25-cm² flask from step 1. Rock plate gently by hand to distribute the cells evenly and incubate 2 to 3 hr at 27°C until cells have attached.
4. Remove old medium and replace with 3 ml fresh TNM-FH/10% FBS. Continue incubation, feeding culture every 3 days (by removing old medium and replacing with fresh) until cells reach confluency (form a packed monolayer).

It is important to hold the plates at an angle and remove and add medium at one corner so as not to dislodge the cells from the monolayer.

Maintain and subculture monolayer cultures

5. Remove the old medium from a confluent plate or flask of *Sf9* cells and resuspend cells by gently spritzing them with medium from a pipet.
6. Count the cells using a hemacytometer designed for tissue culture cells (*APPENDIX 3F*).
Each cell in a small square of the hemacytometer is equivalent to 10⁴ cells/ml.
7. Seed 1–2 × 10⁶ cells from step 5 in new 60-mm plates or 25-cm² flasks and rock to evenly distribute the cells (or use a larger plate or flask with more medium if preparing a larger culture of cells). Add fresh TNM-FH/10% FBS to bring the volume to 3 ml.
8. Incubate at 27°C, feeding the culture every 3 days with TNM-FH/10% FBS, until the cells reach confluency.

Maintain and subculture suspension cultures

9. Remove medium and resuspend cells from confluent monolayer culture as described in step 5. Count the cells using a hemacytometer (*APPENDIX 3F*).
- 10a. Seed cells in a spinner culture flask at ~4–5 × 10⁵ cells/ml. Incubate at 27°C with constant stirring on a stir plate set at 60 to 80 rpm. Leave the side-arm caps slightly loosened to ensure adequate aeration.
- 10b. Alternatively, seed cells in a shaker flask and incubate at 27°C in a shaking incubator. When initially adapting cells, shake at 80 to 90 rpm. Increase by 10 rpm each time cells are split until cultures are shaking at 150 rpm.

The volume of cell suspension in the shaker flasks should not exceed 40% of the total capacity.

11. Count cells every 2 to 3 days using a hemacytometer (*APPENDIX 3F*). Subculture when cells reach a concentration of 2–2.5 × 10⁶ cells/ml by transferring the appropriate number of cells to a new flask containing fresh TNM-FH/10% FBS to achieve a final density of 4–5 × 10⁵ cells/ml.

Alternatively, pour out the appropriate volume of cell suspension and replace it with fresh medium.

12. Determine cell viability by adding 0.1 ml of 0.4% trypan blue to 1 ml log-phase cells and examining the cells under a microscope at low power. Count the number of cells that take up trypan blue (dead cells) and count the total number of cells, then calculate the percentages of dead cells and viable cells.

A healthy culture of cells should be >97% viable. To maintain a sufficient transfer of oxygen to the cells in suspension, a minimum ratio of surface area to volume of culture must be maintained (Maiorella et al., 1988). If this value decreases, the cells will not grow exponentially and will stop growing at a lower cell density. This ratio is adequate for a 100-ml culture grown in a 100-ml spinner flask, but decreases when larger spinner flasks are used at maximum volume. Thus, to maintain sufficient oxygen transfer, smaller volumes should be used in the larger flasks unless an outside source of air is introduced into the flask (UNIT 16.11).

Adapt cells to serum-free medium

13. Subculture monolayer cells (from step 5) or suspension cells (from step 11) into medium composed of one part complete TNM-FH/10% FBS and one part serum-free medium (BaculoGold, Sf-900 II, or ExCell 401). Allow cells to grow to confluency (monolayer cultures) or to a density of $2\text{--}3 \times 10^6$ cells/ml (suspension cultures).

Other commercially available serum- or protein-free media besides BaculoGold, Sf-900 II, and ExCell 401 may be used. The final choice of serum-free medium should be based on a comparison of cell growth curves and production of recombinant protein in different media.

14. Repeat the subculture and growth procedure as in step 13 using a medium composed of one part FBS-containing complete medium and three parts serum-free medium.
15. Repeat the subculture and growth procedure as in step 13 using a medium composed of one part FBS-containing complete medium and between 7 and 9 parts serum-free medium.
16. Subculture the cells into serum-free medium.

Cells may adapt slowly to serum-free medium and may require several passages before growth rates and viability return to normal.

Freeze cells

17. Count cells to be frozen from an exponentially growing culture using a hemacytometer (APPENDIX 3F).
18. Centrifuge cells 10 min at $1000 \times g$ (2000 rpm in a GH-3.7 rotor), room temperature, and discard supernatant.
19. Resuspend cell pellet at $1\text{--}2 \times 10^7$ cells/ml in TNM-FH/10% FBS. Add an equal volume of TNM-FH/10% FBS containing 20% DMSO and place cells on ice. Dispense 1-ml aliquots of this cell suspension into screw-top cryostat freezing vials and incubate 1 hr at -20°C , then overnight at -70°C .

Alternatively, cells can be frozen using serum-free medium (with and without DMSO) if the cells have been adapted to serum-free medium as described in steps 13 to 16.

20. Transfer frozen cells to a liquid nitrogen freezer for long-term storage.

COTRANSFECTION OF INSECT CELLS USING LINEARIZED BACULOVIRAL DNA

There are several protocols available for preparing baculoviral supernatant, including transfection with bacmid DNA, transfection with linearized baculoviral DNA, and cotransfection with wild-type baculoviral DNA. Reagents and protocols for these methods are available commercially, and are summarized below. In general, the authors have found that preparation of bacmid DNA is the simplest means of generating baculovirus (Bac-to-Bac system, Invitrogen; see Alternate Protocol 1); however, cotransfection with linearized baculoviral DNA is still also used widely.

**BASIC
PROTOCOL 2**

**Protein
Expression**

16.10.3

One of the most commonly used methods of introducing baculovirus and transfer-plasmid DNA into susceptible insect cells is to coprecipitate the DNA with calcium phosphate and present the mixture to insect cells. For cotransfection, prepare $\geq 10 \mu\text{g}$ of purified plasmid DNA. Care must be taken that the plasmid be as clean as possible. With impure plasmids, cells may lyse shortly after transfection, resulting in a lower viral titer. At ~ 24 hr post-transfection, *Sf9* cell viability should be greater than 97%.

It should be noted that this protocol is optimized for use with Pharmingen linear DNA. Clontech and Invitrogen (see APPENDIX 4) have their own protocols for use with the linear DNA that they offer.

Materials

Spodoptera frugiperda (*Sf9*) cells growing in tissue culture at 50% to 70% confluence *or* growing in suspension culture at $1\text{--}1.5 \times 10^6$ cells/ml (see Basic Protocol 1)

TNM-FH insect medium (see recipe) containing 10% fetal bovine serum (FBS)

Linearized ORF 1629–deleted AcMNPV DNA (e.g., BaculoGold from Pharmingen; see Fig. 16.9.3)

Recombinant baculovirus transfer vector containing gene of interest (UNIT 16.9)

Transfection buffer B (see recipe)

Control transfer vector: pVL1392-Xyle (Pharmingen)

Transfection buffer A (see recipe)

500 mM catechol/50 mM sodium bisulfate

60-mm tissue culture plates

27°C incubator (humidification optional)

Inverted microscope

15-ml conical centrifuge tubes

Beckman GPR centrifuge with GH-3.7 horizontal rotor (or equivalent), 4°C

Additional reagents and equipment for amplification of viral supernatants to produce baculoviral stocks (see Basic Protocol 3) and plaque assay of baculovirus (see Basic Protocol 4)

Prepare cells and DNA

1. In each of three 60-mm tissue culture plates seed 2×10^6 *Sf9* cells in TNM-FH medium containing 10% FBS. Incubate in a 27°C incubator until cells attach.

Cell attachment should be done on a flat and even surface, allowing the cells to attach firmly, which usually takes ~ 5 min. If cells do not attach after that time, either they are not healthy or the wrong plates (e.g., petri dishes that have not been tissue culture treated) have been used.

The quality of the insect cells is very important and only rapidly dividing cells should be used.

2. While cells are attaching, combine in a microcentrifuge tube 0.5 μg linearized ORF 1629–deleted AcMNPV DNA and 2 to 5 μg recombinant baculovirus transfer vector containing the gene of interest. Mix well by gently vortexing or flicking the tube. Let mixture sit 5 min, then add 1 ml transfection buffer B.
3. Prepare positive control by combining, in a microcentrifuge tube, 0.5 μg linearized ORF 1629–deleted AcMNPV DNA and 2 μg pVL1392-Xyle control transfer vector DNA. Mix well by gently vortexing or flicking the tube. Let mixture sit 5 min, then add 1 ml transfection buffer B.

4. Label the first plate (from step 1) as the cotransfection plate. Aspirate old medium and replace with 1 ml transfection buffer A, making sure that the entire surface of the plate is covered to prevent the cells from drying out.
5. Label the second plate (from step 1) as the positive control. Aspirate old medium and replace with 1 ml of transfection buffer A, as in step 4.
6. Label the third plate (from step 1) as negative control. Aspirate old medium and replace with 3 ml fresh TNM-FH/10% FBS, without adding any DNA.

Transfect cells

7. Add 1 ml of the solution prepared in step 2 (containing the vector with the gene of interest) drop-by-drop to the cotransfection plate. After every three to five drops, gently rock the plate back and forth to mix the drops with the medium.

During this procedure, a fine calcium phosphate/DNA precipitate should form. The quality of the precipitate that is optimal for transfection can be assessed visually. It should be of a fine white milky appearance.

8. Add 1 ml of the solution prepared in step 3 (containing the positive control vector) drop-by-drop to the positive control plate, repeating the procedure in step 7.
9. Incubate all three plates 4 hr in a 27°C incubator.

The time of exposure to the calcium precipitate is critical for optimal transfection results. If the incubation time is too long, cell viability will be dramatically reduced. For different cell lines, the optimal incubation time varies. For Sf9 cells, the optimal time is 4 hr.

10. After 4 hr, remove the medium from the cotransfection plate and the positive control plate (but not the negative control plate). Add 3 ml fresh TNM-FH/10% FBS to each plate, rock the plate back and forth several times, then remove all the medium again. Add 3 ml of fresh TNM-FH/10% FBS to each plate and incubate all three plates 4 to 5 days at 27°C.

It is not necessary to change the medium of the negative control plate.

Check for successful transfection

11. After 4 days, check the three plates for signs of infection using an inverted microscope. Compare the negative and positive controls to the cotransfection plate.

Infected cells are much larger than uninfected cells and have enlarged nuclei. Because they stop dividing early in infection, their cell density will be much lower as compared to the uninfected population. Furthermore, infected cells do not attach well to the plate and a high percentage of them will float in the medium. Many of these infection signs may not be visible at this time because the virus titer during the cotransfection is usually low. A further amplification step (see below) may be needed to visualize these changes.

12. After 5 days, collect the supernatants of the cotransfection and positive control plates. Determine viral titer by plaque screening (see Basic Protocol 4).

Alternatively, cotransfection efficiency may be assessed by endpoint dilution assay (see Basic Protocol 4, step 7 annotation).

13. Check the expression of the protein of interest by lysing the transfected cells (for recombinant proteins that are not secreted) or using an aliquot of the supernatant (for recombinant proteins that are secreted), and performing an appropriate assay.

Unless a sensitive assay is available for the protein of interest, expression of the recombinant protein may not be detectable at this stage.

14. Assay for cells expressing the XylE protein in the positive control plate by adding 100 μ l of 500 mM catechol/50 mM sodium bisulfate.

Infected cells expressing XylE protein will turn bright yellow in ~5 min.

15. Transfer the transfection supernatants from each plate to sterile conical 15-ml centrifuge tubes and centrifuge 10 min at $1000 \times g$ (2000 rpm in GH-3.7 rotor), 4°C. Transfer viral supernatant to new sterile tubes and store at 4°C in the dark.

16. Amplify viral supernatant to produce a high-titer virus stock for production of the recombinant protein by infection of insect cells (see Basic Protocol 3).

Alternatively, a single recombinant virus, obtained by plaque purification (see Basic Protocol 4), may be used for virus amplification.

ALTERNATE PROTOCOL 1

GENERATION OF RECOMBINANT BACULOVIRUS BY PREPARATION AND TRANSFECTION OF BACMID DNA USING THE BAC-TO-BAC SYSTEM

The Bac-to-Bac manual from Invitrogen provides detailed protocols for preparation and transfection of bacmid DNA. The following is a brief overview of the protocol starting from the cDNA being cloned into a pFastBac vector. All protocols and reagents are available from Invitrogen (<http://www.invitrogen.com>).

Materials

cDNA of interest (e.g., UNIT 5.5)

Bac-to-Bac Baculovirus Expression System (Invitrogen) including:

pFastBac plasmid vector and control plasmid

DH10Bac competent *E. coli* cells

Bac-to-Bac manual (available online at http://invitrogen.com/content/sfs/manuals/bactobac_man.pdf)

Cellfectin lipid transfection reagent

Selection plates for transposition (see recipe)

Bacmid selection medium (see recipe)

Sf9 cells in serum-free medium

Serum-free medium

27°C incubator

6-well tissue culture plates

Additional reagents and equipment for cloning DNA into plasmids (UNIT 3.16), introduction of plasmid DNA into *E. coli* cells (UNIT 1.8), growing *E. coli* on solid (UNIT 1.3) and in liquid (UNIT 1.2) medium, alkaline lysis miniprep (UNIT 1.6), and preparation of insect cells for transfection (see Basic Protocol 2, steps 1 to 6, in this unit), and amplification of recombinant baculovirus to produce stocks (see Basic Protocol 3 in this unit)

1. Transpose cDNA into bacmid to create recombinant bacmid as follows (also see Bac-to-Bac manual).
 - a. Clone the cDNA of interest into the pFastBac plasmid vector (also see UNIT 3.16).
 - b. Transform plasmid DNA into DH10Bac competent cells (also see UNIT 1.8).

The DH10Bac competent cells carry the bacmid DNA and helper plasmid (encoding proteins required for transposition).
 - c. Incubate 4 hr at 37°C to allow transposition of the gene of interest into the bacmid.

2. Grow colonies (also see *UNIT 1.2*) for 24 hr on selection plates for transposition. Identify recombinant bacmid-containing clones by blue-white selection.

Transposition disrupts the bacmid lacZ α gene, enabling selection via Bluo-Gal, which is included in the selection plates. Also see UNIT 1.4.

3. Pick colonies and culture 24 hr in bacmid selection medium.
4. Prepare bacmid DNA by alkaline lysis miniprep (*UNIT 1.6*).
5. Prepare Sf9 cells (see Basic Protocol 2, steps 1 to 6, in this unit). Use Cellfectin lipid transfection reagent to transfect cells with the bacmid DNA (see Bac-to-Bac manual), then incubate for 3 days at 27°C.
6. Collect transfection supernatant (containing recombinant baculovirus) and amplify to produce baculovirus stock (see Basic Protocol 3).

GENERATION OF RECOMBINANT BACULOVIRUS USING WILD-TYPE BACULOVIRAL DNA

**ALTERNATE
PROTOCOL 2**

As an alternative to linearized ORF 1629–deleted baculoviral DNA, circular wild-type AcMNPV DNA can be used for cotransfection of insect cells with baculoviral transfer plasmids (also see Basic Protocol 2). However, recombination efficiency is dramatically lower (usually around 0.1% to 0.2%) as compared to that obtained with linearized DNA. The technique described here thus requires the identification and purification of recombinants by multiple rounds of plaque assay.

The following protocol describes how to isolate and purify AcMNPV wild-type baculoviral DNA, which can then be used to cotransfect susceptible insect cells (e.g., Sf9) with an appropriate plasmid vector to generate recombinant baculoviruses. The cotransfection of wild-type baculovirus DNA and recombinant transfer plasmid should be performed as described in Basic Protocol 2 (using wild-type baculoviral DNA instead of linearized ORF 1629–negative AcMNPV DNA). To facilitate screening between wild-type and recombinant viral plaques, pBlueBacIII (Invitrogen) or pVL1393-XylE (Pharmingen) can be used as a positive control, making it possible to visualize differences between nonrecombinant wild-type viral plaques and recombinant viral plaques.

Wild-type baculovirus available upon request from Dr. Max D. Summers, Department of Entomology, Texas Agricultural Experiment Station, Texas A & M University, College Station, Texas 77843 (Phone: 409-845-9730).

Additional Materials (also see *Basic Protocol 2*)

- Wild-type baculovirus (Dr. Max D. Summers, see above, or Pharmingen)
- Sucrose cushion solution (see recipe)
- 0.1 \times and 1 \times TE buffer, pH 7.4 (*APPENDIX 2*)
- 25% and 56% (w/v) sucrose (ultrapure) in 0.1 \times TE buffer (filter sterilize and store up to 1 month at 4°C)
- Extraction buffer (see recipe)
- 10 mg/ml proteinase K (prepare fresh)
- 10% *N*-lauroylsarcosine (sodium salt; filter sterilize and store up to 1 year at 4°C)
- 25:24:1 phenol/chloroform/isoamyl alcohol (*UNIT 2.1*)
- 100% (ice-cold) and 70% (room temperature) ethanol
- 3 M sodium acetate, pH 5.2 (*APPENDIX 2*)

**Protein
Expression**

16.10.7

150-mm tissue culture dishes
50-ml conical centrifuge tubes
Beckman GPR centrifuge with GH-3.7 horizontal rotor (or equivalent), 4°C
Beckman ultracentrifuge with SW-27 or SW-28 rotor *and* SW-41 rotor (or other ultracentrifuge with equivalent rotors), 4°C, and appropriate tubes
Sucrose gradient maker
15-ml polypropylene centrifuge tubes
50°C water bath
5- or 10-ml wide-mouth pipets

Additional reagents and equipment for phenol/chloroform extraction and ethanol precipitation of DNA (*UNIT 2.1*) and quantitating DNA by absorbance spectrometry (*APPENDIX 3D*)

Prepare viral supernatant and collect virus

1. Seed at least ten 150-mm plates with 2.0×10^7 Sf9 cells/plate in 30 ml TNM-FH medium containing 10% FBS. Incubate 1 hr at 27°C to allow the cells to attach firmly, then infect them with AcMNPV wild-type virus at a multiplicity of infection (MOI) of 0.1. Incubate 3 to 5 days at 27°C, examining plates periodically with an inverted microscope for the presence of occlusion bodies.

Dr. Max D. Summers provides a manual with his materials that includes details for these procedures. Pharmingen provides similar documentation with its wild-type baculovirus. Also see O'Reilly et al. (1992).

MOI is equal to plaque-forming units (pfu; see Basic Protocol 4) divided by the number of cells (pfu/cell). If the supplier of the virus does not provide adequate information regarding pfu, this can be determined by plaque assay (see Basic Protocol 4).

Occlusion bodies are highly refractile, giving them a yellowish-green crystalline appearance that is readily detected under a light microscope.

2. When occlusion bodies are observed in most cells, pool the viral supernatant (~30 ml per plate) in six 50-ml conical tubes. Centrifuge 10 min at $1000 \times g$ (2000 rpm in GH-3.7), 4°C, then pour viral supernatant into six new tubes. Repeat centrifugation to completely remove any remaining cells.

Sterile technique is not required in this step nor for any of the remaining steps of this protocol.

3. Place 6 to 35 ml of viral supernatant in each of an appropriate number of ultracentrifuge tubes for the SW-27 or SW-28 rotor and balance the tubes. Underlay with 3 ml sucrose cushion solution. Precool the ultracentrifuge and rotor to 4°C.

The amount of sucrose cushion may have to be increased for larger volumes of viral supernatant.

4. Centrifuge 60 min at $100,000 \times g$ (24,000 rpm in an SW-27 or -28 rotor), 4°C, to pellet the virus. Pour off supernatants and invert tubes on a Kimwipe to drain as much liquid as possible.

Separate virus from cellular contaminants (if necessary)

5. Examine the viral pellets carefully. If the viral pellet is pure (i.e., has a light bluish appearance), proceed to DNA isolation (step 11). If the pellet appears yellowish, separate the virus from cellular contaminants by steps 6a to 10a or 6b to 10b.

To purify viral pellet by sucrose-gradient fractionation

- 6a. Add 2 ml of $0.1 \times$ TE buffer to one of the viral pellets and repeatedly pipet up and down with a Pasteur pipet to resuspend. Transfer the buffer with the resuspended virus to a second tube containing a pellet and repeat the resuspension, then repeat in turn for each pellet until all pellets are pooled in the same 2 ml.

If pellet is difficult to resuspend, incubate preparation overnight at 4°C.

- 7a. Place 25% and 56% sucrose solutions in $0.1 \times$ TE buffer in the reservoirs of a gradient maker and prepare two linear 25% to 56% sucrose gradients in SW-41 ultracentrifuge tubes.

If a gradient maker is unavailable, simply layer the 25% sucrose carefully atop the 56% sucrose to form a step gradient. Some investigators report that a step gradient gives a sharper band.

- 8a. Carefully layer 1 ml of the pooled viral suspension on top of each sucrose gradient. Centrifuge 90 min at $100,000 \times g$ (28,000 rpm in an SW-41 rotor), 4°C.

After centrifugation, virus should be visible as a broad bluish-white band inside the gradient.

- 9a. Using a Pasteur pipet, transfer the viral bands to a new SW-41 ultracentrifuge tube. Add enough $0.1 \times$ TE buffer to fill the tube (~35 ml), then centrifuge 30 min at $100,000 \times g$ (28,000 rpm in an SW-41 rotor), 4°C, to pellet the virus. Decant supernatant and invert tube on a Kimwipe to drain as much liquid as possible.

- 10a. Resuspend the virus pellet in 9 ml extraction buffer and transfer 4.5-ml aliquots to two 15-ml polypropylene centrifuge tubes. Proceed to step 11.

To purify viral pellet by microcentrifugation

- 6b. Add 3 ml extraction buffer to one of the viral pellets and repeatedly pipet up and down with a Pasteur pipet to resuspend. Transfer the buffer with the resuspended virus to a second tube containing a pellet and repeat the suspension, then repeat in turn for each pellet until all pellets are pooled in the same 3 ml.

If pellet is difficult to resuspend, incubate preparation overnight at 4°C.

- 7b. Transfer 1.5 ml viral suspension into each of two 1.5-ml microcentrifuge tubes. Microcentrifuge 5 min at maximum speed and pool supernatants in one 15-ml polypropylene centrifuge tube.

- 8b. Resuspend each pellet in 1 ml extraction buffer.

- 9b. Microcentrifuge the pellets 5 min at maximum speed and combine the two supernatants with the pooled supernatants in the 15-ml tube.

- 10b. Bring volume in the 15-ml tube to 9 ml with extraction buffer and transfer 4.5-ml aliquots to two new 15-ml polypropylene centrifuge tubes. Proceed to step 11.

Isolate DNA from purified virions

11. Add 200 μ l of 10 mg/ml proteinase K to each tube and incubate 1 to 2 hr at 50°C.
12. Add 0.5 ml of 10% N-lauroylsarcosine to each tube and incubate 2 hr or overnight at 50°C.
13. Extract DNA twice with an equal volume of 25:24:1 phenol/chloroform/isoamyl alcohol (see UNIT 2.1 for additional details on extraction of DNA).

Extreme care should be taken to be as gentle as possible to avoid shearing the DNA at this point. Use a wide-bore Pasteur pipet and mix DNA solutions by inverting the tubes rather than by vortexing.

14. Transfer the aqueous phase containing the DNA to another 15-ml tube using a wide-mouth 5- to 10-ml pipet. Add 10 ml of ice-cold 100% ethanol to each tube and mix gently by inverting the tubes several times. Incubate 10 min at -80°C (see *UNIT 2.1* for additional details on ethanol precipitation of DNA).
15. Centrifuge 20 min at $1500 \times g$ (2500 rpm in GH-3.7 rotor), 4°C , and discard supernatant. Rinse the DNA pellet with 70% ethanol and air dry pellet for 30 to 60 min. Resuspend pellet in 800 μl of $1\times$ TE buffer.
16. Transfer 400 μl of the resuspended DNA to each of two microcentrifuge tubes and reprecipitate the DNA by adding 40 μl of 3 M sodium acetate and 2 vol of ice-cold 100% ethanol to each tube. Incubate 10 min at -80°C .
17. Microcentrifuge 10 min and discard supernatant. Rinse DNA pellet with 70% ethanol and air dry pellet. Resuspend DNA in 0.3 to 1.0 ml of $1\times$ TE buffer.
18. Quantitate DNA by measuring A_{260} (*APPENDIX 3D*) and calculate yield. Store the circular wild-type baculoviral DNA at 4°C (stable for several months).

This method should yield 50 to 100 μg viral DNA per ten 150-mm dishes. If difficulty is encountered resuspending the DNA, heat mixture for ~ 15 min at 65°C .

19. Cotransfect Sf9 cells with the circular wild-type baculoviral DNA and a baculoviral transfer plasmid containing a gene of interest as in Basic Protocol 2, substituting the wild-type DNA for the linearized ORF 1629–deleted DNA in step 2 of that protocol.

Recombination efficiency is dramatically lower with wild-type DNA as compared to that obtained with linearized DNA. The technique described here thus requires the identification and purification of recombinants by multiple rounds of plaque assay (see Basic Protocol 4).

ALTERNATE PROTOCOL 3

GENERATION OF RECOMBINANT BACULOVIRUSES BY DIRECT CLONING

This protocol describes the generation of recombinant virus by direct cloning methods (see Fig. 16.10.1), which may be applicable to the production of high-diversity expression libraries in baculovirus. Two modified AcMNPV baculoviruses—vEHuni and vECuni—have been constructed that have two *Bsu36I* sites downstream, respectively, of the *hsp70* promoter and the synthetic promoter *PcapminXIV* (Lu and Miller, 1996). Cleavage of the *Bsu36I* sites produces overhanging TTA ends, which are filled in by incubation with the Klenow fragment of DNA polymerase I in presence of dTTP, thus leaving a TT overhang. Ready-to-use vEHuni and vECuni baculoviral DNA can be purchased from Pharmingen. The gene to be cloned must be flanked by *EcoRI* sites and the *EcoRI* ends produced by cleavage at these sites have to be partially filled in with Klenow fragment in the presence of dATP to leave AA overhangs. The gene can then be cloned directly into the vEHuni or vECuni genome.

Additional Materials (also see Basic Protocol 2)

- Purified vector containing gene of interest, flanked by *EcoRI* sites
- Klenow fragment of DNA polymerase I (*UNIT 3.5*)
- Reaction buffer for Klenow fragment (see recipe)
- 25:24:1 phenol/chloroform/isoamyl alcohol (*UNIT 2.1*)
- $1\times$ TE buffer (*APPENDIX 2*)
- vEHuni or vECuni (Pharmingen), linearized, partially filled in by Klenow fragment treatment, and containing TTA ends
- T4 DNA ligase (*UNIT 3.14*)
- 15°C water bath

Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5A*) and phenol/chloroform extraction and ethanol precipitation of DNA (*UNIT 2.1*)

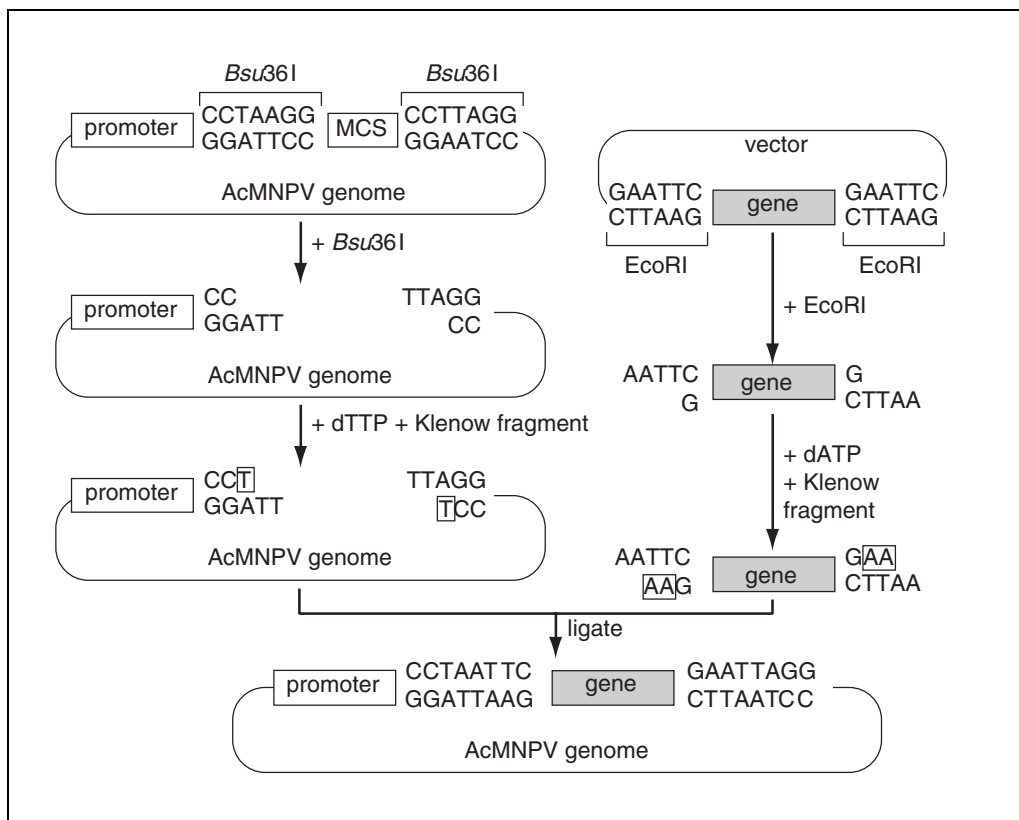


Figure 16.10.1 Direct cloning of a gene of interest into baculovirus DNA. At the left is a diagram representing an AcMNPV recombinant containing two different *Bsu361* sites. Digestion of this viral DNA with *Bsu361* followed by a partial fill-in reaction with dTTP and Klenow fragment of DNA polymerase I generates a linear viral DNA with TT overhanging ends. At the right, a foreign gene with flanking *EcoRI* sites is digested with *EcoRI* to generate overhanging ends, which are then partially filled in using dATP. The resulting AA overhanging ends are then compatible with the TT overhanging ends of the viral DNA. The viral DNA and foreign gene DNAs are then combined, ligated, and transfected into insect cells.

1. Digest vector containing gene of interest with *EcoRI* in a total volume of 50 μ l for 16 hr at 37°C and isolate gene of interest that contains flanking *EcoRI* sites by agarose gel electrophoresis (UNIT 2.5A).

If gene of interest does not contain flanking EcoRI sites, a polymerase chain reaction with specific primers containing EcoRI sites can be used to insert flanking EcoRI sites. Alternatively, EcoRI linkers can be ligated to the purified gene.

2. In a total volume of 50 μ l, incubate up to 5 μ g of *EcoRI*-digested and purified gene of interest with 5 U Klenow fragment in reaction buffer for Klenow fragment for 20 min at 37°C.
3. Extract reaction mix with 1 vol of 25:24:1 phenol/chloroform/isoamyl alcohol, then ethanol precipitate DNA and resuspend in 10 μ l of 1 \times TE buffer (UNIT 2.1).
4. In a total volume of 50 μ l, mix 0.5 μ g vEHuni or vECuni DNA with 0.1 to 1 μ g of the treated gene fragment from step 3 (to obtain an ~1:60 molar ratio of baculoviral DNA to gene fragment), and add 2 U T4 DNA ligase. Incubate overnight at 15°C.

The ligated AcNPV DNA is now ready to be transfected into susceptible insect cells.

5. Transfect the entire ligation mixture from step 4 into 2×10^6 *Sf9* cells as described in Basic Protocol 2.

The ligation mixture from this protocol replaces the reaction mixture composed of ORF 1629–deleted AcMNPV DNA and recombinant baculovirus transfer vector used in step 2 of Basic Protocol 2. As in Basic Protocol 2, use pVLI392-XylE as a control.

To purify recombinant plaques obtained by direct cloning, see Basic Protocol 4.

PREPARATION OF BACULOVIRUS STOCKS

This protocol describes how to prepare a large-scale stock of wild-type or recombinant AcMNPV virus from either monolayer or suspension culture. Growing insect cells (*Sf9*) are infected with virus at a low MOI (<1.0). A viral stock is obtained by harvesting the culture supernatant when the majority of cells show cytopathic effects (~4 to 5 days postinfection). Viral stocks can be stored at 4°C, but must be shielded from light to maintain the viral titer. Liquid nitrogen freezing is recommended for long-term storage. Plaque assay should be used to determine the titer of viral stocks.

Materials

Sf9 cells in monolayer culture (see Basic Protocol 1, step 5) or suspension culture (see Basic Protocol 1, step 11)

TNM-FH insect medium (see recipe) containing 10% fetal bovine serum (FBS)

Baculoviral inoculum, wild-type (*UNIT 16.9*) or recombinant (e.g., supernatant from cotransfected *Sf9* cells; see Basic Protocol 2), pfu determined by plaque assay (see Basic Protocol 4)

150-mm tissue culture dishes

27°C incubator (humidification optional)

Inverted microscope

Beckman GPR centrifuge with GH-3.7 horizontal rotor (or equivalent), 4°C

Screw-top cryostat freezing vials

Liquid nitrogen freezer

Spinner culture flasks (for suspension cultures) and stir plate for multiple spinner flasks (all available from Techne or Bellco) *or* disposable (plastic) shaker flasks (Corning) and shaking incubator set to 27°C, 90 to 150 rpm

To amplify virus from monolayer cultures

- 1a. Seed two 150-mm plates with 1.8×10^7 *Sf9* cells/dish maintained in TNM-FH medium containing 10% FBS. Incubate 1 hr in a 27°C incubator to allow cells to attach.
- 2a. While cells are attaching, add the viral inoculum to 30 ml of fresh TNM-FH/10% FBS in sufficient quantity to achieve an MOI of 0.1 to 1. When cells have attached, remove the medium from the plates and add 15 ml of this virus-containing medium to each plate.

MOI is defined as pfu/cell. The volume of viral inoculum needed to infect a given number of cells equals $MOI \times (\text{number of cells/titer of viral stock in pfu/ml})$.

- 3a. Incubate cells with virus-containing medium for several days at 27°C. Examine the cells daily under an inverted microscope for signs of infection—i.e., cytopathic effects (if occlusion body–negative recombinant viruses are used) or occlusion bodies (if wild-type baculovirus or occlusion body–positive recombinant viruses are used).

Occlusion bodies are highly refractile, giving them a yellowish-green crystalline appearance that is readily detected under a light microscope.

- 4a. When the majority of cells show cytopathic effects or occlusion bodies (usually 4 to 5 days postinfection) transfer the cells to sterile tubes, centrifuge 10 min at $1000 \times g$ (2000 rpm in a GH-3.7 rotor), 4°C , and transfer supernatant to new, sterile tubes. Dispense several 1-ml aliquots into screw-top cryostat freezing vials and freeze them in a liquid nitrogen freezer for long-term storage. Keep the remaining stock at 4°C in the dark for short-term storage.

This protocol should result in ~40 ml of a virus stock at 1×10^8 to 3×10^8 pfu/ml.

If the virus to be amplified is a single plaque isolate (see Basic Protocol 4), place the agarose plug in a microcentrifuge tube containing 0.5 ml TNM-FH/10%FBS and rotate the tube overnight at 4°C . Infect fresh Sf9 cells as in step 1a using 2×10^6 cells in a 100-mm flask and incubate 1 hr at 27°C . Add 8 ml of TNM-FH/10% FBS medium and incubate them for 4 days at 27°C . Harvest the virus as in step 4a. Determine the viral titer by plaque assay (see Basic Protocol 4) and amplify the virus by repeating steps 1a to 3a to obtain a high-titer stock. Virus stocks can also be prepared in serum-free insect cell culture medium if the cells have been adapted to serum-free medium (see Basic Protocol 1). Follow steps 1a to 4a but use serum-free culture medium instead of TNM-FH/10% FBS medium.

To amplify virus from suspension cultures

- 1b. Grow Sf9 cells in a 500-ml spinner culture bottle or a 250-ml shaker flask with 50 ml TNM-FH/10% FBS to a density of $\sim 1 \times 10^6$ cells/ml (see Basic Protocol 1).

Sf21 cells (see UNIT 16.9) may be used instead of Sf9 cells. They are grown in the same way as Sf9 cells.

- 2b. Centrifuge cells 10 min at $1000 \times g$, room temperature, and discard supernatant. Resuspend cell pellet in 10 to 20 ml fresh TNM-FH/10% FBS. Add viral inoculum at an MOI of 0.1 to 0.5.

See annotation to step 2a regarding how to calculate the volume of viral inoculum needed.

- 3b. Add cells back to spinner flask and bring volume to 100 ml with TNM-FH/10% FBS. Incubate 3 to 4 days at 27°C with constant stirring on a stir plate set at 60 to 80 rpm. Leave the side-arm caps slightly loosened to ensure adequate aeration. Periodically remove aliquots of the suspension and examine microscopically for cytopathic effects and occlusion bodies.

- 4b. When the majority of cells show cytopathic effects or occlusion bodies (usually 4 to 5 days postinfection) transfer the cells to sterile tubes, centrifuge 10 min at $1000 \times g$, 4°C , and transfer supernatant to new sterile tubes. Dispense several 1-ml aliquots into screw-top cryostat freezing vials and freeze them in a liquid nitrogen freezer for long-term storage. Keep the remaining stock at 4°C in the dark for short-term storage.

Virus stocks can also be prepared in suspension by growing insect cells (Sf9, Sf21, and High Five) in serum-free or protein-free insect cell culture medium in either a 100-ml spinner flask or a 500-ml spinner flask filled only to 250 ml until the cell density is $1\text{--}2 \times 10^6$ cells/ml. Virus should be added at an MOI of 0.1 to 0.5 directly to the suspension culture, the cells incubated 4 to 5 days, and the virus harvested as in step 4b.

TITERING BACULOVIRUS STOCKS USING PLAQUE ASSAY

This protocol describes how to titer viral stocks using the plaque-assay procedure. When preparing new viral stocks or when carrying out infections for protein production and protein analyses, it is important to know the titer of a viral stock, which is expressed in plaque-forming units per milliliter (pfu/ml). In this procedure, serial dilutions of viral stock are used to infect exponentially growing Sf9 cells. After removing the viral supernatant, the cells are covered with an agarose overlay and incubated 6 to 10 days. Plaques are counted visually to determine the viral titer. To become familiar with the morphological differences between plaques produced by wild-type virus versus those

BASIC PROTOCOL 4

Protein Expression

16.10.13

produced by recombinant virus, it is helpful to first practice plaquing β -Gal recombinant virus alongside wild-type virus (in a manner analogous to that used with plaque purification of bacteriophage; *UNIT 1.11*). This greatly facilitates screening of recombinant viruses encoding the gene of interest. It is possible to distinguish a wild-type plaque from a recombinant plaque by holding the plate overhead and looking at the bottom of the plate directly. Regions surrounding wild-type plaques will look grayish-white, whereas recombinant plaques will not. After identifying putative recombinant viral plaques by one of these methods, pick several plaques and place the agarose plugs in 1-ml serum-free medium. Vortex and store up to several months at 4°C until needed.

As a simple and rapid alternative to plaque assays, commercially available titer kits that use antibodies to viral proteins can be used (e.g., BacPAK viral titer kit, BD Biosciences).

Materials

Exponentially growing *Sf9* cells (see Basic Protocol 1) in monolayer culture
TNM-FH insect cell medium (see recipe) with and without 10% fetal bovine serum (FBS)

Baculoviral stock

Agarose overlay (prepare 30 min before use in step 5; see recipe)

Trypan blue overlay (optional; see recipe)

60-mm tissue culture plates

27°C incubator (humidification optional)

1.5-ml screw-top cryostat tube

1. Dilute a culture of exponentially growing *Sf9* cells to $\sim 5 \times 10^5$ cells/ml in TNM-FH medium containing 10% FBS. Seed cells onto 60-mm tissue culture plates at two different densities— 2×10^6 and 1.5×10^6 cells/plate—several hours before plaquing. Set up duplicate plates for each dilution of viral stock. Incubate in 27°C incubator.

Duplicates of each dilution are required because of the variability encountered using plaque assays. Thus, there will be four plates for each dilution: two with 2×10^6 cells and two with 1.5×10^6 cells. Generally, a total of three viral dilutions are plaqued, necessitating a total of twelve dishes for each stock being titered (six with 2×10^6 cells and six with 1.5×10^6 cells).

The density of the cell monolayer is critical to the success of the plaque assay. If the plaques are too small after 5 days, the initial cell density was too high. If the plaques are large and diffuse, the initial cell density was too low. Use a series of cell densities to find the optimal plaquing cell density for the particular cell line being used.

2. When ready to plaque, make 5-ml serial dilutions of the baculoviral stocks in TNM-FH medium/10% FBS as follows, depending on the nature of the viral stock:

high-titer viral stock— 10^{-6} , 10^{-7} , and 10^{-8} dilutions

transfection supernatant from circular AcMNPV DNA— 10^{-4} , 10^{-5} , and 10^{-6} dilutions

transfection supernatant from linear AcMNPV DNA— 10^{-1} and 10^{-2} dilutions

single plaque pick-ups— 10^{-1} , 10^{-2} , and 10^{-3} dilutions.

Making serial dilutions is analogous to titering phage stocks (see UNIT 1.11).

3. Remove medium from cells (from step 1) with a sterile pipet. Add 1 ml of each viral dilution to duplicate plates and incubate 1 hr at 27°C. Rock the plates when adding the virus inoculum to ensure even infection of the cells.

It is important to hold the plates at an angle and remove and add medium at one corner so not to dislodge the cells from the monolayer.

4. After 1 hr incubation, prepare the agarose overlay.

The agarose overlay keeps the virus released from the infected cells from diffusing far from the site where the initial infection took place. Thus, only cells immediately neighboring the cell of primary infection are subsequently infected by progeny virus. When enough cells in the immediate vicinity of the primary infection are lysed, a plaque (or hole in the cell monolayer) results.

5. Remove the viral supernatant from the cells with a sterile pipet and add 4 ml agarose overlay. Allow the agarose to harden on the plates for 10 to 20 min at room temperature (to allow condensation to escape). Wrap the plates individually with Parafilm (to avoid desiccation) and incubate 6 to 8 days at 27°C.

If the recombinant virus contains a lacZ gene—e.g., as in recombinants derived from pBlueBacIII (Invitrogen) or pAC360 β-Gal (available from Dr. Max Summers; see Alternate Protocol 2)—150 μg/ml of Xgal (from 20 mg/ml stock prepared in sterile dimethylformamide; stable for several months at –20°C) should be added to the agarose before it hardens. Recombinant plaques will then develop a bright blue color and will be easily distinguished from nonrecombinants.

If a humidified 27°C incubator is used, it is not necessary to wrap the plates with Parafilm.

If plaques are not clearly visible after 1 week, the plates can be incubated longer, because plaques continue to form for up to 2 weeks. If no plaques are visible after 2 weeks, the cells were probably plated at too high a density and should be replated at a lower density (between 1×10^6 and 1.3×10^6 cells/plate).

6. On plates containing plaques that are well formed and easily visualized, count the number of plaques at each dilution within a set. Calculate the viral titer (pfu/ml).

Ten plaques at a 10^{-7} dilution or one plaque at a 10^{-8} dilution gives a titer of 10^8 pfu/ml.

7. If difficulties are encountered visualizing plaques, stain with trypan blue as follows. Prepare the trypan blue overlay and dispense 1 ml on plates that have been incubated for 6 to 8 days so that plaques are well formed. Incubate the plates overnight at 27°C to allow the dye to diffuse into the dead cells. Count the number of blue plaques and determine the viral titer.

The dead cells within a plaque will take up the trypan blue dye but the surrounding live cells will exclude the dye.

Another procedure for determining virus titer is by endpoint dilution (O'Reilly et al., 1992). In this method, a series of viral dilutions are made and used to infect cells in microtiter wells. Each well is then scored for the presence or absence of viral infection and a 50% endpoint is determined. However, results from this method are often more difficult to interpret than plaque assays when titrating recombinant virus stocks. Wild-type virus is very easy to score because of the accumulation of occlusion bodies, but recombinant-virus infection can sometimes be difficult to score because of the lack of occlusion bodies.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Agarose overlay

- 40 ml 2× Grace's insect medium, supplemented (Life Technologies)
- 10 ml fetal bovine serum (FBS), heat-inactivated 30 min at 56°C
- 0.25 ml 10 mg/ml gentamicin (Life Technologies; optional)
- 50 ml 1% (w/v) agarose stock (sterile; see recipe)

continued

**Protein
Expression**

16.10.15

Mix the first three ingredients together, filter sterilize, and equilibrate to 42°C. Microwave the 1% agarose ~2 min until liquid and equilibrate to 42°C. Mix the two solutions just prior to the end of the 1-hr viral incubation (see Basic Protocol 4, step 3). Use immediately (this solution cannot be stored).

This is enough for 24 60-mm tissue culture dishes, using 4 ml/plate. The optimal temperature for equilibration is 42°C. If the overlay is too hot, the cells will lyse; if it is too cool the agarose will solidify into clumps, making visual screening difficult.

Agarose stock, 1% (w/v)

Place 0.5 g SeaKem agarose (FMC Bioproducts) into each of ten 100-ml Wheaton bottles. Add 50 ml distilled water to each bottle, close loosely, and autoclave 20 min. When the solutions have cooled, tighten lids. Store indefinitely at room temperature.

SeaKem agarose is recommended over SeaPlaque agarose for plaque assays because it can be used at a lower concentration. High concentrations of agarose tend to obscure the plaques or even kill the underlying insect cells.

Bacmid selection medium

LB medium (UNIT 1.1) containing:

50 µg/ml kanamycin

7 µg/ml gentamicin

5 µg/ml tetracycline

Extraction buffer

100 mM Tris·Cl, pH 7.5 (APPENDIX 2)

90 mM EDTA

200 mM KCl

Filter sterilize and store up to 1 year at 4°C

Reaction buffer for Klenow fragment

10 mM Tris·Cl, pH 7.5 (APPENDIX 2)

10 mM MgCl₂

10 mM dATP

Store up to 2 years at -20 °C

Selection plates for transposition

LB plates (UNIT 1.1) containing:

50 µg/ml kanamycin

7 µg/ml gentamicin

10 µg/ml tetracycline

100 µg/ml Bluo-Gal (Invitrogen)

40 µg/ml IPTG

Sucrose cushion solution

25% (w/v) sucrose

5 mM NaCl

10 mM EDTA

Store up to 1 month at 4°C

TNM-FH medium

To 500 ml 1× Grace's insect medium supplemented with yeastolate and lactalbumin hydrolysate (Life Technologies), add 5 ml 10 mg/ml gentamicin (optional), and 50 ml heat-inactivated fetal bovine serum (FBS; 10% final; see below). Filter sterilize. Incubate a 5-ml sample 2 days at 37°C to check sterility; store remainder at 4°C until manufacturer's expiration date.

continued

For medium with 20% DMSO, add sterilized (autoclaved) DMSO to 20% (v/v).

TNM-FH medium with and without FBS is needed for the protocols throughout this unit. TNM-FH medium including FBS and antibiotics can be obtained from several vendors (e.g., Pharmingen).

To heat-inactivate FBS, incubate 30 min at 56°C and store in 50-ml aliquots at -20°C (stable ≥1 year).

Transfection buffer A

To unsupplemented 1× Grace's insect medium (Life Technologies) add fetal bovine serum (FBS) to 10% and filter sterilize. Prepare fresh for each transfection and store in 10-ml aliquots.

Transfection buffer A is also sold by Pharmingen.

Transfection buffer B

25 mM HEPES, pH 7.1

125 mM CaCl₂

140 mM NaCl

Filter sterilize and store in 10-ml aliquots up to 6 months at 4°C

Transfection buffer B is also sold by Pharmingen.

Trypan blue overlay

Prepare a 1% (w/v) trypan blue solution in distilled water and filter sterilize. Microwave 1% agarose (see recipe) 2 min until liquid and equilibrate solution to 42°C. In parallel, equilibrate the 1% trypan blue solution to 42°C. Add 4 ml of the trypan blue solution to 50 ml of 1% agarose and mix well. Use immediately (this solution cannot be stored).

COMMENTARY

Background Information

Although recent research has provided more knowledge of the molecular biology of protein expression in insect cells using the baculovirus system, the ability of a given recombinant virus to produce large quantities of foreign proteins must still be determined empirically. Levels of expression have been reported to vary from 1 to 500 mg of recombinant protein/liter (assuming that 1 liter contains 2×10^9 cells). In general, recombinant proteins comprise 1% to 5% of total cell protein.

There are, however, several factors that may influence gene expression in the baculovirus system that should be taken into account when selecting a transfer plasmid and constructing the recombinant gene. Plasmids derived from the vector pVL941 (e.g., pVL1393 and pBlue-BacIII) contain an ATT in place of the original polyhedrin translation initiation codon ATG. Beames et al. (1991) have reported that translation can initiate at the ATT in these vectors. This results in the fusion of polyhedrin amino-terminal sequences to some of the heterologous proteins produced by these vectors if the foreign gene is cloned into the same reading frame

as the ATT. The lengths of the 5' and 3' untranslated regions of foreign genes expressed using this system have varied greatly. The significance of these sequences for transcriptional and translational efficiencies is unknown, and the standard practice is to keep their length to a minimum. One advantage of this expression system, however, is that it is unnecessary to precisely engineer sequences 5' to the initiation codon for efficient expression, as often must be done for bacterial expression systems, although the presence of an A at -3 (relative to AUG at +1, +2, and +3) may be important (O'Reilly et al., 1992). In addition, the inserted gene does not have to carry its own polyadenylation signal because the polyhedrin-gene polyadenylation signal is present in most expression vectors.

Because no intron-containing genes have been identified for baculovirus, it is recommended that inserted genes be derived either from cDNAs or genomic clones lacking introns. Reports in the literature, however, indicate that proper splicing may occur in some cases. For example, Jeang et al. (1987) reported preferential and proper processing of the small T antigen splicing signals when the SV40 virus

early region was cloned into the baculovirus expression vector.

Finally, if the recombinant protein contains a signal peptide, it may be expressed more efficiently in insect cells if it is fused to a heterologous signal peptide in place of the homologous one. The human immunodeficiency virus envelope protein gp120 is expressed and secreted at a much higher level when fused to signal peptides derived from baculovirus proteins (Murphy et al., 1993). Some transfer vectors that contain signal peptide coding regions are available commercially (see Table 16.9.1). However, this approach does not necessarily result in higher levels of protein production for every recombinant protein (Jarvis et al., 1993).

Critical Parameters and Troubleshooting

Sf9 cells have a doubling time of 18 to 24 hr in TNM-FH medium/10% FBS and should be subcultured two to three times per week. *Sf9* cells should be grown at 27°C. As carbon dioxide is not essential, the cells can actually be grown on the benchtop as long as room temperature remains within the range of 25° to 30°C. Medium should be equilibrated to room temperature before use. Gentamicin is optional and can be added to the medium if contamination is a problem.

Sf9 cells are adapted for growth in monolayer or suspension cultures. Trypsin is not required to detach them from monolayer cultures. *Sf9* cells may be repeatedly transferred between monolayer and suspension cultures without noticeable changes in viability or growth rate. It is easiest to carry cells and maintain their viability when they are carried in suspension rather than in monolayer cultures. A 50-ml culture of cells in a 100-ml spinner flask is routinely carried and used to subculture and seed monolayer cultures when needed.

For preparing viral stocks, low MOIs (e.g., 0.1) should be used. When this is the case the virus goes through multiple rounds of replication and infection and there is less chance for mutant viruses or defective interfering particles to form. When infecting for high levels of protein production, a high MOI (between 5 and 10) is recommended in monolayer cultures. For large suspension cultures (>100 ml), an MOI of 1 or 2 is sufficient. The virus is stable for 6 months at 4°C but must be shielded from light to maintain the titer (Jarvis and Garcia, 1994). Placement of 1 ml of each stock in liquid

nitrogen is recommended for long-term storage.

It is highly recommended that linear wild-type DNA be used for cotransfection or that bacmid DNA be prepared in *E. coli* and used for transfection. This greatly simplifies the screening procedure and saves time in obtaining a pure recombinant viral stock. If circular wild-type DNA is used, it is important to purify the recombinant viral DNA through several rounds of plaque assay.

Degradation by mechanical shearing and by nuclease contamination are the major problems encountered when trying to purify viral DNA from intact virions. Use of wide-mouth 5- or 10-ml pipets or Pasteur pipets with the necks broken off to handle the viral DNA helps to avoid shearing. In addition, tubes containing viral DNA should be mixed by inversion rather than by vortexing when performing phenol extractions and ethanol precipitations, and the purified DNA should be stored at 4°C rather than frozen and thawed.

It is very useful to practice visualizing recombinant plaques using the β -Gal recombinant virus as described in Basic Protocol 4. As mentioned, screening for recombinants is the most time-consuming and variable part of the entire expression system. The cell density at the time of plaquing seems to be the most critical parameter for achieving good plaques in the shortest period of time. Thus, it is helpful to seed duplicate plates at a couple of different cell densities. A perfect plaque assay will yield easily visualized plaques within 6 to 8 days.

Anticipated Results

Maintenance and culture of insect cells

When grown at 27°C, healthy insect cells should double every 18 to 24 hr. Healthy, logarithmically growing cells will be maintained in suspension cultures most readily when they are cultured at densities between 5×10^5 and 2×10^6 cells/ml. Cell viability should be ~97%.

Cotransfection

After 5 days, 20% to 100% of the cells should be showing signs of infection. Infected cells are of increased size with enlarged nuclei. They stop dividing and do not attach well to the plate. Many of the signs of infection may not be obvious if the transfection efficiency results in a low viral titer. In this case, it is necessary to amplify the virus in order to verify the success of the cotransfection. When generating recombinant baculovirus using wild-type bacu-

loviral DNA (see Alternate Protocol 2), the recombination frequency will be 0.1% to 0.2%. This method requires selection based upon identification of viral occlusion bodies. The direct cloning method (see Alternate Protocol 3) results in ~99% recombinant virus with ~50% containing the foreign gene in the correct orientation.

Plaques can generally be identified within 5 to 10 days post-infection and counted to determine viral titers. When individual plaques are picked to isolate recombinant virus, the resulting viral titer is generally low (5×10^5 pfu/ml) and will need to be amplified several times to produce a high-titer stock.

Time Considerations

Purification of wild-type viral DNA takes ~7 days to complete. Subcloning and large-scale production of recombinant DNA should take 1 to 2 weeks. The transfection procedure takes 1 day, and the transfection supernatant is harvested after 4 to 5 days. Plaquing takes 1 day, and it generally takes 1 week before the plaques are sufficiently formed to begin screening. Screening takes an additional day. It takes 4 to 5 days from the time of infecting Sf9 cells with a plaque to the time of harvesting the culture supernatant for an expanded virus stock. Purifying recombinants takes from <1 week to 3 weeks, depending on whether linear DNA was used in the transfection and whether several rounds of plaquing are required to purify recombinant virus from wild-type virus.

Using the Bac-to-Bac system (Invitrogen) to generate and transfect bacmid DNA, virus can be generated from a cloned cDNA in 7 days.

Literature Cited

- Beames, B., Braunagel, S., Summers, M.D., and Lanford, R.E. 1991. Polyhedrin initiator codon altered to AUU yields unexpected fusion protein from a baculovirus vector. *BioTechniques* 11:378-383.
- Jarvis, D.L. and Garcia, A. 1994. Long-term stability of baculoviruses stored under various conditions. *BioTechniques* 16:508-513.
- Jarvis, D.L., Summers, M.D., Garcia, A., and Bohlmeier, D.A. 1993. Influence of different signal peptides and prosequences on expression and secretion of human tissue plasminogen activator in the baculovirus system. *J. Biol. Chem.* 268:16754-16762.
- Jeang, K., Giam, C., Nerenberg, M., and Khoury, G. 1987. Abundant synthesis of functional T cell leukemia virus type I p40x protein in eucaryotic cells by using a baculovirus expression vector. *J. Virol.* 61:708-713.
- Lu, A. and Miller, L.K. 1996. Generation of recombinant baculoviruses by direct cloning. *BioTechniques* 21:63-68.
- Maiorella, B., Inlow, D., Shauger, A., and Harano, D. 1988. Large-scale insect cell culture for recombinant protein production. *BioTechnology* 6:1406-1410.
- Murphy, C.L., McIntire, J.R., Davis, D.R., Hodgdon, H., Seals, J.R., and Young, E. 1993. Enhanced expression, secretion, and large-scale purification of recombinant HIV-1 gp120 in insect cells using the baculovirus egt and p67 signal peptides. *Protein Expr. Purif.* 4:349-357.
- O'Reilly, D.R., Miller, L.K., and Luckow, V.A. 1992. *Baculovirus Expression Vectors*. W.H. Freeman, New York.

Key Reference

O'Reilly et al., 1992. See above.

A guide assembled to aid researchers using the baculoviral expression system, containing detailed protocols for using this system effectively.

Contributed by Cheryl Isaac Murphy
Aquila Biopharmaceuticals
Worcester, Massachusetts

Helen Piwnica-Worms
Washington University School of Medicine
St. Louis, Missouri

Stefan Grünwald
Pharming
San Diego, California

William G. Romanow
Corning Costar
Portsmouth, New Hampshire

Nicole Francis and Hua-Ying Fan
Massachusetts General Hospital
Boston, Massachusetts

Expression and Purification of Recombinant Proteins Using the Baculovirus System

UNIT 16.11

This unit describes how to analyze protein expression in cells infected with recombinant baculovirus on a small scale for optimizing protein production (see Basic Protocol 1 and Support Protocols 1 and 2), how to maximize and scale up recombinant protein production (see Basic Protocol 2), and how to purify recombinant proteins (see Basic Protocol 3; see Alternate Protocol). Before proceeding with large-scale expression, it is recommended that the putative recombinants obtained be assayed for their ability to produce the protein of interest. This will exclude the possibility of using a recombinant virus that does not produce the protein of interest, and will, in the long run, save time. Basic Protocol 1 details the expression of protein on a small scale for further analysis. In order to optimize protein expression, it is very important to determine the time course of maximum protein expression (see Support Protocol 1). If the recombinant protein is expressed in small quantities and cannot be visualized without radiolabeling, a metabolic labeling of infected insect cells is advised (see Support Protocol 2). After the recombinant protein expression has been characterized on a small scale, Basic Protocol 2 describes how to scale up protein expression. Methods are also described for co-infection with multiple baculoviruses for purification of recombinant protein complexes (see Alternate Protocol 1). If purification of the recombinant protein is needed and it is expressed as a fusion protein with either a polyhistidine (6×His), glutathione-*S*-transferase (GST), or FLAG tag, it can be purified in a single step according to Basic Protocol 3, or Alternate Protocols 2 or 3.

SMALL-SCALE EXPRESSION FOR INITIAL ANALYSIS

In this procedure, *Sf9* cells (UNITS 16.9 & 16.10) are infected with an expanded recombinant virus stock and can be analyzed 2 to 3 days later. The assays employed depend on the nature of the protein being produced. This protocol gives some suggested approaches but is certainly not comprehensive. Screening should be individually tailored to the properties of the protein being overproduced and the availability of detection reagents.

Materials

- Spodoptera frugiperda* (*Sf9*) cells (UNIT 16.10)
- TNM-FH insect medium (see recipe in UNIT 16.10) with or without 10% fetal bovine serum (FBS)
- High-titer recombinant baculovirus stocks (UNIT 16.10)
- Phosphate-buffered saline (PBS; APPENDIX 2)
- 1× SDS sample buffer (UNIT 10.2)
- 60-mm tissue culture plates
- 27°C incubator (humidification optional)
- 15-ml polypropylene centrifuge tubes
- Beckman GPR centrifuge with GH-3.7 rotor (or equivalent), 4°C
- Boiling water bath or 100°C heating block
- Sonicator
- Additional reagents for quantitating protein using the Bradford method (UNIT 10.1) and preparing insect cell cultures and viral stocks (UNIT 16.10)

Prepare and process cultures

1. Seed 2.5×10^6 *Sf9* cells into 60-mm tissue culture plates containing 3 ml TNM-FH insect medium with or without 10% FBS. Incubate 1 hr at 27°C to allow cells to attach.

Prepare one plate for each viral stock to be tested and one plate as an uninfected control.

BASIC
PROTOCOL 1

Protein
Expression

16.11.1

Contributed by Cheryl Isaac Murphy, Helen Piwnica-Worms, Stefan Grünwald, William G. Romanow, Nicole Francis, and Hua-Ying Fan

Current Protocols in Molecular Biology (2004) 16.11.1-16.11.14

Copyright © 2004 by John Wiley & Sons, Inc.

Supplement 65

2. Replace the medium with fresh TNM-FH/10% FBS medium and add 0.1 ml high-titer baculovirus stock to the appropriate plates at an MOI of 1 to 10. Incubate 3 days at 27°C.
3. Harvest cells by gently dislodging them from the plates, and transfer cells and culture medium to 15-ml polypropylene centrifuge tubes.
4. Centrifuge 10 min at 1000 × *g* (2000 rpm in a GH-3.7 rotor), 4°C.

For secreted proteins

- 5a. Transfer the culture supernatants to new tubes.
- 6a. Determine the protein concentrations in the supernatants using the Bradford method (UNIT 10.1). Proceed to analysis (step 7).

Supernatants may be frozen up to several months at –80°C.

For intracellular proteins

- 5b. Discard the supernatants. Rinse cells by resuspending the cell pellets gently in PBS, centrifuging again as in step 4, then discarding the supernatants.
- 6b. Add 500 µl of 1× SDS sample buffer to each pellet and boil 5 to 10 min by placing the tube in a boiling water bath or 100°C heating block. Sonicate samples if they are too viscous because of the presence of DNA. Continue to sonicate until viscosity clears, then determine the protein concentration in each sample using the Bradford method. Proceed to analysis (step 7).

Sonication times will vary with individual sonicators. DNA can also be disrupted by passing the lysate through an 18- or 20-G needle approximately ten times.

Alternatively, lyse cell pellets in 0.5 ml of an appropriate lysis buffer, which may differ according to cell type (e.g., insect cell lysis buffer; see Reagents and Solutions) supplemented with protease inhibitors. Lysis buffers containing EDTA must not be used with 6×His fusion products. Microcentrifuge 10 min at 4°C to clarify the lysates and transfer supernatants to new tubes. Add 100 µl of each lysate to 100 µl of 2× SDS sample buffer and boil 3 min in a boiling water bath. Freeze remaining lysate up to several months at –80°C.

Analyze proteins

7. Analyze the proteins in each sample by one of the following methods.
 - a. *Immunoblotting* (UNIT 10.8): Load 20 to 40 µg total cell protein per lane on a one-dimensional SDS-polyacrylamide gel. Remember to include the uninfected control.
 - b. *Coomassie brilliant blue staining* (UNIT 10.6): Load 20 to 40 µg total cell protein per lane on a one-dimensional SDS-polyacrylamide gel. If the recombinant virus is not pure, recombinant protein will be detected only if it is produced at very high levels in the infected cells.
 - c. *Functional assays*: Use any assay that is typically used to monitor the protein of interest—e.g., mobility-shift DNA-binding assays (for a DNA-binding protein; UNIT 12.2), in vitro kinase assays (for a protein kinase), nucleotide-binding assays (for a protein that binds nucleotides), or thymidine-incorporation assays (for a protein that is a growth factor).
 - d. *Metabolic labeling of recombinant proteins*: Perform as described (see Support Protocol 2).

If there are no easy assays for monitoring the recombinant protein produced by the baculovirus, then the putative recombinants should be monitored for the presence or

absence of the foreign gene. A simple dot hybridization technique is given in Summers and Smith (1987); alternatively PCR amplification can be used (Chapter 15 of this manual and O'Reilly et al., 1992). Several companies (e.g., Invitrogen and Clontech) sell PCR primers that will work for most baculovirus vectors.

8. Interpret results to identify which of the putative recombinant stocks is an actual recombinant that produces the desired protein. Plaque-purify recombinants so they are free from any contaminating wild-type virus (if they were not produced with linearized baculoviral or bacmid DNA; see *UNIT 16.10*). Prepare a large viral stock and titer the recombinant virus (*UNIT 16.10*).

DETERMINING TIME COURSE OF MAXIMUM PROTEIN PRODUCTION

Because expression of the recombinant protein is regulated by the polyhedrin promoter, which is activated very late in the lytic cycle of the virus, the recombinant protein will be expressed late as well. Recombinant proteins are usually detected between 15 and 24 hr postinfection and accumulate until ~40 hr postinfection, at which time their accumulation levels off. Because individual proteins display differences in their stability within insect cells, it is recommended that the time course of protein accumulation be charted for each protein expressed using this system. The following protocol describes how to determine when an intracellular recombinant protein is maximally produced by harvesting and analyzing cells at various times after infection.

Additional Materials (also see Basic Protocol 1)

Wild-type baculovirus (available from Dr. Max D. Summers; see *UNIT 16.10* for instructions on generating recombinant baculoviruses using wild-type baculoviral DNA)

1. Seed 3×10^6 Sf9 cells/plate into fifteen 60-mm tissue culture plates, each containing 3 ml TNM-FH medium with or without 10% FBS. Incubate 1 hr at 27°C to allow cells to attach, then infect seven plates with wild-type baculovirus and seven plates with recombinant virus, each at an MOI of 10. Leave one plate as an uninfected control.

Alternatively, Sf9 cells can be cultured in suspension in three 100-ml spinner or shaker flasks. When cells reach a density of 1.5×10^6 cells/ml, infect one flask with recombinant virus and one with wild-type virus, leaving one as an uninfected culture. Use an MOI of 1 to 2.

2. Harvest cells at various times (from ~15 to 72 hr) postinfection by transferring cells and culture supernatants to centrifuge tubes and centrifuging 10 min at $1000 \times g$ (2000 rpm in a GH-3.7 rotor), 4°C. Harvest the uninfected cells at 15 hr postinfection.

If cells are growing in suspension, remove 2-ml aliquots at each time point and process the cells and supernatants in the same way as for the monolayer cultures discussed here.

3. Process and analyze the cells (if the recombinant protein is not secreted) or supernatants (if the recombinant protein is secreted) as described (see Basic Protocol 1).

When staining with Coomassie brilliant blue (see Basic Protocol 1, step 7), look for a protein that appears as a function of time postinfection with the recombinant virus but not with the wild-type virus. The protein must be reasonably abundant for this method to be successful.

SUPPORT PROTOCOL 1

Protein Expression

16.11.3

METABOLIC LABELING OF RECOMBINANT PROTEINS

Metabolic labeling *in vivo* is a sensitive way to detect recombinant proteins, because at the time the recombinant protein is expressed, host protein synthesis is essentially terminated. Thus, all label is incorporated into late-viral-specific proteins, including the protein of interest. The most commonly used procedure for radiolabeling proteins is the incorporation of [³⁵S]methionine and [³⁵S]cysteine, both of which are essential amino acids. For better results, the intracellular pool of these two amino acids should be depleted prior to radiolabeling. This can be achieved by preincubating the cells in methionine/cysteine-free medium for 30 min. The efficiency of incorporation depends on the number of methionines and cysteines in the particular protein of interest. After labeling, the cells are lysed. The proteins are resolved by SDS-PAGE and visualized by autoradiography.

Additional Materials (also see Basic Protocol 1)

Wild-type baculovirus (available from Dr. Max D. Summers; see *UNIT 16.10* for instructions on generating recombinant baculoviruses using wild-type baculoviral DNA)

Methionine-free or methionine-free/cysteine-free Grace's insect cell culture medium (Invitrogen)

EXPRE³⁵S³⁵S, containing [³⁵S]methionine and [³⁵S]cysteine (>1000 Ci/mmol; Du Pont NEN)

Additional reagents and equipment for one-dimensional SDS-PAGE (*UNIT 10.2*) and autoradiography (*APPENDIX 3A*)

1. Seed 2.5×10^6 cells into 60-mm tissue culture plates containing 3 ml TNM-FH medium with 10% FBS. Prepare one plate to be infected with each putative recombinant virus and one control plate to be infected with wild-type baculovirus.
2. Incubate 1 hr at 27°C to allow cells to attach. Aspirate medium, then add 1 ml recombinant virus or 1 ml medium containing wild-type virus at an MOI of 5 to 10. Incubate 1 hr at room temperature.
3. Remove the medium from each plate by aspiration. Add 3 ml TNM-FH/10% FBS medium to the cells and incubate 24 to 48 hr at 27°C.
4. Carefully remove medium from each plate, then rinse cells once with methionine-free or methionine-free/cysteine-free medium. Add 1 ml methionine-free or cysteine-free medium to each plate. Incubate cells 30 min at 27°C, then add 0.25 to 0.5 μ Ci EXPE³⁵S³⁵S per plate and incubate 3 to 4 hr at 27°C.

It is important to hold the plates at an angle and remove and add medium at one corner so as not to dislodge the cells from the monolayer.

5. Transfer cells and culture supernatant from each plate to a separate 15-ml polypropylene centrifuge tube and centrifuge 10 min at $1000 \times g$ (2000 rpm in a GH-3.7 rotor), 4°C.
6. Process and analyze the cells (if the recombinant protein is not secreted) or the supernatant (if the recombinant protein is secreted) according to the appropriate procedures in Basic Protocol 1 (see Basic Protocol 1, steps 5a or 5b, 6a or 6b, and 7).

*If an antibody is available, it is recommended that the labeled lysates be immunoprecipitated (*UNIT 10.16*) prior to boiling in SDS sample buffer and resolution by SDS-PAGE. Immunoprecipitation will help detect the recombinant protein if it comigrates with a labeled host or late-viral-specific protein.*

*Polyhedra (*UNIT 16.9*) are solubilized only under very alkaline conditions (0.1 M final NaOH concentration). Without prior disruption under alkaline conditions, <10% of polyhedra will be solubilized in SDS sample buffer.*

7. Visualize proteins by autoradiography (APPENDIX 3A). Inspect the autoradiogram for protein of the expected molecular weight that appears in cells infected with recombinant baculovirus but not with wild-type baculovirus.

LARGE-SCALE PRODUCTION OF RECOMBINANT PROTEINS

This protocol describes how to maximize and scale up recombinant protein production in *Sf9* cells. The cells are grown in suspension in serum-free medium in 1- to 10-liter spinner flasks (Techne or Bellco) or in shaker flasks (Bellco). They are then infected with recombinant virus and harvested and analyzed at the time of optimum protein accumulation (see Support Protocol 1).

Materials

Spodoptera frugiperda (*Sf9*) cells (UNIT 16.10)

Serum-free insect cell culture medium (e.g., BaculoGold medium from Pharmingen, *Sf*-900 II from Invitrogen, or ExCell 401 from JRH Biosciences)

High-titer recombinant baculovirus (UNIT 16.10)

1- to 10-liter spinner (Techne or Bellco) or shaker (Bellco) flasks

Two-port cap assemblies for spinner flasks (Techne or Bellco; optional)

Silicone tubing (Cole-Palmer) with $\frac{3}{16}$ -in. (0.48-cm) inner diameter (i.d.), $\frac{5}{16}$ -in. (0.8-cm) outer diameter (o.d.), and $\frac{1}{16}$ -in. (0.16-cm) wall

0.2- μ m filter units (Millipore)

4-in. (10.16-cm) cable ties (Cole-Palmer)

Tension tool (Cole-Palmer)

Stir plate for multiple spinner flasks (Techne or Bellco; optional)

27°C incubator

Air-supply pump (Bellco)

Additional reagents and equipment for preparing insect cell cultures and viral stocks (UNIT 16.10)

- 1a. Grow *Sf9* cells in suspension culture and adapt to serum-free medium (see UNIT 16.10, Basic Protocol 1). Perform steps 2 through 4.
- 1b. Alternatively, incubate cells in a 27°C orbital shaker set at 150 rpm for 500-ml flasks or 100 rpm for 2.8-liter Fernbach flasks. Proceed to step 5.
2. Prepare spinner flasks to be used for scale-up of *Sf9* cells by attaching the appropriately sized two-port cap to one side arm and a plain cap to the other sidearm. Put a short (~6-in. or 15-cm) piece of tubing on the air-vent port and attach a filter to the end. Secure the tubing to the port and filter with cable ties using a tension tool.

Refer to Figure 16.11.1 for these procedures.

3. Put a longer piece of tubing (1 to 2 feet or 30 to 60 cm) on the air-supply port and attach a 0.2- μ m filter unit to the end. Secure with cable ties. Attach another piece of tubing to the other end of the filter with a cable tie. Cover the end of the tubing with aluminum foil.

The length of the silicone tubing attached to the air-supply port will depend on the distance between the sidearm and the air-supply pump in the incubator.

Several sizes of two-port cap assemblies are available for Bellco spinner flasks. The smallest size fits the 1-liter flask and the next larger size fits the 3-liter flask. The largest size fits the 6-, 8-, and 15-liter flasks.

4. Loosen the cap on the side arm opposite to the two-port assembly a quarter turn and autoclave flask 1 hr.

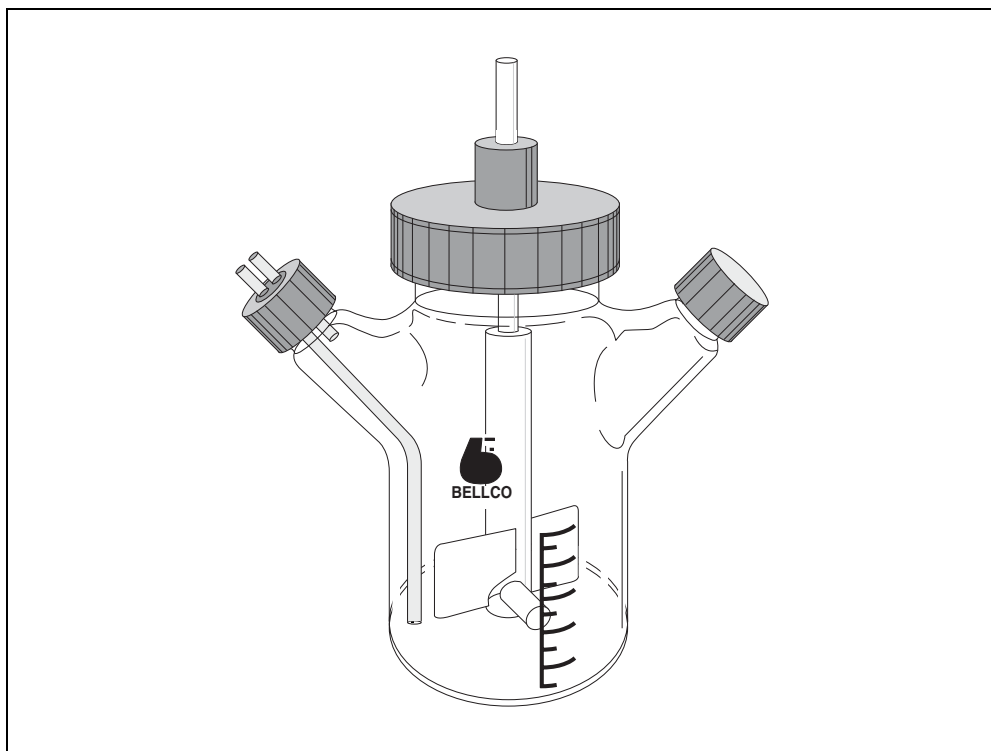


Figure 16.11.1 Bellco spinner flask with two-port cap assembly.

- 5a. *For spinner flasks:* Seed the autoclaved flask with *Sf9* cells adapted to serum-free medium (from step 1). Fill the flask to between half full and full (e.g., 1.5 to 3 liters in a 3-liter flask), adding enough serum-free medium to make a final cell density of $5\text{--}6 \times 10^5$ cells/ml.

If enough cells are available, flasks can be seeded at densities up to 1.5×10^6 cells/ml.

- 5b. *For shaker flasks:* Adapt cells to larger flasks for 1 to 2 days prior to infection. Once the culture is established, proceed to step 8.

A 2.8-liter shaker flask will hold 1 liter of culture.

6. Place the flask on a stir plate for spinner flasks in a 27°C incubator. Set the stir speed at 80 rpm. Remove the aluminum foil from the air-intake tube and attach the end of the tube to the air-supply pump. Turn on the pump and set to the lowest setting.

A flow rate of 500 to 700 ml/min works well with these flasks; the larger flasks require a higher flow rate than the smaller ones. It is recommended that the pump flow rate be measured empirically before hooking it up to the flask or that a flow meter be attached between the pump and the flask.

7. Grow the cells to a density of $\sim 1.5 \times 10^6$ cells/ml. Working in a laminar-flow hood, add virus at a multiplicity of infection (MOI) of 1 to 2 directly to the flask through the side arm.

8. Place the flask on the magnetic stir plate at 27°C and connect the air supply. For shaker cultures, place flask in the shaker with a loosened cap. Incubate the culture for the optimal amount of time as determined in Support Protocol 1.

9. Process the supernatant for secreted proteins or the cells for intracellular proteins.

There are many protein purification methods available and the choice of which ones to use depends on the nature of the recombinant protein expressed. For additional information see Chapter 10 or Coligan et al. (1997).

CO-INFECTION WITH MULTIPLE BACULOVIRUSES FOR PURIFICATION OF RECOMBINANT PROTEIN COMPLEXES

ALTERNATE
PROTOCOL 1

In this procedure, *Sf9* cells are co-infected with multiple baculoviruses that express subunits of multiprotein complexes. Complexes can then be purified using an affinity tag on a single subunit. This protocol is essentially identical to the one described above (see Basic Protocol 2), except that multiple viruses are used to infect each culture. The two parameters that need to be optimized are the ratio of viruses encoding different subunits of the complex and the timing of infection. Two examples of successful expression and purification of recombinant complexes are Phelan et al. (1999) and Zhang et al. (1999).

Additional Materials (see *Basic Protocol 2*)

Recombinant baculoviruses for complex subunits with one subunit carrying an affinity tag for purification
Shaker or spinner flask cultures (see *Basic Protocol 2*)

1. Grow *Sf9* cells in suspension culture in serum-free medium (*UNIT 16.10*). Prepare flasks for infection as described (see *Basic Protocol 2*, steps 2 to 6).

*Infect cells with multiple viruses at an $MOI \geq 2$ to 10 for each virus. *Sf9* cells will be multiply infected so that all subunits can be expressed in each cell. Ratios of various viruses can be adjusted to maximize stoichiometric complex formation. In general, more virus should be used for larger subunits and less for smaller subunits. For example, for a four-protein complex with two 170-kDa and two ~50-kDa proteins, the authors use a 3:1 ratio of virus (large:small protein). Epitope tagging the subunit (*UNIT 16.13*) expressed at the lowest levels should mainly result in purification of complex rather than free subunit. Some experimentation may be required to determine optimal ratios for complex production. This will require production of enough protein to purify at least a small amount of complex to assess subunit stoichiometry. If obtaining reproducible stoichiometries proves to be difficult, changing which subunit is epitope-tagged or adding a second type of epitope tag to a different subunit (and using a two-step affinity purification; *UNIT 10.15*) may help.*

Certain proteins do not co-express well; mixing extracts expressing the two proteins prior to purification sometimes yields successful complex formation (e.g., Phelan et al., 1999).

2. Incubate culture for an optimal amount of time (see *Support Protocol 1*).

In general, larger proteins tend to be more unstable and timing determined for single large proteins frequently work best for multiprotein complexes. The time of maximal expression is not always optimal for maximal intact protein—shorter infection times that may yield less protein tend to give more intact and active complex.

3. Harvest cells and process for intracellular proteins.

*To maintain complex integrity, it may be necessary to adjust lysis conditions or use subcellular fractionation (e.g., nuclear extracts for nuclear proteins, *UNIT 12.1*).*

PURIFICATION OF RECOMBINANT PROTEINS CONTAINING A POLYHISTIDINE (6×HIS) TAG

BASIC
PROTOCOL 3

The pAcHLT-A, -B, and -C transfer vectors (Pharming) and the pBlueBacHis A, B, and C vectors (Invitrogen) contain DNA that encodes an N-terminal tag of six histidine residues, followed by an extended multiple cloning site (MCS). The MCS is in a different reading frame in each of the vectors to simplify cloning. The expressed recombinant protein will be a 6×His fusion protein suitable for affinity purification on Ni-NTA agarose. Approximately 1 to 2 mg of 6×His recombinant fusion protein is routinely obtained per liter insect cell culture.

Protein
Expression

16.11.7

Materials

Insect cell lysis buffer (see recipe) containing 1× protease inhibitors (see recipe for 50×)

Ni-NTA agarose (Qiagen)

6×His wash buffer (see recipe)

6×His elution buffer (see recipe) containing imidazole as either a step or linear gradient

Beckman GPR centrifuge with GH-3.7 horizontal rotor (or equivalent), 4°C

Sorvall centrifuge and SS-34 rotor (or equivalent) or 0.2-μm filter

Suitable chromatography column

Additional reagents and equipment for preparing insect cell cultures and viral stocks (UNIT 16.10) and quantitation of protein by absorbance spectrometry (APPENDIX 3D)

Harvest and lyse cells

1. Infect the desired amount of cells (e.g., 5×10^6 cells per 100-mm plate) with the amplified recombinant virus at an MOI of 5 to 10. Cultivate 3 to 5 days and examine for typical signs of infection (UNIT 16.10).
2. Harvest the cells and supernatant from infected plate(s) by gently spritzing the cells with a sterile pipet to resuspend. Transfer to a centrifuge tube. Centrifuge 10 min at $1000 \times g$ (2000 rpm in a GH-3.7 rotor), 4°C.
- 3a. *For secreted proteins:* Transfer the supernatant to a new tube and proceed to purification (step 6).
- 3b. *For intracellular proteins:* Discard the supernatant, then rinse the cells by resuspending the cell pellet gently in PBS, repeating the centrifugation, and discarding the supernatant.
4. Lyse the cells by adding 1 ml insect cell lysis buffer containing 1× protease inhibitors for every 1×10^7 cells and incubating 45 min on ice.

Choice of lysis buffer may vary with cell type, but EDTA-containing lysis buffers must not be used with 6×His fusion proteins.
5. Clarify the lysate by centrifuging 30 min at $40,000 \times g$ (18,000 rpm in an SS-34 rotor), 4°C to pellet the cellular debris or filter the lysate through a 0.2-μm filter.
6. Gently resuspend the Ni-NTA agarose and pour into a suitable chromatography column. Allow the beads to settle and the column to drain, then wash twice with 3 to 5 column volumes of 6×His wash buffer to remove the ethanol preservative. Allow the column to drain.

One milliliter of Ni-NTA agarose will bind ~5 to 10 mg of 6×His fusion protein.
7. Apply clarified lysate to column. Adjust the column flow rate to a maximum of 5 column volumes per hour.

Keep the flowthrough fraction to run on SDS-PAGE (UNIT 10.2) in case the binding capacity of the Ni-NTA agarose was exceeded.
8. Wash the column with 10 bed volumes of 6×His wash buffer and allow the column to drain while periodically monitoring the A_{280} . Repeat washing (~4 times) until the A_{280} of the column effluent is <0.01 . Discard the washes.
9. Add 3 bed volumes of the 6×His elution buffer (including imidazole either as a step or a linear gradient) to the column. Adjust the column flow rate to a maximum of 1 ml/min per milliliter of resin beads. Allow the column to drain completely while collecting the eluted fractions.

The optimal amount of imidazole (0.1 M to 0.5 M) needed for elution will vary on the basis of the properties of the bound protein, and should be empirically determined by the researcher.

Pharmingen has also introduced the BioColors-His baculovirus transfer vector set, which allows the production of a fusion protein consisting of the desired gene product along with one of several derivatives of the *Aquorea victoria* green fluorescent protein (GFP; see UNIT 9.7C) and a histidine affinity tag. Under UV light, such a recombinant fusion protein can be visualized during the entire purification process, which facilitates the establishment of an optimized purification procedure.

PURIFICATION OF RECOMBINANT PROTEINS CONTAINING A GST TAG

ALTERNATE PROTOCOL 2

Several baculovirus vectors are commercially available that allow the expression of glutathione-S-transferase (GST)-tagged fusion proteins, which greatly facilitate the purification of recombinant proteins. Pharmingen's pAcGHLT-A-, -B, and -C transfer vectors, for example, encode N-terminal 6×His and GST tags followed by an extended MCS. The MCS is in a different reading frame in each of the vectors to simplify cloning. Because the GST vectors also contain a 6×His sequence, the expressed recombinant protein will be a 6×His-containing GST fusion protein. This feature allows affinity purification using either glutathione agarose beads or Ni-NTA agarose. A thrombin cleavage site following the affinity tags allows for the proteolytic cleavage of the GST-6×His fusion partner from the protein of interest.

The GST purification method is based on the remarkable selectivity and affinity of recombinant proteins equipped with a GST affinity tag toward glutathione immobilized on a resin. The expressed GST fusion proteins are authentically processed, and may be purified without the use of detergents under completely nondenaturing conditions. Purifications to >90% homogeneity are easily achieved in a single step by affinity chromatography using glutathione agarose beads. The affinity of GST for glutathione is so strong that it allows a highly efficient separation of GST fusion proteins from contaminating polypeptides even under nondenaturing conditions.

Although addition of a 6×His or GST tag to a recombinant protein may greatly simplify purification, these methods do not work for all proteins and should be tried on a small scale first.

Additional Materials (also see Basic Protocol 3)

- Glutathione agarose beads (Pharmingen or Sigma)
- PBS wash buffer (Pharmingen)
- GST elution buffer (see recipe)
- 50 mM Tris·Cl, pH 8.0 (APPENDIX 2)
- Thrombin, bovine (e.g., Sigma or Boehringer Mannheim)
- Additional reagents and equipment for dialysis (APPENDIX 3C)

Purify GST fusion protein

1. Prepare clarified solution of recombinant protein (see Basic Protocol 3, steps 1 to 5).
2. Gently resuspend the glutathione agarose beads and pour slurry into a suitable chromatography column. Allow the beads to settle and the column to drain. Wash twice with 3 to 5 bed volumes of PBS wash buffer to remove the ethanol preservative, then allow the column to drain.
3. Apply the clarified solution of recombinant proteins to the column. Adjust column flow rate to a maximum of 5 ml/min per milliliter of beads.

Keep the flowthrough fraction to run on SDS-PAGE (UNIT 10.2) in case the binding capacity of the glutathione beads was exceeded.

Protein Expression

16.11.9

4. Wash the column twice with 5 bed volumes PBS wash buffer. Allow the column to drain and discard the washes.
5. Add 3 bed volumes GST elution buffer to the column. Adjust the column flow rate to a maximum of 1 ml/min per milliliter beads. Allow the column to drain completely while collecting the eluted fraction.

The addition of 150 mM NaCl, 5 mM CaCl₂ (or for some proteins 5 mM MgCl₂), and 0.1% 2-mercaptoethanol is optional but may be required for the solubility of some proteins.

6. Remove the free glutathione by dialysis (*APPENDIX 3C*) for ≥4 hr against 100 vol of 50 mM Tris-Cl (pH 8.0), 4°C. Change dialysis buffer after 2 hr.

Perform thrombin cleavage

7. Add 200 μg (10 thrombin units) of bovine thrombin per milligram purified GST or 6×His fusion protein containing a thrombin cleavage site.
8. Mix, then incubate at room temperature up to 12 hr.

In many cases a 20- to 60-min incubation will be sufficient.

9. Add 2 vol of a 50% (v/v) slurry of glutathione agarose beads at the end of the cleavage reaction.

This directly removes GST and uncleaved GST fusion protein. Similarly, 6×His and uncleaved 6×His fusion protein may be removed by directly adding Ni-NTA agarose at the end of the cleavage reaction.

10. Incubate 30 min at 4°C, then centrifuge 10 min at 5000 × g (2250 rpm in an SS-34 rotor), 4°C. Retain the supernatant and freeze at –80°C.

The supernatant will contain the purified protein as well as thrombin and can be stored frozen at –80°C. Some proteins may require the addition of BSA or glycerol (50% final concentration) for stability.

Thrombin cleaves in 50 mM Tris-Cl buffer and does not require specific metal ions for its activity (Haun and Moss, 1993; Wu et al., 1992). However, Guan and Dixon (1991) have recommended using a buffer containing 50 mM Tris-Cl (pH 8.0), 150 mM NaCl, 2.5 mM CaCl₂, and 0.1% 2-mercaptoethanol for efficient cleavage.

*An efficient thrombin cleavage primarily depends on the sequence of the thrombin consensus site and the three-dimensional structure surrounding that site. The thrombin cleavage consensus site is *XXP(K/R)*BB*, where *X* stands for hydrophobic apolar amino acids, *P* stands for proline, (*K/R*) symbolizes that both lysine and arginine work in this position, and *B* stands for nonacidic amino acids. The (*) symbolizes the cleavage position, which is at the carboxy-terminal side of the arginine or the lysine residue. The thrombin site used in the pAcGHLT and pAcHLT vectors is LVPR*GS. A BamHI site codes for the amino acids GS within the cloning site and allows the insertion of the desired gene within the thrombin coding sequence. This cloning strategy allows release of nearly authentic proteins after thrombin cleavage.*

ALTERNATE PROTOCOL 3

PURIFICATION OF RECOMBINANT PROTEINS CONTAINING A FLAG TAG

Baculovirus vectors allowing the expression of FLAG-tagged fusion proteins, which generally retain the biochemical activities of untagged proteins, are not commercially available; however, the FLAG epitope can be conveniently added to either the N- or C-terminus of proteins of interest by PCR. FLAG-tagged proteins can be affinity purified using anti-FLAG M2 affinity beads (Sigma) and eluted with FLAG peptide (Sigma). The resulting proteins are often pure enough for many biochemical assays.

Additional Materials (see also *Basic Protocol 3*)

Anti-FLAG M2 affinity beads (Sigma)

BC buffers (*UNIT 16.13*)

FLAG peptide (Sigma)

Additional reagents and equipment for lysing cells (*UNIT 16.7*) and SDS-PAGE (*UNIT 10.2A*)

Purify FLAG-tagged protein

1. Prepare the clarified solution of recombinant protein as described (see *Basic Protocol 3*, steps 1 to 5). Alternatively, if detergent inhibits subsequent assays, prepare the cell lysate by repeated cycles of freezing in dry ice and thawing at 37°C, followed by sonication (*UNIT 16.7*).

Additionally, a subcellular fraction containing recombinant protein may be used as starting material (e.g., nuclear extract, S100 fraction, nuclear pellet prepared according to UNIT 12.1).

2. Equilibrate anti-FLAG M2 affinity beads with 10 bead volumes of the buffer used in step 1. Collect beads by centrifuging 5 min at 1000 × g, 4°C.
3. Mix recombinant protein-containing solution with beads and incubate for 3 hr at 4°C.
4. Centrifuge M2-beads 5 min at 1000 × g, 4°C. Save a portion of supernatant to assess binding and discard the remainder.
5. Resuspend M2-beads with BC300 and pour into a suitable chromatography column (e.g., Bio-Rad Econo-Column). Allow beads to settle and column to drain.
6. Sequentially wash column with 10 bead volumes buffer containing increasing salt concentrations up to 2 M KCl (e.g., BC600, BC1200, and BC2000).
7. Sequentially wash column with 10 bead volumes buffer containing decreasing salt until desired concentration is obtained (e.g., BC100).
8. Allow column to drain completely. Add 0.5 bead volumes of 0.5 to 2 mg/ml FLAG peptide in elution buffer (e.g., BC100). Collect flow-through as pre-elution.
9. Add 0.25 bead volumes elution buffer and let sit over column surface for 1 hr at 4°C.
10. Collect elution in fractions of 1/3 to 1/2 of the bead volume for a total of 3 to 5 bead volumes.

Beads can be incubated with peptide for a second hour and a second elution collected; this can be repeated a third time. FLAG peptide may be omitted for the final elution.

11. Analyze unbound and eluted fractions by SDS-PAGE (*UNIT 10.2A*)

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

GST elution buffer

Dissolve 42 mg reduced dry glutathione in 40 ml of 50 mM Tris·Cl, pH 8.0 (*APPENDIX 2*). Divide into small aliquots and store up to 6 months at –20°C.

6×His elution buffer

50 mM sodium phosphate, pH 8.0 (APPENDIX 2)
300 mM NaCl
10% (v/v) glycerol
Imidazole to make either a step or linear gradient for elution
Prepare using distilled water and filter sterilize
Store up to 6 months at 4°C.

A maximum concentration of 0.5 M imidazole is recommended for the final elution. The concentration of imidazole should be determined empirically by the researcher.

6×His wash buffer

50 mM sodium phosphate, pH 8.0 (APPENDIX 2)
300 mM NaCl
10% (v/v) glycerol
Prepare using distilled water and filter sterilize
Store up to 6 months at 4°C.

For a more stringent wash, the pH may be adjusted to 6.0 and 0.8 to 40 mM imidazole may be added. The concentration of imidazole should be determined empirically by the researcher.

Insect cell lysis buffer, 1×

10 mM Tris-Cl, pH 7.5 (APPENDIX 2)
130 mM NaCl
1% (v/v) Triton X-100
10 mM NaF
10 mM sodium phosphate, pH 7.5 (APPENDIX 2)
10 mM sodium pyrophosphate
Prepare using distilled water and filter sterilize
Store up to 6 months at 4°C.
Add 50× protease inhibitor cocktail (see recipe) to 1× before use

Protease inhibitor cocktail, 50×

Prepare in 100% ethanol
800 µg/ml benzamidine
500 µg/ml phenanthroline
500 µg/ml aprotinin
500 µg/ml leupeptin
500 µg/ml pepstatin A
50 mM PMSF
Vortex to resuspend insoluble constituents before use
Store up to 6 months at -20°C

COMMENTARY

Background Information

There are many protein purification methods available, and the choice of which to use depends on the nature of the recombinant protein expressed (for additional information see Chapter 10). Suspension (spinner or shaker) cultures are often more convenient to use for protein expression than monolayer cultures and densities of up to 10^7 cells/ml with >97% viability can be achieved. Suspension cultures are also preferred when growing large quantities of cells (e.g., to prepare large viral stocks or large

quantities of recombinant protein), because the entire process can be carried out in one flask. For small-scale procedures, by contrast, monolayer cultures are preferred because they seem less susceptible to contamination and can more easily be assessed to monitor the progress of viral infections. Monolayer cultures are absolutely required for transfections as well as for plaquing viral supernatants, and they are also preferred for preparing viral stocks or producing recombinant protein.

Most cell lines are readily adaptable to serum-free medium when in suspension. The use of serum-free medium has several advantages. First, cells can be grown to higher cell densities in serum-free medium. Second, serum-free medium is more consistent and less expensive than serum-supplemented medium. In addition, serum-free medium has a low protein content, which aids in the purification of secreted recombinant proteins. It should be determined empirically which cell line and medium formulation is most productive for the particular recombinant protein of interest. *Spodoptera frugiperda* (*Sf9*, *Sf21*) cells grown in serum-free medium are more sensitive to centrifugation, stick more tightly to plasticware, undergo a growth lag when seeded at too low a density, and cannot be passaged more than ~50 times.

In regard to purification, secreted recombinant proteins are much easier to purify than nonsecreted proteins, because the ratio between recombinant protein and host proteins in the medium is much higher than in lysates, especially when protein-free medium has been used. The general strategy for purifying protein from the medium depends on the nature of the recombinant protein. If an antibody against the desired protein is available in large quantities, it can be used for affinity purification. Otherwise, conventional ion-exchange chromatography may perform equally well.

Critical Parameters

The efficiency of heterologous gene expression in the baculovirus system can differ 1000-fold as a result of the intrinsic nature of the gene and the encoded protein. Modifying the heterologous gene will generally influence gene expression by only 2- to 5-fold. Researchers should not feel compelled to modify their gene excessively. For some general rules regarding the improvement of gene expression, see O'Reilly et al. (1992). When expressing protein, a time-course experiment should be conducted to determine peak protein production.

To optimize for protein production, it is recommended that log-phase *Sf9* cells be infected at high multiplicity of infections (MOIs)—e.g., 5 to 10 for monolayer culture and 1 to 2 for suspension culture. Cells to be infected should be >97% viable. Use of high-quality media and FBS also seems to contribute to high protein yields.

The most common method for radiolabeling protein involves incorporating [³⁵S]methionine or [³⁵S]cysteine. To improve results, the intra-

cellular pool of these essential amino acids should be depleted prior to radiolabeling by incubating the cells in methionine/cysteine-free medium. If an antibody is available, the labeled lysate should be immunoprecipitated to help detect the recombinant protein if it comigrates with a labeled host protein.

To maximize large-scale production of recombinant proteins, *Sf9* cells should be grown in suspension in serum-free medium. Flasks should remain <50% full or be vented to ensure proper oxygenation. Cells should be grown to a density of 1.5×10^6 cells/ml before infecting at an MOI of 1 to 2.

Purification of recombinant proteins can be simplified by expressing them with either a GST or 6×His tag. When purifying these proteins via affinity chromatography, the flow-through fractions should be retained for analysis by SDS-PAGE. Analysis of these fractions may help in troubleshooting poor affinity binding. Other affinity tags such as FLAG, myc, or HA—for which high-quality affinity resins are commercially available, and for which gentle elutions with peptide can be used—work well for insect cell expression. For purification of complexes, only a single subunit should be tagged. In general, epitope tagging the protein expressed at the lowest levels results in optimal stoichiometries of purified complexes.

The addition of a 6×His or GST tag to a recombinant protein may greatly simplify purification; however, these methods do not work for all recombinant proteins and should be tried on a small scale first.

Anticipated Results

When expressing proteins in the baculovirus system, one can typically expect to produce up to 1 to 5 mg protein per liter insect-cell culture. However, it is extremely important to note that gene expression can vary 1000-fold as a result of the intrinsic nature of the protein. Recombinant protein levels can vary from 1% to 50% of the total insect-cell protein. When using an MOI >1, it is expected that most cells in the culture will become infected at the same time and thus produce protein within the same time frame. A time course can be run to test for peak protein production. Cells infected during log-phase growth with a high surface-to-volume ratio to allow for proper gas exchange will produce protein optimally.

Approximately 1 to 2 mg of 6×His and GST fusion protein are routinely obtained per liter of insect-cell culture. Affinity purifications to

>90% homogeneity are easily achieved using single-step affinity purification with Ni-NTA, glutathione agarose, or other epitope tags.

Time Considerations

It takes 1.5 to 2 weeks to obtain a spinner culture from a frozen vial of *Sf9* cells. A viral stock can be obtained in ~4 days and plaqued in 1 day. It takes ~1 week before the plaques are formed and ready to be counted.

It takes an additional 2 days to infect *Sf9* cells to obtain cells or medium for protein analysis. Analyzing for the recombinant protein after harvesting should take 1 to 3 days. Finally, it will take ~4 days to complete the time course and another 2 to 3 days to harvest and analyze the infected cell pellets or culture supernatants.

Literature Cited

- Coligan, J.E., Dunn, B.M., Ploegh, H.L., Speicher, D.W., and Wingfield, P.T. (eds.) 1997. *Current Protocols in Protein Science*. John Wiley & Sons, New York.
- Guan, K. and Dixon, J.E. 1991. Eukaryotic proteins expressed in *Escherichia coli*: An improved thrombin cleavage and purification procedure of fusion proteins with glutathione S-transferase. *Anal. Biochem.* 192:262-267.
- Haun, R.S. and Moss, J. 1993. Ligation-independent cloning of glutathione S-transferase fusion genes for expression in *Escherichia coli*. *Gene* 112:37-43.
- O'Reilly, D.R., Miller, L.K., and Luckow, V.A. 1992. *Baculovirus Expression Vectors*. W.H. Freeman, New York.
- Phelan, M.L., Sif, S., Narlikar, G.J., and Kingston, R.E. 1999. Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol. Cell* 3:247-53.

Summers, M.D. and Smith, G.E. 1987. *A manual of methods for baculovirus vectors and insect cell culture procedures*. Texas Agricultural Experiment Station Bulletin No. 1555. College Station, Texas.

Wu, H., White, G.C., Workman, E.F., Jenzano, J.W., and Lundblad, R.L. 1992. Affinity chromatography of platelets on immobilized thrombin: retention of catalytic activity by platelet-bound thrombin. *Thrombosis Res.* 67:419-427.

Zhang, Y., Ng, H.-H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. 1999. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* 13:1924-1935.

Key Reference

O'Reilly et al., 1992. See above.

A laboratory manual that will aid researchers in the expression and purification of recombinant proteins using the baculovirus system.

Contributed by Cheryl Isaac Murphy
Aquila Biopharmaceuticals
Worcester, Massachusetts

Helen Piwnica-Worms
Washington University School of Medicine
St. Louis, Missouri

Stefan Grünwald
Pharmingen
San Diego, California

William G. Romanow
Corning Costar
Portsmouth, New Hampshire

Nicole Francis and Hua-Ying Fan
Massachusetts General Hospital
Boston, Massachusetts

EXPRESSION OF PROTEINS IN MAMMALIAN CELLS

SECTION III

Transient Expression of Proteins Using COS Cells

UNIT 16.12

This unit describes the use of COS cells to efficiently produce a desired protein in a short period of time. These cells express high levels of the SV40 large tumor (T) antigen, which is necessary to initiate viral DNA replication at the SV40 origin. Three factors contribute to make COS cell expression systems appropriate for the high-level, short-term expression of proteins: (1) the high copy number achieved by SV40 origin-containing plasmids in COS cells 48 hr post-transfection, (2) the availability of good COS cell expression/shuttle vectors, and (3) the availability of simple methods for the efficient transfection of COS cells. Each COS cell transfected with DNA encoding a cell-surface antigen (in the appropriate vector) or cytoplasmic protein will express several thousand to several hundred thousand copies of the protein 72 hr post-transfection. If the transfected DNA encodes a secreted protein, up to 10 μg of protein can be recovered from the supernatant of the transfected COS cells 1 week post-transfection. COS cell transient expression systems have also been used to screen cDNA libraries, to isolate cDNAs encoding cell-surface proteins, secreted proteins, and DNA binding proteins, and to test protein expression vectors rapidly prior to the preparation of stable cell lines (UNIT 9.5).

BASIC
PROTOCOL

This transfection protocol is a modification of that presented in UNIT 9.2 and gives conditions for optimal transfection of COS cells. The main difference between this procedure and that in UNIT 9.2 is the composition of the DEAE-dextran/chloroquine solution, which is prepared here in PBS, not TBS, and contains chloroquine to prevent the acidification of endosomes presumed to carry the DEAE-dextran/DNA into the cell. (This acidification results in acid hydrolysis of the DNA, giving rise to mutations and destruction of the DNA.) With this protocol, 40% to 70% of the cells can be routinely transfected.

Materials

- Appropriate vector (see Background Information)
- COS-7 cells to be transfected (see Critical Parameters), adapted to growth in 2% serum (see annotation to step 2)
- Dulbecco's minimum essential medium (DMEM) with 2% calf serum (DMEM-2 CS)
- DMEM with 2% NuSerum (Collaborative Research) (DMEM-2 NS), 37°C
- Phosphate-buffered saline (PBS; APPENDIX 2)
- DEAE-dextran/chloroquine solution: PBS containing 10 mg/ml DEAE-dextran (Sigma) + 2.5 mM chloroquine (Sigma)
- 10% dimethyl sulfoxide (DMSO; Sigma) in PBS (see APPENDIX 2 for PBS)
- 0.5 mM EDTA in PBS
- 100-mm tissue culture dishes
- Humidified 37°C, 6% CO₂ incubator
- Phase-contrast microscope
- Sorvall RT-6000B rotor (or equivalent)

Protein
Expression

Contributed by Alejandro Aruffo

Current Protocols in Molecular Biology (1998) 16.12.1-16.12.7

Copyright © 2002 by John Wiley & Sons, Inc.

16.12.1

Supplement 60

Additional reagents and equipment for subcloning of DNA (UNIT 3.16), preparation of miniprep DNA (UNIT 1.6), purification of DNA by CsCl/ethidium bromide equilibrium centrifugation (UNIT 1.7), and flow cytometric analysis (Otten et al., 1995)

NOTE: All cell culture incubations should be carried out in a humidified 37°C, 6% CO₂ incubator unless otherwise stated.

1. Subclone the gene of interest into the appropriate vector to obtain the desired recombinant DNA (UNIT 3.16). Purify the recombinant DNA by a miniprep procedure (5-ml culture UNIT 1.6), or by CsCl/ethidium bromide centrifugation (UNIT 1.7).
2. Seed COS-7 cells in DMEM-2 CS at ~20% confluence per 100-mm dish the day prior to transfection (so they will be ~50% confluent the next day). Grow cells overnight to ~50% confluence.

The COS-7 cells should be adapted to 2% serum (from the usual 10% serum) by passaging them through decreasing concentrations of serum, in 2% increments, over ~1 month. The procedure can be carried out in 10% serum, but at much greater expense.

A confluent dish of COS-7 cells (~10⁶ cells) is usually split 1:5 on the day prior to transfection to give 2 × 10⁵ cells/100-mm dish in 10 ml of DMEM-2 CS. In some cases cells may benefit from an extra day of growth prior to transfection (see step 6 annotation).

3. Just before use (for each 100-mm dish of COS cells to be transfected), thoroughly mix 5 ml of 37°C DMEM-2 NS with 0.2 ml of DEAE-dextran/chloroquine solution. Add 5 to 10 µg recombinant DNA and mix.

It is important that the DEAE-dextran be well mixed with the medium before the DNA is added—otherwise, the DNA, a negatively charged polymer, will form large precipitates with DEAE-dextran, a positively charged polymer. These large precipitates cannot be taken up by the cell, resulting in a reduced transfection efficiency. When larger dishes are used, the amount of medium/DEAE/DNA should be sufficient to easily cover the cells and should include 400 µg/ml DEAE-dextran, 100 µM chloroquine, and 1 to 2 µg/ml DNA in DMEM-2 NS.

Either CsCl-purified (UNIT 1.7) or miniprep (UNIT 1.6) plasmid DNA can be used for the transfections. If miniprep DNA is used, use one-fifth of the miniprep per transfection.

4. Aspirate medium from COS cells and add the DMEM-2 NS/DEAE-dextran/DNA prepared in step 3. Incubate cells 3 to 4 hr. Observe the cells using a phase-contrast microscope.

The DEAE-dextran will cause cells to retract and become vacuolated. Efficiency of transfection increases with longer incubation periods; on the other hand, so does cell death. The 3- to 4-hr incubation suggested here is a good starting point. However, several time points should be tried to optimize transfection of the particular population of cells used.

5. Aspirate DMEM/DEAE-dextran/DNA and add 5 ml of 10% DMSO (prepared in PBS). Incubate cells 2 min at room temperature.

The DMSO shock results in increased transfection efficiencies. Without this step, transfection efficiencies might be lower by a factor of two or more.

6. Aspirate DMSO and add 10 ml DMEM-2 CS. Grow cells overnight (12 to 20 hr).

If a significant number of the cells have lifted from the plate after overnight growth, the procedure will need to be repeated, but cells should be allowed to grow an extra day at step 2.

7. Passage (split and replat) each 100-mm dish of transfected COS cells into two new 100-mm dishes.

After transfection, the COS cells will look unhealthy. Passaging them the day after transfection facilitates recovery, resulting in better levels of protein expression. In addition,

DEAE-dextran treatment makes the cells sticky, and passaging the cells the morning after transfection restores their adhesion characteristics so that they may be once again lifted by a gentle treatment with PBS and EDTA (see step 8b).

- 8a. *When expressing secreted proteins:* 96 hr (4 days) after completing step 7, add 5 ml DMEM-2 CS and incubate 4 days. Harvest the medium, remove dead cells and debris by centrifuging 10 min in a Sorvall RT-6000B rotor at ~2000 rpm (~1000 × g), room temperature, and save the supernatant (see Anticipated Results). Detect secreted proteins by bioassay (UNIT 9.5) or by metabolic labeling and immunoprecipitation (UNITS 10.18 & 10.16), immunoaffinity chromatography (UNIT 10.11), radioimmunoassay (UNIT 11.17), or western blotting (UNIT 10.8).

Do not aspirate the old medium prior to addition of 5 ml DMEM-2 CS because this medium contains the secreted protein. Addition of extra medium 96 hr post-transfection results in better yield of expressed protein; however, it also increases the level of total protein (since the medium contains 2% serum), which could complicate protein purification. To eliminate this problem, COS cells can be placed in serum-free medium 10 to 12 hr after they have been replated, although (in our hands) this results in a 10-fold lower yield of expressed protein than in the presence of serum. Thus, unless it is absolutely necessary to remove additional contaminating protein, serum should be present in the medium even at reduced levels (1%).

- 8b. *When expressing cell-surface or intracellular proteins:* Aspirate medium from cells 72 hr (3 days) after transfection in step 6. Add 5 ml PBS, swirl, and aspirate PBS. Add 5 ml of 0.5 mM EDTA in PBS and incubate 15 min. Lift cells from the dish by gently dislodging them with a Pasteur pipet. Stain cell-surface proteins with the appropriate fluorescent antibody and detect by microscopy or flow cytometry (Otten et al., 1995).

Transfected COS cells will tend to clump when lifted from the dish. Pipetting the cells up and down will tend to disrupt these clumps. More effective dispersion of the clumps can be obtained by forcing the cells through a 100- μ M nylon mesh.

COMMENTARY

Background Information

COS cells

COS cells are African green monkey kidney cells (CV-1) that have been transformed with an origin-defective SV40 virus, which has integrated into COS cell chromosomal DNA. Therefore, COS cells produce wild-type SV40 large T antigen but no viral particles. Since SV40 large T antigen is the only viral protein required *in trans* for viral replication (i.e., its coding sequence need not be located on the DNA molecule on which it acts), SV40 origin-containing plasmids replicate in these cells to a high copy number (10,000 to 100,000 copies/cell) 48 hr post-transfection. If the plasmid carries a cDNA or genomic insert encoding a desired protein (under the control of the appropriate promoter), COS cells will express the protein at relatively high levels over a short period of time. Transfected COS cells produce protein in a burst that starts ~24 hr post-trans-

fection and can last for up to a week. However, due to the excessive burden placed on the transfected cell by the replicating plasmid and the high levels of protein production, the transfected cells typically either die or lose the plasmid a week after transfection.

COS cells were developed by Yakov Gluzman (1981) as a host for the propagation of SV40 virus early-region mutants. The first SV40 origin-containing plasmids to be used in conjunction with COS cells were made by Lusky and Botchan (1981). Short-term expression systems using both COS cells and SV40 origin-containing plasmids were initially used to identify DNA sequences required for transcription of the human α 1-globin gene (Mellon et al., 1981).

COS cells were first used to produce cell-surface and secreted proteins by Rose and Bergmann (1982), who looked at the expression of wild-type and mutant vesicular stomatitis virus glycoprotein in transfected cells. This

technique was subsequently used to study the expression of insulin (Laub and Rutter, 1983), somatostatin (Warren and Shields, 1984), and acetylcholine receptors (Mishina et al., 1984). These experiments demonstrated that COS cells could be used to express biologically active cell-surface and secreted proteins. Furthermore, these proteins were correctly processed although they are normally not produced by COS cells.

COS cell expression was initially used to screen a cDNA library to isolate a cDNA encoding human granulocyte/macrophage colony-stimulating factor (Wong et al., 1985). This was subsequently extended to isolation of cDNAs encoding cell-surface proteins (Aruffo and Seed, 1987a; Seed and Aruffo, 1987) and DNA-binding proteins (Tsai et al., 1989).

The expressed protein produced in COS cells is in most cases biologically active. However, although COS cells are able to carry out some post-translational modifications, they may not modify the expressed protein in exactly the same way as the cell that would normally produce it. For example, COS cells do not express the α -(1,3)fucosyltransferase, which is capable of transferring fucose to either sialyl or asialyl precursors (Goelz et al., 1990). In addition, insufficient post-translational modification occurs in the case of lymphocyte cell-surface proteins, which tend to be underglycosylated in COS cells (Aruffo and Seed, 1987b). This might be due to an overburdening of the COS cell glycosylation machinery by the high levels of protein expression and/or by the lack of enzymes required to carry out the full post-translational modifications.

An alternative to COS cells is provided by WOP cells (Dailey and Basilico, 1985), which are mouse 3T3 cells transformed by an origin-defective polyoma virus. Like COS cells, they produce no viral particles. However, they produce polyoma large T antigen and are therefore capable of replicating a plasmid containing a polyoma origin of replication to a copy number that is typically ten times lower than that obtained in COS cells. WOP cells are also more delicate than COS cells, making them harder to transfect. For these two reasons, COS cells should be used whenever possible. However, WOP cells should be used in those cases where a monoclonal antibody that is used to identify and/or purify the protein being expressed cross-reacts with COS cell proteins. This situation has arisen when transient expression in COS cells was carried out in order to clone a human

protein with a mouse anti-human monoclonal antibody that also recognizes the equivalent monkey protein (Seed and Aruffo, 1987). In this situation, the mouse cell line presented a useful alternative to COS cells by avoiding monoclonal antibody cross-reactivity.

Vectors

The main requirements of any COS cell expression/shuttle vector are: (1) an SV40-derived origin of replication, (2) appropriate eukaryotic transcription regulatory elements (i.e., enhancer, promoter, and polyadenylation signal sequences), (3) a prokaryotic origin of replication, and (4) a prokaryotic genetic marker for selection in *Escherichia coli*. A particularly useful example of such a vector is CDM8 (Fig. 16.12.1; Seed, 1987). The eukaryotic transcription element of CDM8 is composed of the cytomegalovirus (CMV) enhancer-promoter, with an SV40 virus-derived intron and polyadenylation signal; the CMV promoter is a *cis* element (i.e., one that must be located adjacent to the DNA it acts on) that directs transcription of the DNA subcloned downstream from it. The prokaryotic genetic marker in CDM8 is provided by the *supF* (amber suppressor) gene. CDM8 is propagated in host bacterial cells containing helper plasmid P3, which contains amber mutations in the genes responsible for tetracycline and ampicillin resistance. (P3 has been introduced into many *E. coli* strains.) When CDM8 is transformed into an *E. coli* strain containing P3, the amber mutations are suppressed, rendering the host resistant to tetracycline and ampicillin. In addition to these elements, the CDM8 expression/shuttle vector contains an M13 origin of replication so that it can be used for the production of single-stranded DNA, a T7 RNA polymerase promoter for preparation of mRNA *in vitro*, and a polyoma virus-derived origin of replication which permits plasmid replication in WOP cells.

Other vectors commonly used for COS cell transient expression include pXM (Yang et al., 1986) and pDC201 (Sims et al., 1988). These two plasmids contain the adenovirus-2 major late promoter and tripartite mRNA leader. This element acts in conjunction with the adenovirus VA RNA (also produced by the vectors) to increase the translatability of the mRNA encoding the desired protein. It is thought that adenovirus VA RNAs increase translation efficiency of mRNAs containing the major late promoter tripartite leader by facilitating the

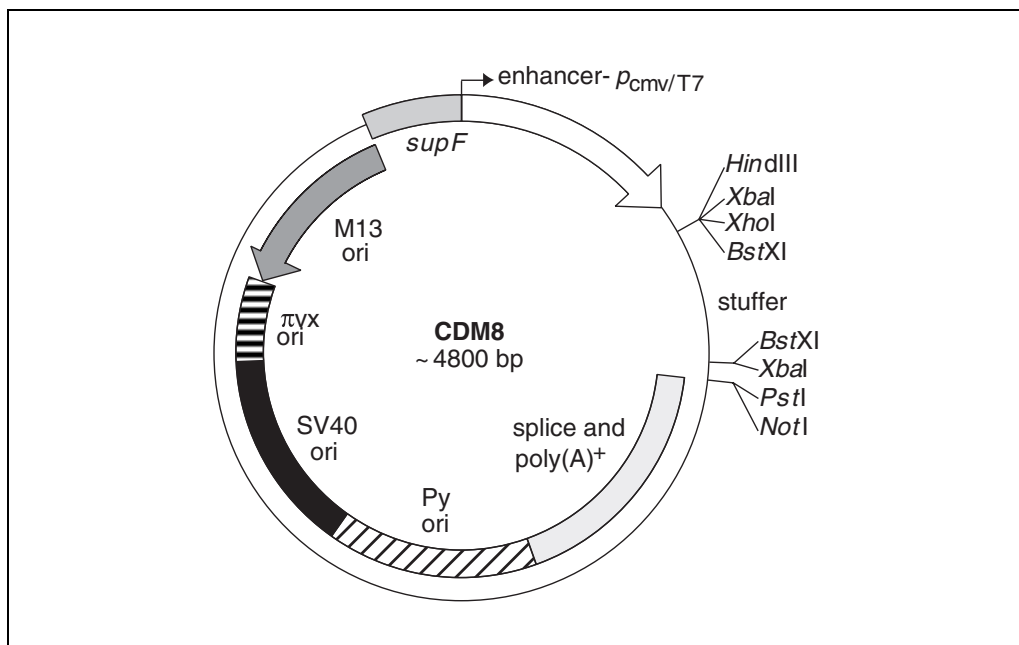


Figure 16.12.1 CDM8. CDM8 (Seed, 1987) is a COS cell expression/shuttle vector that contains an SV40-derived origin of replication (SV40 ori), eukaryotic transcription regulatory elements [splice and poly(A)⁺], a prokaryotic origin of replication (π vX; derived from the pBR322 ori), and a prokaryotic genetic marker (*supF*). In addition, CDM8 contains an M13 origin of replication (M13 ori), a T7 RNA polymerase promoter ($p_{CMV/T7}$), and a polyoma-derived origin of replication (Py ori). Any restriction endonuclease sites shown can be used for cloning, but the inserted fragment must have its 5' end nearest the enhancer $p_{CMV/T7}$. The stuffer sequence is used to detect a size difference after restriction digestion.

interaction between mRNA and a 43S ribosomal protein translation preinitiation complex (Kaufman, 1985).

Critical Parameters

Efficiency of transfection depends critically on the length of time that the cells (COS or other cells) are incubated in the presence of DEAE-dextran/DNA. Longer periods of time result in higher transfection efficiencies. However, the DEAE-dextran/chloroquine solution is quite toxic to cells and in general, cells should not be in its presence for >4 hr. In the past, DEAE-dextran transfections were carried out in the absence of serum, because a precipitate of unknown composition that seemed to be very toxic formed in DEAE-dextran/calf serum mixtures. Medium containing NuSerum, on the other hand, does not form this precipitate and tends to enhance the ability of the cells to tolerate DEAE-dextran; thus NuSerum should always be included in transfection medium.

Efficiency of transfection can also be affected by the quality of the DNA and the age of the DEAE-dextran/chloroquine solution. It is preferable to use CsCl-purified or other highly purified DNA whenever possible (*UNIT*

1.7). However, miniprep DNA (*UNIT* 1.6) or DNA purified using a pZ523 column (5'→3') or by other methods can also be used. The DEAE-dextran chloroquine solution can be kept at 4°C for several months but it is wise to prepare it fresh about every 3 months.

COS cells can be obtained from the American Type Culture Collection; several sublines exist, including COS-1 and COS-7. The COS-7 subline is recommended because it produces a higher plasmid copy number. These cells grow as a monolayer in DMEM-2 CS in a humidified 37°C, 5% CO₂ incubator; however, the ATCC grows its COS cells in DMEM-10 fetal bovine serum (FBS). Since serum is expensive, changing the growth medium is worthwhile. The change from FBS to CS should be done slowly over 1 to 2 weeks; the concentration of CS should then be reduced to 2% in 2% increments over ~1 month.

Because the growth characteristics, transfectability, and protein expression properties of COS cells change with time and with repeated subculturing, and because these changes tend not to favor the production of high levels of proteins, it is prudent to freeze aliquots of the original COS cell stock in DMEM-10 CS/10%

DMSO for later use in a -70°C freezer for 24 hr and then transfer them to a -150°C (liquid nitrogen) freezer. COS cells grow rapidly requiring passage every 4 to 5 days; typically, a confluent plate of cells is split 1 to 10.

To obtain good levels of transient protein production from transfected COS cells, it is very important to replat the transfected cells onto new dishes with fresh medium the morning after transfection. In addition to enhancing protein production, replating the transfected cells allows lifting of the cells from the dish using only PBS/EDTA (without trypsin). This is very important when transient expression by COS cells is used to produce cell-surface proteins.

Anticipated Results

The Basic Protocol should yield transfection efficiencies of 40% to 70%. When COS cells are being used to produce cell-surface or intracellular proteins, it can be expected that each transfected cell will express several thousand copies of this protein (10,000 to 100,000 copies/cell) 72 hr post-transfection. If COS cells are used to produce secreted proteins, up to 1 $\mu\text{g}/\text{ml}$ of protein can be recovered from the supernatant of a 100-mm dish of transfected cells 1 week post-transfection. However, the amount of protein produced by COS cells can vary dramatically depending on the protein being produced. This was the case when COS cells were used to produce soluble immunoglobulin fusion forms of cell-surface proteins (Aruffo et al., 1990). In this case, one of the fusion proteins, CD8 immunoglobulin, was secreted from COS cells at high levels (1 $\mu\text{g}/\text{ml}$) while the other, CD44 immunoglobulin, was secreted very poorly if at all. It was found that the CD44 fusion protein was sequestered inside the cell. To obtain efficient secretion of the CD44 fusion protein, it was necessary to change the amino-terminal signal sequence of the CD44 fusion protein. Interestingly, the native cell-surface forms of both CD8 and CD44 are expressed equally efficiently on the surface of transfected COS cells. For some of these immunoglobulin fusion proteins, 0.5 ml of medium contained plenty of protein (~500 ng) after concentration using a protein A affinity matrix. In some cases, it is possible to use such COS cell supernatants directly without further purification (Aruffo et al., 1990).

Time Considerations

It is important not to transfect the cells too soon after replating; >8 to 12 hr should pass

between the time the cells are seeded on the plate and the time of transfection. Once the transfection has started, the DEAE-dextran/DNA mixture should be left on the cells for a minimum of 2 hr and a maximum of 4 hr; because the mixture increases transfection efficiency it should remain in contact with the cells as long as they appear viable. After transfection, the cells look quite unhealthy, and 12 to 24 hr post-transfection they should be replated in new dishes with fresh medium.

The peak of plasmid replication in transfected COS cells occurs 48 to 72 hr post-transfection. Protein production starts 24 hr post-transfection but peaks 72 to 96 hr post-transfection. Thus, when expressing cell-surface or cytoplasmic proteins, the cells should be harvested 72 to 96 hr post-transfection. However, transfected cells continue to produce protein for up to a week post-transfection and when expressing secreted proteins, the supernatants should be harvested a week post-transfection.

Literature Cited

- Aruffo, A. and Seed, B. 1987a. Molecular cloning of a CD28 cDNA by a high-efficiency COS cell expression system. *Proc. Natl. Acad. Sci. U.S.A.* 84:8573-8577.
- Aruffo, A. and Seed, B. 1987b. Molecular cloning of two CD7 (T cell leukemia antigen) cDNAs by a COS cell expression system. *EMBO J.* 6:3313-3316.
- Aruffo, A., Stamenkovic, I., Melnick, M., Underhill, C.B., and Seed, B. 1990. CD44 is the principal cell-surface receptor for hyaluronate. *Cell* 61:1303-1313.
- Dailey, L. and Basilico, C.J. 1985. Sequences in the polyoma virus DNA regulatory region involved in viral DNA replication and early gene expression. *J. Virol.* 54:739-749.
- Gluzman, Y. 1981. SV-40 transformed simian cells support the replication of early SV40 mutants. *Cell* 23:175-182.
- Goelz, S.E., Hession, C., Goff, D., Griffiths, B., Tizard, R., Newman, B., Chi-Rosso, G., and Lobb, R. 1990. ELFT: A gene that directs the expression of an ELAM-1 ligand. *Cell* 63:1349-1356.
- Kaufman, R.J. 1985. Identification of the components necessary for adenovirus translational control and their utilization in cDNA expression vectors. *Proc. Natl. Acad. Sci. U.S.A.* 82:689-693.
- Laub, O. and Rutter, W.J. 1983. Expression of the human insulin gene and cDNA in a heterologous mammalian system. *J. Biol. Chem.* 258:6043-6050.
- Lusky, M. and Botchan, M. 1981. Inhibition of SV40 replication in simian cells by specific pBR322 DNA sequences. *Nature* 293:79-81.

- Mellon, P., Parker, V., Gluzman, Y., and Maniatis, T. 1981. Identification of DNA sequences required for transcription of the human $\alpha 1$ -globin gene in a new SV40 host-vector system. *Cell* 27:279-288.
- Mishina, M., Kurosaki, T., Tobimatsu, T., Morimoto, Y., Noda, M., Yamamoto, T., Terao, M., Lindstrom, J., Takahashi, T., Kuno, M., and Numa, S. 1984. Expression of functional acetylcholine receptor from cloned cDNAs. *Nature* 307:604-608.
- Otten G., Yokoyama, W.M. and Holmes, K.H. 1995. Flow cytometry analysis using the Becton Dickinson FACScan. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. 5.4.1-5.4.19. John Wiley & Sons, New York.
- Rose, J.K. and Bergmann, J.E. 1982. Expression from cloned cDNA of cell-surface, secreted forms of the glycoprotein of vesicular stomatitis virus in eucaryotic cells. *Cell* 30:753-762.
- Seed, B. 1987. An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. *Nature* 329:840-842.
- Seed, B. and Aruffo, A. 1987. Molecular cloning of the CD2 antigen, the T cell erythrocyte receptor, by a rapid immunoselection procedure. *Proc. Natl. Acad. Sci. U.S.A.* 84:3365-3369.
- Sims, J.E., March, C.J., Cosman, D., Widmer, M.B., MacDonald, H.R., McMahan, C.J., Grubin, C.E., Wignall, J.M., Jackson, J.L., Call, S.M., Friend, D., Alpert, A.R., Gillis, S., Urdal, D.L., and Dower, S.K. 1988. cDNA expression cloning of the IL-1 receptor, a member of the immunoglobulin superfamily. *Science* 241:585-589.
- Tsai, F., Martin, D.I.K., Zon, L.I., D'Andrea, D., Wong, G.G., and Orkin, S.H. 1989. Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature* 339:446-451.
- Warren, T.G. and Shields, D. 1984. Expression of preprosomatostatin in heterologous cells: Biosynthesis, posttranslational processing, and secretion of mature somatostatin. *Cell* 39:547-555.
- Wong, G.G., Witek, J.S., Temple, P.A., Wilkens, K.M., Leary, A.C., Luxenber, D.P., Jones, S.S., Brown, E.L., Kay, R.M., Orr, E.C., Shoemaker, C., Golde, D.W., Kaufman, R.J., Hewick, R.M., Wang, E.A., and Clark, S.C. 1985. Human GM-CSF: Molecular cloning of the complementary DNA and purification of the natural and recombinant proteins. *Science* 228:810-815.
- Yang, Y., Ciarletta, A.B., Temple, P.A., Chung, M.P., Kovacic, S., Witek-Giannotti, J.S., Leary, A.C., Kriz, R., Donahue, R.E., Wong, G.G., and Clark, S.C. 1986. Human IL-3 (multi-CSF): Identification by expression cloning of a novel hematopoietic growth factor related to murine IL-3. *Cell* 47:3-10.

Key Reference

Warren and Shields, 1984. See above.

This article shows that COS cells can be used as an efficient, short-term, mammalian expression system for the production of proteins.

Contributed by Alejandro Aruffo
Bristol-Myers Squibb
Princeton, New Jersey

Expression and Purification of Epitope-Tagged Multisubunit Protein Complexes from Mammalian Cells

Biochemical characterization and functional studies of mammalian proteins are often hampered by the availability of the purified protein, in particular, when the functional entity is present as a multisubunit protein complex in the cell. To overcome the difficulties in the purification of multisubunit protein complexes from mammalian cells, one may create stable cell lines containing epitope-tagged protein. In general, a stable cell line that expresses an epitope-tagged subunit of a protein complex is first established either by retrovirus-mediated gene transfer (for constitutive expression) or by a tetracycline-regulated system (for inducible expression). Immunoaffinity purification using epitope-specific monoclonal antibody-conjugated beads is then performed to pull down the epitope-tagged multisubunit protein complex, which is finally recovered by peptide elution under neutral pH or physiological conditions and is readily available for functional assays.

Several commonly used epitopes are listed in Table 10.15.1 (UNIT 10.15). For simplicity, only protocols using the FLAG epitope (Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys, or DYKDDDDK) are described here. The authors normally use human HeLa cells, which can grow in monolayer as well as in suspension, as parental cells for establishing stable cell lines, although other cells, such as human 293, can also be used for the experiments. By and large, HeLa-derived cell lines expressing an epitope-tagged protein are initially isolated in monolayer culture and then adapted to suspension culture for the purpose of collecting large quantities of cells for protein purification. The first protocol in this unit (see Basic Protocol 1) describes the procedures involved in the establishment of a stable cell line constitutively expressing the FLAG-tagged protein by retrovirus-mediated gene transfer and immunoaffinity purification of the epitope-tagged multisubunit protein complex. The next protocol (see Basic Protocol 2) outlines the steps involved in the establishment of an inducible cell line conditionally expressing the FLAG-tagged protein by a tetracycline-regulated system, and the one-step immunoaffinity purification of the multisubunit protein complex.

Since an epitope-tagged protein may exist as different complexes in the cell, it is also important to isolate various forms of multisubunit protein complexes. This can sometimes be achieved simply by careful selection of the starting material (e.g., using nuclear extract, cytoplasmic S100 fraction, or the chromosomal fraction) for immunoaffinity purification or by manipulating the wash conditions with different concentrations of salt, detergent, and chaotropic agents during immunoaffinity purification. The Alternate Protocol provides an excellent example for the purification of different forms of human RNA polymerase II complexes, achieved simply by choosing the appropriate starting material and by varying wash conditions. Frequently, multiple forms of the epitope-tagged protein complexes coexist in the same starting material. In that case, a simple chromatographic column, such as the P11 ion-exchange resin, can be used to separate multiple forms of protein complexes into different fractions prior to immunoaffinity purification. The isolation of various human TATA-binding protein (TBP)-containing complexes, described in the Support Protocol, is a good example of combining the P11 column and immunoaffinity purification. These protocols, collectively, illustrate a powerful methodology in applying epitope tagging and stable cell line approaches for the purification of multisubunit protein complexes from mammalian cells.

PURIFICATION OF MULTISUBUNIT PROTEIN COMPLEXES FROM CLONAL CELL LINES CONSTITUTIVELY EXPRESSING A FLAG-TAGGED PROTEIN

To purify a multisubunit protein complex, the authors usually begin with a constitutive expression system using retrovirus-mediated gene transfer for the establishment of clonal cell lines expressing the FLAG-tagged subunit of a protein complex. Retrovirus-mediated gene transfer (*UNITS 9.9-9.14*) provides a gene delivery system with few copies of the viral integrant to ensure the expressed FLAG-tagged protein can be efficiently assembled into a large protein complex, in a stoichiometric ratio, with endogenous cellular proteins. A retroviral vector containing a drug-selection marker and the FLAG sequence introduced either at the N-terminus or at the C-terminus of the protein-coding region is first created and delivered into a packaging cell line that provides all the viral proteins essential for virion assembly. Viral particles are then collected from the cell culture medium and used to infect HeLa or other chosen cells. Cellular clones potentially expressing the tagged protein are initially selected by drug resistance and further identified by immunoblotting from individually expanded colonies. Clonal cell lines expressing the FLAG-tagged protein are then adapted to suspension culture and used for immunoaffinity purification of the epitope-tagged protein complexes.

*NOTE:*All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

Materials

FLAG-tagged protein-coding sequence
Retroviral vector (e.g., pBabe neo, Morgenstern and Land, 1990; also see *UNITS 9.9-9.14*)
2× HEPES-buffered saline (HBS), pH 7.12 (*UNIT 16.14*)
0.25 M CaCl₂
Retrovirus-packaging cells (e.g., Ψ_{CRIP} , Danos and Mulligan, 1988)
Dulbecco's modified Eagle's medium (DMEM) with and without 10% calf serum
15% (v/v) glycerol/DMEM
DMEM with 10% fetal bovine serum (FBS; *APPENDIX 3F*)
2.5 mg/ml polybrene stock solution
50% confluent HeLa cells
Drug-containing selection medium (e.g., 1 mg/ml of G418 (total weight), 0.2 mg/ml of hygromycin, or 0.5 $\mu\text{g/ml}$ of puromycin; see *UNIT 9.5*)
Trypsin-EDTA (*APPENDIX 3F*)
1× PBS (*APPENDIX 2*)
1× SDS-PAGE protein sample buffer (*UNIT 10.2A*)
Anti-FLAG M2 monoclonal antibody (and/or anti-protein antibodies; Sigma)
Joklik's medium (Sigma) with 10% calf serum and 5% calf serum
Anti-FLAG M2 monoclonal antibody-conjugated beads (Sigma)
BC100 (see recipe)
BC300 (see recipe)/0.1% Nonidet P-40
FLAG peptide (100 mg/ml in water)
Liquid nitrogen
5-ml round-bottom Falcon tube
60-mm and 100-mm tissue culture plates
15-ml sterile tube
0.45- μm cellulose acetate syringe filter
24-well tissue culture plate
1.5-ml and 0.5-ml microcentrifuge tubes

250-ml, 500-ml, 1-liter, 3-liter, and 12-liter spinner flasks
End-over-end tube rotator
Rotor (e.g., Sorvall H-6000A or equivalent)
Microcentrifuge spin column

Additional reagents and equipment for use of retroviral vectors (*UNITS 9.9-9.14*),
preparation of nuclear extracts (*UNIT 12.1*), and immunoblotting (*UNIT 10.8*)

Establish stable cell lines by retrovirus-mediated gene transfer

1. Clone the FLAG-tagged protein-coding sequence into a retroviral vector containing a drug-selection marker (e.g., neomycin).

UNITS 9.9-9.14 describe the use of retroviral vectors.

2. Transfer 20 μg plasmid DNA into a 5-ml round-bottom Falcon tube and mix with 0.5 ml of 2 \times HBS, pH 7.12.

The following transfection procedure works well for Ψ_{CRIP} and many other packaging cell lines.

3. Add 0.5 ml of 0.25 M CaCl_2 dropwise while vortexing. Leave for 20 to 30 min at room temperature.
4. Vortex the calcium phosphate–DNA coprecipitates vigorously and then add the solution to 50% confluent Ψ_{CRIP} cells, which are maintained in DMEM plus 10% calf serum in a 100-mm tissue culture plate. Leave cells for 3 to 4 hr in a 37°C, 5% CO_2 incubator.
5. Remove medium and add 3 ml of 15% (v/v) glycerol/DMEM for 3 min at room temperature.
6. Add 8 ml DMEM to quickly dilute glycerol, mix well, and remove solution.
7. Repeat washes three additional times with 3 ml DMEM for each wash.
8. Add 10 ml DMEM with 10% calf serum to transfected cells and leave plates for 18 hr in a 37°C, 5% CO_2 incubator.
9. Collect viral particles in a 15-ml sterile tube after filtering cell culture supernatant (18-hr post glycerol shock) through a 0.45- μm cellulose acetate syringe filter to remove floating cells and large cellular debris.
10. Mix 1 ml of filtered viral stock with 1 ml of DMEM plus 10% FBS and polybrene (final concentration 8 $\mu\text{g}/\text{ml}$), and incubate mixture with 50% confluent HeLa cells, which are grown in DMEM with 10% FBS in 100-mm cell culture plates and have been prewashed with medium two times, for 2.5 hr at 37°C with intermittent mixing.
11. Add an additional 8 ml of DMEM with 10% FBS, without removing the original infection medium, and leave cells for 1 to 1.5 days in a 37°C, 5% CO_2 incubator.
12. Split cells 1:5 in drug-containing medium and change the selection medium every 3 to 4 days.

The drug concentrations normally used for HeLa cells, depending on individual drug-selection markers, are: 1 mg/ml of G418 (total weight), 0.2 mg/ml of hygromycin, or 0.5 $\mu\text{g}/\text{ml}$ of puromycin. See UNIT 9.5 for additional details.

Drug-resistant colonies are usually visible after 2 to 3 weeks.

Identify clonal cell lines expressing FLAG-tagged protein

13. Prepare a 24-well plate and add 2 drops (~100 μ l) of trypsin-EDTA to each well at alternate positions to avoid cross contamination between neighboring wells.
14. Circle the drug-resistant colonies (~2- to 3-mm diameter), to be picked up, on the bottom of the plate with a black (or blue) marker pen, remove medium, wash cells once with 1 \times PBS, and then remove solution.
15. Take ~50 μ l of trypsin-EDTA from the first well using a 200- μ l pipet tip and transfer a circled colony back to the first well after pipetting up and down several times.
16. Repeat the transfer process for a total of 12 colonies.
17. Add 1 to 2 ml of selection medium to each well and put the 24-well cell culture plate back to a 37°C, 5% CO₂ incubator for cells to attach and grow.

Depending on the number of viable cells transferred, it may take 3 to 14 days for cells to become confluent in the wells.

18. Transfer cells to a 60-mm cell culture plate when they become confluent in the 24-well plate and continue the expansion for individual cellular clones.
19. When cells are nearly confluent on the 60-mm cell culture plate, split cells again to one 60-mm cell culture plate and two 100-mm cell culture plates.

The cells grown on the 100-mm plates will be frozen down as individual cloned cell lines.

20. Prepare whole-cell lysates from 80% to 90% confluent cells still in log phase grown on the 60-mm cell culture plate by adding 300 μ l of 1 \times SDS-PAGE protein sample buffer to plate, after washing cells two times with 1 \times PBS. Pipet lysate up and down many times until the sample is no longer viscous, and then transfer the lysate to a 1.5-ml microcentrifuge tube.
21. Perform immunoblotting (*UNIT 10.8*) on individually collected whole-cell lysates with anti-FLAG M2 monoclonal antibody and/or anti-protein antibodies to identify the cellular clones that express FLAG-tagged protein.

Purify FLAG-tagged protein complexes from established cell lines

22. Select a positive clone expressing FLAG-tagged protein for further expansion by adapting it to suspension culture. Combine cells from twelve 100-mm cell culture plates into a 250-ml spinner flask and adjust the cell density to 0.5×10^6 cells/ml using Joklik's medium with 10% FBS. Maintain cells at the same density using Joklik's medium with 10% FBS for ≥ 3 days. If cells are healthy and nearly grow exponentially, start to feed cells using Joklik's medium with 10% calf serum. Continue feeding cells using Joklik's medium with 10% calf serum for at least three days. If cells still look healthy and grow well, begin to feed cells using Joklik's medium with 7.5% (or 5%) calf serum.

During the expansion and serum-switching/reducing process, if cells look unhealthy or cell density drops, centrifuge cells for 5 min at $300 \times g$ and resuspend cells in fresh Joklik's medium with the same concentration of calf serum, this helps to remove cell debris. Readjust the cell density to 0.5×10^6 cells/ml. Once cells are healthy and continue to grow, begin to lower the concentration of serum as previously described. Repeat the process until cells are maintained in Joklik's medium with 5% calf serum.

23. Grow and expand the cells growing in Joklik's medium with 5% calf serum continually to 20 liters or any other desired volume.

24. Prepare nuclear extract, cytoplasmic S100, and nuclear pellet (or chromosomal fraction) according to *UNIT 12.1* (Dignam et al., 1983).
25. Perform immunoblotting (*UNIT 10.8*) to identify the cellular fractions containing FLAG-tagged protein.
26. Add 0.4 ml of anti-FLAG M2 monoclonal antibody-conjugated beads, which have been prewashed several times with BC100, to a 15-ml tube containing 10 to 14 ml of nuclear extract (or S100, or solubilized nuclear pellet). Incubate the sample for 6 to 12 hr at 4°C on an end-over-end tube rotator.
27. Centrifuge for 2 min at $300 \times g$ (Sorvall H-6000A rotor 1000 rpm), 4°C, to separate supernatant from beads.
28. Remove the supernatant and wash protein-bound beads five times with BC300/0.1% Nonidet P-40.
Inclusion of Nonidet P-40 prevents protein/bead aggregation and also helps the immunoaffinity beads rotate better during the elution step.
29. Transfer the M2-conjugated beads with bound proteins to a microcentrifuge spin column.
30. Microcentrifuge for 10 sec at maximum speed (10,000 rpm) to remove the residual liquid.
31. Add 0.5 ml of BC300/0.1% Nonidet P-40 plus 0.2 mg/ml of FLAG peptide to the dried beads. Mix well and incubate for 20 to 60 min at 4°C with continuous rotation.
32. Collect the eluate (i.e., the first elution) by centrifuging for 10 sec at 4°C.
33. Repeat the elutions (steps 31 and 32) for a total of four times to recover the majority of FLAG-tagged protein complexes.
34. Dispense samples into small aliquots (e.g., 20 μ l) in 0.5-ml microcentrifuge tubes and freeze in liquid nitrogen. Store aliquots of the purified complexes at -80°C until use.

PURIFICATION OF MULTISUBUNIT PROTEIN COMPLEXES FROM CLONAL CELL LINES CONDITIONALLY EXPRESSING THE FLAG-TAGGED PROTEIN

Sometimes, the introduced FLAG-tagged protein-coding sequence is not expressed in the established clonal cell lines following retrovirus-mediated gene transfer and drug selection. This is likely due to cytotoxicity resulting from unregulated or constitutive expression of the ectopic gene product. To overcome this problem, an inducible mammalian expression system can be used to turn “on” or “off” expression of the FLAG-tagged protein based on the use of tetracycline. In the process of establishing FLAG-tagged protein-expressing cell lines, tetracycline is continuously added to silence ectopic gene expression. The inducibility of individual cellular clones is then screened by immunoblotting using whole-cell lysates collected in the presence and absence of tetracycline. Once identified, the inducible cell line is adapted to suspension culture and used for immunoaffinity purification of FLAG-tagged multisubunit protein complexes after removing tetracycline.

**BASIC
PROTOCOL 2**

**Protein
Expression**

16.13.5

Materials

Protein-coding sequence
Tetracycline-regulated expression plasmid (e.g., pTetCMV-F^o(S), Wu and Chiang, 1996; UNIT 16.14)
Sheared calf thymus DNA
SacI-linearized plasmid containing a drug-selection marker (e.g., pREP4, Invitrogen)
Cell line expressing a tetracycline-controlled transactivator (e.g., HtTA-1, Gossen and Bujard, 1992)
DMEM with 10% FBS (APPENDIX 3F)
G418 sulfate
Trypsin-EDTA (APPENDIX 3F)
1 M BES buffer, pH 7.2 (see recipe)
Hygromycin B
Tetracycline
1× SDS-PAGE protein sample buffer (UNIT 10.2A)
1× PBS (APPENDIX 2)
Anti-FLAG M2 monoclonal antibody and/or anti-protein antibodies
Joklik's medium (Sigma) with 10% calf serum
Joklik's medium with 5% calf serum
HeLa cells
Anti-FLAG M2-conjugated beads
BC100 and BC300 (see recipe)
10% Nonidet P-40
FLAG peptide
Liquid nitrogen

0.4-cm cuvettes for electroporation
24-well, 60-mm, and 100-mm tissue culture plates
15-ml and 50-ml Falcon tubes
Rotor (e.g., Sorvall H-6000A or equivalent)
Electroporator (e.g., Bio-Rad Gene Pulser II)
Disposable glass pipet
1.5-ml and 0.5-ml microcentrifuge tubes
250-ml, 500-ml, 1-liter, 3-liter, and 12-liter spinner flasks
250-ml sterile conical centrifuge tubes
Microcentrifuge spin column

Additional reagents and equipment for tissue culture (APPENDIX 3F) and immunoblotting (UNIT 10.8)

Establish tetracycline-regulated cell lines by electroporation

1. Clone the protein-coding sequence into a tetracycline-regulated expression plasmid containing the FLAG epitope sequence, such as pTetCMV-F^o(S).
2. Mix 10 μg of linearized tetracycline-regulated FLAG-tagged protein-expressing plasmid (e.g., *PvuI* used to linearize most pTetCMV-F^o(S)-derived constructs), 50 μg of sheared calf thymus DNA, and 0.5 μg of *SacI*-linearized pREP4, which contains a hygromycin-selection marker, into an electroporation cuvette.
3. Combine HtTA-1 cells (or other cells expressing a tetracycline-controlled transactivator), which are grown in DMEM with 10% FBS containing 0.6 mg/ml of G418, from several 100-mm cell culture plates into a 50-ml Falcon tube after trypsin-EDTA treatment, centrifuge for 5 min at 300 × *g* (Sorvall H-6000A rotor 1000 rpm), and

resuspend cell pellets in DMEM with 10% FBS and 5 mM BES buffer, pH 7.2, at a density of 2×10^7 cells/ml.

4. Add 250 μ l of resuspended cells, using a 200- μ l pipet tip to transfer 125 μ l of cells two times, to the electroporation cuvette containing plasmids and carrier DNA (see step 2). Mix well by pipetting up and down several times.
5. Electroporate DNAs into cells using the Bio-Rad Gene Pulser II with setting at 960 μ F and 200 V. Leave cuvette for 10 min at room temperature and then transfer cells, using a disposable glass pipet, to a 15-ml centrifuge tube containing 10 ml DMEM with 10% FBS.
6. Centrifuge for 5 min at $300 \times g$. Aspirate and discard the supernatant, which contains untransfected DNAs and cell debris generated by electroporation.
7. Resuspend cells in 5 ml DMEM with 10% FBS and distribute cells onto five 100-mm cell culture plates, each containing 9 ml DMEM with 10% FBS, and leave for 2 days in a 37°C , 5% CO_2 incubator.
8. Initiate drug selection by changing medium to DMEM with 10% FBS containing 400 $\mu\text{g/ml}$ of G418, 200 $\mu\text{g/ml}$ of hygromycin B, and 2 $\mu\text{g/ml}$ of tetracycline.
9. Change medium every 3 or 4 days and continue drug selection for ~ 3 weeks until drug-resistant colonies are clearly visible.

Identify clonal cell lines conditionally expressing FLAG-tagged protein

10. Pick up 12 to 24 drug-resistant colonies and expand them into cell lines following the procedures described in Basic Protocol 1, steps 13 to 18.
11. When cells are nearly confluent on the 60-mm cell culture plate during expansion, split cells to one 100-mm cell culture plate and two 60-mm cell culture plates.

The cells grown on the 100-mm cell culture plate will be frozen down as individual cloned cell lines. The cells grown on the two 60-mm cell culture plates are plated in selection medium, one with and the other without tetracycline, for 48 to 72 hr.

12. Prepare whole-cell lysates from cells grown on the two 60-mm cell culture plates by adding 300 μ l of $1\times$ SDS-PAGE protein sample buffer to each plate, after washing the cells two times with $1\times$ PBS. Pipet the lysate up and down many times until the sample is no longer viscous and then transfer lysate to a 1.5-ml microcentrifuge tube.
13. Perform immunoblotting (UNIT 10.8) on an aliquot of the individually collected whole cell lysates with anti-FLAG M2 monoclonal antibody and/or anti-protein antibodies to identify the cellular clones that express FLAG-tagged protein only in the absence of tetracycline.

Purify FLAG-tagged protein complexes from established cell lines

14. Select a positive clone expressing FLAG-tagged protein for further expansion by adapting it to suspension culture. Combine cells from twelve 100-mm cell culture plates into a 250-ml spinner flask (at a density of 0.5×10^6 cells/ml) and gradually replace medium from DMEM with 10% FBS to Joklik's medium with 10% calf serum, and eventually to Joklik's medium with 5% calf serum, but include 1 $\mu\text{g/ml}$ of tetracycline and 0.6 mg/ml of G418 to the medium throughout the expansion process until culture volume reaches 250 ml.
15. Continue to expand cells in tetracycline-containing medium, but omit G418 after 250 ml culture.

16. After expansion of cell culture to 12 liters, remove tetracycline by centrifuging cells for 5 min at $300 \times g$ (Sorvall H-6000A rotor 1000 rpm) using 250-ml sterile conical tubes. Repeat the process until all of the cells are consolidated into two 250-ml conical tubes.
17. Wash cell pellets with 250 ml of $1 \times$ PBS in each tube. Centrifuge for 5 min at $300 \times g$ (Sorvall H-6000A rotor 1000 rpm). Repeat washes for a total of 4 to 6 times.
18. Resuspend cell pellets in ~ 120 ml of Joklik's medium with 5% calf serum without tetracycline, and dispense cells into six 12-liter flasks, each containing 1.5 liters of Joklik's medium with 5% calf serum.
19. Four days after protein induction with medium doubling every day, prepare nuclear extract, cytoplasmic S100, and nuclear pellet (or chromosomal fraction) according to *UNIT 12.1*.

Protein induction in suspension culture usually takes longer than induction in monolayer culture. From the time course study, protein induction after 4 days maximizes the quantity of the induced FLAG-tagged protein, in part due to a reduced doubling time in a larger volume of suspension culture using less expensive and lower concentrations of calf serum. Induction also appears to work better under a slightly dilute condition with medium doubling every day.

20. Perform immunoblotting (*UNIT 10.8*) to identify the cellular fractions containing FLAG-tagged protein.
21. For immunoaffinity purification of FLAG-tagged protein complexes, follow the same procedures as described in Basic Protocol 1, steps 26 to 34.

**ALTERNATE
PROTOCOL**

PURIFICATION OF MULTIPLE FORMS OF EPITOPE-TAGGED PROTEIN COMPLEXES BY VARYING THE STARTING MATERIAL AND WASH CONDITIONS

Using human RNA polymerase II (pol II) as an example, this protocol describes the purification of epitope-tagged protein complexes. RNA pol II is a multisubunit protein complex containing 12 polypeptides (RPB1-12). The largest subunit (RPB1) of pol II has a unique C-terminal domain (CTD) containing heptapeptide sequence, Tyr-Ser-Pro-Thr-Ser-Pro-Ser (YSPTSPS), which occurs 26 times in yeast and 52 times in humans. This CTD can be phosphorylated by many protein kinases, resulting in three forms of pol II commonly found in eukaryotes. The IIO form of pol II contains a highly or hyper-phosphorylated CTD, whereas the IIA form of pol II has a non- or hypo-phosphorylated CTD. The IIB form of pol II does not have the CTD, likely generated due to protease cleavage. In addition, pol II can form an extremely large protein complex, so-called pol II holoenzyme, by association with a subset of general transcription factors and other cellular proteins involved in chromatin remodeling, mRNA processing, and DNA repair. This Alternate Protocol exemplifies an application of the epitope-tagging and stable cell line approaches for the simultaneous purification of the IIO, IIA, and IIB forms of pol II as well as pol II holoenzyme, which contains only the nonphosphorylated form of RPB1. This is achieved simply by careful selection of the starting material and by modifying the composition of protein buffer used for immunoaffinity purification (see Fig. 16.13.1). Please note that the conditions outlined herein will vary with other tagged protein systems.

Additional Materials (also see Basic Protocols 1 and 2)

- 10 M urea
- BC850 (see recipe)
- Buffer B (see recipe)

**Purification of
Epitope-Tagged
Multisubunit
Protein
Complexes**

16.13.8

Buffer D (see recipe)
 FLAG peptide elution buffer (see recipe)
 Ammonium sulfate
 Sonicator
 Rotor (e.g., Bechman 45-Ti or equivalent)

Purify FLAG-tagged human pol II holoenzyme

- 1a. Establish a tetracycline-regulated HeLa-derived cell line conditionally expressing a FLAG-tagged subunit of human pol II, and prepare nuclear extract, cytoplasmic S100, and nuclear pellet according to the procedures described in Basic Protocol 2, steps 1 to 20.
- 2a. For immunoaffinity purification of FLAG-tagged human pol II holoenzyme from nuclear extract or S100, follow Basic Protocol 1, steps 26 to 34, except use BC100/0.1% Nonidet P-40 as wash and elution buffer.

Purify IIA form of FLAG-tagged human pol II

- 1b. For immunoaffinity purification of FLAG-tagged human pol II from nuclear extract or S100, follow Basic Protocol 1, steps 26 to 34, except wash the protein-bound beads sequentially five times each with BC850/0.1% Nonidet P-40 containing 1.0 M urea,

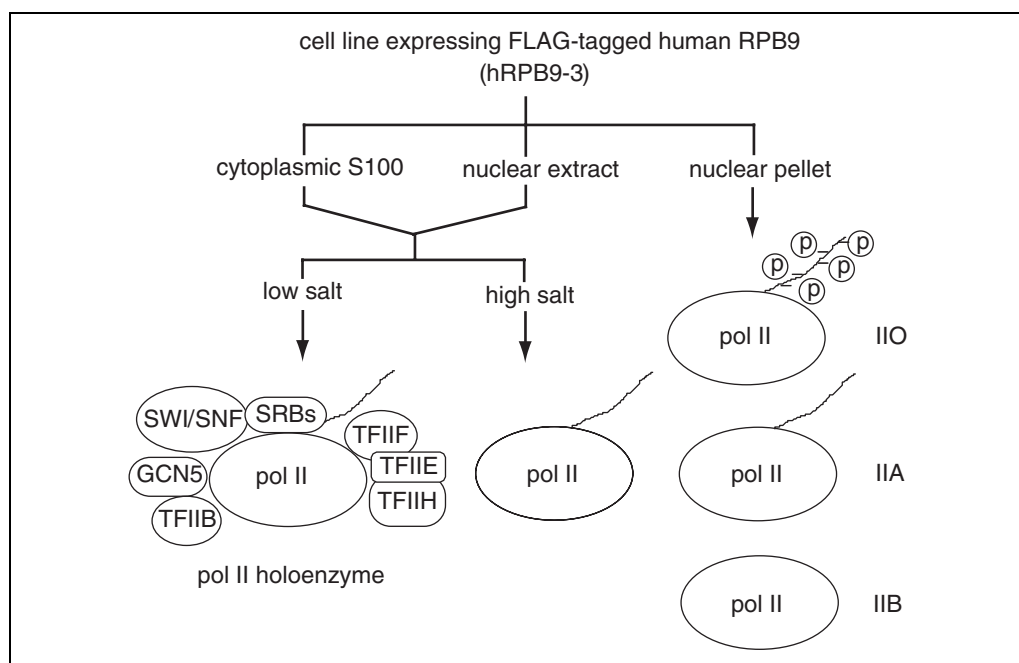


Figure 16.13.1 Purification of human RNA polymerase II complexes. The hRPB9-3 cells (Wu and Chiang, 1998), derived from human HeLa cells that conditionally express the FLAG-tagged RPB9 subunit of human RNA polymerase II, are first separated into cytoplasmic S100, nuclear extract, and nuclear pellet. Immunoaffinity purification is then performed with S100 or nuclear extract under either low salt (100 mM KCl-containing buffer) or high salt (850 mM KCl and 1.0 M urea-containing buffer) wash conditions, which results in the purification of human RNA polymerase II (pol II) holoenzyme and pol II, respectively. The pol II holoenzyme complex purified from S100 or nuclear extract contains pol II, a subset of general transcription factors (TFIIB, TFIIH, TFIIF, and TFIIIE), SRBs, histone acetyltransferase GCN5, and chromatin remodeling factor SWI/SNF (Wu and Chiang, 1998; Wu et al., 1999). Additional forms of pol II complexes with different phosphorylation status on the largest subunit (RPB1) of pol II can also be purified from the nuclear pellet under the high salt wash condition. The IIO form of pol II contains the highly phosphorylated carboxy-terminal domain (CTD), indicated by a tail, in RPB1, whereas the IIA and IIB forms of pol II have non-phosphorylated or truncated CTD in RPB1, respectively.

and BC100, then finally elute the bound proteins with BC100 containing 0.2 mg/ml of FLAG peptide (FLAG peptide elution buffer).

Inclusion of high salt and urea during the washes disrupts the interactions between pol II and other weakly associated polypeptides found in the pol II holoenzyme, thereby allowing the recovery of only the IIA form of the core pol II, which forms a tight complex resistant to high salt and urea washes.

Purify FLAG-tagged human pol II containing a mixture of IIA, IIB, and IIO forms

1c. Slowly agitate 35 ml of nuclear pellet isolated from hRPB9-3, which is a HeLa-derived cell line conditionally expressing the FLAG-tagged RPB9 subunit of human pol II, with 70 ml of buffer B at 4°C (or on ice) and then with 11 ml of 3 M (NH₄)₂SO₄ to adjust the ammonium sulfate concentration to 0.3 M. Stir for an additional 30 min.

The use of buffer B together with ammonium sulfate provides both a magnesium ion and high salt concentration to help dissociate pol II complexes from the chromosomal fraction and at the same time preserve pol II activity.

2c. Sonicate five times, for 1 min each, on an ice water slurry with 20-sec intervals between bursts.

3c. Centrifuge the mixture for 90 min at 185,000 × g (Beckman 45-Ti rotor 40,000 rpm), 4°C, and then pour the supernatant into a 500-ml beaker.

4c. Add 2 volumes of buffer B dropwise with a syringe to gradually adjust the ammonium sulfate concentration to 0.1 M.

5c. Remove precipitated material by centrifuging for 60 min at 185,000 × g (Beckman 45-Ti rotor 40,000 rpm) and pour supernatant to another 500-ml beaker.

6c. Precipitate pol II by slowly adding solid ammonium sulfate to 65% saturation (i.e., 0.42 g/ml of suspension) and stir for an additional 30 min.

7c. Centrifuge for 60 min at 142,000 × g (Beckman 45-Ti rotor 35,000 rpm) and resuspend the pellet in 40 ml of buffer D.

8c. Dialyze the sample against 4 liters of BC100 for 6 hr with at least one change of buffer.

The final recovered sample is ~46 ml.

9c. For immunoaffinity purification of FLAG-tagged human pol II complexes containing a mixture of IIA, IIB, and IIO forms, follow Basic Protocol 1, steps 26 to 34, except sequentially wash the protein-bound beads with five times each of BC850/0.1% Nonidet P-40 with 1.0 M urea, and BC100, before final elutions with BC100 containing 0.2 mg/ml of FLAG peptide (FLAG peptide elution buffer).

The inclusion of high salt and urea during the initial washes disrupts the interactions between pol II and other weakly associated polypeptides found in the pol II holoenzyme, thereby allowing the recovery of core pol II complexes with different phosphorylation states. Equilibration of the immobilized complexes to BC100 is necessary before final elutions with FLAG peptide BC100 elution buffer.

PURIFICATION OF MULTIPLE FORMS OF EPITOPE-TAGGED PROTEIN COMPLEXES FOLLOWING A P11 ION-EXCHANGE CHROMATOGRAPHIC COLUMN

Human TATA-binding protein (TBP) can associate with different classes of cellular proteins to form distinct protein complexes, such as SL1, TFIID, and TFIIB, which are required for transcription by RNA polymerase I, II, and III, respectively. These distinct protein complexes are all found in nuclear extract. A simple chromatographic step, such as a P11 ion-exchange column, is then used to separate various TBP-containing complexes into different fractions prior to immunoaffinity purification (see Fig. 16.13.2). This Support Protocol illustrates an application of conventional chromatography together with immunoaffinity purification for the isolation of multiple protein complexes containing a common subunit.

Additional Materials (also see Basic Protocol 1)

- P11 ion-exchange resin
- Various BC buffers (BC100, BC300, BC500, BC850, and BC1200; see recipe)
- Chromatographic column (~100 ml capacity)
- Flow adapter
- Chart recorder
- Fraction collector
- Dialysis tubing (MWCO 12,000; Sigma)
- Additional reagents and equipment for dialysis (APPENDIX 3C)

Purify TFIID and TFIIB from a cell line expressing FLAG-tagged human TBP

1. Establish a HeLa-derived cell line expressing FLAG-tagged human TBP by retrovirus-mediated gene transfer (UNITS 9.9.9.14) and prepare nuclear extract, S100, and nuclear pellet from the established cell line using the procedures described in Basic Protocol 1, steps 1 to 24.
2. Pack a P11 ion-exchange column containing ~60 ml of resin in a chromatographic column.
3. Connect the column to a flow adapter, a chart recorder, and a fraction collector, and equilibrate the whole system with BC100 overnight. Set the flow rate at ~1 column volume (CV) per hr (i.e., 1 ml/min).

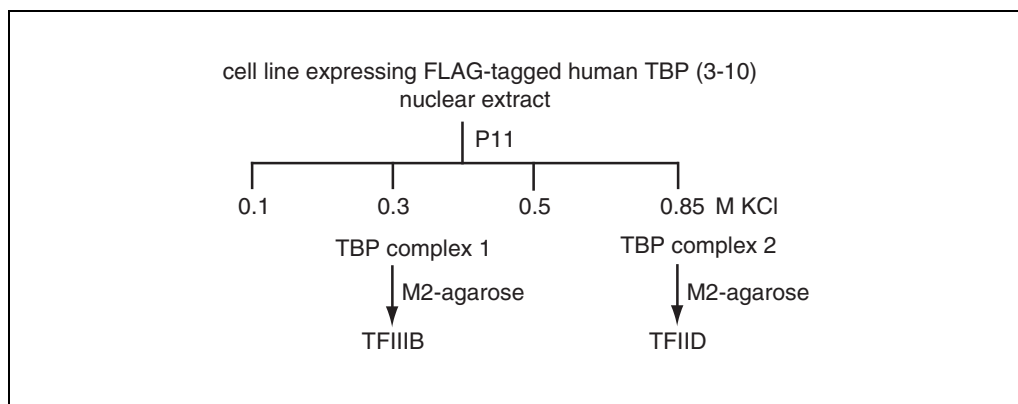


Figure 16.13.2 Purification of different FLAG-tagged TBP complexes. Nuclear extract prepared from 3-10 (Chiang et al., 1993), which is a HeLa-derived cell line constitutively expressing FLAG-tagged TBP, is loaded onto a P11 ion-exchange column. Proteins bound to the P11 resin are sequentially eluted off the column by step elutions with 0.1, 0.3, 0.5, and 0.85 M KCl-containing buffer. The FLAG-tagged TBP complexes are mainly detected at the 0.3 and 0.85 M KCl fractions, which are then used for immunoaffinity purification by anti-FLAG M2 monoclonal antibody-conjugated beads to isolate TFIIB and TFIID complexes, respectively.

4. Thaw ~100 ml of nuclear extract in an ice bucket overnight at 4°C.
5. Centrifuge nuclear extract for 20 min at 105,000 × g (Beckman 45-Ti rotor 30,000 rpm), 4°C.
6. Pour supernatant into 50-ml tubes and load the sample onto an equilibrated P11 column at a flow rate of 1 CV/hr. Collect the eluate, ~12 ml/tube, with a fraction collector.

Remember to save a small aliquot (~100 μl) for immunoblotting analysis.

7. Wash the column with ~2 CV of BC100 or until the absorption profile, monitored at 280 nm, almost reaches the baseline.
8. Step elute the bound proteins by sequentially switching to BC300, BC500, BC850, and BC1200, when the protein curve of each elution buffer nearly reaches the baseline.
9. Combine the fractions representing each of the BC100, BC300, BC500, and BC850 protein peaks.

The BC1200 fractions, containing a minor amount of SL1, are not collected for further purification.

10. Dialyze BC500 and BC850 samples against 4 liters of BC100 for 5 hr at 4°C.
11. Centrifuge samples for 15 min at 46,000 × g (Beckman 45-Ti rotor 20,000 rpm), 4°C.
12. Aliquot the supernatant of BC500 and BC850, as well as the BC100 and BC300 fractions, individually, into several 15-ml tubes (for immunoaffinity purification) and 1.5-ml microcentrifuge tubes (for immunoblotting and protein analysis).
13. Freeze the aliquots in liquid nitrogen and store samples at –80°C until use.
14. Perform immunoblotting (*UNIT 10.8*) to locate the P11 fractions containing FLAG-tagged TBP.
15. Purify TFIIB from the P11 0.3 M KCl fraction and TFIID from the P11 0.85 M KCl fraction according to the procedures described in Basic Protocol 1, steps 26 to 34.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

BES buffer, 1 M (pH 7.2)

21.32 g of BES (*N, N*-bis[2-hydroxyethyl]-2-aminoethanesulfonic acid) (Sigma)
 60 ml H₂O
 Adjust pH to 7.12 with 1 M NaOH and bring up total volume to 100 ml
 Autoclave and store up to 6 months at room temperature

BC100

20 mM Tris·Cl, pH 7.9 at 4°C (*APPENDIX 2*)
 20% glycerol
 0.2 mM EDTA, pH 8.0 (*APPENDIX 2*)
 100 mM KCl
 1 mM DTT (added prior to use)
 0.5 mM PMSF (added prior to use)
 Store up to 6 months at 4°C

BC300, BC500, BC850, and BC1200

Prepare as for BC100 (see recipe) except the KCl concentration is adjusted to 300, 500, 850, and 1200 mM, respectively.

Buffer B

50 mM Tris·Cl, pH 7.9 at 4°C (*APPENDIX 2*)
25% glycerol
5 mM MgCl₂
5 mM EDTA, pH 8.0 (*APPENDIX 2*)
5 mM EGTA, pH 8.0
5 mM DTT (added prior to use)
0.5 mM PMSF (added prior to use)
Store up to 6 months at 4°C

Buffer D

50 mM Tris·Cl, pH 7.9 at 4°C (*APPENDIX 2*)
25% glycerol
5 mM EDTA, pH 8.0 (*APPENDIX 2*)
5 mM EGTA, pH 8.0
2 mM DTT (added prior to use)
0.5 mM PMSF (added prior to use)
Store up to 6 months at 4°C

FLAG peptide elution buffer

Diluting FLAG peptide from 100 mg/ml of FLAG peptide (in water) in BC100 or BC300 (see recipe) prior to use.

COMMENTARY

Background Information

The FLAG epitope consists of 8 amino acids that are mostly charged residues (Hopp et al., 1988). The popularity of this epitope lies in the following facts.

1. It is synthetic, so that very few, if any, cross-reacting species are detected in bacterial, yeast, insect, or mammalian cells.

2. It is very hydrophilic and tends to help proteins stay soluble in aqueous solution.

3. It is small and highly antigenic. The introduction of the FLAG sequence rarely changes protein conformation and has a tendency to stimulate the immune response when injected into animals for antibody production (Chiang and Roeder, 1995).

4. A heart muscle kinase site linked to the FLAG sequence offers a unique way to label proteins with ³²P after purification (Blonar and Rutter, 1992). The labeled proteins can be used for interaction studies, expression library screening and protein tracking (Chiang and Roeder, 1995).

5. A peptide elution method has been developed to elute the FLAG-tagged protein (Chiang and Roeder, 1993) or protein complexes (Chiang et al., 1993) off the anti-FLAG M2

monoclonal antibody-conjugated beads under neutral pH or physiological conditions, thereby allowing the recovery of fully active proteins ready for functional assays.

6. The fold purification achieved by M2-agarose beads, which are commercially available, is exceedingly high, making it possible to purify FLAG-tagged multisubunit protein complexes to near homogeneity via one-step immunoaffinity purification (Chiang et al., 1993; Fondell et al., 1996; Kershner et al., 1998; Ogryzko et al., 1998; Sif et al., 1998; Wu et al., 1998; Gu et al., 1999; Ikura et al., 2000).

Although purification of recombinant FLAG-tagged protein can be easily achieved by using bacterial and baculovirus expression systems (Chiang and Roeder, 1993; Wu et al., 1999), the reconstitution of a fully functional protein complex from individually purified subunits or from insect cells coinfecting with several recombinant baculoviruses expressing different subunits remains challenging (Bell et al., 1995; Chen and Tjian, 1996; Guermah et al., 1998; Phelan et al., 1999). Therefore, we and others have applied retrovirus-mediated gene transfer to deliver an epitope-tagged subunit of a large protein complex into human

HeLa cells for the natural assembly of multisubunit protein complexes in a mammalian cell. Isolation of the *in vivo*-assembled large protein complexes is then facilitated by immunoaffinity purification using epitope-specific monoclonal antibody-conjugated beads and peptide elution (Zhou et al., 1992; Chiang et al., 1993). This methodology, combining epitope-tagging and stable cell line approaches, has made it possible to purify many multisubunit protein complexes, as long as a clonal cell line expressing the tagged subunit of a protein complex is available.

Sometimes, cytotoxicity caused by constitutive expression of the tagged protein prevents the establishment of a stable cell line by retrovirus-mediated gene transfer. The application of an inducible mammalian expression system, in combination with FLAG epitope-tagging and peptide elution methods, has further expanded our ability to establish stable cell lines expressing essentially any epitope-tagged protein (Wu and Chiang, 1996). The purification of human pol II and TFIIF from tetracycline-regulated clonal cell lines conditionally expressing FLAG-tagged RPB9 and the FLAG-tagged p62 subunit of human TFIIF, respectively, was the first documentation detailing the procedures for biochemical purification of multisubunit protein complexes from an inducible mammalian expression system (Kershnar et al., 1998). The protocols described here not only greatly simplify the purification steps, but also dramatically reduce the cost of preparation.

Critical Parameters and Troubleshooting

Location of the epitope tag. In general, the FLAG epitope sequence is introduced at either the N-terminus or the C-terminus of the protein-coding region. A critical parameter is to keep the tag away from essential regions of the protein. In many cases, however, it is hard to predict where the best location is in a protein and which subunit may be exposed on the surface of a protein complex. The suggestion is to tag different subunits in parallel to enhance the successful rate.

Isolation of drug-resistant colonies. The cellular colonies are ready to be picked up when their diameter is between 2-3 mm. However, some slow-growing colonies should not be ignored, as expression of potentially toxic proteins may reduce the growth rate of the cells. The suggestion is to pick up colonies of various

sizes, especially if there are clear differences in the diameter of the colonies.

Identification of positive clones. It is very important to use anti-FLAG M2, not M1, monoclonal antibody for immunoblotting, as anti-FLAG M2 monoclonal antibody can recognize the FLAG sequence at any location on the protein, whereas the M1 monoclonal antibody only recognizes the FLAG sequence at the very N-terminal region of the protein. In addition, a negative result with the M2 monoclonal antibody does not necessarily indicate a failure in establishing clonal cell lines. The FLAG epitope may sometimes be masked by the fusion partner, thereby preventing its detection by anti-epitope antibodies. In this case, it is essential to further confirm the results by anti-protein antibodies.

Induction in suspension culture. It often happens that the tagged protein is not successfully induced after continued expansion in suspension culture. This is probably caused by residual tetracycline, which prevents tetracycline-dependent transactivator (tTA) from binding to the regulated gene, or by a reduced expression of tTA in the clonal cell line. To overcome these problems and to ensure a high level of protein induction, we usually follow these four guidelines. First, preselect cells with antibiotic (G418) in a small culture (~100 ml) in the presence of tetracycline for at least three days before further expansion in the absence of antibiotic (G418). Presumably, this will boost the expression of tTA in the cell, thereby enhancing the level of transactivation from the tetracycline-regulated promoter. Second, wash cells at least 4 to 6 times with 1×PBS to remove tetracycline prior to induction. Third, use 250-ml conical centrifuge tubes instead of 1-liter flat bottom bottles to spin down cells when washing with 1×PBS. Conical centrifuge tubes allow for better cell adhesion; thus medium and PBS washes can be completely removed after each successive centrifugation without losing cell volume. The 1-liter flat bottom centrifuge bottles, although convenient to use, do not permit complete removal of the original medium without disturbing the cell pellet; a minor residual tetracycline contaminant is sufficient to prohibit the induction. Fourth, use tetracycline-free serum as a growth supplement when protein induction is desired. The authors have discovered that some batches of calf serum contain endogenous tetracycline, likely resulting from the use of tetracycline in cattle feed to protect animals from bacterial infection.

Removal of FLAG peptide. The presence of the free FLAG peptide in the purified protein complexes, following elution from the M2-conjugated beads, is unlikely to cause problems in functional assays. In fact, the free peptide can enhance the stability of the purified protein complexes, which are often present in dilute concentrations after the final step of purification. However, if the presence of FLAG peptide presents a problem, e.g., by masking the surfaces of protein complexes or by competing for resin binding to an ion-exchange chromatographic column, a dialysis step can be included to remove the free peptide. Thus, depending on the purpose, FLAG peptide may or may not be removed after immunoaffinity purification.

Specificity of purified protein complexes. Normally, a control purification with extract prepared from the parental cells is conducted in parallel to identify polypeptides binding non-specifically to the M2-conjugated beads. This is critical, especially for the characterization of the protein composition in a purified protein complex. It is also a good idea to perform a small-scale purification, e.g., by using 200 μ l of protein sample and 10 μ l of M2-conjugated beads in each tube (Chiang et al., 1993), in order to compare different washing conditions. Once the condition is optimized, a large-scale purification is then conducted. In this way, the specificity of the protein complexes is enhanced, while the nonspecific background is minimized.

Consideration of elution buffer. The elution conditions also play an important role in immunoaffinity purification. Ideally, proteins are stored in low salt-containing buffers which allow for better manipulation of the salt conditions in later functional assays. But when comparing the elution conditions, we find that some proteins, such as the FLAG-tagged human papillomavirus E2 protein, elute better in 300 mM KCl-containing buffer (Hou et al., 2000), while the majority of other proteins elute well in 100 mM KCl-containing buffer. Inclusion of a minor amount (0.1% to 0.01%) of detergents, such as Nonidet P-40 or Triton X-100, in the elution buffer prevents protein/bead aggregation, allowing the immunoaffinity beads to rotate better during the elution step. A caveat in using detergent, however, is that the activity of some proteins (or complexes) may be sensitive to different concentrations of detergent. Luckily, protein complexes purified with these procedures, including human TFIID, TFIIF,

RNA polymerase II, and RNA polymerase II holoenzyme, are fully functional in both basal and activated transcription (Chiang et al., 1993; Chiang and Roeder, 1995; Kershnar et al., 1998; Wu and Chiang, 1998; Wu et al., 1998, 1999).

Combination of immunoaffinity purification and conventional column chromatography. Additional steps of conventional column chromatography, such as ion-exchange and gel filtration, can be included prior to or following immunoaffinity purification to further enhance the purity of the target protein complex.

Anticipated Results

The percentage of drug-resistant colonies expressing FLAG-tagged protein via retrovirus-mediated gene transfer is theoretically 100%. However, integration of the coding cassette into a transcriptionally silenced region of the cellular chromosomes may prevent the expression of FLAG-tagged protein, thereby reducing the number of positive clones. Even so, there are usually several clonal cell lines expressing FLAG-tagged protein from a dozen cellular isolates. In contrast, only 5 to 10% of clonal cell lines expressing FLAG-tagged protein were identified when the tetracycline-regulated expression system was used to establish stable cell lines. This is in part due to random integration of the tetracycline-regulated expression plasmid into cellular chromosomes. Thus, more drug-resistant colonies (>12) should be expanded when using the inducible expression system.

The level of induced FLAG-tagged protein, which can vary over three orders of magnitude, is modulated by the amount of tetracycline included in the growth medium (Wu and Chiang, 1996). A maximal induction is usually achieved by complete removal of tetracycline. Normally the concentration of tetracycline is kept at either zero or 1 μ g/ml to turn "on" or "off" the expression of the tagged protein throughout the entire experiment.

The purified protein complexes are finally analyzed, in comparison with a control sample purified in parallel, by silver staining and immunoblotting. A purification table listing variables, including protein amount of the starting material, units, specific activity, fold purification, and yield, is typically generated. This is helpful for estimating the efficiency of immunoaffinity purification and other chromatographic steps (Kershnar et al., 1998).

Time Considerations

It usually takes three to four weeks to see drug-resistant colonies and an additional three weeks to identify clonal cell lines expressing FLAG-tagged protein by either retrovirus-mediated gene transfer or the tetracycline-regulated expression system. Expansion of a clonal cell line in suspension culture ready for extract preparation may take another 3 to 4 weeks. Once cells are ready for harvesting, it takes only one day to prepare nuclear extract, S100, and nuclear pellet, and another two days for immunoaffinity purification. If a chromatography step is involved, only one additional day is needed. Taken together, it may take 2 to 3 months to purify a multisubunit protein complex when beginning with an expression plasmid containing the FLAG-tagged protein-coding sequence.

Literature Cited

- Bell, S.P., Mitchell, J., Leber, J., Kobayashi, R., and Stillman, B. 1995. The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing. *Cell* 83:563-568.
- Blanar, M.A. and Rutter, W.J. 1992. Interaction cloning: identification of a helix-loop-helix zipper protein that interacts with c-Fos. *Science* 256:1014-1018.
- Chen, J.L. and Tjian, R. 1996. Reconstitution of TATA-binding protein-associated factor/TATA-binding protein complexes for in vitro transcription. *Methods Enzymol.* 273:208-217.
- Chiang, C.-M. and Roeder, R.G. 1993. Expression and purification of general transcription factors by FLAG epitope-tagging and peptide elution. *Peptide Res.* 6:62-64.
- Chiang, C.-M. and Roeder, R.G. 1995. Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. *Science* 267:531-536.
- Chiang, C.-M., Ge, H., Wang, Z., Hoffmann, A., and Roeder, R.G. 1993. Unique TATA-binding protein-containing complexes and cofactors involved in transcription by RNA polymerases II and III. *EMBO J.* 12:2749-2762.
- Danos, O. and Mulligan, R.C. 1988. Safe and efficient generation of recombinant retroviruses with amphotropic and ecotropic host ranges. *Proc. Natl. Acad. Sci. U.S.A.* 85:6460-6464.
- Dignam J.D., Lebovitz, R.M., and Roeder, R.G. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 11:1475-1489.
- Fondell, J.D., Ge, H., and Roeder, R.G. 1996. Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proc. Natl. Acad. Sci. U.S.A.* 93:8329-8333.
- Gossen, M. and Bujard, H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U.S.A.* 89:5547-5551.
- Gu, W., Malik, S., Ito, M., Yuan, C.X., Fondell, J.D., Zhang, X., Martinez, E., Qin, J., and Roeder, R.G. 1999. A novel human SRB/MED-containing cofactor complex, SMCC, involved in transcription regulation. *Mol. Cell* 3: 97-108.
- Guermah, M., Malik, S., and Roeder, R.G. 1998. Involvement of TFIID and USA components in transcriptional activation of the human immunodeficiency virus promoter by NF- κ B and Sp1. *Mol. Cell. Biol.* 18:3234-3244.
- Hopp, T.P., Prickett, K.S., Price, V.L., Libby, R.T., March, C.J., Cerretti, D.P., Urdal, D.L., and Conlon, P.J. 1988. A short polypeptide marker sequence useful for recombinant protein identification and purification. *Bio-Technology* 6:1204-1210.
- Hou, S.Y., Wu, S.-Y., Zhou, T., Thomas, M.C., and Chiang, C.-M. 2000. Alleviation of human papillomavirus E2-mediated transcriptional repression via formation of a TATA binding protein (or TFIID)-TFIIB-RNA polymerase II-TFIIF preinitiation complex. *Mol. Cell. Biol.* 20:113-125.
- Ikura, T., Ogryzko, V.V., Grigoriev, M., Groisman, R., Wang, J., Horikoshi, M., Scully, R., Qin, J., and Nakatani, Y. 2000. Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis. *Cell* 102:463-473.
- Kershnar, E., Wu, S.-Y., and Chiang, C.-M. 1998. Immunoaffinity purification and functional characterization of human transcription factor IIIH and RNA polymerase II from clonal cell lines that conditionally express epitope-tagged subunits of the multiprotein complexes. *J. Biol. Chem.* 273:34444-34453.
- Morgenstern, J.P. and Land, H. 1990. Advanced mammalian gene transfer: High titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucleic Acids Res.* 18:3587-3596.
- Ogryzko, V.V., Kotani, T., Zhang, X., Schlitz, R.L., Howard, T., Yang, X.J., Howard, B.H., Qin, J., and Nakatani, Y. 1998. Histone-like TAFs within the PCAF histone acetylase complex. *Cell* 94:35-44.
- Phelan, M.L., Sif, S., Narlikar, G.J., and Kingston, R.E. 1999. Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol. Cell* 3:247-253.
- Sif, S., Stukenberg, P.T., Kirschner, M.W., and Kingston, R.E. 1998. Mitotic inactivation of a human SWI/SNF chromatin remodeling complex. *Genes Dev.* 12:2842-2851.
- Wu, S.-Y. and Chiang, C.-M. 1996. Establishment of stable cell lines expressing potentially toxic proteins by tetracycline-regulated and epitope-tagging methods. *BioTechniques* 21:718-725.
- Wu, S.-Y. and Chiang, C.-M. 1998. Properties of PC4 and an RNA polymerase II complex in

- directing activated and basal transcription *in vitro*. *J. Biol. Chem.* 273:12492-12498.
- Wu, S.-Y., Kershnar, E., and Chiang, C.-M. 1998. TAF_{II}-independent activation mediated by human TBP in the presence of the positive cofactor PC4. *EMBO J.* 17:4478-4490.
- Wu, S.-Y., Thomas, M.C., Hou, S.Y., Likhite, V., and Chiang, C.-M. 1999. Isolation of mouse TFIID and functional characterization of TBP and TFIID in mediating estrogen receptor and chromatin transcription. *J. Biol. Chem.* 274:23480-23490.
- Zhou, Q., Lieberman, P.M., Boyer, T.G., and Berk, A.J. 1992. Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Genes Dev.* 6:1964-1974.
- Chiang and Roeder, 1993. See above.
This paper documents the development of the FLAG peptide elution method for purification of recombinant FLAG-tagged protein from bacteria.
- Kershnar, et al., 1998. See above.
This paper is the first documentation detailing the procedures for biochemical purification of multisubunit protein complexes from an inducible mammalian expression system.
- Wu and Chiang, 1996. See above.
This paper describes the construction of tetracycline-regulated FLAG-tagged expression plasmids, and the isolation and characterization of the inducible cell lines.

Key References

Chiang et al., 1993. See above.

This paper describes the purification of human TFIID and TFIIB complexes from a stable cell line constitutively expressing FLAG-tagged TBP.

Contributed by Shwu-Yuan Wu and
Cheng-Ming Chiang
Case Western Reserve University School
of Medicine
Cleveland, Ohio

Inducible Gene Expression Using an Autoregulatory, Tetracycline-Controlled System

Tetracycline-regulated gene expression systems have been developed to overcome some of the obstacles encountered using other strategies for inducible gene expression in mammalian cells. These difficulties include pleiotropic, nonspecific effects or toxicity of inducing agents or treatments, and high uninduced background levels of expression. This unit describes protocols for using a modified tetracycline-regulated system in which a transcriptional transactivator drives expression of itself and a target gene in cultured cells and, to some extent, in transgenic mice. This transactivator (tTA) is a fusion protein consisting of the tetracycline-repressor of *E. coli* and the transcriptional activation domain of the VP16 protein of herpes simplex virus. In the absence of tetracycline, tTA binds to and activates genes preceded by a heptamerized version of the tetracycline-resistance operator of Tn10 plus a minimal CMV promoter (here collectively referred to as Tet P). Binding of tTA to Tet P and subsequent gene activation are blocked in the presence of tetracycline. The plasmid pTet-Splice (Fig. 16.14.1A) contains Tet P upstream, and SV40 splice and polyadenylation signals downstream, of a multiple cloning site into which sequences encoding the open reading frame (ORF) of a target gene of choice is easily inserted. Autoregulatory tTA expression is driven from the plasmid pTet-tTAK (Fig. 16.14.1B), in which the tTA ORF (including an optimal sequence for initiation of translation according to Kozak) has been inserted into pTet-Splice.

The protocols in this unit describe the transfection of adherent cells and the testing of resultant clones for inducible transactivator or target gene protein expression. Stably transfected fibroblast cell lines expressing transactivator and target gene(s) can be derived by first cotransfecting pTet-tTAK and a plasmid encoding a selectable marker and obtaining stable lines with inducible transactivator expression (see Basic Protocol). These lines are subsequently stably cotransfected with plasmids encoding the target gene(s) and a second selectable marker. The procedure may also be used to cotransfect pTet-tTAK with the target gene–encoding plasmid(s) and a single selectable marker plasmid. The choice of method depends upon the feasibility of screening for the protein products of the target genes. While the consecutive method is more systematic, cotransfection may be faster given a relatively straightforward screening method for expression of the target gene (see Critical Parameters).

A Support Protocol also describes methods to test stably transfected cell lines for inducible gene expression, for transient transfection and induction of tet-regulated plasmids, and for detection of the tTAK gene in cells (or transgenic mice).

CALCIUM PHOSPHATE-MEDIATED STABLE TRANSFECTION OF NIH3T3 CELLS WITH pTet-tTAK AND TETRACYCLINE-REGULATED TARGET PLASMIDS

This protocol describes the stable transfection of adherent cells with pTet-tTAK for the derivation of cell lines expressing inducible tTA. In the first round of transfection stable cell lines expressing inducible tTA alone are produced. The single transfection procedure may also be used for stable cotransfection of pTet-tTAK and plasmids expressing the target gene(s). In the second round of transfection tTA expressing lines are transfected with plasmids expressing the target gene(s).

**BASIC
PROTOCOL**

**Protein
Expression**

16.14.1

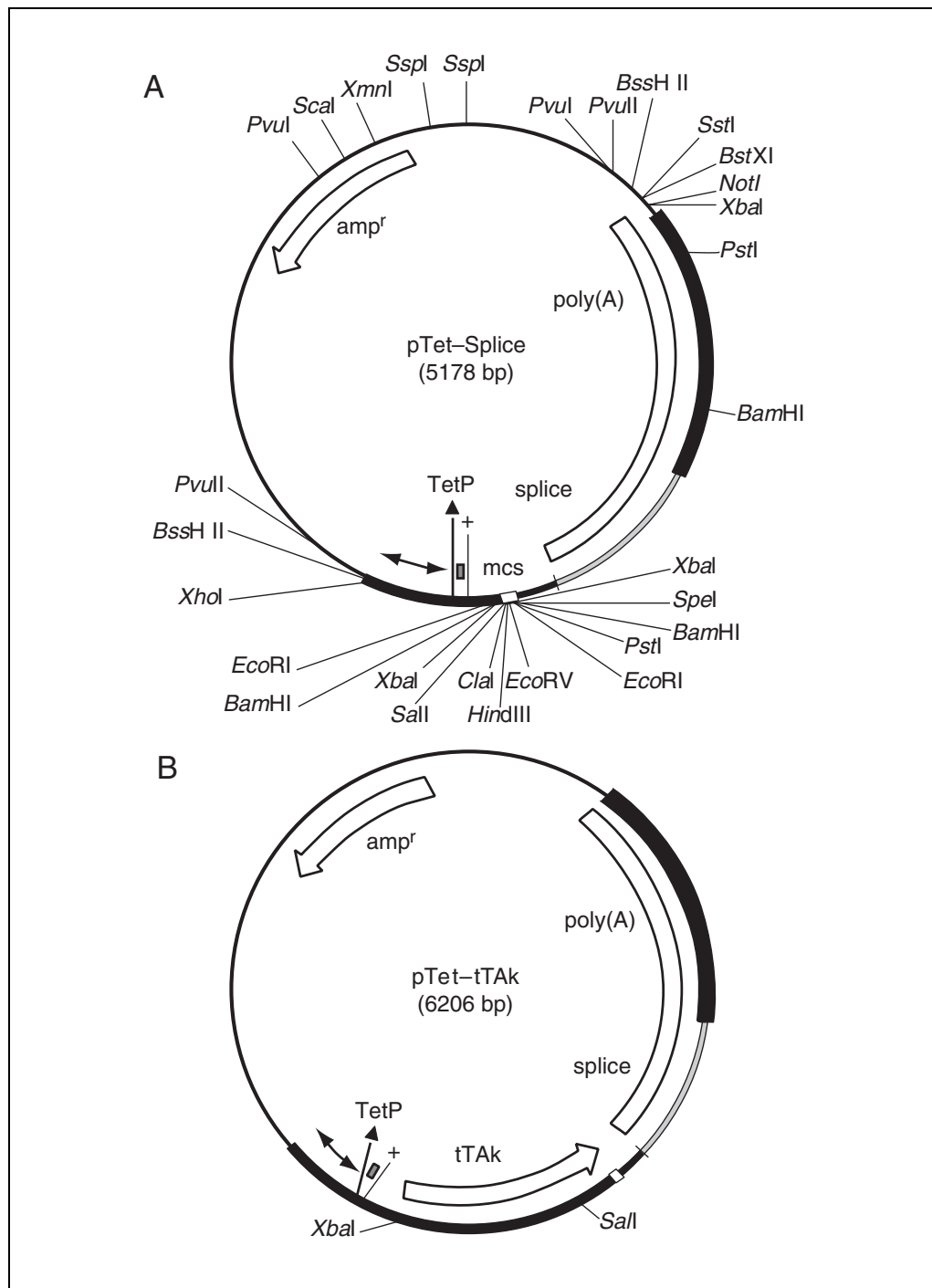


Figure 16.14.1 (A) The plasmid pTet-Splice (Shockett et al., 1995) is designed to drive tetracycline-regulated expression of a target gene inserted into the multiple cloning site (mcs). The tetracycline-regulated promoter (TetP) consists of a heptamerized tetracycline operator (double-ended arrow) upstream of a minimal human CMV promoter that includes bases -53 (triangle) to +75. The transcriptional start site (+) and TATAA box (small rectangle) are also indicated. This TetP fragment is an *XhoI-Sall* fragment derived from pUHC13-3 (Gossen and Bujard, 1992). SV40-derived sequences downstream of the MCS drive mRNA splicing and polyadenylation. The backbone of the plasmid is from Bluescript II KS+ (Stratagene) and carries the ampicillin resistance gene (*amp^r*). (B) pTet-tTak (Shockett et al., 1995) consists of the tTak open reading frame inserted into the *HindIII-EcoRV* sites of pTet-Splice.

Materials

NIH3T3 cells
Complete DMEM-10 medium (see recipe)
Complete DMEM/tet: complete DMEM-10 medium (see recipe) containing 0.5 µg/ml tetracycline hydrochloride (Sigma; dilute 10 mg/ml stock in 70% ethanol and store protected from light at -20°C)
Selection medium (see recipe) containing 125 µM, 250 µM, or 500 µM L-histidinol
Plasmids for first-round or cotransfection procedure: pTet-tTAk (Life Technologies) and plasmids containing target gene ORF(s) cloned into pTet-Splice (Life Technologies), pSV2-His, or another selectable marker plasmid; purified by CsCl banding (UNIT 1.7) or anion-exchange chromatography (UNIT 2.1B)
Plasmids for second round transfection procedure: plasmids containing target gene ORF(s) cloned into pTet-Splice, pPGKpuro, or another selectable marker plasmid; purified by CsCl banding (UNIT 1.7) or anion-exchange chromatography (UNIT 2.1B)
2 M CaCl₂
HEPES-buffered saline (HeBS; see recipe)
10 mg/ml chloroquine (19 mM; optional; Sigma); dilute in water and store at -20°C
85% HeBS/15% glycerol, prewarmed to 37°C
3 mg/ml puromycin (Sigma) diluted in PBS (APPENDIX 2)
Phosphate-buffered saline (PBS; APPENDIX 2)
1× trypsin/EDTA (Life Technologies)
10-cm and 6-cm tissue culture plates
4-ml polystyrene tubes (Falcon)
24-well and 6-well tissue culture plates

NOTE: All tissue culture incubations are performed in a humidified 37°C, 5% CO₂ incubator.

Grow the cells

1. *First round only:* Grow cells in complete DMEM-10 medium. The day before transfection split cells into 10-cm tissue culture plates in complete DMEM/tet to achieve one-third confluence on the day of the transfection.

From this point on cells are kept in the presence of 0.5 µg/ml tet.

One plate per transfection is needed at this stage. A typical experiment might include one plate for tTA only, one for tTA plus target gene, and one to serve as the untransfected control plate.

Second round only: Grow stable cell lines that inducibly express autoregulatory tTA in selection medium/500 µM L-histidinol. The day before transfection split into 10-cm plates in this same medium to achieve one-third confluence on the day of transfection.

Transfect the cells

2. Linearize plasmids prior to transfection and adjust concentration to ≥0.5 mg/ml.

See Damke et al. (1995) for discussion of other selectable markers. All plasmids should be purified by CsCl banding (UNIT 1.7) or on a Qiagen column (UNIT 2.1B).

3. *First round only:* Mix 10 to 20 µg of pTet-tTAk (in the presence or absence of an equimolar amount of target gene plasmids) plus 1 to 2 µg pSV2-His (a molar ratio

of ~10:1 of each tet plasmid to selectable marker plasmid) with 500 μ l HeBS in a clear 4-ml polystyrene tube.

A control mock transfection should be performed with no DNA added to the transfection. All of these cells should die in the selection medium/125 μ M L-histidinol introduced in step 14.

Second round only: Mix 10 to 20 μ g each of target gene plasmid(s) plus 1 to 2 μ g pPGKPuro (a molar ratio of ~10:1 of each tet plasmid to selectable marker plasmid) with 500 μ l HeBS in a clear 4-ml polystyrene tube.

A control mock transfection should be performed with no DNA added to the transfection. All of these cells should die in the presence of the puromycin introduced in step 14. The optimal killing concentration for puromycin (lowest dose between 0.1 μ g/ml to 10 μ g/ml that kills all untransfected cells within a few days) should be determined empirically prior to the transfection and varies with the cell type.

4. Add 32.5 μ l of 2 M CaCl₂ to plasmid DNA and mix immediately by gentle vortexing. With occasional gentle mixing, allow precipitate to form for 15 to 30 min at room temperature or until solution is visibly cloudy when compared to a tube containing water.
5. Aspirate all of the medium from cells, doing one plate at a time. Mix precipitate a few times by pipetting with a Pasteur pipet, and apply dropwise and evenly over cells.
6. Incubate 30 min, gently rocking the plate after 15 min to ensure even coverage over entire plate.
7. *First round only:* Add 10 ml complete DMEM/tet, with or without 25 μ M chloroquine (final), to each plate.

Although the use of chloroquine may further reduce cell integrity during the glycerol shock (step 9), it can improve transfection efficiency.

Second round only: Add 10 ml selection medium/500 μ M L-histidinol, with or without 25 μ M chloroquine (final), to each plate.

8. Incubate 4 to 5 hr.

The optimal length of incubation may vary for different cell types.

9. Gently aspirate medium from cells with minimal disruption of the precipitate that has settled onto the cells. Shock cells by adding dropwise 2.5 ml of prewarmed 85% HeBS/15% glycerol.

It is normal for the cells to look somewhat ragged before and especially after glycerol shock. Two to four plates may be shocked at one time, depending on the speed of the researcher.

10. Aspirate HeBS/glycerol after exactly 2.5 min. Work quickly, as glycerol can be very toxic to the cells.

The length of time cells are exposed to glycerol solution can be varied and increased up to 4 to 5 min to optimize transfection efficiency for different cell types. Cells should be shocked the maximal length of time which results in the least cell death.

11. *First round only:* Immediately, gently, and quickly wash cells twice by adding 10 ml complete DMEM/tet and immediately aspirating.

Because cells tend to come loose from the plate after glycerol addition, add all medium to a single spot on the plate.

Second round only: Immediately, gently, and quickly wash cells twice by adding 10 ml selection medium/500 μ M L-histidinol and immediately aspirating.

Again, add medium to a single spot on the plate to avoid loosening the cells.

12. *First round only:* Add 10 ml complete DMEM/tet. Incubate cells overnight.
- Second round only:* Add 10 ml selection medium/500 μM L-histidinol. Incubate cells overnight.
13. *First round only:* The morning after the transfection, aspirate the medium and replace with 10 ml complete DMEM/tet. Continue incubation.
- Second round only:* The morning after the transfection, aspirate the medium and replace with 10 ml selection medium/500 μM L-histidinol. Continue incubation.

Select and clone transfected cells

14. *First round only:* At 48 hr posttransfection, split cells into selection medium/125 μM L-histidinol at several dilutions ranging from 3×10^4 to 1×10^6 cells per 10-cm plate. Make more than one plate in the mid-range that corresponds to an approximate split from one confluent plate of 1:16 to 1:32.

Second round only: At 48 hr posttransfection, split cells as above, using selection medium/500 μM L-histidinol containing 3 $\mu\text{g/ml}$ puromycin (final).

The optimal killing concentration for puromycin (lowest dose between 0.1 $\mu\text{g/ml}$ to 10 $\mu\text{g/ml}$ that kills all untransfected cells within a few days) should be determined empirically prior to the transfection and varies with the cell type. The concentration of 3 $\mu\text{g/ml}$ puromycin is sufficient for selection of transfected NIH3T3 cells.

15. *First round only:* Refeed cells 4 days later with selection medium/125 μM L-histidinol. When colonies have formed, increase the concentration of L-histidinol in the selection medium to 250 μM .

L-histidinol is normally toxic to cells. The concentration of L-histidinol in the selection medium is therefore kept low initially and is raised as the number of cells expressing pSV₂-His at high levels reaches a critical mass.

Second round only: Refeed cells 4 days later with selection medium/500 μM L-histidinol/puromycin.

16. When colonies are well established (at about day 12 to 14 of selection), circle their borders with a marker. Aspirate medium from plate and place a plastic cloning ring (autoclaved upright in vacuum grease) on the plate to surround an individual clone. Wash clones quickly with ~ 100 μl PBS and add 2 drops of trypsin (~ 100 μl) for 30 sec to 1 min.

Pick cells from plates on which individual colonies are moderately spaced and can easily be distinguished.

17. *First round only:* Loosen cells by pipetting up and down with a Pasteur pipet and transfer colonies to wells of a 24-well plate into 1 ml selection medium/250 μM L-histidinol.

Second round only: Loosen cells as for first round, transferring them into 1 ml selection medium/500 μM L-histidinol/puromycin.

18. *First round only:* When cells are heavy in wells, split into 6-cm dishes in selection medium/500 μM L-histidinol.

Second-round only: When cells are heavy in wells, split into 6-cm dishes in selection medium/500 μM L-histidinol/puromycin.

All trypsinization is performed by standard methods (APPENDIX 3F), involving a quick PBS wash, a 1 to 3 min trypsin/EDTA incubation (2 ml per confluent 10-cm plate), and using 3rd selection medium/500 μM L-histidinol (\pm puromycin) and containing 10% calf serum to dilute and stop the trypsin.

19. *First round only:* Expand cells for testing in selection medium/500 μM L-histidinol. Freeze aliquots of cells for storage in liquid nitrogen and grow in selection medium/500 μM L-histidinol from this point on. Test for tTA or target gene expression (if applicable; see Support Protocol for methods that may be used). Or, if applicable, repeat transfection procedure with target gene plasmid(s), following steps 1 to 18 and using the options listed for second-round transfection.

Second round only: Test for target gene expression by northern or immunoblotting after induction (see Support Protocol). Freeze aliquots for storage in liquid nitrogen and grow in selection medium/500 μM L-histidinol/puromycin from this point on.

ANALYSIS OF TARGET GENE PROTEIN EXPRESSION

This protocol outlines methods for the analysis of target gene expression and inducibility. Instructions for inducing stable cell lines, for examining transient target gene expression with and without induction, and for PCR amplification of the tTA gene are included, with references to detection procedures such as Southern, northern, and immunoblotting techniques.

Induction of Stable Cell Lines

Stable cell lines can be tested for tTA or target gene expression by comparing induced to uninduced cells for tTA mRNA or target gene mRNA (see Detection of tTA Transgene in Cellular or Tail DNA by Southern Blotting), or protein expression or protein activity. Multiple lines may be screened at a time.

The night before induction, the cells are plated in selection medium/500 μM L-histidinol (see recipe) containing 3 $\mu\text{g/ml}$ puromycin at an appropriate density such that cells will be subconfluent to confluent at the time of harvest. Cells are washed three times with PBS (*APPENDIX 2*), with gentle swirling. Immediately, the medium is replaced with selection medium *without* 0.5 $\mu\text{g/ml}$ tetracycline hydrochloride (tet). (For tet⁺ controls, simply aspirate medium and replace with fresh selection medium containing tet.) Cells are incubated 6 to 48 hr in a humidified 37°C, 5% CO₂ incubator, then trypsinized (*APPENDIX 3F*) and harvested at 4°C, and an aliquot of 0.15–0.4 $\times 10^6$ cells is analyzed by immunoblotting (see *UNIT 10.8*).

Alternatively, cells may be grown in selection medium in the presence of tet, transferred to tubes [with a quick wash with cold PBS followed by trypsinization (*APPENDIX 3F*) and stopping of the trypsin by addition of selection medium containing tet], washed three times with PBS (or just pelleted, for tet⁺ controls), and replated into selection medium with and without tet at an appropriate density such that the cells will be subconfluent to confluent at the time of harvest.

Induction of Gene Expression in Transiently Transfected Cells

Transient transfection of tet-regulated plasmids is useful in several situations, including the initial testing of the autoregulatory system in a given cell line, screening stable tTA expressors for inducible expression, and biological applications where transient expression is specifically desired.

The night before the transfection, cells are split into medium containing 0.5 $\mu\text{g/ml}$ tetracycline hydrochloride; the following day they are then transfected by methods appropriate for the cells being used (*UNITS 9.1-9.4*). Cells are induced by washing them three times in medium without tet. For CaPO₄ transfection, washes are incorporated into those normally performed after glycerol shock (see Basic Protocol, step 11). Uninduced cell

controls are washed with medium containing tet. Medium with and without tet is added to the appropriate plates, then the cells are incubated for 12 to 48 hr in a humidified 37°C, 5% CO₂ incubator. The cells are harvested at 4°C and, if trypsinized (*APPENDIX 3F*), cold medium containing 10% FBS (with and without tet, as appropriate) is used to stop the action of the trypsin. Cells are pelleted for freezing or lysis, and tTA or target gene (experimental or reporter) expression can be analyzed by northern blotting (*UNIT 4.9*), immunoblotting (*UNIT 10.8*), or by an appropriate activity assay (see Commentary).

Detection of tTA Transgene in Cellular or Tail DNA by PCR

PCR is routinely used to detect the Tet-tTA transgene in candidate transgenic mouse tail DNA. The forward primer derives from the minimal human CMV promoter, CMV-F1:

5'-TGACCTCCATAGAAGACACC-3'

The reverse primer, TTA-REV1, is specific for the tTA ORF:

5'-ATCTCAATGGCTAAGGCGTC-3'

Hot-start PCR (*UNIT 15.1*) is performed on 150 ng of each tail DNA to be analyzed in a reaction mix containing 1.5 mM MgCl₂, 0.5 μM each primer, and 0.2 mM each dNTP. PCR cycling conditions are as follows:

1 cycle:	3 min	94°C	
		80°C	(pause) add <i>Taq</i> polymerase
30 cycles:	45 sec	94°C	(denaturation)
	45 sec	58°C	(annealing)
	90 sec	72°C	(extension)
1 cycle	10 min	72°C	(extension)
		8°C	(end).

Products are analyzed on a 1% to 1.3% agarose gel; the main product of interest is visible as a 290-bp band after ethidium bromide staining.

Detection of tTA Transgene in Cellular or Tail DNA by Southern Blotting

The tTA transgene may also be detected by Southern blot analysis (*UNIT 2.9*). Tail DNA is digested with *EcoRI* and blots are probed with a 761-bp *XbaI-SalI* tTA insert from pTet-tTA. This fragment detects a 1094-bp tTA fragment of the transgene. This probe may also be used to detect tTA mRNA by northern blotting (*UNIT 4.9*).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

Complete DMEM-10

Dulbecco's minimal essential medium containing:

10% donor bovine calf serum (JRH Biosciences)

100 U/ml penicillin/100 μg/ml streptomycin (Life Technologies)

2 mM glutamine (Life Technologies)

All DMEM complete medium used in this unit (with or without selection reagents or 0.5 μg/ml tetracycline hydrochloride) may be stored protected from light ~1 month at 4°C.

Fetal bovine serum (FBS) may also be used in place of donor bovine calf serum, but the latter is less expensive.

HEPES-buffered saline (HBS)

6 mM dextrose
137 mM NaCl
5 mM KCl
0.7 mM Na₂HPO₄·7H₂O
21 mM HEPES (free acid)
Adjust final pH to 7.05 with NaOH
Filter sterilize and store in aliquots at –20°C

Selection medium

Complete histidine-free DMEM (Irvine Scientific, purchased without glutamine), containing:

10% donor bovine calf serum (JRH Biosciences)
100 U/ml penicillin/100 µg/ml streptomycin (Life Technologies)
2 mM glutamine (Life Technologies)
0.5 µg/ml tetracycline-HCl (Sigma; dilute 10 mg/ml stock in 70% ethanol and store protected from light at –20°C)
125 µM, 250 µM, or 500 µM L-histidinol (Sigma, dilute in water as a 125 mM stock and store at –20°C)

COMMENTARY

Background Information

Inducible, tetracycline-regulated gene expression systems were initially developed to allow the controlled expression in eukaryotic cells of foreign genes not tolerated constitutively in cultured cells or during the development of transgenic animals. The general features of tetracycline-regulated gene expression strategies and their improvements over previous inducible expression systems have been addressed in current review articles (Gossen et al., 1993; Barinaga, 1994; Damke et al., 1995; Shockett and Schatz, 1996). The autoregulatory tTA system used in this protocol derives directly from a constitutive tTA system described by Gossen and Bujard (1992). Although tight regulatory control and high inducibility was achieved with the original system in HeLa cells, the inability to detect clones expressing moderate to high levels of tTA by immunoblotting suggested that the tTA was toxic when expressed constitutively. The autoregulatory tTA system was designed to overcome possible toxic effects of constitutive tTA expression by making tTA expression itself tetracycline regulated. Autoregulated tTA expression theoretically allows for the selection of clones expressing higher levels of tTA via an autoregulatory feed-forward mechanism that is activated only in the absence of tetracycline. In the presence of tetracycline, low-level tTA and target gene expression are driven from the minimal human CMV promoter. However, any tTA produced is

unable to bind to tet operators upstream of the tTA or target gene. Conversely, when tetracycline is removed from the system, the small amounts of tTA protein expressed from the minimal promoter can bind the *tet* operators upstream of the tTA gene, driving higher levels of tTA (for controlled periods of time) and, subsequently, target gene expression.

The theoretical benefits of the autoregulatory tTA system have been confirmed by experiments in stably transfected NIH3T3 cell lines (Shockett et al., 1995). In these experiments, expression of the recombination activating genes *RAG-1* and *RAG-2*, and subsequent DNA recombination activated by these proteins, was higher and more frequently detected among stable transfectants expressing autoregulatory tTA than in constitutive tTA expressors. In transgenic mice expressing a luciferase reporter target transgene, the levels of expression appear to be 1 to 2 orders of magnitude greater with the autoregulatory system, although the uninduced levels also appear to be higher.

Since the description of the early tTA systems, several laboratories have created modified vectors, including streamlined versions containing both tTA and the target gene, viral vectors, and vectors in which expression of two different target genes may be differentially or co-regulated. Some of these systems and their applications have recently been reviewed (Shockett and Schatz, 1996).

Critical Parameters and Troubleshooting

Cell lines stably expressing both autoregulatory tTA and target genes have been derived at fairly high efficiencies by simultaneous transfection of all plasmids. This method may be faster, but it may require the screening of more clones than if stable lines with low basal and high induced levels of tTA are first derived and subsequently transfected with plasmids encoding the target genes. For the derivation of these clones, any selectable marker combination should theoretically work for consecutive cotransfection. Additionally, although the Basic Protocol describes calcium phosphate-mediated transfection of adherent fibroblast cell lines, the procedure can be adapted for other cell types using their optimal methods of transfection and selection. The protocol can also be scaled down to require fewer cells by using smaller dishes or wells and reducing all components proportionately.

Using the autoregulatory tTA system, tTA mRNA induction appears to be a good indicator of induced tTA expression (see Support Protocol). Alternatively, the vector pUHC13-3 (Life Technologies) encoding luciferase under tet control may be transiently transfected into putative stable tTA expressors as previously described (see Support Protocol and Damke et al., 1995). Cells are then cultured for 12 to 48 hr in the presence and absence of tetracycline. Luciferase activity is easily measured in cell lysates using a kit (Luciferase Assay System and Dual-Luciferase Reporter Assay System; Promega) in which luciferase activity in cell lysates is normalized either to total protein determined using a Bradford protein assay (*UNIT 10.1*), or to a transfection control, respectively. Although basal expression of target plasmids tends to be higher when transiently transfected and luciferase detection is extremely sensitive, this method can be useful for the initial testing of the system in a given cell type (Damke et al., 1995).

It is imperative after stable transfection with pTet-tTA that cells be maintained in medium containing 0.5 µg/ml tetracycline to prevent any toxic effects of tTA expression and subsequent selection against clones expressing high levels of tTA.

Anticipated Results

In the authors' experience with stably transfected NIH3T3 cells, expression of induced

tTA and target gene has been observed by 6 hr and peaks at ~12 hr after induction. In cells that stably express tTA, transient target gene expression has been observed by 12 hr. In cells transiently expressing tTA and a tet-sensitive luciferase reporter (pUHC13-3), luciferase activity induced by 2 orders of magnitude has been observed by 20 hr.

Time Considerations

Starting with the plasmid vectors and following the transfection protocols above, stable clones expressing tTA (or tTA + target gene(s) if cotransfecting) are obtained in ~12 to 14 days. Approximately 2 additional weeks are required for expansion and testing of candidate clones. Subsequent transfection of a stable inducible tTA clone with vectors expressing target genes will require the same amount of time. Transient transfection and inducible gene expression may be achieved within 48 hr.

Literature Cited

- Barinaga, M. 1994. Researchers devise a master gene control switch. *Science* 265:26-28.
- Damke, H., Gossen, M., Freundlieb, S., Bujard, H., and Schmid, S.L. 1995. Tightly regulated and inducible expression of dominant interfering dynamin mutant in stably transformed HeLa cells. *Methods Enzymol.* 257:209-220.
- Gossen, M. and Bujard, H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U.S.A.* 89:5547-5551.
- Gossen, M., Bonin, A.L., and Bujard, H. 1993. Control of gene activity in higher eukaryotic cells by prokaryotic regulatory elements. *Trends Biochem. Sci.* 18:471-475.
- Shockett, P.E. and Schatz, D.G. 1996. Commentary: Diverse strategies for tetracycline-regulated inducible gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 93:5173-5176.
- Shockett, P., Difilippantonio, M., Hellman, N., and Schatz, D. 1995. A modified tetracycline-regulated system provides autoregulatory, inducible gene expression in cultured cells and transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* 92:6522-6526.

Contributed by Penny Shockett
Yale University School of Medicine
New Haven, Connecticut

David Schatz
Howard Hughes Medical Institute and
Yale University School of Medicine
New Haven, Connecticut

Overview of the Vaccinia Virus Expression System

Vaccinia virus was introduced in 1982 as a vector for transient expression of genes in mammalian cells (Mackett et al., 1982; Panicali and Paoletti, 1982). This expression system differs from others in that transcription occurs in the cytoplasm of the cell rather than in the nucleus. As a vector, vaccinia virus has a number of useful characteristics, including a capacity that permits cloning large fragments of foreign DNA (>20 kbp) with retention of infectivity, a wide host range, a relatively high level of protein synthesis, and “appropriate” transport, secretion, processing, and posttranslational modifications as dictated by the primary structure of the expressed protein and the cell type used. For example, *N*- and *O*-glycosylation, phosphorylation, myristylation, and cleavage, as well as assembly of expressed proteins, occur in an apparently faithful manner.

Laboratory applications of vaccinia virus vectors include production of biologically active proteins in tissue culture, analysis of mutant forms of proteins, and determination of transport and processing signals. In addition, recombinant vaccinia viruses have been important for immunological studies (Benink and Yewdell, 1990). Infected cells can serve as targets to analyze the antigenic specificity of cytotoxic T cells. Recombinant viruses can be used to infect animals in order to determine cell-mediated and humoral responses to specific proteins, and are being evaluated as candidate vaccines for human and veterinary uses.

Several variations of the vaccinia vector system have been developed and either standard or more attenuated, host-restricted strains may be used. After obtaining the virus stock (UNIT 16.16), the gene of interest is placed under control of a vaccinia virus promoter and integrated into the genome of vaccinia so as to retain infectivity (UNIT 16.17). Alternatively, expression can be achieved by transfecting a plasmid containing the vaccinia promoter-controlled gene into a cell that has been infected with vaccinia virus. These recombinant viruses are then characterized using various methods (UNIT 16.18). In still another variation, the bacteriophage T7 RNA polymerase gene can be integrated into the genome

of vaccinia so that a gene controlled by a T7 promoter, either in a transfected plasmid or a recombinant vaccinia virus, will be expressed.

VACCINIA REPLICATION CYCLE

Vaccinia is the prototypic member of the *Orthopoxvirus* genus of the *Poxviridae* family. Poxviruses differ from other eukaryotic DNA viruses in that they replicate in the cytoplasm rather than in the nucleus. Vaccinia virus has a linear, double-stranded DNA genome of nearly 200,000 bp that encodes most of the proteins needed for replication and transcription in the cytoplasm.

The replication cycle of poxviruses is represented in Figure 16.15.1. The virus particle or virion consists of a complex core structure surrounded by a lipoprotein envelope. Remarkably, all proteins necessary for transcription of the early class of genes are packaged with the genome in the core. These include the following virus-encoded proteins: a multisubunit, DNA-dependent RNA polymerase, an early transcription factor, capping and methylating enzymes, and a poly(A) polymerase. The transcription system is activated upon infection, and early mRNAs and proteins can be detected within the first hour. The early mRNAs closely resemble their eukaryotic counterparts—they are capped, methylated, polyadenylated, and of discrete size. Termination of transcription occurs ~50 bases after the sequence TTTTNT (where N can be any nucleotide; the termination signal is actually recognized in the nascent RNA as UUUUUNU). There is no evidence for splicing or other kinds of processing involving RNA cleavage.

DNA replication begins within a few hours postinfection and leads successively to the intermediate and late phases of gene expression. The promoters associated with intermediate and late genes differ in sequence from early promoters and have different transcription factor requirements. The intermediate factors are synthesized early in order to facilitate transcription of newly replicated viral DNA. Some of the late factors are products of intermediate genes and hence are made and used after viral DNA replication.

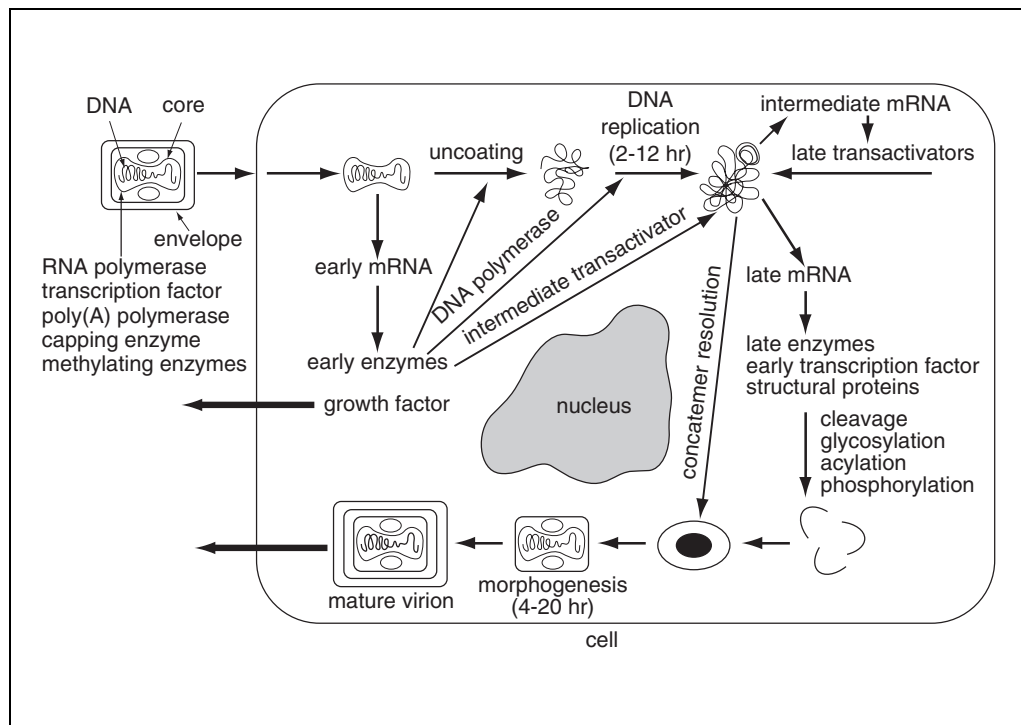


Figure 16.15.1 Replication cycle of vaccinia virus. After entry of vaccinia virus into the cells, early genes are expressed, leading to secretion of several proteins (including a growth factor), uncoating of the virus core, and the synthesis of DNA polymerase (and other replication proteins), RNA polymerase subunits, and transcriptional transactivators of the intermediate class of genes. After DNA replication, intermediate mRNAs are made, some of which encode late transcriptional transactivators. This in turn leads to expression of late genes, which encode structural proteins, enzymes, and early transcription factors that are packaged in the assembling virus particles. Some of the mature virions are wrapped in Golgi-derived membranes and are released from the cell. The bold arrows indicate products that exit the cell. Reprinted with permission from Raven Press (Moss, 1996a).

The mRNAs transcribed from intermediate and late genes differ from typical early mRNAs as follows. First, the termination signal UUUUUNU is not recognized; consequently the mRNAs are long and heterogeneous in length, making procedures like northern blotting virtually useless. Second, the 5' ends of intermediate and late mRNAs contain a capped poly(A) leader of ~35 nucleotides that is probably the result of an RNA polymerase slippage mechanism within the conserved AAA at the initiation site.

Many early proteins are not synthesized beyond ~6 hr postinfection (unless an inhibitor of DNA replication such as cytosine arabinoside is added) because of the cessation of early gene transcription and the relatively short half-life of all mRNAs at late times. Some genes, however, have tandem early and late promoters so that they are expressed throughout the growth cycle. Intermediate

genes are expressed for a relatively short period after DNA replication. Either because of intrinsic promoter strength, DNA copy number, or prolonged expression (>20 hr), much more protein is made from the strongest late promoters than from early or intermediate promoters.

EFFECTS OF VACCINIA INFECTION

Standard vaccinia virus strains can productively infect most mammalian and avian cell lines, with a few exceptions such as Chinese hamster ovary (CHO) cells, but may not complete the replication cycle in primary lymphocytes or macrophages. The host-restricted modified vaccinia virus Ankara (MVA) strain has been shown to replicate efficiently only in chick embryo fibroblasts and BHK-21 cells, but viral and recombinant protein expression occurs in all cell lines infectable by standard strains. Vaccinia infection generally results in

rapid inhibition of host nucleic acid and protein synthesis. Inhibition of host protein synthesis is dramatic and probably results from several factors whose identities are uncertain; the relative contribution of each factor may depend on the virus multiplicity, cell type, and time of analysis. At the time of maximal late gene expression, host protein synthesis has been largely suppressed, facilitating the identification of viral or recombinant proteins by pulse-labeling with radioactive amino acids (*UNIT 10.18*).

In fibroblasts, the initial cytopathic effect is cell rounding, and is obvious by several hours postinfection. Nevertheless, the majority of cells remain intact for ≥ 48 hr. Some virus strains may cause less cytopathic effects than others, but mutants that efficiently express recombinant proteins but do not inhibit host gene expression have not been identified. Approximately 100 to 200 plaque-forming units (pfu), equivalent to ~ 2500 to 5000 particles, are made per cell within a 20- to 40-hr period. With the commonly used vaccinia virus Western Reserve (WR) strain, $>95\%$ of the infectious virus remains cell-associated. With some other vaccinia virus strains, notably IHD-J, larger amounts of extracellular virus are produced.

VACCINIA VECTOR EXPRESSION SYSTEM

Genes or cDNAs containing open reading frames derived from prokaryotic, eukaryotic, or viral sources have been expressed using vaccinia virus vectors. The gene of interest is usually placed next to a vaccinia promoter and this expression cassette is then inserted into the virus genome by homologous recombination or direct ligation (*UNIT 16.17*). Use of poxvirus promoters is essential because cellular and other viral promoters are not recognized by the vaccinia transcriptional apparatus. Strong late promoters are preferable when high levels of expression are desired. An early promoter, however, may be of use if it is desirable to express proteins prior to the occurrence of major cytopathic effects, or when the purpose is to make cells that express antigens in association with major histocompatibility class I molecules so they form cytotoxic T cell targets or prime animals for a cytotoxic T cell response. (The ability of vaccinia viral vectors to direct this type of antigen presentation seems to diminish late in infection.) The most versatile and widely used promoters contain early and late promoter elements. Transcripts originating early will terminate after the sequence

TTTTTNT; thus, any cryptic TTTTTNT termination motifs within the coding sequence of the gene should be altered by mutagenesis if an early poxvirus promoter is used (Earl et al., 1990). To mimic vaccinia virus mRNAs, untranslated leader and 3'-terminal sequences are usually kept short.

A number of plasmids have been designed with restriction endonuclease sites for insertion of foreign genes downstream of vaccinia promoters (*UNIT 16.17*). The expression cassette is flanked by vaccinia DNA to permit homologous recombination when the plasmid is transfected into cells that have previously been infected with a vaccinia virus. The flanking vaccinia virus DNA is chosen so that recombination will not interrupt an essential viral gene.

Without selection, the ratio of recombinant to parental vaccinia virus is usually $\sim 1:1000$. Although this frequency is high enough to permit the use of plaque hybridization (*UNITS 6.3 & 6.4*) or immunoscreening (*UNIT 6.7*) to pick recombinant viruses, a variety of methods have been employed to facilitate identification of recombinant viruses. Some widely used selection or screening techniques are described in *UNIT 16.17*. Commonly, the expression cassette is flanked by segments of the vaccinia thymidine kinase (TK) gene so that recombination results in inactivation of TK. Virus with a TK⁻ phenotype can then be distinguished from those with a TK⁺ phenotype by infecting a TK⁻ cell line in the presence of 5-bromodeoxyuridine (BrdU), which must be phosphorylated by TK to be lethally incorporated into the virus genome. Alternatively, recombinant viruses can be selected by the co-expression of a bacterial antibiotic resistance gene such as guanine phosphoribosyltransferase (*gpt*). Co-expression of the *Escherichia coli lacZ* gene or other color markers allows rapid screening of recombinant virus plaques. Complementation of host-range and plaque-forming defects are also useful for isolation of recombinant viruses (*UNIT 16.17*).

STEPS FOR EXPRESSION OF GENES USING VACCINIA VECTORS

The expression of genes using the vaccinia expression system is presented in detail in *UNITS 16.16-16.18* and is outlined in the flowchart in Figure 16.15.2. A brief overview is presented below.

1. Prepare a stock of standard, host-range-restricted, or bacteriophage T7 RNA polym-

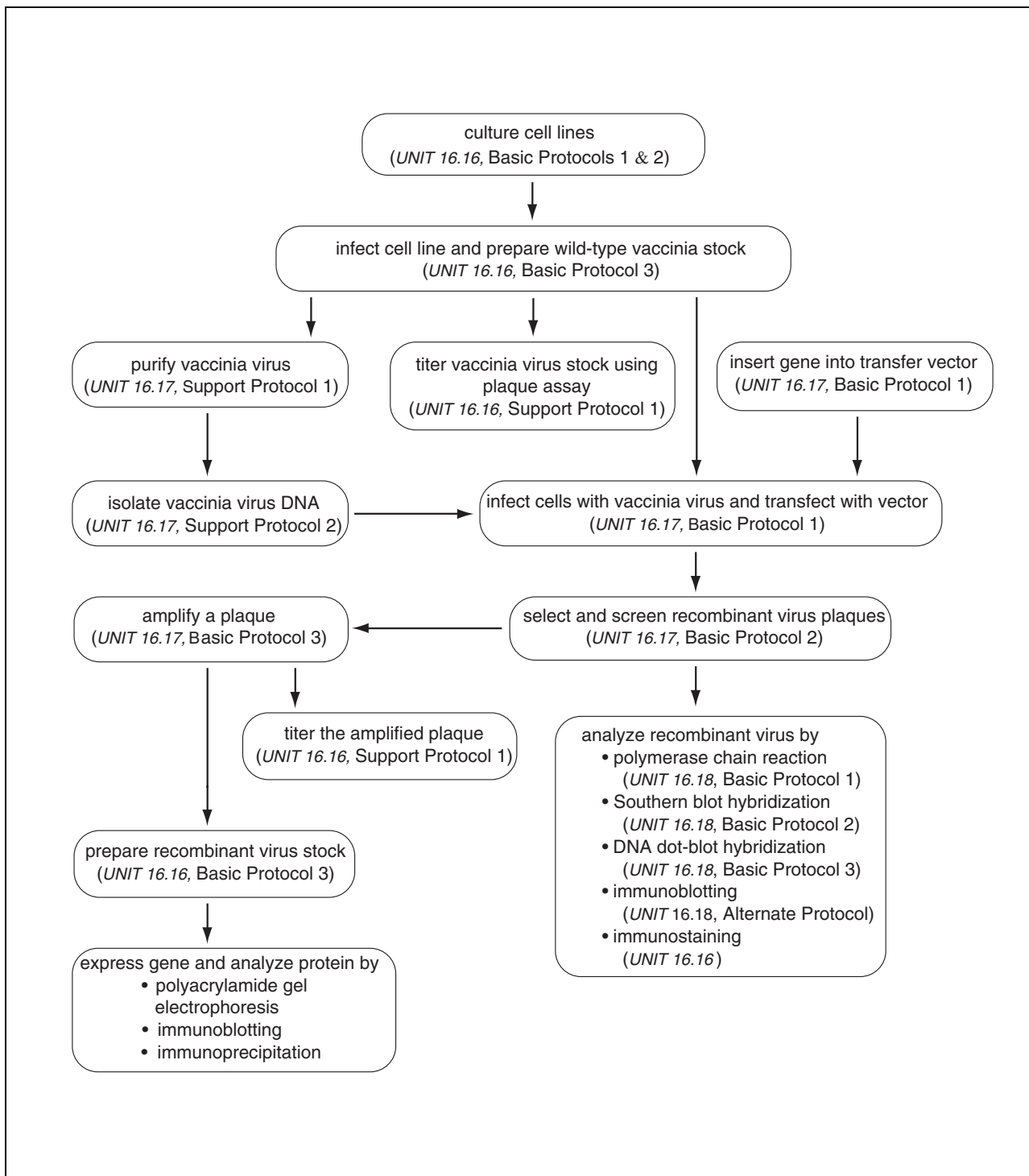


Figure 16.15.2 Flowchart showing protocols for gene expression using the recombinant vaccinia virus system.

erase-expressing vaccinia virus. At the same time, subclone the gene of interest into a plasmid transfer vector (UNITS 16.16 & 16.17).

2. Infect cells with vaccinia virus and transfect with the recombinant plasmid (UNIT 16.17).

3. Lyse the cells and plaque the virus under suitable selection or screening conditions (UNIT 16.17).

4. Pick plaques and confirm the presence and/or expression of the foreign gene (UNIT 16.18).

5. Amplify the plaque and prepare recombinant virus stock (UNIT 16.17).

6. Infect cells and analyze proteins synthesized (UNIT 16.18).

SAFETY PRECAUTIONS FOR USING VACCINIA

Vaccinia virus is not to be confused either with variola virus, another member of the *Orthopoxvirus* genus that caused smallpox prior to its eradication, or with varicella virus, a herpes virus that causes chicken pox. Until 1972, vaccinia virus was routinely used in the United States as a live vaccine to prevent smallpox, and a residual scar, commonly on the upper arm, is evidence of that vaccination.

To prevent laboratory infections, the Centers for Disease Control (CDC) and the National Institutes of Health (NIH) recommend that individuals who come into contact with vaccinia virus receive vaccinations at 10-year intervals (Richmond and McKinney, 1993). The CDC has supplied vaccine for such purposes when requested by qualified health workers. Eczema or an immunodeficiency disorder in the laboratory worker or a close contact, however, may be a contraindication to vaccination, which should only be given under medical supervision. The benefits of routine vaccination for healthy investigators have also been questioned (Baxby, 1989; Wenzel and Nettelmann, 1989). Vaccinia virus is very stable and parenteral inoculation, ingestion, and droplet or aerosol exposure of mucous membranes are the primary hazards to laboratory or animal care personnel. Standard biosafety level 2 (BL-2) practices and class I or II biological safety cabinets should be employed (Richmond and McKinney, 1993). The NIH intramural program has lowered the safety requirements for highly attenuated vaccinia virus strains such as modified vaccinia virus Ankara (MVA; UNIT 16.6) and NYVAC (Tartaglia et al., 1992). However, local institutional biosafety offices should be contacted to determine current policy regarding vaccination and physical containment.

Additional precautions may be necessary for expression of certain genes such as toxins or large segments of other viral genomes, and guidelines for recombinant DNA work should be consulted. Approval of local biosafety committees may be necessary.

LITERATURE CITED

- Baxby, D. 1989. Smallpox vaccination for investigators. *Lancet* 2:919.
- Bennink, J.R. and Yewdell, J.W. 1990. Recombinant vaccinia viruses as vectors for studying T lymphocyte specificity and function. *Curr. Topics Microbiol. Immunol.* 163:153-184.
- Earl, P.L., Hügin, A.W., and Moss, B. 1990. Removal of cryptic poxvirus transcription termination signals from the human immunodeficiency virus type 1 envelope gene enhances expression and immunogenicity of a recombinant vaccinia virus. *J. Virol.* 64:2448-2451.
- Mackett, M., Smith, G.L., and Moss, B. 1982. Vaccinia virus: A selectable eukaryotic cloning and expression vector. *Proc. Natl. Acad. Sci. U.S.A.* 79:7415-7419.
- Moss, B. 1996a. *Poxviridae: The viruses and their replication.* In *Virology* (B.N. Fields, D.M. Knipe, and P. M. Howley, eds.) pp. 2637-2671. Raven Press, New York.
- Panicali, D. and Paoletti, E. 1982. Construction of poxviruses as cloning vectors: Insertion of the thymidine kinase gene from herpes simplex virus into the DNA of infectious vaccinia virus. *Proc. Natl. Acad. Sci. U.S.A.* 79:4927-4931.
- Richmond, J.Y. and McKinney, R.W. 1993. *Biosafety in Microbiological and Biomedical Laboratories*, 3rd ed. U.S. Government Printing Office, Washington, D.C.
- Tartaglia, J., Perkus, M.E., Taylor, J., Norton, E.K., Audonnet, J.C., Cox, W.I., Davis, S.W., Vanderhoeven, J., Meignier, B., Riviere, M., Languet, B., and Paoletti, E. 1992. NYVAC-A highly attenuated strain of vaccinia virus. *Virology* 188:217-232.
- Wenzel, R.P. and Nettelmann, M.D. 1989. Smallpox vaccination for investigators using vaccinia recombinants. *Lancet* 2:630-631.

KEY REFERENCES

- Johnson, G.P., Goebel, S.J., and Paoletti, E. 1993. An update on the vaccinia virus genome. *Virology* 196: 381-401.
- Moss, B. 1996b. Genetically engineered poxviruses for recombinant gene expression, vaccination, and safety. *Proc. Natl. Acad. Sci. U.S.A.* 93:11341-11348.

Contributed by Bernard Moss and
Patricia L. Earl
National Institute of Allergy and
Infectious Diseases
Bethesda, Maryland

Protein
Expression

16.15.5

Preparation of Cell Cultures and Vaccinia Virus Stocks

This unit describes the maintenance of cell lines used with vaccinia virus, both in monolayer cultures (see Basic Protocol 1) and in suspension (see Basic Protocol 2). The suspended cell culture is then used in the preparation of vaccinia virus stocks (see Basic Protocol 3). The preparation of chick embryo fibroblasts (CEF) is also presented (see Basic Protocol 4), for use in the production of the highly attenuated and host range-restricted modified vaccinia virus Ankara (MVA) strain of vaccinia virus (see Basic Protocol 5). Additionally, support protocols are presented for the titration of standard and MVA vaccinia virus stocks (see Support Protocols 1 and 2, respectively).

Because standard vaccinia virus strains have a broad host range, there is considerable latitude in the selection of cell lines; those described below (see Basic Protocols 1 and 2) have been found to give good results. BS-C-1 cells give the best results for a plaque assay, whereas HeLa cells are preferred for preparation of virus stocks. CV-1 cells can be used for both procedures, but they are generally used for transfection (UNIT 16.17). Human thymidine kinase-negative (TK⁻) 143B cells are used when TK selection is employed (UNIT 16.17), but they can be used for transfection as well as for a plaque assay. Either CEF or BHK-21 cells are used to propagate MVA and prepare recombinant MVA. Table 16.16.1 presents a summary of the uses for specific cell lines.

NOTE: Carry out all procedures in this unit using sterile technique, preferably in a biosafety cabinet.

Table 16.16.1 Cell Lines Used in Specific Vaccinia Protocols

Cell line	Use ^a	Procedure
HeLa S3	Virus stock preparation	UNIT 16.16 Basic Protocol 3
	Virus purification	UNIT 16.17 Support Protocol 1
	Plaque amplification	UNIT 16.17 Basic Protocol 3
BS-C-1	Plaque assay	UNIT 16.16 Support Protocol 1
	Transfection (optional)	UNIT 16.17 Basic Protocol 1
	XGPRT selection	UNIT 16.17 Basic Protocol 2
	Plaque amplification (optional)	UNIT 16.17 Basic Protocol 3
CV-1	Transfection	UNIT 16.17 Basic Protocol 1
	Virus stock preparation (optional)	UNIT 16.16 Basic Protocol 3
	Plaque assay (optional)	UNIT 16.16 Support Protocol 1
HuTK ⁻ 143B	TK selection	UNIT 16.17 Basic Protocol 2
	Plaque assay (optional)	UNIT 16.16 Support Protocol 1
	Transfection (optional)	UNIT 16.17 Basic Protocol 1
CEF	MVA procedures	UNIT 16.16 Basic Protocol 5, Support Protocol 2; UNIT 16.17 Basic Protocol 4
BHK-21	MVA procedures	UNIT 16.16 Basic Protocol 5, Support Protocol 2; UNIT 16.17 Basic Protocol 4

^aThe preferred use(s) for each cell line is listed first; if optional is indicated, the cell line can be used for the indicated procedure, but the results may not be as good as those from the preferred cell line.

CULTURE OF MONOLAYER CELLS

Frozen cells are thawed and grown in appropriate complete medium containing twice the maintenance amount of serum (see below). When the cells are confluent, they are treated with trypsin/EDTA, diluted, and maintained in appropriate complete medium containing 10% FBS (see Table 16.16.2).

Table 16.16.2 Media Used for Growth and Maintenance of Cell Lines^a

Cell line	Maintenance medium	Start-up medium ^a
BHK-21	Complete MEM-10	Complete MEM-20
BS-C-1	Complete MEM-10	Complete MEM-20
CEF	Complete MEM-10	Complete MEM-10
CV-1	Complete DMEM-10	Complete DMEM-20
HeLa S3	Complete spinner medium-5	Complete MEM-10
HuTK ⁻ 143B	Complete MEM-10/BrdU	Complete MEM-20/BrdU

^aSee Reagents and Solutions for recipes.

Materials

Frozen ampule of cells (Table 16.16.2): BS-C-1 (ATCC no. CCL26), CV-1 (ATCC no. CCL70), HuTK⁻ 143B (ATCC no. CRL8303), or BHK-21 (ATCC no. CCL10) cells

70% ethanol

Start-up medium (Table 16.16.2): complete MEM-20, complete DMEM-20, or complete MEM-20/BrdU (see recipes), 37°C

Maintenance medium (Table 16.16.2): complete MEM-10, complete DMEM-10, or complete MEM-10/BrdU (see recipes), 37°C

PBS (optional; APPENDIX 2)

Trypsin/EDTA: 0.25% (w/v) trypsin/0.02% (w/v) EDTA, 37°C

25-cm² and 150-cm² tissue culture flasks

Humidified, 37°C, 5% CO₂ incubator

Begin the culture

1. Thaw a frozen ampule of cells in a 37°C water bath.
2. Sterilize the ampule tip with 70% ethanol, break the neck, and transfer the cells with a pipet into a 25-cm² tissue culture flask containing 5 ml start-up medium. Rotate the flask to evenly distribute the cells and place overnight in a humidified, 5% CO₂ incubator at 37°C.
3. Aspirate the start-up medium and replace with appropriate maintenance medium. Return cells to the CO₂ incubator at 37°C and check daily for confluency.

Cells should be passaged when they become confluent. Generally, if cells are split 1:20, they reach confluence in 1 week and need not be counted.

Maintain the culture

4. When the cells are a confluent monolayer, aspirate medium.
5. Wash cells once with PBS or trypsin/EDTA to remove remaining serum from the cells by covering cells with the solution and pipetting it off.

6. Overlay cells with 37°C trypsin/EDTA using a volume that is just enough to cover the monolayer (e.g., 0.3 ml for a 25-cm² flask). Allow to sit 30 to 40 sec (cells should become detached) and shake the flask to completely detach cells.
7. Add 1.4 ml appropriate maintenance medium. Pipet the cell suspension up and down several times to disrupt clumps (these cells are ready for passage).
8. Remove 0.5 ml cell suspension and add it to a new 150-cm² tissue culture flask containing 30 ml maintenance medium. Rotate the flask to evenly distribute the cells and place in a CO₂ incubator at 37°C until the cells are confluent (~1 week). Maintain the cells by splitting ~1:20 in maintenance medium at approximately weekly intervals.

Cells can be maintained in smaller flasks if desired. If so, volumes should be adjusted proportionately.

CULTURE OF CELLS IN SUSPENSION

HeLa S3 cells are maintained in complete spinner medium-5.

Materials

Frozen ampule of HeLa S3 cells (ATCC no. CCL2.2)
 70% ethanol
 Complete MEM-10 (see recipe), 37°C
 Trypsin/EDTA: 0.25% (w/v) trypsin/0.02% (w/v) EDTA, 37°C
 Complete spinner medium-5 (see recipe), 37°C
 25-cm² tissue culture flask
 Humidified, 37°C, 5% CO₂ incubator
 Sorvall H-6000A rotor (or equivalent)
 100- or 200-ml vented spinner bottles and caps with filters (Bellco)
 Additional reagents and equipment for counting cells with a hemacytometer
 (APPENDIX 3F)

Begin the culture

1. Thaw a frozen ampule of HeLa S3 cells in a 37°C water bath.
2. Sterilize the ampule tip with 70% ethanol, break the neck, and transfer the cells with a pipet to a 25-cm² tissue culture flask containing 5 ml complete MEM-10. Rotate the flask to evenly distribute the cells and place overnight in a humidified, 5% CO₂ incubator at 37°C.
3. Aspirate medium. Overlay cells with 0.5 ml of 37°C trypsin/EDTA and let sit 30 to 40 sec.

Since these cells do not attach firmly to the flask, they should not be washed prior to trypsinization.
4. Add 10 ml complete spinner medium-5 and transfer cells to a 50-ml centrifuge tube. Centrifuge 5 min in a Sorvall H-6000A rotor at 2500 rpm (1800 × g), room temperature, and discard supernatant.
5. Suspend cell pellet in 5 ml complete spinner medium-5 by pipetting up and down to disrupt clumps.
6. Add 50 ml complete spinner medium-5 to a 100- or 200-ml vented spinner bottle and transfer the cell suspension to this bottle.

**BASIC
 PROTOCOL 2**

**Protein
 Expression**

16.16.3

7. Remove 1 ml cell suspension and count the cells using a hemacytometer (*APPENDIX 3F*). Add complete spinner medium-5 to adjust the cell density to $3\text{--}4 \times 10^5$ cells/ml. Place cells in a 37°C incubator without CO_2 and stir continuously.

The initial high density is used because some cells are not viable.

8. Grow cells for two successive days, counting cells daily and adding complete spinner medium-5 as necessary to maintain a concentration of $3\text{--}4 \times 10^5$ cells/ml.

Maintain the culture

9. Remove 1 ml cell suspension and monitor the cells using a hemacytometer.
10. When the density is $4\text{--}5 \times 10^5$ cells/ml, dilute the cells to 1.5 or 2.5×10^5 cells/ml with complete spinner medium-5 for alternate day or daily feeding, respectively.
11. Place a 100- or 200-ml vented spinner bottle containing 50 ml or 100 ml cells, respectively, in 37°C incubator without CO_2 and stir continuously. Passage every 1 to 2 days.

HeLa S3 cells are grown and maintained in complete spinner medium-5 in vented spinner bottles at 37°C without CO_2 . Cells are diluted with fresh medium at 1- to 2-day intervals to keep the cell density between 1.5×10^5 and 5×10^5 cells/ml.

BASIC PROTOCOL 3

PREPARATION OF A VACCINIA VIRUS STOCK

To prepare a vaccinia virus stock, HeLa S3 cells from a spinner culture (see Basic Protocol 2) are plated the day before infection and allowed to attach. They are then infected with trypsinized virus. After several days, the infected cells are harvested and lysed during repeated freeze-thaw cycles. The virus stock is then aliquoted and stored at -70°C . To titer this stock, see Support Protocol 1. The protocol can be modified for monolayer cultures.

Materials

- HeLa S3 cells from suspension culture (see Basic Protocol 2)
- Complete MEM-10 and -2.5 (see recipe), 37°C
- Vaccinia virus (ATCC no. VR1354 or equivalent)
- 0.25 mg/ml trypsin (2 \times crystallized and salt-free; Worthington; filter sterilize and store at -20°C)
- Sorvall H-6000A rotor (or equivalent)
- 150-cm² tissue culture flask
- Humidified, 37°C , 5% CO_2 incubator
- Additional reagents and equipment for counting cells with a hemacytometer (*APPENDIX 3F*)

Prepare cells

1. Count HeLa S3 cells from a suspension culture using a hemacytometer (*APPENDIX 3F*).
2. Centrifuge 5×10^7 cells 5 min in a Sorvall H-6000A rotor at 2500 rpm (1800 \times g), room temperature, and discard supernatant.
3. Resuspend cells in 25 ml of 37°C complete MEM-10, dispense in a 150-cm² tissue culture flask, and place overnight in a humidified, 5% CO_2 incubator at 37°C .

Increase the number of HeLa cells proportionately if more than one 150-cm² flask is to be infected.

As an alternative to HeLa suspension cells, prepare 150-cm² flasks of monolayer cells as described (see Basic Protocol 1) and continue with step 4 below.

Trypsinize virus

4. Just prior to use, mix an equal volume of vaccinia virus stock and 0.25 mg/ml trypsin, and vortex vigorously. Incubate 30 min in a 37°C water bath, vortexing at 5- to 10-min intervals throughout the incubation.

Virus stocks are usually at a titer of $\sim 2 \times 10^9$ pfu/ml, but may be significantly lower depending on the source.

Vortexing usually breaks up any clumps of cells. However, if there are still visible clumps, chill to 0°C and sonicate 30 sec on ice (UNIT 16.17). Sonication can be repeated several times but the sample should be allowed to cool on ice between sonications.

Infect cells

5. Dilute trypsinized virus in complete MEM-2.5 at $2.5\text{--}7.5 \times 10^7$ pfu/ml. Decant or aspirate medium from the 150-cm² flask of cells and add 2 ml diluted, trypsinized virus. Place 2 hr in a CO₂ incubator at 37°C, rocking flask by hand at 30-min intervals.

The optimal multiplicity of infection (MOI) is 1 to 3 pfu/cell. Multiplicities of 0.1 pfu/cell may be necessary if the titer of the initial virus stock is low. The trypsinized virus must be diluted ≥ 10 -fold to avoid detaching the cells.

6. Overlay cells with 25 ml complete MEM-2.5 and place 3 days in a CO₂ incubator at 37°C.
7. Detach the infected cells from the flask by shaking and pour or pipet into a sterile plastic screw-cap tube. Centrifuge 5 min at $1800 \times g$, 5° to 10°C, and discard supernatant.
8. Resuspend cells in 2 ml complete MEM-2.5 (per initial 150-cm² flask) by gently pipetting or vortexing.

Harvest virus stock

9. Lyse the cell suspension by freeze-thaw cycling as follows: freeze in dry ice/ethanol, thaw in a 37°C water bath, and vortex. Carry out the freeze-thaw cycling a total of three times.
10. Keep the virus stock on ice and divide it into 0.5- to 2-ml aliquots. Store the aliquots indefinitely at -70°C.

TITRATION OF VACCINIA VIRUS STOCKS BY PLAQUE ASSAY

Serial dilutions of the trypsinized virus stock (see Basic Protocol 3) are used to infect the appropriate cell line. After several days growth, the medium is removed and the cells are stained with crystal violet. Plaques appear as 1- to 2-mm-diameter areas of diminished staining due to the retraction, rounding, and detachment of infected cells.

Additional Materials (also see *Basic Protocols 1 and 3*)

- BS-C-1 cells from confluent monolayer culture (see Basic Protocol 1)
- Virus stock (see Basic Protocol 3)
- 0.1% (w/v) crystal violet (Sigma) in 20% ethanol (store indefinitely at room temperature)
- 6-well, 35-mm² tissue culture dishes
- Additional reagents and equipment for counting cells with a hemacytometer (*APPENDIX 3F*)

SUPPORT PROTOCOL 1

Protein Expression

16.16.5

Prepare cell and virus stocks

1. Trypsinize confluent monolayer of BS-C-1 cells as described (see Basic Protocol 1, steps 4 to 7).
2. Count the cells using a hemacytometer (*APPENDIX 3F*).
3. Seed wells of a 6-well, 35-mm² tissue culture dish with 5×10^5 cells per well BS-C-1 cells in 2 ml complete MEM-10. Place overnight in a humidified, 5% CO₂ incubator at 37°C to reach confluency.
4. Trypsinize virus stock (see Basic Protocol 3, step 4).
5. Make nine 10-fold serial dilutions (*UNIT 1.11*) of the trypsinized virus in complete MEM-2.5, using a fresh pipet for each dilution.

Perform assay

6. Remove medium from BS-C-1 cells and infect cells in duplicate wells with 0.5 ml of the 10⁻⁷, 10⁻⁸, and 10⁻⁹ trypsinized virus dilutions. Place 1 to 2 hr in a CO₂ incubator at 37°C, rocking dish at 15- to 30-min intervals to spread virus uniformly and keep cells moist.
7. Overlay cells in each well with 2 ml complete MEM-2.5 and place 2 days in a CO₂ incubator at 37°C.
8. Remove medium and add 0.5 ml of 0.1% crystal violet to each well. Incubate 5 min at room temperature.
9. Aspirate crystal violet and allow wells to dry.
10. Determine the titer by counting plaques within the wells and multiplying by the dilution factor.

Most accurate results are obtained from wells with 20 to 80 plaques. In determining the titer, take into account the 1:1 dilution of the virus stock with trypsin.

BASIC PROTOCOL 4

PREPARATION OF CHICKEN EMBRYO FIBROBLASTS

The attenuated, replication-deficient modified vaccinia virus Ankara (MVA) can be used as an alternative to standard strains of vaccinia virus, for use as a vector to express foreign genes. MVA was derived from the vaccinia virus strain Ankara by multiple (>570) passages in chick embryo fibroblasts (CEF). Although MVA expresses recombinant proteins efficiently, the assembly of infectious particles is interrupted in human and most other mammalian cells lines, providing an added degree of safety to laboratory personnel. The NIH Intramural Biosafety Committee has determined that individuals do not need to be vaccinated to work with MVA, and can do so under biosafety level 1 (BL-1) conditions if no other vaccinia virus strains are being manipulated at the same location by other investigators. Recombinant MVA expressing bacteriophage T7 RNA polymerase, which can be used for high level transient expression, has also been constructed (*UNIT 16.19*)

CEF or Syrian hamster kidney cells (BHK-21, see Basic Protocol 1) are permissive for growth of MVA. MVA is titered by immunostaining (see Support Protocol 2), because it does not form distinct plaques for accurate quantitation.

In this protocol, nine-day-old chicken embryos are trypsinized and plated in appropriate medium. After the cells form a confluent monolayer, they are transferred directly to a 31°C incubator where they can be held for 2 to 3 weeks before making secondary CEF for virus infection. With each passage, CEF require longer to become confluent; therefore use of the second passage is recommended.

Materials

Nine-day-old embryonated eggs (Specific Pathogen Free Eggs, SPAFAS)
70% ethanol
MEM with no additives, 37°C
Trypsin/EDTA: 0.25% (w/v) trypsin/0.02% (w/v) EDTA, 37°C
Complete MEM-10 (Table 16.16.2; see recipe), 37°C

Sterile dissecting scissors and forceps
100-cm² sterile petri dishes
10-ml syringes
Sterile trypsinization flask with magnetic stir bar
Humidified, 37° and 31°C, 5% CO₂ incubators
500-ml beakers with two layers of gauze taped over tops
Sorvall RC-3B centrifuge and 250-ml centrifuge bottles (or equivalent)
150-cm² tissue culture flasks

Prepare tissue

1. Position ten 9-day-old embryonated eggs with air space (blunt end) up and spray with 70% ethanol.
2. Crack the top of an egg with sterile dissecting scissors, and cut off shell to just above the membrane while keeping the latter intact. Remove membrane with sterile forceps. Repeat with remaining eggs.

NOTE: Healthy eggs have well-formed blood vessels.

3. Remove the embryo from each egg and combine in a 100-cm² sterile petri dish.
4. Remove head and feet from each embryo and place the rest of the body in a 100-cm² sterile petri dish containing 10 ml MEM with no additives. Mince embryos by squeezing through 10-cc syringes (~5 embryos/syringe) into a sterile trypsinization flask.

Dissociate cells

5. Add 100 ml of 37°C trypsin/EDTA and incubate 5 min in a humidified, 5% CO₂ incubator at 37°C, with stirring.
6. Decant fluid from the trypsinization flask into a 500-ml beaker covered with gauze. Transfer filtrate to a 250-ml centrifuge bottle.
7. Add 100 ml fresh trypsin/EDTA to the remaining tissue in the trypsinization flask, and incubate 5 min at 37°C.
8. Pour digest into a second 500-ml beaker covered with gauze, and add filtrate to the first filtrate in the 250-ml centrifuge bottle.
9. Centrifuge 10 min at 1200 × *g*, 4°C, in a Sorvall RC-3B centrifuge.
10. Aspirate and discard supernatant from pellet, add 10 ml complete MEM-10, resuspend by pipetting ~10 to 15 times, and adjust volume to 100 ml.
11. Transfer to a 250-ml centrifuge bottle and centrifuge 10 min at 1200 × *g*, 4°C.
12. Resuspend pellet in 5 ml complete MEM-10 and adjust volume to 30 ml.

Prepare confluent culture

13. Add 1 ml cell suspension to each of thirty 150-cm² tissue culture flasks containing 30 ml complete MEM-10.

14. Incubate at 37°C for several days until confluent, and move flasks to a humidified, 31°C, 5% CO₂ incubator for storage.

The CEF cell cultures can be held for 2 to 3 weeks at 31°C without further attention. Primary CEF can also be used directly for virus growth, or the trypsinized stock of CEF can be frozen in liquid nitrogen for future use.

PREPARATION OF AN MVA STOCK

To prepare an MVA stock, CEF (see Basic Protocol 4) or BHK-21 (see Basic Protocol 1) cells are plated several days before infection and allowed to approach confluency. They are then infected with MVA. After several days, the infected cells are harvested, lysed by repeated freezing and thawing, sonicated, dispensed in small aliquots, and stored at -70°C.

NOTE: BHK-21 cells should be acquired from ATCC, as cells from alternative sources may support lower levels of MVA replication.

Materials

150-cm² tissue culture flasks of nearly confluent CEF (see Basic Protocol 4) or BHK-21 cells (see Basic Protocol 1)

Complete MEM-10 and -2.5 (see recipe), 37°C

Modified vaccinia virus Ankara (MVA; A. Mayr, Institut für Med. Mikrobiologie, Munich, Germany, or B. Moss, email *bmoss@nih.gov*)

150-cm² tissue culture flasks

Humidified, 37°C, 5% CO₂ incubator

Sorvall RC-3B centrifuge and sterile 250-ml centrifuge bottles (or equivalent)

Additional reagents and equipment for trypsinizing cells (see Basic Protocol 1)

1. Trypsinize seven 150-cm² tissue culture flasks of nearly confluent CEF *or* two of BHK-21 cells as described (see Basic Protocol 1, steps 4 to 7), using 1.5 ml trypsin/EDTA and 8.5 ml maintenance medium for the larger flask. Distribute to twenty 150-cm² flasks. Incubate at 37°C in a humidified, 5% CO₂ incubator until nearly confluent monolayers have formed (usually 2 days).

CEF and BHK-21 cells are split 1:3 and 1:10, respectively. It is advisable to use monolayers of BHK-21 cells at 90% confluency, as completely confluent cultures degrade within 48 hr.

2. Remove medium and add 30 ml complete MEM-2.5 to each 150-cm² flask.

3. Thaw MVA virus and sonicate 30 sec on ice to break up clumps.

Trypsinization of virus is avoided with MVA because it may lower the titer.

4. Add 1 to 3 infectious units of virus/cell to the medium and place cells in a CO₂ incubator for 3 days at 37°C.

A monolayer of CEF or BHK-21 cells in a 150 cm² flask contains ~1–2 × 10⁷ cells. If there is insufficient virus, then use fewer or smaller flasks of confluent cells.

5. Detach the cells with a cell scraper or, if possible, by shaking. Transfer cells and medium to a sterile 250-ml centrifuge bottle and centrifuge 10 min at 1200 × *g*, 4°C, in a Sorvall RC-3B centrifuge. Discard supernatant.

6. Resuspend cells in 1 ml complete MEM-2.5 (per 150-cm² flask) by pipetting or vortexing.

7. Lyse cells by freeze-thaw cycling: freeze in dry ice/ethanol, thaw in a 37°C water bath, and vortex. Carry out the freeze-thaw cycling a total of three times.

8. Sonicate stock on ice for 1 min, pause for 1 min, and repeat sonication. Dispense in 0.5-ml aliquots and store indefinitely at -70°C .

TITRATION OF MVA STOCKS BY IMMUNOSTAINING

The titer of MVA is routinely determined by immunostaining, because MVA does not form distinct plaques in CEF or BHK-21 cells. To determine the MVA titer, serial dilutions of virus stock are prepared and used to infect CEF or BHK-21 cells. At 24 hr after infection, the cells are fixed with acetone/methanol and immunostained using a polyclonal vaccinia virus antiserum and a secondary peroxidase-conjugated antibody. A focus of infected cells appears as a 0.3- to 0.4-mm reddish brown-stained area. If the cells are left for two days before immunostaining, secondary foci may form by virus spread giving inaccurate titers.

Materials

- CEF (see Basic Protocol 4) or BHK-21 cells (see Basic Protocol 1) in a 150-cm² tissue culture flask
- Complete MEM-10 and -2.5 (see recipe), 37°C
- MVA stock (see Basic Protocol 5)
- 1:1 (v/v) acetone/methanol
- PBS (*APPENDIX 2*), with and without 3% FBS
- Rabbit anti-vaccinia antibody (e.g., Access Biomedical & Diagnostic Research Labs or see Linscott, 1998)
- Horseradish peroxidase-conjugated whole anti-rabbit Ig antibody (HRP-anti-rabbit; Amersham)
- Dianisidine, or premade peroxidase substrate kit (Sigma)
- PBS/H₂O₂ (add 10 μl 30% H₂O₂ to 10 ml PBS immediately before use)
- 6-well, 35-mm² tissue culture dishes
- Humidified, 37°C, 5% CO₂ incubator
- Additional reagents and equipment for trypsinizing cells (see Basic Protocol 1)

Infect and fix cells

1. Trypsinize a 150-cm² tissue culture flask of CEF or BHK-21 cells as described (see Basic Protocol 1, steps 4 to 6), using 1.5 ml trypsin/EDTA for the larger flask. Resuspend cells by repeated pipetting in 5 ml complete MEM-10, and add an additional 60 ml complete MEM-10.
2. Add 2 ml cell suspension to each well of a 6-well, 35-mm² tissue culture dish. Incubate overnight in a humidified, 5% CO₂ incubator at 37°C to reach near confluency.
3. Remove medium and replace with 2 ml complete MEM-2.5.
4. Thaw MVA stock and sonicate 30 sec on ice.
5. Make eight 10-fold serial dilutions of the virus in complete MEM-2.5, using a fresh pipet for each dilution.
6. Add 0.1 ml of each of the 10⁻⁶, 10⁻⁷, and 10⁻⁸ dilutions to cells in duplicate wells. Swirl gently to mix, and incubate 24 hr in a CO₂ incubator at 37°C.
7. Remove fluid and fix cells 2 min with 1 ml of 1:1 acetone/methanol. Remove fixative and add 2 ml PBS to each well.

At this point, the plates can be immunostained immediately or stored at 4°C for several weeks.

Perform immunostaining

8. Dilute rabbit anti-vaccinia antibody in PBS containing 3% FBS and add 1 ml/well. Incubate 1 hr at room temperature, with gentle rocking, if desired.

The optimal dilution for each antibody must be determined empirically, but a 1:500 or 1:1000 dilution is a good place to start.

9. Wash twice with 2 ml PBS.

10. Dilute HRP-anti-rabbit secondary antibody in PBS containing 3% FBS and add 1 ml/well. Incubate for 30 to 45 min at room temperature, with gentle rocking, if desired.

The optimal dilution for each antibody must be determined, but a 1:500 or 1:1000 dilution is a good place to start. If another anti-vaccinia antibody is used in step 8, the appropriate HRP-conjugated anti-species secondary antibody should be substituted here.

11. Wash twice with 2 ml PBS.

12. Make a saturated solution of dianisidine in 0.5 ml ethanol, vortex, incubate 5 min at 37°C, and clarify by microcentrifugation at maximum speed for 1 min. Add 0.2 ml dianisidine solution to 10 ml PBS/H₂O₂. Alternatively, use premade substrate tablets according to manufacturer's instructions.

CAUTION: Dianisidine is carcinogenic and the powder should be manipulated with gloves in a fume hood. 30% H₂O₂ is caustic to skin and should be handled with gloves.

13. Add 0.5 ml dianisidine substrate solution to each well. Rotate dish gently and let stand ~10 min. Check microscopically for foci of infected cells. When development is complete, wash dishes with water and overlay with 1 ml water to preserve stain.

Weaker antibody may take longer to develop.

14. Count the number of stained foci and multiply by the dilution factor to express titer as infectious units/ml.

Most accurate results are obtained from wells with 20 to 100 plaques. In determining the titer as infectious units/ml, be sure to multiply by 10 to take into account the 0.1 ml addition of virus to the wells.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Complete DMEM-10 or -20

Dulbecco's minimum essential medium (DMEM) containing:

10% or 20% fetal bovine serum (FBS)

0.03% glutamine

100 U/ml penicillin

100 µg/ml streptomycin sulfate

Store up to several months at 4°C

Add supplements from stock solutions prepared in water at the following initial concentrations: 3% glutamine (100×), 20,000 U/ml penicillin (200×), and 20 mg/ml streptomycin (200×). Filter sterilize stock solutions. Store 100× glutamine 4 months at 4°C, and 200× penicillin/streptomycin 4 months at -20°C.

Fetal bovine serum is added at 10% to complete medium for maintenance of cells and at 20% to encourage initial growth upon thawing. See Chapter 9 introduction for a full discussion concerning media preparation and use of serum (e.g., heat-inactivation, screening).

Complete MEM-2.5, -10, or -20

Minimum essential medium (MEM) containing:

2.5%, 10%, or 20% FBS

0.03% glutamine

100 U/ml penicillin

100 µg/ml streptomycin sulfate

Store up to several months at 4°C

Add supplements from stock solutions as described for complete DMEM (see recipe). Also see annotation to DMEM recipe concerning growth versus maintenance levels of serum.

Complete MEM-10/BrdU or -20/BrdU

Prepare as for complete MEM-10 or -20 (see recipe) and add a 5 mg/ml solution of 5-bromodeoxyuridine (BrdU) to 25 µg/ml final. Store up to several months at 4°C.

Prepare the 5 mg/ml BrdU stock solution (200×) in water and filter sterilize. Store in the dark at -20°C. After thawing 5 mg/ml BrdU, vortex to be sure it is in solution before adding to MEM.

Complete spinner medium-5

MEM spinner medium

5% horse serum

Store up to several months at 4°C

Supplements are in MEM spinner medium; no addition is necessary. Horse serum is used because it is cheaper than FBS and may give less cell clumping.

COMMENTARY

Background Information

An overview of the vaccinia life cycle and expression system is presented in *UNIT 16.15*.

Because HeLa cells consistently give high yields of virus, they are routinely used for preparation of virus stocks. HeLa S3 cells, as obtained from the ATCC, grow well in monolayer culture but can also be put into suspension culture. After repeated passages in suspension, they do not adhere well to flasks and grow poorly in monolayer cultures. The authors prefer HeLa cells adapted to suspension culture because large numbers can be grown in a single bottle. However, suspension cultures require more maintenance than monolayer cultures and the latter may be more convenient.

Suspension cells are allowed to form a monolayer before infection to increase the chances of cell-to-cell spread if not all cells are initially infected with the viral inoculum. Thus, good yields of virus may be obtained even if the inoculum is <1 pfu/cell. HeLa cells—even those adapted to monolayer culture—are fairly round to begin with and therefore show little visible evidence of infection. BS-C-1 cells, by contrast, are long and spindle-shaped but round up dramatically a few hours after infection. This property accounts for the highly visible plaques obtained with BS-C-1 cells, and their preference for titration by plaque assay.

The Western Reserve (WR) strain of vaccinia virus (ATCC no. VR1354) is widely used for laboratory studies. It gives high yields of cell-associated virus, discrete plaques, and is well adapted to mice and other laboratory animals. Other strains of vaccinia virus are available from the ATCC and private sources.

Horse serum is used for growth of HeLa cells in suspension because it is cheaper than FBS and may give less cell clumping. Horse serum is used in spinner medium for growth of HeLa cells in suspension. All monolayer cells are grown in medium with FBS.

Modified vaccinia virus Ankara (MVA; Mayr et al., 1975) was one of several highly attenuated strains of vaccinia virus that were developed but were not extensively used for smallpox vaccination, in part because of the successful eradication of the disease with the conventional vaccine strains. Restriction enzyme analysis demonstrated that MVA had suffered multiple deletions (Meyer et al., 1991) during its long passage history in CEF that may account for its severe host restriction, including an inability to replicate efficiently in human and most other mammalian cells (Carroll and Moss, 1997). It was somewhat surprising to find that the replication defect occurred at a late stage of virion assembly and consequently that viral or recombinant gene expression was unimpaired

(Sutter and Moss, 1992). Therefore, MVA can be viewed as an efficient single-round expression vector with little potential to spread. Although MVA causes less cytopathic effects than standard vaccinia virus in some cell lines, in others the effect is quite severe (Carroll and Moss, 1997).

Because of its host-range restriction, recombinant MVA were initially made and propagated exclusively in primary or secondary CEF. A recent survey demonstrated that a stable Syrian hamster kidney cell line, BHK-21, was also permissive for MVA growth and could be used for making recombinant viruses (Carroll and Moss, 1997). Although many laboratories may prefer to use a stable cell line, the authors continue to employ CEF because they give slightly higher MVA titers, can be stored in a dormant state for several weeks (whereas BHK-21 cells must be passaged every 5 days), and are better suited to immunostaining because of greater adherence than BHK-21 cells to plastic dishes.

Critical Parameters

Proper maintenance of actively growing cell lines is important in order to achieve efficient infections and high yields of virus. This is especially true for HeLa S3 cell suspension cultures, which should be maintained at $1.5\text{--}5 \times 10^5$ cells/ml (requiring passaging every 1 to 2 days). HeLa cells can achieve a density of $8\text{--}10 \times 10^5$ cells/ml, but exhibit a lag period upon dilution before optimal growth is resumed; prolonged maintenance at high density will lead to cell death. HeLa S3 cells adapted to grow in suspension cultures do not grow well in monolayer cultures. However, the cells will adhere to flasks, which allows the cells to be infected on the following day (see Basic Protocol 3).

Since most progeny viruses remain cell-associated, infected cells must be disrupted by freeze-thaw cycling and trypsinization. These procedures are important for releasing virus from the host cells. An entire stock of virus can be subjected to freeze-thaw cycling, but only the portion to be used should be trypsinized. Sonication also helps in disaggregating virus but is usually unnecessary if the virus stock has been trypsinized. Unlike standard replication-competent vaccinia protocols, trypsin is not used as a pretreatment to enhance the infectivity of MVA. Pretreatment of MVA with trypsin can decrease virus titer.

Although vaccinia virus is relatively stable, stocks should be kept on ice while in use and

should be stored at -70°C . A vaccinia virus stock should have a stable titer for many years when stored at this temperature. Although the stock can be frozen and thawed several times without loss of infectivity, storing in aliquots is recommended.

The preparation of a seed stock of virus that is sufficient for future experiments is strongly recommended. A common mistake is to continually passage the virus.

Anticipated Results

A vaccinia stock should have a titer of $\geq 1\text{--}2 \times 10^9$ pfu/ml. MVA stocks can approach a titer of 1×10^{10} infectious units/ml when prepared as described (see Basic Protocol 5), and recombinant virus can reach $1\text{--}3 \times 10^9$ infectious units/ml. The viral stock should have a stable titer for many years when stored as indicated above (see Critical Parameters).

Time Considerations

Actively growing cells should be prepared in advance, as it will take ≥ 1 week to revive frozen cells. For preparation of a vaccinia virus stock, infected cells should be incubated for 3 days. During this period, the BS-C-1 cells for plaque titration can be prepared. For a small volume of stock (i.e., from one 150-cm² flask), harvesting and freeze-thaw cycling can be done in <1 hr. Determination of the titer of a virus stock requires 2 days of incubation to allow for development of plaques.

Preparation of CEF cells from embryonated eggs requires ~ 2 hr. Thereafter, they can be stored for several weeks with no maintenance. Starting BHK-21 cells from frozen cultures can take >1 week and requires passaging every 5 days. For preparation of MVA stocks, 3 days are required for virus growth, 2 hr for harvesting, freeze-thaw cycling, and dispensing, and 1 day of incubation for determination of the virus titer by immunostaining. The immunostaining procedure takes ~ 3 hr.

Literature Cited

- Carroll, M.W. and Moss, B. 1997. Host range and cytopathogenicity of the highly attenuated MVA strain of vaccinia virus: Propagation and generation of recombinant viruses in a nonhuman mammalian cell line. *Virology* 238: 198-211.
- Linscott, W.D. 1998. Linscott's Directory of Immunological and Biological Reagents, 10th ed. W.D. Linscott, Santa Rosa, CA.
- Mayr, A., Hochstein-Mintzel, V., and Stickl, H. 1975. Abstammung, eigenschaften and verwendung des attenuierten vaccinia-stammes MVA. *Infection* 3:6-14.

Meyer, H., Sutter, G., and Mayer, A. 1991. Mapping of deletions in the genome of the highly attenuated vaccinia virus MVA and their influence on virulence. *J. Gen. Virol.* 72:1031-1038.

Sutter, G. and Moss, B. 1992. Nonreplicating vaccinia virus efficiently expresses recombinant genes. *Proc. Natl. Acad. Sci. U.S.A.* 89:10847-10851.

Contributed by Patricia L. Earl, Norman
Cooper, Linda S. Wyatt, and Bernard Moss
National Institutes of Allergy and
Infectious Diseases
Bethesda, Maryland

Miles W. Carroll
Oxford BioMedica
Oxford, United Kingdom

This unit first describes how to infect cells with vaccinia virus and then transfect them with a plasmid-transfer vector to generate a recombinant virus (see Basic Protocol 1). Methods are also presented for purifying vaccinia virus (see Support Protocol 1) and for isolating viral DNA (see Support Protocol 2), which can be used during transfection. Also presented are selection and screening methods used to isolate recombinant viruses (see Basic Protocol 2) and a method for the amplification of recombinant viruses (see Basic Protocol 3). Finally, a method for live immunostaining that has been used primarily for detection of recombinant modified vaccinia virus Ankara (MVA) is presented (see Basic Protocol 4).

HeLa S3 cells are used for large-scale growth of vaccinia virus. However, several other cell lines may be required for plaque purification and amplification. For thymidine kinase (TK) selection, HuTK⁻ 143B cells are used; for xanthine-guanine phosphoribosyltransferase (XGPRT) selection, BS-C-1 cells are used. CV-1 or BS-C-1 cells are used for transfection, and either can be used for determination of virus titer (UNIT 16.16). With MVA, all steps are carried out in CEF or BHK-21 cells (see Basic Protocol 4).

CAUTION: Proceed carefully and follow biosafety level 2 (BL-2) practices when working with standard vaccinia virus (see UNIT 16.15 for safety precautions).

NOTE: Carry out all procedures for preparation of virus in a biosafety cabinet.

TRANSFECTION OF INFECTED CELLS WITH A VACCINIA VECTOR

The foreign gene of interest is subcloned into a plasmid transfer vector (Fig. 16.17.1 and Fig. 16.17.2) so that it is flanked by DNA from a nonessential region of the vaccinia genome. This recombinant plasmid is then transfected into cells that have been infected with vaccinia virus. Homologous recombination occurs between the vaccinia virus genome and homologous sequences within the plasmid (Fig. 16.17.1). The recombinant virus is obtained in a cell lysate, which is then subjected to several rounds of plaque purification using appropriate selection and/or screening protocols (see Basic Protocol 2). For modified vaccinia virus Ankara (MVA), the same transfection procedure is used, except that the virus is not trypsinized and either CEF (UNIT 16.16) or BHK-21 cells are substituted for CV-1 or BS-C-1 cells. For live immunostaining to detect recombinant MVA or standard virus, see Basic Protocol 4.

Materials

- pSC11, pRB21, pSC65, pLW9, or other suitable vector (Table 16.17.1)
- CV-1, BS-C-1, BHK-21, or CEF cells (UNIT 16.16)
- Complete MEM-10 and -2.5 media (UNIT 16.16)
- Vaccinia virus stock (UNIT 16.16)
- 0.25 mg/ml trypsin (2× crystallized and salt-free; Worthington; filter sterilize and store at -20°C)
- Transfection buffer (see recipe)
- 2.5 M CaCl₂
- 20 mM HEPES, pH 7.4
- DOTAP liposomal transfection reagent (Boehringer Mannheim)
- OptiMEM medium (Life Technologies)
- Dry ice/ethanol bath
- 25-cm² tissue culture flask

BASIC PROTOCOL 1

Protein Expression

12 × 75-mm polystyrene tubes
 Disposable scraper or rubber policeman, sterile
 15-ml conical centrifuge tubes
 Sorvall centrifuge with H-6000A rotor (or equivalent)

Additional reagents and equipment for subcloning (UNIT 3.16), isolation of plasmid (UNIT 1.7), and tissue culture (APPENDIX 3F)

NOTE: All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

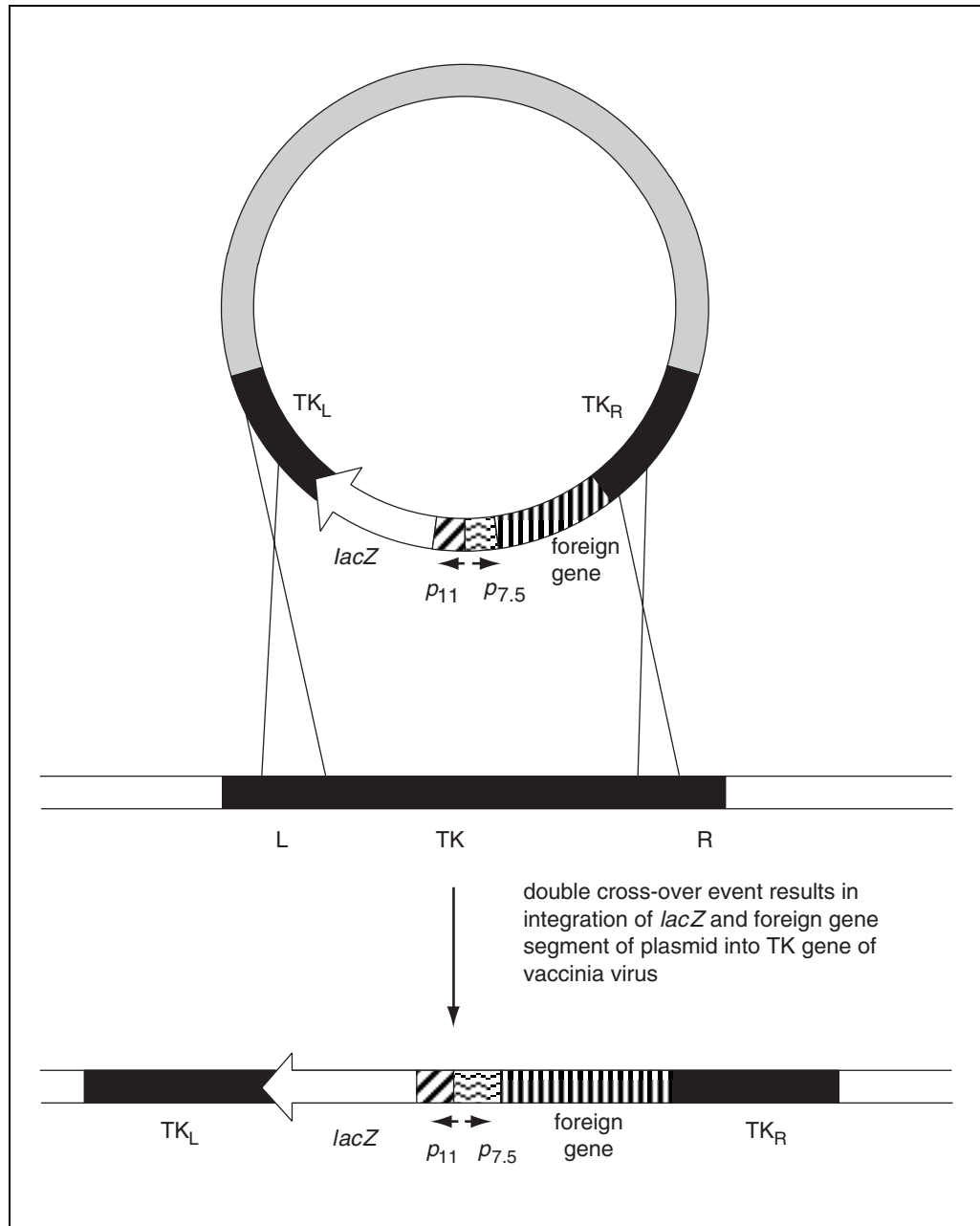


Figure 16.17.1 Homologous recombination between a transfected plasmid and the vaccinia virus genome. TK_L and TK_R are vaccinia virus DNA sequences flanking the foreign gene. p₁₁ and p_{7.5} are promoters.

Prepare recombinant plasmid DNA

1. Subclone the gene of interest into the multiple cloning site (MCS) in pSC11, pRB21, pSC65, pLW9, or other suitable vector (UNIT 3.16) and isolate plasmid (UNIT 1.7).

Prepare vaccinia-infected cells

2. Seed a 25-cm² flask with 1×10^6 CV-1, BS-C-1, BHK-21, or CEF cells in complete MEM-10 medium. Incubate to near confluency (usually overnight).

UNIT 16.16 details the culture of these cells.

3. Just prior to use, mix equal volumes of vaccinia virus stock and 0.25 mg/ml trypsin, and vortex vigorously. Incubate 30 min in a 37°C water bath, vortexing at 5- to 10-min intervals during the incubation. For MVA, substitute 30 sec sonication on ice for trypsinization to break up clumps (also see UNIT 16.16).

Virus stocks are usually at a titer of $\sim 2 \times 10^9$ pfu/ml, but may be significantly lower depending on the source.

Vortexing usually breaks up any clumps of cells. However, if there are still visible clumps, chill to 0°C and sonicate 30 sec on ice. Sonication can be repeated several times but the sample should be allowed to cool on ice between sonications. Trypsinization of MVA is avoided because it may lower the titer.

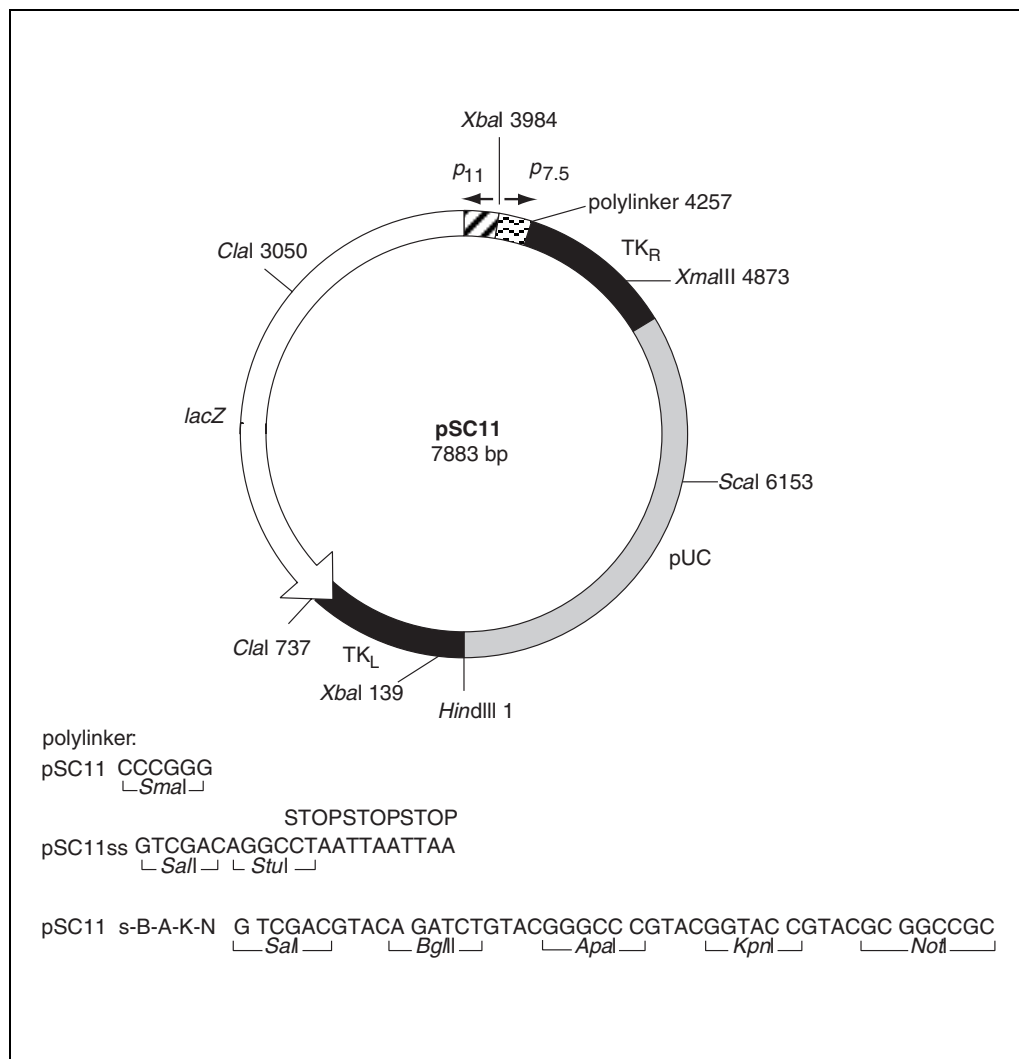


Figure 16.17.2 Plasmid transfer vector pSC11.

4. Dilute trypsinized or sonicated virus in complete MEM-2.5 to 1.5×10^5 pfu/ml. Aspirate medium from confluent monolayer of cells and infect with 1 ml diluted vaccinia virus (0.05 pfu/cell). Incubate 2 hr with rocking at ~15-min intervals.

Approximately 30 min before the end of the infection period, prepare DNA according to the CaCl_2 or DOTAP method as described below.

Prepare DNA

For CaCl_2 method

- 5a. Place 1 ml transfection buffer into a 12×75 -mm polystyrene tube and add 5 to 10 μg (in $<50 \mu\text{l}$) of the recombinant plasmid containing the gene of interest (from step 1).

Inclusion of vaccinia DNA (see Support Protocol 2) in the transfection is not required but results in a higher efficiency of recombination. If it is added, combine 1 μg vaccinia DNA

Table 16.17.1 Vaccinia Virus Transfer Vectors

Vector ^a	Promoter ^b	Cloning sites ^c	Insertion sites ^d	Selection/screening	Reference
pGS20	$p_{7.5}$ (E/L)	<i>Bam</i> HI; <i>Sma</i> I	TK	TK ⁻	Mackett et al., 1984
pSC11	$p_{7.5}$ (E/L)	<i>Sma</i> I; MCS	TK	TK ⁻ , β -gal	Chakrabarti et al., 1985; Earl et al., 1990; Bacik et al., 1994
pMJ601, pMJ602	p_{syn} (L)	MCS	TK	TK ⁻ , β -gal	Davison and Moss, 1990
pRB21	p_{syn} (E/L)	MCS	F12L/F13L	Plaque	Blasco and Moss, 1995
pMC02	p_{syn} (E/L)	MCS	TK	TK ⁻ , GUS	Carroll and Moss, 1995
pSC59	p_{syn} (E/L)	MCS	TK	TK ⁻	Chakrabarti et al., 1997
pSC65	p_{syn} (E/L)	MCS	TK	TK ⁻ , β -gal	Chakrabarti et al., 1997
pJS4	p_{syn} (E/L) $\times 2$	MCS	TK	TK ⁻	Chakrabarti et al., 1997
pJS5	p_{syn} (E/L) $\times 2$	MCS	TK	TK ⁻ , gpt	Chakrabarti et al., 1997
pG06	p_{syn} (E/L) $\times 2$	MCS	Del III	Transient gpt ^e	Sutter et al., 1994
pLW-7	p_{syn} (E/L)	MCS	Del III	Transient gpt ^e	Wyatt et al., 1996
pMC03	p_{syn} (E/L)	MCS	Del III	GUS	Carroll and Moss, 1995
pLW-9	p_{H5} (E/L)	MCS	Del III	Transient gpt ^e	Wyatt et al., 1996
pLW-17	p_{H5} (E/L)	MCS	Del II	None	L. Wyatt and B. Moss, unpub. observ.
pLW-21	p_{syn} (E/L)	MCS	Del II	None	L. Wyatt and B. Moss, unpub. observ.
pLW-22	p_{syn} (E/L)	MCS	Del II	β -gal	L. Wyatt and B. Moss, unpub. observ.
pLW-24	$p_{7.5}$ (E/L)	MCS	Del II	None	L. Wyatt and B. Moss, unpub. observ.

^apRB21 was specifically designed for use with vaccinia virus vRB12, which has a deletion in the F13L gene. The plasmids pG06, pLW-7, pMC03, pLW-9, pLW-17, pLW-21, pLW-22, and pLW-24 were designed for MVA.

^bAbbreviations: E, early; L, late; E/L, early and late. The designation " $\times 2$ " refers to two oppositely oriented promoters that can be used for expression of two genes.

^c*Sma*I digestion gives a blunt end for cloning any fragment that has been blunt-ended. MCS signifies multiple cloning sites.

^dAbbreviations: TK, thymidine kinase locus; F12L/F13L, between F12L and F13L open reading frames; Del III, site of natural deletion in MVA.

^eTransient selection in which XGPRT gene is deleted from recombinant vaccinia virus during recombination; see Background Information.

with 5 to 10 μg recombinant plasmid DNA and vortex vigorously to shear the high-molecular-weight DNA before adding CaCl_2 or DOTAP. A discussion of calcium phosphate transfection can be found in UNIT 9.1; also see Critical Parameters.

- 6a. Slowly add 50 μl of 2.5 M CaCl_2 and mix gently. Leave 20 to 30 min at room temperature.

Gentle mixing is essential; see Critical Parameters. A fine precipitate should appear.

For DOTAP method

- 5b. Add 70 μl of 20 mM HEPES, pH 7.4, and 30 μl DOTAP to a 12 \times 75-mm polystyrene tube. In a second tube, add 5 μg recombinant plasmid containing the gene of interest (from step 1) to 20 mM HEPES for a final volume of 100 μl .

See step 5a regarding optional addition of vaccinia DNA.

- 6b. Add the DNA solution to the DOTAP solution and leave 15 min at room temperature, then add 1 ml OptiMEM medium.

Transfect cells

7. Aspirate virus inoculum from monolayer of cells (step 4).
- 8a. *For CaCl_2 method:* Add the precipitated DNA suspension (step 6a) to the cells and leave 30 min at room temperature, then add 9 ml complete MEM-10 and incubate 3 to 4 hr at 37°C.
- 8b. *For DOTAP method:* Add the DNA/DOTAP solution (step 6b) to the cells and incubate 4 hr at 37°C.
9. Aspirate medium, replace with 5 ml complete MEM-10, and incubate 2 days at 37°C.
10. Dislodge cells with a disposable scraper or sterile rubber policeman and transfer to a 15-ml conical centrifuge tube. Centrifuge 5 min at 1800 $\times g$ (2500 rpm in a Sorvall H-6000A rotor), 5° to 10°C, then aspirate and discard medium.
11. Resuspend cells in 0.5 ml complete MEM-2.5. Lyse the cell suspension by performing three freeze-thaw cycles, each time by freezing in a dry ice/ethanol bath, thawing in a 37°C water bath, and vortexing.
12. Store the cell lysate at -70°C until needed in the selection and screening procedure (see Basic Protocol 2).

PURIFICATION OF VACCINIA VIRUS

Vaccinia is usually purified by zonal sucrose gradient centrifugation. Purified virus is useful for preparation of vaccinia DNA (see Support Protocol 2), in studies in which contaminating infected-cell proteins are undesirable, and as a very high titer stock. For large-scale purification (as in this protocol, which is for 1-liter cultures or multiples thereof) it is preferable to use HeLa cell suspensions for infection rather than monolayer cultures. If monolayer cells are used, follow the alternative procedure for monolayers (see UNIT 16.16, Basic Protocol 3) then continue with step 9 of this protocol. For many purposes, virus that is partially purified (through step 14 of this protocol) may suffice. For purification of MVA, virus is not trypsinized and either CEF or BHK-21 cells are used instead of HeLa.

SUPPORT PROTOCOL 1

Protein Expression

16.17.5

Materials

Vaccinia virus stock (UNIT 16.16)
0.25 mg/ml trypsin (2× crystallized and salt-free; Worthington; filter sterilize and store at -20°C)
HeLa S3 cells growing in suspension culture (UNIT 16.16)
Complete spinner medium-5 (UNIT 16.16)
10 mM and 1 mM Tris·Cl, pH 9.0 (APPENDIX 2)
36% (w/v) sucrose solution in 10 mM Tris·Cl, pH 9.0
40%, 36%, 32%, 28%, and 24% (w/v) sucrose solutions in 1 mM Tris·Cl, pH 9.0
Sorvall centrifuge with H-6000A rotor (or equivalent)
2-liter vented spinner flasks (microcarrier type; Bellco)
Dounce homogenizer, glass, with tight pestle
Cup sonicator (e.g., Ultrasonic Processor VC-600 from Sonics and Materials)
Ultracentrifuge with Beckman SW 27 or SW 28 rotor (or equivalent) and sterile centrifuge tubes
Spectrophotometer
Additional reagents and equipment for tissue culture and counting cells (APPENDIX 3F), and titering virus (UNIT 16.16)

NOTE: All culture incubations should be performed in a humidified 37°C , 5% CO_2 incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

Infect the cells

1. Just prior to use, mix equal volumes of vaccinia virus stock and 0.25 mg/ml trypsin and vortex vigorously. Incubate 30 min in a 37°C water bath, vortexing at 5- to 10-min intervals during the incubation.

Virus stocks are usually at a titer of $\sim 2 \times 10^9$ pfu/ml, but may be significantly lower depending on the source.

Vortexing usually breaks up any clumps of cells. However, if there are still visible clumps, chill to 0°C and sonicate 30 sec on ice. Sonication can be repeated several times but the sample should be allowed to cool on ice between sonications.

2. Count HeLa S3 cells using a hemacytometer (APPENDIX 3F).
3. Transfer 5×10^8 HeLa S3 cells into a centrifuge tube, then centrifuge 10 min at $1800 \times g$ (2500 rpm in an H-6000A rotor), room temperature, and discard supernatant.
4. Resuspend cells in complete spinner medium-5 at 2×10^7 cells/ml.
5. Add trypsinized virus (from step 1) at an MOI of 5 to 8 pfu/cell and incubate 30 min with stirring.
6. Transfer cells to a vented spinner flask containing 1 liter complete spinner medium-5 and incubate 2 to 3 days with stirring.
7. Centrifuge cells 5 min at $1800 \times g$, 5° to 10°C , and discard supernatant.
8. Resuspend cells in 14 ml of 10 mM Tris·Cl, pH 9.0. Keep samples on ice for the remainder of the protocol.

Lyse the cells

9. Homogenize cell suspension with 30 to 40 strokes in a glass Dounce homogenizer with tight pestle. Check for cell breakage by light microscopy.
10. Centrifuge 5 min at $300 \times g$ (900 rpm in H-6000A rotor), 5° to 10°C , to remove nuclei. Save the supernatant.
11. Resuspend cell pellet in 3 ml of 10 mM Tris·Cl, pH 9.0. Centrifuge 5 min at $300 \times g$, 5° to 10°C . Save supernatant and pool with supernatant from step 10.
12. Sonicate pooled supernatants (lysate), keeping the lysate cold the entire time, using a cup sonicator as follows.
 - a. Split the sample into 3-ml aliquots to be sonicated separately.
 - b. Fill the cup with ice-water (~50% ice). Place the tube containing the lysate in the ice-water and sonicate 1 min at full power.
 - c. Repeat this three to four times, placing the lysate on ice for ≥ 30 sec between sonications.

Because sonication melts the ice, it is necessary to replenish the ice in the cup.

Obtain purified virus

13. Layer the sonicated lysate onto a cushion of 17 ml of 36% sucrose in a sterile SW 27 (or SW 28) centrifuge tube. Centrifuge 80 min at $32,900 \times g$ (13,500 rpm in an SW 27 rotor), 4°C . Aspirate and discard the supernatant.
14. Resuspend the viral pellet in 1 ml of 1 mM Tris·Cl, pH 9.0.

At this stage, the virus may be sufficiently pure for some purposes—e.g., isolation of DNA.
15. Sonicate once for 1 min as in step 12.
16. Prepare a sterile 24% to 40% continuous sucrose gradient in a sterile SW 27 centrifuge tube the day before it is needed by carefully layering 6.8 ml each of 40%, 36%, 32%, 28%, and 24% sucrose. Let sit overnight in refrigerator.
17. Overlay the sucrose gradient with 1 ml of sonicated viral pellet from step 15. Centrifuge 50 min at $26,000 \times g$ (12,000 rpm in an SW 27 rotor), 4°C .
18. Observe the virus as a milky band near the middle of the tube. Aspirate the sucrose above the band and discard. Carefully collect the virus band (~10 ml) with a sterile pipet, place in a sterile tube, and save.
19. Collect aggregated virus from the pellet at the bottom of the sucrose gradient after aspirating the remaining sucrose from the tube. Resuspend the viral pellet by pipetting up and down in 1 ml of 1 mM Tris·Cl, pH 9.0.
20. Sonicate resuspended pellet once for 1 min as in step 12.
21. Reband the virus from the pellet as in steps 16 to 18 and pool band with band from step 18. Add 2 vol of 1 mM Tris·Cl, pH 9.0, and mix. Transfer to sterile SW 27 centrifuge tubes.

The total volume should be ~60 ml, which is enough to fill two SW 27 centrifuge tubes. If less volume is obtained, fill the tubes with 1 mM Tris·Cl, pH 9.0.
22. Centrifuge 60 min at $32,900 \times g$, 4°C , then aspirate and discard supernatant.
23. Resuspend the virus pellets in 1 ml of 1 mM Tris·Cl, pH 9.0. Sonicate as in step 12 and divide into 200- to 250- μl aliquots. Save one aliquot for step 24 and freeze remainder at -70°C .

Quantitate virus

24. On the unfrozen aliquot, estimate the amount of virus spectrophotometrically at 260 nm. Use this to determine the amount of virus needed in the protocol for isolation of vaccinia virus DNA (see Support Protocol 2).

One optical density unit is $\sim 1.2 \times 10^{10}$ virus particles, which is $\sim 2.5\text{--}5 \times 10^8$ pfu. This value is due to light scattering, not absorbance, and therefore may vary slightly with different spectrophotometers.

25. Titer the virus by sonicating 20 to 30 sec on ice (see step 12), preparing 10-fold serial dilutions down to 10^{-10} , and infecting, in duplicate, confluent BS-C-1 cell monolayers in 6-well tissue culture plates using the 10^{-8} , 10^{-9} , and 10^{-10} dilutions (UNIT 16.16).

SUPPORT PROTOCOL 2

ISOLATION OF VACCINIA VIRUS DNA

If vaccinia DNA is used in the transfection protocol (see Basic Protocol 1, step 5a annotation), it is isolated after a proteinase K digestion and phenol extraction. The extracted DNA is precipitated and then dissolved in water. The DNA concentration is determined by measuring A_{260} .

Materials

Purified vaccinia virus (see Support Protocol 1)
50 mM and 1 M Tris·Cl, pH 7.8 (APPENDIX 2)
10% (w/v) sodium dodecyl sulfate (SDS)
60% (w/v) sucrose
10 mg/ml proteinase K
Buffered phenol: phenol equilibrated with 50 mM Tris·Cl, pH 7.8 (UNIT 2.1)
1:1 (v/v) phenol/chloroform
1 M sodium acetate, pH 7.0
100% and 95% ethanol

Spectrophotometer
Sorvall centrifuge with H-6000A rotor (or equivalent)

Additional reagents and equipment for phenol extraction of DNA (UNIT 2.1) and measuring DNA concentration (APPENDIX 3D)

NOTE: Avoid vortexing the DNA throughout this protocol.

1. Measure the optical density of purified vaccinia virus at 260 nm and bring 20 optical density units of virus (also see Support Protocol 1, step 24) to 1.2 ml final volume with 50 mM Tris·Cl, pH 7.8.

Because the vaccinia virus genome is large, care must be taken to avoid shearing (i.e., by not vortexing) if full-length DNA (such as for restriction digestion analysis) is desired.

2. Add the following to the vaccinia suspension (for 2 ml final volume):

0.1 ml 1 M Tris·Cl, pH 7.8
0.1 ml 10% SDS
0.2 ml 60% sucrose
0.4 ml 10 mg/ml proteinase K

Incubate 4 hr at 37°C.

3. Extract twice with phenol, each time by adding an equal volume of buffered phenol, mixing gently by rocking the tube, centrifuging 10 min at $300 \times g$ (900 rpm in an H-6000A rotor), room temperature, then aspirating and saving the aqueous phase using a pipet with the tip cut off (also see UNIT 2.1).

4. Using the technique described in step 3, extract once with 1:1 phenol/chloroform.
5. Add $\frac{1}{10}$ vol of 1 M sodium acetate, pH 7.0 and 2.5 vol of 100% ethanol. Mix gently and chill several hours at -20°C .
6. Microcentrifuge 10 min at maximum speed, 4°C . Aspirate and discard supernatant.
7. Wash pellet twice with 95% ethanol, each time by adding the ethanol, microcentrifuging at maximum speed, and aspirating the supernatant. Air dry and dissolve in 100 μl water.
8. Prepare dilutions (using only a small amount of the total material) and measure A_{260} to determine DNA concentration (APPENDIX 3D).

SELECTION AND SCREENING OF RECOMBINANT VIRUS PLAQUES

For standard vaccinia virus strains, procedures are described involving xanthine-guanine phosphoribosyltransferase (XGPRT; Falkner and Moss, 1988) or thymidine kinase (TK; Mackett et al., 1984) for selecting virus plaques that contain recombinant DNA. A newer and simpler drug-free method (plaque selection) is based on repair of a mutation that caused the parental virus to form pin-point plaques (Blasco and Moss, 1995; see Background Information regarding choice of procedure). In addition, β -galactosidase screening (Chakrabarti et al., 1985) or β -glucuronidase (GUS) screening (Carroll and Moss, 1995) can be used alone or in conjunction with TK selection to discriminate TK^- recombinants from spontaneous TK^- mutants. For each method, recombinant virus (obtained in the transfection; see Basic Protocol 1) is used to infect a monolayer culture of cells—for XGPRT selection, BS-C-1 cells are used because large plaques are obtained; for TK selection, it is necessary to use a cell line such as HuTK⁻ 143B that is deficient in thymidine kinase. Molten agarose with medium containing the appropriate selective drugs is then pipetted onto the infected cell monolayer. Because of cell-to-cell spread of virus, each productively infected cell gives rise to a plaque. After two days, a second agarose overlay containing neutral red is placed on top of the first agarose overlay, whereupon neutral red is taken up by viable cells. In infected areas of the monolayer, the cells are rounded and dead, and thus appear as colorless plaques after neutral red staining. If β -galactosidase (or GUS) screening is used, the substrate Xgal (or Xgluc) is included in the second agarose overlay. Plaques containing infected cells that have expressed β -galactosidase or GUS turn blue; thus, blue (recombinant) plaques can be distinguished from clear (parental) plaques. A Pasteur pipet is used to aspirate infected cells from the plaques and the virus is released by freeze-thaw cycling and sonication. Several rounds of plaque purification are used to ensure the absence of residual non-recombinant virus.

For MVA, the virus is not trypsinized and the XGPRT, β -galactosidase, and GUS methods are used with CEF or BHK-21 cells. TK selection has not been used with MVA. Basic Protocol 4 describes live immunostaining of recombinant MVA.

Materials

BS-C-1, HuTK⁻ 143B, BHK-21, or CEF confluent monolayer cells (UNIT 16.16) and appropriate complete medium

Complete MEM-2.5 medium (UNIT 16.16)

Selective agents (for XGPRT selection; filter sterilize, and store at -20°C):

10 mg/ml (400 \times) mycophenolic acid (MPA; Calbiochem) in 0.1 N NaOH

10 mg/ml (40 \times) xanthine in 0.1 N NaOH

10 mg/ml (670 \times) hypoxanthine in 0.1 N NaOH

Transfected cell lysate (see Basic Protocol 1)

2% LMP agarose (Life Technologies) in H_2O , sterilized by autoclaving

BASIC PROTOCOL 2

Protein Expression

16.17.9

Complete 2× plaque medium-5 (see recipe)
5 mg/ml 5-bromodeoxyuridine (BrdU) in H₂O (for TK selection; filter sterilize and store at -20°C)
10 mg/ml neutral red in H₂O
4% Xgal in dimethylformamide (optional, for β-galactosidase screening; Table 1.4.2)
2% Xgluc in dimethylformamide (optional, for GUS screening)
Dry ice/ethanol bath
6-well, 35-mm tissue culture plates
Cup sonicator (e.g., Ultrasonic Processor VC-600 from Sonics and Materials)
45°C water bath
Cotton-plugged Pasteur pipets, sterile
Additional reagents and equipment for tissue culture and counting cells (*APPENDIX 3F*) and serial dilution of virus (*UNIT 1.11*)

NOTE: All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

Prepare the cells

1. Trypsinize confluent monolayer culture and resuspend in appropriate complete medium as in *UNIT 16.16*, steps 4 to 7 of Basic Protocol 1.
 - a. For XGPRT selection, plaque selection, or color screening, use BS-C-1 cells.
 - b. For TK selection, use HuTK⁻ 143B cells.
 - c. For MVA, use BHK-21 or CEF cells.
2. Count cells using a hemacytometer (*APPENDIX 3F*).
3. Plate 5 × 10⁵ cells/well in a 6-well tissue culture plate (2 ml/well final). Incubate until confluent (this should take <24 hr).
4. Prepare cells as follows.
 - a. For XGPRT selection, preincubate monolayer for 12 to 24 hr in filter-sterilized complete MEM-2.5 containing 1/400 vol 10 mg/ml MPA, 1/40 vol 10 mg/ml xanthine, and 1/670 vol 10 mg/ml hypoxanthine.
 - b. For plaque or TK selection or color screening methods, do not preincubate.

Prepare the lysate and infect cells

5. Just prior to use, mix 100 μl of transfected cell lysate and 100 μl of 0.25 mg/ml trypsin, and vortex vigorously. Incubate 30 min in a 37°C water bath, vortexing at 5- to 10-min intervals during the incubation, then sonicate 20 to 30 sec on ice. For MVA, omit trypsinization but sonicate 20 to 30 sec on ice to break up clumps.
6. Make four 10-fold serial dilutions (ranging from 10⁻¹ to 10⁻⁴; *UNIT 1.11*) of the trypsinized and/or sonicated cell lysate in complete MEM-2.5 as follows.
 - a. For XGPRT selection, add MPA, xanthine, and hypoxanthine at the concentrations indicated in step 4, substep a, above.
 - b. For plaque or TK selection or color screening methods, make no additions.

7. Aspirate medium from the cell monolayers (from step 3) and infect with 1.0 ml diluted lysate per well. Incubate 2 hr, rocking at 30-min intervals.

Use dilutions between 10^{-2} and 10^{-4} .

Perform agarose overlays

8. Before the 2-hr infection is finished, melt 2% LMP agarose (1.5 ml \times number of wells) and place in a 45°C water bath to cool (be sure it cools to 45°C before using it to overlay cells). Prepare and warm to 45°C the necessary amount of selective plaque medium (1.5 ml \times number of wells) by making the following additions to complete 2 \times plaque medium-5.
 - a. For XGPRT selection, include MPA, xanthine, and hypoxanthine at twice the concentrations indicated in step 4, substep a.

Twice the concentration is necessary because the 2 \times plaque medium will be mixed 1:1 with agarose.
 - b. For TK selection, include $\frac{1}{100}$ vol of 5 mg/ml BrdU.
 - c. For plaque selection or color screening, make no additions.
9. Prepare appropriate selective agarose by mixing equal volumes of 2% LMP agarose and selective plaque medium from step 8, substep a or b.
10. Aspirate the viral inoculum from the infected cells (from step 7). Overlay each well with 3 ml appropriate selective agarose and allow to solidify at room temperature or 4°C. Incubate 2 days.
11. Prepare the second agarose overlay by mixing equal volumes of 2% LMP agarose (1 ml \times number of wells, melted and cooled to 45°C as in step 8) and 2 \times plaque medium-5 (1 ml \times number of wells, warmed to 45°C) with $\frac{1}{100}$ vol of 10 mg/ml neutral red. If β -galactosidase screening is to be used, add $\frac{1}{120}$ vol of 4% Xgal to the agarose/plaque medium. If GUS screening is to be used, add $\frac{1}{100}$ vol of 2% Xgluc. Overlay each well with 2 ml of this second agarose overlay, allow to solidify, and incubate overnight.

There is no need to add any FBS, glutamine, penicillin/streptomycin, or any selection drugs to this overlay medium.

Obtain the plaques

12. Add 0.5 ml complete MEM-2.5 to sterile microcentrifuge tubes. When incubation period is complete (step 11), pick well-separated plaques by squeezing the rubber bulb on a sterile, cotton-plugged Pasteur pipet and inserting the tip through the agarose to the plaque. Scrape the cell monolayer and aspirate the agarose plug into the pipet. Transfer to a tube containing 0.5 ml complete MEM-2.5. Repeat for six to twelve plaques with separate pipets, placing each in a separate tube.
13. Vortex each virus-containing tube, then perform three freeze-thaw cycles, each time by freezing in a dry ice/ethanol bath, thawing in a 37°C water bath, and vortexing.
14. Place tube containing virus into a cup sonicator containing ice-water and sonicate 20 to 30 sec at full power.

If TK selection only has been used, plaque isolates should be tested by PCR, DNA dot-blot hybridization, or immunostaining (all procedures described in UNIT 16.18), because some plaques will contain spontaneous TK⁻ mutations and not recombinant virus.

Carry out several rounds of plaque purification

15. Prepare monolayers of the appropriate cell line as described in steps 1 to 4.

One 6-well plate is needed for each plaque isolate.

16. Make three 10-fold serial dilutions at 10^{-1} , 10^{-2} , and 10^{-3} , of each of several plaque isolates as described in step 6.

If XGPRT selection is used, cells must be preincubated with selective drugs and serial dilutions of the viral isolates must also contain selective drugs (step 4, substep a).

17. Aspirate medium from cell monolayers and infect two wells with 1.0 ml of each dilution of virus. Incubate 2 hr, rocking by hand at 30-min intervals.
18. Repeat steps 8 to 14 for three or more rounds of plaque purification to ensure a clonally pure recombinant virus.

AMPLIFICATION OF A PLAQUE

A recombinant plaque isolate (obtained after the selection and screening protocol; see Basic Protocol 2) is amplified by infection of successively larger numbers of cells. Medium containing drugs for XGPRT or TK selection is usually used, up to and including the infection of cells in a 25-cm² flask. Freeze-thaw cycling is carried out to release the recombinant virus from the cells. The titer of the virus stock is determined as described in UNIT 16.16. For amplification of MVA, see Basic Protocol 4.

Materials

- Resuspended recombinant plaque (see Basic Protocol 2)
- Confluent monolayer cultures of appropriate cells in both a 12-well, 22-mm tissue culture plate and a 25-cm² tissue culture flask (UNIT 16.16)
- Complete MEM-2.5 and -10 media (UNIT 16.16)
- Selective agents (for XGPRT selection; filter sterilize, and store at -20°C):
 - 10 mg/ml (400 \times) mycophenolic acid (MPA; Calbiochem) in 0.1 N NaOH
 - 10 mg/ml (40 \times) xanthine in 0.1 N NaOH
 - 10 mg/ml (670 \times) hypoxanthine in 0.1 N NaOH
- 5 mg/ml 5-bromodeoxyuridine (BrdU) in H₂O (for TK selection; filter sterilize and store at -20°C)
- Dry ice/ethanol bath
- Spinner culture of HeLa S3 cells (UNIT 16.16)
- Cup sonicator (e.g., Ultrasonic Processor VC-600 from Sonics and Materials)
- 15-ml conical centrifuge tubes
- Sorvall centrifuge with H-6000A rotor (or equivalent)
- 150-cm² tissue culture flask
- Additional reagents and equipment for tissue culture and counting cells (APPENDIX 3F)

NOTE: All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

Infect a monolayer culture of cells with a plaque

1. Place tube containing resuspended recombinant plaque into a cup sonicator containing ice-water and sonicate 20 to 30 sec at full power.
2. Infect appropriate confluent monolayer culture in 12-well plate with 250 μl ($\frac{1}{2}$) of each plaque isolate. Incubate 1 to 2 hr, rocking at \sim 15-min intervals.

If XGPRT selection is used, the monolayer culture is preincubated for 12 to 24 hr in complete MEM-2.5 containing MPA, xanthine, and hypoxanthine (see Basic Protocol 2, step 4, substep a). Infection should also be done in the presence of these drugs.

3. Overlay with 1 ml complete MEM-2.5 medium containing the appropriate selective agents.
 - a. For XGPRT selection, include $\frac{1}{400}$ vol 10 mg/ml MPA, $\frac{1}{40}$ vol 10 mg/ml xanthine, and $\frac{1}{670}$ vol 10 mg/ml hypoxanthine.
 - b. For TK selection, include $\frac{1}{200}$ vol of 5 mg/ml BrdU.

Incubate 2 days or until cytopathic effect (cell rounding) is obvious.

4. Scrape cells, transfer to microcentrifuge tube, and microcentrifuge 30 sec at maximum speed. Aspirate and discard medium.
5. Resuspend cells in 0.5 ml complete MEM-2.5 then perform three freeze-thaw cycles, each time by freezing in a dry ice/ethanol bath, thawing in a 37°C water bath, and vortexing.
6. Place tube containing cell suspension into a cup sonicator and sonicate 20 to 30 sec at full power.

Scale up the culture

7. Dilute 0.25 ml of lysate from step 6 with 0.75 ml complete MEM-2.5 containing selective agents (see steps 2 and 3). Infect appropriate confluent monolayer culture in a 25-cm² flask and incubate 30 min.
8. Overlay with 4 ml complete MEM-2.5 containing the appropriate selective agents (step 3). Incubate 2 days or until cytopathic effect is obvious.
9. Scrape cells, transfer to 15-ml conical centrifuge tube, and centrifuge 5 min at 1800 × g (2500 rpm in H-6000A rotor), 5° to 10°C. Resuspend cells in 0.5 ml complete MEM-2.5 and repeat freeze-thaw cycling and sonication as described in steps 5 and 6.
10. Count HeLa S3 cells from a spinner culture using a hemacytometer (*APPENDIX 3F*).

Alternatively, prepare monolayer cells as described in UNIT 16.16 and proceed to step 13, below.

11. Place 5×10^7 HeLa S3 cells in a centrifuge tube. Centrifuge 5 min at 1800 × g, room temperature, and discard supernatant.
12. Resuspend cells in 25 ml complete MEM-10, dispense in one 150-cm² flask, and incubate overnight (for infection the following day).
13. Remove medium from cells and replace with a mixture of 0.25 ml lysate (from step 9) and 1.75 ml complete MEM-2.5. Incubate 1 hr, rocking the flask at 15- to 30-min intervals.
14. Overlay with 25 ml complete MEM-2.5 (selection is not required at this step) and incubate 3 days.
15. Detach cells from the flask by shaking or scraping if necessary. Transfer to centrifuge tube by pipetting, then centrifuge 5 min at 1800 × g, 5° to 10°C. Aspirate and discard supernatant.
16. Resuspend cells in 2 ml complete MEM-2.5 and carry out freeze-thaw cycling three times as described in step 5.
17. Determine titer of the viral stock as described in *UNIT 16.16*. Freeze viral stock at -70°C.

LIVE IMMUNOSTAINING OF MVA RECOMBINANTS

Live immunostaining was developed for modified vaccinia virus Ankara (MVA) because this strain does not form discrete, easily recognizable plaques and because this technique helps to avoid the incorporation of selectable or screening marker genes. Immunostaining can be used for recombinant proteins that are expressed on the cell surface or in the cytoplasm. These protocols are also applicable to standard strains of vaccinia virus.

The strong adherence of chicken embryo fibroblasts (CEF) to concavalin A-coated plastic dishes make them superior to BHK-21 or other cell lines that the authors have tried.

Materials

- 150-cm² flask of confluent CEF (*UNIT 16.16*)
- Complete MEM-2, -2.5, and -10 media (*UNIT 16.16*)
- Transfected cell lysate (see Basic Protocol 1)
- Dry ice/ethanol bath
- Primary antibody to protein product of foreign gene
- Horseradish peroxidase-conjugated secondary antibody (to species of primary antibody)

- Concanavalin A-coated 6-well tissue culture plates (see Support Protocol 3)
- Cup sonicator (e.g., Ultrasonic Processor VC-600 from Sonics and Materials)
- Inverted microscope
- Sterile toothpicks
- Cell scraper or plunger of 1-ml syringe
- 75- and 150-cm² tissue culture flasks

- Additional reagents and equipment for culture, trypsinization, and immunostaining of CEF cells, titering of MVA, and preparation of MVA stocks (*UNIT 16.16*)

NOTE: All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

Prepare, infect, and immunostain CEF cells

1. Trypsinize confluent CEF monolayer cultures, seed cells in concanavalin A-coated 6-well tissue culture plates, and incubate until nearly confluent (*UNIT 16.16*).
2. Thaw transfected cell lysate and sonicate in a cup sonicator for 30 sec at full power, on ice. Add 100, 10, 1, or 0.1 µl of lysate to duplicate wells of CEF containing 2 ml of complete MEM-2.5 medium. Gently swirl to mix. Incubate 2 days.
3. Immunostain the unfixed cells using antibody to the protein product of the foreign gene (see *UNIT 16.16*, Support Protocol 2, steps 8 to 13).

If the protein of interest is expressed intracellularly, remove the fluid from the wells and place plate at -70°C for 1 hr. Allow to thaw and begin staining as described in UNIT 16.16. This procedure ruptures the cells in situ and allows the antibody to penetrate.

Isolate recombinant virus

4. Aspirate the medium and examine the cell monolayer using an inverted microscope. Touch immunostaining foci with sterile toothpicks. Place toothpicks individually in small vials containing 0.5 ml of complete MEM-2.5, and break off each toothpick so that the sterile part is inside the tube.

5. Perform three freeze/thaw cycles on each tube, each time by freezing in a dry ice/ethanol bath, thawing in a 37°C water bath, and vortexing. Place tube into a cup sonicator and sonicate 30 sec at full power.
6. Replate onto new CEF monolayers and plaque purify the MVA recombinant a second time as in steps 2 to 5. Plaque purify an additional three times.

In this protocol, plaque-purification is done under a liquid overlay instead of agarose to allow immunostaining. To check stability and purity, an extra plate can be held for 3 days and then fixed and immunostained as in UNIT 16.16, Support Protocol 2. The presence of nonstaining foci, which may be detected at this time because of cytopathic effects, are due to wild-type virus or an unstable recombinant.

Amplify plaque-purified virus

7. Infect 1 well of a concanavalin A-coated 6-well plate of CEF or BHK-21 cells with 0.25 ml of plaque-purified MVA recombinant in a total volume of 2 ml MEM-2.5. Incubate 2 days.
8. Remove and discard 1 ml of medium covering the cell monolayer. Dislodge cells into the 1 ml of remaining medium with a cell scraper (or plunger of a 1 ml syringe) and transfer to a vial. Carry out three freeze/thaw cycles as in step 5 to produce a lysate.
9. Amplify the virus by inoculating 0.5 ml of lysate into 75-cm² flask of CEF or BHK-21 cells containing 15 ml MEM-2.5. After 2 days harvest and lyse cells as in step 8.
10. Inoculate a 150-cm² flask of CEF or BHK-21 cells containing 30 ml MEM-2.5 with 0.5 ml of lysate from step 9.
11. Determine titer and prepare large stock of recombinant MVA (UNIT 16.16, Basic Protocol 5). Freeze viral stock in aliquots at -70°C.

COATING PLATES WITH CONCAVALIN A

This protocol describes the preparation of concanavalin A-coated plates to be used growing CEF monolayers for live immunostaining as in Basic Protocol 4.

Materials

Concanavalin A (Sigma)
Phosphate-buffered saline (PBS; APPENDIX 2)
6-well, 35-mm tissue culture dishes
Plastic bags for storage

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

1. Add 12 mg concanavalin A to 120 ml sterile PBS for a final concentration of 100 µg/ml.
2. Add 1 ml of the PBS/concanavalin A solution to each well of twenty 6-well plates. Incubate 1 hr at room temperature.
3. Aspirate the liquid from each well and rinse with 2 ml PBS.
4. Remove the fluid from each well and let plates dry by storing open in a biological safety hood (to keep sterile).
5. Store plates in a plastic bag (to preserve sterility) at room temperature (will keep for months).

SUPPORT PROTOCOL 3

**Protein
Expression**

16.17.15

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Complete 2× plaque medium-5

2× plaque medium (Life Technologies) containing:
5% fetal bovine serum (FBS; APPENDIX 3F)
0.03% (w/v) glutamine
100 U/ml penicillin
100 µg/ml streptomycin
Store up to 3 months at 4°C

Transfection buffer

0.14 M NaCl
5 mM KCl
1 mM Na₂HPO₄·2H₂O
0.1% (w/v) dextrose
20 mM HEPES
Adjust to pH 7.05 with 0.5 M NaOH and filter sterilize
Store indefinitely at -20°C

The pH of this buffer is critical and should be between 7.0 and 7.1.

COMMENTARY

Background Information

Plasmid-transfer vectors have three essential components: an expression cassette consisting of a natural or synthetic vaccinia promoter, restriction endonuclease sites for insertion of foreign genes, and flanking vaccinia virus DNA that determines the site of homologous recombination. An additional component may provide antibiotic selection or color screening. Various transfer vectors are listed in Table 16.17.1 and one is presented in detail in Figure 16.17.2.

Homologous recombination (Fig. 16.17.1) is the usual way of generating recombinant vaccinia viruses. Initially, recombination may result from a single cross-over event, resulting in integration of the circular plasmid and the creation of a tandem duplication; however, this intermediate is highly unstable. A second recombination event then occurs between the tandem repeats, resolving the structure into a small circular DNA molecule and either a wild-type or recombinant viral genome. An alternative method of direct ligation using an unique restriction endonuclease site in the vaccinia genome has also been described (Pfleiderer et al., 1995; Merchlinsky et al., 1997). This method avoids an *E. coli* cloning step and can potentially be used for direct cloning of cDNA libraries in vaccinia virus.

Considerable attention has been devoted to promoters because they affect both the time and

level of expression (UNIT 16.15). The most widely used promoter, *p*_{7.5}, actually contains tandem late and early promoters (Cochran et al., 1985) and provides a moderate level of expression throughout infection. The modified *p*_{H5} promoter provides higher levels of both early and late expression than *p*_{7.5} (Wyatt et al., 1996). Strong natural late promoters include *p*₁₁ (Bertholet et al., 1985) and *p*_{CAE} (Patel et al., 1988). About a two-fold increase in expression may be achieved with the synthetic late promoter in pMJ601 and pMJ602 (Davison and Moss, 1990) or the synthetic early/late compound promoter in pSC59 and pSC65 (Chakrabarti et al., 1997; Table 16.17.1).

The ability to synthesize many different kinds of proteins, including those with transmembrane domains, is one advantage of the vaccinia virus expression system. Nevertheless, very high expression of certain genes might adversely affect virus replication. If difficulty is experienced in obtaining recombinant vaccinia with strong promoters, the weaker *p*_{7.5} or *p*_{H5} promoter or the inducible vaccinia virus/bacteriophage T7 hybrid system should be tried (UNIT 16.19).

Several different methods for selecting recombinant viruses are available. The widely used TK selection method is based on the insertional inactivation of the thymidine kinase gene (UNIT 9.5). In the presence of active TK, added BrdU is phosphorylated and incorpo-

rated into viral DNA, where it causes lethal mutations. If TK⁻ cells are used, then TK⁻ virus will replicate normally in the presence of BrdU, whereas TK⁺ virus will not. Because the product of a single cross-over event is still TK⁺, only double cross-over events are selected. However, not all TK⁻ viral plaques will be recombinants because spontaneous TK⁻ mutants arise at a frequency of 1:10,000. Depending on the transfection efficiency, recombinants may comprise 10% to >90% of the TK⁻ plaques.

β-galactosidase or β-glucuronidase (GUS) screening is based on the coinserion of the *E. coli lacZ* (UNIT 1.4) or *GUS* gene, under the control of a vaccinia virus promoter, into the vaccinia virus genome along with the gene of interest. If the TK gene is insertionally inactivated by such an event, recombinant viruses will be TK⁻ and will make blue plaques on medium containing Xgal or Xgluc. This combination of color screening and TK selection will discriminate TK⁻ recombinants from spontaneous TK⁻ mutants. An example of an insertion vector that provides both TK⁻ and β-galactosidase screening is shown in Figure 16.17.2.

XGPRT selection uses mycophenolic acid (MPA), an inhibitor of purine metabolism (see UNIT 9.5). Because MPA blocks the pathway for GMP synthesis, it interferes with the replication of vaccinia and severely reduces the size of vaccinia plaques. This effect can be overcome, however, by expression of the *E. coli gpt* gene in the presence of xanthine and hypoxanthine (i.e., XGPRT can use either of these as substrate for synthesis of GMP). Thus, coexpression of XGPRT provides a useful selection system. Usually the *XGPRT* gene is placed adjacent to the gene of interest and within the vaccinia virus DNA flanking sequences. Unlike the TK⁻ situation, the viral products arising from both single and double cross-over events will be selected. Therefore, it is important to do successive plaque isolations so that the single cross-over events will be resolved. A reverse-selection procedure can be used to delete the *XGPRT* gene from a recombinant vaccinia virus or replace it with another gene (Isaacs et al., 1990). Another procedure, transient XGPRT selection, is performed using a transfer vector with the *XGPRT* gene outside of the vaccinia virus DNA flanking sequences (Falkner and Moss, 1990). Under these conditions, only the single cross-over recombinant virus will express XGPRT, providing substantial enrichment over the parental virus. However, when MPA selection is removed, the desired double cross-over recombinant virus without the

XGPRT gene will have to be differentiated from the parental virus by PCR, immunostaining, or other methods.

Plaque selection (Blasco and Moss, 1995) provides an extremely simple method of selecting recombinant vaccinia viruses that involves neither drugs nor reporter genes. The parental virus (vRB12) contains a deletion of most of the F13L gene, which prevents vaccinia virus from making normal-size plaques. The transfer vector pRB21 contains a segment of the F13L gene adjacent to the gene of interest, so that recombinant viruses will have a functional F13L gene and will make normal-size plaques that are easily distinguished from the pin-point plaques of vRB12. Therefore, one merely needs to pick large plaques to isolate recombinant viruses. However, it is important to check that the virus is a double cross-over recombinant.

To avoid further impairment of MVA, foreign genes have been targeted to existing deletion sites. Deletion II encompasses parts of the *HindIII* N-K fragments; deletion III lies within the *HindIII* A fragment (Meyer et al., 1991). The TK gene has been successfully used as an insertion site for MVA (Carroll and Moss, 1997; T. Blanchard, pers. comm.), although there have been reports of some difficulties in isolating stable TK⁻ MVA (Scheiflinger et al., 1996). Since there are no MVA-permissive TK⁻ cell lines, the main attraction for using the TK site is the large number of available transfer vectors that target this site. When some foreign genes were placed under the control of the very strong synthetic early/late promoter, the resulting recombinant MVA were difficult to isolate or unstable (Wyatt et al., 1996). To date, the examples have been membrane proteins. In most cases, stable recombinants could be made by using a promoter of lower strength, e.g., *p*_{H5} or *p*_{7.5}.

Critical Parameters

During purification of vaccinia virus (see Support Protocol 1) it is important to check the extract after Dounce homogenization to be sure that cells are broken; if they are not, repeat with more force for ten more strokes and check again. In addition, be sure to perform all steps on ice (this is particularly important when sonicating). Use sterile technique throughout.

Calcium phosphate transfections are described in detail in UNIT 9.1. As discussed there, obtaining a very fine precipitate of calcium phosphate and DNA is critical for efficient transfection. Several factors can influence the quality of the precipitate. First, the pH of the

transfection buffer should be between 7.0 and 7.1. Second, if wild-type vaccinia DNA is used, it should be vortexed vigorously to shear the high-molecular-weight DNA prior to precipitation. Although coprecipitation of viral DNA with the insertion vector increases the efficiency of recombination, it is not required if a selection procedure is used. Third, the CaCl_2 should be added slowly and mixed gently (not vortexed), only until full mixing is achieved. The tube should then be left undisturbed until the solution is layered on the cells. Cationic lipids are simpler to use and provide similar or better transfection efficiencies. Although the DOTAP procedure was described in this unit, other cationic lipids are also effective.

During plaque purification and amplification of a new recombinant virus, it is important to maintain the appropriate selective pressure to prevent growth of any contaminating wild-type virus. After isolation, selection is not required. Check the purity of a recombinant virus by PCR or Southern blot analysis of a *Hind*III digest of the virus; the fragment into which the foreign gene has been cloned (the 5.1-kb *Hind*III J fragment for TK selection) should have increased by the size of the insert.

Freezing vaccinia stocks causes clumping of virus particles; thus, stocks should be trypsinized and/or sonicated after thawing. This is particularly important when plaque-purifying a virus, as each plaque should have arisen from a single virus. Trypsinization has been avoided with MVA because preliminary experiments suggested a loss in titer. Some lots of agarose contain contaminants that are toxic to cells (the authors have used low-melting-point agarose from Life Technologies successfully).

It is prudent to save half of each stock at each step in case of contamination or failure at a succeeding step. The authors strongly recommend preparation of a seed stock of the final recombinant virus that will be adequate for future needs. A common mistake is to continually passage the virus.

If immunostaining is used instead of selection or color screening, the antibody must give a good signal to initially pick out the few small foci of cells infected with recombinant virus among a great excess of cells infected with the parental nonstaining virus. Before using an antibody for isolation of recombinant MVA, it can be tested by staining the MVA-infected cells after the first transfection step.

See *UNIT 16.15* for safety precautions that need to be taken when working with vaccinia virus.

Anticipated Results

Approximately $1-5 \times 10^{10}$ pfu of purified virus should be obtained per liter of 5×10^8 HeLa cells.

Depending on the efficiency of the transfection, single, well-isolated plaques should be visible in cells infected with one of the recommended virus dilutions. With TK selection, from 10% to 90% of the plaques will contain recombinant virus. If β -galactosidase or GUS screening is also used, only blue recombinant virus plaques should be picked. With XGPRT selection, all plaques picked should contain recombinant virus. If the titer of recombinant virus is low, amplification can be achieved by a round of growth in the presence of MPA prior to plaquing (use the procedure described in amplification of a plaque).

Cytopathic effects should be clearly visible at each step of amplification except with final infection (infected HeLa cells do not exhibit clear cytopathic effects). The titer of the final crude stock should be $1-2 \times 10^9$ pfu/ml.

Time Considerations

During purification of vaccinia virus, the viral amplification takes 2 to 3 days. After harvesting the infected cells, the entire purification can be done in 1 day. However, the protocol can be stopped after resuspension at step 12 or after collecting the band at step 16. The virus can be stored overnight at 4°C or at -70°C for longer periods.

The infection/transfection procedure for generation of recombinant virus takes ~ 6 hr and cells are harvested after 2 days.

Each round of selection and plaque purification takes 3 days, although little working time is required. Plaques can be picked and reinfections performed on the same day.

Amplification of a single plaque isolate to a small high-titer crude stock will take ~ 7 days. Large stocks should then be prepared as described in *UNIT 16.17*.

Transfection of cells with the plasmid transfer vector requires ~ 2 hr to complete, and the cells are harvested after 2 days. Each round of isolation and plaque purification requires about 30 min for infection, 2 days for incubation, 3 hr for color or immunostaining, and 2 hr for isolating new recombinants and freeze-thaw cycling. Recombinant isolates can be picked and reinfections performed on the same day. From completion of plaque purification to the preparation and titering of a virus stock takes another 10 days.

Literature Cited

- Bacik, I., Cox, J.H., Anderson, R., Yewdell, J.W., and Bennink, J.R. 1994. TAP (transporter associated with antigen processing)-independent presentation of endogenously synthesized peptides is enhanced by endoplasmic reticulum insertion sequences located at the amino- but not carboxyl-terminus of the peptide. *J. Immunol.* 152:381-387.
- Bertholet, C., Drillien, R., and Wittek, R. 1985. One hundred base pairs of 5' flanking sequence of a vaccinia virus late gene are sufficient to temporally regulate late transcription. *Proc. Natl. Acad. Sci. U.S.A.* 82:2096-2100.
- Blasco, R. and Moss, B. 1995. Selection of recombinant vaccinia viruses on the basis of plaque formation. *Gene* 158:157-162.
- Carroll, M.W. and Moss, B. 1995. *E. coli* β -glucuronidase (GUS) as a marker for recombinant vaccinia viruses. *BioTechniques* 19:352-355.
- Carroll, M.W. and Moss, B. 1997. Host range and cytopathogenicity of the highly attenuated MVA strain of vaccinia virus: Propagation and generation of recombinant viruses in a nonhuman mammalian cell line. *Virology* 238:198-211.
- Chakrabarti, S., Brechling, K., and Moss, B. 1985. Vaccinia virus expression vector: Coexpression of beta-galactosidase provides visual screening of recombinant virus plaques. *Mol. Cell. Biol.* 5:3403-3409.
- Chakrabarti, S., Sisler, J.R., and Moss, B. 1997. Compact, synthetic, vaccinia virus early/late promoter for protein expression. *BioTechniques* 23:1094-1097.
- Cochran, M.A., Puckett, C., and Moss, B. 1985. In vitro mutagenesis of the promoter region for a vaccinia virus gene: Evidence for tandem early and late regulatory signals. *J. Virol.* 54:30-37.
- Davison, A.J. and Moss, B. 1990. New vaccinia virus recombination plasmids incorporating a synthetic late promoter for high level expression of foreign proteins. *Nucl. Acids Res.* 18:4285-4286.
- Earl, P., Koenig, S., and Moss, B. 1990. Biological and immunological properties of human immunodeficiency virus type 1 envelope glycoprotein: Analysis of proteins with truncations and deletions expressed by recombinant vaccinia viruses. *J. Virol.* 65:31-41.
- Falkner, F.G. and Moss, B. 1988. *Escherichia coli gpt* gene provides dominant selection for vaccinia virus open reading frame expression vectors. *J. Virol.* 62:1849-1854.
- Falkner, F.G. and Moss, B. 1990. Transient dominant selection of recombinant vaccinia viruses. *J. Virol.* 64:3108-3111.
- Isaacs, S.N., Kotwal, G.J., and Moss, B. 1990. Reverse guanine phosphoribosyltransferase selection of recombinant vaccinia viruses. *Virology* 178:626-630.
- Mackett, M., Smith, G.L., and Moss, B. 1984. General method for production and selection of infectious vaccinia virus recombinants expressing foreign genes. *J. Virol.* 49:857-864.
- Meyer, H., Sutter, G., and Mayr, A. 1991. Mapping of deletions in the genome of the highly attenuated vaccinia virus MVA and their influence on virulence. *J. Gen. Virol.* 72:1031-1038.
- Merchinsky, M., Eckert, D., Smith, E., and Zauderer, M. 1997. Construction and characterization of vaccinia direct ligation vectors. *Virology* 238:444-451.
- Patel, D.D., Ray, C.A., Drucker, R.P., and Pickup, D.J. 1988. A poxvirus-derived vector that directs high levels of expression of cloned genes in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 85:9431-9435.
- Pfleiderer, M., Falkner, F.G., and Dorner, F. 1995. A novel vaccinia virus expression system allowing construction of recombinants without the need for selection markers, plasmids and bacterial hosts. *J. Gen. Virol.* 76:2957-2962.
- Scheiflinger, F., Falkner, F.G., and Dorner, F. 1996. Evaluation of the thymidine kinase (tk) locus as an insertion site in the highly attenuated vaccinia MVA strain. *Arch. Virol.* 141:663-669.
- Sutter, G., Wyatt, L.S., Foley, P.L., Bennink, J.R., and Moss, B. 1994. A recombinant vector derived from the host range-restricted and highly attenuated MVA strain of vaccinia virus stimulates protective immunity in mice to influenza virus. *Vaccine* 12:1032-1040.
- Wyatt, L.S., Shors, S.T., Murphy, B.R., and Moss, B. 1996. Development of a replication-deficient recombinant vaccinia virus vaccine effective against parainfluenza virus 3 infection in an animal model. *Vaccine* 14:1451-1458.

Key References

Mackett et al., 1984. See above.

Piccini, A., Perkus, M.E., and Paoletti, E. 1987. Vaccinia virus as an expression vector. *Methods Enzymol.* 153:545-563.

Contributed by Patricia L. Earl,
Bernard Moss, and Linda S. Wyatt
National Institute of Allergy and
Infectious Diseases
Bethesda, Maryland

Miles W. Carroll
Oxford BioMedica
Oxford, United Kingdom

Characterization of Recombinant Vaccinia Viruses and Their Products

UNIT 16.18

After a recombinant vaccinia virus is made (UNIT 16.17), its DNA and protein products can be analyzed in several ways. Identification of the recombinant virus can be carried out by PCR (see Basic Protocol 1), with verification of correct insertion of the DNA by Southern blotting (see Basic Protocol 2). Vaccinia DNA can also be detected by dot-blot hybridization (see Basic Protocol 3). Finally, when antibodies are available, protein expression can be analyzed by immunological methods such as dot blotting with an antibody (see Alternate Protocol) or immunoblotting (see Basic Protocol 4) and/or immunoprecipitation (see Basic Protocol 5). In addition, immunostaining can be used for identification of recombinant plaques as well as for determination of the purity of a recombinant virus stock. For these purposes, the support protocol for immunostaining in UNIT 16.16 should be used; however an antibody to the foreign protein should be substituted for anti-vaccinia antibody.

All of the protocols in this unit can be used for characterization of modified vaccinia virus Ankara (MVA) recombinant viruses (UNIT 16.17).

CAUTION: Proceed carefully and follow biosafety level 2 (BL-2) practices when working with vaccinia virus (see UNIT 16.15 for safety guidelines).

NOTE: Carry out all procedures for growth of vaccinia virus using sterile technique in a biosafety cabinet.

DETECTION OF VACCINIA DNA USING PCR

The presence of recombinant virus in isolated plaques can be determined by infecting cells with the virus, extracting the DNA, and performing PCR analysis utilizing oligonucleotide primers in the flanking vaccinia DNA. Although PCR analysis can be done directly on DNA in picked virus plaques (Zhang and Moss, 1991), individual researchers sometimes have problems with this shorter procedure. PCR is also useful for determining the purity of a viral stock, as both wild-type and recombinant DNAs are amplified using this protocol. Here a method is described for isolation of DNA from infected cells. An alternative to this protocol is to use a commercial kit such as Qiagen QIAamp Blood Kit for extraction of DNA.

Materials

- Confluent BS-C-1 or HuTK⁻ 143B cell monolayer (UNIT 16.16)
- Phosphate-buffered saline (PBS; APPENDIX 2)
- Trypsin/EDTA (0.25%:0.02%), 37°C
- Complete MEM-10, -2.5, and -5 media (UNIT 16.16)
- Complete MEM-10/BrdU medium (UNIT 16.16)
- Recombinant virus plaques, picked and resuspended in tubes (UNIT 16.17)
- Selective agents (filter sterilize and store at -20°C; also see UNIT 16.17):
 - 10 mg/ml (400×) mycophenolic acid (MPA; Calbiochem) in 0.1 N NaOH (for XGPRT selection)
 - 10 mg/ml (40×) xanthine in 0.1 N NaOH (for XGPRT selection)
 - 10 mg/ml (670×) hypoxanthine in 0.1 N NaOH (for XGPRT selection)
 - 5 mg/ml (200×) 5-bromodeoxyuridine (BrdU) in water (for TK selection)
- Dry ice/ethanol bath
- DNA extraction buffer (see recipe)
- 3 M sodium acetate, pH 6.0 (APPENDIX 2)

**BASIC
PROTOCOL 1**

**Protein
Expression**

95% and 70% ethanol, ice-cold
10× MgCl₂-free PCR amplification buffer (*UNIT 15.1*)
10 mM 4dNTP mix (*UNIT 15.1*)
100 μM primer (sequence depends on where foreign gene is inserted; see *UNIT 2.11*
for oligonucleotide synthesis strategies)
15 mM MgCl₂
24-well tissue culture plates
Cup sonicator
Additional reagents and equipment for tissue culture and counting cells
(*APPENDIX 3F*), phenol/chloroform extraction of DNA, PCR amplification
(*UNIT 15.1*), and agarose gel electrophoresis (*UNIT 2.5*)

NOTE: All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

Set up cells for infection

1. Aspirate medium from a confluent cell monolayer.
 - a. For XGPRT selection, use BS-C-1 cells.
 - b. For TK selection, use HuTK⁻ 143B cells.
 2. Wash cells once with 37°C PBS or trypsin/EDTA to remove the remaining serum.
 3. Overlay cells with 37°C trypsin/EDTA using a volume that is just enough to cover the monolayer (1.5 ml for a 150-cm² flask). Allow to sit ~30 sec (the cells should become detached). Shake to detach cells completely.
 4. Add 8.5 ml of the appropriate medium.
 - a. For BS-C-1 cells, use complete MEM-10.
 - b. For HuTK⁻ 143B cells, use complete MEM-10/BrdU.
- Pipet cell suspension up and down several times to disrupt clumps.
5. Count cells using a hemacytometer (*APPENDIX 3F*).
 6. Plate 1.25×10^5 cells/well in 24-well tissue culture dishes (0.5 ml/well final). Incubate until confluent (this should take <24 hr).

Infect cells with virus

7. Place tube containing virus into a cup sonicator containing an ice/water mixture and sonicate 20 to 30 sec at full power.
8. Aspirate medium from each confluent well and infect each well with one recombinant virus plaque, using half the volume in which each plaque is suspended (i.e., 0.25 ml where 0.5 ml is used to resuspend the plaque). Incubate 1 to 2 hr, rocking at 15- to 30-min intervals.

Perform selection

9. Overlay with 1 ml complete MEM-2.5 to which appropriate selection drugs have been added. Incubate until cytopathic effect (cell rounding) is evident (usually 24 to 48 hr).
 - a. For XGPRT selection, add $\frac{1}{400}$ vol of 10 mg/ml MPA, $\frac{1}{40}$ vol of 10 mg/ml xanthine, and $\frac{1}{670}$ vol of 10 mg/ml hypoxanthine.
 - b. For TK selection, add $\frac{1}{200}$ vol of 5 mg/ml BrdU.

10. Scrape cells and transfer to a microcentrifuge tube. Microcentrifuge 30 sec at maximum speed. Aspirate medium, add 1 ml PBS, vortex to resuspend, then microcentrifuge again, 30 sec at maximum speed. Aspirate PBS, then resuspend pellet in 0.5 ml PBS.

Cells in a 24-well plate can be scraped using the plunger of a 1-ml syringe.

Prepare DNA for PCR amplification

11. Lyse the cell suspension by carrying out three freeze-thaw cycles, each time by freezing in a dry ice/ethanol bath, thawing in a 37°C water bath, and vortexing.
12. Place tube containing lysate into a cup sonicator containing an ice/water mixture and sonicate 20 to 30 sec at full power.
13. Place 50 to 100 µl of the lysate into a microcentrifuge tube. Add 1/10 vol of DNA extraction buffer. Mix by gently inverting the tube several times. Incubate 2 hr to overnight at 37°C.
14. Extract once with an equal volume of buffered phenol and once with an equal volume of chloroform (*UNIT 2.1*).
15. Add 1/10 vol of 3 M sodium acetate, pH 6.0 (0.3 M final) and 2.5 vol of 95% ethanol. Place at -20°C for 20 min.
16. Microcentrifuge 15 min at maximum speed, 4°C. Aspirate and discard supernatant. Wash pellet with ice-cold 70% ethanol, then air dry.
17. Dissolve pellet in 20 µl deionized water. Prepare the following reaction mix on ice:

10 µl redissolved (template) DNA pellet
5 µl 10× MgCl₂-free PCR amplification buffer
5 µl 10 mM 4dNTP mix
0.5 µl 100 µM primer
5 µl 15 mM MgCl₂
24.5 µl H₂O

Overlay with mineral oil.

A PCR kit such as the Expand High Fidelity PCR System (Boehringer Mannheim) is appropriate for this step.

18. Carry out PCR (*UNIT 15.1*) using the following thermal cycling parameters:

35 cycles:	20 sec	95°C	(denaturation)
	20 sec	55° to 60°C	(annealing)
	1 to 2 min	72°C	(extension)
1 cycle:	5 min	72°C	(extension)
Final step:	indefinitely	4°C	(hold).

These conditions are for a 1-kb fragment.

19. Analyze PCR product on agarose gel (*UNIT 2.5*).

DETECTION OF VACCINIA DNA USING SOUTHERN BLOT HYBRIDIZATION

Historically, vaccinia virus DNA has been characterized by its *Hind*III restriction endonuclease digestion pattern. The TK gene, into which most foreign genes are inserted in recombinant viruses, is located in the 5.1-kbp *Hind*III J fragment. Thus in a recombinant virus, if the inserted gene has no *Hind*III sites, this fragment should be increased in size. The 5.1-kbp *Hind*III J fragment should be absent, indicating lack of contamination with wild-type vaccinia virus. Southern blot analysis can be done with DNA from purified virus (UNIT 16.17) or, more easily, from a crude lysate of infected cells as described below.

Materials

Confluent BS-C-1 cell monolayer (UNIT 16.16)
Complete MEM-10 and -5 media (UNIT 16.16)
Recombinant vaccinia virus stock (UNIT 16.16)
0.25 mg/ml trypsin (2× crystallized and salt-free; Worthington; filter sterilize and store at -20°C)
Phosphate-buffered saline (PBS; APPENDIX 2)
Low-salt buffer (see recipe)
20 mg/ml proteinase K
Buffered phenol (UNIT 2.1)
1:1 phenol/chloroform
3 M sodium acetate, pH 6.0 (APPENDIX 2)
95% and 70% ethanol, ice-cold
TE buffer, pH 7.8 (APPENDIX 2)
*Hind*III restriction endonuclease (UNIT 3.1) and appropriate buffer
12-well tissue culture plates

Additional reagents and equipment for preparing BS-C-1 cells for vaccinia infection (see Basic Protocol 1), tissue culture and counting cells (APPENDIX 3F), phenol/chloroform extraction of DNA (UNIT 2.1), restriction endonuclease digestion of DNA (UNIT 3.1), Southern blotting (UNIT 2.9A), and hybridization (UNIT 2.10)

NOTE: All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

Prepare cell cultures and infect cells

1. Set up cell cultures for infection (see Basic Protocol 1, steps 1 to 5) starting with a confluent BS-C-1 cell monolayer.
2. Plate 2.5×10^5 cells/well in 12-well tissue culture plates (1 ml/well final) using complete MEM-10 medium. Incubate until confluent (this should take <24 hr).
3. Just prior to use, mix equal volumes of recombinant vaccinia virus stock and 0.25 mg/ml trypsin and vortex vigorously. Incubate 30 min in a 37°C water bath, vortexing at 5- to 10-min intervals during the incubation.

Virus stocks are usually at a titer of $\sim 2 \times 10^9$ pfu/ml but may be significantly lower depending on the source.

Vortexing usually breaks up any clumps. However, if clumps still remain, chill to 0°C and sonicate 30 sec on ice. Sonication can be repeated several times but the sample should be allowed to cool on ice between sonications.

4. Aspirate medium and infect monolayer culture at a multiplicity of infection (MOI) of 10 to 20 pfu/cell in 250 μ l complete MEM-5. Incubate 1 to 2 hr, rocking at ~30-min intervals.
5. Overlay with 1 ml complete MEM-5 and incubate ~24 hr.

Harvest cells

6. Scrape cells and transfer to a microcentrifuge tube. Microcentrifuge 30 sec at maximum speed, room temperature, then aspirate and discard supernatant.

The plunger of a 1-ml syringe can be used to scrape cells.

7. Resuspend pellet in 50 μ l PBS and vortex.
8. Add 300 μ l of low-salt buffer and 10 μ l of 20 mg/ml proteinase K to a microcentrifuge tube. Add cell suspension to this same tube. Vortex, then incubate 5 hr to overnight at 37°C.

Extract DNA and analyze by Southern blotting

9. Extract once with an equal volume of buffered phenol, then once with an equal volume of 1:1 phenol/chloroform (UNIT 2.1).

If the solution is viscous after phenol extraction, pass it through a 25-G needle to shear DNA and reduce viscosity.

10. Add $\frac{1}{10}$ vol 3 M sodium acetate, pH 6.0 (0.3 M final) and 2.5 vol 95% ethanol. Place 30 min on dry ice or overnight at -20°C.
11. Microcentrifuge 10 min at maximum speed, 4°C. Aspirate and discard supernatant.
12. Wash pellet with ice-cold 70% ethanol, then air dry.
13. Dissolve pellet in 100 μ l TE buffer, pH 7.8.
14. Digest 15 μ l of DNA solution with *Hind*III (UNIT 3.1) and analyze by Southern blotting (UNIT 2.9A) and hybridization (UNIT 2.10).

DETECTION OF VACCINIA DNA USING DOT-BLOT HYBRIDIZATION

In addition to PCR and Southern blotting, DNA dot-blot hybridization, as described in this protocol, may be used for identifying recombinant virus in isolated plaques.

Materials

0.5 N NaOH
1 M Tris·Cl, pH 7.5 (APPENDIX 2)
2 \times SSC (APPENDIX 2)

Nitrocellulose membrane (see Table 2.9.1)
Whatman 3MM filter paper
80°C vacuum oven

Additional reagents and equipment for preparing cells, infecting with vaccinia virus, and performing selection (see Basic Protocol 1), dot blotting (UNIT 2.9B), and hybridization (UNIT 2.10)

NOTE: All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

**BASIC
PROTOCOL 3**

**Protein
Expression**

16.18.5

Prepare and infect cells and harvest cell suspension

1. Infect cells with virus and perform selection (see Basic Protocol 1, steps 1 to 9).
2. Scrape cells and transfer to a microcentrifuge tube. Microcentrifuge 30 sec at maximum speed, then aspirate and discard medium. Resuspend cell pellet in 200 μ l PBS.

Cells in a 24-well plate can be scraped using the plunger of a 1-ml syringe.
3. Freeze thaw and sonicate cell suspension (see Basic Protocol 1, steps 11 and 12).

Perform blotting procedure

4. Cut nitrocellulose membrane to the appropriate size to fit in the dot-blot apparatus. Wet the membrane by immersing it in water and place in the apparatus. Add 20 to 100 μ l cell lysate (from step 3) to each well. Draw the liquid through the membrane by applying suction, then remove the nitrocellulose from the apparatus.
5. Cut six pieces of Whatman 3MM filter paper so they are larger than the nitrocellulose membrane. Soak one piece of 3MM paper with 0.5 N NaOH. Lay the nitrocellulose membrane on top of the wet 3MM paper and let sit for 1 min. Remove membrane and place it on a separate piece of dry 3MM paper for 1 min. Repeat this procedure two times using the same piece of 3MM paper.

Forceps should be used to transfer the membrane. See UNIT 2.1B for additional discussion of dot blotting.
6. Repeat step 5 (three times per membrane) with Whatman 3MM paper soaked in 1 M Tris·Cl, pH 7.5.
7. Repeat step 5 (three times per membrane) with 3MM Whatman paper soaked in 2 \times SSC.
8. Bake membrane 2 hr at 80°C under vacuum.

Baking fixes the DNA to the filter.
9. Proceed with DNA hybridization as described in *UNIT 2.10*.

ALTERNATE PROTOCOL

DETECTION OF EXPRESSED PROTEIN BY A DOT-BLOT PROCEDURE

This protocol is an alternative to DNA dot-blot hybridization for checking plaques to identify recombinant virus. Infected cell lysates are spotted onto a nitrocellulose membrane as in DNA dot-blotting (see Basic Protocol 3). The membrane is then incubated with an antibody that recognizes the expressed foreign gene product. ¹²⁵I-labeled protein A is used to detect the bound antibody. Alternatively, a chemiluminescent detection system such as Amersham ECL can be used. The membrane is then exposed to X-ray film.

Note that infected cells comprising a plaque can be stained in situ by following the protocol in *UNIT 16.16* (Basic Protocol 4, titration of MVA by immunostaining) except that antibody to the recombinant protein should be substituted for antibody to vaccinia virus.

Additional Materials (also see *Basic Protocol 3*)

- PBS/Tween (see recipe) with and without 4% (w/v) BSA
- Antibody to foreign protein
- 100 μ Ci/ml [¹²⁵I]-labeled protein A (30 mCi/mg; Amersham)
- Additional reagents and equipment for autoradiography (*APPENDIX 3A*)

1. Prepare cells and blot cell lysate (see Basic Protocol 3, steps 1 to 4).
2. Soak membrane ~30 min in PBS/Tween containing 4% BSA. Wash once with PBS/Tween without BSA.

Washing is done by pouring liquid off and replacing it.

3. Dilute antibody to foreign protein 1:50 to 1:5000 in a minimal volume of PBS/Tween. Incubate membrane in this solution ≥ 1 hr at room temperature.
4. Wash approximately five times with an excess volume of PBS/Tween.
5. Incubate membrane for 30 min in a minimal volume of PBS/Tween containing 1 μ Ci of [125 I]-labeled protein A.

Alternatively, a chemiluminescent detection kit (e.g., Amersham ECL) may be used to detect the antibody

6. Wash membrane approximately five times with an excess volume of PBS/Tween.
7. Wrap membrane in plastic wrap and autoradiograph by exposing to X-ray film (APPENDIX 3A).

DETECTION OF EXPRESSED PROTEIN USING IMMUNOBLOTTING

After a recombinant virus has been plaque purified and amplified, immunoblotting can be used to determine if the recombinant protein is of the expected size and whether or not it is secreted from infected cells. The steps below are for nonsecreted (intracellular) proteins, but annotations describe procedures when working with cells in which the recombinant protein is secreted into the medium.

Materials

- Confluent BS-C-1 cell monolayer (UNIT 16.16)
- Complete MEM-5 medium (UNIT 16.16)
- Recombinant vaccinia virus stock (UNIT 16.16)
- 0.25 mg/ml trypsin (2 \times crystallized and salt-free; Worthington; filter sterilize and store at -20°C)
- Cell lysis buffer (see recipe)
- 5 \times SDS sample buffer (UNIT 10.2)
- 6-well tissue culture plates
- Sorvall centrifuge with H-6000A rotor (or equivalent)
- 95 $^{\circ}\text{C}$ water bath
- Additional reagents and equipment for preparing cell cultures for infection (see Basic Protocol 1) and immunoblotting (UNIT 10.8)

NOTE: All culture incubations should be performed in a humidified 37 $^{\circ}\text{C}$, 5% CO_2 incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

1. Set up cell cultures for infection (see Basic Protocol 1, steps 1 to 5) starting with a confluent BS-C-1 cell monolayer.
2. Plate 5×10^5 cells/well in a 6-well tissue culture plate (2 ml/well final) using MEM-5 medium. Incubate until confluent (this should take <24 hr).
3. Just prior to use, mix equal volumes of recombinant vaccinia virus stock and 0.25 mg/ml trypsin and vortex vigorously. Incubate 30 min in a 37 $^{\circ}\text{C}$ water bath, vortexing at 5- to 10-min intervals during the incubation.

BASIC PROTOCOL 4

Protein Expression

16.18.7

Virus stocks are usually at a titer of $\sim 2 \times 10^9$ pfu/ml but may be significantly lower depending on the source.

Vortexing usually breaks up any clumps. However, if clumps still remain, chill to 0°C and sonicate 30 sec on ice. Sonication can be repeated several times but the sample should be allowed to cool on ice between sonications.

4. Aspirate medium and infect monolayer culture at a multiplicity of infection (MOI) of 10 to 30 pfu/cell in 1 ml (final volume) complete MEM-5. Incubate 1 to 2 hr, rocking at ~ 30 -min intervals.

If purified virus is used, sonicate 30 sec on ice prior to use (see Basic Protocol 1, step 7).

If analyzing proteins that are secreted into the medium, it is important to infect and overlay the cells with medium that does not require the addition of serum or to use only 1% serum, since a high concentration of serum proteins interferes with gel analysis.

5. Overlay with 2 ml complete MEM-5 and incubate 24 to 48 hr.
6. Scrape cells and transfer to a centrifuge tube. Centrifuge 5 min at $1800 \times g$ (2500 rpm in a H-6000A rotor), 5° to 10°C. Aspirate and discard supernatant.

For secretory proteins, concentrate the medium by spinning in a Centricon 10 microconcentrator (Amicon) until the volume is reduced ~ 5 -fold. Alternatively, proteins can be precipitated from the medium by trichloroacetic acid (TCA) precipitation as follows: (1) Add Triton X-100 to the sample to 0.03% (v/v) final; (2) add an equal volume of 20% TCA; (3) place 15 min on ice; (4) microcentrifuge 15 min at maximum speed, 4°C; (5) aspirate supernatant and wash the pellet with 70% ethanol, (6) air dry the pellet, add 5 μ l 1 M Tris base, and dissolve in 1 \times SDS/sample buffer.

7. Resuspend cell pellet in 200 μ l cell lysis buffer and transfer to a microcentrifuge tube. Vortex and let sit 10 min on ice.

Cell lysis buffer contains nonionic detergent only. This may not be sufficient for extraction of some proteins. In such cases, SDS-containing buffer (UNIT 16.11) may be necessary.

8. Microcentrifuge 10 min at maximum speed, 4°C. Separate supernatant (cytoplasm) and pellet (nuclei).
9. To 20 μ l of the supernatant, add 5 μ l of 5 \times SDS/sample buffer and heat 5 min at 95°C. Dissolve pellet in 200 μ l of 1 \times SDS/sample buffer and heat 5 min at 95°C.
10. Proceed with immunoblotting (UNIT 10.8).

BASIC PROTOCOL 5

DETECTION OF EXPRESSED PROTEIN USING IMMUNOPRECIPITATION

When strong late promoters are used, the expressed recombinant protein can be labeled with radioactive amino acids and detected by polyacrylamide gel electrophoresis. Immunoprecipitation of the labeled proteins can be used to increase sensitivity and specificity when early or late promoters are used.

Materials

Complete MEM-5 medium (UNIT 16.16)

Complete methionine- or cysteine-free MEM-5 medium (see recipe in UNIT 16.16 but use methionine- or cysteine-free MEM, e.g., from Life Technologies) prepared with dialyzed FBS (e.g., HyClone)

[³⁵S]methionine or [³⁵S]cysteine (depending on which amino acid was omitted from the medium; see Critical Parameters)

Phosphate-buffered saline (PBS; APPENDIX 2)

Cell lysis buffer (see recipe)

Additional reagents and equipment for preparing cells and infecting with vaccinia virus (see Basic Protocol 4) and immunoprecipitation (UNIT 10.16)

NOTE: All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

1. Set up cell cultures and infect cells (see Basic Protocol 4, steps 1 to 4).
2. Overlay with 2 ml complete MEM-5 medium and continue incubation for the appropriate period of time.

The time after infection at which metabolic labeling is performed will depend upon which promoter is used to express the gene of interest. If an early promoter only is used, labeling should be done between 1 and 6 hr post infection. If a late promoter is used, labeling should be done between 4 and 20 hr post infection. With a compound early/late promoter, either time may be used.

3. At the appropriate time after infection, aspirate medium and replace with 0.5 ml complete methionine- or cysteine-free MEM-5 containing 25 to 50 µCi of [³⁵S]methionine or [³⁵S]cysteine and 35 µl complete MEM-5. Incubate 2 to 3 hr to overnight.
4. Add 0.15 ml complete MEM-5 as chase. Continue incubation for 1 hr.
5. Remove medium and save if the expressed protein is secreted.
If the medium is to be analyzed, remove any free cells by microcentrifuging 3 min.
6. Overlay cells with 1 ml PBS, scrape, and transfer to a microcentrifuge tube. Microcentrifuge 3 min at maximum speed, room temperature. Aspirate and discard supernatant.
7. Resuspend cells in 200 µl cell lysis buffer. Vortex and place 10 min on ice.
8. Microcentrifuge 10 min at maximum speed, 4°C. Remove supernatant and proceed with immunoprecipitation (beginning with Basic Protocol, step 4, in UNIT 10.16).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Cell lysis buffer

100 mM Tris·Cl, pH 8.0 (APPENDIX 2)
100 mM NaCl
0.5% (v/v) Triton X-100 or NP-40
Store indefinitely at room temperature

Add PMSF (or other protease inhibitors; optional) from a 100× stock to 0.2 mM final just before use. Store 100× inhibitor stock at -20°C.

DNA extraction buffer

20 mM EDTA
20 mM Tris·Cl, pH 7.5 (APPENDIX 2)
0.5% (w/v) SDS
0.5 to 1.0 mg/ml proteinase K (add immediately before use)
Store components without proteinase K indefinitely at room temperature

Low-salt buffer

20 mM Tris·Cl, pH 8.0 (APPENDIX 2)
10 mM EDTA
0.75% (w/v) sodium dodecyl sulfate (SDS)
Store indefinitely at room temperature

PBS/Tween

Add 0.5% (v/v) Tween-20 to PBS (APPENDIX 2). Store indefinitely at room temperature.

COMMENTARY

Background Information

Several different methods of characterizing recombinant vaccinia viruses are available. Each is suitable at a different time during isolation and characterization of a recombinant virus. During plaque purification of a new virus, PCR, DNA dot-blot hybridization, or immunoblotting of individual plaque isolates are quick and easy ways of determining which plaque contains the foreign gene. This is particularly important if TK selection only is used, since not all plaques will contain recombinant virus.

After an isolate has been plaque purified several times and amplified, analysis of the DNA by PCR or Southern blotting can be used to determine whether or not a pure isolate has been obtained. Restriction endonuclease digestion of vaccinia DNA with *Hind*III gives a specific pattern. The *Hind*III fragment into which the foreign gene has been inserted should be increased in size by the size of the foreign gene (unless the latter contains a *Hind*III site). The TK gene, which is the most commonly used region for insertion of a foreign gene, is located in the 5.1-kbp *Hind*III J fragment.

Western blotting or radioimmunoprecipitation using an antibody specific for the gene of interest will determine whether the correct gene product is made. Analysis of the extracellular medium will show whether the protein is secreted. If a nuclear localization signal is present, the protein may be translocated into the nucleus. In this case, the nuclei should also be analyzed (see Basic Protocol 4, step 8).

For large-scale production of protein, HeLa cell suspension cultures may be infected as described in the support protocol for purification of vaccinia virus in UNIT 16.17, and harvested after 48 to 72 hr. Alternatively, microcarrier cultures of Vero cells may be used (Barrett et al., 1989).

Critical Parameters

In all protocols for identifying plaques containing recombinant virus, it is important to include positive and negative controls. For PCR or DNA dot-blot hybridization, the plasmid used for transfection is an appropriate positive control, while cells infected with wild-type vaccinia virus can serve as a negative control. For immunoblotting, an extract of cells known to express the foreign gene can be used (e.g., influenza virus-infected cells could be used as a positive control for analysis of a recombinant vaccinia virus expressing an influenza gene). Cells infected with wild-type vaccinia virus can be used as a negative control.

In planning an immunoprecipitation, check the amino acid composition of the expressed protein for numbers of methionine and cysteine residues and then label with the appropriate amino acid(s).

If lysates are frozen prior to immunoprecipitation, it is important to microcentrifuge 5 min at maximum speed to pellet subcellular debris before mixing a sample of the supernatant with antibody.

Anticipated Results

The Southern blotting protocol in this unit is a quicker, easier method than the one in UNIT 2.9. Since the most common region used for insertion of foreign genes into vaccinia virus is the TK locus within the 5.1-kbp *Hind*III J fragment, the size of this fragment should be larger in a recombinant virus (if there is no *Hind*III site within the inserted DNA). Hybridization of a second blot or rehybridization of the original blot with ³²P-labeled vaccinia virus DNA will show the overall *Hind*III restriction digestion pattern. The 5.1-kbp *Hind*III J fragment should be absent, indicating that the virus isolate does not contain any contaminating wild-type recombinant virus. A control of wild-type *Hind*III-digested vaccinia DNA is helpful for a

comparison to the original restriction endonuclease pattern.

Visualization of protein bands from an immunoblotting procedure is usually accomplished with an X-ray film exposure of 1 to 24 hr. Additional time may be required if the antibody recognition is poor; if this is the case, cells can be lysed in a smaller volume, allowing more lysate to be analyzed in one lane. Additional incubation time or a lower dilution of antibody may also help to intensify the signal. A western blot that has been exposed to X-ray film but has not been completely dried can be reincubated with the same antibody (if the original antibody concentration was too low) or with a different antibody. Negative and positive controls should be included to help identify specific protein bands.

For an immunoprecipitation, the gel exposure time will depend on several factors, including the amount of label incorporated into the protein and the affinity of the antibody for the protein. It may be necessary to optimize results by adjusting the amount of labeled amino acid used, the amount of antibody used, or the time of incubation of the lysate with the antibody.

Time Considerations

The PCR, DNA dot-blot hybridization, and immunoblotting protocols require 2 days for amplification of the virus. An additional day is necessary for PCR amplification and agarose gel analysis or for hybridization and washing. Time of exposure to X-ray film for DNA dot-blot and immunoblotting protocols varies depending upon the hybridization probe or antibody used.

Literature Cited

- Barrett, N., Mitterer, A., Mundt, W., Eibl, J., Eibl, M., Gallo, R.C., Moss, B., and Dorner, F. 1989. Large-scale production and purification of a vaccinia recombinant-derived HIV-1 gp 160 and analysis of its immunogenicity. *AIDS Res. Hum. Retroviruses* 5:159-171.
- Zhang, Y. and Moss, B. 1991. Inducer-dependent conditional-lethal mutant animal viruses. *Proc. Natl. Acad. Sci. U.S.A.* 88:1511-1515.

Contributed by Patricia L. Earl and
Bernard Moss
National Institute of Allergy and
Infectious Diseases
Bethesda, Maryland

Gene Expression Using the Vaccinia Virus/T7 RNA Polymerase Hybrid System

UNIT 16.19

This unit describes a transient cytoplasmic expression system that relies on the synthesis of the bacteriophage T7 RNA polymerase in the cytoplasm of mammalian cells. To begin, a gene of interest is inserted into a plasmid such that it comes under the control of the T7 RNA polymerase promoter (p_{T7}). Using liposome-mediated transfection, this recombinant plasmid is introduced into the cytoplasm of cells infected with vTF7-3, a recombinant vaccinia virus encoding bacteriophage T7 RNA polymerase (see Basic Protocol 1). During incubation, the gene of interest is transcribed with high efficiency by T7 RNA polymerase. This transfection protocol is adequate for analytical purposes and is simple because no new recombinant viruses need to be made.

For large-scale work, the p_{T7} -regulated gene can be inserted into a second recombinant vaccinia virus by homologous recombination (UNIT 16.17) and used in combination with vTF7-3 to coinfect cells grown in suspension (see Basic Protocol 2) or used to infect OST7-1 cells (a stable cell line that constitutively expresses the T7 RNA polymerase; see Basic Protocol 3). Expressed protein is then analyzed by pulse-labeling (see Support Protocol) or by the methods detailed in UNIT 16.18, and purified as described in UNIT 10.10.

There have been several innovations to this vaccinia virus/T7 RNA polymerase hybrid expression system (see Background Information). One new development is the VOTE inducible expression system, which eliminates the need to use two recombinant viruses or a special cell line (see Basic Protocol 4).

NOTE: All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All reagents and equipment coming into contact with live cells must be sterile, and aseptic technique should be used accordingly.

LIPOSOME-MEDIATED TRANSFECTION FOLLOWING RECOMBINANT VACCINIA VIRUS (vTF7-3) INFECTION

**BASIC
PROTOCOL 1**

This protocol begins with a plasmid vector containing the gene of interest at a site between p_{T7} and the T7 terminator (Figs. 16.19.1 & 16.19.2; also see Critical Parameters). This recombinant plasmid is then introduced via liposome-mediated transfection (UNIT 9.4) into cells already infected with vaccinia virus vTF7-3 (a recombinant virus expressing the T7 RNA polymerase gene). After harvesting, the cells are analyzed for expression of the gene product as described in UNIT 16.18 or by pulse-labeling (see Support Protocol).

Materials

Recombinant plasmid: gene of interest subcloned (UNIT 3.16) into pTF7-5 or pTM1 vectors (available from B. Moss, e-mail bmoss@nih.gov; Figs. 16.19.1 and 16.19.2) or other plasmid containing the T7 promoter (e.g., pBluescript, Stratagene; see UNIT 16.2 Critical Parameters and Fig. 1.10.8)

Confluent CV-1 cell monolayer (ATCC #CCL70; UNIT 16.16)

Complete DMEM-10 (UNIT 16.16)

vTF7-3 vaccinia virus stock (ATCC #VR-2153)

Opti-MEM I reduced serum medium (Life Technologies)

Liposome suspension (Lipofectin or TransfectAce; Life Technologies)

6-well tissue culture dishes with 35-mm-diameter wells

**Protein
Expression**

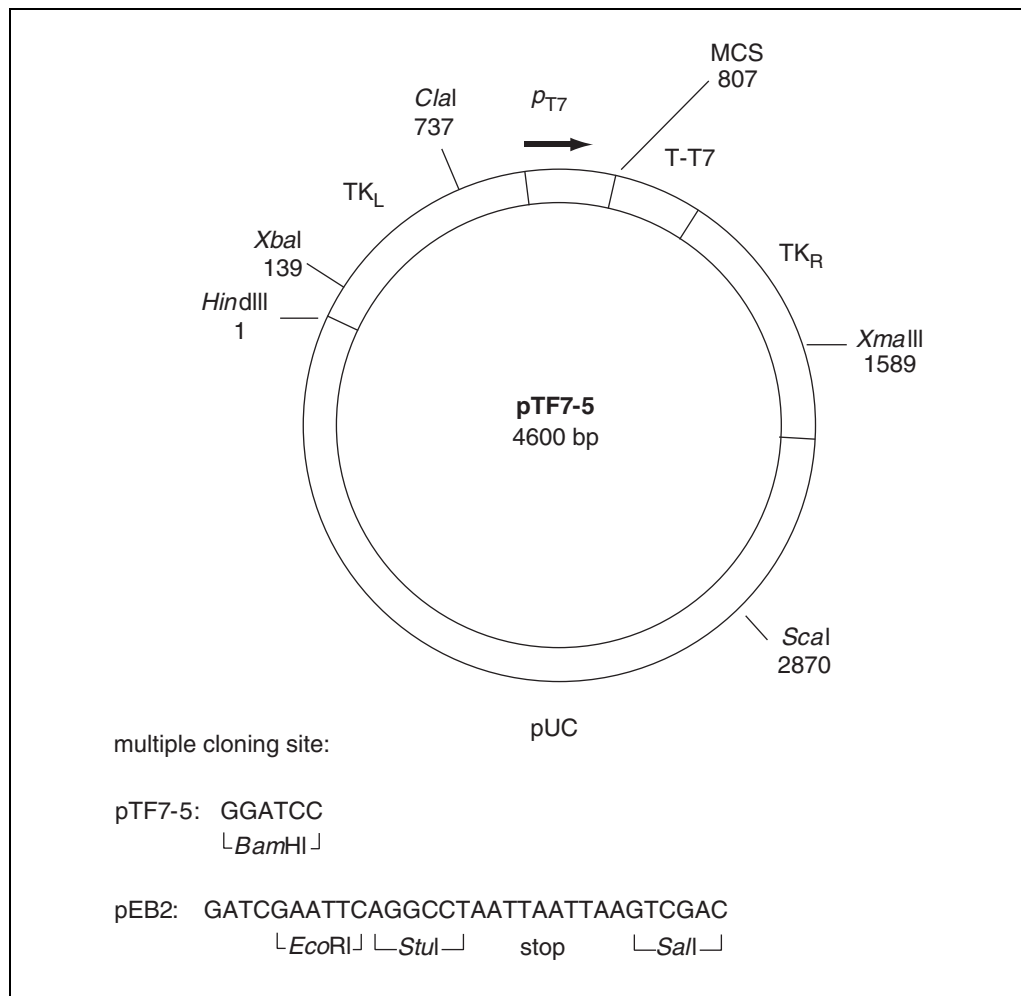


Figure 16.19.1 pTF7-5 (Fuerst et al., 1986) contains p_{T7} and a terminator, between which is a unique *Bam*H1 site for insertion of a gene. The derivative, pEB2 (Berger et al., 1988) contains unique *Eco*RI and *Stu*I sites followed by translation stop codons in all three reading frames. The expression cassette is flanked by segments of the vaccinia virus TK gene; thus, TK⁻ selection can be used for isolation of recombinant plaques (UNIT 16.17).

Cup sonicator (385-W)
12 × 75-mm polystyrene tubes (Falcon)

Additional reagents and equipment for purification of plasmid (UNIT 1.7)

1. Purify recombinant plasmid for transfection by CsCl/ethidium bromide centrifugation or PEG precipitation (UNIT 1.7).

Transfection will require 5 μg DNA/well of cells.

When using the pTM1 plasmid (Fig. 16.19.2), make sure that the initiator ATG of the gene of interest is the ATG of the pTM1 NcoI site (see Critical Parameters). Determine that the correct construction has been obtained by restriction mapping and DNA sequencing (UNIT 3.2).

2. The day before infection, seed wells of a 6-well tissue culture dish with 5×10^5 CV-1 cells/well in complete DMEM-10. Incubate until cells are near confluency (usually overnight).

Cells should be counted with a hemacytometer (APPENDIX 3F).

3. Vortex stock of vTF7-3 to disperse clumps. Sonicate the lysate using a cup sonicator as follows: (a) fill the cup with ice water (~50% ice); (b) place tube containing vTF7-3

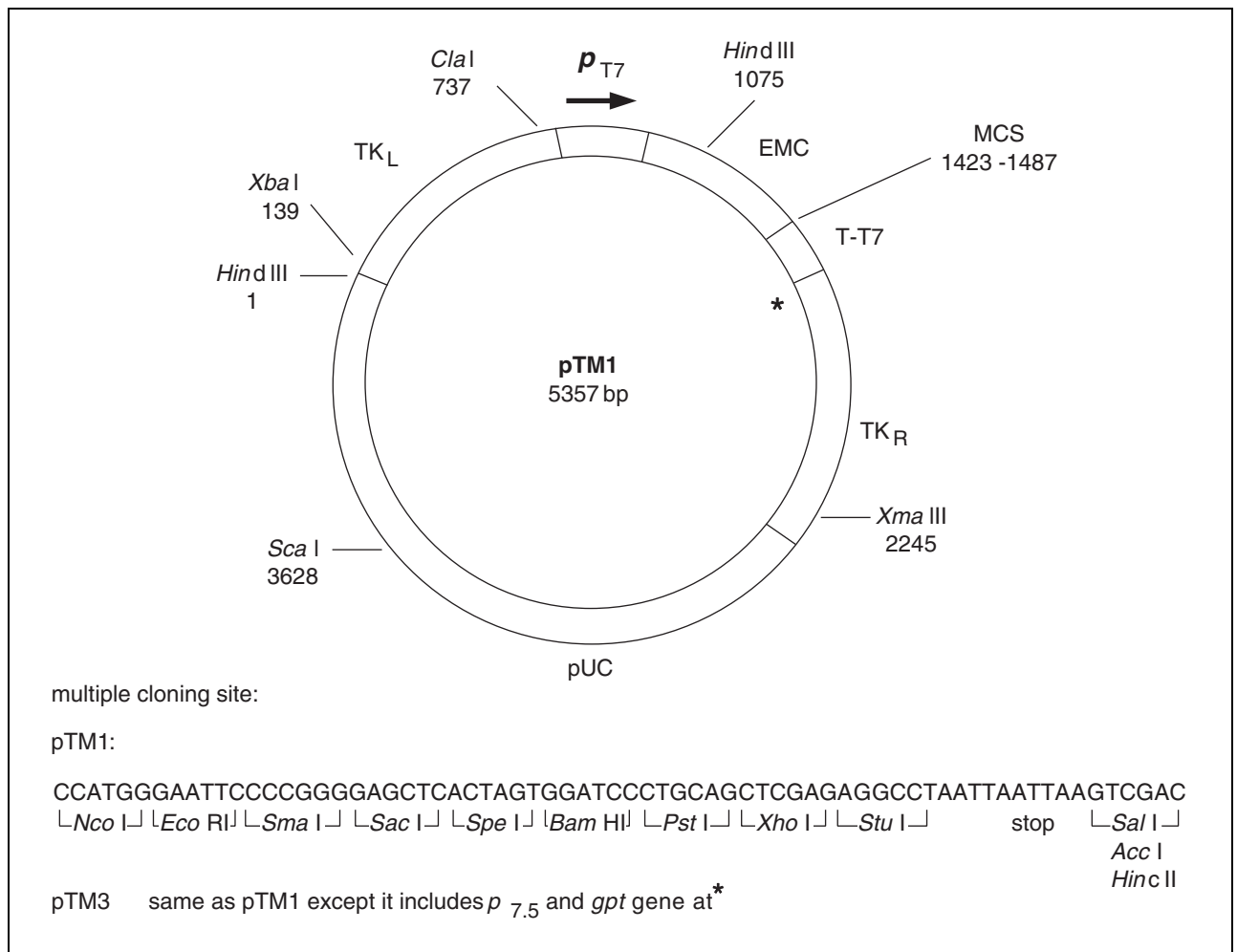


Figure 16.19.2 pTM1 contains the encephalomyocarditis virus untranslated leader region (EMCV-UTR) downstream of p_{T7} . The *NcoI* site contains the translation initiation codon and should be used for insertion of the 5' end of the protein-coding DNA segment. The 3' end of the DNA can be inserted into any downstream sites preceding the T7 terminator. The expression cassette is flanked by segments of the vaccinia virus TK gene; thus, TK⁻ selection can be used for isolation of recombinant virus (UNIT 16.17). pTM3 is the same as pTM1 except it includes the guanine phosphoribosyltransferase (*gpt*) gene to permit mycophenolic acid selection for recombinant virus (Moss et al., 1990; UNIT 9.5).

stock in the ice water; and (c) sonicate full power, 30 sec. Place lysate on ice ≥ 30 sec, then repeat sonication.

Because sonication melts the ice, it is necessary to replenish the ice in the cup.

4. Dilute an aliquot of the virus to 2×10^7 pfu/ml in Opti-MEM I.
5. Infect subconfluent monolayer of CV-1 cells from step 2 with 0.5 ml/well diluted virus stock (i.e., 10 pfu/cell). Incubate 30 min, rocking every 5 to 10 min.
6. Approximately 5 min before the end of the infection, prepare liposome suspension. Place 1 ml Opti-MEM I in separate 12×75 -mm polystyrene tubes (one for each well to be tested). Vortex liposome suspension and add 15 μ l to the medium; vortex briefly. Add 5 μ g recombinant plasmid (from step 1) and mix gently.

Vortex the liposome suspension before use to suspend the liposomes, which gradually settle during storage. Do not use a polypropylene tube, as the DNA/liposome complex may adhere to the surface.

7. Aspirate virus inoculum from CV-1 cells and add the DNA/liposome complex directly to the cells. Incubate 5 to 24 hr, then analyze expressed protein by pulse labeling (see Support Protocol) or by methods described in *UNIT 16.18*.

The time needed for incubation depends on the analysis that will be done. Although expression is detectable within a few hours after transfection, pulse-labeling at late times is preferred because host protein synthesis is inhibited and synthesis of the recombinant protein is increased (UNIT 16.15).

COINFECTION WITH TWO RECOMBINANT VACCINIA VIRUSES

The recombinant plasmid containing the gene of interest under control of p_{T7} (Basic Protocol 1) is incorporated into vaccinia virus by homologous recombination (*UNIT 16.17*) and a stock is made (*UNIT 16.16*). This stock is used together with vTF7-3 to coinfect cells in monolayer or suspension cultures. During incubation, the gene of interest is efficiently transcribed by T7 RNA polymerase and translated in the cytoplasm of the infected cells. The protein is then analyzed (*UNIT 16.18*) or purified (*UNIT 10.10*).

This coinfection protocol is recommended for mass production of proteins. For this purpose, cells in suspension (i.e., Vero or HeLa S3 cells) are used.

Materials

- vTF7-3 vaccinia virus stock (ATCC #VR-2153)
- Stock of recombinant vaccinia virus encoding the gene of interest under control of p_{T7} (*UNITS 16.16 & 16.17*)
- Confluent monolayer culture of CV-1 cells (*UNIT 16.16*) or HeLa S3 cells from a spinner culture (*UNIT 16.16*)
- 0.25 mg/ml trypsin (Worthington 2× crystalline and salt-free; filter sterilize and stored at -20°C)
- Complete DMEM-10, MEM spinner-5, or MEM-2.5 (depends on cell type; see below and *UNIT 16.16*)
- Sorvall H-6000 rotor or equivalent
- 6-well tissue culture dishes with 35-mm-diameter wells or larger tissue culture flasks for scaled-up analysis
- Additional reagents and equipment for counting cells with a hemacytometer (*APPENDIX 3F*), vaccinia virus analysis (*UNIT 16.18*), and protein purification (*UNIT 10.10*)

Prepare the cells

- 1a. *For infection of a CV-1 monolayer:* The day before infection, seed wells of a 6-well tissue culture dish with 5×10^5 cells in complete DMEM-10. Incubate until cells are near confluency (usually overnight).

Increase volumes proportionally if large-scale production is desired.

- 1b. *For infection of HeLa S3 suspension cells:* Just prior to infection, count cells using a hemacytometer (*APPENDIX 3F*) and centrifuge the desired amount 5 min at $1800 \times g$ (2500 rpm in Sorvall H-6000 rotor), room temperature. Discard supernatant, resuspend cells in MEM spinner-5 medium at 2×10^7 cells/ml, and transfer to a smaller sterile tissue culture flask or Erlenmeyer flask.

2. Trypsinize virus stocks. Place an aliquot of each stock in a separate tube and add an equal volume of 0.25 mg/ml trypsin. Vortex vigorously. Incubate 30 min, vortexing at 5- to 10-min intervals.

The highest level of expression using this system is obtained with an MOI of 10 pfu/cell for each virus.

Vortexing usually breaks up any clumps of cells. However, if any visible clumps remain, chill to 0°C and sonicate 30 sec on ice. Sonication can be repeated several times but the sample should cool on ice between sonications (see Basic Protocol 1).

3. Combine the two trypsinized virus stocks.

Infect the cells

For cells in monolayer culture

- 4a. Dilute trypsinized virus mixture with complete MEM-2.5 to 4×10^7 pfu/ml.

This represents 2×10^7 pfu/ml from each virus stock.

- 5a. Aspirate medium from cells and infect wells of a 6-well dish with 0.5 ml virus mixture per well. Incubate 1 to 2 hr, rocking dish at 15- to 30-min intervals.
- 6a. Overlay cells with 2 ml DMEM-10 and incubate. Analyze gene expression 5 to 24 hr after infection, depending on the method used. Alternatively, harvest 2 to 3 days after infection for protein purification.

Although expression is detectable within a few hours after infection, analysis by immunoprecipitation, immunoblotting, or northern blotting is preferred at late times because the T7-regulated gene product accumulates to a higher level. Earlier analysis may be preferred for assaying the biological activity of the expressed protein.

For suspension cells

- 4b. Add trypsinized virus mixture to the concentrated cells. Stir with a sterile magnetic stir bar 30 min at 37°C.
- 5b. Transfer infected cells to a larger vented spinner flask and dilute with MEM spinner-5 to bring the cell density to 5×10^5 cells/ml. Stir 1 to 3 days at 37°C.
- 6b. Harvest cells and analyze or purify the expressed protein (see annotation to step 6a above).

INFECTION OF OST7-1 CELLS WITH A SINGLE VIRUS

OST7-1 cells are derived from mouse L929 cells and constitutively express the bacteriophage T7 RNA polymerase in the cytoplasm. Thus, using this cell line eliminates the need for infection with ν TF7-3. This protocol describes the infection of OST7-1 cells with a single recombinant virus containing the gene of interest.

Materials

Confluent monolayer of OST7-1 cells (available from B. Moss, e-mail bmoss@nih.gov)

Complete DMEM-10 (UNIT 16.16) containing 400 potent μ g/ml Geneticin (G418, UNIT 9.5; Life Technologies; prepared from 80 mg/ml stock in PBS, filter sterilized and stored at -20°C)

Stock of recombinant vaccinia virus containing the gene of interest under the control of p_{T7} promoter (UNITS 16.16 & 16.17)

0.25 mg/ml trypsin (Worthington 2 \times crystalline and salt-free; filter sterilize and store at -20°C)

Complete MEM-2.5 (UNIT 16.16)

6-well tissue culture dish with 35-mm-diameter wells

**BASIC
PROTOCOL 3**

**Protein
Expression**

16.19.5

**BASIC
PROTOCOL 4**

1. Seed wells of a 6-well tissue culture dish with 1×10^6 OST7-1 cells/well in complete DMEM-10 containing 400 potent $\mu\text{g/ml}$ Geneticin. Incubate until cells reach confluency (usually overnight).
Complete DMEM-10 containing Geneticin is used to maintain the culture, but Geneticin is not required for the expression experiment itself.
2. Mix an equal volume of the virus stock and 0.25 mg/ml trypsin by vortexing vigorously. Incubate 30 min at 37°C, vortexing at 5- to 10-min intervals.
Use an MOI of 10 pfu/cell.
3. Dilute the trypsinized virus to 2×10^7 pfu/ml with complete MEM-2.5.
4. Infect confluent monolayer of OST7-1 cells with 0.5 ml virus mixture/well. Incubate 1 to 2 hr, rocking at 15- to 30-min intervals.
5. Add 1.5 ml/well complete MEM-2.5 and continue to incubate.
6. Analyze the cells 3 to 24 hr after infection (see Basic Protocol 2, annotation to step 6a).

GENE EXPRESSION USING THE VOTE SYSTEM

In this more recently developed method for vaccinia virus/T7 RNA polymerase hybrid expression, the gene of interest is cloned into the plasmid pVOTE.1 or pVOTE.2 and then incorporated into the genome of vaccinia virus vT7lacOI by homologous recombination (UNIT 16.17) and a stock is made (UNIT 16.16). Cells are then infected with the recombinant vaccinia virus in the presence of isopropyl-1-thio- β -D-galactoside (IPTG) and the gene of interest is efficiently transcribed by T7 RNA polymerase and translated in the cytoplasm of infected cells. Figure 16.19.3 illustrates this system.

Materials

Vaccinia virus vT7lacOI (available from B. Moss, e-mail *bross@nih.gov*)
DNA containing gene of interest
pVOTE.1 or pVOTE.2 (available from B. Moss)
CV-1 or BS-C-1 confluent monolayer (UNIT 16.16)
Cells for protein expression (e.g., HeLa)
Isopropyl-1-thio- β -D-galactoside (IPTG)

Additional reagents and equipment for PCR (UNIT 15.1), restriction enzyme digestion (UNIT 3.1), preparation of vaccinia virus stock (UNIT 16.16), infection of cells with vaccinia virus, transfection of infected cells with a vaccinia vector, harvest of vaccinia virus, selection of recombinant vaccinia virus plaques, and plaque purification of virus (UNIT 16.17), and determination of protein expression in the vaccinia system (UNIT 16.18; also see Support Protocol in this unit)

Insert gene into pVOTE.1 or pVOTE.2

1. Amplify the open reading frame from the gene of interest by PCR (UNIT 15.1) so that the translation-initiation codon is part of an *NcoI* (CCATGG) site for cloning into pVOTE.1 or a *NdeI* (CATATG) site for cloning into pVOTE.2.
*The other end of the PCR product should be compatible with a unique restriction endonuclease site present in the MCS (*XhoI*, *SacI*, *SmaI*, *EcoRI*, *PstI*, *Sall*, or *BamHI*).*
For high expression, the translation start codon must be adjacent to the encephalomyocarditis virus ribosome entry site.
2. Digest the PCR product and either pVOTE.1 or pVOTE.2 with the chosen restriction endonucleases (UNIT 3.1) and ligate the PCR product and plasmid together.

Transfer gene to vT7lacOI

3. Prepare a stock of vT7lacOI (UNIT 16.16) and use the stock to infect CV-1 or another cell line (UNIT 16.17).
4. Transfect recombinant plasmid derived from pVOTE.1 or pVOTE.2 (from step 1) into the cells infected with vT7lacOI (from step 3). Harvest the virus.

All of these procedures are described in UNIT 16.17.

5. Select recombinant virus for expression of XGPRT and plaque purify virus (UNIT 16.17).

Because the gene is inserted into the hemagglutinin site of vT7lacOI, the cells infected with recombinant virus can be recognized with the light microscope by their formation of syncytia.

6. Prepare a stock of the plaque-purified virus (UNIT 16.16).

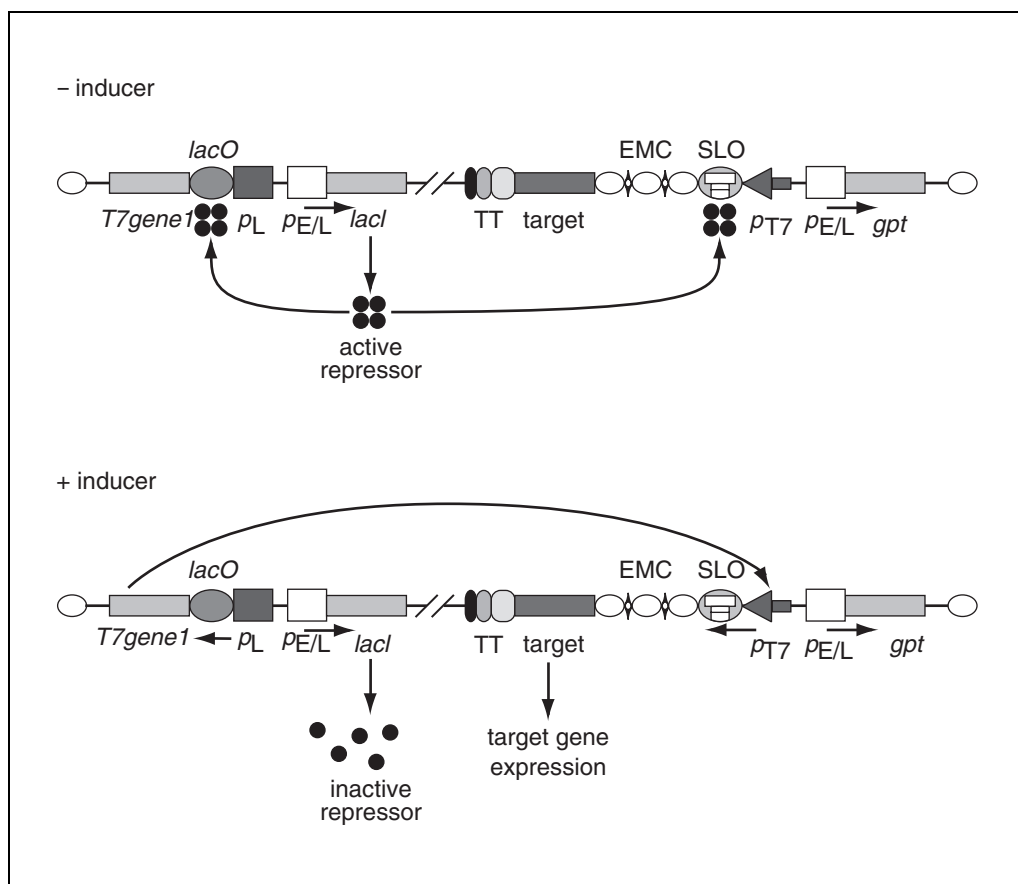


Figure 16.19.3 Schematic representation of the VOTE expression system. Portions of a recombinant vaccinia virus genome containing regulatory elements are diagrammed. Transcription of the *lacI* gene under control of a vaccinia virus early/late promoter ($p_{E/L}$) results in continuous synthesis of repressor monomers (represented by solid circles) which assemble into tetramers and bind to the *lacO* adjacent to a vaccinia virus late (p_L) promoter, preventing transcription of the bacteriophage T7 RNA polymerase gene (*T7gene1*), and to the modified *lacO* (SLO) next to the T7 promoter (p_{T7}), preventing transcription of the target gene should leaky synthesis of the T7 RNA polymerase occur. Upon induction, the repressor is inactivated, vaccinia virus RNA polymerase transcribes T7 gene 1, and T7 RNA polymerase is made. The latter binds to the T7 promoter and transcribes the target gene from the SLO to the triple terminator (TT). Translation of the target gene mRNA is enhanced by the EMC leader. From Ward et al., 1995; reprinted with permission of the National Academy of Sciences.

**SUPPORT
PROTOCOL**

Express the gene by infecting cells in the presence of IPTG

7. Infect monolayer or suspension cultures of HeLa or other cells with 10 pfu/cell of the recombinant virus and add 5 μ M to 2 mM IPTG to the medium at the start of infection.
8. Determine expressed proteins by standard methods (*UNIT 16.18*; also see Support Protocol in this unit).

DETECTION OF EXPRESSED PROTEIN USING PULSE LABELING

Pulse labeling infected cells is a quick and sensitive way to monitor expression of the T7-regulated gene relative to expression of vaccinia late proteins. When performed 24 hr post-infection under optimal conditions, up to 80% of the [³⁵S]methionine is incorporated into the product of the T7-regulated gene. This protein is detected using SDS-PAGE. If the EMC untranslated leader is present (e.g., if the pTM1 plasmid is used), hypertonic conditions may be used to selectively enhance cap-independent translation.

Additional Materials

- Infected cells expressing the desired T7-regulated gene of interest (see Basic Protocols 1, 2, 3, or 4) in a 6-well tissue culture dish
- Methionine- or cysteine-free, serum-free MEM (Life Technologies, Select-Amine Kit)
- 10 mCi/ml [³⁵S]methionine (1175 Ci/mmol; Amersham) *or* 15 mCi/ml [³⁵S]cysteine (600 Ci/mmol; Amersham)
- Phosphate-buffered saline (PBS; *APPENDIX 2*), ice-cold
- Cell lysis buffer (*UNIT 16.18*)
- 6 \times SDS/sample buffer (*UNIT 10.2*)
- Fixing solution (*UNIT 10.6*)
- Fluorographic solution (EN³HANCE from Du Pont NEN; *or* Amplify from Amersham)
- Cell scraper
- 95°C water bath
- Additional reagents and equipment for denaturing (SDS) gel electrophoresis (*UNIT 10.2*) and autoradiography (*APPENDIX 3A*)

Label the cells

1. Obtain infected cells expressing the desired T7-regulated gene of interest 24 hr after infection. Aspirate medium and replace it with 0.3 ml methionine- or cysteine-free MEM per well. Incubate 20 min.

Hypertonic conditions selectively enhance cap-independent translation. Therefore, if the EMC leader is present in the plasmid used in the infection (e.g., pTM1; Fig. 16.19.2), add 4.8 μ l of 5 M NaCl to each well (190 mM final).

2. Add 30 μ Ci of [³⁵S]methionine or [³⁵S]cysteine to each well. Incubate 30 min.

Lyse the cells

3. Aspirate medium and add 1 ml ice-cold PBS to each well.
4. Scrape cells with a disposable cell scraper and transfer to a 1.5-ml microcentrifuge tube. Microcentrifuge 1 min at high speed. Aspirate and discard supernatant.
5. Resuspend cell pellet in 100 μ l cell lysis buffer. Vortex briefly and incubate 5 min on ice.

6. Microcentrifuge 5 min at maximum speed and transfer supernatant to a clean microcentrifuge tube.

Analyze proteins

7. Remove 10 μ l lysate and place in another microcentrifuge tube with 2 μ l of 6 \times SDS/sample buffer. Heat 5 min at 95°C.
8. Load sample on single well of denaturing SDS-polyacrylamide gel and perform electrophoresis (UNIT 10.2).

A 10% polyacrylamide gel is recommended for separation of vaccinia late proteins that will appear as background bands on the autoradiogram.

It is important to use a control of infected cells that do not express the gene of interest.

9. Fix gel 20 min in fixing solution.
10. Incubate gel in fluorographic solution according to manufacturer's instructions.
11. Dry gel and autoradiograph by exposing to X-ray film.

If the pTM1 vector is used, the gene product is usually at a high enough level to appear as a major band among the vaccinia virus late gene products. If the T7-regulated gene product co-migrates with a vaccinia gene product, it will be necessary to analyze the expression using other methods as described in UNIT 16.18.

COMMENTARY

Background Information

The most efficient procedures for expression of genes in the cytoplasm of mammalian cells utilize a recombinant vaccinia virus encoding the bacteriophage T7 RNA polymerase. Many of the properties of T7 RNA polymerase that contribute to its successful use in prokaryotic gene expression (UNIT 16.2) are also relevant to its application for expression in eukaryotic cells. These properties include (1) the single subunit structure of T7 RNA polymerase, (2) the processive transcriptase activity of the enzyme, and (3) stringent promoter specificity in the eukaryotic cytoplasm.

In one version of the hybrid expression system described in this unit, plasmid DNA containing a gene of interest under the control of the p_{T7} promoter is transfected into cells infected with the recombinant vaccinia virus, after which the gene is transcribed efficiently by T7 RNA polymerase (Fuerst et al., 1986). Alternatively, DNA can be transfected into a cell line that constitutively expresses T7 RNA polymerase in the cytoplasm. However, for reasons that remain unclear, high-level expression in this cell line still requires infection with wild-type vaccinia virus (even when the cap-independent EMC virus untranslated region is used; Elroy-Stein and Moss, 1990). Therefore, the T7 RNA polymerase–cell line does not offer a clear advantage over the T7 RNA polym-

erase–vaccinia virus recombinant for transfection experiments.

In the vaccinia/T7 system, mRNA derived from the transfected gene can be as much as 10% to 30% of total cytoplasmic RNA. However, although mRNA levels of the T7-regulated gene are often quite high, only moderate amounts of protein are made. This limited translation is presumably due to the low efficiency of capping the transcripts made by T7 RNA polymerase (Fuerst and Moss, 1989). This difficulty was addressed by incorporating a feature of picornavirus mRNAs—a long untranslated leader region (UTR or “ribosomal landing pad”) that facilitates cap-independent ribosome binding—just downstream of the T7 promoter in certain vectors. The use of such vectors (e.g., pTM1; Fig. 16.19.2) in the vaccinia/T7 system enhances expression 5- to 10-fold (Elroy-Stein et al., 1989).

Transfection of DNA can be mediated by calcium phosphate or cationic liposomes. The latter method has reproducibly allowed 80% to 90% of the cells to be transfected (Felgner et al., 1987; Elroy-Stein and Moss, 1990; Rose et al., 1991; see also UNIT 9.4).

There have been several more recent innovations in the vaccinia virus/T7 RNA polymerase hybrid expression system. One is the use of the highly attenuated MVA strain of vaccinia virus to express the T7 RNA polymerase gene.

Overall, the levels of expression of transfected plasmids are similar using vTF7-3 and MVA/T7pol, but one may be higher than the other in certain cells (Wyatt et al., 1995). The main advantages of MVA/T7pol over vTF7-3 are lower biosafety requirements and the expression of genes in the absence of virus replication in nonpermissive cells (UNIT 16.16). Stocks of MVA/T7pol are prepared as described for MVA in UNIT 16.16. Liposome-mediated transfection is carried out as described in Basic Protocol 1 of this unit.

Recombinant vaccinia viruses that express SP6 RNA polymerase have been made, and are used in conjunction with plasmid vectors that use SP6 promoters for gene regulation (Usdin et al., 1993).

Another innovation is the development of the VOTE inducible expression system (see Basic Protocol 4 and Ward et al., 1995). This system should supersede the two-virus system and the need for cell lines expressing T7 RNA polymerase (UNIT 16.19). The VOTE system is tightly regulated, because there is an *E. coli lacO* regulating the T7 RNA polymerase gene and another one regulating the T7 promoter adjacent to the gene of interest. The level of expression can be titrated with IPTG. Another version of the VOTE system has a thermolabile *E. coli lac* repressor, so that induction occurs by temperature elevation (Ward et al., 1995).

Critical Parameters

To maximize protein expression, the gene of interest is cloned into a plasmid vector (e.g., pTF7-5 or pTM1) that contains both p_{T7} and the T7 terminator (Figs. 16.19.1 & 16.19.2). Better expression (5- to 10-fold) is obtained by using the encephalomyocarditis virus (EMCV)-UTR downstream of p_{T7} in pTM1. Alternatively, any plasmid that carries the T7 promoter (e.g., pBluescript; see UNIT 16.2 critical parameters) can be used but may give lower expression levels.

When the “ribosomal landing pad” within the EMCV-UTR is used, translation initiation occurs at the ATG, which is in the *NcoI* site of pTM-1. If native protein is to be expressed, the initiator ATG codon must be in the *NcoI* site. If necessary, the *NcoI* site may be mutated but the position of the initiator ATG should not be changed. If a fusion protein is desired, the coding sequence is inserted into one of the multiple cloning sites, in frame with the ATG of the *NcoI* site.

For some methods of detection (e.g., immunofluorescence; Earl et al., 1990), it is pref-

erable to harvest the cells ≤ 24 hr after transfection. If the incubation continues much beyond this point, the cells may detach from the monolayer and be lost during washing or staining.

For analytical purposes, expression can be achieved by transfecting the recombinant plasmid into cells that have been infected with a recombinant vaccinia virus, such as vTF7-3, which expresses the T7 RNA polymerase gene. For mass production of proteins—e.g., 8 μg CAT/ 10^6 cells (Elroy-Stein et al., 1989) or 11 mg HIV envelope protein/liter (Barret et al., 1989)—a recombinant vaccinia virus that contains the p_{T7} -regulated target gene should be generated by homologous recombination (UNIT 16.17). It should be introduced into cells constitutively expressing the T7 RNA polymerase, or used with vTF7-3 to coinfect any mammalian or avian cell line. Alternatively, the VOTE system eliminates the need for special cell lines or vTF7-3.

Anticipated Results

In most cases, 60% to 90% of transfected cells (following infection with vTF7-3) or 99% of infected cells (when the gene of interest is incorporated in a vaccinia virus) are observed to express the desired protein. If protein is being detected by pulse labeling, an overnight exposure should result in the protein product appearing as a dominant band over the background of vaccinia proteins.

Time Considerations

Starting with a plasmid containing the gene of interest cloned in the appropriate vector, the entire procedure from infection or infection/transfection to harvesting and analyzing the expressed product requires 2 to 3 days.

Literature Cited

- Barret, N., Mitterer, A., Mundt, W., Eibl, J., Eibl, M., Gallo, R.C., Moss, B. and Dorner, F. 1989. Large-scale production and purification of a vaccinia recombinant-derived HIV-1 gp160 and analysis of its immunogenicity. *AIDS Res. Hum. Retroviruses* 5:159-171.
- Berger, E.A., Fuerst, T.R., and Moss, B. 1988. A soluble recombinant polypeptide comprising the amino-terminal half of the extracellular region of the CD4 molecule contains an active binding site for human immunodeficiency virus. *Proc. Natl. Acad. Sci. U.S.A.* 85:2357-2361.
- Earl, P.L., Koenig, S. and Moss, B. 1990. Biological and immunological properties of human immunodeficiency virus type 1 envelope glycoprotein: Analysis of proteins with truncations and deletions expressed by recombinant vaccinia viruses. *J. Virol.* 65:31-41.

- Elroy-Stein, O. and Moss, B. 1990. Cytoplasmic expression system based on constitutive synthesis of bacteriophage T7 RNA polymerase in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 87:6743-6747.
- Elroy-Stein, O., Fuerst, T.R., and Moss, B. 1989. Cap-independent translation of mRNA conferred by encephalomyocarditis virus 5' sequence improves the performance of the vaccinia virus/bacteriophage T7 hybrid expression system. *Proc. Natl. Acad. Sci. U.S.A.* 86:6126-6130.
- Felgner, P.L., Gadek, T.R., Holm, H., Roman, R., Chan, H.W., Wenz, M., Northrop, J.P., Ringold, G.A., and Danielsen, G.A. 1987. Lipofection: A highly efficient, lipid-mediated DNA transfection procedure. *Proc. Natl. Acad. Sci. U.S.A.* 84:7413-7417.
- Fuerst, T.R. and Moss, B. 1989. Structure and stability of mRNA synthesized by vaccinia virus-encoded bacteriophage T7 RNA polymerase in mammalian cells. Importance of the 5' untranslated leader. *J. Mol. Biol.* 206:333-348.
- Fuerst, T.R., Niles, E.G., Studier, F.W., and Moss, B. 1986. Eukaryotic transient expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 83:8122-8126.
- Moss, B., Elroy-Stein, O., Mizukami, T., Alexander, W.A., and Fuerst, T.R. 1990. New mammalian expression vectors. *Nature (Lond.)* 348:91-92.
- Rose, J.K., Buonocore, L., and Whitt, M.A. 1991. A new cationic liposome reagent mediating nearly quantitative transfection of animal cells. *Biotechniques* 10:520-525.
- Usdin, T.B., Brownstein, M.J., Moss, B., and Isaacs, S.N. 1993. SP6 RNA polymerase containing vaccinia virus for rapid expression of cloned genes in tissue culture. *BioTechniques* 14:222-224.
- Ward, G.A., Stover, C.K., Moss, B., and Fuerst, T.R. 1995. Stringent chemical and thermal regulation of recombinant gene expression by vaccinia virus vectors in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 92:6773-6777.
- Wyatt, L.S., Moss, B., and Rozenblatt, S. 1995. Replication-deficient vaccinia virus encoding bacteriophage T7 RNA polymerase for transient gene expression in mammalian cells. *Virology* 210:202-205.

Contributed by Orna Elroy-Stein
Tel Aviv University
Tel Aviv, Israel

Bernard Moss
National Institute of Allergy and
Infectious Diseases
Bethesda, Maryland

Expression of Proteins Using Semliki Forest Virus Vectors

UNIT 16.20

Semliki Forest virus (SFV) vectors have been developed to provide a convenient system to express protein-encoding sequences in virtually any animal cell. This unit presents two strategies for protein expression using SFV vectors. In both cases the protein-coding sequence of interest is cloned into a plasmid vector, which is subsequently used to produce recombinant RNA *in vitro*. This RNA, which is of positive polarity, is transfected into cells and there is amplified by virtue of its self-encoded RNA replicase. The same replicase also produces a shorter RNA species that encodes the protein of interest.

Protein expression can be achieved in two ways. In the first strategy, cells are transfected and directly analyzed for expression of the heterologous protein. The other strategy employs *in vivo* packaging of the RNA into SFV particles; recombinant RNA is cotransfected with a special helper RNA that codes for the structural proteins needed for virus assembly. SFV particles carrying only recombinant RNA are formed and are used to infect cells for analysis of protein expression. The first expression strategy is useful when studying the phenotypic expression of several mutant genes, in preliminary gene expression analyses, and when only a limited amount of expression studies are planned for one construct. The second strategy is more useful when many expression analyses are planned with the same construct and when large-scale production of the gene product is planned.

The first basic and alternate protocols employ the first strategy and differ only in transfection procedure (electroporation and liposome-mediated transfection, respectively). Accompanying support protocols provide methods for checking expression and transfection through β -galactosidase assays of transfected cells and cell lysates. The second basic protocol uses the second strategy and is accompanied by support protocols providing methods for titrating and purifying recombinant virus stocks. Although the protocols presented here are designed for use with BHK (baby hamster kidney) cells, the virus has a very broad host range and can be used with many different cell types (see Background Information).

STRATEGIC PLANNING: CHOICE OF SFV VECTOR

The available SFV vectors are all pGEM-based, confer ampicillin resistance, and carry the SFV replicase-encoding genes *nsP1-4* followed by a promoter for subgenomic transcription (Fig. 16.20.1). A unique *SpeI* restriction site is placed at the 3' end of the SFV-derived sequence to permit linearization of the plasmid prior to runoff transcription *in vitro*. Downstream from the subgenomic promoter (which overlaps with the 3' end of the *nsP4* gene) is a polylinker site (*Bam*HI-*Sma*I-*Xma*I) where the foreign DNA sequence is inserted. Following the polylinker are translational stop codons in all three reading frames (in case the inserted sequence lacks a stop codon).

The basic vector is pSFV1, which is used when the inserted DNA fragment contains an initiation codon (AUG) preceded by its own Kozak box (sequence specifying ribosome binding). pSFV3 contains a Kozak box and a start codon and is used when the inserted sequence lacks these elements; it provides proper translation as long as the inserted sequence is placed in frame with the initiating AUG of the vector. pSFV3-lacZ contains an added copy of the *E. coli* β -galactosidase gene and is used as a convenient control.

NOTE: Consult UNIT 4.1 for recommendations for minimizing contamination of RNA preparations by ribonucleases and Critical Parameters regarding precautions for working with SFV particles.

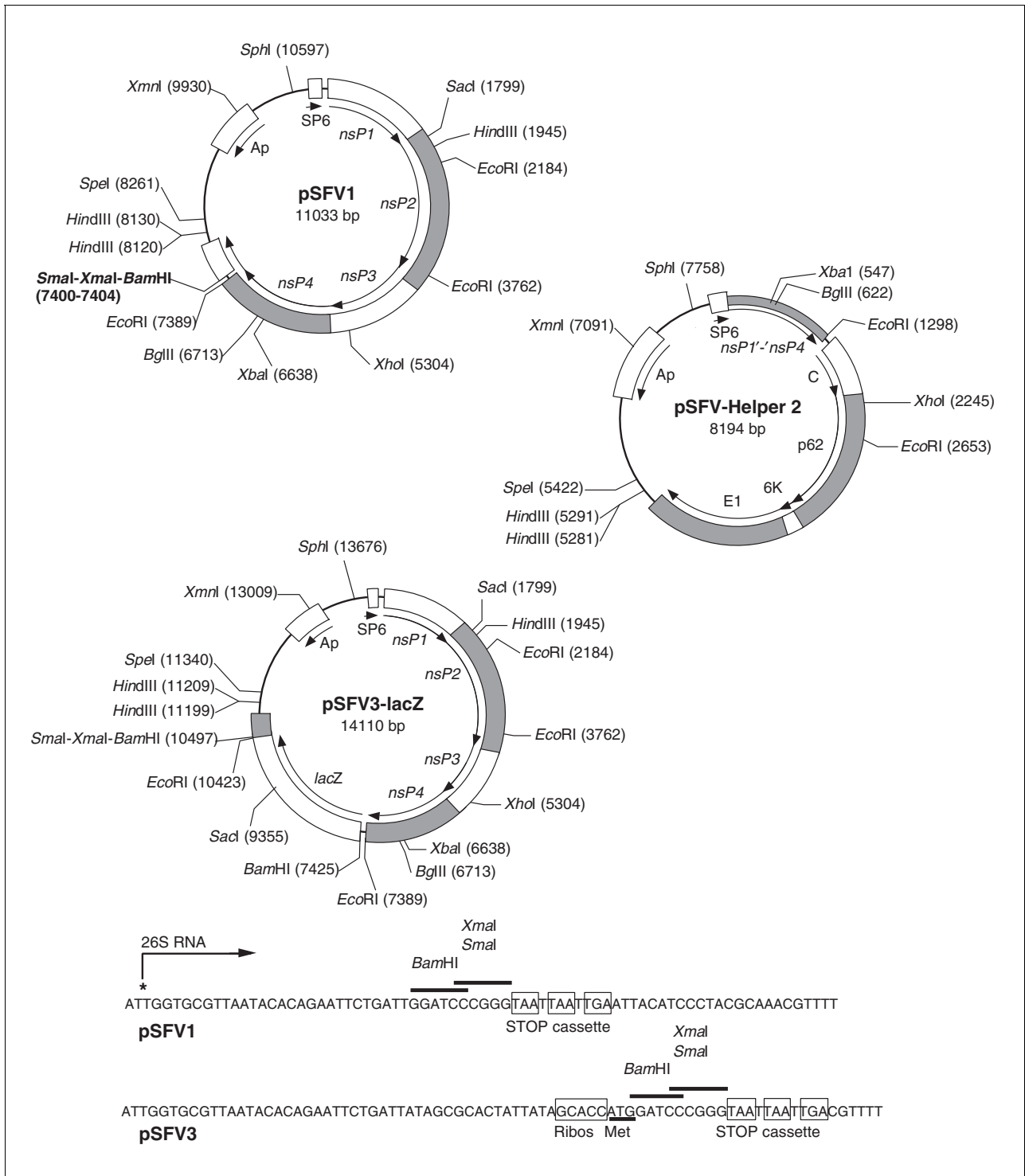


Figure 16.20.1 SFV expression vectors and their polylinker regions. pSFV1 and pSFV3 are almost identical in size and structure, differing only in their polylinker regions; pSFV3-lacZ is a variant of pSFV-3 containing an inserted copy of the *E. coli* β -galactosidase gene. These vectors carry the four genes encoding the SFV replicase (*nsP1-4*) but lack the structural genes of SFV (*C*, *p62*, *6K*, and *E1*). In contrast, the Helper plasmid lacks the *nsP* genes (the corresponding deletion flanking the remaining *nsP* sequences is indicated as *nsP1'-nsP4*, where the primes indicate the position of the deletion) but carries the structural genes. For all three plasmids, in vitro transcription is driven by the SP6 promoter. The unique *SpeI* site is used to linearize the plasmids prior to transcription. Sequences flanking the polylinker regions of pSFV1 and pSFV3 are shown below the plasmid maps. The first nucleotide transcribed from the subgenomic promoter, indicated above the pSFV1 sequence, is the same for all three plasmids. (Consult Liljeström and Garoff, 1991b; Liljeström et al., 1991; and Berglund et al., 1993 for further details).

16.20.2

EXPRESSION OF PROTEINS FROM RECOMBINANT SFV RNA USING ELECTROPORATIVE TRANSFECTION

BASIC PROTOCOL

SFV vector plasmids carrying heterologous sequences are used as templates for RNA synthesis *in vitro*. The RNA is transfected into BHK cells by electroporation, where the self-encoded RNA replicase amplifies the RNA molecules. In conjunction with efficient translational start signals resident on the RNA molecules, this ensures a high production rate for the heterologous protein. Production is easily monitored by metabolic labeling of transfected cells and analysis of the protein products by SDS-PAGE and autoradiography.

Materials

DNA fragment encoding protein of interest
pSFV1, pSFV3, or pSFV3-lacZ expression vector (GIBCO/BRL; see Fig. 16.20.1)
SpeI restriction endonuclease and buffer (UNIT 3.1)
10× SP6 RNA polymerase buffer
50 mM dithiothreitol (DTT)
10 mM ^{m7}G(5')ppp(5')G
rNTP mix
40 U/μl RNasin (Promega) or other RNase inhibitor
60 U/μl SP6 RNA polymerase (UNIT 3.8)
5× TD solution
λ DNA molecular weight markers (e.g., λ digested with *EcoRI* + *HindIII*; UNIT 2.5A)
BHK-21 cells (ATCC)
Complete BHK-21 medium
Phosphate-buffered saline (PBS; Reagents and Solutions), 37°C, room temperature, and ice cold
Trypsin/EDTA solution: 0.5 mg/ml trypsin/0.2 mg/ml EDTA in PBS
Starvation medium
15 mCi/ml [³⁵S]methionine (>1000 Ci/mmol)
Chase medium
NP-40 lysis buffer
75-cm² tissue culture flask
Electroporator (e.g., Bio-Rad)
0.2- or 0.4-cm electroporation cuvette
35-mm tissue culture plate

Additional reagents and equipment for subcloning (UNIT 3.16), preparation of plasmid DNA (UNIT 1.7), restriction endonuclease digestion (UNIT 3.1), phenol extraction and ethanol precipitation of DNA (UNIT 2.1), spectrophotometric quantitation of RNA and DNA (APPENDIX 3), agarose gel electrophoresis (UNIT 2.5A), SDS-PAGE for protein analysis (UNIT 10.2), and autoradiography (APPENDIX 3)

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly. All incubations are performed in a humidified, 37°C, 5% CO₂ incubator unless otherwise specified.

Subclone fragment and prepare RNA by *in vitro* transcription

1. Subclone the DNA fragment of interest into an SFV expression vector. Grow up resulting plasmid and prepare purified DNA.
2. Linearize recombinant plasmid DNA by restriction digestion with *SpeI*. Phenol extract and ethanol precipitate the DNA. Determine DNA concentration and resuspend in water at 0.3 μg/μl.

Protein Expression

16.20.3

3. Set up *in vitro* transcription in a microcentrifuge tube by mixing the following:

- 5 μ l DNA (1.5 μ g)
- 5 μ l 10 \times SP6 buffer
- 5 μ l 50 mM DTT
- 5 μ l 10 mM m 7G(5')ppp(5')G
- 5 μ l rNTP mix
- 23 μ l H₂O
- 1.5 μ l 40 U/ μ l RNasin
- 0.5 μ l 60 U/ μ l SP6 RNA polymerase.

Because the spermidine in the SP6 buffer may precipitate DNA at 0°C, the reaction should be set up at room temperature. Under the conditions described (1.5 μ g of linear DNA and 30 U of SP6 RNA polymerase per reaction), the transcription mixture is saturated for DNA and should yield ~50 μ g of RNA. The yield can be increased by increasing the amount of SP6 RNA polymerase used.

4. Incubate 60 min in a 37°C water bath. Transfer a 1- μ l aliquot into 10 μ l water and add 3 μ l of 5 \times TD solution. Run sample on a 0.5% agarose gel, using digested λ DNA as a molecular weight marker. Freeze the remainder of the reaction in 10- μ l aliquots at -80°C.

Although the gel is nondenaturing and thus does not reflect the molecular weight of the RNA produced, it is nevertheless useful for checking RNA quality and quantity. The RNA band should be defined (no smearing) and relatively thick in comparison to DNA bands. If desired, RNA samples of known concentration can be run as a rough quantitative comparison.

Transfect BHK-21 cells by electroporation

5. Inoculate BHK-21 cells into 10 ml complete BHK-21 medium in a 75-cm² tissue culture flask and grow to late log phase (this will usually take 1 to 2 days).
6. Aspirate medium and wash cells once with room-temperature PBS.
7. Add 2 ml trypsin/EDTA solution, incubate until cells detach (~1 min), and briefly pipet cell solution back and forth, monitoring under a microscope, to obtain a single-cell suspension. Add 10 ml complete BHK-21 medium to stop trypsinization.
8. Centrifuge cells 5 min at 400 \times g, room temperature, in a tabletop centrifuge and remove supernatant. Resuspend cells in 10 to 20 ml PBS.
9. Centrifuge cells as in the previous step and resuspend in PBS at 10⁷ cells/ml.
10. Thaw an aliquot of the RNA to be transfected (transcription mixture from step 4), place 10 to 15 μ l in a microcentrifuge tube, and add 0.8 ml of cell suspension. Mix thoroughly by pipetting and transfer to a 0.2- or 0.4-cm electroporation cuvette.

Using 4 μ g RNA per 10⁷ cells will give a transfection frequency close to 100%. Thus, using 10 to 15 μ l transcription mixture (~5 to 7 μ g RNA) per 75-cm² flask works well.

11. Pulse twice at room temperature using a voltage of 850 V, a capacitor setting of 25 μ F, and a time constant after the pulse of 0.4.

If a large number of samples are being manipulated and cells start to sediment in the cuvettes before the pulse, cap each cuvette tightly and invert once before shocking.

The Bio-Rad gene pulser uses capacitor discharges to produce pulses, generating an exponential-decay waveform. For this electroporator the pulse-controller should be disconnected when using 0.4-cm cuvettes; in this case the set value of 850 V will represent the actual voltage given to cells in the cuvette (i.e., field strength = 2125 V/cm). When using 0.2-cm cuvettes, the pulse-controller must be connected to prevent

arc. In this case, however, the set *V* value will not represent what is given to the cells, so the actual *V* required to obtain the proper field strength should be determined by measuring the pulse directly in the cuvette (generally ~1700 V). If an electroporator other than the Bio-Rad model is used, settings may need to be modified; consult UNITS 9.3 & 9.9 for recommendations.

12. Dilute transfected cells 20-fold in complete BHK-21 medium, rinsing the cuvette with the medium to collect all cells, and transfer to a 35-mm tissue culture plate. Incubate 7 to 9 hr.

Cells will reattach in ~1 hr and metabolic labeling can be performed as early as 4 hr after transfection. However, for maximum labeling of the protein of interest and minimum labeling of host proteins, incubation for 7 to 9 hr after transfection is recommended.

Analyze heterologous protein by metabolic labeling of cells

13. Aspirate medium and wash cells twice with 3 ml prewarmed PBS. Overlay cells with 2 ml starvation medium and incubate plate 30 to 45 min.
14. Prepare 500 μ l of pulse medium: starvation medium containing 25 to 50 μ Ci [³⁵S]methionine. Aspirate medium from cells, replace with pulse medium, and incubate for the desired pulse time.

Consult UNIT 10.18 for a discussion of how to select appropriate pulse and chase times.

15. Remove pulse medium and wash cells once with 2 ml chase medium. Overlay cells with 2 ml chase medium and incubate for desired chase time.
16. Remove chase medium and wash cells with 3 ml ice-cold PBS.
17. Add 300 μ l NP-40 lysis buffer and incubate 10 min on ice.
18. Resuspend cells and transfer to a microcentrifuge tube. Microcentrifuge 5 min at 6000 rpm at 4°C or room temperature to pellet nuclei. Store supernatant (cytoplasmic proteins) or pellet (nuclear proteins) at -80°C.
19. Assay for protein expression by SDS-PAGE and autoradiography.

If pSFV3-lacZ was used, expression and transfection efficiencies can be checked using the β -galactosidase expression assays described in the first and second support protocols.

EXPRESSION OF PROTEINS FROM RECOMBINANT SFV RNA USING LIPOSOME-MEDIATED TRANSFECTION

ALTERNATE PROTOCOL

If a gene pulser is not available, liposome-mediated transfection can replace electroporation (steps 5 to 12 of the first basic protocol). With this method, the transcription mixture must be purified to remove the NTPs from the reaction mixture, either by gel filtration or by isopropanol precipitation of the RNA with a subsequent ethanol wash.

Additional Materials

- TE buffer, pH 7.5 (APPENDIX 2)
- 3 M sodium acetate, pH 4.8 (APPENDIX 2)
- Opti-MEM transfection medium (GIBCO/BRL)
- Isopropanol, -20°C
- 75% (v/v) ethanol
- Lipofectin (GIBCO/BRL)
- Nu-Clean R50 RNA spin columns (IBI)
- Centrifuge with swinging-bucket rotor

Protein Expression

16.20.5

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly. All incubations are performed in a humidified, 37°C, 5% CO₂ incubator unless otherwise specified.

1. Prepare RNA by in vitro transcription as described in steps 1 to 4 of the preceding basic protocol.

To purify RNA by gel filtration:

- 2a. Warm column to room temperature and invert several times to resuspend gel. Remove top and bottom caps and allow excess buffer to drain out of the column.
- 3a. Position the column on top of a 1.5-ml microcentrifuge tube placed within bucket of a swinging-bucket rotor and prespin 2 min at 1100 × *g*. Discard collection tube and replace with a fresh one.
- 4a. Thaw transcription mixture. If necessary, dilute with TE buffer to a final volume of 50 µl (volumes between 50 and 100 µl can be used). Add sample directly to top of column and spin again as in step 3a. Discard column. Aliquot 20 µl of the purified RNA sample (from collection tube) and freeze at −80°C.

To purify RNA by isopropanol precipitation:

- 2b. Thaw transcription reaction mixture and dilute with water to 100 µl. Add 20 µl of 3 M sodium acetate, pH 4.8.
- 3b. Add 72 µl of −20°C isopropanol, mix, and incubate 10 min at room temperature.
- 4b. Microcentrifuge 15 min at top speed, room temperature. Remove supernatant, rinse with 500 µl of 75% ethanol, respin, remove ethanol, and briefly dry. Resuspend pellet in 50 µl TE buffer, pH 7.5.

Only DNA or RNA >100 bp is precipitated by this method. However, it is important that the precipitation step be performed at room temperature.

5. Grow BHK-21 cells on 35-mm tissue culture plates to ~80% confluency.
6. Remove medium and replace with 2 ml Opti-MEM medium. Incubate 5 to 10 min.
7. While cells are incubating, mix 1 ml Opti-MEM with 9 µg Lipofectin (per plate of cells) and vortex mixture for 10 sec. Add 2.3 µg purified RNA (step 4a or 4b) and vortex again.
8. Remove the Opti-MEM covering the cells and replace with Opti-MEM/Lipofectin/RNA mixture. Incubate 2 hr, tilting the plate every 15 min.

Up to 90% transfection frequencies can be obtained using Lipofectin. However, it is imperative that the transfection mixture always cover the cells as evenly as possible. Frequent tilting of plates is recommended for consistent results.

9. Remove transfection mix and wash cells once with 2 ml complete BHK-21 medium. Add 2 to 3 ml complete BHK-21 medium and continue incubation 7 to 9 hr.
10. Analyze heterologous protein by metabolic labeling of cells as described in steps 13 to 19 of the first basic protocol.

SCREENING FOR GENE EXPRESSION USING β -GALACTOSIDASE

The pSFV3-lacZ vector can be used to check expression and transfection efficiencies. A quick, convenient method to screen cells for lacZ expression is by direct staining with Xgal. Alternatively, an infected cell lysate can be assayed for β -galactosidase activity.

Screening of Cells for β -Galactosidase Activity**Materials**

BHK-21 cells transfected with pSFV-lacZ vector (first basic protocol, step 12, or first alternate protocol, step 9)
Methanol, -20°C
Xgal stain solution
Stereomicroscope

1. Following infection or transfection to permit expression of β -galactosidase, incubate cells 7 hr to overnight.
2. Wash cells twice with PBS. Add 3 ml of -20°C methanol and incubate plate 5 min at -20°C to fix the cells.
3. Wash cells three times with PBS at room temperature and overlay with 1 ml Xgal stain solution. Monitor the ratio of positive (blue) to negative (white) cells using a stereomicroscope.

Staining is usually quite fast (beginning within minutes) and is complete after 2 to 4 hr at 37°C . Alternatively, plates can be left at room temperature overnight.

The ratio of blue to white cells serves as an indication of transfection efficiency.

Screening of Cell Lysates for β -Galactosidase Activity**Additional Materials**

Lysate of BHK-21 cells transfected with pSFV3-lacZ vector (first basic or alternate protocol)
Z buffer with and without 4 mg/ml *o*-nitrophenyl- β -D-galactopyranoside (ONPG)
Purified β -galactosidase
1 M sodium carbonate

1. Place 1 μl cell lysate (~200 to 300 cell equivalents) in a 1.5-ml microcentrifuge tube and add Z buffer to 400 μl . Incubate 5 min.
2. Add 100 μl of 4 mg/ml Z buffer ONPG, mix well by vortexing, and incubate until sample becomes yellow.
3. Stop reaction by adding 500 μl of 1 M sodium carbonate. Mix well. Measure A_{420} .

To permit determination of the actual amount of lacZ protein (β -galactosidase) produced in the cell lysate, a standard curve should be prepared using purified β -galactosidase (a series of concentrations ranging from 0.2 to 50 ng β -galactosidase is usually suitable). These samples will generally need to be incubated ~15 min. The exact quantitation will vary with the preparation, but is usually ~1 U/3 μg protein.

EXPRESSION OF PROTEIN FROM IN VIVO-PACKAGED RECOMBINANT SFV PARTICLES

Recombinant and helper RNAs are cotransfected into BHK cells. As a result of trans complementation, recombinant RNAs are packaged into SFV particles which are released into the growth medium (Fig. 16.20.2). The virus produced is used to infect cells in culture which are then assayed for expression of the cloned sequences.

Materials

- pSFV1, pSFV3, or pSFV3-lacZ vector (GIBCO/BRL; see Fig. 16.20.1)
- pSFV-Helper 2 DNA (GIBCO/BRL; see Fig. 16.20.1)
- BHK-21 cells (ATCC)
- Minimum essential medium (MEM), supplemented
- Phosphate-buffered saline (PBS; Reagents and Solutions)
- α -chymotrypsin solution
- 2 mg/ml aprotinin (Sigma)
- 35-mm tissue culture plate

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly. All incubations are performed in a humidified, 37°C, 5% CO₂ incubator unless otherwise specified.

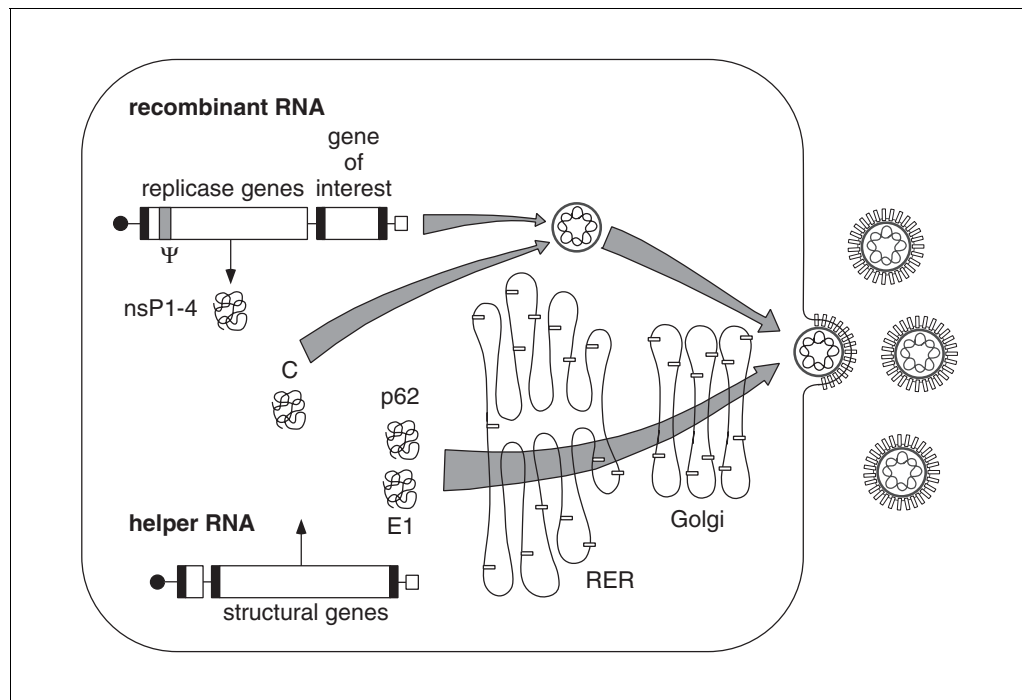


Figure 16.20.2 Schematic presentation of in vivo packaging of recombinant RNA into SFV particles. Recombinant and helper RNAs are cotransfected into BHK cells, where the recombinant RNA-encoded replicase amplifies both RNA species. The capsid protein (C) packages only recombinant RNA molecules into nucleocapsids, as the helper RNAs lack a packaging signal (Ψ ; this resides in the *nsP1* gene). Assembly and budding of new recombinant SFV virions occurs at the cell surface through an interaction between cytoplasmically preformed nucleocapsids and the viral spike membrane proteins p62 and E1. Stippled boxes in the recombinant RNA represent packaging signals; black boxes represent replication sequences recognized by the SFV replicase; black circles at the 5' ends of the RNAs denote cap structures; and open boxes at the 3' ends of the RNAs denote poly-A stretches.

Package recombinant RNA into pSFV

1. Subclone DNA fragment of interest into pSFV plasmid. Prepare RNA by in vitro transcription from the recombinant pSFV plasmid and the pSFV-Helper 2 plasmid using the procedure described in steps 1 to 4 of the first basic protocol.
2. Mix the two RNAs in a 1:1 molar ratio. Transfect RNA mixture into BHK-21 cells as described in steps 5 to 12 of the first basic protocol or steps 5 to 9 of the alternate protocol. Incubate 24 hr.

In practice, equal amounts (usually 10 μ l each) of the two transcription mixtures or purified RNAs can be used, because the transcription mixture was set up to contain a saturated amount of DNA (and thus promoter copies). Deviations from a 1:1 ratio will lead to lower stock titers.

3. Collect the medium and clarify by centrifuging 15 min at $2000 \times g$ in a tabletop centrifuge at 4°C .
4. Aliquot and freeze 0.5-ml aliquots quickly on dry ice or in liquid nitrogen. Store at -80°C . Titrate one aliquot as described in the third support protocol.

At this stage virus can also be purified and/or concentrated (fourth support protocol).

Infect cells with SFV

5. Grow fresh BHK-21 cells to 80% to 100% confluency on a 35-mm tissue culture plate.
6. Aspirate medium and wash cells thoroughly with PBS.

Activate Helper 2-packaged recombinant virus

7. Add $\frac{1}{20}$ vol α -chymotrypsin solution to virus preparation and incubate 30 min at room temperature.
8. Inactivate α -chymotrypsin by adding 0.5 vol aprotinin.
9. Dilute activated virus preparation as appropriate in supplemented MEM using a multiplicity of infection of five. Transfer solution to ice.

Activation should be done just before infection, as the activated stock cannot be stored frozen or for longer periods of time on ice.

10. Add 500 μ l of the virus solution to the cells and incubate 45 min to 1 hr.
11. Remove virus solution, add 3 ml complete BHK-21 medium (or other suitable medium required for the experiment), and continue incubation as required.
12. Prepare cell lysates and analyze protein expression as described in steps 17 to 19 of the first basic protocol.

DETERMINATION OF RECOMBINANT VIRUS TITER

The recombinant virus can initiate only one round of intracellular replication. Therefore, the titer of the packaged stock cannot be determined by conventional plaque assay, but is done using indirect immunofluorescence. Different dilutions of the virus stock are used to infect cells and heterologous protein expression is detected by specific antibodies.

Materials

Recombinant virus stock (second basic protocol)
PBS-Eisen
Methanol, -20°C
Blocking buffer: 0.5% (w/v) gelatin/0.2% (w/v) BSA in PBS-Eisen

SUPPORT PROTOCOL

Protein Expression

16.20.9

Primary antibody in blocking buffer
Secondary antibody in blocking buffer
Moviol 4-88 solution containing 2.5% DABCO (1,4-diazobicyclo-[2.2.2]-octane)
Glass coverslips
Additional reagents and equipment for immunofluorescent labeling of monolayer cells (UNIT 14.6)

1. Grow BHK-21 cells on glass coverslips to ~70% confluency and count with a hemacytometer to determine number of cells plated.
2. Activate virus stock as in steps 7 to 10 of second basic protocol and prepare a series of dilutions in supplemented MEM. Infect the cells as described in the second basic protocol, steps 5 and 6. Incubate 7 to 15 hr to allow for expression of the protein of interest.

A good initial dilution series is 1:10, 1:100, and 1:1000 (Critical Parameters, UNIT 14.6).

3. Rinse each coverslip twice with PBS-Eisen to cover. Fix cells 5 min in -20°C methanol for 5 min.
4. Remove methanol and wash coverslip three times with PBS.
5. Add blocking buffer and incubate 30 min at room temperature to block nonspecific binding.
6. Replace blocking buffer with primary antibody solution. Incubate 30 min at room temperature.
7. Wash three times with PBS. Repeat step 6 using secondary antibody solution.
8. Wash three times with PBS and once with water. Drain and allow coverslip to air dry.
9. Mount coverslip on glass slide using 10 to 20 μl Moviol 4-88 solution containing 2.5% DABCO. Carry out immunofluorescent staining, count cells, and calculate the percentage of cells expressing the protein of interest.

For example, if 10^6 cells were plated, and 1% are observed expressing protein, then the viral titer is 10^4 per coverslip. This value should then be multiplied by the dilution factor to obtain the true titer.

DABCO reduces the fading of FITC.

SUPPORT PROTOCOL

PURIFICATION OF SFV PARTICLES

To concentrate or purify the recombinant virus from the medium, the particles are sedimented by ultracentrifugation onto a sucrose cushion.

Additional Materials

Transfected BHK-21 cells (first basic protocol, step 12)
20% (w/v) and 55% (w/v) sucrose in TNE buffer
TNE buffer
Beckman SW-40 or SW-41 centrifuge tubes and rotor, or equivalent

1. Centrifuge medium from transfected BHK-21 cells 15 min at $2000 \times g$ in a tabletop centrifuge at 4°C to separate growth medium from remaining cells and cell debris.
2. Set up a step gradient consisting of 1 ml of 55% sucrose and 3 ml of 20% sucrose in an SW-40 or SW-41 ultracentrifuge tube. Layer cleared medium on top of gradient (9 ml for a SW-40 tube or 8 ml for a SW-41 tube). Centrifuge 90 min at 30,000 rpm

(~160,000 × g) in a Beckman SW-40 or SW-41 rotor to sediment the virus onto the 55% cushion (see also *UNIT 5.3*).

3. Aspirate top fraction containing medium and remove (from above) 0.8 ml of the 55% sucrose from the bottom of the tube. Collect the virus band from the bottom of the tube in a total volume of 1 ml. Divide band into 50- to 100- μ l aliquots, diluting 1:2 or more with TNE buffer if desired. Quickly freeze on dry ice or in liquid nitrogen and store at -80°C .

Aliquoting is recommended because repeated freezing and thawing can markedly reduce the infectivity of the virus preparation.

REAGENTS AND SOLUTIONS

α -chymotrypsin solution

Prepare in PBS (see recipe):

10 mM MgCl_2

20 mM CaCl_2

10 mg/ml α -chymotrypsin

Store several months at -20°C

Chase medium

MEM (GIBCO/BRL)

2 mM glutamine

20 mM HEPES

150 $\mu\text{g/ml}$ methionine

Store ≤ 1 month at 4°C

Complete BHK-21 medium

G-MEM (GIBCO/BRL)

5% (v/v) fetal calf serum

10% (w/v) tryptose phosphate

20 mM HEPES

2 mM glutamine

0.1 U/ml penicillin (optional)

0.1 $\mu\text{g/ml}$ streptomycin (optional)

Store ≤ 1 month at 4°C

Minimum essential medium (MEM), supplemented

Prepare MEM (GIBCO/BRL) containing 0.2% (w/v) BSA, 2 mM glutamine, and 20 mM HEPES. Store ≤ 1 month at 4°C .

NP-40 lysis buffer

1% (v/v) Nonidet P-40 (NP-40; from 10% stock)

50 mM Tris-Cl, pH 7.6

150 mM NaCl

2 mM EDTA

Store several months at 4°C

PBS-Eisen

0.257 g NaH₂PO₄·H₂O
2.250 g Na₂HPO₄·2H₂O
8.767 g NaCl
H₂O to 1000 ml
Store indefinitely at 4°C

Phosphate-buffered saline (PBS)

2.68 mM KCl
1.47 mM KH₂PO₄
136.8 mM NaCl
8.0 mM Na₂HPO₄·7H₂O
Adjust pH to 7.0 to 7.2 with HCl
Autoclave 30 min
Store indefinitely at 4°C

PBS is also available commercially (GIBCO/BRL) and other formulations exist (e.g., see APPENDIX 2); however, not all can be successfully used for these protocols. PBS for these applications must be without MgCl₂ or CaCl₂.

rNTP mix

10 mM ATP
10 mM CTP
10 mM UTP
5 mM GTP
Aliquot and store indefinitely at -20°C

SP6 RNA polymerase buffer, 10×

400 mM HEPES·KOH, pH 7.4
60 mM magnesium acetate
20 mM spermidine·HCl
Store indefinitely at -20°C

Starvation medium

Methionine-free MEM (GIBCO/BRL)
2 mM glutamine
20 mM HEPES
Store ≤1 month at 4°C

TD solution, 5×

20% (w/v) Ficoll 400
25 mM EDTA, pH 8.0
0.05% (w/v) bromphenol blue
0.03% (w/v) xylene cyanol
Store several months at 4°C

TNE buffer

50 mM Tris·Cl, pH 7.4
100 mM NaCl
0.5 mM EDTA
Store indefinitely at room temperature

Xgal stain solution

Prepare in PBS:

5 mM $K_4Fe(CN)_6 \cdot 3H_2O$ (potassium ferrocyanide)

5 mM $K_3Fe(CN)_6$ (potassium ferricyanide)

2 mM $MgCl_2$

Store several months at 4°C

Just before use, add Xgal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside; dimethylformamide stock) to 1 mg/ml.

CAUTION: Avoid coming into contact with or inhaling cyanide and discard waste using appropriate precautions.

This recipe is based on Sanes et al. (1986); see also UNIT 9.11 for further discussion.

Z buffer

60 mM Na_2HPO_4

40 mM NaH_2PO_4

10 mM KCl

1 mM $MgSO_4$

40 mM 2-mercaptoethanol (2-ME)

Store \leq 1 month at 4°C

COMMENTARY

Background Information

Alphavirus replication. Alphaviruses such as Semliki Forest virus (SFV) are small-enveloped viruses with single-stranded RNA genomes of positive polarity that replicate in virtually all animal cells. Upon infection, the 5' end of the viral genome is used for translation of four nonstructural proteins (*nsP1-4*). These are responsible for the production of genome-length negative strands, which are subsequently used for production of new positive-strand genomes as well as a smaller subgenome species. The latter corresponds to the last one-third of the viral genome and is used as the mRNA for virus structural proteins—i.e., the capsid protein, C, and the two membrane proteins, p62 and E1 (Schlesinger and Schlesinger, 1990). The C protein complexes with new viral genomes to form cytoplasmic nucleocapsid structures, while the two membrane proteins are translocated into the endoplasmic reticulum to form heterodimers; these heterodimers are routed to the cell surface, where virus budding occurs. The membrane-protein spikes mediate virus binding and uptake into new host cells by receptor-mediated endocytosis. Inside the endosomes, the acidic milieu induces structural changes in the spike protein, thereby inducing fusion of the viral and endosomal membranes. This results in release of the nucleocapsid with the viral genome into the cell cytoplasm (Bron et al., 1992; Garoff et al., 1990; Liljeström and Garoff, 1991a; Lobigs and Garoff, 1990; Lo-

bigs et al., 1990a, b; Salminen et al., 1992; Suomalainen et al., 1992; Wahlberg and Garoff, 1992; Wahlberg et al., 1989, 1992).

Alphaviruses as expression vectors. Several different alphavirus genomes have recently been cloned as cDNAs and used to produce infectious RNA by in vitro transcription (Davis et al., 1989; Kuhn et al., 1991; Liljeström et al., 1991; Rice et al., 1987). The SFV and Sindbis cDNAs have been further developed into general expression systems for insertion of foreign DNA in place of the viral structural genes (Bredenbeek and Rice, 1992; Liljeström, 1993; Liljeström and Garoff, 1991b; Rice, 1992; Xiong et al., 1989). The basic difference between the two systems is the type of helper used. The Sindbis system uses a helper vector containing the sequence of a Sindbis DI (defective interfering) particle, and therefore contains an RNA packaging signal. When cotransfected with Sindbis recombinant RNA into host cells, the DI-helper RNA provides structural proteins for packaging of both the recombinant genome and the DI genome into virus particles. A mixed DI/recombinant virus stock is obtained after amplification through several passages. In contrast, the SFV system uses a helper construct which, although coding for the structural proteins, cannot itself be packaged due to the lack of a proper RNA packaging signal. Its use together with SFV recombinant RNA results in production of a virus stock that contains only recombinant virus genomes; therefore,

vector structural proteins are not produced when cells are infected by these particles. High-efficiency transfection conditions such as electroporation permit the generation of a high-titer recombinant SFV stock.

Advantages of the SFV expression system.

In comparison to those of other expression systems, the SFV replication cycle exhibits many features that are advantageous for the practical performance of the expression system. (1) Extremely high levels of production can be obtained because the heterologous gene is expressed using the viral RNA replication and translation signals. (2) Recombinant RNA molecules are expressed in the host cell cytoplasm. The virus *nsp* region encodes the machinery required for RNA replication as well as capping of the 5' ends, thereby circumventing many problems that may occur in nuclear gene expression (such as limitation of transcription factors and problems with RNA splicing or RNA transport). (3) The system is very fast and easy to use, employing standard plasmid subcloning techniques. A high-titer recombinant virus stock can be produced by a single passage in electroporated cells, without need for tedious amplification or selection procedures. (4) The recombinant virus stock does not include any helper virus, and thus there is no interference of coexpressed virus structural proteins with the heterologous gene product when the stock is used in expression studies. (5) SFV has a very wide natural host range. Consequently, virus particles produced by the SFV expression system are capable of delivering a heterologous gene into probably any higher eukaryotic tissue culture cell type used today. (6) Cell morphology is preserved. Although use of the SFV expression system eventually leads to cell death, there is a variable but considerable time lag before morphological changes begin to occur (10 to 20 hr depending on the host cell line). In particular, this expression system can be used in epithelial cell cultures (MDCK) and primary neuronal cells (rat ganglion cells) within these incubation time spans without loss of differentiated morphological features.

Critical Parameters

Transfection efficiency. The success of the SFV expression system is highly dependent on the RNA transfection efficiency. Practically all BHK cells can be transfected using the conditions described in the electroporation protocol. Similar results can probably be obtained in other cell types if transfection efficiency is optimized as function of voltage, capacitance,

time constant of electrical pulse, and number of pulses (see *UNITS 9.3 & 9.9*). Under optimal conditions, transfection with Lipofectin also gives good results. Although Lipofectin can be used with a wide variety of cells, the exact optimal conditions may vary slightly. Parameters to consider are final amounts of lipid and RNA, lipid:RNA ratio, and time of lipid-RNA administration (see *UNITS 9.4 & 9.9*, detailing liposome-mediated transfection with DNA, for further discussion of Lipofectin). For in vitro expression of the recombinant protein, an important consideration is the quality of the RNA preparation (see *Critical Parameters in UNIT 10.17*).

Virus particle formation. Another factor of importance is the ability of a given cell type to support virus particle formation. This can be checked using wild-type SFV. Recombinant virus stock should be aliquoted and stored at -80°C after quick freezing in dry ice or liquid nitrogen. Original culture supernatant or step gradient-purified virus can be stored in this fashion. Repeated freezing and thawing should be avoided, as it reduces virus infectivity; pelleting of the virus is not recommended, as it also may lead to considerable (10- to 100-fold) reduction in infectivity.

Safety considerations. When using the original pSFV-Helper 1 and recombinant RNA for virus stock preparation, recombination (strand-switching by the replicase) of the RNA species may occur (Geigenmüller-Gnirke et al., 1991; Weiss and Schlesinger, 1991). It is estimated that the frequency of recombinant particles found in one packaging mixture is $\sim 10^{-6}$ (i.e., $\leq 10^3$ spreading competent SFV particles). Although this would be of little importance for the outcome of an expression experiment involving, for instance, 10^6 cells infected with 5 infectious units (IU) per cell, the existence of such particles constitutes a potential biosafety risk. Therefore, a new vector, pSFV-Helper 2, was designed for use in producing conditionally infectious particles (Berglund et al., 1993). These particles must be activated by chymotrypsin treatment to become infectious. Due to the design of the new Helper system, no replication-proficient virus has been found. The new Helper has been approved by the NIH Recombinant DNA Advisory Committee for use at Biosafety Level-2 (BL-2).

Anticipated Results

In general, 10^7 BHK cells should yield $\sim 10^9$ to 10^{10} infectious recombinant particles when

cotransfected with recombinant and pSFV-Helper 2 RNA and incubated for 24 hr. In cells transfected with ~5 IU/cell, after ~4 hr of incubation the majority of newly synthesized polypeptide chains should represent heterologous protein products, meaning that the labeled heterologous protein should be seen as the major band on SDS gel analysis and autoradiography. After 20 to 25 hr of incubation the heterologous protein product should be the predominant protein in the cell, and Coomassie brilliant blue staining of the SDS gel should reveal the heterologous protein as a major band.

Time Considerations

Recombinant RNA can be used directly for phenotype analysis in cells after transfection. In this case phenotype analysis can be done after only a 4-hr incubation. Stock virus preparation requires 24-hr incubation of cotransfected cells. The supernatant can then be used directly for infection-mediated gene transduction into new hosts.

Literature Cited

- Berglund, P., Sjöberg, M., Garoff, H., Atkins, G.J., Sheahan, B.J., and Liljeström, P. 1993. SFV expression system: Production of conditionally infectious recombinant particles. *Bio/Technology* 11:916-920.
- Bredenbeek, P.J. and Rice, C.M. 1992. Animal RNA virus expression systems. *Semin. Virol.* 3:297-310.
- Bron, R., Wahlberg, J.M., Garoff, H., and Wilschut, J. 1992. Membrane fusion of Semliki Forest virus in a model system: Correlation between fusion kinetics and structural changes in the envelope glycoprotein. *EMBO J.* 12:693-701.
- Davis, N.L., Willis, L.V., Smith, J.F., and Johnston, R.E. 1989. In vitro synthesis of infectious Venezuelan Equine Encephalitis virus RNA from a cDNA clone: Analysis of a viable deletion mutant. *Virology* 171:189-204.
- Garoff, H., Huylebroeck, D., Robinson, A., Tillman, U., and Liljeström, P. 1990. The signal sequence of the p62 protein of Semliki Forest virus is involved in initiation but not in completing chain translocation. *J. Cell Biol.* 111:867-876.
- Geigenmüller-Gnirke, U., Weiss, B., Wright, R., and Schlesinger, S. 1991. Complementation between Sindbis viral RNAs produces infectious particles with a bipartite genome. *Proc. Natl. Acad. Sci. U.S.A.* 88:3253-3257.
- Kuhn, R.J., Niesters, H.G.M., Hong, Z., and Strauss, J.H. 1991. Infectious RNA transcripts from Ross River virus cDNA clones and the construction and characterization of defined chimeras with Sindbis virus. *Virology* 182:430-441.
- Liljeström, P. 1993. Virally based transient expression systems. *Curr. Opin. Ther. Pat.* 3:375-402.
- Liljeström, P. and Garoff, H. 1991a. Internally located cleavable signal sequences direct the formation of Semliki Forest virus membrane proteins from a polyprotein precursor. *J. Virol.* 65:147-154.
- Liljeström, P. and Garoff, H. 1991b. A new generation of animal cell expression vectors based on the Semliki Forest virus replicon. *Bio/Technology* 9:1356-1361.
- Liljeström, P., Lusa, S., Huylebroeck, D., and Garoff, H. 1991. In vitro mutagenesis of a full-length cDNA clone of Semliki Forest virus: The 6000-molecular-weight membrane protein modulates virus release. *J. Virol.* 65:4107-4113.
- Lobigs, M. and Garoff, H. 1990. Fusion function of the Semliki Forest virus spike is activated by proteolytic cleavage of the envelope glycoprotein p62. *J. Virol.* 64:1233-1240.
- Lobigs, M., Wahlberg, J.M., and Garoff, H. 1990a. Spike protein oligomerization control of Semliki Forest virus fusion. *J. Virol.* 64:5214-5218.
- Lobigs, M., Zhao, H., and Garoff, H. 1990b. Function of Semliki Forest virus E3 peptide in virus assembly: Replacement of E3 with an artificial signal peptide abolishes spike heterodimerization and surface expression of E1. *J. Virol.* 64:4346-4355.
- Rice, C.M. 1992. Examples of expression systems based on animal RNA viruses: Alphaviruses and influenza virus. *Curr. Opin. Biotechnol.* 3:522-532.
- Rice, C.M., Levis, R., Strauss, J.H., and Huang, H.V. 1987. Production of infectious RNA transcripts from Sindbis virus cDNA clones: Mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. *J. Virol.* 61:3809-3819.
- Salminen, A., Wahlberg, J.M., Lobigs, M., Liljeström, P., and Garoff, H. 1992. Membrane fusion process of Semliki Forest virus II: Cleavage dependent reorganization of the spike protein complex controls virus entry. *J. Cell Biol.* 116:349-357.
- Sanes, J., Rubenstein, J.L.R., and Nicolas, J.-F. 1986. Use of recombinant retrovirus to study post-implantation cell lineage in mouse embryos. *EMBO J.* 5:3133-3142.
- Schlesinger, S. and Schlesinger, M.J. 1990. Replication of Togaviridae and Flaviviridae. In *Virology* (B.N. Fields and D.M. Knipe, eds.) pp. 697-711. Raven Press, N.Y.
- Suomalainen, M., Liljeström, P., and Garoff, H. 1992. Spike protein-nucleocapsid interactions drive the budding of alphaviruses. *J. Virol.* 66:4737-4747.
- Wahlberg, J. and Garoff, H. 1992. Membrane fusion process of Semliki Forest virus I: Low pH-induced rearrangement in spike protein quaternary structure precedes virus penetration into cells. *J. Cell Biol.* 116:339-357.

- Wahlberg, J.M., Boere, W.A., and Garoff, H. 1989. The heterodimeric association between the membrane proteins of Semliki Forest virus changes its sensitivity to mildly acidic pH during virus maturation. *J. Virol.* 63:4991-4997.
- Wahlberg, J.M., Bron, R., Wilschut, J., and Garoff, H. 1992. Membrane fusion of Semliki Forest virus involves homotrimers of the fusion protein. *J. Virol.* 66:7309-7318.
- Weiss, B.G. and Schlesinger, S. 1991. Recombination between Sindbis virus RNAs. *J. Virol.* 65:4017-4025.
- Xiong, C., Levis, R., Shen, P., Schlesinger, S., Rice, C.M., and Huang, H.V. 1989. Sindbis virus: An efficient, broad host range vector for gene expression in animal cells. *Science (Wash. DC)* 243:1188-1191.

Key References

- Liljeström et al., 1991. See above.
Describes construction of a cDNA clone of the SFV genome, which was a basis for the SFV expression vectors.
- Liljeström and Garoff, 1991b. See above.
Describes use of SFV vectors to express various proteins.
- Berglund et al., 1993. See above.
Describes use of a second-generation Helper vector for producing conditionally infectious recombinant SFV particles.

Contributed by Peter Liljeström and
Henrik Garoff
Karolinska Institute, Novum Research Center
Huddinge, Sweden

Overview of the HIV-1 Lentiviral Vector System

Replication-defective oncoretroviral vectors have been the most widely used vehicles for gene-transfer studies because of their capacity to efficiently introduce and stably express transgenes in mammalian cells (Hawley, 1996, 2001; *UNITS 9.9-9.14*). A limitation of oncoretroviral vectors, however, is that cell division is required for proviral integration into the host genome (Miller et al., 1990). By comparison, lentiviruses such as human immunodeficiency virus type 1 (HIV-1) have evolved a nuclear-import machinery that allows them to infect nondividing as well as dividing cells (Lewis et al., 1992; Lewis and Emerman, 1994). This unique property has led to the development of lentiviral vectors for gene delivery to a variety of nondividing or slowly dividing cells including neurons and glial cells of the central nervous system, cells of differentiated epithelial tissues, hepatocytes, muscle cells, retinal cells, and hematopoietic stem cells.

Since HIV-1 is a known human pathogen, the development of HIV-1-based lentiviral vectors has been especially challenging. Before they could be considered safe for general (BSL-2) laboratory use, lentiviral vector systems required numerous modifications to preclude the generation of wild-type or replication-competent retroviruses (RCR). The complex genome of HIV-1 was fully characterized to determine the minimal amount of viral sequences that could constitute a functional transfer vector. Production of replication-defective vectors is accomplished by providing only essential virion structural and enzymatic proteins *in trans*. The latest versions of HIV-1-based lentiviral vectors carry as little as 10% of the viral genomic RNA.

Lentiviral vectors based on HIV-2 (Poeschela et al., 1998), simian immunodeficiency virus (Kim et al., 2001), equine infectious anemia virus (Mitrophanous et al., 1999), feline immunodeficiency virus (Curran et al., 2000), and visna virus (Berkowitz et al., 2001) have also been developed. However, since HIV-1-based lentiviral vectors are the most advanced, they are the primary focus of this chapter. This unit is intended to provide an overview of HIV-1 molecular biology and an introduction to successive generations of HIV-1-based lentiviral vectors. In *UNIT 16.22*, detailed protocols outlining the methodology and techniques in-

involved in the construction and application of HIV-1-based lentiviral vector systems are presented. That unit also describes procedures that can be used to concentrate and purify high-titer recombinant lentiviral vector preparations, as well as protocols for transduction of adherent and suspension cells.

MOLECULAR BIOLOGY OF HIV-1

Like other members of the *Retroviridae* family of viruses, HIV-1 reverse transcribes its RNA genome into a double-stranded DNA form, hence the prefix “retro.” HIV-1 belongs to the subfamily *Lentivirinae*, which comprises primate and nonprimate retroviruses that cause slow, progressive diseases affecting the immune system. In humans, HIV-1 infection ultimately leads to the development of acquired immunodeficiency syndrome (AIDS). The molecular biology of HIV-1 has been reviewed extensively elsewhere (Frankel and Young, 1998). The salient features of the structure and life cycle of HIV-1 relevant to lentiviral vector-mediated gene transfer are summarized below.

HIV-1 Genes

The HIV-1 genome (Fig. 16.21.1) is an ~9.3-kb RNA that encodes nine open reading frames (ORFs). Three of these ORFs, *gag*, *pol*, and *env*, which are common to all retroviruses, specify the Gag, Pol, and Env polyproteins, respectively (Table 16.21.1). These polyproteins are subsequently cleaved into the following protein subunits: Gag is cleaved to form the structural proteins matrix (MA), capsid (CA), nucleocapsid (NC), P6, P2, and P1; Pol is cleaved into the replication enzymes protease (PR), reverse transcriptase (RT), and integrase (IN); and Env (or gp160) is cleaved to form the transmembrane (TM or gp41) and surface (SU or gp120) glycoproteins required for viral binding and entry into the cells. Of the remaining six genes, *tat* and *rev* code for the regulatory proteins transactivator of transcription (Tat) and regulator of expression of virion proteins (Rev). The other four genes *vif*, *vpr*, *vpu*, and *nef*, which are referred to as accessory genes since they are not essential for viral replication, encode the virion infectivity factor (Vif), the viral protein R (Vpr), the viral protein U (Vpu), and the negative factor (Nef), respectively. The accessory genes are also called the virulence factors,

Table 16.21.1 HIV-1 Genes, Gene Products, and Their Function

Gene	Size of mRNA	Encoded protein(s)	Function
Regulatory genes			
<i>tat</i>	1.7-2 kb	Tat (16 kDa)	<i>Trans</i> -activation of gene expression
<i>rev</i>	1.7-2 kb	Rev (19 kDa)	Nuclear export of late mRNAs; promotion of polysomal binding to RRE-containing RNAs
Accessory genes			
<i>vif</i>	5 kb	Vif (23 kDa)	Enhancement of virus transmission
<i>vpr</i>	4.5 kb	Vpr (10-15 kDa)	Nuclear transport of proviral DNA, induction of G ₂ arrest in dividing cells
<i>vpu</i>	4 kb	Vpu (15-20 kDa)	CD4 degradation; virus maturation and release
<i>nef</i>	1.7-2 kb	Nef (25-27 kDa)	CD4 and MHC-I down-regulation; enhancement of virus replication
Structural genes			
<i>gag</i>	9.3 kb	Pr55 ^{gag} : MA (p17), CA (p24), NC (p9), p6	Formation of viral particles, packaging of viral genomic RNA
<i>pol</i>	9.3 kb	Pr160 ^{gag-pol} : PR (p10), RT (p61/p52), IN (p31)	Reverse transcription, integration, and virion maturation
<i>env</i>	4 kb	gp160:SU (gp120), TM (gp41)	Binding and entry into the host cell

as they are associated with HIV-1 virulence and pathogenesis in vivo (Cullen, 1991).

HIV-1 Virion

The HIV-1 virion (Fig. 16.21.1B) has a diameter of ~110 nm and consists of a nucleocapsid core surrounded by a lipid membrane (Aloia et al., 1988). Inserted into the lipid membrane are the viral SU and TM glycoproteins which are noncovalently linked (Willey et al., 1988). Proteins within the inner shell of a mature virion are the cleavage products of the Gag (Pr55^{gag}) and Gag-Pol (Pr160^{gag-pol}) precursors (Mervis et al., 1988). The condensed inner core is formed by the capsid protein (CA), p24. In the space between the inner core and the lipid membrane is the N-terminal myristylated matrix protein (MA), p17, which remains associated with the lipid membrane. The virion core contains two copies of the single-stranded genomic RNA to which the NC protein is bound. Also packaged into the virions are host tRNA₃^{Lys}, RT, PR, IN, Vif (Liu et al., 1995), and Vpr (Cohen et al., 1990).

HIV-1 Life Cycle

A schematic representation of the HIV-1 life cycle is shown in Figure 16.21.2. The HIV-1 replication cycle can be divided into the following steps: (i) virus entry; (ii) reverse transcription; (iii) nuclear localization and integration; (iv) viral RNA transcription; (v) RNA processing and viral protein synthesis; and (vi) viral particle assembly, release and maturation.

Virus entry

HIV-1 infection is initiated by the binding of gp120 to the CD4 receptor which is found on the surface of T helper lymphocytes, macrophages, and glial cells (Dalglish et al., 1984; Maddon et al., 1986). Binding of gp120 to the CD4 receptor induces conformational changes that expose the N-terminal hydrophobic domain of gp41. This hydrophobic domain is then inserted into the cell membrane, initiating fusion in a pH-independent manner (Stein et al., 1987). Fusion of gp41 with the cell membrane has been shown to utilize a cellular coreceptor. The chemokine receptors CCR-5 (Alkhatib et al., 1996; Choe et al., 1996) and CXCR-4 (Feng et al., 1996) are the major coreceptors required for entry of macrophage-tropic and T cell-tropic strains of HIV-1, respectively.

Reverse transcription

Following entry, the outer lipid membrane of the virion is removed and its core is delivered into the cytoplasm. Within the core, genomic RNA is reverse transcribed by RT into double-stranded DNA using tRNA₃^{Lys} as a primer (Goff, 1990; Das et al., 1994). The tRNA₃^{Lys} primer allows first strand DNA synthesis by binding to the PBS (primer binding site) located near the 5' end of the HIV-1 genomic RNA. As the minus-strand DNA synthesis proceeds to the 5' end, RNase H degrades the template from the RNA/DNA hybrid allowing the minus-strand to undergo an intermolecular strand switch ("jump") to complementary sequences

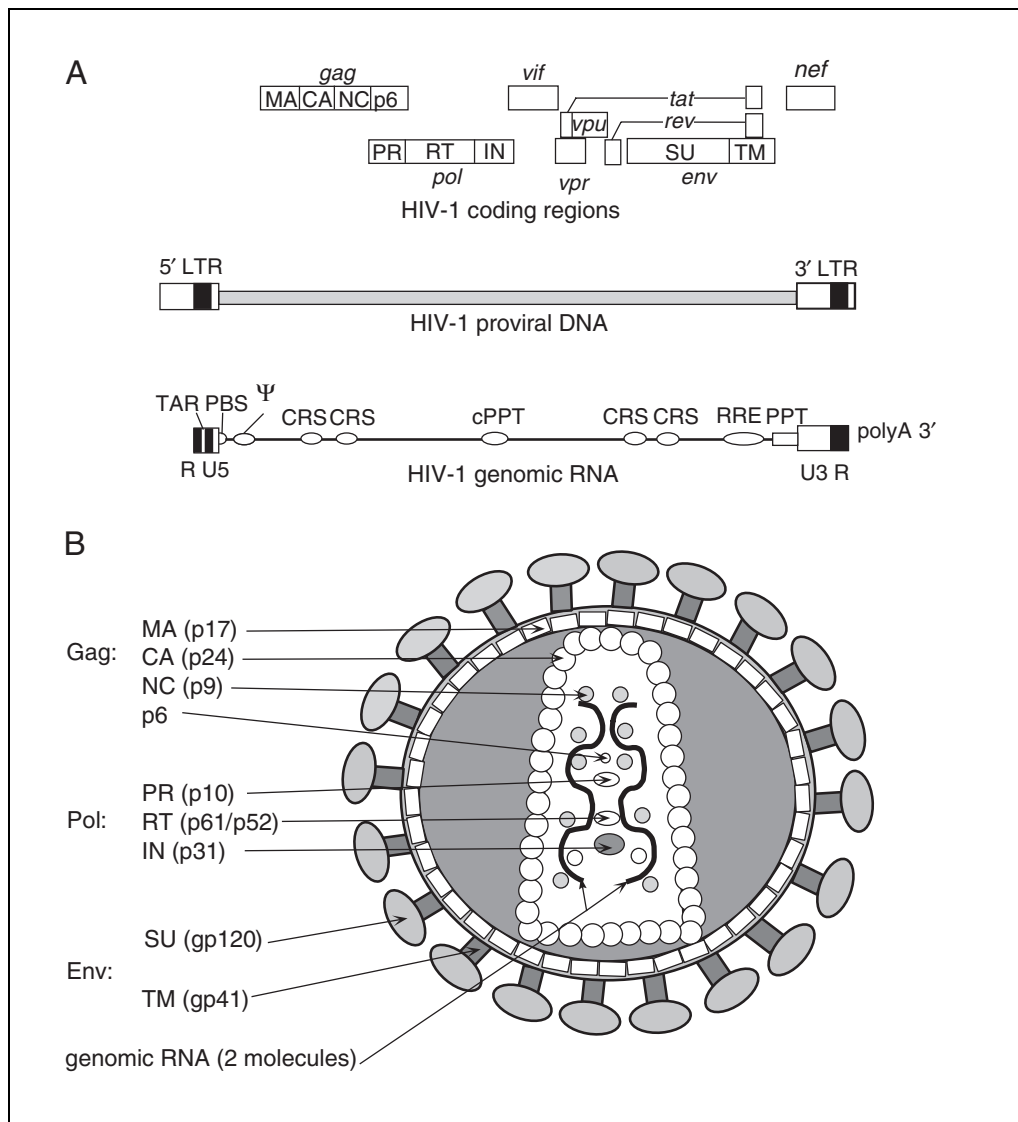


Figure 16.21.1 (A) Structural features of HIV-1 proviral DNA and genomic RNA. The proviral DNA is flanked by LTRs which control the expression of the viral genes. The relative positions of genes that encode the structural and enzymatic proteins (Gag, Pol and Env), the regulatory proteins (Tat and Rev), and the accessory proteins (Vif, Vpu, Vpr and Nef) are indicated. The various *cis*-acting responsive sequences (CRS) within the *gag*, *pol*, and *env* coding regions that negatively regulate HIV-1 gene expression are shown on the HIV-1 genomic RNA. (B) Schematic diagram of the HIV-1 virion. The mature virion contains an inner core surrounded by a host-cell derived lipid membrane. Inserted within the lipid membrane are the viral Env glycoproteins. The condensed inner core is made of the capsid protein. The space between the inner core and the lipid membrane contains the MA protein, which remains attached to the lipid membrane. The inner core encompasses two copies of the genomic RNA which are bound by NC proteins. Also packaged within the viral particles are the RT, PR, IN, Vif and Vpr proteins, as well as the primer tRNA₃^{Lys}, which is of the host origin.

at the 3' end of viral RNA (Panganiban and Fiore, 1988). RNase H activity also generates short template fragments at the 3' polypurine tract (PPT) located upstream of the unique 3' (U3) sequence and an internal PPT located at the 3' end of the *pol* gene (central PPT or cPPT) which serve as primers for plus-strand DNA synthesis (Charneau et al., 1992). A second jump (intramolecular) between the PBS se-

quences allows RT to extend to the ends of the templates, leading to complete synthesis of double-stranded HIV-1 DNA containing flanking long terminal repeats (LTRs).

Nuclear localization and integration

The newly synthesized linear double-stranded DNA is then translocated to the nucleus in the absence of mitosis while still part

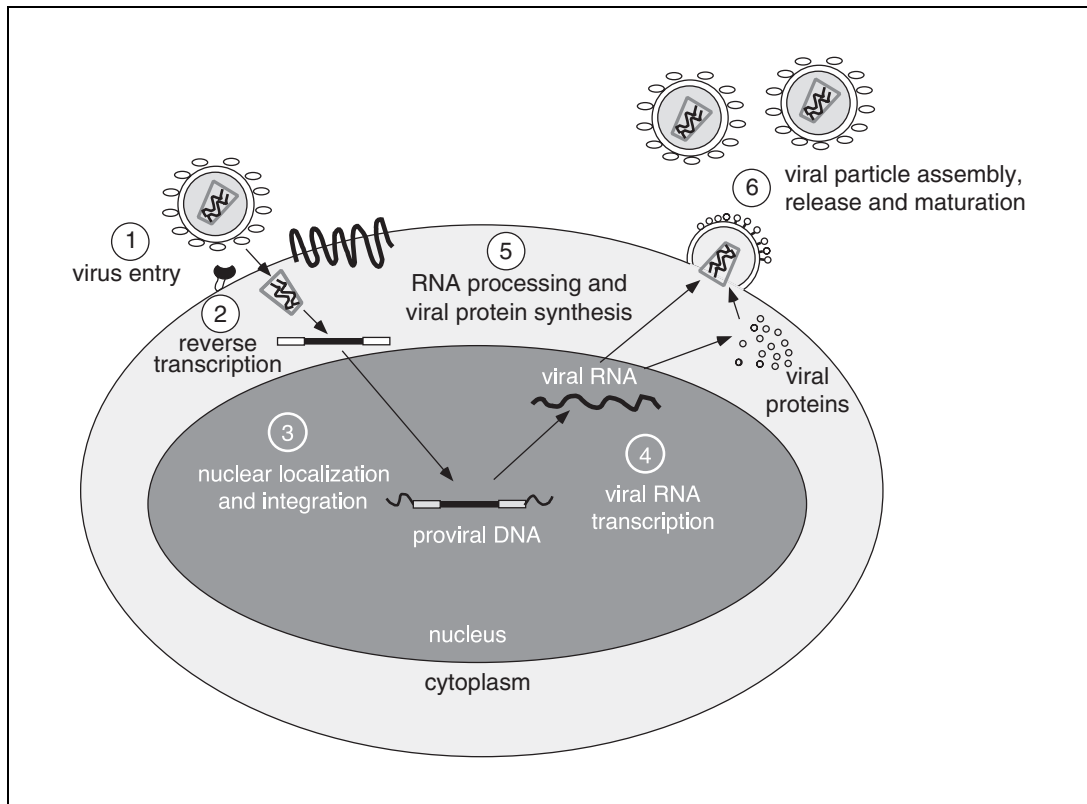


Figure 16.21.2 Schematic diagram of the HIV-1 life cycle. Following entry into the cell, HIV-1 genomic RNA is reverse transcribed; the DNA intermediate, as part of a nucleoprotein complex, is then transported to the nucleus where it is inserted into the cellular genome. The full-length 9.3-kb viral RNA is differentially spliced to give rise to various HIV-1 mRNAs. Virus assembly takes place at the plasma membrane, where two copies of HIV-1 genomic RNA are incorporated into the budding viral particles. The steps indicated by numbers 1 to 6 are described in the text.

of the nucleoprotein complex, which contains the viral proteins Vpr, RT, MA and IN (Farnet and Haseltine, 1991; Bukrinsky et al., 1993). Nuclear translocation is mediated via interaction with the cellular nuclear import pathway (Trono, 1995). Interaction involves Vpr, MA and IN, all of which contain nuclear localization signals (Heinzinger et al., 1994). As described below, studies of lentiviral vectors identified a “central DNA flap” formed during reverse transcription as a result of the cPPT that facilitates nuclear translocation of the HIV-1 preintegration complex in both dividing and nondividing cells (Follenzi et al., 2000; Zenou et al., 2000).

Within the nucleus, the linear double-stranded DNA is integrated into the host chromosome by the action of viral IN to form the provirus. IN recognizes short inverted termini at the ends of the viral DNA and cleaves a dinucleotide from each 3' end (LaFemina et al., 1991). The chromosomal DNA is also cleaved by IN, which subsequently ligates the 5' ends of chromosomal DNA to the 3' ends of the

resulting proviral structure (Bushman and Craigie, 1991).

Transcription

Two identical direct 634-base pair (bp) repeats—the LTR sequences—flank the provirus at either end. The LTRs contain the enhancer and promoter elements for transcription of viral genes. Each LTR consists of the U3 region, the R (repeat) region and the unique 5' (U5) region. The U3 region contains binding sites for several cellular transcription factors including SP-1 and NF- κ B (Cullen, 1991). The R region of the LTR contains the 5' untranslated leader sequence, which provides the binding site for a number of factors involved in transcription. In particular, the 5' untranslated leader sequence contains the *trans*-activation response (TAR) element, which is a 59-nt-long region forming a stem-loop structure at the 5' noncoding region of all HIV-1 RNAs (Selby et al., 1989). Transcription of the HIV-1 provirus initiates at the 5' end of the R region within the 5' LTR (Cullen, 1991), with the site of polyadenylation deline-

ating the 3' end of the R region within the 3' LTR (Valsamakis et al., 1991).

High-level transcription from the HIV-1 provirus requires the viral Tat protein (Feng and Holland, 1988). Activation of HIV-1 transcription by Tat relies on the TAR RNA element. In the absence of Tat, early transcriptional events generate mostly short transcripts. Interaction of Tat with TAR enhances formation of RNA polymerase II complexes. As the Tat protein accumulates, the efficiency of transcriptional elongation is enhanced as it suppresses random terminations and increases the processivity of RNA polymerase II (Keen et al., 1996). In addition to enhancing the transcription of viral RNAs, Tat is also known to block the expression of major histocompatibility complex (MHC) class II genes in HIV-1-infected cells (Kanazawa et al., 2000).

RNA processing and viral protein synthesis

Newly transcribed HIV-1 viral RNA is doubly spliced by the cellular splicing machinery, producing 1.7- to 2.0-kb mRNAs which code for the Tat, Rev, and Nef proteins. The Rev protein regulates transport of HIV-1 transcripts from the nucleus to the cytoplasm by binding to the Rev response element (RRE) located in the *env* coding region (Daly et al., 1989). By binding to RRE, Rev allows the RRE-containing viral RNAs to access a pre-existing cellular-export pathway. The nuclear-export signal of the Rev protein (Fischer et al., 1995; Wen et al., 1995) recognizes and binds to the nuclear-export factor exportin 1 (Ullman et al., 1997). As Rev accumulates in the nucleus, viral RNA transport from the nucleus to the cytoplasm is enhanced. This in turn increases the probability that the viral RNA will escape the splicing machinery, resulting in nuclear export of 4- to 5-kb singly spliced and 9.3-kb unspliced viral RNAs.

The Nef protein is encoded by an open reading frame that overlaps the 3' LTR. Nef is a myristylated plasma membrane-associated protein that is expressed early during virus infection from several multiply spliced mRNAs (Robert-Guroff et al., 1990). Nef reduces the cell surface levels of CD4 (Aiken et al., 1994). This may permit increased incorporation of Env into virions, enhancing viral replication and pathogenesis (Aiken and Trono, 1995). Nef has also been shown to down-regulate cell surface expression of MHC class I molecules (Swann et al., 2001).

Once in the cytoplasm, the singly spliced viral mRNAs are translated to produce the Vif, Vpr, Vpu and Env proteins (Malim et al., 1989). The Vif protein is localized at the inner face of the cytoplasmic membrane of infected cells (Goncalves et al., 1994). It is also incorporated within viral particles in association with the nucleoprotein complex (Liu et al., 1995; Khan et al., 2001). Vif plays a critical role in the production of infectious virions by affecting viral particle assembly or maturation (Gabuzda et al., 1992; Kishi et al., 1992).

Both Vpu and Env proteins are expressed from a singly spliced bicistronic mRNA due to leaky ribosome scanning at the upstream Vpu initiation codon (Schwartz et al., 1992). The viral Env protein is translated as a gp160 precursor polyprotein. Gp160 is directed to the endoplasmic reticulum by its hydrophobic N-terminal signal peptide where it is extensively glycosylated (Robey et al., 1985). Subsequently, gp160 is cleaved within the endoplasmic reticulum–Golgi apparatus by cellular proteases (Stein and Engleman, 1990) into the SU (gp120) and TM (gp41) subunits (Freed et al., 1989), which are then inserted within the plasma membrane. Vpu is a hydrophobic membrane-associated protein, which mediates degradation of CD4 complexes in infected cells (Chen et al., 1996), crucial for maturation and release of fully infectious viral particles (Klimkait et al., 1990).

The unspliced HIV-1 viral RNA is translated into the Pr55^{gag} and Pr160^{gag-pol} polyprotein precursors. The *gag* and *pol* ORFs overlap by 241 nucleotides, with the *pol* ORF in the –1 translational reading frame with respect to the *gag* ORF. The translation of the *pol* ORF thus occurs as a consequence of ribosomal frameshifting, resulting in the synthesis of a Gag-Pol fusion protein, Pr160^{gag-pol} (Jacks et al., 1988). The N-termini of both Pr55^{gag} and Pr160^{gag-pol} polyprotein precursors are myristylated, which mediates their transport to the cellular membrane where they are packaged into the viral particles (Bryant and Ratner, 1990). Within the viral particles, the viral PR cleaves Pr55^{gag} to generate the mature virion structural proteins, MA, CA, and p15 (Peng et al., 1989). Subsequently, p15 is cleaved to produce the NC (p9) and p6 proteins (Sheng and Erickson-Viitanen, 1994). The Pr160^{gag-pol} polyprotein is also cleaved within the mature viral particles to produce PR, RT, and IN by the viral PR, which is itself derived from the Pr160^{gag-pol} precursor (Debouck, 1991).

Viral particle assembly, release, and maturation

HIV-1 virus assembly initiates with interactions between the Pr55^{gag} and Pr160^{gag-pol} polyproteins, producing an immature virion core. In addition to serving as mRNA for Pr55^{gag} and Pr160^{gag-pol} synthesis, the unspliced 9.3-kb RNA also serves as HIV-1 genomic RNA. For efficient incorporation of HIV-1 genomic RNA into viral particles, a ψ packaging signal located within the 5' untranslated region and the *gag* ORF plus a 1.1-kb RNA sequence within the *env* ORF are required (Lever et al., 1989). Dimerization of HIV-1 genomic RNAs is mediated by the annealing of palindromic sequences within the dimer linkage structure located at the 5' end of the RNA (Skripkin et al., 1994). Packaging of two copies of HIV-1 genomic RNA into viral particles then ensues, a process that involves an interaction between the NC domains of the Pr55^{gag} and Pr160^{gag-pol} protein precursors and stem-loop structures within the ψ packaging signal (Clever et al., 1995; Rice et al., 1995). The RT and NC domains of the Pr160^{gag-pol} polyprotein facilitate the incorporation of the cellular tRNA₃^{Lys} primer into the virions (Mak et al., 1994). PR is activated during the final stages of budding within the newly released immature particles, leading to proteolytic processing of the Pr55^{gag} and Pr160^{gag-pol} polyproteins (Peng et al., 1989) and viral particle maturation.

REPLICATION-DEFECTIVE HIV-1-BASED VECTORS

As with oncoretroviral vectors, generation of replication-defective lentiviral vectors requires segregation of *cis*-acting sequences necessary for the transfer of a functional viral genome to target cells and those sequences encoding essential viral structural and enzymatic proteins onto separate plasmids. The transfer vector consists of *cis*-acting sequences—the LTRs, the PBS, the packaging signal, the PPTs and the RRE—linked to a transgene of interest in the context of a transcriptional unit. Following cotransfection of the transfer vector together with packaging and envelope expression plasmids lacking most if not all of the *cis*-acting sequences into an appropriate “recipient” cell, viral proteins provided in *trans* assemble into virions encapsidating the replication-defective transfer vector RNA.

First-Generation HIV-1 Vector Systems

Although earlier HIV-1-derived vector configurations have been reported in the literature

(Page et al., 1990; Landau et al., 1991; Poznan-sky et al., 1991), only the more recent safety-modified versions of HIV-1 vector systems will be described here. An example of a “first generation” HIV-1 vector system comprising three expression plasmids—a transfer vector, a packaging construct, and an envelope gene—is depicted in Figure 16.21.3A (Naldini et al., 1996a). The transfer vector contains intact HIV-1 LTRs and the *cis*-acting sequences described above—including 350-bp of *gag* encompassing the ψ packaging signal and *env* sequences encompassing the RRE—with the exception of the cPPT. The transgene is expressed from an internal human cytomegalovirus (CMV) immediate early region enhancer-promoter. The packaging plasmid encodes all of the HIV-1 proteins except Vpu and Env. It consists of the HIV-1 genome with the following modifications: the 5' LTR was replaced with the CMV promoter to drive expression of the viral proteins required in *trans*; the ψ packaging signal, the *env* gene, and the ORF encoding the Vpu protein were deleted; and the 3' LTR was replaced at the end of the *nef* ORF with a polyadenylation [poly(A)] site from the insulin gene. In the envelope-encoding plasmid, the G glycoprotein gene of vesicular stomatitis virus (VSV-G) is transcribed from the CMV promoter. Transient cotransfection of human embryonic kidney 293T cells (DuBridge et al., 1987) with the three-plasmid combination generates replication-defective VSV-G-pseudotyped vector particles with titers on the order of 10⁵ transducing units (TU)/ml. The lentiviral particles efficiently transduced nondividing cells including human HeLa cells growth-arrested at G₁-S or G₂ phases of the cell cycle, 4-day contact-inhibited rat 208F fibroblasts, and terminally differentiated adult rat neurons.

Second-Generation HIV-1 Vector Systems

It was shown subsequently that none of the four HIV-1 accessory genes *vif*, *vpr*, *vpu*, or *nef* were required for HIV-1 replication in immortalized cell lines (Miller and Sarver, 1997) or for efficient generation of VSV-G-pseudotyped vector particles (Zufferey et al., 1997). This led to the development of a “second generation” of HIV-1 vector systems (Fig. 16.21.3B), which utilized a multiply attenuated packaging construct containing only the HIV-1 *gag*, *pol*, *rev*, and *tat* genes. Although elimination of all four accessory genes had no effect on the titers of VSV-G-pseudotyped vector particles produced in 293T cells, macrophage transduction effi-

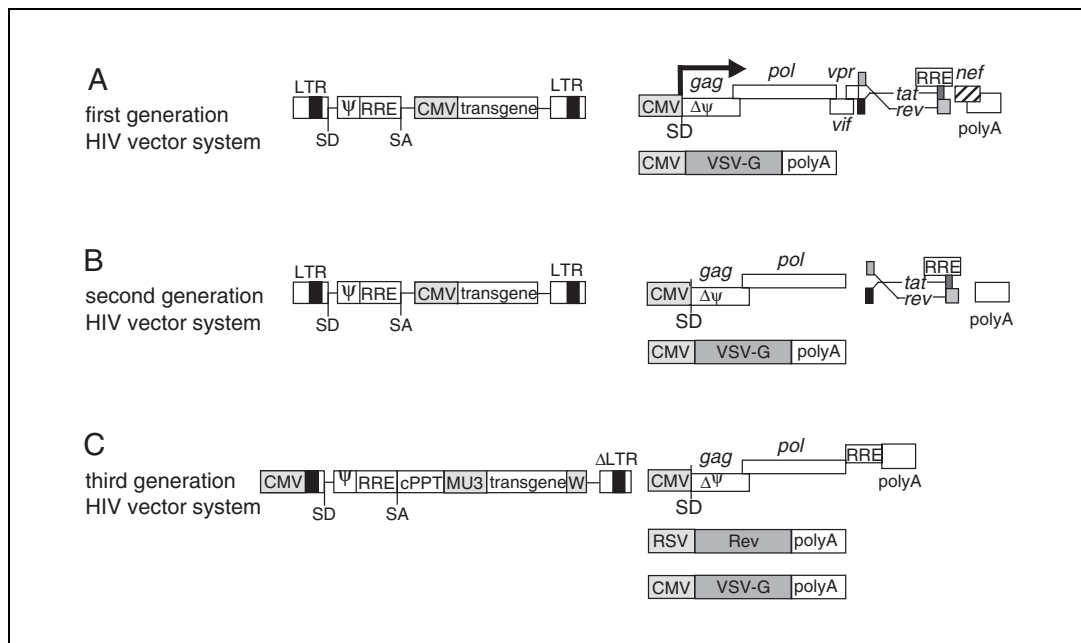


Figure 16.21.3 Schematic representation of various HIV-1 vector systems. **(A)** A first-generation HIV-1 vector system expressing all of the HIV-1 proteins except Vpu and Env. The vector particles are pseudotyped with the VSV-G envelope glycoprotein expressed from a separate plasmid. **(B)** A second-generation HIV-1 vector system having all of the accessory genes removed. **(C)** A third-generation Tat-independent HIV-1 vector system. The SIN transfer vector illustrated contains the cPPT for efficient nuclear import and utilizes an internal MSCV LTR promoter (MU3) and the WPRE (W) for high-level transgene expression. The two packaging constructs encode only the HIV-1 Gag, Pol and Rev proteins.

iciencies were reduced by ~50%, which was apparently due to the lack of the Vpr protein (Zufferey et al., 1997). Efficient transduction of resting hepatocytes, skin fibroblasts, and lymphocytes has also been reported to require HIV-1 accessory proteins (Kafri et al., 1997; Gasmi et al., 1999; Chinnasamy et al., 2000), suggesting that some of these accessory proteins might be necessary for efficient gene transfer of certain cell types.

Third-Generation HIV-1 Vector Systems

The main concern regarding the use of any HIV-1 vector is the potential for generation of replication-competent retroviruses (RCRs). Numerous additional modifications have therefore been made to ensure the biosafety of current “third” generation HIV-1-based delivery systems (Dull et al., 1998; Kim et al., 1998; Miyoshi et al., 1998; Zufferey et al., 1998; also see Fig. 16.21.3C). Since the U3 region of the 3′ HIV-1 LTR serves as a template for the U3 regions of both LTRs in the resulting provirus, third-generation HIV-1 transfer vectors have been developed that contain a 400-bp deletion within this region. The consequence of this modification is that the 5′ LTR of the integrated

vector is almost completely inactivated (Miyoshi et al., 1998; Zufferey et al., 1998). Inability to efficiently transcribe full-length vector RNA by these so-called self-inactivating (SIN) HIV-1 vectors in transduced target cells considerably minimizes the possibility of RCR generation. The SIN configuration also reduces the possibility of oncogene activation by promoter insertional mutagenesis. By design, SIN vectors require an internal promoter for transgene expression. Fortunately, in contrast to what was observed previously for SIN oncoretroviral vectors, this modification does not result in a significant drop in vector titers.

To prevent reconstitution of U3 sequences within the deleted 3′ HIV-1 LTR by homologous recombination with the intact 5′ HIV-1 LTR during transient cotransfection of 293T cells, the U3 region of the 5′ LTR of third-generation transfer vectors has been replaced with the CMV or Rous sarcoma virus LTR promoters (Dull et al., 1998; Kim et al., 1998). Notably, titers as high as 10^7 TU/ml can be obtained from the chimeric LTRs irrespective of whether Tat is present, allowing deletion of the *tat* gene from the packaging construct. This finding permitted further refinement of the packaging system such that the *gag-pol* and *rev* genes can be

expressed from two separate nonoverlapping plasmids (Dull et al., 1998).

Efforts to eliminate the Rev dependence of HIV-1-based lentiviral vectors have generally involved replacing the RRE with autonomous RNA export signals termed constitutive RNA transport elements—typically from the simian type D retroviruses, Mason-Pfizer monkey virus, and the simian retroviruses type 1 and 2 (Bray et al., 1994; Ernst et al., 1997; Kim et al., 1998; Gasmi et al., 1999; Kotsopoulou et al., 2000; Mautino et al., 2000). Collectively, the studies have indicated that these elements can partially substitute for the Rev/RRE combination, although there is usually a reduction in vector titer when the Rev/RRE components are completely removed from the packaging system. For this reason, third-generation HIV-1 vectors generally retain the RRE sequence.

As noted above, after HIV-1 reverse transcription a 99 nucleotide-long overlap—the central DNA flap—is formed that is involved in normal import of the HIV-1 preintegration complex into the nucleus (Follenzi et al., 2000; Zennou et al., 2000). The sequences specifying the central DNA flap—a cPPT and a central termination sequence located within the *pol* ORF—were omitted from earlier generations of HIV-1 vectors. These sequences have been restored in more recent versions of third-generation HIV-1 vectors and shown to facilitate higher transduction efficiencies of several types of cells, both proliferating and growth-arrested.

For high-level transgene expression in a broad range of cell types *in vitro* and *in vivo*, a variety of viral and cellular promoters have been inserted into third generation HIV-1 vector backbones. In addition to the CMV promoter, other promoters that have been demonstrated to perform well include the murine stem cell virus (MSCV) vector LTR (Hawley et al., 1994), the gibbon ape leukemia virus LTR, the human elongation factor 1 α promoter, the CAG promoter (composed of the CMV immediate early enhancer linked to chicken β -actin promoter sequences) and the human X chromosome phosphoglycerate kinase 1 promoter (Ramezani et al., 2000).

Finally, other genetic elements have been identified that stimulate transgene expression post-transcriptionally. Chief among these is the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE). This element has been postulated to exert a positive influence on transgene expression by augmenting 3'-end processing and polyadenylation (Loeb et al., 2000). Inclusion of the WPRE has been shown

to significantly enhance transgene expression levels from several types of promoters in both oncoretroviral and lentiviral vectors, and is thus routinely included in third-generation HIV-1-based gene delivery systems (Zufferey et al., 1999; Deglon et al., 2000; Ramezani et al., 2000).

Packaging Limit of HIV-1 Vectors

Another advantage of HIV-1-based vectors over oncoretroviral vectors is their greater packaging limit. The packaging limit is the length of vector RNA that can be encapsidated into viral particles. Although an HIV-1 vector greater than 12 kb in size has been reported to give reasonable titers (Uchida et al., 1998), a systematic determination of the packaging limit of lentiviral vectors demonstrated that the viral titers decrease semilogarithmically with increasing vector length (Kumar et al., 2001). In that study, VSV-G-pseudotyped HIV-1 vectors with a proviral length of 10.5 to 11 kb yielded titers comparable to that of a control vector 5.5 kb in size. The third-generation vector illustrated in Figure 16.21.3C contains ~1.9 kb of *cis*-acting HIV-1 sequences, an MSCV-based internal promoter (~0.3 kb) and the WPRE (~0.6 kb). It can therefore readily accommodate transgenes of ~8 kb.

HIV-1 Vector Packaging Cell Lines

HIV-1 lentiviral vectors with tropism limited to CD4⁺ cells have been produced from stable packaging cell lines, albeit at titers of only ~10⁴ TU/ml (Corbeau et al., 1996; Srinivasakumar et al., 1997). Because of the toxicity associated with the VSV-G envelope glycoprotein (Yang et al., 1995) as well as high-level constitutive expression of the HIV-1 PR (Kaplan and Swanstrom, 1991) and Vpr proteins (Planelles et al., 1995), the most common method of producing VSV-G-pseudotyped HIV-1 vectors has been by transient cotransfection. Titers in the range of 10⁶ to 10⁷ TU/ml can be obtained in this manner. Nonetheless, transient cotransfection of transfer, packaging, and VSV-G envelope plasmids can yield variable results and the process is labor-intensive when large amounts of vector stocks are required. Thus, although this method is satisfactory for most experimental applications, a stable producer cell line is highly desirable.

The requirement for a stable producer cell line has led to the development of inducible HIV-1-based packaging systems (Yu et al., 1996; Kaul et al., 1998; Kafri et al., 1999; Klages et al., 2000; Pacchia et al., 2001). Many

of these cell lines utilize a tetracycline-regulatable expression system to control protein synthesis from the packaging and VSV-G constructs (Fig. 16.21.4). Tetracycline-regulatable systems are based on the tetracycline resistance (*tet*) operon of *E. coli* (Gossen and Bujard, 1992). In one version of the system, commonly referred to as “tet-off” (Fig. 16.21.4A), the tetracycline repressor protein (TetR) is fused to the activation domain of the herpes simplex virus VP16 protein to generate a tetracycline-repressed transactivator (tTA). In the absence of tetracycline or tetracycline analogs such as doxycycline, tTA binds to and stimulates transcription from a minimal promoter sequence, usually derived from the human CMV immediate early region, linked to seven copies of the *tet* operon sequence (tetO). In the presence of low concentrations of tetracycline/doxycycline, tTA is prevented from binding to the promoter and transcription ceases. In the complementary “tet-on” system (Gossen et al., 1995), a “reverse” transactivator (rtTA) with the opposite properties of tTA binds to the minimal promoter and stimulates transcription only in the presence of doxycycline (Fig. 16.21.4B).

An example of a tTA-based HIV-1-packaging cell line that is capable of producing VSV-G pseudotyped vectors with titers on the order of 10^6 TU/ml for at least 3 to 4 days (Kafri et al., 1999) is illustrated in Figure 16.21.4C. The packaging cell line was generated by stable transfection of 293T cells with an expression plasmid that allows constitutive expression of tTA from the CMV promoter, a packaging construct containing all of the HIV-1 viral genes except the *env* gene under the control of a tetracycline-responsive promoter, and a VSV-G envelope plasmid also under tetracycline-responsive promoter control. More recently, a tTA-regulatable HIV-1 vector packaging cell line based on the third-generation split-genome, multiply attenuated packaging constructs has been developed (Klages et al., 2000). In this system, the HIV-1 proteins have been minimized to Gag, Pol and Rev. The Gag and Gag-Pol proteins are expressed from the CMV promoter. Advantage was taken of the fact that HIV-1 *gag/pol* transcripts are retained in the nucleus and rapidly degraded in the absence of Rev; consequently, Rev was placed under tetracycline-regulated control along with the VSV-G construct. Titers as high as 5×10^6 TU/ml can be obtained after 7 to 10 days following removal of doxycycline, and the cells can be maintained in culture for about another week.

Another group has utilized the ecdysone-inducible expression system to regulate expression of the HIV-1 Gag, Pol, and Rev proteins from a packaging construct and the VSV-G protein from a second plasmid (Pacchia et al., 2001). The system is based on the ability of the ecdysone analog, ponasterone A, to induce transcription in mammalian cells through a modified heterodimeric insect ecdysone receptor (No et al., 1996). In the presence of ponasterone A, the insect ecdysone receptor, modified by fusion to the herpes simplex virus VP16 transactivation domain, forms a heterodimer with the retinoid X receptor. The heterodimeric complex activates gene expression from a hybrid ecdysone response element linked to a minimal promoter. Compared to the tetracycline-regulatable system, the ecdysone-inducible system appears to have a lower basal level of expression and faster induction kinetics (No et al., 1996). Accordingly, the peak production of HIV-1 vector particles in this system was achieved in 3 to 5 days, although the maximum titers obtained were only 10^5 TU/ml.

TARGET CELLS OF HIV-1 VECTORS

An understanding of the molecular mechanisms mediating HIV-1 infection of nondividing cells has permitted HIV-1-derived lentiviral vectors to be designed to retain this property as well. As discussed previously, the nucleophilic properties mainly reside in the MA, IN, and Vpr proteins of the preintegration complex (Bukrinsky et al., 1993; Heinzinger et al., 1994; Gallay et al., 1997). Although the Vpr protein seems to be important for transduction of macrophages and hepatocytes (Naldini et al., 1996a; Kafri et al., 1997), it has been found to be dispensable for transduction of many other cell types (reviewed in Vigna and Naldini, 2000). Notably, however, while HIV-1 vectors do not require cell division, like HIV-1 they are unable to successfully transduce T lymphocytes that are in the G_0 stage of cell cycle. Inefficient transduction of metabolically inactive T cells is due to blocks at the levels of reverse transcription (Korin and Zack, 1998) and nuclear import (Sun et al., 1997). By comparison, T cells in the G_1/S or G_2 stages of cell cycle have permissive cytoplasmic conditions that allow HIV-1 and HIV-1-derived vectors to complete reverse transcription and nuclear localization.

As with oncoretroviral vector systems (Hawley, 1996, 2001), a major factor determining the host range and target cell type of HIV-1

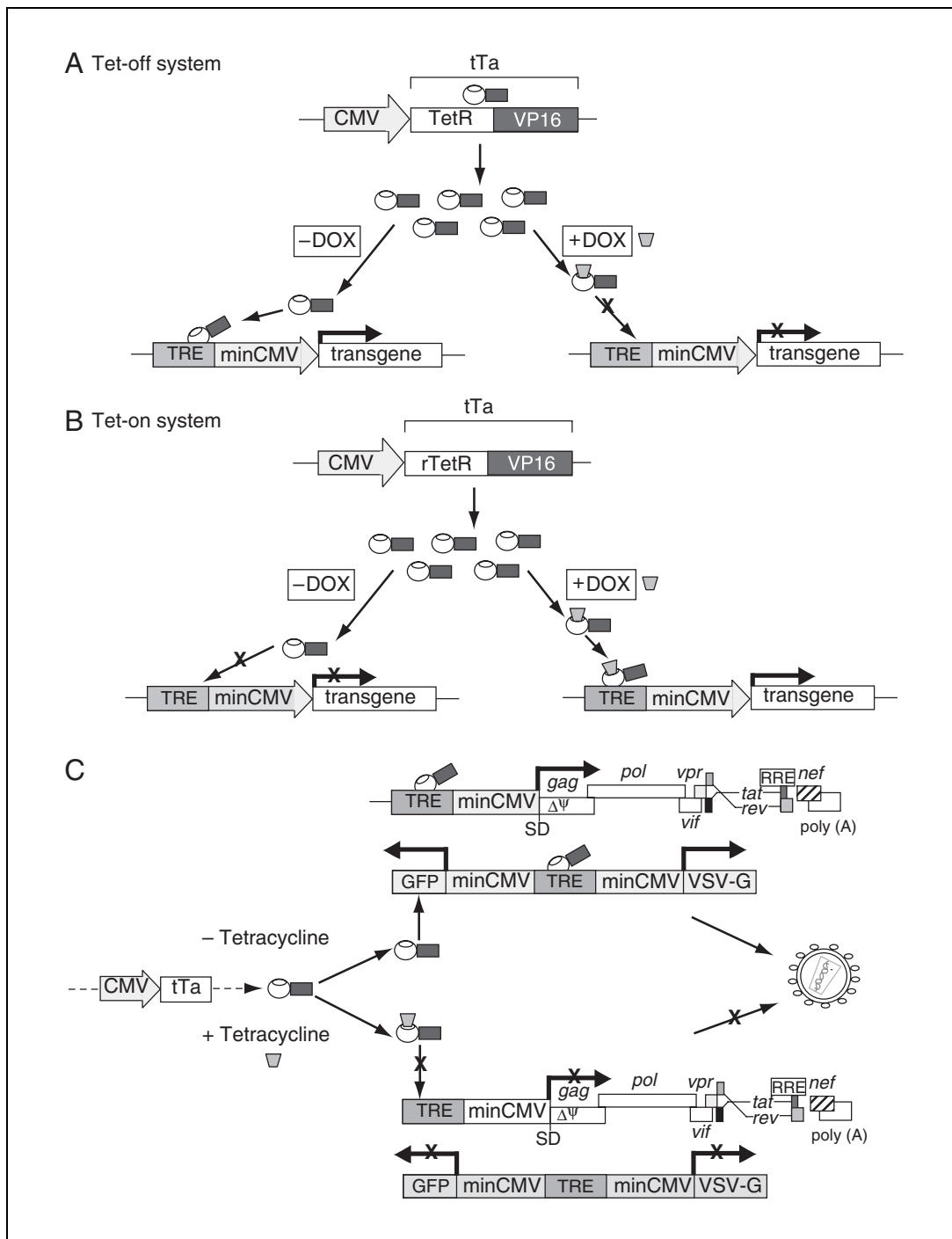


Figure 16.21.4 Schematic representation of a tetracycline-regulatable HIV-1 packaging system. **(A)** Principle of the “tet-off” gene expression system. The tetracycline-controlled transactivator (tTA) was generated by fusion of the tetracycline repressor protein (TetR) to the activation domain of the herpes simplex virus VP16 protein. In the absence of tetracycline or doxycycline, tTA binds to the tetracycline-responsive promoter element (TRE) and stimulates transcription of the transgene. **(B)** Principle of the “tet-on” gene expression system. A mutant TetR (rTetR) was created by four amino acid changes in TetR and fused to the activation domain of the herpes simplex virus VP16 protein to generate a reverse tetracycline-controlled transactivator (rtTA). rtTA binds to the TRE in the presence of doxycycline and stimulates transcription. **(C)** An example of a tetracycline-regulated HIV-1 packaging system. The packaging cells constitutively express tTA which allows expression of the HIV-1 packaging construct and the VSV-G glycoprotein envelope gene in the absence of tetracycline. In order to monitor the induction process, a green fluorescent protein (GFP) reporter gene is coexpressed with the VSV-G envelope glycoprotein gene from a bidirectional tetracycline-responsive promoter.

vectors is the envelope protein. While the Env protein of wild-type HIV-1 limits target cell tropism to CD4⁺ lymphocytes, macrophages, and glial cells, VSV-G-pseudotyped HIV-1 vectors have a greatly expanded tropism, and, with the exceptions noted above, can transduce virtually all types of cells. The expanded tropism of VSV-G-pseudotyped HIV-1 vectors is due to the ability of the VSV-G protein to bind to a phosphatidyl serine component of the lipid bilayer present in the cell membrane of most eukaryotic cells (Schlegel et al., 1983). Thus VSV-G-pseudotyped HIV-1 vectors have been used to transduce and mediate transgene expression in differentiated cells and tissues as diverse as neurons, liver, muscle, and retina (Naldini et al., 1996a,b; Kafri et al., 1997; Miyoshi et al., 1997; Kordower et al., 1999) as well as undifferentiated stem cells and embryos (Hamaguchi et al., 2000; May et al., 2000; Hawley, 2001; Pawliuk et al., 2001; Wolfgang et al., 2001; Lois et al., 2002; Pfeifer et al., 2002).

LITERATURE CITED

- Aiken, C. and Trono, D. 1995. Nef stimulates human immunodeficiency virus type 1 proviral DNA synthesis. *J. Virol.* 69:5048-5056.
- Aiken, C., Konner, J., Landau, N.R., Lenburg, M.E., and Trono, D. 1994. Nef induces CD4 endocytosis: Requirement for a critical dileucine motif in the membrane-proximal CD4 cytoplasmic domain. *Cell* 76:853-864.
- Alkhatib, G., Combadiere, C., Broder, C.C., Feng, Y., Kennedy, P.E., Murphy, P.M., and Berger, E.A. 1996. CC CKR5: A RANTES, MIP-1 α , MIP-1 β receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science* 272:1955-1958.
- Aloia, R.C., Jensen, F.C., Curtain, C.C., Mobley, P.W., and Gordon, L.M. 1988. Lipid composition and fluidity of the human immunodeficiency virus. *Proc. Natl. Acad. Sci. U.S.A.* 85:900-904.
- Berkowitz, R.D., Ilves, H., Plavec, I., and Veres, G. 2001. Gene transfer systems derived from Visna virus: Analysis of virus production and infectivity. *Virology* 279:116-129.
- Bray, M., Prasad, S., Dubay, J.W., Hunter, E., Jeang, K.T., Rekosh, D., and Hammariskjold, M.L. 1994. A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. *Proc. Natl. Acad. Sci. U.S.A.* 91:1256-1260.
- Bryant, M. and Ratner, L. 1990. Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proc. Natl. Acad. Sci. U.S.A.* 87:523-527.
- Bukrinsky, M.I., Sharova, N., McDonald, T.L., Pushkarskaya, T., Tarpley, W.G., and Stevenson, M. 1993. Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc. Natl. Acad. Sci. U.S.A.* 90:6125-6129.
- Bushman, F.D. and Craigie, R. 1991. Activities of human immunodeficiency virus (HIV-1) integration protein in vitro: Specific cleavage and integration of HIV-1 DNA. *Proc. Natl. Acad. Sci. U.S.A.* 88:1339-1343.
- Charneau, P., Alizon, M., and Clavel, F. 1992. A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication. *J. Virol.* 66:2814-2820.
- Chen, B.K., Gandhi, R.T., and Baltimore, D. 1996. CD4 down-modulation during infection of human T cells with human immunodeficiency virus type 1 involves independent activities of vpu, env, and nef. *J. Virol.* 70:6044-6053.
- Chinnasamy, D., Chinnasamy, N., Enriquez, M.J., Otsu, M., Morgan, R.A., and Candotti, F. 2000. Lentiviral-mediated gene transfer into human lymphocytes: Role of HIV-1 accessory proteins. *Blood* 96:1309-1316.
- Choe, H., Farzan, M., Sun, Y., Sullivan, N., Rollins, B., Ponath, P.D., Wu, L., Mackay, C.R., LaRosa, G., Newman, W., Gerard, N., Gerard, C., and Sodroski, J. 1996. The β -chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* 85:1135-1148.
- Clever, J., Sasseti, C., and Parslow, T.G. 1995. RNA secondary structure and binding sites for gag gene products in the 5' packaging signal of human immunodeficiency virus type 1. *J. Virol.* 69:2101-2109.
- Cohen, E.A., Dehni, G., Sodroski, J.G., and Haseltine, W.A. 1990. Human immunodeficiency virus vpr product is a virion-associated regulatory protein. *J. Virol.* 64:3097-3099.
- Corbeau, P., Kraus, G., and Wong-Staal, F. 1996. Efficient gene transfer by a human immunodeficiency virus type 1 (HIV-1)-derived vector utilizing a stable HIV-1 packaging cell line. *Proc. Natl. Acad. Sci. U.S.A.* 93:14070-14075.
- Cullen, B.R. 1991. Regulation of HIV-1 gene expression. *FASEB J.* 5:2361-2368.
- Curran, M.A., Kaiser, S.M., Achacoso, P.L., and Nolan, G.P. 2000. Efficient transduction of non-dividing cells by optimized feline immunodeficiency virus vectors. *Mol. Ther.* 1:31-38.
- Dagleish, A.G., Beverley, P.C., Clapham, P.R., Crawford, D.H., Greaves, M.F., and Weiss, R.A. 1984. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* 312:763-767.
- Daly, T.J., Cook, K.S., Gray, G.S., Maione, T.E., and Rusche, J.R. 1989. Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element in vitro. *Nature* 342:816-819.
- Das, A.T., Koken, S.E., Essink, B.B., van Wamel, J.L., and Berkhout, B. 1994. Human immunodeficiency virus uses tRNA(Lys,3) as primer for reverse transcription in HeLa-CD4⁺ cells. *FEBS Lett.* 341:49-53.

- Debouck, C. 1991. Substrate specificity of the human (type 1) and simian immunodeficiency virus proteases. *Adv. Exp. Med. Biol.* 306:407-415.
- Deglon, N., Tseng, J.L., Bensadoun, J.C., Zurn, A.D., Arsenijevic, Y., Pereira, D.A., Zufferey, R., Trono, D., and Aebischer, P. 2000. Self-inactivating lentiviral vectors with enhanced transgene expression as potential gene transfer system in Parkinson's disease. *Hum. Gene Ther.* 11:179-190.
- DuBridge, R.B., Tang, P., Hsia, H.C., Leong, P.M., Miller, J.H., and Calos, M.P. 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol. Cell Biol.* 7:379-387.
- Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D., and Naldini, L. 1998. A third-generation lentivirus vector with a conditional packaging system. *J. Virol.* 72:8463-8471.
- Ernst, R.K., Bray, M., Rekosh, D., and Hammar-skjold, M.-L. 1997. A structured retroviral RNA element that mediates nucleocytoplasmic export of intron-containing RNA. *Mol. Cell Biol.* 17:135-144.
- Farnet, C.M. and Haseltine, W.A. 1991. Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. *J. Virol.* 65:1910-1915.
- Feng, S. and Holland, E.C. 1988. HIV-1 tat trans-activation requires the loop sequence within tar. *Nature* 334:165-167.
- Feng, Y., Broder, C.C., Kennedy, P.E., and Berger, E.A. 1996. HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 272:872-877.
- Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W., and Luhrmann, R. 1995. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* 82:475-483.
- Follenzi, A., Ailles, L.E., Bakovic, S., Geuna, M., and Naldini, L. 2000. Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nat. Genet.* 25:217-222.
- Frankel, A.D. and Young, J.A. 1998. HIV-1: Fifteen proteins and an RNA. *Annu. Rev. Biochem.* 67:1-25.
- Freed, E.O., Myers, D.J., and Risser, R. 1989. Mutational analysis of the cleavage sequence of the human immunodeficiency virus type 1 envelope glycoprotein precursor gp160. *J. Virol.* 63:4670-4675.
- Gabuzda, D.H., Lawrence, K., Langhoff, E., Terwilliger, E., Dorfman, T., Haseltine, W.A., and Sodroski, J. 1992. Role of vif in replication of human immunodeficiency virus type 1 in CD4⁺ T lymphocytes. *J. Virol.* 66:6489-6495.
- Gallay, P., Hope, T., Chin, D., and Trono, D. 1997. HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. *Proc. Natl. Acad. Sci. U.S.A.* 94:9825-9830.
- Gasmi, M., Glynn, J., Jin, M.-J., Jolly, D.J., Yee, J.-K., and Chen, S.-T. 1999. Requirements for efficient production and transduction of human immunodeficiency virus type 1-based vectors. *J. Virol.* 73:1828-1834.
- Goff, S.P. 1990. Retroviral reverse transcriptase: Synthesis, structure, and function. *J. Acquir. Immune Defic. Syndr.* 3:817-831.
- Goncalves, J., Jallepalli, P., and Gabuzda, D.H. 1994. Subcellular localization of the Vif protein of human immunodeficiency virus type 1. *J. Virol.* 68:704-712.
- Gossen, M. and Bujard, H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U.S.A.* 89:5547-5551.
- Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W., and Bujard, H. 1995. Transcriptional activation by tetracyclines in mammalian cells. *Science* 268:1766-1769.
- Hamaguchi, I., Woods, N.B., Panagopoulos, I., Andersson, E., Mikkola, H., Fahlman, C., Zufferey, R., Carlsson, L., Trono, D., and Karlsson, S. 2000. Lentivirus vector gene expression during ES cell-derived hematopoietic development in vitro. *J. Virol.* 74:10778-10784.
- Hawley, R.G. 1996. Therapeutic potential of retroviral vectors. *Transfus. Sci.* 17:7-14.
- Hawley R.G. 2001. Progress toward vector design for hematopoietic stem cell gene therapy. *Curr. Gene Ther.* 1:1-17.
- Hawley, R.G., Lieu, F.H.L., Fong, A.Z.C., and Hawley, T.S. 1994. Versatile retroviral vectors for potential use in gene therapy. *Gene Ther.* 1:136-138.
- Heinzinger, N.K., Bukinsky, M.I., Haggerty, S.A., Ragland, A.M., Kewalramani, V., Lee, M.A., Gendelman, H.E., Ratner, L., Stevenson, M., and Emerman, M. 1994. The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells. *Proc. Natl. Acad. Sci. U.S.A.* 91:7311-7315.
- Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J., and Varmus, H.E. 1988. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* 331:280-283.
- Kafri, T., Blomer, U., Peterson, D.A., Gage, F.H., and Verma, I.M. 1997. Sustained expression of genes delivered directly into liver and muscle by lentiviral vectors. *Nat. Genet.* 17:314-317.
- Kafri, T., van Praag, H., Ouyang, L., Gage, F.H., and Verma, I.M. 1999. A packaging cell line for lentivirus vectors. *J. Virol.* 73:576-584.
- Kanazawa, S., Okamoto, T., and Peterlin, B.M. 2000. Tat competes with CIITA for the binding to P-TEFb and blocks the expression of MHC class II genes in HIV-1 infection. *Immunity* 12:61-70.
- Kaplan, A.H. and Swanstrom, R. 1991. Human immunodeficiency virus type 1 Gag proteins are processed in two cellular compartments. *Proc. Natl. Acad. Sci. U.S.A.* 88:4528-4532.

- Kaul, M., Yu, H., Ron, Y., and Dougherty, J.P. 1998. Regulated lentiviral packaging cell line devoid of most viral *cis*-acting sequences. *Virology* 249:167-174.
- Kotsopoulou, E., Kim, V.N., Kingsman, A.J., Kingsman, S.M., Mitrophanous, K.A. 2000. A Rev-independent human immunodeficiency virus type 1 (HIV-1)-based vector that exploits a codon-optimized HIV-1 gag-pol gene. *J. Virol.* 74:4839-4852.
- Keen, N.J., Gait, M.J., and Karn, J. 1996. Human immunodeficiency virus type-1 Tat is an integral component of the activated transcription-elongation complex. *Proc. Natl. Acad. Sci. U.S.A.* 93:2505-2510.
- Khan, M.A., Aberham, C., Kao, S., Akari, H., Gorelick, R., Bour, S., Strebel, K. 2001. Human immunodeficiency virus type 1 Vif protein is packaged into the nucleoprotein complex through an interaction with viral genomic RNA. *J. Virol.* 75:7252-7265.
- Kim, V.N., Mitrophanous, K., Kingsman, S.M., and Kingsman, A.J. 1998. Minimal requirement for a lentivirus vector based on human immunodeficiency virus type 1. *J. Virol.* 72:811-816.
- Kim, S.S., Kothari, N., You, X.J., Robinson, W.E. Jr., Schnell, T., Uberla, K., and Fan, H. 2001. Generation of replication-defective helper-free vectors based on simian immunodeficiency virus. *Virology* 282:154-167.
- Kishi, M., Nishino, Y., Sumiya, M., Ohki, K., Kimura, T., Goto, T., Nakai, M., Kakinuma, M., and Ikuta, K. 1992. Cells surviving infection by human immunodeficiency virus type 1: vif or vpu mutants produce non-infectious or markedly less cytopathic viruses. *J. Gen. Virol.* 73:77-87.
- Klages, N., Zufferey, R., and Trono, D. 2000. A stable system for the high-titer production of multiply attenuated lentiviral vectors. *Mol. Ther.* 2:170-176.
- Klimkait, T., Strebel, K., Hoggan, M.D., Martin, M.A., and Orenstein, J.M. 1990. The human immunodeficiency virus type 1-specific protein vpu is required for efficient virus maturation and release. *J. Virol.* 64:621-629.
- Kordower, J.H., Bloch, J., Ma, S.Y., Chu, Y., Palfi, S., Roitberg, B.Z., Emborg, M., Hantraye, P., Deglon, N., and Aebischer, P. 1999. Lentiviral gene transfer to the nonhuman primate brain. *Exp. Neurol.* 160:1-16.
- Korin, Y.D. and Zack, J.A. 1998. Progression to the G_{1b} phase of the cell cycle is required for completion of human immunodeficiency virus type 1 reverse transcription. *J. Virol.* 72:3161-3168.
- Kumar, M., Keller, B., Makalou, N., and Sutton, R.E. 2001. Systematic determination of the packaging limit of lentiviral vectors. *Hum. Gene Ther.* 12:1893-1905.
- LaFemina, R.L., Callahan, P.L., and Cordingley, M.G. 1991. Substrate specificity of recombinant human immunodeficiency virus integrase protein. *J. Virol.* 65:5624-5630.
- Landau, N.R., Page, K.A., and Littman, D.R. 1991. Pseudotyping with human T-cell leukemia virus type 1 broadens the human immunodeficiency virus host range. *J. Virol.* 65:162-169.
- Lever, A., Gottlinger, H., Haseltine, W., and Sodroski, J. 1989. Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. *J. Virol.* 63:4085-4087.
- Lewis, P., Hensel, M., and Emerman, M. 1992. Human immunodeficiency virus infection of cells arrested in the cell cycle. *EMBO J.* 11:3053-3058.
- Lewis, P.F. and Emerman, M. 1994. Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *J. Virol.* 68:510-516.
- Liu, H., Wu, X.N.M., Shaw, G.M., Hahn, B.H., and Kappes, J.C. 1995. The Vif protein of human and simian immunodeficiency viruses is packaged into virions and associates with viral core structures. *J. Virol.* 69:7630-7638.
- Loeb, J., Harris, M., and Hope, T. 2000. The woodchuck hepatitis virus posttranscriptional regulatory element increases transgene expression by enhancing the 3'-end metabolism of mRNAs. *Mol. Ther.* 1:S142.
- Lois, C., Hong, E.J., Pease, S., Brown, E.J., Baltimore, D. 2002. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* 295:868-872.
- Maddon, P.J., Dalgleish, A.G., McDougal, J.S., Clapham, P.R., Weiss, R.A., and Axel, R. 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* 47:333-348.
- Mak, J., Jiang, M., Wainberg, M.A., Hammarskjold, M.L., Rekosh, D., and Kleiman, L. 1994. Role of Pr160gag-pol in mediating the selective incorporation of tRNA(Lys) into human immunodeficiency virus type 1 particles. *J. Virol.* 68:2065-2072.
- Malim, M.H., Hauber, J., Le, S.Y., Maizel, J.V., and Cullen, B.R. 1989. The HIV-1 rev *trans*-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* 338:254-257.
- Mautino, M.R., Keiser, N., and Morgan, R.A. 2000. Improved titers of HIV-1-based lentiviral vectors using the SRV-1 constitutive transport element. *Gene Ther.* 7:1421-1424.
- May, C., Rivella, S., Callegari, J., Heller, G., Gensler, K.M., Luzzatto, L., and Sadelain, M. 2000. Therapeutic haemoglobin synthesis in β -thalassaemic mice expressing lentivirus-encoded human beta-globin. *Nature* 406:82-86.
- Mervis, R.J., Ahmad, N., Lillehoj, E.P., Raum, M.G., Salazar, F.H., Chan, H.W., and Venkatesan, S. 1988. The gag gene products of human immunodeficiency virus type 1: Alignment within the gag open reading frame, identification of posttranslational modifications, and evidence

- for alternative gag precursors. *J. Virol.* 62:3993-4002.
- Miller, D.G., Adam, M.A., and Miller, A.D. 1990. Gene transfer by retrovirus vectors occurs only in cells that are actively replicating at the time of infection. *Mol. Cell Biol.* 10:4239-4242.
- Miller, R.H. and Sarver, N. 1997. HIV-1 accessory proteins as therapeutic targets. *Nature Med.* 3:389-394.
- Mitrophanous, K., Yoon, S., Rohll, J., Patil, D., Wilkes, F., Kim, V., Kingsman, S., Kingsman, A., and Mazarakis, N. 1999. Stable gene transfer to the nervous system using a non-primate lentiviral vector. *Gene Ther.* 11:1808-1818.
- Miyoshi, H., Takahashi, M., Gage, F.H., and Verma, I.M. 1997. Stable and efficient gene transfer into the retina using an HIV-1-based lentiviral vector. *Proc. Natl. Acad. Sci. U.S.A.* 94:10319-10323.
- Miyoshi, H., Blomer, U., Takahashi, M., Gage, F.H., and Verma, I.M. 1998. Development of a self-inactivating lentivirus vector. *J. Virol.* 72:8150-8157.
- Naldini, L., Blomer, U., Gally, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M., and Trono, D. 1996a. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272:263-267.
- Naldini, L., Blomer, U., Gage, F.H., Trono, D., and Verma, I.M. 1996b. Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc. Natl. Acad. Sci. U.S.A.* 93:11382-11388.
- No, D., Yao, T.-P., and Evans, R.M. 1996. Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* 93:3346-3351.
- Pacchia, A.L., Adelson, M.E., Kaul, M., Ron, Y., and Dougherty, J.P. 2001. An inducible packaging cell system for safe, efficient lentiviral vector production in the absence of HIV-1 accessory proteins. *Virology* 282:77-86.
- Page, K.A., Landau, N.R., and Littman, D.R. 1990. Construction and use of a human immunodeficiency virus vector for analysis of virus infectivity. *J. Virol.* 64:5270-5276.
- Panganiban, A.T. and Fiore, D. 1988. Ordered interstrand and intrastrand DNA transfer during reverse transcription. *Science* 241:1064-1069.
- Pawliuk, R., Westerman, K.A., Fabry, M.E., Payen, E., Tighe, R., Bouhassira, E.E., Acharya, S.A., Ellis, J., London, I.M., Eaves, C.J., Humphries, R.K., Beuzard, Y., Nagel, R.L., and Leboulch, P. 2001. Correction of sickle cell disease in transgenic mouse models by gene therapy. *Science* 294:2368-2371.
- Peng, C., Ho, B.K., Chang, T.W., and Chang, N.T. 1989. Role of human immunodeficiency virus type 1-specific protease in core protein maturation and viral infectivity. *J. Virol.* 63:2550-2556.
- Pfeifer, A., Ikawa, M., Dayn, Y., and Verma, I.M. 2002. Transgenesis by lentiviral vectors: Lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos. *Proc. Natl. Acad. Sci. U.S.A.* 99:2140-2145.
- Planelles, V., Bachelier, F., Jowett, J.B., Haislip, A., Xie, Y., Banooni, P., Masuda, T., and Chen, I.S. 1995. Fate of the human immunodeficiency virus type 1 provirus in infected cells: A role for vpr. *J. Virol.* 69:5883-5889.
- Poeschela, E., Gilbert, J., Li, X., Huang, S., Ho, A., and Wong-Staal, F. 1998. Identification of a human immunodeficiency virus type 2 (HIV-1-2) encapsidation determinant and transduction of nondividing human cells by HIV-1-2-based lentivirus vectors. *J. Virol.* 72:6527-6536.
- Poznansky, M., Lever, A., Begeron, L., Haseltine, W., and Sodroski, J. 1991. Gene transfer into human lymphocytes by a defective human immunodeficiency virus type I vector. *J. Virol.* 65:532-536.
- Ramezani, A., Hawley, T.S., and Hawley, R.G. 2000. Lentiviral vectors for enhanced gene expression in human hematopoietic cells. *Mol. Ther.* 2:458-469.
- Rice, W.G., Supko, J.G., Malspeis, L., Buckheit, R.W., Jr., Clanton, D., Bu, M., Graham, L., Schaeffer, C.A., Turpin, J.A., and Domagala, J. 1995. Inhibitors of HIV-1 nucleocapsid protein zinc fingers as candidates for the treatment of AIDS. *Science* 270:1194-1197.
- Robert-Guroff, M., Popovic, M., Gartner, S., Markham, P., Gallo, R.C., and Reitz, M.S. 1990. Structure and expression of tat-, rev-, and nef-specific transcripts of human immunodeficiency virus type 1 in infected lymphocytes and macrophages. *J. Virol.* 64:3391-3398.
- Robey, W.G., Safai, B., Oroszlan, S., Arthur, L.O., Gonda, M.A., Gallo, R.C., and Fischinger, P.J. 1985. Characterization of envelope and core structural gene products of HTLV-III with sera from AIDS patients. *Science* 228:593-595.
- Schlegel, R., Tralka, T.S., Willingham, M.C., and Pastan, I. 1983. Inhibition of VSV binding and infectivity by phosphatidylserine: Is phosphatidylserine a VSV-binding site? *Cell* 32:639-646.
- Schwartz, S., Felber, B.K., and Pavlakis, G.N. 1992. Mechanism of translation of monocistronic and multicistronic human immunodeficiency virus type 1 mRNAs. *Mol. Cell Biol.* 12:207-219.
- Selby, M.J., Bain, E.S., Luciw, P.A., and Peterlin, B.M. 1989. Structure, sequence, and position of the stem-loop in tar determine transcriptional elongation by tat through the HIV-1 long terminal repeat. *Genes Dev.* 3:547-558.
- Sheng, N. and Erickson-Vitonen, S. 1994. Cleavage of p15 protein in vitro by human immunodeficiency virus type 1 protease is RNA dependent. *J. Virol.* 68:6207-6214.
- Skripkin, E., Paillart, J.C., Marquet, R., Ehresmann, B., and Ehresmann, C. 1994. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 91:4945-4949.
- Srinivasakumar, N., Chazal, N., Helga-Maria, C., Prasad, S., Hammarskjold, M.L., and Rekosh, D.

1997. The effect of viral regulatory protein expression on gene delivery by human immunodeficiency virus type 1 vectors produced in stable packaging cell lines. *J. Virol.* 71:5841-5848.
- Stein, B.S. and Engleman, E.G. 1990. Intracellular processing of the gp160 HIV-1 envelope precursor: Endoproteolytic cleavage occurs in a *cis* or medial compartment of the Golgi complex. *J. Biol. Chem.* 265:2640-2649.
- Stein, B.S., Gowda, S.D., Lifson, J.D., Penhallow, R.C., Bensch, K.G., and Engleman, E.G. 1987. pH-independent HIV-1 entry into CD4-positive T cells via virus envelope fusion to the plasma membrane. *Cell* 49:659-668.
- Sun, Y., Pinchuk, L.M., Agy, M.B., and Clark, E.A. 1997. Nuclear import of HIV-1 DNA in resting CD4⁺ T cells requires a cyclosporin A-sensitive pathway. *J. Immunol.* 158:512-517.
- Swann, S.A., Williams, M., Story, C.M., Bobbitt, K.R., Fleis, R., and Collins, K.L. 2001. HIV-1 Nef blocks transport of MHC class I molecules to the cell surface via a PI 3-kinase-dependent pathway. *Virology* 282:267-277.
- Trono, D. 1995. HIV-1 accessory proteins: Leading roles for the supporting cast. *Cell* 82:189-192.
- Uchida, N., Sutton, R.E., Frieria, A.M., He, D., Reitsma, M.J., Chang, W.C., Veres, G., Scollay, R., and Weissman, I.L. 1998. HIV-1, but not murine leukemia virus, vectors mediated high efficiency gene transfer into freshly isolated G₀/G₁ human hematopoietic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 95:11939-11944.
- Ullman, K.S., Powers, M.A., and Forbes, D.J. 1997. Nuclear export receptors: From importin to exportin. *Cell* 90:967-970.
- Valsamakis, A., Zeichner, S., Carswell, S., and Alwine, J.C. 1991. The human immunodeficiency virus type 1 polyadenylation signal: A 3' long terminal repeat element upstream of the AAUAAA necessary for efficient polyadenylation. *Proc. Natl. Acad. Sci. U.S.A.* 88:2108-2112.
- Vigna, E. and Naldini, L. 2000. Lentiviral vectors: Excellent tools for experimental gene transfer and promising candidates for gene therapy. *J. Gene Med.* 2:308-316.
- Wen, W., Meinkoth, J.L., Tsien, R.Y., and Taylor, S.S. 1995. Identification of a signal for rapid export of proteins from the nucleus. *Cell* 82:463-473.
- Wolfgang, M.J., Eisele, S.G., Browne, M.A., Schotzko, M.L., Garthwaite, M.A., Durning, M., Ramezani, A., Hawley, R.G., Thomson, J.A., and Golos, T.G. 2001. Rhesus monkey placental transgene expression after lentiviral gene transfer into preimplantation embryos. *Proc. Natl. Acad. Sci. U.S.A.* 98:10728-10732.
- Wiley, R.L., Bonifacino, J.S., Potts, B.J., Martin, M.A., and Klausener, R.D. 1988. Biosynthesis, cleavage, and degradation of the human immunodeficiency virus 1 envelope glycoprotein gp160. *Proc. Natl. Acad. Sci. U.S.A.* 85:9580-9584.
- Yang, Y., Vanin, E.F., Whitt, M.A., Fornerod, M., Zwart, R., Schneiderman, R.D., Grosveld, G., and Nienhuis, A.W. 1995. Inducible, high-level production of infectious murine leukemia retroviral vector particles pseudotyped with vesicular stomatitis virus G envelope protein. *Hum. Gene Ther.* 6:1203-1213.
- Yu, H., Rabson, A.B., Kaul, M., Ron, Y., and Dougherty, J.P. 1996. Inducible human immunodeficiency virus type 1 packaging cell lines. *J. Virol.* 70:4530-4537.
- Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L., and Charneau, P. 2000. HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell* 101:173-185.
- Zufferey, R., Nagy, D., Mandel, R. J., Naldini, L., and Trono, D. 1997. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nature Biotechnol.* 15:871-875.
- Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L., and Trono, D. 1998. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J. Virol.* 72:9873-9880.
- Zufferey, R., Donello, J.E., Trono, D., and Hope, T.J. 1999. Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *J. Virol.* 73:2886-2892.

Contributed by Ali Ramezani
American Red Cross
Rockville, Maryland

Robert G. Hawley
American Red Cross
Rockville, Maryland
and The George Washington University
Washington, D.C.

Generation of HIV-1-Based Lentiviral Vector Particles

Generation of recombinant HIV-1 vector particles (see *UNIT 16.21* for overview) involves subcloning the gene of interest into an HIV-1 transfer vector backbone, cotransfection of the plasmid into a recipient cell line together with packaging and envelope plasmids (or directly into a stable packaging line), and collection of vector supernatants.

The transfer vector can be selected from currently available vector backbones (Ramezani et al., 2000; Vigna and Naldini, 2000) or custom-tailored for a particular application. Modification of an existing vector—e.g., changing the promoter or the marker gene and/or insertion of a new transgene—involves standard recombinant DNA techniques as described in other units of this book and elsewhere (Sambrook and Russell, 2001).

Once the transfer vector containing the transgene has been constructed and confirmed by restriction enzyme analysis and DNA sequencing, it can be used for the production of recombinant HIV-1 vector particles. Among the many techniques used for transient transfection of adherent mammalian cells, the CaPO₄ method of Graham and van der Eb (1973) remains one of the most cost-effective and reproducible. Because transient transfection of 293T cells is the primary method used to produce HIV-1 vector particles, the CaPO₄ method is used with this cell line in the procedure described below (see Basic Protocol 1) for the production of the vector particles. A protocol for production of vector particles from a stable producer cell line can be found in Klages et al. (2000). Basic Protocol 1 outlines small-scale vector preparation in 100-mm dishes and will yield ~5 ml of vector supernatant with titers of 10⁶ to 10⁷ TU/ml (depending on the vector). When larger volumes of vector supernatant are required, this method can be easily scaled up by using multiple 150-mm dishes. When 150-mm dishes are used, the quantities of all of the components must be tripled to account for the 3-fold increase in surface area compared to that of 100-mm dishes. By concentrating large volumes of vector supernatants obtained in this manner, titers as high as 10⁹ TU/ml can be achieved. A modification of this procedure using sodium butyrate to stimulate production of vector particles is also presented (see Basic Protocol 2).

In addition to the gene-transfer protocols routinely employed for production of recombinant vector supernatants (see Basic Protocols 1 and 2), this unit describes related techniques. When required, vector supernatants can be concentrated by centrifugation procedures (see Basic Protocol 3 and Alternate Protocol 1). Also provided are protocols for further purification and titration of vector particles (see Basic Protocols 4 and 5, and Alternate Protocol 2). Methods for transduction of both adherent and nonadherent cells are included (see Basic Protocol 7). Although the probability of de novo generation of a new infectious recombinant lentivirus (referred to as a replication-competent retrovirus or RCR; see *UNIT 16.21*) is extremely low due to the extensive modifications that have been made to the HIV-1 genome, it is important to stress that caution should still be exercised during the course of recombinant lentiviral vector particle production and application. It is therefore recommended that the RCR detection methods included in this unit (see Basic Protocol 6, and Alternate Protocols 3 and 4) be used regularly in order to definitively rule out the presence of RCR variants in any vector stocks.

CAUTION: When working with human blood, cells, or infectious agents, appropriate biosafety practices must be followed.

NOTE: All culture incubations should be performed in a 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper aseptic technique should be used accordingly.

**BASIC
PROTOCOL 1**

COTRANSFECTION OF 293T CELLS AND PRODUCTION OF LENTIVIRAL VECTOR SUPERNATANTS

VSV-G-pseudotyped HIV-1-based lentiviral vector particles are prepared by transiently transfecting the transfer vector plasmid, the packaging plasmid or plasmids (e.g., pCMVΔR8.91), and the VSV-G protein envelope plasmid (e.g., pMD.G) into subconfluent 293T cells by a modification of the calcium phosphate precipitation method (Hawley et al., 1989). Note that procedures for producing lentiviral supernatants are essentially the same as those used for producing oncoretroviral supernatants; the user is thus also directed to *UNIT 9.11* for additional details on culturing and transfecting 293T cells.

Materials

293T human embryonic kidney cell line (ATCC #CRL-11268)
293T cell growth medium (see recipe) with and without 10 mM HEPES [add HEPES from 100× (1 mM) stock (see recipe)]
Lentiviral vector plasmid DNA containing the gene of interest (see *UNIT 16.21*; Ramezani et al., 2000; Vigna and Naldini, 2000)
Lentiviral packaging construct DNA (e.g., pCMVΔR8.91; Zufferey et al., 1997)
VSV-G expressing plasmid DNA (e.g., pMD.G; Naldini et al., 1996)
2.5 M CaCl₂ (see recipe)
2× HEPES-buffered saline (HeBS; see recipe)
Phosphate-buffered saline (PBS; *APPENDIX 2*)
100-mm tissue culture dishes
Tabletop centrifuge
0.45-μm pore-size filter

Additional reagents and equipment for cell culture and counting cells (*APPENDIX 3F*)

1. Grow 293T cells in 293T cell growth medium.

*293T cells are split every 3 to 4 days between 1:4 to 1:8. A trypsin-EDTA solution (e.g., Life Technologies) is used to dissociate the cells. To trypsinize, remove medium and rinse the monolayer with PBS (without Ca²⁺ and Mg²⁺; see *APPENDIX 2*). Remove the PBS and add enough trypsin-EDTA solution to cover the monolayer. Incubate the cultures at room temperature until the cells round up and detach. Add an equal volume of fresh 293T cell growth medium containing 10% FBS to inactivate the trypsin. Collect the cells by low-speed centrifugation (3 min at 500 × g), resuspend in fresh culture medium, and plate. Do not allow the cells to grow above 80% confluence.*

*See *APPENDIX 3F* for general cell culture procedures.*

2. At a point 24 hr before transfection, count cells (*APPENDIX 3F*), and seed 293T cells into 100-mm tissue culture dishes, each containing 7 ml of 293T cell growth medium, at 4 × 10⁶ cells per dish.
3. For each plate to be used, prepare cotransfection precipitate as follows:
 - a. Mix 15 μg of the vector plasmid containing the gene of interest, 10 μg of the packaging plasmid pCMVΔR8.91, and 5 μg of the VSV-G protein envelope plasmid pMD.G.
 - b. Bring the volume up to 400 μl with sterile water.

- c. Add 100 μ l of 2.5 M CaCl_2 and mix.
- d. Add the DNA/ CaCl_2 solution dropwise to 500 μ l of 2 \times HeBS in a 15-ml conical tube. Use a battery-powered mechanical pipettor and a 2-ml pipet to bubble the 2 \times HeBS as the DNA/ CaCl_2 solution is added.

See Figure 9.1.1 for an illustration of the technique used.

- e. Vortex immediately for 5 sec and incubate at room temperature for 20 min.
4. Add the precipitate dropwise to the plates from step 2 so as to distribute it evenly over the cells and mix gently. Incubate the cells at 37°C overnight (16 hr).
5. The next day, remove medium from each dish, rinse cells with 5 ml PBS, and add 5 ml fresh 293T cell growth medium containing 10 mM HEPES.
6. Incubate an additional 48 hr, then collect the vector-containing supernatant medium. Centrifuge 10 min at 2000 \times g, 4°C, to remove cellular debris, then filter through a 0.45- μ m pore-size filter. Divide into 1- to 5-ml aliquots and freeze at -80°C, keeping a small amount for determination of vector titer (see Basic Protocol 5).

STIMULATION OF VECTOR PARTICLE PRODUCTION USING SODIUM BUTYRATE

**BASIC
PROTOCOL 2**

Sodium butyrate ($\text{CH}_3\text{CH}_2\text{CH}_2\text{COONa}$) is a short-chain fatty acid that has been demonstrated to inhibit histone deacetylase (McKnight et al., 1980). Hyperacetylation of nucleosomal histones results in a more open and accessible chromatin configuration within the promoter regions of genes, leading to higher transcriptional activity (Norton et al., 1989). Notably, expression of exogenous sequences introduced into mammalian cells is also enhanced by sodium butyrate treatment (Gorman and Howard, 1983; Laughlin et al., 1995). Inclusion of sodium butyrate at concentrations of 4 to 10 mM during transient production of lentiviral (and oncoretroviral) vector particles has been reported to result in 10- to 15-fold increases in titers (Soneoka et al., 1995; Gasmi et al., 1999).

Sodium butyrate treatment may be most effective for vectors that have low titers. Since sodium butyrate is toxic to the cells, it must be removed after the initial treatment and replaced with fresh medium if the supernatants are not going to be concentrated.

Materials

Cells containing calcium phosphate–DNA coprecipitate (see Basic Protocol 1, step 5)

500 mM sodium butyrate stock solution (see recipe)

293T cell growth medium (see recipe)

Phosphate-buffered saline (PBS; *APPENDIX 2*)

Additional reagents and equipment for cotransfection of 293T cells and preparation of lentiviral supernatants (see Basic Protocol 1)

1. After the overnight incubation of the cotransfected cells (see Basic Protocol 1, step 5):
 - a. Add sodium butyrate (from 500 mM stock) to 293T cell growth medium to a final concentration of 10 mM.
 - b. Remove the medium with the calcium phosphate–DNA coprecipitate from each dish and replace with an equal volume of the sodium butyrate–containing growth medium
 - c. Incubate the cells for 12 to 14 hr.

**Protein
Expression**

16.22.3

- d. Remove medium and wash with PBS to remove the sodium butyrate. Add 5 ml of fresh 293T cell growth medium.

If vector particles are to be concentrated (see Basic Protocol 3 and Alternate Protocol 1), this medium change is not necessary, as the sodium butyrate will be removed from the vector particle preparations during the subsequent concentration step.

2. Incubate an additional 24 hr, then collect the vector-containing supernatants.

See Basic Protocol 1, step 6, for treatment and storage of the vector-containing medium.

CONCENTRATION OF HIV-1 VECTOR PARTICLES BY ULTRACENTRIFUGATION

Lentiviral vector transduction efficiency depends to a large extent on the ratio of vector particles to target cells (multiplicity of infection, MOI). For example, when using VSV-G-pseudotyped lentiviral vectors, MOI values of 300 to 1000 have been used to achieve efficient gene delivery to hematopoietic progenitor cells (Miyoshi et al., 1999; Case et al., 1999). Achieving these MOI values requires highly concentrated vector particle preparations.

Several methods have been developed to concentrate lentiviral vector particles. The choice of concentration protocol depends on the envelope selected to pseudotype the particle and the quantities of particles to be produced. The stability of the VSV-G envelope protein allows generation of high-titer lentiviral vector stocks by ultracentrifugation, as described here. On the other hand, if the HIV-1 Env, the amphotropic envelope, or similar envelope glycoproteins are preferred, low-speed centrifugation methods should be used (see Alternate Protocol 1), since the TM and SU subunits of these envelope glycoproteins are not covalently linked and they tend to dissociate upon high-speed centrifugation.

Although convenient and efficient, the volume of the vector supernatant that can be concentrated by this ultracentrifugation is limited by the capacity of the ultra-high-speed rotors (six bottles, each containing 70 ml).

Materials

Vector-containing supernatants (see Basic Protocols 1 and 2)
Appropriate medium to resuspend concentrated vector particles
Ultracentrifuge and rotor
Polycarbonate 70-ml ultracentrifuge bottles and caps
Tabletop centrifuge

1. Ultracentrifuge vector supernatants 90 min at $50,000 \times g$, 4°C .

Small (~2- to 5-mm-diameter) pale-yellow pellets should be visible after centrifugation.

2. Remove supernatants and resuspend pellets in ~100 μl of medium appropriate for the downstream application by gentle pipetting. Vortex gently overnight at 4°C to allow complete resuspension.
3. Centrifuge 5 min at $10,000 \times g$, 4°C , in a tabletop centrifuge to bring down insoluble debris. Gently remove the supernatant containing the concentrated vector particles and freeze in 0.5-ml aliquots at -80°C , keeping a small amount for determination of vector titer (see Basic Protocol 5).

Vector particles are usually frozen as 0.5-ml aliquots. However, this volume could vary between 0.1 ml to 1.0 ml depending on the downstream application (i.e., the number of cells to be transduced). The volumes should be chosen so that repeated freezing and thawing of the vector particles is avoided, since this process leads to a reduction in vector titers.

If desired, the vector particles can be further concentrated by another round of ultracentrifugation. See Basic Protocol 4 for purifying the concentrated lentiviral vector particles by anion-exchange chromatography.

CONCENTRATION OF HIV-1 VECTOR PARTICLES BY POLY-L-LYSINE PRECIPITATION

**ALTERNATE
PROTOCOL 1**

Certain cationic molecules such as poly-L-lysine (PLL) have been used to concentrate lentiviral vectors (Zhang et al., 2001). The positive electrostatic charge of PLL binds to the vector particles and forms complexes that can be precipitated by low-speed centrifugation. This method is particularly suitable for concentrating large volumes of lentiviral vector supernatants—taking advantage of the capacity of the rotors (accommodating six bottles, each containing 250 ml) available for low-speed centrifugation—which is higher than the capacity of the ultracentrifuge rotors used in Basic Protocol 3. Furthermore, the relatively low centrifugation speed also prevents the precipitation of dyes such as the phenol red in the medium (Miller et al., 1996), marker gene products (Liu et al., 1996), or other small molecules that are often found to be toxic or cause the differentiation of primary cells.

Additional Materials (also see *Basic Protocol 3*)

100 mg/ml poly-L-lysine hydrobromide (PLL; mol. wt. based on viscosity, 27,400) in PBS (see *APPENDIX 2* for PBS), freshly prepared
Appropriate medium to resuspend concentrated vector particles
250-ml centrifuge bottles

1. Add the 100 mg/ml PLL solution to the vector supernatants (from Basic Protocol 1 or 2) to a final concentration of 0.005% (w/v), then incubate 30 min at 4°C.
2. Precipitate the vector particles by centrifuging 2 hr at $10,000 \times g$, 4°C.
3. After the centrifugation, gently remove the supernatant and resuspend each pellet in 100 μ l of medium appropriate for the downstream application. Gently mix the concentrated vector particles by vortexing overnight at 4°C to allow complete resuspension.
4. Centrifuge 5 min at $5000 \times g$, 4°C, to bring down the debris. Divide the supernatant containing the concentrated vector particles into aliquots and store at -80°C , keeping a small amount for determination of vector titer (see Basic Protocol 5).

PURIFICATION OF CONCENTRATED LENTIVIRAL VECTOR PARTICLES BY ANION-EXCHANGE CHROMATOGRAPHY

**BASIC
PROTOCOL 4**

The relatively low centrifugation speed used in PLL precipitation (see Alternate Protocol 1) to concentrate the lentiviral vector particles should generally preclude the coconcentration of low-molecular-weight toxic impurities. However, lentiviral vector particles concentrated by ultracentrifugation (see Basic Protocol 3) may contain contaminants that can be toxic to certain target cells and should be purified prior to use.

Contaminating toxic proteins can be removed from lentiviral vector preparations by anion-exchange chromatography either before or after the concentration step (Scherr et al., 2002). The yield of purified lentiviral vector particles obtained with this method ranges between 30% and 60%.

**Protein
Expression**

16.22.5

Materials

Vector particles concentrated by ultracentrifugation (see Basic Protocol 3), resuspended in PBS

Phosphate-buffered saline (PBS; *APPENDIX 2*) containing 1 M NaCl

80 × 6-mm anion-exchange column: Fractoflow 80-6 column (Merck) packed with 40- to 90- μ m mesh size resin according to manufacturer's instructions (also see *UNIT 10.10*)

Ultrafree-15 centrifugal filter devices (Millipore; 100,000 MWCO; also see *APPENDIX 3C*)

1. Add 1 to 10 ml of vector particles suspended in PBS to an anion-exchange column.
2. Elute the vector particles from the column by adding 30 ml PBS containing 1 M NaCl at a flow rate of 4 ml/min.
3. Desalt and concentrate the vector particles by centrifuging through a 100,000 MWCO Ultrafree-15 centrifugal filter device 15 min (or until vector supernatants are concentrated to the desired volume) at $2000 \times g$, 4°C (also see manufacturer's instructions).
4. Dilute the concentrated vector particles with the appropriate medium for transduction (Basic Protocol 7) and culture of the target cells. Divide into aliquots and store at -80°C, keeping a small amount for determination of vector titer (see Basic Protocol 5).

ALTERNATE PROTOCOL 2

PURIFICATION OF CONCENTRATED LENTIVIRAL VECTOR PARTICLES BY THE DUAL-FILTER METHOD

Dual-filter purification is a simple and efficient alternative that can be used to purify concentrated lentiviral vector particle preparations. This procedure is much less labor-intensive and time-consuming than anion-exchange chromatography (Basic Protocol 4), and has the advantages of being cost-effective and providing higher yields of vector particle recovery. The protocol is based on the size of the HIV-1 lentiviral vector particle, which is ~100 nm in diameter. The concentrated vector particles are first passed through a 0.22- μ m filter, commonly used to sterilize solutions. While the HIV-1 vector particles readily pass through this filter, large-molecular-weight cellular debris is trapped. The filtrate is then added to an Ultrafree-15 centrifugal filter with a 100-kDa molecular weight cut-off, which allows most small proteins, dyes, and empty vector particles (~50 nm in diameter) to be removed.

The dual-filter method reduces the processing time required to obtain purified vector particles to ~30 min, with a 50% and 70% recovery yield.

Materials

Vector supernatants concentrated by ultracentrifugation (see Basic Protocol 3)

Appropriate serum-free medium

10-ml syringes and 0.22- μ m syringe filters

Ultrafree-15 centrifugal filter devices (Millipore; 100,000 MWCO; also see *APPENDIX 3C*)

1. Add fresh serum-free medium to dilute the concentrated vector particles (from Basic Protocol 3, step 3) to 5 ml.
2. Using a 10-ml syringe, slowly pass the vector particle-containing medium through a 0.22- μ m filter into an Ultrafree-15 centrifugal filter device.

3. Centrifuge the Ultrafree-15 centrifugal filter device 15 min (or until vector supernatants are concentrated to desired volume) at $2000 \times g$, 4°C (also see manufacturer's instructions).
4. Collect the purified vector particles and freeze in aliquots at -80°C . Keep a small aliquot for determination of vector titers (see Basic Protocol 5).

TITRATION OF LENTIVIRAL VECTOR STOCKS

The HT1080 fibrosarcoma cell line (Rasheed et al., 1974) can be used to determine the titer of the lentiviral vector particle preparations.

Materials

HT1080 cell line (ATCC #CCL-121)
 293T cell growth medium (see recipe)
 Vector preparations (Basic Protocols 3 or 4 or Alternate Protocols 1 or 2)
 6 mg/ml (1000 \times) polybrene stock solution
 G418 (UNIT 9.5)
 0.3% (w/v) crystal violet in 70% (v/v) methanol
 6-well tissue culture plates

Additional reagents and equipment for flow cytometric analysis of EGFP expression (Rasko, 1999), Xgal staining for *lacZ* (β -galactosidase) expression (UNIT 9.10), or enumeration of *neo*-resistant cells (UNIT 9.5)

1. Grow HT1080 cells in 293T cell growth medium (same medium used in Basic Protocol 1). At a point 4 to 6 hr before the titration of the vector preparation, seed 2.5×10^5 HT1080 cells in each well of a 6-well plate.
2. At a time point 4 to 6 hr after cells have been seeded, prepare serial dilutions of an aliquot of each vector preparation (for unconcentrated vector supernatants, 10^0 , 10^{-1} , and 10^{-2} ; for concentrated supernatants, 10^{-2} , 10^{-3} , and 10^{-4}), using 293T cell growth medium, in a final volume of 1 ml. To each of the dilutions, add 1 μl of 6 mg/ml (1000 \times) polybrene for a final concentration of 6 $\mu\text{g/ml}$ polybrene. Add each serial dilution to the corresponding cell-containing well and return to the incubator for 4 hr.
Basic Protocol 7 contains additional information concerning stable transduction with recombinant lentiviral vectors.
3. After 4 hr of transduction, remove the vector-containing medium and replace with 2 ml fresh 293T cell growth medium. Continue incubating for 48 hr.
- 4a. *For vectors containing EGFP or lacZ marker:* After 48 hr, determine the relative end-point vector titers (in TU/ml) by flow cytometric analysis for EGFP expression (if the vector contains the EGFP gene; see, e.g., Rasko, 1999) or by Xgal staining (if the vector contains the *lacZ* gene; see UNIT 9.10). Use the following equation:

Vector titer = number of HT1080 cells \times % of EGFP⁺ or *lacZ*⁺ cells \times dilution factor.

- 4b. *For vectors carrying the neo drug-resistance gene:* After 48 hr, split cells 1:10 and seed into 100-mm dishes in medium containing 400 $\mu\text{g/ml}$ G418. Replace medium every 4 to 5 days. At a time point 2 weeks later, fix and stain cells with 0.3% crystal violet in 70% methanol and enumerate the G418-resistant colonies (see UNIT 9.5). Use the following equation to determine the vector titer:

Vector titer = number of G418-resistant colonies \times 10 \times dilution factor.

The factor of 10 must be used because of the 1:10 split.

**DETECTION OF REPLICATION-COMPETENT VIRUSES BY THE
REVERSE TRANSCRIPTASE ASSAY**

Absence of replication-competent HIV-1-derived virus in recombinant lentiviral vector stocks can be determined by the methods here and in Alternate Protocols 3 and 4. This assay measures the enzymatic activity of HIV-1 reverse transcriptase.

Materials

- HT1080 cell line (ATCC #CCL-121)
- Undiluted vector particle supernatant (Basic Protocol 3 or 4 or Alternate Protocol 1 or 2)
- 293T cell growth medium (see recipe)
- 35-, 100-, and 150-mm tissue culture dishes
- Tabletop centrifuge
- Ultracentrifuge and rotor
- Reverse transcriptase assay kit (e.g., Roche Diagnostics, cat. no. 1468120)
- Additional reagents and equipment for stable transduction of cells (see Basic Protocol 5, steps 1 and 2)

1. Prepare a culture supernatant from stably transduced HT1080 cells as follows:
 - a. Transduce 2.5×10^5 cells in a 35-mm tissue culture dish with 1 ml of undiluted vector particle supernatant (see Basic Protocol 5, steps 1 and 2).
 - b. Grow cells until confluent; transfer to a 100-mm tissue culture dish and then to a 150-mm tissue culture dish. Use 293T cell growth medium to split.
A confluent 150-mm dish should contain $1-2 \times 10^7$ cells.
 - c. Centrifuge 10 min at $2000 \times g$, 4°C , to remove cellular debris and collect the supernatant from these cells (~20 ml) for use in the assay.
2. Concentrate supernatants by ultracentrifuging 2 hr at $40,000 \times g$, 4°C , and resuspending the pellet in 40 μl of the lysis buffer provided with the reverse transcriptase assay kit.
3. Test the concentrated culture supernatants for HIV-1 reverse transcriptase activity using a commercial reverse transcriptase assay kit according to the manufacturer's instructions.

**DETECTION OF REPLICATION-COMPETENT VIRUSES BY THE
TAT-INDUCTION ASSAY**

This assay uses a reporter T cell line (CEM-GFP) which expresses the GFP gene driven by the HIV-1 LTR promoter to measure Tat activity. Infection of CEM-GFP cells with wild-type HIV-1 results in a 100-to 1000-fold increase in GFP fluorescence over 2 to 4 days compared with uninfected cells (Gervaix et al., 1997).

Additional Materials (also see Basic Protocol 6)

- CEM-GFP cell line (NIH AIDS Research and Reference Reagent Program)
- Growth medium for CEM-GFP cells: IMDM containing 10% FBS (see recipe for serum-containing growth media for transduction)
- 6 mg/ml polybrene stock solution
- HIV packaging plasmid DNA (e.g., pCMV Δ R8.91)
- Additional reagents and equipment for flow cytometric analysis of GFP expression (Rasko, 1999)

1. Prepare lentiviral supernatant (see Basic Protocol 6, step 1).
2. Culture and passage CM-GFP cells using IMDM containing 10% FBS according to the instructions provided by the NIH.

CM-GFP cells are suspension cells. They are cultured at 1×10^6 cells/ml and split every 3 days at a ratio of 1:3 with fresh medium.

3. Resuspend 5×10^5 CEM-GFP cells in a 1-ml aliquots of the supernatant from step 1 in a well of a 6-well plate, add polybrene to a concentration of 2 μ g/ml (from 6 mg/ml polybrene stock), and incubate overnight to achieve transduction. Transduce CEM-GFP cells with the pCMV Δ R8.91 packaging plasmid as a positive control.
4. Remove polybrene-containing medium and replace with 1 ml fresh medium. Return plates to incubator.

Maintain the cell density below 1×10^6 cells/ml by adjusting the volume of the medium as the cells proliferate.

5. Analyze GFP fluorescence 3 days post-incubation by flow cytometry (Rasko, 1999).

DETECTION OF REPLICATION-COMPETENT VIRUSES BY THE GAG TRANSFER ASSAY

**ALTERNATE
PROTOCOL 4**

The lentiviral vector stocks can be tested for the presence of replication-competent HIV-1-derived virus by serial transfer and performing an enzyme-linked immunosorbent assay (ELISA) to detect the HIV-1 p24 antigen. This assay measures the presence of HIV-1 p24 protein in the medium obtained from vector-transduced cells following several passages in culture (Naldini et al., 1996). Using this method, p24 concentrations as low as 1 pg/ml (~1 to 2 TU/ml) can be detected.

Additional Materials (also see Basic Protocol 6)

p24 ELISA assay kit (e.g., Alliance HIV-1 p24 ELISA kit from Perkin-Elmer Life Sciences; cat. no. NEK050001KT)

1. Prepare lentiviral supernatant from stably transduced cells (see Basic Protocol 6, step 1).
2. Test the culture supernatants for HIV-1 p24 by using a commercial p24 ELISA assay kit according to the manufacturer's instructions.

TRANSDUCTION OF TARGET CELLS

**BASIC
PROTOCOL 7**

Several protocols have been described for transducing target cells with lentiviral vector particles. In general, these are the same protocols that are used to transduce cells with oncoretroviral vector particles (UNIT 9.11). Although it has been reported that polycations such as polybrene or protamine sulfate, which reduce the electrostatic repulsive interactions between the negatively charged surfaces of the cell and the vector particles, do not improve transduction efficiencies with VSV-G-pseudotyped lentiviral vectors, these agents are routinely included in the transduction cocktails. Protamine sulfate is recommended for use when transducing primary cells since polybrene has been found to be somewhat more toxic to these cell types. Likewise, transduction efficiencies of nonadherent hematopoietic cells expressing the fibronectin receptors very late antigen 4 (VLA-4; CD49d) and very late antigen 5 (VLA-5; CD49e) with amphotropic envelope-containing oncoretroviral vector particles have been shown to be increased by incubating the cells on recombinant fibronectin fragments through simultaneous binding and colocalization of the vector particles and target cells (Moritz et al., 1994; Hanenberg et al.,

**Protein
Expression**

16.22.9

1997). While fibronectin does not seem to similarly facilitate transduction of these target cells by VSV-G-pseudotyped vectors, it has been demonstrated to have growth-supporting properties for hematopoietic progenitor cells (Dao et al., 1998; Yokota et al., 1998). It is thus frequently used during lentiviral gene delivery to these cells (Donahue et al., 2001).

Transduction of Adherent Cells

Although lentiviral vectors do not require cell division, optimal transduction efficiencies are obtained if the cells are actively dividing or have entered the G_{1b} phase of the cell cycle. A major limiting factor affecting the transduction efficiency is the local concentration of vector particles (Chuck et al., 1996; Haas et al., 2000). The transduction efficiency is therefore not increased under static incubation conditions by increasing the volume of the vector supernatant (Morgan et al., 1995).

Transduction of Nonadherent Cells

Nonadherent cells are transduced by incubating vector particles with the cells in presence of polybrene or protamine sulfate on CH-296 recombinant fibronectin fragment (Moritz et al., 1994; Donahue et al., 2001). In addition, transduction of nonadherent cells has been shown to be significantly enhanced by centrifugation, i.e., spinoculation (Kotani et al., 1994; Bahnson et al., 1995).

Materials

- Adherent *or* nonadherent target cells
- Appropriate serum-containing growth medium for transduction (see recipe)
- Vector preparation (Basic Protocols 3 or 4 or Alternate Protocols 1 or 2)
- 6 mg/ml (1000×) polybrene stock
- 4 mg/ml (1000×) protamine sulfate stock
- Recombinant fibronectin fragment CH-296 (RetroNectin; Takara Shuzo)
- 2% (w/v) bovine serum albumin (BSA) in phosphate-buffered saline (PBS; APPENDIX 2)
- 30-mm tissue culture dishes (for adherent cells), 35-mm suspension culture dishes (for nonadherent cells), or 6-well tissue culture plates (for transduction of nonadherent cells using recombinant fibronectin fragment)
- 14-ml round-bottom polypropylene centrifuge tubes
- Tabletop centrifuge
- Additional reagents and equipment for cell culture and counting cells (APPENDIX 3F)

To transduce adherent cells

- 1a. Culture the target cells in appropriate serum-containing growth medium (e.g., DMEM containing 10% heat-inactivated FBS).
See APPENDIX 3F for general cell culture procedures.
- 2a. Split cells 8 hr prior to transduction and seed 2.5×10^5 cells in a 35-mm tissue culture dish.
- 3a. Remove supernatant and add vector particles in the presence of 6 µg/ml polybrene (added from 1000× polybrene stock). Incubate for 4 hr.
- 4a. Add 3 ml fresh medium to the cells. Allow 24 to 48 hr before measuring transduction efficiencies.

For most cell types the polybrene-containing medium does not need to be removed, as the addition of the fresh medium dilutes out the polybrene. However, if one is working with a cell type that is very sensitive to polybrene, the polybrene-containing medium can be removed before adding the fresh medium.

To transduce nonadherent cells by spinoculation

- 1b. Culture the target cells in appropriate serum-containing growth medium (e.g., IMDM containing 10% heat-inactivated FBS).
- 2b. Resuspend 5×10^5 cells in 1 ml of lentiviral vector particles in the presence of 6 $\mu\text{g/ml}$ polybrene or 4 $\mu\text{g/ml}$ protamine sulfate (added from corresponding 1000 \times stock) in a 14-ml round-bottom polypropylene centrifuge tube.
- 3b. Centrifuge 2 hr at $2000 \times g$, room temperature, to perform spinoculation. Resuspend the cell pellet in 1 ml of fresh medium and incubate overnight.
- 4b. Repeat the spinoculation (steps 2b and 3b) once a day for 2 days.

To transduce nonadherent cells using recombinant fibronectin fragment

- 1c. Prepare a stock of 9.6 mg/ml recombinant fibronectin fragment in sterile deionized, distilled water. Dilute the stock to final concentration of 9.6 $\mu\text{g/ml}$ with PBS.
- 2c. Place 1 ml of the diluted solution of recombinant fibronectin fragment in each well of a 6-well tissue culture plate and incubate for 2 hr at room temperature.
- 3c. Remove the fibronectin solution and add 1 ml of 2% BSA in PBS. Incubate for 30 min at room temperature to block. Remove the BSA solution and rinse with PBS once.

The blocking step prevents nonspecific binding.

- 4c. Add $0.5\text{--}1 \times 10^6$ cells per well. Allow at least 2 hr for the cells to attach, carefully remove the supernatant, and then add 1 ml of growth medium containing the vector particles (to completely cover the cells). Incubate overnight.

Spinoculation and overnight transduction on fibronectin can be combined to maximize transduction efficiency. Cells can be transduced by spinoculation followed by overnight culture in fresh vector-containing medium on fibronectin-coated plates, and the combined process can be repeated if necessary.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

293T cell growth medium

Dulbecco's modified Eagle's medium (DMEM) supplemented with:

10% heat-inactivated fetal bovine serum (FBS)

2 mM L-glutamine

50 U/ml penicillin

50 $\mu\text{g/ml}$ streptomycin

Store up to 1 month at 4°C

CaCl₂, 2.5 M

Dissolve 183.7 g CaCl₂ dihydrate (tissue culture grade) in deionized, distilled water and make volume up to 500 ml. Filter sterilize through a 0.22- μm nitrocellulose filter and store indefinitely at -20°C in 50-ml aliquots (can be frozen and thawed repeatedly).

HEPES-buffered saline (HeBS) (2 \times)

16.4 g NaCl (0.283 M final)

11.9 g HEPES acid (0.023 M final)

continued

**Protein
Expression**

16.22.11

0.21 g Na₂HPO₄ (1.5 mM final)
H₂O to 1 liter
Titrate to pH 7.05 with 5 M NaOH
Filter sterilize and store indefinitely at –20°C

HEPES stock solution, 1 M (100×)

Weigh 238.3 g HEPES and add water to 1 liter. Filter sterilize and store up to 6 months at 4°C.

Sodium butyrate stock solution, 500 mM

Prepare 500 mM sodium butyrate. Adjust pH to 7.0 with 10 N NaOH. Sterilize the solution by 0.22-μm filtration and store up to 6 months at 4°C.

Serum-containing growth media for transduction (Basic Protocol 7 and Alternate Protocol 3)

Dulbecco's modified Eagle medium (DMEM; for adherent cells) or Iscove's modified Dulbecco's medium (IMDM; for nonadherent cells), supplemented with:

10% (v/v) heat-inactivated FBS

2 mM L-glutamine

50 U/ml penicillin

50 μg/ml streptomycin

Store up to 1 month at 4°C

These media are given as examples; depending upon the cell type other serum-containing media may be appropriate.

COMMENTARY

Background Information

Although stable lentiviral producer cell lines have been developed, transient transfection of human 293T cells is the most widely used strategy to generate lentiviral vector particles. The 293T cell line is a derivative of 293 cells, human embryonic kidney cells transformed by the early region of adenovirus type 5 (Graham et al., 1977), into which the gene for a temperature-sensitive version of the simian virus 40 (SV-40) large T tumor antigen has been inserted (DuBridge et al., 1987). SV40 T antigen stimulates replication of plasmids containing the SV40 origin of replication. The facile transfectability of 293T cells coupled with SV40 T antigen-mediated extrachromosomal amplification of HIV-1 vector plasmids makes these cells particularly well suited for transient production of HIV-1-based lentiviral vector particles. In addition to obviating toxicity problems associated with high-level stable expression of HIV-1 proteins, this transient means of producing VSV-G-pseudotyped lentiviral vectors also circumvents difficulties associated with constitutive expression of the VSV-G envelope glycoprotein, which is fusogenic and causes syncytium formation.

As discussed in *UNIT 16.23*, most current-generation HIV-1-based vectors contain a strong constitutive internal promoter to drive transgene expression. For certain applications, it may be desirable to substitute tissue-specific or regulatable promoter sequences. In addition, a selectable marker gene is frequently incorporated into the transfer vector. Inclusion of a selectable marker gene as a reporter allows identification and/or selection of stably transduced cells. Among the commonly used selectable marker genes are the neomycin-resistance (*neo*) gene, the β-galactosidase (*lacZ*) gene, and the enhanced green fluorescent protein (EGFP) gene (Hawley et al., 2001a). Recently, a red fluorescent protein gene (drFP583) was isolated from a nonbioluminescent reef coral (Matz et al., 1999). The subsequent development of rapidly maturing versions of the drFP583 gene (Bevis and Glick, 2002) and flow cytometry protocols to simultaneously analyze up to four fluorescent proteins (Hawley et al., 2001b) provides the opportunity for facile comparative analysis of multiple constructs or other experimental variables. To express a selectable marker gene in conjunction with a gene of interest, internal ribosome entry site (IRES)

sequences are frequently utilized to generate bicistronic transcripts. The IRES sequence, which is inserted between the two genes, confers cap (m7pppGN)-independent translational initiation of the second gene through formation of a stem-loop structure and a stretch of pyrimidine-rich sequences near the internal AUG codon (Jang and Wimmer, 1990). The most commonly used IRES sequences are those derived from the picornaviruses encephalomyocarditis virus and poliovirus (Adam et al., 1991; Ghattas et al., 1991).

Critical Parameters

During transient transfection, vector particle production peaks on days 2 and 3 post-transfection and drops significantly after day 4. If the vector particles are to be used without further concentration, the optimal time to collect the supernatant is 48 hr post-transfection. If vector particles are to be concentrated by the ultracentrifugation method, the supernatants should be collected on two consecutive days. The supernatants collected on day 1 should be stored at 4°C until they are pooled with the day 2 supernatants for concentration.

When using EGFP as a reporter gene, allow ~72 hr before determining the transduction efficiencies by flow cytometry to ensure that all transduced cells are expressing EGFP at detectable levels. Analysis at this time point also minimizes the contribution of false positive signals due to pseudotransduction, which is the direct transfer of EGFP protein present in the vector supernatants or incorporated into the vector particles to the target cells (Liu et al., 1996). When determining lentiviral vector titers with reporter genes, it is also important to be aware of the fact that transgenes can be transiently expressed from unintegrated lentiviral vectors for 8 to 10 days, although generally at lower levels than from integrated vectors (Haas et al., 2000).

Troubleshooting

Low vector titers are often caused by inefficient transfection of 293T cells. Transfection efficiency can vary significantly and depends on several factors including the following. (1) When the 293T cells are plated for transfection, they must be trypsinized well to prevent clumping and allow formation of a uniform monolayer. Cells should be 50% to 70% confluent at the time of transfection. (2) Plasmid DNA of high purity and quality is crucial to achieve high-efficiency transfection. DNA concentra-

tion and quality should be determined by spectrophotometric analysis and gel electrophoresis. Commercially available plasmid DNA extraction kits can be used to obtain highly purified supercoiled plasmid DNA. Prior to transfection, the DNA should be sterilized by ethanol precipitation. The air-dried pellet is then resuspended in sterile deionized, distilled water. (3) The pH of the 2× HeBS solution is particularly critical and should be exactly 7.05. It is important to bubble the 2× HeBS solution with air as the DNA/CaCl₂ solution is added dropwise. Once added to the cells, a fine precipitate should develop that is readily visible under the microscope.

Vector titers are also influenced by the stability of the vector particles, which is affected by factors such as temperature, pH, freeze-and-thaw frequency, and incubation conditions. VSV-G-pseudotyped HIV-1 vector particles have a half-life of 10.4 ± 1.2 hr at 37°C, 1 to 2 days at room temperature, and ~1 week at 4°C (Higashikawa and Chang, 2001). Changes in pH can dramatically affect vector stability. Although VSV-G-pseudotyped HIV-1 vectors are stable at pH 7.0, their half-life at pH 6.0 or 8.0 is <10 min (Higashikawa and Chang, 2001). In order to prevent pH changes in vector supernatants, add HEPES buffer to the 293T medium at a final concentration of 10 mM.

Anticipated Results

Transfection efficiencies of 293T cells should be in the range of 60% to 80%. A typical vector carrying a 3- to 4-kb transgene should produce a titer of $\sim 10^6$ to 10^7 TU/ml following transient transfection of 293T cells. Titers can be routinely increased to 10^9 TU/ml by concentrating the vector supernatants. Expect a 50% to 75% recovery following vector concentration. Titers will generally be lower with inserts greater than 5 kb. A titer of 10^8 TU/ml should allow efficient transduction of most cell types.

Time Considerations

Although developing and testing new lentiviral vectors can be considerably time-consuming, it should not take more than a few hours a day over a period of less than a week to subclone a DNA fragment into an existing lentiviral vector backbone. Generation of vector particles by transient transfection requires 4 to 5 consecutive days. Concentrating and purifying the vector particles requires 2 to 4 days, titering takes 3 to 10 days, and testing for RCR entails a couple of weeks.

Literature Cited

- Adam, M.A., Ramesh, N., Miller, A.D., and Osborne, W.R. 1991. Internal initiation of translation in retroviral vectors carrying picornavirus 5' nontranslated regions. *J. Virol.* 65:4985-4990.
- Bahnson, A.B., Dunigan, J.T., Baysal, B.E., Mohney, T., Atchison, R.W., Nimgaonkar, M.T., Ball, E.D., and Barranger, J.A. 1995. Centrifugal enhancement of retroviral mediated gene transfer. *J. Virol. Methods* 54:131-143.
- Bevis, B.J., and Glick, B.S. 2002. Rapidly maturing variants of the *Discosoma* red fluorescent protein (DsRed). *Nature Biotechnol.* 20:83-87.
- Case, S.S., Price, M.A., Jordan, C.T., Yu, X.J., Wang, L., Kohn, D.B., and Crooks, G.M. 1999. Stable transduction of quiescent CD34⁺CD38⁻ human hematopoietic cells by HIV-1-based lentiviral vectors. *Proc. Natl. Acad. Sci. U.S.A.* 96:2988-2993.
- Chuck, A.S., Clarke, M.F., and Palsson, B.O. 1996. Retroviral infection is limited by Brownian motion. *Hum. Gene Ther.* 7:1527-1534.
- Dao, M.A., Hashino, K., Kato, I., and Nolta, J.A. 1998. Adhesion to fibronectin maintains regenerative capacity during ex vivo culture and transduction of human hematopoietic stem and progenitor cells. *Blood* 92:4612-4621.
- Donahue, R.E., Sorrentino, B.P., Hawley, R.G., An, D.S., Chen, I.S., and Wersto, R.P. 2001. Fibronectin fragment CH-296 inhibits apoptosis and enhances ex vivo gene transfer by murine retrovirus and human lentivirus vectors independent of viral tropism in nonhuman primate CD34⁺ cells. *Mol. Ther.* 3:359-367.
- DuBridge, R.B., Tang, P., Hsia, H.C., Leong, P.M., Miller, J.H., and Calos, M.P. 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol. Cell Biol.* 7:379-387.
- Gasmi, M., Glynn, J., Jin, M.-J., Jolly, D.J., Yee, J.-K., and Chen, S.-T. 1999. Requirements for efficient production and transduction of human immunodeficiency virus type 1-based vectors. *J. Virol.* 73:1828-1834.
- Gervais, A., West, D., Leoni, L.M., Richman, D.D., Wong-Staal, F., and Corbeil, J. 1997. A new reporter cell line to monitor HIV-1 infection and drug susceptibility in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 94:4653-4658.
- Ghatts, I.R., Sanes, J.R., and Majors, J.E. 1991. The encephalomyocarditis virus internal ribosome entry site allows efficient coexpression of two genes from a recombinant provirus in cultured cells and in embryos. *Mol. Cell Biol.* 11:5848-5859.
- Gorman, C.M. and Howard, B.H. 1983. Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate. *Nucleic Acids Res.* 11:7631-7648.
- Graham, F.L. and van der Eb, A.J. 1973. Transformation of rat cells by DNA of human adenovirus 5. *Virology* 54:536-539.
- Graham, F.L., Smiley, J., Russell, W.C., and Nairn, R. 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* 36:59-74.
- Haas, D.L., Case, S.S., Crooks, G.M., and Kohn, D.B. 2000. Critical factors influencing stable transduction of human CD34⁺ cells with HIV-1-derived lentiviral vectors. *Mol. Ther.* 2:71-80.
- Hanenberg, H., Hashino, K., Konishi, H., Hock, R.A., Kato, I., and Williams, D.A. 1997. Optimization of fibronectin-assisted retroviral gene transfer into human CD34⁺ hematopoietic cells. *Hum. Gene Ther.* 8:2193-2206.
- Hawley, T.S., Sabourin, L.A., and Hawley, R.G. 1989. Comparative analysis of retroviral vector expression in mouse embryonal carcinoma cells. *Plasmid* 22:120-131.
- Hawley, T.S., Telford, W.G., and Hawley, R.G. 2001a. "Rainbow" reporters for multispectral marking and lineage analysis of hematopoietic stem cells. *Stem Cells* 19:118-124.
- Hawley, T.S., Telford, W.G., Ramezani, A., and Hawley, R.G. 2001b. Four-color flow cytometric detection of retrovirally expressed red, yellow, green and cyan fluorescent proteins. *BioTechniques* 30:1028-1034.
- Higashikawa, F. and Chang, L. 2001. Kinetic analyses of stability of simple and complex retroviral vectors. *Virology* 280:124-131.
- Jang, S.K. and Wimmer, E. 1990. Cap-independent translation of encephalomyocarditis virus RNA: Structural elements of the internal ribosomal entry site and involvement of a cellular 57-kD RNA-binding protein. *Genes Dev.* 4:1560-1572.
- Klages, N., Zufferey, R., and Trono, D. 2000. A stable system for the high-titer production of multiply attenuated lentiviral vectors. *Mol. Ther.* 2:170-176.
- Kotani, H., Newton III, P.B., Zhang, S., Chiang, Y. L., Otto, E., Weaver, L., Blaese, R.M., Anderson, W.F., and McGarrity, G.J. 1994. Improved methods of retroviral vector transduction and production for gene therapy. *Hum. Gene Ther.* 5:19-28.
- Laughlin, M.A., Chang, G.Y., Oakes, J.W., Gonzalez-Scarano, F., and Pomerantz, R.J. 1995. Sodium butyrate stimulation of HIV-1 gene expression: A novel mechanism of induction independent of NF- κ B. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* 9:332-339.
- Liu, M.L., Winther, B.L., and Kay, M.A. 1996. Pseudotransduction of hepatocytes by using concentrated pseudotyped vesicular stomatitis virus G glycoprotein (VSV-G)-Moloney murine leukemia virus-derived retrovirus vectors: Comparison of VSV-G and amphotropic vectors for hepatic gene transfer. *J. Virol.* 70:2497-2502.
- Matz, M.V., Fradkov, A.F., Labas, Y.A., Savitsky, A.P., Zarausky, A.G., Markelov, M.L., and Lukyanov, S.A. 1999. Fluorescent proteins from nonbioluminescent *Anthozoa* species. *Nature Biotech.* 17:969-973.
- McKnight, G.S., Hager, L., and Palmiter, R.D. 1980. Butyrate and related inhibitors of histone deacetylation block the induction of egg white genes by steroid hormones. *Cell* 22:469-477.

- Miller, D.L., Meikle, P.J., and Anson, D.S. 1996. A rapid and efficient method for concentration of small volumes of retroviral supernatant. *Nucleic Acids Res.* 24:1576-1577.
- Miyoshi, H., Smith, K.A., Mosier, D.E., Verma, I.M., and Torbett, B.E. 1999. Transduction of human CD34⁺ cells that mediate long-term engraftment of NOD/SCID mice by HIV-1 vectors. *Science* 283:682-686.
- Morgan, J.R., LeDoux, J.M., Snow, R.G., Tompkins, R.G., and Yarmush, M.L. 1995. Retrovirus infection: Effect of time and target cell number. *J. Virol.* 69:6994-7000.
- Moritz, T., Patel, V.P., and Williams, D.A. 1994. Bone marrow extracellular matrix molecules improve gene transfer into human hematopoietic cells via retroviral vectors. *J. Clin. Invest.* 93:1451-1457.
- Naldini, L., Blomer, U., Gally, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M., and Trono, D. 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272:263-267.
- Norton, V.G., Imai, B.S., Yau, P., and Bradbury, E.M. 1989. Histone acetylation reduces nucleosome core particle linking number change. *Cell* 57:449-457.
- Ramezani, A., Hawley, T.S., and Hawley, R.G. 2000. Lentiviral vectors for enhanced gene expression in human hematopoietic cells. *Mol. Ther.* 2:458-469.
- Rasheed, S., Nelson-Rees, W.A., Toth, E.M., Arnstein, P., and Gardner, M.B. 1974. Characterization of a newly derived human sarcoma cell line (HT-1080). *Cancer* 33:1027-1033.
- Rasko, J.E.J. 1999. Reporters of gene expression: Autofluorescent proteins. In *Current Protocols in Cytometry* (J.P. Robinson, Z. Darzynkiewicz, P.N. Dean, A.R. Hibbs, A. Orfao, P.S. Rabinovich, and L.L. Wheeless, eds.) pp. 9.12.1-9.12.16. John Wiley & Sons, New York.
- Sambrook, J., and Russell, D. 2001. *Molecular Cloning: A Laboratory Manual*, 3rd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Scherr, M., Battmer, K., Blomer, U., Schiedlmeier, B., Ganser, A., Grez, M., and Eder, M. 2002. Lentiviral gene transfer into peripheral blood-derived CD34⁺ NOD/SCID-repopulating cells. *Blood* 99:709-712.
- Soneoka, Y., Cannon, P.M., Ramsdale, E.E., Griffiths, J.C., Romano, G., Kingsman, S.M., and Kingsman, A.J. 1995. A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucleic Acids Res.* 23:628-633.
- Vigna, E. and Naldini, L. 2000. Lentiviral vectors: Excellent tools for experimental gene transfer and promising candidates for gene therapy. *J. Gene Med.* 2:308-316.
- Yokota, T., Oritani, K., Mitsui, H., Aoyama, K., Ishikawa, J., Sugahara, H., Matsumura, I., Tsai, S., Tomiyama, Y., Kanakura, Y., and Matsuzawa, Y. 1998. Growth-supporting activities of fibronectin on hematopoietic stem/progenitor cells in vitro and in vivo: Structural requirement for fibronectin activities of CS1 and cell-binding domains. *Blood* 91:3263-3272.
- Zhang, B., Xia, H.Q., Cleghorn, G., Gobe, G., West, M., and Wei, M.Q. 2001. A highly efficient and consistent method for harvesting large volumes of high-titre lentiviral vectors. *Gene Ther.* 8:1745-1751.
- Zufferey, R., Nagy, D., Mandel, R.J., Naldini, L., and Trono, D. 1997. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat. Biotechnol.* 15:871-875.

Contributed by Ali Ramezani
American Red Cross
Rockville, Maryland

Robert G. Hawley
American Red Cross
Rockville, Maryland
and The George Washington University
Washington, D.C.

The ability to select for integration of plasmid DNA into the host chromosome allows the generation of stably transfected cell lines. With transfection of a selectable marker linked to a nonselectable target gene (or by cotransfection of the two unlinked genes), high-level expression of the desired gene is obtained by selecting for amplification of the selectable marker.

This unit presents two systems for gene amplification and expression. The first (see Basic Protocol 1) describes the dihydrofolate reductase (DHFR) selection system while the second (see Basic Protocol 2) is based on selection of the glutamine synthetase (GS) gene. The DHFR system is probably more widely used, and results in very high levels of amplification (up to 1000 copies per cell in some cases) and expression; however, the DHFR amplification process is lengthy and may require several months to isolate and characterize a stable, amplified line. In contrast, the GS system typically requires only a single round of selection for amplification to achieve maximal expression levels. In this system the length of time necessary to isolate stably amplified clones from the primary transfectants is dramatically reduced to 2 months or less.

AMPLIFICATION USING DIHYDROFOLATE REDUCTASE

The pED series of dicistronic vectors (Fig. 16.23.1) can be used to obtain high-level expression of a targeted gene in stably transfected cells. These vectors carry a cloning sequence for insertion of the target gene followed by the selectable and amplifiable marker gene, dihydrofolate reductase (DHFR). Alternatively, a plasmid expressing the gene of interest and a plasmid expressing DHFR can be cotransfected. DHFR-deficient CHO cells transformed with the appropriate vector(s) are selected by ability to grow in nucleoside-free medium. Subsequent selective cycling in the presence of increasing concentrations of methotrexate (MTX)—a potent inhibitor of DHFR function—results in amplification of the integrated DNA and increased expression of the desired gene product.

Materials

pED (Kaufman et al., 1991) expressing appropriate cDNA; *or* pCVSVEII-DHFR or pAd26SV(A) (Kingston et al., 1984; Kaufman and Sharp, 1982a) and a separate vector expressing appropriate cDNA

CHO DXB11 or CHO DG44 cell lines (available from Lawrence Chasin, Columbia University) or CHO GRA (available from Randal Kaufman, University of Michigan)

Complete ADT medium (see recipe)

10% glycerol

Dialyzed fetal bovine serum (FBS; see recipe)

Complete α^- medium (Life Technologies) with 10% dialyzed FBS

Sterile vacuum grease

0.05% trypsin/0.6 mM EDTA in PBS (see APPENDIX 2 for PBS), 37°C

2% methylene blue in 50% ethanol (optional)

5 mM methotrexate (see recipe)

Cloning cylinders (see recipe)

Additional reagents and equipment for subcloning (UNIT 3.16), and either CaPO₄-mediated transfection (UNIT 9.1), electroporation (UNIT 9.3), or liposome-mediated transfection (UNIT 9.4)

NOTE: All tissue culture incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise indicated.

BASIC PROTOCOL 1

Protein Expression

16.23.1

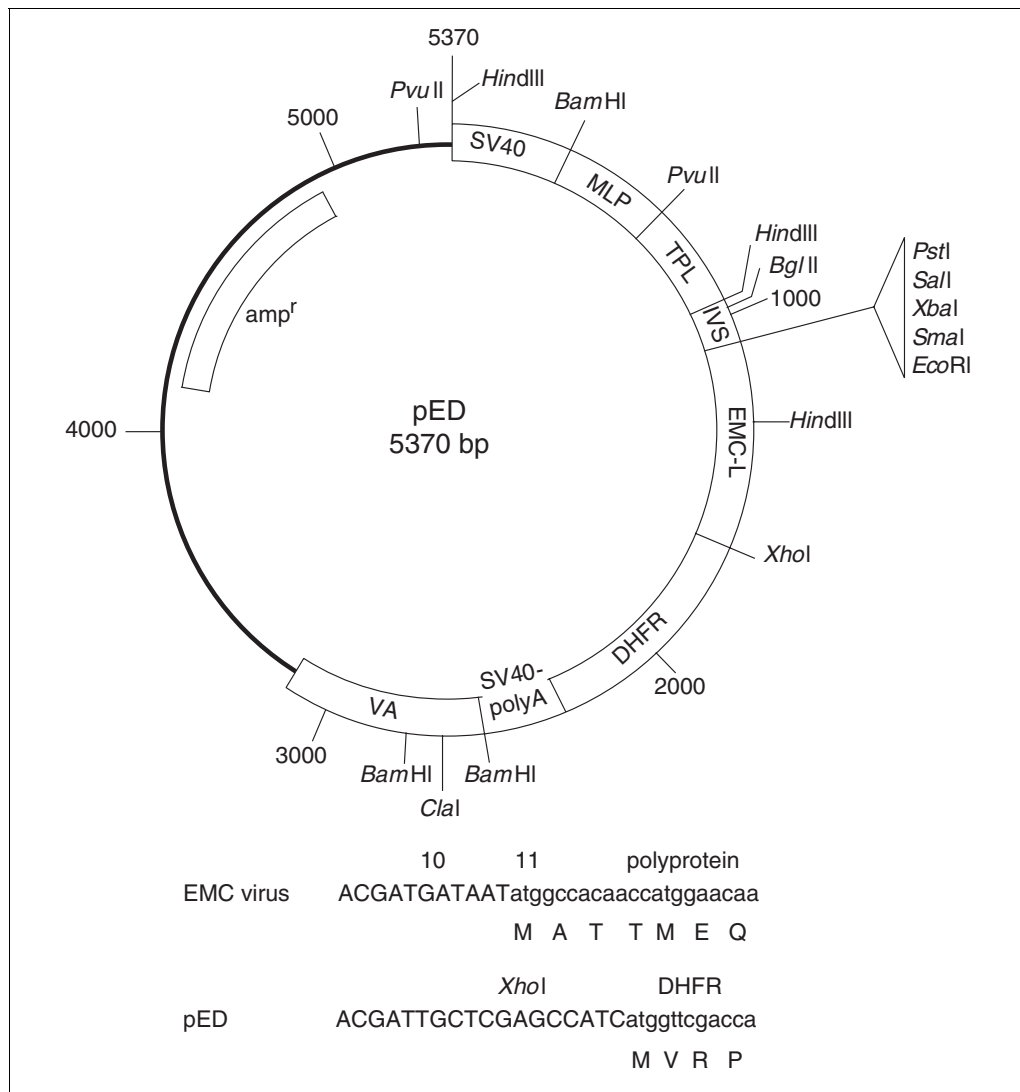


Figure 16.23.1 Map of dicistronic mRNA expression vector pED. The components of the 5360-bp pED expression vector in the pUC18 background are indicated as follows: SV40, *HindIII*-*PvuII* fragment containing the SV40 origin of replication and enhancer element; MLP, adenovirus major late promoter from the *XhoI* site (15.83 map units, m.u.) to the 5' cap site (16.55 m.u.); TPL, 180 bp of the first two and $\frac{2}{3}$ of the third leaders from adenovirus major late mRNAs; IVS, a hybrid intron composed of the 5' splice site from the first leader of adenovirus major late mRNAs and a 3' splice site from an immunoglobulin gene; *PstI* and *EcoRI* unique cloning sites; EMC-L, the 5' untranslated leader from EMC virus (nucleotides 260-827); DHFR, a murine DHFR coding region; SV40-polyA, the SV40 early polyadenylation signal; VA, the adenovirus VAI RNA gene from the *HpaI* (28.02 m.u.) to the *BalI* (29.62 m.u.); and β -lactamase, a selectable gene for propagation in *E. coli*. Below is indicated the sequence junction of the EMC-L and DHFR as compared to the context of the AUG 11 which is initiation codon for the EMC virus polyprotein. A unique *XhoI* restriction site is available for insertion of other coding regions to be translated from the EMC virus leader. Adapted with permission from IRL Press.

Transfect CHO cells

1. The day before transfection, split a confluent dish of CHO cells 1:15 in complete ADT medium.
2. Transfect the cells with 5 to 10 μg plasmid DNA (per dish) from step 1, using either electroporation (UNIT 9.3), the calcium phosphate technique (UNIT 9.1), or liposome-mediated transfection (UNIT 9.4).

The calcium phosphate treatment followed by glycerol shock works well for CHO DXB11 cells. A 3-min shock with 10% glycerol should be performed 4 to 6 hr after the DNA precipitate is placed on the cells. The gene can then be introduced into pED that also contains a DHFR gene. Alternatively, if a pED vector is unavailable, it is possible to introduce the desired gene by cotransfection using two separate plasmids. In this latter case, transfected DNA should contain the plasmid whose amplification is desired and a plasmid expressing the DHFR gene from a strong promoter [e.g., pCVSVEII-DHFR or pAdd26SV(A)]. Use a 5:1 molar ratio of the gene of interest to the DHFR gene.

It is not necessary to physically link the DHFR gene to the gene whose amplification is desired. The two genes will integrate in the same region of the chromosome and will coamplify (Kaufman and Sharp, 1982b). Using one-fifth the molar amount of the DHFR gene makes it likely that most of the selected cell lines will contain the gene of interest in an intact form as well as the DHFR gene.

- 3a. *For cells transfected by electroporation or calcium phosphate:* Allow the cells to reach confluence after transfection (this should occur after 2 to 3 doublings in 2 days.) Split each dish 1:15 into α^- medium containing 10% dialyzed FBS (complete α^- medium).
- 3b. *For cells transfected using liposomes:* Add 5 ml complete α^- medium and incubate overnight. Remove medium, wash twice with 37°C PBS, add 5 ml complete α^- medium, and incubate 2 days. Dilute cells 1:10 or 1:15 into complete α^- medium without ADT.

Complete α^- medium (containing no added nucleosides) is a selective medium, as cells need DHFR to synthesize necessary nucleosides. Use of dialyzed serum is necessary to avoid addition of nucleosides present in normal FBS. Note that methotrexate is not needed for selection.

4. Incubate cells 10 to 12 days if proceeding from step 3a or 14 days if proceeding from step 3b. Move the dishes as infrequently as possible during this time to prevent formation of sibling colonies.

Cells can float away from their original colonies, land elsewhere on the dish, and produce a colony of their own. This should be minimized, as picking and analyzing two such sibling colonies is inefficient (they presumably are identical).

Pick stable colonies

5. At a time point 10 to 14 days after placing the cells in selective medium, check the dishes for colonies by holding the dish above one's head at an angle with respect to the overhead lights and looking for opaque patches. Circle such patches with a laboratory marker so that they can easily be located and examined in the phase-contrast microscope.

In order to determine how well the transfection worked, one of the transfected dishes can be stained with methylene blue. To stain, first aspirate off the medium, then place ~2 ml of a 2% methylene blue solution (made up in 50% ethanol) on each dish. Wait 2 min, then pour the dye solution off and wash off the residual methylene blue by dipping the dish in a bucket of cold water. In order to have good success picking colonies for stable lines, each dish should have several heavily staining colonies as well as 10 more smaller colonies.

6. Select the colonies to be picked. Circle the chosen colonies with a laboratory marker to determine where to place the cloning cylinders.

Choose only large, healthy colonies. Colonies should have ~500 cells, and the cells should appear to be compact and polygonal. Colonies with many flat and spread-out cells should be avoided, as this morphology indicates that they are not making very much DHFR. Pick ~20 colonies so that in the end there will be many stably transfected cell lines available for amplification. Keep track of which dish a colony comes from, as colonies from the same dish may be siblings. This is most easily done by numbering the dishes and using that number in the name of the colony that has been picked.

It is also possible at this point to pool large numbers of transformants that have integrated the vector into different sites. Because different integration sites have quite different potentials for amplification, one can use sequential increases in MTX resistance to select cells rapidly that have amplified the gene to high copy number.

7. Coat one end of a cloning cylinder with sterile vacuum grease by touching the cylinder to grease that has been autoclaved in a glass petri dish. Gently place the cylinder around the colony to be picked (Fig. 16.23.2).

Make certain that there is not too much grease on the end of the cloning cylinder—use a sufficient amount of grease to form a thin film between the cloning cylinder and the tissue culture dish, but do not allow the grease to cover any of the colony. The cloning cylinder is most easily positioned using bent forceps sterilized by flaming in ethanol immediately prior to use.

8. Using a sterile Pasteur pipet, rinse the colony with 37°C 0.05% trypsin/0.6 mM EDTA by filling and emptying the cloning cylinder (Fig. 16.23.2).

Care must be taken to avoid knocking the cylinder over, and also to avoid scraping cells off the colony with the end of the Pasteur pipet. Hold the Pasteur pipet as close to vertically as possible in order to avoid knocking over the cloning cylinder. A little trypsin can be left in the well after rinsing.

9. Add 3 drops of 37°C trypsin/EDTA to the cloning cylinder. Wait 1 min. Fill the cloning cylinder with medium and repeatedly run the contents of the cylinder in and out through a Pasteur pipet in order to remove the trypsinized cells from the dish and disperse them. Plate the cells in a 40-mm dish.

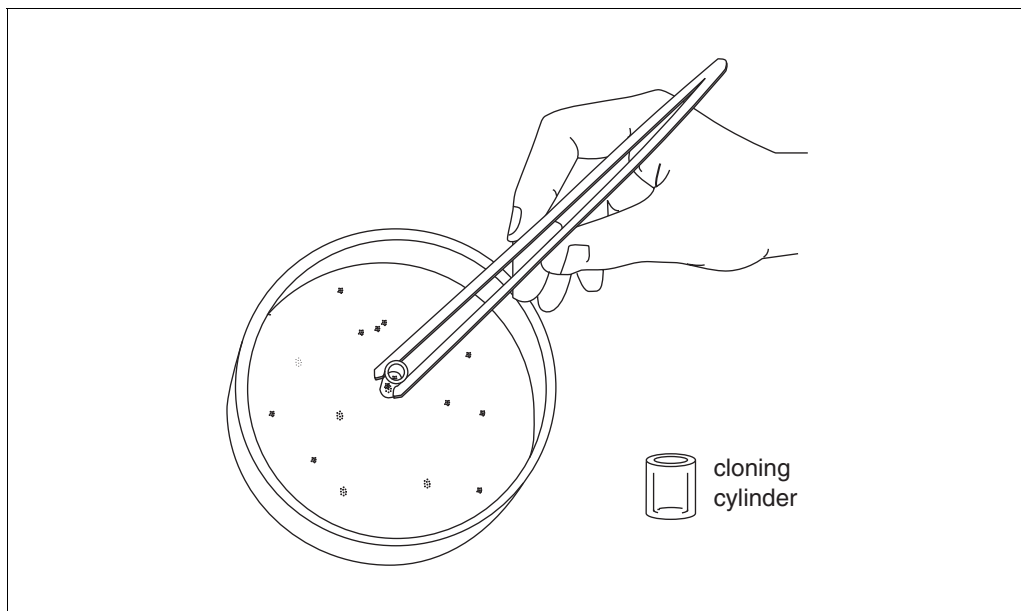


Figure 16.23.2 Placement of cloning cylinder around CHO colony.

10. As the cells grow out, split them frequently (every 4 to 5 days or so) so that they do not form large colonies.

The central cells in large colonies do not fare well.

Amplify stable transfectant

Amplification is a long process. Before amplifying a stable transfectant, one should be sure that the gene of interest has indeed been integrated into the cell in a functional form. This can be done by examining the cellular DNA by Southern analysis (UNIT 2.9), by examining the cellular RNA (Chapter 4), or most easily in many cases by using a functional assay for the introduced protein. Some cell lines amplify more readily than others (chromosome location of the introduced DNA appears to play a role), and the gene of interest can rearrange or mutate during the process. These considerations make it prudent to amplify six or more stable transfectants or pools of cells containing the gene of interest at the same time.

11. Split a confluent dish of cells growing in complete α^- medium 1:6 into two dishes complete α^- medium supplemented with 0.005 μM methotrexate (added from 5mM methotrexate stock).

CAUTION: *Use gloves when handling methotrexate as it is carcinogenic.*

The addition of methotrexate to the medium increases the level of selection, as methotrexate is a potent inhibitor of DHFR. Empirically, 0.005 μM methotrexate requires ~4-fold more DHFR to be made in the cell than does complete α^- medium with no methotrexate. By splitting the cells into this medium, one is selecting for cells making elevated levels of DHFR. This is generally accomplished by increasing the copy number of the transfected DHFR gene.

12. The cells should grow to confluence fairly readily. When they do, split them 1:6 again.

The cells will probably grow more slowly and take on a flat, spread-out morphology. This indicates that they are starved for DHFR.

Cells grow well immediately after splitting into a higher level of methotrexate because they have an endogenous reserve of nucleosides that needs to be depleted before the selection takes place.

13. Keep splitting the cells 1:6 into complete α^- medium supplemented with 0.005 μM methotrexate. When their rate of growth increases and when they begin to take on a more normal morphology, increase the degree of the split to 1:8, then 1:10, then 1:15. When the cells grow to confluence in 3 days from a 1:15 split and have regained a polygonal morphology, perform the next amplification step.

It is possible that cells will immediately grow well in 0.005 μM methotrexate, indicating that they are already making enough DHFR to survive at this level of selection. If so, switch immediately to 0.02 μM methotrexate. Many researchers start out by placing the DHFR-containing cell line in 0.005, 0.02, and 0.05 μM methotrexate (at step 11) to see at what level the initial recombinant lines can survive.

14. Repeat the above process (steps 12 and 13) using complete α^- medium supplemented with 0.02 μM methotrexate.

15. Continue amplifying by increasing the level of methotrexate in the medium by 4-fold increments. Continue until the cells are growing in 20 to 80 μM methotrexate (the cells should now contain 500 to 2000 copies of the transfected gene).

Be sure to freeze samples of the cells at each amplification step to avoid having to go back to the beginning in the event of contamination. Each amplification step should take 3 to 4 weeks. Use of methotrexate levels above 80 μM does not result in much further amplification, as the ability of the cell to transport the drug becomes limiting.

AMPLIFICATION BY CLONING AT EACH SELECTIVE STEP

The above procedure is straightforward and does not require very much hands-on time. It does not, however, necessarily result in a clonal cell line at the end of each passage. The protocol requires a long time (9 to 12 months) to generate an amplified line. One may clone lines from the final amplified line to see whether some clones express more of a desired protein than others. Alternatively, one may amplify by cloning cells at every step and selecting for those that retain high levels of production of the protein of interest. This is a particularly attractive approach if the desired protein is secreted and easily assayed from the medium. This second approach may result in an overproducing line more rapidly, but it requires substantially more effort. If this approach is desired, substitute the following for steps 11 to 14 of Basic Protocol 1.

- 11a. Split a confluent dish of cells growing in complete α^- medium 1:15 into complete α^- supplemented with 0.02 μM methotrexate (8 dishes) and into complete α^- medium supplemented with 0.08 μM methotrexate (7 dishes). Feed every 4 days with the appropriate selective medium.

The goal of this step is to find individual cells that have amplified the DHFR gene enough to grow in a significantly higher level of methotrexate. These cells will expand into colonies during the 10- to 12-day incubation. As described in step 4, sibling colonies must be avoided, so dishes should be disturbed as little as possible.

- 12a. Check the dishes for colonies after 10 to 12 days and pick healthy colonies, as in steps 5 to 10 of the basic protocol.

- 13a. Expand the colonies and check for the level of expression of the desired product. Choose a colony that is producing good levels of the desired product, and repeat steps 11a and 12a using levels of methotrexate 16- and 64-fold higher than the level of methotrexate in which the colony is growing.

Each of these rounds will take ~1 month, and after three to four rounds, the cells should be growing in 80 μM methotrexate and contain highly amplified sequences.

AMPLIFICATION USING GLUTAMINE SYNTHETASE

In the glutamine synthetase (GS) gene amplification system, a cDNA or genomic coding sequence is inserted into the multilinker cloning site of the plasmid pEE14 (Fig. 16.23.3) such that it is expressed from the powerful hCMV promoter-enhancer. pEE14 also contains a glutamine synthetase gene that can be used as a dominant selectable marker in a variety of cell lines including CHO K1. The GS gene expressed from the plasmid confers resistance to a low level of the GS inhibitor methionine sulfoximine (MSX). CHO cells transformed with the vector are selected for lines containing increased numbers of copies of the vector using increased levels of MSX in a single round of amplification.

Materials

- Plasmid vector pEE14 (Celltech)
- Complete Glasgow modified Eagle medium containing 10% dialyzed FBS (complete GMEM-10; see recipe)
- CHO K1 cell line (ATCC #CCL61)
- 100 mM L-methionine sulfoximine (MSX; Sigma) prepared in PBS (see APPENDIX 2 for PBS; filter sterilize MSX solution and store in aliquots at -20°C ; handle carefully)

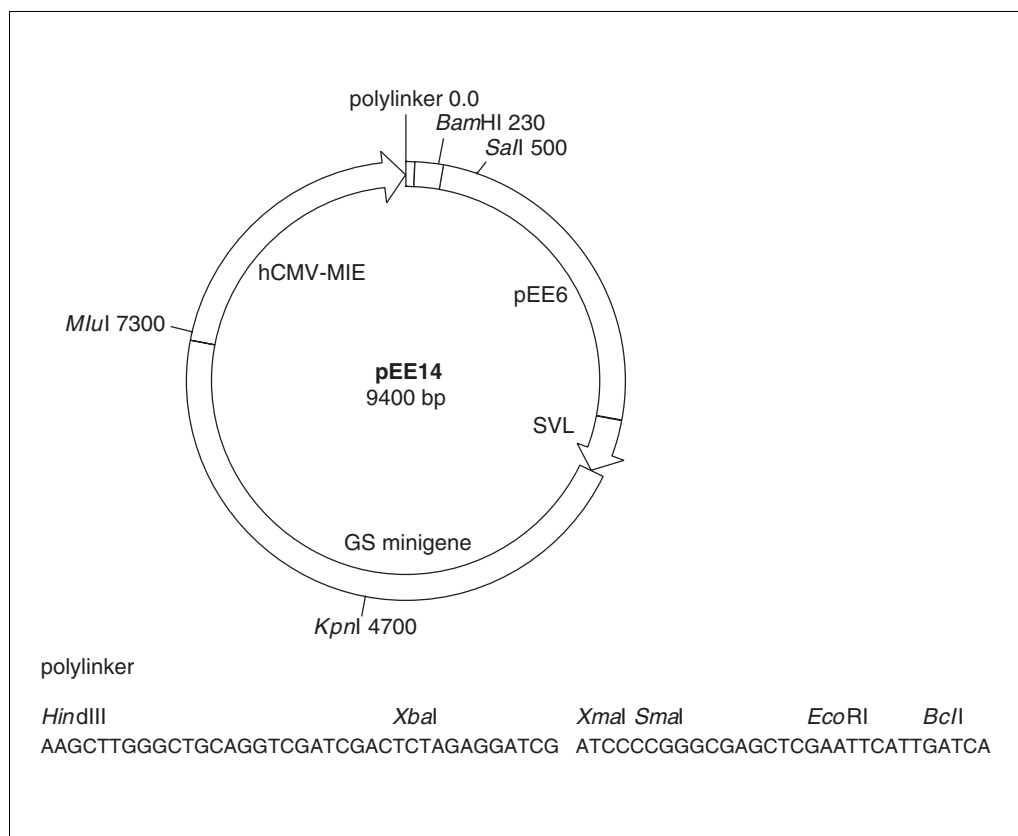


Figure 16.23.3 Map of pEE14 GS expression vector. pEE14 (~9.4 kb in length) contains a GS minigene as the selectable marker which has a single intron and GS polyadenylation signals and is driven from an SV40 late promoter. The hCMV-MIE promoter-enhancer and 5' untranslated region are used to express the gene of interest and the remainder of the plasmid contains an ampicillin-resistance gene and replication origin for replication in *E. coli*. The plasmid was constructed as follows. A 900-bp *EcoRI* fragment from the cDNA clone λ gs1.1 (Hayward et al., 1986) was assembled with a 3.4-kb *EcoRI*-*SacI* hamster GS genomic fragment from pGS1 (Sanders and Wilson, 1984), which provides the 3' end of the minigene. (The *SacI* site was converted to a *BamHI* site to facilitate vector construction.) The *EcoRI* site within the GS coding sequence was destroyed by site-directed mutagenesis without altering the amino acid sequence and a *HindIII* site in GS 3'-flanking DNA was destroyed by digestion with *HindIII*, filling in the single-stranded ends, and religation. A 340-bp SV40 late promoter (Cockett et al., 1990) was added to the 5' end as a *BamHI*-*EcoRI* fragment and the *EcoRI* site between the promoter and the GS sequences was destroyed by filling in. The resulting 4.5-kb *BamHI* fragment was inserted into pEE6hCMV (Stephens and Cockett, 1989) at a single *BglII* site upstream of the hCMV enhancer (hence destroying the *BglII* and *BamHI* sites) to form pEE14. The resulting SV40-GS minigene in pEE14 is functionally equivalent to that in pSVLGS.1 (Bebbington and Hentschel, 1987) but has been deleted of *EcoRI* and *HindIII* sites. Polylinker sequence of pEE14 is shown below.

Additional reagents and equipment for subcloning (UNIT 3.16), CaPO_4 -mediated transfection and glycerol shock (UNIT 9.1), and cloning by limiting dilution (UNIT 11.8)

NOTE: All tissue culture incubations are performed in a humidified, 37°C, 5% CO_2 incubator unless otherwise indicated.

1. Subclone the target gene into the appropriate site within the polylinker of plasmid pEE14 (Fig. 16.23.3).
2. Maintain CHO K1 cells growing exponentially in complete GMEM-10. The day

before transfection, trypsinize the cells and seed several 9-cm petri dishes at 10^6 cells per dish.

3. Introduce 10 μg circular plasmid DNA (per dish) from step 1 into the cells using calcium phosphate-mediated transfection followed by glycerol shock. “Mock”-transfect several plates without added DNA.
4. After 24 hr, replace the medium with fresh complete GMEM-10 containing MSX at a final concentration of 25 μM (selective medium).

CAUTION: MSX is toxic and should be handled carefully.

5. After 4 to 5 days, refeed the plates with fresh selective medium and wait for MSX-resistant colonies to appear, typically two weeks after infection.
6. Score the number of MSX-resistant colonies on transfected and “mock”-transfected plates.

There should be 20 to 30 colonies per plate on transfected dishes and <5 colonies per “mock”-transfected plate.

7. Isolate several independent transfected cell lines producing significant amounts of the desired product (see Basic Protocol 1, steps 6 to 11) for amplification using DHFR. Plate out each cell line on several petri dishes at a density of $\sim 10^6$ cells per dish in complete GMEM-10. Incubate 24 hr.

Whenever trypsinizing GS-selected cells, leave the cells for 24 hr to recover before reapplying MSX. In our experience, independent transfectants amplify more efficiently than pools of transfectants.

8. Replace the medium with fresh selective medium containing various concentrations of MSX, ranging between 100 μM and 1 mM.
9. Incubate the dishes 10 to 14 days, changing the medium once during this time.

After this time considerable cell death should have occurred and colonies resistant to the higher levels of MSX should have appeared. The maximum concentration of MSX at which colonies survive will depend on the particular initial transfectant, but is typically between 250 μM and 500 μM .

10. Isolate the colonies at the highest MSX concentration yielding several discrete colonies. The colonies can either be picked and assayed individually or all colonies from one initial cell line can be pooled and assayed together.

The increased production rate can be up to 10-fold in this first round of amplification. It is not normally appropriate to select for subsequent rounds of amplification because the production rate does not usually increase significantly at higher levels of MSX.

11. Clone the amplified cells with high production rates by limiting-dilution cloning.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Cloning cylinders

Glass or metal cloning cylinders can be purchased (e.g., stainless steel Penicylinders, Fisher). Place clean cylinders well rinsed in distilled water in 95% ethanol. Sterilize by flaming immediately prior to use. Alternatively, disposable cloning cylinders can be prepared by cutting off the fat end of a 200- μl pipettor tip ~ 1 cm from the large opening with a razor blade. Sterilize by autoclaving.

Complete ADT medium

α^- medium (Life Technologies) supplemented with:

- 10 μ g/ml adenosine (Sigma)
- 10 μ g/ml deoxyadenosine (Sigma)
- 10 μ g/ml thymidine (Sigma)
- 10% FBS

It is convenient to prepare stocks of the nucleosides for addition to the medium. Adenosine (1 mg/ml), deoxyadenosine (1 mg/ml), and thymidine (4 mg/ml) stocks are prepared using distilled water, and are filter sterilized.

Complete GMEM-10

Add the following in order given, using aseptic technique:

- 500 ml Glasgow modified Eagle medium (GMEM) without tryptose phosphate broth (Life Technologies, but made *without* glutamine; must be made to order as it is not a stock item)
- 5 ml 100 \times nonessential amino acids (Life Technologies)
- 5 ml G+A (see recipe)
- 5 ml 100 mM sodium pyruvate (Life Technologies)
- 10 ml 50 \times nucleoside mix (see recipe)
- 50 ml dialyzed FBS (see recipe)
- 5 ml of 5000 U/ml penicillin/streptomycin (Life Technologies)

It is essential to use dialyzed FBS when performing GS selection because serum contains significant amounts of glutamine.

Dialyzed fetal bovine serum (FBS)

Purchase dialyzed FBS from commercial supplier (e.g., Life Technologies or J.R.H. Biosciences) or prepare as follows:

1. Heat inactivate FBS at 56°C for 60 min.
2. Soak Spectrapor dialysis tubing (MWCO 6000 to 8000) in PBS (*APPENDIX 2*). Remove, rinse tubing, clip one end closed, and fill with the heat-inactivated FBS.
3. Dialyze (also see *APPENDIX 3C*) 6 to 8 hr in cold room against PBS. Change dialysis solution at least once.
4. Filter sterilize using a 0.02-mm filter and store frozen (–20°C) in 50-ml aliquots.

Glutamate + asparagine (G+A)

- 600 mg L-glutamic acid (Sigma)
- 600 mg L-asparagine (Sigma)
- H₂O to 100 ml
- Filter sterilize using a 2- μ m filter and store at 4°C

Methotrexate, 5 mM stock

Dissolve methotrexate to 5 mM in α^- medium (Life Technologies) and filter sterilize. Store at –20°C in a foil-wrapped container.

It is important to make a large 5 mM stock solution before starting amplification, then dilute that stock solution in α^- medium for the various levels of selective media. The potency of methotrexate can vary somewhat from lot to lot, making it desirable to use one stock throughout the amplification process.

CAUTION: *Methotrexate is a carcinogen and should be handled only with gloves and in a fume hood.*

Nucleoside mix, 50×

35 mg adenosine (Sigma)

35 mg guanosine (Sigma)

35 mg cytidine (Sigma)

35 mg uridine (Sigma)

35 mg thymidine (Sigma)

H₂O to 100 ml

Filter sterilize and store at -20° in 10-ml aliquots

COMMENTARY

Background Information

Gene amplification

When mammalian cells are placed in an environment that requires an increase in a normally constitutive gene product, cells that survive in many cases do so because of an increase in copy number of the gene. This process has been termed gene amplification and involves large regions of the chromosome, so that not only the selected gene becomes amplified, but the surrounding regions as well. Amplification was first detected when cells were treated with increasing concentrations of methotrexate, and the copy number of the resident DHFR gene was analyzed in the surviving cells (Alt et al., 1978). This observation, combined with the observation that cotransfected segments of DNA tend to integrate in the same chromosomal location (Wigler et al., 1978), has resulted in the ability to amplify any desired gene. Notable among the many reports of amplified genes in mammalian cells are *E. coli* XGPRT (Ringold et al., 1981), hepatitis B surface antigen (Christman et al., 1982), mouse *c-myc* (Wurm et al., 1986), tissue inhibitor of metalloproteinases using glutamine synthetase (GS) selection (Cockett et al., 1990), CD4 T lymphocyte glycoprotein (Davis et al., 1990), and human initiation factor 2 α using DHFR selection (eIF-2 α ; Kaufman et al., 1991).

Until recently, gene amplification was achieved by cotransfection of a vector carrying a selectable marker gene with another vector carrying the nonselectable gene of interest. In this process, called cotransformation, separate DNA molecules become ligated and cointegrate as a unit by nonhomologous recombination into the host chromosome (Wigler et al., 1978). Amplification of the region of DNA containing the selectable gene and the target gene is accomplished by incubating the cells in increasing amounts of a specific inhibitory drug. However, varying DNA transfection methods and cell lines can yield dramatically

different frequencies of cotransformation with different plasmids. Consequently, expression vectors have been developed that contain both the selectable gene and a transcription unit for the inserted target gene linked on the same plasmid. Various approaches in constructing these plasmids have been used to ensure adequate expression of both the selectable and nonselectable marker.

CHO cells

Many amplifiable selection markers are now available for use in mammalian cells (see Table 16.23.1; Kaufman, 1989, 1990). Although various cell lines (including monkey COS-1 and NIH 3T3) can be used for gene amplification, many protocols rely upon the use of Chinese hamster ovary (CHO) cells. The advantages of CHO cells for heterologous gene expression are (1) amplified genes that are integrated into host chromosome may be stably maintained, even in the absence of continued drug selection; (2) a variety of proteins have been expressed at high levels (Ringold et al., 1981; Cockett et al., 1990; Davis et al., 1990; Kaufman et al., 1991); and (3) volumes of CHO cells have been scaled up to >5000 liters. Two of the most successful strategies employing CHO cells—the dihydrofolate reductase (DHFR) and glutamine synthetase (GS) gene amplification systems—are described below.

Amplification using dihydrofolate reductase vectors

DHFR catalyzes the conversion of folate to tetrahydrofolate, which is required for purine, amino acid, and nucleoside biosynthesis. The folic acid analog methotrexate (MTX) binds and inhibits DHFR, causing cell death. Surviving populations of cells exposed to sequentially increasing concentrations of MTX contain increased levels of DHFR that result from gene amplification. The pED (DHFR) vector produces a transcript containing the target gene in the 5' position and the selectable marker in the

Table 16.23.1 Amplifiable Markers for Mammalian Cells^a

Selection	Gene
Adenosine, alanosine, and 2'-deoxycoformycin	Adenosine deaminase
Adenine, azaserine, and coformycin	Adenylate deaminase
β-aspartyl hydroxamate or albizzin	Asparagine synthetase
PALA	Aspartate transcarbamoylase
Methotrexate	Dihydrofolate reductase
Methionine sulfoximine	Glutamine synthetase
Cadmium	Metallothionein
α-difluoromethylornithine	Ornithine decarboxylase
Multiple drugs	P-glycoprotein 170
6-azauridine or pyrazofuran	UMP synthetase
Mycophenolic acid	Xanthine-guanine phosphoribosyltransferase

^aAdapted with permission from Academic Press.

3' position of the transcript. This vector has been optimized for translation of the selectable gene in the 3' position by use of specific sequences from the encephalomyocarditis (EMC) virus that promote internal ribosome binding and translation initiation (Kaufman et al., 1991). The development of CHO cell lines (e.g., CHO DG44) that are deleted for the endogenous DHFR genes greatly increases the ease of amplification using DHFR (Urlaub and Chasin, 1980; Urlaub et al., 1983). Amplification of exogenous genes has been accomplished either by using cells and vectors that express a normal DHFR gene (Ringold et al., 1981; Kaufman and Sharp, 1982a; Kaufman et al., 1991), or by using vectors that encode a DHFR gene partially resistant to methotrexate and normal CHO cells (Christman et al., 1982; Kaufman et al., 1991).

Amplification using glutamine synthetase vectors

GS provides the only pathway for synthesis of glutamine in mammalian cells (using glutamate and ammonia as substrates); thus, in a glutamine-free medium, GS is an essential enzyme. CHO cells contain endogenous GS enzyme, but concentrations of methionine sulfoximine (MSX) in excess of 20 to 25 μM are sufficient not only to inhibit wild-type levels of GS but also to prevent the growth of the majority of natural variants that arise by amplification of the endogenous GS genes. Hence, essentially all nontransfected cells are killed when grown in media containing MSX at these levels. The

GS vector pEE14 carries both the selectable GS marker and the nonselectable target gene, each transcribed from a separate promoter. The target gene is cloned into the multilinker cloning site with transcription initiating from the powerful human cytomegalovirus (hCMV-MIE) promoter-enhancer sequence, while transcription of the GS gene is driven from an SV40 late promoter (Cockett et al., 1989). The GS “minigene” in pEE14 permits sufficient GS expression to allow transformants to survive at low levels of MSX, and increased levels of MSX select cells that have undergone amplification of vector sequences integrated into the cell genome. The appropriate choice of MSX concentration for initial selection together with the use of a weakly expressed GS gene typically selects cell lines containing multiple copies of the vector. The combination of the hCMV promoter to drive the gene of interest and the GS-selection system usually provides relatively high levels of expression after only a single round of selection following gene amplification.

Critical Parameters

For this approach to be successful, the gene of interest must be integrated in a functional form in the original cell line. It is also crucial that the gene of interest not rearrange during amplification. These two criteria may be difficult to achieve if the expressed gene is cytotoxic when overproduced. This problem can be circumvented by expressing the gene of interest from an inducible promoter and thus amplify-

ing the gene in an "off" state (Wurm et al., 1986).

The key to obtaining high-level expression in primary transfectants is to screen as many as possible (typically ~100 lines) since the site of integration of the vector in the host cell genome has a profound influence on expression levels. To get good results on amplification, it is also important to screen a number of lines (e.g., 5 to 10 high producers), as again the integration site influences the frequency of amplification. Once high-producing lines have been isolated, it is often useful to reclone these to ensure a homogeneous population of cells, in order to increase the likelihood that the productivity will be maintained over long periods of culture. The selective agent should be present throughout these procedures and only when frozen stocks of recloned amplified cells have been secured may selection be relaxed, if it is important to choose lines that are stable without selective agent. Finally, it may again be necessary to screen a number of lines to identify ones that are stable without selective agent.

Anticipated Results

Dihydrofolate reductase. Cell lines containing >100 copies of an exogenous gene can be produced. The levels of mRNA and protein obtained depend upon the target gene to be expressed, but can constitute up to 5% of total protein synthesis (Kaufman, 1991).

Glutamine synthetase. Introduction of pEE14-based vectors using the CaPO₄-mediated transfection usually leads to multiple copies of the vector becoming integrated in the genome (up to 200 copies). The copy number can increase up to 30-fold in one round of selection for amplification (Cockett et al., 1989). The amount of product made depends both on the individual transfectant and on the protein being expressed, but can be up to 10 µg protein/10⁶ cells per 24-hr period from primary transfectants for some proteins. On amplification, protein expression can parallel the increase in copy number but is likely to reach a plateau, usually after a single round of amplification for many secreted proteins, probably because the secretion apparatus is saturated. Final yields of secreted proteins from overgrown cultures have been 180 mg/liter for tissue inhibitor of metalloproteinases (TIMP), and up to 120 mg/liter for secreted variants of the rat CD4 protein (Davis et al., 1990). pEE14 has also been used successfully to express an integral membrane protein (Harfst et al., 1992).

Time Considerations

DHFR. Transfection and detection of stably transformed colonies takes 2 weeks. It will take ~2 additional weeks to expand the colonies and analyze them to ensure that they contain the gene of interest. Amplification of the resultant stable lines may take up to 6 months by the standard protocol. If cells are cloned at every step, as described in the alternate protocol, the amplification process can take as little as 3 months.

GS. Transfection and analysis of initial cell lines takes ~4 weeks (as for DHFR selection). Selection for vector amplification and screening of resultant cell lines typically takes an additional 4 to 6 weeks; there is usually no advantage in undergoing further rounds of selection.

Literature Cited

- Alt, F.W., Kellems, R.E., Bertino, J.R., and Schimke, R.T. 1978. Selective multiplication of dihydrofolate reductase genes in methotrexate-resistant variants of cultured mammalian cells. *J. Biol. Chem.* 253:1357-1370.
- Bebbington, C.R. and Hentschel, C.C.G. 1987. The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells. *In DNA Cloning, Volume III* (D. Glover, ed.). Academic Press, San Diego.
- Christman, J.K., Gerber, M., Price, P.M., Flordellis, C., Edelman, J. and Acs, G. 1982. Amplification of expression of hepatitis B surface antigen in 3T3 cells cotransfected with a dominant-acting gene and cloned viral DNA. *Proc. Natl. Acad. Sci. U.S.A.* 79:1815-1819.
- Cockett, M.I., Bebbington, C.R., and Yarranton, G.T. 1990. High-level expression of tissue inhibitor of metalloproteinases in Chinese hamster ovary cells using glutamine synthetase gene amplification. *Bio/Technology* 8:662-667.
- Davis, S.J., Ward, H.A., Puklavec, M., Willis, A.C., Williams, A.F., and Barclay, A.N. 1990. High-level expression in Chinese hamster ovary cells of soluble forms of CD4 T lymphocyte glycoprotein including glycosylation variants. *J. Biol. Chem.* 265:10410-10418.
- Harfst, E., Johnstone, A.P., Gout, I., Taylor, A.H., Waterfield, M.D., and Nussey, S.S. 1992. The use of amplifiable high-expression vector pEE14 to study the interactions of autoantibodies with recombinant human thyro trophin receptor. *Mol. Cell Endocrinol.* 83:117-123.
- Hayward, B.E., Hussain, A., Wilson, R.H., Lyons, A., Woodcock, V., McIntosh, B., and Harris, T.J.R. 1986. The cloning and nucleotide sequence of cDNA for an amplified glutamine synthetase gene from the Chinese hamster. *Nucl. Acids Res.* 14:999-1008.
- Kaufman, R.J. 1989. Selection and coamplification. *Meth. Enzymol.* 185:537-566.

- Kaufman, R.J. 1990. Strategies for expressing high-level expression in mammalian cells. *Technique* 2:221-236.
- Kaufman, R.J. and Sharp, P.A. 1982a. Amplification and expression of sequences cotransfected with a modular dihydrofolate reductase complementary DNA gene. *J. Mol. Biol.* 159:601-621.
- Kaufman, R.J. and Sharp, P.A. 1982b. Construction of a modular dihydrofolate reductase cDNA gene: Analysis of signals utilized for efficient expression. *Mol. Cell. Biol.* 2:1304-1319.
- Kaufman, R., Davies, M., Wasley, L., and Michnik, D. 1991. Improved vectors for stable expression of foreign genes in mammalian cells by use of internal ribosomal entry site from EMC virus. *Nucl. Acids Res.* 19:4485-4490.
- Kingston, R.E., Kaufman, R.J., and Sharp, P.A. 1984. Regulation of transcription of the adenovirus E1 α promoter by E1 α gene products: Absence of sequence specificity. *Mol. Cell. Biol.* 4:1970-1977.
- Ringold, G., Dieckman, B., and Lee, F. 1981. Coexpression and amplification of dihydrofolate reductase cDNA and the *Escherichia coli* XGPRT gene in Chinese hamster ovary cells. *J. Mol. Appl. Genet.* 1:165-175.
- Sanders, P.G. and Wilson, R.H. 1984. Amplification and cloning of the Chinese hamster glutamine synthetase gene. *EMBO J.* 3:65-71.
- Stephens, P.E. and Cockett, M.I. 1989. The construction of a highly efficient and versatile set of mammalian expression vectors. *Nucl. Acids Res.* 17:7110.
- Urlaub, G. and Chasin, L.A. 1980. Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proc. Natl. Acad. Sci. U.S.A.* 77:4216-4220.
- Urlaub, G., Kas, E., Carothers, A.M., and Chasin, L.A. 1983. Deletion of the diploid dihydrofolate locus from cultured mammalian cells. *Cell* 33:405-412.
- Wigler, M., Silverstein, S., Lee, L., Pellicer, A., Cheng, Y., and Axel, R. 1977. Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. *Cell* 11:223-232.
- Wurm, F.W., Gwinn, K.A., and Kingston, R.E. 1986. Inducible overproduction of the mouse *c-myc* gene in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 83:5414-5418.

Key References

Kaufman et al., 1991. See above.

Describes the construction and application of dicistronic DHFR vectors that allow stable, high-level expression of inserted cDNAs by selection for methotrexate resistance in both DHFR-containing and DHFR-deficient cells.

Cockett et al., 1990. See above.

Describes the construction of vectors that provide for high-level expression using the GS gene-amplification system.

Contributed by Robert E. Kingston
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

Randal J. Kaufman (DHFR)
University of Michigan
Ann Arbor, Michigan

C.R. Bebbington and M.R. Rolfe (GS)
Celltech Ltd.
Slough, United Kingdom

Helper-Dependent Adenoviral Vectors

UNIT 16.24

Helper-dependent adenoviral vectors (HDAds) have improved the safety profile and greatly prolonged transgene expression compared with early-generation adenoviral vectors (Ads). HDAds contain only Ad *cis* elements necessary for vector genome replication and packaging into the capsid. Other factors are supplied in *trans* by a helper virus. In its preparation, it is important to ensure that the HDAd is free of contamination by helper virus, a first-generation Ad.

A summary of the HDAd strategy is shown in Figure 16.24.1. The first step is to subclone the expression cassette of the gene of interest into a shuttle vector (e.g., p Δ 21, p Δ 28, pC4HSU), which contains stuffer DNA in order to keep vector genome within a size that allows for efficient packaging (27 to 38 kb; Sandig et al., 2000; Kim et al., 2001; Oka et al., 2001). The second step is to rescue the HDAd by transfecting a linearized shuttle vector into a packaging cell line expressing a sequence-specific recombinase (i.e., Cre or FLP). Infection with a helper virus follows, to supply the necessary components for HDAd genome packaging into the capsid. The most commonly used protocols depend on the popular Cre-*loxP* recombinase system to minimize packaging of the helper. The third step is serial vector amplification. The final step is the large-scale vector production followed by purification.

Basic Protocol 1 provides detailed instructions for generating recombinant HDAds using the Cre-*loxP* system. Basic Protocol 2 describes large-scale production and purification of HDAd. Support Protocol 1 describes monitoring of vector amplification by Southern blot analysis. Support Protocol 2 provides a DNA-based method for determining HDAd titer, as HDAds do not form plaques upon infection of 293 cells. Helper virus contamination severely compromises the advantages of HDAds, and can be monitored by a real-time PCR-based method described in Support Protocol 3.

The components of a standard Cre-*loxP* based HDAd production system include: HDAd shuttle vector, pC4HSU, a helper virus H14, 293Cre4 packaging cell line. These components, and 293 and 293N3S cells are available for academic researchers through Microbix (<http://www.microbix.com/>). The 293 cells can also be obtained from ATCC.

293, 293N3S, PERC6, or N52.E6 cells are used only for helper virus amplification or plaque purification. They are not used to produce HDAd vectors. 293Cre66 cells can be used throughout HDAd vector production (transfection, serial passage, and large-scale production). 293Cre66S (suspension) cells are used only in large-scale production. 116 cells can be used throughout vector production.

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper aseptic technique should be used accordingly.

NOTE: All incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified. Some media (e.g., DMEM) may require altered levels of CO₂ to maintain pH 7.4.

BIOSAFETY

According to the National Institutes of Health biosafety guidelines based on risk assessment, manipulations on human adenoviruses should be performed in a laboratory operating at Biosafety Level 2 (BL-2) as approved by the user's Institutional Biosafety

Protein
Expression

16.24.1

Contributed by Kazuhiro Oka and Lawrence Chan

Current Protocols in Molecular Biology (2005) 16.24.1-16.24.23

Copyright © 2005 by John Wiley & Sons, Inc.

Supplement 69

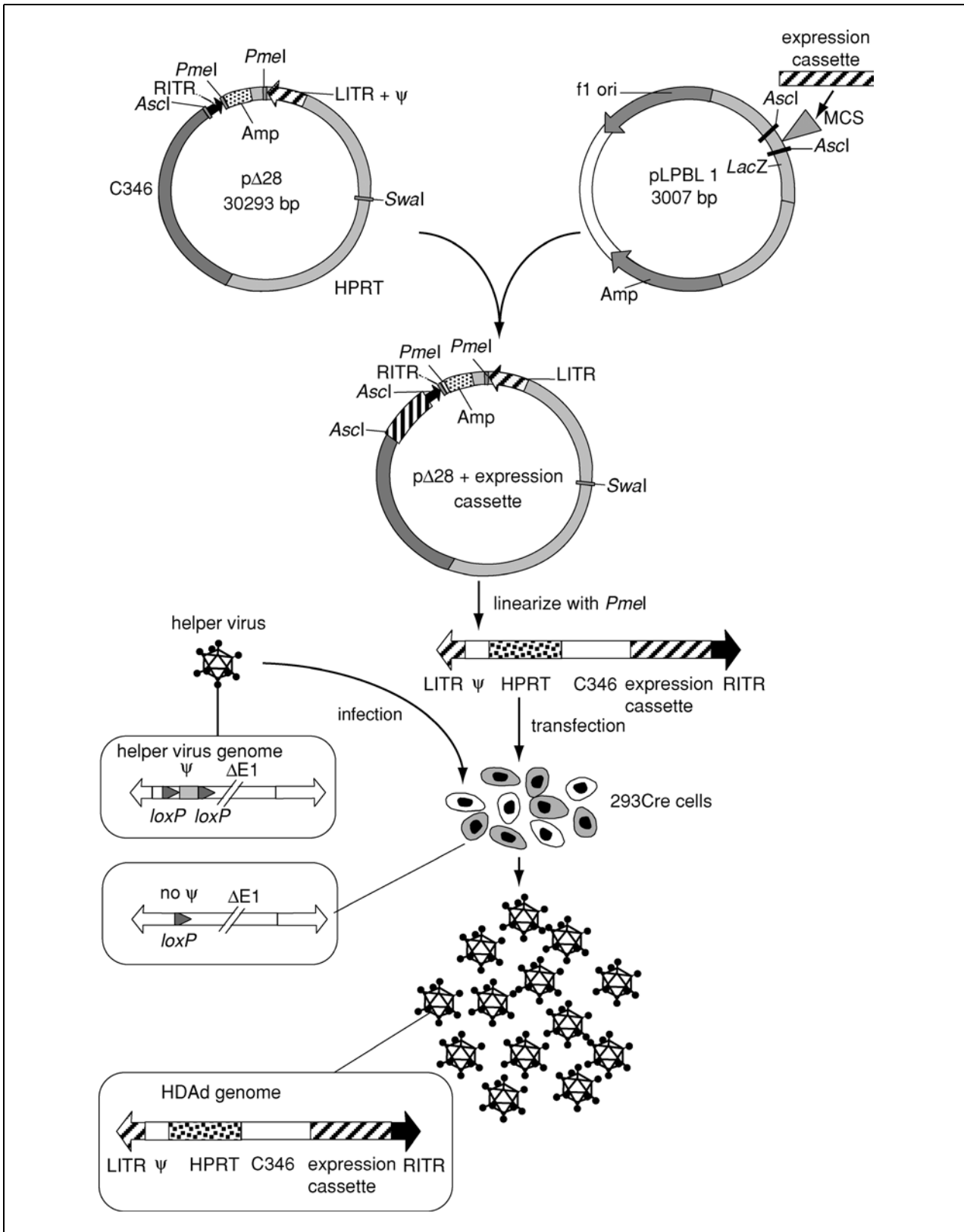


Figure 16.24.1 Legend at right.

Helper-Dependent Adenoviral Vectors

16.24.2

Committee. The requirements include the use of laminar flow hoods, the establishment of proper procedures for decontamination and disposal of liquid and solid waste, and the disinfection of contaminated surfaces and equipment.

STRATEGIC PLANNING

Vector Design

The size of the transgene insert of HDAd must be between 27 and 38 kb for efficient packaging. Smaller inserts can be packaged as head-to-tail and tail-to-tail concatamers within this size limit (Parks and Graham, 1997). If necessary, stuffer DNA is inserted into the vector construct to bring the insert size to the packageable range (Fig. 16.24.1). The nature of stuffer DNA influences efficiency of transgene expression. Fully characterized DNA of mammalian origin that does not contain any genes or repetitive sequences should be used, in order to avoid rearrangement during serial passages and to prevent any possible effects on transgene expression. The presence of matrix attachment sites in the HDAd enhanced transgene expression in one study (Schiedner et al., 2002), but not in another study (Ehrhardt and Kay, 2002). Since HDAd contain no viral coding sequence, transgene expression is largely independent of the orientation and only marginally influenced by the location of the transgene cassette within the vector genome (Schiedner et al., 2002). HDAd containing natural genes appear to have more sustained transgene expression than those containing cDNAs driven by heterologous promoters (Kim et al., 2001; Schiedner et al., 2002). However, long-term transgene expression (>1 year) has been reported in constructs using a heterologous promoter (Soudais et al., 2003; Nomura et al., 2004). Factors that regulate duration of expression are complex and poorly understood. Target-tissue-specific promoters are preferable to ubiquitous promoters for sustained transgene expression, and the use of promoters active in antigen-presenting cells should be avoided to minimize the generation of antibodies against transgene products. The inclusion of the first 400 bp from the right end containing the E4 promoter without any coding sequence into an HDAd vector has been reported to increase vector production efficiency either by enhancing replication or packaging of the vector DNA (Sandig et al., 2000). The authors of this unit have found that inclusion of the element gives inconsistent results, though its presence does not negatively impact viral production, and they usually include it in the standard vector design. Another popular element to increase transgene expression is the

Figure 16.24.1 (at left) Strategy for the construction of helper-dependent adenoviral vector (HDAd) using a helper virus. A simple strategy starts with the subcloning of expression cassette of a gene of interest into an intermediate cloning vector (e.g., pLPBL1), which contains the multiple cloning site (MSC) flanked by *Ascl* sites (Oka et al., 2001). The expression cassette should contain either promoter, cDNA, and polyadenylation signal, or the entire gene. The second step is to subclone the resulting expression cassette into HDAd shuttle vector (e.g., p Δ 28, p Δ 21, and pC4HSU; Sandig et al., 2000; Kim et al., 2001; Oka et al., 2001). It is important to maintain the size of HDAd shuttle vector between 27 and 38 kb (excluding the 3-kb plasmid backbone) for efficient packaging. In addition, the nature of stuffer DNA influences the degree and duration of transgene expression, as well as vector rearrangement during serial passages (Parks and Graham, 1997; Sandig et al., 2000). The HDAd shuttle vector containing the expression cassette is linearized with *PmeI* to release the plasmid backbone, and transfected into 293 or an equivalent E1-complementing cell line expressing Cre recombinase (293 Cre cells). After transfection, 293Cre cells are infected with helper virus, a first-generation E1-deleted Ad vector containing the packaging signal flanked by *loxP* sequences. Infection of 293Cre cells with the helper virus results in excision of its packaging signal. Although the helper virus genome continues to provide all of the components necessary for replication and packaging, the helper virus genome cannot be packaged. Instead, the genome of HDAd vector containing essential *cis*-acting elements (viral packaging signal and inverted terminal repeats) introduced into cells by transfection is replicated and packaged into Ad particles. Abbreviations: RITR, right inverted terminal repeat; Ψ , packaging signal; LITR, left inverted terminal repeat; HPRT, intron region of human genomic HPRT stuffer sequence; C346, cosmid C346 human genomic stuffer sequence.

woodchuck hepatitis virus post-transcriptional regulatory element (WPRE), a *cis*-acting element that increases the amount of unspliced RNA in both the nuclear and cytoplasmic compartments (Donello et al., 1998). The insertion of WPRE in the 3' untranslated region upstream to the polyadenylation signal is a useful strategy to increase transgene expression without affecting promoter specificity. However, the authors of this unit have obtained inconsistent results with this strategy.

Cloning Genes of Interest into Shuttle Vectors

The transgene construct containing the appropriate promoter is cloned into an HDAd shuttle vector. HDAd shuttle vectors contain several unique cloning sites. A simple method to clone the gene into a shuttle vector is by using an additional cloning vector (e.g., pLPBL-1; Oka et al., 2001) in which the expression cassette is constructed and excised using compatible unique restriction enzyme sites (e.g., *AscI* or *SwaI* for p Δ 28; *AscI* for p Δ 21; and *AscI*, *SwaI*, or *NotI* for pC4HSU). In this approach, the expression cassette should not have an internal *AscI*, *SwaI*, *NotI* or *PmeI* site. *PmeI* is used to linearize the shuttle vector prior to transfection.

If one cannot apply such an approach, the cloning can be done by employing partial digestions, blunt-end ligation, or recombination in *E. coli* (Toietta et al., 2002). HDAds can be rescued without linearizing the shuttle vectors. However, the efficiency is very low. Although the plasmid shuttle vector construct is relatively large (30 to 41 kb), subcloning is usually straightforward in experienced molecular biology laboratories. The preferred bacterial strain is DH5 α . Some constructs may require incubation at 30°C to increase DNA stability during vector propagation.

GENERATION OF RECOMBINANT HELPER-DEPENDENT ADENOVIRAL VECTORS USING Cre-*loxP* SYSTEM

To generate recombinant HDAd vectors, linearized shuttle vector DNA is transfected into a Cre-expressing packaging cell line followed by infection with a helper virus. It is important to liberate both ends of the shuttle vector for efficient vector rescue.

The following basic protocol is based on the use of 293Cre66 cells (Schiedner et al., 2002), AdLC8cluc helper virus, and pC4HSU shuttle vector. It is also generally applicable to other cell lines with appropriate modifications.

293Cre66 cells can be transfected by one of two methods: calcium phosphate precipitation or liposome-mediated transfection. The calcium phosphate precipitation method is slightly more effective than liposome-mediated transfection; however, the latter method requires less DNA and appears to be more reproducible. Only the latter method is described below. It is recommended that a shuttle vector containing a reporter gene be included for the first time to easily monitor vector amplification and to ensure a properly functioning system.

A schematic diagram of this protocol is shown in Figure 16.24.2.

Materials

Expression cassette: gene of interest with promoter active in the target tissue and poly(A) signal

Restriction endonucleases *AscI* and *PmeI*

Calf intestinal phosphatase (UNITS 3.10 & 3.16)

Shuttle vector: pC4HSU (available from Microbix), p Δ 21, p Δ 28, pSTK129 or other similar shuttle vectors, with stuffer DNA of mammalian DNA origin; Sandig et al., 2000; Kim et al., 2001; Oka et al., 2001; Schiedner et al., 2002)

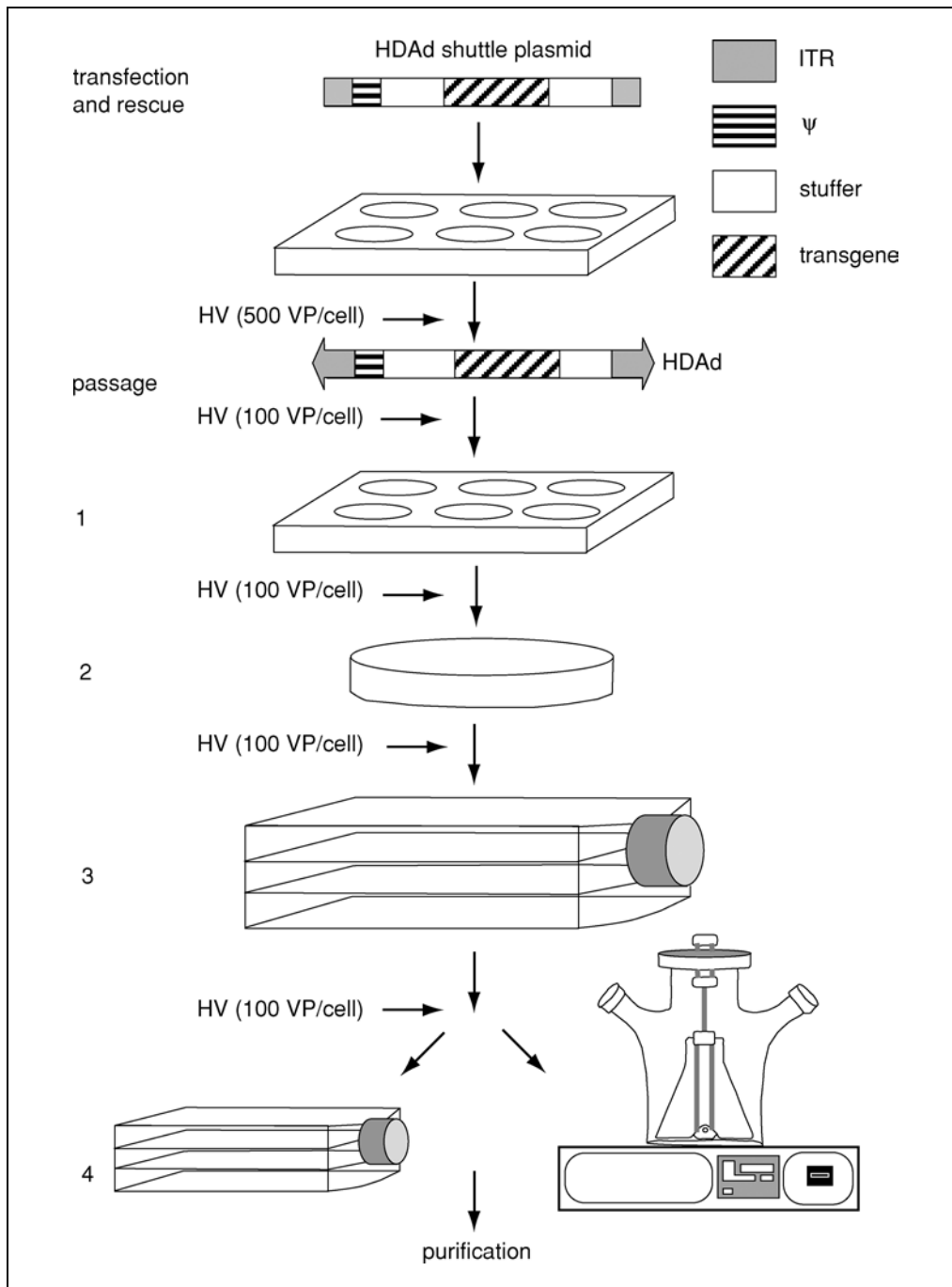


Figure 16.24.2 Schematic overview of helper-dependent adenoviral vector (HDAd) production using Cre-loxP system. The packaging cell line expressing Cre (293Cre66) is transfected with a linearized HDAd shuttle vector (e.g., pC4HSU containing a transgene cassette) and then infected with a helper virus (HV) at 500 vector particles (VP)/cell. Abbreviations: ITR, inverted terminal repeat; ψ , packaging signal. Since the packaging signal of a helper virus is flanked by loxP sequences, it is excised in 293Cre66 cells leaving a helper virus unpackageable. Only HDAd vector is packaged. The titer of HDAd vector is increased by serial passages. The 293Cre66 cells in a 6-well plate are coinfecting with HDAd vector and a helper virus (100 VP/cell) in passage 1. Coinfection is repeated in a 10-cm dish for passage 2, and in a TripleFlask (TF) for passage 3. The vector titer is determined before large-scale production. There are two ways to produce HDAd vectors on a large-scale (see Basic Protocol 2): adherent cell culture system using multiple TFs and suspension cell culture system. In the TF system, typically twelve TFs are infected. For suspension system, 3 liters of suspension cells are infected. After passage 4, HDAd vector is purified by CsCl density gradient centrifugation. The typical yield of HDAd vector is $3\text{--}5 \times 10^{12}$ VP by the adherent cell protocol and $5\text{--}10 \times 10^{12}$ VP by the suspension system. If necessary, the purified HDAd vector is used as a seed stock to produce more vectors.

Subcloning-efficiency competent *E. coli* DH5 α cells (Invitrogen) or other cells not prone to recombination, with transformation efficiency of $\sim 1 \times 10^6$ transformants/ μg DNA/50 μl of competent cells in 10-cm plate

10-cm LB/ampicillin/IPTG/Xgal plates (see recipe)

LB/ampicillin liquid medium: LB medium (APPENDIX 2) containing 100 $\mu\text{g}/\text{ml}$ ampicillin

LB medium (APPENDIX 2)

Plasmid miniprep kit (optional; available from many major molecular biology suppliers)

Plasmid midiprep kit (preferably with ion-exchange column using gravitational flow; available from many major molecular biology suppliers)

1% agarose gel (UNIT 2.5A)

25:24:1 phenol/chloroform/isoamyl alcohol (UNIT 2.1A)

3 M sodium acetate, pH 5.2 (APPENDIX 2)

70% and 100% ethanol

Transfection-grade H₂O

E1-function-complementing cell lines expressing Cre recombinase, e.g., 293Cre4 (Parks et al., 1996; available from Microbix), 293Cre66 and suspension 293Cre66S (Schiedner et al., manuscript in preparation), or 116 (293N3S based suspension cell line; Palmer and Ng, 2003)

E1-function-complementing cell lines (for amplification of helper virus) such as 293 (ATCC, Microbix), 293N3S (suspension cells, Microbix), PERC6, or N52.E6

Complete medium: α -MEM supplemented with 10% heat-inactivated FBS and 1 \times antibiotic-antimycotic (ingredients available from Invitrogen)

Minimal essential medium alpha (α -MEM; Invitrogen) without FBS or antibiotic-antimycotic

0.25% trypsin-EDTA (Invitrogen)

SuperFect transfection reagent (Qiagen)

Helper virus (see recipe)

Maintenance medium for adherent cells: α -MEM supplemented with 5% heat inactivated horse serum and 1 \times antibiotic-antimycotic (ingredients available from Invitrogen)

40% (w/v) sucrose (tissue culture grade) in H₂O, autoclave and filter through 0.22- μm filter; store at room temperature

15-ml and 50-ml conical tubes

37°C orbital shaker

6-well tissue culture plates

10-cm tissue culture plates

Cryotubes

TripleFlask (TF, Nunc) or 15-cm tissue culture dishes

Additional reagents and equipment for restriction endonuclease digestion (UNIT 3.1) and subcloning of DNA (UNIT 3.16), transformation (UNIT 1.8) and growth (UNITS 1.2 & 1.3) of *E. coli*, alkaline lysis miniprep of plasmid DNA (UNIT 1.6), agarose gel electrophoresis (UNIT 2.5A), extraction and precipitation of DNA (UNIT 2.1A), mammalian cell culture (APPENDIX 3F), and Southern blotting (UNIT 2.9A)

Clone gene into HDAd shuttle vector

1. Linearize shuttle vector using restriction endonuclease *Asc*I (UNIT 3.1) and treat with calf intestinal phosphatase (UNIT 3.16), then subclone expression cassettes of interest directly into the shuttle vector (UNIT 3.16), taking into consideration the issues described above (see Strategic Planning). Perform the ligation in a total volume of 10 μl , using 100 to 200 ng of calf intestinal phosphatase-treated shuttle vector.

Confirm the final plasmid construct by diagnostic restriction endonuclease digestions (see Strategic Planning) before transfection.

2. Transform $>2 \mu\text{l}$ of ligation mixture into $50 \mu\text{l}$ of competent *E. coli* cells (UNIT 1.8) with heat shock at 42°C for 30 sec. Spread all of the competent cells directly (without prior amplification in LB medium) onto a 10-cm LB/ampicillin/IPTG/Xgal plate and incubate 15 to 18 hr.
3. Pick smaller white bacterial colonies (four to six “candidate clones”), streak on a 10-cm LB/ampicillin/IPTG/X-gal master plate, and then inoculate them into 4 ml of LB/ampicillin liquid medium in separate 15-ml conical tubes.

Colonies are individually (i.e., separately) streaked in different grids on a master plate first, and then inoculated into LB medium.

4. Incubate the master plate in a 37°C bacterial incubator, and at the same time grow the corresponding candidate clones that were inoculated into the LB/ampicillin liquid medium overnight (15 to 18 hr) at 37°C with agitation on an orbital shaker.

Some constructs require 30°C incubation to avoid vector recombination.

5. Collect the cells from the candidate clones growing in the liquid medium and purify plasmid DNA using a commercial miniprep kit or the alkaline lysis procedure (UNIT 1.6).

The version of miniprep kit using a spin column (Qiagen) is most useful. The amount of DNA obtained from the miniprep may be low due to the size of the plasmid, but should be sufficient for screening.

6. Digest the DNA with diagnostic restriction enzyme *AscI* (UNIT 3.1) and perform agarose gel electrophoresis (UNIT 2.5A).

*The goal of diagnostic enzyme digestion is to confirm the presence and orientation of the expression cassette in an HDAd shuttle vector. *AscI* digestion is used to verify recombinant plasmid with two DNA bands corresponding to the expression cassette insert and the shuttle vector backbone. Additional digestion with an appropriate diagnostic enzyme that cuts into both the expression cassette insert and a site in the shuttle vector is needed to determine the orientation of the insert. It is also important to completely map the final construct to ensure that only a single copy of the cassette is included in the vector.*

7. Pick one correct clone from the master plate (i.e., one whose corresponding cells grown in the liquid medium have shown the correct structure) and grow in 100 ml LB/ampicillin overnight at 37°C with agitation on an orbital shaker.

8. Purify DNA using a commercial midiprep kit.

*A yield of 50 to 100 μg of DNA can be obtained from 100 ml culture. If the yield is poor, it is recommended that the plasmid be retransformed into *E. coli* DH5 α and that a well isolated single colony be picked directly into the 100-ml (or larger volume) culture and grown overnight.*

9. Digest 20 to 40 μg of DNA with *PmeI* (UNIT 3.1) to release the plasmid backbone in a total volume of 100 μl .
10. Confirm the digestion by running 1 to 2 μl of reaction mixture on a 1% agarose minigel (UNIT 2.5A).
11. Extract DNA (UNIT 2.1A) once with 100 μl of 25:24:1 phenol/chloroform/isoamyl alcohol. Precipitate DNA in a 1.5-ml microcentrifuge tube with 10 μl of 3 M sodium acetate, pH 5.2, and 220 μl of 100% ethanol.

- Microcentrifuge 3 min at 12,000 rpm and remove the supernatant. Rinse the DNA pellet with 500 μ l of 70% ethanol, microcentrifuge again for 3 min at 12,000 rpm, then remove the supernatant under the tissue culture hood.

DNA should be kept sterile from this step forward.

- Briefly air dry DNA and resuspend DNA in 20 μ l of transfection-grade water.

Transfect and rescue helper-dependent adenoviral vector

- Trypsinize and seed 293Cre66 cells (or other E1-function-complementing cells expressing Cre recombinase) with complete medium into a 6-well plate to reach 50% to 60% confluency on the day of transfection.

APPENDIX 3F provides basic cell culture protocols. 293Cre66 cells should not be rinsed with PBS before adding trypsin since this cell line is loosely attached to the flask and tends to become detached by doing so.

To obtain the level of confluency specified above, ~80% confluent 293Cre66 cells are split 1:4 (based on surface area). It then takes ~2 days to reach 50% to 60% confluency.

- For each well to be transfected, dissolve 2.5 μ g of the *PmeI*-linearized shuttle vector (from step 13) in 75 μ l plain α -MEM (no FBS or other additives).

If possible, it is recommended that the transfection be performed in duplicate. Proceed to passage 3, below, in duplicate, then pick the CVL (crude viral lysate) containing the higher titer for subsequent large-scale vector production.

- Add 12.5 μ l per well to be transfected of SuperFect (liposome-based) transfection reagent to the DNA solution and mix gently by pipetting up and down several times. Incubate for 5 to 10 min at room temperature. During incubation, wash the cells once with 2 ml of plain α -MEM (no additives) prewarmed to 37°C.

- Add 0.25 ml per well to be transfected of plain α -MEM to DNA solution, mix by pipetting up and down. Remove medium from cells. Overlay 337.5 μ l of the DNA/liposome/ α -MEM mixture per well, on top of cells.

- Incubate cells for 2 to 3 hr, swirling every 30 min, then add 1 ml of complete medium.

- The next day, wash the cells once with 2 ml of plain α -MEM. Remove medium and infect with 0.25 ml of plain α -MEM containing 5×10^8 vector particles (VP) of helper virus (500 VP/cell).

The amount of helper virus is based on the ratio of infectious particles to vector particles being between 1:20 and 1:100. The exact amount of helper virus should be optimized for each helper virus preparation based on the infectious titer. However, the amount described in the protocol is applicable in most cases without modification.

- Incubate cells for 1 hr, swirling every 10 to 15 min. Add 1 ml of maintenance medium for adherent cells.

- Incubate until >90% cells show cytopathic effect (CPE; cells are rounded and detached from the plate), which should occur between 48 to 72 hr post-infection.

- Transfer cells and medium (crude viral lysate, CVL) to a cryotube. Add 0.1 vol of 40% sucrose.

The resulting CVL in 4% sucrose can be stored up to 1 year or more at -80°C.

Perform passage 1

- Seed 293Cre66 cells in a 6-well plate for 80% to 90% confluence on day of infection.

24. On day of infection, thaw CVL (from step 22) in a 37°C water bath. Remove medium and incubate cells with 0.5 ml of CVL together with helper virus at 100 VP/cell (1×10^8 VP). Continue infection for 1 hr, swirling every 15 min.
25. Add 1 ml maintenance medium for adherent cells. Incubate until >90% of cells show CPE.

If cells do not show CPE after 72 hr, helper virus is insufficient. Thus, more helper virus must be added.
26. Transfer CVL (see step 22) to a cryotube. Add 0.1 vol of 40% sucrose and store at -80°C.

Perform passage 2

27. Seed 293Cre66 cells in a 10-cm dish for 80% to 90% confluence on day of transfection.
28. On day of infection, thaw CVL (from step 26) in a 37°C water bath. Remove medium and incubate cells with 1 ml of CVL supplemented with 7×10^8 VP of helper virus. Incubate 1 hr.
29. Add 5 ml maintenance medium for adherent cells, then continue incubating until >90% of cells show CPE (usually 2 to 3 days).
30. Retain a 0.2-ml of CVL for vector DNA analysis (see Support Protocol 1). Transfer the remaining CVL to a 15-ml conical tube, add 0.1 vol of 40% sucrose, and store at -80°C.

Perform passage 3

31. Seed 293Cre66 cells in a Nunc TripleFlask or in three 15-cm tissue culture dishes. Thaw CVL (from step 30) in a 37°C water bath.
32. Dilute CVL with plain α -MEM to 10 ml. Add 5×10^9 VP of helper virus to CVL and incubate 1 hr to infect cells.
33. Add 20 ml maintenance medium for adherent cells. Incubate until >90% of cells show CPE (2 to 3 days).
34. Transfer 0.2 ml of CVL to a 1.5 ml tube for DNA extraction. Transfer the remaining CVL into a 50-ml tube. Add 0.1 vol of 40% sucrose and store at -80°C.
35. Before proceeding to large-scale vector production, analyze vector DNA for vector amplification as described in Support Protocol 1, since an appropriate vector titer is necessary for a large-scale infection. If no clear vector amplification is detected at passage 3, repeat one or two more serial passages in TripleFlasks using the steps described under passage 3, and analyze vector amplification again as described in Support Protocol 1.

With efficient vector amplification, the intensity of the band corresponding to the left arm of HDAd vector is comparable to that of helper virus lacking the packaging signal at passage 3 ($\sim 6.5 \times 10^{10}$ VP/ml). If no HDAd vector is detected after additional passages, it is likely due to inefficient transfection, and repeating the transfection is recommended. The authors have found that one or two more serial passages increase the titer without detectable vector rearrangement. However, unnecessary amplification should be avoided since it increases the chance of vector rearrangement. The efficiency of vector amplification can be influenced by vector structures. If the vector contains a transgene that is cytotoxic and that is expressed during vector production, it will be poorly amplified. If such a poor amplification is anticipated, use packaging cell lines in which the transgene promoter is silent, or consider the use of a regulatable expression system.

**LARGE-SCALE HELPER-DEPENDENT ADENOVIRAL VECTOR
PRODUCTION**

There are two methods for producing HDAd vectors at a large scale. The first method is vector amplification using multiple flasks of adherent cells, as described in Basic Protocol 1 for passage 3 (see Basic Protocol 1, steps 31 to 34). The alternate steps utilize a suspension culture system. Because of ease of handling and scale-up, this is the preferred method for large-scale Ad vector production. Each laboratory should choose the method that best suits its needs. For example, a suspension system is preferred for large animal experiments with a small number of vectors. However, adherent cell culture is more economical if small animal experiments are performed with many vectors.

The following protocol is based on the use of 293Cre66S, 293Cre66 cells adapted for suspension *in vitro* (G. Shiedner et al., manuscript in preparation). The growth medium can be either MEM- α or Joklik MEM supplemented with 10% FBS. This cell line can be maintained in suspension as well as in adherent culture. It grows as adherent after several passages in a culture dish. 293N3SCre8 cells are available free for academic institutions through Merck Research Laboratories. Alternatively, contact Frank Graham at graham@mcmaster.ca. For 293Cre66 and 293Cre66S, contact Stefan Kochenek at Stefan.kochenek@medizin.uni-ulm.de. For 116 cells, contact Philip Ng at png@bcm.tmc.edu.

The protocol includes cell disruption by deoxycholate and CsCl gradient centrifugation using the NVT rotor. The advantage of using the NVT rotor is that the running time is short and purification can be completed in a day.

Materials

- Adherent 293Cre66 or 293Cre4 cells
- Minimal essential medium alpha (α -MEM; Invitrogen)
- CVL from passage 3 or additional passage (see Basic Protocol 1)
- Helper virus (see recipe)
- Maintenance medium for adherent cells: α -MEM supplemented with 5% heat inactivated horse serum and 1 \times antibiotic-antimycotic (ingredients available from Invitrogen)
- 100 mM Tris-Cl, pH 8.0 (APPENDIX 2)
- 40% (w/v) sucrose (tissue culture grade) in H₂O: autoclave and filter through 0.22- μ m filter; store at room temperature
- Cre-expressing, E1-complementing cell lines adapted for suspension: e.g., 293N3SCre8, 293Cre66S (Schiedner et al., manuscript in preparation), or 116 cells
- Growth medium for suspension cells (see recipe)
- Maintenance medium for suspension cells (see recipe)
- 5% sodium deoxycholate, filter-sterilized
- 2 M MgCl₂
- DNase I solution (see recipe)
- RNase A solution (see recipe)
- CsCl density gradient solutions (density = 1.25, 1.35, and 1.41 g/ml; see recipe)
- PBS⁺⁺ : PBS (APPENDIX 2) containing 0.68 mM CaCl₂ and 0.5 mM MgCl₂
- Dialysis buffer (see recipe)
- TripleFlasks (Nunc)
- 250-ml centrifuge tubes (Corning)
- 15-cm culture dishes
- Low-speed centrifuge (e.g., Sorvall RC5C)

250- and 3000-ml spinner culture flasks and spinner culture base (Bellco or Corning)
Refrigerated centrifuge (e.g., Sorvall centrifuge with SS-34 rotor)
NVT 65 ultracentrifuge tubes (Beckman)
2- and 5-ml pipets
Beckman LE-80K ultracentrifuge (or equivalent ultracentrifuge with swinging-bucket rotor)
1- or 3-ml syringes and 21-G needles
Dialysis cassette (Pierce)
Additional reagents and equipment for mammalian cell culture and counting cells
(APPENDIX 3F)

Expand virus

For adherent cells

- 1a. Seed adherent 293Cre66 (see materials for Basic Protocol 1) cells in 12 Nunc Triple-Flasks. Incubate to 80% to 90% confluence (~3 days) in growth medium for adherent cells.

See APPENDIX 3F for basic cell culture techniques.

- 2a. Thaw the CVL in 50-ml tube from passage 3 or additional passage in a 37°C water bath. Bring volume to 36 ml with plain α -MEM in each 50-ml tube. Add helper virus (6×10^{10} VP), and mix by inverting the tube several times.
- 3a. Remove the medium from the TripleFlasks and add 7 ml of plain α -MEM to each TripleFlask. Add 3 ml of CVL containing a helper virus (from step 2a) to each TripleFlask. Continue incubating 1 hr with occasional mixing to infect the cells.
- 4a. Add 20 ml of maintenance medium for adherent cells (see materials for Basic Protocol 1) and continue the incubation. Check the cells every 12 to 24 hr and collect CVL in 250-ml centrifuge tubes when CPE is nearly complete (i.e., when 90% of cells show CPE, 48 to 72 hr post-infection).
- 5a. Centrifuge 10 min at $800 \times g$, 15°C. Discard the supernatant.
- 6a. Resuspend the cells with 100 mM Tris·Cl, pH 8.0, in a total volume of 20 ml. Add 0.1 vol of 40% sucrose, and proceed to purification (step 7). Alternatively, divide into two aliquots in 15-ml tubes, and store at -80°C.

For suspension cells

- 1b. Split suspension-adapted 293Cre66S cells from one ~80% confluent 162-cm² flask into four 15-cm dishes. Culture for 3 days in a 37°C incubator using growth medium for suspension cells.
- 2b. Transfer all cells into a 3000-ml spinner flask and make up the volume to 1000 ml with growth medium for suspension cells (day 0). Culture cells in a 37°C incubator with stirring at 60 rpm. Check the cell density (APPENDIX 3F) the next day (day 1).
Cell density should be $\sim 2 \times 10^5$ cells/ml.
- 3b. Add 1000 ml of growth medium on day 2. Add an additional 1000 ml of growth medium for suspension cells on day 3. Check the cell density (APPENDIX 3F) on day 4.
Cell density should be $\sim 3-4 \times 10^5$ cells/ml.
- 4b. Collect cells by centrifuging 5 min at $800 \times g$. Save 500 ml of the supernatant (conditioned medium). Resuspend cells with 50 ml of conditioned medium, transfer to a 250-ml culture flask and bring the final volume to 100 ml.

- 5b. Thaw the CVL in a 50-ml tube from passage 3 or additional passage in a 37°C water bath. Add helper virus (100 VP/cell × total number of cells from step 3b) to the CVL. Add the above CVL/helper virus mixture to the 250-ml culture flask containing cells. Incubate for 2 hr in a 37°C incubator while stirring at 60 rpm.
- 6b. Transfer into a 3000-ml flask and make up the total volume to 500 ml with the conditioned medium. Add 1500 ml of fresh maintenance medium for suspension cells and continue incubation. Harvest the cells on day 7 (3 days after infection) by centrifuging at 10 min at 800 × g. Store as described in step 6a.

Purify and characterize virus

7. If cells were frozen, thaw by incubation in a 37°C water bath. Add 2 ml of 5% sodium deoxycholate to the lysate, and incubate for 30 min at room temperature while mixing occasionally.

The solution becomes gelatinous.

This step calls for deoxycholate disruption of the cell membranes. Although this requires more hands-on time, the yield of HDAd released from the cells is higher than if several rounds of freezing/thawing (freezing in a dry ice/methanol bath and thawing at 37°C) are used to release the viruses from the cells.

8. Add 0.2 ml of 2 M MgCl₂, 0.15 ml of DNase I solution, and 0.15 ml of RNase A solution to the lysate. Mix well and incubate at 37°C for 1 hr while mixing occasionally.
9. Centrifuge 10 min at 5000 × g (6500 rpm in Sorvall SS-34 rotor), 4°C for 10 min. Recover the clear supernatant.

First ultracentrifugation

10. For each viral preparation to be purified, sterilize six NVT 65 ultracentrifuge tubes either by autoclaving, irradiation under UV light for several hours, or soaking in 100% ethanol followed by washing with sterile water and air drying in the hood (six tubes are required for purification of each viral preparation).
11. In each of the ultracentrifuge tubes, place 3 ml of low-density CsCl gradient solution (1.25 g/ml), then underlay 3 ml of high-density CsCl density gradient solution (1.41 g/ml) using a 2-ml pipet.
12. Carefully overlay 5 ml of clear supernatant collected at step 9 using a 5-ml pipet.
13. Fill the tubes to the neck (using a 2-ml or Pasteur pipet) with either the clear supernatant or PBS⁺⁺, and cap the tubes.

The tubes do not need to be weighed.

14. Ultracentrifuge 30 min at 200,000 × g (50,000 rpm in Beckman LE-80K), 20°C, with the brake on.

Any swinging-bucket rotors can be used, but may require different centrifugation time and/or rpm. Check for the appropriate speed and centrifugation time at <http://www.beckman.com/resourcecenter/labresources/centrifuges/rotorcalc.asp>.

15. Remove caps from the tubes. Collect the lower opalescent band with a 1- or 3-ml syringe and 21-G needle by side puncture (clean the area for needle puncture with 70% ethanol prior to collection).

It is useful to remove the empty capsid band (upper opalescent band) before collecting the infectious virus band, since the two bands are physically close together and are difficult to isolate one from the other.

Second centrifugation

16. Place 8 ml of 1.35 g/ml CsCl density gradient solution in each of a series of new NVT 65 ultracentrifuge tubes prepared as in step 10. Overlay the bands recovered from first centrifugation in the respective tubes, then fill in the tubes to the neck with 1.35 g/ml CsCl density gradient solution.
17. Ultracentrifuge as in step 14, except for 3.5 hr. Recover the opalescent (virus) band with a syringe as in step 15.
18. Transfer the virus into a dialysis cassette. Dialyze against 3 liters of DB without sucrose once for 1 hr at 4°C, and then against 3 liters of DB at 4°C overnight.
19. Characterize the vector.

Routine vector characterization includes, determination of infectious titer (see Support Protocol 2); detection of helper virus contamination (see Support Protocol 3); verification of vector structure (see Support Protocol 4) and determination of vector concentration (VP/ml) using optical density. To read OD, add 20 μ l virus or 20 μ l blank solution and 380 μ l TE buffer (APPENDIX 2)/0.1% SDS, vortex 30 sec, centrifuge 5 min, and measure OD₂₆₀. One OD unit contains $\sim 10^{12}$ viral particles/ml (particles:infectious particles = $\sim 20:1$). However, the OD₂₆₀ calculation is based on an estimate of viral DNA content and does not imply either competent viral packaging or transgene expression. Therefore, it is advisable to combine this estimate with the other approaches noted above.

Efficient HDAd vector amplification results in less than 1% helper virus contamination without purification. The approximate yield and helper virus contamination of HDAd vectors can be determined before CsCl purification by retaining a 0.2-ml aliquot of cell suspension at the time of harvest and extracting DNA for Southern blot analysis as described in Support Protocol 1. The typical yield of HDAd vector on pC4HSU backbone at passage 4 with adherent cells is $3\text{--}5 \times 10^{12}$ VP with $<0.1\%$ helper virus contamination, and the ratio of infectious particles to vector particles is 1:20. With suspension cells, the yield is $\sim 0.5\text{--}1 \times 10^{13}$ VP. For future vector amplification, the purified virus is used as a seed stock.

MONITORING VECTOR AMPLIFICATION

The efficiency of HDAd vector amplification is somewhat variable and often limiting. It is therefore important to monitor vector amplification during serial passages. Southern blot analysis is a dependable method for analyzing vector amplification. By this analysis, vector amplification can be monitored and at the same time DNA rearrangement can be detected. In general, significant vector amplification is detected after passage 2. Southern blotting also allows determination of Cre-mediated excision of the helper virus packaging signal.

Additional Materials (also see Basic Protocol 1)

Aliquots of CVL from serial passages of the virus (see Basic Protocol 1)

Proteinase K reaction mixture (see recipe)

TE buffer, pH 8.0 (APPENDIX 2)

*Hind*III restriction endonuclease (or other appropriate restriction enzyme; see UNIT 3.1)

QIAamp DNA purification kit (Qiagen)

6 \times gel loading buffer (see recipe)

1.2% agarose gel (UNIT 2.5A)

1 kbPlus DNA marker (Invitrogen)

400-bp fragment containing the left ITR plus the packaging signal (contact Dr. K. Oka at kazuhiro@bcm.tmc.edu) as probe for Southern blotting, labeled with ³²P using a random priming DNA labeling kit

Additional reagents and equipment for Southern blotting (UNIT 2.9A) and autoradiography or phosphor imaging (APPENDIX 3A)

SUPPORT PROTOCOL 1

Protein
Expression

16.24.13

1. Thaw the 0.2-ml aliquots of CVL collected from serial passages in a 37°C water bath.
2. Add 0.2 ml proteinase K reaction mixture to each aliquot. Incubate ≥ 1 hr at 56°C.
Not more than 0.2 ml of CVL should be used for extraction of DNA, as excess cellular debris causes poor phase separation for subsequent phenol/chloroform/isoamyl alcohol extraction.
3. Add 400 μ l of 25:24:1 phenol/chloroform/isoamyl alcohol and vortex vigorously.
4. Microcentrifuge 2 min at maximum speed. Transfer the supernatant to a new 1.5-ml microcentrifuge tube.
5. Precipitate DNA by adding 40 μ l of 3 M sodium acetate buffer, pH 5.2, and 880 μ l of 100% ethanol.
A DNA clump should be visible soon after adding ethanol due to the presence of chromosomal DNA. See UNIT 2.1A for additional details on DNA precipitation.
6. Incubate at room temperature for 10 to 20 min to prevent salt precipitation, then microcentrifuge at 3 min at 12,000 rpm. Discard supernatant.
7. Rinse DNA pellet with 500 μ l of 70% ethanol. Briefly air dry pellet, then reconstitute in 30 μ l TE buffer, pH 8.0.
It is important not to excessively dry DNA, since dried DNA is difficult to dissolve.
8. Purify helper virus DNA using QIAamp DNA purification kit. Linearize shuttle vector using restriction endonuclease *PmeI* as in Basic Protocol 1, step 1.
9. Digest 6 μ l of the vector DNA from step 7 with an appropriate enzyme (the authors routinely use *HindIII*; UNIT 3.1) in a total volume of 20 μ l for 1 hr at 37°C. In separate tubes, digest 0.1 μ g (corresponding to 2.58×10^9 VP) of purified helper virus DNA and 0.1 μ g of linearized shuttle vector with the same enzyme used to digest the vector DNA.
10. To each tube, add 4 μ l of 6 \times gel loading buffer, then vortex and microcentrifuge briefly to collect solution at the bottom of the tube. Load contents of tubes in separate lanes of a 1.2% agarose gel (24 μ l/well). Also load 4 μ l of 1kbPlus DNA marker diluted to 50 ng/ μ l in 1 \times loading buffer as a molecular size standard. Perform agarose gel electrophoresis (UNIT 2.5A).
11. Perform Southern blot analysis according to standard protocols (UNIT 2.9A), using [³²P-labeled] 400-bp fragment containing the left ITR plus the packaging signal as the probe. Perform a final wash with 0.1 \times SSC/0.1% SDS at 65°C.
12. Expose the membrane to either a film or a phosphor imager (APPENDIX 3A) to detect the radioactive bands.

The 400-bp fragment containing the left ITR plus the packaging signal is used as a probe, since it differentiates the left and right arms of the HDAd vector. The size for intact helper virus left arm is 3.3 kb; left arm of helper virus lacking the packaging signal, 3.1 kb; the right arm of helper virus, 1 kb; the left arm of pC4HSU based HDAd, 2.4 kb; the right arm of pC4HSU based HDAd, 0.4 kb. When the amplification of HDAd is efficient, the band intensity of HDAd vectors becomes more prominent than that of helper virus.

DETERMINATION OF INFECTIOUS TITER

The infectious titer of a first-generation Ad vector is determined by plaque assay, which relies on vector replication and subsequent viral gene expression in E1-complementing cells. HDAds do not contain any viral coding sequences, and therefore do not form plaques. Hence, it is necessary to perform an assay based on internalization of vector DNA. This assay (based on the method described by Kreppel et al., 2002) can be used to titer early-generation Ads as well.

Additional Materials (also see *Basic Protocol 1*)

293 cells (ATCC)

Helper-dependent adenoviral vector (HDAd; see *Basic Protocol 1*)

HDAd vector containing a reporter gene (e.g., *lacZ*, green fluorescence protein) whose infectious titer is determined independently based on the reporter gene expression on 293 cells

PBS⁺⁺:PBS (*APPENDIX 2*) containing 0.68 mM CaCl₂ and 0.5 mM MgCl₂

10× citric-saline (see recipe)

0.8 M NaOH

ITR or packaging signal (contact Dr. K. Oka at kazuhiro@bcm.tmc.edu) as probe for slot blotting, labeled with ³²P using a random priming DNA labeling kit

24-well tissue culture plates

Additional reagents and equipment for slot blotting (*UNIT 2.9B*) and phosphor imaging (*APPENDIX 3A*)

1. Seed 293 cells in a 24-well plate with complete medium for adherent cells and incubate until cells become 80% to 90% confluent.
2. Dilute HDAd vectors and reference HDAd vector containing a reporter gene to 5 × 10⁹ VP/ml with plain α-MEM.

The infectious titer of the reference vector is determined by counting plaques for each dilution, averaging the duplicate wells (foci expressing the reporter gene are identified and counted by microscopy). The average will determine the titer of adenoviral stock as expressed in blue-forming units per milliliter (bfu/ml) for lacZ reporter gene or green fluorescence-forming units per milliliter (gfu/ml) for GFP reporter gene.

3. Remove medium from each well and replace with 0.2 ml of plain α-MEM.
4. Add the following volumes of diluted HDAd vector to duplicate wells: 0, 2, 4, 8, 16, and 32 μl. Add the same volumes of the reference vector to separate duplicate wells. Incubate 1 hr to infect.
5. Add 0.3 ml of maintenance medium for adherent cells (see materials for *Basic Protocol 1*), and incubate the plate overnight.
6. Wash the cells twice, each time with 1 ml of PBS⁺⁺, then add 0.2 ml of 1× citric saline to each well. Incubate at room temperature for ~20 min or until cells start to round up and detach from the well.
7. Transfer the cell suspension to a 1.5 ml microcentrifuge tube by pipetting up and down. Add 0.2 ml of 0.8 M NaOH, vortex vigorously, and incubate at room temperature for 15 min.
8. Assemble the slot blot apparatus and load 200 μl of the above cell suspension to each slot. Proceed to slot-blot hybridization (*UNIT 2.9B*), using the ³²P-labeled ITR or packaging signal as a probe.

9. Quantify the radioactive bands with a phosphor imager. Calculate the infectious titer of HDAd vector is calculated from the standard curve obtained by HDAd containing the reporter gene.

An alternative method, based on competition of an HDAd vector with a reference Ad vector for infection of 293 cells, has been reported by Palmer and Ng (2004). This method is potentially simple and accurate; however, the authors of this unit have not established a working protocol.

REAL-TIME PCR TO DETECT HELPER VIRUS CONTAMINATION

Helper virus contamination can be determined by semiquantitative Southern blot analysis using known amounts of helper virus DNA as a standard. The radioactive bands corresponding to the left arm are quantified by a phosphor imager and the amount of helper virus present in HDAd vector is calculated from the standard. Real-time PCR is a more accurate way to quantify helper virus contamination. Although TaqMan PCR, originally developed for this purpose, may have greater precision, the following protocol using SYBR Green I as a detection dye also works quite well and is simple and economical. For this analysis, DNA should be purified by a QIAamp DNA kit or an equivalent method.

Additional Materials (also see Basic Protocol 1)

Helper-dependent adenoviral vectors (HDAds) to be analyzed (see Basic Protocol 1)

2× SYBR Green QPCR master mix (Stratagene, Bio-Rad, or equivalent)

10 pmol/μl PCR primers (see recipe)

96-well reaction plates

Real time QPCR platform (Stratagene Mx3000 or equivalent)

1. Purify HDAd vector to be analyzed using QIAamp DNA purification kit.
2. Linearize shuttle vector using restriction endonuclease *PmeI* (UNIT 3.1). Make seven serial dilutions of helper virus and *PmeI*-linearized pC4HSU shuttle vector DNA by 10-fold increments starting at 0.1 ng/μl and ending at 10⁻⁷ ng/μl, with water.
3. Make four serial dilutions of HDAd vectors to be analyzed by 10-fold increments starting at 0.1 ng/μl and ending at 10⁻⁴ ng/μl with water.
4. Prepare PCR-primer mix for helper virus. For 24 wells, mix:

25 × 26 μl (24 wells + 2 extra wells) of 2× SYBR Green QPCR master mix

26 μl of 10 pmol/μl primer 1

26 μl of 10 pmol/μl primer 2

13 × 26 μl of H₂O.

Mix gently by pipetting up and down. Minimize exposure to light since SYBR Green is light sensitive.

5. Prepare PCR-primer mix for HDAd vector by repeat the procedure in step 3 with primers 3 and 4 or primers 5 and 6.
6. Dispense 40 μl of PCR primer mix to duplicate wells of a 24-well reaction plate.
7. Add 10 μl of DNA solution (steps 1 and 2) to each well in duplicate. Add 10 μl of water to separate duplicate wells as a no-template control (NTC).

Therefore, two rows include NTC, standards (seven dilutions), and HDAd vector to be assayed (four dilutions) in duplicate.

8. Perform real-time PCR according to the manufacturer's instruction for the Stratagen Mx3000 using the following cycling programs.

1 cycle:	10 min 95°C	(initial denaturation)
40 cycles:	30 sec 95°C	(denaturation)
	30 sec 55°C	(annealing)
	60 sec 72°C	(extension).

9. Calculate helper virus contamination in percent:

$$\% \text{ contamination} = (\text{ng of helper virus}/\text{ng of HDAd vector})/(\text{ng of shuttle vector DNA}/\text{ng of HDAd vector}) \times 100 \times \text{size correction factor}$$

The size correction factor is equal to $[31 \times (\text{size of pC4HSU based HDAd vector in kb})]/[35 \times (\text{size of AdLC8cluc in kb})]$.

VERIFICATION OF VECTOR STRUCTURES

HDAdS are produced in serial passages by coinfection with helper virus. The existence of overlapping sequences between HDAd and helper virus could result in recombination between the two viruses. Such vector recombination can be detected by Southern blot analysis, in which the vector undergoing recombination appears as extra bands of varying sizes other than HDAd vector or helper virus. Therefore, it is important to routinely compare the structures of HDAd and shuttle vector plasmid by Southern blot analysis. The protocol for doing this is described in Support Protocol 1, except that in this case it utilizes a probe corresponding to the plasmid shuttle vector without the plasmid backbone.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Citric-saline, 10×

1.35 M KCl

0.15 M sodium citrate

Autoclave and store up to 3 months at room temperature

CsCl density gradient solutions

For each density gradient solution to be prepared, pour 350 ml of 1× Tris density buffer (see recipe for 10×) into a beaker. Add the following quantities of CsCl to prepare the gradient solutions of the respective density:

135 g CsCl for $d = 1.25$

160 g CsCl for $d = 1.33$

180 g CsCl for $d = 1.41$.

Mix with a stir bar, and equilibrate to room temperature. Place exactly 1 ml of the CsCl solution in a small weighing boat using a 1-ml pipet calibrated such that 1 ml of water is exactly equal to 1.00 g, obtain an accurate weight, and return the 1-ml aliquot to the beaker. If the density is higher than desired, add more 1× Tris density buffer to the respective beaker; if it is lower, add more CsCl powder. Repeat this process until the exact density concentration is reached. Sterilize the density gradient solutions using a 0.22- μm filter. Store up to 6 months at room temperature.

SUPPORT PROTOCOL 4

Protein
Expression

16.24.17

Dialysis buffer

10 mM Tris·Cl, pH 8.0 (*APPENDIX 2*)
2 mM MgCl₂
4% (w/v) sucrose
Store up to 2 weeks at 4°C

DNase I solution

10 mg/ml bovine pancreatic deoxyribonuclease I
20 mM Tris·Cl, pH 7.4 (*APPENDIX 2*)
50 mM NaCl
1 mM dithiothreitol
0.1 mg/ml BSA
50% (v/v) glycerol
Store up to 1 year in aliquots at –20°C

Gel loading buffer, 6×

1× TAE buffer (*APPENDIX 2*) containing:
0.25% (w/v) bromphenol blue
0.25% (w/v) xylene cyanol
30% (v/v) glycerol
Store up to 1 year at 4°C

Growth medium for suspension cells

Joklik MEM or S-MEM containing:
10% FBS
1% Pluronic F-68
1× penicillin/streptomycin
1× L-glutamine
80 µg/ml G418
Store up to 2 months at 4°C

The above ingredients are available from Invitrogen.

Helper virus

Helper virus with the packaging signal flanked by *loxP* sites—e.g., AdLC8cluc (Parks et al., 1996; available free to academic institutions through Merck Research Laboratories) or H14 (Sandig et al., 2000; available commercially from Microbix)—is required. Helper virus should be plaque purified prior to large-scale production on 293 cells or other E1-function-complementing cell lines and its structure verified before its use. This is essential since any variant (e.g., helper virus without *loxP* sequences) can outgrow HDAd during serial passages. Purified helper virus is stored at –80°C in small aliquots in 10 mM Tris·Cl, pH 8.0 (*APPENDIX 2*)/2 mM MgCl₂/4% (w/v) sucrose at concentrations of 1–3 × 10¹² viral particle (VP)/ml. Freezing/thawing several times does not affect the function of helper virus.

See Critical Parameters and Troubleshooting for more information.

LB/ampicillin/IPTG/Xgal plates

LB plates (*APPENDIX 2*) containing:
100 µg/ml ampicillin
0.4 mM isopropylthio-β-galactoside (IPTG)
0.004% (w/v) 5-bromo-4-chloro-3-indolyl-β-D-galactoside (Xgal)
Store up to 2 weeks at 4°C

Maintenance medium for suspension cells

Joklik MEM (Invitrogen) containing:
5% FBS
1% Pluronic F-68
1× penicillin/streptomycin
1× L-glutamine
Store up to 2 months at 4°C

PCR primers, 10 pmol/μl

The PCR primer sequences to detect helper virus and HDAd vector genome (see Support Protocol 3) are as follows. Dilute to 10 pmol/μl using sterile water.

Helper virus (AdLC8cluc):

Primer 1: 5′-CCTACACCAACACAACAACCTC-3′

Primer 2: 5′-ATCCACCTCAAAAGTCATGTC-3′

268-bp PCR product

HDAd on pC4HSU backbone:

Primer 3: 5′-CCATCAGCATAACATACGCAG-3′

Primer 4: 5′-CCAGTCAGTCACAGTCAAGAG-3′

260-bp PCR product

HDAd on pΔ21 or pΔ28 backbone:

Primer 5: 5′-CCCATCTCCTTCATCACATCTC-3′

Primer 6: 5′-TCACAGTTCCTCCAGCCTC-3′

273-bp PCR product

See UNIT 2.11 for oligonucleotide synthesis. HPLC-purified oligonucleotides are preferred, but not essential.

Proteinase K reaction mixture

Prepare a 10 mg/ml proteinase K stock solution and store in small aliquots up to 1 year at −20°C. Just before use, mix the following at room temperature to prepare the reaction mixture:

176 μl H₂O

20 μl 10% (w/v) SDS

4 μl 10 mg/ml proteinase K stock solution

RNase A solution

10 mg/ml RNase A

10 mM Tris·Cl, pH 7.4 (APPENDIX 2)

15 mM NaCl

Store up to 1 year in small aliquots at −20°C

Tris density buffer, 10×

80.0 g NaCl

3.8 g KCl

30.0 g Tris base

1.87 g Na₂HPO₄·7H₂O

Add H₂O to 1000 ml

Adjust pH to between 7.4 and 7.5 with 5 M NaOH

Filter sterilize and store up to 6 months at room temperature

COMMENTARY

Background Information

Early-generation adenoviral (Ad) vectors retain many viral genes. In vivo toxicity and immune response to transduced cells associated with these Ad vectors are caused in part by leaky viral-gene expression. These problems are partially solved in HDAds by removing all viral sequences from the Ad except the essential *cis* elements. This approach also increased the cloning capacity to the packaging limit (~38 kb); see review by Kochanek (1999). This most recently developed Ad vector has been referred to as “helper-dependent,” “high-capacity,” “gutless,” or “guttled” Ad vectors. All the necessary functions and proteins for replication and packaging are supplied in *trans* by a helper virus. The term “helper-dependent” reflects the current production method described herein. To take full advantage of this system, the helper virus, a first-generation Ad, must be separated from HDAd vectors. The most popular method utilizes the Cre-*loxP* (Parks et al., 1996) or FLP-*frt* (Ng et al., 2001; Umana et al., 2001) sequence-specific recombinase systems to remove the packaging signal of a helper virus during vector production. HDAd vectors have been reported to exhibit superior performance over early-generation Ads in many in vivo applications, due to their low toxicity and long-term transgene expression (Morral et al., 1999; Kim et al., 2001; Soudais et al., 2003; Nomura et al., 2004). Ad vectors are essentially a nonintegrating viral vector system. Transgenes delivered by Ad vectors are eventually eliminated from the host by cell division or cell turnover. HDAd vector is no exception. Readministration of the same vector is not effective because the Ad vectors are neutralized by antibodies against capsid proteins. However, neutralizing antibodies are serotype-specific, which permits the readministration of the same HDAd vector of different serotype (Parks et al., 1999; Kim et al., 2001). A major advantage of this strategy using HDAd vector is that HDAd vectors with various serotypes can be readily produced by coinfection of the same vector backbone and a helper virus of the alternate serotypes during the large-scale production phase of the vector (serotype switch). Another strategy to overcome the transient nature of transgene expression is to incorporate elements that permit stable integration of the vector into the host chromosome. Taking advantage of the cloning capacity and broad tropism of HDAd vectors

and the capability of DNA integration of other systems, many laboratories have developed HDAd hybrid viral vectors (Recchia et al., 1999, 2004; Goncalves et al., 2001; Soifer et al., 2001, 2002; Yant et al., 2002). In most of the hybrid HDAd vectors, the transgene is integrated at random sites, except for HDAd/AAV hybrid viruses that coexpress Rep78 gene (Recchia et al., 1999, 2004). However, many of these systems require coexpression of necessary components for vector integration, in addition to therapeutic genes. Such components may have adverse effects on the host, though the use of regulatable systems to express such proteins could minimize such effects. Optimization of such integrating systems could produce even more powerful HDAd variants in the future.

Critical Parameters and Troubleshooting

The critical requirements for a good HDAd preparation are a helper virus with the correct structure, a packaging cell line with high Cre expression, and shuttle vectors of appropriate structure. High-titer helper virus is prepared using plaque-purified helper virus. It is important to have enough helper virus stock for vector production; also, the structure of the helper virus must be fully characterized. Pick four to six well isolated plaques and propagate to obtain high-titer virus seed stock, which allows for viral DNA extraction and subsequent diagnostic enzyme digestions. Large-scale helper virus preparation is carried out using the fully characterized seed stock. Helper virus passages should not exceed 5 for working helper virus. Helper virus contamination in HDAd vector preparation is determined by the efficiency of Cre-mediated excision of the packaging signal of helper virus. The efficiency of excision is determined by Southern blot analysis, as described in Support Protocol 1. The band corresponding to the left arm of intact helper virus should not be detectable in CVL. Helper virus contamination <0.2% is desirable, though for some experiments, levels <1% may be acceptable. The overall structures of HDAd vectors influence the efficiency of vector amplification as well as propensity for vector DNA rearrangement. In the authors' experience, vectors 31 to 35 kb in size are efficiently amplified. Vectors containing genomic fragments tend to be amplified more efficiently, and they may also be more stable during vector production. Thus, it is desirable

Table 16.24.1 Troubleshooting Guide for Production of Helper-Dependent Adenoviral Vectors

Problems	Possible causes(s)	Solution(s)
Low number or no colonies after transformation in DH5 α	Transformation efficiency is too low	Use commercial Subcloning Efficiency DH5 α (Invitrogen)
	Ligation did not work.	Run 1-2 μ l of ligation mixture on a 1% agarose gel to make sure that the banding patterns are different before and after the ligation. If not, use more T4 DNA ligase.
Too many blue colonies after transformation	The DNA insert to be subcloned contains plasmid backbone.	Purify the insert twice and run on an agarose gel to make sure the band corresponding to the plasmid backbone is not detectable. Qiagen spin column gel extraction kit works well.
Too many colonies of the intact shuttle vector	Alkaline phosphatase reaction is insufficient	Treat linearized shuttle vector with 1 U of calf intestinal alkaline phosphatase for 30 min at 37°C, and extract with phenol/chloroform twice.
The structure of recombinant DNA does not match either the shuttle vector or the plasmid containing expression cassette	Most likely, recombination occurred during bacterial culture at 37°C	Incubate bacterial culture at 30°C
CPE occurs too early or too late	Amount of helper virus is inappropriate	Adjust helper virus to obtain 90% CPE 2-3 days after infection
No HDAd vector is detected by Southern blot analysis after passage 3	Transfection efficiency is low	Check the transfection efficiency using the reporter gene construct. One day after the transfection, at least 20%-30% cells should be positive for the reporter gene expression.
	HDAd vector is not amplified efficiently	Perform one or more passages in 6- or 10-cm dish to determine the presence of HDAd vector. If the vector is not detectable, repeat transfection.
	Helper virus does not provide the necessary functions	Perform comprehensive restriction analysis of the helper virus. If the structure differs, plaque purify the original helper virus.
HDAd vector amplification is poor	Transgene product is toxic and is expressed during vector production	Use a promoter not active in the packaging cell line, or use a regulatable system
	Vector genome size is close or out of the size of limit	Change the vector genome size for efficient packaging (31-35 kb)
In Southern blot analysis, extra band(s) in addition to the 4 expected bands from a helper virus and HDAd vector are detected	Recombination between the helper virus and the HDAd vector, or within the HDAd vector occurred	It is possible to obtain HDAd vector with the correct structure by repeating the transfection. The authors have obtained a few HDAd vectors with the correct structure after 30 transfections.
		Change the vector design or the shuttle vector. HDAd vector with pC4HSU-based backbone is less prone to recombination.

*continued***Protein
Expression****16.24.21**

Table 16.24.1 Troubleshooting Guide for Production of Helper-Dependent Adenoviral Vectors, *continued*

Problems	Possible causes(s)	Solution(s)
After infection of packaging cell line with a helper virus, the band corresponding to the left arm of an intact helper virus is detected by Southern blot analysis	Cre expression is not optimum for the amount of helper virus used	Use less helper virus or change to another Cre-expressing packaging cell line
Multiple bands other than empty capsids or helper virus after CsCl density gradient ultracentrifugation	Potential difference of HDAd vector genome conformation	Collect each band separately and characterize for DNA structures, and transgene expression. They are likely the same.
No transgene expression detected	The integrity of transgene expression cassette is not maintained	Make sure the transgene cassettes intact by restriction enzyme analysis
	Transgene is not efficiently expressed	Include a Kozak sequence at the translation initiation site, or WPRE after the termination codon but before the polyadenylation signal. Alternatively, use a more potent promoter.
	Infectious titer is too low	Determine infectious titer by slot blot analysis described in Support Protocol 2. Harvest virus within 3 days after infection and use sodium deoxycholate method to release virus during purification.

to use genomic DNA containing all the necessary elements for gene expression, if possible.

Table 16.24.1 lists potential problems associated with HDAd vector production and their possible causes and solutions.

Anticipated Results

HDAd technology is still in its infancy. Despite an increasing number of publications on the use of HDAd vectors, large-scale production of HDAd vectors free from helper virus contamination has been technically demanding. The protocols described herein work well for pC4HSU-derived HDAds. A key step may be transfection efficiency. If transfection efficiency is low, HDAd amplification is poor and more passages are required to achieve enough titer for large-scale vector production. This increases the chance of vector DNA rearrangement. The typical yield of HDAd vectors on pC4HSU backbone is 10,000 to 20,000 VP/cell with 293Cre66 adherent cells, and 3000 to 8000 VP/cell with 293Cre66S suspension cells.

Time Considerations

Each serial passage takes 2 to 3 days. If every step works well and there is enough titer after passage 3, one can expect to get the pu-

rified HDAd vector from the first large-scale vector production in 2 to 3 weeks. Cloning the desired expression cassette into a shuttle vector requires some experience with large plasmid vectors, but should be straightforward for experienced molecular biology laboratories. Therefore, the total time commitment needed for each experiment is determined largely by the efficiency of transfection and subsequent vector amplification.

Literature Cited

- Donello, J.E., Loeb, J.E., and Hope, T.J. 1998. Woodchuck hepatitis virus contains a tripartite posttranscriptional regulatory element. *J. Virol.* 72:5085-5092.
- Ehrhardt, A., and Kay, M.A. 2002. A new adenoviral helper-dependent vector results in long-term therapeutic levels of human coagulation factor IX at low doses in vivo. *Blood* 99:3923-3930.
- Goncalves, M.A., Pau, M.G., de Vries, A.A., and Valerio, D. 2001. Generation of a high-capacity hybrid vector: Packaging of recombinant adenoassociated virus replicative intermediates in adenovirus capsids overcomes the limited cloning capacity of adenoassociated virus vectors. *Virology* 288:236-246.
- Kim, I.H., Jozkowicz, A., Piedra, P.A., Oka, K., and Chan, L. 2001. Lifetime correction of genetic deficiency in mice with a single injection of helper-dependent adenoviral vector. *Proc. Natl. Acad. Sci. U.S.A.* 98:13282-13287.

- Kochanek, S. 1999. High-capacity adenoviral vectors for gene transfer and somatic gene therapy. *Hum. Gene Ther.* 10:2451-2459.
- Kreppel, F., Biermann, V., Kochanek, S., and Schiedner, G. 2002. A DNA-based method to assay total and infectious particle contents and helper virus contamination in high-capacity adenoviral vector preparations. *Hum. Gene Ther.* 13:1151-1156.
- Morral, N., O'Neal, W., Rice, K., Leland, M., Kaplan, J., Piedra, P.A., Zhou, H., Parks, R.J., Velji, R., Aguilar-Cordova, E., Wadsworth, S., Graham, F.L., Kochanek, S., Carey, K.D., and Beaudet, A.L. 1999. Administration of helper-dependent adenoviral vectors and sequential delivery of different vector serotype for long-term liver-directed gene transfer in baboons. *Proc. Natl. Acad. Sci. U.S.A.* 96:12816-12821.
- Ng, P., Beauchamp, C., Eveleigh, C., Parks, R., and Graham, F.L. 2001. Development of a FLP/rtt system for generating helper-dependent adenoviral vectors. *Mol. Ther.* 3:809-815.
- Nomura, S., Merched, A., Nour, E., Dieker, C., Oka, K., and Chan, L. 2004. Low-density lipoprotein receptor gene therapy using helper-dependent adenovirus produces long-term protection against atherosclerosis in a mouse model of familial hypercholesterolemia. *Gene Ther.* 11:1540-1548.
- Oka, K., Pastore, L., Kim, I.H., Merched, A., Nomura, S., Lee, H.J., Merched-Sauvage, M., Arden-Riley, C., Lee, B., Finegold, M., Beaudet, A., and Chan, L. 2001. Long-term stable correction of low-density lipoprotein receptor-deficient mice with a helper-dependent adenoviral vector expressing the very low-density lipoprotein receptor. *Circulation* 103:1274-1281.
- Palmer, D. and Ng, P. 2003. Improved system for large-scale production of helper-dependent adenoviral vectors. *Mol. Ther.* 7:S164.
- Palmer, D. J. and Ng, P. 2004. Physical and infectious titers of helper-dependent adenoviral vectors: A method of direct comparison to the adenovirus reference material. *Mol. Ther.* 10:792-798.
- Parks, R.J. and Graham, F.L. 1997. A helper-dependent system for adenovirus vector production helps define a lower limit for efficient DNA packaging. *J. Virol.* 71:3293-3298.
- Parks, R.J., Chen, L., Anton, M., Sankar, U., Rudnicki, M.A., and Graham, F.L. 1996. A helper-dependent adenovirus vector system: Removal of helper virus by Cre-mediated excision of the viral packaging signal. *Proc. Natl. Acad. Sci. U.S.A.* 93:13565-13570.
- Parks, R., Eveleigh, C., and Graham, F. 1999. Use of helper-dependent adenoviral vectors of alternative serotypes permits repeat vector administration. *Gene Ther.* 6:1565-1573.
- Recchia, A., Parks, R.J., Lamartina, S., Toniatti, C., Pieroni, L., Palombo, F., Ciliberto, G., Graham, F.L., Cortese, R., La Monica, N., and Colloca, S. 1999. Site-specific integration mediated by a hybrid adenovirus/adeno-associated virus vector. *Proc. Natl. Acad. Sci. U.S.A.* 96:2615-2620.
- Recchia, A., Perani, L., Sartori, D., Olgiati, C., and Mavilio, F. 2004. Site-specific integration of functional transgenes into the human genome by adeno/AAV hybrid vectors. *Mol. Ther.* 10:660-670.
- Sandig, V., Youil, R., Bett, A.J., Franlin, L.L., Oshima, M., Maione, D., Wang, F., Metzker, M.L., Savino, R., and Caskey, C.T. 2000. Optimization of the helper-dependent adenovirus system for production and potency in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 97:1002-1007.
- Schiedner, G., Hertel, S., Johnston, M., Biermann, V., Dries, V., and Kochanek, S. 2002. Variables affecting in vivo performance of high-capacity adenovirus vectors. *J. Virol.* 76:1600-1609.
- Soifer, H., Higo, C., Kazazian, H.H. Jr., Moran, J.V., Mitani, K., and Kasahara, N. 2001. Stable integration of transgenes delivered by a retrotransposon-adenovirus hybrid vector. *Hum. Gene Ther.* 12:1417-1428.
- Soifer, H., Higo, C., Logg, C.R., Jih, L.J., Shichinohe, T., Harboe-Schmidt, E., Mitani, K., and Kasahara, N. 2002. A novel, helper-dependent, adenovirus-retrovirus hybrid vector: Stable transduction by a two-stage mechanism. *Mol. Ther.* 5:599-608.
- Soudais, C., Skander, N., and Kremer, E.J. 2003. Long-term in vivo transduction of neurons throughout the rat central nervous system using novel helper-dependent CAV-2 vectors. *FASEB J.* 18:391-393.
- Toietta, G., Pastore, L., Cerullo, V., Finegold, M., Beaudet, A.L., and Lee, B. 2002. Generation of helper-dependent adenoviral vectors by homologous recombination. *Mol. Ther.* 5:204-210.
- Umana, P., Gerdes, C.A., Stone, D., Davis, J.R., Ward, D., Castro, M.G., and Lowenstein, P.R. 2001. Efficient FLPe recombinase enables scalable production of helper-dependent adenoviral vectors with negligible helper-virus contamination. *Nat. Biotechnol.* 19:582-585.
- Yant, S.R., Ehrhardt, A., Mikkelsen, J.G., Meuse, L., Pham, T., and Kay, M.A. 2002. Transposition from a gutless adeno-transposon vector stabilizes transgene expression in vivo. *Nat. Biotechnol.* 20:999-1005.

Key References

Kochanek, 1999. See above.

An excellent comprehensive review for development of helper-dependent adenoviral vectors.

Contributed by Kazuhiro Oka and
Lawrence Chan
Baylor College of Medicine
Houston, Texas

CHAPTER 17

Preparation and Analysis of Glycoconjugates

INTRODUCTION

OVERVIEW

The modern revolution in molecular biology has been driven to a large extent by advances in methods for analysis and manipulation of DNA, RNA, and proteins. Although oligosaccharides (sugar chains) are also major macromolecules of the typical cell, they did not initially share in this molecular revolution. The reasons for this are, to a large extent, technical. Whereas DNA, RNA, and proteins are linear polymers that can usually be directly sequenced, oligosaccharides show substantially more complexity, having branching and anomeric configurations (α and β linkages). Thus, while three amino acids or nucleotides can be combined into six possible sequences, three hexose monosaccharides can theoretically generate 1056 possible oligosaccharides. Also, the synthesis of DNA, RNA, and proteins are template-driven, and the sequence of one can generally be predicted from that of another. In contrast, the biosynthesis of oligosaccharides, termed glycosylation, is extremely complex, is not template-driven, varies among different cell types, and cannot be easily predicted from simple rules.

It is clear that oligosaccharides have important, albeit varied, effects upon the biosynthesis, folding, solubility, stability, subcellular trafficking, turnover, and half-life of the molecules to which they are attached. These are matters of great importance to the cell biologist, protein chemist, biotechnologist, and pharmacologist. On the other hand, the successful growth of several glycosylation mutants as permanent tissue culture cell lines indicates that the precise structure of many oligosaccharides is not critical for the growth and viability of a single cell in the protected environment of the culture dish. Thus, until recently, it was possible for many molecular biologists working with in vitro single-cell systems to ignore the existence of oligosaccharides. However, with the increasing emphasis on studying cell-cell interactions in normal development, tissue morphogenesis, immune reactions, and pathological conditions such as cancer and inflammation, the study of

oligosaccharide structure and biosynthesis has become very important.

The term “glycobiology” has found widespread acceptance for denoting studies of the biology of glycoconjugates in both simple and complex systems. Many technical advances have occurred in the analysis of oligosaccharides, making it now feasible to study them in detail. These advances include the development of sensitive and specific assays and the availability of numerous purified enzymes (glycosidases) with a high degree of specificity (the glycobiologist’s equivalent of restriction enzymes). Thus, the time has come to include oligosaccharide molecules in the overall lexicon of molecular biology.

In spite of all these recent advances, many analytical techniques in glycobiology have remained in the domain of the few laboratories that specialize in the study of oligosaccharides. Likewise, published compendia of carbohydrate methods are designed mainly for use by experts. This chapter attempts to “democratize” this technology and place some of it within easy reach of any laboratory with basic capabilities in biochemistry and molecular biology. The techniques described here include modern versions of time-honored methods and recently developed methods, both of which have widespread applications. However, it is important to emphasize that the protocols presented here are by no means comprehensive. Rather, they serve as a starting point for the uninitiated scientist who wishes to explore the structure, biosynthesis, and biology of oligosaccharide chains. In most cases, further analysis using more sophisticated techniques will be required to obtain final and definitive results. Nonetheless, armed with results obtained by techniques described here, the average molecular biologist can make intelligent decisions about the need for such analyses.

The appearance of many commercial kits for analysis of glycoconjugates is another sign that the technology has arrived and that many laboratories have developed an interest in glycobiology. It is worth noting that while some of these kits seek to simplify the use of well-es-

Contributed by Ajit Varki, Hudson H. Freeze, and Adriana E. Manzi

Current Protocols in Molecular Biology (1995) 17.0.5-17.0.14

Copyright © 2000 by John Wiley & Sons, Inc.

**Preparation
and Analysis of
Glycoconjugates**

17.0.5

Supplement 22

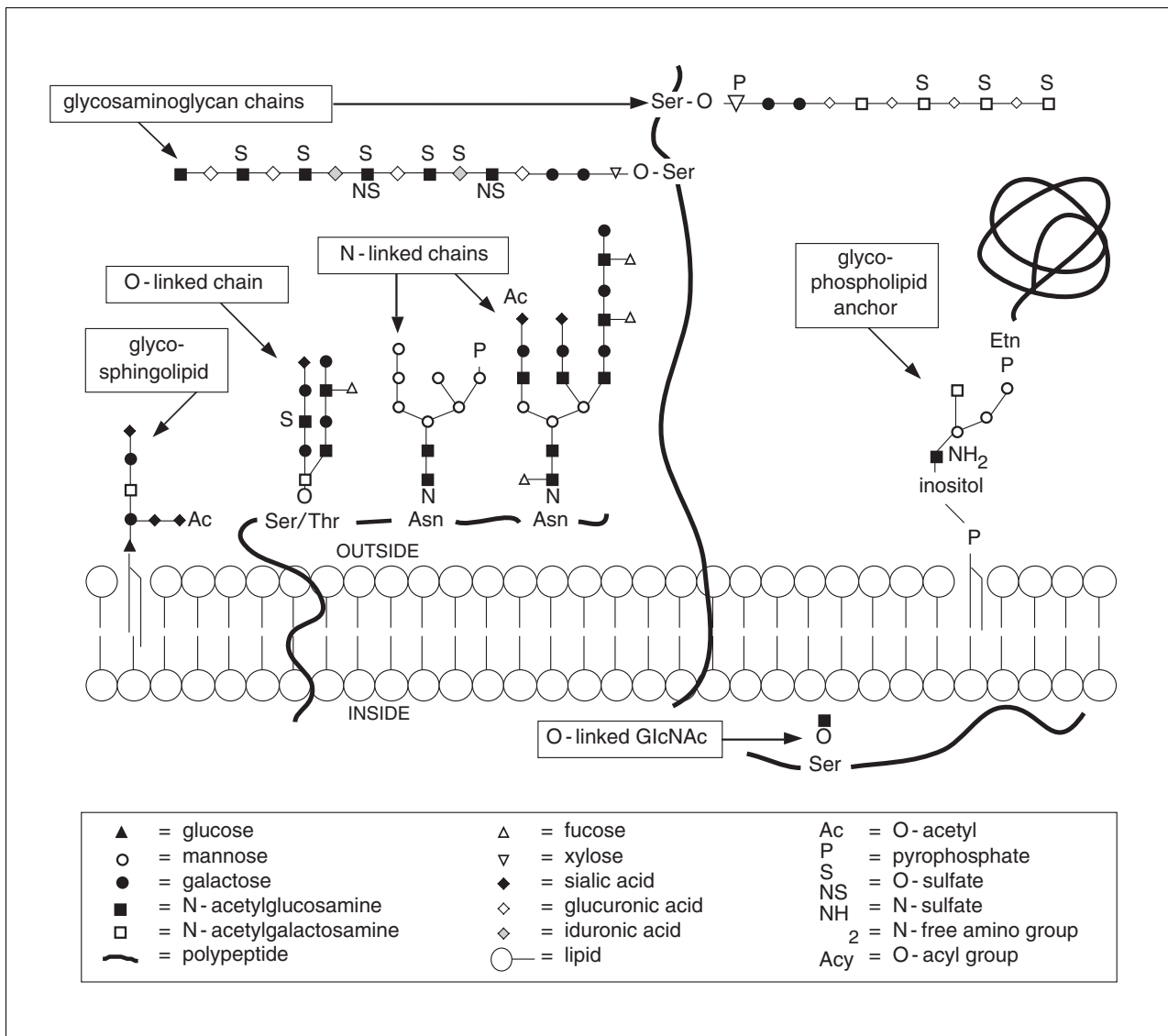


Figure 17.0.1 Common oligosaccharide linkage regions on animal cell glycoconjugates. The most common types of oligosaccharides found in animal glycoconjugates are shown, with an emphasis upon the linkage region between the oligosaccharide and the protein or lipid. Other rarer types of linkage regions and free oligosaccharides that can exist naturally (e.g., hyaluronan) are not shown.

established techniques, others contain methodologies that have been newly developed by the companies themselves. Experience with the latter methodologies may be limited in academic scientific laboratories, and the techniques in question may therefore not be represented in this chapter. However, although the admonition *caveat emptor* is appropriate, some of these kits may well become useful adjuncts to the methods presented in this chapter.

The different types of glycosylation are perhaps best defined by the nature of the linkage region of the oligosaccharide to a lipid or protein (Fig. 17.0.1). Although the linkage regions of these molecules are unique, the sugar chains frequently tend to share common types of outer

sequences. It is important to note that this chapter deals only with the major forms of glycosylation found on higher-animal glycoconjugates—*N*-acetylglucosamine (GlcNAc)-*N*-Asn-linked, *N*-acetylgalactosamine (GalNAc)-*O*-Ser/Thr-linked oligosaccharides on glycoproteins, xylose-*O*-Ser-linked glycosaminoglycans on proteoglycans, ceramide-linked glycosphingolipids, phosphatidylinositol-linked glyco-phospholipid anchors, and *O*-linked *N*-acetylglucosamine (GlcNAc-*O*-Ser). These structures are depicted in Figure 17.0.1. There are many other less common forms of glycosylation that one may need to consider, especially if prior literature suggests their existence in a given situation. Examples

of rarer sugar chains include (1) O-linked glucose, mannose, and fucose; (2) N-linked glucose; (3) glucosyl-hydroxylysine; and (4) GlcNAc-P-Ser (phosphoglycosylation).

Likewise, this chapter deals only with the most common forms of shared outer sequences (e.g., sialylated and fucosylated lactosamines, polylactosamines, *O*-glycosaminoglycan chains, and blood group sequences) and does not deal with rarer sequences (e.g., bisecting xylose residues, N-linked glycosaminoglycans, and β -linked GalNAc residues on N-linked oligosaccharides). Some time-honored techniques of somewhat lesser utility have been excluded, as have some requiring very sophisticated instrumentation. Likewise, several emerging technologies with great promise (such as fluorescent tagging and gel electrophoresis) have not been included. For further information, the reader is directed to the key references section at the end of this introduction.

CHOICE OF TECHNIQUES

For the novice experimenter in glycoconjugate analysis, the greatest difficulty is in deciding which protocols are applicable to the question at hand, are sensitive enough to yield results, and are most likely to give useful answers. The following tables are therefore provided as a general guide to glycoconjugate analysis.

Suggestions are made for the protocols that are most likely to be useful based upon the questions being asked (Table 17.0.1), the amount of material that is available for analysis (Table 17.0.2), and the type of glycoconjugate that is being studied (Table 17.0.3). The user is advised to select protocols based on the information in these tables and then to consult the commentary section of the selected protocols for further information regarding practicality and applicability to a particular experimental situation. Choice of technique may also be dictated by whether the glycoconjugate to be studied is being monitored with a radioactive tracer or by direct chemical analysis, and by what fraction of a precious purified sample can be expended for oligosaccharide analysis.

The techniques in this chapter are not sufficiently inclusive to give final confirmation of the complete structure of an oligosaccharide. In such cases, additional studies of the glycoconjugate may be necessary using more advanced techniques. In some cases, if the sample is highly purified and available in large amounts, use of nondestructive methods such as NMR should be considered before committing the sample to destructive analysis. Alternatively, the information generated by the techniques described in this chapter may be sufficient to answer the biological question(s) at hand.

Table 17.0.1 Protocol Choice for Glycoconjugate Analysis Based on Question Being Asked

Question	Suggested protocols (unit no.)
Is there anything special about purifying my glycoconjugate?	17.1, 17.2, 17.3
Is my protein glycosylated?	17.1, 17.2, 17.4, 17.5, 17.6, 17.7, 17.8, 17.9, 17.10, 17.12, 17.13, 17.17, 17.18
How much of my glycoconjugate consists of sugar chains?	17.9, 17.10, 17.12, 17.13, 17.17, 17.18
What monosaccharides are in my glycoconjugate, and in what ratio?	17.4, 17.9, 17.12, 17.16, 17.18, 17.19
How many glycosylation sites are there on my protein?	17.10, 17.13, 17.14B
Can I specifically label the sugar chains on my glyconjugate?	17.4, 17.5, 17.6
Does my antibody recognize sugar chains on the glycoconjugate?	17.7, 17.8, 17.12, 17.13, 17.17
Can I selectively release the sugar chains from my glycoconjugate?	17.8, 17.13, 17.15, 17.17
Does my protein have a glycopospholipid anchor?	17.4, 17.8
What type of glycosphingolipids does my cell have?	17.3, 17.4, 17.7
Are there glycosaminoglycan chains on my protein?	17.2, 17.13, 17.17, 17.22
Can I alter glycosylation during biosynthesis?	17.10, 17.11
Can I alter glycosylation on the surface of intact cells?	17.6, 17.10, 17.12, 17.13, 17.17
Can I release and isolate intact or fragmented oligosaccharides from my glycoconjugate?	17.8, 17.12, 17.13, 17.14, 17.15, 17.17
What are the basic structural characteristics of the released oligosaccharides?	17.5, 17.6, 17.12, 17.13, 17.20, 17.21, 17.22, 17.23

Table 17.0.2 Protocol Choice Based on Amount of Glycoconjugate Available

Unit no.	Sensitivity range for typical usage ^{a,b}				
	>100 nmol	10–100 nmol	1–10 nmol	1–0.1 nmol	<0.1 nmol
17.4	—	P*	U*	U*	U*
17.5	U	U	U*	U*	U*
17.6	P	U*	U*	U*	U*
17.7	—	—	U	U	U
17.8	—	P	U	U	U*
17.9	U	U	P	—	—
17.10	P	U	U*	U*	—
17.11	P	U	U*	U*	—
17.12	U	U	U	U	U
17.13	P	U	U	U*	U*
17.14	U	U	U	U*	U*
17.15	U	U	U	U*	U*
17.16	U	U	U	U	U
17.17	P	U	U	U*	U*
17.18	U	U	P	U	U
17.19A	—	—	U	U	P
17.19B	—	—	P*	U*	U*
17.20	—	U	U*	U*	U*
17.21	—	—	U	U*	U*
17.22A	U	U	U	U	U*
17.22B	—	—	U	U	U*
17.23	U	U	U	U	U*

^aAbbreviations: P, possibly useful; U, useful; * radioactive samples.

^bThe average molecular weight of a monosaccharide unit of sugar chains is ~200. To start, assume the glycoconjugate contains 5 to 10 mol monosaccharide/mol glycoconjugate. Radioactive samples should have a specific activity from 0.1 to 1 Ci/mmol.

STEREOCHEMISTRY AND DIAGRAMMATIC REPRESENTATION

Basic Stereochemistry of Monosaccharides

In addition to this discussion, the reader is referred to Allen and Kisalius (1992) for a more detailed discussion of the principles of carbohydrate structure. The italicized terms in the discussion below are defined in the glossary at the end of this unit.

Glyceraldehyde, the simplest monosaccharide, has only one chiral (asymmetric) carbon (C-2). Therefore, it is a chiral molecule that shows optical isomerism; it can exist in the form of two nonsuperimposable mirror images called *enantiomers* (see Fig. 17.0.2). These enantiomers have identical physical properties, except for the direction of rotation of the plane

of polarized light (–, left hand; +, right hand). Historically, the (+)-glyceraldehyde was arbitrarily assigned the prefix **D** (for dextrorotatory), and the (–)-glyceraldehyde, the prefix **L** (for levorotatory). The pair of enantiomers also have identical chemical properties, except toward optically active reagents. This fact is particularly important in biological systems, because most enzymes and the compounds they work on are optically active.

The configuration of the highest-numbered asymmetric carbon atom in the chain of higher monosaccharides is determined by comparison with the configuration of the chiral (or asymmetric) carbon of glyceraldehyde. Thus, a prefix **D**- is added to the name of each monosaccharide having the configuration of **D**-glyceraldehyde at the highest-numbered asymmetric carbon, and the prefix **L**- to those having the configuration of **L**-glyceraldehyde at the high-

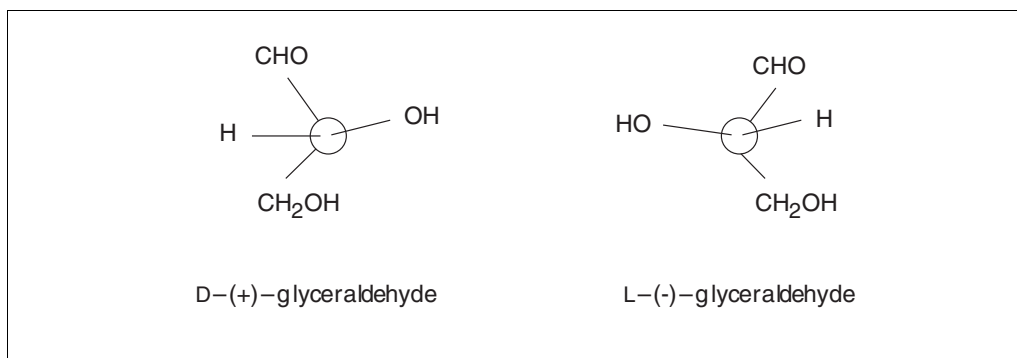


Figure 17.0.2 Enantiomers of glyceraldehyde.

est-numbered asymmetric carbon. The direction of rotation of the plane of polarized light for the compound needs to be determined experimentally, and is indicated with the – or + sign, between parentheses, immediately after the prefix D or L (e.g., D-(+)-glucose). In biological systems, where stereochemical specificity is the rule, D molecules may be completely active and L molecules may be completely inert, or vice versa. All known naturally

occurring monosaccharides in animal cells are in the D configuration, except for fucose and iduronic acid, which are in the L configuration.

Each *aldose* is usually present in a cyclic structure, because the *hemiacetal* produced by reaction of the aldehyde group at C-1 with the hydroxyl group at C-5 gives a six-membered ring, called a *pyranose* (see Fig. 17.0.3). When five-membered rings are formed by reaction of the C-1 aldehyde with the C-4 hydroxyl, they

Table 17.0.3 Protocol Choice Based on Type of Glycoconjugate Studied^a

Unit no.	Type of glycoconjugate ^b					
	<i>N</i> -Glc-NAc-linked glyco-protein	<i>O</i> -GalNAc-linked glyco-protein	<i>O</i> -Xylose-linked proteo-glycan	Glyco-sphingo-lipid	Glyco-phospho-lipid anchor	O-linked GlcNAc
17.4	U ^b	U	U	U	U	P
17.5	U	U	—	U	P	—
17.6	U	U	—	—	P	U
17.7	U	U	—	U	P	U
17.8	—	—	—	—	U	—
17.9	U	U	—	P	P	—
17.10	U	—	—	—	—	—
17.11	—	U	U	—	—	—
17.12	U	U	—	U	P	—
17.13	U	—	—	—	—	—
17.14A	U	P	P	—	—	—
17.14B	U	P	—	—	P	—
17.15	—	U	—	—	—	—
17.16	U	U	—	U	P	P
17.17	—	—	—	U	—	—
17.18	U	U	P	P	P	P
17.19A	U	U	—	U	U	P
17.19B	—	—	U	—	—	—
17.20	U	P	—	P	—	—
17.21	U	U	—	P	—	—
17.22A	P	P	U	—	U	—
17.22B	—	—	U	—	—	—
17.23	U	U	U	P	—	—

^aThis chapter deals only with the most common forms of glycosylation on glycoconjugates.

^bAbbreviations: P, possibly useful; U, useful.

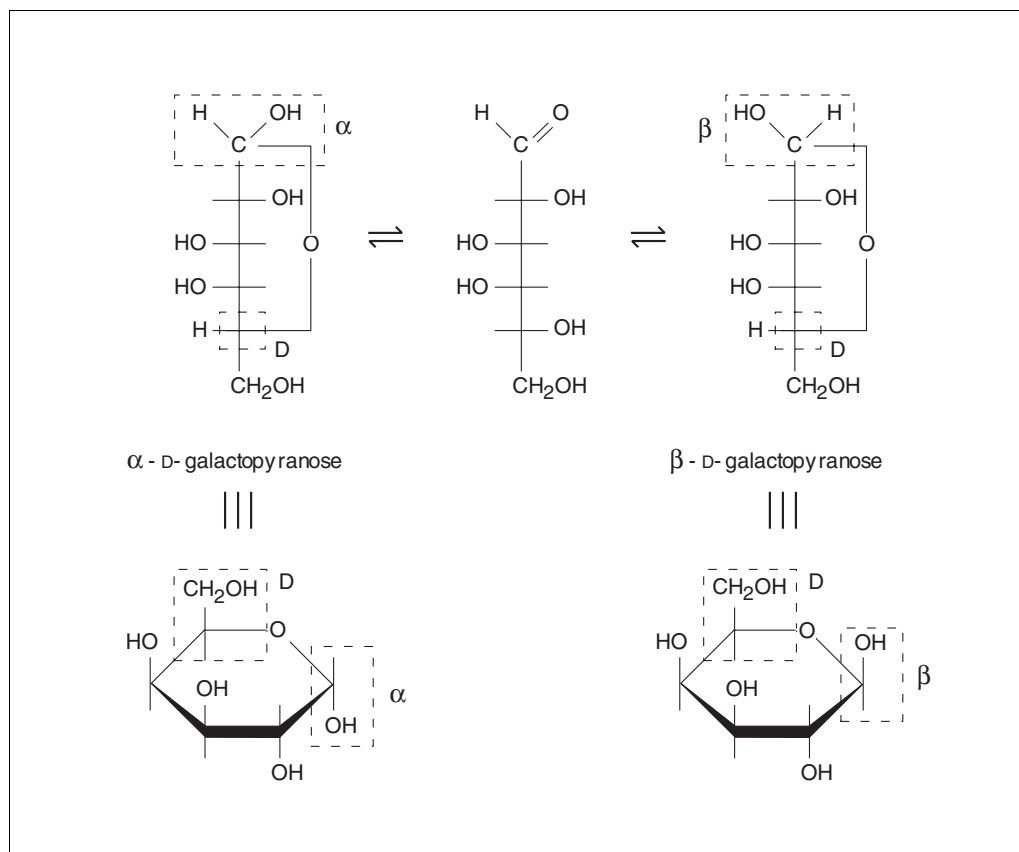


Figure 17.0.3 Formation of cyclic structures: open and ring forms of galactose.

are called *furanoses*. Hexoses commonly form pyranosidic rings, and pentoses form furanosidic rings, although the reverse is possible. Similarly, ketoses (e.g., sialic acids) can form *hemiketals* by reaction of the keto group at C-2 with the hydroxyl group at C-6.

These ring forms are the rule in intact oligosaccharides. The formation of the ring produces an additional chiral center at C-1 (or C-2 for keto sugars). Thus, two additional isomers are possible, and these are called *anomers*. These anomers are designated α and β . Because glycosidic linkages occur via anomeric centers, they are called α linkages and β linkages.

Diagrammatic Representations of Monosaccharides

Monosaccharides are conventionally presented in the Fisher projection or the Haworth representation (see Fig. 17.0.4).

The Fisher projection

The carbon chain of the monosaccharide is written vertically with carbon atom 1 at the top. Horizontal lines represent bonds projecting out

from the plane of the paper, whereas vertical lines represent bonds projecting behind the plane. When the hydroxyl group at the highest-numbered asymmetric carbon is on the right, the monosaccharide belongs to the D series. When this group is on the left, the monosaccharide belongs to the L series.

The Haworth representation

The six-membered ring with the oxygen at the upper-right corner is approximately perpendicular to the plane of the paper, with the groups attached to the carbons above or below the ring. All groups that appear to the right in the Fisher projection appear below the plane of the ring in the Haworth representation. The carboxymethyl group (C-6) lies above the ring in D series hexoses, below it in L series. The only difference between the α and β anomers of a given hexose is the relative position of the hydroxyl and the hydrogen at C-1. In the α anomer, the hydrogen is below the ring for the D series (above it for L series); in the β anomer the hydrogen is above the ring for the D series (below for L series).

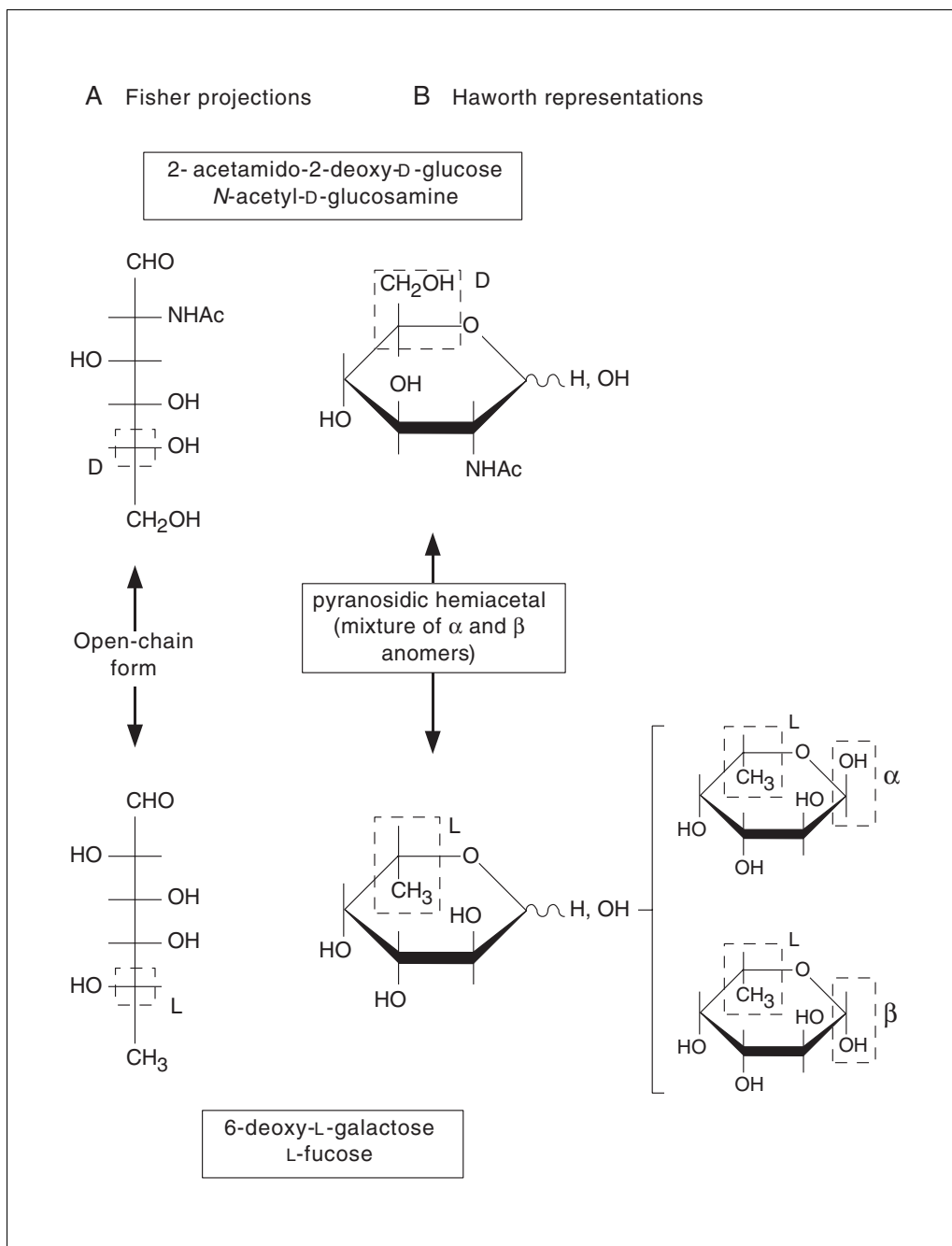


Figure 17.0.4 Fisher projections and Haworth representations of monosaccharides.

Conformation of Monosaccharides in Solution

The Haworth representation does not depict the real conformation of these molecules. The preferred conformation in solution of the six-membered ring monosaccharides is a chair conformation, in which each group can adopt either an *equatorial* or *axial* position.

Two possible chair conformations called *conformers* exist in equilibrium. The direction of equilibrium varies among monosaccharides depending on the relative position of hydroxyl

groups or other substituents. The preferred conformation has the lowest number of bulky groups in an axial position (Fig. 17.0.5).

Formulas for Representation of Oligosaccharide Chains

Full and correct representation of the formula of an oligosaccharide chain requires notation of the complete stereochemistry of the component monosaccharides. Thus, the simple tetrasaccharide commonly called sialyl-Lewis^x

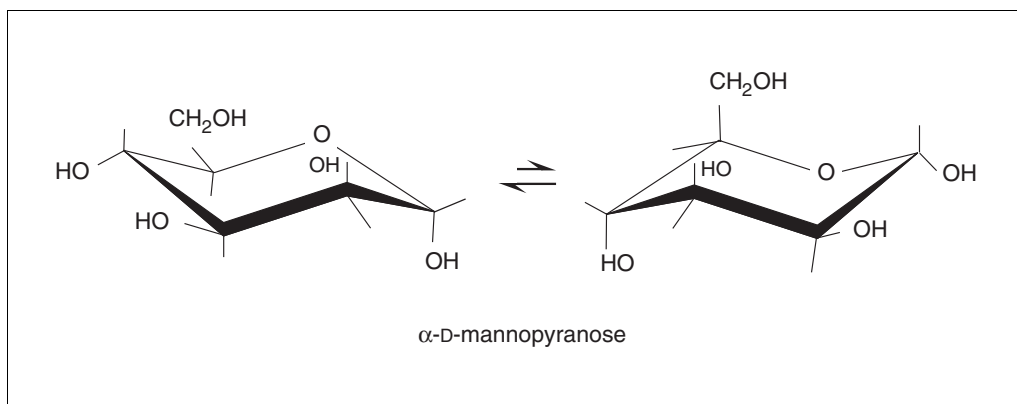
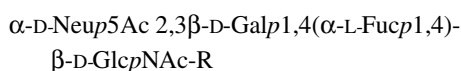
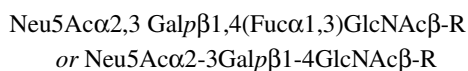


Figure 17.0.5 D-Mannose chair conformation.

(attached to an underlying oligosaccharide, R), would be written as follows:



In more common practice, the D and L configurations and nature of ring structures are assumed, and the formula is written as:



3
|
1
Fuc α

Higher-Order Structures

The presence of monosaccharides with different types of glycosidic linkages in oligo- and polysaccharides creates further complexity, with the development of secondary and tertiary structures in these molecules. The approximate shape adopted in solution by a given carbohydrate chain can be predicted according to the type of glycosidic linkages involved. These higher-order structures may be critical for the biological roles of glycoconjugates.

GLOSSARY

aldose monosaccharide with a carbonyl group at the end of the carbon chain (aldehyde group); the carbonyl group is assigned the lowest possible number (i.e., carbon 1).

carbohydrates polyhydroxyaldehydes or polyhydroxyketones, or compounds that can be hydrolyzed to them.

diastereoisomer compound with identical formulas that have a different spatial distribution of atoms (e.g., galactose and mannose).

enantiomers nonsuperimposable mirror images of any compound (e.g., D- and L-glucose).

epimers two monosaccharides differing only in the configuration of a single chiral carbon.

Fuc L-fucose type of deoxyhexose (see also *types of monosaccharides*).

Gal D-galactose type of hexose (see also *types of monosaccharides*).

GalNAc (N-acetyl-D-galactosamine) type of hexosamine (see *types of monosaccharides*).

ganglioside anionic glycolipid containing one or more units of sialic acid.

Glc D-glucose type of hexose (see also *types of monosaccharides*).

GlcNAc (N-acetyl-D-glucosamine) type of hexosamine (see also *types of monosaccharides*).

glyceraldehyde simplest monosaccharide (an aldotriose; i.e., with three carbon atoms).

glycoconjugate natural compound with one or more mono- or oligosaccharide units covalently linked to a noncarbohydrate moiety.

GLUA or GIA (D-glucuronic acid) type of uronic acid (see also *types of monosaccharides*).

glycolipid or glycosphingolipid oligosaccharide attached via glucose or galactose to the terminal primary hydroxyl group of ceramide, which is composed of a long-chain base (i.e., sphingosine) and a fatty acid. Glycolipids can be neutral or anionic (negatively charged).

glycophospholipid anchor glycan bridge between phosphatidylinositol and a phosphoethanolamine in amide linkage to the C terminus of a protein; constitutes the sole membrane anchor for such proteins.

glycoprotein glycoconjugate in which a protein carries one or more oligosaccharide chains covalently attached to a polypeptide backbone via N-GlcNAc- or O-GalNAc-linkages.

glycosaminoglycan linear copolymers of disaccharide-repeating units, each composed of hexosamine and hexose or hexuronic acid; these oligosaccharide chains define the proteoglycans. The type of disaccharide unit can define glycosaminoglycans as chondroitin or dermatan sulfate, heparan or heparin sulfate, hyaluronic acid, and keratan sulfate. Glycosaminoglycans (except hyaluronic acid) also contain sulfate esters substituting hydroxyl or amino groups (*N*- or *O*-sulfate groups).

glycosidic linkage linkage of a monosaccharide to another residue via the anomeric hydroxyl group.

IdUA or IdA (*L*-iduronic acid) type of uronic acid (see also *types of monosaccharides*).

ketose monosaccharide with a carbonyl group in an inner carbon; the carbonyl group is assigned the lowest possible number (i.e., carbon 2).

Man *D*-mannose type of hexose (see also *types of monosaccharides*).

monosaccharide carbohydrate that cannot be hydrolyzed into simpler units (see also *types of monosaccharides*).

mucin large glycoproteins that contain many (up to several hundred) *O*-GalNAc-linked oligosaccharide chains that are often closely spaced.

***N*-linked oligosaccharide** oligosaccharide covalently linked to an asparagine residue of a polypeptide in the consensus sequence: -Asn-X-Ser/Thr. The many different kinds of *N*-linked oligosaccharides have certain common features: (1) a common core pentasaccharide containing two mannosyl residues α -linked to a third mannosyl unit, in turn β -linked to a chitobiosyl group; (2) a chitobiosyl group β -linked to the asparagine amide nitrogen. *N*-linked oligosaccharides can be divided into three main classes: high-mannose, complex, and hybrid.

Neu5Ac (*N*-acetyl-*D*-neuraminic acid) type of sialic acid (Sia; see also *types of monosaccharides*).

***O*-linked oligosaccharide** oligosaccharide linked to the polypeptide via *N*-acetylgalactosamine (GalNAc) to serine (or threonine). Note that other types of *O*-linked oligosaccharides also exist (e.g., *O*-GlcNAc) but that the *O*-GalNAc linkage, being the most well-known, is often described by this generic term.

oligosaccharide branched or linear chain of monosaccharides attached to one another via glycosidic linkages (see also *polysaccharides*).

polylactosaminoglycan long chain of repeating units of the disaccharide β -Gal(1-4)-GlcNAc. These chains may be modified by sialylation, fucosylation, or branching. When sulfated, they are called keratan sulfate (see *glycosaminoglycans*).

polysaccharides branched or linear chain of monosaccharides attached to each other via glycosidic linkages that usually contain repetitive sequences. The number of monosaccharide residues that represents the limit between an oligosaccharide and a polysaccharide is not defined. A tetra-fucosylated, sialylated tetraantennary carbohydrate moiety containing eighteen monosaccharide units, present in a glycoprotein, is considered an oligosaccharide, while a carbohydrate composed of eighteen glucose residues linked β -(1-4) present in a plant extract is considered a polysaccharide.

polysialic acid homopolymer of sialic acid selectively expressed on a few vertebrate proteins and on the capsular polysaccharides of certain pathogenic bacteria.

proteoglycan glycoconjugate having one or more *O*-xylose-linked glycosaminoglycan chains (rather than *N*-GlcNAc- or *O*-GalNAc-linked oligosaccharides) linked to protein. The distinction from a glycoprotein is otherwise arbitrary, because some proteoglycans can have both glycosaminoglycan chains and *N*- or *O*-linked oligosaccharides attached to them.

reducing sugar sugar that undergoes typical reactions of aldehydes (e.g., is able to reduce Ag^+ or Cu^{++}). Mono-, oligo-, or polysaccharides can be reducing sugars when the aldehyde group in the terminal monosaccharide residue is not involved in a glycosidic linkage.

saccharide modifications hydroxyl groups of different monosaccharides can be subject to phosphorylation, acetylation, sulfation, methylation, or fatty acylation. Amino groups can be free, *N*-acetylated, or *N*-sulfated. Carboxyl groups are occasionally subject to lactonization to nearby hydroxyl groups.

Sia (*sialic acid*) generic name for a family of acidic nine-carbon monosaccharides (e.g., Neu5Ac).

types of monosaccharides monosaccharides may have a carbonyl group at the end of the carbon chain (aldehyde group), or in an inner carbon (ketone group). The carbonyl group is assigned the lowest possible number, e.g., carbon 1 (C-1) for the aldehyde group; carbon 2 (C-2) for the most common ketone groups. These two types are named aldoses and ketoses.

The simplest monosaccharide is glyceraldehyde (see Fig. 17.0.2), an aldotriose (containing three carbon atoms). Natural aldoses with different number of carbon atoms in their chain are named accordingly (e.g., aldohexoses, containing six carbon atoms). Two monosaccharides differing only in the configuration of a single chiral carbon are called epimers. For example, glucose and galactose are epimers of each other at C-4. Common monosaccharides present in animal glycoconjugates are (1) deoxyhexoses (e.g., L-Fuc); (2) hexosamines, usually N-acetylated (e.g., D-GalNAc and D-GlcNAc); (3) hexoses (e.g., D-Glc, D-Gal, and D-Man); (4) pentoses (e.g., D-Xyl); (5) sialic acids (Sia; e.g., Neu5Ac); (6) uronic acids (e.g., D-GlA and L-IdA).

Xyl D-xylose type of pentose (see also *types of monosaccharides*).

Further detailed information on the analysis of oligosaccharides is provided in the specific literature cited in individual protocol units. The following sources also can be used for general information, details on specific methods, and for many methods that are not included in this chapter.

LITERATURE CITED

- Allen, H.J. and Kisalius, E.C. (eds.). 1992. Glycoconjugates: Composition, Structure, and Function. pp. 167-202. Marcel Dekker, N.Y.
- Beeley, J.G. 1985. Glycoprotein and Proteoglycan Techniques. Elsevier/North-Holland, Amsterdam and New York.
- Chaplin, M.F. and Kennedy, J.F. (eds.). 1987. Carbohydrate Analysis: A Practical Approach. IRL Press, Oxford.
- Clamp, J.R., Bhatti, T., and Chambers, R.E. 1971. The determination of carbohydrates in biological materials by gas-liquid chromatography. *Methods Biochem. Anal.* 19:229-344.
- Cummings, R.D., Merkle, R.K., and Stults, N.L. 1989. Separation and analysis of glycoprotein oligosaccharides. *Methods Cell Biol.* 32:141-183.
- Dell, A. 1987. F.A.B.-mass spectrometry of carbohydrates. *Adv. Carbohydr. Chem. Biochem.* 45:19-72.

- Esko, J.D. 1991. Genetic analysis of proteoglycan structure, function and metabolism. *Curr. Opin. Cell Biol.* 3:805-816.
- Fleischer, S. and Fleischer, B. (eds.). 1983. Biomembranes, Part L. *Methods Enzymol.* Vol. 98.
- Fukuda, M. (ed.). 1993. Glycobiology: A Practical Approach. IRL Press, Oxford. In press.
- Ginsburg, V. (ed.). 1972. Complex Carbohydrates, Part B. *Methods Enzymol.* Vol. 28.
- Ginsburg, V. (ed.). 1978. Complex Carbohydrates, Part C. *Methods Enzymol.* Vol. 50.
- Ginsburg, V. (ed.). 1982. Complex Carbohydrates, Part D. *Methods Enzymol.* Vol. 83.
- Ginsburg, V. (ed.). 1987. Complex Carbohydrates, Part E. *Methods Enzymol.* Vol. 138.
- Ginsburg, V. (ed.). 1989. Complex Carbohydrates, Part F. *Methods Enzymol.* Vol. 179.
- Hart, G. and Lennarz, W. 1993. Guide to Techniques in Glycobiology. *Methods Enzymol.* In press.
- Lane, D.G. and Lindahl, U. (eds.). 1990. Heparin: Chemical and Biological Properties. CRC Press, Boca Raton, Fla.
- McCloskey, J.A. (ed.). 1990. Mass Spectrometry. *Methods Enzymol.* Vol. 193.
- Stanley, P. 1992. Glycosylation engineering. *Glycobiology* 2:99-107.
- Varki, A. 1991. Radioactive tracer techniques in the sequencing of glycoprotein oligosaccharides. *FASEB J.* 5:226-235.
- Whistler, R. and Wolfram, F. (eds.). 1968-1980. Methods in Carbohydrate Chemistry, Vols I-VIII. Academic Press, San Diego.

Contributed by Ajit Varki
Contributing editor (Chapter 17)
University of California San Diego
La Jolla, California

Hudson H. Freeze (choice of
technique)
La Jolla Cancer Research Foundation
La Jolla, California

Adriana E. Manzi (stereochemistry
and glossary)
University of California San Diego
La Jolla, California

SPECIAL CONSIDERATIONS OF GLYCOCONJUGATES AND THEIR PURIFICATION

Because of their unique composition and structural characteristics, there are instances in which glycoconjugates can be studied in biological mixtures directly, without purification. However, as Arthur Kornberg has said, to answer a biological question the best course of action is usually to “purify, purify, purify.” Indeed, the complete elucidation of any biological system involving a glycoconjugate will require that it be purified for structural and functional characterization.

This section describes general principles for purifying three major classes of glycoconjugates—glycoproteins, glycolipids, and proteoglycans. These classes have been described in the chapter introduction and glossary. Each is a large and polydisperse family of molecules with widely ranging structural and functional characteristics. Needless to say, many of the classic principles used for purification of other macromolecules (e.g., size fractionation and selective precipitation) apply to each of these types of glycoconjugates; on the other hand, an individual situation may provide unique opportunities for selective purification of a particular molecule.

The units presented in this section take the middle road and outline the principles particularly applicable to the purification of each class of glycoconjugate, taking advantage of group-specific properties. Glycoproteins usually have sugar chains that can be recognized by specific plant lectins; thus, affinity chromatography using immobilized lectins is a powerful technique, when properly applied (*UNIT 17.1*). Proteoglycans are characterized by a high density of negative charge and frequently by unusual buoyant densities and sedimentation properties, which can be used to advantage in their purification (*UNIT 17.2*). Glycolipids as a group are more hydrophobic than the other glycoconjugates, yet have amphipathic properties that allow them to be purified away from other lipids (*UNIT 17.3*).

Selective purification of these groups of molecules facilitates their further fractionation and final characterization. However, it is important to realize that exceptions can be encountered (e.g., glycoproteins with unusually high negative charges or glycolipids with extremely hydrophilic properties). Also, the phenomenon of microheterogeneity in oligosaccharide structure can cause even a single glycoconjugate to manifest a range of properties. For example, if a single glycoprotein from a single cell type has heterogeneity in N-linked oligosaccharide processing, some (but not all) of the protein may be recovered during affinity chromatography on wheat germ agglutinin (WGA)-Sepharose (*UNIT 17.1*).

At the present time there is no generalized purification scheme that will provide good yields of all of the major classes of glycoconjugates from a single starting sample. For example, the organic extractions required for glycolipids (*UNIT 17.3*) irreversibly denature most glycoproteins and proteoglycans. Conversely, the glycolipids are usually irretrievable from the detergent extractions used for membrane-bound glycoproteins (*UNIT 17.1*). Thus, if there is a need to examine all of the major glycoconjugates in a given cell or tissue, it is advisable to purify the major classes of molecules in parallel with one another, using separate aliquots of the same batch of starting material.

Even when studying a specific molecule, it is usually worthwhile to carry out a group-specific purification as a first step—e.g., to extract total cellular lipids before purifying away a specific disialoganglioside (*UNIT 17.3*). If the glycoconjugate in question has been radiolabeled (see *UNIT 17.4*), one also has the advantage that purification to “radiometric homogeneity” is sufficient for many types of analyses. Thus, if a glycoprotein has been

metabolically labeled with [2-³H]mannose and the intent is to study the N-linked oligosaccharides, it may not matter that various unlabeled proteins or lipids are contaminating the final preparation. In the final analysis, the broad principles outlined in *UNITS 17.1-17.3* must be adapted by the investigator to the situation at hand.

UNIT 17.1

Special Considerations for Glycoproteins and Their Purification

Glycoproteins contain a variety of different sugar chains. Some glycoproteins have many large and highly charged chains while others have only a single neutral sugar residue that is easily overlooked. This enormous diversity makes it impossible to recommend any single procedure for purifying glycoproteins. Most procedures used to purify nonglycosylated proteins are equally useful for glycoproteins. However, the presence of large amounts of carbohydrate on glycoproteins often endows them with higher charge, increased density, larger apparent size, and greater solubility than nonglycosylated proteins. Decades ago, these physical properties were cleverly exploited to purify and characterize many of the most abundant glycoproteins (Gibbons, 1972). More recently, greater knowledge of carbohydrate structure and biosynthesis has led to development of techniques that rely on more subtle and specific aspects of carbohydrate structure. Some of these approaches can also be used to detect a single sugar on a few picomoles of protein (see *UNIT 17.7*).

This unit begins by describing some properties of glycoproteins—e.g., subcellular location and solubility—that may be useful in determining which purification techniques to try. This discussion is followed by two protocols describing preparative glycoprotein purification using lectin-affinity chromatography, as well as an outline for a small-scale pilot procedure designed to check lectin binding and elution conditions. Lectins are often used for purifying glycoproteins because, in contrast to conventional purification procedures (e.g., gel filtration and ion-exchange chromatography; *UNITS 10.9 & 10.10*) that exploit general physical properties of glycoproteins, lectins recognize specific three-dimensional structures created by a cluster of sugar residues. Conventional purification procedures are generally tried before applying lectin-affinity chromatography.

SPECIAL CONSIDERATIONS IN THE PURIFICATION OF GLYCOPROTEINS

Subcellular Location of Glycosylated Proteins

Most protein glycosylation occurs in the lumen of the endoplasmic reticulum and the Golgi apparatus, but certain types of glycosylation also occur in the cytoplasm (see *UNIT 17.13*). Secreted proteins, membrane proteins, and proteins that enter vesicles such as lysosomes are good candidates for having some type of carbohydrate component. The most common types of sugar chains are found in N- and O-linkages to proteins and as components of glycopospholipid anchors. Each of these types of glycosylation can occur individually or in any combination. The physical properties they impart can be effectively used for some purifications. Knowledge of the type of glycosylation is thus necessary in order to determine the optimal purification method. The methods outlined below may be helpful for determining glycosylation type.

Secreted glycoproteins. It is relatively simple to determine if the protein is secreted by analyzing conditioned medium used to culture the cells. This may require dialysis and concentration of the conditioned medium. If the cells must be grown in the presence of added serum components during the time when the protein is secreted, it may be very difficult to use any specific glycoprotein purification approach to enrich for your protein because competing serum glycoproteins will probably be in great excess. Thus it is important to determine if the target molecule can be secreted into serum-free medium.

Cytoplasmic glycoproteins. If the protein is not secreted, cells can be ruptured in an iso-osmotic medium (e.g., sucrose) and centrifuged to separate insoluble membranes and vesicles

from soluble cytoplasmic components (UNIT 17.6). Glycoproteins are not commonly found in the soluble fraction. If the protein of interest is detected in the soluble fraction and is later found to be glycosylated, it may have been proteolyzed from a membrane. This is the most likely explanation for N-linked glycoproteins that are found in the soluble fraction. It is also possible that the protein is one of the recently described cytoplasmic or nuclear glycoproteins that contain O-linked *N*-acetylglucosamine (GlcNAc), in which case it may bind to wheat germ agglutinin (WGA, from *Triticum vulgare*; alternate protocol).

Organelle glycoproteins. If the protein of interest is not present in the soluble cytoplasmic fraction, the membrane and vesicle pellet is then resuspended in the presence of protease inhibitors and briefly exposed to a hypo-osmotic medium (or frozen and thawed) to lyse sensitive organelles and release their soluble contents. Membranes are again sedimented by centrifugation. Glycoprotein solubilized by this procedure will be accompanied by lysosomal enzymes. These proteases are potentially troublesome, but protease inhibitors will lessen or eliminate their effect. Various exo- and endoglycosidases may also be released by this procedure, but they are usually present at such a low concentration that they do not degrade the sugar chains on other glycoproteins.

Membrane-bound and glycopospholipid-anchored proteins. If the protein of interest is not solubilized by any of the above procedures, it is probably a peripheral or integral membrane protein, or it is anchored to glycopospholipid. Different nonionic detergents or increasing concentrations of a single detergent can be used to solubilize these proteins. Detergents that have high critical micelle concentrations (CMC) and can be easily dialyzed are preferred, e.g., hexyl-, heptyl-, and octyl-glucosides or glucamide detergents (MEGA-8, Calbiochem). However, a glucose-based detergent such as octyl-glucoside may block protein binding on some lectin columns such as Concanavalin A (Con A)-Sephacrose, or may produce a very high background in direct sugar analysis (UNIT 17.9) even after extensive dialysis. If the protein is glycopospholipid-anchored, the lipid component will cause it to partition into the detergent phase of a Triton X-114/water two-phase system (UNIT 17.8) and to interact strongly with hydrophobic matrices such as phenyl-Sephacrose. In some instances, the lipid portion of the anchor may be removed by digestion with phosphatidylinositol-specific phospholipase C

(UNIT 17.8). This cleavage leaves the carbohydrate portion of the anchor associated with the protein and converts the lipophilic protein into a soluble, more hydrophilic protein. Phospholipase digestion reduces the protein's affinity for hydrophobic matrices and results in its extraction into the aqueous phase of Triton X-114/water mixture (UNIT 17.8).

Solubility Properties

The solubility properties of a glycoprotein with a small amount of carbohydrate (~5%) will not be influenced much by the sugar component. If the carbohydrate is part of a glycopospholipid anchor, detergent solubilization may be required. Standard ammonium sulfate or ethanol precipitation (UNIT 10.9) works well for many glycoproteins, but those with high sugar content may remain soluble, even in relatively high concentrations of ethanol. Very anionic proteins with abundant sialic acid or sulfate esters may be resistant to precipitation with trichloroacetic acid or perchloric acid (APPENDIX 2). This unusual solubility is sometimes exploited as a purification step because most other proteins will be precipitated under these conditions. However, a combination of 10% trichloroacetic acid and 2% phosphotungstic acid will precipitate even highly charged proteins. Prolonged exposure to low pH may cause partial loss of sialic acid and deamidation of Asn and Gln, resulting in unwanted protein microheterogeneity. Extremes of pH should be avoided because of the possibility of hydrolysis of labile groups such as phosphodiester or acetyl esters.

Nonglycosylated proteins band in cesium chloride gradients at ~1.3 g/ml and polysaccharides at ~1.6 to 2.0 g/ml. This substantial difference in density can be used to separate very highly glycosylated proteins from those containing lesser amounts of carbohydrate. This treatment may permanently denature or inactivate some proteins, and is useful only if the protein is stable, is active in the presence of CsCl₂, or renatures once the CsCl₂ is removed.

Chromatographic and Staining Properties

Gel-filtration columns and SDS-PAGE. Proteins with high carbohydrate content (>20 to 30%) elute much earlier on gel-filtration columns (UNIT 10.9) than nonglycosylated proteins of similar size. If the amount of carbohydrate is variable, it may broaden the peak as well. Migration of proteins in SDS-PAGE (UNIT 10.2) is also affected by the extent of glycosylation. A protein with multiple anionic sugar

chains, each with a variable number of sialic acids or sulfate esters, can broaden a single protein band into a smear on gels. Closely clustered, short, O-linked sugar chains can appear to contribute as much as ten times their actual mass to the apparent molecular weight of a protein.

Very anionic glycoproteins may not stain with the normal Coomassie brilliant blue or silver staining procedures (UNIT 10.6). Some glycoproteins may even be seen as negative bands against the brownish background of silver-stained gels. Other stains, such as the Coomassie-based ProBlue stain (Integrated Separation Systems) or the cationic dyes toluidine blue and alcian blue, are often used for staining proteoglycans and other highly anionic proteins.

Ion-exchange and isoelectric focusing chromatography. The presence of sialic acids and sulfate esters makes some glycoproteins bind well to ion-exchange columns (e.g., DEAE-Sephadex, DEAE-Sephacel; UNIT 10.10). High-salt solutions elute these glycoproteins well after most other proteins have been eluted. The variable number of negative charges can also give multiple separate or very broad and poorly resolved protein peaks (UNITS 10.3 & 10.10). Multiple peaks or broad bands are also encountered in isoelectric focusing gels. Sialidase or endoglycosidase digestions can sometimes simplify these patterns by removing negative charges.

In the case of “mucins” (highly charged glycoproteins with closely clustered or extended O-linked GalNAc chains) or proteoglycans, ion-exchange columns may need to be run in 6 M urea to prevent protein aggregation.

Preparation of Protein Samples for Carbohydrate Analysis

Carbohydrate contamination. A highly purified protein that gives only a single band by SDS-PAGE analysis may still be unsuitable for carbohydrate analysis. Glucose is only very rarely found on sugar chains of proteins. On the other hand, it is almost always a readily available contaminant. DEAE- and CM-cellulose, agarose, Sephadex, Kimwipes, sucrose, glycerol, starch powder used on latex gloves, and glucoside-based detergents are all carbohydrates and can make a relatively large contribution to the sugar content of a sample, depending upon exactly how the material is analyzed. If proteins are purified on sucrose gradients, extracted with glucose-based detergents, or normally stored in glycerol, these sugar contaminants must first be completely removed by

dialysis. Acrylamide-based gels (e.g., Bio-Gel series) are preferred over the dextran-type gels (Sephadex) for gel-filtration purification because the latter can contribute fragments of beads as soluble but nondialyzable material (UNIT 10.9). As a final precaution, protein-containing solutions should be filtered through a 0.2- μ m filter to remove any particles that may have come from ion-exchange or affinity columns. If glucose is found by compositional analysis even after taking these precautions, it is still most likely to be a contaminant and not a component of the glycoprotein.

Lectin-affinity chromatography. The most frequently used specific purification procedure for glycoproteins is lectin-affinity chromatography. Dozens of plant lectins have been identified and in many cases their sugar specificity is known. This sugar specificity is the basis for the separation and structural analysis of many individual oligosaccharides and glycopeptides described in this chapter, and for the identification of lectin-binding glycoproteins in UNIT 17.7. Glycoproteins often contain multiple chains of a given sugar type and multiple interactions with the immobilized lectin may make it difficult to elute some of the proteins from the lectin column. The presence of multiple low-affinity sugar chains can still give substantial interaction with a lectin-containing column even though the individual sugar chains may themselves have a low affinity. Thus, one should be careful not to interpret too much about sugar chain structure from a protein's binding to a lectin column. This by no means belittles their usefulness for glycoprotein purification; lectin-affinity chromatography may be used to partially purify glycoproteins and to provide rather limited qualitative information about the nature of their carbohydrate components. Other preparative purification steps, such as ammonium salt precipitation (UNIT 11.3) and gel-filtration and ion-exchange chromatographies (UNITS 10.9 & 10.10), are usually done prior to using lectins, but this is not necessary.

Many different lectins are commercially available in an immobilized form suitable for glycoprotein purification, but some can be rather expensive. Con A-Sepharose is the most commonly used lectin for glycoprotein purification. It is relatively inexpensive, it is stable, and it can bind to many different glycans. Bound proteins can be eluted with α -methyl-D-mannoside (α MM). WGA is the next most popular lectin for the same reasons although it is somewhat more expensive.

Other notable lectins used to purify glycoproteins include *Ricinus communis* lectin (RCA I) for proteins that carry Gal-terminated sugar chains, and pea or lentil lectins for those N-linked oligosaccharides that have a fucose residue in the chitobiosyl core region. The utility of these lectins for characterizing glycopeptides is described in *UNIT 17.7*. Limited availability, higher price, limited track records, and, in some cases, the requirement for exotic sugars to elute bound proteins make some of these lectins less appropriate for routine use in glycoprotein purification. However, the principles of Con A–Sepharose and WGA-agarose lectin-affinity chromatography should apply to these other lectins as well. A broad range of immobilized lectins is available from E-Y Laboratories, Pharmacia Biotech, Vector Labs, and Sigma (*APPENDIX 4*).

Examples of lectin-affinity chromatography purifications. The basic and alternate protocols presented below describe the use of lectins for preparative glycoprotein purification. Con A–Sepharose and WGA-agarose were chosen for convenience and availability. The support protocol describes a small-scale pilot procedure to test for lectin binding and to determine elution conditions. There are many variations on the basic procedure in the literature, but all use the same principles: bind the protein to immobilized lectin through its sugar chain, wash away unbound protein, and elute bound protein with a simple sugar that resembles the sugar ligand of the bound protein. Because many proteins have sugar chains that can bind to a specific lectin, this procedure seldom yields a pure protein.

CON A–SEPHAROSE AFFINITY CHROMATOGRAPHY

Con A–Sepharose chromatography is used to partially purify glycoproteins that contain terminal mannose or glucose residues; the steps presented below are typical conditions for this type of chromatography, although a variety of approaches have been utilized. In this protocol, bound glycoproteins are eluted with α -methyl-D-mannoside (α MM) after the column is first washed to remove unbound and weakly bound proteins. Before proceeding, it is advisable to conduct a pilot study to test the protein of interest for lectin binding and elution conditions (see support protocol).

Materials

- 10 mg/ml Con A–Sepharose (Pharmacia Biotech or Sigma)
- Column buffer
- 0.5 M α MM in column buffer
- Protein sample in column buffer
- Glass wool
- 1.5 \times 30–cm glass or disposable chromatographic column
- Additional reagents and equipment for degassing solutions (*UNIT 10.12*), phenol–sulfuric acid assay for sugars (*UNIT 17.9*) and specific assay for detecting the protein of interest

NOTE: This procedure should be carried out at room temperature if the protein to be isolated will tolerate this condition. If not, carry it out in a cold room, and prechill all solutions to maintain temperature.

1. Gently resuspend 50 ml settled Con A–Sepharose (10 mg lectin/ml packed resin) in 50 ml column buffer to make a slurry. Degas the slurry.

This volume of lectin beads should be sufficient to bind ~100 mg glycoproteins. If a 1.5 \times 30–cm column is not available, either the dimensions of the column or the total amount of resin can be modified. A 1.0 \times 30–cm column will bind ~50 mg total glycoprotein. If the amount of glycoprotein in the sample is considerably less than 50 to 100 mg, decrease the volume of the column accordingly. The ratio of input protein to lectin does not seem to matter as long as the column is not overloaded.

BASIC PROTOCOL

Preparation and Analysis of Glycoconjugates

17.1.5

2. Pack a glass wool plug over a scintered glass or polypropylene frit at the bottom of the column and pour the degassed slurry into the column.

The glass wool is important to prevent the beads from clogging the frit and slowing the flow rate.

3. Continue packing the column until the desired level is reached; for a 50-ml volume this is ~28 cm. Wash the gel with 2 to 3 column volumes of column buffer to remove any loosely bound or degraded Con A.
4. Wash with 2 to 3 column volumes of 0.5 M of α MM in column buffer or the highest concentration of α MM that will be used.
5. Wash the column with >5 column volumes of column buffer *without* α MM to reequilibrate.

It is important to prewash the column with the eluting sugar and then to reequilibrate the column to remove any materials which might have previously bound to the column. Check the completeness of washing by the phenol-sulfuric acid assay (UNIT 17.9). Estimate the amount of residual sugar using α MM as a standard. An acceptable level is <0.1 mM or ~20 μ g/ml.

6. Slowly load the protein sample on the column to permit binding without disturbing the surface. A flow rate of ~1 ml/min (~0.5 ml/min per cm² of area) is desirable.
7. Wash column with column buffer and monitor the flowthrough and subsequent wash fractions by measuring the A_{280} until it approaches baseline value.

If the A_{280} does not rapidly return to baseline, it may indicate weak interaction of some proteins with the lectin, or overloading of the column.

8. Assay flowthrough and wash fractions for the presence of the protein of interest.

To check the lectin-binding properties, see support protocol below. To determine whether the column was overloaded, add a small amount of the flowthrough (1%) to a small volume of beads as described in the support protocol. If the sample is bound to the beads as determined by the appropriate assay, overloading of the original column is indicated. In this case, reapply the flowthrough with excess glycoprotein to the larger column after the bound proteins of the first run are eluted with α MM and the column is reequilibrated.

9. Elute the column with 0.5 M α MM in column buffer and monitor fractions for A_{280} and activity. Pool peak-activity fractions.

Very broad peaks may result during elution with the sugar if the protein dissociates very slowly. When this happens, sample recovery can be improved by filling the column with 0.5 M α MM in column buffer and allowing the column to stand for a few hours. This allows dissociation of the bound material; when the flow is started again, it should give a very sharp peak. Other possible remedies include warming a cold column to room temperature in the presence of α MM, increasing the concentration of α MM to 1 M, or increasing the NaCl concentration to 1 M. A combination of these may be needed to elute a tightly bound protein.

10. Regenerate the column by washing it with 10 vol column buffer or until the α MM concentration is <20 μ g/ml.

Column may now be used for another run or stored indefinitely at 4°C in column buffer containing 0.02% (w/v) NaN₃. Reequilibrate in column buffer before using again.

PILOT STUDY TO DETERMINE LECTIN BINDING AND ELUTION CONDITIONS

If the target protein can be detected easily, it is worthwhile to test a small sample to establish binding and elution conditions before applying the entire sample to a large column. The easiest approach is to mix a small amount of sample with a measured amount of Con A–Sepharose beads in a series of microcentrifuge tubes. If the material in the sample binds, it can then be eluted from the washed beads with various amounts of competing α MM to determine when the activity is eluted into the supernatant. This shows whether binding occurs and what concentration of competing sugar (or other conditions) will be required to elute it from the column. An abbreviated version of this procedure can be done with a single sample to determine if the protein binds and if it can be eluted with only one concentration (0.75 M) of α MM. It is important to keep the temperature constant, because changes can affect binding and dissociation of the ligands.

Additional Materials

Sepharose 4B (Pharmacia Biotech) or other beaded gel to fill space in the tubes
 α MM: 0, 0.1, 0.2, 0.4, 0.8, 1.0, and 1.5 M concentrations, in column buffer

1. Prepare a 50% slurry of Con A–Sepharose in column buffer. Cut ~2 to 3 mm off the small end of a 1000- μ l pipet tip. Resuspend the slurry immediately before each pipetting, and use the truncated tip to dispense 200 μ l of slurry into a series of seven microcentrifuge tubes. To one of these tubes—the no-lectin control—add an equal volume of 50% slurry of Sepharose 4B. Allow gel to settle to see if the dispensing was reasonably accurate.

Gel material is needed in the controls to account for the volume changes caused by adding the gel. Other nonionic beaded gels, such as Sephadex (Pharmacia Biotech) or Bio-Gel (Bio-Rad) may be substituted.

2. Add the protein sample in 50 μ l column buffer and allow 15 min for binding (shake the tube occasionally). Microcentrifuge the gel beads 1 min at 1000 \times g to sediment. Remove the supernatant and save for analysis in step 6.
3. Use an appropriate assay for the protein of interest to determine if it binds to Con A–Sepharose and not to control beads.
4. Wash the gel beads with three 1.4-ml washes of column buffer; after each wash microcentrifuge the beads using conditions in step 2.
5. Resuspend gel beads in 200 μ l of buffer containing 0, 0.1, 0.2, 0.4, 0.8, 1.0, or 1.5 M α -MM for each of the seven tubes containing Con A–Sepharose. Add the same volume of column buffer or 1.5 M α MM in column buffer to the control.
6. Incubate 15 min and centrifuge the gel beads as in step 2. Analyze the supernatants using the appropriate detection assay to determine if the protein has been eluted and at what concentration of competing sugar.

Increasing the concentration of α MM should elute the target protein from the Con A–Sepharose. If it is not eluted, longer incubation times (10 hr) in the presence of α MM or higher temperature may be needed to elute the protein from the beads. Although the entire procedure can be done in the cold, binding is tighter at low temperature, and this may make elution more difficult. Once the conditions are established for a particular sample or a lectin, scaling up the preparation for the basic protocol should proceed smoothly.

WHEAT GERM AGGLUTININ (WGA)–AGAROSE AFFINITY CHROMATOGRAPHY

WGA-agarose chromatography is used to purify proteins that contain terminal *N*-acetylglucosamine (GlcNAc) or sialic acid residues. A protein sample is applied to the column, the column is washed to remove unbound and weakly bound proteins, and bound glycoproteins are eluted with GlcNAc. Most of the advice for purification of glycoproteins on Con A columns applies to similar columns containing immobilized WGA. It is best to test the binding of the target protein to the lectin using a pilot study like that for Con A (support protocol) before running a large column, substituting GlcNAc for the α MM.

Additional Materials

- 5 mg/ml wheat germ agglutinin (WGA)–agarose (E-Y Laboratories, Pharmacia Biotech, Sigma)
- Phosphate-buffered saline (PBS; APPENDIX 2)
- 0.1 M *N*-acetylglucosamine (GlcNAc) in PBS
- Protein sample in PBS
- 1.0 × 10–cm glass or disposable column

1. Resuspend 1 vol WGA-agarose in 2 to 3 vol PBS to make a slurry.

One milliliter of gel should be enough to bind 1 to 2 mg glycoprotein. To be safe, assume ~5% to 10% of total cell lysate can bind. If target protein requires detergent to solubilize it, WGA is still active in 25 mM Tris-Cl (pH 7.5)/1% Lubrol PX/0.1% sodium deoxycholate. Triton X-100 or NP-40 (~2% to 3% v/v final) may also be used.

2. Place a glass wool plug over the frit at the bottom of a column whose length is 10 times its diameter. Pour the slurry into the column. Wash the column with ~2 column volumes of PBS to remove any unbound or degraded WGA, then wash with 1 column volume PBS containing 0.1 M GlcNAc. Finally, wash with 5 column volumes PBS.

The length of the column should be ~10 times its diameter so that weakly bound proteins can be eluted by prolonged PBS wash in the absence of any competing sugar. This will separate them from proteins that do not interact at all, and differs from using Con A where continued buffer wash in the absence of α MM does not usually elute weakly bound protein. The choice of these column dimensions for WGA and for most other lectins is based on the successful fractionation of individual glycopeptides by serial lectin-affinity chromatography.

3. Add protein sample in ~0.1 column volumes at a flow rate of ~2 ml/cm² per hour, and wash the column with PBS until the A_{280} returns to baseline.

With a small sample volume it is easier to detect proteins that do not bind to the column at all and elute promptly as a sharp peak in the flowthrough. Proteins that bind weakly to the column are gradually washed off by continued elution with PBS alone; these proteins would “smear” as a trailing peak. If a preliminary test run of the target protein shows that it binds strongly to the lectin, and if the column is not overloaded, the sample can be added to the column in a larger volume. The flow rate can be varied. Sometimes the protein of interest can be absorbed to the lectin simply by gently mixing the two components for several hours before pouring the loaded gel into the column. The protein can then be eluted after washing as determined by the pilot study.

4. Elute the column with 2 to 3 column volumes of 0.1 M GlcNAc in PBS, and assay the fractions for the protein of interest.

Generally, 0.1 M GlcNAc is sufficient to elute the protein. If this does not work, try a higher GlcNAc concentration (e.g., up to 0.25 M), temporarily turn off the column, or raise the temperature to elute the sample, as with Con A. Also, adding 0.5 M NaCl to the GlcNAc-containing elution buffer can sometimes improve recovery. A pilot study to determine elution conditions, similar to that described for Con A (support protocol), is highly recommended.

5. Regenerate the column by washing with ≥ 10 column volumes of PBS or until reducing sugar content is below 20 $\mu\text{g/ml}$. The column can be reused immediately or stored at 4°C in column buffer containing 0.02% (w/v) NaN_3 .

REAGENTS AND SOLUTIONS

Column buffer

0.01 M Tris-Cl, pH 7.5

0.15 M NaCl

1 mM CaCl_2

1 mM MnCl_2

Store indefinitely at room temperature

Prepare the MnCl_2 within a day or two of use and add it to the buffer only after the pH has been adjusted. The final concentration of metal ions can be decreased 10-fold or more if needed.

If detergents are required to solubilize the protein, Triton X-100 or Nonidet P-40 at $\leq 2\%$ have a negligible effect on lectins. Do not use glucoside-based detergents because they may interfere with binding.

COMMENTARY

Background Information

See the introductions to Chapter 17 and to this unit for discussions concerning properties of glycoproteins that relate to their purification.

Critical Parameters

Like most protein purification procedures, successful lectin-affinity chromatography is empirical, so few parameters are “written in stone.” However, in using lectins, it is important to first determine the binding and elution conditions in a pilot study. Blindly loading a sample onto a column and then trying different elution conditions may give very low yields, either because the protein may become inactivated or because it cannot be eluted under those particular conditions. It is usually easy to bind the protein to the column, but elution may be more difficult. There are a number of ways to elute tightly bound protein: increase the concentration of sugar in the buffer used for elution, increase temperature, increase the salt (NaCl) concentration, or stop the column flow for an extended period during the elution with sugar. A combination of these may be needed.

These modifications might cause a protein to elute differently from the same column. Thus, it may be useful to perform two consecutive runs with different elution conditions to separate the protein of interest from different contaminants. Ketcham and Kornfeld (1992) describe the purification of a rare glycosyl transferase using WGA, whereby the lectin-affinity chromatography step alone yielded a 160-fold purification. Another important point

to remember is that environmental carbohydrate contamination may contribute significantly to the total quantity of carbohydrate in a preparation. Follow the suggestions given within the protocols to minimize this effect.

Anticipated Results

Lectins can give 2- to >100 -fold purifications of glycoproteins and nearly quantitative yields, depending upon the source of the sample and elution conditions. If the protein of interest is reasonably stable at high salt concentrations and room temperature, it is likely that specific elution conditions can be found that will produce good yields.

Time Considerations

Initial determination of the binding and elution conditions can easily be done in a day, assuming that the detection assay for the target protein is simple and fast. The amount of time needed to conduct a single run once the elution conditions are known can vary depending upon the size of the sample, the column, and whether elution requires long-term incubation in the presence of the sugar. Even assuming the extremes of sample size and elution conditions, a single run should be completed in 2 to 3 days.

Literature Cited

Gibbons, R.A. 1972. Physico-chemical methods for the determination of the purity, molecular size and shape of glycoproteins. *In* Glycoproteins: Their Composition, Structure and Function (A. Gottschalk, ed.) pp. 31-140. Elsevier Science Publishing, New York.

Ketcham, C.M. and Kornfeld, S. 1992. Purification of UDP-*N*-acetylglucosamine: Glycoprotein *N*-acetylglucosamine-1-phosphotransferase from *Acanthamoeba castellanii* and identification of a subunit of the enzyme. *J. Biol. Chem.* 267:11645-11653.

Key References

Beeley, J.G. 1984. Glycoprotein and Proteoglycan Techniques. *In* Laboratory Techniques in Biochemistry and Molecular Biology, pp. 29-99. Elsevier Science Publishing, New York.

Good discussions about general properties of glycoconjugates.

Dulaney, J.T. 1979. Binding interactions of glycoproteins with lectins. *Mol. Cell. Biochem.* 21:43-62.

Lists many references and conditions used for lectins in protein purification.

Montreuil, J., Bouquelet, S., Debary, H., Fournat, B., Spik, G., and Strecker, G. 1986. Glycoproteins. *In* Carbohydrate Analysis: A Practical Approach (M.F. Chaplin and J.F. Kennedy, eds.) pp. 166-173. IRL Press, Washington, D.C.

Lists several conditions for selected lectin-affinity purifications of proteins.

Contributed by Hudson H. Freeze
La Jolla Cancer Research Foundation
La Jolla, California

Special Considerations for Proteoglycans and Glycosaminoglycans and Their Purification

UNIT 17.2

Proteoglycans (PG) contain long linear glycosaminoglycan (GAG) chains that consist of repeats of disaccharides and therefore differ from the short branched oligosaccharides found on glycoproteins and glycolipids. The principal GAGs of animal tissue PGs include chondroitin sulfate, dermatan sulfate, heparan sulfate, or keratan sulfate. Heparin, the commonly used anticoagulant, is a highly modified form of heparan sulfate; it is made solely by connective tissue mast cells. Repeating disaccharide units composed of glucuronic acid (GIUA) or iduronic acid (IdUA) and an *N*-acetylhexosamine residue (GlcNAc or GalNAc) make up the backbone of most GAGs. Keratan sulfate, however, contains disaccharides of Gal and GlcNAc. All GAGs, except hyaluronan, contain a large number of sulfate residues that, together with the uronic acids, impart a large negative charge to the chains. The high density of negative charges on GAGs distinguishes these polysaccharides from the oligosaccharides found on glycoproteins and glycolipids, and is the basis for their physical separation. This property facilitates GAG and PG purification by anion-exchange chromatography and precipitation with cetylpyridinium chloride (CPC) and ethanol.

The synthesis of GAGs occurs while they are attached to core proteins; free chains can accumulate because of proteolysis and endoglycolytic cleavage of the PG. For example, heparin chains are initiated and elongated on a core protein called serglycin. Soon after assembly of the heparin proteoglycan, proteases degrade the serglycin core, and an endoglucuronidase generates short heparin oligosaccharides that sort to secretory granules. Most cells constitutively secrete specific PGs into extracellular fluids and matrix. They also express several membrane-intercalated PGs that are shed from the cell surface by proteolysis. Fragmentation of matrix proteoglycans often occurs and gives rise to GAG chains attached to short peptides. Hyaluronan, a nonsulfated GAG, does not assemble while linked to protein, but associates noncovalently with PGs that contain specific binding domains.

In general, the problems encountered in purifying PGs are those encountered in purification of other protein glycoconjugates. The purification protocol must ensure efficient extraction of tissue or cells, high recovery in separation and purification steps, and negligible degradation. This unit describes two simple methods for extracting PGs and GAGs—one using high-salt/detergent extraction (first basic protocol) and one using alkali treatment (first alternate protocol). These are complemented by two techniques for concentrating samples that also yield a significant degree of purification—anion exchange chromatography (second basic protocol) and precipitation with CPC (second alternate protocol).

HIGH-SALT/DETERGENT EXTRACTION OF PROTEOGLYCANS AND GLYCOSAMINOGLYCANS

**BASIC
PROTOCOL**

Efficient extraction of PGs requires the use of chaotropic salts in an appropriate buffer. Detergent is included to ensure extraction of membrane-intercalated PGs, to prevent aggregation, and to decrease adsorption to glassware. In this protocol, samples (tissue, conditioned medium, biological fluid, or cultured cells) are treated with a high-salt/detergent (guanidine·HCl/Zwittergent 3-12) extraction buffer to extract PGs as well as free GAG chains and GAGs with a short peptide chain. Extracted material can be concentrated and partially purified by anion-exchange chromatography or CPC and ethanol precipitation (second basic and alternate protocols, respectively). Studies involving proteoglycans are conducted at 4°C in the presence of protease inhibitors to minimize degradation of core proteins; work with glycosaminoglycans can be done at room temperature.

**Preparation
and Analysis of
Glycoconjugates**

Contributed by Jeffrey D. Esko

Current Protocols in Molecular Biology (1993) 17.2.1-17.2.9

Copyright © 2000 by John Wiley & Sons, Inc.

17.2.1

Supplement 22

Materials

Tissue sample, conditioned medium, biological fluids, *or* cultured cells
Guanidine-HCl/Zwittergent 3-12 extraction buffer *or* Triton X-100 extraction buffer, at 4°C

200× protease inhibitor stock solutions

Centrifuge and rotor (e.g., Sorvall SS-34) *or* microcentrifuge, at 4°C

Additional reagents and equipment for measurement of uronic acids (UNIT 17.9) and metabolic radiolabeling (UNIT 17.4)

1a. *To extract from tissue samples:* Chill, mince, and mix tissue with 5 to 10 vol of 4°C guanidine-HCl/Zwittergent 3-12 extraction buffer per gram of tissue (wet weight). Stir the sample overnight at 4°C.

1b. *To extract from conditioned medium or biological fluids:* Chill samples to 4°C and dissolve solid guanidine-HCl to 4 M (final), Zwittergent 3-12 detergent to 0.2% w/v (final) and EDTA to 10 mM (final). Add protease inhibitors from stock solutions to achieve final concentrations of 10 mM NEM, 1 mM PMSF, 1 µg/ml pepstatin A, and 0.5 µg/ml leupeptin.

The reagents are added in this way to minimize the increase in volume.

1c. *To extract from cultured cells:* Add ~1 ml of guanidine-HCl/Zwittergent 3-12 or extraction buffer per 10⁷ cells (in a pellet), or 3 to 5 ml per 75 cm² of cell monolayer. Cell pellets and monolayers will dissolve within 1 hr at 4°C.

The PGs of cultured cells, conditioned medium, and biological fluids will also solubilize in Triton X-100 extraction buffer (see reagents and solutions), but the efficiency of extraction should be tested (critical parameters). The use of Triton X-100 extraction buffer eliminates the need to remove guanidine-HCl before further purification.

2. Centrifuge 20 min at 12,000 × g (10,000 rpm in a Sorvall SS-34 rotor), 4°C, or microcentrifuge 10 min at maximum speed, 4°C, to remove insoluble residue. Collect the supernatant. If necessary, reextract residue with fresh extraction buffer, centrifuge as before, and combine the supernatants.

For conditioned medium, biological fluids, and cultured cells, the initial treatment with extraction buffer is usually sufficient. Tissue samples may need to be reextracted (see critical parameters).

3. Purify the GAGs by anion-exchange chromatography or by treating with CPC and ethanol (second basic and alternate protocols, respectively).

4. Quantitate the amount of GAG present by measuring the uronic acid content chemically, or by measuring the amount of radioactive material present if the starting material was radiolabeled with ³⁵S₄ or [³H]GlcN.

The measurement of uronic acid is also described in Bitter and Muir (1962) and Filisetti-Cozzi and Carpita (1991).

ALTERNATE PROTOCOL

Special Considerations for Proteoglycans and Glycosamino- glycans and Their Purification

ALKALI EXTRACTION OF PROTEOGLYCANS AND GLYCOSAMINOGLYCANS

If the primary objective is simply to determine whether a sample contains PGs or GAGs, an aliquot can be extracted with alkali. Alkali extraction results in some elimination of chains from core proteins and possible cleavage of the core, so this technique is unacceptable for isolating intact PGs. Alkali-extracted samples are usually subjected to exhaustive proteolysis or complete β-elimination (UNIT 17.15) to liberate GAG chains. The GAGs are then purified by anion-exchange chromatography or precipitation with CPC and ethanol

17.2.2

(second basic and alternate protocols, respectively) and quantitated chemically or radiochemically. This protocol is suitable for cultured cells, conditioned medium, or biological fluids; tissue samples should be extracted using the high-salt/detergent procedure (first basic protocol).

Additional Materials

0.1 N or 10 N sodium hydroxide (NaOH; *APPENDIX 2*)
10 N acetic acid

Additional reagents and equipment for β -elimination (*UNIT 17.15*)

1. Solubilize samples in 0.1 N NaOH at room temperature. Incubate 10 min at 25°C.
Generally 0.5 ml of 0.1 N NaOH is sufficient for 10^7 cells as a pellet or 2 ml per 75 cm^2 of cell monolayer. Cells suspended in culture medium can be extracted by adding 10 μl of 10 N NaOH per milliliter medium.
2. Neutralize sample with 10 μl of 10 N acetic acid per milliliter sample. Remove any precipitate that forms by centrifuging 20 min at $12,000 \times g$ (10,000 rpm in a Sorvall SS-34 rotor), 4°C, or microcentrifuging 10 min at maximum speed, 4°C.
3. Treat neutralized samples with a nonspecific protease, such as Pronase or proteinase K or subject samples to β -elimination conditions to liberate individual GAG chains.
4. Purify the GAGs by anion-exchange chromatography or by treating with CPC and ethanol (second basic and alternate protocols, respectively).
5. Quantitate the amount of GAG present by measuring the uronic acid content chemically or by measuring the amount of radioactive material present if the starting material was radiolabeled with $^{35}\text{SO}_4$ or $[^3\text{H}]\text{GlcN}$.

ANION-EXCHANGE CHROMATOGRAPHY OF PROTEOGLYCANS AND GLYCOSAMINOGLYCANS

BASIC PROTOCOL

PGs and GAGs can be purified by anion-exchange chromatography on weak anion-exchange resins (*UNIT 10.10*). This step allows collection of very small amounts of material from a large volume of solution, thus concentrating the sample. Anion-exchange chromatography allows separation of different types of GAGs, but the low-salt elution conditions described here will yield a mixture of all GAGs and PGs regardless of composition.

Materials

PG-GAG extract (first basic or alternate protocol)
Urea/Zwittergent 3-12 buffer
20 mg/ml chondroitin sulfate or heparin in 20 mM Tris·Cl, pH 7.0
DEAE-Sephacel (Pharmacia Biotech)
0.2 M NaCl/50 mM sodium acetate, pH 6.0
DEAE wash buffer
DEAE elution buffer
Sephadex G-25 (Pharmacia Biotech)
10% ethanol
20 mM Tris·Cl, pH 7.4 (*APPENDIX 2*)
Centrifuge and rotor (e.g., SS-34) or microcentrifuge, at 4°C

Additional reagents and equipment for dialysis (*APPENDIX 3*) and gel filtration and ion-exchange chromatography (*UNITS 10.9 & 10.10*)

1. If the extract contains guanidine-HCl/Zwittergent 3-12, dialyze it against urea/Zwittergent 3-12 buffer. Remove any precipitate that forms during the exchange of urea

Preparation and Analysis of Glycoconjugates

17.2.3

for guanidine by centrifuging 20 min at $12,000 \times g$ (10,000 rpm in a Sorvall SS-34 rotor), 4°C, or microcentrifuging 10 min at maximum speed, 4°C.

Gel-filtration chromatography or ultrafiltration also can be used to change buffers. Extracts prepared in Triton X-100 extraction buffer (see critical parameters) can be chromatographed without further manipulation. Protease inhibitors can be included as long as they do not contribute significantly to the ionic charge of the sample solution.

2. Add 1 to 2 mg chondroitin sulfate as carrier if the total GAG content of the sample is $<100 \mu\text{g}$ and if the addition will not interfere with subsequent analysis.

If the amount of PG or GAG is to be quantitated by measuring uronic acid, carrier GAG must be omitted. Chondroitin sulfate is preferred because it is less expensive than heparin or heparan sulfate. Heparin can inhibit chondroitinases that are used to establish the presence and composition of chondroitin sulfate chains.

3. Prepare the DEAE-Sepharose column. Wash the resin three times with 0.2 M NaCl/50 mM sodium acetate, pH 6.0. Pour a small column by placing 0.5 to 1 ml resin in a polypropylene pipet tip (P-1000) plugged with a small piece of glass wool. Cut the tip slightly (1 to 2 mm) to widen the opening and improve flow rate. Remove the bottom one-third of a second tip with a razor blade and insert it into the first tip. This provides ~ 2.5 ml of additional head space above the resin bed.

The upper tip will accommodate a disposable 10-ml pipet for adding solutions. Larger column beds can be poured in commercial columns.

4. Equilibrate the column with ~ 5 column volumes of DEAE wash buffer.
5. Apply the sample to the equilibrated column. Elute contaminants with 20 to 30 column volumes of DEAE wash buffer. Elute the PGs and GAGs with 5 column volumes of DEAE elution buffer and collect as one fraction.

Avoid exceeding the binding capacity of the resin (~ 5 mg of chondroitin sulfate, heparan sulfate or heparin per milliliter of resin).

Buffers prepared without urea, detergent, and protease inhibitors can be substituted if these reagents interfere with subsequent analysis (e.g., enzymatic digestions by polysaccharide lyases). Chromatography of free GAGs does not require urea or protease inhibitors, but detergent should be included to improve recovery. Fast flow rates can be used.

6. Desalt sample using a Sephadex G-25 gel-filtration column run in 10% ethanol.

If the sample volume is ≤ 2.5 ml, disposable PD-10 columns (Pharmacia Biotech) can be used.

7. Lyophilize the sample and rehydrate it in 20 mM Tris·Cl, pH 7.4 or other suitable buffer. Store at 4°C until further use.

CPC/ETHANOL PRECIPITATION OF PROTEOGLYCANS AND GLYCOSAMINOGLYCANS

ALTERNATE PROTOCOL

This protocol describes the concentration and partial purification of extracted PGs and GAGs with the lipophilic cation cetylpyridinium chloride (CPC). The CPC binds to charged groups along the GAGs, rendering them insoluble. This method of precipitation should be used only when the concentration of PGs and GAGs in the original sample is large enough to avoid use of carrier or when the addition of carrier is of no consequence—e.g., when the sample is extracted from cells biosynthetically labeled with $^{35}\text{SO}_4$ or [^3H]GlcN (UNIT 17.4).

Additional Materials

- 20 mM Tris·Cl/0.2 M NaCl, pH 7.4
- 5% (w/v) cetylpyridinium chloride hydrate (CPC; Aldrich), in water
- 0.5 M sodium acetate (APPENDIX 2)
- 95% ethanol, 4°C
- 0.5 M sodium acetate in 10% (v/v) ethanol (sodium acetate/ethanol)
- Additional reagents and equipment for ethanol precipitation (UNIT 2.1)

1. If the extract contains guanidine·HCl/Zwittergent 3-12 or high salt (≥ 0.2 M), dialyze it against 20 mM Tris·Cl/0.2 M NaCl (pH 7.4) or use gel-filtration chromatography to exchange the buffers.

Electrolytes and detergents compete with CPC precipitation of GAGs. All GAGs will precipitate if the salt concentration is ≤ 0.2 M. Protease inhibitors can be included if they do not contribute significantly to the ionic charge of the sample.

2. Add chondroitin sulfate as carrier so that the final GAG content of the sample is ≥ 0.1 mg/ml.

Precipitation with CPC requires that the GAG concentration be ≥ 0.1 mg/ml for high recovery. If carrier is added, it is not possible to quantitate GAGs by measuring uronic acid content.

3. Add 1 ml of 5% CPC/4 ml sample. Incubate 1 hr at 37°C.
4. Centrifuge 20 min at $12,000 \times g$ (10,000 rpm in a Sorvall SS-34 rotor), or microcentrifuge 10 min at maximum speed, room temperature. Remove the supernatant and resuspend the pellet in a small volume of 0.5 M sodium acetate so that the final GAG concentration is ~ 10 mg/ml.

Centrifugation must be done at room temperature or the CPC will precipitate. If necessary, briefly incubate the sample in a boiling water bath to dissolve the pellet.

5. Add 4 vol of ice-cold 95% ethanol, mix, and incubate ≥ 2 hr at 4°C.

Precipitation with ethanol requires salt and adequate mass for high recovery. The samples should contain 0.1 M to 0.5 M salt and 0.1 mg/ml GAG. If adequate material is not present, add additional chondroitin sulfate as carrier.

6. Centrifuge 20 min at $12,000 \times g$ (10,000 rpm in a Sorvall SS-34 rotor), or microcentrifuge 10 min at maximum speed, to pellet the white precipitate that forms. Remove and discard supernatant.
7. Dissolve pellet in a small volume of 0.5 M sodium acetate/10% ethanol such that the final concentration is ~ 1 mg/ml. Repeat the ethanol precipitation.
8. Dry precipitate under vacuum and resuspend in 20 mM Tris·Cl, pH 7.4 or other suitable buffer. Store at 4°C until further use.

Preparation and Analysis of Glycoconjugates

17.2.5

REAGENTS AND SOLUTIONS

10 N acetic acid

Dilute 65.4 ml glacial acetic acid (15.3 M, ACS reagent grade) with water to a total volume of 100 ml. Store indefinitely at room temperature.

DEAE elution buffer

Add 58.5 g NaCl (1 M final) to 1 liter urea/Zwittergent 3-12 buffer. Store at 4°C for several months.

DEAE wash buffer

Add 11.7 g NaCl (0.2 M final) to 1 liter urea/Zwittergent 3-12 buffer. Store at 4°C for several months.

Guanidine·HCl/Zwittergent 3-12 extraction buffer

382.1 g guanidine·HCl (4 M)
2 g Zwittergent 3-12 (0.2% w/v)
4.1 g sodium acetate (50 mM)
3.8 g EDTA (10 mM)
5 ml NEM stock solution (10 mM)
5 ml PMSF stock solution (1 mM)
5 ml pepstatin A stock solution (1 µg/ml)
5 ml leupeptin stock solution (0.5 µg/ml)

Dissolve guanidine·HCl Zwittergent 3-12, sodium acetate, and EDTA in 800 ml water. Add protease inhibitors just before use. Adjust pH to 6.0 and bring volume to 1 liter.

200× protease inhibitor stock solutions

Leupeptin (0.1 mg/ml). Dissolve 5 mg leupeptin (Ac-Leu-Leu-Arg-al hemisulfate, MW 475.6) in water and adjust volume to 50 ml. Store in aliquots at -20°C.

Leupeptin inhibits cathepsin B.

N-ethylmaleimide (NEM; 2 M). Dissolve 25 g NEM in ethanol and adjust volume to 100 ml. Store in aliquots at -20°C.

NEM inhibits thiol proteases and prevents nonspecific disulfide exchange.

Pepstatin A (0.2 mg/ml). Dissolve 5 mg pepstatin A (Isovaleryl-Val-Val-Sta-Ala-Sta) in ethanol and adjust volume to 25 ml. Store in aliquots at -20°C.

Pepstatin A inhibits cathepsin D.

Phenylmethylsulfonyl fluoride (PMSF; 0.2 M). Dissolve 3.5 g PMSF in ethanol and adjust volume to 100 ml. Store in aliquots at -20°C.

PMSF inhibits serine proteases.

0.5 M sodium acetate/10% (v/v) ethanol.

Dissolve 4.1 g sodium acetate in water and adjust volume to 90 ml. Add 10 ml ethanol.

Triton X-100 extraction buffer

5 g Triton X-100 (0.5% w/v)
8.77 g NaCl (0.15 M)
2.42 g Tris base (20 mM Tris·Cl)
3.8 g EDTA (10 mM)
5 ml NEM stock solution (10 mM)

continued

5 ml PMSF stock solution (1 mM)
5 ml pepstatin A stock solution (1 µg/ml)
5 ml leupeptin stock solution (0.5 µg/ml)

Dissolve Triton X-100, NaCl, Tris base, and EDTA in ~800 ml water. Add protease inhibitors just before use. Adjust pH to 7.4 and bring volume to 1 liter.

Urea/Zwittergent 3-12 buffer

240 g urea (ultrapure grade, 4 M)
2 g Zwittergent 3-12 (0.2% w/v)
4.1 g sodium acetate (50 mM)
3.8 g EDTA (10 mM)
5 ml NEM stock solution (10 mM)
5 ml PMSF stock solution (1 mM)
5 ml pepstatin A stock solution (1 µg/ml)
5 ml leupeptin stock solution (0.5 µg/ml)

Dissolve urea, Zwittergent 3-12, sodium acetate, and EDTA in ~800 ml water. Add protease inhibitors just before use. Adjust pH to 6.0 and bring volume to 1 liter.

COMMENTARY

Background Information

Proteoglycan (PG) extraction methods were first developed for studying PGs in cartilage. These methods use high salt and guanidine-HCl to disrupt ionic interactions and they are considered to be “dissociative” because their effectiveness depends on the dissociation of PGs from aggregates containing hyaluronan and matrix proteins (Sajdera and Hascall, 1969). PGs and glycosaminoglycans (GAG) present in other tissues, biological fluids, and cultured cell lines (Kjellén and Lindahl, 1991) differ significantly in structure and composition from those found in cartilage, but the original extraction methods remain effective. Many PGs associate with membranes through hydrophobic domains or glycosyl phosphatidylinositol anchors (UNIT 17.8). Membrane PGs require detergent to disrupt hydrophobic interactions and to produce a micellar dispersion. Current extraction and purification protocols employ both denaturants and detergents (Yanagishita et al., 1987; Hascall and Kimura, 1982).

Critical Parameters and Troubleshooting

Extraction efficiency. The guanidine-HCl/Zwittergent 3-12 extraction buffer should render the majority of macromolecules soluble. However, some tissues (e.g., aorta) may resist dispersion, leaving some PGs in the insoluble residue. Mechanical homogenization of samples in extraction buffer may improve recovery. To determine the overall efficiency of extraction, measure the release of PGs over time. Incubate aliquots of sample in extraction buffer

and at certain intervals, separate the PGs from other constituents by anion-exchange chromatography (second basic protocol) or precipitation with cetylpyridinium chloride (CPC) and ethanol (second alternate protocol). If sufficient material is present (≥ 50 µg of uronic acid), monitor the uronic acid content of a sample to measure PG and GAG release. Alternatively, if cultured cells are labeled biosynthetically with $^{35}\text{SO}_4$ or $[6\text{-}^3\text{H}]\text{GlcN}$ (UNIT 17.4), extraction of PGs from cultured cells is easily followed radiochemically. When the yield of uronic acid or radioactivity in the extract reaches a plateau, no additional material can be obtained under these conditions. Any insoluble residue should be dissolved in 0.1 N NaOH and residual GAGs measured. If more than 10% of the material remains in the residue, the extractions condition may be varied by using a different detergent (e.g., 1% Triton X-100 or 1% deoxycholate), by adding other salts (e.g., LiCl), by treating samples with collagenase, by altering the pH, or by adding a reductant (e.g., 2-mercaptoethanol). Kimura et al. (1981) showed that sequential addition of detergent and guanidine-HCl extracts more proteoglycan from chondrosarcoma than adding both agents simultaneously.

Recovery. As in any purification scheme, adequate starting material is required to ensure high recovery. Samples should ideally contain at least 0.1 mg of GAG as determined by uronic acid content. Because cell cultures tend to yield small amounts of material (0.1 to 1 µg of GAG uronic acid per 10^6 cells), carrier GAGs should be added. Commercial preparations of chon-

droitin sulfate are inexpensive and provide a suitable carrier for purifying PGs as well as GAGs. However, addition of carrier makes it impossible to monitor recovery by chemically measuring uronic acid, so carrier should only be added to samples labeled biosynthetically with $^{35}\text{SO}_4$ or $[^3\text{H}]\text{GlcN}$. Sufficient mass (≥ 0.1 mg/ml) is critical for efficient precipitation with CPC and ethanol.

Inclusion of protease inhibitors is critical to maintain the integrity of PG core proteins. The buffers used in these protocols contain leupeptin (to inhibit cathepsin B), NEM (to inhibit thiol proteases and prevent nonspecific disulfide exchange), Pepstatin A (to inhibit cathepsin D), PMSF (to inhibit serine proteases), and EDTA (to inhibit metalloproteases). In addition, samples should be chilled rapidly and processed immediately, and the pH should be maintained at pH 6.0 to diminish the activity of acidic lysosomal proteases and neutral tissue proteases.

If degradation of core protein occurs under the conditions described in the protocols (as detected by variable recovery), other inhibitors should be used. For example, benzamidine-HCl at 1 mM inhibits serine proteases; 6-aminohexanoic acid at 0.1 M inhibits cathepsin D; phosphoramidon at 0.5 mM inhibits collagenases, metalloendoproteinases, and stromelysin; soy bean trypsin inhibitor at 0.1 mg/ml inhibits trypsin-like proteases; and bestatin at 0.1 mM inhibits several aminopeptidases. Sodium dodecylsulfate (SDS) can also inhibit proteases because of its ability to denature proteins. However, SDS is difficult to remove; its incompatibility with anion-exchange chromatography and CPC precipitation complicates further purification steps. Nevertheless, SDS helps maintain solubility of aggregating proteins and should be considered as an additive in purification steps after CPC precipitation or anion-exchange chromatography.

Zwitterionic (e.g., Zwittergent 3-12) and nonionic detergents (e.g., Triton X-100) help solubilize membrane PGs and improve recovery by blocking nonspecific binding and aggregation. These detergents remain soluble both in the presence of high concentrations of salts and at low temperature, and are compatible with most purification methods.

Purity. Purification of PGs to chemical homogeneity (a single species of PG with no contaminants) presents a challenging problem. The procedures described in this unit should yield material enriched in PGs because few macromolecules exhibit the charge charac-

teristics imparted by the polyanionic GAG chains. Nucleic acids tend to copurify with PGs and GAGs, but nucleic acids and weakly charged or neutral oligosaccharides and polysaccharides are not precipitated efficiently by CPC. If necessary, samples can be treated with DNase and RNase (*UNIT 3.12*) to reduce the content of nucleic acids.

Some contaminants associate with PGs through ionic interactions with the GAG chains. Other purification methods should be tried, including density-gradient centrifugation in CsCl (Hascall and Kimura, 1982), gel-filtration chromatography in high salt or guanidine-HCl (*UNIT 10.9*), reversed-phase chromatography on octyl-Sepharose (Yanagishita et al., 1987), and affinity purification using monoclonal antibodies (*UNIT 10.11A*; Hascall and Kimura, 1982; Yanagishita et al., 1987). With the exception of affinity purification protocols that use immobilized proteins or antibodies, most of these methods are compatible with dissociative conditions.

Some protein contaminants may aggregate with core proteins through interchain disulfide bonds. Some matrix PGs may not extract efficiently if they are linked by disulfides to insoluble matrix proteins (Parthasarathy and Spiro, 1981). Reduction with 10 mM dithiothreitol at elevated temperature followed by alkylation (e.g., with iodoacetamide) should help dissociate these proteins.

The purification protocol described above does not separate free GAG chains from PGs. Free chains usually have different Stokes radii than PGs, and gel-filtration chromatography frequently separates GAGs from PGs. Shift of material into more included fractions after β -elimination or proteolysis confirms that PGs were present in the original sample. If sufficient material is present, colorimetric quantitation of proteins (*UNIT 10.1*) can also help distinguish PGs from GAGs, but the presence of contaminating proteins may complicate this analysis. Labeling PGs biosynthetically with radioactive amino acids and $^{35}\text{SO}_4$ or $[^3\text{H}]\text{GlcN}$ provides another way to monitor elution of PGs and GAGs.

Several excellent reviews describe detailed procedures for further characterization of intact PGs, their protein cores, and the GAG chains (Rodén et al., 1972; Hascall and Kimura, 1982; Yanagishita et al., 1987).

Anticipated Results

The high-salt/detergent and alkali extraction methods will reliably extract the majority

of PGs from most tissues and cells. Ion-exchange chromatography and precipitation with CPC and ethanol will yield a preparation of PGs or GAGs of adequate concentration and purity for compositional analysis.

Time Considerations

Tissues usually require overnight extraction, but extraction of cultured cells occurs very rapidly, usually within an hour. Complete dialysis requires several buffer changes over a period of 24 to 48 hr. Anion-exchange chromatography and CPC precipitation take ~1 day to complete. Thus, it should take 3 to 4 days to extract, partially purify, and concentrate PGs and GAGs from biological samples.

Literature Cited

- Bitter, T. and Muir, H.M. 1962. A modified uronic acid carbazole reaction. *Anal. Biochem.* 4:330-334.
- Filisetti-Cozzi, T.M. and Carpita, N.C. 1991. Measurement of uronic acids without interference from neutral sugars. *Anal. Biochem.* 197:157-162.
- Hascall, V.C. and Kimura, J.H. 1982. Proteoglycans: Isolation and characterization. *Methods Enzymol.* 82:769-800.
- Kimura, J.H., Caputo, C.B., and Hascall, V.C. 1981. The effect of cycloheximide on synthesis of proteoglycans by cultured chondrocytes from the Swarm rat chondrosarcoma. *J. Biol. Chem.* 1256:4368-4376.
- Kjellén, L. and Lindahl, U. 1991. Proteoglycans: Structure and interactions. *Ann. Rev. Biochem.* 60:443-475.

Parthasarathy, N. and Spiro, R.G. 1981. Characterization of the glycosaminoglycan component of the renal glomerular basement membrane and its relationship to the peptide portion. *J. Biol. Chem.* 256:507-513.

Rodén, L., Baker, J.R., Cifonelli, J.A., and Matthews, M.B. 1972. Isolation and characterization of connective tissue polysaccharides. *Methods Enzymol.* 28:73-140.

Sajdera, S.W. and Hascall, V.C. 1969. Protein polysaccharide complex from bovine nasal cartilage. A comparison of low and high shear extraction procedures. *J. Biol. Chem.* 244:77-87.

Yanagishita, M., Midura, R., and Hascall, V.C. 1987. Proteoglycans: Isolation and characterization from tissue cultures. *Methods Enzymol.* 138:279-289.

Key References

Rodén et al., 1972. See above.

This review details conventional methods for isolating and characterizing the composition, size, and fine structure of GAG chains.

Hascall et al., 1982. See above.

This review describes detailed methods for purifying and characterizing tissue proteoglycans.

Yanagishita et al., 1987. See above

This review extends accepted methodology for purifying and characterizing tissue proteoglycans to those found in cultured cells.

Contributed by Jeffrey D. Esko
University of Alabama
Birmingham, Alabama

Special Considerations for Glycolipids and Their Purification

The term “glycolipid” refers to a large class of heterogeneous and rather ill-defined conjugates composed of sugar and lipid. They range from the large oligosaccharide conjugates of polyprenols and dolichols (intermediates in the biosynthesis of bacterial-cell-wall polysaccharides and animal-cell glycoproteins, respectively) to membrane components of plants, yeast, mycobacteria, and animal cells. This unit presents methods for preparing glycosphingolipids, the most commonly studied glycolipids, for analysis. In general, these studies are used for comparison of glycosphingolipids from different cells or cell lines or for comparison of experimentally treated and control cells or cell lines.

The first basic protocol describes the extraction of all glycosphingolipids, including those carrying large oligosaccharides, from small amounts of tissue or cultured cells, and their subsequent separation based on their relative solubilities in polar and nonpolar solvents (Folch partitioning). A support protocol describes the preparation and use of Sep-Pak C₁₈ cartridges for use in the first basic protocol. The second basic protocol describes preparation of gangliosides from the final product of the previous protocol (Folch upper phase) and their separation into mono-, di-, and polysialylated gangliosides by DEAE-Sephadex chromatography. The alternate protocol describes a more rapid method for separation of gangliosides from the total lipid extract generated partway through the first basic protocol.

STRATEGIC PLANNING

Whether a class of glycolipids is involved in biosynthesis of other glycoconjugates or in the anchoring of specific oligosaccharides or proteins to cell membranes, the unique hydrophilic and hydrophobic properties that play a significant role in their function also present unique challenges for their isolation and purification. Because of their diversity, no single protocol is convenient for purification of all classes of glycolipids; the method of choice will depend on what glycolipids are of interest. Discussion of the purification of all of the diverse classes of glycolipids is beyond the scope of a single unit. However, similarities in physical properties do permit the use of many similar or related techniques for the different classes.

Most research has focused on cell-surface glycolipids of animal cells, particularly glycosphingolipids and glycosyl phosphatidylinositol (GPI) membrane anchors (UNIT 17.8). Glycosphingolipids are glycolipids composed of a ceramide (a fatty acyl derivative of the long-chain base sphingosine) glycosidically linked to a monosaccharide, usually glucose or galactose, that may in turn be modified with additional sugar molecules (see Fig. 17.3.3). Glycosphingolipids are heterogeneous in their properties due to variation in the fatty acid and/or sphingosine structure of the ceramide component as well as the diversity of the oligosaccharide component. For example, glycosphingolipids composed of ceramides with short fatty-acid side chains and large oligosaccharide chains will be much more soluble in polar solvents than glycosphingolipids composed of ceramides with long fatty-acid side chains and short oligosaccharide chains. In addition, charged residues on oligosaccharides will decrease their hydrophobicity.

The two classes of glycosphingolipids to be considered in this unit are (1) the neutral glycosphingolipids—from monohexosylceramides up to very large neutral glycolipids that have been referred to as polyglycosyl ceramides; and (2) the negatively charged glycosphingolipids—primarily gangliosides, which are glycosphingolipids containing negatively charged sialic acid residues.

The first basic protocol separates the lipid extract into a nonpolar lower-phase fraction containing neutral glycolipids smaller than tetrahexosylceramides along with contaminating neutral and zwitterionic lipids, and a polar upper-phase fraction containing gangliosides and neutral glycolipids generally larger than tetrahexosylceramides. The upper-phase fraction is desalted by elution from a Sep-Pak C₁₈ cartridge. The lower phase can subsequently be fractionated using methods unaffected by contaminating neutral lipids and phospholipids. These lower-phase separations are not affected by neutral lipid and phospholipid contaminants if they are labeled with radioactive sugar precursors, because only the sugars will be detected by the separation technique. Contaminated lower phases can also be analyzed with carbohydrate-specific reagents if the cells used as starting material were metabolically radiolabeled (UNIT 17.4). Thin-layer chromatography (TLC; UNITS 10.12-10.14) combined with autoradiography (APPENDIX 3) can be used for separation and detection. This permits detection and characterization of extremely small amounts of glycolipids without the use of expensive instrumentation, such that single culture dishes of radiolabeled cells can be used to obtain significant structural information on purified, radiolabeled glycosphingolipids. Alternatively, high-performance liquid chromatography (HPLC; UNIT 9.6) or overlay analysis of thin-layer chromatograms using ¹²⁵I-labeled carbohydrate-binding proteins may be employed. The upper phase can be treated as described in the second basic protocol to purify the gangliosides. If only the gangliosides (and not the neutral glycosphingolipids) are of interest, the process can be shortened by carrying out just the first nine steps of the first basic protocol followed by the rapid ganglioside extraction method described in the alternate protocol. Like the first basic protocol, both protocols for purifying gangliosides are suitable for analysis of radiolabeled glycoconjugates (UNIT 17.4). All of these protocols are compatible with subsequent extraction and analyses of total membrane glycopeptides or glycoproteins (UNITS 17.1, 17.4 & 17.5) from the same sample. Figure 17.3.1 diagrams the fractionation process and relates it to the individual protocols.

CAUTION: If cells being used are metabolically radiolabeled (UNIT 17.4), appropriate procedures for manipulation and disposal of radioactive samples should be employed.

NOTE: Deionized, distilled water or its equivalent should be used throughout this unit.

EXTRACTION AND PURIFICATION OF GLYCOLIPIDS FROM CELLS OR TISSUES

BASIC PROTOCOL

Because glycosphingolipids contain both oligosaccharide and ceramide components, they are amphipatic—i.e., present the contradictory properties of being both hydrophilic and hydrophobic. Depending on the relative content of polar, nonpolar (neutral sugar), and charged substituent groups, glycosphingolipids behave differently in aqueous and nonaqueous solvents. Thus, a wide variety of methods for extracting and separating these compounds based on solubility have been developed. This protocol details a method for extracting not only the nonpolar glycosphingolipids but also the charged and highly glycosylated derivatives from cultured cells or small amounts of tissue (≤ 0.5 g). Total cell lipids are first extracted with organic solvents, then separated into fractions by the Folch partition method (Fig. 17.3.2).

Materials

- Tissue sample or cell culture from one to several culture plates
- Methanol (HPLC grade)
- Chloroform (HPLC grade)
- 4:8:3 (v/v/v) chloroform/methanol/water
- 1:1 and 2:1 (v/v) chloroform/methanol
- 100% ethanol

Preparation and Analysis of Glycoconjugates

17.3.2

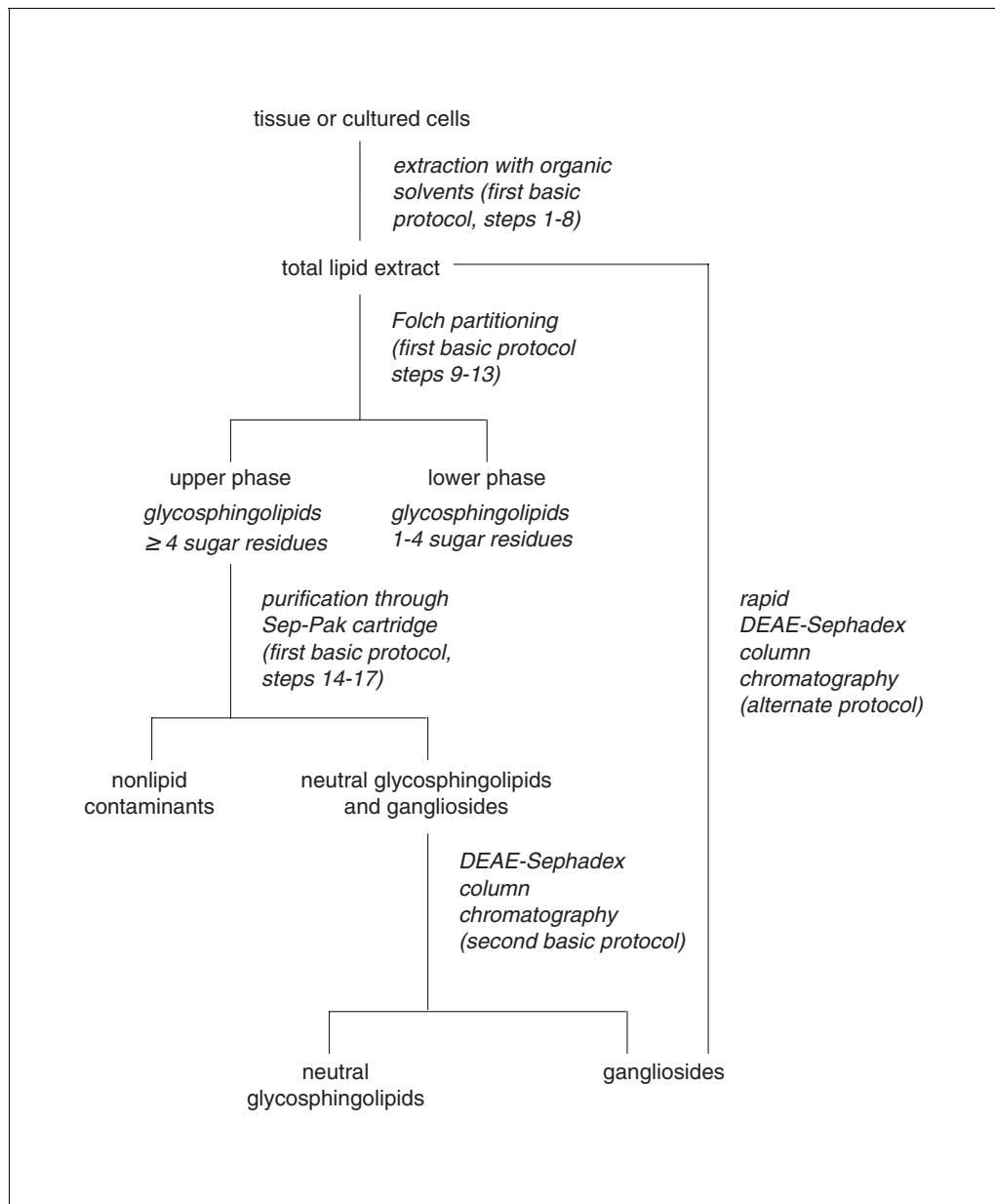


Figure 17.3.1 Flow chart showing fractionation of glycosphingolipids from total lipid extract of cells.

0.1 M KCl

1:1 (v/v) methanol/0.1 M KCl

1:1 (v/v) methanol/H₂O

Sep-Pak C₁₈ cartridge (support protocol)

Blender (Waring), tip sonicator (Branson), or Tissuemizer (Tekmar) equipped with a probe of appropriate size

Bath sonicator

Rotary evaporator, Speedvac (Savant), Vortex-Evaporator (Labconco), or nitrogen dryer

1.5-ml polypropylene microcentrifuge tubes resistant to chloroform or 12-ml conical glass centrifuge tubes

Additional reagents and equipment for quantitation of protein (UNIT 10.1), and protein electrophoresis (UNITS 10.2-10.5)

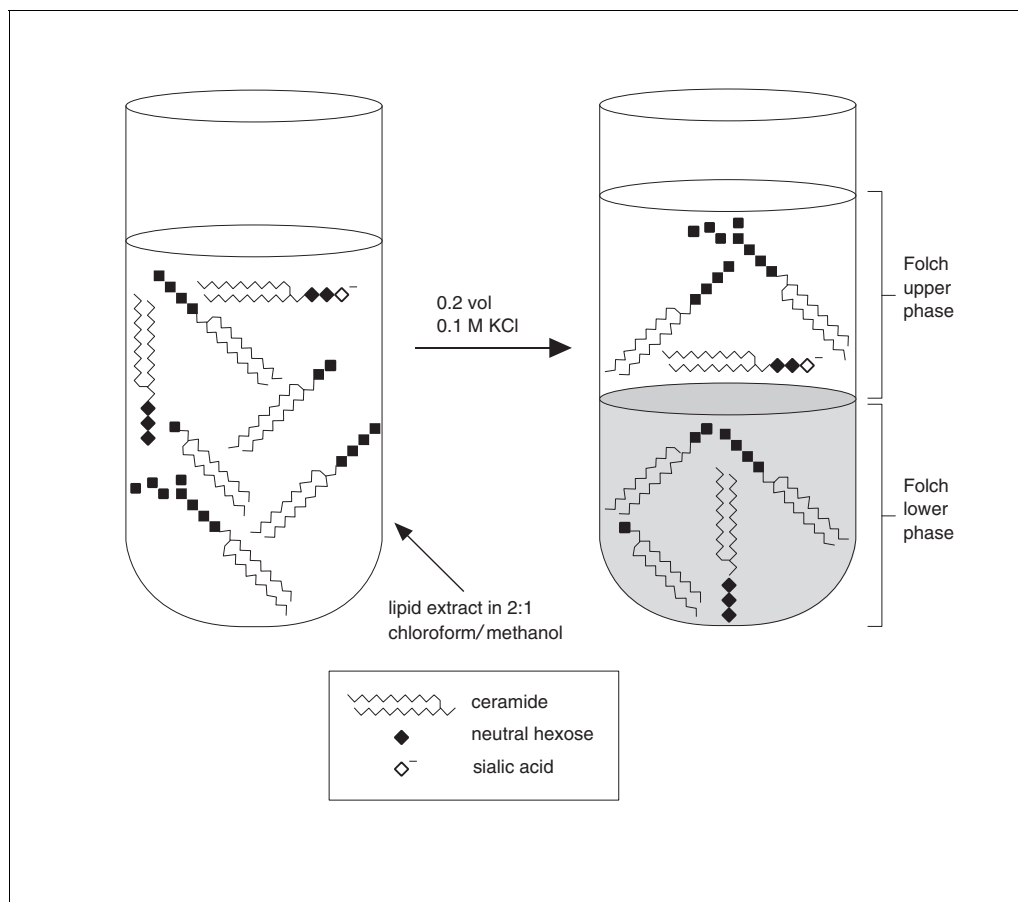


Figure 17.3.2 Folch partitioning of glycosphingolipids.

Extract sample with organic solvents

1. If using cultured cells, harvest and wash by centrifuging, removing supernatant, resuspending pellet in PBS or a suitable isotonic solution, centrifuging again, and discarding supernatant, according to the usual procedure for the cell type (this removes serum components and, in the case of metabolic labeling experiments, unincorporated radioactive precursor). Suspend tissue or cell pellet in 2° to 4°C water, using 2 to 3 ml water per gram of tissue (wet weight) or enough water to make a suspension of cultured cells containing 0.1 to 6 mg/ml protein. Homogenize using a blender, sonicator, or Tissuemizer.

Efficient extraction of glycolipids from lyophilized cell preparations is very difficult and should be avoided. For cultured cells, a pellet from one to several 60-mm plates of cells may be suspended in 0.24 ml of water in a 1.5-ml microcentrifuge tube. With microcentrifuge tubes, homogenization can be carried out using a bath sonicator. For larger samples (up to 1 ml), 12-ml conical glass centrifuge tubes should be used.

Samples should be kept on ice between homogenization and extraction to prevent digestion of the carbohydrates by endogenous glycosidases.

2. To 3 vol aqueous homogenate, add 8 vol methanol followed immediately by 4 vol chloroform, mixing continuously while adding the solvents. For small volumes, this step can be carried out in a bath sonicator.

In the case of a sample initially suspended in 0.24 ml, 0.64 ml methanol should be added, followed by 0.32 ml chloroform, for a total extract volume of ~1.2 ml.

3. Incubate sample in a bath sonicator 30 min at 20° to 30°C.

Temperatures of 20° to 30°C may destroy labile O-acetyl groups and/or induce transacylation of O-acetylated derivatives of sialic acids. Lower temperatures may be required if information on sialic acid acetylation is critical to subsequent analyses; in this case solvents should be chilled to the appropriate temperature before use. Temperatures at or below room temperature can be maintained by adding ice to the bath sonicator.

4. Centrifuge the sample 1 min at $8800 \times g$ (e.g., 11,000 rpm in a microcentrifuge) and collect supernatant.
5. Extract the pellet from step 4 with a volume of 4:8:3 chloroform/methanol/water roughly approximating the total extract volume in step 2 (e.g., ~1 ml may be used for a sample initially suspended in 0.24 ml water). Centrifuge sample 1 min at $8800 \times g$. Collect supernatant and pool with that from step 4.
6. Extract the pellet from step 5 three more times as described in step 4, using 1:1 chloroform/methanol, 2:1 chloroform/methanol, and 100% ethanol, respectively. Pool all the supernatants with that from steps 4 and 5.
7. If desired, analyze the ethanol-washed pellet (containing primarily lipid-free proteins and glycoproteins) for total proteins by one- or two-dimensional gel electrophoresis or digestion with protease for isolation of total glycopeptides. Do not allow pellet to dry; resuspend in a buffer containing SDS and store at -20° or -80°C prior to further manipulation if desired.

The two chloroform/methanol washes, using solvents with increasingly higher concentrations of nonpolar solvents compared to the sample, ensure the extraction of less polar glycolipids (i.e., mono- or disaccharide derivatives of ceramide that are more hydrophobic than glycolipids with larger oligosaccharide components). The ethanol wash removes the chloroform and methanol to avoid any detrimental effects these solvents may have on subsequent treatment of the cell pellet.

8. Evaporate the pooled supernatants from steps 4 to 6 to dryness using a rotary evaporator, Vortex-Evaporator, Speedvac, or nitrogen dryer.

Evaporating the sample under a stream of nitrogen requires considerably more time to remove the final traces of H_2O than does evaporation with a rotary evaporator, Vortex-Evaporator, or Speedvac. Dry crude lipid extract can be stored desiccated at -20°C prior to further manipulation.

Carry out Folch partitioning of glycosphingolipids

9. Dissolve the crude lipid extract from step 8 in 1.0 ml of 2:1 chloroform/methanol (or 10% to 25% the volume of the initial suspended sample if this is larger than 1 ml).
10. Partition by adding 0.2 ml (or 0.2 vol) of 0.1 M KCl and mixing vigorously by vortexing. Centrifuge 1 min at $8800 \times g$ (it will separate into two phases).

For small-scale preparations, this can be done in a 1.5-ml polypropylene microcentrifuge tube. For larger volumes, phase separation may be accomplished by allowing the mixture to stand at room temperature until separated (or in a separatory funnel overnight at 4°C).
11. Carefully remove the upper phase (~50% of the total volume) with a Pasteur pipet. Wash lower phase by adding 1:1 methanol/0.1 M KCl in a volume equivalent to the removed upper phase, mixing, and centrifuging 1 min at $8800 \times g$ to separate the phases (this removes contaminating upper phase).
12. Collect the upper phase and pool with the first upper phase. Wash the lower phase once more with 1:1 methanol/0.1 M KCl and pool with upper phases from step 11.

13. Wash the combined upper phases with 0.5 vol of 2:1 chloroform/methanol. Combine the lower phases from the washes with the pooled lower phases from step 12.

The combined lower-phase fractions contain neutral glycolipids with oligosaccharides smaller than tetrasaccharides, neutral lipids, and phospholipids. Tetrasaccharide-ceramides usually partition between the two phases, causing slight contamination. The combined upper-phase fractions contain all of the gangliosides and most of the larger neutral glycolipids (>tetrahexosylceramides) as well as all water-soluble materials from the original extract and inorganic salts introduced during washes of the lower phase. When metabolically radiolabeled cells are used, the upper phase will also contain radioactive sugar precursors and their intermediates.

The combined lower-phase fractions can now be analyzed for glycosphingolipids using methods unaffected by contaminating neutral and phospholipids—e.g., TLC and autoradiography of glycolipids (if cells were radiolabeled with monosaccharide precursors), overlay analysis of thin-layer chromatograms using ¹²⁵I-labeled carbohydrate-binding proteins, or HPLC (Kannagi, 1987). The combined upper-phase fractions can be processed as described in the following steps and further purified as described in the last two protocols.

Prepare combined upper-phase fractions for analysis

14. Dilute combined upper-phase fraction to 5 to 10 ml with 1:1 methanol/water and apply it to a Sep-Pak C₁₈ cartridge prepared as described in the support protocol below. Collect effluent in a test tube and reapply to the cartridge; repeat once.

The liquid should be applied to the Sep-Pak C₁₈ cartridge as described in the support protocol. Passing the upper-phase fraction through the cartridge removes salts and other nonlipid contaminants. Collecting and reapplying the effluent to the cartridge ensures complete absorption of all lipid-containing compounds in the sample.

15. Wash the cartridge five times with 10-ml aliquots of water (50 ml total) to ensure removal of all polar contaminants.

When radiolabeled cells are used, these washes can be pooled for further analysis of radiolabeled precursors, but should contain no glycolipid. They should be discarded as radioactive waste.

16. Elute gangliosides and high-molecular-weight neutral glycolipids from the column by washing five times with 1:1 chloroform/methanol and three times with 4:8:3 chloroform/methanol/water.

The final wash with more polar solvent elutes glycolipids containing large oligosaccharides (when present in the sample) that were not eluted with 1:1 chloroform/methanol.

17. Combine solvent washes and evaporate to dryness. Store residue desiccated at –20°C until needed.

PREPARATION OF SEP-PAK C₁₈ CARTRIDGES

This protocol details preparation of Sep-Pak C₁₈ cartridges for use in removing salts and other nonlipid contaminants from the upper-phase glycolipid fraction of Folch-partitioned material. All lipid-containing compounds will be retained by the Sep-Pak C₁₈ cartridge, water-soluble contaminants are eluted with polar solvents, and lipid-containing material is eluted with nonpolar solvent.

Additional Materials

- 0.1 M ammonium acetate in 1:1 (v/v) methanol/water
- Sep-Pak C₁₈ cartridges (Waters)
- 10-ml glass syringe with locking hub (e.g., Becton Dickinson Luer-Lok)

SUPPORT PROTOCOL

Preparation and Analysis of Glycoconjugates

17.3.6

1. Remove plunger from the barrel of the syringe and attach the Luer-Lok hub of the syringe to the long end of a Sep-Pak C₁₈ cartridge. Pump 10 ml methanol through the cartridge by adding the liquid to the barrel of the syringe, inserting the plunger into the syringe, and slowly pushing the liquid through the cartridge. Collect eluant methanol in a beaker or other appropriate container. Remove the cartridge from the syringe *before* removing the plunger. Remove the plunger and reattach the cartridge.
2. In the same fashion, carry out the following further washes (10 ml each):
 - 3 washes with 1:1 chloroform/methanol
 - 1 wash with methanol
 - 2 washes with 0.1 M ammonium acetate in 1:1 methanol/water.

Sep-Pak cartridges may be reused at least five times. They should be prepared for storage by flushing the last of the methanol/water wash out with air and stored dry at room temperature. Preparation steps should be repeated after storage prior to use.

BASIC PROTOCOL

PREPARATION OF GANGLIOSIDES

Fractionation of the glycolipid extract into neutral glycolipids and gangliosides can be easily accomplished by separating charged glycolipids from neutral glycolipids of the upper-phase fraction obtained by Folch partitioning (the material from step 17 above). A simple column chromatographic procedure is described here, but more sophisticated methods, including anion-exchange HPLC (UNIT 10.13), may be used. This protocol is adapted from a method originally developed to isolate the gangliosides from the upper phase of extracts from 50 g of tissue but also works well on a small scale for isolating gangliosides from the upper phase of extracts of metabolically radiolabeled cultured cells.

Materials

Methanol
 DEAE-Sephadex equilibrated in 100% methanol,
 50% slurry (reagents and solutions)
 Purified Folch upper-phase fraction (first basic protocol)
 0.01, 0.2, and 0.5 M ammonium acetate in methanol
 Glass wool

1. Set up a DEAE-Sephadex column in a 5-in. Pasteur pipet as follows: secure the Pasteur pipet to a ring stand with a small clamp so that the column effluent can be collected into a beaker or into tubes in a test tube rack for manual processing. Use a 9-in. pipet to position a small ball of glass wool in the narrow portion of the Pasteur pipet.

If the amount of glass wool is too large, or if it is packed too tightly into the tip of the pipet, the flow rate will be restricted. Alternatively, a column with appropriate solvent-resistant fittings can be used with a fraction collector for automatic sample collection.

2. Wash the glass wool with methanol to determine an appropriate flow rate in the empty column (this also removes any contamination from the glass wool). If the flow rate is less than 1 drop/sec, repack the glass wool to ensure a reasonable flow rate when the column is packed with DEAE-Sephadex.
3. Add sufficient 50% slurry of DEAE-Sephadex in methanol to the pipet to obtain a column of resin 2 to 4 cm high.

A 4-cm column has a bed volume of ~1 ml; half that volume is normally sufficient capacity for radiolabeled material.

4. Wash the column with 2 to 3 column volumes of methanol.
5. Dissolve the upper-phase fraction in 0.1 to 1 ml methanol and apply to the column of DEAE-Sephadex.

The upper-phase fraction is free of salts and other nonlipid contaminants. Because the charged glycolipids (gangliosides) will be retained by the column under these conditions, the sample volume is not critical.

6. Elute neutral glycolipids in the upper phase by washing the column with 5 column volumes of methanol. Evaporate the eluant to dryness (e.g., using a nitrogen dryer).

This fraction contains desalted neutral glycolipids that are larger than trihexosylceramides. These desalted neutral glycolipids can be combined with the lower-phase fraction that contains the remainder of the neutral glycolipids. This is a convenient fraction for many analyses including thin-layer chromatography (TLC) and TLC overlay analyses.

7. Elute gangliosides that contain a single sialic acid (monosialylgangliosides) from the column with 5 column volumes of 0.01 M ammonium acetate in methanol.

Eluant may be collected in equal-sized fractions if the elution profile is important, or it may be collected in a single container. Concentrate monosialylgangliosides (e.g., by evaporation under nitrogen). If this fraction is to be used for TLC analysis, it must be desalted. This can conveniently be done by adding water to 50% and applying the material to a Sep-Pak C₁₈ cartridge (support protocol).

8. Elute the gangliosides that contain two sialic acids per oligosaccharide with 5 column volumes of 0.2 M ammonium acetate in methanol.

9. Elute polysialylgangliosides (tri- and tetrasialylgangliosides) with 5 column volumes of 0.5 M ammonium acetate in methanol.

These fractions can be concentrated and/or desalted on Sep-Pak cartridges as described for monosialylgangliosides in step 7.

If the sample contains sulfated glycolipids, it may be necessary to elute them with 0.8 to 1.0 M ammonium acetate in methanol.

RAPID PURIFICATION OF GANGLIOSIDES

For projects that focus on the analysis of only gangliosides, this protocol describes a method for rapid preparation of total gangliosides from total cellular lipid extract using an ion-exchange column equilibrated in a more polar solvent.

Additional Materials

DEAE-Sephadex equilibrated in 60:30:8 methanol/chloroform/water (reagents and solutions)

Total lipid extract (step 8 of first basic protocol)

60:30:8 (v/v/v) methanol/chloroform/water

60:30:8 (v/v/v) methanol/chloroform/0.8 M aqueous KCl

1. Prepare a DEAE-Sephadex column as described in step 1 of the second basic protocol, but using DEAE-Sephadex equilibrated in 60:30:8 methanol/chloroform/water.
2. Dissolve the total lipid extract from cells or tissues from step 8 of the first basic protocol in 60:30:8 methanol/chloroform/water.

The wash volume is not critical; a volume equivalent to the volume of the original extract is generally appropriate.

**ALTERNATE
PROTOCOL**

**Preparation
and Analysis of
Glycoconjugates**

17.3.8

3. Apply total lipid extract to the column slowly (<1 ml/min) to ensure complete absorption of charged lipids.
4. Elute the column with 10 column volumes of 60:30:8 methanol/chloroform/water to remove uncharged and zwitterionic lipids.

The flow rate can be increased for this and subsequent elution steps.

5. Elute gangliosides and other acidic lipids with 10 column volumes of 60:30:8 methanol/chloroform/0.8 M aqueous KCl.

REAGENTS AND SOLUTIONS

DEAE-Sephadex, equilibrated

Suspend 5 g dry DEAE-Sephadex resin in 60 ml of 60:30:8 methanol/chloroform/0.8 M aqueous KCl. Allow resin to settle and remove supernatant by aspiration. Repeat this procedure twice and allow the final slurry to equilibrate overnight (this prepares the resin in its acetate form). Remove the supernatant and equilibrate the resin by repeated washing with 60:30:8 methanol/chloroform/water through filter paper or a coarse-porosity sintered-glass funnel.

The equilibrated resin can be stored as a 50% slurry in 60:30:8 methanol/chloroform/water for immediate application in the alternate protocol for rapid preparation of gangliosides. For separation of upper-phase gangliosides, equilibrate the resin in 100% methanol and store as a 50% slurry in methanol.

COMMENTARY

Background Information

Most research on glycolipids has focused on the analysis of cell surface glycolipids of animal cells which include the glycosphingolipids and the glycosyl phosphatidylinositol (GPI) membrane anchors. Glycosphingolipids are composed of ceramide, which is a fatty acyl derivative of the long-chain base sphingosine, glycosidically linked to a monosaccharide, usually glucose or galactose (Fig. 17.3.3). Other sugars may be added to the monosaccharide. One of the GPI anchors first described attaches *Trypanosoma brucei* variant surface glycoprotein (VSG) to the plasma membrane. This glycolipid anchor is composed of phosphatidylinositol linked glycosidically to a tetrasaccharide that terminates with a phosphoethanolamine moiety (Doering et al., 1990). The ethanolamine is linked by an amide linkage to the α -carboxyl group of a protein to form the covalent attachment of VSG to a membrane glycolipid. Variations of this GPI structure have been found in protozoans, yeast, slime molds, and most animal cells.

Because glycosphingolipids are found in the mammalian-cell plasma membrane, they are

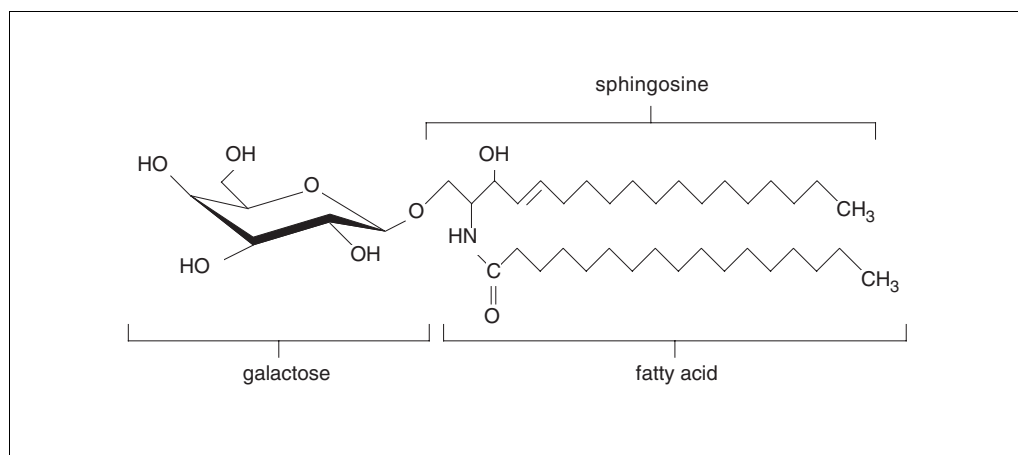


Figure 17.3.3 The structure of galactosylceramide.

the most-studied and best-characterized of all the glycolipids. They are currently of interest to molecular and cellular biologists investigating structure and function of cell surface carbohydrates. In animal cells, glycosphingolipids function as anchors for surface carbohydrate antigens and as potential receptors for attachment of bacteria, viruses, and toxins (Karlsson, 1989). Glycosphingolipids may also have more dynamic functions as modulators of cellular interactions and transmembrane signaling (Hakomori, 1990).

The oligosaccharides of glycosphingolipids can be very complex; their heterogeneity is reviewed by Stults et al. (1990), who have cataloged the structures of 267 glycosphingolipids. In general glycosphingolipids can be divided into three classes based on their carbohydrate structures: neutral glycolipids with uncharged oligosaccharides; gangliosides with oligosaccharides containing one or more negatively charged sialic acid residues; and sulfato-glycosphingolipids with oligosaccharides derivatized with sulfate esters. The oligosaccharide “core” structure of most animal-cell glycolipids consists of lactose linked glycosidically to ceramide. Elongation of this lactosylceramide by monosaccharide addition results in several families of glycolipids based on the sequence and linkages of sugars. Table 17.3.1 (based on Machler and Sweeley, 1978; Weigandt, 1985; and Makaanu et al., 1992) lists prefixes used in naming glycolipids and their associated structures.

Variations on the basic glycolipid structure are found throughout nature. Plants, yeast, and

fungi express inositol-containing sphingolipids, and methods for their isolation and characterization have been reviewed by Laine and Hsieh (1987). These glycolipids are glycosphingolipids made up of a common core of inositol joined in a phosphodiester linkage to ceramide, with longer oligosaccharide chains glycosidically linked to the inositol moiety. Mycobacteria contain several novel classes of glycolipids including the lipo-oligosaccharides, glycopeptidolipids, and phenolic glycolipids; the structure, isolation, and antigenicity of these compounds have been reviewed by McNeil et al. (1989).

Glycosphingolipids are minor components in animal cells and much of the difficulty in their characterization is associated with their limited availability. This unit describes protocols that can be used to characterize the small amounts of material available in cultured cells. Glycosphingolipids are unique among glycoconjugates in that, regardless of their oligosaccharide complexity, each species carries only a single oligosaccharide moiety. A chromatographically pure sample generally represents a glycosphingolipid with a single oligosaccharide. The oligosaccharide structures of purified glycosphingolipids are determined using many of the structural analytical techniques described in this chapter. The apparent fidelity of structure (one glycosphingolipid, one oligosaccharide) makes it possible to assign structures based on co-chromatography of unknown structures with known standards on TLC, using chemical methods for sugar detection to locate individual glycosphingolipids. This approach

Table 17.3.1 Oligosaccharide Structures and Prefixes^a

Prefix	Structure
Arthro	GlcNAc β 1-3Man β 1-4GlcCer
Gala	Gal α 1-4GalCer
Ganglio	Gal β 1-3GalNAc β 1-4Gal β 1-4GlcCer
Globo	GalNAc β 1-3Gal α 1-4Gal β 1-4GlcCer
Globoiso	GalNAc β 1-3Gal α 1-3Gal β 1-4GlcCer
Lacto	Gal β 1-3GlcNAc- β 1-3Gal β 1-4GlcCer
Lactoneo	Gal β 1-4GlcNAc- β 1-3Gal β 1-4GlcCer
Muco	Gal β 1-3Gal β 1-4Gal β 1-4GlcCer
Mollu	Man α 1-3Man β 1-4GlcCer
Schisto	GalNAc β 1-4GlcCer

^aThese prefixes are used in designations of structures within each family. For example, the structure listed for the Lactoneo series is lactoneotetra-glycosylceramide and that listed for the Globo series is globotetra-glycosylceramide. Extensive discussion of nomenclature is beyond the scope of this unit; for more information, refer to Weigandt, 1985, and Macher and Sweeley, 1978.

has generated considerable information on the specific tissue and species distribution of glycosphingolipids. As more sensitive techniques have become available, glycosphingolipids, like glycoproteins, have been found to possess far greater structural variety than originally assumed. Research in this field is shifting from structural to functional analysis. The development of methods that define receptor function or immunological reactivity of glycolipids prior to their structural analysis has been important to functional studies.

The protocols in this unit are based on several procedures for analysis of glycosphingolipids extracted from mammalian tissues. In Folch partitioning, lipid extracts are partitioned into upper and lower phases based on differential solubility of glycolipids with different-sized oligosaccharide chains (Folch et al., 1957; Esselman et al., 1972; Ledeen and Yu, 1982). The extraction procedure is fundamentally that described by Svennerholm and Fredman (1980) for the quantitative extraction of gangliosides from animal tissues—a procedure selected to ensure complete extraction of glycolipids. The application of this procedure to the extraction and analysis of glycolipids and glycoproteins from the same sample was suggested by Finne and Krusius (1982), who used the method to obtain lipid-free glycopeptides. Isolation and fractionation of gangliosides by DEAE-Sephadex chromatography is adapted from Momoi et al. (1976). These protocols have been used successfully in the isolation, purification, and identification of immunologically reactive glycosphingolipids from both unlabeled (Magnani et al., 1982) and metabolically radiolabeled (Clark et al., 1991) cultured animal cells.

Critical Parameters and Troubleshooting

The use of metabolically radiolabeled cells as described in *UNIT 17.4* permits analysis of a small number of cells to obtain significant information on the carbohydrate composition and structure of glycosphingolipids.

The success of the separation and purification procedures will be obvious upon analysis of the glycosphingolipid fractions obtained. Poor resolution of glycolipids during most chromatographic analyses is generally associated with the presence of salts in the sample or with poorly equilibrated columns. Use of the Sep-Pak C₁₈ cartridge is the most convenient method for removal of salt from the sample

(Williams and McClure, 1980).

The combined lower-phase neutral glycolipid fraction will contain non-glycolipid contaminants of the lower phase, including neutral and zwitterionic lipids such as cholesterol, lecithin, ethanolamine, phosphoglycerides, etc. For many analyses this contamination does not contribute to, or interfere with, separation or detection methods. For example, in the case of material from cells metabolically labeled with monosaccharide precursors, essentially all of the detected label will be in the sugars. Label can, however, appear in non-glycolipid components of the lower phase of a Folch extract. In extracts of Chinese hamster ovary cells, such material was tentatively identified as phosphatidylinositol (Smith et al., 1990). *UNIT 17.16* describes how to determine the composition of radiolabeled monosaccharides of isolated glycolipids. When specific reactions such as sugar-specific stains or carbohydrate-specific antibodies are used to identify glycolipids on TLC plates, these contaminants will not react. If removal of the contaminants is necessary, it can be accomplished by saponification followed by chromatography on silicic acid (Ledeen and Yu, 1982). However, these procedures will destroy the base-labile substitutions of sialic acid on the gangliosides.

Contaminating phospholipids can also be removed by saponification followed by dialysis. Conditions that affect sialic acids should be considered when planning extraction and purification of gangliosides. To preserve any *O*-acetyl derivatives of the sialic acid residues on gangliosides, which may be very important in functional analyses, it is important to avoid the use of bases. The *O*-acetyl derivatives of sialic acid are extremely labile and migrate via transacylation reactions in these cases.

The plastic microcentrifuge tubes used in this unit must be resistant to chloroform-containing solvents. Most polypropylene tubes claim resistance to chloroform, but it is worth testing the tubes to be sure. Centrifugation steps involving microcentrifuge tubes should be carried out in a cold room, because heat buildup during centrifugation can increase pressure inside capped tube (containing volatile solvents) and cause them to open during centrifugation, resulting in loss of solvent and possibly sample. For large-scale preparations, and where physical methods such as mass spectroscopy are to be used for structural analyses, use a solvent-resistant apparatus and substitute glass for plastic to avoid contamination with plasticizers.

Anticipated Results

The protocols described in this section are intended to provide the investigator with methods for preparing neutral and/or acidic glycosphingolipids from cultured animal cells or small samples of tissue. Because the glycosphingolipids make up such a small amount of total cell mass, it is very difficult to obtain information related to the theoretical yield of glycolipid; however, these protocols will provide a representative sample of the neutral glycolipids and gangliosides present in the starting material. The results of each isolation step and how each sample may be used for further analysis are described throughout the unit. The final samples of neutral glycolipid and/or ganglioside are to be used for further analysis or characterization.

Time Considerations

The total extract can be obtained and dried in ~3 hr, depending upon the method used for evaporation. Folch partitioning of the total extract into upper and lower phases will require 30 to 60 min. The preparation of the DEAE-Sephadex requires an overnight equilibration step; if large amounts of resin are equilibrated and stored, only a brief wash of the resin with the starting solvent is required before each time the procedure is carried out. Preparation of a Sep-Pak C₁₈ cartridge according to the support protocol will require ~30 min. The desalting step using a Sep-Pak cartridge is somewhat cumbersome and time-consuming because multiple samples must be processed individually and the procedure requires considerable manipulation of syringe and cartridge. If multiple samples are being processed, commercial vacuum manifolds that accommodate large numbers of cartridges may be used. Several hours are required to run the DEAE-Sephadex column with a flow rate of 1 to 2 ml/min and multiple solvent systems to fractionate gangliosides. From extraction to fractionation of gangliosides can take up to 2 days. Rapid isolation of the ganglioside fraction (steps 1 to 12 of the first basic protocol) for use in the alternate protocol can be accomplished within 1 to 2 hr; the alternate protocol will produce neutral and acidic fractions within 1 day.

Literature Cited

- Clark, G.F., Cummings, R.D., Mattox, S., Gorbea, C., and Smith, D.F. 1991. Decreased biosynthesis of Forssman glycolipid after retinoic acid-induced differentiation of mouse F9 teratocarcinoma cells: Lectin affinity chromatography of the glycolipid-derived oligosaccharides. *Carbohydr. Res.* 213:155-168.
- Doering, T.L., Masterson, W.J., Hart, G.W., and Englund, P.T. 1990. Biosynthesis of glycosyl phosphatidylinositol membrane anchors. *J. Biol. Chem.* 265:611-614.
- Esselman, W.J., Laine, R.A., and Sweeley, C.C. 1972. Isolation and characterization of glycosphingolipids. *Methods Enzymol.* 28:140-159.
- Finne, J. and Krusius, T. 1982. Preparation and fractionation of glycopeptides. *Methods Enzymol.* 83:269-277.
- Folch, J., Lees, M.B., and Sloane Stanley, G.H. 1957. A simple method for the isolation and purification of total lipides from animal tissues. *J. Biol. Chem.* 226:497-503.
- Hakomori, S. 1990. Bifunctional role of glycosphingolipids. *J. Biol. Chem.* 265:18713-18716.
- Kannagi, K., Watanabe, K., and Hakomori, S. 1987. Isolation and purification of glycosphingolipids by high performance liquid chromatography. *Methods Enzymol.* 138:3-12.
- Karlsson, K.-A. 1989. Animal glycosphingolipids as membrane attachment sites for bacteria. *Ann. Rev. Biochem.* 58:309-350.
- Laine, R.A. and Hsieh, T.C.-Y. 1987. Inositol-containing sphingolipids. *Methods Enzymol.* 138:186-195.
- Ledeen, R.W. and Yu, R.K. 1982. Gangliosides: Structure, isolation and analysis. *Methods Enzymol.* 83:139-191.
- McNeil, M., Chatterjee, D., Hunter, S.W., and Brennan, P.J. 1989. Mycobacterial glycolipids: Isolation, structures, antigenicity, and synthesis. *Methods Enzymol.* 179:215-242.
- Macher, B.A. and Sweeley, C.C. 1978. Glycosphingolipids: Structure, biological source, and properties. *Methods Enzymol.* 50:236-250.
- Magnani, J.L., Nilsson, B., Brockhaus, M., Zopf, D., Steplewski, Z., Koprowski, H., and Ginsburg, V. 1982. A monoclonal antibody-defined antigen associated with gastrointestinal cancer is a ganglioside containing sialylated lacto-*N*-fucopentaose II. *J. Biol. Chem.* 257:14365-14369.
- Makaaru, C.K., Damian, R.T., Smith, D.F., and Cummings, R.D. 1992. The human blood fluke *Schistosoma mansoni* synthesizes a novel type of glycosphingolipid. *J. Biol. Chem.* 267:2251-2257.

- Momoi, T., Ando, S., and Nagai, Y. 1976. High resolution preparative chromatography system for gangliosides using DEAE-Sephadex and the new porous silica, iatrobeads. *Biochim. Biophys. Acta* 441:488-497.
- Smith, D.F., Larsen, R.D., Mattox, S., Lowe, J.B., and Cummings, R.D. 1990. Transfer and expression of murine UDPGal:β-D-Galα1,3Galactosyltransferase gene in transfected Chinese hamster ovary cells. *J. Biol. Chem.* 265:6225-6234.
- Stults, C.L.M., Sweeley, C., and Macher, B.A. 1990. Glycosphingolipids: Structure, biological source, and properties. *Methods Enzymol.* 179:167-214.
- Svennerholm, L. and Fredman, P. 1980. A procedure for the quantitative isolation of brain gangliosides. *Biochim. Biophys. Acta* 617:97-109.
- Weigandt, H. 1985. Glycolipids. In *New Comprehensive Biochemistry* (A. Neuberger and L.L.M. van Deenen, eds.) Vol. 10. pp. 2-4. Elsevier Science Publishing, New York.
- Williams, M.A. and McClure, R.H. 1980. The use of Sep-Pak C₁₈ Cartridge during the isolation of gangliosides. *J. Neurochem.* 35:266-269.

Contributed by David F. Smith
The University of Georgia
Atlanta, Georgia

Pedro A. Prieto
Ross Laboratories
Columbus, Ohio

DETECTION OF SACCHARIDES ON GLYCOCONJUGATES

The first step towards analyzing the structure and function of oligosaccharides on a glycoconjugate is the detection of the sugar chains themselves. The units presented in this section describe a variety of ways in which this can be achieved. In some cases, such studies may provide the first evidence that the molecule being examined is a glycoconjugate. In other cases, these methods may be used to qualitatively or quantitatively analyze the nature of the sugar chains on a molecule that is already known to be a glycoconjugate. For help in deciding which protocols to use, the user is referred to the section entitled “Choice of Techniques” found in the introduction to this chapter.

Introducing a label into the sugar chains of the glycoconjugate of interest is often informative, and can be very convenient for further analyses. This can be achieved by metabolic radiolabeling (*UNIT 17.4*), chemical labeling (*UNIT 17.5*), or transferring label from a radioactive sugar nucleotide using a glycosyltransferase (*UNIT 17.6*). Each of these approaches has its advantages and disadvantages, in terms of specificity, sensitivity, and practical utility. The user is advised to read the commentaries of each of these units before choosing an approach to take.

It has been recently recognized that many membrane glycoproteins acquire a glycospholipid anchor during biosynthesis. A set of methods designed to detect and analyze such anchors is presented in *UNIT 17.8*. There are also many well-established methods for the direct chemical analysis of the oligosaccharides present in glycoconjugates. In this supplement, the classic phenol–sulfuric acid assay for hexoses and pentoses is presented (*UNIT 17.9*); others will follow in future updates. Finally, oligosaccharides on many glycoconjugates can be altered with a variety of commercially available inhibitors and competitive substrates that can work in whole cells or organisms. In this supplement, the prevention of N-linked glycosylation on glycoproteins is described (*UNIT 17.10*).

There are, of course, many other approaches to the direct analysis of oligosaccharides on glycoconjugates that could have been presented in this section. The current and upcoming choices for inclusion (see Chapter 17 table of contents) were dictated by the need to “keep it simple” and to meet the criterion of broad general utility to the average molecular biologist. As “glycotechnology” improves and simplifies, further additions to this section are anticipated.

Metabolic Radiolabeling of Animal Cell Glycoconjugates

Some useful information about glycoconjugates can be obtained by labeling their aglycone (noncarbohydrate) portions—e.g., labeling proteins with radioactive amino acids (*UNIT 10.18*)—then using techniques described elsewhere in this chapter to infer indirectly the presence, type, and nature of oligosaccharide chains. This unit describes metabolic labeling techniques that provide more specific information about the structure, sequence, and distribution of the sugar chains of glycoconjugates. Although these techniques provide less information than complete sequencing of the sugar chains, the partial structural information derived is sufficient for many purposes. Following metabolic labeling, the radioactive glycoconjugate of interest is isolated, individual glycosylation sites are identified and separated if necessary, and the labeled oligosaccharides are subjected to structural analysis. Metabolic labeling has the advantages of being simple and easy to perform and of requiring no sophisticated instrumentation other than a scintillation counter. Also, purification of the glycoconjugate to radiometric homogeneity is sufficient for further analysis. Important practical considerations for metabolic labeling experiments include selecting the type of experiment and the labeled precursor, understanding the specificity of labeling, and maximizing uptake and incorporation (see critical parameters). Before proceeding with experiments to label glycoconjugates metabolically, labeling conditions should be evaluated to ensure that cell viability and metabolism are not altered, and optimized for uptake and incorporation of the precursor (this should be done for each combination of precursor and cell line).

In the basic protocol, actively growing cell cultures are grown through several population doublings in complete medium supplemented with radiolabeled glycoconjugate precursors to reach a steady-state level of incorporation. In the alternate protocols, cells are cultured for a short period of time in a deficient medium that contains a high concentration of radiolabeled precursor. A pulse or pulse-chase labeling procedure (first alternate protocol) can be used to analyze precursor-product relationships. With sequential pulse-labeling (second alternate protocol), it is possible to obtain quantities of labeled glycoconjugates with the use of a minimal amount of labeled precursor by using the same medium to pulse-label a series of cultures. A support protocol describes the preparation of multiply deficient medium (MDM) for use in making appropriate deficient media.

NOTE: Conventional protocols for handling, monitoring, shielding, and disposing of radioactivity and radioactive waste (*APPENDIX 1*) and for tissue culture of cells and sterile handling of media (Chapter 9) should be followed throughout these protocols.

BASIC PROTOCOL

STEADY-STATE LABELING WITH RADIOACTIVE PRECURSORS

When studying glycoconjugates in an established tissue culture cell line, molecules of interest can be labeled for structural characterization. In the following protocol, actively growing cultures of cells are grown to a steady-state level of incorporation in complete medium supplemented with radiolabeled glycoconjugate precursor.

Optimizing Conditions

An attempt should be made to label the molecules as close as possible to metabolic steady state—i.e., a constant level of radioactivity per mass unit of a given monosaccharide in all glycoconjugates in the cell—under conditions of normal growth. In practice, this is somewhat difficult to achieve with most cell types. The maximum possible number of cells should be grown for the longest possible period of time in the minimum possible volume of complete medium supplemented with the maximum possible amount of

radioactive label. Before labeling is attempted, the following growth characteristics should be determined for the cell type of interest: (1) population doubling time, (2) maximum degree of dilution upon splitting that is compatible with proper regrowth, (3) maximum cell density compatible with healthy growth and metabolism (confluence), and (4) minimum volume of medium that will sustain regrowth from a full split to confluence. If these parameters are properly defined, growth from full split to confluence will allow at least three population doublings during the labeling period for most cell lines. In some cases, however (e.g., with very slowly growing cell lines, lines requiring frequent medium changes, or cells that cannot be diluted to low density when splitting), this may not be feasible.

Materials

Radioactive precursor: ^3H - or ^{14}C -labeled monosaccharide, [^{35}S]sulfate, [^3H]acetate, or [^{32}P]orthophosphate; at highest available specific activity
Complete tissue culture medium appropriate for long-term growth of tissue culture cell line, supplemented as necessary
Established tissue culture cell line, either suspension or monolayer
Phosphate-buffered saline (PBS; *APPENDIX 2*), pH 7.2

Sterile plastic syringes with Luer-/Lok fittings
Disposable 0.22- μm filters with Luer-Lok fittings
Disposable flasks fitted with 0.22- μm filters
Tissue culture plates or flasks
Tabletop centrifuge, 4°C, and appropriate screw-cap centrifuge tubes
Rubber policeman or disposable cell scraper

Additional reagents and equipment for trypsinization (e.g., Chapter 9 introduction)

1. If necessary, dry radioactive precursor to remove organic solvents. Dissolve in a small volume of complete tissue culture medium.
2. Filter sterilize radioactive medium, using a standard disposable 50-ml vacuum suction filter device (filter flask fitted with 0.22- μm filter) for volumes ≥ 10 ml or a 0.22- μm filter attached directly to the tip of a disposable sterile syringe (both syringe and filter with Luer-Lok fittings) for smaller volumes. Wash out the original container with additional medium and use this fluid to wash through the filter (this maximizes recovery of label).

CAUTION: Dispose of the radioactive filter and/or syringe appropriately.

The radioactive precursor may be available in a sterile aqueous medium, ready for use. However, such preparations are more expensive, and after repeated opening of such packages, filtering is recommended to ensure continued sterility.

3. Add additional medium if necessary to bring up the volume to the amount needed for the experiment. Warm the medium to the temperature at which labeling will be done in an incubator or water bath.
4. Using a sterile pipet tip, aliquot a small amount of radioactive medium and save for scintillation counting in step 8.
5. Split cells according to standard method (e.g., detach by trypsinization), resuspend in radioactive medium, and plate out onto tissue culture plates. For suspension culture, dilute directly in radioactive medium or concentrate as follows: place culture in an appropriately sized screw-cap centrifuge tube and centrifuge in a tabletop centrifuge 5 min at $\sim 500 \times g$. Discard supernatant and resuspend cell pellet in radioactive medium. Incubate under standard conditions for the cell line being studied.

The number of cells needed, and therefore the culture vessels and volume of medium used, will depend on the growth characteristics of the particular cell line being studied and should be optimized as described above.

6. When the cells reach maximum growth (near confluence or late log phase), chill culture on ice and harvest cells, scraping monolayer cells from plate using a rubber policeman. Pellet cells 5 min at $\sim 500 \times g$, 4°C. Decant the labeled medium and save an aliquot for scintillation counting in step 8.

If the glycoconjugates in the conditioned medium are to be studied, the medium should be filtered through a 0.22- μm filter to remove any broken cell debris remaining after centrifugation.

7. Wash the cell pellet twice in a >50 -fold excess of ice-cold PBS, pH 7.2.

At this stage, the labeled cell pellet can be frozen at -20°C to -80°C for later analysis if desired.

8. Using a scintillation counter, determine the efficiency of incorporation by measuring the radioactivity incorporated into an aliquot of the cells and the sample of conditioned medium from step 6. As a control, measure the radioactivity of the sample of sterile radioactive medium from step 4.

It may be necessary to dilute the medium to get an accurate count.

ALTERNATE PROTOCOL

PULSE OR PULSE-CHASE LABELING WITH RADIOACTIVE PRECURSORS

If steady-state labeling (basic protocol) does not yield sufficient incorporation of label into the glycoconjugate of interest, it may be desirable to provide labeled precursor in the absence of or with the decreased amounts of the unlabeled form. The pulse procedure described in the following protocol—in which cells are cultured briefly in deficient medium supplemented with a radioactive precursor—may be used in these cases. Alternatively, the procedure can be expanded into a pulse-chase by adding a chase of nonradioactive medium; this permits brief labeling of the glycoconjugate of interest to establish precursor-product relationships. These procedures are most useful for labeling monosaccharides that compete with glucose for uptake—Gal, Glc, Man, and GlcNH₂ (see commentary). They are of limited value for monosaccharides that are not usually taken up efficiently by cells—GlcNAc, GalNAc, ManNAc, Neu5Ac, Fuc, Xyl, and GlcA—unless very large quantities of such labeled molecules can be used for pulse labeling. Because pulse labelings are typically done for a short time (e.g., minutes or hours), cells are usually used in a nearly confluent state, to maximize uptake and incorporation.

Optimizing Conditions

Before the experiment is performed, several pilot experiments should be carried out to determine the optimal conditions for labeling.

Determining base incorporation level and the effect of glucose. MDM medium appropriate for growing cultures of the cell line of interest should be reconstituted (support protocol) to 100% levels of all components except the one being presented as a radiolabeled precursor or competing molecule (e.g., glucose). If required by the cells, dialyzed serum should be added to the appropriate final concentration. Finally, radiolabeled precursor should be added. A small-scale pilot labeling should be carried out by culturing cells in the labeled medium and monitoring incorporation of the label into the macromolecule of interest or whole-cell glycoproteins. If the radiolabeled precursor is a monosaccharide, the effect of glucose on incorporation of radiolabeled precursor should be

assessed by carrying out this pilot experiment in duplicate using one lot of medium prepared as described and one prepared the same way but without glucose.

Determining optimal concentration of radiolabeled precursor. Enough MDM medium for several pilot experiments should be reconstituted to 100% levels of all components except the one being presented as a radiolabeled precursor. Radiolabeled precursor should be added and the medium divided into aliquots. The aliquots should be supplemented with an unlabeled 100× stock solution of the same precursor to yield a series of media containing different concentrations of the unlabeled precursor (e.g., 0%, 5%, 10%, 20%, 50%, and 100% of the concentration in normal medium). These should be used for small-scale pilot labelings of the cell line of interest for defined periods of time and incorporation of label monitored as described above. Incorporation should be plotted against the total precursor concentration. The point at which the curve breaks—i.e., where the percentage of label incorporated is markedly decreased by further addition of unlabeled compound—is the point at which the precursor concentration is no longer limiting for biosynthetic reactions (see Fig. 17.4.1). The optimal concentration of unlabeled precursor will be a little higher than the break point for the curve, where unlabeled precursor is not limiting but incorporation is still good.

Checking cell growth at optimal precursor concentration. Reconstituted MDM medium containing radioactive precursor plus the optimal concentration of unlabeled precursor should be prepared and cells grown in this medium for the desired period of time. Incorporation of radiolabeled precursor should be monitored and linearity of uptake and incorporation of label over time should be assessed. Measures of cell growth, stability, and general biosynthetic capability [e.g., cell counts, Strober (1991a); trypan blue exclusion, *UNIT 11.5*; and radioactive amino acid incorporation, *UNIT 10.18*] should also be assayed to determine if reducing the concentration of the precursor into this range has

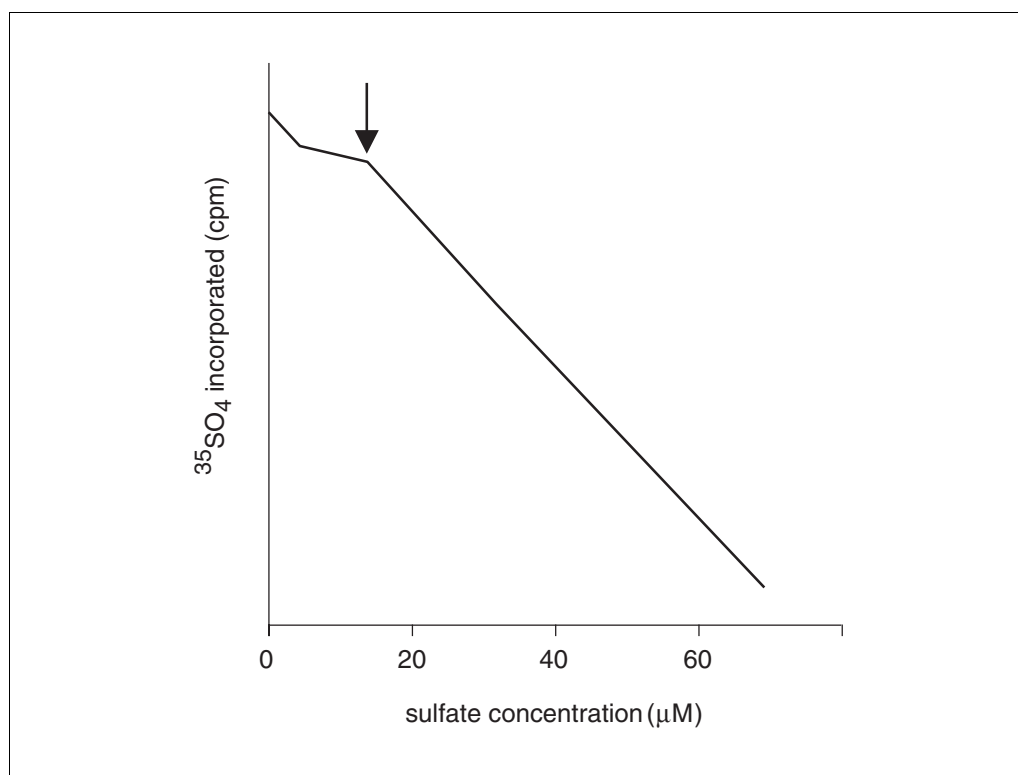


Figure 17.4.1 Incorporation of radiolabeled sulfates into macromolecules into cells labeled in sulfate-free medium and the effect of adding increasing amounts of unlabeled sulfate. The arrow indicates the point at which the sulfate is no longer limiting.

any detrimental effect upon the cells. It is sometimes necessary to reach a compromise between the opposing factors of lowered concentration, labeling time, and cell viability. An attempt should be made to determine the lowest concentration of the unlabeled precursor that can be used for the desired period of time without affecting cell growth, viability, and biosynthesis of macromolecules in general. It should be kept in mind that the use of partially deficient medium may selectively alter glycoconjugate makeup (e.g., lowering sulfate concentration too much can result in undersulfation of glycosaminoglycans in some cell types).

Additional Materials

Multiply deficient medium (MDM; Table 17.4.1 and support protocol),
supplemented as appropriate
Fetal calf serum (GIBCO/BRL), dialyzed (*APPENDIX 3*) against sterile
0.15 M NaCl until glucose concentration reaches $\sim 270 \mu\text{M}$, to remove
small molecules that could dilute the radioactivity

1. Determine optimal conditions for labeling as described above. Grow monolayer cells in complete medium to near confluence or suspension cells to late log phase.
2. Based on the pilot experiments for optimizing conditions described above, reconstitute MDM to prepare a stock of medium partially deficient in only the precursor of interest.
3. Add an appropriate quantity of radioactive precursor. Add dialyzed fetal calf serum if required for growth of the cell line being studied. Filter sterilize, warm the medium to 37°C , and reserve an aliquot of radioactive medium as described in steps 2 to 4 of the basic protocol.
4. For monolayer cells, aspirate culture medium from plate; for suspension cells, microcentrifuge or centrifuge briefly and decant the supernatant medium. Add an appropriate quantity of radioactive labeling medium (step 3). Incubate cells for the desired period of time.

Labeling time should be based on the length of time the cells can survive in deficient medium as determined in the optimizing pilot experiment.

Save radioactive medium for analysis, if necessary.

5. For pulse-labeling, proceed to step 6. For pulse-chase labeling, remove medium as described in step 4, add a chase of nonradioactive complete medium, and incubate for the desired period of time.

For a pulse-chase experiment, several vessels of each cell culture should be set up and chased for varying periods of time.

6. Harvest and wash the cells. Determine efficiency of incorporation as described in steps 6 to 8 of the basic protocol.

SEQUENTIAL PULSE OR PULSE-CHASE LABELING WITH REUSE OF RADIOACTIVELY LABELED MEDIUM

ALTERNATE PROTOCOL

In some cases, incorporation of label into the glycoconjugate of interest may be inadequate, even under defined conditions with selectively deficient medium. Also, prolonged exposure to deficient medium may result in alterations in synthesis of proteins or other macromolecules. In these situations, it may be desirable to expose a series of plates of cells sequentially to a small volume of medium containing a high concentration of label. This also conserves potentially expensive radiolabeled precursor. In the following protocol, multiple sets of cell cultures are sequentially pulse-labeled. One set is incubated in labeling medium for the pulse period, then the labeling medium is removed and added to the next set. For a pulse experiment (designed to obtain a sizable quantity of material with a high level of incorporation), the cells are harvested at this point; for a pulse-chase, a chase of nonradioactive complete medium is added to the first set of cultures. This process is repeated until all cells have been labeled (Fig. 17.4.2). Because the incubation period is relatively short (e.g., a few hours), it is assumed that only a very small fraction of the radiolabeled precursor is consumed from the medium during each labeling cycle and that the concentrations of other essential components are not substantially changed. To further conserve labeled medium, some investigators freeze it after limited usage and thaw it to reuse (the pH of thawed medium should be adjusted before use). When reusing medium, experimental results should be interpreted carefully, especially with regard to the specific activity of labeled products and the possibility of metabolic effects of secreted molecules.

Additional Materials

Multiply deficient medium (MDM; Table 17.4.1 and support protocol),
supplemented as appropriate

Fetal calf serum (GIBCO/BRL), dialyzed (*APPENDIX 3*) against sterile 0.15 M NaCl
until glucose concentration reaches $\sim 270 \mu\text{M}$, to remove small molecules that
could dilute the radioactivity

1. Determine optimal conditions for labeling as described in the introduction to the previous protocol. Set up several identical sets of monolayer or suspension cells to be labeled sequentially and grow to confluence or late log phase.

For suspension cultures in which it is possible to use <1 ml of labeling medium, cells may be grown in tightly capped 1.5-ml screw-cap microcentrifuge tubes and incubated at 37°C on a rotating end-over-end mixer.

2. Prepare radioactive labeling medium and label the first set of cell cultures as described in steps 2 to 4 of the previous protocol for the appropriate length of time.

Labeling time will vary from minutes to hours, depending upon the cell type and the objective of the labeling.

3. For monolayer cells, aspirate the radioactive medium from the first set of cultures; for suspension cells, microcentrifuge briefly (~ 10 sec at top speed) and decant the supernatant radioactive medium. Transfer the radioactive medium to the second set of cultures.
4. For pulse-labeling, harvest cells from first set of cultures. For pulse-chase labeling, add a chase consisting of an adequate quantity of nonradioactive complete culture medium to the first set of cultures and continue incubating.
5. Repeat transfers until all sets of cells have been labeled.
6. Harvest all plates. For pulse-labeling, pool all the pellets together. For pulse-chase labeling, keep the individual cell pellets separate (see Fig. 17.4.2).

Preparation and Analysis of Glycoconjugates

17.4.7

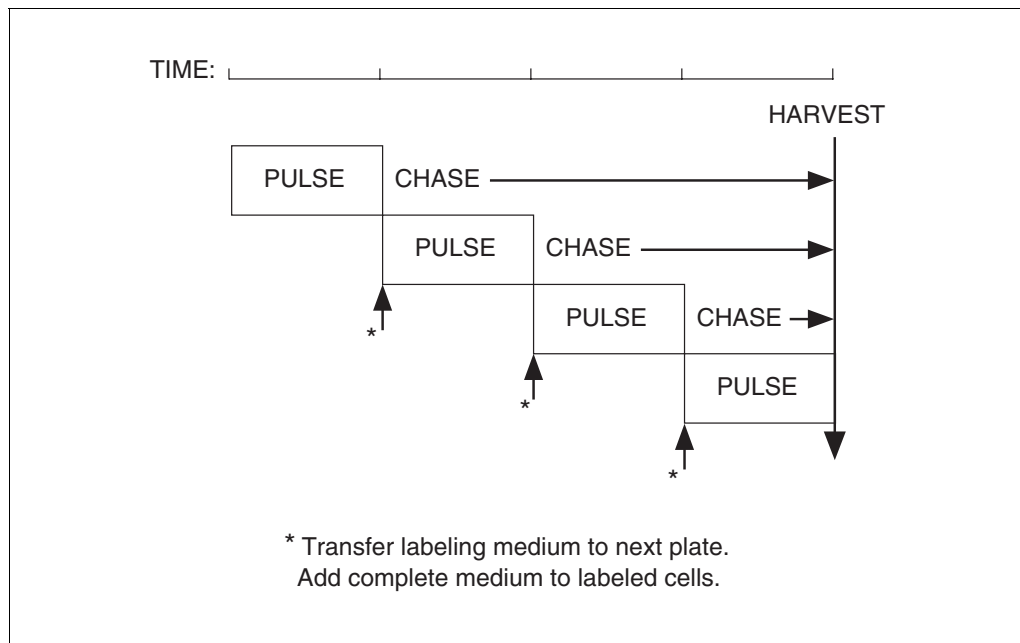


Figure 17.4.2 Scheme for sequential pulse-labeling of cells reutilizing radioactive medium.

**SUPPORT
PROTOCOL**

**PREPARATION AND SUPPLEMENTATION OF MULTIPLY
DEFICIENT MEDIUM (MDM)**

Some types of selectively deficient media are available commercially at reasonable prices, and some companies will custom-prepare selectively deficient media on request. For a laboratory where labeling experiments with a variety of deficient media are frequently carried out, it is convenient to prepare a basic multiply deficient medium (MDM) that is completely lacking in several commonly studied components. This medium can be used to make up different selectively deficient media as needed. The following MDM, based on α -MEM medium (GIBCO/BRL), supports growth of most tissue culture cells.

For labeling experiments, MDM can be reconstituted with ^3H - or ^{14}C -labeled monosaccharides, [^{35}S]sulfate, ^{35}S -labeled methionine or cysteine, or [^3H]serine. If desired, other specific components can also be omitted and replaced with their radiolabeled forms.

NOTE: Tissue culture grade reagents (including distilled, deionized water) should be used in making MDM. To ensure that reagents do not become contaminated by compounds such as endotoxin as portions are being removed from stock bottles, it is important to pour liquids carefully and use disposable spatulas for removing solids. For the same reason, clean glassware must be used; preferably, a set of glassware should be set aside for this purpose. 100 \times stock solutions of many components can be purchased commercially in sterile form, or may be made up from solids and filter sterilized individually.

Additional Materials

- Stock solutions for multiply deficient medium (MDM; Table 17.4.1)
- 100 \times stock solutions for reconstituting MDM (Table 17.4.2)

1. Based on the experiment to be performed, determine appropriate components for MDM (see box below for example).
2. Make up MDM according to the ingredient list in Table 17.4.1 by dissolving salts and phenol red one by one in ~800 ml water, adding sufficient pyruvate, amino acid, and vitamin stocks for a 1 \times final concentration in 1 liter, and adding water to 1 liter.

SELECTIVE RECONSTITUTION OF MDM MEDIA WITH MISSING COMPONENTS

The following example gives 50 ml of a supplemented MDM for optimal labeling of endothelial cells with [³⁵S]sulfate and [6-³H]glucosamine, which requires a medium with no cysteine, low glucose, and low concentrations of methionine and sulfate (see Roux et al., 1988, for detailed rationale). Stock solutions are as listed in Table 17.4.2.

To 50 ml MDM, add:

500 μl 100× NaHCO₃ stock solution

500 μl 100× glutamine stock solution

500 μl 100× serine stock solution

50 μl 100× glucose stock solution (200 mg/L final)

10 μl 100 mM inorganic sulfate stock solution (20 μM final)

50 μl 100× methionine stock solution (0.1× final)

Just before use, add 0.5 mCi each of [³⁵S]sulfate and [6-³H]glucosamine and mix. Filter sterilize directly into culture vessel.

The total volume of reconstituted medium will be slightly more than 50 ml; for practical purposes, this minor discrepancy can be ignored. The medium should be orange-red, indicating the correct pH range.

3. Filter sterilize (the solution will be yellow). Do not adjust the pH. Freeze 25-ml aliquots in 50-ml tubes at -20°C.
4. Add appropriate components (from those listed in Table 17.4.2) to reconstitute MDM for the labeling experiment (see box for example).

COMMENTARY

Background Information

If sufficient quantities of pure molecules are available (i.e., in the nmol range), complete sequencing of oligosaccharides is best performed by conventional physical techniques. However, isolation of sufficient quantities may not be practical (e.g., in analysis of biosynthetic intermediates or rare molecules). Alternatively, the biological questions underlying the experiment may be adequately answered by partial structural analysis. In these cases, metabolic labeling with radioactive sugars or donors that can transfer label to sugars may provide sufficient structural information about the oligosaccharide chains (Cummings et al., 1989).

Sugar chains of glycoconjugates can be successfully labeled by short-term labeling experiments using radioactive precursors (Tabas and Kornfeld, 1980; Goldberg and Kornfeld, 1981; Roux et al., 1988; Muchmore et al., 1989). Although sugar nucleotides are the immediate donors for glycosylation reactions, they cannot be taken up by cultured cells. Hence, metabolic labeling of sugar chains is accomplished by

providing the cells with radiolabeled monosaccharide precursors (Yurchenco et al., 1978; Yanagishita et al., 1989; Varki, 1991). These are taken up by the cells, activated to sugar nucleotides, and transported into the Golgi apparatus, where lumenally oriented transferases add the monosaccharides to lumenally oriented acceptors (Hirschberg and Snider, 1987). The few known exceptions to this topology are cytosolic and nuclear forms of glycosylation (Hart et al., 1989). For oligosaccharides that contain modifications such as sulfate, acetate, and phosphate ester groups, an alternative metabolic labeling technique is to use a [³⁵S]sulfate, [³H]acetate, or [³²P]orthophosphate label, respectively.

While metabolic labeling can provide useful information regarding glycoprotein oligosaccharides, there are significant limitations to its use. First, it is difficult to determine when true steady-state labeling of a cell is reached (as a rule of thumb, 3 to 4 doublings are usually assumed to be sufficient); thus the numerical ratio between labeled glycoconjugates can be

Table 17.4.1 Composition of Multiply Deficient Medium (MDM)

Component	Final concentration in complete medium	Stock ^a
CaCl ₂	200 mg/l	Salt, reagent grade
KCl	400 mg/l	Salt, reagent grade
MgCl ₂ ^b	75 mg/l	Salt, reagent grade
NaCl	6800 mg/l	Salt, reagent grade
NaH ₂ PO ₄ ·H ₂ O	140 mg/l	Salt, reagent grade
Phenol red	10 mg/l	10×
Sodium pyruvate	110 mg/l	100×
L-Alanine	25 mg/l	100×
L-Arginine	126 mg/l	100×
L-Asparagine	50 mg/l	100×
L-Aspartic acid	30 mg/l	100×
L-Cysteine	NONE ^c	—
L-Glutamic acid	75 mg/l	100×
L-Glutamine	NONE ^c	—
L-Glycine	50 mg/l	100×
L-Histidine	42 mg/l	100×
L-Isoleucine	52 mg/l	100×
L-Leucine	52 mg/l	100×
L-Lysine	72 mg/l	100×
L-Methionine	NONE ^c	—
L-Phenylalanine	32 mg/l	100×
L-Proline	40 mg/l	100×
L-Serine	NONE ^c	—
L-Threonine	48 mg/l	100×
L-Tryptophan	10 mg/l	100×
L-Tyrosine	36 mg/l	100×
L-Valine	46 mg/l	100×
MEM vitamins (mixture)	1×	100×

^a10× and 100× stock solutions are available commercially (e.g., from GIBCO/BRL).

^bDo not use MgSO₄ in place of MgCl₂.

^cNot present in MDM; add as needed for specific experiments.

Table 17.4.2 Stock Solutions for Reconstitution of MDM^a

Component	Final concentration in complete medium	Stock ^b
NaHCO ₃ ^c	1× (2200 mg/l)	100×
HEPES·HCl	20 mM (4.76 g/l)	2 M, pH 7.3
Na ₂ SO ₄	0.81 mM (115 mg/l)	100 mM, sterile
D-Glucose	1× (1000 mg/l)	100×
L-Cysteine	1× (100 mg/l)	100×
L-Glutamine	1× (292 mg/l)	100×
L-Methionine	1× (15 mg/l)	100×
L-Serine	1× (25 mg/l)	100×

^aIndividual components are added to MDM at full strength, or lowered concentration or may be left out altogether, depending upon the experiment planned.

^b100× stock solutions are available commercially (e.g., from GIBCO/BRL).

^cUse either NaHCO₃/CO₂ or HEPES·HCl (not both); these control the pH of the final medium.

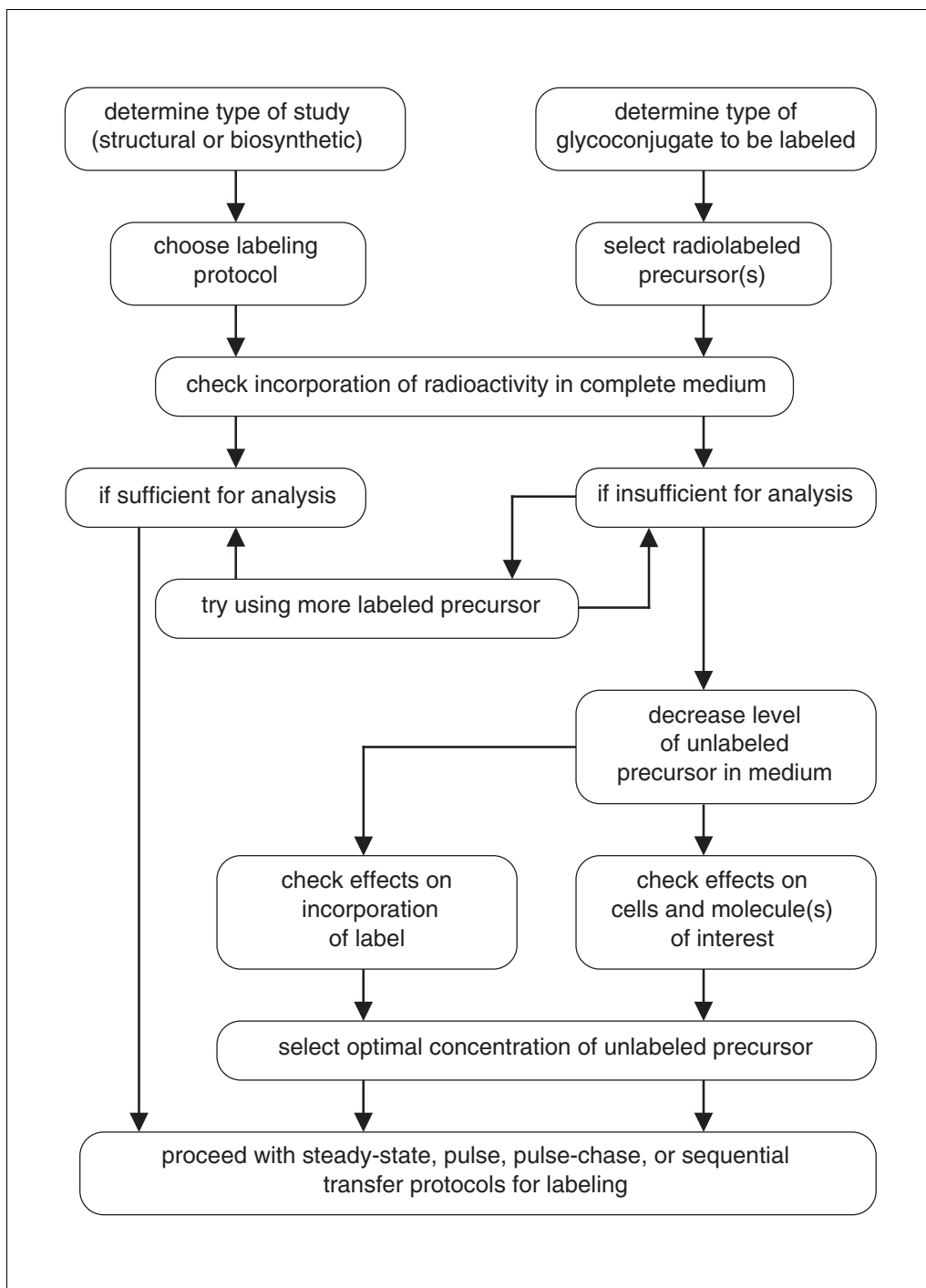


Figure 17.4.3 Strategy for planning metabolic labeling of animal cell glycoconjugates.

misleading. Second, individual precursors show greatly differing uptake and incorporation in different cell types. Third, almost all labeled precursors are only partially specific for certain monosaccharides, and the degree of this specificity can vary depending on cell type. Fourth, while lowering glucose concentration improves the incorporation of monosaccharides that compete with glucose for uptake, the

low glucose supply may also directly affect oligosaccharide precursors in some cell types.

Critical Parameters

As diagrammed in Figure 17.4.3, what labeling protocol is most suitable for a particular study depends on several different considerations. Some of these arise from the objectives of the study and must be investigated by per-

Table 17.4.3 Distribution of Monosaccharides and Modifications in Glycoconjugates^a

Precursor	Type of glycoconjugate					
	<i>N</i> -GlcNAc-linked glycoprotein	<i>O</i> -GalNAc-linked glycoprotein	Xylose-linked proteoglycan	Glyco-sphingolipid	Glyco-phospholipid anchor	O-linked GlcNAc
Man	+++	–	–	–	+++	–
Fuc	++	++	+/-	+	–	–
Gal	+++	+++	++	+++	+/- (side chain)	–
Glc	+ (precursor)	–	–	++	–	–
GlcNAc	+++	+++	++	++	+ (free amine)	++
GalNAc	+/-	+++	++	++	+/- (side chain)	–
Sia	+++	+++	–	+++	–	–
GlcA	+?	–	+++	+	–	–
SO ₄	+	+	++++	+	–	–
P _i esters	Man-6-P	–	Xyl-P	–	M6P (in core)	–
<i>O</i> -Acetyl Acyl	Sia-OAc –	Sia-OAc –	– –	Sia-OAc ++++	– +/- (on inositol)	– –

^aThe relative distribution of the different monosaccharides and modifications in the commonly occurring glycoconjugates, indicated by a relative scale from + to +++, is generally valid over many cell types. However, an uncommon monosaccharide may be commonly found in some cell types or glycoconjugate (e.g., most pituitary glycoprotein hormones have GalNAc as a major component of their *N*-linked oligosaccharides). Other symbols: +?, possible distribution; +/-, variably present; –, not identified to date.

forming pilot experiments to determine optimal conditions (see optimizing conditions section of introductions to the first two protocols).

General considerations

If the goal of a study is to obtain labeled material for structural identification and partial characterization, the yield of radioactivity should be maximized by either steady-state or pulse-labeling. For quantitating the relative mass of a molecule in two different samples or comparing the masses of two molecules in the same sample, the oligosaccharide should be labeled to constant specific activity—i.e., steady-state distribution (however, as mentioned above, this may not always be achievable). For establishing precursor-product relationships between molecules, a pulse-chase protocol should be used. If the radioactive precursor is expensive, a sequential procedure may be employed in pulse and pulse-chase experiments to minimize the quantity needed.

Selecting a labeled precursor

Selection of the labeled precursor to be used is based upon several factors, including efficiency of uptake (see below) and type of gly-

coconjugate to be labeled. In most cases, the precursor will be a monosaccharide; for an oligosaccharide chain containing ester modifications of sugars, it is also possible to use a precursor such as acetate, sulfate, or orthophosphate.

Factors affecting specificity and final distribution. In selecting a monosaccharide precursor, it should be noted that the distribution of monosaccharides among different types of vertebrate oligosaccharides is nonrandom (Table 17.4.3). It is also important to consider metabolic pathways for uptake, activation, utilization, and interconversion of the various monosaccharides and their nucleotide sugars, which have been studied extensively (Fig. 17.4.4 and Table 17.4.4). The final distribution and specific activity of label from a given radiolabeled monosaccharide can be significantly affected by dilution from unlabeled compound generated by endogenous pathways, the characteristics of the particular cell type being labeled, and the labeling conditions (Kim and Conrad, 1976; Yurchenco et al., 1978; Yanagishita et al., 1989). The exact position of the tritium label within a monosaccharide can affect the ultimate fate of the label

Table 17.4.4 Cellular Monosaccharides Labeled with Radioactive Precursors^a

Product identified as:	Radiolabeled precursor						
	[2- ³ H]Man	[6- ³ H]Gal	[1- ³ H]Gal	[1- ³ H]Glc	[6- ³ H]GlcNH ₂	[6- ³ H]ManNAc	[6- ³ H]GalNAc
[³ H]Man	+++	+/-	+/-	++	+/-	+/-	+/-
[³ H]Fuc	+++	+/-	+/-	++	+/-	+/-	+/-
[³ H]Gal	-	++++	++++	+++	+/-	+/-	+/-
[³ H]GlcA	-	-	++	++	-	-	-
[³ H]Glc	-	++	++	++++	+/-	+/-	+/-
[³ H]GlcNAc	-	+/-	+/-	++	++++	++	+++
[³ H]GalNAc	-	+/-	+/-	++	+++	+	++++
[³ H]Sia	-	-	-	++	++	++++	+/-

^aThe extent of conversion from the original monosaccharide into minor pathways will vary considerably, depending upon the cell type studied, and the length of the labeling (longer time periods tend to result in more conversion into other monosaccharides). The extent of incorporation is indicated by a relative scale from + to +++++. Other symbols: +/-, variable low-level incorporation; -, not identified to date.

(Diaz and Varki, 1985). Interconversion between monosaccharides (e.g., conversion to glucose; Fig. 17.4.4), which can be expected to occur with prolonged labeling, will eventually result in labeling of nonoligosaccharide components. An exception is labeling with [2-³H]mannose: this is extremely specific, because conversion of the labeled mannose can yield only [2-³H]fucose or unlabeled fructose-6-phosphate. In the latter case, the label is lost as tritiated water, which is diluted into the pool of cellular water. Thus, in most instances, label from [2-³H]mannose remains confined to mannose and fucose residues, regardless of how long the labeling continues.

Factors affecting uptake and incorporation. With radioactive amino acids, high specific activity labeling can be obtained by omitting the unlabeled molecule from the medium (see UNIT 10.18). Labeling with radioactive sugars or other oligosaccharide precursors, however, is usually less efficient. Thus, incorporation into the glycoconjugate of interest must be optimized empirically, taking into consideration the following factors.

Experimental parameters. Amount of label, concentration of label in the medium, cell number, duration of labeling, and number of cell doublings during labeling are obviously important. The goal is to expose the maximum number of cells to the maximum amount of label in the minimum volume of medium for the longest period of time. Some of these factors are at odds with one another, and the correct balance must

be tailored to the particular cell type and the question being investigated.

Competition with glucose. When labeling using a radioactive monosaccharide precursor, it is important to determine whether the monosaccharide competes with glucose for uptake. The high concentration of glucose in normal tissue culture medium (~5 mM) means that monosaccharides competing with glucose for active transport cannot be taken up very well. Thus, lowering glucose concentration can improve uptake. On the other hand, monosaccharides that do not compete with glucose seem to be taken up only by inefficient noncompetitive, passive mechanisms; thus, glucose concentration has little effect upon their already poor uptake. Experience indicates that glucosamine, galactosamine, galactose, and mannose can compete with glucose for uptake into most cells, whereas *N*-acetylglucosamine, *N*-acetylmannosamine, mannosamine, fucose, and xylose do not (Yurchenco et al., 1978). However, members of the glucose transporter family of gene products are tissue-specific in expression; a given cell type may express more than one, each with distinctive kinetic and stereospecific uptake properties (Gould and Bell, 1990). Thus, each monosaccharide must be tested in the specific cell type under study to see if it competes with glucose. A final point that should be considered is that in some cell types, lipid-linked oligosaccharide precursors may become altered when glucose-free medium is used (Rearick et al., 1981).

Other intracellular metabolic factors. Dilution of label by endogenously synthesized monosaccharides, pool size of individual monosaccharides and nucleotides, and flux rates between interconverting pathways can all affect the final specific activity and distribution of the label. If cells are cultured in labeled medium for prolonged periods of time, the medium may become depleted of glucose (Kim and Conrad, 1976; Varki and Kornfeld, 1982). The specific activity of the labeled sugar nucleotides in the cells may then actually rise as the medium glucose concentration falls below the K_m of the cellular glucose transporter(s).

These factors must be considered in designing the labeling protocol, although they can be controlled only to a limited extent. The goal is to obtain sufficient label in the glycoconjugate of interest without significantly altering the metabolic state of the cell.

Metabolic labeling using nonmonosaccharide precursors. Acetylation, sulfation, and phosphorylation can significantly affect the behavior of oligosaccharides both in biological systems and during analysis. Labeled precursors such as [^3H]acetate, [^{35}S]sulfate, and [^{32}P]orthophosphate can be used to label these modified oligosaccharides. These

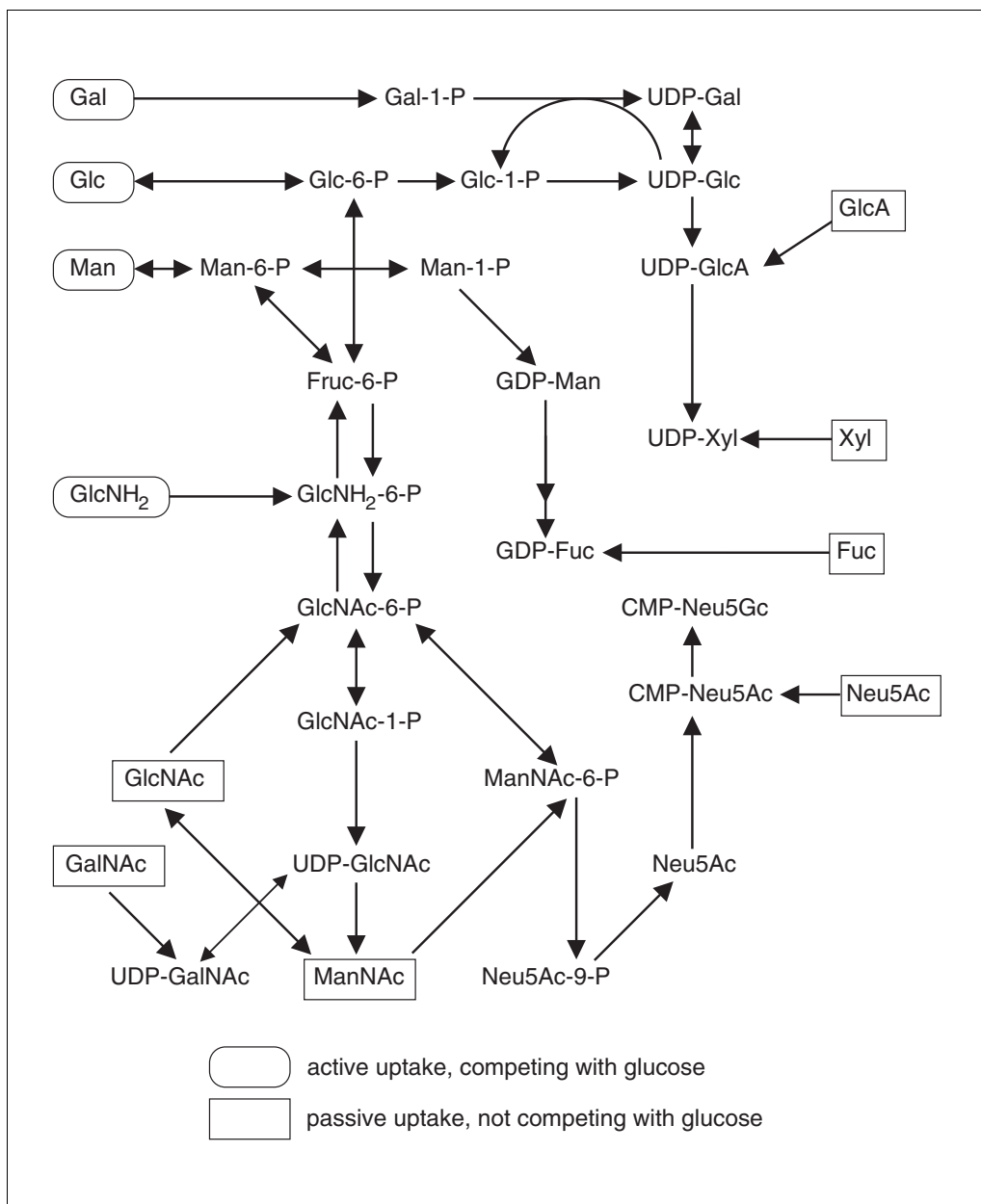


Figure 17.4.4 Cytosolic pathways for the interconversion of monosaccharides and their nucleotide sugar forms.

precursors are expected to enter a variety of other cellular macromolecules, and release and/or isolation of the labeled oligosaccharides with specific endoglycosidases is often required before further analysis. Double-labeling of both the modified oligosaccharide and the underlying sugar chain (e.g., with a ^{14}C - or ^3H -labeled monosaccharide precursor) can be quite useful in monitoring purification and in subsequent structural analysis. With [^{35}S]sulfate, labeling efficiency varies widely with cell type. In many cells, if the endogenous pool of the precursor 3'-phosphoadenosine-5'-phosphosulfate (PAPS) is very small, the specific activity of the exogenously added [^{35}S]sulfate can be practically unaltered inside the cell (Yanagishita et al., 1989). In some cell types, however, the endogenous sulfate pool is constantly diluted by breakdown of the sulfur-containing amino acids methionine and cysteine (Esko et al., 1986; Roux et al., 1988). In this situation, reducing the cysteine and methionine concentrations in the medium can improve incorporation of label into macromolecules (Roux et al., 1988). Sulfate levels can also be affected by certain antibiotics (e.g., gentamicin) that are sulfate salts.

Determining the specific activity of incorporated label

In most instances, the precise specific activity of each monosaccharide pool need not be determined. If steady-state labeling is attempted, a plateau in the rate of incorporation of label per milligram of cell protein can be taken as a rough indication that a steady state has been reached. In some cases, however, the precise specific activity of a given monosaccharide may be of interest. Discussions of how this can be determined can be found in Kim and Conrad (1976), Yurchenco et al. (1978), Yanagishita et al. (1989), and Varki (1991). Endogenous glucose is the normal precursor for hexoses and hexosamines (see Fig. 17.4.4). Experimental manipulations that alter the concentration of glucose can thus alter the concentration of the internal pool of other hexoses and hexosamines. Exogenously added labeled hexosamines become diluted within the cell, making specific radioactivity in the cell lower than that of the starting material.

With the advent of HPAE-PAD (High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection), measurement of monosaccharide levels in the low picomole range is now possible. Thus, it is now feasible to measure the specific activity of a

labeled monosaccharide component of a purified oligosaccharide by acid hydrolysis of a portion of the oligosaccharide (UNIT 17.15) followed by HPAE-PAD. HPLC-based methods for accurate measurement of low quantities of sulfate are also now available.

Troubleshooting

Ideally, the radiochemical purity of the precursor should be checked using an appropriate chromatographic procedure (see UNIT 17.16). As with all tissue culture work, cells should be checked for mycoplasma infection (Strober, 1991b) before the experiments are carried out. If bacterial contamination poses a problem for long-term labeling experiments, appropriate antibiotics should be added to the medium.

Anticipated Results

Levels of incorporation will be highly variable, depending upon the cell type, the culture conditions, and the precursor used. If adequate incorporation into the glycoconjugate of interest is obtained, analyses can be carried out.

Time Considerations

Optimizing conditions for uptake and incorporation of label for each precursor and cell line combination (introductions to the first two protocols) may take one to two weeks, but should not be omitted: the time will be well-spent. The time required for the actual labeling protocols will vary widely depending on the goal of the experiment. Preparation of multiply deficient media can take an entire day, especially if some components must be prepared from scratch. If the amount required for a series of experiments is planned ahead of time, however, this need only be done occasionally. Reconstituting multiply deficient medium takes up to 1 hr (assuming all the stock solutions are already available).

Literature Cited

- Cummings, R.D., Merkle, R.K., and Stults, N.L. 1989. Separation and analysis of glycoprotein oligosaccharides. *Methods Cell Biol.* 32:141-183.
- Diaz, S. and Varki, A. 1985. Metabolic labeling of sialic acids in tissue culture cell lines: Methods to identify substituted and modified radioactive neuraminic acids. *Anal. Biochem.* 150:32-46.
- Esko, J.D., Elgavish, A., Prasthofer, T., Taylor, W.H., and Weinke, J.L. 1986. Sulfate transport-deficient mutants of Chinese hamster ovary cells. Sulfation of glycosaminoglycans dependent on cysteine. *J. Biol. Chem.* 261:15725-15733.

- Goldberg, D. and Kornfeld, S. 1981. The phosphorylation of beta-glucuronidase oligosaccharides in mouse P388D1 cells. *J. Biol. Chem.* 256:13060-13067.
- Gould, G.W. and Bell, G.I. 1990. Facilitative glucose transporters: An expanding family. *Trends Biochem. Sci.* 15:18-23.
- Hart, G.W., Haltiwanger, R.S., Holt, G.D., and Kelly, W.G. 1989. Glycosylation in the nucleus and cytoplasm. *Annu. Rev. Biochem.* 58:841-874.
- Hirschberg, C.B. and Snider, M.D. 1987. Topography of glycosylation in the rough endoplasmic reticulum and Golgi apparatus. *Annu. Rev. Biochem.* 56:63-87.
- Kim, J.J. and Conrad, E.H. 1976. Kinetics of mucopolysaccharide and glycoprotein biosynthesis by chick embryo chondrocytes. Effect of D-glucose concentration in the culture medium. *J. Biol. Chem.* 251:6210-6217.
- Muchmore, E.A., Milewski, M., Varki, A., and Diaz, S. 1989. Biosynthesis of *N*-glycolyl-neuraminic acid: The primary site of hydroxylation of *N*-acetylneuraminic acid is the cytosolic sugar nucleotide pool. *J. Biol. Chem.* 264:20216-20223.
- Rearick, J.I., Chapman, A., and Kornfeld, S. 1981. Glucose starvation alters lipid-linked oligosaccharide biosynthesis in Chinese hamster ovary cells. *J. Biol. Chem.* 256:6255-6261.
- Roux, L., Holoyda, S., Sunblad, G., Freeze, H.H., and Varki, A. 1988. Sulfated N-linked oligosaccharides in mammalian cells I: Complex-type chains with sialic acids and O-sulfate esters. *J. Biol. Chem.* 263:8879-8889.
- Strober, W. Monitoring cell growth. 1991a. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. A.3.1-A.3.3. John Wiley & Sons, New York.
- Strober, W. Diagnosis and treatment of mycoplasma-contaminated cell cultures. 1991b. *In* Current Protocols in Immunology (J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, and W. Strober, eds.) pp. A.3.9-A.3.12. John Wiley & Sons, New York.
- Tabas, I. and Kornfeld, S. 1980. Biosynthetic intermediates of beta-glucuronidase contain high mannose oligosaccharides with blocked phosphate residues. *J. Biol. Chem.* 255:6633-6639.
- Varki, A. 1991. Radioactive tracer techniques in the sequencing of glycoprotein oligosaccharides. *FASEB J.* 5:226-235.
- Varki, A. and Kornfeld, S. 1982. The spectrum of anionic oligosaccharides released by endo- β -*N*-acetylglucosaminidase H from glycoproteins. *J. Biol. Chem.* 258:2808-2818.
- Yanagishita, M., Salustri, A. and Hascall, V.C. 1989. Specific activity of radiolabeled hexosamines in metabolic labeling experiments. *Methods Enzymol.* 179:435-445.
- Yurchenco, P.D., Ceccarini, C., and Atkinson, P.H. 1978. Labeling complex carbohydrates of animal cells with monosaccharides. *Methods Enzymol.* 50:175-204.

Contributed by Sandra Diaz and Ajit Varki
University of California San Diego
La Jolla, California

Chemical Labeling of Carbohydrates by Oxidation and Sodium Borohydride Reduction

UNIT 17.5

This unit describes a collection of methods for chemical labeling of carbohydrates—free oligosaccharides or oligosaccharides conjugated to proteins, peptides, or lipids—by oxidation followed by reduction or by direct reduction. Oligosaccharides can be labeled with either radioisotopes or nonradioactive fluorescent molecules. These labelings allow one to follow the oligosaccharides during chromatography and in cells if labeled by fluorescent molecules.

Selective oxidation with mild periodate followed by reduction with tritiated sodium borohydride ($\text{NaB}[^3\text{H}]_4$) results in selective radiolabeling of sialic acid residues on oligosaccharides or glycoproteins (first basic protocol). Alternatively, treatment of samples with galactose oxidase (first alternate protocol) results in oxidation of galactose or *N*-acetylgalactosamine residues at nonreducing termini, rendering these residues susceptible to labeling with $\text{NaB}[^3\text{H}]_4$. Oxidized glycoconjugates can also be labeled using the fluorescent probe lucifer yellow CH (second alternate protocol). Free oligosaccharides can be labeled by reduction with $\text{NaB}[^3\text{H}]_4$ (second basic protocol). The third alternate protocol describes the release and simultaneous labeling of *O*-glycan oligosaccharides by alkaline beta-elimination in the presence of $\text{NaB}[^3\text{H}]_4$.

CAUTION: This procedure should be performed only by personnel trained in the proper use of ^3H isotope and in NRC-licensed sites. Standard precautions to prevent excessive exposure and radioactive contamination of personnel and equipment should be followed at all times.

CAUTION: All $\text{NaB}[^3\text{H}]_4$ manipulations must be performed in a well-vented fume hood.

RADIOLABELING OLIGOSACCHARIDES AFTER MILD PERIODATE OXIDATION

**BASIC
PROTOCOL**

This protocol describes radiolabeling oligosaccharides by mild periodate oxidation followed by reduction with $\text{NaB}[^3\text{H}]_4$, selectively labeling sialic acid residues at nonreducing termini. The reaction is illustrated in Figure 17.5.1. Labeled oligosaccharides are precipitated and separated from unincorporated label by gel-filtration chromatography.

Materials

For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

- Oligosaccharide or glycopeptide containing sialic acids, desalted
- 0.2 mM sodium periodate (prepare fresh at room temperature and keep on ice)
- 10 mM sodium acetate, pH 5.5
- 10 mM glycerol (store at 4°C)
- 0.2 M sodium borate buffer, pH 9.5 (UNIT 11.16; adjust quantities of boric acid and 10 M NaOH appropriately)
- 0.5 mCi/ μl tritiated sodium borohydride ($\text{NaB}[^3\text{H}]_4$; 25 to 30 Ci/mmol) in 0.01 M NaOH (see recipe)
- 0.2 M sodium borohydride (NaBH_4)/0.2 M sodium borate buffer, pH 9.5
- 1 M acetic acid in methanol
- Nitrogen (N_2) stream
- Methanol

**Preparation and
Analysis of
Glycoconjugates**

Contributed by Minoru Fukuda

Current Protocols in Molecular Biology (1994) 17.5.1-17.5.8

Copyright © 2000 by John Wiley & Sons, Inc.

17.5.1

Supplement 26

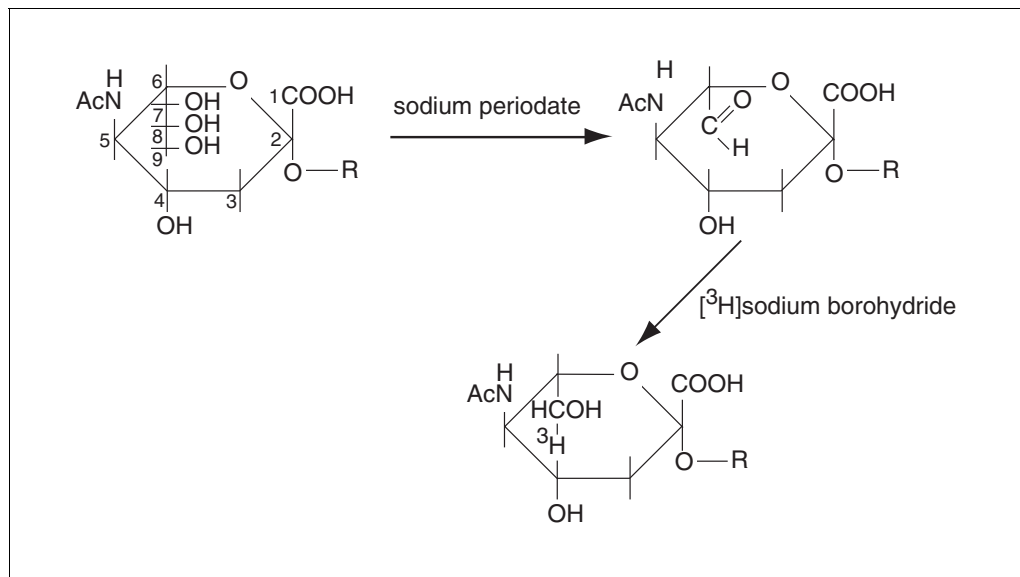


Figure 17.5.1 Labeling oligosaccharides by periodate oxidation followed by $\text{NaB}[^3\text{H}]_4$ reduction.

7-ml conical glass (Pyrex) test tubes with Teflon-covered caps
 1.0 × 40-cm Sephadex G-15 or G-25 column in water (*UNIT 10.9*), calibrated to identify void volume
 25° to 30°C water bath

CAUTION: Steps 3 to 8 must be performed in a well-vented fume hood to prevent radioactive contamination. Excess $\text{NaB}[^3\text{H}]_4$ should be stored away in a radioactive waste bottle for appropriate disposal.

Label oligosaccharides

- Mix the following in a 7-ml conical glass test tube (200 μl total volume):
 - 5 to 500 μg oligosaccharide or glycopeptide
 - 100 μl 0.2 mM sodium periodate
 - 100 μl 10 mM sodium acetate.

Loosely cover with Teflon-covered cap and incubate 20 min on ice.

- Add 10 μl of 10 mM glycerol and incubate 20 min at 37°C.

The glycerol consumes any excess periodate that remains.

- Add 100 μl of 0.2 M sodium borate buffer and 10 μl of 0.5 mCi/ μl $\text{NaB}[^3\text{H}]_4$ (5 mCi). Incubate 1 hr at room temperature in a fume hood.
- Add 100 μl of 0.2 M NaBH_4 /0.2 M borate buffer and incubate 30 min at room temperature.

Destroy excess $\text{NaB}[^3\text{H}]_4$ and remove borate as methyl borate

- Add 3 ml of 1 M acetic acid in methanol and dry under a nitrogen stream in a 25° to 30°C water bath, in a well-vented hood.
- Repeat step 5.
- Add 3 ml methanol (without acetic acid) and dry sample as in step 5.

Purify labeled oligosaccharides

- Dissolve sample in ~200 μl water and apply to a calibrated $1.0 \times 40\text{-cm}$ Sephadex G-15 or G-25 column. Elute with water and collect 1-ml fractions with a fraction collector.

This step separates the labeled oligosaccharides from the low-molecular-weight breakdown products of $\text{NaB}[^3\text{H}]_4$.

- Count 10- μl aliquots of each fraction in a scintillation counter.
- Pool and dry fractions that contain radioactive oligosaccharides or glycopeptides. Store at -20°C until used (the label is stable for >12 months).

Oligosaccharides or glycopeptides should elute at or near the void volume, followed by a larger peak of breakdown products of $\text{NaB}[^3\text{H}]_4$. Be careful to avoid the second peak.

RADIOLABELING OLIGOSACCHARIDES AFTER GALACTOSE OXIDASE TREATMENT

ALTERNATE PROTOCOL

Oxidation with galactose oxidase and reduction with $\text{NaB}[^3\text{H}]_4$ specifically labels galactose and *N*-acetylgalactosamine residues that are located at nonreducing termini of oligosaccharides. Residues at internal positions are also oxidized and labeled, but much less efficiently. This reaction is illustrated in Figure 17.5.2. Sialic acid attached to galactose hinders oxidation by galactose oxidase. If desired, sialic acid can be removed by treating the sample with sialidase (neuraminidase).

Additional Materials

For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

- 1 U/ μl galactose oxidase (partially purified, Sigma; store aliquots in PBS at -20°C)
- 10 mM sodium phosphate buffer, pH 7.0
- 15-ml conical glass centrifuge tube

CAUTION: Steps 2 to 4 must be performed in a well-vented fume hood to prevent radioactive contamination.

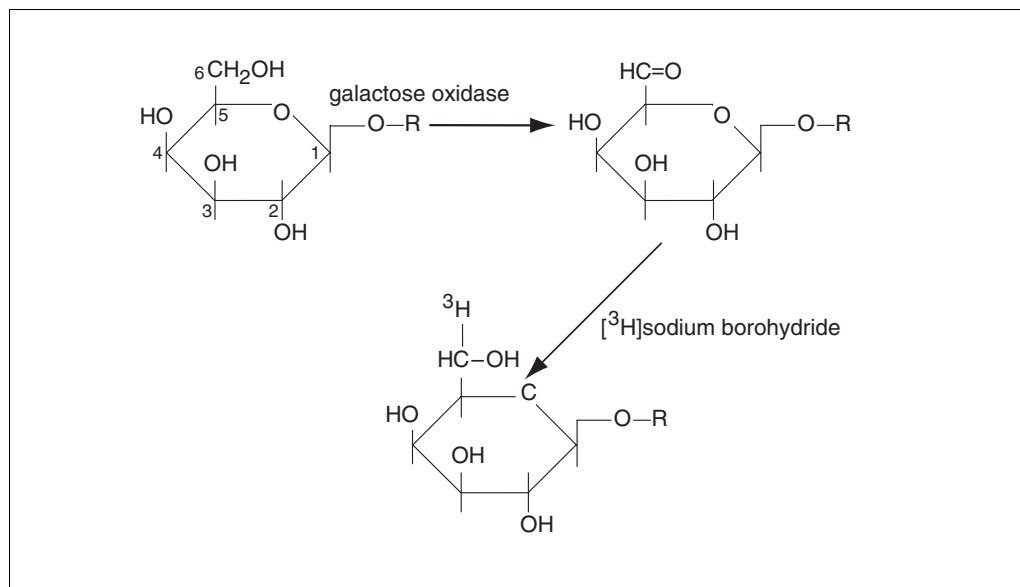


Figure 17.5.2 Labeling oligosaccharides by galactose oxidase treatment followed by $\text{NaB}[^3\text{H}]_4$ reduction.

Preparation and Analysis of Glycoconjugates

17.5.3

1. Mix the following in a 15-ml conical centrifuge tube (110 μ l total volume):

50 to 500 μ g oligosaccharide or glycopeptide
10 μ l 1 U/ μ l galactose oxidase
100 μ l 10 mM sodium phosphate buffer.

Incubate 2 hr at 37°C.

Sialylated oligosaccharides or glycopeptides should be treated with sialidase (UNIT 17.12) before treatment with galactose oxidase because sialic acid hinders oxidation by galactose oxidase. If the sample is treated with sialidase, either the sample should be desalted or the volume of the galactose oxidase reaction should be increased to ensure that the pH is 7.0.

2. Add 100 μ l of 0.2 M sodium borate buffer and 10 μ l of 0.5 mCi/ μ l NaB[³H]₄ (5 mCi). Incubate 1 hr at room temperature in a fume hood.
3. Add 100 μ l of 0.2 M NaBH₄/0.2 M sodium borate buffer and incubate 1 hr at room temperature.
4. Acidify, dry, and purify labeled oligosaccharides or glycopeptides as in steps 5 to 10 of the first basic protocol.

ALTERNATE PROTOCOL

FLUORESCENCE LABELING OLIGOSACCHARIDES AFTER PERIODATE OR GALACTOSE OXIDASE TREATMENT

Following oxidation, oligosaccharide-containing samples may be labeled with the non-isotopic fluorescent probe lucifer yellow CH, instead of being radioactively labeled with NaB[³H]₄. This technique can be used to label oligosaccharides on glycopeptides or glycolipids. Labeled oligosaccharides are separated from unincorporated dye by gel-filtration chromatography.

Additional Materials

For recipes, see Reagents and Solutions in this unit (or cross-referenced unit); for common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Phosphate-buffered saline (PBS; APPENDIX 2)

Lucifer yellow CH (Aldrich)

0.1 M sodium cyanoborohydride (NaBH₃CN)

0.5 \times 15-cm Sephadex G-15 and 1.0 \times 25-cm Sephadex G-25 columns in water (UNIT 10.9), calibrated

25°C water bath

Oxidize oligosaccharides

1. Oxidize oligosaccharide or glycopeptide sample (steps 1 to 2 of the first basic protocol or step 1 of the first alternate protocol).
2. Apply sample to a 0.5 \times 15-cm Sephadex G-15 column and elute with water. Collect 0.5-ml fractions.
3. Pool and dry fractions that contain the oxidized oligosaccharides or glycopeptides.
The appropriate fractions should be identified based on column calibration with glycopeptide standards (see UNIT 10.9). Oligosaccharides of glycopeptides elute at or near the void volume.
4. Dissolve sample in 100 μ l PBS.

Label oligosaccharides with fluorescent dye

5. Add lucifer yellow CH to 5 mM final concentration and incubate 12 hr at room temperature.
6. Apply reaction mixture to a 1.0 × 25-cm Sephadex G-25 column. Elute with water and collect 1-ml fractions. Wash until the unincorporated dye elutes.

Purify labeled oligosaccharides

7. Pool and dry fractions containing first fluorescent peak (labeled oligosaccharides).

It may be necessary to read UV absorbance or fluorescence of the eluted fractions, as fluorescence is not easily visible.

8. Dissolve sample in 0.1 ml of 0.1 M NaBH₃CN and incubate 15 min at 25°C. Immediately apply to the same Sephadex G-25 column used in step 6 and elute with water (avoid acidic conditions).
9. Collect, pool, and dry fluorescent oligosaccharides—i.e., the fractions at or near the void volume. Avoid light exposure as much as possible after this point.

Fractions that contain fluorescently labeled oligosaccharides or glycopeptides should elute in the void volume.

Remaining solutions should be discarded as toxic waste.

10. Dissolve sample in a minimal amount of water and store at –20°C until used.

RADIOLABELING FREE OLIGOSACCHARIDES

Free oligosaccharides can be obtained from Asn-linked oligosaccharides by hydrazinolysis (UNIT 17.15) or PNGaseF (*N*-glycanase) treatment (UNIT 17.13). In the following procedure, free oligosaccharides are purified by a simple gel filtration and then directly reduced with NaB[³H]₄. After reduction, the aldehyde group at C-1 of the reducing terminus becomes an alcohol with one radioactive tritium (see Fig. 17.5.2). Introduction of ³H at the reducing terminal allows the oligosaccharides to be followed by monitoring radioactivity.

Materials

*For recipes, see **Reagents and Solutions** in this unit (or cross-referenced unit); for common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.*

Free oligosaccharides (UNITS 17.15 & 17.12)

0.3 M sodium borate buffer, pH 9.5 (UNIT 11.16; adjust quantities of boric acid and 10 M NaOH appropriately)

0.5 mCi/μl tritiated sodium borohydride (NaB[³H]₄; 25 to 35 Ci/mmol) in 0.05 M NaOH (see recipe)

1 M sodium borohydride (NaBH₄)

1 M acetic acid in methanol

Nitrogen (N₂) stream

Methanol

7-ml conical glass (Pyrex) tubes

0.5 × 25-cm Sephadex G-25 or G-15 column (UNIT 10.9), calibrated to determine void volume

Additional reagents and equipment for gel-filtration chromatography (UNIT 10.9)

CAUTION: Steps 4 to 8 must be performed in a well-vented fume hood to prevent radioactive contamination.

BASIC PROTOCOL

Preparation and Analysis of Glycoconjugates

17.5.5

1. Desalt 1 to 500 µg free oligosaccharides by gel-filtration chromatography on a calibrated 0.5 × 25-cm Sephadex G-25 or G-15 column. Elute in water. Collect 0.5-ml fractions.

Oligosaccharides must be free from peptides or amino acids that may incorporate label during reduction.

2. Pool and dry fractions that contain free oligosaccharides. Dissolve in 50 µl water.
3. Add 0.3 M sodium borate buffer and water to a 100-µl final volume and a 0.2 M-sodium borate final concentration.

If oligosaccharides are dissolved in water, volume and concentration of buffer should be adjusted to give the desired final concentration.

4. Immediately add 10 to 20 µl of 0.5 mCi/µl NaB[³H]₄ in 0.05 M NaOH (5 to 10 mCi) and incubate 2 hr at room temperature in a fume hood.
5. Add 100 µl of 1 M NaBH₄ and incubate 2 hr at room temperature to complete the reduction.
6. Add 3 ml of 1 M acetic acid in methanol and dry under a nitrogen stream in a 30° to 35°C water bath.
7. Repeat step 6.
8. Add 3 ml methanol and dry under a nitrogen stream in a 30° to 35°C water bath.
9. Dissolve sample in 50 to 200 µl water and store at -70°C until used.

ALTERNATE PROTOCOL

RADIOLABELING O-GLYCAN OLIGOSACCHARIDES

During release from glycoproteins or glycopeptides by beta-elimination, the reducing termini of O-glycans can be labeled by alkaline borohydride treatment in the presence of a large amount of NaB[³H]₄. The reaction, however, must be carried out in the presence of 1 M NaBH₄ to prevent oligosaccharide degradation after release.

Additional Materials

For recipes, see Reagents and Solutions in this unit (or cross-referenced unit); for common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

100 to 1000 µg glycopeptides or ≤5 mg glycoproteins
0.1 to 0.2 mCi/µl tritiated sodium borohydride (NaB[³H]₄; 25 to 30 Ci/mmol) in
1 M NaBH₄/0.05 M NaOH (see recipe)
1 M acetic acid in methanol
Methanol
30° to 35°C and 45°C water baths

CAUTION: Steps 2 to 4 must be performed in a well-vented fume hood to avoid radioactive contamination.

1. Place 100 to 1000 µg glycopeptides or ≤5 mg glycoproteins in a 7-ml conical glass tube. Add 200 µl of 0.1 to 0.2 mCi/µl NaB[³H]₄ in 1 M NaBH₄/0.05 M NaOH (20 to 40 mCi). Incubate 24 to 48 hr at 45°C in a fume hood.

For larger amounts of sample, the volume of solution should be increased to dissolve sample completely and the amount of Na[³H]₄ should be increased to 100 mCi.

2. Add 3 ml of 1 M acetic acid in methanol and dry sample under a nitrogen stream in a 30° to 35°C water bath.

3. Repeat step 2.
4. Add 3 ml methanol and dry under a nitrogen stream in a 30° to 35°C water bath.
5. Dissolve sample in 100 to 200 µl water and store at -20°C until used.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Tritiated sodium borohydride (NaB[³H]₄), 0.5 mCi/µl

Dissolve 100 mCi NaB[³H]₄ (highest available specific activity, usually 25 to 35 Ci/mmol) in 200 µl ice-cold 0.01 or 0.05 M NaOH. Store in 10-µl aliquots for several months at -70°C until use. To make NaB[³H]₄ in 1 M NaBH₄/0.05M NaOH, dissolve solid NaBH₄ in tritiated solution in 0.05 M NaOH.

Alternatively, NaB[³H]₄ can be dissolved in freshly distilled anhydrous dimethylformamide and stored at -20°C in well-sealed tubes. Some investigators report stability for longer periods than with NaOH solutions.

COMMENTARY

Background Information

When sufficient quantities of purified glycoconjugates (e.g., glycoproteins or glycopeptides) are available, their sugar chains can be labeled and used for structural or sequence analysis (Morell and Ashwell, 1972; Takasaki and Kobata, 1978). These sugar chains can be radioactively labeled by periodate or enzymatic oxidation followed by reduction with tritiated sodium borohydride (NaB[³H]₄). Sialic acid residues can be preferentially oxidized by mild periodate oxidation because they contain *cis*-glycols in exocyclic carbons (Fig. 17.5.1). Such glycols are much more reactive with periodate than glycols in cyclic structures because of free rotation at C-8 and C-9. The aldehyde formed at C-7 or C-8 can be reduced by NaB[³H]₄ and radioactively labeled in the process (Fig. 17.5.1). Galactose oxidase treatment oxidizes galactose or *N*-acetylgalactosamine residues at the nonreducing termini, at C-6, forming aldehydes. The resultant aldehydes are reduced and labeled with NaB[³H]₄ (Fig. 17.5.2).

Oxidized samples can also be labeled with fluorescent molecules (Wilchek et al., 1980; Spiegel, 1987). After oxidation by periodate or galactose oxidase treatment, the sample is reacted with a fluorescent dye that contains a hydrazide group. After forming a Schiff base, the linkage can be stabilized by reduction. Once a fluorescent label is attached to an oligosaccharide, it can be detected by a fluorescence or UV detector. Fluorescence-tagged oligosac-

charides can be incubated with cells and incorporated oligosaccharides can be detected by fluorescence microscopy.

These methods have been used to examine the biological significance of oligosaccharide structures and their distribution on glycoconjugates. In one study, the structures of O-linked oligosaccharides were found to vary with the different maturation stages along a cell lineage (Fukuda et al., 1986). In another study, glycosylation of erythropoietin, which is essential for its function, was examined and compared with that of a recombinant erythropoietin (Sasaki et al., 1987).

Critical Parameters

There are two critical parameters for these procedures. First, it is essential that each reaction take place at the designated pH. If the pH of a reaction mixture is not properly adjusted, the reaction will not occur. Second, the quality of the NaB[³H]₄ added is critical. When NaB[³H]₄ is purchased, the reagent should be dark blue or slightly violet in color. When NaB[³H]₄ absorbs water and becomes degraded, it appears white and should not be used for labeling. Free oligosaccharides must be desalted by gel-filtration chromatography to remove salts, free amino acids, and peptides before chemical labeling.

An alternative method is to label the *N*-acetyl groups of *O*-glycans that have already been released from glycoproteins by alkaline beta-

elimination. The reduced *O*-glycans can be treated with hydrazine to remove *N*-acetyl groups and reacylated with radioactive acetic anhydride. This procedure is useful if anhydrous hydrazine is readily available (see Amano and Kobata, 1989).

Anticipated Results

Chemical labeling of carbohydrates by these procedures generally depends on the amount of samples available. For example, NaB[³H]₄ reduction of oligosaccharides is one of the most efficient methods to label oligosaccharides, but it still requires 0.05 to 0.1 μM for efficient labeling. Since NaB[³H]₄ reduction and fluorescence labeling are dependent on reactions with aldehyde groups, it is also an advantage to have a reasonable mass of sample, reducing the contribution from contaminants. Once there is a sufficient amount of sample, however, total radioactivity incorporated will allow many manipulations on radioactively labeled samples.

Time Considerations

In general, all procedures in this unit are simple and rapid. Each of the protocols described should take 1 or 2 days to finish.

Literature Cited

- Amano, J. and Kobata, A. 1989. Quantitative conversion of mucin-type sugar chains to radioactive oligosaccharides. *Methods Enzymol.* 179:261-270.
- Fukuda, M., Carlsson, S.R., Klock, J.C., and Dell, A. 1986. Structures of O-linked oligosaccharides isolated from normal granulocytes, chronic myelogenous leukemia cells, and acute myelogenous leukemia cells. *J. Biol. Chem.* 261:12796-12806.
- Morell, A.G. and Ashwell, G. 1972. Tritium labeling of glycoproteins that contain terminal galactose residue. *Methods Enzymol.* 28:205-208.
- Sasaki, H., Bothner, B., Dell, A., and Fukuda, M. 1987. Carbohydrate structure of erythropoietin expressed in Chinese hamster ovary cells by a human erythropoietin cDNA. *J. Biol. Chem.* 262:12059-12076.
- Spiegel, S. 1987. Fluorescent gangliosides. *Methods Enzymol.* 138:313-318.
- Takasaki, S. and Kobata, A. 1978. Microdetermination of sugar composition by radioisotope labeling. *Methods Enzymol.* 50:50-54.
- Wilchek, M., Spiegel, S., and Spiegel, Y. 1980. Fluorescent reagents for the labeling of glycoconjugates in solution and on cell surfaces. *Biochem. Biophys. Res. Commun.* 92:1215-1222.

Contributed by Minoru Fukuda
La Jolla Cancer Research Foundation
La Jolla, California

Detection and Analysis of Proteins Modified by O-Linked N-Acetylglucosamine

UNIT 17.6

The modification of Ser and Thr residues with O-linked β -N-acetylglucosamine (O-GlcNAc) is a common, dynamic, and essential modification of nuclear and cytoplasmic proteins. O-GlcNAc is ubiquitous, having been identified in all complex eukaryotes studied, including filamentous fungi and plants. O-GlcNAc has been found on a diverse range of proteins including cytoskeletal proteins, nuclear pore proteins, RNA polymerase II (RNA Pol II), transcription factors, proto-oncogene products, tumor suppressors, hormone receptors, phosphatases, and kinases (Comer and Hart, 2000).

The functional consequences of modifying proteins with O-GlcNAc is unclear, but it is required for survival at the single-cell level (Shafi et al., 2000). Three features suggest that O-GlcNAc performs a regulatory role: (1) O-GlcNAc occurs at sites on the protein backbone that are similar to those modified by protein kinases; (2) O-GlcNAc is reciprocal with phosphorylation on some well-studied proteins, such as RNA Pol II, estrogen receptor- β , SV40 large T-antigen, and the *c-Myc* proto-oncogene product; and (3) like phosphorylation, O-GlcNAc is highly dynamic, with rapid cycling in response to cellular signals or cellular stages. Perturbations in the metabolism of GlcNAc, which alter the regulation of many O-GlcNAc proteins, have been implicated in Alzheimer's disease, diabetes, and cancer (Wells et al., 2001).

This unit concentrates on techniques for the detection and analysis of proteins modified by O-GlcNAc, as well as methods for the analysis of enzymes responsible for the addition and removal of O-GlcNAc. We have focused on methods that require standard laboratory equipment. However, in some cases we also discuss more specialized technology.

The unit is set out in a stepwise manner. First, a protocol for increasing the stoichiometry of O-GlcNAc on proteins is given (see Basic Protocol 1). This is followed by simple techniques for the detection/screening of O-GlcNAc modified proteins either by western blotting or lectin affinity chromatography (see Basic Protocols 2 to 4). Separate protocols verify that the glycan is O-linked GlcNAc (see Support Protocols 1 and 2). These methods are followed by protocols for more comprehensive analysis of O-GlcNAc modified proteins, including labeling of O-GlcNAc residues with [3 H]Gal, and subsequent product analysis (see Alternate Protocol 1, see Basic Protocols 5 to 8, and see Support Protocols 3 and 4). The final two protocols assay for O-GlcNAc transferase and O-GlcNAcase activity, respectively (see Basic Protocols 9 and 10 and Support Protocol 5).

INCREASING THE STOICHIOMETRY OF O-GlcNAc ON PROTEINS BEFORE ANALYSIS

**BASIC
PROTOCOL 1**

In many cultured mammalian cells, the number of O-GlcNAc moieties per protein molecule can be increased by treating cells with O-(2-acetamido-2-deoxy-D-gluco-pyranosylidene)-amino-N-phenyl-carbamate (PUGNAc), a potent and cell-permeable inhibitor of O-GlcNAcase (Haltiwanger et al., 1998). Alternatively, streptozotocin (STZ; Roos et al., 1998) and glucosamine (Han et al., 2000) have been used to increase the stoichiometry of O-GlcNAc on proteins. However, STZ has been shown to induce poly-(ADP-ribose) polymerase-mediated apoptosis in Min6 cells (Gao et al., 2000) and should be used with caution. In addition, STZ is only effective in cells that express the glucose transporter GLUT-2 (Schnedl et al., 1994).

**Preparation and
Analysis of
Glycoconjugates**

Materials

Cells of interest growing in monolayer culture, and appropriate culture medium
20 mM PUGNAc (CarboGen Laboratories) stock in Milli-Q water (filter sterilize and store in aliquots up to 6 to 12 months at -80°C)
500 mM glucosamine stock in 500 mM HEPES, pH 7.5 (make just prior to use; filter sterilize)
500 mM streptozotocin (STZ; Sigma) in 100 mM citrate buffer, pH 4.5 (make just prior to use; filter sterilize)
100-mm tissue culture dishes

1. Grow cells in monolayer culture in a sufficient number of 100-mm dishes.
2. Add (or replace growth medium with fresh medium containing) 40 to 100 μM PUGNAc (added from 20 mM stock), 5 mM glucosamine (added from 500 mM stock), or 2 to 5 mM STZ (added from 500 mM stock). Incubate cells in an incubator for 6 to 18 hr.

PUGNAc is taken up by both dividing and stationary cells.

If necessary, cells can be treated by PUGNAc for several days without any apparent cell toxicity. However, prolonged treatment does not appear to result in additional increase in O-GlcNAc compared to the 6 to 18 hr treatment.

When using glucosamine, mannitol is often added at the same concentration. This controls for changes in osmolarity due to the additional sugar in the medium (Heart et al., 2000).

3. At the end of treatment, take the dishes out of the incubator and place on ice. Extract as desired. Separate proteins by SDS-PAGE (UNIT 10.2A) and electroblot onto appropriate membrane (UNIT 10.8).

Alternatively, extract proteins and proceed with protein purification (UNITS 10.9-10.14) or immunoprecipitation (UNIT 10.16).

BASIC PROTOCOL 2

DETECTION OF PROTEINS MODIFIED BY O-GlcNAc USING ANTIBODIES

Recently Comer et al. (2001) showed that an antibody (CTD 110.6) raised against the glycosylated C-terminal domain of the RNA polymerase II large subunit was a general O-GlcNAc antibody. Unlike many lectins, CTD 110.6 shows little cross-reactivity toward terminal GlcNAc on complex glycans. Several other antibodies, HGAC 85 (Turner et al., 1990) and RL2 (Snow et al., 1987), have been reported as O-GlcNAc specific antibodies. However, these antibodies only recognize a subset of O-GlcNAc modified proteins, and RL2 in particular is dependent on the structure of the peptide backbone around the glycosylation site (Holt et al., 1987).

It is important to include an appropriate negative control (100 ng ovalbumin) for cross-reactivity toward N-linked glycans. Using the standard Amersham Pharmacia Biotech enhanced chemiluminescent (ECL) system, the authors have found that 10 μg of cytoplasmic, nuclear, or total cell extract is sufficient. For purified proteins, Comer and co-workers found that 25 to 50 ng of a neoglycoconjugate was sufficient (Comer et al., 2001).

Materials

Purified or crude protein (e.g., Basic Protocol 1) separated by SDS-PAGE (UNIT 10.2A) and electroblotted to polyvinylidene difluoride (PVDF; UNIT 10.8) or nitrocellulose (duplicate blots are needed)
TBS-HT (see recipe)

Antibody: CTD 110.6 ascites (Covance) diluted 1/2500 in TBS-HT (see recipe for TBS-HT)

N-acetylglucosamine (GlcNAc; Sigma)

HRPO-conjugated anti-mouse IgM diluted 1/5000 in TBS-HT

TBS-HD (see recipe)

ECL kit (Amersham Pharmacia Biotech)

Additional reagents and equipment for visualization with chromogenic and luminescent substrates (UNIT 10.8)

1. Block blots by incubating with TBS-HT for 60 min at room temperature.

Batteiger et al. (1982) have shown that high concentrations of Tween 20 substitute for blocking with milk or bovine serum albumin (BSA).

2. Incubate blots with CTD 110.6 (1/2500 dilution in TBS-HT), in duplicate, with and without 10 mM GlcNAc, overnight at 4°C.

To control for specificity it is important to perform a control blot. Here, the antibody is preincubated with 10 mM GlcNAc for ~5 min on ice before being applied to the control blot. Note that the concentration of antibody should be optimized with each new preparation. We find that 1/2500 is a good place to start with ascites (containing antibody in the mg/ml range).

3. Wash blots in TBS-HD twice, each time for 10 min at room temperature.

As the blots are washed under these conditions, prestained markers will fade. This does not appear to be the result of proteins being solubilized off the membranes, but of the dye used to stain the markers. While the markers fade, they do not completely disappear. The authors usually double the amount of marker used to compensate for fading.

4. Wash blots in TBS-HT three times, each time for 10 min, room temperature.

5. Incubate blots with HRPO-conjugated anti-mouse IgM (1/5000 dilution in TBS-HT) for 50 min, room temperature.

The concentration of secondary antibody varies from lot to lot and should be optimized each time with each new preparation.

6. Wash blots in TBS-HD twice, each time for 10 min, room temperature.

7. Wash blots in TBS-HT three times, each time for 10 min, room temperature.

8. Develop the HRPO reaction using, e.g., the ECL system (UNIT 18.4).

Note that the antibody often cross-reacts with pre-stained markers.

DETECTION OF PROTEINS MODIFIED BY O-GlcNAc USING THE LECTIN sWGA

Many lectins are reportedly specific for β -GlcNAc residues. The authors have typically used succinylated wheat germ agglutinin (sWGA), which is widely available and is derivatized with a number of useful functional groups including horseradish peroxidase (HRPO). Before succinylation, WGA will recognize both sialic acid and GlcNAc (Monsigny et al., 1980). For additional information concerning lectin chromatography, see UNIT 17.1.

The amount of “test” protein used is dependent on the technique(s) used to develop the HRPO reaction. Using the standard Amersham Pharmacia Biotech ECL system the authors find that 7.5 μ g of cytoplasmic or nuclear extract is sufficient.

BASIC PROTOCOL 3

Preparation and Analysis of Glycoconjugates

17.6.3

It is important to include an appropriate positive (100 ng ovalbumin) and negative (100 ng of BSA) control. As a control, a portion of the sample should also be treated with hexosaminidase (see Support Protocol 2), to show that reactivity is toward GlcNAc. In addition, the sample should be subjected to reductive β -elimination to verify that lectin/antibody reactivity is towards O-linked glycans (see Support Protocol 1).

Materials

Purified or crude protein separated by SDS-PAGE (UNIT 10.2A) and electroblotted to polyvinylidene difluoride (PVDF; UNIT 10.8) or nitrocellulose (duplicate blots are needed)

5% (w/v) BSA in TBST (see recipe for TBST)

TBST (see recipe)

0.1 μ g/ml HRPO-conjugated sWGA (EY Labs) in TBST (see recipe for TBST): the lectin can be stored at 1 mg/ml in 0.01 M PBS, pH 7.4 (APPENDIX 2), at -20°C for at least 1 year

N-acetylglucosamine (GlcNAc; Sigma)

High-salt TBST (HS-TBST): TBST (see recipe) containing 1 M NaCl

Tris-buffered saline (TBS; see recipe)

ECL kit (Amersham Pharmacia Biotech)

Additional reagents and equipment for visualization with chromogenic and luminescent substrates (UNIT 10.8)

1. Wash duplicate blots for 10 min in 5% BSA/TBST, room temperature.
2. Block by incubating blots in 5% (w/v) BSA/TBST for at least 60 min at room temperature.

IMPORTANT NOTE: Milk cannot be used as the blocking agent, since many of the proteins in milk are modified by glycans that react with sWGA.

3. Wash blots three times, each time for 10 min in TBST, room temperature.
4. Incubate blots in 0.1 μ g/ml sWGA-HRPO in TBST, in duplicate, with and without 1 M GlcNAc, overnight at 4°C .

To control for lectin specificity it is important to perform a control blot. Here, the lectin is preincubated with 1 M GlcNAc for ~ 5 min on ice before being applied to the control blot.

5. Wash blots six times, each time for 10 min, in HS-TBST.
6. Wash blots once in TBS for 10 min.
7. Develop the HRPO-reaction using, e.g., the the ECL system (UNIT 10.8).

Using the ECL system described, 100 ng of ovalbumin should be visualized in 5 to 15 sec.

CONTROL FOR O-LINKED GLYCOSYLATION

Traditionally, mild alkaline reduction (reductive β -elimination) has been used to release O-linked carbohydrates from proteins (Amano and Kobata, 1990). This method has been adapted for blots to show that lectin/antibody reactivity is toward O-linked rather than N-linked glycans (Duk et al., 1997). Proteins blotted to PVDF are treated with 55 mM NaOH overnight (releasing O-linked sugars) and then probed using lectins or antibodies.

There are a number of reasons why lectin/antibody reactivity could be lost after NaOH treatment, e.g., the sugars were destroyed instead of being released, or the protein was degraded. To control for these, it is important to have control proteins with N- and O-linked sugars, and to stain one blot for protein after treatment preferably with an antibody. The

SUPPORT PROTOCOL 1

Detection and Analysis of Proteins Modified by O-Linked N- Acetylglucosamine

17.6.4

authors suggest a control blot of bovine asialofetuin (Sigma) which contains both N- and O-linked sugars terminating in GlcNAc, treated and not treated with PNGase F (*UNIT 17.13A*).

Materials

Protein samples and controls blotted to PVDF (triplicate blots are needed; nitrocellulose is not suitable as it dissolves in 55 mM NaOH)

Tris-buffered saline (TBS; see recipe)

55 mM NaOH

3% (w/v) BSA in TBST (see recipe for TBST)

40°C water bath

Additional reagents and equipment for probing protein blots with protein-specific antibodies (see Basic Protocol 2) or lectins (see Basic Protocol 3)

1. Wash blots once in TBS for 10 min.
2. Incubate two blots in 55 mM NaOH at 40°C overnight; incubate the control blot in Milli-Q water at 40°C overnight.

The blots treated with NaOH will yellow slightly.

3. Wash blots three times, each time for 10 min at room temperature, in TBST.
4. Block by incubating blots in 3% w/v BSA/TBST for 60 min at room temperature.
5. Probe blots (one treated and one untreated) with carbohydrate-specific lectins (see Basic Protocol 3) or antibodies (see Basic Protocol 2). Probe the second NaOH-treated blot with a protein-specific antibody (see Basic Protocol 2).

On the untreated blot, asialofetuin ± PNGase F should react with sWGA, as both the N- and O-linked sugars contain terminal GlcNAc residues. On the treated blot only the asialofetuin – PNGase F should react with sWGA.

DETECTION AND ENRICHMENT OF PROTEINS USING sWGA-AGAROSE

sWGA lectin affinity chromatography provides a convenient method for enriching and detecting O-GlcNAc modified proteins. This procedure has been adapted for detecting proteins that are difficult to purify or are present in low copy number, such as transcription factors. In this protocol, the protein of interest is synthesized in a rabbit reticulocyte lysate (RRL) *in vitro* transcription translation (ITT) system (Promega) and labeled with either [³⁵S]Met, [³⁵S]Cys, or [¹⁴C]Leu. After desalting, the proteins are tested for their ability to bind sWGA agarose in a GlcNAc-specific manner (Roquemore et al., 1994).

Alternatively, the lectin *Ricinus communis* agglutinin 1 (RCA1) has been used to select for O-GlcNAc proteins that have previously been labeled by galactosyltransferase (see Alternate Protocol 1). Proteins modified by terminal Gal are specifically retained on a RCA1 affinity column. Labeled O-GlcNAc proteins are released under mild conditions, while those containing N-linked structures require lactose addition to the buffer before elution results (Hayes et al., 1995; Greis and Hart, 1998). The method described in this protocol can be adapted for RCA1 affinity chromatography by substituting RCA1-agarose (EY Labs) for sWGA-agarose and changing the order of the Gal and GlcNAc elution buffer.

Materials

cDNA subcloned into an expression vector with an SP6 or T7 promoter (~0.5 to 1 µg/µl)

Kit for RRL ITT system (Promega)

BASIC PROTOCOL 4

Preparation and Analysis of Glycoconjugates

17.6.5

Label: [³⁵S]Met, or [³⁵S]Cys, or [¹⁴C]Leu
sWGA-agarose (Vector Laboratories)
sWGA wash buffer: PBS (APPENDIX 2) containing 0.2% (v/v) NP-40
sWGA Gal elution buffer (see recipe)
sWGA GlcNAc elution buffer (see recipe)
1-ml tuberculin syringe with glass wool plug at bottom to support chromatography matrix or Bio-Rad Bio-Spin disposable chromatography column
Additional reagents and equipment for digesting proteins with hexosaminidase (see Support Protocol 2), desalting (see Support Protocol 5), SDS-PAGE (UNIT 10.2A), and autoradiography (APPENDIX 3A)

Prepare proteins

1. Synthesize proteins to incorporate the desired label ([³⁵S]Met, [³⁵S]Cys, or [¹⁴C]Leu) using the RRL ITT system according to the manufacturer's instructions. Include the protein of interest, a positive control for sWGA binding (for example, the nuclear pore protein p62), a negative control (luciferase, supplied with kit), and a no-DNA control.
2. Treat half of each sample with hexosaminidase (see Support Protocol 2).
3. Desalt samples using spin filtration (e.g., Amersham Pharmacia Biotech Microspin G-50 columns) or a 1-ml G-50 desalting column (as for desalting O-GlcNAc transferase; see Support Protocol 5).

The following procedure is carried out at 4°C.

Apply protein samples to chromatography columns

4. Equilibrate sWGA-agarose and pack column as follows:
 - a. If resin is supplied as 50% slurry (i.e., 50% resin/50% storage solution) remove 300 µl (double the volume required) and pipet into a 1-ml tuberculin syringe or disposable chromatography column.
 - b. Let storage solution drain from resin.
 - c. Equilibrate resin by washing column four times, each time with 1 ml of sWGA wash buffer. Cap column.

The volumes given are appropriate for a sample derived from an ITT. For enrichment of other protein samples, the volume of sWGA should be optimized for the protein sample applied. The authors find that 50 mg of cell extract requires 1 ml of sWGA-agarose, assuming that 1% to 2% of the total cell extract is modified by O-GlcNAc.

5. Apply sample (~30 µl of an ITT reaction) to the column and let stand at 4°C for 30 min, or cap and incubate at 4°C for 30 min with rotating or rocking.

Wash column and elute GlcNAc

6. At the end of the 30-min incubation, uncap the column and allow the sample to "run through" the resin. Collect this as the "run through" fraction. Wash column with 15 ml of sWGA wash buffer at 10 ml/hr, collecting 0.5-ml fractions.
7. Load the column with 300 µl of sWGA Gal elution buffer and let stand at 4°C for 20 min.
8. Wash column with 5 ml of sWGA Gal elution buffer, collecting 0.5 ml fractions.
9. Repeat steps 6 to 8 using GlcNAc elution buffer.
10. Count 25 µl of each fraction using a liquid scintillation counter.

11. Pool positive fractions that elute in the presence of GlcNAc and precipitate using TCA or methanol.

To precipitate proteins with methanol, mix 1 vol of sample with 10 vol of ice-cold methanol. Incubate overnight at -20°C . Recover protein by microcentrifuging 10 min at $16,000 \times g$, 4°C , in a microcentrifuge tube (which is the most efficient procedure) or in 15-ml conical centrifuge tubes for 10 min at $3000 \times g$, 4°C . Resuspend samples in SDS-PAGE sample buffer (UNIT 10.2A).

As many proteins in rabbit reticulocyte lysate contain O-GlcNAc and bind WGA, the authors do not recommend the addition of carrier proteins at this point. Typically, a fraction of the GlcNAc elution containing a total of 1000 to 2000 dpm [^{35}S]Met is precipitated and analyzed by SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A).

The use of acetone to precipitate proteins is not recommended, as free GlcNAc will also precipitate.

12. Analyze pellet by SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A).

DIGESTION OF PROTEINS WITH HEXOSAMINIDASE

Terminal GlcNAc and O-GlcNAc can be removed from proteins using commercially available hexosaminidases; these enzymes will also cleave terminal GalNAc residues. Unlike O-GlcNAcase, commercial hexosaminidases have low pH optima, typically pH 4.0 to 5.0.

Materials

Protein sample for digestion (include a positive control, e.g., ovalbumin)
2% (w/v) SDS
2 \times hexosaminidase reaction mixture (see recipe)

Additional reagents and equipment for SDS-PAGE (UNIT 10.2A) and electroblotting (UNIT 10.8)

1. Mix sample 1:1 with 2% SDS and boil for 5 min.
2. Mix sample 1:1 with 2 \times hexosaminidase reaction mixture and incubate at 37°C for 4 to 24 hr.
3. To assess completeness of the digestion, separate an aliquot of the reaction by SDS-PAGE (UNIT 10.2A) and electroblot (UNIT 10.8) onto an appropriate membrane. Probe blots with carbohydrate-specific lectins or antibodies.

DETECTION OF PROTEINS MODIFIED BY O-GlcNAc USING GALACTOSYLTRANSFERASE

The enzyme β -1,4-galactosyltransferase (from bovine milk) will label any terminal GlcNAc residue with Gal, using uridine diphospho-D-Gal (UDP-Gal) as a donor substrate (Brew et al., 1968). Hart and colleagues have exploited this property, using the enzyme to label terminal GlcNAc residues on proteins with [^3H]Gal, forming a [^3H]- β Gal1-4 β GlcNAc (Torres and Hart, 1984; Roquemore et al., 1994; Greis and Hart, 1998). The labeled sugar can be chemically released (via β -elimination) and analyzed by size-exclusion chromatography on a BioGel-P4 column, using the ^3H radiolabel to detect the fraction of interest (Roquemore et al., 1994). Labeling the O-GlcNAc allows for the subsequent detection of the proteins and peptides of interest during SDS-PAGE, HPLC, protease digestion, and Edman degradation steps. Researchers have been able to identify glycosylation sites on as little as 10 pmol using these methods (Greis et al., 1996).

SUPPORT PROTOCOL 2

ALTERNATE PROTOCOL 1

Preparation and Analysis of Glycoconjugates

17.6.7

To achieve efficient labeling of some proteins, it is necessary to denature samples, for example by boiling in the presence of 10 mM DTT and 0.5% (w/v) SDS. Galactosyltransferase has been shown to be active in solutions containing 5 mM DTT, 0.5 M NaCl, up to 2% (v/v) Triton X-100, up to 2% (v/v) NP-40, and 1 M urea. Up to 0.5% (w/v) SDS can be used, if it is titrated with a 10-fold molar excess of either Triton X-100 or NP-40 in the final reaction mixture. Digitonin, which is commonly used to solubilize cells, should be used with caution, as it is a substrate for galactosyltransferase. The total ionic strength should be less than 0.2 M.

Galactosyltransferase requires 1 to 5 mM Mn^{2+} for activity, but is inhibited by Mg^{2+} and concentrations of Mn^{2+} >20 mM. EDTA (or analogs) should be avoided unless titrated with appropriate levels of Mn^{2+} . Note that 1 mol of EDTA binds 2 mol of Mn^{2+} .

Free UDP is also an inhibitor of galactosyltransferase. For studies where complete labeling of the GlcNAc is preferable, such as site mapping, calf intestinal alkaline phosphatase is included in the reaction, as it degrades UDP (Unverzagt et al., 1990). While this increases the efficiency of the reaction, it is important to add this to the control as some preparations of alkaline phosphatase contain proteins that will label with galactosyltransferase (R.N. Cole, pers. commun.).

NOTE: Protease inhibitors, such as PIC 1, PIC 2, and PMSF (see recipe for 1000× protease inhibitors in Reagents and Solutions), can be included (final concentrations, 1×), but GlcNAc and 1-amino GlcNAc should be removed prior to labeling by spin filtration or another method of desalting.

Materials

Protein sample(s)
Dithiothreitol (DTT)
Sodium dodecyl sulfate (SDS)
Label: 1.0 mCi/ml UDP-[³H]Gal, (17.6 Ci/mM; Amersham Pharmacia Biotech) in 70% v/v ethanol
Nitrogen source
25 mM 5'-adenosine monophosphate (5'-AMP), in Milli-Q water, pH 7.0
Buffer H (see recipe)
10× galactosyltransferase labeling buffer (see recipe)
Galactosyltransferase, autogalactosylated (see Support Protocol 3)
Calf intestinal alkaline phosphatase
Unlabeled UDP-Gal
Stop solution: 10% (w/v) SDS/0.1 M EDTA

100°C water bath
30 × 1-cm Sephadex G-50 column equilibrated in 50 mM ammonium formate/0.1% (w/v) SDS

Additional reagents and equipment for acetone precipitation of protein (*UNIT 17.10A*), PNGase F digestion of proteins (*UNIT 17.13A*), SDS-PAGE (*UNIT 10.2A*), and product analysis (see Basic Protocol 6)

Prepare the reaction

1. Denature protein sample by adding DTT to 10 mM and SDS to 0.5% (w/v), then boiling the sample for 10 min.
2. Decide how many reactions are going to be carried out and thus how much label will be needed (~1 to 2 μ Ci/reaction).

A positive control (ovalbumin, 2 µg), a negative control (because galactosyltransferase can label itself), and a sample-minus-enzyme control will be needed.

3. Remove solvent from label in a Speed-Vac evaporator or under a stream of nitrogen.

Ethanol can inhibit the galactosyltransferase reaction, but if <4 µl is required the label can be added directly to the reaction (final reaction volume, 500 µl).

4. Resuspend appropriate amount of label for each reaction, respectively, in 50 µl of 25 mM 5'-AMP.

The AMP is included to inhibit possible phosphodiesterase reactions, which might compete for label during the labeling experiment.

5. Set up reactions as follows:

Up to 50 µl protein sample (final concentration 0.5 to 5 mg/ml)
350 µl buffer H
50 µl 10× galactosyltransferase labeling buffer
50 µl UDP-[³H]Gal/5'-AMP mixture from step 4
30 to 50 µl autogalactosylated galactosyltransferase
1 to 4 U calf intestinal alkaline phosphatase
Milli-Q water to final volume of 500 µl

Reaction volumes can be scaled down to 50 µl.

6. Labeling at 37°C for 2 hr or at 4°C overnight.

These are the typical conditions. Galactosyltransferase is active over a range of temperatures.

7. Add unlabeled UDP-Gal to a final concentration of 0.5 to 1.0 mM and another 2 to 5 µl of galactosyltransferase.

For studies where complete labeling of the GlcNAc is required, such as site mapping, the reactions are chased with unlabeled UDP-Gal and fresh galactosyltransferase.

8. Add 50 µl of stop solution to each sample and heat to 100°C for 5 min in a water bath.

Isolate the product

9. Resolve the protein from unincorporated label using a Sephadex G-50 column equilibrated in 50 mM ammonium formate/0.1% w/v SDS. Collect 1-ml fractions.

Size-exclusion chromatography using Sephadex G-50 is traditionally used to desalt samples. However, TCA precipitation, spin filtration/buffer exchange, or other forms of size-exclusion chromatography (e.g., Pharmacia PD-10 desalting column) can be used. The addition of carrier proteins such as BSA (~67 kDa) and cytochrome c (~12.5 kDa) to samples and buffers will reduce the amount of protein lost due to nonspecific protein adsorption.

10. Count a 50-µl aliquot of each fraction using a liquid scintillation counter.

Approximately 2×10^6 dpm of [³H]Gal should be incorporated into 2 µg of ovalbumin.

11. Combine the void volume and lyophilize to dryness.

12. Resuspend samples in Milli-Q water and precipitate with acetone (UNIT 17.10A).

13. Treat samples with PNGase F (UNIT 17.13A), separate by SDS-PAGE (UNIT 10.2A) and detect by autoradiography (APPENDIX 3A). Alternatively, subject samples to “product analysis” to confirm that the label was incorporated onto O-GlcNAc (see Basic Protocol 6).

AUTOGALACTOSYLATION OF GALACTOSYLTRANSFERASE

As galactosyltransferase contains N-linked glycosylation sites, it is necessary to block these before using this enzyme to probe other proteins for terminal GlcNAc.

Materials

- 10× galactosyltransferase labeling buffer (see recipe)
- 10,000 U/ml aprotinin
- 2-mercaptoethanol
- UDP-Gal
- Saturated ammonium sulfate: >17.4 g (NH₄)₂SO₄ in 25 ml Milli-Q water
- 85% ammonium sulfate: 14 g (NH₄)₂SO₄ in 25 ml Milli-Q water
- Galactosyltransferase storage buffer (see recipe)
- 30- to 50-ml centrifuge tubes
- Refrigerated centrifuge

1. Resuspend 25 U of galactosyltransferase in 1 ml of 1× galactosyltransferase labeling buffer.
2. Transfer sample to 30- to 50-ml centrifuge tube.
The centrifuge tubes should be able to withstand a centrifugal force of 15,000 × g.
3. Remove a 5-μl aliquot for an activity assay.
This is the “Pre-Gal” sample to be used in Support Protocol 4.
4. Add 10 μl of 10,000 U/ml aprotinin, 3.5 μl of 2-mercaptoethanol, and 1.5 to 3.0 mg of UDP-Gal.
5. Incubate the sample on ice for 30 to 60 min.
6. Add 5.66 ml of prechilled saturated ammonium sulfate in a dropwise manner. Incubate on ice for 30 min.
7. Centrifuge 15 min at >10,000 × g, 4°C. Pour off supernatant.
8. Resuspend pellet in 5 ml cold 85% ammonium sulfate and incubate on ice for 30 min.
9. Centrifuge 15 min at >10,000 × g, 4°C, and pour off supernatant.
10. Resuspend pellet in 1 ml of galactosyltransferase storage buffer and divide into 50-μl aliquots, saving 5 μl for an activity assay as the “Auto-Gal” sample. Assay that aliquot for activity (see Support Protocol 4).
11. Store remaining aliquots up to 1 year at –20°C pending use in Alternate Protocol 1.

ASSAY OF GALACTOSYLTRANSFERASE ACTIVITY

As sample and activity may be lost during the autogalactosylation procedure, it is important to assess the activity of the enzyme.

Materials

- 1.0 mCi/ml UDP-[³H]Gal, (17.6 Ci/mM; Amersham Pharmacia) in 70% v/v ethanol
- Nitrogen source
- 1× galactosyltransferase dilution buffer: galactosyltransferase storage buffer (see recipe) supplemented with 5 mg/ml BSA
- 10× galactosyltransferase labeling buffer (see recipe)

25 mM 5'-adenosine monophosphate (5'-AMP) in Milli-Q water, pH 7.0
 "Pre-Gal" sample aliquot (see Support Protocol 3, step 3) and "Auto-Gal" sample aliquot (see Support Protocol 3, step 10)
 200 mM GlcNAc
 Dowex AG1-X8 resin (PO₄ form) slurry in 20% (v/v) ethanol
 Glass wool

1. Dry 40 µl of 0.1 µCi/µl of UDP-[³H]Gal in a Speed-Vac evaporator or under a stream of nitrogen.
2. Resuspend in 90 µl of 25 mM 5'-AMP.
3. Make 1/1000, 1/10,000, and 1/100,000 serial dilutions of the "Pre-Gal" and "Auto-Gal" sample aliquots, in 1× galactosyltransferase dilution buffer. Using these dilutions, 200 mM GlcNAc, and 10× galactosyltransferase labeling buffer, prepare reaction mixtures as described in Table 17.6.1.
4. Start the reaction by adding 10 µl of 0.05 µCi/µl UDP-[³H]Gal (see step 2) to each tube.
5. Incubate samples at 37°C for 30 min.
While the samples are incubating, prepare the columns.
6. Pour 1 ml of Dowex AG1-X8 slurry (PO₄ form) into 13 Pasteur pipets, each plugged with a small amount of glass wool.
The glass wool prevents the resin from flowing out of the column. If too much glass wool is used, it will reduce the flow rate of the column. The glass wool plug should be 0.3 to 0.5 mm long and should not be over-compressed.
7. Wash with at least 3 ml of Milli-Q water. Do not let the columns run dry.
8. When almost all the Milli-Q water has eluted, place each column over a separate 15-ml scintillation vial.
9. Stop the reaction (still incubating from step 5) by adding 500 µl of Milli-Q water.
10. Load each sample onto the corresponding column and add a 500 µl water wash of the tube. Collect eluate as fraction A.
11. Elute with two 1-ml additions of Milli-Q water. Collect eluates as fractions B and C, respectively.

Table 17.6.1 Reaction Mixtures for Assay of Galactosyltransferase Activity

Sample	Dilution	Vol. (µl) of diluted sample	Vol. (µl) of 200 mM GlcNAc	Vol. (µl) 10× Gal labeling buffer	Vol. (µl) Milli-Q water
Blank		0	10	10	70
Pre-Gal	1/1000	10	10	10	60
	1/10,000	10	10	10	60
	1/100,000	10	10	10	60
Auto-Gal	1/1000	10	10	10	60
	1/10,000	10	10	10	60
	1/100,000	10	10	10	60

12. Count 100 μl of the sample and 10 μl of the UDP-[^3H]Gal using a liquid scintillation counter.
13. Calculate activity.
 - a. Calculate the total moles of UDP-Gal transferred to GlcNAc as follows:

$$\{ \text{cpm}_{(\text{with GlcNAc})} - \text{cpm}_{(\text{no GlcNAc})} / \text{total cpm} \} \times 1.76 \text{ nmol}$$

Note, $\text{cpm}_{(\text{with GlcNAc})}$ represents counts in the sample, $\text{cpm}_{(\text{no GlcNAc})}$ represents counts in the no-enzyme control, and total cpm represents the total counts available to transfer. $\{ \text{cpm}_{(\text{with GlcNAc})} - \text{cpm}_{(\text{no GlcNAc})} / \text{total cpm} \}$ represents the portion of [^3H]Gal transferred in the reaction. This proportion is multiplied by the total number of moles of UDP-Gal in the reaction. If the specific activity of the UDP-[^3H]Gal is 17.6 Ci/mM and 10 μl of label was used, then there are 1.76 nmol of UDP-Gal in 10 μl .
 - b. Calculate the activity; one unit of activity (U) is defined as 1 μM of Gal transferred to GlcNAc per minute at 37°C.

**BASIC
PROTOCOL 5**

DETECTION OF PROTEINS MODIFIED BY O-GlcNAc USING METABOLIC LABELING

Metabolic labeling of O-GlcNAc-bearing proteins provides a useful means to test if a protein of interest is modified by O-GlcNAc, to observe the gross dynamic changes in O-GlcNAc levels and to study the subcellular localization in response to stimulation or during cell cycle. In the protocol described below, cells are labeled with [^3H]glucosamine, which is metabolized to UDP-[^3H]GlcNAc in the hexosamine synthetic pathway. For labeling, it is critical that the labeled sugar compete with glucose import; this ensures efficient uptake of the label. While glucosamine is a good competitor, N-acetylglucosamine is not. For further discussions on metabolic labeling of glycoconjugates, readers are encouraged to consult Varki (1994) for details.

Materials

Cells of interest, growing in culture
 Biosynthetic labeling medium (see recipe)
 Phosphate-buffered saline (PBS; APPENDIX 2)

Additional reagents and equipment for PNGase F digestion of proteins (UNIT 17.13A), SDS-PAGE (UNIT 10.2A), and autoradiography (APPENDIX 3A)

1. Replace growth medium with biosynthetic labeling medium. Label cells for 5 to 24 hr in an incubator.

The labeling is done in a glucose-free medium to maximize labeling efficiency. The addition of nonessential amino acids to the medium reduces the influx of glucosamine into the amino acid biosynthetic pathways (Medina et al., 1998).

2. Wash labeled cells twice, each time by resuspending in 5 ml of PBS and centrifuging 10 min at $500 \times g$, 4°C.
3. Extract as desired.

See Critical Parameters for a discussion on reducing O-GlcNAcase activity.

There are many types of cell fractionation and extraction techniques available. Those that minimize plasma membrane and organelle contamination are preferable. If a particular study concerns one or a few proteins, and specific antibodies are available, the protein(s) can be isolated by immunoprecipitation after metabolic labeling.

4. Treat sample with PNGase F (UNIT 17.13A).

5. Separate proteins by SDS-PAGE (UNIT 10.2A) and visualize labeled proteins by autoradiography (APPENDIX 3A).

CHARACTERIZATION OF LABELED GLYCANS BY β -ELIMINATION AND CHROMATOGRAPHY

BASIC
PROTOCOL 6

This protocol has three steps: (1) the release of carbohydrates as sugar alditols by reductive β -elimination; (2) desalting the sample, while confirming the size of the labeled sugar alditol(s); and (3) confirmation that the product is [^3H] β Gal1-4 β GlcNAcol (from galactosyltransferase labeling) or [^3H]GlcNAcol (from metabolic labeling).

Materials

Labeled proteins
 β -elimination reagent: 1 M NaBH_4 /0.1 M NaOH (prepare fresh)
4 M acetic acid
Screw-cap microcentrifuge tubes

Additional reagents and equipment for acetone or methanol precipitation of proteins (UNIT 17.10A), size-exclusion (gel-filtration) chromatography (UNIT 10.9), and Dionex chromatography (Townsend et al., 1990; Hardy and Townsend, 1994)

1. Acetone or methanol precipitate labeled proteins (UNIT 17.10A) in screw-cap microcentrifuge tubes.

The β -elimination reaction is performed in screw-cap microcentrifuge tubes to prevent the lids from popping open during the lengthy incubation, which would result in evaporation of the sample. The tubes should be tightly sealed. Take care in opening the tubes at the end of the reaction, as gas generated during the reaction will escape.

2. Resuspend sample in 500 μl of β -elimination reagent and incubate at 37°C for 18 hr.
After several hours check that the pH is >13. Add more β -elimination reagent if needed.
3. Cool the sample on ice.
4. Neutralize the reaction by adding 5 μl cold 4 M acetic acid in a stepwise manner. Check that the pH is between pH 6 and 7.

Samples can be desalted either by chromatography on a Sephadex G-50 column (1 \times 30 cm, equilibrated in 50 mM ammonium formate, 0.1% SDS) or by anion-exchange chromatography on a 1-ml Bio-Rad Dowex AG 50W-X2 200-400 mesh (H^+ form) column equilibrated in water. Fractions containing [^3H]GlcNAc or [^3H]Gal are pooled and lyophilized. Residual NaBH_4 is removed by washing the sample with methanol; NaBH_4 is volatile in the presence of methanol and is removed in a Speed-Vac evaporator or under a stream of nitrogen (Fukuda, 1990).

5. Resuspend the sugar alditols in Milli-Q water. Analyze by size-exclusion chromatography or by Dionex chromatography.

To determine the size of the oligosaccharide, labeled glycans released by β -elimination are subjected to size exclusion chromatography. Readers are referred to several standard methods using BioGel P4 (Kobata, 1994) or TSK Fractogel (Fukuda, 1990) chromatography. Alternatively, the Amersham Pharmacia Biotech Superdex Peptide column, equilibrated in 30% v/v CH_3CN , 0.1% v/v TFA, has been used to size oligosaccharides (R.N. Cole, pers. commun.).

To determine the nature of the monosaccharide alditol or disaccharide alditol generated from either metabolic labeling or galactosyltransferase labeling, samples released by β -elimination can be analyzed by high-voltage paper electrophoresis or high-pH anion exchange chromatography (HAPEC) with pulsed amperometric detection on a Dionex CarboPac PA100 column (Townsend et al., 1990; Hardy and Townsend, 1994).

Preparation and
Analysis of
Glycoconjugates

17.6.13

SITE MAPPING BY MANUAL EDMAN DEGRADATION

This protocol describes methods for mapping sites on peptides that have been labeled with [³H]Gal (see Alternate Protocol 1) or [³H]glucosamine (see Basic Protocol 5). Proteins of interest can be digested in solution or in gel (Riviere and Tempst, 1995; Stone and Williams, 1995) and the peptides separated by conventional reversed-phase HPLC (UNIT 10.14).

Glycopeptides are covalently attached to arylamine-derivatized PVDF disks via the activation of peptide carboxyl groups using water-soluble *N*-ethyl-*N'*-dimethylamine-nopropylcarbodiimide (EDC; Kelly et al., 1993). Note that both the C-terminus and any acidic residue will couple to the membrane and as such, cycles with acidic amino acids will yield blanks.

Some researchers have shown that glycosylated amino acids can be visualized during automated Edman degradation. However, this technique requires at least 20 pmol of starting material, where at least 20% of the sample is glycosylated. For more details, readers are referred to Zachara and Gooley (2000).

Materials

- Acetonitrile (CH₃CN; HPLC-grade)
- Trifluoroacetic acid (TFA; sequencing grade)
- Sequelon-AA Reagent Kit (Millipore) containing:
 - Mylar sheets
 - Carbodiimide
 - Coupling buffer
 - N*-ethyl-*N'*-dimethylamine-nopropylcarbodiimide (EDC)
- Labeled sample and control peptides (must contain ≥1000 dpm and not be in amine-containing buffers such as Tris)
- Methanol (HPLC-grade)
- Sequencing reagent (see recipe)
- 100 mM Tris·Cl, pH 7.4 (APPENDIX 2)
- Heating block
- Screw-capped polypropylene microcentrifuge tubes

Couple the sample to a PVDF disk

1. Resuspend peptide in 10% to 30% (v/v) CH₃CN and up to 0.1% (v/v) TFA.
2. Place disk on a Mylar sheet (from Sequelon kit) on a heating block at 55°C.
3. Wet disk with 10 μl of methanol and allow excess methanol to evaporate.
4. Apply sample in 10-μl aliquots, allowing the membrane to come to near-dryness between aliquots.
5. Dry the membrane after all of the sample has been applied. Place the disc in the lid of a screw-cap microcentrifuge tube for subsequent reactions.
6. Prepare coupling reagent by combining 1 mg carbodiimide per 100 μl of coupling buffer (both provided with Sequelon kit). Add ~1 mg of EDC per 100 μl of coupling reagent and carefully pipet ~50 μl onto each disk.
7. Incubate samples at 4°C for 30 min. At the end of the reaction, remove the coupling reagent to a microcentrifuge tube.

Incubation of the sample at 4°C increases the yield. The membrane should not be incubated for >30 min.

8. Place the membrane in a microcentrifuge tube. Wash the membrane alternately three times with 1 ml methanol and 1 ml Milli-Q water, vortexing briefly after each addition. Combine each wash with the coupling reagent from step 7.
9. Dry the membrane. Samples are stable at -20°C on disks for at least 6 months.
10. Count an aliquot of the pooled coupling reagent and washes (step 8) and determine the coupling efficiency.

Sequencing peptides

The times given in this procedure are critical. Steps 12 to 17 represent a “cycle.” This method is only efficient for 10 to 20 cycles.

11. Place disk in a screw-capped polypropylene microcentrifuge tube.
12. Add 0.5 ml of the sequencing reagent to the disk and incubate 10 min at 50°C in a heating block.

This step derivatizes the N-terminus of the peptide, forming the phenylthiocarbamyl derivative.

13. Wash disk five times in microcentrifuge tube, each time with 1 ml of methanol, vortexing after each addition. Dry disk in a Speed-Vac evaporator for 5 min.
14. Add 0.5 ml TFA and incubate at 50°C for 6 min. Remove and save the supernatant.

In this step, the derivatized amino acid is released from the peptide.

15. Wash the disk with 1 ml of methanol. Save the supernatant and combine the supernatant from step 14.
16. Wash disk five times, each time with 1 ml methanol.
17. Return to step 12 and repeat the cycle.

Steps 12 to 17 represent a “cycle,” with one amino acid being released per cycle from the N-terminus of the protein or peptide. This method is only efficient for 10 to 20 cycles. As each cycle represents the release of an amino acid, it is not necessary to perform 20 cycles if the peptide in question is 10 amino acids long. The authors usually perform 15 cycles, as the data become hard to interpret after this.

18. Dry the combined supernatants from steps 14 and 15 on a 50°C heating block, or on the bench overnight.
19. Add 0.5 ml of 100 mM Tris·Cl, pH 7.4 and 15 ml liquid scintillation fluid. Count on a liquid scintillation counter.

Edman degradation is a sequential process, with one amino acid being released from the N-terminus of the protein or peptide per cycle. In the method described here, glycosylated amino acids are being detected by following the [^3H]Gal label.

To interpret the data, plot the dpm per cycle as a bar graph. A cycle with a glycosylated amino acid should have much higher counts than either the preceding or following cycle, unless there are two sites in a row.

Manual Edman degradation is often performed in concert with other techniques that will identify the peptide being sequenced, e.g., conventional Edman degradation or MALDI-TOF MS (UNITS 10.21 & 10.22). When looking at data sets from several techniques, the cycle with the counts should line up with a Ser or Thr residue.

SITE MAPPING BY MASS SPECTROMETRY AFTER β -ELIMINATION

As mass spectrometry (MS; *UNITS 10.21 & 10.22*) has become more readily available, several MS techniques have been applied to the characterization of O-GlcNAc proteins and peptides (Reason et al., 1992; Greis et al., 1996; Haynes and Aebersold, 2000). When analyzing proteins or peptides modified by O-GlcNAc via mass spectrometry, several challenges need to be overcome. Some researchers have shown that when using matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF; *UNIT 10.22*) with a conventional matrix (such as α -cyano-4-hydroxycinnamic acid) the addition of a single GlcNAc to a peptide may reduce the signal by approximately five-fold. In addition, the presence of the unglycosylated peptide further suppresses the signal (Hart et al., 2000). Secondly, the β -O-GlcNAc bond is more labile in the mass spectrometer than other bonds formed by protein glycosylation. Thus, in conventional electrospray ionization-MS, O-GlcNAc is released at lower orifice potentials and collision energies than is required to sequence peptides.

Greis et al. (1996) developed a technique where peptides are analyzed before and after β -elimination. During β -elimination Ser (89 atomic mass units) and Thr (101 atomic mass units) residues are converted to 2-aminopropenoic acid (69 atomic mass units) and α -aminobutyric acid (83 atomic mass units), respectively. These amino acids have different masses than their parent amino acids and can be used to map the site of attachment. It should be noted that this method will also release phosphate linked to Ser and Thr residues. In the protocol described here, glycans are released from peptides by β -elimination, and then analyzed using conventional MS-MS strategies.

Materials

Samples from β -elimination (see Basic Protocol 7)

200 mM NaOH

300 mM acetic acid

Screw-cap polypropylene microcentrifuge tubes

45°C heating block

Additional reagents and equipment for ESI-MS (*UNIT 10.21*), reversed-phase HPLC of peptides (*UNIT 10.14*), and MALDI-TOF MS (*UNIT 10.22*)

1. Combine the following in a screw-cap polypropylene microcentrifuge tube:

≤ 25 μ l sample from β -elimination

25 μ l 200 mM NaOH

Milli-Q water to a final volume of 50 μ l

2. Incubate at 45°C for 4 hr in a heating block.

Time is critical, as increased incubation times will result in the peptide backbone hydrolyzing.

3. Place sample on ice and add 25 μ l of 300 mM acetic acid in 5- μ l increments.

The reaction mixture will boil over unless the acetic acid is added in a stepwise manner.

4. Dry sample in a Speed-Vac evaporator and store at -80°C until analyzed by ESI-MS (*UNIT 10.21*). Alternatively, analyze peptides by MALDI-TOF MS (*UNIT 10.22*) after desalting using reversed-phase chromatography (*UNIT 10.14*).

Zip-Tips (Millipore) are ideal for desalting small quantities of peptide before MALDI-TOF MS.

ASSAY FOR OGT ACTIVITY

The detection and analysis of O-GlcNAc on proteins is only the first step in the analysis of O-GlcNAc and the protein(s) of interest. More important is determining the function of the modification. Protocols for the analysis of the enzymes that add and remove O-GlcNAc have been included, as they may aid in understanding the role of O-GlcNAc. Recent examples where studies such as this have been critical include those which have shown the reciprocity between O-GlcNAc and O-phosphate on the C-terminal domain of RNA Pol II; studies showing elevated activity of enzymes in certain tissue/cell lines and tissue fractions; and, finally, studies which have indicated that the enzymes responsible for the addition and removal of O-GlcNAc copurify with kinases and phosphatases.

O-GlcNAc transferase (OGT), or uridine diphospho-*N*-acetylglucosamine:polypeptide β -*N*-acetylglucosaminyltransferase, transfers GlcNAc to the hydroxyl groups of Ser and Thr residues of proteins and peptides using UDP-GlcNAc as a donor substrate (Haltiwanger et al., 1992; Kreppel et al., 1997). OGT activity is assayed by determining the rate at which [³H]GlcNAc is transferred to an acceptor peptide. A number of peptides have been identified as substrates for OGT in vitro, but a peptide (³⁴⁰PGGSTPVSSANMM³⁵²) from the α -subunit of casein kinase II (CKII) is the most efficient in vitro substrate known to date (Kreppel and Hart, 1999).

OGT activity can be assayed in crude preparations (Haltiwanger et al., 1992) or using recombinant protein (Kreppel and Hart, 1999). OGT activity is sensitive to salt inhibition, so it is important to desalt the preparation before assaying if high salt concentrations are present (see Support Protocol 5). For pure preparations 0.2 to 1 μ g is typically used per assay, in crude preparations 20 to 50 μ g, though the latter is precipitated using ammonium sulfate (40% to 60%).

Materials

0.1 mCi/ml UDP-[³H]GlcNAc (20 to 45 Ci/mmol; NEN Life Science Products) in 70% ethanol

25 mM 5'-adenosine monophosphate (5'-AMP), in Milli-Q water, pH 7.0

Nitrogen source

Crude or purified OGT sample, desalted (see Support Protocol 5)

10 \times OGT assay buffer (see recipe)

CKII peptide substrate (⁺H₂N-PGGSTPVSSANMM-COO⁻): dissolve in H₂O to 10 mM and adjust to pH 7 if necessary

50 mM formic acid

Methanol (HPLC-grade)

Waters Sep-Pak C₁₈ cartridges

1. Dry down an aliquot of UDP-[³H]GlcNAc in a Speed-Vac evaporator or under a stream of nitrogen just prior to use. Resuspend in an appropriate volume of 25 mM 5'-AMP, so that the concentration is 0.02 μ Ci to 0.1 μ Ci/ μ l.

AMP is included in the assay to competitively inhibit any pyrophosphatase in the sample that will hydrolyze the UDP-GlcNAc.

2. Set up assay reactions as follows:

5 μ l of 10 \times OGT assay buffer

10 μ l of 10 mM CKII peptide substrate

5 μ l of 0.02 to 0.1 μ Ci/ μ l UDP-[³H]GlcNAc (from 0.1 to 0.5 μ Ci total)

\leq 25 μ l of desalted OGT to be analyzed

H₂O to 50 μ l

It is critical to include a negative control. A mimic of the CKII peptide where the Ser and Thr residues are replaced with Ala is appropriate, or, simply, a “no-enzyme” control can be included.

The results generated are variable and the reactions should be set up in duplicate.

3. Incubate at room temperature for 30 min.

An incubation time of 15 to 30 min at room temperature is usually sufficient.

4. Stop reaction by adding 450 μl of 50 mM formic acid.
5. Wet a Waters Sep-Pak C_{18} cartridge with methanol, then wash the cartridge with 5 ml H_2O .
6. Load the reaction (500 μl total) onto the cartridge with a syringe. Wash with 5 ml H_2O .

The CKII peptide binds to the matrix of the C_{18} cartridge. Unincorporated UDP- ^3H GlcNAc is eliminated by the wash.

7. Elute the peptide with 2 to 4 ml methanol, directly into a 15-ml liquid scintillation counter tube.
8. Add 10 ml scintillation fluid and count ^3H . Calculate OGT activity according to the following equations.

$$\mu\text{Ci of GlcNAc incorporated} = (\text{dpm in sample} - \text{dpm in blank}) / (2.22 \times 10^6 \text{ dpm} \cdot \mu\text{Ci}^{-1})$$

$$\text{mmol of GlcNAc incorporated} = (\mu\text{Ci of GlcNAc incorporated}) / (\text{specific activity in } \mu\text{Ci/mmol})$$

This number should be expressed in terms of mg of OGT or cell extract. If the assay is done at several time points, it can be expressed as mmol/min.

The activity can be expressed either as dpm ^3H incorporation into the peptide, or as $\mu\text{mol } ^3\text{H}$ incorporation ($1 \mu\text{Ci} = 2.22 \times 10^6 \text{ dpm}$).

SUPPORT PROTOCOL 5

DESALTING THE O-GlcNAc TRANSFERASE

OGT activity is sensitive to salt inhibition ($\text{IC}_{50} = 40$ to 50 mM NaCl). It is important to desalt the enzyme preparation before assay if high concentration of salt is present.

Materials

- Sephadex G-50 slurry (Pharmacia Biotech)
- OGT desalting buffer (see recipe)
- Protein sample for OGT assay in volume $\leq 200 \mu\text{l}$
- 1-ml tuberculin syringe
- 1.5-ml tubes, prechilled

1. Pack a column containing exactly 1 ml of Sephadex G-50 slurry in a 1-ml tuberculin syringe. Wash the column with 5 ml of OGT desalting buffer.

Sephadex G-50 is usually supplied in 20% (v/v) ethanol.

2. Load protein sample onto column.

The volume of sample can be up to $200 \mu\text{l}$.

3. Wash column with desalting buffer so that the total volume of this wash and the protein sample is $350 \mu\text{l}$. For example, if sample volume is $150 \mu\text{l}$, add $200 \mu\text{l}$ desalting buffer to the column at this step.

4. Transfer the syringe column to a clean, prechilled 1.5-ml tube. Elute protein with 200 μ l desalting buffer. Keep on ice.

This is the desalted sample.

Alternatively, PD-10 desalting columns (Amersham Pharmacia Biotech) can be used if O-GlcNAc transferase is in larger volume (1 to 2.5 ml).

ASSAY FOR O-GlcNAcase ACTIVITY

O-GlcNAcase, also known as N-acetylglucosaminidase or hexosaminidase C (EC 3.2.1.52), is a cytosolic glycosidase specific for O-linked β -GlcNAc. The activity of O-GlcNAcase can be conveniently assayed in vitro with a synthetic substrate, *p*-nitrophenol N-acetylglucosaminide (*p*NP- β -GlcNAc). The cleavage product, *p*NP, has an absorbance peak at 400 nm.

Materials

Partially purified O-GlcNAcase (0.2 to 1 μ g) or cell extract sample (20 to 50 μ g, precipitated with 30% to 50% ammonium chloride)

10 \times O-GlcNAcase assay buffer (see recipe)

100 mM (50 \times) *p*-nitrophenol N-acetylglucosaminide (*p*NP-GlcNAc) in DMSO

500 mM Na₂CO₃

96-well flat bottom plates or 1.5 ml microcentrifuge tubes

Plate reader or spectrophotometer

1. Prepare O-GlcNAcase.

Native or recombinant O-GlcNAcase can be partially purified from animal tissues or cultured cells by several chromatographic steps (Dong and Hart, 1994; Gao et al., 2001). Alternatively, a crude enzyme preparation can be generated by passing cell extract over a 1-ml Con-A column. Most of the interfering acidic hexosaminidases are modified by N-linked sugars and bind to Con-A, while neutral O-GlcNAcase is in the flow-through (Izumi and Suzuki, 1983).

2. Precool 96-well plate or microcentrifuge tubes on ice.

3. Set up reactions in the precooled plate wells or tubes as follows:

1 to 50 μ l partially purified O-GlcNAcase enzyme or cell extract

10 μ l 10 \times O-GlcNAcase assay buffer

2 μ l 100 mM *p*NP-GlcNAc

H₂O to 100 μ l

The total reaction volume can be scaled up to 500 μ l in microcentrifuge tubes.

**p*NP-GlcNAc breaks down chemically. A blank reaction without enzyme should be included to determine the background.*

50 mM GalNAc is included in the reaction to inhibit lysosomal hexosaminidases A and B which may be present in the enzyme preparation. O-GlcNAcase is not inhibited by 50 mM GalNAc.

4. Mix well and cover.

5. Incubate at 37°C for 30 min to 4 hr.

*Yellow color will develop as *p*NP-GlcNAc is hydrolyzed by O-GlcNAcase. Reactions should be optimized to keep the absorbance within the linear range of the spectrophotometer. The authors find that 20 to 50 μ g of cell extract used in a reaction of 100 μ l, with a 1 to 2 hr incubation time is appropriate.*

6. At the end of incubation, add an equal volume of 500 mM Na₂CO₃ to each well (100 μl or 500 μl).

Na₂CO₃ raises the pH to >pH 9.0, intensifying the yellow color and stopping the reaction, as O-GlcNAcase has little activity at pH 9 to 10.

7. Read the absorbance at 400 nm on a plate reader or spectrophotometer.
8. Calculate O-GlcNAcase activity according to the following equation:

$$\text{mM of GlcNAc released} = A_{400} / (17.4 \times 10 \text{ mM}^{-1} \cdot \text{cm}^{-1} \times \text{pathlength})$$

One unit is the amount of enzyme catalyzing the release of 1 μmol/min of pNP from pNP-GlcNAc.

The molar extinction coefficient for pNP is 17.4 × 10 mM⁻¹·cm⁻¹ at pH 10. The path length for 200 μl on a 96-well plate is 0.71 cm.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water or equivalent for the preparation of all buffers. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Biosynthetic labeling medium

Glucose-free culture medium containing:

50 μCi/ml D-[6-³H]glucosamine (22 Ci/mmol; Amersham Pharmacia Biotech)

10% (v/v) FBS

Prepare fresh

Buffer H

50 mM HEPES, pH 6.8

50 mM NaCl

2% (v/v) Triton X-100

Store up to 1 month at room temperature

Citrate-phosphate buffer, pH 4.0, 2×

Dissolve 12.9 g citric acid monohydrate (mol. wt. 210) and 20.6 g disodium hydrogen phosphate heptahydrate (Na₂HPO₄·6H₂O) in 300 ml Milli-Q water. Bring volume to 500 ml.

Galactosyltransferase labeling buffer, 10×

100 mM HEPES, pH 7.5

100 mM galactose

50 mM MnCl₂

Store up to 1 month at 4°C

Galactosyltransferase storage buffer

2.5 mM HEPES, pH 7.4

2.5 mM MnCl₂

50% (v/v) glycerol

Store up to 1 month at room temperature

Hexosaminidase reaction mixture, 2×

Per reaction:

25 μ l 2 \times citrate-phosphate buffer (see recipe)
1 U *N*-acetyl- β -D-glucosaminidase (V-Labs)
0.01 U aprotinin
1 μ g leupeptin
1 μ g α_2 -macroglobulin

O-GlcNAcase assay buffer, 10×

500 mM sodium cacodylate, pH 6.4
500 mM *N*-acetylgalactosamine (GlcNAc)
3% (w/v) bovine serum albumin (BSA)
Prepare fresh

OGT assay buffer, 10×

500 mM sodium cacodylate, pH 6.0
10 mg/ml bovine serum albumin (BSA)
10 mM 1-amino-GlcNAc (2-acetamido-1-amino-1,2-dideoxy- β -D-glucopyranose;
Sigma)
Prepare fresh

OGT desalting buffer

20 mM Tris·Cl, pH 7.8 (APPENDIX 2)
1 mg/ml bovine serum albumin (BSA)
20% (v/v) glycerol
0.02% (w/v) NaN₃
Store up to 1 week at 4°C

Protease inhibitors, 1000×

PIC 1, 1000 \times :

Dissolve the following in 10,000 U/ml aprotinin solution (Sigma)
1 mg/ml leupeptin
2 mg/ml antipain
10 mg/ml benzamide

PIC 2, 1000 \times :

Prepare in DMSO
1 mg/ml chemostatin
2 mg/ml pepstatin

PMSF, 1000 \times :

0.1 M phenylmethylsulfonyl fluoride in 95% ethanol

Sequencing reagent

7 ml methanol (HPLC-grade)
1 ml triethylamine (TEA; sequencing grade)
1 ml phenylisothiocyanate (PITC; sequencing grade)
1 ml Milli-Q water
Stable 24 hr at 4°C

sWGA Gal elution buffer

Phosphate-buffered saline (PBS; APPENDIX 2) containing:
0.2% (v/v) NP-40
1 M D-(+)-galactose (Gal)
Store up to 1 week at 4°C

sWGA GlcNAc elution buffer

Phosphate-buffered saline (PBS; *APPENDIX 2*) containing:

0.2% (v/v) NP-40

1 M N-acetylglucosamine (GlcNAc)

Store up to 1 week at 4°C

TBS-HD

10 mM Tris-Cl, pH 7.5 (*APPENDIX 2*)

150 mM NaCl

1% (v/v) Triton X-100

0.1% (w/v) SDS

0.25% (w/v) deoxycholic acid

Store up to 1 month at 4°C

TBS-HT

10 mM Tris-Cl, pH 7.5 (*APPENDIX 2*)

150 mM NaCl

0.3% (v/v) Tween 20

Store up to 1 month at room temperature

TBST

10 mM Tris-Cl, pH 7.5 (*APPENDIX 2*)

150 mM NaCl

0.05% (v/v) Tween 20

Store up to 1 month at room temperature

Tris-buffered saline (TBS)

10 mM Tris-Cl, pH 7.5 (*APPENDIX 2*)

150 mM NaCl

Store up to 1 month at room temperature

COMMENTARY

Background Information

(β)-D-1-4-galactosylaminyltransferase from bovine milk recognizes terminal N-acetylglucosamine (GlcNAc) residues and modifies them by the addition of a single Gal residue. Torres and Hart (1984) first used this enzyme in combination with UDP-[3 H]Gal to demonstrate that bovine lymphocytes contain proteins modified by O-linked GlcNAc. Further refinements of this experiment led them to propose that the product, β Gal1-4 β GlcNAc, was the result of the galactosyltransferase recognizing and modifying a single GlcNAc residue O-linked to Ser/Thr residues of nuclear and cytoplasmic proteins (Holt and Hart, 1986). Since this report, many cytosolic and nuclear proteins from mammalian cells were shown to be modified by O-GlcNAc. This method has remained the “gold standard” technique for the detection of O-GlcNAc-modified proteins, as the label provides a “tag” for subsequent analyses, such as those described under Product Characterization, below (see Critical Parameters and Troubleshooting).

In subsequent years, methods such as WGA affinity and western blotting with GlcNAc-specific lectins and antibodies have become popular as simple techniques for the initial characterization of target proteins.

Critical Parameters and Troubleshooting

Extraction of proteins from cells

The O-GlcNAc modification can be removed from proteins either by cytosolic O-GlcNAcase or lysosomal hexosaminidases. The inclusion of inhibitors during the extraction and purification process will preserve the levels of O-GlcNAc on proteins. Commonly used inhibitors (Dong and Hart, 1994) include 1-amino-GlcNAc (1 mM), GlcNAc (100 mM), and PUGNAc (5 μ M). Note that these may have to be removed, as they will act as inhibitors in other methods.

Product characterization

Product characterization is a critical step showing that a protein is modified by O-GlcNAc, and not other glycans. While many proteins modified by O-GlcNAc have been identified, there is evidence based on metabolic labeling (Medina et al., 1998) and lectin labeling studies (Hart et al., 1989) that indicate that O-GlcNAc is not the only intracellular carbohydrate post-translational modification. In addition, at least one peptide mimic of O-GlcNAc has been identified in cytokeratins (Shikhman et al., 1994).

Moreover, many techniques used for breaking open cells also release proteins that are modified by complex N- and O-linked sugars, which may contain terminal GlcNAc. Many of the techniques described in this unit will recognize any GlcNAc residue, and it is important to perform the described controls such as PNGase F digestion to show specificity.

Product analysis is critical for metabolic labeling with glucosamine. While UDP-GlcNAc is the major product, glucosamine can enter other biosynthetic pathways, such as those used for amino acid synthesis. This issue was highlighted by studies of the SV40 large T-antigen. Some researchers have found that the SV40 large T-antigen labels with a number of different tritiated carbohydrates. However, O-GlcNAc is the only carbohydrate post-translational modification of the SV40 large T-antigen. The incorporation of glucosamine into amino acid biosynthetic pathways could be reduced by growing cells in the presence of excess nonessential amino acids (Medina et al., 1998).

Lastly, while galactosyltransferase is specific for terminal GlcNAc residues, researchers (Elling et al., 1999) have shown that galactosyltransferase will modify GlcNAc linked in either the α - or β -anomeric conformation. The authors of this unit have shown that proteins modified by α -O-GlcNAc will be labeled using the procedure described (N. Zachara, unpub. observ.). While α -O-GlcNAc has not been identified in complex eukaryotes, it is a common modification of cell surface proteins of simple eukaryotes such as trypanosomes and *Dictyostelium*. Product analysis, such as HPAEC of the sugar alditols, will resolve many of the issues discussed.

Time Considerations

Detection of O-GlcNAc proteins using antibodies (see Basic Protocol 2) and lectins (see

Basic Protocol 3) will take approximately 2 to 3 days after the extraction of the proteins from cells. Samples and controls must be treated with PNGase F and/or hexosaminidase (1 hr to overnight), before SDS-PAGE and blotting. In either case, overnight incubation at 4°C provides the best signal-to-noise ratio.

WGA affinity chromatography of low-copy-number proteins will take 2 to 3 days. The ITT and WGA affinity chromatography can be completed in 1 day; subsequent analysis of the product by SDS-PAGE will take 1 to 2 days depending on the label used and the amount of label incorporated into proteins eluting from the WGA-agarose.

Autogalactosylation of the galactosyltransferase and subsequent analysis of the activity will take 1 to 2 days. As the enzyme is stable for 6 to 12 months at -20°C , autogalactosylation does not need to be repeated for each analysis. Labeling of the proteins can take several hours to overnight, though optimization of the conditions may take a few days. The subsequent analysis, as well as desalting (dependent on the technique used), PNGase F digestion (1 hr to overnight), precipitation of protein (3 hr to overnight), SDS-PAGE, and autoradiography (1 to 10 days), can take up to 2 weeks. The length of time allotted to product analysis is dependent on the methods chosen, but will almost certainly require 7 to 10 days. Further analysis, including digestion of labeled proteins and subsequent purification of peptides, will take at least 3 days.

Literature Cited

- Amano, J. and Kobata, A. 1990. Quantitative conversion of mucin-type sugar chains to radioactive oligosaccharides. *Methods Enzymol.* 179:261-270.
- Batteiger, B., Newhall, W.J., and Jones, R.B. 1982. The use of Tween 20 as a blocking agent in the immunological detection of proteins transferred to nitrocellulose membranes. *J. Immunol. Methods* 55:297-307.
- Brew, K., Vanaman, T.C., and Hill, R.L. 1968. The role of alpha-lactalbumin and the A protein in lactose synthetase: A unique mechanism for the control of a biological reaction. *Proc. Natl. Acad. Sci. U.S.A.* 59:491-497.
- Comer, F.I. and Hart, G.W. 2000. O-Glycosylation of nuclear and cytosolic proteins: Dynamic interplay between O-GlcNAc and O-phosphate. *J. Biol. Chem.* 275:29179-29182.
- Comer, F.I., Vosseller, K., Wells, L., Accavitti, M.A., and Hart, G.W. 2001. Characterization of a mouse monoclonal antibody specific for O-linked N-acetylglucosamine. *Anal. Biochem.* 293:169-177.

- Dong, D.L. and Hart, G.W. 1994. Purification and characterization of an O-GlcNAc selective N-acetyl- β -D-glucosaminidase from rat spleen cytosol. *J. Biol. Chem.* 269:19321-19330.
- Duk, M., Ugorski, M., and Lisowska, E. 1997. Beta-elimination of O-glycans from glycoproteins transferred to immobilized P membranes: Method and some applications. *Anal. Biochem.* 253:98-102.
- Elling, L., Zervosen, A., Gallego, R.G., Nieder, V., Malissard, M., Berger, E.G., Vliegthart, J.F., and Kamerling, J.P. 1999. UDP-N-Acetyl-alpha-D-glucosamine as acceptor substrate of beta-1,4-galactosyltransferase: Enzymatic synthesis of UDP-N-acetylglucosamine. *Glycocon. J.* 16:327-336.
- Fukuda, M. 1990. Characterization of O-linked saccharide structures from cell surface glycoproteins. *Methods Enzymol.* 179:17-29.
- Gao, Y., Parker, G.J., and Hart, G.W. 2000. Streptozotocin-induced β -cell death is independent of its inhibition of O-GlcNAcase in pancreatic Min6 cells. *Arch. Biochem. Biophys.* 383:296-302.
- Gao, Y., Wells, L., Comer, F.I., Parker, G.J., and Hart, G.W. 2001. Dynamic O-glycosylation of nuclear and cytosolic proteins: Cloning and characterization of a neutral, cytosolic β -N-acetylglucosaminidase from human brain. *J. Biol. Chem.* 276:9838-9845.
- Greis, K.D. and Hart, G.W. 1998. Analytical methods for the study of O-GlcNAc glycoproteins and glycopeptides. *Methods Mol. Biol.* 76:19-33.
- Greis, K.D., Hayes, B.K., Comer, F.I., Kirk, M., Barnes, S., Lowary, T.L., and Hart, G.W. 1996. Selective detection and site-analysis of O-GlcNAc-modified glycopeptides by beta-elimination and tandem electrospray mass spectrometry. *Anal. Biochem.* 234:38-49.
- Haltiwanger, R.S., Blomberg, M.A., and Hart, G.W. 1992. Glycosylation of nuclear and cytoplasmic proteins: Purification and characterization of a uridine diphospho-N-acetylglucosamine:polypeptide β -N-acetylglucosaminyltransferase. *J. Biol. Chem.* 267:9005-9013.
- Haltiwanger, R.S., Grove, K., and Philipsberg, G.A. 1998. Modulation of O-linked N-acetylglucosamine levels on nuclear and cytoplasmic proteins in vivo using the peptide O-GlcNAc- β -N-acetylglucosaminidase inhibitor O-(2-acetamido-2-deoxy-D-glucopyranosylidene) amino-N-phenylcarbamate. *J. Biol. Chem.* 273:3611-3617.
- Han, I., Oh, E.-S., and Kudlow, J.E. 2000. Responsiveness of the state of O-linked N-acetylglucosamine modification of nuclear pore protein p62 to the extracellular glucose concentration. *Biochem. J.* 350:109-114.
- Hardy, M.R. and Townsend, R.R. 1994. High-pH anion exchange chromatography of glycoprotein-derived carbohydrates. *Methods Enzymol.* 230:208-225.
- Hart, G.W., Holt, G.D., Haltiwanger, R.S., and Kelly, W.G. 1989. Cytoplasmic and nuclear glycosylation. *Annu. Rev. Biochem.* 58:841-874.
- Hart, G.W., Cole, R.N., Kreppel, L.K., Arnold, C.S., Comer, F.I., Iyer, S., Cheng, X., Carroll, J., and Parker, G.J. 2000. Glycosylation of proteins: A major challenge in mass spectrometry and proteomics. In *Proceedings of the 4th International Symposium on Mass Spectrometry in the Health and Life Sciences* (A. Burlingame, S. Carr, and M. Baldwin, eds.) pp. 365-382. Humana Press, Totowa, N.J.
- Hayes, B.K., Greis, K.D., and Hart, G.W. 1995. Specific isolation of O-linked N-acetylglucosamine glycopeptides from complex mixtures. *Anal. Biochem.* 228:115-122.
- Haynes, P.A. and Aebersold, R. 2000. Simultaneous detection and identification of O-GlcNAc-modified glycoproteins using liquid chromatography-mass spectrometry. *Anal. Chem.* 72:5402-5410.
- Heart, E., Choi, W.S., and Sung, C.K. 2000. Glucosamine-induced insulin resistance in 3T3-L1 adipocytes. *Am. J. Physiol. Endocrinol. Metab.* 278:E103-E112.
- Holt, G.D. and Hart, G.W. 1986. The subcellular distribution of terminal N-acetylglucosamine moieties: Localization of a novel protein-saccharide linkage, O-linked GlcNAc. *J. Biol. Chem.* 261:8049-8057.
- Holt, G.D., Snow, C.M., Senior, A., Haltiwanger, R.S., Gerace, L., and Hart, G.W. 1987. Nuclear pore complex glycoproteins contain cytoplasmically disposed O-linked N-acetylglucosamine. *J. Cell Biol.* 104:1157-1164.
- Izumi, T. and Suzuki, K. 1983. Neutral β -N-Acetylhexosaminidases of rat brain: Purification and enzymatic and immunological characterization. *J. Biol. Chem.* 258:6991-6999.
- Kelly, W.G., Dahmus, M.E., and Hart, G.W. 1993. RNA polymerase II is a glycoprotein: Modification of the COOH-terminal domain by O-GlcNAc. *J. Biol. Chem.* 268:10416-10424.
- Kobata, A. 1994. Size fractionation of oligosaccharides. *Methods Enzymol.* 230:200-208.
- Kreppel, L.K. and Hart, G.W. 1999. Regulation of a cytosolic and nuclear O-GlcNAc transferase: Role of the tetratricopeptide repeats. *J. Biol. Chem.* 274:32015-32022.
- Kreppel, L.K., Blomberg, M.A., and Hart, G.W. 1997. Dynamic glycosylation of nuclear and cytosolic proteins: Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats. *J. Biol. Chem.* 272:9308-9315.
- Medina, L., Grove, K., and Haltiwanger, R.S. 1998. SV40 large T antigen is modified with O-linked N-acetylglucosamine but not with other forms of glycosylation. *Glycobiology* 8:383-391.
- Reason, A.J., Morris, H.R., Panico, M., Marais, R., Treisman, R.H., Haltiwanger, R.S., Hart, G.W., Kelly, W.G., and Dell, A. 1992. Localization of O-GlcNAc modification on the serum response

- transcription factor. *J. Biol. Chem.* 267:16911-16921.
- Riviere L.R., and Tempst, P. 1995. Enzymatic digestion of proteins in solution. *In Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 11.1.1-11.1.19. John Wiley & Sons, New York.
- Roos, M.D., Xie, W., Su, K., Clark, J.A., Yang, X., Chin, E., Paterson, A.J., and Kudlow, J.E. 1998. Streptozotocin, an analog of N-acetylglucosamine, blocks the removal of O-GlcNAc from intracellular proteins. *Proc. Assoc. Am. Physicians* 110:422-432.
- Roquemore, E.P., Chou, T.Y., and Hart, G.W. 1994. Detection of O-linked N-acetylglucosamine (O-GlcNAc) on cytoplasmic and nuclear proteins. *Methods Enzymol.* 230:443-460.
- Schnedl, W.J., Ferber, S., Johnson, J.H., and Newgard, C.B. 1994. STZ transport and cytotoxicity: Specific enhancement in GLUT-2 expressing cells. *Diabetes* 43:1326-1333.
- Shafi, R., Iyer, S.P., Ellies, L.G., O'Donnell, N., Marek, K.W., Chui, D., Hart, G.W., and Marth, J.D. 2000. The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proc. Natl. Acad. Sci. U.S.A.* 97:5735-5739.
- Shikhman, A.R., Greenspan, N.S., and Cunningham, M.W. 1994. Cytokeratin peptide SFGSGFGGGY mimics N-acetyl-beta-D-glucosamine in reaction with antibodies and lectins, and induces in vivo anti-carbohydrate antibody response. *J. Immunol.* 153:5593-5606.
- Stone, K. L., and Williams, K.R. 1995. Digestion of proteins in gels for sequence analysis. *In Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 11.3.1-11.3.13. John Wiley & Sons, New York.
- Snow, C.M., Senior, A., and Gerace, L. 1987. Monoclonal antibodies identify a group of nuclear pore complex glycoproteins. *J. Cell Biol.* 104:1143-1156.
- Torres, C.R. and Hart, G.W. 1984. Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. *J. Biol. Chem.* 259:3308-3317.
- Townsend, R.R., Hardy, M.R., and Lee, Y.C. 1990. Separation of oligosaccharides using high-performance anion-exchange chromatography with pulsed amperometric detection. *Methods Enzymol.* 179:65-76.
- Turner, J.R., Tartakoff, A.M., and Greenspan, N.S. 1990. Cytologic assessment of nuclear and cytoplasmic O-linked N-acetylglucosamine distribution by using anti-streptococcal monoclonal antibodies. *Proc. Natl. Acad. Sci. U.S.A.* 87:5608-5612.
- Unverzagt, C., Kunz, H., and Paulson, J.C. 1990. High-efficiency synthesis of sialyloligosaccharides and sialoglycopeptides. *J. Am. Chem. Soc.* 112:9308-9309.
- Varki, A. 1994. Metabolic radiolabeling of glycoconjugates. *Methods Enzymol.* 230:16-32.
- Wells, L., Vosseller, K., and Hart, G.W. 2001. Glycosylation of nucleocytoplasmic proteins: Signal transduction and O-GlcNAc. *Science* 291:2376-2378.
- Zachara, N.E. and Gooley, A.A. 2000. Identification of glycosylation sites in mucin peptides by Edman degradation. *Methods Mol. Biol.* 125:121-128.

Contributed by Natasha E. Zachara,
Robert N. Cole, and Gerald W. Hart
The Johns Hopkins University
Medical School
Baltimore, Maryland

Yuan Gao
Deakin University
Victoria, Australia

Lectin Analysis of Proteins Blotted onto Filters

UNIT 17.7

Lectins are proteins that bind with great specificity to certain carbohydrate structures. Plant lectins are widely used for investigations of carbohydrate structure (UNIT 17.3) and for fractionation and purification of individual oligosaccharides and glycopeptides (UNIT 17.18). This unit describes the use of lectins as sensitive indicators for the presence of certain carbohydrate structures linked to proteins blotted onto filters. A tagged lectin is incubated with a blot containing the target protein and binding of the lectin is detected by one of several different procedures. Direct approaches include using lectins labeled with ^{125}I or conjugated to horseradish peroxidase or alkaline phosphatase, which can be detected by chromogenic or luminescent visualization systems. Indirect approaches involve using lectins conjugated to biotin or digoxigenin followed by a second incubation with alkaline phosphatase–conjugated avidin or antibodies specific for the haptenic digoxigenin group and then by visualization. Several commercial kits are available that provide labeled lectins, control proteins, and developing reagents needed for visualization. These systems can also be adapted for use with lectins other than those supplied with kits. The following protocol is easy to perform with or without a kit. However, the results, while suggestive of carbohydrate structure, are not definitive.

Materials

- Purified protein sample containing 1 to 10 μg of target protein
- Lectin blot kit (e.g., E-Y Laboratories Lectin Staining Kit; Genzyme Lectin Link Kit; Boehringer Mannheim Glycan Differentiation Kit) *or* equivalent materials:
 - Glycoprotein known to bind with the chosen lectins (positive control; Table 17.7.1)
 - Blocking solution: e.g., 1% to 2% (w/v) gelatin in TTBS or 3% periodate-oxidized bovine serum albumin (BSA)
 - Incubation buffer: e.g., 0.1% (v/v) Tween 20 in Tris-buffered saline (TTBS; UNIT 10.8)
 - Labeled or conjugated lectin (Tables 17.7.2 & 17.7.3) in TTBS or other appropriate incubation buffer
 - Enzyme- or antibody-linked antibody (secondary reagent), if needed
 - Chromogenic or luminescent visualization reagent for detecting tagged lectin or antibody (Table 10.8.1)
- Additional reagents and equipment for SDS-PAGE (UNIT 10.2), immunoblotting and immunodetection (UNIT 10.8), and autoradiography (APPENDIX 3)

**BASIC
PROTOCOL**

Table 17.7.1 Lectin-Binding Glycoproteins^a

Glycoprotein	Molecular weight (kDa)	Lectin bound ^b
RNase B	17	Con A
Ovalbumin	45	Con A
Transferrin	80	SNA
Fetuin	68	SNA, MAA
Asialofetuin	61	RCA I
α 1-acid glycoprotein	45	DSA

^aGlycoproteins known to bind with the chosen lectins are used as positive controls and as standards for quantitative analysis. They are supplied with some lectin blot kits (e.g., Boehringer Mannheim's) or available separately from Sigma.

^bAbbreviations: Con A, concanavalin A; SNA, *Sambucus nigra* agglutinin; MAA, *Maackia amurensis* agglutinin; RCA I, *Ricinus communis* agglutinin I; DSA, *Datura stramonium* agglutinin.

**Preparation
and Analysis of
Glycoconjugates**

17.7.1

Contributed by Hudson H. Freeze

Current Protocols in Molecular Biology (1993) 17.7.1-17.7.8

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 23

Table 17.7.2 Lectins Suitable for Analysis of Proteins Blotted onto Filters

Lectin	Specificity ^a	Notes	Kit ^b
Con-A (Concanavalin A)	α Man, α Glc	Strongly indicates N-linked oligosaccharide chains. Binds high mannose, hybrid, or biantennary complex chains	EY-L GN-LL
GNA (<i>Galanthus nivalis</i> agglutinin)	Man α Man	Binds terminal mannose-linked, α 1-3, α 1-6 or α 1-2 linked to mannose. It will identify "high mannose" N-glycan chains or O-glycosidically linked mannose chains in yeast glycoproteins	BM-GD
WGA (Wheat germ agglutinin)	β -GlcNAc or sialic acid	Binds terminal β GlcNAc or sialic acids on various glycans	GN-LL, EY-L
RCA I (<i>Ricinus communis</i> agglutinin)	β Gal	Binds Gal-terminated N-linked oligosaccharides, but not exclusive	GN-LL
DSA (<i>Datura stramonium</i> agglutinin)	Gal β 1-4GlcNAc	Binds Gal β -1-4GlcNAc in complex and hybrid N-glycans and GlcNAc in O-glycans, also polylectosamines	BM-GD GN-LL
MAA (<i>Maackia amurensis</i> agglutinin)	SA α 2-3 Gal- β -GlcNAc	Binds sialic acid linked α 2-3 to galactose in N- and O-linked glycans	BM-GD GN-LL
SNA (<i>Sambucus nigra</i> agglutinin)	SA α 2-6 Gal or GalNAc	Binds sialic acid linked α 2-6 to Gal or GalNAc; identifies sialylated O-glycan, N-glycan chains	BM-GD GN-LL
PNA (Peanut agglutinin)	Gal β 1-3 GalNAc	Recognizes the core disaccharide galactose β 1-3 N-acetylgalactosamine O-glycosidically linked chains (except of yeast glycoproteins)	BM-GD

^aAbbreviations: Man, mannose; Glc, glucose; GlcNAc, N-acetylglucosamine; Gal, galactose; SA, sialic acid; GalNAc, N-acetylgalactosamine.

^bEY-L, E-Y Laboratories Lectin Staining Kit; BM-GD, Boehringer Mannheim Glycan Differentiation Kit; GN-LL, Genzyme Lectin Link Kit.

1. Separate the proteins in the sample by one-dimensional SDS-PAGE and transfer onto nitrocellulose by electroblotting. Include positive control (a glycoprotein known to bind to the lectins) and extra sample lanes for incubation controls (see step 5 annotation).

Inclusion of a positive control is particularly important in early experiments when incubation conditions are being optimized. These procedures are useful for relatively pure proteins. Crude mixtures of proteins are not suitable for these tests.

If some samples are to be digested with exo- or endoglycosidases to facilitate interpretation of results (see critical parameters), digestions should be carried out prior to running the sample on the gel. In this case a gel with separate wells must be used and a glycosidase-only control lane included for each digestion. Many of the glycosidases are themselves glycoproteins and will bind lectins.

2. Incubate nitrocellulose blot in blocking solution with agitation \geq 30 min at room temperature.

Table 17.7.3 Commercial Sources of Lectins for Detecting Blotted Glycoproteins

Supplier	Conjugate				
	Biotin	Peroxidase	Alkaline phosphatase	Digoxigenin	Lectin conjugates in kits
E-Y Labs	+	+	+	—	Avidin/horseradish peroxidase Avidin/alkaline phosphatase
Boehringer Mannheim	+	—	—	+	Anti-digoxigenin Alkaline phosphatase
Genzyme	+	—	—	—	Avidin/alkaline phosphatase
Sigma	+	+	—	—	—
ICN	+	+	—	—	—

Gelatin is used as the blocking reagent because it is not a glycoprotein. Blocking reagents that contain even small amounts of glycoproteins (e.g., nonfat dry milk) may bind to the lectins, potentially depleting the available lectin or producing a high background signal. Alternatively, periodate oxidized BSA (3%) can be used as a blocking reagent (Glass et al., 1981).

3. Wash each filter three times in 50 ml TTBS or other buffer appropriate for incubation with the labeled lectin, 15 min per wash.

Other buffers and washing conditions may be suitable, as discussed in the literature accompanying lectin blot kits, but TTBS is routinely used.

4. Place the blot on a smooth, clean surface, such as a glass plate, and use a clean razor blade to cut the nitrocellulose into multiple strips approximately the width of a single lane.

It is not necessary to cut the strips exactly along the interlane boundaries. If photographs are required, the lane can be trimmed to a desirable width later.

5. Place each strip in a separate container and add sufficient TTBS or other appropriate buffer containing ~1 µg/ml labeled lectin to cover and suspend strip. Incubate 1 hr at room temperature with gentle shaking.

Incubation conditions must be optimized for each lectin and for each detection method. This information is available in kit instructions or from the supplier. Chelators such as EDTA should be avoided, as some lectins (e.g., Con A) require tightly bound metal ions for activity.

A sample incubated without lectin should be prepared as a control. Regardless of the method used to detect lectin binding, it is also important to include controls in which either the primary or secondary reagents are eliminated. If a sufficient amount (50 to 100 mM) of the appropriate competing sugar is available, lectin binding should be carried out in the presence of the sugar as a negative control (the sugar should eliminate specific binding).

6. Remove the lectin solution and store it at 4°C. Wash filter three times in 50 ml TTBS, 10 min each time.

Lectin solutions may be reused several times within one week of their first use if they are stored at 4°C. If a lectin solution is reused, the positive control will detect whether the lectin is still active.

- 7a. For lectins directly tagged with ^{125}I , biotin, horseradish peroxidase, or alkaline phosphatase: Visualize by autoradiography or using the appropriate chromogenic or luminescent visualization solution.
- 7b. For lectins conjugated to biotin or digoxigenin: Incubate filter in the appropriate secondary reagent 1 hr at room temperature. Wash filter 3 times with 50 ml TTBS, 10 min each time. Detect bound lectin using the appropriate chromogenic or luminescent visualization solution.

Lectins conjugated to these groups are detected by incubation with avidin–alkaline phosphatase conjugates or anti-digoxigenin–alkaline phosphatase.

COMMENTARY

Background Information

Lectins have been used for many years to detect the presence of various types of sugar chains. Lectin blots provide quick glimpses of the types of sugars attached to purified proteins. Even though the blots provide useful information, they cannot provide proof of a sugar structure and hence cannot substitute for the more specific and quantitative analytical approaches presented in this chapter. This is because lectin binding and its quantitation depend upon many factors that vary with different individual proteins. The advantage of using lectins is that many different ones are available and analysis requires only modest amounts of protein and routine laboratory equipment. Also, many detection methods are available. In order to have confidence in the results of lectin-binding studies, it is necessary to show that their binding specificity is abolished by coincubation with a competing sugar (usually a monosaccharide) or by prior enzymatic digestion to remove some or all of the protein-linked sugar chains. Lectins labeled with fluorescent indicators can be used to visualize glycoconjugates directly in hydrated gels (West and McMahon, 1977). However, this approach requires several days. This unit describes a less time-consuming method (based on Tanner and Anstee, 1976) in which the proteins are blotted onto filters prior to detection by one of several methods.

Critical Parameters and Troubleshooting

Information concerning critical parameters and troubleshooting for protein blotting, immunoprobeing, and chromogenic and luminescent detection is provided in *UNIT 10.8*.

Experimental design

Two points are especially important in designing lectin-binding experiments: the choice of lectin and the choice of method for detecting

lectin bound to glycoprotein. If the goal of the experiment is to obtain preliminary evidence that a protein contains carbohydrate groups, a battery of lectins should be used along with appropriate controls (e.g., competition by appropriate sugars). If the experiment is meant to serve as a guide to further structural analysis, the effects of various chemical or enzymatic treatments must be combined with analysis using one or more lectins known to bind to the glycoconjugate. Table 17.7.2 shows the specificity of the most popular lectins, some of which are supplied in the various kits.

The choice of detection method is affected by the amount of protein available and the extent of glycosylation. If the sample contains 1 to 10 μg of target protein or if the material is extensively glycosylated, lectins labeled directly with ^{125}I , alkaline phosphatase, or horseradish peroxidase can be used. If either the protein quantity or extent of glycosylation is minimal, an indirect system using lectins that are biotinylated or coupled to digoxigenin and a secondary enzyme- or antibody-linked reagent has the advantage of signal amplification.

Kits

The various reagents can be purchased separately or as prepackaged kits from the suppliers listed in Table 17.7.3. E-Y Laboratories offers the largest selection of purified lectins and lectins tagged with various conjugates.

The E-Y Laboratories kit uses lectins coupled to horseradish peroxidase. The Genzyme and Boehringer Mannheim kits use alkaline phosphatase-coupled molecules to detect the tagged lectins. Alternatively, tagged lectins can easily be prepared in the laboratory (*UNIT 11.1*). The major advantages of using kits are that they are self-sufficient, containing all of the reagents and positive controls needed, and that the lectins they employ are among those with the best characterized specificities. When coupled

with glycosidase digestions, the kits probably provide the most information about the sugar chains in the least amount of time for the least amount of money. Because there are limitations to the information that can be obtained, the decision about which individual lectins or kits to buy will depend upon the precision of information required.

The Glycan Differentiation Kit (Boehringer Mannheim) supplies lectins that are coupled to the steroid digoxigenin using an *N*-hydroxy-succinimide ester. An alkaline phosphatase-conjugated antiserum against this steroid is used to visualize bound lectin on target glycoproteins (Haselbeck et al., 1990). Digoxigenin is a naturally occurring steroid found only in plants of the *Digitalis* family and not in animals, which minimizes background. The sheep Fab fragment to the hapten is affinity purified and coupled to alkaline phosphatase. The company has used the digoxigenin system extensively in different types of kits to detect proteins, glycoproteins, and nucleic acids (UNIT 3.18). Lectin binding is simply an extension of this technology. The system has also been adapted to visualize glycoconjugates by light and electron microscopy (Sata et al., 1990).

The Lectin Link Kit (Genzyme) uses a biotinylated lectin and an alkaline phosphatase-avidin conjugate to visualize glycoproteins on the blot. The basic ABC (avidin-biotin conjugate) technology is well established in the literature for many assays. This kit as presented is primarily intended to identify and differentiate between various types of N-linked oligosaccharides. The conclusion that binding of any lectin is exclusively due to the presence of a particular structure on an N-linked oligosaccharide chain is valid only if all lectin binding is destroyed by PNGase F digestion (UNIT 17.13), because, except for Con A, no lectin is specific for N-linked oligosaccharides. It is important to note that not all N-linked oligosaccharides are sensitive to PNGase F (UNIT 17.13); resistance to PNGase F does not prove that a glycan is, by elimination, O-linked.

Anticipated Results

Even though lectins are quite specific, results are more qualitative than quantitative. The results obtained using lectins should be regarded as guides and preliminary indications of likely structures. Lectin-binding studies, especially when combined with a battery of enzymatic digestions, are useful experiments for preliminary characterization of glycoprotein structure because they are quick, reasonably

priced, and use only a few micrograms (or less) of purified sample material. However, they cannot substitute for the more precise structural and chemical analyses presented in other units in Chapter 17.

Inclusion of a positive control provides a qualitative and rough quantitative estimate for the amount of reactive material in the sample relative to the standards. This may be especially important in assessing the effects and completeness of exo- or endoglycosidase digestions. The problem with attempting accurate quantitation is that the intensity of staining will depend upon the binding affinity, the amount of protein, and the number of accessible glycans on the protein. All of these factors are important, and a high number of low-affinity binding sites present on a large quantity of protein can give an intensity on a blot similar to that of a single high-affinity binding site on a smaller amount of protein.

Two examples illustrate useful approaches for maximizing the information from lectin binding. In both, lectin-binding results first suggest the presence of certain types of carbohydrate chains on the protein. This information is used to design glycosidase digestions (UNIT 17.13) that should alter the binding of some or all of the lectins. Digested glycoproteins are probed with the lectins to confirm the predictions. All of the data are combined into a composite picture of the sugar chains present. To be conclusive, results of lectin binding and glycosidase digestions should be internally consistent. Because some of the glycosidases are themselves glycoproteins, controls should include a lane with glycosidase alone to identify bands due to the presence of glycosidase in the digested sample.

Example 1

A hypothetical purified glycoprotein reacts with wheat germ agglutinin (WGA) and *Sambucus nigra* agglutinin (SNA; see Table 17.7.2). WGA binds to terminal GlcNAc residues or sialic acids. Terminal GlcNAc may occur in complex N-linked oligosaccharides, such as those in ovomucoid, or as multiple residues of O- β -GlcNAc or GlcNAc-1-P directly bound to protein. Sialic acid could be present in α 2-3 or α 2-6 linkages to underlying β Gal residues on N- or O-linked chains. SNA binds to sialic acid linked α 2-6 to β Gal. One way to distinguish among the possibilities is to treat the sample with sialidase. Loss of WGA binding following sialidase treatment indicates that WGA binding is due to sialic acid and not terminal GlcNAc.

If both WGA binding and SNA binding are lost with PNGase F digestion, all of the WGA and SNA binding are due to the presence of sialic acids linked α 2-6 to β Gal residues on N-linked oligosaccharides.

Example 2

Experimental results from lectin and glycosidase digestion analysis of a hypothetical pu-

rified glycoprotein, glycoprotein, are shown in Figure 17.7.1. Undigested glycoprotein (panel A) produces a diffuse band when incubated with *Datura stramonium* agglutinin (DSA), *Ricinus communis* agglutinin (RCA I), and *Maackia amurensis* agglutinin (MAA) and a less diffuse band with SNA, but no band when incubated with Con A. Thus, all the lectins but Con A bind to glycoprotein. These data sug-

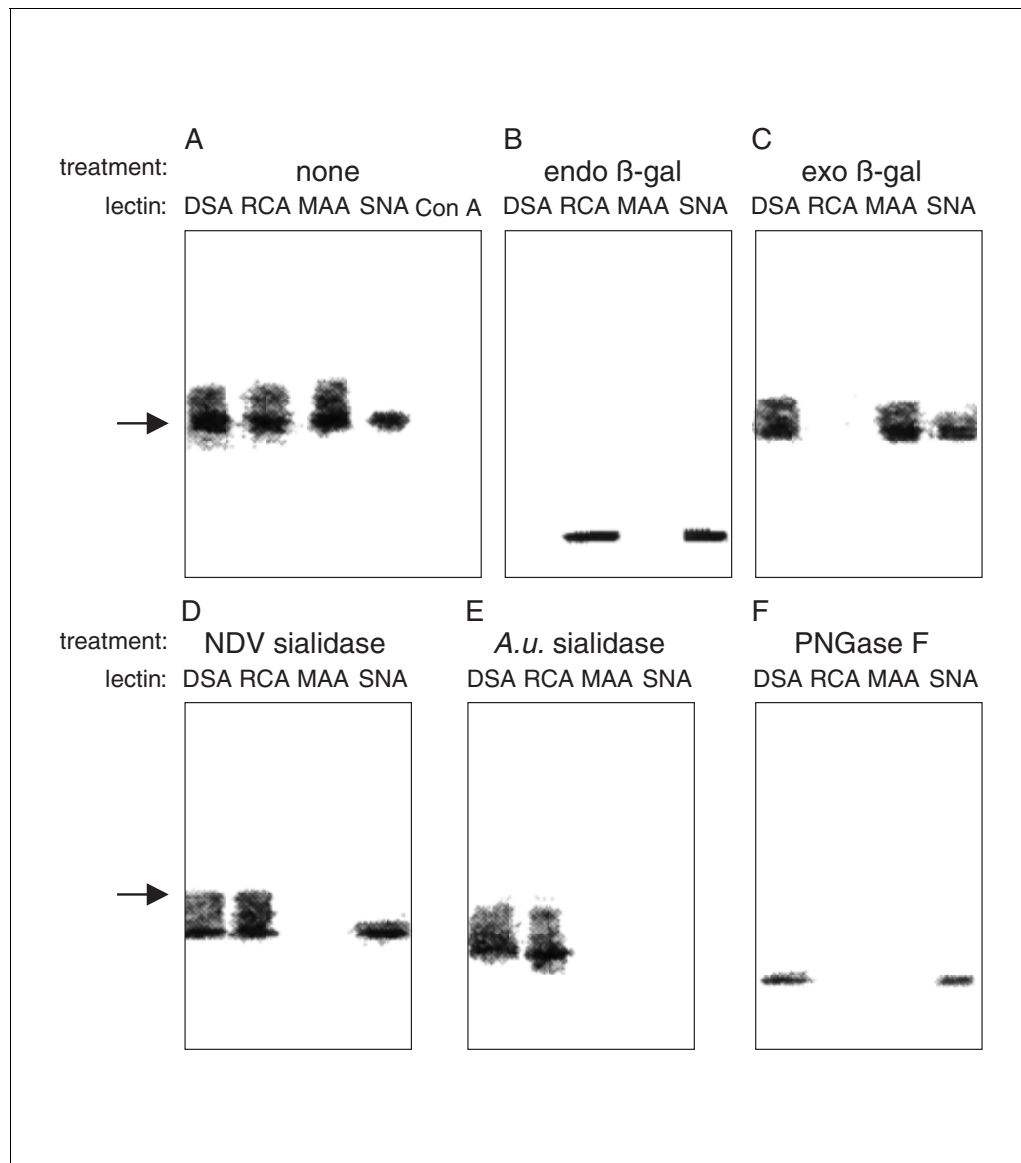


Figure 17.7.1 Example: Lectin blot analysis of a hypothetical purified glycoprotein, glycoprotein, using one of the procedures or kits described in this unit. All enzymatic digestions were conducted prior to lectin binding, and each lectin was previously tested on a positive control protein to check the reagents, developing conditions, etc. All samples of the test protein were run on SDS-PAGE, blotted onto nitrocellulose filters, blocked, and probed with the indicated lectins after the enzymatic treatment. Control lanes with glycosidase only control are not shown, but would normally be run. **(A)** Undigested glycoprotein. **(B)** Glycoprotein digested with endo- β -galactosidase. **(C)** Glycoprotein digested with exo- β -galactosidase. **(D)** Glycoprotein digested with Newcastle Disease Virus (NDV) sialidase. **(E)** Glycoprotein digested with the broad-spectrum sialidase from *Arthrobacter ureafaciens* (A.U.). **(F)** Glycoprotein digested with PNGase F. Arrow indicates the densest portion of the band for undigested glycoprotein.

gest that the glycoprotein may contain poly-lactosamine chains (DSA), terminal β -Gal residues (RCA I), and both α 2-3- and α 2-6-linked sialic acids (MAA and SNA, respectively). High-mannose type chains are probably absent (Con A).

This is useful information, but performing a few enzyme digestions and then probing with all of the lectins will provide much more. First, specific enzyme digestions should be done to show that each of the lectins is, in fact, binding to the expected sugar structure. Digestion with endo- β -galactosidase (panel B), which removes poly-lactosamine chains (UNIT 17.13), abolishes all bands in the DSA lane, showing that binding to DSA was due entirely to poly-lactosamine chains. The apparent M_r of the band has dropped and its boundaries have sharpened, showing that the poly-lactosamines contributed to the diffuse appearance of the original band. Glycoproteins often produce diffuse bands in gels and on blots because of microheterogeneity in the carbohydrate chains (UNIT 17.1). This digestion also destroys all of the MAA binding, suggesting that the poly-lactosamine chains, and only those chains, were terminated by α 2-3-linked sialic acids (UNIT 17.12). Bands are still observed in the RCA I lane, showing that terminal β -Gal residues were not digested, and in the SNA lane, showing that the poly-lactosamine chains do not terminate in α 2-6 sialic acids. If the digestion is done with exo- β -galactosidase (panel C), which removes terminal β -Gal residues (UNIT 17.18), it destroys the band in the RCA I lane but has minimal effects on the position and appearance of the other bands. This shows that the lectin was bound to the predicted terminal sugar residues. There are probably only a small number of these and they did not contribute much to the heterogeneity of the protein. Digestion with Newcastle Disease Virus (NDV) sialidase (panel D), which specifically removes α 2-3-linked sialic acids, abolished the binding of MAA, showing that MAA bound the α 2-3 sialic acid, as predicted. The significant reduction in the apparent size of the protein size shows that the sialic acids make a large contribution to the apparent size. Digestion with *Arthrobacter ureafaciens* sialidase (panel E), a broad-spectrum sialidase (UNIT 17.12), abolishes the bands in both the MAA and SNA lanes. Diffuse bands are still seen in the DSA and RCA I lanes; the apparent protein size is slightly smaller than in panel D, showing that the α 2-6-linked sialic acids make a relatively small contribution to the apparent size compared to those linked α 2-3. For each digestion

the bands are reduced in M_r and/or become sharper compared to those in the undigested glyco-biologin.

The final digestion (panel F) is with peptide N:glycosidase F (PNGase F), which removes essentially all N-linked chains; however, some are resistant (UNIT 17.13). Any lectin binding that is completely destroyed by this digestion is due only to components found on N-linked chains. PNGase F digestion of glyco-biologin completely abolishes the bands in the RCA I and MAA lanes but only partially reduces the intensity of the bands in the DSA and SNA lanes. The lack of binding to RCA I and MAA indicates that all of the β -Gal-terminated chains and those with α 2-3 sialic acids are exclusively N-linked (RCA I and MAA). Reduced binding to SNA indicates that some of the N-linked chains also contain α 2-6 sialic acid, which may also be found on O-GalNAc-linked chains. Reduced binding to DSA suggests either that the protein also contains poly-lactosamine units on both N- and O-linked chains or possibly that the PNGase F digestion was incomplete. The band is still diffuse because of residual poly-lactosamine chains. The apparent protein size is considerably smaller, showing that most of the heterogeneity was due to poly-lactosamine chains on N-linked oligosaccharides.

It is important to stress again that lectin blots give only preliminary evidence for the presence of the suggested structures. The results should be viewed as useful guides.

Time Considerations

Most enzymatic digestions are carried out overnight. If the proteins are separated on minigels, the entire analysis including blotting and development can be done in a single day. The procedure can be interrupted following protein transfer to nitrocellulose. Incubation in blocking buffer can be extended to overnight at 4°C.

Literature Cited

Detailed references regarding detection of the various types of conjugates are given in UNIT 10.8. Kit manufacturers will supply relevant references upon request. Much of their information may have been obtained exclusively from data produced in company laboratories, although some has also appeared in the peer-reviewed literature.

Glass, W.F., Briggs, R.C., and Hnilica, L.S. 1981. Use of lectins for detection of electrophoretically separated glycoproteins transferred onto nitrocellulose. *Anal. Biochem.* 115:219-224.

- Sata, T., Zuber, C., and Roth, J. 1990. Lectin-digoxigenin conjugates: A new hapten system for glycoconjugate cytochemistry. *Histochem.* 94:1-11.
- West, C.M. and McMahon, D. 1977. Identification of concanavalin A receptors and galactose-binding proteins in purified plasma membranes of *Dictyostelium discoideum*. *J. Cell Biol.* 74:264-273.

Key References

- Haselbeck, A., Schickaneder, E., von der Eltz, H., and Hösel, W. 1990. Structural characterization of glycoprotein carbohydrate chains by using digoxigenin-labeled lectins on blots. *Anal. Biochem.* 191:25-30.

Description of techniques for digoxigenin-labeled lectins, showing uses for highly purified glycoproteins on blots.

- Tanner, M.J. and Anstee, D.J. 1976. A method for the direct demonstration of the lectin-binding components of the human erythrocyte membrane. *Biochem. J.* 153:265-270.

Good overview of procedure.

Contributed by Hudson H. Freeze
La Jolla Cancer Research Foundation
La Jolla, California

Detection of Glycophospholipid Anchors on Proteins

Many eukaryotic proteins are tethered to the plasma membrane by glycosyl phosphatidylinositol (GPI) membrane anchors. Figure 17.8.1 schematically depicts the minimal (core) structure common to all GPI protein anchors characterized to date. Detailed structural information is described in reviews by Cross (1990) and Ferguson (1991).

This unit provides a general approach for detecting GPI-anchored proteins. First, the detergent-partitioning behavior of a protein of interest is examined for characteristics of GPI-linked species. The partitioning of total cellular and isolated proteins with Triton X-114 is described in the first basic and alternate protocols, respectively. Precondensation of Triton X-114, necessary to remove hydrophilic contaminants before partitioning, is outlined in the first support protocol.

The protein may also be subjected to specific enzymatic or chemical cleavages to release the protein from its GPI anchor. Phospholipase cleavage is detailed in the second basic and alternate protocols, and chemical cleavage with nitrous acid is described in the third basic protocol.

If GPI-anchored proteins are radiolabeled with fatty acids, it facilitates the detection of the GPI protein products following the cleavage reactions. Separation of lipid moieties is described in the third support protocol and base hydrolysis of proteins is presented in the fourth basic protocol. Figure 17.8.2 is a flowchart depicting the relationships between the various protocols.

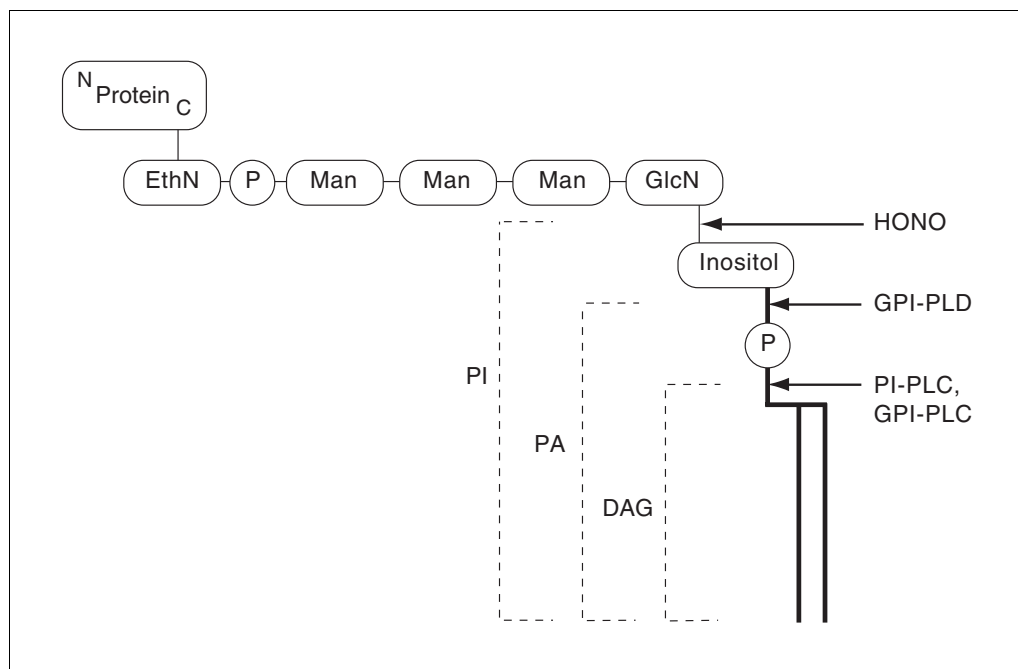


Figure 17.8.1 Schematic representation of the glycan core structure common to all GPI anchors. The sites of cleavage of phospholipase C enzymes (PI-PLC, GPI-PLC), phospholipase D (GPI-PLD), and nitrous acid (HONO) are indicated, as are lipid products resulting from these cleavages: PI (phosphatidylinositol), PA (phosphatidic acid), DAG (diacylglycerol). Other abbreviations: Man, mannose; GlcN, glucosamine; EthN, ethanolamine; P, phosphate group. Phospholipase C treatment generates inositol cyclic phosphate, the major epitope contributing to the cross-reacting determinant (CRD). Lipid products may vary from those depicted, depending on features of the anchor (see background information).

Contributed by Tamara L. Doering, Paul T. Englund, and Gerald W. Hart
Current Protocols in Molecular Biology (1993) 17.8.1-17.8.13
 Copyright © 2000 by John Wiley & Sons, Inc.

Preparation
and Analysis of
Glycoconjugates

17.8.1

Supplement 22

Strategic Planning

The techniques described below require that a method for identifying the protein of interest be available. Proteins may be visualized after one-dimensional gel electrophoresis (UNIT 10.2) and Coomassie blue or silver staining (UNIT 10.6), or they may be identified by immunoprecipitation (UNIT 10.16), or immunoblotting (UNIT 10.8) using available specific antibodies. Enzymes may also be detected by their known activities (e.g., see Kodukula et al., 1992).

It is extremely convenient, particularly for low-abundance species, to use proteins that are labeled with radioactive amino acids (UNIT 10.18) or GPI components (see Fig. 17.8.1). The ability to radiolabel proteins with GPI precursors such as fatty acids, ethanolamine, *myo*-inositol, and certain monosaccharide precursors (UNIT 17.4) is in itself suggestive of GPI anchorage. Because GPI-anchored proteins commonly represent a small fraction of eukaryotic glycoconjugates, metabolic radiolabeling may be inefficient. Tunicamycin (UNIT 17.10) may be used to increase relative incorporation of monosaccharide precursors into GPIs.

Rigorous identification of specific GPI anchor structures requires detailed analysis. These methods, which are beyond the scope of this unit, are described in Ferguson (1992).

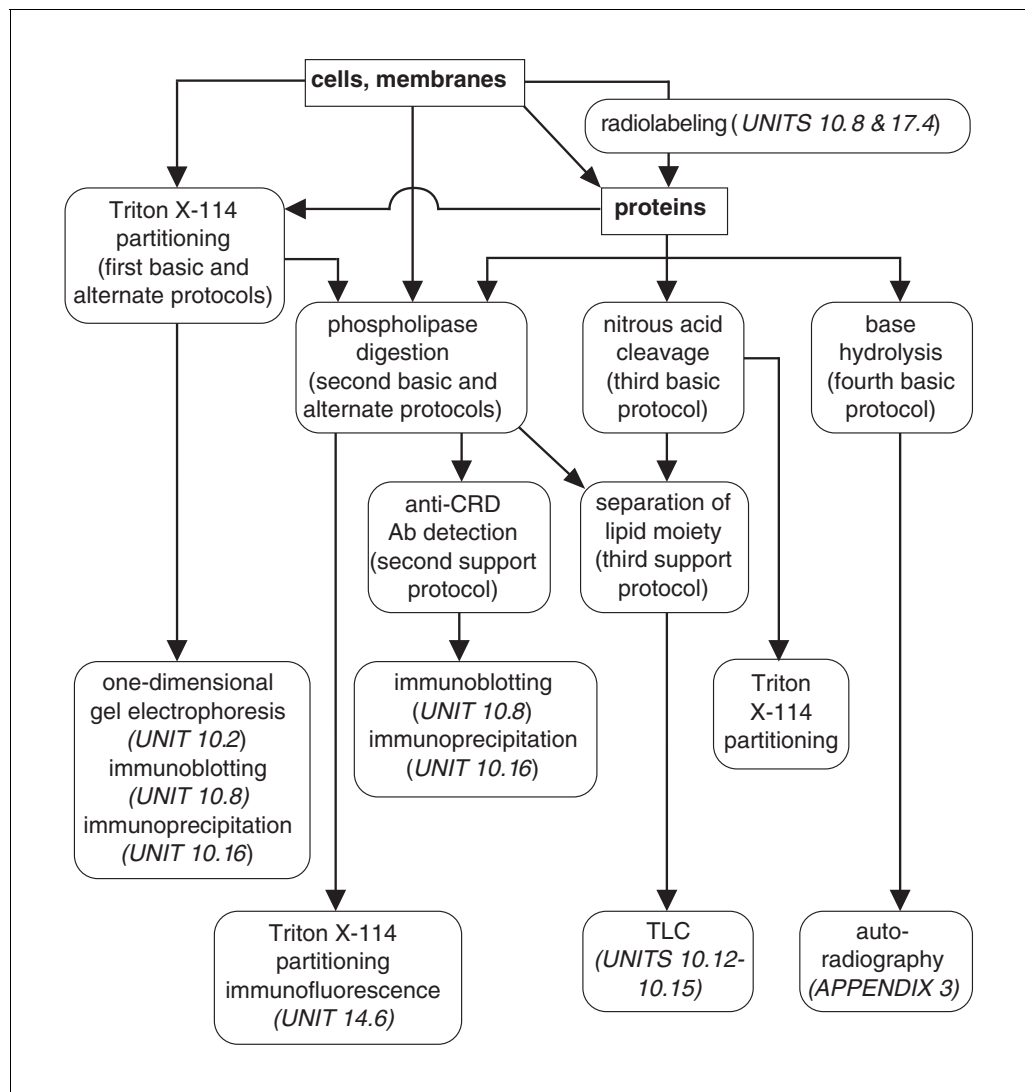


Figure 17.8.2 Flowchart for the detection of GPI-anchored proteins.

EXTRACTION AND PARTITIONING OF TOTAL PROTEINS FROM CELLS OR MEMBRANES WITH TRITON X-114

BASIC PROTOCOL

Extraction of cells or tissues with the detergent Triton X-114 provides a profile (and potentially an initial purification) of amphiphilic proteins, including integral membrane proteins and species bearing GPI anchors. Once a GPI-containing protein is released from the lipid component of its anchor, it will no longer partition into the detergent-enriched phase. This alteration of partitioning behavior provides a rapid assay for presence or absence of anchors. It may be used to monitor the changes in proteins produced by the other procedures.

In this protocol, cells or membrane fractions are first resuspended in TBS, then extracted with a Triton X-114 incubation on ice. The mixture is then warmed and separates into two protein-containing phases due to aggregation of detergent micelles (see background information). The upper (detergent-depleted) phase contains hydrophilic proteins; amphiphilic proteins (including those with GPI anchors) partition to the lower (detergent-enriched) phase. Both phases are then analyzed to clarify the group to which each protein belongs.

Materials

Cells, membrane fraction, or other source of protein
Tris-buffered saline (TBS), ice-cold
Precondensed Triton X-114 stock solution in TBS (support protocol), ice-cold
15-ml polypropylene centrifuge tubes
Centrifuges: low-speed (tabletop) and equipped with appropriate rotor (e.g., SS-34), at 4°C and room temperature

1. Resuspend cells (or membranes) in ice-cold TBS to a final protein concentration of ≤ 4 mg/ml in a 15-ml tube.

The amount of protein required is that which will give detectable levels at the end of the procedure and will depend on the system the individual researcher is using. As a guideline, use several times the amount of protein that is believed to be needed, to account for partitioning and losses (see critical parameters).

2. Add $\frac{1}{5}$ vol precondensed Triton X-114 stock solution ($\sim 2\%$ final concentration).
3. Extract cells by incubating 15 min on ice with occasional mixing.

At this point in the procedure there will only be one phase.

4. Centrifuge cell mixture 10 min at $10,000 \times g$ (9000 rpm in an SS-34 rotor), 4°C, and transfer supernatant to a fresh tube. Resuspend pellet in ice-cold TBS and save on ice for analysis in step 7.

The cold supernatant fraction contains both soluble and detergent-extracted proteins; the latter group includes proteins anchored by GPI structures or by transmembrane polypeptide domains. The pellet contains additional GPI species, as well as cell elements insoluble in nonionic detergents.

5. Warm supernatant fraction to 37°C in a water bath until the solution becomes cloudy.
6. Centrifuge the solution in a tabletop centrifuge 10 min at $1000 \times g$, room temperature. Collect upper and lower phases in separate tubes.

The upper phase (detergent-depleted) contains soluble proteins; the lower (detergent-enriched) phase contains proteins anchored by GPI structures or by transmembrane domains.

7. Analyze the resuspended pellet from step 4 and each phase from step 6 for the proteins of interest.

Preparation and Analysis of Glycoconjugates

17.8.3

As discussed in strategic planning, analysis may be by an activity assay, one-dimensional gel electrophoresis and staining of proteins (UNITS 10.2 & 10.6), or immunoprecipitation (UNIT 10.16) or immunoblotting (UNIT 10.8) using specific antibodies.

**ALTERNATE
PROTOCOL**

PARTITIONING OF ISOLATED PROTEINS WITH TRITON X-114

A modification to the basic protocol for Triton X-114 extraction of proteins from whole cells or membranes is used to partition isolated proteins.

1. Dilute or dissolve the protein in 1 ml ice-cold TBS in a 1.5-ml microcentrifuge tube. Add 0.2 ml Triton X-114 and mix.

See critical parameters for guidelines on the total amount of protein to use.

2. Warm protein mixture 15 min (or until it becomes cloudy) in a 37°C water bath. Microcentrifuge 5 min at maximum speed.
3. Collect the upper and lower phases and analyze each phase for the protein(s) of interest.

**SUPPORT
PROTOCOL**

PRECONDENSATION OF TRITON X-114 DETERGENT

Triton X-114 often contains hydrophilic contaminants, and must be precondensed by several rounds of phase separation before use.

Materials

Triton X-114 detergent
Tris-buffered saline (TBS)
50-ml centrifuge tubes
Tabletop centrifuge

1. Dissolve 1.5 g Triton X-114 in 50 ml TBS, place in a 50-ml centrifuge tube, then chill on ice.

At this point, the solution will be clear.

2. Warm the solution to 37°C in a water bath.

Following this warming, the solution will be turbid.

3. Centrifuge the solution in a tabletop centrifuge 10 min at 1000 × g, room temperature.
4. Remove and discard the upper (detergent-depleted) phase and redissolve the lower (detergent-enriched) phase in an equal volume of ice-cold TBS.

Hydrophilic contaminants partition into the upper phase.

5. Repeat partitioning (steps 2 to 4) three times.

The final detergent-enriched phase contains ~12% detergent. It may be stored at 4°C for routine use or frozen for long-term storage.

IDENTIFICATION OF GPI-ANCHORED PROTEINS BY PI-PLC DIGESTION OF INTACT CELLS

BASIC PROTOCOL

Phospholipase C (PLC) cleaves within the GPI membrane anchors (see Fig. 17.8.1) of glycoproteins, causing the release of protein from cell membranes. In this protocol, intact cells are digested with phosphatidylinositol-specific phospholipase C (PI-PLC). Following incubation with the enzyme, the cell mixture is centrifuged and both the resulting pellet and supernatant are analyzed for the presence of the protein of interest. Proteins released from GPI anchors are found in the supernatant; lack of release may indicate absence of a GPI anchor, modification of the GPI protein, or other factors (see background information).

Materials

- Cells (or membrane preparation)
- Bacterial phosphatidylinositol-specific phospholipase C (PI-PLC; e.g., from *Bacillus thuringiensis*, Oxford Glycosystems)
- Hanks balanced salt solution (HBSS; GIBCO/BRL), buffered saline, or culture medium
- Tabletop centrifuge and appropriate centrifuge tubes

1. Obtain cells (or membranes) of interest in dispersed suspension. Wash duplicate aliquots twice in HBSS, buffered saline, or culture medium, then resuspend in any of these solutions.

The final protein concentration of the samples (typically 0.1 to 0.5 mg/ml) should be based on supplier specifications for the enzyme preparation used. See critical parameters for guidelines on the total amount of protein to use.

2. Add bacterial PI-PLC to one aliquot. Add no enzyme to the second aliquot (control). Incubate tubes 1 hr at 37°C.

The amount of enzyme should be determined by supplier recommendations.

3. Centrifuge cells in a tabletop centrifuge 5 min at 1000 × g. Remove each supernatant to a fresh tube and save for analysis in step 4. Resuspend pellet in a volume of buffered saline equal to the volume of supernatant removed.

4. Assay supernatant and pellet fractions for protein(s) of interest.

The proteins may be assayed by Triton X-114 detergent partitioning (first basic protocol), one-dimensional gel electrophoresis (UNIT 10.2) followed by Coomassie blue or silver staining (UNIT 10.6), immunoprecipitation (UNIT 10.16) or immunoblotting (UNIT 10.8)

IDENTIFICATION OF GPI ANCHORAGE BY PHOSPHOLIPASE TREATMENT OF ISOLATED PROTEINS

ALTERNATE PROTOCOL

Several specific phospholipases that effectively degrade only GPI species are useful for analyzing GPI anchorage of isolated proteins (see background information). In this protocol, aliquots of the protein are digested in parallel with the following phospholipases: bacterial phosphatidylinositol-specific phospholipase C (PI-PLC), a GPI-specific phospholipase C (GPI-PLC) from *Trypanosoma brucei*, and a GPI-specific phospholipase D (GPI-PLD) from mammalian serum (rat, rabbit, or human). Proteins are partitioned by Triton X-114 extraction (first alternate protocol), then analyzed. Specific cleavage of proteins by phospholipase treatment strongly indicates GPI anchorage.

Preparation and Analysis of Glycoconjugates

17.8.5

Materials

Isolated protein

Acetone, -20°C

Appropriate enzyme buffer: GPI-PLD buffer, GPI-PLC buffer, and PI-PLC buffer

Phospholipase enzyme: GPI-PLD from rat, rabbit or human whole serum, GPI-PLC from *Trypanosoma brucei* (Oxford GlycoSystems), and PI-PLC from *Bacillus thuringiensis* (Oxford GlycoSystems) or *B. cereus* (Boehringer Mannheim or Sigma)

Tris-buffered saline (TBS)

Precondensed Triton X-114 solution (first support protocol)

Centrifuge and rotor (e.g., SS-34)

1. If the protein sample is lyophilized or concentrated in a buffer similar to the phospholipase buffers, begin at step 3. If it is not, acetone-precipitate the protein by adding 8 vol cold (-20°C) acetone to the protein solution and incubating it ≥ 3 hr at -20°C .

See critical parameters for guidelines on the total amount of protein to use. For low-abundance material, efficiency of precipitation may be increased by including 15 μg of a carrier protein (e.g., cytochrome c or bovine serum albumin).

2. Centrifuge the sample 20 min at $15,000 \times g$ ($\sim 11,000$ rpm in an SS-34 rotor), 4°C . Aspirate supernatant and allow pellet to air dry.
3. For each enzyme to be tested, dissolve two aliquots of the protein in 100 μl of the appropriate enzyme buffer in microcentrifuge tubes.
4. To one aliquot from each pair add the appropriate enzyme. Add an equal volume of the appropriate enzyme buffer to the second aliquot (control). Mix tubes and incubate 1 hr at 37°C .

The amount of enzyme added should be based on the activity of the lipase preparation (see supplier specifications) and the estimated amount of protein present. It may be desirable to titrate the enzyme, using various amounts in parallel reactions, to be sure that all sensitive proteins are completely digested.

GPI-PLD is not generally available, but whole serum (rat, rabbit or human) may serve as a source of this activity. Aliquots of serum may be stored frozen and thawed before use. For the assay, use 2 μl whole serum in a 100- μl reaction.

5. Dilute reaction with 200 μl ice-cold TBS, then add 60 μl Triton X-114. Chill 15 min on ice.
6. Partition the proteins as in steps 2 and 3 of the first alternate protocol.

GPI-anchored proteins partition into the detergent-enriched phase of Triton X-114 in the control reaction (uncleaved) and into the detergent-depleted phase after cleavage of the GPI anchor by phospholipase.

7. Analyze each phase for the presence of protein(s) of interest as in step 7 of the first basic protocol.

Alternatively, cleavage products can be detected by reactivity with α -CRD antibody (second support protocol) and, if the anchor is radiolabeled with fatty acids (UNIT 17.4), by separation of radiolabeled lipid products.

Migration of phospholipase-digested proteins on gel electrophoresis is not predictable and should not be used alone to assess cleavage. No shift, or a shift in either direction (relative to uncleaved protein migration) may be observed upon GPI anchor removal, depending on the properties of the protein.

DETECTION OF PRODUCTS AFTER PHOSPHOLIPASE TREATMENT BY REACTIVITY WITH ANTI-CRD ANTIBODY

SUPPORT
PROTOCOL

Cleavage of GPI structures with phospholipase C reveals a cryptic antigen termed the cross-reacting determinant (CRD). The major epitope contributing to the CRD is an inositol 1,2-(cyclic) monophosphate formed during phospholipase C cleavage; other structural features may also contribute to the antigenicity. Reactivity of a protein with anti-CRD antibody after solubilization by phospholipase C strongly indicates the presence of a GPI anchor. Anti-CRD antibody (Oxford GlycoSystems) may be used after phospholipase C treatment to immunoprecipitate (UNIT 10.16) or immunoblot (UNIT 10.8) proteins of interest. It should be noted that phospholipase D cleavage does not create the epitope required for α CRD activity (see Fig. 17.8.1).

NITROUS ACID CLEAVAGE OF GPI-ANCHORED PROTEINS

BASIC
PROTOCOL

An unusual property common to all GPI structures is the presence of a nonacetylated glucosamine that is glycosidically linked to inositol. The glycosidic bond is specifically cleaved when the glucosamine undergoes nitrous acid deamination, thereby cleaving the GPI anchor (see Fig. 17.8.1). Because nonacetylated glucosamine rarely occurs outside of GPIs, nitrous acid deamination may be used to identify GPI linkage to protein. Due to side reactions, deamination is not always quantitative or precisely reproducible, but it does serve a useful diagnostic purpose.

In this protocol, the protein is incubated in nitrous acid to cleave the GPI anchors. Cleaved and uncleaved proteins are then separated by Triton X-114 partitioning.

Materials

Protein(s)
0.1 M acetate buffer, pH 3.5
0.5 M NaNO₂, made fresh
0.5 M NaCl

1. Dry two aliquots of the protein of interest in microcentrifuge tubes and dissolve each in 0.1 ml of 0.1 M acetate buffer, pH 3.5.

See critical parameters for guidelines on the total amount of protein to use.

Nonidet P-40 (NP-40) detergent (0.1% v/v final) may be added to aid solubilization.

2. Add 0.1 ml of 0.5 M NaNO₂ to one tube. To the second tube (control), add 0.1 ml of 0.5 M NaCl. Incubate 3 hr at room temperature.
3. Detect the release of protein from GPI anchor by partitioning and analyzing the reaction products as in the first alternate protocol (Triton X-114 partitioning of isolated proteins).

If the anchor is radiolabeled with fatty acids (UNIT 17.4), the support protocol for separation of lipid moiety may be used.

Preparation
and Analysis of
Glycoconjugates

17.8.7

**SEPARATION OF LIPID MOIETY TO DETECT CLEAVAGE OF
GPI-ANCHORED PROTEINS**

If GPI-anchored proteins are radiolabeled with fatty acids, the efficacy of GPI anchor cleavage may be conveniently assessed by examining the released radiolabeled lipid products (see Fig. 17.8.1). In this protocol, radiolabeled proteins that have been cleaved by phospholipase or nitrous acid treatment are extracted with butanol. Released lipid moieties are extracted into the butanol phase and are analyzed and quantitated.

Materials

Radiolabeled protein, cleaved by phospholipase or nitrous acid treatment (second alternate and third basic protocols)

Water-saturated *n*-butanol

1. Add 100 μ l water to each of two 1.5-ml microcentrifuge tubes. To one tube, add 100 μ l cleaved radiolabeled protein. To the second, add 100 μ l control (uncleaved) protein.
2. Add 200 μ l water-saturated *n*-butanol to each tube and vortex vigorously.
3. Microcentrifuge 30 sec at maximum speed, room temperature, to separate the two phases.
4. Remove the upper (butanol) phase to a fresh tube, and set aside. Add another 200 μ l of water-saturated butanol to the original tubes.
5. Vortex and centrifuge as in steps 2 and 3 above.
6. Remove the second butanol phase and pool it with the first, retaining the lower (aqueous) phase for analysis.
7. Quantitate the percentage of lipid moieties released by the cleavage reactions by liquid scintillation counting of both butanol and aqueous phases. Compare counts recovered in butanol phases of the cleaved material with those recovered from control reactions to assess specific release.

To characterize released products fully, examine the butanol-extracted products by thin layer chromatography; compare to standards (Doering et al., 1990b). Products of phospholipase C digestion include diacylglycerol, alkyl-acyl glycerol, or ceramide, depending upon the original structure. Phosphatidylinositol (from nitrous acid treatment) or phosphatidic acid (released by phospholipase D) may also be analyzed and quantitated. This assay has been used for purification of phospholipase activity (Hereld et al., 1986).

BASE HYDROLYSIS OF RADIOLABELED PROTEINS

For proteins radiolabeled with fatty acids, it is useful to ascertain whether incorporated fatty acids are actually part of a GPI structure or are present in amide linkage (N-myristoylation) or thioester linkage (as in palmitoylation of cysteine residues). This can be characterized by assessing the lability of the fatty acid linkage to various chemical treatments.

In this protocol, aliquots of radiolabeled proteins are resolved by electrophoresis and the various lanes are treated with KOH in methanol or hydroxylamine. Methanol and Tris buffer are used to treat the control lanes. The lanes are then processed for fluorography. If the chemical bonds in the protein are susceptible to these treatments, a radiolabeled signal will no longer be produced at the protein band of interest.

Materials

Protein radiolabeled with fatty acid
0.2 M KOH in methanol
Methanol
1 M hydroxylamine-HCl, pH 7.5, made fresh
1 M Tris·Cl, pH 7.5 (APPENDIX 2)

Additional reagents and equipment for one-dimensional gel electrophoresis (UNIT 10.2), staining of gels (UNIT 10.6), and autoradiography (APPENDIX 3)

1. Resolve five identical samples of protein radiolabeled with fatty acid by SDS-PAGE electrophoresis. Include appropriate markers.
2. Cut apart the five gel lanes, and stain one to visualize protein and markers.
3. Soak the other four lanes in one of the following solutions for 1 hr at room temperature:
 - Lane 1: 0.2 M KOH in methanol
 - Lane 2: methanol alone
 - Lane 3: 1 M hydroxylamine-HCl, pH 7.5
 - Lane 4: 1 M Tris·Cl, pH 7.5
4. Soak the treated gel slices in three washes (10 min each) of water.
5. Process gel pieces for autoradiography.

If bonds are susceptible to the treatments above, radiolabel from fatty acids will no longer produce a signal at the position of the protein band of interest. KOH in methanol cleaves thio- and oxyesters but not amide linkages. Hydroxylamine treatment cleaves thioesters only. Fatty acids that are part of a GPI moiety will be released by KOH in methanol, but not by hydroxylamine treatment. The methanol and Tris·Cl treatments serve as controls. Ceramide-containing anchors are not sensitive to mild base treatment (see anticipated results and Conzelmann et al., 1992).

REAGENTS AND SOLUTIONS

0.1 M acetate buffer, pH 3.5

Dissolve 0.82 g sodium acetate in 95 ml distilled water. Adjust pH to 3.5 with acetic acid, then bring volume to 100 ml with distilled water. Store indefinitely at room temperature.

GPI-PLC buffer

50 mM Tris·Cl, pH 8.0
5 mM EDTA
1% NP-40
Store indefinitely at 4°C

GPI-PLD buffer

50 mM Tris·Cl, pH 7.4
10 mM NaCl
2.5 mM CaCl₂
0.1 % NP-40
Store indefinitely at 4°C

1 M hydroxylamine-HCl, pH 7.5

Dissolve 6.95 g hydroxylamine-HCl in 95 ml distilled water. Adjust pH to 7.5 with NaOH, then increase volume to 100 ml with water. Make fresh for each use.

PI-PLC buffer (for *B. thuringiensis* enzyme)

25 mM Tris acetate, pH 7.4
0.1% (w/v) sodium deoxycholate (Na-DOC)
Store indefinitely at 4°C

Tris-buffered saline (TBS)

10 mM Tris-Cl, pH 7.5
150 mM NaCl
Store indefinitely at room temperature

Water-saturated n-butanol

Mix equal volumes of distilled water and *n*-butanol. Cap tightly and shake vigorously. Let stand to allow phases to separate; the upper phase is the butanol. Store indefinitely at room temperature.

COMMENTARY

Background Information

Glycosyl phosphatidylinositol (GPI) membrane anchors are one mode of anchoring proteins to cell surfaces. These structures consist of a glycan bridge between phosphatidylinositol and phosphoethanolamine; the phosphoethanolamine is in amide linkage to the C-terminus of the protein (see Fig. 17.8.1). The glycan core of GPIs (shown schematically in Fig. 17.8.1) is remarkably conserved throughout evolution, although it is modified in many cell types by the addition of side chains. The lipid moieties of GPIs, however, are quite heterogeneous. This portion may consist of diacyl glycerol, alkyl acyl glycerol, *lyso*-acyl compounds, or ceramide. The hydrocarbon chains also vary in length and degree of saturation, and may be present as mixtures of species. GPI structures are reviewed in Ferguson (1991).

Much of our understanding of GPI anchors is derived from studies of the variant surface glycoprotein (VSG) of trypanosomes. Early work showed that VSG is associated with the membrane via a lipid moiety that contains dimyristoyl phosphatidylinositol (Cardoso de Almeida and Turner, 1983; Ferguson and Cross, 1984; Ferguson et al., 1985, 1986). Other investigations had shown that bacterial phospholipase C releases certain proteins from cell membranes (reviewed by Low, 1989). Together, these studies led to the definition of a class of proteins bearing a C-terminal glycolipid anchor.

Further investigations in this field have led to the elucidation of the pathway of GPI biosynthesis (reviewed by Doering et al., 1990a; Field and Menon, 1991), the purification of a GPI-specific phospholipase C (GPI-PLC), and the development of anti-cross-reacting determinant (anti-CRD) antibodies as probes for

GPIs. Phosphatidylinositol-specific phospholipase C (PI-PLC) and GPI-specific phospholipase D (GPI-PLD) are also useful in identification of GPI-anchored proteins.

The protocols included here are designed to allow identification of GPI-anchored proteins without detailed structural analysis. Because these techniques are rather diverse, further explanatory information is outlined below.

Triton X-114 partitioning. Triton X-114 forms a clear micellar solution at low temperatures; the solution separates into two phases when warmed above 20°C, due to aggregation of detergent micelles (Bordier, 1981). When cellular material is extracted in Triton X-114 at low temperatures, the solution contains soluble proteins, integral membrane proteins, and GPI-anchored proteins. Centrifugation results in a pellet containing some of the GPI-anchored species as well as cellular components that are insoluble in nonionic detergents (Hooper and Bashir, 1991). When the detergent solution is warmed, amphiphilic proteins (including those with GPI anchors) associate with the detergent-enriched phase, while hydrophilic ones partition to the detergent-depleted phase. Some GPI-anchored proteins exhibit detergent insolubility that develops shortly after synthesis (Brown and Rose, 1992); this is thought to be due to aggregation of the anchored proteins with glycosphingolipids, possibly as a step in directed protein transport.

Once a GPI-containing protein is released from the lipid component of its anchor, it will no longer partition into the detergent-enriched phase. This alteration of partitioning behavior provides a rapid assay for the presence or absence of anchors. It may be used to monitor the changes in proteins produced by the other procedures.

Phospholipase digestion. Bacterial phosphatidylinositol-specific phospholipase C (PI-PLC) efficiently cleaves GPI anchors, releasing proteins in aqueous soluble forms. This enzyme also cleaves phosphatidylinositol (PI), *lyso*-PI, and PI-linked glycans. Proteins specifically released from membranes by PI-PLC have all been found to be GPI-anchored, making this an extremely reliable mode of analysis (for review see Ikezawa, 1991).

Several other specific phospholipases that effectively degrade only GPI species are also useful for analysis. These enzymes are generally more expensive than PI-PLC. One of these, initially purified from *Trypanosoma brucei*, is a GPI-specific phospholipase C (GPI-PLC). Another, present in mammalian serum (rat, rabbit, or human), is a GPI-specific phospholipase D (GPI-PLD). These enzymes are convenient tools for identifying the presence of a GPI anchor.

Release of a protein from its GPI anchor is followed by using appropriate protein detection methods—e.g., enzyme activity assays, electrophoretic separation (UNIT 10.2) followed by protein staining (UNIT 10.6), and, if specific antibodies are available, immunoprecipitation (UNIT 10.16) or immunoblotting (UNIT 10.8). Alternatively, treated cells may be studied for loss of a specific protein using immunofluorescence detection methods (UNIT 14.6), including fluorescence-activated cell sorting. Also, because release by PLC reveals a specific epitope (termed the cross-reacting determinant, or CRD), antibodies directed against this epitope may be employed to identify cleavage products resulting from PLC treatment (second support protocol).

Lack of release may indicate absence of a GPI anchor, although some caution is required in interpreting negative results. Certain GPI molecules, bearing an additional fatty acid on inositol, are not susceptible to cleavage by either bacterial PI-PLC or trypanosomal GPI-PLC. Some proteins are partly sensitive, suggesting that only a fraction of these polypeptides bear anchors with acyl-inositol. Serum phospholipase D (GPI-PLD) cleaves GPI structures that contain acylated inositols, and is useful in their identification. However, it is not effective on membrane-bound proteins in the absence of detergent.

When intact cells are treated with phospholipase, partial or total resistance to cleavage can also occur due to inaccessibility of the protein to the enzyme, expression of a protein on the cell surface in both GPI-anchored and trans-

membrane-anchored forms, or tight association of a protein with a nonsusceptible protein on the cell surface (reviewed by Rosenberry, 1991).

Critical Parameters

Some details of these procedures will be affected by the method used to detect each protein under study. For example, specific quantities of starting materials are not provided because detection methods vary widely in terms of sensitivity. Also, for activity assays, the conditions of each procedure and whether the activity would survive such treatment must be considered. In general, plan to use two to ten times the amount of protein that would be reasonable for the mode of detection employed, depending upon the availability of material. This should allow for potential losses sustained during the procedure, as well as for the fact that proteins may be recovered in several fractions.

Triton X-114 partitioning provides a first hint of GPI anchorage, although transmembrane domains are also concentrated in the detergent phase of this partition. For example, if a protein is known from sequence data to have no membrane-spanning domain, its detection in the detergent phase of a Triton X-114 partition would be suggestive of lipid modification. In general, however, this method is most useful as a means of assessing release of a protein from its anchor, as shown by an alteration of partitioning behavior. Many of the other methods described specifically release proteins from their anchors to demonstrate GPI-linkage (e.g., nitrous acid deamination or phospholipase digestion); Triton X-114 is a convenient tool for assessing such release. Triton X-114 often contains hydrophilic contaminants, and should be precondensed by several rounds of phase separation before use (Bordier, 1981).

Phospholipase digestion is useful for determining GPI linkage, and bacterial PI-PLC is perhaps the best enzyme to try initially for reasons of cost and efficacy. Although GPI-PLC is more specific, it is generally more expensive.

Some GPI anchors are not susceptible to cleavage by phospholipase C. These structures have an additional fatty acid covalently attached to the inositol of the anchor (Rosenberry, 1991). For this reason, failure of cleavage by PLC does not necessarily imply that a protein is not GPI-anchored. In such cases GPI-PLD may be useful, although this enzyme does not expose the epitope for anti-CRD detection. Ex-

traction and analysis of radiolabeled lipid products from phospholipase digestions provides rapid assessment of cleavage. As mentioned earlier, radiolabeling the protein of interest is often helpful and permits less material and reagents to be used. In particular, if a protein may be radiolabeled with fatty acids, base hydrolysis is a helpful analytical tool, especially if lipid modification is suspected but other methods have yielded negative results.

Anticipated Results

Proteins anchored by GPI structures without acylated inositols should require detergent for solubilization. They may be extracted by Triton X-114 or remain in the cell pellet after such extraction (depending on solubility properties of the protein). In a warmed Triton X-114 solution, anchored protein(s) partition into the detergent-enriched phase. These proteins can be chemically cleaved by nitrous acid deamination, and enzymatically hydrolyzed by treatment with GPI-PLC, PI-PLC, or GPI-PLD. The protein cleaved with PLC will react with α -CRD antibody, and appropriate lipid products will be generated from all cleavage reactions (Fig. 17.8.1). Radiolabel incorporated into the anchor as fatty acid is released by treatment with KOH in methanol, but not by hydroxylamine treatment, if the fatty acid is hydroxy-esterified to glycerol (as in diacyl or alkyl-acyl anchors). Some PI-PLC and nitrous acid-sensitive anchors have been found to contain ceramide instead of diacyl or alkyl-acyl glycerol (e.g., in yeast). These anchors are resistant to mild base treatment, but are sensitive to conditions that hydrolyze amides (Conzelmann et al., 1992).

If the GPI anchor includes an acylated inositol, the above characteristics will hold true except that PLC treatment will not cleave the anchor. In these cases, nitrous acid deamination, GPI-PLD digestion, and radiolabel studies are used to characterize the linkage.

Time Considerations

In general, the methods described above are fairly rapid, with enzyme incubations or chemical reactions and subsequent partitioning requiring 1 to 2 hr. The time required for completion of these analyses depends primarily on the method used to detect the protein(s) of interest. For example, if a protein is radiolabeled in the fatty acid portion of the anchor, cleavage may

be assessed by lipid extraction and scintillation counting in ≤ 1 hr. If immunoprecipitation or immunoblotting is used for product detection, analysis will be more lengthy.

Literature Cited

- Bordier, C. 1981. Phase separation of integral membrane proteins in Triton X-114 solution. *J. Biol. Chem.* 256:1604-1607.
- Brown, D.A. and Rose, J.K. 1992. Sorting of GPI-anchored proteins to glycolipid-enriched membrane subdomains during transport to the apical surface. *Cell* 68:533-544.
- Cardoso de Almeida, M.L. and Turner, M.J. 1983. The membrane form of variant surface glycoproteins of *Trypanosoma brucei*. *Nature* 302:349-352.
- Conzelmann, A., Puoti, A., Lester, R.L., and Desponds, C. 1992. Two different types of lipid moieties are present in glycoposphoinositol-anchored membrane proteins of *Saccharomyces cerevisiae*. *EMBO J.* 11:457-466.
- Cross, G.A.M. 1990. Glycolipid anchoring of plasma membrane proteins. *Annu. Rev. Cell Biol.* 6:1-39.
- Doering, T.L., Masterson, W.J., Hart, G.W., and Englund, P.T. 1990a. Biosynthesis of glycosyl phosphatidylinositol membrane anchors. *J. Biol. Chem.* 265:611-614.
- Doering, T.L., Raper, J., Buxbaum, L.U., Hart, G.W., and Englund, P.T. 1990b. Biosynthesis of glycosyl phosphatidylinositol protein anchors. *Methods* 1:288-296.
- Ferguson, M.A.J. and Cross, G.A.M. 1984. Myristylation of the membrane form of *Trypanosoma brucei* variant surface glycoprotein. *J. Biol. Chem.* 259:3011-3015.
- Ferguson, M.A.J., Haldar, K., and Cross, G.A.M. 1985. *Trypanosoma brucei* variant surface glycoprotein has a sn-1,2-dimyristoyl glycerol membrane anchor at its COOH-terminus. *J. Biol. Chem.* 260:4963-4968.
- Ferguson, M.A.J., Duszenko, M., Lamont, G.S., Overath, P., and Cross, G.A.M. 1986. Biosynthesis of *Trypanosoma brucei* variant surface glycoprotein: N-glycosylation and addition of a phosphatidylinositol membrane anchor. *J. Biol. Chem.* 261:356-362.
- Ferguson, M.A.J. 1991. Lipid anchors on membrane proteins. *Current Opinion in Structural Biology.* 1:522-529.
- Ferguson, M.A.J. 1992. The chemical and enzymic analysis of GPI fine structure. In *Lipid Modifications of Proteins: A Practical Approach*. (A.J. Turner and N. Hooper, eds.) pp. 191-230, IRL Press, Oxford.
- Field, M.C. and Menon, A.K. 1991. Biosynthesis of glycolipid anchors in *Trypanosoma brucei*. *Trends Genet.* 3:107-115.

- Hereld, D., Krakow, J.L., Bangs, J.D., Hart, G.W., and Englund, P.T. 1986. A phospholipase C from *Trypanosoma brucei* which selectively cleaves the glycolipid on the variant surface glycoprotein. *J. Biol. Chem.* 261:13813-13819.
- Hooper, N.M. and Bashir, A. 1991. Glycosyl-phosphatidylinositol-anchored membrane proteins can be distinguished from transmembrane polypeptide-anchored proteins by differential solubilization and temperature-induced phase separation in Triton X-114. *Biochem. J.* 280:745-751.
- Ikezawa, H. 1991. Bacterial PIPLCs—unique properties and usefulness in studies on GPI anchors. *Cell Biol. Int. Rep.* 15:1115-1131.
- Kodukula, K., Gerber, L.D., Amthauer, R., Brink, L., and Udenfriend, S. 1992. Biosynthesis of glycosylphosphatidylinositol (GPI)-anchored membrane proteins in intact cells: Specific amino acid requirements adjacent to the site of cleavage and GPI attachment. *J. Cell Biol.* 120:657-664.
- Low, M.G. 1989. The glycosyl-phosphatidylinositol anchor of membrane proteins. *Biochim. Biophys. Acta* 988:427-454.
- Rosenberry, T.L. 1991. A chemical modification that makes glycoinositol phospholipids resistant to phospholipase C cleavage: Fatty acid acylation of inositol. *Cell Biol. Int. Rep.* 15:1133-1149.

Key References

- Ferguson, M.A.J., Homans, S.W., Dwek, R.A., and Rademacher, T.W. 1988. Glycosyl-phosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane. *Science* 239:753-759.

This landmark paper first described the core glycan of GPI anchors, providing a scheme for their structural analysis.

- Post-translational Modifications of Proteins by Lipids: A Laboratory Manual. 1988. (U. Brodbeck and C. Bordier, eds.) Springer-Verlag, N.Y.

A useful laboratory manual of methods dealing with various post-translational modifications of proteins by lipids; many methods relate to GPI-linked proteins.

- Methods: A Companion to Methods in Enzymology, Vol. 1, Number 3. (P. Casey, ed.) Academic Press, San Diego, Calif.

This issue of Methods, titled 'Covalent modification of proteins by lipids,' contains two chapters specifically about GPIs—Doering et al., 1990b (see above); Mayor and Menon (pp. 297-305)—as well as related topics.

Contributed by Tamara L. Doering
University of California
Berkeley, California

Paul T. Englund and Gerald W. Hart
Johns Hopkins Medical School
Baltimore, Maryland

Direct Chemical Analysis of Glycoconjugates for Carbohydrates

The following unit presents protocols for detection of different sugars bound to glycoconjugates. The solution containing the carbohydrate material is treated with a specific reagent, generating a colored reaction product that can be detected spectrophotometrically. For each assay, the absorbance of the colored solution is proportional to the amount of sugar present in the glycoconjugate.

BASIC PROTOCOL 1

PHENOL–SULFURIC ACID ASSAY FOR HEXOSES AND PENTOSES

This protocol describes the use of the phenol–sulfuric acid colorimetry to quantitatively determine the total level of hexoses or pentoses and their derivatives in an experimental sample. Monosaccharide samples of known concentration are used to construct a standard curve for comparison. The assay is optimized for use with small samples.

NOTE: Distilled or deionized water and high-quality sulfuric acid must be used for this procedure, as contamination of these reagents may lead to a false-positive result. It is advisable to check each new bottle of sulfuric acid and the water being used as described in steps 2 to 8 before using them for the assay. Both reagents should show an absorbance <0.05 (measured against water that has not been submitted to the assay).

Materials

- Sample to be assayed
- Monosaccharide standard: e.g., D-(+)-galactose (Sigma, purified-grade anhydrous crystals; Critical Parameters)
- 5% (w/v) phenol (see recipe)
- Concentrated sulfuric acid (ACS reagent-grade)
- Thick-walled, 16 × 125–mm Pyrex test tubes, either brand new and washed with distilled water, or acid washed
- Automatic glass dispenser *or* glass pipet with a portion of the tip removed for rapid flow, suitable for use with concentrated sulfuric acid
- 1.0-ml glass cuvettes (1-cm path length)
- Visible-light spectrophotometer

1. Dissolve sample in water or buffer at an appropriate concentration to 300 or 450 μ l final volume (for duplicate or triplicate samples, respectively).

The sample is dissolved in a new or acid-washed tube.

The quantity of sample dissolved must be such that the sugar concentration is ≥ 10 ng/ml. If insufficient data are available about the sample to estimate its sugar concentration, it is advisable to assay two solutions having at least a ten-fold difference in concentration. Depending on the results, it may be necessary to increase or reduce the concentration to obtain absorbance in the linear range of the assay (0.1 to 0.8). Volumes can be varied according to the quantity of material available as long as the proportions of reagents described are maintained.

2. Transfer two 150- μ l aliquots of the dissolved sample (or three if possible) into new or acid-washed 16 × 125–mm Pyrex test tubes.

If the sample volume is scaled up or down, the test tube used must have a volume considerably higher than the final volume of reaction mixture to reduce the chance of spilling concentrated sulfuric acid. The use of wide-mouth tubes slows dissipation of the heat (produced by the exothermic reaction between sulfuric acid and water) that is required for color development.

3. Set up three blank tubes using 150 μ l sample solvent (water or buffer; see step 1) in each.
4. In triplicate, prepare a series of tubes containing 10, 20, 30, 40, 50, 60, 70, and 80 ng/ μ l of the monosaccharide standard in sample solvent. Adjust all volumes to 150 μ l with water.

To minimize errors, prepare a stock solution of the standard and make sequential dilutions. Keep stock solutions at -20°C .
5. Add 150 μ l of 5% (w/v) phenol solution to each tube and vortex briefly to mix.

CAUTION: *Rubber gloves and eye protection must be worn while performing steps 6 to 8, which involve handling of concentrated sulfuric acid and generation of heat.*
6. Rapidly add 750 μ l concentrated sulfuric acid to each tube, using an automatic glass dispenser or a glass pipet with a portion of the tip removed. Ensure that the acid stream hits the liquid surface directly to produce rapid mixing and even heat distribution. Let tubes stand 10 min at room temperature.
7. Vortex tubes again briefly to mix. Let stand 30 min at room temperature while the color develops.

The color is stable for an additional several hours.
8. Transfer the solution from each tube to a 1.0-ml (1-cm path length) glass cuvette. Using the spectrophotometer, read the absorbance of each tube at 480 nm (to detect pentoses, uronic acids, and their methylated derivatives) and/or 490 nm (to detect hexoses and their methylated derivatives).

Care must be taken when transferring the solutions to the cuvettes. Any spill inside the spectrophotometer should be cleaned up immediately, as the corrosive sulfuric acid will damage the instrument.
9. Average the absorbances of the triplicate tubes for each standard. Subtract the absorbance of the blank from each average. Prepare a standard curve of absorbance versus concentration in ng/ μ l for each absorbance (A_{480} or A_{490}) that was monitored.
10. Average the absorbance of the duplicate or triplicate sample tubes and subtract the absorbance of the blank. Determine the amount of sugar in the sample by reference to the standard curve.

Reagents and Solutions

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

Phenol, 5% (w/v)

Dissolve 5 g of 99% phenol crystals in water and bring the volume to 100 ml with water. Store solution in a dark bottle at room temperature (stable indefinitely).

CAUTION: *Use gloves and eye protection when weighing phenol, as it is very corrosive. Store the crystals in a dark bottle at 4°C to prevent oxidation.*

Commentary

Background Information

The phenol–sulfuric acid assay for detecting sugars has been widely used since it was first developed by Dubois et al. (1956). The method is simple, rapid, sensitive, accurate, specific for carbohydrates, and widely applicable. Most

sugars (except hexosamines; see **UNIT 17.11**), sugar derivatives, oligosaccharides, and polysaccharides react with these reagents. The assay is based on the fact that when carbohydrates are heated in the presence of strong acids, they are converted into furfural and its homologs by

oxidation, reduction, and condensation processes. These products can react with organic substances such as phenol to form colored compounds. The type and ratio of the reaction products varies from carbohydrate to carbohydrate and also depends on reagent concentrations, temperature, and time of heating. Although the exact composition of the colored products is indeterminate, the assay has been standardized and the results obtained for each type of carbohydrate are highly reproducible. The reagents are inexpensive, readily available, and stable.

Critical Parameters and Troubleshooting

The monosaccharide used as a standard for constructing the standard curve should be chosen to match the expected composition of the sample, as different individual monosaccharides will give different results. For glycoproteins of animal origin, galactose can be used as the standard. The assay will certainly indicate whether carbohydrates are present in a sample, and will give an idea of its percentage (quantitation). However, when assaying unknown complex carbohydrates, it is not possible to select the exact matching standard, and the values obtained will only be approximate.

Triplicate measurement minimizes errors. For accurate and reproducible results, it is important to use brand-new test tubes that have been rinsed with distilled water to eliminate any lint or paper fiber, or tubes that have been acid-cleaned and rinsed. Accidental contamination with lint may still occur, producing high readings, but the use of triplicate tubes makes it possible to identify and discard such false-positive tubes.

If the assay is used to check the elution profile of a carbohydrate-containing sample

eluted from a column, the buffers and column eluant must first be tested for positive reactions. With columns packed with carbohydrate-based polymers (e.g., Sephadex, Pharmacia), any breakdown of the packing material will cause an extremely intense color in the assay. Buffers used for elution may also contain sugar-bearing contaminants such as bacteria, in which case fresh lots should be prepared before attempting this kind of monitoring. The blank should contain buffer eluted from the column prior to the sample. If organic solvents are present, they must be removed by evaporation prior to performing the assay.

Anticipated Results

The phenol-sulfuric acid assay provides a quantitative measure of the total content of pentoses or hexoses and their derivatives in the sample analyzed, expressed as equivalents of the monosaccharide used as a standard (e.g., galactose). When carried out in triplicate as described, the assay yields results accurate to $\pm 2\%$.

Time Considerations

In general, the complete assay can be performed in <1 hr. However, if a considerable number of experimental samples are to be assayed (e.g., when monitoring effluent from a column), the time required may be longer.

Literature Cited

Dubois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A., and Smith, F. 1956. Colorimetric method for determination of sugars and related substances. *Anal. Chem.* 28:350-356.

Original description of this technique, containing the most complete study to date of the reactivity of different types of carbohydrates.

BASIC PROTOCOL 2

FERRIC ORCINOL ASSAY FOR SIALIC ACIDS

When sialic acids are oxidized in concentrated acid in the presence of orcinol, a blue-purple chromophore is formed. The reaction provides an estimation of the sialic acid content.

Materials

- Sample to be assayed
- Sialic acid standard: 1 mM *N*-acetylneuraminic acid (Neu5Ac; Boehringer Mannheim; store frozen at -20°C)
- Bial reagent (see recipe)
- Isoamyl alcohol
- Thick-walled 16 \times 125-mm Pyrex test tubes
- Heating block or boiling water bath

Glass marbles
1.0-ml glass cuvettes (1-cm path length)
Visible-light spectrophotometer

1. Dissolve sample in water to an appropriate concentration.

When no information regarding the sialic acid content of a sample is available, it is a good practice to assay two aliquots differing in concentration ~5- to 10-fold.

2. Transfer three 150- μ l aliquots to 16 \times 125-mm Pyrex test tubes. Adjust volume to 200 μ l with water.
3. In triplicate, prepare a series of tubes containing 5, 10, 20, 30, and 40 nmol Neu5Ac standard (5 to 40 μ l of 1 mM Neu5Ac). Adjust to 200 μ l water. Prepare a blank using 200 μ l of water.
4. Add 200 μ l Bial reagent, vortex, and cover tubes with glass marbles. Heat tubes 15 min at 100°C.
5. Cool tubes by immersing in room temperature tap water.
6. Add 1 ml isoamyl alcohol to each tube. Vortex tubes and let sit 5 min in an ice bath.
7. Centrifuge tubes 3 min in a tabletop centrifuge to separate the phases.
8. Transfer upper phase to 1.0-ml cuvettes with a Pasteur pipet. Measure the absorbance at 570 nm and subtract the absorbance of the blank.
9. Prepare standard curve of absorbance versus nmol Neu5Ac using the average of the triplicate standard solutions.
10. Determine the amount of sialic acid in each unknown sample by comparing to the standard curve. Quantity is expressed as nmol Neu5Ac.

Reagents and Solutions

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Bial reagent

0.2 g orcinol (5-methylresorcinol)
81.4 ml concentrated HCl
2 ml 1% (w/v) ferric chloride (analytical grade)
H₂O to 100 ml
Store \leq 1 week at 4°C

Commentary

Background Information

Glycosidically bound and free sialic acids form a blue-purple chromophore when heated in the presence of orcinol/Fe³⁺ (Bial reagent) in concentrated HCl. The chromophore is soluble in organic solvents. The assay is accurate and reproducible, and the minimum detection limit is ~5 nmol (1.5 μ g). Determination of free sialic acids is best achieved by the more sensitive 2'-thiobarbituric acid (TBA) assay (see UNIT 17.18). A more accurate determination can be carried out later, after release of the sugar with sialidase or mild acid and purification as de-

scribed in UNIT 17.18. *N*-acetylneuraminic acid (Neu5Ac) is used as a standard. However, it should be noted that the extinction coefficient is different for different sialic acid derivatives, such as *N*-glycolylneuraminic acid. Any *O*-acylated sialic acids will have the same extinction coefficients as their parent molecules because, under the strong acidic conditions, all ester groups are rapidly hydrolyzed. Even crude biological materials can be tested, and therefore the assay is adequate for monitoring purification and fractionation protocols.

The ferric orcinol assay was introduced by

Klenk and Langerbeins (1941) and has been used since then with some adaptations to reduce the amount of sample required (Schauer, 1978). Interference caused by other saccharides has been observed (Svennerholm, 1963; Spiro, 1966; Veh et al., 1977).

Critical Parameters

Free or glycosidically bound pentoses, hexoses, and uronic acids can interfere with the assay, limiting its quantitative value. Therefore, some error can be expected when analyzing biological materials that, purified or not, have a high content of these sugars. Nevertheless, this assay provides a simple way to detect sialic acids.

Anticipated Results

When working with a crude biological mixture, a clearly positive ferric orcinol assay will indicate the presence of sialic acids in the preparation; however, an accurate estimation of the total content of sialic acids cannot be obtained by this method. A negative assay, on the contrary, will not rule out the presence of sialic acid, because small amounts of these may be masked by other substances (Critical Parameters).

Time Considerations

The content of sialic acids in a set of up to ten samples can be obtained in 1 hr, provided standards and reagents are already prepared.

Literature Cited

- Klenk, E. and Langerbeins, H. 1941. Über die Verteilung der Neuraminsäure im Gehirn (Mit einer Mikromethode zur quantitativen Bestimmung der Substanz im Nervengewebe). *Hoppe-Seyler's Z. Physiol. Chem.* 270:185-193.
- Schauer, R. 1978. Characterization of sialic acids. *Methods Enzymol.* 50:64-89.
- Spiro, R.G. 1966. Analysis of sugars found in glycoproteins. *Methods Enzymol.* 8:3-26.
- Svennerholm, L. 1963. Sialic acids and derivatives: Estimation by the ion-exchange method. *Methods Enzymol.* 6:459-462.
- Veh, R.W., Corfield, A.P., Sander, M., and Schauer, R. 1977. Neuraminic acid-specific modification and tritium labelling of gangliosides. *Biochim. Biophys. Acta* 486:145-160.

Key Reference

Schauer, R. 1978. See above.

Describes the method and comments on its specificity and sensitivity.

BASIC PROTOCOL 3

MBTH ASSAY FOR HEXOSAMINES AND ACETYLHEXOSAMINES

Most methods reported for determining the concentration of hexosamines or *N*-acetylhexosamines are based on the procedures proposed by Elson and Morgan (see *UNIT 17.18*). However, the Elson-Morgan reaction is not specific for hexosamines and requires strict attention to reaction conditions. Although several modifications have been developed, in each case the formation of chromophore requires free amino sugars (not bound within an oligosaccharide), *N*-acetylated or not. Quantitative liberation of amino sugars from oligosaccharides is difficult because of their resistance to acid hydrolysis (see *UNIT 17.16*). In this protocol, the concentration of free hexosamines as well as those bound within an oligosaccharide (including *N*-acetylated species) can be determined after deamination of the sample with 3-methyl-2-benzothiazolinone hydrazone hydrochloride (MBTH) under mild acidic conditions.

Materials

- Sample to be analyzed
- 1 M HCl (ACS reagent-grade)
- Hexosamine standard: 1 mM *N*-acetyl-D-glucosamine or *N*-acetyl-D-galactosamine (Sigma)
- 2.5% (w/v) sodium nitrite (Sigma; store in dark bottle at room temperature; prepare fresh weekly)
- 12.5% (w/v) ammonium sulfamate (ACS reagent-grade; store ≥ 2 months at room temperature)
- 0.25% (w/v) 3-methyl-2-benzothiazolinone hydrazone hydrochloride (MBTH; Kodak; store in dark bottle at 4°C and replace if yellowish)

0.5% (w/v) ferric chloride (certified ACS reagent-grade; store indefinitely at room temperature)

13 × 100-mm Pyrex test tubes with Teflon-lined screw caps

Heating block, 110°C

Water baths, room temperature and 37°C

1.0-ml glass cuvettes (1-cm path length)

Visible-wavelength spectrophotometer

1. In triplicate, add 100 µl of sample to be analyzed to 13 × 100-mm Pyrex test tubes. Add 100 µl of 1 M HCl. Mix well by vortexing.
2. Set up a blank tube containing 100 µl water and 100 µl of 1 M HCl.
3. Prepare a series of tubes containing 5, 10, 15, 20, 25, and 30 nmol *N*-acetyl-D-glucosamine or *N*-acetyl-D-galactosamine standard in 100 µl water. Add 100 µl of 1 M HCl. Prepare each determination in triplicate.
4. Cap tubes and place in heating block 2 hr at 110°C.
5. Transfer tubes to a rack and place in a water bath. Allow samples to cool to room temperature.
6. In a fume hood, add 400 µl of 2.5% sodium nitrite and vortex. Let stand 15 min at room temperature.
CAUTION: Steps 6 to 8 should be done in a fume hood.
7. Add 200 µl of 12.5% ammonium sulfamate and vortex. Let stand 5 min at room temperature.
Allow the nitrogen oxides (brownish fumes) to dissipate.
8. Remove tubes from fume hood. Add 200 µl of 0.25% MBTH and vortex. Cap tubes and incubate 30 min at 37°C.
9. Add 200 µl of 0.5% ferric chloride, cap tubes, and incubate 5 min at 37°C.
10. Allow tubes to cool to room temperature and carefully transfer solutions to 1.0-ml cuvette with a Pasteur pipet.
11. Measure absorbance of samples at 650 nm, using reagent blank to zero spectrophotometer.
Absorbances between 0.1 and 1.0 U should be expected for the standards.
12. Prepare standard curve of absorbance versus nmol hexosamine using the average value for each hexosamine concentration.
13. Determine the amount of hexosamine in the sample by comparing to standard curve.

Commentary

Background Information

The MBTH assay is highly specific for amino sugars. This method is based on the formation of an intense blue-colored complex between the 2,5-anhydrohexoses produced upon deamination of hexosamines and 3-methyl-2-benzothiazolinone hydrazone hydrochloride (MBTH) under mild acidic conditions. Strong acidic conditions are not required because complete cleavage of the glycosidic linkage is not necessary for the MBTH color reac-

tion to occur. Heating in mild acid is sufficient for the complete de-*N*-acetylation of *N*-acetylhexosamines (i.e., *N*-acetylglucosamine and *N*-acetylgalactosamine) that is required for the reaction to proceed. Sodium nitrite is added to the acidic solution and the liberated nitrous acid produces the deamination. Sodium sulfamate is used to destroy the excess of nitrous acid. Final color is developed upon addition of ferric chloride to the samples. The proposed reaction mechanism is shown in Figure 17.9.1. The

method is useful for determination of free or glycosidically bound hexosamines, including *N*-acetylated species.

The MBTH method was introduced by Tsuji et al. (1969a), who proposed a mechanism for the formation of the blue chromophore. They also found that methyl glucosaminide gives an extinction coefficient almost identical to that of free hexosamine, showing that complete hydrolysis is not required (Tsuji et al., 1969b). Several modifications were introduced (Smith and Gilkerson, 1979) to reduce the difficulty of the original protocol developed by Tsuji et al. (1969a).

Critical Parameters

The color reaction is highly specific for amino sugars because *N*-acetylated hexosamines, neutral monosaccharides, glucuronic

acid, ascorbic acid, and most amino acids yield no color. Some amino acids (e.g., tryptophan, threonine, and methionine) exhibit color, but the intensities are far lower than those obtained with hexosamines. The complete de-*N*-acetylation of *N*-acetylhexosamines necessary for the reaction is achieved by heating the sugar-containing samples in 0.5 M HCl for 2 hr at 110°C. Glucosamine and galactosamine give practically identical color yields ($\epsilon_{\max} = 3.86$ and 3.62×10^{-4} , respectively). Both monosaccharides show linear absorbances in the range of 1 to 30 $\mu\text{g/ml}$. The response of mannosamine is very low ($\epsilon_{\max} = 0.80 \times 10^{-4}$), although it is also linear. This color reaction also allows the estimation of free hexosamines in the presence of *N*-acetylhexosamines, if the de-*N*-acetylation step is omitted.

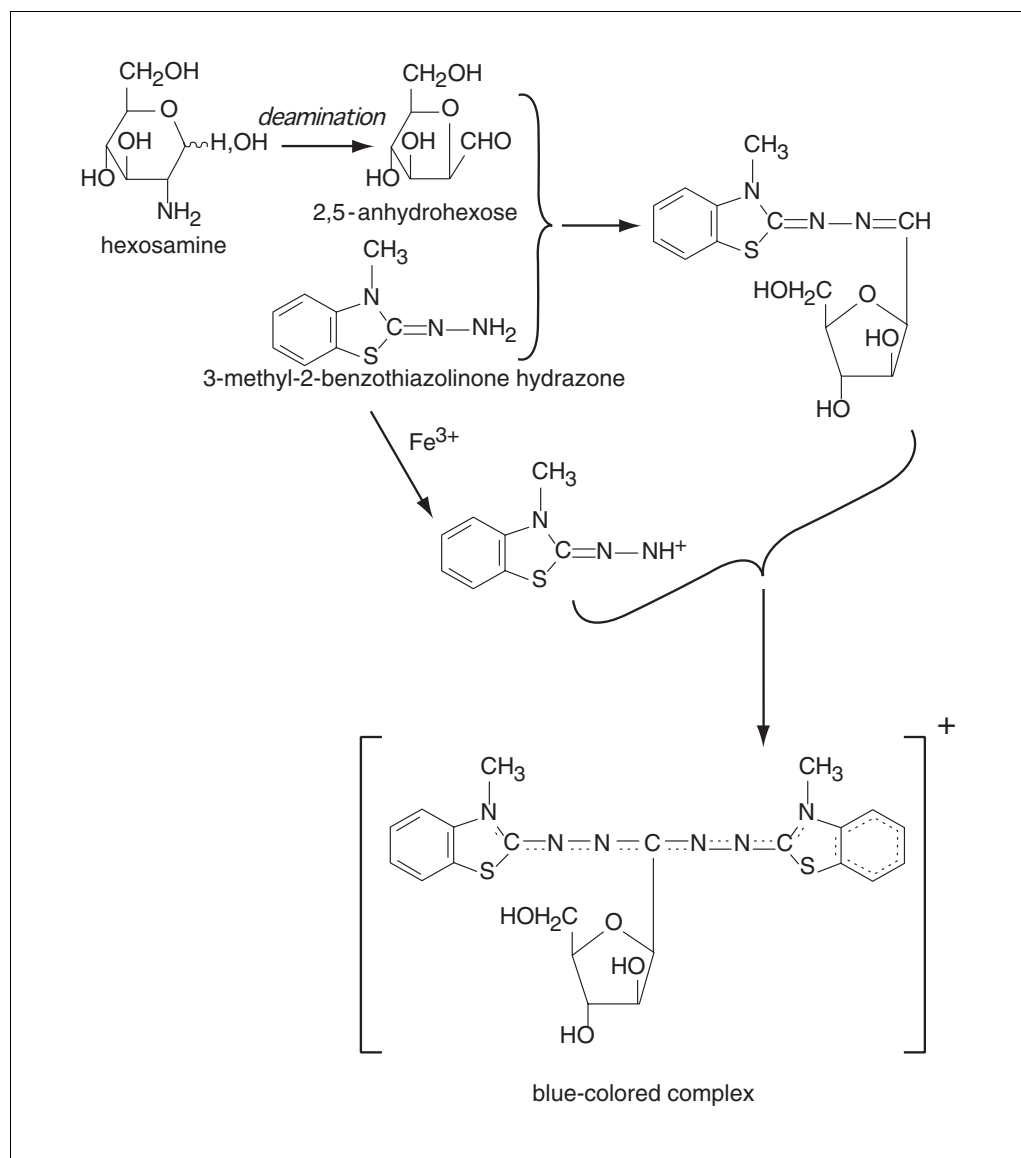


Figure 17.9.1 Probable mechanism of MBTH color reaction.

Anticipated Results

The application of this protocol will indicate the total content of glucosamine or galactosamine (*N*-acetylated or free) in the sample under analysis.

Time Considerations

The complete assay requires 4 hr, provided all reagents and solutions are prepared in advance.

Literature Cited

Smith, R.L. and Gilkerson, E. 1979. Quantitation of glycosaminoglycan hexosamine using 3-methyl-

2-benzothiazolone hydrazone hydrochloride. *Anal. Biochem.* 98:478-480.

Tsuji, A., Kinoshita, T., and Hoshino, M. 1969a. Microdetermination of hexosamines. *Chem. Pharm. Bull.* 17:217-218.

Tsuji, A., Kinoshita, T., and Hoshino, M. 1969b. Analytical chemical studies on amino sugars. II. Determination of hexosamines using 3-methyl-2-benzothiazolone hydrazone hydrochloride. *Chem. Pharm. Bull.* 17:1505-1510.

Key Reference

Smith, R.L. and Gilkerson, E. 1979. See above.

Describes the simplified protocol and its application to glycosaminoglycans.

ASSAYS FOR URONIC ACIDS

Hexuronic acids occur in polysaccharides of plants, fungal cell walls, bacterial capsules, animal tissues, and cultured cells. In animal tissues and cells, most uronic acid is found in glycosaminoglycans and proteoglycans. Considerable effort has been expended to develop specific, rapid methods for detecting and quantifying free and polymeric uronic acids. The simplest procedures involve hydrolysis of polysaccharides with mineral acid, dehydration of the uronic acid, and conjugation to a chromogen such as carbazole (Basic Protocol 4) or *meta*-hydroxybiphenyl (Alternate Protocol) to form a chromophore that can be quantitated by spectrophotometry. The reactions are generally specific for hexuronic acids, sufficiently sensitive for most laboratory situations (≥ 10 nmol), and require only routine laboratory equipment.

Carbazole Assay for Uronic Acids

Uronic acid is detected by monitoring the formation of a chromophore following reaction with the chromogen carbazole. The reaction is carried out in two steps: samples are first treated with strong acid to hydrolyze polysaccharides and dehydrate the sugars, and the products are then treated with carbazole or *meta*-hydroxybiphenyl to generate the chromophore.

Materials

Sample to be assayed

Uronic acid standard: 1 mM glucuronolactone (see recipe)

4 M ammonium sulfamate (Sigma)

25 mM sodium tetraborate in sulfuric acid (see recipe)

0.1% (w/v) carbazole (Kodak) in 95% ethanol (store at 4°C)

16 × 125-mm Pyrex tubes (open, glass-stoppered, or screw-cap culture tubes fitted with Teflon-lined caps)

Water bath, 100°C, or heating block

Disposable glass or plastic 1.0-ml cuvettes

UV/VIS spectrophotometer

1. Prepare duplicates of unknown sample in 16 × 125-mm Pyrex tubes. Adjust volume to 200 μ l with water. Set up a blank reaction tube containing 200 μ l water.
2. Prepare a series of tubes containing 10, 20, 30, 40, 50, 75, and 100 nmol glucuronolactone standard in 200 μ l water. Include a blank containing only 0.2 ml water.
3. Add 20 μ l of 4 M ammonium sulfamate and vortex.

BASIC PROTOCOL 4

Preparation and Analysis of Glycoconjugates

17.9.8

4. Add 1 ml of 25 mM sodium tetraborate in H_2SO_4 and mix carefully. Cover tubes with caps or marbles and heat 5 min at 100°C. Cool samples to room temperature.

CAUTION: Mixing sulfuric acid with the sample liberates considerable heat. Some investigators chill the samples on ice and use cold ammonium sulfamate. Make sure the tubes are vented during the heating procedure and wear eye protection.

5. Add 40 μ l of 0.1% carbazole and heat 15 min at 100°C. Cool samples to room temperature.

Color develops during this step.

6. Transfer samples to 1.0-ml disposable cuvettes using a Pasteur pipet and read absorbance at 520 nm.

Absorbance of blank sample should be ≤ 0.025 when read against water.

NOTE: Plastic disposable cuvettes can be used. Do not spill the sulfuric acid-containing samples on the outside of the cuvette or into the spectrophotometer.

7. Prepare a standard curve of A_{520} versus nmol glucuronic acid. Determine the amount of uronic acid in unknown sample by reference to the standard curve.

Amount of uronic acid will be expressed as glucuronic acid equivalents.

ALTERNATE PROTOCOL

Meta-Hydroxybiphenyl Assay for Uronic Acids

Substitution of *meta*-hydroxybiphenyl for the carbazole used in Basic Protocol 4 allows color development at room temperature.

Additional Materials (also see Basic Protocol 4)

0.15% (w/v) *meta*-hydroxybiphenyl (Kodak) in 0.5% (w/v) NaOH

1. Prepare samples and treat with ammonium sulfamate and sodium tetraborate in sulfuric acid (see Basic Protocol 4, steps 1 to 4).
2. Add 40 μ l of 15% *meta*-hydroxybiphenyl and incubate 15 min at room temperature.
3. Transfer samples to disposable cuvettes, read absorbance, and determine the amount of uronic acid by reference to the standard curve (see Basic Protocol 4, steps 6 and 7).

Reagents and Solutions

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Glucuronolactone, 1 mM

Dissolve 88 mg of glucuronolactone in 1 ml water and freeze in 0.2-ml aliquots. Before assay, dilute one 0.2-ml aliquot to 100 ml with water to obtain 1 mM solution. Store 3 to 6 months at 4°C.

Sodium tetraborate, 25 mM in H_2SO_4

Working in a fume hood, dissolve 0.95 g sodium tetraborate ($Na_2B_4O_7 \cdot 10H_2O$; Sigma) in 100 ml concentrated sulfuric acid. Stir at room temperature with magnetic stir bar.

A large volume of solution can be prepared in advance because the reagent is quite stable. Store at room temperature.

CAUTION: Sulfuric acid will cause severe burns.

Commentary

Background Information

Carbazole reacts with hexuronic acids treated with strong acid. The reaction was first described by Dische (1947) and subsequent modifications have improved sensitivity, specificity, and rapidity. The original reaction conditions gave varying color yields with different uronic acids, but the inclusion of borate (Bitter and Muir, 1962) lessened this problem. Nevertheless, it should be noted that different uronic acids give somewhat different responses in the assay. When uronic acids constitute a small proportion of the total sugar, the accuracy of the assay is reduced due to the Browning reaction of other sugars in hot sulfuric acid. Inclusion of sulfamate (Galambos, 1967) reduces the color yield from neutral sugars to $\leq 5\%$ that of uronic acids. Substitution of *meta*-hydroxybiphenyl for carbazole further reduces interference from neutral sugars (Blumenkrantz and Asboe-Hansen, 1973; Filisetti-Cozzi and Carpita, 1991), because color development can be carried out at room temperature. Thus, the Alternat Protocol should be used if uronic acid represents $<20\%$ of the total hexose in the sample.

Critical Parameters

The three most substantial problems encountered in quantitating uronic acids are the varying color yields of different uronic acids, presence of colored reaction products from other components in mixtures, and interference by certain salts. The inclusion of borate substantially improves the reactivity of recalcitrant uronic acids such as mannuronic acid. Uronic acids in glycosaminoglycan chains can react differently than free sugars, e.g., heparin gives an anomalous high color yield (Kosakai and Yosizawa, 1978). The differential color yields of uronic acids require use of appropriate standards for accuracy. When the identity of the glycosaminoglycan (GAG) chains is known, preparation of a standard curve using the same GAG rather than glucuronolactone will give more accurate results.

The substitution of *meta*-hydroxybiphenyl for carbazole lowers the production of nonspecific color due to hexoses beyond that obtained by inclusion of 0.1 M sulfamate. Hexosamines and pentoses do not produce color in either reaction. The amount of sulfamate included assures maximum reduction of interference. Larger amounts will result in decreased produc-

tion of the uronic acid chromophore. Carbazole produces somewhat more chromophore than *meta*-hydroxybiphenyl, but the difference is small. Addition of sulfamate reduces the color yield by only $\sim 30\%$.

Interference by salts and other cellular components has been surveyed in greater detail in the carbazole reaction than in the *meta*-hydroxybiphenyl reaction. Samples containing sodium chloride (<2 M), urea (<6 M), guanidine-HCl (<2 M), and protein (<0.2 mg/ml) yield the expected color. Sodium nitrite (>5 μ M) and hydrogen peroxide (>10 nM) give a green color that overwhelms the signal from uronic acids. Sulfhydryl reagents accelerate color formation. Because the *meta*-hydroxybiphenyl assay works under milder conditions, it is preferable to the carbazole assay. However, interference by contaminants should be checked by adding them to a standard solution.

Anticipated Results

The assay will provide an estimate of total glycosyluronic acids in a sample. The extinction coefficient at 520 nm for a 1% solution is ~ 1000 . Thus, the assay is sufficiently sensitive to detect ~ 10 to 25 nmol uronic acid.

Time Considerations

The assay can be completed in ~ 20 to 30 min if the reagents and solution are prepared in advance.

Literature Cited

- Bitter, T. and Muir, H.M. 1962. A modified uronic acid carbazole reaction. *Anal. Biochem.* 4:330-334.
- Blumenkrantz, N. and Asboe-Hansen, G. 1973. New method for quantitative determination of uronic acids. *Anal. Biochem.* 54:484-489.
- Dische, Z. 1947. A new specific color reaction of hexuronic acids. *J. Biol. Chem.* 167:189-198.
- Filisetti-Cozzi, T.M. and Carpita, N.C. 1991. Measurement of uronic acids without interference from neutral sugars. *Anal. Biochem.* 197:157-162.
- Galambos, J.T. 1967. The reaction of carbazole with carbohydrates. I. Effect of borate and sulfamate on the carbazole color of sugars. *Anal. Biochem.* 19:119-132.
- Kosakai, M. and Yosizawa, Z. 1978. Study on the factors yielding high color in the carbazole reaction with hexuronic acid-containing substances. *J. Biochem.* 84:779-785.

Key References

Bitter and Muir, 1962. See above.

Filisetti-Cozzi and Carpita, 1991. See above.

These two references describe the most recent modifications of the methods.

Contributed by Adriana Manzi
University of California San Diego
La Jolla, California

Jeffrey Esko
University of Alabama at Birmingham
Birmingham, Alabama

Inhibition of N-Linked Glycosylation

UNIT 17.10A

**BASIC
PROTOCOL**

Treatment of cells with inhibitors of the enzymes that synthesize N-linked oligosaccharide chains results in production of glycoproteins (this unit) or glycolipids (UNIT 17.10B) containing missing or altered chains. This approach is useful for examining potential functional roles of oligosaccharides on specific proteins or intact cells. Table 17.10.1 lists the enzymes blocked by each inhibitor and the consequences; Figure 17.10.1 diagrams the early processing steps in the assembly of N-linked oligosaccharides, showing the enzymes blocked by each inhibitor except tunicamycin (which affects an earlier stage). These inhibitors can be used to prevent N-linked glycosylation in cultured cells. First, the optimal concentration of inhibitor (i.e., highest nontoxic concentration) is determined by monitoring [³⁵S]methionine incorporation as a measure of protein biosynthesis (UNIT 10.18). The inhibitor's ability to inhibit oligosaccharide processing is then determined by analyzing cells labeled with [³H]mannose using TCA precipitation (UNIT

Table 17.10.1 Inhibitors of Enzymes that Synthesize N-linked Oligosaccharides

Type of inhibitor	Target enzyme	Effect on oligosaccharide structure	Effective concentration range
Glycosylation inhibitor			
Tunicamycin	GlcNAc transferase	Prevents assembly of GlcNAc-PP-dolichol and thus assembly of G ₃ M ₉ GlcNAc ₂ -PP-dolichol ^a ; glycosylation of Asn residues does not occur; proteins migrate faster on SDS-PAGE and generally show less size heterogeneity; proteins will not shift to a faster mobility on SDS-PAGE after PNGase F treatment	0.5-10 µg/ml
Processing inhibitors			
Deoxynojirimycin Castanospermine	Glucosidase I and/or II	Prevents removal of first glucose residue, thereby inhibiting any further processing of the oligosaccharide chain; proteins may migrate with a larger or smaller size on SDS-PAGE, depending on extent of processing that normally occurs; sensitive to endo H ^b	0.5-200 mM 1-50 µg/ml
Deoxymanno- jirimycin	α-mannosidase I	Prevents removal of mannose residues on the α1-3 arm of the high mannose structure, thereby blocking the activity of GlcNAc T I and thus α-mannosidase II; proteins generally run as a smaller size on SDS-PAGE, depending on size of the normal oligosaccharide; structures remain sensitive to endo H ^c	1-5 mM
Swainsonine	α-mannosidase II	Prevents removal of mannose residues on the α1-6 arm of the high mannose structure, preventing the activities of GlcNAc T II and V (which add GlcNAc to the α1-6 mannose residue); addition of GlcNAc, galactose, and sialic acid to the α1-3 mannose can occur normally, although the structures remain sensitive to endo H digestion	1-10 µg/ml

^aAbbreviations: G₃M₉GlcNAc₂-PP-dolichol, Glc₃Man₉GlcNAc-PP-dolichol; GlcNAc T I, II, V, *N*-acetylglucosaminyltransferase I, II, V; PNGase F, peptide N:glycosidase F.

^bMature processed oligosaccharide structures cannot be cleaved from the protein by endoglycosidase H; structures retaining mannose residues on the α1-6 arm can be removed, generally resulting in a difference in size easily detected by SDS-PAGE.

^cEndomannosidases can bypass this block.

Contributed by Leland D. Powell

Current Protocols in Molecular Biology (1995) 17.10.1-17.10.9

Copyright © 2000 by John Wiley & Sons, Inc.

17.10.1

Supplement 32

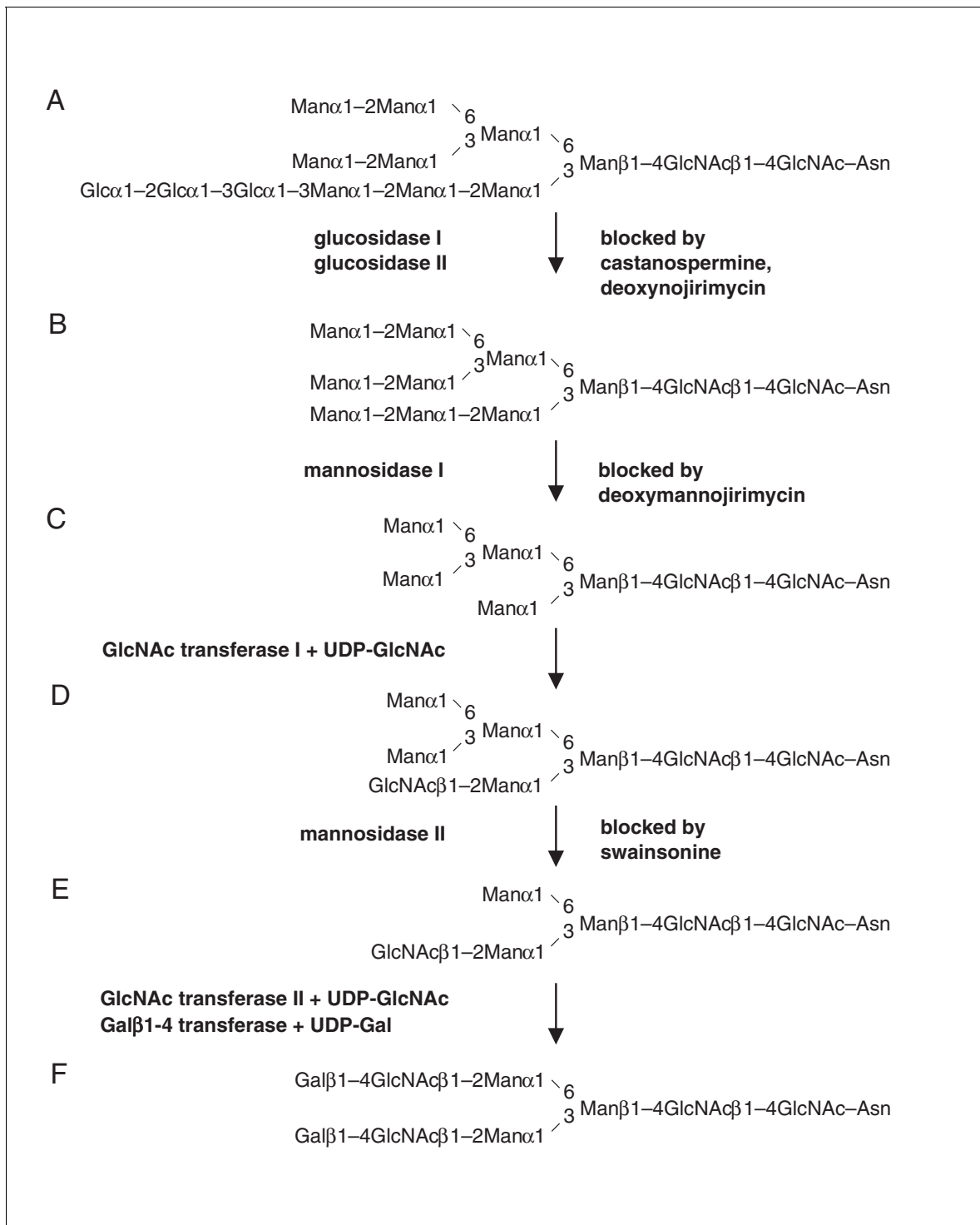


Figure 17.10.1 Structures of early processing intermediates in the assembly of N-linked oligosaccharides. Structure A is initially assembled on the lipid carrier dolichol phosphate, a process blocked by tunicamycin, and then transferred to the Asn residue. The initial trimming glucosidases I and II convert A into B; these glucosidases are inhibited by castanospermine and deoxynojirimycin. Deglucosidation is followed by removal of the outer four mannose residues by mannosidase I (inhibited by deoxymannojirimycin), to form C. Removal of these mannose residues is the signal for the first *N*-acetylglucosaminyltransferase (GlcNAc Transferase I) to form D. Structure D, but not C, is a substrate for mannosidase II, resulting in the formation of E, which can be further extended into bi-, tri- and tetraantennary structures (sialylated and neutral) by a series of *N*-acetylglucosaminyltransferases, galactosyltransferases, and sialyltransferases plus the appropriate sugar-nucleotide donors.

17.10.2

10.18) or endo H digestion (UNIT 17.13). Further suggestions are given on how to use methods for identifying a specific glycoprotein (if available) to measure the effect of the inhibitor on its N-linked oligosaccharide chains. A support protocol details a method for concentrating proteins by acetone precipitation.

NOTE: All media and solutions should be made with distilled, deionized water. All media and equipment coming into contact with cells should be sterile. Incubations of cells should be performed in a humidified 5% CO₂, 37°C incubator.

Materials

Cultured cell line, either adherent or suspension
Complete culture medium
Inhibitor of N-linked glycosylation (one or more of the following: tunicamycin, deoxynojirimycin, castanospermine, deoxymannojirimycin, or swainsonine; see reagents and solutions and Table 17.10.2)
Solvent used for making inhibitor solution (see reagents and solutions)
Multiply deficient medium (MDM; UNIT 17.4) without glucose
 [³H]mannose (5 to 20 Ci/mmol)
Phosphate-buffered saline (PBS; APPENDIX 2), ice-cold
Lysis buffer (UNIT 10.16) without 1% bovine hemoglobin
0.5 U/ml endoglycosidase H (endo H) and endo H digestion buffer (reagents and solutions)
20% (w/v) SDS
Sephacryl S-200 column (Table 10.9.2)
25 mM ammonium formate/0.1% (w/v) SDS
24-well tissue culture plate
100-mm tissue culture plates
Disposable plastic scraper or rubber policeman
1.5-ml, 15-ml, or 50-ml conical polypropylene centrifuge tube
Additional reagents and equipment for tissue culture (Chapter 9 introduction), counting viable cells (UNIT 11.5), short-term labeling of suspension or adherent cells with [³⁵S]methionine (UNIT 10.18), metabolic radiolabeling of glycoconjugates (UNIT 17.4), lysis of cells (as for immunoprecipitation; UNIT 10.16), TCA precipitation (UNIT 10.18), endo H digestion (UNIT 17.13), gel-filtration chromatography (UNIT 10.9), and one-dimensional gel electrophoresis of proteins (UNIT 10.2)

Determine optimal inhibitor concentration

1. For each inhibitor tested, set up 15 wells of cells in a 24-well tissue culture plate by splitting the cells (by trypsinization for adherent cells or dilution for suspension cells) into 1.8 ml final per well of complete medium. Incubate 24 hr.

A ratio of splitting should be used to yield confluent (but not overgrown) cultures by the third day (i.e., 48 hr after splitting).

2. While cells are incubating, make a series of 1:1 (v:v) dilutions of inhibitor(s) as follows. Place the volume of inhibitor stock solution indicated in Table 17.10.2 in a sterile microcentrifuge tube and dilute to 800 µl final volume with complete culture medium. Transfer 400 µl to a second tube and dilute with 400 µl complete culture medium. Repeat for a total of 7 tubes. As a control, prepare an identical series of dilutions of the solvent used for making the inhibitor solution.

Table 17.10.2 Range of Inhibitor Concentrations for Dilution Series

Inhibitor	Stock solution	Stock added to first dilution (μ l)	Concentration range
Tunicamycin	1 mg/ml	80	0.15-10 μ g/ml
Swainsonine	1 mg/ml	80	0.15-10 μ g/ml
Deoxynojirimycin	400 mM	200	0.15-10 mM
Deoxymannojirimycin	400 mM	200	0.15-10 mM
Castanospermine	400 mM	200	0.15-10 mM

3. Add 200 μ l from each tube (from step 1) to a well of the tissue culture plate containing cells (2 ml final). Add 200 μ l complete culture medium alone to the remaining well (as a zero-inhibitor control point). Incubate plate 24 hr. Store unused inhibitor-containing media at 4°C.

If the inhibitor is tunicamycin, dilutions should be prepared fresh each day, as it may precipitate out of solution during storage.

4. Perform short-term labeling of cells with [³⁵S]methionine (UNIT 10.18) using 0.2 mCi/ml label and 1.8 ml final volume per well.
5. Add to wells the remaining 200 μ l of each dilution of inhibitor (from step 2) and incubate a further 4 hr.
6. Harvest cells and determine the incorporation of [³⁵S]methionine into macromolecules by TCA precipitation (UNIT 10.18). Plot label incorporation versus inhibitor concentration.

CAUTION: Safety precautions for disposing of radioactive waste should be followed.

Comparing incorporation by cells incubated with and without inhibitor will indicate the maximum concentration of inhibitor that does not inhibit protein synthesis. This is the optimal concentration for the following steps.

Label cells in the presence of N-linked glycosylation inhibitor

7. Split a new cell sample into complete medium in a 100-mm tissue culture plate and incubate 24 hr.
8. Add inhibitor to the optimal concentration determined above. Set up a control plate containing the same quantity of stock solution solvent but no inhibitor. Incubate for 24 hr.
9. Wash the cells using the same procedure as employed during labeling (step 4) and, to each plate, add sufficient glucose-free MDM to cover the cells (5 ml for a 100-mm plate). Add inhibitor or solvent as in step 8. Add [³H]mannose to 0.02 to 0.1 mCi/ml. Incubate 4 to 12 hr.

The duration of incubation will depend on how well the cells label with [³H]mannose. The amount of [³H]mannose used will depend on the cell line examined. However, for most cells, a single 100-mm plate with 5 ml media incubated with 0.02 mCi/ml [³H]mannose for 6 hr should yield sufficient radioactive material for the analyses described below. Extended incubations in glucose-free labeling medium should be avoided.

10. Harvest the cells by scraping with a disposable scraper or rubber policeman (for adherent cells) or centrifugation (for suspension cells) into ice-cold PBS.

CAUTION: Safety precautions for disposing of radioactive waste should be followed.

Measure incorporated macromolecular radioactivity

11. If tunicamycin was used as the inhibitor, determine the amount of incorporated macromolecular radioactivity by TCA precipitation and proceed to step 18. If another inhibitor was used, carry out steps 12 to 17.

The amount of radioactivity incorporated into macromolecules reflects the effect of the inhibitor(s) on oligosaccharide biosynthesis: the presence of inhibitor should block incorporation of [³H]mannose into macromolecules.

12. Wash the cells twice by resuspending the pellet in ice-cold PBS, centrifuging, and removing the supernatant. Resuspend the pellet in lysis buffer without bovine hemoglobin and lyse as described in UNIT 10.16 (steps 1 to 4, basic protocol).
13. Divide each sample into two equal aliquots in conical polypropylene centrifuge tubes and precipitate protein with acetone (support protocol).

Use 1.5-ml, 15-ml, or 50-ml centrifuge tubes depending on the quantity of sample (8 vol acetone must be added).

14. Resuspend pellets in 20 to 40 μ l endo H digestion buffer per 10^7 cells and boil 10 min to inactivate endogenous hydrolases.
15. Add 5 μ l endo H per 100 μ l sample to one tube of each sample pair and incubate both overnight at 37°C.
16. Add $\frac{1}{10}$ vol of 20% (w/v) SDS to each sample and boil 3 to 5 min to terminate the digestion.
17. Apply one sample at a time to a calibrated Sephacryl S-200 column equilibrated in 25 mM ammonium formate/0.1% SDS, in a sample volume that is <5% of the bed volume. Collect the eluate and use a scintillation counter to measure the radioactivity associated with glycoproteins (eluting in the void volume of the column, V_o) and with free oligosaccharides (eluting after V_o) for each of the four samples. Wash column with 2 to 4 vol 25 mM ammonium formate/0.1% SDS between sample applications.

In the absence of endo H digestion, all of the radioactivity from both the inhibitor-treated and untreated samples should elute in V_o . Elution of any radioactive material after V_o may indicate that the wash steps following acetone precipitation did not completely remove low-molecular-weight contaminants present in the cell lysate.

For samples digested with endo H, elution profiles should demonstrate a marked increase in the amount of endo H-releasable radioactivity in samples from inhibitor-treated cells relative to treated cells. For samples not treated with inhibitor, the oligosaccharides cleaved from the peptide backbone by endo H represent immature oligosaccharides still being processed in the Golgi apparatus and a small population of endo H-sensitive structures found on some mature glycoproteins. For samples treated with inhibitor, a much higher percentage of the total radioactivity should be released, as the inhibitors prevent the processing of the oligosaccharides into mature oligosaccharide structures (Fig. 17.10.1) and the underprocessed oligosaccharides contain more mannose residues than the processed ones. Theoretically, 100% of the radioactivity should be releasable by endo H digestion of the inhibitor-treated samples (see critical parameters), although in reality no inhibition is ever complete.

To save time, analyze only the first and last tubes of the control cells in this fashion, i.e., those not treated at all and those treated with the highest concentration of the inhibitor solvent (from step 2 above). If there is no difference in the results from these two groups of cells, it is reasonable to assume that those cells treated with lower concentrations of the inhibitor solvent will yield the same results. However, if the solvent is found to be toxic, then the cells exposed to lower concentrations will need to be examined to find the maximally tolerated concentration of inhibitor solvent.

18. If desired, and if methods are available for purifying and identifying a specific glycoprotein from cells radiolabeled with [³⁵S]methionine, examine the effect of a given inhibitor by SDS-PAGE and autoradiography.

Each N-linked carbohydrate chain contributes ~2000 to 4000 Da to a protein's mass, so synthesis of a protein in the presence of tunicamycin will result in faster mobility on SDS-PAGE. The effect of other inhibitors can be demonstrated by an increased sensitivity to digestion by endo H, which is also detectable as an increase in migration rate in SDS-PAGE.

SUPPORT PROTOCOL

ACETONE PRECIPITATION

Acetone precipitation is a useful step for concentrating proteins and for exchanging them from one buffer to another. As little as 10 ng of protein can be precipitated successfully.

Additional Materials

100% acetone (HPLC or ACS grade), -20°C

1.5-ml, 15-ml, or 50-ml conical polypropylene or other centrifuge tubes

1. Place the protein solution on ice, add 8 vol acetone (-20°C), mix gently, and precipitate at -20°C overnight.

If the sample volume is too large to be conveniently diluted with 8 vol, concentrate it by lyophilizing and resuspending in a smaller volume. If the sample contains NP-40 or Triton X-100 detergent, however, there is a limit to how much it can be concentrated, as high concentrations of detergent (e.g., >1%) will partition out of acetone solutions, appearing as an oily pellet after centrifugation (step 2).

If a large amount of protein is present (i.e., a precipitate is readily visible after the acetone is added), it can be precipitated after standing at -20°C for only a few hours instead of overnight. Leaving the material in acetone for longer than 1 day should be avoided, as it may become increasingly difficult to redissolve the protein.

2. Collect the precipitate by centrifuging 15 min at 3000 × g, 4°C (a small pellet should be visible). Carefully invert the tube and pour out the acetone into a clean tube.

If the pellet is oily due to the presence of detergent, it should be extracted by adding 85% (v/v) acetone (-20°C), mixing, incubating at -20°C for >1 hr, and recentrifuging.

3. Centrifuge the tube containing the pellet briefly to concentrate the remaining acetone and remove it with a micropipettor. Allow pellet to dry until moist but not powder-dry, and resuspend in an appropriate quantity of the desired solvent.

Care must be taken in drying the pellet, as great difficulty may be encountered in resuspending some proteins if the pellet dries completely. This insolubility depends on both the individual protein and the total amount present.

REAGENTS AND SOLUTIONS

NOTE: Deionized, distilled water should be used for buffer and inhibitor solutions.

Endo H digestion buffer

15 µl 0.5 M sodium citrate, pH 5.5

75 µl water

5 µl 10% PMSF in isopropanol

This is sufficient for 5 µl of 0.5 U/ml endo H (100 µl digestion volume). The PMSF prevents proteolysis. The presence of nonionic detergent is not required. Prepare just before use; PMSF is unstable in water.

Inhibitor stock solutions

Dissolve castanospermine (mol. wt. 189.2), 1-deoxymannojirimycin (mol. wt. 199.6), or 1-deoxynojirimycin (mol. wt. 163.2) in water to 400 mM; dissolve swainsonine (mol. wt. 173.2) in water to 1 mg/ml. Sterilize by filtration through a 0.2- μ m filter. Store frozen at -20°C (stable indefinitely).

Dissolve tunicamycin (mol. wt. 840) in dimethylsulfoxide (DMSO), dimethylformamide (DMF), 95% ethanol, or 25 mM NaOH to 1 mg/ml. Store frozen at -20°C (stable \sim 1 yr).

If NaOH is used, store stock solution in a small tube that is tightly capped to prevent adsorption of carbon dioxide, which will lower the pH. Tunicamycin's solubility in neutral aqueous solution is <1 mg/ml (attempting to make such a solution will produce a visible precipitate).

COMMENTARY

Background Information

Assembly of N-linked oligosaccharide chains involves the sequential action of several glycosidases and glycosyltransferases (reviewed by Hubbard and Ivatt, 1981, and Kobata and Takasaki, 1992). Growth of cells in the presence of inhibitors of these enzymes produces glycoproteins with underprocessed or missing oligosaccharide chains. Depending on the protein, these changes may alter biological functions and/or rates of transport, secretion, or turnover (Elbein, 1987; McDowell and Schwarz, 1988).

N-linked glycosylation begins with the assembly of the lipid-linked precursor glucose₃mannose₉N-acetylglucosamine₂-pyrophosphate-dolichol (Glc₃Man₉GlcNAc₂-PP-dolichol; similar to Structure A, Fig. 17.10.1). Assembly of this precursor requires N-acetylglucosamine-pyrophosphate-dolichol (GlcNAc-PP-dolichol), whose formation is blocked by tunicamycin (Tkacz and Lampen, 1975). Thus, the presence of tunicamycin results in proteins missing some or all of their N-linked side chains; however, those that are added due to incomplete inhibition are processed normally. Tunicamycin may also inhibit protein synthesis; not all proteins will be affected to the same degree. Tunicamycin also blocks the assembly of type II keratan sulfate chains, because these glycans also utilize GlcNAc-PP-dolichol (Hart and Lennarz, 1978). Glycolipid biosynthesis may also be inhibited, although the mechanism underlying this has not been established (Yusuf et al., 1983; Guarnaccia et al., 1987).

The glycosidase inhibitors castanospermine, deoxynojirimycin, deoxymannojirimycin, and swainsonine are considerably less toxic than tunicamycin. Castanospermine inhibits glucosidase I, and deoxynojirimycin in-

hibits both glucosidases I and II (Fig. 17.10.1; Pan et al., 1983; Saunier et al., 1982). In some studies, the metabolic block produced by these two drugs is not complete, yielding a mixture of normally processed and underprocessed structures. More recently, an endo- α -D-mannosidase has been described that cleaves the tetrasaccharide Glc₃-Man₁ from Glc₃Man₉-GlcNAc₂, effectively bypassing the block produced by the glucosidase inhibitors (Lubas and Spiro, 1987). Although the question has not been examined fully, the level of this endo- α -D-mannosidase may vary between different cell lines.

Deoxymannojirimycin and swainsonine are inhibitors of mannosidases I and II, respectively (Fig. 17.10.1; Fuhrmann et al., 1984; Tulsiani et al., 1982). By inhibiting mannosidase I, deoxymannojirimycin prevents addition of the α 1-2 N-acetylglucosamine (GlcNAc) residue required to produce Structure D, Figure 17.10.1, which is a necessary step for the action of mannosidase II. Thus, deoxymannojirimycin prevents addition of any GlcNAc residues, blocking galactosylation and sialylation as well. Swainsonine blocks mannosidase II. However, it does not prevent the single GlcNAc residue (Structure D, Fig. 17.10.1) from being galactosylated and sialylated, resulting in hybrid structures (Tulsiani and Touster, 1983).

Some of the glycosidase inhibitors are also active against lysosomal enzymes, although the significance of this effect in tissue culture cells has been little explored. Deoxynojirimycin can inhibit synthesis of the lipid-linked precursor in cell-free extracts (Romero et al., 1985); this also has not been thoroughly investigated with intact cells. No information is available on the metabolism or half-lives of these inhibitors in cultured cells.

Critical Parameters and Troubleshooting

The two most critical variables are duration of exposure and inhibitor concentration. The treatment times suggested in this unit (24 hr) will permit a cell to replace most of its endogenous glycoproteins with glycoproteins synthesized in the presence of inhibitor. For proteins that turn over slower or faster than average, longer or shorter incubation times may be appropriate. If pulse-chase studies with radiolabeled sugar or amino acid precursors (*UNITS 17.4 & 10.18*) are being carried out, it is necessary to incubate cells in inhibitor for only 1 to 2 hr before adding label.

N-linked glycosylation inhibitors may be toxic to the cells, as indicated by decreased cell viability and/or reduced protein synthesis. The concentration of tunicamycin that can be used is generally limited by its toxicity, which may be considerable in some cell lines (more so with tumor cells than nontransformed cells). Other glycosidase inhibitors are usually not toxic in the concentration ranges suggested in Table 17.10.1 (the highest concentrations listed represent those used in most published studies).

It is important to realize that in a given cell line, not all glycoproteins will be affected to the same extent by a given inhibitor. Moreover, it is possible that not all sites on the same glycoprotein will be affected equally. These caveats should be kept in mind when interpreting the results from studies employing these inhibitors.

Failure to observe an effect from an inhibitor may indicate that concentration of inhibitor was too low and/or length of exposure was too short. Cell lines differ in their susceptibility to these compounds.

Safety precautions and experimental parameters for labeling—including labeling time, label concentration and specific activity, cell density, and culture conditions—are discussed in the critical parameters of *UNIT 10.18*.

Anticipated Results

These experiments will demonstrate the chosen inhibitor's toxicity, as assessed by protein synthesis, and its ability to alter oligosaccharide biosynthesis, as assessed by total [³H]mannose incorporation (for tunicamycin) or endo H release of the oligosaccharides (for the other inhibitors listed in Table 17.10.1).

Time Considerations

Setting up the initial round of cell cultures for determining the optimal inhibitor concentration takes 2 hr, after which they are incubated 24 hr. On the second day, the dilution series of medium containing inhibitor is set up, the cells are transferred to these media (<1 hr), and the plates incubated 24 hr. On the third day, the incubation of the cells with the radiolabeled precursor is set up (~4 hr); the actual incubation takes 4 to 16 hours. Following incubation, the cells are assayed for radiolabel incorporation by TCA precipitation (2 to 4 hr).

The main part of the experiment follows a similar time course if the inhibitor used is tunicamycin. If a different inhibitor is used, analysis of endo H susceptibility by Sephacryl S-200 chromatography will take ~1 day, including the overnight incubation.

Literature Cited

- Elbein, A.D. 1987. Inhibitors of the biosynthesis and processing of N-linked oligosaccharide chains. *Annu. Rev. Biochem.* 56:497-534.
- Fuhrmann, U., Bause, E., Legler, G., and Ploegh, H. 1984. Novel mannosidase inhibitor blocking conversion of high mannose to complex oligosaccharides. *Nature (Lond.)* 307:755-758.
- Guarnaccia, S.P., Shaper, J.H., and Schnaar, R.L. 1987. Tunicamycin inhibits ganglioside biosynthesis in neuronal cells. *Proc. Natl. Acad. Sci. U.S.A.* 80:1551-1555.
- Hart, G.W. and Lennarz, W.J. 1978. Effects of tunicamycin on the biosynthesis of glycosaminoglycans by embryonic chick cornea. *J. Biol. Chem.* 253:5795-5801.
- Hubbard, S.C. and Ivatt, R.J. 1981. Synthesis and processing of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* 50:555-583.
- Kobata, A. and Takasaki, S. 1992. Structure and biosynthesis of cell surface carbohydrates. In *Cell Surface Carbohydrates and Cell Development* (M. Fukuda, ed.) pp. 1-24. CRC Press, Boca Raton, Fla.
- Lubas, W.A. and Spiro, R.G. 1987. Golgi endo- α -D-mannosidase from rat liver, a novel N-linked carbohydrate unit processing enzyme. *J. Biol. Chem.* 262:3775-3781.
- McDowell, W. and Schwarz, R.T. 1988. Dissecting glycoprotein biosynthesis by the use of specific inhibitors. *Biochimie (Paris)* 70:1535-1549.
- Pan, Y.T., Hori, H., Saul, R., Sanford, B.A., Molyneux, R.J., and Elbein, A.D., 1983. Castanospermine inhibits the processing of the oligosaccharide portion of the influenza viral hemagglutinin. *Biochemistry* 22:3975-3984.

Romero, P., Friedlander, P., and Herscovics, A. 1985. Deoxynojirimycin inhibits the formation of Glc₃Man₉GlcNAc₂-PP-dolichol in intestinal epithelial cells in cultures. *FEBS Lett.* 183:29-32.

Saunier, B., Kilker, R.D., Tkacz, J.S., Quaroni, A., and Herscovics, A. 1982. Inhibition of N-linked complex oligosaccharide formation by 1-deoxynojirimycin, an inhibitor of processing glucosidases. *J. Biol. Chem.* 257:14155-14161.

Tkacz, J.S. and Lampen, J.O. 1975. Tunicamycin inhibition of polyisoprenyl *N*-acetylglucosaminyl pyrophosphate formation in calf liver microsomes. *Biochem. Biophys. Res. Commun.* 65:248-55.

Tulsiani, D.R.P., Harris, T.M., and Touster, O. 1982. Swainsonine inhibits the biosynthesis of complex glycoproteins by inhibition of Golgi mannosidase II. *J. Biol. Chem.* 257:7936-7939.

Tulsiani, D.R.P. and Touster, O. 1983. Swainsonine causes the production of hybrid glycoproteins by human skin fibroblasts and rat liver Golgi preparations. *J. Biol. Chem.* 258:7578-7585.

Yusuf, H.K.M., Pohlentz, G., Schwarzmann, G., and Sandhoff, K. 1983. Ganglioside biosynthesis in Golgi apparatus of rat liver. *Eur. J. Biochem.* 134:47-54.

Key References

Elbein, A.D. 1987. See above.

Useful general review.

McDowell, W. and Schwarz, R.T. 1988. See above.

Useful general review.

Hubbard, S.C. and Ivatt, R.J. 1981. See above.

Good review on early work on the processing pathway.

Kobata, A. and Takasaki, S. 1992. See above.

Good review of recent work on some of the complexities in oligosaccharide processing.

Contributed by Leland D. Powell
University of California San Diego
La Jolla, California

Inhibition of Glycolipid Biosynthesis

BASIC PROTOCOL

Adequate inhibition of glycolipid biosynthesis allows the study of their biological functions. The method presented in this unit employs a synthetic analog of ceramide, PDMP (1-phenyl-2-decanoylamino-3-morpholino-1-propanol), that inhibits glycolipid biosynthesis in cultured cells. Optimum conditions for inhibition of glycolipid biosynthesis are determined, glycolipids extracted from cultures grown with and without inhibitor, and the patterns of glycolipids analyzed by HPTLC (UNIT 9.7A). Detection is achieved using colorimetric reactions, or by monitoring radioactivity when cells have been metabolically radiolabeled. The effects of these changes in biological functions can then be studied later.

Materials

Adherent cell culture
 10 mM PDMP in ethanol (see recipe)
 Culture medium appropriate for cell line, with or without serum
 Ethanol
 Radiolabeled precursors (UNIT 17.4; for metabolic labeling)
 PBS (APPENDIX 2)
 2:1, 1:1, and 1:2 (v/v) chloroform/methanol
 10:10:1 (v/v/v) chloroform/methanol/water
 HPTLC standards: mixture of gangliosides or neutral lipids from bovine brain (Sigma or Accurate Chemical) or mixture of neutral glycolipids (Accurate Chemical)
 Orcinol/sulfuric acid reagent (see recipe)
 Resorcinol/hydrochloric acid reagent (see recipe)
 Screw-cap glass tube, sterile
 Conical glass tubes (10- or 50-ml, depending on volume of pellet to be extracted)
 Sonicator bath
 Nitrogen stream
 Rotary or shaker evaporator
 Silica-gel HPTLC plates
 Enhance spray (DuPont NEN; for metabolic labeling)
 Glass spray unit
 Oven adjustable up to 140°C
 Additional reagents and equipment for HPTLC (UNIT 9.7A) and for metabolic radiolabeling (UNIT 17.4) and autoradiography (APPENDIX 3A; both optional)

NOTE: All incubations should be carried out under the conditions normally used for the cell line to be analyzed.

Incubate cells with inhibitor

1. Prepare cell cultures and maintain as usual until subconfluent. Set up sufficient flasks or plates for the set of experiments using different POMP concentrations and incubation times (see steps 4 and 5).

This procedure can be performed in either the presence or absence of fetal bovine serum, and with adherent cells or (with appropriate modifications) suspension cultures.

2. Allow 10 mM PDMP in ethanol to warm up to room temperature. Transfer to a sterile screw-cap glass tube a sufficient amount of this solution to set up a series of incubations with 5 to 25 μ M final concentration of PDMP and different time points (see steps 4 and 5). Evaporate to dryness with a nitrogen stream.

3. Add culture medium and sonicate 10 min in a sonicator bath to suspend the drug.
4. Add PDMP solution to aliquots of the cell culture to obtain final concentrations of 5, 10, 15, 20, and 25 μM . Add ethanol as needed to achieve the same ethanol concentration in all samples, and add the same total volume of ethanol without PDMP to a control culture.
5. Incubate the cells for selected durations.

Periods ranging from 10 hr to 4 days have been employed. Optimum inhibition times may differ from one particular cell line to another. It is necessary to check the cells for viability as well as for morphological changes, particularly after prolonged incubations with the inhibitor.

6. Remove medium and add fresh medium containing PDMP as before. If the cells are being metabolically labeled, add the radiolabeled precursor at this point (see *UNIT 17.4*).

Prepare total lipid extract

7. Harvest cells into a 10- or 50-ml conical glass tube, pellet the cells, and wash three times with PBS.
8. Add 10 vol of 2:1 (v/v) chloroform/methanol to the cell pellet. Extract the lipids by homogenizing or sonicating 5 min. Centrifuge 10 min at $\sim 350 \times g$, 4°C . Transfer the supernatant (extract) to a clean glass tube using a Pasteur pipet.
9. Repeat the extraction as in step 8 with, successively, 1:1 (v/v) chloroform/methanol, 1:2 (v/v) chloroform/methanol, and 10:10:1 (v/v/v) chloroform/methanol/water.
10. Pool all extracts and dry them under reduced pressure, using a rotary or shaker evaporator or a nitrogen stream depending on the volume of extract to be dried.

Analyze extract by HPTLC and (optionally) autoradiography

11. Dissolve the pooled dry extract (total lipid extract) in a small volume of 2:1 (v/v) chloroform/methanol and spot one aliquot on silica-gel HPTLC plates (*UNIT 9.7A*). Spot a mixture of ganglioside and/or neutral glycolipid standards on one side of the plate. Perform HPTLC analysis. Allow the plates to dry completely after developing.

The amount of total lipid required for positive resulting in observable bands (~ 1 nmol/band) will vary between cell lines because of different compositions and ratios of components. It may be necessary to do several trials until adequate results are obtained. When analyzing radiolabeled extracts, it is advisable to spot the same number of total counts in each lane.

When analyzing nonradiolabeled glycolipids, and if sufficient material is available, two different HPTLC plates can be spotted in parallel with the same total lipid extracts. Staining one with orcinol/sulfuric acid and the other with resorcinol/hydrochloric acid will allow determination of which bands correspond to neutral glycolipids and which to gangliosides.

12. When analyzing nonradiolabeled glycolipids, proceed directly to step 13. When analyzing metabolically radiolabeled glycolipids, visualize them by autoradiography (*APPENDIX 3A*) using Enhance scintillant spray, then proceed to step 13 to visualize the nonradiolabeled standards.
- 13a. Using a glass spray unit, spray plates with orcinol/sulfuric acid reagent. Heat in a 100°C oven until bands are observed (~ 10 min, but check the plate often to avoid overstaining).

This step detects all glycolipids, which give brown bands.

In steps 13a and 13b, it is possible to stain only the side of the plate containing the standards by covering the rest of the plate with a piece of glass held in position with binder clips.

- 13b. Alternatively, spray the plate with resorcinol/hydrochloric acid reagent and heat 15 min in a 140°C oven.

This step detects only gangliosides, which give blue bands.

14. Compare the pattern of glycolipids obtained for the control cells with those obtained with different concentrations of inhibitor after different incubation periods.
15. If cells were metabolically radiolabeled, quantitate the individual components by scraping off the corresponding areas of silica gel and measuring radioactivity by liquid scintillation counting. If a scanning densitometer is available, quantitation can be achieved by scanning the autoradiograph or the stained glass plate.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

PDMP stock solution, 10 mM in ethanol

Dissolve 3.86 mg *d,l-threo*-1-phenyl-2-decanoylamino-3-morpholino-1-propanol-HCl (PDMP; FW 390.6; available from Matreya) in 1.0 ml ethanol. Store at -20°C until used.

PDMP is quite stable in aqueous solvents. No breakdown was observed after incubating a 1 mM solution in PBS (pH 7.4) for 10 days at 41°C. Stability in ethanol is even higher.

Orcinol/sulfuric acid reagent

Dissolve 200 mg orcinol in 33 ml water. Dilute to 100 ml with concentrated sulfuric acid. Store in dark bottle at 4°C. Discard when the solution turns brownish (8 to 10 days).

Orcinol from Sigma (FW 142.2) is suitable for making this reagent.

Resorcinol/hydrochloric acid reagent

80 ml concentrated HCl

10 ml 2% resorcinol (1,3-benzenediol) in H₂O

250 μl 0.1 M copper sulfate in H₂O (2.5 mM final concentration)

Store in a dark bottle at 4°C

Discard if the solution turns brown

Resorcinol from Sigma (crystalline; ~98% pure; FW 110.1) is suitable for making this reagent.

COMMENTARY

Background Information

Many studies have shown that glycosphingolipids and gangliosides are involved in a great variety of biological phenomena. As constituents of mammalian cell membranes, gangliosides significantly affect cell-surface and transmembrane signalling events such as cellular immune responses, cell growth, and differentiation. Other observed effects of gangliosides include those of “shed” gangliosides in malignancy; the effects of gangliosides and their products on protein kinase C; calcium

binding; and their interaction with various bioeffectors, including bacterial toxins, bacterial binding proteins, interferons, serotonin, fibronectin, plasma low-density lipoproteins, lectins, and glycoprotein hormones.

Many studies of the role of these molecules have been carried out by adding specific exogenous glycosphingolipids or gangliosides, or mixtures thereof, to cultured cells. Interpretation of these studies is difficult for various reasons. True incorporation of the added molecules into the cell membrane is hard to prove,

and the possibility of abnormal effects caused by micelles associated with the cell is hard to rule out. Also, exogenous material may displace the controlling factor from its normal location in the cell membrane or influence its metabolism.

Another approach to clarifying the biological roles of glycosphingolipids and gangliosides is the use of enzyme inhibitors to study the effects of their depletion in cultured cells and animals. This inhibition can be achieved using a synthetic analog of ceramide, PDMP. This method was developed by Radin and Inokuchi, whose idea that cancer results from microsphingolipidosis led them to explore ways to change the biological processes in which glycolipids are involved. One such process is the synthesis of glycolipids from nonglycolipid precursors and lower glycolipids. Most neutral glycolipids and gangliosides have a common precursor, GlcCer, which is formed enzymatically from ceramide and UDP-Glc. The synthetase responsible for this transfer is UDP-glucose:*N*-acylsphingosine glycosyltransferase (EC 2.4.1.80). Because rapid liver growth was observed in mice when glucosylceramide accumulated (Datta and Radin, 1986), the authors explored the possibility that inhibiting the synthesis of the common precursor GlcCer would decrease the amount of glycolipids due to normal turnover (Hospattankar and Radin, 1982). After a few attempts, they obtained an active inhibitor, PDMP (*D*-*threo*-1-phenyl-2-decanoylamino-3-morpholino-1-propanol; Inokuchi and Radin, 1987). This compound was shown to inhibit ganglioside biosynthesis effectively in intact cells, without any major toxic effects (Inokuchi and Radin, 1987; Inokuchi et al., 1987). The effects of the enzyme inhibitor were counteracted by including GlcCer in the culture medium (Inokuchi et al., 1989). For a review of the available data, see Radin and Inokuchi (1991).

Critical Parameters

It is critical to understand that very little is known about the regulation of glycosphingolipid biosynthesis, turnover, or even which steps are regulation points. On the other hand, the composition of glycosphingolipids varies tremendously between different types of cells. The time required to obtain optimum inhibition of any glycolipid or collection of glycolipids depends on how completely the enzyme is inhibited, the activity of the specific hydrolases involved, the susceptibility of each glycolipid to attack by its hydrolase, and the rates of the

synthases that recycle smaller glycolipids into larger ones. Although the effects on simple gangliosides such as GM₃ are readily observed, complex molecules such as GD1b require more time to become affected. Complete inhibition, as well as complete recovery from the inhibitor, should be expected to take at least 24 hr.

In consequence, the conditions required for inhibition and recovery of each particular system need to be carefully established. Various concentrations of inhibitor should be tried to establish the minimum amount required for inhibition. It is also important to understand that an inhibitor may have pleiotropic effects, including the accumulation of precursors (e.g., ceramide), expansion of alternate pathways (e.g., sphingomyelin synthesis), and other non-specific effects (e.g., on protein synthesis). Depending on the experiment, such alternate explanations for the observed effects may need to be explored.

Anticipated Results

The protocol allows optimization of the conditions for maximum inhibition of glycolipid biosynthesis with minimum effects on the morphology and viability of cells. Once these conditions are established, different studies of the biological roles of glycosphingolipids and gangliosides can be performed.

Time Considerations

Complete inhibition of glycolipid biosynthesis may require up to 4 days of cell culture. Further extraction of glycolipids and analysis of the resulting data are expected to take a maximum of 2 days, including quantitation. When metabolic labeling and autoradiography are included, extra time is required for detection, and will depend on the total counts available and their distribution among components.

Literature Cited

- Datta, S.C. and Radin, N.S. 1986. Glucosylceramide and level of glucosidase-stimulating proteins. *Lipids* 21:702-709.
- Hospattankar, A.V. and Radin, N.S. 1982. Changes in liver lipids after administration of 2-decanoylamino-3-morpholinopropiophenone and chlorpromazine. *Lipids* 17:538-543.
- Inokuchi, J.I. and Radin, N.S. 1987. Preparation of the active isomer of 1-phenyl-2-decanoylamino-3-morpholino-1-propanol, inhibitor of murine glucocerebrosidase synthetase. *J. Lipid Res.* 28:565-571.
- Inokuchi, J.I., Mason, I., and Radin, N.S. 1987. Antitumor activity via inhibition of glycosphingolipid biosynthesis. *Cancer Ltrs.* 38:23-30.

Inokuchi, J., Momosaki, K., Shimeno, H., Nagamatsu, A., and Radin N.S. 1989. Effects of D-threo-PDMP, an inhibitor of glucosylceramide synthetase, on expression of cell surface glycolipid antigen and binding to adhesive proteins by B16 melanoma cells. *J. Cell. Physiol.* 141:573-583.

Radin, N.S. and Inokuchi, J. 1991. Use of PDMP for the study of glycosphingolipid functions. *Trends Glycos. Glycotech.* 3:200-213.

Excellent review, providing very useful references for comparing conditions used for different systems along with specific protocols for studying the different biological roles affected by inhibition of glycolipid biosynthesis.

Contributed by Adriana E. Manzi
University of California San Diego
La Jolla, California

Key Reference

Radin and Inokuchi, 1991. See above.

Synthetic Glycosides as Primers of Oligosaccharide Biosynthesis and Inhibitors of Glycoprotein and Proteoglycan Assembly

UNIT 17.11

With the exception of hyaluronic acid, all mammalian saccharides assemble while attached to a lipid or protein primer. Several cases are now known in which oligosaccharide synthesis will occur on synthetic glycoside primers added to cells. β -D-xylosides initiate glycosaminoglycan (GAG) synthesis (Okayama et al., 1973) by substituting for endogenous xylosylated core proteins (Basic Protocol 1). At high concentration xylosides will also prime oligosaccharides that resemble glycolipids (Freeze et al., 1993). *N*-acetyl- α -D-galactosaminides initiate the synthesis of O-linked oligosaccharides found on mucins and other glycoproteins (Basic Protocol 2) in an analogous manner. Even disaccharides, such as peracetylated *N*-acetylglucosaminide, can act as primers (Sarkar et al., 1995). Because these primers compete with endogenous substrates, they also act as inhibitors of proteoglycan (PG) and glycoprotein synthesis. Thus, primers have utility for studying the biological activity of glycoconjugates in cells, tissues, and animals. This unit describes procedures for using glycoside primers in cell culture.

XYLOSIDE INITIATION OF GAG SYNTHESIS AND INHIBITION OF PG ASSEMBLY

BASIC
PROTOCOL 1

Materials

- 0.2 M *p*-nitrophenyl- β -D-xyloside in DMSO (see recipe)
- Tissue culture growth medium appropriate for cell line
- Cultured cells, in suspension or adherent
- Dimethylsulfoxide (DMSO, C₂H₆SO; mol. wt. 78.13)
- Radioactive precursors: H₂³⁵SO₄, [6-³H]GlcNH₂, or [1-³H]Gal

Additional reagents and equipment for metabolic radiolabeling (UNIT 17.4), isolating proteoglycans and glycosaminoglycans (UNIT 17.3), and gel filtration (UNIT 17.20) and reversed-phase chromatography (UNIT 17.21)

NOTE: All incubations should be carried out in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

1. Add 0.2 M *p*-nitrophenyl- β -D-xyloside in DMSO to aliquots of growth medium to achieve final concentrations of 0, 0.01, 0.03, 0.1, 0.3, and 1 mM. Add DMSO as needed to achieve the same concentration of the solvent in all samples.

Alternatively, dissolve enough inhibitor directly in growth medium to achieve a final concentration of 10 mM, then prepare the indicated serial dilution series with fresh growth medium.

DMSO can cause some cells to differentiate, which can change the pattern of glycosylation.

2. Add each supplemented growth medium to a previously established cell culture by replacing the spent growth medium with the supplemented medium. Incubate 1 hr.

Do not add concentrated stocks of glycosides directly to cell cultures, as high local concentrations of solvent and glycoside can kill cells.

3. Add radioactive precursors to each culture (generally 10 μ Ci/ml ³⁵SO₄²⁻, 20 μ Ci/ml [6-³H]GlcNH₂, or 50 μ Ci/ml [1-³H]Gal should be sufficient; see UNIT 17.4). Continue to incubate cells \geq 1 hr.

Preparation and
Analysis of
Glycoconjugates

17.11.1

4. Isolate the PGs and GAGs made by the cells and the oligosaccharides primed on the glycoside from the growth medium (UNIT 17.3).
5. Analyze the material by gel filtration and reversed-phase chromatography (see UNITS 17.20 & 17.21).

Oligosaccharides generated on primers are generally smaller than glycoproteins and proteoglycans.

N-ACETYL- α -D-GALACTOSAMINIDE INITIATION OF OLIGOSACCHARIDE SYNTHESIS AND INHIBITION OF O-LINKED GLYCOPROTEIN ASSEMBLY

Additional Materials (also see Basic Protocol 1)

0.2 M *p*-nitrophenyl-*N*-acetyl- α -D-galactosaminide in DMSO (see recipe)

NOTE: All incubations should be carried out in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

1. Add 0.2 M *p*-nitrophenyl-*N*-acetyl- α -D-galactosaminide in DMSO to aliquots of growth medium to achieve final concentrations of 0, 1, 2, 5, and 10 mM. Add DMSO as needed to achieve the same concentration of the solvent in all samples.

Alternatively, dissolve enough inhibitor directly in growth medium to achieve a final concentration of 10 mM, then prepare the indicated serial dilution series with fresh growth medium.

2. Add each supplemented growth medium to a previously established cell culture. Incubate 1 hr.

Do not add concentrated stocks of glycosides directly to cell cultures, as high local concentrations of solvent and glycoside can kill the cells.

3. Add radioactive precursors to cultures (generally 20 μ Ci/ml [6-³H]GlcNH₂ or 50 μ Ci/ml [1-³H]Gal should be sufficient; see UNIT 17.4). Incubate cells a further 3 hr.
4. Harvest the glycoproteins from the cells and growth medium (UNIT 17.3).
5. Analyze the material by gel filtration and reversed-phase chromatography (UNITS 17.20 & 17.21). Small oligosaccharides linked to hydrophobic aglycones will bind to C18 reversed-phase resins, aiding the separation of primed material from endogenous glycoproteins.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

***p*-nitrophenyl-*N*-acetyl- α -D-galactosaminide, 0.2 M in DMSO**

Dissolve 68.5 mg *p*-nitrophenyl-*N*-acetyl- α -D-galactosaminide (C₁₄H₁₈N₂O₈; mol. wt. 342.3) in 1 ml dimethylsulfoxide. Warm if necessary to dissolve.

***p*-nitrophenyl- β -D-xyloside, 0.2 M in DMSO**

Dissolve 54.2 mg *p*-nitrophenyl- β -D-xyloside (C₁₁H₁₃NO₇; mol. wt. 271.2) in 1 ml dimethylsulfoxide. Warm if necessary to dissolve.

Other β -D-xylosides are also commercially available, including the fluorescent compounds naphthol- β -D-xyloside and 4-methylumbelliferyl- β -D-xyloside. Using these latter compounds, the fluorescence of the aglycone can be used to monitor the recovery of oligosaccharides primed on the glycoside.

COMMENTARY

Background Information

Proteoglycan (PG) biosynthesis is initiated by the transfer of D-xylose from UDP-xylose to specific serine residues in core proteins. This intermediate acts as a natural primer for the assembly of heparan sulfate, heparin, chondroitin sulfate, and dermatan sulfate chains, depending on the tissue. In 1973, Okayama et al. (1973) reported that synthetic β -D-xylosides would prime glycosaminoglycan (GAG) synthesis, apparently by substituting for xylosylated core proteins. Since then, a variety of β -D-xylosides have been studied as primers with the following results. (1) Priming requires the β -anomer of D-xylose (Galligani et al., 1975). Other pentosides (β -D-lyxoside, β -D-ribose, and α -L-arabinoside) and hexosides of the linkage region generally lack activity (Robinson and Robinson, 1985; F.N. Lugemwa, A.K. Sarkar, and J.D. Esko, unpub. observ.). (2) Priming activity correlates with hydrophobicity of the aglycone (Robinson et al., 1975; Robinson and Robinson, 1981; Sobue et al., 1987; Kolset et al., 1990; Lugemwa and Esko, 1991). Thus, methyl- β -D-xyloside is a weak primer compared to *p*-nitrophenyl- β -D-xyloside or 4-methylumbelliferyl- β -D-xyloside. (3) The most active primers contain O or S in the glycosidic linkage, but *N*-xylosides and homo-*C*-xylosides also have activity (Sobue et al., 1987). (4) Priming is dose dependent (Esko et al., 1987; Lugemwa and Esko, 1991). (5) β -D-xylosides prime GAGs in most cells and can cause more GAG to be made than normally occurs on natural core proteins. (6) Cells secrete the majority of material generated on β -D-xylosides into the growth medium. (7) β -D-xylosides prime chondroitin sulfate or dermatan sulfate in virtually all cells, but priming of heparan sulfate occurs poorly, except with appropriate aglycones (Lugemwa and Esko, 1991; Fritz et al., 1994). (8) The fine structure of the chains depends on the concentration of the primer and may not precisely duplicate the structure of chains on natural core protein substrates (Fransson et al., 1992; Fritz et al., 1994). (9) β -D-xylosides compete with endogenous xylosylated core proteins for precursors and the various biosynthetic enzymes. Thus, β -D-xylosides act as inhibitors of PG synthesis. The inhibitory activity of xylosides has been exploited to examine the biological activity of PGs in cells.

The enormous success of β -D-xylosides in altering PG synthesis suggested that other gly-

cosides might act as artificial primers. Kuan et al. (1989) showed that aryl-*N*-acetyl- α -D-galactosaminides inhibited mucin assembly in colon carcinoma cells. Benzyl- α -D-galactosaminide inhibits O-linked oligosaccharide synthesis on cell-surface glycoproteins as well, which can have profound effects on cell adhesion (Kojima et al., 1992). Zhuang et al. (1991) showed that α -D-galactosaminides prime the synthesis of O-linked oligosaccharides. Thus, the inhibition of glycoprotein assembly is most likely due to competition between endogenous core protein substrates and the added glycoside. Recently, Neville et al. (1995) showed that *N*-acetylglucosaminides will prime poly-lactosamines. Sarkar et al. (1995) extended these studies by showing that more complex glycosides consisting of two sugar residues will prime oligosaccharides and inhibit the formation of glycoproteins involved in cell adhesion. These latter compounds are not commercially available, but they can be made relatively easily and inexpensively. The activity of disaccharide-based primers creates the opportunity to customize compounds that act at later points in glycoprotein synthesis.

Critical Parameters

Glycosides containing hydrophobic aglycones are moderately soluble in aqueous solutions. Stock solutions can be prepared in DMSO or ethanol and appropriate dilutions made into growth medium. β -D-xylosides are stable in both organic and aqueous solvents. It is important to add glycosides to growth medium before adding them to cells, because the concentrated solvent can cause cells to lyse. To avoid this problem, the glycosides can be dissolved directly into growth medium at a final concentration of ≤ 10 mM. A serial dilution series prepared in fresh growth medium can then be added to previously established cell cultures. This method minimizes the use of solvent and provides more accurate control over the concentration of glycoside.

Priming of oligosaccharides occurs in a concentration-dependent manner. Priming efficiency is easily assessed by adding a radioactive precursor, such as [6 - 3 H]GlcNH $_2$ or [1 - 3 H]Gal. 35 SO $_4^{2-}$ can be used to measure GAG synthesis. After brief incubation, the mixtures of radioactive oligosaccharides can be isolated from the growth medium, because cells secrete the majority of material generated on primers (UNITS 17.1 & 17.3). Usually this will require separation

of endogenous glycoproteins and PGs from the primed oligosaccharides and GAGs, e.g., by gel filtration or reversed-phase chromatography (UNITS 17.20 & 17.21).

In general, GAG biosynthesis peaks in most cells when the β -D-xyloside concentration reaches 10 to 100 μ M. However, this value varies in different cells. Thus, it is important to vary the concentration from 1 μ M to 1 mM and assess the minimal concentration needed for maximally priming GAGs. The priming of glycolipid-like compounds on β -D-xylosides requires relatively higher concentrations of primer (Freeze et al., 1993). Similarly, the priming of O-linked oligosaccharides by *N*-acetyl- α -D-galactosaminides depends on concentration, but the effective dose is considerably higher (1 to 10 mM) than that observed for β -D-xylosides (Kuan et al., 1989). In contrast, priming by modified disaccharides occurs in the 10 to 100 μ M range (Sarkar et al., 1995). These differences may relate to the relative abundance of endogenous substrates, solubility of the glycoside, its uptake through plasma membranes into the Golgi, its relative affinity for the glycosyltransferases, enzyme concentration and composition, or its susceptibility to hydrolysis.

Excessive concentrations of glycosides can be deleterious (Kanwar et al., 1986; Lugemwa and Esko, 1991). All glycosides exhibit detergent properties, because the molecules have both a hydrophilic end (the glucose unit) and a hydrophobic end (the aglycone). To test for toxic effects, one can employ glycosides with the wrong anomeric stereochemistry (e.g., α -D-xylosides; Farach et al., 1988), other pentosides (e.g., *p*-nitrophenyl- α -L-arabinoside), or the aglycone alone (e.g., *p*-nitrophenol).

Specificity exists with respect to the composition of GAG chains generated on β -D-xylosides (Lugemwa and Esko, 1991). Most xylosides (e.g., those containing linear aliphatic chains, single aromatic rings, and saturated multicyclic ring systems) tend to prime chondroitin sulfate and dermatan sulfate. The priming of chondroitin sulfate does not seem to depend on the aglycone: D-xylose and methyl- β -D-xyloside will initiate chondroitin sulfate (Okayama et al., 1973; Schwartz et al., 1974; Galligani et al., 1975; Lugemwa and Esko, 1991). Recent studies show that priming of heparan sulfate depends on aglycone structure (Lugemwa and Esko, 1991; Fritz et al., 1994). For example, naphthol- β -D-xyloside and estradiol- β -D-xyloside will prime heparan sulfate in Chinese hamster ovary cells. At low concentra-

tions, these primers initiate mostly chondroitin sulfate. As the concentration increases beyond 10 μ M, the proportion of heparan sulfate increases. Similar dose-dependent effects have been observed on the priming of different O-linked oligosaccharides by *N*-acetyl- α -D-galactosaminides (Zhuang et al., 1991; Kuan et al., 1989). The partitioning of intermediates into different biosynthetic pathways may reflect molecular recognition between the aglycone and specific enzymes.

Feeding cells glycosides provides an easy method to determine if a relationship exists between glycosylation and a particular biological activity. Glycoside primers subvert the cellular machinery for making oligosaccharides from the assembly of chains on endogenous core protein primers. Thus, glycosides inhibit the glycosylation of glycoproteins and PGs. Inhibition occurs in a dose-dependent fashion, but the inhibition is rarely complete, probably due to the inability of glycosides to mimic the endogenous core protein substrates. It is important to remember the following points. (1) Glycosides inhibit the formation of entire classes of glycoconjugates. The effects on individual proteins and PGs may vary. (2) Inhibition occurs in the midst of priming. The primed material may also have a different composition than the oligosaccharides produced on core proteins. Thus, any effect on a biological phenomenon may be due to the decline in glycoprotein or PG assembly or the increase in free oligosaccharide chains. (3) One should always consider the possibility that the enhanced rate of synthesis of oligosaccharides on primers may deplete pools of nucleotide sugars and thus alter the assembly of other types of glycoconjugates. (4) Primers may not always be as specific as once thought. For example, Freeze and coworkers have shown that β -D-xylosides will prime compounds resembling ganglioside GM₃ and an unusual pentasaccharide containing a terminal α -linked GalNAc (Freeze et al., 1993; Salimath et al., 1995).

Anticipated Results

The addition of glycosides to cultured cells primes the synthesis of free oligosaccharide chains and inhibits the formation of mature glycoconjugates. Priming and inhibition occur in a dose-dependent manner, and the appropriate concentration required for optimal priming and inhibition must be determined empirically. Figure 17.11.1 shows gel filtration profiles of GAG and PG synthesized by a murine mastocytoma cell line (Montgomery et al.,

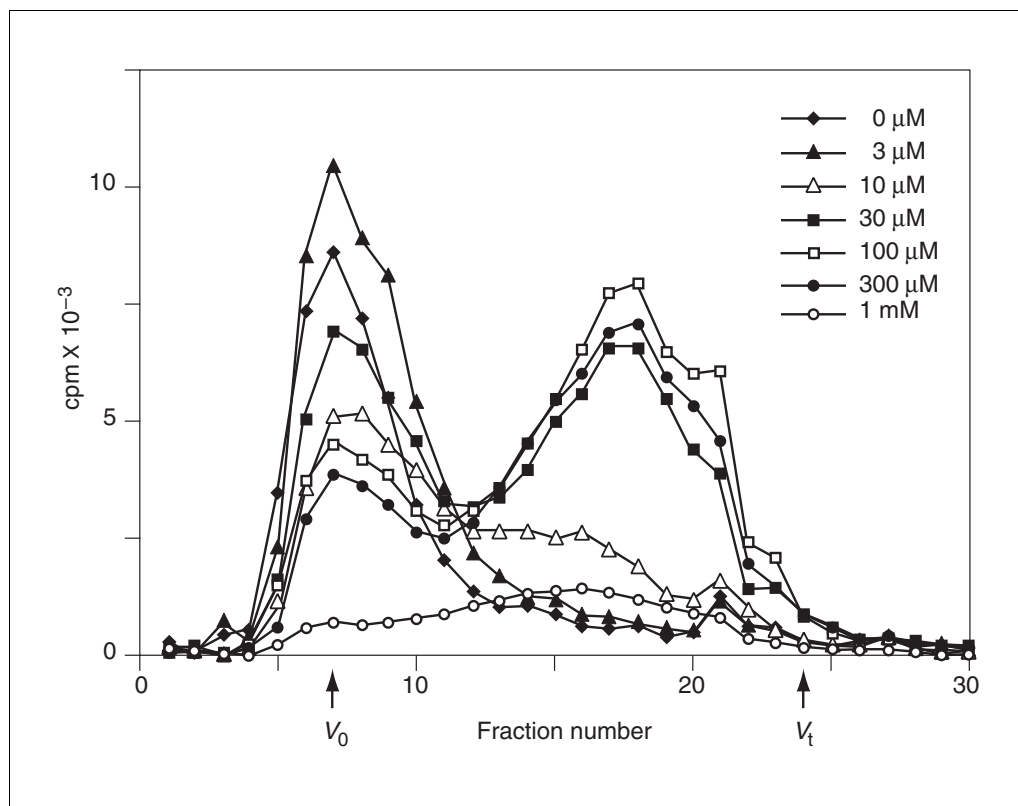


Figure 17.11.1 Gel filtration chromatography of glycosaminoglycans (GAGs) and proteoglycans (PGs) generated in the presence of naphthol- β -D-xyloside. Mastocytoma cells (Montgomery et al., 1992) were incubated 4 hr with the indicated concentration of naphthol- β -D-xyloside and 20 μ Ci/ml ^{35}S SO_4^{2-} . The [^{35}S]GAGs and [^{35}S]PGs in the cells and growth medium were solubilized in Triton buffer and fractionated over a small DEAE-Sephacel column as described in UNIT 17.5. An aliquot was analyzed by gel filtration HPLC (TSK G4000 SW, 30 cm \times 7.5 mm i.d., Pharmacia Biotech). Samples were eluted with 0.5 M NaCl/0.1 M KH_2PO_4 (pH 6.0)/0.2% (w/v) Zwittergent 3-12 (Calbiochem) at a flow rate of 0.5 ml/min.

1992). As the concentration of β -D-xyloside rises, the amount of free GAG chains increases, as measured by material eluting in more retarded fractions. The size of these chains decreases as the primer concentration continues to rise. Presumably, this reflects the enhanced initiation of GAG chains and the competition of primers with endogenous xylosylated core proteins for limited precursors. The amount of mature PG decreases further as primers divert more resources away from the formation of GAG chains on core proteins. *N*-acetyl- α -D-galactosaminides and acetylated disaccharides have similar effects on the assembly of O-linked oligosaccharides on glycoproteins (Kojima et al., 1992; Sarkar et al., 1995). The relatively small oligosaccharides generated on these primers can be purified by reversed-phase and gel filtration chromatography (Zhuang et al., 1991).

Literature Cited

- Esko, J.D., Weinke, J.L., Taylor, W.H., Ekborg, G., Rodén, L., Anantharamaiah, G., and Gawish, A. 1987. Inhibition of chondroitin sulfate and heparan sulfate synthesis in a Chinese hamster ovary cell mutant defective in Galactosyltransferase I. *J. Biol. Chem.* 262:12189-12195.
- Farach, M.C., Tang, J.P., Decker, G.L., and Carson, D.D. 1988. Differential effects of *p*-nitrophenyl-D-xylosides on mouse blastocysts and uterine epithelial cells. *Biol. Reprod.* 39:443-455.
- Fransson, L.-A., Havsmark, B., Sakurai, K., and Suzuki, S. 1992. Sequence analysis of *p*-hydroxyphenyl-*O*- β -D-xyloside initiated and radioiodinated dermatan sulfate from skin fibroblasts. *Glycoconj. J.* 9:45-55.
- Freeze, H.H., Sampath, D., and Varki, A. 1993. α - and β -xylosides alter glycolipid synthesis in human melanoma and Chinese hamster ovary cells. *J. Biol. Chem.* 268:1618-1627.
- Fritz, T., Lagemwa, F.N., Sarkar, A.K., and Esko, J.D. 1994. Priming of heparan sulfate on β -D-xylosides depends on aglycone structure. *J. Biol. Chem.* 269:300-308.

- Galligani, L., Hopwood, J., Schwartz, N.B., and Dorfman, A. 1975. Stimulation of synthesis of free chondroitin sulfate chains by β -D-xylosides in cultured cells. *J. Biol. Chem.* 250:5400-5406.
- Kanwar, Y.S., Rosenzweig, L.J., and Jakubowski, M.L. 1986. Xylosylated-PG-induced Golgi alterations. *Proc. Natl. Acad. Sci. U.S.A.* 83:6499-6503.
- Kojima, N., Handa, K., Newman, W., and Hakomori, S.-I. 1992. Inhibition of selectin dependent tumor cell adhesion to endothelial cells and platelets by blocking O-glycosylation of these cells. *Biochem. Biophys. Res. Comm.* 182:1288-1295.
- Kolset, S.O., Sakurai, K., Ivhed, I., Overvatn, A., and Suzuki, S. 1990. The effect of β -D-xylosides on the proliferation and PG biosynthesis of monoblastic U-937 cells. *Biochem. J.* 265:637-645.
- Kuan, S.-F., Byrd, J.C., Busbaum, C., and Kim, Y.S. 1989. Inhibition of mucin glycosylation by aryl-*N*-acetyl- α -galactosaminides in human colon cancer cells. *J. Biol. Chem.* 264:19271-19277.
- Lugemwa, F.N. and Esko, J.D. 1991. Estradiol β -D-xyloside, an efficient primer for heparan sulfate biosynthesis. *J. Biol. Chem.* 266:6674-6677.
- Montgomery, R.I., Lidholt, K., Flay, N.W., Liang, J., Vertel, B., Lindahl, U., and Esko, J.D. 1992. Stable heparin-producing cell lines from the Furth murine mastocytoma. *Proc. Natl. Acad. Sci. U.S.A.* 89:11327-11331.
- Neville, D.C.A., Field, R.A., and Ferguson, M.A.J. 1995. Hydrophobic glycosides of *N*-acetylglucosamine can act as primers for polyactosamine synthesis and can affect glycolipid synthesis in vivo. *Biochem. J.* 307:791-797.
- Okayama, M., Kimata, K., and Suzuki, S. 1973. The influence of *p*-nitrophenyl- β -D-xyloside on the synthesis of proteochondroitin sulfate by slices of embryonic chick cartilage. *J. Biochem.* 74:1069-1073.
- Robinson, H.C., Brett, M.J., Tralagga, P.J., Lowther, D.A., and Okayama, M. 1975. The effect of D-xylose, β -D-xylosides and β -D-galactosides on chondroitin sulphate biosynthesis in embryonic chicken cartilage. *Biochem. J.* 148:25-34.
- Robinson, J.A. and Robinson, H.C. 1981. Control of chondroitin sulphate biosynthesis. β -D-xylopyranosides as substrates for UDP-galactose: D-xylose transferase from embryonic-chicken cartilage. *Biochem. J.* 194:839-846.
- Robinson, J.A. and Robinson, H.C. 1985. Initiation of chondroitin sulphate synthesis by β -D-galactosides. Substrates for galactosyltransferase II. *Biochem. J.* 227:805-814.
- Salimath, P.V., Spiro, R.C., and Freeze, H.H. 1995. Identification of a novel glycosaminoglycan core-like molecule II. α -GalNAc-capped xylosides can be made by many cell types. *J. Biol. Chem.* 270:9164-9168.
- Sarkar, A.K., Fritz, T.A., Taylor, W.H., and Esko, J.D. 1995. Disaccharide uptake and priming in animal cells: Inhibition of sialyl Lewis X by acetylated Gal β 1 \rightarrow 4GlcNAc β -O-naphthalene-methanol. *Proc. Natl. Acad. Sci. U.S.A.* In press.
- Schwartz, N.B., Galligani, L., Ho, P.L., and Dorfman, A. 1974. Stimulation of synthesis of free chondroitin sulfate chains by β -D-xylosides in cultured cells. *Proc. Natl. Acad. Sci. U.S.A.* 71:4047-4051.
- Sobue, M., Habuchi, H., Ito, K., Yonekura, H., Oguri, K., Sakurai, K., Kamohara, S., Ueno, Y., Noyori, R., and Suzuki, S. 1987. β -D-xylosides and their analogues as artificial initiators of GAG chain synthesis. Aglycone-related variation in their effectiveness in vitro and in ovo. *Biochem. J.* 241:591-601.
- Zhuang, D., Grey, A., Harris-Brandts, M., Higgins, E., Kashem, M.A., and Dennis, J.W. 1991. Characterization of O-linked oligosaccharide biosynthesis in cultured cells using *para*-nitrophenyl- α -D-GalNAc as an acceptor. *Glycobiology* 1:425-433.

Key References

Fritz et al., 1994. See above.

Demonstrates that the type of glycoaminoglycan chain produced on a β -D-xyloside depends on the structure of the aglycone.

Kuan et al., 1989. See above.

*First description of *N*-acetyl- α -galactosaminides as primers and inhibitors of O-linked glycosylation of glycoproteins.*

Salimath et al., 1995. See above.

β -D-xylosides will also prime unusual oligosaccharides—e.g., GalNAc α 1-4GlcA β 1-3Gal β 1-3Gal β 1-4Xyl-R, GM₃-like compounds, and other oligosaccharides.

Sarkar et al., in press. See above.

Demonstrates that cells will take up suitably modified disaccharides and use them as primers, opening up the possibility of designing more complex and selective glycosylation inhibitors.

Contributed by Jeffrey D. Esko
University of Alabama at Birmingham
Birmingham, Alabama

Rebecca I. Montgomery
Northwestern University
Chicago, Illinois

RELEASE OF SACCHARIDES FROM GLYCOCONJUGATES

In many instances, it is useful to release a part or all of the oligosaccharide chains from a glycoconjugate before analyzing their structure, and the units presented in this section describe a variety of ways in which this can be achieved. Some of the methods described here (particularly the enzymatic ones) can also be applied to the study of free oligosaccharides or to oligosaccharides that are still attached to macromolecules.

Enzymes that degrade sugar chains fall into two general classes. Endoglycosidases cleave at defined sites within a sugar chain, with a specificity often based on certain features of the adjacent monosaccharide units. Thus, they are analogous to the restriction endonucleases that work on nucleic acids (*UNIT 3.1*). On the other hand, exoglycosidases specifically release single monosaccharides only when they are present as terminal units on sugar chains. Unlike the cases with exonucleases, the types of possible terminal units on oligosaccharide chains are numerous. Correspondingly, the number of available exoglycosidases with different specificities is quite large. In this section, one of the most commonly used group of exoglycosidases, the sialidases (also known as neuraminidases), is discussed in *UNIT 17.12*. Some commonly used endoglycosidases and glycoamidases for N-linked oligosaccharides (*UNIT 17.13A*) and polysaccharide lyases (*UNIT 17.13B*) are then considered.

Additional methods will be provided in future supplements (see Chapter 17 table of contents). One of the simplest and most direct ways to free an oligosaccharide chain from a glycoprotein or proteoglycan is to digest away the peptide with a broad-spectrum protease such as pronase or proteinase K (*UNIT 17.14*). In these cases, most of the peptide is removed except for the amino acids immediately surrounding the glycosylation site(s). If needed, more specific proteases (e.g., trypsin, chymotrypsin) can be used to generate glycopeptides with large stretches of amino acids still attached to the glycosylation sites. This approach is particularly useful for the detection of individual glycosylation sites on glycoproteins (*UNIT 17.14*).

There are also many classic and well-established methods for the chemical release of saccharides from glycoconjugates (*UNIT 17.15*). Some are designed to release intact sugars chains and others release the individual monosaccharide units. Although these methods tend to be somewhat less specific and potentially more destructive than the enzymatic approaches, they are cheap, convenient, and generally easy to use. These methods will be presented in future supplements.

Of course, there are many other techniques for the release of saccharides from glycoconjugates that could have been presented in this section. The current and upcoming selections were based on the criterion of broad general utility to the average molecular biologist. As methods in “glycotechnology” improve and simplify, further additions to this section may become appropriate.

Sialidases

Sialic acids are a family of nine-carbon acidic sugars found at the nonreducing terminus of many glycoconjugates. Sialidases (a term preferred to neuraminidases) can remove these sugar units selectively from cell surfaces, membranes, or purified glycoconjugates. This may be done to analyze the sialic acids or to study the consequences of their removal. The more sensitive and specific colorimetric assays for sialic acids work on free but not on glycosidically bound molecules (UNIT 17.16). Although careful acid hydrolysis also can be used to release sialic acids (UNIT 17.15), partial destruction of modifications and/or incomplete release can be problematic. Furthermore, acid hydrolysis may alter the underlying glycoconjugate or may be incompatible with functional studies.

In this unit, sialidase digestion of purified glycoproteins is described in the basic protocol and treatment of intact cells is outlined in the alternate protocol. The physical properties of the four most useful sialidases are listed in Table 17.12.1; their relative activities against sialic acids with different modifications and in different linkages are listed in Table 17.12.2 (see also critical parameters). The choice of enzyme depends upon the nature of the sample and knowledge of the type of sialylated glycoconjugates present.

BASIC PROTOCOL

SIALIDASE TREATMENT OF PURIFIED GLYCOPROTEINS

A purified glycoprotein sample is dissolved in digestion buffer and digested with a sialidase. Controls of sample alone and enzyme alone are also prepared. Digestion of gangliosides (glycolipids with sialic acid) requires detergent (e.g., deoxycholate, cholate, or taurocholate). After digestion, the reaction is terminated and desialylation is monitored.

Materials

Sialic acid-containing sample
Sialidase digestion buffer
Sialidase (Table 17.12.1)

Additional reagents and equipment for quantitating sialic acid (UNIT 17.16)

1. Dissolve the sialic acid-containing sample in sialidase digestion buffer at ~0.1 to 1 mM sialic acid concentration (final).

If the amount of sialic acid is not known, it can be estimated by acid hydrolysis (UNIT 17.15) or by quantitating the amount of sialic acid released with increasing amounts of

Table 17.12.1 General Properties of Commercially Available Sialidases

Feature	<i>Vibrio cholerae</i>	<i>Clostridium perfringens</i>	<i>Arthrobacter ureafaciens</i>	Newcastle disease virus
Form	Purified protein	Purified protein	Purified protein	Whole virion
Ca ⁺⁺ requirement	Yes	No	No	No
pH optimum	5.6 ^a	4.5 ^b /5.2 ^c	5.0 ^d	5.0–6.0
pH range	4.0–8.0	4.0–7.0	4.0–7.0	NA ^e
Maximal activity				
% at pH 5.0	>90 ^a	~95 ^b / ^{>} 90 ^c	>90 ^d	NA ^e
% at pH 7.0	>70 ^a	ND ^e /~40 ^c	~30 ^d	NA ^e

^aIn 0.1 M Tris/maleic buffer (Ada et al., 1961) and sodium acetate (Cassidy et al., 1965).

^bIn 0.1 M sodium acetate (Cassidy et al., 1965).

^cIn 0.1 M citrate-phosphate (Cassidy et al., 1965).

^dIn 0.1 M sodium acetate buffer (Uchida et al., 1979).

^eAbbreviations: NA, not available (pH profile not published, although enzyme remains active at pH 7.0); ND, not determined.

Sialidases

17.12.2

sialidase. If the amount of sialidase added is suboptimal, free sialic acid measured (by the TBA or DMB assays, UNIT 17.16) will increase with increasing amounts of sialidase.

If the sample contains gangliosides, include $1/50$ vol of 10% sodium deoxycholate (0.2% final) and place tubes in a sonicator bath for 1 to 2 min to ensure complete dispersal. If large amounts of gangliosides are present, the amount of detergent in the sample should be kept equal to or greater than the amount of ganglioside on a weight-to-weight basis. If no direct information is available concerning the amount of ganglioside present, a series of digestions with increasing amounts of detergent (e.g., 0.1%, 0.3%, 1.0%) may need to be done. Excess detergent may actually inhibit digestion. The *Arthrobacter ureafaciens* sialidase (see Table 17.12.1) is most active with gangliosides. Deoxycholate solutions may become cloudy in the presence of salts.

2. Add 1 to 20 mU sialidase to the sample and enzyme control tubes. Mix well.

Select the enzyme to be used based on the properties in Table 17.12.1. The amount of enzyme to use depends on the amount of releasable sialic acid in the sample. Initially, use ~ 1 mU/nmol of sialic acid. Further experiments may indicate that more or less enzyme is required.

3. Prepare blank tubes containing either sample alone or enzyme alone (in sialidase digestion buffer).

Alternative controls could include reactions incubated with inactivated sialidase (prepared by boiling 5 min) or active sialidase plus 1 mM 2,3-dehydro-2-deoxy-N-acetylneuraminic acid (Boehringer Mannheim), a sialidase inhibitor.

Table 17.12.2 Sialidase Action on Types of Sialic Acids^a

Sialic acid		<i>Vibrio cholerae</i>	<i>Clostridium perfringens</i>	<i>Arthrobacter ureafaciens</i>	Newcastle disease virus
Type	Linkage(s)	Relative susceptibility			
Neu5Ac	$\alpha 2-3$	++++	++++	++	++++
	$\alpha 2-6$	+++	+++	++++	R
	$\alpha 2-8$	++	++	+++	+++
Neu5Gc	$\alpha 2-3$	++++	++++	++	++
	$\alpha 2-6$	+++	+++	++++	R
	$\alpha 2-8$	++	++	+++	++
7(9) mono- <i>O</i> -acetyl	$\alpha 2-3$	++	++++	++	++++
	$\alpha 2-6$	+	+++	++++	R
	$\alpha 2-8$	+	++	+++	+++
4-mono- <i>O</i> -acetyl	$\alpha 2-3$	R	R	R	R
	$\alpha 2-6$	R	R	R	R
	$\alpha 2-8$	R	R	R	R
7(8)9 di- <i>O</i> -acetyl	$\alpha 2-3$	R?	R?	++	?
	$\alpha 2-6$	R?	R?	++++	R
	$\alpha 2-8$	R?	R?	+++	?
7,8,9 tri- <i>O</i> -acetyl	$\alpha 2-3$	R?	R?	++?	?
	$\alpha 2-6$	R?	R?	++?	R
	$\alpha 2-8$	R?	R?	++?	?
Periodate-oxidized	$\alpha 2-3$	+	++	R?	?
	$\alpha 2-6$	+	++	R?	?
Neu5Ac-7(8)	$\alpha 2-8$?	?	R?	?
<i>O</i> -methyl,	$\alpha 2-3$?	?	?	?
<i>O</i> -sulfate,	$\alpha 2-6$?	?	?	?
combinations	$\alpha 2-8$?	?	?	?

^aAbbreviations: Neu5Ac, *N*-acetylneuraminic acid; Neu5Gc, *N*-glycolylneuraminic acid; R, practically resistant under typical digestion conditions; ?, not known. See Varki (1992) for additional details of terminology.

Table 17.12.3 Sialic Acid Standards

Glycoconjugate	Molecule	Linkages	Sialic acids ^a	Sialic acid concentration (nmol/mg)	Stock solution preparation ^b
Sialylactose (mixed isomers ^c)	Oligosaccharide	α 2-3 α 2-6	Neu5Ac	1600	31 mg/ml in water (100 \times)
Fetuin ^d	Glycoprotein (N- and O-linked)	α 2-3 α 2-6	Neu5Ac Neu5Gc(?)	280	20 mg/ml in water (10 \times)
Bovine submaxillary mucin	Mucin (O-linked)	α 2-6	mono- and di- <i>O</i> -acetylated Neu5Ac and Neu5Gc	150-400	20 mg/ml in water (10 \times)
Colominic acid	Polysaccharide	α 2-8	Neu5Ac	3400	15 mg/ml in water (100 \times)
Mixed brain gangliosides	Glycolipid	α 2-3 α 2-8	Neu5Ac ^f	950 ^e	5 mg/ml in 2:1:0.1 (v/v/v) chloroform/methanol/water (10 \times)

^aAbbreviations: Neu5Ac, *N*-acetylneuraminic acid; Neu5Gc, *N*-glycolylneuraminic acid.

^bStore all stock solutions at -20°C . Dilution to 1 \times will yield ~ 0.5 mM sialic acid.

^cRatio of α (2-3) to α (2-6) $\sim 4:1$, although lot-to-lot variation probably occurs.

^dRatio of α (2-3) to α (2-6) $\sim 1.6:1.4$ in commercial preparations, including Sigma (#F-3004) or GIBCO/BRL (Townsend et al., 1989).

^eDepending on the manufacturer.

^fCommercial preparations of bovine brain gangliosides have been treated with base during purification, resulting in loss of *O*-acetyl esters.

4. Incubate 3 to 4 hr at 37°C .

Digestion of gangliosides may require 24 to 48 hr.

For selective digestion of α (2-3) and α (2-8) linkages by the Newcastle Disease virus sialidase, use 1 to 5 mU in a 25 to 50 μl reaction and incubate 15 to 30 min at 37°C . These conditions are appropriate when the enzyme is present in excess of substrate, as is often the case when radiochemical amounts of substrate are present. The enzyme has a low level of activity against α (2-6) linkages; this becomes evident with prolonged incubations. If large amounts of sialic acids are present (e.g., >1 nmol), it will be necessary to titrate the enzyme. This can be done by setting up digests of standards containing α (2-3) or α (2-6)-linked sialic acids (Table 17.12.3) at the same concentration as the unknown sample and determining the amount of Newcastle disease virus sialidase that releases $>90\%$ of the α (2-3)-linked and $<5\%$ of the α (2-6)-linked sugar.

5. Terminate the reaction by boiling for 5 min. Check for release of sialic acid by monitoring:

- Sialic acid by direct measurement (UNIT 17.16)
- Shift in pI of a glycoprotein by isoelectric focusing (UNIT 10.4)
- Shift in TLC pattern for glycolipids (UNIT 17.3)
- Shift in apparent molecular weight on SDS-PAGE (UNIT 10.2; generally a loss of several sialic acid residues will be visible by small shifts on SDS-PAGE to either lower or higher apparent molecular weights)
- Alteration in lectin-binding patterns for some lectins (UNIT 17.7)
- Shift in negative charge of isolated oligosaccharides or glycolipids (UNIT 17.17).

Sialidases

17.12.4

SIALIDASE TREATMENT OF INTACT CELLS

Although the pH optima of bacterial and viral sialidases range from 4.5 to 5.5, these enzymes can be used to treat intact, viable cells at pH 7.0. Cells are washed and resuspended in isotonic serum-free buffer, then treated with the sialidase. The reaction is terminated by centrifugation and washing.

Additional Materials

Cells, prepared as a single-cell suspension
HEPES-buffered saline (HeBS)
Culture medium appropriate for cells, with serum

1. Wash cell suspension free from serum-containing medium by centrifuging 10 min at $500 \times g$ and resuspending twice in HeBS. After the last wash, resuspend in HeBS at $0.2\text{--}1 \times 10^7$ cells/ml. Divide the cells between two tubes.

Use HeBS containing 1 mM CaCl₂ if Vibrio cholerae sialidase is to be used.

Cell number can be determined by trypan blue exclusion as described in UNIT 11.5.

2. To one tube, add sialidase. To the second tube, add no enzyme. Incubate 30 min at 37°C.

Select the enzyme to be used based on the properties in Table 17.12.1. Initially use the sialidase in vast excess, e.g., 100 mU/ml for 10^7 cells/ml (presuming 10^7 cells will yield ~1 nmol), then titrate down for the optimum amount.

Alternative controls could include reactions incubated with inactivated sialidase (prepared by boiling 5 min) or active sialidase plus 1 mM 2,3-dehydro-2-deoxy-N-acetylneuraminic acid (Boehringer Mannheim), a sialidase inhibitor.

- 3a. *To characterize the sialic acid molecules released from the cells:* Chill reactions to 4°C and centrifuge 10 min at $2000 \times g$, 4°C. Collect supernatant, boil 5 min to inactivate sialidase, and proceed with characterization of the released sialic acid (UNIT 17.16).

It is important to inactivate the sialidase as otherwise it will continue to work on any glycoproteins shed or secreted into the medium during the 30-min incubation.

- 3b. *To characterize the glycoconjugates on the cell surface:* Wash cells three times in culture medium containing serum (to remove free sialidase) before proceeding with biochemical or functional assays.

REAGENTS AND SOLUTIONS**HEPES-buffered saline (HeBS)**

476 mg HEPES (20 mM final)
812 mg NaCl (140 mM final)
H₂O to 100 ml
Adjust pH to 7.0 with HCl
Stable <1 year at 4°C.

Sialidase digestion buffers

Clostridium perfringens and Arthrobacter ureafaciens sialidases: Dissolve 8.6 g sodium acetate in 100 ml water to obtain a 0.1 M sodium acetate stock solution. Adjust pH to 5.5 with 0.1 M acetic acid. Store <1 year at 4°C.

Vibrio cholerae sialidase: Add 14.7 mg CaCl₂ dihydrate and 580 mg NaCl to 100 ml of 0.1 M sodium acetate, pH 5.5. Store <1 year at 4°C.

Newcastle disease virus sialidase: Add 2 mg of fatty acid-free BSA (Sigma A7511)

per milliliter of 0.1 M sodium acetate, pH 5.5. Any high-purity BSA can be used (contaminating glycosidases may be present in low levels in some BSA preparations). Store indefinitely at -20°C .

Sialidases

Commercial sources for sialidases include Sigma, Calbiochem, Boehringer Mannheim, and Oxford GlycoSystems (see Table 17.12.1 and APPENDIX 4).

***Clostridium perfringens* sialidase:** Supplied as a lyophilized powder. Reconstitute to 10 mU/ μl in digestion buffer (see above) and store for several months at 4°C .

***Arthrobacter ureafaciens* sialidase:** Supplied as a lyophilized powder. Reconstitute to 1 to 10 mU/ μl per manufacturer's instructions. Store for ~ 6 months at 4°C .

***Vibrio cholerae* sialidase:** Supplied in solution. Store per manufacturer's instructions.

Newcastle disease virus sialidase: Available commercially. Alternatively, prepare as per Paulson et al. (1982). Store stocks in 5- to 10- μl aliquots at -70°C . Thaw only once and store on wet ice at 4°C until use.

COMMENTARY

Background Information

Sialidases from many viral and bacterial strains have been described (Drzeniek, 1973). In general, bacterial enzymes have the broadest spectrum of activity, although the limited specificity of the viral enzymes can be an advantage for certain analyses.

Sialic acid is found in $\alpha(2-3)$ or $\alpha(2-6)$ linkages to neutral sugars (e.g., galactose, *N*-acetylgalactosamine, *N*-acetylglucosamine) or in $\alpha(2-8)$ linkage to another sialic acid. The $\alpha(2-3)$ - and $\alpha(2-6)$ -linked sialic acid residues are found in glycoproteins (both N- and O-linked) and in glycolipids. $\alpha(2-8)$ -linked residues are found in glycolipids, colominic acid (from *E. coli*), and a few mammalian glycoproteins (e.g., neural cell adhesion molecule). As outlined in Table 17.12.2, not all sialidases are active towards all linkages, and these differences can be exploited in structural analyses. Sialidases are only active against α -linked sialic acids. The only natural β -linked sialic acid is in CMP (cytidine monophosphate)-sialic acid.

The common sialic acid *N*-acetylneuraminic acid (Neu5Ac) can carry modifications (primarily *O*-acetylation but also glycolylation, sulfation, and methylation; see Varki, 1992, and Table 17.12.2). Sialidases differ in their activities towards these modified sialic acids. Unfortunately, the distribution of sialic acids containing these modifications has not been well-characterized in most mammalian systems.

Critical Parameters

Several caveats must be kept in mind in using sialidases.

Cleavage. The relative catalytic rates of the different enzymes for the various linkage types—i.e., $\alpha(2-3)$, $\alpha(2-6)$, and $\alpha(2-8)$ —are not of much consequence in most instances, because the enzymes are usually used in excess. Important exceptions to this include: (1) The Newcastle disease virus sialidase preferentially cleaves $\alpha(2-3)$ and $\alpha(2-8)$ linkages. However, $\alpha(2-6)$ -linked sialic acid will be hydrolyzed slowly if an excessive amount of enzyme is employed or if incubations are continued >30 min. (2) The *Vibrio cholerae* enzyme does not cleave the $\alpha(2-3)$ -linked “internal” sialic acid of extended gangliosides (e.g., G_{M1}), although it is active against the oligosaccharide if ceramide is removed. (3) The $\alpha(2-8)$ linkages of colominic acid and b series gangliosides are relatively resistant to release. Combined treatments with sialidases and endosialidases (Troy, 1992) may be used for polysialic acids with more than five sialic acid units. (4) Ganglioside-bound sialic acids may require detergents for complete release. Deoxycholate and cholate work the best.

Effect of substitution. Substitutions have variable effects upon release depending on the enzyme used, as summarized in Table 17.12.2. The decrease in rate with the *N*-glycolyl modification is not relevant in using any of the enzymes. The decrease in rates for the 9- or 7-mono-*O*-acetyl substitutions are also not practically relevant, except for the *Vibrio* enzyme, and particularly with gangliosides. On the other hand, the 4-mono-*O*-acetyl substitution causes complete resistance to all known sialidases. The effects of di- and tri-*O*-acetyl

substitutions have not yet been carefully studied, and it is not safe to extrapolate from data on mono-*O*-acetyl molecules. Likewise, combinations of substitutions (e.g., *N*-glycolyl and *O*-acetyl) have not been studied carefully. Finally, there is no information on the rarer types of sialic acids (e.g., *O*-methyl and *O*-sulfate substituted). If *O*-acetylation is suspected, prior treatment of the glycoconjugates with base (UNIT 17.15) will cause de-*O*-acetylation and improve release.

Controls. Whenever possible, controls such as the sample without enzyme and the enzyme alone must be studied in parallel to be sure that the sialic acids detected are actually released from the sample. This is particularly important when ultrasensitive methods (such as TBA with HPLC detection or DMB derivatization; UNIT 17.16) are employed. These methods are so sensitive that “environmental contamination” with sialic acids can become a problem.

Analysis of results. It is useful to compare the amount of sialic acid released (UNIT 17.16) by sialidase with the amount released by mild acid hydrolysis (UNIT 17.15). If the amount released by the sialidases is significantly less than that released by mild acid, then the sialic acid residues may be resistant due to (1) incorrect choice of sialidase, (2) modification of the sialic acid residues, (3) steric hindrance, or (4) suboptimal conditions for digestion.

It may be important to demonstrate that the observed effect produced by sialidase treatment (e.g., shift in isoelectric point on a focusing gel or a functional response) is in fact due to the sialidase activity. Treatment with heat-inactivated enzyme (prepared by boiling for 5 min) or inhibition with 1 mM 2,3-dehydro-2-deoxy-*N*-acetylneuraminic acid, a specific inhibitor of sialidases, can be used. The reaction should be titrated so that just enough sialidase is being used for the observed result, and then the effect of the inhibitor assessed.

Sialidase inhibitors. Inhibitors of sialidases do exist. Dextran sulfate inhibits the *Vibrio cholerae* sialidase. Other polyanions (e.g., DNA, RNA, and proteoglycans) may also inhibit this or other sialidases. Free sialic acid also can inhibit the *Vibrio cholerae* enzyme, with a K_i of ~5 mM.

Agglutination. Treating intact cells with Newcastle disease virus can cause agglutination of cells. This effect should not be confused with clumping due to cell death.

Buffer systems. The sialidases are compatible with several different buffer systems, with some differences in pH optima noted (Table

17.12.1). *Vibrio cholerae* sialidase works well in sodium acetate and Tris/maleate at pH 5.0 to 6.0 but is considerably more active in Tris/maleate at pH >5.0. *Clostridium perfringens* sialidase has pH optima of 4.5 in sodium acetate but 5.5 in citrate phosphate. *Arthrobacter ureafaciens* sialidase digestion proceeds maximally at pH 5.5, yet retains >50% of activity at pH 4.5 and 7.0. For all sialidases, alternative choices include sodium acetate, ammonium acetate, Tris, HEPES, phosphate, and cacodylate. Acetate interferes with HPLC on AX-5 resin (UNIT 17.12) and β -hexosaminidase digestion (UNIT 17.13).

Troubleshooting

Failure to digest a known sialoglycoconjugate to completion could be due to failure to add sufficient sialidase, use of an old or improperly reconstituted and/or stored enzyme aliquot, or inactivation of the enzyme after addition. If the sample is an unknown but potentially contains >1 nmol of releasable sialic acid, a titration comparing increasing amounts of sialidase versus sialic acid released is useful. Sialidases can lose activity rapidly (e.g., <1 hr) after dilution into an assay mix; inclusion of BSA will help prevent this inactivation. Some manufacturers recommend that the *Arthrobacter ureafaciens* and *Clostridium perfringens* sialidases be reconstituted with BSA for stability on storage.

Anticipated Results

If the type of sialidase used is tailored to the type of sialic acid and linkage, it should be possible to obtain nearly complete release of these residues.

Time Considerations

Once the appropriate buffers and standards are made, the actual digestions will take 30 min to 4 hr or longer (especially for gangliosides). Time required for analysis of the digested product(s) will depend on the specific analytical method employed (as listed in step 6 of the basic protocol).

Literature Cited

- Ada, G.L., French, E.L., and Lind, P.E. 1961. Purification and properties of neuraminidase from *Vibrio cholerae*. *J. Gen. Microbiol.* 24:409-421.
- Cassidy, J.T., Jourdan, G.W., and Roseman, S. 1965. The sialic acids IV: Purification and properties of sialidase from *Clostridium perfringens*. *J. Biol. Chem.* 240:3501-3506.
- Drzeniek, R. 1973. Substrate specificity of neuraminidases. *Histochem. J.* 5:271-290.

- Townsend, R.R., Hardy, M.R., Cumming, D.A., Carver, J.P., and Bendiak, B. 1989. Separation of branched sialylated oligosaccharides using HPAE-PAD. *Anal. Biochem.* 182:1-8.
- Troy, F.A. 1992. Polysialylation: From bacteria to brains. *Glycobiology* 2:5-24.
- Uchida, Y., Tsukada, Y., and Sugimori, T. 1979. Enzymatic properties of neuraminidases from *Arthrobacter ureafaciens*. *J. Biochem.* 86:1573-1585.
- Varki, A. 1992. Diversity in the sialic acids. In *Glycobiology* Vol. 2 (1): 25-40.
- Corfield, A.P., Sander Wewer, M., Veh, R.W., Wember, M., and Schauer, R. 1986. The action of sialidases on substrates containing *O*-acetyl sialic acids. *Biol. Chem. Hoppe-Seyler* 367:433-439.

Describes the activity of different sialidases on O-acetylated sialic acids.

Varki, A., and Diaz, S. 1984. The release and purification of sialic acids from glycoconjugates: Methods to minimize the loss and migration of *O*-acetyl groups. *Anal. Biochem.* 137:236-247.

Describes essential steps necessary to preserve O-acetyl groups during release and characterization of sialic acids.

Key References

- Corfield, A.P., Wember, M., Schauer, R., and Rott, R. 1982. The specificity of viral sialidases. *Eur. J. Biochem.* 124:521-525.

Describes biochemical properties of the Newcastle disease virus sialidase.

- Corfield, A.P., Higa, H.H., Paulson, J.C., and Schauer, R. 1983. The specificity of viral and bacterial sialidases for alpha (2-3) and alpha (2-6) linked sialic acids in glycoproteins. *Biochim. Biophys. Acta* 744:121-126.

Provides characterization of sialidases on $\alpha(2-3)$ and $\alpha(2-6)$ linked sialic acids.

Contributed by Leland D. Powell and
Ajit P. Varki
University of California San Diego
La Jolla, California

Endoglycosidase and Glycoamidase Release of N-Linked Oligosaccharides

Carbohydrate chain modifications are often used to monitor glycoprotein movement through the secretory pathway. This is because stepwise sugar-chain processing is unidirectional and generally corresponds to the forward or anterograde movement of proteins. This unit offers a group of techniques that will help analyze the general structure of carbohydrate chains on a protein and, therefore, oligosaccharide processing mileposts. The minimum requirements are that the protein can be labeled metabolically (UNIT 10.18) and immunoprecipitated (UNIT 10.16) and clearly seen on a gel or blot (UNIT 10.8). The sugar chains themselves are not analyzed, but their presence and structure are inferred from gel mobility differences after one or more enzymatic digestions. This approach is most often used in combination with [³⁵S]Met pulse-chase metabolic labeling protocols, but they can be applied to any suitably labeled protein (e.g., biotinylated or ¹²⁵I-labeled). As the oligosaccharide chains mature, they become either sensitive or resistant to highly specific glycosidases. Some of these enzymes cleave intact oligosaccharide chains from the protein—e.g., endo H, endo F₂, endo F₃, peptide:N-glycosidase F (PNGase F), endo D, and O-glycosidase. Others strip only terminal sugars (e.g., sialidase) or degrade a selected portion of the chain (e.g., endo-β-galactosidase). The techniques can be adapted to count the number of N-linked oligosaccharide chains on a protein. One unusual protease, O-sialoglycoprotease, degrades only proteins containing tight clusters of O-linked sialylated sugar chains. These techniques work best on average size proteins (<100 kDa) that contain a few percent carbohydrate by weight, where a gel shift of 1 kDa can be seen. A summary of the enzymes and their applications is shown in Table 17.13A.1.

This unit provides information on how to measure changes in carbohydrate structure and how these changes relate to protein trafficking. Fortunately, the techniques are independent of mechanistic views, although it should be borne in mind that the organization and distribution of many of these indicator enzymes are cell type dependent.

Table 17.13A.1 Enzymes Described in This Unit

Enzyme	Indications and uses	Monitors ^a
Endo D	Transient appearance of highly processed, sensitive forms prior to addition of GlcNAc by GlcNAc transferase I	Cis to medial Golgi
Endo F ₂	Presence of biantennary chains ± core fucose	Medial Golgi
Endo F ₃	Presence of core fucosylated biantennary α chains and/or triantennary chains ± core fucosylation	Medial Golgi
Endo-β-galactosidase	Presence of poly lactosamines	Trans-Golgi and TGN
Endo H	Conversion of high mannose to complex type N-linked chains	Cis to medial Golgi
O-Glycosidase	Presence of Galβ1,3GalNAc-α-Thr/Ser O-linked chains	Cis to medial Golgi
PNGase F	Presence of N-linked chains cleaves; nearly all N-linked chains; only enzyme that cleaves tetrantennary chains	Medial Golgi
Sialidase	Acquisition of sialic acids	Trans-Golgi and TGN
O-Sialoglycoprotease	Presence of mucin-like proteins with cluster of sialylated oligosaccharides	Trans-Golgi and TGN

^aAbbreviation: TGN, trans-Golgi network.

Contributed by Hudson H. Freeze

Current Protocols in Molecular Biology (1999) 17.13A.1-17.13A.26

Copyright © 1999 by John Wiley & Sons, Inc.

The starting material for these protocols is assumed to be [³⁵S]Met-labeled, immunoprecipitated protein bound to ~20 μl of protein A–Sepharose beads (as described in *UNIT 10.16*). The trace amount of protein is eluted by heating in a small volume of 0.1% SDS, diluted in the appropriate buffer and then digested with one or more enzymes in a small volume. The digest is analyzed on an appropriate SDS-PAGE system that can detect a 1- to 2-kDa size change. A change in the mobility of the protein after digestion is evidence that the carbohydrate chain was sensitive to the enzyme and therefore, that the protein had encountered a certain enzyme in the processing pathway. Alternatively, the analysis can be done by two-dimensional isoelectric focusing (IEF)/SDS-PAGE or two-dimensional nonequilibrium pH gradient electrophoresis (NEPHGE)/SDS-PAGE (*UNIT 10.4*) to see the loss of charged sugar residues or of anionic oligosaccharide chains. The same digestions and SDS-PAGE analysis also apply to proteins that are radioiodinated or biotinylated, or to immunoprecipitates derived from subcellular fractions separated on sucrose or Percoll gradients.

It is important to present the glycosylation pathways, as a detailed description of the pathways is needed to appreciate how they will be used in this unit. A single protein can have more than one kind of oligosaccharide (N-linked and O-linked), and each individual N-linked chain can mature into a different final form. The same is true for O-linked chains. Each is described below.

THE N-LINKED PATHWAY

The N-linked oligosaccharide maturation pathway is most frequently used for tracking protein movement through the Golgi complex. A common feature of all N-linked chains is the core region pentasaccharide shown in Figure 17.13A.1, which consists of three mannose units and two *N*-acetylglucosamine units. The mannose units comprise the trimannosyl core, and two of these residues are α-linked to the only β-linked mannose in the molecule. The β-linked mannose is bound to one of the two *N*-acetylglucosamines. Because they are β1-4 linked to each other, resembling the polysaccharide chitin, this is called a chitobiose disaccharide. Initially, all *N*-glycosylated proteins begin life when a preformed, lipid-associated oligosaccharide is transferred within the lumen of the endoplasmic reticulum (ER) to Asn of proteins having an Asn-X-Thr/Ser sequence. This precursor oligosaccharide contains three glucose (Glc), nine mannose (Man), and two *N*-acetylglucosamine (GlcNAc) sugar residues, and has the structure shown in Figure 17.13A.1. There are several ways to depict this structure. The short-hand symbol method used in Figure 17.13A.1 is the most convenient, but be sure to note the linkages of the individual sugars, as they are important. The α and β symbols denote the anomeric configuration of the sugar, and the number indicates which hydroxyl group of the next sugar is involved in the glycosidic linkage. In all cases, the anomeric position is 1, except in sialic acid where it is 2.

The details of the pathway are presented in Figures 17.13A.2 and 17.13A.3, along with the sensitivity to each endoglycosidase or glycoamidase. The figures show the steps between high-mannose and hybrid types (Fig. 17.13A.2) and complex types (Fig. 17.13A.3). The three Glc residues (filled triangles) are removed from properly folded proteins within the ER by two different oligosaccharide-processing α-glucosidases. The first α1-2Glc is cleaved by α-glucosidase I, and the next two α1-3Glc residues by α-glucosidase II (Fig. 17.13A.2, step 1). An ER-associated α-mannosidase removes one Man residue (open circle; Fig. 17.13A.2, step 2). The protein then moves on to the first step in Golgi-localized processing—the removal of the three remaining α1-2 Man units by Golgi α-mannosidase I to produce Man₅GlcNAc₂ (Fig. 17.13A.2, steps 3 and 4). Many proteins have only high-mannose-type oligosaccharides with five to nine Man residues,

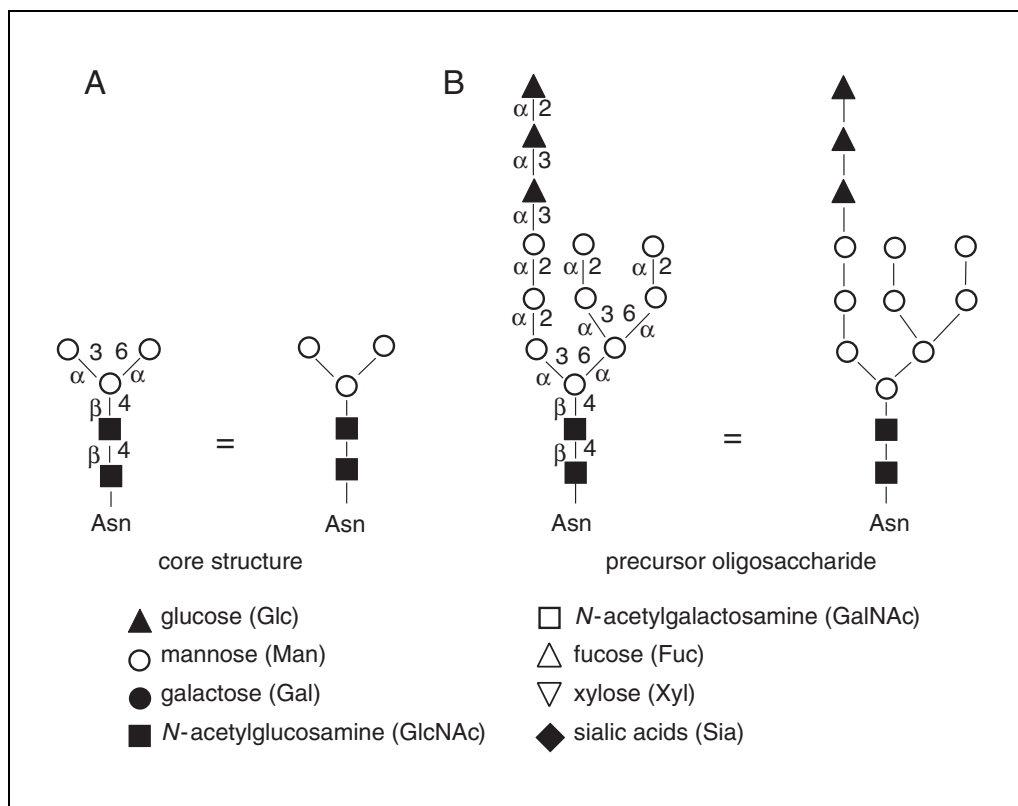


Figure 17.13A.1 Symbol structures for the core region and precursor of N-linked sugar chains. Each sugar is given a symbol and abbreviation at the bottom of the figure. Each one except sialic acid uses its anomeric carbon (C-1) for linking to other sugars. Sialic acid uses C-2 for glycosidic linkage to other sugars. Glycosidases and glycosyltransferases are anomeric specific and distinguish α or β configurations of each sugar. The core structure (**A**) is common to all N-linked chains and is composed of three Man and two GlcNAc residues. The α or β configuration of each sugar is indicated, and the OH group to which that sugar is linked is shown on the bar linking the two symbols. Thus, GlcNAc β 1-4GlcNAc β is represented by two filled squares with β and 4 between them. When a structure is first presented, it will have full display such as that on the left side; if it is repeated, only the symbols will be used as shown immediately to the right. The precursor oligosaccharide (**B**) for all N-linked chains is synthesized in the ER and transferred cotranslationally to the peptide containing an available Asn-X-Thr/Ser sequon.

and no further processing occurs. Alternatively, one to five GlcNAc residues (filled squares) can be added to the trimannosyl core, and these are usually extended with galactose (Gal; filled circles) and sialic acid (Sia; filled diamonds) residues. These extensions, called antennae, are the hallmarks of complex-type oligosaccharides. The transformation of the precursor sugar chain into various high-mannose or complex types is called oligosaccharide processing (Kornfeld and Kornfeld, 1985).

Man₅GlcNAc₂ is an important intermediate because it can have several fates. The first is the well-established addition of one GlcNAc residue by GlcNAc transferase I (Fig. 17.13A.2, step 6). This is the first step toward the formation of complex chains. However, simply adding Gal and Sia to the terminal GlcNAc of this oligosaccharide forms a hybrid structure (Fig. 17.13A.2, steps 12 and 13), where the left side of the molecule looks like a complex chain having one antenna, and the right side still resembles a high-mannose chain. The GlcNAc₁Man₅GlcNAc₂ structure is the *required* substrate for α -mannosidase II, which removes the two terminal Man units from the upper branch of the chain (i.e., the α 1-3Man and α 1-6 Man units; Fig. 17.13A.2, step 7). This enzyme only works after the addition of the first GlcNAc.

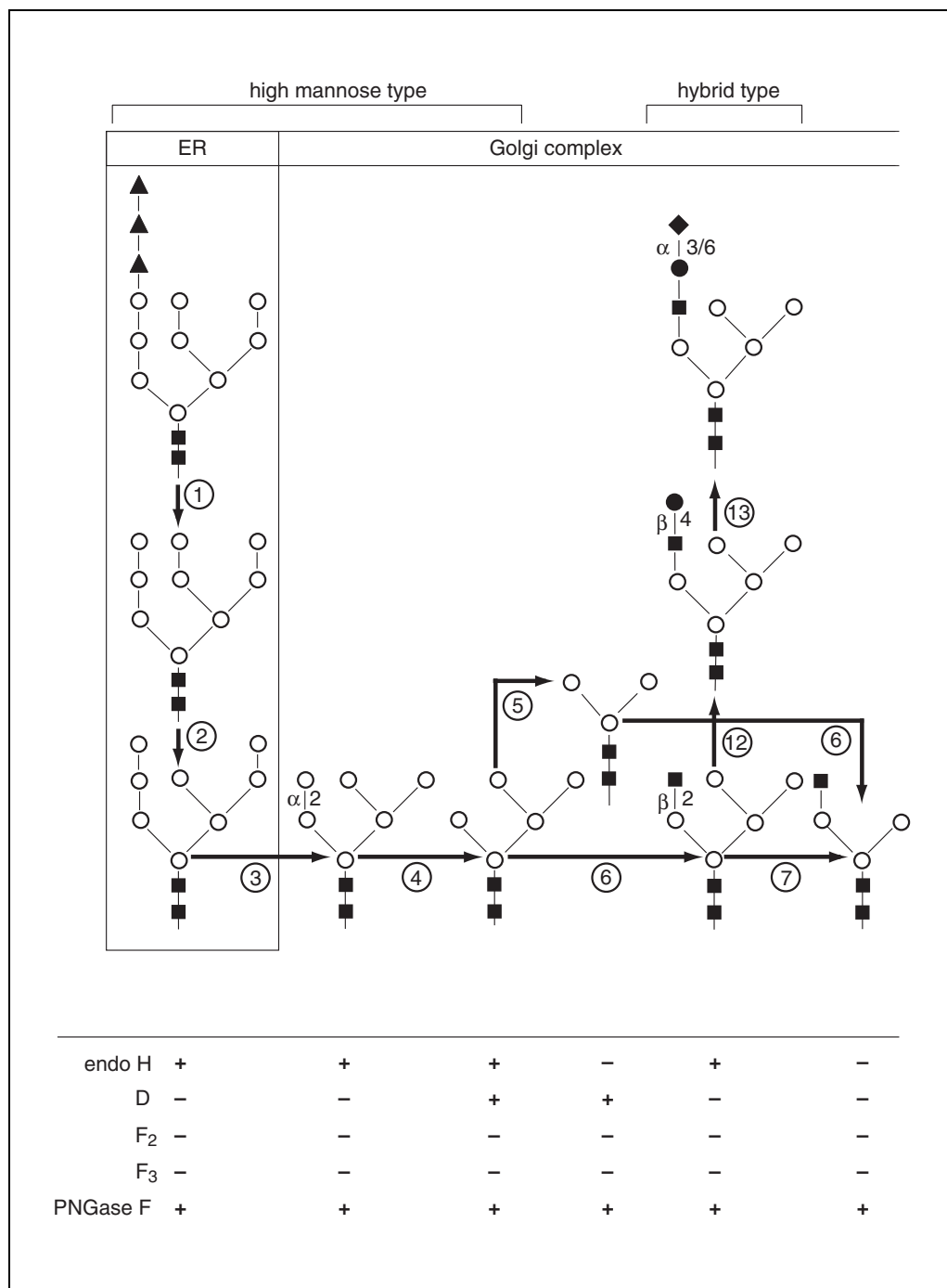


Figure 17.13A.2 N-linked oligosaccharide maturation pathway for high-mannose and hybrid types, and sensitivities to various enzymes (see Fig. 17.13A.1 for key). Brackets (top) show the structures designated as high-mannose and hybrid chains. The boxes indicate ER or Golgi localization. The pathway begins with the precursor oligosaccharide (see Figure 17.13A.1). Each successive numbered step in circles represents a glycosidase or glycosyl transferase that generates a new sugar chain with different sensitivities to the various endoglycosidases or PNGase F. (1) precursor oligosaccharide is trimmed by α -glucosidases I and II, removing three Glc. (2) ER mannosidase removes one Man. (3) α -Mannosidase I in Golgi complex removes two Man to make $\text{Man}_6\text{GlcNAc}_2$, with a single remaining α 1-2Man. (4) The final α 1-2Man is removed by a Golgi complex α -mannosidase I. (5) α -Mannosidase III removes the α 1-3 and α 1-6Man units to make $\text{Man}_3\text{GlcNAc}_2$. (6) GlcNAc transferase I adds GlcNAc to either $\text{Man}_5\text{GlcNAc}_2$ or $\text{Man}_3\text{GlcNAc}_2$. (7) α -Mannosidase II removes the α 1-3 and α 1-6Man units to make $\text{GlcNAc}_1\text{Man}_3\text{GlcNAc}_2$. Sensitivity to various enzymes (bottom) changes when moving from left to right, but remains the same within vertical columns. *NOTE:* This continued maturation to form complex chains is shown in Figure 17.13A.3. Additionally, these figures are not comprehensive; many glycosylation steps have not been included, but they do not affect the sensitivities to the enzymes listed.

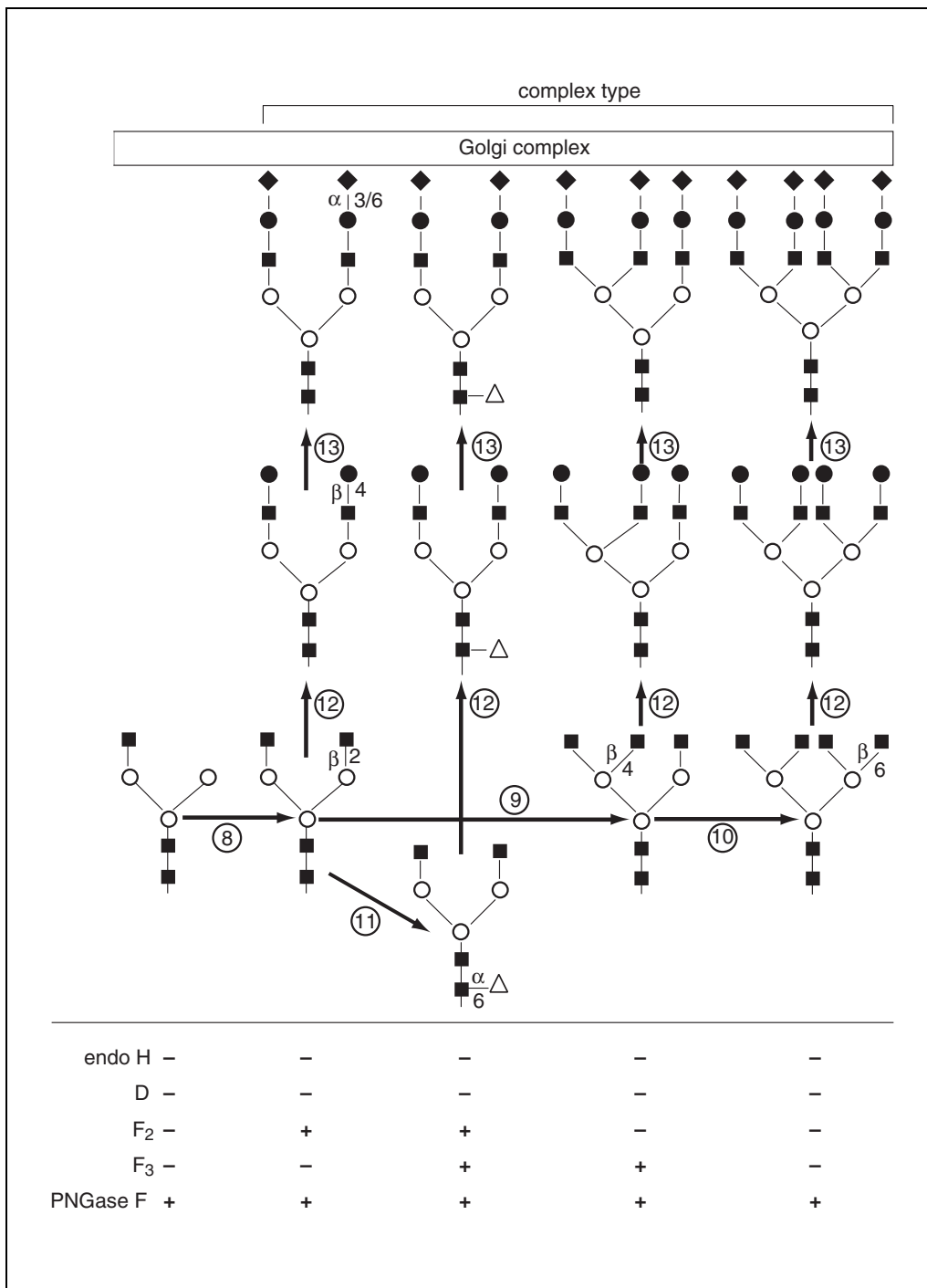


Figure 17.13A.3 N-linked oligosaccharide maturation pathway for complex types, and sensitivities to various enzymes (see Fig. 17.13A.1 for key; see Fig. 17.13A.2 for additional details). **(8)** GlcNAc transferase II adds a second GlcNAc to initiate a biantennary chain. **(9)** GlcNAc transferase IV adds a third GlcNAc to initiate a triantennary chain. **(10)** GlcNAc transferase V adds a fourth GlcNAc to initiate a tetraantennary chain. **(11)** Fucosyltransferase adds α 1-6Fuc to the core region of complex chains. **(12)** β 1-4Gal is added to available GlcNAc residues of hybrid and complex chains. **(13)** α 2-3 or α 2-6Sia is added to Gal residues of hybrid and complex chains.

The other fate for the $\text{Man}_5\text{GlcNAc}_2$ chain is to act as a substrate for the newly identified α -mannosidase III (Chui et al., 1997), which removes the same two Man units as α -mannosidase II, but does not require the prior addition of the first GlcNAc (Fig. 17.13A.2, step 5). α -Mannosidases II and III show partial overlap in the Golgi complex, but α -mannosidase III may occur preferentially in an earlier compartment. The $\text{Man}_3\text{GlcNAc}_2$ product of α -mannosidase III is also a substrate for GlcNAc transferase I. Thus, the same product, $\text{GlcNAc}_1\text{Man}_3\text{GlcNAc}_2$ can be formed in two ways: first, by the sequential action of α -mannosidase III and GlcNAc transferase I (Fig. 17.13A.2, steps 5 and 6) or, second, by GlcNAc transferase I and α -mannosidase II (Fig. 17.13A.2, steps 6 and 7). Sensitivity to specific enzyme digestions (endo H and endo D) can distinguish which route was taken (Fig. 17.13A.2).

GlcNAc transferase II now adds a second GlcNAc to the α 1-6-linked Man (Fig. 17.13A.3, step 8). This molecule can also have several fates. First, fucose (Fuc) can be added to GlcNAc residue linked to the Asn of the protein (Fig. 17.13A.3, step 11). Second, one to three more GlcNAc residues can be added to the core mannose residues to initiate tri- and tetraantennary chains (Fig. 17.13A.3, steps 9 and 10), and even pentaantennary chains (not shown). GlcNAc additions are considered to occur in the medial Golgi regions. Each GlcNAc-based branch can be individually modified, but they are usually extended by one Gal (Fig. 17.13A.3, step 12) and terminated by a Sia (Fig. 17.13A.3, step 13). Both of these sugars are usually thought to be added in *trans*-Golgi cisternae or in the *trans*-Golgi network (TGN). Sometimes selected antennae are also fucosylated in the TGN. One or more terminal Gal residues can be extended by variable-length polylactosamines (Gal β 4-GlcNAc repeats) capped by a Sia. GlcNAc and Gal can be sulfated as a late, perhaps even final, step of processing. These extensions/modifications are thought to occur in the late Golgi complex and TGN, but their order and compartmental segregation are not well understood. Other modifications of N-linked sugar chains are known, but there are fewer tools available to analyze their biosynthetic localization.

THE O-LINKED PATHWAY

For practical purposes, only a portion of the O-linked pathway—i.e., the addition of the first few sugars—will be presented. However, it is very important to remember that some of the same outer chain structures such as Sia, polylactosamines, and Fuc residues are common to both N- and O-linked oligosaccharides.

α -N-Acetylgalactosamine (α -GalNAc; open square) is the lead-off sugar for the O-linked pathway (Fig. 17.13A.4; also see Fig. 17.13A.1 for symbols). It is added to Ser/Thr residues that occur in the proper configuration, generating a broad variety of acceptor sequences. These sequences often cluster as repeats within mucin-like domains. GalNAc is added in the earliest parts of the Golgi complex, not cotranslationally. GalNAc can be further extended by at least six different sugars. The most common is the addition of a β 1-3Gal (Fig. 17.13A.4, step 1), forming a disaccharide that is one of the few O-linked chains that can be diagnosed by enzymatic digestions. This disaccharide is often capped by a Sia (Fig. 17.13A.4, step 2). Additional sugars such as Sia (Fig. 17.13A.4, step 3) or GlcNAc followed by Gal (Fig. 17.13A.4, steps 4 and 5) can be added. Structural analysis can be done by sequential exoglycosidase digestion, but given the complexity and heterogeneity of the sugar chains, such analysis is not a very useful indicator for tracking protein movement through the Golgi complex. Many O-linked chains have terminal Sia residues and, when tightly clustered on Ser/Thr residues, these chains promote proteolysis by O-sialoglycoprotease regardless of the structure of the underlying sugar chain.

Another type of O-linked glycosylation is the addition of glycosaminoglycan (GAG) chains to form proteoglycans. This occurs by a different pathway than the α -GalNAc

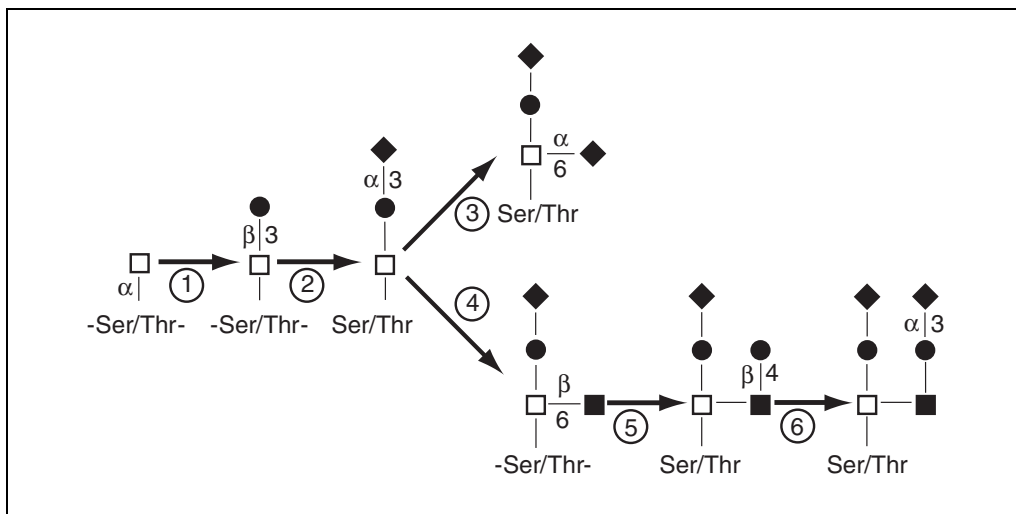


Figure 17.13A.4 A small portion of the O-GalNAc pathway (see Fig. 17.13A.1 for key). The first step of the O-linked pathway occurs in the early Golgi complex with the addition of α -GalNAc. There are at least six other sugars that can be added at this point in this complex pathway. Often β 1-3Gal is added (1), quickly followed by α 2-3Sia (2). The presence of these structures can be detected with a combination of O-glycosidase and sialidase. Additional sugars can be added as shown. α 2-6Sia (3) or β 1-6GlcNAc (4) followed by β 1-4Gal (5) and α 2-3Sia (6) on Gal. Each of these sugars must be removed before O-glycosidase can cleave the disaccharide.

linkage. Instead, the chains begin by addition of a β -Xylose (Xyl; open inverted triangle) residue to Ser and are then elongated by two Gal residues and a glucuronic acid (GlcA; half-filled diamond) residue. This core structure can be further elongated by the addition of GlcA β 1-3GalNAc β disaccharides to form the backbone of chondroitin/dermatan sulfate chains, or by GlcA β 1-3GlcNAc α to form the backbone of heparan sulfate chains. Biosynthesis and movement of these proteins have also been followed through the Golgi complex. Initiation begins in late ER/early Golgi complex, and the core tetrasaccharide is probably finished within the medial Golgi, but the addition of chondroitin chains appears to be confined to the TGN. In addition to the well-known O-linked GAG chains, there is clear evidence for the existence of a class of N-linked GAG chains.

ENDOGLYCOSIDASE H DIGESTION

Endoglycosidase H (endo H) cleaves N-linked oligosaccharides between the two *N*-acetylglucosamine (GlcNAc) residues (Fig. 17.13A.5) in the core region of the oligosaccharide chain (Fig. 17.13A.1) on high-mannose and hybrid, but not complex, oligosaccharides. In this protocol, a fully denatured protein is digested with endo H to obtain complete release of sensitive oligosaccharides.

Materials

- Immunoprecipitated protein of interest (UNIT 10.16)
- 0.1 M 2-mercaptoethanol (2-ME)/0.1% (w/v) SDS (ultrapure electrophoresis grade; prepare fresh)
- 0.5 M sodium citrate, pH 5.5
- 1% (w/v) phenylmethylsulfonyl fluoride (PMSF) in isopropanol
- 0.5 U/ml endoglycosidase H (endo H; natural or recombinant; Sigma, Glyko, or Boehringer Mannheim)
- 10 \times SDS sample buffer (UNIT 10.2A)
- Water baths, 30 $^{\circ}$ to 37 $^{\circ}$ C and 90 $^{\circ}$ C

BASIC PROTOCOL 1

Preparation and Analysis of Glycoconjugates

17.13A.7

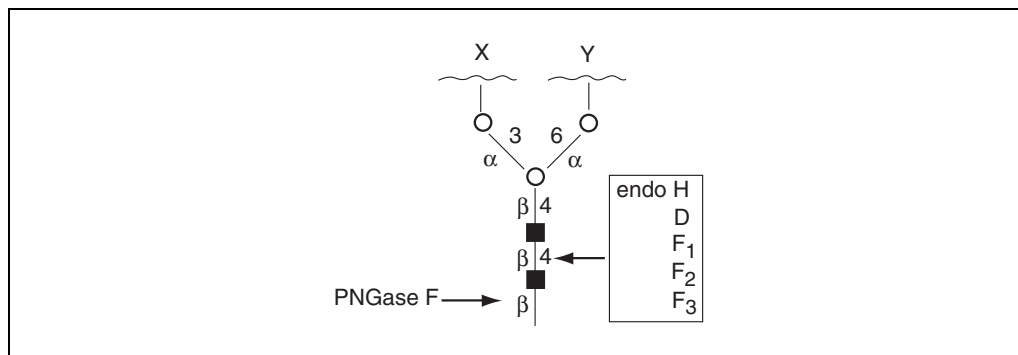


Figure 17.13A.5 PNGase F and endoglycosidase-sensitive bonds in the core of N-linked oligosaccharides (see Fig. 17.13A.1 for key). PNGase F is a glycoamidase that severs the bond between GlcNAc and Asn, liberating the entire sugar chain and converting Asn into Asp. The endoglycosidases (H, D, and Fs) cleave the bond between the two GlcNAc residues in the core region, leaving one GlcNAc still bound to the protein. The differential specificity of the endoglycosidases is based on the structure of the sugar chain in a fully denatured protein. Incomplete denaturation may not expose all sensitive linkages. X and Y are unspecified sugar residues.

Additional reagents and equipment for SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A)

1. Add 20 to 30 μ l of 0.1 M 2-ME/0.1% SDS to immunoprecipitate in a microcentrifuge tube, mix well, and heat denature 3 to 5 min at 90°C.

Use the larger amount of reagent ($\geq 30 \mu$ l) for more complete recovery. This treatment may also release some (unlabeled) antibody molecules.

Protein solubilization in nonionic detergents such as Triton X-100 or Nonidet P-40 is not always sufficient to completely expose all susceptible cleavage sites. Only strong denaturation with SDS exposes all sites for maximum cleavage.

2. Cool and microcentrifuge for 1 sec at 1000 \times g to collect condensed droplets in the bottom of the tube.
3. Place 10- μ l aliquots of solubilized, denatured protein (supernatant) in each of two clean microcentrifuge tubes, one for a control (no enzyme) and the other to digest (plus enzyme).
4. Add in the following order, mixing after each addition:

6 μ l 0.5 M sodium citrate, pH 5.5
 20 μ l H₂O
 2 μ l 1% PMSF (in isopropanol)
 1 μ l 0.5 U/ml endo H (enzyme digest only; substitute with water in control).

The PMSF prevents proteolysis. Nonionic detergent is not required to prevent inactivation of endo H as long as high-purity SDS is used.

The amount of water can be varied. Other solutions (or an increased amount of sample) may be substituted for water if needed, but potassium buffers should be avoided because they precipitate SDS as a potassium salt. The reaction volume can be scaled up proportionally if required, but in nearly all cases the enzyme will be in vast excess.

5. Incubate overnight at 30°C to 37°C.
6. Immediately prior to electrophoresis, inactivate endo H by adding 4 μ l of 10 \times SDS sample buffer and heating 5 min at 90°C.

7. Analyze protein by one-dimensional SDS-PAGE (UNIT 10.2) and autoradiography (APPENDIX 3A).

The presence of high mannose and/or hybrid N-linked oligosaccharide chains will be evidenced by increased mobility of the digested proteins on SDS-PAGE.

ENDOGLYCOSIDASE D DIGESTION

Like endo H, endo D also cleaves between the two GlcNAc residues in the core of the N-linked sugar chains (Fig. 17.13A.5). However, its narrow substrate specificity makes it useful for detecting the transient appearance of just a few early processing intermediates. It requires that the 2 position of the α 1-3-linked core Man be unsubstituted. This intermediate arises after processing by either α -mannosidase I or III, but prior to addition of the first GlcNAc or action of α -mannosidase II (see Fig. 17.13A.2, steps 3 to 5). Cells with a defect in GlcNAc I transferase (e.g., Lec 1 CHO cells) do not add the first GlcNAc residue (Fig. 17.13A.2, step 6), and N-linked oligosaccharides will remain sensitive to endo D because they cannot modify the α 1-3Man residue.

Materials

Immunoprecipitated protein of interest (UNIT 10.16)
0.1 M 2-mercaptoethanol (2-ME)/0.1% (w/v) SDS (ultrapure electrophoresis grade; prepare fresh)
0.5 M NaH₂PO₄, pH 6.5
10% (w/v) Triton X-100 or Nonidet P-40 (NP-40)
0.5 U/ml endoglycosidase D (endo D; Boehringer Mannheim)
10× SDS sample buffer (UNIT 10.2A)
Water baths, 37° and 90°C

Additional reagents and equipment for SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A)

1. Denature immunoprecipitated protein in 20 to 30 μ l of 0.1 M 2-ME/0.1% SDS by heating for 3 to 5 min at 90°C.

Use the larger amount of reagent ($\geq 30 \mu$ l) for more complete recovery. This treatment may also release some (unlabeled) antibody molecules.

2. Cool and microcentrifuge at 1000 \times g for 1 sec to collect condensed droplets in the bottom of the tube.
3. Transfer 10- μ l aliquots of supernatant into two microcentrifuge tubes, one for a control (no enzyme) and the other to digest (plus enzyme).
4. Add in the following order, mixing after each addition:

2 μ l 10% Triton X-100 or NP-40 (20-fold excess over SDS)
2 μ l 0.5 M NaH₂PO₄, pH 6.5
5 μ l H₂O
1 μ l 1 IU/ml endo D (enzyme digest only; substitute with water in control).

The 20-fold excess of nonionic detergent is essential to prevent inactivation of endo D by SDS.

The amount of water can be varied. Other solutions (or an increased amount of sample) may be substituted for water if needed, but potassium buffers should be avoided because they precipitate SDS as a potassium salt. The reaction volume can be scaled up proportionally if required, but in nearly all cases the enzyme will be in vast excess.

5. Incubate at 37°C overnight.

6. Immediately prior to electrophoresis, inactivate by adding 2 μ l of 10 \times SDS sample buffer and heating for 5 min at 90°C to 95°C.
7. Analyze protein by one-dimensional SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A).

Endo D sensitivity is detected by increased electrophoretic mobility of the digested proteins on SDS-PAGE.

ENDOGLYCOSIDASE F₂ DIGESTION

Endo F₂, like endo H and endo D, cleaves between the two GlcNAc residues in the chitobiose core (Fig. 17.13A.5). It preferentially releases biantennary complex-type oligosaccharide chains from glycoproteins, but does not cleave tri- or tetraantennary chains.

Materials

Immunoprecipitated protein of interest (UNIT 10.16)
0.1 M 2-mercaptoethanol (2-ME)/0.1% (w/v) SDS (ultrapure electrophoresis grade; prepare fresh)
0.5 M sodium acetate, pH 4.5 (APPENDIX 2)
10% (w/v) Triton X-100 or Nonidet P-40 (NP-40)
0.1 M 1,10-phenanthroline in methanol
200 mU/ml endoglycosidase F₂ (endo F₂; Glyko)
4 \times SDS sample buffer (UNIT 10.2A)
Water baths, 30° to 37°C and 90°C

Additional reagents and equipment for SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A)

1. Denature immunoprecipitated protein in 20 to 30 μ l of 0.1 M 2-ME/0.1% SDS by heating for 3 to 5 min at 90°C.

Use the larger amount of reagent ($\geq 30 \mu$ l) for more complete recovery. This treatment may also release some (unlabeled) antibody molecules.

2. Cool and microcentrifuge at 1000 \times g for 1 sec to collect condensed droplets in the bottom of the tube.
3. Transfer 10- μ l aliquots of supernatant into two microcentrifuge tubes, one for a control (no enzyme) and the other to digest (plus enzyme).
4. Add in the following order, mixing after each addition:

15 μ l 0.5 M sodium acetate, pH 4.5
3 μ l 0.1 M 1,10-phenanthroline in methanol
2 μ l 10% Triton X-100 or NP-40 (20-fold excess over SDS)
1 μ l 200 mU/ml endo F₂ (enzyme digest only; substitute 0.5 M sodium acetate in control).

A 10- to 20-fold excess of nonionic detergent is required to stabilize the enzyme.

The amount of water can be varied. Other solutions (or an increased amount of sample) may be substituted for water if needed, but potassium buffers should be avoided because they precipitate SDS as a potassium salt. The reaction volume can be scaled up proportionally if required, but in nearly all cases the enzyme will be in vast excess.

5. Incubate the mixture at 30° to 37°C overnight.

Some inactivation of the enzyme occurs at 37°C, even with nonionic detergent present; however, if the enzyme is present in sufficient excess, incubation can generally be carried out successfully at 37°C.

6. Immediately before electrophoresis, inactivate by adding 8 μ l of 4 \times SDS sample buffer and heating for 5 min at 90°C to 95°C.
7. Analyze the protein by one-dimensional SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A).

Sensitivity to endo F₂ is detected by increased electrophoretic mobility on SDS-PAGE.

ENDOGLYCOSIDASE F₃ DIGESTION

Endoglycosidase F₃ (endo F₃) is another endoglycosidase with a narrow substrate range and, therefore, high specificity: it cleaves triantennary chains, but not high-mannose, hybrid, nonfucosylated biantennary or tetraantennary chains. A core-fucosylated biantennary chain is the only other demonstrated substrate. When both endo F₃ and endo F₂ digestions are done in parallel on a sample, it can provide evidence for chain branching and core fucosylation. The approach is essentially the same as for the other endoglycosidases.

Materials

Immunoprecipitated protein of interest (UNIT 10.16)
0.1 M 2-mercaptoethanol (2-ME)/0.1% (w/v) SDS (ultrapure electrophoresis grade; prepare fresh)
0.5 M sodium acetate, pH 4.5 (APPENDIX 2)
10% (w/v) Triton X-100 or NP-40
0.1 U/ml endoglycosidase F₃ (endo F₃; Glyko)
10 \times SDS sample buffer (UNIT 10.2A)
Water baths, 37° and 90°C

Additional reagents and equipment for SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A)

1. Denature immunoprecipitated protein in 20 to 30 μ l of 0.1 M 2-ME/0.1% SDS and heat denature 3 to 5 min at 90°C.

Use the larger amount of reagent ($\geq 30 \mu$ l) for more complete recovery. This treatment may also release some (unlabeled) antibody molecules.

2. Cool and microcentrifuge for 1 sec at 1000 \times g to collect condensed droplets at the bottom of the tube.
3. Transfer 10- μ l aliquots of supernatant into two microcentrifuge tubes, one for a control (no enzyme) and the other to digest (plus enzyme).
4. Add in the following order, mixing after each addition:

2 μ l 10% Triton X-100 or NP-40 (20-fold excess over SDS)
4 μ l 0.5 M sodium acetate, pH 4.5
5 μ l H₂O
1 μ l 0.1 U/ml endo F₃ (enzyme digest only; substitute with water in control).

The amount of water can be varied. Other solutions (or an increased amount of sample) may be substituted for water if needed, but potassium buffers should be avoided because they precipitate SDS as a potassium salt. The reaction volume can be scaled up proportionally if required, but in nearly all cases the enzyme will be in vast excess.

5. Incubate at 37°C overnight.

BASIC PROTOCOL 4

Preparation and Analysis of Glycoconjugates

17.13A.11

6. Immediately prior to electrophoresis, inactivate by adding 2 μ l of 10 \times SDS sample buffer and heating for 5 min at 90°C to 95°C.
7. Analyze by one-dimensional SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A).

Sensitivity to endo F₃ is detected by increased mobility on SDS-PAGE.

PEPTIDE:N-GLYCOSIDASE F DIGESTION

PNGase F is a glycoamidase that cleaves the bond between the Asp residue of the protein and the GlcNAc residue that joins the carbohydrate to the protein (Fig. 17.13A.5). Because it liberates nearly all known N-linked oligosaccharides from glycoproteins, it is the preferred enzyme for complete removal of N-linked chains. It is the only enzyme that releases tetra- and pentaantennary chains. The glycoprotein sample must be denatured and digested with PNGase F to remove N-linked oligosaccharides completely.

Materials

Immunoprecipitated protein of interest (UNIT 10.16)
0.1 M 2-mercaptoethanol (2-ME)/0.1% (w/v) SDS (ultrapure electrophoresis grade; prepare fresh)
0.5 M Tris·Cl, pH 8.6 as determined at 37°C (APPENDIX 2)
10% (w/v) Triton X-100 or Nonidet P-40 (NP-40)
200 to 250 mU/ml peptide:N-glycosidase F (PNGase F; Sigma or Glyko)
10 \times SDS sample buffer (UNIT 10.2A)
Water baths, 30° to 37°C and 90°C

Additional reagents and equipment for SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A)

1. Denature immunoprecipitated protein in 20 to 30 μ l of 0.1 M 2-ME/0.1% SDS by heating 3 to 5 min at 90°C.

Use the larger amount of reagent ($\geq 30 \mu$ l) for more complete recovery. This treatment may also release some (unlabeled) antibody molecules.

2. Cool and microcentrifuge for 1 sec at 1000 \times g to collect condensed droplets in the bottom of the tube.
3. Transfer 10- μ l aliquots of supernatant into two microcentrifuge tubes, one for a control (no enzyme) and the other to digest (plus enzyme).
4. Add in the following order, mixing after each addition:

3 μ l 0.5 M Tris·Cl, pH 8.6
5 μ l H₂O
2 μ l 10% NP-40 or Triton X-100
5 μ l 200 to 250 mU/ml PNGase F (enzyme digest only; substitute with 0.5 M Tris·Cl in control).

Sodium phosphate or HEPES buffer, pH 7.0, can be used instead of Tris·Cl. Avoid potassium buffers because these may cause precipitation of a potassium SDS salt. Use of a nonionic detergent is essential, because SDS inactivates PNGase F. A 10-fold weight excess of any of the above nonionic detergents over the amount of SDS will stabilize the enzyme.

The amount of water can be varied. Other solutions (or an increased amount of sample) may be substituted for water if needed, but potassium buffers should be avoided because they precipitate SDS as a potassium salt. The reaction volume can be scaled up proportionally if required, but in nearly all cases the enzyme will be in vast excess.

5. Incubate overnight at 30° to 37°C.

6. Immediately prior to electrophoresis, inactivate the enzyme by adding 2.5 μl of 10 \times SDS sample buffer and heating 3 to 5 min at 90°C.
7. Analyze the protein by one-dimensional SDS-PAGE (UNIT 10.2) and autoradiography (APPENDIX 3A).

The presence of N-linked oligosaccharide chains will be evidenced by increased electrophoretic mobility on SDS-PAGE.

ESTIMATING THE NUMBER OF N-LINKED OLIGOSACCHARIDE CHAINS ON A GLYCOPROTEIN

SUPPORT PROTOCOL

One widely used application of endo H or PNGase F digestion is estimation of the number of N-linked oligosaccharide chains on a given glycoprotein. This is done by creating a ladder of partially digested molecules that each differ by only one N-linked sugar chain. The number of separate bands in a one-dimensional polyacrylamide gel (less one for the totally deglycosylated protein) provides an estimate of the number of N-linked chains. The conditions used to generate partially deglycosylated protein must be determined for each protein studied, because the sensitivity of each chain may be different, even when all of them are completely exposed by denaturation. For this protocol, either the incubation time or the amount of enzyme can be varied to determine the best conditions to produce a ladder of partial digests. Usually five or six points are enough to provide a reasonable estimate (Fig. 17.13A.6). Of course, it is important to use enough enzyme to obtain complete deglycosylation. This is best done by monitoring the effects of endo H or PNGase F on newly synthesized [^{35}S]Met pulse-labeled protein just after synthesis, but before any N-linked oligosaccharide processing has occurred. Pulse labeling of protein

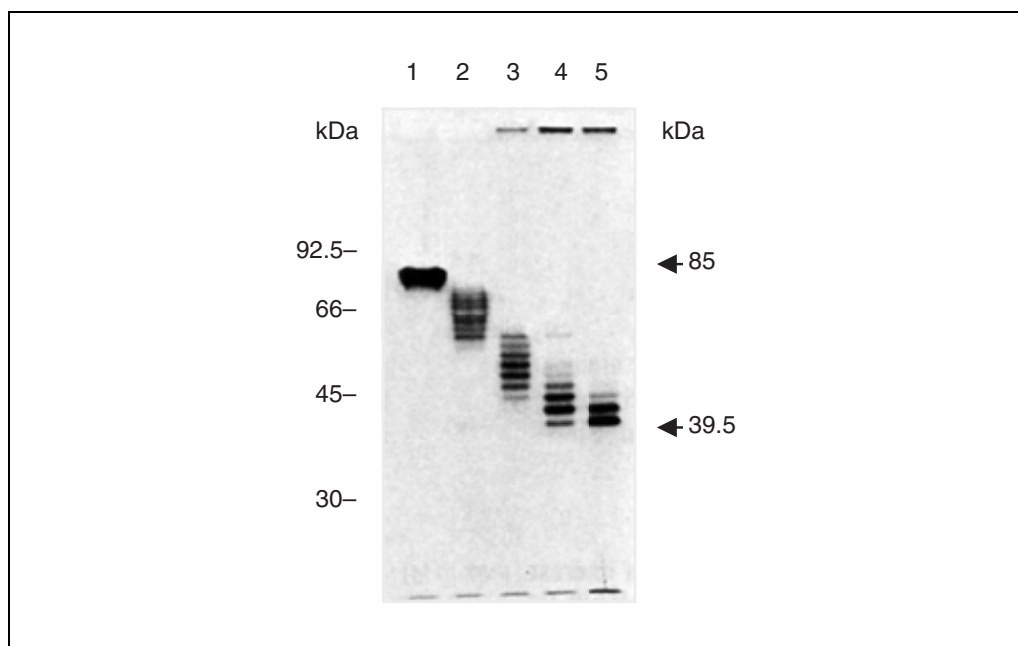


Figure 17.13A.6 Data from the estimation of the number of glycosylation sites on lysosome-associated membrane protein 1 (LAMP-1; Viitala et al., 1988). LAMP-1 contains eighteen potential N-linked sites. Graded digestion with increasing amounts of PNGase F was used to generate this ladder of glycoforms. Each band contains at least one less N-linked chain than the band above it. An average N-linked carbohydrate chain has an apparent mass of ~1.5 to 3 kDa. Lysosomal membrane glycoprotein was immunoprecipitated from [^{35}S]Met-labeled cells and the sample was digested with PNGase F for 0 min (lane 1), 5 min (lane 2), 20 min (lane 3), 45 min (lane 4), and 24 hr (lane 5). Figure courtesy of Dr. Minoru Fukuda.

Preparation and Analysis of Glycoconjugates

17.13A.13

for 10 min with [³⁵S]Met followed by digestion is the best way to be sure that all chains are removed.

1. Add 0.1 M 2-ME/0.1% SDS solution to the total volume of immunoprecipitated protein required and heat denature by incubating 3 to 5 min at 90°C.

Each digestion reaction requires 20 μl of immunoprecipitate. Thus, 120 to 140 μl is sufficient for one control plus five or six digests.

2. Cool and centrifuge for 1 sec at 1000 × g to collect condensed droplets at the bottom of the tube.
3. Aliquot 10 μl supernatant to the number of microcentrifuge tubes required to cover the concentration range (e.g., 0.01 to 1 mU/ml PNGase F) or incubation times (e.g., 5 to 60 min) plus one for an undigested control.
4. Add remaining reagents as specified for endo H (see Basic Protocol 1, step 4) or PNGase F (see Basic Protocol 5, step 4), adjusting the enzyme concentration as desired.
5. Incubate at 30°C for the desired length of time.

High enzyme concentration (10 mU/ml) and prolonged incubation (16 hr) must be among the conditions included, in order to ensure that there is a data point for maximum deglycosylation.

For varying enzyme concentrations, incubate for the same amount of time, but the duration of incubation should be shorter than what would give complete digestion because the goal is to obtain increasing extent of incomplete cleavage.

6. After the desired incubation time, inactivate enzyme by adding 0.1 volume of μl of 10× SDS sample buffer and heating 5 min at 90° to 95°C.
7. Analyze the sample from each concentration/time point, including undigested sample, by one-dimensional SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A).

Most newly formed N-linked chains will have a molecular weight in the range of 1500 to 2200, and loss of one chain is sufficient to change the migration of a protein. This procedure has been used to count up to eighteen N-linked sites on one molecule. A sample result is shown in Figure 17.13A.6.

BASIC PROTOCOL 6

SIALIDASE (NEURAMINIDASE) DIGESTION

Sialic acids are the terminal sugars on many N- and O-linked oligosaccharides. The great majority are released with broad-specificity sialidases (neuraminidases) such as that from *Arthrobacter ureafaciens*. Because sialic acids are charged, their loss usually changes the mobility on one-dimensional SDS polyacrylamide gels, but it will always change the mobility on a two-dimensional gel. Since one-dimensional analysis is easier, it can be tried first.

Materials

- Immunoprecipitated protein of interest (UNIT 10.16)
- 0.1 M 2-mercaptoethanol (2-ME)/0.1% (w/v) SDS (ultrapure electrophoresis grade; prepare fresh)
- 10% (w/v) Triton X-100 or Nonidet P-40 (NP-40)
- 0.5 M sodium acetate, pH 5.0 (APPENDIX 2)
- 1 IU/ml neuraminidase from *Arthrobacter ureafaciens* (Sigma or Glyco)
- 10× SDS sample buffer (UNIT 10.2A)
- Water baths, 37° and 90°C

Additional reagents and equipment for SDS-PAGE (UNIT 10.2A), for IEF/SDS-PAGE (UNIT 10.3), or NEPHGE/SDS-PAGE (UNIT 10.4), and for autoradiography (APPENDIX 3A)

1. Denature immunoprecipitated protein in 20 to 30 μl of 0.1 M 2-ME/0.1% SDS by heating 3 to 5 min at 90°C.

Denaturation is less important here, because the sialic acids are exposed at the ends of the sugar chains. In most instances, the denaturation step can probably be omitted and the digestion done while the protein is still bound to the beads.

2. Cool and microcentrifuge for 1 sec at 1000 \times g to collect condensed droplets in the bottom of the tube.
3. Transfer 10- μl aliquots of supernatant to two clean microcentrifuge tubes, one for a control (no enzyme) and the other to digest (plus enzyme).
4. Add in the following order, mixing after each addition:

2 μl 10% Triton X-100 or NP-40 (20-fold excess over SDS)

4 μl 0.5 M sodium acetate, pH 5.0

5 μl H₂O

1 μl 1 IU/ml neuraminidase (enzyme digest only; substitute with water for control).

This amount of neuraminidase should be in great excess. Addition of nonionic detergent is not needed if the digestion is done while the protein was still bound to the beads.

The amount of water can be varied. Other solutions (or an increased amount of sample) may be substituted for water if needed, but potassium buffers should be avoided because they precipitate SDS as a potassium salt. The reaction volume can be scaled up proportionally if required, but in nearly all cases the enzyme will be in vast excess.

5. Incubate at 37°C overnight.

This time can be shortened to 2 hr, if necessary, but longer incubations are better.

6. Immediately prior to electrophoresis, inactivate the enzyme by adding 2 μl of 10 \times SDS sample buffer and heating 3 to 5 min at 90°C.

If the protein will be analyzed by IEF or NEPHGE, addition of sample buffer is replaced by lysis buffer used for these techniques.

7. Analyze the protein using either the appropriate one-dimensional SDS-PAGE system (UNIT 10.2A) or a two-dimensional IEF/SDS-PAGE or NEPHGE/SDS-PAGE system (UNITS 10.3 & 10.4), and detect by autoradiography (APPENDIX 3A).

Removal of sialic acids usually results in a decrease in apparent molecular weight on one-dimensional gel analysis, or an increase in the isoelectric point of the protein analyzed by two-dimensional gel analysis.

ENDO- β -GALACTOSIDASE DIGESTION

The endo- β -galactosidase from *Bacillus fragilis* degrades polylactosamine chains (Gal β 1-4GlcNAc β 1-3)_n found on both N- and O-linked oligosaccharides. The variable length of these repeating units usually causes the protein to run as a broad band or a smear on the gel. Although not all linkages are equally cleaved by this enzyme (see Fig. 17.13A.7), sensitive proteins that often run as broad bands or smears on gels—e.g., lysosome-associated membrane protein 1 (LAMP-1)—produce both sharper bands and lower molecular weight species after digestion.

**BASIC
PROTOCOL 7**

**Preparation and
Analysis of
Glycoconjugates**

17.13A.15

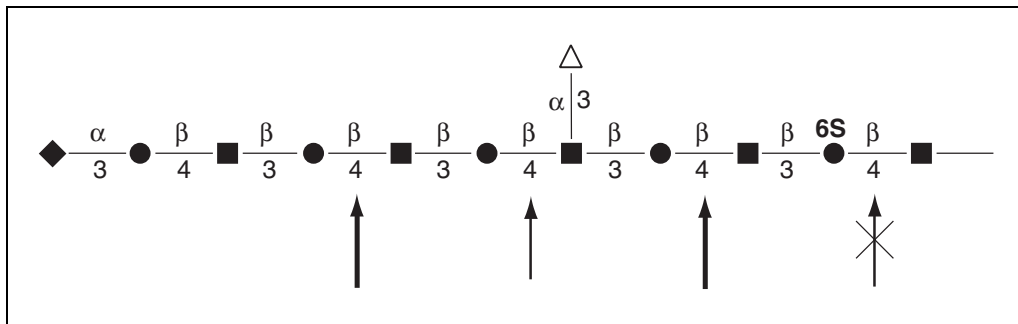


Figure 17.13A.7 Endo- β -galactosidase-sensitive linkages in polylectosamines (see Fig. 17.13A.1 for key). Linear, unsubstituted polylectosamine units (GlcNAc β 1-3Gal β 1-4) are sensitive to digestion with endo- β -galactosidase, while substitutions—such as sulfate esters (S) or branches starting with GlcNAc (not shown)—completely block digestion. Substitution of neighboring GlcNAc with Fuc or sulfate esters slows the rate, but does not block cleavage. Sensitive sites are shown with bold arrows, slowly hydrolyzed sites with a lighter arrow, and resistant bonds are struck out. Various substitutions are possible, leading to broad bands on gels. This will create variable sensitivities, but even partial sensitivity should give a sharper, more defined band.

Materials

Immunoprecipitated protein of interest (*UNIT 10.16*)
 0.1 M 2-mercaptoethanol (2-ME)/0.1% (w/v) SDS (ultrapure electrophoresis grade; prepare fresh)
 0.5 M sodium acetate buffer, pH 5.8 (*APPENDIX 2*)
 10% (w/v) Triton X-100 or Nonidet P-40 (NP-40)
 100 mU/ml endo- β -galactosidase (Boehringer Mannheim, Oxford GlycoSystems, Sigma)
 10 \times SDS sample buffer (*UNIT 10.2A*)
 Water baths, 37 $^{\circ}$ and 95 $^{\circ}$ C

Additional reagents and equipment for SDS-PAGE (*UNIT 10.2A*) and autoradiography (*APPENDIX 3A*)

- Denature immunoprecipitated protein in 20 to 30 μ l of 0.1 M 2-ME/0.1% SDS by heating 3 to 5 min at 95 $^{\circ}$ C.

Use the larger amount of reagent ($\geq 30 \mu$ l) for more complete recovery. This treatment may also release some (unlabeled) antibody molecules.

- Cool and microcentrifuge 1 sec at 1000 \times g to collect condensed droplets in the bottom of the tube.
- Transfer 10- μ l aliquots of supernatant into two microcentrifuge tubes, one for a control (no enzyme) and the other to digest (plus enzyme).

- Add in the following order, mixing after each addition:

2 μ l 10% Triton X-100 or NP-40 (20-fold excess over SDS)
 4 μ l 0.5 M sodium acetate, pH 5.8
 5 μ l H₂O
 1 μ l 100 mU/ml endo- β -galactosidase (enzyme digest only; substitute with water in control).

The amount of water can be varied. Other solutions (or an increased amount of sample) may be substituted for water if needed, but potassium buffers should be avoided because they precipitate SDS as a potassium salt. The reaction volume can be scaled up proportionally if required, but in nearly all cases the enzyme will be in vast excess.

5. Incubate at 37°C overnight.
6. Immediately prior to electrophoresis, inactivate by adding 3 μ l of 10 \times SDS sample buffer and heating for 5 min at 90°C.
7. Analyze protein by one-dimensional SDS-PAGE (UNIT 10.2) and autoradiography (APPENDIX 3A).

If the protein has poly lactosamine chains, its mobility should increase after digestion.

ENDO- α -N-ACETYL GALACTOSAMINIDASE DIGESTION

This enzyme (also known as O-glycosidase or O-glycanase) has limited utility because it is highly specific for cleaving only one O-linked disaccharide, Gal β 1-3GalNAc α -Ser/Thr. Adding any more sugars, including sialic acid, renders the molecule resistant to cleavage and requires removal of each residue before the enzyme will work. Prior sialidase digestion is sometimes used (see Basic Protocol 6), and this can be done while the protein is still bound to the immunoprecipitation beads.

Materials

- Immunoprecipitated protein of interest (UNIT 10.16)
- 0.1 M 2-mercaptoethanol (2-ME)/0.1% (w/v) SDS (ultrapure electrophoresis grade; prepare fresh)
- 0.5 M sodium citrate phosphate buffer, pH 6.0, containing 500 μ g/ml BSA (complete buffer supplied with enzyme)
- 10% (w/v) Triton X-100 or Nonidet P-40 (NP-40)
- 300 mU/ml endo- α -N-acetylgalactosaminidase (5 \times concentrate from Glyko; use according to directions)
- 10 \times SDS sample buffer (UNIT 10.2A)
- Water bath, 37° and 95°C
- Additional reagents and equipment for SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A)

1. Denature immunoprecipitated protein in 20 to 30 μ l of 0.1 M 2-ME/0.1% SDS by heating 3 to 5 min at 95°C.

Use the larger amount of reagent ($\geq 30 \mu$ l) for more complete recovery. This treatment may also release some (unlabeled) antibody molecules.

2. Cool and microcentrifuge 1 sec at 1000 \times g to collect condensed droplets in the bottom of the tube.
3. Transfer 10- μ l aliquots of supernatant into two microcentrifuge tubes, one for a control (no enzyme) and the other to digest (plus enzyme).
4. Add in the following order, mixing after each addition:

- 2 μ l 10% Triton X-100 or NP-40 (20-fold excess over SDS)
- 4 μ l 0.5 M sodium citrate phosphate buffer, pH 6.0, with 500 μ g/ml BSA
- 3 μ l H₂O
- 1 μ l 300 mU/ml endo- α -N-acetylgalactosidase (enzyme digest only; substitute with water in control).

The amount of water can be varied. Other solutions (or an increased amount of sample) may be substituted for water if needed, but potassium buffers should be avoided because they precipitate SDS as a potassium salt. The reaction volume can be scaled up proportionally if required, but in nearly all cases the enzyme will be in vast excess.

5. Incubate at 37°C overnight.
6. Immediately prior to electrophoresis, inactivate by adding 2 µl of 10× SDS sample buffer and heating for 5 min at 90°C.
7. Analyze protein by one-dimensional SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A).

If the protein contains the disaccharide unit, the mobility of the protein should increase. Presence of only a single unit (mol. wt. ~400 Da) may be difficult to detect unless a high-resolution gel is used.

O-SIALOGLYCOPROTEASE DIGESTION

Digestion with O-sialoglycoprotease requires that the substrate have a tight cluster of sialylated O-linked oligosaccharides. Proteins with a single O-linked chain or a few widely spaced chains will not be cleaved. This property makes the enzyme a valuable diagnostic tool.

Materials

- Immunoprecipitated protein of interest (UNIT 10.16)
- 0.1 M 2-mercaptoethanol (2-ME)/0.1% (w/v) SDS (ultrapure electrophoresis grade; prepare fresh)
- 0.4 M HEPES buffer, pH 7.4
- 10% (w/v) Triton X-100 or Nonidet P-40 (NP-40)
- 2.4 mg/ml O-sialoglycoprotease (O-sialoglycoprotein endoglycoprotease; Accurate Chemical & Scientific; reconstituted according to directions)
- 10× SDS sample buffer (UNIT 10.2A)
- Water baths, 37° and 95°C

Additional reagents and equipment for SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A)

1. Denature immunoprecipitated protein in 20 to 30 µl of 0.1 M 2-ME/0.1% SDS by heating 3 to 5 min at 95°C.

Use the larger amount of reagent (≥30 µl) for more complete recovery. This treatment may also release some (unlabeled) antibody molecules.
2. Cool and microcentrifuge 1 sec at 1000 × g to collect condensed droplets in the bottom of the tube.
3. Transfer 10-µl aliquots of supernatant into two microcentrifuge tubes, one for a control (no enzyme) and the other to digest (plus enzyme).
4. Add in the following order, mixing after each addition:
 - 2 µl 10% Triton X-100 or NP-40 (20-fold excess over SDS)
 - 4 µl 0.4 M HEPES buffer, pH 7.4
 - 5 µl H₂O
 - 2 µl 2.4 mg/ml O-sialoglycoprotease (enzyme digest only; substitute with water in control).

O-Sialoglycoprotein endopeptidase is a partially purified enzyme, and the specific activity is relatively low. A quantity of 1.0 µg of this enzyme preparation will cleave 5 µg of sensitive substrate per hour at 37°C. Human glycophorin A can serve as a positive control.

The amount of water can be varied. Other solutions (or an increased amount of sample) may be substituted for water if needed, but potassium buffers should be avoided because they precipitate SDS as a potassium salt. The reaction volume can be scaled up proportionally if required, but in nearly all cases the enzyme will be in vast excess.

5. Incubate at 37°C overnight.
6. Immediately prior to electrophoresis, inactivate by adding 2.5 μl of 10× SDS sample buffer and heating for 5 min at 90°C.
7. Analyze protein by one-dimensional SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A).

If the digestion was successful, the target protein will be undetectable or may be cleaved into small fragments.

COMMENTARY

Background Information

The results of digestion of a hypothetical protein with two N-linked carbohydrate chains as it moves through the ER and Golgi complex with various enzymes are shown in Figure 17.13A.8. At 0 min, both N-linked chains are high-mannose type. They have lost their Glc residues and one Man residue in the ER. Both are sensitive to endo H and PNGase F digestion, yielding a protein with only two remaining GlcNAc residues in the case of endo H diges-

tion, and no carbohydrate at all in the case of PNGase F digestion. These sugar chains are resistant to the other enzyme digestions.

At 45 min, the protein is in the medial Golgi complex and both sugar chains have been processed by Golgi α-mannosidase I. However, one of the chains (left) has been partially processed by Golgi α-mannosidase III to an endo H-resistant/endo D-sensitive chain. The other chain (right) has received a GlcNAc residue from GlcNAc transferase I, but has not encountered

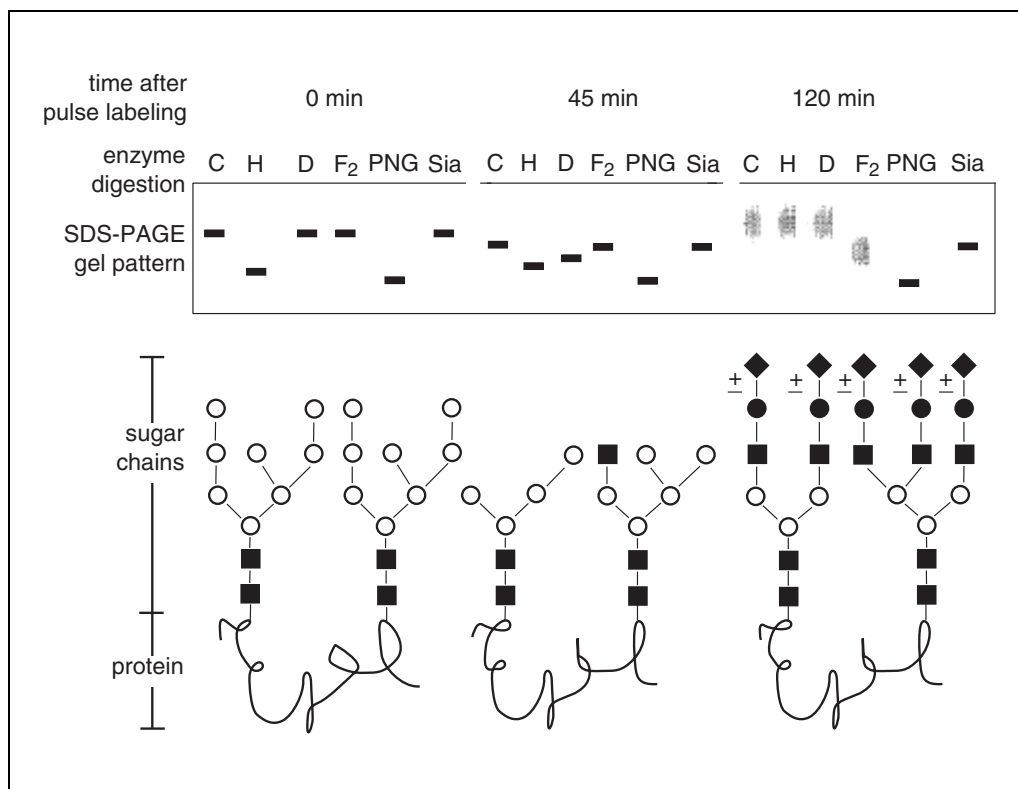


Figure 17.13A.8 Schematic diagram showing results that could be obtained for a hypothetical protein with two N-linked glycosylation sites as it moves through the Golgi complex. Assume that the protein has been biosynthetically labeled with an amino acid precursor (such as [³⁵S]Met) for 10 min and chased in the absence of label for 45 min and 120 min. The protein is then precipitated with a specific antibody. At each time point, equal amounts of the sample are analyzed by fluorography after one-dimensional SDS-PAGE, either without any digestion (control; C) or following digestion with endo H (H), endo D (D), endo F₂ (F₂), PNGase F (PNG), or sialidase (Sia). Oligosaccharide structures consistent with the banding patterns are shown below the schematic gel pattern.

Golgi α -mannosidase II. This chain is still endo H sensitive and endo D resistant. Both chains are sensitive to PNGase F, but neither is sensitive to endo F₂ or sialidase digestion.

At 120 min, both chains have fully matured to complex-type chains. One is a sialylated biantennary chain and the other is a sialylated triantennary chain. Note that each sialic acid is marked \pm , indicating that not all molecules are fully sialylated, accounting for the broader bands. Neither sugar chain is sensitive to endo H or endo D digestion. The biantennary chain (left) is sensitive to endo F₂ cleavage, leaving one GlcNAc residue on the protein, but the triantennary chain (right) is not sensitive to this enzyme. Both chains are cleaved by PNGase F. All sialic acids are removed by sialidase to produce a sharp band, but the underlying sugar chains remain.

Endoglycosidase H

Endo H from *Streptomyces plicatus* cleaves the bond between the two GlcNAc residues in the core of N-linked oligosaccharides. One GlcNAc residue remains attached to the protein or peptide and the remainder of the chain is released as an intact unit (Tarentino et al., 1989). The oligosaccharide structures cleaved by endo H are shown schematically in Figure 17.13A.5. Substrates for endo H include all high-mannose oligosaccharides and certain hybrid types, but not bi-, tri-, tetra- or pentaantennary (complex) chains. Endo H also cleaves oligosaccharides that have α 1-6 fucose residues bound to the reducing (protein to carbohydrate linkage) GlcNAc residue (Tarentino et al., 1989).

Endo H sensitivity is the most common way to trace the movement of newly synthesized glycoproteins from the endoplasmic reticulum (ER) into the Golgi complex. Proteins remain sensitive to endo H while they are in the ER and in early regions of the Golgi complex; they become endo H-resistant after they are processed by enzymes located in the medial Golgi complex. Endo H cleaves the N-linked oligosaccharides from proteins as long as they have not lost the α 1-3Man residue cleaved by Golgi mannosidase II or mannosidase III (Fig. 17.13A.2). After removal of that mannose, the oligosaccharide becomes endo H resistant. Endo H is the best enzyme for identifying high-mannose and hybrid chains and their general transition toward complex chains.

Using endo H alone gives an incomplete map of subcellular trafficking. The resolution of this picture can be enhanced by combining it with endo D digestions. In general, N-linked

chains become endo H resistant at about the same time they become transiently sensitive to endo D. However, the discovery of the alternative pathway for complex N-glycan processing using α -mannosidase III now gives a different perspective. Both α -mannosidase II and III are expressed in most mouse tissues except for hematopoietic cells, which seem to have only α -mannosidase II. So it is conceivable that some proteins, and even different N-linked chains on a single protein, can be processed by α -mannosidase II and others by α -mannosidase III, yielding complex digestion patterns and kinetics of endo H and endo D sensitivity.

Endoglycosidase D

The narrow substrate specificity of endo D makes it useful for detecting a very restricted set of processing intermediates, particularly for distinguishing α -mannosidase II from α -mannosidase III processing. When the results of endo D digestion are combined with the results of an endo H digestion of the same sample, it can help determine which of the two alternate processing pathways is being used. The key requirement for endo D cleavage is an unsubstituted 2 position on the α 1-3Man that forms the trimannosyl core (Fig. 17.13A.1). This window of endo D sensitivity occurs only *after* removal of the α 1-2Man residue by α -mannosidase I and *before* the addition of a GlcNAc residue to this same position by GlcNAc transferase I (Beckers et al., 1987; Davidson and Balch, 1993). The immediate precursor of both pathways is Man₆GlcNAc₂, which is endo D resistant/endo H sensitive. The better-known α -mannosidase II pathway involves the removal of the last α 1-2Man residue by α -mannosidase I (making the oligosaccharide endo D sensitive/endo H sensitive), and then the addition of GlcNAc via GlcNAc transferase I (making it endo D resistant/endo H sensitive). This product, and only this one, is the substrate for α -mannosidase II, which removes two of the remaining five Man residues, specifically the α 1-3Man and α 1-6Man that form the upper branch. Once this occurs, this endo D-resistant molecule also becomes endo H resistant. This structure is also a substrate for GlcNAc transferase II on the way to forming a biantennary chain.

The second pathway involves α -mannosidase III, which was identified as a functionally important enzyme in nearly all tissues when the α -mannosidase II pathway was ablated in mice (Chui et al., 1997). After α -mannosidase I removes the final α 1-2Man from

Man₆GlcNAc₂ to generate Man₅GlcNAc₂, α -mannosidase III cleaves the same two Man residues as α -mannosidase II; however, α -mannosidase III cleavage does not require addition of the first GlcNAc that α -mannosidase II requires. α -Mannosidase III creates the endo D-sensitive/endo H-resistant trimannosyl core. This sugar chain serves as an acceptor for GlcNAc transferase I to form an endo D-resistant/endo H-resistant molecule that can be modified by GlcNAc transferase II, and then on to biantennary chains. As summarized in Figure 17.13A.2, using the two endoglycosidases together will distinguish whether the oligosaccharide was processed by α -mannosidase II or α -mannosidase III. This may be important since there is evidence that distribution of the two enzymes in the Golgi complex is different with α -mannosidase III in an earlier compartment. It is important to point out that all chains of even a single protein are not necessarily processed using all the same enzymes.

Mutant cell line CHO Lec1 lacks GlcNAc transferase I activity and cannot synthesize either complex or hybrid chains. These cells can be used to measure the kinetics of acquiring endo D sensitivity (Beckers et al., 1987). This cell line can be obtained through ATCC (CRL-1735). The chains become permanently endo D sensitive, because they lack GlcNAc transferase I and cannot be converted back to an endo D-resistant form. These chains will also remain endo H sensitive, because α -mannosidase II requires GlcNAc transferase I. However, if any chains are processed by α -mannosidase III, they would become endo H resistant. Again, the combination of endo D and endo H digestions can reveal which pathway was used.

CHO Lec1 cells are also useful for tracking the movement of a protein from the ER into the earliest Golgi compartment where α -mannosidase I is located. Acquisition of endo D sensitivity requires the action of this enzyme. The advantage of using Lec1 cells is that the proteins remain permanently endo D sensitive and there is no risk of kinetically missing that small window of sensitivity before the sugar chain might become endo D resistant once again. Even if α -mannosidase III acts on the protein, it would still remain endo D sensitive, and no further processing would occur.

Endoglycosidase F and peptide N-glycosidase F

Elder and Alexander (1982; Alexander and Elder, 1989) made a landmark discovery when they identified an enzyme in culture filtrates of the bacterium *Flavobacterium meningosepticum* that cleaved N-linked oligosaccharides from glycoproteins. This preparation had a broad substrate specificity. The endoglycosidase activity in this preparation (endo F) was actually due to a set of enzymes, each with a more restricted substrate range. Like endo H, the endo F enzymes cleave the sugar chains between the two core GlcNAc residues (Fig. 17.13A.5). Endo F was originally thought to be a single enzyme, but it is now known that each of the three enzymes has a distinct specificity (Plummer and Tarentino, 1991). The specificity of endo F₁ is very similar to that of endo H, while endo F₂ prefers biantennary chains, and endo F₃ will cleave core fucosylated biantennary and triantennary, but not tetraantennary, chains (Fig. 17.13A.2 and Fig. 17.13A.3). All are commercially available from Glyko.

Plummer et al. (1984) carefully analyzed *Flavobacterium* filtrates and found that the very broad substrate range was actually due to a glycoamidase activity rather than an endoglycosidase activity. The glycoamidase releases the entire carbohydrate chain from the protein by cleaving the Asp-GlcNAc bond (Fig. 17.13A.5). The enzyme is called by various names, including peptide:N-glycosidase F (PNGase F), glycopeptidase F, and N-glycanase (previously available from Genzyme), but the proper name is peptide N-4(N-acetyl- β -glucosaminyl)asparagine amidase F. PNGase F has the broadest specificity, and it releases most of the N-linked oligosaccharide chains from proteins.

Endo F₂ and Endo F₃

Endo F₂ prefers biantennary chains over high-mannose chains by ~20 fold. Thus, endo H and PNGase F are better choices for broadly distinguishing high-mannose from complex chains as described above in the endo H protocol (Tarentino and Plummer, 1994).

Many proteins have core-fucosylated N-linked glycans, and the addition of fucose can be used as an additional trafficking marker. Endo F₃ will hydrolyze triantennary chains, but endo F₂ will not. Endo F₂ hydrolyzes biantennary chains; however, endo F₃ will also hydrolyze core-fucosylated biantennary chains only

a bit more slowly than it does triantennary chains. Thus, if all the chains on a protein are sensitive to endo F₂ (biantennary) and to endo F₃, this is evidence for the presence of a core fucose on those chains.

The enhancement of endo F₃ activity on biantennary chains with a core α 1-6Fuc points out that some specificities are really a matter of relative rates of cleavage. If both endo F₂ (biantennary) and endo F₃ (triantennary and biantennary with core fucose) cleave the protein, they may be acting on different chains on the same molecule. If there is only a single chain, repeat the experiment under the same conditions using 10- to 20-fold dilutions of each enzyme. If both still cleave the chain about equally, it is evidence for core fucosylation of a biantennary chain.

PNGase F

PNGase F has the broadest specificity of all the enzymes that cleave N-linked oligosaccharides. It is indifferent to all extended structures on the chains, such as sulfate, phosphate, polylactosamines, polysialic acids, and even the occasional glycosaminoglycan chain. Most of the modifications in the Man₃GlcNAc₂ core region also make no difference in chain cleavage. The only oligosaccharide structural feature that confers PNGase F resistance is the presence of an α 1-3Fuc on the GlcNAc bound to Asn (Tretter et al., 1991). This modification is commonly found in plants and in some insect glycoproteins, but it is rare in most mammalian cell lines. However, caution is warranted, as there is evidence that some mammalian cells do have the critical α 1-3Fuc transferase, and some studies show that a majority of N-linked chains of bovine lung are actually PNGase F resistant! It is not known how common this resistance may be. It is thus important to document N-glycosylation with proteins still in the ER (see Support Protocol) before they might be processed to a PNGase F-resistant form.

Sialidase

Sialidases are also called neuraminidases because the most common form of sialic acid is *N*-acetylneuraminic acid. The sialic acids are a family with over forty different members, but fortunately the very great majority of them can be removed from the oligosaccharides by the broad-spectrum sialidase from *Arthrobacter ureafaciens* (AUS). It can even digest polysialic acids, a rare modification found on only a few proteins such as neuronal cell adhesion molecule (NCAM). Sialidases with selected speci-

ficities from other sources are available but would not usually be needed. AUS has an optimum pH of 5.0, with ~30% of maximum activity at pH 7.

Because sialic acids are charged, they affect gel mobility of proteins more than would be expected from their nominal molecular weight. The magnitude of the gel shift depends on the number of residues. It is difficult to estimate their number by sialidase digestion, but the mobility change is usually sufficient if there are several sialic acid residues. On the other hand, if a protein has only one sialic acid, its presence could be missed using standard one-dimensional SDS-PAGE. To be certain of the effects of sialidase, the sample can be analyzed by a two-dimensional system, using IEF or NEPHGE in the first dimension (UNITS 10.3 & 10.4). The loss of even a single sialic acid will be evident because it changes the isoelectric point.

AUS will remove sialic acids from both N- and O-linked chains, so the type of chain carrying them must be determined independently using PNGase F or possibly O-glycosidase in combination with sialidase. A protein will generally be partially or completely resistant to O-glycosidase because the required disaccharide, Gal β 1-3GalNAc, is usually extended and often sialylated. Until the sialic acid is removed, it will be resistant.

The presence of sialic acid (sialidase sensitivity) is often used as an indication of the transport of a protein into the *trans*-Golgi network (TGN). This may be true in general, but it is important to remember that the distribution of Golgi enzymes is cell type dependent. For instance, α -mannosidases I and II, which are typically considered *cis*/medial Golgi enzymes, are strongly expressed on the brush border of enterocytes—hardly a Golgi compartment. There are other similar examples of various distributions of sialyl transferase. Moreover, there are different sialyl transferases and each may have its own unique distribution. Although one should be cautious, it is probably safe to place sialyl transferase in the late Golgi compartment rather than an early one.

Endo- α -N-acetylgalactosaminidase

This enzyme from *Diplococcus pneumoniae* also goes by various names, including O-glycosidase and O-Glycanase. The last name is a trade name from Genzyme, which no longer sells enzymes; the enzyme is now available from Glyko. This enzyme has a narrow substrate range and cleaves only Gal β 1-3GalNAc α -Ser/Thr. These are only the first two

sugars added in the diverse O-linked pathway that can produce glycans with a dozen or more sugar units. A portion of the pathway is shown in Figure 17.13A.4. Fortunately, many, but far from all, O-linked chains have the simple trisaccharide structure and would be sensitive to cleavage after removing the Sia. Thus, sequential individual digestions or mixed digestions can be used. As both Gal and GalNAc (and probably Sia) are added in the early Golgi complex, sensitivity to the enzyme shows that the protein carries O-linked chains, but matching enzyme sensitivity and a Golgi compartment to further chain extension is difficult. Combining a battery of exoglycosidases (sialidase, α -fucosidase, α -*N*-acetylgalactosaminidase, and β -hexosaminidase) with endo- α -*N*-acetylgalactosaminidase will probably remove most O-linked sugar chains, except sulfated ones. The bottom line is that it is easy to use the enzyme in combination with sialidase to show that a protein has simple O-linked chains, but it is difficult to conclude much more concerning either the structure of the sugar chain or intracellular trafficking.

Endo- β -galactosidase

Bacterioides fragilis endo- β -galactosidase is one of several enzymes that specifically degrade polyactosamines by cleaving linear chains of GlcNAc β 1-3Gal β 1-4 repeats at the Gal β 1-4 linkage. Any substitution on the galactose itself blocks cleavage; however, modifications of the neighboring sugars can slow hydrolysis (Fig. 17.13A.7). For instance, fucosylation and/or sulfation of nearby GlcNAc slows cleavage, but chain branching or sulfation at Gal block it. Even with these potential complexities, digestion with endo- β -galactosidase will sharpen a broad band even if it does not cleave every linkage. The repeating GlcNAc β 1-3Gal β 1-4 units can be found on both N- and O-linked chains, so sensitivity to PNGase F digestion can potentially distinguish the location. Lysosome-associated membrane protein 1 (LAMP-1) has polyactosamine repeats on N-linked chains. Remember that glycosylation is not template driven, so oligosaccharides often exist as a continuum of different structures on individual proteins. For example, heavily sulfated polyactosamine repeats are also known as keratan sulfate and are degraded by keratanases.

O-Sialoglycoprotease

O-Sialoglycoprotease (also called O-glycoprotease or O-sialoglycoprotein endopepti-

dase; Mellors and Lo, 1995) is a neutral metalloprotease produced by *Pasteurella haemolytica*. This enzyme requires clusters of sialylated oligosaccharides on Ser or Thr residues (Norgard et al., 1993). Having a single sialylated O-linked sugar chain on a protein will not lead to degradation, nor will having a nonsialylated sugar chain. Therefore, this enzyme can be used in a typical pulse-chase experiment to indicate whether there are tightly grouped O-linked chains and when they are sialylated. Adding the initial α -GalNAc in the early Golgi complex is insufficient to cause proteolysis; sialylation is specifically required, and this may occur in a later Golgi compartment. Proteolysis can generate smaller fragments of a target protein that are still visible on gels, or the fragments may be so small that they are not even seen on the gels. It depends on the protein. Many leukocyte antigens such as the P-selectin ligand, or others such as CD43, CD44, CD45, and CD34 found on hematopoietic stem cells, are all substrates. As sialic acids are found on both N- and O-linked chains (in clusters and not), sequential digestions using PNGase F, sialidase, and O-sialoglycoprotease endopeptidase in different orders can reveal different kinds of sialylated glycans. If used as part of a pulse-chase protocol, they can reveal different kinetic subcompartments. Not all sialyl transferases are necessarily within the same Golgi compartment.

Critical Parameters

For nearly all of the digestions, complete denaturation of the protein can be important, as maximum deglycosylation occurs only when the sugar chains of glycoproteins are completely exposed. This is not important for sialidase digestions since sialic acids are exposed. Usually endo H digestions will work without full denaturation, but the unprocessed high-mannose chains remain unprocessed because they are often less exposed to the processing enzymes. This may or may not be true of the target protein, but it is better to be safe and denature it completely before digestion. Heating the immunoprecipitate with 0.1 M 2-mercaptoethanol/0.1% SDS is the best way to release the protein in a denatured state. Assuming that the immunoprecipitate contains <10 μ g of protein, 30 μ l of 0.1% SDS still provides a three-fold excess over the protein. Adding SDS presents another problem: free SDS may denature the digesting enzyme before it has a chance to finish its job. For most enzymes, be sure that the free SDS concentration is <0.01%. The best

way to do this is to add a 10- to 20-fold weight excess of nonionic detergent with a low critical micellar concentration (e.g., Triton X-100 or NP-40). These detergents will form mixed micelles with the free SDS and keep it from denaturing the added enzymes.

The amount of enzyme and the incubation time recommended in the protocols are in excess and should be sufficient to cleave any of the sensitive linkages. The incubation times can be shortened, if necessary, but it is better to keep the enzyme concentration as indicated.

Many of the digestions (e.g., sialidase, O-sialoglycoprotease) can be adapted for use on membrane preparations or on live cell surfaces by simply omitting the ionic and nonionic detergents and decreasing the incubation time. The problem is that some linkages may not be exposed and/or sensitive to the digestion. Thus, the usefulness of this approach needs to be determined on a case-by-case basis.

Endo H

Endo H has a broad pH optimum between 5.5 and 6.5, and phosphate or citrate/phosphate buffers can be used in place of citrate. Endo H is very stable to proteases, freezing and thawing, and prolonged incubations. No additives are required for storage of the enzyme. At concentrations below 5 to 10 $\mu\text{g/ml}$ (200 to 400 mU/ml), endo H will bind to glass, so it should be stored in plastic vials (e.g., screw-cap microcentrifuge tubes).

Endo D

Endo D has a broad pH optimum of 4 to 6.5. One unit of enzyme activity will cleave 1 μmole of a $\text{Man}_5\text{GlcNAc}_3$ glycopeptide per min at 37°C. It is supplied from Boehringer Mannheim as a powder containing 0.1 U of activity. Adding 0.1 ml of water gives a 20 mM phosphate buffer, pH 7, 0.05% sodium azide, and 5 mg/ml BSA. It is stable for 3 months at 4°C or at -20°C, but freezing and thawing should be avoided. There may be a slight contamination (<0.2%) with β -N-acetylglucosaminidase activity.

Endo F

Commercial endo F preparations are mixtures of endo F₁, F₂, and F₃. Endo F preparations should not be used for routine deglycosylation or to draw conclusions about the structure of the released oligosaccharides unless the specificity is clearly defined.

Endo F₂

Endo F₂ has a broad pH optimum of 4 to 6 and retains >50% of its activity at pH 7. The enzyme is sensitive to SDS, but adding nonionic detergents prevents denaturation of the enzyme by SDS. Although the enzyme is stable at 4°C for months, it can be frozen in aliquots at -70°C as long as repeated freeze/thaw cycles are avoided. The 1,10-phenanthroline can be used to inhibit a trace of a zinc metalloprotease that may be present.

PNGase F

The pH optimum for PNGase F is 8.6, but 80% of full activity occurs between 7.5 to 9.5 with a range of buffers including phosphate, ammonium bicarbonate, Tris-Cl, and HEPES. Borate buffers inhibit the enzyme. Commercial PNGase F is endo F free and is stable for 6 months at 4°C, or indefinitely at -70°C. However, it should be stored in small aliquots and repeated freeze/thaw cycles should be avoided. PNGase F will bind to glass and plastic surfaces and should not be stored in dilute solutions (<0.1 mU/ml). All of the unit activities of commercial preparations are based on cleavage of dansylated glycopeptides; they are expressed in nmoles/min, which are actually mU, not true International Units (1 International Unit = 1 $\mu\text{mole/min}$). SDS inactivates PNGase F, but adding a ten-fold weight excess of nonionic detergents protects the enzyme (Tarentino et al., 1989; Tarentino and Plummer, 1994).

Sialidase

AUS is available as a lyophilized powder from typical commercial sources such as Sigma, Boehringer Mannheim, and Glyko. It should be reconstituted in water at 1 to 10 mU/ μl according to manufacturer's directions. It is stable for 6 months at 4°C. Treatment with sialidase is also used in assays of protein transport to the cell surface.

Endo- α -N-acetylgalactosaminidase (O-glycosidase)

Endo- α -N-acetylgalactosaminidase has a pH optimum of 6.0 and has 50% activity at 5.5 to 7.0. The thiol inhibitor parachloromercuric benzoate (PCMB; 1 mM) inactivates the enzyme, and 1 mM EDTA inhibits it (63%), as do Mn^{2+} and Zn^{2+} (50%). Chloride also inhibits the enzyme, so HCl-containing buffers should be avoided. The enzyme will have full access to the sugar chains only after denaturing the protein with SDS, but the excess SDS needs to be removed by forming mixed micelles with

nonionic detergents. The enzyme is stable at 4°C and at -20°C, but freeze/thaw cycles should be avoided.

Endo-β-galactosidase

The enzyme is supplied by several commercial sources, including Glyko and Sigma. The enzyme is free of contaminating endo- and exoglycosidases. It has a pH optimum of 5.8 and should be stored at -20°C, but is stable for ≥2 months at 4°C. Glyko supplies the enzyme as a lyophilized powder with BSA for stability. EDTA, Ca²⁺, Mn²⁺, and Mg²⁺ do not affect stability or activity, but PCMB inactivates it.

O-Sialoglycoprotease

The partially purified enzyme is supplied by Accurate Chemical and Scientific as a lyophilized powder containing nonsubstrate bovine serum proteins and HEPES buffer. The enzyme should be reconstituted according to the manufacturer's instructions, divided into aliquots appropriate for a single use, and stored at -20°C. Freeze/thaw cycles inactivate the enzyme, and it is inhibited by EDTA or 1,10-phenanthroline. It is possible to check the activity with a positive control of glycoporphin A, which is available through Sigma.

Troubleshooting

Most of the procedures should work as described, but there is a chance that the enzyme is inactive because of a variety of factors such as age, poor storage, or excess SDS. To check activity, it is worthwhile to run a positive control digestion using the same solutions including SDS and nonionic detergents as for the samples. Since the positive controls are simply glycoproteins that are visualized by Coomassie or silver staining, this requires running a separate gel for staining. This should not be required on a routine basis if the enzymes are used and stored as directed. The most likely culprit in failed digestions is using SDS solutions that are too old or too impure.

Anticipated Results

If the digestions are effective, the labeled band will usually show increased mobility on the gel. In rare instances, digestions can actually decrease mobility. The amount of change will depend on the contribution of that component to the overall mass of the protein. As mentioned before, a gel system that allows visualization of a 1-kDa change should be used. Proteins that are >100 kDa may cause problems for fine resolution. Here are a few numbers to keep in mind.

1. The smallest N-linked chain (Man₃GlcNAc₂) will have a mass of ~0.9 kDa.

2. Two sialic acids on a single N-linked biantennary chain will have a mass of ~0.6 kDa, but their loss may appear larger. If they occur on clustered O-linked chains (sialoglycoprotease sensitive), the apparent size difference will be even larger.

3. Most polylectosamines are three or more repeats, and therefore their mass would be ~1 kDa. The protein will probably run as a heterogeneous smear or broad band before digestion.

4. A single O-glycosidase-sensitive disaccharide (0.4 kDa) may be below detection limits.

Time Considerations

All digestions can be done overnight for convenience, but the amount of enzymes should be sufficient for complete digestion in less time. The gels are run the next day, but the time needed for the development of autoradiograms will depend on the strength of the signal. A low-abundance protein labeled for 10 to 30 min with [³⁵S]Met may give a weak signal and require long exposures (e.g., 2 weeks). Trafficking of abundant glycoproteins such as viral coat proteins require only short exposure times (e.g., a few hours).

Literature Cited

- Alexander, S. and Elder, J.H. 1989. Endoglycosidases from *Flavobacterium meningosepticum*: Application to biological problems. *Methods Enzymol.* 179:505-518.
- Beckers, C.J.M., Keller, D.S., and Balch, W.E. 1987. Semi-intact cells permeable to macromolecules: Use in reconstitution of protein transport from the endoplasmic reticulum to the Golgi complex. *Cell* 50:523-534.
- Chui, D., Oh-Eda, M., Liao, Y.-F., Panneerselvam, K., Lal, A., Marek, K.W., Freeze, H.H., Moremen, K.W., Fukuda, M.N., and Marth, J.D. 1997. Alpha-mannosidase-II deficiency results in dyserythropoiesis and unveils an alternate pathway in oligosaccharide biosynthesis. *Cell* 90:157-167.
- Davidson, H.W. and Balch, W.E. 1993. Differential inhibition of multiple vesicular transport steps between the endoplasmic reticulum and trans Golgi network. *J. Biol. Chem.* 268:4216-4226.
- Elder, J.H. and Alexander, S. 1982. Endo-β-N-acetylglucosaminidase F: Endoglycosidase from *Flavobacterium meningosepticum* that cleaves both high mannose and complex glycoproteins. *Proc. Natl. Acad. Sci. U.S.A.* 79:4540-4544.
- Kornfeld, R. and Kornfeld, S. 1985. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* 54:631-664.

- Mellors, A. and Lo, R.Y. 1995. *O*-Sialoglycoprotease from *Pasteurella haemolytica*. *Methods Enzymol.* 248:728-740.
- Norgard, K.E., Moore, K.L., Diaz, S., Stults, N.L., Ushiyama, S., McEver, R.P., Cummings, R.D., and Varki, A. 1993. Characterization of a specific ligand for P-selectin on myeloid cells. A minor glycoprotein with sialylated *O*-linked oligosaccharides. *J. Biol. Chem.* 268:12764-12774.
- Plummer, T.H. Jr. and Tarentino, A.L. 1991. Purification of the oligosaccharide-cleaving enzymes of *Flavobacterium meningosepticum*. *Glycobiology* 1:257-263.
- Plummer, T.H. Jr., Elder, J.H., Alexander, S., Phelan, A.W., and Tarentino, A.L. 1984. Demonstration of peptide:*N*-glycosidase F activity in endo- β -*N*-acetylglucosaminidase F preparations. *J. Biol. Chem.* 259:10700-10704.
- Tarentino, A.L. and Plummer, T.H. Jr. 1994. Enzymatic deglycosylation of asparagine-linked glycans: Purification, properties, and specificity of oligosaccharide-cleaving enzymes from *Flavobacterium meningosepticum*. *Methods Enzymol.* 230:44-57.
- Tarentino, A.L., Trimble, R.B., and Plummer, T.H. Jr. 1989. Enzymatic approaches for studying the structure, synthesis, and processing of glycoproteins. *Methods Cell Biol.* 32:111-139.
- Tretter, V., Altmann, F., and März, L. 1991. Peptide- N^4 -(*N*-acetyl- β -glucosaminy) asparagine amidase F cannot release glycans with fucose attached $\alpha 1 \rightarrow 3$ to the asparagine-linked *N*-acetylglucosamine residue. *Eur. J. Biochem.* 199:647-652.
- Viitala, J., Carlsson, S.R., Siebert, P.D., and Fukuda, M. 1988. Molecular cloning of cDNAs encoding lamp A, a human lysosomal membrane glycoprotein with apparent Mr approximately equal to 120,000. *Proc. Natl. Acad. Sci. U.S.A.* 85(11):3743-3747.

Key References

- Beckers et al., 1987. See above.
Describes the use of Lec 1 CHO cells and endo D to study processing.
- Chui et al., 1997. See above.
Demonstrates the importance of α -mannosidase III.
- Kornfeld and Kornfeld, 1985. See above.
Landmark review of processing.
- Tarentino and Plummer, 1994. See above.
Best and most recent review of the use of these enzymes.

Contributed by Hudson H. Freeze
The Burnham Institute
La Jolla, California

Analysis of Glycosaminoglycans with Polysaccharide Lyases

Polysaccharide lyases are a class of enzymes useful for analysis of glycosaminoglycans (GAGs) and the glycosaminoglycan component of proteoglycans (PGs). These enzymes cleave specific glycosidic linkages present in acidic polysaccharides and result in depolymerization (Linhardt et al., 1986). These enzymes act through an eliminase mechanism resulting in unsaturated oligosaccharide products that have UV absorbance at 232 nm. The lyases are derived from a wide variety of pathogenic and nonpathogenic bacteria and fungi (Linhardt et al., 1986). This class of enzymes includes heparin lyases (heparinases), heparan sulfate lyases (heparanases or heparitinases), chondroitin lyases (chondroitinases), and hyaluronate lyases (hyaluronidases), all of which are described in this unit.

Polysaccharide lyases can be used, alone or in combinations, to confirm the presence of GAGs in a sample as well as to distinguish between different GAGs (see Table 17.13B.1 and Commentary). The protocols given for heparin lyase I are general and, with minor modifications (described for each lyase and summarized in Table 17.13B.2), can be used for any of the polysaccharide lyases.

The basic protocol describes depolymerization of GAGs in samples containing 1 μg to 1 mg of GAGs. The alternate protocol describes depolymerization of GAGs in samples containing <1 μg of radiolabeled GAG. Two support protocols describe assays to confirm and quantitate the activity of heparin and chondroitin ABC lyases. It is recommended that enzyme activity be assayed before the enzyme is used in an experiment to be sure it is active and has been stored properly.

The standard definition of a unit (U), 1 μmol product formed/min, is used throughout this article. Some lyases are sold in nonstandard units (e.g., 0.1 $\mu\text{mol/hr}$, $\Delta A_{232}/\text{min}$), and these should either be converted to standard units or the activity should be determined using the appropriate support protocol.

Table 17.13B.1 Polysaccharide Lyases Used to Identify Glycosaminoglycans

Glycosaminoglycan	Enzyme
Heparin, heparan sulfate, chondroitin sulfate, dermatan sulfate, hyaluronate	Heparin lyases I, II, III, and chondroitin ABC, AC lyases
Heparin, heparan sulfate, chondroitin sulfate, dermatan sulfate	Heparin lyases I, II, III, and chondroitin ABC lyase
Heparin, heparan sulfate	Heparin lyases I, II, III
Chondroitin sulfate, dermatan sulfate, hyaluronate	Chondroitin ABC, AC lyases
Chondroitin sulfate, dermatan sulfate	Chondroitin ABC lyase
Heparin	(Heparin lyase I–heparin lyase III) ^a
Heparan sulfate	Heparin lyase III
Chondroitin sulfate	Chondroitin AC lyase (or chondroitin ABC lyase–chondroitin B lyase) ^a
Dermatan sulfate	Chondroitin B lyase (or chondroitin ABC lyase–chondroitin AC lyase) ^a
Hyaluronate	Hyaluronate lyase

^aSample is divided into two portions and each is treated with a different lyase. The amount of depolymerization (i.e., counts moving from V_o to V_t of a gel-filtration column) is determined for each portion, and the difference gives a measure of the amount of glycosaminoglycan.

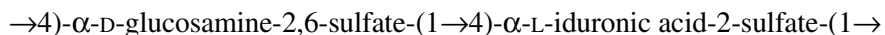
Contributed by Robert J. Linhardt

Current Protocols in Molecular Biology (1999) 17.13B.1-17.13B.16

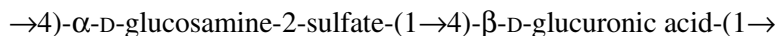
Copyright © 1999 by John Wiley & Sons, Inc.

OVERVIEW OF HEPARIN LYASES

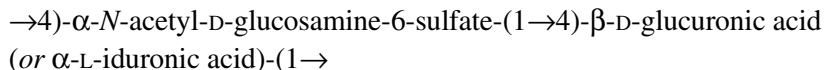
There are three well-characterized polysaccharide lyases that act endolytically on heparin and heparan sulfate (Jandik et al., 1994); these enzymes are called heparin lyases. Heparin and heparan sulfate GAGs are structurally related, linear sulfated polysaccharides. Heparin's major sequence (representing 70% to 90% of its structure) is as follows:



Heparan sulfate is composed primarily of equal proportions of the following:



and



These disaccharide sequences are found in differing amounts in both heparin and heparan sulfate (Desai et al., 1993a,b). Substrates for heparin lyases are illustrated in Figure 17.13B.1.

The nomenclature of the three heparin lyases is somewhat confusing. However, heparin lyase I and III have enzyme commission (EC) numbers to facilitate their identification. The decision as to which enzyme should be used for a particular application is based on both the specificity desired and the reaction conditions required. The activity of these lyases toward specific glycosidic linkages has been determined using structurally characterized oligosaccharide substrates (Desai et al., 1993a). The primary linkages cleaved by these enzymes and their relative activities toward heparin and heparin sulfate (Desai et al., 1993a, b) are presented in Figure 17.13B.1 and Table 17.13B.3. Information on the optimal conditions for the activity and stability for these enzymes is given under the description of each enzyme and summarized in Table 17.13B.2.

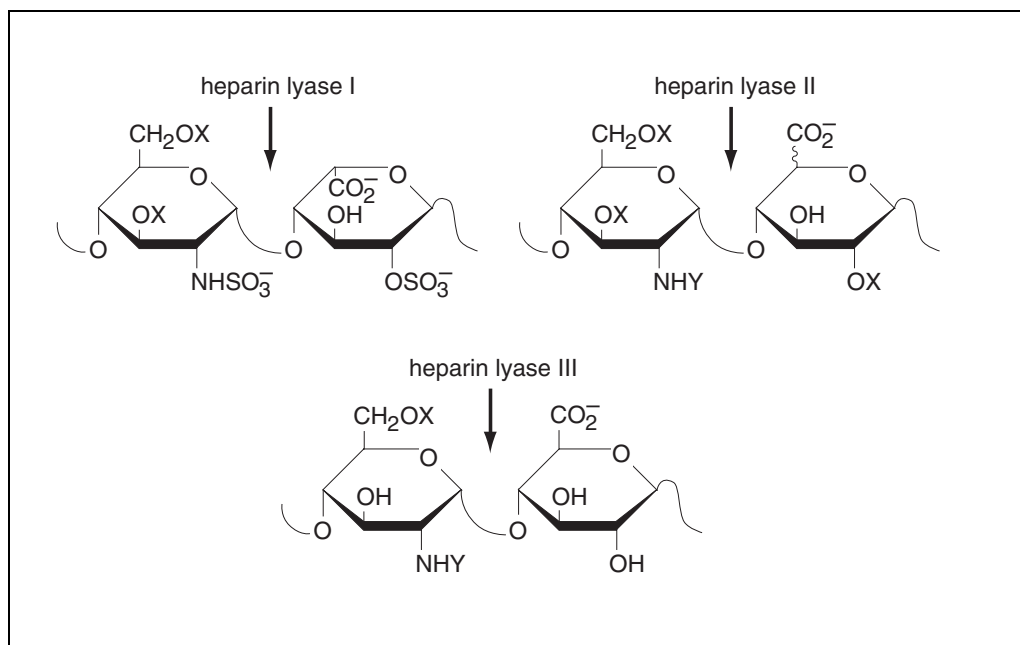


Figure 17.13B.1 Primary glycosidic linkages cleaved by heparin lyases. Abbreviations: X, H or SO_3^- ; Y, CH_3CO or SO_3^- . Heparin lyase II cleaves at either glucuronic or iduronic acid residues.

Table 17.13B.2 Reaction Conditions for Polysaccharide Lyases with Optimum Buffers and Reaction Temperatures

Lyase	Buffer ^a	Optimum temperature
Heparin I	Sodium phosphate/NaCl, pH 7.1	30°C
Heparin II	Sodium phosphate, pH 7.1	35°C
Heparin III	Sodium phosphate, pH 7.6	35°C
Chondroitin ABC	Tris-Cl/sodium acetate, pH 8	37°C
Chondroitin AC	Tris-Cl/sodium acetate, pH 8	37°C
Chondroitin B	Ethylenediamine/acetic acid/NaCl, pH 8	25°C
Hyaluronate	Sodium acetate/NaCl, pH 5.2	>30°C

^a See Reagents and Solutions for buffer recipes.

Table 17.13B.3 Activity of Heparin Lyases

Activity and substrate conversion	Heparin lyase I	Heparin lyase II	Heparin lyase III
<i>Heparin</i> ^a			
% Activity ^b	100	58	<1
% Conversion ^c	58 (76) ^d	85	6
<i>Heparan sulfate</i> ^e			
% Activity	13	100	100
% Conversion	19	39	94

^aPorcine mucosal heparin.

^b Percent activity = [initial rate on the substrate examined/initial rate on substrate giving the highest activity] × (100).

^cPercent conversion = [moles of linkages cleaved/total moles of hexosamine→uronic acid linkages] × (100).

^dBovine lung heparin.

^eBovine kidney heparan sulfate.

HEPARIN LYASE I (Lohse and Linhardt, 1992)

Heparin lyase I (EC 4.2.2.7), from *Flavobacterium heparinum* (*Cytophagia heparinia*), is commonly referred to as heparinase. The enzyme has a molecular weight of 42,800 Da and a pI of 9.1 to 9.2. Heparin lyase I has a random endolytic action pattern—i.e., it randomly acts on any site with the appropriate primary structure within the polymeric substrate (Fig. 17.13B.1; Jandik et al., 1994).

Complete Heparin Lyase–Catalyzed Depolymerization of an Unlabeled Sample

Samples consisting of tissues, biological fluids, PGs, and GAGs (UNIT 17.2) that contain microgram quantities of heparin and are not metabolically labeled can be analyzed using heparin lyase (see Critical Parameters for method to distinguish between heparin and heparan sulfate).

ENZYME

BASIC PROTOCOL

Preparation and Analysis of Glycoconjugates

17.13B.3

Materials

Heparin- or heparan sulfate-containing sample
Sodium phosphate/NaCl buffer (see recipe)
Heparin lyase I solution (see recipe)

Spectropor dialysis membrane, MWCO 1000 (Spectrum)
500- μ l polypropylene microcentrifuge tubes
30° and 100°C water baths

Additional reagents and equipment for polysaccharide dialysis (APPENDIX 3D),
HPLC (UNIT 17.18), and gel-filtration chromatography (UNITS 10.9 & 17.17)

1. Dissolve sample, containing 1 μ g to 1 mg heparin, in 50 μ l sodium phosphate/NaCl buffer. Dialyze sample against sodium phosphate/NaCl buffer using 1000 MWCO dialysis membrane.
2. Thaw and assay activity of a frozen aliquot of enzyme.
3. Thaw 10 μ l heparin lyase I solution at room temperature and add 40 μ l sodium phosphate/NaCl buffer to the 500- μ l tube containing enzyme. Add 50 μ l sodium phosphate/NaCl buffer to another 500- μ l tube to serve as a blank control.

Additional enzyme (10- to 100-fold) may be required to break down small, resistant oligosaccharides (Rice and Linhardt, 1989; Desai et al., 1993a).

For heparin lyase II and III, the chondroitin lyases, and hyaluronate lyase, see alternative conditions listed in Table 17.13B.2 and in descriptions of individual enzymes below.

4. Add 50 μ l sample to each tube and incubate 8 to 12 hr at 30°C.
5. Heat tube 2 min at 100°C to terminate the reaction. Analyze the product by a method appropriate for its purity and concentration.

A pure sample containing >10 μ g of heparin can be analyzed by measuring the difference in absorbance at 232 nm (ΔA_{232}) in 30 mM HCl between enzyme-treated sample and blank ($\epsilon = 5500 M^{-1}$ for oligosaccharide products in 30 mM HCl). Enzyme-treated sample is diluted with measured amounts of 30 mM HCl until the A_{232} is between 1 and 2. The blank is diluted with the same quantity of 30 mM HCl and its A_{232} is measured. The difference between the two measured A_{232} values, ΔA_{232} , is used to calculate the moles of oligosaccharide product formed. Ten moles of oligosaccharide product is obtained for each mole of heparin. If the treated sample contains substantial amounts of protein or other substances that absorb at or near 232 nm, disappearance of polysaccharide substrate can be measured using a dye-binding assay (Grant et al., 1984). Smaller quantities of samples or samples of lower purity can be analyzed by HPLC or gel-filtration chromatography using UV or conductivity detection methods.

Complete Heparin Lyase–Catalyzed Depolymerization of Very Small Amounts of Radiolabeled Glycosaminoglycans

ALTERNATE
PROTOCOL

When attempting to use heparin lyase to depolymerize radiolabeled samples that contain very small quantities of heparin, it is often useful to add cold substrate as a carrier so the activity of heparin lyase can be distinguished from that of trace amounts of chondroitin lyases that are often present in heparin lyase preparations. Chondroitin lyase may pose a problem when using a heparin lyase to distinguish between heparin/heparan sulfate and chondroitin/dermatan sulfate in radiolabeled samples (see Critical Parameters and Table 17.13B.3). Therefore, it is recommended that cold carrier chondroitin/dermatan sulfate be added to block the action of minor chondroitin lyase contaminants. Alternatively, chondroitin ABC lyase from *Proteus vulgaris* (an organism free of enzymes acting on heparin and heparan sulfate) can be used to detect chondroitin sulfate chains specifically (see discussion of chondroitin lyases).

Additional Materials (also see Basic Protocol)

Radiolabeled heparin-containing sample (UNIT 17.4)
20 mg/ml chondroitin sulfate A solution (see recipe)
20 mg/ml chondroitin sulfate C solution (see recipe)
20 mg/ml dermatan sulfate solution (see recipe)

1. Dissolve GAG sample containing radiolabeled heparin in 50 μ l sodium phosphate/NaCl buffer. Dialyze sample against sodium phosphate/NaCl buffer using 1000 MWCO dialysis membrane.
2. Thaw 10 μ l of heparin lyase I solution at room temperature, immediately prior to use.
For heparin lyase II and III, the chondroitin lyases, and hyaluronate lyase, see alternative conditions listed in Table 17.13B.2 and in descriptions of individual enzymes below.

3. Add 30 μ l sodium phosphate/NaCl buffer to the 500- μ l tube containing enzyme.
4. *Optional:* Add 1.7 μ l each of 20 mg/ml chondroitin sulfate A, chondroitin sulfate C, and dermatan sulfate substrate solutions (34 μ g each) to the enzyme in buffer.

This step should be performed if there is any concern that the enzyme contains chondroitin lyase impurities.

5. Add 50 μ l radiolabeled heparin solution and incubate 8 to 12 hr at 30°C.
6. Heat 2 min at 100°C to terminate the reaction.
7. Analyze depolymerized radioactive sample by HPLC or gel-filtration chromatography using radioisotope detection methods.

In gel-filtration chromatography following treatment with heparin lyase, counts in fractions corresponding to an apparent molecular weight <1500 Da confirm the presence of heparin/heparan sulfate. Similar results are obtained with heparin lyase II and III except heparin lyase III does not act on heparin (Table 17.13B.3).

Preparation and
Analysis of
Glycoconjugates

17.13B.5

Assay of Heparin Lyase Activity

Commercial preparations of heparin lyase should be assayed before use, particularly for applications in which the level of enzyme activity is critical—e.g., when trying to distinguish between heparin and heparan sulfate or measure their relative amounts in a sample containing both.

Materials

Sodium phosphate/NaCl buffer (see recipe)
Heparin lyase I solution (see recipe)
20 mg/ml heparin solution (see recipe)
UV spectrophotometer, temperature controlled
1-ml quartz cuvette with 1-cm pathlength

Reaction conditions

A 700- μ l reaction should contain:
50 mM sodium phosphate, pH 7.1
100 mM sodium chloride
10 mU heparin lyase I
1 mg heparin

Protocol

1. Add 640 μ l sodium phosphate/NaCl buffer to a 1-ml cuvette. Warm the cuvette to 30°C in a temperature-controlled UV spectrophotometer.

If a temperature-controlled spectrophotometer is not available, activity can be measured at room temperature, or sample can be incubated in a water bath and absorbance measured at fixed time points.

2. Thaw a 10- μ l aliquot of heparin lyase I solution at room temperature.
3. Take cuvette out of the spectrophotometer, remove 90 μ l warm buffer and transfer it to the tube containing heparin lyase I solution. Immediately transfer the entire 100 μ l of buffer and enzyme back into the warm cuvette.
4. Place cuvette in the spectrophotometer and set the A_{232} to zero.
5. Remove cuvette from the spectrophotometer and add 50 μ l of 20 mg/ml heparin substrate solution to initiate the reaction. Seal cuvette with Parafilm and invert once or twice to mix. Remove Parafilm and return cuvette to spectrophotometer.
6. Within 30 sec after addition of substrate, begin to measure absorbance continuously or at 30 sec intervals for 2 to 10 min. Plot A_{232} versus time.

At room temperature (~20°C) a two-fold decrease in reaction rate is observed; this requires a 4- to 20-min assay time.

7. Calculate the enzyme activity (1 U = 1 μ mol product formed/min) from the initial rate (<5% reaction completion) using $\epsilon = 3800 \text{ M}^{-1}$ for the reaction products in sodium phosphate/NaCl buffer. Each product formed has an unsaturated uronic acid residue at its nonreducing terminus. Enzyme activity is calculated as

$$\text{Enzyme activity} = \frac{(\Delta A_{232}/\text{min}) (700 \mu\text{l})}{3800 \text{ M}^{-1}}$$

The slope of the linear portion of the curve is used to calculate the initial rate of reaction.

HEPARIN LYASE II (Lohse and Linhardt, 1992)

Heparin lyase II (no EC number), an endolytic enzyme from *Flavobacterium heparinum*, has a molecular weight of 84,100 Da and a pI of 8.9 to 9.1. The substrate for heparin lyase II is indicated in Table 17.13B.3 and Figure 17.13B.1.

Protocols for the assay and use of heparin lyase II are identical to those for heparin lyase I, with the following modifications (see Table 17.13B.2):

1. Prepare enzyme and substrate solutions in sodium phosphate buffer, pH 7.1 (see recipes).
2. Carry out the reaction at 35°C.
3. Use 20 mg/ml heparan sulfate solution (see recipe) in Support Protocol 1 to assay for enzyme activity.

HEPARIN LYASE III (Lohse and Linhardt, 1992)

Heparin lyase III (EC 4.2.2.8), an endolytic enzyme from *Flavobacterium heparinum*, is commonly referred to as heparitinase and has a molecular weight of 70,800 Da and a pI of 9.9 to 10.1. Heparin lyase III can be used to confirm the presence of heparan sulfate in a sample (Fig. 17.13B.1 and Tables 17.13B.1 and 17.13B.3).

Protocols for the assay and use of heparin lyase III are identical to those of heparin lyase I, with the following modifications (see Table 17.13B.2):

1. Prepare enzyme and substrate solutions in sodium phosphate buffer, pH 7.6 (see recipes).
2. Carry out the reaction at 35°C.
3. Use 20 mg/ml heparan sulfate solution (see recipe), instead of heparin, as the substrate to assay for enzyme activity in Support Protocol 1.

OVERVIEW OF CHONDROITIN SULFATE LYASES

There are several polysaccharide lyases that act on chondroitin sulfates, dermatan sulfate, and hyaluronate. Chondroitin sulfate galactosaminoglycans are structurally related, sulfated, alternating 1→3, 1→4 linked, linear polysaccharides. The structure of the major disaccharide linkage found in each chondroitin sulfate and the enzyme that acts at each linkage are shown in Figure 17.13B.2. Hyaluronate has the same backbone structure, except that it is not sulfated and contains *N*-acetylglucosamine in place of *N*-acetylgalactosamine (Fig. 17.13B.3).

The decision of which chondroitin lyase to use should be based on both the specificity desired and the reaction conditions required (Tables 17.13B.1 & 17.13B.2).

CHONDROITIN ABC LYASE (Yamagata et al., 1968)

Chondroitin ABC lyase (chondroitinase ABC, EC 4.2.2.4), from *Proteus vulgaris*, has a molecular weight of 150,000 Da. This enzyme acts endolytically on chondroitin sulfates A-E (Fig. 17.13B.2; Jandik et al., 1994), slowly on hyaluronic acid, and not at all on heparin, heparan sulfate, or keratan sulfate.

ENZYME

ENZYME

ENZYME

**Preparation and
Analysis of
Glycoconjugates**

17.13B.7

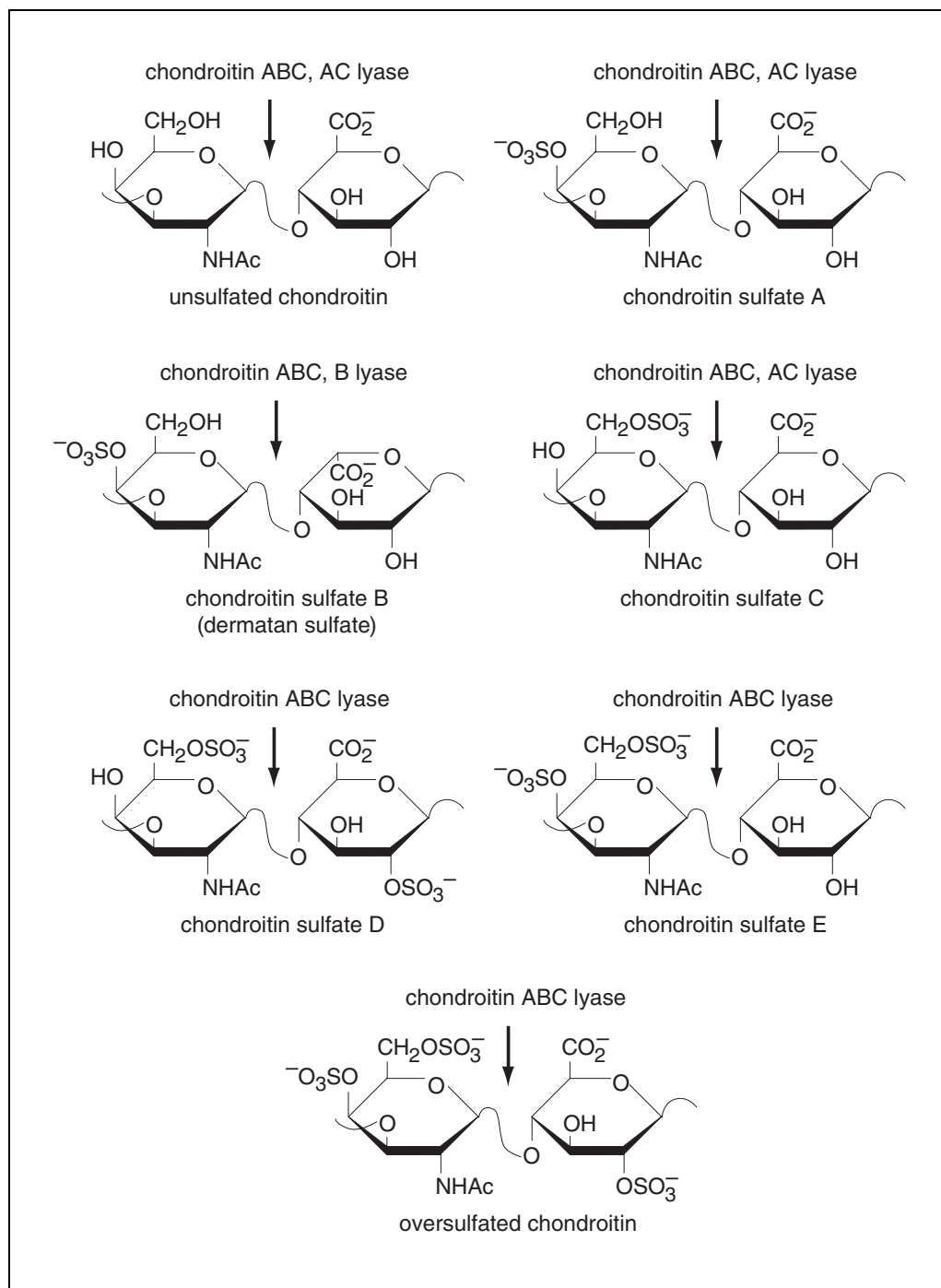


Figure 17.13B.2 Glycosidic linkages cleaved by chondroitin lyases. Abbreviation: Ac, CH_3CO .

The basic and alternate protocols for chondroitin ABC lyase are identical to those described for heparin lyase I with the following modifications:

1. Prepare enzyme and substrate solutions in Tris-Cl/sodium acetate buffer, pH 8 (see recipe).
2. Carry out the reaction at 37°C .
3. When using the alternate protocol for depolymerization of very small amounts of radiolabeled GAGs, omit step 4.

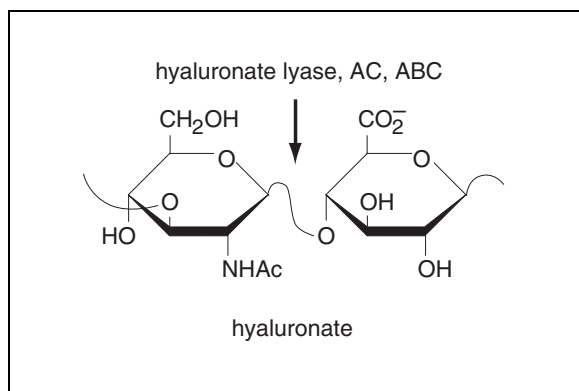


Figure 17.13B.3 Glycosidic linkage in hyaluronate cleaved by hyaluronate lyase and chondroitin AC and ABC lyases. Abbreviation: Ac, CH₃CO.

A radioactive sample containing chondroitin sulfate or dermatan sulfate can be analyzed, following chondroitin ABC lyase treatment, by HPLC (UNIT 17.18) or gel-filtration chromatography (UNITS 10.9 & 17.17) using radioisotope detection methods. In gel-filtration chromatography following treatment with chondroitin ABC lyase, counts in fractions corresponding to a molecular weight <1000 Da confirm the presence of chondroitin sulfate or dermatan sulfate.

Assay of Chondroitin ABC Lyase Activity

Commercial preparations of chondroitin lyase should be assayed before use, particularly for applications in which the level of enzyme activity is critical.

Materials

- Tris·Cl/sodium acetate buffer, pH 8.0 (see recipe)
- Chondroitin ABC lyase solution (see recipe)
- 20 mg/ml chondroitin sulfate A (see recipe), chondroitin sulfate C solution (see recipe), or dermatan sulfate solution (see recipe)
- UV spectrophotometer, temperature controlled
- 1-ml quartz cuvette with 1-cm path length

Reaction conditions

- A 700- μ l reaction should contain:*
- 50 mM Tris·Cl, pH 8
 - 60 mM sodium acetate
 - 10 mU chondroitin lyase
 - 1 mg chondroitin sulfate A or C, or dermatan sulfate

Protocol

1. Add 640 μ l Tris·Cl/sodium acetate buffer to a 1-ml quartz cuvette. Warm the cuvette to 37°C in a temperature-controlled spectrophotometer.

If a temperature-controlled spectrophotometer is not available, activity can be measured at room temperature or the sample can be incubated in a water bath and the absorbance measured at fixed time points.

2. Thaw a 10- μ l aliquot of chondroitin ABC lyase solution at room temperature.
3. Take cuvette out of the spectrophotometer, remove 90 μ l warm buffer, and transfer it to enzyme solution. Immediately transfer entire 100 μ l buffer plus enzyme back to the warm cuvette.
4. Place cuvette in spectrophotometer and set the absorbance at 232 nm (A_{232}) to zero.

SUPPORT PROTOCOL 2

Preparation and Analysis of Glycoconjugates

17.13B.9

5. Remove cuvette from spectrophotometer and add 50 μl of 20 mg/ml chondroitin A or C or dermatan sulfate solution to initiate reaction. Seal cuvette with Parafilm and invert once or twice to mix. Remove Parafilm and return cuvette to spectrophotometer.

To assay for chondroitin AC lyase activity, use chondroitin A or C as substrate. To assay for chondroitin B lyase activity, use dermatan sulfate as substrate (see Table 17.13B.1 and description of individual enzymes below).

6. Within 30 sec after adding substrate begin to measure the absorbance continuously or at 30-sec intervals for 2 to 10 min. Plot A_{232} versus time.
7. Calculate the enzyme activity (1 U = 1 μmol product formed/min) from the initial rate (<5% reaction completion) using $\epsilon = 3800 \text{ M}^{-1}$ for reaction products at pH 8. Enzyme activity is calculated as

$$\text{Enzyme activity} = \frac{(\Delta A_{232}/\text{min})(700 \mu\text{l})}{3800 \text{ M}^{-1}}$$

Calculate the number of product molecules formed per substrate molecule from the A_{232} measured at reaction completion.

ENZYME

CHONDROITIN AC LYASE (Yamagata et al., 1968; Hiyama and Okada, 1975; Michelacci and Dietrich, 1975)

Two chondroitin AC lyases from *Arthrobacter aurescens* and *Flavobacterium heparinum* share the same enzyme commission number (EC 4.2.2.5). Chondroitin AC lyase from *A. aurescens* has a molecular weight of 76,000 Da (Hiyama and Okada, 1975) and a pI of 5.46. It acts exolytically on chondroitin sulfate A and C (Jandik et al., 1994) and exhibits a three-fold higher activity on hyaluronate. It can act at the glucuronic acid residues in dermatan sulfate (Fig. 17.13B.2) but does not act on heparin or heparan sulfate (Fig. 17.13B.1). Chondroitin AC lyase from *F. heparinum* acts endolytically on chondroitin sulfates A and C (Jandik et al., 1994), on hyaluronate, and at the glucuronate residues of dermatan sulfate (Fig. 17.13B.2; Gu et al., 1993). It does not act on heparin, heparan sulfate, or keratan sulfate (Yamagata et al., 1968).

The protocols for chondroitin AC lyase are identical to those described for chondroitin ABC lyase except chondroitin A or C should be used as substrate in Support Protocol 2 for assaying chondroitin lyase activity.

ENZYME

CHONDROITIN B LYASE (Michelacci and Dietrich, 1975)

Chondroitin B lyase (no EC number), from *Flavobacterium heparinum*, has a molecular weight of 55,000 Da. Chondroitin B lyase acts only on dermatan sulfate and not on the glucuronate residues of chondroitin sulfates A, C, and hyaluronate, and not on heparin or heparan sulfate (Fig. 17.13B.2).

This enzyme can be assayed as described for chondroitin ABC lyase using dermatan sulfate as substrate. The heparin lyase I protocols are used for sample analysis with this enzyme, with the following modifications (see Table 17.13B.2):

1. Prepare enzyme and substrate solutions in ethylenediamine/acetic acid/NaCl buffer, pH 8 (see recipe).
2. Carry out the reaction at 25°C.
3. Use dermatan sulfate solution (see recipe) as substrate in Support Protocol 2.

HYALURONATE LYASE (Vesterberg, 1968; Rautela and Abramson, 1973; Sasaki et al., 1982)

ENZYME

Hyaluronate lyases (EC 4.2.2.1 and EC 4.2.99.1) act only on hyaluronate (Hiyama and Okada, 1975) allowing hyaluronate (Fig. 17.13B.3) to be distinguished from chondroitin sulfate (Fig. 17.13B.2 and Table 17.13B.1). The enzyme from *Streptomyces hyalurolyticus* has been purified and has a pI <7 (Sasaki et al., 1982). Homogeneous *Staphylococcus aureus* hyaluronate lyase has a molecular weight of 84,000 Da (Rautela and Abramson, 1973) and a pI of 7.4 to 7.9 (Vesterberg, 1968).

Reaction conditions

A 700- μ l reaction should contain:

- 50 mM sodium acetate, pH 5.2
- 125 mM sodium chloride
- 10 mU hyaluronate lyase
- 1 mg hyaluronate

The basic, alternate, and support protocols for hyaluronate lyase are identical to those described for heparin lyase with the following modifications (see Table 17.13B.2):

Protocol

1. Use sodium acetate/NaCl buffer, pH 5.2 (see recipe), to prepare enzyme (hyaluronate lyase; see recipe) and substrate solutions.
2. Perform the reaction at 30° to 60°C.

Higher temperatures can be useful to reduce substrate viscosity—hyaluronic acid is very viscous. Higher temperatures will also increase the reaction rate, and they can be used to inhibit the activity of other lyases that may be present as contaminants of the hyaluronate lyase preparation.

3. Use hyaluronate solution (see recipe), instead of heparin solution, as substrate to assay enzyme activity (see Support Protocol 1).
4. When using the Alternate Protocol for depolymerization of very small amounts of radiolabeled GAGs, omit step 4.

Following treatment with hyaluronate lyase, a radioactive sample containing hyaluronate can be analyzed by HPLC (UNIT 17.18) or gel-filtration chromatography (UNIT 17.17) using radioisotope detection methods. In gel-filtration following treatment with hyaluronate lyase, counts in fractions corresponding to a molecular weight <500 Da confirm the presence of hyaluronate.

Testicular hyaluronidase is a hydrolase (EC 3.2.1.35) that also acts on hyaluronate (also on chondroitin sulfates A and C). Because it is not a polysaccharide lyase, testicular hyaluronidase cannot be assayed with the support protocol, but it can be used in the basic and alternate protocols with the modifications described above.

Preparation and
Analysis of
Glycoconjugates

17.13B.11

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Chondroitin lyase solutions, 1 mU/ μ l

Dissolve 0.1 U lyophilized enzyme in 100 μ l Tris-Cl/sodium acetate buffer (see recipe) for chondroitin ABC lyase (chondroitinase ABC; Sigma or Seikagaku) and chondroitin AC lyase (chondroitinase AC from *Arthrobacter aurescens* and *Flavobacterium heparinum*; Sigma or Seikagaku) or ethylenediamine/acetic acid/NaCl buffer (see recipe) for chondroitin B lyase (chondroitinase B; Sigma or Seikagaku). Store in 10-mU aliquots in 500- μ l polypropylene tubes <1 year at -70°C .

Different suppliers use different definitions for units of enzyme activity. Enzyme activity levels should be assayed and standardized, as described in the support protocols, in the laboratory before the enzymes are used in analytical studies.

Chondroitin sulfate A or C solution, 20 mg/ml

Dissolve 20 mg chondroitin sulfate A or C, sodium salt (Sigma or Seikagaku) in 1 ml Tris-Cl/sodium acetate buffer, pH 8 (see recipe) or other buffer appropriate to enzyme/substrate pair (see Table 17.13B.2). Store 1 year at $<0^{\circ}\text{C}$.

Dermatan sulfate solution, 20 mg/ml

Dissolve 20 mg dermatan sulfate, sodium salt (Sigma or Seikagaku) in 1 ml ethylenediamine/acetic acid/NaCl buffer, pH 8 (see recipe) or other buffer appropriate to enzyme/substrate pair (see Table 17.13B.2). Store 1 year at $<0^{\circ}\text{C}$.

Ethylenediamine/acetic acid/NaCl buffer

Dissolve 3.0 g ethylenediamine (50 mM final) and 1.7 g NaCl (30 mM final) in 900 ml H_2O . Adjust pH with glacial acetic acid to pH 8, and bring volume to 1 liter with H_2O . Store <1 month at 4°C .

Heparan sulfate solution, 20 mg/ml

Dissolve 20 mg heparan sulfate, sodium salt from bovine kidney (Seikagaku) in 1 ml sodium phosphate buffer, pH 7.6 (see recipe). Store 1 year at 0°C .

Heparin lyase solutions, 1 mU/ μ l

Dissolve 0.1 U lyophilized enzyme in 100 μ l sodium phosphate/NaCl buffer (see recipe) for heparin lyase I or sodium phosphate buffer (see recipe) for heparin lyase II and III adjusted to the appropriate pH. Store enzyme in 10 mU aliquots in 500- μ l polypropylene tubes <1 year at -70°C .

*Heparin lyase I from *Flavobacterium heparinum* is sold as heparinase I (Sigma) and heparinase (Seikagaku). Heparin lyase II is sold as heparinase II (Sigma) and heparitinase II (Seikagaku). Heparin lyase III from *Flavobacterium heparinum* is sold as heparinase III (Sigma) and heparatinase or heparatinase I (Seikagaku).*

Different suppliers use different definitions for units of enzyme activity. Enzyme activity levels should be assayed and standardized, as described in the support protocols, in the laboratory before the enzymes are used in analytical studies.

Heparin solution, 20 mg/ml

Dissolve 20 mg heparin, sodium salt (140 to 180 USP U/mg), from porcine intestinal mucosa or bovine lung (Sigma) in sodium phosphate/NaCl buffer, pH 7.1 (see recipe). Store 1 year at 0°C .

Different suppliers use different definitions for units of enzyme activity. Enzyme activity levels should be assayed and standardized, as described in the support protocols, in the laboratory before the enzymes are used in analytical studies.

Hyaluronate lyase solution, 1 mU/μl

Dissolve 0.1 U lyophilized enzyme (hyaluronidase from *Streptomyces hyalurolyticus*; Sigma or Seikagaku) in 100 μl sodium acetate/NaCl buffer, pH 5.2 (see recipe). Aliquot 10 mU to 500-μl polypropylene tubes and store <1 year at -70°C.

Different suppliers use different definitions for units of enzyme activity. Enzyme activity levels should be assayed and standardized, as described in the support protocols, in the laboratory before the enzymes are used in analytical studies.

Hyaluronate solution, 1.4 mg/ml

Dissolve 1.4 mg hyaluronate in 1 ml sodium acetate/NaCl buffer, pH 5.2 (see recipe). Store 1 year at <0°C.

Hyaluronate of high molecular weight is very viscous. It is best to use a hyaluronate of low to medium molecular weight to measure activity.

Sodium acetate/NaCl buffer

Dissolve 4.1 g sodium acetate (30 mM final) and 7.3 g NaCl (125 mM final) in 900 ml H₂O. Adjust pH with glacial acetic acid to pH 5.2 and bring volume to 1 liter with H₂O. Store <1 month at 4°C.

Sodium phosphate buffer

Dissolve 7.1 g dibasic sodium phosphate (50 mM final) in 900 ml H₂O. Adjust pH with concentrated phosphoric acid to pH 7.1 for heparin lyase II and pH 7.6 for heparin lyase III. Bring volume to 1 liter with H₂O. Store <1 month at 25°C.

Sodium phosphate/NaCl buffer

Dissolve 7.1 g dibasic sodium phosphate (50 mM final) and 5.8 g sodium chloride (100 mM final) in 900 ml H₂O. Adjust pH with phosphoric acid to pH 7.1 and bring volume to 1 liter with H₂O. Store <1 month at 25°C.

Tris·Cl/sodium acetate buffer

Dissolve 6.05 g Tris base (50 mM final) and 8.17 g sodium acetate (60 mM final) in 900 ml H₂O. Adjust pH to 8.0 with concentrated hydrochloric acid, and bring volume to 1 liter with H₂O. Store <1 month at 4°C.

COMMENTARY

Background Information

Proteoglycans (PGs) are primarily found in the extracellular matrix and are important in cell-cell interaction. Glycosaminoglycans (GAGs) are the dominant physical, chemical, and biological features of PGs. Lyases cleave specific glycosidic linkages at C4 of uronic acid residues present in GAGs (Figs. 17.13B.1, 17.13B.2, and 17.13B.3) through an eliminase mechanism, resulting in unsaturated oligosaccharide products that have UV absorbance spectra maxima at 232 nm (Linhardt et al., 1986). An eliminase catalyzes an elimination reaction that results in the formation of a double bond. In contrast, a hydrolase (the most common class of enzymes acting on polysaccharides) breaks down polysaccharides through the addition of water and does not result in the formation of a double bond. Lyases are useful for identifying and distinguishing the GAGs

present in unlabeled or radiolabeled samples (see Table 17.13B.1).

Critical Parameters and Troubleshooting

Lyases are primarily of microbial origin and often a single organism produces multiple lyases acting on a variety of PGs and GAGs. Thus, it must be recognized that despite the high level of purity of the commercially available enzymes, they may contain small amounts of enzymatic impurities. These impurities can cause misleading results, particularly when large quantities of enzyme are used to treat very small quantities of sample—e.g., when these enzymes are used to analyze radiolabeled samples.

Protease contamination can also be present in the enzyme preparation. Commercial enzymes often contain bovine serum albumin

(BSA) for stabilization during lyophilization, and this greatly reduces potential problems associated with proteolytic contamination.

Complete degradation of a GAG. Table 17.13B.1 shows how the polysaccharide lyases can be used, alone or in combination, to determine the content of a GAG. All three heparin lyases used in combination (at equal-unit concentrations in 50 mM sodium phosphate buffer, pH 7, at 30°C) will degrade heparin or heparan sulfate to disaccharides, although a small quantity of lyase-resistant tetrasaccharide may remain. Chondroitin ABC lyase can be used to completely digest a mixture of chondroitin sulfates. If hyaluronate and chondroitin sulfate are both present, it is advisable to use an equal-unit mixture of chondroitin ABC and AC lyases. For complete degradation of GAGs, except keratan sulfate (see UNIT 17.2), it is advisable to treat the sample with an equal-unit mixture of heparin lyases I, II, and III and chondroitin ABC and AC lyases in sodium phosphate/NaCl buffer, pH 7, 30°C.

Heparin lyase III can be used to distinguish between samples containing heparan sulfate and those containing heparin. Although heparin lyase I (heparinase) has been used by investigators to demonstrate the presence of heparin in a sample, it also acts at some linkages present in heparan sulfate (Table 17.13B.3). Gel-filtration analysis can help distinguish between heparin and heparan sulfate because the large oligosaccharides formed from heparan sulfate elute in the void volume. In addition, oligosaccharide product compositional analysis (UNIT 17.19) is useful in distinguishing heparin from heparan sulfate when using heparin lyase I.

Caution is required when attempting to distinguish heparin/heparan sulfate from chondroitin/dermatan sulfate using heparin lyases. *Flavobacterium heparinum* produces both heparin and chondroitin/dermatan sulfate lyases; thus, minor contaminating activities can result in false positives. Step 4 of the Alternate Protocol for the use of heparin lyase on radiolabeled samples—incubation in the presence of nonspecific substrate—is included to eliminate false positives. The use of chondroitin ABC lyase, which contains no heparin lyase activity, is complicated by heparin's inhibition of chondroitin ABC lyase (Nakada and Wolfe, 1961). This inhibition is overcome by using excess enzyme (Linhardt et al., 1991).

Chondroitin sulfate is routinely distinguished from dermatan sulfate using chondroitin AC and ABC lyases (Saito et al., 1968). This can give slightly different results from those

obtained using chondroitin AC and B lyases (Linhardt et al., 1991). Chondroitin AC lyases from *F. heparinum* and *Arthrobacter aureescens* both act at glucuronate-containing linkages in dermatan sulfates (Gu et al., 1993) and are useful for analysis of the glucuronic acid content of dermatan sulfate (Linhardt et al., 1991).

Hyaluronate content can be determined using hyaluronate lyase, which is specific for hyaluronate. Although hyaluronate lyases are inhibited by chondroitin sulfates, the addition of salt (150 mM NaCl) overcomes this inhibition (Nakada and Wolfe, 1961; Yamagata et al., 1968).

Contaminants in samples. The presence of certain metals (particularly divalent metals), proteases, polyanions, detergents (SDS and Triton X-100), and denaturants (urea and guanidine) can interfere with the activity of lyases. Before using these enzymes, detergents should be removed by precipitation with potassium chloride or by using a detergent-removal column such as Biobeads (Bio-Rad). Urea and guanidine should be removed by exhaustive dialysis using controlled-pore dialysis membrane (MWCO 1000).

Reaction conditions. Lyases are compatible with a wide range of buffers including succinate, acetate, ethylenediamine acetate, Tris-Cl, bis-Trispropane-HCl, sodium phosphate, MOPS, TES, and HEPES (Lohse and Linhardt, 1992). The presence of calcium may either enhance or reduce lyase activity. Its effect is probably due to changes calcium causes in the GAG substrate conformation and not through direct interaction with the enzyme. Because calcium is incompatible with certain buffers, e.g., phosphate, and can lead to variable results, its use is not recommended. The pH optima for lyases are broad—between pH 5 and 9. Hyaluronate lyase works best at pH <7, and chondroitin and heparin lyases work best at pH ≥7. Lyases can be used at temperatures between 20° and 40°C. Optimum temperatures are always a compromise between activity and stability considerations. If enzyme instability is a concern due to elevated temperature or prolonged incubation time, heparin lyase II should be used to degrade heparin or heparan sulfate, chondroitin ABC lyase should be used to degrade chondroitin sulfates and dermatan sulfates, and hyaluronate lyase should be used for hyaluronate. Optimal reaction conditions for the polysaccharide lyases are summarized in Table 17.13B.2.

Removal of lyase and lyase activity. Following the use of a lyase, residual lyase activity can

be destroyed by heating the reaction mixture to 100°C or by adding denaturants or detergents. Most lyases are cationic proteins and can be removed from anionic oligosaccharide products by passing the reaction mixture through a small cation-exchange column, such as SP-Sephadex (Sigma), adjusted to an acidic pH. The oligosaccharide products (void volume) are then readjusted to neutral pH and analyzed (UNITS 17.17 & 17.19). This method can also be used to remove BSA, an excipient found in many of the commercial enzymes, from the oligosaccharide products.

Enzyme stability during storage. These enzymes can be stored in their lyophilized or reconstituted states at -20° or -70°C for ≤1 year. Once an enzyme is reconstituted, it should be aliquoted and frozen immediately. Single aliquots can be thawed to assay the enzyme or for use in an experiment. Heparin lyase II is very stable but heparin lyase III is unstable and should be used immediately after removing from frozen storage. The heparin lyases, particularly heparin lyase III, are sensitive to freeze-thawing and lyophilization. Among the chondroitin lyases, chondroitin AC lyase is most susceptible to thermal inactivation (Michelacci and Dietrich, 1975). Lyase storage stability is enhanced by high (>2 mg/ml) protein concentrations. This is often accomplished by addition of BSA.

Anticipated Results

When used correctly, an active lyase should specifically catalyze the breakdown of its GAG substrates. The types of GAGs present in a sample can be easily identified by using multiple lyases. The amount of product formed, and thus the amount of GAG in the sample, can be determined. Accuracy of analysis depends on a number of factors including the complexity of the sample and the types and concentrations of GAGs present. For a sample containing one GAG in high concentration with no contaminating proteins, salts, detergents, etc., GAG concentration can be determined to ±5%. For a complex sample containing many different GAGs at low concentrations in the presence of high levels of contaminants, it may only be possible to estimate GAG concentration to ±100%.

Time Considerations

It takes about half a day to prepare buffers and solutions and to reconstitute, aliquot, and freeze the enzyme. One aliquot of frozen enzyme can be thawed and assayed to ensure that

the enzyme is active and has been stored properly. Application of the enzyme to determine the presence or type of GAG in a sample requires ~3 days for sample preparation, overnight treatment with enzyme, and analysis.

Literature Cited

- Desai, U.R., Wang, H.H., and Linhardt, R.J. 1993a. Substrate specificity of the heparin lyase from *Flavobacterium heparinum* Arch. Biochem. Biophys. 306:461-468.
- Desai, U.R., Wang, H.M., and Linhardt, R.J. 1993b. Specificity studies on the heparin lyases from *Flavobacterium heparinum*. Biochemistry 32:8140-8145.
- Grant, A.C., Linhardt, R.J., Fitzgerald, G., Park, J.J., and Langer, R. 1984. Metachromatic activity of heparin and heparin fragments. Anal. Biochem. 137:25-32.
- Gu, K., Liu, J., Pervin, A., and Linhardt, R.J. 1993. Comparison of the activity of two chondroitin AC lyases on dermatan sulfate. Carbohydr. Res. 244:369-377.
- Hiyama, K. and Okada, S. 1975. Crystallization and some properties of chondroitinase from *Arthrobacter aureus*. J. Biol. Chem. 250:1824-1828.
- Jandik, K.A., Gu, K., and Linhardt, R.J. 1994. Action pattern of polysaccharide lyases on glycosaminoglycans. Glycobiology 4:000-000.
- Linhardt, R.J., Cooney, C.L., and Galliher, P.M. 1986. Polysaccharide lyases. Appl. Biochem. Biotechnol. 12:135-177.
- Linhardt, R.J., Al-Hakim, A., Liu, J., Hoppensteadt, D., Mascellani, G., Bianchini, P., and Fareed, J. 1991. Structural features of dermatan sulfates and their relationship to anticoagulant and antithrombotic activities. Biochem. Pharmacol. 42:1609-1619.
- Lohse, D.L. and Linhardt, R.J. 1992. Purification and characterization of heparin lyases from *Flavobacterium heparinum*. J. Biol. Chem. 267:24347-24355.
- Michelacci, Y.M. and Dietrich, C.P. 1975. A comparative study between a chondroitinase B and a chondroitinase AC from *Flavobacterium heparinum*. Biochem. J. 151:121-129.
- Nakada, H.I. and Wolfe, J.B. 1961. Studies on the enzyme chondroitinase: Product structure and ion effects. Arch. Biochem. Biophys. 94:244-251.
- Rautela, G.S. and Abramson, C. 1973. Crystallization and partial characterization of *Staphylococcus aureus* hyaluronate lyase. Arch. Biochem. Biophys. 158:687-693.
- Rice, K.G. and Linhardt, R.J. 1989. Study of defined oligosaccharide substrates of heparin and heparan monosulfate lyases. Carbohydr. Res. 190:219-233.
- Saito, H., Yamagata, T., Suzuki, S. 1968. Enzymatic methods for the determination of small quantities of isomeric chondroitin sulfates. J. Biol. Chem. 243:1536-1542.

Sasaki, I., Gotoh, H., Yamamoto, R., Tanaka, H., Takami, K., Yamashita, K., Yamashita, J., and Horio, T. 1982. Hydrophobic-ionic chromatography: Its application to microbial glucose oxidase, hyaluronidase, cholesterol oxidase, and cholesterol esterase. *J. Biol. Chem.* 91:1555-1561.

Vesterberg, O. 1968. Studies on extracellular proteins from *Staphylococcus aureus* III. Investigations on the heterogeneity of hyaluronate lyase using the method of isoelectric focusing. *Biochim. Biophys. Acta* 168:218-226.

Yamagata, T., Saito, H., Habuchi, O., and Suzuki, S. 1968. Purification and properties of bacterial chondroitinases and chondrosulfatases. *J. Biol. Chem.* 243:1523-1535.

Key References

Desai et al., 1993a, b. See above.

Describes the specificity of the heparin lyases in detail.

Lohse and Linhardt, 1992. See above.

Describes physical and catalytic properties of the heparin lyases.

Michelacci and Dietrich, 1975. See above.

Describes kinetic properties of chondroitin B lyase and chondroitin AC lyase.

Seikagaku Company Product Literature, 1991.

Describes general properties and assay conditions for all of the polysaccharide lyases.

Saito et al., 1968. See above.

Good example of how different chondroitin sulfates can be distinguished using chondroitin ABC lyase and chondroitin AC lyase.

Contributed by Robert J. Linhardt
University of Iowa
Iowa City, Iowa

Preparation of Glycopeptides

UNIT 17.14A

Generation of glycopeptides from glycoproteins is frequently useful when analyzing a protein's oligosaccharide side chains. Freed from the bulk of the polypeptide backbone by proteolysis, glycopeptides can be characterized by a variety of techniques. This approach is especially useful for proteins with multiple glycosylation sites, but it is also applicable when concern exists that the protein's structure and solubility will interfere with particular techniques such as gel-filtration and lectin-affinity chromatography (UNIT 17.1), glycosyltransferase assays (UNIT 17.6), and colorimetric assays (UNIT 17.9).

Extensive proteolysis with pronase or proteinase K results in oligosaccharides with one or a few amino acid residues attached. These glycopeptides are suitable for compositional analyses (UNIT 17.16), exo- and endoglycosidase digestions (UNITS 17.12 & 17.13), and lectin-affinity chromatography (UNIT 17.1). This technique, detailed in the first basic protocol, is often employed as a first step in characterizing oligosaccharides on very large glycoproteins such as proteoglycans and mucins.

Limited proteolysis with a specific endoproteinase (e.g., trypsin, α -chymotrypsin, and V8 protease), as described in the second basic protocol, leaves a larger peptide attached to the oligosaccharide. The resulting glycopeptides are generally suitable substrates for Peptide:N-glycosidase F (PNGase F; UNIT 17.13), an enzyme useful in defining oligosaccharide-peptide linkages. Additionally, they can be separated by C₁₈ reversed-phase chromatography, resulting in a glycopeptide map that is analogous to a peptide map (UNIT 10.12), and used for detection of glycosylation sites (UNIT 17.14B).

Alternative approaches for removing oligosaccharide side chains from the peptide include β -elimination (UNIT 17.15), endoglycosidases (UNIT 17.13), phospholipases (for GPI-anchored proteins; UNIT 17.8), or hydrazinolysis (Takasaki et al., 1982).

EXTENSIVE PROTEOLYTIC DIGESTION OF GLYCOPROTEINS

In this protocol, glycoproteins are applied to a Sephacryl S-200 gel-filtration column to separate them from any lower-molecular-weight material that may be present (e.g., degradation products or other low-molecular-weight contaminants). Denatured glycoproteins in SDS will generally elute in the void volume (V_0); the contaminants elute after the V_0 . The glycoproteins pooled from the V_0 are then digested with pronase or proteinase K and rechromatographed on the same column. Glycopeptides are recovered in the material that now elutes after the V_0 and are desalted before further use.

If the sample has already been purified by techniques such as immunoprecipitation (UNIT 10.16), dialysis (APPENDIX 3), or gel filtration (UNIT 10.9) which reliably remove low-molecular weight contaminants, it can be digested directly with pronase or proteinase K. After boiling to terminate the protease reaction, the glycopeptides can be desalted. Thus, the use of SDS and passage over Sephacryl S-200 can be completely avoided.

Generation of glycopeptides from a sample metabolically labeled with a radioactive sugar precursor (UNIT 17.4) is described and radioactivity is followed for each gel-filtration run. However, the same approach can be employed with unlabeled glycoproteins, following the column profiles with a carbohydrate-specific colorimetric assay (UNITS 17.9 & 17.16) if sufficient mass is present (see specific units for detection limits). Suitable samples range from total cellular lysates (containing multiple different glycoproteins together with degradation products and precursors) to highly purified protein samples.

BASIC
PROTOCOL

Preparation
and Analysis of
Glycoconjugates

17.14.1

Contributed by Leland D. Powell

Current Protocols in Molecular Biology (1995) 17.14.1-17.14.9

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 32

Materials

Radiolabeled glycoprotein sample (UNIT 17.4)
20% (w/v) sodium dodecyl sulfate (SDS)
1 M 2-mercaptoethanol (2-ME)
Sephacryl S-200 gel-filtration column (50- to 75-ml bed volume; Sigma or Pharmacia Biotech), equilibrated (UNIT 10.9) in ammonium formate/azide solution containing 0.1% (w/v) SDS (see recipe)
100% and 85% (v/v) acetone (HPLC or ACS grade), ice-cold
0.1 M Tris-Cl, pH 7.5 (APPENDIX 2), without or with 10 mM CaCl₂ (Tris/CaCl₂)
Pronase or proteinase K stock solution
Phenylmethylsulfonyl fluoride (PMSF), in 100% ethanol
Sephadex G-15 (Sigma or Pharmacia Biotech) or Bio-Gel P-2 (Bio-Rad) columns (30- to 60-ml bed volume), equilibrated in ammonium formate/azide solution
15- or 50-ml conical polypropylene tubes
Beckman TJ-6 centrifuge or equivalent
Water bath, 50°C

Additional reagents and equipment for preparation and standardization of gel-filtration columns and sample desalting (UNIT 10.9), quantitation of proteins (UNIT 10.1), and acetone precipitation (UNIT 17.10)

Purify the protein sample

1. Place radiolabeled glycoprotein sample in a 15-ml conical polypropylene tube. Add $\frac{1}{10}$ vol of 20% SDS and $\frac{1}{50}$ vol of 1 M 2-ME and boil 5 min. Cool to room temperature and centrifuge 10 min at $1000 \times g$ (2000 rpm in a Beckman TJ-6), room temperature to remove any precipitate. Save the supernatant.

The protein sample should contain <1 μg to several mg, depending on the source. The sample volume should be ≤1 ml. If the volume of the initial sample is >1 ml, lyophilize and resuspend it in 1 ml water (or less, if solubility permits). Large amounts of protein may require a larger volume, depending on the protein's solubility. If a larger volume is used, the volume of the Sephacryl S-200 column should be increased so that the sample volume is no more than 5% of the bed volume of the column.

Boiling ensures that any endogenous proteases or glycosidases are inactivated. For proteins known to be soluble without SDS, the addition of SDS can be eliminated. After boiling, the sample can be freed of low-molecular-weight contaminants, if necessary, by chromatography directly on Sephacryl S-200 (step 2) using a buffer in which the sample is known to be soluble.

2. Apply supernatant to Sephacryl S-200 column (50- to 75-ml bed-volume) equilibrated in ammonium formate/azide solution containing 0.1% (w/v) SDS. Collect 1.0- to 1.5-ml fractions and count 2% to 5% of each eluted fraction in a scintillation counter to determine the elution profile.

Sephacryl S-200 is utilized because almost all glycoproteins will elute in the V_0 in the presence of SDS. The profile should contain a peak eluting at V_0 . Additional peak(s) eluting after the V_0 , resulting from degradation products or unincorporated isotope, may or may not be present, depending on the source of the sample and its preparation (see critical parameters).

3. Pool the V_0 into a 15- or 50-ml polypropylene tube and determine the amount of protein present. Lyophilize the sample, resuspend in $\frac{1}{10}$ the original volume with water, and add 8 vol of ice-cold 100% acetone. Precipitate overnight at -20°C .

It is only necessary to determine the amount of protein in the sample if >20 μg may be present, so that sufficient protease may be added.

4. Centrifuge 15 min at $3000 \times g$, 4°C . Wash pellet gently with ice-cold 85% acetone and centrifuge as before. Remove supernatant with a Pasteur pipet and save pellet.

The second wash step is necessary to remove residual SDS which may interfere with the digestion. If desired, transfer the sample to a smaller polypropylene or a 5- or 10-ml screw-cap glass tube prior to the addition of acetone.

Digest with protease

- 5a. *For pronase digestion:* Add 200 μl of Tris/ CaCl_2 buffer to the pellet. If the total protein is $<20 \mu\text{g}$, add 20 μl pronase stock solution. Tightly cap the tube and incubate 2 to 4 hr at 50°C .

- 5b. *For proteinase K digestion:* Add 200 μl of 0.1 M Tris·Cl, pH 7.5, to the pellet. If the total protein is $<20 \mu\text{g}$, add 20 μl proteinase K stock solution. Tightly cap the tube and incubate 2 to 4 hr at 50°C .

See commentary for a discussion on the choice and uses of the two proteases. If the sample contains $>20 \mu\text{g}$ glycoprotein, add sufficient protease stock to give a protease:sample ratio of 1:20 (w/w). If larger amounts of protein are present, add larger amounts of digestion buffer. Depending on the protein sample, concentrations $\leq 100 \text{ mg/ml}$ are acceptable.

6. Add a second aliquot of protease (equal to that added in steps 5a or 5b) and continue the 50°C digestion for 12 to 16 hr.

The pellet should be completely solubilized during the digestion. If not, continue digestion for 48 or 72 hr. Inclusion of 2-ME to 20 mM (final) may increase solubility.

If a 50°C water bath is unavailable, the digestion may proceed equally well at 37°C . However, because microbial growth may be a problem at 37°C , 0.02% sodium azide should be added to the reaction if the digestion is done at this temperature.

7. Add $\frac{1}{10}$ vol of 20% SDS to completely solubilize any residual aggregated material. Boil 10 min, cool, and centrifuge 10 min at $1000 \times g$, room temperature.

Any protease not destroyed by boiling may contaminate the Sephacryl S-200 column and act to digest samples applied subsequently, resulting in obvious confusion if nondigested samples are applied. Inactivation of proteinase K can be augmented by phenylmethylsulfonyl fluoride (PMSF). Add $\frac{1}{100}$ vol of 0.1 M solution in 100% ethanol to the sample and incubate 20 to 30 min at room temperature prior to boiling.

8. Apply the supernatant to the Sephacryl S-200 column, collect fractions, and determine the elution profile as in step 2.

Recover glycopeptides

9. Compare the elution profiles before and after protease digestion. Material that elutes after the V_0 (often a broad peak) represents glycopeptides released from the glycoproteins that eluted in the V_0 of the first column. Pool the glycopeptides and add $\frac{1}{50}$ vol saturated KCl. Incubate 4 to 12 hr at 4°C .

The KCl treatment removes the bulk of the SDS.

10. Remove the insoluble K^+ -SDS precipitate by centrifuging 10 min at $1000 \times g$, 4°C . Collect the supernatant, concentrate by lyophilizing, and resuspend in 0.5 to 1.0 ml water.

Some residual K^+ -SDS precipitate may be apparent after resuspension and can be removed by repeating the centrifugation step as above.

11. Desalt glycopeptides on a Sephadex G-15 or Bio-Gel P-2 column (30- to 60-ml bed volume) equilibrated in ammonium formate/azide solution. Collect fractions that are

~5% of the total bed volume and determine the elution profile as in step 2. Pool the glycopeptides eluting in the V_0 and lyophilize. Store at -20°C .

The samples may be stored frozen indefinitely.

BASIC PROTOCOL

PROTEOLYTIC DIGESTION OF PURIFIED GLYCOPROTEINS WITH ENDOPEPTIDASES

Glycoproteins labeled metabolically with a radioactive sugar (UNIT 17.4) are first reduced and alkylated to increase solubility and prevent the cross-linking of peptides by disulfide bonds, and then digested with a site-specific endoproteinase. Unlike digestion with pronase or proteinase K, the resulting glycopeptides will generally be attached to a larger peptide of a defined length.

Materials

Radiolabeled protein sample (UNIT 17.4)
Tris/SDS solution
Dithiothreitol (DTT)
Iodoacetamide
1N sodium hydroxide (NaOH), prepared fresh every 4 weeks
2-mercaptoethanol (2-ME), undiluted
Sephadex G-50 column (30- to 60-ml bed volume; Sigma or Pharmacia Biotech) column, equilibrated in ammonium formate/azide solution containing 0.1% (w/v) SDS (UNIT 10.9)
100% and 85% (v/v) acetone, ice-cold
Nitrogen (N_2) stream
Protease digestion buffer
Protease stock solution
5- to 15-ml Pyrex or polypropylene tubes
Aluminum foil
pH paper, range pH 7 to 12
10-ml round-bottomed polypropylene tubes
Beckman TJ-6 centrifuge or equivalent
1.5-ml polypropylene microcentrifuge tubes
Additional reagents and equipment for preparation and standardization of gel-filtration columns and sample desalting (UNIT 10.9) and acetone precipitation (UNIT 17.10)

1. Place the purified radiolabeled protein sample in a 5- to 15-ml Pyrex or polypropylene tube and dissolve in 700 μl Tris/SDS solution.

Recovery of radiolabeled glycopeptides proteins can be aided by the addition of a carrier protein (e.g., 50 μg of ovalbumin, bovine immunoglobulin, or fetuin), which should be added before the reduction step.

In principle, any denaturing and strongly buffered solution similar to the reducing buffer can be employed. For instance, if the protein is in a PBS/NP-40 buffer, $1/10$ vol of 3 M Tris-Cl, pH 8.6/20% SDS stock solution could be added to approximate the reduction buffer. Alternatively, the protein sample can be exchanged into reducing buffer by gel-filtration chromatography (UNIT 10.9), dialysis (APPENDIX 3), or acetone precipitation (UNIT 17.10).

2. Add 11.9 mg DTT and boil 3 min. Incubate 1 hr at 37°C .
3. Remove the sample from the water bath and wrap tube in aluminum foil to protect from light. Add 32.3 mg of iodoacetamide. *Immediately* check the pH with pH paper

and adjust to pH 8 to 9 by adding 50- μ l aliquots of 1 N NaOH (5 to 10 aliquots may be needed). Continue the incubation for 1 hr at room temperature, shielded from light.

Alkylation of the free sulfhydryl groups prevents the reformation of disulfide bonds. Improper pH during this procedure will inhibit the alkylation. Exposure to light can produce unwanted side reactions, but dim light is acceptable to use.

4. Terminate the reaction with 50 μ l of undiluted 2-ME and apply to a Sephadex G-50 column (30- to 60-ml bed volume) equilibrated in ammonium formate/azide solution containing 0.1% (w/v) SDS. Collect 1.0- to 1.5-ml fractions and count 2% to 5% of each eluted fraction in a scintillation counter to locate the glycoprotein eluting in the V_0 . Pool the appropriate fractions into a 10-ml round-bottomed polypropylene tube and lyophilize.
5. Redissolve the reduced and alkylated protein in 0.3 to 0.4 ml water. Add 8 vol of ice-cold 100% acetone and precipitate overnight at -20°C .
6. Centrifuge 15 min at $3000 \times g$, 4°C , and discard the acetone. Rinse the walls of the tube and the pellet with 1 ml of 85% ice-cold acetone, and centrifuge as before. Use a Pasteur pipet to carefully aspirate most of the acetone, leaving all precipitated material behind, and dry with a gentle stream of N_2 .

The pellet must be dry because any remaining acetone will interfere with the endopeptidase reaction.

7. Add 200 μ l of protease digestion buffer and a small stirbar. Add 10 μ l protease stock, rotate the tube gently to coat the liquid over the walls where the sample has been in contact, and stir gently at room temperature. Add another 10 μ l protease after 8 to 12 hr and digest for a total of 24 hr.

See commentary for choice of protease. For radiolabeled proteins, generally present in microgram quantities or less, the amount of protease (2 mg) used here should represent a considerable excess over the sample. If carrier protein has been added, or the mass of the sample is known, sufficient protease should be added to ensure at least a 1:1 protease:sample (w/w) ratio. For nonradiolabeled proteins, protease:sample ratios of 1:10 to 1:30 are employed.

8. Terminate the reaction by boiling 5 min. Use a large magnetic stirbar on the outside of the tube to pull the small stirbar halfway up the side. Rinse the small stirbar with 0.2 to 0.4 ml water, then remove it. Centrifuge the sample 10 min at $2000 \times g$, room temperature, and transfer the supernatant to a 1.5-ml microcentrifuge tube. Wash the large tube and pellet once with water, centrifuge as before, and combine supernatants.
9. Resuspend the pellet in 0.5 ml digestion buffer and count an aliquot ($\sim 5\%$) in a scintillation counter. Likewise, count an aliquot ($\sim 5\%$) of the supernatant. Calculate the percent release of radioactivity as soluble peptides by:

$$100 \times \text{cpm in supernatant} / (\text{cpm in supernatant} + \text{cpm in pellet})$$

If a significant amount of material remains insoluble, redigest with more of the same or another protease (e.g., α -chymotrypsin or V8 protease). Store samples frozen.

Soluble material will represent both digested sample (i.e., peptides and glycopeptides) and possibly undigested sample (e.g., if no appropriate sites of proteolysis are present or accessible). Likewise, insoluble material may represent undigested sample (insoluble because of denaturation and acetone precipitation) and perhaps glycopeptides and peptides as well. Some peptides and glycopeptides are insoluble in aqueous solutions at neutral pH and are only soluble in dilute acid (e.g., 25 mM sodium phosphate, pH 2.5), alkali (e.g., ammonium hydroxide), organic solvents (e.g., acetonitrile), or denaturants (e.g., SDS or guanidine). The choice of buffer used to resuspend the digested material in this step may

be altered depending on the next analytical step planned. For example, if the sample is to be analyzed by SDS-PAGE (UNIT 10.2), the digested material can be resuspended directly in sample buffer containing SDS. If reverse-phase liquid chromatography is planned, the sample can be resuspended in TFA or TFA containing 6M guanidine buffer. Thus, the demonstration that the endopeptidase added in step 7 has cleaved the sample depends not on the solubility of the sample but rather on direct biochemical evidence of proteolysis, such as a change in size as detected by SDS-PAGE, gel-filtration chromatography, or reverse-phase peptide mapping. The calculation of percent release of soluble radioactivity is essential to determine if insolubility and thus sample loss at this point is a problem.

REAGENTS AND SOLUTIONS

Ammonium formate/azide solution

1.58 g ammonium formate (25 mM)
0.2 g sodium azide (NaN_3 ; 0.2% w/v)
 H_2O to 1 liter
Store indefinitely at room temperature

Prepare the Sephacryl S-200 and Sephadex G-50 column equilibration buffers by adding 1.0 g sodium dodecyl sulfate (SDS; 0.1% w/v) to this solution.

Protease digestion buffers

Trypsin or α -chymotrypsin digestion buffers

0.79 g ammonium bicarbonate
 H_2O to 100 ml
Titrate with NH_4OH to pH 8.0. Filter through a 0.45- μm membrane and store at -20°C .

V8 protease digestion buffer

74.4 mg Na_2EDTA (2 mM)
0.39 g ammonium bicarbonate (50 mM) or 0.71 g sodium phosphate (50 mM)
 H_2O to 100 ml
Titrate with NH_4OH to pH 7.8
Store at -20°C

In the presence of ammonium acetate or bicarbonate, V8 protease cleaves on the carboxyl side of Asp residues; in the presence of phosphate, it cleaves after both Asp and Glu residues (Drapeau, 1977).

Protease stock solutions

For α -chymotrypsin: Dissolve 10 mg in 100 μl of 1 mM HCl. Store unused portion at 4°C for the second addition (step 8, alternate protocol). Do not use >1 day old.

For pronase: Dissolve 10 mg/ml pronase in Tris/ CaCl_2 and incubate 2 hr at 50°C to destroy contaminating glycosidases. Store at 4°C ; use within 48 hr.

For proteinase K: Resuspend 2 mg/ml proteinase K in 0.1 ml of 0.1 M Tris-Cl, pH 7.5, containing 2 mM calcium chloride (22 mg anhydrous CaCl_2 in 100 ml buffer). Store stock solution ≤ 2 weeks at 4°C or ≤ 1 month at -20°C without loss of activity.

Calcium reduces autolysis of the protease. Proteinase K is also available in solution.

For trypsin: Dissolve 10 mg TPCK-treated trypsin (Worthington) or DPCC-trypsin (Sigma) in 100 μl of 1 mM HCl. Store unused portion at 4°C for the second addition (step 8, alternate protocol). Do not use TPCK-treated trypsin if the second digestion is planned to be done with α -chymotrypsin. Do not use >1 day old.

For V8 protease (endoproteinase Glu-C or staphylococcal V8 protease): Make a 10 mg/ml stock in V8 protease digestion buffer. Store at 4°C . Do not use >1 day old.

Saturated KCl

45 g KCl
100 ml H₂O

Mix to saturation by stirring overnight, or by boiling for 5 min and allowing solution to cool to room temperature. Store indefinitely at room temperature.

Tris/CaCl₂ solution

0.11 g anhydrous calcium chloride (CaCl₂; 10 mM)
100 ml 0.1 M Tris·Cl, pH 7.5
Store indefinitely at room temperature

Tris/SDS solution

To 90 ml H₂O add:
6.05 g Tris·Cl (0.5 M)
2.0 g sodium dodecyl sulfate (SDS; 2% w/v)
Titrate with HCl to pH 8.6
Add H₂O to 100 ml
Store indefinitely at room temperature

COMMENTARY

Background Information

Pronase is widely utilized to generate glycopeptides with minimal peptide content, irrespective of the nature of the oligosaccharide-peptide linkage (Spiro, 1966). This is of obvious advantage when the nature of the glycosidic linkage is not known. Thus, the total oligosaccharide population from a sample can be released and characterized either functionally or structurally by serial lectin-affinity chromatography (UNIT 17.18) or endoglycosidase digestion (UNIT 17.13). Gel-filtration chromatography is limited by potential heterogeneity of the peptides remaining attached to the oligosaccharide. For the same reason, ion-exchange chromatography can not be used to determine the number of charged groups, such as sialic acid or sulfate, although similarly prepared samples can be compared with one another.

Pronase is a mixture of several proteases of differing specificities that are capable (in principle) of removing all amino acid residues except the linkage residue from a glycopeptide. However, in some applications, more than one amino acid residue is left attached. Although there are few published applications of proteinase K in the generation of glycopeptides, it offers several advantages over pronase. These include equally broad activity, wider pH range (pH 4 to 12), greater tolerance of SDS and urea, lack of divalent cation requirement, and inactivation by PMSF (Ebeling et al., 1974). Also, contaminating glycosidases and phosphatases do not appear to be a problem, eliminating the need for the predigestion step required for pro-

nase. Like pronase, proteinase K will digest a protein into small fragments, showing a preference for cutting after hydrophobic amino acid residues. The generation of oligosaccharides linked to single amino acids probably does not occur with great efficiency because proteinase K is not active against dipeptide substrates with unblocked amino groups.

Studies specifically directed at proteoglycan analyses often start with pronase digestion of [³H]glucosamine and ³⁵SO₄⁻²-labeled material. A useful overview of this approach, including steps on sample preparation, can be found in Britz and Hart (1983).

Glycopeptides prepared with pronase or proteinase K are not reliable substrates for PNGase F (UNIT 17.13), which requires a peptide linkage on both sides of the glycosylated Asn residue for activity. Glycopeptides derived from a site-specific endopeptidase are generally suitable substrates for PNGase F. Occasionally glycopeptides resistant to PNGase F may be encountered, if the endoprotease happens to cleave adjacent to the glycosylated Asn residue. Digestion with a different endoprotease should circumvent this problem.

Endopeptidases (Judd, 1990; Swiedler et al., 1983) that require a slightly alkaline pH are suitable for work with glycoproteins. Digestion buffers requiring an acidic pH and chemical cleavage techniques may degrade some oligosaccharide structures. Trypsin, α-chymotrypsin, and V8 protease are active between pH 7.5 to 8.5 and in 0.1% SDS, and represent a reasonable first choice. O-acetyl groups on

sialic acid, however, are stable under acidic but not alkaline conditions (UNIT 17.12).

Critical Parameters and Troubleshooting

The details of sample preparation will largely depend on the source of material, its purity, and the detection method. Total cellular extracts, containing both secreted and integral membrane proteins, require detergents for solubility, as may denatured proteins. In contrast, secreted and plasma glycoproteins do not require detergents. [³H]Mannose will label only N-linked oligosaccharides and GPI anchors (UNITS 17.4 & 17.8), and therefore proteoglycans, glycolipids, and O-linked chains will be mostly invisible. [³H]Glucosamine will label all of these structures, and they are all detected (although with variable sensitivities) by hexose-specific colorimetric assays (UNIT 17.9).

The steps required in sample preparation will depend on the sample source and the specific questions being asked. The following points should be considered:

1. Samples need to be purified away from low-molecular-weight glycopeptides, degradation products, and sugar precursors prior to analysis.

2. Glycolipids can cause problems. Alone or with nondenaturing detergents, they migrate at a high molecular weight by gel filtration (as micelles), but at a low molecular weight in the presence of excess SDS or denaturing solvents. On SDS-PAGE, they can migrate as small proteins (~10,000 MW). They may precipitate with acid, acetone, or ethanol (depending on the amount of salts and protein present), but can be difficult to extract from tissues with acetone. If glycolipids will be a problem, specific steps for their removal should be employed (Finne and Krusius, 1982; Britz and Hart, 1983).

3. The solubility of denatured glycoproteins can be a limiting factor in any of the steps. Occasionally, prolonged incubations at elevated temperatures or the addition of 2-mercaptoethanol to 20 mM (final) will solve problem.

Careful bookkeeping of recoveries at each step is important to detect unexpected losses. The use of polypropylene tubes and a carrier protein with radioactive samples may help decrease these losses.

Both protocols assume that digestion will yield glycopeptides containing single glycosylation sites. This is generally the case for N-linked oligosaccharides, as their sites are spaced far apart on the peptide backbone. O-linked glycosylation sites can be clustered in

regions of high serine (Ser) and/or threonine (Thr) content and be resistant to proteolysis. Heterogeneity in glycosylation (both in terms of the size of the oligosaccharide and whether a particular Ser/Thr residue is glycosylated) may cause a partial block in protease accessibility to certain regions of the peptide, resulting in incomplete and variable proteolysis and heterogeneity in the resulting glycopeptides.

Anticipated Results

Digestion of radiolabeled samples with pronase or proteinase K will yield molecules consisting of the oligosaccharide structure found at a given glycosylation site, plus one or a few amino acids. Such structures can be studied functionally (e.g., in blocking, binding, or adhesion assays) or structurally (UNITS 17.6, 17.13, 17.16, 17.18, & 17.19).

Digestion of purified glycoproteins with site-specific proteases will generate glycopeptides with a larger peptide portion attached. These glycopeptides can be fractionated by reversed-phase chromatography because the peptides are of a defined length (UNIT 10.12), used as substrates for endoglycosidases, or subjected to other biochemical analyses.

Time Considerations

Each individual step (e.g., enzyme digestion, gel filtration, and desalting) will require less than half a day. Lyophilization and acetone precipitation are overnight procedures. The procedure from starting sample to purified glycopeptides can easily be done in five calendar days.

Literature Cited

- Britz, J.S. and Hart, G.W. 1983. Biosynthesis of glycosaminoglycans by epithelial and lymphocytic components of murine thymus. *J. Immunol.* 130:1848-1855.
- Drapeau, G.R. 1977. Cleavage at glutamic acid with Staphylococcal protease. *Methods Enzymol.* 47:189-191.
- Ebeling, W., Hennrich, N., Clockow, M., Metz, H., Orth, H.D., and Lang, H. 1974. Proteinase K from *Tritirachium album* Limber. *Eur. J. Biochem.* 47:91-97.
- Finne, J. and T. Krusius. 1982. Preparation and fractionation of glycopeptides. *Methods Enzymol.* 83:269-277.
- Judd, R.C. 1990. Peptide mapping. *Methods Enzymol.* 182:613-626.
- Spiro, R.G. 1966. Characterization of carbohydrate units of glycoproteins. *Methods Enzymol.* 8:26-52.

- Swiedler, S.J., Hart, G.W., Tarentino, A.L., Plummer, T.H., and Freed, J.H. 1983. Stable oligosaccharide microheterogeneity at individual glycosylation sites of a murine major histocompatibility antigen derived from a B-cell lymphoma. *J. Biol. Chem.* 258:11515-11523.
- Takasaki, S., Mizuochi, T., and Kobata, A. 1982. Hydrazinolysis of asparagine-linked sugar chains to produce free oligosaccharides. *Methods Enzymol.* 83:263-268.

Key References

- Judd, 1990. See above.
Useful review on specific endopeptidases.
- Spiro, 1966. See above.
Describes use of pronase and techniques for handling glycopeptides.
- Swiedler et al., 1983. See above.
Technique from which the basic protocol on endo-proteinase digestion was adapted.

Contributed by Leland D. Powell
University of California San Diego
La Jolla, California

Detection of Individual Glycosylation Sites on Glycoproteins

Many glycoproteins contain multiple sites of glycosylation. They may all have the same oligosaccharide–amino acid linkage (GlcNAc–Asn, GalNAc–Ser/Thr, GlcNAc–Ser/Thr, Xyl–Ser/Thr, or GPI anchor), or different types may be present. Although many of the methods in this unit are directed towards characterizing the total oligosaccharide population on a glycoprotein, for some investigations it may be necessary to determine (1) how many different sites exist, (2) the sugar–amino acid linkage of each site, and (3) the structural characteristics of the oligosaccharides found at the different sites.

Although such an in-depth study is not feasible for all sites on all glycoproteins, a general approach (described in the Basic Protocol) based on peptide mapping techniques has been employed successfully. Glycopeptides generated by endopeptidase digestion (UNIT 17.14A) are separated by reversed-phase chromatography (UNIT 10.12) using C₁₈ resin. Elution times are determined primarily by hydrophobic interactions between the peptide and the resin. However, the presence of hydrophilic, negatively charged oligosaccharides shortens retention times, causing glycopeptides to elute in considerably broader peaks than do peptides. By following the elution profile either radiochemically (for material prepared with radioactive sugar precursors or sulfate; UNIT 17.4) or colorimetrically (using hexose-specific assays; UNITS 17.9 & 17.16), the peaks corresponding to unique glycopeptides can be identified. With proper controls, the number of peaks will correspond to the number of different glycosylation sites. The eluted fractions are suitable for analysis by lectin chromatography (for N-linked oligosaccharides; UNIT 17.1), and the peptide sugar linkage can be defined either by endoglycosidase digestion (see Support Protocol and UNIT 17.13) or chemical cleavage (β -elimination; UNIT 17.15). Oligosaccharides freed from the peptide as described in the Support Protocol can be characterized by size or charge (UNIT 17.17), techniques not generally applicable with glycopeptides.

BASIC PROTOCOL

FRACTIONATION OF GLYCOPEPTIDES BY REVERSED-PHASE HPLC

A highly purified glycoprotein metabolically labeled with a radioactive sugar precursor (UNIT 17.4) is digested with a site-specific protease (UNIT 17.14A). The resulting glycopeptides, each of which contains a single or just a few glycosylation site(s), are fractionated by reversed-phase HPLC. The individual glycopeptide fractions are pooled separately and desalted in preparation for further analytical studies (see Support Protocol).

Materials

Highly purified glycoprotein metabolically labeled with a radioactive sugar precursor (UNIT 17.4)

HPLC buffers A and B (see recipe)

1 M Tris·Cl, pH 8.4 (APPENDIX 2)

Organic solvents: chloroform, DMSO, acetonitrile, and methanol

10 mM ammonium formate in water and in 50% acetonitrile (see recipe)

Ventilated oven, 40° to 45°C

Nitrogen tank

C₁₈ cartridge (e.g., Sep-Pak cartridge, Waters)

1-cc and glass 10-cc syringes

50-ml polypropylene tube

Additional reagents and equipment for peptide isolation by reversed-phase HPLC (UNIT 10.12)

Detection of Individual Glycosylation Sites on Glycoproteins

17.14.10

Contributed by Leland D. Powell

Current Protocols in Molecular Biology (1995) 17.14.10-17.14.15

Copyright © 2000 by John Wiley & Sons, Inc.

1. Prepare glycopeptides from purified, labeled glycoprotein sample by reduction, alkylation, and specific endopeptidase digestion (*UNIT 17.14A*). Lyophilize the digested samples.

If the primary sequence of the protein is known, choose a protease or combination of proteases that will result in peptides having a unique glycosylation site. Even if one is only interested in N-linked sites and is working with material labeled with [2-³H]mannose, the potential of other glycosylation sites to interfere with proteolysis must be kept in mind.

If sequence data are not available, several proteases or combinations thereof may need to be compared.

2. Immediately prior to HPLC analysis, resuspend the lyophilized sample in 100 to 200 μ l HPLC buffer A, wetting the walls of the tube exposed to sample to ensure maximal recovery.
3. Microcentrifuge the sample 10 min at 10,000 \times g, 4°C. Using a pipet tip with 2 to 4 mm of the tip cut off, transfer solution and as much of the particulate material as possible to a 1.5-ml microcentrifuge tube on ice. Rinse out the digestion tube again and combine with the first wash. Keep the total volume to less than that of the injector sample loop.

As the HPLC buffers are acidic and sialic acid residues potentially acid labile, minimize the time of exposure of the sample to the buffer and keep it on ice until just before injecting onto the HPLC.

4. Microcentrifuge the sample material 10 min at 10,000 \times g, 4°C. Carefully remove the supernatant fluid, measure the radioactivity of an aliquot by scintillation counting, and inject the rest on the HPLC column. If minimal amounts of the sample are available, wash the insoluble material once in HPLC buffer A, spin as before, and pool the supernatant fluids.

Do not inject particulate material onto the HPLC column.

Depending on the complexity of the pattern and purity of the sample, anywhere from 5,000 to 50,000 cpm of sample are necessary for an adequate analysis.

5. Determine the amount of glycopeptide insoluble in HPLC buffer A (the “core”) by washing the insoluble particulate material once in water, resuspending it in water, and measuring the radioactivity of an aliquot by scintillation counting. If a significant percentage of the total sample remained insoluble after this first proteolytic digestion, redigest the material with the same protease (using a larger amount) or another protease.

If TPCK-treated trypsin was used, chymotrypsin digestion cannot follow. If digestion with chymotrypsin is anticipated, use the highest-grade trypsin available (not TPCK-treated).

6. Perform HPLC exactly as described in *UNIT 10.12*, except starting with the column equilibrated in HPLC buffer A and eluting the peptides with a linear gradient to 100% HPLC buffer B over 90 min, using a flow rate of 1 ml/min, and collecting 1-min fractions.

With experience, adjustments may need to be made in the shape of the gradient used to optimize the fractionation of individual glycopeptides.

7. Collect the effluent into tubes containing 0.2 to 0.5 ml of 1 M Tris·Cl, pH 8.4 (ensure final pH >5). As the run proceeds, remove tubes and mix to assure timely neutralization. Determine the elution profile by measuring the radioactivity of an aliquot of each fraction by liquid scintillation counting.

Analytical runs do not require neutralization. Instead, collect the fractions directly into scintillation vials and dry by placing them (in a cardboard or metal rack) overnight in a ventilated oven at 40° to 45°C, resuspending the salt pellet in 0.2 ml water followed by scintillation cocktail, and counting.

8. Pool the appropriate peaks and dry under a nitrogen stream.

Once most of the organic solvent is removed, the sample can be frozen, dried under vacuum, and stored at -20°C. Samples should be stable at least 1 year.

9. Prepare a C₁₈ cartridge by washing sequentially with chloroform, DMSO, acetonitrile, methanol, and water (10 ml per wash), and finally with 20 ml of 10 mM ammonium formate in water.
10. Resuspend the sample in 1 to 2 ml water and apply to the cartridge using a 1-cc syringe. Rinse the sample tube with 1 ml of 10 mM ammonium formate and apply to the cartridge. Repeat once more, then wash the cartridge with an additional 20 ml of 10 mM ammonium formate, saving all the aqueous washes.
11. Using a glass syringe, elute the cartridge with 6 ml of 10 mM ammonium formate/50% acetonitrile into a 50-ml polypropylene tube and dry under nitrogen as before. Store frozen.

The fractionated glycopeptides can now be analyzed by further digestion (see Support Protocol).

12. Follow the recovery of glycopeptide throughout by measuring the radioactivity of a fixed percentage (1% to 5%) of the initial sample, the aqueous washes, and the acetonitrile-eluted sample by scintillation counting, being careful to mix each sample before removing the aliquots for counting.

Most glycopeptides will be soluble in aqueous solutions in the absence of detergents or organic solvents. However, poor recoveries or unexpected losses when pipetting suggest that solubility may be limiting. To check for "sticky" glycoprotein, rinse tubes and pipet tips that have been exposed to samples with 0.1% SDS and measure the radioactivity of the wash by scintillation counting. Resuspension of the sample, after removal of the acetonitrile, in 0.1% NP-40, 10 mM β-octylglucoside, or another nondenaturing detergent may be necessary.

SUPPORT PROTOCOL

ENDOGLYCOSIDASE DIGESTION OF PURIFIED GLYCOPEPTIDES

Glycopeptides pooled and desalted from a reversed-phase chromatogram (see Basic Protocol) are digested with a specific endoglycosidase (UNIT 17.13A). Sensitivity to digestion is determined by rechromatographing over the same C₁₈ column.

Additional Materials (also see Basic Protocol)

Glycopeptide sample (see Basic Protocol, step 11)

HPLC apparatus with 500-μl injector sample loop

Additional reagents and equipment for endoglycosidase digestion (UNIT 17.13)

1. Resuspend glycopeptide sample in 10 mM ammonium formate in water and transfer to a microcentrifuge tube. Dry in a vacuum concentrator.
2. Add 10 to 50 ml of an appropriate buffer for digestion with PNGase F, endo F, or endo H (UNIT 17.13A) or endo-O-α-N-acetylgalactosaminidase (UNIT 17.17). Add the appropriate enzyme and allow digestion to proceed as specified in the respective unit.

Detergents can be omitted from the enzyme buffer formulations, as most glycopeptides are soluble in aqueous solutions.

3. Terminate the reaction by boiling 3 min.
4. Just prior to HPLC chromatography, dilute with 5 to 10 vol HPLC buffer A (but not exceeding the capacity of the injector loop), microcentrifuge 10 min at $10,000 \times g$, 4°C , and perform HPLC (see Basic Protocol, steps 6 and 7).

A 500- μl injector sample loop is required. Whenever possible, do not load over 400 μl in the loop, as the presence of a stationary phase along the walls of microbore tubing limits their usable volume.

5. Determine the elution time of the sample.

A decrease in elution time relative to that of the undigested glycopeptide indicates that its oligosaccharide was a substrate for the chosen endoglycosidase.

As in Basic Protocol step 7, glycopeptides are retained by the C_{18} column. After removal of the peptide, the free oligosaccharide will elute in the first few minutes of the column run (with or just after the breakthrough volume). Nondigested material will elute at the same time as seen initially in Basic Protocol step 7. The percentage release of the oligosaccharide can be calculated directly. These runs can be performed either analytically or preparatively.

An alternative to HPLC is to fractionate the digested oligosaccharide on a Sep-Pak C_{18} cartridge; glycopeptides will be retained and released oligosaccharides will elute in the water wash (unless the original glycopeptide elutes with water or very early in the HPLC gradient).

Nondigested glycopeptides may be recovered as in Basic Protocol steps 7 to 12. Released oligosaccharides may be desalted on a $1 \times 20\text{-cm}$ column of Sephadex G25 or Bio-Gel P-2, eluted in 10 mM ammonium formate for addition analyses.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

HPLC buffers A and B

2 \times HPLC buffer stock
28.1 g NaClO_4 (HPLC grade)
2.35 ml phosphoric acid
 H_2O to 1 liter

The pH will be 1.7 to 1.9 (do not adjust).

Buffer A: Mix 500 ml of 2 \times stock and 500 ml water.

Buffer B: Mix 500 ml of 2 \times stock and 500 ml HPLC-grade acetonitrile.

Use deionized, distilled water throughout. Filter and degas solvents.

Ammonium formate, 10 mM, in water and in 50% acetonitrile

Prepare 20 mM ammonium formate in water; do not adjust the pH, which will be ~ 6.7 to 6.9 . Dilute 1:1 (v/v) with water (to make 10 mM ammonium formate) or with acetonitrile (to make 10 mM ammonium formate/50% acetonitrile).

COMMENTARY

Background Information

Analysis of protein glycosylation on a site-by-site basis is not widely done, as it requires more time and material than analysis of total protein oligosaccharide. However, as more refined questions in glycobiology are addressed, such detailed structural information will increasingly be needed. Most published reports on this topic have focused on N-linked oligosaccharides, as these are in general more structurally complex than O-linked (mucin-like) oligosaccharides. Additionally, owing to their spacing on the peptide backbone, N-linked oligosaccharides are relatively easily separated into unique peptides by endopeptidase digestion and thus lend themselves to study by this technique. In principle, this approach could be used to identify a C-terminal peptide with a GPI anchor (released by PI-PLC digestion; *UNIT 17.8*). In contrast, O-linked mucin-type structures are frequently, although not always, clustered on the peptide backbone and cannot be separated by this approach. As [2-³H]mannose labels predominantly N-linked and GPI structures, it is often the radiolabel of choice for these analyses.

Two major questions are addressed in a site-by-site analysis.

1. How many sites are glycosylated? The consensus sequence for glycosylation of O-linked sites and proteoglycans is not clearly established. The sequence for N-linked glycosylation (Asn-X-Ser/Thr) is easily identified, yet not all possible acceptor sequences are glycosylated. Thus, direct biochemical methods must be employed to determine the number of glycosylation sites.

2. How do the structures found at the different sites compare? It is widely accepted that for a glycoprotein with a single glycosylation site, multiple different oligosaccharide structures will be found. The various factors responsible for this heterogeneity have been recently reviewed (Kobata and Takasaki, 1992). Moreover, site-specific microheterogeneity is also present: at the different glycosylation sites on a given glycoprotein, different populations of oligosaccharides will be found. These differences can only be assessed if the different sites are first separated and then analyzed independently.

Site-specific glycosylation patterns have been determined for several glycoproteins, including α 1-acid glycoprotein (Treueheit et al.,

1992), murine histocompatibility antigens (Swiedler et al., 1983), and macrophage adhesion molecules (Dahms et al., 1985).

Critical Parameters

Detailed studies such as those outlined in this unit require very clean glycoprotein preparations. Even if, based on SDS-PAGE and fluorographic analysis, a sample appears to be clean, with just low levels of radioactivity migrating diffusely outside of the band of interest, it is still possible for this low-level background to be concentrated by the C₁₈ column and thus turn into a significant contaminant. Thus, it may be useful to purify the radiolabeled sample by SDS-PAGE before proceeding with proteolysis and HPLC. Proteolysis of gel-purified samples can be performed directly in the acrylamide gel slices, on material blotted onto nitrocellulose, or after eluting from the gel slice.

The choice of radioactive precursor for labeling will depend on a variety of factors; see *UNIT 17.4* Commentary for discussion. As indicated in Background Information, [2-³H]mannose is often used for the analysis of N-linked oligosaccharides.

The potential presence of multiple glycosylation sites on each glycopeptide, either from clustered sites or from incomplete proteolysis, is a significant concern. This is most worrisome for O-linked sites, which frequently (but not always) appear in clusters on the peptide backbone.

Large proteins with multiple sites (e.g., >6) can be difficult to analyze owing to the complexity of the resulting chromatogram. An alternative strategy entails separating the protein into large glycopeptides by cleaving with CNBr or by limited proteolysis of nondenatured proteins, resulting in domains that can be fractionated by size, reversed-phase HPLC (*UNIT 10.12*), or ion-exchange chromatography (*UNIT 10.10*) and subsequently subjected to more extensive proteolysis.

Troubleshooting

The possibility of obtaining overly complex reversed-phase chromatograms is one of the most significant limitations of this methodology. Heterogeneously glycosylated peptides will migrate as broad clusters of peaks rather than the discrete peaks seen with peptides; incomplete proteolytic digestion can also result in increased peak complexity. If confronted with a complex or uninterpretable profile, in-

tially rule out incomplete digestion by repeating the process with more protease, longer digestion times, and/or combinations of proteases (trypsin and chymotrypsin). Alternatively, reduce the oligosaccharide heterogeneity (as with sialidase digestion) and repeat the C₁₈ analysis.

Recovery of glycopeptides at each step should be monitored closely to detect unexpected losses and/or solubility problems. As a rule, glycopeptides are very soluble in aqueous buffers, but large proteolytic fragments from denatured proteins may exhibit problematic solubility behavior.

Some large N-linked glycopeptides are resistant to cleavage by PNGase F. The basis of this inhibition has not yet been determined.

Anticipated Results

The C₁₈ chromatogram should contain discrete clusters of peaks representing single peptides with heterogeneous oligosaccharide side chains. Comparison of glycopeptides radiolabeled with different radioactive sugar precursors should indicate the type of oligosaccharide present (N-, O-, or GPI-linked structures). Comparison of the amount of material in each peak will indicate if some sites are underglycosylated. If the procedure is done on a preparative scale, the different glycopeptide fragments are suitable for more detailed structural analysis.

Time Considerations

Preparation of glycopeptides is described (including time considerations) in *UNIT 17.14A*.

Performing the HPLC analysis will require a day's effort. Endoglycosidase digestions require ~1 hr to set up, an overnight digestion, and 2 to 4 hr to analyze. Multiple sites will require proportionally more time.

Literature Cited

- Dahms, N.M. and Hart, G.W. 1985. Lymphocyte function-associated antigen 1 contains sulfated N-linked oligosaccharides. *J. Immunol.* 134:3978-3986.
- Kobata, A. and Takasaki, S. 1992. Structure and biosynthesis of cell surface carbohydrates. *In Cell Surface Carbohydrates and Development* (M. Fukuda, ed.) pp. 1-24.
- Swiedler, S.J., Hart, G.W., Tarentino, A.L., Plummer, T.H., and Freed, J.H. 1983. Stable oligosaccharide microheterogeneity at individual glycosylation sites of a murine major histocompatibility antigen derived from a B-cell lymphoma. *J. Biol. Chem.* 258:11515-11523.
- Treuheit, M.J., Costello, C.E., and Halsall, H.B. 1992. Analysis of the five glycosylation sites of human α 1-acid glycoprotein. *Biochem. J.* 283:105-112.

Key Reference

Swiedler et al., 1983.

The reference upon which most of these methods are based.

Contributed by Leland D. Powell
University of California San Diego
School of Medicine
La Jolla, California

β -Elimination for Release of O-Linked Glycosaminoglycans from Proteoglycans

UNIT 17.15A

**BASIC
PROTOCOL**

O-linked glycosaminoglycan (GAG) chains in proteoglycans are readily released from their core proteins by treatment with alkali at room temperature. This β -elimination is the same type of reaction as that for releasing O-linked oligosaccharides from their core proteins. Under the reaction conditions described here, N-linked oligosaccharides remain attached to the core protein, but any O-linked oligosaccharides will be released along with the GAG chains. The procedure can be used to isolate the free GAG chains, the free O-linked oligosaccharides, and the core protein (which will still have any N-linked oligosaccharides that were originally present).

Materials

Proteoglycan solution to be analyzed
Alkaline borohydride reagent (see recipe)
2 M HCl
1 M NaOH

Additional reagents and equipment for dialysis (*APPENDIX 3*)

1. Dialyze proteoglycan solution exhaustively against water to remove all buffer salts. Transfer to small test tube or centrifuge tube and dry in gentle stream of air or by lyophilization.
2. Dissolve dried proteoglycan sample in minimal volume of alkaline borohydride reagent and incubate 24 hr at room temperature.

The reaction tube should not be capped because evolution of hydrogen gas may result in pressure buildup.

3. Add 2 M HCl at room temperature, monitoring acidity with pH paper until the pH of the solution is 0 (to convert excess borohydride to hydrogen gas).
4. When evolution of hydrogen has ceased, add 1 M NaOH until pH is 6 to 8 as measured with pH paper.

For detailed discussion of further treatment of the products, see Commentary.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Alkaline borohydride reagent

Dissolve 22.4 mg sodium borohydride (NaBH_4) in 2 ml of 0.4 M NaOH just before use.

The resulting solution contains 0.3 M NaBH_4 .

**Preparation and
Analysis of
Glycoconjugates**

17.15.1

Contributed by H. Edward Conrad

Current Protocols in Molecular Biology (1995) 17.15.1-17.15.3

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 31

COMMENTARY

Background Information

Chondroitin sulfate, dermatan sulfate, and heparan sulfate are glycosaminoglycans (GAGs) occurring in cell-surface proteoglycans, in which they are O-linked to their core proteins through a xylosyl→serine linkage. Cartilage keratan sulfate is also O-linked to proteins in proteoglycans (through *N*-acetylgalactosamine, or GalNAc), whereas corneal keratan sulfate occurs as an N-linked polysaccharide. Hyaluronic acid, however, does not occur linked to a protein. Although heparin is synthesized in the mast cell as a proteoglycan, it is cleaved from its core protein in the final stages of synthesis and occurs, for the most part, as a mixture of free polysaccharide chains that are released from the cells only when the cells degranulate.

The basic protocol gives essentially complete release of O-linked GAGs, along with any O-linked oligosaccharides attached to the core protein. This is accomplished in 24 hr without removal of N-linked oligosaccharides (Mayo and Carlson, 1970; Seno and Sekizuka, 1978). The serine residues to which the GAGs are O-linked are converted to dehydroalanine residues, but the core protein is not cleaved. Thus, the molecular weight of both the core protein and the released GAG chains can be estimated by gel electrophoresis or gel filtration (Bienkowski and Conrad, 1985; Fedarko and Conrad, 1986). Sodium borohydride (NaBH_4) is included in the reaction mixture to prevent alkaline cleavage of the newly formed xylosyl or GalNAc reducing terminals from the rest of the GAG chain (known as the peeling reaction). If desired, tritiated sodium borohydride (NaB^3H_4) can be used in the reaction to label the reducing-terminal xylose of the GAG chains (as well as the reducing-terminal residues of the O-linked oligosaccharides, if present). If NaB^3H_4 is used, all reactions must be carried out in a designated radioactive fume hood because oxidation of NaB^3H_4 yields $^3\text{H}_2$ gas. N-linked oligosaccharides remain attached to the core protein under the mild reaction conditions described here.

This protocol is useful for demonstrating the proteoglycan nature of an unknown sample containing GAGs, as the molecular weight of proteoglycans will be reduced, but that of free GAG chains will not (Bienkowski and Conrad, 1984; Fedarko and Conrad, 1986). It is also a necessary step for characterizing the proteoglycan and its constituent core protein and GAG

chain(s) after the GAGs are separated from the smaller O-linked oligosaccharides that are also released in the reaction. If NaB^3H_4 is used in the reaction to label the reducing-terminal xylosyl, the molar ratio of total disaccharides in the GAG chains to [^3H]xylosyl can be used to estimate the size of the GAG chains. In addition, the linkage region can be recovered from an ^3H -labeled GAG chain by trimming away the main chain with lyases (*UNIT 17.14B*) or nitrous acid (for heparan sulfate), leaving the labeled linkage region. If the starting proteoglycan contains both O-linked GAGs and O-linked oligosaccharides, the products will consist of core protein, polysaccharides, and oligosaccharides. After the reaction the products can be separated from each other by gel filtration to resolve the oligosaccharides from the core protein and GAG chains. The polymeric fraction can then be resolved by DEAE-cellulose chromatography in a salt gradient to separate the core protein from the more highly charged GAG chains (Bienkowski and Conrad, 1984).

Critical Parameters

Although the reaction can be run at higher temperatures (35° to 50°C) to increase the reaction rate, loss of sulfate residues, cleavage of N-linked oligosaccharides (Ogata and Lloyd, 1982), or cleavage of the polypeptide chain may result. Under the milder conditions described here, these side-reactions appear to be minimal, and both the carbohydrate chains and the core proteins can be recovered intact. These conditions release both xylose-linked glycosaminoglycans (GAGs), GalNAc-linked keratan sulfate, and O-linked oligosaccharides, and there seems to be little difference in the rates of β -elimination of these different carbohydrates. No comparisons of these rates, however, have been reported. Sodium borohydride must be present to reduce the new reducing terminals of the carbohydrate chains as they are released from the core protein, so that alkaline degradation of the carbohydrate chains cannot occur. It is advisable to free the proteoglycan from oxyanions such as phosphate prior to reaction because these will catalytically destroy sodium borohydride (Conrad et al., 1973).

The procedure described in this unit does not include palladium chloride (PdCl_2), which has been used by other workers (Seno and Sekizuka, 1978) to reduce dehydroalanine to alanine as it is formed during the elimination.

PdCl_2 may be useful to stabilize the core protein when more harsh conditions are used for the β -elimination, and to allow amino acid analysis of the resulting core protein by observing the conversion of serine to alanine. It is not, however, necessary for the reaction, and the colloidal Pd formed in the reaction may interfere with recovery of the reaction products. It should be noted that the number of moles of serine converted to alanine is a reflection of the number of moles of carbohydrate chains originally O-linked to the core protein.

Anticipated Results

The β -elimination reaction is virtually stoichiometric. Assuming that the proteoglycan is pure before the reaction is run, the products of this reaction will be the intact core protein, the intact glycosaminoglycan chains, and any O-linked oligosaccharides that were also attached to the core protein.

Time Considerations

This reaction can be run ≥ 24 hr without any serious side reactions. Normally, a 24-hr incubation is sufficient to obtain complete cleavage of the proteoglycan.

Literature Cited

Bienkowski, M.J. and Conrad, H.E. 1984. Kinetics of proteoglycan sulfate synthesis, secretion, endocytosis, and catabolism by a hepatocyte cell line. *J. Biol. Chem.* 259:12989-12996.

Conrad, H.E., James, M.E., and Varboncoeur, E. 1973. Qualitative and quantitative analysis of reducing carbohydrates by radiochromatography on ion exchange papers. *Anal. Biochem.* 51:486-500.

Fedarko, N.F. and Conrad, H.E. 1986. A unique heparan sulfate in the nuclei of hepatocytes. Structural changes with the growth state of the cells. *J. Cell Biol.* 102:587-599.

Mayo, J.W. and Carlson, D.M. 1970. Effect of alkali and sodium borohydride at alkaline pH on *N*-acetylchondrosine: Reduction versus cleavage. *Carbohydr. Res.* 15:300-303.

Ogata, S.-I. and Lloyd, K.O. 1982. Mild alkaline borohydride treatment of glycoproteins—A method for liberating both N- and O-linked carbohydrate chains. *Anal. Biochem.* 119:351-359.

Seno, N. and Sekizuka, E. 1978. Quantitative β -elimination of O-glycosyl linkages in chondroitin sulfates. *Carbohydr. Res.* 62:271-299.

Key Reference

Seno and Sekizuka, 1978. See above.

Describes the basic procedure used here and includes the use of PdCl_2 to reduce the dehydroalanine residues to alanine residues.

Contributed by H. Edward Conrad
University of Illinois
Urbana, Illinois

β -Elimination for Release of O-GalNAc-Linked Oligosaccharides from Glycoproteins and Glycopeptides

This unit describes release of oligosaccharides that are attached to polypeptides through an *N*-acetylgalactosamine (GalNAc) linkage to the hydroxyl groups of serine or threonine. The β -elimination procedures described here can be used to recover the oligosaccharide chains (also called glycans) and/or identify the serine or threonine residues involved in the linkage.

β -elimination (Fig. 17.15.1) is usually carried out in the presence of a reducing agent such as sodium borohydride (NaBH_4) under alkaline conditions (Basic Protocol; see also UNIT 17.15A). The reducing agent immediately converts released oligosaccharides into reduced oligosaccharides, which is necessary because free oligosaccharides can be degraded under alkaline conditions by further β -elimination (peeling reaction; see Background Information). This method provides quantitative release of the oligosaccharides but does not yield intact polypeptides (see Critical Parameters).

An alternative method is to employ only alkaline conditions without a reducing agent (Alternate Protocol 1). The extent of conversion of serine and threonine residues in the polypeptide after oligosaccharide release indicates the number of such residues involved in linkage to O-glycans. However, this method does not allow quantitative recovery of the released glycans.

Another alternative is to use sodium sulfite (Na_2SO_3 ; see Alternate Protocol 2), which converts serines and threonines that are attached to O-glycans into cysteic acid and α -amino- β -sulfonylbutyric acid, respectively. This method, however, does not allow quantitative recovery of released glycans because further β -elimination of O-glycans (known as the peeling reaction; Fig. 17.15.2) cannot be prevented. The two products of the second alternate protocol (cysteic acid and α -amino- β -sulfonylbutyric acid) cannot be resolved using an amino acid analyzer, as they will elute in the same peak, so they must be quantitated using paper chromatography (Support Protocol).

β -elimination is not restricted to oligosaccharides attached through GalNAc but may also be used for release of oligosaccharides linked through other residues, including *N*-acetylglucosamine and xylose. For release of glycosaminoglycans linked through xylose, see UNIT 17.15A.

BASIC PROTOCOL

β -ELIMINATION IN THE PRESENCE OF SODIUM BOROHYDRIDE

β -elimination is achieved by reducing the sample under alkaline conditions using NaBH_4 (Carlson, 1968; Fukuda, 1989). Products are resolved by passing the reaction mixture through a Sephadex column. This procedure allows quantitative recovery of the O-glycans that are released.

Materials

- Glycopeptide or glycoprotein sample
- 1 M NaBH_4 /0.05 M NaOH (prepare fresh)
- 1 M acetic acid in methanol (prepare fresh, cool to room temperature before use)
- Nitrogen gas stream
- Methanol (HPLC grade)
- 0.2% orcinol/2 M H_2SO_4 (store at 4°C)
- 9-ml screw-cap conical glass test tubes

β -Elimination for Release of Oligosaccharides

17.15.4

Contributed by Minoru Fukuda

Current Protocols in Molecular Biology (1995) 17.15.4-17.15.9

Copyright © 2000 by John Wiley & Sons, Inc.

25° to 30°C and 45°C water baths
1 × 30–cm Sephadex G-15 column (Pharmacia Biotech) equilibrated in H₂O
Silica gel–coated thin-layer chromatography (TLC) plate
150° to 200°C oven

1. Mix 1 to 500 µg glycopeptide or glycoprotein sample and 5 to 100 µl of 1 M NaBH₄/0.05 M NaOH in a 9-ml screw-cap conical glass test tube.

The volume of solution used should be minimal but enough to dissolve the sample.

2. Close test tube loosely with screw cap and incubate 24 to 36 hr in a 45°C water bath.
3. Open test tube in a vented hood and slowly add 3 ml of 1 M acetic acid in methanol. Place tube in a 25° to 30°C water bath and dry contents in a nitrogen gas stream.
4. Repeat step 3.
5. Add 3 ml methanol, dry as in step 3, then dissolve in ~200 µl water. Either proceed to step 6 or freeze at –20°C until ready to continue.
6. Apply to a 1.0 × 30–cm Sephadex G-15 column and elute with water. Collect 1-ml fractions using a fraction collector.
7. Spot an aliquot from each fraction on a silica gel–coated TLC plate and allow to dry. Spray plate with orcinol/H₂SO₄ reagent and heat 4 min in a 150° to 200°C oven. Identify fractions containing carbohydrates by appearance of a brown color.

Alternatively, other methods for detection of carbohydrates can be used. It is also possible to calibrate the column before the experiment so that the fractions where the oligosaccharides will elute are known in advance. Those fractions are then collected and pooled.

8. Dry carbohydrate-containing fractions by evaporation and freeze at –20°C until ready for further analysis.

RAPID β-ELIMINATION IN DIMETHYL SULFOXIDE SOLUTION

This procedure facilitates release of O-glycans so that β-elimination can be completed in a short time. It yields oligosaccharides that are not reduced at the reducing terminals, making it possible to add various functional groups such as fluorescent probes.

Additional Materials (also see *Basic Protocol*)

Dimethyl sulfoxide (DMSO)
0.425 M KOH
100% ethanol
0.4 M HCl

1. Dissolve glycoproteins or glycopeptides in 500 µl to 1 ml of DMSO to a final concentration of 5 mg/ml in a 9-ml screw-cap conical glass test tube.
2. Add sufficient 0.425 M KOH and 100% ethanol to give a 0.17 M KOH concentration and a DMSO/water/ethanol ratio of 50:40:10 (v/v/v).

For example, to treat a sample dissolved in 1 ml DMSO, add 800 µl of 0.425 M KOH and 200 µl of 100% ethanol.

3. Incubate 1 hr at 45°C, then neutralize with 0.4 M HCl using pH paper as indicator.
4. Proceed with Basic Protocol, steps 6 to 8, or freeze sample at –20°C until ready to continue.

ALTERNATE PROTOCOL 1

Preparation and Analysis of Glyconjugates

17.15.5

β -ELIMINATION IN THE PRESENCE OF SODIUM SULFITE

This procedure is used (in combination with the Support Protocol) to determine the number of serine and threonine residues involved in O-glycan attachment, which are converted to cysteic acid and α -amino- β -sulfonylbutyric acid, respectively, by Na_2SO_3 .

Additional Materials (also see Basic Protocol)

0.5 M Na_2SO_3 /0.1 M NaOH (prepare fresh)

1 M and 6 M HCl

7-ml conical glass test tube

Vacuum evaporator, 40°C

110°C oven

1. In a 7-ml conical glass test tube, dissolve glycopeptides to a concentration of 1.0 mM or glycoprotein to a concentration of 8 mg/ml in 0.5 M Na_2SO_3 /0.1 M NaOH. Incubate 48 hr at 37°C.

Glycoprotein samples can first be treated by oxidation with performic acid to facilitate determination of serine and threonine residues (see Critical Parameters).

2. Terminate reaction by adjusting pH to ~3 with 1 M HCl. Evaporate to dryness in a 40°C vacuum evaporator.
3. Dissolve in 200 μ l of 6 M HCl and transfer to a test tube. Hydrolyze 24 hr in a 110°C oven with tube sealed.
4. Evaporate hydrolysis reaction mixture (total amino acid mixture) to dryness in a 40°C vacuum evaporator, then proceed to support protocol for quantitation of O-linked serine and threonine residues (see Support Protocol).

The reaction mixture may be dissolved in water and stored at -20°C until ready for analysis.

SEPARATION OF β -ELIMINATION PRODUCTS BY PAPER CHROMATOGRAPHY

In the second alternate protocol, the serine and threonine residues that are linked to O-glycans are converted, respectively, to cysteic acid and α -amino- β -sulfonylbutyric acid. Conventional amino acid analyzers combine the total yield of these two compounds in the same peak. This support protocol employs paper chromatography to separate cysteic acid and α -amino- β -sulfonylbutyric acid for the purpose of determining the number of O-linked serine and threonine residues in a glycoprotein or glycopeptide (Spiro, 1972).

Materials

Dried β -elimination reaction product (see Alternate Protocol 2)

0.05 M NaOH

Dowex AG-1-X8 resin (200 to 400 mesh; formate form; Bio-Rad)

5 mM pyridine formate buffer, pH 4.0 (see recipe)

4 M formic acid

4:1:5 (v/v/v) 1-butanol/acetic acid/water

0.5 \times 5-cm chromatography column

Whatman 3 MM chromatography paper

Paper chromatography chamber

1. Dissolve dried β -elimination product in water and titrate to pH 4.0 with 0.05 N NaOH.

2. Equilibrate Dowex AG-1-X8 resin by suspending in pyridine formate buffer, pH 4.0 for several hours. Pack 0.5 equilibrated resin into a 0.5 × 5-cm column and wash with pyridine formate buffer.
3. Apply sample from step 1 to column. Wash column with 10 ml pyridine formate buffer, then elute acidic amino acids with 5 vol of 4 M formic acid.
4. Dilute eluate with 2 ml water, then lyophilize to remove formic acid and pyridine formate.
5. Separate acidic amino acids in 4:1:5 (v/v/v) 1-butanol/acetic acid/water for 5 days using Whatman 3 MM paper.

The migration relative to aspartic acid (R_{Asp}) in this system is 0.42 for both cysteic acid and α -amino- β -sulfonylbutyric acid.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 5.

Pyridine formate buffer (5 mM), pH 4.0

Prepare 5 mM pyridine formate and adjust to pH 4.0 with 6 M formic acid. Store at room temperature in a sealed glass bottle.

COMMENTARY

Background Information

O-linked oligosaccharides, which are attached to polypeptides through the β -carbon-borne hydroxyl groups of serine or threonine, are abundant in mucin-type glycoproteins. O-linked oligosaccharides can be released from glycoproteins by β -elimination, after which the oligosaccharides can be fractionated and their structures determined. The method relies on β -elimination, shown in Figure 17.15.1, and takes advantage of the presence of the hydroxyl group of serine or threonine at the amino acid's β carbon. Under alkaline conditions, the peptide-group oxygen becomes negatively charged, and a double bond is formed between the peptide carbon and the α carbon. This double bond in turn migrates to make a double bond between the α and β carbons, and the β -carbon oxygen and linked oligosaccharide are released.

During this β -elimination, the released O-glycans are exposed to strong alkaline conditions. Such treatment tends to induce another β -elimination reaction within the released O-glycans (known as a peeling reaction). In particular, if the next sugar residue is attached to C-3 (also known as the β carbon) of *N*-acetylgalactosamine (GalNAc), this additional β -elimination occurs as in Fig. 17.15.2—i.e., presence of alkali introduces a double bond and this induces another β -elimination reaction in which the sugar chain is cleaved and released

from GalNAc. It is therefore necessary to introduce a reducing agent to prevent such further β -elimination. Thus, to prevent further degradation of released oligosaccharides and permit quantitative recovery of the released O-glycans, alkaline degradation is usually carried out in the presence of large amounts of the reducing agent NaBH₄ (Basic Protocol). β -elimination employing NaBH₄, however, does not provide quantitative recovery of alanine or α -aminobutyric acid (produced, respectively, from serine or threonine residues that are involved in O-glycan attachment).

Critical Parameters

The method employing 1 M NaBH₄ in 0.05 N NaOH (Basic Protocol) is most suitable for quantitative release of O-linked oligosaccharides from glycoproteins. This method is, however, not suitable for obtaining intact polypeptides because the sodium borohydride also cleaves peptide bonds and may also often lead to release of N-linked oligosaccharides (Ogata and Lloyd, 1982).

The need for a technique that kept the polypeptide portion as intact as possible led to the invention of the procedure using NaOH in DMSO (Alternate Protocol 1; Downs et al., 1973). However, this method is better suited for release of intact oligosaccharides than intact polypeptides, as cleavage of polypeptides cannot be completely avoided.

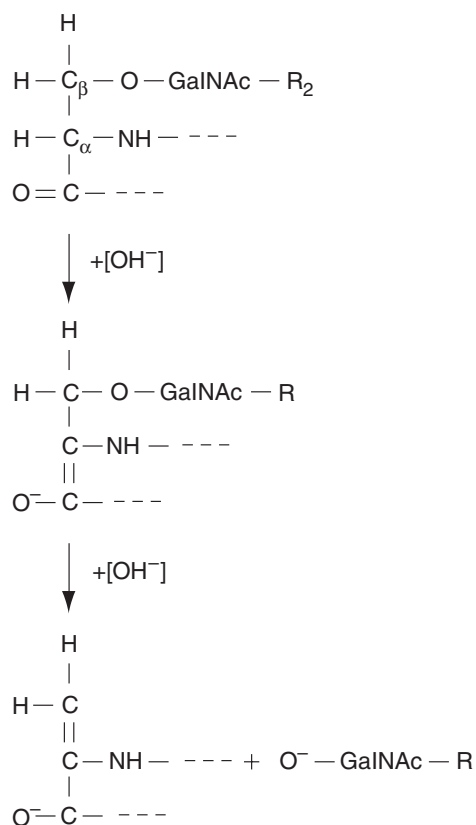


Figure 17.15.1 β -elimination for release of *O*-GalNAc-linked oligosaccharides from glycoproteins and glycopeptides. R represents oligosaccharide chain; broken lines represent remainder of polypeptide chain. The double bond formed can be reduced in the presence of sodium borohydride and palladium chloride (also see *UNIT 17.15A*).

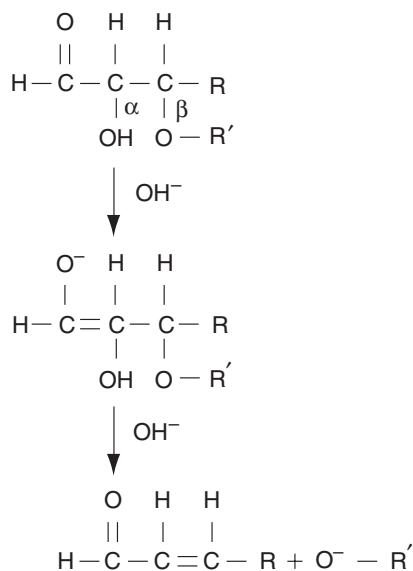


Figure 17.15.2 Peeling reaction. R, remainder of GalNAc moiety; R', oligosaccharide chain.

The alkaline sodium sulfite method using 0.5 M Na₂SO₃ in 0.1 M NaOH (Alternate Protocol 2), is best suited for estimating the number of serine and threonine residues involved in O-glycan attachment. However, this method does not give a 100% yield of α -aminobutyric acid because the conversion of threonine to α -amino- β -sulfonylbutyric acid is usually ~75%. This method also does not prevent the peeling reaction involving the O-glycans that are released. It is useful only for determining the amount of serine and threonine involved in O-glycosylation. In order to determine accurately the amount of cysteic acid and α -aminobutyric acid produced in the alkaline sodium sulfite reaction, it is desirable to prepare a control in which the glycoproteins are oxidized with performic acid before β -elimination. This treatment cleaves disulfide cross-linkages in the peptide chain, converting, for example, cystine to cysteine residues. The difference between the amounts of cysteic acid and α -aminobutyric acid obtained before and after the alkaline sodium sulfite reaction represents the amount of serine and threonine involved in O-glycan attachment.

In order to release O-glycans without cleavage of the polypeptide moiety, other chemical methods must be employed—e.g., those employing trifluoromethanesulfonic acid or hydrogen fluoride (Sojar and Bahl, 1987). Attempts to characterize the oligosaccharides released using these methods have not yet been made.

Troubleshooting

Problems occur in β -elimination if the sample is treated with reducing agent under acidic conditions. This lowers the efficiency of the β -elimination reaction and also decreases the effective concentration of reducing agent. If it is not known whether or not the sample is acidic, it should be desalted before carrying out the reaction.

Anticipated Results

β -elimination using the three methods presented in this unit is virtually stoichiometric for release of O-linked oligosaccharides. Of these methods, only the Basic Protocol yields intact

oligosaccharides. Using Alternate Protocol 1, the released oligosaccharides may be degraded, although relatively intact polypeptides may be released. Using Alternate Protocol 2, it is possible to determine the number of serine and threonine residues involved in O-glycan attachment; however, here too the released oligosaccharides may be degraded. Neither of the alternate protocols is therefore suited for oligosaccharide analysis.

Time Considerations

The alkaline borohydride reaction (Basic Protocol and Alternate Protocol 1) can be run 24 to 48 hr without any serious side reactions. Including the time for chromatography, 48 hr is sufficient to obtain a mixture of O-linked oligosaccharides from glycopeptide or glycoprotein. Alternate Protocol 2 (β -elimination with sodium sulfite) will take an additional 24 hr.

Literature Cited

- Carlson, D.M. 1968. Structures and immunochemical properties of oligosaccharides isolated from pig submaxillary mucins. *J. Biol. Chem.* 243:616-626.
- Downs, F., Herp, A., Moschera, J., and Pigman, W. 1973. β -elimination and reduction reactions and some applications of dimethylsulfoxide on submaxillary glycoproteins. *Biochim. Biophys. Acta* 328:182-192.
- Fukuda, M. 1989. Characterization of O-linked saccharides from cell surface glycoproteins. *Methods Enzymol.* 179:17-29.
- Ogata, S.-I. and Lloyd, K.O. 1982. Mild alkaline borohydride treatment of glycoproteins—A method for liberating both N- and O-linked carbohydrate chains. *Anal. Biochem.* 119:351-359.
- Sojar, H.T. and Bahl, O.P. 1987. Chemical deglycosylation of glycoproteins. *Methods Enzymol.* 138:350-359.
- Spiro, R.G. 1972. Study of carbohydrates of glycoproteins. *Methods Enzymol.* 28:3-43.

Contributed by Minoru Fukuda
La Jolla Cancer Research Foundation
La Jolla, California

Acid Hydrolysis for Release of Monosaccharides

The first step in obtaining a compositional analysis of a glycoconjugate is the release of the individual monosaccharide constituents. How these are released and processed depends on what methods are available for their analysis and identification. When the analysis is to be performed by high-performance liquid chromatography (HPLC), the monosaccharides must be released by acid hydrolysis of the glycoconjugate. Free glucose residues can then be separated using a pellicular-resin anion-exchange column and detected by pulsed amperometry (see UNIT 17.19A). When a gas-liquid chromatography (GLC) approach is to be taken in the analysis—and the glycoconjugate contains amino sugars and/or uronic acids—methanolysis, followed by trimethylsilylation of hydroxyl and carboxyl groups, is the method of choice (UNIT 17.19A). Often a prediction can be made of the type of monosaccharides present in a glycoconjugate by analogy with other glycoconjugates from closely related sources. If the sample is not expected to contain amino sugars or uronic acids—or if only neutral sugars must be measured—acid hydrolysis, followed by reduction and peracetylation (UNIT 17.19A), is preferred.

When an HPLC approach is taken, neutral sugars, hexosamines, and uronic acids can be directly analyzed after strong acid hydrolysis (see Basic Protocol 3). If sialic acids are involved, they must be analyzed separately by submitting an aliquot of the sample to mild acid conditions (see Basic Protocol 2), after which the sialic acids are purified and analyzed by HPLC without derivatization (see UNIT 17.19A) or after derivatization (see UNIT 17.18). In some cases, enzymatic release of sialic acids (see UNIT 17.12) may be chosen to preserve the *O*-acyl substitutions.

Mild acid hydrolysis to release fucosyl residues from glycoconjugates is described in Basic Protocol 1. Mild acid hydrolysis to release sialic acids from glycoconjugates, along with dialysis and column chromatography procedures to purify the sialic acids, are described in Basic Protocol 2. Strong acid hydrolysis to release all monosaccharides from glycoconjugates is described in Basic Protocol 3. Conditions required for successful release of all the different monosaccharides in a glycoconjugate vary to a great extent when hydrochloric or sulfuric acid is used for hydrolysis. However, the conditions required for hydrolysis using trifluoroacetic acid (TFA; see Basic Protocol 3) are considerably more uniform. Apart from sialic acids, which are known to be destroyed under strong acid conditions, all sugar residues—including fucose, hexosamines, and uronic acids—can be confidently determined after optimized treatment with TFA. TFA also has the advantage of being easily eliminated by evaporation without leaving a residue. For the release of sialic acids, however, mild acid hydrolysis using acetic acid (see Basic Protocol 2) and/or enzymatic release (UNIT 17.12) is recommended. The Alternate Protocol describes a harsher acid hydrolysis procedure for 4-*O*-acetylated sialic acids, which are more difficult to release.

A precise quantitative ratio of the different types of monosaccharides in a glycoconjugate is very difficult to obtain. This arises from the fact that all glycosidic bonds must be split while avoiding destruction of the liberated monosaccharides. Susceptibility to acid hydrolysis, as well as stability of the released unit, is different for each monosaccharide—e.g., pyranosides and furanosides hydrolyze at different rates; the presence of substituents affects the rate of hydrolysis; and α and β anomers have different rates of hydrolysis. In some cases, therefore, more than one set of conditions may be required to achieve a complete analysis (see Critical Parameters).

MILD ACID HYDROLYSIS FOR RELEASE OF FUCOSE RESIDUES

This protocol describes mild acid hydrolysis conditions appropriate for the release of terminal nonreducing fucose from a glycoconjugate. These conditions should release the fucose residues without releasing other monosaccharides (except sialic acids). Strong acid hydrolysis (see Basic Protocol 3) should be used to analyze for total monosaccharide composition, including fucose.

Materials

Glycoconjugate-containing sample to be analyzed

Standard: 1 $\mu\text{mol/ml}$ L-fucose (mol. wt. 164.2) in H_2O

0.05 M HCl prepared from constant-boiling HCl (sequencer-grade, Pierce)

99% methanol (anhydrous; e.g., Aldrich)

Small glass culture tubes or 3.5-ml glass vials with Teflon-lined screw caps

Heating block or oven

Nitrogen (N_2) or vacuum evaporation system: e.g., Speedvac (Savant), shaker-evaporator (Baxter Scientific), or lyophilizer

Dialysis tubing (500 MWCO)

Additional reagents and equipment for phenol–sulfuric acid assay for monosaccharides (UNIT 17.9), dialysis (APPENDIX 3C), and compositional analysis of monosaccharides from glycoproteins (UNIT 17.19A)

1. Estimate the total carbohydrate composition of the sample to be analyzed by the phenol–sulfuric acid assay (UNIT 17.9).
2. Using N_2 , a Speedvac system, or a lyophilizer, evaporate three aliquots of the sample to be analyzed (each containing ≥ 10 nmol of carbohydrate) to dryness in small glass culture tubes or 3.5-ml glass vials. Simultaneously evaporate to dryness three 10- μl aliquots of 1 $\mu\text{mol/ml}$ L-fucose for use as standards.

Use brand-new precleaned glass tubes or vials, or clean used ones by heating ≥ 3 hr at 50°C first in concentrated nitric acid, then in 6 M HCl—each time by placing them inside a glass container with a lid and filling the container with acid. Care should be taken to ensure that each tube or vial is completely filled with acid. Rinse the vials thoroughly with Milli-Q purified water or equivalent, then with ethanol, and finally dry in an oven. When < 5 μg total carbohydrates are to be analyzed, it is recommended that the glass vials be silanized by incubating 15 min at room temperature with 2% dichlorodimethylsilane in toluene. The solution is then decanted and the silanized vials rinsed successively with methanol and hot distilled water and allowed to dry.

3. Redissolve each dried sample and standard by adding 0.4 ml of 0.05 M HCl for up to 1 mg of material. Flush samples and standards with nitrogen and cap tubes.

When glass vials are used, nitrogen can be introduced through a needle inserted in the septum and evacuated through a second needle.

4. Using a 100°C heating block or oven, heat one sample and one standard 2 hr, another sample and standard 4 hr, and the remaining sample and standard 8 hr.

When a heating block is used, care must be taken that transfer of heat to the tubes is adequate. This can be achieved by filling the wells with pump oil or using a sand or metallic bath. Most of the tube or vial should be covered with the heating element to avoid distillation of the solvent in the upper part of the tube. When oil is used in the wells, excess oil must be wiped from the vials and the vials washed with hexane and dried before opening.

5. After each sample and standard has been heated for the required time period, evaporate the solution using a nitrogen or vacuum evaporation system. Wash residues twice with 99% methanol, then evaporate again.

CAUTION: This procedure should be carried out in a chemical fume hood.

If this is of interest, a defucosylated glycoconjugate can be prepared following this protocol. After acid hydrolysis under optimal conditions, neutralize reaction mixture with 0.05 M NaOH, dilute, and dialyze at 4°C against 20 vol water using 500-MWCO dialysis tubing to recover defucosylated glycoconjugate. Alternatively, the defucosylated product can be recovered by gel filtration chromatography (UNIT 10.9) on Bio-Gel P2 or Sephadex G-15, using water for elution.

6. Store the dried hydrolysates at -20°C until ready for compositional analysis. Results will indicate the best time for complete hydrolysis with minimal decomposition.

Procedures for compositional analysis are found in UNIT 17.19.

RELEASE OF SIALIC ACIDS (EXCLUDING 4-O-ACETYLATED SPECIES) BY MILD ACID HYDROLYSIS AND PURIFICATION OF THE PRODUCT

**BASIC
PROTOCOL 2**

This procedure allows the quantitative release of most types of known sialic acids, other than those acetylated at the 4 position, with minimal destruction. Subsequent handling of the sialic acid hydrolysate, including purification, is also designed to avoid migration or loss of *O*-acetyl groups. After purification, the mixtures of sialic acids can be analyzed by HPLC with (UNIT 17.18) or without derivatization (UNIT 17.19A). For release of sialic acids acetylated at the 4 position, see Alternate Protocol. This protocol can be used to prepare a mixture of sialic acids from commercially available glycoconjugate standards (e.g., bovine submaxillary mucin) for use as a standard in sialic acid analysis.

Materials

Glycoconjugate-containing sample to be analyzed
Sodium formate buffer, pH 5.5 (see recipe)
0.1 N H₂SO₄
10 M and 2 M acetic acid
1% butylated hydroxytoluene (BHT; Sigma) in ethanol
Dowex AG 50W-X2 ion-exchange resin (200 to 400 mesh, hydrogen form; Bio-Rad)
Dowex AG 3-X4A ion-exchange resin (100 to 200 mesh, hydroxyl form; Bio-Rad)
10 mM and 1 M formic acid (ACS certified), ice-cold
Dialysis tubing (1,000 and 12,000 MWCO; Spectrapor, Spectrum)
Heating block
5-ml glass culture tubes with Teflon-lined screw caps
0.5 × 10-cm glass chromatography columns *or* Pasteur pipets plugged with glass wool
20-ml glass test tubes
Lyophilizer or shaker-evaporator (Baxter Scientific)
Additional reagents and equipment for dialysis (APPENDIX 3C), de-*O*-acetylation of sialic acids and TBA assay (UNIT 17.18), and ion-exchange chromatography (UNIT 10.10)

Prepare sample and determine optimal conditions

1. Thoroughly homogenize sample to be analyzed and dialyze (APPENDIX 3C) overnight at 4°C against a 100-fold excess of sodium formate buffer, pH 5.5, using 12,000-

**Preparation and
Analysis of
Glycoconjugates**

17.16.3

MWCO tubing. Lyophilize dialysate and store frozen at -20°C until ready for analysis.

Dialysis is not necessary for relatively purified biological samples that do not contain low-molecular-weight contaminants. If not to be analyzed immediately, samples should be lyophilized and stored frozen at -20°C .

Crude biological samples may contain O-acetyl esterases, which could cause de-O-acetylation of sialic acids during the isolation procedure. These enzymes can be inactivated by adding a 100 mM stock solution of diisopropyl fluorophosphate (DFP) in isopropanol to a final concentration of 1 mM, then incubating 15 min on ice. As DFP is extremely toxic, safety recommendations should be studied carefully; gloves should be worn and the compound should be used in the fume hood.

2. Dissolve the sample in water. De-O-acetylate an aliquot of sample, release sialic acids with 0.1 N H_2SO_4 , and determine total sialic acid content by the TBA assay (UNIT 17.18).

Although the TBA assay uses harsher conditions for releasing the sialic acids and will cause some destruction, it is used here in a comparative fashion. Therefore, it is a valid indication of the recoveries from step to step.

3. Transfer three aliquots of the sample from step 1 (each containing ~ 10 nmol of sialic acids) to 5-ml glass culture tubes.
4. Add 10 M acetic acid to a final acetic acid concentration of 2 M, then add 1 vol of 1% BHT per 100 vol sample.

The BHT retards lipid peroxidation.

5. Using an 80°C heating block or oven, heat one aliquot 2 hr, another aliquot 4 hr, and the remaining aliquot 5 hr.

When a heating block is used, care must be taken that the transfer of heat to the tubes is adequate. See Basic Protocol 1, step 4 annotation, for proper technique.

6. After each sample has been heated for the required time period, lyophilize and de-O-acetylate, then determine total free sialic acid by the TBA assay (UNIT 17.18). Compare the numbers obtained to determine the time required for maximal release of sialic acids.

Release and purify sialic acids

7. In a 5-ml glass culture tube add 2 M acetic acid to the lyophilized sample from step 1 (0.4 ml for up to 1 mg) and incubate at 80°C for the time determined in steps 5 and 6. Cool to room temperature.
8. Transfer reaction mixture into 1000-MWCO dialysis tubing, then wash the test tube twice with water, each time transferring the washings into the dialysis tubing. Dialyze (APPENDIX 3C) overnight at 4°C against 10 vol water and retain the dialysate (solution outside the tubing).

An MWCO of 1000 is sufficient to allow the passage of free sialic acids. Bigger pores may allow more contamination by permitting bigger molecules to dialyze out.

When total volumes are small (100 μl to 1 ml), the dialysis tubing can be placed inside a test tube. In this case, it must be closed with knots instead of plastic clamps.

Assuming that equilibrium is achieved, a 10% loss (sialic acids remaining in the tubing) should be expected when dialyzing against 10 vol water. The recovery can be improved by repeating the dialysis.

- Using a 0.5 × 10-cm chromatography column or Pasteur pipet plugged with glass wool, prepare a 1-ml column of Dowex AG 50W-X2 ion-exchange resin (see UNIT 10.10) in a cold room or cold box, using ice-cold water to suspend the resin and wash the column.

The resin must be extensively washed with ice-cold water before use. All chromatography steps must be performed at 4°C and eluants must be ice cold.

- Load the dialysate from step 8 directly to the column. Wash column with 6 ml cold water and collect eluant in a 15-ml glass test tube on ice.

For large preparations, every step can be scaled up, but it may be necessary to lyophilize the dialysate and dissolve it in a small volume of water before loading it onto the column. The content of the dialysis tubing (the retentate, containing the desialylated glycoconjugates under study) can also be frozen, lyophilized, and reserved for other analyses.

- Check pH of eluant with pH paper. If >3, go to step 12. If <3, add 10 mM sodium formate buffer, pH 5.5, until pH is >3.

- Using a 0.5 × 10-cm chromatography column or Pasteur pipet plugged with glass wool, prepare a 1-ml column of Dowex 3-X4A ion-exchange resin (hydroxyl form). Convert resin to formate form by equilibrating in 3 vol of 1 M formic acid. Let column stand in contact with formic acid for 15 min, then wash with 10 mM sodium formate (pH 5.5) until effluent pH is stable at 5.5 (see UNIT 10.10).

Large batches of resin can be prepared and stored at 4°C. Always take into consideration the capacity of the resin and calculate the volume of the column required to bind the total sialic acids present in the sample as determined by the TBA assay in step 2.

- Load sample from step 11 on column. Wash column immediately with 7 ml ice-cold 10 mM formic acid, and discard washings.

- Elute sialic acids with 10 ml ice-cold 1 M formic acid. Collect eluant on ice in a 20-ml glass test tube, then evaporate it to dryness by lyophilization or in a shaker-evaporator with the temperature of the water bath maintained at <37°C. Store at -20°C until analysis.

The purified mixture of sialic acids can be analyzed by HPLC (UNIT 17.19A) or derivatized and analyzed by reversed-phase HPLC with fluorometric detection (UNIT 17.18).

RELEASE AND PURIFICATION OF 4-O-ACETYLATED SIALIC ACIDS

A harsher acid hydrolysis procedure is used to liberate 4-O-acetylated sialic acids, which are resistant to release by the method indicated in Basic Protocol 2.

Additional Materials (also see Basic Protocol 2)

23.4 M (concentrated) formic acid
Heating block

- Prepare the sample and determine total sialic acid content in an aliquot of the material by TBA assay (see Basic Protocol 2, steps 1 and 2).
- Adjust dialyzed sample to pH 2.0 by dropwise addition of concentrated formic acid. Heat at 70°C for 1 hr. Cool to room temperature, and centrifuge at room temperature using a tabletop centrifuge to separate supernatant and pellet. Separate supernatant by aspiration and retain both supernatant and pellet.

When a heating block is used, care must be taken that the transfer of heat to the tubes is adequate. See Basic Protocol 2 step 4 annotation for proper technique.

ALTERNATE PROTOCOL

Preparation and Analysis of Glycoconjugates

17.16.5

3. Dialyze supernatant overnight at 4°C against 10 vol water using 1000-MWCO tubing, then lyophilize the solution outside the tubing. Repeat the dialysis twice again and pool the recovered material.

When total volumes are small (100 µl to 1 ml), the dialysis tubing can be placed inside a test tube. In this case, it must be closed with knots instead of plastic clamps.

4. Treat the pellet obtained in step 2, resuspended in a similar volume of water, twice more at pH 2 with heating as described in step 2, centrifuging and dialyzing the supernatants each time as described in step 3 to recover the hydrolysis products.
5. Pool the lyophilized dialysates and purify the mixture of sialic acids by ion-exchange chromatography (see Basic Protocol 2, steps 9 to 14). The annotation to Basic Protocol 2, step 14 describes options for analysis.

STRONG ACID HYDROLYSIS FOR QUANTITATIVE RELEASE OF HEXOSES, PENTOSES, HEXOSAMINES, AND URONIC ACIDS FROM GLYCOCONJUGATES

This protocol describes several sets of conditions for the acid hydrolysis of monosaccharides from glycoconjugates. When no data is available from previous experiments regarding optimal conditions, it is recommended that one try more than one set of conditions to determine which ones are optimal. It is also important to prove the reproducibility of the method under the conditions chosen. The mixtures of free monosaccharides obtained by acid hydrolysis can be directly analyzed by high-pH anion-exchange chromatography with pulsed amperometric detection (HPAE-PAD) or by derivatization followed by GLC (see *UNIT 17.19A* for both techniques).

Materials

- Glycoconjugate-containing sample to be analyzed
- Internal standard (i.e., a monosaccharide that does not occur naturally in the sample)
- Standard mixture (see recipe)
- 2 M and 4 M trifluoroacetic acid prepared from concentrated TFA (HPLC/spectra grade sequanal quality, Pierce)
- 99% methanol (anhydrous; e.g., Aldrich)
- Heating block or oven
- Small glass culture tubes or 3.5-ml glass vials with Teflon-lined screw caps
- Nitrogen (N₂) or vacuum evaporation system: e.g., Speedvac (Savant) or shaker-evaporator (Baxter Scientific)
- Additional reagents and equipment for phenol–sulfuric acid assay for monosaccharides (*UNIT 17.9*) and compositional analysis of monosaccharides released from glycoproteins (*UNIT 17.19A*)

1. If sufficient material is available, estimate the total carbohydrate content of the sample to be analyzed by the phenol–sulfuric acid assay (*UNIT 17.9*).

Analysis can be done without having this information when sample is precious, as the analytical tools used for the compositional analysis of acid hydrolysates (i.e., HPAEC-PAD, GLC) are far more sensitive than the colorimetric assay.

2. Using N₂, a Speedvac system, or a lyophilizer, evaporate three aliquots of the sample to be analyzed (each containing 5 to 50 µg carbohydrate) to dryness in small glass tubes or 3.5-ml glass vials. If the planned compositional analysis will be done by GLC, add 1 to 10 µg of internal standard to the sample before evaporating it.

Rhamnitol and arabinitol are good choices as internal standards for analysis of animal glycoconjugates, and inositol for plant polysaccharides.

3. Prepare (as in step 2) three aliquots of the standard mixture appropriate for the monosaccharides to be determined and the analytical method to be used, and three blanks to detect possible contamination. For GLC analysis, also add the internal standard.

The time involved in HPLC or GPC analysis of each sample is quite long (see UNIT 17.19A); therefore, preparation of a mixture of the appropriate standards in equimolar amounts helps reduce the analysis time. On the other hand, when identification of one particular component is doubtful, coinjection of one aliquot of the sample with one aliquot of the suspected standard into the GLC or HPLC column gives a definitive answer on the presence of that particular monosaccharide. In such cases, processing each monosaccharide standard individually in parallel can be helpful.

An appropriate blank is a material that has been processed exactly as the sample was (e.g., a pool of fractions from a chromatography step that showed only background response to the method used for monitoring). Blanks are especially useful when very small amounts of material are processed, as in such cases the limit of sensitivity will be determined by the accumulated contaminants (e.g., from Sephadex columns or dirty glassware).

4. Dissolve the three evaporated aliquots of each sample, standard, and blank in 100 to 500 μ l of 2 M TFA. Flush with N_2 and cap the tubes.
5. Using a 100°C heating block or oven, heat one sample, one standard, and one blank 3 hr; another sample, standard, and blank 4 hr; and the remaining sample, standard, and blank 6 hr.

This time course will determine the best conditions for maximum recovery with minimum destruction. When insufficient material is available for a time-course experiment, use 3 hr of hydrolysis for glycoproteins, glycopeptides, or free oligosaccharides of the N- and/or O-linked type that contain hexosamines and neutral monosaccharides. Use 6 hr of hydrolysis when the sample is suspected to contain glycosyluronic acids (e.g., proteoglycans or glycosaminoglycan chains).

When a heating block is used, care must be taken that the transfer of heat to the tubes is adequate. See Basic Protocol 1 step 4 annotation for proper technique.

6. After heating, evaporate solution in each tube using a nitrogen or vacuum evaporation system. Wash residues twice with 99% methanol, then evaporate again.

Store the dry hydrolysates at -20°C until ready for compositional analysis.

7. Analyze the hydrolysate by HPAE-PAD or derivatize and analyze by GLC (see UNIT 17.19A).

8. Compare the recovery of each monosaccharide obtained using the different hydrolysis times and decide upon the most suitable hydrolysis time for the sample under study. Repeat the analysis using such conditions and check for reproducibility.

If quantitation remains uncertain, hydrolysis can be repeated with the addition of a known amount of an internal standard, even when HPLC is used for the analysis. Because the constituents will be known at this point, any noninterfering monosaccharide or alditol can be used as internal standard. The choice will depend on the method used for analysis. When the analysis is done by HPAE-PAD, monosaccharide alditols cannot be used as internal standards because they are not retained by the Carbo-Pak PA-1 column used for monosaccharide separation. In this case, 2-deoxyrhamnose or 2-deoxyglucose are appropriate as internal standards. Prepare an aqueous solution of the standard and add a known amount to the sample, evaporating them together before dissolving in the acid at step 4. The standard must be present in an amount comparable to that of the major constituents of the sample.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Sodium formate buffer, pH 5.5

Weigh 68 mg sodium formate, dissolve in water, then bring volume to 1 liter. Adjust to pH 5.5 by adding 1 M formic acid dropwise.

Standard mixture

For GLC of alditol acetates: Prepare an equimolar mixture of the following:

D-(+)-glucose (anhydrous, mixed anomers; Sigma)

D-(+)-galactose (crystalline; Sigma)

D-(+)-mannose (crystalline; Sigma)

D-(+)-xylose (99%; Sigma)

For HPAE-PAD: To the above mixture (without xylose), add equimolar amounts of:

N-acetyl-D-galactosamine (Sigma)

N-acetyl-D-glucosamine (Sigma)

D-glucuronolactone (Sigma; when uronic acid is expected)

If xylose is expected, include it in the mixture, but be careful to use the right conditions for HPAEC-PAD (UNIT 17.19A).

COMMENTARY

Background Information

Fucose release

Fucosyl residues are typically in nonreducing terminal positions and are therefore susceptible to acid hydrolysis under relatively mild conditions. However, specific sites of attachment can increase the stability of the glycosidic bond (e.g., a positively charged amino group produced by de-*N*-acetylation of an *N*-acetylhexosamine residue will lead to a stable glycosidic linkage at the 3 position). Thus, it is necessary to determine for each specific glycoconjugate the conditions required for complete liberation with minimal destruction. The parameters to be varied are the type of acid, concentration of acid, and time and temperature used for the hydrolysis. The set of conditions suggested in Basic Protocol 1 are the best starting point.

Most of the problems encountered when trying to determine the best hydrolysis conditions for releasing fucose residues from a given glycoconjugate have been reviewed by Gottschalk (1972). Typical conditions are 0.5 M hydrochloric acid at 100°C for 16 hr (Schrager and Oates, 1968), although 15 min at 100°C in 1 M HCl are enough to release fucose from orosomucoid (Gottschalk, 1972), and 0.6 N sulfuric acid at 100°C for 1 hr is required to release all the fucose from total serum proteins (Gyorky and Houck, 1965).

Sialic acid release

More than 25 different kinds of modified sialic acids have now been reported to be found in nature. Most arise from substitution of the parent molecule with a variety of different groups. These modifications have been shown to affect a wide spectrum of biological phenomena. Many studies of sialic acids have failed to recognize the complexity of this family of sugars, because conventional methods for analyzing glycoconjugates—e.g., methanolysis, hydrazinolysis, methylation analysis, and β -elimination—result in the destruction of *O*-substituted species, especially *O*-acetyl esters.

The classical method for releasing sialic acids from glycoconjugates (incubation with 0.1 N H₂SO₄ for 1 hr at 80°C) results in extensive destruction of *O*-acetyl groups (Schauer, 1978). In addition, the presence of *O*-acetyl groups makes the sialic acid molecule partially or completely resistant to release by all available bacterial and viral neuraminidases (Schauer, 1978, 1982; Schauer and Corfield, 1982). The use of milder conditions (incubation with 0.5 M formic acid for 1 hr at 80°C) allowed the release and identification of many previously undetected *O*-acetylated sialic acids (Schauer 1978, 1982). However, even under these conditions quantitative analysis is not possible as a result of significant destruction and incomplete release of *O*-acetylated species using formic acid (pH 2.1) for acid hydrolysis (Schauer, 1978; Varki and Kornfeld, 1980;

Varki and Diaz, 1983). Prolonged hydrolysis in 2 M acetic acid (pH 2.4 to 2.5) at 80°C was found to achieve maximal release of *O*-acetylated sialic acids with minimal loss of *O*-acetyl groups (Varki and Diaz, 1983). These latter conditions have been successfully used to release and purify labile substituted sialic acids from different biological sources (Manzi et al., 1990). Hydrolysis times vary for different glycoconjugates. When sialic acids are acetylated at the 4 position, complete hydrolysis is difficult to achieve, but once released, the acetyl group becomes very labile, producing a 4,7-anhydro species.

Further complexity arises from the migration of *O*-acetyl groups from the 7 or 8 positions to the thermodynamically more stable 9 position of the sialic acid exocyclic side chain. This migration occurs rapidly, with $t_{1/2}$ ranging from minutes to hours, depending on the pH and temperature (Kamerling and Vliegthart, 1989; Varki and Diaz, 1983). Migration is minimal between pH 3 and 5.

A recent review (Varki, 1992) covers the diversity in the sialic acids, including the different problems encountered during their analysis and the methods available to overcome these problems.

Strong acid hydrolysis with TFA

Hydrolysis of glycosidic linkages involving hexoses requires more vigorous conditions than the ones used for fucose or sialic acids. Typically, hydrochloric (1 or 2 M, 100°C, 1 to 6 hr), sulfuric (0.1 to 2 N, 100°C, 4 to 12 hr), or trifluoroacetic acid (2 M, 121°C, 1 to 2 hr) have been used for this purpose, under conditions that vary depending on the substrate. Provided that the most labile species—e.g., sialic acids—are quantitated separately, the most severe problem encountered when determining the monosaccharide composition is the incomplete release of 2-acetamido sugars and glycosyluronic acids. The presence of these sugars can often be predicted based on the origin of the glycoprotein under study. In the case of acetamido sugars, if *N*-deacetylation occurs first, the stability of the glycosidic linkage is increased. Therefore, the conditions must be such that hydrolysis of the glycosidic linkage is substantially faster than the de-*N*-acetylation reaction. The glycosidic linkage of glycosyluronic acids is particularly resistant to acid hydrolysis, and further inhibition of their complete release occurs when the adjacent sugar in the chain is an *N*-sulfated or *N*-acetylated hexosamine. This problem is routinely encoun-

tered in glycosaminoglycan chains, where quantitative release requires a previous reduction of the carboxylic acid. On the other hand, liberated glycosyluronic acids are very susceptible to degradation, particularly by decarboxylation. Therefore, analysis of glycosaminoglycan chains is better achieved using enzymes (see UNIT 17.19B).

When enough material is available, recovery and reproducibility can be assessed under several different sets of conditions. In many cases this is not possible, and incubation with 2 M TFA for 3 hr at 100°C is recommended under such circumstances.

A review of the different conditions used for the release of monosaccharides from glycoproteins, including a detailed analysis of the mechanisms of acid hydrolysis of glycosides, is provided by Gottschalk (1972). In most cases, hydrolysis with TFA will produce as good a recovery of fucose as hydrolysis with 0.5 M HCl, although in the former case other sugars are also released. Where the sample is incubated with 2 M TFA at 121°C, 25% of the fucose decomposes after 6 hr. However, because only 1 hr is required for complete release, this rate of decomposition is tolerable (Albersheim et al., 1967). Where the sample is heated at 100°C, 98% recovery of fucose is achieved even after 6 hr (Biermann, 1988). Usually, estimation of the content of hexosamines in glycoproteins has been based on HCl hydrolysis at different concentrations (ranging from 2 M to 6 M) at temperatures close to 100°C for varying periods of time (ranging from 1 to 24 hr; Gottschalk, 1972). Also, incubation with trifluoroacetic acid (0.5 M, 100°C, 19 hr; 4 M, 125°C, 1 hr; or 4 M, 100°C, 4 hr) has been shown to give good results (Keene et al., 1983; Neeser and Schweizer, 1984). The variety of methods clearly indicates that the effectiveness of the hydrolysis greatly depends on the glycoconjugate—e.g., the type of linkage, the type of aglycone (noncarbohydrate portion), and the presence of substituents such as *N*-acetyl groups. Mannosamine is less acid-stable than glucosamine and galactosamine (Ludowieg and Benmaman, 1967). Analysis of the rates of decomposition of monosaccharides in the presence of 2 M (10 mg/ml) TFA—by continuing hydrolysis of polysaccharides after most of the monosaccharides have been liberated—have indicated that after 6 hr at 121°C, >50% of the xylose and arabinose and >25% of the galactose, rhamnose, and fucose are decomposed (Albersheim et al., 1967). However, as most of the monosaccharides are liberated within 1 hr,

these rates of decomposition are tolerable. Hydrolysis of urinary glycoconjugates with 2 M TFA for up to 15 hr at 100°C indicated that the maximum yield of aldoses and uronic acids is obtained after 6 hr. The recovery of the individual monosaccharides added at the beginning of the treatment was >93% in all cases, including that of glucuronic acid (Honda et al., 1981). A comparison between the action of different acids on each type of carbohydrate-containing molecule was done by Biermann (1988). Conditions used for acid hydrolysis of glycoconjugates before GLC analysis have also been reviewed by Kamerling and Vliegthart (1989).

Critical Parameters

Fucose release. Use of high-quality acids and highly purified deionized water is recommended. In the case of glycoproteins, destruction arising from the interaction between liberated sugars and amino acids can be minimized by using a low concentration of glycoprotein to perform the hydrolysis (<1 mg/ml) and excluding oxygen (by displacing it with dry nitrogen or performing the hydrolysis in a tube sealed under vacuum).

Sialic acid release. All the parameters involved in the protocol (e.g., temperature, pH, and concentration) must be carefully controlled. Because pH values outside the range of 3 to 7 should be avoided, strongly basic anion-exchange resins (e.g., Dowex 1 or 2) should not be used. Samples must be fresh, and must be processed and analyzed immediately. It is also critical that purified samples be kept dry at -70° or -80°C while stored. However, even storage at these temperatures for extended periods of time can result in migration or loss of *O*-acetyl groups. It should be pointed out that optimal methods for the release and purification of the rare forms of sialic acids (e.g., *O*-methylated, *O*-sulfated, or multiply modified) have not been adequately worked out.

Strong acid hydrolysis with TFA. High-quality acids must be used and highly purified deionized water must be used to dilute the acid, because trace iron can destroy hexosamines. Degradation is also avoided by excluding oxygen (usually by flushing with nitrogen, although it is also possible to work under vacuum). Working at a low concentration of glycoconjugate (<1 mg/ml) reduces the destruction of sugars. A competing sugar not present in the glycoprotein can also be added to reduce destruction (ribose has been used for this purpose). It must be kept in mind that—

even when treating the standards in parallel—they may not reflect the decomposition suffered by the glycosidically bound monosaccharides in the sample, because in the latter case they spend less time as free glycoses. Therefore, a method that yields less decomposition is always preferred to a correction factor worked up on the basis of monosaccharide standards.

Anticipated Results

Fucose release. Basic Protocol 1 will cleave the glycosidic linkages of all terminal nonreducing fucosyl residues. Once free, the fucosyl residues will remain intact, giving a reasonable estimation of the amount of fucose in the glycoconjugate (UNIT 17.19). Complete defucosylation of a sample can be determined by analysis of the amount of fucose in the hydrolysate at different time points. However, the use of TFA (strong acid hydrolysis; Basic Protocol 3) permits an accurate estimation of fucose together with the rest of the sugars. Therefore, strong acid hydrolysis is recommended when an analysis of the total glycosyl composition is required. The most valuable use of Basic Protocol 1 is therefore the defucosylation of glycoconjugates for further analysis.

Sialic acid release. The conditions for release and purification of sialic acids used in these protocols ensure maximum release of most modified sialic acids (with the exemption of methylated or sulfated species), prevent migration of *O*-acetyl groups from the 7 position to the 9 position, and prevent de-*O*-acylation. Therefore, careful use of the suggested protocols will result in the most accurate possible quantitation of the different sialic acid species present in a glycoconjugate. Sialic acids obtained by these procedures can be analyzed by HPAEC-PAD (UNIT 17.19A) or other HPLC techniques with (UNIT 17.18) or without derivatization (UNIT 17.19A). Therefore, an accurate quantitation also depends on the careful use of the final method of analysis. Basic Protocol 2 can be used to obtain a mixture of Neu5Ac, Neu5Gc, and of some of the mono-, di-, and tri-*O*-acetylated derivatives of these sialic acids from commercially available standards such as bovine submaxillary mucin. A mixture of standards will be very useful when analyzing the presence of sialic acid modifications in a glycoconjugate.

Strong acid hydrolysis with TFA. Careful application of Basic Protocol 3 will result in hydrolysis of the glycosidic linkages of hexoses, hexosamines, and uronic acids in glycoconjugates with minimal destruction.

Time Considerations

Fucose release. Because an 8-hr reaction is involved, it is better to prepare the samples the previous day. All hydrolysis and recovery procedures can be completed during the second day. Any additional time involved depends on the method used for analysis.

Sialic acids. After the time course to determine the time required for complete hydrolysis, release and purification require 2 days. Analysis must then be done immediately.

Strong acid hydrolysis with TFA. Preparation, hydrolysis, and drying of the hydrolysate should take ≤ 8 hr. The time required for analysis (by HPLC or GLC) is indicated in UNIT 17.19A.

Literature Cited

- Albersheim, P., Nevins, D.J., English, P.D., and Karr, A. 1967. A method for the analysis of sugars in plant cell-wall polysaccharides by gas-liquid chromatography. *Carbohydr. Res.* 5:340-345.
- Biermann, C.H. 1988. Hydrolysis and other cleavages of glycosidic linkages in polysaccharides. *Adv. Carbohydr. Chem. Biochem.* 46:251-271.
- Gottschalk, A. 1972. Glycoproteins: Their Composition, Structure, and Function (A. Gottschalk, ed.), 2nd ed., Part A, pp. 225-299. Elsevier/North-Holland, Amsterdam.
- Gyorky, G. and Houck, J.C. 1965. The determination of terminal protein-bound fucose. *Can. J. Biochem.* 43:1807-1811.
- Honda, S., Suzuki, S., Kakahi, K., Honda, A., and Takai, T. 1981. Analysis of the monosaccharide compositions of total nondialyzable urinary glycoconjugates by the dithioacetal method. *J. Chromatogr.* 226:341-350.
- Kamerling, J.P. and Vliegthart, J.F.G. 1989. Carbohydrates. In *Clinical Biochemistry: Principles, Methods, Applications*. Vol. 1, Mass Spectrometry (A.M. Lawson, ed.) pp. 175-263. Walter de Gruyter & Co., Berlin.
- Keene, L., Lindberg, B., Madden, J.K., Lindberg, A.A., and Gemski, P. 1983. Structural studies of the *Escherichia coli* O-antigen 25. *Carbohydr. Res.* 122:249-256.
- Ludowieg, J. and Benmaman, J.D. 1967. Colorimetric differentiations of hexosamines. *Anal. Biochem.* 19:80-88.
- Manzi, A.E., Dell, A., Azadi, P., and Varki, A. 1990. Studies of naturally occurring modifications of sialic acids by fast-atom bombardment-mass spectrometry. Analysis of positional isomers by

periodate cleavage. *J. Biol. Chem.* 265:8094-8107.

- Neeser, J.R. and Schweizer, T.F. 1984. A quantitative determination by capillary gas-liquid chromatography of neutral and amino sugars (as *O*-methyloxime acetates), and a study on hydrolytic conditions for glycoproteins and polysaccharides in order to increase sugar recoveries. *Anal. Biochem.* 142:58-67.
- Schauer, R. 1978. Biosynthesis of sialic acids. *Methods Enzymol.* 50:64-89.
- Schauer, R. 1982. Chemistry, metabolism, and biological functions of sialic acids. *Adv. Carbohydr. Chem. Biochem.* 40:131-234.
- Schauer, R. and Corfield, A.P. 1982. In *Sialic Acids: Chemistry, Metabolism and Function* (R. Schauer, ed.) pp. 51-57. Springer-Verlag, New York.
- Schrager, J. and Oates, M.D.G. 1968. The carbohydrate components of the hydrolysates of gastric secretion and extracts from mucous glands of the gastric body mucosa and antrum. *Biochem. J.* 106:523-529.
- Varki, A. 1992. Diversity in the sialic acids. *Glycobiology* 2:25-40.
- Varki, A. and Diaz, S. 1983. The release and purification of sialic acids from glycoconjugates: Methods to minimize the loss and migration of *O*-acetyl groups. *Anal. Biochem.* 137:236-247.
- Varki, A. and Kornfeld, S. 1980. An autosomal dominant gene regulates the extent of 9-*O*-acetylation of murine erythrocyte sialic acids. *J. Exp. Med.* 152:532-544.

Key References

- Biermann, 1988. See above.
- Gottschalk, 1972. See above.

These references discuss the mechanisms of acid hydrolysis of the different monosaccharides, and the influence of structure in determining alternative mechanisms. They also review many different hydrolysis conditions, and illustrate application of particular conditions to different types of glycoconjugates and the results obtained.

Varki and Diaz, 1983. See above.

Describes the parameters that must be varied to arrive at the optimal conditions used in the method described here.

Contributed by Adriana Manzi
University of California San Diego
La Jolla, California

Enzymatic Release of Oligosaccharides from Glycolipids

UNIT 17.17A

USE OF ENDOGLYCOCERAMIDASE FOR SELECTIVE CLEAVAGE OF GLYCOLIPIDS

BASIC
PROTOCOL

Selective cleavage of the oligosaccharide moiety of glycolipids for further structural analysis can be achieved by means of endoglycoceramidase (EGCase), an enzyme specific for the linkages between oligosaccharide and ceramide residues in glycolipids. The method can be used for analysis of glycolipids present in macroscopic amounts or for glycolipids that have been radiolabeled in their carbohydrate moiety. A variety of different methods can be used for analysis of the released oligosaccharides.

Materials

Sample to be analyzed—neutral glycolipids (UNIT 17.2), gangliosides (UNIT 17.17B), or desialylated gangliosides (see UNIT 17.12)—in organic solvent

1 mg/ml disialoganglioside-G_{D1a} (Sigma) in 2:1 (v/v) chloroform/methanol

0.5% (w/v) sodium taurocholate in 0.05 M sodium acetate buffer, pH 6.0

Endoglycoceramidase (EGCase; from *Rhodococcus sp.*; for purification see Ito and Yamagata, 1989a)

Toluene

Methanol

60:40:9 (v/v/v) chloroform/methanol/0.2% (w/v; aqueous) calcium chloride

Orcinol/sulfuric acid reagent (see recipe)

2:1 (v/v) chloroform/methanol

Bath sonicator

Precoated high-performance thin-layer (HPTLC) silica gel 60 plates (10 × 10-cm; Merck)

Glass chromatography tank

140°C oven

Tabletop centrifuge

Digest sample with EGCase

1. Evaporate an aliquot of the sample to be analyzed to dryness in a glass vial. Simultaneously evaporate 10 µl of 1 mg/ml disialoganglioside-G_{D1a} in the same manner, as a positive control.

When analyzing complex mixtures of unknown composition, estimate the amount of sample to be aliquotted by carrying out this protocol with two or three different amounts of sample.

Evaporation can be done by simply leaving the glass vial containing the chloroform/methanol solution open to the atmosphere or by using a Speedvac evaporator.

2. Dissolve the evaporated residues in 50 µl 0.5% sodium taurocholate in 0.05 M sodium acetate buffer, pH 6.0 by brief sonication in a bath sonicator.

0.4% Triton X-100 can be used in place of sodium taurocholate. EGCase molecules aggregate in the absence of detergents.

3. Add 12 mU of EGCase to each sample. Place the vials containing the samples, with caps loosened, inside a closed container in which an open tube of toluene is also placed to create a toluene atmosphere. Begin incubating at 37°C.

Preparation and
Analysis of
Glycoconjugates

17.17.1

Contributed by Adriana E. Manzi

Current Protocols in Molecular Biology (1995) 17.17.1-17.17.5

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 32

- At intervals of 0, 1, 3, 6, 15, and 20 hr of incubation, take an aliquot of the reaction mixture with a volume corresponding to ≥ 1 nmol of each glycolipid. Dilute each aliquot with 9 vol methanol.

Determine optimal incubation time for cleavage

- Apply an aliquot of each diluted sample hydrolysate and G_{D1a} hydrolysate (1 to 2 nmol) along a line drawn 1 cm from the bottom of an HPTLC silica gel 60 plate.

Depending on how many points are used, multiple plates may be required. Five to seven samples can be spotted per plate.

The plates must be marked in advance with pencil lines 1 cm from the top and bottom of the plate. The bottom line should be further divided with pencil marks into 1-cm lanes with ~ 0.5 -cm spacings between lanes. The lanes will correspond to the different incubation times used in step 4.

- Air dry plate and place in a glass chromatography tank that has been equilibrated with 60:40:9 chloroform:methanol:0.02% aqueous calcium chloride.

The HPTLC is to separate the glycolipids. Released oligosaccharides can be analyzed by HPTLC on silica gel 60 plates using 2:1:1 (v/v/v) 1-butanol:glacial acetic acid:water as the developing solvent.

- Let the solvent ascend until it reaches the upper line, then take the plate out of the tank and allow to thoroughly air dry in a fume hood.

- Spray with orcinol-sulfuric acid reagent and heat at 140°C for 5 min.

Alternatively, if the carbohydrate moiety has been radiolabeled, the plate can be submitted to autoradiography.

- Note the lane in which complete release of oligosaccharides is indicated by the disappearance of the glycolipid bands that are present in the starting material.

If no lane indicating complete reaction is seen, longer incubation periods should be tried. The time points indicated to step 4 are for analysis of purified glycolipids for which some information is available. A time course up to 72 hr is suggested for unknown glycolipid structures. In these cases, add 36 mU of EGCase instead of 12 mU.

Analyze released oligosaccharides

- Set up reaction as described in steps 1 to 3, incubate for the time determined to be optimal for complete release of oligosaccharides, then stop the reaction by adding 2 vol of 2:1 chloroform:methanol.

- Shake the mixture of aqueous cleaved glycolipid and organic solvent, then centrifuge 10 min at 2000 rpm in a tabletop centrifuge. Remove the upper (aqueous) phase containing the oligosaccharides and lyophilize. Store at -20°C .

The remaining glycolipids and ceramides will partition into the lower (organic) phase.

The dried sample can be dissolved in water and submitted to a variety of different methods of oligosaccharide analysis. The method to be chosen will depend on the availability of instrumentation, the complexity of the sample, the amount of material available, and the questions to be answered. Possible methods for the analysis of the released oligosaccharides include: ion-exchange HPLC of anionic oligosaccharides (UNIT 17.21A) and sizing HPLC of neutral oligosaccharides (UNIT 17.21B). High-performance anion-exchange chromatography on pellicular-resin columns can also be used for both anionic and neutral oligosaccharides. Radiolabeled samples are monitored by liquid scintillation counting of collected fractions, and nonradiolabeled ones by pulsed amperometric detection (PAD).

In most cases, dilution of the reaction mixture before injection into an HPLC system will suffice. When desalting by gel-filtration chromatography (UNIT 10.9) is required prior to further analysis, the released oligosaccharides can be separated from remaining glycolipids and ceramides and at the same time desalted by passing the reaction mixture from step 10 (without adding organic solvent) through a Dowex 50 (hydrogen form) gel-filtration column and washing with water. Long-chain nonpolar molecules will be retained by hydrophobic interactions and cations will be exchanged. The oligosaccharides that are recovered should be lyophilized. Alternatively, gel-filtration chromatography may be performed on the upper phase (step 11) using an 0.6 × 20-cm Sephadex G-10 column, eluting with water. The elution may be monitored by spotting aliquots of each fraction on a piece of silica gel-coated TLC plate and staining the carbohydrates with orcinol-sulfuric acid reagent (i.e., by spraying the dried plate and heating 15 min at 110°C). All fractions containing carbohydrates are then pooled and lyophilized.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Orcinol/sulfuric acid reagent

Weigh 200 mg orcinol (5-methylresorcinol·H₂O, formula weight 142.15; Aldrich). Dissolve in 100 ml ice-cold 2 M sulfuric acid. Store in a glass container at 4°C; discard when solution turns brownish.

COMMENTARY

Background Information

Endoglycosylceramidase (EGCase) is an enzyme capable of hydrolyzing the linkage between the oligosaccharides and ceramides of acidic and neutral glycolipids to produce free, intact oligosaccharide and ceramide residues. The enzyme does not cleave the linkages between monosaccharides and ceramides (in cerebroside) or between oligosaccharides and diacylglycerol (in glycoglycerolipids). Commercially available EGCase is isolated from the culture supernatant of *Rhodococcus* sp. and has an apparent molecular weight >16 kDa as determined by gel-filtration chromatography on Sephadex G-100. This preparation contains three isozymes with molecular weights of 60,000 to 75 kDa as determined by SDS-PAGE.

One unit of enzyme is defined as the amount needed to catalyze the release of 1 μmol/min of reducing power (as glucose) from a substrate of mixed bovine brain gangliosides at 37°C, pH 6.0. The specific activity is not less than 15 mU/mg of protein.

Chemical treatments—e.g., ozonolysis and periodate oxidation—have been used to obtain free oligosaccharides from glycosphingolipids. However, these reactions cause disruption of the ceramide residues.

Ito and Yamagata (1986) reported a novel glycosphingolipid-specific endoglycosidase—endoglycosylceramidase—or “endoglyco-

ceramidase” for short. The enzyme, isolated from the culture supernatant of *Rhodococcus* sp. strain G-74-2, was capable of cleaving the linkage between oligosaccharides and ceramide, releasing both parts intact. Later, similar enzymes called ceramide-glycanases were found in leeches (Li et al., 1986) and earthworms (Li et al., 1987). The crude enzyme, however, could not be used to release oligosaccharides from living cells because contamination with hemolysin caused cell death.

A variety of methods for the analysis of the released oligosaccharides have been developed. Some of them include a derivatization step to incorporate a fluorescent or hydrophobic tag, followed by HPLC analysis of the products (Shimamura et al., 1988; Rasilo et al., 1989). Analysis of the released oligosaccharides by HPAE-PAD (see UNIT 17.19A) has also been reported (Ito et al., 1991), as well as the use of the enzyme in the structural characterization of metabolically radiolabeled gangliosides (Manzi et al., 1990).

In 1989, a mutant strain of *Rhodococcus* sp., M-750, was isolated. This organism was capable of producing three forms of endoglycosylceramidase, and two of these forms were purified (Ito and Yamagata, 1989a). The three isoenzymes were separated from each other and their substrate-specificity examined with various glycosphingolipids. EGCase I (with a pI of 5.3)

and EGCase II (with a pI of 4.5) hydrolyze the glucosyl-ceramide linkage of globo-, lacto-, and ganglio-type glycosphingolipids. EGCase I hydrolyzes globo-type glycosphingolipids much faster than EGCase II. Neither of these two isoenzymes hydrolyzes the galactosyl-ceramide linkage for which EGCase III is specific (Ito and Yamagata, 1989b). A new mutant, M-777, produced EGCases at levels 5-fold higher than M-750 (Ito et al., 1991).

In spite of extensive purification of the enzymes, effective EGCase cleavage of glycosphingolipids in vivo while maintaining viability of the cells still presents two problems. First, the activity of the enzymes is very low in the absence of detergents, and a detergent capable of stimulating activity without impairing cell viability has not been found. Second, the optimal pH is quite acidic, which is another factor that prevents the study of the biological functions of glycosphingolipids. Recently, Ito et al. (1991) reported finding two activator proteins in the culture supernatant of *Rhodococcus* sp. strain M-777. Activator protein II was purified and shown to stimulate the activity of EGCase II much more than that of EGCase I, but was nonspecific with respect to the glycolipid. Activator protein I seems to be specific for stimulating the activity of EGCase I. Following the addition of activator protein II, EGCase II hydrolyzed cell-surface glycosphingolipids quite efficiently at a neutral pH where almost no hydrolysis occurred with EGCase II alone (Ito et al., 1991). In the near future, the use of activator proteins might permit the study of the biological functions of glycosphingolipids in living cells using EGCase.

Critical Parameters

If the activity of EGCase towards disialoganglioside-G_{D1a} is defined as 100% at pH 5 to 6, only 50% of that activity is observed at pH 7, and 20% at pH 4. The enzyme is completely inactive at pH 3. Maximum stability of the enzyme is observed between pH 6 and 8.

The enzyme is inhibited by Cu²⁺, Ba²⁺, and Hg²⁺. It is stable in buffered solution for 6 months at 4°C. Excessive freeze-thawing should be avoided.

The commercially available enzyme does not contain sphingomyelinase, β-galactosidase, β-mannosidase, β-galactosaminidase, β-glucosidase, α-fucosidase, or sialidase activity. Trace amounts of β-glucosaminidase, or α-mannosidase, as well as protease activity, can

sometimes be detected—this is indicated in the data sheet provided by the manufacturer.

Sonication of the substrate prior to the addition of enzyme is recommended.

It is always necessary to determine empirically the time required for complete cleavage, as different glycolipids exhibit different rates of hydrolysis.

Anticipated Results

The use of this protocol, in combination with appropriate analysis of the released oligosaccharides, will allow the structural characterization of the types of carbohydrate chains that are linked to glycolipids in a given system. However, as mixtures of glycolipids may be complicated, the usefulness of this protocol can be driven one step forward by studying purified glycolipids. Where this is done, the released oligosaccharides can be submitted to acid hydrolysis for composition analysis (UNITS 17.19A & 17.19B).

Time Considerations

Up to 3 days of incubation with the enzyme may be required for complete release. The time required for analysis of the released oligosaccharides depends on the method used.

Literature Cited

- Ito, M. and Yamagata, T. 1986. A novel glycosphingolipid-degrading enzyme cleaves the linkage between the oligosaccharide and ceramide of neutral and acidic glycosphingolipids. *J. Biol. Chem.* 261:14278-14282.
- Ito, M. and Yamagata, T. 1989a. Endoglycoceramidase from *Rhodococcus* species G-74-2. *Methods Enzymol.* 179:488-495.
- Ito, M. and Yamagata, T. 1989b. Purification and characterization of glycosphingolipid-specific endoglycosidases (endoglycoceramidases) from a mutant strain of *Rhodococcus* sp. Evidence for three species of endoglycoceramidase with different specificities. *J. Biol. Chem.* 264:9510-9519.
- Ito, M., Ikegami, Y., and Yamagata, T. 1991. Activator proteins for glycosphingolipid hydrolysis by endoglycoceramidases. Elucidation of biological functions of cell-surface glycosphingolipids in situ by endoglycoceramidases made possible using these activator proteins. *J. Biol. Chem.* 266:7919-7926.
- Li, S.-C., Degasperis, R., Muldrey, J.E., and Li, Y.-T. 1986. A unique glycosphingolipid-splitting enzyme (ceramide-glycanase from leech) cleaves the linkage between the oligosaccharide and the ceramide. *Biochem. Biophys. Res. Commun.* 141:346-352.

- Li, Y.-T., Ishikawa, Y., and Li, S.-C. 1987. Occurrence of ceramide-glycanase in the earthworm, *Lumbricus terrestris*. *Biochem. Biophys. Res. Commun.* 149:167-172.
- Manzi, A.E., Sjoberg, E.R., Diaz, S., and Varki, A. 1990. Biosynthesis and turnover of O-acetyl and N-acetyl groups in the gangliosides of human melanoma cells. *J. Biol. Chem.* 265:13091-13103.
- Rasilo, M.-L., Ito, M., and Yamagata, T. 1989. Liberation of oligosaccharides from glycosphingolipids on PC12 cell surface with endoglycoceramidase. *Biochem. Biophys. Res. Commun.* 162:1093-1099.
- Shimamura, M., Hayase, T., Ito, M., Rasilo, M.-L., and Yamagata, T. 1988. Characterization of a major neutral glycolipid in PC12 cells as III³Gal α -globotriaosylceramide by the method for determining glycosphingolipid saccharide sequence with endoglycoceramidase. *J. Biol. Chem.* 263:12124-12128.

Key References

Ito and Yamagata, 1986. See above.

Describes the purification of the enzyme, its specificity, and examples of its use.

Ito and Yamagata, 1989a. See above.

Describes the purification and properties of the enzyme and contains many references regarding its application.

Contributed by Adriana E. Manzi
University of California San Diego
La Jolla, California

Endo- β -Galactosidases and Keratanase

INTRODUCTION

Endo- β -galactosidases were discovered during the search for a keratan sulfate-degrading enzyme in the culture filtrates of *Coccobacillus* sp. (Hirano and Meyer, 1971), *Pseudomonas* sp. (Nakasawa and Suzuki, 1975), and *Escherichia freundii* (Kitamikado et al., 1970a,b,c). Later studies on the *E. freundii* keratanase demonstrated that this enzyme is capable of hydrolyzing a wide range of glycoconjugates—including nonsulfated oligosaccharides isolated from milk as well as the carbohydrate moieties of glycoproteins and glycolipids—thus expanding the applicability of this enzyme as an endo- β -galactosidase to a variety of complex glycoconjugates (Fukuda and Matsumura, 1975, 1976). Endo- β -galactosidase produced by *Bacterioides fragilis* has similar properties to that from *E. freundii* (Scudder et al., 1983). On the other hand, keratanase isolated from *Pseudomonas* sp. is specific to the sulfated poly lactosamines, and thus only hydrolyzes keratan sulfates (Nakazawa and Suzuki, 1975; Fukuda, 1985; Nakasawa et al., 1984). An endo- β -galactosidase found in *Diplococcus pneumoniae* showed strict specificity toward blood-group A and B structures (Takasaki and Kobata, 1976). More recently, an endo- β -galactosidase specific to Gal α 1 \rightarrow 3Gal was found in the culture filtrate of *Clostridium perfringens* (Fushuku et al., 1987).

ENDO- β -GALACTOSIDASE FROM *E. FREUNDII*

Source

The *E. freundii* line designated also as *Citrobacter freundii* EGI (Paoletti et al., 1990) can be obtained from American Type Culture Collection (ATCC; APPENDIX 4). The endo- β -glycosidase enzyme can be purified from the culture filtrate of this organism by affinity chromatography using side chain-free mucins.

Properties

Estimates of the molecular weight of the endo- β -galactosidase from *E. freundii* range from 28 to 32 kDa, depending on the method employed. The enzyme is stable in a solution of pH 4.5 to ~5.5, but unstable below pH 4.0 or above pH 6.0. It is stabilized by addition of 450 μ g/ml BSA to the enzyme solution. The activity of this enzyme is optimal between pH 5.5 and

5.8, using a sodium acetate buffer. The isoelectric point is estimated to be pH 8.0 (Nakagawa et al., 1980).

Substrates and Specificity

E. freundii endo- β -galactosidase hydrolyzes high-molecular-weight substrates at the middle of the carbohydrate chain by random depolymerization (Fukuda and Matsumura, 1976).

Cleavage of keratan sulfates

E. freundii endo- β -galactosidase hydrolyzes the β -galactosidic bond of the keratan sulfates as shown in Figure 17.17.1A, but does not cleave the bond if the adjacent galactose residue is sulfated at the C-6 position, as indicated in Figure 17.17.1B. The sulfate group on GlcNAc is not a necessary factor for *E. freundii* endo- β -galactosidase, as a nonsulfated disaccharide is released from the keratan sulfate. Sulfation of the galactose residue at the C-6 position, however, hinders hydrolysis of the galactosidic bond by endo- β -galactosidase.

Cleavage of glycoproteins

The usefulness of the *E. freundii* endo- β -galactosidase has been shown in detection of poly lactosaminoglycans in glycoproteins. The endo- β -galactosidase was found to be active upon human erythrocyte band 3 carbohydrate (Fukuda et al., 1979), large glycopeptides from mouse embryonal carcinoma F9 cells (Muramatsu et al., 1979), and glycopeptides isolated from a GM1 gangliosidosis patient (Tsay et al., 1975)—all of which were found to contain keratan-like poly lactosaminoglycans with the aid of this enzyme. Poly lactosaminoglycans are present in various glycoproteins—including a glucose-transporter protein, lysosomal membrane proteins, precursor of Rh antigen, fibronectin from human placenta, lymphocyte-associated antigen 1 (LFA-1) α and β -chains, and laminin.

Endo- β -galactosidase from *E. freundii* can hydrolyze galactose as shown in Figure 17.17.2. The galactose at the branch point is also susceptible, although with a low efficiency (Fukuda et al., 1984).

The endo- β -galactosidase can hydrolyze fucosylated lactosaminoglycans as shown in Figure 17.17.3, but the galactosyl linkage attached to the fucosylated N-acetylglucosamine

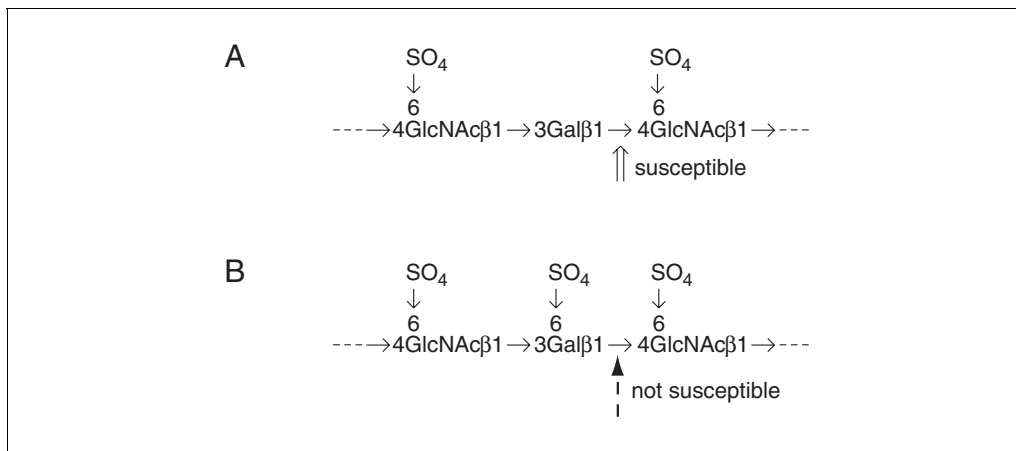


Figure 17.17.1 Hydrolysis of keratan sulfates by *E. freundii* endo- β -galactosidase. **(A)** Endo- β -galactosidase hydrolyzes β -galactosidic bond of keratan sulfates. **(B)** Hydrolysis hindered by sulfation of galactose residue at C-6 position.

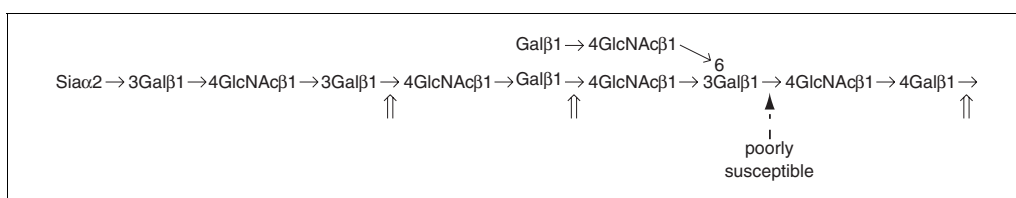


Figure 17.17.2 Hydrolysis of galactose in glycoproteins by *E. freundii* endo- β -galactosidase.

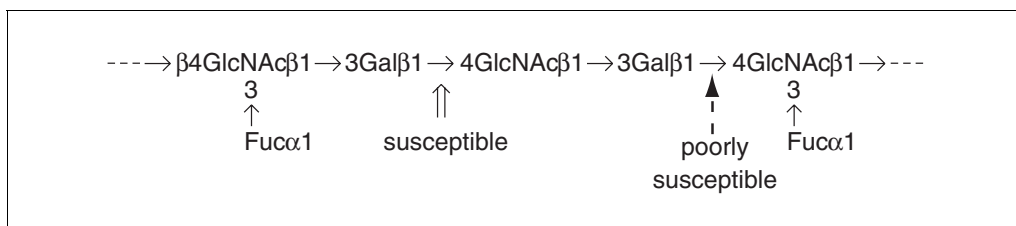


Figure 17.17.3 Hydrolysis of fucosylated lactosaminoglycans by *E. freundii* endo- β -galactosidase. The galactosyl residue attached to the fucosylated N-acetylglucosamine is poorly susceptible.

is poorly susceptible to the enzyme (Sponcer et al., 1984; Kannagi et al., 1982a).

Cleavage of glycolipids

The so-called lacto-series glycolipids ($R_3 \rightarrow \text{GlcNAc}\beta 1 \rightarrow 3\text{-Gal}\beta 1 \rightarrow 4\text{Glc} \rightarrow \text{Cer}$) are susceptible to endo- β -galactosidase (Fukuda et al., 1976). For example, endo- β -galactosidase hydrolyzes lacto-N-neotetraosylceramide, producing trisaccharide and glucosylceramide as shown in Figure 17.17.4A and lacto-N-neohexaosylceramide is hydrolyzed to produce two oligosaccharides—lacto-N-triaosylceramide and glucosylceramide—as shown in Figure 17.17.4B. Branched lacto-series glycolipids—e.g., lacto-iso-N-neooctaosylceramide—are hydrolyzed to a large

oligosaccharide and glucosylceramide as shown in Figure 17.17.4C.

Although the β -galactosidic linkage at the branch point in the erythrocyte polylectosaminoglycan-peptides is cleaved by endo- β -galactosidase, under practical conditions the hydrolysis of glycolipids at the branch point is too poor to be detected.

Hydrolysis of α_2 -glycolipid (Kannagi et al., 1982b) indicates that not only the $\text{GlcNAc}\beta 1 \rightarrow 3\text{Gal}\beta 1 \rightarrow$ structure, but also the $\text{GalNAc}\beta 1 \rightarrow 3\text{Gal}\beta 1 \rightarrow$ structure is susceptible to endo- β -galactosidase.

Cleavage of oligosaccharides

E. freundii endo- β -galactosidase can hydrolyze oligosaccharides isolated from human

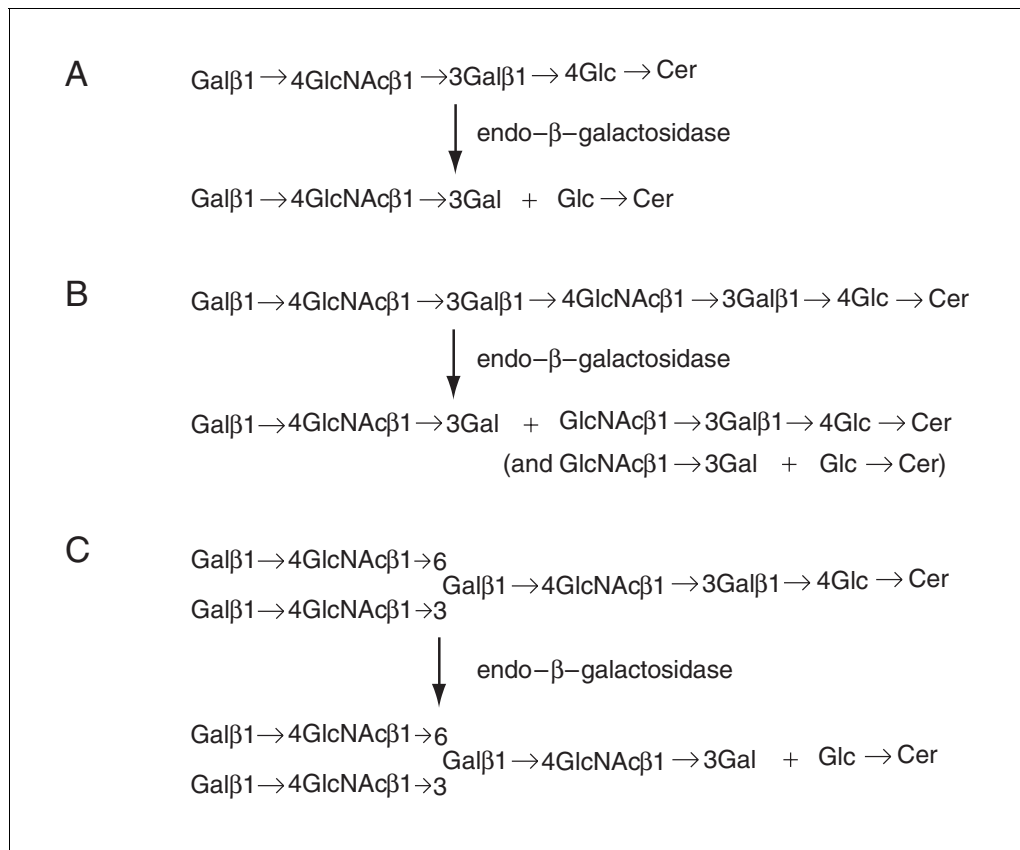


Figure 17.17.4 Hydrolysis of glycolipids by *E. freundii* endo-β-galactosidase. **(A)** Enzyme hydrolyzes lacto-N-neotetraosylceramide producing trisaccharide and glucosylceramide. **(B)** Enzyme hydrolyzes lacto-N-neohexaosylceramide to produce lacto-N-triaosylceramide and glucosylceramide. **(C)** Enzyme hydrolyzes branched lacto-series glycolipid to produce large oligosaccharide and glucosylceramide.

milk (Fukuda, 1981; Nakagawa et al., 1980). Affinity of the enzyme for oligosaccharide substrates is weaker than that for polylactosaminoglycan, keratan sulfate, and glycolipid substrates. A high concentration of the enzyme is needed for an efficient enzymic hydrolysis of oligosaccharides.

Application of Endo-β-Galactosidase to Various glycoconjugates

Glycopeptides

To digest glycopeptides with endo-β-galactosidase, mix 10 to 100 μg of the glycopeptide—metabolically labeled (UNIT 17.4) or unlabeled—and 25 mU of endo-β-galactosidase with 200 μl of 0.2 M sodium acetate buffer, pH 5.0 (APPENDIX 2), then incubate at 37°C overnight. The oligosaccharides released by endo-β-galactosidase are usually analyzed by gel filtration (UNIT 10.9). In preparation for this, the digest of unlabeled glycopeptides is reduced with NaB[³H]₄ (see UNIT 17.5). The resulting

[³H]-labeled oligosaccharide alcohols are then applied to a gel-filtration column packed with Sephadex G-50 (Pharmacia Biotech) and/or a column packed with Bio-Gel P-4 (Bio-Rad).

Upon Bio-Gel P-4 gel filtration, the oligosaccharide components are eluted at the elution volumes of disaccharides, trisaccharides, and tetrasaccharides; larger (or sialylated) molecules are eluted in the void volume. The results obtained for erythrocyte band 3 showed that the oligosaccharide alcohols released are: GlcNAcβ1→3GalOH, Galβ1→4GlcNAcβ1→3GalOH, and Fucα1→2Galβ1→4GlcNAcβ1→3GalOH.

The oligosaccharide components eluted in the void volume from a Bio-Gel P-4 column are sialylated oligosaccharides—because after desialylation these components are often eluted at the elution volumes of trisaccharides or larger oligosaccharides. The structure of the smallest sialylated oligosaccharide component released from band 3 was: NeuNAcα2→3Galβ1→4GlcNAcβ1→3GalOH.

Because disaccharides should be produced from the internal moieties of poly lactosamine in which tri- and tetrasaccharides represent nonreducing terminals. Thus, in the absence of fucosylation or sialylation the ratio of disaccharides to tri- and tetrasaccharides indicates the length of the poly lactosamine repeats. These analyses, combined with methylation analyses of intact poly lactosamines have yielded substantial structural information on glycoproteins.

Glycolipids

To hydrolyze glycolipids with endo- β -galactosidase, take an aliquot corresponding to 10 μ g of glycolipid dissolved in 2:1 (v/v) chloroform/methanol and place in a small glass test tube or microcentrifuge tube. Dry the glycolipid sample under N_2 gas. To the dried sample, add 10 μ l of 0.2 M sodium acetate buffer, pH 5.8 (APPENDIX 2) containing 20 μ g sodium deoxytaurocholate, as well as 10 μ l of 0.25 mU/ μ l endo- β -galactosidase. Incubate 2 hr to overnight to achieve enzymatic hydrolysis. Sialylated glycolipid—e.g., sialylparagloboside and linear lacto-series glycolipids—can be hydrolyzed easily under these conditions; stronger conditions—e.g., 10 \times more enzyme—are required for hydrolysis of branched and/or highly fucosylated glycolipids.

Intact cells (Fukuda et al., 1979)

Endo- β -galactosidase can directly modify cell-surface glycoproteins and glycolipids. When human erythrocytes are treated with endo- β -galactosidase, Ii antigens and ABO blood-group antigens are abolished or decreased. In addition to these antigenic changes, poly lactosamines on band 3 and glycoproteins on band 4.5 are hydrolyzed, though O-linked oligosaccharides in glycoporphins remain unaffected. Glycolipids on the cell surface are also hydrolyzed by endo- β -galactosidase, but only lacto-series glycolipids with long carbohydrate chains are hydrolyzed by this treatment. The innermost galactosyl linkage of glycolipids, R \rightarrow Gal β 1 \rightarrow 4Glc \rightarrow Cer, is not susceptible to the enzyme treatment on the cell surface. Endo- β -galactosidase treatment of intact cells, combined with cell-surface labeling using galactose oxidase or galactosyltransferase, has been useful for identifying poly lactosamine-bearing glycoproteins.

To perform cell-surface labeling and endo- β -galactosidase digestion, the galactose and

N-acetyl galactosamine residues on the cell surface can be specifically labeled by the galactose oxidase/ $NaB[^3H]_4$ method (Gahmberg and Hakomori, 1973). Wash the erythrocytes three times with cold PBS (APPENDIX 2), then suspend $\sim 2 \times 10^7$ cells (or 200 μ l of packed cells) in 1 ml PBS pH 7.4. Add 10 μ l of 1 U/ml galactose oxidase (Sigma) in PBS, pH 7.4 and incubate 2 hr at room temperature. After incubation, add 2.5 mCi $NaB[^3H]_4$ in 50 μ l 0.01 N NaOH and incubate 30 min at room temperature. Finally, add ~ 0.5 mg nonradioactive $NaBH_4$ and wash 3 \times with cold PBS (pH 7.0), centrifuging between washes. Resuspend 0.1 ml of these surface-labeled cells (as packed cells) in PBS pH 7.0 to a final volume of 200 μ l. To this cell suspension, add 25 μ l of 1 U/ml endo- β -galactosidase, then incubate 2 hr at room temperature. Wash the cells with 3 ml cold PBS (pH 7.0) and centrifuge. The pelleted cells can be used for glycoprotein and glycolipid analysis, and the supernatant can be analyzed for released oligosaccharides. This procedure can be used not only for erythrocytes, but also for nucleated cells.

To analyze glycoproteins and glycolipids of surface-labeled cells, prepare erythrocyte membranes from the surface-labeled cells that have been treated with or without endo- β -galactosidase as described above. Membrane lysates from nucleated cells—e.g., lymphocytes and fibroblasts—can be prepared by vortexing the cells in an equal volume of cold PBS containing 0.5% NP-40 and 1 mM PMSF. The nuclei are then separated out by centrifugation, after which the supernatant containing the cell lysate may be dissolved in SDS sample buffer and subjected to SDS-polyacrylamide gel electrophoresis (UNIT 10.2), followed by fluorography (APPENDIX 3A) to detect radioactive protein components.

To extract glycolipids from erythrocyte membranes or intact cells, these are dissolved in 10 vol of 2:1 (v/v) chloroform/methanol and the extract fractionated by Folch partition (UNIT 17.3) into lower and upper phases. If necessary, upper-phase glycolipids can be further separated into neutral and acidic glycolipids by DEAE-Sephadex. Each glycolipid fraction is finally analyzed by thin-layer chromatography followed by fluorography. Oligosaccharides released from the cell surface by endo- β -galactosidase treatment can also be analyzed by gel-filtration chromatography on Sephadex G-50 and by paper chromatography.

ENDO-β-GALACTOSIDASE FROM BACTEROIDES FRAGILIS (Scudder et al., 1983)

Source

Endo-β-galactosidase from *B. fragilis* is commercially available from Genzyme (APPENDIX 4).

Properties

The properties and specificities of the *B. fragilis* endo-β-galactosidase are similar to those of the corresponding *E. freundii* enzyme. The enzyme can be stored at -20°C for long periods of time; storage at a protein concentration of 44 μg/ml at 4°C for 2 to 3 months did not affect the activity, but storage at a low concentration (<3 μg/ml) causes a significant loss of activity. In the presence of 0.2 mg/ml of albumin, the enzyme retains its activity over a wide range of pH.

Substrates and Specificity (Scudder et al., 1983, 1984, 1986)

B. fragilis endo-β-galactosidase hydrolyzes keratan sulfate and produces sulfated and non-sulfated oligosaccharides, which are identical to the digestion products of keratan sulfate with *E. freundii* endo-β-galactosidase. Hydrolysis of oligosaccharides by endo-β-galactosidases from *B. fragilis* and *E. freundii* shows that both enzymes hydrolyze the internal Galβ1→4Glc linkage of type 1 and type 2 oligosaccharides. The type 2 oligosaccharide lacto-N-neotetraose was hydrolyzed to a greater extent than its type 1 isomer, lacto-N-tetraose. *B. fragilis* endo-β-galactosidase hydrolyzes the linear but not the branched carbohydrate chain of lacto-series glycolipids. Because the *B. fragilis* endo-β-galactosidase is quite similar to that from *E. freundii*, this enzyme can be applied widely to the analysis of glycoconjugates, as described for the *E. freundii* enzyme.

ENDO-β-GALACTOSIDASE FROM FLAVOBACTERIUM KERATOLYTICUS (Kitamikado et al., 1981)

Endo-β-galactosidase from *F. keratolyticus* has been purified, and the specificity of this enzyme has been compared with that of the *E. freundii* enzyme. *F. keratolyticus* endo-β-galactosidase is similar to the *E. freundii* enzyme, except that the *F. keratolyticus* enzyme can hydrolyze oligosaccharide substrates more efficiently.

BLOOD-GROUP TYPE A and B-SPECIFIC

ENDO-β-GALACTOSIDASE FROM DIPLOCOCCUS PNEUMONIAE (Takasaki and Kobata, 1976)

Culture filtrate from *Diplococcus pneumoniae* (also called *Streptococcus pneumoniae*) is known as a rich source of endo- and exoglycosidases including β-galactosidase, β-N-acetylglucosaminidase, endo-α-N-acetylglucosaminidase, endo-N-acetylglucosaminidase (endo-D), endo-β-galactosidase DII, and blood group A and B-specific endo-β-galactosidase. The last enzyme hydrolyzes the structure shown in Figure 17.17.5. It hydrolyzes the blood-group A determinant faster than the B determinant, and hydrolyzes the oligosaccharides with A and B determinants that are composed of type 2 chains, but not those composed of type 1 chains.

ENDO-β-GALACTOSIDASE DII FROM DIPLOCOCCUS PNEUMONIAE (Fukuda, 1985)

An endo-β-galactosidase that hydrolyzes the internal β-galactosidic linkages of the structure shown in Figure 17.17.6 was found in the culture supernatant of *D. pneumoniae*. This enzyme was named endo-β-galactosidase DII to differentiate it from the previously isolated blood-group type A and B-specific endo-β-

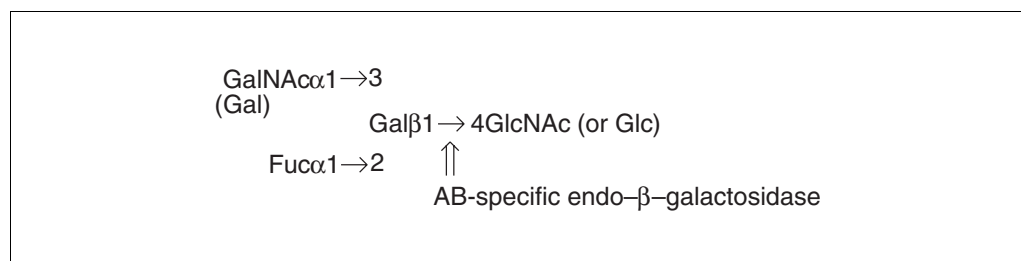


Figure 17.17.5 Structure hydrolyzed by blood-group AB-specific endo-β-glycosylase from *D. pneumoniae*.

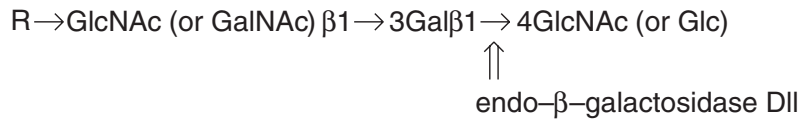


Figure 17.17.6 Structure hydrolyzed by endo- β -galactosidase DII from *D. pneumoniae*. R = H, Gal β 1 \rightarrow 4, or NeuNAc α 2 \rightarrow 3/6Gal β 1 \rightarrow 4.

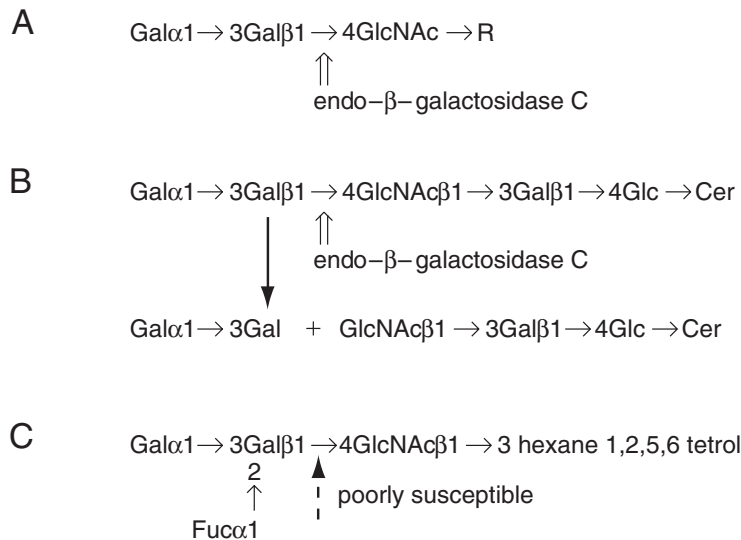


Figure 17.17.7 Hydrolysis by endo- β -glycosidase C from *C. perfringens*. **(A)** Enzyme hydrolyzes β -galactosidic linkages of glycoproteins and glycolipids. **(B)** Enzyme hydrolyzes the glycolipid pentaglycosylceramide. **(C)** Enzyme does not hydrolyze blood-group B oligosaccharide.

galactosidase from the same bacterium. The specificity of endo- β -galactosidase DII is similar to that of *E. freundii* endo- β -galactosidase, as both recognize the N-acetylglucosaminyl structure and hydrolyze internal β -galactosyl linkages. However, endo- β -galactosidase DII differs from the *E. freundii* enzyme in that it does not degrade keratan sulfate. Endo- β -galactosidase DII hydrolyzes glycolipids and polyglucosaminoglycans endoglycosidically, but from the nonreducing-terminal end.

ENDO- β -GALACTOSIDASE C FROM CLOSTRIDIUM PERFRINGENS (Fushuku et al., 1987)

Endo- β -galactosidase C hydrolyzes β -galactosidic linkages of glycoproteins and glycolipids that have the terminal structure Gal α 1 \rightarrow 3Gal β 1 \rightarrow 4GlcNAc as shown in Figure 17.17.7A. This terminal Gal α 1 \rightarrow 3Gal β 1 \rightarrow structure is present in mice, rabbits, cattle, and new world monkeys, but not in humans, apes, or old world monkeys.

The carbohydrate moiety of a glycoprotein that has a Gal α 1 \rightarrow 3Gal β 1 \rightarrow 4GlcNAc \rightarrow terminal structure can be efficiently hydrolyzed by endo- β -galactosidase C, which releases the disaccharide Gal α 1 \rightarrow 3Gal. Glycolipids having this terminal structure can also be hydrolyzed. For example, pentaglycosylceramide isolated from rabbit erythrocyte membranes is hydrolyzed as shown in Figure 17.17.7B.

The enzyme, however, does not hydrolyze blood-group B oligosaccharide (as shown in Figure 17.17.7C) although it hydrolyzes Gal α 1 \rightarrow 3Gal β 1 \rightarrow 4GlcNAc β 1 \rightarrow 3 hexane 1,2,5,6 tetrol. Endo- β -galactosidase C does not hydrolyze polyglucosaminoglycans from human erythrocytes of blood-group B. Thus, the enzyme is different from the endo- β -galactosidase from *D. pneumoniae*, which acts on blood group A and B antigens, as well as the endo- β -galactosidases from *E. freundii*, *B. fragilis*, and *F. keratolyticus*, which act on polyglucosaminoglycans. In addition, endo- β -

galactosidase C does not hydrolyze keratan sulfate.

KERATANASE FROM *PSEUDOMONAS* SP. (Nakazawa and Suzuki, 1975)

Source

Keratanase (or keratan sulfate-specific endo- β -galactosidase) has been purified from *Pseudomonas* sp. cell extract and is commercially available from Miles (APPENDIX 4).

Properties

The purified enzyme has been stored in ice. The pH optimum is between 7.2 and 7.4 using corneal keratan sulfate as substrate. At a 1 mM concentration, Ca^{2+} , Mg^{2+} , and Mn^{2+} have no effect on the enzymatic activity. Acetate and maleate elicit a slight increase in activity (~15%) at concentrations of 0.05 M.

Substrates and Specificity

Keratanase from *Pseudomonas* sp. hydrolyzes keratan sulfate but not chondroitin sulfates, hyaluronic acid, dermatan sulfate, or heparin. In contrast to *E. freundii* endo- β -galactosidase, this enzyme does not hydrolyze lacto-N-tetraose, lacto-series glycolipids, or erythrocyte polylectosaminoglycans—thus *Pseudomonas* endo- β -galactosidase is specific for keratan sulfate. Two sulfated oligosaccharides—

6-sulfo-GlcNAc β 1 \rightarrow 3Gal and
6-sulfo-GlcNAc β 1 \rightarrow 3-6-sulfo-GlcNAc β 1 \rightarrow
4-6-sulfo-GlcNAc β 1 \rightarrow 3Gal—

have been identified in the digestion products from keratan sulfates produced by this enzyme. However, the nonsulfated disaccharide GlcNAc β 1 \rightarrow 3Gal is not produced.

Applications

This enzyme has been quite useful in distinguishing keratan sulfates or sulfated polylectosaminoglycans from nonsulfated polylectosaminoglycans (Nakazawa et al., 1984).

ENDO- β -GALACTOSIDASE SPECIFIC TO THE REGION LINKING CHONDROITIN SULFATE AND THE CORE PROTEIN (Takagaki et al., 1990a,b)

Endo- β -galactosidase that degrades the linkage between a chondroitin sulfate and core protein has been reported in rabbit liver and in the midgut gland of the mollusk *Patinopecten* sp. Chondroitin sulfate-containing proteogly-

cans were digested with endo- β -xylosidase from the midgut gland of *Patinopecten* sp., and the released chondroitin sulfate (with an R-Gal β 1 \rightarrow 3Gal β 1 \rightarrow 3Gal β 1 \rightarrow 4Xyl terminus) was conjugated with pyridylamine (PA) by reductive amination. This chondroitin sulfate-PA was incubated with a crude enzyme solution from the midgut gland of *Patinopecten* sp. HPLC analysis of the PA-oligosaccharide product identified Gal β 1 \rightarrow 4Xyl-PA, suggesting strongly that an endo- β -galactosidase had specifically hydrolyzed the Gal β 1 \rightarrow 3Gal linkage in the linkage region of chondroitin sulfate.

Literature Cited

- Fukuda, M.N. 1981. Purification and characterization of endo- β -galactosidase from *Escherichia freundii* induced by hog gastric mucin. *J. Biol. Chem.* 256:3900-3905.
- Fukuda, M.N. 1985. Isolation and characterization of a new endo- β -galactosidase from *Diplococcus pneumoniae*. *Biochemistry* 24:2154-2163.
- Fukuda, M.N. and Matsumura, G. 1975. Endo- β -galactosidase of *Escherichia freundii*. Hydrolysis of pig colonic mucin and milk oligosaccharides by endoglycosidic action. *Biochem. Biophys. Res. Commun.* 64:465-471.
- Fukuda, M.N. and Matsumura, G. 1976. Endo- β -galactosidase of *Escherichia freundii*. Purification and endoglycosidic action on keratansulfates, oligosaccharides and blood group active glycoprotein. *J. Biol. Chem.* 251:6218-6225.
- Fukuda, M.N., Fukuda, M., and Hakomori, S. 1979. Cell surface modification by endo- β -galactosidase, change of blood group activities and release of oligosaccharides from glycoproteins and glycosphingolipids of human erythrocytes. *J. Biol. Chem.* 254:5458-5465.
- Fukuda, M., Dell, A., Oates, J.E., and Fukuda, M.N. 1984. Structure of branched lactosaminoglycan, the carbohydrate moiety of band 3 isolated from adult human erythrocytes. *J. Biol. Chem.* 259:8260-8273.
- Fushuku, N., Muramatsu, H., Uezono, M.M., and Muramatsu, T. 1987. A new endo- β -galactosidase releasing Gal α 1 \rightarrow 3Gal from carbohydrate moieties of glycoproteins and from a glycolipid. *J. Biol. Chem.* 262:10086-10092.
- Gahmberg, C.G. and Hakomori, S. 1973. External labeling of cell surface galactose and galactosamine in glycolipid and glycoprotein of human erythrocytes. *J. Biol. Chem.* 248:4311-4317.
- Hirano, S. and Meyer, K. 1971. Enzymatic degradation of corneal and cartilaginous keratosulfates. *Biochem. Biophys. Res. Commun.* 44:1371-1375.
- Kannagi, R., Nudelman, E., Levery, S.B., and Hakomori, S. 1982a. A series of human erythrocyte glycosphingolipids reacting to the monoclonal antibody directed to a developmentally regulated antigen, SSEA-1. *J. Biol. Chem.* 257:14865-14874.

- Kannagi, R., Fukuda, M.N., and Hakomori, S. 1982b. A new glycolipid antigen isolated from human erythrocyte membranes reacting with antibodies directed to globo-N-tetraosylceramide (globoside). *J. Biol. Chem.* 257:4438-4442.
- Kitamikado, M., Ueno, R., and Nakamura, T. 1970a. Enzymatic degradation of Whale cartilage keratosulfate-I, isolation of keratosulfate-degrading bacterium. *Bull. Jap. Soc. Sci. Fish.* 36:592-596.
- Kitamikado, M., Ueno, R., and Nakamura, T. 1970b. Enzymatic degradation of Whale cartilage keratosulfate-II, identification of a keratosulfate-degrading bacterium. *Bull. Jap. Soc. Sci. Fish.* 36:1172-1174.
- Kitamikado, M. and Ueno, R. 1970c. Enzymatic degradation of Whale cartilage keratosulfate-III, purification of a bacterial keratosulfate-degrading enzyme. *Bull. Jap. Soc. Sci. Fish.* 36:1175-1180.
- Kitamikado, M. Ito, M., and Li, Y.-T. 1981. Isolation and characterization of a keratan sulfate-degrading endo- β -galactosidase from *Flavobacterium Keratolyticus*. *J. Biol. Chem.* 256:3906-3909.
- Muramatsu, T., Gachelin, G., Damonville, M., Delarbre, C., and Jacob, F. 1979. Cell surface carbohydrates of embryonal carcinoma cells: Polysaccharidic side chains of F9 antigens and of receptors to two lectins, FBP and PNA. *Cell* 18:183-191.
- Nakagawa, H., Yamada, T., Chien, J.-L., Gardas, A., Kitamikado, M., Li, S.-C., and Li, Y.-T. 1980. Isolation and characterization of an endo- β -galactosidase from a new strain of *Escherichia freundii*. *J. Biol. Chem.* 255:5955-5959.
- Nakazawa, K. and Suzuki, S. 1975. Purification of keratansulfate-endogalactosidase and its action on keratan sulfates of different origin. *J. Biol. Chem.* 250:912-917.
- Nakazawa, K., Hassell, J.R., Hascal, V.C., Lohmander, S., Newsome, D.A., and Krachmer, J. 1984. Defective processing of keratan sulfate in macular corneal dystrophy. *J. Biol. Chem.* 259:13751-13757.
- Paoletti, L.C., Kasper, D.L., Michon, F., DiFabio, J., Holme, K., Jennings, H.J., and Wessels, M.R. 1990. An oligosaccharide-tetanus toxoid conjugate vaccine against type III group B streptococcus. *J. Biol. Chem.* 265:18278-18283.
- Scudder, P., Uemura, K., Kolby, J., Fukuda, M.N., and Feizi, T. 1983. Isolation and characterization of an endo- β -galactosidase from *Bacterioides fragilis*. *Biochem. J.* 213:485-494.
- Scudder, P., Hanfland, P., Uemura, K., and Feizi, T. 1984. Endo- β -galactosidase of *Bacterioides fragilis* and *Escherichia freundii* hydrolyze linear but not branched oligosaccharide domains of glycolipids of the neolactoseries. *J. Biol. Chem.* 259:6586-6592.
- Scudder, P., Tang, P.W., Hounsell, E.F., Lawson, A.M., Mehmet, H., and Feizi, T. 1986. Isolation and characterization of sulfated oligosaccharides released from bovine corneal keratansulfate by the action of endo- β -galactosidase. *Eur. J. Biochem.* 157:365-373.
- Spooner, E., Fukuda, M., Klock, J.C., Oates, J.E., and Dell, A. 1984. Isolation and characterization of polyfucosylated lactosaminoglycan from human granulocytes. *J. Biol. Chem.* 259:4792-4801.
- Takagaki, K., Kon, A., Kawasaki, H., Nakamura, T., Tamura, S., and Endo, M. 1990a. Presence of an endo- β -galactosidase degrading the linkage region between the chondroitin sulfate chain and core peptide of proteoglycan. *Biochem. Biophys. Res. Commun.* 169:15-21.
- Takagaki, K., Kon, A., Kawasaki, H., Nakamura, T., Tamura, S., and Endo, M. 1990b. Isolation and characterization of Patnopecten mid-gut gland endo- β -xylosidase active on peptide chondroitin sulfate. *J. Biol. Chem.* 265:854-860.
- Takasaki, S. and Kobata, A. 1976. Purification and characterization of an endo- β -galactosidase produced by *Diplococcus pneumoniae*. *J. Biol. Chem.* 251:3603-3609.
- Tsay, G.C., Dawson, G., and Li, Y.-T. 1975. Structure of the glycopeptide storage material in Gm1 gangliosides. Sequence determination with specific endo- and exoglycosidases. *Biochim. Biophys. Acta* 385:305-311.

Contributed by Michiko N. Fukuda
La Jolla Cancer Research
La Jolla, California

ANALYSIS OF SACCHARIDES RELEASED FROM GLYCOCONJUGATES

Many of the more specific and sensitive methods for analyzing the structure of oligosaccharides require that the oligosaccharides be first released from the glycoconjugate to which they are attached and, if necessary, purified away from the protein moiety. The units presented in this section describe various ways in which such released oligosaccharides can be analyzed. It must be recognized that no matter how good such a method is, it is highly dependent upon the quality of the preceding steps—i.e., the specific and quantitative release of the oligosaccharide from the glycoconjugate and the use of an appropriate purification scheme that does not result in selective losses. The methods presented in this section include those designed for the analysis of monosaccharides (*UNIT 17.18*) and those designed for the study of oligosaccharides (or glycopeptides; *UNITS 17.19-17.23*). The first group primarily provide monosaccharide compositional information and are a good starting point for the analysis of a previously unexplored glycoconjugate. Methods in the latter group range from simple techniques for determining basic properties of an oligosaccharide, to complex and sophisticated methodologies requiring the use of expensive instruments.

There are of course many other well-established techniques and approaches to the structural analysis of oligosaccharides, some of which require specialized instrumentation (e.g., mass spectrometry and nuclear magnetic resonance). The choice of techniques for inclusion in this section was based partly on the criteria of ease of use and broad general utility to the average molecular biologist.

Analysis of Monosaccharides

UNIT 17.18

This unit presents methods for assaying sialic acids, reducing sugars, and hexosamines. The BCA assay (Basic Protocol 1) detects free reducing termini in sugars released from glycoconjugates by appropriate treatments. Assays employing Ehrlich reagent (DMAB) detect hexosamines and *N*-acetylhexosamines (Basic Protocol 2 and Alternate Protocol); Basic Protocol 2 also includes a method for hydrolyzing the glycosidic linkages of the hexosamines and a method for re-*N*-acetylation is described in Support Protocol 1). The TBA (Basic Protocol 3) and DMB assays (Basic Protocol 4) can be used to quantitate and fractionate free forms of many types of sialic acids; techniques for liberating the sialic acids from the parent glycoconjugates are provided in Support Protocols 2 and 3. (Another alternative, the ferric orcinol method presented in *UNIT 17.9*, can measure both free and glycosidically bound sialic acids irrespective of modifications, but with much reduced sensitivity.)

Note that the acid hydrolysis conditions used to release sialic acids are much milder than those used for other monosaccharides—indeed, sialic acids are destroyed under the harsher conditions. Therefore, a sialic acid-containing molecule needs to be subjected to at least two separate analyses—one for sialic acids and one for the other monosaccharides.

BCA ASSAY FOR REDUCING SUGARS WITH SPECTROPHOTOMETRIC DETECTION

BASIC PROTOCOL 1

This protocol presents the most sensitive method developed to date to measure free reducing termini in sugars. This type of assay is useful for determining the total amount of oligosaccharides released from N-linked or O-linked glycoconjugates by enzymatic treatment (*UNIT 17.13A*) or from glycolipids by endoglycoceramidase (*UNIT 17.17A*). It may also be used to determine the total sugars present after acid hydrolysis.

Preparation and Analysis of Glycoconjugates

17.18.1

Materials

Sample in water or aqueous buffer
Standard: e.g., glucosamine hydrochloride or *N*-acetyl-D-glucosamine (both available from Sigma)
Bovine serum albumin (BSA)
BCA reagent (reagent C; see recipe)
16 × 115-mm Pyrex test tubes
Clean glass marbles
Heating block, 80°C
Spectrophotometer and 1.0-ml glass cuvettes

1. In 16 × 115-mm Pyrex test tubes, prepare duplicates containing aliquots of sample and of 2 to 25 nmol of an appropriate standard in 300 µl water. Also prepare a blank consisting of 300 µl of water or the buffer in which sample is dissolved, and a control containing BSA at a concentration equal to the maximum protein concentration expected in the samples.

Avoid borate ions in the buffer; but if these cannot be avoided, make sure to adjust the pH to 10.6 to avoid interference.

*The expected reducing sugar for the type of sample under analysis should be used as standard. *N*-acetylglucosamine should be used for oligosaccharides released from *N*-linked glycoproteins; glucose for oligosaccharides released from gangliosides.*

2. Add 700 µl BCA reagent to each tube, vortex, and cover with glass marbles to prevent evaporation. Heat 30 min at 80°C.
3. Allow tubes to cool to room temperature. Transfer solutions to 1.0-ml glass cuvettes and measure the absorbance at 560 nm, adjusting against the values obtained for the blank.
4. Prepare a standard curve of absorbance vs. nmol reducing sugar.
5. Determine the amount of reducing sugar in each sample by reference to the standard curve.

Reagents and Solutions

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

BCA reagent

Reagent A: Dissolve 1.5 g disodium 2,2'-bicinchoninate (4,4'-dicarboxy-2,2'-biquinoline; FW 388.3; Sigma) in 1 liter water. Add 71.6 g anhydrous sodium carbonate (ACS reagent grade, FW 105.99; Sigma) while stirring. Dilute to 1.15 liters with water. Store in a dark bottle at room temperature (stable several months).

Reagent B: Dissolve 3.5 g aspartic acid (FW 133.1; Sigma) and 5.0 g anhydrous sodium carbonate in 100 ml water, shaking to dissolve (this produces foam). In a separate container, dissolve 1.09 g copper sulfate (CuSO₄·5 H₂O, FW 249.6; ACS reagent grade) in 40 ml water. Mix the two solutions and dilute to 150 ml with water (a temporary fluffy blue precipitate may appear). Store at 4°C (stable several months).

Reagent C (BCA reagent): Mix 23 ml reagent A and 1 ml reagent B with 6 ml of 100% ethanol. The solution will have an intense blue color and strong fluorescence. Store in a brown bottle up to 2 weeks at 4°C (discard earlier if purple background color increases). Let stand 2 hr before use.

Commentary

Background Information

Most assays proposed for reduction of sugars make use of inorganic oxidants that accept electrons from sugar aldehyde groups. Aldoses are converted to aldonic acids (and higher oxidation products) and either the metal ions directly formed, metal oxides, or the complexation products of the latter with chromogenic chelators are quantitated. The appropriate oxidation/reduction potential for this reaction depends on the pH and complexation constants.

The BCA reagent forms colored complexes with copper(I); it is possible to couple this reaction to the reduction from copper(II) to copper(I) produced by reducing sugars in alkaline solution.

The most commonly used reduction assay for sugars is the one developed by Park and Johnson (1949) as a modification of the colorimetric method of Folin and Malmros (1929). This method, based on the production of Prussian blue following the reduction of ferricyanide ions in alkaline solution, has a detection limit of ~1 µg (5 nmol) reducing sugar.

Mopper and Gindler (1973) later proposed a new colorimetric reagent for automatic sugar chromatography that was sensitive to ~10⁻¹⁰ mol of sugar in the ethanol/water mobile phase. The dye is based on the formation of a deep-blue complex with Cu(I) ions and 2,2'-bichinchoninic acid in alkaline solution. The reaction is first order over a wide range. However, the sensitivity of the reaction is lower and the linearity range is reduced in water solution. Moreover, when borate ions are present, the Mopper-Gindler reagent produces only a faint color. Sinner and Puls (1978) achieved increased color formation by adjusting the reagent/borate buffer mixture, when used, to the optimal pH of the reaction (10.6), and increased the linear detection range by doubling the BCA concentration. Using these conditions, 55 pmol can be detected in the effluent from an HPLC column with an on-line UV spectrophotometer. McFeeters (1980) adapted the method to manual sugar determinations using ethylene glycol instead of ethanol to increase color formation. The conditions used for on-line detection can also be applied to manual analysis provided special care is taken to compensate for the volatility of ethanol.

Waffenschmidt and Jaenicke (1987) have suggested that for the manual spectrophotomet-

ric quantitation of sugar samples, the sensitivity of the reaction could be improved by replacing the aspartic acid in the BCA reagent with serine.

Critical Parameters

Using manual spectrophotometry, reliable detection is achieved with <3 nmol of mono-, oligo-, or polysaccharides. Buffering salts, sulfuric acid, sodium hydroxide, hydrochloric acid, and acetic acid do not interfere. It is also possible to work in the presence of borate ions if the pH is adjusted. The expected reducing sugar for the type of sample under analysis is used as a standard. However, linear responses may not be obtained with increasing size of the oligosaccharide (Doner and Irwin, 1992). Proteins cause a low level of interference of ~50:1 (w/w) protein/sugar.

Anticipated Results

With this protocol it is possible to detect with confidence 3 nmol of reducing sugar (mono-, oligo- or polysaccharide) in aqueous solution.

Time Considerations

The assay can be completed in ~1 hr.

Literature Cited

- Doner, L.W. and Irwin, P.L. 1992. Assay of reducing end-groups in oligosaccharide homologues with 2,2'-bichinchoninate. *Anal. Biochem.* 202:50-53.
- Folin, O. and Malmros, H. 1929. An improved form of Folin's micro method for blood sugar determinations. *J. Biol. Chem.* 83:115.
- McFeeters, R.F. 1980. A manual method for reducing sugar determinations with 2,2'-bichinchoninate reagent. *Anal. Biochem.* 103:302-306.
- Mopper, K. and Gindler, E.M. 1973. A new noncorrosive dye reagent for automatic sugar chromatography. *Anal. Biochem.* 56:440-442.
- Park, J.T. and Johnson, M.J. 1949. A submicrodetermination of glucose. *J. Biol. Chem.* 181:149-151.
- Sinner, M. and Puls, J. 1978. Noncorrosive dye reagent for detection of reducing sugars in borate complex ion-exchange chromatography. *J. Chromatogr.* 156:197-204.
- Waffenschmidt, S. and Jaenicke, L. 1987. Assay of reducing sugars in the nanomole range with 2,2'-bichinchoninate. *Anal. Biochem.* 165:337-340.

Key Reference

Waffenschmidt and Jaenicke, 1987. See above.

Describes the modification of the original method for improved sensitivity that is used in this protocol.

ASSAYS FOR FREE HEXOSAMINES WITH SPECTROPHOTOMETRIC DETECTION

Free hexosamines can be quantitated spectrophotometrically by a series of procedures derived from the Elson-Morgan and Morgan-Elson reactions that can be used for both hexosamines and N-acetylated hexosamines. These assays require a hydrolysis step (see Basic Protocol 2) to cleave the glycosidic linkages of the hexosamines. This hydrolysis also produces de-N-acetylation. Thus, a re-N-acetylation step is required (see Support Protocol 1) for the Morgan-Elson assay.

Acetylacetone/DMAB (Elson-Morgan) Assay for Free Hexosamines

Materials

Hexosamine-containing sample ($\geq 6 \mu\text{g}$ hexosamine) in aqueous solution
1 mM hexosamine standard stock solution(s): *N*-acetylglucosamine and/or
N-acetylgalactosamine (see recipe)
4 M HCl (diluted from ACS reagent-grade concentrated HCl)
Acetylacetone reagent (see recipe)
95% ethanol
Ehrlich reagent (DMAB)/HCl (see recipe)
13 \times 100-mm Pyrex test tubes with Teflon-lined screw caps
Nitrogen or vacuum evaporation system (Speedvac or shaker-evaporator)
Heating block or oven
Boiling water bath
Spectrophotometer and 1.0-ml glass cuvettes

Liberate hexosamines from glycosidic linkages

1. Dry triplicate aliquots of the sample (containing $\geq 6 \mu\text{g}$ hexosamine) in 13 \times 100-mm Pyrex test tubes with Teflon-lined screw caps.
2. From hexosamine standard stock solution(s), prepare triplicate standards containing 10, 20, 30, 40, 50, and 60 μg *N*-acetylglucosamine and/or *N*-acetylgalactosamine in water. Prepare a blank containing the same total volume of water.
3. Dissolve samples, standards, and blank in 4 M HCl (0.4 ml/mg dry sample). Flush with nitrogen and cap tubes. Heat 4 hr at 100°C.

Because the stability of this type of linkage depends on the structure of the glycoconjugate, optimal conditions for hydrolysis should be determined for each particular case (see UNIT 17.16). When using HCl, concentrations from 2 to 6 M and hydrolysis times from 2 to 10 hr should be tried.

4. Evaporate the HCl in a vacuum evaporation system or under nitrogen. Place the tubes overnight in a desiccator containing sodium hydroxide pellets under vacuum.

Quantitate hexosamines by condensation with alkaline acetylacetone and DMAB

5. Dissolve the hydrolysates in 250 μl water.
6. Add 500 μl acetylacetone reagent to each sample, standard, and blank. Vortex well and cap. Heat 10 min in a boiling water bath.
7. Cool to room temperature. Add 2.5 ml of 95% ethanol and vortex. Incubate 5 min at $75 \pm 2^\circ\text{C}$.
8. Add 500 μl DMAB/HCl and heat 30 min at $75 \pm 2^\circ\text{C}$.
9. Cool tubes to room temperature and add 2.5 ml of 95% ethanol. Let stand 30 min.

- Carefully transfer the solutions to 1.0-ml glass cuvettes and measure the absorbance at 520 nm, adjusting against the values obtained for the blank.
- Prepare a standard curve of absorbance vs. nmol hexosamine using the average value for each concentration.
- Determine the amount of hexosamine in the sample by reference to the standard curve.

Acidic DMAB (Morgan-Elson) Assay for Free *N*-Acetylhexosamines

**ALTERNATE
PROTOCOL**

Additional Materials (also see *Basic Protocol 2*)

20 mM potassium tetraborate
Ehrlich reagent (DMAB)/HCl/acetic acid (see recipe)

- Liberate hexosamines from their glycosidic linkages (see *Basic Protocol 2*, steps 1 to 4). Re-*N*-acetylate the samples (see *Support Protocol 1*).
- Dissolve each sample in 250 μ l water.
- Add 50 μ l of 20 mM potassium tetraborate to each sample, standard, and blank. Vortex well and cap. Heat 3 min in a boiling water bath.
- Cool rapidly to room temperature in a water bath.
- Add 1.5 ml DMAB/HCl/acetic acid, washing down any condensate that has formed, and vortex. Incubate 20 min at 37°C.
- Cool to room temperature. Carefully transfer the solutions to 1.0-ml glass cuvettes and measure the absorbance at 585 nm, adjusting against the values obtained for the blank.
- Prepare a standard curve of absorbance vs. nmol of hexosamine using the average value for each concentration.
- Determine the amount of hexosamine in the sample by reference to the standard curve.

RE-N-ACETYLATION OF FREE HEXOSAMINES

**SUPPORT
PROTOCOL 1**

Hexosamines that have been released from their glycosidic linkages are re-*N*-acetylated prior to detection using potassium tetraborate and DMAB/HCl/acetic acid (see *Alternate Protocol*).

Additional Materials (also see *Basic Protocol 2*)

1.5% (v/v) acetic anhydride (99%, Aldrich) in acetone (HPLC grade, Fisher)
Aqueous sample containing hexosamines liberated from glycosidic linkages by acid hydrolysis, and glucosamine and galactosamine standards treated in the same fashion (see *Basic Protocol 2*, step 2)

- Add 100 μ l of 1.5% acetic anhydride in acetone to 800 μ l of the aqueous solution of the samples containing the liberated hexosamines obtained by acid hydrolysis.
- Repeat the procedure for aqueous solutions of standard glucosamine and galactosamine hydrochlorides.
- Incubate samples and standards 5 min at room temperature.
- Dry under vacuum or with a nitrogen stream.

**Preparation and
Analysis of
Glycoconjugates**

17.18.5

Reagents and Solutions

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Acetylacetone reagent

Prepare 48 ml of 0.625 M sodium carbonate (FW 106; anhydrous, ACS reagent grade) in water. Add 2 ml acetylacetone (2,4-pentanedione, FW 100.1; Sigma). Prepare just prior to use.

Ehrlich reagent (DMAB)/HCl/acetic acid

Add 1.5 ml water to 11 ml concentrated HCl. Add 87.5 ml glacial acetic acid (FW 60.05; 99.99%; Aldrich). Dissolve 10 g *p*-dimethylaminobenzaldehyde (DMAB)—4-(*N,N*-dimethylamino)benzaldehyde, DMAB; crystalline, ACS reagent grade, FW 149.2—in the HCl/acetic acid mixture. Store in a dark bottle at room temperature (stable several weeks). Dilute 10 ml of mixture to 100 ml with glacial acetic acid immediately before using.

Ehrlich reagent (DMAB)/HCl

Weigh 1.6 g *p*-dimethylaminobenzaldehyde (DMAB)—4-(*N,N*-dimethylamino)benzaldehyde, DMAB; crystalline, ACS reagent grade, FW 149.2. Dissolve in 30 ml concentrated HCl. Store in a dark bottle at room temperature (stable several weeks).

Hexosamine standard stock solutions, 1 mM

Glucosamine: Dissolve 9 mg D-glucosamine hydrochloride (2-deoxy-D-glucose hydrochloride, FW 179.2) or 11 mg of *N*-acetyl-D-glucosamine (2-acetamido-2-deoxy-D-glucose, FW 221.2) in water. Dilute to 50 ml with more water. Store in aliquots at -20°C and thaw as needed.

Galactosamine: Dissolve 9 g *N*-acetyl-D-galactosamine hydrochloride (2-deoxy-D-galactose hydrochloride, FW 179.2) or 11 g *N*-acetyl-D-galactosamine (2-acetamido-2-deoxy-D-galactose, FW 221.2) in water. Dilute to 50 ml with more water. Store in aliquots at -20°C and thaw as needed.

Hexosamines may be obtained from Sigma.

Commentary

Background Information

The hexosamines glucosamine and galactosamine are the only amino sugars detected to date as structural components of glycoproteins and glycolipids. They generally occur in *N*-acetylated form. Provided they are quantitatively released from their glycosidic linkages, hexosamines can be quantitated colorimetrically. The primary problem, however, is achieving this quantitative liberation, because amino sugars are extremely stable under strongly acidic conditions (see UNIT 17.16).

Most methods reported for colorimetric quantitation of glucosamine and galactosamine, *N*-acetylated or not, are based on the procedures described by Elson and Morgan (1933) and Morgan and Elson (1934).

The basis of the Morgan-Elson reaction for quantitating *N*-acetylglucosamine (2-acetami-

do-2-deoxy-D-glucose) and *N*-acetylgalactosamine (2-acetamido-2-deoxy-D-galactose) is the red color produced by various *N*-acetylhexosamine derivatives on treatment with acidic Ehrlich reagent (*p*-dimethylaminobenzaldehyde, or DMAB) following pretreatment with alkali. Detection of free amino sugars using this reaction can be achieved by including a re-*N*-acetylation step in the procedure (Roseman and Ludowieg, 1954; Levvy and McAllan, 1959). A more sensitive method was developed based on the observation that the presence of borate buffer increases the intensity of color in the reaction (Reissig et al., 1955). Different structures have been proposed for the chromogens formed during the alkaline treatment of *N*-acetylhexosamines in the Morgan-Elson reaction (Foster and Horton, 1959).

The Elson-Morgan reaction for quantitating glucosamine (2-amino-2-deoxy-D-glucose) and galactosamine (2-amino-2-deoxy-D-galactose) involves the production of a red color upon successive treatment with alkaline acetylacetone (2,4-pentanedione) and the Ehrlich reagent. Both hexosamines produce similar color intensities (Rondle and Morgan, 1955). Several red compounds with different ultraviolet absorption maxima are produced in the Elson-Morgan reaction. The proportion of these products depends on the conditions used for the reactions with acetylacetone and the Ehrlich reagent.

Different methods for quantitating both hexosamines individually (when both are present), based on their different reactivities and the differential effects of borate ions in the reactions, have also been developed (Roseman and Daffner, 1956; Tracey, 1955; Ludowieg and Benmaman, 1967).

The protocols described here are those suggested by Chaplin (1986)—i.e., direct reaction with Ehrlich reagent (Alternate Protocol)—and by Montreuil (1986)—i.e., condensation with acetylacetone/Ehrlich reagent (Basic Protocol 2).

Critical Parameters

The presence of neutral sugars and amino acids interferes with the Elson-Morgan assay (Basic Protocol 2), because these react to produce orange chromophores. Therefore, when working with crude hydrolysis mixtures, interfering compounds should be removed before the assay is performed. The reactivity of all 2-amino-2-deoxy sugars in this assay is equivalent. Very careful control of the reaction conditions is required to obtain reproducible results, however, as different chromophores may play the significant role and some chromophores are volatile (e.g., 2-methylpyrrole). An *O*-methyl substitution at position 4 or 6 has little effect in the reaction, but 6-*O*-phosphate and 3-*O*-methyl both cause a shift in the absorption maximum, and the latter also increases the color intensity.

In the Morgan-Elson reaction (Alternate Protocol), furanosidic intermediates are produced in the course of chromogen formation. Therefore, some alkali-stable substituents (e.g., 4-*O*-methyl groups) can prevent color production, while other groups that are easily eliminated (e.g., 3-*O*-acetyl groups) can cause increased chromogen formation. *N*-acetylgalactosamine yields only 23% of the color produced by *N*-acetylglucosamine. The presence of bo-

rate buffer increases the intensity of color in the reaction. With borate, the interference of hexoses and amino acids is less pronounced. The use of different alkaline buffers and heating times during prochromogen formation produces differences in color yield between the different 2-acetamido sugars.

When the Alternate Protocol is used, the absorption curve of hexosamines shows a maximum at 520 nm but the extinction coefficients for glucosamine and galactosamine are different. Sialic acids react to produce a blue coloration, and muramic acids and *N*-acetylhexosamines interfere. Sodium tetraborate and sodium chloride decrease the intensity of the coloration.

Anticipated Results

Any of the protocols described here will give a fair estimation of the total content of hexosamines in a given glycoconjugate, provided hexosamine release is complete.

Time Considerations

The most time-consuming step of this procedure is hydrolysis of the hexosamines, particularly when conditions for complete release need to be determined. The colorimetric reactions take 1 hr (Alternate Protocol) or 2 hr (Basic Protocol 2), respectively. If re-*N*-acetylation is required, an additional hour will be necessary.

Literature Cited

- Chaplin, M.F. 1986. Monosaccharides. *In* Carbohydrate Analysis: A Practical Approach (M.F. Chaplin and J.F. Kennedy, eds.) pp. 1-36. IRL Press, Oxford.
- Elson, L.A. and Morgan, W.T.J. 1933. A colorimetric method for the determination of glucosamine and chondrosamine. *Biochem. J.* 27:1824-1828.
- Foster, A.B. and Horton, D.H. 1959. Aspects of the chemistry of the amino sugars. *Adv. Carbohydr. Chem.* 14:213-281.
- Levy, G.A. and McAllan, A. 1959. The *N*-acetylation and estimation of hexosamines. *Biochem. J.* 73:127-132.
- Ludowieg, J. and Benmaman, J.D. 1967. Colorimetric differentiation of hexosamines. *Anal. Biochem.* 19:80-88.
- Montreuil, J., Bouquelet, S., Debray, H., Fournet, B., Spik, G., and Strecker, G. 1986. Glycoproteins. *In* Carbohydrate Analysis: A Practical Approach. (M.F. Chaplin and J.F. Kennedy, eds.) pp. 143-204. IRL Press, Oxford.
- Morgan, W.T.J. and Elson, L.A. 1934. A colorimetric method for the determination of *N*-acetylglucosamine and *N*-acetylchondrosamine. *Biochem. J.* 27:988-995.

- Reissig, J.L., Strominger, J.L., and Leloir, L. 1955. A modified colorimetric method for the estimation of *N*-acetyl amino sugars. *J. Biol. Chem.* 217:959-966.
- Rondle, C.J.M. and Morgan, W.T.J. 1955. The determination of glucosamine and galactosamine. *Biochem. J.* 61:586-589.
- Roseman, S. and Ludowieg, J. 1954. *N*-acetylation of the hexosamines. *J. Am. Chem. Soc.* 76:301-302.
- Roseman, S. and Daffner, I. 1956. Colorimetric method for determination of glucosamine and galactosamine. *Anal. Chem.* 28:1743-1747.
- Tracey, M.V. 1955. A rapid colorimetric distinction between glucosamine and galactosamine. *Biochim. Biophys. Acta.* 17:159-160.

Key References

- Reissig et al., 1955. See above.
- Rondle and Morgan, 1955. See above.
- Roseman and Ludowieg, 1954. See above.
- Descriptions of the different modifications to the original Elson-Morgan and Morgan-Elson reactions.*

THIOBARBITURIC ACID ASSAY FOR SIALIC ACIDS WITH SPECTROPHOTOMETRIC OR HPLC DETECTION

Unconjugated sialic acids can be measured by oxidation with periodate followed by reaction with the chromogen 2'-thiobarbituric acid (TBA) to form a salmon-pink compound. This chromophore can be detected either by conventional spectrophotometry, with a sensitivity of 1 to 10 nmol, or by HPLC, with a sensitivity of 2 to 10 pmol. The latter approach requires more equipment, but provides substantially improved sensitivity and eliminates interference by some compounds. The most common types of sialic acid, *N*-acetyl- and *N*-glycolylneuraminic acid (Neu5Ac and Neu5Gc, respectively), can be accurately quantitated by this assay. Modified (e.g., *O*-acetylated) sialic acids give variably reduced color response. Glycosidically bound molecules must be released by treatment with base (see Support Protocol 2) followed by mild acid (see Support Protocol 3) or sialidase treatment (*UNIT 17.12*) before detection.

Materials

- Sialic acid-containing sample in solution
- Periodate reagent (see recipe)
- 1 mM *N*-acetylneuraminic acid (Neu5Ac; store at -20°C ; stable indefinitely)
- 1 mM 2-deoxyribose (optional; store at -20°C ; stable indefinitely)
- Arsenite reagent (see recipe)
- TBA reagent (see recipe)
- Cyclohexanone (for spectrophotometric detection)
- HPLC elution buffer (see recipe; for HPLC detection)
- Heating block, 100°C
- For spectrophotometric detection:
 - New 12 \times 75-mm glass tubes
 - Clean glass marbles
 - Tabletop centrifuge
 - Spectrophotometer and 1.0-ml glass cuvettes
- For HPLC detection:
 - 0.4 \times 25-cm C_{18} reversed-phase HPLC column
 - Equipment for HPLC (*UNIT 10.2*)
 - HPLC UV detector with flow detector cell of volume $<20\ \mu\text{l}$, and integrating chart recorder if available

Prepare sample, standards, and controls

1. Prepare two identical 50- μl samples of free or released sialic acid in sample buffer. Add 25 μl periodate reagent to the sample without touching the sides of the tube. Vortex briefly at low speed to mix.

Free sialic acid samples may be derived from biochemical samples (e.g., purified proteins, total serum glycoproteins, or lipid extracts) either by sialidase digestion (UNIT 17.12) or by acid treatment (see Support Protocol 2 or UNIT 17.16).

Use new 12 × 75-mm glass tubes if later detection will be by spectrophotometric assay or 1.5-ml polypropylene microcentrifuge tubes for later HPLC assay.

2. Prepare a series of dilutions of Neu5Ac in the same buffer as that in which the sample is dissolved. Use a range of 1, 4, 10, and 15 nmol for the spectrophotometric assay and 2, 5, 20, 100, and 500 pmol for the HPLC assay. Also prepare two blank samples containing the same total quantity of buffer only. Treat standards and blanks with periodate as described in step 1.

Optionally, samples with 5 to 10 nmol 2-deoxyribose may also be prepared to be used as standards for interfering compounds in the sample.

Oxidize and react with TBA

3. Incubate all tubes 20 min at room temperature.
4. Slowly add 250 μ l arsenite reagent along the rim of each tube so it flows down the sides. Watch for a brown color to begin to develop in the sample and standard tubes (up to 2 min), then vortex vigorously.

The brown color results from the reduction of the periodate, which forms iodine. If any periodate remains unquenched, the TBA chromophore will not be produced. Occasionally the brown color will not appear (Troubleshooting).

5. Incubate tubes 5 min at room temperature.

At this point the tubes can be stored overnight at 4°C if desired.

6. Add 1 ml TBA reagent to each tube, mix well, and heat 15 min at 100°C. Cover each glass tube with a clean glass marble to minimize evaporation, or pierce microcentrifuge tube caps with a needle after closing. A pink color will develop in samples containing >0.5 nmol sialic acid. Chill 2 to 3 min on ice and keep on ice until chromophore detection is performed (beginning at step 7a or 7b).

Volumes of sample, periodate, and arsenite can be changed as long as the proportions remain the same. However, the volume of TBA can be decreased independently if necessary, to accommodate a small tube size (e.g., microcentrifuge tube).

Detect chromophore by spectrophotometric measurement

- 7a. Add 1 ml cyclohexanone to each tube. Vortex well to extract the pink chromophore into the upper (cyclohexanone) layer, then centrifuge 5 min at 500 to 2000 \times *g* (e.g., 1500 rpm in a tabletop centrifuge), room temperature, to separate the phases.

Depending on the size of the cuvettes, the volume of cyclohexanone can be reduced to improve sensitivity.

- 8a. Transfer an appropriate volume of the cyclohexanone layer (0.3 to 0.8 ml, depending on cuvette size) to a 1.0-ml glass cuvette. Read its absorbance at both 549 nm and 532 nm, adjusting against the values obtained for the blank.

The chromophore generated from sialic acids has an absorption maximum of 549 nm. Interfering substances, including 2-deoxyribose and peroxidation products of fatty acids, produce a chromophore with an absorption maximum of 532 nm.

- 9a. Calculate the ratios of A_{549}/A_{532} for the sialic acid standards and the sample. If the ratios match, this indicates that there is no interference. If the ratio is lower for the sample than the standard, perform the calculation indicated in steps 10a to 12a to compensate for interference.

- 10a. Calculate the value of $(0.9 \times A_{549} - 0.3 \times A_{532})$ for samples and standards.
- 11a. Plot a standard curve with these values and determine amount of sialic acid in the samples.
- 12a. Apply the calculations to the 2-deoxyribose sample to assure validity (this value should be close to zero after correction).

These correction coefficients are based on the original publications of Warren (1959) and Aminoff (1959) and generally work for most spectrophotometers. However, variations in slit width between instruments may decrease the correction coefficient accuracy. If unacceptable variance is found, correction coefficients should be determined directly as described in those original publications. The correction becomes less accurate if A_{532} is $>0.7 \times A_{549}$, and the result is practically meaningless if A_{532} is $>1.5 \times A_{549}$.

Detect chromophore by HPLC measurement

- 7b. Equilibrate an 0.4×25 -cm C_{18} reversed-phase HPLC column in HPLC elution buffer running at 0.7 to 1.0 ml/min. Connect column to a UV detector set at 549 nm. Connect the UV detector to an integrating data module, if available.

Faster flow rates, within the pressure limits of the column, will reduce the time required and will not adversely affect sensitivity. Columns 10 to 15 cm long may also be utilized.

All HPLC buffers must be filtered and degassed. A flow detector cell with a volume of $<20 \mu\text{l}$ is essential for maximal sensitivity.

- 8b. Keep the tubes on ice. Immediately prior to performing HPLC, microcentrifuge all tubes 2 to 4 min at $>5,000 \times g$, 4°C , to remove any precipitated reagents.

Keeping the samples chilled maximizes the precipitation of unused TBA, reducing the risk that it will precipitate in the HPLC apparatus.

- 9b. Inject aliquots of the derivatized Neu5Ac standards (and 2-deoxyribose standard, if included) and collect data using UV detector with flow recorder.

Chromophores resulting from these two sugars will elute at 5 to 10 min and 10 to 20 min, respectively, depending on the particular column and the methanol concentration in the elution buffer.

The concentration of methanol in the HPLC elution buffer should result in elution of the TBA chromophore at 6 to 10 min. If the chromophore elutes too early on the particular column in use, decrease the methanol concentration; likewise, increase methanol concentration if the peak elutes too late.

- 10b. Quantitate the signal from each standard, and blank by integration and construct a standard curve.

If a data module capable of integration is not available, the signal can be quantitated by measuring the peak height from a chart recorder.

- 11b. Inject 10 to $400 \mu\text{l}$ of each sample and collect data as in step 10b. Calculate the amount of sialic acid present based on the standard curve.

The amount of sample injected should be based on the amount of sialic acid believed to be present. Quantitation is not affected by the volume injected. Total volumes of sample and reagents can be decreased proportionately to further improve sensitivity by increasing the amount of chromophore present in the injected volume.

When setting up the column for the first time, inject sample of buffer blank or (derivatized) Neu5Ac (100 to 500 pmol) to verify elution times.

At the end of each day wash the column thoroughly with 50% acetonitrile or 50% methanol.

DE-O-ACETYLATION OF SIALIC ACIDS

O-acetylated sialic acids are detected with reduced sensitivity by the TBA assay. Furthermore, O-acetylation may reduce the release of sialic acids from glycoconjugates by either acid hydrolysis (see Support Protocol 3) or some sialidases (*UNIT 17.12*). Because sialic acids from some biological source may be as much as 50% to 100% O-acetylated, de-O-acetylation with base treatment is necessary for accurate quantitation. This fact can be useful, as the amount of color yield in the TBA assay before and after base treatment can be used to recognize the presence of O-acetylated sialic acids. Alternatively, specific O-acetylated sialic acid species can be identified by the DMB assay (see Basic Protocol 4).

Materials

Sialic acid-containing sample in solution
0.2 M NaOH (store at room temperature)
0.9 M H₂SO₄ (for acid release) or 0.2 M acetic acid (for enzymatic release)

1. Mix 20 μ l sialic acid-containing sample with 40 μ l of 0.1 M NaOH and incubate 30 min at 37°C.
- 2a. *For acid release:* Add 4.5 μ l of 0.9 M H₂SO₄ to bring to a final concentration of 0.1 M H₂SO₄ (in excess of Na₂SO₄) and proceed directly to BHT treatment (see Support Protocol 3, steps 2 to 4).

If the sample is free of potential lipid contaminants, the BHT step may be omitted and the sample utilized directly in the TBA or DMB assay.

- 2b. *For enzymatic release:* Neutralize with 20 μ l of 0.2 M acetic acid (check pH on a blank sample to verify it is between 5 and 7). Proceed to sialidase treatment (*UNIT 17.12*).

ACID RELEASE OF DE-O-ACETYLATED SIALIC ACIDS

Acid treatment (0.1 M H₂SO₄) can be used as an alternative to sialidase digestion (*UNIT 17.12*) to release glycosidically bound sialic acid molecules for subsequent quantitation by the TBA assay. O-acetylation is incompletely and variably destroyed by this type of acid treatment. On the other hand, O-acetylation can itself hinder the release of sialic acids. Thus, it is best to eliminate all O-acetyl esters prior to acid hydrolysis by base release (see Support Protocol 2). Alternatively, to preserve O-acetylation (for study with the DMB assay; see Basic Protocol 4), sialic acid may be released by mild acid treatment as described in *UNIT 17.16*.

Materials

Sialic acid-containing sample in solution, preferably de-O-acetylated (see Support Protocol 2)
0.2 M H₂SO₄
1% (w/v) BHT in ethanol (see recipe)
Clean glass marbles
Heating block, 80°C

1. Mix 25 μ l sample with 25 μ l of 0.2 M H₂SO₄ in the bottom of a clean glass tube.

If the sample has been de-O-acetylated in 40 μ l of 0.1 M NaOH and 4.5 μ l of 0.9 M H₂SO₄ added (in Support Protocol 2), this will result in a final excess acid concentration of 0.1 M and the same final volume.

2. Add 0.5 μ l of 1% BHT in ethanol to sample.
BHT prevents interference caused by peroxidation of hydroxylated fatty acids.
3. Cover the tube with a clean glass marble to minimize evaporation and heat 1 hr at 80°C.
4. Cool sample on ice. Proceed with TBA assay (see Basic Protocol 3, starting at step 1) without further adjustment.

Reagents and Solutions

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Arsenite reagent

Dissolve 10.0 g sodium arsenite and 7.1 g sodium sulfate in 70 ml of 0.1 M sulfuric acid. Add additional acid to bring volume to 100 ml. Store indefinitely at room temperature.

The final sodium arsenite concentration is 10% (w/v).

Butylated hydroxytoluene (BHT) in ethanol

Dissolve 100 mg BHT (2,6-di-*tert*-butyl-4-hydroxyanisole, FW 220.4; Sigma) in 10 ml of 100% ethanol. Store tightly capped at room temperature indefinitely.

HPLC elution buffer

Prepare a 2:3:5 (v/v/v) mixture of water/HPLC-grade methanol/0.2% (v/v) ortho-phosphoric acid (85%, ACS reagent grade). Filter mixture and degas.

Periodate reagent

128 mg sodium *meta*-periodate (Sigma; 0.2 M final)

1.3 ml H₂O

1.7 ml phosphoric acid (*ortho*-phosphoric acid, 85%, ACS reagent grade)

Mix well to dissolve

Store indefinitely at 4°C in a dark, foil-covered bottle

TBA reagent

Dissolve 14.2 g sodium sulfate in 170 ml water. Add 1.2 g 2-thiobarbituric acid (4,6-dihydroxypyrimidine-2-thiol, 98%; Sigma). Complete the dissolution of the TBA by adding 1 M NaOH dropwise to ~pH 9, then dilute with water to 200 ml. Store at room temperature.

The solution is stable almost indefinitely—even when a yellow precipitate appears, the reagent can still be used if care is taken to avoid the particulate material.

If HPLC detection is used, sodium sulfate may be omitted from both the TBA and arsenite solutions, as its function is to “salt out” the TBA chromophore into the cyclohexanone layer.

Commentary

Background Information

The original assay for sialic acids described independently by Warren (1959) and Aminoff (1959) is based upon the formation of an adduct between the periodate oxidation product of free unsubstituted sialic acids and 2'-thiobarbituric acid (TBA). The procedure has been modified

over the years (Schauer, 1978; Varki and Kornfeld 1980; Powell and Hart, 1986) to maximize its sensitivity and specificity and to minimize interference. The most useful modifications have been the following:

(1) Prior de-O-acetylation to eliminate the interference of *O*-acetyl esters with the perio-

date oxidation of sialic acid. An increase in color yield after de-O-acetylation indicates the presence of O-acetylation.

(2) An increase in the acid concentration in the arsenite reagent. This seems to avoid subsequent reaction failure by ensuring complete destruction of the periodate.

(3) Addition of BHT (butylated hydroxytoluene) during acid or enzyme release of sialic acids. This minimizes lipid peroxidation and formation of interfering malonaldehyde (Critical Parameters).

(4) Use of reversed-phase HPLC to remove interfering chromophores and improve sensitivity.

Critical Parameters

Complete oxidation of sialic acids in the presence of high concentrations of phosphoric acid and periodate yields β -formyl pyruvic acid. One molecule of this compound condenses with two molecules of TBA (under elevated temperature) to produce the brightly colored TBA adduct (Paerels & Schut, 1965). *O*-acetyl esters on the hydroxyl groups of sialic acids can cause varying degrees of interference, ranging from none (4-*O*-acetylation) or 50% (9-*O*-acetylation) to >90% (7-*O*-acetylation, multiple *O*-acetylations). Prior de-*O*-acetylation prevents this interference in the color yield. Less common substitutions (e.g., *O*-methyl groups) can also cause varying degrees of lower reactivity in comparison to Neu5Ac. For unknown reasons, the color yield also varies according to the substitution at the 5 position, in the approximate ratio of 1:0.8:2.0 for Neu5Ac, Neu5Gc, and 2-keto-3-deoxyoctonate (KDO—as well as, presumably, 2-keto-3-deoxynonosonic acid, KDN) respectively. For working purposes, Neu5Ac is initially used as a standard. Once the approximate amount of sialic acids in the sample is known, more detailed analysis of the types present can be performed (see Basic Protocol 4).

Compounds which, under the conditions of strong acid and periodate, yield β -formylpyruvic acid will yield a salmon-pink TBA chromophore of λ_{\max} 549 nm. These compounds include Neu5Ac, Neu5Gc, KDO, KDN, and the 3-keto-2-deoxyaldulosonic acids. Under these conditions other compounds yield malonaldehyde which, following condensation with TBA, yields an orange-yellow chromophore of λ_{\max} 532 nm. These compounds include 2-deoxyribose and the oxidation products of unsaturated lipids. Spontaneous peroxidation of lipids to malonaldehyde frequently occurs during sam-

ple handling and is accelerated by heat, light, acid pH, and some cations (lipids are frequently encountered, as sialic acids are often associated with membranes). Spontaneous peroxidation can be reduced or prevented by adding antioxidants such as BHT prior to storage or hydrolysis of samples. The presence of malonaldehyde-yielding substances can be detected and corrected for by measuring A_{549} and A_{532} (see Basic Protocol 1). Moreover, the two different chromophores arising from β -formylpyruvic acid and malonaldehyde are easily resolved by HPLC.

Troubleshooting

Failure of brown color to develop after addition of arsenite reagent may be caused by overly rapid addition of the arsenite or inadequate acid in the sample. This latter indicates that the periodate has not been fully consumed and will interfere with formation of the TBA chromophore. Alternatively, even if the brown color is seen, traces of periodate remaining on the walls of the reaction vessel (and therefore not exposed to the arsenite) may still also interfere. Certain buffers (particularly formate, which is frequently used during sialic acid purification) interfere when present in large quantities, giving a bright yellow color instead of the expected salmon pink. The predicted elution times of the chromophores from the C_{18} column are based on experience with several different columns from different suppliers. If trouble is encountered in detecting elution of the chromophores (from sialic acid or 2-deoxyribose), inject 10 to 50 ml of a visibly pink standard solution (e.g., 5 nmol). Two peaks should be seen eluting after the breakthrough peak (V_0), corresponding to the TBA adducts formed from β -formylpyruvic acid (eluting first) and malonaldehyde (eluting second). An even later-eluting reagent peak may also be seen, at an elution time 2 to 3 times that of the TBA chromophore, depending on the injected volume. If these peaks are not seen, or elute later than the indicated times, increase the concentration of methanol and allow the column to equilibrate at least 30 min between changes in methanol concentration.

Anticipated Results

For both the spectrophotometric and HPLC assays, the standard curve for A_{592} vs. nmol sialic acid should be linear over the suggested concentration ranges. Significant deviation from linearity indicates either improper dilution of the sialic acid standard or improper

concentration of one of the reagents. If modifications are made in the outlined protocol, the ratios of sample/periodate/arsenite must be kept constant, as these are optimized for concentrations of periodate, phosphoric acid, and arsenite. An excessive amount of any reagent can be as detrimental as an insufficient amount.

Time Considerations

Preparation of the standard curve and samples should take <3 hr. Determining the absorbance of each sample by HPLC will take ~15 to 20 min. Setup time for HPLC is indicated in *UNIT 10.12*.

Literature Cited

Aminoff, D. 1959. The determination of free sialic acid in the presence of the bound compound. *Virology* 7:355-357.

Paerels, G.B. and Schut, J. 1965. The mechanism of the periodate-thiobarbituric acid reaction of sialic acids. *Biochem. J.* 96:787-792.

Powell, L.D. and Hart, G.W. 1986. Quantitation of picomole levels of sialic acid by a HPLC-adaptation of the thiobarbituric acid assay. *Anal. Biochem.* 157:179-185.

Schauer, R. 1978. Characterization of sialic acids. *Methods Enzymol.* 50:64-89.

Varki, A. and Kornfeld, S. 1980. An autosomal dominant gene regulates the extent of 9-*O*-acetylation of murine erythrocyte sialic acids: A probable explanation for the variation in capacity to activate the alternate complement pathway. *J. Exp. Med.* 152:532-544.

Warren, L. 1959. The thiobarbituric acid assay of sialic acids. *J. Biol. Chem.* 234:1971-1975.

Key References

Aminoff, D. 1959. See above.

Warren, L. 1959. See above.

Original descriptions of the TBA assay; outline factors affecting its validity and reproducibility, including reaction times, reagent concentrations, and interfering compounds.

BASIC PROTOCOL 4

DMB ASSAY FOR SIALIC ACIDS WITH HPLC DETECTION

The method of choice for characterizing the individual components of a sialic acid mixture is the preparation of fluorescent derivatives. Mixtures of free sialic acids obtained using the protocols described in *UNIT 17.16* or *UNIT 17.12* can be reacted directly with 1,2-diamino-4,5-methylenedioxybenzene dihydrochloride (DMB) without the need for further purification. DMB derivatives of sialic acids are analyzed by HPLC on a reversed-phase column and detected fluorometrically.

Materials

N-acetylneuraminic acid (Neu5Ac, MW 309.3; Boehringer Mannheim)

Standard mixture: sialic acids prepared from bovine submaxillary mucin (BSM) using the procedure in *UNIT 17.16*, or mixture of 100 pmol each Neu5Ac and *N*-glycolylneuraminic acid (Neu5Gc, MW 325.3; Sigma)

Sample containing mixture of free sialic acids in solution (5 to 1000 pmol; *UNIT 17.16* or *UNIT 17.12*)

DMB reagent (see recipe)

Acetonitrile (HPLC grade)

50% methanol (HPLC grade) in Milli-Q-purified water

Milli-Q-purified water: deionized water passed through a five-stage Milli-Q Plus system (Millipore)

Helium (for HPLC apparatus)

Heating block, 50°C

HPLC apparatus (preferably ternary system) including pump and fluorescent detector (*UNIT 10.12*)

TSK gel ODS-120T column (250 mm × 4.6 mm i.d., 5 μm particle size; TosoHass)

Guard cartridge TSK gel ODS-120T (1.5 × 3.2 mm)

Filtering unit with 0.45-μm Nylon 66 membranes (Alltech)

Prepare fluorescent derivatives of sialic acids

1. Transfer sample containing 5 to 1000 pmol free sialic acids to a 1.5-ml microcentrifuge tube and dry in a Speedvac. Prepare 1 nmol of each standard (Neu5Ac and standard mixture) in the same fashion.
2. Dissolve the samples in 1 to 5 μ l water.
3. Add 30 μ l DMB reagent and wrap each tube in aluminum foil. Heat 2.5 hr in a heating block at 50°C.
4. Remove and place on ice.

Analysis should be done as soon as possible, but the DMB derivatives can be kept overnight at 4°C.

Analyze the fluorescent derivatives by HPLC with fluorometric detection

5. Turn on the fluorescent detector. Set emission wavelength to 448 nm, excitation wavelength to 373 nm, high voltage to 900, range to 500, and response time to 0.5 sec.
6. Fill the three HPLC bottles with (respectively) acetonitrile, 50% methanol, and filtered Milli-Q-purified water. Degas for 15 min and filter through 0.45- μ m membranes.
7. Run each solvent through the attached lines of the HPLC apparatus to eliminate any bubbles.
8. Set the pump control unit for an isocratic elution with a ratio of 9:7:84 (v/v/v) acetonitrile/methanol/water.

Because 50% methanol is used (to avoid the production of bubbles that occurs when mixing water and pure methanol), 14% of this solvent needs to be run to obtain 7% methanol (i.e., run 9 parts acetonitrile, 14 parts methanol/water mix, and 77 parts water). When a ternary HPLC system is not available, the eluant can be prepared ahead in the ratio stated above.

9. Begin running the solvent at 0.3 ml/min, then connect the TSK gel column to the Guard cartridge. Increase the flow 0.1 ml/min at a time, waiting for stable backpressure between steps, until 0.9 ml/min flow is attained. Wait until a stable baseline is obtained.

The following settings are adequate for most integrators: attenuation, 512; peak threshold, 10000.

10. Inject 100 pmol DMB-derivatized Neu5Ac standard. Check the sensitivity and adjust the integrator settings as required (the way this step is performed will depend on the particular integrator used). Repeat the run if necessary.
11. Inject DMB-derivatized standard mixture (100 pmol each Neu5Ac and Neu5Gc, or mixture containing O-substituted standards prepared from BSM, if available).
12. Inject the derivatized samples. Adjust the integrator settings as necessary, and repeat to obtain all peaks in scale. Calculate the retention time of each peak relative to Neu5Ac and compare with those listed in Table 17.18.1.

Keep the volume of DMB reaction mixture injected as low as possible to maintain column performance. In other words, it is better to use lower attenuation in the integrator and inject a smaller volume of reagent into the HPLC column. When the total amount of sialic acids in the sample is not known, first inject 10% of the total reaction mixture with a low attenuation setting, then if necessary inject more or change the settings.

Table 17.18.1 Relative Elution Times of DMB-Derivatized Sialic Acids from TSK-ODS Column

Sialic acid	t_{Neu5Ac} (min)
Neu5Ac	1.00
Neu5Gc	0.84
Neu5-7Ac ₂	1.06
Neu5-9Ac ₂	1.57
Neu4-5Ac ₂	1.68
Neu5-7(8)-9Ac ₃	1.90
Neu5-7-8-9Ac ₄	1.98
Neu5Gc9Ac	0.88
Neu4Ac5Gc	1.53

13. Determine the units of area per picomole for Neu5Ac and/or Neu5Gc. Assuming equal detector responses, calculate the number of picomoles represented by each peak produced by the samples.

This system can be used to analyze radioactive samples by routing the column effluent to a radioisotope flow detector or fraction collector. The appropriate ratio of scintillation fluid needs to be checked because of the quenching effect of acetonitrile.

14. Verify the presence of the O-acetylated species (see Table 17.18.1 for relative retention times) by collecting each peak, lyophilizing, subjecting to de-O-acetylation (see Support Protocol 2), and reanalyzing. Alternatively, another aliquot of the starting material can be submitted to de-O-acetylation prior to derivatization. In this case, all peaks corresponding to O-acetylated species will collapse into the peak of their parent molecule (Neu5Ac or Neu5Gc).
15. At the end of the day, wash the column with 50% methanol for 30 min at a flow of 0.5 ml/min.

When resolution becomes poor, the column can be cleaned by running it in reverse flow at 0.5 ml/min with a gradient from 10% to 100% acetonitrile in 0.05% trifluoroacetic acid over 1 hr. The column must then be extensively washed with water at 0.5 ml/min, turned around, and reequilibrated with 50% methanol.

Reagents and Solutions

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

DMB reagent

1.6 g 1,2-diamino-4,5-methylenedioxybenzene dihydrochloride (DMB, FW 225.1; Sigma)

80.4 μ l glacial acetic acid (HPLC grade; 1.4 M final)

52.8 μ l 2-mercaptoethanol (FW 78.13; Bio-Rad)

72 μ l 0.25 M sodium hydrosulfite

Store in a container covered with aluminum foil 1 week at 4°C

DMB reagent is light sensitive.

The 0.25 M sodium hydrosulfite solution is made by dissolving the compound (FW 174.1; ~80% pure, Sigma) in water at 43.5 mg/ml.

Commentary

Background Information

Sialic acids released by hydrolysis in 2 M acetic acid or by enzymatic treatment are converted with 1,2-diamino-4,5-methylenedioxybenzene (DMB, a fluorogenic reagent for α -keto acids) to highly fluorescent derivatives (see Fig. 17.18.1) without the occurrence of *O*-acetyl migration and de-*O*-acetylation.

The derivatives are separated using a reversed-phase TSK ODS-120T column eluted isocratically with 9:7:84 (v/v/v) acetonitrile/methanol/water at room temperature in ~25 min. Fluorescence of the eluate is monitored at an excitation wavelength of 373 nm and an emission wavelength of 448 nm. The average limits of detection are 1 to 2 pmol at a signal-to-noise ratio of 3.

Fluorogenic reagents may be used for detection of α -keto acids in very dilute solutions because of their quantitative conversion into the corresponding substituted quinoxalines

(Spiker and Towne, 1962; Hara et al., 1985a). Because sialic acids are α -keto acids, Hara et al. (1985b) tried several of these reagents and found that 1,2-diamino-4,5-dimethoxybenzene reacts with Neu5Ac and Neu5Gc under conditions different than those for other α -keto acids and forms fluorescent compounds. They developed a method for detecting small amounts of Neu5Ac and Neu5Gc in serum and urine using HPLC (Hara et al., 1986) and further extended the method to the analysis of glycoproteins and glycolipids (Hara et al., 1987a). Screening different diamino benzenes as fluorogenic reagents revealed that 1,2-diamino-4,5-methylenedioxybenzene (DMB) is the best reagent for sensitive and rapid detection of α -keto acids (Hara et al., 1987b).

The DMB method could not be applied to *O*-acetylated sialic acids because of de-*O*-acetylation and *O*-acetyl migration in the sulfuric acid used in the reaction. However, *O*-acety-

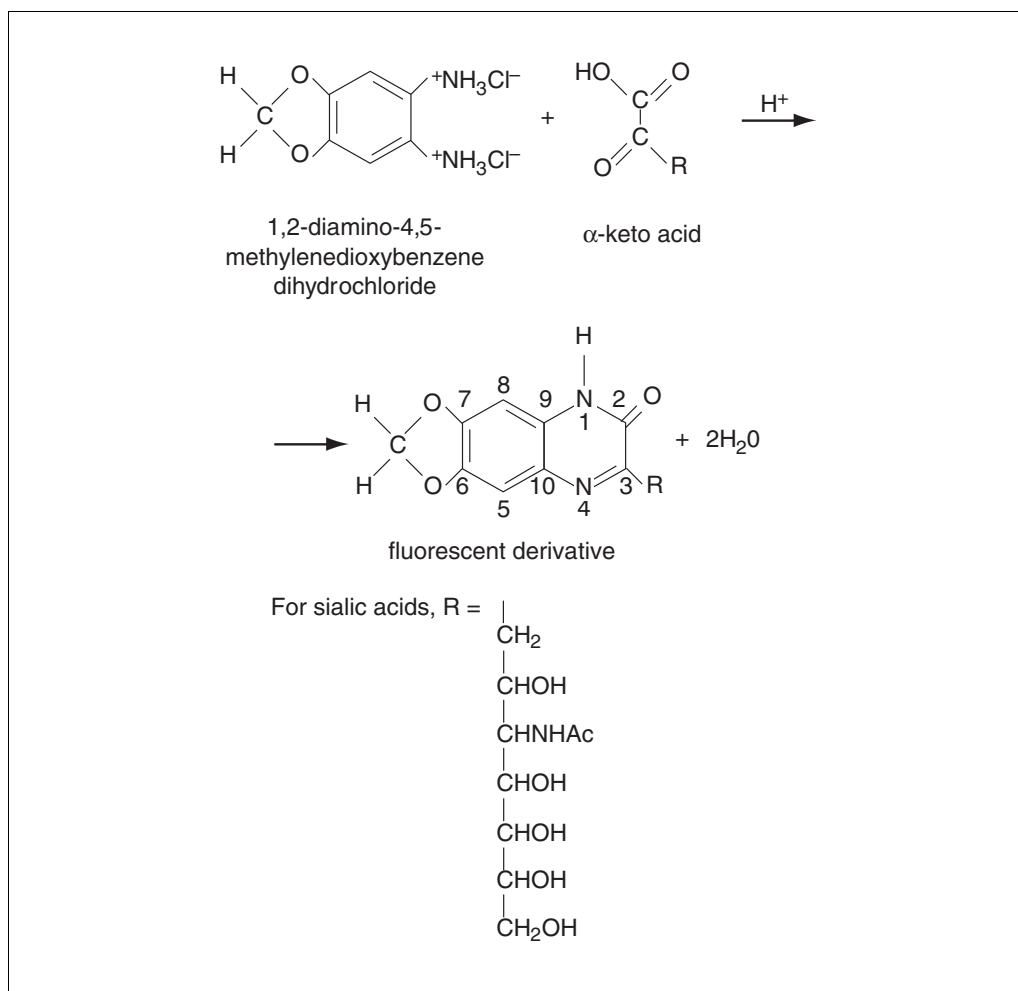


Figure 17.18.1 Reaction of sialic acids with 1,2-diamino-4,5-methylenedioxybenzene (DMB) to form fluorescent derivatives.

lated sialic acids react with DMB in acetic acid at lower temperatures to form fluorescent derivatives, completely avoiding those reactions (Hara et al., 1989). More recently, the value of this technique has been confirmed and extended to the analysis of several additional sialic acid molecules (Manzi et al., 1990).

Critical Parameters

The reaction of sialic acids with DMB is highly specific under controlled conditions (other α -keto acids that may be present in biological samples give very weak responses and/or very delayed retention times under the conditions recommended for the analysis). Most known naturally occurring substituted sialic acids can be derivatized and separated (see Fig. 17.18.2). Exceptions are the 2,3-dehydro sialic acids and glycosides that do not have an α -keto group available for the reaction. The presence of *O*-acetyl esters can be confirmed by repeating the analysis after base treatment (see Support Protocol 2).

Reliable detection and quantitation is possible with 2.5 pmol of any sialic acid. To maintain column performance it is advisable to keep the injection volume of DMB reaction mixture as low as possible. Evidence obtained using ra-

dioactively labeled sialic acids indicates that different degree of derivatization probably occur with molecules with different substitutions. Therefore, it is possible to accurately quantitate a given peak only when a corresponding standard is available, and quantitation of substituted species by comparison with Neu5Ac is only approximate.

Anticipated Results

This protocol will give a qualitative and semiquantitative composition analysis of sialic acids in a mixture release from glycoconjugates using picomole amounts (Critical Parameters has a more detailed discussion).

Time Considerations

When several samples need to be analyzed, because the DMB reaction takes 2.5 hr it is best to do it at the end of the day and keep the derivatized samples at 4°C overnight before analysis. The total time required for each HPLC run is ~30 min. It is advisable to run standards at the beginning and the end of each day to compare relative retention times and area/pmol. A total of 10 to 12 runs can be done in one 8-hr day.

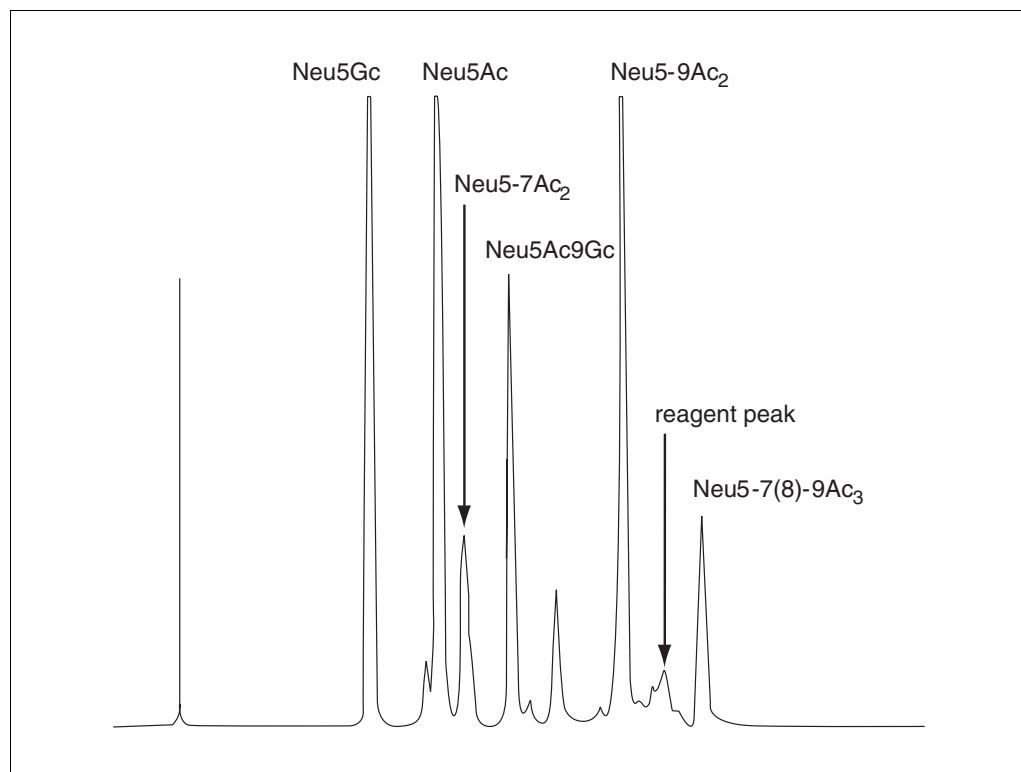


Figure 17.18.2 HPLC profile (fluorescence vs. time) of the DMB derivatives of total sialic acids released from BSM (bovine submaxillary mucin) with 2 M acetic acid and purified according to the protocol described in UNIT 17.16.

Literature Cited

- Hara, S., Yamaguchi, M., Nakamura, M., and Okhura, Y. 1985a. Fluorescent products of reaction between α -keto acids and 1,2-diamino-4,5-dimethoxybenzene. *Chem. Pharm. Bull.* 33:3493-3498.
- Hara, S., Takemori, Y., Yamaguchi, M., Nakamura, M., and Okhura, Y. 1985b. Determination of α -keto acids in serum and urine by high-performance liquid chromatography with fluorescence detection. *J. Chromatogr.* 344:33-39.
- Hara, S., Yamaguchi, M., Takemori, Y., and Nakamura, M. 1986. Highly sensitive determination of *N*-acetyl- and *N*-glycolylneuraminic acids in human serum and urine and rat serum by reversed-phase liquid chromatography with fluorescence detection. *J. Chromatogr.* 377:111-119.
- Hara, S., Takemori, Y., Yamaguchi, M., Nakamura, M., and Okhura, Y. 1987a. Fluorometric high-performance liquid chromatography of *N*-acetyl- and *N*-glycolylneuraminic acids and its application to their microdetermination in human and animal sera, glycoproteins, and glycolipids. *Anal. Biochem.* 164:138-145.
- Hara, S., Yamaguchi, M., Takemori, Y., Nakamura, M., and Okhura, Y. 1987b. 1,2-diamino-4,5-methylendioxybenzene as a highly sensitive fluorogenic reagent for α -keto-acids. *Chem. Pharm. Bull.* 35:687-692.
- Hara, S., Yamaguchi, M., Takemori, Y., Furuhashi, K., Ogura, H., and Nakamura, M. 1989. Determination of mono-*O*-acetylated *N*-acetylneuraminic acids in human and rat sera by fluorometric

high-performance liquid chromatography. *Anal. Biochem.* 179:162-166.

- Manzi, A.E., Diaz, S., and Varki, A. 1990. High pressure liquid chromatography of sialic acids on a pellicular resin anion-exchange column with pulsed amperometric detection: A comparison with six other systems. *Anal. Biochem.* 188:20-32.
- Spiker, J.E. and Towne, J.C. 1962. Fluorometric microdetermination of α -keto acids. *Anal. Chem.* 34:1468-1471.

Key References

Hara et al., 1989. See above.

Describes the use of the DMB method for the detection and quantitation of Neu5Ac, Neu5Gc, and mono-O-acetylated derivatives of Neu5Ac.

Manzi et al., 1990. See above.

Describes the application of the DMB method to other members of the sialic acid family; discusses differences in reactivity with the fluorophore and in response factors between different DMB adducts.

Contributed by Adriana E. Manzi
University of California San Diego
La Jolla, California

Leland D. Powell and Ajit Varki
(thiobarbituric acid assay; de-*O*-acetylation)
University of California San Diego
La Jolla, California

Total Compositional Analysis by High-Performance Liquid Chromatography or Gas-Liquid Chromatography

UNIT 17.19A

Once the presence of carbohydrate in a glycoprotein has been confirmed, the next step is to determine the precise molar ratio of its monosaccharide constituents using protocols in this unit and in the accompanying *UNIT 17.19B* (which presents the compositional analysis of labeled monosaccharides from glycosaminoglycans). This information helps to predict the type of oligosaccharide present and to determine the approach to detailed structural characterization. In some cases, it can also provide the first clues to the presence of a new type of sugar chain. The analysis involves two major phases. First, the release of the individual monosaccharides is achieved by methanolysis (Support Protocol 1), total acid hydrolysis (*UNIT 17.16*), or enzymatic release of sialic acids (*UNIT 17.12*). The second phase involves analysis by fractionation, characterization, and quantitation of the mixtures of monosaccharides by high-performance liquid chromatography using anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD; Basic Protocols 1, 2, and 3) and other HPLC systems (e.g., Alternate Protocol 1) or by gas-liquid chromatography with flame ionization detection (GLC-FID; Basic Protocol 4). The identity of the individual monosaccharides is determined by comparison with known standards processed and analyzed in the same way. Use of these protocols requires availability of instrumentation and understanding of setup maintenance and use.

STRATEGIC PLANNING

The flow chart in Figure 17.19.1 illustrates the alternative pathways and sequence of steps involved in compositional analysis. The protocols presented in this unit describe alternative approaches for the analysis of monosaccharide mixtures obtained by total acid hydrolysis (*UNIT 17.16*) or methanolysis (Support Protocol 1). The following considerations may help the investigator to plan the approach most appropriate for the starting material and intended results.

The choice of protocol to be used to release the individual monosaccharides from the starting glycoconjugate sample depends upon the nature of the material to be analyzed as well as the analysis method to be used. In general, glycoconjugates may be cleaved to their constituent monosaccharides by acid hydrolysis (*UNIT 17.16*) or methanolysis (Support Protocol 1). For the analysis of the constituents of neutral glycoconjugates, total acid hydrolysis of the sample is recommended. Glycoconjugates that may contain hexosamines or uronic acid can be broken down either by acid hydrolysis or methanolysis, and sialic acid residues can be released either enzymatically (as described in *UNIT 17.12*) or by hydrolysis in the presence of mild acid (*UNIT 17.16*). Note that the harsher acid conditions used to release the other monosaccharides cause destruction of sialic acids. For most glycoconjugates, the use of methanolysis to cleave glycosidic linkages is quite effective and generally results in less destruction of the monosaccharides than does acid hydrolysis (see Commentary for a more explicit discussion).

Frequently, the availability of instrumentation determines the choice between comparable options. For most HPLC approaches, such as those described in Basic Protocols 1, 2, and 3, the mixtures of monosaccharides released by acid hydrolysis is directly analysed without further treatment. Mixtures of monosaccharides, released by acid hydrolysis or by methanolysis, being subjected to analysis by GLC-FID (Basic Protocol 4), however, require an initial derivatization step (or series of steps) to convert the mixture to volatile

Preparation and
Analysis of
Glycoconjugates

17.19.1

Contributed by Adriana E. Manzi

Current Protocols in Molecular Biology (1995) 17.19.1-17.19.25

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 32

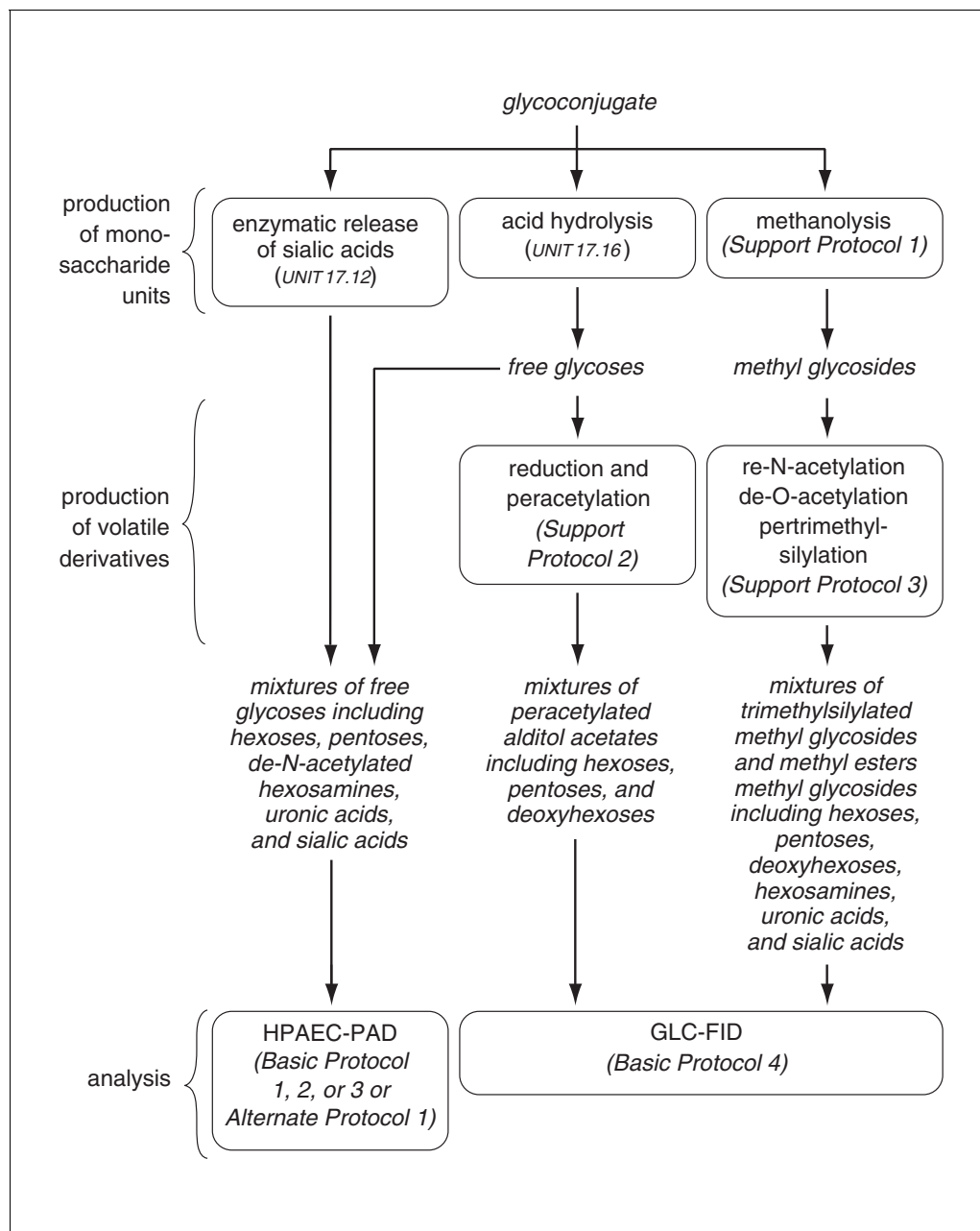


Figure 17.19.1 Flow chart showing the different methods included in this section for compositional analysis of the carbohydrate moiety of glycoconjugates.

compounds. These volatile derivatives can be produced from monosaccharide mixtures using one of the two alternative methods described in the support protocols. In the first derivatization approach, monosaccharide mixtures obtained from acid hydrolysis undergo reduction and peracetylation to produce volatile alditol acetates (Support Protocol 2). In the second approach, the products of methanolysis are converted to volatile trimethylsilyl derivatives (Support Protocol 3). In both cases, the resulting volatile compounds are then analyzed by GLC-FID.

Many different methods have been developed for each strategy, each with its own advantages and disadvantages (a more complete discussion and comparison of alternatives is presented in the Commentary). This unit presents reliable methods that have been proven to be widely applicable.

COMPOSITIONAL ANALYSIS OF FREE MONOSACCHARIDES BY HPAEC-PAD

The mixtures of monosaccharides obtained by acid hydrolysis (*UNIT 17.16*) of a glycoconjugate can be directly analyzed by high-performance liquid chromatography using ion-exchange pellicular resin columns with pulsed amperometric detection (HPAEC-PAD). Figure 17.19.2 provides a schema of a HPAEC-PAD system equipped with a reservoir for postcolumn addition of base. The following protocol (for separating mixtures of neutral monosaccharides and hexoamines) refers to this basic system but may also be adapted for use with newer instruments such as the Dionex DX-300 or DX-500, in which the arrangement of components is somewhat different. In the DX-500, for example, vacuum rather than helium sparging is used to degas the eluants, and because the PAD detector is more sensitive, postcolumn addition of base is not needed. Furthermore, the system can be completely controlled from a computer so that a procedure, including the elution conditions and the detection and integration parameters, can be created and stored as a method file. This method can later be called up and used. Consult the manufacturers' instructions for details of the use of these newer systems. The basic principles of operation and detection and the columns used for separation are the same.

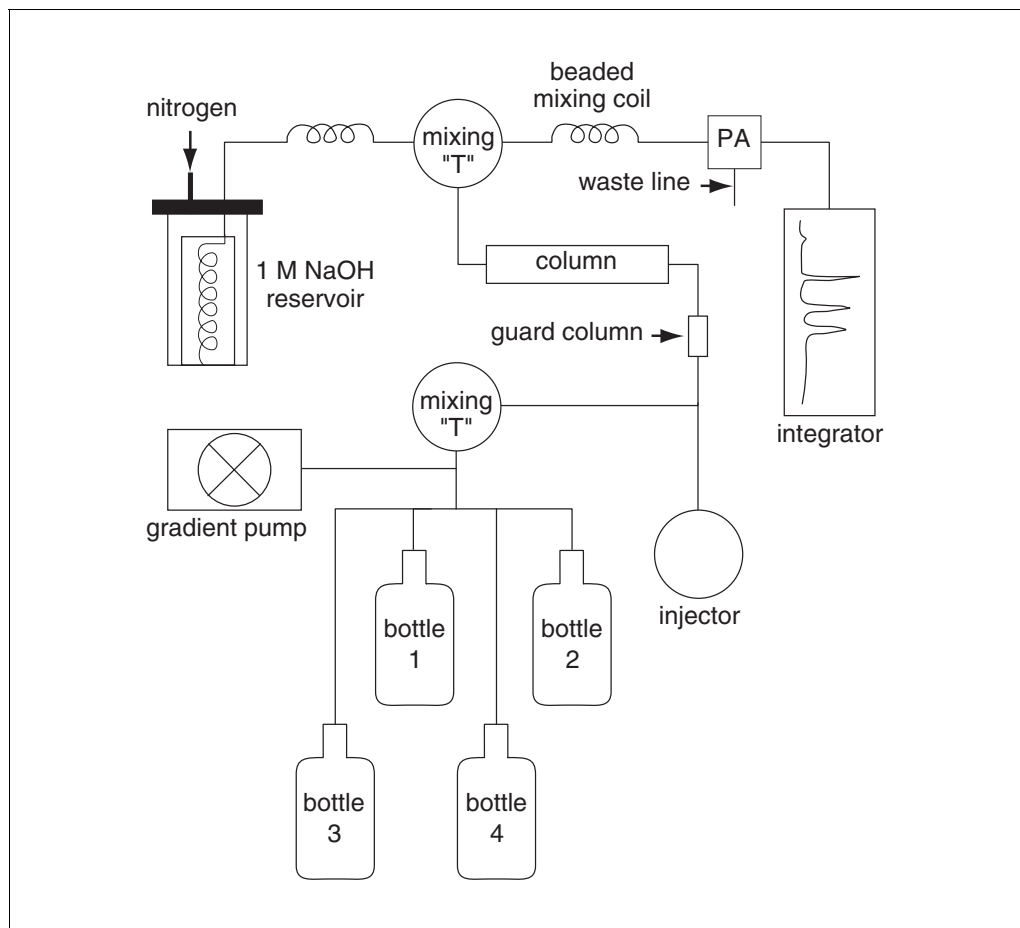


Figure 17.19.2 Schema of a high-performance anion-exchange chromatography system with pulsed-amperometric detector and post-column delivery system.

Analysis of Mixtures of Neutral Monosaccharides and Hexosamines

Materials

Milli-Q-purified water: water deionized by passage through a five-stage Milli-Q Plus system (Millipore)

Certified 50% (19.3 M) sodium hydroxide containing <0.1% sodium carbonate (Fisher)

Nitrogen

Monosaccharide standard mixture (see Support Protocol 4)

Monosaccharide sample prepared by acid hydrolysis (UNIT 17.16: Basic Protocol 3, Basic Protocol 1 when analyzing only fucose)

Filtering unit with 0.45- μ m Nylon 66 membranes (Alltech)

High-pH anion-exchange chromatography system with pulsed amperometric detector consisting of gradient pump, eluant degassing module, and PAD cell equipped with thin gasket (Dionex)

CarboPac PA-1 column (250 \times 4 mm; Dionex)

CarboPac PA Guard column (3 \times 25 mm; Dionex)

Prepare the HPAEC-PAD system

1. Filter 6 liters Milli-Q-purified water through a 0.45- μ m membrane. Transfer 2 liters to bottle 1, 1997 ml to bottle 2, and 1980 ml to bottle 3. Connect to helium sparge line and degas 15 min.

Always rinse the bottles before using them and make sure the lines are equipped with polyethylene inlet frit filters. If 4-liter bottles are available, double the volumes of liquid in the bottles.

2. Prepare 25 mM and 200 mM sodium hydroxide solutions by adding (using disposable plastic pipets) 2.6 and 20.8 ml of 50% sodium hydroxide solution to bottles 2 and 3, respectively. Avoid bubbling air when adding. Degas for another 15 min.

IMPORTANT NOTE: *It is critical that the sodium hydroxide solution be free of carbonate. Do not use sodium hydroxide pellets, as these are usually coated with a film of sodium carbonate produced by absorption of CO₂ from the air; instead, use only 50% (19.3 M) NaOH solution that is free of carbonate. Take the solution from the middle of the bottle. Any CO₂ absorbed will precipitate in 50% NaOH.*

Keep all eluants under helium or vacuum (depending on the instrument used) at all times. Discard any remaining eluants after ~2 weeks and prepare fresh ones. Do not store the Milli-Q-purified water to be used with this system in plastic containers. Rather, use freshly processed water taken directly from the Milli-Q purification system. Do not use plastic tubing extensions out of the Milli-Q system because these can become contaminated with microorganisms.

3. Run solvent from each bottle individually through the attached lines to remove any bubbles.
4. Connect the 4 \times 250-mm CarboPac PA-1 column to the 3 \times 25-mm CarboPac PA Guard column.
5. Set the pump control unit (or computer, if using a computerized system) to deliver a flow rate of 1.0 ml/min with an appropriate elution regimen (see Table 17.19A.1 for examples).

When both Xyl and Man are present in the sample, a very diluted eluent is used to achieve resolution. This considerably increases the time of analysis. Thus, when Xyl is not expected, an isocratic elution using 16 mM NaOH is preferred. These conditions are appropriate for the analysis of N- and O-linked oligosaccharides, and considerably shorter. However, it is important to keep in mind that Xyl and Man will coelute if both are present.

Table 17.19A.1 Sample HPAEC Elution Regimens for Compositional Analysis of Neutral Monosaccharides and Hexosamines^a

Time (min)	%			Comments
	Water	25 mM NaOH	200 mM NaOH	
Separation of Fuc, GalNH₂, GlcNH₂, Gal, Glc, Xyl, and Man				
Initial	92	8	0	
0.0	92	8	0	2 mM isocratic
40.0	92	8	0	
42.0	0	0	100	200 mM regeneration
47.0	0	0	100	End of regeneration
49.0	92	8	0	Reequilibration at initial conditions
90.0	92	8	0	
Separation of Fuc, GalNH₂, GlcNH₂, Gal, Glc, and Man				
Initial	92	0	8	
0.0	92	0	8	16 mM isocratic
22.0	92	0	8	
23.0	0	0	100	200 mM regeneration
33.0	0	0	100	End of regeneration
34.0	92	0	8	Reequilibration at initial conditions
50.0	92	0	8	

^aThe regeneration step between runs ensures reproducibility. When eluting with very diluted solutions such as those used in the first regimen above, at least 30 min of equilibration at the initial conditions is required to achieve good reproducibility between runs.

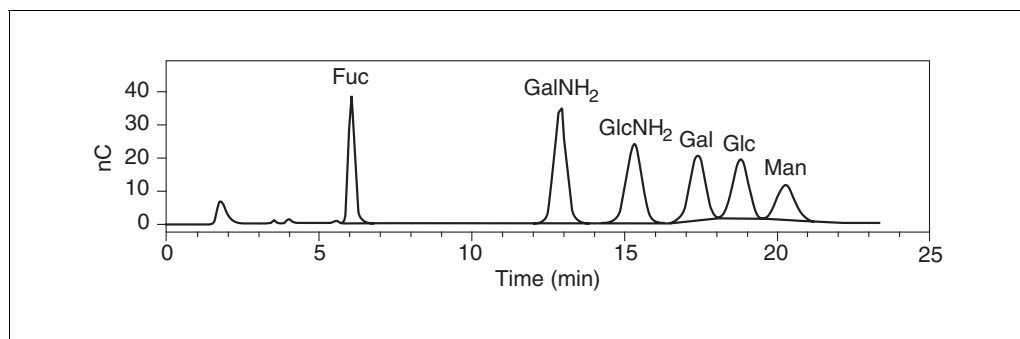


Figure 17.19.3 HPAEC-PAD profile of a mixture of monosaccharide standards containing Fuc, GalNH₂, GlcNH₂, Gal, Glc, and Man. The standards (10 nmol each) were hydrolyzed with 2 M trifluoroacetic acid (TFA) for 4 hr at 100°C, dried, and dissolved in 1.0 ml water. A 100- μ l aliquot (containing 1 nmol of each monosaccharide) was injected on the column and eluted isocratically using 16 mM NaOH.

Figure 17.19.3 shows a typical chromatogram of the mixture of standards used for compositional analysis of N- and O-linked oligosaccharides. Note that glucose is included, not because it is present in these oligosaccharides but because it is a very common contaminant.

6. Start the system and wait until a stable backpressure is attained (usual working pressures are between 1200 and 1500 psi and the maximum allowable pressure is 2000 psi).
7. Enter the settings for the PAD detector (or PED detector working in PAD mode) as indicated in Table 17.19A.2. When using a PADII detector, set the response time to 1 sec, the applied potential range to 1, and the output range to 300 nA. When using

Table 17.19A.2 PAD Detector Settings for Detection of Neutral Monosaccharides and Hexosamines

Time (sec)	Potential (V)	Integration
0.00	0.05	
0.20	0.05	Begin
0.40	0.05	End
0.41	0.75	
0.60	0.75	
0.61	-0.15	
1.00	-0.15	

an ED40 detector in a DX300 or DX500 instrument, select the carbohydrate waveform.

In the DX300 and DX500 systems, each method includes elution conditions, detector settings, and integration parameters. These can be manually selected from the control panels or entered using the software provided by the manufacturer. In both cases, when a method is loaded, the pump starts and the detector is turned on.

Perform postcolumn addition of base (when using a PADI or early versions of the PADI detector only)

8. Pressurize the postcolumn reservoir to deliver 0.3 to 0.4 ml/min of 1 M NaOH solution (helium pressure should be ~32 psi, resulting in a 360 mM sodium hydroxide concentration at the electrode). Check the amount delivered by measuring the total flow out of the PAD cell with a small graduated cylinder.

A regular piston pump cannot be used because this produces a wavy baseline due to pulsation.

When using newer versions of the PADI cell, this step (designed to improve the sensitivity of detection with the older models) is unnecessary. In this case proceed directly to step 10.

9. Make sure that there are no bubbles inside the reference electrode. Wait 15 min for the detector to stabilize, then press the Auto Offset control.

Check the baseline

10. Check the detector baseline using an integrator or the software provided by the manufacturer, and do not proceed until it is stable. Offset as necessary. When using an integrator, the appropriate settings are attenuation, 1024 and peak threshold, 10,000.

Analyze the standards

11. Dissolve dried acid hydrolyzed (UNIT 17.16) monosaccharide standard mixture in 100 to 1000 μ l water. Inject 10 to 100 μ l of mixture onto column.

Use 100 μ l when the mixture contains 2.5 nmol of each monosaccharide and 1000 μ l when the mixture contains 10 nmol of each monosaccharide. Sufficient quantities of the standard mixture should be hydrolyzed to allow for several injections if necessary.

The standard mixture should contain 250 pmol to 1 nmol each of Fuc, GalNH₂, GlcNH₂, Gal, Glc, Xyl, and Man; 250 pmol to 1 nmol each of Fuc, GalNH₂, GlcNH₂, Gal, Glc, and Man; or 250 pmol to 1 nmol Fuc (when only analyzing fucose). The

amount of each monosaccharide standard injected depends on the sensitivity of the particular detector and needs to be determined empirically in each case. The amount of standard mixture injected will depend on the amount of sample available. When the sample is limited, it is necessary to work at high sensitivity, and in this case will be more difficult to obtain a good baseline; also, the amount of standards injected should be adjusted to be in the same range for improved quantitation.

Always prepare standards with extreme care (see Support Protocol 3), as their quality will determine the accuracy of sample quantitation. Always use standards that have been hydrolyzed under the same conditions as the samples. This will correct for the different rates of hydrolysis and different recoveries of the different monosaccharides.

12. Check the resolution between monosaccharides (resolution of Gal/GlcN and Xyl/Man are usually critical, and inversion of the relative order of the components in these pairs may be observed). If resolution is poor, change the concentration of base (when retention times are too short, less base is required).

When using very dilute eluants (e.g., 2 mM NaOH), the concentration required for complete resolution may change when new 25 mM NaOH solution is prepared. This optimum concentration is also different for different columns. Therefore, when a new solution is prepared or a new column is used, always check the resolution obtained for the standard mixture. When using higher concentrations of NaOH as eluant, this sort of variation is not observed.

Analyze the samples

13. When good resolution of the standards is achieved, inject the samples.
14. When using PADI or PADII detectors, it may be necessary (depending on the response obtained for the samples) to change the working range to 100 nA or 1 μ A. Wait until a good baseline is obtained, then inject another sample.

When using newer detectors, these adjustments are made automatically.

15. Quantitate the standard run and use these data (area per nmol or pmol of each monosaccharide) to calculate the amount in nmol or pmol of each component in the sample.

It is advisable to average the results from at least three runs of standards to obtain more accurate response factors for each monosaccharide and improve the accuracy of the determination. This can be easily accomplished when using the software provided by the manufacturer by doing an average calibration and using those parameters to analyze the sample runs.

Maintain the system running at maximum performance

16. If resolution is lost and retention times become too short, wash the column for 30 min with 1 M sodium hydroxide. Wash extensively with water, then return to the working eluant. Wait for the baseline to stabilize under running conditions and check the performance of the column by injecting 10 μ l standard mixture. If performance is not recovered, wash the column with 1 M HCl for 30 min at 1.0 ml/min. Wash extensively with water, then with 1 M NaOH for 30 min at the same flow rate. Wash extensively with water, reequilibrate with the working eluant, and check again.

Prepare high-concentration eluants used for regeneration in separate bottles, or wash the bottles extensively after they have been used to hold these solutions. Also, remember to wash the lines that have carried concentrated solutions before connecting them to the column for regular analysis.

**Analysis of Mixtures of Neutral Monosaccharides, Hexosamines,
and Uronic Acids**

The use of a sodium acetate/sodium hydroxide gradient allows the elution of anionic saccharides, such as uronic acids, that would otherwise be retained by the column.

Additional Materials (also see *Basic Protocol 1*)

Sodium acetate ($\text{CH}_3\text{CO}_2\text{Na}\cdot 3\text{H}_2\text{O}$, FW 136.08; ACS certified grade)

1. Filter 8 liters Milli-Q-purified water through 0.45- μm membrane. Transfer 2 liters to bottle 1. Connect to helium sparge line and degas 15 min.
2. Prepare the following solutions in bottles 2, 3, and 4:

Bottle 2 (200 mM sodium hydroxide): Transfer 1980 ml filtered Milli-Q-purified water to bottle and degas for 15 min. Using a disposable plastic pipet, add 20.8 ml of 50% sodium hydroxide, then continue degassing for 15 min.

Bottle 3 (1 M sodium acetate): Weigh 272 g sodium acetate and dissolve in filter through 0.45- μm membrane, Milli-Q-purified water, then dilute to 2000 ml in a graduated cylinder. Filter through a 0.45- μm membrane. Transfer to bottle, and degas for 15 min.

Bottle 4 (1 M sodium hydroxide): Transfer 1896 ml filtered Milli-Q-purified water to bottle 4 and degas for 15 min. Add 104 ml of 50% sodium hydroxide measured in a plastic cylinder, and continue degassing for 15 min.

3. Run solvent from each bottle individually through the attached lines to remove any bubbles.
4. Connect 4 \times 250-mm CarboPac PA-1 column to the 3 \times 25-mm CarboPac PA Guard column.
5. Set the pump control unit or computer to deliver a flow rate of 1.0 ml/min with an appropriate elution regimen (see Table 17.19A.3 for an example).

A typical resulting chromatogram is shown in Fig. 17.19.4.

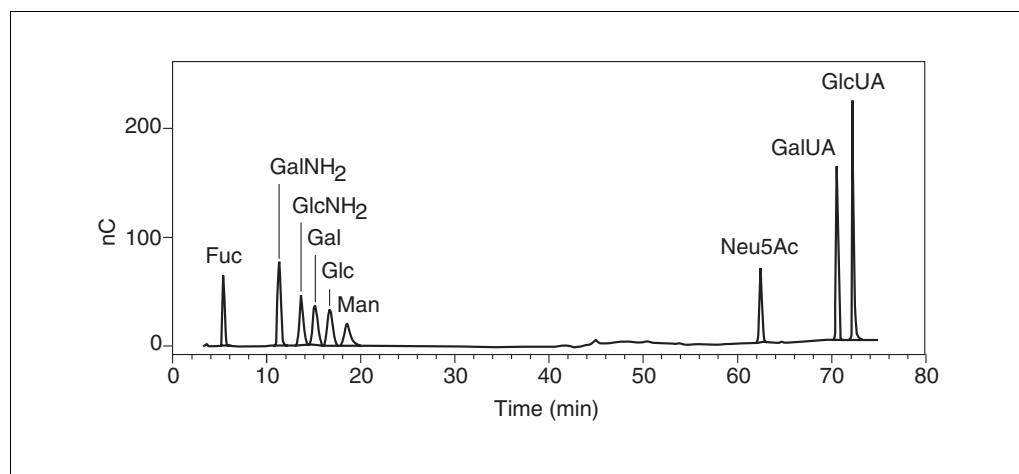


Figure 17.19.4 HPAEC-PAD profile of a mixture of monosaccharide standards containing Fuc, GalNH₂, GlcNH₂, Gal, Glc, Man, Neu5Ac, GalUA, and GlcUA. The standards (20 nmol of each neutral and amino sugar and 50 nmol of each hexouronic acid) were hydrolyzed with 2 M trifluoroacetic acid (TFA) for 6 hr at 100°C and dried; 20 nmol Neu5Ac was then added and the mixture dissolved in 1.0 ml water. A 100- μl aliquot (containing 2 nmol of each neutral monosaccharide, hexosamine, and Neu5Ac, and 5 nmol of each glycosyluronic acid) was injected on the column and eluted using the gradient described in Table 17.19A.3.

Table 17.19A.3 Sample HPAEC Elution Regimen for Compositional Analysis of Neutral Monosaccharides, Hexosamines, and Uronic Acids^a

Time (min)	%				Comments
	Water	200 mM NaOH	1 M NaC ₂ H ₃ O ₂	1 M NaOH	
0	92.0	8.0	0.0	0.0	
0.01	92.0	8.0	0.0	0.0	16 mM NaOH
25	92.0	8.0	0.0	0.0	
35	80.0	20.0	0.0	0.0	40 mM NaOH
40	70.0	30.0	0.0	0.0	60 mM NaOH
45	88.0	0.0	2.0	10.0	100 mM NaOH/ 20 mM NaC ₂ H ₃ O ₂
50	87.2	0.0	3.0	9.8	98 mM NaOH/ 30 mM NaC ₂ H ₃ O ₂
55	87.2	0.0	3.0	9.8	
75	70.4	0.0	20.0	9.6	96 mM NaOH/ 200 mM NaC ₂ H ₃ O ₂
95	22.5	0.0	70.0	7.5	75 mM NaOH/ 700 mM NaC ₂ H ₃ O ₂
97	80.0	0.0	0.0	20.0	200 mM NaOH
107	80.0	0.0	0.0	20.0	
109	92.0	8.0	0.0	0.0	16 mM NaOH
120	92.0	8.0	0.0	0.0	Reset ^b

^aThis gradient separates Fuc, GalNH₂, GlcNH₂, Gal, Glc, Man, GalUA, and GlcUA. A typical chromatogram is shown in Figure 17.19.4.

^bNext injection follows reset (at 130 min).

6. Analyze the standards and samples (see Basic Protocol 1, steps 12 to 15).

The standard mixture should contain 250 pmol to 1 nmol of each neutral monosaccharide and hexosamine and of the appropriate glycosyluronic acid. As in Basic Protocol 2, inject 10 to 100 μ l of this mixture on the column.

Analysis of Mixtures of Sialic Acids

Sialic acids can be separated using the same anion-exchange pellicular resin column without base in the eluant; this permits analysis of O-acetylated species. HPAEC-PAD can be used when ≥ 0.5 nmol are available. Therefore, it is possible to determine the molar ratios of individual sialic acids in mixtures that have been released from glycoconjugates and purified using the protocols described in *UNITS 17.12 & 17.16*. This procedure can also be used for preparative purposes using a larger chromatography column.

Additional Materials (also see *Basic Protocol 2*)

Glacial acetic acid (17.4 M)

N-acetylneuraminic acid (Neu5Ac, MW 309.3; Boehringer Mannheim)

N-glycolylneuraminic acid (Neu5Gc, MW 325.3; Sigma)

Standard mixture: sialic acids prepared from bovine submaxillary mucin (BSM) using the procedures in *UNIT 17.16* (Basic Protocol 2 or Alternate Protocol) or *UNIT 17.12*, or equimolar mixture of Neu5Gc and Neu5Ac processed as for the sample

BASIC PROTOCOL 3

Preparation and Analysis of Glycoconjugates

17.19.9

Prepare the HPAEC-PAD system

1. Filter 2 liters Milli-Q-purified water through 0.45- μ m membrane. Transfer 1 liter to bottle 1, connect to helium sparge line, and degas 15 min.
2. Prepare 1 M sodium hydroxide solution (see Basic Protocol 1, step 2).
3. Weigh 0.68 g sodium acetate and dissolve in Milli-Q-purified water, diluting the solution to 1 liter final. Filter the resulting 5 mM sodium acetate buffer, transfer to bottle 2, connect to helium sparge line, and degas for 15 min.
4. Prepare 1 liter of 5 mM acetic acid using filtered Milli-Q-purified water and glacial acetic acid. Transfer to bottle 3 and degas for 15 min.
5. Run each solvent through the attached lines to remove any bubbles.
6. Connect the 4 \times 250 mM CarboPac PA-1 column with the 3 \times 25-mm CarboPac PA Guard column.
7. Set the pump control unit or computer to deliver a flow rate of 1.0 ml/min with an appropriate two-step elution regimen (see Table 17.19A.4 for an example).

For preparative purposes, a 9 \times 250-mm column at a flow rate of 5 ml/min should be used.

8. Start the system and wait until a stable back pressure is attained (usual working pressures are between 950 and 1400 psi and the maximum allowable pressure is 2000 psi).
9. Perform postcolumn addition of base (if using PADI or earlier PADII system) and determine detector baseline (see Basic Protocol 1, steps 8 to 11).

Analyze the standards

10. Inject 10 μ l of a water solution containing 5 nmol Neu5Ac, working in the 300 nA range.

If the samples to be analyzed have been submitted to acid hydrolysis for the release of sialic acids, use a Neu5Ac standard treated in the same manner for quantitation.

Table 17.19A.4 HPAEC Elution Regimen for Compositional Analysis at Sialic Acids

Time (min)	%				Comments
	Water	5 mM NaC ₂ H ₃ O ₂	5 mM C ₂ H ₄ O ₂	1 M NaOH	
Initial	0	100	0	0	
0.0	0	100	0	0	5 mM NaC ₂ H ₃ O ₂
0.01	0	100	0	0	
5	0	100	0	0	
35	0	50	50	0	2.5 mM NaC ₂ H ₃ O ₂ / 2.5 mM C ₂ H ₄ O ₂
36	0	0	100	0	5 mM AcOH
46	0	0	100	0	
48	80	0	0	20	200 mM NaOH
68	80	0	0	20	
70	0	100	0	0	5 mM NaC ₂ H ₃ O ₂
100	0	100	0	0	

Table 17.19A.5 Relative Elution Times of Sialic Acids Found in Glycoproteins with Different HPLC Systems

Sialic acid	t_{Neu5Ac} in system	
	HPAEC	Amino column ^a
Neu5Ac	1.00	1.00
Neu5Gc	1.17	1.5
Neu5-7Ac ₂	0.74	0.36
Neu5-9Ac ₂	0.95	0.35
Neu4-5Ac ₂	0.76	0.39
Neu5-7(8)-9Ac ₃	0.74	0.23
Neu5-7-8-9Ac ₄	0.62	— ^b
Neu5Gc9Ac	1.06	0.49
Neu4Ac5Gc	0.86	0.54

^aUsing condition B.

^bElutes close to the salt peak.

- Change the range to 100 nA, wait until a good baseline is obtained again, and inject 500 pmol of Neu5Ac in the same volume.

When using a newer version of the PADII detector, these adjustments will be done automatically.

- If a standard mixture containing several sialic acids (obtained from BSM) is available, inject 10 μl containing 500 pmol of each and check the resolution. If this is not available, an equimolar mixture of Neu5Ac and Neu5Gc can be used.

The total running time is under 30 min.

Analyze the samples

- When good resolution and adequate retention times are achieved, inject the sialic acid samples. Compare the relative retention times with those listed in Table 17.19A.5.

It is advisable to average the results from at least three runs of standards to obtain more accurate response factors for each monosaccharide and improve the accuracy of the determination. This can be easily accomplished when using the software provided by the manufacturer by doing an average calibration and using those parameters to analyze the sample runs.

- Characterization of individual peaks can be achieved by repeating the analysis after saponification of *O*-acetyl groups (see UNIT 17.18).
- Determine the response obtained for standard Neu5Ac (area per pmol) and calculate the molar ratio of sialic acids in the mixture considering the following response factors:

<i>Sialic acid</i>	<i>Response factor</i>
Neu5Ac	1.00
Neu5Gc	1.15
Neu5-9Ac ₂	0.46
Neu5-7(8)-9Ac ₃	0.45

Detector responses for other sialic acids have not been determined, but lower response factors for more highly substituted species should be expected.

COMPOSITIONAL ANALYSIS OF FREE SIALIC ACIDS BY AMINE- ADSORPTION ION-SUPPRESSION HPLC WITH UV DETECTION

As an alternative to Basic Protocol 3, mixtures of free sialic acids can be analyzed by amine-adsorption ion-suppression HPLC with UV detection. The choice depends on the instrumentation available and the level of sensitivity required for a particular sample.

This system can be used to analyze radioactive samples if the column effluent is routed to a flow detector or fraction collector. When a flow detector is used, the appropriate ratio of scintillation fluid to be used has to be checked because of the quenching effect of acetonitrile. When fractions are collected and 72% acetonitrile is used, it is necessary to dry them, redissolve them in water, and then measure their radioactivity in a β -counter after addition of scintillation cocktail. When 64% acetonitrile is used, the addition of 1 vol water followed by 10 vol scintillation cocktail yields good sensitivity.

Materials

- Monobasic sodium phosphate (anhydrous, MW 120.0; reagent grade, Sigma)
- Milli-Q-purified water: water deionized by passage through five-stage Milli-Q Plus system (Millipore)
- Acetonitrile (HPLC grade, Fisher)
- N*-acetylneuraminic acid (Neu5Ac, MW 309.3; Boehringer Mannheim)
- N*-glycolylneuraminic acid (Neu5Gc, MW 325.3; Sigma)
- Standard mixture: sialic acids from bovine submaxillary mucin (BSM) prepared using procedures in *UNIT 17.16* or *UNIT 17.12*, or equimolar mixture of Neu5Ac and Neu5Gc processed as for the sample
- Sialic acid sample prepared by acid hydrolysis (*UNIT 17.16*, Basic Protocol 2 or Alternate Protocol) or enzymatic release (*UNIT 17.12*)

- HPLC apparatus with pump and UV detector
- Filtering unit with 0.45- μ m Nylon 66 membranes (Alltech)
- Micropak AX-5 (300 mm \times 7.8 mm i.d., particle size 9 μ m; Varian)
- Micropak AX-5 Guard column (4 cm \times 4 mm)

- Additional reagents and equipment for de-O-acetylation (*UNIT 17.18*)

Prepare the HPLC system

1. Turn on the UV detector and set the wavelength to 200 nm.
2. Prepare 1 liter of 0.25 M monobasic sodium phosphate (30 g/liter) in Milli-Q-purified water and filter through 0.45- μ m membrane.
3. Fill the bottles in the HPLC system with (1) acetonitrile, (2) Milli-Q-purified water filtered through 0.45- μ m membrane, and (3) filtered 0.25 M monobasic sodium phosphate from step 2. Degas 15 min by helium sparging.
4. Run each solvent through the attached lines to eliminate any bubbles.

CAUTION: Do not allow direct mixing of stock acetonitrile and sodium phosphate, as this will cause precipitation of salts in the lines.

5. Set the pump control unit to deliver a flow rate of 1 ml/min with an appropriate isocratic elution regimen as indicated in Table 17.19A.6.
6. Start running the solvent, then connect the 300 \times 7.8-mm Micropak AX-5 column to the 4 cm \times 4 mm Guard column.
7. Wait until the back pressure stabilizes and check the baseline. Wait until a stable baseline is obtained.

Table 17.19A.6 Amine Absorption/Ion Suppression HPLC Elution Regimen for Compositional Analysis of Sialic Acids

Conditions ^a	Flow (ml/min)	%		
		1	2	3
A	1	64	26	10
B	1	72	18	10

^aCondition A is used when only separation of Neu5Ac and Neu5Gc is required. Condition B is used when a mixture containing substituted species will be analyzed.

Analyze the standards

8. Inject 10 μ l of a solution of Neu5Ac in water (10 nmol Neu5Ac total). Check the sensitivity and adjust the integrator settings as required (a good signal-to-noise ratio is normally obtained using attenuation range 8). Repeat if necessary.

It is possible to work in attenuation range 4 to detect 5 nmol of Neu5Ac, but the background is considerably higher.

9. Inject 10 μ l of a mixture of Neu5Ac and Neu5Gc standards containing 10 nmol of each monosaccharide. If using condition B, use a mixture containing substituted sialic acid standards instead obtained from BSM (if available). Check for resolution and adjust the elution conditions if required.

CAUTION: *Remember that sodium phosphate will start precipitating at >72% acetonitrile. Condition B constitutes the maximum compatible percentages of acetonitrile and phosphate.*

Analyze the samples

10. Inject the sialic acid samples. Adjust the integrator settings as required and repeat if necessary.
11. Determine the percentage of each component, assuming equal detector responses.
12. Confirm the presence of O-acetylated species by subjecting the sample to de-O-acetylation (UNIT 17.18), then repeating the analysis. Compare the relative retention times to those listed in Table 17.19A.1.

Maintain the system running with maximum performance

13. Wash the column 30 min with water. Run 100% acetonitrile for 10 min. Disconnect the column.

CAUTION: *Never stop the flow if the lines contain phosphate, as they may become clogged. Also wash the filter and the line coming from the reservoir after using any salt.*

14. If resolution is lost and retention times become reduced, clean the column by running 0.5 M phosphoric acid for at least 1 hr at 1 ml/min flow, then washing extensively with water before changing to the working buffer.

More complete regeneration of the column can be achieved by injecting three 1-ml volumes of 3-aminopropyltriethoxysilane while running the column in 100% acetonitrile, washing extensively with the same solvent, with water, and then with 0.5 M phosphate before use. The performance of the column should be fully recovered after this treatment; however, the retention times may not be identical.

COMPOSITIONAL ANALYSIS BY GLC-FID

This protocol describes the quantitative analysis of derivatized methyl glycosides and alditols through the use of a gas-liquid chromatograph with a flame-ionization detector (see Fig. 17.19.5). With this apparatus, it is quite straightforward to determine the molar ratio of the monosaccharide components of a given glycoconjugate. This analysis can be performed after complete acid hydrolysis (*UNIT 17.16*) and conversion of the free monosaccharides to volatile peracetylated alditols (see Support Protocol 2). It is also possible to cleave the glycosidic linkages with methanol/HCl to produce methyl glycosides, which are converted into volatile trimethylsilylethers (see Support Protocol 3) for further GLC analysis. The last method is preferred when hexosamines or uronic acid are expected to be found in the sample.

Materials

- Monosaccharide standard mixture (see Support Protocol 3) prepared in the same manner as the sample
- Monosaccharide sample prepared by acid hydrolysis (*UNIT 17.16*, Basic Protocol 3) followed by derivatization by peracetylation (see Support Protocol 2) and/or by methanolysis (Support Protocol 1) followed by derivatization by pertrimethylsilylation (see Support Protocol 3)
- Internal standard prepared in the same manner as the sample
- 0.25 mm × 30 m 5% DB-5 fused silica capillary column (J&W Scientific)
- Gas chromatograph with dual-flame ionization detector
- Helium (research grade)
- Hydrogen and air lines (for FID detector only)

Gas-liquid chromatograph with flame-ionization detector (GLC-FID)

1. Connect the capillary column to the gas chromatograph and begin running the carrier gas at a flow rate of 1.5 ml/min. Set the makeup gas flow rate of 35 ml/min and the split ratio to 10:1.

When <25 μg of total carbohydrates are processed, a splitless injection is required.

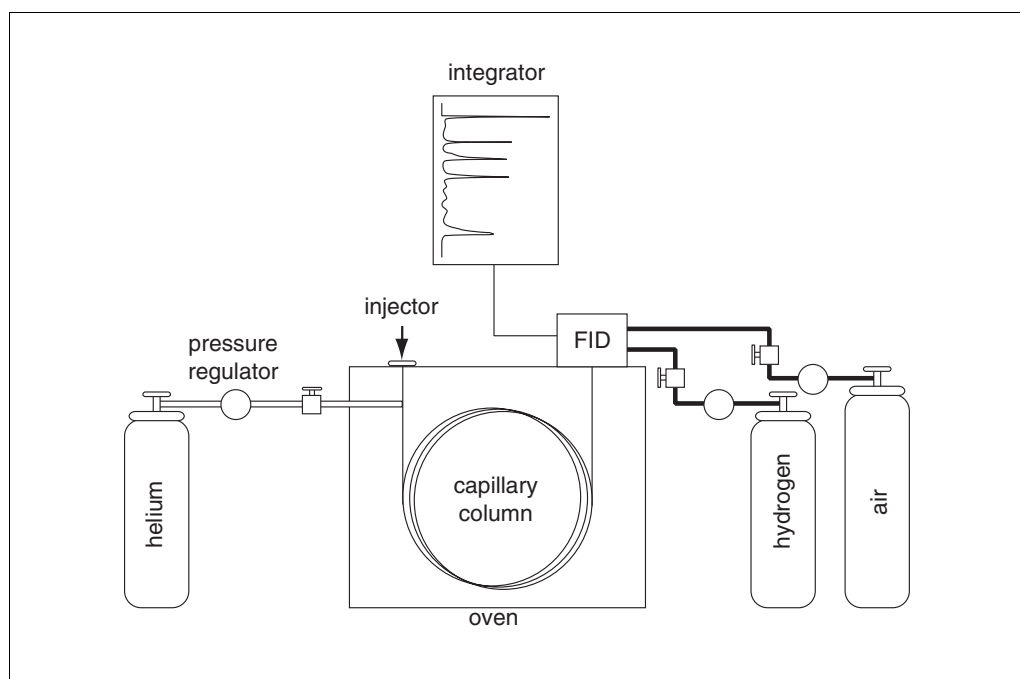


Figure 17.19.5 Schema of a gas-liquid chromatograph with flame-ionization detector.

Several other types of stationary phases can be used for the analysis of acetylated alditols: e.g., CP-Sil 5 WCOT, SP-1000, Silar 10C, OV-1, SE-54, 5% DB-5, DB-1, Carbowax 20M, SE-30, OV-101, and OV-275. In the less polar stationary phases, amino sugar alditols can also be analyzed. Other stationary phases used for the analysis of trimethylsilylated glycosides and methyl glycoside methyl esters include CP-Sil 5 WCOT, 5% DB-5, DB-1, SE-30, and OV-101. Appropriate conditions vary from one to another. Typical chromatograms and relative retention times can be found in Kammerling and Vliegthart (1989) and references therein.

2. Set the GLC oven temperature to 50°C, the injector-port temperature to 150°C, and the FID detector temperature to 250°C, and wait for stabilization.
3. Open the air and hydrogen lines and adjust the flows. Ignite the detector. Check the background.
4. Set up an appropriate temperature program:
 - a. *Alditol acetates*: Start at 50°C and keep at 50°C for 2 min; then go from 50°C to 150°C at 20°C/min (5 min); then go from 150°C to 250°C at 4°C/min (25.0 min); then keep at 4°C for 5 min. (Total running time is 37 min.)
 - b. *Trimethylsilylethers of methyl glycosides*: Start at 50°C and keep at 50°C for 3 min; then go from 50°C to 170°C at 20°C/min (6 min); then go from 170°C to 250°C at 6°C/min (13.3 min); then keep at 250°C for 2.7 min. (Total running time is 25 min.)
5. Inject 1 µl of the mixture of derivatized standards (see Support Protocol 2 and Support Protocol 3). Adjust the settings of the integrator and repeat as necessary to obtain all the peaks on scale.
6. Determine the retention times of each peak relative to the internal standard.
7. Inject the standards (containing internal standard) one at a time to assign the peaks and determine the response factors. In the case of TMS derivatives, also determine the contribution of each peak to the total (%), and select the peak of each monosaccharide to be used for quantitation:

$$R = \frac{\text{nmol internal standard} \times \text{area monosaccharide}}{\text{nmol monosaccharide} \times \text{area internal standard}}$$
8. Inject 1 µl of sample (containing internal standard). Repeat as needed to obtain a good profile.
9. Tentatively assign the peaks by comparison of the relative retention times with those of the standards. If a discrepancy arises, repeat analysis, coinjecting the sample with the individual standard. The questioned peak should appear separated from the standard if it is not the same compound.
10. Determine the molar percentage of individual components by reference to the internal standard:

$$\text{nmol sugar}/\mu\text{g sample} = \frac{\text{area of monosaccharide peak} \times \text{nmoles internal standard}}{\text{area of internal standard peak} \times R \times \mu\text{g of sample}}$$

**QUANTITATIVE RELEASE OF NEUTRAL MONOSACCHARIDES,
HEXOSAMINES, AND URONIC ACIDS BY METHANOLYSIS**

Glycoconjugate samples are hydrolyzed by heating in the presence of methanolic HCl, followed by treatment with pyridine and acetic anhydride.

Materials

- Glycoconjugate sample containing 5 to 50 µg total sugar
- Internal monosaccharide standard: monosaccharide that does not occur naturally in the sample (e.g., monosaccharide alditol that yields only one peak upon gas chromatography, see *UNIT 17.16*; see Support Protocol 4 for monosaccharide preparation)
- Monosaccharide standards: monosaccharides expected to be found in the sample (see Support Protocol 4 for preparation)
- Phosphorus pentoxide (Fisher)
- Acetyl chloride (FW 78.50, 98% pure; Aldrich)
- Methanol (anhydrous, 99% pure, Aldrich; store in desiccator)
- Pyridine (anhydrous silylation grade, Pierce; store in desiccator)
- Acetic anhydride
- 1.5-ml glass Reacti-Vials (Pierce) with Teflon-lined screw caps, either new (rinsed with water, then with ethanol, and dried) *or* acid-cleaned (see recipe)
- Heating block or oven
- Nitrogen or vacuum evaporation system (Speedvac or shaker-evaporator)

1. Combine the glycoconjugate sample and 1 to 10 µg internal monosaccharide standard in a 1.5-ml precleaned or acid-cleaned Reacti-Vial. Lyophilize mixture.
2. Prepare a set of standards consisting of separate solutions containing 50 nmol of each monosaccharide that is expected to be found in the sample, each one also containing 5 nmol of the same internal standard used for the sample. Also prepare a mixture of all the expected monosaccharides plus internal standard. Treat all standards in parallel with the sample.
3. Dry the lyophilized sample and standards overnight in a vacuum desiccator over phosphorus pentoxide.
4. Prepare methanolic HCl (either reagent A or reagent B) by adding acetyl chloride to anhydrous methanol in the following proportions:

	Acetyl chloride	Anhydrous methanol	Final concentration of HCl in methanol
Reagent A	3.50 vol	100 vol	0.50 M
Reagent B	7.00	100 vol	1.00 M

CAUTION: *Add the acetyl chloride carefully, as the resulting reaction is highly exothermic.*

Methanolic HCl reagent A is used except when glycosyluronic acids are present; in such cases reagent B is preferred. Alternatively, an initial acid hydrolysis may be performed using 200 µl of 2 M trifluoroacetic acid (TFA) for 1 hr at 121°C. TFA should then be evaporated under a nitrogen stream or in Speedvac system and samples washed twice with methanol, then evaporated again. This procedure ensures complete recovery of hexosamines but destroys sialic acids. Afterwards, proceed to step 4, and use reagent A.

Prepare a large enough volume of methanolic HCl to minimize pipetting errors, but only in the range required for the experiment, as it cannot be stored.

- Dissolve dried sample and standards (from step 3) in 200 μ l methanolic HCl (reagent A or B), then cap the vials.
- Heat 16 hr at 65°C if using reagent A or 24 hr at 85°C if using reagent B. Check the screw caps for tightness after the first 15 min of heating, then vortex and continue heating.

This procedure yields methyl glycosides of each monosaccharide residue in the glycoconjugate. Hexosamines will be de-N-acetylated.

- Evaporate sample and standards using a stream of nitrogen at room temperature or a Speedvac system.
- Add 100 μ l anhydrous pyridine and vortex to resuspend the residue. Add 100 μ l acetic anhydride, vortex 1 min, and evaporate to dryness under a stream of nitrogen.

This treatment produces re-N-acetylation of hexosamines. Some O-acetyl groups may be incorporated.

- Dissolve the residue in 200 μ l of 0.5 M methanolic HCl, then heat 1 hr at 65°C.

This treatment assures the cleavage of any O-acetyl groups incorporated in step 8.

- Evaporate under a stream of nitrogen, then dry further overnight in a vacuum desiccator over phosphorus pentoxide.

PREPARATION OF VOLATILE DERIVATIVES OF FREE GLYCOSES

Volatile derivatives of free glycoses are prepared by first reducing the free sugars to alditols with sodium borohydride, then peracetylation with acetic anhydride. The resulting alditol acetates can be analyzed immediately by GLC-FID (see Basic Protocol 2).

Materials

- 10 mg/ml sodium borohydride in 1 M ammonium hydroxide (ACS reagent grade)
- Monosaccharide sample prepared by acid hydrolysis (UNIT 17.16, Basic Protocol 3, or Basic Protocol 1 when analyzing only fucose)
- Neutral sugar standard mixture (see Support Protocol 4) submitted to acid hydrolysis (UNIT 17.16, Basic Protocol 3)
- Sodium borohydride (FW 37.8; store in desiccator)
- Glacial acetic acid (FW 60.05; 99.99% pure)
- 1% acetic acid in methanol (anhydrous, 99% pure)
- Acetic anhydride (99% pure)
- Toluene
- Chloroform (anhydrous, 99% pure)
- Acetone (99.9% pure, HPLC grade)
- Phosphorous pentoxide (as desiccant)
- 3.5-ml glass vials with Teflon-lined screw caps
- Nitrogen evaporation unit (Reacti-Vap Evaporator, Reacti-Therm Heating Module, and Reacti-Block B-1 or S-1, Pierce)
- Heating block or oven, 100°C
- Tabletop centrifuge
- Reacti-Vials (Pierce) or other glass test tubes

Reduce mixtures of free glycoses to alditols

- Add 250 μ l of 1 M ammonium hydroxide/10 mg/ml sodium borohydride to the monosaccharide sample and the neutral sugar standard mixture in 3.5-ml glass vials with Teflon-lined screw caps. Let stand 10 min, then check for persistence of

SUPPORT PROTOCOL 2

Preparation and Analysis of Glycoconjugates

17.19.17

microbubbles in the reaction mixture. If these cannot be seen, add a few micrograms of solid sodium borohydride.

It is important that no residue of acid from the hydrolysis be present or the hydride will be rapidly destroyed. If excessive bubbling occurs when adding the borohydride, add methanol and evaporate two or three times, then repeat the reduction.

3. Incubate 2 hr at room temperature.

This reaction can also be run overnight.

4. Add glacial acetic acid dropwise until no further bubbling is observed.

Care must be taken to avoid losing sample with the bubbling, as the reaction is violent. This problem may be circumvented by using a 20% (v/v) solution of acetic acid in methanol instead of glacial acetic acid.

5. Add 250 μ l of 1% acetic acid in methanol to each tube. Evaporate to dryness using a nitrogen evaporation unit or under vacuum using a Speedvac system with the heating element set at 40°C. Repeat four more times.

This treatment removes volatile methyl borate, producing sodium acetate. Complete removal of borate is indicated by colorless crystals, as compared with white powder previously present. When a nitrogen evaporation unit is employed, care must be taken not to blow away the solid residue, particularly in the final steps.

6. Dry sample and standards in a vacuum dessicator overnight over phosphorus pentoxide.

Peracetylate the alditols

7. Add 200 μ l acetic anhydride, vortex, and cap the tubes.

8. Heat 3 hr at 100°C, mixing after the first 15 min of heating to make sure that the solid residue is suspended. Let cool to room temperature.

9. Add 200 μ l toluene. Evaporate to dryness at 40°C using a shaker-evaporator or Speedvac or under nitrogen. Add toluene as needed to dry completely.

Toluene helps to eliminate the reagents by forming an azeotrope.

Clean reaction mixtures by extraction

10. Add 0.5 ml water and 0.5 ml chloroform to each tube, cap, and vortex well. Centrifuge briefly at low speed to clearly separate the two phases.

Salts will be extracted into the water layer, leaving the peracetylated alditols in the organic layer.

11. Remove the upper (water) layer with a Pasteur pipet. Wash the organic layer with water two more times.

12. Transfer the chloroform layer to a clean tube or Reacti-Vial and dry under nitrogen at room temperature or in a Speedvac system.

13. Dissolve in 5 to 50 μ l acetone, depending on the amount of starting material, and use 1 μ l for GLC-FID analysis (Basic Protocol 2).

PREPARATION OF VOLATILE DERIVATIVES OF METHYL GLYCOSIDES

SUPPORT
PROTOCOL 3

Volatile derivatives of methyl glycosides (trimethylsilylethers of neutral sugars and hexosamines and methyl esters of uronic acids) are prepared from the dried methyl glycosides (see Support Protocol 1) using a commercial silylation reagent. The silylated samples must then be analyzed immediately by GLC-FID (see Basic Protocol 2).

Materials

Dried methanolized samples and standards (see Support Protocol 1, step 10)
Tri-Sil Reagent [2:1:10 (v/v/v) hexamethyldisilazane/trimethylchlorosilane/
pyridine; Pierce]
96% hexane (HPLC grade)
Nitrogen stream
Tabletop centrifuge
Reacti-Vials (Pierce)

1. Suspend each dried sample and standard in 100 μ l Tri-Sil and cap vial. Incubate 30 min at room temperature.

Samples and all required standards must be processed in parallel.

2. Evaporate the silylating reagent under a stream of dry nitrogen.

Nitrogen can be dried by including a drying tube in the line.

3. Add 0.5 ml hexane, dissolve the derivatives, and centrifuge briefly at low speed to separate insoluble salts.
4. Transfer supernatant to a clean Reacti-Vial and evaporate under dry nitrogen stream as in step 2.
5. Dissolve the residue 5 to 50 μ l hexane, depending on the amount of starting material. Use 1 μ l of solution for GLC analysis.

PREPARATION OF MONOSACCHARIDE STANDARD SOLUTIONS

SUPPORT
PROTOCOL 4

Proper preparation and storage of monosaccharide standard solutions is critical to achieve good quantitation in a compositional analysis (Hardy and Townsend, 1994).

Materials

NaOH pellets
Monosaccharide standards
Milli-Q-purified water: water deionized by passage through a Milli-Q Plus system
(Millipore)
Vacuum desiccator
Glass vials
100-ml volumetric flasks

1. Prepare a vacuum desiccator containing a plastic beaker with NaOH pellets.
2. Tare a group of glass vials (twice the number of standards to be used). Weight ~500 mg of each monosaccharide standard in separate vials and cover each vial with aluminum foil or a loosely placed clean cap. Cover the extra vials in the same manner and place all the vials in the desiccator.
3. Put the desiccator under continuous vacuum for 1 day.

Preparation and
Analysis of
Glycoconjugates

17.19.19

4. Weigh the vials and return them to the desiccator for another day under vacuum. Continue to do this every day until the change in weight from one day to the next is <1%.

This is normally achieved in 3 to 4 days.

5. Transfer a sample (e.g., 15 mg) of each standard to one of the empty, desiccated tared vials and determine the exact weight of the monosaccharides.
6. Add Milli-Q-purified water to each vial to give a 100 mM solution. Mix well to dissolve.
7. Prepare a dilution of each of these primary standard solutions by transferring 100 μ l of each solution to a 100-ml volumetric flask (100 pmol/ μ l). Dilute to 100 ml with Milli-Q-purified water and mix.
8. Divide each standard into 1-ml aliquots and store these at -20°C until needed.
9. Prepare a mixture of all the required monosaccharides by transferring 100 μ l of each primary standard solution to a 100-ml volumetric flask (100 pmol/ μ l). Bring up to 100 ml with Milli-Q-purified water and mix.
10. Divide the standard mixture into 1-ml aliquots and store these at -20°C until needed.
11. When required, defrost one aliquot of each desired standard solution and use them within 1 or 2 days. Do not refreeze the standards.

Standards prepared in this manner can be used directly for HPAEC or HPLC analysis or lyophilized and used for GLC analysis.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Acid-cleaned glass vials

Heat 1.5-ml glass vials (e.g., Reacti-Vials, Pierce) with Teflon-lined screw caps at 50°C for at least 3 hr each time in concentrated nitric acid and then in 6 M HCl. Rinse thoroughly with water, then with ethanol, and dry in an oven. When $<5\mu\text{g}$ of total carbohydrates is analyzed, it is recommended to silanize the glass vials by incubating them 15 min at room temperature with a 2% solution of dichlorodimethylsilane in toluene, decanting the solution, rinsing successively with methanol and hot distilled water, and allowing the vials to dry.

New vials should be at least rinsed with water, then with ethanol, and dried.

COMMENTARY

Background Information

High-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) and gas-liquid chromatography with gas-liquid chromatography with mass spectroscopic detection (GLC-MS) are the methods most widely used today for compositional analysis of the carbohydrate moieties of glycoproteins. Both methods are able to detect all the possible known monosaccharides. With the HPAEC-PAD system it is possible to analyze qualitatively and quantitatively a mixture of free monosaccharides containing as little as

50 pmol of each individual component. With GLC methods, the use of internal standards helps improve quantitation because the injection of reproducible proportions of the sample into the column is not easy.

HPAEC-PAD and other HPLC techniques

Different high-performance liquid chromatography (HPLC) systems have been used for the analysis of free glycoses obtained by acid hydrolysis, but sensitivity and resolution between individual monosaccharides were traditionally poor. The technique became popular

for analyzing glycoproteins and other glycoconjugates available in limited amounts from biological sources only after pellicular resin anion-exchange columns and the pulse-amperometric detector were developed. Neutral saccharides are analyzed at high pH (12 to 14), at which they are partially or completely ionized. Separation is achieved by adjusting the eluant pH. Anionic saccharides can be separated using sodium acetate/sodium hydroxide gradients. Therefore, with appropriate adjustment of the elution conditions, this system permits not only the compositional analysis of free monosaccharide mixtures but also the analysis of oligo- and polysaccharides (Lee, 1990).

Pulsed amperometric detection (PAD) utilizes a repeating sequence of three potentials applied for specific durations. The resulting total current is the sum of (1) the carbohydrate oxidation current, (2) the current due to the charging of the electrode surface, and (3) the current caused by the oxidation of the gold electrode. The current is only measured during the first applied potential and has been shown to be linear over $>4^{1/2}$ orders of magnitude. Recently, the parameters used for these three steps were optimized (LaCourse and Johnson, 1991). Molecules with similar size and structure are reported to have similar response factors. On the other hand, substituents on hydroxyl groups are expected to affect the sensitivity of detection. The detector works only at high pH. Thus, when base-labile species that are separated without the use of base in the eluant are analyzed, it is necessary to add alkali to the effluent before it enters the PAD cell.

Isocratic elution with sodium hydroxide (2 to 16 mM; the optimal concentration varies from column to column) allows for the separation of neutral sugars and hexosamines. Different sodium acetate gradients have been employed for the analysis of acidic monosaccharides and of mixtures of these and other sugars. The system is also useful for sialic acids, with a limit of detection of ~ 200 pmol (Manzi et al., 1990). Oligo- and polysaccharides can also be analyzed with this system using appropriate conditions (Hardy and Townsend, 1988; 1989). Although monosaccharide alditols are not retained by this column (they will come out in the void volume), oligosaccharide alditols can be analyzed (Reddy and Bush, 1991). Maintenance of a stable column temperature is critical to obtaining reproducible results; failure to do so can cause lack of separation of critical pairs of monosaccharides (Van Riel and Olieman, 1991).

Because only 1% of the sample is oxidized in the electrode, the system can be used preparatively without the need for a bypass. When analyzing base-labile sialic acids, this method cannot be used, because the postcolumn addition of base required for detection will cause saponification of the *O*-acetyl esters. However, the pellicular resin column can be used for fractionation and the eluate routed to a fraction collector. After the fractions are monitored by the TBA assay (UNIT 17.18), they can be pooled and desalted. This method has proved to be the best HPLC technique available for preparative fractionation of mixtures of sialic acids containing labile substituents (Manzi et al., 1990).

Another useful system for the analysis of free sialic acids is amine adsorption-ion suppression HPLC. This is based upon hydrogen bonding between the hydroxyl groups of saccharides and the amine moieties of the stationary phase (Mellis and Baenziger, 1981; Bergh et al., 1981). To fractionate anionic molecules, phosphate is added to the mobile phase to suppress the ionic effects while retaining hydrogen bonding (Mellis and Baenziger, 1983). Good separation of sialic acids is obtained in isocratic mode with a mixture of acetonitrile/water/0.25 M monobasic sodium phosphate at 1 ml/min (Diaz and Varki, 1985). It is necessary to maintain a minimum concentration of 10% phosphate buffer with a maximum working percentage of acetonitrile of 72% (above this the phosphate starts to precipitate). A 64:26:10 (v/v/v) ratio of acetonitrile/water/0.25 M sodium phosphate can be used to separate Neu5Ac and Neu5Gc. For *O*-acetylated sialic acids, however, it is necessary to increase the percentage of acetonitrile to a ratio of 72:18:10 (v/v/v) so as to increase the retention times of the sialic acids (Manzi et al., 1990). Chromatography is monitored by absorption at 200 nm, which requires extremely high-purity samples and reagents. The same analytical column can be used for preparative purposes (up to 400 nmol of sialic acids with a good resolution of the peaks). Recovery of sialic acids is $\sim 65\%$, but some loss and migration of *O*-acetyl groups occur.

GLC-FID

To analyze mixtures of glycoses or methylglycosides by gas-liquid chromatography, it is necessary to prepare volatile derivatives. Alditol acetates are the most widely used derivatives for monosaccharides obtained by acid hydrolysis, whereas trimethylsilylethers of methyl glycosides or methyl glycoside methyl esters are

the preferred derivatives for methanolysis products. These are the derivatives of choice for any glycoconjugate-derived sugar because of the amount of information that can be obtained from them (e.g., retention times in different columns). Other volatile compounds, such as aldononitrile acetates and trifluoroacetylated alditols, have also been successfully employed (as reviewed by Kammerling and Vliegenthart, 1989).

Acetylestere of alditols

When only neutral sugars need to be analyzed, the most common derivatization method is the conversion of the free monosaccharides into their alditol acetates. Each monosaccharide produces a single derivative, giving a simple chromatogram that can be easily interpreted and quantitated. The total analysis time is also shorter because all required standards can be processed as a mixture, requiring only one GLC run.

The procedure involves the reduction of the glycoses to alditols with sodium borohydride; elimination of the excess of hydride with acid; elimination of the boric acid produced as trimethylborate by coevaporation with acidified methanol; and peracetylation of all free hydroxyl groups. Different protocols to achieve this last step have been developed (Kammerling and Vliegenthart, 1989; Albersheim et al., 1967; York et al., 1985). It is possible to simply use the sodium acetate produced in the reaction as the catalyst, and achieve complete acetylation by heating with acetic anhydride. Because leftover borate could inhibit the acetylation, some protocols add pyridine to drive the reaction to completion.

Cleavage of glycosidic linkages by methanolysis

Cleavage of all glycosidic linkages by methanolysis is very effective and causes less destruction of the monosaccharides than acid hydrolysis. Monosaccharides are converted to their methyl glycosides and glycosyluronic acids to their methyl esters methyl glycosides. The *N*-acetyl group of amino sugars are completely cleaved by extensive methanolysis, and it is advisable to incorporate a re-*N*-acetylation step to avoid the production of additional peaks in the chromatogram. A drawback to the reacylation step is that some *O*-acetylation occurs (in some primary hydroxyl groups), but a short additional treatment with methanolic HCl eliminates these groups (Rickert and Sweeley,

1978). Because of anomerization, each monosaccharide produces several methylglycosides (α and β anomers of the pyranosidic and furanosidic forms), which yields a characteristic pattern of peaks in the GLC run. When the sample contains several monosaccharides, however, the pattern is quite complex, with some peaks superimposed, and this makes quantitation difficult. Nevertheless, by processing each of the individual monosaccharides that are expected in parallel, running each one separately under the same conditions, and determining the contribution of each peak to the total, an accurate quantitation can still be achieved.

Different conditions have been used for methanolysis of carbohydrates in glycoproteins using methanolic HCl, including 0.5 M for 16 hr at 65°C (Reinhold, 1972) and 0.75 M for 3 hr at 80°C (Wong et al., 1980). Conditions for the simultaneous release of neutral monosaccharides, acetamido sugars, uronic acids, octulosonic acids, and sialic acids (2 M TFA for 12 hr at 100°C under vacuum) have been reported (Clarke et al., 1991). Nevertheless, the cleavage of the uronic acid glycosidic linkage may be incomplete in some cases, and several glycosyluronic acids produce a certain percentage of 3,6-lactones. The glycosidic linkages of sialic acids are cleaved with good yield, but a significant proportion of the released molecules are destroyed. Although special conditions that minimize this destruction have been developed (Schauer, 1978), *N*-acyl groups and most *O*-acetyl esters are eliminated. Thus, release of intact sialic acids requires much milder acid hydrolysis of a separate aliquot of the sample.

Volatile derivatives of methyl glycosides

The content of neutral sugars, amino sugars, and glycosyluronic acids can be simultaneously determined by this method. Different reagents and procedures have been used to silylate mixtures of methyl glycosides (Kammerling and Vliegenthart, 1989). The most widely used procedure uses 5:1:5 (v/v/v) pyridine/hexamethyldisilazane/trimethylchlorosilane and an incubation that can range from 15 min at room temperature to 20 min at 80°C.

Sialic acids have also been analyzed as trimethylsilyl esters trimethylsilyl ethers or methyl esters trimethylsilyl ethers. Although conditions have been worked out for preparing these two types of derivatives while preserving labile natural substituents such as *O*-acyl groups (Schauer and Corfield, 1982), the derivatization of sialic acids is frequently in-

complete. Further problems are encountered during the subsequent chromatographic separation because the high temperatures required to elute these molecules result in partial destruction of the labile substituents. Therefore, this approach to sialic acid analysis has been replaced by HPLC methods with or without derivatization (see *UNIT 17.18*).

Qualitative and quantitative analysis

Analysis of volatile monosaccharide derivatives is achieved by gas-liquid chromatography. A combined adsorption/partition mechanism takes place between the volatile derivatives and the particular stationary phase that is a liquid at working temperatures. Through the large number of theoretical plates available in the length of the fused-silica capillary columns, the individual components of the mixtures under analysis are separated according to their particular structural characteristics. The mobile phase (helium) carries the molecules throughout the column and to the detector, where the ionization of the flame changes upon their passage, producing an electrical output. This signal is amplified and registered.

Initial characterization of individual peaks is done by comparison with standards run in the same conditions. The retention times of each standard relative to the component used as internal standard in the sample are calculated and compared with the relative retention times of each peak in the chromatogram of the sample. Typical elution profiles as well as relative retention times in the different columns under different conditions can be found in the literature (Kammerling and Vliegthart, 1989, and references therein).

A known amount of an internal standard (a monosaccharide or alditol that does not occur naturally) is added to each glycoconjugate at the beginning of its processing. Losses during the preparation of the sample for GLC will be proportionally the same for each component of the sample and for the internal standard. Thus, the ratio of analyte to internal standard will remain the same, even though the absolute amounts of both may be lower at the end. Quantitative analysis is performed by determining the detector responses for individual monosaccharides relative to the internal standard (known amounts are injected and the areas determined).

Critical Parameters

HPAEC-PAD and other HPLC techniques

When HPAEC-PAD is used for analysis of sialic acids, the sialic acids (which have a pK_a of ~ 2.0) bind to the resin at neutral pH, and it is necessary to use an anion to elute them. Samples are loaded in water to ensure binding and submitted to a pH gradient elution with dilute sodium acetate and acetic acid. This permits the separation of different *O*-acetylated sialic acid species, whose binding to the resin is relatively weaker with higher numbers of substituents. The relative detector responses vary significantly with the presence of different numbers of free hydroxyl groups. The response observed for Neu5Gc is 15% higher than that for Neu5Ac, which bears one fewer hydroxyl group. In contrast, the responses of the detector for Neu5-9Ac₂ and Neu5-7(8)-9Ac₃ are only 46% and 45% of the response for Neu5Ac, respectively. However, because of the possibility of partial de-*O*-acetylation between the point of addition of postcolumn alkali and the PAD detector, the detector response factors may vary somewhat in different laboratories. For preparative purposes, a maximum of 2 μ mol of total sialic acids can be injected without overloading the 9 mm \times 25 cm semipreparative column. After removal by Dowex 50 (H⁺) chromatography of the small amount of acetate required for separation using the HPAEC column, an excellent recovery (>90%) of acetylated species is possible.

When gradients that end in high concentrations of acetate (>100 mM) are used, the slow reequilibration of the column with low concentrations of hydroxide and acetate results in lower retention times and impaired separation. Resolution of some pairs of monosaccharides (e.g., Gal-GlcN and Xyl-Man) during the first isocratic step is extremely sensitive to minor changes in solvent composition. Therefore, extended equilibration periods between runs are necessary to obtain reproducible results. A shorter delay between runs, resulting in a column that is not fully equilibrated but that still achieves adequate separation for a set of monosaccharides, can produce repeatable results if an automatic sample injector with constant intervals between injections is employed.

When using amine adsorption-ion suppression chromatography, the relative proportion of solvents can be varied according to the kind of sialic acids to be separated. Because the eluant is monitored by UV absorbance at 200 nm, residual amounts of purification reagents pro-

Table 17.19A.7 Molecular Weights of Monosaccharides Found in Glycoproteins and Their Volatile Derivatives^a

Monosaccharide	Molecular weight of:				
	Aldoses	Alditols	Acetylated alditols	Methyl glycosides and methyl ester methyl glycosides	TMS ethers of methyl glycosides and methyl ester methyl glycosides
Hexoses	180	182	434	194	482
Pentoses	150	152	362	164	380
Deoxyhexoses	164	166	376	178	394
Hexosamines	179	181	431	193	481
<i>N</i> -acetylhexosamines	221	223	473	235	523
Hexosyluronic acids	194	NA	NA	222	438

^aNA, not applicable.

duce an absorption peak close to the void volume. The lower limit of reliable detection is 2 nmol, and the detector responses are equal for all sialic acids. For preparative purposes, the eluant is monitored by UV and the peaks collected directly on ice, diluted 10-fold with cold water, and purified by ion-exchange chromatography on Dowex 3-X4A (formate; see *UNIT 17.16*) at 4°C. The column is washed with 10 vol of 10 mM formic acid and eluted with 7 vol of 1 M formic acid, and sialic acids are recovered by lyophilization.

GLC with FID detection

When the assignment of some peaks in gas-liquid chromatography (a GLC) run remains uncertain, coinjection with the suspected monosaccharide allows confirmation.

When analyzing trimethylsilylated methyl glycosides, the sample is quantitated by choosing one nonsuperimposed peak for each monosaccharide and considering how much of the compound this area represents. Absolute molar content of monosaccharide in the sample can be calculated when a known amount of an internal standard was added to the sample before processing.

Because the linkage between *N*-acetylglucosamine and asparagine is barely cleaved by methanolysis, this fact has to be considered when quantitation is required. Dehydration of monosaccharides can also occur to a small extent.

Anticipated Results

When analyzing a neutral glycoconjugate (one that does not contain hexosamines, uronic acids, or sialic acids), total acid hydrolysis followed by HPAEC-PAD or reduction/peracetylation and GLC-FID will indicate the relative composition or absolute composition (when an

internal standard is added) of individual monosaccharides. When fucose is present, mild acid hydrolysis conditions can be tried to assure minimal destruction.

When hexosamines or uronic acids are present, acid hydrolysis followed by HPAEC-PAD or methanolysis followed by re-*N*-acetylation/pertrimethylsilylation and GLC-FID can be used. Optimal conditions for maximum release (for both types of sugars) and minimal destruction (for uronic acids) must be employed for both approaches.

When sialic acids are present, HPAEC-PAD or amine adsorption-ion suppression HPLC with UV detection can be employed for compositional analysis of the different species after mild acid hydrolysis or enzymatic release. Alternatively, released sialic acids can be derivatized with DMB and submitted to HPLC (see *UNIT 17.18*).

In some cases it is necessary to use more than one protocol to obtain the complete compositional analysis of sugars.

Analysis of monosaccharides using one of the HPLC-PAD methods described in Basic Protocols 1, 2, or 3 results in the successful separation of the mixture into constituent monosaccharides. Figure 17.19.3 and Figure 17.19.4. illustrate separation profiles obtained from a mixture of monosaccharids subjected to HPAC-PAD.

As mentioned before, a characteristic pattern is obtained for each monosaccharide in the case of trimethylsilylated methyl glycosides. By doing individual runs of each monosaccharide and including an internal standard, it is possible to calculate the percentage of the total monosaccharide that each peak represents. Generally a minimum of 100 ng of each component is required for confident detection. Mo-

lecular weights of the monosaccharides found in glycoproteins, their alditols and methyl glycosides, and their volatile derivatives used in the protocols included in this section are listed in Table 17.19A.7.

Time Considerations

The total time of analysis for each particular method depends on how many runs of HPLC or GLC are required per sample. This depends in turn on the complexity of the sample, as sometimes coinjection with known standards may be required.

Literature Cited

- Albersheim, P., Nevins, D.J., English, P.D., and Karr, A. 1967. A method for the analysis of sugars in plant cell-wall polysaccharides by gas-liquid chromatography. *Carbohydr. Res.* 5:340-345.
- Bergh, M.L., Koppen, P., and Van den Eijnden, D.H. 1981. High-performance liquid chromatography of sialic acid-containing oligosaccharides. *Carbohydr. Res.* 94:225-229.
- Clarke, A.J., Sarabia, V.S. Keenleyside, W., MacLachlan, P.R., and Whitfield, C. 1991. The compositional analysis of bacterial extracellular polysaccharides by high-performance-anion-exchange chromatography. *Anal. Biochem.* 199:68-74.
- Diaz, S. and Varki, A. 1985. Metabolic labeling of sialic acids in tissue culture cell lines: Methods to identify substituted and modified radioactive neuraminic acids. *Anal. Biochem.* 150:32-46.
- Hardy, M.R. and Townsend, R.R. 1988. Separation of positional isomers of oligosaccharides and glycopeptides by high-performance anion-exchange chromatography with pulsed amperometric detection. *Proc. Natl. Acad. Sci. U.S.A.* 85:3289-3293.
- Hardy, M.R. and Townsend, R.R. 1989. Separation of fucosylated oligosaccharides using high pH anion-exchange chromatography with pulsed amperometric detection. *Carbohydr. Res.* 188:1-7.
- Hardy, M.R. and Townsend, R.R. 1994. High-pH anion exchange chromatography of glycoprotein-derived carbohydrates. *Methods Enzymol.* 230:208-225.
- Kamerling, J.P. and Vliegthart, J.F.G. 1989. In *Clinical Biochemistry*, Vol. 1: Principles, Methods, Applications. Mass Spectrometry. (A.M. Lawson, ed.) pp. 175-263. Walter de Gruyter, Berlin.
- LaCourse, W.R. and Johnson, D.C. 1991. Optimization of waveforms for pulsed amperometric detection (p.a.d.) of carbohydrates following separation by liquid chromatography. *Carbohydr. Res.* 215:159-178.
- Lee, Y.C. 1990. High-performance ion-exchange chromatography for carbohydrate analysis. *Anal. Biochem.* 189:151-162.

- Manzi, A.E., Diaz, S., and Varki, A. 1990. High-pressure liquid chromatography of sialic acids on a pellicular resin anion-exchange column with pulsed amperometric detection: A comparison with six other systems. *Anal. Biochem.* 188:20-32.
- Mellis, S.J. and Baenziger, J.U. 1981. Separation of neutral oligosaccharides by high-performance liquid chromatography. *Anal. Biochem.* 114:276-280.
- Mellis, S.J. and Baenziger, J.U. 1983. Size fractionation of anionic oligosaccharides and glycopeptides by high-performance liquid chromatography. *Anal. Biochem.* 134:442-449.
- Reddy, G.P. and Bush, C.A. 1991. High-performance anion-exchange chromatography of neutral milk oligosaccharides and oligosaccharide alditols derived from mucin-type glycoproteins. *Anal. Biochem.* 198:278-284.
- Reinhold, V.N., 1972. Gas-liquid chromatographic analysis of constituents carbohydrates in glycoproteins. *Methods Enzymol.* (Part B) 25:244-260.
- Rickert., S.J. and Sweeley, C.C. 1978. Quantitative analysis of carbohydrate residues of glycoproteins and glycolipids by gas-liquid chromatography. An appraisal of experimental details. *J. Chromatogr.* 147:317-326.
- Schauer, R. 1978. Characterization of sialic acids. *Methods Enzymol.* 50:64-89.
- Schauer, R. and Corfield, A.P. 1982. In *Sialic Acids: Chemistry, Metabolism and Function* (R. Schauer, ed.) pp. 51-57. Springer-Verlag, New York.
- Van Riel, J. and Olieman, C. 1991. Selectivity control in the anion-exchange chromatographic separation of saccharides in dairy products using pulsed amperometric detection. *Carbohydr. Res.* 215:39-46.
- Wong, C.G., Sung, S.-S., J., and Sweeley, C.C. 1980. Analysis and structural characterization of amino sugars by gas-liquid chromatography and mass spectrometry. *Methods Carbohydr. Chem.* 8:5565.
- York, W.S., Darvill, A.G., McNeil, M., Stevenson, T.T., and Albersheim, P. 1985. Isolation and characterization of plant cell wall and cell wall components. *Methods Enzymol.* 118:3-54.

Key References

- Kamerling and Vliegthart, 1989. See above.
- Hardy and Townsend, 1994. See above.
- Manzi et al., 1990. See above.
- Descriptions of the principles of each method with examples of their application.*

Contributed by Adriana E. Manzi
University of California San Diego
La Jolla, California

**Preparation and
Analysis of
Glycoconjugates**

17.19.25

Composition of Labeled Monosaccharides from Glycosaminoglycans

BASIC PROTOCOL

Proteoglycans can be labeled metabolically with [³H]glucosamine (UNIT 17.4) to incorporate label into glucosamine and galactosamine. After isolation of proteoglycans (UNIT 17.2), glycosaminoglycans can be released from the proteoglycans by β-elimination (UNIT 17.15) and the distribution of labeled glucosamine and galactosamine in the isolated glycosaminoglycans can be determined. The glycosaminoglycans are hydrolyzed in acid until all the N-acetyl groups and N- and O-sulfates are removed. This procedure yields oligosaccharides in which some but not all of the uronic acid and amino sugar glycosidic bonds have been hydrolyzed. This partially hydrolyzed product is then treated with nitrous acid at pH 4 (UNIT 17.22A) to obtain a deaminative cleavage of the remaining amino sugar glycosidic bonds in which the reducing terminal of each of the resulting products is an anhydro-D-mannose (from D-GlcN) or an anhydro-D-talose (from D-GalN). The final depolymerization mixture, which contains disaccharides (uronosyl-anhydro sugars), free uronic acids, and free anhydro sugars, is reduced with NaBH₄ to convert the anhydromannose and androtalose residues to anhydromannitol and anhydrotalitol residues, respectively. The mixtures are then separated by paper chromatography and the amount of each labeled component is quantified by scintillation counting.

This procedure can also be used to determine the composition of glycosaminoglycans that have not been labeled metabolically. Samples of unlabeled glycosaminoglycans can be depolymerized in the same manner and the hydrolysis/deamination mixture can be reduced with NaB³H₄ to give stoichiometric ³H labeling of each reducing sugar component of the mixture.

The protocol has been described under the assumption that the user has a working knowledge of the chromatographic procedures required and of the use of radioisotopes in quantitative procedures.

Materials

Glycosaminoglycan sample (optionally, metabolically labeled with [³H]glucosamine; UNIT 17.4)
 20 μCi/ml [¹⁴C]glucose solution (>40 mCi/mmol; optional internal standard)
 3 M and 0.5 M H₂SO₄
 White mineral oil
 5.5 M NaNO₂
 1 M Na₂CO₃
 0.25 M NaBH₄ or 0.25 M ~500 mCi/mmol NaB³H₄, in 0.25 M NaOH
 0.1 M and 1 M NaOH
 Paper chromatography System 1 and System 2 (see recipes)
 Scintillation cocktail (see recipe)
 6 × 150-mm test tubes
 Sand baths: heating elements (Pierce) filled with sand, 99° and 50°C
 Hamilton syringe
 Whatman no. 3 chromatography paper (grade 3 Chr)
 Whatman cellulose phosphate chromatography paper (grade P 81)
 Paper chromatography jars for descending chromatography
 Scintillation counter suitable for dual-label counting

Composition of
Labeled
Monosaccharides
from
Glycosamino-
glycans

17.19.26

Contributed by H. Edward Conrad

Current Protocols in Molecular Biology (1995) 17.19.26-17.19.32

Copyright © 2000 by John Wiley & Sons, Inc.

Hydrolyze glycosaminoglycans with acid

1. Dissolve glycosaminoglycan sample in water and dialyze exhaustively against water to remove salts.

If glycosaminoglycan sample is already radiolabeled (UNIT 17.4), its specific activity must be sufficient to permit detection following the procedure. If unlabeled glycosaminoglycans are to be analyzed after labeling by NaB^3H_4 reduction, a glycosaminoglycan concentration of 1 to 10 mg/ml is desirable. More specifically, the glycosaminoglycan sample should yield 0.1 to 1000 μmol of each mono- or disaccharide to be assayed. This amount of material will incorporate $>10^4$ cpm into each product after reduction with NaB^3H_4 with a specific activity of ~ 500 mCi/mmol.

2. Transfer 40 μl of dissolved sample to a 6 \times 150–mm test tube, add 20 μl of 20 $\mu\text{Ci/ml}$ [^{14}C]glucose, and mix.

The [^{14}C]glucose is an optional addition that provides an internal standard. It can be replaced with 20 μl water if desired. The [^{14}C]glucose introduced should give $\sim 50,000$ cpm/ μl glycosaminoglycan solution and should have a specific activity high enough (>40 mCi/mmol) so that the amount of ^3H incorporated upon reduction with NaB^3H_4 will not be significant.

3. Add 12 μl of 3 M H_2SO_4 to the sample, cover the meniscus with a layer of white mineral oil to prevent evaporation of water (which would concentrate the acid during hydrolysis), and heat the mixture 6 hr at 99°C in a sand bath.

The sample should not be heated at a temperature that would cause the water (acid) to boil during the hydrolysis. Boiling will disrupt the mineral oil layer and allow water to evaporate, concentrating the acid.

Cleave with nitrous acid

4. Cool the hydrolysate. Remove 8 μl with a Hamilton syringe and place in a 6 \times 150–mm test tube, then add 20 μl of 5.5 M NaNO_2 (a large excess) to the sample. Allow deamination reaction to proceed 10 min at room temperature.

If a [^{14}C]glucose internal standard is used, the volume of sample taken should be reflected by the amount of ^{14}C in the aliquot, but the same ratio of sample to NaNO_2 solution must be used in order to bring the pH to 4 (the NaNO_2 neutralizes the acid). The sample pH, which may be measured with pH paper, should be ~ 4 during the deamination reaction.

Reduce with sodium borohydride

5. Add 8 μl of 1 M Na_2CO_3 (pH ~ 8.5) followed by 5 μl of 0.25 M NaBH_4 or NaB^3H_4 .

Unlabeled NaBH_4 is used if the sample has previously been metabolically radiolabeled, and NaB^3H_4 is used if the sample has not been labeled.

CAUTION: *All NaB^3H_4 manipulations (steps 5 to 9) must be carried out in a well-ventilated fume hood.*

6. Cork the reaction tube and heat 30 min at 50°C in a sand bath to reduce the aldehyde groups formed in the deamination reaction.
7. Cool the reduced sample to room temperature. Working in the fume hood, add sufficient 0.5 M H_2SO_4 (~ 20 μl) to the sample to reduce the pH to <3 (this destroys excess NaBH_4 or NaB^3H_4).
8. After ~ 30 min, dry the sample in a stream of air. Redissolve sample in water and dry again to remove residual H_2 or $^3\text{H}_2$ gas.
9. Dissolve the sample in 50 μl water and add 1 M NaOH until the solution become slightly alkaline (pH 8 to 10), as observed with pH paper.

Spot samples and separate by chromatography

10. Mark sheets of Whatman no. 3 and Whatman cellulose phosphate chromatography paper with pencil lines $\frac{1}{2}$ in. apart, then cut the sheets perpendicularly to the lines to create strips $1 \times \sim 22$ in. (2.5×57 cm).
11. Spot 5- to 10- μ l aliquots of each reduced sample from step 8 onto two strips of Whatman no. 3 paper and one strip of cellulose phosphate paper. Place the aliquots in the middle of the sixth $\frac{1}{2}$ -in. segment from one end of each strip; this end is the wick. Allow the spots to dry.

IMPORTANT NOTE: *No more than 5 to 10 μ l of the original hydrolysate should be analyzed on a chromatogram, because excess salt in the sample may interfere with the resolution of the components of the mixture. The borate remaining in the sample after the destruction of the NaBH_4 has no effect on the migration of the products.*

12. Fold the wick ends so that they will fit properly into a descending paper chromatography trough. Place the two Whatman no. 3 strips in chromatography jar containing paper chromatography System 1 reagent and allow them to develop for 15 and 40 hr, respectively. Place the cellulose phosphate strip in a jar containing paper chromatography System 2 reagent and allow it to develop for 10 to 12 hr.

Determine radioactivity by scintillation counting

13. Remove developed strips from solvents after the appropriate times have elapsed. Allow the strips to dry, then cut strips into $\frac{1}{2}$ -in. segments along the pencil markings made in step 10. Place the segments in separate scintillation vials.
14. Add scintillation cocktail to each scintillation vial to cover the paper segment. Determine the radioactivity of each segment by counting in a scintillation counter under optimal conditions for dual-label counting.
15. Plot the data for each strip as cpm vs. segment number.

Analyze data

16. If a [^{14}C]glucitol internal standard was used, compare the migration position of each peak with that of the standard to identify the ^3H -labeled peaks (see Table 17.19B.1).

^3H is counted with an efficiency of $\sim 5\%$. ^3H peaks that overlap with the [^{14}C]glucose peak must be corrected for the ^{14}C spillover into the ^3H channel.

Table 17.19B.1 shows the R_{glucitol} values (the distance that each reduced component on the chromatogram migrates relative to the distance that the [^{14}C]glucitol internal standard migrates) for the expected hydrolysis products in the two chromatography systems. The 15-hr run in System 1 gives a value for the fast-moving anhydromannose and anhydrotalose formed from glucosamine and galactosamine, respectively. The 40-hr run in System 1 separates unhydrolyzed disaccharides, while the [^{14}C]glucitol peak migrates to the end of the strip. Separation of glucuronic and iduronic acids, which migrate together in System 1, is obtained on the System 2 chromatogram.

17. Sum the ^3H cpm in each peak and normalize the result to the same number of [^{14}C]glucitol cpm (the value for the ^{14}C cpm used for the normalization may be chosen arbitrarily, but it must be the same for each paper strip used for the glycosaminoglycan sample being analyzed).
18. Calculate the percentage of each monosaccharide in the hydrolysate as [(total cpm for the monosaccharide in question)/(total cpm in all components)] $\times 100$. Calculate the relative amounts of each monosaccharide component from Table 17.19B.2.

Table 17.19B.1 R_{glucitol} Values for [³H]Alditols^a

Product ^b	Chromatography system	
	System 1	System 2
Gal _R (KS)	1.0	—
<i>Deamination products</i>		
AMan _R (GlcN)	2.0	—
ATal _R (GalN)	2.0	—
<i>Uronic acids</i>		
GlcA _R (Hep, HS, DS)	0.95	0.55
IdoA _R (Hep, HS, DS)	0.95	0.33
<i>Disaccharides</i>		
αIdoA1-4AMan _R (Hep and HS)	0.70	1.15
βGlcA1-4AMan _R (Hep and HS)	0.50	1.00
αIdoA1-3ATal _R (DS)	0.64	0.64
βGlcA1-3ATal _R (CS and DS)	0.45	0.48
βGlcA1-3AMan _R (HA)	—	1.00

^aAll values are for aldehyde-reduced sugars, as indicated by the subscript R. Abbreviations: AMan, anhydro-D-mannose; ATal, anhydro-D-talose; CS, chondroitin sulfate; DS, dermatan sulfate; Gal, D-galactose; GlcA, D-glucuronic acid; GlcN, D-glucosamine; Hep, heparin, HS; heparan sulfate; IdoA, L-iduronic acid; KS, keratan sulfate.

^bThe glycosaminoglycan from which each product is derived is given in parentheses. Standards for AMan_R and ATal_R can be obtained by treating GlcN and GalN, respectively, with nitrous acid at pH 4 and reducing the products with NaB³H₄ (Shively and Conrad, 1970). Standards for IdoA and disaccharides must be obtained by hydrolysis and nitrous acid deamination of heparin, chondroitin sulfate, dermatan sulfate or hyaluronic acid followed by NaB³H₄ reduction of the products (Conrad, 1980).

Table 17.19B.2 Cpm Equivalentents for Monosaccharides^a

Amount of monosaccharide	Sum of total normalized ³ H cpm in:
GlcA _R	GlcA _R + βGlcA1-4AMan _R + βGlcA1-3ATal _R + βGlcA1-3AMan _R
IdoA _R	IdoA _R + αIdoA1-4AMan _R + αIdoA1-3ATal _R
AMan _R	AMan _R + βGlcA1-4AMan _R + βGlcA1-3AMan _R
ATal _R	ATal _R + βGlcA1-3ATal _R + αIdoA1-3ATal _R

^aThe subscript R refers to the aldehyde-reduced form of each mono- or disaccharide. See Table 17.19.8 for abbreviations.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Paper chromatography System 1

18:3:1:4 (v/v/v/v) ethyl acetate/glacial acetic acid/88% formic acid/water.

System 1 is used with Whatman no. 3 chromatography paper.

Paper chromatography System 2

3:2:1 (v/v/v) ethyl acetate/pyridine/0.005 M boric acid.

System 2 is used with cellulose phosphate chromatography paper.

Scintillation cocktail

4 g diphenyloxazole in 1 liter toluene.

COMMENTARY

Background Information

With the exception of keratan sulfate, which contains alternating D-galactose and N-acetyl-D-glucosamine residues, all glycosaminoglycans are sequences of disaccharide units containing a uronic acid (D-glucuronic acid and/or L-iduronic acid) and an amino sugar (D-glucosamine or D-galactosamine). Disaccharides of chondroitin sulfate contain D-GalNAc and D-GlcA, whereas those from dermatan sulfate contain D-GalNAc and either D-GlcA or L-IdoA. Disaccharides from heparin and heparan sulfate contain D-GlcN, which may be either N-acetylated or N-sulfated, and either D-GlcA or L-IdoA. Thus, if one has identified a glycosaminoglycan that is free of other glycosaminoglycans, the primary purpose of monosaccharide analysis is to determine the relative proportions of the two uronic acid types. However, when glycosaminoglycans are isolated from natural sources, they are often recovered as mixtures of several types of glycosaminoglycans that are difficult to resolve completely. In such cases, monosaccharide analysis can often demonstrate the nature of the mixtures by showing the relative proportions of GlcNAc and GalNAc as well as the specific disaccharide unit in which they are found. Individual glycosaminoglycans in a mixture can be recognized by their susceptibility to cleavage by specific enzymes (UNIT 17.13B) or nitrous acid (UNIT 17.22A). In the present method, the glycosaminoglycans are cleaved by acid hydrolysis, which results in loss of the sulfates. Procedures for determination of glycosaminoglycan disaccharides that have retained their sulfate substituents are described in UNITS 17.13B & 17.22A.

Following acid hydrolysis and nitrous acid cleavage of unlabeled glycosaminoglycans, it would be equally feasible to analyze the products of this cleavage reaction sequence by gas chromatography or HPLC without use of NaB³H₄. Although such an analytical scheme has not been developed, it would represent a way to avoid use of radioisotopes. The paper chromatographic procedure used here does offer the possibility of analyzing many samples simultaneously, which cannot be done with the GC or HPLC approaches where samples must be analyzed sequentially. Furthermore, when unknown peaks occur in this analysis, it is clear that the relative amounts of the unknowns are reflected by the amounts of ³H in the peaks, because (1) ³H labeling of carbohydrates is

identical for all reducing carbohydrates when expressed on a molar basis and (2) carbohydrates are virtually the only naturally occurring compounds that can be reduced by NaBH₄ or NaB³H₄. Moreover, it is easy to elute labeled products from the paper segments after analysis for further characterization. This is readily accomplished by rinsing the segments in toluene (to remove the diphenyloxazole remaining from the scintillation cocktail), then eluting the dried segment with water or buffer.

Critical Parameters

The challenge in analyzing the monosaccharide composition of glycosaminoglycans is not in the quantitation of the monosaccharides but in the depolymerization step. The impossibility of obtaining complete hydrolysis of glycosaminoglycans to monosaccharides without decomposition of some of the monosaccharide residues cannot be emphasized too strongly. The labilities of the glycosidic bonds of N-acetylated amino sugars (Shively and Conrad, 1970) and L-iduronic acid (Conrad, 1980) are similar to those for neutral hexoses, but the well-known stability of glucuronic acid glycosides virtually precludes the complete hydrolysis of these bonds without destruction of these residues as well as the other monosaccharide components of the glycosaminoglycan samples. Furthermore, hydrolytic release of the N-acetyl groups of the amino sugars occurs at a rate similar to that for the GlcNAc or GalNAc glycosides, so that many of the acetyl groups are released prior to complete hydrolysis of the amino sugar glycosidic bonds. This precludes complete hydrolysis of the amino sugar glycosides, as the glycosidic bonds of the N-unsubstituted GlcN and GalN are extremely stable to acid (Shively and Conrad, 1970). For the N-sulfated GlcN residues in heparin and heparan sulfate, this problem is particularly serious, as the N-sulfate substituents are very labile and are removed during the first few minutes of hydrolysis.

The tack taken here is to circumvent these hydrolysis problems by taking advantage of the elimination reaction that occurs when N-unsubstituted amino sugars are treated with nitrous acid (in spite of the misstatements often seen in the literature, this is not a hydrolysis reaction), thus cleaving these amino sugar glycosides under extremely mild conditions. This approach requires only that hydrolysis be continued long enough for complete removal of N-sulfate and N-acetyl groups. This avoids the

destruction of monosaccharide residues that occurs when forcing conditions of hydrolysis are used and allows almost stoichiometric recovery of these glycosaminoglycans as mono- and disaccharides, which can be analyzed together as described. Identical results for the analytical procedure described here are obtained whether the acid hydrolysis is carried out for 4, 6, or 8 hr before adding the nitrous acid reagent (Conrad, 1980). This is a reflection of the fact that all N- and O-sulfate groups and all N-acetyl groups will have been completely removed after 4 hr of hydrolysis. During the subsequent period (from the fourth to the eighth hour) there is some further hydrolysis, which reduces the yields of disaccharides and increases the yield of monosaccharides in the assay mixture, but this does not change the final analytical results.

For accurate results, it is advisable to first separate the glycosaminoglycan chains from the core protein and to isolate the free glycosaminoglycan chains for the analysis (UNIT 17.15). This avoids the possibility of loss of monosaccharides due to the Browning reaction, which occurs when carbohydrates and amino acids are heated together in acid.

The nitrous acid reaction is virtually stoichiometric for β -linked amino sugars (Shively and Conrad, 1970). Consequently, disaccharides are obtained in stoichiometric yields from glycosaminoglycans that contain β -linked amino sugars (chondroitin sulfate, dermatan sulfate, hyaluronic acid, and keratan sulfate). However, a side reaction, the ring contraction reaction, occurs at a frequency of 10% to 15% for α -linked amino sugars such as those found in heparin and heparan sulfate (Shively and Conrad, 1970; 1976). The ring contraction reaction results in deamination without bond cleavage and may convert a heparin tetrasaccharide into a new tetrasaccharide instead of two disaccharides. Thus, heparin and heparan sulfate yield significant amounts of ring contraction tetrasaccharides (Bienkowski and Conrad, 1985). However, there seems to be little structural selectivity for the ring contraction: the ring contraction tetrasaccharides contain disaccharide units in the same proportions as are found in the original glycosaminoglycan. Thus, even for heparin and heparan sulfate, the disaccharides formed during nitrous acid cleavage are obtained in the same proportions that are found in the original glycosaminoglycan. Consequently, when yields are expressed in percent of total uronic acids or percent of total amino

sugars, the ring contraction reaction does not interfere with the quantitation, and the yields of di- and monosaccharides are truly representative of their proportions in the starting glycosaminoglycan. Some of these heparin ring contraction tetrasaccharides have been identified (UNIT 17.22B).

Troubleshooting

In order to obtain reproducible migration of peaks, it is necessary to have chromatography tanks well-equilibrated with freshly prepared solvents. If many strips are placed in the tank, it may be necessary to add more chromatography solvent after the first day of the 40-hr development.

Anticipated Results

The results obtained here yield the distribution of labeled monosaccharides in metabolically labeled starting material or the monosaccharide compositions of the unlabeled glycosaminoglycan starting material. This is possible even though complete hydrolysis is not achieved, because one can separate and quantify disaccharides of known composition.

Time Considerations

The analytical procedure can be completed in 3 days, with 1 day for the depolymerization and NaB^3H_4 reduction, two overnight periods for paper chromatography, and finally scintillation counting and calculations. The procedure can be interrupted at any stage. A major advantage of the paper chromatography procedure over HPLC or GC methods is that many samples can be analyzed simultaneously. The only sequential phase of the analysis is scintillation counting, but the conditions described are designed to yield the mono- and disaccharides in sufficient yields (i.e., total ^3H cpm in each peak) so that 30-sec counting times for each paper segment result in good counting statistics.

Literature Cited

- Bienkowski, M.J., and Conrad, H.E. 1985. Structural characterization of the oligosaccharides formed by depolymerization of heparin with nitrous acid. *J. Biol. Chem.* 260:356-365.
- Conrad, H.E. 1980. The acid lability of the glycosidic bonds of L-iduronic acid residues in glycosaminoglycans. *Biochem. J.* 191:355-363.
- Shively, J.E., and Conrad, H.E. 1970. Stoichiometry of the nitrous acid deaminative cleavage of model amino sugar glycosides and glycosaminoglycuronans. *Biochemistry* 9:33-41.

Shively, J.E., and Conrad, H.E. 1976. Formation of anhydrosugars in the chemical depolymerization of heparin. *Biochemistry* 15:3932-3942.

Describes the basic procedure used here.

Key Reference

Conrad, 1980. See above.

Contributed by H. Edward Conrad
University of Illinois
Urbana, Illinois

Analysis of Oligosaccharide Negative Charge by Anion-Exchange Chromatography

UNIT 17.20

This unit presents the analysis of negative charge on labeled N- or O-linked oligosaccharides. These protocols may be used in the initial screening of oligosaccharides to detect negative charge, for analytical or preparative separation of oligosaccharides based on their negative charge, or to analyze the type of negative charge found on the oligosaccharides. The Basic Protocol describes the use of anion-exchange (QAE-Sephadex) chromatography with stepwise elution for estimating the number of negative charges on an oligosaccharide sample derived from glycosidase treatment of a glycoprotein. In the Alternate Protocol, gradient elution is used for the preparative separation of oligosaccharides based on negative charge. The Support Protocol describes a method for measuring loss of or change in negative charge after treatment of the oligosaccharide sample with mild acid and/or phosphatases.

SEPARATION AND ANALYSIS OF ANIONIC OLIGOSACCHARIDES BY QAE-SEPHADEX CHROMATOGRAPHY WITH STEPWISE ELUTION

**BASIC
PROTOCOL**

Negative charges on peptide-free radioactive N- or O-linked oligosaccharides can be detected by binding the molecules to the strong anion exchanger QAE-Sephadex (see also UNIT 10.10). Stepwise batch elution (as described in this protocol) or gradient elution (see Alternate Protocol) with salt separates the bound molecules according to the approximate number of negative charges. Loss or change of negative charge following treatment with sialidases (UNIT 17.12), solvolysis (UNIT 17.23), mild acid (UNIT 17.16), or phosphatase (see Support Protocol) can also be monitored (see Fig. 17.20.1). Because of charge-to-mass effects, this method gives only a general estimate of the number of negative charges on an oligosaccharide. If more precise analysis is desired, protocols using high-performance liquid chromatography (HPLC; UNIT 17.21) are recommended.

Materials

Radiolabeled mixture of oligosaccharides released from glycoprotein
(UNITS 17.12-17.17)

Equilibrated QAE-Sephadex chromatography resin (see recipe)

2 mM Tris base

Elution buffers (see recipe)

1- to 2-ml Pasteur pipets plugged with glass wool *or* 1- to 2-ml disposable plastic columns

Sintered-glass funnel

Additional reagents and equipment for metabolic radiolabeling (UNIT 17.4) and autoradiography (APPENDIX 3A)

1. Pour 0.75-ml column of equilibrated QAE-Sephadex in a 1- to 2-ml Pasteur pipet plugged with glass wool or in a 1- to 2-ml disposable plastic column. Wash with 5 ml of 2 mM Tris base.

It is acceptable to let the top of the column bed go dry during washing.

2. Dilute or dissolve the sample in 0.75 ml of 2 mM Tris base and load onto the column. Wash with seven 0.75-ml aliquots of 2 mM Tris base and collect 1.5-ml fractions.

Try not to disturb the top of the column bed during application. Allow the top of the column bed to go dry between aliquots.

The concentration of salts in the applied sample should be <5 mM and the pH must be >7. If the original sample contains acidic buffers or other salts, it may be necessary

**Preparation and
Analysis of
Glycoconjugates**

17.20.1

Contributed by Ajit Varki

Current Protocols in Molecular Biology (1995) 17.20.1-17.20.7

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 32

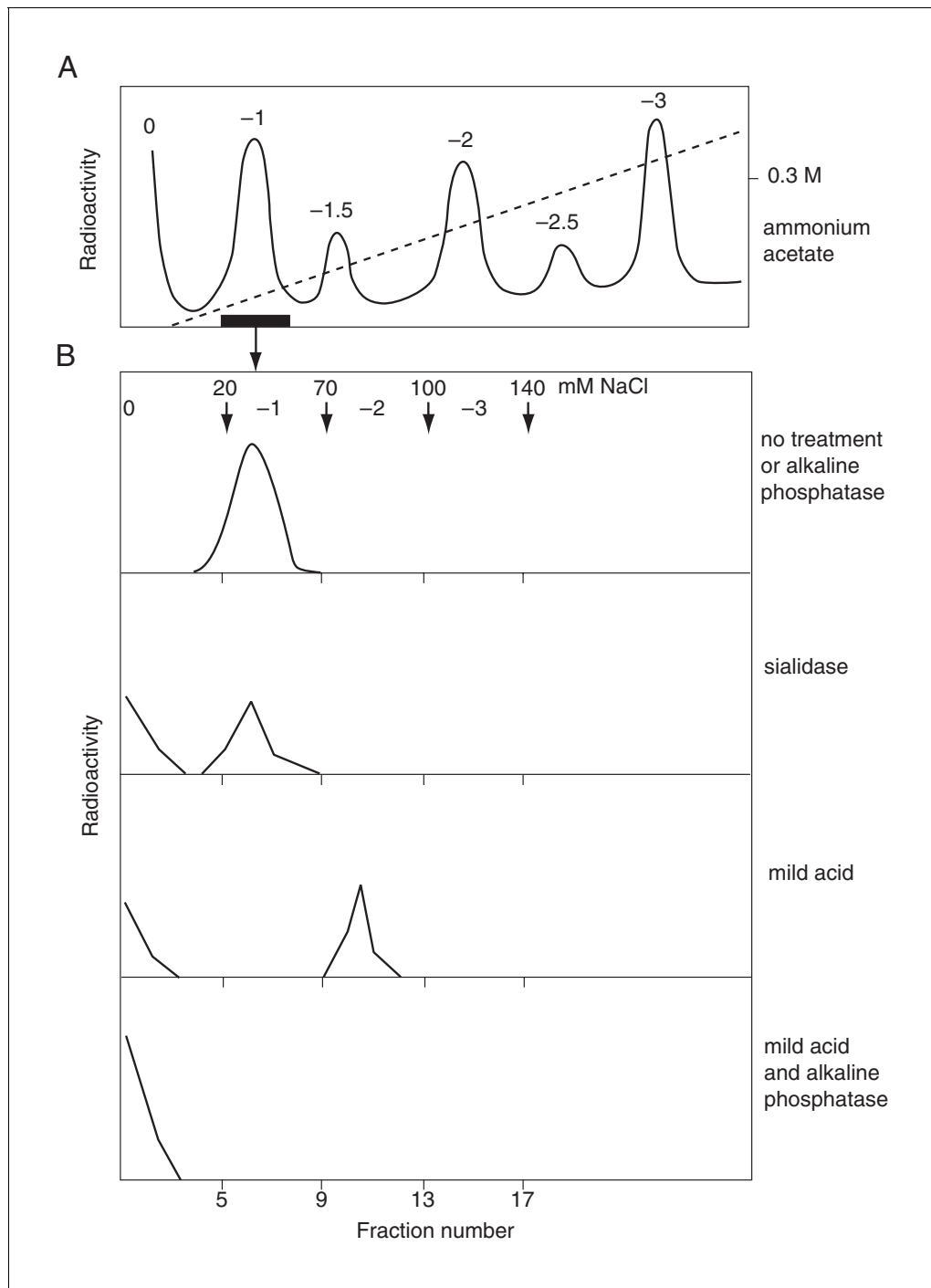


Figure 17.20.1 Schematic examples of QAE-Sephadex analysis (based on Varki and Kornfeld, 1983). Numbers in italics indicate the charges on the high mannose-type oligosaccharides eluting at that position. **(A)** Gradient fractionation of mixture of N-linked oligosaccharides with different combinations of sialic acid phosphomonoesters or phosphodiester. **(B)** Peak marked “-1” consists of a mixture of oligosaccharides with one negative charge due to either one sialic acid residue (acid- or sialidase-sensitive) or one phosphodiester (increased negative charge after mild acid, sensitive to alkaline phosphatase only after mild acid treatment). The procedures used to evaluate oligosaccharide mixtures are listed on the right with examples of resulting shifts in the elution peaks for each type of treatment illustrated on the left.

to neutralize with Tris base and/or dilute the sample to >0.75 ml to reduce the salt concentration. In these cases, the first wash-through tube will contain >1.5 ml fluid.

3. Elute sample from the column with elution buffers containing increasing concentrations of salt (i.e., 20, 70, 100, 140, 200, 400, and 1000 mM NaCl) in 2 mM Tris base. For each salt concentration, use eight 0.75-ml aliquots and collect four 1.5-ml fractions.

The salt elution series suggested above was devised for analysis of phosphorylated high mannose-type oligosaccharides from lysosomal enzymes. Salt concentrations may need to be adjusted to optimize the separation of other types of oligosaccharides (see Critical Parameters).

4. Monitor elution of the oligosaccharides by analyzing aliquots of the collected fractions (or the entire fractions) for radioactivity. In the latter case, collect the eluate directly into large-sized scintillation vials. If the first wash fraction is >1.5 ml (see step 2 annotation), monitor only an aliquot.

When comparing different treatments of the same oligosaccharide—e.g., sialidase (UNIT 17.12), mild acid, and alkaline phosphatase (see Support Protocol)—it is convenient to run all samples and controls in parallel. Approximately ten columns can be easily run at the same time. A repeating pipettor makes the process more reproducible and less taxing.

- 5a. The first time the experiment is done, and anytime it is possible that some of the oligosaccharides may not elute with the highest concentration of salt, add 1.5 ml of 2 mM Tris base to the column and transfer the contents of the column into a scintillation vial. Add scintillation fluid and monitor for any remaining radioactivity.
- 5b. Otherwise, discard the column contents (do not reuse).

SEPARATION AND ANALYSIS OF ANIONIC OLIGOSACCHARIDES BY QAE-SEPHADEX CHROMATOGRAPHY WITH GRADIENT ELUTION

**ALTERNATE
PROTOCOL**

After the initial studies of a mixture of oligosaccharides on a QAE-Sephadex column, it may be desirable to preparatively separate the oligosaccharides with better resolution. Although HPLC methods are preferable (see UNIT 17.21), the following protocol can be used if an HPLC unit or column is not available or if a large capacity is needed.

Additional Materials (also see Basic Protocol)

5- to 20-ml disposable plastic column
Gradient mixer

Additional materials and equipment for salt gradient preparation (UNIT 10.10)

1. Prepare a 3- to 10-ml column of QAE-Sephadex (see Basic Protocol, step 1, but use 5- to 20-ml disposable plastic column). Wash column with 10 column volumes of 2 mM Tris base.
2. Prepare and load the oligosaccharide sample (see Basic Protocol, step 2).
3. Place 50 ml of 2 mM Tris base in the first chamber of the gradient mixer; place 50 ml of an appropriate concentration of NaCl dissolved in 2 mM Tris base in the second chamber (choose salt concentration based on the results of the Basic Protocol).

An ammonium acetate gradient at pH 5.3 can be used to fractionate phosphomonoesters and phosphodiester (see Critical Parameters).

The concentration of NaCl or ammonium acetate in the second chamber is selected based on prior knowledge of the concentration required to elute the most anionic species in the mixture.

**Preparation and
Analysis of
Glycoconjugates**

17.20.3

**SUPPORT
PROTOCOL**

4. Elute the sample from the column with the selected linear gradient regimen. Collect 1-ml fractions and monitor for radioactivity.

**DETECTION AND REMOVAL OF PHOSPHODIESTERS OR
PHOSPHOMONOESTERS**

Oligosaccharide samples thought to contain phosphodiester and phosphomonoesters (e.g., released from lysosomal enzymes) can be treated with mild acid and alkaline phosphatase to detect and remove them. Phosphomonoester-containing molecules will lose negative charge upon phosphatase treatment. Phosphodiester-containing molecules are resistant to alkaline phosphatase alone, but will increase in negative charge following mild acid treatment because of the generation of phosphomonoesters. A combination of mild acid and phosphatase treatment will neutralize phosphodiester; the mild acid conditions used here will also result in removal of sialic acids (see Fig. 17.20.1).

Additional Materials (also see *Basic Protocol*)

Radiolabeled mixture of oligosaccharides released from glycoprotein (*UNITS* 17.12-17.17), desalted and lyophilized
10 U/ml *E. coli* alkaline phosphatase
2 M acetic acid
200 mM Tris·Cl, pH 8.0 (*APPENDIX 2*)

Strong 10-ml conical glass tubes *or* 1-ml Reacti-Vials (Pierce)
Water bath or heating block 80°C

1. Dissolve the oligosaccharide sample in 0.5 ml of 2 M acetic acid in a strong 10-ml conical glass tube or 1-ml Reacti-Vial. Tightly cap or stopper the tube and heat 120 min at 80°C.

This step is designed to detect phosphodiester. A second, identical sample should be prepared and processed starting with step 3; if phosphodiester are present, the acid-treated sample will show a higher negative charge than this non-acid-treated control upon analysis.

2. Flash-freeze and lyophilize the sample.
3. Dissolve in 20 µl water, using a pipettor to wash the walls of the tube. Vortex well and centrifuge briefly at room temperature to get sample to bottom of tube.
4. Remove a 10-µl aliquot of sample and dilute to 1.5 ml with 2 mM Tris base. Apply to a 0.75-ml QAE-Sephadex column for analysis (see *Basic Protocol*).
5. To the remaining sample, add 10 µl of 200 mM Tris·Cl (pH 8.0) and mix. Add 1 µl of 10 U/ml *E. coli* alkaline phosphatase (10 mU). Incubate 1 hr at 37°C.
6. Dilute sample to 1.5 ml with water and apply to a 0.75-ml QAE-Sephadex column (see *Basic Protocol*).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Ammonium acetate (pH 5.3), 1 M

Dissolve 38.54 g ammonium acetate in H₂O. Add H₂O to 500 ml. Adjust pH to 5.3 with 1 M acetic acid. Store at 4°C.

Elution buffers

Prepare stock solutions of 1 M Tris base (pH unadjusted; store at 4°C) and 2 M NaCl (store at room temperature). Use these stocks to make elution buffers containing 2 mM Tris base and different concentrations of NaCl. For gradient elution, make appropriate dilutions of the 1 M ammonium acetate (pH 5.3) stock solution (see recipe). It is not necessary to check the final pH of the diluted solutions.

QAE-Sephadex chromatography resin, equilibrated

Swell QAE-Sephadex (Pharmacia Biotech) in 1 M Tris base (using 1 g dry resin/20 ml Tris) overnight at 4°C, then decant. Repeat once. Pour slurry over a sintered-glass funnel and wash with >20 vol of 2 mM Tris base. Check the pH of the final wash with pH paper to make sure it is similar to that of the Tris base; if not, wash again until appropriate pH is achieved. Store equilibrated resin in 2 mM Tris base at 4°C.

There is no need to remove “fines” (small particles) from the resin because the columns used are small and flow rates are not limiting. With storage and exposure to air, the pH of the Tris base will drift towards neutral; however, this is not of serious concern.

COMMENTARY

Background Information

Negatively charged monosaccharides and highly anionic oligosaccharides bind well to most anion-exchange resins. However, large oligosaccharides with very few negative charges have a poor charge-to-mass ratio and may fail to bind to some anion-exchange resins. The strong anion exchanger QAE-Sephadex was first used (Tabas and Kornfeld, 1980) to permit analysis of high mannose-type N-linked oligosaccharides with single phosphodiester. The column was equilibrated in 2 mM Tris base to ensure that the pH was above neutral (maximizing the deprotonation of all anions) and to limit the amount of salt competing for binding sites. The protocol described here is based upon modifications of this method (Varki and Kornfeld, 1980, 1983; Goldberg and Kornfeld, 1981).

In the time since the original description, modifications have been developed to separate a wide variety of anionic oligosaccharides, including O-linked oligosaccharides, N-linked glycopeptides, and sulfated oligosaccharides (Roux et al., 1988; Fukuda, 1989; Cummings et al., 1989). In each case, column sizes and elution conditions were adjusted to allow appropriate separations and yield. In recent years, Mono-Q FPLC columns (Pharmacia Biotech) have permitted the use of the same QAE exchanger group for more rigorous and complete separations (van Pelt et al., 1987; *UNIT 10.10*).

Critical Parameters

The QAE-Sephadex resin must be thoroughly equilibrated in 2 mM Tris base. The initial loading must be carried out in a minimum

amount of salt (<2.5 mM) to ensure that weakly charged molecules do not escape binding. This can be achieved by desalting samples thoroughly (*UNIT 10.9*) and/or diluting them sufficiently with water or 2 mM Tris base before loading. To make the dilution volume reasonable, the preceding steps (e.g., enzyme reactions) should be carried out in the minimum volumes and lowest salt concentrations possible. Acid treatments are carried out with volatile acids that can be lyophilized before analysis. The charge-to-mass ratio of an oligosaccharide affects the amount of NaCl required for elution. Thus, a high mannose-type N-linked oligosaccharide with a single mannose-6-phosphate residue will elute with 50 to 70 mM NaCl, while free mannose-6-phosphate (with the same anionic group) requires 100 to 120 mM NaCl. The amount of salt needed to elute a particular charge or size class of molecules must be determined empirically. Thus, this technique is most useful when a related series of oligosaccharides with a limited range of size and charge (e.g., phosphorylated N-linked oligosaccharides with one to four negative charges) are being separated and analyzed. However, even when there is very extensive heterogeneity in size and charge (e.g., sulfated sialylated N-linked oligosaccharides), information can be obtained by following relative shifts in the elution position of groups of molecules (Roux et al., 1988). Gradient elution gives more defined separation between individual types of molecules, and is used primarily for preparative separations. If the molecules contain mixtures of phosphomonoesters and phosphodiester, separation can be obtained with a

gradient of ammonium acetate adjusted to pH 5.3 (close to the pK_a of the second negative charge of a phosphomonoester).

Analysis of glycopeptides by QAE-Sephadex chromatography is less satisfactory, because the peptide portion of the molecules can carry variable amounts of negative and/or positive charge. However, if complete proteolytic cleavage is performed (see UNIT 17.14A), each glycopeptide should have only one carboxyl group and separations can be made on the basis of the additional negative charges on the oligosaccharide. Molecules with very high charge density (e.g., heparin chains) are not well resolved on QAE columns because they are difficult to elute.

Troubleshooting

If molecules known to be anionic do not bind to the QAE-Sephadex column, the most likely problems are failure to properly equilibrate the column in 2 mM Tris base and the presence of excessive salt in the sample. If peaks are broad and slurred, the most likely problem is excessive disturbance of the top of the column bed during the application of elution buffers.

Anticipated Results

The method presented in this unit can be used for the following analyses.

1. *Initial screening of oligosaccharides for the presence of negative charge.* The sample is applied and the column washed and then directly eluted with 1 M NaCl in 2 mM Tris base. As a control, a known neutral oligosaccharide and an anionic oligosaccharide should be run under exactly the same conditions to obtain background numbers. A small quantity of the anionic oligosaccharide (<5%) can be expected to escape binding to the column, and a small amount of label from the neutral molecule (<2%) will be nonspecifically found in the salt eluate.

2. *Analytical separation of oligosaccharides by negative charge.* A mixture of oligosaccharides is applied and sequentially eluted batchwise with a defined series of salt elution steps (e.g., 20, 70, 100, 140, 200, 400, and 1000 mM NaCl), and the profile is monitored.

3. *Preparative separation of oligosaccharides by negative charge.* This is based upon the profile obtained by analytical separation. Gradient elution using NaCl or ammonium acetate may be preferred. Peaks are pooled, dried, and desalted on columns of Sephadex G-25 in water.

4. *Analysis of the type of negative charge on oligosaccharides.* Individual peaks obtained by preparative separation are analyzed before and after treatment or combinations of treatments with mild acid (see Support Protocol), with sialidase (UNIT 17.12), by solvolysis (UNIT 17.23), and/or with alkaline phosphatase (see Support Protocol). Fig. 17.20.1 schematically depicts an example of such analyses.

Time Considerations

The initial setup of small columns should take 10 to 20 min. The setup for gradient elution takes somewhat longer. Stepwise elutions of approximately ten samples can usually be carried out in parallel in ~1 hr. If sample preparation for scintillation counting is performed during time gaps in the elution procedure, the entire process can be completed in 2 to 3 hr. The time for scintillation counting will depend upon the amount of radioactivity available in each sample.

Literature Cited

- Cummings, R.D., Merkle, R.K., and Stults, N.L. 1989. Separation and analysis of glycoprotein oligosaccharides. *Methods Cell Biol.* 32:141-183.
- Fukuda, M. 1989. Characterization of O-linked saccharides from cell surface glycoproteins. *Methods Enzymol.* 179:17-29.
- Goldberg, D.E. and Kornfeld, S. 1981. The phosphorylation of β -glucuronidase oligosaccharides in mouse P388D1 cells. *J. Biol. Chem.* 256:13060-13067.
- Roux, L., Holoyda, S., Sundblad, G., Freeze, H.H., and Varki, A. 1988. Sulfated N-linked oligosaccharides in mammalian cells I: Complex-type chains with sialic acids and O-sulfate esters. *J. Biol. Chem.* 236:8879-8889.
- Tabas, I. and Kornfeld, S. 1980. Biosynthetic intermediates of β -glucuronidase contain high mannose oligosaccharides with blocked phosphate residues. *J. Biol. Chem.* 255:6633-6639.
- van Pelt, J., Damm, J.B., Kamerling, J.P., and Vliegthart, J.F.G. 1987. Separation of sialyl-oligosaccharides by medium pressure anion-exchange chromatography on Mono Q. *Carbohydr. Res.* 169:43-51.
- Varki, A. and Kornfeld, S. 1980. Structural studies of phosphorylated high mannose-type oligosaccharides. *J. Biol. Chem.* 255:10847-10858.
- Varki, A. and Kornfeld, S. 1983. The spectrum of anionic oligosaccharides released by endo- β -N-acetylglucosaminidase H from glycoproteins. Structural studies and interactions with the phosphomannosyl receptor. *J. Biol. Chem.* 258:2808-2818.

Key Reference

Varki, A. and Kornfeld, S. 1980. See above.

Figures in the miniprint section of this paper provide several examples of the types of analyses described here.

Contributed by Ajit Varki
University of California San Diego
La Jolla, California

HPLC Methods for the Fractionation and Analysis of Negatively Charged Oligosaccharides and Gangliosides

UNIT 17.21A

This unit describes the fractionation and analysis of anionic oligosaccharides (Basic Protocol 1) and gangliosides (Basic Protocol 2) using anion-exchange high-performance liquid chromatography (HPLC). The sample is loaded onto the column in a low-ionic-strength solvent. The anionic saccharides or gangliosides are retained by the positively charged matrix by ionic interactions; the more negative charge on the compound, the more tightly it binds to the column. Neutral molecules pass through the column unretained. Elution is effected by a gradient of increasing ionic strength that progressively disrupts the ionic interactions between the analyte and the column. Saccharides or gangliosides are eluted in order of the number of negative charges they possess, although the charge-to-mass ratio can also contribute to elution position. Hence, a molecule with a high charge-to-mass ratio may elute later than one with the same net charge but a smaller charge-to-mass ratio. Those unfamiliar with the basic principles of HPLC are referred to *UNITS 10.12-10.14* or to Snyder et al. (1988).

FRACTIONATION OF RELEASED N-LINKED OLIGOSACCHARIDES ON DEAE HPLC COLUMNS

**BASIC
PROTOCOL 1**

The use of DEAE HPLC columns to separate sialylated N-linked oligosaccharides is described. The DEAE column is equilibrated in water and eluted with a gradient of increasing NaCl concentration. Sialic acids are detected based on their radioactivity (if previously radiolabeled) or by the TBA assay for sialic acids (if unlabeled) and comparison with known standards.

Materials

HPLC-grade or high-quality deionized water (e.g., 18 M Ω Milli-Q type)

0.5 M sodium chloride (NaCl) in HPLC-grade water (see recipe)

Oligosaccharide sample (Peptide N:glycosidase F–released and desalted;
UNIT 17.13A)

Sialylated oligosaccharide standards (e.g., Dionex or Oxford Glycosystems)

HPLC apparatus capable of two-component gradient formation

DEAE HPLC column (TosoHaas TSK-gel DEAE-2SW, 25-cm \times 4.6-mm i.d.,
5-mm particle size, or equivalent)

Additional reagents and equipment for desalting sample (*UNITS 17.13A & 17.14A*) and
(if analyzing unlabeled sample) TBA assay (*UNIT 17.18*)

1. Before using any column, wash the column thoroughly. First, wash with at least 10 column volumes of water. Next, wash the column with a linear gradient from 0 to 0.5 M NaCl over 100 min, and then from 0.5 M NaCl to 100% water over 100 min. Use a flow rate of 0.4 to 0.6 ml/min for all washes.

The water and all buffers used should be prefiltered through a 0.45- μ m filter and degassed. Column prefilters and/or guard columns provide additional protection for the HPLC column and increase its productive lifetime.

A flow rate of 0.4 to 0.6 ml/min is recommended for all column washings and elutions. Do not exceed the manufacturer's recommendations for either the flow rate or maximum pressure.

**Preparation and
Analysis of
Glycoconjugates**

17.21.1

2. Thoroughly clean the injection loop with water.
3. Equilibrate the column at room temperature with 100% water at a flow rate of 0.6 ml/min. Wash with at least 10 column volumes of water.
4. Check the baseline obtained with the detection method to be used for the sample. Do not proceed until a stable baseline is obtained.

For analytical work, on-line radioactivity detectors or UV detectors may be used. For UV detection, absorbance at 190 to 210 nm is measured. Because many compounds absorb at these wavelengths, this method of detection demands high purity of the sample. Alternatively, individual fractions may be collected and analyzed by one of several methods (see step 8). Collect a portion of the eluate and establish that the solvents do not interfere with the detection method chosen and that the baseline is stable.

5. Set up a program from the pump control unit for washing unbound material, eluting the bound species, and bringing the column back to the starting conditions, as indicated in Table 17.21A.1.

The NaCl gradient should be shallow for two reasons. First, the difference in NaCl concentration needed to elute oligosaccharides differing by a single negative charge is ~20 mM NaCl, depending on the column. Second, the DEAE resin will shrink and swell with increasing and decreasing NaCl concentrations, respectively. Therefore, rapid changes in ionic strength may damage the column and degrade its performance.

6. Make a “sham” run (injecting water in place of sample), monitoring the eluant, if an on-line detection system is used. If fractions are collected and individually analyzed, collect them every 1 min. Load the water into the injection loop using a syringe with a flat-tip needle compatible with the HPLC injector. Inject the water onto the column and start the gradient program.

If contaminant peaks are observed (peaks >5% full scale when using the detector settings required for the sample), the column must be washed until the peaks disappear. (When using UV detection, a peak will be observed at the void volume, indicating the change in absorbance produced when the water injected passes through the column.) To wash, repeat the gradient and hold at the maximum concentration of NaCl for at least 10 column volumes. Wash extensively with water, then repeat the blank run.

7. Inject and analyze the sialylated oligosaccharide standards individually as indicated in step 6. Determine the elution time for each of the standards. If the separation of the individual standards is adequate, proceed to the analysis of the samples. If not, it may be necessary to adjust the rate of increase in the NaCl concentration to obtain better separation.

The column should be routinely standardized. At least one standard run, preferably with a mixture of sialylated oligosaccharide standards, should be performed at the start of each day.

Table 17.21A.1 DEAE HPLC Elution Regimen for Analysis of N-Linked Oligosaccharides

Time (min)	Eluant
0	100% H ₂ O
5	100% H ₂ O
95	50% 0.25 M NaCl
100	50% 0.25 M NaCl
140	100% H ₂ O
155	100% H ₂ O

- Analyze the desalted samples using the optimum conditions determined in step 7. If the amount of sample is small or the column is being used to obtain individual species for subsequent analysis, collect and analyze fractions every 0.5 to 1.0 min. Analyze the individual fractions by using either liquid scintillation counting or the TBA assay for sialic acids (UNIT 17.18).

The sample must be desalted (as described in UNITS 17.13A & 17.14A) prior to analysis, because even low concentrations of NaCl can prevent the initial binding of the anionic oligosaccharides to the cationic matrix. The desalted oligosaccharides can be concentrated by lyophilization and resuspended in a small volume of water. The main factor in determining the appropriate sample volume is the size of the injection loop, although injection loops of $\leq 500 \mu\text{l}$ are preferred.

- Estimate the number of sialic acid residues present by comparing the elution position of each peak to that of the sialylated oligosaccharide standards run in the same conditions.

*The linkage of the sialic acid residues can also be determined using sialidase digestion (UNIT 17.12) in combination with repeated anion-exchange HPLC. If, for instance, a sample that originally comigrated with a trisialylated triantennary oligosaccharide standard is treated with Newcastle disease virus sialidase, under conditions in which only $\alpha 2$ -3-linked sialic acids are released, and now comigrates with a monosialylated species, it can be inferred that the oligosaccharide in question has two $\alpha 2$ -3-linked sialic acid residues. Additionally, if the oligosaccharide in question is not retained by the DEAE column following treatment with the broad-spectrum sialidase from *Arthrobacter ureafaciens*, it can be inferred that the negative charge on the oligosaccharide in question was due solely to sialic acid residues. Although this type of analysis can be performed on a mixture of oligosaccharides, it is preferable to first isolate the individual HPLC peaks and then to treat with sialidase and reanalyze each separately.*

CAUTION: Remember to desalt the samples submitted to sialidase digestion prior to ion-exchange HPLC analysis.

FRACTIONATION OF GANGLIOSIDES ON DEAE HPLC COLUMNS

The use of DEAE HPLC columns to separate gangliosides from neutral lipids and glycolipids and fractionate them according to the number of sialic acid residues is described. Total lipid extracts, partially fractionated negatively charged glycolipids (total lipid extracts or the fraction enriched in polar lipid obtained by Folch partitioning of the total lipid extracts; see UNIT 17.3), or the total ganglioside fraction eluted from a DEAE anion-exchange column (see UNIT 17.3) can be further fractionated and analyzed by this method. The DEAE column is equilibrated in organic solvent and eluted with a gradient of increasing ammonium acetate concentration. Fraction components are detected based on their radioactivity (if previously radiolabeled) or by HPTLC analysis with chemical or immunological detection (if unlabeled) and gangliosides identified by comparison with known standards.

Materials

- HPLC-grade or high-quality deionized (e.g., 18 M Ω Milli-Q type) water
- HPLC-grade chloroform
- HPLC-grade methanol
- 1:8:1 (v/v/v) chloroform/methanol/1 M aqueous ammonium acetate (see recipe)
- Ganglioside sample (extracted and desalted; UNIT 17.3)
- Ganglioside standards (G_{M1} , G_{D1a} and G_{T1b} ; e.g., Sigma, Boehringer Mannheim, or Calbiochem)
- HPLC apparatus capable of two-component gradient formation
- DEAE HPLC column (TosoHaas TSK DEAE-2SW, or comparable)

BASIC PROTOCOL 2

Preparation and Analysis of Glycoconjugates

17.21.3

Additional reagents and equipment for high-performance thin-layer chromatography (HPTLC; if analyzing unlabeled sample, UNIT 17.17A) with chemical detection (UNIT 17.10B) and for TBA assay (UNIT 17.18)

1. Equilibrate the column in organic solvent by running a linear gradient from 100% water to 1:8:1 (v/v/v) chloroform/methanol/water at 1 ml/min, over 1 hr. If the column was stored in buffer, first wash extensively with water.

The water and all aqueous buffers used should be prefiltered through a 0.45- μ m filter and degassed. Do not degas mixtures containing organic solvents prior to chromatography; moreover, for systems that require continuous helium bubbling during chromatography, a low flow is sufficient. Column prefilters and/or guard columns provide additional protection for the HPLC column and increase its productive lifetime. When using a guard column, be sure to connect it prior to performing the equilibration with organic solvent.

2. Convert the column to the acetate form by running a 1-hr linear gradient from 1:8:1 (v/v/v) chloroform/methanol/water to 1:8:1 (v/v/v) chloroform/methanol/1 M aqueous ammonium acetate at 1 ml/min, and holding at this concentration of salt for another 1 hr. Wash the column extensively with 1:8:1 (v/v/v) chloroform/methanol/water to remove the salt. Use a flow rate of 1 ml/min throughout this step.
3. Flush out the injection loop with water, then methanol, then chloroform, and finally with 1:8:1 (v/v/v) chloroform/methanol/water.
4. For columns previously used to analyze radioactive samples, check the background with an on-line radioactivity detector or by collecting fractions and counting.
5. Set up a program from the pump control unit for washing unbound material, eluting the bound species, and bringing the column back to the starting conditions at 1 ml/min flow as indicated in Table 17.21A.2.

The DEAE resin will shrink and swell with increasing and decreasing salt concentrations, respectively. Therefore, transition from high salt concentration to no salt has to be done slowly. Rapid changes in ionic strength may cause damage to the column and degrade column performance.

Table 17.21A.2 DEAE HPLC Elution Regimen for Analysis of Gangliosides

Step	Time (min)	Eluant
Wash	0	1:8:1 (v/v/v) chloroform/methanol/water
Elution	10	1:8:1 (v/v/v) chloroform/methanol/water
	70	1:8:1 (v/v/v) chloroform/methanol/aqueous 1 M ammonium acetate
Reequilibration	100	1:8:1 (v/v/v) chloroform/methanol/water
	130	1:8:1 (v/v/v) chloroform/methanol/water

- Dissolve equivalent amounts of the ganglioside standards in 1:8:1 (v/v/v) chloroform/methanol/water. Load the solution of the standards into the injection loop using a syringe with a flat-tip needle compatible with the HPLC injector.

Solutions of these standards in 2:1, 1:1, or 1:2 (v/v) chloroform/methanol can also be used. When using a solvent other than the running solvent, keep the injection volume as small as possible.

- Inject the solution onto the column and start the gradient program. When analyzing radiolabeled standards, use an on-line radioactivity detector or route the eluate to a fraction collector (collecting 1-min fractions), remove aliquots of each fraction, add scintillation fluid, and measure radioactivity by counting in a scintillation counter. When analyzing nonradioactive standards, collect fractions and analyze one aliquot of each fraction by HPTLC, detecting the bands chemically (UNIT 17.10B) or by overlaying with monoclonal antibodies when available.

When using an on-line radioactivity detector, determine the required ratio of scintillation cocktail to eluant to avoid excessive quenching by the organic solvents.

- Determine if the separation of the ganglioside standards obtained is as expected (see Table 17.21A.3).

Separation is typically very reproducible, and it is not necessary to run standards each time the column is used.

Table 17.21A.3 Expected Retention Times of Ganglioside Standards

Ganglioside standard	Retention time (min)
G _{M1}	26–30
G _{D1a}	39–43
G _{T1b}	60–64

- Analyze the individual fractions by liquid scintillation counting, by the TBA assay (UNIT 17.18), or by HPTLC as indicated in step 7.

Normally, total lipid extract is obtained free of salts (UNITS 17.3 or UNIT 17.10B) and can be directly analyzed. Samples submitted to previous fractionation steps using salts must be desalted by dialysis or by loading onto a Sep-Pak C₁₈ cartridge prior to analysis (salts are washed off with water and anionic molecules eluted with organic solvents), because even low concentrations of salt can prevent binding of the anionic oligosaccharides to the cationic matrix (UNIT 17.3). Desalted gangliosides can be concentrated by lyophilization and redissolved in a small volume of 1:8:1 (v/v/v) chloroform/methanol/water.

When analyzing radioactive labeled samples, a peak in the void volume indicates the presence of neutral species (e.g., neutral lipids or glycolipids) that are also radioactive. These neutral species should not react in the TBA assay (UNIT 17.18) or with resorcinol spray (UNIT 17.10B). If TBA- or resorcinol-positive species are present at the void volume, the column capacity has been exceeded.

CAUTION: Dry down the organic solvent prior to performing the TBA assay.

When analyzing a large portion of each fraction by HPTLC, the water in the elution solvent will take some time to evaporate. Transfer the aliquot to a microcentrifuge tube, evaporate the solvent to dryness (i.e., use a Speedvac evaporator system), and redissolve in 2:1, (v/v) chloroform/methanol to spot the plate.

- Determine the type of ganglioside (e.g., mono-, di-, or trisialylated) present by comparing the elution position of each peak to that of known ganglioside standards run in the same conditions.
- Determine the number of gangliosides in each fraction by analyzing the HPTLC patterns obtained with resorcinol staining (*UNIT 17.10B*).
- When doing semipreparative chromatography, pool the fractions according to the profile obtained. Concentrate under vacuum, dissolve in a small volume of 1:8:1 (v/v/v) chloroform/methanol/water, dilute with 10 vol water, and desalt by overnight dialysis against water using 500- or 1000-MWCO tubing. Recover the gangliosides by lyophilization.

Alternatively, dry sample, resuspend in water, and desalt using a Sep-Pak C₁₈ cartridge previously washed successively with 20 ml each chloroform, dimethyl sulfoxide, acetonitrile, methanol, and water (see also UNIT 17.3).

IMPORTANT NOTE: *Semipreparative chromatography is defined as the fractionation of the individual species obtained by HPLC with a standard-sized column.*

REAGENTS AND SOLUTIONS

Use HPLC-grade or other high-quality deionized water (e.g., 18 MΩ Milli-Q type) in all recipes and protocol steps, and filter through a 0.45-μm filter and degas before use. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Chloroform/methanol/1 M aqueous ammonium acetate, 1:8:1 (v/v/v)

Dissolve 77 g ammonium acetate in HPLC-grade water. Adjust the volume to 1 liter and filter through a 0.45-μm filter. Combine 1 vol of this solution with 1 vol chloroform and 8 vol methanol, mix, and use for HPLC. Store in a tightly capped bottle.

Sodium chloride (NaCl), 0.5 M

Dissolve 29.25 g NaCl in 900 ml water. Adjust the volume to 1 liter and filter through a 0.45-μm filter before use.

COMMENTARY

Background Information

Ion-exchange HPLC using weak anion exchange chromatography columns has been used to analyze and/or fractionate a variety of anionic oligosaccharides and glycoconjugates. Negatively charged saccharides bind to the positively charged resin of the column and are eluted in order of increasing negative charge with a gradient of increasing ionic strength. DEAE HPLC columns have been run in water for the fractionation of released N- and O-linked oligosaccharides and in organic solvents for the fractionation of gangliosides. Although HPLC analysis requires more equipment than chromatography on QAE-Sephadex (*UNIT 17.19*), the profiles obtained are sharper and saccharides containing few negative charges can be better resolved.

DEAE and amine-bonded (Baenziger and Natowicz, 1981) HPLC columns have been used to fractionate released N-linked oligosac-

charides that differ in their number of sialic acids. Sialidase digestion (*UNIT 17.12*) followed by repeated anion-exchange chromatography yields information about the number and linkage of the sialic acid residues on the oligosaccharide. The decrease in negative charge following sialidase treatment indicates the number of sialic acid residues removed, and the linkage is inferred from knowledge of sialidase specificity. Similarly, oligosaccharides can be analyzed for the presence of phosphate by subjecting them to ion-exchange HPLC analysis following alkaline phosphatase digestion with and without prior mild acid treatment (Varki and Kornfeld, 1983). If the phosphate is “covered” by *N*-acetyl-D-glucosamine (GlcNAc), it will be resistant to alkaline phosphatase. Treatment with mild acid removes the GlcNAc residue and exposes the phosphate for removal by alkaline phosphatase.

Sulfated and sialylated/sulfated oligosaccharides have also been separated using weak anion-exchange HPLC columns. Amine-bonded columns (Micropak AX-5, Varian) have been used to analyze sulfated/sialylated oligosaccharides (Green and Baenziger, 1986). When the column eluant is buffered to pH 4, both sialic acid and sulfate have one negative charge. However, at pH 1.7 sialic acid is only partially ionized and, therefore, has a net negative charge of <1. Thus, performing the HPLC analysis at two different pHs, near the pK of one of the anionic species, permits the fractionation of sulfated and sulfated/sialylated oligosaccharides. This approach has been adapted to DEAE HPLC columns by buffering the column with phosphate buffer at pH 2.5 for the separation of dermatan sulfate fragments (Hayes and Varki, 1993a).

The separation of more negatively charged saccharides requires the use of strong anion-exchange columns. Polysialic acids with 2 to 20 sialic acids have been separated using a Mono-Q column (Hallenbeck et al., 1987).

Several methods to detect oligosaccharides are available. The choice depends upon the nature of the sample. Radiolabeled oligosaccharides can be detected using an on-line radioactivity detector or by collecting fractions and analyzing all or a portion of the individual fractions by liquid scintillation counting. Oligosaccharides can be radiolabeled, either metabolically (UNIT 17.4), enzymatically (UNIT 17.6), or chemically (UNIT 17.5). Oligosaccharides can also be detected spectrophotometrically, as many monosaccharides absorb between 190 and 205 nm. This method, though simple, requires quantities of oligosaccharide that are often prohibitive (>5 nmol). Moreover, the sample and HPLC solvents must be of high purity and the elution medium cannot contain salts that absorb at these wavelengths (e.g., acetate). An alternative approach is to assay the fractions for the presence of sialic acid using the TBA assay (UNIT 17.18), or, in the case of gangliosides, by HPTLC with chemical detection (UNIT 17.10B).

Interpretation of data depends on the use of appropriate standards. Several sialylated N- and O-linked oligosaccharide standards are commercially available. The same standards used to calibrate the HPLC column can be used as controls for sialidase digestions. It should be noted that several distinct oligosaccharide structures may have equivalent net negative charge, resulting in coelution; therefore, DEAE

HPLC is only one facet of oligosaccharide structural analysis.

DEAE HPLC is also useful for separating gangliosides according to the number of sialic acids they contain. All monosialylated gangliosides coelute, although some fractionation can occur because of other structural differences. The same is true for the higher-order (e.g., di- or trisialylated) gangliosides. Complete purification of individual species usually requires a further chromatographic step based on other principles (e.g., HPLC columns of a special silica gel called Iatrobeads, from Iatron Laboratories). This method has been used for the analysis and semipreparative fractionation of labeled and unlabeled gangliosides (Manzi et al., 1990; Sjoberg et al., 1992). When gangliosides contain sulfate groups, interaction is stronger and, in some cases, it is necessary to keep eluting at maximum salt concentration for long periods (A. Manzi, unpub. observ.).

Critical Parameters

The separation of anionic oligosaccharides is greatly affected by the steepness of the NaCl gradient. If the concentration of NaCl is increased too rapidly, the resolution of the differently charged species will be reduced. Additionally, the column may be damaged because of shrinking and swelling of the matrix, which occur with changes in ionic strength. A final concentration of 125 mM NaCl should be sufficient to elute saccharides with up to five negative charges. More negatively charged species can be eluted by increasing the final salt concentration. The highest concentration of NaCl compatible with a given column is indicated by the supplier (e.g., some DEAE columns should not be exposed to NaCl concentrations >1 M).

Failure of acidic saccharides to bind to the column can be caused by several factors. Salt in the sample may impair the ionic interactions responsible for binding to the column. Desalting the sample prior to analysis should prevent this. Alternatively, the acidic moieties may be protonated and hence uncharged. This is unlikely, because the pK_a of the carboxylic acids of acidic sugars is typically <3, with sulfates having even lower pK_a values. On the other hand, phosphodiester have pK_a values around 5, and CO₂ dissolved in column buffers may lower the pH of the column eluant sufficiently to protonate them. This is easily addressed by buffering the column to a pH >6 with 2 to 5 mM Tris base, pH 9.5. In this case, UV cannot be used for detection.

It is not advisable to change the solvent from aqueous to organic and back to analyze anionic oligosaccharides and gangliosides, because the performance of the column in both applications is markedly affected. If both types of analyses are regularly done, two different columns must be used.

Anticipated Results

Oligosaccharides or gangliosides with sialic acid as their only anionic moiety can be separated using the above protocols. When one to four sialic acids/molecule are present, the species are readily resolved. Sialidase treatment of released N-linked oligosaccharides with repeated HPLC analysis provides information about the linkage of the sialic acid residues. In the case of gangliosides, the method is particularly useful for semipreparative purposes.

Time Considerations

A typical separation of released N-linked oligosaccharides requires 1 to 2 hr, depending on the length and steepness of the gradient and the length of time the column is allowed to wash. At least one standard run, preferable with a mixture of standards, should be performed at the start of each day. Thus three to six samples can be analyzed in an 8-hr day. When analyzing gangliosides, one run and equilibration of the column back to the starting conditions usually takes 2 hr.

Literature Cited

- Baenziger, J.U. and Natowicz, M. 1981. Rapid separation of anionic oligosaccharide species by high performance liquid chromatography. *Anal. Biochem.* 112:357-361.
- Green, E.D. and Baenziger, J.U. 1986. Separation of anionic oligosaccharides by high-performance liquid chromatography. *Anal. Biochem.* 158:42-49.
- Hallenbeck, P.C., Yu, F., and Troy, F.A. 1987. Rapid separation of polysialic acid by high-performance liquid chromatography. *Anal. Biochem.* 161:181-186.

Hayes, B.K. and Varki, A. 1993a. Biosynthesis of oligosaccharides in intact Golgi preparations from rat liver. Analysis of N-linked and O-linked glycans labeled by UDP-[6-³H]N-acetylglucosamine. *J. Biol. Chem.* 268:16170-16184.

Manzi, A.E., Sjoberg, E.R., Diaz, S., and Varki, A. 1990. Biosynthesis and turnover of O-acetyl and N-acetyl groups in the gangliosides of human melanoma cells. *J. Biol. Chem.* 265:13091-13103.

Sjoberg, E.R., Manzi, A.E., Dell, A., Khoo, K.-H., and Varki, A. 1992. Structural and immunological characterization of O-acetyl GD2: Evidence that GD2 is an acceptor for ganglioside O-acetyltransferase in human melanoma cells. *J. Biol. Chem.* 267:16200-16211.

Snyder, L.R., Glajch, J.L., and Kirkland, J.J. 1988. *Practical HPLC Method Development*. John Wiley & Sons, New York.

Varki, A. and Kornfeld, S. 1983. The spectrum of anionic oligosaccharides released by endo- β -N-acetylglucosaminidase H from glycoproteins. *J. Biol. Chem.* 258:2808-2818.

Key References

Hayes, B.K. and Varki, A. 1993b. Biosynthesis of oligosaccharides in intact Golgi preparations from rat liver. Analysis of N-linked glycans labeled by UDP-[6-³H]galactose, CMP-[9-³H]N-acetylneuraminic acid and [³H-acetyl]AcCoA. *J. Biol. Chem.* 268:16155-16169.

Describes analysis of sialylated N-linked oligosaccharides with and without prior sialidase treatment.

Manzi et al., 1990. See above.

Describes qualitative and semipreparative analysis of mono-, di-, and trisialylated gangliosides.

Contributed by Adriana E. Manzi
University of California San Diego
La Jolla, California

Bradley K. Hayes
University of Alabama at Birmingham
Birmingham, Alabama

Fractionation and Analysis of Neutral Oligosaccharides by HPLC

UNIT 17.21B

This unit describes the fractionation and analysis of neutral oligosaccharides by high-performance liquid chromatography (HPLC) on bonded amine columns. Separation is achieved with gradients of acetonitrile and water and is based upon hydrogen bonding between the NH_2 groups of the column and the hydroxyl groups of the oligosaccharides, as described in the Basic Protocol. The Support Protocol describes the reduction and desalting of neutral oligosaccharides with sodium borohydride. The basic principles of HPLC are discussed in *UNITS 10.12-10.14* and in Snyder et al. (1988).

SEPARATION OF NEUTRAL OLIGOSACCHARIDES ON A BONDED AMINE HPLC COLUMN

**BASIC
PROTOCOL**

Neutral oligosaccharides bind to amine-bonded HPLC columns because of hydrogen bonding via hydroxyl groups, which is promoted by acetonitrile and disrupted by water. The sample is loaded in a high concentration of acetonitrile and the column is developed with a gradient of increasing water. Neutral oligosaccharides are initially retained by the column and elute in order of increasing size (increasing numbers of hydroxyl groups). Oligosaccharides are detected based on their radioactivity (if previously radiolabeled) or by physical or chemical means (if unlabeled). Information about the number and size of oligosaccharide species in a mixture is thus obtained.

Materials

- Oligosaccharide sample
- Oligosaccharide standards (e.g., Dionex or Oxford Glycosystems)
- HPLC-grade water
- HPLC-grade acetonitrile
- HPLC apparatus capable of two-component gradient formation
- Bonded amine HPLC column (e.g., Varian Micropak AX-5 or Rainin LC-NH₂)
- Additional reagents and equipment for exoglycosidase digestion (optional; *UNITS 17.13 & 17.18*), and reduction and desalting of neutral oligosaccharides (see Support Protocol)

1. Reduce and desalt the sample prior to analysis (see Support Protocol). If desired, prepare parallel samples treated by sequential or combined exoglycosidase digestion (*UNIT 17.13 & 17.18*) chosen to help elucidate the structure.

Reduction of samples gives sharper peaks and better baseline resolution.

The column is also a weak anion exchanger. Salts or buffers can compete with the oligosaccharide for binding sites, altering elution time and resolution. Anionic oligosaccharides will bind to the column and will not elute with water. However, the method can be adapted to study anionic oligosaccharides.

2. Equilibrate the column in 70% acetonitrile/30% water at a flow rate of 1 ml/min.

*Water and acetonitrile should be prefiltered through a 0.45- μm filter and degassed. See *UNIT 10.12* support protocol for preparation and degassing of solvents. Column prefilters and/or guard columns provide additional protection and increase the productive lifetime of the column.*

All common neutral oligosaccharides should bind to the column in 70% acetonitrile. If an oligosaccharide fails to bind, the starting concentration of acetonitrile should be increased.

**Preparation and
Analysis of
Glycoconjugates**

17.21.9

3. Dissolve the sample in a small volume of water and add 3 vol acetonitrile.

The main factor determining the appropriate sample volume is the size of the injection loop. Injection loops of $\leq 500 \mu\text{l}$ are preferred.

4. Load and inject the sample. Immediately start a 70% to 30% acetonitrile gradient over 80 min, maintaining the flow rate of 1 ml/min.

Typical neutral N-linked oligosaccharides elute at water concentrations of $< 70\%$. A second gradient of 30% to 0% acetonitrile over 30 min may be included to elute unexpectedly large neutral oligosaccharides and to help keep the column clean.

5. Detect the oligosaccharides by physical, chemical, or radiometric methods. For analytical work, in-line radioactivity detectors or UV detectors may be used. If the amount of radioactive sample is low, or if the column is being used to obtain individual species for subsequent analysis, collect and analyze fractions every 0.3 to 0.5 min. For UV detection, measure absorbance at 190 to 210 nm. Determine the number of oligosaccharide species present and compare their retention times to those of known standards prepared similarly.

Because many compounds absorb at 190 to 210 nm, UV detection demands a sample of high purity.

6. Infer features of the oligosaccharide structure from the retention time relative to standards and from changes in retention caused by sequential or combined exoglycosidase digestion.

Because retention time is also a reflection of column performance, it is important to properly care for the column. Routine washing with water (step 4 annotation) helps remove contaminants that are hydrogen bonded to the column. Salts or other ionic contaminants can be removed by periodic washing with 0.5 M potassium phosphate, pH 1.7. Column performance will deteriorate over time, even with washing, due to gradual dissolution of the bonded phase. Standards must therefore be run at the beginning and end of each set of samples.

SUPPORT PROTOCOL

SODIUM BOROHYDRIDE REDUCTION AND DESALTING OF NEUTRAL OLIGOSACCHARIDES

Oligosaccharides with a reducing sugar at their inner terminus elute as broader peaks than those that have been reduced to alditols. Whenever possible, therefore, oligosaccharides should be reduced prior to HPLC analysis. Samples and standards should be similarly treated. Recoveries of oligosaccharides after reduction and desalting are typically $\sim 70\%$. Note that a radioactive label can be introduced at this stage if ^3H -labeled sodium borohydride ($[^3\text{H}]\text{NaBH}_4$) is used (UNIT 17.5) and that *O*-*N*-acetyl-D-galactosamine (*O*-GalNAc)-linked oligosaccharides are reduced during their release by β -elimination (UNIT 17.15).

Additional Materials (also see Basic Protocol)

1 M sodium borohydride (NaBH_4) in 0.2 M sodium borate ($\text{Na}_2\text{B}_4\text{O}_7$; see recipe)
5% acetic acid/95% methanol
Mixed-bed anion and cation resin (Amberlite MB-3 or equivalent)

1. Dissolve NaBH_4 to 1 M final concentration in 0.2 M $\text{Na}_2\text{B}_4\text{O}_7$ (pH 9.5) immediately before use. Dissolve the dried oligosaccharides in 100 μl of this reagent and incubate 3 hr at room temperature. Because hydrogen gas may be evolved by decomposing NaBH_4 , cautiously vent the tube occasionally or use a tall, loosely capped tube.

Sodium borate buffer at pH 9.5 is preferred to more basic NaOH solutions, because the latter promote epimerization of the core N-glucosamine to N-mannitosamine, which can result in doublet peaks in the HPLC analysis.

2. Stop the reduction by adding dropwise a 5% acetic acid/95% methanol mixture. The sample will bubble due to evolution of hydrogen gas. Continue until no more gas is evolved.
3. Dry the sample under a stream of N₂ or using a shaker/evaporator, Speedvac evaporator, or similar device. Take to complete dryness.
4. Repeat steps 2 and 3 three times using 0.5 ml of 5% acetic acid/95% methanol each time.

Under acidic conditions the methanol and Na₂B₄O₇ form volatile methyl borates. The salt left behind is sodium acetate.

5. Prewash a 2 to 5 ml column of mixed-bed resin with at least 10 column volumes of water. Dissolve the dried sample in water, load, and elute with water. Collect a single 10 to 15 ml fraction.

Only neutral reduced oligosaccharides should pass through the column unretarded.

6. Lyophilize or dry the sample. Store the oligosaccharide at -20°C, either dry or in water.

Presence of white powder following lyophilization indicates that not all the salt was removed by the resin. If so, repeat the desalting. A 26-G needle added to the column outlet can slow the flow and increase the time for the salts to interact with the resin. Alternatively, an anion-exchange column (e.g., Dowex 3×4a, hydroxide form, prewashed with ≥10 vol water) can be used followed by a cation-exchange column (e.g., Dowex 50, hydrogen form, prewashed with ≥10 vol water).

REAGENTS AND SOLUTIONS

Use HPLC-grade or other high-quality deionized water (e.g., 18 MΩ Milli-Q type) in all recipes and protocol steps, and filter through a 0.45-μm filter and degas before use. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

1 M sodium borohydride in 0.2 M sodium borate

Prepare 0.2 M sodium borate by dissolving 7.63 g sodium borate (Na₂B₄O₇·10H₂O) in 90 ml H₂O. Adjust pH to 9.5 with NaOH and volume to 100 ml. Store at room temperature.

Immediately before use, dissolve 37.8 mg sodium borohydride (NaBH₄) per ml 0.2 M sodium borate, pH 9.5. Do not mix excessively.

COMMENTARY

Background Information

The fractionation and analysis of neutral oligosaccharides on the basis of size using this HPLC method yield several important pieces of information. A minimum estimate of the number of species present is obtained from the number of peaks observed. The size of an N-linked oligosaccharide may be indicative of the extent to which it has been processed. Single or sequential exoglycosidase digestion(s) (UNIT 17.18), in conjunction with repeated HPLC analysis, yield information about the identity and linkage of the monosaccharides present. Although HPLC analysis requires more equipment than conventional chromatography on Bio-Gel P-4, the elution profiles obtained are

sharper and larger oligosaccharides can be better resolved.

Bonded amine HPLC columns have free NH₂ groups that hydrogen bond to the hydroxyl groups of oligosaccharides. Elution is accomplished by increasing the water content, which progressively disrupts the hydrogen bonds. Other factors affecting elution time include the monosaccharide composition and the branching pattern of the oligosaccharide. Several methods of detecting the oligosaccharides are available. Radiolabeled molecules can be detected with an in-line radioactivity detector or by liquid scintillation counting of fractions. Oligosaccharides can be radiolabeled, either metabolically (UNIT 17.4), enzymatically (UNIT

17.6), or chemically (UNIT 17.5). Because some monosaccharides absorb UV light between 190 and 205 nm, they can also be detected spectrophotometrically. However, this requires quantities of oligosaccharide that are sometimes prohibitive (>5 nmol) and a sample of high purity. An alternative approach is to label the reducing terminus of the oligosaccharide with a fluorescent tag (see Tomiya et al., 1988, and references therein). However, some hydrophobic tags can produce anomalous behavior on these HPLC columns.

Interpretation of the data is dependent on the use of appropriate standards. Complex, hybrid, and high-mannose N-linked oligosaccharides and several O-linked oligosaccharide standards are commercially available. The same standards used to calibrate the HPLC column can also be used as controls for exoglycosidase digestions.

The use of bonded amine HPLC columns to separate neutral oligosaccharides was introduced in the early 1980s. Since then several articles have been published describing the behavior of model oligosaccharides on these columns and elucidating the factors that affect elution time (Mellis and Baenziger, 1981; Blanken et al., 1985). Such systems have been used to study the substrate specificity of glycosyltransferases (e.g., Schachter et al., 1989; Brockhausen et al., 1988; Koenderman et al., 1989) and to deduce the order of release of monosaccharides from oligosaccharides by exoglycosidase digestion (e.g., Tomiya et al., 1991). Sequential exoglycosidase digestion coupled with HPLC analysis has been used extensively to elucidate oligosaccharide structure (e.g., Sampath et al., 1992).

A modification of this method has been described for the fractionation of anionic oligosaccharides on the basis of size (Mellis and Baenziger, 1983). Triethylamine acetate is used to suppress the charge of sialic acid, phosphate, and sulfate, thus allowing separation of oligosaccharides on the basis of monosaccharide content alone. The reader is referred to the original description (and to UNIT 17.17B) for further details.

Critical Parameters

The most important parameter is column performance. There is some variability in the quality of separation by columns from different manufacturers and between columns from a given manufacturer. Decreases in column performance can be expected to occur over time due to the gradual dissolution of the bonded

phase (Blanken et al., 1985). Improper care can greatly accelerate the decay of column performance. Inadequate cleaning, the use of unfiltered solvents, and the presence of contaminants in the sample can lead to higher back pressure, broader peaks, and decreased retention time.

The nature of the gradient also affects the ability of the column to resolve individual oligosaccharide species (Mellis and Baenziger, 1981). Most oligosaccharides are retained by the column if the starting acetonitrile/water ratio is 70:30. Typical gradients increase the water content by 0.5%/min. However, for separation of larger oligosaccharides the water content may need to be increased by only 0.3%/min (Blanken et al., 1985). Oligosaccharides that have not been reduced and desalted give broader peaks than reduced molecules. Reduction should be performed in sodium borate buffer and not in NaOH, as strongly alkaline conditions promote the epimerization of the GlcNAcitol to ManNAcitol, which can lead to doublet peaks (Mellis and Baenziger, 1981). Reduced oligosaccharides may have longer retention times than their unreduced counterparts (Koenderman et al., 1989). Thus, the standards used to calibrate the column should be treated the same way as the samples.

Anticipated Results

Homopolymers of increasing size of GlcNAc (chitin) and glucose (dextrans) have been separated into a series of well-spaced peaks, corresponding to oligomers that differ in size by one monosaccharide unit (Mellis and Baenziger, 1981). The elution time of oligosaccharides depends on the number of monosaccharides present and their composition. GlcNAc- and fucose-containing oligosaccharides elute sooner than oligosaccharides having an equal number of hexoses (Blanken et al., 1985), presumably because GlcNAc and fucose have one fewer hydroxyl group. The branching pattern of an oligosaccharide also affects its elution time. Oligosaccharides containing a β 1-6 linkage elute later, presumably because the flexibility of this linkage allows a stronger interaction with the column (Blanken et al., 1985). Thus, it is possible to separate isomers of the same size given good column performance.

Time Considerations

A typical separation requires 1 to 2 hr, depending upon the length and steepness of the gradient and the length of time the column is

allowed to wash. At least one standard run, preferably with a mixture of standards, should be performed at the start of each day. Thus three to six samples can be analyzed in an 8-hr day.

Literature Cited

- Blanken, W.M., Bergh, M.L.E., Koppen, P.L., and van den Eijnden, D.H. 1985. High pressure liquid chromatography of neutral oligosaccharides: Effects of structural parameters. *Anal. Biochem.* 145:322-330.
- Brockhausen, I., Narasimhan, S., and Schachter, H. 1988. The biosynthesis of highly branched N-glycans: Studies on the sequential pathway and functional role of N-acetylglucosaminyltransferases I, II, III, IV, V and VI. *Biochimie* 70:1521-1533.
- Koenderman, A.H.L., Koppen, P.L., Koeleman, A.M., and van den Eijnden, D.H. 1989. N-acetylglucosaminyltransferases III, IV and V activities in Novikoff ascites tumour cells, mouse lymphoma cells and hen oviduct. *Eur. J. Biochem.* 181:651-655.
- Mellis, S.J. and Baenziger, J.U. 1981. Separation of neutral oligosaccharides by high pressure liquid chromatography. *Anal. Biochem.* 114:276-280.
- Mellis, S.J. and Baenziger, J.U. 1983. Size fractionation of anionic oligosaccharides and glycopeptides by high-pressure liquid chromatography. *Anal. Biochem.* 134:442-449.
- Sampath, D., Varki, A., and Freeze, H.H. 1992. The spectrum of incomplete N-linked oligosaccharides synthesized by endothelial cells in the presence of Brefeldin A. *J. Biol. Chem.* 279:4440-4455.

- Schachter, H., Brockhausen, I., and Hull, E. 1989. High-performance liquid chromatography assays of N-acetylglucosaminyltransferases involved in N- and O-linked synthesis. *Methods Enzymol.* 179:351-396.
- Snyder, L.R., Glajch, J.L., and Kirkland, J.J. 1988. Practical HPLC Method Development. John Wiley & Sons, New York.
- Tomiya, N., Awaya, J., Kurono, M., Endo, S., Arata, Y., and Takahashi, N. 1988. Analysis of N-linked oligosaccharides using a two-dimensional mapping technique. *Anal. Biochem.* 171:73-90.
- Tomiya, N., Yuan, Y.C., Yoshida, T., Wada, Y., Away, J., Kurono, M., and Takahashi, N. 1991. Calculated two-dimensional sugar map of pyridylaminated oligosaccharides: Elucidation of the Jack Bean α -mannosidase digestion pathway of Man₉GlcNAc₂. *Anal. Biochem.* 193:90-100.

Key References

- Blanken et al., 1985. See above.
Describes important parameters for optimal resolution of neutral oligosaccharides.
- Mellis and Baenziger, 1981. See above.
Describes important parameters for optimal resolution of neutral oligosaccharides.

Contributed by Bradley K. Hayes
University of Alabama at Birmingham
Birmingham, Alabama

Ajit Varki
University of California San Diego
La Jolla, California

Nitrous Acid Degradation of Glycosaminoglycans

UNIT 17.22A

**BASIC
PROTOCOL**

Glycosaminoglycans (GAGs) are made up of disaccharide units that are distinguished from each other by the monosaccharide units of which they are composed and by the degree and position of sulfation. These disaccharide units represent the monomeric units of the GAG; thus, measurement of the disaccharide composition of a GAG represents the first step in the characterization of the polymer—just as an amino acid analysis represents the first step in characterization of a protein.

In this protocol, alternative sets of steps are presented—the first set for cleavage of the glycosidic bonds of the N-sulfated GlcN residues in heparin and heparan sulfate, and the second set for cleavage of the bonds between the N-acetylated amino sugar residues in heparin, heparan sulfate, chondroitin sulfate, dermatan sulfate, and keratan sulfate, as well as hyaluronic acid. The glycosidic bonds of N-sulfated GlcN residues can be rapidly cleaved with nitrous acid at pH 1.5 at room temperature to yield oligosaccharides. N-acetylated amino sugar residues do not react with nitrous acid. However, these residues can first be deacetylated by hydrazinolysis to generate GalN or GlcN residues, and the glycosidic bonds of these N-unsubstituted amino sugar residues can then be cleaved at room temperature with nitrous acid at pH 4. The N-sulfate groups in heparin and heparan sulfate are stable under the hydrazinolysis conditions. Thus, following N-deacetylation by hydrazinolysis, all GAGs can be converted completely to their constituent disaccharides by treatment with nitrous acid at pH 4 alone (for chondroitin, dermatan, , and keratan sulfates, as well as hyaluronic acid) or with nitrous acid at pH 4 and then at pH 1.5 (for heparin and heparan sulfates). The resulting cleavage products contain reducing-terminal 2,5-anhydro-D-mannose residues (from GlcN) or 2,5-anhydro-D-talose residues (from GalN), which can be reduced stoichiometrically with tritiated sodium borohydride ($\text{NaB}[^3\text{H}]_4$) to ^3H -anhydromannitol or ^3H -anhydrotalitol residues—thus labeling the disaccharides so that they can be assayed qualitatively or quantitatively. The procedure described in the Basic Protocol may also be used to analyze metabolically labeled GAGs (with or without the use of $\text{NaB}[^3\text{H}]_4$).

Materials

- Reference standard (see recipe)
- Unknown sample suspected of containing N-sulfated *or* N-acetylated GAG
- Internal standard (see recipe)
- Nitrous acid reagent, pH 1.5 (see recipe)
- 1 M Na_2CO_3
- 10 $\mu\text{g}/\mu\text{l}$ hydrazine sulfate in anhydrous hydrazine
- 3 M H_2SO_4
- Nitrous acid reagent, pH 4.0 (see recipe)
- 0.5 M tritiated sodium borohydride ($\text{NaB}[^3\text{H}]_4$; ~500 mCi/mmol) in 0.1 M NaOH
- 100- μl Reacti-Vials (Pierce)
- Sand bath: heating block (e.g., Pierce) with wells filled with sand

Cleave GAGs

To cleave N-sulfated GAGs

- 1a. Mix 20 μl of reference standard solution for N-sulfated GAG assay with 5 μl D- ^{14}C glucose internal standard. Prepare the same reaction mixture using 20 μl of unknown sample suspected of containing N-sulfated GAG in place of the standard, and carry both mixtures through all steps in parallel.

**Preparation and
Analysis of
Glycoconjugates**

17.22.1

Contributed by H. Edward Conrad

Current Protocols in Molecular Biology (1995) 17.22.1-17.22.5

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 32

- 2a. Cool a 5- μ l aliquot of the reaction mixture to 0°C and add 20 μ l nitrous acid reagent, pH 1.5.
- 3a. Let mixture warm to room temperature for 10 min to complete the deamination reaction.
- 4a. Bring cleaved reaction mixture to pH 8.5 with 1 M Na₂CO₃.

The cleaved product is then subjected to borohydride reduction and analysis (steps 5 and 6). After borohydride reduction but prior to analysis, it may be desirable to separate the reaction products (a mixture of di- and tetrasaccharides) according to size by gel filtration (see UNIT 10.9) on a Bio-Gel P-10 column, eluting with 1 M Na₂CO₃.

To cleave N-acetylated GAGs

- 1b. Mix 200 μ l of reference standard for N-acetylated GAG assay with 50 μ l of [¹⁴C]glucitol internal standard. Prepare the same reaction mixture using 20 μ l of unknown sample suspected of containing N-acetylated GAG in place of the standard, and carry both mixtures through all steps in parallel.
- 2b. Place a 15- μ l aliquot of the reaction mixture in a 100- μ l Reacti-Vial and evaporate to dryness in a stream of air. Redissolve dried sample in 20 μ l of 10 μ g/ μ l hydrazine sulfate in anhydrous hydrazine, then cap the vial and incubate 6 hr at 100°C in a sand bath.

CAUTION: Hydrazine is a toxic, corrosive, and flammable reagent and should be handled accordingly.

- 3b. Cool reaction mixture, evaporate until as dry as possible in a stream of air, then lyophilize partially dried sample to remove as much hydrazine as possible. Add 20 μ l nitrous acid reagent, pH 4 to the dried sample. Check pH with pH paper and adjust to pH 4 using 5 to 10 μ l of 3 M H₂SO₄, if necessary. Let stand 15 min at room temperature.

Because of residual hydrazine sulfate and hydrazine, the solution is usually at ~pH 6 before adjustment.

The hydrazinolysis procedure removes N-acetyl substituents from the N-acetylated amino sugar residues, but the N-sulfated substituents in heparin or heparan sulfate remain. Hence, steps 1b to 3b result in cleavage at the N-deacetylated hexosamine residues but not at the N-sulfated hexamines. Deamination (step 4b) cleaves at the N-sulfated residues as well; thus, steps 1b to 4b result in the complete conversion of heparin and heparan sulfate (as well as all of the other GAGs) to disaccharides.

- 4b. Adjust reaction mixture to pH 1.5 with 3 M H₂SO₄, then cool solution to 0°C, carry out the deamination reaction, and adjust the reaction mixture to pH 8.5 as in steps 2a to 4a.

The cleaved product is then subjected to borohydride reduction and analysis (steps 5 and 6). All of the products obtained after hydrazinolysis and nitrous acid cleavage are disaccharides, and therefore may be analyzed without further sample purification.

Reduce cleavage products with sodium borohydride

5. Mix cleaved sample with 10 μ l of 0.5 M Na[B³H]₄ in 0.1 M NaOH and incubate 15 min at 50°C.

CAUTION: Carry out all NaB³H₄ manipulations in a well-ventilated fume hood designated for radioactive use.

6. Add 5 μl of 3 M H_2SO_4 and evaporate samples to dryness in a stream of air, then redissolve sample in water and again evaporate to dryness in a stream of air to remove as much $[\text{}^3\text{H}]_2$ as possible.

The sample is now ready for analysis. For analysis by paper chromatography or paper electrophoresis, it should be redissolved in 60 μl of water. For analysis by HPLC, it should be dissolved in the starting buffer to be used the chromatographic run (see UNIT 17.22B for both procedures).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Internal standards

D- $[\text{}^{14}\text{C}]$ glucose standard for N-sulfated GAG assay: Prepare a stock solution containing 20 $\mu\text{Ci}/\text{ml}$ D- $[\text{}^{14}\text{C}]$ glucose (specific activity >40 mCi/mmol; Amersham or Du Pont NEN). Dilute an aliquot of the stock solution in water to obtain a solution containing $\sim 1 \times 10^5$ cpm/ μl of radioactivity.

$[\text{}^{14}\text{C}]$ glucitol standard for N-acetylated GAG assay: Reduce D- $[\text{}^{14}\text{C}]$ glucose with sodium borohydride as in steps 5 and 6 of the Basic Protocol, except use cold NaBH_4 . Dilute the resulting $[\text{}^{14}\text{C}]$ glucitol in water to obtain a solution containing $\sim 1 \times 10^5$ cpm/ μl of radioactivity.

Because D- $[\text{}^{14}\text{C}]$ glucose reacts with hydrazine, it cannot be used directly as an internal standard in the hydrazinolysis reaction. Consequently, it must be reduced for use in the N-acetylated GAG assay.

Nitrous acid reagent, pH 1.5

Prepare 0.5 M H_2SO_4 and 0.5 M $\text{Ba}(\text{NO}_2)_2$ (114 mg/ml) and cool separately to 0°C in an ice bath. Prepare a mixture containing 1 ml of each solution—i.e., 0.5 mmol each of H_2SO_4 and $\text{Ba}(\text{NO}_2)_2$ —at 0°C , then pellet the BaSO_4 precipitate in a clinical centrifuge. Draw off the supernatant with a Pasteur pipet, and keep on ice until ready to use.

This reagent should be prepared when needed and used immediately.

Nitrous acid reagent, pH 4.0

Add 5 ml of 5.5 M NaNO_2 to 2 ml of 0.5 M H_2SO_4 .

This reagent should be prepared when needed and used immediately.

Reference standards

For N-sulfated GAG assay: 25 mg/ml aqueous solution of heparin or heparan sulfate.

For N-acetylated GAG assay: 25 mg/ml aqueous solution of chondroitin-4-sulfate (chondroitin sulfate type A), chondroitin-6-sulfate (chondroitin sulfate type C), dermatan sulfate (chondroitin sulfate type B), or keratan sulfate.

All of the above GAGs are available from Sigma. The standard solutions are stored frozen at -20°C .

COMMENTARY

Background Information

Glycosaminoglycans (GAGs) are families of structurally related polymers composed of sequences of different disaccharides that represent the monomeric units of these polymers. Chondroitin sulfates, dermatan sulfates, and hyaluronic acid represent one family of GAGs.

Chondroitin sulfates contain only GlcA and GalNAc residues, and are sulfated at positions 4 and/or 6 of the GalNAc (and on some of the uronic acid residues as well). In dermatan sulfates, the amino sugar is GalNAc, but both GlcA and IdoA are present in varying proportions. Hyaluronic acid contains GlcNAc and

GlcA. Keratan sulfate contains Gal and GlcNAc, both of which may be sulfated at the 6 position. Thus, in contrast to heparin and heparan sulfate, which contain both N-acetylated and N-sulfated amino sugars as described below, chondroitin sulfate, dermatan sulfate, hyaluronic acid, and keratan sulfate contain only N-acetylated amino sugars. The glycosidic bonds of these N-acetylated amino sugar residues are not cleaved by nitrous acid.

Heparan sulfate and heparin represent another family of glycosaminoglycans, which contain GlcN as the amino sugar and both GlcA and IdoA in varying proportions. A unique structural feature of heparin and heparan sulfate is that a large proportion of their GlcN residues are N-sulfated (85% to 90% in heparin; 40% to 60% in a typical heparan sulfate); the rest are N-acetylated. The glycosidic bonds of these N-sulfated GlcN residues can be cleaved rapidly at room temperature with nitrous acid at pH 1.5 to yield oligosaccharides. This nitrous acid susceptibility distinguishes heparin and heparan sulfate from all other glycosaminoglycans. Because of the low percentage of GlcNAc residues in heparin, the GlcNAcs are found in isolated disaccharide units, surrounded by disaccharides that have N-sulfated GlcNs. Thus, the pH 1.5 nitrous acid cleavage reaction products of heparin are primarily disaccharides, with much smaller quantities of tetrasaccharides from sequences in which the GlcNAcs are located. However, heparan sulfate contains sequences with multiple contiguous GlcNAc-containing disaccharide units, and these sequences yield larger oligosaccharides among the nitrous acid cleavage products. Thus, to analyze the disaccharides formed following cleavage with the pH 1.5 nitrous acid reagent, the saccharides are first separated from the larger oligosaccharides by gel filtration on a BioGel P-10 column using 1 M Na₂CO₃ for elution. The products are then analyzed by paper chromatography, paper electrophoresis, or high-performance liquid chromatography (HPLC) as described in *UNIT 17.22B*.

The N-acetylated amino sugars can be deacetylated by hydrazinolysis to generate GalN or GlcN residues, and the glycosidic bonds of these N-unsubstituted amino sugars can then be cleaved at room temperature with the pH 4 nitrous acid reagent. Thus, following N-deacetylation by hydrazinolysis, those glycosaminoglycans containing only N-acetylated amino sugars are converted completely to disaccharides by the pH 4 nitrous acid reagent (in contrast to heparin and heparan sulfate,

which yield higher oligosaccharides in addition to disaccharides). To convert heparin and heparan sulfate completely to disaccharides, these GAGs must first be N-deacetylated and then treated with nitrous acid at both pH 4 and pH 1.5.

Critical Parameters

The nitrous acid reagents are prepared by mixing nitrite salts with acid to yield the desired pH. Once the nitrite and acid are mixed, a series of complex reactions occurs involving the resulting oxides of nitrogen. Within a short time, the active species of "nitrous acid" undergoes changes that result in the loss the reagent's ability to cleave the glycosidic bonds of the amino sugar (Shively and Conrad, 1976). Consequently, the nitrous acid reagents must be used within a few minutes after preparation. This is of less concern for the pH 4 reagent than for the pH 1.5 reagent, because the pH 4 reagent has a higher concentration of nitrite.

Also, pH is important for maintaining the selectivity of the cleavage. Although there is good selectivity for N-unsubstituted GlcNs and N-sulfated GlcNs at pH 4 and pH 1.5, respectively, the glycosidic bonds of both types of GlcN residues are cleaved at pHs between these values. Thus, it is desirable to let the reactions at the respective pHs proceed only for 10 to 15 min. Also, when samples are derived from buffered solutions, it is necessary to check the pH with pH paper before addition of the nitrous acid reagent. In fact, it is desirable to dialyze the glycosaminoglycan solution to remove all salts before beginning the cleavage step. Another reason for the pre-dialysis is that the NaB[³H]₄ reagent is catalytically destroyed by oxyanions, such as PO₄³⁻ (Conrad et al., 1973).

Anticipated Results

When samples containing mixtures of different glycosaminoglycans are treated with nitrous acid at pH 1.5, only heparin and heparan SO₄ are cleaved; thus, the total amount of glycosaminoglycan that is reduced in molecular size by pH 1.5 nitrous acid represents the heparin/heparan sulfate family. Just as the nitrous acid cleavage is diagnostic for the heparin/heparan sulfate family, the chondroitin sulfate/dermatan sulfate/hyaluronic acid family is specifically cleaved by chondroitinases (or chondroitin lyases) ABC or AC (*UNIT 17.13B*). These selective cleavages can be used for characterization of glycosaminoglycans that are metabolically labeled (e.g., with [³⁵S]O₄²⁻ or [³H]glucosamine). The use of nitrous acid and

the chondroitinases for more definitive characterization of glycosaminoglycans is described in *UNIT 17.22B*.

Time Considerations

The nitrous acid cleavage is complete within a few minutes after the reagent is added to the sample. When hydrazinolysis precedes the nitrous acid cleavage, an additional 6 to 8 hr is required for the hydrazinolysis step and preparation of the sample for nitrous acid cleavage. The overall procedure can be interrupted after hydrazinolysis, nitrous acid cleavage, or $\text{NaB}[\text{}^3\text{H}]_4$ reduction without any loss of precision in the analytical stage.

Literature Cited

Conrad, H.E., Varboncoeur, E., and James, M.E. 1973. Qualitative and quantitative analysis of reducing carbohydrates by radiochromatography on ion exchange papers. *Anal. Biochem.* 51:486-500.

Shively, J.E., and Conrad, H.E. 1976. Formation of anhydrosugars in the chemical depolymerization of heparin. *Biochemistry* 15:3932-3942.

Key Reference

Guo, Y. and Conrad, H.E. 1989. The disaccharide composition of heparins and heparan sulfates. *Anal. Biochem.* 176:96-104.

Shaklee, P.N. and Conrad, H.E. 1984. Hydrazinolysis of heparin and other glycosaminoglycans. *Biochemistry* 217:187-197.

Shaklee, P.N. and Conrad, H.E. 1986. The disaccharides formed by deaminative cleavage of N-deacetylated glycosaminoglycans. *Biochem. J.* 235:225-236.

These references describe the basic procedures in this unit and the background for those procedures.

Contributed by H. Edward Conrad
University of Illinois
Urbana, Illinois

Analysis of Disaccharides and Tetrasaccharides Released from Glycosaminoglycans

Glycosaminoglycans (GAGs) are converted to disaccharides by cleavage with lyases or by treatment with nitrous acid either before or after N-deacetylation. Enzymatic (lyase) cleavage of GAGs to disaccharides is described in *UNIT 17.13B*. Nitrous acid cleavage of GAGs to release disaccharides (or tetrasaccharides, from heparin and heparan sulfate), and the $\text{NaB}[^3\text{H}]_4$ reduction of these products is described in *UNIT 17.22A*.

As described in this unit, individual disaccharides are separated by paper chromatography or paper electrophoresis (Basic Protocol 1) or high-performance liquid chromatography (HPLC; Basic Protocol 2). Lyase-released disaccharides can also be monitored by UV absorbance as described in Background Information. The disaccharides can be radiolabeled for subsequent detection either by metabolic labeling of the cells that produce them (*UNIT 17.4*) or by reduction of the disaccharides from unlabeled GAGs with $\text{NaB}[^3\text{H}]_4$ to incorporate one gram atom of ^3H per mole of disaccharide (*UNIT 17.22A*). The label incorporated into the separated disaccharides can then be used for qualitative or quantitative assay of each disaccharide.

Support Protocol 1 describes mild conditions for borohydride reduction of alkali-labile disaccharides obtained by cleavage of GAGs with lyases. Support Protocol 2 describes scintillation counting of samples obtained from HPLC separation of radiolabeled saccharides, and Support Protocol 3 describes calculations for the quantitation of disaccharides.

BASIC PROTOCOL 1

ANALYSIS OF DISACCHARIDES AND OLIGOSACCHARIDES FROM GLYCOSAMINOGLYCANS BY PAPER CHROMATOGRAPHY AND PAPER ELECTROPHORESIS

Materials

- Sample of lyase-degraded (*UNIT 17.13B*; also see Support Protocol 1) or nitrous acid-degraded glycosaminoglycan (*UNIT 17.22A*)
- Paper chromatography or paper electrophoresis system appropriate to saccharide mixture to be analyzed (see Table 17.22B.1 and Table 17.22B.2; also see recipe in Reagents and Solutions)
- Apparatus for paper electrophoresis

NOTE: [^{14}C]glucose can be added as an internal standard. See *UNIT 17.22A* for details.

1. Run descending paper chromatograms and paper electrophoretograms on 1×22 -in. strips of chromatography paper in the appropriate solvent system for the saccharide mixture to be analyzed. Spot the desired volume of each sample on a 0.5-in. segment of the strip that has been marked off in advance. For paper electrophoresis, dampen the paper strips with the solvent up to the point where the sample was spotted and electrophorese the samples toward the anode at 25 V/cm for 2.5 hr.
2. Count radioactivity on 0.5-in. segments from the paper strips.

Table 17.22B.1 Separation of Disaccharides Released from Chondroitin Sulfate, Dermatan Sulfate, and Keratan Sulfate by Hydrazinolysis/Nitrous Acid Treatment^a

Compound ^b	R _{glucitol} with paper chromatography system ^c		V _E with paper electrophoresis system ^d		HPLC system ^e			
	1	2	1	2	SAX		WAX	
					Eluant (mM KH ₂ PO ₄)	Ret. time (min)	Eluant (mM KH ₂ PO ₄)	Ret. time (min)
Chondroitin SO₄ / Dermatan SO₄ Products								
GlcA-ATal _R	0.63	0.46	0.45	0.19	5	17.0	5	25.0
IdoA-ATal _R	0.77	0.61	0.41	0.18	5	16.5	5	25.0
ATal _R (4-SO ₄)	0.68	3.29	0.70	0.76	5	19.0		
ATal _R (6-SO ₄)	0.61	2.38	0.67	0.80	5	22.5		
GlcA-ATal _R (6-SO ₄)	0.23	0.42	0.78	0.62	40	24.5	25	32.5
IdoA-ATal _R (6-SO ₄)	0.31	0.58	0.74	0.64	40	25.0	25	34.5
GlcA-ATal _R (4-SO ₄)	0.32	0.67	0.72	0.58	40	26.5	25	38.5
IdoA-ATal _R (4-SO ₄)	0.38	0.67	0.69	0.62	40	27.0	25	43.5
IdoA-ATal _R (4,6-diSO ₄)	0.15	0.36	0.93	0.92	180	24.0	135 115	27.5 35.5
GlcA-ATal _R (4,6-diSO ₄)	0.36					135 115	23.5 30.5	
ATal _R (4,6-diSO ₄)	0.27	1.64	1.15	1.23				
IdoA(2-SO ₄)-ATal _R (4-SO ₄)	0.15	0.61	0.96	1.0	180	19.5	135 115	30.0 40.0
GlcA(2-SO ₄)-ATal _R (4-SO ₄)	0.61						135 115	25.0 33.5
GlcA(2-SO ₄)-ATal _R (6-SO ₄)	0.11	0.33	1.03	0.93	180	10.0	135 115	17.5 23.5
IdoA(2-SO ₄)-ATal _R (6-SO ₄)	0.33					135 115	21.0 27.5	
Keratan SO₄ Products								
AMan _R (6-SO ₄)	0.69	2.38	0.72	0.81	5	22.0		
Gal _R (6-SO ₄)	0.42	1.46	0.64	0.71	5	26.0		
Gal-AMan _R	0.84	1.24	0.11	0.13				
Gal(6-SO ₄)-AMan _R	0.38	1.54	0.52	0.62	5	17.0	5	21.0
Gal-AMan _R (6-SO ₄)	0.38	1.50	0.54	0.63	5	20.0	5	22.5
Gal(6-SO ₄)-AMan _R (6-SO ₄)	0.11	1.0	0.86	1.04	70	12.5		

^aAbbreviations: AMan, anhydro-D-mannose; ATal, anhydro-D-talose; Gal, D-galactose; GlcA, D-glucuronic acid; GlcN, D-glucosamine; IdoA, L-iduronic acid; ret. time, retention time; SAX, strong ion-exchange; WAX, weak ion-exchange.

^bAll values are for aldehyde-reduced sugars, indicated by the subscript, R.

^cR_{glucitol} = ratio of migration distance of compound to migration distance of the [¹⁴C]glucitol internal standard. See recipes for paper chromatography systems 1 and 2 in Reagents and Solutions.

^dV_e = ratio of migration distance of compound to migration distance of αIdoA(2-SO₄) 1-4 AMan_R(6-SO₄). This compound is the major disaccharide obtained from heparin; see Table 17.22.3. See recipes for paper electrophoresis systems 1 and 2 in Reagents and Solutions; see Shaklee and Conrad, 1986 for details on paper chromatography.

^eHPLC separations are performed at a flow rate of 1 ml/min under the isocratic salt conditions indicated; see Shaklee and Conrad (1986) for details.

Table 17.22B.2 Separation of Dissaccharides Released from Chondroitin Sulfate by Lyase Treatment^a

Compound	R_{glucitol} with paper chromatography system ^b		HPLC system ^c	
	1	2	SAX retention time (min)	WAX
$\Delta\text{Di-OS}_R$	1.0	0.95		
$\Delta\text{Di-4S}_R$	0.59	0.67	32.0	49.0
$\Delta\text{Di-6S}_R$	0.56	0.82	44.0	49.0
$\Delta\text{Di-diS}_{BR}$	0.26	0.46	61.0	67.0
$\Delta\text{Di-diS}_{DR}$	0.30		66.5	66.5
$\Delta\text{Di-diS}_{ER}$	0.26	0.41	68.0	70.0

^aAbbreviations are: $\Delta\text{Di-OS}_R$, 2-acetamido-2-deoxy-3-*O*-(β -D-glucopyranosyluronic acid)-D-galactitol; $\Delta\text{Di-4S}_R$, 2-acetamido-2-deoxy-3-*O*-(β -D-glucopyranosyluronic acid)-4-*O*-sulfo-D-galactitol; $\Delta\text{Di-6S}_R$, 2-acetamido-2-deoxy-3-*O*-(β -D-glucopyranosyluronic acid)-6-*O*-sulfo-D-galactitol; $\Delta\text{Di-diS}_{BR}$, 2-acetamido-2-deoxy-3-*O*-(2-*O*-sulfo- β -D-glucopyranosyluronic acid)-4-*O*-sulfo-D-galactitol; $\Delta\text{Di-diS}_{DR}$, 2-acetamido-2-deoxy-3-*O*-(2-*O*-sulfo- β -D-glucopyranosyluronic acid)-6-*O*-sulfo-D-galactitol; $\Delta\text{Di-diS}_{ER}$, 2-acetamido-2-deoxy-3-*O*-(β -D-glucopyranosyluronic acid)-4,6-di-*O*-sulfo-D-galactitol; SAX, strong ion-exchange; WAX, weak ion-exchange.

^b R_{glucitol} = ratio of migration distance of compound to migration distance of [¹⁴C]glucitol internal standards.

^cThe HPLC gradient for both SAX and WAX columns consisted of a 30-min isocratic segment at 40 mM KH₂PO₄, followed by a 40-min linear gradient of 400 mM KH₂PO₄. HPLC separations are performed at a flow rate of 1 ml/min.

**BASIC
PROTOCOL 2**

**ANALYSIS OF DISACCHARIDES AND OLIGOSACCHARIDES FROM
GLYCOSAMINOGLYCANS BY HPLC**

Samples to be analyzed by high-performance liquid chromatography (HPLC) are obtained by nitrous acid cleavage of N-deacetylated glycosaminoglycans at pH 4 (UNIT 17.22A), by direct cleavage of heparin or heparan sulfate with nitrous acid at pH 1.5 (UNIT 17.22A), or by enzymatic cleavage with lyases (UNIT 17.13B). For HPLC, all samples must be free of particulate matter and should be dissolved in the initial solvent used to elute the HPLC column.

Materials

Sample: mixture of saccharides obtained from glycosaminoglycan by nitrous acid cleavage (UNIT 17.22A) or lyase cleavage (UNIT 17.13B)

4.5-mm × 25-cm Partisil SAX strong anion-exchange column (Whatman) or 4-mm × 30-cm Varian MicroPak AX-5 weak anion-exchange column (Varian Analytical)

HPLC solvents (see recipe and Tables 17.22B.1-17.22B.5)

Gradient solutions for HPLC (Tables 17.22B.1-17.22B.5)

Bio-Gel P-10 gel filtration column (Bio-Rad)

4.6 × 250-mm Hi-Chrom S-5 ODS C-18 column (Regis Technology)

Fraction collector or in-line radioactivity flow detector (e.g., Packard Instrument)

Additional reagents and equipment for gel-filtration chromatography (UNIT 10.9), ion-exchange HPLC (UNIT 10.13), and reversed-phase HPLC (UNIT 10.12)

**Analysis of
Disaccharides and
Tetrasaccharides
Released from
Glycosamino-
glycans**

17.22.8

For disaccharides obtained from chondroitin sulfate, dermatan sulfate, and keratan sulfate by nitrous acid cleavage

- 1a. Run HPLC on a 4.5-mm × 25-cm Whatman Partisil SAX strong ion-exchange column or 4-mm × 30-cm Varian MicroPak AX-5 weak anion exchange column using the appropriate isocratic elution conditions (see Table 17.22B.1). Proceed to step 3.

For disaccharides obtained from chondroitin sulfate and dermatan sulfate by lyase cleavage

- 1b. Reduce the disaccharide mixture with borohydride as in Support Protocol 1.
- 2b. Run HPLC on a 4.5-mm × 25-cm Whatman Partisil SAX strong ion-exchange column or 4-mm × 30-cm Varian MicroPak AX-5 weak anion exchange column using the appropriate elution conditions.

The maximum sample volume for injection is 500 μl for both nitrous acid- and lyase-cleaved GAGs.

For strong ion-exchange separation of di- and tetrasaccharides obtained from heparin and heparan sulfate by nitrous acid cleavage

- 1c. Prior to chromatographic separation of individual oligosaccharides, separate the oligosaccharides according to size by gel-filtration chromatography (UNIT 10.9) on a Bio-Gel P-10 column in 1 M Na₂CO₃.
- 2c. Analyze di- and tetrasaccharides on a 4.5-mm × 25-cm Whatman Partisil SAX strong ion-exchange column using step gradients at a flow rate of 1 ml/min (see Table 17.22B.3). For disaccharide separation, use an initial 40 mM KH₂PO₄ isocratic segment for 25 min to resolve unsulfated and monosulfated disaccharides, followed by a 1-min linear step gradient to 185 mM KH₂PO₄, and a 40-min isocratic segment at 185 mM KH₂PO₄ to resolve the disulfated disaccharides. For samples with high salt content, it is advisable to run an initial 10-min isocratic segment with 20 mM KH₂PO₄ to wash the excess salt through the column, thus avoiding any possible effect of the high salt on the elution profiles. For tetrasaccharide separation, use a step gradient with five isocratic segments: 40 mM KH₂PO₄ for 25 min; 100 mM KH₂PO₄ for 35 min; 175 mM KH₂PO₄ for 30 min; 360 mM KH₂PO₄ for 25 min; and 400 mM KH₂PO₄ for 25 min. Connect each isocratic segment by a 1-min step gradient from one isocratic segment to the next.

For reversed-phase ion-pairing HPLC separation of di- and tetrasaccharides obtained from heparin and heparan sulfate by nitrous acid cleavage

- 1d. Prior to chromatographic separation of individual oligosaccharides, separate the oligosaccharides according to size by gel-filtration chromatography (UNIT 10.9) on a Bio-Gel P-10 column in 1 M Na₂CO₃.
- 2d. Analyze di- and tetrasaccharides on a 4.6 × 250-mm Hi-Chrom S-5 ODS column at a flow rate of 1 ml/min (Tables 17.22B.4 and 17.22B.5). Isocratic separations of groups of monosaccharides and disaccharides are obtained by isocratic elution with the solvent mixtures shown in Table 17.22B.4. Connect the isocratic segments shown in Table 17.22B.4 as in Table 17.22B.5 to achieve the separation of all of the disaccharides in a single run. Note that GlcA-AManR(SO₄) and IdoA-AManR(SO₄) are not separated under these conditions; for separation of these two disaccharides, the strong-anion exchange HPLC must be used. Isocratic separations of groups of tetrasaccharides are obtained by isocratic elution with the solvent mixtures shown in Table 17.22B.4. Connect the isocratic segments shown in Table 17.22B.4 as in Table 17.22B.5 to achieve the separation of all of the tetrasaccharides.

For details of the reversed-phase ion-pairing procedure, see Guo and Conrad (1988).

Table 17.22B.3 Separation by Strong Ion-Exchange HPLC of Oligosaccharides Released from Heparin by Nitrous Acid Treatment With or Without Prior Hydrazinolysis^{a,b}

Oligosaccharide	Eluant (mM KH ₂ PO ₄)	Retention time (min)
Disaccharides		
GlcA-AMan _R and IdoA AMan _R	40	4.5
AMan _R	40	4.5
AMan _R (SO ₄)	40	7.0
GlcA(SO ₄)-AMan _R	40	19.5
GlcA-AMan _R (SO ₄)	40	23.0
IdoA-AMan _R (SO ₄)	40	26.5
IdoA(SO ₄)-AMan _R	40	30.0
GlcA(SO ₄)-AMan _R (SO ₄)	185	14.0
IdoA(SO ₄)-AMan _R (SO ₄)	185	21.5
GlcA-AMan _R (3,6diSO ₄)	185	25.5
Tetrasaccharides^c		
t1 GlcA-GlcNAc-GlcA-AMan _R	20	28
t2 IdoA-GlcNAc-GlcA-AMan _R	20	33.5
t3 GlcA-GlcNAc(SO ₄)-GlcA-AMan _R	100	31.5
t4 IdoA(SO ₄)-GlcNAc-GlcA-AMan _R	100	37.0
t5 IdoA-GlcNAc(SO ₄)-GlcA-AMan _R	100	42.5
t6	200	25.0
t7 IdoA(SO ₄)-GlcNAc-GlcA-AMan _R (SO ₄)	200	30.0
t8 IdoA-GlcNAc(SO ₄)-GlcA-AMan _R (SO ₄)	200	32.5
t9 IdoA-GlcNAc(SO ₄)-GlcA-AMan _R (3-SO ₄)	185	53.0
t10	320	23.0
t11	320	30.0
t12	320	32.0
t13	350	23.5
t14 IdoA-GlcNAc(SO ₄)-GlcA-AMan _R (3,6-diSO ₄)	350	38.5
t15 IdoA(SO ₄)-GlcNAc(SO ₄)-GlcA-AMan _R (SO ₄)	350	42.0
t16	400	30.0

^aAbbreviations: AMan, anhydro-D-mannose; GlcA, D-glucuronic acid; GlcN, D-glucosamine; GlcNAc, N-acetyl-D-glucosamine; IdoA, L-iduronic acid.

^bHPLC separations are obtained using step gradients. Details of gradients used for disaccharides and oligosaccharides are given in Basic Protocol 2; the KH₂PO₄ concentrations are those at which each oligosaccharide emerges during the gradient. Columns are run at a flow rate of 1 ml/min; see Guo and Conrad (1989) for details.

^cTetrasaccharide designations (t1-t16) are described in Bienkowski and Conrad, 1985. Those tetrasaccharides for which no monosaccharide sequences are given (t6, t10-t13, and t16) are "ring-contraction tetrasaccharides," which are formed in relatively low yields (see Commentary).

Table 17.22B.4 Separation by Isocratic Ion-Pairing HPLC of Oligosaccharides Released from Heparin by Nitrous Acid Treatment, With or Without Prior Hydrazinolysis^a

Oligosaccharide	Solvent ^b	Retention time ^c (min)
Mono- and Disaccharides		
AMan _R	A:B = 94:6	5.0
GlcA-AMan _R	A:B = 94:6	6.5
IdoA-AMan _R	A:B = 94:6	9.0
AMan _R (SO ₄)	A:B = 94:6	12.0
IdoA(SO ₄)-AMan _R	A:B = 94:6	17.0
GlcA(SO ₄)-AMan _R	A:B = 86:14	22.0
GlcA-AMan _R (SO ₄)	A:B = 86:14	22.0
IdoA-AMan _R (SO ₄)	A:B = 86:14	31.0
GlcA-AMan _R (3,6diSO ₄)	A:B = 86:14	23.5
IdoA(SO ₄)-AMan _R (SO ₄)	A:B = 61:39	26.0
GlcA(SO ₄)-AMan _R (SO ₄)	A:B = 61:39	31.5
Tetrasaccharides^d		
t1 GlcA-GlcNAc-GlcA-AMan _R	A:C = 94:6	9.0
t2 IdoA-GlcNAc-GlcA-AMan _R	A:C = 94:6	9.5
t4 IdoA(SO ₄)-GlcNAc-GlcA-AMan _R	A:C = 84:15	29.0
t5 IdoA-GlcNAc(SO ₄)-GlcA-AMan _R	A:C = 84:15	33.0
t3 GlcA-GlcNAc(SO ₄)-GlcA-AMan _R	A:C = 84:15	38.0
t9 IdoA-GlcNAc(SO ₄)-GlcA-AMan _R (3-SO ₄)	A:C = 74:26	29.0
t8 IdoA-GlcNAc(SO ₄)-GlcA-AMan _R (SO ₄)	A:C = 74:26	33.0
t13 IdoA(SO ₄)-RC(SO ₄)IdoA(SO ₄)-AMan _R	A:C = 60:40	12.0
t14 IdoA-GlcNAc(SO ₄)-GlcA-AMan _R (3,6-diSO ₄)	A:C = 60:40	15.0
t16 IdoA-RC(SO ₄)IdoA(SO ₄)-AMan _R (SO ₄)	—	16.0

^aAbbreviations: AMan, anhydro-D-mannose; GlcA, D-glucuronic acid; GlcN, D-glucosamine; IdoA, L-iduronic acid; RC, ring-contraction product.

^bHPLC separations are obtained using a C-18 reversed-phase column using isocratic elution conditions obtained by mixing solvents A, B, and C (see recipes in Reagents and Solutions) in the ratios shown.

^cAll elution are performed at a flow rate of 1 ml/min. For separations of di- or tetrasaccharide mixtures containing the total mixtures of these oligosaccharides, see gradient conditions in Table 17.22.5.

^dTetrasaccharide designations are described in Bienkowski and Conrad, 1985.

Table 17.22B.5 Gradient Conditions for Oligosaccharide Separation by Reversed-Phase Ion-Pairing HPLC^a

Disaccharides		Tetrasaccharides	
Time interval for solvent change (min)	% Solvent B in time interval	Time interval for solvent change (min)	% Solvent C
0→5	6	0→13	6
5→6	6→14	13→14	6→11
6→20	14	14→64	11
20→40	14→20	64→65	11→25
40→41	20→39	65→110	25
41→81	39	110→111	25→31
81→83	39→60	111→156	31 83→100
		157→192	39
		192→193	39→70
		193→220	70

^aOligosaccharides are chromatographed on a C-18 reversed-phase column using the ion-pairing systems described in Basic Protocol 2. Mixtures of standards are chromatographed with increasing percentages of solvent B (disaccharides) or solvent C (tetrasaccharides) in solvent A (see recipes for solvents A, B, and C in Reagents and Solutions). Details of the gradients used are described in Basic Protocol 2.

- Count radioactivity in 5-ml fractions collected in scintillation vials (see Support Protocol 2) or use an in-line radioactivity flow detector. Identify disaccharides on the basis of their chromatographic migration properties by referring to Tables 17.22B.1-17.22B.5. Quantify di- and tetrasaccharides (see Support Protocol 3).

**SUPPORT
PROTOCOL 1**

**BOROHYDRIDE REDUCTION OF ALKALI-LABILE DISACCHARIDES
OBTAINED BY CLEAVAGE WITH LYASES**

Disaccharides formed by cleavage of glycosaminoglycans with specific lyases (*UNIT 17.13B*) can be analyzed by reduction with NaB^3H_4 and separation of the labeled products either by paper chromatography (Basic Protocol 1; Glaser and Conrad, 1979) or HPLC (Basic Protocol 2; Delaney et al., 1980; Shaklee and Conrad, 1986). However, the lyase-generated disaccharides are quite labile under the alkaline conditions required for their reduction with NaB^3H_4 . Thus, a milder procedure than the 50°C reduction procedure described in *UNIT 17.22A* must be used for the reduction. These labile disaccharides (or higher oligosaccharides that may be obtained when the lyase does not convert the glycosaminoglycan completely to disaccharides) can be reduced at 0°C without cleavage.

Materials

Sample of lyase-degraded glycosaminoglycan
1 M Na_2CO_3 , pH 9.0 (ice-cold)
Sodium borohydride reagent (see recipe)
3 M H_2SO_4
6 × 150-mm test tubes

- Place lyase-degraded GAG sample in a 6 × 150-mm test tube, dry sample in a stream of air, then chill tube in an ice bath.
- Add 5 μl ice-cold 1 M Na_2CO_3 (pH 9.0) to sample, then add 10 μl of sodium borohydride reagent. Cork or cap the tube and let it sit on ice 90 min.

IMPORTANT NOTE: If samples are already labeled by metabolically labeling of the GAG, 0.5 M cold NaBH_4 is used in preparing the sodium borohydride reagent instead of 0.5 M $\text{NaB}[^3\text{H}]_4$.

Carry out all $\text{NaB}[^3\text{H}]_4$ manipulations in a well-ventilated fume hood designated for radioactive use.

- Add 2.5 μl of 3 M H_2SO_4 to destroy excess NaB^3H_4 and let sample stand ~30 min.
- Neutralize sample with 1 M Na_2CO_3 and dry it in a stream of air. Redissolve sample in water, then dry again to remove excess $[\text{H}^3]_2$ gas. Finally dissolve sample in 50 μl water.

Aliquots of the sample may be analyzed by paper chromatography (Basic Protocol 1) or HPLC (Basic Protocol 2).

SCINTILLATION COUNTING OF FRACTIONS FROM HPLC ANALYSIS OF SACCHARIDES RELEASED FROM GLYCOSAMINOGLYCANS

SUPPORT
PROTOCOL 2

Effluent from the HPLC column is collected in 0.5-ml fractions in scintillation vials using a fraction collector. Alternatively the effluent may be run directly through an in-line radioactivity flow detector (e.g., Radiomatic). If the total cpm in each peak are high enough to obtain good counting statistics in short counting intervals, a stream splitter may be used to collect the bulk of each fraction (e.g., 0.4 ml) in a test tube while only a portion of each fraction (e.g., 0.1 ml) is taken in scintillation vials for counting. Add an appropriate volume (4 to 5 ml) of any commercial scintillation fluid designed to accommodate the salt and solvent concentrations in the eluants. Alternatively, add 4 ml of a scintillation cocktail containing 3 g diphenyloxazole in a mixture of 250 ml Triton X-114, 750 ml xylene, and 10 ml concentrated HCl. For KH_2PO_4 concentrations >10 mM, add 1 drop concentrated HCl to the xylene-based scintillation fluid to redissolve the salt precipitate that begins to form in the scintillation vial.

CALCULATIONS FOR QUANTITATION OF DISACCHARIDES

SUPPORT
PROTOCOL 3

Procedures described here allow calculation of the number of μmol of each disaccharide in the original sample taken for analysis. These calculations apply primarily to analyses in which disaccharides obtained from unlabeled glycosaminoglycans are mixed with a $[^{14}\text{C}]\text{D}$ -glucose internal standard and reduced with NaB^3H_4 (see UNIT 17.22A), but can also be applied to metabolically labeled materials if the specific radioactivity of the radiolabeled precursor that appears in the metabolically labeled glycosaminoglycan is determined independently.

Calibrate NaB^3H_4 used in the analysis of unknowns

Because the counting efficiency for ^3H varies in each type of counting experiment, cpm/ μmol of a standard reducing sugar must be determined separately for each type of separation procedure. Mix a measured amount of a standard sugar (usually D-glucose) with the $[^{14}\text{C}]\text{D}$ -glucose internal standard (see UNIT 17.22A). Enough $[^{14}\text{C}]\text{glucose}$ should be added so that $\sim 1 \times 10^5$ ^{14}C cpm will be recovered in the final separation step. Reduce an aliquot of this mixture with the same NaB^3H_4 reagent to be used in the analysis (see UNIT 17.22A). If NaB^3H_4 with a specific activity of ~ 500 mCi/mmol (recommended) is used, the $[^3\text{H}]\text{D}$ -glucitol formed by reduction of the D-glucose should have a specific activity of 125 mCi/mmol, and 1 μmol of the standard D-glucose should yield 778×10^6 dpm per μmol of $[^3\text{H}]\text{glucitol}$. Counting efficiency for ^3H in some of the separation steps may be as low as 5%. Chromatograph the resulting $[^3\text{H}]\text{D}$ -glucitol/ $[^{14}\text{C}]\text{glucitol}$ mixture in the same type of analytical separation procedure that is being used for the unknowns (paper chromatography or HPLC). Determine the number of $[^{14}\text{C}]\text{glucitol}$ and $[^3\text{H}]\text{glucitol}$ cpm in the separated peak by scintillation counting under conditions optimized for dual-label counting. The ^3H cpm must be corrected for the number of ^{14}C cpm that spill over into the ^3H counting channel. From the number of $[^{14}\text{C}]\text{glucitol}$ cpm, calculate the number of μmol of D-glucitol in the peak as described below.

When a known number of μmol of glucose are mixed with a known number of cpm of the $[^{14}\text{C}]\text{glucose}$ internal standard, the number of ^{14}C cpm recovered in the analytical stage (e.g., in paper chromatography or HPLC) can be used to calculate the proportion of original μmol of the standard glucose that are recovered in the analytical stage. For example, if the original D-glucose solution contained 1×10^5 μmol of the D-glucose standard and 5×10^5 cpm of $[^{14}\text{C}]\text{glucose}$, and if 1×10^5 cpm of $[^{14}\text{C}]\text{glucitol}$ were recovered in the analytical run, then 2×10^4 μmol of the D-glucitol standard would be recovered in the analytical run. Thus, if 1×10^9 ^3H cpm in the $[^3\text{H}]\text{glucitol}$ were recovered in the analytical run, the number of ^3H cpm/ μmol for glucose would be 5×10^4 .

Preparation and
Analysis of
Glycoconjugates

17.22.13

Using the [^3H]glucitol cpm in the separated peak and the number of μmol of D-glucitol in the peak, calculate the number of [^3H]glucitol cpm/ μmol , which will be the same as the number of ^3H cpm/ μmol for all mono-, di- or tetrasaccharides obtained in the analytical run, as all reducing sugars incorporate the same number of equivalents of ^3H per μmol from the $\text{NaB}[^3\text{H}]_4$. Consequently, when the total number of ^3H cpm is determined for any peak in the analytical run, division of this number by the ^3H cpm/ μmol of the D-glucose standard yields the value for the number of μmol of the compound in the peak.

Calibrate internal standard [^{14}C]glucose

Because the counting efficiency for ^{14}C varies in each type of counting experiment, the number of ^{14}C cpm in a given aliquot of the [^{14}C]glucose internal standard must be determined separately for each type of separation procedure. Reduce an aliquot of the [^{14}C]glucose internal standard using the same reduction conditions used for the unknowns (see UNIT 17.22A for nitrous acid cleavage products; see Basic Protocol 2 for lyase cleavage products). To calibrate [^{14}C]glucose for paper chromatography, transfer the resulting [^{14}C]glucitol solution quantitatively to a paper chromatography strip and run the chromatogram in the solvent for which the [^{14}C]glucose calibration is being made. Develop the chromatogram, cut the strip into 0.5-in. segments for scintillation counting, then determine the total number of ^{14}C cpm in the original aliquot taken for analysis.

To calibrate [^{14}C]glucose for HPLC, inject a measured aliquot of the [^{14}C]glucitol solution quantitatively into an HPLC column injector, elute the column as in a typical analytical run, then determine the number of ^{14}C cpm in the [^{14}C]glucitol peak. The [^{14}C]glucose will emerge in the void volume of the column in the HPLC procedures described here.

Chromatographic and electrophoretic properties of various glycosaminoglycan oligosaccharides are listed in Tables 17.22B.1-17.22B.5.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

HPLC solvents

For anion-exchange chromatography: HPLC solvents for chromatography on weak anion-exchange or strong ion-exchange columns range from 0 mM to 400 mM KH_2PO_4 . Degas and filter 1 liter distilled water by gently heating it to 80°C in a 2-liter vacuum-filtration flask, sealing the flask with a stopper, then placing it in an ice-water bath and evacuating with a water aspirator while gently stirring with a magnetic stirrer. Continue aspirating until no further bubbles are apparent (15 min), then transfer to a volumetric flask containing the amount of KH_2PO_4 required to give the desired concentration. Gently stir the solution, then filter through a 0.45- μm Millipore filter under vacuum into a clean 500-ml bottle, which will serve as a reservoir for the solvent.

For HPLC systems that allow continuous degassing during the run, it is not necessary to degas solvents before making the solutions.

For reversed-phase ion-pairing chromatography: Prepare solvents A, B, and C as follows. Mix or deliver in gradient as described in Table 17.22B.4 and Table 17.22B.5.

Solvent A: Prepare 38 mM $\text{NH}_4\text{H}_2\text{PO}_4$, 2 mM H_3PO_4 , and 1 mM tetrabutylammonium phosphate in water. Adjust pH to 3.6 with 0.1 M NaOH or HCl.

Solvent B: Prepare 38 mM $\text{NH}_4\text{H}_2\text{PO}_4$, 2 mM H_3PO_4 , and 1 mM tetrabutylammonium phosphate in 30/70 (v/v) acetonitrile/water. Adjust pH to 4.2 with 0.1 M NaOH or HCl.

continued

Solvent C: Prepare 38 mM $\text{NH}_4\text{H}_2\text{PO}_4$, 2 mM H_3PO_4 , and 1 mM tetrabutylammonium phosphate in 50:50 (v/v) acetonitrile /water. Adjust pH to 4.4 with 0.1 M NaOH or HCl.

Paper chromatography and paper electrophoresis systems

Use the following paper chromatography or paper electrophoresis systems in conjunction with the data in Table 17.22B.1 or Table 17.22B.2:

Paper chromatography system 1: 3:2:1 (v/v/v) 1-butanol/glacial acetic acid/1 M NH_4OH . Use 1 × 22-in. Whatman no. 3 paper chromatography strips.

Paper chromatography system 2: 3:2:1 (v/v/v) ethyl acetate/pyridine/5 mM boric acid. Use 1 × 22-in. Whatman cellulose phosphate paper chromatography strips.

Paper electrophoresis system 1: 50 ml glacial acetic acid and 10 ml pyridine in 4 liters water (pH 4.0). Use 1 × 22-in. cellulose phosphate paper chromatography strips.

Paper electrophoresis system 2: 350 ml of glacial acetic acid and 100 ml of 88% formic acid in 4 liters water (pH 1.7). Use 1 × 22-in. cellulose phosphate paper chromatography strips.

Sodium borohydride reagent

Just prior to use, dissolve 1.89 mg $\text{NaB}[^3\text{H}]_4$ in 100 μl chilled 0.2 M Na_2CO_3 , pH 10.2, to give a 0.5 M $\text{NaB}[^3\text{H}]_4$ solution. Keep on ice until use.

COMMENTARY

Background Information

A number of high-performance liquid chromatography procedures have been developed to separate and quantitate the disaccharides formed by deaminative cleavage of N-deacetylated glycosaminoglycans with nitrous acid (*UNIT 17.22A*) or cleavage with lyases (*UNIT 17.13B*). These include chromatography on strong anion-exchange columns (SAX; Bienkowski and Conrad, 1985); weak anion exchange columns (WAX; Shaklee and Conrad, 1986); and reversed-phase ion-pairing columns (Guo and Conrad, 1988). Alternatively, some of the disaccharides can be separated by paper chromatography (Hopwood and Elliot, 1983; Hopwood and Muller, 1983), paper electrophoresis (Hopwood and Elliot, 1983; Hopwood and Muller, 1983), thin-layer chromatography (Edge and Spiro, 1985), or capillary-zone electrophoresis (Al-Hakim and Linhardt, 1991; Desai et al., 1993). When lyases are used to cleave the glycosaminoglycan (*UNIT 17.13B*), the lyase-derived di- and oligosaccharides can be assayed by the absorbance of their unsaturated uronic acid residues at 233 nm. This UV assay has been used for detection of oligosaccharides eluted from SAX HPLC columns with a linear sodium chloride gradient running from 0 to 1 M (Rice et al., 1985; Linhardt et al., 1988). Glycosaminoglycans that have been metabo-

lically labeled with ^3H D-glucosamine or $^{35}\text{S}]\text{O}_4^{2-}$ yield disaccharides with the corresponding labels. Disaccharides from both types of cleavage can also be detected and quantified following their reduction with $\text{NaB}[^3\text{H}]_4$ to incorporate one gram atom of ^3H per mol of disaccharide. In this approach, the elution positions of the ^3H -labeled disaccharide peaks separated by HPLC, paper chromatography, or paper electrophoresis are used to identify the disaccharides, and the ^3H cpm recovered in each peak are used to obtain ratios of disaccharides or quantitative measures of the amounts of each disaccharide present in the glycosaminoglycan.

Critical Parameters

The important parameters for the separations described in this unit are those that must normally be observed for chromatographic separations. The critical parameters for the analytical procedures are primarily those involved in the treatment with nitrous acid and $\text{NaB}[^3\text{H}]_4$. These are described in *UNIT 17.22A*. It may be noted that the elution times given in Tables 17.22B.1-17.22B.5 may vary somewhat, but the relative positions of elution remain the same. Positive identification may require chromatographic runs of mixtures of previously characterized disaccharides before or

after the sample run. As columns age, the elution times become shorter and resolution becomes poorer. When resolution becomes unsatisfactory, the column must be replaced.

Anticipated Results

The nitrous acid reaction (see UNIT 17.22A) is virtually stoichiometric for β -linked amino sugars (Shively and Conrad, 1970). Consequently, the yield of disaccharides is stoichiometric from those glycosaminoglycans that contain β -linked amino sugars (i.e., chondroitin sulfate, dermatan sulfate, hyaluronic acid, and keratan sulfate). However, a side reaction called the "ring-contraction reaction" occurs to the extent of 10% to 15% for α -linked amino sugars—e.g., those found in heparin and heparan sulfate (Shively and Conrad, 1970, 1976). The ring-contraction reaction results in deamination without bond cleavage, and may convert a segment in heparin that contains two disaccharides having N-sulfated GlcN residues into a tetrasaccharide fragment instead of the two disaccharides that would be expected. Thus, heparin and heparan sulfate yield small but significant amounts of the ring-contraction tetrasaccharides (Bienkowski and Conrad, 1985). However, there seems to be little structural selectivity for the ring-contraction reaction, as the ring-contraction tetrasaccharides contain disaccharide units in the same proportions as those found in the original glycosaminoglycan. Thus, even for heparin and heparan sulfate, the disaccharides formed during nitrous acid cleavage are obtained in the same proportions that are found in the original glycosaminoglycan. Consequently, when yields are expressed in percent of total uronic acids or percent of total amino sugars, the ring-contraction reaction does not interfere with the quantitation, and the yields of di- and monosaccharides are truly representative of their proportions in the starting glycosaminoglycan. Some of these heparin ring-contraction tetrasaccharides have been identified (Bienkowski and Conrad, 1985).

Time Considerations

The times for chromatography are apparent from the retention times reported in Tables 17.22B.1 to 17.22B.3. Time for processing of the data will depend in great part on the type of scintillation counter used and the data-processing equipment available on site where the methods will be used.

Literature Cited

- Al-Hakim, A. and Linhardt, R.J. 1991. Capillary Electrophoresis for the analysis of chondroitin sulfate and dermatan sulfate-derived disaccharides. *Anal. Biochem.* 195:68-73.
- Bienkowski, M.J. and Conrad, H.E. 1985. Structural characterization of the oligosaccharides formed by depolymerization of heparin with nitrous acid. *J. Biol. Chem.* 260:356-365.
- Delaney, S.R., Conrad, H.E., and Glaser, J.H. 1980. A new approach for isolation and sequencing of pure chondroitin SO₄ oligosaccharides. *Anal. Biochem.* 108:25-34.
- Desai, U.R., Wang, H.-M., Ampofo, S.A., Linhardt, R.J. 1993. Oligosaccharide composition of heparin and low-molecular-weight heparins by capillary electrophoresis. *Anal. Biochem.* 213:120-127.
- Edge, A.S.B., and Spiro, R.G. 1985. Structural Elucidation of glycosaminoglycans through characterization of disaccharides obtained after fragmentation by hydrazine-nitrous acid treatment. *Arch. Biochem. Biophys.* 240:560-572.
- Glaser, J.G., and Conrad, H.E. 1979. Chick embryo liver B-glucuronidase. Comparison of activity on natural and artificial substrates. *J. Biol. Chem.* 254:6588-6597.
- Guo, Y. and Conrad, H.E. 1988. Analysis of oligosaccharides from heparin by reversed phase ion-pairing high performance liquid chromatography. *Anal. Biochem.* 168:54-62.
- Guo, Y. and Conrad, H.E. 1989. The disaccharide composition of heparins and heparan sulfates. *Anal. Biochem.* 176:96-104.
- Hopwood, J.J. and Elliott, H. 1983. Selective depolymerisation of keratan sulfate: Production of radiolabeled substrates for 6-O-sulfogalactose sulfatase and β -D-galactosidase. *Carbohydr. Res.* 117:263-274.
- Hopwood, J.J. and Muller, V.J. 1983. Selective depolymerisation of dermatan sulfate: Production of radiolabelled substrates for α -L-iduronidase, sulfiduronate sulfatase, and β -D-glucuronidase. *Carbohydr. Res.* 122:227-239.
- Linhardt, R.J., Rice, K.G., Kim, Y.S., Lohse, D.L., Wang, H.M., and Loganathan, D. 1988. Mapping and quantification of the major oligosaccharide components of heparin. *Biochem. J.* 254:781-787.
- Rice, K.G., Kim, Y.S., Grant, A.C., Merchant, Z.M., and Linhardt, R.J. 1985. High-performance liquid chromatographic separation of heparin derived oligosaccharides. *Anal. Biochem.* 150:325-331.
- Shaklee, P.N. and Conrad, H.E. 1986. The disaccharides formed by deaminative cleavage of N-deacetylated glycosaminoglycans. *Biochem. J.* 235:225-236.

Shively, J.E. and Conrad, H.E. 1970. Stoichiometry of the nitrous acid deaminative cleavage of model amino sugar glycosides and glycosaminoglycuronans. *Biochemistry* 9:33-41.

Shively, J.E. and Conrad, H.E. 1976. Formation of anhydrosugars in the chemical depolymerization of heparin. *Biochemistry* 15:3932-3942.

Key References

Bienkowski and Conrad, 1985. See above.

Describes the separation of heparin di- and tetrasaccharides on SAX columns.

Guo and Conrad, 1988. See above.

Describes the separation of heparin di- and tetrasaccharides by reversed-phase ion-pairing HPLC.

Shaklee and Conrad, 1986. See above.

Describes the separation of disaccharides from chondroitin sulfate, dermatan sulfate, and keratan sulfate formed by both nitrous acid and lyase cleavage.

Contributed by H. Edward Conrad
University of Illinois
Urbana, Illinois

Analysis of Sulfate Esters by Solvolysis or Hydrolysis

UNIT 17.23

Sulfate esters are found on N- and O-linked sugar chains or glycosaminoglycan (GAG) chains. Few sulfatases are available that can enzymatically remove them, so chemical procedures must be used. These procedures rely on the differential sensitivity of sulfates located in different linkages on the sugar. In comparison to the conditions used for enzymatic digestion, those used for chemical digestion are very harsh and cannot be used on protein-bound carbohydrates (except for analytical purposes, as in Basic Protocol 2). With protein-bound carbohydrates, for preparative purposes, the chains must first be released by one of the procedures described in *UNITS 17.13 & 17.15*. Basic Protocol 1 describes release of sulfate esters by solvolysis. A Support Protocol describes a method for monitoring the efficiency of the solvolysis reaction, and Alternate Protocol 1 provides a scale-up method for using the solvolysis reaction with large amounts of material. Basic Protocol 2 presents a technique for acid hydrolysis to release sulfate esters, and Alternate Protocol 2 describes basic hydrolysis for the same purpose.

SOLVOLYSIS FOR RELEASE OF SULFATE ESTERS FROM GLYCANS

**BASIC
PROTOCOL 1**

Solvolysis is analogous to hydrolysis, except that a nonaqueous solvent is used. Both solvolysis procedures in this unit (also see Alternate Protocol 1) require conversion of the sulfated molecule to the pyridinium salt before heating in dimethyl sulfoxide (DMSO) to remove sulfate. This procedure is for $^{35}\text{SO}_4^-$ - or ^3H -radiolabeled GAG samples, and applies to microgram amounts of material. For scale-up, see Alternate Protocol 1.

Materials

Sulfated sample (*UNIT 17.15*), radiolabeled (*UNIT 17.4* or *UNIT 17.5*)
Glucose-6-sulfate (Sigma)
0.5 × 50-cm Sephadex G-25 gel-filtration column (Pharmacia Biotech)
2-cm column of Dowex-50 X8 (hydrogen form; see recipe for washing procedure)
packed in a Pasteur pipet plugged with glass wool
Pyridine
1 M NaCl
9:1 (v/v) dimethyl sulfoxide (DMSO; reagent grade)/methanol
QAE-Sephadex (Pharmacia Biotech)
Thick-walled conical glass tubes
Heating block
Lyophilizer or shaker-evaporator
Additional reagents and equipment for desalting by gel-filtration chromatography
(*UNIT 10.9*)

1. Desalt sample on a 0.5 × 50-cm Sephadex G-25 column, eluting with water.

It is essential to remove all nonvolatile salts from the sample. If free $^{35}\text{SO}_4$ is present in the sample, sulfate will run through the water-eluted Sephadex G-25 column with an apparent molecular weight much larger than its actual molecular weight, and may overlap with small $^{35}\text{SO}_4$ -labeled oligosaccharides. It may be therefore be necessary to remove free $^{35}\text{SO}_4$ by dialysis from the protein-bound $^{35}\text{SO}_4$ prior to the release of the carbohydrate chains from the protein.

2. Add 5 μg of glucose-6-sulfate to the sample, which will be in a volume of 1 to 5 ml of water. Pass the sample through a 2-cm column of water-washed Dowex-50 X8 (hydrogen form) packed in a glass wool-plugged Pasteur pipet. Wash column with

**Preparation and
Analysis of
Glycoconjugates**

17.23.1

Contributed by Hudson H. Freeze

Current Protocols in Molecular Biology (1995) 17.23.1-17.23.8

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 32

5 ml water and combine washings with the original flowthrough of the column. Check pH of the combined washings with pH paper; it should be within one unit of the pH of the combined sample and wash water that will be used in the subsequent steps.

The Dowex-50 column removes any residual cations and replaces them with protons. The glucose-6-sulfate provides further insurance that any residual cations will not interfere with the solvolysis reaction, but the amount of this reagent must be kept low to prevent unwanted side reactions.

3. Add 50 μ l pyridine to the sample and split it into two aliquots. Add 2 μ l of 1 M NaCl to one sample and nothing to the other. Transfer samples to thick-walled glass conical tubes and lyophilize or dry on a shaker-evaporator.

The sample with the NaCl serves as a negative control for desulfation and for the side effects of the DMSO-treatment step (step 4), because solvolysis does not proceed when the salt is present. In the experimental (i.e., no-salt) sample, most of the pyridine evaporates during lyophilization, except for that which forms the pyridinium salt with the sample or glucose-6-sulfate.

4. Add 0.2 ml of 9:1 (v/v) DMSO/methanol to each sample. Using a heating block, heat 2 hr at 85° to 90°C for O-sulfate esters or at 55°C for N-sulfate esters.
5. Remove DMSO under vacuum using a lyophilizer or shaker-evaporator.

Depending upon the next step—e.g., in the case of QAE-Sephadex chromatography (UNIT 17.20)—it may not be necessary to remove the solvent. If it is necessary to do this, it should be noted that DMSO is not very volatile and will take several hours to evaporate under vacuum. In some cases the sample may need to be diluted.

SUPPORT PROTOCOL 1

MONITORING SOLVOLYSIS REACTION

If it is important to monitor the efficiency of the reaction in Basic Protocol 1, this can be done by running a sample of [³H]glucitol-6-sulfate in parallel with the experimental samples.

Additional Materials (also see Basic Protocol 1)

[³H]glucitol-6-sulfate: glucose-6-sulfate (Sigma) reduced with NaB[³H]₄ (UNIT 17.5)
2-cm QAE-Sephadex column (Pharmacia Biotech)
35 mM and 200 mM NaCl
Scintillation fluid compatible with aqueous samples
Additional reagents and equipment for QAE-Sephadex analysis (UNIT 17.20)

1. Carry out steps 1 and 2 of Basic Protocol 1, running ~10,000 cpm of [³H]glucitol-6-sulfate in parallel with the experimental sample. In step 3, split the pyridine treated column effluent from this glucitol-6-sulfate sample into three aliquots instead of two. Add 2 μ l of 1 M NaCl to one aliquot. Dry all three aliquots as described in Basic Protocol step 3. Keep one of the non-salt-treated aliquots in reserve, then carry out steps 4 and 5 (solvolysis) on the remaining salt-treated and non-salt-treated aliquots.

Alternatively, this ³H-labeled material may be added directly to the sample itself as an internal control. This cannot, however, be done if it interferes with subsequent analysis of the sample products (e.g., if the sample itself is labeled with ³H).

2. Dilute the three [³H]glucitol samples to 1.5 ml with water and treat each of them as follows. Load on a 2-cm QAE-Sephadex column. Wash column three times using 1.5 ml of water each time and collect each effluent in a scintillation vial. Wash three more times using 1.5 ml of 35 mM NaCl each time, and collect the effluents in three more

separate vials. Finally, wash once with 1.5 ml of 200 mM NaCl and collect the effluent in another vial.

Ten vials will thus be obtained for each of the three [³H] glucitol samples.

3. Add 15 ml of an aqueous-compatible scintillation fluid to each vial and count on a scintillation counter.

The [³H] glucitol in untreated sample and the solvolysis control incubation should elute in the 35mM NaCl wash. The [³H] glucitol in solvolysed desulfated sample should elute with water showing that it no longer carries a charge. Very little radioactivity will be seen in the high-salt wash. The efficiency of sulfation is then calculated as the percentage of [³H] glucitol eluted with water compared to the total [³H] glucitol recovered.

SCALE-UP OF SOLVOLYSIS FOR LARGE AMOUNTS OF MATERIAL

Basic Protocol 1 is used to desulfate trace amounts of material, but it can be applied to larger amounts of material such as glycosaminoglycans (GAGs). As with Basic Protocol 1, the extent of the reaction can be monitored by the desulfation of [³H]glucitol-6-sulfate if the presence of this labeled compound will not interfere with the subsequent use of the product. The steps and the preparations of the procedure are essentially the same as described for the Basic Protocol 1, except that no control incubation (i.e., with NaCl; see Basic Protocol 1, step 3) is done on the material. This procedure describes the solvolysis of 10 mg of GAG (or other sulfated sugar chains) and can be further scaled up if necessary.

Additional Materials (also see Basic Protocol 1)

[³H]glucitol-6-sulfate: glucose-6-sulfate (Sigma) reduced with NaB[³H]₄ (UNIT 17.5)
Dialysis tubing (3000 MWCO)

Additional reagents and equipment for dialysis (APPENDIX 3C)

1. Desalt material containing 10 mg GAGs or other sulfated molecules on a Sephadex G-25 column and pool the fractions that contain GAGs.

Free GAG chains can be monitored by the carbazole reaction (see UNIT 17.12) or by amino sugar tests. Follow the normal procedures for gel filtration and pool conservatively to avoid any salt.

2. Mix material with an appropriate amount of [³H]glucitol-6-sulfate and pass the mixture over a 5 × 0.5-cm column of well-washed Dowex-50 X8 (hydrogen form). Wash with 2 to 3 column volumes water, then add 0.1 ml pyridine to the combined eluates. Lyophilize.

The amount of [³H]glucitol-6-sulfate to be added will depend upon the fraction of the total sample that can be sacrificed for the QAE-Sephadex analysis described in Support Protocol 1. The pH of the eluate may be lower than that found in Basic Protocol 1 because the chains, which will be in protonated form, are found at a higher concentration.

3. Dissolve the lyophilized material in 1.0 ml of 90% dimethyl sulfoxide/10% methanol and heat 2 hr at 90°C.

The [³H]glucitol-6-sulfate can be used to confirm desulfation of the GAG chains and to follow the course of the reaction. Take aliquots at different times and monitor its elution from QAE-Sephadex as described in Support Protocol 1. Use <1% of the sample (100 μg) so that the 2-cm QAE-Sephadex column is not overloaded (see Support Protocol 1). The GAG chains may remain bound to the column, but this does not matter because only the desulfation of the [³H]glucitol-6-sulfate is monitored.

ALTERNATE PROTOCOL 1

- Cool the reaction tube. Dialyze the sample against 3 liters of water using 3000-MWCO dialysis tubing.

Alternatively, remove the reagents on a Sephadex G-25 gel-filtration column. Either procedure will work, but the choice is determined by the future plans for the sample. The apparent size of the molecules will decrease because the sulfate contributes disproportionately to their apparent size, and some breakdown of chains may also occur.

ACID HYDROLYSIS FOR RELEASE OF SULFATE ESTERS FROM GLYCOSAMINOGLYCANS

This protocol measures the amount of bound $^{35}\text{SO}_4$ that remains soluble after the precipitation of hydrolyzed sulfate as the barium salt. In this procedure, we assume that the released oligosaccharide is $^{35}\text{SO}_4$ -labeled, but the sugars may also be labeled with ^3H . If the sugars are not labeled with tritium, a tritiated molecule—e.g., ^3H -labeled glucose-6-sulfate should be included as an internal standard for normalizing recovery of the material.

Materials

Sample: $^{35}\text{SO}_4$ -labeled oligosaccharide or glycopeptide
0.50 N HCl (freshly diluted from 12 N reagent grade acid)
0.50 N NaOH (accurately titrated)
4 M NaCl
0.01 N HCl in 1 M NaCl (diluted from 0.50 N HCl above)
2 mM Na_2SO_4 in 2 M KCl
20 mM barium acetate
Heating block with wells filled with oil

- Dispense ≥ 1000 to 2000 cpm of $^{35}\text{SO}_4$ -labeled oligosaccharide or glycopeptide in 50 μl water into a series of 8 to 10 microcentrifuge tubes. Add 50 μl of 0.50 N HCl to each tube.
- Incubate the tubes in a heating block with wells filled with oil at 98° to 100°C for a series of different time periods.

The amount of time will vary depending upon the sample. A preliminary experiment with time points at 1 and 3 hr will help determine the best time points to try. All sulfate esters will eventually be hydrolyzed and even the most resistant ones will be 75% degraded after 3 hr. A reasonable set of time points to try is 0, 5, 15, 30, 60, 90, 120, 150, and 180 min.

- At each time point, remove the tube from the oil, wipe it clean and place it on ice for 1 min. Microcentrifuge 2 sec at maximum speed to collect the liquid in the bottom of the tube and add sufficient 0.5 N NaOH to neutralize the sample. Also add 50 μl of 4 M NaCl to maintain high salt concentration. Keep completed samples on ice while remaining samples are incubating.

Determine the exact volume of 0.5 N NaOH (~50 μl) to add to the samples by first neutralizing a dummy sample containing 50 μl of 0.50 N HCl and 0.50 N water and checking with pH paper. Acceptable pH after neutralization is 3 to 9.

- After the time course has been completed for all samples and all samples have been neutralized, add sufficient 0.01 N HCl in 1 M NaCl to each sample to adjust the pH to 2 to 3. Check with pH paper using 1 to 2 μl of sample.

Determine the exact amount of 0.01 N HCl (~100 μl) to add by adjusting the pH of the dummy sample neutralized above (see step 3, annotation). It is still necessary, however, to check the pH of the actual samples.

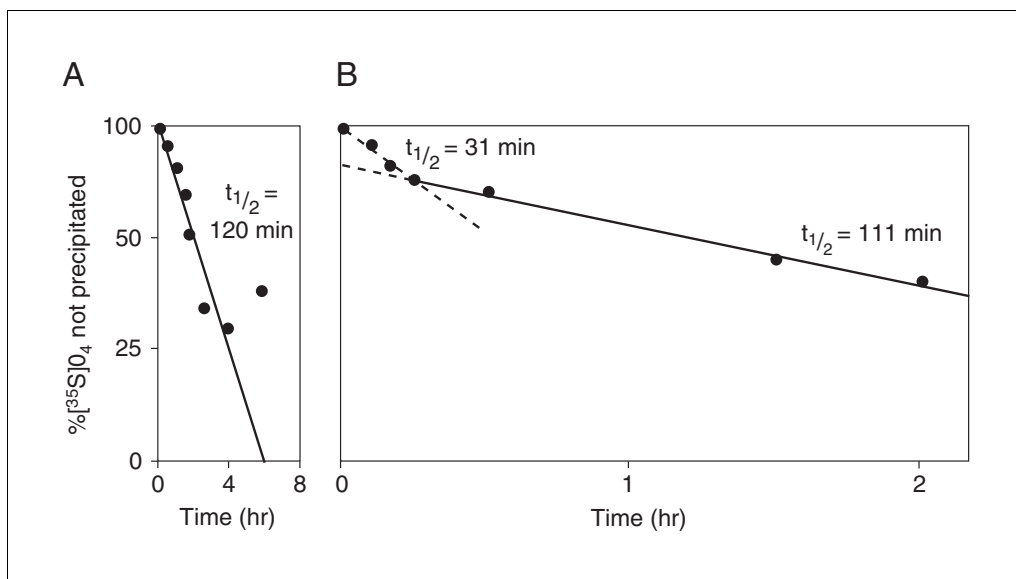


Figure 17.23.1 Examples of kinetics of loss of $^{35}\text{SO}_4$ from N-linked oligosaccharides of *Dicystostelium discoideum*. Two different types of N-linked oligosaccharides were isolated from secreted glycoproteins. **(A)** One type was sensitive to endo H digestion and consisted entirely of a single kinetic class, probably a 6-O sulfate ester with a $t_{1/2}$ of 120 min. **(B)** The other kinetic class was released only by peptide:N-glycosidase F (PNGase F) digestion and not by endo H. It consists of two classes of sulfate esters: the major class (~80%) has a $t_{1/2}$ similar to that of the endo H-sensitive oligosaccharide type (in panel A), but a minor class (~20%) shows a different $t_{1/2}$ (and may be in equatorial linkage).

5. Add 100 μl of 2 mM Na_2SO_4 in 2 M KCl (for a total volume of 400 μl) and mix. Remove 20% of the sample and determine total counts of $^{35}\text{SO}_4$.
6. Add 150 μl of 20 mM barium acetate to the remaining material in each sample and incubate 15 min on ice. Microcentrifuge 5 min at $10,000 \times g$, then remove 90% of the solution (~400 μl) and determine total counts of $^{35}\text{SO}_4$.
7. Normalize the amount of $^{35}\text{SO}_4$ radioactivity that remains soluble as a percentage of the total counts present in the original sample.

Initially, without hydrolysis essentially all of the $^{35}\text{SO}_4$ should remain soluble. Hydrolysis yields fewer soluble counts (more precipitation of $\text{Ba}^{35}\text{SO}_4$).

8. Using the "0" time point as 100% soluble material (this should be >90% of the total counts), plot the decrease in soluble $^{35}\text{SO}_4$ on a log scale for typical first-order decay kinetics. Determine if single or multiple kinetic classes exist, as indicators of the number of types of sulfate esters in the sample.

See Figure 17.23.1 for examples of single and multiple kinetic classes.

BASE HYDROLYSIS FOR RELEASE OF SULFATE ESTERS FROM GLYCOSAMINOGLYCANS

This method will release sulfate esters that are sensitive to base hydrolysis (see Background Information).

Additional Materials (also see Basic Protocol 2)

Sample: $^{35}\text{SO}_4$ -labeled, borohydride-reduced (*UNIT 17.5*) oligosaccharide or glycopeptide

2 N NaOH

2 N HCl (freshly diluted from 12 N reagent-grade acid)

1. Dispense 1000 to 2000 cpm $^{35}\text{SO}_4$ -labeled, borohydride-reduced oligosaccharide or glycopeptide in 50 μl water into a series of at least six microcentrifuge tubes. Add 50 μl of 2 N NaOH to each tube and mix.
2. Incubate the tubes at 80°C in a heating block with wells filled with oil for 0, 1, 2, 4, 6, and 10 hr.

IMPORTANT NOTE: Use prerduced oligosaccharides; otherwise the base may degrade the oligosaccharide chain. The base hydrolysis occurs at about the same rate as the loss of primary-linked sulfates ($T_{1/2} = \sim 1$ to 2 hr).

3. At each time point remove the tube, wipe it clean, and place it on ice for a few minutes. Microcentrifuge 2 sec at $10,000 \times g$ to collect the dispersed droplets. Add 50 μl of 2 N HCl to adjust the pH to 3 to 9, checking 1 to 2 μl of solution with pH paper.
4. After the time course has been completed for all samples and all samples have been neutralized, perform steps 5 to 8 of the acid hydrolysis protocol (see Basic Protocol 2) to precipitate the solubilized $^{35}\text{SO}_4$.

Completeness of any hydrolysis will be indicated by a flattening of the decay curve.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Dowex-50 X8 resin

Prewash Dowex-50 resin by stirring 50 to 100 g into 2 liters of water, allowing it to stand 1 hr, then discarding water. Repeat twice, then leave overnight at room temperature. Pour off the final water and resuspend the resin in 3 to 4 vol water. The resin may be stored at 4°C for months under these conditions.

COMMENTARY

Background Information

Solvolysis

The conditions of the solvolysis reaction can be adjusted to preferentially remove either N- or O-linked sulfate esters, because N-sulfates (e.g., GlcNSO_4 in heparin; not to be confused with sulfated N-linked oligosaccharides) are more sensitive than O-sulfate esters (Inoue and Nagasawa, 1976; Nagasawa et al., 1977).

Acid hydrolysis

It is possible to get an accurate compositional analysis of "sulfated sugars" by doing

an acid hydrolysis because the sulfate esters are cleaved faster than most glycoside bonds. By selecting the right conditions, acid hydrolysis can give information about the number and locations of sulfate esters because they are cleaved at different rates depending upon their location on the individual sugar residues (Freeze and Wolgast, 1984). The kinetics of sulfate liberation (see Fig. 17.23.1) can help to determine the number of classes of sulfate esters present and what their likely locations may be. These properties are not strongly influenced by other substituents on the sulfated sugar or by the position of the sugar within an oligosac-

charide chain. Since the hydrolysis can also cleave glycosidic linkages and partially destroy the sugar chain, only the release of sulfate can be measured.

The acid hydrolysis procedure differentiates between sulfates linked to different hydroxyl groups on the sugar ring. Under the hydrolysis conditions in Basic Protocol 2, the $T_{1/2}$ of sulfates linked to equatorial hydroxyl groups (e.g., glucose-3-sulfate, galactose-3-sulfate, *N*-acetylglucosamine-3-sulfate, and mannose-3-sulfate) is 6 to 25 min; that of those linked to axial hydroxyl groups (e.g., galactose-4-sulfate, *N*-acetylglucosamine-4-sulfate, and mannose-2-sulfate) is 60 to 84 min, and that of those linked to primary hydroxyl groups (e.g., glucose-6-sulfate, galactose-6-sulfate, *N*-acetylglucosamine-6-sulfate, and mannose-6-sulfate) is 90 to 120 min (Rees, 1963). All 6-OH positions are primary, and all of the hydroxyl groups in glucose are equatorial. Any hexose isomer of glucose (e.g., mannose or galactose) will change one of the equatorial positions to an axial position. Even though the time ranges for hydrolysis listed above have not been extensively documented for sulfate-labeled oligosaccharides, they are useful because they will clearly differentiate between multiple classes of sulfate es-

ters. Using the acid hydrolysis technique, it is not possible to determine which sugar is sulfated or where it is found in the chain. If other data are available on the structure of the oligosaccharide, this information may provide likely possibilities for the location of sulfate groups, especially if certain residues are resistant to exoglycosidase digestion.

Base hydrolysis

Only certain types of sulfate esters are sensitive to base hydrolysis (Freeze and Wolgast, 1984; Percival, 1978). Sensitivity is determined by the location of the sulfate on the sugar residue and—in contrast to acid hydrolysis described above—by the other substituents linked to that sugar residue.

Base treatment will release sulfate esters from intact oligosaccharides under two different conditions. The first is when sulfate is located in the 6 (or 3)-OH position and the 3 (or 6)-OH group of the same sugar residue is not substituted. The second is when the sulfate ester is adjacent to an unsubstituted *trans* OH group. In the first case, the sulfate is eliminated and an acid-labile 3,6-anhydrosugar is formed (see Fig. 17.23.2). In the second case, the sulfate is released with the formation of an epoxide

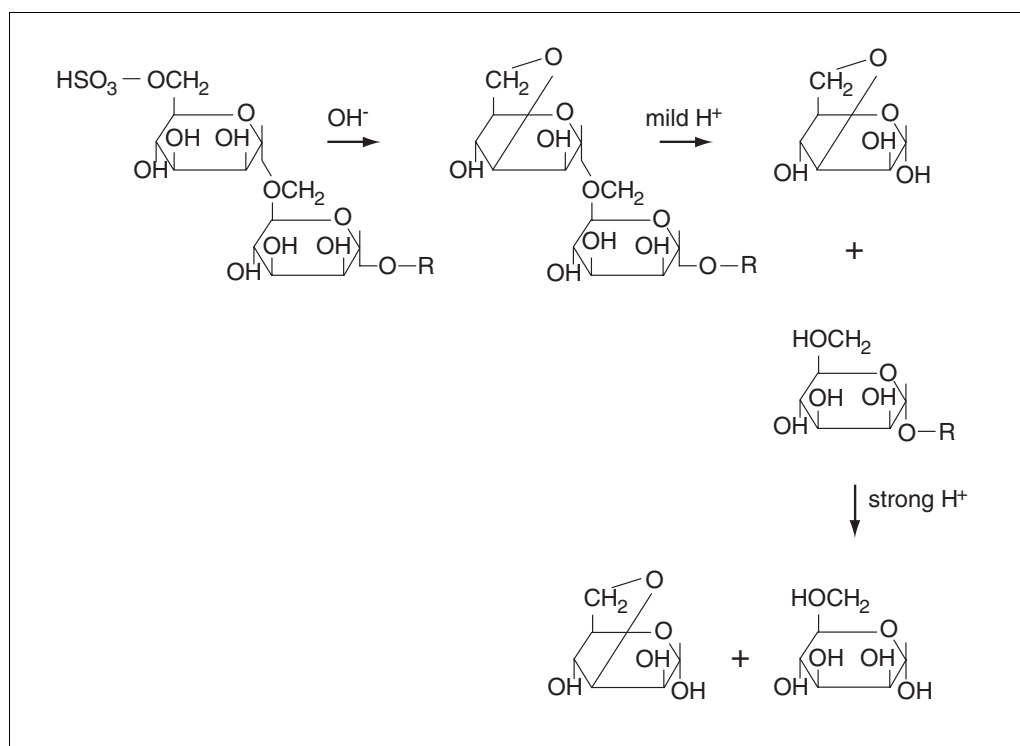


Figure 17.23.2 Formation of 3,6-anhydromannose by base treatment of an oligosaccharide containing mannose-6-sulfate. The formation of the anhydro sugar makes the glycosidic linkage very sensitive to acid; mild hydrolysis causes cleavage. Stronger hydrolysis degrades the usual glycosidic linkages, but leaves the anhydrosugar. When reduced with sodium borohydride, the modified sugar can be quantified to show how much mannose-6-sulfate was originally present.

ring on the sugar. This is destroyed during subsequent acid hydrolysis of the oligosaccharide. The modified sugar can be identified if it is metabolically labeled—e.g., as in the discovery of mannose-6-sulfate (Freeze and Wolgast, 1984)—but this is not routinely done. On the other hand, the loss of sulfate is easily measured. When data on the acid and base hydrolysis of sulfate esters are combined with other structural information on the molecules, several potential structures can be eliminated or supported, but these procedures will not provide proof of the structure.

Critical Parameters

Solvolytic desulfation is a harsh treatment that also removes sialic acids and probably fucose residues. If these sugars are normally present, their loss may be an unacceptable side effect of this procedure. On the other hand, hexoses, hexuronic acids, and hexosamines in glycosidic linkages are mostly stable to solvolysis. The incubated control is very important because it will monitor for any nonspecific destruction of the sugar chain.

For acid hydrolysis, the most critical points to observe are that the pH should be kept between 2 and 3 and that the salt concentration be kept very high during the precipitation of sulfate with barium. Only then is the precipitation specific for sulfate. Under less stringent conditions, the oligosaccharides or even monosaccharide sulfates can be precipitated.

Anticipated Results

Expect ~80% to 90% efficiency in desulfation, and <20% in the control incubation (Freeze and Wolgast, 1984).

Time Considerations

Solvolytic desulfation itself takes only a few hours to complete. The rate-limiting step is the time needed to remove the water after the sample has passed through the Dowex-50 column and the pyridine has been added. This will only take a few minutes using a shaker-evaporator, but will require an overnight lyophilization.

Beginning with the hydrolysis step, the procedure can be completed within a day. If necessary, the procedure can be stopped after taking all of the time points and neutralizing them with base prior to precipitation. Where this is done, the reaction mixtures can be stored at 0°-4°C for 24 hr or more.

Literature Cited

- Inoue, Y. and Nagasawa, K. 1976. Selective N-desulfation of heparin with dimethyl sulfoxide containing water or methanol. *Carbohydr. Res.* 46:87-95.
- Nagasawa, K., Inoue, Y., and Kamata, T. 1977. Solvolytic desulfation of glycosaminoglycuronan sulfates with dimethyl sulfoxide containing water or methanol. *Carbohydr. Res.* 58:47-55.
- Freeze, H. and Wolgast, D. 1984. Structural analysis of N-linked oligosaccharides from glycoproteins secreted by *Dictyostelium discoideum*. *J. Biol. Chem.* 261:127-133.
- Rees, D.A. 1963. A note on the characterization of carbohydrate sulphates by acid hydrolysis. *Biochem J.* 88:343-346.
- Percival, E. 1978. Sulfated polysaccharides of the Rhodophyceae. In *Carbohydrate Sulfates*. (R.G. Schweiger, ed.) pp. 213-224. American Chemical Society, Washington, D.C.

Contributed by Hudson H. Freeze
La Jolla Cancer Research Foundation
La Jolla, California

CHAPTER 18

Analysis of Protein Phosphorylation

INTRODUCTION

Most proteins undergo covalent modification of their amino acid side chains during or after synthesis. This modification, particularly if it is reversible, can provide an extraordinarily sensitive means by which the activity of a protein can be enhanced or diminished. Protein phosphorylation is probably the most common and important form of protein modification.

Thirty years ago, only a few protein kinases had been identified and only a limited number of processes were suspected of being regulated by protein phosphorylation. The increasing power of molecular biology has facilitated the identification and characterization of the genes of literally thousands of protein kinases and protein phosphatases in diverse species, while advances in cell and biochemical techniques have allowed the detailed elucidation of complex cellular regulatory pathways. Today, many aspects of the regulation of cell function are known to be controlled by protein phosphorylation. Proliferation, differentiation, signal transduction, and metabolism are all regulated by the balance of activity of protein kinases and protein phosphatases upon critical target proteins. Both extracellular ligands, such as polypeptide growth factors, and key intracellular signaling elements such as the concentration of cyclic AMP or calcium closely regulate the activities of kinases and phosphatases, which are themselves subject to modification by phosphorylation.

This chapter presents an overview of protein phosphorylation and provides detailed methods designed to detect and identify phosphorylated proteins and analyze their specific sites of modification using both isotopic and nonisotopic approaches. The overview (*UNIT 18.1*) provides specific examples of regulatory processes controlled by phosphorylation and provides a description of and rationale for the various approaches to studying protein phosphorylation. Succeeding units describe the labeling of eukaryotic cells with inorganic phosphate and preparation of lysates for immunoprecipitation (*UNIT 18.2*), phosphoamino acid analysis (*UNIT 18.3*), and detection of protein phosphorylation using immunologic or enzymatic techniques (*UNIT 18.4*).

To examine the functional effects of phosphorylation patterns, phosphoproteins can be dephosphorylated *in vitro*. *UNIT 18.5* presents specific strategies and protocols for the use of enzymes that allow either general or selective reversal of phosphorylation, as well as methods for subsequent detection by analysis of loss of ^{32}P -labeled phosphate or change in electrophoretic migration. The next unit provides yet another, somewhat more sophisticated, approach to the detection of phosphorylation. The state of phosphorylation of a given sequence can be detected quite specifically by antibodies that recognize a protein only in its phosphorylated state, and that do not cross-react with either the unphosphorylated form or other phosphoproteins, thus providing a good measure both of the amount of a specific protein and of its functional state. Strategies and methods for producing and purifying these anti-phosphopeptide antibodies are presented in *UNIT 18.6*.

A survey of the most common protein kinase assays (*UNIT 18.7*) provides readers with the conditions needed to detect and measure their activities. Tyrosine kinase,

cyclic-nucleotide dependent kinases, casein kinases, protein kinase C, and the Ca²⁺/calmodulin-dependent kinases are the major enzymes described here.

Brief chemical permeabilization of cells allows the introduction of labeled cell-impermeable reagents into the extracellular component to study signal-transduction processes. This process, described in *UNIT 18.8*, complements the labeling procedure described in *UNIT 18.2*. Standard approaches for more detailed studies of the labeled products are described in *UNIT 18.9*, which includes methods to generate and analyze phosphopeptide products from a particular protein of interest, as well as the use of manual Edman degradation to define precise positions of phosphorylated residues within labeled peptides.

UNIT 18.10 focuses on protein phosphatase inhibitors as tools for the study of protein phosphorylation. The ability of these reagents to amplify the transient effects of protein kinases can facilitate the elucidation of signaling pathways *in vivo*. Phosphatase inhibitors are also used to promote the isolation of desired phosphoproteins from cell homogenates and in other *in vitro* applications. This unit provides background on the different protein phosphatase enzymes as well as both *in vivo* and *in vitro* methods for specific inhibition of protein phosphatases 1 and 2A, protein phosphatase 2B/calcineurin, and protein tyrosine phosphatases.

UNIT 18.11 describes a powerful technology for studying the function of individual kinases. The ATP binding site of the kinase of interest is mutated in such a way that bulkier ATP analogs can be utilized by the enzyme. Such “analog-sensitive” mutant kinases are uniquely able to phosphorylate their substrates when provided with an appropriate radiolabeled ATP analog that cannot bind to other kinases. The analog-sensitive alleles can also bind highly specific inhibitors that do not affect other kinases. These inhibitors can be used for functional studies in genetically tractable organisms in which the wild-type allele can be replaced with the analog-sensitive allele. This unit describes not only the strategy for generation of desired analog-sensitive alleles, but also protocols for the synthesis of inhibitors and radiolabeled ATP analog.

UNIT 18.12 presents several approaches for the analysis of kinase activation, with a particular emphasis on the important mitogen-activated protein kinases (MAPKs). Protocols are described for measurement of kinase activity after either specific immunoprecipitation of activated MAPKs with phospho-specific antibodies or affinity purification with a specific kinase substrate, such as c-jun. These approaches depend on substantial characterization of the kinase, but this is not necessary for the *in-gel* assay, in which a particular substrate of interest is incorporated into an SDS-polyacrylamide gel. After the samples of interest are resolved, proteins are renatured and kinase activity is detected by soaking the gel in [γ -³²P]ATP. This provides information on the activity of one or more kinases of interest and also reveals their molecular weights.

David D. Moore and
Bartholomew M. Sefton

Overview of Protein Phosphorylation

UNIT 18.1

HISTORY

Phosphorylation is the most common and important mechanism of acute and reversible regulation of protein function. Studies of mammalian cells metabolically labeled with [^{32}P]orthophosphate suggest that as many as one-third of all cellular proteins are covalently modified by protein phosphorylation. Protein phosphorylation and dephosphorylation function together in signal transduction pathways to induce rapid changes in response to hormones, growth factors, and neurotransmitters. Most polypeptide growth factors (platelet-derived growth factor and epidermal growth factor are among the best studied; Heldin, 1995) and cytokines (e.g., interleukin 2, colony stimulating factor 1, and γ -interferon; Ihle et al., 1994) stimulate phosphorylation upon binding to their receptors. Induced phosphorylation in turn activates cytoplasmic protein kinases, such as raf, MEK, and MAP kinases (Marshall, 1995). Additionally, in all nucleated organisms, cell cycle progression is regulated at both the G1/S and the G2/M transitions by cyclin-dependent protein kinases (Doree and Galas, 1994).

Differentiation and development are also controlled by phosphorylation. Development of the R7 cell in the *Drosophila* retina (Simon, 1994) and of the vulva in *Caenorhabditis elegans* (Eisenmann and Kim, 1994) are both dependent on the function of receptor and cytoplasmic protein kinases. Finally, metabolism—in particular, the interconversion of glucose and glycogen and the transport of glucose—is regulated by phosphorylation (Cohen, 1985). Biologists of all stripes therefore find, often unexpectedly and occasionally reluctantly, that they must study protein phosphorylation in order to understand the regulation and function of their favorite gene and its product.

In the early 1940s, soon after their discovery that glycogen phosphorylase existed in two distinct forms in skeletal muscle, Carl and Gerty Cori identified a cellular activity that converted active phosphorylase *a* into a less active phosphorylase *b*. Initial studies of what they called the prosthetic-group removing (PR) enzyme suggested that the inactivation of phosphorylase *a* was mediated via proteolysis (Cori and Cori, 1945). More than ten years later, this converting enzyme was shown to be a protein phosphatase. Thus, a protein phosphatase was studied nearly 20 years prior to the identifica-

tion of the first protein kinase. The finding that phosphorylase phosphatases were more widely expressed in tissues than phosphorylase *a* led to the current realization that these enzymes have a broad substrate specificity and regulate many physiological processes. The use of substrates other than phosphorylase *a* also identified several other protein serine/threonine phosphatases.

LABELING STUDIES

Protein phosphorylation is usually studied by biosynthetic labeling with ^{32}P -labeled inorganic phosphate ($^{32}\text{P}_i$; UNIT 18.2). This is intrinsically quite simple—the label is just added to growth medium. It is this step of an experiment, however, that makes many investigators the most nervous, given the perceived danger of radioactive exposure and the real danger of contamination of laboratory equipment with radioactivity. Neither problem is insurmountable. With proper shielding and technique, exposure of the investigator can be limited to the hands and contamination of the laboratory can be avoided. A general protocol for biosynthetic labeling with $^{32}\text{P}_i$ that maximizes incorporation and minimizes radioactive exposure of workers in the lab and contamination of lab equipment is described in UNIT 18.2. (For a general discussion of radiation safety consult Safe Use of Radioisotopes, APPENDIX 1F.)

Many protein kinases can be assayed by the transfer of radiolabel from [^{32}P]ATP to surrogate protein and peptide substrates. Synthetic peptide substrates have been particularly useful in establishing consensus motifs that define the substrate for protein kinases. By comparison, protein phosphatases show very low activity against synthetic phosphopeptides, greatly preferring the phosphoprotein substrates (Shenolikar and Ingebritsen, 1984). Therefore, attempts to define structural requirements for protein dephosphorylation using synthetic phosphopeptides have failed to establish the molecular basis for substrate specificity of the major cellular protein phosphatases.

The rate of ^{32}P incorporation into proteins *in vivo*, at either basal or hormone-stimulated levels, depends heavily on the rate of turnover of phosphate in the protein. For example, low phosphatase activity against specific sites or proteins will reduce the turnover of protein-bound phosphates and hinder their metabolic labeling in the intact cell. Conversely, physi-

Contributed by Bartholomew M. Sefton and Shirish Shenolikar

Current Protocols in Molecular Biology (1996) 18.1.1-18.1.5

Copyright © 2000 by John Wiley & Sons, Inc.

Analysis of
Protein
Phosphorylation

18.1.1

Supplement 33

ological stimuli that increase phosphatase activity may facilitate metabolic labeling of proteins, allowing rapid replacement of endogenous “cold” phosphate with the radiolabel. Thus, ^{32}P incorporation by itself gives no indication about the extent of phosphorylation of a particular protein. As most phosphoproteins are phosphorylated on multiple sites, this also makes it difficult to correlate changes in protein function with ^{32}P labeling of a given protein in the intact cell.

SITES OF PHOSPHORYLATION

Most proteins are found to be phosphorylated at serine or threonine residues, and many proteins involved in signal transduction are also phosphorylated at tyrosine residues. These three hydroxyphosphoamino acids exhibit sufficient chemical stability at acidic pH that they can be recovered after acid hydrolysis and identified in a straightforward manner. Proteins that contain covalently bound phosphate at histidine, cysteine, and aspartic acid residues, either as phosphoenzyme intermediates or as stable modifications, have also been described. Each of these phosphoamino acids is chemically labile and impossible to study with the standard techniques used for the acid-stable phosphoamino acids. Indeed, they are often identified by inference or elimination. A technique for identifying phosphoserine, phosphothreonine, and phosphotyrosine by acid hydrolysis and two-dimensional thin-layer electrophoresis is described in *UNIT 18.3*. Techniques for analyzing acid-labile forms of protein phosphorylation are described in Ringer, 1991; Kamps, 1991; and Duclos et al., 1991.

Phosphotyrosine is not an abundant phosphoamino acid. Therefore, its detection in samples labeled with $^{32}\text{P}_i$ is often difficult, especially if the samples contain large quantities of proteins phosphorylated at serine residues or are contaminated with RNA. Detection of phosphotyrosine, as well as of phosphothreonine, can be enhanced considerably by incubation of gel-fractionated samples in alkali. This hydrolyzes RNA and dephosphorylates phosphoserine, allowing visualization of minor tyrosine- and threonine-phosphorylated proteins. A simple procedure for alkaline treatment is described in the alternate protocol of *UNIT 18.3*.

DETECTION OF UNLABELED PHOSPHOAMINO ACIDS

If a protein is modified by phosphorylation, identification of the phosphoamino acid can

often be accomplished without resorting to biosynthetic labeling. For example, tyrosine phosphorylation can be studied because proteins containing this rare phosphoamino acid can be detected with great specificity and sensitivity by antibodies to phosphotyrosine (*UNIT 18.4*). By comparison, the quantitation of phosphoserine and phosphothreonine is more difficult and requires partial acid hydrolysis of the phosphoprotein and subsequent separation of phosphoamino acids by high-voltage electrophoresis on thin-layer cellulose acetate plates. Attempts to generate antibodies that recognize phosphoserine or phosphothreonine have failed to produce reagents with the required specificity and/or sufficient sensitivity to be useful. However, once the primary sequence around a phosphorylation site containing phosphoserine or phosphothreonine has been determined, it is possible to make antibodies against synthetic phosphopeptides modeled on these phosphorylation sites (Czernik et al., 1991). Such anti-phosphopeptide antibodies have been very useful tools for monitoring the phosphorylation of the parent protein at specific sites. The phospho-specific antibodies have also been used to inhibit the dephosphorylation of individual sites and demonstrate their role in protein function. Several recent studies have produced phospho-specific antibodies against phosphorylation sites predicted solely by the primary sequence of a protein. These reagents have provided new insights into phosphorylations that control protein function.

More generally, because phosphorylation often alters the mobility of a protein during SDS–polyacrylamide gel electrophoresis and almost always alters its isoelectric point, the presence of phosphorylated residues in an unlabeled protein can be deduced from altered gel mobility after incubation of the protein with a phosphatase. Most phosphorylation sites are thought to reside near the surface of phosphoproteins. Therefore, they are equally accessible to phosphatases and proteases. Indeed, limited proteolysis often produces functional changes in phosphoproteins that closely mimic their *in vitro* dephosphorylation by phosphatase. As monitoring the release of radioactivity from a ^{32}P -labeled phosphoprotein does not distinguish between a phosphatase and a protease reaction, additional steps must be taken to differentiate orthophosphate from small phosphopeptides. This is particularly important when dephosphorylation reactions are carried out in crude extracts (*UNIT 18.5*).

PROTEIN KINASES

Protein kinases exhibit a strict specificity for phosphorylation of either serine/threonine or tyrosine residues. Yet these enzymes share structural homology in several domains such as the ATP-binding and catalytic sites, emphasizing their common function as phosphotransferases. A third group of protein kinases more closely resembles the serine/threonine kinases in their primary sequence but phosphorylate both serine/threonine and tyrosine residues. MEK, a MAP kinase activator (Crews and Erickson, 1992) and *wee1*, an inhibitor of *cdc2* kinase (Parker et al., 1992) are examples of such dual-specificity protein kinases that phosphorylate threonine and tyrosine residues which are closely located in substrate proteins.

Kinases have been used in a number of ways to analyze protein phosphorylation. Mutations that abrogate ATP-binding inactivate protein kinases. Such inactive kinases retain the ability to recognize substrates and therefore act as dominant-negative reagents to analyze the physiological function of kinases in the intact cell (Thorburn et al., 1994; Alberola-Ila et al., 1995). Cell-permeable compounds modeled on ATP also inhibit selected kinases in the intact cell.

Growth factors promote dimerization and activation of transmembrane receptor kinases. Mutant receptors that can dimerize in the absence of the natural ligand offer a novel approach to activate receptor kinases and initiate the cellular responses to growth factors (Spencer et al., 1993). Several cytoplasmic kinases are also activated by binding of intracellular second messengers. The allosteric activation of these kinases results from conformation changes that eliminate internal autoinhibitory interactions which silence the kinases in the absence of ligand. Deletion of the autoinhibitory sequences generates a constitutively active kinase that provides important insights into its physiological role (Planas-Silva and Means, 1992). Synthetic peptides modeled on the autoinhibitory sequences can also be used to selectively inhibit these kinases and reveal their role in controlling protein phosphorylation.

Still other kinases, which are not directly regulated by ligand binding, are activated by phosphorylation. Some of these kinases participate in cascades of sequential phosphorylations that amplify physiological signals. Substitution of the amino acids phosphorylated during the activation of these kinases with nonphosphorylated residues generates dominant-negative enzymes that can block the functions of such

kinases in intact cells. Alternately, acidic amino acids may be substituted for phosphoamino acids to produce a constitutively active kinase (Brunet et al., 1994). Such reagents have been used to validate the role of protein kinases in specific phosphorylation events.

PROTEIN PHOSPHATASES

Many protein phosphatases also show a strict specificity for either phosphoserine/phosphothreonine or phosphotyrosine residues. However, unlike kinases, the serine/threonine and tyrosine phosphatases are not evolutionarily related and exhibit no primary sequence homology (Shenolikar and Nairn, 1991; Charbonneau and Tonks, 1992). Acid and alkali phosphatases can also dephosphorylate phosphoproteins *in vitro* but share no structural homology with protein phosphatases. A new family of phosphatases, such as *cdc25* and *CL100*, are distantly related to the tyrosine phosphatases but dephosphorylate both phosphotyrosine and phosphothreonine in their target substrates. Mutation of an essential cysteine in tyrosine phosphatases has produced dominant-negative enzymes that have been shown to shield their substrates from dephosphorylation by endogenous phosphatases, thereby prolonging the biological effects of protein phosphorylation (Sun et al., 1993). Mutating several conserved amino acids can inactivate a serine/threonine phosphatase (Zhou et al., 1994), but the ability of such mutant phosphatases to modulate serine/threonine phosphorylation in cells has not yet been demonstrated.

A number of potent phosphatase inhibitors have been identified in recent years. Okadaic acid and several other toxins inhibit protein (serine/threonine) phosphatase 1 and 2A, and vanadate and phenylarsenoxide inhibit tyrosine phosphatase. These compounds have implicated reversible phosphorylation as a regulatory mechanism in many physiological processes (Cohen, 1989; Hardie et al., 1991; Shenolikar and Nairn, 1991; Shenolikar, 1994); under certain conditions, they may be the dominant regulators of these cellular processes. These phosphatase inhibitors can also be used to distinguish protein dephosphorylation from proteolysis in crude tissue extracts (Cohen, 1991). However, by far the most important contribution of these reagents has been that they have allowed assessment of the role played by phosphorylation in cellular processes where neither the identity of the phosphoprotein involved nor that of the kinase(s) that regulates its function are known.

The units in this chapter describe techniques that detect protein phosphorylation and identify amino acids that have been covalently modified (UNITS 18.2, 18.3, & 18.4). The next step is to use phosphatases to dephosphorylate the substrate protein in vitro and address the functional relevance of the covalent modification (UNIT 18.5). Once the regulatory role of phosphorylation has been clearly established, the more sophisticated approaches discussed above can be used to identify the specific kinases and phosphatases involved and thereby begin to elucidate the physiological mechanism that regulates the substrate in the intact cell.

LITERATURE CITED

- Alberola-Ila, J., Forbush, K.A., Seger, R., Krebs, E.G., and Perlmutter, R.M. 1995. Selective requirement for MAP kinase activation in thymocyte differentiation. *Nature* 373:620-623.
- Brunet, A., Pages, G., and Poussegur, J. 1994. Constitutively active mutants of MAP (MEK1) induce growth factor-relaxation and oncogenicity when expressed in fibroblasts. *Oncogene* 9:3379-3387.
- Charbonneau, H. and Tonks, N.K. 1992. 1002 protein phosphatases? *Annu. Rev. Cell Biol.* 8:463-493.
- Cohen, P. 1985. The role of protein phosphorylation in the hormonal control of enzyme activity. *Eur. J. Biochem.* 15:439-448.
- Cohen, P. 1989. The structure and regulation of protein phosphatases. *Annu. Rev. Biochem.* 58:453-508.
- Cohen, P. 1991. Classification of protein serine/threonine phosphatases: Identification and quantitation in cell extracts. *Methods Enzymol.* 201:389-398.
- Cori, G.T. and Cori, C.F. 1945. Enzymatic conversion of phosphorylase *a* to *b*. *J. Biol. Chem.* 158:321-345.
- Crews, C.M. and Erickson, R.L. 1992. Purification of a murine protein-tyrosine/threonine kinase that phosphorylates and activates the erk-1 gene product: Relationship to the fission yeast byr1 gene product. *Proc. Natl. Acad. Sci. U.S.A.* 89:8205-8209.
- Czernik, A.J., Girault, J.A., Nairn, A.C., Chen, J., Snyder, G., Kebabian, J., and Greengard, P. 1991. Production of phosphorylation state-specific antibodies. *Methods Enzymol.* 201:264-283.
- Doree, M. and Galas, S. 1994. The cyclin-dependent protein kinases and the control of cell division. *FASEB J.* 8:1114-1121.
- Duclos, B., Marcandier, S., and Cozzone, A.J. 1991. Chemical properties and separation of phosphoamino acids by thin-layer chromatography and/or electrophoresis. *Methods Enzymol.* 201:10-21.
- Eisenmann, D.M. and Kim, S.K. 1994. Signal transduction and cell fate specification during *Caenorhabditis elegans* vulval development. *Curr. Opin. Genet. Dev.* 4:508-516.
- Hardie, D.G., Haystead, T.A.J., and Sim, A.T.R. 1991. Use of okadaic acid to inhibit protein phosphatases in intact cells. *Methods Enzymol.* 201:469-477.
- Heldin, C.-H. 1995. Dimerization of cell surface receptors in signal transduction. *Cell* 80:213-223.
- Ihle, J.N., Witthuhn, B.A., Quelle, F.W., Yamamoto, K., Thierfelder, W.E., Kreider, B., and Silvennoinen, O. 1994. Signalling by the cytokine receptor superfamily: JAKS and STATS. *Trends Biochem. Sci.* 19:222-227.
- Kamps, M.P. 1991. Determination of phosphoamino acid composition by acid hydrolysis of protein blotted to Immobilon. *Methods Enzymol.* 201:21-27.
- Marshall, C.J. 1995. Specificity of receptor tyrosine kinase signalling: Transient versus sustained extracellular signal-regulated kinase activation. *Cell* 80:179-185.
- Parker, L.L., Atherton-Fessler, S., and Piwinica-Worms, H. 1992. p107 wee1 is a dual-specificity kinase that phosphorylates p34cdc2 on tyrosine 15. *Proc. Natl. Acad. Sci. U.S.A.* 89:2917-2921.
- Planas-Silva, M.D. and Means, A.R. 1992. Expression of a constitutive form of calcium/calmodulin-dependent protein kinase II leads to arrest of the cell cycle in G2. *EMBO J.* 11:507-517.
- Ringer, D.P. 1991. Separation of phosphotyrosine, phosphoserine, and phosphothreonine by high-performance liquid chromatography. *Methods Enzymol.* 201:3-10.
- Shenolikar, S. 1994. Protein serine/threonine phosphatases: New avenues for cell regulation. *Annu. Rev. Cell Biol.* 10:55-86.
- Shenolikar, S. and Ingebritsen, T.S. 1984. Protein (serine, threonine) phosphate phosphatases. *Methods Enzymol.* 107:102-130.
- Shenolikar, S. and Nairn, A.C. 1991. Protein phosphatases: Recent progress. *Adv. Second Messenger Phosphoprotein Res.* 23:1-121.
- Simon, M. 1994. Signal transduction during the development of the *Drosophila* R7 photoreceptor. *Dev. Biol.* 166:431-442.
- Spencer, D.M., Wandless, T.J., Schreiber, S.L., and Crabtree, G.R. 1993. Controlling signal transduction with synthetic ligands. *Science* 262:1019-1024.
- Sun, H., Charles, C.H., Lau, L.F., and Tonks, N.K. 1993. MKP-1 (3CH134), an immediate early gene product, is a dual-specificity phosphatase that dephosphorylates MAP kinase in vivo. *Cell* 75:487-493.

- Thorburn, J., McMahon, M. and Thorburn, A. 1994. Raf-1 kinase activity is necessary and sufficient for gene expression changes but not sufficient for cellular morphology changes associated with cardiac myocyte hypertrophy. *J. Biol. Chem.* 269: 30580-30586.
- Zhou, S., Clemens, J.C., Stone, R.L., and Dixon, J.E. 1994. Mutational analysis of a ser/thr phosphatase. Identification of residues important in phosphatase substrate binding and catalysis. *J. Biol. Chem.* 269:26234-26238.

Contributed by Bartholomew M. Sefton
The Salk Institute
San Diego, California

Shirish Shenolikar
Duke University Medical Center
Durham, North Carolina

Labeling Cultured Cells with $^{32}\text{P}_i$ and Preparing Cell Lysates for Immunoprecipitation

UNIT 18.2

This unit describes $^{32}\text{P}_i$ labeling and lysis of cultured cells to be used for subsequent immunoprecipitation of proteins. The approach is appropriate, however, for labeling any cellular constituent with ^{32}P . This procedure is suitable for insect, avian, and mammalian cells and can be used with both adherent and nonadherent cultures. The same general approach—biosynthetic labeling with $^{32}\text{P}_i$ in medium containing a reduced concentration of phosphate—can also be applied to bacteria and yeast; however, specific techniques to accomplish this are not presented here. This approach can also be used to label any cellular constituents with ^{32}P ; in these cases the basic protocol must be modified beginning with step 8.

The first procedure described (see Basic Protocol) is $^{32}\text{P}_i$ labeling of adherent or nonadherent (e.g., hematopoietic) cells with subsequent lysis in a detergent buffer containing either Nonidet P-40 (NP-40), 3[(3-cholamidopropyl)-dimethylammonio]-1-propane-sulfonate (CHAPS), or a combination of NP-40, sodium deoxycholate and SDS (RadioImmunoPrecipitation Assay or RIPA buffer). More rigorous lysis conditions to be used for working with proteins that are difficult to solubilize are also described (see Alternate Protocol).

CAUTION: Unshielded ^{32}P will penetrate ~1 cm into flesh. Exposure to the skin and eyes is, therefore, of concern. Gloves and protective eyewear should always be worn when handling significant amounts of ^{32}P . A 1-in.-thick (2.5-cm) Plexiglas shield, tall enough to look through when seated or standing comfortably, should be used when handling samples containing ^{32}P . For more specific precautions, refer to the section on Safe Use of Radioisotopes in *APPENDIX 1*.

LABELING CULTURED CELLS WITH $^{32}\text{P}_i$ AND LYSIS USING MILD DETERGENT

**BASIC
PROTOCOL**

The first six steps of this protocol describe culturing and labeling procedures for cultures of adherent cells; modifications appropriate for nonadherent cells are described in alternate steps. The remaining steps describe lysis of cells in detergent buffer to prepare the sample for immunoprecipitation. For a more detailed discussion of mammalian cell culture conditions and reagents see the introduction to Chapter 9.

Materials

- Cell culture to be labeled
- Labeling medium: phosphate-free tissue culture medium (e.g., DMEM, *APPENDIX 3F*) supplemented with the usual concentration of serum or serum dialyzed against phosphate-free saline, 37°C
- 1 Ci/ml $\text{H}_3^{32}\text{PO}_4$ in HCl (carrier free ICN)
- Tris-buffered or phosphate-buffered saline (TBS or PBS; see recipes), cold
- Mild lysis buffer or RIPA lysis buffer (see recipes)
- 1-in.-thick Plexiglas shield (*APPENDIX 1*)
- Plugged, aerosol-resistant pipet tips
- Plexiglas box (*APPENDIX 1*), warmed to 37°C
- Screw-cap microcentrifuge tubes
- Plugged disposable pipet or disposable one-piece transfer pipet
- Rubber policeman

**Analysis of
Protein
Phosphorylation**

Sorvall refrigerated centrifuge with SM 24 rotor and rubber adaptors, refrigerated microcentrifuge, or equivalent, 4°C

Plexiglas sheet (10 × 10 × ¼-in.) or Plexiglas tube holder, 4°C (APPENDIX 1)

Additional reagents and equipment for cell culture (see Chapter 9, introduction) and gel electrophoresis (UNIT 10.2), immunoprecipitation (UNIT 10.16), or protein purification (UNITS 10.9-10.11)

NOTE: All culture incubations are performed in a humidified 37°C, 10% CO₂ incubator unless otherwise specified.

Label the cell culture

1. Culture the cells to be labeled to an appropriate stage of growth.

Phosphate transport is maximal in rapidly growing cells. Therefore, except in those cases where phosphorylation of a protein in quiescent cells is to be examined, cells to be labeled should be subconfluent (adherent cells) or at less than maximal density (nonadherent cells). It is useful to change the medium to fresh growth medium 3 to 18 hr prior to labeling.

Brief cultivation of cells in low-phosphate medium (to reduce the phosphate pool by starvation prior to labeling) is of minimal value.

For adherent cells:

- 2a. Remove growth medium by aspiration. Wash away any residual phosphate-containing medium by adding 37°C labeling medium supplemented with serum, but lacking the label, and removing the wash medium by aspiration.

Phosphate-free DMEM is routinely used for labeling medium because RPMI has a very high concentration of phosphates.

- 3a. Add prewarmed labeling medium to cultures, using 0.5 to 1 ml per 35-mm dish, 1 to 2 ml per 50-mm dish, or 2 to 4 ml per 100-mm dish of adherent cells.

For nonadherent cells:

- 2b. Gently centrifuge the culture 1 min at 1800 × g and aspirate the medium away from the cell pellet. Resuspend the cells in labeling medium supplemented with serum, but lacking the label. Centrifuge and remove the medium.

- 3b. Add 2 ml medium per 10⁷ cells and transfer to an appropriate size petri dish.

4. Working behind a Plexiglas shield, use a micropipettor with plugged, aerosol-resistant pipet tips to add ³²P_i to a final concentration of 0.1 to 2 mCi/ml.

Use of plugged, aerosol-resistant pipet tips will minimize contamination of the micropipettor. ³²P_i is usually shipped in HCl and is generally sterile, so it is not necessary to sterilize the labeling medium after adding the radioisotope. Additionally, except in the rare cases when steady-state labeling with ³²P is desired, the labeling interval is usually so short (<6 hr) that microbial growth resulting from added isotope is undetectable. Therefore, addition of ³²P to cells can be performed on the lab bench rather than in a tissue culture hood.

5. Place dishes in a warmed Plexiglas box, and put box in the incubator.

Labeling for 1 to 2 hr is usually sufficient, but cells will tolerate as much as 2 mCi/ml for 6 hr and lower concentrations (0.1 to 0.5 mCi/ml) for 18 hr.

Wash and lyse the cells

6. At the end of the labeling period, carry the labeled cells, still in the Plexiglas box, into the cold room. Place the cells behind a Plexiglas shield.

For adherent cells:

- 7a. Take the dish out of the box and remove the labeling medium manually, using either a plugged disposable pipet or a disposable one-piece plastic Pasteur transfer pipet. Discard the medium and pipet as radioactive wastes.

Labeling medium should be removed without using a vacuum aspirator. Vacuum aspiration generates radioactive aerosols and leaves a radioactive film on the equipment.

- 8a. Wash cells once with 2 to 10 ml cold TBS. Remove the wash buffer manually, as in step 7, and discard as radioactive waste.

The uptake of phosphate is efficient and often a majority of the added $^{32}\text{P}_i$ is found within the labeled cells. This necessitates continued shielding of the labeled cells following removal of the labeling medium.

- 9a. Add lysis buffer to cells, using 0.3 ml per 35-mm dish, 0.6 ml per 50-mm dish, or 1.0 ml per 100-mm dish of adherent cells. Dislodge adherent cells by scraping with a rubber policeman, but leave the lysate in the dish. Incubate 20 min at 4°C. With a rubber policeman, scrape the lysate of adherent cells to the side of the dish and transfer lysate to a screw-cap microcentrifuge tube.

If a low background of nonspecific contaminants is critical, use RIPA lysis buffer. If maintenance of enzymatic activity or the structure of protein complexes is critical, use a milder lysis buffer containing either 3[(3-cholamidopropyl)-dimethylammonio]-1-propane-sulfonate (CHAPS) or Nonidet P-40 (NP-40) as the only detergent. If complete solubilization of the cells and denaturation of the protein is desired, use SDS for lysis (see Alternate Protocol).

For nonadherent cells:

- 7b. Take the dish out of the box, transfer the cells to a screw-cap centrifuge tube, pellet them by centrifugation (1 min at 1800 × g), and remove the medium.

Do not use a vacuum aspirator to remove medium.

- 8b. Resuspend pelleted cells gently in a small volume of cold TBS, transfer to a screw-cap microcentrifuge tube, and pellet the cells by microcentrifuging 1 min at 1800 × g.

Continue to use shielding with cells and lysate.

- 9b. Add 0.5 to 1 ml lysis buffer per 10⁷ cells and resuspend the pellet by gentle agitation with a disposable plastic Pasteur pipet. Incubate 20 min at 4°C.

Choose the lysis buffer as in annotation to step 9a.

10. Cap the tube, and clarify the lysate by centrifuging 30 min at 26,000 × g (17,000 rpm in Sorvall SM24 rotor), 4°C.

Use an ice bucket with a sheet of Plexiglas over it, a chilled Plexiglas tube holder, or a prechilled centrifuge rotor to transport the tube of lysate.

It is best to half-fill the tube to prevent spilling. A Sorvall SM24 rotor with rubber adaptors for microcentrifuge tubes is ideal for clarifying lysates. Alternatively, a refrigerated microcentrifuge can be used for 30 min at maximum speed. Do not use a nonrefrigerated microcentrifuge in the cold room because the centrifuge will warm to 20°C during a 30-min spin.

*Lysates prepared with RIPA buffer often become viscous due to lysis of nuclei. If this occurs, increase the time of centrifugation to 90 min, or add 50 μl of fixed *Staphylococcus aureus* bacteria (Pansorbin, Calbiochem) in RIPA buffer to the lysate prior to centrifugation. Either modification will cause the solubilized DNA to pellet.*

11. After centrifugation, transfer the supernatant (lysate) to a new tube and discard the tube and pellet in radioactive waste.

If the pellet is very viscous, so that it is impossible to remove the supernatant cleanly, suck part of the pellet into a micropipet tip, lift the viscous material out of the tube, and discard it in the radioactive waste. The residual liquid in the tube is the supernatant and can be used for immunoprecipitation.

12. Analyze the labeled lysate using gel electrophoresis, immunoprecipitation, or protein purification. Carry out all analytical procedures at 4°C using adequate shielding.

ALTERNATE PROTOCOL

LYSIS OF CELLS BY BOILING IN SDS

Some proteins, such as eukaryotic RNA polymerase II, are difficult to solubilize with mild lysis buffer or RIPA lysis buffer, and some analytical procedures use antibodies that recognize epitopes exposed only in denatured proteins. In these cases, it is useful to solubilize labeled cells completely in SDS and then adjust the composition of the lysate solution to match that of RIPA buffer for immunoprecipitation. To avoid the formation of spurious disulfide bonds, lysis and washing during immunoprecipitation are carried out in the presence of fresh 1 mM dithiothreitol (DTT). The protocol describes the procedure for adherent cells; modifications for working with nonadherent cells are described as alternate steps.

Additional Materials (also see Basic Protocol)

SDS lysis buffer (see recipe)

RIPA correction buffer (see recipe)

Immunoprecipitate wash buffer (see recipe)

Fixed *Staphylococcus aureus* bacteria (Pansorbin, Calbiochem; optional)

Boiling water bath

1. Label and wash cells (see Basic Protocol steps 1 to 7).
- 2a. *For adherent cells:* Add SDS lysis buffer using 0.1 ml for a 35-mm dish, 0.25 ml for a 50-mm dish, or 0.5 ml for an 100-mm dish. Immediately scrape off the dish with a rubber policeman and transfer the cell lysate to a screw-cap microcentrifuge tube.
- 2b. *For nonadherent cells:* Vortex briefly to loosen the cell pellet, add 1 ml SDS lysis buffer per 5×10^7 cells and vortex again.
3. Boil the samples 2 to 5 min, then add 4 vol RIPA correction buffer and mix well.
4. Clarify the cell lysate by centrifuging 90 min at $26,000 \times g$ (17,000 rpm in Sorvall SM24), 4°C or at maximum speed, 4°C in a refrigerated microcentrifuge.

*Lysate may also be clarified by adding 50 μ l fixed *Staphylococcus aureus* and centrifuging 30 min at $26,000 \times g$, 4°C.*

5. Carry out immunoprecipitation as usual using immunoprecipitate wash buffer for washes.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Immunoprecipitate wash buffer for boiled sample

RIPA lysis buffer (see recipe)

1 mM DTT, added fresh

DTT is added from a 1 M stock solution stored at -20°C.

Mild lysis buffer

10 mM 3[(3-cholamidopropyl)-dimethylammonio]-1-propane-sulfonate (CHAPS) or 1% (w/w) Nonidet P-40 (NP-40)

0.15 M NaCl

0.01 M sodium phosphate, pH 7.2 (APPENDIX 2A)

2 mM EDTA

50 mM sodium fluoride

0.2 mM sodium vanadate added fresh from 0.2 M stock solution

100 U/ml aprotinin (Trasylol, Pentex/Miles)

Store buffer without vanadate at 4°C up to 1 year

CHAPS is a milder detergent than NP-40, but yields precipitates with a higher background and may solubilize some proteins less efficiently.

Sodium vanadate stock solution can be stored in plastic at room temperature.

Phosphate-buffered saline (PBS)

Dissolve the following in 800 to 900 ml H₂O:

8 g sodium chloride (136.8 mM final)

0.2 g potassium chloride (2.5 mM final)

0.115 g dibasic sodium phosphate (anhydrous; 0.8 mM final)

0.2 g monobasic potassium phosphate (anhydrous; 1.47 mM final)

0.1 g calcium chloride (anhydrous; 0.9 mM final)

0.1 g magnesium chloride hexahydrate (0.5 mM final)

Add H₂O to 1 liter

Dispense 50- or 100-ml aliquots into glass bottles

Autoclave and store indefinitely at room temperature

RIPA (RadioImmunoProtection Assay) correction buffer for boiled sample

1.25% (w/w) Nonidet P-40 (NP-40)

1.25% (w/v) sodium deoxycholate

0.0125 M sodium phosphate, pH 7.2

2 mM EDTA

0.2 mM sodium vanadate added fresh from 0.2 M stock solution

50 mM sodium fluoride

100 U/ml aprotinin (Trasylol, Pentex/Miles)

Store buffer without vanadate at 4°C up to 1 year

Sodium vanadate stock solution can be stored in plastic at room temperature.

RIPA (RadioImmunoPrecipitation Assay) lysis buffer

1% (w/w) Nonidet P-40 (NP-40)

1% (w/v) sodium deoxycholate

0.1% (w/v) SDS

0.15 M NaCl

0.01 M sodium phosphate, pH 7.2

2 mM EDTA

50 mM sodium fluoride

continued

0.2 mM sodium vanadate added fresh from 0.2 M stock solution
100 U/ml aprotinin (Trasylol, Pentex/Miles)
Store buffer without vanadate at 4°C up to 1 year

Sodium vanadate stock solution can be stored in plastic at room temperature.

SDS lysis buffer

0.5% (w/v) SDS
0.05 M Tris·Cl, pH 8.0 (APPENDIX 2)
1 mM DTT, added fresh

SDS and Tris·Cl solutions can be made in advance and stored at room temperature. DTT is added from a 1 M stock solutions stored at -20°C.

Tris-buffered saline (TBS)

Dissolve the following salts in 800 to 900 ml H₂O:
8 g sodium chloride (136.8 mM final)
0.38 g potassium chloride (5.0 mM final)
0.1 g calcium chloride (anhydrous; 0.9 mM final)
0.1 g magnesium chloride hexahydrate (0.5 mM final)
0.1 g dibasic sodium phosphate (anhydrous; 0.7 mM final)
Add 25 ml 1 M Tris·Cl, pH 7.4 (APPENDIX 2)
Add H₂O to 1 liter
Dispense 50- or 100-ml aliquots into glass bottles
Autoclave and store indefinitely at room temperature

COMMENTARY

Background Information

The extent to which a protein becomes radiolabeled via biosynthetic labeling with ³²P_i depends on the rate of transport of phosphate into the cells being labeled, abundance of the protein, stoichiometry of phosphorylation of the protein, and rate of phosphate turnover in the protein. The rate of turnover of the protein itself is less important because phosphate in protein usually turns over at a much faster rate than does the protein. Most of the phosphate incorporated into cells is not incorporated into phosphoproteins—the vast majority is incorporated into phospholipid, RNA, and DNA. The efficiency of incorporation of ³²P_i into proteins is therefore low—only ~1% of incorporated radioactivity is found in phosphoproteins.

An obvious and immediate question is how many cells and how much isotope should be used. The answer is that it depends on the protein being studied. If the protein is abundant (i.e., it constitutes 0.5% to 1% of total cellular protein) and highly phosphorylated (i.e., it contains 1 mole of phosphate per mole of protein), it can be labeled by incubating 10⁶ cells with 0.2 to 0.5 mCi ³²P_i in 2 ml labeling medium for 1 hr. In contrast, if the protein is rare or non-abundant, it may be necessary to label 10⁷ cells with 5 to 10 mCi of ³²P_i in 3 to 5 ml labeling medium for 4 to 6 hr. A pilot experiment can

be useful: if 50 cpm are recovered in the protein of interest, use more cells and more isotope; if 200,000 cpm are recovered in the protein of interest, use fewer cells with less isotope.

Critical Parameters

Incorporation of ³²P_i during biosynthetic labeling is greatest if labeling is done in medium lacking any phosphate except the added radioisotope. Labeling can also be accomplished at a somewhat reduced efficiency in complete growth medium. Phosphate-free medium is commercially available but can also be easily formulated in-house by omitting sodium and potassium phosphate from a recipe and replacing them with either sodium chloride or potassium chloride or both. Serum dialyzed against phosphate-free saline can be used instead of complete serum to further reduce the level of phosphate in labeling medium. Dialysis of serum against phosphate-free saline reduces but never entirely eliminates the phosphate in serum. Washing or rinsing cells in reduced-phosphate medium just prior to labeling significantly increases labeling efficiency, but starvation of cells by incubation in reduced-phosphate medium prior to labeling is of very limited value.

Labeling cells with ³²P almost certainly induces radiation damage in the labeled cells and

will affect or arrest cell-cycle progression. If phosphorylation of the protein under study varies during the cell cycle, labeling with ^{32}P may alter its phosphorylation.

At the end of the labeling interval, the radioactive medium should be discarded, except in experiments where labeled virions or labeled secreted proteins are being studied. Labeling medium should be removed *without* using a vacuum aspirator. Vacuum aspiration generates radioactive aerosols and leaves a radioactive film inside the aspirator hose. As a result the hose can become an intense source of radiation exposure. It is best to use a plugged disposable pipet or disposable plastic Pasteur pipet for removing labeling medium. The medium and pipet should be discarded immediately as radioactive wastes.

In general, labeled cells are lysed and subjected to gel electrophoresis, immunoprecipitation, or protein purification. If the labeled cells are to be centrifuged, they should be transferred to capped, disposable centrifuge tubes. Tubes should be no more than half full if they are to be spun in a fixed-angle rotor; tubes that are too full may leak radioactive lysate and contaminate the centrifuge. Screw-cap microcentrifuge tubes are ideal for centrifuging lysates. If large volumes are being handled, multiple partially filled tubes are preferable to tubes filled to the top.

An ice bucket can be used for storage and transport of samples to keep the samples cold and provide considerable shielding. Radiation exposure from the tops of the tubes can be minimized by covering the ice bucket with a 1/4-in. Plexiglas sheet.

RadioImmunoPrecipitation Assay (RIPA) buffer (Brugge and Erikson, 1977) is the lysis buffer of choice because it solubilizes proteins well, gives a low background of nonspecific proteins, and is tolerated by most antigens and antibodies. It does however denature some antigens and disrupt some protein:protein complexes. If 1% (w/w) Nonidet P-40 (NP-40) is used as the only detergent, most of these problems are solved without increasing background unacceptably (Sefton et al., 1980). Some workers like to use digitonin as an extremely mild detergent for cell lysis. Digitonin sometimes is tricky and idiosyncratic, giving erratic results and high backgrounds during immunoprecipitation. 3[(3-cholamidopropyl)-dimethylammonio]-1-propane-sulfonate (CHAPS; Chen et al., 1990) and Brij 96 (Osman et al., 1992) are mild detergents that appear to give more reproducible results.

Besides the usual concern about proteolysis following cell lysis, the cells must be handled carefully to prevent protein phosphorylation and dephosphorylation following lysis of the cells. Both problems can be minimized by proper formulation of the lysis buffer and by keeping the sample at 4°C. Inclusion of 2 mM EDTA in the lysis buffer will minimize phosphorylation in the lysate by chelating both Mg^{2+} and Mn^{2+} , which are essential for protein kinase activity. Addition of phosphatase inhibitors to the lysis buffer will reduce dephosphorylation significantly. For example, 50 mM sodium fluoride is used to inhibit serine and threonine phosphatases, and 0.2 mM sodium vanadate is used to inhibit all known tyrosine phosphatases. Sodium vanadate appears to lose its effectiveness if it is stored diluted in lysis buffer, so it should be added fresh each time. The concentrated stock solution should be stored in plastic at room temperature.

Anticipated Results

This protocol can be used to label phosphoproteins or other phosphorylated cellular constituents. Phosphoproteins will contain ~1% of incorporated label.

Time Considerations

Proteins can and should be labeled biosynthetically with $^{32}\text{P}_i$ and isolated by immunoprecipitation in a single day. As is the case with all ^{32}P -labeled samples, it is important to work quickly because the specific activity of the sample declines 5% per day. If it is absolutely necessary to stop during the immunoprecipitation, it is best to leave the samples as precipitated pellets after aspiration of the wash buffer. Such samples can be stored at 4°C or -20°C overnight. Freezing the cell lysate tends to increase the background. Frozen lysates should be reclarified by centrifugation prior to immunoprecipitation.

Literature Cited

- Brugge, J.S. and Erikson, R.L. 1977. Identification of a transformation-specific antigen induced by an avian sarcoma virus. *Nature* 269:346-348.
- Chen, J., Stall, A.M., Herzenberg, L.A., and Herzenberg, L.A. 1990. Differences in glycoprotein complexes associated with IgM and IgD on normal murine B cells potentially enable transduction of different signals. *EMBO J.* 9:2117-2124.
- Osman, N., Ley, S.C., and Crumpton, M.J. 1992. Evidence for an association between the T cell receptor/CD3 antigen complex and the CD5 antigen complex in human T lymphocytes. *Eur. J. Immunol.* 22:2995-3000.

Sefton, B.M., Hunter, T., and Beemon, K. 1980.
Temperature-sensitive transformation by Rous
sarcoma virus and temperature-sensitive protein
kinase activity. *J. Virol.* 33:220-229.

Contributed by Bartholomew M. Sefton
The Salk Institute
San Diego, California

**Labeling Cells
with $^{32}\text{P}_i$ and
Preparing Lysates**

18.2.8

Phosphoamino Acid Analysis

UNIT 18.3

It is often valuable to identify the phosphorylated residue in a protein. In the case of proteins phosphorylated at serine, threonine, or tyrosine, this is readily accomplished by partial acid hydrolysis in HCl followed by two-dimensional thin-layer electrophoresis of the labeled phosphoamino acid (see Basic Protocol). Phosphothreonine and phosphotyrosine are more stable to hydrolysis in alkali than are RNA and phosphoserine. Therefore, mild alkaline hydrolysis of protein samples can be used to enhance the detection of phosphothreonine and phosphotyrosine (see Alternate Protocol).

Although this procedure can be carried out with a protein eluted from a preparative gel and concentrated by trichloroacetic acid (TCA; UNIT 10.18) or acetone (UNIT 17.10) precipitation, it is most easily accomplished by transfer of the protein of interest to a PVDF membrane. This technique is obviously not ideal if the protein being studied does not transfer efficiently.

NOTE: Wear gloves and use blunt-end forceps to handle membranes.

ACID HYDROLYSIS AND TWO-DIMENSIONAL ELECTROPHORETIC ANALYSIS OF PHOSPHOAMINO ACIDS

**BASIC
PROTOCOL**

The protein to be acid hydrolyzed is transferred to a PVDF membrane using the same technique used for immunoblotting (UNIT 10.8) or for microsequencing (UNIT 10.19). It is useful, but not absolutely essential, to keep the filter wet following transfer. Following acid hydrolysis, phosphoamino acids are separated by two-dimensional thin-layer electrophoresis. Because electrophoresis equipment differs considerably in design, the details of the assembly and placement of the plate are not discussed here. It is assumed that a suitable apparatus is available for use by an experienced operator. Electrophoresis conditions are described for using the HTLE 7000 (CBS Scientific). They are almost certainly not correct for other equipment and will need to be altered according to the equipment manufacturer's directions.

Materials

³²P-labeled phosphoprotein (UNIT 18.2)

India ink solution: 1 μl/ml India ink in TBS (UNIT 18.2)/0.02% (v/v) Tween 20, pH 6.5 (prepare fresh or store indefinitely at room temperature); *or* radioactive or phosphorescent alignment markers

6 M HCl

Phosphoamino acid standards mixture (see recipe)

pH 1.9 electrophoresis buffer (see recipe)

pH 3.5 electrophoresis buffer (see recipe)

0.25% (w/v) ninhydrin in acetone in a freon (aerosol, gas-driven) atomizer/sprayer

PVDF membrane (Immobilon-P, Millipore)

110° oven

Screw-cap microcentrifuge tubes

20 cm × 20 cm × 100 μm glass-backed cellulose thin-layer chromatography plate (EM Sciences)

Large blotter: two 25 × 25-cm layers of Whatman 3MM paper sewn together at the edges, with four 2-cm holes that align with the origins on the TLC plate

Glass tray or plastic box

Whatman 3MM paper

Thin-layer electrophoresis apparatus (e.g., HTLE 7000, CBS Scientific)

Fan

Small blotters: 4 × 25-cm, 5 × 25-cm, and 10 × 25-cm pieces of Whatman 3MM paper

Contributed by Bartholomew M. Sefton

Current Protocols in Molecular Biology (1997) 18.3.1-18.3.8

Copyright © 1997 by John Wiley & Sons, Inc.

**Analysis of
Protein
Phosphorylation**

18.3.1

Supplement 40

50° to 80°C drying oven
Sheets of transparency film for overhead projector

Additional reagents and equipment for SDS-PAGE (UNIT 10.2), immunoblotting (UNIT 10.8 or UNIT 10.19), staining filters (UNIT 10.7), and autoradiography (APPENDIX 3A)

Prepare sample

1. Run radiolabeled phosphoprotein on a preparative SDS-polyacrylamide gel.
2. Transfer proteins electrophoretically to a PVDF membrane. Wash the membrane several times with water. Do not let the membrane dry.

These washes remove buffer and detergent.

3. Locate the band of interest by staining the filter 5 to 10 min in 30 to 50 ml India ink solution with shaking until bands are detectable, *or* by wrapping the filter in plastic wrap, applying radioactive or phosphorescent alignment markers, and performing autoradiography.
4. Excise the piece of filter containing the band of interest with a clean razor blade. Rewet the piece of filter with methanol for 1 min and then rewet it in >0.5 ml water. Place the piece of filter paper in a screw-cap microcentrifuge tube.

Keep the excised piece as small as possible.

Hydrolyze sample

5. Add enough 6 M HCl to submerge the piece of filter. Screw the cap on the tube tightly and incubate 60 min in 110°C oven.
6. Let cool. Microcentrifuge 2 min at maximum speed. Transfer the liquid hydrolysate to a fresh microcentrifuge tube and dry with a Speedvac evaporator.

Drying takes ~2 hr. Simultaneous drying of the hydrolysate and deblocked oligonucleotides in NH₄OH must be avoided, as this will generate a cloud of ammonium chloride that will collect in the centrifuge tube and render the hydrolysate unsuitable for thin-layer electrophoresis.

7. Dissolve the sample in 6 to 10 µl water by vortexing vigorously. Microcentrifuge 5 min at maximum speed.

Prepare plate for first-dimension electrophoresis

8. Spot 25% to 50% of the sample, in 0.25- to 0.50-µl aliquots, on one origin of a 20 cm × 20 cm × 100 µm glass-backed cellulose thin-layer chromatography plate (see Fig. 18.3.1 for arrangement of samples). Between each application, dry the sample spot with compressed air delivered through a Pasteur pipet plugged with cotton.

Use long, thin plastic micropipet loading tips for loading, and do not let the tip touch the plate.

Four samples can be analyzed simultaneously. The complete hydrolysate can be spotted on a single origin, but some streaking in the first dimension may be observed due to overloading. This problem can be avoided by using a fraction of the sample.

9. Spot 1 µl nonradioactive phosphoamino acid standards mixture (containing phosphoserine, phosphothreonine, and phosphotyrosine) on top of each sample in 0.25- to 0.50-µl aliquots as above.
10. Wet the large blotter (with four holes) by submerging it in pH 1.9 electrophoresis buffer in a large glass tray or plastic box. Briefly allow the excess buffer to drain off. Lower the wet blotter onto the prespotted plate with the origins on the plate in the centers of the four holes in the blotter (Fig. 18.3.1). Press on the blotter gently to

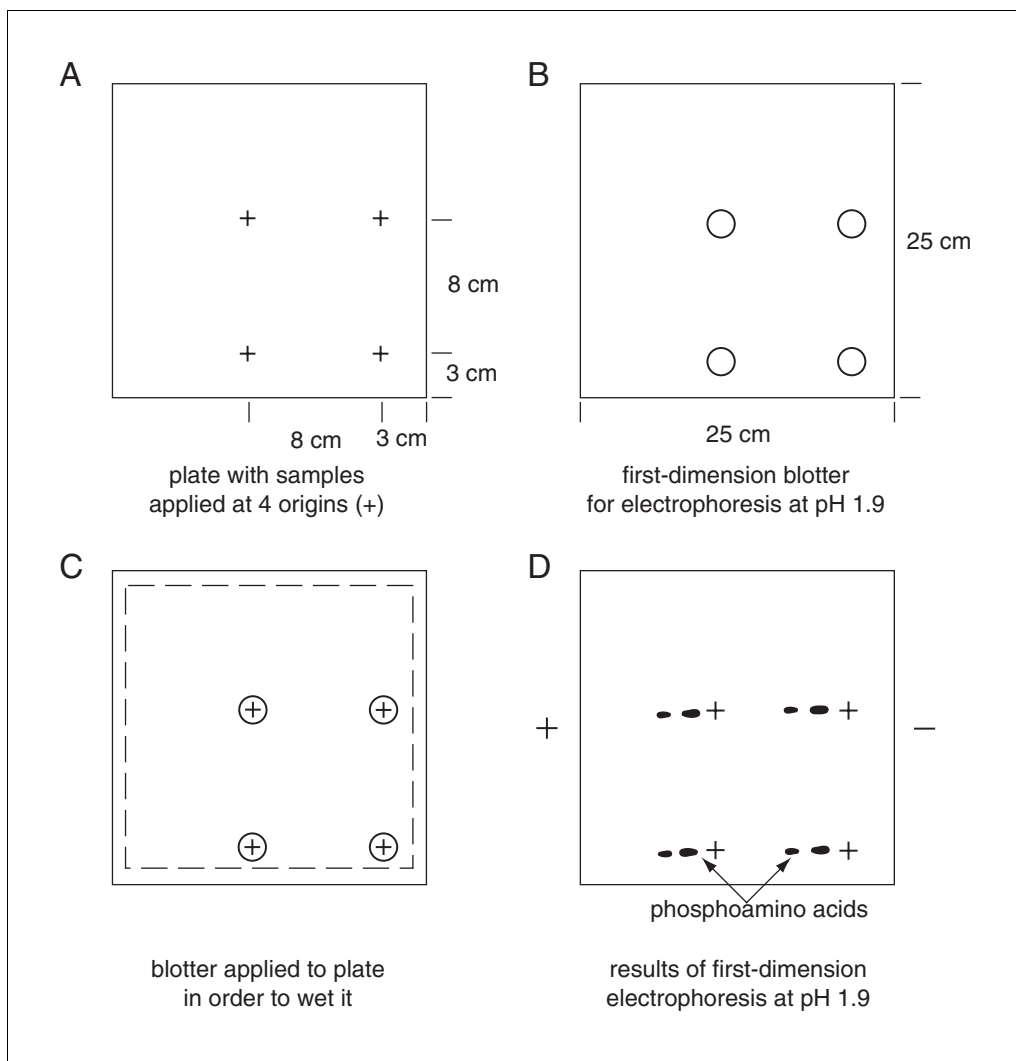


Figure 18.3.1 First-dimension electrophoretic separation of phosphoamino acids at pH 1.9. **(A)** Positions of the four origins on a single 20 × 20-cm plate; **(B)** blotter used for wetting the plate with pH 1.9 electrophoresis buffer; **(C)** placement of the blotter on the plate (underneath; indicated by dashed outline); and **(D)** orientation of the plate between the + and – electrodes with the positions of the phosphoamino acids after electrophoresis.

achieve even wetting of the cellulose and concentration of the samples. When the plate is uniformly wet, remove the blotter.

The blotter should be quite damp but not sopping wet. Excess buffer can be wicked off onto filter paper.

Areas of the plate that are too dry can be seen through the blotter and will appear to be whiter than the rest of the plate. If this happens, dab the blotter with a Kimwipe wetted with pH 1.9 electrophoresis buffer. If there are puddles of buffer on the plate, let them dry before carrying out electrophoresis.

- Place the thin-layer plate in the electrophoresis apparatus and overlap 0.5 cm of the right and left sides of the plate with wicks made of Whatman 3MM paper. If the apparatus has an air bag, be sure to inflate it. Close the cover and start electrophoresis. With an HTLE 7000, double-thickness Whatman 3MM wicks, and a plate with four samples, electrophorese 20 min at 1.5 kV.

For the HTLE 7000 apparatus, use folded-over Whatman 3MM wicks that are 20 cm wide (the same as the plate) and not overly wet. Overly wet wicks will flood the plate and cause sample diffusion.

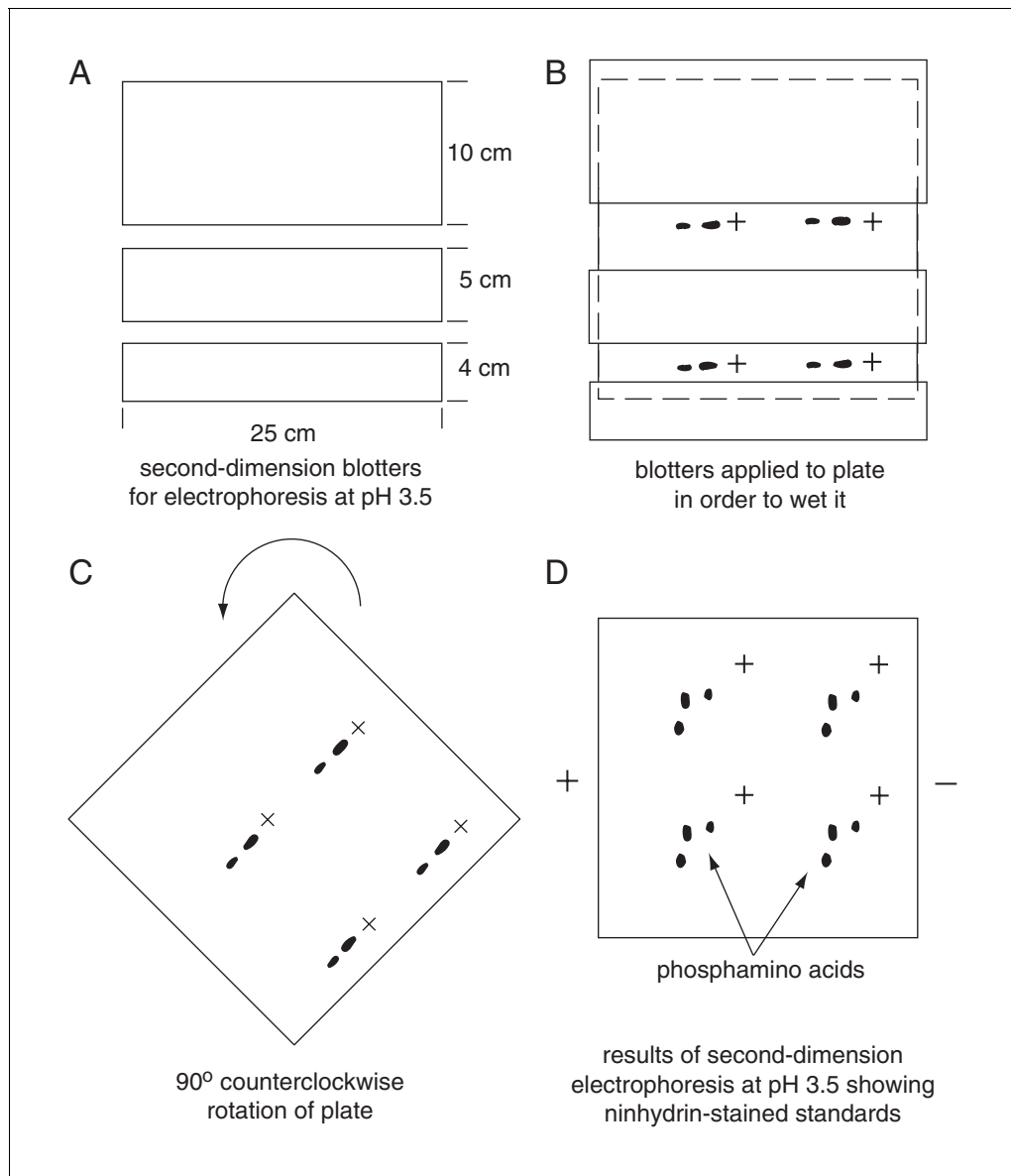


Figure 18.3.2 Second-dimension electrophoretic separation of phosphoamino acids at pH 3.5. **(A)** The three pieces of Whatman 3MM paper used for wetting the plate with pH 3.5 electrophoresis buffer; **(B)** proper placement of the blotters on the plate (underneath; indicated by dashed outline); **(C)** reorientation of the plate for electrophoresis in the second dimension; and **(D)** orientation of the plate between the + and – electrodes with the position of the phosphoamino acids after electrophoresis.

For other electrophoresis apparatuses the appropriate duration of electrophoresis can be determined empirically by examining the rate of migration of the phosphoamino acid standards.

12. Following electrophoresis, remove the plate and quickly air dry with a fan without heating.

It takes ~20 min to dry the plate.

Perform second-dimension electrophoresis

13. Wet the small blotters in pH 3.5 electrophoresis buffer and use them to wet the plate using the method described in step 10 to achieve even wetting without puddling (Fig. 18.3.2).

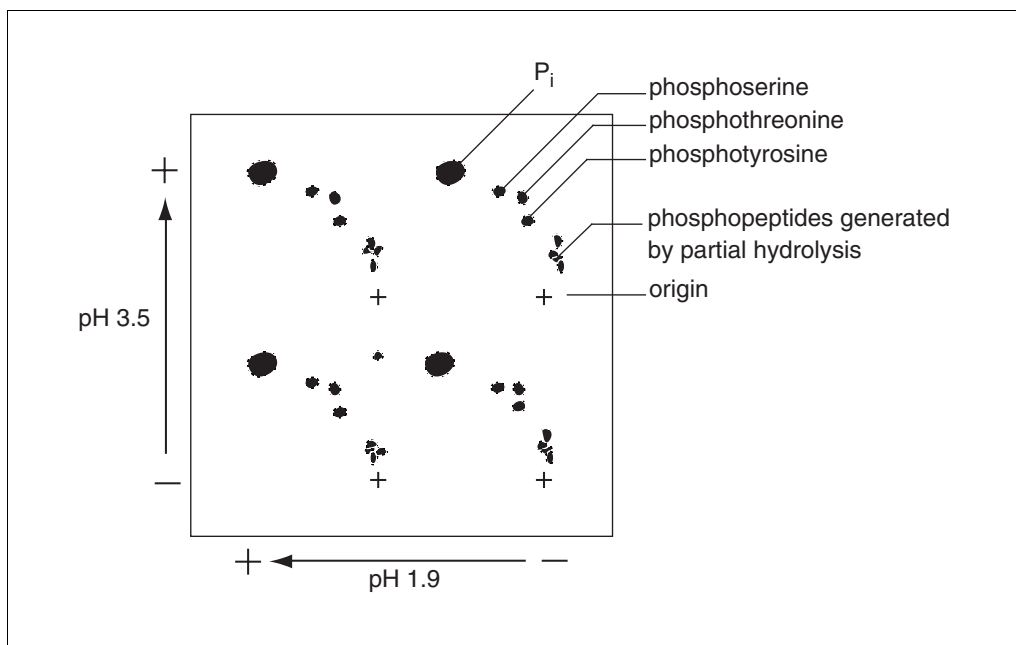


Figure 18.3.3 Hypothetical autoradiogram of a two-dimensional separation. Four samples of acid-hydrolyzed, ^{32}P -labeled proteins are applied at the origins, one in each of the four quadrants. This diagram shows the origins, the directions of electrophoresis, the positions of phosphoserine, phosphothreonine, and phosphotyrosine, the position of P_i , and the position of partially hydrolyzed fragments of the proteins for the upper right-hand sample. Every protein generates different partial hydrolysis peptide fragments.

After electrophoresis at pH 1.9, phosphoamino acids are present as a streak extending from the origin towards the + electrode. Blotters are not applied directly over the phosphoamino acids to prevent sample blurring or smearing.

14. Remove the blotters, rotate the plate 90° counterclockwise, and electrophorese 16 min at 1.3 kV in pH 3.5 electrophoresis buffer if using the HTLE 7000 apparatus.
15. At the end of the electrophoresis run, remove the plate and dry 20 to 30 min in an oven at 50° to 80°C . When dry, spray with 0.25% ninhydrin in acetone, then reheat in the oven 5 to 10 min to visualize the phosphoamino acid standards.
16. Place radioactive or phosphorescent alignment marks on the plate and autoradiograph with an intensifying screen overnight to 10 days at -70°C .
17. Following autoradiography, trace the alignment markers and the stained phosphoamino acid markers onto a transparent sheet used for overhead projectors. Save this template. Align the film with the plate and identify radioactive phosphoamino acids (Fig. 18.3.3).

ALKALI TREATMENT TO ENHANCE DETECTION OF TYR- AND THR-PHOSPHORYLATED PROTEINS BLOTTED ONTO FILTERS

Because phosphothreonine and phosphotyrosine are much more stable to hydrolysis in alkali than RNA or phosphoserine, detection of proteins containing phosphothreonine and phosphotyrosine can often be enhanced by mild alkaline hydrolysis of gel-fractionated samples. Although this technique was first developed for the treatment of fixed polyacrylamide gels, it is much more easily performed with proteins that have been first transferred to a PVDF membrane.

ALTERNATE PROTOCOL

Analysis of Protein Phosphorylation

18.3.5

Alkaline hydrolysis does not preclude subsequent phosphoamino acid analysis. A band from a blot that has been treated with alkali can be excised and subjected to acid hydrolysis as described in the Basic Protocol.

Additional Materials (also see *Basic Protocol*)

1 M KOH

TN buffer: 10 mM Tris·Cl (pH 7.4 at room temperature)/0.15 M NaCl

1 M Tris·Cl, pH 7.0 at room temperature (*APPENDIX 2*)

Covered plastic container (e.g., Tupperware box)

55°C oven or water bath

1. Run radiolabeled phosphoprotein on a preparative SDS–polyacrylamide gel and transfer proteins electrophoretically to a PVDF membrane (see *Basic Protocol* steps 1 and 2).

A nylon membrane may be used in place of a PVDF membrane, but in that case, the bands cannot subsequently be analyzed by acid hydrolysis, as nylon membrane will dissolve in 6 M HCl.

2. Wash membrane thoroughly with water: three 2-min incubations in 1 liter water are sufficient.

These washes remove buffer and detergent.

3. Incubate membrane 120 min at 55°C in an oven or water bath in sufficient 1 M KOH to cover the filter in a covered Tupperware container.

4. Discard KOH. Wash membrane and neutralize remaining KOH by rinsing once for 5 min in 500 ml TN buffer, once for 5 min in 500 ml of 1 M Tris·Cl (pH 7.0), and twice for 5 min in 500 ml water. Wrap the membrane in plastic wrap and autoradiograph overnight with flashed film and an intensifying screen at –70°C.

REAGENTS AND SOLUTIONS

*Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.*

pH 1.9 electrophoresis buffer

50 ml 88% formic acid (0.58 M final concentration)

156 ml glacial acetic acid (1.36 M final concentration)

1794 ml H₂O

Store indefinitely in a sealed bottle at room temperature

pH 3.5 electrophoresis buffer

100 ml glacial acetic acid (0.87 M final concentration)

10 ml pyridine [0.5% (v/v) final concentration]

10 ml 100 mM EDTA (0.5 mM final concentration)

1880 ml H₂O

Store indefinitely in a sealed bottle at room temperature

Phosphoamino acid standards mixture

Prepare a solution of phosphoserine, phosphothreonine, and phosphotyrosine (Sigma) in water at a final concentration of 0.3 µg/ml each. Store in 1-ml aliquots indefinitely at –20°C.

COMMENTARY

Background Information

Phosphoamino acid analysis by the two-dimensional electrophoretic technique described in the basic protocol was first carried out with proteins isolated by elution from unfixed SDS-polyacrylamide gels (Hunter and Sefton, 1980). However, this technique is laborious, especially if it involves grinding up pieces of high-percentage acrylamide gels, and the yields can be disappointing. Additionally, because the eluted protein must be precipitated in the presence of a carrier protein, spotting the whole sample on a single origin usually yields a badly smeared pattern. The grind-and-elute technique is, however, advantageous with proteins that are very refractory to electrophoretic transfer to PVDF membranes.

The alkaline treatment of protein described in the alternate protocol was first developed by Jon Cooper and Tony Hunter, who treated fixed gels with KOH (Cooper and Hunter, 1981). The original technique is tricky because the gel becomes extremely sticky during incubation with KOH and swells. Additionally, the manipulations needed to recover proteins from the gel following treatment are very involved because the proteins are contaminated with products of the hydrolysis of polyacrylamide.

Critical Parameters and Troubleshooting

It is essential to use PVDF membranes to immobilize proteins for acid hydrolysis rather than nylon or nitrocellulose membranes, both of which dissolve in 6 M HCl. Proteins immobilized on either PVDF or nylon membranes may be subjected to alkaline hydrolysis with KOH (Contor et al., 1987), but nitrocellulose membranes are not suitable. Proteins immobilized on nylon cannot subsequently be analyzed by acid hydrolysis because nylon is dissolved by 6 M HCl.

Two-dimensional thin-layer electrophoresis is required for unambiguous identification of phosphorylated residues, as some spots after one-dimensional electrophoresis do not represent pure species. For example, uridine monophosphate, which is generated during acid hydrolysis of RNA (a frequent contaminant of phosphoproteins), comigrates with phosphotyrosine during one-dimensional electrophoresis at pH 3.5.

Streaking of the sample in the first dimension is a symptom of overloading, either with the phosphoprotein itself or with contaminants

in the sample. This problem can be corrected by loading less sample. Streaking in the second dimension is usually the result of problems with wetting or running the plate and cannot be corrected by loading less sample.

Some batches of blotting paper contain calcium, which interferes with electrophoresis of phosphoamino acids (probably by precipitating them). In the author's experience Whatman 3MM paper is quite reliable; other blotting papers are probably suitable as well. Inclusion of EDTA in the pH 3.5 buffer alleviates this problem.

This unit calls for glass-backed cellulose thin-layer plates rather than the plastic-backed variety, which are lighter and less expensive. This is because plastic-backed plates can under some circumstances cause sample streaking. They are, however, probably satisfactory for most experiments. If use of plastic-backed plates results in streaking, try glass-backed plates to see if that corrects the problem.

Anticipated Results

To detect a phosphoamino acid by autoradiography, a minimum of 10 cpm must be spotted and the plate exposed for a week with flashed film and an intensifying screen. Only 15% to 20% of the radioactivity in a phosphoprotein is recovered as phosphoamino acids. The majority is present as $^{32}\text{P}_i$, which is released by dephosphorylation of phosphoamino acids, with the remainder being peptide products resulting from partial acid hydrolysis. As a result, the thin-layer plates will contain a number of radioactive spots that are not phosphoamino acids (see Fig. 18.3.3). Partial hydrolysis products remain near the origin during electrophoresis at pH 1.9, but exhibit some mobility at pH 3.5 (see Fig. 18.3.3). After two-dimensional electrophoresis, they are found above the origin and below the phosphoamino acids. $^{32}\text{P}_i$ has a high mobility at both pH 1.9 and pH 3.5 and is found in the upper left-hand corner of each quadrant of the plate (see Fig. 18.3.3). Because of these additional radioactive spots, it is essential to localize internal phosphoamino acid standards by staining with ninhydrin.

Time Considerations

After the preparative gel has been run and the protein transferred to the membrane, isolation of the membrane fragment containing the protein, followed by acid hydrolysis, takes <2 hr. Two to three hours are required for drying

the sample with a Speedvac evaporator. First- and second-dimension electrophoresis and staining the internal phosphoamino acid standards takes no more than ~2 hr. Overnight autoradiographic exposure with flashed film and an intensifying screen is usually sufficient; however, exposures can be carried out for up to 10 days before the background from the screen becomes objectionable.

Detection of phosphorylated protein after alkali treatment, as described in the alternate protocol, is a very quick procedure requiring little more time than it takes to carry out the alkaline hydrolysis. In general, samples of this sort can be detected after overnight exposure with flashed film and an intensifying screen at -70°C .

Literature Cited

Contor, L., Lamy, F., and Lecocq, R.E. 1987. Use of electroblotting to detect and analyze phosphotyrosine containing peptides separated by two-dimensional gel electrophoresis. *Anal. Biochem.* 160:414-420.

Cooper, J.A. and Hunter, T. 1981. Four different classes of retroviruses induce phosphorylation of tyrosines present in similar cellular proteins. *Mol. Cell. Biol.* 1:394-407.

Hunter, T. and Sefton, B.M. 1980. The transforming gene product of Rous sarcoma virus phosphorylates tyrosine. *Proc. Natl. Acad. Sci. U.S.A.* 77:1311-1315.

Key Reference

Kamps, M.P. and Sefton, B.M. 1989. Acid and base hydrolysis of phosphoproteins bound to Immobilon facilitates the analysis of phosphoamino acids in gel-fractionated proteins. *Anal. Biochem.* 176:22-27.

Discusses all of the variables involved in subjecting filter-bound proteins to acid and base hydrolysis.

Contributed by Bartholomew M. Sefton
The Salk Institute
San Diego, California

Analysis of Phosphorylation of Unlabeled Proteins

UNIT 18.4

Phosphorylation of unlabeled proteins can be detected using immunological or enzymatic techniques. Anti-phosphotyrosine antibodies are used with immunoblots to detect tyrosine phosphorylation. These antibodies can be detected by ^{125}I -labeled protein A (see Basic Protocol 1) or enhanced chemiluminescence (ECL; see Alternate Protocol). Enzymatic dephosphorylation using a general or phosphoamino acid-specific phosphatase is detected by an SDS-PAGE mobility shift (see Basic Protocol 2) or a change in activity. The production of antibodies that recognize specific tyrosine-phosphorylated peptides is described in UNIT 18.6.

IMMUNOBLOTTING WITH ANTI-PHOSPHOTYROSINE ANTIBODIES AND DETECTION USING [^{125}I]PROTEIN A

BASIC
PROTOCOL 1

Initial determination of whether a protein is phosphorylated on tyrosine is carried out by immunoblotting with antibodies to phosphotyrosine. This technique is often superior to the use of biosynthetic labeling with $^{32}\text{P}_i$ because of its sensitivity and low background, and it can be used to detect tyrosine phosphorylation of proteins in tissues that do not incorporate $^{32}\text{P}_i$ efficiently.

Materials

Protein sample: cultured cells, tissue, lysate, or immunoprecipitate

Transfer buffer (UNIT 10.8) containing 100 μM sodium vanadate

Blocking buffer (see recipe)

Anti-phosphotyrosine antibody: 2 $\mu\text{g}/\text{ml}$ rabbit polyclonal anti-phosphotyrosine (UBI) or mouse monoclonal anti-phosphotyrosine—e.g., py20 (Leinco, ICN Biomedicals, Zymed, Transduction Laboratories) or 4G10 (UBI)—in blocking buffer

TN buffer: 10 mM Tris-Cl (pH 7.4 at room temperature)/0.15 M NaCl

TNA solution: TN buffer (recipe above)/0.01% (w/v) sodium azide (store at room temperature)

NP-40 wash solution: 0.05% (v/v) Nonidet P-40 (NP-40)/TNA solution (store at room temperature)

0.5 $\mu\text{Ci}/\text{ml}$ ^{125}I -labeled protein A (30 mCi/mg, ICN) in blocking buffer

India ink solution: 1 μl India ink in TBS (UNIT 18.2)/0.02% (w/v) Tween 20, pH 6.5 (prepare fresh or store indefinitely at room temperature; optional)

Plastic container with lid (e.g., Tupperware box)

Blotting paper

Additional reagents and solutions for SDS-PAGE (UNIT 10.2), immunoblotting (UNIT 10.8), and fluorography (APPENDIX 3A)

NOTE: Wear gloves and use blunt-end forceps to handle membrane.

1. Dissolve cultured cells, tissue, or lysate immunoprecipitate in an equal volume of 2 \times SDS sample buffer and boil 5 min.

Lysates or immunoprecipitates in sample buffer can be stored at -70°C indefinitely. Storage >1 week at -20°C will lead to deterioration. Cells or tissues can be lysed by other means, but care must be taken to prevent phosphorylation or dephosphorylation from occurring following cell lysis. This can be accomplished by using 50 mM sodium fluoride to inhibit serine and threonine phosphatases, 0.2 mM sodium vanadate to inhibit tyrosine phosphatases, and 2 mM EDTA to inhibit kinases. Viscosity resulting from DNA can be reduced

Analysis of
Protein
Phosphorylation

18.4.1

Contributed by Bartholomew M. Sefton

Current Protocols in Molecular Biology (1997) 18.4.1-18.4.7

Copyright © 1997 by John Wiley & Sons, Inc.

Supplement 40

somewhat by boiling the sample 5 min and/or by shearing the sample by passing it five to ten times through a 22-G needle and five to ten times through a 27-G needle.

2. Electrophorese sample on an SDS-polyacrylamide gel. Transfer proteins to a membrane suitable for immunoblotting using transfer buffer that contains 100 μ M sodium vanadate.

The lysate from $\sim 10^5$ cells can be analyzed in a single 0.5-cm wide lane of a 1-mm-thick SDS-polyacrylamide slab gel.

Sodium vanadate should be included in the transfer buffer to prevent tyrosine dephosphorylation.

3. Incubate membrane in 30 to 50 ml blocking buffer ≥ 30 min at room temperature.

Use of Blotto or other blocking mixtures containing dry milk is unsuitable because anti-phosphotyrosine antibodies bind to a number of the protein constituents of milk (see UNIT 10.8 for a discussion of blocking buffers). Blocking buffer can be reused several times.

4. Incubate the membrane in anti-phosphotyrosine antibody 60 to 120 min at room temperature with occasional agitation.

This step is most conveniently done in a covered plastic box using 30 to 50 ml antibody solution for a typical 15×15 -cm blot. Although the volume of antibody solution seems large, it reduces the background. The antibody solution can be saved in the plastic box, stored at 4°C for at least a month, and reused up to five to ten times.

5. Remove the membrane from the staining box and wick off excess liquid with blotting paper. In a new plastic box, wash the membrane twice for 10 min with 50 ml TBSA, twice for 10 min with 50 ml NP-40 wash solution, and twice for 5 min with 50 ml TBSA. Discard washes.

6. Incubate membrane in 30 to 50 ml of 0.5 μ Ci/ml 125 I-labeled protein A in blocking buffer 60 min at room temperature with gentle agitation in a covered plastic box.

125 I-labeled protein A is used to detect bound antibody.

7. Remove the membrane and wick off excess liquid with blotting paper. Wash the membrane as described in step 5 in a clean plastic box.

The blotting paper and first wash should be discarded as radioactive wastes. The amount of radioactivity in the subsequent washes should not be great, and it is generally acceptable to pour them down the drain.

Protein A solution can be stored at least 1 month at 4°C and reused ~ 3 to 5 times.

8. If desired, stain the membrane 5 to 10 min with 30 to 50 ml India ink solution until bands are detectable. Wash the membrane with water to remove excess ink.

9. Remove excess moisture with blotting paper, wrap in plastic wrap, attach radioactive or fluorescent alignment markers, and expose by fluorography using an intensifying screen and preflashed film for 18 hr to 10 days.

DETECTION OF BOUND ANTIBODIES BY ENHANCED CHEMILUMINESCENCE (ECL)

ALTERNATE PROTOCOL

Bound anti-phosphotyrosine antibodies can also be detected using enhanced chemiluminescence (ECL). This technique uses a second antibody coupled to horseradish peroxidase to detect anti-phosphotyrosine antibodies. The bound second antibody is visualized by ECL and the image is captured on film. The technique is faster than that employing [¹²⁵I]-labeled protein A (Basic Protocol 1) because it requires autoradiographic exposure of only minutes rather than days. It also avoids the use of radioactivity. However, results obtained with the two techniques are not always identical and exhibit quantitative differences. Additionally, if immunoprecipitated proteins are being analyzed, the secondary antibody may stain the precipitating antibodies intensely, making detection of proteins of a size similar to those of the heavy and light chains of immunoglobulin difficult.

The technique is fundamentally the same as that using [¹²⁵I]-labeled protein A, except that azide must be omitted from all solutions because it interferes with chemiluminescence chemistry. This makes storage and reuse of antibodies more difficult. Additionally, the technique appears to work somewhat better with nitrocellulose membranes than with Immobilon-P. Therefore, handling of the membranes requires somewhat more delicacy.

Additional Materials (also see Basic Protocol 1)

- Blocking buffer (see recipe) *without* azide
- 0.05% (v/v) Nonidet P-40 (NP-40)/TN buffer (see Basic Protocol 1)
- Horseradish peroxidase–conjugated secondary antibody: anti-rabbit or anti-mouse antibody diluted 1/1000 to 1/2000 in blocking buffer *without* azide
- ECL detection reagents A and B (Amersham)
- Luminol reagent for ECL detection (Amersham)
- Oxidizing reagent for ECL detection (Amersham)
- Nitrocellulose membrane
- Plastic container slightly larger than the membrane
- Plastic sheet protector

NOTE: The reagents for ECL may be purchased as a kit, the Enhanced Chemiluminescence Western Blotting Detection System, from Amersham.

NOTE: Wear gloves and use blunt-end forceps to handle membrane.

Prepare membrane

1. Electrophoresis sample prepared as described in Basic Protocol step 1 on an SDS-polyacrylamide gel. Transfer proteins to a nitrocellulose membrane using a transfer buffer that contains 100 μ M sodium vanadate.
2. In a plastic container, incubate membrane in blocking buffer without azide \geq 30 min at room temperature.

Expose membrane to antibodies

3. Incubate membrane in 30 to 50 ml anti-phosphotyrosine antibody 60 to 120 min at room temperature with occasional agitation.

The anti-phosphotyrosine antibody can be either polyclonal rabbit serum or mouse monoclonal antibodies.

4. In a clean plastic container, wash the membrane twice for 10 min with TN buffer, twice for 10 min with 0.05% NP-40/TN buffer, and twice for 5 min with TN buffer. Discard the washes.

5. In a plastic container, incubate the membrane in 30 to 50 ml of horseradish peroxidase-conjugated secondary antibody 60 min at room temperature with gentle agitation.

Use a secondary antibody appropriate for the primary antibody.

6. Remove membrane and wash it in a clean plastic container as described in step 4.

Detect bound antibodies with ECL

7. In a plastic container only slightly larger than the membrane, mix equal volumes of the Luminol reagent, and the oxidizing reagent for ECL detection.

Amersham recommends that 0.125 ml of each reagent be used per square centimeter of membrane. This is probably excessive with large blots.

8. Put the membrane, with the side containing protein facing up, into the mixed reagents. Agitate gently 60 sec.

9. Remove the blot, wick off excess moisture, and place it, protein-side up, in a sheet protector.

Alternatively, place the blot on a sheet of transparency film used for overhead projectors and cover with a second sheet of transparency film.

10. In the darkroom, expose the blot to X-ray film.

Exposure times can range from 15 sec to 30 min.

BASIC PROTOCOL 2

IDENTIFICATION OF PHOSPHORYLATED PROTEINS BY PHOSPHATASE DIGESTION

Phosphorylation of a protein often alters its mobility during SDS-polyacrylamide gel electrophoresis. An effect of phosphatase digestion on the gel mobility is therefore diagnostic of phosphorylation and is a useful means to determine whether an unlabeled protein is phosphorylated. Additionally, the effect of phosphorylation on the activity of a protein can be assessed by measuring changes in activity following enzymatic dephosphorylation (see *UNIT 18.5* for detailed protocols).

Enzymatic protein dephosphorylation has traditionally been accomplished with alkaline phosphatase or potato acid phosphatase. Both enzymes dephosphorylate phosphoserine, phosphothreonine, and phosphotyrosine. However, phosphoamino acid-specific phosphatases are now commercially available and offer the ability to dephosphorylate a protein specifically at only serine and threonine residues, or at only tyrosine residues (Anderson et al., 1990).

Enzymatic dephosphorylation is intrinsically quite simple. The protein is dissolved in, or in the case of an immunoprecipitated protein suspended in, a buffer suitable for the phosphatase of interest and incubated with the phosphatase for 30 to 60 min. Table 18.4.1 describes the reaction conditions for a number of phosphatases.

Dephosphorylated protein present in an immunoprecipitate can be recovered by centrifugation and analyzed by gel electrophoresis or an enzymatic assay. Alternatively, soluble protein can be analyzed by adding an equal volume of 2× SDS sample buffer (*UNIT 10.2*) to the reaction mix and subjecting the digested protein to gel electrophoresis. Soluble protein can be assayed enzymatically after dephosphorylation if an inhibitor of the phosphatase is added to the reaction mixture. Okadaic acid (20 μM) is a potent and very specific inhibitor of the serine/threonine-specific phosphatase protein phosphatase 2A (PP2A). Sodium vanadate (200 μM) is an efficient inhibitor of all known tyrosine phosphatases. Sodium phosphate (100 mM) is an inhibitor of all protein phosphatases.

Table 18.4.1 Phosphatase Reaction Conditions^a

Phosphatase (supplier ^b)	Amino acid specificity	Buffer	Amount	Incubation conditions	Inhibitor
Bacterial alkaline phosphatase (Pharmacia Biotech)	Phosphoserine, phosphothreonine, and phosphotyrosine	20 mM HEPES, pH 7.0 150 mM NaCl 0.1% (v/v) Triton X-100 10% (v/v) glycerol	2-4 U per reaction	60 min at 30°C	100 mM sodium phosphate
Potato acid phosphatase (Sigma)	Phosphoserine, phosphothreonine, and phosphotyrosine	40 mM PIPES, pH 6.0 1 mM DTT 20 µg/ml aprotinin 20 µM leupeptin	100 µg/ml	10 min at 30°C or 30 min at 4°C	100 mM sodium phosphate
Protein phosphatase 2A (UBI or Calbiochem)	Phosphoserine and phosphothreonine	20 mM HEPES, pH 7.0 1 mM DTT 1 mM MnCl ₂ 100 µg/ml BSA 50 µM leupeptin	10 U/ml	30 to 60 min at 30°C	20 µM okadaic acid
PTP-1B (UBI) or <i>Yersinia</i> PTP (Calbiochem)	Phosphotyrosine	25 mM imidazole-HCl, pH 7.0 1 mg/ml BSA 0.1% (v/v) 2-ME	0.5 mg per reaction	60 min at 30°C	200 µM sodium vanadate

^aAbbreviations: BSA, bovine serum albumin; DTT, dithiothreitol; HEPES, *N*-[2-hydroxyethyl]piperazine-*N'*-[2-ethanesulfonic acid]; 2-ME, 2-mercaptoethanol; PIPES, piperazine-*N,N'*-bis[2-ethanesulfonic acid]; PTP, protein tyrosine phosphatase.

^bSee APPENDIX 4 for supplier contact information.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Blocking buffer

5% (w/v) BSA
1% (w/v) hen ovalbumin
10 mM Tris-Cl, pH 7.4 at room temperature
0.15 M NaCl
0.01% (w/v) sodium azide
Store ≤6 months at 4°C

For enhanced chemiluminescence, omit the sodium azide.

COMMENTARY

Background Information

Immunoblotting with antibodies to phosphotyrosine is an extremely useful technique for the study of tyrosine phosphorylation. It appears to be very specific. Although there is a report that antibodies to phosphotyrosine recognize phosphohistidine (Frackelton et al., 1983), nonspecific cross-reactivity is rarely seen. It is possible to be misled, however, if a large amount of a single protein is subjected to immunoblotting. For example, nonspecific staining of molecular weight standards is often

observed. This almost certainly does not indicate that they contain phosphotyrosine. Rather, it illustrates the fact that the primary antibodies and the secondary detection reagents exhibit a low level of nonspecific binding that can be detectable if a large amount of a single protein is present in a sample.

Critical Parameters

Not all anti-phosphotyrosine antibodies give the same staining patterns for immunoblots of the same sample. Each antibody ap-

**Analysis of
Protein
Phosphorylation**

18.4.5

pears to react with some proteins more strongly than with others. This presumably reflects the amino acid sequence surrounding the phosphotyrosine residue, but the cause has not been examined carefully. Correspondingly, a given anti-phosphotyrosine antibody will not react equally strongly with all sites of tyrosine phosphorylation. For example, rabbit polyclonal anti-phosphotyrosine antibodies raised against phosphoTyr-Gly-Ala react more strongly with phosphorylated Tyr-394 in the Lck tyrosine protein kinase than they do with phosphorylated Tyr-505. The differential reactivity of phosphotyrosine should be kept in mind when drawing quantitative conclusions.

Diluted polyclonal anti-phosphotyrosine antibodies can be reused several times. Their properties, however, change subtly with reuse. The staining of some bands becomes less intense and the staining of others becomes apparently more prominent. This may be a consequence of preferential removal of high-affinity antibodies during early uses. It should not be a problem with monoclonal antibodies.

Analysis of immunoprecipitated proteins by immunoblotting can be problematic. The reagent used to detect the staining antibody on the membrane will often bind to the heavy and light chains of the precipitating antibodies. This occurs to a varying degree with [¹²⁵I]-labeled protein A, which under some circumstances can bind detectably to the heavy chain of rabbit antibodies even after they have been denatured on a membrane. It is a greater problem when second antibodies reactive with the species of antibody used for immunoprecipitation are used as detection reagents. For example, horseradish peroxidase-conjugated anti-mouse antibodies will, in general, bind strongly to both the heavy and light chains of mouse monoclonal antibodies used for immunoprecipitation. This can complicate the use of enhanced chemiluminescence (ECL) as a means for analyzing immunoblots of immunoprecipitated proteins.

Immunoprecipitates can be run on nonreducing gels to cause interfering antibody bands to migrate near the top of the gel. Unreduced IgG runs as a 150-kDa band. Nonreducing sample buffer is the same as SDS sample buffer except that 50 mM iodoacetamide replaces the reducing agents 2-mercaptoethanol (2-ME) and DTT. The mobility of many, but not all, proteins is altered profoundly under nonreducing conditions and the suitability of nonreducing gel electrophoresis must be evaluated carefully for each case. It is often difficult to analyze

reduced and nonreduced samples on the same gel because 2-ME, which is present in most sample buffers at very high concentrations, diffuses rapidly and can reduce nonreduced samples, even in the presence of 50 mM iodoacetamide.

The enhanced chemiluminescent (ECL) technique (Alternate Protocol) is very rapid and sensitive. It is unclear, however, whether it exhibits the same linearity of response as does autoradiography with [¹²⁵I]-labeled protein A. Immunoblotting should be used with caution for quantitative comparisons. For accurate quantitation, it is essential to determine the linearity of the detection system by measuring the signal obtained from multiple known dilutions of a standard sample.

Not all proteins exhibit altered gel mobility when they undergo phosphorylation or dephosphorylation. Lack of effect of phosphatase digestion on the mobility of a protein is therefore not evidence of lack of phosphorylation.

Potato acid phosphatase (Sigma or Calbiochem) is often supplied as a suspension in ammonium sulfate. It can be collected by centrifugation from ammonium sulfate suspension, dissolved at 1 mg/ml in 10 mM HEPES (pH 7.4)/0.5 mM MgCl₂/0.5 mM DTT/50% (v/v) glycerol for storage, and stored at -20°C. Because of the low pH used with this phosphatase, some proteins are denatured during incubation. For example, p60^{c-src} loses considerable activity by incubation under these conditions (Cooper and King, 1986). This will obviously complicate analysis of the effects of phosphatase on enzymatic activity.

Protein phosphatase 2A (PP2A) is generally considered not to be dependent on divalent cations for activity. Nevertheless, it often exhibits considerably more activity on some substrates in the presence of magnesium and manganese (Scheidtmann et al., 1991).

Protein tyrosine phosphatases (PTP) possess an essential cysteine in their active sites and for this reason 2-ME is included in the reaction mixture. These enzymes, however, exhibit considerable activity in buffers that lack reducing agents.

A serious concern when using phosphatases to detect protein phosphorylation is that the protein of interest will be altered or degraded by contaminants such as proteases in the phosphatase preparation. Inclusion of protease inhibitors such as Trasylol (10,000 IU/ml), leupeptin (40 µg/ml), phenylmethylsulfonyl chloride (PMSF; 400 µM), and soy bean trypsin inhibitor (40 µg/ml) will minimize the effect of

contaminating proteases. It is essential, however, also to carry out control phosphatase reactions in the presence of a specific inhibitor of the phosphatase being used. Alterations in gel mobility or enzyme activity that occur under these conditions are almost certainly nonspecific. Sodium phosphate (100 mM) inhibits most phosphatases but adds a lot of salt to the reaction mixture. Okadaic acid (20 μ M) specifically inhibits protein phosphatase 2A (PP2A); sodium vanadate (200 μ M) inhibits all tyrosine protein phosphatases.

Anticipated Results

Immunoblotting with anti-phosphotyrosine antibodies is very sensitive. Even the extremely low levels of tyrosine phosphorylated proteins present in unactivated lymphoid cells can be detected when total lysates of cells are analyzed by gel electrophoresis and immunoblotting.

Identification of phosphoproteins by digestion with phosphatase is sometimes problematical. The gel mobility of some proteins is unaffected by phosphorylation and therefore, is unaffected by dephosphorylation. Thus, only positive results are meaningful for phosphatase digestions.

Time Considerations

Immunoblotting with enhanced chemiluminescence is the most rapid technique described in this unit, because staining can be accomplished in less than a day and bound antibodies visualized within 30 min. Staining with [125 I]-labeled protein A can also be accomplished in a single day, but autoradiography usually requires 18 to 100 hr.

Phosphatase digestion can be performed in <2 hr. Analysis of the effects of dephosphorylation by gel electrophoresis or enzyme assay can be accomplished in 2 to 4 hr.

Literature Cited

- Anderson, N.G., Maller, J.L., Tonks, N.K., and Sturgill, T.W. 1990. Requirement for integration of signals from two distinct phosphorylation pathways for activation of MAP kinase. *Nature* 343:651-653.
- Cooper, J.A. and King, C.S. 1986. Dephosphorylation or antibody binding to the carboxy terminus stimulates p60^{c-src}. *Mol. Cell. Biol.* 6:4467-4477.
- Frackelton, A.R., Ross, A.H., and Eisen, H.N. 1983. Characterization and use of monoclonal antibodies for isolation of phosphotyrosyl proteins from retrovirus-transformed and growth factor-stimulated cells. *Mol. Cell. Biol.* 3:1343-1352.
- Scheidtmann, K.H., Mumby, M.C., Rundell, K., and Walter, G. 1991. Dephosphorylation of simian virus 40 large-T antigen and p53 protein by protein phosphatase 2A: Inhibition by small-t antigen. *Mol. Cell. Biol.* 11:1996-2003.

Key References

- Kamps, M.P. and Sefton, B.M. 1988. Identification of multiple novel polypeptide substrates of the v-src, v-yes, v-fps, v-ros, and v-erb-B oncogenic tyrosine protein kinases utilizing antisera against phosphotyrosine. *Oncogene* 2:305-315.

Examines many of the variables that affect immunoblotting with antibodies to phosphotyrosine.

- Morla, A. and Wang, J.Y.J. 1986. Protein tyrosine phosphorylation in the cell cycle of BALB/c 3T3 fibroblasts. *Proc. Natl. Acad. Sci. U.S.A.* 83:8191-8195.

Establishes the ability of immunoblotting with antibodies to phosphotyrosine to detect the low levels of tyrosine-phosphorylated proteins in normal cells.

Contributed by Bartholomew M. Sefton
The Salk Institute
San Diego, California

Detection of Phosphorylation by Enzymatic Techniques

Reversible protein phosphorylation has been increasingly recognized as an important mechanism for regulating physiological processes in both plant and animal cells. There are a number of techniques to demonstrate the presence of covalently bound phosphate in proteins. Metabolic labeling of cells with ^{32}P and subsequent isolation of the ^{32}P -labeled protein is the most direct method for demonstrating protein phosphorylation. This approach is difficult when the specific polypeptides involved in the physiological phosphorylation event have not been well characterized or when these proteins are present in such low abundance in the cell that significant incorporation of [^{32}P]phosphate cannot be achieved. Metabolic labeling of proteins (UNIT 18.2) may also be compromised by a slow turnover of the protein-bound phosphate in the intact cell. As the extent of phosphorylation of given protein reflects the relative activities of the protein kinases and phosphatases that act on it, knowledge of physiological stimuli that modulate the protein function might be needed to establish the experimental conditions which increase the metabolic labeling of the substrate protein. In the absence of this information, it may be difficult to optimize specific protein phosphorylation in the intact cell. This situation can be further exacerbated by the lack of experimental tools, such as monospecific antibodies or affinity probes, that permit the rapid isolation of the substrate protein and preserve its phosphorylation.

An approach that eliminates the need for metabolic labeling of cells can sometimes be used to monitor *in vivo* protein phosphorylation. This procedure, known as “back” phosphorylation, requires that the protein of interest be at least partially separated from other cellular phosphoproteins under conditions that preserve covalently bound phosphates. The protein is then phosphorylated *in vitro* using radiolabeled ATP. By comparing the extent of radiolabel incorporated into the isolated, phosphorylated protein with that incorporated into the same protein after it has been treated with phosphatases to strip the bound phosphate, it is possible to estimate the phosphorylation state of that protein in the intact cell. This technique requires the identification and availability of the cellular kinases that phosphorylate the protein of interest. Because many phosphoproteins are phosphorylated at multiple sites by several different kinases, back phosphorylation usually monitors only a subset of the endogenous phosphorylation events.

If significant ^{32}P -labeling of the desired protein cannot be easily demonstrated, yet another approach may be taken. Recent studies have utilized one or more phosphatases to reverse existing covalent modifications in proteins, resulting in changes in the function or activity of the substrate. *In vitro* dephosphorylation by one or more commercially available phosphatases has implicated phosphorylation as a regulatory mechanism in many physiological processes. These studies have been undertaken using cell extracts and subcellular fractions, as well as partially purified proteins, and do not require prior knowledge of the phosphorylation event. The methodologies discussed in this unit can be used to detect protein phosphorylation where actual incorporation of [^{32}P]phosphate into a specific polypeptide cannot be demonstrated.

The overall strategy is to first examine the functional effects elicited by nonspecific acid or alkaline phosphatases that dephosphorylate many phosphoproteins *in vitro* (Basic Protocol 1 and Alternate Protocol 1). Nonspecific phosphatases isolated from plant and animal tissues show significant activity against a wide variety of phosphoester-containing compounds, including many phosphoproteins. These enzymes reverse all phosphorylations whether they occur as phosphotyrosine, phosphothreonine, phosphoserine, or phosphohistidine so they can define the regulatory consequence of protein dephosphorylation

without necessarily identifying the specific phosphoamino acid. Protein phosphatases that selectively hydrolyze phosphoserine and phosphothreonine or phosphotyrosine residues can then be used to identify the functionally important covalent modification. Additional protocols describe digestion of phosphoproteins with a protein serine/threonine phosphatase (Basic Protocol 2) and protein tyrosine phosphatase (Alternate Protocol 2). A support protocol has been included to identify the radiolabel as $^{32}\text{P}_i$ based on its ability to form a complex with ammonium molybdate.

Protein dephosphorylation can be analyzed as the loss of [^{32}P]phosphate from metabolically labeled proteins following their separation by SDS-PAGE. Some phosphoproteins also show changes in their electrophoretic migration following the loss of covalently bound phosphate. This can provide an additional monitor of protein dephosphorylation that can be used even in the absence of metabolic labeling. Finally, measurements of protein function (e.g., opening or closing of ion channels, binding of nuclear factors to specific DNA elements, or altered enzyme activity) can define the consequences of dephosphorylation even when little information is available about the specific phosphopeptides involved.

BASIC PROTOCOL 1

DIGESTION OF PHOSPHOPROTEINS WITH NONSPECIFIC ACID PHOSPHATASES

Potato acid phosphatase shows a broad substrate specificity and hydrolyzes many metabolites containing phosphoester bonds. This enzyme shows comparatively low activity against phosphoproteins. Moreover, metabolites present in cell extracts or some subcellular fractions may inhibit its ability to dephosphorylate protein substrates. In this protocol, a sample containing a phosphoprotein is digested with potato acid phosphatase and analyzed for evidence of structural and functional changes.

Materials

Sample containing 100 to 200 μg total protein
50 mM piperazine-*N,N'*-bis(2-hydroxypropanesulfonic acid) (PIPES), pH 6.0
Sephadex G-25 column (optional)
PIPES/2-ME or PIPES/DTT buffer, pH 6.0: 50 mM PIPES containing 15 mM
2-mercaptoethanol or 1 mM dithiothreitol (prepare fresh)
Potato acid phosphatase
2 \times SDS-PAGE sample buffer: 50 mM Tris·Cl (pH 7.5; APPENDIX 2)/0.4 M glycine
(pH 8.3)/0.2% (w/v) SDS
100 mM sodium pyrophosphate or other general phosphatase inhibitor
90°C water bath or heating block
Additional reagents and equipment for electrophoresis (UNIT 10.2)

1. Prepare a protein sample containing 100 to 200 μg total protein. Remove inhibitory metabolites or particulate material.

If the substrate is particulate, soluble contaminants can be removed by washing 2 to 3 times with 50 mM PIPES, pH 6.0. Inhibitory metabolites can be removed by dialyzing against the same buffer, or if the substrate is a soluble protein, it can be desalted on a Sephadex G-25 column.

2. Incubate the sample in 100 μl PIPES/2-ME or PIPES/DTT, pH 6.0, for 10 min at 30°C.

The pH optimum for potato acid phosphatase hydrolysis is between pH 4.0 and 5.0. However, many phosphoprotein substrates may be irreversibly inactivated at low pH, so incubation at pH 6.0 is recommended.

3. Add 5 to 10 U potato acid phosphatase and incubate 15 min at 30°C.

If it is not possible to incorporate [³²P]phosphate into the substrate protein by metabolic labeling, control reactions should include a general phosphatase inhibitor—e.g., sodium vanadate, sodium pyrophosphate, sodium glycerophosphate, sodium orthophosphate, or p-nitrophenyl phosphate (PNPP) at a final concentration ≥10 mM—to distinguish the effects produced by protein dephosphorylation by potato acid phosphatase from those produced by contaminating proteases.

One unit of acid phosphatase hydrolyzes 1 nmol PNPP in 1 min at 30°C and pH 4.0.

4. Remove 10 µl of the sample and add an equal volume of 2× SDS-PAGE sample buffer. Heat 5 min at 90°C.
5. Electrophorese control and dephosphorylated material on an SDS polyacrylamide gel (UNIT 10.2). For a radiolabeled sample, analyze the loss of radiolabel in the substrate protein by autoradiography or phosphoimage analysis; for unlabeled sample, analyze the change in substrate mobility by immunoblot analysis.
6. Add 10 µl of 100 mM sodium pyrophosphate (10 mM final concentration) to the remaining dephosphorylated material to terminate the dephosphorylation reaction. Use an appropriate assay of protein function to characterize the changes that accompany substrate dephosphorylation.

DIGESTION OF PHOSPHOPROTEINS WITH NONSPECIFIC ALKALINE PHOSPHATASE

ALTERNATE PROTOCOL 1

Calf intestine alkaline phosphatase has been widely utilized to dephosphorylate both proteins and nucleic acids. With the widespread use of this enzyme in molecular biology most commercial preparations are essentially free of contaminating proteases. Alkaline phosphatase has a broad pH profile with an optimum between pH 8.0 and 8.5. In contrast to acid phosphatases, this enzyme can be utilized for in vitro dephosphorylation of proteins under conditions that do not denature the substrate proteins. Calf intestine alkaline phosphatase effectively dephosphorylates proteins containing phosphoserine and phosphothreonine residues, which together account for >97% of the protein-bound phosphate in eukaryotic cells. Under defined conditions, preferential dephosphorylation of phosphotyrosines by alkaline phosphatases has also been reported (Swarup et al., 1981). Alkaline phosphatase activity is stimulated by Mg²⁺ ions, although significant activity is seen even in the absence of divalent cations.

Additional Materials (also see Basic Protocol 1)

Tris/MgCl₂, pH 7.5 or HEPES/MgCl₂, pH 7.5 buffer: 50 mM Tris·Cl (pH 7.5; APPENDIX 2)/1 mM MgCl₂ or 50 mM N,-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES; pH 7.5)/1 mM MgCl₂

Calf intestine alkaline phosphatase (molecular biology grade)

1. Incubate the sample (containing 100 to 200 µg total protein in 100 µl Tris/MgCl₂ or HEPES/MgCl₂) for 10 min at 30°C.
2. Add 20 to 30 U calf intestine alkaline phosphatase and incubate 15 min at 30°C.

One unit of calf intestine alkaline phosphatase hydrolyzes 1 nmol p-nitrophenyl phosphate (PNPP) in 1 min at 30°C and pH 8.5.

3. Terminate the dephosphorylation reaction by adding an equal volume of 2× SDS-PAGE sample buffer or a general phosphatase inhibitor. Analyze changes in substrate mobility and function (see Basic Protocol 1).

Analysis of Protein Phosphorylation

18.5.3

DIGESTION OF PHOSPHOPROTEINS WITH PROTEIN SERINE/THREONINE PHOSPHATASES

Protein phosphatase 1, 2A, and 2B (PP1, PP2A, and PP2B, respectively) represent three major classes of protein serine/threonine phosphatases found in eukaryotic cells. These enzymes are currently available from several suppliers; they will retain full enzymatic activity for many months when stored below -20°C (see supplier's recommendations). This protocol describes the use of PP2A, a broad-specificity phosphatase, to dephosphorylate the phosphoproteins.

Materials

Sample containing 100 μg total protein
Tris/DTT/ MnCl_2 buffer, pH 7.5: 50 mM Tris-Cl (pH 7.5; *APPENDIX 2*)/
1 mM dithiothreitol (DTT)/1 mM MnCl_2 (prepare fresh)
Microcystin-LR
Protein phosphatase 2A (PP2A), catalytic subunit

1. Prepare the sample (containing 100 μg total protein) in 100 μl Tris/DTT/ MnCl_2 buffer, pH 7.5. Prepare a control reaction in the same buffer with 1 μM microcystin-LR. Incubate 10 min at 37°C .

Microcystin-LR is a potent inhibitor of PP1 and PP2A.

2. Add 0.2 to 0.5 U PP2A and incubate 10 to 30 min at 37°C .

One unit of PP2A or PP1 activity hydrolyzes 1 nmol phosphorylase A in 1 min at 30°C and pH 7.5.

3. Add microcystin-LR to 1 μM final concentration to terminate the dephosphorylation reaction.

Unlike the general phosphatases discussed above, protein phosphatases show some selectivity for the dephosphorylation of different sites in substrate phosphoproteins and can be used to gain insight into the functional role of individual protein-bound phosphates.

4. Analyze the dephosphorylated material by SDS-PAGE or functional assays (see Basic Protocol 1).

DIGESTION OF PHOSPHOPROTEINS WITH PROTEIN TYROSINE PHOSPHATASES

Transmembrane tyrosine phosphatases such as CD45 possess low in vitro activity against many substrates, leading to the suggestion that association with extracellular ligands activates these receptor phosphatases. In this regard, protein tyrosine phosphatase 1B (PTP-1B) and src homology domain-containing protein tyrosine phosphatases (SH-PTP1 or SH-PTP2) are highly active phosphotyrosine phosphatases that demonstrate no measurable activity against phosphoserine and phosphothreonine and may be the preferred reagents to investigate the regulatory role of tyrosine phosphorylation.

Materials

Sample containing 10 to 100 μg total protein
50 mM imidazole, pH 7.5
Protein tyrosine phosphatase (e.g., PTP-1B or SH-PTP)
2 \times SDS-PAGE sample buffer (see Basic Protocol 1) *or* 100 mM sodium vanadate

1. Incubate the sample (containing 10 to 100 μg total protein) in 50 mM imidazole, pH 7.5, for 10 min at 37°C.
2. Add 1 to 5 U protein tyrosine phosphatase and incubate 15 to 30 min at 37°C.
One unit of phosphatase activity hydrolyzes 1 nmol p-nitrophenyl phosphate (PNPP) in 1 min at 37°C and pH 7.0.
3. Terminate the dephosphorylation reaction by adding an equal volume of 2 \times SDS-PAGE sample buffer *or* 100 mM sodium vanadate to a final concentration of 0.1 mM.
Recombinant PTP-1B and SH-PTP1 are also available as GST fusion proteins that can be bound to glutathione-agarose and conveniently removed by centrifugation or filtration following dephosphorylation of the substrate.
4. Analyze digested material by gel electrophoresis, immunoblotting, or a functional assay (see Basic Protocol 1, steps 5 and 6).

MEASUREMENT AND IDENTIFICATION OF RELEASED ^{32}P

Time-dependent dephosphorylation of radiolabeled proteins is usually monitored by precipitating the protein substrate with 10% to 15% (w/v) trichloroacetic acid (TCA) at 4°C. Following removal of the TCA-denatured phosphoprotein by centrifugation, radioactivity that accumulates in the TCA-soluble fraction provides a monitor of [^{32}P]phosphate release. However, this procedure cannot distinguish between the release of orthophosphate and TCA-soluble phosphopeptides. To define a dephosphorylation event, the TCA-soluble radioactivity must be characterized as $^{32}\text{P}_i$. This procedure takes advantage of the unique ability of $^{32}\text{P}_i$ orthophosphate to form a complex with ammonium molybdate. The [^{32}P]phosphomolybdate complex is soluble in organic solvent and can be readily separated from ^{32}P -labeled phosphopeptides. Therefore, this assay specifically monitors protein dephosphorylation (Shenolikar and Ingebritsen, 1984).

Materials

Trichloroacetic acid
 Radiolabeled protein/phosphatase reaction mixture (see Basic Protocol 1)
 1.25 mM potassium phosphate (KH_2PO_4)/1 N H_2SO_4
 1:1 (v/v) isobutanol/toluene
 5% (w/v) ammonium molybdate
 Scintillation fluid
 Liquid scintillation counter

1. Add trichloroacetic acid (TCA) to dephosphorylated radiolabeled protein to give a final concentration of 10% to 15% (w/v). Incubate 10 min at 4°C.
2. Microcentrifuge 5 min at maximum speed. Remove 0.1 ml of the TCA supernatant and place in a microcentrifuge tube. Add 0.2 ml of 1.25 mM KH_2PO_4 /1 N H_2SO_4 and mix.
3. Add 0.5 ml of 1:1 isobutanol/toluene to the microcentrifuge tube and mix thoroughly.
4. Add 0.1 ml of 5% ammonium molybdate and mix.
5. Microcentrifuge 5 min at maximum speed to separate aqueous and organic phases.
6. Remove 0.3 ml of the upper (organic phase), mix with 1 ml scintillation fluid, and count in a liquid scintillation counter.

SUPPORT PROTOCOL

Analysis of Protein Phosphorylation

18.5.5

COMMENTARY

Background Information

A key feature of protein phosphorylation (as compared to other protein modifications, such as proteolysis) is its reversibility. The ability to phosphorylate a substrate *in vitro* requires that the kinase responsible for phosphorylation be known. However, many cellular protein kinases are themselves activated via protein phosphorylation, making it difficult to use them *in vitro* to phosphorylate the substrate. Establishing the physiological relevance of many *in vitro* phosphorylations is also made difficult because these studies often do not take account of the physiological concentrations of the kinase and its substrate. *In vitro* experiments also tend to ignore the importance of the subcellular localization of the substrate and the kinase, sometimes leading to modifications that do not occur *in vivo*. For these reasons, knowledge of the physiological stimuli that regulate the substrate's function can be used to promote substrate phosphorylation as well as to provide important clues in the identification of protein kinases that phosphorylate the substrate.

Most cellular phosphoproteins are subject to covalent modification at multiple sites by more than one protein kinase. Therefore, *in vitro* studies with purified protein kinases often fail to recreate the spectrum of phosphorylations seen in the intact cell, and this also makes it difficult to define the role of individual covalent modifications. Thus, a new strategy of analyzing the role of multisite phosphorylation has been developed using proteins that have been modified in the intact cell. However, it is particularly important to isolate such substrates in the presence of phosphatase inhibitors and to preserve the protein-bound phosphates during lengthy purification procedures.

Biochemical studies of protein phosphatases have established the *in vitro* substrate specificity of these enzymes, which distinguish between different phosphoamino acids (phosphoserine/phosphothreonine versus phosphotyrosine) and different phosphoprotein substrates. These enzymes can also show preferential dephosphorylation of specific sites (Shenolikar and Nairn, 1991; Charbonneau and Tonks, 1992; Shenolikar, 1994). *In vitro* dephosphorylation by selected phosphatases may not only define the functional role of selected covalent modifications but also identify "silent" or "structural" phosphorylations that do not directly regulate the protein function. These phosphorylations may play an in-

direct role in modulating protein phosphorylation or dephosphorylation of other sites.

This unit has focused on commercially available phosphatases that have been used to characterize the functional importance of protein phosphorylation. For example, several growth-associated kinases, including the mitogen-activated protein kinases (MAP kinases) and the cyclin-dependent kinases such as *cdc2*, are phosphorylated on adjacent threonine and tyrosine residues. Phosphorylation of these enzymes is controlled by kinases and phosphatases that are themselves highly regulated in the intact cell. *In vitro* dephosphorylation by serine/threonine-specific or tyrosine-specific phosphatases both established the functional importance for phosphorylations even prior to the identification of the kinases and phosphatases that regulate these proteins.

The most direct method for detecting protein phosphorylation is clearly metabolic labeling of cells with $^{32}\text{P}_i$ (see *UNIT 18.2*) and subsequent isolation of the ^{32}P -labeled substrate protein. However, it is often difficult to develop a rapid purification procedure for the specific protein of interest without selective reagents such as monospecific antibodies or other affinity ligands. Significant labeling of proteins is also a problem given the low abundance of many regulatory proteins and the slow turnover of their protein-bound phosphate under "basal" conditions. In cases where [^{32}P]phosphate can be readily incorporated into the protein, *in vitro* dephosphorylation by selected phosphatases can establish the functional importance of the modification. Ability to correlate loss of [^{32}P]phosphate from a specific polypeptide with a functional change provides the best evidence for the physiological importance of protein phosphorylation.

Initial experiments should use a general phosphatase such as potato acid phosphatase or calf intestine alkaline phosphatase which can successfully dephosphorylate a wide variety of phosphoproteins modified on several different amino acids. It should be noted that acid phosphatases from animal tissues (e.g., bone, prostate, liver, and heart) represent a class of low-molecular-weight (18,000- to 25,000-kDa) protein tyrosine phosphatases that selectively dephosphorylate phosphotyrosine residues. These enzymes were defined as acid phosphatases because they hydrolyzed *p*-nitrophenyl phosphate (PNPP) at a pH optimum between 4.0 and 5.0. However, at physiological pH,

these enzymes appear to specifically dephosphorylate proteins and peptides containing phosphotyrosine residues. The mammalian acid phosphatases show no significant activity against phosphoserine/phosphothreonine-containing proteins.

With phosphoprotein substrates, potato acid phosphatase shows a preference for dephosphorylating phosphotyrosine residues, followed by phosphothreonine residues. By comparison, this enzyme shows very low activity against phosphoserine-containing substrates. Wheat germ acid phosphatase can dephosphorylate phosphoserines but still shows significantly higher activity against phosphotyrosine and phosphothreonine (Van Etten and Waymack, 1991). In this regard, alkaline phosphatase is often preferred over potato or wheat germ acid phosphatase for dephosphorylation of phosphoserine, the predominant protein modification in eukaryotic cells.

When a general phosphatase, such as alkaline phosphatase, is used, effects independent of protein dephosphorylation must be considered. For instance, alkaline phosphatase hydrolyses ATP and inhibits ATP-dependent processes, such as protein kinases and ATP-dependent ion channels that are also inhibited by protein dephosphorylation (Berger et al., 1993). To distinguish between these differing effects, use of immobilized alkaline phosphatase is recommended. Agarose or dextran beads covalently linked to alkaline phosphatase should be washed extensively with the assay buffer to remove storage solutions and any unbound alkaline phosphatase. Following incubation with the substrate protein, immobilized alkaline phosphatase can be removed by centrifugation or filtration. Functional analysis of the substrate can then be undertaken without concern for the potential effects of the phosphatase on such analysis. However, if the substrate is present in a particulate fraction (e.g., in subcellular organelles), it cannot be easily separated from the beads by centrifugation or filtration. In that case, the phosphatase should be inhibited with a general inhibitor such as sodium pyrophosphate. In contrast to the general phosphatases, protein phosphatases are inefficient at hydrolyzing ATP and do not hinder analysis of ATP-dependent processes.

Once phosphorylation of a substrate is established using a general phosphatase, the next step is to use protein phosphatases that have much higher activities against phosphoprotein substrates than either acid or alkaline phosphatase and are highly selective for specific

phosphoamino acids. Protein phosphatases 1, 2A, and 2B (PP1, PP2A, and PP2B) represent three major classes of protein serine/threonine phosphatases in eukaryotic cells. These enzymes differ in their subunit structure, substrate specificity, and regulation by divalent cations. PP1 and PP2A account for >90% of the total protein serine/threonine phosphatase activity in many cell extracts (Shenolikar and Nairn, 1991) and demonstrate a broad *in vitro* substrate specificity. Therefore, PP1 and PP2A are the most commonly used reagents for *in vitro* dephosphorylation of phosphoserine- and phosphothreonine-containing proteins. PP1 and PP2A are inhibited by microcystin-LR as well as several other recently discovered inhibitors, including okadaic acid, calyculin A, and tautomycin, albeit with differing IC_{50} values.

Inhibitors, such as okadaic acid (with IC_{50} for PP2A of 0.1 to 1.0 nM and for PP1 of 10 to 100 nM), calyculin A (with an IC_{50} for PP1 of 1 nM and for PP2A of 10 nM), and microcystin-LR (with an IC_{50} for both PP1 and PP2A of 0.1 nM) are used to exclude nonspecific effects of PP1 and PP2A. These inhibitors represent important experimental tools for establishing the physiological importance of reversible protein phosphorylation.

PP2B, also called calcineurin, is a Ca^{2+} /calmodulin-dependent phosphatase and demonstrates a narrow *in vitro* substrate specificity compared to PP1 or PP2A. The enzyme purified from mammalian tissues consists of a catalytic subunit that is tightly associated with a calcium-binding regulatory subunit. PP2B has no activity in the absence of 1 μ M calmodulin and 1 mM $CaCl_2$ and can be activated by either Ca^{2+} /calmodulin or 1 mM $MnCl_2$. Many highly purified PP2B preparations show no activity in the presence of Ca^{2+} /calmodulin but are fully activated by Mn^{2+} . The immunosuppressive drugs cyclosporin A and FK506 selectively inhibit PP2B *in vivo* and *in vitro*.

PP2B in the presence of Ni^{2+} ions and PP2A in association with an endogenous regulator protein or viral antigens can dephosphorylate proteins and peptides containing phosphotyrosines. However, under the conditions described in Basic Protocol 2, PP1, PP2A, and PP2B are highly selective for dephosphorylation of phosphoserine and phosphothreonine residues. PP1 and PP2A represent the major phosphohistidine phosphatase activity measured in tissue extracts, but the regulatory importance of this modification has not yet been established.

Critical Parameters

Some commercial preparations of potato acid phosphatase contain contaminating protease activity. Because phosphorylation sites reside near the surface of the substrate protein, they are often accessible to proteases, so functional changes that result from limited proteolysis can be incorrectly attributed to dephosphorylation. To limit the effects of proteases, high enzyme concentrations or prolonged incubations with the general phosphatases should be avoided. Aside from including protease inhibitors in the reaction, two strategies should be considered to distinguish proteolysis from dephosphorylation. Formation of a complex between trichloroacetic acid-soluble [³²P]phosphate and ammonium molybdate (Support Protocol) is particularly useful in distinguishing [³²P]phosphate from ³²P-labeled phosphopeptides. The use of phosphatase inhibitors in control reactions also identifies those effects that can be attributed to reactions in which there are contaminating proteolytic enzymes.

Commercial preparations of PP2A and PP1 contain only the free catalytic subunits. In vivo these subunits associate with regulatory subunits, which modulate their substrate specificity. However, the catalytic subunits are the most readily purified to homogeneity and can be stored for prolonged periods without significant changes in their enzymatic properties. Both PP1 and PP2A catalytic subunits retain high activity in the absence of divalent cations; PP2A activity may be further stimulated by 1 mM Mn²⁺ in the enzyme reaction. Long-term storage of PP1 and PP2A results in an apparent loss of enzyme activity. However, these inactive enzymes can be fully reactivated by including 1 mM MnCl₂ in the assay. The Mn²⁺-dependent PP1 enzyme shows slightly altered substrate specificity when compared with PP1 purified from tissues. Moreover, PP1 is potentially inhibited by several endogenous inhibitors. Thus, PP2A may be the preferred reagent for in vitro dephosphorylation of substrates in many cell extracts. Recent studies show that some tissues may contain endogenous PP2A inhibitors (Li et al., 1995). These inhibitory proteins could hinder the ability of PP2A to dephosphorylate proteins in some tissue preparations.

Troubleshooting

The absence of a functional change following incubation of the substrate with one or more phosphatases may indicate either that the protein was not phosphorylated or that the dephosphorylation of existing phosphates has no effect

on protein function. It is also possible that components in the reaction mixture inhibited the phosphatases. The latter possibility can be readily established if the substrate protein has been metabolically labeled with [³²P]phosphate. SDS-PAGE followed by autoradiography or phosphoimage analysis can directly monitor the time-dependent loss of radiolabel from substrate protein. Coincidence between the removal of radiolabeled phosphate and change of function of the substrate protein can be used to distinguish between important regulatory modifications and “silent” phosphorylations. In analyses of unlabeled substrates, the reaction should include an internal standard or a known ³²P-labeled protein or peptide. Dephosphorylation of the radiolabeled marker in the presence or absence of the sample should establish the extent to which activity of the added phosphatase has been compromised by components in the sample. If several different phosphatases fail to elicit a change in protein function under conditions that dephosphorylate the marker, the substrate protein may contain no protein-bound phosphate or phosphorylation may not be functionally important for the protein.

Anticipated Results

Most phosphoproteins should be dephosphorylated by the general phosphatases, leading to complete removal of protein-bound phosphate and functional changes indicating that reversible phosphorylation controls the protein function. In a number of cases, metabolic labeling and phosphoamino acid analysis or the use of specific protein (phosphoserine/phosphothreonine or phosphotyrosine) phosphatases has successfully identified the amino acids modified. Such studies have then located the phosphorylated residues, especially in proteins whose primary structures have been determined by amino acid sequencing or cDNA cloning. A combination of biochemical studies and site-directed mutagenesis has then been used to establish the functional role of phosphorylation at individual sites. Increasing evidence points to the presence of both positive and negative regulatory phosphorylations in proteins. Thus, it is difficult to predict the functional outcome of complete in vitro dephosphorylation of many substrates.

Time Considerations

Approximately 1 hr is required to complete the dephosphorylation reaction. SDS-PAGE

and analysis of the gel require 1.5 hr. The time required for functional analysis depends on the particular assay used.

Literature Cited

- Berger, H.A., Travis, S.M., and Welsh, M.J. 1993. Regulation of the cystic fibrosis transmembrane conductance regulator Cl^- channel by specific protein kinases and protein phosphatases. *J. Biol. Chem.* 268:2037-2047.
- Charbonneau, H. and Tonks, N.K. 1992. 1002 protein phosphatases? *Annu. Rev. Cell Biol.* 8:463-493.
- Li, M., Guo, H., and Damuni, Z. 1995. Purification and characterization of two potent heat-stable inhibitors of protein phosphatase 2A from bovine kidney. *Biochemistry* 34:1988-1996.
- Shenolikar, S. 1994. Protein serine/threonine phosphatases: New avenues for cell regulation. *Annu. Rev. Cell Biol.* 10:55-86.
- Shenolikar, S. and Ingebritsen, T.S. 1984. Protein (serine and threonine) phosphate phosphatases. *Methods Enzymol.* 107:102-129.
- Shenolikar, S. and Nairn, A.C. 1991. Protein phosphatases: Recent progress. *Adv. Second Messenger Phosphoprotein Res.* 23:1-123.
- Swarup, G., Cohen, S., and Garbers, D.L. 1981. Selective dephosphorylation of proteins containing phosphotyrosine by alkaline phosphatases. *J. Biol. Chem.* 256:1447-1452.
- Van Etten, R.L. and Waymack, P.P. 1991. Substrate specificity and pH dependence of homogeneous wheat germ acid phosphatase. *Arch. Biochem. Biophys.* 288:634-645.

Contributed by Shirish Shenolikar
Duke University Medical Center
Durham, North Carolina

Production of Antibodies That Recognize Specific Tyrosine-Phosphorylated Peptides

It is possible to produce anti-phosphopeptide antibodies (i.e., antibodies recognizing phosphorylated peptides) that recognize a protein only in its phosphorylated state, and that do not cross-react with either the cognate unphosphorylated protein or other phosphoproteins. Because the state of phosphorylation is often indicative of a protein's functional state or activity, such antibodies provide a convenient probe of the functional state of a protein. Thus, unlike conventional antibodies, anti-phosphopeptide antibodies provide information regarding not only the abundance of a protein but also its activity. Unlike general anti-phosphoamino acid (e.g., anti-phosphotyrosine) antibodies, which have broad reactivity, anti-phosphopeptide antibodies may have unique specificity toward the cognate proteins. Such reagents not only facilitate conventional *in vitro* analysis of phosphoproteins, but also allow heretofore impossible applications—e.g., differential isolation of species of a particular protein that have been phosphorylated at individual phosphorylation sites, as well as analysis of the functional state of a protein *in situ* by immunohistochemical techniques.

Basic Protocol 1 describes the production of polyclonal anti-phosphopeptide antibodies and Basic Protocol 2 describes the production of monoclonal anti-phosphopeptide antibodies. Both of these procedures are based upon immunizing animals with oligopeptides that have been synthesized containing phosphorylated tyrosine; hence, a knowledge of the sequence and phosphorylation site of the protein of interest is a prerequisite for either protocol. Such an immunization will generate an immune response with at least four components: (1) anti-carrier protein reactivity, (2) general anti-phosphotyrosine reactivity, (3) phosphorylation-independent anti-peptide reactivity, and (4) phosphorylation-dependent anti-peptide reactivity. For the production of polyclonal antibodies (see Basic Protocol 1), a multiple-step affinity chromatographic purification with several negative-selection steps is carried out to produce a final antibody preparation having the desired reactivity. For production of monoclonal antibodies (see Basic Protocol 2), ELISAs are performed to screen candidate hybridoma supernatants against the cognate phosphopeptide, as well as against related phosphorylated and nonphosphorylated peptides, until one is found with the desired reactivity and specificity. Although monoclonal antibodies have a number of advantages, production of polyclonal antibodies is likely to be more predictable, as in the authors' experience a monoclonal hybridoma clone of stringent specificity may occur with very low frequency (see Critical Parameters). Production of monoclonal antibodies is also generally more time-consuming and expensive. The relative merits of monoclonal and polyclonal antibodies are discussed in the Chapter 11 introduction and in general antibody guides (see Key References). Support Protocols are provided for the coupling of peptides (Support Protocol 2) and phosphotyrosine (Support Protocol 3) to the affinity matrix (Affi-Gel 10); BSA-agarose affinity matrix is commercially available. General methods for the coupling of peptides to carrier proteins for use in immunization are described in *UNIT 11.16*.

This unit describes production of antibodies against tyrosine-phosphorylated peptides, with which the authors have the most expertise, but the principles discussed here also apply to peptides phosphorylated on serine and threonine (see Key References).

Strategic Planning

The first step of the planning phase is design and synthesis of the required peptides (see Support Protocol 1 and Background Information). The authors employ a peptide containing fifteen amino acids with the phosphorylation site Tyr in the center (residue 8). This peptide consists of fourteen amino acids from the actual sequence of the protein plus an N-terminal Lys to enhance coupling of the peptide to a carrier protein (BSA is used as a carrier in this protocol). The peptide, coupled to BSA, is then used to immunize rabbits. The cognate phosphopeptide is also immobilized on an affinity-chromatography column support (see Support Protocol 2); this affinity matrix will ultimately be used to purify the antibody from the rabbit serum. In addition, a series of negative-selection affinity-chromatography columns are used to adsorb antibodies from the serum that cross-react with epitopes other than the cognate phosphopeptide. These columns are composed of matrix coupled to various synthetic peptides—principally the cognate nonphosphorylated peptide (to remove phosphorylation-independent antibodies) and phosphotyrosine itself (to remove indiscriminant anti-phosphotyrosine reactivity; see Support Protocol 3 for preparation of phosphotyrosine affinity columns). Additional negative-selection steps may be used if needed, depending upon the particular target chosen, to remove antibodies that cross-react with phosphorylation sites of other proteins having known homology to the target site. This cross-reactivity occurs particularly when the target protein is a member of a family of homologous proteins—e.g., a subgroup of related receptor tyrosine protein kinases. If such cross-reactivity is anticipated, it is possible to synthesize not only the cognate phosphorylated and nonphosphorylated peptides, but also phosphopeptides based upon the related homologous sequences, to be used for negative selection. The authors also pass the serum over a column consisting of immobilized carrier protein (commercially available BSA-agarose) to remove the majority of the antibodies generated against the much larger carrier protein. Thus, a typical purification scheme might consist of the following:

Negative-selection affinity-purification columns

- BSA (carrier)
- phosphotyrosine
- cognate nonphosphopeptide
- homologous phosphopeptides (if desired)

Positive-selection affinity-purification column

- cognate phosphopeptide.

The second step in planning is to raise and test antisera in rabbits. To prepare the phosphopeptide-carrier conjugate for immunization of rabbits, the authors couple a 30-fold molar excess of phosphopeptide (14.4 mg dissolved in water and neutralized with 1 N NaOH) to BSA (20 mg dissolved in 0.4 M sodium phosphate buffer, pH 7.5), using glutaraldehyde as the cross-linking reagent as described in Doolittle (1986). The reaction is allowed to proceed 30 min at room temperature; the development of a yellow color in the coupling solution is an indication that the glutaraldehyde has reacted. General methods for coupling of peptides to carrier proteins for use in immunization are described in *UNIT 11.16*. Conjugation can be confirmed by performing an anti-phosphotyrosine immunoblot (*UNIT 18.4*) on a small test aliquot of the conjugate (although a smear should be expected on the blot, resulting from variation in the amount of peptide coupled per BSA molecule as well as from possible multimers of BSA). Guidelines for immunization of rabbits are provided in *UNIT 11.12*, and many institutions have core facilities for the immunization of

rabbits and production of antisera. The authors immunize with 1 mg of immunogen (the peptide-BSA conjugate prepared above, dissolved in PBS) in 1 ml of a 50% emulsion with complete Freund's adjuvant. This is administered subcutaneously at multiple sites, followed by four boosts at 2-week intervals using same quantity of immunogen in incomplete Freund's adjuvant. An adequate immune response can be seen as early as week 6.

The specificity of the final antibody is demonstrated most simply by immunoblot analysis (*UNIT 10.8*) using a panel of relevant phosphorylated and nonphosphorylated proteins. Electrophoresis is carried out with purified cognate protein in both its phosphorylated and unphosphorylated states, as well as with any proteins toward which there could conceivably be cross-reactivity. These are most easily isolated by immunoprecipitation using conventional antibodies. The blot is probed with the purified anti-phosphopeptide antibody to demonstrate appropriate reactivity. This immunoblot assay may also be used after each chromatographic step to monitor the success of the purification.

An assay also must be devised for estimating the titer of the immune response in the sera of the immunized rabbits. Convenient screening assays for this purpose include tests for the ability of crude sera to immunoprecipitate the target phosphoprotein from a cell lysate known to contain it, as well as immunoblotting of purified phosphorylated target protein with the crude serum.

The relationship among the above procedures is illustrated by the flow chart in Figure 18.6.1.

Purification Method

The purification of polyclonal antibodies consists of multiple affinity-chromatography steps for the negative and positive selection of antibodies of the appropriate reactivity (see Fig. 18.6.1). All of the following chromatographic steps may be carried out at room temperature, with the solutions applied to the columns by gravity.

Materials

- BSA-agarose affinity matrix (Sigma) packed (as in Support Protocol 2) in a 10-ml bed volume column
- Phosphotyrosine affinity matrix column (10-ml bed volume; see Support Protocol 3)
- Crude serum from rabbit immunized with phosphopeptide-BSA conjugate (refer to Strategic Planning, above)
- PBS/azide: PBS (*APPENDIX 2*) containing 0.02% (w/v) sodium azide (store indefinitely at 4°C or room temperature)
- Cognate nonphosphopeptide affinity matrix column (3-ml bed volume; see Support Protocols 1 and 2 and Strategic Planning)
- 3 M NaSCN
- Homologous phosphopeptide affinity matrix columns (optional; 3-ml bed volume; see Support Protocols 1 and 2 and Strategic Planning)
- Positive-selection phosphopeptide affinity matrix column (3-ml bed volume; see Support Protocols 1 and 2 and Strategic Planning)
- 3.5 M and 4.5 M MgCl₂ (optional; see Critical Parameters)
- Spectrophotometer (optional)
- Dialysis tubing (MWCO 12,000 to 14,000; 10 mm width, 6.4 mm diameter; e.g., Spectra/Por 4 from Spectrum)
- Additional reagents and equipment for dialysis (*APPENDIX 3C*), protein quantitation (*UNIT 10.1*), and analysis of antibodies by ELISA (*UNIT 11.2*) or immunoblotting (*UNIT 10.8*)

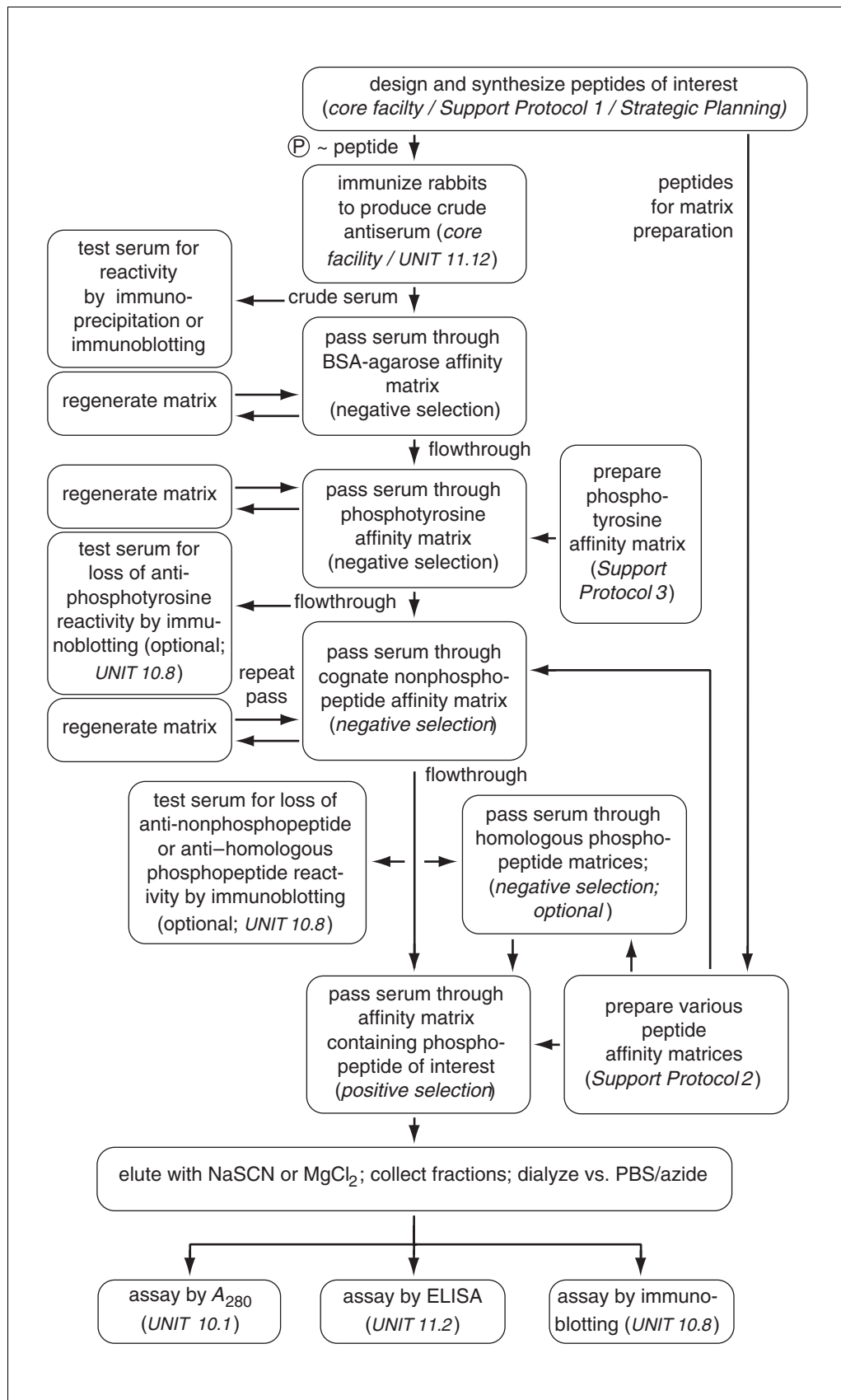


Figure 18.6.1 Flow chart for production of polyclonal anti-phosphopeptide antibodies.

Deplete reactivity of serum toward phosphotyrosine and carrier protein

1. Connect the washed BSA-agarose and phosphotyrosine affinity matrix columns in series.
2. Pass ~15 ml crude serum through the two columns, allowing it to flow through by gravity. Wash with PBS/azide until all of the yellow color of the serum has passed through the columns (or monitor the A_{280} of the column effluent spectrophotometrically until baseline absorbance is reached), then wash with an additional 5 to 10 ml PBS/azide, collecting all washings in the flowthrough fraction, which will contain the antibody of interest.

Aliquots of the crude serum as well as this first flowthrough may be saved for subsequent analysis and comparison of individual purification fractions. The serum may be analyzed at this time for elimination of anti-phosphotyrosine reactivity by performing immunoblotting (UNIT 10.8) of samples containing a variety of phosphotyrosyl proteins; alternatively it is possible to save an aliquot for analysis and proceed to the next step (see Troubleshooting).

Expect that with each passage through a column, the volume of the serum will increase, which will prolong the amount of time needed for subsequent columns. Note that if flowthrough is monitored by the passage of the yellow color of the serum, this will become more difficult as the serum becomes more dilute.

After the serum passes through, regenerate the columns by washing with 10 bed volumes of 3 M NaSCN followed by 10 bed volumes of PBS/azide, and store at 4°C in PBS/azide for future use.

Deplete reactivity of serum toward cognate nonphosphopeptide

3. Pass the flowthrough serum through the nonphosphopeptide affinity matrix column by gravity as many times as necessary to deplete cross-reactivity, regenerating the column between passes by washing with 10 bed volumes of 3 M NaSCN followed by 10 bed volumes PBS/azide.

A 15-ml starting serum sample may have to be passed over this column several times to quantitatively deplete cross-reactivity (see Critical Parameters). The serum may be analyzed after each or several passes, or an aliquot can be saved for analysis at a later time.

4. If cross-reactivity with homologous phosphoprotein(s) is anticipated, pass the flowthrough serum through the homologous phosphopeptide affinity matrix column(s) using the methodology described in step 3.

Refer to Strategic Planning, above, for discussion of cross-reactivity with homologous phosphoproteins.

Purify antibodies by positive-selection affinity chromatography and dialysis

5. Hydrate and wash ~25-cm strips of dialysis tubing in advance for collection of fractions from the positive-selection affinity purification. Secure one end of each length of tubing with a dialysis clamp and check for leaks (APPENDIX 3C). Also prepare 6 liters PBS/azide and cool to 4°C in advance to use as the dialysis solution.

To minimize the time that the antibodies are exposed to the elution solution, fractions (~3 ml each) from the positive-selection affinity column will be collected directly into preprepared dialysis tubing.

6. Pass the flowthrough serum from the previous chromatography step through the positive-selection phosphopeptide affinity column three times, without washing between passes.

Three passes are performed to maximize the interaction of the antibody with the affinity matrix.

7. Collect the flowthrough serum from the final pass, then wash the column with 5 to 20 ml PBS/azide (depending upon the precolumn volume and column bed volume) and combine the washings with the flowthrough serum. Wash with an additional 20 ml PBS/azide, and collect the washings separately as the “wash” fraction.
8. Elute with chaotropic agent of choice: either 20 ml of 3 M NaSCN, or 10 ml of 3.5 M MgCl₂ followed by 10 ml of 4.5 M MgCl₂. Upon starting the elution, immediately begin collecting ~3-ml fractions directly into the dialysis bags prepared in step 5, securing the proximal end of the tubing with dialysis clamps and dropping the bag immediately into the PBS/azide dialysis solution. Collect at least six fractions.

Alternatively, ~3-ml fractions can be collected in tubes and the liquid immediately placed into dialysis tubing upon completion of each fraction. Most of the antibody will elute in the first three fractions from a column of 2- to 3-ml bed volume. Regenerate the column as in step 3.

See Critical Parameters for a discussion of the relative merits of NaSCN versus MgCl₂ as eluants.

9. Dialyze all fractions collected in step 8 exhaustively against PBS/azide at 4°C (APPENDIX 3C).

Analyze and store purified antibody

10. Determine protein concentration of dialyzed fractions by measuring absorbance at 280 nm or by a colorimetric protein assay (UNIT 10.1) and calculate yield.

The authors have generally recovered ≥1 mg of purified antibody from a 15-ml serum sample.

11. Store the antibody in aliquots at -70°C; store working aliquot at 4°C.
12. Assay the reactivity and cross-reactivity of the final samples, as well as any aliquots saved from previous steps, using ELISA (UNIT 11.2) or immunoblotting (UNIT 10.8).

BASIC PROTOCOL 2

PRODUCTION OF MONOCLONAL ANTI-PHOSHOPEPTIDE ANTIBODIES

Strategic Planning

As in the production of polyclonal anti-phosphopeptide antibodies (see Basic Protocol 1), the first step of the planning phase for anti-phosphopeptide monoclonal antibodies is the design and synthesis of the required peptides (see Support Protocol 1 and Background Information). Here too, the authors employ a peptide containing fifteen amino acids with the phosphorylation site Tyr in the center (residue 8), consisting of fourteen amino acids from the actual sequence of the protein plus an N-terminal Lys to enhance coupling to a carrier protein. A different carrier protein must be employed for immunization than will be used as substrate for the ELISA screen, to avoid detection of anti-carrier antibodies in the ELISA (see Purification Method, below). The authors use keyhole limpet hemocyanin (KLH) as the carrier for immunization and BSA as the carrier for the ELISA substrate. General methods for coupling of peptides to carrier proteins for use in immunization are described in UNIT 11.16. The cognate phosphopeptide can also be immobilized to an affinity-chromatographic column support (see Support Protocol 2) and used ultimately to affinity-purify the antibody produced by the desired hybridoma clone. A number of additional peptides must be produced for use in ELISA screening of candidate hybridoma supernatants, to screen both for the presence of the desired reactivity and the absence of undesired cross-reactivity. For ELISA screening, the peptides are coupled to a carrier protein different from the one used for immunization. Bovine serum albumin (BSA) is

Production of Antibodies That Recognize Phosphorylated Peptides

18.6.6

used for screening, because many of the clones are likely to react against the carrier used in the immunization of the mice (KLH) and because unconjugated peptides have not been found to function well as ELISA substrates (Doolittle, 1986). In addition, because antibodies may be generated against the cross-linking reagent, the method used for coupling the phosphopeptide to the BSA to produce the substrate for ELISAs should be different from that used to link peptides to the KLH carrier for immunization (Czernik et al., 1991; Doolittle, 1986).

The second step of planning is to raise antibodies in mice. General guidelines for the immunization of mice and the production of monoclonal antibodies are described in *UNITS 11.4-11.11*, and many institutions have core facilities that will perform that task. The authors have immunized BALB/c mice by intraperitoneal injection with 1 mg/ml of immunogen (dissolved in PBS) in a 50% emulsion with complete Freund's adjuvant on day 1 and boosted intraperitoneally on days 15 and 37 with immunogen in incomplete Freund's adjuvant. Test bleeds from day 47 were analyzed by ELISA. On day 57 the mouse with the best titer was boosted intravenously, and its spleen was harvested on day 60. Freshly harvested spleen cells were prepared for cell fusion to generate hybridoma lines, which were subsequently screened by ELISA (*UNIT 11.7*). The purpose of the ELISA screening is to identify clones that react with the cognate phosphopeptide and to eliminate clones that cross-react with other epitopes in addition to the cognate phosphopeptide. Epitopes screened by ELISA generally consist of the cognate nonphosphorylated peptide (to eliminate clones with phosphorylation-independent reactivities) and at least one unrelated phosphopeptide (to eliminate clones with indiscriminant anti-phosphotyrosine reactivity). In addition, depending upon the particular target chosen, it may be necessary to screen for reactivity with phosphopeptides from proteins of known homology. This occurs particularly when the target protein is a member of a family of homologous proteins—e.g., the subgroups of related receptor tyrosine protein kinases. If such cross-reactivity is anticipated, it is essential to synthesize not only the cognate phosphorylated and nonphosphorylated peptides and at least one unrelated phosphopeptide, but also phosphopeptides based upon the related homologous sequences. Thus, a typical ELISA screening scheme might consist of the following steps.

1. Screen for clones that exhibit reactivity against the cognate phosphopeptide. These clones, which will likely be <10% of all the clones, are potential candidates for the desired clones.
2. Screen for clones with reactivity against the cognate nonphosphopeptide. These clones, which are phosphorylation state-independent, are then eliminated.
3. Screen for clones with reactivity against at least one unrelated phosphopeptide. These clones, which are likely to represent indiscriminant anti-phosphotyrosine reactivity, are also eliminated.
4. Screen for and eliminate clones with reactivity against potential cross-reacting homologous peptides, if any are anticipated.

General guidelines for ELISA screening are described in *UNIT 11.2*. In the ELISAs performed in the authors' laboratory, 96-well polyvinyl chloride assay plates are generally used and bound monoclonal antibody is detected using biotinylated horse anti-mouse secondary antibody followed by horseradish peroxidase (HRPO)-conjugated Avidin D (Vector Labs) and 4-aminoantipyrine(Sigma)/H₂O₂ as the substrate. This gives a deep red color in positive wells and absence of color in negative wells, allowing easy visual scoring without the need for quantitative measurements. The clones that pass all the screening steps are candidates for desirable hybridoma cell lines. These are expanded, frozen down as a pool, subcloned by limiting dilution, then rescreened against cognate phosphopeptide

to ensure continued production of antibody through the early passages. Those continuing to produce are expanded further and a larger quantity of either culture supernatant or ascites fluid is screened in further immunoassays. Reactivity is demonstrated most simply by immunoblot analysis using a panel of relevant phosphorylated and nonphosphorylated proteins. Electrophoresis is carried out with purified (generally by immunoprecipitation using conventional antibodies) cognate protein in both its phosphorylated and unphosphorylated states, as well as with any proteins toward which there could conceivably be cross-reactivity. The blot is probed with the monoclonal antibody preparation to demonstrate appropriate reactivity. The antibody preparation can ultimately be purified on a cognate phosphopeptide affinity column.

The relationship among the above procedures is illustrated by the flow chart in Figure 18.6.2.

Purification Method

This protocol presents procedures for isolating hybridoma clones of the desired reactivity, starting from the point at which candidate hybridoma clones have been seeded in 96-well culture plates (see Fig. 18.6.2). Candidate hybridoma supernatants are screened in ELISA assays against a series of peptides (as described in Strategic Planning). Identified clones are expanded, subcloned by limiting dilution, and further purified.

Materials

- Candidate hybridoma cell lines from fusion (refer to Strategic Planning and *UNITS 11.4-11.11*)
- HT medium (*UNIT 11.7*)
- Screening diluent (see recipe)
- BSA-conjugated cognate phosphopeptide (for use as ELISA antigen; refer to Strategic Planning, Support Protocol 1, *UNIT 11.2* & *UNIT 11.15*)
- Negative control: preimmune serum from mouse used to produce hybridoma line (*UNIT 11.4*)
- Positive control: immune serum from mouse used to produce hybridoma line (*UNIT 11.4*)
- BSA-conjugated cognate nonphosphopeptide (for use as ELISA antigen; refer to Strategic Planning, Support Protocol 1, and *UNIT 11.2*)
- BSA-conjugated noncognate phosphotyrosyl peptide (for use as ELISA antigen; refer to Strategic Planning, Support Protocol 1, and *UNIT 11.2*)
- BSA-conjugated homologous phosphopeptides (optional, for use as ELISA antigens; refer to Strategic Planning, Support Protocol 1, and *UNIT 11.2*)
- 96-well polystyrene tissue culture plates
- Grid note sheets
- Additional reagents and equipment for cloning by limiting dilution (*UNIT 11.8*), ELISA (*UNIT 11.2*), and freezing and recovery of hybridoma cell lines (*UNIT 11.9*)

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

Identify and sample conditioned media from candidate hybridomas

1. Seed hybridoma cells from fusion at a low density (aiming for one cell per three wells) in multiple 96-well polystyrene tissue culture plates using HT medium (see *UNIT 11.7*). Examine at ~2 weeks and identify all single-colony hybridoma culture wells as potential candidates by recording the number of colonies per well on a 96-grid note sheet. Mark the underside of each well containing one colony for easy identification.

The supernatant fluid from wells possessing a single colony will be screened for reactivity.

- Using sterile technique, remove an aliquot of supernatant (e.g., 100 μ l from a 200- μ l culture) from each potential candidate well and transfer to a separate 96-well polystyrene plate (the “screen plate”). Record the original plate number and well position and the corresponding plate number and well position for the screen plate (e.g., on a grid sheet representing the screen plate, record the original plate number and well position from which the supernatant in each well was transferred). Replace the aliquot removed from each original well with 100 μ l 37°C fresh HT medium.

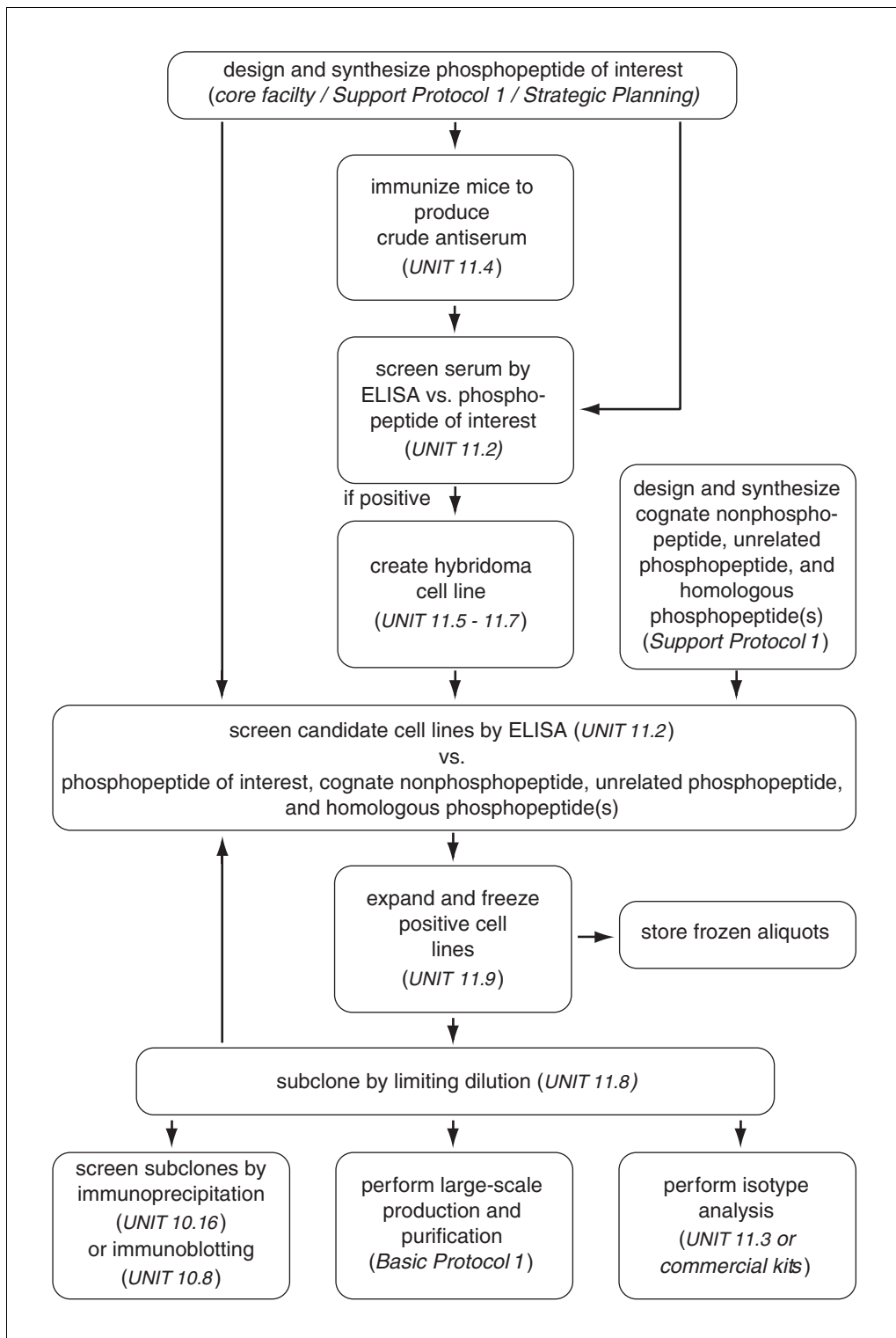


Figure 18.6.2 Flow chart for production of monoclonal anti-phosphopeptide antibodies.

A brightly colored dot sticker the size of one well, attached to the tissue culture hood surface, will help to keep the plate in register, as the plate may be moved one well at a time along the sticker. If the supernatants are not to be tested immediately, store the screen plates at 4°C, wrapped in plastic wrap.

3. To each supernatant in the screen plates, add 150 µl screening diluent

The components of the screening diluent prevent microbial contamination; this reagent also serves to expand the volume should multiple rounds of screening be required.

Screen candidate hybridomas by ELISA

4. Screen an aliquot (e.g., 50 µl) of each candidate hybridoma supernatant from the screen plate in an ELISA (see *UNIT 11.2* as a general guide) using the cognate phosphopeptide–BSA conjugate as antigen. Record the screen-plate number and well position and corresponding original culture plate number and well position for each positive sample. Also assay preimmune serum from the mouse used for fusion (negative control) and immune serum from the same mouse (positive control), each at a 1:100 dilution in a 1:1 mixture of HT medium and screening diluent. Include an ELISA with screening diluent only (additional negative control).

It is important to use a different carrier protein in the ELISAs (BSA) than that used for the immunizations (KLH), to avoid detecting reactivity against the carrier (refer to Strategic Planning, above).

5. Screen an aliquot of each positive sample from step 4 in an ELISA (*UNIT 11.2*) against the cognate nonphosphopeptide–BSA conjugate to eliminate clones with phosphorylation-independent reactivity, against the unrelated phosphotyrosylpeptide–BSA conjugate to eliminate clones with indiscriminant phosphotyrosine reactivity, and, if cross-reactivity with homologous phosphoprotein(s) is anticipated, against phosphopeptide-BSA conjugates corresponding to any potentially cross-reacting homologous phosphoproteins.

These screens may be done sequentially or, if the number of candidate hybridomas has diminished sufficiently, simultaneously. Isolation of clones with phosphorylation-independent cognate peptide specificity may also be useful.

6. Expand clones that satisfy all screening requirements and freeze aliquots (*UNIT 11.9*) as backup; use remainder for subcloning by limiting dilution (*UNIT 11.8*).
7. Screen culture supernatants from subclones as in steps 4 and 5 to check for continued production and specificity of the antibody.

Keep representative independent subclones from each original parental colony, as it cannot be predicted whether each parental monoclonal antibody will recognize the cognate holophosphoprotein or function in all types of immunoassays desired in the future. Candidate subclones should then also be expanded and frozen. It is a good idea to continue to check for production of the appropriate antibody after several rounds of serial passaging and after freezing, thawing, and reexpansion.

8. To characterize the subclones further and identify the most useful clones, assay the culture supernatants for reactivity with cognate holophosphoprotein in individual immunoassays (e.g., immunoprecipitations, as described in *UNIT 10.16*, or immunoblot analyses, as described in *UNIT 10.8*).

*Antibody from the desired clones can be produced on a large scale by harvesting large volumes of culture supernatant and performing affinity purification (see Basic Protocol 1), or by producing ascites fluid (*UNIT 11.10*), which can be purified by ammonium sulfate precipitation followed by affinity purification. Isotype analysis can be performed as described in *UNIT 11.3* or by using commercially available kits (Pierce or Sigma).*

SYNTHESIS OF PEPTIDES

Many institutions have core facilities for synthesis of peptides using automated solid-phase peptide synthesizers. For phosphotyrosine-containing peptides, the authors originally used standard Merrifield solid-phase procedures with *t*-butyloxycarbonyl (Boc) amino acids and Boc-*O*-(dibenzylphosphono)-L-tyrosine as the phosphorylated residue (Barany and Merrifield, 1979). For cleavage of phosphopeptide from the resin, 1 g of phosphopeptide resin was stirred 10 min at 0°C with 2 ml dimethyl sulfide; then an ice-cold mixture of 4 ml trifluoromethanesulfonic acid (TFMSA) and 10 ml of trifluoroacetic acid was added slowly with stirring. The resulting reaction mixture was stirred under a nitrogen atmosphere 4 hr at 0°C and the cleaved phosphopeptide precipitated with methyl *t*-butyl ether at -30°C. The precipitate was dissolved in 40% acetic acid and lyophilized several times from water. This original method suffers from low yields. Currently, most phosphotyrosyl-containing peptides are synthesized with 9-fluorenylmethyloxycarbonyl (Fmoc) phosphotyrosine. The lack of a side chain protecting group on the phosphotyrosine (Kitas et al., 1994) allows standard Fmoc synthesis and cleavage procedures to be employed (Chang and Meienhofer, 1978). However, a higher molar excess of activated Fmoc amino acids than that used in standard procedures may be required for efficient coupling of amino acids after unprotected phosphotyrosine has been added. In some circumstances, a double coupling procedure may be employed to add the few amino acids following phosphotyrosine to the growing peptide chain. This alternative requires the sequential, repeated addition of the next particular amino acid to be added to the peptide. A support protocol for the coupling of peptides to an affinity matrix is provided (see Support Protocol 2); general methods for coupling peptides to carrier proteins are described in *UNIT 11.16*.

COUPLING OF PEPTIDES TO AFFI-GEL 10 AFFINITY MATRIX

This protocol describes coupling of phosphopeptides or nonphosphopeptides to Affi-Gel 10 affinity matrix for use in affinity column chromatographic purification of antibodies. Anhydrous coupling, detailed here, is the most efficient coupling method for peptides. The steps here result in production of 3 ml final bed volume of affinity resin; 3 μmol of peptide is coupled per milliliter Affi-Gel 10. According to the manufacturer's specifications, Affi-Gel 10 resin contains ~15 μmol of active ester per milliliter of gel, and the gel has a capacity for 35 mg of protein or 15 to 20 μmol of a low-molecular-weight ligand per milliliter of gel. Note that these capacities are dependent upon the molecular weight and nature of the protein. After coupling, the matrix may be poured into columns of the appropriate size.

Materials

Synthetic oligopeptide for coupling (see Support Protocol 1)
Dimethyl sulfoxide (DMSO)
N-methylmorpholine (99% purity; Acros Organics)
Affi-Gel 10 (Bio-Rad) or equivalent activated support matrix
Ethanolamine
0.1 M ethanolamine·HCl, pH 8.0
High-salt/high-pH solution: 0.5 M NaCl/0.4% (w/v) sodium bicarbonate
High-salt/low-pH solution: 0.5 M NaCl/100 mM sodium acetate, pH 4.2
PBS/azide: PBS (*APPENDIX 2*) containing 0.02% (w/v) sodium azide
0.5 M NaCl
3 M NaSCN

Polypropylene screw-cap centrifuge tubes (do not use polystyrene tubes with DMSO)

SUPPORT PROTOCOL 1

SUPPORT PROTOCOL 2

Analysis of Protein Phosphorylation

18.6.11

End-over-end rotator
IEC Clinical centrifuge (or equivalent)
Vacuum aspirator
Glass chromatography columns, ≥ 5 -ml capacity

Prepare peptide solution

1. Dissolve peptide in DMSO at a volume ~ 0.5 to 4 times the desired bed volume.

For a 3-ml bed volume, dissolve a total of 9 μmol peptide, which for an average 15-mer peptide of mol. wt. ~ 1700 would be 15.3 mg.

2. Neutralize the peptide solution by titration as follows: add peptide synthesis-grade *N*-methylmorpholine (full-strength or diluted to 1:1 to 1:9 in DMSO) in 1 to 2 μl increments, remove a 2 to 3 μl aliquot after each addition, dilute this with 50 μl water, then spot it onto a strip of pH paper. Repeat this process until a pH in the range of 7 to 8 is reached.

*IMPORTANT NOTE: Titrate carefully, as it is very easy to overshoot the desired pH. Typically 7 to 8 μl of *N*-methylmorpholine is required per 3 μmol peptide, although the precise amount will depend greatly upon the particular amino acid sequence. The aliquots must be mixed with water because pH is a measure of hydrogen ion concentration, and hence the pH of a nonaqueous solution cannot be measured.*

Add peptide to Affi-Gel 10

3. Mix contents of the bottle of Affi-Gel 10 matrix well to resuspend the resin, then transfer the required amount of resin to a screw-cap polypropylene centrifuge tube and wash three times in DMSO, each time by centrifuging 5 min at $700 \times g$ (2000 rpm in an IEC Clinical centrifuge, setting 5), room temperature, aspirating the supernatant, resuspending the resin in 5 vol DMSO, then centrifuging again at $700 \times g$.

IMPORTANT NOTE: Be sure to take at least twice as much Affi-Gel 10 as ultimately will be needed, as the volume will shrink. The resin is supplied as a 50% slurry; therefore, it is necessary to take at least four times the required bed volume. Avoid centrifugation at higher than recommended speed as this may result in collapse of the resin.

4. Aspirate the excess DMSO from the resin and add the peptide solution prepared in step 2. Incubate overnight at room temperature with end-over-end rotation.

Quench and wash the resin

5. Add 2 μl pure ethanolamine per ml resin and incubate 2 hr at room temperature with end-over-end rotation.

Ethanolamine is added directly to the coupling reaction to quench unreacted ester groups.

6. Wash twice in DMSO as in step 3.
7. Wash twice in 0.1 M ethanolamine-HCl, pH 8.0, as in step 3 (performing the first wash on ice as a great deal of heat will be generated). Remove the 0.1 M ethanolamine-HCl from the second wash and replace with fresh, incubate tube overnight at 4°C with end-over-end rotation, then centrifuge at low speed and aspirate the supernatant.
8. Wash three times with high-salt/high-pH solution using the technique described in step 3.

9. Wash three times with high-salt/low-pH solution using the technique described in step 3.
10. Wash three times with PBS/azide using the technique described in step 3 and store washed resin at 4°C until ready to pack in column.
11. Mix matrix into a slurry with 0.5 M NaCl and pour into glass chromatography columns.
Commercially available glass chromatography columns can be used; alternatively, 5-ml plastic syringes with 1-cc plugs of silanized glass wool (UNIT 5.6) in the bottom may function as columns.
12. Wash each column with 10 column volumes 3 M NaSCN followed by 10 column volumes PBS/azide.

COUPLING OF PHOSPHOTYROSINE TO AFFI-GEL 10 AFFINITY MATRIX

SUPPORT PROTOCOL 3

This protocol describes coupling of phosphotyrosine to Affi-Gel 10 affinity matrix for use in affinity column chromatographic purification of antibodies. Phosphotyrosine is insoluble under the anhydrous conditions used for coupling peptides (see Support Protocol 2); thus coupling is done under aqueous conditions. The steps here result in production of 10 ml final bed volume of affinity resin; as for peptides, 3 μmol phosphotyrosine is coupled per milliliter of Affi-Gel 10. According to the manufacturer's specifications, Affi-Gel 10 resin contains ~ 15 μmol of active ester per ml of gel, and the gel has a capacity for 35 mg of protein per ml gel or 15 to 20 μmol of a low-molecular-weight ligand per ml gel. Note that these capacities may vary depending upon the molecular weight and nature of the protein. Phosphotyrosine-agarose affinity matrix is now also available commercially from Sigma.

Materials

- Phosphotyrosine
- 1 M NaOH (optional)
- 0.4% (w/v) sodium bicarbonate
- Fritted-glass funnel and vacuum aspirator
- Glass chromatography column, ≥ 14 -ml capacity
- Additional reagents and equipment for coupling peptides to Affi-Gel 10 affinity matrix (see Support Protocol 2)

Prepare phosphotyrosine solution

1. Dissolve phosphotyrosine in 0.4% sodium bicarbonate at a volume ~ 0.5 to 4 times the desired bed volume.
For a 10-ml bed volume, dissolve a total of 30 μmol phosphotyrosine. Based on its molecular weight of 261, this amounts to 7.8 mg, although it is acceptable (and may be easier) to use an excess, as phosphotyrosine is relatively inexpensive.
2. Check that the pH is in the range of 7 to 8 using pH paper. If adjustment is necessary, neutralize the solution by titration—i.e., adding small amounts of 1 M NaOH, removing a 2- to 3- μl aliquot after each addition, then spotting the aliquots onto a strip of pH paper. Repeat this process until a pH in the range of 7 to 8 is reached.

Add phosphotyrosine to Affi-Gel 10

3. Mix contents of the bottle of Affi-Gel 10 matrix well to resuspend resin, then transfer the required amount of resin to a fritted-glass funnel attached to a vacuum aspirator.

Analysis of Protein Phosphorylation

18.6.13

Wash three times, each time by pouring cold distilled water over the resin in the funnel and aspirating, then perform one final wash in the same manner using ice-cold 0.4% sodium bicarbonate.

IMPORTANT NOTE: *Be sure to take at least twice as much Affi-Gel 10 as ultimately will be needed, as the volume will shrink. The resin is supplied as a 50% slurry; therefore, it is necessary to take at least four times the required bed volume. Do not allow >20 min to elapse from the time the resin is removed from the bottle to the time it is mixed with the phosphotyrosine solution.*

4. Remove most of the excess liquid from the gel by vacuum aspiration without letting it dry completely, transfer it to a screw-cap centrifuge tube, and add the phosphotyrosine solution from step 2.
5. Incubate overnight at 4°C with end-over-end rotation.

Quench and wash resin

6. Add 2 µl pure ethanolamine per ml resin and incubate 2 hr at room temperature with end-over-end rotation.

Ethanolamine is added directly to the coupling reaction to quench unreacted ester groups.

7. Wash resin twice, each time by centrifuging at low speed (see Support Protocol 2, step 3), aspirating the supernatant, resuspending the resin in 5 vol of 0.4% sodium bicarbonate, then centrifuging again at low speed.
8. Wash resin twice with 0.1 M ethanolamine-HCl, pH 8.0, as in step 7. Remove the 0.1 M ethanolamine from the second wash and replace with fresh, incubate tube overnight at 4°C with end-over-end rotation, then centrifuge at low speed and aspirate the supernatant.
9. Wash, store, and pack resin in columns (see Support Protocol 2, steps 8 to 12).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Screening diluent

- 6.18 g boric acid (100 mM final)
- 9.54 g sodium borate (47 mM final)
- 4.38 g NaCl (75 mM final)
- 10 g bovine serum albumin (BSA; 1% w/v final)
- 1 g sodium azide (1% w/v final)
- H₂O to 1 liter

The pH of the solution will be 8.4 to 8.5.

COMMENTARY

Background Information

Conventional antibodies have proven to be invaluable tools in numerous techniques for the biochemical analysis of proteins. In the past antibodies with specificity toward the phosphorylated forms of proteins were produced serendipitously. Using techniques such as those described in this unit, such antibodies may now be produced by design. Whereas

phosphorylated holoproteins might in general be poor candidates for immunogens in the production of such antibodies as a result of their susceptibility to dephosphorylation, it has been suggested that short phosphopeptides are relatively resistant to phosphatases (Czernik et al., 1991), thereby providing a better chance of success.

A dramatic advance in the analysis of pro-

tein tyrosine phosphorylation and the regulation of signal transduction pathways by such phosphorylation occurred with the production of polyclonal and monoclonal anti-phosphotyrosine antibodies (Frackelton et al., 1983; Ross et al., 1981; reviewed in Stern, 1991). These antibodies proved capable of recognizing phosphorylated tyrosine residues in the context of virtually any flanking peptide sequence, and were shown to be exquisitely specific in the requirement for a phosphorylated tyrosine yet at the same time remarkably promiscuous in their acceptance of any peptide sequence. Techniques using polyclonal and monoclonal anti-phosphotyrosine antibodies have supplanted the standard methodology of metabolically labeling proteins with $^{32}\text{P}_i$, which was cumbersome, time-consuming, hazardous, and of relatively lower sensitivity.

Antibodies that recognize a specific tyrosyl-phosphorylated peptide as described in this unit represent a marriage of anti-phosphotyrosine technology with the stringent sequence-dependent specificity of conventional antibodies (Bangalore et al., 1992; DiGiovanna and Stern, 1995; Epstein et al., 1992; Roussel, 1990). Conceptually, they may be thought of either as anti-phosphotyrosine antibodies that have strict sequence specificity, or as conventional antibodies possessing the additional specificity of phosphorylation dependence.

Anti-phosphotyrosine antibodies have been most useful in the analysis of tyrosine phosphorylation of proteins in a technique using a combination of immunoprecipitation and immunoblotting (known as the "IP-western" method). In this procedure, a protein typically is immunoprecipitated either with conventional anti-protein antibody or with anti-phosphotyrosine antibody, then immunoblotted with whichever of these two antibodies was not used for the immunoprecipitation. Because anti-phosphopeptide antibodies recognize the cognate protein, but only when it is phosphorylated, they are capable of performing both functions in a single step—i.e., either an immunoprecipitation alone or an immunoblot alone is sufficient to supply the desired information.

The most useful functions for such antibodies are likely to be those that in fact are not possible using traditional reagents. For example, using anti-phosphopeptide antibodies, it is possible to identify and isolate distinct phosphorylated species of a phosphoprotein containing multiple phosphorylation sites. Thus, phosphorylation at each site can be examined

independently, and a preparative separation of individual phosphospecies of a holoprotein can be achieved. Another unique application is analysis of the abundance and phosphorylation state of individual proteins *in situ* in preparations of cells or tissues. Because phosphorylation is a major mode of regulation of protein function, the phosphorylation state is often an indicator of the functional status of a protein. The mere identification of a protein in a cell or tissue specimen gives no indication of its functional status. The ability of anti-phosphopeptide antibodies to demonstrate the phosphorylation state, and by extrapolation the functional state, of a single protein with high specificity places these reagents in a unique class. Hence, identity and functional status may be probed simultaneously using one simple assay. In tissue sections, one may probe to determine whether a particular protein is present, and, if present, in what functional state it is found. The authors have employed this strategy in immunohistochemical staining of formalin-fixed, paraffin-embedded human breast tumors using antibody to the phosphorylated form of the receptor tyrosine kinase Neu (DiGiovanna and Stern, 1995), and were able to demonstrate that Neu is phosphorylated, and therefore functionally active, in only a subset of the tumors that overexpress this protein.

Anti-phosphopeptide antibodies possess unique properties that render them capable of performing functions not possible with conventional reagents. From the technical considerations regarding their production, it would follow that antibodies with specificity for the non-phosphorylated state should also be achievable, and the production of such antibodies has also been reported (Czernik et al., 1991; Epstein, 1995; Kawakatsu et al., 1996; Nairn et al., 1982; Roussel, 1990; Tzartos et al., 1995).

Critical Parameters

In the production of polyclonal anti-phosphopeptide antibodies, a major challenge is the depletion of cross-reactivity. The first consideration in achieving this is the anticipation of potential cross-reacting proteins in the planning phase of the procedure. Cross-reactivity is most likely to occur when the target protein is a member of a family of closely related homologous proteins—e.g., of the subfamilies of tyrosine protein kinases. The authors have produced a monoclonal anti-phosphopeptide antibody with specificity for the Tyr-1248 autophosphorylation site of Neu (DiGiovanna and Stern, 1995). A homologous site exists in

the epidermal growth factor receptor (EGFR). The antibody produced in the authors' laboratory does not cross-react with the homologous EGFR site in spite of having identical residues in seven of the 14 amino acids of the peptide sequence (five of which are N-terminal to the phosphotyrosine, one C-terminal, and the phosphotyrosine itself). Thus, specificity is achievable even with at least up to 50% identity. Nevertheless, such a close relationship among the phosphorylation sites of different proteins highlights the importance of considering all known related proteins.

The second consideration in depleting cross-reactivity is the quantity of cross-reactive antibodies in a given aliquot to be purified relative to the capacity of the peptide columns used for the negative selection. Because of the expense of the peptides, the authors have generally prepared small columns (with 2- to 3-ml bed volumes) and passed the sera over the columns multiple times, empirically determining the number of passes necessary to quantitatively deplete cross-reacting material. In these columns, 3 μmol of ligand is coupled per milliliter of Affi-Gel 10. The manufacturer states that the resin contains $\sim 15 \mu\text{mol}$ of active ester per milliliter of gel, and that the gel has a capacity for 35 mg of protein or 15 to 20 μmol of a low-molecular-weight ligand per milliliter of gel. The BSA and phosphotyrosine columns are relatively inexpensive to produce; thus, a single large column of each is practical and sufficient.

Because the authors elute the antibodies from the positive-selection affinity column in strongly chaotropic solutions that are potentially deleterious to the stability of the antibody (3 M NaSCN or 3.5 M followed by 4.5 M MgCl_2), care is taken to collect the fractions in preprepared dialysis tubing so that they may be immediately placed into the PBS dialysate, thus minimizing the time that the antibodies are exposed to the eluting solutions. Although MgCl_2 is thought to be "gentler," the authors have found that gravity-driven flow rates from phosphopeptide columns eluted with this salt quickly become extremely slow, possibly because of precipitation of the salt in the column or an interaction of the Mg^{2+} ion with the phosphate groups. Thus use of NaSCN is preferred, and this salt appears to permit recovery of comparable activity.

In the production of monoclonal anti-phosphotyrosyl peptide antibody, which the authors have carried out once, a major technical hurdle was the low frequency with which clones of the

desired specificity were produced (DiGiovanna and Stern, 1995). In that work, in which antibody was produced to the phosphorylated form of the receptor tyrosine kinase Neu, >1200 hybridomas (obtained from a single fusion) were screened to obtain a single clone that satisfied all requirements. The authors found 68 candidate clones that recognized the cognate phosphopeptide, of which only 20 were unreactive toward the cognate nonphosphopeptide. Of those, seven cross-reacted with an unrelated phosphopeptide (i.e., exhibited indiscriminant phosphotyrosine activity), and of the remainder three cross-reacted with the homologous EGFR. The remaining ten were subcloned by limiting dilution, after which only five continued to produce antibody. Of these five, only one reliably detected the phosphorylated holoprotein in immunoblots and immunoprecipitations. Thus, the "hit rate" for an antibody that satisfied all requirements was <1 in 1000 clones. As this is the only monoclonal anti-phosphotyrosyl peptide antibody produced by the authors to date, it is impossible to be sure whether this low frequency will be a general phenomenon. It is crucial to devise a convenient and reliable ELISA assay for such extensive screening. Other assays may require extensive optimization. For example, the monoclonal antibody produced in the above project performed poorly in immunoblotting when any detection system other than a very sensitive chemiluminescence-based assay was used. This was possibly due to a low affinity of the antibody for the cognate protein. In addition, extensive optimization, particularly aimed toward enhancing the strength of the signal, was required for use of the antibody in immunohistochemistry. Finally, it is important to use different carrier proteins for the immunizing conjugate and the ELISA conjugate, so that antibodies to the carrier used for the immunization will not give positive reactions in the ELISA. Some authorities also advocate using different cross-linking agents, as antibodies against the cross-linker itself could theoretically be generated as well (Czernik et al., 1991; Doolittle, 1986).

Troubleshooting

In the production of polyclonal anti-phosphopeptide antibodies, the most common unfavorable outcomes are persistence of cross-reactivity, loss of specific reactivity, and poor yield. Although it may be tempting, in the interest of saving time, to perform the entire purification and postpone all analyses until the

end (a reasonable course of action once a scheme is known to work), it may be prudent, especially with a new protocol, to analyze the serum at each step. For example, the authors have produced antibody to phosphorylated EGFR (M. Digiovanna, M.A. Lerman, and D. Stern, unpub. observ.). The crude antiserum cross-reacted with the related receptors Neu and HER-4, which had been anticipated. The purification was monitored by probing, after each column, an immunoblot consisting of four lanes containing lysates from cells overexpressing (1) phosphorylated EGFR, (2) non-phosphorylated EGFR, (3) phosphorylated Neu, and (4) phosphorylated HER-4. In following the depletion of cross-reactivity, such stepwise analysis demonstrates whether undesired cross-reactivity has been eliminated or whether further passes through a particular negative-selection column are necessary.

If loss of specific reactivity is also a problem, stepwise analyses can also demonstrate where the loss is occurring. Loss of specific reactivity can occur for three major reasons. The first possible reason is that the antibody is not stable over the course of the purification, which may be the case if the procedure is very prolonged. If this is suspected, the pace of the purification may be hastened or the procedure may be carried out at 4°C instead of room temperature. With prolonged purifications, it is especially important to include sodium azide in the PBS used to dilute the serum, to prevent microbial growth. Activity may also be lost nonspecifically as a result of the multiple manipulations in a protocol requiring multiple columns and/or multiple passes through each column. In this case, fraction analysis may indicate which columns are absolutely essential, and also indicate the minimum necessary number of passes through each column. Also, creating larger columns with greater capacity should reduce the number of times the serum will need to be passed over each to deplete cross-reactivity. Finally, specific activity may be lost as a result of the presence in one of the negative-selection columns of a cross-reacting peptide that is so similar to the peptide of interest that essentially any antibody that recognizes the cognate phosphopeptide will recognize this phosphopeptide as well. In this last theoretical scenario, ultimate unique specificity may not be achievable.

The columns used in the purifications can be regenerated and used repeatedly if stored at 4°C in sodium azide-containing buffer to prevent microbial contamination and consequent damage to the matrix. A theoretical considera-

tion is that with repeated use, phosphatases in serum may eventually cleave enough phosphate groups from the peptide moieties on the column to render it ineffective as a phosphopeptide column. Thus, it may be prudent to periodically prepare fresh columns. An alternative would be to add phosphatase inhibitors to the serum prior to purification. Similarly, proteases in serum may eventually cleave peptides from the resin, and this may be prevented by adding protease inhibitors to the serum.

As discussed in Critical Parameters, the main technical hurdle in the production of monoclonal antibodies appears to be the large number of clones that it may be necessary to screen to find one with the desired reactivity. Here, time and persistence are the primary defenses. Another obstacle may be the failure of hybridoma cells to continue to produce antibody upon subcloning, which is likely due to the continuing genetic instability of hybridomas at early passage. If the parental hybridoma colony has been expanded and frozen prior to subcloning, this will serve as a potential source for additional subclones, although it is possible that these additional daughter clones may also be nonproducing. The use of the final monoclonal antibody in a variety of immunoassays is more likely to require potentially extensive optimization of protocols and higher concentrations (because of lower affinities) than the use of polyclonal antibodies.

Anticipated Results

In the production of polyclonal antibodies, the authors have obtained yields of ~1 mg of purified anti-phosphopeptide antibody from a single 15-ml serum sample (M. Digiovanna, M.A. Lerman, and D. Stern, unpub. observ.). These antibodies have been used in the immunoprecipitation of phosphoproteins from both denaturing and nondenaturing solutions. They have also been used in immunoblotting, in immunofluorescence on fixed cultured cells, and in immunohistochemistry on formalin-fixed, paraffin-embedded tissue sections. Specificity can be demonstrated when such assays are inhibited by cognate phosphopeptide but not by cognate nonphosphopeptide or unrelated phosphopeptides.

Through the production of monoclonal antibodies a permanent supply of a uniform reagent is ensured. Many milligrams can be obtained from the ascites produced by a single mouse. The authors have used such an antibody in all of the same assays as described above for polyclonal antibodies. More extensive optimi-

zation of protocols is likely to be required with monoclonal antibodies, and these may be expected to have a lower affinity than typical polyclonal antibodies. Isotype analysis can easily be performed as described in *UNIT 11.3* or by use of commercially available kits (Pierce or Sigma).

Time Considerations

For both protocols, the principal time-consuming step is the production of an immune response in the immunized animals. For the production of polyclonal antibodies, the authors have obtained high titers of reactive sera as early as 6 to 8 weeks in animals that have been boosted every 2 weeks. Later bleeds with continued boosts have yielded even higher titers. Once adequate bleeds have been obtained and the necessary affinity columns prepared, purification can generally be carried out within 1 week, with the exact length of time depending upon the number of columns and the necessity of analyzing the product after each step.

In the production of monoclonal antibodies, the authors have harvested splenocytes at ~2 months from the initial immunization, with several boosts in between. The screening of hybridoma supernatants can take 1 to 2 weeks, and their expansion and subcloning several weeks more.

Literature Cited

- Bangalore, L., Tanner, A.J., Laudano, A.P., and Stern, D.F. 1992. Antiserum raised against a synthetic phosphotyrosine-containing peptide selectively recognizes p185^{neu/erbB-2} and the epidermal growth factor receptor. *Proc. Natl. Acad. Sci. U.S.A.* 89:11637-11641.
- Barany, G. and Merrifield, B. 1979. Solid phase peptide synthesis. In *The Peptides*, Vol. I (E. Gross and J. Meienhofer, eds.) pp. 1-284. Academic Press, New York.
- Chang, C.D. and Meienhofer, J. 1978. Solid-phase peptide synthesis using mild base cleavage of *N* alpha-fluorenylmethoxycarbonylamino acids, exemplified by a synthesis of dihydrosomatostatin. *Int. J. Pept. Protein Res.* 11:246-249.
- Czernik, A.J., Girault, J.-A., Nairn, A.C., Chen, J., Snyder, G., Kebabian, J., and Greengard, P. 1991. Production of phosphorylation state-specific antibodies. *Methods Enzymol.* 201:264-283.
- DiGiovanna, M.P. and Stern, D.F. 1995. Activation state-specific monoclonal antibody detects tyrosine phosphorylated p185^{neu/erbB-2} in a subset of human breast tumors overexpressing this receptor. *Cancer Res.* 55:1946-1955.
- Doolittle, R.F. 1986. Of urfs and orfs: A primer on how to analyze derived amino acid sequences. University Science Books, Mill Valley, Calif.

- Epstein, R.J. 1995. Preferential detection of catalytically inactive c-erbB-2 by antibodies to unphosphorylated peptides mimicking receptor tyrosine autophosphorylation sites. *Oncogene* 11:315-323.
- Epstein, R.J., Druker, B.J., Roberts, T.M., and Stiles, C.D. 1992. Synthetic phosphopeptide immunogens yield activation-specific antibodies to the c-erbB-2 receptor. *Proc. Natl. Acad. Sci. U.S.A.* 89:10435-10439.
- Frackelton, A.R. Jr., Ross, A.H., and Eisen, H.N. 1983. Characterization and use of monoclonal antibodies for isolation of phosphotyrosyl proteins from retrovirus-transformed cells and growth factor-stimulated cells. *Mol. Cell. Biol.* 3:1343-1352.
- Kawakatsu, H., Sakai, T., Takagaki, Y., Shinoda, Y., Saito, M., Owada, M.K., and Yano, J. 1996. A new monoclonal antibody which selectively recognizes the active form of src tyrosine kinase. *J. Biol. Chem.* 271:5680-5685.
- Kitas, E., Kung, E., and Bannwarth, W. 1994. Chemical synthesis of *O*-thiophosphotyrosyl peptides. *Int. J. Pept. Protein Res.* 43:146-153.
- Nairn, A.C., Detre, J.A., Casnellie, J.E., and Greengard, P. 1982. Serum antibodies that distinguish between the phospho- and dephospho-forms of a phosphoprotein. *Nature* 299:734-736.
- Ross, A.H., Baltimore, D., and Eisen, H.N. 1981. Phosphotyrosine-containing proteins isolated by affinity chromatography with antibodies to a synthetic haptent. *Nature* 294:654-656.
- Roussel, R. 1990. Synthetic phosphopeptides modeled on the carboxyl terminus of pp60^{C-SRC}: Specific antibodies, binding and effects on kinase activity. M.S. thesis, University of New Hampshire, Durham.
- Stern, D.F. 1991. Antiphosphotyrosine antibodies in oncogene and receptor research. *Methods Enzymol.* 198:494-501.
- Tzartos, S.J., Kouvatso, R., and Tzartos, E. 1995. Monoclonal antibodies as site-specific probes for the acetylcholine-receptor δ -subunit tyrosine and serine phosphorylation sites. *Eur. J. Biochem.* 228:463-472.

Key References

- Bangalore et al., 1992. See above.
- A description of the production of polyclonal anti-phosphopeptide antibody by the method described in this unit.*
- Czernik et al., 1991. See above.
- A general discussion of the production of phosphorylation-dependent antibodies.*
- DiGiovanna and Stern, 1995. See above.
- A description of the production of monoclonal anti-phosphopeptide antibody by the method described in this unit.*

Doolittle, 1986. See above.

A general discussion of the analysis of protein sequences and the production of anti-peptide antibodies.

Epstein et al., 1992. See above.

A description of the production of polyclonal anti-phosphopeptide antibody by a variation of the method described in this unit.

Harlow, E. and Lane, D. 1988 Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

A comprehensive text describing the production and utilization of antibodies.

Contributed by Michael P. DiGiovanna
Yale University School of Medicine
New Haven, Connecticut

Robert R. Roussel
Dartmouth College
Hanover, New Hampshire

David F. Stern
Yale University School of Medicine
New Haven, Connecticut

Assays of Protein Kinases Using Exogenous Substrates

In studies of the regulation of specific biochemical events by reversible phosphorylation, assaying the protein kinases themselves can often lead to significant progress in understanding the mechanistic details of a system under study. This unit describes assays for a variety of protein kinases that require different conditions to detect and measure their activities: cyclic nucleotide-dependent kinases (see Basic Protocol 1), protein kinase C and isoforms (see Basic Protocol 2), casein kinases (see Basic Protocol 3 and Alternate Protocol), Ca²⁺/calmodulin-dependent kinases (see Basic Protocol 4), and tyrosine kinase (see Basic Protocol 5). The unit is not meant to be a catalog of individual protein kinase assays; however, the general principles of these assays should apply to most if not all known protein kinases. It is also possible to perform in-gel assays for specific kinase activity (see Basic Protocol 6).

To assay the phosphotransfer reactions catalyzed by protein kinases, it is necessary first to identify a target substrate for the transfer reaction. Essentially this means a substrate that is quite specific—i.e., is not phosphorylated very well or at all by other protein kinases under the appropriate assay conditions (see Strategic Planning). The source of the enzyme activity can be crude cell lysate (see Support Protocol 1), a cellular fraction, an immunoprecipitate, a partially purified protein, or a purified enzyme. The results of the assay can be evaluated using trichloroacetic acid (TCA) precipitation to measure incorporation of radioactivity (see Support Protocol 2), adsorption of the labeled material onto P81 phosphocellulose paper (see Support Protocol 3), or electrophoretic separation and autoradiography (see Support Protocol 4).

STRATEGIC PLANNING

The basic strategy for protein kinase assays involves the use of a labeled donor substrate so that when phosphotransferase activity is present in the enzyme sample, accumulation of the label in the protein or peptide acceptor substrate can be easily detected. The most often used protocol requires [γ -³²P]ATP as the donor substrate and a specific protein or peptide as the acceptor substrate. Phosphotransfer is detected as the accumulation of [³²P]-labeled protein or peptide. The main problem with this assay is that it requires ³²P, a high-energy β emitter. There are a few examples of kinase assays that do not require [γ -³²P]ATP, but they are far more specialized and generally not as widely applicable. Hopefully, further development of nonradioactive kinase assays will provide good alternatives and reduce the use of radioisotopes.

Enzyme Sources

Of all the choices to be made in designing protein kinase assays, of paramount importance is the source of enzyme activity: very crude sources such as whole-cell lysates (see Support Protocol 1 and UNIT 18.2), immunoprecipitates, partially purified preparations, and homogeneous purified protein can be the source of the enzyme activity. The first requirement is that the kinase activity be relatively stable under both the conditions used to prepare the enzyme and those used for the assay. Often enzymes are unstable in whole-cell lysates or homogenates and become very stable after only an initial purification step. Protein kinases commonly exist in intact cells as phosphoproteins themselves, and the phosphorylation state is often important for activity. Thus, it is essential to attempt to preserve the phosphorylation state, e.g., by inhibiting protein phosphatases. Including specific kinase inhibitors in the lysis or homogenation buffer is a common procedure to inhibit serine/threonine kinases (sodium fluoride, β -glycerophosphate, okadaic acid) or

tyrosine kinases (sodium vanadate, phenylarsinine oxide). Other common inclusions in lysis buffer are chelating agents and protease inhibitors. Many enzymes, including kinases, are sensitive to limited proteolysis. Chelators, such as EDTA and EGTA, act to chelate calcium and reduce the activity of calcium-activated proteases. The most commonly used protease inhibitors include phenylmethylsulfonyl fluoride (PMSF), leupeptin, pepstatin A, antipain, and benzamidine. It is always a good idea to add these inhibitors to the lysis buffer just before it is used to lyse cells, especially in the early stages of a new project; the need for specific inhibitors should be investigated at a later time. The oxidation state of cysteines and the state of disulfide linkages may influence enzyme activity, so inclusion of a reducing agent—e.g., 2-mercaptoethanol (2-ME) or dithiothreitol (DTT)—is sometimes essential to preserve enzyme activity. Low concentrations of a detergent such as Brij35, at or just above the critical micelle concentration, may also stabilize enzyme activity.

Choice of Substrate

Selecting a substrate normally depends on the type of experiment and the enzyme preparation, i.e., whether there are any contaminating kinases in the enzyme sample that phosphorylate the same or similar substrates. The type of experiment may range from following a particular activity through purification to determining substrate specificity or kinetics for the kinase-catalyzed reaction. Most often specific substrate information is available from previously published work or from initial experiments to identify a suitable substrate. It is always helpful to evaluate the suitability of several substrates in parallel experiments in order to identify the best substrate.

For crude extracts, the best substrate will be one that is phosphorylated by the kinase of interest and by no other kinase. In practice this is rarely the case, so a compromise must be reached. If an antibody that recognizes the kinase is available, it is easier to use that antibody to immunoprecipitate the protein and use the immunoprecipitate as the source of enzyme activity. Immunoprecipitation is useful because it provides a quick one-step partial purification of the kinase that also concentrates the activity. With a partially purified enzyme, it may be easier to select an appropriate specific substrate.

Histones are very basic proteins found complexed with DNA in the nucleus of the cell. They have been used for many years as protein kinase substrates because of their relative abundance, ease of purification, and ability to function as very good substrates for a large number of protein kinases. Distinct fractions of mixed histones and purified histones are available commercially from a wide variety of sources. Different kinases have different specificities with respect to phosphorylation of histones, hence the type of histone preparation that should be used depends on the type of kinase to be assayed. Histones have been used to assay many protein kinases and may be used for cAMP- and cGMP-dependent protein kinases. However, phosphorylation of histone substrates is generally not enzyme specific—e.g., histone H1 is phosphorylated by cAMP-dependent protein kinases, most cyclin/cdk complexes, and some protein kinase Cs albeit on different sites, and histone 2B is phosphorylated by both cAMP- and cGMP-dependent protein kinases and protein kinase B—so more specific substrates are generally preferred.

Initially, it is much easier to use intact proteins as substrates for kinase assays. However, once a substrate is identified, specific determinant(s) within the protein that allow phosphorylation and specific primary sequence requirements can be identified, and the appropriate peptide sequences modeled on the phosphorylation site can be synthesized and used as substrate.

When synthetic peptide substrates are used to assay protein kinase activity, it is possible to study the specificity of different kinases. Synthetic peptide substrates are available from

a number of suppliers, e.g., Peninsula Labs, Bachem Biosciences, and Calbiochem-Novabiochem. Peptides are easy to use: they are potentially available in large quantities as chemically defined molecules, their purification is usually straightforward, and they provide a single phosphorylation site. In addition, it is relatively simple to synthesize a series of peptides with single amino acid changes to further define the specificity of an enzyme. However, peptides do have their disadvantages. They are expensive, \$10 to \$20 per residue. The kinase may not recognize the disordered structure usually adopted by small peptides in solution, or it may promiscuously phosphorylate the peptide. Finally, it is very difficult to use peptides to model reactions in which prior derivitization (i.e., phosphorylation) is required for specific kinase activity.

Table 18.7.1 summarizes the sequence specificity of a number of protein serine/threonine kinases, and Table 18.7.2 provides similar information for protein tyrosine kinases.

Table 18.7.1 Substrate Specificities of Some Protein Serine/Threonine Kinases

Enzyme	Consensus sequence ^{a,b}
cAMP-dependent protein kinase	XArgArgX <u>Ser</u> Φ
cGMP-dependent protein kinase	XArg/LysX <u>Ser</u> Φ
Protein kinase Cα	XLys/ArgXX <u>Ser</u> X
Cyclin B/cdc2	X <u>Ser</u> ProXLys/ArgX
Casein kinase I	Glu/Ser(P) XX <u>Ser</u>
Ca ²⁺ /calmodulin-dependent protein kinase type II	XArgXX <u>Ser</u> X
MAP kinase	XPro <u>Ser</u> XX
Casein kinase II	X <u>Ser</u> XGluGlu/AspX

^aSee Table A.1C.1 for amino acid codes. X indicates any amino acid, Φ indicates a hydrophobic residue, underlining indicates the sites of phosphorylation (which, although marked Ser in each case, may be either Ser or Thr residues), and (P) indicates a phosphorylated residue.

^bSequence information based on Pearson and Kemp (1991).

Table 18.7.2 Substrate Specificities of Some Protein Tyrosine Kinases

Enzyme	Optimal sequence ^{a,b}
MT/c-src	GluGluIle <u>Tyr</u> GlyGluPheGlu
c-abl	AlaGluValIle <u>Tyr</u> AlaAlaProPhe
c-fps/fes	GluGluGluIle <u>Tyr</u> GluGluIleGlu
lck	XGluXIle <u>Tyr</u> GlyValLeuPhe
EGF receptor	GluGluGluGlu <u>Tyr</u> PheGluLeuVal
PDGF receptor	GluGluGluGlu <u>Tyr</u> ValPheIleGlu
Insulin receptor	XGluGluGlu <u>Tyr</u> MetMetMetMet

^aThese sequences were determined through the use of peptide libraries (Songyang et al., 1995), and their relevance to in vivo phosphorylation specificities is unclear.

^bSee Table A.1C.1 for amino acid codes. X indicates any amino acid; underlining indicates the site of phosphorylation.

Assay Conditions

The assay conditions can have profound effects on any enzyme activity: if pepsin is assayed at pH 7, it is essentially inactive, but at pH 3, it has good protease activity. Similarly T4 DNA ligase is inhibited by salt concentrations >150 mM. For kinases, it is a good idea to use preliminary experiments to identify the effects of pH, salt concentration, concentrations of cations such as Mg^{2+} or Mn^{2+} , and temperature to optimize the assay conditions for the kinase of interest. It is often advantageous to perform kinase assays at 30°C rather than 37°C because the lower temperature makes it easier to stay within the linear range of the kinase, thus providing more control of the assay.

Another factor to be considered when optimizing an assay for a newly identified kinase is the concentration of ATP in the reaction mixture. In these protocols and in general, detection of kinase activity is based on the transfer of radiolabeled phosphate from ATP to the substrate, so the concentration of ATP can be varied only a little. Also, the specific activity of the $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ must be known in order to actually measure phosphotransfer. Most kinases have a K_m for ATP of 1 to 100 μM . If there is too much ATP in the reaction mixture, it will be difficult to measure phosphotransfer. An ATP concentration of 50 to 100 μM tends to work well. At that concentration, the enzyme should be working at $\geq 50\%$ of maximum, depending on its apparent K_m for ATP, and addition of sufficient $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ to measure phosphotransfer is not prohibitively expensive. Usually the substrate concentration is high so that the enzyme is working at or close to V_{\max} .

Controls

Including the correct controls for a kinase assay is critical to the success of the assay, especially when the enzyme source is a cell or tissue extract. The appropriate controls should always include a no substrate control, a no enzyme source control, a heat-denatured enzyme control, and, for enzymes that require an activator or cofactor, controls that use irrelevant activator or cofactor and controls without activator or cofactor.

CAUTION: These assays use $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ which should be handled and disposed of according to safety regulations. See *APPENDIX 1F* and the institutional Radiation Safety Office for guidelines for proper handling and disposal of ^{32}P .

ASSAY FOR CYCLIC NUCLEOTIDE-DEPENDENT PROTEIN KINASES

Cyclic adenosine monophosphate (cAMP)-and cyclic guanosine monophosphate (cGMP)-dependent protein kinases (PrKs) are quite similar in structure and require similar assay conditions. cAMP-PrK is a heterotetramer consisting of two regulatory (R) subunits and two catalytic (C) subunits, R_2C_2 . In each case the regulatory subunits have binding sites for two molecules of cAMP. When the cAMP concentration is elevated in cells after activation of receptor-linked adenylyl cyclases, cAMP binds to the R subunits, causing the affinity of the C subunits for R subunits to drop by about four orders of magnitude. The formation of $R_2\text{cAMP}_4$ dimers and free C-subunits is favored, and the C-subunits are now active.

The case of cGMP-PrK is slightly different in that there are no free catalytic subunits released. cGMP-PrK is composed of two identical subunits. Each subunit has a regulatory domain and a catalytic domain, which are homologous to the R and C subunits of cAMP-PrK. On binding four molecules of cGMP the enzyme is activated, presumably by a conformational change.

The assay procedures for these two kinases are very similar and the same protein substrate can be used for each one as detailed below. More recently, both of these kinases have been assayed using peptide substrates.

BASIC PROTOCOL 1

Assays of Protein Kinases Using Exogenous Substrates

18.7.4

Materials

5× cyclic nucleotide–dependent kinase reaction buffer (see recipe)
10 mg/ml histone 2B in H₂O
[γ-³²P]ATP solution (see recipe)
20× cyclic nucleotide solution: 20 μM cyclic AMP in H₂O/20 μM cyclic GMP in H₂O
Enzyme sample containing cyclic nucleotide–dependent kinase activity (see Support Protocol 1), kept on ice until use
30° water bath

Additional reagents and equipment for TCA precipitation (see Support Protocol 2), adsorption onto P81 phosphocellulose paper (see Support Protocol 3), or electrophoretic analysis (see Support Protocol 4)

1. For each assay reaction, add the following to a 1.5-ml microcentrifuge tube kept on ice:

4 μl 5× cyclic nucleotide–dependent kinase reaction buffer
1 μl of 10 mg/ml histone 2B
1 μl [γ-³²P]ATP solution (to give 5 μCi/μl and 5 μM ATP final)
1 μl 20× cyclic nucleotide solution
0 to 13 μl H₂O.

Cap tube and warm the mixture in a 30°C water bath.

Perform each assay in triplicate and include no substrate, no enzyme, and no cyclic nucleotide controls.

The total reaction mix volume is 20 μl per reaction. The amount of water required depends on how much enzyme sample is used in step 2.

Sufficient reaction mix can be prepared in a single tube for all the reactions by multiplying the quantities for a single reaction by the total number of reactions + 1. Then the total amount of reaction mixture (minus enzyme) for one reaction can be added to each tube. This is often very convenient and reduces the potential for pipetting errors.

2. Start the reaction by adding 1 to 14 μl ice-cold enzyme sample containing cyclic nucleotide–dependent kinase activity.

The volume of enzyme used depends on the amount of activity in the enzyme sample. In a preliminary experiment, the maximum indicated amount can be used to gauge the extent of the phosphotransfer reaction. The volume can be reduced as appropriate in order to allow for linear incorporation of phosphate into the substrate during the assay.

For enzyme samples that are immunoprecipitates adsorbed onto Sepharose beads (UNIT 10.16), prepare the reaction mix minus enzyme source first, warm it, then dispense it into the immunoprecipitate. For immunoprecipitates, it is advisable to scale up the reaction for a total volume of 100 μl: add 75 μl reaction mix to 25 μl immunoprecipitate bound to beads.

3. Incubate 10 min in a 30°C water bath.
4. Stop the reaction using the reagent appropriate for the analytic method to be used—20 μl ice-cold 10% TCA for TCA precipitation (see Support Protocol 2), or 10 μl or 20 μl ice-cold 2× SDS-PAGE sample buffer for electrophoretic analysis (see Support Protocol 4). Use 10 μl of the reaction mix for adsorption to P81 phosphocellulose paper (see Support Protocol 3). Proceed with analysis by one of those methods.

ASSAY FOR PROTEIN KINASE C ISOFORMS

There are presently at least ten members of the protein kinase C (PKC) family of kinases. Protein kinase C was originally identified as a kinase activity that was reversibly activated by calcium (Ca^{2+}), phospholipid (usually phosphatidylserine), and the neutral lipid diacylglycerol (DAG; Takai et al., 1979). Subsequently, these kinases were shown to be the sites of action for the tumor-promoting phorbol esters, which bind to some PKC isoforms and activate the kinase activity by substituting for DAG.

The initial isoforms of PKC that were purified were called α , β , and γ , and it is these forms that are the most clearly regulated by Ca^{2+} , phospholipid, and DAG or phorbol esters. They are termed the “conventional” subfamily of PKCs (cPKCs). Another subfamily contains δ , ϵ , ν , and θ , the “novel” PKCs (nPKCs); these kinases are not regulated by Ca^{2+} but are activated by phorbol esters and phospholipids. A third subfamily, the “atypical” PKCs (aPKCs), consist of ζ , λ , and τ ; these are not regulated by Ca^{2+} nor do they bind phorbol esters or DAG, but there is still a requirement for phospholipids. When the isoforms are analyzed, the reaction buffer must include the appropriate activating cofactors.

Materials

- 5 \times PKC reaction buffer (see recipe)
- 10 mg/ml histone H1 in H_2O
- $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ solution (see recipe)
- Enzyme sample containing PKC activity (see Support Protocol 1)
- 30 $^\circ\text{C}$ water bath

Additional reagents and equipment for TCA precipitation (see Support Protocol 2), adsorption onto P81 phosphocellulose paper (see Support Protocol 3), or electrophoretic analysis (see Support Protocol 4)

1. For each assay reaction, add the following to a 1.5-ml microcentrifuge tube:

- 4 μl 5 \times PKC reaction buffer
- 1 μl 10 mg/ml histone H1
- 1 μl $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ solution (to give 5 $\mu\text{Ci}/\mu\text{l}$ and 5 μM ATP final)
- 0 to 14 μl H_2O .

Cap tube and warm the mixture 10 min in a 30 $^\circ\text{C}$ water bath.

Perform each assay in triplicate and include no substrate and no enzyme controls.

The total reaction mix volume is 20 μl per reaction. The amount of water required depends on how much enzyme sample is used in step 2.

Sufficient reaction mix can be prepared in a single tube for all the reactions by multiplying the quantities for a single reaction by the total number of reactions + 1. Then the total amount of reaction mixture (minus enzyme) for one reaction can be added to each tube. This is often very convenient and reduces the potential for pipetting errors.

2. Start reaction by adding 1 to 14 μl enzyme sample containing PKC activity to the warmed reaction mixture.

The volume of enzyme used depends on the amount of activity in the enzyme sample. In a preliminary experiment, the maximum indicated amount can be used to gauge the extent of the phosphotransfer reaction. The volume can be reduced as appropriate in order to allow for linear incorporation of phosphate into the substrate during the assay.

For enzyme samples that are immunoprecipitates adsorbed onto Sepharose beads (UNIT 10.16), prepare the reaction mix minus enzyme source first, warm it, then dispense it into the immunoprecipitate. For immunoprecipitates, it is advisable to scale up the reaction for a total volume of 100 μl : add 75 μl reaction mix to 25 μl immunoprecipitate bound to beads.

3. Incubate 10 min in a 30°C water bath.
4. Stop the reaction using the reagent appropriate for the analytic method to be used—20 μ l ice-cold 10% TCA for TCA precipitation (see Support Protocol 2), 10 μ l, or 20 μ l ice-cold 2 \times SDS-PAGE sample buffer for electrophoretic analysis (see Support Protocol 4). Use 10 μ l of the reaction mix for adsorption to P81 phosphocellulose paper (see Support Protocol 3). Proceed with analysis by one of those methods.

ASSAY FOR CASEIN KINASES USING β -CASEIN

Casein is a very acidic milk protein available in reasonable purity and quantity. Casein kinases were named for their ability to phosphorylate casein rather than other substrates available at the time. Because of its acidic nature, the use of casein as a substrate presents a number of problems, one of which is the fact that it does not carry a net positive charge at low pH and thus does not bind to P81 phosphocellulose paper. In recent years this problem has been circumvented by using short acidic peptides engineered to bind to P81 paper as substrates for the casein kinases (see Alternate Protocol).

There are currently two known mammalian casein kinases, casein kinase I and casein kinase II, named for their order of elution from anion-exchange chromatography columns. Casein kinases show a marked preference for substrates containing acidic residues close to the serine or threonine to be phosphorylated.

Casein kinase I is a 37-kDa monomer that phosphorylates serine and threonine residues (see Table 18.7.1) and utilizes ATP only as phosphoryl group donor. The basic phosphorylation consensus sequence is Glu/Ser(P)XXSer (where the underlined residue is the phosphate acceptor, and may be either Ser or Thr).

Mammalian casein kinase II is a tetramer of $\alpha_2\beta_2$ structure: the α -subunit is 24 to 28 kDa and the β -subunit is 37 to 44 kDa. Its preferred substrate has aspartate or glutamate residues C-terminal to the phosphorylation site serine or threonine and a consensus of SerGluGlu/Asp. Casein kinase II is unusual in that it can utilize both ATP and GTP as phosphoryl group donor with similar K_m values, $\sim 10 \mu\text{M}$ for ATP and $\sim 30 \mu\text{M}$ for GTP. Mammalian casein kinase II is also inhibited by heparin, whereas casein kinase I is not. The IC_{50} for the inhibition of casein kinase II by heparin is in the 10 to 20 nM range and at these concentrations casein kinase I is not inhibited.

Assays of either casein kinase I or II can be performed using casein as substrate. When isolated from milk, casein is a phosphoprotein. However, it is preferable to use dephosphorylated casein as substrate; this is commercially available in a relatively pure state as either the α or β isoform.

Materials

- 5 \times casein kinase reaction buffer (see recipe)
- 10 mg/ml β -casein in H_2O
- $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ solution (see recipe)
- Enzyme sample containing casein kinase activity (see Support Protocol 1), kept on ice
- 30°C water bath
- Additional reagents and equipment for TCA precipitation (see Support Protocol 2) or electrophoretic analysis (see Support Protocol 4)

1. For each assay reaction, add the following to a 1.5-ml microcentrifuge tube kept on ice:

4 μ l 5 \times casein kinase reaction buffer
1 μ l 10 mg/ml β -casein
1 μ l [γ - 32 P]ATP solution (to give 5 μ Ci/ μ l and 5 μ M ATP final)
1 to 14 μ l H₂O.

Cap tube and warm the mixture 10 min in a 30°C water bath.

This reaction mix can also be used for α -casein.

The total reaction mix volume is 20 μ l per reaction. The amount of water required depends on how much enzyme sample is used in step 2.

Sufficient reaction mix can be prepared in a single tube for all the reactions by multiplying the quantities for a single reaction by the total number of reactions + 1. Then the total amount of reaction mixture (minus enzyme) for one reaction can be added to each tube. This is often very convenient and reduces the potential for pipetting errors.

2. Start the reaction by adding 1 to 14 μ l enzyme sample containing casein kinase activity.

The volume of enzyme used depends on the amount of activity in the enzyme sample. In a preliminary experiment, the maximum indicated amount can be used to gauge the extent of the phosphotransfer reaction. The volume can be reduced as appropriate in order to allow for linear incorporation of phosphate into the substrate during the assay.

For enzyme samples that are immunoprecipitates adsorbed onto Sepharose beads (UNIT 10.16), prepare the reaction mix minus enzyme source first, warm it, then dispense it into the immunoprecipitate. For immunoprecipitates, it is advisable to scale up the reaction for a total volume of 100 μ l: add 75 μ l reaction mix to 25 μ l immunoprecipitate bound to beads.

3. Incubate 10 min in a 30°C water bath.
4. Stop the reaction using the reagent appropriate for the analytic method to be used—20 μ l ice-cold 10% TCA for TCA precipitation (see Support Protocol 2), or 10 μ l or 20 μ l ice-cold 2 \times SDS-PAGE sample buffer for electrophoretic analysis (see Support Protocol 4). Proceed with analysis by one of those methods.

ALTERNATE PROTOCOL

ASSAY FOR CASEIN KINASES USING A PEPTIDE SUBSTRATE

Casein kinase I can be assayed with the peptide AspAspAspGluGluSerIleThrArgArg. Casein kinase II has most recently been assayed using specific peptide substrates, e.g., ArgArgArgGluGluGluThrGluGluGlu (where the underlined residue is the phosphate acceptor). In both cases the arginine residues allow binding to P81 phosphocellulose paper. The first peptide is relatively specific for casein kinase I, but the kinetics of its phosphorylation are not ideal; the K_m is in the millimolar range. Thus, it is preferable to use this assay for casein kinase I only if the kinase is highly purified.

Additional Materials (also see Basic Protocol 3)

10 mM synthetic peptide substrate solution (see recipe) for casein kinase: e.g.,
AspAspAspGluGluSerIleThrArgArg (for casein kinase I) or
ArgArgArgGluGluGluThrGluGluGlu (for casein kinase II)

Additional reagents and equipment for TCA precipitation (see Support Protocol 2), adsorption onto P81 phosphocellulose paper (see Support Protocol 3), or electrophoretic analysis (see Support Protocol 4)

1. For each assay, add the following to a microcentrifuge tube:

10 μ l 5 \times casein kinase reaction buffer
5 μ l 10 mM synthetic peptide substrate solution for casein kinase
5 μ l [γ - 32 P]ATP solution (to give 5 μ Ci/ μ l and 5 μ M ATP final)
0 to 30 μ l H₂O.

Cap the tube and warm 10 min in a 30°C water bath.

The total reaction mix volume is 50 μ l per reaction. The amount of water required depends on how much enzyme sample is based.

Sufficient reaction mix can be prepared in a single tube for all the reactions by multiplying the quantities for a single reaction by the total number of reactions + 1. Then the total amount of reaction mix minus enzyme for one reaction can be added to each tube. This is often very convenient and reduces the potential for pipetting errors.

2. Start the reaction by adding 1 to 30 μ l enzyme sample containing casein kinase activity.

The volume of enzyme used depends on the amount of activity in the enzyme sample. In a preliminary experiment, the maximum indicated amount can be used to gauge the extent of the phosphotransfer reaction. The volume can be reduced as appropriate in order to allow for linear incorporation of phosphate into the substrate during the assay.

For enzyme samples that are immunoprecipitates adsorbed onto Sepharose beads (UNIT 10.16), prepare the reaction mix minus enzyme source first, warm it, then dispense it into the immunoprecipitate. For immunoprecipitates, it is advisable to scale up the reaction for a total volume of 100 μ l: add 75 μ l reaction mix to 25 μ l immunoprecipitate bound to beads.

3. Incubate 10 min in a 30°C water bath.
4. Stop the reaction using the reagent appropriate for the analytic method to be used—20 μ l ice-cold 10% TCA for TCA precipitation (see Support Protocol 2), or 10 μ l or 20 μ l ice-cold 2 \times SDS-PAGE sample buffer for electrophoretic analysis (see Support Protocol 4). Use 10 μ l of the reaction mix for adsorption to P81 phosphocellulose paper (see Support Protocol 3). Proceed with analysis by one of those methods.

ASSAY FOR Ca²⁺/CALMODULIN-DEPENDENT KINASES

As the name implies, Ca²⁺/calmodulin-dependent kinases are dependent on Ca²⁺ and calmodulin for their activity. Ca²⁺/calmodulin-dependent kinase I (CaM kinase I) is a monomeric protein of ~40 kDa. Ca²⁺/calmodulin-dependent kinase II (CaM kinase II) is composed of the products of four distinct genes for α , β , γ , and δ subunits. In its native form, CaM kinase II is an oligomer of M_r ~500,000 to 600,000. As its name suggests, both Ca²⁺ and calmodulin are required for activity.

Little is known of the physiological function of CaM kinase I. Activity can be assayed using the protein synapsin 1 as substrate, but both CaM kinase I and CaM kinase II phosphorylate full-length synapsin 1. This problem can be overcome by using synthetic peptide substrates that include the specific phosphorylation site for the enzyme being assayed. CaM kinase I can be assayed using a peptide containing site 1 of synapsin I, TyrLeuArgArgArgLeuSerAspSerAsnPhe (where the underlined residue is the phosphate acceptor). CaM kinase II can be assayed using autocalmodulin, LysLysAlaLeuArgGlnGlu-ThrValAspAlaLeu (Hanson et al., 1989), which is modeled on the autophosphorylation site of the kinase itself.

BASIC PROTOCOL 4

Analysis of Protein Phosphorylation

18.7.9

Materials

- 5× CaM kinase reaction buffer (see recipe)
- 10 mM synthetic peptide substrate solution (see recipe): e.g.,
 - TyrLeuArgArgArgLeuSerAspSerAsnPhe (for CaM kinase I) or
 - LysLysAlaLeuArgGlnGluThrValAspAlaLeu (autocamtide for CaM kinase II)
- [γ -³²P]ATP solution (see recipe)
- 1 mg/ml calmodulin in Milli-Q-purified water (store small aliquots at -70°C)
- Enzyme sample containing CaM kinase activity (see Support Protocol 1), kept on ice
- 30°C water bath
- Additional reagents and equipment for adsorption onto P81 phosphocellulose paper (see Support Protocol 3)

1. Set up and label duplicate 1.5-ml microcentrifuge tubes. For each reaction, prepare a mixture containing:

- 5 μ l 5× CaM kinase reaction buffer
- 5 μ l 10 mM synthetic peptide substrate solution
- 5 μ l [γ -³²P]ATP solution (to give 5 μ Ci/ μ l and 5 μ M ATP final)
- 5 μ l 1 mg/ml calmodulin *or* 5 μ l H₂O
- 0 to 30 μ l H₂O.

Cap tube and warm the mixture 10 min in a 30°C water bath.

The reaction mix is described for 50 μ l per reaction, but it can be carried out in volumes of 25 to 100 μ l by adjusting the amounts of the ingredients proportionally.

Control reactions should include one reaction without calmodulin so that any background phosphotransfer reactions can be accounted for and only calcium²⁺/calmodulin-dependent phosphotransfer is measured.

2. Start each set of reactions sequentially every 30 sec by adding 1 to 25 μ l enzyme sample containing CaM kinase activity to one of the tubes.

The volume of enzyme used depends on the amount of activity in the enzyme sample. In a preliminary experiment, the maximum indicated amount can be used to gauge the extent of the phosphotransfer reaction. The volume can be reduced as appropriate in order to allow for linear incorporation of phosphate into the substrate during the assay.

For enzyme samples that are immunoprecipitates adsorbed onto Sepharose beads (UNIT 10.16), prepare the reaction mix minus enzyme source first, warm it, then dispense it into the immunoprecipitate. For immunoprecipitates, it is advisable to scale up the reaction for a total volume of 100 μ l: add 75 μ l reaction mixture to 25 μ l immunoprecipitate bound to beads.

3. Incubate for 10 min in a 30°C water bath.
4. Once the desired incubation time has elapsed, pipet 30 μ l of the contents of the tube quickly onto the surface of prepared P81 paper squares (see Support Protocol 3).

Only adsorption onto P81 phosphocellulose (see Support Protocol 3) is used to measure CaM kinase-mediated transfer because peptides are not efficiently precipitated by TCA, and they require specialized techniques for analysis by PAGE.

ASSAY FOR TYROSINE KINASES

Acid-denatured rabbit muscle enolase was first used as an exogenous tyrosine kinase substrate in the early 1980s, and it is still a useful substrate in a number of experimental situations. Usually enolase is used as a substrate for cytosolic tyrosine kinases such as pp60^{c-src}. Once the reaction is complete, phosphorylation of enolase is detected by separating proteins in the reaction mixture by SDS-PAGE (UNIT 10.2).

The following method is optimized for pp60^{c-src}, but may be applied to other tyrosine kinases. It is also useful as a framework for the assay of any tyrosine kinase, but optimization of the immunoprecipitation and assay buffers is required for the tyrosine kinase of interest.

Materials

- Rabbit muscle enolase (enolase EC 4.2.1.11; ammonium sulfate precipitate or suspension purified from rabbit muscle)
- 1 mM DTT/50 mM HEPES, pH 7.0
- Glycerol
- 100 mM acetic acid
- 5× tyrosine kinase reaction buffer (see recipe)
- [γ -³²P]ATP solution (see recipe)
- Enzyme sample containing tyrosine kinase activity (see Support Protocol 1), kept on ice
- 2× SDS-PAGE sample buffer (UNIT 10.2), ice-cold
- 30°C and boiling water bath

- Additional reagents and equipment for immunoprecipitation (UNIT 10.16) or SDS-PAGE (see Support Protocol 4 and UNIT 10.2)

1. Microcentrifuge the equivalent of 100 μ g rabbit muscle enolase 5 min at maximum speed, 4°C. Discard the supernatant.
2. Add 10 μ l of 1 mM DTT/50 mM HEPES, pH 7.0, to the pellet, mix thoroughly, and incubate 30 to 60 min on ice.
3. Add 10 μ l glycerol, mix, and then keep on ice if assays are to be performed the same day.

This preparation can also be stored at -70°C until required.

4. Immediately prior to assay, add 20 μ l of 100 mM acetic acid to the enolase solution and mix thoroughly. Incubate this mixture 5 min in a 30°C water bath, and then keep on ice until the reaction mixture is prepared.

Do not store acid-denatured enolase on ice in this state for >1 hr.

5. For each reaction, add the following to a 1.5-ml microcentrifuge tube:

- 4 μ l 5× tyrosine kinase reaction buffer
- 1 μ l acid-denatured enolase (step 4; 2.5 μ g enolase total)
- 1 μ l [γ -³²P]ATP (to give 5 μ Ci/ μ l and 5 μ M ATP final)
- 0 to 14 μ l H₂O.

Cap the tube and warm the mixture 10 min in a 30°C water bath.

The amount of water required depends on the amount of enzyme sample to be added; the total reaction volume should be 20 μ l.

Sufficient reaction mix can be prepared in a single tube for all the reactions by multiplying the quantities for a single reaction by the total number of reactions + 1. Then the total amount of reaction mix minus enzyme for one reaction can be added to each tube. This is often very convenient and reduces the potential for pipetting errors.

6. Start the reactions by adding 1 to 14 μl enzyme sample containing tyrosine kinase activity.

The volume of enzyme used depends on the amount of activity in the enzyme sample. In a preliminary experiment, the maximum indicated amount can be used to gauge the extent of the phosphotransfer reaction. The volume can be reduced as appropriate in order to allow for linear incorporation of phosphate into the substrate during the assay.

For enzyme samples that are immunoprecipitates adsorbed onto Sepharose beads (UNIT 10.16), prepare the reaction mix minus enzyme source first, warm it, then dispense it into the immunoprecipitate. For immunoprecipitates, it is advisable to scale up the reaction for a total volume of 100 μl : add 75 μl reaction mixture to 25 μl immunoprecipitate bound to beads.

7. Incubate the tube 10 min in a 30°C water bath.
8. Stop the reaction by adding 20 μl ice-cold 2 \times SDS-PAGE sample buffer. Mix thoroughly and heat the tube 3 min in a boiling water bath. Analyze 20 μl of the reaction by SDS-PAGE (UNIT 10.2).

Most tyrosine kinases autophosphorylate on a number of tyrosine residues and in in vitro assays are promiscuous with respect to substrate. Backgrounds are often very high so TCA precipitation and adsorption to P81 phosphocellulose paper are not very useful for analysis of results. SDS-PAGE (see Support Protocol 4) gives more specific and cleaner results.

BASIC PROTOCOL 6

IN-GEL PROTEIN KINASE ASSAYS

In-gel protein kinase assays (Kameshita and Fujisawa, 1989) involve the copolymerization of a kinase substrate in the separating gel layer of an SDS-PAGE gel. Protein samples are run on the modified gel as usual (UNIT 10.2). After electrophoresis is complete, the separated proteins are allowed to refold by various gel treatments. The gel is then incubated with [γ - ^{32}P]ATP at 30°C for enough time to allow the ATP to penetrate the entire gel. After further washes to remove unincorporated [γ - ^{32}P]ATP, the gel is stained and fixed as normal. ^{32}P -containing proteins are revealed by autoradiography (APPENDIX 3A). Although in general these assays are not particularly sensitive, they are sometimes very useful for identifying kinases in complex mixtures.

Materials

- 10 mg/ml kinase substrate: e.g., myelin basic protein
- Enzyme sample containing kinase activity (see Support Protocol 1), kept on ice
- 20% (v/v) 2-propanol/50 mM Tris·Cl (pH 8.0 at room temperature; APPENDIX 2)
- 1 mM DTT/50 mM Tris·Cl (pH 8.0 at room temperature)
- 6 M guanidine-HCl/1 mM DTT/50 mM Tris·Cl (pH 8.0 at room temperature) *or* 8 M urea/1 mM DTT/50 mM Tris·Cl (pH 8.0 at room temperature)
- 1 mM DTT/0.05% (v/v) Tween 20/50 mM Tris·Cl (pH 8.0 at 4°C)
- Appropriate kinase reaction buffer (see recipes)
- 10 mM Mg/ATP solution (see recipe)
- 10 mCi/ml [γ - ^{32}P]ATP (3000 Ci/mmol; Amersham, DuPont NEN, or ICN Biomedicals)
- 5% (w/v) trichloroacetic acid (TCA)
- 1% (w/v) sodium pyrophosphate/5% (w/v) TCA
- Container for radioactive incubation: e.g., small tray with tight cover, or heat-sealable polyethylene bag (Seal-a-Meal)
- Seal-a-Meal apparatus (optional)

Additional reagents and equipment for SDS-PAGE (UNIT 10.2) and autoradiography (APPENDIX 3A)

CAUTION: Because the samples are radioactive, great care should be exercised in sample preparation and loading. Perform the reactions and subsequent manipulations in screw-cap microcentrifuge tubes to minimize the risk of ^{32}P contamination. Run gels until the dye front passes out of the gel, as all residual $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ will migrate with the dye front. This means that the SDS-PAGE running buffer will be contaminated with ^{32}P , and it should be disposed of as radioactive waste according to safety regulations. See *APPENDIX 1F* for proper handling and disposal.

Separate proteins on modified gel

1. Prepare the mixture for a polyacrylamide gel of the appropriate %T (see *UNIT 10.2*). Add 10 mg/ml kinase substrate to give a final concentration of 1 mg/ml in the gel mixture. Polymerize as usual in a cassette prepared for casting the gel.

Myelin basic protein has been used with some success for assays of the MAP kinase family of enzymes. However, any protein is suitable if it is a substrate for the kinase of interest. This assay can also be used to identify potential substrates.

2. Prepare the enzyme samples containing kinase activity for SDS-PAGE in the usual way and carry out electrophoresis (see *UNIT 10.2*).
3. Once separation is complete, place the gel in 200 ml of 20% 2-propanol/50 mM Tris·Cl, pH 8.0. Wash the gel for 20 min, change to fresh buffer, and repeat twice for a total of three 20-min washes.

Washing with propanol removes SDS from the gel.

4. Wash the gel with 200 ml of 1 mM DTT/50 mM Tris·Cl (pH 8.0) three times, 20 min each.
5. Wash the gel with 250 ml of either 6 M guanidine·HCl/1 mM DTT/50 mM Tris·Cl (pH 8.0), or 8 M urea/1 mM DTT/50 mM Tris·Cl (pH 8.0) twice, 30 min each.

Either guanidine-HCl or urea can be used to denature the proteins. Depending on the kinase, one or the other may be more successful; guanidine-HCl should be tried first, because it seems to work best for most kinases.

6. Wash the gel eight to ten times at 4°C over a period of 18 hr with 250 ml of 1 mM DTT/0.05% Tween 20/50 mM Tris·Cl (pH 8.0), to renature proteins.

Assay for kinase activity

7. Incubate the gel with 250 ml of an appropriate kinase reaction buffer containing 10 mM MgCl_2 , for 20 min at 30°C.
8. Remove all traces of buffer and place the gel in a small tray or a Seal-a-Meal bag (the container is to be used for radioactive incubation). Add as small an amount of kinase reaction buffer as possible to just cover the gel. Add 1/4 of the reaction buffer volume of 10 mM Mg/ATP solution to give a final concentration of 50 μM ATP. Add 20 $\mu\text{Ci/ml}$ $[\gamma\text{-}^{32}\text{P}]\text{ATP}$.
9. Incubate 1 hr at 30°C.

The duration of this incubation can be adjusted after viewing the results of the initial experiment.

10. Incubate the gel in 250 ml 5% TCA twice, 15 min each. Wash 10 min in 500 ml of 1% sodium pyrophosphate/5% TCA. Repeat this wash until the solution contains little or no radioactivity.
11. Dry the gel onto filter paper and autoradiograph (*APPENDIX 3A*).

PREPARING A CELL LYSATE FOR KINASE ASSAYS

This protocol describes a method for detergent-induced cell lysis to prepare a crude extract containing kinase enzyme activity. Other methods of cell lysis may be appropriate (see *UNIT 12.1*). Hypotonic lysis and isolation of P100 and S100 fractions can also provide useful data on the recovery of a protein kinase as a soluble or membrane-bound activity.

Materials

Cultured cells: adherent cells at ~70% confluence in 100-mm tissue culture dishes or suspension cells at 10^6 cells/ml
PBS (*APPENDIX 2*), ice-cold
Lysis buffer (see recipe)
Protease inhibitor stock solutions (see recipe)
Microcentrifuge, 4°C

1. Wash the cultured cells twice in ice-cold PBS. Completely aspirate the final wash solution.
2. Add 0.75 ml lysis buffer to $\sim 2 \times 10^7$ cells; for adherent cells check that the monolayer is completely covered. Add protease inhibitors from stock solutions to the appropriate final concentration: 1 mM PMSF, 100 μ M benzamidine, 5 μ g/ml leupeptin, 5 μ g/ml pepstatin A, and 5 μ g/ml antipain.

A 0.75- μ l lysate should provide sufficient material for immunoprecipitation (use 25% to 50% of the lysate for an initial experiment and scale up or down depending on the results) or simple purification. If the purification procedure involves more than one chromatography step, prepare a larger lysate because only 10% to 20% of the activity is recovered from each chromatography step.

3. Incubate the cells 10 min at 4°C. Scrape the cells from the dish and transfer the lysate to a labeled microcentrifuge tube.
4. Microcentrifuge the lysate 10 min at maximum speed, 4°C. Carefully remove the supernatant and transfer it to a clean microcentrifuge tube.

Lysates should be used immediately until proper storage conditions (e.g., -70°C, liquid nitrogen) can be determined experimentally. For some kinases, it may not be possible to store the lysate at all.

TCA PRECIPITATION TO DETERMINE INCORPORATION OF RADIOACTIVITY

One of the classical methods for separating a reaction product from the reactants is differential precipitation. In the case of protein kinase assays using a protein substrate and [γ - 32 P]ATP, it is very easy to precipitate the protein and leave the [γ - 32 P]ATP in the soluble fraction. Most proteins are quantitatively precipitated by trichloroacetic acid (TCA), so TCA can be used to precipitate proteins phosphorylated during a kinase assay. Protein precipitate can be captured by filtration and the filter can be washed with TCA. TCA precipitation is a quick and reproducible way to determine the extent of [32 P]phosphate transfer to protein substrates. However, peptides are not precipitated by TCA, so adsorption of labeled peptides to P81 phosphocellulose paper (see Support Protocol 3) is a more suitable method for analyzing the results of kinase assays that use a peptide as substrate.

Materials

Assay samples (see Basic Protocols 1 to 5 and Alternate Protocol)
5% and 10% (w/v) trichloroacetic acid (TCA), ice-cold
95% ethanol
Diethyl ether

Whatman GF-C glass-fiber filters
Vacuum manifold (e.g., Fisher)
20-ml scintillation vials
Scintillation counter

1. Stop the reaction by adding 20 μ l ice-cold 10% TCA. Mix thoroughly and incubate 10 min on ice to precipitate protein.
2. Pipet sample onto Whatman GF-C glass-fiber filter held in a vacuum manifold. Allow the sample to pass through the filter. Rinse the tube with 500 μ l ice-cold 5% TCA. Add this wash to the appropriate filter.
3. Wash each filter with 5 ml ice-cold 5% TCA solution four times. Wash once with 10 ml of 95% ethanol, then three times with 10 ml diethyl ether. Allow each filter to dry for a few minutes.

CAUTION: Diethyl ether is extremely flammable and washes should be performed in a fume hood away from any flame or heat source. Ether washes should be disposed of in an appropriate manner.

4. Place the dry filters in a 20-ml scintillation vial and count in a scintillation counter by detection of Cerenkov radiation.

Alternatively, if dpm quantitation is desired, place dry filter in 10 ml scintillation fluid and count.

The filter can also be counted by Cerenkov radiation after the washes with 95% ethanol. If the filters are to be counted in scintillant, however, the diethyl ether washes are essential, because TCA (which is removed by the ether) causes significant chemiluminescence when placed in scintillation fluid, making the counts produced by the scintillation counter meaningless.

ADSORPTION ONTO P81 PHOSPHOCELLULOSE PAPER

One of the main methods to separate [32 P]-labeled proteins or peptides from [γ - 32 P]ATP after a protein kinase assay reaction is by adsorbing the protein or peptide to P81 phosphocellulose paper. P81 paper is an ion-exchange matrix with net negative charge at most pHs. At low pH (such as in 75 mM orthophosphoric acid, which is used to wash the paper in this protocol), the excess [γ - 32 P]ATP left after a kinase assay will not bind. Under the same conditions, however, the phosphorylated peptide will bind to the P81 paper if it carries a net positive charge at low pH. In practice, this is usually achieved by tagging a peptide with two or three arginine or lysine residues at the N- or C-terminus. At the pH of 75 mM orthophosphoric acid the arginyl or lysyl side chains are positively charged and will bind to P81 paper. It is necessary, of course, that the activity of the kinase in question is not affected by the addition of basic residues to the substrate. Usually five or more residues are placed between the hydroxy-amino acid phosphorylation site and the basic residues required for binding to P81 paper.

When proteins are used as substrates for kinases, the net charge of the protein is also a consideration for adsorption to P81 paper. As for peptides, adsorption to P81 paper is more suitable for basic proteins than for acidic proteins. Under the conditions described in this protocol, histones bind very well to P81 paper because they are highly basic, but casein binds very poorly or not at all because it is a very acidic protein.

SUPPORT PROTOCOL 3

Analysis of Protein Phosphorylation

18.7.15

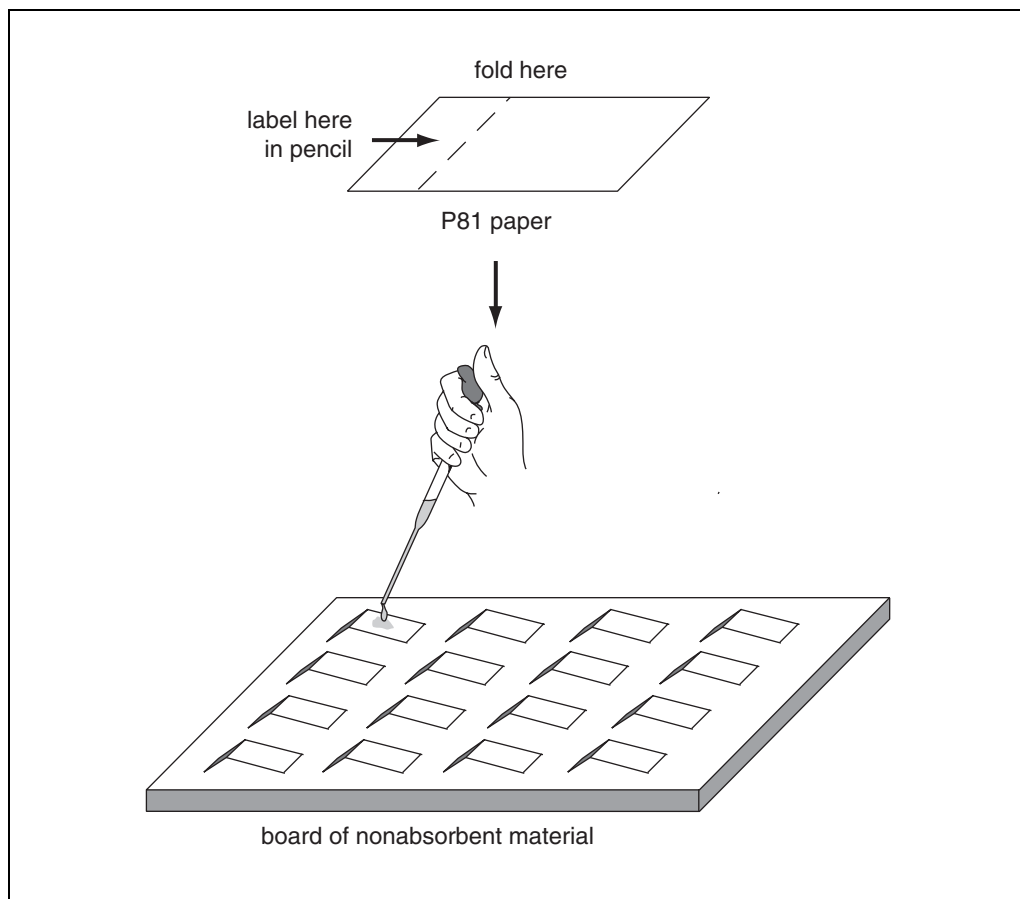


Figure 18.7.1 Addition of sample to P81 phosphocellulose paper. The 2×2 -cm squares of P81 paper are labeled using a pencil (ink is dissolved by the acetone wash), and 50% to 75% of the reaction product is applied to the paper.

Materials

Assay samples (see Basic Protocols 1 to 5 and Alternate Protocol)
 75 mM orthophosphoric acid (stored up to 6 months at room temperature)
 Acetone
 2×2 -cm squares of P81 phosphocellulose paper (Whatman)
 500-ml plastic beaker with the bottom replaced with solvent-resistant plastic or wire mesh (see Fig. 18.7.2)
 20-ml scintillation vial
 Scintillation counter

1. Cut 2×2 -cm squares of P81 paper, fold the end over, and label the folded end with pencil as shown in Figure 18.7.1. Place each square on a large piece of nonabsorbent material, such as plastic or acrylic sheeting or an aluminum foil-covered fiberboard.

Use a soft pencil to label the squares of P81 paper, as the final acetone wash will remove or destroy ink labels.

2. Pipet 15 or 30 μ l of the contents of each assay tube quickly onto the surface of the prepared P81 paper squares. Immediately place the paper squares in a 500-ml plastic beaker with the bottom replaced with solvent-resistant plastic or wire mesh (Fig. 18.7.2).

Usually 50% to 75% of the final reaction mixture is spotted onto P81 paper to ensure that there is sufficient signal to be detected.

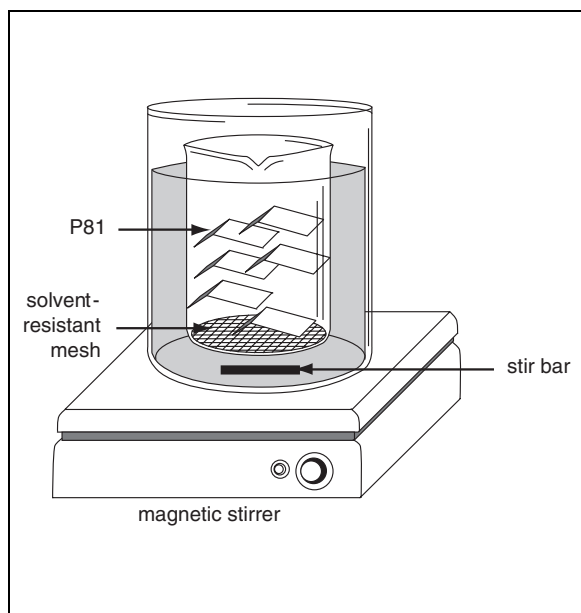


Figure 18.7.2 Setup for washing P81 phosphocellulose paper with reaction samples. The paper squares are washed five times for 5 min each in 75 mM orthophosphoric acid and once in acetone. They are then air dried and counted in a scintillation counter.

If the plastic mesh used in the beaker is not solvent-resistant, the final acetone washes will dissolve it!

3. Place this beaker into a larger beaker containing 75 mM orthophosphoric acid and incubate 5 min with stirring.
4. Pour the used orthophosphoric acid wash into a container suitable for ^{32}P -containing liquid waste. Refill the large beaker with 75 mM orthophosphoric acid and incubate 5 min more. Repeat for a total of five 5-min washes.
5. Fill the large beaker with acetone and wash the P81 papers for 5 min. Discard the acetone and allow the papers to dry.
6. Place each square of P81 paper in a 20-ml scintillation vial and count by detection of Cerenkov radiation.

Alternatively, if the absolute dpm value is desired and a ^{32}P quenching curve is available for the scintillation counter, add scintillation fluid to the vials and count.

ELECTROPHORETIC ANALYSIS OF PHOSPHORYLATION

Electrophoretic analysis of phosphorylation allows a more specific determination of protein substrate phosphorylation, because other phosphoproteins are separated by electrophoresis. It is advantageous to use the resolving power of SDS-PAGE to assay relatively crude enzyme samples; this method of analysis may also give useful information on kinase autophosphorylation.

Stop the reaction (see Basic Protocols 1 to 5 or Alternate Protocol) by adding 20 μl ice-cold $2\times$ SDS-PAGE sample buffer (UNIT 10.2). Mix thoroughly, heat the sample 3 min in a boiling water bath, then analyze by SDS-PAGE (UNIT 10.2). Stain, fix, dry, and autoradiograph (APPENDIX 3A) the gel.

SUPPORT PROTOCOL 4

Analysis of Protein Phosphorylation

18.7.17

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

[γ -³²P]ATP solution

Dilute 10 mM Mg/ATP solution (see recipe) 1/10 in water to give a 1 mM solution. For the assay, add 10 μ l of 1 mM Mg/ATP to 80 μ l water and then add 10 μ l of 10 mCi/ml [γ -³²P]ATP (3.3 pmol ATP).

The final ATP concentration is \sim 100 μ M (plus 33 nM from the [γ -³²P]ATP, which is negligible), and the 100 μ Ci (\sim 2.2 \times 10⁸ dpm) in 100 μ l gives a specific activity of 1 μ Ci/ μ l or \sim 2.2 \times 10⁷ dpm/nmol.

Suppliers for [γ -³²P]ATP include Amersham, NEN, and ICN; it is usually available at a variety of specific activities. A good specific activity to use is 3000 Ci/mmol, normally provided by the manufacturer at a concentration of 10 mCi/ml. Amersham's Redivue is a convenient [γ -³²P]ATP solution to use as it is stored at 4°C instead of -20° C.

As 1 Ci is \cong 2.2 \times 10¹² dpm, then for a specific activity of 3000 Ci/mmol

$$\begin{aligned} &= 3 \text{ Ci}/\mu\text{mol} \\ &= 6.6 \times 10^{12} \text{ dpm}/\mu\text{mol} \\ &= 6.6 \times 10^6 \text{ dpm}/\text{pmol} \\ &= 2.2 \times 10^6 \text{ dpm} \cong 1 \mu\text{Ci} \cong 0.33 \text{ pmol}. \end{aligned}$$

The stock Mg/ATP solution prepared earlier can be diluted to a working ATP solution of the desired concentration (in this example, 0.1 mM).

CaM kinase reaction buffer, 5 \times

100 mM Tris·Cl, pH 7.5 (**APPENDIX 2**)
25 mM MgCl₂
1 mM CaCl₂
Store up to 6 months at -20° C

Casein kinase reaction buffer, 5 \times

100 mM Tris·Cl, pH 7.5 at 30°C (**APPENDIX 2**)
25 mM MgCl₂
2.5 mM DTT
750 mM KCl
Store up to 6 months at -20° C

Cyclic nucleotide-dependent protein kinase reaction buffer, 5 \times

250 mM Tris·Cl, pH 7.5 at 30°C (**APPENDIX 2**)
25 mM MgCl₂
Store up to 6 months at -20° C

Lysis buffer

50 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid), pH 7.4
100 mM NaCl
50 mM sodium fluoride
5 mM β -glycerophosphate
2 mM EDTA
2 mM EGTA
1 mM sodium vanadate (Na₃VO₄)
1% (v/v) Nonidet P-40 (NP-40) or Triton X-100
Store up to 6 months at -20° C

To stabilize kinase activity, specific kinase inhibitors, protease inhibitors, and reducing agents may be added to the lysis buffer (see *Strategic Planning, Enzyme Sources*).

Mg/ATP solution, 10 mM

Stock solutions:

20 mM MgCl₂ or MgSO₄

20 mM Na₂ATP

Working solution: Mix equal volumes of the two stock solutions and measure the pH. Adjust the pH slowly to pH 7.4 with 100 mM HCl or 100 mM NaOH. Store up to 1 year in small aliquots at –20°C.

If a precipitate forms, the solution should be stirred continually and the pH maintained at 7.4.

Adenosine triphosphate is normally purchased as a sodium salt. Most kinases require ATP to be complexed with magnesium for its efficient utilization as a substrate. In some cases manganese can substitute for Mg²⁺.

PKC reaction buffer, 5×

Solvent:

100 mM Tris·Cl, pH 7.5 at 30°C (APPENDIX 2)

25 mM MgCl₂

1 mM CaCl₂

Store up to 6 months at –20°C

Lipids:

Just before the buffer is to be used, transfer enough 5 µg/ml phosphatidylserine stock solution in chloroform to provide 100 µg phosphatidylserine/ml buffer (final concentration) to a clean tube. Add sufficient 0.5 µg/ml Diolein stock solution (Sigma) in chloroform to provide 10 µg Diolein/ml buffer (final concentration) to the same tube and mix. Dry the lipids under a stream of nitrogen.

Working solution:

Add the required volume solvent (see above) and place on ice for 10 min. Sonicate the mixture at medium power for 1 min while the sample is on ice. Let the sample sit on ice 2 min. Repeat sonication 4 more times for a total of 5 min. Discard any unused buffer.

Diolein is the old name for various preparations that contain mixed isomers of diacylglycerol, the cofactor for PKC. Diolein is a suitable source for the cofactor in these assays, and it is less expensive than pure preparations of 1,2-diacylglycerol.

Protease inhibitor stock solutions

100 mM phenylmethylsulfonyl fluoride (PMSF) in 100% ethanol

100 mM benzamidine

1 mg/ml leupeptin

1 mg/ml pepstatin A

1 mg/ml antipain

Store up to 6 months at –20°C

Synthetic peptide substrate solution

Prepare 10 mM solutions of each peptide (highest quality available) in Milli-Q water and adjust pH to 7.4 before storage at –70°C or use.

If the peptide is insoluble, check the sequence—a peptide with a high proportion of acidic residues may be more soluble at alkaline pH, and one with a high proportion of basic residues may be more soluble at acidic pH. Hydrophobic peptides may be solubilized by adding up to 20% (v/v) acetonitrile. Whatever is added must not compromise the assay conditions: e.g., if a peptide must be dissolved at pH 10, adding the solution to the reaction mix must not change the pH of the reaction.

Most synthetic peptides are quite stable when stored at –70°C; the stability of individual peptide substrates should be determined for the storage conditions used, however. If peptides are custom synthesized they should ideally be purified by HPLC before use.

Tyrosine kinase reaction buffer, 5×

100 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid), pH 7.4

5 mM MnCl₂

5 mM DTT

500 μM sodium vanadate (Na₃VO₄)

Store up to 6 months at -20°C

COMMENTARY

Background Information

In recent years there has been an explosion of interest in protein kinases. It is now obvious that this class of enzyme plays many essential roles in a large array of biological processes, including the cellular responses to hormones, control of most stages of the cell cycle, and development of the nervous system in systems as diverse as fruit fly and man. Because of this interest, a very large number of protein kinase and protein phosphatase genes have been identified; current estimates are that there are well over 2000 genes encoding these enzymes in mammalian genomes. Obviously the study of protein kinases and phosphatases can provide fundamental information regarding the regulation of many diverse systems; the protocols provided in this unit include enough information to enable researchers new to this type of experiment to embark on the study of protein kinases.

Protein kinases can be assayed using two types of substrate, intact proteins or synthetic peptides, and there are advantages and disadvantages to the use of each. A number of intact proteins may be used as general protein kinase substrates, and those in common use are available in relatively pure form from a number of manufacturers at reasonable prices. The main disadvantage to the use of proteins as substrates is the lack of specificity, sometimes making interpretation of experimental data difficult. Intact proteins such as myelin basic protein (MBP) contain many different phosphorylation sites. MBP can be phosphorylated by MAP kinase, some protein kinase C isoforms, and Ca²⁺/calmodulin kinase II, each at one or more sites. Hence, it is often impossible to differentiate which kinase is responsible for phosphorylating MBP in a complex mixture like a whole cell lysate. Some proteins cause headaches because they cannot be used with simple experimental protocols, e.g., casein.

The use of synthetic peptides improves the specificity of the reaction by providing a single, well-defined phosphorylation site. They also allow experiments to define the consensus se-

quence for phosphorylation by a particular protein kinase and make it possible to design a substrate that has excellent kinetic properties for the kinase of interest. Peptides are often easier to use than their protein counterparts as in the case of casein kinases, where peptides can be designed that allow the use of phosphocellulose paper rather than other more involved methods for assessment of experimental results. Use of peptides is not without disadvantages though, the most significant being the cost of producing reasonable quantities of large numbers of peptides. When this is coupled with the observation that sometimes peptides are not recognized by the kinase because some structural element is missing, using peptide substrates can be a frustrating and expensive business. However, this is the exception rather than the rule, and the use of specific peptides to assay protein kinases is ultimately the best choice.

One unit of kinase activity is usually defined as the amount of enzyme required to transfer 1 μmol of phosphate to a substrate from ATP per minute. From the following example it is possible to calculate a number of the reaction parameters:



For this example, 1 mg protein (based on protein assay; see *UNIT 10.1A*) was used for the assay. The [γ-³²P]ATP in the assay mixture had a specific activity of 10⁶ dpm/nmol. After the 10-min reaction the ³²P-labeled product had a specific activity of 10⁵ dpm; therefore, there were (10⁵/10⁶) nmol of ³²P-labeled product, and the concentration was 10⁻¹ nmol or 100 pmol. Hence, the enzyme transferred 10 pmol of phosphate per minute, and represents 0.01 U of enzyme activity when measured with this substrate. Put another way, this represents an enzyme specific activity of 10 pmol/min/mg of protein. This latter expression of enzyme activity is a useful measure when comparing different preparations of a kinase or, during purification of a kinase, as a measure of the extent of purification through different steps.

Critical Parameters

Although it may seem obvious, the most critical element in performing kinase assays is using the correct controls. This is especially true when performing assays on tissue samples or cell extracts. Controls should always include addition of no substrate, addition of no enzyme source, addition of heat-denatured enzyme source, and, for reactions that require activator or cofactor, addition of any activator or cofactor and addition of no activator or cofactor.

Another important aspect is the concentration of ATP in the assay. The ATP concentration can vary depending on the type of experiment to be performed. For derivation of kinetic data, the concentrations of both ATP and peptide or protein substrate are critical. If kinase detection during purification is required, assay linearity is not as important as long as the right controls are present. For mapping *in vitro* phosphorylation sites on a kinase substrate, very often the ATP concentration is kept extremely low but the specific activity very high.

When designing potential peptide substrates for individual protein kinases, there are a number of considerations to remember. For example, phosphorylated peptide will bind to the P81 phosphocellulose paper. To be able to analyze the assay with this paper, it is necessary that the peptide carry a net positive charge at low pH. In practice this is usually achieved by tagging a peptide with two or three arginine or lysine residues at the N- or C-terminus. At the pH of 75 mM orthophosphoric acid, the arginyl or lysyl side chains are positively charged and will bind to the P81 paper. It is important to determine that the activity of the kinase of interest is relatively unaffected by the addition of basic residues to its peptide substrates in this way. To help ensure this is the case, it is a good idea to place five or more residues between the hydroxy-amino acid to be phosphorylated and the basic residues required for binding to P81 paper.

Net charge considerations also pertain to some extent to proteins used as substrates for protein kinases. Under the conditions described in Support Protocol 3, histones will bind very well to P81 paper, because they are highly basic, but casein, which is very acidic, binds very poorly or not at all. In these cases, the P81 paper assay is very good when histones are used to assay kinase activity, but worthless when casein is used.

When crude enzyme sources, such as cell lysates, are analyzed, many other proteins are present and phosphorylation of multiple proteins in the crude mixture is possible, as is

autophosphorylation of the protein kinase. Because many proteins will be precipitated by trichloroacetic acid (TCA) or will bind to P81 paper, there can be elevated backgrounds and severe overestimates of peptide phosphorylation when Support Protocols 2 and 3 are used. When performing kinetic studies, this overestimation of phosphorylation can be very misleading, as substrates other than the peptide under study will undoubtedly be phosphorylated and with different kinetics.

Troubleshooting

Any troubleshooting of kinase assay procedures should be very straightforward. If the assay does not work, verify that both the [γ - 32 P]ATP and the peptide or protein substrate have been added at the correct concentrations. It is also very important to keep the enzyme source on ice before use and to store it appropriately between experiments. Very often when this type of experiment does not work it is because of a labile enzyme, poor storage of the enzyme, or a combination. It is important to know how long a particular enzyme is stable under the conditions used and the rate of inactivation or denaturation, if relevant. It is also essential to know if the enzyme activity is comparable between different preparations, especially for purified proteins.

Anticipated Results

Depending on the requirements, information can be obtained about the specific activity of a particular enzyme preparation, kinetic parameters for the reaction catalyzed, or whether a particular kinase has been identified or substantially purified.

Time Considerations

Kinase assays are quite simple to set up and perform. If the samples for assay are available, it is a matter of only ~3 to 4 hr to go from assay set up to data interpretation when either TCA precipitation (Support Protocol 2) or adsorption to phosphocellulose paper (Support Protocol 3) is used to separate phosphorylated protein from contaminating [γ - 32 P]ATP. With SDS-PAGE analysis (Support Protocol 4), the assays can be set up, performed, and the samples run on a SDS-PAGE gel in ~4 to 5 hr. Gel drying and autoradiography can take from 2 hr to a number of days depending on the level of protein phosphorylation. In general, if individual protein bands have ~20,000 cpm of 32 P associated with them, they will be visible with 1 hr of autoradiography.

Literature Cited

- Hanson, P.I., Kapiloff, M.S., Lou, L.L., Rosenfeld, M.G., and Shulman, H. 1989. Expression of a multifunctional Ca^{2+} /calmodulin-dependent protein kinase and mutational analysis of its autoregulation. *Neuron* 3:59-70.
- Kameshita, I. and Fujisawa, H. 1989. A sensitive method for detection of calmodulin-dependent protein kinase II activity in sodium dodecyl sulfate polyacrylamide gel. *Anal. Biochem.* 183:139-143.
- Pearson, R.B., and Kemp, B.E. 1991. Protein kinase phosphorylation site sequences and consensus specificity motifs. *Methods Enzymol.* 200:62-81.
- Songyang, Z., Carraway, K.L. III, Eck, M.J., Harrison, S.C., Feldman, R.A., Mohammadi, M., Schlessinger, J., Hubbard, S.R., Smith, D.P., Eng, C., Lorenzo, M.J., Ponder, B.A.J., Mayer, B.J., and Cantley, L.C. 1995. Catalytic specificity of protein tyrosine kinases is critical for selective signalling. *Nature* 373:536-539.
- Takai, Y., Kishimoto, A., Kikkawa, U., Mori, T., and Nishizuka, Y. 1979. Unsaturated diacylglycerol as a possible messenger for the activation of calcium-activated, phospholipid-dependent protein kinase system. *Biochem. Biophys. Res. Commun.* 91:1218-1224.

Contributed by A. Nigel Carter
The Salk Institute for Biological Studies
La Jolla, California

Permeabilization Strategies to Study Protein Phosphorylation

This unit deals with the use of nucleotide triphosphates to label proteins in vitro in permeabilized cells and isolated cellular fractions. These experiments generally utilize $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ as an exogenously added phosphate donor, although $[\gamma\text{-}^{32}\text{P}]\text{GTP}$ can be used in specific cases. The method is very straightforward, although numerous considerations must be made before applying it to each new system.

UNIT 18.2 describes labeling of phosphoproteins in intact cells with ^{32}P -labeled inorganic phosphate ($^{32}\text{P}_i$). After ascertaining that the protein of interest is a phosphoprotein using the protocols detailed in that unit, the phosphorylation event itself may be studied either in cell-free systems or in a more intact cellular system. In vitro, most protein kinases are capable of phosphorylating many substrates in addition to those phosphorylated in vivo, thus making it desirable to identify and study a protein kinase-catalyzed phosphorylation reaction in a more intact system than a cell lysate or homogenate. Permeabilization of intact cells provides a powerful experimental approach for accomplishing this. Brief permeabilization of cells is possible using a variety of agents. This procedure gives access to the intracellular milieu but still allows signal transduction events to be initiated by hormone stimulation of the cells via cell-surface receptors.

Brief chemical permeabilization of cells allows the introduction of cell-impermeable reagents into the intracellular compartment in order to study particular signal-transduction processes. In this type of experiment, relatively large molecules such as the Fab' fragments of specific antibodies may be introduced into the permeabilized cell, although the kinetics of entry may be too slow for the study of some processes. This technique is often more useful when introducing relatively small molecules into the permeabilized cell, for example non-cell-permeable activators (e.g., hydrophilic or lipophobic small molecules that cannot partition into the membrane) or inhibitors of protein kinases or phosphatases (see, e.g., Calbiochem catalog). The kinetics of entry into permeabilized cells for small molecules such as $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ are fast, and the system equilibrates the cellular ATP and $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ with intracellular processes within 2 to 3 min. Figure 18.8.1 shows the

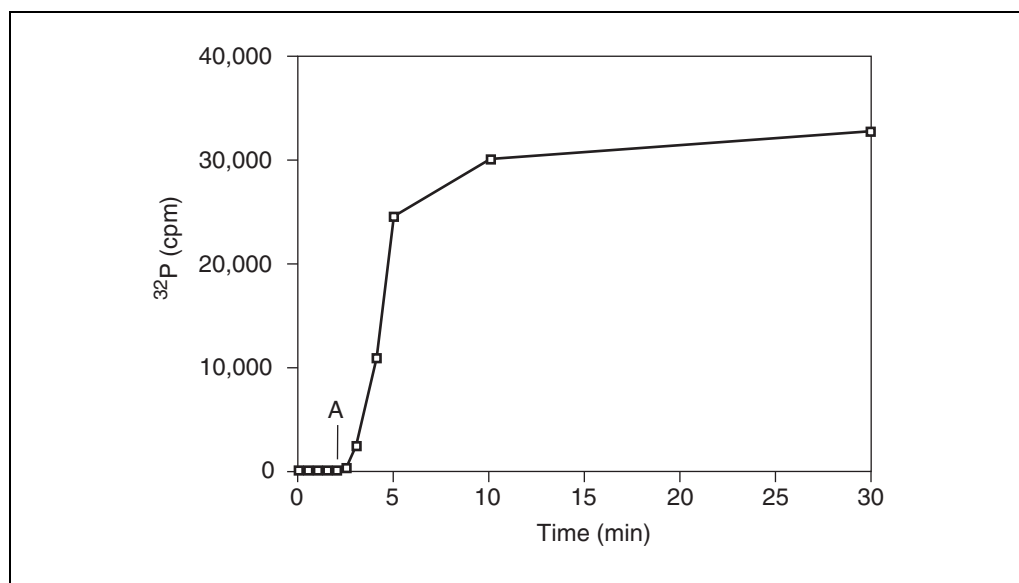


Figure 18.8.1 Equilibration of the monoester phosphates of the lipid phosphatidylinositol-4,5-bisphosphate with the ATP pool after permeabilization of human platelets. The letter A marks the point of addition of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$.

Contributed by A. Nigel Carter

Current Protocols in Molecular Biology (1997) 18.8.1-18.8.19

Copyright © 1997 by John Wiley & Sons, Inc.

equilibration of the monoester phosphates of the lipid phosphatidylinositol-4,5-bisphosphate with the ATP pool after permeabilization of human platelets. As these monoester phosphates are known to equilibrate with the ATP pool extremely quickly, this experiment provides a good indication of the speed of equilibration of cellular ATP pools with intracellular processes after addition of [γ - ^{32}P]ATP to permeabilized cells.

Procedures are outlined for performing a protein phosphorylation experiment using permeabilized cells (see Basic Protocol 1) and isolated intracellular organelles (see Basic Protocol 2). Both of these procedures result in lysates from which the protein of interest may be easily immunoprecipitated; however alternative techniques are described for preparing the final lysate from either Basic Protocol 1 or Basic Protocol 2 for electrophoretic analysis (see Alternate Protocols 1, 2, 3, and 4). A related procedure that does not involve permeabilization is outlined for direct analysis of cytosolic or membrane-bound kinases (see Alternate Protocol 5). Two different methods for determining the specific radioactivity of ^{32}P -containing compounds are also included (see Support Protocols 1 and 2).

CAUTION: It cannot be emphasized enough that each investigator should be fully familiarized with their local regulations regarding the safe handling and disposal of ^{32}P and also have the appropriate equipment for dealing with the containment of ^{32}P . See *APPENDIX 1F* for more information on the safe use of radioisotopes.

NOTE: Milli-Q purified water or its equivalent should be used to prepare all solutions used throughout this unit.

BASIC PROTOCOL 1

ANALYSIS OF PROTEIN PHOSPHORYLATION IN PERMEABILIZED CELLS

When setting out to study a protein phosphorylation event via a permeabilization protocol, the general procedure is as follows. First, cells are incubated in a buffer which has been designed to mimic the composition of the intracellular compartment and which contains a permeabilizing agent. The preferred permeabilization agent of choice is streptolysin O, as this agent allows many signal-transduction systems to remain intact for a number of minutes after its addition (Stephens et al., 1994). This is certainly not true of all permeabilization agents or protocols; if another agent (e.g., saponin) is to be tried, the methods detailed below should be used as a general guideline only. Depending on the experiment to be performed, the permeabilization buffer may also contain [γ - ^{32}P]ATP and the particular reagents under study. The [$^{32}\gamma$ -P]ATP provides the “hot” phosphate, and will equilibrate with the intracellular processes utilizing ATP within 2 to 3 min (A.N. Carter, unpub. observ.). After a brief time the experimental reagents, such as inhibitors, can be introduced into the permeabilization medium and hence into the inside of the cell. At this point the cells may be stimulated with the appropriate agent in the same way as for intact cell preparations. The phosphorylation of the protein under study is initially investigated by immunoprecipitation as described in *UNIT 10.16* and *UNIT 18.2*, followed by phosphopeptide mapping as described in *UNIT 18.3*.

Materials

- Cell culture to be labeled
- Cell culture medium appropriate for cells being studied, buffered with 20 mM HEPES, pH 7.2, prewarmed to 37°C
- Cells under study (normally cultured cell lines or isolated primary cells in culture)
- Extracellular buffer (see recipe), 37°C
- Permeabilization buffer (see recipe), 37°C
- 60 U/ml streptolysin O working solution (see recipe), freshly prepared
- 10 mM cold ATP stock solution (see recipe)
- 10 mCi/ml [γ - ^{32}P]ATP (3000 Ci/mmol; e.g., DuPont NEN)

Permeabilization Strategies to Study Protein Phosphorylation

18.8.2

Pharmacological agents, peptides, or Fab' fragments to be studied
Appropriate receptor agonists
2× lysis buffer for immunoprecipitation (see recipe), ice-cold
Protease inhibitor stock solutions (in absolute ethanol; store up to 6 months at
–20°C):
100 mM phenylmethylsulfonyl fluoride (PMSF)
100 mM benzamidine
1 mg/ml pepstatin A
1 mg/ml leupeptin
1 mg/ml antipain
100 mM DTT
Dry ice/ethanol bath
Noncirculating 37°C water bath containing a flat shelf (preferably made of wire
mesh) with space for the number of cell plates in the experiment
50-ml centrifuge tubes
Tabletop centrifuge
Cell scrapers
Screw-cap microcentrifuge tubes

Permeabilize cells

For monolayer cultures

- 1a. Approximately 1 hr before starting the experiment, change the medium in the cell culture plates to be labeled to equivalent medium containing 20 mM HEPES. Return cells to incubator and allow the cells to equilibrate with the new medium before use.

The number of cells used for any experiment of this type is totally dependent on how easy it is to detect the end product being studied. In some cases, a 60-mm dish of cells that has just reached confluency, with $\sim 1 \times 10^6$ cells, is adequate; in others a 100-mm dish of cells with $\sim 1 \times 10^7$ cells is required. It is advisable to start with $\sim 1 \times 10^6$ cells and see if the signal is detectable. Use of a 60-mm dish allows a smaller volume of permeabilization medium (normally of the order of 0.5 to 2 ml) to be used.

Standard cell culture medium is buffered by bicarbonate ions in the medium and the CO₂ in the atmosphere of the incubator. Outside the incubator this buffering system is ineffective and the pH will quickly rise if the culture is left in the normal atmosphere. The pH of HEPES buffers are minimally affected by temperature (unlike Tris buffers), hence they provide good pH control for cell culture medium. One drawback is that HEPES may be toxic to some cells above a concentration of ~ 20 mM, so it is important to check possible cellular toxicity before the use of HEPES-buffered culture media.

- 2a. Place the required number of plates containing cells in the 37°C water bath and allow 10 min for the medium to reequilibrate.

The cells are stable in this state for an hour or so, depending on the rate of evaporation of the culture medium.

- 3a. Aspirate the culture medium from each plate and quickly rinse the cells twice, each time with 10 ml of 37°C extracellular buffer. Aspirate the extracellular buffer, then add 0.5 to 2 ml permeabilization buffer to each plate.

Be careful when rinsing, as some cell lines are less adherent than others. Usually, most epithelial or mesenchymal lines adhere well and will not slough off the dish unless extreme force is used. Transformed cells usually adhere poorly and can easily be rinsed off the dish if care is not taken.

- 4a. Add streptolysin O (from the 60 U/ml working solution) to each plate to a final concentration of 0.6 U/ml. Add cold ATP (from the 10 mM stock solution) to a final concentration of 100 μ M.

For suspension cultures

- 1b. Approximately 1 hr before starting the experiment, transfer cells to sterile 50-ml centrifuge tubes and centrifuge 5 min at $1000 \times g$, room temperature. Remove the supernatant, then rinse cells twice, each time by adding ~40 ml of 37°C extracellular buffer, centrifuging 5 min at $1000 \times g$, room temperature, and removing the supernatant.

The cell suspension may be split into aliquots, if desired, after the addition of the final rinse of extracellular medium. To do this, add the final (40-ml) rinse of extracellular medium, resuspend cells gently, and dispense aliquots into fresh tubes. Centrifuge the aliquots and remove the supernatants to complete the final rinse before proceeding to step 2b.

- 2b. Resuspend the cells in 37°C extracellular buffer at a density of 1×10^7 cells/ml. Allow the cells to equilibrate 15 to 30 min at 37°C with gentle swirling.

This step may have to be modified if a cell is used that could be activated by this procedure, (e.g., platelets). If a sensitive cell is used, equilibration time may be kept down to 5 to 10 min and the agitation may be omitted.

- 3b. Centrifuge cells again for 5 min at $1000 \times g$, room temperature, and aspirate the buffer.
- 4b. Add streptolysin O (from the 60 U/ml working solution) to fresh 37°C permeabilization buffer to a final concentration of 0.6 U/ml. Add cold ATP (from the 10 mM stock solution) to a final concentration of 100 μ M. Add this permeabilization solution containing streptolysin O and ATP to the cells and mix gently but thoroughly.

Addition of cold ATP at the same time as streptolysin O allows cellular ATP levels to be maintained and will therefore preserve cell viability. One important consideration is whether the cells under study possess cell-surface receptors for ATP. The ATP acts as a full agonist for some types of purinergic receptors, and this can obviously cause confusion when analyzing signal-transduction mechanisms. Prior to performing permeabilization studies, one should first determine, using intact cells, whether extracellular ATP activates signal-transduction mechanisms that interfere with the system under investigation. Stephens et al. (1994) have developed a way of inactivating extracellular receptors for ATP in neutrophils prior to permeabilization. The utility of this method for other cell types is not known to date, although it may be used as a guideline if required. Readers are referred to the original paper for full discussions of this method.

Permeabilize and label cells

5. Permeabilize cells by incubating 1 to 5 min at 37°C.

The time of incubation depends very much on the particular system under study and the cell type, and must be determined empirically. Permeabilization for >5 min usually reduces hormone responsiveness of cells dramatically. This is probably due to a number of processes, including the loss of cytosolic components into the extracellular permeabilization buffer and physical uncoupling of receptors from signal-transduction processes.

6. Add 10 mCi/ml [γ - 32 P]ATP to a final concentration of 50 to 500 μ Ci/ml.

[γ - 32 P]ATP is usually purchased at a specific activity of ~3000 Ci/mmol, and a concentration of 10 mCi/ml. Hence 5 μ l of the stock is 50 μ Ci [γ - 32 P]ATP. The label may be used at this concentration. If the stock [γ - 32 P]ATP is to be diluted into another solution it is better to do this directly before use.

The criterion for the amount of radiolabeled ATP to add (between 50 and 500 μ Ci/ml) is the detection limit. Initial experiments must be performed to see how much label must be added.

The alternative to this separate labeling step is to add the [γ - 32 P]ATP at the same time as the streptolysin O and cold ATP in step 4a or b. The type of experiment to be performed will determine at which point the [γ - 32 P]ATP is added to the cells.

7. Add the pharmacological agents, peptides, or Fab' fragments under study and allow time for their action as required.

Many pharmacological agents amenable to testing by this protocol are listed in the Calbiochem catalog. For example, a peptide containing residues 280 to 305 of the CaM kinase II α subunit can be a potent inhibitor of CaM kinase II and is non-cell-permeable. There are other peptides modeled on pseudosubstrate inhibitors of a variety of protein kinases that can be used as specific inhibitors, all of which are non-cell-permeable. Any Fab' fragment may be tested by this protocol as long as it has an effect on enzyme activity, either positive or negative. Similarly, if an antibody stops an enzyme from reaching its site of action, it may also be tested via this procedure.

The time for a particular agent to act as required should be determined experimentally as part of the initial experimental protocols. As with the [γ - 32 P]ATP addition, these agents may also be added at the same time as streptolysin O and cold ATP.

8. Add receptor agonists as required.

The same considerations are true for receptor agonist additions as for inhibitor and [γ - 32 P]ATP additions to the cells. One must perform preliminary experiments to empirically determine the timings for these additions for each cell type under study.

The remaining steps of this protocol are to prepare the cells for analysis by immunoprecipitation. If electrophoretic analysis is to be performed, see Alternate Protocols 1 and 2.

Prepare cell lysates for immunoprecipitation

- 9a. *For monolayer cultures:* Stop the reactions by rapidly aspirating the permeabilization buffer and adding 0.75 ml ice-cold 2 \times lysis buffer for immunoprecipitation to each plate. Add protease inhibitors and DTT (as stock solutions) to the following final concentrations:

PMSF to 1 mM final
Benzamidine to 1 mM
Pepstatin A to 5 μ g/ml final
Leupeptin to 5 μ g/ml final
Antipain to 5 μ g/ml final
DTT to 1 mM final.

Place plates on ice for 10 min, then scrape the cell debris from the bottoms of the plates into the buffer using a separate cell scraper for each plate.

- 9b. *For suspension cultures:* Stop the reactions by adding an equal volume of ice-cold 2 \times lysis buffer for immunoprecipitation. Add protease inhibitors and DTT to the final concentrations in step 9a, then place the tubes of cells on ice for 10 min.
10. Transfer lysates to screw-cap microcentrifuge tubes (one plate or cell aliquot/tube). Microcentrifuge lysates at maximum speed, 4 $^{\circ}$ C, then transfer the supernatants to fresh screw-cap microcentrifuge tubes and freeze immediately in a dry ice/ethanol bath. Store at -70° C until needed.

The lysates are now ready for immunoprecipitation (UNIT 10.16).

CAUTION: *The permeabilization buffer in the latter steps of this protocol contains [γ - 32 P]ATP, and great care should be taken to remove the medium safely and dispose of it correctly as liquid radioactive waste. The lysates and cell scrapers are also very radioactive at this point. Screw-cap tubes should be used at all subsequent stages of processing the cell lysates as these provide the best containment of the 32 P-containing samples. The same microcentrifuge should be used for all centrifugation steps; after use it will undoubtedly be contaminated with 32 P. It can then be set aside for future use with 32 P-containing samples with the appropriate safety precautions.*

INTACT CELL SAMPLE PREPARATION FOR ELECTROPHORETIC ANALYSIS OF PROTEIN PHOSPHORYLATION

The preceding protocol (see Basic Protocol 1) is set up to allow immunoprecipitation of specific proteins after the permeabilization experiment. Two alternatives to this approach, described below, involve using the separating power of polyacrylamide gel electrophoresis (PAGE) to fractionate proteins, instead of immunoprecipitation. Fractionation can be either by one-dimensional SDS-PAGE (UNIT 10.2), or by two-dimensional PAGE in which isoelectric focusing is performed as the first dimension, followed by SDS-PAGE (UNIT 10.3). In each case, sample loading must be optimized so that best resolution of the protein of interest is achieved.

These protocols are written for both monolayer cells and suspension cells. To perform the procedure with suspension cells, the cells must be centrifuged down as quickly as possible after reacting with the appropriate pharmacological agents/receptor agonists, then treated as described for monolayer cultures. Obviously, the timing of the reactions is less precise this way, but experiments can be done successfully and reproducibly if reaction and centrifugation times are kept exactly the same.

CAUTION: Samples for SDS-PAGE or isoelectric focusing will still contain a great deal of ^{32}P at the latter stages of the protocol. Great care should be exercised in sample preparation. In particular do not use a probe sonicator to shear DNA, as ^{32}P -containing aerosols will be produced. In the case of SDS-PAGE, the gels should be run until the dye front passes out of the gel. This means that the SDS-PAGE running buffer will be also be contaminated with ^{32}P . In the case of isoelectric focusing the ^{32}P will focus out of the gel during electrophoresis; hence the running buffer will be contaminated with ^{32}P . These solutions should be disposed of as radioactive waste as local regulations dictate.

ALTERNATE PROTOCOL 1

Intact Cell Sample Preparation for SDS-PAGE

Additional Materials (also see Basic Protocol 1)

2× SDS-PAGE sample buffer (see recipe), 4°C

Bath sonicator

Boiling water bath

1. Prepare cells and perform permeabilization and labeling (see Basic Protocol 1, steps 1 through 8).
- 2a. *For monolayer cultures:* Stop the reactions by rapidly aspirating the permeabilization buffer and adding 100 to 250 μl of 4°C 2× SDS-PAGE sample buffer to each plate. Quickly scrape the cells into the buffer using a separate cell scraper for each plate and quantitatively transfer the sample to a screw-cap microcentrifuge tube.
- 2b. *For suspension cultures:* Microcentrifuge the reaction tubes 1 min at maximum speed, room temperature. Quickly aspirate the supernatants, taking care not to disturb the cell pellets. Add 100 to 250 μl of 2× SDS-PAGE sample buffer to each tube and vortex thoroughly.

When performing this step for suspension cultures it is very important to use exactly the same times for centrifugations if reproducible data are to be obtained.

3. If samples are very viscous as a result of DNA being released from cell nuclei, shear the DNA by placing the capped sample tubes in a bath sonicator filled with ice water and sonicating 5 min.
4. Boil samples 3 min in a boiling water bath.

To perform the analysis, carefully apply up to 50 μl of the samples to the sample wells of a precast SDS-PAGE gel. Proceed as described in UNIT 10.2.

Intact Cell Sample Preparation for Isoelectric Focusing

ALTERNATE
PROTOCOL 2

Additional Materials (also see Basic Protocol 1)

Two-dimensional-PAGE lysis buffer (see recipe), 4°C

Two-dimensional-PAGE sample buffer (see recipe)

Bath sonicator

Dry ice/ethanol bath

Lyophilizer

1. Prepare cells and perform permeabilization and labeling (see Basic Protocol 1, steps 1 through 8).
- 2a. *For monolayer cultures:* Stop the reactions by rapidly aspirating the permeabilization buffer and adding 100 to 250 μ l of 4°C two-dimensional-PAGE lysis buffer. Scrape the cells into the buffer using a separate cell scraper for each plate and transfer the sample to a screw-cap microcentrifuge tube.
- 2b. *For suspension cultures:* Microcentrifuge the reaction tubes 1 min at maximum speed, room temperature. Quickly aspirate the supernatants, taking care not to disturb the cell pellets. Add 100 to 250 μ l of two-dimensional PAGE lysis buffer to each tube and vortex thoroughly.

When performing this step for suspension cultures it is very important to use exactly the same times for centrifugations if reproducible data are to be obtained.

3. If samples are very viscous as a result of DNA being released from cell nuclei, shear the DNA by placing the samples in a bath sonicator filled with ice-water and sonicating 5 min. Freeze samples in a dry ice/ethanol bath, then lyophilize.
4. Redissolve the lyophilized samples in 50 to 100 ml of two-dimensional-PAGE sample buffer. Warm the samples to 37°C for 3 to 5 min.

To perform the analysis, apply 10 to 20 μ l of each sample to a prepared isoelectric focusing gel of the required pH gradient (see UNIT 10.3). Samples for isoelectric focusing must have a low salt concentration because a high salt concentration often causes band broadening during isoelectric focusing.

ANALYSIS OF PROTEIN PHOSPHORYLATION USING ISOLATED SUBCELLULAR FRACTIONS

BASIC
PROTOCOL 2

When the protein under study exists in a particular organelle and is phosphorylated in intact cells, one consideration when studying the phosphorylation event *in vitro* in isolated organelles is whether the protein is accessible to exogenously added kinases or phosphatases, or whether it is accessible to endogenous kinases or phosphatases. For instance, if it is a protein of the endoplasmic reticulum, is its location transmembrane or intraluminal? If transmembrane, is it normally phosphorylated on the cytosolic side or luminal side of the membrane?

Permeabilization protocols can also be used with intact intracellular organelles as well as intact cells. This is obviously an important consideration when dealing with proteins found inside the organelle in question. Another consideration here is whether the organelle can be isolated in an intact condition, or whether its contents will be partially or completely lost during purification. In the case of mitochondria this does not pose a problem, but for more fragile organelles such as the Golgi apparatus this may be more problematic.

The study of protein phosphorylation in isolated organelles (see Castle, 1995 for isolation procedures) can be a straightforward extension of the permeabilization protocols de-

Analysis of
Protein
Phosphorylation

18.8.7

scribed above (see Basic Protocol 1 and Alternate Protocols 1 and 2). The isolated organelle is treated in the same way as one would treat cultured cells in suspension. However, as there is little or no requirement for the maintenance of signal-transduction mechanisms in isolated organelles, the permeabilization procedure can be more prolonged than for intact cells. This allows the introduction of larger molecules such as purified protein kinases into the organelles under study.

Materials

Suspension of isolated intracellular organelles
Intracellular buffer (see recipe), 4°C
Permeabilization buffer (see recipe), 37°C
60 U/ml streptolysin O working solution (see recipe)
10 mM cold ATP stock solution (see recipe)
10 mCi/ml [γ -³²P]ATP (3000 Ci/mmol; e.g., DuPont NEN)
Reagents under study: e.g., kinase/phosphatase inhibitors, protein kinases, or phosphatases
2× lysis buffer for immunoprecipitation (see recipe), 4°C

Screw-cap microcentrifuge tubes
Benchtop ultracentrifuge (e.g., Beckman TL-100 or Airfuge)
37°C circulating water bath

1. Wash the organelle preparation three times with 4°C intracellular buffer, each time by ultracentrifuging 10 min at $\geq 100,000 \times g$. Transfer the equivalent of 1 mg of protein to individual screw-cap microcentrifuge tubes and keep on ice.

The best way to wash the organelle preparations is to use a Beckman TL-100 benchtop ultracentrifuge or the older Beckman Airfuge. Both of these instruments will produce g-forces $> 100,000 \times g$. Because of the difficulty in centrifuging some organelles (e.g., plasma membrane preparations) it is essential to use ultracentrifugation for all organelles except nuclei, which can be recovered by microcentrifuging 30 sec at maximum speed. See Castle (1995) for details of organelle isolation.

2. Ultracentrifuge the organelle preparation 10 min at $\geq 100,000 \times g$, 4°C, and aspirate the supernatant. Add 200 ml 37°C permeabilization buffer, mix gently, and transfer to the 37°C circulating water bath.

Work quickly at this point to minimize any proteolytic and other unwanted side reactions.

3. Add streptolysin O (from the 60 U/ml working solution) to each tube to a final concentration of 0.6 U/ml. Add cold ATP (from the 10 mM stock solution) to a final concentration of 100 μ M.

Exactly the same considerations apply when performing permeabilization of organelles as when performing permeabilization of intact cells (see Basic Protocol 1). However, because there is no intact signal-transduction machinery in isolated organelle preparations, the range of possible experiments is fewer, so the system is usually more robust than for intact-cell experiments.

4. Add 10 mCi/ml [γ -³²P]ATP to a final concentration of 50 to 500 μ Ci/ml.

The criterion for the amount of radiolabeled ATP to add (between 50 and 500 μ Ci/ml) is the detection limit. Initial experiments must be performed to see how much label must be added.

5. Add the reagents under study—e.g., kinase/phosphatase inhibitors, protein kinases, or phosphatases.

6. Add 200 ml of 4°C 2× lysis buffer to each tube to lyse the organelles. Mix well and place on ice for 5 min.
7. Microcentrifuge the samples 10 min at maximum speed, 4°C. Store lysates at –70°C.

The lysates are now ready for immunoprecipitation (UNIT 10.16).

ORGANELLE SAMPLE PREPARATION FOR ELECTROPHORETIC ANALYSIS OF PROTEIN PHOSPHORYLATION

Exactly the same considerations apply to the electrophoretic analysis of protein phosphorylation in isolated organelles as apply to permeabilized cells (see Alternate Protocols 1 and 2). Organelle preparations must first be ultracentrifuged and the supernatant removed before adding the appropriate buffer for electrophoresis. For this, the use of a Beckman TL-100 benchtop ultracentrifuge, which can provide g -forces $>300,000 \times g$, is recommended. Complete removal of the supernatant is especially important if two-dimensional-PAGE (Alternate Protocol 4) is used, as the salt concentration must be kept low.

CAUTION: Samples for SDS-PAGE or isoelectric focusing will contain a great deal ^{32}P at the latter stages of the protocol. Great care should be exercised in sample preparation. In the case of SDS-PAGE, the gels should be run until the dye front passes out of the gel. This means that the SDS-PAGE running buffer will be also be contaminated with ^{32}P . In the case of isoelectric focusing, the ^{32}P will focus out of the gel during electrophoresis; hence the running buffer will be contaminated with ^{32}P . These solutions should be disposed of as radioactive waste as local regulations dictate.

Organelle Sample Preparation for SDS-PAGE

Additional Materials (also see Basic Protocol 2)

2× SDS-PAGE sample buffer (see recipe)
Boiling water bath

1. Prepare organelles and perform permeabilization and labeling (see Basic Protocol 2, steps 1 through 5).
2. Ultracentrifuge samples 10 min at $100,000 \times g$, 4°C, and aspirate the supernatant. Add 50 μl of 2× SDS-PAGE sample buffer and mix well.
3. Boil the samples 3 min in a boiling water bath.

The samples are now ready for analysis by SDS-PAGE (UNIT 10.2).

Organelle Sample Preparation for Isoelectric Focusing

Additional Materials (also see Basic Protocol 2)

Two-dimensional-PAGE lysis buffer (see recipe)
Lyophilizer

1. Prepare organelles and perform permeabilization and labeling (see Basic Protocol 2, steps 1 through 5).
2. Ultracentrifuge samples 10 min at $>100,000 \times g$, 4°C, and aspirate the supernatant. Add 200 μl of two-dimensional-PAGE lysis buffer and mix well.
3. Lyophilize the samples overnight.

The samples are now ready for analysis by isoelectric focusing (UNIT 10.3). Samples for isoelectric focusing must have a low salt concentration because a high salt concentration often causes band broadening during the isoelectric focusing dimension.

**ALTERNATE
PROTOCOL 3**

**ALTERNATE
PROTOCOL 4**

**Analysis of
Protein
Phosphorylation**

18.8.9

DIRECT ANALYSIS OF CYTOSOLIC OR MEMBRANE-BOUND KINASES

This protocol utilizes isolated subcellular fractions in straightforward kinase assays by adding exogenous, purified protein kinases to the reaction mixture containing the organelle in question. This works only if the protein under study is phosphorylated by a cytosolic or membrane-bound kinase and its substrate is on the cytosolic face of the organelle. One can then determine the phosphorylation of proteins by analyzing the reaction mixture by SDS-PAGE. It is a good idea to perform the reactions and subsequent manipulations in screw-cap microcentrifuge tubes to minimize the risk of [^{32}P] contamination.

This protocol may be used to extend findings based on the use of the permeabilization strategies and assays with isolated organelles detailed elsewhere in this unit. Hence, if a protein of interest is phosphorylated, it can be useful to determine the subcellular localizations of the kinase(s) responsible, as this may make it possible to infer clues as to the identity of the kinase(s).

Materials

Appropriate kinase assay buffer (UNIT 18.7), 4°C
[γ - ^{32}P]ATP solution (see recipe in UNIT 18.7)
Suspension of isolated intracellular organelles, 4°C
Purified protein kinase of interest (keep on ice)
2× SDS-PAGE sample buffer (see recipe), ice-cold

Screw-cap microcentrifuge tubes
30°C and boiling water baths

1. Prepare the following reaction mixture in a screw-cap microcentrifuge tube at 4°C:

10 μl 5× kinase assay buffer
5 μl [γ - ^{32}P]ATP solution (5 μCi /5 μM ATP final)
1 to 35 μl organelle suspension (i.e., volume equivalent to 1 mg protein)
 H_2O to 50 to 100 μl .

After reaction mix has been assembled, warm to 30°C for 5 min.

2. Start reaction by addition of the purified protein kinase of interest. Incubate 15 min at 30°C.

In the case of immunoprecipitated enzymes adsorbed onto beads, the assay mix minus the enzyme source can be made up complete first. This can then be dispensed, prewarmed, onto the immunoprecipitates.

Incubation time can be adjusted as necessary depending on the extent of protein phosphorylation observed and the linearity of the reaction.

3. Stop the reactions by the addition of 100 μl of ice-cold 2× SDS-PAGE sample buffer. After thorough mixing, heat the samples 3 min in a boiling water bath.

The samples are then ready for resolution by SDS-PAGE (UNIT 10.2).

Samples for SDS-PAGE contain all the ^{32}P originally in the assays. Great care should be exercised in sample preparation and loading. The gels should be run until the dye front passes out of the gel. This means that the SDS-PAGE running buffer will be contaminated with ^{32}P , and this should be disposed of as radioactive waste as local regulations dictate.

DETERMINATION OF SPECIFIC RADIOACTIVITY

The specific radioactivity (specific activity) of any radioactive compound is defined as the number of atoms of the radioactive isotope in a molecule as a proportion of the total number of atoms of the same type. The structure of ATP is shown in Figure 18.8.2, and the form used in the study of protein kinases is $[\gamma\text{-}^{32}\text{P}]\text{ATP}$, indicating that only the γ phosphate group has the radioactive ^{32}P atom. There are, however, two other atoms of ^{31}P . Hence, to attain the maximum possible specific activity, one should replace 100% of all phosphorus atoms with the ^{32}P isotope. This is obviously not necessary in the case of kinase phosphotransfer reactions, as only the γ phosphate is transferred to a substrate; therefore in practice it is irrelevant whether the α and β phosphate groups have the ^{32}P label. The units used by vendors of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ are curies per mole (Ci/mol), even though the curie is no longer the international unit of radioactivity.

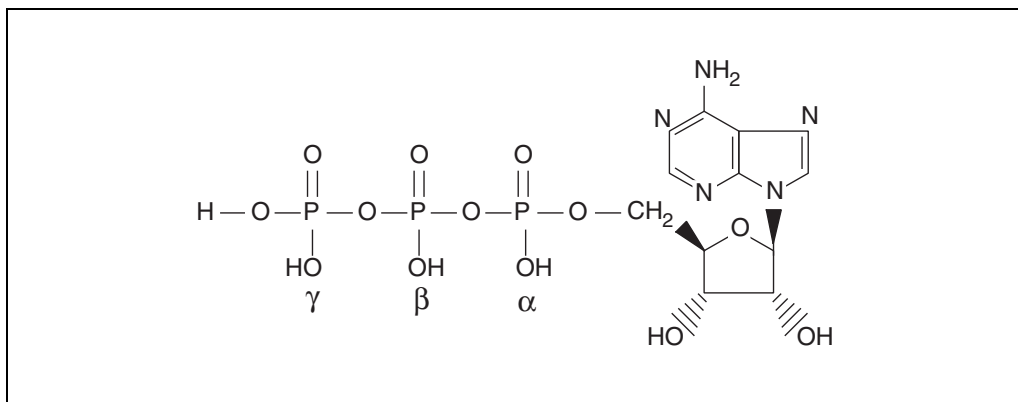


Figure 18.8.2 Structure of ATP highlighting the α , β , and γ phosphates/phosphorus atoms.

It is important to know the specific radioactivity of the ATP in a kinase assay, as important data can be derived from this information. Consequently this can be determined as required, although one can make assumptions based on the volume of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ added to cold ATP prior to the assay. Obviously, the specific activity of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ cannot be assumed when adding $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ to permeabilized cells or after labeling intact cells with $^{32}\text{P}_i$. In these cases, one can determine the specific activity experimentally as described below.

The specific radioactivity of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ can be determined by a number of methods, but the following are reliable and relatively simple. Methodologies for the extraction of ATP from either $^{32}\text{P}_i$ -labeled cells or from cell lysates are detailed below (Stephens and Downes, 1990). The basis for the extraction is the same in each case, but slightly different alternative steps are required. Protein or cell debris is first precipitated with perchloric acid and removed by centrifugation. The supernatants are then neutralized with a mixture of tri-*n*-octylamine and 1,1,2-trichlorotrifluoroethane (Freon). The aqueous phase from this step can then be directly analyzed for ATP concentration and ^{32}P content. ATP may be quantified using anion-exchange HPLC to separate nucleotides and a UV detector to detect nucleotides via UV absorption at 254 nm. The peak areas associated with injection of different amounts of ATP onto the column are then calculated and a standard curve constructed. The radioactivity associated with the ATP peak is easily determined by scintillation counting.

Materials

- ³²P-labeled cells or cell lysates
- Phosphate-buffered saline (PBS; APPENDIX 2)
- Standard ATP samples of known concentration in 4% perchloric acid 4% or 8% (v/v) perchloric acid, ice-cold
- Tri-*n*-octylamine (Sigma)
- 1,1,2-trichlorotrifluoroethane (Freon; e.g., Aldrich, Sigma)
- Linear gradient consisting of:
 - Solution A: Milli-Q water filtered through 0.2-mm filter
 - Solution B: 1.25 M NH₄H₂PO₄, pH 3.8 (adjust pH with H₃PO₄; filter through 0.2-mm filter)
- Nucleotide standard mix: e.g., 1 mM each of AMP, ADP, ATP, GMP, GDP, and GTP

- Cell scrapers
- Polypropylene screw-cap microcentrifuge tubes or 5-ml polypropylene tubes with very tight-fitting caps
- Tabletop centrifuge
- HPLC system consisting of:
 - Whatman Partisphere SAX HPLC column and guard column
 - Apparatus for producing linear gradient
 - UV detector set at 254 nm
 - Fraction collector
- Whatman low-dead-volume 0.2-mm syringe filters

- Additional reagents and equipment for HPLC (UNITS 10.12-10.14)

Extract ATP from cells or cell lysates

For labeled cells

- 1a. Aspirate the medium from the labeled cells and rinse twice with PBS. Aspirate the final wash thoroughly and place the cells on ice. Quickly add 0.5 ml of ice-cold 4% perchloric acid and mix gently to lyse the cells. Incubate 5 min on ice.
- 2a. Scrape the cells into the acid solution using a cell scraper and quantitatively transfer the lysate to a screw-cap microcentrifuge tube. Microcentrifuge 5 min at maximum speed, 4°C, to bring down the cell debris.
- 3a. Transfer the supernatant to a fresh tube and keep on ice. Freshly prepare a mixture of 1.2 parts tri-*n*-octylamine to 1 part Freon (on a v/v basis) and mix well. Add 0.6 ml of this mixture to the perchloric acid supernatants, cap the tubes tightly, and vortex for at least 1 min. Extract four to six standard ATP samples of known concentration in parallel with the test supernatants.

It is very important that the samples be mixed extremely thoroughly at this point to ensure complete neutralization.

For labeled cell lysates

- 1b. In a 5-ml polypropylene tube add an equal volume of ice-cold 8% perchloric to the amount of labeled cell lysate required and place on ice. Mix thoroughly and incubate 5 min on ice.
- 2b. Centrifuge 5 min (or until protein is completely pelleted) at 3800 × *g*, 4°C.
- 3b. Transfer the supernatant to a fresh tube and keep on ice. Freshly prepare a mixture of 1.2 parts tri-*n*-octylamine to 1 part Freon (on a v/v basis) and mix well. Add 1.2 vol of this mixture to the volume of sample remaining after protein precipitation and

vortex extremely well. Extract four to six standard ATP samples of known concentration in parallel with the test supernatants.

It is very important that the samples be mixed extremely thoroughly at this point to ensure complete neutralization.

Recover neutralized aqueous sample

4. After complete mixing, centrifuge 5 min at $3800 \times g$, room temperature.

*Three layers should be evident following centrifugation. The bottom layer is excess tri-*n*-octylamine/Freon, the central layer is octylamine perchlorate (which is often yellow in color), and the upper layer is the neutralized aqueous sample.*

5. Transfer the upper layer to a fresh tube and measure the pH using pH paper.

*The pH should be pH 6 to 7. If the pH of the sample is <6 , repeat the treatment with tri-*n*-octylamine/Freon, although this is rarely necessary.*

Quantitate ATP concentration

6. Equilibrate the SAX column in water for 30 min at 1 ml/min flow rate.

As an alternative to this step and subsequent steps it is possible to quantitate ATP using a luminescent assay. A very easy, sensitive and convenient assay for ATP is available from Sigma and is based on using the enzyme luciferase, which produces light from luciferin and ATP. Hence, at constant luciferin/luciferase concentrations, standard curves can be constructed for the quantum light yield at different ATP concentrations. Unknown samples derived from cell lysates can thus be easily quantitated. If, however, one needs to determine the specific radioactivity of the γ phosphate of ATP (as for $^{32}\text{P}_i$ -labeled cells), proceed as described in this protocol.

7. Subject the column to a blank gradient run—i.e., run the linear gradient to be used (from 100% solution A to 100% solution B) through the column but do not inject the sample.

A typical simple gradient is shown in Table 18.8.1. The initial 5 min allows the contents of the injection loop to be cleared and passed onto the column, assuming use of up to a 5-ml-volume loop. It is very important that the ionic strength of the sample be <200 mM for the inorganic phosphate to bind to the column. In practice, sample loops of up to 5-ml volume can be used successfully, allowing 0.5-ml samples to be diluted 10-fold with water before analysis.

Table 18.8.1 Gradient Profile for the Separation of Nucleotides by HPLC and Approximate Retention Times for Selected Nucleotides^{a,b}

Time (min)	%A	%B	Approximate retention time (min)		
			AMP/GMP	ADP/GDP	ATP/GTP
0	100	0			
5	100	0			
70	40	60	30/35	45/52	75/85
85	40	60			
90	0	100			
120	0	100			
125	0	0			

^aA = Milli-Q water; B = 1.25 M $\text{NH}_4\text{H}_2\text{PO}_4$ buffer, pH 3.8.

^bColumn = Whatman Partisphere SAX, 25 cm \times 4.6-mm i.d. The retention times for the nucleotides are approximate, as different HPLC systems produce slight differences in gradient delivery and different HPLC columns give different separations.

8. Filter standard mixture of nucleotides through a low-dead-volume syringe filter, then run on the SAX column.

A good mix to use is AMP, ADP, ATP, GMP, GDP, and GTP, all at a concentration of 1mg/ml. A 10-ml injection will result in easily detected peaks at 254 nm using the UV detector. Once consistent separation of the required standard nucleotides is made, the samples can be analyzed.

9. Filter the four to six ATP samples of known concentration (from step 5) through a low-dead-volume syringe filter and inject them into the SAX column. Construct a standard curve.

The unknown samples should fall around the center of the curve. In practice, this can be estimated according to the fact that the intracellular concentration of ATP is normally ~1 mM, so the starting cell volume can be used to roughly estimate the ATP concentration in the cell lysates.

10. Filter the unknown samples (from step 5) through a low-dead-volume syringe filter, then run them on the SAX column and collect the ATP peak (which absorbs at 254 nm) using a fraction collector connected in series after the UV monitor.

The concentration of ATP is then calculated.

11. Transfer an aliquot of the ATP peak to a scintillation vial and determine cpm.

The amount of [γ - ^{32}P]ATP is then calculated from the amount of ^{32}P associated with the peak absorbing at 254 nm.

SUPPORT PROTOCOL 2

DETERMINATION OF THE SPECIFIC ACTIVITY OF THE γ -PHOSPHATE OF [^{32}P]ATP

The above technique (see Support Protocol 1) details methods for the determination of the total ATP concentration and the amount of ^{32}P associated with that concentration. For the case of permeabilized cells where [γ - ^{32}P]ATP has been introduced specifically as the phosphate donor, the amount of radioactivity associated with ATP will give the specific activity of the γ phosphate. However, when cells are labeled with [^{32}P]inorganic phosphate ($^{32}\text{P}_i$) as in UNIT 18.2, the ^{32}P will be distributed between all of the phosphorus atoms in ATP and hence the total radioactivity will not give a meaningful result regarding the specific activity of the γ phosphate of ATP. Also, in permeabilization studies such as the ones described in this unit, it is good to know the final specific activity of the ATP because the [γ - ^{32}P]ATP is diluted with the ATP contained within the cells. A good method for obtaining this specific activity information is the one developed by Hawkins et al. (1983), which uses cyclic AMP-dependent protein kinase (cA-PrK) to phosphorylate the substrate histone 2A under conditions where the purified [^{32}P]ATP is the sole ATP in the assay. The ^{32}P -labeled ATP is completely used up during the phosphorylation of histone. Phosphorylated histone is easily separated from other ^{32}P -containing metabolites; hence the amount of ^{32}P associated with the γ phosphate of the labeled ATP is easily determined. The method detailed below is a slight modification of the Hawkins method and uses a peptide substrate in place of histone to allow greater specificity. This procedure is the basis for many protein kinase assays and is described in greater detail in UNIT 18.7. The details of assays for cA-PrK are dealt with in detail in that unit, along with specifics about assay buffers and substrate preparation.

Materials

- Cell lysates containing unknown [^{32}P]ATP concentrations
- cA-PrK (purified subunit from bovine heart; Calbiochem) and 10 \times cA-PrK assay buffer (see recipe)
- 0.5 mM cyclic AMP (cAMP; Calbiochem)

10 mg/ml kemptide (Sigma)
1 M DTT (store at -20°C)
0.5 M MgCl_2
75 mM orthophosphoric acid
Acetone

Screw-cap microcentrifuge tubes

P 81 phosphocellulose paper, precut and labeled (Support Protocol 3 in *UNIT 18.7*)

30°C circulating water bath

Additional reagents and equipment for precipitating ^{32}P -labeled peptide on P 81 paper (Support Protocol 3 in *UNIT 18.7*)

1. For each lysate to be assayed, prepare the following reaction mixture in a screw-cap microcentrifuge tube on ice:

25 μl [γ - ^{32}P]ATP-containing lysate
5 μl $10\times$ cA-Prk assay buffer
1 μl 0.5 mM cAMP
5 μl 10 mg/ml kemptide
1 μl 1 M DTT
1 μl 0.5 M MgCl_2
2 μl H_2O .

Once all assay tubes have been set up and labeled, cap tubes and keep on ice.

2. Start one reaction every 30 sec by directly adding 15 U cA-PrK to the reaction mixture and transferring the tube to a 30°C water bath. Incubate for the optimal period of time.

Because the objective is to transfer all of the γ -phosphate of ATP if possible, a relatively large amount of cA-PrK is required. Initially, 15 U per tube can be used; this can be adjusted for later experiments.

The optimal assay time must be determined empirically by performing a time course. Times of 60, 120, 180, and 240 min should be tried; the time chosen should be one after the incorporation of ^{32}P into kemptide has reached a plateau.

3. Pipet 30 μl of the contents of each tube onto the surface of prepared P 81 paper squares (see *UNIT 18.7*, Support Protocol 1). Immediately place the squares in a beaker of 75 mM orthophosphoric acid.
4. Once all the assay tubes have been treated as in step 3, wash the P 81 paper squares five times with 75 mM orthophosphoric acid as described in *UNIT 18.7*, Support Protocol 1.
5. Fill the beaker with acetone and wash the papers an additional 5 min as described in *UNIT 18.7*, Support Protocol 1. Discard the acetone and allow the papers to dry at room temperature.
6. Place each square of P 81 paper in a separate 20-ml scintillation vial and determine the associated radioactivity using a scintillation counter.

Once the number of dpm of ^{32}P transferred is known, it is a simple task to calculate the specific activity of the ATP input into the experiment. For example—if the amount of total ATP has been determined to be 10 μmol , the radioactivity associated with ATP is 220,000 dpm, the amount of ^{32}P transferred to kemptide is 110,000 dpm, and the amount of ^{32}P remaining is 110,000 dpm—then there are 125,000 dpm of ^{32}P in the γ -phosphate of ATP as there are $\sim 2.2 \times 10^6$ dpm per μCi . Hence, 110,000 dpm is equivalent to 0.05 μCi , so the specific activity of the γ -phosphate of ATP is 0.5 mCi/mmol.

REAGENTS AND SOLUTIONS

Use Milli-Q purified water or equivalent in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

ca-PrK assay buffer, 10×

500 mM Tris·Cl, pH 7.5 at 30°C (APPENDIX 2)

50 mM MgCl₂

Store up to 6 months at -20°C

Cold ATP stock solution, 10 mM

Mix equal volumes of 20 mM MgCl₂ and 20 mM disodium ATP. Slowly raise the pH of the solution to 7.4 with 1 M NaOH using constant stirring. Filter sterilize and store in small aliquots up to 1 year at -20°C.

Extracellular buffer

110 mM NaCl

10 mM KCl

1 mM MgCl₂

1.5 mM CaCl₂

30 mM HEPES

10 mM glucose

Filter sterilize and store up to 1 month at 4°C

Intracellular buffer

25 mM NaHCO₃

120 mM KCl

1 mM KH₂PO₄

10 mM MgCl₂

20 mM HEPES

10 mM glucose

Filter sterilize and store up to 1 month at 4°C

Lysis buffer for immunoprecipitation, 2×

100 mM HEPES

200 mM NaCl

40 mM EDTA

4 mM EGTA

100 mM NaF

20 mM β-glycerophosphate

2 mM Na₃VO₄

2% (v/v) NP-40 or Triton X-100

Filter sterilize and store up to 6 months at 4°C

This buffer has a 10× concentration of EDTA to chelate the excess magnesium used in kinase reactions.

PBS/BSA

Dissolve 10 mg bovine serum albumin (BSA) in 100 ml PBS (APPENDIX 2). Filter sterilize using a low-protein-binding membrane (e.g., 0.2-μm Millipore GV filters) and store up to several months at 4°C.

Permeabilization buffer

120 mM KCl
25 mM NaHCO₃
5 mM HEPES
10 mM MgCl₂
1 mM KH₂PO₄
1 mM EGTA
300 mM CaCl₂
Filter sterilize and store up to several months at 4°C

This buffer has a free ionized calcium concentration of ~100 nM.

SDS-PAGE sample buffer, 2×

4 ml 0.125 M Tris·Cl, pH 6.8 (APPENDIX 2)
4 ml 10% (w/v) SDS
1 ml glycerol
0.4 ml 2-mercaptoethanol
0.6 ml H₂O
20 mg bromphenol blue
Filter sterilize and store up to 6 months at -20°C

The pH of the Tris buffer is measured at room temperature.

Streptolysin O solutions

600 U/ml stock solution: Make up a stock solution of 600 U/ml streptolysin O in PBS (APPENDIX 2) and snap-freeze in liquid nitrogen. Store at -70°C in small aliquots up to several months.

60 U/ml working solution: Immediately before the experiment, dilute the 600 U/ml stock solution to 60 U/ml with PBS/BSA (see recipe). Keep on ice until ready to use.

Two-dimensional-PAGE lysis buffer

20 mM Tris·Cl, pH 8.0 (APPENDIX 2)
0.3% (w/v) SDS
1% (v/v) 2-mercaptoethanol
Filter sterilize and store up to 6 months at -20°C

Two-dimensional-PAGE sample buffer

6 g urea
4 ml NP-40
0.5 ml 40% ampholytes of the required pH range (UNIT 10.3)
0.154 g DTT
1.12 ml H₂O
Filter sterilize and store in small aliquots up to 1 month at -70°C

IMPORTANT NOTE: *Use the best-purity urea available so that the concentrations of the breakdown products of urea—carbamic acid and ammonia—are minimized. This is important because proteins can be carbamylated in solution by these breakdown products, thereby altering the native protein's net charge and hence its isoelectric point.*

COMMENTARY

Background Information

The permeabilization of cells has become a popular method for the study of the regulation of signal-transduction mechanisms. In practice, similar methodologies can be used to study the different types of receptor-regulated intracellular communication mechanisms such as those regulated by heterotrimeric G-protein-coupled receptors, receptor tyrosine kinases, and cytokine receptors (Hawkins et al., 1983; Stephens and Downes, 1990; Alexander et al., 1992; Stutchfield and Cockcroft, 1993; Slowiejko et al., 1994; Stephens et al., 1994; Cunningham et al., 1995; Martys et al., 1995; Taylor et al., 1995). Although permeabilization protocols are very simple in theory, they are very much more problematic in practice. This is because different cells have different permeabilization properties; hence each time one applies permeabilization protocols to a new cell type, one must reexamine the experimental protocol in detail before reproducible data may be obtained. Similarly, different permeabilization-inducing agents are available, but they are not easily interchangeable, nor do they show predictable properties in the maintenance of signal-transduction processes in different cell types. The current agent of choice is streptolysin O, a bacterial toxin from the organism *Streptococcus pyogenes*. Streptolysin O has been used to study G-protein-coupled receptors and receptor tyrosine kinases; for these systems it appears to work quite reproducibly (Hawkins et al., 1983; Stephens and Downes, 1990; Alexander et al., 1992; Stutchfield and Cockcroft, 1993; Slowiejko et al., 1994; Stephens et al., 1994; Cunningham et al., 1995; Martys et al., 1995; Taylor et al., 1995).

A series of experiments utilizing cell permeabilization will not produce quick results. It will, however, produce data not normally obtainable using other experimental approaches.

A number of preliminary experiments must be performed to determine the optimal conditions for the permeabilization of intact cells, as well as for the time of addition for various reagents. Aspects to be considered include the following.

1. The kinetics of phosphorylation of the protein under study in intact cells must be determined.

2. The time for cold ATP/[γ - 32 P]ATP equilibration with intracellular processes—e.g., phosphatidic acid formation and phosphatidylinositol-4-phosphate or phosphatidylinositol-

4,5-bisphosphate synthesis—must be determined. The timing of the addition of [γ - 32 P]ATP to the permeabilization medium must be optimized—e.g., should the [γ - 32 P]-ATP be added at the same time as the permeabilization agent, or some time later?

3. If possible, the integrity and time course of the receptor activation of a known signal-transduction mechanism should be determined—e.g., one might determine whether MAP kinase is activated under the conditions of permeabilization indicated in the experiments performed under point 2, above, as well as the optimal time for activation and how long the response is detectable.

4. One should experiment with the timing of the addition of receptor agonists—e.g., can a receptor agonist be added at the same time as the permeabilization agent, at the same time as the [γ - 32 P]ATP, or some time after these reagents?

5. It should be determined whether the protein under study is phosphorylated under the same conditions as the reporter activity observed under point 3, above, in the permeabilized cells.

6. The timing of when inhibitors/activators or other reagents need to be added should be determined, as well as their concentration for optimal function in the permeabilized cells.

Once these issues have been resolved, experiments can easily be optimized and the phosphorylation event under study examined in detail.

Critical Parameters and Troubleshooting

A number of problems can be experienced when using permeabilization protocols. High-quality streptolysin O must be used, and each preparation of streptolysin O should be tested according to the supplier's instructions before use.

The timing for addition of all reagents to the cells must be strictly reproduced within experiments and between experiments. If large variability is noted between experiments and the various timings have been strictly observed, then the problem is almost always due to the streptolysin O preparation used. The solution should then be checked per the supplier's instructions.

Another problem that may be encountered is hydrolysis of [γ - 32 P]ATP by endogenous ATPases. This should always be assayed in any

experiment by determining the production of [³²P] inorganic phosphate during the time course of the experiment. If the ATP concentration is depleted by >10% during the course of the experiment, one can include general ATPase inhibitors in the permeabilization buffer. One such inhibitor is ouabain; however the potential exists for the ATPase inhibitor to inhibit whatever process is being studied; hence more experiments are required to determine the result empirically for each inhibitor to be used.

Anticipated Results

Once the optimal conditions for permeabilization, ATP addition, agonist addition, and inhibitor addition have been determined it is relatively straightforward to study a novel phosphorylation event. Initially, one can detect the phosphorylation of the protein under study as a result of receptor activation. If pharmacological tools are available—e.g., inhibitors/activators of specific protein kinases—it is possible to make a provisional identification of the kinase responsible for the phosphorylation event being examined. It is also possible to determine the proximity of the phosphorylation event to receptor activation.

Time Considerations

The entire permeabilization procedure can be completed in ~2 hr. After this time, the samples generated can be prepared for analysis by immunoprecipitation or directly analyzed by SDS-PAGE. Immunoprecipitation times are dependent on the antibodies being used, but high-affinity interactions should only require 1 to 2 hr of incubation at 4°C. Harvesting the immunoprecipitates with protein A-Sepharose will take 1 hr at 4°C, and the washing procedure then takes 30 min before the samples can be prepared for SDS-PAGE.

One-dimensional SDS-PAGE gels can generally be run in ~3 hr for a regular-sized gel or 1 hr for a minigel. The gels can then be stained to the required level in 1 to 2 hr before placing on a gel dryer for 1 to 2 hr. Once the gel has been stained and dried, it can be analyzed by either standard autoradiography or using a Phosphorimager. The exposure times required are dependent on the radioactivity contained in the protein bands. For regular autoradiography using an intensifying screen, one can see a band containing a few hundred cpm with an overnight exposure at -70°C. A Phosphor screen will give similar data with ~4 hr exposure.

For two-dimensional-PAGE analysis, the samples can be placed in the lyophilizer over-

night prior to electrophoresis. The two-dimensional-PAGE procedure will take one day to complete before the gel can be dried and analyzed by autoradiography or with a Phosphorimager.

Literature Cited

- Alexander, D.R., Brown, M.H., Tutt, A.L., Crumpton, M.J., and Shivnan, E. 1992. CD3 and CD2 antigen-mediated CD3 γ -chain phosphorylation in permeabilized human T cells: Regulation by cytosolic phosphatases. *Biochem. J.* 288:69-77.
- Castle, D.J. 1995. Purification of organelles from mammalian cells. *In Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 4.2.1-4.2.55. John Wiley & Sons, New York.
- Cunningham, E., Thomas, G.M., Ball, A., Hiles, I., and Cockcroft, S. 1995. Phosphatidylinositol transfer protein dictates the rate of inositol triphosphate production by promoting the synthesis of PIP₂. *Curr. Biol.* 5:775-783.
- Hawkins, P.T., Michell, R.H., and Kirk, C.J. 1983. A simple assay method for determination of the specific radioactivity of the γ -phosphate group of ³²P-labeled ATP. *Biochem. J.* 210:717-720.
- Martys, J.L., Shevell, T., and McGraw, T.E. 1995. Studies of transferrin recycling reconstituted in streptolysin O-permeabilized Chinese hamster ovary cells. *J. Biol. Chem.* 270(43):25976-25984.
- Slowiejko, D.M., Levey, A.I., and Fisher, S.K. 1994. Sequestration of muscarinic cholinergic receptors in permeabilized neuroblastoma cells. *J. Neurochem.* 62:1795-1803.
- Stephens, L.R. and Downes, C.P. 1990. Product-precursor relationships amongst inositol polyphosphates: Incorporation of [³²P]Pi into myo-inositol 1,3,4,6-tetrakisphosphate, myo-inositol 1,3,4,5-tetrakisphosphate, myo-inositol 3,4,5,6-tetrakisphosphate and myo-inositol 1,3,4,5,6-pentakisphosphate in intact avian erythrocytes. *Biochem. J.* 265:435-452.
- Stephens, L.R., Jackson, T., and Hawkins, P.T. 1994. Synthesis of phosphatidylinositol-3,4,5-trisphosphate in permeabilized neutrophils regulated by receptors and G-proteins. *J. Biol. Chem.* 268:17162-17172.
- Stutchfield, J. and Cockcroft, S. 1993. Correlation between secretion and phospholipase D activation in differentiated HL60 cells. *Biochem. J.* 293:649-655.
- Taylor, J.A., Karas, J.L., Ram, M.K., Green, O.M., and Seidel-Dugan, C. 1995. Activation of the high-affinity immunoglobulin E receptor Fc ϵ RI in RBL-2H3 cells is inhibited by Syk SH2 domains. *Mol. Cell. Biol.* 15:4149-4157.

Contributed by A. Nigel Carter
The Salk Institute for Biological Studies
La Jolla, California

**Analysis of
Protein
Phosphorylation**

18.8.19

Phosphopeptide Mapping and Identification of Phosphorylation Sites

UNIT 18.9

Many proteins in the cell are modified by phosphorylation. Protein phosphorylation can affect catalytic activity, localization of a protein in the cell, protein stability, and the ability of a protein to dimerize or form a stable complex with other molecules. There are several techniques available to find out whether or not a protein is modified by phosphorylation. To understand exactly why a particular protein becomes phosphorylated, it may be necessary to identify precisely which amino acid residues are phosphorylated. These residues can then be changed by site-directed mutagenesis, and the mutant protein can be examined for changes in activity, intracellular localization, and association with other proteins in the cell.

Studies geared towards understanding the phosphorylation of a particular protein usually start with labeling of the protein in intact cells followed by phosphoamino acid analysis (UNITS 18.2 & 18.3). Proteolytic digestion of a ^{32}P -labeled protein, followed by separation of the digestion products in two dimensions on a TLC plate, will give rise to a phosphopeptide map (see Basic Protocol 1). Phosphopeptides present on the TLC plate are visualized by autoradiography. These maps give information about the number of phosphate-containing peptides in the digest, and this is related to the number of phosphorylation sites present in the protein. Phosphopeptide maps can also be used to find out whether the sites of phosphorylation on a protein change upon treatment of cells with certain agents. Treatment of cells could lead to a reduction in the labeling of certain peptides and an increase in the labeling of other peptides present on the peptide map. This suggests that treatment results in a loss of phosphorylation at certain sites and an increase in phosphorylation at other sites. The identification of the sites of phosphorylation requires further analysis of the phosphopeptides present on these maps (see Support Protocol 1 and Basic Protocols 2 and 3). If 10 pmol of phosphorylated material can be generated, phosphopeptides can be purified by HPLC and identified directly by mass spectrometry or peptide microsequencing (see Support Protocol 2).

TRYPTIC PHOSPHOPEPTIDE MAPPING OF PROTEINS ISOLATED FROM SDS-POLYACRYLAMIDE GELS

**BASIC
PROTOCOL 1**

^{32}P -labeled proteins are resolved by SDS-PAGE and visualized following autoradiography. Protein bands are cut out of the dried gel and the protein of interest is isolated by extraction from the gel and TCA precipitation in the presence of carrier protein. The precipitated protein is oxidized in performic acid and digested with trypsin. The bicarbonate buffer is evaporated by several rounds of lyophilization, the tryptic peptide mix is spotted on a TLC plate and peptides are resolved by electrophoresis and chromatography in two dimensions and visualized by autoradiography.

Materials

- Samples containing ^{32}P -labeled proteins of interest (UNIT 18.2)
- Fluorescent ink or paint (can be obtained from most arts and crafts supply stores)
- 50 mM ammonium bicarbonate, pH 7.3 to 7.6 (when freshly prepared the buffer has a pH of ~7.5), and pH 8.0 (the pH drifts overnight to ~8.0, ideal for digestion with trypsin or chymotrypsin as in step 18)
- 2-mercaptoethanol
- 20% (w/v) SDS
- 50 mM ammonium bicarbonate, pH 7.3 to 7.6, containing 0.1% (w/v) SDS and 1.0% (v/v) 2-mercaptoethanol

**Analysis of
Protein
Phosphorylation**

2 mg/ml carrier protein (RNase A, BSA, or immunoglobulins) in deionized water
(store in aliquots at -20°C or -70°C)
100% (w/v) trichloroacetic acid (TCA)
96% ethanol, ice-cold
30% (w/v) hydrogen peroxide
98% (w/v) formic acid
1 mg/ml TPKK-treated trypsin (e.g., Worthington) in deionized water or 0.1 mM
HCl (store in aliquots at -70°C or under liquid nitrogen)
Electrophoresis buffers (see recipe): pH 1.9, 3.5, 4.72, 6.5, and 8.9
Green marker dye (see recipe)
Chromatography buffer (see recipe)

Single-edge razor blades or surgical blades
Scintillation counter appropriate for Cerenkov counting
1.7-ml screw-cap microcentrifuge tubes (Sarstedt)
Disposable tissue grinder pestles (Kontes)
Platform rocker
Tabletop centrifuge with swinging-bucket rotor
Glass-backed TLC plates (20×20 cm, 100 μm cellulose; EM Science)
Low volume adjustable pipet with long disposable tips made of flexible plastic,
e.g., gel-loading tips
Air line fitted with filter to trap aerosols and particulate matter
HTLE 7000 electrophoresis apparatus (CBS Scientific)
Polyethylene sheeting (35×25 cm; CBS Scientific)
Electrophoresis wicks (20×28 cm sheet of Whatman 3MM paper folded
lengthwise to give double thickness sheets of 20×14 cm)
Chromatography tank (CBS Scientific)
Fan for drying TLC plates
65 $^{\circ}\text{C}$ drying oven

Additional reagents and equipment for SDS-PAGE (UNIT 10.2) and autoradiography
(APPENDIX 3A)

Isolate ^{32}P -labeled protein by SDS-PAGE and recover from the gel

1. Resolve the samples containing the ^{32}P -labeled protein of interest by SDS-polyacrylamide gel electrophoresis (SDS-PAGE; UNIT 10.2).
2. Following electrophoresis, dry the gel, mark it around the edges with fluorescent ink, and expose to X-ray film (autoradiography; APPENDIX 3A).
3. Localize the protein of interest in the gel by aligning fluorescent markers around the gel precisely with their images on the film. Staple the film and gel together and place this sandwich film-side-down on a light box. Mark the position of ^{32}P -labeled protein bands on the back of the gel using a soft pencil or ballpoint pen (do not use a felt-tip marker).
4. Separate the gel from the film and cut out the protein bands from the individual lanes of the gel using a single-edge razor blade. Strip the paper backing from the gel slices and remove residual bits of paper by scraping gently with a razor blade. Try not to shave pieces from the gel because this will reduce recovery of the protein of interest. Place each gel slice in a 1.7-ml screw cap tube and determine the amount of radioactivity by Cerenkov counting in a scintillation counter.
5. Rehydrate each dry gel slice in 500 μl of 50 mM ammonium bicarbonate, pH 7.3 to 7.6, for 5 min at room temperature. Mash the gel slice using a Kontes tissue grinder pestle until no bits are seen when the tube is held up to the light. Add 500 μl of 50

mM ammonium bicarbonate, pH 7.3 to 7.6, 10 μ l of 2-mercaptoethanol, and 10 μ l of 20% SDS. Boil 2 to 3 min.

6. Extract the protein from the gel by incubation on a rocker for at least 4 hr at room temperature or for at least 90 min at 37°C.

For convenience, extractions can be done overnight.

7. Collect the gel slurry at the bottom of the tube by centrifuging 5 min at 500 \times g, room temperature, in a tabletop centrifuge with a swinging-bucket rotor. Transfer the supernatant to a new 1.5-ml microcentrifuge tube.

IMPORTANT NOTE: The brand of tube is important. From this point on use a brand of tubes that does not produce unwanted side reactions or retain too many cpm at the final transfer step (see Critical Parameters). The authors use microcentrifuge tubes from Myriad Industries.

8. Before starting the second elution, measure the volume of the first eluate and calculate the volume to be used for the second elution so that the volume of the combined eluates will measure \sim 1300 μ l. For the second elution, resuspend the gel pellet in the calculated appropriate volume of 50 mM ammonium bicarbonate containing 0.1% SDS and 1.0% 2-mercaptoethanol. Vortex, then boil 2 to 3 min and extract again by incubation on a rocker for at least 90 min.
9. Separate the gel from the eluate again by centrifugation as in step 7 and transfer the supernatant to the tube containing the first eluate.
10. To clear the combined eluate of gel slurry that has been inadvertently transferred, microcentrifuge 5 to 10 min at full speed, then transfer the supernatant to a new microcentrifuge tube. Before discarding the gel bits, monitor by Cerenkov counting to ensure that 60% to 90% of the 32 P-labeled protein has been extracted.

It is important to remove all gel fragments, and it may be worthwhile to repeat the last centrifugation step one more time.

11. Cool the eluates by placing them on ice. Add 20 μ g carrier protein (10 μ l of a 2 mg/ml stock), mix well, add 250 μ l ice-cold 100% TCA, mix well, and incubate for 1 hr on ice.
12. Collect the protein precipitate by microcentrifuging 5 to 10 min at full speed, 4°C. Decant the supernatant, then microcentrifuge again for 3 min at 4°C and aspirate the last traces of TCA.
13. Wash the TCA precipitate by adding 500 μ l ice-cold 96% ethanol, invert the tube a few times, and microcentrifuge 5 min at full speed, 4°C. Decant the bulk of the supernatant, microcentrifuge again for 3 min at 4°C, and aspirate the residual liquid. Air dry the protein pellet (do not lyophilize).
14. Monitor the precipitate by Cerenkov counting to make sure that the majority of the 32 P-labeled protein has been recovered.

There should be as many or slightly more cpm in the sample at this point as compared to the eluate (see step 10), since the liquid of the eluate will have quenched the counting somewhat.

Incubate with performic acid to achieve oxidation of the 32 P-labeled protein

15. Generate performic acid by incubating 9 parts formic acid with 1 part 30% hydrogen peroxide for 60 min at room temperature. Cool the performic acid by placing it on ice.

16. Resuspend the TCA-precipitated protein pellet in 50 μ l of the ice-cold performic acid and incubate 60 min on ice.
17. Add 400 μ l deionized water, mix, and freeze on dry ice. Evaporate the performic acid under vacuum in a SpeedVac evaporator.

It is extremely important to dilute and then freeze the sample before evaporating it, otherwise the elevated temperature of the SpeedVac evaporator may cause acid hydrolysis of the sample.

A sample (5% to 10% of the digest, at least 200 cpm) can be taken at this point for phosphoamino acid analysis. Lyophilize and proceed as described in UNIT 18.3.

Perform proteolytic digestion with trypsin

18. Resuspend the protein pellet in 50 μ l of 50 mM ammonium bicarbonate, pH 8.0, and add 10 μ g trypsin (10 μ l of a 1 mg/ml stock). Digest for 3 to 4 hr or overnight at 37°C.
19. Add a second 10 μ g aliquot of trypsin and digest again for 3 to 4 hr or overnight at 37°C.
20. Add 400 μ l deionized water, and lyophilize in a SpeedVac evaporator. Resuspend the pellet in 400 μ l deionized water and lyophilize again. Repeat these steps until at least four lyophilizations have been achieved.

At this stage there should be no visible pellet.

21. Resuspend the tryptic digest in 400 μ l electrophoresis buffer or deionized water.

The authors use pH 1.9 buffer or pH 4.72 buffer for samples that will be analyzed by electrophoresis at pH 1.9 and pH 4.72, respectively, and deionized water for samples that will be analyzed by electrophoresis at pH 8.9.

22. Clear the peptide mix of all particulate matter by microcentrifuging 5 to 10 min at full speed, transfer the supernatant to a new microcentrifuge tube, and lyophilize. Measure the amount of 32 P-radioactivity in the final sample by Cerenkov counting.

It is very important that there be no particulate matter in this final supernatant, and it may be worthwhile to repeat this centrifugation step one more time.

23. Resuspend the digest in at least 5 μ l of pH 1.9 electrophoresis buffer, pH 4.72 electrophoresis buffer, or deionized water, and collect the sample at the bottom of the tube by microcentrifuging 2 to 5 min at full speed.

Perform first-dimension electrophoresis on a TLC plate

24. Mark the sample and dye origins on the cellulose side of a glass-backed TLC plate with a small cross, using an extra-soft, blunt pencil, making sure not to perturb the cellulose layer (or alternatively mark on the reverse side with a permanent marker).

This is most easily done by placing the plate on top of a "life-size" marking template on a light box (Fig. 18.9.1).

25. Spot each sample onto the corresponding origin using an adjustable low-volume pipet fitted with a long flexible tip (round, gel-loading tips work well). Apply 0.2 to 0.5 μ l drops, and dry between applications using an air line fitted with a filter to trap aerosols and particulate matter, and a 1 ml syringe or a Pasteur pipet to focus the air flow.

Avoid touching the plate with the air nozzle or the pipet tip, since gouges on the cellulose may affect electrophoresis or chromatography. Ideally, at least 1000 cpm should be loaded onto the TLC plate. A brown ring around the circumference of the spot is normal.

26. Spot 0.5 μl of green marker on the dye origin at the top of the plate (Figs. 18.9.1 and 18.9.2).

This marker dye is green, but separates into its blue (xylene cyanol FF) and yellow (e-dinitrophenyllysine) components during electrophoresis (Fig. 18.9.2D).

27. Prepare the HTLE 7000 apparatus as described below, referring to Figure 18.9.3.
- Fill the buffer tanks so that the level of the electrophoresis buffer is ~ 5 cm deep. Place a sheet of polyethylene on the Teflon cover that protects and insulates the base and tuck the ends down between the base and the buffer tanks to hold the sheet in place.
 - Wet the electrophoresis wicks in electrophoresis buffer and slide them into the slots of the buffer tank with the folded edge up. Fold the ends of the wicks over the polyethylene sheet on the base. Place the second polyethylene sheet over the base, the electrophoresis wicks, and part of the buffer tank.

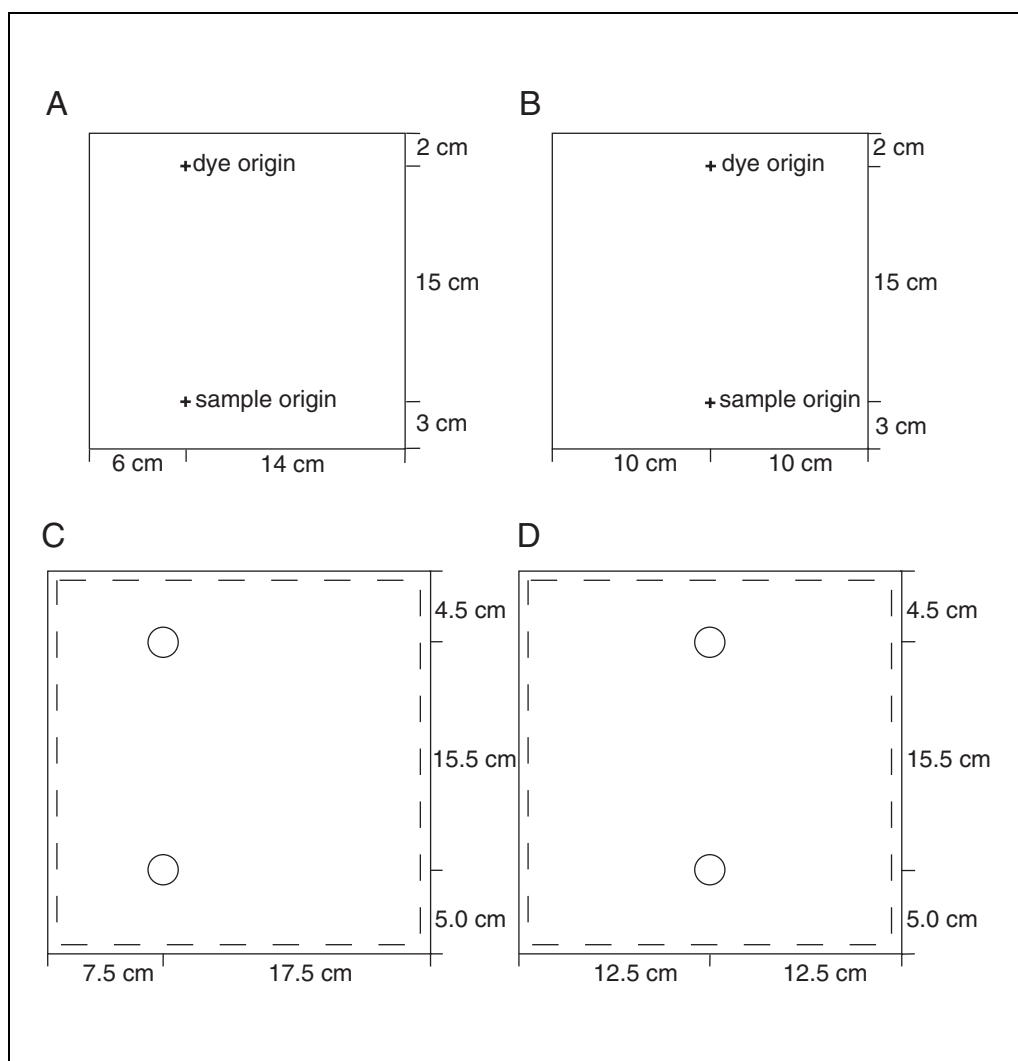


Figure 18.9.1 Sample and dye origins and blotter dimensions for separation of peptides at different pH values. **(A)** Location of the sample and dye origins for electrophoresis at pH 1.9 and pH 4.72 and **(B)** at pH 8.9. To mark a TLC plate, the plate is placed on top of a life-size template. This is then placed on top of a light box and the origins are marked on the cellulose side using a very blunt extra soft pencil. Dimensions of the blotter and the location of two holes that fit over the sample and dye origins are shown **(C)** for at pH 1.9 and 4.72 and **(D)** for pH 8.9.

- c. Place the second Teflon sheet and the neoprene pad on top, close the apparatus, and secure the lid with the two pins. Inflate the airbag by turning up the air pressure to 10 psi to squeeze out excess buffer from the electrophoresis wicks. Keep the air pressure on until ready to start the first run.
- d. After the samples have been loaded on the TLC plates (steps 25 and 26) and one is ready to start the electrophoresis, shut off the air pressure and open the apparatus. Remove the neoprene pad and the top Teflon and polyethylene sheets, and fold back the electrophoresis wicks. Wipe both polyethylene sheets dry with tissue paper.

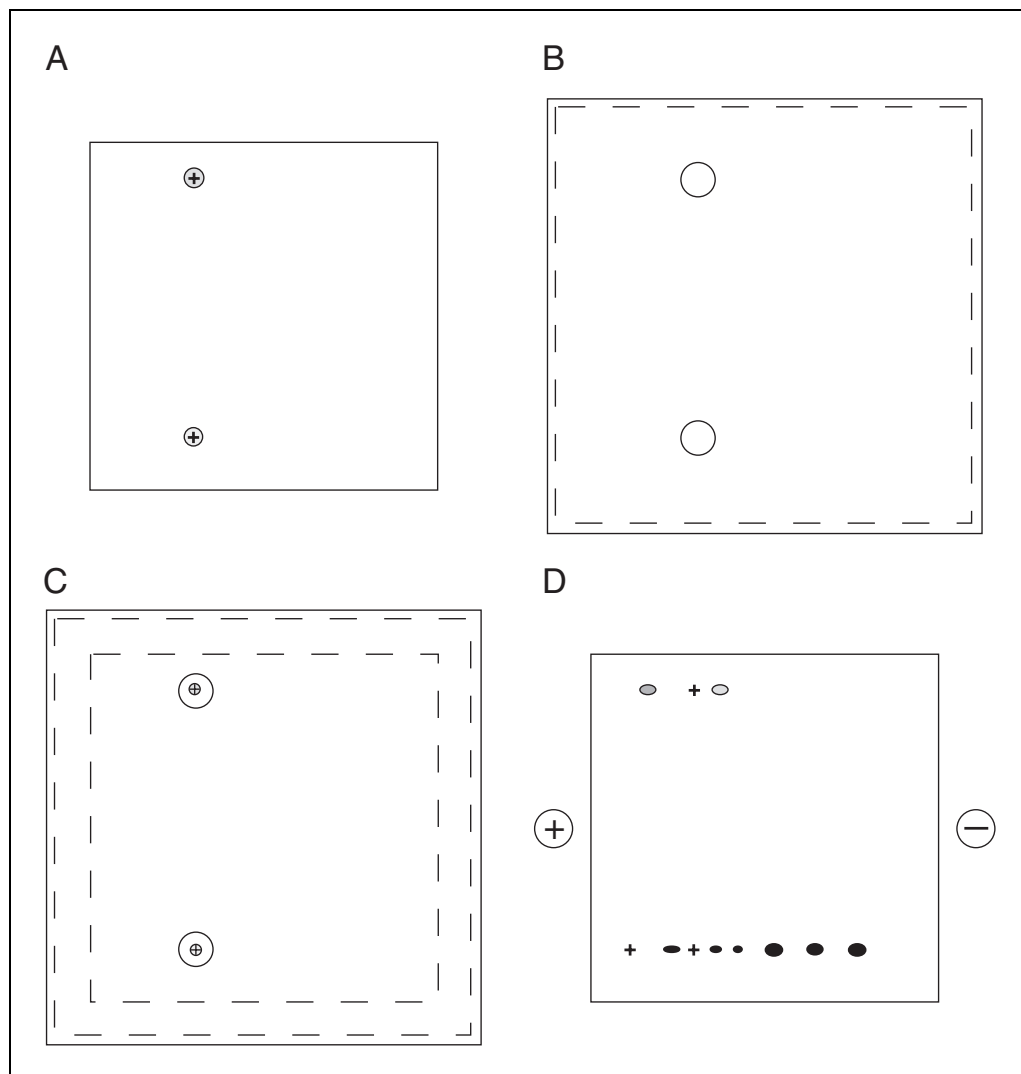


Figure 18.9.2 Separation of peptides by electrophoresis. **(A)** The sample and dye are spotted on their respective origin at the bottom and the top of the TLC plate as described in the text. **(B)** The blotter is soaked briefly in the electrophoresis buffer, and excess liquid is removed by blotting briefly on a piece of 3MM filter paper. **(C)** The TLC plate is wetted by placing the wetted blotter on top, with the sample and marker origins in the centers of the two holes. The blotter is pressed onto the TLC plate around the sample and marker origins to ensure uniform flow of electrophoresis buffer from the blotter towards the sample and marker origins. This will result in concentration of the sample and marker dye on their respective origins, and will improve resolution. The rest of the blotter is pressed with a flat hand onto the TLC plate, the blotter is removed and the plate is examined ; it should be dull-gray with no shiny puddles of buffer. Excess buffer should be allowed to evaporate or be blotted very carefully with tissue paper. The plate is placed on the apparatus and the electrophoresis run for 20 to 30 min at 1 kV. This results in separation of the peptides in the first dimension **(D, peptides shown in black)**. The position of the anode and cathode are indicated in panel D.

- e. Wet the TLC plate containing a sample as described in Figure 18.9.2 and place the plate on the polyethylene sheet covering the base. Fold the wicks over the plate so they cover ~1 cm of the plate at each end and carefully reassemble the apparatus as described above. Avoid lateral movement of the polyethylene sheet when it is in contact with the TLC plate. Secure the lid with the pins, inflate the airbag to 10 psi, turn on the cooling water flow, and start the electrophoresis.

At this point the authors concentrate the sample by wetting the TLC plates with electrophoresis buffer using a blotter with holes around the origin (Figs. 18.9.1 and 18.9.2). The blotter is made from two layers of Whatman 3MM paper, stitched together around the edges. The 1.50-cm holes that surround the origins (Fig. 18.9.1) are cut with a sharp cork borer. These blotters can be reused many times; it is best to keep a separate blotter for each buffer.

The buffer has to move with similar speed from the entire circumference towards the origin. The sample will inevitably streak if the buffer takes a long time to wet the spot, or moves unevenly through the spot.

- f. Perform electrophoresis for 20 to 30 min at 1.0 kV.

Electrophoresis results in separation of the peptides in one dimension (Fig. 18.9.2).

28. After completing the run, disassemble the apparatus and air dry the plate with the help of a fan for at least 30 min after electrophoresis is completed.

Do not dry in an oven as this will bake the peptides onto the cellulose, thereby interfering with the separation in the chromatography dimension.

Perform second-dimension separation by chromatography

29. Apply a drop (~0.5 μ l) of green marker dye in the left or right hand margin of the plate at the same level as the sample origin, avoiding the area that has been compressed by contact with the electrophoresis wick (Fig. 18.9.4). Place the dried plates in an almost upright position in the chromatography tanks with the appropriate chromatography buffer (see Critical Parameters), and replace the lid. Do not disturb or open a tank while chromatography is in progress. Allow the buffer to run to within 1 to 2 cm of the top of the plate.

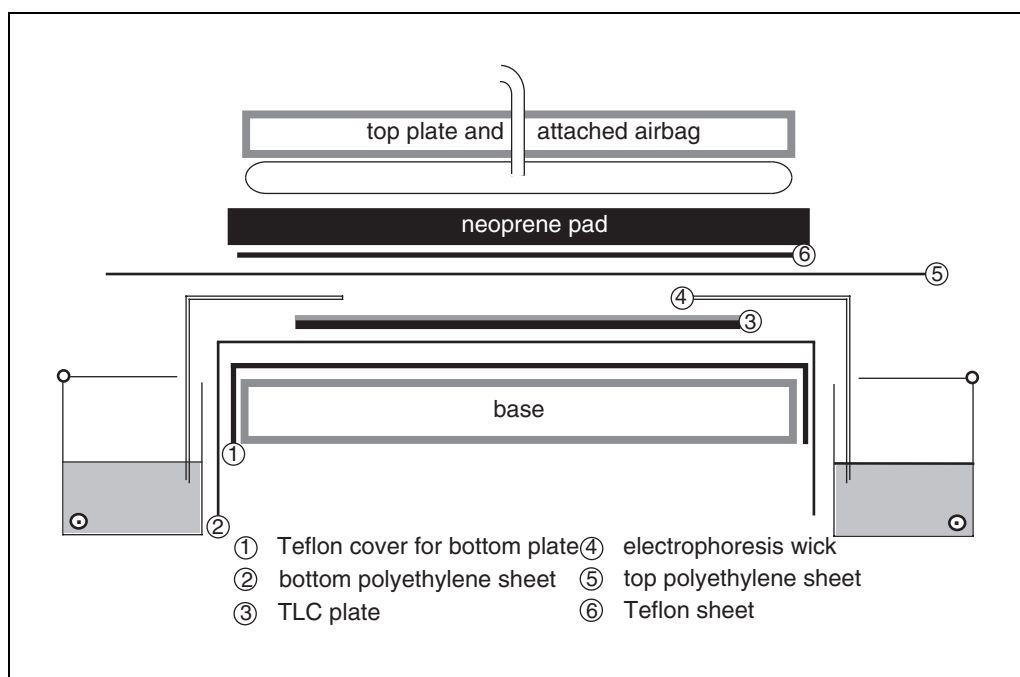


Figure 18.9.3 Preparation of the HTLE 7000 electrophoresis system.

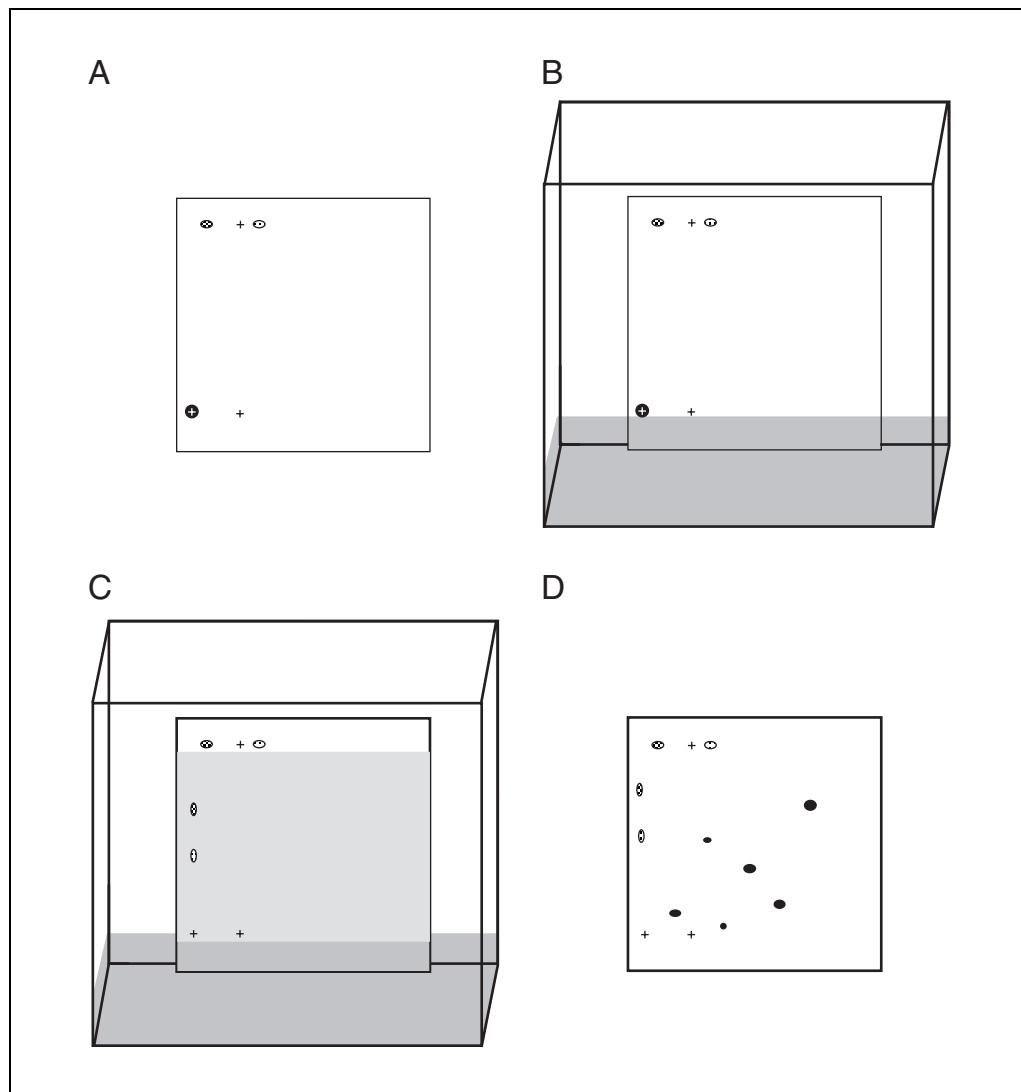


Figure 18.9.4 Separation of peptides in the second dimension by chromatography. After electrophoresis, air-dry the plate. A fan may be used to facilitate this. Add a small amount of green marker dye in the left hand (A) or right hand margin at the same level as the sample origin. Place the plate(s) almost upright in a chromatography tank, replace the lid and run the chromatography until the buffer front reaches to within 1 to 2 cm from the top of the TLC plate (B and C). This results in separation of the peptides in the second dimension (D, peptides shown in black). Open the tank, take out all plates, let the plates air dry, apply fluorescent ink at the margins of the plate and expose to X-ray film.

See Figure 18.9.4 for illustrations of these procedures. For information on selecting an appropriate chromatography buffer, see Critical Parameters.

30. Remove all TLC plates from the chromatography tank at the time the tank is opened. Allow the plates to dry for 1 hr in a fume hood or for 15 min in a 65°C oven.

Do not use the oven if peptides are to be extracted from these plates for further analysis.

31. Mark the dried plates with fluorescent ink around the edge; these reference marks can be used later to align the autoradiogram with the TLC plate. Expose the plates to X-ray film in the presence of an intensifier screen for autoradiography (APPENDIX 3A), or to a phosphorimager screen. If needed recover peptides for further analysis (see Support Protocol)

X-ray film may be presensitized for increased sensitivity (see APPENDIX 3A).

PROTEOLYTIC DIGESTION OF IMMOBILIZED PROTEINS

ALTERNATE PROTOCOL

Isolation of proteins from polyacrylamide gels is a lengthy and laborious procedure (see Basic Protocol 1). In addition, recoveries can be poor. Saving time can be important if one is working with the limited amounts of ^{32}P radioactivity present in proteins labeled in intact cells. As an alternative, proteins can be transferred to nitrocellulose or PVDF membranes, followed by digestion of the immobilized protein. The peptides are oxidized after digestion and elution from the membrane. Obviously, this approach is not a good choice for proteins that transfer with poor efficiency.

Additional Materials (also see Basic Protocol 1)

Methanol

0.5% (w/v) PVP-360 in 100 mM acetic acid (see recipe)

50 mM ammonium bicarbonate, pH 8.0

PVDF membrane (Immobilon P, Millipore) or nitrocellulose membrane (*UNIT 10.8*)

Saran Wrap or Mylar

Additional reagents and equipment for wet or semidry protein transfer (*UNIT 10.8*)

1. Resolve the ^{32}P -labeled samples by SDS-polyacrylamide gel electrophoresis (*UNIT 10.2*) and transfer the proteins to a PVDF or nitrocellulose membrane using a standard wet or semidry protein-transfer protocol (*UNIT 10.8*).
2. Air dry the membrane and wrap it in Saran Wrap or Mylar to prevent the membrane from sticking to the film, mark with fluorescent ink (see Basic Protocol 1, step 2), and expose the blot to X-ray film (autoradiography; *APPENDIX 3A*).
3. Align the film with the membrane using the fluorescent markers and their images on the film to identify the exact position of the protein of interest on the membrane (see Basic Protocol 1, step 3).
4. Cut out the strips of membrane containing the protein of interest with a single-edge razor blade, then cut this strip into several smaller pieces. Place all pieces of membrane containing a particular phosphate-labeled protein in a single microcentrifuge tube. Quantify the amount of radioactivity present on these strips of membrane by Cerenkov counting.
5. Rewet the membrane by adding 500 μl methanol, wash the membrane strips several times with deionized water, and incubate for 30 min at 37°C with 0.5% PVP-360 in 100 mM acetic acid.
6. Wash the membrane at least five times, each time with 1 ml deionized water, then two times, each time with 1 ml of 50 mM ammonium bicarbonate, pH 8.0.
7. Add enough 50 mM ammonium bicarbonate to cover the pieces of membrane (usually 200 to 400 μl), then add 10 μg TPCK-trypsin (10 μl of a 1 mg/ml stock). Incubate for at least 2 hr at 37°C.
8. Add another 10 μl aliquot of 1 mg/ml TPCK-trypsin and incubate again for 2 hr at 37°C.
9. Vortex briefly, then microcentrifuge briefly at full speed to collect all liquid at the bottom of the tube and transfer the supernatant to a fresh microcentrifuge tube. Rinse the membrane pieces with 500 μl of deionized water, microcentrifuge briefly, and add the rinse to the supernatant.

Analysis of Protein Phosphorylation

18.9.9

**SUPPORT
PROTOCOL 1**

10. Lyophilize in a SpeedVac evaporator and quantitate the elution of ^{32}P -labeled peptides by Cerenkov counting.

80% to 90% of the radioactivity should be in the eluate.

11. Oxidize the peptides by incubation in performic acid (see Basic Protocol, steps 15 to 17).
12. Add 500 μl of deionized water, lyophilize, and proceed with electrophoresis on TLC plate (see Basic Protocol, steps 21 to 31).

ISOLATION OF PHOSHOPEPTIDES FROM THE CELLULOSE PLATE

Individual phosphopeptides can be isolated from the TLC plate for further analysis. The location of the phosphopeptides on the TLC plate is determined by aligning the autoradiogram with the TLC plate. The cellulose containing the phosphopeptide of interest is scraped off the plate and sucked into a pipet tip fitted with a 25 μm filter. Peptides are eluted from the cellulose and lyophilized in a SpeedVac evaporator.

Materials

TLC plate with resolved phosphopeptides and corresponding autoradiogram (see Basic Protocol 1 or Alternate Protocol)
Electrophoresis buffer, pH 1.9 (see recipe)

Single-edge razor blades
1000- μl (blue) pipet tips
Small sintered polyethylene disk to fit inside blue tips (Kontes)
Glass rod or similar instrument to push filters into tips

Prepare the elution tips

1. Using a sharp razor blade, carefully remove the collar portion of the wide end of the blue tip. Trim ~ 3 mm off the small end of the tip as well.
2. Using a glass rod, push a sintered polyethylene disk in through the wide end of the blue tip until it fits snugly in the tip.

Use of a glass rod with the same diameter as the disk helps keep the disk straight, i.e., perpendicular to the length of the tip. Do not push the disk down too far or it will cause the tip to bulge out and crack. The sintered disk will serve as a barrier across the pipet tip to catch the cellulose as it is scraped from the plate. It therefore must fit securely in the tip, able to withstand the pull of the vacuum line.

This is the most difficult part of this protocol—but be consoled by the fact that once made, a good tip will last forever!

3. Test the placement of the disk in the tip by attaching a piece of tubing to the wide end of the tip, with the other end of the tubing connected to a vacuum line. Now apply a strong vacuum and use one finger to block off the small end of the elution tip. Examine to make sure that the sintered disk stays in place.

If the sintered disk stays in place, the elution tip is ready to use.

Mark the location of peptides to be eluted

4. Hold the TLC plate, cellulose side up, over a light box. Place the autoradiogram directly onto the cellulose layer of the plate, precisely aligning the reference marks on the plate with their images on the film.
5. Using a dark laboratory marker, trace the outline of the spot(s) of interest on the (glass) underside of the TLC plate.

Be conservative—when vacuuming two adjacent spots there should always be cellulose left on the plate between the two.

6. Remove the film, put the plate down on the lightbox, and use a soft lead pencil to trace the marker outline, this time on the cellulose side so it will be possible to see the outline without the benefit of the lightbox.

Vacuum the cellulose and elute the peptide(s)

7. Connect the elution tip to a vacuum via a piece of tubing and turn the vacuum on.
8. Use the small end of the elution tip to scrape the cellulose off a spot of interest; the cellulose will be sucked up against the filter barrier in the tip as it is scraped from the plate. When the spot is completely removed from the plate, ease the tubing off the wide end of the elution tip, keeping the small end upright.

The same spot can be vacuumed from multiple plates into one elution tip. However, after repeated use, the small end of the elution tip becomes “dull,” and it becomes increasingly difficult to scrape the cellulose from the plate. When this happens, use a razor blade to trim a thin sliver of plastic from the small end of the tip to recreate the sharp edge.

9. Place the elution tip into a 1.5-ml microcentrifuge tube with the wide end down and the side of the sintered disk containing the vacuumed cellulose up.

The elution tip now becomes a little column.

10. Immediately pipet 100 μ l of electrophoresis buffer, pH 1.9 (elution buffer) onto the cellulose; let this soak in while other spots are vacuumed from the plate.

The elution buffer used here should be pH 1.9. If pH 1.9 buffer fails to elute all the phosphopeptide, try deionized water.

11. When all spots have been vacuumed and the last one has been left to soak in buffer for ~5 min, place the microcentrifuge tubes, tips and all, into a microcentrifuge. Run the microcentrifuge at full speed for ~3 sec, then shut it off. Pipet another 100 μ l elution buffer onto the cellulose in each tube and let it sit and soak for another 5 min before centrifuging it through the column. Repeat the elutions five times to give 600 μ l of eluate in each tube.

This is usually enough to elute >90% of the radioactivity from the cellulose.

12. Remove the elution tip from each of the microcentrifuge tubes, being careful to leave all the eluate in the tube (some may cling to the sides of the tip as drops, which should be removed and added back to the contents of the tube). Save the elution tip. If eluting more than one spot, keep track of which tip was used for which peptide.
13. Clarify the eluate(s) by microcentrifuging 5 min at full speed (a small cellulose pellet will be visible after centrifugation; its size will depend on how snugly the sintered disk fits into the elution tip). Transfer the supernatant to a fresh microcentrifuge tube.

It is very important to remove all traces of cellulose at this point, as contamination of the phosphopeptide with cellulose can ruin further analyses.

14. Count both the eluate and the “empty” elution tips by Cerenkov counting.

~90% of the radioactivity should be in the eluates, with little remaining in the cellulose left in the tips.

Given the pain and frustration involved in their manufacture, a good elution tip should be saved and reused. To clean these tips, apply a vacuum to the small end of the tip and suck the cellulose out (into a vacuum flask) while aspirating ~10 ml elution buffer or deionized water through the tip to rinse it. Dry and then count the tips on a scintillation counter before reusing them.

15. Lyophilize the eluates in a SpeedVac, then count them by Cerenkov counting.

The counts here should be slightly higher than those of the liquid eluate. The number of cpm in this final sample of eluted peptide will often determine how it can be analyzed further.

DETERMINATION OF THE POSITION OF THE PHOSPHORYLATED AMINO ACID IN THE PEPTIDE BY MANUAL EDMAN DEGRADATION

If insufficient material is available for direct sequencing, a manual Edman degradation of the peptide can be performed to determine at which position the phosphorylated amino acid is present in the peptide. During each cycle of Edman degradation, the most amino-terminal amino acid residue is released from the peptide, and a sample from the reaction mixture is taken after each cycle. Phosphoserine or phosphothreonine is released as a derivative of serine or threonine and free phosphate; in contrast, phosphotyrosine is released as the anilinothiazolinone derivative of phosphotyrosine. Free phosphate and the PTH derivative of phosphotyrosine can be separated from the peptide by electrophoresis on a TLC plate. This approach indicates at which cycle the radioactivity and thus the phosphorylated amino acid is released from the peptide.

Materials

- Eluted phosphopeptide (see Support Protocol 1)
- 5% (v/v) phenylisothiocyanate (PITC) in pyridine
- 10:1 (v/v) heptane/ethyl acetate—mix 10 parts heptane with 1 part ethyl acetate
- 2:1 (v/v) heptane/ethyl acetate—mix 2 parts heptane with 1 part ethyl acetate
- 100% (w/v) trifluoroacetic acid (TFA)
- Electrophoresis buffer, pH 1.9 (see recipe)
- 200 to 500 cpm ^{32}P (prepared by diluting ^{32}P orthophosphate with deionized water)
or 2 mg/ml PTH-phosphotyrosine (see recipe)
- Microcentrifuge tubes (Myriad Industries)
- 45°C water bath
- Scintillation counter appropriate for Cerenkov counting
- Glass-backed TLC plates (20 × 20 cm, 100 μm cellulose; EM Science)
- 65°C drying oven or fan
- Additional reagents and equipment for electrophoresis of peptides on a TLC plate (see Basic Protocol 1 and Figure 18.9.3) and autoradiography (APPENDIX 3A)

Determine experimental parameters

1. Decide the number of cycles to be run based on the list of candidate peptides.

The number of cycles is designated as X. The starting volume for each cycle will be 20 μl.
2. Dissolve the eluted peptide in 20 μl deionized water in what will be called the reaction tube (a microcentrifuge tube).
3. Remove a sample equal to 20/(X + 1) μl to a new tube; this is the starting material sample. Store this at 4°C.

This sample will be lyophilized with the other cycle fractions at a later point.

Perform the Edman reactions

4. Add enough deionized water to the reaction tube to restore the volume to 20 μl. Count the sample at this point:
 - a. to insure that the expected number of cpm have in fact been removed from the initial sample (as the starting material sample);

- b. to check the cpm at the start of each given cycle.
5. Add 20 μl of 5% phenylisothiocyanate in pyridine to each reaction tube, vortex well, spin briefly in a microcentrifuge to collect the sample at the bottom, and incubate at 45°C for 30 min.
 6. Add 200 μl of 10:1 heptane/ethyl acetate to each reaction tube and vortex for 15 sec. Microcentrifuge 1 min at full speed to separate the two phases.
The pyridine will partition into the (upper) organic phase.
 7. Carefully remove the upper organic phase using a plastic transfer pipet. Reextract the (bottom) aqueous phase a second time with 10:1 heptane/ethyl acetate as in step 6.
 8. Extract the aqueous phase two more times as in step 6, this time using 2:1 heptane/ethyl acetate.
 9. Freeze the aqueous phase on dry ice and lyophilize in a SpeedVac evaporator.
 10. Dissolve the dried sample in 50 μl of 100% trifluoroacetic acid (TFA) and incubate this at 45°C for 10 min.
 11. Lyophilize the sample in a SpeedVac evaporator.
 12. Count the sample by Cerenkov counting.
There should be the same number of cpm as at the beginning of the cycle (i.e., at step 4 in this case).
 13. Add 20 μl deionized water to the reaction tube, vortex, and microcentrifuge briefly. Remove a portion for analysis of the first-cycle products that is equal to $20/X$. Store this at 4°C with the starting material sample.
 14. Add deionized water to restore the sample volume to 20 μl to start the second cycle. Repeat steps 5 to 12.
 15. After the second cycle, add 20 μl deionized water, resuspend the remaining sample, and remove $20/(X-1)$ μl to a new tube for analysis of the second-cycle products. Repeat steps 4 to 12.
 16. Continue repeating steps 4 to 12 until the desired number of cycles have been run.
For each new cycle, the amount of the sample to be removed is $20/X-Y$ where Y equals the cycle number minus 1.

Analyze the Edman products

17. Lyophilize all samples to dryness in a SpeedVac evaporator. Count all final samples by Cerenkov counting,
18. If lyophilized, dissolve the samples in 5 μl of pH 1.9 electrophoresis buffer or deionized water. Microcentrifuge 2 min at maximum speed to bring down any insoluble material.
Alternatively, if the sample volumes removed after each cycle are small enough, skip steps 17 and 18 and load the samples directly onto the TLC plate.
19. Spot all samples from the analysis of a given phosphopeptide at least 1 cm apart on a line of origins running vertically through the center of the TLC plate (Fig. 18.9.5). As a marker, depending on the phosphoamino acid content of the peptide under investigation, spot 50 to 200 cpm of [^{32}P]phosphate or 1 to 2 μg PTH-phosphotyrosine (0.5 to 1.0 μl of 2 mg/ml PTH-phosphotyrosine) at an origin on that same vertical

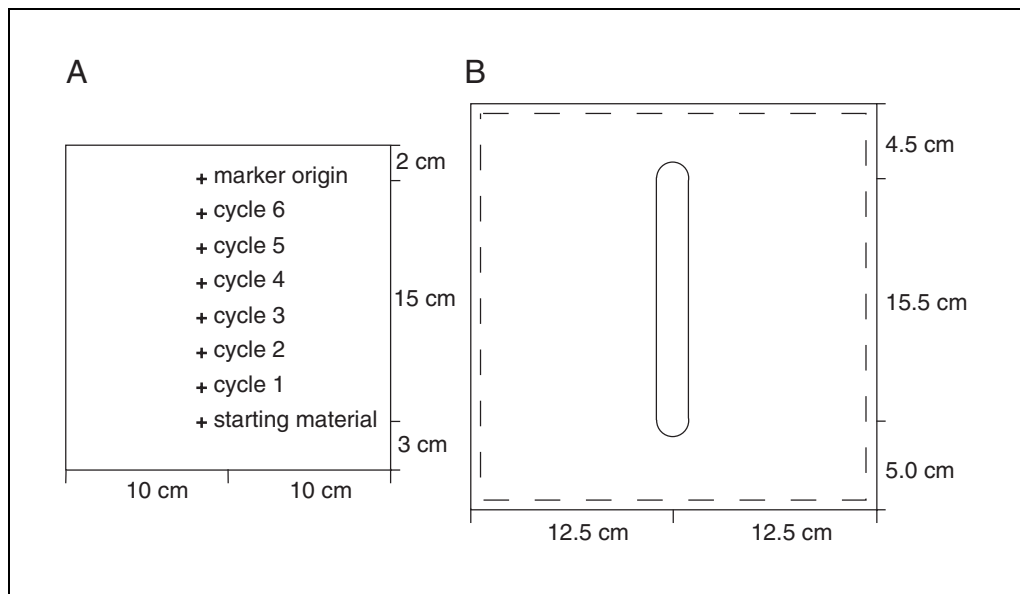


Figure 18.9.5 Sample and dye origins and blotter dimensions for analysis of manual Edman degradation products at pH 1.9. **(A)** Location of the sample and standard origins. To mark a TLC plate, the plate is placed on top of a life-size template on top of a light box and the origins are marked on the cellulose side using a very blunt extra-soft pencil. **(B)** Dimensions of the blotter and the location of the slot that fits over multiple sample and marker origins. The blotter is soaked in electrophoresis buffer, blotted with a sheet of Whatman paper to remove most of the buffer, and placed on top of the TLC plate so that the origins are in the middle of the slot.

line. Load $\frac{1}{3}$ to $\frac{1}{2}$ μl of sample at a time, air drying the sample between applications (see Basic Protocol 1, step 25).

20. Wet the plate as described in Figure 18.9.5.
21. Prepare the HTLE 7000 apparatus and electrophorese the samples for 25 min at 1.0 kV in pH 1.9 electrophoresis buffer (see Basic Protocol 1 and Figure 18.9.3).
22. After drying the plate (either in a 65°C oven or with a fan) mark it appropriately with radioactive or fluorescent markers and expose it to presensitized film with an intensifying screen at -70°C (autoradiography; APPENDIX 3A).

**BASIC
PROTOCOL 3**

**DIAGNOSTIC SECONDARY DIGESTS TO TEST FOR THE PRESENCE OF
SPECIFIC AMINO ACIDS IN THE PHOSHOPEPTIDE**

Further information about a phosphopeptide of interest can be obtained by digestion with sequence-specific proteases or cleavage by site-specific chemicals. After incubation with a protease or chemical, the peptide is analyzed by separation in two dimensions on a TLC plate. A change in mobility upon treatment with a particular reagent indicates that the peptide was susceptible to cleavage, and consequently that the amino acid or amino acid sequence that confers susceptibility to cleavage by this reagent must be present in the peptide.

Materials

- Eluted phosphopeptide (see Support Protocol 1)
- Enzyme to be used for digestion and appropriate buffer (see Table 18.9.1)
- 2-mercaptoethanol

Table 18.9.1 Specificities and Digestion Conditions for Enzymes and Other Cleavage Reagents

Enzyme or reagent	Specificity ^a	Digestion conditions	Comments
TPCK-trypsin	K—X; R—X	pH 8.0-8.3	Does not cut K/R-P; cuts inefficiently at K/R-X-P.Ser/P.Thr and K/R-D/E; cuts wells at K/R-P.Ser/P.Thr; cuts X-K/R-K/R-K/R incompletely
α-Chymotrypsin	F—X; W—X; Y—X	pH 8.3	Does not cleave F/W/Y-P or P.Tyr-X
Thermolysin	X—L; X—I; X—V	pH 8.0, 1 mM CaCl ₂ , 55°C	Will recognize most apolar residues to some extent; CaCl ₂ may affect the electrophoretic mobility
Proline-specific endopeptidase	P—X	pH 7.6	—
Cyanogen bromide (CNBr)	M—X	50 mg/ml CNBr in 70% formic acid, 90 min, 23°C	CNBr is toxic; will only cleave unoxidized methionine
V8 protease	E—X	pH 7.6	V8 will not cleave at every E in whole proteins; does give a consistent pattern
Endoproteinase Asp-N	X—CSO ₃ H; X—D	pH 7.6	Will cleave X-E at high enzyme concentrations
Formic acid	D—P	70% formic acid, 37°C, 24-48 hr	—

^aThe dash indicates the cleavage site. See APPENDIX 1C for definitions of the one-letter abbreviations for amino acids.

Electrophoresis buffer of appropriate pH (see recipe)

Water bath at appropriate temperature for enzyme digestion

Glass-backed TLC plates (20 × 20 cm, 100 μm cellulose; EM Science)

Additional reagents and equipment for chromatography and electrophoresis of phosphopeptides (see Basic Protocol, steps 24 to 31)

1. Dissolve the eluted phosphopeptide in 50 μl of the appropriate buffer in a microcentrifuge tube and microcentrifuge briefly to collect all solution at the bottom of the tube. Check the pH of the peptide solution by spotting 1 μl on a piece of pH paper to be sure that this final pH will allow enzyme activity. If the buffer's pH has been altered dramatically by addition of the peptide, adjust it before adding enzyme.
2. Remove a portion of the sample (usually 50%) to be run both as an undigested control and as a mix with a portion of the digested sample.
3. Add 1 to 2 μg enzyme to the portion of the sample to be digested, vortex, and concentrate the sample in the bottom of the tube by microcentrifuging briefly.
4. Incubate all tube(s) in a water bath at the appropriate temperature for at least 1 hr.
5. Add another aliquot of enzyme and continue the incubation step for an additional hour.
6. Add 1 μl of 2-mercaptoethanol to each sample and boil 5 min to inactivate the enzyme

Do this to all samples to ensure uniformity of sample preparation.

It is necessary to completely inactivate the enzyme prior to loading the sample on the plate when analyzing a mix of digested and undigested peptide, since the undigested sample may be rapidly digested when the two samples are mixed.

7. Lyophilize the samples in a SpeedVac evaporator.
8. Resuspend the samples by vortexing vigorously in 6 μ l of electrophoresis buffer of the appropriate pH. Microcentrifuge at full speed to bring down insoluble material.
9. Load half of the undigested sample on a single TLC plate. Load half of the digested sample on each of two TLC plates; on one of these load the other half of the corresponding undigested sample as a mix.
10. Perform electrophoresis and chromatography on the plate as described above in Basic Protocol 1, steps 24 to 31. Based on the position where the particular phosphopeptide being analyzed ran in the original map, choose a pH and running time that will allow good separation of the peptide from its potential cleavage products but will ensure retention of the smaller cleavage products on the plate.

**SUPPORT
PROTOCOL 2**

**PREPARATION OF PHOSHOPEPTIDES FOR MICROSEQUENCE
DETERMINATION OR MASS SPECTROMETRY**

The following is a general protocol and list of considerations for generating enough material for analysis by mass spectrometry or microsequencing starting with either intact cells or an in vitro system.

1. Optimize ^{32}P labeling of the protein of interest.

If the site of interest is seen only in stimulated cells, a time course of phosphorylation following stimulation may be helpful, as would determination of the optimal concentration of agonist.

If an in vitro system is being employed, determine the optimum conditions (time, and ratio of kinase and substrate concentrations) for the kinase reaction. Include 1 mM cold ATP in the reactions to maximize the stoichiometry. UNIT 18.7, which deals with in vitro phosphorylation reactions, provides a detailed discussion of these parameters and how to manipulate them.

2. Calculate the number of cells or amount of substrate needed to isolate 10 pmol phosphorylated material.

Even under optimal conditions, it is often not possible to achieve more than 25% stoichiometry of phosphorylation in vitro. In intact cells, the stoichiometry may be even less. It cannot hurt to overestimate the amount of starting material required, as the losses taken during the isolation procedures will always exceed expectation.

3. When calculating how to scale up the reactions, consider the following points.

- a. The radioactivity of these samples is only used for visualization purposes—i.e., to determine which gel band to isolate, which phosphopeptide to isolate from the TLC plate, and which HPLC fraction(s) to use for final analysis. Thus, the majority of the material can be unlabeled, as only \sim 1000 cpm per map spot are necessary for analysis at the time when the preparative HPLC is run. When isolating overexpressed protein from cells, labeling only 2 or 3 dishes of the 20 needed to generate enough material may be sufficient. When using an in vitro system, perform only one reaction using γ [^{32}P]-ATP (include only the very minimum amount of cold ATP necessary for kinase activity) to generate the labeled material. To generate enough material for further analysis, perform an additional kinase reaction with unlabeled ATP only. For visualization, mix the labeled and unlabeled samples before resolving them by SDS-PAGE.

- b. The efficiency of protein elution decreases as the amount of gel increases, so try to keep the number of lanes on the preparative gel(s) to a minimum. About four lanes/slices of acrylamide gel can be successfully extracted per tube.
- c. If ~20 μg substrate protein is present in each elution sample, it will not be necessary to add carrier protein at the TCA precipitation step (see Basic Protocol 1 step 11). This will result in a cleaner sample, as the tryptic fragments of the carrier protein will be eliminated from the mix of fragments run on the TLC plate.
- d. The 20 μg trypsin used to digest map samples in Basic Protocol 1 is in vast excess. While it is important that the digestion go as far to completion as possible, it is probably not necessary to scale up the amount of trypsin used. Instead, consider pooling several like samples at the performic acid digestion step (at the end of the 60 min incubation in order to give the protein the maximum time to dissolve). Adding 1 to 5 μg (total) trypsin to even 50 μg of protein for digestion is not unreasonable. Minimizing the amount of trypsin used will in turn minimize the amount of “extra” protein loaded on each TLC plate and ensure that the sample does not streak due to overloading.
- e. Determine the number of TLC plates to be run based on the amount of total protein to be analyzed—total protein includes the amount of trypsin and the amount (if any) of carrier protein used as well as the amount of the protein of interest. Even though the capacity of the TLC plates is ~100 μg , to ensure good separation, no more than 60 μg of protein should be run on each plate.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Chromatography buffers

Phosphochromatography buffer:

750 ml *n*-butanol
500 ml pyridine
150 ml glacial acetic acid
600 ml deionized water
Store at room temperature

Isobutyric acid buffer:

1250 ml isobutyric acid
38 ml *n*-butanol
96 ml pyridine
58 ml acetic acid
558 ml deionized water

Regular chromatography buffer:

785 ml *n*-butanol
607 ml pyridine
122 ml glacial acetic acid
486 ml deionized water

Store all of the above buffers up to 6 months at room temperature.

Electrophoresis buffers

For each of these buffers, mix well and check the pH. Record the pH and the date on the bottle; if the pH is more than a tenth of a unit off, remake the solution. Do not adjust the pH. Store all buffers at room temperature.

pH 1.9 buffer:

50 ml formic acid (88% w/v)
156 ml glacial acetic acid
1794 ml deionized water

pH 3.5 buffer

100 ml glacial acetic acid
10 ml pyridine
1890 ml deionized water

pH 4.72 buffer

100 ml *n*-butanol
50 ml pyridine
50 ml glacial acetic acid
1800 ml deionized water

pH 6.5 buffer

8 ml glacial acetic acid
200 ml pyridine
1792 ml deionized water

pH 8.9 buffer

20 g ammonium carbonate
2000 ml deionized water

Green marker dye

Prepare a solution containing 5 mg/ml ϵ -dinitrophenyllysine (yellow) and 1 mg/ml xylene cyanol FF (blue) in pH 4.72 electrophoresis buffer (see recipe) diluted 1:1 with deionized water. Store up to 1 year at room temperature.

PTH-phosphotyrosine, 2 mg/ml

Combine 20 μ l of 100 mg/ml phosphotyrosine with 20 μ l of 5% (v/v) phenylisothiocyanate in pyridine. Incubate 30 min at 45°C. Extract twice with 200 μ l of 10:1 (v/v) heptane/ethyl acetate, then twice with 200 μ l of 2:1 (v/v) heptane/ethyl acetate (see Basic Protocol 2, steps 6 to 8, for extraction technique). Freeze the aqueous phase and lyophilize in a SpeedVac evaporator. Dissolve the sample in 0.1 N HCl, incubate 20 min at 80°C, and lyophilize again in a SpeedVac evaporator. Dissolve in 1 ml pH 1.9 buffer (see recipe for electrophoresis buffers).

PVP-360 in 100 mM acetic acid

0.5 g PVP-360 (Sigma)
575 μ l glacial acetic acid
99.4 ml deionized water
Store up to 1 year at room temperature

COMMENTARY

Background Information

Phosphopeptide mapping is a very sensitive technique that can help the investigator answer a variety of questions about a protein of interest. For some, phosphopeptide mapping is a tool to find out whether a particular protein is phosphorylated on one or more sites. This question can be answered by simply running a phosphopeptide map of the protein labeled in living cells. Other investigators want to know whether the increase in phosphorylation seen when cells are treated with a particular agent is restricted to one or more specific sites or whether it is evenly distributed over all phosphorylation sites present in the protein. Finally, detailed analysis of phosphopeptides isolated from a TLC plate can be used to identify the residues that are phosphorylated in a protein of interest.

Several different strategies may be followed to identify the phosphorylation site represented by a particular spot on a phosphopeptide map. This is most definitively accomplished by eluting phosphopeptides from cellulose plates for either direct sequencing or for analysis by mass spectrometry (see Support Protocol; Fischer et al., 1991; Wang et al., 1993; Mitchelhill et al., 1997). While these two techniques require expensive instruments and expertise not found in most laboratories, such analysis can often be arranged by collaboration. However, sometimes it is not possible to take advantage of these techniques, since they require 1 to 10 pmole of material for analysis. For a 50-kDa protein one would need 0.1 to 1.0 μg starting material, assuming 50% recovery and 100% stoichiometry of phosphorylation at the site of interest. Use of an *in vitro* phosphorylation system that mimics the situation in intact cells will simplify matters greatly. Further considerations and strategies for preparation of samples for these two techniques are discussed at the end of this chapter in Support Protocol 2.

Another approach to phosphorylation-site identification is to make an educated guess as to the identity of the site. Clues to a site's identity include phosphoamino acid analysis of the individual phosphopeptide (UNIT 18.3); the result of manual Edman degradation of a phosphopeptide providing the cycle at which the phosphate is released and thus the position of the phosphorylated residue in the peptide (see Basic Protocol 2); and secondary enzymatic digests of the phosphopeptide that can be used in a diagnostic sense to determine the presence of other specific amino acids in the peptide (see

Basic Protocol 3). All three of these techniques are easily accomplished in a laboratory that is already set up for phosphopeptide mapping. The first step in all three is the isolation of the phosphopeptide from the cellulose plate (see Support Protocol 1). The validity of one's guess can be tested by phosphopeptide mapping of a mutant protein lacking a phosphate acceptor at the site in question. Alternatively the guess can be substantiated by synthesizing the tryptic phosphopeptide and testing it for comigration with the phosphopeptide isolated from the peptide map.

Critical Parameters and Troubleshooting

Generating phosphopeptide maps

Keep in mind that the sort of analyses presented throughout this unit will give information regarding only the acid-stable forms of phosphoamino acids (i.e., phosphoserine, phosphothreonine, and phosphotyrosine) and will essentially ignore other forms such as phosphohistidine and phosphoaspartate, should they be present.

Carrier Protein. The authors prefer to use RNase as carrier protein during TCA precipitation, particularly when analyzing proteins labeled in intact cells, because it degrades ^{32}P -labeled RNA species that may have copurified with the protein of interest. The nucleotides generated by the degradation of RNA do not precipitate in TCA.

Cleaving the protein. In order to generate a phosphopeptide map, the ^{32}P -labeled protein needs to be cleaved into smaller fragments that can be separated by electrophoresis and chromatography on TLC plates. To do this requires an enzyme or chemical agent that cleaves reproducibly and with a certain frequency. If not enough cleavage sites are present, the fragments generated will be too large and will not be separated easily by electrophoresis and chromatography on TLC plates. In addition, large fragments may contain multiple phosphorylation sites. This leads to maps that are less informative and more difficult to analyze. The authors routinely use trypsin and chymotrypsin. Other reagents are available (Table 18.9.1), but most of them cut much less frequently and some of them do not work very efficiently on full-length proteins.

Removing ammonium bicarbonate. Following digestion, repeated cycles of lyophilization

are carried out to evaporate all ammonium bicarbonate. The presence of salts in the sample will interfere with the electrophoretic separation of the peptides. After lyophilization, the protein digest appears as an invisible film at the bottom of the tube. The presence of any crystalline material indicates the presence of salts, most likely ammonium bicarbonate that can be removed by additional rounds of lyophilization.

Controlling oxidation. Both cysteine and methionine can give rise to several oxidized derivatives. The oxidation state of these amino acids affects the mobility of peptides during chromatography, resulting in separation of oxidation state isomers. This complicates the interpretation of the phosphopeptide map. To prevent this, the protein or peptides are oxidized by incubation in performic acid at 4°C. Incubation at higher temperatures may give rise to unwanted side reactions and should be avoided.

Elution and TCA precipitation or transfer to a membrane? In Basic Protocol 1, the ³²P-labeled protein is isolated from a small slice of a dried polyacrylamide gel by rehydrating and grinding up the gel followed by elution in a buffer containing SDS and 2-mercaptoethanol. The protein is subsequently TCA precipitated, oxidized, and digested with trypsin. This is a time-consuming and laborious procedure. Yields are variable and usually not better than 50%. The alternative is to transfer the protein to a PVDF membrane; any unoccupied protein-binding sites on the strips of membrane containing the protein of interest are blocked by incubation with PVP-360 in acetic acid before incubation with trypsin. Most peptides dislodge from the membrane during the digestion. This protocol is much faster and less laborious, and does not require the use of additional carrier proteins that may lead to overloading of the TLC plate and to streaky maps. Obviously this method is a poor choice for proteins that transfer poorly from the polyacrylamide gel to a membrane. In addition, it is possible that certain peptides that are generated during protease digestion retain a high affinity for the membrane and therefore fail to elute. If those peptides contain a phosphorylation site, this site will not be represented on the peptide map. This can lead to misinterpretations of the results. It is therefore advisable to first compare maps generated with Basic Protocol 1 and the Alternate Protocol. If these maps are identical, and if the protein transfers well from the gel to the membrane, the Alternate Protocol should be the protocol of choice.

Amount of sample. The authors like to load at least 1000 cpm on a plate for a peptide map. If the final sample has many more than 1000 cpm and a “pretty-looking” map is desired, it may be better to load only a fraction of the sample. Remember that overloading can lead to streaky maps. If a preparative map from which a particular peptide will be isolated is being run, it may be best to run the entire sample on two (or more) separate plates. Theoretically, it should be possible to separate 100 µg of material on a single TLC plate; this is often not the case in practice. Check the rate at which the first drop spotted sinks into the cellulose; as more sample is spotted, this rate will decrease. If, while spotting, one gets to a point where the sample drop just sits on the origin and does not spread into the cellulose, stop loading.

Peptide diffusion. Peptides diffuse during the electrophoresis and chromatography, and this leads to a reduction in resolution and sensitivity. To counteract this, the authors try to keep the area on the TLC plate onto which the sample is spotted as small as possible by spotting only a small amount at a time (i.e., less than 1 µl) and drying the sample origin between spottings. In addition, the sample is concentrated by wetting the TLC plates with electrophoresis buffer using a blotter with holes cut out around the origin (Figs. 18.9.1 and 18.9.3). Pressing the edges of the hole onto the plate results in buffer moving from the blotter towards the center of the hole. This concentrates the sample on the origin. For this process to work well, the origin has to be precisely in the center of the hole. In addition, the buffer has to move with similar speed from the entire circumference towards the origin. The sample will inevitably streak if the buffer takes a long time to wet the spot, or moves unevenly through the spot.

Electrophoresis system. In the authors' laboratories the HTLE 7000 electrophoresis system (Fig. 18.9.2) is used. This system features water cooling and an inflatable airbag that presses the TLC plate against the cooling plate. Water cooling prevents overheating during electrophoresis. The inflatable airbag presses excess buffer from the TLC plate; this limits diffusion of the peptides and improves resolution.

Buffers. Three different buffers are typically used for electrophoresis: pH 1.9, pH 4.72, and pH 8.9. Less often used are pH 3.5 and pH 6.5 electrophoresis buffer. To find out which buffer gives the best separation of peptides generated from a particular protein, all three buffers should be tested. If possible, the authors prefer

to work with pH 1.9 buffer. Most peptides dissolve well at this pH. In addition, use of this buffer results less often in streaky maps. The authors usually spot the digest on the origins as marked in Figures 18.9.1 and 18.9.3. For optimal separation of the phosphopeptides generated from a particular protein, the position of the origin and the electrophoresis time may need to be changed. We prefer to change the time of electrophoresis rather than changing the voltage.

Chromatographic process. Chromatography usually takes 12 to 15 hr, but the exact time may vary depending on the age of the chromatography buffer, the batch of plates, the buffer system, the quality of reagents used in the buffer, and the temperature in the room. Three different chromatography buffers are commonly used: isobutyric acid buffer, regular chromatography buffer, or phosphochromatography buffer (see Reagents and Solutions). Pyridine, which is present in all three chromatography buffers, oxidizes and turns yellow over time. Do not use oxidized pyridine to make up chromatography buffers.

To find out which buffer gives the best separation of the peptides generated from a particular protein, all three chromatography buffers should be compared. Most investigators prefer not to use isobutyric acid buffer because it is particularly foul smelling. During the chromatography run, the air space in the tank saturates with buffer and this makes it possible for the volatile chromatography buffer to run all the way to the top of the TLC plate. When the chromatography tank is opened, most of the buffer-saturated air will escape from the tank. This makes it counterproductive to extend the run after the tank has been opened. Therefore do not open the tank when chromatography is in progress, and take all plates out when the chromatography tank is opened.

Separation of the yellow and blue dye functions as a control for successful electrophoresis and allows one to follow the progress during chromatography. The dyes can also be used as standards relative to which the mobility of phosphopeptides of interest can be described, and can be used as markers for the comparison of one plate to another. The yellow compound is neutral at pH 4.72 and pH 8.9 and defines the position to which neutral peptides migrate; at pH 1.9 ϵ -dinitrophenyllysine is positively charged.

Phosphopeptide identification

After running several phosphopeptide maps, it may become apparent that particular phosphopeptides present on the map change in intensity upon treatment of the cells with specific reagents. Such observations often lead to the next question—what is the identity of peptide “A” that becomes phosphorylated following treatment of the cells with factor “B”? If approximately 1 to 10 pmole of phosphorylated peptide can be generated, the peptide is isolated from the TLC plate, purified by HPLC, and identified by mass spectrometry or microsequencing.

In many cases, it is not possible to obtain a phosphorylated peptide in such quantities. The investigator is then forced to learn as much about the phosphopeptide as possible before making an educated guess. We find it useful to make a list of all possible candidate peptides including some of their properties. The next step is to eliminate as many candidates as possible using mobility predictions and the results of relatively simple experiments that can be performed on the minute amounts of labeled peptides isolated from TLC plates.

Making a list of candidate peptides and eliminating the first candidates

Make a list of all possible phosphopeptides that could be generated from the protein of interest given the enzyme used in the primary digest; be sure to include partial cleavage products on this list. This list of peptides should include the nature and position of amino acids that can be phosphorylated and the peptides' susceptibility to further cleavage by proteases or chemicals (for an example, see van der Geer and Hunter, 1990)

After making such a list, first calculate and then plot the predicted mobilities of all candidate phosphopeptides. See Table 18.9.2 for values that can be used to do this.

To calculate electrophoretic mobility. The mobility of a peptide in the electrophoresis dimension is dependent on its charge (e) to mass (M) ratio, as $M_r = keM^{-2/3}$. When calculating relative mobilities (M_r) the simplified equation $M_r = eM^{-1}$ can be used with good success. The net charge on a peptide is calculated by summing the charges of the N and C termini and those of the side chains of its amino acids at a given pH, and dividing by either the actual mass of the peptide or simply by the number of amino acids in it. Approximate charge values at the specific pHs commonly

Table 18.9.2 Approximate Charge Values at Specific pHs Commonly Used for Electrophoresis^a

	pH 1.9	pH 3.5	pH 4.7	pH 6.5	pH 8.9
Amino-terminal NH ₂	+1	+1	+1	+1	+0.5
Carboxy-terminal COOH	N	-0.5	-1	-1	-1
Arginine	+1	+1	+1	+1	+1
Aspartate	N	N	-0.7	-1	-1
Cysteine (oxidized)	-1	-1	-1	-1	-1
Histidine	+1	+1	+1	+0.5	N
Glutamate	N	N	-0.5	-1	-1
Lysine	+1	+1	+1	+1	+1
Phosphoserine	-1	-1	-1	-1.3	-2
Phosphothreonine	-1	-1	-1	-1.3	-2
Phosphotyrosine	-1	-1	-1	-1.3	-2

^aN indicates neutral.

used for electrophoresis are given in Table 18.9.2.

To calculate chromatographic mobility. A peptide's mobility in the chromatographic dimension is dependent on its hydrophobicity, and thus on its amino acid sequence. The order of the amino acids will also change the peptide's mobility; thus, two peptides of identical sequence which are phosphorylated at one or the other of two possible sites may migrate different distances in the second map dimension, although they migrate identically in the first dimension since their charge:mass ratio is the same. While it is not possible to exactly predict chromatographic mobilities, relative mobilities can be plotted with some success by calculating an average mobility for the peptide based on migratory values of its constituent amino acids. This is not ideal, since the calculated R_f values of individual amino acids are significantly influenced by the presence of their charged amino and carboxy termini, which of course are noncontributory in the context of a peptide. This accounts in part for the compression of the calculated maps compared with the observed peptide migrations. Values for chromatographic mobilities of amino acids have been published in (Boyle et al., 1991); these were determined for each individual amino acid relative to the ϵ -DNP-lysine (yellow) marker using cellulose plates available twenty years ago. The quality of the cellulose used in such plates has changed markedly over the years; contact Ned Lamb for values that have been empirically determined more recently (<http://www.genestream.org>).

Bear in mind that these calculations can also be accomplished using a computer program. Ned Lamb (CNRS, Paris) has constructed a Web site for analysis of phosphopeptide maps, which may be found at <http://www.genestream.org>. Phospepsort 4, the program that he has developed based on an earlier version which originated at the Salk Institute, gives the biophysical characteristics as well as the electrophoretic and chromatographic mobilities of each proteolytic fragment. Alternatively, predicted peptide mobilities can be visualized using the graphical interface to Phospepsort 4: Mobility. In addition, Ned Lamb is working on a program that fits the calculated mobility values to the actual values for peptides of known composition. The Resolve program then reads the position of a spot on the actual map and calculates which peptide(s) derived from the protein being mapped could have the mobility of that spot.

It is imperative to note that to date there is no program that accurately predicts the mobility of all peptides of a protein. This may be explained by the fact that mobilities are calculated using values established for single amino acids rather than for peptides. Plotting predicted phosphopeptide mobilities on a graph using linear axes results in a greatly compressed map as compared to that observed in one's autoradiograms, especially in the chromatographic dimension. Therefore, do not despair if the predicted map of all phosphopeptides in the protein of interest looks nothing like the actual map that was generated experimentally. The value of such predictions comes from the fact that the *relative* positions of phosphopep-

tides are predicted with great accuracy by such programs/calculations. Thus, if the predicted mobility of a hypothetical phosphopeptide places it on the anode side of a cluster of phosphopeptide candidates, while the phosphopeptide of interest on the actual map is present on the cathode side of the cluster, the hypothetical phosphopeptide may be eliminated from further consideration. This example illustrates how careful use of predicted peptide mobility maps may lead to elimination of candidate peptides.

Isolating peptides from TLC plates

Phosphopeptides isolated by elution from cellulose, as described above, can be used without further purification for certain types of analysis. What many people overlook, however, is that this sample is by no means necessarily pure. It includes, in addition to the radioactive phosphopeptide in question, any unlabeled tryptic fragments that may have comigrated with it on the cellulose plate—these peptides are generated from trypsin itself and from the carrier protein used in the TCA precipitation during sample preparation. For this reason, the sample is usually further purified by HPLC to clean it up before analysis by mass spectrometry or microsequencing; the column fractions are counted in a scintillation counter to determine which ones to use for further analysis. However, for manual Edman degradation (see Basic Protocol 2), secondary cleavage (see Basic Protocol 3), and phosphoamino acid analysis (*UNIT 18.3*) no further purification is necessary, as the interpretation of the results relies solely on the visualization of the resultant ³²P-containing reaction products. The presence of unlabeled contaminants does not interfere with the interpretation of the results.

Phosphoamino acid analysis

Perhaps the most obvious step to take with an unidentified phosphopeptide in hand is to determine the phosphoamino acid content of the peptide of interest. This will eliminate many candidate peptides from consideration; this step is obviously not necessary if a long exposure of the phosphoamino acid analysis of the labeled protein in question indicates that only one species of phosphorylated amino acid is present. For phosphoamino acid analysis only 50 cpm of purified phosphopeptide is needed (though in peptide mapping and related protocols one can never have enough cpm). The details of such analysis have been discussed in detail in *UNIT 18.3*. Briefly, the phosphopeptide

eluted from the TLC plate is hydrolyzed by incubation for 60 min at 110°C in 30 μ l of 6 N HCl. The appearance of yellow to brown color in the sample during hydrolysis indicates that some cellulose remained despite efforts to clarify the phosphopeptide eluate. This sometimes causes the sample to streak. After the sample is lyophilized, it is resolved with stainable standards in two dimensions by electrophoresis as described (*UNIT 18.3*). The phosphoamino acid composition is determined by matching the resultant spot(s) on the autoradiogram with the ninhydrin-stained standards on the cellulose plate.

Manual Edman degradation

At pH 8 to 9, phenylisothiocyanate reacts with the free amino group(s) of a peptide to form a corresponding phenylthiocarbonyl peptide. Treatment of these PTC-peptides with acid (TFA) results in the cleavage of the derivatized amino-terminal amino acid and its release as an anilinothiazolinone molecule. This latter species is not stable, and will cyclize to yield the phenylthiohydantoin (PTH) derivative of the amino acid in aqueous acid. If a phosphoserine or phosphothreonine residue is present, a β -elimination during the cyclization releases free phosphate. Phosphotyrosine, however, is simply released as the anilinothiazolinone derivative. This may be converted to the phenylhydantoin form for analysis by incubating it in 0.1 N HCl at 80°C for 20 min. PTH-phosphotyrosine to use as a marker is easily synthesized by reacting phosphotyrosine with phenylisothiocyanate and then heating it in acid (see Basic Protocol 2, step 5); it can be visualized as a dark spot using a hand-held UV light.

While the protocol is relatively simple, each cycle takes ~2 hr to complete and requires at least 100 cpm for an unambiguous result. It is important to run a portion of the starting material out on the TLC plate to show how much, if any, free phosphate is there, since some hydrolysis of the peptide may have occurred during its isolation. Because at each cycle the reaction may not go to completion, one should plan to do at least one more cycle than is predicted to be necessary to release the phosphate (i.e., if all the candidate peptides are phosphorylated at or before the third residue from the N terminus, at least 4 cycles should be run). Thus, how many cpm are in the eluted map spot may determine how many cycles are run. In any case, we generally do not attempt to perform more than 5 or 6 cycles. If no phosphate is released during the course of these

cycles, this conclusion in itself will usually greatly reduce the number of candidate sites.

In addition to the position of the phosphorylated residue, clues to a peptide's sequence may also be gleaned from Edman degradation. The way in which the residual peptide shifts its electrophoretic position on the plate after each cycle will indicate whether an acidic or basic amino acid has just been removed. If the tryptic peptide's carboxy-terminal residue is a lysine, the lysine's ϵ -amino group will be derivatized in the first cycle, and so a positive charge will be lost in that instance as well.

If an automated peptide sequencer is available, 20 or more Edman cycles may be analyzed by coupling the phosphopeptide via carboxyl groups to a Sequelon membrane (Millipore) and letting the machine do the work. At the end, the fractions are counted to see where the radioactivity is released. This method obviously requires far fewer cpm in a phosphopeptide sample than does the manual Edman protocol, since one only analyzes the released material for radioactivity rather than a portion of the whole sample at the end of each cycle. A method for adapting an automated sequencer for such a purpose is discussed in Mitchelhill et al. (1997).

An adaptation of the protocol presented above is found in Fischer et al. (1997). This protocol uses a volatile isothiocyanate (trifluoroethyl isothiocyanate) and volatile buffers; as a consequence the extraction steps can be eliminated and this results in shorter cycle times (~45 min).

Secondary digests

In the past, secondary digests of tryptic phosphopeptides represented a large part of the further analysis of these peptides. The utility of these digests, however, is totally dependent on the sequence of the protein in question. A list of enzymes commonly used in such digests is found in Table 18.9.1, along with their cleavage site(s) and optimal pH and temperature. As is true for trypsin, these enzymes sometimes cleave inefficiently when encountering a cleavage site in a certain sequence context—some of these problematic sites are listed as well. In our experience, the enzymes/reagents that cleave only one amino acid (i.e., proline-specific endopeptidase, V8, or cyanogen bromide) tend to be more useful and give less ambiguous results.

Do not expect that the results of one enzyme digest will eliminate more than a few candidate sites. As it is most likely that the enzyme in question will not cleave the phosphopeptide, it

is imperative that a positive control be included in the digest—i.e., a peptide whose sequence is known and which the enzyme will cleave. If such a peptide is not available, a portion of the primary digest might be used as a positive control—with luck this will contain at least one peptide whose migration will be changed by this secondary digestion.

Running out the mix of undigested and digested peptides is very important, since failure to comigrate with the original peptide will unequivocally demonstrate a change in the phosphopeptide's mobility. A sample treated exactly as the digested sample, but without the cleavage reagent, should be analyzed in parallel. This is to ensure that changes in mobility seen in the digested peptide are truly due to the presence of the cleavage reagent.

When deciding where to spot the samples and what conditions to use to run the plates, keep in mind that if the peptide did not migrate very far in the original map, it may be possible to load two sample on a single plate. Also keep in mind, however, that free phosphate may be released during the enzymatic digestion (due to elevated temperatures and resulting hydrolysis). Free ^{32}P -phosphate originating from the sample loaded to the right may complicate the interpretation of the sample loaded on the left (anode) side of the plate.

It should be noted that additional information about the other amino acids in a phosphopeptide may be gleaned just by running a tryptic digest in the electrophoretic dimension at another pH—identification of the different peptides by their mobility in the chromatography dimension will allow one to compare the mobility of peptides when run at different pHs and decide if a particular spot has changed its migration in the first dimension. Such a change in migration would be affected only if the phosphopeptide contained an amino acid whose charge was changed at the second pH used.

Comigration of a synthesized phosphopeptide with a phosphopeptide isolated from a peptide map

After eliminating all but a few of the candidate peptides, comigration of a synthetic phosphopeptide with the ^{32}P -labeled peptide generated by digesting a protein from labeled cells may provide convincing evidence as to the latter's identity. Obviously, care must be taken that the synthetic peptide be quite pure (generating only one spot on the cellulose plate) and that, while enough is loaded to be easily visualized, the plate not be overloaded (which

would cause streaking and therefore render the comigration ambiguous). If the peptide is synthesized as a phosphopeptide, ~5 to 25 μg of pure peptide is needed for each comigration, as it will need to be ninhydrin stained for visualization. This is by far the easier approach—starting with an unphosphorylated peptide entails not only finding a kinase that will phosphorylate it, but also purifying the phosphopeptide first before running the comigration. In either case, synthesis of peptides, and in particular those with special residues such as phosphorylated amino acids, is not inexpensive, and so this type of experiment is usually attempted after one has accumulated several other clues regarding a site's identity.

The alternative is to mutate the remaining candidate phosphorylation sites. An epitope-tagged version of the mutant protein can then be expressed in cells and mapped to check whether the spot in question disappears from the map. This approach of course assumes that the sequence of the protein in question is known and that one has a clone in hand for mutagenesis.

While satisfying to many, this mutagenesis approach is not definitive, as it can be argued that the mutated form of the protein may not fold correctly and therefore may not be phosphorylated correctly, resulting in a misleading loss of map spots. Another potential problem is that, if the phosphorylation of the site in question is an ordered event, dependent on another site's state of phosphorylation, mutation of this other site could lead to the erroneous conclusion that the disappearance of a particular spot is the direct result of the mutation introduced, rather than an event of secondary consequence. Mutagenesis then should be taken as a supporting argument for the presumed identity of a phosphorylation site, rather than as definitive proof. Taken in conjunction with other evidence, it can nevertheless be quite convincing.

Scaling up phosphopeptide mapping to isolate peptides for mass spectrometry and microsequencing

The most definitive methods for determining the site(s) of phosphorylation in a protein are the determination of the amino acid sequence or mass of the phosphopeptides. While the sensitivity of the instruments used in both these techniques has improved over the past 5 years, both methods require on the order of 1 to 10 pmole of material for analysis. When analyzing the phosphorylation of a receptor protein-tyrosine kinase of which ~10,000

molecules are present on the cell surface, one would need to grow 60 10-cm dishes of cells to isolate the 6×10^{11} molecules (1 pmol) required for a successful analysis. In this calculation, we have assumed 100% recovery and 100% stoichiometry of phosphorylation at the sites of interest. However, since it is very possible that not every molecule is phosphorylated at the sites in question, and that one is likely to take at least a 50% loss of material over the entire protocol, it would be best to start then with at least 240 dishes of cells. The use of overexpressed protein in cells would obviously facilitate the accumulation of sufficient amounts of protein for analysis.

Isolation of enough material for analysis becomes less difficult if an *in vitro* system can be used to generate the phosphopeptides in question. Protein kinases are notoriously promiscuous *in vitro*, so it may be possible to generate the sites of interest using a kinase that is actually not the one responsible for the phosphorylation *in vivo*. The protein of interest produced in bacteria makes a good substrate, as it is most likely not phosphorylated to start with. It is first necessary to show that the phosphopeptides generated by incubation of recombinant protein with a purified kinase *in vitro* comigrate exactly with those purified from labeled cells. If there is any doubt about this, it is reassuring to run the maps for a longer time than usual in the electrophoresis dimension, or at a different pH to further demonstrate comigration. Given the specificity of the two-dimensional separation on these maps, it is unlikely that two phosphopeptides that truly comigrate will not be identical.

Anticipated Results

Upon developing the first autoradiogram of an initial peptide map, depending on the number of spots seen, the researcher will most likely be left wondering what it all means. There are several things to keep in mind when studying the pattern of spots on a peptide map. First, migration in the electrophoresis dimension is a function of the charge to mass ratio of a peptide. Migration in the chromatography dimension is related to the hydrophobicity of a peptide. The more hydrophobic a peptide, the further it migrates in the chromatographic dimension. In general, most phosphopeptides of the protein studied will be represented by one spot in the map. It is possible then to determine the relative stoichiometry of phosphorylation at different sites by comparing the intensity of each spot on the autoradiogram. This, however, relies on the

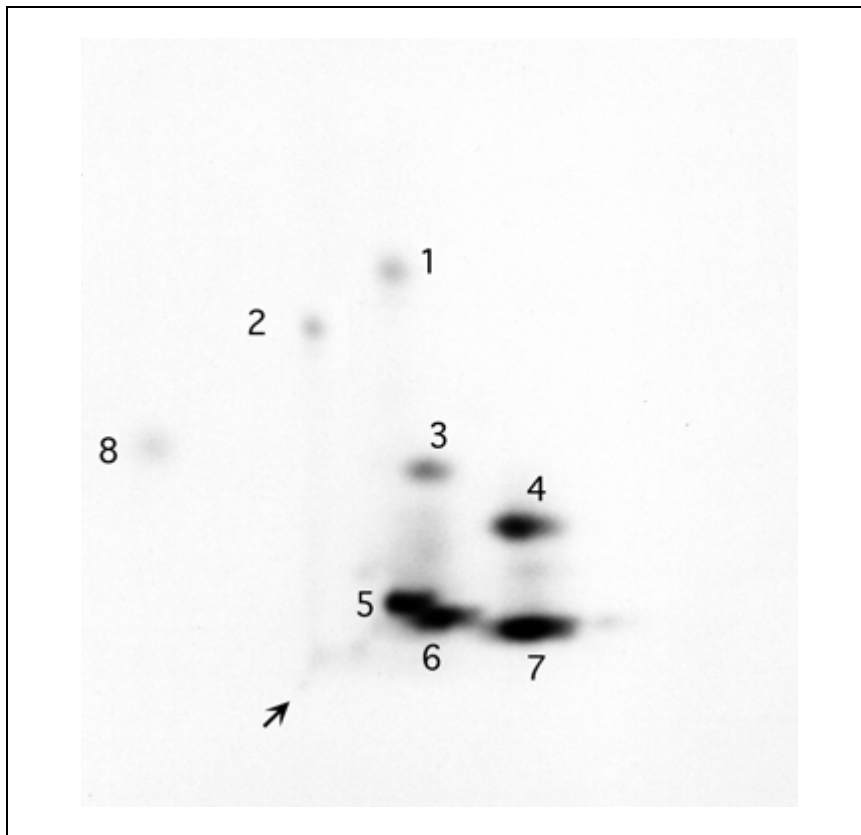


Figure 18.9.6 An example of a tryptic phosphopeptide map based on that of human Nck-alpha. For the first (horizontal) dimension, electrophoresis was run at pH 1.9 for 25 min at 1.0 kV; the anode is at the left. Ascending chromatography was run for 15 hr in phosphochromo buffer. The sites represented by spots 1 to 7, with the exception of spot 2, have been identified. Spot 1 is an 11-amino-acid, phosphotyrosine-containing peptide. While spot 2 also contains phosphotyrosine, and runs in a position likely to be the doubly phosphorylated version of this peptide, it turns out to be unrelated to spot 1. Spot 3 represents a 5-amino-acid, phosphoserine-containing peptide; this same peptide with an amino-terminal arginine runs as spot 4; thus it can be seen that in this case the tryptic cleavage is largely incomplete. Spot 5 represents a 20-amino-acid phosphoserine-containing peptide which, with an amino-terminal lysine, runs as spot 6. Spot 7 represents a peptide that is unrelated to 5 and 6. Spot 8 represents free phosphate, released during sample preparation.

assumption that all tryptic peptides are recovered during the the entire protocol with similar efficiencies.

There are several cases, however, in which this one phosphopeptide:one spot rule does not hold. One such case is when the enzyme used to digest the protein of interest has not worked to completion, yielding both partial and complete digestion products seen in the map.

Multiple digestion products are also generated at sites where a run of basic residues is present. Trypsin works very efficiently as an endopeptidase, hydrolysing peptide bonds following basic residues. In contrast, trypsin works poorly as an exopeptidase. Trypsin will

cleave randomly within the run of basic residues and is unable to take off additional basic residues that may have been left. This results in a series of digestion products differing by the number of basic residues at their amino- or carboxy-terminus. Addition of such a basic residue to a phosphopeptide changes its migration in both the electrophoresis and the chromatography dimension (it is more positively charged and runs further towards the cathode, it is also more hydrophilic and as a consequence runs less far relative to the buffer front in phosphochromo buffer).

Similarly, a given peptide will migrate differently if phosphorylated at two positions

rather than one. Addition of another negative charge will again make the peptide more hydrophilic and thus it will migrate less far in the chromatography dimension. This time, however, it will migrate further towards the anode in the electrophoresis dimension, giving a diagonal pattern descending in the opposite direction to that seen with the addition of a lysine or an arginine.

It is important to remember that the electrophoretic mobility of a peptide is dependent on its mass. For larger peptides, the slope of the diagonal seen with the addition of either positive or negative charges will be steeper, since the addition of another charge when divided by the mass will make less of a difference to the distance the peptide travels.

While trypsin has been traditionally used in peptide mapping, it may not be the enzyme of choice for proteins phosphorylated by PKA, PKC, or other protein kinases whose recognition sequence involves multiple arginines or lysines, as trypsin often fails to cleave after all such residues when they are present in runs. In addition, trypsin cleaves inefficiently at arginines or lysines two residues amino-terminal of a phosphoserine or phosphothreonine (i.e., R/K-X-P.Ser).

Sometimes an individual peptide may appear to have an electrophoretic partner that migrates directly above or below it in the chromatographic dimension. This sort of pattern may be observed as the result of two different scenarios: (1) it may be the result of incomplete oxidation of the peptide if it contains a methionine residue (in which case the lower spot is the oxidized form), or (2) it may be the result of methylation of the peptide running in the lower position. Such a methylation may occur during the performic acid oxidation and is dependent on the 1.5-ml microcentrifuge tubes being used. Historically we have found that certain tubes are more apt than others to produce such unwanted side reactions; for this reason it is advisable to stock certain lots of tubes that do not produce such artifacts in the final maps. Similarly, brands and batches of tubes appear to differ in the extent to which peptides “stick” to them during the final steps of the protocol.

An exemplary tryptic phosphopeptide map based on that of a real protein (Nck) is shown in Figure 18.9.6. This map illustrates the points mentioned above. Perhaps most importantly, it also illustrates the fact that just because two spots appear to be on a diagonal it is not a foregone conclusion that they are related. Although peptides 1 and 2 appear to represent the

singly and doubly phosphorylated forms of a single tryptic peptide, in this case it turned out that they represent two completely different peptides. Peptides 3 and 4 and 5 and 6, respectively, represent two sets of peptides that are related and differ only by the addition of a basic residue. Spot 8 represents free phosphate, liberated by hydrolysis of phosphoester bonds that has occurred during sample preparation. It is useful to both compare the amount of free phosphate generated in different samples and to use the phosphate spot as another standard marker when comparing peptide mobilities on different plates.

Time Considerations

To generate a two-dimensional phosphopeptide map, at least 9 days will elapse from the time the ^{32}P label is added to the cells until the autoradiogram of the map is in hand. The typical researcher, intrigued by one or more particular spots that appear or disappear from such maps depending on how the cells or samples were treated, may rush to attempt to identify the phosphorylation site represented by such spot(s). Please be advised that this will take at least 4 months of hard work and effort, assuming that everything goes well. There are several different strategies to follow, which are outlined throughout the unit (especially see Background Information). Choice of a particular course will depend on the reagents and the equipment available for analysis.

Literature Cited

- Boyle, W.J., van der Geer P., and Hunter, T. 1991. Phosphopeptide mapping and phosphoamino acid analysis by two-dimensional separation on thin-layer cellulose plates. *Methods Enzymol.* 201:110-148.
- Fischer, W.H., Karr, D., Jackson, B., Park, M., and Vale, W. 1991. Microsequence analysis of proteins purified by gel electrophoresis. *Methods Neurosci.* 6:69-84.
- Fischer, W.H., Hoeger, C.A., Meisenhelder, J., Hunter, T., and Craig, A.G. 1997. Determination of phosphorylation sites in peptides and proteins employing a volatile Edman reagent. *J. Protein Chem.* 16:329-333.
- Mitchellhill, K.I., Michell, B.J., House, C.M., Stapleton, D., Dyck, J., Gamble, J., Ullrich, C., Witters, L.A., and Kemp, B.E. 1997. Posttranslational modifications of the 5'-AMP-activated protein kinase β 1 subunit. *J. Biol. Chem.* 272:24475-24479.
- van der Geer, P. and Hunter, T. 1990. Identification of tyrosine 706 in the kinase insert as the major colony-stimulated factor 1 (CSF-1)-stimulated autophosphorylation site in the CSF-1 receptor

in a murine macrophage cell line. *Mol. Cell. Biol.* 10:2991-3002.

Wang, Y.K., Liao, P.-C., Allison, J., Gage, D.A., Andrews, P.C., Lubman, D.M., Hanash, S.M., and Strahler, J.R. 1993. Phorbol 12-myristate 13-acetate-induced phosphorylation of op18 in Jurkat T cells. *J. Biol. Chem.* 268:14269-14277.

Key References

Boyle et al., 1991. See above.

van der Geer, P., Luo, K., Sefton, B.M., and Hunter, T. 1993. Phosphopeptide mapping and phosphoamino acid analysis on cellulose thin-layer plates. *In* Protein Phosphorylation; a Practical Approach (D.G. Hardie, ed.) pp. 31-59, IRL Press, Oxford.

Both of these papers discuss many of the protocols described in this unit.

Internet Resources

<http://www.genestream.org>

This Web site contains a program for calculating the mobility of a peptides of known composition and a program that reads the position of a spot on the actual map and calculates which peptide(s) derived from the protein being mapped could have the mobility of that spot.

Contributed by Jill Meisenhelder
and Tony Hunter
The Salk Institute for Biological Studies
La Jolla, California

Peter van der Geer
University of California, San Diego
La Jolla, California

Use of Protein Phosphatase Inhibitors

UNIT 18.10

Reversible protein phosphorylation is recognized as a major mechanism regulating the physiology of plant and animal cells. Virtually every biochemical process within eukaryotic cells is controlled by the covalent modification of key regulatory proteins. This in turn dictates the cellular response to a variety of physiological and environmental stimuli; errors in signals transduced by phosphoproteins contribute to many human diseases. Thus, defining protein phosphorylation events, and specifically, the phosphoproteins involved, is crucial for obtaining a better understanding of the physiological events that distinguish normal and diseased states.

In studying protein phosphorylation, two common experimental problems arise that mandate the use of protein phosphatase inhibitors. The first problem arises when the goal of the experiment is to decipher physiological events regulated by reversible protein phosphorylation but the hormonal stimuli or signaling pathways involved are not known. This situation is further complicated by the fact that many protein phosphorylations are rapid and transient, thereby eluding investigation. One solution is to employ cell-permeable compounds that inhibit cellular protein phosphatases in order to amplify the basal activity of protein kinases. Inhibition of protein phosphatases also prevents the turnover of protein-bound phosphate, thereby enhancing the detection of phosphoproteins by standard techniques such as SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A) or phosphorimaging.

In other situations, the aim may be to analyze the impact of hormones and other physiological stimuli on the function of a specific phosphoprotein. Such detailed biochemical analysis is only made possible after the protein of interest has been separated from other cellular components that potentially interfere with these studies. In this circumstance, the cells or tissues exposed to hormones must be homogenized in the presence of protein phosphatase inhibitors to preserve the cellular phosphorylation events promoted by the hormones. The goal is to utilize the most effective inhibitors to abolish the activity of protein phosphatases that may dephosphorylate the proteins of interest during the lengthy protein isolation procedures. This unit describes the general or basic protocols for inhibiting the cellular PP1/PP2A activity with okadaic acid (see Basic Protocol 1) as well as in vitro with microcystin-LR (see Basic Protocol 2). Basic Protocols 3 and 4 describe approaches recommended for inhibiting cellular and in vitro activity, respectively, of PP2B/calcineurin. Finally, Basic Protocol 5 outlines a widely utilized strategy for inhibiting protein tyrosine phosphatases. The reader should be aware that these Basic Protocols are deliberately designed to cover the widest range of experimental situations and therefore represent prototypic procedures that can be readily used with many phosphatase inhibitors that are not specifically mentioned in this unit.

A variety of very effective inhibitors, particularly those targeting protein serine/threonine phosphatases, are currently commercially available. Some of these compounds are cell-permeable and yet others are better utilized in vitro to suppress phosphatase activity in lysates and/or cell fractions. This unit describes methods for defining physiologically relevant protein phosphorylation events. Thus, in some cases, cells in culture are treated with cell-permeable phosphatase inhibitors to detect and establish the in vivo phosphorylation events. In other cases, the cells are stimulated with known physiological stimuli that modify cellular protein phosphorylation. In either situation, subsequent analyses require that the cells be homogenized in buffers containing phosphatase inhibitors that preserve the protein modifications through the steps of protein purification and analysis. This in turn facilitates the identification of the particular amino acids modified and sets

Analysis of
Protein
Phosphorylation

18.10.1

the stage for more defined molecular approaches, such as site-directed mutagenesis, that will allow the investigator to establish the functional role of individual protein phosphorylation events.

PROTEIN SERINE/THREONINE PHOSPHATASES

Serines and threonines constitute the major sites of protein phosphorylation in all eukaryotic cells. Varying estimates suggest that even in cells transformed using oncogenic viral protein tyrosine kinases, >97% of the protein-bound phosphate appears on serines and threonines. Thus, it is not surprising that numerous microorganisms produce toxins and natural products that elicit their deleterious effects by inhibiting cellular protein serine/threonine phosphatases (McCluskey et al., 2002). Many of these compounds show some degree of selectivity for subgroups of protein serine/threonine phosphatases and may provide insight into the cellular phosphatases that catalyze protein dephosphorylation events. The specificity of some of these compounds has been determined in vitro as differences in their concentration-response curves for the inhibition of purified protein phosphatases. When used in vivo, the chemical properties of the compounds, many of which are hydrophobic, do not allow accurate estimates of cellular concentration or predict their subcellular distribution. Thus, the in vitro specificity of the phosphatase inhibitors may be lost and other more direct molecular approaches may be required to define the protein phosphatases that regulate specific cellular events. The principal use of phosphatase inhibitors, as described in this unit, is to complement the in vitro use of purified phosphatases (UNIT 18.5) to define the regulatory importance of reversible protein phosphorylation, rather than to identify the physiologically relevant protein phosphatases.

Early biochemical studies classified eukaryotic protein serine/threonine phosphatases into two major groups, termed type-1 (PP1) and type-2 protein phosphatases (Shenolikar and Nairn, 1991). The type-2 phosphatases are further separated into three distinct groups, PP2A, PP2B, and PP2C, showing very different profiles of inhibition by the available compounds (Table 18.10.1). Molecular cloning of mammalian phosphatase catalytic subunits further expanded this family of enzymes, which to date includes up to five additional phosphatase catalytic subunits numbered PP3 to PP7. Of these five, all but PP7 share a highly conserved catalytic site, and therefore are similarly inhibited by various compounds. Yet others, such as PP2C or PP7, show complete insensitivity to the known phosphatase inhibitors and there are no currently available tools to inhibit these phosphatases.

Table 18.10.1 Commercially Available PP1/PP2A Inhibitors^a

Compound	Phosphatases	Recommended use	Supplier(s)
Okadaic acid	PP2A > PP1	In vivo	Alexis Biochemicals, Sigma
Calyculin A	PP1 ≥ PP2A	In vivo	Alexis Biochemicals, Sigma
Tautomycin	PP1 > PP2A	In vivo	Biomol Research Laboratories
Cantharidin A	PP2A > PP1	In vivo	—
Microcystin-LR	PP1 = PP2A	In vitro	Alexis Biochemicals, Sigma
Nodularin	PP1 = PP2A	In vitro	—
Fostreicin	PP2A>>PP1	In vitro	A.B. Scientific, Sigma, Alexis Biochemicals

^aAll compounds listed in this table also inhibit PP3, PP4, PP5, and PP6. Their shared sensitivity to these compounds results in their inclusion in the category of PP2A-like enzymes. In contrast, these compounds show no significant inhibitory activity towards PP2B, PP2C, and PP7. As the compounds possess little structural homology, several different inhibitors may be used to validate the physiological effects of inhibiting their shared targets, serine/threonine phosphatases.

Table 18.10.2 Commercially Available PP2B (Calcineurin)^a Inhibitors

Compound	Phosphatases	Recommended use	Suppliers
Cyclosporin A ^b	PP2B	In vivo	Alexis Biochemicals, Sigma
FK506	PP2B	In vivo	Alexis Biochemicals
Cypermethrin ^c	PP2B	In vivo	Alexis Biochemicals
Deltamethrin	PP2B	In vitro	Alexis Biochemicals

^aPP2B (calcineurin) is a Ca²⁺/calmodulin-activated protein serine/threonine phosphatase and, like many Ca²⁺/calmodulin-activated enzymes, is inhibited by the calcium chelator, EGTA. Inhibition of PP2B activity using EGTA is not a very viable approach in cells and the compounds listed in this table more effectively suppress cellular PP2B activity.

^bA number of inactive and partially active compounds related to cyclosporin are available. Rapamycin, which shares common target proteins with FK506, namely FK506-binding proteins or FKBP, does not inhibit PP2B activity and is used as control for FK506.

^cCompounds related to cypermethrin, such as resmethrin, show weak or no activity as PP2B inhibitors and can function controls for the pyrethroid inhibitors.

The compounds listed in Table 18.10.1 inhibit several classes of protein serine/threonine phosphatases. Enzymes such as PP3, PP4, PP5, and PP6 share structural homology to PP2A and therefore show similar sensitivity to these compounds. In contrast, PP2B is inhibited by a distinct group of compounds (Table 18.10.2) that show no activity against either PP1 or the PP2A-like phosphatases.

PP1/PP2A Classes of Protein Serine/Threonine Phosphatases

PP1 and PP2A together account for >90% of protein serine/threonine phosphatase activity in most eukaryotic cells. With the exception of mammalian skeletal muscle, where PP1 predominates, most cells and tissues possess equivalent amounts of PP1 and PP2A. This has made okadaic acid and calyculin A, both potent cell-permeable PP1/PP2A inhibitors, the most widely utilized compounds for studying cellular protein phosphorylation.

INHIBITION OF CELLULAR PP1/PP2A ACTIVITY WITH OKADAIC ACID

Okadaic acid (OA), a polyether produced by marine dinoflagellates, is concentrated by shellfish and is a common cause of diarrhetic shellfish poisoning in humans. OA is commercially available as a free acid or as a potassium, sodium, or ammonium salt. Prevailing evidence favors the sodium salt of OA as the most readily permeable across cell membranes.

Biochemical studies demonstrate that OA inhibits the purified PP2A catalytic subunit with an IC₅₀ value of 0.2 to 1.0 nM (Cohen et al., 1989). This concentration of OA, however, has little effect on PP2A activity in cell lysates, which typically exhibit 30- to 150-fold higher IC₅₀ values than the purified enzyme. Extensive dilution of cell extracts enhances the efficacy of OA as a PP2A inhibitor. This may suggest that OA binds, albeit with low affinity, to other cellular components in cell extracts. This may be more of an issue in intact cells where significantly higher OA concentrations (10 to 100 nM) are required for effective PP2A inhibition.

PP1, unlike PP2A, has a higher IC₅₀ value for OA (10 to 100 nM; Cohen et al., 1989), whether assayed with purified enzyme or cell lysates, and the sensitivity of PP1 activity to OA is not modified by dilution of cell lysates. In some cellular studies, OA concentrations between 10 and 50 nM resulted in significant enhancement of protein phosphorylation, presumably reflecting the inhibition of PP2A and other PP2A-like enzymes. In contrast, OA concentrations >10 μM are needed to effectively inhibit cellular PP1 activity.

**BASIC
PROTOCOL 1**

**Analysis of
Protein
Phosphorylation**

18.10.3

This difference in affinity can provide an opportunity to discriminate between the physiological functions of the two major classes of cellular protein serine/threonine phosphatases, namely PP1 and PP2A.

Numerous studies suggest that the predominant site of OA accumulation is the lipid bilayer of the plasma membrane (Suganuma et al., 1989; Nam et al., 1990; Namboodiripad and Jennings, 1996). OA transfers slowly from this site to other compartments and this may account for the slow onset of phosphatase inhibition by OA in most cells (up to 90 min) as well as the requirement of 10- to 100-fold higher concentrations of OA to inhibit PP2A activity in intact cells compared to cell lysates. Thus, it is surprising to find that OA can also be washed out of cells, albeit thorough washing can take hours. This reversible behavior may allow investigators to analyze the reversal of phosphorylation events triggered by OA.

Duration of OA treatment to elicit maximal changes in cellular protein phosphorylation is largely empirically determined and must be defined for each individual cell type. In general, cells display a maximal response to micromolar concentrations of OA within 10 to 30 min. Lower OA concentrations require significantly longer exposure times (up to 2 hr). Due to their remarkable ability to activate multiple cellular signaling pathways, OA and other cell-permeable phosphatase inhibitors display some cytotoxicity. This is partially alleviated by lowering OA concentrations and reducing exposure times. Thus, OA treatments for several hours or days are not recommended. If longer times are necessary, calyculin A (CA), which inhibits PP1 and PP2A with near equal potency, may be a better substitute. While the literature on calyculin A is less extensive, this compound can be used at much lower concentrations (0.1 to 1.0 μM) than OA to yield similar physiological effects.

OA induces errors in DNA repair that are not mediated by phosphatase inhibition, (Nakagama et al., 1997) which suggests the presence of additional targets for OA. This behavior mandates that appropriate controls be used in cellular studies. In this regard, the treatment of cells with vehicle (DMSO) is commonly used as a control for OA, CA, and other phosphatase inhibitors that are dissolved in this solvent. There are, however, a number of commercially available inactive OA analogs that may serve as better controls. The best of these is 1-norokadaone, whose structure closely matches that of OA and maintains some of the non-phosphatase-mediated effects elicited by OA. Structure-function studies that established a critical role for the acidic side-chain of OA in phosphatase inhibition also generated a number of inactive OA analogs in which the critical carboxylate group is esterified (Nashiwaki et al., 1991). Such compounds, like methyl okadaate, show low or no activity in vitro as PP1/PP2A inhibitors. However, these compounds may not be effective negative controls, as they are slowly de-esterified in cells, yielding the active OA.

Finally, OA, CA, and several other phosphatase inhibitors show significant activity as tumor promoters in mouse skin and gut (Fujuki et al., 1988; Suganuma et al., 1990). Thus, it is important that the investigator takes extreme care in handling and disposing of solutions containing these compounds.

Materials

Cultured mammalian cells grown to <50% confluence (*APPENDIX 3F*)

0.1 or 1.0 mM okadaic acid sodium salt (Alexis Biochemicals or Sigma), in DMSO (store at -20°C in a light-proof container under N_2)

0.1 or 1.0 mM 1-norokadaone (Calbiochem or Sigma) in DMSO (store at -20°C in a light-proof container under N_2)

NOTE: Okadaic acid (OA), either as sodium, potassium, or ammonium salt, shows increased solubility in aqueous solutions compared to the free acid. The solutions of the OA salts are also more stable during storage. However, for reasons that are not entirely clear, aqueous solutions of OA lose their activity as phosphatase inhibitors over time. By comparison, the stock solutions of OA in DMSO (or ethanol or methanol) can be stored in a dark glass container under nitrogen at -20°C for many months without significant loss of activity.

1. Prepare cultured mammalian cells to $\sim 50\%$ confluence (*APPENDIX 3F*).

Okadaic acid activates a variety of cell signaling pathways and physiological events in mammalian cells (Schonthal, 1998). Thus, responses of different mammalian cells to OA vary widely. In general, changes in protein phosphorylation in quiescent or confluent cells induced by OA are less profound. Moreover, quiescent cells often initiate apoptosis following OA exposure (Kiguchi et al., 1994; Morimoto et al., 1997). In contrast, programmed cell death appears less of a problem when analyzing actively growing cultured cells.

2. Add okadaic acid (final concentration $10\ \mu\text{M}$) to the culture medium or exchange with fresh medium containing $10\ \mu\text{M}$ okadaic acid. Expose control cells to medium containing vehicle (DMSO) or the inactive analog, 1-norokadaone. Incubate cells for 30 min prior to harvesting (*APPENDIX 3F*) for analysis of physiological functions (Schonthal, 1998) and/or phosphoproteins (Chapter 18).

The maximum (final) concentration of DMSO in the assay should not exceed 1.0% (v/v).

Inactive analogs, such as 1-norokadaone, should also be analyzed at a final concentration of $10\ \mu\text{M}$.

INHIBITION OF PP1/PP2A ACTIVITY IN VITRO WITH MICROCYSTIN-LR

Microcystin-LR, a cyclic heptapeptide produced by the blooms of cyanobacteria in freshwater lakes, is a member of a family of cyclic peptide products that inhibit PP1 and PP2A at nanomolar or subnanomolar concentrations. Many members of this family, including the cyclic pentapeptides, nodularins, show similar phosphatase target specificity. Liver cells possess a unique uptake mechanism for these compounds. Hence, microcystins are potently hepatotoxic but in general, microcystins do not enter most tissues and cells.

The broad specificity of microcystin-LR for PP1, PP2A, and other PP2A-like phosphatases as well as its remarkable potency and stability in buffered solution (compared to OA or CA) makes this compound an ideal reagent for preservation of cellular protein phosphorylations in lysates and during lengthy chromatographic purification of phosphoproteins. Microcystin-LR is also significantly less expensive than CA or OA and thus well suited for inclusion in large volumes of buffers used for dialysis and protein chromatography. Moreover, in contrast to OA and CA, no special precautions are needed in the storage of microcystin-LR, although prolonged storage of stock solutions should be at -20°C . It is, however, necessary to take some steps to avoid ingestion of this compound as toxicological studies show that very small quantities of microcystins can lead to serious liver failure and extensive tissue damage.

X-ray crystallography defined the binding site for microcystin-LR in the catalytic center of PP1 α . Other studies suggest that microcystin-LR forms a stable adduct with a highly conserved cysteine adjacent to the PP1 catalytic site. This process is spontaneous and slow and plays no role in phosphatase inhibition. However, the possibility remains that prolonged exposure (days to weeks) of phosphatases to microcystins generates the enzyme-toxin adduct and limits subsequent studies. For example, following short-term

exposure to microcystins, the observed phosphatase inhibition is readily reversed by dilution or dialysis. Significant adduct formation may prevent enzyme reactivation and reversal of covalent modifications.

Materials

Tissue or cells

0.1 or 1.0 mM microcystin-LR (Alexis Biochemicals or Sigma) in DMSO (store at -20°C)

1. Homogenize tissue or cells in buffers containing 1 μM microcystin-LR and process samples for functional or phosphoprotein analyses (Chapter 18).

BASIC PROTOCOL 3

INHIBITION OF CELLULAR PP2B/CALCINEURIN ACTIVITY WITH CYCLOSPORIN A

PP2B, also known as calcineurin, is a Ca^{2+} /calmodulin-activated protein serine/threonine phosphatase. PP2B was thus the first phosphatase determined to be directly controlled by a second messenger. PP2B is the therapeutic target for the two major immunosuppressants, cyclosporin A and FK506. These drugs, however, do not directly interact with the PP2B catalytic subunit but function through intermediate protein receptors collectively known as immunophilins to inhibit enzyme activity. Distinct immunophilins bind cyclosporin A and FK506. However, both drug/receptor complexes are effective PP2B inhibitors. The unique activity of these compounds on T-lymphocytes is due to their ready access to T-cells, the low abundance of their target, PP2B, in T-lymphocytes, and the key role of PP2B in activating cytokine genes that promote T-cell proliferation. PP2B and the immunophilins are, however, widely expressed in mammalian tissues and PP2B also plays an important role in regulating the physiology of mammalian heart, brain, and other tissues. While the reduced access within various tissues may limit the utility of these compounds in animals, both cyclosporin A and FK506 are excellent inhibitors of PP2B functions in experimental systems such as tissue slices and cultured cells.

Cyclosporin A has an apparent IC_{50} value for PP2B inhibition ranging from 10 to 100 nM. The use of cyclosporin A is limited by its poor solubility in aqueous solutions and its slow entry into mammalian cells. In contrast, FK506 is a ten-fold more potent PP2B inhibitor and enters cells rapidly. However, commercial availability of FK506 is still restricted, therefore, this protocol focuses on cyclosporin A, which is widely available from many different commercial sources. There are also numerous inactive analogs of cyclosporin available for use as controls. Many of these inactive analogs bind effectively to the target cyclophilin but fail to inhibit PP2B activity. In contrast, another immunosuppressant, rapamycin, binds to the same immunophilin as FK506 but does not inhibit PP2B and can be used as a control for FK506. Rapamycin is a potent inhibitor of the signaling pathway controlled by the protein kinase, TOR (target of rapamycin), and therefore, has its own effects on cellular protein phosphorylation.

Materials

0.1 mM cyclosporin A (Alexis Biochemical or Sigma) in DMSO

Cultured cells

1. Add 0.1 mM cyclosporin A (final concentration 0.1 to 1 μM) to culture medium or exchange with fresh medium containing cyclosporin A. Incubate cells for 60 min prior to harvesting for analysis.

The slow entry of cyclosporin A into cells often requires incubating cells and tissue slices for periods >1 hr, as very few physiological events respond to this drug in periods under 30 min.

IN VITRO INHIBITION OF PP2B/CALCINEURIN ACTIVITY WITH CYPERMETHRIN

BASIC
PROTOCOL 4

Cypermethrin and other type II pyrethroids are synthetic insecticides that show activity as PP2B inhibitors and are available from many commercial sources. In addition, a number of pyrethroid analogs, resmethrin, allethrin, and permethrin, show reduced or no activity as PP2B inhibitors and can be used as controls. This is particularly important for cellular studies using cypermethrin, as many of the type-II pyrethroids have a wide range of physiological effects including oxidative stress, cytotoxicity, and tumorigenesis. Only a few of these are likely to be mediated by PP2B inhibition. Given the availability of cyclosporin A, a better characterized and more potent PP2B inhibitor, the value of using pyrethroids in mammalian cells is questionable.

Materials

- 1.0 mM cypermethrin (Alexis Biochemicals) in DMSO and store in dark glass containers pretreated with 1% polyethylene glycol (PEG)
- Sample cell extracts or subcellular fractions

NOTE: Cypermethrin and other type II pyrethroids are sold in both solid and liquid form. The liquid form represents cypermethrin dissolved in organic solvents, such as DMSO, acetone, or ethanol. Hydrophobicity of these compounds means that long-term storage in plastic or polyethylene containers should be avoided because these compounds adhere to such surfaces and may be slowly extracted from solution. Many pyrethroids are also light-sensitive. If stored in dark glass containers pretreated with 1% (w/v) PEG to prevent surface adherence, stock solutions of cypermethrin are stable for many months. For cellular studies using cypermethrin, the compound should be added directly to the culture medium as its PP2B inhibitory activity is reduced when made up in fresh medium. While the underlying basis for a time-dependent loss of cypermethrin activity is not clear, this provides a strong argument for using cypermethrin and the pyrethroid inhibitors solely for in vitro studies of PP2B function.

1. Add 1.0 mM cypermethrin (10 to 20 μ M final concentration) to cell extracts or subcellular fractions and analyze phosphoproteins within a period not exceeding 60 min.

Cypermethrin addition in vitro should result in rapid PP2B inhibition that is evident within minutes. Prolonged incubations with this compound are not recommended.

INHIBITION OF PROTEIN TYROSINE PHOSPHATASES WITH SODIUM ORTHOVANADATE

BASIC
PROTOCOL 5

Protein tyrosine phosphatases (PTPases) represent a larger family of enzymes derived from a common ancestral gene and share no structural homology with the protein serine/threonine phosphatases (Sefton and Shenolikar, 1995). Thus, PTPases are not sensitive to any of the compounds described above. PTPases share a common catalytic mechanism that involves the formation of an enzyme phospho-intermediate with the modification of a conserved cysteine within the catalytic cleft. While cysteine-modifying reagents, such as phenyl arsene oxide, and metals such as zinc chloride, potently inhibit PTPases, these reagents are non-specific and likely target many proteins other than PTPases and therefore, will not be further discussed here.

Currently, the most useful reagent for inhibiting PTPases in vivo and in vitro is vanadate, a phosphate analog. One might predict that vanadate would also be non-specific and inhibit many phosphotransferases and phosphohydrolases, including protein serine/threonine phosphatases. However, the remarkable chemistry of the vanadium metal

Analysis of
Protein
Phosphorylation

18.10.7

ion (Gordon, 1991) can be exploited to produce a preferential and effective PTPase inhibitor.

Materials

Sodium orthovanadate (Sigma)

0.1 M NaOH

Sample cells

Boiling water bath or heating block

Additional reagents and equipment for metabolic labeling (Sefton, 1997) or immunoblotting with an anti-phosphotyrosine antibody (DiGiovanna et al., 1996)

1. Dissolve sodium orthovanadate (200 mM final concentration) in sterile, distilled water (often a yellowish solution) and adjust pH of the solution to 10.0 with 0.1 M NaOH. Place solution in a boiling water bath until the solution clarifies or becomes translucent. Readjust pH to 10.0 and store the stock solution in a flint glass container at room temperature or dispense into aliquots in plastic containers and store at -20°C .

Sodium orthovanadate, whether solid or in solution at neutral pH, slowly oxidizes and acquires a yellow color. The presence of thiols, often used to maintain protein structure and function following the homogenization of tissues or cells, promotes a rapid (<30 min) conversion to the oxycation, vanadyl. Finally, vanadate, particularly in solutions below 0.1 mM at neutral pH, also undergoes polymerization, which is also associated with increased yellow color. All of these factors result in reduced vanadate activity as a PTPase inhibitor. While vanadate slowly reverts to a colorless active monovalent form, especially on dilution, this process is greatly accelerated by increasing pH and temperature, as described above.

2. Incubate cells with 0.1 mM orthovanadate for periods not to exceed 2 to 3 hr before analyzing tyrosine-phosphorylated proteins by metabolic labeling (Sefton, 1997) or immunoblotting with an anti-phosphotyrosine antibody (DiGiovanna et al., 1996).

Maximal inhibition of cellular PTPase activity often requires up to 1.0 mM vanadate. This concentration of vanadate is also chosen to preserve protein tyrosine phosphorylation following cell disruption. However, much lower concentrations of vanadate are recommended for the treatment of intact cells to limit or avoid cytotoxicity. Even 0.1 mM vanadate is detrimental to cell function if the duration of treatment extends for several hours or days. This, however, makes the complete inhibition of PTPases achieved in cell lysates and other fractions by 1.0 mM vanadate easier to interpret than the results of cellular studies where changes in cell function cannot always be attributed to PTPase inhibition. Finally, some protein serine/threonine phosphatases display low but measurable PTPase activity in vitro and are not inhibited by 1.0 mM vanadate. These enzymes may account for slow dephosphorylation of tyrosine-phosphorylated proteins during lengthy isolation procedures.

COMMENTARY

Background Information

In contrast to the large number (300 to 500) of genes encoding mammalian protein kinases, many fewer genes encode the catalytic subunits of protein serine/threonine phosphatases (Cohen, 1997). Thus, the functional diversity of protein serine/threonine phosphatases arises through other mechanisms that include a rapidly expanding number of regulatory subunits that may dictate the substrate specificity and subcellular localization of the phosphatase catalytic subunits (Cohen, 2002). This means that the approaches outlined in this unit for

inhibiting protein serine/threonine phosphatases, while highly effective, may not distinguish between the functions of various intracellular pools of these enzymes. In addition, many phosphatase inhibitors are highly hydrophobic and may display significant differences in accessing intracellular pools of phosphatases. Hence, these protocols are most appropriately used for detecting and defining the physiological role of the phosphorylated protein substrate as opposed to making conclusions regarding the types or pools of phosphatases involved.

While more genes encode protein tyrosine phosphatases than serine/threonine phosphatases, these too fall far short of the numbers representing the complimentary protein tyrosine kinases. Mammalian PTPases can be separated into two broad groups, transmembrane “receptor” PTPases and soluble PTPases (Fischer et al., 1991). Many transmembrane PTPases contain two catalytic domains and emerging studies suggest that only one of these domains plays an active role in enzyme catalysis. The “inactive” catalytic domain is thought to regulate the function of receptor PTPases. It is unclear whether the second catalytic site functions as a decoy for PTPase inhibitors or if other mechanisms account for the apparent reduced sensitivity of transmembrane PTPases to vanadate. Considerable anecdotal evidence points to the preferential inhibition of soluble PTPases in cells treated with vanadate.

While the protocols discussed in this unit focus on a few pharmacological reagents, more potent and selective phosphatase inhibitors are likely to become available to investigators in the near future. The driving force behind the development of new phosphatase inhibitors is the desire by pharmaceutical companies to provide new therapies for human diseases associated with defects in signal transduction. The proof-of-principle for phosphatase inhibitors as potential medicines has already been established by the two PP2B/calcineurin inhibitors, cyclosporin A and FK506, which are the most effective immunosuppressants currently available. However, both compounds function by an indirect mechanism (i.e., via cellular receptors or immunophilins) to inhibit PP2B activity and both cause side-effects or toxicity in humans. Thus, there is a continued need for novel PP2B/calcineurin inhibitors as safe and improved immunosuppressants. With the availability of three-dimensional structures for PP2B/calcineurin and other phosphatases with bound inhibitors, improvements in structure-activity profiles for phosphatase inhibitors are inevitable.

The current paucity of reagents to inhibit protein tyrosine phosphatases (PTPases) should change very rapidly in the coming years as many pharmaceutical companies are currently embarked on the search for PTPase inhibitors (Shenolikar and Brautigan, 2000). With growing evidence that a specific PTPase, PTP1B, antagonizes insulin signaling in many mammalian tissues, several compounds have been identified that selectively target this phosphatase (Goldstein, 2001). Whether these com-

pounds eventually emerge as therapies for diabetes is uncertain, but they will undoubtedly yield excellent tools for the investigation of signaling pathways regulated by PTP1B and other PTPases.

Despite having a common catalytic mechanism, some phosphatases have been defined as “dual-specificity” in that they dephosphorylate adjacent phosphotyrosines and phosphothreonines in their target substrates, many of which control gene expression and cell proliferation. Perhaps due to differences in the architecture of their catalytic centers, dual-specificity phosphatases also display a decreased sensitivity to vanadate. Recent studies show that some members of the PTPase family do not act on phosphoproteins but rather they dephosphorylate phospholipids (Maehama et al., 2001). One example of this is PTEN, a tumor-suppressor gene that is mutated in many human cancers. PTEN functions as a lipid phosphatase that antagonizes insulin signaling. Development of selective PTEN inhibitors may also yield new therapies for diabetes. The ability of vanadate to inhibit both PTP1B and PTEN may account for its ability to elicit physiological effects similar to insulin and other growth factors.

The focus of this unit on a few selected protein serine/threonine phosphatase inhibitors should not minimize the value of a multitude of other protein serine/threonine phosphatase inhibitors currently available to researchers. Several hundred natural products appear to elicit their effects on mammalian cell physiology by inhibiting phosphatases. Such compounds include antibiotics such as tautomycin, insecticides such as cantharidins, and other structurally diverse compounds (McCluskey et al., 2002). Among these, one particular compound bears special mention, fostreicin, produced by a strain of *Streptomyces*, may be the most specific currently available inhibitor of the PP2A-like enzymes. Fostreicin demonstrates an IC_{50} value for PP2A of 1 to 3 nM while $>100 \mu\text{M}$ fostreicin is needed to inhibit the structurally related serine/threonine phosphatase, PP1. Use of fostreicin is, however, limited by its propensity to be inactivated by oxidation and thus must be stored in the presence of a large excess of antioxidants. Commercially available fostreicin, therefore, contains numerous additives, making it very difficult to use as a specific reagent in cellular studies. Nevertheless, it is likely that selective PP1 and PP2A inhibitors can and will be developed (see Key References). Compared to most

current serine/threonine phosphatase inhibitors, which display cytotoxicity and tumor promoter activity, clinical trials established fostreicin as a potential anti-neoplastic drug. This also hints at the possibility that more selective PP1 or PP2A inhibitors may be better experimental tools than currently available compounds whose cytotoxicity may be a result of inhibiting many different protein serine/threonine phosphatases. On the other hand, it is the remarkable chemistry of the vanadium metal ion that makes the current compounds such effective tools for analyzing highly labile and difficult-to-detect cellular phosphorylation events.

Finally, a number of broad-acting or non-specific reagents have been used to inhibit protein phosphatases. These include sodium fluoride (50 mM) and sodium pyrophosphate or glycerophosphate (10 mM). These compounds suppress protein dephosphorylations in cell lysates and fractions, although not as effectively as other inhibitors. They can serve as alternatives for inhibiting phosphatases *in vitro* but show no utility in phosphoprotein analysis in intact cells.

Critical Parameters

Treatment of cells with the cell-permeable phosphatase inhibitors described in this unit is often so effective that it saturates and occludes the physiological signaling pathways, making studies that utilize combinations of hormones and phosphatase inhibitors difficult to interpret. Many mammalian proteins are regulated both positively and negatively by phosphorylation, generally occurring at distinct sites. Thus, the treatment of cells with phosphatase inhibitors may facilitate both types of modifications and yield a compound functional readout, which may bear little or no relationship to the physiological mechanisms that orchestrate the orderly control of protein function. Even when utilizing phosphatase inhibitors *in vitro*, artifactual or post-homogenization phosphorylations can result from inappropriate exposure of proteins to kinases normally located in distinct subcellular compartments and therefore are not normally physiological regulators of these proteins. Thus, this unit stresses the detection of cellular protein phosphorylation and events bearing on the functional role of such modifications.

The pharmacological approach to studying cellular protein phosphorylation as described in this unit can be employed by investigators from many backgrounds. However, further op-

timization of the outlined experimental strategy may be necessary for individual cell types, where permeability of the compounds and the array of phosphatases inhibited may dictate unique physiological consequences. Thus, the investigator is encouraged to undertake pilot studies that determine the optimal concentrations of the compound and the duration of cell exposure required to obtain a measurable and meaningful cellular response. Also, as emphasized above, while the cell-permeable phosphatase inhibitors display some target specificity, they cannot unequivocally define the phosphatases that regulate specific phosphoproteins or physiological events. Finally, the investigator should be aware that many of the compounds described above display some capacity to bind proteins other than phosphatases albeit with much lower affinity. Thus, control experiments should utilize compounds, such as the inactive okadaic acid or cyclosporin analogs, that fail to inhibit the target phosphatases. These controls should provide a more compelling argument supporting the biological effects uniquely observed with the active compounds as reflecting the inhibition of the target protein phosphatases.

To define the specific phosphatases, the investigator must turn to more molecular approaches that may involve the overexpression of phosphatases (currently a very challenging task), or the endogenous protein inhibitors that show exquisite selectivity for individual phosphatases (Oliver and Shenolikar, 1998). An alternate to overexpression may be the microinjection of phosphatases and the inhibitor proteins, an approach that has been successful in modulating cellular phosphatase activity (Alberts et al., 1993; Mulkey et al., 1994; Morishita et al., 2001). The difficulty of introducing protein inhibitors into cells, whether by overexpression or microinjection, is that most, if not all, of these proteins are tightly regulated by covalent modification. The functions of various inhibitors may be enhanced or dampened by reversible protein phosphorylation, making it difficult to predict their effects on phosphatases. Yet another approach is the use of antisense oligonucleotides or interference RNAs to “knock-down” expression of specific protein phosphatases. While this approach harbors tremendous specificity, it is difficult to control the effectiveness of these inhibitors which lower phosphatase expression, and elicit appropriate physiological responses that cannot be readily predicted. The use of dominant-negative reagents, including inactive phos-

phatase catalytic subunits or fragments of phosphatases, has shown some promise in suppressing phosphatase functions, but these reagents most likely rely on their competition with endogenous phosphatases for regulators rather than substrates, thereby making it difficult to predict the physiological outcome. However, catalytically inactive phosphatases have proven very useful in deciphering the functions of PTPases. These inactive phosphatases function as substrate traps and form stable complexes with cellular phosphoproteins, allowing identification of the potential physiological substrates of the PTPase. Even in this situation, there are serious concerns that overexpression of inactive PTPases may distort the natural specificity of these enzymes and yield spurious or unreliable data. Protein serine/threonine phosphatases themselves also occasionally form stable complexes with cellular substrates and such complexes can be isolated using affinity chromatography based on phosphatase inhibitors, such as microcystin-LR, immobilized on Sepharose or other dextran surfaces. Such affinity matrices are not commercially available but a number of publications describe the synthesis and use of such resins. The reader may refer to Campos et al. (1996) or Damer et al. (1998). However, this approach has thus far been more successful in identifying phosphatase regulators than substrates.

Troubleshooting

The use of phosphatase inhibitors as described in this unit is associated with several problematic issues regarding compound stability and permeability. First and foremost, many of these compounds are highly hydrophobic. Some are light-sensitive and others difficult to store for long periods. These factors contribute to the problem of defining their precise concentration in culture media and buffers and predicting their effects on target phosphatases. In cellular studies, there are additional problems in that external inhibitor concentration is not an accurate predictor of the intracellular concentration of compounds, which are not evenly distributed throughout the cell. Thus, the literature reports widely varying IC_{50} values for inhibition of purified phosphatases and the sensitivity of known PP1 and/or PP2A-regulated events in cells. Thus, the protocols described in this unit deliberately err on the side of using excess phosphatase inhibitors. While this ensures effective phosphatase inhibition, additional problems may arise in terms of the cytotoxicity elicited by these reagents, which varies

widely among different cell types. Quiescent, confluent and highly differentiated cells appear more susceptible to programmed cell death in response to phosphatase inhibitors such as OA and CA, although the prolonged overactivation of multiple signaling pathways by these inhibitors may also contribute to the cytotoxicity observed in actively growing cells. Thus, investigators are encouraged to determine the lowest concentrations of inhibitors and shortest duration of cell exposure necessary for a consistent and measurable physiological response.

Anticipated Results

Phosphatase inhibitors that inhibit the major mammalian protein serine/threonine or tyrosine phosphatases unmask an array of kinase cascades and thereby enhance the phosphorylation of a wide range of proteins. This in turn permits the investigator to detect and analyze novel phosphorylation events. By selecting specific phosphatase inhibitors and restricting their concentrations, the investigator can also inhibit a subset of cellular phosphatases yielding more defined results by not only limiting the panel of phosphoproteins visualized but also the sites modified on the given phosphoprotein. For an example of a specific phosphatase inhibition experiment, refer to Haysstead et al. (1989). Functional effects associated with the rapid phosphorylation-dephosphorylation of cellular proteins are more readily visualized in the presence of phosphatase inhibitors. These experiments can establish the importance of protein phosphorylation in physiological events where the precise protein target or phosphorylation event has not been identified. However, readouts such as gene expression that result from increased phosphorylation of nuclear proteins often require long exposure of cells to phosphatase inhibitors and thereby risk the potentially confounding effects of cytotoxicity. Here, a compromise must be reached and inhibitor concentrations lowered to levels that maintain cell viability and still promote the transcription activation of target genes.

Time Considerations

A common scenario for the experiments described above is that cells are metabolically labeled using ^{32}P -orthophosphate for periods between 90 and 120 min (Sefton, 1997). Phosphatase inhibitors are then added to the media and the cells incubated for a further 30 min. Following homogenization, the phosphoproteins are analyzed using SDS-PAGE (UNIT 10.2A)

and autoradiography or phosphorimaging, which adds an additional 2 to 3 hr. Thus, the total time required may be minimally estimated at 6 hr.

In the absence of metabolic labeling, SDS-PAGE may be followed by electrophoretic transfer and analysis of changes in protein mobility reflecting increased phosphorylation or the modification of specific proteins. Phosphorylation sites may be identified by western immunoblotting using phosphospecific antibodies. These approaches provide different pieces of information but carry similar total time considerations.

Literature Cited

- Alberts, A.S., Thorburn, A.M., Shenolikar, S., Mumby, M.C., and Feramisco, J.R. 1993. Regulation of cell cycle progression and nuclear affinity of the retinoblastoma protein by protein phosphatases. *Proc. Natl. Acad. Sci. U.S.A.* 90:388-392.
- Campos, M., Fadden, P., Alms, G., Qian, Z., and Haystead, T.A. 1996. Identification of protein phosphatase-1-binding proteins by microcystin-biotin affinity chromatography. *J. Biol. Chem.* 271:28478-28484.
- Cohen, P.T.W. 1997. Novel protein serine/threonine phosphatases: Variety is the spice of life. *Trends Biochem. Sci.* 22:245-251.
- Cohen, P.T. 2002. Protein phosphatase 1- targeted in many directions. *J. Cell Sci.* 115:241-256.
- Cohen, P., Klumpp, S., and Schelling, D.L. 1989. An improved procedure for identifying and quantitating protein phosphatases in mammalian tissues. *FEBS Lett.* 250:596-600.
- Damer, C.K., Partridge, J., Pearson, W.R., and Haystead, T.A. 1998. Rapid identification of protein phosphatase 1-binding proteins by mixed peptide sequencing and data base searching. Characterization of a novel holoenzymic form of protein phosphatase 1. *J. Biol. Chem.* 273:24396-24405.
- DiGiovanna, M.P., Roussel, R.R., and Stern, D.F. 1996. Production of antibodies that recognize specific tyrosine-phosphorylated peptides. In *Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, D.W. Speicher, and P.T. Wingfield, eds.), pp. 13.6.1-13.6.13. John Wiley & Sons, New York.
- Fischer, E.H., Charbonneau, H., and Tonks, N.K. 1991. Protein tyrosine phosphatases: A diverse family of intracellular and transmembrane enzymes. *Science* 253:401-406.
- Fujiki, H., Sukanuma, M., Suguri, H., Yoshizawa, S., Takagi, K., Uda, N., Wakanatsu, K., Yamada, K., Murata, M., and Yasumoto, T. 1988. Diarrhetic shellfish toxin, dinophysistoxin-1, is a potent tumor promoter on mouse skin. *Jap. J. Cancer Res.* 79:1089-1093.
- Goldstein, B.J. 2001. Protein-tyrosine phosphatase 1B (PTP1B): A novel therapeutic target for type 2 diabetes mellitus, obesity and related states of insulin resistance. *Curr. Drug Target-Immun. Endocrin. Metab. Disorders* 1:265-275.
- Gordon, J.A. 1991. Use of vanadate as protein-phosphotyrosine phosphatase inhibitor. *Methods Enzymol.* 201:477-482.
- Haystead, T.A., Sim, A.T., Carling, D., Honnor, R.C., Tsukitani, Y., Cohen, P., and Hardie, D.G. 1989. Effects of the tumour promoter okadaic acid on intracellular protein phosphorylation and metabolism. *Nature* 337:78-81.
- Kiguchi, K., Glesne, D., Chubb, C.H., Fujiki, H., and Huberman, E. 1994. Differential induction of apoptosis in human breast tumor cells by okadaic acid and related inhibitors of protein phosphatases 1 and 2A. *Cell Growth & Differentiation* 5: 995-1004.
- Maehama, T., Taylor, G.S., and Dixon, J.E. 2001. PTEN and myotubularin: Novel phosphoinositide phosphatases. *Annu. Rev. Biochem.* 70:247-79.
- McCluskey, A., Sim, A.T.R., and Sakoff, J.A. 2002. Serine-threonine protein phosphatase inhibitors: Development of potential therapeutic strategies. *J. Medicinal Chem.* 45:1151-1175.
- Morimoto, Y., Ohba, T., Kobayashi, S., and Haneli, T. 1997. The protein phosphatase inhibitors okadaic acid and calyculin A induce apoptosis in human osteoblastic cells. *Exp. Cell Res.* 230:181-186.
- Morishita, W., Connor, J.H., Xia, H., Quinlan, E.M., Shenolikar, S., and Malenka, R.C. 2001. Regulation of synaptic strength by protein phosphatase 1. *Neuron* 32:1133-1148.
- Mulkey, R.M., Endo, S., Shenolikar, S., and Malenka, R.C. 1994. Involvement of a calcineurin/inhibitor-1 phosphatase cascade in hippocampal long-term depression. *Nature* 369:486-488.
- Nakagama, H., Kaneko, S., Shima, H., Fukuda, H., Kominami, R., Sugimura, T., and Nagao, M. 1997. Induction of minisatellite mutation in NIH3T3 cells by treatment with the tumor promoter okadaic acid. *Proc. Natl. Acad. Sci. U.S.A.* 94:10813-10816.
- Nam, K.Y., Hiro, M., Kimura, S., Fujiki, H., and Imanishi, Y. 1990. Permeability of a non-TPA-type tumor promoter, okadaic acid, through lipid bilayer membrane. *Carcinogenesis* 11:1171-1174.
- Namboodiripad, A.N. and Jennings, M.L. 1996. Permeability characteristics of erythrocyte membrane to okadaic acid and calyculin A. *Am. J. Physiol.* 270:C449- C456.
- Nashiwaki, S., Fujiki, H., Sukanuma, M., Furuya-Suguri, H., Matsushima, R., Iida, Y., Ojika, M., Yamada, K., Uemura, D., and Yasumoto, T. 1991. Structure-activity relationship within a series of okadaic acid derivatives. *Carcinogenesis* 11:1837-1841.
- Oliver, C.J. and Shenolikar, S. 1998. Physiologic importance of protein phosphatase inhibitors. *Frontiers in Biosci.* 3:D961-D972.

- Schonthal, A.H. 1998. Analyzing gene expression with the use of serine/threonine phosphatase inhibitors. *Methods Mol. Biol.* 93:35-40.
- Sefton, B.M. 1997. Labeling cultured cells with $^{32}\text{P}_i$ and preparing cell lysates for immunoprecipitation. *In* Current Protocols in Protein Science (J.E. Coligan, B.M. Dunn, D.W. Speicher, and P.T. Wingfield, eds.), pp. 13.2.1-13.2.8. John Wiley & Sons, New York.
- Sefton, B.M and Shenolikar, S. 1995. Overview of protein phosphorylation. *In* Current Protocols in Protein Science (J.E. Coligan, B.M. Dunn, D.W. Speicher, and P.T. Wingfield, eds.), pp. 13.1.1-13.1.5. John Wiley & Sons, New York.
- Shenolikar, S. and Nairn, A.C. 1991. Protein phosphatases: Recent progress. *Adv. Second Messenger Phosphoprotein Res.* 23:1-121.
- Shenolikar, S. and Brautigan, D.L. 2000. Targeting protein phosphatases: Medicines for the new millennium. *Science (STKE)* Nov. 2000. <http://stke.sciencemag.org/>.
- Suganuma, M., Suttajit, M., Suguri, H., Ojika, M., Yamada, K., and Fujiki, H. 1989. Specific binding of okadaic acid, a new tumor promoter in mouse skin. *FEBS Lett.* 250:615-618.
- Suganuma, M., Fujiki, H., Furuya-Suguri, H., Yoshizawa, S., Yasumoto, S., Kato, Y., Fusetani, N., and Sugimura, T. 1990. Calyculin A, an inhibitor of protein phosphatases, is a potent tumor promoter on CD-1 mouse skin. *Cancer Res.* 50:3521-3525.

Key References

Current Medicinal Chemistry. November, 2002, vol. 9.

A special "hot topics" issue, which is devoted to the synthesis of serine/threonine phosphatase inhibitors.

Contributed by Douglas C. Weiser and
Shirish Shenolikar
Duke University Medical Center
Durham, North Carolina

Design and Use of Analog-Sensitive Protein Kinases

Many protein kinases can be engineered to accept analogs of ATP that are not efficiently used by wild-type kinases. These engineered kinases, which are referred to as “analog-sensitive” or “*-as*” alleles, are also often sensitive to protein kinase inhibitor variants that do not block the activity of nonmutant kinases. Selective *in vitro* use of radiolabeled ATP analogs by *-as* kinases can be exploited to identify the direct phosphorylation targets of individual kinases in complex extracts. In organisms in which it is practical to replace wild-type kinase genes with engineered alleles, the *in vivo* activity of an *-as* kinase can be reversibly blocked with an allele-specific inhibitor. Thus, analog-sensitive kinases can be effective tools for discovery of the cellular functions and phosphorylation targets of individual enzymes. A theoretical background for the design and use of these alleles is discussed in the Commentary section at the end of this unit, as are strategies for construction of candidate *-as* alleles of any kinase.

In vivo use of analog-sensitive kinase alleles requires access to kinase inhibitor analogs. This unit describes protocols for synthesis of 1NM-PP1 and 1NA-PP1 (see Basic Protocols 1 and 2), which are among the most effective allele-specific inhibitors of *-as* kinases. These protocols assume knowledge of organic synthesis, and it is advisable to work with a chemist that can provide assistance if one is not familiar with organic chemistry methods. Under no circumstances should one attempt these protocols if not comfortable with safety procedures fundamental to synthetic organic chemistry. Much of the equipment required for these procedures is uncommon in molecular biology groups but readily accessible in chemistry laboratories. The authors have provided information relevant to analysis of NMR spectra for researchers familiar with this technique. A detailed description of product analysis is, however, beyond the scope of this section.

The authors have also supplied a protocol for production of γ -³²P-labeled *N*⁶(benzyl)ATP, an ATP analog that is commonly selectively used by *-as* kinases, from the ADP analog (see Basic Protocol 3 and Support Protocols 1 and 2). This procedure employs enzymatic phosphorylation of *N*⁶(benzyl)ADP by immobilized nucleoside diphosphate kinase (NDPK), and can be used to produce other γ -labeled ATP analogs from ADP analogs (Mourad and Parks, 1966; Kraybill et al., 2002). Commercial availability of *N*⁶(benzyl)ADP, as well as other ADP analogs, is still somewhat limited. A method for synthesis of *N*⁶(benzyl)ADP has been published (Shah et al., 1997; Shah and Shokat, 2002). Labeling of ADP analogs by the procedure provided here requires moderately large quantities (~2.5 mCi) of [γ -³²P]ATP, and it is therefore advisable to ensure that this raises no issues with one's license to use radioactive materials.

A final protocol (see Basic Protocol 4) describes the assays for testing inhibition of analog-sensitive kinases in yeast.

NOTE: Although the protocols described in this unit are directly derived from published material, it is important to note that the use of analog-sensitive kinases, analog-sensitive kinase inhibitors, and ATP analogs is protected by U.S. patents 6,610,483; 6,521,417; 6,390,821 and 6,383,790, as well as patents in other countries, which are owned by Princeton University and exclusively licensed to Cellular Genomics Incorporated (CGI) of Branford, Connecticut. Information regarding accessing rights to make and use the technology can be obtained by contacting CGI.

SYNTHESIS OF 1NM-PP1

This protocol and the one that follows for 1NA-PP1 (see Basic Protocol 2) are based on previously published methods (Bishop et al., 1998; Bishop et al., 1999). The synthetic schemes are diagrammed in Figure 18.11.1. A representative NMR spectrum is presented in Figure 18.11.2.

Materials

1-naphthylacetic acid (Acros Organics)
Hexane
N,N-dimethylformamide (DMF; Aldrich)
Oxalyl chloride (Aldrich)
Tetrahydrofuran (THF; Aldrich)
Sodium hydride (Aldrich)
Malononitrile (Aldrich)
1 N H₂SO₄
Ethyl acetate
MgSO₄, anhydrous
1,4-dioxane (Aldrich)
Sodium bicarbonate (Aldrich)
Dimethyl sulfate (Aldrich)
Diethyl ether (Et₂O)
Silica Gel 60 (pore size 0.040 to 0.063 mm; Merck)
Ethanol (Aldrich)
Triethylamine (Aldrich)
tert-butylhydrazine hydrochloride (Aldrich)
Chloroform (CHCl₃)
1% (v/v) methanol in CHCl₃
Formamide (Aldrich)
Activated charcoal
Celite

Büchi Rotavapor Model R-200 or equivalent rotary evaporator
Ice bath
Separatory funnel
Oil baths, 80°, 100°C, 180°C
10-in. (25.4-cm) length × 2-in. (5.0-cm) i.d. chromatography column
TLC plates (Silica Gel F₂₅₄; EM Science) and tank
Reflux condenser
Filter paper
Buchner funnel

Synthesize 1-naphthylacetylmalononitrile

1. Dissolve 3.72 g 1-naphthylacetic acid (20.0 mmol) in 100 ml hexane.
2. Add 146 mg *N,N*-dimethylformamide (DMF; 2.0 mmol).
3. Add, dropwise, 12.7 g oxalyl chloride (100 mmol).
4. Wait for gas evolution to stop (~30 min), then remove solvent under reduced pressure in a Büchi Rotavapor or equivalent rotary evaporator.
5. Redissolve the residue (containing the acid chloride) in 30 ml tetrahydrofuran (THF).
6. Suspend 1.06 g sodium hydride (44.0 mmol) in 10 ml THF.

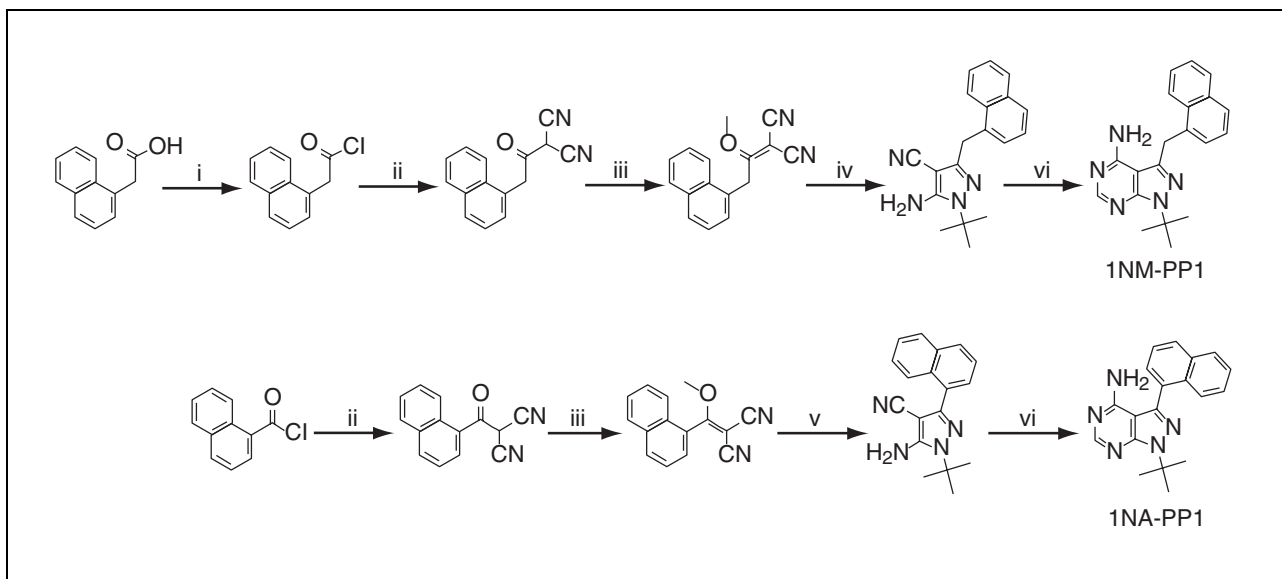


Figure 18.11.1 Synthetic schemes of 1NM-PP1 and 1NA-PP1. Conditions: (i) 5 eq oxalyl chloride, 0.1 eq DMF, room temperature, 1 hr; (ii) 2 eq NaH, 1 eq malononitrile, THF, 0°C to room temperature, 1 hr; (iii) 8 eq NaHCO₃, 5 eq dimethyl sulfate, dioxane/water (6:1), reflux, 1 hr; (iv) 2 eq triethylamine, 1 eq *tert*-butylhydrazine hydrochloride, EtOH, reflux, 1 hr; (v) 1 eq *tert*-butylhydrazine, DMF, room temperature, 1 hr; (vi) formamide, 180°C, 10 hr.

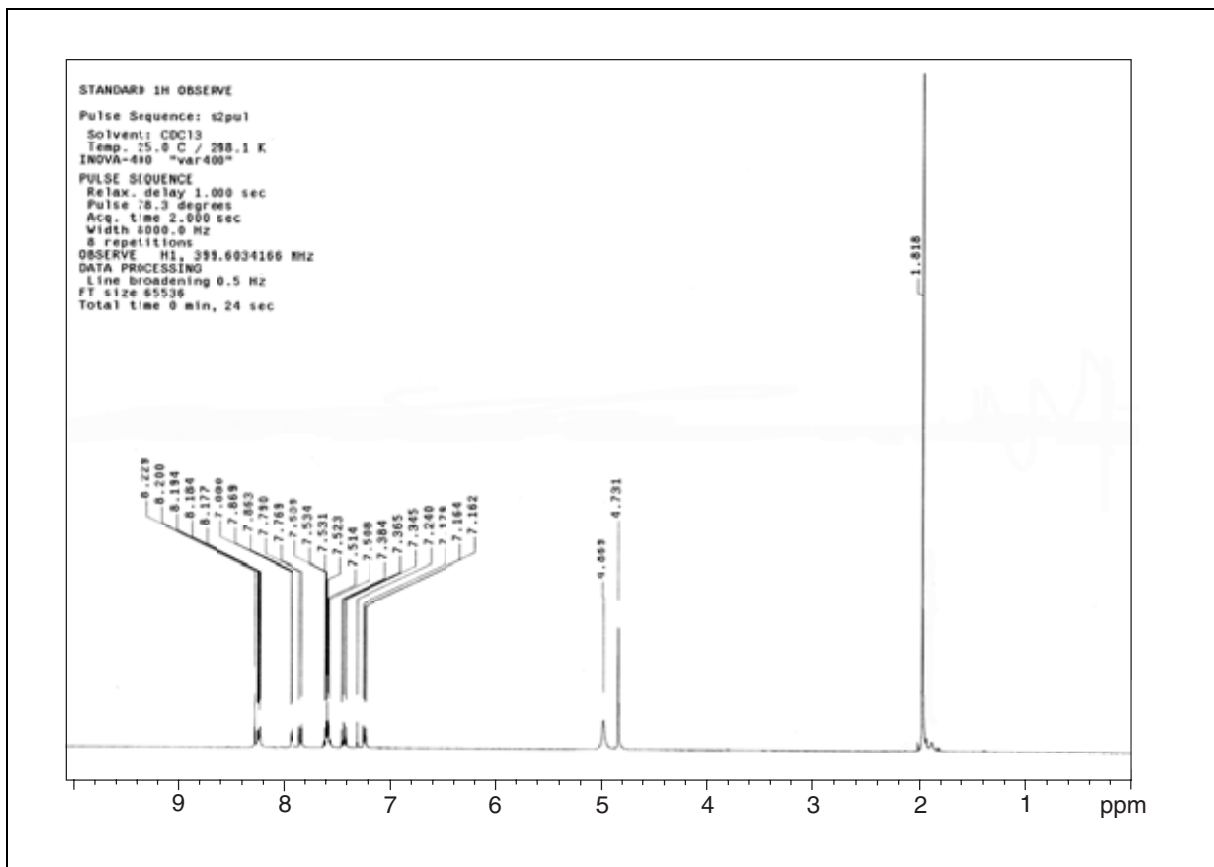


Figure 18.11.2 NMR data for 1-*tert*-butyl-3-naphthalen-1-ylmethyl-1*H*-pyrazolo[3,4-*d*]pyrimidin-4-ylamine (1NM-PP1). Compound is white powder; ¹H NMR (CDCl₃, 400 MHz) δ 1.82 (s, 9H), 4.73 (s, 2H), 4.87 (br s, 2H), 7.17 (d, J = 7 Hz, 1H), 7.37 (t, J = 8 Hz, 1H), 7.52 (m, 2H), 7.78 (d, J = 8 Hz, 1H), 7.87 (m, 1H), 8.19 (m, 1H), 8.23 (s, 1H); ¹³C NMR (CDCl₃, 100 MHz) δ 29.2, 32.7, 60.0, 101.1, 123.5, 125.6, 125.8, 126.2, 126.6, 128.2, 128.9, 131.9, 133.9, 134.0, 140.5, 154.5, 154.7, 157.6, HRMS (EI) molecular ion calculated for C₂₀H₂₁N₅ 331.17993, found 331.17951.

7. Dissolve 1.45 g malononitrile (22.0 mmol) in 10 ml THF and add dropwise to the NaH/THF suspension (step 6) on an ice bath.
8. Add the THF solution of acid chloride from step 5 dropwise to the malononitrile solution with vigorous stirring. Stir for 1 hr.
9. Add ~15 ml of 1 N H₂SO₄ and 50 ml water.
10. Extract three times, each time with 100 ml ethyl acetate, in a separatory funnel.

The ethyl acetate phase is the top phase.

11. Combine ethyl acetate extracts and dry by addition of ~5 g anhydrous MgSO₄.
12. Evaporate the solvent under reduced pressure in a Büchi Rotavapor or equivalent rotary evaporator, to yield 1-naphthylacetylmalononitrile, a yellow solid.

Yield should be ~70%.

Synthesize 1-naphthylmethyl(methoxy)methylidenemalononitrile

13. Dissolve 2.23 g 1-naphthylacetylmalononitrile (from step 12; 10.0 mmol) in a mixture of 18 ml dioxane and 3 ml water.
14. Add 6.8 g sodium bicarbonate.
15. Add 4.9 ml of dimethyl sulfate with vigorous stirring. Heat the reaction mixture at 100°C on an oil bath for 2 hr.
16. Cool the reaction mixture to room temperature.
17. Dilute with 70 ml water and extract three times, each time with 100 ml Et₂O, in a separatory funnel.
18. Combine the Et₂O extracts and dry over anhydrous MgSO₄.
19. Evaporate the Et₂O under reduced pressure in a Büchi Rotavapor or equivalent rotary evaporator to give an oil.
20. Prepare a 10-in. (25.4-cm) × 2-in. (5.0-cm) column containing 200 g Silica Gel 60, saturated with 1:1 Et₂O/hexane (the mobile phase).
21. Dissolve the oil from step 19 in a small amount of mobile phase and apply to column. Elute with mobile phase. Sample the eluate as it emerges from the column and analyze by thin-layer chromatography (TLC) using 1:1 Et₂O/hexanes as the mobile phase. Pool fractions containing a species with R_f of ~0.24.
22. Evaporate solvent under reduced pressure in a Büchi Rotavapor or equivalent rotary evaporator to produce a white crystalline solid 1-naphthylmethyl(methoxy)methylidenemalononitrile.

Yield should be ~75%.

Synthesize 5-amino-1-tert-butyl-3-(1'-naphthylmethyl)-4-cyano-1H-pyrazole

23. Dissolve 1.74 g of 1-naphthylmethyl(methoxy)methylidenemalononitrile (from step 22; 7.0 mmol) in 50 ml ethanol.
24. Add 1.51 g triethylamine (15.0 mmol) and 0.92 g *tert*-butylhydrazine hydrochloride (7.4 mmol). Reflux at 80°C on an oil bath for 1 hr.
25. Cool the mixture and evaporate the solvent under reduced pressure.
26. Suspend the residue in water and extract three times, each time with 50 ml CHCl₃, in a separatory funnel.

The CHCl₃ phase is the top phase.

27. Combine the CHCl_3 extracts, dry over anhydrous MgSO_4 , and evaporate the solvent under reduced pressure in a Büchi Rotavapor or equivalent rotary evaporator, to give a yellow solid.
28. Prepare a 10-in. (25.4-cm) \times 2-in. (5.0-cm) column containing 100 g Silica Gel 60, saturated with CHCl_3 (the mobile phase).
29. Dissolve the yellow solid in a minimal volume of CHCl_3 and apply to the column. Elute with 1% methanol in CHCl_3 . Sample the eluate as it emerges from the column and analyze by thin-layer chromatography (TLC) using 1% methanol in CHCl_3 as the mobile phase. Pool fractions containing a species with R_f of ~ 0.1 .
30. Evaporate solvent to obtain a white solid, 5-amino-1-*tert*-butyl-3-(1'-naphthylmethyl)-4-cyano-1*H*-pyrazole.

Yield should be $\sim 75\%$.

Synthesize 4-amino-1-*tert*-butyl-3-(1'-naphthylmethyl)-1*H*-pyrazolo[3,4-*d*]-pyrimidine (1NM-PP1)

31. Suspend 0.9 g 5-amino-1-*tert*-butyl-3-(1'-naphthylmethyl)-4-cyano-1*H*-pyrazole (from step 30; 3.4 mmol) in 20 ml of formamide.
32. Heat this suspension at 180°C on an oil bath, stirring with a reflux condenser for 10 hr.
33. Cool reaction mixture and dilute with 80 ml water.
34. Collect precipitate by filtration on paper in a Buchner funnel.
35. Dissolve precipitate in 40 ml room temperature ethanol.
36. Add 1 g powdered activated charcoal. Boil at 80°C for 10 min.
37. Pour a celite pad on a Buchner funnel using a slurry of celite in ethanol. Filter the hot mixture through this celite pad. Collect the filtrate.
38. Evaporate the solvent to obtain 1NM-PP1 as a white powder.

Expect yields of $\sim 75\%$.

SYNTHESIS OF 1NA-PP1

1NA-PP1 can be synthesized in the same manner as 1NM-PP1, with two differences. First, the acid chloride (1-naphthoyl chloride) is directly used in the first step of the synthesis, since it is commercially available. Second, a different condition is used in the reaction that forms the pyrazole ring, giving much better yields than the original procedure. A representative NMR spectrum is presented in Figure 18.11.3.

Materials

1-naphthoyl chloride (Aldrich)
 Tetrahydrofuran (THF; Aldrich)
 Sodium hydride (Aldrich)
 Malononitrile (Aldrich)
 1 N H_2SO_4
 Ethyl acetate
 MgSO_4 , anhydrous
 1,4-dioxane (Aldrich)
 Sodium bicarbonate (Aldrich)
 Dimethyl sulfate (Aldrich)

**BASIC
 PROTOCOL 2**

**Analysis of
 Protein
 Phosphorylation**

18.11.5

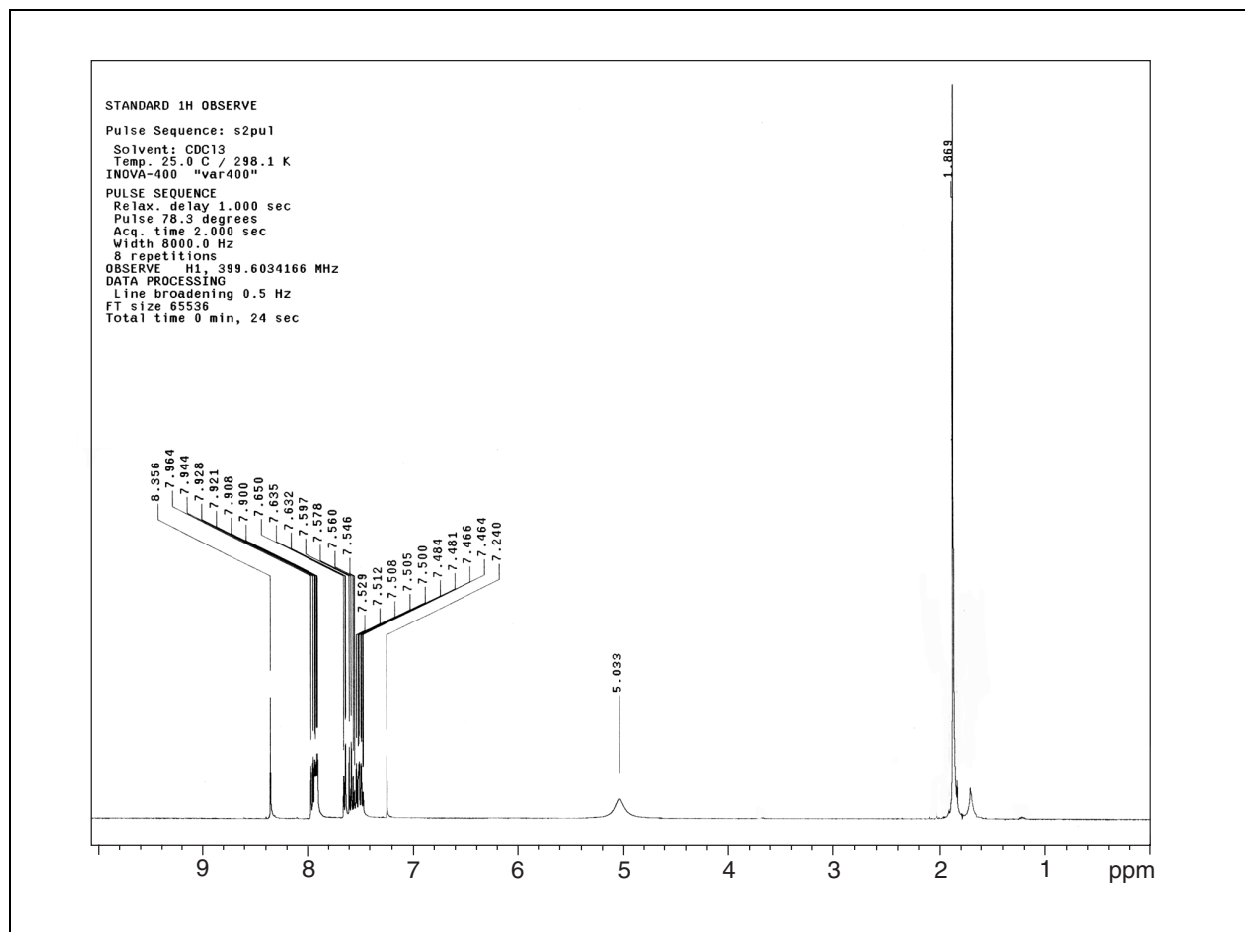


Figure 18.11.3 NMR data for 1-*tert*-butyl-3-naphthalen-1-yl-1*H*-pyrazolo[3,4-*d*]pyrimidin-4-ylamine (1NA-PP1). Compound is white powder; ¹H NMR (CDCl₃, 400 Mhz) δ 1.87 (s, 9H), 5.04 (br s, 2H), 7.05 (m, 2H), 7.58 (t, J = 8 Hz, 1H), 7.64 (d, J = 7 Hz, 1H), 7.92 (m, 2H), 7.95 (d, J = 8 Hz, 1H), 8.36 (s, 1H); ¹³C NMR (CDCl₃, 100 Mhz) δ 29.3, 60.6, 101.5, 125.5, 125.6, 126.5, 127.1, 128.4, 128.4, 129.6, 130.6, 131.8, 134.0, 140.4, 153.9, 154.7, 157.7, HRMS (EI) molecular ion calculated for C₁₉H₁₉N₅ 317.16247, found 317.16247.

Diethyl ether (Et₂O)
 Silica Gel 60 (pore size 0.040 to 0.063 mm; Merck)
 Sodium ethoxide (Acros Organics)
tert-butylhydrazine hydrochloride (Aldrich)
N,N-dimethylformamide (DMF; Aldrich)
 Chloroform (CHCl₃)
 1% (v/v) methanol in CHCl₃
 Formamide (Aldrich)
 Ethanol (Aldrich)
 Activated charcoal
 Celite
 Separatory funnel
 Oil baths, 80°, 100°, and 180°C
 10-in. (25.4-cm) length × 2-in. (5.0-cm) i.d. and 8-in. (20.3-cm) length × 1.5-in. (3.8-cm) i.d. chromatography columns
 TLC plates (Silica Gel 60 F₂₅₄; EM Science) and tank
 Reflux condenser
 Filter paper
 Buchner funnel

Synthesize 1-naphthoymalononitrile

1. Dissolve 3.80 g (20 mmol) 1-naphthoyl chloride in 30 ml tetrahydrofuran (THF).
2. Dissolve 1.06 g sodium hydride (44.0 mmol) in 10 ml THF.
3. Dissolve 1.45 g malononitrile (22.0 mmol) in 10 ml THF and add dropwise to the NaH/THF suspension on an ice bath.
4. Add the THF solution of 1-naphthoyl chloride from step 1 dropwise to the malononitrile solution with vigorous stirring. Stir for 1 hr.
5. Add ~15 ml of 1 N H₂SO₄ and 50 ml water.
6. Extract three times, each time with 100 ml ethyl acetate in separatory funnel.

The ethyl acetate phase is the top phase.

7. Combine ethyl acetate extracts and dry by addition of anhydrous MgSO₄.
8. Evaporate the solvent to yield 1-naphthoymalononitrile, a yellow solid.

Yield should be approximately 70%.

Synthesize 1-naphthyl(methoxy)methylidenemalononitrile

9. Dissolve 2.20 g 1-naphthoymalononitrile (from step 8; 10.0 mmol) in a mixture of 18 ml dioxane and 3 ml water.
10. Add 6.8 g sodium bicarbonate.
11. Add 4.9 ml of dimethyl sulfate with vigorous stirring. Heat the reaction mixture at 100°C on an oil bath for 2 hr.
12. Cool the reaction mixture to room temperature.
13. Dilute with 70 ml water and extract three times, each time with 100 ml Et₂O.
14. Combine the Et₂O extracts and dry over anhydrous MgSO₄.
15. Evaporate the Et₂O under reduced pressure to give an oil.
16. Prepare a 10-in. (25.4-cm) × 2-in. (5.0-cm) column containing 200 g Silica Gel 60, saturated with 1:1 Et₂O/hexane (the mobile phase).
17. Dissolve the oil from step 15 in a small amount of mobile phase and apply to column. Elute with mobile phase. Sample the eluate as it emerges from the column and analyze by thin-layer chromatography (TLC) using 1:1 Et₂O/hexanes as the mobile phase. Pool fractions containing a species with R_f of ~0.22.
18. Evaporate solvent under reduced pressure in a Büchi Rotavapor or equivalent rotary evaporator, to produce a white crystalline solid, 1-naphthyl(methoxy)methylidenemalononitrile.

Yield should be ~75%.

Synthesize 5-amino-1-tert-butyl-3-(1'-naphthyl)-4-cyano-1H-pyrazole

19. Add 0.38 g sodium ethoxide (5.5 mmol) and 0.68 g *tert*-butylhydrazine hydrochloride (5.5 mmol) to 10 ml DMF. Stir mixture at room temperature for 0.5 hr.
20. Filter the mixture. Save the filtrate, which contains *tert*-butylhydrazine.
21. Dissolve 1.17 g 1-naphthyl(methoxy)methylidenemalononitrile (from step 18; 5.0 mmol) in 20 ml DMF. Add the *tert*-butylhydrazine-containing filtrate from step 20. Stir the reaction mixture at room temperature for 1 hr.
22. Evaporate the solvent under reduced pressure.

23. Suspend the residue in water and extract three times, each time with 50 ml CHCl₃, in a separatory funnel.
24. Combine the CHCl₃ extracts, dry over anhydrous MgSO₄, and evaporate the solvent to give a yellow solid.
25. Prepare an 80-in. (20.3-cm) × 1.5-in. (3.8-cm) column containing 100 g Silica Gel 60, saturated with CHCl₃ (the mobile phase).
26. Dissolve the yellow solid in a minimal volume of CHCl₃ and apply to the column. Elute with 1% methanol in CHCl₃. Sample the eluate as it emerges from the column and analyze by thin-layer chromatography (TLC) using 1% methanol in CHCl₃ as the mobile phase. Pool fractions containing a species with R_f of ~0.1.
27. Evaporate solvent to produce a white solid, 5-amino-1-tert-butyl-3-(1'-naphthyl)-4-cyano-1H-pyrazole.

Yield should be ~75%.

Synthesize 4-amino-1-tert-butyl-3-(1'-naphthyl)-1H-pyrazolo[3,4-d]pyrimidine (1NA-PP1)

28. Suspend 0.99 g 5-amino-1-tert-butyl-3-(1'-naphthyl)-4-cyano-1H-pyrazole (3.4 mmol) in 20 ml of formamide.
29. Heat this suspension at 180°C on an oil bath, stirring with a reflux condenser for 10 hr.
30. Cool the reaction mixture and dilute with 80 ml water.
31. Collect the precipitate by filtration on paper in a Buchner funnel.
32. Dissolve precipitate in 40 ml room temperature ethanol. Add 1 g powdered activated charcoal. Boil at 80°C for 10 min.
33. Pour a celite pad on a Buchner funnel using a slurry of celite in ethanol. Filter the hot mixture through this celite pad. Collect the filtrate.
34. Evaporate the solvent to obtain 1NA-PP1 as a white powder.

Expect yields of ~75%.

PREPARATION OF γ -³²P-LABELED N⁶(BENZYL)ATP

NDPK is a ubiquitous enzyme that catalyzes the exchange of γ phosphate between nucleoside triphosphate species and nucleoside diphosphate species by means of a bi-bi-ping-pong mechanism involving a phospho-enzyme intermediate (Weaver, 1962; Mourad and Parks, 1966). In this protocol, 6× His-tagged NDPK is immobilized in the solid phase by metal affinity chromatography, and the column-bound NDPK is equilibrated with [γ -³²P]ATP in low magnesium to generate the labeled phospho-enzyme. The column is then washed to remove residual ATP, and a ten-fold molar excess of N⁶(benzyl)ADP is added in the presence of larger amounts of magnesium. Eluate from this stage contains a mixture of [γ -³²P]N⁶(benzyl)ATP and N⁶(benzyl)ADP, and can be used directly in kinase reactions. This protocol can also be used to generate γ -³⁵S-labeled nucleoside triphosphates.

Materials

- 1:1 slurry of cobalt affinity resin (IDA-Co²⁺-Sepharose) in HBS/0.02% sodium azide (see Support Protocol 1)
- HEPES buffered saline (HBS): 150 mM NaCl/100 mM HEPES, pH 7.4
- Purified NDPK-6×His (see Support Protocol 2)
- Phosphate-buffered saline (PBS): 150 M NaCl/100 mM sodium phosphate pH 7.4 (see APPENDIX 2 for sodium phosphate buffer)
- [γ -³²P]ATP (~3000 Ci/mmol or ~7000 Ci/mmol)
- 1 mM N⁶(benzyl)ADP (Shah et al., 1997; Shah and Shokat, 2002)
- PBS (see above) containing 5 mM MgCl₂
- 2-mm glass beads (VWR)
- 1-ml disposable pipet tip
- Stand and clamp to accommodate 1-ml disposable pipet tip

NOTE: All steps are performed at room temperature.

Prepare cobalt affinity column

1. Place a 2-mm glass bead inside a 1-ml disposable pipet tip and flick the tip to make sure the bead is firmly lodged near the end. Use scissors to remove the end of the tip where it extends beyond the glass bead. Mount this miniature column in a clamp on a stand.
2. Add 200 μ l of 1:1 slurry of cobalt affinity resin in HBS/0.02% sodium azide (prepared as in Support Protocol 1) to the column prepared in step 1.
3. Add 1 ml HBS and allow the buffer to drain until level with the top of the beads.

If desired, a 1-ml pipet can be used to drive fluid flow through the beads. If this is done, be careful to avoid introducing a negative relative pressure in the column when removing the pipet, as this can disturb the beads.

Add NDPK-6×His and radiolabeled ATP to column

4. Add ~165 μ g purified NDPK-6×His.

If the NDPK is in a buffer other than the recommended HKG buffer (see Support Protocol 2), it must be freed of any chelators or reducing agents, including EGTA, before addition to the column.

5. When this has flowed into the column, follow with 1 ml PBS. If desired, retain the eluate to quantify protein retention.

PBS is used in place of HBS in all steps subsequent to addition of NDPK. IDA-Co²⁺ appears to have a weak affinity for the various nucleoside phosphate species; PBS competitively blocks these interactions. A slight lavender color may be noticed on the initial addition of PBS. This is insoluble cobaltous phosphate. It has not been found to interfere with any subsequent steps, nor has the use of PBS caused any significant loss of enzyme from the column.

NDPK-6×His has an A₂₈₀ extinction coefficient of 1.31 for a 1 mg/ml solution, as calculated using the Gill equation.

6. Dilute 833 pmol [γ -³²P]ATP (2.5 mCi at 3000 Ci/mmol, 5.8 mCi at 7000 Ci/mmol) with PBS to contain no greater than 1 mM reducing agent.

Commercial [γ -³²P]ATP may contain an appreciable quantity of DTT or other reducing agent. This should be diluted so that enzyme is not lost from the column due to reduction of cobalt.

The reactions can be scaled for use of smaller quantities of label.

7. Add the radioisotope to the column and allow it to flow through, driving the flow if necessary.

A moderate darkening of the column due to a low level of metal reduction may be observed, and a small amount of reduced metal may be present in the eluate. This should not normally be cause for concern. Small samples of the load and eluate may be taken to quantify label retention on the column by liquid scintillation counting.

8. Wash the column at least twice, each time with 1 ml PBS.

The eluate should be sampled to quantify label release in these steps, to ensure that residual [γ - ^{32}P]ATP has been cleared from the column.

Add N^6 (benzyl)ADP to column and elute N^6 (benzyl)ATP

9. Add 8 μl of 1 mM N^6 (benzyl)ADP (8 nmol) to 32 μl PBS containing 5 mM MgCl_2 .

This ADP analog and others can be synthesized using published procedures (Shah et al., 1997; Shah and Shokat, 2002).

10. Carefully add this to the column, and then gently force it into the beads until the liquid is again level with the top of the beads.

11. Add 250 μl PBS containing 5 mM MgCl_2 and gently force it through the beads until level, collecting the eluate.

A small sample of the eluate should be taken for liquid scintillation counting to determine yield.

SUPPORT PROTOCOL 1

PREPARATION OF COBALT AFFINITY RESIN (IDA- Co^{2+} -SEPHAROSE)

Materials

Iminodiacetic acid (IDA)–Sepharose slurry (Sigma)
200 mM cobalt chloride
HEPES-buffered saline (HBS): 150 mM NaCl/100 mM HEPES, pH 7.4
HBS (see above) containing 0.02% sodium azide
15-ml conical tubes

1. Place ~3 ml of IDA-Sepharose bead slurry in a 15-ml conical tube.
2. Wash several times with water, each time by filling the 15-ml conical tube and then decanting.
3. Wash once with 200 mM cobalt chloride.
The binding is essentially instantaneous.
4. Wash four times with HBS.
5. Store beads at 4°C in a bead volume equivalent of HBS containing 0.02% sodium azide.

SUPPORT PROTOCOL 2

EXPRESSION AND PURIFICATION OF NDPK

A 6 \times His–tagged NDPK from *S. cerevisiae* is expressed in BL21(DE3) *E. coli* and purified by metal-affinity chromatography.

Materials

pJDB1 expression plasmid: available from the Shokat Laboratory
(justinb@itsa.ucsf.edu) or Weiss Laboratory (elweiss@northwestern.edu)
E. coli strain BL21(DE3) (Novagen)
IPTG
Cobalt affinity resin (IDA- Co^{2+} -Sepharose; see Support Protocol 1)
HIK200 buffer (see recipe), 4°C

HEK10 buffer (see recipe), 4°C

HKG buffer (see recipe), 4°C

Liquid nitrogen

30°C shaking incubator

5-ml chromatography column

MWCO 15,000 dialysis membrane

Additional reagents and equipment for transformation of *E. coli* (UNIT 1.8), growing bacterial cultures (UNIT 1.2), preparation of bacterial lysates (UNIT 1.7), dialysis (APPENDIX 3C), and determination of protein concentration (UNIT 10.1A)

Grow transformed bacteria

1. Transform (UNIT 1.8) the pJDB1 expression plasmid into *E. coli* strain BL21(DE3).

This plasmid is a derivative of pET19b containing the S. cerevisiae gene YNK1 with a C-terminal 6×-His tag. This plasmid confers resistance to kanamycin. Cells carrying this plasmid should be cultured in medium containing 30 μg/ml kanamycin for the transformation procedure (see UNIT 1.2).

2. Grow transformed strain to mid-log phase ($OD_{600} = 0.5$ to 1.0) at 37°C (UNIT 1.2).
3. Add IPTG to the culture for a final concentration of 0.4 mM.
4. Incubate with shaking at 37°C for 3 hr.

Prepare bacterial lysate

The following steps are performed at 4°C, using chilled buffers.

5. Isolate *E. coli* cells by centrifugation (UNIT 1.7).
6. Produce clarified lysate by method of choice (UNIT 1.7), ensuring that the salt concentration is not less than 150 mM in the final extract.
7. Prepare a 5-ml column of IDA-Co²⁺-Sephrose.

A 5-ml column is appropriate for a 1-liter-scale expression.

Purify lysate by cobalt affinity chromatography and dialysis

8. Apply the lysate (from step 7) to the column.
9. Wash with 4 column volumes HIK200 buffer.
10. Elute with 2.5 column volumes HEK10 buffer.
11. Dialyze (APPENDIX 3C) against three changes of HKG buffer, each time for 2 hr, using a MWCO 15,000 dialysis membrane.
12. Determine protein concentration by measuring A_{280} (UNIT 10.1A), divide into aliquots, and freeze using liquid nitrogen. Store at -80°C.

NDPK-6×His has an A_{280} extinction coefficient of 1.31 for a 1 mg/ml solution, as calculated using the Gill equation (also see UNIT 10.1A).

The purified protein has a tendency to precipitate in buffers containing <250 mM salt. Highly concentrated solutions may exhibit some precipitation at even higher salt concentrations. If precipitation is observed during dialysis, the precipitated protein may be resolubilized by slow addition of 2.5 M KCl. If necessary; a final KCl concentration as high as 800 mM is acceptable.

Storage buffer should be free of DTT or other reducing agents, as these will complicate subsequent use of the enzyme.

IN VIVO INHIBITION OF ANALOG-SENSITIVE KINASES IN YEAST

If the activity of a protein kinase of interest is required for growth under some or all conditions, halo assays can be used to evaluate inhibition of *-as* alleles of this kinase. These simple experiments involve placement of a small disc of filter paper saturated with a compound of interest onto a solid-medium plate that has been uniformly spread with yeast cells. If the compound inhibits growth of the cells, a clear zone or “halo” will be evident around the filter disc. The size of this halo is a function of the sensitivity of the cells to the compound. These assays are a convenient way to evaluate sensitivity to a number of different compounds. Alternatively, a traditional liquid culture and inhibitor assay may be performed. Both methods are included here.

Materials

Inhibitors: INM-PP1 (see Basic Protocol 1) and INA-PP1 (see Basic Protocol 2)

DMSO

Yeast cells of interest containing wild-type (control) and analog-sensitive kinase alleles

Appropriate liquid and solid yeast media (*UNIT 13.1*)

5-mm sterile filter discs

Additional reagents and equipment for growing yeast cells (*UNIT 13.2*)

For the halo assay

- 1a. Make 10 mM, 1 mM, and 0.5 mM working solutions of inhibitors (1NM-PP1 and 1NA-PP1) in DMSO.
- 2a. Grow yeast cells of interest (containing *-as* analog-sensitive kinase gene; also include wild-type control) to log phase in liquid medium (*UNIT 13.2*).
- 3a. Dilute cultures to $OD_{600} = 0.2$. Place 2 to 3 ml of a diluted culture onto a solid medium plate (*UNITS 13.1 & 13.2*) such that the entire surface of the plate is covered with excess liquid medium. Allow cells to settle for 2 to 5 min.
- 4a. Tilt the plate so excess liquid pools along one edge. Remove most of this liquid with a pipet. Rock the plate a few times to ensure that remaining liquid is evenly distributed. Allow the plate to dry.
- 5a. Place 5-mm sterile filter discs onto the plate.
Four or five evenly spaced discs can generally be placed on a single plate.
- 6a. Gently pipet 10 μ l of inhibitor solution onto the middle of a filter disc. Include a DMSO-only control. Allow cells to grow and measure diameter of growth inhibition zones.

For the liquid culture assay

- 1b. Make a stock solution of the inhibitor of interest (1NM-PP1 or 1NA-PP1) dissolved in DMSO, containing between 10 mM and 25 mM inhibitor.

The solution can be stored at room temperature in a dark box.

- 2b. Grow yeast cells of interest (containing *-as* analog-sensitive kinase gene; also include wild-type control) to mid-log phase in liquid medium (*UNIT 13.2*).

Remember to grow enough cells to perform inhibitor treatment and a DMSO-only control. If planning to use conditioned medium, grow enough excess culture to provide medium for subsequent incubations with and without inhibitor.

Fresh and conditioned media generally differ in pH and nutrient concentration. The use of conditioned medium avoids effects caused by adaptation to new medium.

- 3b. Add inhibitor stock solution to the liquid medium. When assaying a range of inhibitor concentrations, ensure that the final concentration of DMSO is the same in all incubations. If fresh nonconditioned medium is used, add the inhibitor stock to fresh medium while actively mixing, either by swirling in a flask or stirring with a magnetic bar. Add the same amount of DMSO to a second portion of fresh medium for a no-inhibitor control. Harvest the cells from the culture grown in step 2 by filtration or centrifugation, and resuspend them in the inhibitor-containing and control medium, respectively. If conditioned medium is used, remove cells from an appropriate volume of culture by filtration or centrifugation. Collect the cleared conditioned medium (filtrate or supernatant), add inhibitor to one portion and DMSO for control to a second portion, then resuspend the cells back into the inhibitor- (or DMSO)-containing conditioned medium.

Direct addition of concentrated inhibitor to growing cultures is not recommended.

- 4b. Incubate and perform phenotypic analysis as desired.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

HEK10 buffer

20 mM HEPES, pH 7.4
10 mM EDTA
250 mM KCl
Store indefinitely at room temperature

HIK200 buffer

20 mM HEPES, pH 7.4
200 mM imidazole
50 mM KCl
Store indefinitely at 4°C

HKG buffer

20 mM HEPES, pH 7.4
250 mM KCl
20% (v/v) glycerol
Store indefinitely at room temperature

COMMENTARY

Background Information

Understanding the function of any protein kinase requires identification of its direct phosphorylation targets and evaluation of effects caused by its inactivation. This kind of characterization is among the most challenging in contemporary molecular biology. The study of protein kinase function has been greatly advanced through the use of cell-permeable compounds that inhibit the enzymatic activity of these proteins. Protein kinase inhibitors with relatively high specificity have been identified in screens of chemical compound libraries (Gray et al., 1998). Such compounds have already proven useful in treat-

ment of cancers caused by deranged protein kinase signaling (Druker et al., 2001). However, kinase inhibitors generally block the activity of a group of closely related enzymes. This lack of perfect specificity can prove problematic if closely related kinases control different processes in the cell type being studied. Acquisition of specific kinase inhibitors through exhaustive screening of compound libraries is furthermore a laborious process, and is beyond the capabilities of most molecular biology laboratories.

In the absence of perfectly specific protein kinase inhibitors, it is necessary to use alternative means to inactivate a kinase of interest.

Despite their power, each of these approaches has drawbacks. Depletion techniques (such as RNAi) do not address the function of the kinase's enzymatic activity. Furthermore, the kinetics of depletion can be very slow, which may allow homeostatic mechanisms to compensate for loss of enzyme function through activation of alternate pathways. Endogenous kinase genes can be replaced with catalytically inactive alleles in organisms that are easy to manipulate genetically. This is not straightforward with essential kinases, and is also subject to concerns about phenotypes that arise through adaptation. It is possible to generate conditional alleles that are rapidly inactivated under restrictive conditions (most often elevated temperature). These alleles are usually generated through random mutagenesis of the kinase of interest, and the biochemical reason for their effect on activity is generally obscure. More importantly, the phenotype of any conditional allele must be weighed against the effects of the restrictive conditions on wild-type cells. Temperature shifts can cause major perturbation of cell physiology, significantly complicating interpretation of conditional mutant phenotypes.

Given limitations in the above approaches, chemical and genetic approaches have been combined to allow highly specific inhibition of many protein kinases. The ATP-binding site of many protein kinases can be mutated to create kinase variants that are inhibited by cell-permeable compounds that do not appreciably affect the activity of wild-type protein kinases (Bishop et al., 2000; Weiss et al., 2000; Carroll et al., 2001; Sekiya-Kawasaki et al., 2003; Wang et al., 2003; reviewed in Bishop et al., 2001). Allele-specific inhibition of these kinases *in vivo* can be used to determine an individual protein kinase's physiological function without the confounding effects of temperature shift and adaptation. These "analog-sensitive" or "*-as*" mutant enzymes can often use derivatized forms of ATP that are poorly used by wild-type kinases (Shah et al., 1997; Bishop et al., 1998; Polson et al., 2001; Shah and Shokat, 2002). Specific use of ATP analogs by *-as* kinases *in vitro* can be exploited to directly identify kinase phosphorylation targets in complex protein mixtures (Shah and Shokat, 2002; Ubersax et al., 2003). Substrate specificity of *-as* kinase alleles does not appear to differ from corresponding wild-type alleles (Witucki et al., 2002). This is consistent with the physical separation between the mutated region of the

ATP-binding pocket and the substrate-recognition domain.

Design of analog-sensitive protein kinase alleles

The primary sequences and three-dimensional structures of protein kinase ATP-binding sites are conserved (Hanks and Hunter, 1995), and it is therefore straightforward to design candidate analog-sensitive kinase alleles. Analog-sensitive protein kinase alleles are constructed by substitution of glycine or alanine for an amino acid in kinase subdomain V, which is typically bulky and hydrophobic. This "gatekeeper" residue forms part of a hydrophobic pocket that interacts with the adenosine moiety of ATP; many competitive inhibitors of protein kinases bind in this site. Replacing the large side chain at this position with a smaller one alters the shape of the pocket, allowing molecules that do not interact with wild-type kinases to bind. Glycine substitutions at this position are referred to as "*-as1*" alleles and alanine substitutions as "*-as2*" alleles. The rationale behind this approach is supported by analysis of crystal structures of Hck-PP1 complex and the c-Src(*-as1*)-*N*⁶(benzyl)ATP (Schindler et al., 1999; Witucki et al., 2002). The nature of these alleles is illustrated in Figure 18.11.4, which schematizes some of the kinase-inhibitor contacts that are important for association of PP1 with the ATP binding site of Hck, as well as the engineered extra space present in Hck1-*as1* that allows binding of 1NA-PP1.

Figure 18.11.5 shows an alignment of amino acids in subdomains V and VII that are relevant to the construction of analog-sensitive kinases. The gatekeeper amino acid, which is mutated to alanine or glycine to create the *-as1* and *-as2* alleles, is shaded in gray. The amino acid that must be changed to construct *-as* kinase alleles can be most easily identified using a freely available online resource at <http://kinase.ucsf.edu/ksd/>. This Web site includes thorough instructions for use of the database. It is advisable to use the GenBank GI number of the protein kinase of interest to search for the appropriate alignment using the "Search by: family name" option. Alignments are generated as Microsoft Excel format files, for download to a computer, or in HTML format for viewing using a Web browser. In these alignments, ATP-contacting amino acids are highlighted green, and the gatekeeper residue is highlighted red.

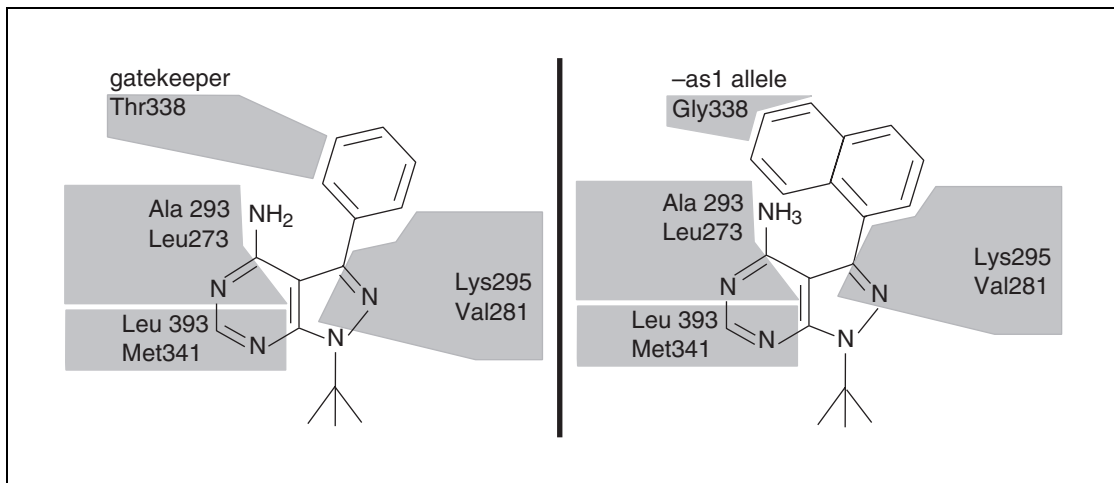


Figure 18.11.4 Inhibitor binding. Important contacts between PP1 and the ATP binding site of the Src-family protein kinase Hck (left) and the space created in the *-as1* allele that allows 1NA-PP1 binding (right). The right panel is inferred from the structure of c-Src(*-as1*).

	Subdomain V Gatekeeper	Subdomain VII -as3 position
v-Src [332]	EPIYIVTLEYMSKGSLLDFLKGEMG---//---	LVCKVADFGLARLIEDNEY
Cdk2 [074]	NKLYLVFEFLHQ-DLKKFMDASAL---//---	GAIKLADFGLARAFGVVPR
Cdc28 [082]	HKLYLVFEFLDL-DLKRYMEGIPK---//---	GNLKLGDFGLARAFGVPLR
Fus3 [093]	NEVYIIQELMQT-DLHRVISTQML---//---	CDLKVCDFGLARIIDESAA
Cla4 [634]	DDLWVVMFMEGGSLTDI IENSPT---//---	ARVKITDFGF CARLTDKRS
JNK [102]	QDVYLVMEELMDA-NLCQVIHMELD---//---	CTLKILDFGLARTACTNFM
PKA α [115]	SNLYMVM EYVAGGEMF SHLRRIGR---//---	GYIQVTDGFGAKRVKGRTW

Figure 18.11.5 Subdomain V and VII alignments. Analog-sensitive alleles of these protein kinases, whose wild-type sequences are shown, have been engineered. The gatekeeper residue in each kinase and the site of the second mutation (immediately amino terminal to the DFG motif) in Cla4 and JNK are shaded in gray.

Troubleshooting

Specific activity of *-as* kinases

The ideal analog-sensitive kinase allele is functionally silent (i.e., still recognizes ATP, is not a hypermorph or hypomorph, is localized to the same cellular compartments as the wild-type kinase, etc.) and yet has features not found in any wild-type protein kinases: it preferentially uses unnatural ATP analogs, and is selectively inhibited by low concentrations of compounds that do not affect wild-type kinases. In practice, *-as* kinase alleles may require relatively high concentrations of inhibitor to block activity, and they may exhibit substantially reduced catalytic activity. It is advisable to construct both *-as1* and *-as2* alleles: reduction of enzyme activity is generally more severe in *-as1* alleles than in *-as2* alleles, while *-as2* alleles are sometimes less

sensitive to inhibitor than *-as1* alleles. Sensitivity to inhibitor can be enhanced in some cases by an additional mutation outside the adenosine-binding pocket: kinases mutable to optimal analog sensitivity generally have the sequence (L/I)ADFG in subdomain VII. In Figure 18.11.4, this residue is shown shaded in gray. If in the sequence XDFG the amino acid at position X is anything other than glycine or alanine, introduction of alanine at this position may greatly enhance sensitivity to inhibitor. This amino acid was identified as an important determinant of PP1 binding to Hck (Schindler et al., 1999). Changing the sequence TDFG in the yeast PAK Cla4p ADFG greatly enhances sensitivity of the *-as2* allele (Weiss et al., 2000); such double-mutant kinases are called *-as3* alleles. The *prk1-as3* allele exhibits similarly enhanced sensitivity to inhibitor (Sekiya-Kawasaki et al., 2003).

While *-as* alleles of some protein kinases exhibit minor reduction in enzymatic activity, *-as1* and *-as2* alleles of some kinases have greatly reduced catalytic activity. The structural features that determine tolerance for the *-as* alleles are presently poorly understood. However, alignment of the ATP-binding sites of *-as* kinases that work well with those that do not may suggest compensatory mutations that ameliorate loss of activity in *-as1* and *-as2* alleles. MEKK1-*as1* (I1288G), for example, exhibits low kinase activity using both ATP and *N*⁶-phenethyl-ATP as substrates. MEKK1 contains cysteine at position 1238 on the β 2 strand, which should be in direct contact with the oxygen atom in the ribose ring of ATP; the vast majority of kinases contain Val at corresponding positions. A Cys1238 to Val mutation introduced into MEKK1-*as1* restored the kinase activity of MEKK1-*as1* (using ATP) to wild-type level, and this double-mutant enzyme was also able to use *N*⁶-(phenethyl)-ATP more efficiently than MEKK1-*as1* (Dennis Templeton, pers. comm.). If one constructs a *-as* allele that is severely compromised, it is advisable to change highly divergent residues in the ATP binding site to the consensus amino acid.

Heterologous expression systems for protein kinases (e.g., bacterial or insect cell expression) can produce active wild-type forms of mammalian kinases. However, the authors have found that these heterologous expression systems often do not provide active forms of *-as1* or *-as2* alleles of protein kinases. The reasons for this are particular to each protein kinase, but some common problems have been encountered. Many protein kinases require phosphorylation of an activation site (often in the activation segment). These residues may or may not be phosphorylated in a *trans*-auto-phosphorylation-dependent manner, or may be phosphorylated by an "upstream" kinase. In heterologous expression systems, such upstream kinases are absent and the function of activation-site phosphorylation must be carried out by the kinase itself in a nonoptimal reaction. Since the *-as* alleles of a given kinase may already be slightly less active with ATP as a phosphodonor than wild-type enzyme, this non-natural phosphorylation reaction may proceed extremely inefficiently, resulting in production of an inactive *-as* kinase allele. This can be corrected by incubation of the *-as* kinase with 1 mM ATP for 1 to 3 hr prior to assay, to greatly favor any phosphorylation activity to "autoactivate" the kinase.

Naturally, if the normal upstream kinase is known, it is preferable to activate the *-as* kinase by direct incubation with this enzyme. Interestingly, a number of *-as* kinases that cannot be activated by these means and never shown sufficient activity via heterologous expression systems are fully able to complement the function when expressed in the homologous cell where it normally carries out its function. Therefore, heterologous expression results with *-as* alleles should be interpreted with caution.

Phosphorylation target identification by allele-specific ATP analog use

While ATP analogs like *N*⁶(benzyl)ATP are poorly used by most wild-type kinases, they are apparently accepted moderately well by some (Kraybill et al., 2002). This can result in background labeling that is independent of *-as* kinase addition. It is therefore important to characterize the level of background labeling that occurs when γ -³²P labeled ATP analogs are added to the protein mixture of interest; in some cases it will be necessary to reduce the level of background. The nature and magnitude of endogenous kinase activity in cell extracts and fractions varies significantly depending on the extract source, preparation, and fractionation strategy. Therefore, optimal strategies for background reduction are likely to be case-specific. However, the inclusion of significant concentrations of unlabeled ATP is often helpful. Since *-as* kinases can have higher *K_m*'s for ATP than wild-type kinases, ATP will often more effectively compete for ATP analog binding to wild-type kinases than to the *-as* kinase of interest. Consequently, background labeling will be reduced to a greater extent than *-as* kinase-dependent signal.

ATP is rapidly depleted in many extracts. In most cases, it is preferable to maintain a constant ATP concentration through the use of an ATP regenerating system. Addition of ATP, a molar excess of phosphocreatine, and creatine kinase will serve this purpose: the excess phosphocreatine acts as a substrate for creatine kinase in the conversion of ADP to ATP. If an ATP regenerating system is to be used in *-as* kinase substrate labeling experiments, it is advisable to first determine the ATP concentration that gives the greatest inhibition of background labeling without significantly compromising *-as* kinase-dependent signal. To this end, one should set up a series of extract labeling reactions in which

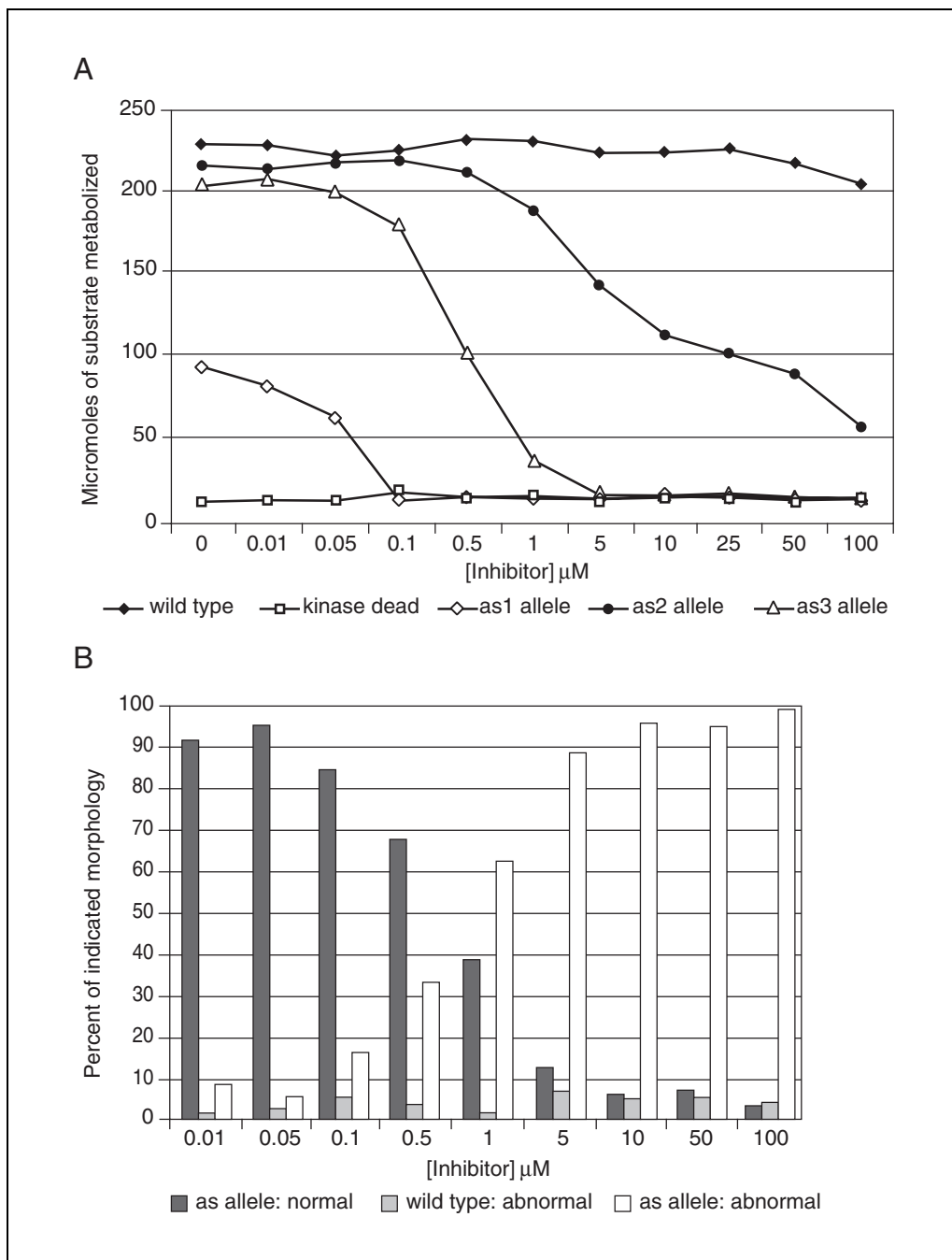


Figure 18.11.6 Hypothetical data from $-as$ kinase inhibition experiments. **(A)** Inhibition curves for a kinase-dependent metabolic process in which the amount of metabolite consumed in a fixed time can be measured. **(B)** Simple morphometric analysis of the inhibition of a kinase required for normal cell morphogenesis. Percent of normal wild-type cells is omitted for clarity.

the ATP concentration is varied from 100 μM to 2 mM.

It may not be possible to detect specifically labeled proteins in complex mixtures directly if the $-as$ kinase of interest has low specific activity, the extract exhibits high background labeling, or the phosphorylation targets of that $-as$ kinase are very rare. If a protein or complex of interest can be affinity purified, how-

ever, specific labeling by $-as$ kinases may be easier to demonstrate. In this approach, labeling reactions are performed using $[\gamma\text{-}^{32}\text{P}]\text{N}^6(\text{benzyl})\text{ATP}$ (or other radiolabeled analog) in a complex protein mixture with and without $-as$ kinase. The protein or complex of interest is then isolated from the two reactions and incorporation of label compared (Ubersax et al., 2003).

Anticipated Results

Analog-sensitive kinase alleles should ultimately be useful for *in vivo* analysis of kinase function in a wide range of model systems. There is, therefore, no general method for their use, but rather a set of principles for good experimental design. Ideally, the endogenous kinase of interest should be replaced by the *-as* allele. Once this is accomplished, two major questions must be answered in any *-as* kinase experiment: (1) how well does the mutant kinase supply function in the absence of inhibitor; and (2) how completely is the *-as* kinase activity of interest blocked by addition of allele-specific inhibitor? These questions can be answered through quantitative evaluation of phenotypic responses to a wide range of inhibitor concentrations.

Figure 18.11.6 shows hypothetical examples of dose-response experiments that evaluate *-as* kinase function and inhibition. Ideally, the *in vivo* function of the *-as* kinase in question should be significantly inhibited at PP1 derivative concentrations $<10 \mu\text{M}$. When a kinase is required for a process or enzyme activity with a measurable steady-state rate, the progress of this process or activity over a fixed time should be measured in the presence of increasing inhibitor concentrations (Fig. 18.11.6A). If the kinase is required for cell morphogenesis, organization of intracellular structures, or cell-cycle progress, the percentage of cells with abnormal (or cell-cycle-arrested) morphology should be scored (Fig. 18.11.6B). Naturally, inhibitor treatment of wild-type cells should be used as a control. When possible, the effects of *-as* kinase inhibition should also be measured relative to the phenotype of a catalytically inactive kinase allele. With extended PP1 analog treatment, the phenotype of a fully inhibited *-as* allele should be identical to a catalytically inactive kinase allele.

Figure 18.11.6A shows hypothetical inhibition curves for wild-type, catalytically inactive, *-as1*, *-as2*, and *-as3* alleles of a kinase required for a cell's consumption of a measurable compound. The measurements involved in such assays should be performed on cells that have been pretreated with inhibitor, allowing time for the treatment to have maximal effect. While the *-as1* allele in this experiment is sensitive to low concentrations of inhibitor, it is also severely compromised relative to wild-type and is therefore of limited utility. The *-as2* allele is less compromised, but is not sufficiently sensitive to inhibitor.

The *-as3* allele is sufficiently sensitive while retaining adequate *in vivo* activity in the absence of inhibitor. The phenotype of this allele "saturates" at $\sim 5 \mu\text{M}$ inhibitor. It may be advisable to use slightly higher concentrations in future assays to ensure inactivation. Figure 18.11.6B shows morphological phenotypes over a range of inhibitor concentrations. A number of researchers have published examples of dose-response analysis with *-as* kinases (Bishop et al., 2000; Weiss et al., 2000; Weiss et al., 2002; Sekiya-Kawasaki et al., 2003).

Thus far, most *in vivo* inhibition studies using *-as* kinases have been performed in the budding yeast *Saccharomyces cerevisiae*. The considerable pace of yeast experiments is in part due to the ease of gene replacement in this organism; *-as* kinase inhibition is also clearly useful in metazoans (Wang et al., 2003). The authors have included guidelines for *-as* kinase inhibition in yeast. The specific assay conditions will, of course, depend on the nature of the process of interest.

Literature Cited

- Bishop, A.C., Shah, K., Liu, Y., Witucki, L., Kung, C., and Shokat, K.M. 1998. Design of allele-specific inhibitors to probe protein kinase signaling. *Curr. Biol.* 8:257-266.
- Bishop, A., Kung, C., Shah, K., Witucki, L., Shokat, K.M., and Liu, Y. 1999. Generation of monospecific nanomolar tyrosine kinase inhibitors via a chemical genetic approach. *J. Am. Chem. Soc.* 121:627-631.
- Bishop, A.C., Ubersax, J.A., Petsch, D.T., Matheos, D.P., Gray, N.S., Blethrow, J., Shimizu, E., Tsien, J.Z., Schultz, P.G., Rose, M.D., Wood, J.L., Morgan, D.O., and Shokat, K.M. 2000. A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* 407:395-401.
- Bishop, A.C., Buzko, O., and Shokat, K.M. 2001. Magic bullets for protein kinases. *Trends Cell Biol.* 11:167-72.
- Carroll, A.S., Bishop, A.C., DeRisi, J.L., Shokat, K.M., and O'Shea, E.K. 2001. Chemical inhibition of the Pho85 cyclin-dependent kinase reveals a role in the environmental stress response. *Proc. Nat. Acad. Sci. U.S.A.* 98:12578-12583.
- Druker, B.J., Sawyers, C.L., Capdeville, R., Ford, J.M., Baccarani, M., and Goldman, J.M. 2001. Chronic myelogenous leukemia. *Hematology (Am. Soc. Hematol. Educ. Program)* pp. 87-112.
- Gray, N.S., Wodicka, L., Thunnissen, A.M., Norman, T.C., Kwon, S., Espinoza, F.H., Morgan, D.O., Barnes, G., LeClerc, S., Meijer, L., Kim, S.H., Lockhart, D.J., and Schultz, P.G. 1998. Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science* 281:533-538.

- Hanks, S.K. and Hunter, T. 1995. Protein kinases 6. The eukaryotic protein kinase superfamily: Kinase (catalytic) domain structure and classification. *FASEB J.* 9:576-596.
- Kraybill, B.C., Elkin, L.L., Blethrow, J.D., Morgan, D.O., and Shokat, K.M. 2002. Inhibitor scaffolds as new allele specific kinase substrates. *J. Am. Chem. Soc.* 124:12118-12128.
- Mourad, N. and Parks, R.E. Jr. 1966. Erythrocytic nucleoside diphosphokinase. II. Isolation and kinetics. *J. Biol. Chem.* 241:271-278.
- Polson, A.G., Huang, L., Lukac, D.M., Blethrow, J.D., Morgan, D.O., Burlingame, A.L., and Ganem, D. 2001. Kaposi's sarcoma-associated herpesvirus K-bZIP protein is phosphorylated by cyclin-dependent kinases. *J. Virol.* 75:3175-3184.
- Schindler, T., Sicheri, F., Pico, A., Gazit, A., Levitzki, A., and Kuriyan, J. 1999. Crystal structure of Hck in complex with a Src family-selective tyrosine kinase inhibitor. *Mol. Cell* 3:639-648.
- Sekiya-Kawasaki, M., Groen, A.C., Cope, M.J., Kaksonen, M., Watson, H.A., Zhang, C., Shokat, K.M., Wendland, B., McDonald, K.L., McCaffery, J.M., and Drubin, D.G. 2003. Dynamic phosphoregulation of the cortical actin cytoskeleton and endocytic machinery revealed by real-time chemical genetic analysis. *J. Cell Biol.* 162:765-772.
- Shah, K. and Shokat, K.M. 2002. A chemical genetic screen for direct v-Src substrates reveals ordered assembly of a retrograde signaling pathway. *Chem. Biol.* 9:35-47.
- Shah, K., Liu, Y., Deirmengian, C., and Shokat, K.M. 1997. Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proc. Natl. Acad. Sci. U.S.A.* 94:3565-3570.
- Ubersax, J.A., Woodbury, E.L., Quang P.N., Paraz, M., Blethrow, J.D., Shah, K., Shokat, K.M., and Morgan, D.O. 2003. Targets of the cyclin-dependent kinase Cdk1. *Nature* 425:859-864.
- Wang, H., Shimizu, E., Tang, Y.P., Cho, M., Kyin, M., Zuo, W., Robinson, D.A., Alaimo, P.J., Zhang, C., Morimoto, H., Zhuo, M., Feng, R., Shokat, K.M., and Tsien, J.Z. 2003. Inducible protein knockout reveals temporal requirement of CaMKII reactivation for memory consolidation in the brain. *Proc. Natl. Acad. Sci. U.S.A.* 100:4287-4292.
- Weaver, R.H. 1962. Nucleoside diphosphokinases. *Enzymes* 6:151-160.
- Weiss, E.L., Bishop, A.C., Shokat, K.M., and Drubin, D.G. 2000. Chemical genetic analysis of the budding-yeast p21-activated kinase Cla4p. *Nat. Cell Biol.* 2:677-685.
- Weiss, E.L., Kurischko, C., Zhang, C., Shokat, K., Drubin, D.G., and Luca, F.C. 2002. The *Saccharomyces cerevisiae* Mob2p-Cbk1p kinase complex promotes polarized growth and acts with the mitotic exit network to facilitate daughter cell-specific localization of Ace2p transcription factor. *J. Cell Biol.* 158:885-900.
- Witucki, L.A., Huang, X., Shah, K., Liu, Y., Kyin, S., Eck, M.J., and Shokat, K.M. 2002. Mutant tyrosine kinases with unnatural nucleotide specificity retain the structure and phospho-acceptor specificity of the wild-type enzyme. *Chem. Biol.* 9:25-33.

Contributed by Justin Blethrow, Chao Zhang, and Kevan M. Shokat
University of California at San Francisco
San Francisco, California

Eric L. Weiss
Northwestern University
Evanston, Illinois

Detection of MAP Kinase Signaling

UNIT 18.12

Transmission of extracellular signals to their intracellular targets is mediated by a network of interacting proteins that relay biochemical messages, thereby controlling many cellular processes. Several related intracellular signaling pathways, collectively known as mitogen-activated protein kinase (MAPK) signaling cascades, have been elucidated in the past decade (Seger and Krebs, 1995). Transmission of signals via these cascades is usually initiated by activation of a small G protein (e.g., Ras) and followed by sequential stimulation of several sets of cytosolic protein kinases. Four distinct MAPK cascades, ERK (extracellular signal-related protein kinase), JNK (c-Jun NH₂-terminal kinase), p38 MAPK, and BMK (big MAPK), also known as ERK5, have been elucidated to date (Fig. 18.12.1). Each is named after the subgroup of its MAPK components and is composed of up to five levels: MAP4K (MAPK kinase kinase), MAP3K (MAPK kinase kinase), MAPKK (MAPK kinase), MAPK, and MAPKAPK (MAPK-activated protein kinase) (Fig. 18.12.1). One or more components in each of these levels phosphorylates and activates components in the next level, until a downstream component phosphorylates a target regulatory molecule. These cascades can cooperate in transmitting signals from most extracellular stimuli and can thus determine a cell's fate in response to the ever-changing environment. For a detailed description of the MAPK cascades, see Background Information.

Since the majority of MAPK cascade components are kinases, the methods used to detect MAPK cascade activation involve determination of protein kinase phosphorylations and activities. Anti-phospho-MAPK antibodies are used as a fast and convenient tool to monitor the activation of MAPKs (see Basic Protocol 1). The actual activity of these kinases is often detected by immunoprecipitation followed by phosphorylation of specific substrates. The activity assay is performed while the enzyme is still bound to the beads for immunoprecipitation, and the amount of incorporated phosphate is monitored by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and autoradiography (see Basic Protocol 1; Fig. 18.12.2). An alternative method involves affinity purification with a specific substrate of the examined kinase, followed by phosphorylation; this assay is described for JNK (see Basic Protocol 3; Fig. 18.12.2). When the identity of the kinase is not known, or there are no reagents available for its determination, an in-gel kinase assay should be used (see Basic Protocol 2; Fig. 18.12.2). Support protocols detail the preparation of EGF-stimulated Rat1 cells (Support Protocol 1) and GST-Jun-glutathione beads (Support Protocol 2), as needed for these assays.

CAUTION: Investigators should wear gloves for all procedures involving radioactivity and should be careful not to contaminate themselves and their clothing. When working with ³²P, investigators should frequently check themselves and the working area for radioactivity using a hand-held monitor. Any radioactive contamination should be cleaned up using appropriate procedures. Radioactive waste should be placed in appropriately designated areas for disposal. Follow the guidelines provided by your local radiation safety adviser (also see APPENDIX 1F).

DETERMINATION OF MAP KINASE (ERK) ACTIVATION USING ANTIBODIES

This protocol describes determination of ERK activity using either anti-phospho-MAPK antibodies or by immunoprecipitating the enzyme with specific antibodies and then performing a phosphorylation reaction in vitro. ERK is used here as an example; with appropriate reagents, this method can be performed with most MAPK isoforms and other components of the MAPK cascade. This protocol results in fast and efficient detection of

**BASIC
PROTOCOL 1**

**Analysis of
Protein
Phosphorylation**

18.12.1

Contributed by Hadara Rubinfeld and Rony Seger

Current Protocols in Molecular Biology (2004) 18.12.1-18.12.23

Copyright © 2004 by John Wiley & Sons, Inc.

Supplement 67

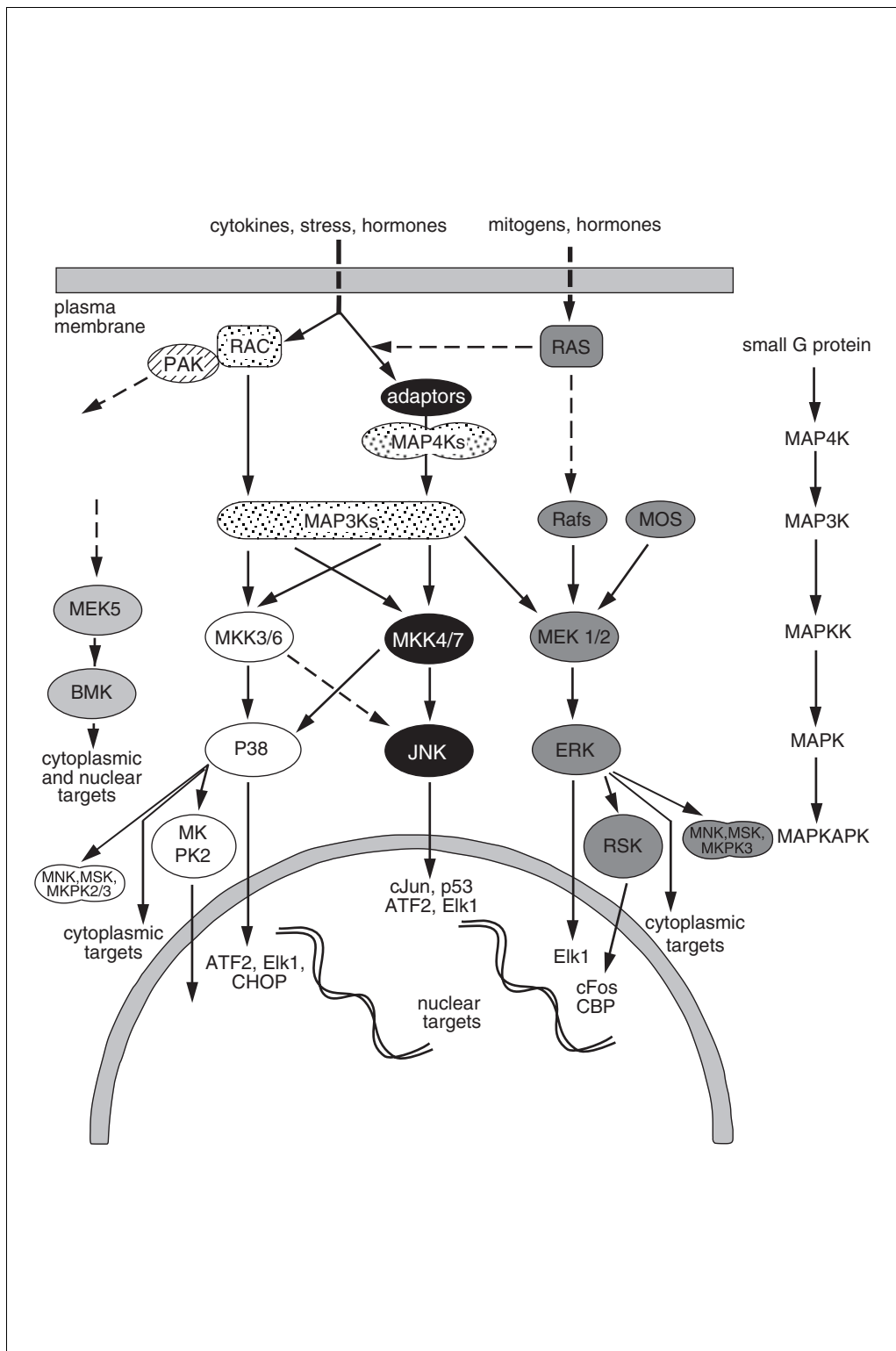


Figure 18.12.1 Schematic representation of mitogen-activated protein kinase (MAPK) cascades. The extracellular signal-related protein kinase (ERK) cascade is represented by light gray shading, the c-Jun NH₂-terminal kinase (JNK) cascade by dark gray shading, the stress-related protein kinase p38 MAPK cascade by white, and the big MAPK (BMK) cascade by stippled shading. Components that are shared by more than one cascade are indicated by combinations of shading. The connections between components from different levels are shown by arrows; the specifics of these interactions have yet to be defined. Abbreviations: MAP4K, MAPK kinase kinase; MAP3K, MAPK kinase kinase; MAPKK, MAPK kinase; MAPKAPK, MAPK-activated protein kinase.

the activity of desired protein kinases and reliable quantification of phosphorylation. Since the solid support (immunoprecipitating beads) may influence the kinase activity, however, the results do not always accurately reflect the specific activity of the examined kinase. As an alternative, it is possible to elute the kinases of interest from the immunoprecipitating beads (or isolate them by other means) and then determine their activity in solution (Seger et al., 1994). These methods can be used with almost all tissue culture cell lines, homogenized animal organs, and even whole lower organisms.

Materials

Six 6-cm tissue culture plates of EGF-stimulated Rat1 cells (Support Protocol 1, steps 1 to 4)
Homogenization buffer (see recipe), ice cold
Kinase buffer (see recipe), ice cold
Coomassie protein assay reagent (Pierce)
Protein standards: 5, 10, 20, 50, 100, and 200 $\mu\text{g/ml}$ BSA in homogenization buffer
Blocking solution: TBS-T (*UNIT 20.3*) containing 2% (w/v) bovine serum albumin (BSA)
Antibodies for immunoblotting:
 Primary antibody: goat monoclonal anti-phospho-MAPK antibody (anti-DP-ERK, Sigma)
 Secondary antibody: alkaline phosphatase (AP)/enhanced chemiluminescence (ECL)-conjugated anti-goat antibody (see *UNIT 10.8*)
 Polyclonal anti-general ERK antibody from different species than original primary antibody (e.g., Sigma; optional, for second immunoblotting)
 Horseradish peroxidase (HRP)/enhanced chemiluminescence (ECL)-conjugated secondary antibody against species of the anti-general ERK antibody (optional, for second immunoblotting)
Tris-buffered saline with Triton X-100 (TBS-T; *UNIT 20.3*)
Protein A-Sepharose beads, swollen and packed (see recipe)
Phosphate-buffered saline (PBS; *APPENDIX 2*)
Antibody for immunoprecipitation: anti-ERK C-terminal antibody (e.g., Sigma)
0.5 M LiCl solution, ice cold
RadioImmunoPrecipitation Assay (RIPA) buffer (see recipe), ice cold
3 \times reaction mixture (RM \times 3; see recipe)
2 mg/ml myelin basic protein (MBP)
4 \times SDS-PAGE sample buffer (see recipe)
10% to 12% or 15% SDS-polyacrylamide gel (*UNIT 10.2A*)
Prestained protein markers, 16 to 175 kDa (New England Biolabs)
Running buffer: 25 mM Tris-Cl (pH 8.3)/188 mM glycine/0.1% SDS
Staining solution: 40% (v/v) methanol/7% (v/v) acetic acid/25 g/liter Coomassie brilliant blue R-250
Destaining solution: 15% (v/v) isopropanol/7% (v/v) acetic acid in water
1.5-ml microcentrifuge tubes, prechilled to 4°C (four sets of six, each labeled 1 to 6)
Sonicator
Stopwatch
1-ml pipet tips, prechilled to 4°C
Microcentrifuge, 4°C
96-well flat-bottomed microtiter plates
Microtiter plate reader, 595 nm wavelength
End-over-end rotator
Nitrocellulose membrane (see *UNIT 10.8*)
Flat plastic or glass box to accommodate gel
Shields for radioactive work

Tube heater/shaker (e.g., Eppendorf Thermomixer) *or* water bath, 30°C

Flat container for washing gel

Whatman 3MM filter paper, cut larger than the gel

X-ray film (not needed if phosphoimager is used)

Film cassette for either X-ray film or phosphoimager

X-ray film developer or phosphoimager

Additional reagents and equipment for preparation of EGF-stimulated Rat1 cells (Support Protocol 1), SDS-PAGE (UNIT 10.2A), transfer of proteins to nitrocellulose membrane, immunoblotting, and chemiluminescent detection (UNIT 10.8), and autoradiography (APPENDIX 3A)

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper aseptic technique should be used accordingly.

NOTE: All incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified. Some media (e.g., DMEM) may require altered levels of CO₂ to maintain pH 7.4.

Prepare cytosolic extracts

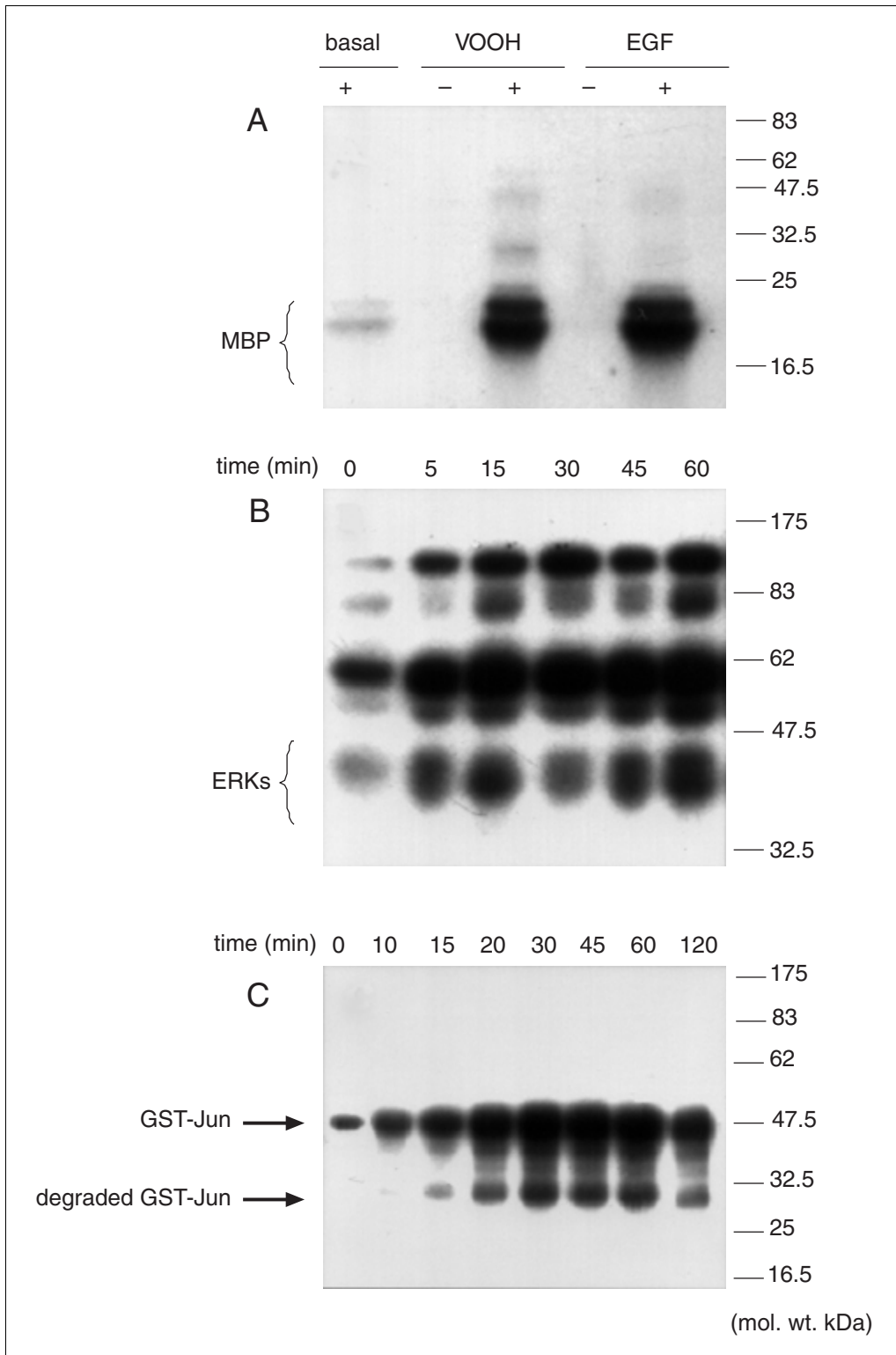
1. Add 350 μ l ice-cold homogenization buffer to each plate of EGF-stimulated Rat1 cells on ice, tilt the plate gently, and scrape the cells into the buffer using a plastic scraper or rubber policeman. Using prechilled pipet tips, transfer the cells and buffer to labeled, prechilled 1.5-ml microcentrifuge tubes.

Special consideration should be given to the composition of homogenization buffer. The authors recommend using β -glycerophosphate, which serves both as a buffer and as a general phosphatase inhibitor; rather than Tris or HEPES. Sodium orthovanadate is used to inhibit tyrosine phosphatases, and the mixture of pepstatin A, aprotinin, leupeptin, and benzamide is used to inhibit proteinases. Ice-cold homogenization buffer blocks most of the phosphatase and proteinase activities in cell extracts.

2. Disrupt the cells by sonication (two 7-sec, 50-W pulses per 0.5-ml sample) on ice.

Sonication allows extraction of proteins from the cytosolic and nuclear fractions of the cells but not from the membrane fraction; therefore, the extract is called a cytosolic extract. Cellular extraction with nonionic detergents, which extract proteins from the membrane, cytosolic, and some nuclear fractions of the cell, makes determination of the protein concentration somewhat difficult but is often used. Extraction with RIPA buffer or by freeze-thawing can be used for some kinases, but these methods are less effective (see Critical Parameters).

Figure 18.12.2 (at right) Detection of mitogen-activated protein kinase (MAPK) activity by the methods described in this unit. **(A)** Detection of extracellular signal-related protein kinase (ERK) activity by immunoprecipitation and myelin basic protein (MBP) phosphorylation. NIH-3T3 cells were grown to subconfluency in 6-cm plates and then starved as described in Support Protocol 1. Cells were then stimulated with 100 μ M sodium orthovanadate and 200 μ M hydrogen peroxide (VOOH) for 15 min, or 50 ng/ml epidermal growth factor (EGF) for 5 min, or left untreated (basal). Cytosolic extracts were prepared by sonication and the resulting proteins (300 μ g) were incubated with either 30 μ l anti-ERK C-terminal antibody (Santa Cruz Biotechnology)-conjugated protein A-Sepharose beads (+) or with unconjugated protein A-Sepharose beads (-). The phosphorylation reaction on MBP was performed as described in Basic Protocol 1. **(B)** Detection of protein kinase activity by the in-gel kinase assay. MCF7 cells overexpressing the ErbB-2 receptor were stimulated with 50 ng/ml EGF for the indicated times. The in-gel kinase assay was performed as described in Basic Protocol 2. ERK1 and ERK2 bands are indicated. The identity of other bands was not determined. **(C)** Detection of c-Jun NH₂-terminal protein kinase (JNK) activity by an affinity assay. α T3-1 cells were treated with 100 nM D-Trp gonadotropin-releasing hormone for the indicated times. The JNK assay was performed as described in Basic Protocol 3. GST-cJun and a degradation product (p30) are indicated. Abbreviation: GST-Jun, glutathione-S-transferase-conjugated cJun.



3. Microcentrifuge the cytosolic extracts 15 min at $14,000 \times g$, 4°C . Transfer supernatants to fresh, prechilled, and prelabeled tubes. Take 5- to 10- μl aliquots of these extracts for protein determination. Store the remainder of each cytosolic extract on ice until needed.

Determine the protein concentration of each sample at this stage so that identical amounts of proteins from the different samples can be compared and the relative amount of protein kinases in each sample determined accurately. Comparing samples based on cell number, rather than protein concentration, can result in differences of up to 20% in the amount of protein. Such differences can cause even larger ones when phosphorylation is assessed immunologically.

Do not leave samples on ice >45 to 90 min, to avoid unnecessary exposure to phosphatases and proteinases (see Critical Parameters).

Determine protein concentration

4. Add 10 μl of each cytosolic extract to 190 μl kinase buffer in labeled 1.5-ml microcentrifuge tubes.

Dilutions of $\geq 1:20$ are usually necessary to ensure that the sample protein concentrations are in the linear range of the protein determination assay. This dilution is not always necessary with some Coomassie protein assay reagents that have extended ranges.

5. Transfer 10 μl of each of the protein standards into two wells each of a 96-well, flat-bottomed microtiter plate.

Protein standards should be prepared in the same buffer as was used for the cell extraction.

6. Transfer 10 μl of each of the diluted cytosolic extracts into two wells of the same microtiter plate. Add 200 μl Coomassie protein assay reagent to all wells.
7. Place the microtiter plate in a microtiter plate reader and measure the absorbance of the samples at 595 nm (A_{595}). Calculate the protein concentrations of the cytosolic extracts by comparing the absorbance of the sample with the standard curve.
8. Based on the calculated protein concentrations, transfer a volume of each cytosolic extract containing 100 to 500 μg protein into a fresh, prechilled 1.5-ml microcentrifuge tube.

Detect activation of ERK

At this stage, activation of ERK can be determined using small amounts of extracted proteins either by immunoblotting with anti-phospho and anti-general-ERK antibodies (steps 9a to 17a) or by immunoprecipitation followed by an in vitro kinase assay (steps 9b to 24b).

To detect ERK activation using anti-phospho antibody

- 9a. Load cytosolic extract on a 10% or 12% SDS-PAGE gel along with a prestained protein marker and electrophorese at 150 V (UNIT 10.2A). Once the dye front has reached the end of the gel, remove the gel from the apparatus and proceed to step 10a.
- 10a. Transfer the separated proteins onto a nitrocellulose membrane for 45 min at 100 V, 4°C (UNIT 10.8). At end of transfer period, rinse membrane with TBS-T and place in flat container.
- 11a. Incubate nitrocellulose membrane in blocking solution for 60 min at room temperature.

Milk is not recommended for blocking because it probably contains phosphorylated casein, which may interfere with the detection of the anti-phospho-ERK antibody.

- 12a. Incubate blot with primary antibody (goat monoclonal anti-DP ERK antibody diluted according to manufacturer's instructions).
- 13a. Wash blot in flat container at least three times, each time for 15 min with TBS-T at 23°C.
- 14a. Incubate blot with secondary antibody (AP/ECL-conjugated anti-goat antibody) for 45 min at room temperature.
- 15a. Wash blot at least three times, each time for 10 min with TBS-T.
- 16a. Detect phosphorylated ERK using chemiluminescent detection (AP/ECL; e.g., *UNIT 10.8*).

After detecting the phosphorylated ERK it is recommended that it be determined whether there is an equal amount of ERK, using anti-general ERK antibody, as described in step 17a.

- 17a. Incubate the blot in blocking solution for 30 min at room temperature. Incubate blot with "new" primary antibody (polyclonal anti-general ERK antibody) from different species than primary antibody used in step 12a. Incubate with HRP/ECL-labeled secondary antibody (if AP/ECL-conjugated secondary antibody was used in step 14a) against species used for the "new" primary antibody and detect using chemiluminescent detection (*UNIT 10.8*).

It should be noted that the different antibodies may interfere with the detection of one another; therefore, the above step must be performed using another, identical blot, or an additional stripping step must be performed on the original blot (UNIT 10.8).

Two or three bands are usually stained by the antibodies. When two bands appear, these are the p42 ERK2 and p44 ERK1. In some cell lines and tissues, a third band at 46 kDa is detected (ERK1b). The intensity of the bands is elevated in the treated samples, reflecting their regulatory phosphorylation upon stimulation, while the amount of ERK as detected by the anti-total ERK antibody is not changed for up to 2 hr of stimulation.

An alternative way to detect ERK activity is by immunoprecipitation followed by in vitro kinase assay (steps 9b to 24b).

To detect ERK activation by immunoprecipitation

- 9b. Place 180 μ l swollen, packed protein A-Sepharose beads in a 1.5-ml microcentrifuge tube. Add 320 μ l PBS and 25 μ l of the antibody to be conjugated, diluted according to the manufacturer's instructions. Rotate the mixture for 1 hr at room temperature on an end-over-end rotator to allow the antibody to bind to the protein A.

This step can be done at 4°C for 16 hr. Ideally, this step should be performed either before or simultaneously with the preparation of cytosolic extracts so that the immunoprecipitation step can proceed without delay.

Anti-C-terminal antibodies are generally used for the determination of kinase activity, because their binding to the kinase does not interfere with its kinase activity. Usually 1 to 5 μ g per reaction is sufficient. Polyclonal antibodies are available from a number of suppliers (e.g., Transduction Laboratories, Sigma, Santa Cruz Biotechnology, Upstate Biotechnology, and Zymed Laboratories).

These volumes are calculated for ten reactions (usually 10 to 20 μ l beads is used per reaction), but because of the density of the beads they will probably only be sufficient for eight reactions. The amounts can be scaled up as long as the proportions are maintained.

For easy handling of the resin, cut the ends of the pipet tips to enlarge their openings.

- 10b. Resuspend beads in 1 ml ice-cold PBS and microcentrifuge 1 min at 14,000 \times g, 4°C. Remove supernatant and add 1 ml ice-cold homogenization buffer. Repeat

centrifugation three more times with homogenization buffer, ending with a final addition of an equal volume of ice-cold homogenization buffer.

Antibody-conjugated beads can be stored in homogenization buffer for ≤ 3 days at 4°C.

- 11b. Add 30 μl antibody-conjugated bead suspension (15 μl beads and 15 μl homogenization buffer) to a 300- μl sample of cytosolic extract containing 50 to 500 μg total protein in prechilled 1.5-ml microcentrifuge tubes. Rotate on end-over-end rotator for 2 hr at 4°C.

Conjugating the antibodies to protein A–Sepharose beads prior to adding them to the cytosolic extracts minimizes the time the samples are incubated with the antibodies, minimizing exposure of the kinases to phosphatases and proteinases in the extracts. This procedure also ensures that only antibodies recognized by protein A will be used for the immunoprecipitation. If polyclonal antibodies are added to the cytosolic extracts, antibodies that are not recognized by protein A can bind to the desired antigen but will not be precipitated when protein A–Sepharose beads are added, and will reduce the efficiency of immunoprecipitation.

- 12b. Microcentrifuge protein A–Sepharose bead/cytosolic extract mixture 1 min at 14,000 $\times g$, 4°C. Discard supernatant, add 1 ml ice-cold RIPA buffer, and centrifuge as before. Repeat centrifugation twice with 1 ml ice-cold 0.5 M LiCl and twice with 1 ml ice-cold kinase buffer.

These stringent washes remove most of the protein kinases that can nonspecifically interact with the protein A–Sepharose beads.

- 13b. After the last wash, completely remove kinase buffer from the conjugated beads by removing supernatant, microcentrifuging 1 min at 14,000 $\times g$, 4°C, without adding more buffer, and gently removing residual buffer from above beads.

- 14b. Resuspend beads in 15 μl water. Prepare the laboratory bench for working with small amounts of radioactivity.

Determination of enzymatic activity is not always accurate when enzymes are bound to beads. The kinase(s) of interest can be released from the beads at step 14b or 15b by the addition of excess immunizing peptide, allowing the phosphorylation reaction (steps 15b to 17b) to be performed without the interference of the beads. The activity can then be measured as described below (steps 18b to 24b) or by paper assay. If by paper assay, terminate the phosphorylation reaction by spotting 20 μl of each cytosolic extract on a phosphocellulose paper square (Whatman P81) and washing immediately with 150 mM phosphoric acid. Measure phosphate incorporation using scintillation cocktail and counter.

- 15b. Add 10 μl RM $\times 3$ to each tube.

The most important components of RM $\times 3$ are the Mg^{2+} and $[\gamma\text{-}^{32}\text{P}]\text{ATP}$, which are essential for the phosphorylation reaction. The use of 100 μM ATP with ~ 4000 cpm/pmol $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ provides a good linear range and reproducible results. When the enzymatic activity of the kinases is low, which makes detection of phosphorylation difficult, the concentration of unlabeled ATP should be reduced to 10 to 20 μM and the amount of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ increased to 50,000 cpm/pmol. Addition of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ alone is not recommended because this will result in a nanomolar concentration of ATP, which is considerably below the K_M for ATP and may lead to nonspecific phosphorylation.

The β -glycerophosphate in the reaction mixture serves as a buffer, but can also inhibit residual phosphatases that may have nonspecifically bound to the beads. BSA serves as a carrier protein; it can be eliminated if purity is required. EGTA chelates Ca^{2+} , which may interfere with some kinase activities, DTT keeps the proteins reduced, and sodium orthovanadate inhibits tyrosine phosphatases.

- 16b. Add 5 μ l of 2 mg/ml MBP to each tube and place the mixture in a 30°C Thermomixer or water bath.

MBP is probably not a physiological substrate for any MAPK, but it is a good general substrate for many kinases, including ERKs, in vitro. Substrates should be well phosphorylated by the desired kinases to allow accurate detection of the phosphorylation reaction.

- 17b. Incubate 20 min at 30°C with either constant or frequent shaking.

If a Thermomixer is not available, a water bath or other heating device can be used.

- 18b. Add 10 μ l of 4 \times SDS-PAGE sample buffer to each tube to stop the phosphorylation reaction.

- 19b. Boil for 5 min. Centrifuge for 1 min at 14,000 \times g, room temperature.

- 20b. Prepare a 15% separating gel with a 3% stacking gel (UNIT 10.2A).

- 21b. Load samples and prestained protein markers on the gel.

Load prestained markers into the first or second lane of the gel so that the first lane can be located on the dried gel and the molecular weights of the detected proteins determined.

- 22b. Run the gel at 150 V, constant voltage, until the bromphenol blue dye is 0.5 cm from the bottom of the gel (~1 hr).

- 23b. Place the gel in a flat container, add 50 ml staining solution, and let stand for 20 min at room temperature. Remove staining solution and add 50 ml destaining solution for 30 min. Repeat destaining three times with 50 ml destaining solution for 30 min each time.

This extensive destaining removes excess free [γ -³²P]ATP which would affect the background radioactivity levels. Extensive destaining is not necessary if the phosphorylation of the desired proteins is very high. Alternatively, the proteins can be transferred to a nitrocellulose membrane (UNIT 10.8) and then exposed as below.

- 24b. Place the gel on Whatman 3MM filter paper, cover with plastic wrap, and dry the gel in a gel dryer for 1.5 hr at 80°C. Expose the gel in a phosphoimager or on X-ray film (UNIT 10.5 or APPENDIX 3A).

Bands should appear at 16 to 21 kDa, which is the molecular weight of the four MBP isoforms.

IN-GEL KINASE ASSAY

If the identity of the kinase is not known or there are no specific antibodies available for the kinase, an in-gel kinase assay can be used in place of immunoprecipitation. This in-gel protocol involves copolymerizing a substrate with SDS and polyacrylamide and running the protein sample on the resulting SDS-polyacrylamide gel. After several rounds of denaturation and renaturation, a phosphorylation reaction is performed on the gel and the phosphorylated bands are visualized by autoradiography or by a phosphoimager (UNIT 10.5). With this method, the molecular weight of the protein kinase is revealed and novel protein kinases can be identified. Not all protein kinases can be renatured under the conditions of this protocol, however, and the linear range of this assay is usually limited; therefore, this method should not be used routinely to monitor and characterize known protein kinases.

BASIC PROTOCOL 2

Analysis of Protein Phosphorylation

18.12.9

Materials

Cell lysates from serum-starved, EGF-stimulated Rat1 cells (Support Protocol 1, steps 1 to 8) or cytosolic extracts (see Basic Protocol 1, steps 1 to 3)
4× SDS-PAGE sample buffer (see recipe)
2 mg/ml myelin basic protein (MBP)
1.5 M Tris·Cl, pH 8.8 (APPENDIX 2)
30% acrylamide/0.8% bisacrylamide (UNIT 10.2A)
10% ammonium persulfate
Stacking gel (UNIT 10.2A)
TEMED
20% isopropanol/50 mM HEPES, pH 7.6
Renaturation buffer: 50 mM HEPES (pH 7.6)/5 mM 2-mercaptoethanol
Renaturation buffer/0.05% Tween 20
6 M urea in kinase buffer (see recipe for kinase buffer)
In-gel kinase buffer: 20 mM HEPES (pH 7.6)/20 mM MgCl₂
In-gel kinase buffer/2 mM DTT/20 μM ATP/100 μCi [γ -³²P]ATP
5% trichloroacetic acid (TCA)/1% sodium pyrophosphate (NaPPi)
Water bath, 30°C, with proper shielding for radioactive work

Additional reagents and equipment for SDS-PAGE (UNIT 10.2A) and autoradiography (APPENDIX 3A)

Procedure

1. Determine protein concentration of cell lysates (Support Protocol 1, steps 9 to 12) or cytosolic extracts (see Basic Protocol 1, steps 4 to 8). Based on the calculated protein concentrations, transfer a volume of each cell lysate containing 30 to 80 μg protein to a fresh 1.5-ml microcentrifuge tube.
2. Add ¼ vol of 4× SDS-PAGE sample buffer to each tube and mix thoroughly. Keep at 4°C.

IMPORTANT NOTE: *Do not boil the samples.*
3. Prepare 8 ml of 12% polyacrylamide gel (separating gel) containing MBP by mixing the following:
 - 0.7 ml H₂O
 - 2.0 ml MBP
 - 2.0 ml of 1.5 M Tris·Cl, pH 8.8
 - 3.2 ml 30% acrylamide/0.8% bisacrylamide
 - 100 μl of 10% ammonium persulfate
 - 6 μl TEMED.

After the separating gel polymerizes, add a 3% stacking gel (UNIT 10.2A).

4. Load samples and prestained markers onto gel and run at 100 V until the dye front has reached 0.5 cm from the bottom of the gel (~90 min).

The gel should not be heated above 30°C; therefore, the voltage used for electrophoresis should be ≤100 V.
5. Remove the gel from the apparatus, cut off the stacking gel, and carefully place the gel in a flat container. Wash twice with 100 ml of 20% isopropanol/50 mM HEPES, pH 7.6, for 30 min each at room temperature. Repeat washing twice with 100 ml renaturation buffer for 30 min each at room temperature and twice with 100 ml of 6 M urea in kinase buffer for 15 min each at room temperature.

The second wash with 20% isopropanol/50 mM HEPES, pH 7.6, can be done overnight at 4°C.

6. Place the gel in 4°C cold room, remove 50 ml of washing solution, add 50 ml renaturation buffer/0.05% Tween 20, and shake for 15 min. Repeat twice, removing 50 ml of the washing solution, adding 50 ml renaturation buffer/0.05% Tween 20, and shaking for 15 min each time. Wash the gel three times with 100 ml renaturation buffer/0.05% Tween 20 for 15 min each time. Shake the gel overnight in the cold room.
7. Remove washing buffer and incubate the gel in 30 ml in-gel kinase buffer for 30 min at 30°C. Remove the buffer, add 20 ml in-gel kinase buffer/2 mM DTT/20 μM ATP/100 μCi [γ -³²P]ATP, and incubate in 30°C water bath for 2 hr.

The amount of radioactive material is very high at this stage and the reaction should be performed with proper shielding. Make sure that the gel is straight in the flat container. Unequal distribution of the phosphorylation buffer can interfere with the phosphorylation reaction.

8. Wash the gel carefully four times with 5% TCA/1% NaPPi for 15 min each at room temperature. If the gel is still very radioactive, continue washing overnight.
9. Dry the gel (see Basic Protocol 1, step 24b) and subject to autoradiography.

Bands should appear where kinases phosphorylated the MBP copolymerized in the gel (see Fig. 18.12.2).

JNK ASSAY

Protein kinases can sometimes be isolated without using specific antibodies. For example, with JNK it is possible to exploit the tight and specific binding of the three forms of JNK (p46, p52, and p54) to their substrate, Jun (Hibi et al., 1993). Thus, glutathione-*S*-transferase (GST)–Jun bound to glutathione-conjugated beads can be used to specifically isolate the JNKs, while stringent washings remove irrelevant protein kinases. Phosphorylation using [γ -³²P]ATP and Mg²⁺ is then performed with the bead-bound kinase and is detected by autoradiography after SDS-PAGE. This affinity purification method has been successfully used in many recent studies. When GST-Jun on glutathione-conjugated beads is used as the substrate, this method gives more accurate results than the purification of JNKs by immunoprecipitation (Basic Protocol 1). The amount of immobilized Jun is the limiting factor, and the linear range of this assay is limited. Preparation of the GST-Jun beads is described below (see Support Protocol 2).

Materials

- Cytosolic extracts from serum-starved, EGF-stimulated Rat1 cells (see Basic Protocol 1, steps 1 to 3)
- 10× binding buffer (see recipe)
- GST-Jun beads: GST-Jun(1–91) bound to glutathione-conjugated beads (see Support Protocol 2)
- HB1B buffer (see recipe), ice cold
- JNK kinase buffer (see recipe), ice cold and room temperature
- 10× ATP mix: 20 mM unlabeled ATP/2 μCi [γ -³²P]ATP, freshly prepared
- 4× SDS-PAGE sample buffer (see recipe)
- 12% SDS-polyacrylamide gel (UNIT 10.2A)
- 1.5-ml microcentrifuge tubes, prechilled
- Microcentrifuge, 4°C
- Tube heater/shaker (e.g., Eppendorf Thermomixer) or water bath, 30°C
- Additional materials and equipment for SDS-PAGE (UNIT 10.2A)

BASIC PROTOCOL 3

Analysis of Protein Phosphorylation

18.12.11

Procedure

1. Determine the protein concentration of cytosolic extracts (see Basic Protocol 1, steps 4 to 8).
2. Place 150 μ l cytosolic extract or cell lysate (50 to 500 μ g protein), 30 μ l of 10 \times binding buffer, 100 μ l water, and 20 μ l GST-Jun beads in prechilled 1.5-ml microcentrifuge tubes.

The amount of GST-Jun beads can vary according to the amount of the protein conjugated to the beads. Amounts of 2 to 4 μ g protein per 20 μ l beads usually give good results. The truncated form of Jun (residues 1 to 91) is recommended for these experiments, but similar results can be obtained with the full-length or 1-to-74 constructs.

3. Incubate 1 hr at 4°C with constant shaking. Microcentrifuge 2 min at 14,000 \times g, 4°C, and discard the supernatant.

The JNKs bind to the GST-Jun on the beads during this incubation.

4. Resuspend pelleted beads in 1 ml ice-cold HB1B buffer and microcentrifuge for 2 min at 14,000 \times g, 4°C. Discard supernatants and repeat centrifugation twice with 1 ml ice-cold HB1B buffer each time and once with 1 ml ice-cold JNK kinase buffer.

Remove the supernatant completely after each centrifugation.

5. Add 30 μ l JNK kinase buffer at room temperature to the pelleted beads and place the tubes in 30°C Thermomixer or water bath. Add 3 μ l of 10 \times ATP mix. Close the tubes and incubate 20 min with shaking.

Take proper precautions when working with radioactive material.

6. Add 11 μ l of 4 \times SDS-PAGE sample buffer. Boil the samples 5 min.
7. Prepare 12% SDS-polyacrylamide gel with a 3% stacking gel (UNIT 10.2A).
8. Load samples and prestained protein markers onto the gel. Run the gel at 150 V, constant voltage, until the bromphenol blue reaches 0.5 cm from the bottom of the gel (~1 hr).
9. Stain, dry, and analyze the gel (see Basic Protocol 1, steps 23b and 24b).

A band should be detected at 46 kDa (the molecular weight of the truncated GST-Jun). In many cases, an additional band is observed at 30 kDa, which represents a degradation product of GST-Jun and is a good indication of JNK activity as well.

SUPPORT PROTOCOL 1

PREPARATION OF EGF-STIMULATED RAT1 CELL EXTRACT

This protocol describes EGF stimulation of Rat1 cells, but with minor changes can be used with almost all tissue culture cell lines, homogenized animal organs, and even whole lower organisms.

Materials

Rat1 cells
Dulbecco's modified Eagle medium containing 10% heat-inactivated fetal bovine serum (DMEM/10% FBS; APPENDIX 3F)
Starvation medium: DMEM/0.1% FBS
Epidermal growth factor (EGF) buffer: phosphate-buffered saline (PBS)/0.5 mg/ml bovine serum albumin (BSA, crystalline)
50 μ g/ml EGF in EGF buffer
PBS (APPENDIX 2), ice cold
Homogenization buffer (see recipe), ice cold
Lysis buffer: homogenization buffer/1% Triton X-100, ice cold
Kinase buffer (see recipe), ice cold

Protein standards: 5, 10, 20, 50, 100, and 200 µg/ml BSA in homogenization buffer/0.03% Triton X-100

Coomassie protein assay reagent (Pierce)

6-cm tissue culture plates

1.5-ml microcentrifuge tubes, prechilled to 4°C (four sets of six, each set labeled 1 to 6)

Stopwatch

1-ml pipet tips, prechilled to 4°C

Microcentrifuge, 4°C

96-well flat-bottomed microtiter plate

Microtiter plate reader, 595-nm wavelength

Additional reagents and equipment for tissue culture (*APPENDIX 3F*)

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper aseptic technique should be used accordingly.

NOTE: All tissue culture incubations should be performed in a humidified 37°C, 5% to 10% CO₂ incubator unless otherwise specified.

Prepare cellular extracts

1. Grow six 6-cm tissue culture plates of Rat1 cells in DMEM/10% FBS to subconfluency ($\sim 0.5 \times 10^6$ cells/plate).
2. Remove culture medium, add 2 ml starvation medium to each plate, and culture a further 18 hr.

Make sure that the plates remain flat and that the medium covers the entire plate. Serum starvation makes the cells quiescent, which can be achieved under these conditions within 14 to 24 hr. Starvation for too long or any change in temperature or pH may be stressful to the cells and may induce activation of one or more signaling pathways.

3. Add 2.5 µl of 50 µg/ml EGF to three plates (stimulated) and 2.5 µl EGF buffer to three plates (control). Return the plates to the incubator for 5, 15, and 45 min.

Add EGF or EGF buffer first to the plates with the longest incubation, then at appropriate intervals to the plates with the next longest incubation. It is useful to make and use a time chart to ensure that stimuli will be given at the appropriate times and the cells harvested within a short period of time (5 to 10 min).

If the influence of the stimulating agent on the particular cells being used is not yet known, a positive control should be included—for example, peroxovanadate (0.1 mM vanadate/0.2 mM hydrogen peroxide), which nonspecifically activates many signaling events in most tissue culture cells.

4. After 5, 15, or 45 min, remove medium from plates. Rinse plates twice with 5 ml ice-cold PBS and once with 5 ml ice-cold homogenization buffer.

Since slowing and arresting of biological processes is desired at this stage, the plates should be placed on ice. Washing and harvesting each plate takes 0.5 to 1.5 min; all six plates should be harvested within 5 to 10 min.

5. Add 350 µl ice-cold lysis buffer to each plate, tilt the plate gently (on ice), and scrape the cells into the buffer using a plastic scraper or rubber policeman. Using prechilled 1-ml pipet tips, transfer the cells and buffer to prelabeled, prechilled 1.5-ml microcentrifuge tubes.

The sample being harvested may become slightly viscous; the 1-ml pipet tips can be cut to make the tip opening wider, allowing easy collection of these samples.

Detergents other than Triton X-100, such as 0.5% Nonidet P-40 (NP-40; Igepal), can also be used.

6. Continue the cell lysis for an additional 10 min on ice.

A nonionic detergent is used for membrane disruption because it extracts proteins from membranal, cytosolic, and some nuclear fractions of the cell, and because the determination of protein concentration in such extracts is relatively easy (steps 9 to 12). Alternatively, a more stringent mix of detergents (RadioImmunoPrecipitation Assay [RIPA] buffer, see Reagents and Solutions), which causes complete extraction of proteins from most cell compartments, can be used for membrane disruption, although it might interfere with the protein determination. Cells can also be disrupted by sonication (two 7-sec, 50-W pulses for a 0.5-ml sample) on ice, which extracts proteins from the cytosolic and nuclear fractions, but not from membranes. The extraction by sonication is performed in homogenization buffer without detergents, and allows easy determination of protein concentration.

7. Microcentrifuge the cellular extracts 15 min at $15,000 \times g$, 4°C . Transfer supernatants to fresh, prechilled, and prelabeled microcentrifuge tubes.

The supernatants contain the protein extracts to be examined for phosphorylation.

8. Transfer 5- to 10- μl aliquots of each extract to labeled microcentrifuge tubes for determination of protein concentration, and store the remainder of the extracts on ice until needed.

The protein concentration is determined at this stage so that identical amounts of proteins from the different samples can be compared and the relative amount of phosphoproteins can be determined accurately. Comparing samples based on cell number, rather than protein concentration, can result in differences of up to 20% in the amount of protein; such large differences can cause even larger ones when phosphorylation is assessed immunologically.

Determine protein concentration

9. Add 145 μl kinase buffer to each 5- μl aliquot of cellular extract.

High concentrations of Triton X-100 can interfere with colorimetric measurements of the Coomassie brilliant blue; therefore, dilutions of $\geq 1:20$ are necessary. Similar or higher dilutions are required for sonicated and RIPA extracts as well, because of high detergent or protein concentrations. The Lowry method of protein determination cannot be used because of the dithiothreitol used to prevent degradative oxidation in the extraction buffer.

10. Transfer 10 μl of each of the protein standards into two wells of a 96-well flat-bottomed microtiter plate.

Protein standards should be prepared in the same buffer as was used for cell extraction.

11. Transfer 10 μl of each of the diluted cellular extracts into two wells of the same microtiter plate. Add 200 μl Coomassie protein assay reagent to all wells.

12. Place the microtiter plate in a microtiter plate reader and measure the absorbance of the standards and samples at 595 nm (A_{595}). Use the absorbance of the standards to construct a standard curve (absorbance versus concentration). Calculate the protein concentrations of the samples by comparing the absorbance of the sample with the standard curve.

SUPPORT PROTOCOL 2

PREPARATION OF GST-JUN-GLUTATHIONE BEADS

Materials

Bacteria transformed with GST-Jun(1–91)-expressing plasmid (plasmid is available from several investigators; alternatively, similar proteins can be obtained from, e.g., Calbiochem)

LB medium with appropriate antibiotics (UNIT 1.1)

100 mM isopropyl- β -D-thiogalactopyranoside (IPTG; store in aliquots <1 year at -20°C)

PBS/protease inhibitors (see recipe), ice cold
Washed glutathione beads (see recipe)
PBS/protease inhibitors/20% (v/v) glycerol, ice cold
End-over-end rotator

1. Grow bacteria transformed with GST-Jun-expressing plasmid at 30°C in 4 liters LB medium until optical density at 600 nm (OD_{600}) is 0.6 (~3 to 4 hr).
2. Add IPTG to 0.4 nM final and grow bacteria an additional 4 hr at 30°C.
3. Centrifuge 10 min at $6000 \times g$, 4°C.
4. Resuspend pellet in 80 ml ice-cold PBS/protease inhibitors, sonicate, and centrifuge 10 min at $15,000 \times g$, 4°C. Transfer supernatant to a fresh tube and repeat centrifugation.
5. Discard pellet, add 3 ml washed glutathione beads to supernatant, and rotate end over end 2 hr at 20 rpm, 4°C.
6. Wash beads with ice-cold PBS three times. Add 8 ml PBS/protease inhibitors/20% glycerol.

Beads can be stored 1 to 2 days at 4°C or <4 months at -20°C in ice-cold PBS/protease inhibitors/20% glycerol before being used in Basic Protocol 3.

REAGENTS AND SOLUTIONS

Use deionized or distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Binding buffer, 10×

1.5 M NaCl
220 mM HEPES, pH 7.7
20 mM $MgCl_2$
0.75% (v/v) Triton X-100
200 mM β -glycerophosphate
1 mM EDTA
1 mM sodium orthovanadate
1 mM phenylmethylsulfonyl fluoride (PMSF)
10 μ g/ml aprotinin
10 μ g/ml leupeptin
2 μ g/ml pepstatin A
Store 2 months at 4°C

This buffer is derived from Hibi et al. (1993).

Glutathione beads, washed

Add 50 ml ice-cold PBS (APPENDIX 2) to 3 ml glutathione beads, centrifuge 5 min at $2000 \times g$, 4°C, and discard supernatant. Repeat centrifugation three times, adding 50 ml ice-cold PBS each time.

HB1B buffer

20 mM HEPES, pH 7.7
50 mM NaCl
0.1 mM EDTA
25 mM $MgCl_2$
0.05% (v/v) Triton X-100
Store 2 months at 4°C

Homogenization buffer

50 mM β -glycerophosphate, pH 7.3
1.5 mM EGTA
1.0 mM EDTA
0.1 mM sodium vanadate
1.0 mM benzamidine
10 μ g/ml aprotinin
10 μ g/ml leupeptin
2.0 μ g/ml pepstatin A
1.0 mM DTT (*APPENDIX 2*)
Store up to 3 months at 4°C without DTT
Add DTT just before use

This buffer is derived from the work of Ahn et al. (1990) and Seger et al. (1994).

JNK kinase buffer

20 mM HEPES, pH 7.7
20 mM MgCl_2
20 mM β -glycerophosphate
0.1 mM sodium orthovanadate
2 mM dithiothreitol (DTT; *APPENDIX 2*), added fresh
Store without DTT for 2 months at 4°C

Kinase buffer

50 mM β -glycerophosphate, pH 7.3
1.5 mM EGTA
1.0 mM EDTA
0.1 mM sodium vanadate
1.0 mM dithiothreitol (DTT; *APPENDIX 2*)
Store up to 3 months at 4°C without DTT
Add DTT just before use

This buffer is derived from the work of Ahn et al. (1990).

PBS/protease inhibitors

PBS containing:

10 μ g/ml leupeptin (from 10 mg/ml stock solution)
10 μ g/ml aprotinin (from 2 mg/ml stock solution)
2 μ g/ml pepstatin A (from 10 mg/ml stock solution)
1 mM phenylmethylsulfonyl fluoride (PMSF)
Prepare fresh

Leupeptin and aprotinin are prepared in water; pepstatin is prepared in ethanol or methanol. All stock solutions are stored <1 year at -20°C.

Protein A-Sepharose beads

Place protein A-Sepharose beads (~150 μ l) in a 1.5-ml microcentrifuge tube, add 1 ml PBS (*APPENDIX 2*), and let the beads swell for 10 min at room temperature. Add 1 ml PBS to swollen beads and microcentrifuge 1 min at 15,000 \times g, room temperature. Discard the supernatant. Repeat for a total of three washes.

Although protein A-Sepharose is recommended for this method, other commercially available protein A-conjugated resins (e.g., agarose, HiTrap) can be used. Protein G-coupled resins are required to immunoprecipitate certain types of monoclonal antibodies.

Some resins are supplied as ready-to-use solutions and will not require this swelling step.

RadioImmunoPrecipitation Assay (RIPA) buffer

20 mM Tris·Cl, pH 7.4 (*APPENDIX 2*)
137 mM NaCl
10% (v/v) glycerol
0.1% (w/v) SDS
0.5% (w/v) deoxycholate
1% (v/v) Triton X-100
2.0 mM EDTA
1.0 mM PMSF (*APPENDIX 2*; add fresh)
20 μ M leupeptin
Store up to 3 months at 4°C

Reaction mixture, 3 \times (RM \times 3)

75 mM β -glycerophosphate, pH 7.3
30 mM MgCl₂
0.9% (w/v) BSA
3 mM DTT (*APPENDIX 2*)
3 mM EGTA
0.3 mM sodium orthovanadate
100 μ M [γ -³²P]ATP (~4000 cpm/pmol)
Store without [γ -³²P]ATP for 1 year at -20°C

SDS-PAGE sample buffer, 4 \times

200 mM Tris·Cl, pH 6.8 (*APPENDIX 2*)
40% (v/v) glycerol
8% (w/v) SDS
0.2% (w/v) bromphenol blue
8% (v/v) 2-mercaptoethanol
Store up to 12 months at -20°C

COMMENTARY

Background Information

The MAPK cascades

Sequential activation of kinases (protein kinase cascades) is a common mechanism of signal transduction in many cellular processes (Campbell et al., 1995). Several related intracellular signaling cascades, collectively known as mitogen-activated protein kinase (MAPK) signaling cascades, have been elucidated so far (Davis, 1994; Seger and Krebs, 1995; Kyriakis and Avruch, 1996; Marshall, 1996; Robinson and Cobb, 1997). Each cascade seems to consist of up to five levels of protein kinases that sequentially activate each other by phosphorylation. The similarities between the enzymes that comprise each level in the various cascades make these kinases part of a superfamily of kinases.

The activation of each of these cascades is initiated either by small GTP-binding proteins or by adaptor proteins that transmit the signal to protein kinases, commonly referred to as MAPK kinase kinase (MAP4Ks). Then

the signal from each of these branches is transmitted downstream through the cascade by enzymes at the next levels that are referred to as MAP3K (MAPK kinase kinase), MAPKK (MAPK kinase), MAPK, and MAPK-activated protein kinases (MAPKAPKs). The existence of four to five levels in each MAPK cascade allows for signal amplification, determination of specificity, and tight regulation of the transmitted signal.

The four distinct MAPK cascades that are currently known are named according to the subgroup of their MAPK components: the ERK (extracellular signal-related protein kinase), JNK (c-Jun NH₂-terminal kinase), p38MAPK (SAPK 2-4), and BMK (big MAPK or ERK 5) cascades. These MAPK cascades cooperate to transmit signals to their intracellular targets, thereby initiating cellular processes such as proliferation, differentiation, development, stress response, and apoptosis. In this section, the various MAPK cascades will be briefly reviewed.

The ERK cascade, also known as the p42, p44 MAPK cascade, was the first MAPK cascade elucidated (Seger and Krebs, 1995). This cascade is initiated by the small G protein Ras, which upon stimulation causes membrane translocation and activation of the protein serine/threonine kinase Raf1. Once activated, Raf1 continues transmission of the signal by phosphorylating two regulatory serine residues located in the activation loop of MEK, causing its full activation. Other kinases that can also activate MEK are A-Raf, B-Raf, Mos, TPL2, and MEKK2 (Seger and Krebs, 1995; Salmeron et al., 1996; Wang et al., 1996), which all seem to phosphorylate the same regulatory residues of MEK. Activated MEK is a dual-specificity protein kinase that appears to be the only kinase capable of specifically phosphorylating and activating the next kinase in this cascade, which is ERK.

ERK activation requires phosphorylation of two regulatory residues, threonine and tyrosine, that reside in a TEY phosphorylation motif (Payne et al., 1991; Canagarajah et al., 1997). Phosphorylation of threonine and tyrosine residues is essential for the activation of all MAPKs. In the other cascades, however, the identity of the middle amino acid in the TXY motif of the MAPK varies and determines the specificity of the signal.

ERK appears to be an important regulatory molecule. ERK, which by itself can phosphorylate regulatory targets in the cytosol such as phospholipase A₂ (PLA₂; Lin et al., 1993), can translocate into and phosphorylate substrates in the nucleus such as ELK1 (Chen et al., 1992; Marais et al., 1993) and can transmit the signal to the MAPKAPK level. The main MAPKAPK of the ERK cascade is RSK, which can also translocate to the nucleus upon activation and phosphorylate a set of nuclear substrates different from those phosphorylated by ERK. MNK and MSK are other MAPKAPKs, but these are activated also by the p38 MAPK cascade (Fukunaga and Hunter, 1997; Waskiewicz et al., 1997; Deak et al., 1998). Activation of the ERK cascade was initially implicated in the transmission and control of mitogenic signals; this cascade is now known to be important also for differentiation, development, stress response, learning and memory, and morphology determination.

ERKs are activated primarily by mitogenic signals, whereas other MAPK cascades are activated mainly by cellular stress, such as heat shock, ischemia, UV irradiation, and cytokines

(Woodgett et al., 1996) and are referred to as stress-activated protein kinase (SAPK) cascades (Kyriakis and Avruch, 1996). The SPK (p38 MAPK) cascade consists of MAPKs that contain a glycine residue in their TXY activation motif (TGY; Han et al., 1994).

Many kinases in the MAP3K and MAP4K levels have been implicated in the SPK cascade (Fig. 18.12.1), but their individual roles are not yet known. GCK1 and HPK1 (Kiefer et al., 1996; Pombo et al., 1995), and probably also PAK1 (Zhang et al., 1995), may belong to the MAP4K level of SPK, although it is not clear whether they are involved in the activation of SAPKs. Twelve distinct kinases have already been implicated in the MAP3K level of this cascade: MEKK1 to 5, MTK1, MLK2, MLK3, ASK1, TPL2, DLK, and TAK1 (Yamaguchi et al., 1995; Blank et al., 1996; Fan et al., 1996; Salmeron et al., 1996; Wang et al., 1996; Deacon and Blank, 1997; Gerwins et al., 1997; Takekawa et al., 1997; Wang et al., 1997). At the MAPKK level, MKK3 (MEK3), MKK6 (MEK6), and, to a lesser extent, also MKK4 (MEK4) seem to play a role in the activation of p38 MAPK (Kyriakis and Avruch, 1996).

The MAPK level components of the p38MAPK cascade are also known as RK, Hog, SAPK2a, and CSBP (Han et al., 1994), SAPK2b, SAPK3, and SAPK4 (Kyriakis et al., 1994; Cuenda et al., 1997; Goedert et al., 1997). Once these p38 MAPKs are activated, they either transmit the signal to the MAPKAPK level components MAPKAPK 2 and 3 (Stokoe et al., 1992; McLaughlin et al., 1996), MNK, and MSK, or phosphorylate regulatory molecules such as PLA₂ (Kramer et al., 1996) and the transcription factors ATF2, ELK1 (Karin, 1995), CHOP (Wang and Ron, 1996), and MEF2C (Han et al., 1997).

JNKs (also called SAPK1), which comprise a third MAPK subgroup, are also SAPKs. However, these enzymes are distinct from the p38 MAPKs, mainly because they contain TPY rather than TGY residues in their activation motif. Like the other MAPK cascades, the JNK cascade is triggered by the small GTPases (Crespo et al., 1997) Rac and CDC42, and can also be activated by specific adaptor proteins. The signals are then transmitted via MAP4K and MAP3K components that are largely shared with the p38 MARK cascades. Since the SPK and JNK cascades are not always simultaneously activated, the signals must be separately regulated to allow separate cascades. These distinct activations are likely to be regulated by

various scaffold proteins, specific to each one of these cascades.

At the MAPKK level, the JNKs can be activated by two dual-specificity enzymes (JNKKs): MKK4 (SKK1, JNKK1; Yan et al., 1994) and MKK7 (Holland et al., 1997; Tournier et al., 1997; Lu et al., 1997). All three JNKKs seem to be able to activate the components at the MAPK level, JNK1 to 3, which have molecular weights 46, 54, and 52 kDa, respectively. No enzymes at the MAPKAPK level and no cytosolic targets have been identified for JNKs, but these enzymes appear to be major regulators of nuclear processes, in particular transcription. Shortly after activation, JNKs translocate into the nucleus where they physically associate with and activate their target transcription factors (e.g., c-Jun, ATF, Elk, etc.).

The BMK comprise another MAPK subgroup (Zhou et al., 1995; Abe et al., 1996) with molecular weights of ~110 kDa. BMK1 (also known as ERK5) and MEK5 are the only known components of this MAPK cascade. Like ERKs, BMK contain a TEY phosphorylation motif. They seem to be involved in stress processes, as well as in regulation of proliferation (Kato et al., 1998). The C terminus of BMK is unique, and appears to act as a transcription factor by itself. Thus, BMK may have a dual enzymatic activity, and this fact singles it out among the other MAPKs.

Kinase cascades other than the MAPK cascades are also activated in response to mitogenic stimulation. These include the NIK-IKK1/2 (DiDonato et al., 1997; Regnier et al., 1997) and PI3K-PDK-AKT-GSK3 (Cohen et al., 1997) Rho-dependent pathways (Leung et al., 1995) and the phosphokinase A (PKA)-phosphorylase kinase pathway (Campbell et al., 1995). Because of their distinct characteristics, these pathways are usually not considered to be genuine MAPK cascades, although they are involved in transmission of many extracellular signals.

All the pathways mentioned are apparently activated to some extent by distinct extracellular agents and, as a result of their action in an elaborate network, determine the outcome of each stimulation. The full dimensions of this network, the mode of regulation of its components, and the mechanisms by which these cascades determine cell fates in response to various stimuli have yet to be fully elucidated.

Analysis of cascade enzymes

The relative strength and the duration of the signals transmitted in each MAPK cascade are thought to be major determinants of signaling specificity. Accurate detection of the signals transmitted via various MAPK cascades towards target molecules is important, therefore, in the study of intracellular signaling. The activity of one component of the MAPK level of each cascade (e.g., ERK, JNK, SPK) is usually a sufficient indicator of the transmitted signal. Sometimes, however, the activity of additional components at upstream or downstream levels must be determined because of cross-talk between various cascades. For example, p38MAPK can be activated by as many as three distinct MAPKKs (MKK3, 4, or 6; Kyriakis and Avruch, 1996), and it is important to determine which of these MAPKKs is the immediate activator in different systems.

Most components of the MAPK cascades belong to the large family of protein kinases, which consists of several hundred distinct members. For studying protein kinases in general and MAPK components in particular, specific detection of the activity or phosphorylation of the desired protein kinase is essential. The activity of a particular protein kinase can be singled out from a multitude of related activities that might mask its activity in two main ways. One method uses a specific substrate that is recognized only by the desired protein kinase. This method is good for detecting kinases like MEK, which seems to specifically and selectively phosphorylate its downstream component, ERK. The other method is to isolate the protein kinase and then use general substrates as indicators of its activity. This method has been used successfully in studies of MAPK cascades. However, in most studies, anti-phospho-MAPKs are used to detect the level of MAPK phosphorylation, which is well correlated to its activity (Yung et al., 1997).

In one of the first methods used for the systematic detection of protein kinases involved in growth factor signaling, protein kinases were isolated using Mono Q fast protein liquid chromatography (FPLC; Ahn et al., 1990). This method involves stimulating tissue culture cells, fractionating the cytosolic extracts of these cells on a Mono Q column (Pharmacia Biotech), and examining the resulting fractions for protein kinase activity. Since fractionation with Mono Q columns is extremely reproducible, kinases that are activated upon stimulation can be detected by comparing

the elution profiles of kinases from activated and unactivated cells. The advantages of this method are (1) the ability to identify novel kinases and measure their activity, (2) the ability to detect the overall activity of many protein kinases, and (3) its good linear range, which allows determination of the ratio between the activities of distinct protein kinases at a given time. The main disadvantage of this method is that separation of various protein kinases is not always complete. In addition, this is a very laborious method and it is difficult to examine more than one sample per day.

Another method that the authors have found useful in detecting novel protein kinases is the in-gel kinase assay (see Basic Protocol 2; Kameshita and Fujisawa, 1989). This technique involves copolymerization of a given substrate in an SDS-polyacrylamide gel, electrophoresis of the samples of interest on the copolymerized gel, and in-gel phosphorylation in the presence of [γ - 32 P]ATP. The advantage of this method is that it reveals the molecular weight of the kinases with the desired specificity, which helps to identify the enzymes of interest. Also, several samples can be examined simultaneously. The main disadvantages are that (1) not all protein kinases can be renatured in the SDS gel, (2) each in-gel assay takes ~3 days, and (3) there is a narrow linear range of protein kinase activities, which can interfere with detection of the increase in induction of protein kinases upon stimulation.

The Mono Q fractionation and in-gel kinase assay methods are mainly used to identify or characterize novel protein kinases. The resolution of these two methods is not always adequate, however, and more specific and convenient methods are recommended for the characterization of a given protein kinase. Such specific methods often require the isolation of the protein kinase of interest, although a specific activator or substrate can sometimes be used (as is the case with PKA or MEK). In studies of MAPKs, the desired protein kinases are often isolated by immunoprecipitation with specific antibodies directed to the C-terminal domain of the kinase or by immunoblot analysis with antibodies to the activated kinase. Antibodies can also be used to detect slower migration on SDS-polyacrylamide gel electrophoresis (PAGE) that occurs upon phosphorylation of regulatory residues of some MAPKs. This gel shift does not always correlate with enzymatic activity, however, as was shown for ERK and for Raf1. Methods for affinity purification that do not involve antibod-

ies can sometime be used to isolate given protein kinases (see Basic Protocol 3). Although affinity techniques (including immunoprecipitation) are often used, it should be noted that the attachment to a solid support that occurs in this method can interfere with the accurate detection of the kinase activity.

Critical Parameters

Several points should be considered when using immunoprecipitation followed by phosphorylation. One of the most important parameters for the success of this procedure is the method of protein extraction used. Since MAPKs are localized within cells, the cellular membranes must be disrupted to access the desired targets. The protein kinases of interest must then be obtained and preserved in their active form, and the amount of irrelevant kinases reduced. For example, activated Raf-1 can be present in mitochondrial membranes, which might not be disrupted by some extraction procedures but are disrupted if RadioImmunoPrecipitation Assay (RIPA) buffer is used for extraction. The method for cellular extraction described in Basic Protocol 1 can be effectively used to detect most MAPKs by immunoprecipitation or other methods. Sonication disrupts the plasma membrane but does not solubilize it; therefore the resulting extracts should contain both cytosolic and some nuclear fractions. Depending on the subcellular localization of the proteins of interest, other extraction methods can be used. For example, cellular extracts obtained with detergents such as Triton X-100 usually contain membrane, cytosolic, and some nuclear components. Cellular extracts obtained with RIPA buffer should contain solubilized proteins from most cellular compartments. Cellular extraction by the addition of hot SDS-PAGE sample buffer to cells is not recommended, because it may free chromatin, which causes formation of a gel that is hard to handle. Extraction by freeze/thawing is also not recommended, because molecular degradation can occur during thawing.

Another consideration for successful detection of phosphoproteins is minimization of protein degradation and dephosphorylation. During extraction, most cellular organelles break, exposing phosphoproteins to phosphatases and proteinases. To minimize the effects of these enzymes, specific inhibitors of phosphatases and proteinases can be added to the extraction buffers and extraction can be performed at low temperature. Phosphatases are usually efficient enzymes, however, and extractions should be

performed as fast as possible even if these precautions are taken.

The success of the immunoprecipitation protocol also depends on the quality and specificity of the antibodies used. The antibodies employed should recognize only the desired protein kinase, not isoforms or irrelevant enzymes. The antibodies should also not interfere with the enzymatic activity of the enzymes tested. In addition, the amount of protein in the different samples and the dilution of antibodies should be optimized to avoid nonspecific recognition of excess proteins. Stringent washing of the immunoprecipitates is necessary to reduce nonspecific precipitation of contaminating kinases. Washing may not completely prevent coimmunoprecipitation of protein kinases other than those desired, however, and these might interfere with the phosphorylation reaction. In this case, it may be necessary to use a specific substrate or direct assaying methods (e.g., in-gel kinase assay).

Other parameters that should be considered for accurate comparison of protein kinases are: (1) starvation of the cells before activation may interfere with activation of the desired protein kinase or may cause activation of some SAPKs; (2) the optimal length of stimulation may vary from cell to cell and from one protein kinase to the other; therefore, appropriate time points for each kinase should be determined; and (3) for accurate comparison of the activities of protein kinases, detection should be performed in the linear range of the phosphorylation reaction. Thus, the amount of protein used for immunoprecipitation, the concentration of antibodies, the length of the phosphorylation reaction, and the exposure to X-ray film or to the phosphoimager should be optimized in order to reach linearity. If necessary, a standard curve with the protein kinases of interest can be made, and serial dilutions of the cytosolic extracts or a time course of the phosphorylation can be used to ensure one is working in a linear range.

Anticipated Results

In Basic Protocol 1, which describes immunoprecipitation followed by phosphorylation, extracts from nonstimulated Rat1 cells contain very little phosphorylated myelin basic protein (MBP; Fig. 18.12.2), which represents the activity of both ERK1 and ERK2 in these cells. Upon addition of EGF to the cells, the amount of phosphate incorporated into MBP should increase with time, peaking at 30 min after stimulation and declining thereafter. This pattern of phosphorylation represents the transient

activation of MAPK by EGF in these cells. When the in-gel assay is used (Basic Protocol 2), both ERK1 and ERK2 should be detected at positions corresponding to molecular weights of 44 and 42 kDa, respectively, and the kinetics of activation of both should be similar to that of the MBP phosphorylation obtained with the immunoprecipitation. JNK activation (Basic Protocol 3) should peak at 30 min and decline thereafter.

Time Considerations

After cell harvesting, the immunoprecipitation protocol requires extraction (0.5 hr), immunoprecipitation and washing (3 to 4 hr), the phosphorylation reaction (0.5 to 1.0 hr), SDS-PAGE (2.5 hr), and processing of the gel (6 to 16 hr). Since this procedure usually takes more than a single working day, it can be interrupted after boiling the samples in sample buffer or at any time during the destaining period. The in-gel kinase assay can take 2 to 3 days and the JNK assay usually takes 24 hr, including SDS-PAGE processing.

Literature Cited

- Abe, J., Kusuhara, M., Ulevitch, R.J., Berk, B.C., and Lee, J.D. 1996. Big mitogen-activated protein kinase 1 (BMK1) is a redox-sensitive kinase. *J. Biol. Chem.* 271:16586-16590.
- Ahn, N.G., Weiel, J.E., Chan, C.P., and Krebs, E.G. 1990. Identification of multiple epidermal growth factor-stimulated protein serine/threonine kinases from Swiss 3T3 cells. *J. Biol. Chem.* 265:11487-11494.
- Blank, J.L., Gerwins, P., Elliott, E.M., Sather, S., and Johnson, G.L. 1996. Molecular cloning of mitogen-activated protein/ERK kinase kinases (MEKK) 2 and 3. Regulation of sequential phosphorylation involving mitogen-activated protein kinase and c-Jun kinase. *J. Biol. Chem.* 271:5361-5368.
- Campbell, J.S., Seger, R., Graves, J.D., Graves, L.M., Jensen, A.M., and Krebs, E.G. 1995. The MAP kinase cascade. *Recent Prog. Horm. Res.* 50:131-159.
- Canagarajah, B.J., Khokhlatchev, A., Cobb, M.H., and Goldsmith, E.J. 1997. Activation mechanism of the MAP kinase ERK2 by dual phosphorylation. *Cell* 90:859-869.
- Chen, R.H., Sarnecki, C., and Blenis, J. 1992. Nuclear localization and regulation of erk- and rsk-encoded protein kinases. *Mol. Cell. Biol.* 12:915-927.
- Cohen, P., Alessi, D.R., and Cross, D.A. 1997. PDK1, one of the missing links in insulin signal transduction? *FEBS Lett.* 410:3-10.
- Crespo, P., Schuebel, K.E., Ostrom, A.A., Gutkind, J.S., and Bustelo, X.R. 1997. Phosphotyrosine-dependent activation of Rac-1 GDP/GTP ex-

change by the vav proto-oncogene product. *Nature* 385:169-172.

- Cuenda, A., Cohen, P., Buee-Scherrer, V., and Goedert, M. 1997. Activation of stress-activated protein kinase-3 (SAPK3) by cytokines and cellular stresses is mediated via SAPKK3 (MKK6); comparison of the specificities of SAPK3 and SAPK2 (RK/p38). *EMBO J.* 16:295-305.
- Davis, R.J. 1994. MAPKs: New JNK expands the group. *Trends Biochem. Sci.* 19:470-473.
- Deacon, K. and Blank, J.L. 1997. Characterization of the mitogen-activated protein kinase kinase 4 (MKK4)/c-Jun NH₂-terminal kinase 1 and MKK3/p38 pathways regulated by MEK kinases 2 and 3. MEK kinase 3 activates MKK3 but does not cause activation of p38 kinase in vivo. *J. Biol. Chem.* 272:14489-14496.
- Deak, M., Clifton, A.D., Lecocq, L.M., and Alessi, D.R. 1998. Mitogen- and stress-activated protein kinase-1 (MSK1) is directly activated by MAPK and SAPK2/p38, and may mediate activation of CREB. *EMBO J.* 17:4426-4441.
- DiDonato, J.A., Hayakawa, M., Rothwarf, D.M., Zandi, E., and Karin, M. 1997. A cytokine-responsive I κ B kinase that activates the transcription factor NF- κ B. *Nature* 388:548-554.
- Fan, G., Merritt, S.E., Kortjenann, M., Shaw, P.E., and Holzman, L.B. 1996. Dual leucine zipper-bearing kinase (DLK) activates p46SAPK and p38MAPK but not ERK2. *J. Biol. Chem.* 271:24788-24793.
- Fukunaga, R. and Hunter, T. 1997. MNK1, a new MAP kinase-activated protein kinase, isolated by a novel expression screening method for identifying protein kinase substrates. *EMBO J.* 16:1921-1933.
- Gerwins, P., Blank, J.L., and Johnson, G.L. 1997. Cloning of a novel mitogen-activated protein kinase kinase kinase, MEKK4, that selectively regulates the c-Jun amino terminal kinase pathway. *J. Biol. Chem.* 272:8288-8295.
- Goedert, M., Cuenda, A., Craxton, M., Jakes, R., and Cohen, P. 1997. Activation of the novel stress-activated protein kinase SAPK4 by cytokines and cellular stresses is mediated by SKK3 (MKK6); comparison of its substrate specificity with that of other SAP kinases. *EMBO J.* 16:3563-3571.
- Han, J., Lee, J.D., Bibbs, L., and Ulevitch, R.J. 1994. A MAP kinase targeted by endotoxin and hyperosmolarity in mammalian cells. *Science* 265:808-811.
- Han, J., Jiang, Y., Li, Z., Kravchenko, V.V., and Ulevitch, R.J. 1997. Activation of the transcription factor MEF2C by the MAP kinase p38 in inflammation. *Nature* 386:296-299.
- Hibi, M., Lin, A., Smeal, T., Minden, A., and Karin, M. 1993. Identification of an oncoprotein- and UV-responsive protein kinase that binds and potentiates the c-Jun activation domain. *Genes Dev.* 7:2135-2148.
- Holland, P.M., Suzanne, M., Campbell, J.S., Noselli, S., and Cooper, J.A. 1997. MKK7 is a stress-activated mitogen-activated protein kinase kinase functionally related to hemipterous. *J. Biol. Chem.* 272:24994-24998.
- Kameshita, I. and Fujisawa, H. 1989. In-gel kinase assay. *Anal. Biochem.* 183:139-143.
- Karin, M. 1995. The regulation of AP-1 activity by mitogen-activated protein kinases. *J. Biol. Chem.* 270:16483-16486.
- Kato, Y., Tapping, R.S., Huang, S., Watson, M.H., Ulevitch, R.J., and Lee, J.D. 1998. BMK1/ERK5 is required for cell proliferation induced by epidermal growth factor. *Nature* 395:713-716.
- Kiefer, F., Tibbles, L.A., Anafi, M., Janssen, A., Zanke, B.W., Lassam, N., Pawson, T., Woodgett, J.R., and Iscove, N.N. 1996. HPK1, a hematopoietic protein kinase activating the SAPK/JNK pathway. *EMBO J.* 15:7013-7025.
- Kramer, R.M., Roberts, E.F., Um, S.L., Borsch-Haubold, A.G., Watson, S.P., Fisher, M.J., and Jakubowski, J.A. 1996. p38 mitogen-activated protein kinase phosphorylates cytosolic phospholipase A₂ (cPLA₂) in thrombin-stimulated platelets. Evidence that proline-directed phosphorylation is not required for mobilization of arachidonic acid by cPLA₂. *J. Biol. Chem.* 271:27723-27729.
- Kyriakis, J.M. and Avruch, J. 1996. Sounding the alarm: Protein kinase cascades activated by stress and inflammation. *J. Biol. Chem.* 271:24313-24316.
- Kyriakis, J.M., Banerjee, P., Nikolakaki, E., Dai, T., Rubie, E.A., Ahmad, M.F., Avruch, J., and Woodgett, J.R. 1994. The stress-activated protein kinase subfamily of c-Jun kinases. *Nature* 369:156-160.
- Leung, T., Manser, E., Tan, L., and Lim, L. 1995. A novel serine/threonine kinase binding the Ras-related RhoA GTPase which translocates the kinase to peripheral membranes. *J. Biol. Chem.* 270:29051-29054.
- Lin, L.L., Wartmann, M., Lin, A.Y., Knopf, J.L., Seth, A., and Davis, R.J. 1993. cPLA₂ is phosphorylated and activated by MAP kinase. *Cell* 72:269-278.
- Lu, X., Nemoto, S., and Lin, A. 1997. Identification of c-Jun NH₂-terminal protein kinase (JNK)-activating kinase 2 as an activator of JNK but not p38. *J. Biol. Chem.* 272:24751-24754.
- Marais, R., Wynne, J., and Treisman, R. 1993. The SRF accessory protein Elk-1 contains a growth factor-regulated transcriptional activation domain. *Cell* 73:381-393.
- Marshall, C.J. 1996. Ras effectors. *Curr. Opin. Cell Biol.* 8:197-204.
- McLaughlin, M.M., Kumar, S., McDonnell, P.C., Van Horn, S., Lee, J.C., Livi, G.P., and Young, P.R. 1996. Identification of mitogen-activated protein (MAP) kinase-activated protein kinase-3, a novel substrate of CSBP p38 MAP kinase. *J. Biol. Chem.* 271:8488-8492.
- Payne, D.M., Rossomando, A.J., Martino, P., Erickson, A.K., Her, J.-H., Shabanowitz, J., Hunt, D.F., Weber, M.J., and Sturgill, T.W. 1991. Iden-

- tification of the regulatory phosphorylation sites in pp42/mitogen-activated protein kinase (MAP kinase). *EMBO J.* 10:885-892.
- Pombo, C.M., Kehrl, J.H., Sanchez, I., Katz, P., Avruch, J., Zon, L.I., Woodgett, J.R., Force, T., and Kyriakis, J.M. 1995. Activation of the SAPK pathway by the human STE20 homologue germinal centre kinase. *Nature* 377:750-754.
- Regnier, C.H., Song, H.Y., Gao, X., Goeddel, D.V., Cao, Z., and Rothe, M. 1997. Identification and characterization of an I κ B kinase. *Cell* 90:373-383.
- Robinson, M.J. and Cobb, M.H. 1997. Mitogen-activated protein kinase pathways. *Curr. Opin. Cell Biol.* 9:180-186.
- Salmeron, A., Ahmad, T.B., Carlile, G.W., Pappin, D., Narsimhan, R.P., and Ley, S.C. 1996. Activation of MEK-1 and SEK-1 by Tpl-2 proto-oncogene, a novel MAP kinase kinase kinase. *EMBO J.* 15:817-826.
- Seger, R. and Krebs, E.G. 1995. The MAPK signaling cascade. *FASEB J.* 9:726-735.
- Seger, R., Seger, D., Reszka, A.A., Munar, E.S., Eldar-Finkelman, H., Dobrowolska, G., Jensen, A.M., Campbell, J.S., Fischer, E.H., and Krebs, E.G. 1994. Over-expression of mitogen-activated protein kinase kinase (MAPKK) and its mutants in NIH-3T3 cells: Evidence that MAPKK's involvement in cellular proliferation is regulated by phosphorylation of serine residues in its kinase subdomains VII and VIII. *J. Biol. Chem.* 269:29876-29886.
- Stokoe, D., Campbell, D.G., Nakielny, S., Hidaka, H., Leever, S.J., Marshall, C., and Cohen, P. 1992. MAPKAP kinase-2, a novel protein kinase activated by mitogen-activated protein kinase. *EMBO J.* 11:3985-3994.
- Takekawa M., Posas, F., and Saito, H. 1997. A human homolog of the yeast Ssk2/Ssk22 MAP kinase kinase kinases, MTK1, mediates stress-induced activation of the p38 and JNK pathways. *EMBO J.* 16:4973-4982.
- Tournier, C., Whitmarsh, A.J., Cavanagh, J., Barrett, T., and Davis, R.J. 1997. Mitogen-activated protein kinase kinase 7 is an activator of the c-Jun NH₂-terminal kinase. *Proc. Natl. Acad. Sci. U.S.A.* 94:7337-7342.
- Wang, X.Z. and Ron, D. 1996. Stress-induced phosphorylation and activation of the transcription factor CHOP (GADD153) by p38 MAP kinase. *Science* 272:1347-1349.
- Wang, X.S., Diener, K., Jannuzzi, D., Trollinger, D., Tan, T.H., Lichenstein, H., Zukowski, M., and Yao, Z. 1996. Molecular cloning and characterization of a novel protein kinase with a catalytic domain homologous to mitogen-activated protein kinase kinase kinase. *J. Biol. Chem.* 271:31607-31611.
- Wang, W., Zhou, G., Hu, M.C.T., Yao, Z., and Tan, T.H. 1997. Activation of the hematopoietic progenitor kinase-1 (HPK1)-dependent, stress-activated c-Jun N-terminal kinase (JNK) pathway by transforming growth factor β (TGF- β)-activated kinase (TAK1), a kinase mediator of TGF- β signal transduction. *J. Biol. Chem.* 272:22771-22775.
- Waskiewicz, A.J., Flynn, A., Proud, C.G., and Cooper, J.A. 1997. Mitogen-activated protein kinases activate the serine/threonine kinases MNK1 and MNK2. *EMBO J.* 16:1909-1920.
- Woodgett, J.R., Avruch, J., and Kyriakis, J. 1996. The stress-activated protein kinase pathway. *Cancer Surv.* 27:127-138.
- Yamaguchi, K., Shirakabe, K., Shibuya, H., Irie, K., Oishi, I., Ueno, N., Taniguchi, T., Nishida, E., and Matsumoto, K. 1995. Identification of a member of the MAPKKK family as a potential mediator of TGF- β signal transduction. *Science* 270:2008-2011.
- Yan, M., Dai, T., Deak, J.C., Kyriakis, J.M., Zon, L.I., Woodgett, J.R., and Templeton, D.J. 1994. Activation of stress-activated protein kinase by MEKK1 phosphorylation of its activator SEK1. *Nature* 372:798-800.
- Yung, Y., Dolginov, Y., Yao, Z., Rubinfeld, H., Michael, D., Hanoch, T., Roubini, E., Lando, Z., Zharhary, D., and Seger, R. 1997. Detection of ERK activation by a novel monoclonal antibody. *FEBS Lett.* 408:292-296.
- Zhang, S., Han, J., Sells, M.A., Chernoff, J., Knaus, U.G., Ulevitch, R.J., and Bokoch, G.M. 1995. Rho family GTPases regulate p38 mitogen-activated protein kinase through the downstream mediator PAK1. *J. Biol. Chem.* 270:23934-23936.
- Zhou, G., Bao, Z.Q., and Dixon, J.E. 1995. Components of a new human protein kinase signal transduction pathway. *J. Biol. Chem.* 270:12665-12669.

Contributed by Hadara Rubinfeld and Rony Seger
The Weizmann Institute of Science
Rehovot, Israel

Isolation of Phosphopeptides by Immobilized Metal Ion Affinity Chromatography

UNIT 18.13

This unit describes mass spectrometric (MS) analysis of phosphopeptides and different approaches to enrich the phosphopeptides from complex protein samples prior to MS analysis. Basic Protocol 1 involves manual or automated prefractionation of peptides by cation-exchange chromatography followed by immobilized metal ion affinity chromatography (IMAC) using commercially available spin columns. It is suitable when larger amounts of complex material are available (at least 100 μg), and can be performed in any molecular biology laboratory because it requires no specialized chromatographic or mass spectrometric equipment. The eluted samples can be sent for MS analysis at a suitable facility. The Alternate Protocol describes the use of anion-exchange chromatography and IMAC. Basic Protocol 2 describes an online IMAC method optimized for highly selective and sensitive detection of phosphopeptides from complex mixtures using methyl esterification of peptides. Basic Protocol 3 describes a new method for analysis of tyrosine phosphorylation from complex mixtures using IMAC. Prefractionation of samples by multidimensional chromatography at either the peptide or protein level considerably increases the number of identified phosphopeptides for Basic Protocols 1 and 2.

MANUAL PEPTIDE FRACTIONATION BY CATION-EXCHANGE CHROMATOGRAPHY AND IMAC

**BASIC
PROTOCOL 1**

This protocol describes the fractionation of a complex peptide mixture by cation-exchange chromatography, followed by immobilized metal ion affinity chromatography (IMAC) of each individual fraction. No specialized large equipment (such as FPLC or HPLC) is needed, and the procedure, up to the elution of the purified phosphopeptides, can be performed in any molecular biology laboratory.

The first consideration is the preservation of the *in vivo* phosphorylation state in the cell extract. Total soluble proteins can be extracted under denaturing conditions to prevent dephosphorylation, but if subcellular fractionation is performed, native buffers with generous amounts of appropriate phosphatase inhibitors must be used (e.g., EDTA, sodium pyrophosphate, sodium fluoride, β -glycerophosphate, and calyculin A for serine/threonine phosphatases; sodium orthovanadate for tyrosine phosphatases). It is also worth considering the inhibition of protein kinases after breaking the cells to avoid nonspecific *in vitro* phosphorylation in the extract; addition of EDTA may be sufficient to block kinase activity by chelating magnesium, but more specific inhibitors like K-252a can also be used.

The use of volatile buffers in all steps from digestion to IMAC has the considerable advantage of preventing sample losses through the desalting steps (Gruhler et al., 2005). Triethylammonium bicarbonate (TEAB) is more volatile than ammonium bicarbonate and is preferred for both trypsin digestion and anion-exchange chromatography (see Alternate Protocol). The salt component in cation-exchange chromatography is ammonium formate. IMAC is very sensitive to salt, and ammonium formate is easily removed from the eluted fractions by lyophilization or by concentration in a Speedvac evaporator.

**Analysis of
Protein
Phosphorylation**

18.13.1

Peptide trap cartridges are offered by a variety of vendors and allow manual sample preparation with a syringe. Optimize Technologies offers cation-exchange cartridges (Opti-Lynx) with different capacities and a holder with fittings. A high-capacity cartridge of this type is also supplied with Applied Biosystem's iTRAQ kit. This protocol using cation exchange is compatible with iTRAQ labeling for quantitative experiments. It is an adaptation of the manufacturer's protocol for manual sample cleanup after labeling with the iTRAQ reagent (an amine-labeling reagent for multiplexed relative and absolute protein quantitation). The method described here uses the ion-exchange cartridge supplied by the manufacturer with the solvents described in Gruhler et al. (2005). This provides a simple manual setup that requires little specialized equipment, but for greater reproducibility of fractionation and greater loading capacity (milligram range), the use of an HPLC or FPLC system with an SCX column is recommended (see Beausoleil et al., 2004; Gruhler et al., 2005).

A variety of IMAC materials have become commercially available, including different combinations of metals and chelating resins and options for batch or spin-column chromatography. In side-by-side tests, these were at least as good as hand-made micro-columns packed with Fe³⁺-charged POROS-MC resin (Nühse et al., 2003). For ease of use and reproducibility, this protocol uses Phos-Select IMAC resin (Sigma).

This protocol is universally applicable to complex soluble or membrane protein samples. In the latter case, washing the membrane preparation with a high-pH buffer removes more soluble proteins and leads to more peptide identifications from integral membrane proteins.

Materials

- Protein sample (100 to 500 µg)
- 100 mM sodium carbonate (Na₂CO₃)
- 50 and 500 mM triethylammonium bicarbonate (TEAB) in ddH₂O, diluted from 1 M stock solution (Sigma)
- Sequencing-grade modified trypsin
- 2.5% (v/v) formic acid/30% (v/v) acetonitrile
- Cation-exchange cartridge (e.g., Optimize Technologies Opti-Lynx or Applied Biosystems iTRAQ), with holder and syringe adapter
- SCX start buffer (see recipe)
- SCX step elution buffers (10, 20, 30, 40, 60, and 100 mM salt; see recipe)
- SCX cleaning buffer (see recipe)
- Sodium azide, optional
- Phos-Select IMAC resin (Sigma)
- IMAC load/wash buffer (see recipe)
- IMAC elution buffer (see recipe)
- 10% (v/v) formic acid
- 11 × 34-mm polycarbonate ultracentrifuge tubes
- Ultracentrifuge (e.g., Beckman Optima tabletop ultracentrifuge with a TLA120 rotor)
- 37°C water bath
- Lyophilizer or Speedvac evaporator
- 1.5- and 2-ml microcentrifuge tubes
- End-over-end rotator
- Fritted spin columns (e.g., Mobicols, MoBiTec; Handee Mini Spin columns, Pierce)

Digest sample

For insoluble, particulate, or membrane protein samples

- 1a. Resuspend protein/membrane pellet in 100 mM Na₂CO₃ in a 11 × 34-mm polycarbonate ultracentrifuge tube and incubate 10 min on ice to remove peripheral proteins. Ultracentrifuge 30 min at 100,000 × g, 4°C.

This step is only necessary when there is a particular focus on integral membrane proteins. Peripheral membrane proteins can make up ≥50% of a microsomal pellet and are efficiently stripped off the membrane with the alkaline carbonate solution. At the same time, integral membrane proteins are partially denatured.

- 2a. Resuspend the pellet in 0.8 ml of 500 mM TEAB and centrifuge again 30 min at 100,000 × g, 4°C.
- 3a. Resuspend the pellet in 0.2 ml of 50 mM TEAB and measure the protein concentration.

Ammonium bicarbonate is the traditional buffer for digestion and can be used, but TEAB evaporates much more readily.

- 4a. Start with 100 to 500 μg protein and add sequencing-grade modified trypsin at a trypsin/protein ratio of 1:50 to 1:20 (2 to 25 μg trypsin). Digest overnight at 37°C.

For larger ion-exchange columns with a higher capacity, up to several milligrams of protein can be digested.

Sequencing-grade modified trypsin is used to avoid the low chymotrypsin-like activity that is present in unmodified trypsin.

The digested sample can be stored for several weeks at -70°C if required, but some sample loss may occur.

- 5a. Lyophilize or concentrate in a Speedvac evaporator to 10 to 20 μl. Proceed to step 6.

For soluble protein samples

- 1b. Digest proteins as described in Basic Protocol 2, steps 1c to 6c, and then desalt according to Basic Protocol 2, steps 12 to 17. Proceed to step 6.

Perform cation-exchange chromatography

6. Redissolve the dried peptide sample in 0.5 ml of 2.5% formic acid/30% acetonitrile.

The sample should ideally have zero salt and be at pH ≤2.7. Since ammonium formate does not buffer at this pH, the peptides are loaded in dilute acid rather than the start buffer.

7. Centrifuge 10 min at 15,000 × g, room temperature, to remove insoluble material.
8. Prepare the cation-exchange cartridge by washing at ~1 ml/min with the following:

2 ml SCX start buffer
2 ml SCX cleaning buffer
2 ml SCX start buffer
1 ml 2.5% formic acid/30% acetonitrile.

9. Slowly (~200 μl/min) load the cleared sample onto the cartridge. Collect flowthrough, if desired.

In the authors' experience, few phosphopeptides are recovered from the unbound fraction, but this may vary in different types of samples.

10. Wash with 2 ml wash buffer. Elute with 0.5 ml each of 10, 20, 30, 40, 60, and 100 mM SCX step elution buffers for highly complex samples, or with 25, 50, and 100 mM SCX step elution buffers for less complex samples.

11. Lyophilize fractions or concentrate in a Speedvac evaporator to ~20 μ l.

Lyophilization is easier if the fractions are first concentrated in the Speedvac to half of the original volume, because the acetonitrile lowers the freezing point of the liquid. Thawed samples will “boil over” in the vacuum of the lyophilizer.

12. Clean cartridge with 2 ml cleaning buffer and then re-equilibrate with 2 ml start buffer or with 2 ml start buffer plus 3 mM sodium azide for long-term storage.

Perform IMAC

13. Resuspend Phos-Select IMAC resin to homogeneity.

14. Pipet an aliquot of the 50% slurry (40 to 80 μ l per SCX fraction per IMAC experiment) into a 2-ml microcentrifuge tube. Centrifuge 2 min at $500 \times g$, room temperature. Remove as much of the manufacturer's glycerol-containing storage buffer as possible.

15. Add 1 ml IMAC load/wash buffer, mix by inverting a few times, centrifuge again, and remove wash buffer. Repeat at least three times.

16. Distribute equal aliquots of the slurry into tubes for each SCX fraction to be analyzed.

17. Redissolve each concentrated or dried peptide fraction from ion-exchange chromatography in 0.3 ml IMAC load/wash buffer. If insoluble material remains, clear samples by centrifuging 5 min at $15,000 \times g$, room temperature.

18. Add dissolved peptide fractions to the tubes with Phos-Select resin. Incubate by end-over-end rotation for 1 to 2 hr at room temperature.

19. Prepare a fritted spin column with a 2-ml collection tube for each sample. Flick down liquid droplets and resin from the lid of the tube, transfer the complete slurry to the spin columns, and centrifuge 30 sec at $8200 \times g$, room temperature. Collect flowthrough (nonphosphorylated peptides) if desired.

20. Add 200 μ l IMAC load/wash buffer to the beads and resuspend by pipetting the slurry up and down. Centrifuge 30 sec at $8200 \times g$, room temperature.

Unlike with prepacked spin columns, it is not sufficient to simply add the wash solution to the top of the tube. The beads should be resuspended by pipetting up and down.

21. Wash one additional time with IMAC load/wash buffer and then one time with ddH₂O.

22. Transfer spin columns to clean 1.5-ml microcentrifuge tubes.

23. Add 30 μ l IMAC elution buffer, resuspend well, and incubate 1 to 2 min at room temperature. Centrifuge 10 sec at $8200 \times g$ and collect the eluate.

24. Repeat elution with IMAC elution buffer and combine eluates.

Take care not to leave the Phos-Select resin in the alkaline buffer for longer than ~5 min total.

25. Concentrate eluates in the Speedvac to ~10 to 20 μ l and acidify with 1 to 2 μ l of 10% formic acid.

The samples are ready for mass spectrometric analysis. Rapid sample losses due to adsorption occur if the tubes are kept at room temperature or 4°C, but it is possible to store them at -70°C for a short time (up to several days).

MANUAL PEPTIDE FRACTIONATION BY ANION-EXCHANGE CHROMATOGRAPHY AND IMAC

ALTERNATE
PROTOCOL

This protocol uses a microcolumn setup with POROS HQ anion-exchange resin and a volatile buffer system at alkaline pH. If the sample digest has been performed in dilute TEAB without urea or other additives (as in the case of membrane or insoluble proteins, which can be washed in trypsin buffer after denaturing as described in Basic Protocol 1), the digest can be applied directly to the column after adding acetonitrile and pelleting the insoluble material.

Additional Materials (also see Basic Protocol 1)

POROS HQ strong anion-exchange resin
SAX start buffer (see recipe)
SAX step buffer (see recipe)
SAX step elution buffers (50, 100, 200, and 500 mM; see recipe)
GELoader tips (Eppendorf)

1. Prepare POROS HQ resin by washing ~100 μ l of settled material in 1 ml SAX start buffer.
2. Prepare a microcolumn in a GELoader tip (see Support Protocol 1). Load the slurry into the prepared column to a packed height of ~2 to 3 cm in the wider part of the tip (i.e., 3 to 10 mm below the wide neck).
3. Wash with 50 μ l each of SAX start buffer, SAX step buffer, and again SAX start buffer.
4. Prepare the sample digest in SAX start buffer: if the digest is in trypsin buffer, by adding 0.75 vol each ddH₂O and acetonitrile; if it is in buffers containing urea or salts, by reversed-phase cleanup (see Basic Protocol 2, steps 12 to 14; use 500 μ l of 20 mM TEAB/50% acetonitrile for elution).
5. Slowly load the sample onto the microcolumn with a syringe and adapter (see Support Protocol 1).
6. Wash three times with 50 μ l SAX start buffer.
7. Elute with 50 μ l SAX step elution buffers (50, 100, 200, and 500 mM salt).

The number of steps depends on the complexity of the sample; most phosphopeptides elute between 100 and 200 mM salt.

8. Lyophilize the eluates as in Basic Protocol 1, step 11.

MANUAL CHROMATOGRAPHY WITH MICROCOLUMNS

SUPPORT
PROTOCOL 1

Conventional sample preparation methods are often not appropriate for the microgram to nanogram amounts of material available for MS experiments. Gobom et al. (1999) introduced a manual microchromatographic method for reversed-phase (RP) purification and concentration of peptides for analysis by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). The RP material is packed into the narrow capillary tip of a gel-loading pipet tip with a constricted end. A column of 2 to 3 cm packing height has a volume of only a few microliters, and thus reduces sample loss from nonspecific adsorption. This format is very versatile, and virtually any chromatographic material with a bead size of ~20 μ m or larger can be used. To create columns with slightly higher capacity, the wider-bore part of the pipet tip can also be packed to a volume of up to ~50 μ l (e.g., for ion-exchange separation of several hundred micrograms of digest). The disadvantage of this method is that preparation of multiple samples is time-consuming and most likely not as reproducible as automated approaches.

Analysis of
Protein
Phosphorylation

18.13.5

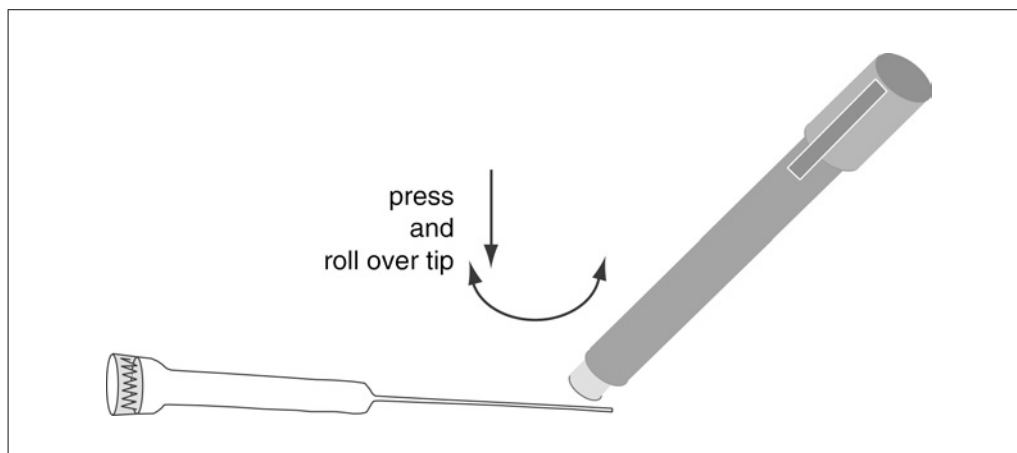


Figure 18.13.1 Constriction of a GELoader tip for use as a microcolumn.

The only pipet tips that work satisfactorily for this technique are original Eppendorf GELoader tips; the capillaries are soft enough to be constricted so they retain the chromatography material. A 1-ml plastic Luer-type syringe is also needed, and a 200- μ l pipet tip trimmed on both ends with a razor blade is used as a tight adapter between the Luer syringe and the GELoader tip.

To constrict the GELoader tip, lay the tip flat on the benchtop and press with a semi-blunt plastic item (e.g., the back end of a pen; Fig. 18.13.1) at a point \sim 1 to 2 mm away from the end of the fine capillary. Use considerable pressure, but be careful not to sever the tip off. Pipet 10 μ l acetonitrile into the tip and hold upright. If the tip is sufficiently compressed, the drop will not move or drop out. If it does, the tip is not narrow enough to hold the POROS beads; try again with a fresh tip. Fit the syringe and adapter onto the tip and apply full pressure. Ideally, there should be resistance at first, and then the liquid should come out. It should now be possible to load the slurry into the tip.

Pipet the slurry, sample, wash, and elution buffers as far down into the tip as possible and then flick the tip down with a quick wrist motion. Fit the syringe onto the tip and press the liquid through. Depending on the backpressure from the medium, full compression of a 1-ml syringe may be necessary. For larger column volumes, take care not to run the material dry. Microcolumns packed only into the capillary end of the tip are less sensitive to drying out briefly, provided no chromatographic resolution is required.

**BASIC
PROTOCOL 2**

**ENRICHMENT OF PHOSHOPEPTIDES THROUGH METHYL ESTER
FORMATION AND IMAC**

Background binding of peptides containing a high percentage of acidic amino acids to the IMAC column can lead to significant contamination of the enriched phosphopeptide samples, and thus difficulty in applying IMAC alone to satisfactorily enrich phosphopeptides from complex cellular lysates (Brill et al., 2004). This protocol describes IMAC enrichment of phosphopeptide methyl esters for increased IMAC selectivity (Ficarro et al., 2002, 2005; Salomon et al., 2003; Brill et al., 2004; Ndassa et al., 2006). Although this procedure requires the use of an HPLC pump and practice working with capillary columns, it provides unparalleled sensitivity and efficiency due to the highly selective enrichment of phosphopeptides from complex cell-derived samples. All separations and sample transfers occur column-to-column (“online”) without exposure to plastic, including microcentrifuge tubes or 96-well plates, thereby minimizing peptide loss. Sensitivity arises from the use of capillary chromatography coupled with nanospray mass spectrometry at flow rates approaching 20 nl/min. This protocol describes preparation of the phosphopeptide methyl esters, IMAC enrichment, and MS analysis. The

**Isolation of
Phosphopeptides
by IMAC**

18.13.6

IMAC enrichment procedure is performed with the use of a pressure bomb and a total of four custom-made capillary columns. This IMAC protocol incorporates newly developed improvements in IMAC buffer solutions, is optimized for reduced competition on the IMAC column between singly- and multiply-phosphorylated peptides, and improves overall yield of phosphopeptides (Ndassa et al., 2006).

Protocols for elution of phosphopeptides from the precolumn over the analytical column and into the mass spectrometer will vary depending on the configuration and type of chromatography and mass spectrometry equipment available. The method described here works well for the LTQ, LTQ-FTICR, or LTQ-Orbitrap and Agilent 1100 quaternary pump.

Materials

Protein sample (1×10^8 cells or ~ 100 fmol purified protein)
Cell lysis buffer A (see recipe), optional
Immunoprecipitation resin: anti-phosphotyrosine agarose (e.g., clone PT66, Sigma), optional
Immunoprecipitation wash buffer (see recipe), optional
Trypsin digestion buffer with urea (see recipe)
Milli-Q water
Sequencing-grade modified trypsin (Promega)
Dialysis buffer (see recipe), optional
Trizol LS (Invitrogen), optional
DC protein assay kit (Bio-Rad) containing:
 Alkaline copper tartrate solution
 Folin reagent
 Surfactant solution
 Bovine gamma-globulin standard
Angiotensin II phosphate (EMD Biosciences) or other synthetic phosphopeptide
1 M HCl
Reversed-phase loading buffer: 0.1% acetic acid in Milli-Q water
Reversed-phase elution buffer (see recipe)
Macrotrap wash buffer 1: 20:80 (v/v) dimethylsulfoxide (DMSO)/methanol
Macrotrap wash buffer 2: 45:45:10 (v/v/v) isopropanol/water/formic acid
Methylation solution (see recipe)
Nitrogen source
Fe(III)-activated IMAC column (see Support Protocol 2)
IMAC preparation buffer 1: 50 mM EDTA, pH 8
IMAC preparation buffer 2: 100 mM FeCl₃ (99.99+%, Sigma)
Desalting column (see Support Protocol 2)
HPLC buffer A (IMAC): 0.01% acetic acid in Milli-Q ddH₂O
HPLC buffer B (IMAC): 0.01% acetic acid in acetonitrile (99.8%; HPLC grade)
IMAC wash buffer (see recipe)
Precolumn (see Support Protocol 2)
IMAC elution buffer: 25 mM Na₂HPO₄, pH 9.0
Analytical column (see Support Protocol 2)
HPLC buffer C (reversed-phase): 0.1 M acetic acid in Milli-Q ddH₂O
HPLC buffer D (reversed-phase): 0.1 M acetic acid in acetonitrile (99.8%; HPLC grade)

1-ml syringes and 27-G needles
50-ml conical tubes
1.5-ml microcentrifuge tubes
95°C heating block
0.45- μ m, 4-mm syringe filters (e.g., Millipore Millex), optional

200- and 1000- μ l pipets (Rainin)
10,000 MWCO Slide-A-Lyzer (Pierce), optional
Reversed-phase C18 peptide macrotrap (Michrom Bioresources)
1.5-ml screw-top microcentrifuge tubes (Sarstedt)
Speedvac evaporator (Thermo Savant)
Inline microfilters (Upchurch Scientific)
HPLC pump (e.g., Agilent 1100 Quaternary Pump)
Fused silica capillaries (360- μ m o.d. \times 50-, 75-, and 200- μ m i.d.; Polymicro Technologies)
Microtee (Upchurch Scientific)
Laser Puller P2000 (Sutter Instruments), optional
Mass spectrometer (e.g., LTQ-FTICR and LTQ available from Thermo)
Valco trap holder (Michrom Bioresources)
Multiple-syringe infusion pump
500- μ l Hamilton syringe
Platinum wire (Sigma)

NOTE: All sample preparation procedures involving intact proteins are carried out at 4°C unless otherwise specified.

Prepare sample

For mammalian cell phosphorylation analysis

- 1a. Lyse cells in 25 ml cell lysis buffer A per 1×10^9 cells.
- 2a. Centrifuge 20 min at $12,000 \times g$, 4°C, and collect supernatant (total cellular lysate).
- 3a. Store samples, if necessary, at -80°C . Proceed to step 11.

For an individual protein or a subset of total cellular phosphorylated proteins

- 1b. Immunoprecipitate phosphorylated proteins from the total cellular lysate using CNBr-agarose-conjugated antibodies. Add 200 μ l of immunoprecipitation resin per 1×10^9 cell equivalents in 25 ml (in a 50-ml tube) and incubate 4 hr with tumbling.

CNBr linkage is preferable for coupling antibodies to agarose beads compared to noncovalent linkages such as protein A/G due to its resistance to the urea used in the elution of proteins in this protocol, and the high density of antibody on the resin resulting in reduced amounts of resin. For isolation of the tyrosine phosphoproteome, anti-phosphotyrosine agarose clone PT66 (Sigma) works well. Clone 4G10 works very poorly for immunoprecipitation of native proteins. Tyrosine phosphorylation represents only a small fraction (~0.1%) of cellular phosphorylation, necessitating an additional affinity enrichment step before IMAC to enrich the population of tyrosine phosphorylated proteins relative to serine and threonine phosphorylated proteins. For serine and threonine phosphorylation, no pre-enrichment step is necessary before IMAC.

- 2b. Wash beads once with 50 ml cell lysis buffer A and once with 50 ml immunoprecipitation wash buffer in a 50-ml conical tube. Centrifuge 5 min at $1700 \times g$, 4°C, and then aspirate the supernatant.

It is critical to be thorough in the removal of the wash buffer, because detergent (Triton X-100) will significantly contaminate the IMAC and reversed-phase columns, resulting in high background detergent peaks in the MS spectra and dramatic reduction of the number of phosphorylation sites identified.

- 3b. Add 1 ml of immunoprecipitation wash buffer to beads and transfer beads into a 1.5-ml microcentrifuge tube using a 1000- μ l pipet tip with 3 mm of the tip cut off.
- 4b. Insert a 27-G needle attached to a 1-ml syringe along the side of the tube and withdraw residual wash buffer to thoroughly dry beads.

- 5b. Elute immunoprecipitated proteins by incubating with 200 μ l trypsin digestion buffer with urea in a 1.5-ml microcentrifuge tube for 5 min at 95°C.
- 6b. Repeat elution with another 200 μ l trypsin digestion buffer with urea for 5 min at room temperature, without additional heating. Insert needle along the side and to the bottom of the tube to withdraw all of the supernatant, being careful not to contaminate with resin.
- 7b. (Optional) Filter the eluted proteins with 0.45- μ m, 4-mm syringe filters into 1.5-ml microcentrifuge tubes to remove any residual resin.
- 8b. Measure volume of eluant with a 200- μ l pipet and dilute precisely 1:1 with Milli-Q water for a final urea concentration of 4 M.
- 9b. Add 10 μ g sequencing-grade modified trypsin per 1×10^9 immunoprecipitated cell equivalents in a 1.5-ml microcentrifuge tube and incubate overnight at 37°C.
Do not reduce and alkylate proteins prior to trypsin digestion and IMAC enrichment because this frequently reduces the number of phosphorylation sites discovered.
- 10b. Store trypsin-digested samples, if necessary, at -80°C. Proceed to step 11.

For total cellular phosphorylation

- 1c. Dialyze total cellular lysate 1:1000 (v/v) in dialysis buffer in a 10,000 MWCO Slide-A-Lyzer three times, 3 hr each time, at 4°C.
- 2c. Extract proteins by Trizol LS according to manufacturer's instructions.
- 3c. Solubilize the Trizol pellet in trypsin digestion buffer with urea for 5 min at 95°C.
- 4c. Dilute 1:1 with Milli-Q water for a final urea concentration of 4 M.
- 5c. Measure the protein concentration using the Bio-Rad DC protein assay kit according to the manufacturer's instructions.
- 6c. Add sequencing-grade modified trypsin to the protein at a ratio of 1:20 (w/w) trypsin/total protein and incubate overnight at 37°C.
Do not reduce and alkylate proteins prior to trypsin digestion and IMAC enrichment because this frequently reduces the number of phosphorylation sites discovered.
- 7c. Store trypsin-digested samples, if necessary, at -80°C. Proceed to step 11.

Prepare peptide methyl esters

11. Add 10 pmol of angiotensin II phosphate or another synthetic phosphopeptide per 1×10^9 cell equivalents to digested peptides as a control for methylation.
Many synthetic tyrosine phosphorylated peptides are available from Anaspec.
12. Acidify peptide solution to pH 3 with 1 M HCl. Load sample on a reversed-phase C18 peptide macrotrap (placed in a Valco trap holder) that has been pre-equilibrated by washing with 500 μ l of reversed-phase loading buffer (delivered by a 500- μ l Hamilton syringe attached to a multiple syringe infusion pump).
13. Wash macrotrap with 1 ml of reversed-phase loading buffer.
This step removes urea and salts from samples.
14. Elute peptides with 500 μ l of reversed-phase elution buffer into 1.5-ml screw-top microcentrifuge tubes.
Loss of peptides due to nonspecific absorption to tubes and plates is a major concern. In the authors' experience, Sarstedt tubes provide the lowest nonspecific peptide binding and should be used for macrotrap elution and methylation.

15. Regenerate macrotrap columns by washing with 500 μl of macrotrap wash buffer 1 followed by 500 μl of macrotrap wash buffer 2.

Carryover of peptides on C18 columns is a persistent concern that must be monitored carefully. This regeneration procedure is designed to minimize this carryover and prevent contamination of subsequent samples.

16. Reequilibrate macrotrap with 500 μl of reversed-phase loading buffer and repeat steps 12 through 16 until all samples are processed.

17. Completely dry peptides in a Speedvac evaporator without heat for 2 hr with rotation.

It is imperative that all water be removed at this point to optimize the yield of the methylation procedure.

18. Mix 400 μl of methylation solution with each dried peptide sample and layer nitrogen on top of the reaction prior to closing cap.

19. Mix on vortex mixer for 5 sec and allow reaction to proceed for 2 hr.

20. Dry methylated peptides for 2 hr in a Speedvac evaporator without heat and with rotation. Make sure methylation solution is fully removed after 2 hr.

21. Store dried, methylated peptides, if desired, at -80°C .

Enrich peptide methyl esters by IMAC

22. Prepare an Fe(III)-activated IMAC column by washing with IMAC preparation buffer 1, then IMAC preparation buffer 2, and finally reversed-phase loading buffer for 5 min each at a rate of 20 $\mu\text{l}/\text{min}$.

23. Connect an inline microfilter as a frit to the outlet of a desalting column and equilibrate it with reversed-phase loading buffer.

For elution of peptides from the desalting column to the IMAC column, an inline microfilter is used as a union to attach the columns in series.

24. Reconstitute the dried, methylated peptides with reversed-phase loading buffer and immediately load on the desalting column.

Typical reconstitution volumes are from 10 to 100 μl , but the actual volume is not critical. Peptides should only be reconstituted immediately before IMAC enrichment.

25. Wash column for 10 min at 7 $\mu\text{l}/\text{min}$ with reversed-phase loading buffer.

Column flow rates are not absolutely critical, but may be controlled by adjusting the air pressure to the pressure bomb and measuring the flow rate using a 5- μl measured capillary and a stopwatch (500 psi is a good starting point).

26. Connect the inlet of the IMAC column to the outlet of the desalting column using the inline microfilter already attached.

27. Attach the inlet of the desalting/IMAC column pair to the outlet of the HPLC pump. Elute to the activated IMAC column by washing 20 min at 1.8 $\mu\text{l}/\text{min}$ with reversed-phase elution buffer or using a 17-min HPLC gradient of 0% to 70% HPLC buffer B at 1.8 $\mu\text{l}/\text{min}$.

A step-elution method would not require use of an HPLC pump, but reveals approximately half the number of phosphopeptide species from a complex cell lysate.

28. Remove IMAC column from desalting column. Wash IMAC column with IMAC wash buffer for 5 min followed by reversed-phase loading buffer for 2.5 min, both at 20 $\mu\text{l}/\text{min}$.

29. Connect precolumn to IMAC column through the inline microfilter.

30. Elute with 12 μl IMAC elution buffer at 2 $\mu\text{l}/\text{min}$.
31. Wash precolumn with 10 μl HPLC buffer A at 5 $\mu\text{l}/\text{min}$.

Analyze by MS

32. Connect the washed precolumn to the analytical column through an inline microtee with one side coupled to a platinum wire for the application of electrospray voltage.
33. Use the “vented column” approach to control flow rate through the precolumn and analytical column. Connect an empty 50- μm capillary (called a restrictor) to a microtee junction between the incoming HPLC capillary and the precolumn. Configure the flow rate through the precolumn and analytical column by adjusting the length of the restrictor such that 0.25 ml/min HPLC pump output produces \sim 200 nl/min flow rate at the electrospray tip.

This configuration allows rapid delivery of the gradient to the precolumn and is preferable to nanoHPLC, which is more expensive and results in long delays for gradient changes to be conveyed to the precolumn and analytical column.

34. Elute peptides into the analytical column with a 30-min gradient from 0% to 70% HPLC buffer D.

Optionally, the flow rate may be reduced during the peptide elution part of the gradient by adjusting the flow rate at the HPLC pump down to 0.05 ml/min. This reduction in flow rate, called peak parking, is useful in increasing the sensitivity for detection of peptide ions and spreading out the peptides so that the time to acquire MS and MS/MS data for individual peptides is increased.

35. Acquire spectra with data-dependent scanning (one MS scan followed by five MS/MS scans of the most abundant ions). Set collision energy at 35% with a 3-Da window, and set dynamic exclusion at a repeat count of one and an exclusion time of 1.5 min.

Dynamic exclusion prevents the redundant acquisition of MS/MS spectra.

36. Assign MS/MS spectra with the Sequest or Mascot search algorithms. Search for a variable addition of 79.9663 Da on Ser, Thr, and Tyr (indicating phosphorylation), and a static modification of 17.0342 Da on Asp, Glu, and the peptide C terminus (indicating D3-methylation).

CONSTRUCTION OF COLUMNS FOR IMAC AND REVERSED-PHASE SEPARATION OF METHYL ESTER-DERIVITIZED PEPTIDES

The enrichment of phosphopeptide methyl esters by IMAC requires the use of a total of four columns. A large-capacity desalting column is used after peptide methylation to remove residual acid from the methylation reaction for better IMAC retention. An IMAC column loaded with Fe^{3+} enriches phosphopeptides. These phosphopeptides are then eluted to a precolumn with phosphate, and the phosphate is removed by washing. Peptides are then eluted by a reversed-phase gradient from the precolumn, through the analytical column, and into the mass spectrometer. These columns should be prepared and utilized in close collaboration with a mass spectrometry facility. Generally, this technique is compatible with any electrospray mass spectrometer capable of nanospray. In the authors' experience, the LTQ-FTICR and LTQ from Thermo are excellent instruments for phosphorylation analysis due to their impressive scan rate and sensitivity. Other instruments such as quadrupole time-of-flight (QTOF) mass spectrometers would likely provide enough sensitivity and mass accuracy, but require longer gradients for complex samples due to their slow scan rate.

Flow rates or volumes are measured with a 5- μl measured capillary and a stopwatch. Collect column flowthrough in a capillary over a 1-min period and then use a ruler to

SUPPORT PROTOCOL 2

Analysis of Protein Phosphorylation

18.13.11

measure the number of microliters collected in the measured capillary over this time period.

Use of the pressure bomb: Pressure bombs are used both to pack slurries of chromatographic resin as well as to load samples into 50- to 250- μm -i.d. capillaries. It is not feasible to provide the necessary pressure of >500 psi to prepare a column using a syringe. Although HPLC pumps can provide these sorts of pressures, it is not convenient to pass an emulsion of resin through the flow path of the HPLC. The pressure bomb (GNF commercial systems; Fig. 18.13.2) works by application of 500 to 1000 psi of nitrogen pressure to the 1/8-in. stainless steel tubing in the side of the bomb. A Swagelok three-position 1/8-in. ball valve is attached to the tubing to allow for the controlled release of pressure from the bomb. The sample is placed inside the bomb and a small glass capillary containing a porous frit is inserted into the sample solution through a 1/16-in. to 0.5-mm Teflon one-hole ferrule (Chromatography Research) in the top. The sample can be a mixture of reversed-phase or IMAC resin in 95% acetonitrile/5% isopropanol, column wash solution, or analytical sample. The pressure bomb is placed on a standard laboratory magnetic stirrer and all emulsions of resins are stirred in the bomb in small glass vials. Various materials may be used for the resin-retaining frit, including LiChrosorb Si60 (5 μm ; EMD Biosciences) or Kasil (PQ Corporation). Most conveniently, an inline microfilter from Upchurch can simply be screwed onto the end of the capillary. Once the bomb is closed with the two top screws, application of high pressure (typically through a Swagelok three-way valve) from a nitrogen gas tank causes the sample to be passed onto

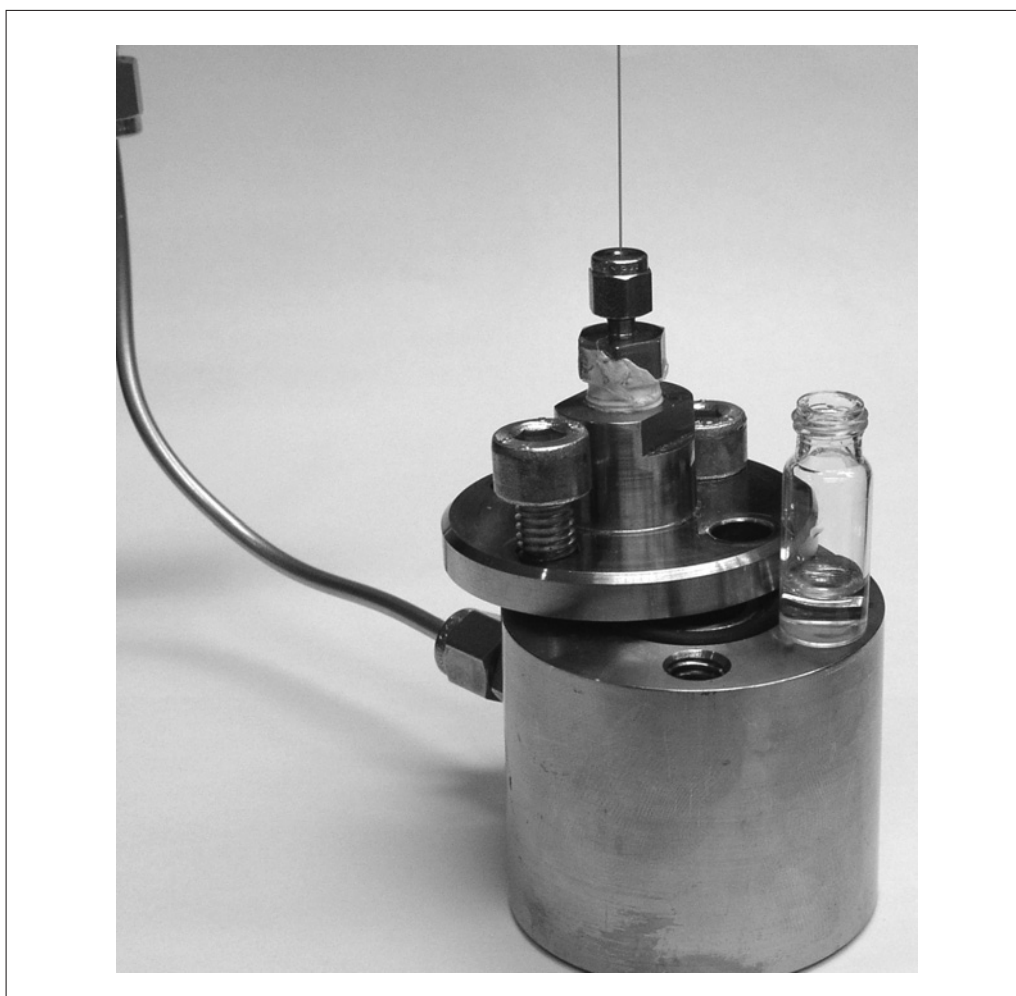


Figure 18.13.2 A pressure bomb used for packing resins and loading samples.

the column. The flow rate or volume through the column is gauged with a Drummond 5- μ l calibrated pipet and can be adjusted by altering the gas pressure at the regulator.

Construction of desalting and IMAC columns: The fused silica columns may be cut to the desired length with a fused silica cutter (Alltech), being sure that all cuts are straight (check under microscope). All columns are packed on a pressure bomb as described above and fritted with inline microfilters. The desalting column is packed with 12 cm of POROS 10R2 (a C18 resin available from Applied Biosystems) in 360- μ m o.d. \times 200- μ m i.d. fused-silica tubing, and the IMAC column is packed with 15 cm of POROS 20MC (Applied Biosystems) in a 360- μ m o.d. \times 200- μ m i.d. fused silica capillary. After packing, both columns are washed once for 5 min under 500 psi of bomb-applied pressure with acetonitrile and then with reversed-phase loading buffer.

Construction of precolumn: The precolumn is prepared by first preparing a LiChrosorb Si60 frit in a 360- μ m o.d. \times 75- μ m i.d. fused silica capillary. To prepare this type of frit, dip the capillary into the powdered LiChrosorb powder ten times and then flame the tip with a nitrous oxide flame three times for 1 sec each pass. A nitrous oxide flame may be created with a microflame gas torch set (VWR) by first starting a butane flame, minimizing its intensity, and then increasing the nitrous oxide until a uniform blue flame is produced. Adjust the capillary flame timing such that the LiChrosorb material is in the flame long enough to cause a brilliant white glow. Pack the precolumn with 2 cm of 5- μ m Monitor C18 (100- Å , Column Engineering) on the pressure bomb and wash with acetonitrile and then reversed-phase loading buffer.

Construction of an analytical column with integrated electrospray emitter tip: There are two possible configurations of the analytical column. For optimal sensitivity, analytical columns may be prepared with integrated electrospray emitter tips (Fig. 18.13.3) on a P2000 laser puller (Sutter Instruments). Useful recommendations for the tuning of the laser puller are provided in the manufacturer's manual. Remove 3 cm of the polyimide coating from a 360- μ m o.d. \times 75- μ m i.d. fused silica capillary by flaming with butane and removing debris with isopropanol. On the exposed section of the capillary, pull a 5- μ m-i.d. tip, being careful to avoid the polyimide. Obtaining a 5- μ m tip requires adjustment of the heat, velocity, delay, and pull settings on the puller. Good starting values are heat = 290, filament = 0, velocity = 15, delay = 127, and pull = 10. To maintain reproducible pulling, clean the puller mirror regularly with spectroscopy-grade isopropanol or acetone. Determine the diameter of the tip by bomb loading 3- μ m silica (Bangs Labs) into the column after pulling the tip. Two 3- μ m beads should have lodged into the tip. If the tip is occluded by one bead, it is too narrow; try decreasing the heat setting. After fritting the analytical column with 3- μ m silica, pack 12 cm of 5- μ m Monitor C18 (100- Å , Column Engineering) into the column.

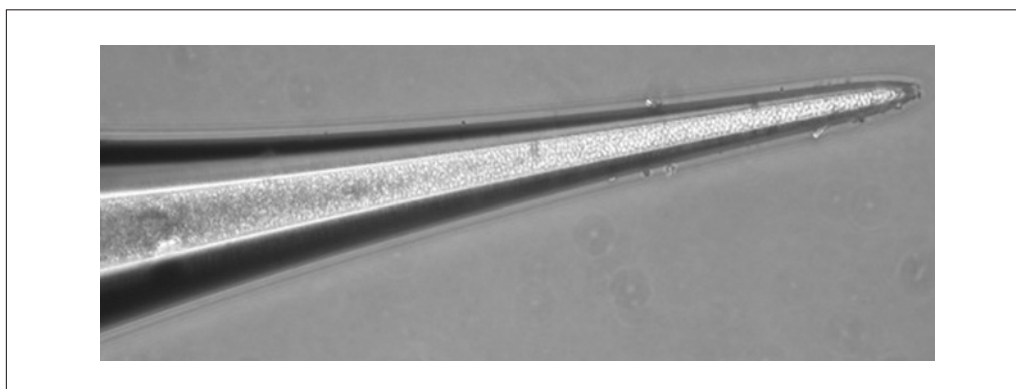


Figure 18.13.3 A 5- μ m emitter tip (under a microscope at 200 \times magnification).

Construction of analytical column attached to a picotip: If a laser puller is not available, picotips (8- μ m; New Objective) may be purchased. This configuration is less sensitive but may be adequate for some analyses. Pack a 360- μ m-o.d. \times 75- μ m-i.d. fused silica capillary, fritted with an inline microfilter, with 12 cm of 5- μ m Monitor C18 (100- \AA , Column Engineering). The analytical column is then attached to the picotip through the inline microfilter.

**BASIC
PROTOCOL 3**

**ENRICHMENT OF TYROSINE-PHOSPHORYLATED PEPTIDES
VIA PEPTIDE IMMUNOPRECIPITATION AND IMAC**

Unlike phosphotyrosine immunoprecipitation (IP) at the protein level (see Basic Protocol 2, steps 1b to 10b), peptide IP employs phosphotyrosine-specific antibodies to isolate tyrosine-phosphorylated peptides directly from a total cell lysate of digested cellular proteins (Rush et al., 2005). The advantage of protein IP is that it permits the more extensive analysis of all serine, threonine, and tyrosine phosphorylation sites on tyrosine-phosphorylated proteins as well as all phosphorylation sites on proteins that interact in a complex with tyrosine-phosphorylated proteins. The advantage of phosphotyrosine peptide IP is that it is more sensitive for detection of tyrosine phosphorylation due to the increased efficiency of capture of the phosphorylated peptides by the PTyr100 antibody relative to the PT66 antibody. Another advantage of the peptide IP procedure is that it provides higher-quality quantitative analysis, because peptide chemical isotopic labeling such as with the popular iTRAQ reagent does not interfere significantly with the IP step. With protein IP, peptides must be labeled after the IP step.

In the peptide IP procedure, cells are lysed and digested, then the digests are desalted and immunoprecipitated by antibody-conjugated protein G beads (Rush et al., 2005). Peptides are eluted from beads by IP elution buffer (Zhang et al., 2005). A second enrichment of phosphopeptides is accomplished by IMAC as described in Basic Protocol 2, but without methylation. IMAC is necessary after peptide IP due to unacceptably high carryover levels of unphosphorylated peptides through the IP. The strategy of enhancing the selectivity of IMAC using peptide methyl esters is unnecessary in this procedure due to the selective enrichment afforded by the phosphotyrosine peptide IP. Although, in theory, methyl ester formation would diminish the low levels of nonspecific unphosphorylated peptides that appear after peptide IP and IMAC, the decrease in overall peptide yield through the methyl ester reaction results in overall net decrease in phosphopeptides identified with this method.

Materials

- Anti-phosphotyrosine antibody (clone PTyr100, Cell Signalling Technology)
- Antibody coupling buffer: 20 mM sodium phosphate, pH 7.0
- Protein G-agarose beads (Roche)
- Protein sample (3×10^7 cells or ~ 100 fmol purified protein)
- Cell lysis buffer B (see recipe)
- Dithiothreitol (DTT)
- Iodoacetamide
- 100 mM ammonia bicarbonate, pH 8.9
- DC protein assay kit (Bio-Rad) containing:
 - Alkaline copper tartrate solution
 - Folin reagent
 - Surfactant solution
 - Bovine gamma-globulin standard
- Affinity purified, TPCK-treated trypsin (Promega)
- Concentrated HCl
- Sep-Pak C18 plus column (Waters)

100% acetonitrile
Sep-Pak loading buffer: 0.1% acetic acid in Milli-Q water
Sep-Pak elution buffer: 25% acetonitrile/0.1% acetic acid in Milli-Q water
Peptide immunoprecipitation (IP) buffer (see recipe), 4°C
Synthetic phosphopeptide external standard
pTyr-100 immobilized antibody
Peptide IP wash buffer 1 (see recipe)
Peptide IP wash buffer 2 (see recipe)
Peptide IP elution buffer: 100 mM glycine, pH 2.5
Tyrosine-phosphorylated standard peptides (e.g., angiotensin II phosphate, EMD Biosciences; or various pTyr peptides, Anaspec)

End-over-end rotator, 4°C
56°C water bath
15-ml conical tubes
1.5-ml screw-top microcentrifuge tubes (Sarstedt)
18- and 27-G needles
Speedvac evaporator (Thermo Savant)

Additional reagents and equipment for IMAC enrichment and MS analysis
(see Basic Protocol 2)

NOTE: All sample preparation procedures involving intact proteins are carried out at 4°C unless otherwise specified.

Prepare antibody

1. Prepare the pTyr-100 immobilized antibody by mixing soluble pTyr-100 in antibody coupling buffer with protein G–agarose beads at a concentration of 2.25 mg antibody/ml resin. Incubate overnight at 4°C with end-over-end rotation.

The pre-immobilized pTyr-100 currently sold from Cell Signaling Technologies is much less efficient than immobilized antibody prepared with protein G–agarose for peptide immunoprecipitation.

Prepare sample for immunoprecipitation

2. On the same day, lyse cells in 2 ml cold cell lysis buffer B per 3×10^7 cells and incubate 20 min on ice.
3. Centrifuge lysate 15 min at $14,000 \times g$, 4°C, and retain the supernatant.
4. Add DTT to a concentration of 10 mM to reduce proteins, and incubate 1 hr in a 56°C water bath.
5. Add iodoacetamide to a concentration of 55 mM and incubate 1 hr at room temperature in the dark.
6. Dilute lysate five times with 100 mM ammonia bicarbonate, pH 8.9, to a final urea concentration of 1.6 M. Measure the protein concentration using the DC protein assay kit according to the manufacturer's instructions.
7. Add affinity-purified, TPCK-treated trypsin to protein at a ratio of 1:100 (w/w) trypsin/total protein and incubate overnight at room temperature.

A trypsin concentration of 1:40 is optimal but may be cost prohibitive. The 1:40 ratio would afford ~10% more phosphopeptides and a lower rate of missed cleavage sites. A 1:300 ratio of trypsin provides unacceptably high rates of miscleavage and greatly diminishes the number of phosphopeptides identified. Promega affinity-purified and TPCK-treated trypsin works best.

8. Acidify tryptic digests to pH 2 with concentrated HCl and centrifuge acidified digests 5 min at $2000 \times g$ to remove any precipitate.

9. Pre-equilibrate a Sep-Pak C18 plus column by washing once with 5 ml of 100% acetonitrile and then twice with 3.5 ml of Sep-Pak loading buffer.

The flow rate for steps 8 through 11 is ~1 to 10 ml/min. Within this range, a lower flow rate gives a better loading efficiency. The control of flow rate can be accomplished by adjusting the intensity of the vacuum.

10. Load precleared digested peptides onto washed Sep-Pak plus column.
11. Wash column with 10 ml Sep-Pak loading buffer to remove salts.
12. Elute peptides from column with 10 ml Sep-Pak elution buffer into a 15-ml conical tube.
13. Dispense eluents into 1.5-ml screw-top microcentrifuge tubes with their tops perforated using an 18-G needle.
14. Dry peptides in a Speedvac evaporator without heat for 6 hr with rotation. Combine eluents from all tubes when the total volume fits in one tube, before peptides are completely dried.

Perform immunoprecipitation

15. Reconstitute peptides with 500 μ l cold (4°C) peptide IP buffer and add 10 pmol of a synthetic phosphopeptide external standard for quantitation of relative phosphopeptide abundance, as well as to control for variations in peptide IP efficiency.

A variety of synthetic tyrosine phosphorylated peptides are available from Anaspec.

16. Add 20 μ l (bed volume) of pTyr-100 immobilized antibody to peptide solution and incubate with gentle tumbling overnight at 4°C.
17. Wash resin two times with 400 μ L peptide IP wash buffer 1 for 10 min at 4°C, and microcentrifuge 5 min at 2000 \times g, 4°C, to bring down the resin. Remove supernatant as completely as possible by inserting a 27-G needle along the side and to the bottom of the microcentrifuge tube to remove any trace amounts of detergent.
18. Wash resin two times with 400 μ l peptide IP wash buffer 2 for 5 min at 4°C.
19. Elute phosphopeptides from resin with 100 μ L of peptide IP elution buffer 1 hr at room temperature. Spin down 5 min at 2000 \times g, room temperature, and collect all eluent by inserting a 27-G needle along the side and to the bottom of the microcentrifuge tube.
20. Add 200 fmol of another tyrosine phosphorylated standard peptide such as angiotensin II phosphate to the eluent as a quantitative control for the remaining steps.
21. Perform IMAC enrichment and MS analysis as in Basic Protocol 2, steps 22 through 36, but replace the database search parameters in step 36 with a static modification of +57.0215 Da on Cys and a variable modification of +79.9663 Da on Ser, Thr, and Tyr.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Cell lysis buffer A

- 20 μ g/ml aprotinin
- 50 mM Tris·Cl, pH 7.5 (APPENDIX 2)
- 100 mM NaCl

1% (v/v) Triton X-100
10% (v/v) glycerol
20 µg/ml leupeptin
1 mM Pefabloc
2 mM Na₃VO₄ (prepared fresh)
1 mM EDTA
10 mM β-glycerophosphate
Prepare fresh and keep at 4°C

Cell lysis buffer B

8 M urea
1 mM sodium orthovanadate
100 mM ammonium bicarbonate, pH 8.0
Prepare fresh and keep at 4°C

Dialysis buffer

50 mM ammonium bicarbonate, pH 8.3
4 M urea
Prepare fresh and keep at 4°C

IMAC elution buffer

0.4 M ammonium hydroxide
30% (v/v) acetonitrile

This buffer is for use with Phos-Select resin (Sigma); for other affinity resins, the manufacturers' instructions should be followed.

IMAC load/wash buffer

250 mM acetic acid
30% (v/v) acetonitrile

This buffer is for use with Phos-Select resin (Sigma); for other affinity resins, the manufacturers' instructions should be followed.

IMAC wash buffer

25% (v/v) acetonitrile
75% (v/v) ddH₂O
0.01% (v/v) acetic acid
100 mM NaCl

Immunoprecipitation wash buffer

20 mM Tris, pH 7.4 (APPENDIX 2)
120 mM NaCl

Methylation solution

Layer nitrogen gas over the reaction at each point in its preparation. Add 0.75 ml of 99.8% methanol-*d*₄ (Cambridge Isotope Labs) to a small dry glass vial with stirring (methanol-*d*₄ is used primarily because it is supplied very dry in ampules). Add 120 µl of 98+% acetyl chloride (Sigma) to methanol-*d*₄ dropwise with stirring (popping occurs if acetyl chloride is added too quickly). Incubate 5 min at room temperature and add immediately to dried peptides. The volume of this solution should be scaled up such that all related samples are methylated with the same methylation solution (use 400 µl of solution for each sample).

This solution must be very dry due to the reversible nature of the methylation reaction.

Peptide immunoprecipitation buffer

30 mM Tris·Cl, pH 7.4 (*APPENDIX 2*)
30 mM NaCl
0.3% (v/v) nonidet P-40 (NP-40)

Peptide immunoprecipitation wash buffer 1

100 mM Tris·Cl, pH 7.4 (*APPENDIX 2*)
100 mM NaCl
0.3% (v/v) nonidet P-40 (NP-40)

Peptide immunoprecipitation wash buffer 2

100 mM Tris·Cl, pH 7.4 (*APPENDIX 2*)
100 mM NaCl

Reversed-phase elution buffer

70% (v/v) acetonitrile
29.9% (v/v) Milli-Q water
0.1% (v/v) acetic acid (final concentration)

SAX start buffer

20 mM triethylammonium bicarbonate (TEAB)
30% (v/v) acetonitrile
The pH is ~8 or 9. No adjustment is necessary.

SAX step buffer

20 mM triethylammonium bicarbonate (TEAB)
1 M ammonium formate
30% (v/v) acetonitrile

SAX step elution buffers

Mix SAX start buffer (see recipe) and SAX step buffer (see recipe) to yield ammonium formate concentrations ranging from 50 to 500 mM.

These concentrations are appropriate for POROS HQ resin. For other anion-exchange resins, the concentrations may need to be adjusted.

SCX cleaning buffer

1 M ammonium formate
30% (v/v) acetonitrile
Adjust to pH 3 with formic acid

This requires a large amount of formic acid, and it is virtually impossible to bring the pH below 3.

SCX start buffer

5 mM ammonium formate
30% (v/v) acetonitrile
Adjust to pH 2.7 with formic acid

SCX step elution buffers

Mix SCX start buffer (see recipe) and SCX cleaning buffer (see recipe) to yield total ammonium formate concentrations ranging from 10 to 100 mM.

In Opti-Lynx cartridges, most phosphopeptides elute between 10 and 50 mM salt. For other column materials, the appropriate salt concentrations must be determined empirically.

Trypsin digestion buffer with urea

100 mM ammonium bicarbonate, pH 8.3

8 M urea

Urea solution MUST be prepared fresh before each use.

COMMENTARY

Background Information

Most cellular processes are regulated at some level by protein phosphorylation, be it developmental, hormonal, stress, or disease signaling. Of all posttranslational modifications, phosphorylation is the best understood, probably most widespread, and arguably the most important. To understand and investigate the role phosphorylation plays in cellular signaling, it is essential to know not only the identity of the phosphorylated proteins but also the exact site of modification. Site-directed mutagenesis of the phosphorylatable residue to a phospho-amino acid mimetic or a non-phosphorylatable residue reveals the biological significance of the modification, and ultimately the circuitry of the signaling pathway.

The determination of phosphorylation sites on individual proteins of interest was once an arduous task involving *in vivo* or *in vitro* labeling with radioactive phosphate, separation of labeled peptides by column or thin-layer chromatography, and Edman sequencing. These tools are still indispensable in some cases, but the rapid development of highly sensitive and accessible mass spectrometry techniques makes this an invaluable addition to the toolbox.

Phosphopeptides have a number of properties that make their analysis by mass spectrometry substantially more difficult than that of other peptides. Small phosphopeptides are so hydrophilic that they may not be retained on the reversed-phase matrices typically used in LC-MS (Larsen et al., 2004). In addition, phosphorylation sites of signaling proteins are often extremely difficult to detect due to their low abundance and low stoichiometry of phosphorylation in a cellular context of abundant background proteins. Several approaches have been developed to overcome these limitations by increasing the scale of the experiment through multidimensional chromatography (MudPIT; Washburn et al., 2001) or low-pH strong-cation exchange (Beausoleil et al., 2004). These unspecific methods have a lower efficiency in purification of phosphopeptides than affinity methods such as immobilized metal ion affinity chromatography (IMAC), making them better suited as a prefraction-

ation step prior to affinity chromatography for extremely complex protein mixtures such as total cell lysates. Phosphopeptide tandem mass spectra frequently have poor quality due to a combination of poor signal-to-noise ratio (arising from the low abundance and poor ionization of phosphorylated peptides) and poor peptide backbone fragmentation (due to abundant neutral loss of phosphate from phosphoserine and phosphothreonine residues). Since a phosphoprotein is often identified from a single phosphopeptide spectrum, it is necessary to “manually” inspect the quality of the database match for every candidate MS/MS spectrum. This time-consuming limitation could potentially be overcome by increasing the quality of the data through the use of high mass accuracy spectrometers that narrow down the database search (Mann and Jensen, 2003), better fragmentation methods such as electron transfer dissociation (Syka et al., 2004), and more highly selective chromatography (Ficarro et al., 2002; Brill et al., 2004; Ndassa et al., 2006), as well as the development of new statistical algorithms (MacCoss et al., 2002). The limited dynamic range of mass spectrometers necessitates selective enrichment of phosphopeptides to obtain the highest numbers of phosphorylation sites from a complex mixture. Furthermore, peptide MS/MS spectra are routinely obtained in the positive ion mode, *i.e.*, from fragmentation of the protonated peptide cations, where the additional negative charge of the phosphate group reduces the ionization efficiency of phosphopeptides relative to unphosphorylated peptides. While it is not possible to change the phosphate group itself when looking for definitive proof of the modification from the mass spectrum, it is possible to overcome the ionization problem by enriching the sample as much as possible for phosphopeptides.

By far the most successful method for enriching phosphopeptides is immobilized metal ion affinity chromatography (IMAC; UNIT 10.11B). A range of variations has been published, using different chromatographic materials, metal ions, and manual or automated setups (Neville et al., 1997; Posewitz and Tempst, 1999; Stensballe et al., 2001;

Ficarro et al., 2005; Olsen et al., 2006). Often, it is much easier to isolate phosphopeptides from complex mixtures than to identify a certain phosphopeptide(s) from individual proteins.

Affinity chromatography of proteins on chelated metal ions has been used for a long time and enjoys widespread popularity in the form of a nickel-nitrilotriacetic (NTA) or iminodiacetic acid (IDA) resin for the purification of hexahistidine-tagged proteins. Andersson and Porath (Andersson and Porath, 1986) were the first to notice that Fe^{3+} -chelated IDA has a strong affinity for highly phosphorylated proteins. While the use of IMAC for full-length proteins has been limited, at least until the recent availability of an IMAC-based commercial kit for enrichment of phosphoproteins, the isolation of phosphopeptides has been successful in a large number of cases. One of the first uses of the technique was in the study of phosphopeptides from casein in a nutritional context (Scanff et al., 1991). A large number of technical publications on new concepts for phosphopeptide isolation with IMAC or other procedures demonstrate feasibility exclusively with digests of β -casein. The reader should be wary of claims of the universality of these approaches until successfully applied to mixtures of proteins. IMAC with Fe^{3+} or other metals such as Ga^{3+} , or more recently TiO_2 (Olsen et al., 2006), is the longest established and most widely, used phosphoenrichment technique. IMAC has allowed the identification of phosphorylation sites in individual proteins (e.g., Michel et al., 1988; Neville et al., 1997; Fuglsang et al., 1999) as well as in complex mixtures (Vener et al., 2001; Ficarro et al., 2002; Nühse et al., 2003; Salomon et al., 2003; Olsen et al., 2006).

The mechanism of phosphoamino acid binding to chelated Fe^{3+} is a combination of ion exchange, i.e., binding of the negatively charged phosphate groups to the positive metal ion, and a more specific coordination of phosphate in the metal ligand sphere (Holmes and Schiller, 1997). It is easy to see why highly acidic peptides can bind to IMAC material as contaminants—the ion exchange function applies to them just as to phosphopeptides. Typical chromatographic conditions for IMAC make it more specific for phosphorylated peptides, and at least for some types of samples, the degree of purity is sufficient for large-scale mass spectrometric analysis (Nühse et al., 2003). For more complex sample analysis such as for total lysates or total cellular tyrosine phosphorylation, methyl esterifi-

cation of carboxyl groups provides remarkable selectivity and efficiency in phosphopeptide isolation (Ficarro et al., 2002; Salomon et al., 2003).

To increase the number of phosphorylation sites discovered, very complex samples may be fractionated at the peptide or protein level prior to IMAC and LC-MS. In the case of phosphoproteins, one-dimensional gel electrophoresis or affinity purification may be helpful. For phosphopeptides, anion- or cation-exchange chromatography is useful. Both have a double effect of “deconvoluting” the sample as well as enriching phosphopeptides to some extent over the bulk of unmodified peptides. In cation exchange chromatography, many phosphopeptides bind weakly and elute very early or do not bind at all, while many unphosphorylated peptides bind well. Although not as selective as IMAC, this reasonably selective enrichment in the early eluting fractions may be sufficient for analysis without IMAC, especially at low pH (Beausoleil et al., 2004). In anion exchange chromatography, the bulk of phosphopeptides binds better than other peptides and is eluted over a wider range of salt concentrations (Nühse et al., 2003). This orthogonal subfractionation of phosphopeptides may provide added benefits over cation exchange chromatography, where the interesting fraction elutes over a very narrow salt range.

Critical Parameters and Troubleshooting

The most typically encountered problems are too many unphosphorylated peptides in the IMAC eluates, too few phosphopeptides, or both.

Following Basic Protocol 1, an abundance of non-phosphopeptides means that the IMAC procedure was not specific enough. If spin columns are used for IMAC, it is critical to resuspend the resin well in the various wash buffers and to not simply add the buffers onto the column. To improve specificity, additional wash steps can be introduced but need to be balanced against the phosphopeptide yield—their affinity to the medium is not high enough to withstand extensive washing. Alternatively, different suppliers of IMAC resin can be tested. Similar sets of peptides have been successfully identified with both Sigma's Phos-Select and Pierce's gallium-loaded IMAC spin columns, but results may vary depending on the sample. If acidic peptides are the major contaminating non-phosphopeptides, methyl ester derivitization

of peptides is recommended (see Basic Protocol 2).

There are many possible reasons for finding too few phosphopeptides. The *in vivo* phosphorylation status of proteins needs to be preserved by harvesting cells rapidly and either isolating proteins under denaturing conditions until digesting with trypsin, or by adding sufficient amounts of phosphatase inhibitors to the native extract. It may be worth increasing inhibitor concentrations above levels published in the literature, and in some cases a substantially higher yield of phosphopeptides can be achieved by adding calyculin A (5 to 20 nM) as a phosphatase inhibitor even during the tryptic digestion. Although performing the protein digestion overnight is the most convenient way, it can be detrimental for smaller amounts of protein to be left at 37°C for too long; a few hours may be sufficient. The addition of RapiGest detergent (Waters) speeds up tryptic digestion and is, in the authors' experience, compatible with IMAC, provided acid hydrolysis of the detergent is performed according to the manufacturer's instructions and the digest is cleaned up by cation exchange separation.

If too few phosphopeptides are recovered from IMAC columns following Basic Protocol 1, the incubation time of the sample with the resin may be increased, washes shortened, and elution times prolonged, although the manufacturers' protocols are generally optimized.

In Basic Protocols 2 and 3, phosphopeptide yield is fully optimized by direct elution of phosphopeptides from the reversed-phase desalt column to the IMAC column to the reversed-phase precolumn and gradient elution into the mass spectrometer via the analytical column. In this method, the peptide never encounters plastic once the phosphopeptides are enriched providing an ideal yield in their separation.

After elution of phosphopeptides from the beads in Basic Protocol 1, it is crucial to bear in mind that even with a very good yield, the phosphopeptide solution is dilute and total amounts are low. Any delay before the analysis should be kept to a minimum; ideally, IMAC purification is scheduled to be directly followed by mass spectrometric analysis. If necessary, isolated phosphopeptides can be frozen at -80°C if they are not too dilute, but freezing total (or fractionated) protein digests is always preferable. Freeze-dried samples should not be refrozen; they are very hygroscopic and degrade when wet.

Finally, it is self-evident that high-sensitivity sequencing of phosphopeptides requires optimal calibration and sensitivity of the mass spectrometer. Optimal sensitivity is achieved through minimization of the electrospray flow rate (<20 nl/min) and electrospray emitter tip diameter.

Anticipated Results

The number of sequenced phosphopeptides depends on the amount and complexity of the sample, the detected abundance of phosphopeptides in the sample (which in turn depends on the natural abundance of the phosphoproteins and on each phosphopeptide's unique and unpredictable ionization efficiency) and the sensitivity and speed of the mass spectrometer. Following Basic Protocol 1, several tens to >100 phosphopeptide sequences can be identified from an enriched sample analyzed in a single LC-MS run, and ~100 to 200 with manual prefractionation of the sample. With a larger input of sample, more extensive fractionation and possibly multiple repeated LC-MS runs for each IMAC-enriched sample, this figure can be increased to several thousand identified peptides. With most types of samples, ~70% to 90% pure phosphopeptides have been obtained using samples prefractionated by ion-exchange chromatography.

With Basic Protocols 2 and 3, hundreds to thousands of phosphopeptides are identified routinely depending on the nature of the sample and the amount of prefractionation of proteins and peptides performed prior to IMAC.

If the isolated phosphopeptides are analyzed on an ion trap mass spectrometer, serine- and threonine-phosphorylated peptides typically display a prominent neutral loss, i.e., an abundant fragment ion corresponding to loss of phosphoric acid (-98 Da for +1; -49 *m/z* for +2; -32.7 Da for +3 precursor ions) that often completely dominates the spectrum and reduces the quantity and quality of useful b and y ion peaks. Rapidly evolving mass spectrometry methodologies such as MS³ (Olsen and Mann, 2004), pseudo MS³ (Schroeder et al., 2004), and ETD (Syka et al., 2004) aim to overcome this significant problem.

It is important to note that tyrosine phosphorylation is substantially less frequent than serine or threonine phosphorylation. Without enrichment of phosphotyrosine-containing proteins or peptides by immunoprecipitation, peptides containing this residue will be outnumbered several hundred fold in

the protein digest by serine- and threonine-phosphorylated peptides.

Time Considerations

The entire procedure from sample preparation to mass spectrometric analysis will take between 2 and 3 days; there are several steps where samples can be frozen before continuing. Protein extracts are prepared on day 1 and will require between a few hours and a full day, depending on whether subcellular fractionation and, optionally, immunoprecipitation is performed. In almost all cases, the digestion is performed overnight. On the second day, the lyophilization before ion exchange chromatography (see Basic Protocol 1) requires several hours, and for ion exchange fractions, >0.5 ml, the lyophilization of these fractions will preferably be done overnight again. On the third day, the IMAC purification will take ~2 hr for up to twelve separate samples. For Basic Protocol 2, peptide methylation or peptide IP takes 1 day and the entire procedure of desalting the peptides, IMAC enrichment, and reversed-phase elution into the instrument can be accomplished in 3 hr.

Literature Cited

- Andersson, L. and Porath, J. 1986. Isolation of phosphoproteins by immobilized metal (Fe^{3+}) affinity chromatography. *Anal. Biochem.* 1:250-254.
- Beausoleil, S.A., Jedrychowski, M., Schwartz, D., Elias, J.E., Villen, J., Li, J., Cohn, M.A., Cantley, L.C., and Gygi, S.P. 2004. Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc. Natl. Acad. Sci. U.S.A.* 101:12130-12135.
- Brill, L.M., Salomon, A.R., Ficarro, S.B., Mukherji, M., Stettler-Gill, M., and Peters, E.C. 2004. Robust phosphoproteomic profiling of tyrosine phosphorylation sites from human T cells using immobilized metal affinity chromatography and tandem mass spectrometry. *Anal. Chem.* 76:2763-2772.
- Ficarro, S.B., McClelland, M.L., Stukenberg, P.T., Burke, D.J., Ross, M.M., Shabanowitz, J., Hunt, D.F., and White, F.M. 2002. Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 20:301-305.
- Ficarro, S.B., Salomon, A.R., Brill, L.M., Mason, D.E., Stettler-Gill, M., Brock, A., and Peters, E.C. 2005. Automated immobilized metal affinity chromatography/nano-liquid chromatography/electrospray ionization mass spectrometry platform for profiling protein phosphorylation sites. *Rapid Commun. Mass Spectrom.* 19:57-71.
- Fuglsang, A.T., Visconti, S., Drumm, K., Jahn, T., Stensballe, A., Mattei, B., Jensen, O.N., Aducci, P., and Palmgren, M.G. 1999. Binding of 14-3-3 protein to the plasma membrane H(+)-ATPase
- AHA2 involves the three C-terminal residues Tyr(946)-Thr-Val and requires phosphorylation of Thr(947). *J. Biol. Chem.* 274:36774-36780.
- Gobom, J., Nordhoff, E., Mirgorodskaya, E., Ekman, R., and Roepstorff, P. 1999. Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* 2:105-116.
- Gruhler, A., Olsen, J.V., Mohammed, S., Mortensen, P., Faergeman, N.J., Mann, M., and Jensen, O.N. 2005. Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics* 4:310-327.
- Holmes, L.D. and Schiller, M.R. 1997. Immobilized iron(III) metal ion affinity chromatography for the separation of phosphorylated macromolecules: Ligands and applications. *J. Liq. Chromatogr. Relat. Technol.* 1:123-142.
- Larsen, M.R., Graham, M.E., Robinson, P.J., and Roepstorff, P. 2004. Improved detection of hydrophilic phosphopeptides using graphite powder microcolumns and mass spectrometry: Evidence for in vivo doubly phosphorylated dynamin I and dynamin III. *Mol. Cell. Proteomics.* 3:456-465.
- MacCoss, M.J., Wu, C.C., and Yates, J.R. III. 2002. Probability-based validation of protein identifications using a modified SEQUEST algorithm. *Anal. Chem.* 74:5593-5599.
- Mann, M. and Jensen, O.N. 2003. Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* 3:255-261.
- Michel, H., Hunt, D.F., Shabanowitz, J., and Bennett, J. 1988. Tandem mass spectrometry reveals that three photosystem II proteins of spinach chloroplasts contain *N*-acetyl-*O*-phosphothreonine at their NH2 termini. *J. Biol. Chem.* 263:1123-1130.
- Ndassa, Y.M., Orsi, C., and Marto, J.A. 2006. Improved immobilized metal affinity chromatography for large-scale phosphoproteomics applications. *J. Prot. Res.* 5:2789-2799.
- Neville, D.C., Rozanas, C.R., Price, E.M., Gruis, D.B., Verkman, A.S., and Townsend, R.R. 1997. Evidence for phosphorylation of serine 753 in CFTR using a novel metal-ion affinity resin and matrix-assisted laser desorption mass spectrometry. *Protein Sci.* 11:2436-2445.
- Nühse, T.S., Stensballe, A., Jensen, O.N., and Peck, S.C. 2003. Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry. *Mol. Cell. Proteomics* 2:1234-1243.
- Olsen, J.V. and Mann, M. 2004. Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. *Proc. Natl. Acad. Sci. U.S.A.* 101:13417-13422.
- Olsen, J.V., Blagoev, B., Gnäd, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. 2006. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127:635-648.

- Posewitz, M.C. and Tempst, P. 1999. Immobilized gallium(III) affinity chromatography of phosphopeptides. *Anal. Chem.* 14:2883-2892.
- Rush, J., Moritz, A., Lee, K.A., Guo, A., Goss, V.L., Spek, E.J., Zhang, H., Zha, X.M., Polakiewicz, R.D., and Comb, M.J. 2005. Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat. Biotechnol.* 23:94-101.
- Salomon, A.R., Ficarro, S.B., Brill, L.M., Brinker, A., Phung, Q.T., Ericson, C., Sauer, K., Brock, A., Horn, D.M., Schultz, P.G., and Peters, E.C. 2003. Profiling of tyrosine phosphorylation pathways in human cells using mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* 2:443-448.
- Scanff, P., Yvon, M., and Pelissier, J.P. 1991. Immobilized Fe³⁺ affinity chromatographic isolation of phosphopeptides. *J. Chromatogr.* 539:425-432.
- Schroeder, M.J., Shabanowitz, J., Schwartz, J.C., Hunt, D.F., and Coon, J.J. 2004. A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. *Anal. Chem.* 76:3590-3598.
- Stensballe, A., Andersen, S., and Jensen, O.N. 2001. Characterization of phosphoproteins from electrophoretic gels by nanoscale Fe(III) affinity chromatography with off-line mass spectrometry analysis. *Proteomics* 2:207-222.
- Syka, J.E., Coon, J.J., Schroeder, M.J., Shabanowitz, J., and Hunt, D.F. 2004. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* 101:9528-9533.
- Vener, A.V., Harms, A., Sussman, M.R., and Vierstra, R.D. 2001. Mass spectrometric resolution of reversible protein phosphorylation in photosynthetic membranes of *Arabidopsis thaliana*. *J. Biol. Chem.* 10:6959-6966.
- Washburn, M.P., Wolters, D., and Yates, J.R. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 19:242-247.
- Zhang, Y., Wolf-Yadlin, A., Ross, P.L., Pappin, D.J., Rush, J., Lauffenburger, D.A., and White, F.M. 2005. Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol. Cell. Proteomics* 4:1240-1250.

Contributed by Thomas Nühse
University of Manchester
Manchester, United Kingdom

Kebing Yu and Arthur Salomon
Brown University
Providence, Rhode Island

CHAPTER 19

Informatics for Molecular Biologists

INTRODUCTION

Biological information—including DNA sequence and gene mapping data as well as their associated literature—is growing exponentially, especially with the advent of large genome projects that generate enormous amounts of DNA sequence data (see Fig. 19.2.1). To many, this flood of information is overwhelming. How can it be collected, sorted, accessed, and manipulated? Recently, significant progress has been made in the development of computer programs that navigate this sea of information. Many databases have been developed that store the data, and much progress has been made toward interconnecting them in ways that are transparent (user-friendly) to the average over-worked molecular biologist. The key to the new information age is the worldwide Internet computer network, and that is where this chapter begins.

Traditionally, the Internet has primarily been used to retrieve software or database files and to allow research centers to communicate with one another. Now, however, new uses for the Internet are being developed and are starting to change the way information is obtained. These new services provide methods for submitting interactive queries to a distant database, making an enormous amount of information freely available to any scientist with access to the Internet. Even a geographically isolated researcher can gain access to the same information that colleagues at a large metropolitan laboratory take for granted, simply by arranging access to an institution that is connected to the Internet. Although the network services are free, the institution or department must buy and maintain the equipment required for the network connection.

This chapter provides an introduction to the vast array of biological information freely available via the Internet and focuses on the software tools required for access to and analysis of data contained in DNA and protein sequence databases (e.g., GenBank, SWISS-PROT) and the associated biomedical literature. Its aim is to facilitate the understanding and use of the most-commonly employed informatic tools used by biologists. The *Current Protocols in Bioinformatics* series provides additional depth and breadth.

UNIT 19.1 provides an overview of the Internet for biologists, and describes basic tools for electronic information retrieval from public data repositories and newsgroups. This unit presents a simple introduction to the tools necessary for exchanging information on the Internet, including a brief description of how to send and receive information by electronic mail (e-mail) and participate in network newsgroups. Navigational strategies and tools are described that allow the user either to browse or conduct a directed search. The World Wide Web (WWW), for example, is an expanding information retrieval system for documents from all over the world, based upon cross-references set up as “hypertext” links. Perhaps the most significant navigational tool for the World Wide Web is the graphical browser Mosaic from the National Center for Supercomputer Applications. Detailed methods are described for searching and downloading information using anonymous file transfer protocol (FTP), Gopher, and the World Wide Web. A final section on so-called “client-server” applications provides introductions to several important programs for biologists. These include BLAST for rapid comparison of experimental sequences against the DNA and protein sequence databases, Entrez for rapid sequence comparison and MEDLINE

literature surveys, GDB Accessor for accessing gene mapping data, and XGRAIL for rapidly scanning experimental sequence data for coding regions.

UNIT 19.2 describes integrated informational retrieval from sequence, structure, and literature databases using the Entrez system (the Entrez client software will actually be included in a future supplement of *CP on CD* with software for MACAW, the Multiple Alignment Construction and Analysis Workbench). Entrez is a user-friendly program that integrates DNA and protein sequence data with its associated abstracts and citations contained in the MEDLINE database. It is therefore possible using Entrez to search the databases by bibliographic associations (e.g. author name, gene family, or map location) as well as by sequence homology. The submission of new data, updates, and corrections to the sequence databases is an important, but often neglected, topic that is also described in this unit.

UNIT 19.3 provides detailed descriptions of methods for homology searching of DNA and protein sequence databases using the popular Basic Local Alignment Search Tool (BLAST) family of programs. The basic algorithm is compared and contrasted with FASTA. Notably, many tricks and tips for avoiding common pitfalls in homology searching are discussed. An explanation of how to judge the statistical significance of search results is also included.

UNIT 19.4 describes how to access a multitude of protein databases to obtain a variety of information about protein sequences. Searching these databases is an important early step in the characterization of a newly identified gene or protein, often providing key information concerning function. In many cases, phylogenetic information is available about the evolution of a particular protein or protein family, often providing additional insights into function. Protein databases include both small and large experimental ones that primarily catalog information about protein sequences and secondary databases that reorganize, annotate, and manipulate the data in the experimental ones. This unit provides a guide to these important resources.

UNIT 19.5 describes basic protein sequence analysis focused on enabling the biologist to make a more informed prediction of the likely molecular function of their protein of interest. It builds on the information presented in *UNITS 19.3 & 19.4*. Specifically, it discusses how to identify structural and functional domains, predict helical transmembrane regions, assess the subcellular localization of the protein, and predict key functional residues and potential sites for post-translational modifications *in silico*. A discussion of the critical parameters and assumptions for these tools, as well as guidelines for assessing the reliability of the findings are presented. These predictions can then be strengthened or confirmed by assessing experimental information. For example, one could determine whether a protein predicted to be plastid-localized has been identified by proteomics as residing in the plastid (e.g., see The Plastid Proteome Database, <http://cbsusrv01.tc.cornell.edu/users/ppdb/>).

A cautionary note: It is critical to remember that these informatic analyses are dependent on an accurate protein sequence. Certain transit peptides (such as chloroplast transit sequences) are often missed using annotation programs. In addition, there may be alternative forms of the enzyme (e.g., due to alternative splicing). Therefore, it is important to assess the validity of the sequence of interest manually by comparison with expressed sequence tag (EST) databases, alternative splicing event databases (e.g., see http://rarge.gsc.riken.jp/a_splicing/index.pl), tiling array expression databases such as the Arabidopsis Tiling Array Transcriptome Express Tool (<http://signal.salk.edu/cgi-bin/atta>), proteomic databases, and the primary literature.

Frederick Ausubel, Nicholas C. Dracopoli, and Mary Wildermuth

With the explosion of sequence and structural information available to researchers, the field of bioinformatics is playing an increasingly large role in the study of fundamental biomedical problems. The challenge facing computational biologists will be to aid in gene discovery and in the design of molecular modeling, site-directed mutagenesis, and experiments of other types that can potentially reveal previously unknown relationships with respect to the structure and function of genes and proteins. This challenge becomes particularly daunting in light of the vast amount of data that has been produced by the Human Genome Project and other systematic sequencing efforts to date.

Before embarking on any practical discussion of computational methods in solving biological problems, it is necessary to lay the common groundwork that will enable users to both access and implement the algorithms and tools discussed in this book. We begin with a review of the Internet and its terminology, also discussing major classes of Internet protocols, without becoming overly engaged in the engineering minutiae underlying these protocols. A more in-depth treatment on the inner workings of these protocols may be found in a number of well-written reference books intended for the lay audience (Krol and Klopfenstein, 1996; Rankin, 1996; Kennedy, 1999). This unit will also discuss matters of connectivity, ranging from simple modem connections to digital subscriber lines (DSL). Finally, we will address one of the most common problems that has arisen with the proliferation of Web pages throughout the world—i.e., finding useful information on the World Wide Web.

INTERNET BASICS

Despite the impression that the Internet is a single entity, it is actually a network of networks, composed of interconnected local and regional networks in more than 100 countries. While work on remote communications began in the early 1960s, the true origins of the Internet lie with a research project on networking at the Advanced Research Projects Agency of the U.S. Department of Defense in 1969 named ARPANET. The original ARPANET connected four nodes on the West Coast, with the immediate goal of being able to transmit information on defense-related research between laboratories. A number of different network projects subsequently surfaced, with the next landmark developments coming over ten years later. In 1981, BITNET (“Because It’s Time”) was introduced, providing point-to-point connections between universities for the transfer of electronic mail and files. In 1982, ARPA introduced the Transmission Control Protocol (TCP) and Internet Protocol (IP); TCP/IP allowed different networks to be connected to and communicate with one another, creating the system that is in place today.

A number of references chronicle the development of the Internet and communications protocols in detail (Quarterman, 1990; Froehlich and Kent, 1991; Krol and Klopfenstein, 1996). Most users, however, are content to leave the details of *how* the Internet works to their systems administrators; the relevant fact to most is that it *does* work.

Once the machines on a network are connected to one another, there needs to be some way to unambiguously specify a single computer so that messages and files find their intended recipient. To accomplish this, all machines directly connected to the Internet have an **IP number**. IP numbers are unique, identifying one and only one machine. The IP number is made up of four numbers separated by periods; for example, the IP number for the main file server at the National Center for Biotechnology Information (NCBI) at the National Institutes of Health (NIH) is *130.14.25.1*. The numbers themselves represent,

from left to right, the domain (*130.14* for the NIH), the subnet (*.25* for the National Library of Medicine at NIH), and the machine itself (*.1*). While the use of IP numbers aids the computers in directing data, they are obviously very difficult for users to remember. Therefore, an IP number often has a **fully-qualified domain name** (FQDN) associated with it, which is dynamically translated in the background by a **domain name server**. Going back to the NCBI example, instead of using *130.14.25.1* to access the NCBI computer, a user could instead use *ncbi.nlm.nih.gov* and achieve the same result. Reading from left to right, the IP number goes from least to most specific, while the FQDN equivalent goes from most specific to least. The name of any given computer can then be thought of as taking the general form *computer.domain*, with the top-level domain (the portion coming after the last period in the FQDN) falling into one of the six broad categories shown in Table 19.1.1. Outside the United States, the top-level domain names may be replaced with a two-letter code specifying the country where the machine is located (e.g., *.ca* for Canada and *.uk* for the United Kingdom).

In an effort to anticipate the needs of Internet users in the future, as well as to try to erase the arbitrary line between top-level domain names based on country, the now-dissolved International Ad Hoc Committee (IAHC) was charged with developing a new framework of generic top-level domains (gTLD). The new, recommended gTLDs were set forth in a document entitled *The Generic Top Level Domain Memorandum of Understanding* (gTLD-MOU); these gTLDs are overseen by a number of governing bodies and are also shown in Table 19.1.1.

The most concrete measure of the size (and, thereby, the success) of the Internet lies in actually counting the number of machines physically connected to it. The Internet Software Consortium conducts an Internet Domain Survey twice each year to count these machines, otherwise known as **hosts**. In performing this survey, ISC considers not only how many hostnames have been assigned, but how many of those are actually in use; a

Table 19.1.1 Top-Level Domain Names

Top-level domain names	
Inside U.S.	
<i>.com</i>	Commercial site
<i>.edu</i>	Educational site
<i>.gov</i>	Government site
<i>.mil</i>	Military site
<i>.net</i>	Gateway or network host
<i>.org</i>	Private (usually not-for-profit) organizations
Examples of top-level domain names used outside the United States	
<i>.ca</i>	Canadian site
<i>.ac.uk</i>	Academic site in the United Kingdom
<i>.co.uk</i>	Commercial site in the United Kingdom
Generic top-level domains proposed by IAHC	
<i>.firm</i>	Firms or businesses
<i>.shop</i>	Businesses offering goods to purchase (stores)
<i>.web</i>	Entities emphasizing activities relating to the World Wide Web
<i>.arts</i>	Cultural and entertainment organizations
<i>.rec</i>	Recreational organizations
<i>.info</i>	Information sources
<i>.nom</i>	Personal names (e.g., <i>yourlastname.nom</i>)

hostname might be issued, but the requestor may be holding the name in abeyance for future use. To test for this, a representative sample of host machines are sent a probe (a “ping”), with a signal being sent back to the originating machine if the host was indeed found. The rate of growth of the number of hosts has been phenomenal; from a paltry 213 hosts in August 1981, the Internet has in excess of 60 million hosts now. The doubling time for the number of hosts is on the order of 18 months. Most of this growth has come from the commercial sector, capitalizing on the growing popularity of multimedia platforms for advertising and communications such as the World Wide Web.

CONNECTING TO THE INTERNET

Of course, before being able to use all of the resources that the Internet has to offer, one needs to actually make a physical connection between one’s own computer and “the information superhighway.” For purposes of this discussion, the elements of this connection have been separated into two discrete parts: the actual, physical connection (meaning the “wire” running from one’s computer to the Internet backbone) and the service provider, who handles issues of routing and content once connected. Keep in mind that, in practice, these are not necessarily treated as two separate parts—for instance, one’s service provider may also be the same company that will run cables or fibers right into one’s home or office.

Copper Wires, Coaxial Cables, and Fiber Optics

Traditionally, users who were attempting to connect to the Internet away from the office have had one and only one option—a modem, which uses the existing copper twisted-pair cables carrying telephone signals to transmit data. Data transfer rates using modems are relatively slow, allowing for data transmission in the range of 28.8 to 56 kilobits per second (Kbps). The problem with using conventional copper wire to transmit data lies not in the copper wire itself, but in the switches that are found along the way that route information to their intended destinations. These switches were designed for the efficient and effective transfer of voice data, but were never intended to handle the high-speed transmission of data. While most people still use modems from their homes, a number of new technologies are already in place and will become more and more prevalent for accessing the Internet away from hard-wired Ethernet networks. The maximum speeds at which each of the services that are discussed below can operate are shown in Figure 19.1.1.

The first of these “new solutions” is the integrated services digital network, or ISDN. While the advent of ISDN was originally heralded as the way to bring the Internet into the home in a speed-efficient manner, it required that special wiring be brought into the home. It also required that users be within a fixed distance from a central office, on the order of 20,000 feet or less. The cost of running this special, dedicated wiring, along with a per-minute pricing structure, effectively placed ISDN out of reach of most individuals. While ISDN is still available in many areas, this type of service is quickly being supplanted by more cost-effective alternatives.

In looking at alternatives that did not require new wiring, cable television providers began to look at ways in which the coaxial cable already running into a substantial number of households could be used to also transmit data. Cable companies are able to use bandwidth that is not being used to transmit television signals (effectively, unused channels) to push data into the home at very high speeds, up to 4.0 megabits per second (Mbps). The actual computer is connected to this network through a cable modem, which uses an Ethernet connection to the computer and a coaxial cable to the wall. Homes in a given area all share a single cable, in a wiring scheme very similar to that by which individual computers are connected via the Ethernet in an office or laboratory setting. While this branching

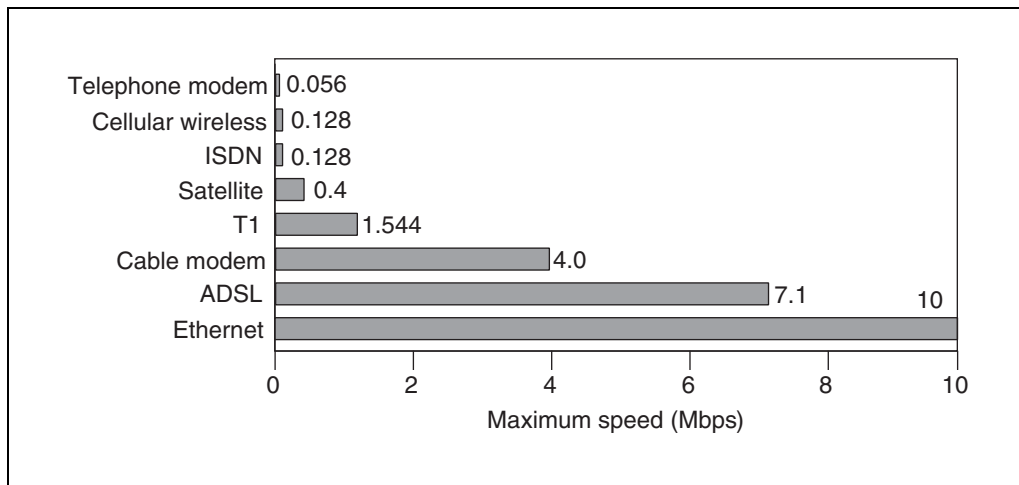


Figure 19.1.1 Performance of various types of Internet connections, by maximum throughput. The numbers indicated in the graph refer to peak performance; often the actual performance of any given method may be on the order of one-half slower, depending on configurations and system conditions.

arrangement can serve to connect a large number of locations, there is one major disadvantage—as more and more homes connect through their cable modems, service effectively slows down as more signal attempts to pass through any given node. One way of circumventing this problem is the installation of more switching equipment and reducing the size of a given “neighborhood.”

Since the local telephone companies were the primary ISDN providers, they quickly turned their attention to ways in which the existing, conventional copper wire already in the home could be used to transmit data at high speed. The solution here is the digital subscriber line, or DSL. By using new, dedicated switches that are designed for rapid data transfer, DSL providers can circumvent the old voice switches that slowed down transfer speeds. Depending on the user’s distance from the central office and whether a particular neighborhood has been wired for DSL service, speeds are on the order of 0.8 to 7.1 Mbps. The data transfers do not interfere with voice signals, and users can use the telephone while connected to the Internet; the signals are “split” by a special modem that passes the data signals to the computer and a microfilter that passes voice signals to the handset. There is a special type of DSL called **asynchronous DSL**, or ADSL. This is the variety of DSL service that is becoming more and more prevalent. Most home users download much more information than they send out, so systems are engineered to provide super-fast transmission in the “in” direction, with transmissions in the “out” direction being 5 to 10 times slower. Using this approach maximizes the amount of bandwidth that can be used without necessitating new wiring. One of the advantages of ADSL over cable is that ADSL subscribers effectively have a direct line to the central office, meaning that they do not have to compete with their neighbors for bandwidth. This, of course, comes at a price; at the time of this writing, ADSL connectivity options were on the order of twice as expensive as cable Internet.

Some of the newer technologies involve wireless connections to the Internet. These include using one’s own cell phone or a special cell phone service (such as Ricochet) to upload and download information. These cellular providers can provide speeds on the order of 28.8 to 128 Kbps, depending on the density of cellular towers in the service area. Fixed-point wireless services can be substantially faster, since the cellular phone does not have to “find” the closest tower at any given time. Along these same lines, satellite providers are also coming on-line. These providers allow for data download directly to a

satellite dish with a southern exposure, with uploads occurring through traditional telephone lines. While the satellite option has the potential to be amongst the fastest of the options discussed, current operating speeds are only on the order of 400 Kbps.

Content Providers versus ISPs

Once an appropriately fast and price-effective connectivity solution is found, users will then need to actually connect to some sort of service that will enable them to traverse the Internet space. The two major categories in this respect are **on-line services** or **Internet service providers (ISPs)**. On-line services such as America Online (AOL) and CompuServe offer a large number of interactive digital services, including information retrieval, electronic mail (e-mail), bulletin boards, and “chat rooms” where users who are on line at the same time can converse with each other about any number of subjects. While the on-line services now provide access to the World Wide Web (see discussion of The World Wide Web), most of the features and services available through these systems reside within a proprietary, closed network; once a connection is made between the user’s computer and the on-line service, accessing the special features, or content, of these systems does not require ever leaving the on-line system’s host computer. Specialized content can range from access to on-line travel reservation systems to encyclopedias that are constantly being updated—items that would not be available to anyone unless they subscribed to that particular on-line service.

Internet service providers, or ISPs, take the opposite tack. Instead of focusing on providing content, the ISPs provide the tools necessary for users to send and receive e-mail, upload and download files, and navigate around the World Wide Web to find information at remote locations. The major advantage of ISPs is connection speed; ISPs often provide faster connections than the on-line services. Most ISPs charge a monthly fee for unlimited use.

The line between on-line services and ISPs has already begun to blur. AOL’s now monthly flat fee pricing structure allows users to obtain all of the proprietary content found on AOL as well as all of the Internet tools available through ISPs, often at the same cost as a simple ISP connection. The extensive AOL network puts access to AOL as close as a local phone call in most of the United States, providing access to e-mail no matter where the user is located—a feature that small, local ISPs cannot match.

Not to be outdone, many of the major national ISP providers now also provide content through the concept of **portals**. Portals are Web pages that can be customized to the needs of the individual user and which serve as a jumping-off point to other sources of news or entertainment on the Net. In addition, many national firms such as Mindspring are able to match AOL’s ease of connectivity on the road, and both ISPs and online providers are becoming more and more generous in providing users the capacity to publish their own Web pages. Developments such as this, coupled with the move of local telephone and cable companies into providing Internet access through new, faster fiber optic networks, foretell major changes in how people will access the Net in the future, changes that should favor the end-user both in price and performance.

ELECTRONIC MAIL

Most people are introduced to the Internet through the use of electronic mail (**e-mail**). The use of e-mail has become practically indispensable in many settings owing to its convenience as a medium for sending, receiving, and replying to messages. Its advantages are many:

It is much quicker than postal, or “snail mail.” Messages tend to be much clearer and more to the point than is the case in typical telephone or face-to-face conversations.

Recipients have more flexibility in deciding whether a response needs to be sent immediately, relatively soon, or at all, giving individuals more control over workflow.

It provides a convenient method by which messages can be filed or stored.

There is little or no cost involved in sending an e-mail message.

While these and other advantages have pushed e-mail to the forefront of interpersonal communication in both industry and the academic community, users should be aware of several major disadvantages. First is the issue of security. As mail travels towards its recipient, it may pass through a number of remote nodes. The message could be intercepted and read at any one of those nodes by someone with high-level access, such as a systems administrator. Second is the issue of privacy. In industrial settings, e-mail is often considered to be an asset of the company for use only in official communication and, as such, is subject to monitoring by supervisors. The opposite is often true in academic, quasi-academic, or research settings; for example, National Institutes of Health policy encourages personal use of e-mail within the bounds of certain published guidelines. The key words here are “published guidelines”; no matter what the setting, users of e-mail systems should always be informed as to their organization’s policy regarding the appropriate use and confidentiality of e-mail so that they may use the tool properly and effectively. An excellent, basic guide to the effective use of e-mail is highly recommended (Rankin, 1996).

Sending E-Mail

E-mail addresses take the general form *user@computer.domain*, where *user* is the name of the individual user and *computer.domain* specifies the actual computer that the e-mail account is located on. Like a postal letter, an e-mail message is comprised of an **envelope** or **header**, showing the e-mail addresses of the sender and recipient, a line indicating the subject of the e-mail, and information about how the e-mail message actually travelled from the sender to the recipient. The header is followed by the actual message, or **body**, analogous to what would go inside the postal envelope. Figure 19.1.2 illustrates all the components of an e-mail message.

E-mail programs vary widely, depending on both the platform and the needs of the users. Usually the characteristics of the local area network (LAN) dictate what types of mail programs can be used, and the decision is often left to systems administrators rather than individual users. Among the most widely used e-mail packages with a graphical user interface are Eudora for the Macintosh and both Netscape Messenger and Microsoft Exchange for Macintosh, Windows, and UNIX platforms. Text-based e-mail programs, which are accessed by logging into a UNIX-based account, include Elm and Pine.

Bulk E-Mail

As with postal mail, there has been an upsurge in “spam” or “junk e-mail,” where companies compile bulk lists of e-mail addresses for use in commercial promotions. Since most of these lists are compiled from on-line registration forms and similar sources, the best defense for remaining off of these bulk e-mail lists is to be selective as to whom e-mail addresses are provided. Most newsgroups keep their mailing lists confidential; if in doubt and this is a concern, one should ask.

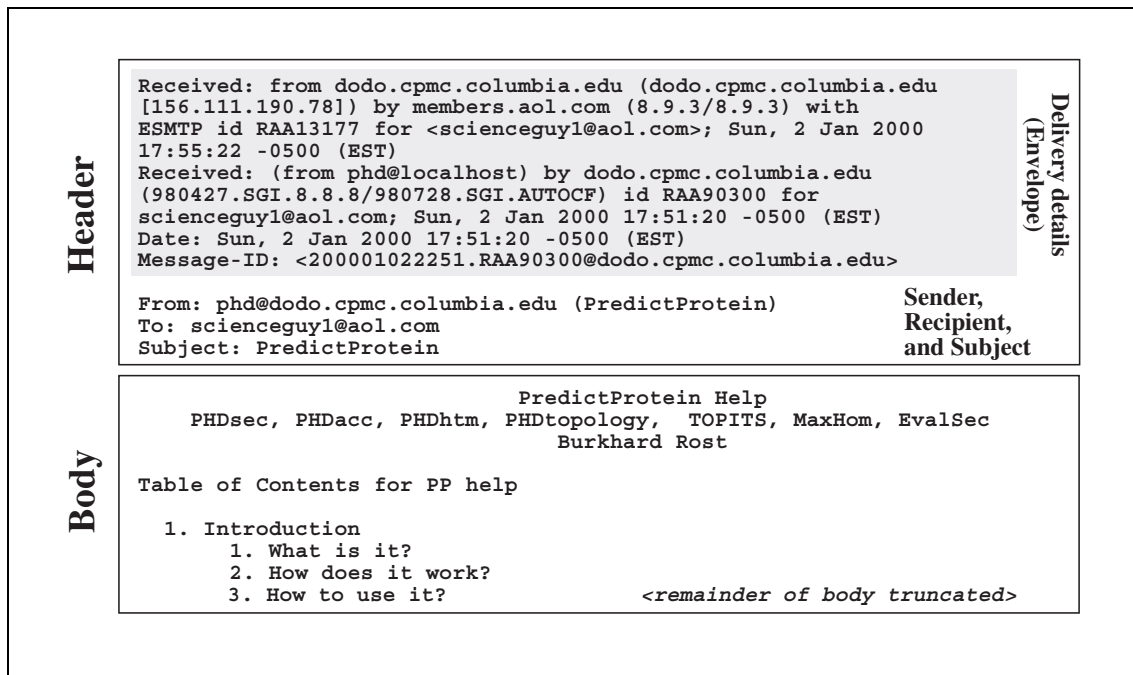


Figure 19.1.2 Anatomy of an e-mail message, with relevant components indicated. This message is an automated reply to a request for help file for the PredictProtein E-mail server.

E-mail Servers

Most often, e-mail is thought of as a way to simply send messages, whether it be to one recipient or many. It is also possible to use e-mail as a mechanism for making biological predictions or retrieving records from biological databases. Users can send e-mail messages in a predefined format, defining the action to be performed for remote computers known as **servers**; the servers will then perform the desired operation and e-mail back the results. While this method is not interactive (in that the user cannot adjust parameters or have control over the execution of the method in real time), it does place the responsibility of hardware maintenance and software upgrades on the persons maintaining the server, allowing users to concentrate on their results instead of on programming. For most servers, sending the message *help* to the server e-mail address will return a detailed set of instructions for using that server, including the way in which queries need to be formatted.

Aliases and Newsgroups

In the example in Fig. 19.1.2, the e-mail message is being sent to a single recipient. One of the strengths of e-mail is that a single piece of e-mail can be sent to a large number of people. The primary mechanism for doing this is through **aliases**; a user can define a group of people within their mail program and give the group a special name, or alias. Instead of using the individual e-mail addresses for all of the people in the group, the user can just send the e-mail to the alias name, and the mail program will handle broadcasting the message to each person in that group. Setting up alias names is a tremendous time-saver even for small groups; it also ensures that all members of a given group actually receive all e-mail messages intended for the group.

The second mechanism for broadcasting messages is through **newsgroups**. This model works slightly differently in that the list of e-mail addresses is compiled and maintained on a remote computer through subscriptions, much like magazine subscriptions. For example, the BIOSCI newsgroups are amongst the most highly-trafficked, offering a

forum for discussion or the exchange of ideas in a wide variety of biological subject areas. To begin receiving the messages posted to the automated sequencing discussion group within BIOSCI, a user would send a message to *biosci-server@net.bio.net* with the wording *subscribe autoseq* in the body of the message. The user would then receive all future postings to that group and be able to participate in the discussions. If a user wished to be removed from the group, a message would be sent to the same address, but this time, the body of the message would read *unsubscribe autoseq*. For more information on BIOSCI, including a complete list of discussion groups, an e-mail message can be sent to *biosci-server@net.bio.net*; in this case the subject line should be left blank and the words *info faq* typed in the body of the message. The BIOSCI server will then return a copy of the Frequently Asked Questions (FAQ) in response, with detailed information on each newsgroup overseen by BIOSCI.

It is also possible to participate in newsgroups without having each and every piece of e-mail flood into one's private mailbox. Instead, interested participants can use newsreading software, such as NewsWatcher for the Macintosh, which provides access to the individual messages making up a discussion. The major advantage is that the user can pick and choose which messages to read by scanning the subject lines; the remainder can be discarded by a single operation. NewsWatcher is an example of what is known as a **client-server application**—the client software (here, NewsWatcher) runs on a client computer (a Macintosh), which in turn interacts with a machine at a remote location (the server). Client-server architecture is interactive in nature, with a direct connection being made between the client and server machines.

Once NewsWatcher is started, the user is presented with a list of newsgroups available to them (Fig. 19.1.3); this list will vary, depending on the user's location, as systems administrators have the discretion to allow or block certain groups at a given site. From the rear-most window in the figure, the user double-clicks on the newsgroup of interest (here, *bionet.genome.arabidopsis*), which spawns the window shown in the center. At the top of the center window is the current unread message count, and any message within the list can be read by double-clicking on that particular line. This, in turn, spawns the last window (in the foreground), showing the actual message. If a user decides not to read any of the messages, or is done reading individual messages, the balance of the messages within the newsgroup (center) window can be deleted by first choosing *Select All* from the *File* menu, then selecting *Mark Read* from the *News* menu. Once the newsgroup window is closed, the unread message count is reset to zero. Every time NewsWatcher is restarted, it will automatically poll the news server for new messages that have been created since the last session.

As with most of the tools that will be discussed in this unit, news-reading capability is built into Web browsers such as Netscape Navigator and Microsoft Internet Explorer.

FILE TRANSFER PROTOCOL

Despite the many advantages afforded by e-mail in transmitting messages, experienced e-mail users have no doubt experienced frustration in trying to transmit files (**attachments**) along with an e-mail message. The mere fact that a file can be attached to an e-mail message and sent does not mean that the recipient will be able to detach, decode, and actually use the attached file. While more cross-platform e-mail packages such as Microsoft Exchange are being developed, the use of different e-mail packages by people at different locations means that sending files via e-mail is not an effective, foolproof method, at least in the short term. One solution to this problem is through the use of a **file transfer protocol** (or **FTP**). The workings of FTP are quite simple—a connection is made

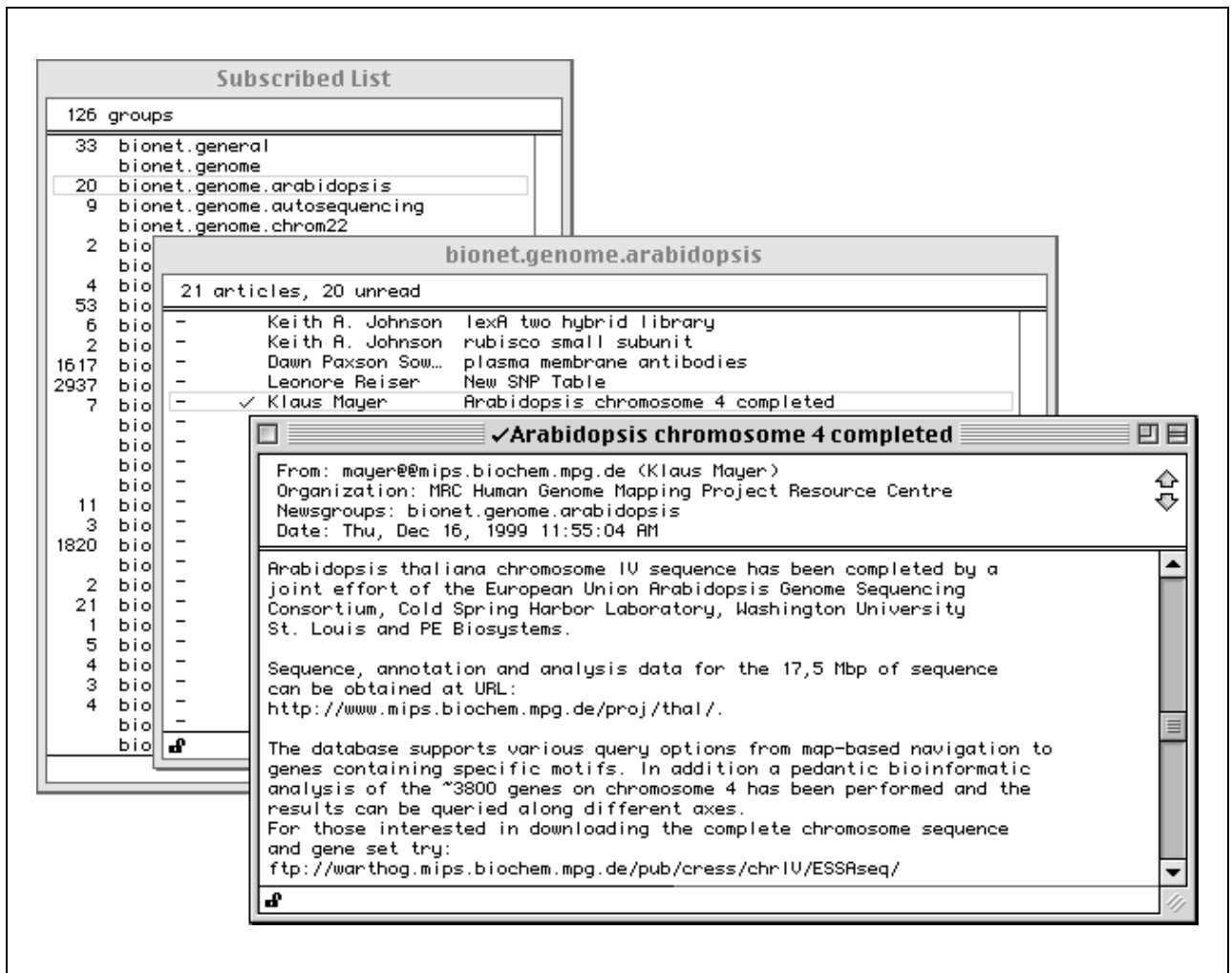


Figure 19.1.3 Using NewsWatcher to read postings to newsgroups. The list of newsgroups that the user has subscribed to is shown in the *Subscribed List* window (*left*). The list of new postings for the highlighted newsgroup (*bionet.genome.arabidopsis*) is shown in the center window. The window in the foreground shows the contents of the posting selected from the center window.

between a user's computer (the **client**) and a remote server, and that connection remains in place for the duration of the FTP session. File transfers are very fast, at rates on the order of 5 to 10 kilobytes per second, with speeds varying with time of day, distance between the client and server machines, and overall traffic on the network.

In the ordinary case, making an FTP connection and transferring files requires that a user have an account on the remote server. However, there are many files and programs that the academic community makes freely available, and access to those files does not require having an account on each and every machine where these programs are stored. Instead, connections are made using a system called **anonymous FTP**. Under this system, the user connects to the remote machine, and instead of entering a username/password pair, types *anonymous* as the username and enters an e-mail address in place of a password. Providing one's e-mail address allows the server's systems administrator to compile access statistics which may, in turn, be of use to those actually providing the public files or programs. An example of an anonymous FTP session using UNIX is shown in Fig. 19.1.4.

```

$ ftp ftp.bio.indiana.edu
Connected to magpie.bio.indiana.edu.
220 iubio.bio.indiana.edu FTP server ready.
Name: anonymous
331 Guest login ok, send your complete e-mail address as password.
Password: *****
230-                Welcome to IUBio archive!
230-
230- This is a user-supported archive for biology software and
data.
230-
230-                See the file Archive.Doc for details of this archive.
230-
230-                See IUBio Bio-Mirror archive of large data sets at
230- ftp to iubio.bio.indiana.edu, user: iubio, password:iubio
230- This includes GenBank, EMBL and DDBJ and other biosequence
data.
230-
230-                Report problems, uploads and other matters via e-mail to
230-                archive@bio.indiana.edu.
230-
230 Guest login ok, access restrictions apply.
Remote system type is UNIX.
Using binary mode to transfer files.
ftp> cd /molbio/align/clustal
250 CWD command successful.
ftp> get clustalw1.75.unix.tar.Z
local: clustalw1.75.unix.tar.Z remote: clustalw1.75.unix.tar.Z
200 PORT command successful.
150 Opening BINARY mode data connection for
clustalw1.75.unix.tar.Z (230379 bytes).
226 Transfer complete.
230379 bytes received in 0.45 seconds (500.75 Kbytes/s)
ftp> quit
221-You have transferred 230379 bytes in 1 files.
221-Total traffic for this session was 231859 bytes in 1
transfers.
221-Thank you for using the FTP service on iubio.bio.indiana.edu.
221 Goodbye.

```

Figure 19.1.4 Using UNIX FTP to download a file. An anonymous FTP session is established with the molecular biology FTP server at the University of Indiana to download the ClustalW alignment program. The user inputs are shown in boldface.

Although FTP occurs within the UNIX environment, Macintosh and PC users can employ programs that utilize graphical user interfaces (GUI, pronounced “gooey”) to navigate through the UNIX directories on the FTP server. Users need not have any knowledge of UNIX commands to download files; they can instead rely on pop-up menus and the ability to point-and-click their way through the UNIX file structure. The most popular FTP program on the Macintosh platform for FTP sessions is Fetch.

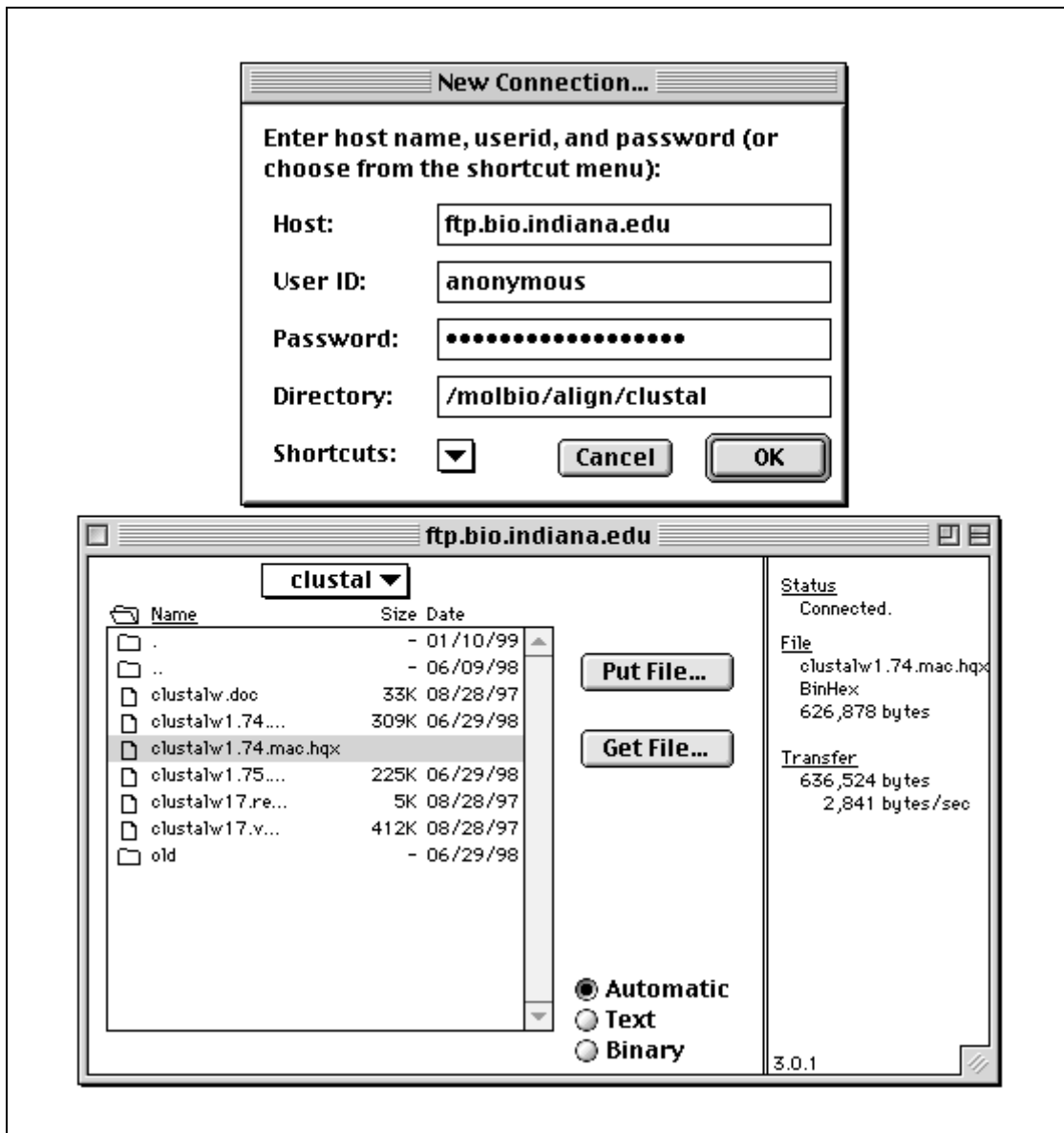


Figure 19.1.5 Using Fetch to download a file. An anonymous FTP session is established with the molecular biology FTP server at the University of Indiana (top) to download the ClustalW alignment program (bottom). Notice the difference between this GUI-based program and the UNIX equivalent illustrated in Figure A.3J.4.

A sample Fetch window is shown in Figure 19.1.5 to illustrate the difference between using a GUI-based FTP program and the equivalent UNIX FTP in Figure 19.1.4. In the figure, notice that the *Automatic* radio button (near the bottom of the second window under the *Get File* button) is selected, meaning that Fetch will determine the appropriate type of file transfer to perform. This may be manually overridden by selecting either *Text* or *Binary*, depending on the nature of the file being transferred. As a rule, text files should be transferred as *Text*, programs or executables as *Binary*, and graphic format files such as PICT and TIFF files as *Raw Data*.

THE WORLD WIDE WEB

While FTP is of tremendous use in the transfer of files from one computer to another, it does suffer from some limitations. When working with FTP, once a user enters a particular directory, they can only see the names of the directories or files. In order to actually view

what is within the files, it is necessary to physically download them onto one's own computer. This inherent drawback led to the development of a number of **distributed document delivery systems (DDDS)**, interactive client-server applications that allowed information to be viewed without having to perform a download. The first generation of DDDS development led to programs like Gopher, which allowed plain text to be viewed directly through a client-server application. From this evolved the most widely known and widely used DDDS, namely the World Wide Web. The Web is an outgrowth of research performed at the European Nuclear Research Council (CERN) in 1989 that was aimed at sharing research data between several locations. That work led to a medium through which text, images, sounds, and videos could be delivered to users on demand, anywhere in the world.

Navigation on the World Wide Web

Navigation on the Web does not require advance knowledge of the location of the information being sought. Instead, users can navigate by clicking on specific text, buttons, or pictures. These clickable items are collectively known as **hyperlinks**. Once one of these hyperlinks is clicked, the user is taken to another Web location, which could be at the same site or halfway around the world. Each document displayed on the Web is called a **Web page**, and all of the related Web pages on a particular server are collectively called a **Web site**. Navigation strictly through the use of hyperlinks has been nicknamed "Web surfing."

Users can take a more direct approach to finding information by entering a specific address. One of the strengths of the Web is that the programs used to view Web pages (appropriately termed **browsers**) can be used to visit Gopher and FTP sites as well, somewhat obviating the need for separate Gopher or FTP applications. As such, a unified naming convention was introduced to indicate to the browser program both the location of the remote site and, more importantly, the type of information at that remote location so that the browser could properly display the data. This standard-form address is known as a **uniform resource locator (URL)**, and takes the general form *protocol://computer.domain*, where *protocol* specifies the type of site and *computer.domain* specifies the location (Table 19.1.2). The *http* used for the protocol in World Wide Web URLs stands for *hypertext transfer protocol*, the method used in transferring Web files from the host computer to the client.

Browsers

Browsers, which are used to look at Web pages, are client-server applications that connect to a remote site, download the requested information at that site, display the information on the user's monitor, then disconnect from the remote host. The information retrieved from the remote host is in a platform-independent format called **hypertext markup language (HTML)**. HTML code is strictly text-based, and any associated graphics or sound for that document exist as separate files in a common format. For example, images

Table 19.1.2 Uniform Resource Locator (URL)
Format for Each Type of Transfer Protocol

Site	URL format (example)
General form	<i>protocol://computer.domain</i>
FTP site	<i>ftp://ftp.ncbi.nlm.nih.gov</i>
Gopher site	<i>gopher://gopher.iubio.indiana.edu</i>
Web site	<i>http://www.nhgri.nih.gov</i>

may be stored and transferred in GIF format, developed by CompuServe for the quick and efficient transfer of graphics; while GIF format is most commonly used for graphics, other formats such as JPEG and BMP may also be used. Because of this, a browser can display any Web page on any type of computer, whether it be a Macintosh, IBM-compatible, Linux, or UNIX machine. The text is usually displayed first, then the remaining elements are placed on the page as they are downloaded. With minor exceptions, a given Web page will look the same when the same browser is used on any of the above platforms.

The two major players in the area of browser software are Netscape, with their Communicator product, and Microsoft, with Internet Explorer. As with many other areas where multiple software products are available, the choice between Netscape and Internet Explorer comes down to one of personal preference. While the computer literati will debate the fine points of difference between these two packages, for the average user, both packages perform equally well and offer the same types of features, adequately addressing the Web-browser needs of most users.

It is worth mentioning that, while the Web is by definition a visually based medium, it is also possible to travel through Web space and view documents without the associated graphics. For users limited to line-by-line terminals, a browser called Lynx is available. Developed at the University of Kansas, Lynx allows users to use their keyboard arrow keys to highlight and select hyperlinks, using their return key the same way that Netscape and Internet Explorer users would click their mouse.

Internet versus Intranet

The World Wide Web is normally thought of as a way to communicate with people at a distance, but the same infrastructure can be used to connect people within an organization. Such **intranets** provide an easily accessible repository of relevant information, capitalizing on the simplicity of the Web interface. It also provides another channel for broadcast or confidential communication within the organization. Having an intranet is of particular value when members of an organization are physically separated, whether it be in different buildings or different cities. Intranets are protected in such a way that people who are not on the organization's network are prohibited from accessing the internal Web pages; additional protections through the use of passwords are also common.

Finding Information on the World Wide Web

Most people find information on the Web the old fashioned way—by word of mouth either by using lists such as Table 19.1.3 or by simply following hyperlinks put in place by Web authors. Continuously clicking from page to page can be a highly ineffective way of finding information, especially when the information sought is of a very focused nature. One way of finding interesting and relevant Web sites is to consult **virtual libraries**, which are curated lists of Web resources arranged by subject. Virtual libraries of special interest to biologists include the WWW Virtual Library, maintained by Keith Robison at Harvard and the EBI BioCatalog, based at the European Bioinformatics Institutes. The URLs for these sites can be found in Table 19.1.3.

It is also possible to directly search the Web by using **search engines** such as Alta Vista and Excite, among others. A search engine is simply a specialized program that can perform full-text or keyword searches on databases that catalog Web content. The result of a search is a hyperlinked list of Web sites fitting the search criteria from which the user can visit any or all of the found sites. However, search engines use slightly different methods in compiling their databases. One variation is the attempt to capture most or all of the text of every Web page that the search engine is able to find and catalog (“Web crawling”). Another technique is to catalog only the title of each Web page rather than its

Table 19.1.3 World Wide Web Sites of Interest

Site	URL
Domain names	
gTLD-MOU	http://www.gtld-mou.org
Internet Software Consortium	http://www.isc.org
Electronic mail and newsgroups	
BIOSCI Newsgroups	http://www.bio.net/docs/biosci.FAQ.html
Eudora	http://www.eudora.com
Microsoft Exchange	http://www.microsoft.com/exchange/
NewsWatcher	ftp://ftp.acns.nwu.edu/pub/newswatcher/
File Transfer Protocol	
Fetch 3.0/Mac	http://www.dartmouth.edu/pages/softdev/fetch.html
FTP Voyager	http://ftpvoyager.deerfield.com
Internet access	
America Online	http://www.aol.com
AT&T	http://www.att.com/worldnet
Bell Atlantic	http://www.bellatlantic.net
Bell Canada	http://www.bell.ca
CompuServe	http://www.compuserve.com
MCI	http://www.mci.com
Ricochet	http://www.ricochet.net
Telus	http://telus.com
Virtual libraries	
EBI BioCatalog	http://www.ebi.ac.uk/biocat/biocat.html
Amos' WWW Links Page	http://www.expasy.ch/alinks.html
NAR Database Collection	http://www3.oup.co.uk/nar/Volume_28/Issue_01/introduction/
WWW Virtual Library	http://mcb.harvard.edu/BioLinks.html
World Wide Web browsers	
Internet Explorer	http://www.microsoft.com/insider/ie5/default.htm
Lynx	ftp://ftp2.cc.ukans.edu/pub/lynx
Netscape Navigator	http://home.netscape.com
World Wide Web search engines	
AltaVista	http://www.altavista.com
Excite	http://www.excite.com
HotBot	http://hotbot.lycos.com
Infoseek	http://infoseek.go.com
Lycos	http://www.lycos.com
Northern Light	http://www.northernlight.com
World Wide Web meta-search engines	
MetaCrawler	http://www.metracrawler.com
Savvy Search	http://www.savvysearch.com

Table 19.1.4 Number of Hits Returned for Four Defined Search Queries on Some of the More Popular Search and Meta-Search Engines

Search term	Search Engine			Meta-search engine		
	HotBot	Excite	Infoseek	Lycos	MetaCrawler	SavvySearch
genetic mapping	478	1,040	4,326	9,395	62	58
human genome	13,231	34,760	15,980	19,536	42	54
positional cloning	279	735	1,143	666	40	52
prostate cancer	14,044	53,940	24,376	33,538	60	57

entire text. A third is to consider words that must appear next to each other or only relatively close to one another. Because of these differences in search-engine algorithms, the results returned by issuing the same query to a number of different search engines can produce wildly different results (Table 19.1.4). It may also be noted from Table 19.1.4 that most of the numbers are exceedingly large, reflecting the overall size of the World Wide Web. Unless a particular search engine ranks its results by relevance (for example, by scoring words in a title higher than words in the body of the Web page), the results obtained may not be particularly useful. It is also necessary to keep in mind that, depending on the indexing scheme that the search engine is using, the found pages may actually no longer exist, leading the user to the dreaded “404 Not Found” error.

Compounding this problem is the issue of **coverage**—the number of Web pages that any given search engine is actually able to survey and analyze. A comprehensive study by Lawrence and Giles (1998) indicates that the coverage provided by any of the search engines studied is both small and highly variable. For example, the HotBot engine produced 57.5% coverage of what was estimated to be the size of the “indexable Web,” while Lycos had only 4.41% coverage, a full order of magnitude less than HotBot. The most important conclusion from this study was the extent of coverage increased as the number of search engines are increased and the results from those individual searches are combined. Combining the results obtained from the six search engines examined in this study produced coverage approaching 100%.

To address this point, a new class of search engines called **meta-search engines** have been developed. These programs will take the user’s query and poll anywhere from five to ten of the “traditional” search engines. The meta-search engine will then collect the results, filter out duplicates, and return a single, annotated list to the user. One big advantage is that the meta-search engines take relevance statistics into account, returning much smaller lists of results. Even though the hit list is substantially smaller, it is much more likely to contain sites that directly address the original query. Since the programs must poll a number of different search engines, searches conducted this way obviously take longer to perform, but the higher degree of confidence in the compiled results for a given query outweighs the extra few minutes (sometimes only seconds) of search time. Reliable and easy-to-use meta-search engines include SavvySearch and MetaCrawler (see Table 19.1.3).

LITERATURE CITED

- Froehlich, F. and Kent, A. 1991. ARPANET, the Defense Data Network, and Internet. *In* Encyclopedia of Communications. Marcel Dekker, New York.
- Krol, E. and Klopfenstein, B.C. 1996. The Whole Internet User’s Guide and Catalog. O’Reilly and Associates, Sebastopol, Calif.
- Kennedy, A.J. 1999. The Internet: Rough Guide 2000. Rough Guides, London.

- Lawrence, S. and Giles, C.L. 1998. Searching the World Wide Web. *Science* 280:98-100.
- Rankin, B. 1996. Dr. Bob's Painless Guide to the Internet and Amazing Things You Can Do with E-mail. No Starch Press, San Francisco.
- Quarterman, J. 1990. The Matrix: Computer Networks and Conferencing Systems Worldwide. Digital Press, Bedford, Mass.
-

Contributed by Andreas D. Baxevanis
National Human Genome Research Institute, NIH
Bethesda, Maryland

B.F. Francis Ouellette
Centre for Molecular Medicine and Therapeutics
University of British Columbia
Vancouver, British Columbia

Sequence Databases: Integrated Information Retrieval and Data Submission

UNIT 19.1 discusses general tools for electronic information retrieval, giving only brief glimpses of the databases of interest to molecular biologists. This unit provides a more thorough overview of biomedical information resources, focusing on sequence data, structure information, and the associated literature—and also discusses how nucleotide sequence data gets into the databases in the first place—an often neglected but nevertheless important topic. UNIT 19.3 covers sequence homology searching in detail.

First it is important to appreciate not only the variety of available data but also how rapidly it is expanding (Fig. 19.2.1A). The biomedical literature, as represented by MEDLINE, is growing at a rate of 400,000 articles per year. MEDLINE, produced by the U.S. National Library of Medicine (NLM), currently contains bibliographic information and journal abstracts from more than 11,000,000 publications indexed from ~4,300 biomedical journals.

Genetic and physical mapping data are expanding rapidly, as are protein crystal and NMR structures; DNA and protein sequence data have been growing exponentially for years (Fig. 19.2.1B to D). Even faster growth rates are anticipated for cDNA and concerted genome sequencing as technology advances. In mid-1997 complete sequences of *Saccharomyces cerevisiae* and seven bacterial genomes were available and the sequencing phase of the Human Genome Project was initiated. By January 2000, there were 24 bacterial and archaeal genomes completed and another 70 in progress. *Drosophila melanogaster* and *Caenorhabditis elegans* were both essentially done. Other eukaryotes for which substantial genomic sequencing had been done included human, mouse, *Arabidopsis thaliana*, rice, corn and the malaria parasite (*Plasmodium falciparum*).

GenBank is a comprehensive repository of sequence data and associated annotation, built and distributed by the National Center for Biotechnology Information (NCBI) at the NLM in Bethesda, Maryland. GenBank is part of an international collaboration with the DNA Database of Japan (DDBJ) in Mishima and the European Molecular Biology Laboratory (EMBL) Data Library at the European Bioinformatics Institute near Cambridge, United Kingdom. Because these partners exchange data daily, what we refer to here as simply “GenBank” actually contains sequences from GenBank, EMBL, and DDBJ.

In the 1990s, new types of sequence data became a major part of the database entries. For example, “expressed sequence tags” or ESTs are partial sequences that are derived from automated, “single-pass” sequencing on clones that are randomly selected from cDNA libraries, usually with the intent of surveying expressed genes (Adams et al., 1991; Kahn et al., 1992; Okubo et al., 1992; Waterston et al., 1992). These data present certain analytic challenges (Boguski et al., 1993) that are dealt with in UNIT 19.3, and they also require special procedures for batch submission to sequence databases (see discussion of submitting expressed sequence tags (EST), sequence tag site (STS), or genome survey sequence (GSS) data). Short genomic sequences, including those derived from “exon trapping” or “exon amplification” and genomic survey sequences (Church et al., 1993; Smith et al., 1994) present analytical challenges similar to those for ESTs. There are special procedures for submission of these “GSS” entries, similar to the procedures for EST submission. Large-scale sequencing of complete genomes has required additional methods for submission and has spurred development of new ways to access and display the data.

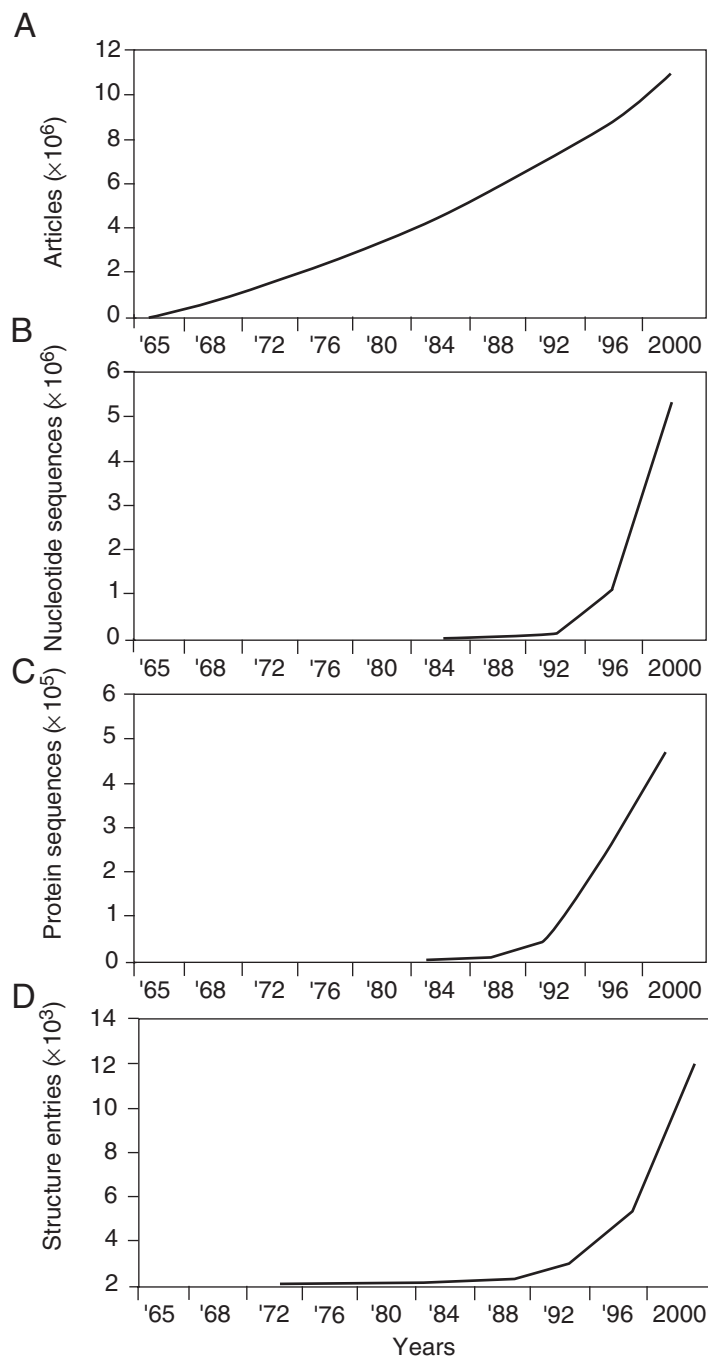


Figure 19.2.1 The cumulative growth of biomedical research data. **(A)** Growth of MEDLINE. **(B)** Growth in the number of nucleotide sequences in GenBank (Benson et al., 2000). **(C)** Growth in the number of protein sequences in the “nonredundant” set of proteins in GenBank. **(D)** Growth of protein and nucleic acid three-dimensional structures represented in the Brookhaven Protein Data Bank, April 2000 (<http://www.rcsb.org/pdb/>). These figures occasionally induce panic or despair in individuals who fear that it is impossible to make sense of so much data; however, modern information retrieval systems such as *Entrez* (Fig. 19.2.2) have succeeded in integrating this data, reducing redundancy, and providing powerful and convenient user interfaces.

Retrieval of records by electronic mail and anonymous FTP are covered in *UNIT 19.1*, as well as access to databases via the World Wide Web. Database homology searching is described in detail in *UNIT 19.3*. The present unit focuses on the wide variety of sequence, structure, and bibliographic databases—as well as how they have been integrated (Fig. 19.2.2) into a single, easy-to-use retrieval environment called *Entrez*, which provides users with “one stop shopping” over the Internet.

INTRODUCTION TO *ENTREZ*

Entrez (Fig. 19.2.2) ties together a diverse set of information resources: it accesses nucleotide and protein sequences from a number of databases as well as genome data from the NCBI Genomes division, three-dimensional structures from MMDB (Molecular Modeling Database of NCBI), human disease data from OMIM (Online Mendelian Inheritance in Man), genetic locus data from LocusLink, and bibliographic citations from PubMed (discussed below). *Entrez* is available through the World Wide Web (<http://www.ncbi.nlm.nih.gov/Entrez/>) and in a client/server version (available by FTP, see <http://ncbi.nlm.nih.gov/Entrez/Network/nentrez.overview.html> or ftp://ncbi.nlm.nih.gov/entrez/README_1). In the case of the client/server application, Network *Entrez* (*NetEntrez*), the client runs on the user’s local machine and interacts with the server at NCBI. Versions of the client program are available for Macintosh, Microsoft Windows, Unix/X-windows and some other systems. *Entrez* for the World Wide Web (*WebEntrez*) works through standard WWW browsers. *NetEntrez* and *WebEntrez* have similar func-

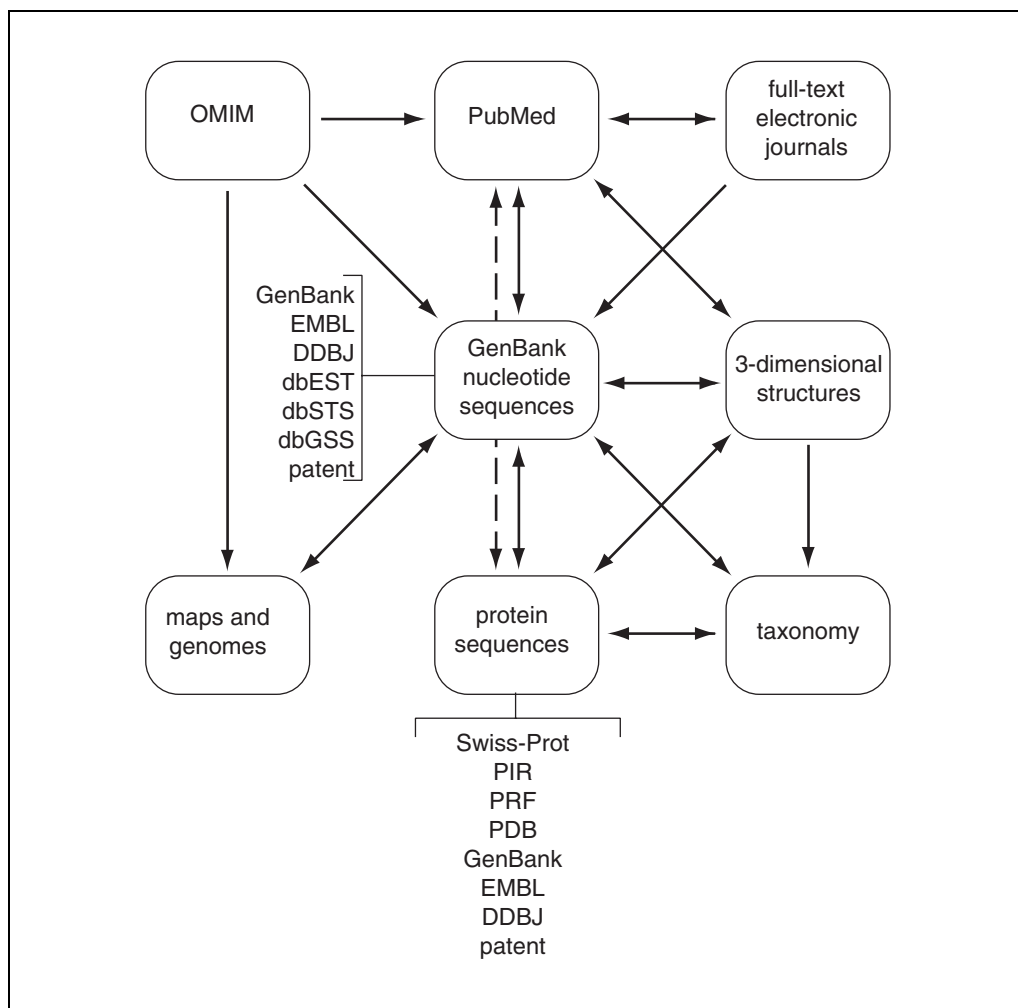


Figure 19.2.2 Data sources and links that constitute the *Entrez* system.

tionality, but there are differences in the way that they are used. *WebEntrez* has the advantage of additional links—e.g., to WebBLAST outputs and to full-text articles for some journals. The *NetEntrez* interface is more intuitive and is often easier to use for complex queries. *Entrez* can be used in many ways and some common scenarios are presented below.

Suppose we have just done a similarity search and have observed a number of promising matches on the “hit list.” The information on this list is often insufficient to allow us to reach even preliminary conclusions about the biological implications of our discovery. Thus, we want to find the publication describing this sequence, which meant, until recently, that we had to use a separate retrieval system to obtain the database record corresponding to the sequence. In this database record we would find the literature citation, and then we would have to take a trip down to the library to actually pull the printed copy of the journal. The “activation energy” for this process is high enough that one sometimes decides it is not worth it for “borderline” matches. Even if our motivation is high enough, perhaps there is no literature citation in the database record, or it involves a journal to which our library does not subscribe. In addition to finding the corresponding publication to confirm our search results, we may also want to retrieve some of the matching sequences and perform additional searches and alignments with these. All together, this is quite an involved process.

Entrez substantially lowers the activation energy by combining all of these activities in an easy-to-use system with a mouse-driven, graphical user interface. *Entrez* establishes all of the requisite links between DNA sequences, the proteins they encode, and the literature that describes their biology. The literature component of *Entrez* comes from the PubMed database, which includes abstracts. While an abstract is not a substitute for the full text of a paper, one can frequently go an amazingly long way in the analysis and interpretation of sequence homology results by the information contained in an abstract. For an increasing list of journals, the Web version of *Entrez* also has links to full-text articles. For many human disease genes, there are also links from PubMed records to the description of the disease in the Online Mendelian Inheritance in Man (OMIM) database. The PubMed database of bibliographic information contains citations from MEDLINE and PreMEDLINE (preliminary citations that eventually become MEDLINE records).

Entrez not only contains explicit links between different data sources, but also implicit or computed links within the data itself. Daily, the entire set of DNA and protein sequence databases are compared among one another and all significant homologies are computed and stored in the system. Thus, it is possible to instantly retrieve the homologs of any sequence in *Entrez* with the click of a mouse, without having to manually perform a new database search. Thus *Entrez* contains within it thousands of answers to questions that have not yet been asked. In one interesting case, this automated system actually “discovered” a significant homology prior to its recognition by a human biologist (Boguski and McEntyre, 1994).

In addition to precomputed sequence homologies, *Entrez* performs a similar operation on PubMed records. It is possible to compute a statistical relationship between two articles based on their frequency of use of significant terms. To indicate what constitutes a significant term, consider two negative examples: i.e., the terms “novel” and “important” are not significant terms because they appear in most scientific papers and thus are not of use in distinguishing one paper from another in terms of specific content. The set of “related articles” (sometimes referred to as “neighbors”) is recomputed daily. This constitutes a very powerful information resource. You can enter *Entrez* via an author name or keyword and retrieve not only a specific paper but an entire bibliography of related material. You can then branch out (or “neighbor out”) to other areas of the information

space and effortlessly go back and forth between the literature and the sequences. Specific examples have appeared in reviews (Baxevanis et al., 1997; Cockerill, 1994; Harper, 1994; Schuler et al., 1996). Further examples can also be found on the NCBI Coffee Break web site (<http://www.ncbi.nlm.nih.gov/Coffeekbreak/>). These pages describe recent biological discoveries and include tutorials demonstrating the linkage between various components of *Entrez* and other bioinformatics tools.

Finding information on sequence homologs using the Web version of *Entrez* is made even easier by the connections between *Entrez* and Web-based BLAST output. For example, it is possible to do a BLAST search using WebBLAST, and the hit list will have hypertext links to protein or nucleotide entries in *Entrez*.

Other features of *Entrez* include taxonomy-based sequence retrieval; from the Web page for the NCBI Taxonomy Browser you can scan through organism names and phylogenetic trees and select the full set of nucleotide or protein sequences for any species. There are now links in *Entrez* from sequences and the literature to protein three-dimensional structures in NCBI's Molecular Modeling Database (MMDB). A three-dimensional structure viewer called Cn3D has also been added to the system. Cn3D can be run as a client/server program or as a "Helper" application integrated into a standard Web browser.

Another special viewer is available for the *Entrez* Genomes division. The genome-level views presented by *Entrez* are built from sequences of complete chromosomes, from composites of sequence fragments, and from integrated genetic and physical maps. *Entrez* allows the user to visualize the sequence information at varying levels of detail, either graphically or as text. The chromosome views are tightly linked to the other *Entrez* component databases, allowing the user to jump effortlessly between maps, sequences, and bibliographic components.

DATA SUBMISSION: GENERAL CONSIDERATIONS

During the late 1970s, when DNA sequencing technology was still new, the *Journal of Biological Chemistry* required that any submitted manuscript containing sequence data be accompanied by the actual autoradiograms, so that reviewers could check the accuracy of the sequence data. We have come a long way since then. Currently, journals prefer not to even publish sequence data because it takes up a great deal of space and is much more usable in electronic form. Most journals, however, require proof in the form of a database accession number that the authors of a publication have sent their data to GenBank, EMBL, or DDBJ. Accession numbers are unique identifiers (consisting of one letter followed by five digits or two letters followed by six digits) that serve as confirmation that the data has been submitted and allow the scientific community to retrieve it. Making sequences available by depositing them in the public databases is recognized as an important responsibility of researchers to the community.

GenBank, the EMBL Data Library, and the DDBJ are partners in the international database collaboration; authors may submit their sequences to whichever database is most convenient without regard to where papers describing those sequences may be published. Sequences should be submitted to a single site; data submitted to one site are exchanged with the others on a daily basis.

Sequences by themselves have little meaning; hence, thorough, accurate biological annotation is very important. One important source of annotation is the published article referring to the sequence; indeed, NCBI has made every effort to link citations to the sequences via MEDLINE and *Entrez*. However, electronic annotation obtained from authors at the time of submission is becoming even more critical as the amount of data

grows (Fig. 19.2.1B), because the only realistic way of searching through this much data is by computer. The database staff are very helpful in assisting submitters in the proper annotation of their data.

Sequences submitted to any of the sequence databases go through a series of validation checks conducted by the database staff. Accession numbers are returned to the authors, usually within 24 hr, or less, if there are no problems with the submission(s). The accession number should be included in any manuscript that deals with the sequence, preferably in a footnote on the first page of the article (or in the manner prescribed by the individual journal). The full text of the provisional record is passed back to the authors for their review prior to being released. Authors may request that the data be kept confidential until publication. See Internet Resources for URLs and e-mail addresses of the sequence databases.

SUBMITTING A SEQUENCE TO THE NUCLEOTIDE DATABASE

Submission Methods

As mentioned above, a sequence can be submitted to any of the three sequence databases (see discussion of Data Submission: General Considerations; also see Internet Resources). A submission to GenBank is described in this section, but readers should keep in mind that the example could apply to DDBJ or EMBL submissions. Researchers should submit their sequences to whichever database is most convenient. The important point is to only submit to one database.

If the researcher has access to the World Wide Web, the easiest way to submit data to GenBank is to use the BankIt form on the NCBI home page (or for DDBJ and EMBL submissions, use Sakura or WebIn, respectively; see Internet Resources). BankIt (<http://www.ncbi.nlm.nih.gov/BankIt/>) allows one to simply copy and paste most of the submission information and data directly out of one's word processor or sequence-analysis package into an electronic form. BankIt was designed to handle the most common types of GenBank submissions, such as mRNAs or short genomic records.

NCBI also has developed a platform-independent submission program called Sequin, which runs stand-alone or over the network. Sequin is suitable for a wide range of sequence lengths and complexities, including traditional (gene-sized) nucleotide sequences, segmented entries (e.g., genomic sequences of a spliced gene where not all of the intronic sequences have been determined), long (genome-sized) sequences with many annotated features, and sets of related sequences (i.e., population, phylogenetic, or mutation studies of a particular gene, region, or viral genome). It also has a number of built-in validation functions for enhanced quality assurance. The validator checks such things as missing organism information, incorrect coding-region lengths (compared to the attached protein sequence), internal stop codons in coding regions, mismatched amino acids, or nonconsensus splice sites. Sequin can present various views of the data, including GenBank-, FASTA-, and EMBL-formatted files, and a graphical view of the sequence including features and alignments. A sequence editor is built into Sequin; this editor automatically adjusts feature intervals as the sequence is edited. Additional capabilities include sequence-analysis functions—e.g., a coding region translator, an ORF finder, and a function that will replace a sequence with the sequence of the complementary strand and simultaneously adjust the positions of the features.

Sequin is available from NCBI by anonymous FTP (UNIT 19.1); see <http://www.ncbi.nlm.nih.gov/Sequin/index.html> for further details on Sequin, including full detailed documentation.

BankIt is best suited for submitting one or a small number of sequences. Sequin provides more of a data “workbench” environment that can accommodate large as well as small contigs with detailed annotation that can be added over time and which can finally be used to submit data to GenBank. In addition to these methods, NCBI provides streamlined submission procedures for large data sets—e.g., expressed sequence tags (ESTs), sequence-tagged sites (STSs), genome survey sequences (GSSs), and high-throughput genome sequences (HTGSs). NCBI provides a means for interacting with the laboratory information-management systems of genome-sequencing centers to ensure the efficient, timely, and accurate submission of their data. These options are discussed in more detail in subsequent sections (see discussion of Submitting EST, STS, or GSS Data, and Submitting High-Throughput Genome Sequences).

Instructions and Tips for Preparing Sequence Submissions

The most important item is, of course, the new DNA sequence itself. It should be free of sequencing errors, cloning artifacts, etc.; standard things to check for in this regard are described below. In the special case of “single-pass” sequences such as ESTs, the data often contain ambiguities and sequencing errors; in these cases it is important to note it as “single-pass” data.

What does one need to know about a sequence to submit it to GenBank? The sequence record contains references to the article in which the sequence is published, information on the submitters, and information on the source organism. The record also should describe the sequence as fully as possible. Significant regions of the sequence, such as coding region, promoter region, and transcription boundaries, are labeled with defined feature names. These features are detailed in the GenBank Feature Table document, available by anonymous FTP (UNIT 19.1) from <ftp://ncbi.nlm.nih.gov> in the directory */genbank/docs*, or on the Web at <http://www.ncbi.nlm.nih.gov/collab/FT/index.html>. This list of features (along with their syntax and qualifiers) is somewhat overwhelming, and it is much easier to familiarize oneself with this system by simply looking at some well-done examples of existing GenBank records (see Appendix to this unit).

The DEFINITION line (see Appendix to this unit for illustrations of this and other elements of the GenBank record) is a short description of the biological sequence in the record and the organism from which it was derived. The LOCUS field contains the length of the sequence, the type of sequence, and the date that it was released to the public (or re-released following any updates, corrections, or modifications). The ORGANISM field in the SOURCE section contains a complete phylogenetic classification. The REFERENCE section contains one or more citations, the name and address of the submitter, and the date that the sequence was originally submitted to the database. The COMMENT field may be used to include annotations that cannot be accommodated in the FEATURES table. REFERENCES and COMMENTS that apply to the complete record are referred to as “descriptors.” Whenever possible, annotations should be expressed as “legal” FEATURES, using the proper syntax, because this is the only manner in which they are computer-readable. A more complex example, illustrating detailed annotation of a human breakpoint cluster region, is the GenBank record with the accession number U07000, which can be obtained from *Entrez* (see Introduction to *Entrez*, above) or the e-mail server. A few other examples are: accession numbers AF010398-AF010400, a segmented set with three segments (see Appendix to this unit, Example 4); AF018073, a bacterial sequence with multiple genes; and AF015226, a mitochondrial sequence with multiple genes (see Appendix to this unit, Example 3).

Further details on submission to the sequence databases can be found on the database Web sites and in Kans and Ouellette (1998). The following are seven important questions that should be asked when preparing a sequence for submission.

1. *What organism is the source of this DNA?*

Taxonomic classification of the organism from which the sequence was isolated is very important when using GenBank to ensure that the correct genetic code is used when translating the DNA to generate the protein product from a coding sequence (CDS) feature (see item 4). If the wrong taxa are used, or if a mitochondrial origin is not indicated, then an erroneous translation will be produced and may go undetected. Correct taxonomic classifications are also important for retrieval purposes via *Entrez*.

Over 60,000 species are represented in GenBank. The ten most common at this time are: *Homo sapiens* (human), *Mus musculus* (mouse), *Drosophila melanogaster* (fruit fly), *Caenorhabditis elegans* (nematode worm), *Arabidopsis thaliana* (thale cress), *Oryza sativa* (rice), *Rattus norvegicus* (lab rat), *Danio rerio* (zebrafish), *Saccharomyces cerevisiae* (baker's yeast), and *Lycopersicon esculentum* (tomato).

A sequence submission in most cases need only include the genus and species, which the database staff use to obtain the complete classification from NCBI's Taxonomy Database. If precise information is not available, the researcher should send in the best conjecture as to the genus, species, and/or common name, and the database staff will attempt to determine the proper classification.

Because there is great biodiversity among microbial organisms used for research, it is often important to include the strain name or number with the DNA sequence for microbial submissions. These will appear as a qualifier, for example, */strain="S288c"*, and are quite useful when doing sequence comparisons or trying to explain certain sequence conflicts. Strain information is also important in the context of genetic mapping data from the mouse (e.g., */strain="BALB/c"*) and other organisms.

2. *Is this a genomic or mRNA (cDNA) sequence?*

The sequenced molecule is most often DNA, but the annotation should indicate the biological identity of the starting material. For example, cDNA (i.e., an mRNA that was made into a DNA molecule using reverse transcriptase so that it could be cloned) should be annotated in the database as an mRNA. HIV is an interesting case. HIV sequences, as well as those from other retroviruses, may be classified in different ways: as cDNA corresponding to genomic RNA (therefore as "RNA"); as nonintegrated, nonproviral DNA (therefore as single-stranded DNA or "ssDNA"); or as integrated, proviral DNA, (therefore as double-stranded DNA or "dsDNA").

3. *If this is a genomic sequence, are there introns for which the sequence has not been determined?*

In many cases only parts of a gene have been sequenced—e.g., the exons with only a few nucleotides of flanking intron. In this case it is necessary to create what is known as a "segmented entry." What this means is that the researcher will create individual submissions for each discrete sequence and indicate the relationships between them—i.e., their linear order. Additionally, if the approximate distances (in base pairs or kilobase pairs) between the segments of a set are known, this information should be included. This will permit the complete gene to be reassembled automatically by software programs. It should be noted that this is an example of a type of entry for which Sequin is the best tool.

4. Is there a coding sequence (CDS) or other RNA product (rRNA or tRNA) on this DNA? If so, what are their positions in the DNA sequence?

If this is an mRNA, or if it is a DNA molecule encoding a protein product, it is important to indicate that a coding region is present and to supply its start and stop codons. RNA products also have boundaries or coordinates that should be indicated. An important check that is done with all submissions is to verify that the coordinates submitted correspond with the indicated feature(s). If there are any discrepancies, these are resolved by the database annotation staff in consultation with the submitter.

As described above under question 1, a common problem occurs when the taxonomic information is incorrect or incomplete, leading to the use of an improper genetic code. Fifteen different genetic codes are presently in use, as outlined in the GenBank Feature Table Document (also see information on BLAST parameters in the BLAST unit, see *UNIT 19.3*.) Any codon usage that is not represented in this list must use the translation exception qualifier (*/transl_except*) to indicate the discrepancy.

Sequin has a built-in tool for analyzing open reading frames. The user can set the minimum size of the open reading frames and the genetic code and the “ORF Finder” will graphically display all open reading frames in a sequence. There is also a separate Web version of the ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>).

5. If there is a CDS, what is gene name and the name of the encoded protein? If the gene product is an enzyme, what is the EC number?

Putting a protein or gene product name on a CDS is the best way of labeling the protein encoded by the DNA. It is possible to add more than one qualifier, but a */product*, */gene*, and */EC_number* (if the product is an enzyme) are quite sufficient. This is the label that will be carried over to the protein databases—e.g., GenPept or Swiss-Prot. Enzyme Commission (EC) numbers can be obtained or verified using the World Wide Web (*UNIT 19.1*) at <http://www.expasy.ch/enzyme/enzyme-search-ec.html>.

6. Is there information about your sequence for which there does not seem to be an appropriate place on the forms?

There are sometimes problems finding the appropriate place to put a given feature in the BankIt or Sequin forms. In such cases, just attach a note to the submission and the database staff will find the correct place or contact you for more information.

7. Is this sequence ready for submission? Has it been tested for the presence of vector contamination, mitochondrial DNA, ribosomal and tRNAs, and repetitive (e.g., Alu) elements?

Potential problems can be detected by using the BLAST programs to compare a new sequence against special data sets that represent common “contaminants.” The network-aware version of Sequin will allow you to run BLASTN against vector and mitochondrial databases. The VecScreen system has recently been developed for quickly finding vector contamination in sequences. VecScreen is automatically run on BankIt sequences and it can be run from within Sequin. It can also be used independently from its own web page (<http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html>). Other validation routines are built into both the Sequin and BankIt submission tools. These routines include a comparison of the conceptual translation of a coding region with the submitted amino acid sequence, a check that actual data length matches the given length, a check for correct use of qualifiers on features, and a check for illegal characters. Although the database staff applies a number of

validation tests to each new sequence, it is more expeditious for authors themselves to catch any problems prior to submission.

SUBMITTING AN UPDATE OR CORRECTION TO AN EXISTING GENBANK ENTRY

Authors are encouraged to submit updates or corrections to their sequence records. In fact anyone who notices a problem, error or omission in a database record should bring this to the attention of the database staff. It is not uncommon for the authors to forget to notify the databases that a previously confidential sequence may now be released. If an accession number is seen in a published article and the sequence cannot be retrieved, this is almost certainly the case. The databases will release the data if the complete journal citation is supplied, including the full title of the paper, or if a copy of the title page plus the page showing the accession number is faxed to the databases (see Internet Resources for contact information).

Update information consisting of a simple revision, such as a citation change or release of information, can be submitted by sending an e-mail message in paragraph form explaining the change. The message must include the accession number of the sequence to be updated along with all update, correction, or publication information.

Updates can also be submitted using BankIt or Sequin. See the BankIt or Sequin Web pages for details (see Submitting a Sequence to the Nucleotide Database, Submission Methods).

SUBMITTING EST, STS, OR GSS DATA

Expressed-sequence tags (ESTs), sequence-tagged sites (STSs), and genome survey sequences (GSSs) are usually submitted to GenBank and the specialized databases dbEST, dbSTS, or dbGSS as batches of dozens to thousands of entries, with a great deal of redundancy in the citation, submitter, and library information. For these data, there are special procedures designed to improve the efficiency of the submission process. A special “tagged flat file” input format is used. Documents describing the format can be obtained by sending a request to *info@ncbi.nlm.nih.gov*. They are also available from the dbEST, dbSTS, or dbGSS Web pages at: <http://www.ncbi.nlm.nih.gov/dbEST/>, <http://www.ncbi.nlm.nih.gov/dbSTS/>, and <http://www.ncbi.nlm.nih.gov/dbGSS/>, respectively. Once the data are ready to submit, they should be e-mailed to *batch-sub@ncbi.nlm.nih.gov*. For large data sets, an account can be set up into which the data can be deposited by FTP. For more information write to *info@ncbi.nlm.nih.gov*.

SUBMITTING HIGH-THROUGHPUT GENOME SEQUENCES (HTGS)

Some groups may find that their needs are not conveniently met by the standard means for data submission—e.g., a genome center that has its own internal information system or a group assembling a very large contig. The staff of NCBI, EBI, or DDBJ are happy to discuss these issues and make special arrangements with such groups so that their data can be conveniently incorporated into the databases. This includes setting up automatic exchange of data, creation of special FTP accounts, and the generation of tools to ensure data exchange in the most useful format. At NCBI, for example, FTP accounts have been set up for all submitting genome sequencing centers and a variety of tools have been created for accelerating the submission of high-throughput genome sequences (HTGS).

For large sequencing operations that wish to develop an integrated direct submission system, or for those researchers who have a data submission problem that does not seem

to be addressed by the standard tools, e-mail can be sent to *info@ncbi.nlm.nih.gov*. The situation should be explained in as much detail as possible and a phone number and postal address should be provided so that the GenBank staff can contact the investigators to discuss their situation.

CONCLUSION

The Handbook of Nucleic Acid Sequences (Barrell and Clark, 1974) contained all known DNA and RNA sequences available at that time. The book was only about half an inch thick and contained 51 tRNA sequences, 11 5S rRNA sequences, some assorted viral and mRNAs, and 9 DNA sequences—the longest of which was an 89-nucleotide sequence from bacteriophage f1. All together, this “database” amounted to ~10,000 nucleotides. If the current data in GenBank (5.8 billion nucleotides in 5.6 million entries) were published in a comparable printed form the stack of books would be about the height of Mount McKinley (20,320 feet).

The early history of electronic sequence databases is covered in part by Smith (1990). Developments since 1990 can be surveyed by consulting the annual database issues of *Nucleic Acids Research* (e.g., Benson et al., 2000) and the authors recommend scanning this special issue occasionally to learn about new databases and services. There has been a great proliferation of electronic data resources and a revolution in their delivery to the scientific community via the Internet, as described in *UNIT 19.1*.

This unit has discussed retrieval and submission of sequence data and associated literature as well as access to precomputed homology results via the *Entrez* system. But the real power of sequences is the possibility of making new and sometimes profound discoveries by computing on the data. The information content of sequences is a fundamental link between genotype and phenotype, and it is possible to infer much about genes and proteins—their functions, structures, evolution, and regulation—by the simple discovery of a sequence homology. For detailed discussion of how to effectively search for homologies in sequence databases, consult the BLAST unit (*UNIT 19.3*).

LITERATURE CITED

- Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R., and Venter, J.C. 1991. Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science* 252:1651-1656.
- Barrell, B.G. and Clark, B.F.C. 1974. *Handbook of Nucleic Acid Sequences*. Joynson-Bruvvers, Oxford.
- Baxevanis, A.D., Boguski, M.S., and Ouellette, B.F.F. 1997. Computational analysis of DNA and protein sequences. In *Genome Analysis: A Laboratory Manual* (B. Birren, E.D. Green, S. Kapholz, R.M. Myers, and J. Roskams, eds.) pp. 533-586. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Rapp, B.A., Wheeler, D.L. 2000. GenBank. *Nucl. Acids. Res.* 28:15-18.
- Boguski, M. and McEntyre, J. 1994. I think therefore I publish. *Trends Biochem. Sci.* 19:71.
- Boguski, M.S., Lowe, T.M., and Tolstoshev, C.M. 1993. dbEST: Database for “Expressed Sequence Tags”. *Nature Genet.* 4:332-333.
- Church, D.M., Stotler, C.J., Rutter, J.L., Murrell, J.R., Trofatter, J.A., and Buckler, A.J. 1993. Isolation of genes from complex sources of mammalian genomic DNA using exon amplification. *Nature Genet.* 6:98-105.
- Cockerill, M. 1994. A versatile tool for retrieving molecular sequences. *Trends Biochem. Sci.* 19:94-96.
- Harper, R. 1994. Access to DNA and protein databases on the Internet. *Current Opin. Biotechnol.* 5:4-18.
- Kahn, A.S., Wilcox, A.S., Polymeropoulos, M.H., Hopkins, J.A., Stevens, T.J., Robinson, M., Orpana, A.K., and Sikela, J.M. 1992. Single pass sequencing and physical and genetic mapping of human brain cDNAs. *Nature Genet.* 2:180-185.

- Kans, J.A. and Ouellette, B.F.F. 1998. Submitting DNA sequences to the databases. *In* Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins (A.D. Baxevanis and B.F.F. Ouellette, eds.) pp. 319-353. John Wiley & Sons, New York.
- Okubo, K., Hori, N., Matoba, R., Niiyama, T., Fukushima, A., Kojima, Y., and Matsubara, K. 1992. Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nature Genet.* 2:173-179.
- Schuler, G.D., Epstein, J.A., Ohkawa, H., and Kans, J.A. 1996. *Entrez*: Molecular biology database and retrieval system. *Methods Enzymol.* 266:141-162.
- Smith, M.W., Holmsen, A.L., Wei, Y.H., Peterson, M., and Evans, G.A. 1994. Genomic sequence sampling: A strategy to high resolution sequence-based physical mapping of complex genomes. *Nature Genet.* 7:40-47.
- Smith, T.F. 1990. The history of the genetic sequence databases. *Genomics* 6:702-707.
- Waterston, R., Martin, C., Craxton, M., Huynh, C., Coulson, A., Hillier, L., Durbin, R., Green, P., Shownkeen, R., Halloran, N., Metzstein, M., Hawkins, T., Wilson, R., Berks, M., Du, Z., Thierry-Mieg, J., and Sulston, J. 1992. A survey of expressed genes in *Caenorhabditis elegans*. *Nature Genet.* 1:114-123.

INTERNET RESOURCES

DNA Data Bank of Japan (DDBJ; Center for Information Biology, National Institute of Genetics), 1111 Yata, Mishima, Shiznoka 411, Japan; Fax 81-559-81-6849.

e-mail submissions: ddbjsub@dbj.nig.ac.jp

updates: ddbjudt@dbj.nig.ac.jp

information: dbj@dbj.nig.ac.jp

home page: <http://www.dbj.nig.ac.jp/>

WWW submissions: <http://sakura.dbj.nig.ac.jp/>

European Molecular Biology Laboratory (EMBL), EMBL Outstation, European Bioinformatics Institutes (EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, United Kingdom.

e-mail submissions: datasubs@ebi.ac.uk

updates: update@ebi.ac.uk

information: datalib@ebi.ac.uk

home page: <http://www.ebi.ac.uk>

WWW submissions: <http://www.ebi.ac.uk/Submissions/index.html>

WebIn: <http://www.ebi.ac.uk/embl/Submission/webin.html>

National Center for Biotechnology Information (NCBI), National Library of Medicine (NLM), Bldg. 38A, Room 8N-803, 8600 Rockville Pike, Bethesda, Maryland 20894; Telephone: 301-496-2475; Fax: 301-480-9241.

e-mail submissions: gb-sub@ncbi.nlm.nih.gov

EST/GSS/STS: batch-sub@ncbi.nlm.nih.gov

updates: update@ncbi.nlm.nih.gov

information: info@ncbi.nlm.nih.gov

home page: <http://www.ncbi.nlm.nih.gov/>

WWW submissions: <http://www.ncbi.nlm.nih.gov/Genbank/index.html>

BankIt: <http://www.ncbi.nlm.nih.gov/BankIt/>

Contributed by Jane M. Weisemann and Mark S. Boguski
National Center for Biotechnology Information
Bethesda, Maryland

B.F. Francis Ouellette
Centre for Molecular Medicine and Therapeutics
University of British Columbia
Vancouver, Canada

APPENDIX: SAMPLE GENBANK RECORDS

Note that these examples present data in a format called the “GenBank report.” This file format is available with Sequin or *Entrez*, but other formats for viewing the data are also available, including the “EMBL” format, FASTA format (see *UNIT 19.3*), and a graphical view.

Example 1: A Human mRNA Sequence Record with the Feature Labels and Other Landmarks Underlined

This is a typical record for an mRNA sequence with a coding region (*CDS*) and *gene* features. The *CDS* feature includes */gene*, *standard_name*, and */function* qualifiers, which are a few of many qualifiers that could be used. There is also the */translation* qualifier with the amino acid sequence for the coding region. The position of the coding region on the nucleotide sequence is given by the interval 731..3616. The record also has a *source* feature. This feature is required for all GenBank entries. A *source* feature must at a minimum include an */organism* qualifier. There are two citations associated with the record. *REFERENCE 1* is the publication in which the sequence is described and the accession number is published. *REFERENCE 2* shows the name and address of the submitter of the record and the date of submission. The *ORGANISM* field contains the name of the organism (the same as that given in the */organism* qualifier) and the phylogenetic classification. The *DEFINITION* gives a brief summary of the identity of the sequence. The typical basic *DEFINITION* consists of the organism name, gene product, gene symbol in parentheses, molecule type (e.g., mRNA), and a statement about the completeness of the coding region. The *LOCUS* field gives the length of the sequence in the record, the molecule type (mRNA, DNA, RNA, etc.), and the date the current version of the record was released to the public. Note that GenBank report qualifiers are underlined in Example 1 for ease of identification, but not in actual GenBank reports.

```
LOCUS      HSU11690      4266 bp      mRNA      PRI      10-DEC-1994
DEFINITION Human facio-genital dysplasia (FGD1) mRNA, complete cds.
ACCESSION   U11690
VERSION     U11690.1 GI:595424
KEYWORDS    .
SOURCE      human.
   ORGANISM Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 4266)
   AUTHORS   Pasteris,G.N., Cadle,A., Logie,L.J., Porteous,M.E.M.,
             Schwartz,C.E., Stevenson,R.E., Glover,T.W., Wilroy,R.S. and
             Gorski,J.L.
   TITLE     Isolation and characterization of the facio-genital dysplasia
             (Aarskog-Scott syndrome) gene: a putative rho/rac guanine
             nucleotide exchange factor
   JOURNAL   Cell 79, 669-678 (1994)
   MEDLINE   95042764
REFERENCE  2 (bases 1 to 4266)
   AUTHORS   Pasteris,N.G.
   TITLE     Direct Submission
   JOURNAL   Submitted (29-JUN-1994) Noe G. Pasteris, Pediatric Genetics,
             University of Michigan, 1150 W. Medical Center Dr., Ann Arbor, MI
             48109, USA
FEATURES
   source      Location/Qualifiers
               1..4266
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="X"
               /map="Xp11.21"
               /clone="pFCF1.1,pFCF3.85"
```



```

/tissue_type="craniofacial"
/clone_lib="lambda ZAP II 16 week human fetal craniofacial
cDNA library"
/dev_stage="fetal"
gene      731..3616
          /gene="FGD1"
CDS       731..3616
          /gene="FGD1"
          /standard_name="faciogenital dysplasia"
          /function="putative guanine nucleotide exchange factor"
          /codon_start=1
          /protein_id="AAA57004.1"
          /db_xref="GI:595425"
          /translation="MHGHRAPGGRRAFGARTPGHEPAGAAPPACADSDPGASEPGLLA
RRSGSALGGPLDPQFVGPSTSLGAAPGHRVLP CGPSPQHHRALRFSYHLEGSQPRP
          -----sequence omitted for brevity-----
          GAPQDVKAQRSLLPLIGFEVGPPEAGERPDRRHVFKITQSHLSWYFSPETEELQRRWMA
          VLGRAGRGDTFPCGPPTLSEDEMEEAPVAALGATAEPPESPQTRDKT"
BASE COUNT  858 a  1442 c  1203 g   763 t
ORIGIN
      1 cagggctcct tcccgccgcg cgccgccgag acccaggctg aagctgggga ggacgggtga
     61 gcccgcccca ggccgttctc ctgcccccgc ggctggagca tcgtctggga ctactggctg
          -----sequence omitted for brevity-----
    4201 tttttgtgt tttttgtgt gtttttttt ttttaagaaa aatacagttt atttcagct
    4261 ttacag
//

```

Example 2: A Bacterial Genomic Sequence Record

This bacterial genomic sequence contains several genes. The record has three *CDS* features for two complete protein coding regions (*mutS* and *mutL*) and one partial coding region (*cotE*). There are additional features in the record: ribosome binding sites (*RBS*), promoter regions (*-10_{signal}*) and *stem_{loop}* regions. The *DEFINITION* includes the name of all three genes and indication of the completeness of the coding regions. *REFERENCE 2* is the main citation for this entry. *REFERENCE 1* cites a paper that pertains only to nucleotides 1 to 81 of the sequence.

```

LOCUS      BSU27343      4740 bp      DNA           BCT           21-NOV-1996
DEFINITION Bacillus subtilis spore coat protein (cotE) gene, partial cds, and
            mismatch repair recognition proteins (mutS) and (mutL) genes,
            complete cds.
ACCESSIO   U27343
VERSION    U27343.1 GI:1002518
KEYWORDS   .
SOURCE     Bacillus subtilis.
            ORGANISM   Bacillus subtilis
            Bacteria; Firmicutes; Bacillus/Clostridium group;
            Bacillus/Staphylococcus group; Bacillus.
REFERENCE  1 (bases 1 to 81)
            AUTHORS    Zheng,L.B., Donovan,W.P., Fitz-James,P.C. and Losick,R.
            TITLE      Gene encoding a morphogenic protein required in the assembly of the
            outer coat of the Bacillus subtilis endospore
            JOURNAL     Genes Dev. 2 (8), 1047-1054 (1988)
            MEDLINE     89006215
REFERENCE  2 (bases 1 to 4740)
            AUTHORS    Ginetti,F., Perego,M., Albertini,A.M. and Galizzi,A.
            TITLE      Bacillus subtilis mutS mutL operon: identification, nucleotide
            sequence and mutagenesis
            JOURNAL     Microbiology 142 (Pt 8), 2021-2029 (1996)
            MEDLINE     96349107
REFERENCE  3 (bases 1 to 4740)
            AUTHORS    Albertini,A.M.
            TITLE      Direct Submission

```

JOURNAL Submitted (22-MAY-1995) Alessandra M. Albertini, Dipartimento di Genetica e Microbiologia, University of Pavia, 207 via Abbiategrasso, Pavia, 27100, Italy

```
FEATURES             Location/Qualifiers
    source             1..4740
                        /organism="Bacillus subtilis"
                        /strain="168"
                        /db_xref="taxon:1423"
                        /map="170 degrees"
                        /clone="pFG2752, pFG2764, pFG2767"
    gene               1..57
                        /gene="cotE"
    CDS                <1..57
                        /gene="cotE"
                        /note="CotE"
                        /codon_start=1
                        /transl_table=11
                        /evidence=experimental
                        /product="spore coat protein"
                        /protein_id="AAB19234.1"
                        /db_xref="GI:1002519"
                        /translation="DEELEDINPEFLVGDPEE"
    -10_signal        38..43
                        /gene="cotE"
                        /note="consensus to canonical -10 sigma A promoter"
    stem_loop          55..102
                        /function="transcription terminator"
    RBS                174..181
                        /gene="mutS"
    gene               174..2766
                        /gene="mutS"
    CDS                208..2766
                        /gene="mutS"
                        function="mismatch-repair recognition"
                        /codon_start=1
                        /transl_table=11
                        /evidence=experimental
                        /product="MutS"
                        /protein_id="AAB19235.1"
                        /db_xref="GI:1002520"
                        /translation="MIQQYLKIKAEHQDAFLFFRLGDFYEMFFEDAKKASQELEITLT
SRDGGAAEKIPMCGVPYHSASAYIEQLIKKGYKVAICEQTEDPKAAKGVVKREVVQLI
TPGTVMDGKGIHESENNFIASVSACSNGYGLALSDLTTGENLAVLIERLEDVISEIYS
VGAREIVVSGSLDADTVQQLRERCATISIEDGETDEHVTIEHLLNNEITKTFRLRY
TYLKRTQKRSLDHLQPQVYELEEAMKIDLYSKRNLELTETIRSKNKGSLWLLEDDET
KTAMGGRLLKQWIDRPLIRVNQIEERQEMVETLMSHFFEREDLRERLKEVYDLERLAG
RVAFGNVNARDLIQLKESLKQVPGIKQLVASLAHDKAKERAKRIDPCGDVLELLEAL
YENPPLSVKEGNLIKDGYNQKLDEYRDASRNGKDWIARLEQQEREYTGIRSLKVGFNK
VFGYYIEVTKANLHLLLEEGRYERNETLTNAERYITPELKEKEALILEANNICELEYE
LFTELREKVKQYIPRLQQLAKQMSELDALQCFATISENRHYTKPEFSKDEVEVIEGRH
PVVEKVMDSQEYVPPNMMGDNRQMLLITGPNMSGKSTYMRQIALISIMAQIGCFVPA
KKAVLPFDQIFTRIGAADDLISGQSTFMVEMLEAKNAIVNATKNLSLIFDEIGRGT
TYDGMALAQAIIEYVHDHIGAKTLFSTHYHELTVLEDKLPQLKNVHVRABEYNGTVVF
LHQIKEGAADKSYGIHVAQLAELPGDLIARAQDILKELEHSGNKPEVPVQKQVKEEP
AQLSFFDEAEKPAETPKLSKKEKQVIDAFKSLNILDMTPLEAMNEMYKLQKKLH"
    gene               2782..4665
                        /gene="mutL"
    CDS                2782..4665
                        /gene="mutL"
                        /function="mismatch-repair recognition"
                        /codon_start=1
                        /transl_table=11
                        /evidence=experimental
                        /product="MutL"
                        /protein_id="AAB19236.1"
```

```

/db_xref="GI:1002521"
/translation="MAKVIQLSDELSNKIAAGEVVERPASVVKELVENAIDADSTVIE
IDIEEAGLASIRVLDNGEGMENEDCKRAFRRHATSKI KDENDLFRVRTLGFRRGALPS
IASVSHLEITTSTGEGAGTKLVLQGGNI ISESRSSSRKGT EIVVSNLFFNTPARLKYM
KTVHTELGNI TDVNRIALAHPEVSI RLRHHGKNLLQTNGNGDVRHVLAAYGTAVAK
KMLPLHVSSLD FEVKG YIALPEITRASRNYMSSV VNGRYIKNFPLVKAVHEGYHTLLP
IGRHPITFIEITMDPILVDVNVHPSKLEVRLSKETELHDLIRDGIKDVFKQQQLIPSA
QVPKKSAPAIKNEQQFITFDEKPPKVPKSTAPSYSPMKLSSVVKPVDABEKLPP
LQFDAPPIVDQEQTLEVSDVSAEQPETFEQECHEEQQPASDRVPIMYPIGQMHGTYI
LAQNGENGLYIIDQHAAQERIKYEYFREKVGVEPEVQEMIVPLTFHYSTNEALIEQH
KQELESVGVFLSEFSGNSYIVRCHPAWFPKGEAELEIEIIQQVLDKSNIDIKKLRREE
AAIMMSCKGSIKANRHLRND EIKALLDDL RSTSDPFTCPHGRPIIIHSHSTYEMEKMPK
RVM"
stem_loop 4675..4715
/function="transcription terminator"
BASE COUNT 1460 a 995 c 1177 g 1108 t
ORIGIN
1 gatgaagagc ttgaagacat caaccggag tttttagtg gagatcctga agaataaaaa
61 agggactagg ggagacagta cccccaagtc tctttttat attgatata cattctttta
-----sequence omitted for brevity-----
4621 cactcgacat atgagatgga aaagatggtc aaacgcgtga tgtagcgggg gtggtaggca
4681 ttgatccatt ctctgaatgg atcaatgctt ttttgtact ccttcgagtg tcagattata

```

Example 3: A Mitochondrial DNA Record

This is a mitochondrial sequence and is similar to the bacterial genomic sequence record (see Example 2). It has two partial protein coding regions (*COI* and *COII*) and also a region encoding a tRNA (*tRNA* feature). The *source* feature includes a */mitochondrion* qualifier to show that the nucleotide sequence is from mitochondrial DNA. This is also indicated in the *ORGANISM* field and in the *DEFINITION*.

```

LOCUS      AF015226      623 bp      DNA              INV      14-APR-1998
DEFINITION Forficula auricularia Roma cytochrome oxidase 1 (COI) gene, partial
           cds; tRNA-Leu gene, complete sequence; and cytochrome oxidase 2
           (COII) gene, partial cds, mitochondrial genes encoding
           mitochondrial products.
ACCESSION  AF015226
VERSION    AF015226.1 GI:2318138
KEYWORDS   .
SOURCE     European earwig.
           ORGANISM  Mitochondrion Forficula auricularia
           Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
           Pterygota; Neoptera; Orthopteroidea; Dermoptera; Forficulidae;
           Forficula.
REFERENCE  1 (bases 1 to 623)
           AUTHORS   Wirth,T., Le Guellec,R., Vancassel,M. and Veuille,M.
           TITLE     Molecular and reproductive characterization of sibling species in
           the european earwig (Forficula auricularia)
           JOURNAL   Evolution 52 (1), 260-265 (1998)
REFERENCE  2 (bases 1 to 623)
           AUTHORS   Wirth,T., Le Guellec,R., Vancassel,M. and Veuille,M.
           TITLE     Direct Submission
           JOURNAL   Submitted (21-JUL-1997) Integrative Biology, University of Basel,
           Section of Conservation Biology, St-Johanns-Vorstadt 10, Basel
           CH-4056, Switzerland
FEATURES   Location/Qualifiers
           source     1..623
                   /organism="Forficula auricularia"
                   /organelle="mitochondrion"
                   /strain="Roma"
                   /db_xref="taxon:13068"
           gene       <1..223
                   /gene="COI"
           CDS        <1..223

```

```

        /gene="COI"
        /codon_start=2
        /transl_table=5
        /product="cytochrome oxidase 1"
        /protein_id="AAC13377.1"
        /db_xref="GI:2318139"
        /translation="YPDAYVCWNVVSSLGSTISLVGILLLVVIVWESMTSQREVLVQ
tRNA      241..308
        /product="tRNA-Leu"
gene      324..>623
        /gene="COII"
CDS      324..>623
        /gene="COII"
        /codon_start=1
        /transl_table=5
        /product="cytochrome oxidase 2"
        /protein_id="AAC13378.1"
        /db_xref="GI:2318140"
        /translation="MATWANLGFQDSAAPLMEQLTFPHDHTLFILLMITVMVGYLLGT
LYFNRFVNRFLLEGQTIIEVVWTVIPAVTLIFIALPSLRLLYLLDEINQPSLTLKVV"
BASE COUNT      137 a      103 c      154 g      229 t
ORIGIN
      1 ttaccagat gcttacgttt gttggaacgt agtctcttct ttaggttcga caatttcctt
      61 ggttgggatt cttttattgg ttgtgattgt ttgagagagt atgaccagtc agcgggaggt
           -----sequence omitted for brevity-----
      541 tgattttcat tgcctacc tcgttgcggt tgttgtacct tttagatgaa attaatacaac
      601 catcccttac cttaaagga gtg
//

```

Example 4: A Human Genomic Sequence in Three Segments

This is a set of three records for a human gene. The sequence in an individual GenBank entry must be contiguous with no gaps. In the case of this genomic sequence there are gaps between exon 1 and exon 2 and between exon 2 and exon 3. Thus the gene sequence is entered as three entries. These are, however, linked together in what is termed a “segmented set”. Each entry includes a *SEGMENT* line showing the total number of segments in the set and the segment number of the individual entry. The three records together include the complete coding region for the gene. The *CDS* feature appears in the file on the last segment containing coding sequence in the set. It indicates which parts of each sequence make up the coding region and gives the full translation.

```

LOCUS      HSTALDR1      1463 bp      DNA      PRI 17-APR-1998
DEFINITION Homo sapiens transaldolase-related protein gene, exon 1.
ACCESSION  AF010398
VERSION    AF010398.1 GI:2318007
KEYWORDS   .
SEGMENT    1 of 3
SOURCE     human.
           ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
           Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1463)
           AUTHORS  Kusuda,J., Hirai,M., Toyoda,A., Tanuma,R., Nomura-Kitabayashi,A.
           and Hashimoto,K.
           TITLE    Cloning and chromosomal localization of a paralog and a mouse
           homolog of the human transaldolase gene
           JOURNAL  Gene 209 (1-2), 13-21 (1998)
           MEDLINE  98192510
REFERENCE  2 (bases 1 to 1463)
           AUTHORS  Kusuda,J. and Hashimoto,K.
           TITLE    Direct Submission
           JOURNAL  Submitted (26-JUN-1997) Division of Genetic Resources, National

```

Institute of Infectious Diseases, 1-23-1, Toyama-chu, Shinjuku-ku,
Tokyo 162, JAPAN

REFERENCE 3 (bases 1 to 1463)

AUTHORS Kusuda,J. and Hashimoto,K.

TITLE Direct Submission

JOURNAL Submitted (12-NOV-1997) Division of Genetic Resources, National
Institute of Infectious Diseases, 1-23-1, Toyama-chu, Shinjuku-ku,
Tokyo 162, JAPAN

REMARK Sequence update by submitter

FEATURES Location/Qualifiers

source 1..1463
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p15"
/clone="cCI11p15-49"

exon 1064..1210
/number=1

BASE COUNT 236 a 504 c 472 g 251 t

ORIGIN

1 cgccactgca ctccagcctg ggcaacaagg gcgagattcc gtctcaaaaa aaaaaaaagt
61 cttttcaggt ttttgcattt ctgacaaccg gatggcccca cctggacctg ccaaccagtt
-----sequence omitted for brevity-----
1381 gaggggagcg gggcccgggt cggccgtgag aggcgcaggg ccagcggcga ccgtgagggg
1441 aacggggccc ggtcggccg tga

//

LOCUS HSTALDR2 1745 bp DNA PRI 17-APR-1998

DEFINITION Homo sapiens transaldolase-related protein gene, exon 2.

ACCESSION AF010399

VERSION AF010399.1 GI:2318008

KEYWORDS .

SEGMENT 2 of 3

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1745)

AUTHORS Kusuda,J., Hirai,M., Toyoda,A., Tanuma,R., Nomura-Kitabayashi,A.
and Hashimoto,K.

TITLE Cloning and chromosomal localization of a paralog and a mouse
homolog of the human transaldolase gene

JOURNAL Gene 209 (1-2), 13-21 (1998)

MEDLINE 98192510

REFERENCE 2 (bases 1 to 1745)

AUTHORS Kusuda,J. and Hashimoto,K.

TITLE Direct Submission

JOURNAL Submitted (26-JUN-1997) Division of Genetic Resources, National
Institute of Infectious Diseases, 1-23-1, Toyama-chu, Shinjuku-ku,
Tokyo 162, JAPAN

REFERENCE 3 (bases 1 to 1745)

AUTHORS Kusuda,J. and Hashimoto,K.

TITLE Direct Submission

JOURNAL Submitted (12-NOV-1997) Division of Genetic Resources, National
Institute of Infectious Diseases, 1-23-1, Toyama-chu, Shinjuku-ku,
Tokyo 162, JAPAN

REMARK Sequence update by submitter

FEATURES Location/Qualifiers

source 1..1745
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p15"
/clone="cCI11p15-49"

exon 963..1086
/number=2

```

BASE COUNT      355 a   473 c   413 g   504 t
ORIGIN
    1 gcatgcaacta ccacacctgg ctaatttagt ctttctatga aacctaaaaa gaatgattgg
    61 tgagtgatag cagtggataa aattagtctt aaagttgcag cagtgcgggt cagcaactat
      -----
      sequence omitted for brevity-----
    1681 ctgcctccca gattcaagcg attctcctgc ctcagcctgc cgagcagctg ggattacagg
    1741 catgc
//
LOCUS      HSTALDR3      6676 bp   DNA                PRI      17-APR-1998
DEFINITION Homo sapiens transaldolase-related protein gene, exons 3-8 and
            complete cds.
ACCESSION  AF010400
VERSION    AF010400.1 GI:2612877
KEYWORDS   .
SEGMENT    3 of 3
SOURCE     human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 6676)
  AUTHORS  Kusuda,J., Hirai,M., Toyoda,A., Tanuma,R., Nomura-Kitabayashi,A.
            and Hashimoto,K.
  TITLE    Cloning and chromosomal localization of a paralog and a mouse
            homolog of the human transaldolase gene
  JOURNAL  Gene 209 (1-2), 13-21 (1998)
  MEDLINE  98192510
REFERENCE  2 (bases 1 to 6676)
  AUTHORS  Kusuda,J. and Hashimoto,K.
  TITLE    Direct Submission
  JOURNAL  Submitted (26-JUN-1997) Division of Genetic Resources, National
            Institute of Infectious Diseases, 1-23-1, Toyama-chu, Shinjuku-ku,
            Tokyo 162, JAPAN
REFERENCE  3 (bases 1 to 6676)
  AUTHOR   Kusuda,J. and Hashimoto,K.
  TITLE    Direct Submission
  JOURNAL  Submitted (12-NOV-1997) Division of Genetic Resources, National
            Institute of Infectious Diseases, 1-23-1, Toyama-chu, Shinjuku-ku,
            Tokyo 162, JAPAN
REMARK     Sequence update by submitter
COMMENT    On Nov 13, 1997 this sequence version replaced gi:2318009.
FEATURES   Location/Qualifiers
    source   1..6676
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="11"
             /map="11p15"
             /clone="cCI11p15-49"
    mRNA     join(AF010398.1:1064..1210,AF010399.1:963..1086,41..148,
             1182..1313,4472..4648,4876..5072,5416..5561,5941..6151)
             /product="transaldolase-related protein"
    CDS      join(AF010398.1:1114..1210,AF010399.1:963..1086,41..148,
             1182..1313,4472..4648,4876..5072,5416..5561,5941..5973)
             /codon_start=1
             /product="transaldolase-related protein"
             /protein_id="AAC52068.1"
             /db_xref="GI:2612879"
             NPSLILAAAQMPAYQELVVEEAIAYGRKLGGSQEDQIKNAIDKLFVLFGAEILKKIPGR
             VSTEVDARLSFDKAMDAMVARARRLIELYKEAGISKDRILIKLSSTWEGIQAGKELEEQH
             GIHCNMTLLFSFAQAVACAEAGVTLISPFVGRILDWHVANTDKKSYEPLDGPVKSVT
             KIYNYKFKFSYKTIIVMGASFRNTGEIKALAGCDFLTISPKLLGELLQDNAKLVPLSA
             KAAQASDLEKIHLDKESFRWLHNEDQMAVEKLSDGIRKFAADAVKLERMLTERMFNAE
             NGK"
    exon     41..148
             /number=3
    exon     1182..1313

```

```

                /number=4
    exon          4472..4648
                /number=5
    exon          4876..5072
                /number=6
    exon          5416..5561
                /number=7
    exon          5941..6151
                /number=8
BASE COUNT      1439 a   1833 c   1802 g   1602 t
ORIGIN
      1 tcagaagctt ctaacctgct tttttccctt tgaatttcag gtcacaagag gaccagatta
      61 aaaatgctat tgataaactt tttgtggtgt ttggagcaga aatactaaag aagattccgg
          -----sequence omitted for brevity-----
     6601 ccaccttggt ggggcagcac gacatgcacc aacctggc cttgaagtcc ccaccccat
     6661 ttttttggtt gtttga
//
```

Sequence Similarity Searching Using the BLAST Family of Programs

Database sequence similarity searching is carried out thousands of times each day by researchers worldwide. Scientists in traditional laboratories use search results, for example, to infer the functions of newly discovered cDNAs, predict new members of gene families, and explore evolutionary relationships between sequences. In turn, they populate sequence databases with biochemically and/or genetically characterized sequences. With the advent of whole-genome sequencing, a new breed of scientist is now using these characterized sequences to predict the location and function of coding and regulatory regions in large segments of genomic DNA. Their contribution to the sequence databases is the submission, at a rapid pace, of generally uncharacterized mRNA and genomic DNA sequences. These data can then be used by the first group of scientists to enhance the understanding of their sequences. Both types of sequencing efforts have greatly increased the size and quality of the sequence databases in recent years. Thus, sequence similarity searching has become a valuable research tool for all molecular biologists (Altschul et al., 1994; Schuler, 1998).

Over the years, a number of algorithms have been implemented that allow searching of sequence databases. The most useful of these tools should share the following characteristics: (1) Speed. Because today's databases are so large, the programs must be fast in order to process megabases of sequence in seconds. (2) Sensitivity. The programs must report all potentially interesting similarities. (3) Rigorous statistics. The programs must provide a way to evaluate the significance of the results. (4) Ease of use. Scientists with no formal training in sequence-analysis algorithms should understand how to use the programs and interpret the results. Advanced users should have the option to tailor the programs to their needs. (5) Access to up-to-date databases. The doubling time of GenBank is currently ~16 months. It is important to search the most recent version of the database.

The BLAST (Basic Local Alignment Search Tool) family of sequence similarity search programs satisfies the above criteria. In short, users input either a nucleotide or amino acid query sequence, and search a nucleotide or amino acid sequence database. The program returns a list of the sequence "hits," alignments to the query sequence, and statistical values. This unit describes how to choose an appropriate BLAST program and database, perform the search, and interpret the results.

ACCESSING BLAST PROGRAMS AND DOCUMENTATION

The National Center for Biotechnology Information (NCBI) currently supports two versions of BLAST free of charge: BLAST 2.0 (gapped BLAST) and Position-Specific Iterated BLAST (PSI-BLAST). BLAST 2.0 is the standard version of BLAST, which allows a user to search a sequence database with a nucleotide or protein sequence of interest. BLAST 2.0 places gaps into the query and target sequences so that separate areas of similarity between the two sequences can be returned as one hit. PSI-BLAST is an iterative BLAST search, which is optimized for finding distantly related sequences.

Other BLAST programs are also available from the NCBI Web page. "BLAST 2 sequences" uses the BLAST search engine to produce an alignment of two sequences entered by the user. On the Specialized BLAST pages, researchers can use the BLAST engine to search sequences that are not in GenBank. At present, these databases include unfinished microbial genomes, *P. falciparum* (the human malaria parasite), and tentative

human consensus (THC) sequences from The Institute for Genomic Research (TIGR). The content of these pages is, however, subject to change.

The easiest and most popular way to access the BLAST suite of programs is through the NCBI World Wide Web site, at <http://www.ncbi.nlm.nih.gov/BLAST/>. All versions of BLAST are accessible from this site, and can be used to query all sequence databases available at the NCBI. Documentation, which includes an overview of BLAST, BLAST frequently asked questions (FAQs), a “What’s New” page, the BLAST manual, and a list of references, is also available here.

For users who want to run BLAST against private local databases or downloaded copies of NCBI databases, the NCBI offers a stand-alone version of the BLAST program. BLAST binaries and documentation are provided for the latest versions of IRIX, Solaris, DEC OSF1, and Win32 systems. BLAST 2.0 executables may be found on the NCBI anonymous FTP server at <ftp://ncbi.nlm.nih.gov/blast/executables/>.

BLAST can also be run as a client-server program, in which the user installs client software on a local machine that communicates across the network with a server at NCBI. This setup is useful for researchers who run large numbers of searches on NCBI databases, because they can automate the process to run on their local computer. The BLAST client may be found on the NCBI anonymous FTP server at <ftp://ncbi.nlm.nih.gov/blast/network/netblast>.

The NCBI BLAST e-mail server is the best option for people without convenient access to the Web. A similarity search can be performed by sending a properly formatted e-mail message containing the nucleotide or protein query sequence to blast@ncbi.nlm.nih.gov. The query sequence is compared against a specified database and the results are returned in an e-mail message. For more information on formulating e-mail BLAST searches, please send a message consisting of the word HELP to the same address, blast@ncbi.nlm.nih.gov.

This unit concentrates on the BLAST searches which can be performed from the NCBI Web site. The other implementations are mainly for advanced users of BLAST, or those with special needs. For any of these services, questions should be directed to blast-help@ncbi.nlm.nih.gov.

INTRODUCTION TO BLAST

Basic Versus Advanced BLAST Searches

BLAST 2.0 is available in a basic or advanced version. In both versions, the user can select the type of BLAST program and the database to be searched, and choose whether to filter the query sequence to mask low-complexity regions (see below). The advanced version allows the user to change parameters as well. For most researchers, the Basic version, which uses the default parameters, is adequate. For a discussion of BLAST parameters, see Appendix A at the end of this unit.

BLAST Programs

Five types of the BLAST program have been developed to support sequence similarity searching using a variety of nucleotide and protein sequence queries and databases. These programs are listed and described in Table 19.3.1. The type of BLAST search to be carried out is dependent on the type of information that is desired.

Table 19.3.1 BLAST Search Programs

Program	Query sequence	Database sequence	Comments
BLASTP	Protein	Protein	Can be run in standard mode or in a more sensitive iterative mode (PSI-BLAST), which uses the previous search results to build a profile for subsequent rounds of similarity searching.
BLASTN	Nucleotide (both strands)	Nucleotide	Parameters optimized for speed, not sensitivity; not intended for finding distantly related coding sequences. Automatically checks complementary strand of query.
BLASTX	Nucleotide (six-frame translation)	Protein	Very useful for preliminary data containing potential frameshift errors (ESTs, HTGs, and other “single-pass” sequences).
TBLASTN	Protein	Nucleotide (six-frame translation)	Essential for searching protein queries against EST database. Often useful for finding undocumented open reading frames or frameshift errors in database sequences.
TBLASTX	Nucleotide (six-frame translation)	Nucleotide (six-frame translation)	Should be used only if BLASTN and BLASTX produce no results. Restricted for search against EST, STS, HTGS, GSS, and Alu databases.

NCBI Databases

One frequent mistake in sequence similarity searching is failure to search an up-to-date database. The NCBI produces GenBank (Benson et al., 1998; <ftp://ncbi.nlm.nih.gov/genbank/gbrel.txt>) and updates it daily. It also shares data on a daily basis with the DNA Data Bank of Japan (DDBJ; Tateno et al., 1998) and the European Molecular Biology Laboratory (EMBL; Stoesser et al., 1998). A search of the NCBI databases using the BLAST Web page, client, or e-mail server guarantees access to the most recent database. The NCBI supports a number of databases for sequence similarity searching. These databases are subject to change, and a current list and description are available on the NCBI Web site at http://www.ncbi.nlm.nih.gov/BLAST/blast_databases.html. This section describes some of the commonly used NCBI databases. Examples demonstrating the utility of various BLAST programs and NCBI databases are given in the next section.

Peptide sequence databases for BLASTP and BLASTX

nr: The *nr* (nonredundant) database is the most comprehensive. GenBank, DDBJ, and EMBL are nucleotide sequence databases. If any nucleotide sequence in any of the databases is annotated with a coding sequence (CDS), this CDS appears in *nr*. *nr* also contains protein sequences obtained from PDB (sequences associated with 3-dimensional structures in the Brookhaven Protein Data Bank), Swiss-Prot (a curated database of protein sequences; Bairoch and Apweiler, 1998), PIR (Protein Identification Resource, a comprehensive collection of protein sequences; Barker et al., 1998), and PRF (Protein Research Foundation). Although *nr* may contain multiple copies of similar sequences, identical sequences are merged into one entry. To be merged, two sequences must have identical lengths and every residue at every position must be the same. There are *nr* databases for both peptide and nucleotide sequences. The peptide database is automatically selected for BLASTP and BLASTX searches.

month. The month database receives its sequences from the same sources as nr, but contains only those sequences released within the last 30 days. All sequences in month are also present in nr.

Alu. The Alu database contains six-frame translations of representative *Alu* repeats from all *Alu* subfamilies (Claverie and Makalowski, 1994). If a query sequence containing an *Alu* repeat is used in a BLAST search of nr or month, many of the resulting high-scoring hits will also contain *Alu* sequences. It may be useful, especially with a genomic sequence query, to perform a search of the Alu database to identify the location of any *Alu* repeats that might produce high-scoring and potentially misleading hits in queries of other databases.

Nucleotide sequence databases for BLASTN, TBLASTN, and TBLASTX

nr. The nr (nonredundant) database contains all nucleotide sequences present in GenBank, EMBL, and DDBJ. It also contains nucleotide sequences obtained from PDB (sequences associated with 3-dimensional structures in the Brookhaven Protein Data Bank). nr comprises only sequences that are normally well annotated, so it does not contain expressed sequence tag (EST), sequence-tagged site (STS), genome survey sequence (GSS), or high-throughput genomic (HTG) sequences. Although nr may contain multiple copies of similar sequences, identical sequences are merged into one entry. To be merged, two sequences must have identical lengths and every nucleotide at every position must be the same.

month. The month database contains all nucleotide sequences present in GenBank, EMBL, DDBJ, and PDB that were released within the last 30 days. Unlike nr, it also contains EST, STS, GSS, and HTG sequences released within the last month.

EST. EST accesses a nonredundant copy of all ESTs present in GenBank, EMBL, and DDBJ (Boguski et al., 1993). ESTs are short sequences, a few hundred nucleotides in length, which are derived by partial, single-pass sequencing of inserts of randomly selected cDNA clones (Adams et al., 1991). Since the number of ESTs is increasing rapidly, it is an important database to search for novel cDNAs. As of August, 1998, ~70% of the sequences in GenBank were ESTs; of these, 61% were from human, 20% from mouse.

STS. STS contains a nonredundant copy of all STSs present in GenBank, EMBL, and DDBJ. An STS is a short unique genomic sequence that is used as a sequence landmark for genomic mapping efforts (Olson et al., 1989). As of August, 1998, 83% of the sequences in the STS database were from human.

HTGS. HTGS contains “unfinished” DNA sequences generated by the high-throughput sequencing centers (Ouellette and Boguski, 1997). A typical HTG record might consist of all the first-pass sequence data generated from a single cosmid, BAC, YAC, or P1 clone. The record is composed of two or more sequence fragments that have a total length of ≥ 2 kb and contain one or more gaps. The sequences are normally updated by the sequencing centers as more data become available. A single accession number is assigned to this collection of sequences. The accession number does not change as the record is updated, and only the most recent version of the record remains in GenBank. Phase 1 HTG sequences are unordered, unoriented contigs with gaps. Phase 2 HTG sequences are ordered, oriented contigs with or without gaps. All HTG records contain a prominent warning that the sequence data is unfinished and may contain errors. When a record is considered finished, it becomes a Phase 3 HTG and is moved to the nr database with the same accession number. HTGS is a valuable source of new genomic sequences not yet in nr.

GSS. GSS includes short, single-pass genomic data identified by various means (Smith et al., 1994). Many of the sequences have been mapped. As of August, 1998, 80% of the sequences in GSS were from human, 14% from *Arabidopsis thaliana*.

Alu. The Alu database contains representative *Alu* repeats from all *Alu* subfamilies (Claverie and Makalowski, 1994). If a query sequence containing an *Alu* repeat is used in a BLAST search of the above nucleotide databases, many of the resulting high-scoring hits will also contain *Alu* sequences. It may be useful, especially with a genomic sequence query, to perform a search of the Alu database to identify the location of any *Alu* repeats that might produce high-scoring and potentially misleading hits in queries of other databases.

Vector. The Vector database contains nucleotide sequences of a number of standard cloning vectors. New sequences should be screened against the Vector database to assure that they do not contain any vector contamination.

Mito. The Mito database contains representative mitochondrial sequences from many families. Nuclear-derived sequences may be screened against the Mito database to assure that they do not contain any mitochondrial contamination.

Formatting the Query Sequence

Users of the BLAST Web page can initiate a search either by entering the sequence itself, or, if the sequence is already in the sequence database, by entering the accession number or gi (see Appendix B at the end of this unit). The preferred format for entering new sequences is the so-called FASTA format; however, if the sequence is not in FASTA format or the sequence is interspersed with numbers and spaces, BLAST will still accept the query. A sequence in FASTA format begins with a single-line description, followed by lines of sequence data. The description line is distinguished from the sequence data by a > symbol (greater than) in the first column. An example sequence in FASTA format is:

```
>aaseq Human choroideremia protein
MADNLPTEFDVVIIGTGLPESILAAACSRSGQRVLIHIDRSRYGGNWFSSFSGLLSWLKEYQQNDIGE
ESTVWQDLIHETEEAITLRKKDETIQHTEAFPYASQDMEDNVVEIGALQKNPSLGVSNFTFEVLDSALP
EESQLSYFNSDEMPAKHTQKSDTEISLEVTDVVEESVEKEKCYGDKTCMHTVSDKDGDKDESKSTVEDKAD
EPIRNRITYSQIVKEGRRFNIDLVSKLLYSQGLLIDLLIKSDVSRVVEFKNVTRILAFREGKVEQVPCSR
ADVFNKELTMVEKRMLMKFLTFCLEYEQHPDEYQAFRQCSFSEYLNKTKKLT PNLQHFVLHSIAMTSESS
CTTIDGLNATKNFLQCLGRFGNTPFLLFPLYGQGEIPQGFRCMCAVFGGIYCLRHVKVQCFVVDKESGRCKA
IIDHFGQRINAKYFIVEDSYLSEETCSNVQYKQISRAVLITDQSI LKTDLDQQTSILIVPPAEPGACAVR
VTELCSSMTMCKDTYLVHLTCSSSKTAREDLSEVVKLFTPYTETINEEELTKPRLWLALYFNMRDSS
GISRSSYNGLP SNVYVCSGPDGCLGNEHAVKQAETL FQEIFPTEEF C P P P P N P E D I I F D G D D K Q P E A P G T
NNVVMAKLESSEESKNLESPEKHLQN
```

Alternatively, the user can enter the database accession number or gi. The above sequence is in fact already in the database with a gi identifier of 116365 and a Swiss-Prot accession number of P26374. Either of these sequence identifiers can be entered on the BLAST page. The type of identifier (for nucleotide or protein sequence) must match the type of query sequence used in that search (Ostell and Kans, 1998; Ouellette, 1998). Sequences in manuscripts are often referred to by their GenBank accession number; however, this accession number refers to a nucleotide (not protein) sequence, and cannot be used to initiate a BLASTP or TBLASTN search, both of which require a protein sequence query. The identifier of the protein sequence encoded by a nucleotide accession number can be obtained by searching the NCBI's Entrez nucleotide database, at <http://www.ncbi.nlm.nih.gov/Entrez>.

Filtering Sequences

Both nucleotide and protein sequences may contain regions of low complexity, i.e., regions with homopolymeric tracts, short-period repeats, or segments enriched in one or only a few residues. Such low-complexity regions commonly give spuriously high BLAST scores that reflect compositional bias rather than significant position-by-position alignment (Altschul et al., 1994). For example, two protein sequences that contain low-complexity regions rich in the same amino acids may produce high-scoring alignments in those regions even though other parts of the proteins are entirely dissimilar. Because these alignments do not reflect a common ancestry, no functional inference is justified despite their high statistical significance. Filtering the query sequence (that is, replacing the repeated sequence with strings of n for nucleotide sequence or X for protein sequence) can eliminate potentially confounding matches, such as hits to low-complexity, proline-rich regions or poly(A) tails present in the database.

By default, all searches performed through the NCBI BLAST Web page automatically filter the query sequence, as do the BLAST clients, e-mail server, and stand-alone programs. However, filtering can be turned off, even on the Basic BLAST page. Filtered sequence is represented in the final BLAST report as a string of n or X (e.g., nnnnnnnnnn or XXXXXXXXXX). BLASTN queries are filtered with DUST (R.L. Tatusov and D.J. Lipman, pers. comm.). Other BLAST queries use SEG (Wootton and Federhen, 1993, 1996).

Viewing the BLAST Results

The NCBI BLAST Web page can return results in one of three ways. The default is to display the results in the browser window from which the user initiated the search. The document will have hypertext links that make it easier to analyze the results. BLAST queries are processed in the order in which they are received, except that less computationally intensive jobs (e.g., BLASTN, BLASTP, and smaller databases such as month) are given priority. In the middle of the afternoon, the BLAST server may be busy. Thus, it is sometimes more efficient to receive the BLAST results by e-mail. E-mail results are sent either as plain text or in HTML format. The HTML-formatted results must be opened in a Web browser, and the resulting document contains hypertext links.

EXAMPLES OF BLAST SEARCHES

In this section, interpretation of BLAST results is explained using examples for BLASTP, BLASTX, TBLASTN, BLASTN, and PSI-BLAST. An example is not shown for TBLASTX, as it is a tool of last resort and not useful for the majority of users. The first example, for BLASTP, contains general information, and should be read by all BLAST users. GenBank increases in size by thousands of sequences every week; thus, results obtained from running the same searches again will differ from those shown in the examples.

BLASTP

The BLASTP program compares a protein query to a protein database. A typical search is shown in Figure 19.3.1. The Swiss-Prot database has been selected and the program has been changed to BLASTP. A query in FASTA format has been entered in the input box. The query is the human choroideremia protein, implicated in hereditary blindness (Seabra et al., 1993).

The top of the results page for this search is shown in Figure 19.3.2. It begins with some header information about the type of program (BLASTP), the version (2.0.5), and a release

NCBI **Basic BLAST** Entrez ?

Clear Input **Advanced BLAST**

Message of the day ...
Sequence submissions to GenBank:
gb-sub@ncbi.nlm.nih.gov

Click here for a [description](#) of the 2.0 version of BLAST

Choose program to use and database to search:

Program Database

Perform ungapped alignment

The query sequence is filtered for low complexity regions by default.

Enter here your input data as

```
>aaseq Human choroideremia protein
MADNLETFEDVYIIIGTGLPESILAAACSRSGQRVLHIDSRSYGGNWAASFSGLLSWLKE
ESTVWQDLIHETEEAITLRKKDET IQHT EAFPYASQDMEDNVEEIGALQKNPSLGVSNITF
EESQLSYFNSDEMPAKHTQKSDTEISLEVTDVESVEKEKYCGDKT CMT VSDKDGDKDES
EPIRNRITYSQIVKEGRRFNIDL VSKLLYSQGLLIDLLIKSDVSRVVEFKNVT RILAFREGI
ADVNSKELTMVEKRMLMKFLTFCLEYEQHPDEYQAFRCFSSEYLTKKLT PNLQHFVLH
CTE TPCNNAIYNTLOCLCPRETEPELELYVCCPRPCCSRMCAITCCITVCLPDKVCCPII
```

Please read about [FASTA](#) format description

The BLAST server may be very busy during the weekday, resulting in delays for users. The email option allows a user to receive the results quickly in a convenient form. If the HTML option is used, the results should be loaded into a web browser for viewing.

Send reply to the Email address: In HTML format

Comments and suggestions to: < blast-help@ncbi.nlm.nih.gov >
Credits to: Tom Madden, Sergei B. Shavirin and Jinghui Zhang

Figure 19.3.1 Submitting a BLASTP search using the NCBI's World Wide Web interface.

```
BLASTP 2.0.5 [ May-5-1998 ]

Reference:
Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer,
Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),
"Gapped BLAST and PSI-BLAST: a new generation of protein database search
programs", Nucleic Acids Res. 25:3389-3402.

Query= aaseq Human choroideremia protein
(656 letters)

Database: Non-redundant SwissProt sequences
74,596 sequences; 26,848,718 total letters

Searching.....done

If you have any problems or questions with the results of this search
please refer to the BLAST FAQs
```

Figure 19.3.2 Example of the top portion of a BLASTP report.

Distribution of 40 Blast Hits on the Query Sequence

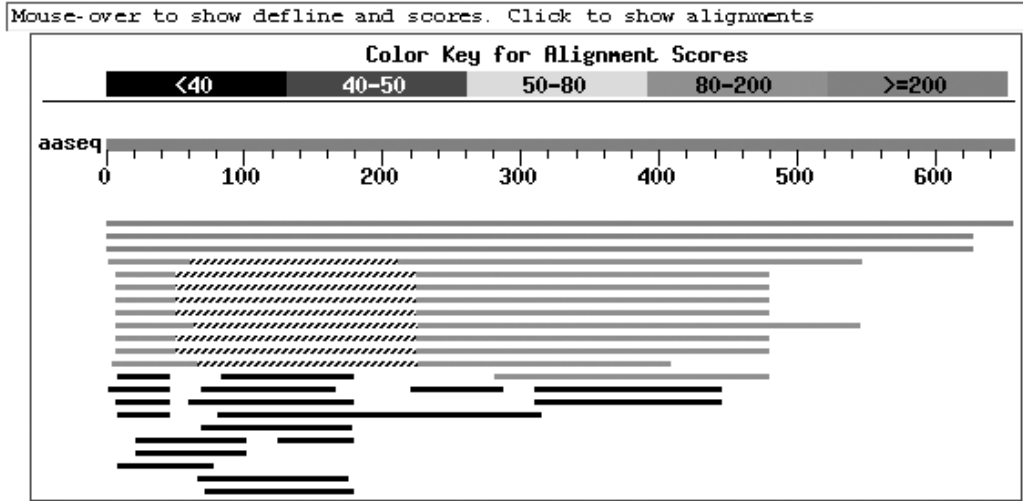


Figure 19.3.3 Example of the graphical view of a BLASTP report. The bars are color coded by the strength of the database match. The strongest matches (those with a bit score >200) are red, followed by pink (bit score 80 to 200), green (50 to 80), blue (40 to 50), and black (<40).

date. The version and release date will change as the program is updated. Also listed are a reference, the query definition line, and a summary of the database used. Figure 19.3.3 presents a graphical overview of the results. The database hits are shown aligned to the query, which is the numbered bar at the top. The next three bars show high-scoring database matches that align to the query sequence throughout its length. The next nine bars show lower-scoring matches that align to two regions of the query—one from residue 3 or 4 of the query up to about residue 60, and the other from about residues 220 to 500. For these nine bars the solid segments (on the left and right) indicate aligned regions; the cross-hatched segment indicates that the two alignments (i.e., the solid portions) are from the same database sequence, but that there is no alignment where the bar is cross-hatched. If there is no cross-hatching between two or more bars in the same row (i.e., the thirteenth row), this denotes different alignments with unrelated database sequences. These bars are presented on the same line to conserve space in the graphical view. Moving the cursor over a bar in the graphic (“mousing over”) causes the identifier and definition line of the database match to be shown in the window. If the results of a BLAST search are sent by e-mail, this graphical view is not included.

The hit list produced by the BLASTP search is shown in Figure 19.3.4. Each line of the hit list is composed of four fields. The first field contains the database designation, accession number, and locus name for the matched sequence, separated by vertical bars (see Appendix B at the end of this unit for more information). The second field contains a brief textual description of the sequence, the definition line. The content of the definition line varies between and within databases, but usually includes information on the organism from which the sequence was derived, the type of sequence (e.g., mRNA or DNA), and some information about function or phenotype. The definition line is often truncated in the hit list to keep the display compact, but is fully displayed in association with the local alignments (see below). The third field contains the alignment score in bits. Higher scoring hits are found at the top of the list. In very general terms, this score is calculated from a formula that takes into account the alignment of similar or identical

Sequences producing significant alignments:				Score	E
				(bits)	Value
sp	P26374	RAE2_HUMAN	RAB PROTEINS GERANYLGERANYLTRANSFERASE COM...	1223	0.0
sp	P24386	RAE1_HUMAN	RAB PROTEINS GERANYLGERANYLTRANSFERASE COM...	881	0.0
sp	P37727	RAE1_RAT	RAB PROTEINS GERANYLGERANYLTRANSFERASE COMPO...	856	0.0
sp	P39958	GD11_YEAST	SECRETORY PATHWAY GDP DISSOCIATION INHIBITOR	127	6e-29
sp	P50397	GDIB_MOUSE	RAB GDP DISSOCIATION INHIBITOR BETA (RAB G...	124	5e-28
sp	P21856	GDIA_BOVIN	RAB GDP DISSOCIATION INHIBITOR ALPHA (RAB ...	122	1e-27
sp	P50398	GDIA_RAT	RAB GDP DISSOCIATION INHIBITOR ALPHA (RAB GD...	122	1e-27
sp	P31150	GDIA_HUMAN	RAB GDP DISSOCIATION INHIBITOR ALPHA (RAB ...	122	2e-27
sp	P50395	GDIB_HUMAN	RAB GDP DISSOCIATION INHIBITOR BETA (RAB G...	121	4e-27
sp	Q10305	YD4C_SCHPO	PUTATIVE SECRETORY PATHWAY GDP DISSOCIATIO...	121	4e-27
sp	P50399	GDIB_RAT	RAB GDP DISSOCIATION INHIBITOR BETA (RAB GDI...	120	7e-27
sp	P32864	RAEP_YEAST	RAB PROTEINS GERANYLGERANYLTRANSFERASE COM...	98	5e-20
sp	P50396	GDIA_MOUSE	RAB GDP DISSOCIATION INHIBITOR ALPHA (RAB ...	80	9e-15
sp	Q49398	GLF_MYCGE	UDP-GALACTOPYRANOSE MUTASE	35	0.35
sp	P24588	AK79_HUMAN	A-KINASE ANCHOR PROTEIN 79 (AKAP 79) (CAMP...	35	0.46
sp	P36225	MAP4_BOVIN	MICROTUBULE-ASSOCIATED PROTEIN 4 (MICROTUB...	34	0.79
sp	Q46337	SOXA_CORSP	SARCOSINE OXIDASE ALPHA SUBUNIT	34	0.79
sp	P30599	CHS2_USTMA	CHITIN SYNTHASE 2 (CHITIN-UDP ACETYL-GLUCO...	33	1.4
sp	P53911	YNN6_YEAST	HYPOTHETICAL 49.4 KD PROTEIN IN NAM9-FPR1 ...	32	2.3
sp	P75499	GLF_MYCPN	UDP-GALACTOPYRANOSE MUTASE	32	2.3
sp	P37747	GLF_ECOLI	UDP-GALACTOPYRANOSE MUTASE	32	3.1
sp	P40142	TKT_MOUSE	TRANSKETOLASE (TK) (P68)	32	4.0
sp	P10587	MYSG_CHICK	MYOSIN HEAVY CHAIN, GIZZARD SMOOTH MUSCLE	32	4.0
sp	P50137	TKT_RAT	TRANSKETOLASE (TK)	32	4.0
sp	Q02455	MLP1_YEAST	MYOSIN-LIKE PROTEIN MLP1	31	5.2
sp	P52538	DNBI_HSV6Z	MAJOR DNA-BINDING PROTEIN (MDBP)	31	6.9
sp	P52338	DNBI_HSV6U	MAJOR DNA-BINDING PROTEIN (MDBP)	31	6.9
sp	P37637	YHIV_ECOLI	HYPOTHETICAL 111.5 KD PROTEIN IN HDED-GADA...	31	9.0
sp	Q02469	FRDA_SHEPU	FUMARATE REDUCTASE FLAVOPROTEIN SUBUNIT PR...	31	9.0
sp	Q01550	TANA_XENLA	TANABIN	31	9.0
sp	P49731	MIS5_SCHPO	MIS5 PROTEIN	31	9.0

Figure 19.3.4 Example of the hit list from a BLASTP report.

residues, as well as any gaps that must be introduced in order to align the sequences. A key element in this calculation is the “substitution matrix,” which assigns a score for aligning any possible pair of residues. The BLOSUM-62 matrix is the default for BLAST, and it works well for most searches. The fourth field contains the Expect (*E*) value, which provides an estimate of statistical significance. *E* values reflect how many times one expects to see such a score occur by chance. A statistician would consider an *E* value <0.05 to be significant. However, an inspection of the alignments (see below) is required to determine biological significance. The maximum number of hits displayed in the hit list is set at a default value of 500. This number can be changed on the advanced BLAST page with the *descriptions* option.

For the first hit in the list, the database designation is sp (for Swiss-Prot), the accession number is P26374, the locus name is RAE2_HUMAN, the definition line begins RAB PROTEINS, the score is 1223, and the *E* value is 0.0.

The higher the score and the lower the *E* value, the more statistically significant is the hit. Note that the first thirteen hits all have very low *E* values (<10⁻¹⁴) and are either RAB proteins or GDP dissociation inhibitors. The next eighteen database matches have much higher *E* values (≥0.35), meaning that roughly one match with that score would be expected by chance. Furthermore, the definition lines no longer show the same consistency. Given the combination of inconsistent definition lines and high *E* values, caution should be used in assuming an evolutionary relationship between the query and the eighteen lower-scoring database hits. It would be necessary to examine the alignments in order to make a final determination.

The pairwise alignments between the query and the database hits are displayed below the hit list. Figure 19.3.5 shows one sequence alignment between the query sequence and a


```

sp|P37727|RAE1_RAT RAB PROTEINS GERANYLGERANYLTRANSFERASE COMPONENT A 1 (RAB ESCORT
PROTEIN 1) (REF-1)
Length = 650

Score = 856 bits (2188), Expect = 0.0
Identities = 433/632 (68%), Positives = 491/632 (77%), Gaps = 13/632 (2%)

Query: 1  MADNLPTEFDVVIIGTGLPESILAAACSRSGQRVLHIDSRSYGGNWAASF SFGLLSWLK 60
MADNLP++FDV++IGTGLPESI+AAACSRSGQRVLH+DSRSYGGNWAASF SFGLLSWLK
Sbjct: 1  MADNLPSEFDVVIIGTGLPESIIAAACSRSGQRVLHVDSRSYGGNWAASF SFGLLSWLK 60

Query: 61  EYQQNNDIGEEST VVWQDLIHETEEAITLRKKDET IQHTEAFFYASQDMEDNVEEIGALQ 120
EYQ+NND+ E+++ WQ+ I E EEAI L KD+TIQH E F YASQD+ +VEE GALQ
Sbjct: 61  EYQENNDVVTENSM-WQEQILENEEAIFLSSKDKT IQHVEVFCYASQDLHKDVVEEAGALQ 119

Query: 121  KNPSLGVSNFTFEVLDSA----LPEESQLSYFNSDEMPAKHTQKSDTEISLEVTDVEESV 176
KN + S E ++A LP + S E+PA+ +Q E S EV D E +
Sbjct: 120  KNHASVT SAQSAEAAEAET SCLPTAVEPLSMSCEIPAEQSQCPGFESSPEVNDAAETG 179

Query: 177  EKEKYCGDKT CMHT VXXXXXXXXXXXXX VEDKADEPIRNRITYSQIVKEGRRFNIDLVS 236
+KE V+D + P +NRITYSQI+KEGRRFNIDLVS+
Sbjct: 180  KKEN-----SDAKSSTEESNVKQVDNTEPKKNRITYSQIIKEGRRFNIDLVSQ 231

Query: 237  LLYSQGLLIDLLIKSDVSRVYEFKNVTRILAFREGKVEQVPCSRADVFNSKELTMVEKRM 296
LLYS+GLLIDLLIKS+VSRY EFKN+TRILAFREG VEQVPCSRADVFNSK+LTMVEKRM
Sbjct: 232  LLYSRGLLIDLLIKSNVSRVYAEFKNITRILAFREGTVEQVPCSRADVFNSKQLTMVEKRM 291

Query: 297  LMKFLTFCLEYEQHPDEYQAFRCQSFSEYLKTKKLT PNLQHFLVLSIAMT SESSCTT IDG 356
LMKFLTFC+EYE+HPDEY+A+ +FSEYLKTKLTPNLQ+FVLHSIAMT SE++ T+DG
Sbjct: 292  LMKFLTFCVEYEEHPDEYRAYEGTTFSEYLKTKLTPNLQYFVLHSIAMT SETT SCT VDG 351

Query: 357  LNATKNFLQCLGRFGNTPFLFPLYGQGEIPOGFCRMCVFGGIYCLRHKVQCFVVDKESG 416
L ATK FLQCLGR+GNTPFLFPLYGQGE+PQ FCRMCAVFGGIYCLRH VQC VVDKES
Sbjct: 352  LKATKKFLQCLGRYGNTPFLFPLYGQGELPQCFCRMCAVFGGIYCLRHVSVQCLVVDKESR 411

Query: 417  RCKAIIDHFGQRINAKYFIVEDSYLSEETCSNVQYKQISR AVLITDQSILKTDLDQQT SI 476
+CKA+ID FGQRI +K+FI+EDSYLSE TCS VQY+QISR AVLITD S+LKT DQQ SI
Sbjct: 412  KCKAVIDQFGQRIISKHFIIEDSYLSENTCSR VQYRQISR AVLITDGSVLKTDADQQVSI 471

Query: 477  LIVPPAEPGACAVRVT ELCSSMT CMKDTYL VHLTCSSSKTAREDES VVKKLFTPYTET 536
L VP EPG+ VRV ELCSSMT CMK TYL VHLTC SSKTAREDELE VV+KLFTPYTE
Sbjct: 472  LAVFAEPPGSGFVGVRIELCSSTMT CMKGTYL VHLTCMSSKTAREDLERVVQKLFPTPYTEI 531

Query: 537  EINEEELTKPRLLWALYFNMRDSSGISRSSYNGLPSNVYVCSGPD CGLGNEHAVKQAETL 596
E E++ KPRLLWALYFNMRDSS ISR YN LPSNVYVCSGPD GLGN++AVKQAETL
Sbjct: 532  EAENEQVEKPRLLWALYFNMRDSSDISRDCYNDLPSNVYVCSGPD SGLGNDNAVKQAETL 591

Query: 597  FQXXXXXXXXXXXXXXXXXDGDDKQPEAP 628
FQ DGD Q E P
Sbjct: 592  FQQICPNEDFCPAPPNPEDIVLDGDSQQQEVF 623

```

Figure 19.3.5 Example of a BLASTP alignment.

database hit; other alignments are omitted for brevity. The alignment is preceded by the sequence identifier, the full definition line, and the length, in amino acids, of the database sequence. Next comes the score of the match in bits (the raw score is in parentheses), the Expect value (*E* value) of the hit, the number of identical residues in the alignment (identities), the number of conservative substitutions (positives) according to the scoring system (e.g., BLOSUM-62), and, if applicable, the number of gaps in the alignment. Finally the actual alignment is shown, with the query on top and the database match labeled as subject (Sbjct). Since both the query and the match are amino acid sequences, the numbers at left and right refer to the position in the amino acid sequence. The center line between the two sequences indicates the similarities between the sequences. If the residue is identical between the query and the subject, the residue itself is shown; conservative substitutions, as judged by the substitution matrix, are indicated with a + sign. One or more dashes (-) within a sequence indicates insertions or deletions. Query residues masked for low complexity are replaced by Xs (see fourth and eleventh blocks, Fig. 19.3.5). By default, a maximum of 500 pairwise alignments are displayed. This number can be changed on the advanced BLAST page with the *alignments* option.

Results generated on the NCBI BLAST Web page contain text hyperlinks that assist the user in navigating around the page. Clicking on a colored bar in the graphical view (Fig. 19.3.3) moves the user to the sequence alignment (Fig. 19.3.5). The database sequence identifiers, visible both in the one-line descriptions of the database matches (Fig. 19.3.4) and in the sequence alignment (Fig. 19.3.5), link to the Entrez record that describes the sequence. Entrez may provide more information about the sequence, including links to relevant publication abstracts in PubMed. The score (to the right of the one-line description; Fig. 19.3.4) provides a link within the page to the alignment between that hit and the query.

BLASTX

BLASTX searches involve a six-frame translation of a nucleotide sequence queried against a protein database (Gish and States, 1993). Researchers most often use a BLASTX search if they have sequenced a DNA molecule and do not know the reading frame of the coding sequence, the beginning or end of the open reading frame, or the function of the encoded protein. It is also useful for analyzing preliminary sequence data containing potential sequencing or frameshift errors, such as ESTs, HTGs, or GSSs. A BLASTX search of the nr protein database can quickly highlight any similarities between all the open reading frames in a nucleotide sequence and any characterized proteins.

In the example illustrated here, the mRNA sequence encoding human ataxia telangiectasia (ATM; accession number U33841; Lavin and Shiloh, 1997) was queried against the nr protein database. The output is similar to that from BLASTP. At the top (not shown) is information about the search, including the query, database, and type of BLAST program used. Next is the graphical overview representing the alignment of the database hits to the query sequence (Fig. 19.3.6), followed by a list of the hits, including their definitions, scores, and *E* values (Fig. 19.3.7 shows the top portion of the list). Pairwise alignments between the query and hit sequences follow (selected alignments are shown in Fig. 19.3.8).

The top-scoring seven hits, which have scores >2600, are to other orthologs of the ATM gene (Fig. 19.3.7). Most of these hits align with the query sequence over its entire length (in Fig. 19.3.6, the first five bars span the entire length of the query sequence; note also the longer bars in the sixth and seventh lines). The 3' end of the ATM mRNA (carboxy terminal end of the resulting protein translation), aligns with the catalytic domain of a number of phosphatidylinositol 3-kinases or similar proteins—such as the yeast TEL1 gene (a phosphatidylinositol kinase involved in controlling telomere length), the yeast ESR1 gene (a putative phosphatidylinositol kinase required for mitotic cell growth, DNA repair, and meiotic recombination), and the yeast TOR1 protein (a phosphatidylinositol kinase required for cell cycle progression). The alignment with these proteins starts in different places depending on the protein; one group starts around nucleotide 4700 of ATM (lines 9, 10, and 15 in Fig. 19.3.6), and other groups start around nucleotide 5800 (lines 17 to 19), 6800 (lines 7, 8, and 12 to 14), and 8000 (lines 20 to 60). Such similarities have provided insight into the role of ATM in response to DNA damage and cell cycle control (Lavin and Shiloh, 1997). The ATM gene also shares similarity with viral capsid proteins around nucleotides 2500 and 6000 (short bars on lines 6 to 14). However, these hits have low scores and high *E* values (not shown) and are not likely to be significant.

The hits to the orthologs of ATM reveal the location of the coding sequence within the query nucleotide sequence. In a BLASTX output, position in the query sequence (the top line of a pairwise alignment) is described as a nucleotide position, while the position in the database hit (the bottom line of the alignment) is described as an amino acid position (Fig. 19.3.8). For example, amino acids 1 to 3066 of the mouse ortholog of ATM, gi

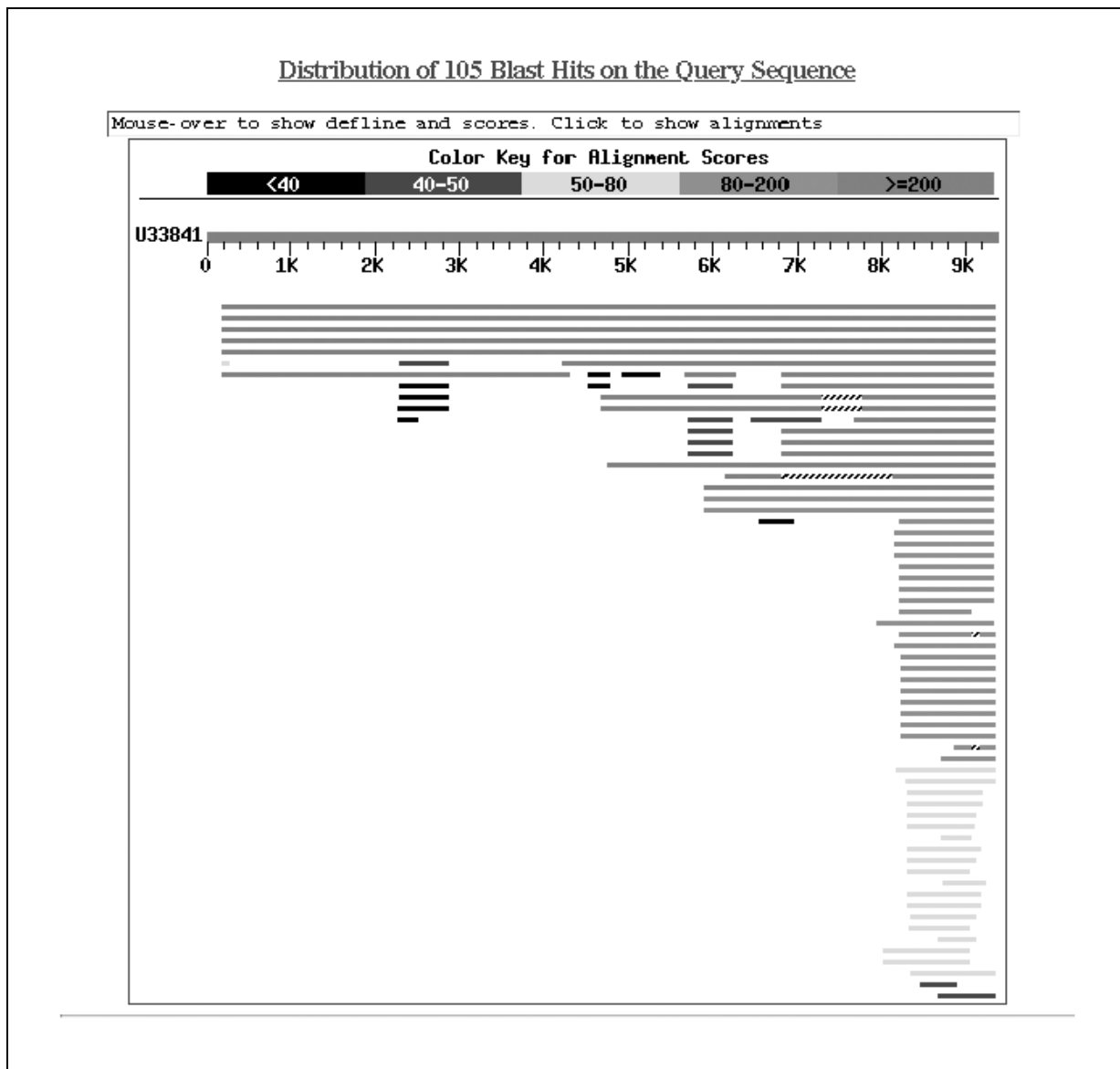


Figure 19.3.6 Example of the graphical view of a BLASTX report.

1469394, align with the translation of nucleotide positions 190 through 9357 of the query sequence. The exact location of the ATM coding sequence cannot be easily determined from the hits to the yeast sequences, as the similarity does not extend throughout the length of the protein (Fig. 19.3.8, bottom). However, the hits to the yeast sequences do provide confirmation about the phase of the reading frame, and possibly the location of specific domains conserved between humans and yeast.

TBLASTN

TBLASTN searches involve a protein sequence queried against the six-frame translation of a nucleotide sequence database. TBLASTN searches are especially useful for finding novel similarities to a protein sequence in the open reading frames from uncharacterized, sometimes lower-quality nucleotide sequences, such as those found in EST, STS, HTGS, and GSS databases. The translations of these sequences are not present in nr because the sequences are not annotated. The most common use of TBLASTN is to search for ESTs that are similar to a protein of interest. These ESTs might represent orthologs of the protein from a different species, or just additional members of a gene family.

Sequences producing significant alignments:	Score (bits)	E Value
gi 1063621 mutated in ataxia telangiectasia	6052	0.0
gi 1497931 (U55757) ataxia-telangiectasia [Homo sapiens]	6048	0.0
prf 2124355A ATM gene [Homo sapiens]	6046	0.0
gi 2304971 (U82828) ATM [Homo sapiens]	6046	0.0
gi 1469394 (U43678) homolog of the human ataxia telangiectasia ...	5085	0.0
pir A43100 ataxia telangiectasia-associated protein - human >g...	3429	0.0
gnl FID c198283 (X91196) A-T [Homo sapiens]	2640	0.0
gnl FID c304507 (Z35849) ORF YBL088c [Saccharomyces cerevisiae]	335	2e-90
sp P38110 TEL1_YEAST TELOMER LENGTH REGULATION PROTEIN TEL1 >gi...	335	2e-90
gi 1235902 (U49844) FRAP-related protein [Homo sapiens] >gi 158...	230	1e-58
gnl FID c276916 (Y09077) atr [Homo sapiens] >gi 1666240 (U76308...	230	1e-58
gnl FID c276898 (Y09076) RAD3 [Schizosaccharomyces pombe] >gi 1...	221	6e-56
sp P38111 ESR1_YEAST ESR1 PROTEIN >gi 626949 pir S46005 ESR1 p...	219	3e-55
gi 950173 (U31109) Mec1p [Saccharomyces cerevisiae]	219	3e-55
gnl FID d1002337 (D11088) Esr1 protein [Saccharomyces cerevisiae]	219	3e-55
gi 998353 (U34925) MEI-41 [Drosophila melanogaster] >gi 1583574...	212	4e-53
gnl FID c330334 (Z97992) putative phosphatidylinositol 3-kinase...	208	6e-52
gi 298028 (X71416) TOR2 [Saccharomyces cerevisiae]	200	1e-49
sp P32600 TOR2_YEAST PHOSPHATIDYLINOSITOL 3-KINASE TOR2 (PI3-KI...	199	2e-49
prf 2010264B TOR2(DRR2) gene [Saccharomyces cerevisiae]	199	2e-49
gi 1938552 (U97016) strong similarity to the P13/P14-kinases fa...	195	3e-48
gi 408956 (L19540) mutant drxl-1 protein [Saccharomyces cerevis...	192	3e-47
sp P35169 TOR1_YEAST PHOSPHATIDYLINOSITOL 3-KINASE TOR1 (PI3-KI...	192	4e-47
gi 468739 (X74857) TOR1 gene product [Saccharomyces cerevisiae]	192	4e-47
sp P42346 FRAP_RAT FKBP-RAPAMYCIN ASSOCIATED PROTEIN (FRAP) (RA...	178	4e-43
sp P42345 FRAP_HUMAN FKBP-RAPAMYCIN ASSOCIATED PROTEIN (FRAP) (...	176	1e-42
prf 2014422A FKBP-rapamycin-associated protein [Homo sapiens]	171	5e-41
bbs 158901 RAPT1=putative novel phosphatidylinositol 3-kinase (...	166	2e-39
gi 1938569 (U97189) strong similarity to thw P13/P14 family of...	150	1e-34
gnl FID c236204 (Z70756) T06E4.3 [Caenorhabditis elegans]	146	2e-33
gnl FID d1020343 (D87521) DNA-PKcs [Mus musculus] >gi 3241860 g...	141	9e-32
gnl FID d1012789 (D83786) mDNAPk3' [Mus musculus]	141	9e-32
gi 2749954 (AF001413) DNA-dependent protein kinase [Xenopus lae...	137	1e-30
gi 1765938 (U47077) DNA-dependent protein kinase catalytic subu...	134	7e-30
gi 1017757 (U35835) DNA-PK [Homo sapiens] >gi 1587037 prf 2205...	132	4e-29
gi 995941 (U34994) DNA dependent protein kinase catalytic subun...	113	2e-23
pir A57099 DNA-activated protein kinase, catalytic subunit - h...	112	4e-23
gnl FID d1025773 (AB007881) KIAA0421 [Homo sapiens]	89	5e-16
gi 1688254 (U78157) DNA-dependent protein kinase [Mus musculus]	83	3e-14
gi 2708741 (AC003952) hypothetical protein [Arabidopsis thaliana]	78	7e-13
gi 987948 (Z46973) phosphatidylinositol 3-kinase [Homo sapiens]	71	9e-11
pir S57219 phosphatidylinositol 3-kinase - human	71	9e-11
gi 1688256 (U78158) DNA-dependent protein kinase [Mus musculus]	66	4e-09
sp P42347 P3K1_SOYBN PHOSPHATIDYLINOSITOL 3-KINASE, ROOT ISOFOR...	66	4e-09
gi 1938573 (U97190) strong similarity to phosphatidylinositol 3...	66	4e-09
gnl FID c258266 (X99912) 1-phosphatidylinositol 3-kinase [Droso...	66	5e-09
sp P42348 P3K2_SOYBN PHOSPHATIDYLINOSITOL 3-KINASE, NODULE ISOFO...	65	6e-09
sp P54676 P3K4_DICDI PHOSPHATIDYLINOSITOL 3-KINASE VPS34-LIKE (...	65	8e-09
sp P54677 PI4K_DICDI PHOSPHATIDYLINOSITOL 4-KINASE (PI4-KINASE)...	63	3e-08
sp P42339 PI3K_ARATH PHOSPHATIDYLINOSITOL 3-KINASE (PI3-KINASE)...	62	4e-08

Figure 19.3.7 Example of the hit list from a BLASTX report.

The following example shows a TBLASTN search of HTGS, a database that contains unfinished and uncharacterized genomic sequences generated by genome sequencing centers. Since HTGs are genomic and often contain introns, the results can be more difficult to interpret than those from an EST search. However, HTGs can also provide insight into the chromosomal organization of genes.

In this case, the six-frame translation of HTGS is searched with the amino acid sequence of the disintegrin-like domain of mouse ADAM 1 (Wolfsberg et al., 1995). TBLASTN output is similar to that of BLASTP and BLASTX. Figure 19.3.9 shows the graphical overview of the alignment, Figure 19.3.10 shows the list of database hits, and Figure

gi|1469394 (U43678) homolog of the human ataxia telangiectasia gene [Mus musculus]
Length = 3066

Score = 5085 bits (13044), Expect = 0.0
Identities = 2526/3056 (82%), Positives = 2751/3056 (89%), Gaps = 14/3056 (0%)

Query: 190 MSLVLNDDLICCRQLEHDRATEREKKEVEKFKRLIRDPET IKHLDRHSDSKQGGKYLNWDAV 369
MSL LNDLLICCRQLEHDRATER+KEV+KFKRLI+DPET++HLDRHSDSKQGGKYLNWDAV
Sbjct: 1 MSLALNDDLICCRQLEHDRATEREKKEVDKFKRLIQDPET VQHLDHRHSDSKQGGKYLNWDAV 60

Query: 370 FRFLQKYIQKETECLEIRIAPNVASATQASRQKKMQEISSLVKYFIKCANRRAPRLKQCQL 549
FRFLQKYIQKE E LR AK NVSA+TQ+SRQKKMQEISSLV+YFIKCAN+RAPRLKQC+L
Sbjct: 61 FRFLQKYIQKEMESLRTAKSNVSATTQSSRQKKMQEISSLVRYFIKCANRAPRLKQCQL 120

Query: 550 LNYIMDTVKDSSNGAIYGADCSNILLKDILSVRKYWCEISQQQWLELFSVYFRLYLKPSQ 729
LNY+MDTVKDSSNG YGADCSNILLKDILSVRKYWCE+SQQQWLELFS+YFRLYLKPSQ
Sbjct: 121 LNYVMDTVKDSSNGLTYGADCSNILLKDILSVRKYWCEVSQQQWLELFSLYFRLYLKPSQ 180

[middle portion of alignment not shown]

Query: 8968 FRRCCCKTMEVMRNSQETLLT IVEVLLYDPLFDWT MNPLKALYLQORPEDET ELHPTLNA 9147
FRRCCCKTMEVMR+SQETLLT IVEVLLYDPLFDWT MNPLKALYLQORPEDE++LH T NA
Sbjct: 2937 FRRCCCKTMEVMRNSQETLLT IVEVLLYDPLFDWT MNPLKALYLQORPEDES DLHSTPNA 2996

Query: 9148 DDQECKRNLSDIDQSFDKVAERVLMLRQEKLGVEEGT VLSVGGQVNLIIQQAIDPKNLS 9327
DDQECK++LSD DQSF+KVAERVLMLRQEKLGVEEGT VLSVGGQVNLIIQQA+DPKNLS
Sbjct: 2997 DDQECKQLSDTDQSFNKVAERVLMLRQEKLGVEEGT VLSVGGQVNLIIQQAIDPKNLS 3056

Query: 9328 RLFFPGWKAWV 9357
RLFFPGWKAWV
Sbjct: 3057 RLFFPGWKAWV 3066

[intervening alignments not shown]

gi|468739 (X74857) TOR1 gene product [Saccharomyces cerevisiae]
Length = 2470

Score = 192 bits (482), Expect = 4e-47
Identities = 124/397 (31%), Positives = 214/397 (53%), Gaps = 19/397 (4%)

Query: 8155 LKNLEDVVVPTMEIKVDHTGEY--GNLVT IQSFKAERLAGGVNLPKIIDCVGSDGKER 8325
L+++ ++T ++++ GY + I F+ F+ P+ GSDGK+
Sbjct: 2060 LQHVSPLLATHDELA VPGTYFP GKPT IRIAKFEPLFSV ISSKQRPRKFSIKGSDGKDY 2119

Query: 8326 RQLVKGRDDL RQDAVMQVQMCNTLLQRNTE TRKRKLT ICTYKVVPLSQRSGVLEWCTG 8505
+ ++KG +D+RQD+++ Q+F + NTLL+ ++E KR L I Y +PLS +SG+L W
Sbjct: 2120 KYVLKGHEDIRQDSLVMQLFGLVNTLLKNDSECFKRHLDIQQYFAIPLSFKSGLLGWVFN 2179

Query: 8506 TVPIGEFLVNNEDGAHKRYRPNDFAFQCCKMMVEVQKKSFEKYEYVFMVQCQNF--QPV 8679
+ + + D + P + + + + + +K EVF N Q +
Sbjct: 2180 SDTFHVLIREHRDA--KKIPLNIEHWMLQMAPDYENLTL LQKIEVFTYALDNTKGQDL 2236

Query: 8680 FRYFCMEKFLDPAIWF EKRLAYTRSVATSSIVGYILGLGDRHVQNILINEQS AELVHIDL 8859
++ + K W E+R YRS+A S+ GYILGLGDRH N++++ + +++HID
Sbjct: 2237 YKILWL-KSRSETWLERRTTYTRSLAVMSMTGYILGLGDRHPSNLMLDRITGKVIHIDF 2295

Query: 8860 GVAFEQKILPT-PETVPFRLTRDIVDGMGITGVEGVFRRCCCKTMEVMRNSQETLLTIV 9036
G FE + PE VPFRLTR + M ++G+EG FR CE M V+R+++E+L+ I+
Sbjct: 2296 GDCFEAAILREKYPEKVPFRLTRMLTYAMEVSGIEGFRITCENVMRVLDRDNKESLMAIL 2355

Query: 9037 EVLLYDPLFDWT MNPLKALYLQORPEDET ELHPTLNADDQECKRNLSDIDQSFDKVAER- 9213
E DPL W + Q+ ++T + L + ++ +++++ AE+
Sbjct: 2356 EAFALDPLIHGFD----LPPQKLTQGTGIPLPLINSELLRKGAITVEEAAANMEAEQQ 2410

Query: 9214 -----VLMRLQEKLG--VEEGT VLSVGGQVNLIIQQAIDPKNLSRLFFGW 9345
VL R+ +KL G ++ L V QV+ LIQQA + L + + GW
Sbjct: 2411 NETRNARAMLVLRRITDKLTGNIDIKRFNELDVPVEQVDKLIQQAATSIERLCQHYIGW 2466

Figure 19.3.8 Example of selected BLASTX alignments.

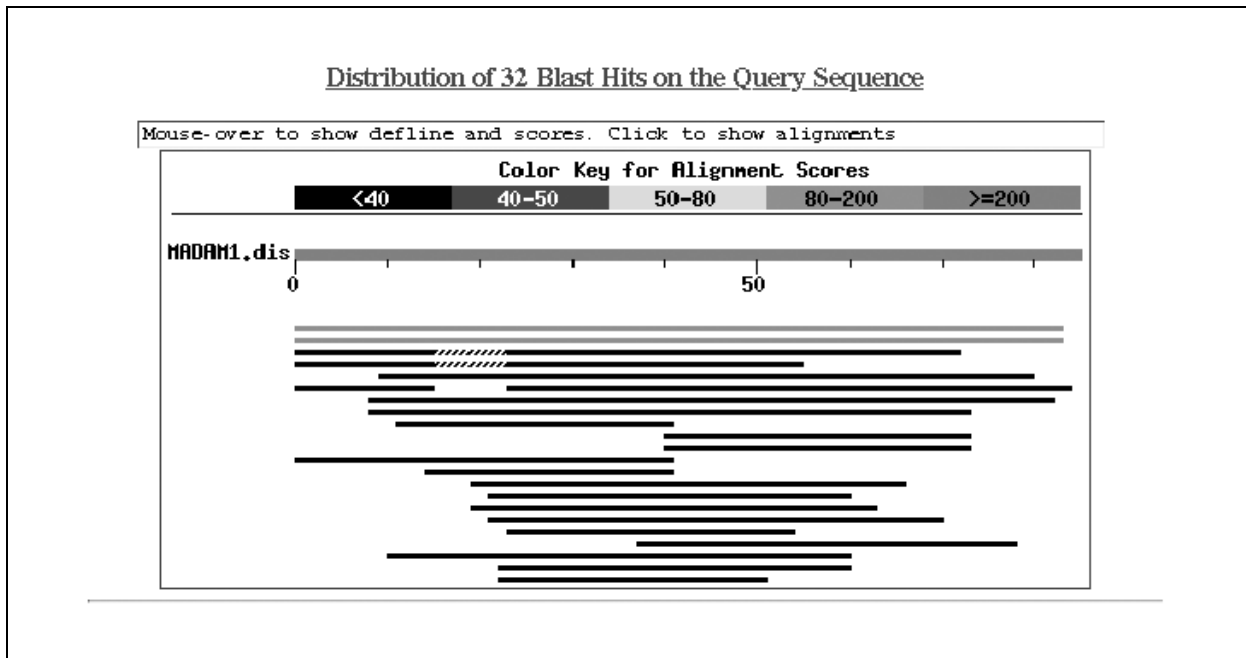


Figure 19.3.9 Example of the graphical view of a TBLASTN report.

Sequences producing significant alignments:

		Score (bits)	E Value
emb Z98857 CEY111B2	Caenorhabditis elegans DNA *** SEQUENCING I...	94	2e-19
emb Z92819 CEY37D8	Caenorhabditis elegans DNA *** SEQUENCING IN...	94	2e-19
qb AC005075 AC005075	Homo sapiens clone RG219E16; HTGS phase 1,...	37	0.027
qb AC002557 AC002557	*** SEQUENCING IN PROGRESS *** Homo sapien...	36	0.035
qb AC003656 AC003656	*** SEQUENCING IN PROGRESS *** EPML/APECED...	34	0.23
qb AC005180 AC005180	*** SEQUENCING IN PROGRESS *** Homo sapien...	33	0.40
qb AC004877 AC004877	Homo sapiens clone DJ0751H13; HTGS phase 1,...	32	0.68
qb AC003066 AC003066	*** SEQUENCING IN PROGRESS *** Mus musculu...	31	1.2
qb AC002042 AC002042	*** SEQUENCING IN PROGRESS *** Homo sapien...	31	1.5
qb AC003097 AC003097	*** SEQUENCING IN PROGRESS *** Human chrom...	31	1.5
qb AC004303 AC004303	*** SEQUENCING IN PROGRESS *** DS06475; D1...	31	1.5
qb AC003095 AC003095	*** SEQUENCING IN PROGRESS *** Human chrom...	31	1.5
qb AC000406 HSAC000406	*** SEQUENCING IN PROGRESS *** Human Chr...	30	3.4
emb AL022343 HS608E8	Human DNA sequence *** SEQUENCING IN PROGR...	30	3.4
emb AL022100 HS1059H15	Human DNA sequence *** SEQUENCING IN PRO...	30	3.4
emb Z92863 CEY64G10	Caenorhabditis elegans DNA *** SEQUENCING I...	30	3.4
emb AL031279 HS163016	Human DNA sequence *** SEQUENCING IN PROGR...	29	5.9
emb Z97634 HS367G8	Human DNA sequence *** SEQUENCING IN PROGRES...	29	5.9
qb AC005025 AC005025	Homo sapiens clone GS464G18; HTGS phase 1,...	29	5.9
qb AC004618 AC004618	*** SEQUENCING IN PROGRESS *** Homo sapien...	29	5.9
qb AC005086 AC005086	Homo sapiens clone RG305I02; HTGS phase 1,...	29	7.8
qb AC004169 AC004169	*** SEQUENCING IN PROGRESS *** Homo sapien...	29	7.8
emb AL008728 HS125N5	Human DNA sequence *** SEQUENCING IN PROGR...	29	7.8

Figure 19.3.10 Example of the hit list from a TBLASTN report.

19.3.11 shows selected pairwise alignments between the query sequence and the hits. For TBLASTN searches, the top line (Query) of the alignment contains protein sequence, and the location refers to amino acids. The bottom line (Sbjct) of the alignment contains a translated nucleotide sequence, and the location refers to the nucleotide location.

Disintegrin-like domains contain a characteristic pattern of cysteines, so it is usually easy to identify the true hits by checking for conservation in the number of cysteine residues, and the spacing between them. The first hit, a 627-kb *C. elegans* clone, aligns with the

emb|Z98857|CEY111B2 *Caenorhabditis elegans* DNA *** SEQUENCING IN PROGRESS *** from clone Y111B2; HTGS phase 1 [*Caenorhabditis elegans*]
Length = 627200

Score = 94.0 bits (230), Expect = 2e-19
Identities = 43/84 (51%), Positives = 56/84 (66%), Gaps = 4/84 (4%)

Query: 1 CGNGVVEDLEECDCGS--DCDSHPCCSP-TCTLKEGAQCS-EGLCYINCTFKKKGSLCRP 56
CGNGVV+ EECDCGS +C PCC P TCTL+ AQC+ CC+ C +K G CR
Sbjct: 322651 CGNGVVDGSEECDCGSRENCQD-PCCDPLTCTLRPHAQCAAHHKCCRCELKRAGDTCRS 322827

Query: 57 AEDVCDLPEYCDGSGTQECFANSIMQDGT 84
++ CD+ E CDG + +CP + + DGT
Sbjct: 322828 SKSPCDVAEQCDGKSGDCPPDGHLLDGT 322911

Score = 46.1 bits (107), Expect = 1e-06
Identities = 30/55 (54%), Positives = 36/55 (64%), Gaps = 21/55 (38%)

Query: 1 CGNGVVEDLEECDCG---SDCD--SHPCCSP-----TCTLKEGAQC--SEGLCCY 43
CGN + E EECDCG +DCD CC P C K GAQC S+G CC
Sbjct: 102880 CGNQIYEPGEECDGSGFSAQDCDQMGDKCCVPHEARGNGGGPCKRKPQAQCSPSQGYCCN 103059

Query: 44 --NCTF--KKKGLSCR 55
C+ K + +CR
Sbjct: 103060 PDTCSLHGKNEEKICR 103107

Score = 24.3 bits (51), Expect = 1e-06
Identities = 9/26 (34%), Positives = 12/26 (45%)

Query: 58 EDVCDLPEYCDGSGTQECFANSIMQDG 83
E C + CDG +CP + DG
Sbjct: 103155 ESECSNLQTCGRNAQCPVSPPKHDG 103232

emb|Z92819|CEY37D8 *Caenorhabditis elegans* DNA *** SEQUENCING IN PROGRESS *** from clone Y37D8; HTGS phase 1 [*Caenorhabditis elegans*]
Length = 353436

Score = 94.0 bits (230), Expect = 2e-19
Identities = 43/84 (51%), Positives = 56/84 (66%), Gaps = 4/84 (4%)

Query: 1 CGNGVVEDLEECDCGS--DCDSHPCCSP-TCTLKEGAQCS-EGLCYINCTFKKKGSLCRP 56
CGNGVV+ EECDCGS +C PCC P TCTL+ AQC+ CC+ C +K G CR
Sbjct: 304601 CGNGVVDGSEECDCGSRENCQD-PCCDPLTCTLRPHAQCAAHHKCCRCELKRAGDTCRS 304425

Query: 57 AEDVCDLPEYCDGSGTQECFANSIMQDGT 84
++ CD+ E CDG + +CP + + DGT
Sbjct: 304424 SKSPCDVAEQCDGKSGDCPPDGHLLDGT 304341

gb|AC005075|AC005075 *Homo sapiens* clone RG219E16; HTGS phase 1, 3 unoxdered pieces [*Homo sapiens*]
Length = 198776

Score = 36.7 bits (83), Expect = 0.027
Identities = 13/26 (50%), Positives = 16/26 (61%)

Query: 24 CSPTCTLKEGAQCSEGLCCYINCTFKK 49
C CTL + +QCS+GLCC C K
Sbjct: 53474 CCKKCTLTQDSQCSDLCCCKKCVNK 53551

Score = 31.7 bits (70), Expect = 0.89
Identities = 11/16 (68%), Positives = 13/16 (80%)

Query: 1 CGNGVVEDLEECDCGS 16
CGNG +E EECDCG+
Sbjct: 53178 CGNGFIETGEECDGT 53225

Score = 29.0 bits (63), Expect = 5.9
Identities = 10/27 (37%), Positives = 18/27 (66%)

Query: 47 FKKKGLSCRPAEDVCDLPEYCDGSGTQE 73
F+ G++CR A + CD+ E C G++ +
Sbjct: 57012 FQPMGTVCREAVMDCDIRETCSGNSSQ 57092

Figure 19.3.11 Example of a TBLASTN alignment.

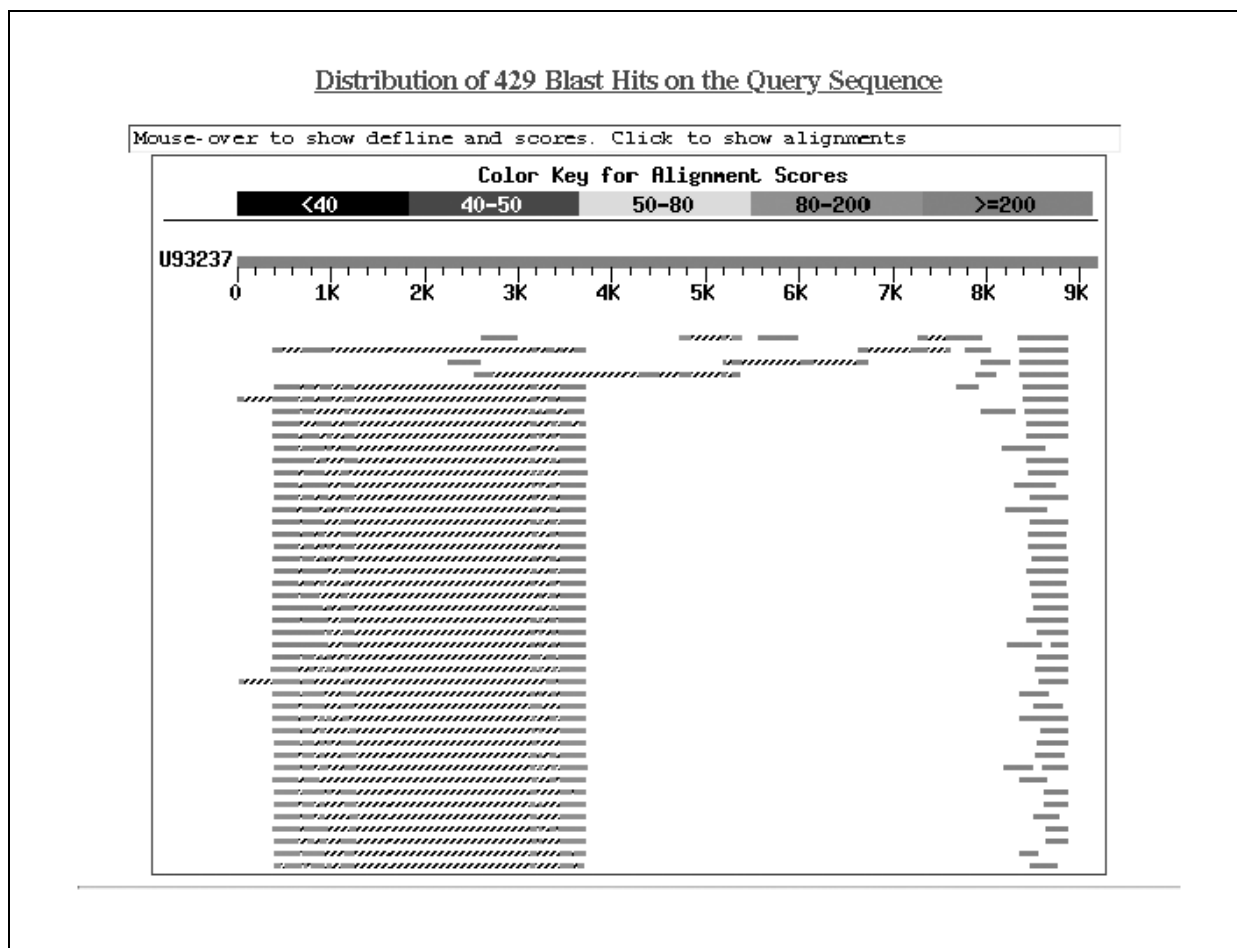


Figure 19.3.12 Example of the graphical view of a BLASTN report.

query sequence in two regions. Both alignments encompass the full length of the query sequence, and both contain a characteristic alignment of cysteine residues. The first region (the first aligned block) comprises nucleotides 322651-322911 of the *C. elegans* clone. The second region consists of the second and third alignment blocks. The first 55 amino acids of the disintegrin domain align with nucleotides 102880-103107 of the clone, and the last 25 amino acids align with nucleotides 103155-103232 of the clone. Thus, this *C. elegans* clone appears to encode two distinct disintegrin-like domains, and these two sequences are likely linked in the *C. elegans* genome. The region between nucleotides 103107-103155 of the clone may contain an intron. The second database hit is to a shorter *C. elegans* clone that, from the sequence of its encoded disintegrin domain, appears to contain some of the same sequence as the first *C. elegans* clone. The third database hit is to a human clone. As amino acids 1 to 16 of the query aligns with nucleotides 53178 to 53225 of the hit, and amino acids 24 to 49 align with nucleotides 53474 to 53551, the human sequence may contain an intron between nucleotides 53225 and 53474. There is a likely second intron between nucleotides 53551 and 57012.

BLASTN

Figure 19.3.12 presents the graphical overview of a BLASTN search of accession number U93237 against the human EST database. U93237 is the 9.2-kb genomic sequence of the human menin (MEN1) gene for multiple endocrine neoplasia type I. This example demonstrates the power of a BLAST search to confirm exons in a genomic sequence, as well as the danger of accepting results without further investigation. The matches on the

Sequences producing significant alignments:			Score	E
			(bits)	Value
gb AA776738 AA776738	ah49f08.s1	Soares testis NHT Homo sapiens ...	950	0.0
gb AI082242 AI082242	ox79g05.x1	Soares_NhHMPu_S1 Homo sapiens c...	942	0.0
gb AA743431 AA743431	ny23f01.s1	NCI_CGAP_GCB1 Homo sapiens cDNA...	940	0.0
gb AA705195 AA705195	zj96e04.s1	Soares fetal liver spleen lNFLS...	922	0.0
gb AA973790 AA973790	oo45g04.s1	NCI_CGAP_Lu5 Homo sapiens cDNA ...	862	0.0
gb AA604513 AA604513	no73c10.s1	NCI_CGAP_AA1 Homo sapiens cDNA ...	850	0.0
emb Z98467 HSZ98467		Homo sapiens mRNA; expressed sequence tag; ...	841	0.0
gb AI056630 AI056630	oz18e01.x1	Soares_fetal_liver_spleen_lNFLS...	809	0.0
gb AA636030 AA636030	nr37h07.s1	NCI_CGAP_Pr22 Homo sapiens cDNA...	797	0.0
gb AA614240 AA614240	np09f12.s1	NCI_CGAP_Pr3 Homo sapiens cDNA ...	791	0.0
gb AA147620 AA147620	zl52d02.s1	Soares pregnant uterus NbHPU Ho...	777	0.0
gb AA262936 AA262936	zr71a02.s1	Soares_NhHMPu_S1 Homo sapiens c...	773	0.0
gb AA147612 AA147612	zl52d02.r1	Soares pregnant uterus NbHPU Ho...	761	0.0
gb AI078460 AI078460	oz13e02.x1	Soares_fetal_liver_spleen_lNFLS...	755	0.0
gb AA209475 AA209475	zq84e11.s1	Stratagene hNT neuron (#937233)...	747	0.0
gb AA459317 AA459317	aa25e05.r1	NCI_CGAP_GCB1 Homo sapiens cDNA...	741	0.0
gb AA805578 AA805578	ob44g06.s1	NCI_CGAP_GCB1 Homo sapiens cDNA...	739	0.0
gb H41351 H41351	yp71b11.s1	Homo sapiens cDNA clone 192861 3'.	716	0.0
gb AA157374 AA157374	zo68h09.s1	Stratagene pancreas (#937208) H...	710	0.0
gb AA775293 AA775293	ad18f11.s1	Soares_NbHFB Homo sapiens cDNA ...	706	0.0
gb H97384 H97384	yw03h05.s1	Homo sapiens cDNA clone 251193 3'.	702	0.0
gb H91639 H91639	yv03g12.s1	Homo sapiens cDNA clone 241702 3'.	666	0.0
gb N51774 N51774	yz01g01.s1	Homo sapiens cDNA clone 281808 3'.	666	0.0
emb Z98466 HSZ98466		Homo sapiens mRNA; expressed sequence tag; ...	654	0.0
gb H38333 H38333	yp70g02.r1	Homo sapiens cDNA clone 192818 5'.	636	e-180

Figure 19.3.13 Example of the hit list from a BLASTN report.

right half of the diagram (base pairs 4000 to 9000) confirm the locations of exons 2 to 10 of the MEN1 gene (Chandrasekharappa et al., 1997; Zhang and Madden, 1997). The large number of matches on the left side of the figure (base pairs 500 to 1000 and 3000 to 3600) appear at first glance to be interesting matches, but are actually caused by the presence of *Alu* elements in the query and EST sequences. One can determine this by checking the definition lines of the matches in this region, as many of these ESTs are annotated as being similar to *Alu* elements. It is important to remember, though, that the definition line summaries (e.g., Fig. 19.3.13) contain truncated definition lines, and that it is necessary to look at the full definition line (provided in the alignment section, Fig. 19.3.14) to see the *Alu* annotation. One can also perform a BLASTN search of U93237 against the *Alu* database.

Fig. 19.3.13 shows the top portion of the list of database hits for the BLASTN search of U93237 against human ESTs. Fig. 19.3.14 presents one alignment of the query sequence with an EST containing an *Alu* repeat. Vertical bars connect bases that are identical in the query and subject. Since both the query and the subject are nucleotide sequences, the numbers shown refer to nucleotide sequence locations. Nucleotides in the query sequence that have been masked for low complexity are replaced by the letter n (e.g., nnnnnnn in the third block). BLASTN checks the plus and minus strand of the query against the plus strand of the database sequence. The alignment shown here presents the plus strand of the query to the minus strand of the database sequence, which is equivalent to showing the minus strand of the query against the plus strand of the database sequence.

BLASTN is optimized for speed, not sensitivity. The algorithm uses a simple system to assess the alignments; matches are given a positive score, mismatches a negative score. BLASTN should only be used when a direct comparison of nucleotide sequences is desired. BLASTN results should not be used to make any predictions about the functions of the encoded proteins. Generally, searches that involve protein comparisons are more sensitive, owing to the larger alphabet (twenty residues versus four nucleotides) and the

```
gb|AA634227|AA634227 ac72cll.sl Stratagene lung (#937210) Homo sapiens cDNA clone 868148
3' similar to contains Alu repetitive element;contains
L1.tl L1 repetitive element ;
Length = 462

Score = 248 bits (125), Expect = 6e-63
Identities = 237/275 (86%), Positives = 237/275 (86%), Gaps = 11/275 (4%)

Query: 3459 ctgtogccaggctggagtgagggtgcaatctggctcactgcaagctctgcoctctg 3518
      ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||| ||| ||| |
Sbjct: 339 ctgtcaccaggctggagtgagggtgcaatctggctcactgcaagctctgcoctctg 280

Query: 3519 otttcaatgaccatttctctgcoctctgcoctcagcoctctgagttagctgggactacaggagc 3578
      ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||| ||| ||| |
Sbjct: 279 ggttccacaccattctcc-----tgcoctcagcoctctgagttagctgggactacaggagc 227

Query: 3579 ctgaccaccatgcoctggctaannnnnnnnnggatttttagtagagacagggtttccaccatg 3638
      ||||||| ||| ||||||| ||||||| ||||||| ||||||| ||||||| ||| ||| ||| |
Sbjct: 226 ctgaccaccatgcoctggctaa---tttttgtagtagagacagggtttccaccatg 171

Query: 3639 ttagccaggatggtctgatctctctgacctgtgatacgcgcoctggcoctccaaagt 3698
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||| ||| ||| |
Sbjct: 170 ttagccaggatggtctgatctctctgacctgtgatacgcgcoctggcoctccaaagt 111

Query: 3699 gctgggattacaggctgagccaaccacacggctc 3733
      ||||||| ||||||| ||||||| ||||| |||||
Sbjct: 110 gctgggattacaggctgagccaaccacacggctc 76
```

Figure 19.3.14 Example of a BLASTN alignment.

degeneracy of the genetic code (two codons, with different third bases often coding for the same protein). The more sophisticated scoring system used with protein comparisons also takes into account similarities between different residues.

PSI-BLAST

A major new feature of the BLAST version 2.0 is the iterative search. Position-specific iterated BLAST (PSI-BLAST) first performs a normal BLASTP run, but then internally computes a profile using the most significant hits (Altschul et al., 1997). A profile may be understood as a table that lists the probability of finding a residue at each position in a conserved protein domain. Conserved regions between the query and the most-significant matches are used to compute the profile. These regions are presumably important to the function of the protein. The profile is then used to perform another search (or iteration), at which point the new hits are also added into the profile. The cycle may then be repeated, with a new profile computed from the results of the last iteration. Each PSI-BLAST iteration takes into account the most significant sequences found in the previous round (i.e., those below the threshold for inclusion in the profile). The algorithm is more sensitive than BLASTP, and more likely to lead to the discovery of distantly related sequences. At present, such iterative searching is only available with the BLASTP program.

The PSI-BLAST output is different from the other programs. Figure 19.3.15 shows the one-line summaries from the first PSI-BLAST iteration (i.e., the first search using the profile) of a mouse adenosine deaminase (Swiss-Prot accession P03958) against the nr database. The one-line summaries are divided into two groups, those with an E value better (lower) than that used to build the profile (Fig. 19.3.15; the default value is 0.001), and those with a worse E value (not shown). Matches better than the threshold, which will be included in the profile, are checked. The user may check (or uncheck) matches, causing them to be included (or excluded) from the profile used in the next iteration. Matches found on a previous iteration are marked with a green ball. New sequences are

Sequences with E-value BETTER than threshold

Sequences producing significant alignments:			Score	E	
			(bits)	Value	
<input checked="" type="checkbox"/>	sp P03958 ADA_MOUSE	ADENOSINE DEAMINASE (ADENOSINE AMINOHYDROLA...	546	e-155	
<input checked="" type="checkbox"/>	gi 1518868 (M34252)	adenosine deaminase (ADA) [Mus musculus]	544	e-154	
<input checked="" type="checkbox"/>	sp P00813 ADA_HUMAN	ADENOSINE DEAMINASE (ADENOSINE AMINOHYDROLA...	543	e-154	
<input checked="" type="checkbox"/>	gi 178075 (K02567)	adenosine deaminase [Homo sapiens]	541	e-153	
<input checked="" type="checkbox"/>	prf j1203203A	deaminase a,adenosine [Homo sapiens]	541	e-153	
<input checked="" type="checkbox"/>	pdb 1ADD _	Adenosine Deaminase (E.C.3.5.4.4) Complexed With 1-...	541	e-153	
<input checked="" type="checkbox"/>	pdb 1FKW _	Murine Adenosine Deaminase (D295e) Zinc Cofactor, T...	539	e-152	
<input checked="" type="checkbox"/>	pdb 1FKK _	Murine Adenosine Deaminase (D296a) Zinc Cofactor, T...	537	e-152	
<input checked="" type="checkbox"/>	pdb 1UIP _	Adenosine Deaminase (His 238 Glu Mutant) Hydrolase,...	537	e-152	
<input checked="" type="checkbox"/>	pdb 1UIO _	Adenosine Deaminase (His 238 Ala Mutant) Hydrolase,...	536	e-152	
<input checked="" type="checkbox"/>	gnl PID el246378 (Z97053)	adenosine deaminase (ADA) [Homo sapiens]	479	e-135	
<input checked="" type="checkbox"/>	gi 1197210 (X02189)	adenosine deaminase [Homo sapiens]	444	e-124	
<input checked="" type="checkbox"/>	sp P22333 ADD_ECOLI	ADENOSINE DEAMINASE (ADENOSINE AMINOHYDROLA...	404	e-112	
<input checked="" type="checkbox"/>	pir jA37943	adenosine deaminase (EC 3.5.4.4) - Escherichia coli...	400	e-111	
<input checked="" type="checkbox"/>	gi 1397267 (U61947)	Similar to adenosine deaminase. [Caenorhabd...	393	e-108	
<input checked="" type="checkbox"/>	gnl PID el251145 (AL021841)	add [Mycobacterium tuberculosis]	368	e-101	
<input checked="" type="checkbox"/>	sp P53984 ADD_STRVG	ADENOSINE DEAMINASE (ADENOSINE AMINOHYDROLA...	345	2e-94	
<input checked="" type="checkbox"/>	gnl PID el310879 (AL031035)	putative adenosine deaminase [Strep...	344	4e-94	
<input checked="" type="checkbox"/>	gnl PID d1016109 (D90807)	Adenosine deaminase (EC 3.5.4.4) (Ade...	340	6e-93	
<input checked="" type="checkbox"/>	pir jS73031	adenosine deaminase (EC 3.5.4.4) - Mycobacterium le...	335	3e-91	
<input checked="" type="checkbox"/>	sp P53909 ADA_YEAST	PROBABLE ADENOSINE DEAMINASE (ADENOSINE AMI...	312	2e-84	
<input checked="" type="checkbox"/>	gi 178079 (K00509)	adenosine deaminase [Homo sapiens]	310	6e-84	
<input checked="" type="checkbox"/>	prf j1001165A	deaminase,adenosine [Homo sapiens]	227	1e-58	
<input checked="" type="checkbox"/>	gi 861309 (U28928)	coded for by C. elegans cDNA yk20f6.3; coded...	227	1e-58	
<input checked="" type="checkbox"/>	gi 3322300 (AE001189)	adenosine deaminase, putative [Treponema ...	160	2e-38	
NEW	<input checked="" type="checkbox"/>	gnl PID d1012479 (D83125)	secretory component [Sarcophaga peregr...	98	9e-20
NEW	<input checked="" type="checkbox"/>	gi 171053 (M30449)	AMP deaminase (EC 3.5.4.6) [Saccharomyces ce...	58	1e-07
NEW	<input checked="" type="checkbox"/>	sp P15274 AMDM_YEAST	AMP DEAMINASE (MIOADENYLATE DEAMINASE) >gi...	58	1e-07
NEW	<input checked="" type="checkbox"/>	gi 2149616 (U90888)	AMP deaminase isoform C [Rattus norvegicus]	57	2e-07
NEW	<input checked="" type="checkbox"/>	sp O09178 AMD3_RAT	AMP DEAMINASE 3 (AMP DEAMINASE ISOFORM E)	57	2e-07
NEW	<input checked="" type="checkbox"/>	sp O08739 AMD3_MOUSE	AMP DEAMINASE 3 (AMP DEAMINASE ISOFORM E) ...	56	3e-07
NEW	<input checked="" type="checkbox"/>	sp Q01432 AMD3_HUMAN	AMP DEAMINASE 3 (AMP DEAMINASE ISOFORM E) ...	56	3e-07
NEW	<input checked="" type="checkbox"/>	pir jS68146	AMP deaminase (EC 3.5.4.6) splice form la - human >...	56	3e-07
NEW	<input checked="" type="checkbox"/>	pir jS68147	AMP deaminase (EC 3.5.4.6) splice form lc - human >...	56	3e-07

Figure 19.3.15 Example of the hit list from a PSI-BLAST report.

marked with a yellow “new”. The page also alerts the user when convergence is achieved (i.e., no new sequences below the threshold for inclusion in the profile were found in that iteration).

In the hit list in Figure 19.3.15, the first iteration of PSI-BLAST finds many adenosine deaminase sequences (marked with a green ball) found in the previous (standard) BLASTP search. In addition, PSI-BLAST identifies a number of AMP deaminase sequences (marked with “new”). Indeed, there is a well-known homology between adenosine deaminase and AMP deaminase (Chang et al., 1991; Holm and Sander, 1997) that would not normally be found by BLASTP, but is found by PSI-BLAST.

If one wants to report the *E* value for a hit found with a PSI-BLAST search, one should use the *E* value calculated during the first iteration that the match is below the threshold

for inclusion in the profile (i.e., when the hit is marked with a “new” flag). However, if the hit is found in the first BLASTP run, then the reported *E* value should be from this run. In either case, in subsequent iterations, the hit itself is used to compute the profile, and the reported *E* value is not representative of the true value.

SEARCHING STRATEGIES

Filtering for Low-Complexity Regions

Low-complexity regions tend to result in misleading database search results. The residue frequencies within such regions differ dramatically from the database as a whole, disrupting the statistics used by BLAST. This can produce alignments that are based purely on compositional bias rather than a significant position-by-position alignment. By default, low-complexity sequences are filtered out of query sequences and replaced by strings of n (nucleotide) or X (protein). A BLASTP search of a proline-rich, DNA-directed RNA polymerase (Swiss-Prot accession P11414) against the nr database illustrates the differences in output resulting from a filtered or unfiltered query sequence. Twenty-four database matches with an *E* value better than 10 are found if filtering is enabled; 4873 are found if filtering is disabled.

The top-scoring database matches for the unfiltered search are shown in Figure 19.3.16; high-scoring matches that are not found in a filtered search are marked. Note that the definition lines of most of the new matches indicate that the sequences are proline-rich, like the query, but do not provide information about the function of the protein. Other hits to RNA polymerases, found near the top of the list in a filtered search, are buried much deeper in the list of matches. The match to “splP14248,” apparently a good match based on the definition line, does not appear in a filtered BLAST search. An examination of the alignment (not shown) indicates that the match is based on solely on proline-rich regions.

Reporting BLAST Results

The most reliable indicator of the importance of a BLAST alignment is the Expect (*E*) value (the number of chance database matches one expects to see at the same score). This number takes into account the length of the query and size of the database as well as the scoring system. Neither the bit nor the raw score is a reliable indicator, as their significance cannot be judged independently of the information used to calculate the *E* value. Normal intuition fails when one is faced with databases of the current size (e.g., the EST database contains ~700 million base pairs at this time), as chance alignments are likely and it is impossible for a user to know at what score matches become significant. Statistical significance also varies from database to database (e.g., a hit with a certain score may be statistically significant in a search of the relatively small month database, but not in a search of nr), so that reporting a certain score without the context of the database size can be misleading. The percentage identity is also a poor indicator of statistical significance for the same reasons; additionally, there is really no way to estimate an *E* value or even a score from a percentage identity, making it impossible to even guess whether a hit could be due to chance or not.

Analyzing mRNA Sequences

1. Identify the open reading frame to obtain the protein sequence. Perform a BLASTX search of the cDNA sequence against the nr protein database. Any resulting hits may identify the correct reading frame for the protein sequence and provide information about its function.

Sequences producing significant alignments:			Score (bits)	E Value
<input type="checkbox"/>	sp P11414 RPB1_CRIGR	DNA-DIRECTED RNA POLYMERASE II LARGEST SUB...	1004	0.0
<input type="checkbox"/>	pir A28490	DNA-directed RNA polymerase (EC 2.7.7.6) II large c...	1002	0.0
<input type="checkbox"/>	dbj D87293_1	(D87293) RNA polymerase II largest subunit [Crice...	982	0.0
<input type="checkbox"/>	sp P24928 RPB1_HUMAN	DNA-DIRECTED RNA POLYMERASE II LARGEST SUB...	982	0.0
<input type="checkbox"/>	gi 825713	(X74874) RNA polymerase II largest subunit [Homo sapi...	982	0.0
<input type="checkbox"/>	sp P08775 RPB1_MOUSE	DNA-DIRECTED RNA POLYMERASE II LARGEST SUB...	980	0.0
<input type="checkbox"/>	gi 2145091	(U37500) RNA polymerase II largest subunit [Mus musc...	966	0.0
<input type="checkbox"/>	gi 1255805	(U53333) <i>C. elegans</i> DNA-directed RNA polymerase II l...	538	e-152
<input type="checkbox"/>	sp P16356 RPB1_CAEEL	DNA-DIRECTED RNA POLYMERASE II LARGEST SUB...	538	e-152
<input type="checkbox"/>	sp P18616 RPB1_ARATH	DNA-DIRECTED RNA POLYMERASE II LARGEST SUB...	505	e-142
<input type="checkbox"/>	pir S14181	DNA-directed RNA polymerase (EC 2.7.7.6) largest ch...	491	e-138
<input type="checkbox"/>	pir S14182	DNA-directed RNA polymerase (EC 2.7.7.6) largest ch...	491	e-138
<input type="checkbox"/>	gi 18734	(X52493) DNA-directed RNA polymerase [Glycine max]	491	e-138
<input type="checkbox"/>	sp P31635 RPB0_ARATH	DNA-DIRECTED RNA POLYMERASE II LARGEST SUB...	489	e-137
<input type="checkbox"/>	sp P04052 RPB1_DROME	DNA-DIRECTED RNA POLYMERASE II LARGEST SUB...	476	e-134
<input type="checkbox"/>	gi 18732	(X52492) DNA-directed RNA polymerase [Glycine max]	451	e-126
<input type="checkbox"/>	gi 158148	(M19537) RNA polymerase II largest subunit (, EC 2.7....	439	e-122
<input type="checkbox"/>	sp P36594 RPB1_SCHPO	DNA-DIRECTED RNA POLYMERASE II LARGEST SUB...	420	e-117
<input type="checkbox"/>	pir S14183	DNA-directed RNA polymerase (EC 2.7.7.6) largest ch...	420	e-117
<input type="checkbox"/>	sp P04050 RPB1_YEAST	DNA-DIRECTED RNA POLYMERASE II LARGEST SUB...	382	e-105
<input type="checkbox"/>	gi 4398	(X03128) put. RNA polymerase II largest subunit [Saccha...	363	1e-99
<input type="checkbox"/>	sp P35084 RPB1_DICDI	DNA-DIRECTED RNA POLYMERASE II LARGEST SUB...	359	2e-98
<input type="checkbox"/>	sp P13983 EXTN_TOBAC	EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLI...	292	3e-78
<input type="checkbox"/>	gi 3172136	(U90211) RNA polymerase II largest subunit [Acantham...	228	7e-59
<input type="checkbox"/>	prf 1814452C	Hyp-rich glycoprotein [Zea diploperennis]	226	2e-58
<input type="checkbox"/>	pir S22456	hydroxyproline-rich glycoprotein - perennial teosin...	222	3e-57
<input type="checkbox"/>	gi 1015937	(X91836) extensin class 1 protein [Vigna unguiculata]	222	3e-57
<input type="checkbox"/>	prf 1814452B	Hyp-rich glycoprotein [Zea mays]	215	3e-55
<input type="checkbox"/>	pir JQ0985	hydroxyproline-rich glycoprotein precursor - maize ...	214	7e-55
<input type="checkbox"/>	pir S23760	polyphenolic adhesive protein - blue mussel (fragme...	214	9e-55
<input type="checkbox"/>	pir S68957	adhesive protein - <i>Mytilus galloprovincialis</i> >gi 96...	212	3e-54
<input type="checkbox"/>	sp P14248 RPB1_PLAFD	DNA-DIRECTED RNA POLYMERASE II LARGEST SUB...	210	2e-53
<input type="checkbox"/>	gi 727264	(U18791) hydroxyproline-rich glycoprotein precursor [...	206	2e-52
<input type="checkbox"/>	pir S20500	hydroxyproline-rich glycoprotein - rice >gi 433816 ...	203	2e-51
<input type="checkbox"/>	gi 3135306	(AF053356) zonadhesin [Homo sapiens]	192	3e-48
<input type="checkbox"/>	sp P14918 EXTN_MAIZE	EXTENSIN PRECURSOR (PROLINE-RICH GLYCOPROT...	190	1e-47

Figure 19.3.16 Example of a hit list from a BLASTP report in which the query sequence was not filtered. Black squares, added manually by the authors, indicate hits that would not appear if the query had been filtered.

- Determine whether the protein sequence is similar to that of any previously identified proteins. Perform a BLASTP and a PSI-BLAST (iterative) search against the nr protein database using the protein sequence as a query.
- Determine whether the protein sequence is similar to the translation products of any uncharacterized DNA sequences whose translations are not in nr. Perform a TBLASTN search of the EST, STS, GSS, and HTGS databases.
- After the initial search, perform a monthly search of the month nucleotide and protein sequence databases to provide information about newly added sequences. A BLASTP

search of the month protein database will reveal newly characterized protein sequences. A TBLASTN search of the month nucleotide database will reveal new open reading frames in nr, EST, HTGS, STS, and GSS. PSI-BLAST searches of nr should also be conducted occasionally, as new sequences may be added to the profile that will, in turn, help identify more distantly related sequences.

5. If the protein sequence contains multiple functional domains, it may be useful to perform the searches with each of these domains individually.

Analyzing Genomic DNA Sequences

1. Identify a potential transcribed segment. Perform a BLASTN search of the nucleotide sequence against the DNA sequences in the EST database to determine if any mRNAs are derived from the genomic sequence. EST hits will highlight the location of exons. A BLASTN search against nr may identify transcripts or previously characterized genomic segments.
2. Identify potential open reading frames. Perform a BLASTX search against the nr protein database. Any resulting hits may identify the correct reading frame of a protein sequence and provide information about its function. If no results are found with the previous methods, a TBLASTX search (translated nucleotide query versus a translated nucleotide database) against EST, GSS, HTGS, and STS may identify other open reading frames.
3. After the initial search, a monthly search of the month nucleotide (with BLASTN) and protein (with BLASTX) sequence databases will provide information about newly added sequences.
4. Any potential transcribed sequences can be analyzed as per the above instructions for mRNA sequences.

Searching Short Sequences

Short sequences can only produce short alignments. Such short alignments need to be relatively strong (i.e., have a high percentage of matching residues) to rise above the background noise. Short, but strong, alignments are more easily detected using a matrix with a higher relative entropy than that of the default BLOSUM-62 (Altschul, 1991). Relative entropy is basically the average information available per position to distinguish the alignment from chance. The BLOSUM series of matrices does not include any that are suitable for the shortest queries, so it is recommended to use the PAM matrices instead (Dayhoff et al., 1978; Schwartz and Dayhoff, 1978). These matrices may be chosen from a menu on the Advanced BLAST page. Suggested matrices for different query lengths are shown in Table 19.3.2 (S.F. Altschul, pers. comm.).

Table 19.3.2 Substitution Matrices for Short Sequences

Query length (amino acids)	Matrix
<35	PAM-30
35-50	PAM-70
50-85	BLOSUM-80
>85	BLOSUM-62

SEQUENCE ALIGNMENT ALGORITHMS

There are two types of pairwise sequence alignments, global and local. Global strategies attempt to align along the entire length of both sequences, while local alignments focus on sequence similarity over shorter regions. Historically, local alignments have been very successful at identifying similarities between proteins. Subsequences of similarity, surrounded by unrelated sequences, can identify, for example, active sites that are conserved between proteins or exons in genomic sequence. Smith and Waterman (1981) were the first to develop a local alignment algorithm. However, such rigorous alignment algorithms are often quite slow and impractical for searching current databases on a routine basis.

Heuristic algorithms implement certain shortcuts to allow a comparison against an entire database within a reasonable time. They are not exhaustive searches in the sense that they do not explore possibilities that seem unlikely to be interesting. Some programs allow the user to adjust parameters that increase the sensitivity of the search (i.e., explore more possibilities), but this must be balanced against the slower speed of the search. Studies have shown that heuristic algorithms, with default settings, approach the sensitivity of exhaustive searches (Altschul et al., 1997). BLAST (Altschul et al., 1990, 1997) and FASTA (Pearson, 1990; Pearson and Lipman, 1988) are two popular heuristic programs. The rest of this discussion will focus on the newest version (2.0) of BLAST.

BLAST calculates a statistical significance for every alignment it produces. Significance is expressed in terms of the *E* value, which is the number of matches with a given score (the calculation of the score is described below) that one would expect by chance. The *E* value depends on the size of the database and the length of the query, as increasing either of these parameters increases the number of chance matches. Calculation of the *E* value is performed with Karlin-Altschul statistics (Karlin and Altschul, 1990, 1993).

The raw score of an alignment is calculated with a scoring system, which is a table of residue substitution scores and penalties for the existence and extension of a gap. For normal protein/protein comparisons, a matrix (i.e., table of residue substitution scores) provides a score for each possible alignment of two amino acids. For example, L aligned with I gets a positive score, while L aligned with E gets a negative score. BLOSUM-62 is the default matrix, and experimentation has shown that it is among the best for detecting weak protein similarities (Henikoff and Henikoff, 1992). PSI-BLAST uses a table that lists the frequency of finding a certain residue at each position in a conserved protein domain. This table is calculated “on the fly” for each iteration from the best matches to the database. BLASTN uses a simple scheme where matches have a positive score and mismatches have a negative score. Gaps are opened or extended when one sequence is longer than another over a certain region. BLAST charges the score $-a$ for the existence of a gap and $-b$ for each letter in a gap, so that a gap of length k letters is penalized $-(a + bk)$. Gap costs of this form are known as affine gap costs. These gap costs may be changed by the user.

The scoring system determines the raw score of an alignment. A bit score can be calculated from the raw score using the following formula:

$$S_{\text{bit}} = [\lambda \times S_{\text{raw}} - \ln(K)] / \ln(2)$$

where λ and K are Karlin-Altschul parameters that depend on the scoring system. The statistical significance corresponding to the resulting bit score is independent of the scoring system used. This allows one to compare the bit scores obtained from two different BLAST runs, each performed using different matrices or gap extension and existence values. Nevertheless, it is more meaningful to compare the results of different BLAST

runs using their *E* values rather than their bit scores, as the *E* value takes into account the size of the database.

A word-based approach is used to find matches in BLAST. An ungapped search, with BLASTN, may be used to demonstrate the concept of a word. First a list is made of all the subsequences in the nucleic acid query sequence that are of a certain length (called the word size), and the positions are recorded. The database is then scanned, looking for words that are in the list. If a match is found between a query and a database word, the alignment is extended (without allowing gaps) until the corresponding score drops a certain amount below the maximum found, at which point that maximum score (and the corresponding alignment) is used. The aligned region is a local alignment. With BLASTN, only exact matches are extended and the default word size is 11 bp.

The procedure is more complicated for the other programs (BLASTP, BLASTX, TBLASTN, and TBLASTX). Exact word matches are not required and the default word size is three residues. A match between a query word and a database word is based on the scoring matrix, which specifies whether two residues result in a positive, negative, or neutral match. If three residues in a query word are compared with three residues in a database sequence, and the score is above a certain threshold, this word warrants further consideration. The default behavior in BLAST version 2.0 is that two words (on the query) must match with two words (on the database sequence) within a window of 40 residues before the match is extended, as described for BLASTN. The effect of looking for two word matches is to eliminate random matches, increasing the speed of BLAST by a factor of three.

Ungapped alignments are used as a starting point for the gapped alignments that BLAST produces. The eleven highest-scoring letters are determined and the center letter is used as the start of the gapped alignment. The alignment is then extended, with gaps allowed. This extension is not exhaustive; if the score drops by more than a certain amount from the maximum found so far, that maximum (and the corresponding alignment) is used (Altschul et al., 1997; Zhang et al., 1998).

The default BLAST parameters are tuned for a balance of speed and sensitivity. It is possible that significant matches may be missed, especially if small databases are searched. In that case, it may be advisable to use a shorter word size for BLASTN or a lower-threshold word for the other programs.

It is important to remember that BLAST calculates similarity. Homology is an evolutionary relationship that must be proved by further analysis.

LITERATURE CITED

- Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., et al. 1991. Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science* 252:1651-1656.
- Altschul, S.F. 1991. Amino acid substitution matrices from an information theoretic perspective. *J. Mol. Biol.* 219:555-565.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Altschul, S.F., Boguski, M.S., Gish, W., and Wootton, J.C. 1994. Issues in searching molecular sequence databases. *Nature Genet.* 6:119-129.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucl. Acids Res.* 25:3389-3402.
- Bairoch, A. and Apweiler, R. 1998. The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1998. *Nucl. Acids Res.* 26:38-42.

- Barker, W.C., Garavelli, J.S., Haft, D.H., Hunt, L.T., Marzec, C.R., Orcutt, B.C., Srinivasarao, G.Y., Yeh, L.S.L., Ledley, R.S., Mewes, H.W., Pfeiffer, F., and Tsugita, A. 1998. The PIR-International Protein Sequence Database. *Nucl. Acids Res.* 26:27-32.
- Benson, D.A., Boguski, M.S., Lipman, D.J., Ostell, J., and Ouellette, B.F. 1998. GenBank. *Nucl. Acids Res.* 26:1-7.
- Boguski, M.S., Lowe, T.M., and Tolstoshev, C.M. 1993. dbEST—database for “expressed sequence tags”. *Nature Genet.* 4:332-333.
- Chandrasekharappa, S.C., Guru, S.C., Manickam, P., Olufemi, S.E., Collins, F.S., Emmert-Buck, M.R., Debelenko, L.V., Zhang, Z., Lubensky, I.A., Liotta, L.A., et al. 1997. Positional cloning of the gene for multiple endocrine neoplasia-type 1. *Science* 276:404-407.
- Chang, Z.Y., Nygaard, P., Chinault, A.C., and Kellems, R.E. 1991. Deduced amino acid sequence of *Escherichia coli* adenosine deaminase reveals evolutionarily conserved amino acid residues: Implications for catalytic function. *Biochemistry* 30:2273-2280.
- Claverie, J.M. and Makalowski, W. 1994. Alu alert. *Nature* 371:752.
- Dayhoff, M.O., Schwartz, R.M., and Orcutt, B.C. 1978. A model of evolutionary change in proteins. In *Atlas of Protein Sequence and Structure*, Vol. 5, suppl. 3. (M.O. Dayhoff, ed.) pp. 345-352. National Biomedical Research Foundation, Washington, D.C.
- Gish, W. and States, D.J. 1993. Identification of protein coding regions by database similarity search. *Nature Genet.* 3:266-272.
- Henikoff, S. and Henikoff, J.G. 1992. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. U.S.A.* 89:10915-10919.
- Holm, L. and Sander, C. 1997. An evolutionary treasure: Unification of a broad set of amidohydrolases related to urease. *Proteins* 28:72-82.
- Karlin, S. and Altschul, S.F. 1990. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci. U.S.A.* 87:2264-2268.
- Karlin, S. and Altschul, S.F. 1993. Applications and statistics for multiple high-scoring segments in molecular sequences. *Proc. Natl. Acad. Sci. U.S.A.* 90:5873-5877.
- Lavin, M.F. and Shiloh, Y. 1997. The genetic defect in ataxia-telangiectasia. *Annu. Rev. Immunol.* 15:177-202.
- Olson, M., Hood, L., Cantor, C., and Botstein, D. 1989. A common language for physical mapping of the human genome. *Science* 245:1434-1435.
- Ostell, J.M. and Kans, J.A. 1998. The NCBI data model. In *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins* (A.D. Baxevanis and B.F.F. Ouellette, eds.) pp. 121-144. John Wiley & Sons, New York.
- Ouellette, B.F.F. 1998. The GenBank sequence database. In *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins* (A.D. Baxevanis and B.F.F. Ouellette, eds.) pp. 16-45. John Wiley & Sons, New York.
- Ouellette, B.F. and Boguski, M.S. 1997. Database divisions and homology search files: A guide for the perplexed. *Genome Res.* 7:952-955.
- Pearson, W.R. 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* 183:63-98.
- Pearson, W.R. and Lipman, D.J. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A.* 85:2444-2448.
- Schuler, G.D. 1998. Sequence alignment and database searching. In *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins* (A.D. Baxevanis and B.F.F. Ouellette, eds.) pp. 145-171. John Wiley & Sons, New York.
- Schwartz, R.M. and Dayhoff, M.O. 1978. Matrices for detecting distant relationships. In *Atlas of Protein Sequence and Structure*, Vol. 5, suppl. 3. (M.O. Dayhoff, ed.) pp. 353-358. National Biomedical Research Foundation, Washington, D.C.
- Seabra, M.C., Brown, M.S., and Goldstein, J.L. 1993. Retinal degeneration in choroideremia: Deficiency of rab geranylgeranyl transferase. *Science* 259:377-381.
- Smith, T.F. and Waterman, M.S. 1981. Identification of common molecular subsequences. *J. Mol. Biol.* 147:195-197.
- Smith, M.W., Holmsen, A.L., Wei, Y.H., Peterson, M., and Evans, G.A. 1994. Genomic sequence sampling: A strategy for high resolution sequence-based physical mapping of complex genomes. *Nature Genet.* 7:40-47.
- Stoesser, G., Moseley, M.A., Sleep, J., McGowran, M., Garcia-Pastor, M., and Sterk, P. 1998. The EMBL nucleotide sequence database. *Nucl. Acids Res.* 26:8-15.

- Tateno, Y., Fukami-Kobayashi, K., Miyazaki, S., Sugawara, H., and Gojobori, T. 1998. DNA Data Bank of Japan at work on genome sequence data. *Nucl. Acids Res.* 26:16-20.
- Wolfsberg, T.G., Straight, P.D., Gerena, R.L., Huovila, A.P., Primakoff, P., Myles, D.G., and White, J.M. 1995. ADAM, a widely distributed and developmentally regulated gene family encoding membrane proteins with a disintegrin and metalloprotease domain. *Dev. Biol.* 169:378-383.
- Wootton, J.C. and Federhen, S. 1993. Statistics of local complexity in amino acid sequences and sequence databases. *Comput. Chem.* 17:149-163.
- Wootton, J.C. and Federhen, S. 1996. Analysis of compositionally biased regions in sequence databases. *Methods Enzymol.* 266:554-571.
- Zhang, J. and Madden, T.L. 1997. PowerBLAST: A new network BLAST application for interactive or automated sequence analysis and annotation. *Genome Res.* 7:649-656.
- Zhang, Z., Berman, P., and Miller, W. 1998. Alignments without low-scoring regions. *J. Comput. Biol.* 5:197-210.

Contributed by Tyra G. Wolfsberg and Thomas L. Madden
National Center for Biotechnology Information
National Library of Medicine, NIH
Bethesda, Maryland

APPENDIX A: BLAST PARAMETERS

The Advanced BLAST page offers the possibility to change a number of BLAST parameters. The most important options supported are listed below.

Descriptions (pull-down menu) restricts the number of one-line descriptions to the number specified. The default is 500.

Alignments (pull-down menu) restricts the number of database hits for which alignments are shown. Several alignments may be associated with one database sequence, so the number of alignments may actually be larger than this number. The default is 500.

Expect value (pull-down menu; also known as the *E* value) is the statistical significance threshold for reporting matches. The default value is 10, such that ten alignments are expected to be found merely by chance, according to the stochastic model of Karlin and Altschul (1990). Lower *E* thresholds are more stringent, leading to fewer chance matches being reported. Fractional values are acceptable.

Filtering (pull-down menu) masks off segments of the query sequence that have low compositional complexity, as determined by the SEG program (Wootton and Federhen, 1993, 1996) or, for BLASTN, by the DUST program (R.L. Tatusov and D.J. Lipman, pers. comm.). Filtering is only applied to the query sequence (or its translation products), not to database sequences. Default filtering is DUST for BLASTN, SEG for other programs.

Genetic code (pull-down menu) selects the genetic code to be used for a BLASTX translation of the query. The default setting is the universal code. The choice of code is determined by the source of the query sequence (e.g., species, mitochondrial).

Graphical overview (checkbox) selects whether an overview of the database sequences aligned to the query sequence will be shown. The score of an alignment is indicated by one of five different colors, which divides the range of scores into five groups. If more than one alignment is displayed for a database sequence, the two alignments are connected by a cross-hatched bar. Mousing over a match causes the definition and score to be shown in the window at the top; clicking on a bar takes the user to the associated alignments. The overview is on by default.

Matrix (menu) allows one to change the matrix from the default of BLOSUM-62. This is useful for searches against short sequences (see Searching Short Sequences). It is also possible to change the gap existence and extension penalties, but the defaults for a given matrix are recommended.

Organism (pull-down menu) limits the search by organism. The menu provides a list of the most common organisms; others may be entered in a text box.

APPENDIX B: SEQUENCE IDENTIFIER SYNTAX

The syntax of sequence header lines used by the NCBI BLAST server depends on the database from which each sequence was obtained. Table 19.3.3 lists the identifiers for the databases from which the sequences were derived. For example, an identifier might be *gb|M73307|AGMA13GT*, where the *gb* tag indicates that the identifier refers to a GenBank sequence, *M73307* is its GenBank accession number, and *AGMA13GT* is its GenBank locus.

NCBI assigns *gi* identifiers for all sequences contained within NCBI's sequence databases (Ostell and Kans, 1998). The *gi* identifier provides a uniform and stable naming convention whereby a specific sequence is assigned its unique *gi* identifier. If a nucleotide or protein sequence changes, a new *gi* identifier is assigned, even if the accession number of the record remains unchanged. Thus, *gi* identifiers provide a mechanism for identifying the exact sequence that was used or retrieved in a given search.

For searches of the *nr* protein database where the sequences are derived from conceptual translations of sequences from the nucleotide databases the *gi* syntax is *gi|gi_identifier*. An example would be *gi|451623 (U04987) env [Simian immunodeficiency...*, where *451623* is the *gi* identifier and *U04987* is the accession number of the nucleotide sequence from which it was derived.

Users may select the *-gi* option for BLAST output, which will produce a header line with the *gi* identifier concatenated with the database identifier of the database from which it was derived. For example, *gi|176485|gb|M73307|AGMA13GT* would be used for a match from a nucleotide database, and *gi|129295|sp|P01013|OVAX_CHICK* for a protein database.

The *gnl* (general) identifier allows databases not listed in Table 19.3.3 to be identified with the same syntax. An example here is the PID identifier *gnl|PID|e1632*. PID stands for Protein-ID, and the *e* in *e1632* indicates that this ID was issued by EMBL. As mentioned above, use of the *-gi* option produces the NCBI *gi* (in addition to the PID), which users can also use to retrieve sequences of interest.

Table 19.3.3 Identifier Syntax for Sequence Databases

Database	Identifier syntax
GenBank	<i>gb accession locus</i>
EMBL Data Library	<i>embl accession locus</i>
DDBJ (DNA Data Bank of Japan)	<i>dbj accession locus</i>
NBRF PIR	<i>pir entry</i>
Protein Research Foundation	<i>prf name</i>
Swiss-Prot	<i>sp accession entry name</i>
Brookhaven Protein Data Bank	<i>pdb entry chain</i>
Patents	<i>pat country number</i>
GenInfo Backbone Id	<i>bbs number</i>
General database identifier	<i>gnl database identifier</i>

Protein databases have become a crucial part of modern biology. Huge amounts of data for protein structures, functions, and particularly sequences are being generated. These data cannot be handled without using computer databases. Searching databases is often the first step in the study of a new protein. Without the prior knowledge obtained from such searches, known information about the protein could be missed, or an experiment could be repeated unnecessarily. Comparison between proteins and protein classification provide information about the relationship between proteins within a genome or across different species, and hence offer much more information than can be obtained by studying only an isolated protein. In this sense, protein comparison through databases allows one to view life as a forest instead of individual trees. In addition, secondary databases derived from experimental databases are also widely available. These databases reorganize and annotate the data or provide predictions. The use of multiple databases often helps researchers understand evolution, structure, and function of a protein.

Protein databases are especially powered by the Internet. Unlike traditional media, such as the CD-ROM, the Internet allows databases to be easily maintained and frequently updated with minimum cost. Researchers with limited resources can afford to set up their own databases and disseminate their data quickly. Notably, many small databases on specific types of proteins, such as the EF-Hand Calcium-Binding Proteins Data Library (http://structbio.vanderbilt.edu/cabp_database/) and O-GlycBase (<http://www.cbs.dtu.dk/databases/OLYCBASE/>), became widely available. Users worldwide can easily access the most up-to-date version through a user-friendly interface. Most protein databases have interactive search engines so that users can specify their needs and obtain the related information interactively. Many protein databases also allow submitters to deposit data interactively, and allow database servers to check the format of the data and provide immediate feedback.

Although some protein databases are widely known, they are far from being fully utilized in the protein science community. This unit provides a starting point for readers to explore the potential of protein databases on the Internet. Databases for different aspects of proteins are discussed, with the focus on sequence, structure, and family. The strengths and weaknesses of the databases will be addressed. For Web addresses of the databases discussed in this unit, see Internet Resources and Table 19.4.1. From hundreds of on-line protein databases, several major databases are discussed as examples to illustrate their features and how they can be used effectively. Most other protein databases can be explored in a similar way.

PROTEIN SEQUENCE DATABASES

Thanks to the Human Genome Project and other sequencing efforts, new sequences have been generated at a prodigious rate. A large number of DNA sequences, which are equivalent to hundreds of new proteins, are determined each day for the human genome alone. It is expected that 10,000 genomes will be sequenced in the next decade. These sequences provide a rich information source and the core of the revolutionary movement toward “large-scale biology.”

Various databases contain protein sequences with different focuses. SwissProt (Bairoch and Apweiler, 1999) provides more annotations than any other sequence database, with a minimal level of redundancy, through human input or integration with other databases. However, the labor-intensive enhancement process prevents many sequences from being added to SwissProt quickly. As a supplement of SwissProt, the TrEMBL database

Contributed by Dong Xu and Ying Xu

Current Protocols in Molecular Biology (2004) 19.4.1-19.4.15

Copyright © 2004 by John Wiley & Sons, Inc.

Table 19.4.1 Web Addresses and Sizes of Selected Protein Databases

Database	Web site	Size ^a
<i>Sequence:</i>		
DDBJ	http://www.ddbj.nig.ac.jp/	25,149,821 sequences
DATA	http://luggagefast.stanford.edu/group/arabprotein/	25,846 sequences
GeneCards	http://bioinfo.weizmann.ac.il/cards/	14,519 genes
Genome Channel	http://compbio.ornl.gov/channel/	NA
KEGG	http://www.genome.ad.jp/kegg/	433,051 genes
nr	http://www.ncbi.nlm.nih.gov/BLAST/	129,6341 sequences
OWL	http://www.bioinf.man.ac.uk/dbbrowser/OWL/	279,796 sequences
PEDANT	http://pedant.mips.biochem.mpg.de	NA
PIR	http://pir.georgetown.edu	1,173,204 sequences
PRF	http://www.genome.ad.jp/hrbin/www_bfind?prf	216,752 sequences
STACK	http://www.sanbi.ac.za/Dbases.html	NA
SwissProt	http://www.expasy.ch/sprot/sprot-top.html	123,192 sequences
SYSTEMS	http://systems.molgen.mpg.de/	290,811 sequences
TIGR Gene Indices	http://www.tigr.org/tdb/tgi/	NA
TrEMBL	http://www.expasy.ch/sprot/sprot-top.html	829,760 sequences
UniGene	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene	NA
<i>Structure:</i>		
3Dee	http://www.compbio.dundee.ac.uk/3Dee/	NA
BioMagResBank	http://www.bmrb.wisc.edu	2494 structures
Decoys 'R' Us	http://dd.stanford.edu/	NA
EBI-MSD	http://msd.ebi.ac.uk	NA
Enzyme Structures	http://www.biochem.ucl.ac.uk/bsm/enzymes/	10208 structures
GRASS	http://trantor.bioc.columbia.edu/GRASS/	NA
Image Library	http://www.imb-jena.de/IMAGE.html	NA
LPFC	http://www-smi.stanford.edu/projects/helix/LPFC/	NA
MolMovDB	http://bioinfo.mbb.yale.edu/MolMovDB/	NA
PDB	http://www.rcsb.org/pdb/	20,473 structures
PDBsum	http://www.biochem.ucl.ac.uk/bsm/pdbsum/	21,361 structures
PRESAGE	http://presage.berkeley.edu/	NA
STING	http://trantor.bioc.columbia.edu/STING/	NA
WPDB	http://www.sdsc.edu/pb/wpdb/	>6000 structures
<i>Sequence family:</i>		
BLOCKS	http://www.blocks.fhcrc.org/	8656 blocks
COG	http://www.ncbi.nlm.nih.gov/COG/	104,101 sequences
DOMO	http://www.infobiogen.fr/services/domo/	83,054 sequences
InterPro	http://www.ebi.ac.uk/interpro/	6383 families
Pfam	http://www.sanger.ac.uk/Pfam/	5193 families
PRINTS	http://www.bioinf.man.ac.uk/dbbrowser/PRINTS/	1750 sites
ProClass	http://pir.georgetown.edu/gfserver/proclass.html	155,868 sequences
ProDom	http://protein.toulouse.inra.fr/prodom.html	481,952 sequences
PROSITE	http://www.expasy.ch/prosite/	1614 sites
SBASE	http://hydra.icgeb.trieste.it/~kristian/SBASE/	338,655 sequences
VIDA	http://www.biochem.ucl.ac.uk/bsm/virus_database/VIDA.html	NA

continued

Table 19.4.1 Web Addresses and Sizes of Selected Protein Databases, *continued*

Database	Web site	Size ^a
<i>Structure family:</i>		
CATH	http://www.biochem.ucl.ac.uk/bsm/cath/	36,480 domains
CE	http://cl.sdsc.edu/ce.html	NA
CL	http://cl.sdsc.edu/cl1.html	NA
Dali Domain Dictionary	http://columba.ebi.ac.uk:8765/holm/ddd2.cgi	771 domain classes
FSSP	http://www2.ebi.ac.uk/dali/fssp/	3242 families
HOMSTRAD	http://www-cryst.bioc.cam.ac.uk/~homstrad/	1033 families
HSSP	http://swift.embl-heidelberg.de/hssp/	NA
PartsList	http://bioinfo.mbb.yale.edu/align/	NA
SCOP	http://scop.mrc-lmb.cam.ac.uk/scop/	17406 structures
SLoop	http://www-cryst.bioc.cam.ac.uk/~sloop/	NA VAST
<i>Function family:</i>		
AARSDb	http://rose.man.poznan.pl.aars.index.html	NA
ABCdb	http://ir2lcb.cnrs-mrs.fr/ABCdb/	NA
AraC/XylS database	http://www.eez.csic.es/arac-xyls/	NA
ASPD	http://wwwmgs.bionet.nsc.ru/mgs/gnw/aspd/	NA
Breast Cancer Gene	http://condor.bcm.tmc.edu/ermb/bcgd/	NA
CSDBase	http://www.chemie.uni-marburg.de/~csdbase/	NA
DEXH/D Family Database	http://www.helicase.net/dexhd/dbhome.htm	NA
EF-hand CaBP	http://structbio.vanderbilt.edu/cabp_database/	NA
EcoCyc	http://ecocyc.org/	NA
ENZYME	http://us.expasy.org/enzyme/	4173 enzymes
Gene Ontology (GO)	http://geneontology.org/	NA
GPCRDB (receptors)	http://www.gpcr.org/7tm/	NA
Homeobox Page	http://www.biosci.ki.se/groups/tbu/homeo.html	NA
MEROPS (peptidase)	http://merops.sanger.ac.uk/	NA
MHCPEP (peptides)	http://wehieh.wehi.edu.au/mhcpep/	13,000 peptides
Nuclear Protein Database (NPD)	http://npd.hgu.mrc.ac.uk/	1227 proteins
O-GlycBase	http://www.cbs.dtu.dk/databases/OGLYCBASE/	242 glycoproteins
PDD	http://www-lecb.ncifcrf.gov/PDD/	NA
PROCAT	http://www.biochem.ucl.ac.uk/bsm/PROCAT/PROCAT.html	NA
Protein Kinase Resource (PKR)	http://pkr.sdsc.edu/html/index.shtml	7588 clusters
RNase P	http://www.mbio.ncsu.edu/RNaseP/home.html	NA
Sentra	http://www-wit.mcs.anl.gov/sentra/	NA
TransportDB	http://66.93.129.133/transporter/wb/index2.html	NA
Tumor Gene	http://condor.bcm.tmc.edu/oncogene.html	>300 genes
WIT2	http://wit.mcs.anl.gov/WIT2/	NA
Wnt gene Homepage	http://www.stanford.edu/~rnuisse/wntwindow.html	NA

continued

Table 19.4.1 Web Addresses and Sizes of Selected Protein Databases, *continued*

Database	Web site	Size ^a
<i>Binding:</i>		
BIND	http://binddb.org	15,141 interactions
DIP	http://dip.doe-mpi.ucla.edu	22,229 interactions
GRID	http://biodata.mshri.on.ca/grid/	13,819 interactions
Het-PDB Navi	http://daisy.bio.nagoya-u.ac.jp/golab/hetpdbnavi.html	NA
MHC-Peptide Interaction Database	http://surya.bic.nus.edu.sg/mpid/	90 complexes
Protein-Protein Interface	http://www-1mmb.ncifcrf.gov:80/-tsai/	NA
ReLiBase	http://relibase.rutgers.edu/	11,938 proteins
<i>Energetics:</i>		
ProTherm	http://www.rtc.riken.go.jp/jouhou/Protherm/protherm.html	13,046 entries
<i>References:</i>		
DBCAT (Catalog of Databases)	http://www.infobiogen.fr/services/dbcat/	511 databases
MEDLINE	http://www.ncbi.nlm.nih.gov/PubMed/	>12,000,000 citations
SeqAnalRef	http://www.expasy.ch/seqanalref/	NA
<i>Combined:</i>		
3DinSight	http://www.rtc.riken.go.jp/jouhou/3dinsight/3dinsight.html	NA
Entrez	http://www.ncbi.nlm.nih.gov/Entrez/	NA
SRS	http://srs.ebi.ac.uk	264 databases

^aNA, size not available at time of printing. The data are as of March, 2003.

(Bairoch and Apweiler, 1999) contains all the translations of EMBL nucleotide sequence entries (Rodriguez-Tom et al., 1996) that have not been integrated into SwissProt. Another protein sequence database, the Protein Identification Resource (PIR; Barker et al., 1999), attempts to build a complete and nonredundant database from a number of protein and nucleic acid sequence databases. Identical and highly similar sequences from the same species are merged into a single entry. Each entry in PIR provides bibliographic and annotated information for the protein. The nr protein database is used for BLAST searching (Altschul et al., 1997), which is described in *UNIT 19.3* of this book. It includes entries from the nonredundant GenBank (Benson et al., 1999) translations, SwissProt, PIR, Protein Research Foundation (PRF) in Japan, and the Protein Data Bank (PDB). Only entries with absolutely identical sequences are merged.

Most of the sequence databases have a sequence search tool and cross-references to entries of other protein and gene databases. Many sequence databases, such as SwissProt and PIR, also provide text searching using, for instance, protein names or key words. To study a new protein, the authors recommend first performing a sequence search using BLAST in nr if the protein sequence is available. The search often gives entry names in the protein databases included in nr. Even when the protein is not found in nr, it is likely that a homologous protein will be hit, which can often lead to some useful information, such as the function of the query protein. If the sequence of the query protein is unavailable, doing a text search in SwissProt or PIR usually identifies the protein. SwissProt is probably the place to obtain the most information about a protein if it can be found in SwissProt. However, some additional information may be found by checking other sequence databases. For example, PIR provides some useful information on protein family classification for each entry, and the Kyoto Encyclopedia of Genes and Genomes (KEGG; Ogata et al., 1999) annotates some gene entries with information about

metabolic and regulatory pathways. If the query protein is not available in nr, one may find a gene model (predicted protein sequence) by searching genetic sequence databases that contain predicted genes from nucleotide sequences (e.g., the Genome Channel; Mural et al., 1999). Although predicted sequences generated by computational gene-finding tools may contain errors, such databases cover a large number of proteins and are often reliable enough to provide useful information.

SwissProt, as a curated protein sequence database, offers a wide range of annotations, covering areas such as function, domain parsing, post-translational modifications, and variants. SwissProt can be accessed at <http://www.expasy.ch/sprot/sprot-top.html> (home server, Switzerland), <http://us.expasy.org/sprot/sprot-top.html> (USA), <http://ca.expasy.org/sprot/sprot-top.html> (Canada), <http://cn.expasy.org/sprot/sprot-top.html> (China), <http://kr.expasy.org/sprot/sprot-top.html> (Korea), and <http://tw.expasy.org/sprot/sprot-top.html> (Taiwan). SwissProt is integrated into many other databases, such as SRS (see SRS below).

Human vitronectin is used here as an example for searching protein sequence databases. To locate the SwissProt entry for this protein, one can search either the entry name (VTNC_HUMAN) or the accession number (P04004) obtained from a BLAST search. Alternatively, one can use the full-text search at the SwissProt Web page to search by protein name (human vitronectin) or key words (e.g., serum spreading, as vitronectin is also called serum spreading factor s-protein). A combination of several entries can be used in a search.

The entry name in SwissProt has the general format X_Y, where X is a mnemonic code of up to four characters indicating the protein name (in this case, VTNC), and Y is a mnemonic species identification code of up to five characters for the biological source of the protein. Some codes used for Y are full English names, e.g., HORSE, HUMAN, MAIZE, MOUSE, PIG, RAT, SHEEP, YEAST (baker's yeast, *Saccharomyces cerevisiae*), and WHEAT. Some are abbreviations, including BOVIN (bovine), CHICK (chicken), ECOLI (*Escherichia coli*), PEA (garden pea, *Pisum sativum*), RABIT (rabbit), SOYBN (soybean, *Glycine max*), and TOBAC (common tobacco, *Nicotina tabacum*).

An entry name may have several accession numbers if they have been merged. An accession number is always conserved from release to release, and therefore allows unambiguous citation.

Each entry contains the following items shown in table format in the NiceProt View layout: (1) general information about the entry, (2) name and origin of the protein, (3) references providing the protein sequence, (4) comments (e.g., annotated functions, subunits, similar proteins), (5) cross-references (links to other databases), (6) key words, (7) features, and (8) sequence information. The text in the comments entry provides a function annotation for the protein (e.g., "Vitronectin is a cell adhesion and spreading factor found in serum and tissues. Vitronectins interact with glycosaminoglycans and proteoglycans. ."). Cross-references lists the annotations of the protein by other databases, such as GeneCards (Rebhan et al., 1998) and ProDom (Corpet et al., 1999). GeneCards, a database of human genes, shows chromosomal location and the involvement of the protein in certain diseases (if applicable). ProDom contains protein domain families by automated sequence comparisons. Clicking the link to ProDom from SwissProt leads to a nice graphic view for domain parsing, as shown in Figure 19.4.1 for vitronectin.

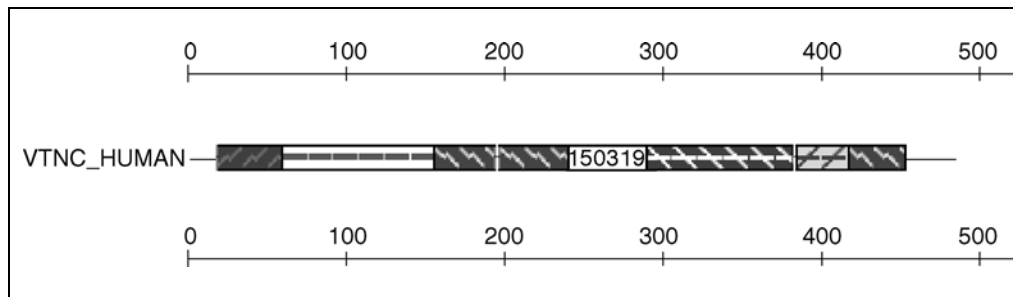


Figure 19.4.1 Domain parsing of human vitronectin by ProDom in SwissProt.

Various research results are given under features. Some of the features items for VTNC_HUMAN are as follows:

SIGNAL	1	19	
CHAIN	20	398	V65 SUBUNIT.
CHAIN	399	478	V10 SUBUNIT.
PEPTIDE	20	63	SOMATOMEDIN B.
DOMAIN	150	287	HEMOPEXIN-LIKE 1.
DOMAIN	288	478	HEMOPEXIN-LIKE 2.
SITE	64	66	CELL ATTACHMENT SITE.
SITE	398	399	CLEAVAGE.
MOD_RES	75	75	SULFATATION.
CARBOHYD	86	86	
DISULFID	293	430	
BINDING	362	395	HEPARIN.
CONFLICT	50	50	C → N (IN REF. 5).

where SIGNAL represents the extent of a signal sequence (prepeptide), MOD_RES indicates a post-translationally modified residue (sulfatation, in this case), CARBOHYD shows the glycosylation site, DISULFID means that a disulfide bond exists between the two indicated residues (293 and 430), and CONFLICT shows that different papers report differing sequences.

PROTEIN STRUCTURAL DATABASES

Searching structure databases is becoming more and more popular in molecular biology. The three-dimensional structures of proteins not only define their biological functions, but also hold a key in rational drug design. Traditionally, protein structures were solved at a low-throughput mode. However, recent advances in new technologies, such as synchrotron radiation sources and high-resolution nuclear magnetic resonance (NMR), accelerate the rate of protein structure determination substantially. There is an overwhelming consensus in the structural biology community that protein structures can be solved en masse (an effort called structural genomics), in a similar fashion as for determining DNA sequences, and that impact of this approach can be compared with that of the Human Genome Project.

The only international repository for the processing and distribution of protein structures is the PDB (Bernstein et al., 1977). The structures in the PDB were determined experimentally by X-ray crystallography (~86%) and NMR (~14%). Theoretical models have been removed from PDB, effective July 2, 2002, based on the new PDB policy. The PDB also contains some structures of chemical ligands and nucleotides. Each PDB entry is represented by a four-character identifier (PDB ID), where the first character is always a number from 0 to 9 (e.g., 1cau, 256b). The PDB can be accessed through the home server (<http://www.rcsb.org/pdb/> or <http://www.pdb.org> in the USA) or through one of

many mirror sites from around the world, as listed at the home server. Some of the mirror sites are as follows:

Rutgers University, Piscataway, NJ, United States: <http://rutgers.rcsb.org/>

NIST, Gaithersburg, MD, United States: <http://nist.rcsb.org/>

Cambridge Crystallographic Data Centre, United Kingdom: <http://pdb.ccdc.cam.ac.uk/>

National University of Singapore, Singapore: <http://pdb.bic.nus.edu.sg/>

Osaka University, Japan: <http://pdb.protein.osaka-u.ac.jp/>

Universidade Federal de Minas Gerais, Brazil: <http://www.pdb.ufmg.br/>

Max Delbrück Center for Molecular Medicine, Germany: <http://www.pdb.mdc-berlin.de/>

The PDB offers three search methods: search by PDB ID, by SearchLite, and by SearchFields. SearchLite is a simple key-word search that uses, for instance, protein name or author's name. A search using SearchFields, as an advanced search engine, allows a user to specify features of the protein, such as Enzyme Commission (EC) number, name of binding ligand, range of protein size, range of resolution in the X-ray structure, and secondary structure content.

The PDB stores structural information in two formats: the PDB file format (Bernstein et al., 1977) and the macromolecular crystallographic information file (mmCIF) format (Bourne et al., 1997). The PDB file format is still the dominant format used in the protein community. It contains three parts: annotations, coordinates, and connectivities. The connectivity part, which shows chemical connectivities between atoms, is optional. It is listed at the end of the PDB file, beginning the line with the key word CONECT. The coordinate part uses each line for a three-dimensional coordinate of an atom, starting from ATOM (for standard amino acids) or HETATM (for nonstandard groups). The following shows an example of the PDB file format:

```

HEADER OXIDOREDUCTASE (OXYGEN(A)) 14-JUN-89 1GOX 1GOX 3
COMPND GLYCOLATE OXIDASE (E.C.1.1.3.1) 1GOX 4
...
ATOM 232 N ALA 29 54.035 4.332 19.352 1.00 23.93 1GOX 374
ATOM 233 CA ALA 29 52.992 65.356 19.569 1.00 24.74 1GOX 375
ATOM 234 C ALA 29 53.519 66.762 19.309 1.00 25.43 1GOX 376
ATOM 235 O ALA 29 54.648 67.179 19.655 1.00 25.66 1GOX 377
ATOM 236 C BALA 29 52.433 65.340 20.993 1.00 24.54 1GOX 378
...
HETATM 3165 O HOH 658 62.480 62.480 0.000 0.50 65.79N 1GOX 3170
CONNECT 2837 2838 2854 1GOX 3171

```

Each line shows the atom serial number, atom type, residue type, chain identifier (in case of multi-chain structure), residue serial number, orthogonal coordinates (three values), occupancy, temperature factor, and segment identifier.

The annotation part of the PDB file format contains dozens of possible record types, including: HEADER (name of protein and release date), COMPND (molecular contents of the entry), SOURCE (biological source), AUTHOR (list of contributors), SS-BOND (disulfide bonds), SLTBRG (salt bridges), SITE (groups comprising important sites), HET (nonstandard groups or residues [heterogens]), MODRES (modifications to standard residues), SEQRES (primary sequence of backbone residues), HELIX (helical substructures), SHEET (sheet substructures), and REMARK (other information and comments).

The PDB allows a user to view a molecule structure interactively through a Virtual Reality Modeling Language (VRML) viewer, RasMol (Sayle and Milner-White, 1995), Chime, or QuickPDB (a Java applet for viewing sequence and structure) when the browser is

configured to support these free rendering tools. The PDB provides related information about the protein, such as secondary structure assignment and geometry. Each PDB entry also links to a wide range of annotations from secondary databases, including (1) summary and display databases such as Graphical Representation and Analysis of Structure Server (GRASS; Nayal et al., 1999), Image Library (Shnel, 1996), Molecular Modelling Database (MMDB; Marchler-Bauer et al., 1999) in Entrez, PDBsum (Laskowski et al., 1997), and Sequence to and within Graphics (STING); (2) domain partition information from 3Dee (Siddiqui and Barton, 1995); (3) the MEDLINE bibliography; (4) structure quality assessment in PDBREPORT from WHAT IF (Vriend, 1990); (5) protein movements recorded in Database of Macromolecular Movement (MolMovDB; Gerstein and Krebs, 1998); (6) structure families (CATH, CE, FSSP, SCOP, and VAST, as discussed later in this unit); and (7) geometry analyses of the protein, such as CSU Contacts of Structural Units (Sobolev et al., 1999) and castP Identification of Protein Pockets & Cavities (Liang et al., 1998).

Several structure databases that are not linked by the PDB can also provide useful information. WPDB (Shindyalov and Bourne, 1995) can be used to visualize and analyze a PDB entry from Microsoft Windows. BioMagResBank (University of Wisconsin, 1999) is a repository for NMR spectroscopy data on proteins, peptides, and nucleic acids. Particularly, it provides partial NMR data (e.g., chemical shifts) before the full structure is solved.

PROTEIN FAMILY DATABASES

Proteins can be classified according to their evolutionary, structural, or functional relationships. A protein in the context of its family is much more informative than the single protein itself. For example, residues conserved across the family often indicate special functional roles. Two proteins classified in the same functional family may suggest that they share similar structures, even when their sequences do not have significant similarity.

There is no unique way to classify proteins into families. Boundaries between different families may be subjective. The choice of classification system depends in part on the problem; in general, the authors suggest looking into classification systems from different databases and comparing them. Three types of classification methods are widely adopted, based upon the similarity of sequence, structure, or function. Sequence-based methods are applicable to any proteins whose sequences are known, while structure-based methods are limited to the proteins of known structures, and function-based methods depend on the functions of proteins being annotated. Sequence- and structure-based classifications can be automated and are scalable to high-throughput data, whereas function-based classification is typically carried out manually. Structure- and function-based methods are more reliable, while sequence-based methods may result in a false positive result when sequence similarity is weak (i.e., two proteins are classified into one family by chance rather than by any biological significance). In addition, since protein structure and function are better conserved than sequence, two proteins having similar structures or similar functions may not be identified through sequence-based methods.

Databases for Sequence-Based Protein Families

Sequence-based protein families are classified according to a profile derived from a multiple-sequence alignment. The profile can be shown across a long domain (typically 100 residues or more) or can be revealed in short sequence motifs. Classification methods based on profiles across long domains tend to be more reliable but less sensitive than those based on short sequence motifs.

Several sequence-based methods focus more on profiles across long domains, including Pfam (Bateman et al., 1999), ProDom (Corpet et al., 1999), SBASE (Murvai et al., 1999),

and Clusters of Orthologous Group (COG; Tatusov et al., 1997). These methods differ in the techniques used to construct families. Pfam builds multiple-sequence alignments of many common protein domains using hidden Markov models. The ProDom protein domain database consists of homologous domains based on recursive PSI-BLAST searches (UNIT 19.3). SBASE is organized through BLAST neighbors and is grouped by standard protein names that designate various functional and structural domains of protein sequences. COG aims toward finding ancient conserved domains by delineating families of orthologs across a wide phylogenetic range.

The following shows an example of Pfam for the GRIP domain (accession number PF01465). Pfam lists some useful information for the entry as follows:

The GRIP (golgin-97, RanBp2alpha, Imh1p and p230/golgin-245) domain is found in many large coiled-coil proteins. It has been shown to be sufficient for targeting to the Golgi. The GRIP domain contains a completely conserved tyrosine residue.

The references of the above annotation are also given. In addition, Pfam gives the alignment between the family members:

```

015045/1511-1558      SAANLEYLKNVLLQFIFLKPG--SERERLLPVINTMLQLSPEEKGLAAV      015045
YNP9_CAEEL/633-681  NEKNMEYLKNVVFQFLKPESVP-AERDQLVIVLQRVLHLSPEVEILKAA      P34562
Q06704/864-909      KNEKIAYIKNVLLGFLEHKE---QRNQLLPVISMILLQLDSTDEKRLVMS      Q06704
Q92805/691-737      REINFEYLKHVVLFKFMSCRES--EAFHLIKAVSVLLNFSQEEENMLKET      Q92805
O42657/703-748      MLIDKEYTRNILFQFLEQRD---RRPEIVNLLSILLDLSEEQKQLLSV      O42657
O70365/1161-1205    EPTEFEYLKVMFEYMMGR----ETKTMKVITTVLKFDPDDQAQKILER      070365
Q21071/692-741D     DPAAEAYLRNVLYRYMTNRESLGKESVTLARVIGTVARFDESQMKNVISS      Q21071
Q18013/574-623      STSEIDYLRNIFTQFLHSMGSPNAASKAILKAMGSVLKVPMAEMKIIDKK      Q18013

```

The alignment shows accession numbers and the range of each sequence. One can identify some features of the family through this pattern (i.e., from particularly conserved residues at specific alignment positions).

Some methods are based on “fingerprints” of small conserved motifs in sequences, as with PROSITE (Hofmann et al., 1999), PRINTS (Attwood et al., 1999), and BLOCKS (Heniko et al., 1999). In protein sequence families, some regions have been better conserved than others during evolution. These regions are generally important for the function of a protein or for the maintenance of its three-dimensional structure, and hence are suitable for fingerprinting. The fingerprints can be used to assign a newly sequenced protein to a specific family. Fingerprints are derived from gapped alignments in PROSITE and PRINTS, but are derived from ungapped alignments (corresponding to the highly conserved regions in proteins) in BLOCKS. A fingerprint in PRINTS may contain several motifs from PROSITE, and thus may be more flexible and powerful than a single PROSITE motif. Therefore, PRINTS can provide a useful adjunct to PROSITE. It should be noted that some functionally unrelated proteins may be classified together due to chance matches in short motifs.

Other sequence-based protein family databases consist of multiple sources. The ProClass database (Wu et al., 1999) is a nonredundant protein database organized according to family relationships as defined collectively by PROSITE patterns and PIR superfamilies. The MEGACLASS server (States et al., 1993) provides classifications by different methods, including Pfam, BLOCKS, PRINTS, ProDom, and SBASE. The MOTIF search engine at <http://motif.genome.ad.jp/> includes PROSITE, BLOCKS, ProDom, and PRINTS.

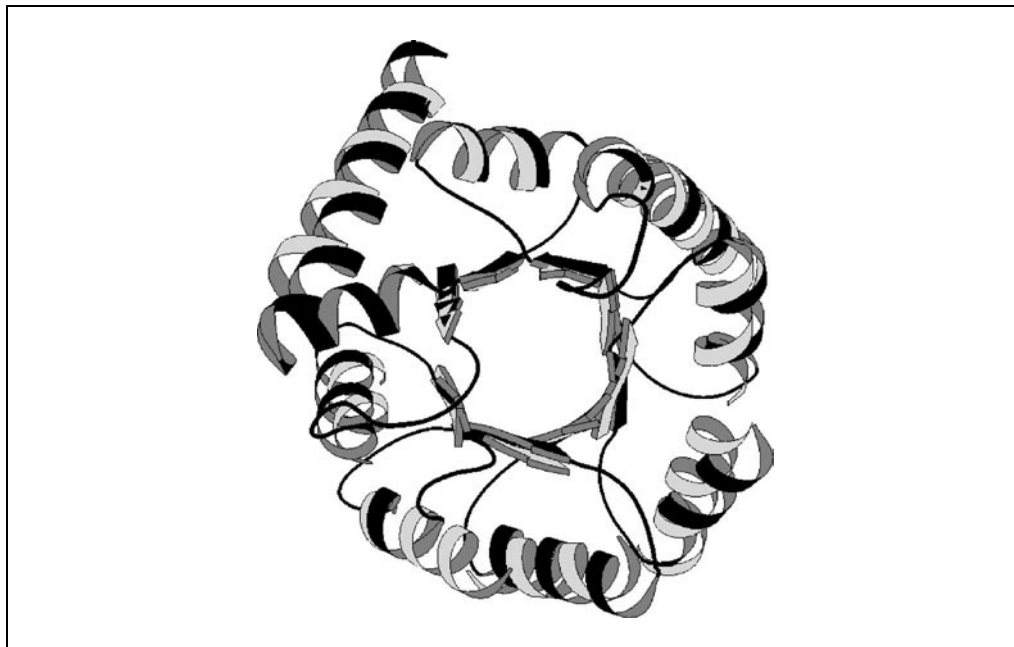


Figure 19.4.2 Structure superposition between glycolate oxidase (1gox, in black) and inosine monophosphate dehydrogenase (1ak5, in gray). This figure was made using MOLSCRIPT (Kraulis, 1991).

Databases for Structure-Based Protein Families

The hierarchical relationship among proteins can be clearly revealed in structures through structure-structure comparison. Structure families often provide more information on the relationship between proteins than what sequence families can offer, particularly when two proteins share a similar structure but no significant sequence identity. Figure 19.4.2 shows an example of a structure-structure alignment between two proteins. Sometimes, sequence similarity between two proteins exists but is not strong enough to produce an unambiguous alignment. In this case, the alignment between two structures can generate better alignment in terms of biological significance, and thus may pinpoint the active sites more accurately.


Different structure-structure comparison methods yield different structure families. CATH (Class, Architecture, Topology and Homologous superfamily; Orengo et al., 1997) is a hierarchical classification of protein domain structures. CE (Combinatorial Extension of the optimal path; Shindyalov and Bourne, 1998) provides structural neighbors of the PDB entries with structure-structure alignments and three-dimensional superpositions. FSSP (Fold classification based on Structure-Structure alignment of Proteins; Holm and Sander, 1996) features a protein family tree and a domain dictionary, in addition to whole-chain-based classification, sequence neighbors, and multiple structure alignments. SCOP (Structural Classification of Proteins; Murzin et al., 1995) uses augmented manual classification, class, fold, superfamily, and family classification. VAST (Vector Alignment Search Tool; Gibrat et al., 1996) contains representative structure alignments and three-dimensional superpositions. Among these five databases, SCOP provides more function-related information. However, due to the manual work involved, SCOP is not updated as frequently as the others (as of June, 2003, it was last updated for the PDB release on March 1, 2003), whereas FSSP and CATH follow the PDB updates closely.

SCOP is used here as an example to show the features of structure-based families. SCOP can be accessed through its home server in the UK (<http://scop.mrc-lmb.cam.ac.uk/scop/>). It is also widely mirrored around the world, including:

<http://pdb.wehi.edu.au/scop/> (Australia)
<http://mdl.ipc.pku.edu.cn/scop/> (China)
<http://www.cdfd.org.in:5555/scop/> (India)
<http://pdb.weizmann.ac.il/scop/> (Israel)
<http://loki.polito.it/scop/> (Italy)
<http://scop.protres.ru/> (Russia)
<http://scop.bic.nus.edu.sg/> (Singapore)
<http://scop.life.nthu.edu.tw/> (Taiwan)
<http://scop.berkeley.edu/> (USA)

SCOP describes the hierarchical relationship among proteins through the major levels of (homologous) family, superfamily, and fold. Proteins are clustered together into a (homologous) family if they have significant sequence similarity. Different families that have low sequence similarity but whose structural and functional features suggest a common evolutionary origin are placed together in a superfamily. Different superfamilies are categorized into a fold if they have the same major secondary structures in the same arrangement and with the same topological connections (the peripheral elements of secondary structure and turn regions may differ in size and conformation). Two superfamilies in the same fold may not have a common evolutionary origin. Their structural similarities may arise from the physics and chemistry of proteins favoring certain packing arrangements and chain topologies (Murzin et al., 1995). Figure 19.4.3 shows the SCOP interface using an example of protein 1gox in the PDB.

Structural Classification of Proteins










































Protein: Glycolate oxidase from spinach (*Spinacia oleracea*)

Lineage:

1. Root: [scop](#)
2. Class: [Alpha and beta proteins](#)
Mainly parallel beta sheets (beta-alpha-beta units)
3. Fold: [beta/alpha \(TIM\)-barrel](#)
contains parallel beta-sheet barrel, closed; n=8, S=8; strand order 12345678
4. Superfamily: [FMN-linked oxidoreductases](#)
5. Family: [FMN-linked oxidoreductases](#)
6. Protein: Glycolate oxidase
7. Species: [spinach \(*Spinacia oleracea*\)](#)

PDB Entry Domains:

1. [1gox](#)                                       

Databases for Function-Based Protein Families

There are various protein functional families classified from different perspectives. The ENZYME data bank (Bairoch, 1993) contains the following data for each enzyme: EC number, recommended name, alternative names, catalytic activity, cofactors, pointers to the SwissProt entry, and pointers to any disease associated with a deficiency of the enzyme. PROCAT is a database of three-dimensional enzyme active site templates (Wallace et al., 1996). PDD (Protein Disease Database; Lemkin et al., 1995; Merril et al., 1995) correlates diseases with proteins observable in serum, urine, and other common human body fluids based on biomedical literature. There are also a growing number of databases dedicated to special types of proteins, such as antibodies, G-protein-coupled receptors, HIV proteases, glycoproteins, and RNases, as shown in Table 19.4.1.

OTHER DATABASES

Protein Binding Databases

Protein binding includes protein-substrate docking and protein-protein association. ReLiBase (Hendlich, 1998) is a database system for analyzing receptor-ligand complexes in the PDB. DIP (Database of Interacting Proteins) records protein pairs that are known to bind with each other. The information in DIP may provide information related to signaling pathways, multiple interactions, and complex systems.

Protein Energetics Databases

There are few databases for protein energetics, due to the low-throughput nature of the data source. One useful energetics database can be found in ProTherm (Thermodynamic Database for Proteins and Mutants; Gromiha et al., 1999). It contains thermodynamic data on mutations, including Gibbs free energy, enthalpy, heat capacity, and transition temperature. These data are important for understanding the structure and stability of proteins.

Bibliographic Databases

Searching for protein information through traditional bibliographic databases, such as MEDLINE or Grateful Med, can be rewarding. In addition, some bibliographic reference databases dedicated to proteins may provide certain information more directly. For example, SeqAnalRef stores papers dealing with sequence analysis.

COMBINED DATABASES

By integrating different types of protein databases together, a database of databases (or a data warehouse) can be built. Such combined databases not only serve as “one-stop shopping,” but also provide cross-references between entries in different databases. Two combined databases, Entrez and SRS, have been very successful.

Entrez

Entrez (Schuler et al., 1996) is a combined database consisting of literature, protein sequence and structure, nucleotide sequence, and taxonomy. Different types of information are interconnected through the grouping of sequences/structures and references by computed similarity scores. Entrez can be used through a variety of media, including CD-ROM, a custom Graphical Interface client, a World Wide Web browser, a command line browser (CLEVER), and the National Center for Biotechnology Information's (NCBI's) toolkit written in C.

SRS

SRS (Sequence Retrieval System; Etzold et al., 1996) is the most comprehensive database for molecular biology. The home server at <http://srs.ebi.ac.uk> supports 264 biological databases (as of March, 2003), including almost all the major protein/genetic databases. As an indexing system, it provides fast access to different databases through searches by sequence or by key words from various data fields. SRS also builds indices using cross-references between databases. An entry from one database can be linked to other databases that contain the entry. However, it should be noted that the contents of SRS may lag behind the other databases in updating (i.e., some new entries in the original databases may not be included in SRS).

SUMMARY

This unit reviews several major protein databases on the Internet, and shows what kind of information users can expect from protein databases. Although all technical procedures cannot be described here, most of the protein databases are easy to use and provide detailed on-line manuals so that even users with little computer skill can learn them quickly.

Protein databases may not always be easily accessible or usable through the Internet. Sometimes a database server may be down or the Internet connection may be interrupted. Some structures or image files are very large (several megabytes), and the download time may be long. It can be helpful to use a mirror site of the database at a close location in order to accelerate the access speed. For a frequent user, it may be worthwhile to install the database on a local machine. On the other hand, it must be kept in mind that a mirror site or a local copy may contain an older version of the database than the one on the home server.

It is important to assess the quality of the data. There are three types of data in protein databases. (1) Experimental data are generally very reliable. However, some entries may contain errors (e.g., some protein sequences) or may be based on low-resolution data (e.g., some protein structures determined by NMR). (2) Annotation data uses computational techniques on experimental data, for example, secondary structure assignment and domain partition in structure. These data depend on the quality of the experimental data and the computational methods used. Different methods may yield different results. (3) Prediction data includes, for example, sequence domain parsing and three-dimensional structure prediction. No matter how good the method, the results are still predictions and should be subjected to experimental verification. In addition, different methods typically give different predictions.

In summary, caution is needed when using the data from databases to draw a conclusion. It is worthwhile to check the same type of data from different databases and compare them. It is sometimes necessary to use additional computational tools (e.g., tools to assess the quality of a structure) for further analysis.

ACKNOWLEDGMENTS

The authors thank Drs. Edward C. Uberbacher, Michael A. Unseren, Jay Snoddy, and Gwo-liang Chen for helpful discussions. This work is supported by the Office of Biological and Environmental Research, U.S. Department of Energy, under Contract DE-AC05-00OR22725, managed by UT-Battelle, LLC.

LITERATURE CITED

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucl. Acids Res.* 25:3389-3402.

- Attwood, T.K., Flower, D.R., Lewis, A.P., Mabey, J.E., Morgan, S.R., Scordis, P., Selley, J., and Wright, W. 1999. PRINTS prepares for the new millennium. *Nucl. Acids Res.* 27:220-225.
- Bairoch, A. 1993. The ENZYME data bank. *Nucl. Acids Res.* 21:3155-3156.
- Bairoch, A. and Apweiler, R. 1999. The SwissProt protein sequence data bank and its supplement TrEMBL in 1999. *Nucl. Acids Res.* 27:49-54.
- Barker, W.C., Garavelli, J.S., McGarvey, P.B., Marzec, C.R., Orcutt, B.C., Srinivasarao, G.Y., Yeh, L.L., Ledley, R.S., Mewes, H., Pfeiffer, F., Tsugita, A., and Wu, C. 1999. The PIR-international protein sequence database. *Nucl. Acids Res.* 27:39-42.
- Bateman, A., Birney, E., Durbin, R., Eddy, S.R., Finn, F.D., and Sonnhammer, E.L.L. 1999. Pfam 3.1:1313 multiple alignments match the majority of proteins. *Nucl. Acids Res.* 27:260-262.
- Benson, D.A., Boguski, M.S., Lipman, D.J., Ostell, J., Ouellette, B.F., Rapp, B.A., and Wheeler, D.L. 1999. Genbank. *Nucl. Acids Res.* 27:12-17.
- Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T., and Tasumi, M. 1977. The protein data bank: A computer based archival file for macromolecular structures. *J. Mol. Biol.* 112:535-542.
- Bourne, P., Berman, H., Watenpaugh, K., Westbrook, J., and Fitzgerald, P. 1997. The macromolecular crystallographic information file (mmCIF). *Methods Enzymol.* 277:571-590.
- Corpet, F., Gouzy, J., and Kahn, D. 1999. Recent improvements of the ProDom database of protein domain families. *Nucl. Acids Res.* 27:263-267.
- Etzold, T., Ulyanov, A., and Argos, P. 1996. SRS: Information retrieval system for molecular biology data banks. *Methods Enzymol.* 266:114-128.
- Gerstein, M. and Krebs, W. 1998. A database of macromolecular motions. *Nucl. Acids Res.* 26:4280-4290.
- Gibrat, J.F., Madej, T., and Bryant, S.H. 1996. Surprising similarities in structure comparison. *Curr. Opinion Struct. Biol.* 6:377-385.
- Gromiha, M.M., An, J., Kono, H., Oobatake, M., Uedaira, H., and Sarai, A. 1999. Protherm: Thermodynamic database for proteins and mutants. *Nucl. Acids Res.* 27:286-288.
- Hendlich, M. 1998. Databases for protein-ligand complexes. *Acta Crystallogr., Sect. D* 1:1178-1182.
- Heniko, J.G., Heniko, S., and Pietrokovski, S. 1999. New features of the blocks database servers. *Nucl. Acids Res.* 27:226-228.
- Hofmann, K., Bucher, P., Falquet, L., and Bairoch, A. 1999. The PROSITE database, its status in 1999. *Nucl. Acids Res.* 27:215-219.
- Holm, L. and Sander, C. 1996. Mapping the protein universe. *Science* 273:595-602.
- Kraulis, P. 1991. MOLSCRIPT—a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24:946-950.
- Laskowski, R.A., Hutchinson, E.G., Michie, A.D., Wallace, A.C., Jones, M.L., and Thornton, J.M. 1997. PDBsum: A web-based database of summaries and analyses of all PDB structures. *Trends Biochem. Sci.* 22:488-490.
- Lemkin, P.F., Orr, G.A., Goldstein, M.P., Creed, G.J., Myrick, J.E., and Merril, C.R. 1995. The protein disease database of human body fluids: I. Computer methods and data issues. *Appl. Theor. Electrophor.* 5:55-72.
- Liang, J., Edelsbrunner, H., and Woodward, C. 1998. Anatomy of protein pockets and cavities: Measurement of binding site geometry and implications for ligand design. *Protein Sci.* 7:1884-1897.
- Marchler-Bauer, A., Address, K.J., Chappay, C., Geer, L., Madej, T., Matsuo, Y., Wang, Y., and Bryant, S.H. 1999. MMDB: Entrez's 3D structure database. *Nucl. Acids Res.* 27:240-243.
- Merril, C.R., Goldstein, M.P., Myrick, J.E., Creed, G.J., and Lemkin, P.F. 1995. The protein disease database of human body fluids: I. Rationale for the development of this database. *Appl. Theor. Electrophor.* 5:49-54.
- Mural, R.J., Parang, M., Shah, M., Snoddy, J., and Uberbacher, E.C. 1999. The Genome Channel: A browser to a uniform first-pass annotation of genomic DNA. *Trends Genet.* 15:38-39.
- Murvai, J., Vlahovicek, K., Barta, E., Szepesvari, C., Acatrinei, C., and Pongor, S. 1999. The SBASE protein domain library, release 6.0: A collection of annotated protein sequence segments. *Nucl. Acids Res.* 27:257-259.
- Murzin, A.G., Brenner, S.E., Hubbard, T., and Chothia, C. 1995. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247:536-540.
- Nayal, M., Hitz, B.C., and Honig, B. 1999. GRASS: A server for the graphical representation and analysis of structures. *Protein Sci.* 8:676-679.
- Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. 1999. KEGG: Kyoto encyclopedia of genes and genomes. *Nucl. Acids Res.* 27:29-34.

- Orengo, C.A., Michie, A.D., Jones, D.T., Swindells, M.B., and Thornton, J.M. 1997. CATH—a hierarchic classification of protein domain structures. *Structure* 5:1093-1108.
- Rebhan, M., Chalifa-Caspi, V., Prilusky, J., and Lancet, D. 1998. GeneCards: A novel functional genomics compendium with automated data mining and query reformulation support. *Bioinformatics* 14:656-664.
- Rodriguez-Tom, P., Stoehr, P.J., Cameron, G.N., and Flores, T.P. 1996. The European Bioinformatics Institute (EBI) databases. *Nucl. Acids Res.* 24:6-13.
- Sayle, R.A. and Milner-White, E.J. 1995. RASMOL: Biomolecular graphics for all. *Trends Biochem. Sci.* 20:374-376.
- Schuler, G.D., Epstein, J.A., Ohkawa, H., and Kans, J.A. 1996. Entrez: Molecular biology database and retrieval system. *Methods Enzymol.* 266:141-162.
- Shindyalov, I.N. and Bourne, P.E. 1995. WPDB: A PC-based tool for analyzing protein structure. *J. Appl. Crystallogr.* 28:847-852.
- Shindyalov, I.N. and Bourne, P.E. 1998. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* 11:739-747.
- Shnel, J. 1996. Image library of biological macromolecules. *Comput. Appl. Biosci.* 12:227-229.
- Siddiqui, A.S. and Barton, G.J. 1995. Continuous and discontinuous domains: An algorithm for the automatic generation of reliable protein domain definitions. *Protein Sci.* 4:872-884.
- Sobolev, V., Sorokine, A., Prilusky, J., Abola, E.E., and Edelman, M. 1999. Automated analysis of interatomic contacts in proteins. *Bioinformatics.* 15:327-332
- States, D.J., Harris, N.L., and Hunter, L. 1993. Computationally efficient representation of classes in protein sequence megaclassification. *Proc. Intl. Syst. Mol. Biol.* 1:387-394.
- Tatusov, R.L., Koonin, E.V., and Lipman, D.J. 1997. A genomic perspective on protein families. *Science* 278:631-637.
- University of Wisconsin. 1999. BioMagResBank. University of Wisconsin, Madison, Wis.
- Vriend, G. 1990. WHAT IF: A molecular modelling and drug design program. *J. Mol. Graphics* 8:52-56.
- Wallace, A.C., Laskowski, R.A., and Thornton, J.M. 1996. Derivation of 3D coordinate templates for searching structural databases: Application to Ser-His-Asp catalytic triads in the serine proteinases and lipases. *Protein Sci.* 5:1001-1013.
- Wu, C., Shivakumar, S., and Huang, H. 1999. Proclass protein family database. *Nucl. Acids Res.* 27:272-274.

INTERNET RESOURCES

The Web addresses of the databases mentioned in this unit are listed in Table 19.4.1. Readers can find more protein databases and their related tools in the following Web pages, which collect a large number of useful links.

<http://compbio.ornl.gov/structure/resource/>

Oak Ridge National Laboratory's resources of protein modeling tools.

<http://www.public.iastate.edu/~pedro/research.tools.html>

Pedro's biomolecular research tools.

<http://www.expasy.org/alinks.html>

Amos' WWW links page.

Contributed by Dong Xu
 Digital Biology Laboratory
 University of Missouri–Columbia
 Columbia, Missouri

Ying Xu
 Department of Biochemistry and Molecular Biology
 University of Georgia
 Athens, Georgia

Basic Protein Sequence Analysis

UNIT 19.5

Protein sequence analysis can be performed from many perspectives and with a vast array of bioinformatics methods. This unit will focus on those analyses that can be performed using only a computer (desktop workstation or laptop), Internet access, and a laboratory notebook to record results. All the resources needed are accessible via publicly available Web servers and databases, and require little or no computational expertise. These tools enable biologists to predict the structure of a protein, the presence of functional motifs or domains, cellular localization (e.g., membrane-bound, secreted, chloroplast, or nuclear), and post-translational modifications. Integration of all these data enables the biologist to make a more informed prediction of the likely molecular functions of proteins of interest.

A plethora of tools are available for these tasks, and in choosing preferred methods for each of the separate protocols, the authors have been guided primarily by two criteria: user-friendliness and utility. Some tools are extremely powerful, but inappropriate for the computational novice. Other tools may be easy to use, but relatively ineffectual. We have kept working biologists in mind, and asked ourselves which tools our biologist colleagues would find easiest to use and understand. With these criteria in mind, we selected one or two protocols for each task. Identifying structural domains using different types of servers are presented (Basic Protocol 1 and Alternate Protocols 1 and 2), as are procedures for predicting transmembrane regions and subcellular localization (Basic Protocol 2 and Alternate Protocol 3) and key functional residues and motifs (Basic Protocol 3). Each procedure also includes a brief discussion of some of the Web servers and databases that we feel deserve attention. Support Protocols for each method present a strategy for the integration of the results from the different procedures (Support Protocols 1, 2, 3, 4, 5, 6, and 7).

IDENTIFYING STRUCTURAL AND FUNCTIONAL DOMAINS USING INTEGRATED META-SERVERS

**BASIC
PROTOCOL 1**

Below is a protocol for the use of integrated meta-servers. Meta-servers include prediction results from several independent prediction tools to produce a consensus prediction. This helps a biologist avoid false positive predictions. Our primary protocol for this task involves the use of the SMART (Simple Modular Architecture Research Tool) server provided by the European Molecular Biology Laboratory (EMBL) at Heidelberg (Schultz et al., 1998; Letunic et al., 2004). This server integrates structure and function prediction tools with pre-calculated analyses (including domain structure prediction, intron/exon boundary detection and ortholog identification) for proteins in the Swiss-Prot and TrEMBL databases (Boeckmann et al., 2003) and all proteomes analyzed by the Ensembl group (Hubbard et al., 2002). The extensive results and appealing interface make a biologist's work much easier, and this is our reason for recommending the use of this server.

Materials

Input data

The sequence to be analyzed should be in FASTA format, or may alternatively be submitted as raw sequence (the single-letter representation of the protein sequence without the definition line). An example FASTA format is shown in Figure 19.5.1. In cases where the sequence of interest is in the Swiss-Prot or

Informatics

19.5.1

```
>sp|P27037|AVR2_HUMAN Activin receptor type II precursor (ACTR-II)
MGAAAKLAFVFLI SCSSGAILGRSETQECLFFNANWEKDR TNQTGVEPCYGDKDKRRHCFATWKNISGS
IEIVKQGCWLD DINCYDR TDCVEKKDSPEVYFCCCEGNMCNEKFSYFP EMEV TQPTS NPVTPKPPYYN IL
LYSLVPLML IAGIVICAFWVYRHHK MAYPPVLVPTQDPGPPPPSPLLGLKPLQLLEVKARGRF GCVWKAQ
LLNEYVAVKIFPIQDKQSWQNEYEVYSLPGMKHENILQFTGAEKRGTSDVDLW LITAFHEKGSLSDFLK
ANVVSWNELCHIAETMARGLAYLHEDI PGLKDGHKPAISHRDIKSKNVLLKNNLTACIADFG LALKFEAG
KSAGDTHGQVGTTRRYMAPEVLEGA INFQRDAFLRIDMYAMGLVLWELASRCTAADGFPVDEYMLPFEEEEIG
QHPSLEDMQEVVVHKKRPVLRDYWQKHAGMAML CETIEECWDHDAEARLSAGCVGERITQMQLTNIIT
TEDIVTVVTMVTNVDFPPKESL
```

Figure 19.5.1 FASTA format of AVR2_HUMAN protein sequence. The Swiss-Prot accession number is P27037 and the Swiss-Prot ID is AVR2_HUMAN.

TrEMBL databases, the accession number is normally sufficient to retrieve precalculated results. Otherwise, the protein sequence is required as described in the protocol below.

1. Point the browser at the SMART Web site (<http://smart.embl-heidelberg.de/>). On the SMART main page that appears, click either SMART MODE: NORMAL or SMART MODE: GENOMIC. In NORMAL SMART, the database contains Swiss-Prot, SP-TrEMBL and stable Ensembl proteomes. In GENOMIC SMART, only the proteomes of completely sequenced genomes are used. We recommend the use of NORMAL SMART. In the page that now appears, paste the sequence of interest in FASTA format in the box provided. If the sequence is from Swiss-Prot, TrEMBL, or the EMBL database, one could alternatively submit the Accession number or Entry name in the box marked Sequence ID or ACC (Fig. 19.5.2).
2. Select additional analyses. At this point, one can either click the Sequence SMART button or select additional analyses. In this example (shown in Fig. 19.5.3), we have selected all available analyses.
 - a. Outlier homologs.

This option will search sequence databases derived from proteins of solved structures and return the ID and E-value (see Table 19.5.1 for definition) of the top hit and the database searched. The search is done on PDB (Protein Data Bank, a database of solved structures) and position-specific iterated BLAST (PSI-BLAST) profiles derived from SCOP (Structural Classification Of Proteins; Table 19.5.1), a system of classifying proteins based on structural and functional similarity (Murzin et al., 1995).
 - b. PFAM domains.

This option searches for the presence of PFAM domains (Bateman et al., 2002) in the query sequence. Since SMART (as of November, 2004) has 667 HMMs, and PFAM (August, 2004) has 7503 hidden Markov models (HMMs), searching PFAM HMMs increases the likelihood of a match.
 - c. Signal peptides.

This gives information on whether the query sequence contains a signal peptide, using SigCleave program.
 - d. Internal repeats.

This option shows any repeated motifs found in the query.
3. Select the Sequence SMART button. This will retrieve precalculated results (when available) or search the sequence against the library of HMMs available.

A

Sequence analysis

You may use either the Swissprot/Sptrembl/Ensembl sequence identifier (**ID**) / accession number (**ACC**) or the protein sequence itself to request the SMART service.

Sequence ID or ACC

Sequence

```
>sp|P27037|AVR2_HUMAN Activin receptor type II precursor (ACTR-II) (ACTRIIA)
MGAAAKLAFAYFLISCSSGAILGRSETQECLFFNANWEKDRTNQT
GVEPCYGDKDKRRHCFATWKNISGS
IEIVKQGCWLDDINCYDRITDCVEKKDSPEVYFCCCEGNMCNEKF
SYFPEMEVTQPTSNPVTPKPPYYNIL
LYSLVPLMLIAGIVICAFWVYRHHKMAYPPVLVPTQDPGPPPPSPL
LGLKPLQLLEVKARGRFGCVWKAQ
```

Sequence SMART Reset

B

Sequence ID or ACC

 ←

Sequence

Figure 19.5.2 SMART sequence submission form. (A) Submission of sequence in FASTA format. (B) Submission using the Swiss-Prot accession number in the box marked with arrow.

Sequence SMART Reset

HMMER searches of the SMART database occur by default. You may also find:

- [Outlier homologues](#) and homologues of known structure
- [PFAM domains](#)
- [signal peptides](#)
- [internal repeats](#)

Figure 19.5.3 Sequence submission options in SMART.

Table 19.5.1 Definitions of Acronyms and Abbreviations Used in this Unit

Acronym/ abbreviation	Full name	Description
3D-PSSM	3-D Position Specific Scoring Matrix	The 3D-PSSM is a protein fold recognition server that uses 1-D and 3-D sequence profiles coupled with secondary structure and solvation potential information
BLAST	Basic Local Alignment Search Tool	BLAST is a sequence comparison algorithm optimized for speed, used to search sequence databases for optimal local alignments to a query. Variants of BLAST enable searches of protein databases using nucleotide queries and vice versa.
CDART	Conserved Domain Architecture Retrieval Tool	For a given protein query sequence, CDART displays the predicted functional/structural domains that make up the protein and lists proteins with similar domain architectures
EMBL	European Molecular Biology Laboratory	EMBL is an international research organization with its main laboratory in Heidelberg, Germany, and four outstations in Hinxton, U.K. (the European Bioinformatics Institute, EBI), Grenoble, France, Hamburg, Germany, and Monterotondo, Italy.
E-value	Expectation value	For a given score <i>S</i> and a database size, the E-value is the number of hits with scores equivalent to or better than <i>S</i> that are expected to occur in a database of that size by chance. The lower the E value, the more significant the score.
HMM	Hidden Markov Model	HMMs for proteins are statistical models (akin to profiles and PSSMs) that represent the amino acid preferences and insertion/deletion likelihoods at each position. They are used to detect remote homologs and to generate alignments of new sequences to a family.
MSA	Multiple Sequence Alignment	A multiple sequence alignment refers to an alignment of more than two sequences; pairwise alignment refers to an alignment of two sequences. Gap characters (typically, dots or dashes) are inserted between amino acids so that all sequences have the same length, and can be arranged in a matrix to display the amino acids conserved across the family. MSAs are used as input to construction of PSSMs and HMMs, to estimate phylogenetic trees, and for numerous other tasks.
NCBI	National Center for Biotechnology Information	The NCBI is the main resource for computational biologists in the U.S., and includes a number of important tools (e.g., BLAST) and databases. Databases include the CDD (Conserved Domain Database), GenBank (a comprehensive sequence database), PubMed (for retrieving papers in biomedical literature), OMIM (Online Mendelian Inheritance in Man), and a host of other important resources.

continued

Table 19.5.1 Definitions of Acronyms and Abbreviations Used in this Unit, *continued*

Acronym/abbreviation	Full name	Description
PFAM	Protein Families	PFAM is a large collection of multiple sequence alignments and hidden Markov models representing many protein domains and families. It has a web interface for identification of domains in a given protein, and includes hyperlinks to relevant literature and functional/structural information for many domains.
PSI-BLAST	Position-Specific Iterated BLAST	An iterative search using the BLAST algorithm and profile construction, enabling identification of remote homologs
PSIPRED		PSIPRED is a simple and reliable secondary-structure prediction method, incorporating two feed-forward neural networks, which perform an analysis on output obtained from PSI-BLAST to predict whether a position is likely to be found in an α helix, β strand, or coil.
PSSM	Position-specific scoring matrix	A PSSM is a statistical representation of the amino acids in a multiple sequence alignment, and is used to provide a log-odds score for each amino acid at each position in the alignment. PSSMs, HMMs, and profiles are related.
RasMol	Raster Display of Molecules.	Molecular visualization freeware to visualize 3-D structure of proteins
RPSBLAST	Reverse PSI-BLAST	RPS-BLAST (Reverse PSI-BLAST) searches a query sequence against a database of profiles (i.e., the reverse of PSI-BLAST that searches a profile against a database of sequences).
SAWTED	Structure Assignment With Text Description	SAWTED compares annotations of the query and closely related sequences with annotations associated with database hits having weak scores. This is included to improve the ability to differentiate between true homologs and spurious matches (i.e., SAWTED enables sequences with weak scores but similar annotations to be given more credence in homology search).
SCOP	Structure Classification of Proteins	The SCOP database classifies protein structural domains into a hierarchy (class, fold, superfamily and family), based on functional and structural data. Proteins in the same SCOP superfamily are expected to be related by divergent evolution from a common ancestral protein and to share a similar (albeit potentially high-level) function. Proteins in different SCOP folds are believed to not be homologous.
SMART	Simple Modular Architecture Research Tool	The SMART Web server allows rapid identification of the domain architecture of a given protein, and includes precomputed results for many sequences

Your sequence is identical to [swissprot|P27037|AVR2_HUMAN](#), displaying precalculated results.

Domains within the query sequence [gi|114722|sp|P27037|AVR2_HUMAN](#) of 513 residues

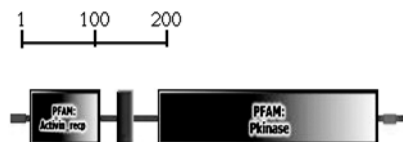


Figure 19.5.4 SMART results for AVR2_HUMAN. Integrated results from several servers show an N-terminal signal peptide, two PFAM domains, a transmembrane domain, and a low-complexity region. This black and white facsimile of the figure is intended only as a placeholder; for full-color version of figure go to <http://www.interscience.wiley.com/c-p/colorfigures.htm>.

4. Examine predicted domains and motifs. The features represented in the example (Fig. 19.5.4) are:
 - a. Signal peptide: the small red box at the N-terminus.
 - b. Two PFAM domains: Activin receptor and Pkinase.
 - c. Low-complexity region: pink box at C-terminus.
 - d. Transmembrane domain: Blue vertical box.
5. Place the mouse cursor over predicted domains and unidentified regions to see their start and end amino acid positions.
6. Click on the domains to get a detailed description of the domain.
7. Scroll down to see the E-value of the predicted matches.
8. Analyze results (see Support Protocol 1).

SUPPORT PROTOCOL 1

GUIDELINES FOR UNDERSTANDING RESULTS OF ANALYSES FROM INTEGRATED META-SERVERS

The domains in SMART (see Basic Protocol 1) are represented by seed alignments (alignments used to build the HMMs). Homologs are gathered by different iterative methods and included in a multiple sequence alignment (MSA) after estimating the statistical significance of sequence similarities and examining experimental biological information, if available. To reduce redundancy in a family of sequences representing a domain, a phylogenetic tree is constructed using the MSA of all the sequences in the family, and, from every branch that is less than a distance of 0.2 on the tree, only a single sequence is chosen. This corresponds to about 80% sequence identity. These sequences are then used in the seed alignment from which the profiles and HMMs representing a domain are constructed (Schulz et al., 1998).

Interpreting the Significance of a Match to a Profile

Most similarity search methods employ E-values (Expect values) to report the statistical significance of a match. How an E-value is calculated varies with each method. Generally, the E-value represents the number of matches expected to be found merely by chance. For instance if an E-value is 10, then 10 matches are expected to be found by chance alone. In general, an E-value <0.001 is a good indicator of the credibility of predicted homology, and the likelihood of a spurious result increases as the E-value approaches 1. A significant E-value implies similarity in structure; the functional specificity may have changed. To confirm functional similarity, it is critical to check for agreement at key functional residues (particularly active site residues in the case of an enzyme), and for a high percent identity (as a conservative rule of thumb, we recommend requiring

a minimum of 40%). If the profile or HMM represents a structural domain, the query should align over a significant fraction of the profile/HMM. As a basic rule of thumb, we recommend requiring a minimum of 75% coverage.

Note that the SMART output does not provide this information (i.e., percent identity, coverage of the profile and agreement with key positions); these criteria will need to be confirmed separately.

Factors Affecting E-Value

E-values are affected by the length of the profile or HMM, with longer profile/HMMs producing more significant E-values for related sequences and shorter profile/HMMs producing weaker E-values. Weak E-values are particularly common when the profile or HMM is based on very short motifs (e.g., the PFAM Leucine Rich Repeat HMMs, having between 20 and 25 residues). Because a near-exact match to a short motif can occur by chance alone, E-values will be correspondingly high (i.e., approaching or exceeding 1) even for closely related sequences. For this reason, many HMM and profile libraries employ empirically determined score cutoffs instead of E-values (see, e.g., the PFAM gathering threshold); these tend to be more relaxed for shorter profile/HMMs than for longer.

Another factor that affects E-values is the size of the database searched. The probability of finding a match by chance increases with the size of the database searched. The upshot of this is that a match between a given query and a hit will have different E-values depending on the database searched. For example, if the E-value of a match between query X and hit Y is 0.01 on searching the National Center for Biotechnology Information (NCBI) nonredundant database (2,182,528 sequences as of November, 2004), then it will be smaller (e.g., 0.001), and hence more significant, on searching Swiss-Prot (159,078 sequences as of November, 2004) using the same search method. The database size becomes an issue especially when E-values are not extremely significant.

IDENTIFYING STRUCTURAL AND FUNCTIONAL DOMAINS USING THE NCBI CD-SEARCH

ALTERNATE PROTOCOL 1

The Conserved Domain Database (CDD; Marchler-Bauer et al., 2003) at NCBI is a database of Position Specific Scoring Matrices (PSSMs) generated from alignments imported from PFAM and SMART and automated alignments from COGs (Cluster of Orthologous Groups; Tatusov, 2001) CD-Search is a tool that uses RPS-BLAST (Reverse PSI-BLAST) to search a sequence against a database of profiles, namely the CDD (Marchler-Bauer et al., 2004). This is the reverse of PSI-BLAST, which searches a profile against a database of sequences.

Materials

See Basic Protocol 1.

1. Point the browser at the CDD (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>).
2. Input a sequence in FASTA format (or, alternatively, the NCBI accession or GenBank accession number) in the box provided (Fig. 19.5.5).
3. Set the Search Database from the pull-down menu of that name. The default option is CDD, but users can also search SMART, PFAM, and COGs.
4. Set other parameters using the “click here for advanced options” link, if desired.
5. Click Submit.

Informatics

19.5.7

Run CD-Search:

Search Database:

Enter Protein query as Accession, Gi, or Sequence in FASTA format:

Read about [FASTA format description](#), click [here](#) for advanced options.

Figure 19.5.5 CD-Search sequence submission form.

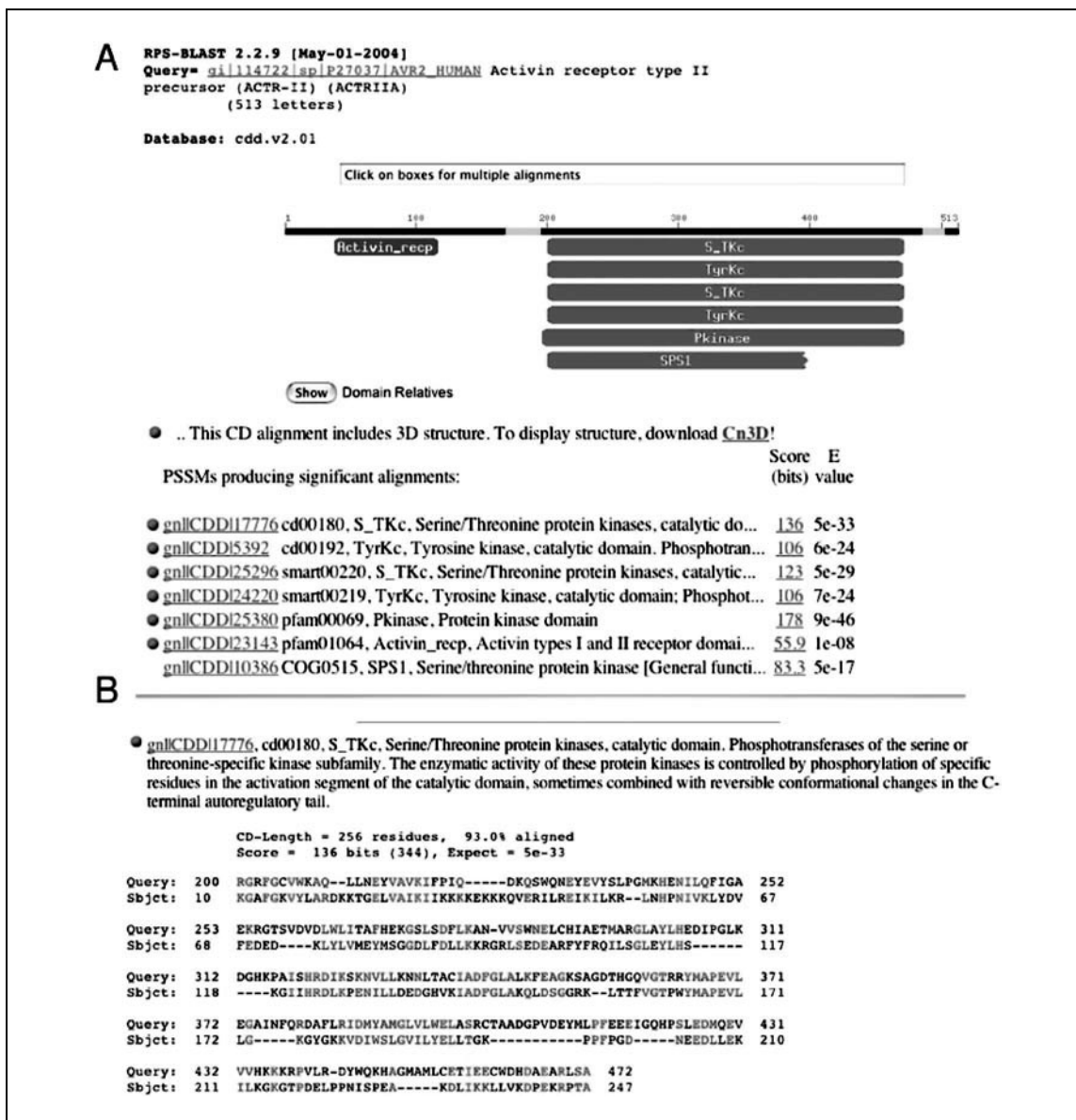


Figure 19.5.6 Results from CD-Search. (A) The output gives a graphic display and E-values of hits. (B) Pairwise alignment of the query with the top hit. This black and white facsimile of the figure is intended only as a placeholder; for full-color version of figure go to <http://www.interscience.wiley.com/c-p/colorfigures.htm>.

6. The CDD results page returns a graphic display of domains in the query, the E-values for each hit, and the alignments (Fig. 19.5.6). Analyze results (see Support Protocol 2).

GUIDELINES FOR UNDERSTANDING RESULTS OF ANALYSES FROM THE NCBI CD-SEARCH

SUPPORT PROTOCOL 2

See Alternate Protocol 1 for steps in carrying out this search.

1. The first few lines give details about the version of RPS-BLAST used, the definition line of the query sequence, and the database searched. Because the CDD has regular updates, we recommend noting these data for record-keeping purposes.
2. In the graphic display, the query sequence is represented as a black bar with a ruler above indicating its length. Low-complexity regions are marked in cyan.
3. The domain matches are shown below in different colors depending on the E-value of the hits. The best hits are marked in red, the second best in blue and so on. The CDD contains multiple PSSMs for the same structural domains; these may have differences in length and coverage of the query, and have correspondingly different E-values. Hits to conserved domains that are related by the Conserved Domain Architecture Retrieval Tool (CDART; Geer et al., 2002; also see Table 19.5.1) resource are indicated by the same color. In the example in Figure 19.5.6, the hits to different PSSMs representing kinase domains are all colored red.
4. A jagged edge seen in the hit display indicates incomplete coverage of the domain (a default minimum of 80% coverage is used by the CD search). For example the SPS1 domain in Figure 19.5.6 contains a jagged right edge, indicating that the alignment to the query does not contain this region of the domain.
5. Below the graphical display is a list of hits along with their E-values. A pink dot before the hit indicates the presence of a structure for the predicted hit. Clicking on the dot will display the structure. Prior installation of Cn3D software is required (<http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml>).
6. Scrolling down the page displays the pairwise alignment between the query and a representative sequence for the PSSM. Identical residues between the query and the hit sequence are colored red, similar residues are colored blue, and masked-out regions are in italics.

Viewing a multiple sequence alignment (MSA) of the query to a conserved domain

7. Clicking on the graphical display of a domain brings up a multiple sequence alignment of the query to selected PSSM training sequences (Fig. 19.5.7). By default, sequences with high similarity to the query protein are shown. This page gives information about the source of the domain alignment, PubMed references, description, and other relevant data. The default MSA viewing option is Hypertext; our guidelines for interpreting results assume this option is used.
8. Interpreting the MSA: aligned positions are in uppercase, unaligned positions are in lowercase, and gaps are indicated by dashes (-). Residues labeled “unaligned” are inserts between positions in the predicted consensus structure. Profiles and hidden Markov models (HMMs) are constructed to represent the conserved structural elements of a protein family or domain. Amino acids corresponding to additional structure (e.g., extended loop regions) will be displayed in lowercase. These lowercase letters are indicative of insertions relative to the consensus structure. Missing parts of the consensus structure are indicated with dash characters (--).

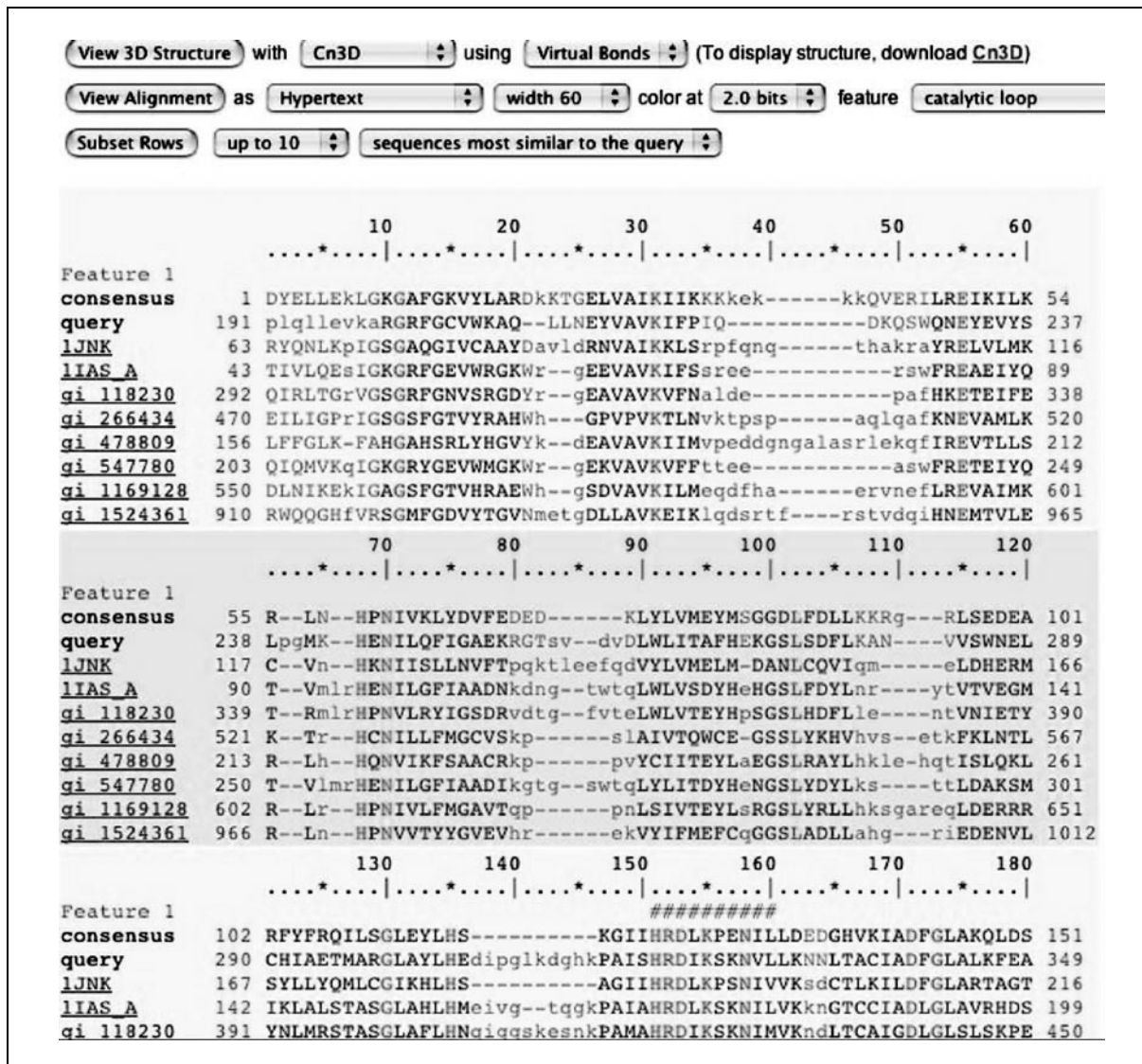


Figure 19.5.7 Multiple sequence alignment of AVR2_HUMAN kinase domain with a conserved kinase domain hit. The columns marked with # have been identified as central for kinase function.

9. For domains with a solved structure, one can view the structure using the View 3D Structure option. The coloring scheme of the alignment can be changed to indicate different levels of conservation. For instance using the option Identity from the “color bits” pull-down menu shows identical residues in the alignment in red, aligned residues in blue, and unaligned residues and columns containing many gaps in gray.
10. Features defining a family are highlighted with # above the corresponding columns in the MSA. The evidence for the feature is presented at the bottom of the MSA. This is useful to predict if the query sequence also contains residues important for the family. In our example, clicking on the S_TKc shows a catalytic loop as feature 1 and the 3-D structure as evidence.

Interpreting the significance of results

11. If E-values are weak, it may be helpful to examine the PSSM hits to see if all refer to the same structural or functional class, or if competing hypotheses for a region are presented by the CDD analysis. Checking for significant percent identity and coverage (as described in the SMART discussion earlier in this unit) is important to avoid spurious and misleading predictions.

12. The Advanced Search mode allows users to change the maximum allowed E-value cutoff. This is generally done if no matches are possible with more restrictive cutoffs. If a permissive threshold (e.g., E-value 1) is used, false positive matches should be expected.
13. Though the domains represented in CD are similar to those in PFAM and SMART, the search and scoring mechanisms are very different. While PFAM uses profile HMM scoring, CD-Search employs RPS-BLAST. Hence, the E-values that are reported from PFAM and CD-Search for the same PFAM domain may not be identical.

To interpret the significance of E-values and the effect of profile/HMM lengths on E-values, see the guidelines described in Support Protocol 1.

PREDICTING STRUCTURAL DOMAINS AND SECONDARY STRUCTURE USING 3D-PSSM

ALTERNATE PROTOCOL 2

3D-PSSM is a Web server designed explicitly and exclusively for protein structure prediction, which incorporates a variety of advanced remote homolog detection and threading methods (Kelley et al., 2000). We present the 3D-PSSM as a supportive procedure in cases where other servers are unable to provide clues to molecular function or structure for regions of the protein. In these cases, the advanced methods incorporated in the 3D-PSSM workflow can often give insights not available via other servers for identifying homologous structures. In contrast to PFAM, the NCBI CDD, and many other servers that model both structural domains and domains having no known structure, any match to a 3D-PSSM domain implies a solved structure for that region. One of the great benefits of 3D-PSSM is the automatic construction of a homology model for any regions in the query matching one or more of the 3D-PSSM models.

The 3D-PSSM server is continually updated. As of February, 2005, 3D-PSSM contains a library of 9864 PSSMs representing structural domains.

Materials

See Basic Protocol 1.

1. Point the browser at <http://www.sbg.bio.ic.ac.uk/servers/3dpssm/>.
2. Click on Recognise a Fold (Fig. 19.5.8A) for the sequence submission form.
3. Enter a valid e-mail address.
4. Paste the query sequence in FASTA format. The query sequence must be <800 residues. The maximum number of residues allowed per submission by 3D-PSSM is restricted to 800. If the protein of interest has >800 residues, we suggest that it be submitted in two parts with overlap. For example, for a protein of 1000 residues, submit region 1 to 600 first, and 400 to 1000 next.
5. Click Submit. Results of the run will be returned by e-mail. The e-mail also contains a hyperlink to the Web version of the results. Results are stored only for 5 days in the server (we recommend downloading results from the Web site).
6. Click on the hyperlink in the e-mail to open a Web page with the results (Fig. 19.5.8B). Analyze results (see Support Protocol 3).

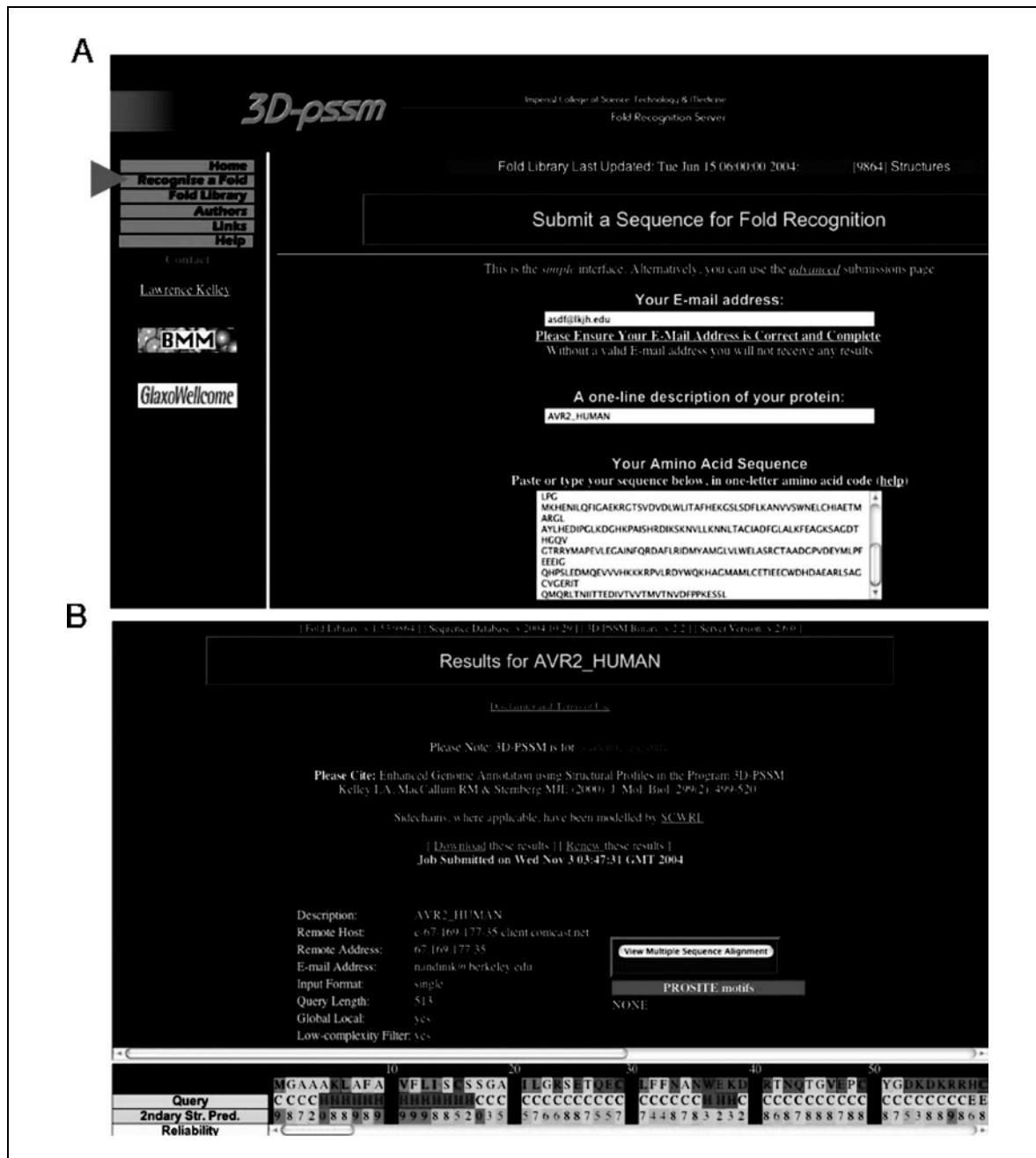


Figure 19.5.8 Submission form and results page from 3D-PSSM server. **(A)** The submission form; **(B)** the results page.

**SUPPORT
PROTOCOL 3**

**GUIDELINES FOR UNDERSTANDING RESULTS OF ANALYSES FROM
THE 3D-PSSM SERVER**

We recommend 3D-PSSM (see Alternate Protocol 2) as an alternative for domain identification for those cases where regions of a protein have no significant matches to domains in the CDD or SMART databases. In these cases, the advanced methods used by 3D-PSSM may be able to identify a weak but potentially significant match to a protein of solved structure and shed light on the functional role of the protein. Caution must be employed when examining the 3D-PSSM results where the E-value is only marginally significant (e.g., >0.001), as the potential for a false positive match increases as the E-value approaches 1.

**Basic Protein
Sequence Analysis**

19.5.12

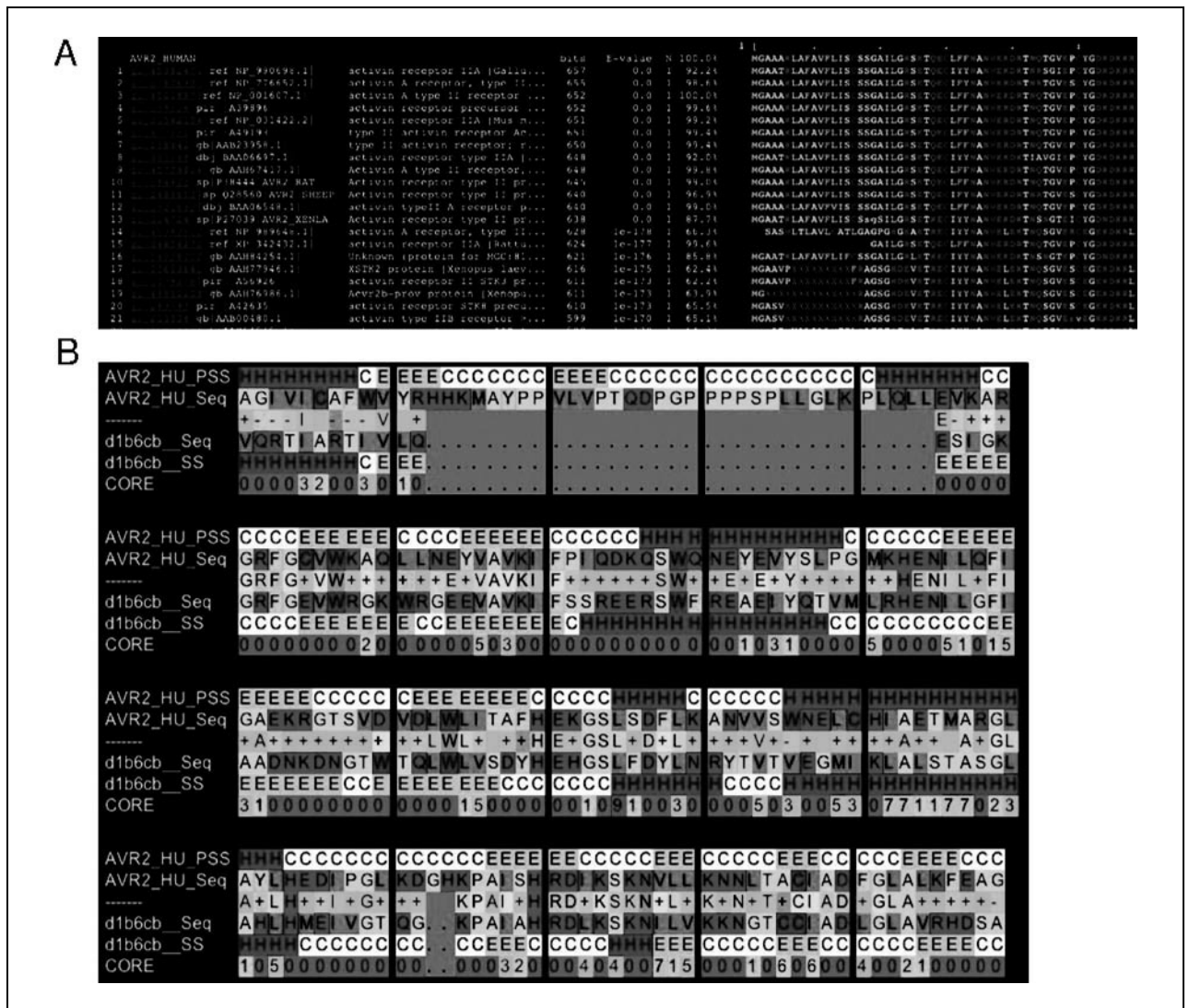


Figure 19.5.9 Sequence alignments produced by 3D-PSSM (A) Alignment of query to its homologs. (B) Pairwise alignment and comparison of secondary structure and solvent accessibility predictions for the query with that of the hit (structural domain). This black and white facsimile of the figure is intended only as a placeholder; for full-color version of figure go to http://www.interscience.wiley.com/c_p/colorfigures.htm.

The 3D-PSSM Web server produces a great deal of interesting and informative results, some of which are beyond the scope of this unit. In this protocol, we highlight a selection of the most interesting results. We encourage users to explore further.

1. The first page of the results is separated into two frames. The upper frame contains a summary of results. The lower frame shows the secondary structure of the query sequence predicted by PSIPRED (Jones, 1999; also see Table 19.5.1). We focus here on results in the upper panel.
2. Click on View Multiple Sequence Alignment to examine the alignment of the query and predicted sequence homologs (Fig. 19.5.9A). The 3D-PSSM Web server uses this alignment to construct a PSSM and score proteins of solved structure, as part of a bidirectional search for structural homologies. Each sequence in the alignment has a hyperlink to the corresponding GenBank file at NCBI. The scores and E-value of each hit, and percent ID to query are also available.

- Return to the original results page and scroll down to view matches to structural domains. The top 20 structural hits are shown in a tabular format. The table contains the following information.

- Alignment.

A link to the sequence alignment of the query to the hit (structural domain) and comparison of the secondary structure and solvent accessibility of residues between query and hit (Fig. 19.5.9B).

- Fold Library.

The ID of the structural hit (d[`pdbservice`] [`chain`] [`region`]), along with the pairwise ID between query and hit, are shown. Clicking on the ID brings up a page of the fold library containing information about the hit structural domain, including an MSA of the hit with its sequence and structural homologs.

- Models.

A link to the predicted 3-D model of the query protein based on the alignment to the hit which can be viewed by the protein visualization software RasMol (see Table 19.5.1) or Chime.

- E-value.

The E-value of the hit is shown, colored according to level of confidence, red indicating very high confidence.

- Other information.

The other information in the table includes SAWTED (Table 19.5.1) E-value for text matching between query and definitions of hits, SCOP classification of hit, and other data.

If the E-value to a structural hit is >.001, the following analyses are recommended to assess the credibility of the predicted homology

- Check the sequence alignment. The query and structure should align with few insertions or gap characters, and have significant pairwise identity. As a rule of thumb, a pairwise identity of 25% or more over 80 amino acids can be used to infer homology (assuming very few gaps). As the percent identity drops below this level, the potential for a spurious match by chance alone increases. In these cases, it is critical to confirm that the alignment shows agreement at known critical residues.
- Compare the agreement between predicted and known secondary structure. Compare the predicted secondary structure of the query with that of hit. The secondary structures are indicated as follows: H for Helix; E for Strand; C for Coil. Agreement between helical and sheet structure is generally more important than in the more variable coil regions.
- Compare the agreement in the conserved core structure. The hydrophobic core region of the structural domain hit is marked in the MSA page as Core (Fig. 19.5.9B), numerically labeled to indicate how buried the residues are (9 for very buried and making many contacts and 0 for exposed and making few contacts) and color coded to indicate important core residues (red for important core residues). Significant gaps or inserts in the conserved core elements of the structure should be considered evidence against the predicted homology.

An important issue to keep in mind is that 3D-PSSM may only identify one structural domain in a multidomain protein. To enable identification of domains in the remaining region(s) of the query, separate submissions of unknown regions are necessary.

There are a few other Web servers and databases enabling integrated searches, as shown in Table 19.5.2.

Table 19.5.2 Web Servers Providing Domain Identification Tools

Web server	Description	URL
SMART	Structural, functional, and localization prediction	http://smart.embl-heidelberg.de/
PFAM	Includes both structural and functional domains	http://www.sanger.ac.uk/Software/Pfam/ http://pfam.wustl.edu
SUPERFAMILY	Based on the Astral PDB90 database of structural domains	http://supfam.mrc-lmb.cam.ac.uk/SUPERFAMILY/hmm.html
SAM-T02	The UCSC structure prediction HMM library	http://www.cse.ucsc.edu/research/compbio/HMM-apps/T02-query.html
TigrFAMs	From TIGR, a Web server of HMMs for structure (and function) prediction	http://tigrblast.tigr.org/web-hmm/
PhyloFacts	Includes subfamily HMMs for structural domains.	http://phylogenomics.berkeley.edu/phylofacts/
NCBI Conserved Domain	Includes protein domains with solved structures	http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
3D-PSSM	Includes a secondary structure prediction to obtain predictions of 3-D fold	http://www.sbg.bio.ic.ac.uk/~3dpssm/

PREDICTING HELICAL TRANSMEMBRANE REGIONS AND SUBCELLULAR LOCALIZATION

The molecular function of a protein is context-dependent; the subcellular localization of a protein plays a large role in determining its partners and pathway interactions. The cellular sorting machinery makes use of sequence signals targeting a protein for secretion, targeting it to different organelles, or targeting it for insertion into the membrane. Bioinformatics methods take advantage of these motifs and patterns as well as other sequence attributes to predict where a protein is localized. We review these in this protocol.

Our primary recommended resource for this protocol makes use of a variety of software tools developed by the Center for Biological Sequence Analysis at the Technical University of Denmark (<http://www.cbs.dtu.dk/services/>). This resource enables the detection of helical transmembrane domains, signal peptides, and other organellar targeting signals. Due to the biological importance of transmembrane proteins, we discuss the detection of these domains first.

Prediction of helical transmembrane regions in protein of interest

Transmembrane predictions involve not only the location of probable membrane helices in a protein, but also the full topology of the protein with respect to the membrane (i.e., which regions of the protein span the membrane, which regions are cytoplasmic, and which are extracellular). Among the several Web servers available for transmembrane prediction, we describe the use of the TMHMM server. The TMHMM server uses sophisticated hidden Markov models methodologies to predict transmembrane regions in proteins, incorporating information on hydrophobicity, helix length, charge bias, and topological constraints into its model (Krogh et al., 2001). It has a user-friendly Web interface, is fast in its searches, and allows batch mode submissions of up to 4000 sequences.

BASIC PROTOCOL 2

Informatics

19.5.15

Materials

Input data

Protein sequence file, in either raw or FASTA format (Fig. 19.5.1), or, if more than one sequence is used, a file containing all the input sequences in FASTA format can be used

1. Point the browser at the TMHMM server (<http://www.cbs.dtu.dk/services/TMHMM/>).
2. Paste in the sequence(s), or upload a file directly from the local computer (recommended when one wants to analyze large numbers of sequences). To upload a file, use the Browse button below the text “Submission of a local file in FASTA format.” Each file can contain at most 4000 sequences.
3. Set display options. If submitting a single sequence, we recommend setting the display as “Extensive, with graphics.” This will provide results in the most information-rich format (see below). The recommended option for multiple sequence search is “One line per protein.” This returns the results in a tab-delimited format that can be saved as a text file to be opened in a program like Excel.
4. Click Submit. After a short time the results will be displayed.
5. Only 10 job (including batch mode) submissions per day are allowed per user.
6. Analyze results (see Support Protocol 4).

SUPPORT PROTOCOL 4

GUIDELINES FOR UNDERSTANDING RESULTS OF PREDICTIONS OF HELICAL TRANSMEMBRANE REGIONS AND SUBCELLULAR LOCALIZATION

For steps in predicting helical transmembrane regions and subcellular localization, see Basic Protocol 2.

1. Depending on the selection during submission, the results could be either in “long” or “short” format. The long format is shown in Figure 19.5.10.
2. View summary of results containing the following information:
 - a. Length.
Length of query protein.
 - b. Number of predicted TMHs.
Number of predicted TM helices in the query.
 - c. Exp number of AAs in TMHs.
Number of amino acids predicted to be in TM helix. If this number is greater than 18, then it is predicted to be a TM.
 - d. Exp number, first 60 AAs.
Number of amino acids expected to be in TM helix in the first 60 residues. A warning POSSIBLE N-term signal sequence is generated if this number exceeds 10.
 - e. Total prob of N-in.
This represents the probability of N-terminal region being on the cytoplasmic side.
3. View plot. The plot summarizes the various signals used by the TM server. The *x* axis corresponds to the amino acid positions while the *y* axis corresponds to the likelihood of an individual position being predicted to be in a TM. Note that the *y* axis continues beyond a value of 1; the region above 1 is used to display the topology. Since the

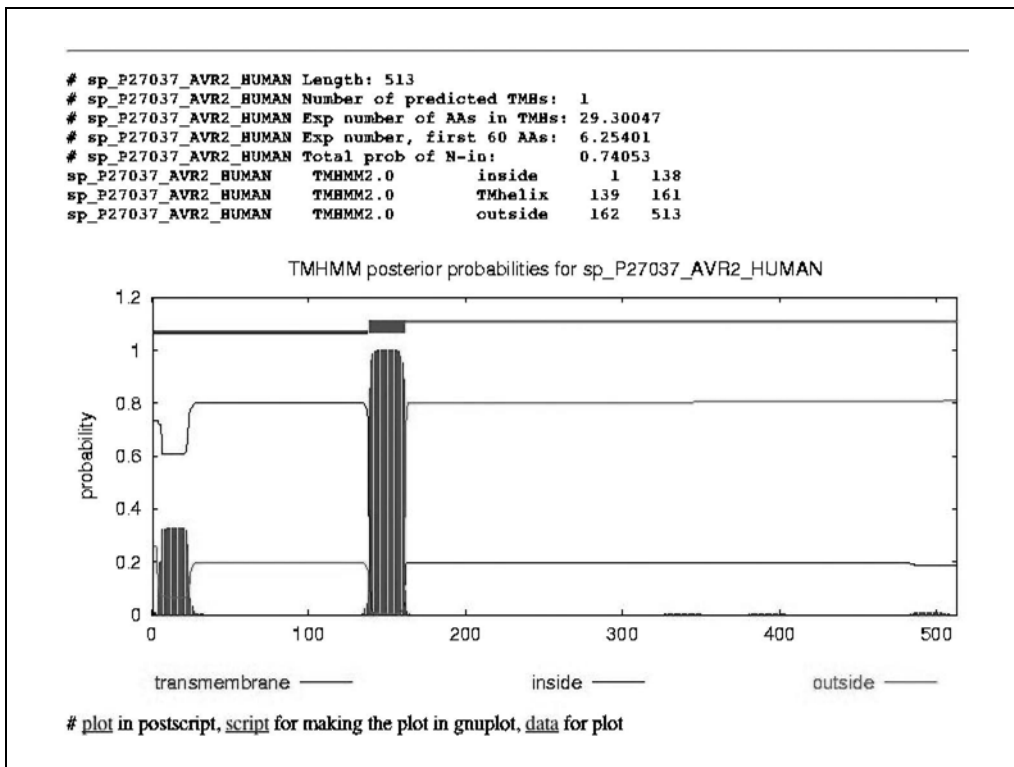


Figure 19.5.10 TMHMM output for AVR2_HUMAN. Shown are results in the “extensive with graphics format.” A summary of results is followed by a graphic display. The x axis on the graph represents amino acid positions in the query sequence and the y axis represents the probability of a residue to be in a TM. The peaks indicate positions with higher probability of being a TM domain. The predictions of TM and topology are indicated above a y axis value of 1. The region predicted to be inside (cytoplasmic) is represented with a blue line (marked as INSIDE in this figure) and the region on the outside (extracellular) is represented in pink (labeled OUTSIDE in this figure). This black and white facsimile of the figure is intended only as a placeholder; for full-color version of figure go to http://www.interscience.wiley.com/c_p/colorfigures.htm.

TMHMM server integrates a variety of signals from the protein, occasionally the predicted number of helices is lower than the number of peaks in the plot.

4. View raw data. Below the plot is a hyperlink to the raw data for the plot.
5. Cautions

- a. Accuracy of topology predictions

The accuracy of topology prediction (i.e., which parts of the proteins are extracellular and which are cytoplasmic) is known to be one of the most challenging aspects of transmembrane predictions. Even when the location of transmembrane domains is predicted correctly, the topology can be reversed.

- b. Potential for false positives and false negatives.

All TM prediction servers are known to have problems with false positive and false negative predictions. False positives arise often for hydrophobic stretches and signal peptides. False negatives can arise for any number of reasons, and are particularly common in noncanonical TM stretches containing some polar or charged residues.

- c. Incorporating prediction of globular structure.

It can be helpful to integrate domain detection, particularly of globular domains, when interpreting TM predictions. If a predicted TM occurs in a region that has a significant match to a protein of solved structure, we recommend examining what is known about that structure. Some structures include transmembrane domains, but

these are few. The vast majority of solved structures represent globular (soluble) domains that do not include transmembrane domains. If a predicted TM domain has a significant match to a globular protein, and that region is not at the extreme N- or C-terminus of the domain, we would discount the likelihood of the TM domain.

**ALTERNATE
PROTOCOL 3**

**PREDICTING THE SUBCELLULAR LOCALIZATION OF A PROTEIN
USING TargetP**

Cells have evolved a complex machinery for enabling proteins to be targeted to specific sub-cellular locations. Targeted proteins have specific signals recognized by the machinery. For secreted proteins and those targeted to the organelles such as mitochondria or chloroplast, the targeting signals are amino-terminal peptides recognized by the translocation machinery (Schatz and Dobberstein, 1996).

The targeting signals for different locations have specific features. For instance, the amino-terminal signal peptide (~20 amino acids) that targets proteins for secretion starts with a positively charged segment followed by a central hydrophobic region and a polar segment containing the cleavage site recognized by the signal peptidase. Different approaches have been developed to predict protein localization (Emanuelsson and von Heijne, 2001). Most methods specialize in predicting one specific location (e.g., SignalP for predicting signal peptides). In this protocol we review TargetP, an integrated approach that distinguishes proteins targeted to different locations in the cell such as mitochondria, chloroplast, and the secretory pathway.

Materials

See Basic Protocol 2.

1. Point the browser at the TargetP server (<http://www.cbs.dtu.dk/services/TargetP/>).
2. Paste in the sequence(s), or upload a file directly from the local computer.
3. Identify the origin of the protein sequence (i.e., plant or animal). If it is of plant origin, TargetP will search for chloroplast localization signals.
4. Select the option to perform a cleavage site prediction.
5. Use default settings for specificity cutoffs. Click the “Submit sequence/file” button.
6. Analyze results (see Support Protocol 5).

**SUPPORT
PROTOCOL 5**

**GUIDELINES FOR UNDERSTANDING RESULTS PREDICTING THE
SUBCELLULAR LOCALIZATION OF A PROTEIN USING TargetP**

Results (see Alternate Protocol 3) include scores for different predicted cellular locations (Fig. 19.5.11).

SP : Signal peptide
mTP : Mitochondrial target peptide
cTP : Chloroplast (if plant sequence is used)
Other: For other locations

Along with the scores, a “reliability class” (RC) is also calculated. This gives information on the difference between the highest score and the second best.

RC 1: $\text{diff} > 0.800$
RC 2: $0.800 > \text{diff} > 0.600$
RC 3: $0.600 > \text{diff} > 0.400$
RC 4: $0.400 > \text{diff} > 0.200$
RC 5: $0.200 > \text{diff}$

```

## ## ## T A R G E T P 1.0 prediction results ## ## ##
Number of input sequences: 1
Cleavage site predictions not included.
Using NON-PLANT networks.

-----
Name      Length  mTP  SP  other  Loc.  RC
-----
sp_P27037_AVR2_HUMAN  513  0.053  0.893  0.054  S     1
-----
cutoff                0.00  0.00  0.00

```

Figure 19.5.11 TargetP prediction of the subcellular localization of AVR2_HUMAN.

Table 19.5.3 Cellular Localization Prediction

Web server	Description	URL
TMHMM	The TMHMM transmembrane helix prediction server	http://www.cbs.dtu.dk/services/TMHMM/
DAS	The DAS transmembrane helix prediction server	http://www.sbc.su.se/~miklos/DAS/
SignalP	The SignalP signal peptide prediction server	http://www.cbs.dtu.dk/services/SignalP/
ChloroP	The ChloroP chloroplast targeting prediction server	http://www.cbs.dtu.dk/services/ChloroP/
PredictNLS	Nuclear localization signal prediction	http://cubic.bioc.columbia.edu/predictNLS/

In interpreting TargetP results, it is important to consider both the predicted location as well as the RC score. Lower RC values imply greater expected accuracy; an RC of 1 or 2 is a good indicator of reliability. In our example, the score for SP is 0.853 and the RC value is 1, both supporting the prediction of presence of a signal peptide in the protein (Fig. 19.5.11). TargetP is known to have greater specificity in prediction compared to other methods in this class, but this comes at a cost of sensitivity. A null prediction does not imply that the protein is cytoplasmic. By contrast, cleavage-site prediction is less precise. Finally it is critical to indicate whether the input sequence is of plant or animal origin so that the appropriate localization tests can be performed; it is otherwise difficult to distinguish between chloroplast and mitochondrial signals.

Other resources for prediction of transmembrane helices or subcellular localization can be found in Table 19.5.3.

PREDICTING KEY FUNCTIONAL RESIDUES AND MOTIFS USING THE PROSITE WEB SERVER

Not all residues in a protein are created equal. Residues in active sites, e.g., the catalytic triad of serine proteases, are critical for enzymatic activity. Other positions may be involved in substrate recognition and binding. Still other positions may be conserved for structural reasons. Over the years, biologists have identified a large number of key residues and conserved motifs that are useful in defining and recognizing protein families.

Our primary recommended resource for this protocol is PROSITE. The PROSITE Web server enables a biologist to determine whether any experimentally characterized motifs are present in a sequence of interest (Sigrist et al., 2002; Hulo et al., 2004). This resource enables the detection of different protein motifs, post-translational modification sites, and domains. Since domain detection is provided through other servers (presented earlier in this unit), this protocol focuses on the use of the PROSITE Web server for detecting the presence of conserved motifs and functional residues.

**BASIC
PROTOCOL 3**

Informatics

19.5.19

Materials

Input data

Sequence in raw or FASTA format (Fig. 15.9.1)

1. Point the browser to the PROSITE Web site (<http://us.expasy.org/prosite/>).
2. Input the sequence in FASTA format in the sequence box provided.
3. Choose the option to exclude patterns with high probability of occurrence. This will decrease the likelihood of spurious predictions.
4. Click the Quick Scan button to search for conserved motifs in the PROSITE database.
5. Analyze results (see Support Protocol 6).

SUPPORT PROTOCOL 6

GUIDELINES FOR UNDERSTANDING RESULTS OF SEARCHES DONE USING THE PROSITE WEB SERVER

The PROSITE Web server (see Basic Protocol 3) includes both pattern and profile search. Profiles are usually constructed to represent domains and protein families, while patterns are confined to the representation of short conserved motifs and critical residues. When a query protein is submitted, both patterns and profiles are searched and the results are returned as “hits by profile” (Fig. 19.5.12A) and “hits by patterns” (Fig. 19.5.12B). For each detected feature, the position and the sequence motif are provided along with a link to detailed information about each feature. Placing the mouse cursor over predicted domains in the graphical display highlights the corresponding region in the sequence.

Pattern detection has certain advantages relative to profile search. First, pattern detection is extremely fast. Second, and potentially of greater interest to biologists, the results are easily interpreted; the motifs are typically quite short, and critical residues defining the motif can be examined easily.

There are two main disadvantages to pattern detection, relative to profile search. First, a strict agreement with a pattern is required, resulting in lower sensitivity than the more permissive profile searches, which allow substitutions. Second, the results from pattern searches are not given scores; the pattern is either present or it is not. This makes it difficult to evaluate the significance of the match. In addition, some patterns have a high associated false-positive rate (this is more common in very short or highly variable patterns). For example, the post-translational modification site for N-glycosylation is Asn-Xaa-Ser/Thr (where Xaa represents any amino acid). However, the presence of this motif may not be sufficient to ensure glycosylation, as the process is also dependent on the structure of the protein (Sigrist et al., 2002).

The PROSITE Web server contains information to help reduce the likelihood of errors. First, each pattern includes information regarding the number of hits obtained while scanning the Swiss-Prot database with that pattern, the number of false positives and false negatives, and results from pattern search using database randomization. These data should be consulted to obtain an understanding of the likely significance of a match to this pattern. In addition, the PROSITE profile searches can provide additional support for a functional classification based on motif and pattern detection. For instance, the presence of a serine/threonine protein kinase signature pattern along with detection of kinase domain for AVR2_HUMAN strengthens the significance of the prediction.

An example of a six-amino-acid PROSITE pattern is shown below.

G-{EDRKHPFYW}-x(2)-[STAGCN]-{P} [G is the N-myristoylation site]

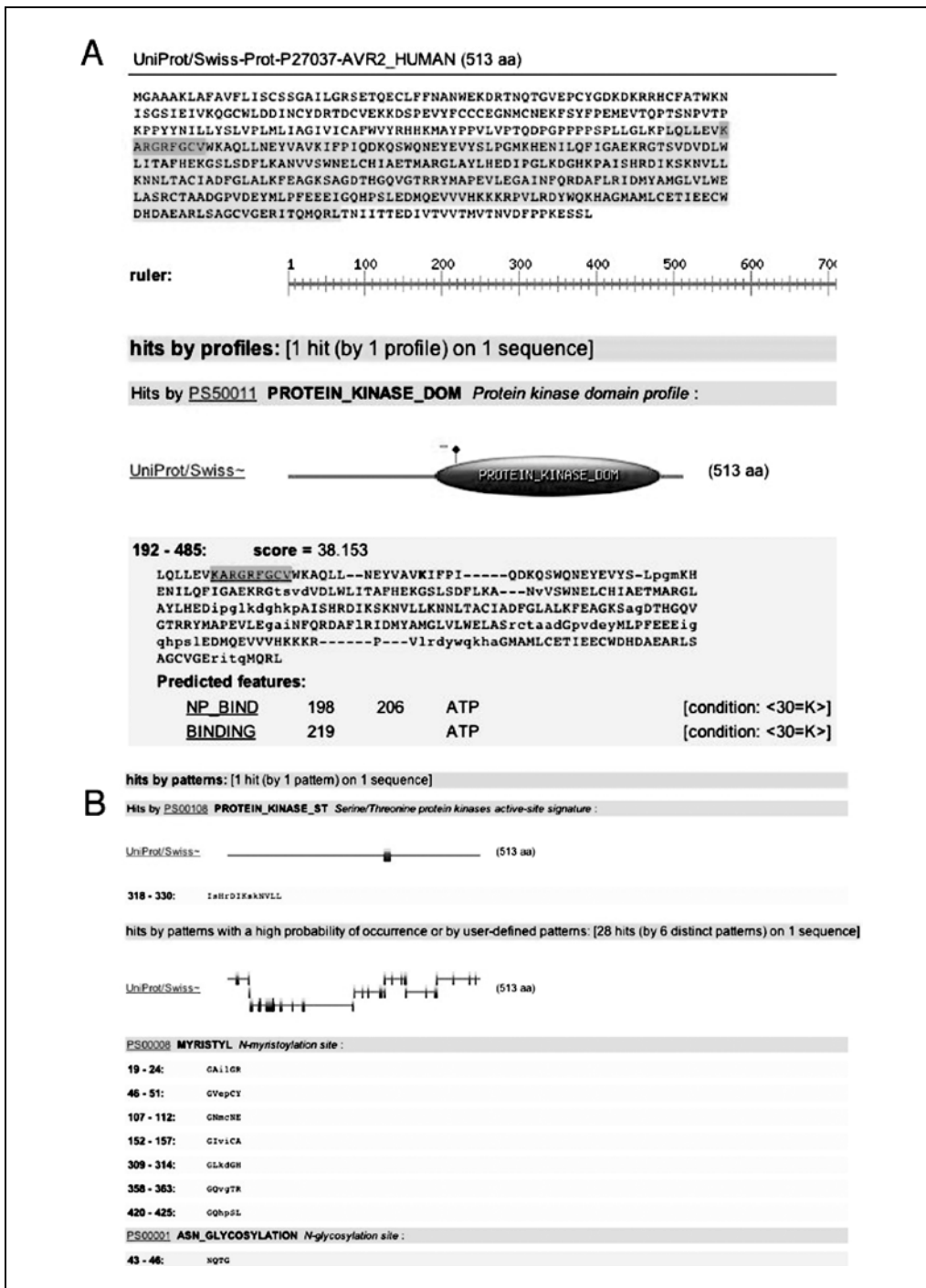


Figure 19.5.12 Prediction of patterns by PROSITE for AVR2.HUMAN. (A) results of profile search; (B) results of pattern search.

Residues in straight brackets ([]) designate residues that are favorable at a position. Residues in curly brackets ({}) designate those amino acids that are *not* allowed at a position (i.e., in this example, the amino acids EDRKHPFYW are not allowed at position 2). The sequence x (2) indicates any two amino acids.

Other resources for similar tasks can be found in Table 19.5.4.

Table 19.5.4 Predicting Key Functional Residues and Motifs

Web server	Description	URL
Evolutionary Trace	University of Cambridge implementation of Evolutionary Trace algorithm	http://www-cryst.bioc.cam.ac.uk/~jiye/evoltrace/evoltrace.html
MEME	The MEME motif discovery software	http://meme.sdsc.edu/meme/Website/intro.html
PROSITE	Detects previously characterized motifs characterizing protein families, as well as post-translational modifications	http://www.expasy.ch/prosite

SUPPORT PROTOCOL 7

HOMOLOG IDENTIFICATION

As a support protocol, we encourage biologists to gather homologs for their sequence of interest and to construct and analyze a multiple sequence alignment of these homologs. The BLAST family of methods (Altschul et al., 1990) is by far the most popular means for identifying homologs in database search; we refer readers to *UNIT 19.3* in this manual or the BLAST tutorial at the NCBI Website (<http://www.ncbi.nlm.nih.gov/Education/BLASTinfo/tut1.html>).

Homolog identification is useful for a variety of reasons. First, homologs can give clues to the overall molecular function of a protein. When inferring molecular function by homology, it is critical to confirm that the query and hit have the same overall fold (i.e., are globally alignable), that they are orthologous, and that the database annotation of the hit is based on experiment and not homology (reviewed in Eisen, 1998, and in Sjölander, 2004).

Gathering homologs is also the first step in construction of a multiple sequence alignment and the (somewhat esoteric) art of alignment gazing. Alignment gazing can help a biologist discriminate between true homologs and spurious inclusions, and assists in the identification of key functional residues. Sequences inserting large numbers of amino acids, that do not align over long regions, or that do not agree at conserved positions (especially longer conserved motifs) should be removed from the alignment. The final set of high-confidence homologs can be used to supplement the structure-prediction efforts (i.e., by parallel analyses using all the servers and approaches described in this unit). In addition, the multiple sequence alignment can be used as input into motif-discovery systems such as MEME. This is particularly helpful when there is limited (or no) success in analysis of the query sequence. Consensus approaches such as these have been shown to dramatically improve prediction accuracy.

COMMENTARY

Background Information

Each of the methods presented in this unit has the potential to produce errors of different types. When examining results, it is important to keep in mind the types of errors that are possible with each method, and to attempt to separate those predictions that are highly credible from those that are potentially misleading.

Homology-based function prediction is particularly prone to systematic error; domain shuffling, gene duplication, and speciation produce families of proteins sharing some regions of similarity but with different functions and

overall folds. Because of these issues, even a significant E-value between a query and a database hit is not sufficient for inference of molecular function. To avoid these errors, phylogenomic inference of a protein's molecular function in the context of its related family members is critical. A detailed description of this approach is beyond the scope of this unit; overviews of the issues and methodologies for phylogenomic analysis are presented in Eisen (1998) and Sjölander (2004).

For these reasons, this unit focuses less on obtaining a prediction of precise molecular

function and more on a structural annotation of a protein. The protocols presented here have been selected to enable a biologist to label regions in the protein as functional or structural domains, to identify key residues and motifs, and to predict a protein's cellular localization.

Most homolog-detection methods use either profile- or HMM-based strategies. Both profiles and HMMs are generalizations of multiple sequence alignment of homologs. They differ in how they estimate gap parameters; profiles have a fixed cost for insertions and deletions, while HMMs learn them from the training sequences used in the MSA, thus varying the cost for different positions. Both methods are dependent on precision of homolog gathering, alignment accuracy, and estimation of amino acid substitutions.

Domain-based approaches to protein annotation have a critical limitation. The profiles (and HMMs and PSSMs) used to detect the presence of these domains are typically optimized for remote homolog detection. This makes them very effective at providing clues to a protein's fold or function that might otherwise not be achievable. However, because some protein folds are able to provide a multiplicity of distinct functions, a significant score to a profile for a domain may mean only that the query has a similar *fold*, but may not actually imply a common molecular function. We strongly advise checking for agreement at key functional residues (where this information is available). The use of consensus approaches to predicting a protein's structure can help avoid errors, provided that the individual methods are orthogonal. For instance, since both SMART and the CDD include PFAM profile HMMs in their structure-prediction process, a consensus structure prediction between the NCBI CDD or SMART and PFAM is much less informative than a consensus of the NCBI CDD and the 3D-PSSM structure prediction Web server (which uses an entirely different approach).

Methods for predicting helical transmembrane domains often rely on weak nonspecific signals such as hydrophobicity, and hence have a fairly high frequency of false-positive as well as false-negative predictions (Chen et al., 2002). To boost accuracy in this task, we strongly encourage biologists to examine alternative hypotheses. Does the predicted transmembrane domain occur in a region predicted (with a significant E-value) to be a globular structure? Is it found at the amino-terminus (where it is more likely to be a signal peptide)? In either case, the probability of the

region being a transmembrane domain should be discounted.

In summary, our recommendation for enhancing prediction accuracy involves integrating results and information from a variety of different resources, and biological data should be incorporated wherever possible.

Critical Parameters and Troubleshooting

On occasion, a Web server will report an error for certain input sequences. In some cases, hidden characters may make identification of the problem difficult. If this happens, try truncating the definition line to the name of the protein, inserting a carriage return after the definition line, and then resubmitting the sequence.

Literature Cited

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Bateman, A., Birney, E., Cerruti, L., Durbin, R., Eddy, S.R., Griffiths-Jones, S., Howe, K.L., Marshall, M., and Sonnhammer, E.L. 2002. The Pfam protein families database. *Nucleic Acids Res.* 30:276-280.
- Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.C., Estreicher, A., Gasteiger, E., Martin, M.J., Michoud, K., O'Donovan, C., Phan, I., Pilbout, S., and Schneider, M. 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* 31:365-370.
- Chen, C.P., Kernytsky, A., and Rost, B. 2002. Transmembrane helix predictions revisited. *Protein Sci.* 11:2774-2791.
- Eisen, J.A. 1998. Phylogenomics: Improving functional predictions for uncharacterized genes by evolutionary analysis. *Genome Res.* 8:163-167.
- Emanuelsson, O. and von Heijne, G. 2001. Prediction of organellar targeting signals. *Biochim. Biophys. Acta* 1541:114-119.
- Geer, L.Y., Domrachev, M., Lipman, D.J., and Bryant, S.H. 2002. CDART: Protein homology by domain architecture. *Genome Res.* 12:1619-1623.
- Hubbard, T., Barker, D., Birney, E., Cameron, G., Chen, Y., Clark, L., Cox, T., Cuff, J., Curwen, V., Down, T., Durbin, R., Eyraas, E., Gilbert, J., Hammond, M., Huminiecki, L., Kasprzyk, A., Lehvaslaiho, H., Lijnzaad, P., Melsopp, C., Mongin, E., Pettett, R., Pocock, M., Potter, S., Rust, A., Schmidt, E., Searle, S., Slater, G., Smith, J., Spooner, W., Stabenau, A., Stalker, J., Stupka, E., Ureta-Vidal, A., Vastrik, I., and Clamp, M. 2002. The Ensembl genome database project. *Nucleic Acids Res.* 30:38-41.
- Hulo, N., Sigrist, C.J., Le Saux, V., Langendijk-Genevaux, P.S., Bordoli, L., Gattiker, A., De Castro, E., Bucher, P., and Bairoch, A. 2004.

- Recent improvements to the PROSITE database. *Nucleic Acids Res.* 32:D134-D137.
- Jones, D.T. 1999. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 292:195-202.
- Kelley, L.A., MacCallum, R.M., and Sternberg, M.J. 2000. Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.* 299:499-520.
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. 2001. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* 305:567-580.
- Letunic, I., Copley, R.R., Schmidt, S., Ciccarelli, F.D., Doerks, T., Schultz, J., Ponting, C.P., and Bork, P. 2004. SMART 4.0: Towards genomic data integration. *Nucleic Acids Res.* 32:D142-D144.
- Marchler-Bauer, A., Anderson, J.B., DeWeese-Scott, C., Fedorova, N.D., Geer, L.Y., He, S., Hurwitz, D.I., Jackson, J.D., Jacobs, A.R., Lanczycki, C.J., Liebert, C.A., Liu, C., Madej, T., Marchler, G.H., Mazumder, R., Nikolskaya, A.N., Panchenko, A.R., Rao, B.S., Shoemaker, B.A., Simonyan, V., Song, J.S., Thiessen, P.A., Vasudevan, S., Wang, Y., Yamashita, R.A., Yin, J.J., and Bryant, S.H. 2003. CDD: A curated Entrez database of conserved domain alignments. *Nucleic Acids Res.* 31:383-387.
- Marchler-Bauer, A. and Bryant, S.H. 2004. CD-Search: Protein domain annotations on the fly. *Nucleic Acids Res.* 32:W327-W331.
- Murzin, A.G., Brenner, S.E., Hubbard, T., and Chothia, C. 1995. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247:536-540.
- Schatz, G. and Dobberstein, B. 1996. Common principles of protein translocation across membranes. *Science* 271:1519-1526.
- Schultz, J., Milpetz, F., Bork, P., and Ponting, C.P. 1998. SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc. Natl. Acad. Sci. U.S.A.* 95:5857-5864.
- Sigrist, C.J., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., Bairoch, A., and Bucher, P. 2002. PROSITE: A documented database using patterns and profiles as motif descriptors. *Brief. Bioinform.* 3:265-274.
- Sjölander, K. 2004. Phylogenomic inference of protein molecular function: Advances and challenges. *Bioinformatics* 20:170-179.
- Tatusov, R.L., Natale, D.A., Garkavtsev, I.V., Tatusova, T.A., Shankavaram, U.T., Rao, B.S., Kiryutin, B., Galperin, M.Y., Fedorova, N.D., and Koonin, E.V. 2001. The COG database: New developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* 29:22-28.

Contributed by Nandini Krishnamurthy
and Kimmen V. Sjölander
University of California
Berkeley, California

Analysis and Management of Microarray Gene Expression Data

Microarray technology has been available to researchers for approximately 10 years, and over time a few aspects of its use have become standard. In this unit, the basic steps and common applications for microarray gene expression studies are covered, including pre-processing, analysis, and data management. Several different array-based platforms are available to analyze transcript abundance, including spotted cDNAs, spotted oligonucleotides, and in situ-synthesized oligonucleotides (i.e., Affymetrix or Nimblegen; see UNIT 22.4 for discussion of different platforms). The information provided in this unit will be generally applicable to the different platforms and special considerations noted. Microarrays are now used for purposes besides monitoring transcript abundance, including detecting mutations and protein binding (see also UNIT 22.1); however, these will not be addressed here, and microarray will always refer to a transcript abundance microarray in this unit.

Microarray practitioners have the benefit of standards that have been developed to specify what information should be recorded with a microarray experiment. The Microarray Gene Expression Data Society (MGED Society, see Internet Resources) has developed the Minimal Information About a Microarray Experiment (MIAME) checklist that has been adopted by many journals as a requirement for publication and has become an internationally accepted standard in this area. These journals typically also require deposition of data in public repositories such as ArrayExpress and the Gene Expression Omnibus (GEO; see Internet Resources). The MGED Society continues to work not only on standards but also on software applications to make them easier to utilize. The MGED effort has also provided standards for terminology (the MGED Ontology). These will be adhered to in this unit as much as possible to avoid ambiguous terms such as “probe,” which is often used in different settings to describe either the material that is labeled (as used in UNIT 22.1) or the material attached to the array. Instead, the material that is labeled is called the *labeled extract* and the material that is attached to the array is called the *reporter*. A

reporter represents a transcript, which in turn refers to a gene. For simplicity, the terms reporter, transcript, and gene may be used interchangeably in this unit. However, the signal measured after hybridization is that coming from a reporter and—unless signals from different reporters representing the same transcripts are combined (i.e., through averaging) and subsequently signals from different transcripts from the same gene are combined—the authors are actually referring to reporters. An *experiment* is a collection of *assays* (hybridizations) performed for a common purpose. An assay can be *one-channel* if only one labeled extract is hybridized to the array (e.g., nylon filter assays or Affymetrix assays) or *two-channel* if two labeled extracts are simultaneously hybridized to the array (e.g., one labeled with a Cy3 dye and one labeled with a Cy5 dye).

This unit starts with a quick overview of what microarrays are good for and what they are not good for. The most common uses are: (1) to find expressed genes in a given sample or experimental condition, (2) to find differentially expressed genes between two samples or experimental conditions, (3) to find expression signatures that define particular conditions or classes of samples, and (4) to find expression signatures that define particular sets of genes. There are many variations of these, but this unit will focus on these standard use cases. There are many things that microarrays are not good for, for example, determining the expression or differential expression of a single gene or small set of genes. In microarray studies, some accuracy and power for any given gene are sacrificed to obtain parallel measurements for thousands of genes. One would never measure a single gene or even a few genes using microarrays, which are generally far less accurate and sensitive than other methods such as PCR-based methods. Microarrays are also not very good at measuring relative expression levels between different genes in the same condition. The variation in hybridization affinity of the different reporters makes their intensities incomparable in principle. In practice, there is generally significant rank correlation between the true expression levels and the microarray

intensities of genes, but the comparison can be quite misleading for any given pair of genes. Therefore, one should only compare the intensity levels of the same reporter between experiments, and not of different reporters in the same experiment. It should also be kept in mind that microarrays measure the steady-state levels of RNA transcripts. Although often used as proxies for gene activity, they do not, in general, accurately reflect protein levels or their functional activity.

Figure 19.6.1 illustrates the flow of steps involved in carrying out a complete microarray experiment and these are addressed in this unit, which is organized in the following manner:

- Experimental Design
- Data Preprocessing
- Expression and Differential Expression
- Classification
- Looking at Gene Sets
- Databases

It should be noted that many aspects of microarray analysis remain issues of ongoing

debate. As such, it is impossible for any researcher to write a primer on microarray data analysis without a significant amount of author bias. Therefore, in order to be forthright before proceeding, it is necessary to point out that this discussion at times represents what the authors have found to be the most efficient and effective methods. This is a result of personal experience and perspective, and should be taken as such. For a discussion of some approaches that have become “consensus” in dealing with different aspects of preprocessing and analysis, most of which coincide with the authors’ current preferred approaches, refer to Allison et al. (2006).

Finally, it is important to keep in mind that biological conclusions should not be based solely on a microarray experiment. However, a microarray experiment can be of significant value in terms of narrowing down hypotheses. As such, microarrays are just a part of the iterative process of discovering important biological relationships and should not be considered the last word.

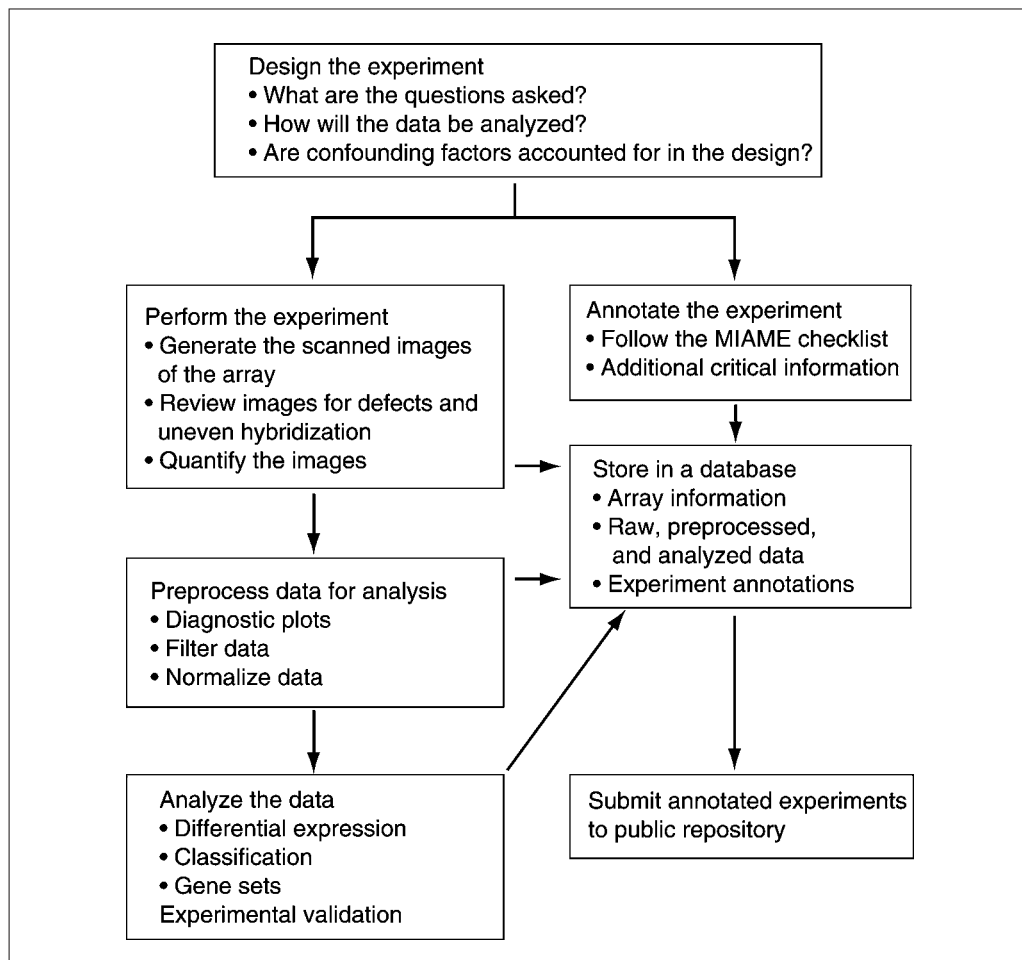


Figure 19.6.1 Steps involved in carrying out a complete microarray experiment.

EXPERIMENTAL DESIGN

The more time and care spent designing the experiment, the fewer headaches there will be during analysis. Experimental design depends on the study goal, but there are a few general principles that apply to all experiments. Perhaps the foremost issues are making sure there is sufficient replication and avoiding the introduction of confounding factors. Special considerations and issues related to design of two-channel assays are also considered below.

Replication

Even if a method delivered a perfect signal without any noise component, replication would be necessary, because one is generally interested in expression behavior in a population of samples, not in just one individual sample. For example, to understand the changes that take place between certain tumor and normal cells would require a representative random sample of both cell populations. As is typical in microarray studies, however, there might be an impossible trade-off in practice. Typically, only a limited number of assays can be performed, not enough for a truly random sample. This does not mean that one should not proceed; it just means that one must be careful about interpreting the results and do further follow-up to make sure the results generalize.

In microarray experiments, technical replicates (i.e., replicates of the same biological sample) should not be treated as independent replicates in the subsequent statistical analyses. Instead, technical replicates should be combined to obtain a single intensity per reporter for the same biological sample, e.g., by averaging the intensities or, in the case of two-channel data, averaging the log ratios. This way technical replicates can be used to provide a more reliable intensity for a given sample, but not as genuine replicates for the sample type under consideration. For Affymetrix, technical variability is usually quite low and (once established with a few technical replicates in a pilot study) technical replicates can be replaced with more biological replicates. Technical variability can be greater for two-channel platforms, especially due to differences in the fluorescent dyes used for labeling. Dye-swaps serving as a form of technical replicate are discussed below.

One can obtain multiple biological replicates and pool the starting materials to hybridize to a single array. This will decrease variation without requiring more arrays. It must be kept in mind, however, that there must be sufficient replicate assays in order to

perform statistical analysis. Thus, two replicates with a pool of five samples for each are not better than ten replicates each of individual samples. One should generate a reasonable number of assays per condition. If there are more samples available than arrays, then it makes sense to pool samples especially if the biological variability between samples in the same condition can be large and not related to the hypothesis being tested. Pooling should of course be done uniformly across all assays.

There is no hard and fast formula for determining the “right” number of replicates. The answer depends on knowing more about the nature of the particular expression levels in the samples, which, if known ahead of time, would eliminate the need to do the experiment. Thus, one is forced to make a judgment call based on the practicality of obtaining the samples and the cost of the arrays. Five or six replicates are always better than two or three. Ten or more is good if it is affordable and does not become so unwieldy that it introduces factors such as a change in technician. This is not to say one cannot ever extract meaningful results with two replicates per condition. If many of the planned assays fail, then one might do such an analysis, but it is generally not a good idea to design the experiment with such poor replication.

Confounding Factors

To illustrate the importance of controlling for confounding factors, first consider the simplest type of experiment where the goal is to compare conditions A and B on one-channel arrays. Suppose that appropriate random samples have been chosen to represent the two conditions. If all assays for condition A are processed together, and then all assays for condition B are processed together, there will be no way to know if the effects seen derive from the condition (A versus B) or from the processing batch. For example, since hybridization date is known to be an important factor in microarrays, one should ideally hybridize all samples together. If this is impossible, then one should hybridize half of the A samples and half of the B samples together in the first batch, and the remaining A and B samples together in the second batch. Although this will increase the variability within both groups, possibly decreasing the power of the downstream analyses from what could be obtained if all assays were done at once, this is necessary to avoid introducing a confounding factor that can lead to false results.

Another confounding factor to be aware of is the “individual.” Suppose for example that one is comparing condition A to condition B, and that the actual transcript abundance of gene *g* is typically twice as high in condition B compared to condition A. Suppose that one is interested in discovering this effect. One could measure condition A from each of *M* individuals and condition B from each of *N* different individuals. If there is very little variability across the individuals in each group, then an unpaired statistical test should have the power to detect the effect for relatively small values of *M* and *N* (see APPENDIX 31 for a discussion of unpaired and paired tests). However, if there is a lot of variability from individual to individual, then, even if the ratio of expression level for gene *g* in condition B versus condition A in the same individual is always two, the more variable the gene is within conditions A and B, the more replicates will be needed to achieve sufficient power to detect the effect. In this case, the distribution of the paired (within individual) intensity ratios for gene *g* will be very tightly centered around 2, while the separate distributions of intensities for gene *g* in conditions A and B will have large variance. If one expects such an effect, then one should design the experiment so as to have pairing by individual if possible.

There are many other factors that can potentially introduce unwanted variability, e.g., harvest conditions, RNA extraction date, sample preparation date, hybridization date. Sometimes it is impossible to randomize for all known effects. The hybridization date is an important factor to avoid confounding. When samples are harvested, extracted, prepared, and labeled all side-by-side but are then hybridized to the arrays on separate dates, the main effect in the data will often turn out to result from the hybridization date. This is not platform dependent, but appears to be the case across the spectrum of platforms from Affymetrix to two-channel spotted glass slide arrays.

Balancing Replication and Confounding Factors

Judgment calls as usual are unavoidable. For example, will it be better to do four replicates all together or to do six replicates but split over two batches? The answer to this question can rarely be determined ahead of time. Sometimes confounding factors are not readily apparent ahead of time, and will only be revealed after the experiment has been completed, for example, by doing quality control

clustering (see section on Classification below). If this step reveals factors that have inadvertently confounded the factors of interest, then a repeat of the experiment might be necessary. Even if an experimental factor (such as hybridization date) is not confounded with the factors of interest, it might still be necessary to take the experimental factor into account during the analysis (“block” analyses; see Expression and Differential Expression section). In any case, the three main concerns in experimental design are: (1) does the experimental design effectively address the questions and/or hypothesis of interest, (2) is there sufficient biological replication, and (3) are there any obvious confounding factors?

Two-Channel Assays

Greater challenges in experimental design come from the use of two-channel assays. The first thing to always keep in mind is that two-channel assays should not be thought of as two one-channel assays. In particular, the intensities coming from the two channels should not be used separately in analysis; instead one should work directly with their ratios. The variability (arising primarily from spot-to-spot inconsistencies) introduced by not taking ratios could be so dramatic as to affect all downstream analyses profoundly. Thus, all of the two-channel experimental designs discussed here will revolve around using ratios. Although a one-channel experiment might be easier to design, one might still opt to use two-channel arrays for practical reasons. For example, two-channel spotted arrays can be very inexpensive, as low as \$20 per slide, and can easily be produced in-house.

Direct comparison

Consider again the case where two experimental conditions A and B are to be compared. In this case, the simplest two-channel approach is to use a *direct comparison* where A is labeled for one channel and B is labeled for the other channel. For the sake of normalization (to be discussed in the Data Preprocessing section below), one might hybridize each pair of samples twice, swapping the channels for half the assays, and taking the reciprocals of the intensity ratios for those. Such *dye-swaps* represent technical replicates and should be combined to form one (more reliable) intensity ratio for each sample pair. In any case, with direct comparison data some form of statistical test can be performed on the ratios to test for differential expression between the two conditions (as described in the Expression

and Differential Expression section below). If, however, the goal is to compare multiple conditions, possibly also to perform a clustering or class prediction analysis (as discussed in the Classification section below), then simply having direct comparison ratios might not be sufficient.

Reference design

An alternative to the direct comparison design is the *reference design*. In this case, to compare conditions A and B, a third condition C is chosen as a reference condition. A number of assays are prepared with a sample of type A labeled for one channel and a sample of type C labeled for the other, and a number of assays are prepared with a sample of type B labeled for one channel and a sample of type C labeled for the other. The resulting A/C ratios and B/C ratios can be analyzed analogously to one-channel data. The reference sample C must be chosen judiciously, because it can have an impact on how well the data can be normalized. If there are a lot of differentially expressed genes between conditions A and C, and similarly for conditions B versus C, then there should not be gross asymmetry between the up-regulated and down-regulated genes. Otherwise, normalization will be problematic, as described in the section on Data Pre-processing below. Often C is taken as a pool of conditions A and B. Another popular approach is to use a universal reference, such as those sold by Stratagene (see Internet Resources). An advantage of a universal reference is that it might allow for comparisons across studies. However, using a universal reference can have a dramatic effect on the genes identified in a differential expression analysis. This is because genes with higher mean intensities also have higher variance. For example, if a gene is differentially expressed between conditions A and B but its abundance in both these conditions is at the low end of detectable signal, and if the abundance of this gene in the reference condition C is very high, the measured intensities of the latter could introduce enough variance to render the differential expression undetectable. This effect would be mitigated by a reference whose abundance for that gene is in a similar range as A and B, which is exactly what would be obtained using a pooled reference. Furthermore, a pooled reference would achieve this for every gene.

When comparing three conditions A, B, and C, one can choose a fourth sample D to use as common reference for all three conditions. It might still be preferable to construct a direct

comparison design, however. If one conducts N assays each of A/B, B/C, and C/A, then all possible comparisons can still be made. Once one has four or more conditions and wants to make all possible comparisons, it starts to become prohibitive to do direct comparisons between each pair of conditions and it becomes necessary to use a common reference for all assays. With four conditions there are six comparisons, with five conditions there are ten, and it grows quickly from there. Some authors suggest “daisy chaining” the samples in a large loop design, i.e., performing A/B assays, B/C assays, C/D assays, and so on (Kerr and Churchill, 2001). However, this kind of design limits the analysis methods available and makes the A/D comparison depend on the quality of the B/C assays. If there are many conditions this can be risky. Yang and Speed (2002) discuss this as well as additional design issues particularly pertinent to two-channel assays.

A variant of the reference design is the *blocked reference design*. Rather than using the same reference sample, independent references are matched (relative to some relevant factor) to the conditions that one wants to compare and are co-hybridized to each respectively. This type of design can be useful in time series experiments with a baseline and in drug experiments with a placebo group (see Steibel and Rosa, 2005).

DATA PREPROCESSING

Upon completion of the bench work which yields hybridized arrays, a series of steps needs to be carried out to generate the final data to be used as input to downstream analyses. The first step after the hybridization procedure is *image acquisition*, which entails generating digital images from the assays, e.g., using a phosphorimager (for radioactively labeled targets) or a scanning laser microscope (for fluorescently labeled targets). Once the images have been generated, it is necessary to derive quantified values from them. This process is referred to as *feature extraction*, *image quantification*, or *image analysis*. Then visual inspection and diagnostic plots can help assess and possibly filter the data. Finally, the data need to be *normalized*. These preprocessing steps, beginning with feature extraction, are discussed in detail below.

Feature Extraction

A variety of freeware and commercial software packages have been developed to address feature extraction and this remains an

active area of research. For spotted arrays, a feature extraction software (e.g., ArrayVision, GenePix, ScanAlyze, Spot, TM4 Spotfinder; see Internet Resources) typically produces the main signal measures and additional measures to help assess background noise and artifacts (e.g., background glow, dust, scratches, smears). Yang et al. (2002a) provide a detailed description of the various image analysis steps for spotted arrays as well as a comparison of some software packages, with a particular focus on two-channel data.

With probe set arrays like the Affymetrix chips, each transcript is represented by a set of short oligonucleotides. The array consists of *probe cells*, each containing (perfect match or single-mismatch) copies of the same short oligonucleotide; the collection of probe cells corresponding to the same transcript is called a *probe set* (i.e., the reporter). In this case, feature extraction entails first getting quantified values for the probe cells, then using these to generate a summary expression measure for each probe set. A variety of different algorithms have been proposed which generate probe cell and probe set values (typically accounting for background) as well as additional measures. To cite a few, the MAS 4.0 and the more recent (and substantially different) MAS 5.0 algorithms were developed at Affymetrix, with version 5.0 intended to subsume version 4.0. Details on this can be found in the Affymetrix statistical algorithms reference guide (see Internet Resources). The dChip software (for details, including publications and tutorials, see Web site indicated in Internet Resources) employs a model-based approach towards feature extraction, contingent on the availability of multiple assays in the experiment. The Robust Multichip Analysis (RMA) is another method to quantify Affymetrix data in the presence of multiple assays. A stand-alone implementation for Windows and Linux is provided by the RMA-Express software (see Internet Resources). Implementations of RMA and its variant gcRMA are also available in the affy and gcrma R packages from the Bioconductor project (see Internet Resources). Affymetrix has released another algorithm, named PLIER (Probe Logarithmic Intensity Error), which builds upon some of the above approaches.

These approaches to summarization of Affymetrix data typically incorporate normalization options (see below) and can produce substantially different results. Benchmark datasets have been made available to facilitate comparisons and performance assessments of

summarization methods for short oligonucleotide platforms, and various papers have been published to address these (e.g., Cope et al., 2003; Rajagopalan, 2003; Barash et al., 2004; Choe et al., 2005; Irizarry et al., 2006). Currently, the authors preferentially use gcRMA as a default Affymetrix quantification software. The gcRMA software package has performed well in the published comparisons just listed. As stated in Irizarry et al. (2006), gcRMA was one of the methods that “appear to be finding a balance between accuracy and precision that permits them to perform well across the range of gene expression.” The original RMA was noted to compress fold change ratios. gcRMA takes into account GC-content in the estimation of non-specific binding and is improved in this area.

Some feature extraction software output data only at a very raw level, leaving to the user further decisions on how to process these prior to analysis; others can automatically perform certain kinds of preprocessing and might offer the user various preprocessing choices. In any case, it is important to be aware of the main preprocessing steps to consider and of the assumptions that these make. There is no hard and fast rule that can be applied in all situations and the best choice of preprocessing depends on the experiment at hand, including the types of samples that are hybridized to the arrays and the collection of transcripts represented on the particular array used. Before using a preprocessing method, especially a normalization method, it is important to be aware of the assumptions made. Using these methods indiscriminately can lead to wrong conclusions, as errors will propagate through the cascade of subsequent steps. Some of the common preprocessing issues, methods, and assumptions are discussed below.

Background Correction

As mentioned above, part of the signal in an array image comes from background noise and does not reflect actual labeled extract hybridized to reporter. This means that each reporter's signal (sometimes referred to as the *foreground signal*) incorporates some background noise. There are various approaches that have been proposed to estimate background and to adjust the foreground signals using these estimates. Often these methods try to estimate background locally (i.e., estimating background as a function of location on the chip) and correct the foreground signal by subtracting background signal for the corresponding location on the chip. Different

quantification packages vary in the approaches used to perform background estimates. For spotted arrays, Yang et al. (2002a) illustrate some of the approaches. Typically, when working with most of the current spotted array quantification software (except for the Spot software), the authors do not apply background subtraction, but simply use the foreground measurements output by the software. This is because subtraction often introduces significant amounts of additional variability, especially at the low end of the intensity spectrum, and adds the complication of negative signal intensities.

Images, Diagnostic Plots, and Quality Measures

It is always recommended to examine the files (e.g., .TIF or .DAT) generated by the image acquisition process, as well as additional diagnostic plots of each assay, prior to proceeding with preprocessing. This can help identify obvious artifacts and eliminate from further analyses those reporters or assays that are of poor quality, as well as identify some of the sources of bias in the data that might have to be dealt with. Among the freeware providing diagnostic plot functionality, some of the Bioconductor packages (see Internet Resources) are very useful. For example, for two-channel arrays, the *marray* package has the capability to plot color-coded images of different measurements (e.g., the individual channels foreground and background, the difference of the log intensities in the two channels) laid over the spatial configuration of the array. The same package also provides functionality for various boxplots and for the very useful *M* versus *A* plots, where *M* represents the difference in \log_2 intensities in the two channels and *A* the average \log_2 intensity in the two channels. Namely, $M = \log_2(R) - \log_2(G)$ and $A = [\log_2(R) + \log_2(G)]/2$, where *R* and *G* are the signals (possibly background corrected) of the Cy5 (red) and Cy3 (green) channels, respectively. Figure 19.6.2A illustrates an example of such a plot. A clean self-self hybridization should yield points tightly clustered around the horizontal line $M = 0$ (intensity is the same in both channels). Those reporters that have higher average intensity between the two channels are further to the right in the graph. Albeit, an *M* versus *A* plot is merely obtained by a rotation and rescaling of the $\log_2(R)$ versus $\log_2(G)$ scatter plot, it makes it easier to visualize intensity-dependent red versus green effects (that is, biases which depend on the average \log_2 intensity *A*), which appear as deviations from the horizontal $M = 0$ line.

The TM4 Spotfinder software (Saeed et al., 2003) is another example of freeware providing some diagnostic plot capabilities.

Affymetrix provides a number of metrics that can be used to assess assay quality. Details can be found in the Data Analysis Fundamentals Manual (see Internet Resources). Hybridization and poly-A controls are spiked into the sample and therefore have expected signal intensities. Two genes, GAPDH and β -actin, are used as controls to assess the 3'-to-5' ratio. A high ratio indicates possible degraded RNA or an inefficient reaction during sample preparation. The other measures of quality are more subjective and one cannot generally say what the expected values should be. However, given replication, these measures should not vary tremendously. This is a general rule that applies to any array platform. One of the basic measures in an Affymetrix array is the percent of genes called present. Outliers from the distribution over all replicates in the experiment, or extremely high or low numbers of genes called present, indicate a poor quality assay. Similarly, the scaling factors (when the software is set to automatically normalize the assays, see below) should not vary tremendously from assay to assay. For more on this, the Affymetrix documentation should be consulted.

Two-channel quantification software packages generally offer fewer metrics; for example, there is no present call. However, with two-channel arrays, it can be useful to graph the correlation between pre- and post-normalized data from each assay and then to graph these correlations to identify outliers. Any assay that undergoes a significantly more extreme normalization than the rest of the assays is likely to have some quality issues.

Another very useful approach for identifying poor quality data is to use various clustering methods. The clustering approach is discussed below (see Classification).

Filtering

Typically, for each assay, some of the reporters on the array might need to be excluded from further processing and analyses, due to a variety of reasons. At times a whole assay might need to be discarded if it is of poor quality overall. Visual inspection of array images and color-coded plot images are the first tools that should be used to detect artifacts like scratches or splotches that might lead to the exclusion of a set of reporters, if not the whole assay. The quality metrics described above can be used to further filter out poor-quality assays. In addition, some of the image analysis

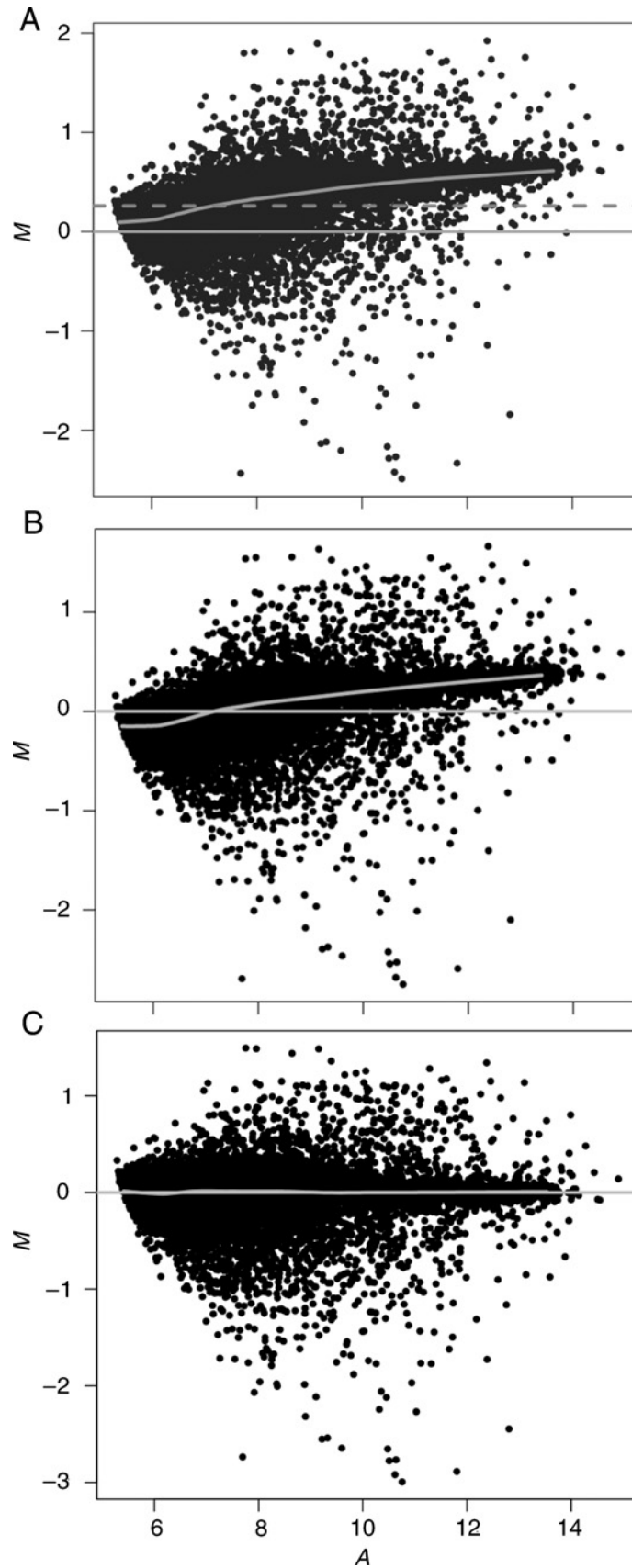


Figure 19.6.2 Legend at right.

packages provide measures that can be used to identify potentially bad reporters in an assay, which should be filtered out. Other reporters that one might want to exclude from further analyses might be specific controls, like anchors that are used to help in the feature extraction step, or blanks.

Normalization

In microarray jargon, *normalization* indicates the attempt to identify and remove systematic sources of variation in the measured intensities due to wide variety of factors including separate reverse transcription and labeling, different scanning parameters, print-tip differences (when multiple printing pins are used to print each array), other spatial effects, different dye labeling efficiencies, the quality of the microarray printing, or the quality of the mRNA used. As a simple example, if two assays use the same labeled extract and the images are scanned with different settings, a reporter on one of the images might be brighter than the corresponding reporter on the other image; hence, the resulting raw quantified intensity will be higher, even though in reality the two reporters should have the same intensity as they represent the RNA abundance level of the same transcript in identical samples. One should never knowingly introduce such variations in the scanner settings, but when it happens, one can attempt some kind of rescaling normalization to try to salvage the data set. In general, normalization puts the data on equal footing before intensity comparisons are made. Normalization is a very delicate step, however, that can have a dramatic effect on the results. If not done carefully, one can easily normalize out biological effects of interest. Methods have been developed for both one-channel and two-channel data and the most common are described below. This does not represent an exhaustive description of all methods published on the subject to date, as this is an active area of research. The goal of this discussion is to examine some of the subtleties of this process and of the common assumptions on which normalization techniques rely.

Two-channel assays

As mentioned in the section on Experimental Design, typically for this kind of assay, one should only work with the ratios of the intensities from the two channels (after possible background subtraction) rather than with the separate intensities in each of the channels, as the former approach can greatly reduce the spot-to-spot variation due to spot size and concentration. It is also very common to work in logarithmic scale (e.g., using the M values defined above). Thus, if two identical samples are hybridized to the two channels of an array, one would expect that, after biases are removed, all points should be close to the line $M = 0$ in the M versus A plot. Typically, however, plots of raw data (i.e., prior to bias removal) do not look like this. For example, the Cy3 dye can show a stronger signal, and this effect can be intensity (A) dependent.

Typical methods to normalize M values include normalization by a global constant, global loess normalization, print-tip loess normalization, and dye-swap normalization (performed alone or on top of the previous methods, if the experiment contains dye-swaps). Additional scaling adjustments are also sometimes applied. These methods and their assumptions are discussed briefly here; for further details, the reader is referred to Yang et al. (2002b) and Quackenbush (2002).

The first three methods involve two choices: (a) which normalization factors to use to adjust the values at each reporter, and (b) which reporters to use in order to compute the normalization factors. The latter could be all reporters on the array or a suitable subset of these, such as a set of appropriate controls (e.g., reporters representing transcripts known to be expressed at the same level in both samples hybridized to the array).

In what follows, it should be kept in mind that what comes out of the normalization step are normalized (log) ratios, not normalized values for the individual channels.

Normalization by a global constant is a very simplistic approach and is not generally recommended, but is described for the sake of motivation. This normalization is performed

Figure 19.6.2 (at left) (A) Plot of M versus A for a two-channel microarray assay (see text). The dashed horizontal line is the line $M = m$ where m is the median of M over all reporters. The solid curve represents a global loess fit. (B) Plot of M versus A for the data from panel A after normalization by the global constant m . The solid curve represents a global loess fit and illustrates that this normalization did not eliminate intensity-dependent biases. (C) Plot of M versus A for the data from panel A after global loess normalization. All global loess fitting and normalization performed with the Bioconductor marray R package.

by subtracting from the M value at each reporter the same constant m , which is obtained by computing the median or mean of M over a chosen set of reporters as per (b) above. In this case, the normalization amounts to a vertical shift in the M versus A plot (see Fig. 19.6.2B). This method assumes that the biases to be corrected are the same at all reporters. It also assumes that the differentially expressed genes in the two samples being hybridized to the two channels show similar amounts of up and down regulation. If this last condition does not hold, one could normalize out a biological effect and not just the dye bias. For example, if one uses a brain-specific array (i.e., an array representing mostly genes expressed in the brain) and co-hybridizes a brain sample versus a pancreas sample to that array, then normalizing the M values by the constant computed from all reporters on the array would not be a wise choice of normalization, as it is unlikely that the above assumptions are satisfied.

The assumption that the biases are the same at all reporters is a very strong one and typically does not hold. In most cases there is a trend in the M versus A plot which indicates that the biases might depend on A , i.e., on the average (log) intensity in the two channels (as is the case for the data in Fig. 19.6.2A). To address this, the *global loess* (or *lowess*) *normalization* method subtracts from the M value for each reporter a value $m(A)$ that depends on the A value of the reporters. This has the effect of shifting points in the plot up or down by different amounts based on their A value, potentially allowing a greater number to lie near $M = 0$. The value $m(A)$ is computed by fitting a curve to an appropriate subset (possibly all) of the reporters in the M versus A plot, using a “locally weighted polynomial regression” approach termed *lowess* or *loess* depending on the implementation (Cleveland, 1979; Cleveland and Devlin, 1988). For each A , $m(A)$ is taken to be the M value of the corresponding point on this curve. The subset to which the curve is fit needs to be chosen so that for any range of A values, the differentially expressed genes in the two samples being hybridized to the two channels show similar amounts of up and down regulation. Loess curves are somewhat robust to outliers, so if there is only a relatively small amount of differential expression then this assumption can be violated to some extent; however, the more differential expression there is between the two groups, the more this assumption matters. Figure 19.6.2C illus-

trates the results of performing global loess normalization on the data from panel A. On the one hand, global loess normalization addresses a more general type of bias (intensity dependent) than normalization by a constant; on the other hand, it requires the samples and arrays to satisfy stronger assumptions, as similar assumptions must now hold not just overall, but at every intensity level. In general, normalization methods that correct for a larger variety of biases also typically need to make more assumptions on the samples and arrays. This underscores the importance of keeping normalization issues in mind from the start of an experiment, i.e., during the experimental design phase.

When an array has been printed using multiple printing tips, there might be systematic differences between printing tips that also need to be corrected. Figure 19.6.3A shows curves fit for each print-tip, where each curve is generated by fitting it to those points in the M versus A plot corresponding to reporters in the same print-tip. In this case, it appears that different print-tips have different trends. The *print-tip loess* (or *lowess*) method is aimed at correcting both intensity-dependent and print-tip-dependent biases by performing a separate loess correction for each of the print-tips. This is done by subtracting from M at each reporter a value $m(p,A)$ which depends on A at that reporter as well as on the print-tip p to which the reporter belongs. This value, for each A , is obtained by fitting a loess curve for each print-tip to an appropriate subset (possibly all) of the reporters on that print-tip in the M versus A plot (see Fig. 19.6.3A). For each print-tip, the subset to which the curve is fitted needs to be chosen so that one can make the following assumption on the corresponding reporters: for any small range of A values and any print-tip, the differentially expressed genes in the two samples being hybridized to the two channels show similar amounts of up and down regulation. Figure 19.6.3B illustrates the results of performing print-tip lowess normalization on the data from panel A. If, due to the nature of the samples and the reporters on the array, one does not have such a subset available for all print-tips, then this normalization should not be applied. Note that the print-tip loess method is basically analogous to the global loess except that the data are first stratified (by print-tip) and loess is applied within each stratum instead of globally. In the same spirit as print-tip loess normalization, one can address similarly other kinds of bias that might be

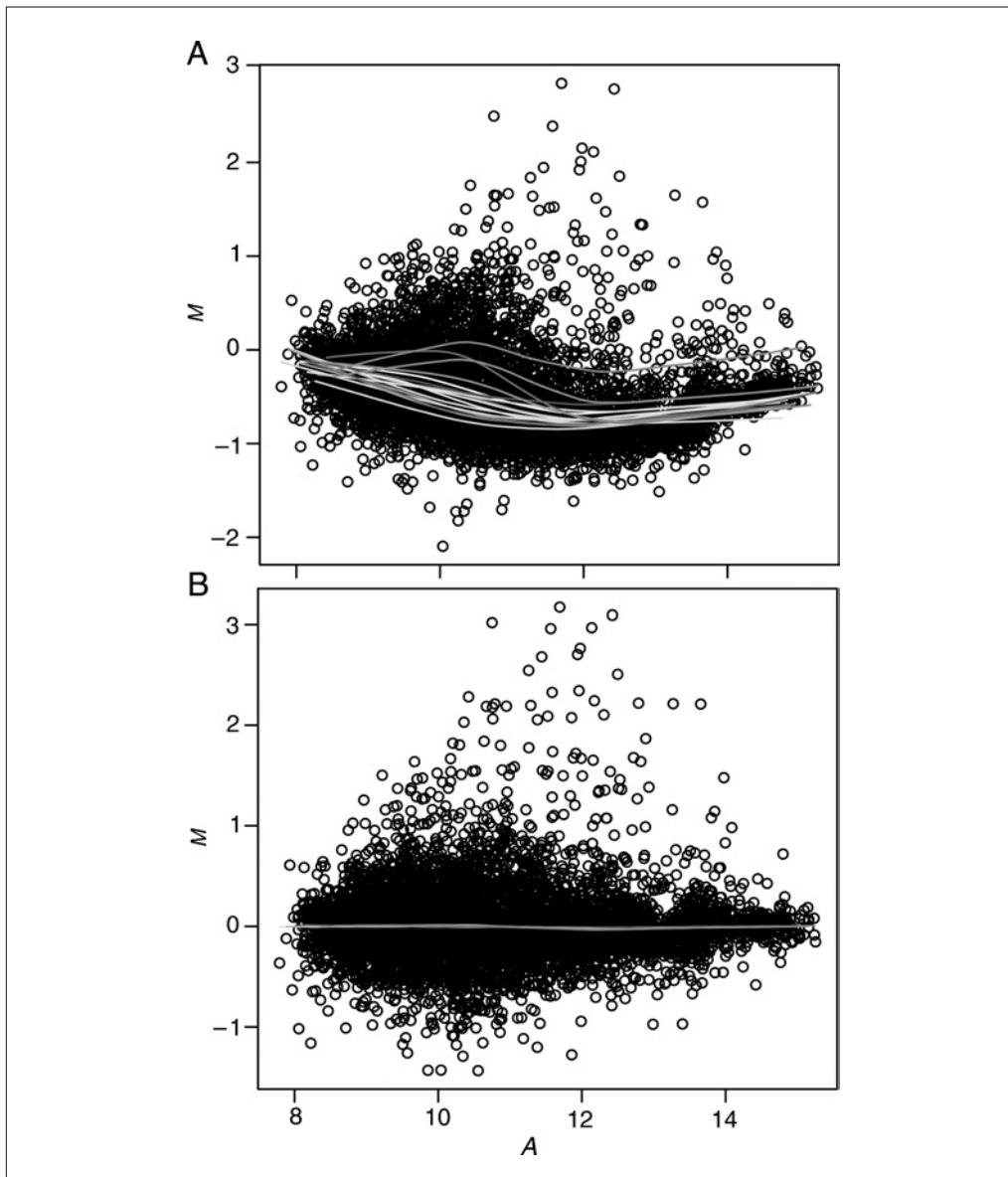


Figure 19.6.3 (A) Plot of M versus A for a two-channel microarray assay. Each curve represents a loess fit to one of the print-tips. Data generated with the Bioconductor *sma* R package, using the *MouseArray* dataset provided therein. (B) Plot of M versus A for the data from panel A after print-tip loess normalization (performed with the Bioconductor *sma* R package). For color version of this figure see <http://www.currentprotocols.com>.

uncovered from diagnostic plots, e.g., plate biases, or spatial biases other than print-tip groups.

In addition to the adjustments described above for a given “stratum” (e.g., print-tip) loess normalization, some methods additionally perform a scale adjustment across strata within a slide. Essentially, the stratum loess normalization will transform the data so that the M values will be centered around 0 for each stratum across all intensities. A *scale adjustment* is aimed at adjusting the spreads of these values, making them equal over all strata. This adjustment makes the further assumption that

the spreads of the M values should be the same over the different strata.

Each of the methods described above is applied within each assay in an experiment with the goal of putting the two channels on equal footing. Often, when multiple assays are to be compared, after normalizing each assay with one of the above methods, it is possible to proceed to the comparisons. In some cases if, after normalizing the M values on each assay, one sees too much variation between their spread across all assays and it is assumed that the samples at hand should roughly have the same kind of variation of M values for the

reporters represented on the array, additional scale adjustments could be performed that will render these spreads roughly the same across all the assays. However, the authors have found global and print-tip loess effective (whenever applicable) and typically refrain from making scale adjustments both within and across assays.

As mentioned in the section on Experimental Design, technical dye-swaps are performed in some experimental designs. For this situation, dye-swap normalization might be applicable. This consists, for each reporter, in taking the average of the M value of one assay with the negative of the M value on the other assay. This is usually referred to as *dye-swap normalization* or *self-normalization*. The assumption made by this method is that the channel biases are very similar in the two assays. As long as this assumption can be made, when dye-swaps are available, this method can be a viable alternative if the assumptions of the previously described normalization methods are not satisfied. Even in cases when one of the previous methods is applicable, dye-swap normalization can be applied following normalization by one of these methods, to combine technical dye-swaps replicates. Moreover, there might be occasional reporter-specific dye incorporation biases which dye-swap normalization can help suppress. Note, however, that application of dye-swap normalization could exaggerate intensity-dependent biases if the assumption that channel biases are similar in two dye-swap assays is not satisfied.

A few comments should be made regarding the typical subset of reporters on an array used to compute the normalization function in the constant or loess normalization methods. This needs to be determined based on the biological assumptions that appear reasonable for the samples at hand and the genes represented on the array. If these assumptions are not valid for the set of all reporters on the array, there might be alternative choices that will still work. For example, if there is set of genes represented on the array which are known to be expressed at roughly the same level in the two samples and which (for intensity dependent normalizations) span the range of A values, then this set could be used. Typically, however, one does not have this kind of prior knowledge. An alternative type of control that might be used is a spiked control. Spiked controls are control reporters that are printed onto the array, representing exogenous transcripts whose corresponding mRNAs can be added to the labeling reaction along with the experimental RNA. A concern with the latter type of control is that

they are not useful for normalization in situations when the quality/quantity of the RNA in the two samples of interest is different. However, spikes are very useful as they can provide an internal quality control for labeling reactions, and can allow rapid identification of problems in terms of mRNA quality along with any potential issues that may have arisen due to an error in the labeling experimental procedure.

The two-channel normalization methods described above are implemented in many freeware and commercial software packages. For example, within the Bioconductor project, one can find implementations both in the *array* and in the *limma* packages. Similarly, the MIDAS component of TM4 implements loess normalization.

Besides the methods described above, other approaches have been developed for two-channel data. Among these, one that is also available from Bioconductor is the *variance stabilizing normalization* (Huber et al., 2002) implemented in the *vsn* package. This method consists of first applying, separately to each channel, a calibration to standardize the data, combined with a variance-stabilizing transformation aimed at rendering the resulting variance for each reporter approximately independent of its mean intensity. (As such, this method could be employed to normalize one-channel data. Moreover, like loess, this can be applied in a stratified fashion.) The difference of the resulting values in each channel is then taken for each reporter, generating the so-called *generalized log-ratios*, which can be employed instead of the traditional M . For two-channel data, however, the authors feel it is best to work solely with ratios without separating the two channels, and thus prefer to utilize approaches which operate directly on the M values, like the loess methods utilizing M versus A plots.

One-channel assays

Affymetrix probe set level data are similar to one-channel data on other platforms in that, for each assay, there is only one value (corresponding to the unique channel) attached to each reporter. Therefore, normalization methods applicable to Affymetrix probe set data are generally also applicable to other one-channel platform data and vice versa. In addition, for Affymetrix data, one could proceed to normalization at the probe cell instead of the probe set level, and some packages implement this.

One very common method of putting different assays on equal footing is to transform the intensities so that their average (or their

median) over an appropriately chosen subset S of the reporters is the same, say T (as target), for all assays. The correction might be additive or multiplicative depending on whether or not the data are in log scale. The assumptions made by this type of normalization are that the biases are the same for all reporters on the assay and that the samples hybridized to the different assays have the same average (median) RNA abundance over the transcripts corresponding to the reporters in S .

A more sophisticated method is the *quantile normalization* method (Bolstad et al., 2003), which makes the assumption that the samples hybridized to the different assays have roughly the same distribution of RNA abundance over the transcripts represented on the array. This method transforms the intensities so that the bulk of the intensity distribution is the same for all assays in an experiment, typically with some differences in the distribution tails (which might reflect actual biological differences). For Affymetrix data, this method is usually applied at the probe cell level. To assess whether or not possible intensity-dependent biases are removed by this kind of processing, one can look at M versus A plots for pairs of one-channel assays. Generally quantile normalization helps remove this type of bias. There are other methods geared at dealing with intensity-dependent biases and specifically based on the M versus A plot idea. The complication is that here one does not have two channels per assay. To deal with this, some of these methods select one of the assays as a baseline and form ratios relative to this for each of the remaining assays. Others (e.g., *cyclic loess* and *contrast-based normalization*) avoid selection of a baseline and use different strategies to overcome this issue. A review and comparison of the above methods can be found in Bolstad et al. (2003). Based on these comparisons, those authors recommend quantile normalization as the preferred choice. The authors of this unit also preferentially utilize this method (applied at the probe cell level in the case of Affymetrix). Quantile normalization is available in many free and commercial packages, including the Bioconductor *affy* and *gcrma* packages.

EXPRESSION AND DIFFERENTIAL EXPRESSION

Expression

Perhaps the most basic question that can be asked is which genes are expressed in the sample being assayed. Unfortunately, most mi-

croarrays are not designed to answer this question. Further, no microarray can determine that a gene is not expressed, because the gene might simply be expressed below the background level of the sensitivity of the array. On the other hand, the signal of a truly nonexpressed gene can be quite high, so in order to have any chance at making a reliable call one needs some kind of control elements on the array. This is provided on probe set arrays that have mismatch probe cells (e.g., Affymetrix arrays) which are used by some quantification software to make presence calls. For example, in the MAS 5.0 software, a statistical test to assess whether the mismatch probes are significantly dimmer than the match probes is typically performed and a p -value is computed. By default, if this p -value is .04 or less, then the software reports a *Present* call (P), if it is 0.6 or higher it reports an *Absent* call (A), and if it is between 0.4 and .06 it reports a *Marginal* call (M). The A call must be interpreted as insufficient *evidence* of expression, and not the absence of expression itself. Furthermore, the P calls are not multiple testing corrected (a correction needed when performing many hypothesis tests simultaneously, see below), so one should expect 4% of the array to be called present even if there are no expressed genes at all.

For arrays that do not have mismatch elements, often genes that are expected to not be expressed are incorporated into the arrays as controls. With enough controls one can get a distribution of nonexpressed elements, and can use some percentile cutoff as a criterion for expression. Since it is often hard to know what should be unexpressed, sequences from highly unrelated species are often used. As such, they might not represent the true null case, so this is subject to a very large margin of error and must be used as only a very crude measure. Often studies use blank spots to generate the “unexpressed” distributions; this is an even rougher estimate and should always be considered as such in any downstream analysis.

Differential Expression

Differential expression analysis involves using microarray data to find genes that are expressed at different levels between two experimental conditions. It should be kept in mind that what is measured for any given transcript is its percentage of the total labeled RNA. Therefore, the differences measured for a given gene are affected both by copy number per cell and by total RNA amount per cell.

Differential expression analysis is more complicated than might be expected and often requires difficult judgment calls. An in-depth look at the problem can be found in Ewens and Grant (2005) as well as Ge et al. (2003). Indeed, every differential expression protocol suffers from a trade-off between advantages and disadvantages, many of which cannot be determined ahead of time. The first issue to be aware of is that of multiple testing. This refers to the phenomenon that, when multiple tests are run in parallel, what might be unlikely to be observed in the data for one test becomes more likely to be observed in at least one of the parallel tests. When thousands of tests are run, what is very unlikely can become very likely. Therefore, any p -values that could be derived from tests on the individual reporters should be very carefully interpreted in the context of many reporters (typically thousands) being tested simultaneously. Usually p -values are modified in order to have sensible error rate control for all of the tests in parallel. The problem is complicated by the large number of tests (typically thousands or tens of thousands) and the small number of replicates available (typically fewer than 10). This renders the traditional and straightforward multiple testing corrections, such as Bonferroni corrections (Ewens and Grant, 2005), virtually useless.

In what follows, it is assumed that all data have been fully preprocessed and, for simplicity, genes are referred to instead of reporters, but often the analyses are performed using the individual reporter values without combining them into genes. Both in the case of one-channel data and in the case of two-channel reference design data, for each gene there is a value in each replicate of the two conditions to be compared (i.e., two sets of replicate values, one for each condition). For reference design data, each of these numeric values represents the (log) ratio of experimental to reference channel intensities, but since the reference should be constant for all samples, one can proceed with differential expression analysis using these values the same way one-channel intensities are used. Thus, the two cases (reference design and one channel) are handled together. In the case of two-channel direct comparison data, there is one set of replicate values, where each such value is a (log) ratio of the intensities in the two conditions being compared, giving the relative abundance between these two conditions (no reference condition). Since the channels are never separated, special considerations of this case will be made along the way.

The popular approaches to differential expression include different schools of thought: e.g., Family-Wise Error Rate methods (FWER), False Discovery Rate methods (FDR), ANOVA-based methods (e.g., Kerr et al., 2000), and Bayesian approaches. There are also nonstatistical methods that could be sufficient for some purposes, e.g., simply creating ranked lists. Each of these approaches has subcategories; for example, the FDR methods divide into Benjamini and Hochberg based methods and permutation estimation methods. The authors are particularly partial to the FDR permutation estimation methods, which will be the focus here, although the FWER methods will be examined in some detail as well.

Consider the scenario of one-channel assays or two-channel reference design assays. If conducting a single test for the equality of means of a gene's expression level in two different conditions, then the null hypothesis is straightforward. H_0 is set to be the hypothesis that the means in both conditions are equal, and a statistical test of this null hypothesis is performed. It is important to keep in mind an asymmetry in this argument. While some genes can be identified as differentially expressed using this approach, that is *not* to say that genes failing to be included in this category are *not* differentially expressed. More technically, statistical measures of confidence are generally ascribed to the rejection of H_0 , while confidence is not ascribed to the nonrejection. From nonrejection of H_0 , one can only conclude that there is insufficient evidence to reject H_0 , not that there is evidence that H_0 is actually true. Therefore, it is never concluded from the nonrejection of H_0 that a gene is not differentially expressed.

The main issue in the single gene case is in determining an appropriate test. The standard T -test will generally not be appropriate, because of the (often severe) violation of the necessary assumption that the gene's intensity is normally distributed in each condition. Instead a permutation based T -test is generally more appropriate. The downside to a permutation-based approach is the need for sufficient replication. For example, with only three replicates per condition, there are only 20 permutations and so the smallest p -value obtainable would be 0.05, and 5% of the genes on the array will achieve this level, on average, regardless of their differential expression. A general algorithm for computing permutation p -values for the T -statistic is as follows. Suppose there are n and m replicates in the two conditions, respectively. A permutation consists of taking

any n of the $n+m$ values and forming a new group 1, and taking the remaining m as a new group 2. Suppose there are N such permutations (including the trivial permutation which keeps the groups unpermuted). Let T_0 be the value of the T -statistic on the unpermuted data. For any permutation b , let T_b be the value of the T -statistic for that permutation of the data. Suppose that $T_0 \geq T_b$ for M of the N permutations b . Then the permutation p -value is defined to be M/N . This tests for whether the gene is up-regulated in the second condition as compared to the first. If one is interested in down-regulation, one would run the test switching the roles of the conditions. If one is interested in differential expression irrespective of direction, then one should work with the absolute value of the T -statistic instead of the T -statistic itself. In the case of the two-channel direct comparison scenario, for one individual gene the investigator would be testing whether or not the mean log ratio is >0 for up-regulation (and <0 for down-regulation). For this, the one-sample T -statistic is used with the appropriate modifications. For example, given n replicate log ratios, a permutation approach would consist, for each permutation b , in selecting a subset of these values and changing their sign, prior to evaluating the value of the statistic for that permutation.

If extending the single-gene procedure to conduct many tests simultaneously, it is not straightforward to set up an appropriate statistical framework. If it is desired to control the probability of making no mistaken calls at all, then this is called controlling the *Family-Wise Error Rate* (FWER). A popular and powerful algorithm for doing this with microarray data is the *Westfall and Young step-down method* (Westfall and Young, 1993; Ge et al., 2003), which assigns an adjusted p -value to each gene, so that choosing all genes with adjusted p -value less than or equal to α guarantees no false positives with probability of error α . There are two variations of the Westfall and Young step-down method, one called minP and one called maxT. The maxT is preferable when there are very few replicates (less than seven or eight per condition). Specific efficient algorithms for the minP and maxT methods can be found in Ge et al. (2003). Implementations of these algorithms can be found in the *multtest* Bioconductor package.

The FWER is generally too strong of an error control for most investigators' purposes, and the authors use it only when it is necessary to be sure there are no mistakes at all

on the resulting gene list. Most of the time, the investigator is not willing to give up the power necessary to assure no false positives and is willing to tolerate some false calls in order not to miss others. For example, suppose there are 10,000 genes on an array. If an investigator would rather be given a set of 150 genes and be told that 75% of them are likely to be true positives than be given a list of 30 genes and be told that they are all true positives, then the FWER is probably too conservative for that investigator's purposes, and this is why the False Discovery Rate (FDR) was introduced. The FDR measures the (expected) *proportion* of false positives among the set of all predictions. The FDR is not this proportion itself, since the methods available only control the *expected value* of this proportion. Intuitively, the FDR is the (expected) proportion of false differential expression calls in the set of all differential expression calls, but there are a number of ways to define the FDR in order to make this rigorous and practical. The original definition of Benjamini and Hochberg (1995) is the expected value of V/R given that $R > 0$, where V is the number of false calls and R is the total number of calls. The 1995 publication introduced the FDR concept and gave a first method for controlling it, which will be described below. The FDR is generally the most natural and useful method for error rate control; however, the endeavor to control it rigorously has been problematic. Strong heuristic assumptions are always relied upon and come in two forms. The first type are assumptions made about the data. The second are assumptions made about the reliability of estimations and other computational heuristics. Given the landscape of available methods, the authors have come to prefer those that make fewer assumptions about the data. The industry that has derived from the original Benjamini and Hochberg method tends to rely on very strong regularity assumptions about the data, e.g., identically and independently distributed t -statistics or, even worse, normally distributed intensities. These methods also tend to be very conservative and start with gene-by-gene p -values which themselves cannot generally be expected to be reliable, particularly when there are few replicates per condition. As a result, the authors rarely, if ever, use Benjamini and Hochberg based methods. More precisely, to apply Benjamini and Hochberg methods for controlling the FDR, one must first generate gene-by-gene p -values. Unless one has a fair number of replicates (at least 8 to 10),

one should not try to generate permutation p -values, as there will not be enough permutations to produce sufficiently small p -values. In the case of insufficient replicates, the best one can do is to apply some flavor of parametric test, typically a two-sample T -test to test for equality of means or, in the case of direct comparison data, a one-sample T -test to test whether the log ratios have means equal to one. Once gene-by-gene p -values are created, the Benjamini and Hochberg correction is applied as follows. Let P_i be the i -th smallest p -value, so that $P_1 \leq P_2 \leq P_3 \leq \dots$. Let $q_i = \alpha i/g$ where g is the total number of genes and α is the desired false discovery rate. Let k be the maximum i such that $P_i < q_i$. If no such k exists, then all null hypotheses are accepted (e.g., one predicts nothing as differentially expressed). If $k > 1$, then one rejects hypotheses H_1 through H_k , predicting those genes to be differentially expressed. It can be proven, in the case that the p -values for all of the genes are independent, that this set of genes has a false discovery rate no greater than α . Note that this says nothing about the actual proportion of false positives in the predicted set; it only says something about the expected value of this proportion. In any given case the actual value could be very different from the expected value. What is outlined above is the most basic version of the Benjamini and Hochberg algorithm. There has been a lot of work to extend the methods to require less strong assumptions about the data (for example, Yekutieli and Benjamini, 1999; Ge et al., 2003; Storey, 2003).

A totally different approach to the FDR has been taken by SAM (Tusher et al., 2001) and PaGE (Grant et al., 2005; see Internet Resources). These algorithms use permutation distributions to estimate certain quantities, so these methods are referred to as *permutation estimation* methods. There is only heuristic theory to support doing these estimations; however, these methods make far fewer assumptions about the data than Benjamini and Hochberg methods, and are thus more reliable and less conservative. It is important to keep in mind that an FDR is quite different from a p -value. As a result, one is often interested in setting the FDR quite high; even 50% FDR gives meaningful results since it produces (on average) a set of genes highly enriched for differentially expressed genes, while on the array 99.9% of the genes might not be differentially expressed. For the purpose of producing sets of genes highly enriched for differential expression, nobody is too concerned about second decimal place accuracy in the FDR. In fact, if

one sets it to 40% but realizes only 50%, the investigator is almost always more than happy with the outcome. Of course one does not want it to be wildly inaccurate. In the authors' experience, SAM and PaGE achieve a practical level of power and accuracy, but any differential expression result set must be carefully interpreted. Often results do not make sense, for example, if every gene on the array is down-regulated. When this happens, the algorithms show that there is such an effect in the data, but it might be due to poorly normalized or poorly generated data.

The PaGE algorithm is based on using a cutoff C for some statistic S (e.g., the T -statistic) and making differential expression calls for each gene based on whether its statistic exceeds C . Suppose there are R such genes. PaGE then estimates the number of nondifferentially expressed genes that are expected to exceed this cutoff, say V . The ratio V/R is then an estimate for the false discovery rate. The quantity V is estimated using permutations. A permutation of the data means the experiments are randomly assigned to conditions. In most permutations no genes should be differentially expressed, so for any such permutation, any gene with T -statistic greater than C should be a false positive, thus giving an estimate of V . To the extent that some genes will be differentially expressed for some permutations, this should only lead to overestimating V and therefore being conservative. By taking many permutations (all of them, if possible) and computing for each the number of statistics that exceed C , a permutation distribution is obtained for V , from which the expected value of V is estimated. This is repeated for a sequence of values of C , and each gene is assigned the lowest FDR realized for any value of C . These gene-by-gene measures are called q -values in the literature. In this way, taking the set of genes with q -value at most α gives a set of genes with FDR estimated to be no more than α . SAM does something similar, and both SAM and PaGE use a modified form of the T -statistic as statistic. Both SAM and PaGE have a multi-class mode. SAM's multi-class mode tests whether the gene has equal distributions in all conditions. PaGE's multi-class mode compares every condition to one fixed reference condition and produces patterns across the conditions for easy visualization of the results. Both methods have a mode for analyzing direct comparison data. SAM also offers a block permutation mode in case of a block design where a block is defined by a factor (see Experimental

Design). In this mode, only permutations that keep samples in the same block are used, which increases power if there is a significant block effect.

CLASSIFICATION

A very popular use of microarray experiments has been in the area of classification problems. Here it is necessary to distinguish between two main scenarios. In both cases the input consists of a two-dimensional matrix of data, representing the preprocessed results of a microarray experiment monitoring the genes represented on a particular array over a set of assays. These are typically one-channel assays or two-channel reference design assays, so that each assay represents one main sample of interest (in a two-channel reference design assay, this would be the nonreference sample). The matrix has as many rows as there are genes on the array (here for simplicity the authors refer to genes, but in reality one might just work at the reporter or at the transcript level), say m , and as many columns as there are samples of interest, say n . The value in the i -th row and j -th column of this matrix represents the preprocessed signal for gene i in sample j (this would be a relative abundance in a two-channel reference design). For each gene a profile exists consisting of the signals for that gene over all the samples (one row of the data matrix); similarly, for each sample a profile exists consisting of the signals for that sample over all the genes (one column of the data matrix). Thus, both genes and samples can be represented by vectors (the profiles) of values where the vector has length n for a gene and m for a sample. In *UNIT 22.5*, the components of such a vector are termed parameters or dimensions; here they are also referred to as features. In some circumstances it might happen that a gene has missing values in some of the assays (e.g., due to filtering of poor-quality spots in an assay). Methods have been proposed to “impute” these values (e.g., Troyanskaya et al., 2001). A classification problem can refer to classification of genes or classification of samples; in both cases the problem reduces to classification of profiles. There are two types of classification problems: *class discovery* and *class prediction*. The former represents an unsupervised classification, because the classes are not known beforehand, whereas the latter represents a supervised one, because relevant classes are known a priori.

In class discovery, the object is to classify the profiles into groups based on their similarity, therefore “discovering” such groups.

For example, class discovery methods could be employed to determine a molecular classification of samples which are otherwise indistinguishable; to determine groups of genes which are expressed similarly in the samples examined and which therefore might share the same type of function or the same regulatory mechanisms; or to reduce the dimensionality of the collection of gene profiles by first grouping them and then studying only a representative profile per group (e.g., the average gene profile in that group). Class discovery methods could also be used for quality control purposes as illustrated further below.

In class prediction, typically the given profiles are already known to belong to certain classes, which are used as a training set to build a classifier to classify novel profiles into one of these known classes. For example, one might have gene expression profiles for a collection of samples partitioned into patients that do and do not respond well to a given treatment. A predictor will be used for a new patient to decide which of the two treatments is likely to perform well, based on his/her expression profile.

The case of time series data, when one seeks a classification of the gene profiles, is somewhat distinguished as there is additional structure in the data (an ordering) that can be exploited. This is discussed separately below.

There are a variety of free and commercial software packages (including Bioconductor) with implementations of some of the classification methods described below. A rich workbench for data mining, comprising implementations of many classification methods, is WEKA (Frank et al., 2004; Witten and Frank, 2005; see also Internet Resources).

Class Discovery

This is probably the area which has seen the greatest number of publications in the microarray literature, especially in the early days of the technology. If one believes there might be interesting subtypes of samples or genes that are not known ahead of time but can be revealed by their expression signatures, then a class discovery analysis is appropriate. Typically, clustering approaches are utilized for class discovery. Existing clustering techniques have been borrowed and applied to microarray data analysis, and variants, refinements, and new approaches have been proposed. Clustering has played and will continue to play an important role in the microarray field as long as it is used when the question of interest is indeed that of class discovery. It is important, however, not to abuse the method (as has often

been done in the literature) when the question at hand is of a different sort and other types of analyses are more appropriate. In clustering, the degree of similarity between profiles is measured and profiles that are most alike are then put into groups (clusters). These clusters can be well defined, as with a definite partition of the profiles, or they can be loosely defined, as represented by the leaves of a hierarchical tree.

There is a vast literature on clustering applied to microarrays and this topic is also addressed in *UNIT 22.5*. It is difficult to impossible to systematically sort out the landscape of methods, as there are no good benchmarks and there is such a variety of methods that one cannot even put them on equal footing. The best one can hope to do is to settle on a few methods that seem to work. This unit summarizes some of the most commonly used approaches, with the criteria that they are well established, readily available, and relatively convenient, and that they have worked reasonably well for the authors.

There are three main choices to make when doing clustering analyses: (1) what features to use, (2) what similarity measure to use (for similarity-based methods), and (3) what clustering method to use. The first consists in determining which components (features) of the profile vector to use. When clustering genes, all features (samples) are often used. When clustering samples, the full profiles have a very high dimension, as there are typically thousands of genes represented on an array. Therefore, one might use only a portion of them, e.g., those that vary the most across the samples or those that are expressed above background level in at least one of the samples. The idea is to remove those features that are not adding information. In other situations, there might be precise subsets of the genes that one wants to focus on, e.g., based on some prior biological knowledge, and only those are used as features. The results will vary depending on what genes one chooses to cluster on, and in each case a different facet of the relationships in the samples might be brought to light. Liu (2005) provides a survey on feature selection for classification problems. Finally, it should be noted that some subsets should never be used for classification. For example, any subset of genes identified from a differential expression analysis will completely bias any classification algorithm, as those genes will naturally distinguish between the two conditions being compared. In this way, any division of the data could be found by using just those

genes that are most different between the two groups.

The second choice when performing clustering analyses using a similarity-based method, like hierarchical clustering or partitioning (described in detail below), has to do with similarity measures—that is, what criterion to use to decide the degree of similarity between two profiles. The most commonly used measures are the Euclidean distance and the Pearson correlation coefficient (refer to *UNIT 22.5* for further discussion). Results can vary tremendously according to the similarity measure chosen. The authors have found that Euclidean distance is more sensitive to artifacts in the data and generally use standard Pearson Correlation, both to cluster genes and samples, also because they are typically more interested in similarity in shape than in magnitude. Often there is a choice between centered and uncentered correlation. Centered is just standard Pearson Correlation, which is what the authors typically use. Another reasonable measure is Spearman's rank correlation, which first ranks the data in order of intensity, and then applies standard correlation to the profiles of the ranks. This can help put data on equal footing. Typically, as the similarity measure and the subset of genes used in the analysis are varied, a number of different results are obtained. One should avoid trying too many combinations so as to find whatever one is looking for. One will inevitably try several ways and should only believe results which are clear and strong effects.

Regarding the third choice, several clustering methods were available prior to the advent of microarrays (e.g., Hartigan, 1975) and have been utilized for this purpose. Microarray technology has also spurred the development of new methods, e.g., CAST (Ben-Dor et al., 1999), PLAID (Lazzeroni and Owen, 2002), and Gene Shaving (Hastie et al., 2003). The two most frequently used approaches in the microarray literature are hierarchical clustering and partitioning approaches, which are described in more detail below.

Hierarchical clustering, borrowed from phylogeny, is based on building a tree with the profiles at the leaves (see Fig. 19.6.4). This tree is a rigid structure representing varying degrees of similarity. Clusters are not well defined, and for any node in the tree one gets a cluster consisting of the leaves below that node. There are many variations found in the available implementations. For example, bottom-up (agglomerative) implementations (e.g., Eisen et al., 1998) start with each profile alone in a group by itself. Similar groups

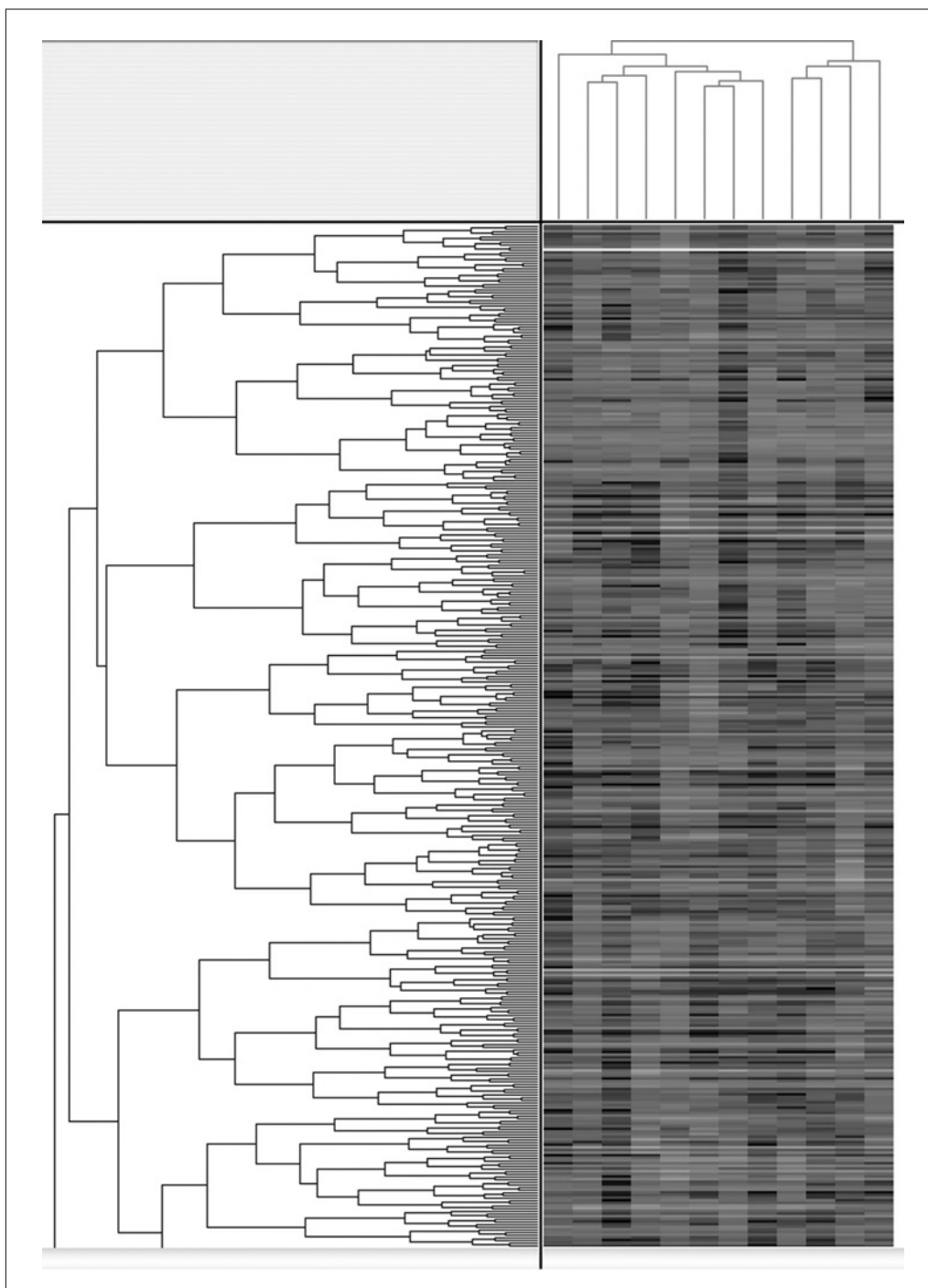


Figure 19.6.4 Result of hierarchical clustering of both genes and samples for twelve one-channel assays, using centered Pearson correlation and the Eisen's TreeView software (<http://rana.lbl.gov/EisenSoftware.htm>). The sample tree is along the top and the gene tree is along the side. Brighter green represents higher signal intensity. For color version of this figure see <http://www.currentprotocols.com>.

are then gradually merged until all profiles are in one group. Agglomerative hierarchical clustering is discussed in *UNIT 22.5*. Top-down (divisive) implementations (e.g., Alon et al., 1999) start with a single group including all profiles. This is split into two groups, which are in turn split, and so on. In either case, the

output of these methods is a tree with the profiles at the leaves.

Partitioning clustering methods (like *k*-means) do not produce trees; instead they are based on the idea of partitioning the profiles into a *prespecified* number of mutually exclusive and exhaustive clusters, then iteratively

reallocating the profiles to clusters until some criterion is met, e.g., sufficiently high within-cluster similarity. Self Organizing Map (SOM; UNIT 22.5) also leads to a partition of the profiles into a prespecified number of clusters, but it imposes a partial structure on the data in that clusters are represented by nodes in some space (typically a plane), and close nodes correspond to “similar clusters.” However this is much less rigid than the structure imposed by hierarchical clustering. An application of SOM to microarray data can be found in Tamayo et al. (1999). Another useful method is Principal Components Analysis (UNIT 22.5), which can be used to project the data into the two or three most important dimensions for easy visualization, which helps reveal the relationships between the profiles.

The literature on this topic is so large that it is impossible to reference all the relevant papers and surveys. The reader is referred to Fasulo (1999), D’haeseleer et al. (2000), and Handl et al. (2005) for surveys on recent work on clustering.

Besides the typical use of clustering to aid in the discovery of novel classes of profiles, another useful application is that of quality control (Coombes et al., 2002, 2003; Grant et al., 2003; Stivers et al., 2003). For an experiment with replicate assays for each of a collection of experimental conditions, upon clustering, profiles of samples from the same experimental condition should be closer to each other than they are to profiles of samples from the other conditions. When this does not occur, it could be a sign of a problem with the experimental procedure or the data annotation. An application of this is illustrated in Coombes et al. (2003), where quality control clustering revealed a problem that was eventually traced back to mis-annotation of the reporters on the array.

One of the biggest issues with clustering is that results can vary significantly when any of the three main choices—feature, similarity measure, and clustering method—is changed. According to the questions and the data at hand, some approaches might be more suitable than others, and this needs to be determined on a case-by-case basis. When the best route to take is not clear, a typical approach is to try a variety of routes and look for those clusters that are consistently picked up, taking the robustness of these as a possible sign of biological relevance. If one has run a number of tree-based methods, it is possible to combine them into a consensus tree which represents the divisions that are found consistently. There

has also been formal work specifically aimed at cluster validation analysis. A survey of this, with a special focus on post-genomic data, can be found in Handl et al. (2005).

Class Prediction

In this type of problem the input is a training set of profiles with a known classification and the goal is to generate a predictor (a.k.a. classifier) to be used to classify novel profiles. A variety of machine learning approaches to build classifiers were in existence prior to the advent of microarrays and have proven useful in the microarray context. Reviews of some these methods as applied to microarray data can be found in Dudoit et al. (2002), Speed (2003), and Statnikov et al. (2005). The issue of feature selection, mentioned in the clustering section, becomes particularly important in this context. Generally, once a predictor is built, one needs an independent data set to validate the predictor. If this is not available, cross-validation techniques are typically used. These consist in iteratively using only a portion of the training set to build a predictor and testing its classification error rate on the rest of the training set.

One of the simplest classifier techniques is *k-nearest neighbors*. This is based on predicting the class of a novel profile by looking at its *k* most similar profiles in the training set and assigning the novel profile to the class to which the majority of these belong. This method requires the choice of a similarity measure. As for the choice of *k*, the typical approach is to try a range of (integer) values and to select that which yields the smallest cross-validation error rate. That is, for each choice of *k* in this range, each profile in the training set is classified using the majority class among its *k* closest profiles, excluding itself. This classification is then compared to the true known class label for that profile, and the error rate for that choice of *k* is computed by looking at the number of profiles which are mis-classified using this approach.

Nearest centroid methods are also fairly intuitive. The simplest application consists in computing the centroid (average profile) for each class in the training set, and then classifying each novel profile into the class whose centroid it is closest to. More sophisticated modifications of the nearest centroid idea include Linear Discriminant Analysis (LDA) methods (e.g., Dabney and Storey, 2005), where the covariance matrices for each class are taken into account. Various proposals have been put forward to deal with the feature

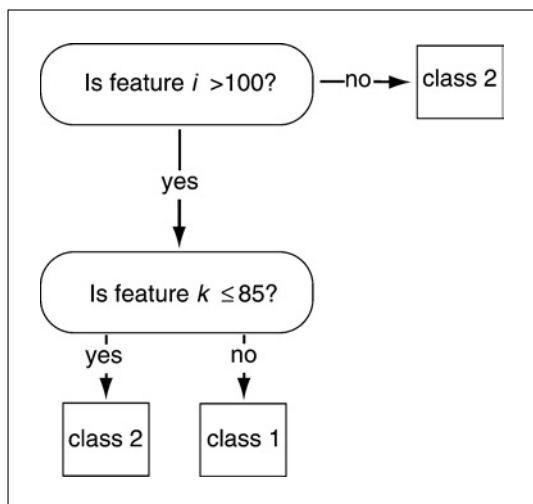


Figure 19.6.5 Simple example of a decision tree.

selection issue for the application of this type of method to microarray experiments where the training sets consist of sample (rather than gene) profiles that have a large number of features (number of genes). One approach in the nearest centroid spirit is Prediction Analysis of Microarray (PAM; Tibshirani et al. 2002, 2003; also see Internet Resources). This uses “shrunk centroids,” which represent a “denoised” version of centroids, down-weighting the contribution of noisy genes to the classifier. Dabney and Storey (2005) propose a different approach which employs a “greedy” algorithm to estimate an optimal subset of features. Greedy algorithms are often used in bioinformatics, when the space of all possible solutions (in this case, subsets of features) to explore is prohibitively large for an exhaustive search. A greedy algorithm does not explore all possible solutions, but always takes the best immediate (local) solution at each step of the search, which can occasionally lead it to miss the optimal global solution.

A very different approach to the construction of a classifier that is frequently utilized in microarray applications is that of Support Vector Machines (SVM; Vapnik, 1998). In its simplest form, SVM has been used for binary classifications (two classes). The idea is to consider each profile as a point in m -dimensional space (where m is the number of features), to map the training profiles into a higher-dimensional space, and then to find a hyperplane in the latter that separates the points corresponding to the two different classes. A novel profile is then classified according to which side of this hyperplane it maps to in this higher-dimensional space. For a review of various extensions of SVMs that have been developed for the case of more than two classes,

see Statnikov et al. (2005). Furlanello et al. (2003) further investigate the issue of feature selection (and feature selection bias) for this type of approach.

Classification trees (Breiman et al., 1984) represent yet another approach to building predictors. The main idea is to build a tree with a rule at every node, typically expressed as a condition on some component(s) of the profile. The leaves of the tree are labeled with class labels. This tree is built using the training set. For a novel profile, one then starts at the root of this tree and moves down. At every node, according to how the condition is satisfied by the appropriate component(s) of the profile, one goes down one branch or another, until a leaf is reached. The label at that leaf corresponds to the classification for that profile (see Fig. 19.6.5). The appeal of this type of class prediction method is that it provides some explicit rules that lead to the classification in each of the classes, thus helping to interpret the class distinctions.

One problem that can occur with decision trees as well as with other machine learning approaches is that the resulting predictors might be unstable, i.e., small perturbations of the training set might lead to very different predictors. This is a sign of *overfitting*, which occurs when the model built for the predictor starts fitting the noise in the training set, thereby leading to a predictor that might perform extremely well on the training set but not as well on the general population of profiles it is supposed to represent. There are various approaches to deal with this within a given machine learning method (for example pruning of decision trees), but an approach that can be applied more generally to any method is to use ensembles of classifiers. The main idea is

to start with the training set and then build a collection of training sets obtained by perturbing the original one, e.g., by resampling. Then a predictor is built using each of these training sets, and the classification of each novel profile is obtained by evaluating it with each predictor and assigning it to the class predicted by majority “vote.” In the case of decision trees, this leads to Random Forests (Breiman, 2001; see also Breiman and Cutler Random Forest Web site in Internet Resources). In this approach, a tree is constructed for each resampling of the original training set (either sampling N profiles with replacement or sampling without replacement about one-third of the profiles each time), yielding a “forest” of decision trees. Each new profile is evaluated according to each of the trees in the forest, and each tree gives a classification, representing a vote for that class. The final classification is the one having the most votes.

As for clustering, there are so many different approaches to class prediction that only some can be illustrated here. The methods presented appear frequently in the microarray literature. The review references cited in this unit contain more pointers and information on other methods.

Time Series

The situation when the experiment consists of a time series, i.e., the samples correspond to different time points (e.g., duration of a certain compound treatment, or different developmental time points in an organism) involves some special considerations as compared to the case of unordered data. In this case, typically one is interested in gene profiles. All of the clustering techniques mentioned above are still applicable to this situation, but one might want to take further advantage of the inherent ordering of the samples according to time and there are approaches that have been designed to this end. Moreover, in many instances one already has in mind some a priori temporal pattern of interest and wants to identify which gene profiles follow those patterns. Another way to take advantage of the temporal component in time series data is to smooth the data across time to eliminate outliers that might be artifacts. A sufficient number of time-points is required to make smoothing reasonable; certainly three or four is not sufficient, while ten or twenty generally is. Smoothing can most easily be done using a local least squares fit to a moving window along the data.

A simple intuitive approach to microarray time series analysis is to first define templates,

e.g., representing an induction or reduction over time or oscillations. Each gene profile is then compared to each of the templates to assess whether or not there is sufficient similarity. As usual, a similarity criterion needs to be established first. An example of an application of this approach can be found in Scearce et al. (2002).

Schliep et al. (2003) propose a more complex approach which utilizes hidden Markov models (HMMs; Rabiner, 1989; Durbin et al., 1999) in a model-based clustering approach. The idea is to start with a seed collection of HMMs, e.g., each representing a prototypical time course pattern. The gene profiles are iteratively assigned to these HMMs and used to refine these HMMs. If at any step an HMM has too few members, it is eliminated; if it has too many members, it is split into two. Thus, the number of clusters is learned from the data. A noise HMM is also used to define a cluster aimed at picking up the noise.

Another example of a model-based approach to time series clustering is that described in Ramoni et al. (2002) and implemented by Cluster Analysis of Gene Expression Dynamics (CAGED; see Internet Resources). This employs a Bayesian approach to determine the number and members of the clusters. In Spellman et al. (1998), time-series microarray experiments are employed to identify cell cycle-regulated genes via a method that combines Fourier transforms with a scaling factor obtained via comparison to gene profiles of genes already known to be regulated by the cell cycle. For a more thorough review of the challenges and techniques geared at analyzing microarray time series data, the reader is referred to Bar-Joseph et al. (2004).

LOOKING AT GENE SETS

One of the advantages of microarray technology is the ability to look in parallel at many transcripts. This enables an investigator to examine sets of genes simultaneously rather than just individual genes. In this respect, it is necessary to distinguish between two kinds of scenarios. The first scenario is that a microarray experiment is conducted in order to discover potentially interesting sets of genes, which has been the focus of the previous section (e.g., genes differentially expressed across two or more conditions or genes that form clusters based on their expression profiles). These sets are typically generated using the analysis methods described in the preceding sections. Once any such potentially interesting set is determined, one typically wants to investigate

its genes further to assess whether or not they share any biologically relevant property such as common pathways, biological processes, or regulators. The second scenario is that an investigator already has in mind particular sets of genes of interest (e.g., genes linked by a particular biological property like a pathway, process, or regulatory mechanism which will define a gene set) or has a set of interesting genes derived from a previous experiment of some sort. In this case, utilizing a microarray study might lead to a better understanding of the molecular profiles of the genes in the set and enable the discovery of any common characteristics of their transcripts with respect to conditions of interest (i.e., the transcriptional profile of each gene set as a whole).

Summarizing, in the first scenario, the starting point is a microarray experiment and the goal is to elucidate common biological trends in gene sets constructed by analyzing that experiment. In what follows, this scenario is referred to as the *a posteriori gene set scenario*, since the gene sets are derived from the microarray experiment. In the second scenario, the starting point is one or more predefined gene sets and the goal is to evaluate their transcriptional behavior utilizing a microarray experiment. This scenario is referred to as the *a priori gene set scenario*, since the gene sets are defined independently of the microarray experiment. Some approaches that have been utilized in each of these scenarios are described below.

A Posteriori Gene Set Scenario

Given a gene set resulting from analyses carried out on a microarray experiment, a typical question of interest is whether or not this set shows enrichment in a predefined biological property. For example, one might be interested in finding whether or not there is a particular biological process or pathway that is common to a large proportion of the genes in the set. Rephrased more rigorously, this question becomes: is there statistical evidence that the proportion of genes in this set having the specified biological property is larger than expected if the set had been chosen at random from the gene population? One of the most popular and powerful associations used for this purpose are Gene Ontology terms. The Gene Ontology (GO) Project (see Internet Resources) provides structured controlled vocabularies for three categories describing gene products: biological processes, cellular components, and molecular functions. Moreover, annotation of genes with terms from these on-

tologies (“Mappings to GO”) is also provided. Given the array used in the microarray study from which a gene set has been derived, and given a GO term of interest (in one of the three GO categories), typical approaches take the genes monitored on the array as a surrogate of the genome and use the GO mapping to determine the fraction that are annotated with that GO term. Then they compare to the latter the fraction of genes in the selected gene set which map to that term, using the hypergeometric distribution (e.g., see section 2.7.3 of Tamhane and Dunlop, 2000) to identify whether or not that GO term is overrepresented in the gene set. Since, for a given gene set of interest, this statistical test is repeated for each GO term for a biological process, it is then necessary to adjust the individual *p*-value obtained from each test for multiple testing (see Expression and Differential Expression for a discussion of multiple testing issues). Of course, if the biological theme of interest is from another category (e.g., KEGG pathways; see Internet Resources), the same approach is applicable. The Expression Analysis Systematic Explorer (EASE), available at the DAVID Web site (see Internet Resources) from the NIAID, is an example of software implementing this kind of analysis.

Although the idea behind this kind of approach is simple, there are some complicating issues. Namely, the *p*-values derived from a hypergeometric distribution make the assumption that the observations are independent, but genes are not independent of each other. Moreover, a set derived from a microarray experiment typically consists of reporters which represent genes, but a given gene might be represented more than once on the array. The latter problem is easier to deal with. For example, EASE does not simply use the identifiers (typically GenBank accessions; see Internet Resources) supplied by the user, but collapses multiple representatives of the same Entrez Gene (see Internet Resources) to avoid the situation in which a gene with multiple accessions receives more than one vote in the overrepresentation analysis. However, the problem of gene dependencies still remains.

Besides EASE, there are several other tools that address the same kind of question, many of them implementing an analogous approach. A summary of tools specifically geared to the biological theme of GO categories as it pertains to an *a posteriori* gene set from microarray experiments, as well as links to the respective Web sites, is available from the Gene Ontology Project Web site (GO tools, microarray).

Another type of biological theme that one might be interested in discovering for a gene set derived from microarray analysis is that of regulatory modules. Namely, given a set of genes that a microarray study indicates might share common regulators (e.g., because they are all differentially expressed under certain conditions or because they have similar expression profiles across multiple conditions), can one identify combinations of regulatory elements (e.g., transcription factor binding sites, TFBS) which appear to be enriched in the promoter regions of these genes? Given the sequences of interest, one might either look for occurrences of known TFBSs (e.g., from TRANSFAC or JASPAR; see Internet Resources) or one might try to find possibly novel motifs. Tools aimed at motif discovery in a group of sequences include MEME (Bailey and Elkan, 1994), AlignAce, Weeder, and the ELPH Gibbs Sampler (see Internet Resources for the corresponding Web sites), which use different algorithmic approaches. There are also approaches that simultaneously combine the discovery of relevant gene sets and of their regulatory modules (e.g., Segal et al., 2003). These citations are not comprehensive as there is a large amount of literature and ongoing research in the area of gene regulation which leverages on the integration of gene expression microarray data with sequence data, as well as other data types (e.g., ChIP-chip data).

A Priori Gene Set Scenario

In this scenario, the starting point consists of a collection of predefined gene sets and a microarray experiment, and the goal is to overlay each such set over the expression data to further characterize these sets in terms of their expression profile as *a whole set*. Here each gene set might be defined based on an ontology (e.g., all genes mapping to a given GO term), a pathway (e.g., all genes in a given KEGG pathway), or on an investigator's prior knowledge and/or experiments.

A first kind of analysis in this scenario is that provided by the public software GenMAPP (see Internet Resources). A MAPP is a grouping of genes defined by some organizing principle. Quoting from <http://www.genmapp.org/GenMAPPHelp2/GenMAPP.htm>:

“Examples of the types of MAPPs that can be represented in GenMAPP are metabolic pathways, signal transduction cascades, sub cellular locations, gene families, or lists of genes associated with Gene Ontology categories. Each MAPP contains gene objects,

which represent biological genes or gene products.”

One of the features of GenMAPP is that of enabling a user to color-code genes within a MAPP according to user-selected criteria on their expression in a microarray experiment, thus providing the means to identify interesting expression trends in that MAPP.

The Gene Set Enrichment Analysis (GSEA; see Internet Resources), first presented in Mootha et al. (2003), is an approach aimed at identifying which gene sets from a given collection show enrichment in genes which are at the top (or bottom) of a ranked list, e.g., where the ranking is based on differential expression between conditions. This method utilizes a score based on the Kolmogorov-Smirnov statistic (see, e.g., Hollander and Wolfe, 1999) and incorporates multiple testing corrections. Details on the algorithm are available in the documentation at the GSEA Web site. Also available at this Web site is MSigDB, a Molecular Signature Database of gene sets.

The general philosophy behind the approaches outlined above is that power can be gained when looking at sets of genes rather than at individual genes. The reason for this is that some characteristics of individual genes might be too weak to overcome noise and to be detected by a statistical analysis if the gene is treated as an individual, but might be identified when the gene is considered as member of a group of genes sharing a common biological theme, since its signal, compounded with those of other genes, might lead to an overall signal whose significance is detectable.

DATABASES

Microarray data are sufficiently large and complex that a database is often required for data management. The bulk of the data will be the intensity values associated with reporters on the array and annotation of the sequences represented by those reporters. Just as important are annotations describing the experimental plan, materials, and protocols used. If the microarray study is to be part of a publication, many journals require that the study be deposited at a public repository adhering to established standards. Topics to be covered in this section are choosing a database system, annotation of microarray data, and submission of data to microarray repositories.

Choosing a Database System

Database systems consist of a schema describing how the data are organized and

software for loading data as well as retrieving data for generating files, reports, or viewing online. Many database systems are currently available covering a wide range of needs. These include freely available systems as well as commercial ones. To help choose a database system, a comparison of microarray databases was published by Gardiner-Garden and Littlejohn (2001) and used as the basis for an online table at the Stanford Microarray Database (SMD; see Internet Resources). Sherlock and Ball (2005) have published a more recent comparison on a few selected freely available systems. Issues to consider include cost (including support), features, and coverage of required information. A database requires a database management system and the most popular ones are relational database management systems (RDMS) such as Oracle, Sybase, Windows SQLServer, IBM DB-2, PostgreSQL, or MySQL. Commercial systems such as Oracle require licenses which can be expensive (although some academic institutions have licenses that cover everyone). These have the advantage of being very powerful, well supported, and well tested. Free RDMSs such as PostgreSQL and MySQL (see Internet Resources) are generally simpler and easier to manage but may not work with the chosen microarray database system. Part of the cost to consider is the staff needed to maintain the database as well as the hardware needed to install it. Generally speaking, choices will be driven by the amount of data anticipated and the expertise available as well as which RDMS is supported by the schema chosen. Any schema used should be *MIAME-compliant*, which means that all the types of information on the MIAME checklist accessible from the MGED Society Web site (see Internet Resources) can be captured in the chosen schema. Schemas vary from those that simply focus on manipulating and viewing microarray results (e.g., SMD), to those that cover the steps required to make a microarray (i.e., include laboratory information management system or LIMS information, e.g., BASE; see Internet Resources), to those covering additional areas outside of microarrays (i.e., include genomics and other functional genomics areas, e.g., RAD as part of GUS; see Internet Resources). Indeed, the reason for the availability of different microarray database systems is that these have been built to address the different needs of groups. Features to consider are tools to load microarray data, annotate the experiments, enable analysis, and export data in a standard format such as MAGE-ML. A

MAGE software tool kit (MAGEstk) to generate MAGE-ML is available through the MAGE working group page at the MGED Society Web site.

Annotation and Data Loading

The intensity values obtained in a microarray experiment are context-dependent; therefore, to interpret the meaning of the intensity value one must know the state of the cells from which the nucleic acid was obtained as well as how the nucleic acid was labeled, hybridized, and so on. While there are many factors both biological and technical that can influence those values, a minimum information about a microarray experiment (MIAME) standard has been established to cover the basics (Brazma et al., 2001). Microarray experiments need to be planned carefully from beginning to end. Although this is true for any experiment, it is perhaps more important here due to the time and cost involved in performing microarray experiments. The MIAME checklist covers experimental design, samples, measurements, normalization controls, and array design. These topics should be considered prior to starting a series of experiments rather than just used to describe what was done at the end. For example, the choice of experimental design will depend on the available samples and, together with normalization controls, determine what kinds of analysis will be most effective.

Capturing the annotation in a structured way for storage in a database is usually done with Web-based tools packaged with the microarray database system employed or custom built for a particular project (or institutional core). For example, the RAD Study Annotator (Manduchi et al., 2004) provides a Web-based tool to annotate biomaterials, protocols, experimental designs and factors, and other MIAME-related information. Array designs and data are loaded using software written for RAD. Wherever possible, the MGED Ontology is used primarily in the form of drop-down menus to choose terms for annotation. This same approach is taken by a number of other microarray database systems, several of which are listed in a recent paper on the MGED Ontology (Whetzel et al., 2006). The use of common terminology by different groups facilitates data retrieval, sharing, and analysis by making it clear when experimental factors and sample characteristics are comparable. It should be noted that for many projects the MIAME checklist does not go far enough to provide sufficient annotation for certain types of

experiments such as those for toxicogenomics. Extension of MIAME have been developed that can be used as guides, and these can be found at the MIAME working group page on the MGED Society Web site.

Submission of Data to Microarray Repositories

Public repositories for microarray data have been established. The primary repositories are ArrayExpress, developed and maintained by the European Bioinformatics Institute, and the Gene Expression Omnibus (GEO), developed and maintained by the National Center for Biotechnology Information (see Internet Resources). Both ArrayExpress and GEO are able to accept submissions in the standard MAGE-ML format and therefore having a database system that generates MAGE-ML addresses this step. Use of MAGE-ML is not mandatory. GEO also has established its own format, Simple Omnibus Format in Text (SOFT). ArrayExpress provides a tool, MIAMExpress, that allows direct submissions. MIAMExpress provides a Web-based tool for annotating experiments using the MGED Ontology as described above. Both repositories do request that submission meet the MIAME standard.

Despite the tools and standards available, it is still possible to submit poorly annotated microarray experiments. The consequences of poor annotation are the failure to live up to the spirit of providing sufficient information to clearly understand the data and the diminished chances of the experimental results being of further use. Therefore, every reasonable effort should be made to fully annotate submissions.

CONCLUSIONS

The landscape of microarray platforms and analysis methods might be overwhelming, but there is no need to master every platform and every method of analysis. Generally a laboratory will use one platform and will assemble a set of protocols that works for them. Optimally, this should be done with the aid of a bioinformatics/statistics consultant. This being said, a lot of mileage can be obtained from just a few tools. For example, a major portion of the information that is desired can be obtained, for two-channel assays, using loess normalization, SAM, or PaGE for differential expression analysis and hierarchical clustering for identifying the major patterns. Analogously, one can use Affymetrix arrays pro-

cessed with gcRMA and then analyzed with SAM or PaGE and via clustering.

To conclude, the following steps are recommended as a general approach for using microarrays (Fig. 19.6.1):

1. Consult a statistician and/or bioinformaticist when designing your experiment. Usually, microarray cores can provide this service or at least recommend someone who can.

- a. Determine what kinds of analysis will be performed and then decide on the number and types of replicates needed to do the experiment.

- b. Make sure that confounding factors like date of hybridization are addressed (i.e., don't process all the controls one day and all the experimental samples another day).

2. Perform the experiment and start collecting information required for the MIAME checklist.

- a. Get the quantified data files (e.g., CEL files for Affymetrix).

- b. Get other files that have MIAME information (e.g., EXP for Affymetrix that provide hybridization date, protocol, array).

3. Preprocess and normalize the data (e.g., use gcRMA for Affymetrix, loess for two-channel assays).

- a. Assess the assays by inspecting images and generating diagnostic plots.

- b. Perform quality control analysis looking for hybridizations that are outliers among replicates in terms of how much they need to be adjusted for normalization and number of signals above background (e.g., P calls for Affymetrix).

- c. Apply quality control clustering to see whether replicates cluster together as expected.

- d. Filter out reporters to be discarded and normalize.

4. Analyze the data (e.g., SAM or PaGE for differential expression, hierarchical clustering).

- a. Look for shared characteristics of differentially expressed genes and genes in a cluster (e.g., look for overenrichment of GO terms using EASE).

- b. Validate differences seen (e.g., qRT-PCR for key genes identified as differentially expressed).

5. Load the data and experiment (MIAME) annotation into a database.

6. Publish your study and submit your microarray data and annotation to ArrayExpress or GEO.

ACKNOWLEDGMENTS

The authors thank John Tobias for his critical reading of this unit and comments, and Mitchell Guttman for generating Figure 19.6.4.

LITERATURE CITED

- Allison, D.B., Cui, X., Page, G.P., and Sabripour, M. 2006. Microarray data analysis: From disarray to consolidation and consensus. *Nat. Rev. Genet.* 7:55-65.
- Alon, U., Barkai, N., Notterman, D.A., Gish, K., Ybarra, S., Mack, D., and Levine, A.J. 1999. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc. Natl. Acad. Sci. U.S.A.* 96:6745-6750.
- Bailey, T.B. and Elkan, C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, pp. 28-36. AAAI Press, Menlo Park, Calif.
- Barash, Y., Dehan, E., Krupsky, M., Franklin, W., Geraci, M., Friedman, N., and Kaminski, N. 2004. Comparative analysis of algorithms for signal quantitation from oligonucleotide microarrays. *Bioinformatics* 20:839-846.
- Bar-Joseph, Z. 2004. Analyzing time series gene expression data. *Bioinformatics* 20:2493-2503.
- Ben-Dor, A., Shamir, R., and Yakhini, Z. 1999. Clustering gene expression patterns. *J. Comput. Biol.* 6:281-297.
- Benjamini, Y. and Hochberg, Y. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B Met.* 57:289-300.
- Bolstad, B.M., Irizarry, R.A., ~strand, M., and Speed, T.P. 2003. A comparison of normalization methods for high-density oligonucleotide array data based on variance and bias. *Bioinformatics* 19:185-193.
- Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., Gaasterland, T., Glenisson, P., Holstege, F.C.P., Kim, I.F., Markowitz, V., Matese, J.C., Parkinson, H., Robinson, A., Sarkans, U., Schulze-Kremer, S., Stewart, J., Taylor, R., Vilo, J., and Vingron, M. 2001. Minimum information about a microarray experiment (MIAME): Toward standards for microarray data. *Nat. Genet.* 29:365-371.
- Breiman L., 2001. Random forests. *Machine Learning* 45:5-32.
- Breiman, L., Friedman, J.H., Olshen, R., and Stone, C.J. 1984. Classification and Regression Trees. The Wadsworth Statistics/Probability Series. Wadsworth International Group, Belmont, Calif.
- Choe, S.E., Boutros, M., Michelson, A.M., Church, G.M., and Halfon, M.S. 2005. Preferred analysis methods for Affymetrix GeneChips revealed by a wholly defined control dataset. *Genome Biology* 6:R16.
- Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.* 74:829-836.
- Cleveland, W.S. and Devlin, S.J. 1988. Locally weighted regression: An approach to regression analysis by local fitting. *J. Am. Stat. Assoc.* 83:596-610.
- Coombes, K.R., Highsmith, W.E., Krogmann, T.A., Baggerly, K.A., Stivers, D.N., and Abruzzo, L.V. 2002. Identifying and quantifying sources of variation in microarray data using high-density cDNA membrane arrays. *J. Comp. Biol.* 9:655-669.
- Coombes, K.R., Wang, J., and Abruzzo, L.V. 2003. Monitoring the quality of microarray experiments. In *Methods of Microarray Data Analysis III* (K.F. Johnson and S.K. Lin, eds.) pp. 25-40. Kluwer Academic Publishers, Boston.
- Cope, L.M., Irizarry, R.A., Jaffee, H., Wu, Z., and Speed, T.P. 2003. A benchmark for Affymetrix GeneChip expression measures. *Bioinformatics* 20:323-331.
- Dabney, A.R. and Storey, J.D. 2005. Optimal Feature Selection for Nearest Centroid Classifiers, With Applications to Gene Expression Microarrays. UW Biostatistics Working Paper Series. Working Paper 267. The Berkeley Electronic Press, Berkeley, Calif.
- D'haeseleer, P., Liang, S., and Somogyi, R. 2000. Genetic network inference: From co-expression clustering to reverse engineering. *Bioinformatics* 16:707-726.
- Durbin, R., Eddy, S.R., Krogh, A., and Mitchison, G. 1999. Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. Cambridge University Press, Cambridge, U.K.
- Dudoit, S., Fridlyand, J., and Speed, T.P. 2002. Comparison of discrimination methods for the classification of tumors using gene expression data. *J. Am. Stat. Assoc.* 97:77-87.
- Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* 95:14863-14868.
- Ewens, W.J. and Grant, G.R. 2005. Statistical Methods in Bioinformatics: An Introduction, 2nd ed. Springer-Verlag, New York.
- Fasulo, D. 1999. An analysis of recent work on clustering algorithms. *Technical Report TR 0103-02*, University of Washington, Department of Computer Science & Engineering, Seattle.
- Frank, E., Hall, M., Trigg, L., Holmes, G., and Witten, I.H. 2004. Data mining in bioinformatics using WEKA. *Bioinformatics* 20:2479-2481.
- Furlanello, C., Serafini, M., Merler, S., and Jurman, G. 2003. Entropy-based gene ranking without selection bias for the predictive classification of microarray data. *BMC Bioinformatics* 4:54.
- Gardiner-Garden, M. and Littlejohn, T.G. 2001. A comparison of microarray databases. *Brief Bioinform.* 2:143-158.

- Ge, Y., Dudoit S., and Speed, T.P. 2003. Resampling-based multiple testing for microarray data hypothesis. *Test* 12:1-44.
- Grant, G.R., Manduchi, E., Pizarro, A., and Stoeckert, C.J. Jr. 2003. Maintaining data integrity in microarray data management. *Biotechnol. Bioeng.* 84:795-800.
- Grant, G.R., Liu, J., and Stoeckert, C.J. Jr., 2005. A practical false discovery rate approach to identifying patterns of differential expression in microarray data. *Bioinformatics* 21:2684-2690.
- Handl, J., Knowles, J., and Kell, D.B. 2005. Computational cluster validation in post genomic data analysis. *Bioinformatics* 21:3201-3212.
- Hartigan, J. 1975. Clustering Algorithms. Wiley, Chichester, U.K.
- Hastie, T., Tibshirani, R., Eisen, M.B., Alizadeh, A., Levy, R., Staudt, L., Chan, W.C., Botstein, D., and Brown, P. 2003. Gene shaving as a method for identifying distinct sets of genes with similar expression patterns. *Genome Biol.* 1:research0003.1-research0003.21.
- Hollander, M. and Wolfe, D.A. 1999. Nonparametric Statistical Methods, 2nd ed. Wiley Interscience, New York.
- Huber, W., von Heydebreck, A., Sültmann, H., Poustka, A., and Vingron, M. 2002. Variance stabilization applied to microarray data calibration and to quantification of differential expression. *Bioinformatics* 18:S96-S104.
- Irizarry, R.A., Wu, Z., and Jaffee, H.A. 2006. Comparison of affymetrix GeneChip expression measures. *Bioinformatics* 22:789-794.
- Kerr, K. and Churchill, G.A. 2001. Experimental design for gene expression microarrays. *Bio-statistics* 2:183-202.
- Kerr, M.K., Martin, M., and Churchill, G.A. 2000. Analysis of variance for gene expression microarray data. *J. Comput. Biol.* 7:819-837.
- Lazzeroni, L.C. and Owen, A. 2002. Plaid models for gene expression data. *Statistica Sinica* 12:61-86.
- Liu, H. 2005. Evolving feature selection. *IEEE Intelligent Systems* 20:64-76.
- Manduchi, E., Grant, G.R., He, H., Liu, J., Mailman, M.D., Pizarro, A.D., Whetzel, P.L., and Stoeckert, C.J. Jr. 2004. RAD and the RAD Study-Annotator: An approach to collection, organization and exchange of all relevant information for high-throughput gene expression studies. *Bioinformatics* 20:452-459.
- Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., Houstis, N., Daly, M., Patterson, N., Mesirov, J.P., Golub, T.R., Tamayo, P., Spiegelman, B., Lander, E.S., Hirschhorn, J.N., Altshuler, D., and Group, L.C. 2003. PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* 34:267-273.
- Quackenbush, J. 2002. Microarray data normalization and transformation. *Nat. Genet.* 32:496-501.
- Rabiner, L.A. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE Inst. Electr. Electron. Eng.* 77:257-286.
- Rajagopalan, D. 2003. A comparison of statistical methods for analysis of high density oligonucleotide array data. *Bioinformatics* 19:1469-1476.
- Ramoni, M., Sebastiani, P., and Kohane, I. 2002. Cluster analysis of gene expression dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 99:9121-9126.
- Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., and Quackenbush, J. 2003. TM4: A free, open-source system for microarray data management and analysis. *Biotechniques* 34:374-378.
- Scearce, L.M., Brestelli, J.E., McWeeney, S.K., Lee, C.S., Mazzarelli, J., Pinney, D.F., Pizarro, A., Stoeckert, C.J. Jr, Clifton, S.W., Permutt, M.A., Brown, J., Melton, D.A., and Kaestner, K.H. 2002. Functional genomics of the endocrine pancreas. The pancreas clone set and PanChip, new resources for diabetes research. *Diabetes* 51:1997-2004.
- Schliep, A., Schönhuth, A., and Steinhoff, C. 2003. Using hidden Markov models to analyze gene expression time course data. *Bioinformatics* 19:i255-i263.
- Segal, E., Yelensky, R., and Koller, D. 2003. Genome-wide discovery of transcriptional modules from DNA sequence and gene expression. *Bioinformatics* 19:273-282.
- Sherlock, G. and Ball, C.A. 2005. Storage and retrieval of microarray data and open source microarray database software. *Mol. Biotechnol.* 30:239-251.
- Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., and Futcher, B. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* 9:3273-3297.
- Speed T.P. (ed.) 2003. Statistical Analysis of Microarray Gene Expression Data. Chapman & Hall/CRC, Boca Raton, Fla.
- Statnikov, A., Aliferis, C.F., Tsamardinos, I., Hardin, D., and Levy, S. 2005. A comprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis. *Bioinformatics* 21:631-643.
- Steibel, J.P. and Rosa, G.J.M. 2005. On reference designs for microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 4(1):Article 36. <http://www.bepress.com/sagmb/vol4/iss1/art36>.

- Stivers, D., Wang, J., Rosner G., and Coombes, K. 2003. Organ specific differences in gene expression and unigene annotations describing source material. *In Methods of Microarray Data Analysis III* (K.F. Johnson and S.K. Lin, eds.) pp. 59-72. Kluwer Academic Publishers, Boston.
- Storey, J.D. 2003. The positive false discovery rate: A Bayesian interpretation and the q -value. *Ann. Stat.* 31:2013-2035.
- Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E., and Golub, T. 1999. Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. U.S.A.* 96:2907-2912.
- Tamhane, A.C. and Dunlop, D.D. 2000. *Statistics and Data Analysis*. Prentice Hall, Upper Saddle River, N.J.
- Tibshirani, R., Hastie, T., Narasimhan, B., and Chu, G. 2002. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 99:6567-6572.
- Tibshirani, R., Hastie, T., Narasimhan, B., and Chu, G. 2003. Class prediction by nearest shrunken centroids, with applications to DNA microarrays. *Stat. Sci.* 18:104-117.
- Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., and Altman, R.B. 2001. Missing value estimation methods for DNA microarrays. *Bioinformatics* 17:520-525.
- Tusher, V.G., Tibshirani, R., and Chu, G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U.S.A.* 98:5116-5121.
- Vapnik, V. 1998. *Statistical Learning Theory*. Wiley Interscience, New York.
- Westfall, P.H. and Young, S.S. 1993. *Resampling-based multiple testing*. Wiley InterScience, New York.
- Whetzel, P.L., Parkinson, H., Causton, H.C., Fan, L., Fostel, J., Fragoso, G., Game, L., Heiskanen, M., Morrison, N., Rocca-Serra, P., Sansone, S.A., Taylor, C., White, J., and Stoeckert, C.J. Jr. 2006. The MGED Ontology: A resource for semantics-based description of microarray experiments. *Bioinformatics*. 22:866-873.
- Witten, I.H., and Frank, E. 2005. *Data mining: Practical machine learning tools and techniques*, 2nd Edition. Morgan Kaufmann, San Francisco.
- Yang, Y.H. and Speed, T.P. 2002. Design issues for cDNA microarray experiments. *Nat. Rev. Genet.* 3:279-588.
- Yang, Y.H., Buckley, M.J., Dudoit, S., and Speed, T.P. 2002a. Comparison of methods for image analysis on cDNA microarray data. *J. Computat. Graph. Stat.* 11:108-136.
- Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J., and Speed, T.P. 2002b. Normalization for cDNA microarray data: A robust composite method addressing single and multiple slide systematic variation. *Nucl. Acids Res.* 30:e15.
- Yekutieli, D. and Benjamini, Y. 1999. Resampling-based false discovery rate controlling multiple test procedures correlated test statistics. *J. Stat. Plan. Inference* 82:171-196.

INTERNET RESOURCES

- http://www.affymetrix.com/support/technical/technotes/statistical_reference_guide.pdf
Affymetrix statistical algorithms reference guide (MAS 5.0).
- http://www.affymetrix.com/support/downloads/manuals/data_analysis_fundamentals_manual.pdf
Affymetrix Data Analysis Fundamentals Manual.
- <http://www.ebi.ac.uk/arrayexpress/>
The ArrayExpress repository
- <http://atlas.med.harvard.edu/>
The AlignACE Web site.
- http://www1.amershambiosciences.com/aptrix/upp01077.nsf/Content/microarrays_analysis
ArrayVision image analysis software.
- <http://base.thep.lu.se/>
The BASE - BioArray Software Environment Web site.
- <http://www.gene-regulation.com/pub/databases.html>
The BIOBASE Web site, which includes TRANSFAC.
- <http://www.bioconductor.org>
The Bioconductor project Web site.
- <http://www.genomethods.org/caged>
The CAGED Web site.
- <http://david.abcc.ncifcrf.gov>
DAVID Web site at NIAID, where EASE is also available.
- <http://www.dchip.org>
dChip software, with links to references and tutorials.
- <http://www.cbcb.umd.edu/software/ELPH>
The ELPH Gibbs Sampler Web site.
- <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>
The Entrez Gene Web site.
- <http://www.ncbi.nlm.nih.gov/Genbank>
The GenBank Web site.
- <http://www.ncbi.nlm.nih.gov/geo/>
The Gene Expression Omnibus (GEO) repository.
- <http://www.geneontology.org/>
The Gene Ontology Project Web site.
- http://www.moleculardevices.com/pages/software/gn_genepix_pro.html
GenePix image analysis software.
- <http://www.broad.mit.edu/gsea>
The Gene Set Enrichment Analysis (GSEA) Web site.

- <http://www.genmapp.org/>
The GenMAPP project Web site.
- <http://www.gusdb.org/>
The GUS (Genomics Unified Schema) Web site containing the RAD software.
- <http://mordor.cgb.ki.se/cgi-bin/jaspar2005/jaspar.db.pl>
The JASPAR Web site.
- <http://www.genome.jp/kegg/pathway.html>
The KEGG Pathway database.
- <http://meme.sdsc.edu>
The MEME Web site.
- <http://www.mged.org>
The MGED Web site with links for MIAME, MAGE, and MGED Ontology.
- <http://www.ebi.ac.uk/miamexpress/>
The MIAMExpress annotation and submission tool.
- <http://www.mysql.com/>
The MySQL open source database Web site.
- <http://www.itl.nist.gov/div898/handbook/pmd/section1/pmd144.htm>
NIST handbook, section on LOESS.
- <http://www.cbil.upenn.edu/PaGE>
The PaGE Web site.
- <http://www-stat.stanford.edu/~tibs/PAM/>
The PAM Web site.
- <http://www.postgresql.org/>
The PostgreSQL open source database Web site.
- http://stat-www.berkeley.edu/users/breiman/RandomForests/cc_home.htm
L. Breiman and A. Cutler Web site on Random Forests.
- <http://rmaexpress.bmbolstad.com/>
RMAExpress Web site, which includes relevant references.
- <http://www-stat.stanford.edu/~tibs/SAM/>
The SAM Web site
- <http://rana.lbl.gov/manuals/ScanAlyzeDoc.pdf>
ScanAlyze image analysis software.
- <http://experimental.act.cmis.csiro.au/Spot/index.php>
Spot image analysis software.
- <http://genome-www5.stanford.edu/download/>
The Stanford Microarray Database (SMD) software download site.
- <http://genome-www5.stanford.edu/resources/databases.shtml>
The Stanford Microarray Database (SMD).
- <http://www.stratagene.com>
The Stratagene Web site.
- <http://www.tm4.org/>
TM4: a package of open source software for microarray analysis, comprising MADAM, SpotFinder, MIDAS, and MeV.
- <http://www.stat.berkeley.edu/users/terry/zarray/TechReport/589.pdf>
Technical report on normalization for two-channel arrays by Y.H. Yang, S. Dudoit, P. Luu, and T.P. Speed.
- <http://159.149.109.16:8080/weederWeb/>
The Weeder Web site.
- <http://www.cs.waikato.ac.nz/ml/weka/>
The WEKA web resource.

Contributed by Gregory R. Grant,
Elisabetta Manduchi, and Christian J.
Stoeckert, Jr.
University of Pennsylvania School of
Medicine
Philadelphia, Pennsylvania

CHAPTER 20

Analysis of Protein Interactions

INTRODUCTION

Most of the work of living organisms is performed by proteins. Proteins do their work by acting on other macromolecules: nucleic acids, carbohydrates, lipids, and, especially, other proteins. Inside cells, protein-protein interactions are instrumental in enzymatic actions upon protein substrates and in the fleeting and lasting protein assemblies that govern signal transduction, cell division, DNA replication, and transcription initiation. Outside cells, protein interactions allow cells to talk with each other: ligands expressed on the cell surface often bind protein receptors expressed on adjacent cells, and secreted protein ligands bind receptors on distant cells. Techniques to detect and study protein-protein interactions have become steadily more easy to use in recent years. This chapter lays them out.

BASIC CONCEPTS

Most of the work of living organisms is performed by proteins. Proteins do their work by acting on other macromolecules: nucleic acids, carbohydrates, lipids, and, especially, other proteins. Inside cells, protein-protein interactions are instrumental in enzymatic actions upon protein substrates and in the fleeting and lasting protein assemblies that govern signal transduction, cell division, DNA replication, and transcription initiation. Outside cells, protein interactions allow cells to talk with each other: ligands expressed on the cell surface often bind protein receptors expressed on adjacent cells, and secreted protein ligands bind receptors on distant cells. Techniques to detect and study protein-protein interactions have become steadily more easy to use in recent years. This unit, and those that follow, lay them out.

Qualitative statements about protein-protein interactions inform much of the current biological literature. This unit presents the basics for understanding such qualitative statements—i.e., the quantitative concepts by which these can be evaluated. Colloquially, protein interactions are often referred to as either **stable** or **transient**. The statement that a certain interaction is **stable** is typically used when the interacting proteins form a complex. Membership in a protein complex is most commonly defined operationally—based on detection of the protein after its coprecipitation with a reagent specific for another member of the complex—but is sometimes defined based on a strong and biologically plausible interaction in a two-hybrid experiment. Proteins that form such complexes with one another are often referred to as **partners**. By contrast, the term **transient** is used to denote, e.g., interactions between an enzyme and a protein substrate.

As the preceding example suggests, there is no particular correlation between the stability of a protein interaction and its biological significance. It is also worth noting that there are exceptions to commonly held assumptions about stable and transient interactions: some interactions observed in coprecipitation experiments have no biological validity, whereas some enzyme-substrate interactions last for minutes. For these reasons, it is often useful to think of protein interactions in more formal terms. The strength, or **affinity**, of an interaction between two proteins is described by **equilibrium parameters**. The speed at which an interaction occurs, and at which interacting proteins come apart, is given by **kinetic parameters**.

Table 20.0.1 Dissociation Constants for Some Biologically Significant Interactions

Interaction	K_d
Streptavidin-biotin binding	10^{-14} M
Antibody-antigen interaction with good antibody	10^{-8} to 10^{-10} M
DNA binding protein with specific site	10^{-8} to 10^{-10} M
Antibody-antigen interaction with weak antibody	10^{-6} M
Enzyme-substrate interactions	10^{-4} to 10^{-10} M
Cooperative interaction between DNA-bound phage repressors	10^{-4} M
Cooperative interaction between phage repressor and <i>E. coli</i> RNA polymerase	10^{-2} M

EQUILIBRIUM PARAMETERS

The strength of an interaction can be given as the **equilibrium dissociation constant**, K_d , or by its reciprocal, K_a (discussed below). When two proteins A and B associate, the K_d is given by

$$K_d = \frac{[A][B]}{[AB]}$$

where $[AB]$ is the concentration of the complexed species and $[A]$ and $[B]$ are the concentrations of the noncomplexed species. Concentrations are given in molar terms, as is the K_d .

There are two useful special cases to keep in mind. First, when A and B are present in equal concentration, the concentration at which half of each protein is present in the form of AB complex is the K_d . Second, when the concentration of A vastly exceeds the concentration of B (say, by 100-fold)—typically indicated $[A] \gg [B]$ —the concentration of A at which half of the B is found in the AB complex approximates K_d .

Typical K_d values for biologically significant interactions are given in Table 20.0.1. Note that many significant interactions are weak.

Alternatively, affinity may be given as an **equilibrium association constant**, K_a . The K_a is simply

$$K_a = \frac{[AB]}{[A][B]}$$

In other words, it is the reciprocal of the equilibrium dissociation constant: $K_a = 1/K_d$. Association constants are used less often than K_d , but do appear in the literature of some subfields—for example, in descriptions of antibody-antigen interactions.

The strength of an interaction is directly proportional to the change in Gibbs free energy (ΔG) when A and B interact, which is given by

$$\Delta G = \Delta H - T\Delta S$$

where T is the temperature in degrees Kelvin, ΔS is the change in entropy (S), and ΔH is the change in enthalpy (H). ΔG is given in units of kilocalories/mole (kcal mol^{-1}). An

increase in the free energy of an interaction by 1 kcal decreases the K_d by a factor of ~7. Note that the temperature affects the entropy term of the equation. As the temperature decreases, reactions driven by enthalpy are less affected by losses in entropy, and are usually favored. By contrast, reactions whose free energy is derived largely from favorable changes in entropy are disfavored at lower temperatures. This means that the dependence of a protein association on temperature can often provide a clue to which term predominately contributes to the free energy change involved in that association.

The relationship between K_a and ΔG is defined as follows:

$$\Delta G^0 = -RT \ln \frac{[AB]}{[A][B]}$$

$$\Delta G^0 = -RT \ln K_a = -RT \ln \left(\frac{1}{K_d} \right) = -RT \ln K_d$$

where ΔG^0 is the free energy change under standard conditions (25°C); R is the universal gas constant (1.9872 calmol⁻¹K⁻¹); and T is the temperature in degrees Kelvin (25°C is 289.1 K). Therefore,

$$\Delta G^0 = 0.588 \ln K_d$$

and, since $\ln x = 2.303 \log_{10} x$, $\Delta G^0 = 1.36 \log_{10} K_d$. For example:

- (a) $K_d = 1 \times 10^{-14}$, then $\Delta G^0 = -19.04 \text{ kcal mol}^{-1}$
- (b) $K_d = 1 \times 10^{-2}$, then $\Delta G^0 = -2.72 \text{ kcal mol}^{-1}$

The streptavidin-biotin reaction (a) is intrinsically more favorable in the direction of binding, [AB] formation, than the low-affinity interaction (b) involving phage repressor (see Table 20.0.1). It should be pointed out that the DG values calculated above assume that the molar ratio of reactants [A] [B] and product [AB] is 1 M (standard state).

Many (but by no means all) biologically important protein interactions seem to be largely driven by ΔH , or changes in enthalpy. That is fortunate, in that changes in entropy on binding are very hard to quantitate, or even think about precisely. For example, proteins in solution are surrounded by water molecules that form hydrogen bonds with surface residues. When two proteins interact, the ordered arrangement of the water molecules that surrounded the interacting surfaces of the proteins is often disrupted, and this loss of order provides an entropically favorable term to the free energy of the interaction. Such changes in the free energy due to entropy are very hard to predict. By contrast, enthalpic changes are easier to understand. If formation of one hydrogen bond liberates about -1 kcal M^{-1} , and formation of a particular ionic contact liberates about -2 kcal M^{-1} , then the energies of these changes are additive such that the formation of both bonds usually liberates -3 kcal M^{-1} —which means that the K_d is decreased >100-fold by the enthalpic changes.

Kinetic Parameters

The above descriptions of protein interactions make no reference to the speed at which association or dissociation occurs. These speeds are given by **kinetic parameters**.

The **dissociation rate constant**, k_{dissoc} , gives the speed at which the AB complex dissociates into A and B ($AB \rightarrow A + B$). k_{dissoc} is a first-order rate constant—i.e., one that

is dependent on the concentration of one species, in this case the AB complex—and is given by the rate of decrease in the concentration of AB:

$$k_{\text{dissoc}}[A][B] = \frac{-d[AB]}{dt}$$

Its units are those of reciprocal time (t in the equation), usually given in sec^{-1} . For example, a dissociation rate constant of 10^{-4}sec^{-1} means that one in 10^4 of the AB complexes present comes apart each second.

Similarly, the **association rate constant**, k_{assoc} , gives the speed at which A and B associate to form an AB complex ($A + B \rightarrow AB$). This is a second-order reaction—i.e., its speed depends on the concentrations of both A and B—and its rate constant is given by

$$k_{\text{assoc}}[A][B] = \frac{+d[AB]}{dt}$$

Its units are those of reciprocal concentration \times reciprocal time, typically given in $\text{M}^{-1} \text{sec}^{-1}$.

For example, suppose that protein A is present at a concentration of 10^{-6}M in a cell, that protein B is injected to a nuclear concentration of 10^{-5}M , and that the rate constant for this antibody-antigen association is $10^{-4} \text{M}^{-1} \text{sec}^{-1}$. After injection, the concentration of AB will be

$$\begin{aligned} [AB] &= [A] \times [B] \times K_{\text{assoc}} \\ &= (10^{-6} \text{M})(10^{-5} \text{M})(10^{-4} \text{M}) \\ &= 10^{-7} \text{M sec}^{-1} \end{aligned}$$

That is, in the first second after mixing, 10^{-7}M of AB complex will form. Every 10-fold increase in the concentration of either reactant increases the rate of product formation 10-fold.

Note that, at equilibrium, by definition,

$$\frac{d[AB]}{dt} = 0$$

and, therefore, for the AB interaction,

$$K_{\text{d}} = \frac{k_{\text{dissoc}}}{k_{\text{assoc}}}$$

The fact that the strength of equilibrium interactions reflects the speed of associations and dissociations has important consequences. To understand this, imagine two pairs of proteins that interact with the same K_{d} . Proteins A and B come together slowly, but, once the AB complex forms, it takes a long time to come apart. By contrast, proteins C and D come together rapidly, and the CD complex dissociates rapidly. There are two common cases in which these kinetic differences in the AB and CD associations would be significant.

One is in measurement. Many techniques, such as the “pulldown” and immunological coprecipitation techniques described in this chapter, rely on the fact that proteins remain associated during some sequence of steps, while they are being separated from other proteins in a mixture and while the isolated complex is being rinsed. No matter how tightly the proteins associate, if they come apart before their complex can be separated and rinsed, the complex will not be detected. Moreover, if the AB and CD interactions

have the same K_d , but CD comes apart more rapidly, then a coprecipitation experiment can falsely suggest that the CD association is weaker.

The second case concerns the biological effects. Many biological phenomena, such as the transcription phenotypes resulting from protein-protein interactions in two-hybrid experiments (*UNIT 19.3*), seem to be well-described by consideration of equilibrium measurements. However, it is worth keeping in mind that any biological process that results from the association of two proteins requires a minimum time to occur. For some enzyme-substrate interactions, the minimum time may be on the order of microseconds, but for others, such as the initiation of DNA replication, it may be measured in seconds. If the complex dissociates faster than the minimum time, then, no matter how tight the interaction, the process will not occur. If AB and CD have the same K_d , but CD dissociates faster, the association may not produce a biological effect.

WHAT THIS CHAPTER DESCRIBES

To a contemporary biologist, the phrase “analysis of protein interactions” encompasses both the identification of interacting proteins and the measurement of the strength and rates of their interactions. This chapter covers both. It does not a number of classical biochemical techniques, including those that separate bound species from free species on chromatographic media, in velocity gradients, or by equilibrium dialysis. It has been designed to present most of the techniques currently used for identification and the most widely used techniques for measurement.

This chapter features several units that provide quite different approaches for identifying interacting proteins. The first approach, described in *UNIT 20.1*, presents recent developments in two-hybrid methods. These approaches allow detection of proteins that interact with a DNA-bound “bait” protein, and isolation of the genes that encode them, based on their biological effect on transcription of reporter genes in yeast.

This chapter also adds a second method (*UNIT 20.6*) for detecting interacting proteins. This method (called the “far-Western”) depends on proteins in a liquid phase (e.g., a cell extract interacting with “bait” proteins separated electrophoretically on a gel). Proteins that interact and whose interaction persists through the washing steps are detected by antibodies against them. This method favors detection of protein interactions with long-off rates.

A third procedure, detailed in *UNIT 20.2*, provides a biochemical method for detecting associations between a GST-fused bait protein and a mixture of proteins that may contain an interacting partner. In this coprecipitation or “pulldown” approach, interacting proteins are identified and can be purified using standard affinity methods for GST-containing proteins.

In the related technique of co-immunoprecipitation (*UNIT 20.5*), cell-free extracts are incubated with an antibody to a desired protein (or to an epitope tag) in order to coprecipitate associated proteins. This approach is used to isolate multiprotein complexes that presumably were present in the intact cell.

Another molecular biological approach that directly yields a clone encoding the interacting protein is described in *UNIT 20.3*. Interaction cloning (also known as expression cloning) is a technique to identify and clone genes which encode proteins that interact with a protein of interest, or “bait” protein. Phage-based interaction cloning requires a gene encoding the bait protein and an appropriate expression library constructed in a bacteriophage expression vector, such as λ gt11.

An in vitro technique called surface plasmon resonance (SPR) can simultaneously detect interactions between unmodified proteins and measure kinetic parameters of the interaction. The availability of user-friendly instruments has facilitated the use of SPR as a means of studying macromolecular interactions. This technology is presented in *UNIT 20.4*.

Roger Brent

Interaction Trap/Two-Hybrid System to Identify Interacting Proteins

To understand the function of a particular protein, it is often useful to identify other proteins with which it associates. This can be done by a selection or screen in which novel proteins that specifically interact with a target protein of interest are isolated from a library. One particularly useful approach to detect novel interacting proteins—the two-hybrid system or interaction trap (see Figs. 20.1.1 and 20.1.2)—uses yeast as a “test tube” and transcriptional activation of a reporter system to identify associating proteins (see Background Information). This approach can also be used specifically to test complex formation between two proteins for which there is a prior reason to expect an interaction.

In the basic version of this method (see Fig. 20.1.2), the plasmid pEG202 or a related vector (see Fig. 20.1.3 and Table 20.1.1) is used to express the probe or “bait” protein as a fusion to the heterologous DNA-binding protein LexA. Many proteins, including transcription factors, kinases, and phosphatases, have been successfully used as bait proteins. The major requirements for the bait protein are that it should not be actively excluded from the yeast nucleus, and it should not possess an intrinsic ability to strongly activate transcription. The plasmid expressing the LexA-fused bait protein (see Table 20.1.1) is used to transform yeast possessing a dual reporter system responsive to transcriptional activation through the LexA operator. In one such example, the yeast strain EGY48 (see Table 20.1.2) contains the reporter plasmid pSH18-34. In this case, binding sites for LexA are located upstream of two reporter genes. In the EGY48 strain, the upstream activating sequences of the chromosomal *LEU2* gene—required in the biosynthetic pathway for leucine (Leu)—are replaced with LexA operators (DNA binding sites). pSH18-34 contains a LexA operator–*lacZ* fusion gene. These two reporters allow selection for transcriptional activation by permitting selection for viability when cells are plated on medium lacking Leu, and discrimination based on color when the yeast is grown on medium containing Xgal (UNIT 13.6).

In Basic Protocol 1, EGY48/pSH18-34 transformed with a bait is characterized for its ability to express protein (Support Protocol 1), growth on medium lacking Leu, and for the level of transcriptional activation of *lacZ* (see Fig. 20.1.2A). A number of alternative strains, plasmids, and strategies are presented which can be employed if a bait proves to have an unacceptably high level of background transcriptional activation.

In an interactor hunt (Basic Protocol 2), the strain EGY48/pSH18-34 containing the bait expression plasmid is transformed (along with carrier DNA made as described in Support Protocol 2) with a conditionally expressed library made in the vector pJG4-5 (see Fig. 20.1.6 and Table 20.1.3). This library uses the inducible yeast *GAL1* promoter to express proteins as fusions to an acidic domain (“acid blob”) that functions as a portable transcriptional activation motif (act) and to other useful moieties. Expression of library-encoded proteins is induced by plating transformants on medium containing galactose (Gal), so yeast cells containing library proteins that do not interact specifically with the bait protein will fail to grow in the absence of Leu (see Fig. 20.1.2B). Yeast cells containing library proteins that interact with the bait protein will form colonies within 2 to 5 days, and the colonies will turn blue when the cells are streaked on medium containing Xgal (see Fig. 20.1.2C). The DNA from interaction trap positive colonies can be analyzed by polymerase chain reaction (PCR) to streamline screening and detect redundant clones in cases where many positives are obtained in screening (see Alternate Protocol 1). The plasmids are isolated and characterized by a series of tests to confirm specificity of the interaction with the initial bait protein (Support Protocols 3 to 5). Those found to be specific are ready for further analysis (e.g., sequencing).

Contributed by Erica A. Golemis, Ilya Serebriiskii, Russell L. Finley, Jr., Mikhail G. Kolonin, Jenő Gyuris, and Roger Brent

Current Protocols in Molecular Biology (1999) 20.1.1–20.1.40

Copyright © 1999 by John Wiley & Sons, Inc.

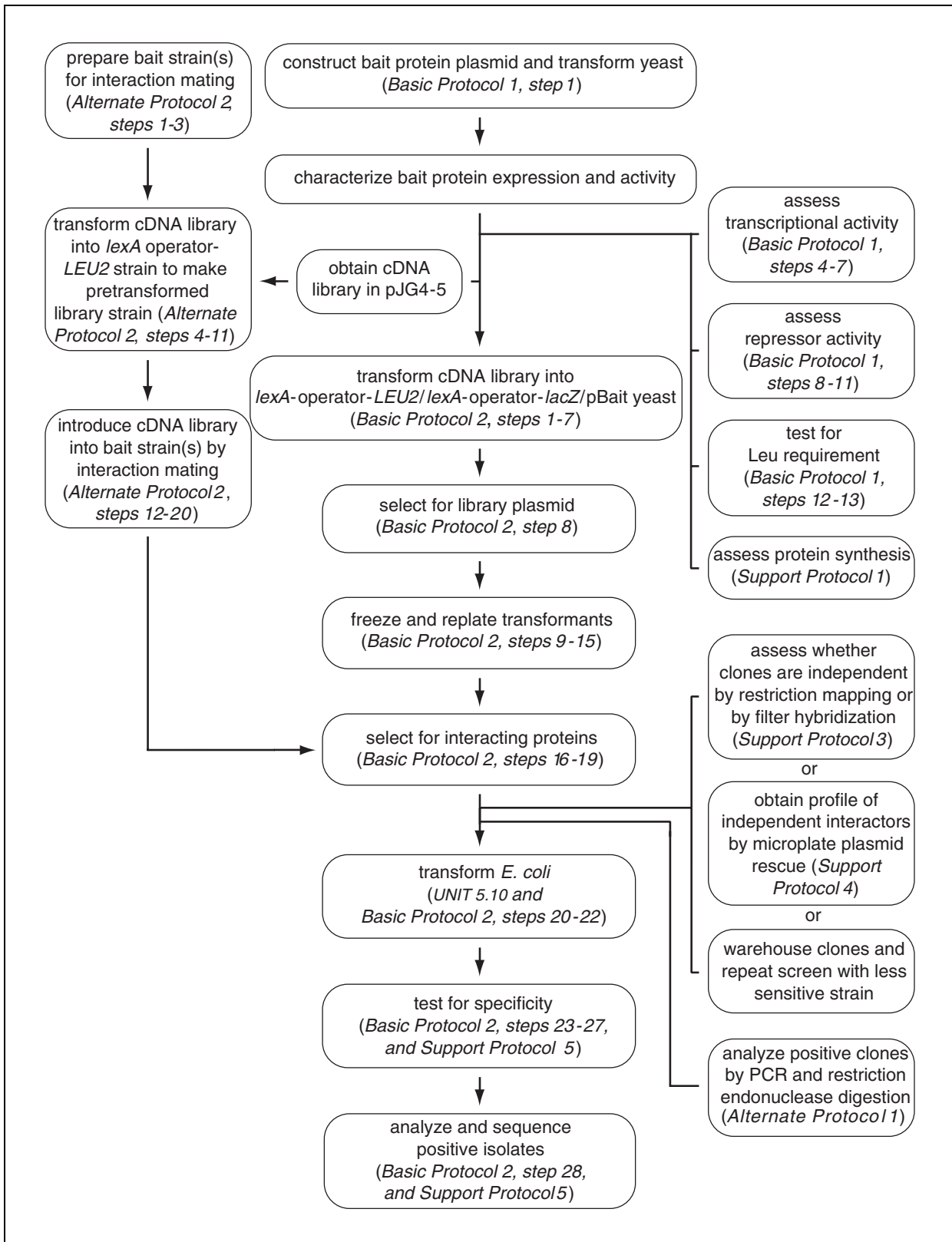


Figure 20.1.1 Flow chart for performing an interaction trap.

20.1.2

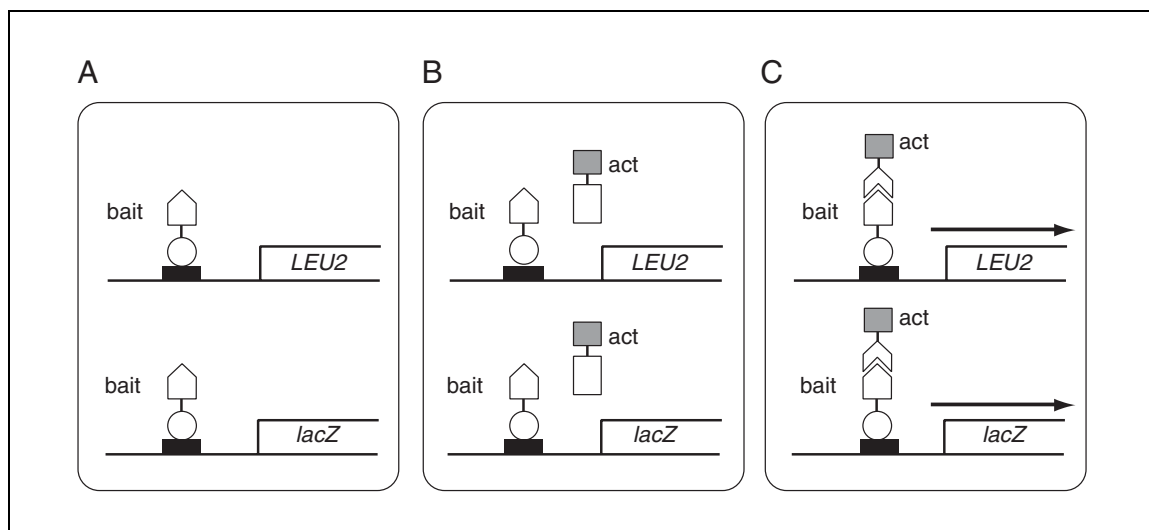


Figure 20.1.2 The interaction trap. **(A)** An EGY48 yeast cell containing two LexA operator–responsive reporters, one a chromosomally integrated copy of the *LEU2* gene (required for growth on –Leu medium), the second a plasmid bearing a *GAL1* promoter–*lacZ* fusion gene (causing yeast to turn blue on medium containing Xgal). The cell also contains a constitutively expressed chimeric protein, consisting of the DNA-binding domain of LexA fused to the probe or bait protein, shown as being unable to activate either of the two reporters. **(B)** and **(C)**, EGY48/pSH18-34/pbait-containing yeast have been additionally transformed with an activation domain (act)–fused cDNA library in pJG4-5, and the library has been induced. In **(B)**, the encoded protein does not interact specifically with the bait protein and the two reporters are not activated. In **(C)**, a positive interaction is shown in which the library-encoded protein interacts with bait protein, resulting in activation of the two reporters (arrow), thus causing growth on medium lacking Leu and blue color on medium containing Xgal. Symbols: black rectangle, LexA operator sequence; open circle, LexA protein; open pentagon, bait protein; open rectangle, library protein; shaded box, activator protein (acid blob in Fig. 20.1.6).

When more than one bait will be used to screen a single library, significant time and resources can be saved by performing the interactor hunt by interaction mating (see Alternate Protocol 2). In this protocol, EGY48 is transformed with library DNA and the transformants are collected and frozen in aliquots. For each interactor hunt, an aliquot of the pretransformed EGY48/library strain is thawed and mixed with an aliquot of a bait strain transformed with the bait expression plasmid and pSH18-34. Overnight incubation of the mixture on a YPD plate results in fusion of the two strains to form diploids. The diploids are then exposed to galactose to induce expression of the library-encoded proteins, and interactors are selected in the same manner as in Basic Protocol 2. The advantage to this approach is that it requires only one high-efficiency library transformation for multiple hunts with different baits. It is also useful for bait proteins that are somewhat toxic to yeast; yeast expressing toxic baits can be difficult to transform with the library DNA.

CHARACTERIZING A BAIT PROTEIN

The first step in an interactor hunt is to construct a plasmid that expresses LexA fused to the protein of interest. This construct is transformed into reporter yeast strains containing *LEU2* and *lacZ* reporter genes, and a series of control experiments is performed to establish whether the construct is suitable as is or must be modified, and whether alternative yeast reporter conditions should be used. These controls establish that the bait protein is made as a stable protein in yeast, that it is capable of entering the nucleus and binding LexA operator sites, and that it does not appreciably activate transcription of the LexA operator–based reporter genes. This last is the most important constraint on use of this system. The LexA-fused bait protein must not activate transcription of either re-

BASIC PROTOCOL 1

Analysis of Protein Interactions

20.1.3

Table 20.1.1 Interaction Trap Components^{a,b}

Plasmid name/source	Selection		Comment/description
	In yeast	In <i>E. coli</i>	
<i>LexA fusion plasmids</i>			
pEG202 ^{c,d,e}	<i>HIS3</i>	Ap ^r	Contains an <i>ADH</i> promoter that expresses LexA followed by polylinker
pJK202	<i>HIS3</i>	Ap ^r	Like pEG202, but incorporates nuclear localization sequences between LexA and polylinker; used to enhance translocation of bait to nucleus
pNLexA ^e	<i>HIS3</i>	Ap ^r	Contains an <i>ADH</i> promoter that expresses polylinker followed by LexA; for use with baits where amino-terminal residues must remain unblocked
pGilda ^d	<i>HIS3</i>	Ap ^r	Contains a <i>GAL1</i> promoter that expresses same LexA and polylinker cassette as pEG202; for use with baits whose continuous presence is toxic to yeast
pEE202I	<i>HIS3</i>	Ap ^r	An integrating form of pEG202 that can be targeted into <i>HIS3</i> following digestion with <i>KpnI</i> ; for use where physiological screen requires lower levels of bait to be expressed
pRFHM1 ^{e,f} (control)	<i>HIS3</i>	Ap ^r	Contains an <i>ADH</i> promoter that expresses LexA fused to the homeodomain of bicoid to produce nonactivating fusion; used as positive control for repression assay, negative control for activation and interaction assays
pSH17-4 ^{e,f} (control)	<i>HIS3</i>	Ap ^r	<i>ADH</i> promoter expresses LexA fused to GAL4 activation domain; used as a positive control for transcriptional activation
pMW101 ^f	<i>HIS3</i>	Cm ^r	Same as pEG202, but with altered antibiotic resistance markers; basic plasmid used for cloning bait
pMW103 ^f	<i>HIS3</i>	Km ^r	Same as pEG202, but with altered antibiotic resistance markers; basic plasmid used for cloning bait
pHybLex/Zeo ^{f,g}	Zeo ^r	Zeo ^r	Bait cloning vector compatible with interaction trap and all other two-hybrid systems; minimal <i>ADH</i> promoter expresses LexA followed by extended polylinker
<i>Activation domain fusion plasmids</i>			
pJG4-5 ^{c,d,e,f}	<i>TRP1</i>	Ap ^r	Contains a <i>GAL1</i> promoter that expresses nuclear localization domain, transcriptional activation domain, HA epitope tag, cloning sites; used to express cDNA libraries
pJG4-5I	<i>TRP1</i>	Ap ^r	An integrating form of pJG4-5 that can be targeted into <i>TRP1</i> by digestion with <i>Bsu36I</i> (New England Biolabs); to be used with pEE202I to study interactions that occur physiologically at low protein concentrations
pYESTrp ^g	<i>TRP1</i>	Ap ^r	Contains a <i>GAL1</i> promoter that expresses nuclear localization domain, transcriptional activation domain, V5 epitope tag, multiple cloning sites; contains f1 ori and T7 promoter/flanking site; used to express cDNA libraries (Invitrogen)
pMW102 ^f	<i>TRP1</i>	Km ^r	Same as pJG4-5, but with altered antibiotic resistance markers; no libraries yet available
pMW104 ^f	<i>TRP1</i>	Cm ^r	Same as pJG4-5, but with altered antibiotic resistance markers; no libraries yet available
<i>LacZ reporter plasmids</i>			
pSH18-34 ^{d,e,f}	<i>URA3</i>	Ap ^r	Contains 8 <i>LexA</i> operators that direct transcription of the <i>lacZ</i> gene; one of the most sensitive indicator plasmids for transcriptional activation
pJK103 ^e	<i>URA3</i>	Ap ^r	Contains two <i>LexA</i> operators that direct transcription of the <i>lacZ</i> gene; an intermediate reporter plasmid for transcriptional activation
pRB1840 ^e	<i>URA3</i>	Ap ^r	Contains 1 <i>LexA</i> operator that directs transcription of the <i>lacZ</i> gene; one of the most stringent reporters for transcriptional activation
pMW112 ^f	<i>URA3</i>	Km ^r	Same as pSH18-34, but with altered antibiotic resistance marker
pMW109 ^f	<i>URA3</i>	Km ^r	Same as pJK103, but with altered antibiotic resistance marker

continued

20.1.4

Table 20.1.1 Interaction Trap Components^{a,b}, continued

Plasmid name/source	Selection		Comment/description
	In yeast	In <i>E. coli</i>	
pMW111 ^f	<i>URA3</i>	Km ^r	Same as pRB1840, but with altered antibiotic resistance marker
pMW107 ^f	<i>URA3</i>	Cm ^r	Same as pSH18-34, but with altered antibiotic resistance marker
pMW108 ^f	<i>URA3</i>	Cm ^r	Same as pJK103, but with altered antibiotic resistance marker
pMW110 ^f	<i>URA3</i>	Cm ^r	Same as pRB1840, but with altered antibiotic resistance marker
pJK101 ^{e,f} (control)	<i>URA3</i>	Ap ^r	Contains a <i>GAL1</i> upstream activating sequence followed by two <i>lexA</i> operators followed by <i>lacZ</i> gene; used in repression assay to assess bait binding to operator sequences

^aAll plasmids contain a 2 μ m origin for maintenance in yeast, as well as a bacterial origin of replication, except where noted (pEE202I, pJG4.5I).

^bInteraction Trap reagents represent the work of many contributors: the original basic reagents were developed in the Brent laboratory (Gyuris et al., 1993). Plasmids with altered antibiotic resistance markers (all pMW plasmids) were constructed at Glaxo in Research Triangle Park, N.C. (Watson et al., 1996). Plasmids and strains for specialized applications have been developed by the following individuals: E. Golemis, Fox Chase Cancer Center, Philadelphia, Pa. (pEG202); J. Kamens, BASF, Worcester, Mass. (pJK202); cumulative efforts of I. York, Dana-Farber Cancer Center, Boston, Mass. and M. Sainz and S. Nottwehr, U. Oregon (pNLexA); D.A. Shaywitz, MIT Center for Cancer Research, Cambridge, Mass. (pGilda); R. Buckholz, Glaxo, Research Triangle Park, N.C. (pEE202I, pJG4-5I); J. Gyuris, Mitotix, Cambridge, Mass. (pJG4-5); S. Hanes, Wadsworth Institute, Albany, N.Y. (pSH17-4); R.L. Finley, Wayne State University School of Medicine, Detroit, Mich. (pRFHM1); S. Hanes, Wadsworth Institute, Albany, N.Y. (pSH18-34); J. Kamens, BASF, Worcester, Mass. (pJK101, pJK103); R. Brent, The Molecular Sciences Institute, Berkeley, Calif. (pRB1840). Specialized plasmids not yet commercially available can be obtained by contacting the Brent laboratory at (510) 647-0690 or brent@molsci.org, or the Golemis laboratory, (215) 728-2860 or EA_Golemis@fcc.edu.

^cSequence data are available for pEG202 (pLexA) accession number pending.

^dPlasmids commercially available from Clontech and OriGene; for Clontech pEG202 is listed as pLexA, pJG4-5 as pB42AD, and pSH18-34 as p8op-LacZ.

^ePlasmids and strains available from OriGene.

^fIn pMW plasmids the ampicillin resistance gene (Ap^r) is replaced with the chloramphenicol resistance gene (Cm^r) and the kanamycin resistance gene (Km^r) from pBC SK(+) and pBK-CMV (Stratagene), respectively. The choice between Km^r and Cm^r or Ap^r plasmids is a matter of personal taste; use of basic Ap^r plasmids is described in the basic protocols. Use of the more recently developed reagents would facilitate the purification of library plasmid in later steps by eliminating the need for passage through KC8 bacteria, with substantial saving of time and effort. Ap^r has been maintained as marker of choice for the library plasmid because of the existence of multiple libraries already possessing this marker. These plasmids are the basic set of plasmids recommended for use.

^gPlasmids commercially available from Invitrogen as components of a Hybrid Hunter kit; this kit also includes all necessary positive and negative controls (not listed in this table). See Background Information for further details on commercially available reagents.

porter—the EGY48 strain (or related strain EGY191) that expresses the LexA fusion protein should not grow on medium lacking Leu, and the colonies should be white on medium containing Xgal. The characterized bait protein plasmid is used for Basic Protocol 2 to screen a library for interacting proteins.

Materials

DNA encoding the protein of interest

Plasmids (see Table 20.1.1): pEG202 (see Fig. 20.1.3), pSH18-34 (see Fig. 20.1.4), pSH17-4, pRFHM1, and pJK101 for basic characterization; other plasmids for specific circumstances as described (Clontech, Invitrogen, OriGene, or R. Brent)

Yeast strain EGY48 (*ura3 trp1 his3 3LexA-operator-LEU2*), or EGY191 (*ura3 trp1 his3 1LexA-operator-LEU2*; Table 20.1.2)

Complete minimal (CM) medium dropout plates (*UNIT 13.1*), supplemented with 2% (w/v) of the indicated sugars (glucose or galactose), in 100-mm plates:

Glu/CM, –Ura, –His

Gal/CM, –Ura, –His

Gal/CM, –Ura, –His, –Leu

Z buffer (*UNIT 13.6*) with 1 mg/ml 5-bromo-4-chloro-3-indolyl- β -D-galactosidase (Xgal)

**Analysis of
Protein
Interactions**

20.1.5

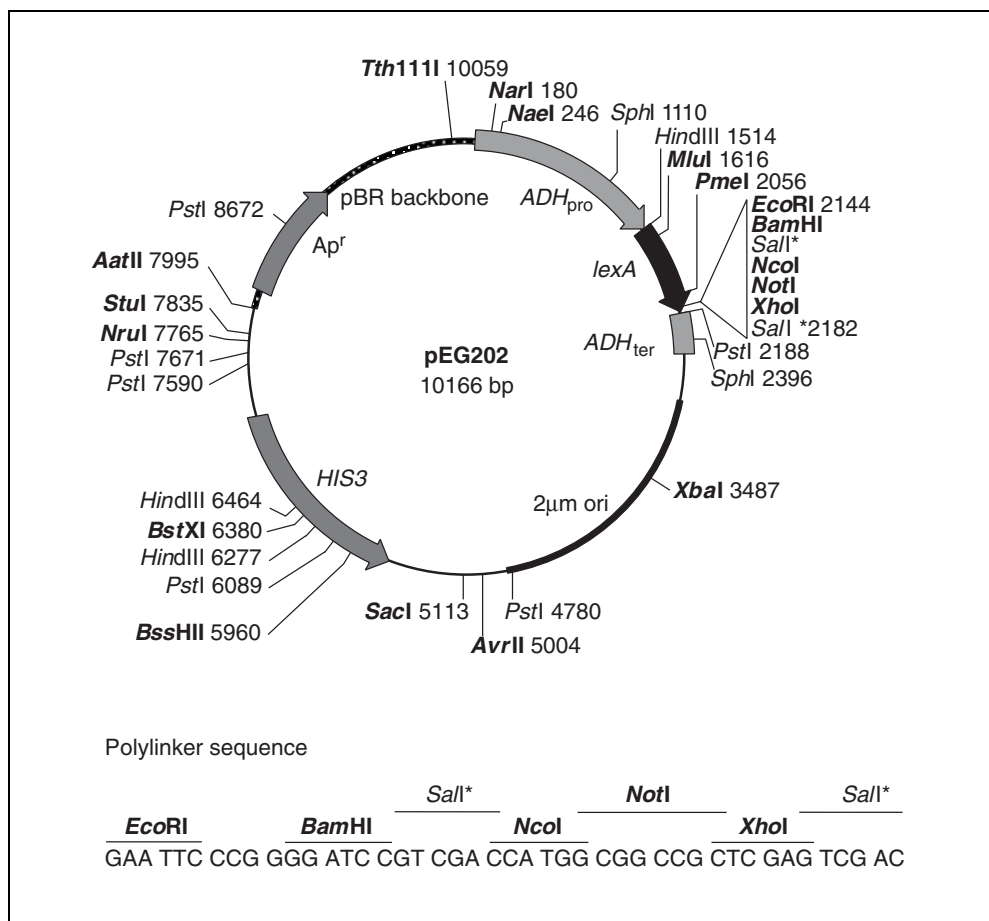


Figure 20.1.3 LexA-fusion plasmids: pEG202. The strong constitutive *ADH* promoter is used to express bait proteins as fusions to the DNA-binding protein LexA. Restriction sites shown in this map are based on recently compiled pEG202 sequence data and include selected sites suitable for diagnostic restriction endonuclease digests. A number of restriction sites are available for insertion of coding sequences to produce protein fusions with LexA; the polylinker sequence and reading frame relative to LexA are shown below the map with unique sites marked in bold type. The sequence 5'-CGT CAG CAG AGC TTC ACC ATT G-3' can be used to design a primer to confirm correct reading frame for LexA fusions. Plasmids contain the *HIS3* selectable marker and the 2μm origin of replication to allow propagation in yeast, and the *Ap^r* antibiotic resistance gene and the pBR origin of replication to allow propagation in *E. coli*. In the recently developed LexA-expression plasmids pMW101 and pMW103, the ampicillin resistance gene (*Ap^r*) has been replaced with the chloramphenicol resistance gene (*Cm^r*) and the kanamycin resistance gene (*Km^r*), respectively (see Table 19.2.1 for details).

Gal/CM dropout liquid medium (*UNIT 13.1*) supplemented with 2% Gal
 Antibody to LexA or fusion domain: monoclonal antibody to LexA (Clontech,
 Invitrogen) or polyclonal antibody to LexA (available by request from R. Brent
 or E. Golemis)
 H₂O, sterile
 30°C incubator
 Nylon membrane
 Whatman 3MM filter paper

Additional reagents and equipment for subcloning DNA fragments (*UNIT 3.16*), lithium acetate transformation of yeast (*UNIT 13.7*), liquid assay for β-galactosidase (*UNIT 13.6*), preparation of protein extracts for immunoblot analysis (see Support Protocol 1), and immunoblotting and immunodetection (*UNIT 10.8*)

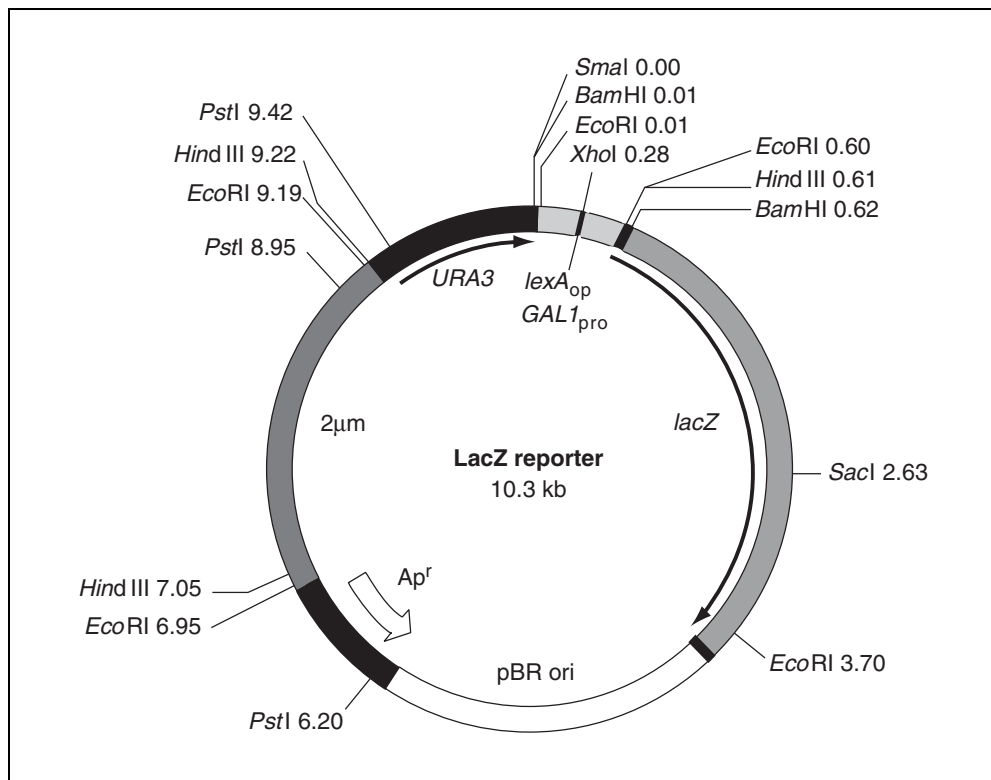


Figure 20.1.4 LacZ reporter plasmid. pRB1840, pJK103, and pSH18-34 are all derivatives of LR1Δ1 (West et al., 1984) containing eight, two, or one operator for LexA (*LexA_{op}*) binding inserted into the unique *XhoI* site located in the minimal *GAL1* promoter (*GAL1_{pro}*; 0.28 on map). The plasmid contains the *URA3* selectable marker, the 2μm origin to allow propagation in yeast, the ampicillin resistance (*Ap^r*) gene, and the pBR322 origin (*ori*) to allow propagation in *E. coli*. Numbers indicate relative map positions. In the recently developed derivatives, the ampicillin resistance gene (*Ap^r*) has been replaced with the chloramphenicol or kanamycin resistance genes (see Table 19.2.1 for details).

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

Transform yeast with the bait protein plasmid

1. Using standard subcloning techniques (*UNIT 3.16*), insert the DNA encoding the protein of interest into the polylinker of pEG202 (see Fig. 20.1.3) or other LexA fusion plasmid to make an in-frame protein fusion.

The LexA fusion protein is expressed from the strong alcohol dehydrogenase (ADH) promoter. pEG202 also contains a HIS3 selectable marker and a 2μm origin for propagation in yeast. pEG202 with the DNA encoding the protein of interest inserted is designated pBait. Uses of alternative LexA fusion plasmids are described in Background Information.

2. Perform three separate lithium acetate transformations (*UNIT 13.7*) of EGY48 using the following combinations of plasmids:

pBait + pSH18-34 (test)
 pSH17-4 + pSH18-34 (positive control for activation)
 pRFHM1 + pSH18-34 (negative control for activation).

Use of the two LexA fusions as positive and negative controls allows a rough assessment of the transcriptional activation profile of LexA bait proteins. pEG202 itself is not a good negative control because the peptide encoded by the uninterrupted polylinker sequences is itself capable of very weakly activating transcription.

Table 20.1.2 Interaction Trap Yeast Selection Strains^a

Strain	Relevant genotype	Number of operators	Comments/description
EGY48 ^{b,c,d}	<i>MATα trp1, his3, ura3, lexAops-LEU2</i>	6	<i>lexA</i> operators direct transcription from the <i>LEU2</i> gene; EGY48 is a basic strain used to select for interacting clones from a cDNA library
EGY191	<i>MATα trp1, his3, ura3, lexAops-LEU2</i>	2	EGY191 provides a more stringent selection than EGY48, producing lower background with baits with intrinsic ability to activate transcription
L40 ^c	<i>MATα trp1, leu2, ade2, GAL4, lexAops-HIS34, lexAops-lacZ8</i>		Expression driven from <i>GAL1</i> promoter is constitutive in L40 (inducible in EGY strains); selection is for <i>HIS</i> prototrophy. Integrated <i>lacZ</i> reporter is considerably less sensitive than pSH18-34 maintained in EGY strains

^aInteraction Trap reagents represent the work of many contributors: the original basic reagents were developed in the Brent laboratory (Gyuris et al., 1993). Strains for specialized applications have been developed by the following individuals: E. Golemis, Fox Chase Cancer Center, Philadelphia, Pa. (EGY48, EGY191); A.B. Vojtek and S.M. Hollenberg, Fred Hutchinson Cancer Research Center, Seattle, Wash. (L40). Specialized strains not yet commercially available can be obtained by contacting the Brent laboratory at The Molecular Sciences Institute, Berkeley, (510) 647-0690 or brent@molsci.org, or the Golemis laboratory, (215) 728-2860 or EA_Golemis@fcc.edu.

^bStrains commercially available from Clontech.

^cStrains commercially available from Invitrogen as components of a Hybrid Hunter kit; the kit also includes all necessary positive and negative controls (not listed in this table). See Background Information for further details on commercially available reagents.

^dStrains commercially available from OriGene.

pSH18-34 contains a 2 μ m origin and a *URA3* selectable marker for maintenance in yeast, as well as a bacterial origin of replication and ampicillin-resistance gene. It is the most sensitive *lacZ* reporter available and will detect any potential ability to activate *lacZ* transcription. *pSH17-4* is a *HIS3* 2 μ m plasmid encoding *LexA* fused to the activation domain of the yeast activator protein *GAL4*. This fusion protein strongly activates transcription. *pRFHM1* is a *HIS3* 2 μ m plasmid encoding *LexA* fused to the N-terminus of the *Drosophila* protein *bicoid*. This fusion protein has no ability to activate transcription.

3. Plate each transformation mixture on Glu/CM –Ura, –His dropout plates. Incubate 2 days at 30°C to select for yeast that contain both plasmids.

Colonies obtained can be used simultaneously in tests for the activation of lacZ (steps 4 to 7) and LEU2 (steps 12 to 13) reporters.

Assay *lacZ* gene activation by β -galactosidase assay

4. Streak a Glu/CM –Ura, –His master dropout plate with at least five or six independent colonies obtained from each of the three transformations in step 3 (test, positive control, and negative control) and incubate overnight at 30°C.

The filter assay described in Steps 5a to 7a (based on Breeden and Nasmyth, 1985) provides a rapid assay for β -galactosidase transcription. Alternatively, a liquid assay (UNIT 13.6) or a plate assay (described in Steps 5b to 7b) may be used.

Perform filter assay for β -galactosidase activity:

- 5a. Lift colonies by gently placing a nylon membrane on the yeast plate and allowing it to become wet through. Remove the membrane and air dry 5 min. Chill the membrane, colony side up, 10 min at –70°C.

Whatman 3MM filters can be cut to the size of the yeast plate as a more economical alternative to nylon membranes for performing lifts. In addition, two or three 5-min temperature cycles (–70°C to room temperature) can be used instead of a single cycle to promote better lysis; this may be worth doing if there is difficulty visualizing blue color.

- 6a. Cut a piece of Whatman 3MM filter paper slightly larger than the colony membrane and soak it in Z buffer containing 1 mg/ml Xgal. Place colony membrane, colony side up, on Whatman 3MM paper, or float it in the lid of a petri dish containing ~2 ml Z buffer with 1 mg/ml Xgal.

Acceptable results may be obtained using as little as 300 µg/ml Xgal.

- 7a. Incubate at 30°C and monitor for color changes.

It is generally useful to check the membrane after 20 min, and again after 2 to 3 hr. Strong activators will produce a blue color in 5 to 10 min, and a bait protein (LexA fusion protein) that does so is unsuitable for use in an interactor hunt using this lacZ reporter plasmid. Weak activators will produce a blue color in 1 to 6 hr (compare versus negative control pRFHMI which will itself produce a faint blue color with time) and may or may not be suitable. Weak activators should be tested using the repressor assay described in steps 8 to 11.

Perform Xgal plate assay for lacZ activation:

- 5b. Prepare Z buffer Xgal plates as described in UNIT 13.1.

For activation assays, plates should be prepared with glucose as a sugar source. For repression assays (steps 8 to 11), galactose should be used as a sugar source. In our experience, when patching from a master plate to Xgal plates, sufficient yeast are transferred that plasmid loss is not a major problem even in the absence of selection; this is balanced by the desire to assay sets of constructs on the same plate to eliminate batch variation in Xgal potency. Hence, plates should be made either with complete minimal amino acid mix, or by dropping out only uracil (-Ura), to make the plates universally useful.

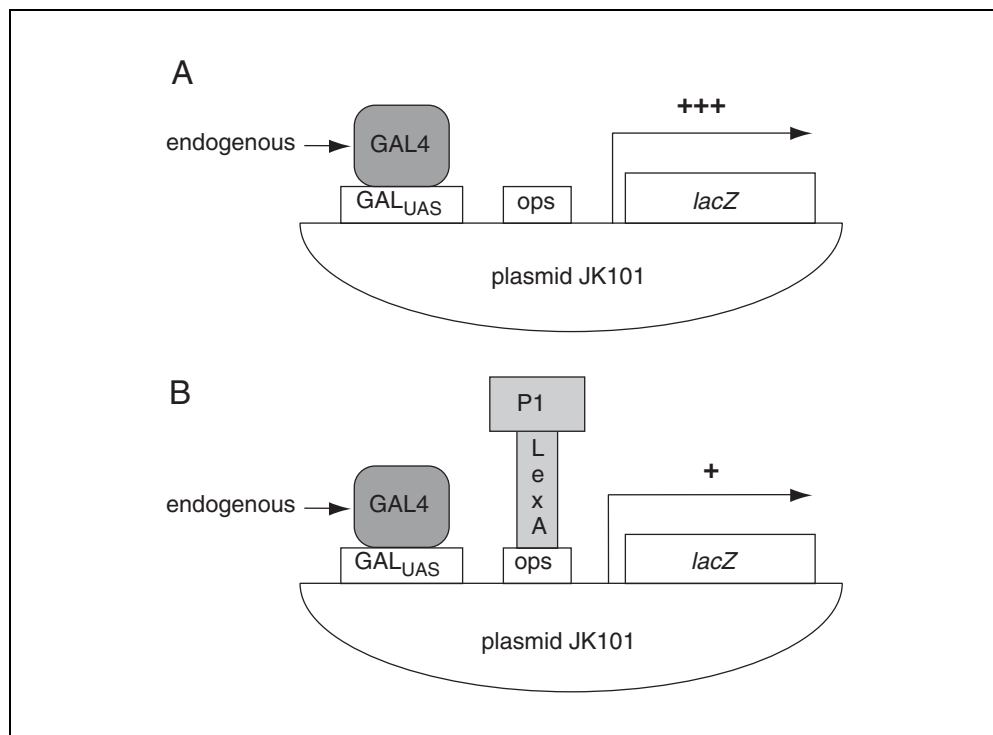


Figure 20.1.5 Repression assay for DNA binding. **(A)** The plasmid JK101 contains the upstream activating sequence (UAS) from the *GAL1* gene followed by *LexA* operators upstream of the *lacZ* coding sequence. Thus, yeast containing pJK101 will have significant β-galactosidase activity when grown on medium in which galactose is the sole carbon source because of binding of endogenous yeast GAL4 to the GAL_{UAS}. **(B)** LexA-fused proteins (P1-LexA) that are made, enter the nucleus, and bind the LexA operator sequences (ops) will block activation from the GAL_{UAS}, repressing β-galactosidase activity (+) 3- to 5-fold. On glucose/Xgal medium, yeast containing pJK101 should be white because GAL_{UAS} transcription is repressed.

6b. Streak yeast from master plate to Xgal plate and incubate at 30°C.

7b. Examine plates for color development at intervals over the next 2 to 3 days.

Strongly activating fusions should be visibly blue on the plate within 12 to 24 hr; moderate activators will be visibly blue after ~2 days.

When a bait protein appreciably activates transcription under these conditions, there are several recourses. The first and simplest is to switch to a less sensitive lacZ reporter plasmid; use of pJK103 and pRB1840 may be sufficient to reduce background to manageable levels. If this fails to work, it is frequently possible to generate a truncated LexA fusion that does not activate transcription.

Confirm fusion-protein synthesis by repression assay

For LexA fusions that do not activate transcription, confirm by performing a repression assay (Brent and Ptashne, 1984) that the LexA fusion protein is being synthesized in yeast (some proteins are not) and that it is capable of binding LexA operator sequences (Fig. 20.1.5). The following steps can be performed concurrently with the activation assay.

8. Transform EGY48 yeast with the following combinations of plasmids (three transformations):

pBait + pJK101 (test)

pRFHM1 + pJK101 (positive control for repression)

pJK101 alone (negative control for repression).

9. Plate each transformation mix on Glu/CM–Ura, –His dropout plates or Glu/CM–Ura dropout plates as appropriate to select yeast cells that contain the indicated plasmids. Incubate 2 to 3 days at 30°C until colonies appear.

10. Streak colonies to a Glu/CM–Ura, –His or Glu/CM–Ura dropout master plate and incubate overnight at 30°C.

11. Assay β -galactosidase activity of the three transformed strains (test, positive control, and negative control) by liquid assay (using Gal/CM dropout liquid medium), filter assay (steps 5a to 7a, first restreaking to Gal/CM plates to grow overnight), or plate assay (steps 5b to 7b, using Gal/CM–Ura XGal plates).

This assay should not be run for more than 1 to 2 hr for membranes, or 36 hr for Xgal plates, as the high basal lacZ activity will make differential activation of pJK101 impossible to see with longer incubations. Use of Xgal plates, and inspection 12 to 24 hr after streaking, is generally most effective.

The plasmid pJK101 contains the galactose upstream activating sequence (UAS) followed by LexA operators upstream of the lacZ coding sequence. Thus, yeast containing pJK101 will have significant β -galactosidase activity when grown on medium in which galactose is the sole carbon source because of binding of endogenous yeast GAL4 to the GAL_{UAS}. LexA-fused proteins that are made, enter the nucleus, and bind the LexA operator sequences block activation from the GAL_{UAS}, repressing β -galactosidase activity 3- to 20-fold. Note that on Glu/Xgal medium, yeast containing pJK101 should be white, because GAL_{UAS} transcription is repressed.

12. If a bait protein neither activates nor represses transcription, perform immunoblot analysis by probing an immunoblot of a crude lysate with antibodies against LexA or the fusion domain to test for protein synthesis (see Support Protocol 1).

Even if a bait protein represses transcription, it is generally a good idea to assay for the production of full-length LexA fusions, as occasionally some fusion proteins will be proteolytically cleaved by endogenous yeast proteases. If the protein is made but does not repress, it may be necessary to clone the sequence into a LexA fusion vector that contains a nuclear localization motif, e.g., pJK202 (see Table 20.1.1), or to modify or truncate the fusion domain to remove motifs that target it to other cellular compartments (e.g., myristoylation signals).

Test for *Leu* requirement

These steps can be performed concurrently with the *lacZ* activation and repression assays.

13. Disperse a colony of EGY48 containing pBait and pSH18-34 reporter plasmids into 500 μ l sterile water. Dilute 100 μ l of suspension into 1 ml sterile water. Make a series of 1/10 dilutions in sterile water to cover a 1000-fold concentration range.
14. Plate 100 μ l from each tube (undiluted, 1/10, 1/100, and 1/1000) on Gal/CM –Ura, –His dropout plates and on Gal/CM –Ura, –His, –Leu dropout plates. Incubate overnight at 30°C.

There will be a total of eight plates. Gal/CM –Ura, –His dropout plates should show a concentration range from 10 to 10,000 colonies and Gal/CM –Ura, –His, –Leu dropout plates should have no colonies.

Actual selection in the interactor hunt is based on the ability of the bait protein and acid-fusion pair, but not the bait protein alone, to activate transcription of the LexA operator-LEU2 gene and allow growth on medium lacking Leu. Thus, the test for the Leu requirement is the most important test of whether the bait protein is likely to have an unworkably high background. The LEU2 reporter in EGY48 is more sensitive than the pSH18-34 reporter for some baits, so it is possible that a bait protein that gives little or no signal in a β -galactosidase assay would nevertheless permit some level of growth on –Leu medium. If this occurs, there are several options for proceeding, the most immediate of which is to substitute EGY191 (see Table 20.1.2), a less sensitive screening strain, and repeat the assay.

As outlined in this protocol, the authors recommend the strategy of performing the initial screening using the most sensitive reporters and then, if activation is detected, screening with increasingly less sensitive reporters (see Critical Parameters for further discussion).

PERFORMING AN INTERACTOR HUNT

An interactor hunt involves two successive large platings of yeast containing LexA-fused probes and reporters and libraries in pJG4-5 (Fig. 20.1.6, Table 20.1.3) with a cDNA expression cassette under control of the *GAL* promoter. In the first plating, yeast are plated on complete minimal (CM) medium –Ura, –His, –Trp dropout plates with glucose (Glu) as a sugar source to select for the library plasmid. In the second plating, which selects for yeast that contain interacting proteins, a slurry of primary transformants is plated on CM –Ura, –His, –Trp, –Leu dropout plates with galactose/raffinose (Gal/Raff) as the sugar source. This two-step selection is encouraged for two reasons. First, a number of interesting cDNA-encoded proteins may be deleterious to the growth of yeast that bear them; these would be competed out in an initial mass plating. Second, it seems likely that immediately after simultaneous transformation and Gal induction, yeast bearing particular interacting proteins may not be able to initially express sufficient levels of these proteins to support growth on medium lacking Leu. Library plasmids from colonies identified in the second plating are purified by bacterial transformation and used to transform yeast cells for the final specificity screen.

A list of libraries currently available for use with this system is provided in Table 20.1.3. The protocol outlined below describes the steps used to perform a single-step screen that should saturate a library derived from a mammalian cell. For screens with libraries derived from lower eukaryotes with less complex genomes, fewer plates will be required.

Occasionally, baits that seemed well-behaved during preliminary tests produce unworkably high backgrounds of “positives” during an actual screen (see Background Information and Critical Parameters). To forestall the waste of time and materials performing a screen with such a bait would entail, an alternative approach is to perform a scaled-back

screen when working with a new bait (e.g., 5 rather than 30 plates of primary transformants). The results can be assessed before doing a full screen; it is then possible to switch to lower-sensitivity reporter strains and plasmids, if appropriate. Although individual baits will vary, the authors' current default preference is to use the *lacZ* reporter pJK103 in conjunction with either EGY48 or EGY191. Polymerase chain reaction (PCR) can also be used in a rapid screening approach that may be preferable if a large number of positions are obtained in a library screen (see Alternate Protocol 1).

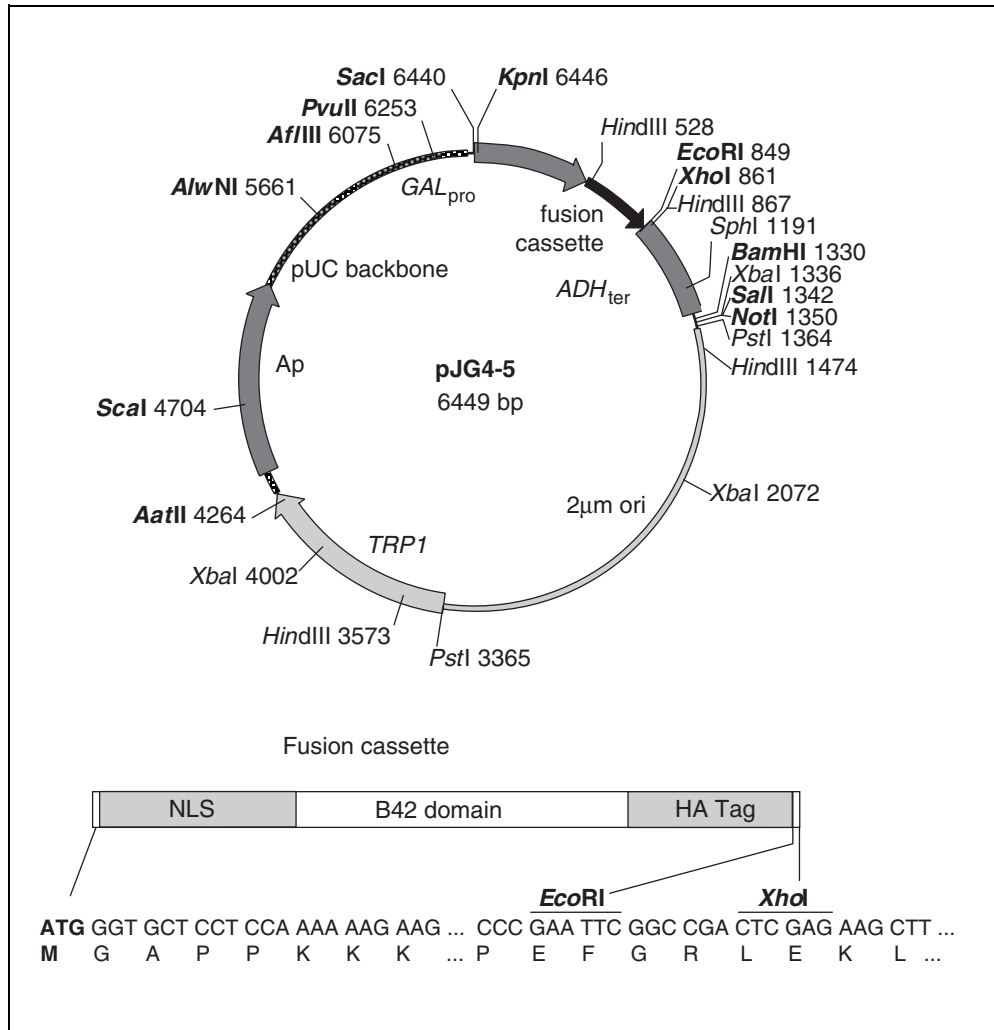


Figure 20.1.6 Library plasmids: pJG4-5. Library plasmids express cDNAs or other coding sequences inserted into unique *EcoRI* and *XhoI* sites as a translational fusion to a cassette consisting of the SV40 nuclear localization sequence (NLS; PPKKKRKVA), the acid blob B42 domain (Ruden et al, 1991), and the hemagglutinin (HA) epitope tag (YPYDVPDYA). Expression of cassette sequences is under the control of the *GAL1* galactose-inducible promoter. This map is based on the sequence data available for pJG4-5, and includes selected sites suitable for diagnostic restriction digests (shown in bold). The sequence 5'-CTG AGT GGA GAT GCC TCC-3' can be used as a primer to identify inserts or to confirm correct reading frame. The pJG4-5 plasmid contains the *TRP1* selectable marker and the 2µm origin to allow propagation in yeast, and the antibiotic resistance gene and the pUC origin to allow propagation in *E. coli*. In the recently developed pJG4-5 derivative plasmids pMW104 and pMW102, the ampicillin resistance gene (*Ap*^r) has been replaced with the chloramphenicol resistance gene (*Cm*^r) and the kanamycin resistance gene (*Km*^r), respectively (see Table 19.2.2 for details). Currently existing libraries are all made in the pJG4-5 plasmid (Gyuris et al., 1993) shown on this figure. Unique sites are marked in bold type.

Materials

Yeast containing appropriate combinations of plasmids (see Table 20.1.1 and Table 20.1.2):

EGY48 containing *LexA*-operator-*lacZ* reporter and pBait (see Basic Protocol 1)

EGY48 containing *LexA*-operator-*lacZ* reporter and pRFHM-1

EGY48 containing *LexA*-operator-*lacZ* reporter and any nonspecific bait

Complete minimal (CM) dropout liquid medium (UNIT 13.1) supplemented with sugars (glucose, galactose, and/or raffinose) as indicated [2% (w/v) Glu, or 2% (w/v) Gal + 1% (w/v) Raff]:

Glu/CM –Ura, –His

Glu/CM –Trp

Gal/Raff/CM –Ura, –His, –Trp

H₂O, sterile

TE buffer (pH 7.5; APPENDIX 2)/0.1 M lithium acetate

Library DNA in pJG4-5 (Table 20.1.3 and Fig. 20.1.6)

High-quality sheared salmon sperm DNA (see Support Protocol 2)

40% (w/v) polyethylene glycol 4000 (PEG 4000; filter sterilized)/0.1 M lithium acetate/TE buffer (pH 7.5)

Dimethyl sulfoxide (DMSO)

Complete minimal (CM) medium dropout plates (UNIT 13.1) supplemented with sugars and Xgal (20 µg/ml) as indicated [2% (w/v) Glu, and 2% (w/v) Gal + 1% (w/v) Raff]:

Glu/CM –Ura, –His, –Trp, 24 × 24-cm (Nunc) and 100-mm

Gal/Raff/CM –Ura, –His, –Trp, 100-mm

Gal/Raff/CM –Ura, –His, –Trp, –Leu, 100-mm

Glu/Xgal/CM –Ura, –His, –Trp, 100-mm

Gal/Raff/Xgal/CM –Ura, –His, –Trp, 100-mm

Glu/CM –Ura, –His, –Trp, –Leu, 100-mm

Glu/CM –Ura, –His, 100-mm

Gal/CM –Ura, –His, –Trp, –Leu, 100-mm

TE buffer (pH 7.5), sterile (optional)

Glycerol solution (see recipe)

E. coli KC8 (*pyrF leuB600 trpC hisB463*; constructed by K. Struhl and available from R. Brent)

LB/ampicillin plates (UNIT 1.1)

E. coli DH5α or other strain suitable for preparation of DNA for sequencing

Bacterial defined minimal A medium plates: 1× A medium plates containing 0.5 µg/ml vitamin B1 (UNIT 1.1) and supplemented with 40 µg/ml each Ura, His, and Leu

30°C incubator, with and without shaking

Low-speed centrifuge and rotor

50-ml conical tubes, sterile

1.5-ml microcentrifuge tubes, sterile

42°C heating block

Glass microscope slides, sterile

Additional reagents and equipment for rapid miniprep isolation of yeast DNA

(UNIT 13.11), transformation of bacteria by electroporation (UNIT 1.8), miniprep isolation of bacterial DNA (UNIT 1.6), restriction endonuclease digestion (UNIT 3.1; optional), and agarose gel electrophoresis (UNIT 2.5A; optional)

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

Table 20.1.3 Libraries Compatible with the Interaction Trap System^a

Source of RNA/DNA	Vector	Independent clones	Insert size (average) ^b	Contact information
Cell lines				
HeLa cells (human cervical carcinoma)	JG	9.6×10^6	0.3-3.5 kb (1.5 kb)	R. Brent, Clontech, Invitrogen, OriGene
HeLa cells (human cervical carcinoma)	Y	3.7×10^6	0.3-1.2 kb	Invitrogen
WI-38 cells (human lung fibroblasts), serum-starved, cDNA	JG	5.7×10^6	0.3-3.5 kb (1.5 kb)	R. Brent, Clontech, OriGene
Jurkat cells (human T cell leukemia), exponentially growing, cDNA	JG	4.0×10^6	0.7-2.8 kb (1.5 kb)	R. Brent
Jurkat cells (human T cell leukemia)	Y	3.2×10^6	0.3-1.2 kb	Invitrogen
Jurkat cells (human T cell leukemia)	Y	3.0×10^6	0.5-4.0 kb (1.8 kb)	Clontech
Jurkat cells (human T cell leukemia)	JG	5.7×10^6	(>1.3)	OriGene
Jurkat cells (human T cell leukemia)	JG	2×10^6	0.7-3.5 kb (1.2 kb)	S. Witte
Be Wo cells (human fetal placental choriocarcinoma)	Y	5.4×10^6	0.3-0.8 kb	Invitrogen
Human lymphocyte	JG	4.0×10^6	0.4-4.0 kb (2.0 kb)	Clontech
CD4 ⁺ T cell, murine, cDNA	JG	> 10^6	0.3-2.5 kb (>0.5 kb)	R. Brent
Chinese hamster ovary (CHO) cells, exponentially growing, cDNA	JG	1.5×10^6	0.3-3.5 kb	R. Brent
A20 cells (mouse B cell lymphoma)	Y	3.11×10^6	0.3-1.2 kb	Invitrogen
Human B cell lymphoma	JG	—	—	H. Niu
Human 293 adenovirus-infected (early and late stages)	JG	—	—	K. Gustin
SKOV3 human Y ovarian cancer	Y	5.0×10^6	(>1.4 kb)	OriGene
MDBK cell, bovine kidney	JG	5.8×10^6	(>1.2 kb)	OriGene
MDCK cells	JG	—	—	D. Chen
HepG2 cell line cDNA	JG	2×10^6	—	M. Melegari
MCF7 breast cancer cells, untreated	JG	1.0×10^7	(>1.5 kb)	OriGene
MCF7 breast cancer cells, estrogen-treated	JG	1.0×10^7	(>1.1 kb)	OriGene
MCF7 cells, serum-grown	JG	1.0×10^7	0.4-3.5 kb	OriGene
LNCAP prostate cell line, untreated	JG	2.9×10^6	(>0.8 kb)	OriGene
LNCAP prostate cell line, androgen-treated	JG	4.6×10^6	(>0.9 kb)	OriGene
Mouse pachytene spermatocytes	JG	—	—	C. Hoog
Tissues				
Human breast	Y	9×10^6	0.4-1.2 kb	Invitrogen
Human breast tumor	Y	8.84×10^6	0.4-1.2 kb	Invitrogen
Human liver	JG	> 10^6	0.6-4.0 kb (>1 kb)	R. Brent
Human liver	Y	2.2×10^6	0.5-4 kb (1.3 kb)	Clontech
Human liver	JG	3.2×10^6	0.3-1.2 kb	Invitrogen
Human liver	JG	1.1×10^7	(> 1 kb)	OriGene
Human lung	Y	5.9×10^6	0.4-1.2 kb	Invitrogen
Human lung tumor	Y	1.9×10^6	0.4-1.2	Invitrogen
Human brain	JG	3.5×10^6	0.5-4.5 kb (1.4 kb)	Clontech
Human brain	Y	8.9×10^6	0.3-1.2 kb	Invitrogen
Human testis	Y	6.4×10^6	0.3-1.2 kb	Invitrogen
Human testis	JG	3.5×10^6	0.4-4.5 kb (1.6 kb)	Clontech
Human ovary	Y	4.6×10^6	0.3-1.2 kb	Invitrogen

*continued***20.1.14**

Table 20.1.3 Libraries Compatible with the Interaction Trap System^a, continued

Source of RNA/DNA	Vector	Independent clones	Insert size (average) ^b	Contact information
Human ovary	JG	4.6 × 10 ⁶	(>1.3 kb)	OriGene
Human ovary	JG	3.5 × 10 ⁶	0.5-4.0 kb (1.8 kb)	Clontech
Human heart	JG	3.0 × 10 ⁶	0.3-3.5 kb (1.3 kb)	Clontech
Human placenta	Y	4.8 × 10 ⁶	0.3-1.2 kb	Invitrogen
Human placenta	JG	3.5 × 10 ⁶	0.3-4.0 kb (1.2 kb)	Clontech
Human mammary gland	JG	3.5 × 10 ⁶	0.5-5 kb (1.6 kb)	Clontech
Human peripheral blood leucocyte	JG	1.0 × 10 ⁷	(>1.3 kb)	OriGene
Human kidney	JG	3.5 × 10 ⁶	0.4-4.5 kb (1.6 kb)	Clontech
Human fetal kidney	JG	3.0 × 10 ⁶	(>1 kb)	OriGene
Human spleen	Y	1.14 × 10 ⁷	0.4-1.2 kb	Invitrogen
Human prostate	Y	5.5 × 10 ⁶	0.4-1.2 kb	Invitrogen
Human normal prostate	JG	1.4 × 10 ⁶	0.4-4.5 kb (1.7 kb)	Clontech
Human prostate	JG	1.4 × 10 ⁶	(>1 kb)	OriGene
Human prostate cancer	JG	1.1 × 10 ⁶	(>0.9 kb)	OriGene
Human fetal prostate	JG	—	—	OriGene
Human fetal liver	JG	3.5 × 10 ⁶	0.3-4.5 kb (1.3 kb)	Clontech
Human fetal liver	Y	2.37 × 10 ⁶	0.3-1.2 kb	Invitrogen
Human fetal liver	JG	8.6 × 10 ⁶	(>1 kb)	OriGene
Human fetal brain	JG	3.5 × 10 ⁶	0.5-1.2 kb (1.5 kb)	R. Brent, Clontech, Invitrogen, OriGene
Mouse brain	JG	6.1 × 10 ⁶	(>1 kb)	OriGene
Mouse brain	JG	4.5 × 10 ⁶	0.4-4.5 kb (1.2 kb)	Clontech
Mouse breast, lactating	JG	1.0 × 10 ⁷	0.4-3.1 kb	OriGene
Mouse breast, involuting	JG	1.0 × 10 ⁷	0.4-7.0 kb	OriGene
Mouse breast, virgin	JG	1.0 × 10 ⁷	0.4-5.5 kb	OriGene
Mouse breast, 12 days pregnant	JG	6.3 × 10 ⁶	0.4-5.3 kb	OriGene
Mouse skeletal muscle	JG	7.2 × 10 ⁶	0.4-3.5 kb	OriGene
Rat adipocyte, 9-week-old Zucker rat	JG	1.0 × 10 ⁷	0.4-5.0 kb	OriGene
Rat brain	JG	4.5 × 10 ⁶	0.3-3.4 kb	OriGene
Rat brain (day 18)	JG	—	—	H. Niu
Rat testis	JG	8.0 × 10 ⁶	(>1.2 kb)	OriGene
Rat thymus	JG	8.2 × 10 ⁶	(>1.3 kb)	OriGene
Mouse liver	JG	9.5 × 10 ⁶	(>1.4 kb)	OriGene
Mouse spleen	JG	1.0 × 10 ⁷	(>1 kb)	OriGene
Mouse ovary	JG	4.0 × 10 ⁶	(>1.2 kb)	OriGene
Mouse prostate	JG	—	—	OriGene
Mouse embryo, whole (19-day)	JG	1.0 × 10 ⁵	0.2-2.5 kb	OriGene
Mouse embryo	JG	3.6 × 10 ⁶	0.5-5 kb (1.7 kb)	Clontech
<i>Drosophila melanogaster</i> , adult, cDNA	JG	1.8 × 10 ⁶	(>1.0 kb)	OriGene
<i>D. melanogaster</i> , embryo, cDNA	JG	3.0 × 10 ⁶	0.5-3.0 kb (1.4 kb)	Clontech
<i>D. melanogaster</i> , 0-12 hr embryos, cDNA	JG	4.2 × 10 ⁶	0.5-2.5 kb (1.0 kb)	R. Brent
<i>D. melanogaster</i> , ovary, cDNA	JG	3.2 × 10 ⁶	0.3-1.5 kb (800 bp)	R. Brent
<i>D. melanogaster</i> , disc, cDNA	JG	4.0 × 10 ⁶	0.3-2.1 kb (900 bp)	R. Brent
<i>D. melanogaster</i> , head	JG	—	—	M. Rosbash

continued

Table 20.1.3 Libraries Compatible with the Interaction Trap System^a, continued

Source of RNA/DNA	Vector	Independent clones	Insert size (average) ^b	Contact information
Miscellaneous				
Synthetic aptamers	PJM-1	>1 × 10 ⁹	60 bp	R. Brent
<i>Saccharomyces cerevisiae</i> , S288C, genomic	JG	>3 × 10 ⁶	0.8-4.0 kb	R. Brent
<i>S. cerevisiae</i> , S288C, genomic	JG	4.0 × 10 ⁶	0.5-4.0 kb	OriGene
Sea urchin ovary	JG	3.5 × 10 ⁶	(1.7 kb)	Clontech
<i>Caenorhabditis elegans</i>	JG	3.8 × 10 ⁶	(>1.2 kb)	OriGene
<i>Agrobacterium tumefaciens</i>	JG	—	—	—
<i>Arabidopsis thaliana</i> , 7-day-old seedlings	JG	—	—	H.M. Goodman
Tomato (<i>Lycopersicon esculentum</i>)	JG	8 × 10 ⁶	—	G.B. Martin
<i>Xenopus laevis</i> embryo	JG	2.2 × 10 ⁶	0.3-4 kb (1.0 kb)	Clontech

^aMost libraries are constructed in either the pJG4-5 vector or the pYESTrp vector (JG or Y in the Vector column); the peptide aptamer library is made in the pJM-1 vector. Libraries available from the public domain were constructed by the following individuals: (1) J. Gyuris; (3) C. Sardet and J. Gyuris; (4) W. Kolanus, J. Gyuris, and B. Seed; (39) D. Krainc; (50-52) R. Finley; (55) P. Watt; (54) P. Colas, B. Cohen, T. Jessen, I. Grishina, J. McCoy, and R. Brent (Colas et al., 1996). All libraries mentioned above were constructed in conjunction with and are available from the laboratory of Roger Brent, (510) 647-0690 or brent@molsci.org. The following individual investigators must be contacted directly: (18) J. Pugh, Fox Chase Cancer Center, Philadelphia, Pa.; (8,9) Vinyaka Prasad, Albert Einstein Medical Center New York, N.Y.; (57, 58) Gregory B. Martin, gmartin@dept.agry.purdue.edu; (11) Huifeng Niu, hn34@columbia.edu; (16) Christer Hoog, christer.hoog@cmb.ki.se; (12) Kurt Gustin, kgus@umich.edu; (6) Stephan Witte, Stephan.Witte@nimbus.rz.uni-konstanz.de.

^bInsert size ranges for pJG4-5 based libraries originally constructed in the Brent laboratory, which are now commercially available from Clontech, were reestimated by the company.

Transform the library

- Grow an ~20-ml culture of EGY48 or EGY191 containing a *LexA*-operator-*lacZ* reporter plasmid and pBait in Glu/CM –Ura, –His liquid dropout medium overnight at 30°C.
For best results, the pBait and lacZ reporter plasmids should have been transformed into the yeast within ~7 to 10 days of commencing a screen.
- In the morning, dilute culture into 300 ml Glu/CM –Ura, –His liquid dropout medium to 2 × 10⁶ cell/ml (OD₆₀₀ = ~0.10). Incubate at 30°C until the culture contains ~1 × 10⁷ cells/ml (OD₆₀₀ = ~0.50).
- Centrifuge 5 min at 1000 to 1500 × g in a low-speed centrifuge at room temperature to harvest cells. Resuspend in 30 ml sterile water and transfer to 50-ml conical tube.
- Centrifuge 5 min at 1000 to 1500 × g. Decant supernatant and resuspend cells in 1.5 ml TE buffer/0.1 M lithium acetate.
- Add 1 μg library DNA in pJG4-5 and 50 μg high-quality sheared salmon sperm carrier DNA to each of 30 sterile 1.5-ml microcentrifuge tubes. Add 50 μl of the resuspended yeast solution from step 4 to each tube.

The total volume of library and salmon sperm DNA added should be <20 μl and preferably <10 μl.

A typical library transformation will result in 2 to 3 × 10⁶ primary transformants. Assuming a transformation efficiency of 10⁵/μg library DNA, this transformation requires a total of 20 to 30 μg library DNA and 1 to 2 mg carrier DNA. Doing transformations in small aliquots helps reduce the likelihood of contamination, and for reasons that are not clear, provides significantly better transformation efficiency than scaled-up versions.

Do not use excess transforming library DNA per aliquot of competent yeast cells because each competent cell may take up multiple library plasmids, complicating subsequent analysis.

6. Add 300 μ l of sterile 40% PEG 4000/0.1 M lithium acetate/TE buffer, pH 7.5, and invert to mix thoroughly. Incubate 30 min at 30°C.
7. Add DMSO to 10% (~40 μ l per tube) and invert to mix. Heat shock 10 min in 42°C heating block.
- 8a. *For 28 tubes:* Plate the complete contents of one tube per 24 \times 24-cm Glu/CM –Ura, –His, –Trp dropout plate and incubate at 30°C.
- 8b. *For two remaining tubes:* Plate 360 μ l of each tube on 24 \times 24-cm Glu/CM –Ura, –His, –Trp dropout plate. Use the remaining 40 μ l from each tube to make a series of 1/10 dilutions in sterile water. Plate dilutions on 100-mm Glu/CM –Ura, –His, –Trp dropout plates. Incubate all plates 2 to 3 days at 30°C until colonies appear.

The dilution series gives an idea of the transformation efficiency and allows an accurate estimation of the number of transformants obtained.

Collect primary transformant cells

Conventional replica plating (UNIT 13.3) does not work well in the selection process because so many cells are transferred to new plates that very high background levels inevitably occur. Instead, the procedure described below creates a slurry in which cells derived from $>10^6$ primary transformants are homogeneously dispersed. A precalculated number of these cells is plated for each primary transformant.

9. Cool all of the 24 \times 24-cm plates containing transformants for several hours at 4°C to harden agar.
10. Wearing gloves and using a sterile glass microscope slide, gently scrape yeast cells off the plate. Pool cells from the 30 plates into one or two sterile 50-ml conical tubes.

This is the step where contamination is most likely to occur. Be careful.

11. Wash cells by adding a volume of sterile TE buffer or water at least equal to the volume of the transferred cells. Centrifuge ~5 min at 1000 to 1500 \times g, room temperature, and discard supernatant. Repeat wash.

After the second wash, pellet volume should be ~25 ml cells derived from 1.5×10^6 transformants.

12. Resuspend pellet in 1 vol glycerol solution, mix well, and store up to 1 year in 1-ml aliquots at –70°C.

Determine replating efficiency

13. Remove an aliquot of frozen transformed yeast and dilute 1/10 with Gal/Raff/CM –Ura, –His, –Trp dropout medium. Incubate with shaking 4 hr at 30°C to induce the GAL promoter on the library.

Raffinose (Raff) aids in growth without diminishing transcription from the GAL1 promoter.

14. Make serial dilutions of the yeast cells using the Gal/Raff/CM –Ura, –His, –Trp dropout medium. Plate on 100-mm Gal/Raff/CM –Ura, –His, –Trp dropout plates and incubate 2 to 3 days at 30°C until colonies are visible.
15. Count colonies and determine the number of colony-forming units (cfu) per aliquot of transformed yeast.

In calculating yeast concentrations, it is useful to remember that 1 OD₆₀₀ unit = $\sim 2.0 \times 10^7$ yeast cells. In general, if the harvest is done carefully, viability will be greater than 90%. Some intrepid investigators perform this step simultaneously with plating out on Leu selective medium (steps 16 and 17).

Screen for interacting proteins

16. Thaw the appropriate quantity of transformed yeast based on the plating efficiency, dilute, and incubate as in step 13. Dilute cultures in Gal/Raff/CM –Ura, –His, –Trp, –Leu medium as necessary to obtain a concentration of 10^7 cells/ml ($OD_{600} = \sim 0.5$), and plate 100 μ l on each of as many 100-mm Gal/Raff/CM –Ura, –His, –Trp, –Leu dropout plates as are necessary for full representation of transformants. Incubate 2 to 3 days at 30°C until colonies appear.

Because not all cells that contain interacting proteins plate at 100% efficiency on –Leu medium (Estojak et al., 1995), it is desirable that for actual selection, each primary colony obtained from the transformation be represented on the selection plate by three to ten individual yeast cells. This will in some cases lead to multiple isolations of the same cDNA; however, because the slurry is not perfectly homogenous, it will increase the likelihood that all primary transformants are represented by at least one cell on the selective plate.

It is easiest to visually scan for Leu^+ colonies using cells plated at $\sim 10^6$ cfu per 100-mm plate. Plating at higher density can contribute to cross-feeding between yeast, resulting in spurious background growth. Thus, for a transformation in which 3×10^6 colonies are obtained, plate $\sim 2 \times 10^7$ cells on a total of 20 selective plates.

17. Carefully pick appropriate colonies to a new Gal/Raff/CM –Ura, –His, –Trp, –Leu master dropout plate. Incubate 2 to 7 days at 30°C until colonies appear.

A good strategy is to pick a master plate with colonies obtained on day 2, a second master plate (or set of plates) with colonies obtained on day 3, and a third with colonies obtained on day 4. Colonies from day 2 and 3 master plates should generally be characterized further. If many apparent positives are obtained, it may be worth making master plates of the much larger number of colonies likely to be obtained at day 4 (and after). See Critical Parameters and annotation to step 19 for additional information about appropriate colony selection for the master plate.

If no colonies appear within a week, those arising at later time points are likely to be artifactual. Contamination that has occurred at an earlier step (e.g., during plate scraping) is generally reflected by the growth of a very large number of colonies (>500 /plate) within 24 to 48 hr after plating on selective medium.

Some investigators omit use of a Gal/Raff/CM –Ura, –His, –Trp, –Leu master plate, restreaking directly to a Glu/CM –Ura, –His, –Trp master plate as in step 19.

Test for Gal dependence

The following steps test for Gal dependence of the Leu^+ insert and lacZ phenotypes to confirm that they are attributable to expression of the library-encoded proteins. The *GALI* promoter is turned off and –Leu selection eliminated before reinducing.

18. Restreak from the Gal/Raff/CM –Ura, –His, –Trp, –Leu master dropout plate to a 100-mm Glu/CM –Ura, –His, –Trp master dropout plate. Incubate overnight at 30°C until colonies form.
19. Restreak or replica plate from this plate to the following plates:

Glu/Xgal/CM –Ura, –His, –Trp
Gal/Raff/Xgal/CM –Ura, –His, –Trp
Glu/CM –Ura, –His, –Trp, –Leu
Gal/Raff/CM –Ura, –His, –Trp, –Leu.

At this juncture, colonies and the library plasmids they contain are tentatively considered positive if they are blue on Gal/Raff/Xgal plates but not blue or only faintly blue on Glu/Xgal plates, and if they grow on Gal/Raff/CM –Leu plates but not on Glu/CM –Leu plates.

The number of positives obtained will vary drastically from bait to bait. How they are processed subsequently will depend on the number initially obtained and on the preference

of the individual investigator. If none are obtained using EGY48 as reporter strain, it may be worth attempting to screen a library from an additional tissue source. If a relatively small number (≤ 30) are obtained, proceed to step 20. However, sometimes searches will yield large numbers of colonies (>30 to 300, or more). In this case, there are several options. The first option is to warehouse the majority of the positives and work up the first 30 that arise; those growing fastest are frequently the strongest interactors. These can be checked for specificity, and restriction digests can be used to establish whether they are all independent cDNAs or represent multiple isolates of the same, or a small number, of cDNAs. If the former is true, it may be advisable to repeat the screen in a less sensitive strain background, as obtaining many different interactors can be a sign of low-affinity nonspecific background. Alternatively, if initial indications are that a few cDNAs are dominating the positives obtained, it may be useful to perform a filter hybridization with yeast (see Support Protocol 3) using these cDNAs as a probe to establish the frequency of their identification and exclude future reisolation of these plasmids. The second major option is to work up large numbers of positives to get a complete profile of isolated interactors (see Support Protocol 4). A third option is to temporarily warehouse the entire results of this first screen, and repeat the screen with a less sensitive strain such as EGY191, on the theory that it is most important to get stronger interactors first and a complete profile of interactors later. Finally, some investigators prefer to work up the entire set of positives initially obtained, even if such positives number in the hundreds. Particularly in this latter case, it is most effective to use Alternate Protocol 1 as a means to identify unique versus common positives.

Isolate plasmid from positive colonies by transfer into *E. coli*

- 20a. Transfer yeast plasmids directly into *E. coli* by following the protocol for direct electroporation (UNIT 1.8, Alternate Protocol 2). Proceed to step 22.
- 20b. Isolate plasmid DNA from yeast by the rapid miniprep protocol (UNIT 13.11) with the following alteration: after obtaining aqueous phase, precipitate by adding sodium acetate to 0.3 M final and 2 vol ethanol, incubate 20 min on ice, microcentrifuge 15 min at maximum speed, wash pellet with 70% ethanol, dry, and resuspend in 5 μ l TE buffer.

Cultures can be grown prior to the miniprep using Glu/CM –Trp to select only for the library plasmid; this may increase the proportion of bacterial colonies that contain the desired plasmid.

21. Use 1 μ l DNA to electroporate (UNIT 1.8) into competent KC8 bacteria, and plate on LB/ampicillin plates. Incubate overnight at 37°C.

Electroporation must be used to obtain transformants with KC8 because the strain is generally refractory to transformation.

22. Restreak or replica plate colonies arising on LB/ampicillin plates to bacterial defined minimal A medium plates containing vitamin B1 and supplemented with Ura, His, and Leu but lacking Trp. Incubate overnight at 37°C.

Colonies that grow under these conditions contain the library plasmid.

The yeast TRP1 gene can successfully complement the bacterial trpC-9830 mutation, allowing the library plasmid to be easily distinguished from the other two plasmids contained in the yeast. It is helpful to first plate transformations on LB/ampicillin plates, which provides a less stringent selection, followed by restreaking to bacterial minimal medium to maximize the number of colonies obtained (E.G., unpub. observ.).

23. Purify library-containing plasmids using a bacterial miniprep procedure (UNIT 1.6).

Some investigators are tempted to immediately sequence DNAs obtained at this stage. At this point, it is still possible that none of the isolated clones will express bona fide interactors, and it is suggested that the following specificity tests be completed before committing the effort to sequencing (also see annotation to step 28).

Because multiple 2 μ m plasmids with the same marker can be simultaneously tolerated in yeast, it sometimes happens that a single yeast will contain two or more different library plasmids, only one of which encodes an interacting protein. The frequency of this occurrence varies in the hands of different investigators and may in some cases account for disappearing positives if the wrong cDNA is picked. When choosing colonies to miniprep, it is generally useful to work up at least two individual bacterial transformants for each yeast positive. These minipreps can then be restriction digested (UNIT 3.1) with EcoRI + XhoI to release cDNA inserts, and the size of inserts determined on an agarose minigel (UNIT 2.5A) to confirm that both plasmids contain the same insert. An additional benefit of analyzing insert size is that it may provide some indication as to whether repeated isolation of the same cDNA is occurring, generally a good indication concerning the biological relevance of the interactor. See Background Information for further discussion.

Assess positive colonies with specificity tests

Much spurious background will have been removed by the previous series of controls. Other classes of false positives can be eliminated by retransforming purified plasmids into “virgin” LexA-operator-LEU2/LexA-operator-lacZ/pBait-containing strains that have not been subjected to Leu selection and verifying that interaction-dependent phenotypes are still observed. Such false positives could include mutations in the initial EGY48 yeast that favor growth on Gal medium, library-encoded cDNAs that interact with the LexA DNA-binding domain, or proteins that are sticky and interact with multiple biologically unrelated fusion domains.

24. In separate transformations, use purified plasmids from step 23 to transform yeast that already contain the following plasmids and are growing on Glu/CM –Ura, –His plates:

EGY48 containing pSH18-34 and pBait
EGY48 containing pSH18-34 and pRFHM-1
EGY48 containing pSH18-34 and a nonspecific bait (optional).

25. Plate each transformation mix on Glu/CM –Ura, –His, –Trp dropout plates and incubate 2 to 3 days at 30°C until colonies appear.
26. Create a Glu/CM –Ura, –His, –Trp master dropout plate for each library plasmid being tested. Streak adjacently five or six independent colonies derived from each of the transformation plates. Incubate overnight at 30°C.
27. Restreak or replica plate from this master dropout plate to the same series of test plates used for the actual screen:

Glu/Xgal/CM –Ura, –His, –Trp
Gal/Raff/Xgal/CM –Ura, –His, –Trp
Glu/CM –Ura, –His, –Trp, –Leu
Gal/CM –Ura, –His, –Trp, –Leu.

True positive cDNAs should make cells blue on Gal/Raff/Xgal but not on Glu/Xgal plates, and should make them grow on Gal/Raff/CM –Leu but not Glu/CM –Leu dropout plates only if the cells contain LexA-bait. cDNAs that meet such criteria are ready to be sequenced (see legend to Fig. 20.1.3 for primer sequence) or otherwise characterized. Those cDNAs that also encode proteins that interact with either RFHM-1 or another nonspecific bait should be discarded.

It may be helpful to cross-check the isolated cDNAs with a database of cDNAs thought to be false positives. This database is available on the World Wide Web as a work in progress at <http://www.fccc.edu:80/research/labs/golemis/InteractionTrapInWork.html>. cDNAs reported to this database are generally those isolated only once in a screen in which obviously true interactive partners were isolated multiple times, cDNAs that may interact with more than one bait, or cDNAs for which the interaction does not appear to make biological sense

in the context of the starting bait. Although some proteins in this database may ultimately turn out in fact to associate with the bait that isolated them, they are by default unlikely to possess a unique and interesting function in the context of that bait if they are well represented in the database.

28. If appropriate, conduct additional specificity tests (see Support Protocol 5). Analyze and sequence positive isolates.

The primer sequence for use with pJG4-5 is provided in the legend to Figure 20.1.4.

DNA prepared from KC8 is generally unsuitable for dideoxy or automated sequencing even after use of Qiagen columns and/or cesium chloride gradients. Library plasmids to be sequenced should be retransformed from the KC8 miniprep stock (step 23) to a more amenable strain, such as DH5 α , before sequencing is attempted.

RAPID SCREEN FOR INTERACTION TRAP POSITIVES

Under some circumstances, it may be desirable to attempt the analysis of a large number of positives resulting from a two-hybrid screen. One such hypothetical example would be a bait with a leucine zipper or coiled coil known to dimerize with partner “A” that is highly expressed. In order to identify the rare novel partner “B”, it is necessary to work through the high background of “A” reisolates. This protocol uses the polymerase chain reaction (PCR) in a strategy to sort positives into redundant (multiple isolates) and unique classes prior to plasmid rescue from yeast, thus greatly reducing the number of plasmid isolations that must be performed. An additional benefit is that this protocol preidentifies positive clones containing one or multiple library plasmids; for those containing only one library plasmid, only a single colony needs to be prepared through KC8/DH5 α .

Additional Materials (also see Basic Protocol 2)

Yeast plated on Glu/CM –Ura, –His, –Trp master plate (see Basic Protocol 2, step 19)

Lysis solution (see recipe)

10 μ M forward primer (FP1): 5'-CGT AGT GGA GAT GCC TCC-3'

10 μ M reverse primer (FP2): 5'-CTG GCA AGG TAG ACA AGC CG-3'

Toothpicks or bacterial inoculating loops (UNIT 1.1), sterile

96-well microtiter plate

Sealing tape, e.g., wide transparent tape

150- to 212- μ m glass beads, acid-washed (UNIT 13.13)

Vortexer with flat plate

Additional reagents and equipment for performing an interactor hunt (see Basic Protocol 2), PCR amplification of DNA (UNIT 15.1), agarose gel electrophoresis (UNITS 2.5A & 2.6), restriction endonuclease digestion (UNIT 3.1), electroporation (UNIT 1.8), and miniprep isolation of bacterial DNA (UNIT 1.6)

1. Perform an interactor hunt (see Basic Protocol 2, steps 1 to 19).
2. Use a sterile toothpick or bacterial inoculating loop to transfer yeast from the Glu/CM, –Ura, –His, –Trp master plate into 25 μ l lysis solution in a 96-well microtiter plate. Seal the wells of the microtiter plate with sealing tape and incubate 1.5 to 3.5 hr at 37°C with shaking.

The volume of yeast transferred should not exceed ~2 to 3 μ l of packed pellet; larger quantities of yeast will reduce quality of the DNA. DNA can be efficiently recovered from master plates that have been stored up to 1 week at 4°C. If yeast have been previously gridded on master plates, transfer to microtiter plates can be facilitated by using a multicolony replicator.

ALTERNATE PROTOCOL 1

Analysis of Protein Interactions

20.1.21

3. Remove tape from the plate, add ~25 μ l acid-washed glass beads to each well, and reseal with the same tape. Firmly attach the microtiter plate to a flat-top vortexer, and vortex 5 min at medium-high power.

The microtiter plate can be attached to the vortexer using 0.25-in (0.64-cm) rubber bands.

4. Remove the tape and add ~100 μ l sterile water to each well. Swirl gently to mix, then remove sample for step 5. Press the tape back firmly to seal the microtiter plate and place in the freezer at -20°C for storage.
5. Amplify 0.8 to 2.0 μ l of sample by standard PCR (UNIT 15.1) in a ~30- μ l volume using 3 μ l each of the forward primer FP1 and the reverse primer FP2. Perform PCR using the following cycles:

Initial step:	2 min	94 $^{\circ}\text{C}$
31 cycles:	45 sec	94 $^{\circ}\text{C}$
	45 sec	56 $^{\circ}\text{C}$
	45 sec	72 $^{\circ}\text{C}$.

These conditions have been used successfully to amplify fragments up to 1.8 kb in length; some modifications, such as extension of elongation time, are also effective.

6. Load 20 μ l of the PCR reaction product on a 0.7% low melting temperature agarose gel (UNIT 2.6) to resolve PCR products. Based on insert sizes, group the obtained interactors in families, i.e., potential multiple independent isolates of identical cDNAs. Reserve gel until results of step 7 are obtained.

*No special precautions are needed for storing the gel. Since *HaeIII* digests typically yield rather small DNA fragments, running the second gel does not take a lot of time. Usually, the delay does not exceed 45 to 60 min, during which time the first gel may be stored in a gel box at room temperature or wrapped in plastic wrap at 4 $^{\circ}\text{C}$.*

7. While the gel is running, use the remaining 10 μ l of PCR reaction product for a restriction endonuclease digestion with *HaeIII* in a digestion volume of ~20 μ l (UNIT 3.1). Based on analysis of the sizes of undigested PCR products in the gel (step 6), rearrange the tubes with *HaeIII* digest samples so that those thought to represent a family are side by side. Resolve the digests on a 2% to 2.5% agarose gel (UNIT 2.5A).

Most restriction fragments will be in the 200-bp to 1.0-kb size range so using a long gel run is advisable. This analysis should produce a distinct fingerprint of insert sizes and allow definition of library cDNAs as unique isolates or related groups.

A single positive yeast will sometimes contain multiple library plasmids. An advantage of this protocol is the ready detection of multiple library plasmids in PCR reactions; thus, following subsequent bacterial transformations, only a single TRP1 colony would need to be analyzed unless multiple plasmids were already known to be present.

8. Isolate DNA fragments from the low melting temperature agarose gel (step 6).

If inspection of the banding pattern on the two gels suggests that a great many reisolates of a small number of cDNAs are present, it may be worthwhile to immediately sequence PCR products representative of these clusters, but it is generally still advisable to continue through specificity tests before doing so. If the PCR products are sequenced, the FP1 forward primer works well in automated sequencing of PCR fragments, but the FP2 primer is only effective in sequencing from purified plasmid.

In general, priming from the AT-rich ADH terminator downstream of the polylinker/cDNA in library plasmid is less efficient than from upstream of the cDNA, and it is hard to design effective primers in this region.

9. Remove the microtiter plate of lysates from the freezer, thaw it, and remove 2 to 4 μ l of lysed yeast for each desired positive. Electroporate DNA into either DH5 α or KC8 *E. coli* as appropriate, depending on the choice of bait and reporter plasmids (see

Table 20.1.1 and see Background Information for further information). Refreeze the plate as a DNA reserve in case bacteria fail to transform on the first pass.

KC8 E. coli should be used for electroporation when the original reagents pEG202/pJG4-5/pJK101 are used for the interaction trap.

An additional strength of this protocol is that it identifies redundant clones before transfer of plasmids to bacteria, thus reducing the amount of work required in cases where plasmid identity can be unambiguously assigned. However, although restriction endonuclease digestion and PCR analysis are generally highly predictive, they are not 100% certain methods for estimating cDNA identity. Thus, if there is any doubt about whether two cDNAs are the same, investigators are urged to err on the side of caution.

10. Prepare a miniprep of plasmid DNA from the transformed bacteria (*UNIT 1.6*) and perform yeast transformation and specificity assessment (see Basic Protocol 2, steps 24 to 28).

PERFORMING A HUNT BY INTERACTION MATING

An alternative way of conducting an interactor hunt is to mate a strain that expresses the bait protein with a strain that has been pretransformed with the library DNA, and screen the resulting diploid cells for interactors (Bendixen et al., 1994; Finley and Brent, 1994). This “interaction mating” approach can be used for any interactor hunt, and is particularly useful in three special cases. The first case is when more than one bait will be used to screen a single library. Interaction mating allows several interactor hunts with different baits to be conducted using a single high-efficiency yeast transformation with library DNA. This can be a considerable savings, since the library transformation is one of the most challenging tasks in an interactor hunt. The second case is when a constitutively expressed bait interferes with yeast viability. For such baits, performing a hunt by interaction mating avoids the difficulty associated with achieving a high-efficiency library transformation of a strain expressing a toxic bait. Moreover, the actual selection for interactors will be conducted in diploid yeast, which are more vigorous than haploid yeast and can better tolerate expression of toxic proteins. The third case is when a bait cannot be used in a traditional interactor hunt using haploid yeast strains (see Basic Protocol 2) because it activates transcription of even the least sensitive reporters. In diploids the reporters are less sensitive to transcription activation than they are in haploids. Thus, the interaction mating hunt provides an additional method to reduce background from transactivating baits.

In the protocol described below, the library DNA is used to transform a strain with a *LEU2* reporter (e.g., EGY48). This pretransformed library strain is then frozen in many aliquots, which can be thawed and used for individual interactor hunts. The bait is expressed in a strain of mating type opposite to that of the pretransformed library strain, and also bearing the *lacZ* reporter. A hunt is conducted by mixing the pretransformed library strain with the bait strain and allowing diploids to form on YPD medium overnight. The diploids are then induced for expression of the library-encoded proteins and screened for interactors as in Basic Protocol 2.

NOTE: Strain combinations other than those described below can also be used in an interaction-mating hunt. The key to choosing the strains is to ensure that the bait and prey strains are of opposite mating types and that both have auxotrophies to allow selection for the appropriate plasmids and reporter genes. Also, once the bait plasmid and *lacZ* reporter plasmid have been introduced into the bait strain, and the library plasmids have been introduced into the library strain, the resulting bait strain and library strain must each have auxotrophies that can be complemented by the other, so that diploids can be selected.

ALTERNATE PROTOCOL 2

Additional Materials (also see *Basic Protocols 1 and 2*)

Yeast strains: either RFY206 (Finley and Brent, 1994), YPH499 (Sikorski and Hieter, 1989; ATCC #6625), or an equivalent *MATa* strain with auxotrophic markers *ura3*, *trp1*, *his3*, and *leu2*

YPD liquid medium (*UNIT 13.1*)

Glu/CM –Trp plates: CM dropout plates –Trp (*UNIT 13.1*) supplemented with 2% glucose

pJG4-5 library vector (Fig. 20.1.6), empty

100-mm YPD plates (*UNIT 13.1*)

Additional reagents and equipment for lithium acetate transformation of yeast (*UNIT 13.7*)

Construct the bait strain

The bait strain will be a *MATa* yeast strain (mating type opposite of EGY48) containing a *lacZ* reporter plasmid like pSH18-34 and the bait-expressing plasmid, pBait.

1. Perform construction of the bait plasmid (pBait; see Basic Protocol 1, step 1).
2. Cotransform the *MATa* yeast strain (e.g., either RFY206 or YPH499) with pBait and pSH18-34 using the lithium acetate method (*UNIT 13.7*). Select transformants on Glu/CM –Ura,–His plates by incubating plates at 30°C for 3 to 4 days until colonies form. Combine 3 colonies for all future tests and for the mating hunt.

The bait strain (RFY206/pSH18-34/pBait or YPH499/pSH18-34/pBait) can be tested by immunoblotting to ensure that the bait protein is expressed (see Support Protocol 1). Synthesis and nuclear localization of the bait protein can also be tested by the repression assay (see Basic Protocol 1, steps 8 to 12).

3. *Optional:* Assay *lacZ* gene activation in the bait strain (see Basic Protocol 1, steps 4 to 7).

If the bait activates the lacZ reporter, a less sensitive lacZ reporter plasmid (Table 20.1.1), or an integrated version of the lacZ reporter should be tried. A bait that strongly activates the lacZ reporters usually cannot be used in a hunt based on selection of interactors with the LEU2 reporter, because the LEU2 reporters are more sensitive than the lacZ reporters. However, both reporters are less sensitive to activation by a bait in diploid cells, as compared to haploid cells. Thus, a more important test of the transactivation potential of a bait is to test the leucine requirement of diploid cells expressing it, as described in steps 6 to 20, below.

Prepare the pretransformed library strain (EGY48 + library plasmids)

4. Perform a large-scale transformation of EGY48 with library DNA using the lithium acetate method (see Basic Protocol 2, steps 1 to 8, except start with EGY48 bearing no other plasmids). To prepare for transformation, grow EGY48 in YPD liquid medium. Select library transformants on Glu/CM –Trp plates by incubating 3 days at 30°C.
5. Collect primary transformants by scraping plates, washing yeast, and resuspending in 1 pellet vol glycerol solution (see Basic Protocol 2, steps 9 to 12). Freeze 0.2 to 1.0 ml aliquots at –70° to –80°C.

The cells will be stable for at least 1 year. Refreezing a thawed aliquot will result in loss of viability. Thus, many frozen aliquots should be made, so that each thawed aliquot can be discarded after use.

Prepare the pretransformed control strain (EGY48 + pJG4-5)

6. Transform EGY48 grown in YPD liquid medium with the empty library vector, pJG4-5, using the lithium acetate method (UNIT 13.7). Select transformants on Glu/CM –Trp plates by incubating 3 days at 30°C.
7. Pick and combine three transformant colonies and use them to inoculate 30 ml of Glu/CM –Trp medium. Incubate 15 to 24 hr at 30°C (to $OD_{600} > 3$).
8. Centrifuge 5 min at 1000 to 1500 $\times g$, room temperature, and remove supernatant. Resuspend in 10 ml sterile water to wash cells.
9. Centrifuge 5 min at 1000 to 1500 $\times g$, room temperature, and remove supernatant. Resuspend in 1 pellet vol glycerol solution and freeze 100- μ l aliquots at –70° to –80°C.

Determine plating efficiency of pretransformed library and pretransformed control strains

10. After freezing (at least 1 hr) thaw an aliquot of each pretransformed strain (from step 5 and step 9) at room temperature. Make several serial dilutions in sterile water, including aliquots diluted 10⁵-fold, 10⁶-fold, and 10⁷-fold. Plate 100 μ l of each dilution on 100-mm Glu/CM –Trp plates and incubate 2 to 3 days at 30°C.
11. Count the colonies and determine the number of colony-forming units (cfu) per aliquot of transformed yeast.

The plating efficiency for a typical library transformation and for the control strain will be $\sim 1 \times 10^8$ cfu per 100 μ l.

Mate the bait strain with the pretransformed library strain and the pretransformed control strain

In steps 12 through 20, an interactor hunt is conducted concurrently with testing *LEU2* reporter activation by the bait itself. For most baits, this approach will be the quickest way to isolate interactors. However, for some baits, such as those that have a high transactivation potential, or those that affect yeast mating or growth, steps 12 through 20 will serve as a pilot experiment to determine the optimal parameters for a subsequent hunt.

12. Grow a 30-ml culture of the bait strain in Glu/CM –Ura, –His liquid dropout medium to mid to late log phase ($OD_{600} = 1.0$ to 2.0, or 2 to 4 $\times 10^7$ cells/ml).

A convenient way to grow the bait strain is to inoculate a 5-ml culture with approximately three colonies from a plate and grow it overnight at 30°C with shaking. In the morning, measure the OD_{600} , dilute into a 30-ml culture to a final $OD_{600} = 0.2$, and grow at 30°C with shaking. The culture should reach mid to late log phase before the end of the day.

13. Centrifuge the culture 5 min at 1000 to 1500 $\times g$, room temperature, to harvest cells. Resuspend the cell pellet in sterile water to make a final volume of 1 ml.

This should correspond to $\sim 1 \times 10^9$ cells/ml.

14. Set up two matings. In one sterile microcentrifuge tube mix 200 μ l of the bait strain with 200 μ l of a thawed aliquot of the pretransformed control strain from step 9. In a second microcentrifuge tube mix 200 μ l of the bait strain with $\sim 1 \times 10^8$ cfu (~ 0.1 to 1 ml) of the pretransformed library strain from step 5.

The library mating should be set up so that it contains a ~ 2 -fold excess of bait strain cfu over pretransformed library strain cfu. Because the bait strain was harvested in log phase, most of the cells will be viable (i.e., cells/ml = \sim cfu/ml), and the number of cfu can be sufficiently estimated from optical density (1 $OD_{600} = \sim 2 \times 10^7$ cells/ml). Under these conditions, $\sim 10\%$ of the cfu in the pretransformed library strain will mate with the bait

strain. Thus, a complete screen of 10^7 library transformants will require a single mating with at least 10^8 cfu of the pretransformed library strain and at least 2×10^8 cfu of the bait strain.

To screen more library transformants, set up additional matings. The number of pretransformed library transformants to screen depends on the size of the library and the number of primary transformants obtained in step 5. If the size of the library is larger than the number of transformants obtained in step 5, the goal will be to screen all of the yeast transformants. In this case, complete screening of the library will require additional transformations of EGY48 and additional interactor hunts. If the size of the library is smaller than the number of transformants obtained in step 5, the goal will be to screen at least a number of transformants equivalent to the size of the library.

15. Centrifuge each cell mixture for 5 min at 1000 to $1500 \times g$, pour off medium, and resuspend cells in 200 μ l YPD medium. Plate each suspension on a 100-mm YPD plate. Incubate 12 to 15 hr at 30°C.
16. Add ~1 ml of Gal/Raff/CM –Ura, –His, –Trp to the lawns of mated yeast on each plate. Mix the cells into the medium using a sterile applicator stick.
17. Transfer each slurry of mated cells to a 500-ml flask containing 100 ml of Gal/Raff/CM –Ura, –His, –Trp dropout medium. Incubate with shaking 6 hr at room temperature to induce the *GAL1* promoter, which drives expression of the cDNA library.
18. Centrifuge the cell suspensions 5 min at 1000 to $1500 \times g$, room temperature, to harvest the cells. Wash by resuspending in 30 ml of sterile water and centrifuging again. Resuspend each pellet in 5 ml sterile water. Measure OD₆₀₀ and, if necessary, dilute to a final concentration of $\sim 1 \times 10^8$ cells/ml.

This is a mixture consisting of haploid cells that have not mated and diploid cells. Under a microscope, the two cell types can be distinguished by size (diploids are $\sim 1.7\times$ bigger than haploids) and shape (diploids are slightly oblong and haploids are spherical). Because diploids grow faster than haploids, this mixture will contain $\sim 10\%$ to 50% diploid cells. The actual number of diploids will be determined by plating dilutions on –Ura, –His, –Trp medium, which will not support the growth of the parental haploids.

19. For each mating make a series of $1/10$ dilutions in sterile water, at least 200 μ l each, to cover a 10^6 -fold concentration range. Plate 100 μ l from each tube (undiluted, 10^{-1} , 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , and 10^{-6} dilution) on 100-mm Gal/Raff/CM –Ura, –His, –Trp, –Leu plates. Plate 100 μ l from the 10^{-4} , 10^{-5} , and 10^{-6} tubes on 100-mm Gal/Raff/CM –Ura, –His, –Trp plates. Incubate plates at 30°C. Count the colonies on each plate after 2 to 5 days.
20. For the mating with the pretransformed library, prepare an additional 3 ml of a 10^{-1} dilution. Plate 100 μ l of the 10^{-1} dilution on each of 20 100-mm Gal/Raff/CM –Ura, –His, –Trp, –Leu plates. Also plate 100 μ l of the undiluted cells on each of 20 100-mm Gal/Raff/CM –Ura, –His, –Trp, –Leu plates. Incubate at 30°C. Pick Leu⁺ colonies after 2 to 5 days and characterize them beginning with step 17 of Basic Protocol 2.

The number of Leu⁺ colonies to pick to ensure that all of the pretransformed library has been screened depends on the transactivation potential of the bait protein itself. The transactivation potential is expressed as the number of Leu⁺ colonies that grow per cfu (Leu⁺/cfu) of the bait strain mated with the control strain, as determined in step 19 of this protocol. It can be calculated as the ratio of the number of colonies that grow on Gal/Raff/CM –Ura, –His, –Trp, –Leu to the number of colonies that grow on Gal/Raff/CM –Ura, –His, –Trp for a given dilution of the mating between the bait strain and the control strain. A bait with essentially no transactivation potential will produce less than 10^{-6} Leu⁺/cfu. For a bait to be useful in an interactor hunt it should not transactivate more than 10^{-4} Leu⁺/cfu.

To screen all of the pretransformed library, it will be necessary to pick a sufficient number of Leu^+ colonies in addition to background colonies produced by the transactivation potential of the bait itself. Thus, the minimum number of Leu^+ colonies that should be picked in step 20 of this protocol is given by:

$$(\text{transactivation potential, } Leu^+/\text{cfu}) \times (\# \text{ library transformants screened}).$$

For example, if 10^7 library transformants were obtained in step 2 (and at least 10^8 cfu of these transformants were mated with the bait strain in step 14, since only ~10% will form diploids), and the transactivation potential of the bait is 10^{-4} Leu^+ /cfu, then at least 1000 Leu^+ colonies must be picked and characterized. In other words, if the rarest interactor is present in the pretransformed library at a frequency of 10^{-7} , to find it one needs to screen through at least 10^7 diploids from a mating of the library strain. However, at least 1000 of these 10^7 diploids would be expected to be Leu^+ due to the bait background if the transactivation potential of the bait is 10^{-4} . The true positives will be distinguished from the bait background in the next step by the galactose dependence of their Leu^+ and $lacZ^+$ phenotypes.

PREPARATION OF PROTEIN EXTRACTS FOR IMMUNOBLOT ANALYSIS

To confirm that the bait fusion protein constructed in Basic Protocol 1 is synthesized properly, a crude lysate is prepared for SDS-PAGE and immunoblot analysis (UNITS 10.2 & 10.8). The presence of the target protein is detected by antibody to LexA or the fusion domain.

Materials

Master plates with pBait-containing positive and control yeast on Glu/CM –Ura, –His dropout medium (see Basic Protocol 1, step 4)

Glu/CM –Ura, –His dropout liquid medium: CM dropout plates –Ura, –His (UNIT 13.1) supplemented with 2% glucose

2× Laemmli sample buffer (see recipe)

Antibody to fusion domain or LexA: monoclonal antibody to LexA (Clontech, Invitrogen) or polyclonal antibody to LexA (available by request from R. Brent or E. Golemis)

30°C incubator

100°C water bath

Additional reagents and equipment for SDS-PAGE (UNIT 10.2) and immunoblotting and immunodetection (UNIT 10.8)

1. From the master plates, start a 5-ml culture in Glu/CM –Ura, –His liquid medium for each bait being tested and for a positive control for protein expression (i.e., RFHMI or SH17-4). Incubate overnight at 30°C.

For each construct assayed, it is a good idea to grow colonies from at least two primary transformants, as levels of bait expression are sometimes heterogenous.

2. From each overnight culture, start a fresh 5-ml culture in Glu/CM –Ura, –His at $OD_{600} = \sim 0.15$. Incubate again at 30°C.
3. When the culture has reached $OD_{600} = 0.45$ to 0.7 (~4 to 6 hr), remove 1.5 ml to a microcentrifuge tube.

For some LexA fusion proteins, levels of the protein drop off rapidly in cultures approaching stationary phase. This is due to a combination of the diminishing activity of the ADHI promoter in late growth phases and the relative instability of particular fusion domains. Thus, it is not a good idea to let cultures become saturated in the hopes of obtaining a higher yield of protein.

SUPPORT PROTOCOL 1

Analysis of Protein Interactions

20.1.27

4. Microcentrifuge cells 3 min at $13,000 \times g$, room temperature. When the pellet is visible, remove the supernatant.

Inspection of the tube should reveal a pellet ~1 to 3 μ l in volume. If the pellet is not visible, microcentrifuge another 3 min.

5. Working rapidly, add 50 μ l of $2\times$ Laemmli sample buffer to the visible pellet in the tube, vortex, and place the tube on dry ice.

Samples may be frozen at -70°C .

6. Transfer frozen sample directly to a boiling water bath or a PCR machine set to cycle at 100°C . Boil 5 min.

7. Microcentrifuge 5 sec at maximum speed to pellet large cellular debris.

8. Perform SDS-PAGE (UNIT 10.2) using 20 to 50 μ l sample per lane.

9. To detect the protein, immunoblot and analyze (UNIT 10.8) using antibody to the fusion domain or LexA.

SUPPORT PROTOCOL 2

PREPARATION OF SHEARED SALMON SPERM CARRIER DNA

This protocol generates high-quality sheared salmon sperm DNA (sssDNA) for use as carrier in transformation (Basic Protocol 2). This DNA is also suitable for other applications where high-quality carrier DNA is needed (e.g., hybridization). This protocol is based on Schiestl and Gietz (1989). For more details of phenol extraction or other DNA purification methods, consult UNIT 2.1A.

Materials

High-quality salmon sperm DNA (e.g., sodium salt from salmon testes, Sigma or Boehringer Mannheim), desiccated
TE buffer, pH 7.5 (APPENDIX 2), sterile
TE-saturated buffered phenol (UNIT 2.1A)
1:1 (v/v) buffered phenol/chloroform
Chloroform
3 M sodium acetate, pH 5.2 (APPENDIX 2)
100% and 70% ethanol, ice cold
Magnetic stirring apparatus and stir-bar, 4°C
Sonicator with probe
50-ml conical centrifuge tube
High-speed centrifuge and appropriate tube
 100°C and ice-water baths

1. Dissolve desiccated high-quality salmon sperm DNA in TE buffer, pH 7.5, at a concentration of 5 to 10 mg/ml by pipetting up and down in a 10-ml glass pipet. Place in a beaker with a stir-bar and stir overnight at 4°C to obtain a homogenous viscous solution.

It is important to use high-quality salmon sperm DNA. Sigma Type III sodium salt from salmon testes has worked well, as has a comparable grade from Boehringer Mannheim. Generally it is convenient to prepare 20- to 40-ml batches at a time.

2. Shear the DNA by sonicating briefly using a large probe inserted into the beaker.

The goal of this step is to generate sheared salmon sperm DNA (sssDNA) with an average size of 7 kb, but ranging from 2 to 15 kb. Oversonication (such that the average size is closer to 2 kb) drastically decreases the efficacy of carrier in enhancing transformation. The original version of this protocol (Schiestl and Gietz, 1989) called for two 30-sec pulses at

Table 20.1.4 Two-hybrid System Variants^a

System	DNA-binding domain	Activation domain	Selection	Reference
Two-hybrid	<i>GAL4</i>	<i>GAL4</i>	Activation of <i>lacZ</i> , <i>HIS3</i>	Chien et al., 1991
Interaction trap	<i>LexA</i>	<i>B42</i>	Activation of <i>LEU2</i> , <i>lacZ</i>	Gyuris et al., 1993
“Improved two-hybrid”	<i>GAL4</i>	<i>GAL4</i>	Activation of <i>HIS3</i> , <i>lacZ</i>	Durfee et al., 1993
Modified two-hybrid	<i>LexA</i>	VP16	Activation of <i>HIS3</i> , <i>lacZ</i>	Vojtek et al., 1993
KISS	<i>GAL4</i>	VP16	Activation of <i>CAT</i> , <i>hyg^r</i>	Fearon et al., 1992
Contingent replication	<i>GAL4</i>	VP16	Activation of T-Ag, replication of plasmids	Vasavada et al., 1991

^aAbbreviations: *CAT*, chloramphenicol transferase gene; *hyg^r*, hygromycin resistance gene; T-Ag, viral large T antigen.

three-quarter power, but optimal conditions vary between sonicators. The first time this protocol is performed, it is worthwhile to sonicate briefly, then test the size of the DNA by running out a small aliquot alongside molecular weight markers on an agarose gel containing ethidium bromide. The DNA can be sonicated further if needed.

- Once DNA of the appropriate size range has been obtained, extract the sssDNA solution with an equal volume of TE-saturated buffered phenol in a 50-ml conical tube, shaking vigorously to mix.
- Centrifuge 5 to 10 min at 3000 × g, room temperature, or until clear separation of phases is obtained. Transfer the upper phase containing the DNA to a clean tube.
- Repeat extraction using 1:1 (v/v) buffered phenol/chloroform, then chloroform alone. Transfer the DNA into a tube suitable for high-speed centrifugation.
- Precipitate the DNA by adding 1/10 vol of 3 M sodium acetate and 2.5 vol of ice-cold 100% ethanol. Mix by inversion. Centrifuge 15 min at ~12,000 × g, room temperature.
- Wash the pellet with 70% ethanol. Briefly dry either by air drying, or by covering one end of the tube with Parafilm with a few holes poked in and placing the tube under vacuum. Resuspend the DNA in sterile TE buffer at 5 to 10 mg/ml.

Do not overdry the pellet or it will be very difficult to resuspend.

- Denature the DNA by boiling 20 min in a 100°C water bath. Then immediately transfer the tube to an ice-water bath.
- Place aliquots of the DNA in microcentrifuge tubes and store frozen at –20°C. Thaw as needed.

DNA should be boiled again briefly (5 min) immediately before addition to transformations.

Before using a new batch of sssDNA in a large-scale library transformation, it is a good idea to perform a small-scale transformation using suitable plasmids to determine the transformation efficiency. Optimally, use of sssDNA prepared in the manner described will yield transformation frequencies of >10⁵ colonies/μg input plasmid DNA.

YEAST COLONY HYBRIDIZATION

This protocol is adapted from a modification of the classic protocol of Grunstein and Hogness (1975; Kaiser et al., 1994). It is primarily useful when a large number of putative interactors has been obtained, and initial minipreps and restriction digests have indicated that many of them derive from a small number of cDNAs; these cDNAs can then be used as probes to screen and eliminate identical cDNAs from the pool.

Materials

Glu/CM –Trp plates: CM dropout plates –Trp (*UNIT 13.1*) supplemental with 2% glucose

Master dropout plate of yeast positive for Gal dependence (see Basic Protocol 2, step 18)

1 M sorbitol/20 mM EDTA/50 mM DTT (prepare fresh)

1 M sorbitol/20 mM EDTA

0.5 M NaOH

0.5 M Tris·Cl (pH 7.5)/6× SSC (*APPENDIX 2*)

2× SSC (*APPENDIX 2*)

100,000 U/ml β-glucuronidase (type HP-2 crude solution from *Helix pomatia*; Sigma)

82-mm circular nylon membrane, sterile

Whatman 3MM paper

80°C vacuum oven *or* UV cross-linker

Additional reagents and equipment for bacterial filter hybridization (*UNITS 6.3 & 6.4*)

1. Place a sterile nylon membrane onto a Glu/CM –Trp dropout plate. From the master dropout plate of Gal-dependent positives, gently restreak positives to be screened onto the membrane and mark the membrane to facilitate future identification of hybridizing colonies. Grow overnight (~12 hr) at 30°C.

Growth for extended periods of time (i.e., 24 hr) may result in difficulty in obtaining good lysis. It is a good idea to streak positive and negative controls for the cDNAs to be hybridized on the membrane.

2. Remove membrane from plate. Air dry briefly. Incubate ~30 min on a sheet of Whatman 3MM paper saturated with 1 M sorbitol/20 mM EDTA/50 mM DTT.

Optionally, before commencing chemical lysis, membranes can be placed at –70°C for 5 min, then thawed at room temperature for one or more cycles to enhance cell wall breakage.

3. Cut a piece of Whatman 3MM paper to fit inside a 100-mm petri dish. Place the paper disc in the dish and saturate with 100,000 U/ml β-glucuronidase diluted 1:500 in 1 M sorbitol/20 mM EDTA (2 μl glucuronidase per ml of sorbitol/EDTA to give 200 U/ml final). Layer nylon membrane on dish, cover dish, and incubate up to 6 hr at 37°C until >80% of the cells lack a cell wall.

The extent of cell wall removal can be determined by removing a small quantity of cells from the filter to a drop of 1 M sorbitol/20 mM EDTA on a microscope slide and observing directly with a phase-contrast microscope at ≥60× magnification. Cells lacking cell wall are nonrefractile.

4. Place membrane on Whatman 3MM paper saturated with 0.5 M NaOH for ~8 to 10 min.
5. Place membrane on Whatman 3MM paper saturated with 0.5 M Tris·Cl (pH 7.5)/6× SSC for 5 min. Repeat with a second sheet of Whatman 3MM paper.

6. Place membrane on Whatman 3MM paper saturated with 2× SSC for 5 min. Then place membrane on dry Whatman paper to air dry for 10 min.
7. Bake membrane 90 min at 80°C in vacuum oven or UV cross-link.
8. Process as for bacterial filter hybridization (UNITS 6.3 & 6.4), hybridizing the membrane with probes complementary to previously isolated cDNAs.

When selecting probes, either random-primed cDNAs or oligonucleotides complementary to the cDNA sequence may be used. If the cDNA is a member of a protein family, it may be advantageous to use oligonucleotides to avoid inadvertently excluding genes related but not identical to those initially obtained.

MICROPLATE PLASMID RESCUE

In some cases, it is desirable to isolate plasmids from a large number of positive colonies (Basic Protocol 2, steps 18 and 19). The protocol described below is a batch DNA preparation protocol developed by Steve Kron (University of Chicago, Chicago, Ill.) as a scale-up of a basic method developed by Manuel Claros (Laboratoire de Génétique Moleculaire, Paris, France).

Materials

- 2× Glu/CM –Trp liquid medium: 2× CM –Trp liquid medium (UNIT 13.1) supplemented with 4% glucose
- Master plate of Gal-dependent yeast colonies (see Basic Protocol 2, step 18)
- Rescue buffer: 50 mM Tris·Cl (pH 7.5)/10 mM EDTA/0.3% (v/v) 2-mercaptoethanol (prepare fresh)
- Lysis solution: 2 to 5 mg/ml Zymolyase 100T/rescue buffer *or* 100,000 U/ml β-glucuronidase (type HP-2 crude solution from *Helix pomatia*; Sigma) diluted 1:50 in rescue buffer
- 10% (w/v) SDS
- 7.5 M ammonium acetate (APPENDIX 2)
- Isopropanol
- 70% ethanol
- TE buffer, pH 8.0 (APPENDIX 2)
- 24-well microtiter plates
- Centrifuge with microplate holders, refrigerated
- Repeating micropipettor
- 37°C rotary shaker

Grow yeast cultures

1. Aliquot 2 ml of 2× Glu/CM –Trp medium into each well of a 24-well microtiter plate. Into each well, pick a putative positive colony. Grow overnight with shaking at 30°C.

The 2× minimal medium is used to maximize the yield of yeast. Four plates can generally be handled conveniently at once, based on the number that can be centrifuged simultaneously.

2. Centrifuge 5 min at 1500 × g, 4°C. Shake off supernatant with a snap and return the plate to upright.
3. Swirl or lightly vortex the plate to resuspend cell pellets in remaining liquid. Add 1 ml water to each well and swirl lightly.

Cell pellets can most easily be resuspended in residual liquid before adding new solutions. Addition of liquid can be accomplished using a repeating pipettor.

SUPPORT PROTOCOL 4

Analysis of Protein Interactions

20.1.31

4. Centrifuge 5 min at $1500 \times g$, 4°C . Shake off supernatant and resuspend pellet. Add 1 ml rescue buffer.
5. Centrifuge 5 min at $1500 \times g$, 4°C . Shake off supernatant and resuspend pellet in the small volume of liquid remaining in the plate.

Lyse cells

6. To each well, add 25 μl lysis solution. Swirl or vortex to mix. Incubate (with cover on) on a rotary shaker ~ 1 hr at 37°C .

Lysis solution need not be completely dissolved before use. By 1 hr, lysis should be obvious as coagulation of yeast into a white precipitate.

Susceptibility of yeast strains to lytic enzymes varies. If lysis occurs rapidly, then less lytic enzyme should be used. If the lysis step is allowed to go too far, too much of the partially dissolved cell wall may contaminate the final material. Lysis can be judged by examining cells with a phase-contrast microscope. Living cells are white with a dark halo and dead cells are uniformly gray. Lysis leads to release of granular cell contents into the medium. Once cells are mostly gray and many are disrupted, much of the plasmid should have been released.

7. To each well, add 25 μl of 10% SDS. Mix gently by swirling to completely disperse the precipitates. Allow plates to sit 1 min at room temperature.

At this point, the wells should contain a clear, somewhat viscous solution.

Purify plasmid

8. To each well, add 100 μl of 7.5 M ammonium acetate. Swirl gently, then incubate 15 min at -70°C or -20°C until frozen.

*Addition of acetate should result in the formation of a massive white precipitate of cell debris and SDS. The freezing step appears to improve removal of inhibitors of *E. coli* transformation.*

9. Remove plate from freezer. Once it begins to thaw, centrifuge 15 min at $3000 \times g$, 4°C . Transfer 100 to 150 μl of the resulting clear supernatants to clean 24-well plates.

In general, some contamination of the supernatant with pelleted material cannot be avoided. However, it is better to sacrifice yield in order to maintain purity.

10. To each well, add ~ 0.7 vol isopropanol. Mix by swirling and allow to precipitate 2 min at room temperature.

A cloudy fine precipitate should form immediately after isopropanol is added.

11. Centrifuge 15 min at $3000 \times g$, 4°C . Shake off supernatant with a snap.
12. To each well, add ~ 1 ml cold 70% ethanol. Mix by swirling, centrifuge 5 min at $3000 \times g$, 4°C . Shake off supernatant with a snap, invert plates and blot well onto paper towel. Allow plates to air dry.
13. To each well, add 100 μl TE buffer. Swirl well and allow to rest on bench several minutes, until the pellets appear fully dissolved. Transfer preps to microcentrifuge tubes or 96-well plates for storage at -20°C .

*One to five microliters of each of the resulting preparations can be used to transform competent *E. coli*: for KC8, electroporation should be used (see Basic Protocol 2, step 21). Sometimes, the yield of transformants is low if *E. coli* carrying plasmids are not permitted time to increase the plasmid copy number above a critical threshold before the cells are placed on selective medium. Allow plenty of time for cells to express antibiotic resistance or the TRP1 gene before plating.*

If insufficient numbers of colonies are obtained by this approach, the final plasmid preparation can be resuspended in 20 μl instead of 100 μl TE buffer to concentrate the DNA stock.

ADDITIONAL SPECIFICITY SCREENING

The three test plasmids outlined (pSH18-34, pRFHM1, and pEG202; see Basic Protocol 2, step 24) represent a minimal test series. If other LexA-bait proteins that are related to the bait protein used in the initial library screen are available, substantial amounts of information can be gathered by additional specificity tests. For example, if the initial bait protein was LexA fused to the leucine zipper of c-Fos, specificity screening of interactor-hunt positives against the leucine zippers of c-Jun or GCN4 in addition to that of c-Fos might allow discrimination between proteins that are specific for fos versus those that generically associate with leucine zippers.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Glycerol solution

- 65% (v/v) glycerol, sterile
- 0.1 M MgSO₄
- 25 mM Tris·Cl, pH 8.0 (APPENDIX 2)
- Store up to 1 year at room temperature

Laemmli sample buffer, 2×

- 10% (v/v) 2-mercaptoethanol (2-ME)
- 6% (w/v) SDS
- 20% (v/v) glycerol
- 0.2 mg/ml bromphenol blue
- 0.025× Laemmli stacking buffer (see recipe; optional)
- Store up to 2 months at room temperature

This reagent can conveniently be prepared 10 ml at a time.

Laemmli stacking buffer, 2.5×

- 0.3 M Tris·Cl, pH 6.8
- 0.25% (w/v) SDS
- Store up to 1 month at 4°C

Lysis solution

- 50 mM Tris·Cl, pH 7.5 (APPENDIX 2)
- 10 mM EDTA
- 0.3% (v/v) 2-mercaptoethanol (2-ME), added just before use
- 2% (v/v) β-glucuronidase from *Helix pomatia* (Type HP-2; Sigma), added just before use

COMMENTARY

Background Information

Interaction-based cloning is derived from three experimental observations. In the first, Brent and Ptashne (1985) demonstrated that it was possible to assemble a novel, functional transcriptional activator by fusing the DNA-binding domain from one protein, LexA, to the activation domain from a second protein, GAL4. This allowed the use of a single reporter system containing a single DNA-binding motif, the LexA operator, to study transcriptional ac-

tivation by any protein of interest. In the second, Ma and Ptashne (1988) built on this work to demonstrate that the activation domain could be brought to DNA by interaction with a DNA-binding domain. In the third, Fields and Song (1989), working independently of Ma and Ptashne, used two yeast proteins, SNF1 and SNF4, to make an SNF1 fusion to the DNA-binding domain of GAL4 and an SNF4 fusion to the GAL4 activation domain. They demonstrated that the strength of the SNF1-SNF4

interaction was sufficient to allow activation through a GAL4 DNA-binding site. From this, they suggested the feasibility of selecting interacting proteins by performing screens of cDNA libraries made so that library-encoded proteins carried activating domains.

Several groups have developed cDNA library strategies along these lines, with some systems using LexA and others using GAL4 as the DNA-binding domain (Table 20.1.4). LexA and GAL4 each have different properties that should be considered when selecting a system. LexA is derived from a heterologous organism, has no known effect on the growth of yeast, possesses no residual transcriptional activity, can be used in GAL4⁺ yeast, and can be used with a Gal-inducible promoter. Because GAL4 is an important yeast transcriptional activator, it has the disadvantage that experiments must be performed in gal4⁻ yeast strains to avoid background due to activation of the reporter system by endogenous GAL4. Such gal4⁻ strains are frequently less healthy and more difficult to transform than wild-type strains, and either libraries must be constitutively expressed or alternate inducible systems must be used. By contrast, the GAL4 DNA-binding domain may be more efficiently localized to the nucleus and may be preferred for some proteins (for a review of GAL4-based systems, see Bartel et al., 1993). Whichever system is used, it is important to remember that the bait protein constitutes a novel fusion protein whose properties may not exactly parallel those of the original unfused protein of interest. Although systems using the two-hybrid paradigm have been developed in mammalian cells (see Table 20.1.4), these have not been used effectively in library screens. It seems likely that the organism of choice for two-hybrid identification of novel partner proteins will remain yeast.

cDNAs that pass specificity tests are referred to as positives, or “true positives.” In interactor hunts conducted to date, anywhere from zero to practically all isolated plasmids passed the final specificity test. If no positives are obtained, the tissue source for the library originally used may not be appropriate, and a different library may produce better results. However, there are some proteins for which no positives are found. Various explanations for this are provided below. Conversely, some library-encoded proteins are known to be isolated repeatedly using a series of unrelated baits, and these proteins demonstrate at least some specificity. One of these, heat shock protein 70, might be explained by positing that it

assists the folding of some LexA-fused bait proteins, or alternatively, that these bait proteins are not normally folded. This example illustrates the point that the physiological relevance of even quite specific interactions may sometimes be obscure.

Because the screen involves plating multiple cells to Gal/CM –Ura, –His, –Trp, –Leu drop-out medium for each primary transformant obtained, multiple reisolates of true positive cDNAs are frequently obtained. If a large number of specific positives are obtained, it is generally a good idea to attempt to sort them into classes—for example, digesting minipreps of positives with *EcoRI*, *XhoI*, and *HaeIII* will generate a fingerprint of sufficient resolution to determine whether multiple reisolates of a small number of clones or single isolates of many different clones have been obtained. The former situation is a good indication that the system is working well.

An important issue that arises in an interactor hunt is the question of how biologically relevant interacting proteins that are isolated are likely to be. This leads directly to the question of what K_d of association two molecules must have to be detected by an interactor hunt. In fact, this is not at all a simple issue. For the system described here, most fusion proteins appear to be expressed at levels ranging from 50 nM to 1 μ M (Golemis and Brent, 1992). Given the strength of the *GAL* promoter, it is likely that many library-encoded proteins are expressed at similarly high levels, $\geq 1 \mu$ M in the nucleus (Golemis and Brent, 1992). At this concentration, which is in considerable excess over the nuclear concentration of operator-bound bait protein, a cDNA library-encoded protein should half-maximally occupy the DNA-bound bait protein if it possesses a K_d of 10^{-6} M, making it theoretically possible that very-low-affinity interactions could be detected. Such interactions have been observed in some cases. In contrast, some interactions that have been previously established using other methods and are predicted by known K_d to be easily detected by these means, either are not detected or are detected only weakly (Finley and Brent, 1994; Estojak et al., 1995). Because of the conservation of many proteins between lower and higher eukaryotes, one explanation for this observation is that either one or both of the partners being tested is being sequestered from the desired interaction by fortuitous association with an endogenous yeast protein. A reasonably complete investigation of the degree of correlation between in vitro determina-

tions of interaction affinity and apparent strength of interaction in the interaction trap is included in Estojak et al. (1995). The result of this investigation suggests it is important to measure the affinity of detected interactions under different conditions, using a second assay system, rather than to draw conclusions about affinity based on detection in the interaction trap.

A number of different plasmids can be used for conducting an interactor hunt. Their properties are summarized in Tables 20.1.1 and 20.1.2. Because of the generous and open scientific exchange between investigators using the system, the number of available plasmids and other components has greatly expanded since the appearance of the initial two-hybrid reagents, facilitating the study of proteins inaccessible by the original system.

The original parent plasmid for generating LexA fusions, pEG202 is a derivative of 202 + PL (Ruden et al., 1991; see Fig. 20.1.3) that contains an expanded polylinker region. The available cloning sites in pEG202 include *EcoRI*, *BamHI*, *SalI*, *NcoI*, *NotI*, and *XhoI*, with the reading frame as described in the legend to Figure 20.1.3. Since the original presentation of this system, a number of groups have developed variants of this plasmid that address specialized research needs. Those currently available, as well as purposes for which they are suited, are listed in Table 20.1.1. pGilda, created by David A. Shaywitz, places the LexA-fusion cassette under the control of the inducible *GAL1* promoter, allowing expression of the bait protein for limited times during library screening, reducing the exposure of yeast to toxic baits. pJK202, created by Joanne Kamens, adds nuclear localization sequences to pEG202, facilitating assay of the function of proteins lacking internal nuclear localization sequences. pNLexA, created by Ian York, places LexA carboxy-terminal in the fusion domain, allowing assay of interactions that require an unblocked amino-terminus on the bait protein. pEE202I, created by Mike Watson and Rich Buckholz, allows chromosomal integration of a pEG202-like bait, thus reducing expression levels so they are more physiological for bait proteins normally present at low levels intracellularly. All of these have been extensively tested by numerous researchers. pGilda, pJK202, and pEE202I work with complete reliability. pNLexA works effectively with ~50% of the fusion domains tried, but synthesizes only very low levels of protein (relative to expression of the same fusion domain as a

pEG202 fusion) with the remaining 50%. Attachment of fusion domains amino-terminal either to *LexA* or *GAL4* has been generally problematic in the hands of many investigators; it may be that appending additional protein sequences to the amino termini of these proteins is destabilizing, although the problem has not been rigorously investigated.

A series of *lacZ* reporters of differing sensitivity to transcriptional activation can be used to detect interactions of varying affinity (see Table 20.1.2). These plasmids are *LexA* operator-containing derivatives of the plasmid LR1Δ1 (West et al., 1984). In LR1Δ1, a minimal *GAL1* promoter lacking the *GAL1* upstream activating sequences (*GAL_{UAS}*) is located upstream of the bacterial *lacZ* gene. In pSH18-34, eight *LexA* operators have been cloned into an *XhoI* site located 167 bp upstream of the *lacZ* gene (S. Hanes, unpub. observ.). pJK103 and pRB1840 contain two and one operator, respectively.

pJK101 is similar to pSH18-34, except that it contains the *GAL1* upstream activating sequences (*GAL_{UAS}*) upstream of two *LexA* operator sites. A derivative of del20B (West et al., 1984), it is used in the repression assay (Brent and Ptashne, 1984; see Fig. 20.1.5) to assess LexA fusion binding to operator.

pSH17-4 is a *HIS3* 2μm plasmid that encodes LexA fused to the activation domain of the yeast activator GAL4. EGY48 cells bearing this plasmid will produce colonies in overnight growth on medium lacking Leu, and yeast that additionally contain pSH18-34 will turn deep blue on plates containing Xgal. This plasmid serves as a positive control for the activation of transcription.

pRFHM1 is a *HIS3* 2μm plasmid that encodes LexA fused to the N-terminus of the *Drosophila* protein bicoid. The plasmid has no ability to activate transcription, so EGY48 cells that contain pRFHM1 and pSH18-34 do not grow on -Leu medium and remain white on plates containing Xgal. pRFHM1 is a good control for specificity testing, because it has been demonstrated to be sticky—that is, to associate with a number of library-encoded proteins that are clearly nonphysiological interactors (R. Finley, Wayne State University, Detroit, Mich., unpub. observ.).

This protocol uses interaction libraries (Table 20.1.3) made in pJG4-5 or its derivatives (see Fig. 20.1.6). pJG4-5 was developed to facilitate isolation and characterization of novel proteins in interactor hunts (Gyuris et al., 1993). The pJG4-5 cDNA library expression

cassette is under control of the *GAL1* promoter, so library proteins are expressed in the presence of galactose (Gal) but not glucose (Glu). This conditional expression has a number of advantages, the most important of which is that many false-positives obtained in screens can be easily eliminated because they do not demonstrate a Gal-dependent phenotype. The expression cassette consists of an ATG to start translation, a nuclear localization signal to extend the interaction trap's range to include proteins that are normally predominantly localized in the cytoplasm, an activation domain (acid blob; Ma and Ptashne, 1987), the hemagglutinin epitope tag to permit rapid assessment of the size of encoded proteins, *EcoRI-XhoI* sites designed to receive directionally synthesized cDNAs, and the alcohol dehydrogenase (*ADH*) termination sequences to enhance the production of high levels of library protein. The plasmid also contains the *TRP1* auxotrophy marker and 2 μ m origin for propagation in yeast. A derivative plasmid, pJG4-5I, was created by Mike Watson and Richard Buckholz to facilitate chromosomal integration of the activation domain fusion expression plasmid.

A series of recently developed derivatives of pEG202, pJG4-5, and *lacZ* reporter plasmids (MW101 to MW112) alter the antibiotic resistance markers on these plasmids from ampicillin (Ap^r) to either kanamycin (Km^r) or chloramphenicol (Cm^r ; Watson et al., 1996). Judiciously mixing and matching these plasmids in conjunction with Ap^r libraries would considerably reduce work subsequent to library screening, because the KC8 transformation, which involves *trpC* complementation in bacteria, could be omitted.

EGY48 and EGY191 (see Table 20.1.2) are both derivatives of the strain U457 (a gift of Rodney Rothstein, Columbia University, New York, N.Y.) in which the endogenous *LEU2* gene has been replaced by homologous recombination with *LEU2* reporters carrying varying numbers of *LexA* operators, using a procedure detailed in Estojak et al. (1995).

Interaction Trap-compatible reagents have recently become commercially available; Clontech and Invitrogen were the first to market such reagents and have recently been joined by OriGene. All suppliers use systems with the most sensitive reporters (EGY48 and pSH18-34), and provide their own positive and negative controls for testing activation or interaction between defined proteins. For expression of bait and library proteins, the Clontech Matchmaker LexA two-hybrid system and the

OriGene Duplex-A system use some of the basic set of plasmids described here (see Table 20.1.1 for availability). Forward sequencing primers for bait and library plasmids are included in the Clontech kit, and Insert Screening Amplimer Sets for both plasmids can be acquired separately. Additional related products from Clontech include KC8 competent cells, anti-LexA monoclonal antibodies, a yeast transformation system, a yeast plasmid isolation kit, and an EGY48 partner strain for yeast mating to facilitate the analysis of interaction specificity. OriGene has a generally similar product line to Clontech. In contrast, Invitrogen has substantially modified the Interaction Trap core reagents to develop its own bait and library plasmids. pHybLex/Zeo, a novel bait plasmid, is ~50% smaller than the original pEG202 (making it easier to clone into), and it has an enriched polylinker. Significantly, it replaces both the *Ap^r* and *HIS3* genes with a novel gene that confers resistance to the antibiotic Zeocin (supplied with the kit), which provides selection in both bacteria and yeast. This elimination of auxotrophic selection for the bait plasmid renders the LexA-fusion construct usable with libraries and strains from all existing two-hybrid systems and additionally facilitates the direct selection of library plasmid in strains other than KC8. Some changes, which are designed to make the vector easier to use, have also been introduced in the library vector pYESTrp (e.g., it uses a V5 epitope tag for protein detection). The Invitrogen kit, termed Hybrid Hunter, includes the bait/library/reporter plasmids and EGY48 yeast strain as noted, and additionally includes primer sets for bait and library plasmids and the L40 yeast strain, should an investigator wish to use a *HIS3* auxotrophy selection. Additional related products from Invitrogen include antibodies for detection of bait and prey fusion proteins (anti-LexA and anti-V5), pJG4-5 library vector primers, and a Transformation Kit.

A significant advantage of the entry of commercial entities into the Interaction Trap field is the rapid increase in the number of compatible cDNA libraries. A list of currently available premade libraries available from these companies is presented in Table 20.1.3, and custom-made libraries are also available upon request. Because new libraries and other related reagents are being constantly added to the line of two-hybrid related products, it is advisable to contact the companies or visit their Web sites (www.clontech.com, www.invitrogen.com, and www.origene.com) for the latest information.

Finally, over the last several years, a number of groups have adapted basic two-hybrid strategies to more specialized applications, and they have devised strategies to broaden their basic functionality. Interaction Mating (Finley and Brent, 1994) has been used to establish extended networks of targeted protein-protein interaction. In this approach, a panel of LexA-fused proteins are transformed into a *MAT α* haploid selective strain (such as RFY206), a panel of activation-domain fused proteins are transformed into a suitable *MAT α* haploid (such as EG448), and the two panels are cross-gridded against each other for mating. Selected diploids are then screened by replica plating to selective medium. This approach complements library screening in large-scale applications, such as proposed definition of interaction maps for entire genomes (Bartel et al., 1996).

Interaction mating has also provided the basis for an alternative two-hybrid hunt protocol (see Alternate Protocol 2), useful in cases when a single library will be screened with different baits. In this approach (Bendixen et al., 1994; Finley and Brent, 1994; Kolonin and Finley, 1998), a library is introduced into a single strain, like EGY48, and aliquots are stored frozen. To conduct a hunt, an aliquot is thawed and mated with a strain expressing a bait. This allows one to avoid repeated high-efficiency transformations, since a single library transformation can provide enough pretransformed yeast to conduct dozens of interactor hunts. Moreover, some yeast strains pretransformed with libraries are becoming commercially available, which may eliminate altogether the need to conduct a high-efficiency library transformation for some researchers.

Two-hybrid approaches have been shown to be effective in identifying small peptides with biological activities on selected baits (Yang et al., 1995; Colas et al., 1996), which may prove to be useful as a guide to targeted drug design. Rapid screening protocols have been devised using custom-synthesized libraries expressing sheared plasmid DNA to facilitate rapid mapping of interaction interfaces (Stagljar et al., 1996). Osborne and coworkers have demonstrated the effectiveness of a tribrid (or tri-hybrid) approach, in which an additional plasmid expresses a tyrosine kinase to specifically modify a bait protein, allowing detection of SH2-domain-containing partner proteins that recognize specific phosphotyrosine residues (Osborne et al., 1995). A variety of more elaborate tribrid approaches, in which a DNA-binding domain fused protein is used to present an

intermediate nonprotein compound for interaction with a library, have been developed and proven effective. These approaches have allowed the identification of proteins binding specific drug ligands (Chiu et al., 1994; Licitra and Liu, 1996), as well as the identification of proteins binding to RNA sequences (SenGupta et al., 1996; Wang et al., 1996). It is expected that the range of utility of these systems will continue to expand.

Critical Parameters and Troubleshooting

To maximize chances of a successful interactor hunt, a number of parameters should be taken into account. Before attempting a screen, bait proteins should be carefully tested to ensure that they have little or no intrinsic ability to activate transcription. Bait proteins must be expressed at reasonably high levels and must be able to enter the yeast nucleus and bind DNA (as confirmed by the repression assay). Optimally, integrity and levels of bait proteins should be confirmed by immunoblot analysis, using an antibody to either LexA or the fused domain. In particular, at this time, bait proteins that have extensive transmembrane domains or are normally excluded from the nucleus are not likely to be productively used in a library screen. Proteins that are moderate to strong activators will need to be truncated to remove activating domains before they can be used.

If a protein neither activates nor represses, the most likely reason is that it is not being made. This can be determined by immunoblot analysis of a crude lysate protein extract of EGY48 (*UNIT 10.8*; Samson et al., 1989) containing the plasmid, using anti-LexA antibodies as primary antiserum. If the full protein is not made, it may be possible to express truncated derivatives of the protein. If the protein is made, but still does not repress, it may not enter the yeast nucleus effectively, although this appears to be a relatively rare problem. In this case, introducing the coding sequence for the fused moiety into a LexA fusion vector containing a nuclear localization motif (e.g., pJK202; J. Kamens, BASF, Worcester, Mass., unpub. observ.) may solve the problem.

The test for the leucine (Leu) requirement is extremely important to determine whether the bait protein is likely to yield an unworkably high background. The *LEU2* reporter in EGY48 is more sensitive than the pSH18-34 reporter for some baits (Estojak et al., 1995). Therefore, it is possible that a bait protein demonstrating little or no signal in a β -galac-

tosidase assay may nevertheless permit some growth on –Leu medium. If this occurs, there are several options. First, a less sensitive strain can be used, as described in the text. Second, background can sometimes be reduced further by making the EGY strain diploid (e.g., D. Krainc, Harvard Medical School, Boston, Mass.; R. Finley and R. Brent, unpub. observ.) or by performing the hunt by interaction mating as described in Alternate Protocol 2. A third option is to attempt to truncate the bait protein to remove activating function. In general, it is useful to extrapolate from the number of cells that grow on –Leu medium to the number that would be obtained in an actual library screen, and determine if this is a background level that can be tolerated. For example, if two colonies arise from 100,000 plated cells on –Leu medium, 200 to 400 would be expected in an actual screen of 10^6 cDNAs. Although this is a high initial number of positives, the vast majority should be eliminated immediately through easily performed controls. This is a judgment call. Finally, very rarely it happens that a bait that appears to be well behaved and negative for transcriptional activation through all characterization steps will suddenly develop a very high background of transcriptional activation following library transformation. The reason for this is currently obscure, and no means of addressing this problem has as yet been found: such baits are hence inappropriate for use in screens.

The protocols described in this unit use initial screening with the most sensitive reporters followed by substitution with less sensitive reporters if activation is detected. An obvious question is, why not start out working with extremely stringent reporters and know immediately whether the system is workable? In fact, some researchers routinely use a combination of pJK103 or pRB1840 with EGY191, and obtain proteins that to date appear to be biologically relevant partners from library screens. However, extensive comparison studies using interactors of defined *in vitro* affinity with different combinations of *LacZ* and *LEU2* reporters (Estojak et al., 1995) have indicated that although the most sensitive reporters (pSH18-34) may in some cases be prone to background problems, the most stringent reporters (EGY191, pRB1840) may miss some interactions that certainly are biologically relevant and occur inside cells. In the end, the choice of reporters devolves to the preference of individual investigators: the bias of the authors is to cast a broad net in the early stages of a screen,

and hence to use more sensitive reporters when practicable.

It is important to move expeditiously through characterization steps and to handle yeast transformed with bait plasmids with care. In cases where yeasts have been maintained on plates for extended periods (e.g., 4 days at room temperature or >2 to 3 weeks at 4°C), unexpected problems may crop up in subsequent library screens.

The transformation protocol is a version of the lithium acetate transformation protocol described by Schiestl and Gietz (1989) and Gietz et al. (1992; see UNIT 13.7) that maximizes transformation efficiency in *Saccharomyces cerevisiae* and produces up to 10^5 colonies/ μ g plasmid DNA. In contrast to *Escherichia coli*, the maximum efficiency of transformation for *S. cerevisiae* is $\sim 10^4$ to 10^5 / μ g input DNA. It is extremely important to optimize transformation conditions before attempting an interactor hunt. Perform small-scale pilot transformations to ensure this efficiency is attained and to avoid having to use prohibitive quantities of library DNA. In addition, as for any effort of this type, it is a good idea to obtain or construct a library from a tissue source in which the bait protein is known to be biologically relevant.

In practice, the majority of proteins isolated by interaction with a LexA fusion turn out to be specific for the fused domain; a smaller number are nonspecifically sticky, and to date there appears to have been only one isolation from a eukaryotic library of a protein specific for LexA. However, it is generally informative to retest positive clones on more than one LexA bait protein; ideally, library-derived clones should be tested against the LexA fusion used for their isolation, several LexA fusions to proteins that are clearly unrelated to the original fusion, and if possible, several LexA fusions that there is reason to believe are related to the initial protein (e.g., if the initial probe was LexA-Fos, a good related set would include LexA-Jun and LexA-GCN4).

Colony selection for master plate production is one of the more variable parts of the procedure. For strong interactors, colonies will grow up in 2 days. However, if plates are left at 30°C, new colonies will continue to appear every day. Those that appear rapidly are most likely to reflect interactors that are biologically relevant to the bait protein. Those that appear later may or may not be relevant. However, many parameters can delay the time of colony formation of cells that contain valid interactions, including the strength of the interaction

and the level of expression of the library-encoded protein.

Anticipated Results

Depending on the protein used as bait, anywhere from zero to hundreds of specific interactors will be obtained from 10⁶ primary transformants.

Time Considerations

If all goes well, once the required constructions have been made it will take ~1 week to perform yeast transformations, obtain colonies, and determine whether bait proteins are appropriate. It will take a second week to perform library transformations, replat to selective medium, and obtain putative positives. A third week will be required to rescue the plasmid from the yeast, passage it through *E. coli*, transform fresh yeast, and confirm specificity.

Literature Cited

- Bartel, P.L., Chien, C.-T., Sternglanz, R., and Fields, S. 1993. Using the two-hybrid system to detect protein-protein interactions. *In Cellular Interactions in Development: A Practical Approach* (D.A. Hartley, ed.) pp. 153-179. Oxford University Press, Oxford.
- Bartel, P.L., Roecklein, J.A., SenGupta, D., and Fields, S. 1996. A protein linkage map of *Escherichia coli* bacteriophage T7. *Nature Genet.* 12:72-77.
- Bendixen, C., Gangloff, S., and Rothstein, R. 1994. A yeast mating-selection scheme for detection of protein-protein interactions. *Nucl. Acids Res.* 22:1778-1779.
- Breeden, L. and Nasmyth, K. 1985. Regulation of the yeast HO gene. *Cold Spring Harbor Symp. Quant. Biol.* 50:643-650.
- Brent, R. and Ptashne, M. 1984. A bacterial repressor protein or a yeast transcriptional terminator can block upstream activation of a yeast gene. *Nature* 312:612-615.
- Brent, R. and Ptashne, M. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* 43:729-736.
- Chien, C.-T., Bartel, P.L., Sternglanz, R., and Fields, S. 1991. The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci. U.S.A.* 88:9578-9582.
- Chiu, M.I., Katz, H., and Berlin, V. 1994. RAPT1, a mammalian homolog of yeast Tor, interacts with the FKBP12/rapamycin complex. *Proc. Nat. Acad. Sci. U.S.A.* 91:12574-12578.
- Colas, P., Cohen, B., Jessen, T., Grishina, I., McCoy, J., and Brent, R. 1996. Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2. *Nature* 380:548-550.
- Durfee, T., Becherer, K., Chen, P.L., Yeh, S.H., Yang, Y., Kilburn, A.E., Lee, W.H., and Elledge, S.J. 1993. The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes & Dev.* 7:555-569.
- Estojak, J., Brent, R., and Golemis, E.A. 1995. Correlation of two-hybrid affinity data with in vitro measurements. *Mol. Cell. Biol.* 15:5820-5829.
- Fearon, E.R., Finkel, T., Gillison, M.L., Kennedy, S.P., Casella, J.F., Tomaselli, G.F., Morrow, J.S., and Dang, C.V. 1992. Karyoplasmic interaction selection strategy: A general strategy to detect protein-protein interactions in mammalian cells. *Proc. Nat. Acad. Sci. U.S.A.* 89:7958-7962.
- Fields, S. and Song, O. 1989. A novel genetic system to detect protein-protein interaction. *Nature* 340:245-246.
- Finley, R.L., Jr., and Brent, R. 1994. Interaction mating reveals binary and ternary connections between *Drosophila* cell cycle regulators. *Proc. Natl. Acad. Sci. U.S.A.* 91:12980-12984.
- Gietz, D., St. Jean, A., Woods, R.A., and Schiestl, R.H. 1992. Improved method for high-efficiency transformation of intact yeast cells. *Nucl. Acids Res.* 20:1425.
- Golemis, E.A. and Brent, R. 1992. Fused protein domains inhibit DNA binding by LexA. *Mol. Cell Biol.* 12:3006-3014.
- Grunstein, M., and Hogness, D.S. 1975. Colony hybridization: A method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci. U.S.A.* 72:3961-3965.
- Gyuris, J., Golemis, E.A., Chertkov, H., and Brent, R. 1993. Cdi1, a human G1- and S-phase protein phosphatase that associates with Cdk2. *Cell* 75:791-803.
- Kaiser, C., Michaelis, S., and Mitchell, A. 1994. *Methods in Yeast Genetics*, a Cold Spring Harbor Laboratory Course Manual, pp.135-136. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Kolonin, M.G. and Finley, R.L., Jr. 1998. Targeting cyclin-dependent kinases in *Drosophila* with peptide aptamers. *Proc. Natl. Acad. Sci. U.S.A.* *In press.*
- Licitra, E.J. and Liu, J.O. 1996. A three-hybrid system for detecting small ligand-protein receptor interactions. *Proc. Nat. Acad. Sci. U.S.A.* 93:12817-12821.
- Ma, J. and Ptashne, M. 1987. A new class of yeast transcriptional activators. *Cell* 51:113-119.
- Ma, J. and Ptashne, M. 1988. Converting an eukaryotic transcriptional inhibitor into an activator. *Cell* 55:443-446.
- Osborne, M., Dalton, S., and Kochan, J.P. 1995. The yeast tribrid system: Genetic detection of transphosphorylated ITAM-SH2 interactions. *Bio/Technology* 13:1474-1478.
- Ruden, D.M., Ma, J., Li, Y., Wood, K., and Ptashne, M. 1991. Generating yeast transcriptional activators containing no yeast protein sequences. *Nature* 350:426-430.
- Samson, M.-L., Jackson-Grusby, L., and Brent, R. 1989. Gene activation and DNA binding by

Drosophila Ubx and abd-A proteins. *Cell* 57:1045-1052.

Schiestl, R.H. and Gietz, R.D. 1989. High-efficiency transformation of intact yeast cells using single-stranded nucleic acids as a carrier. *Curr. Genet.* 16:339-346.

SenGupta, D.J., Zhang, B., Kraemer, B., Pochart, P., Fields, S., and Wickens, M. 1996. A three-hybrid system to detect RNA-protein interactions in vivo. *Proc. Nat. Acad. Sci. U.S.A.* 93:8496-8501.

Sikorski, R.S. and Hieter, P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* 122:19-27.

Stagljar, I., Bourquin, J.-P., and Schaffner, W. 1996. Use of the two-hybrid system and random sonicated DNA to identify the interaction domain of a protein. *BioTechniques* 21:430-432.

Vasavada, H.A., Ganguly, S., Germino, F.J., Wang, Z.X., and Weissman, S.M. 1991. A contingent replication assay for the detection of protein-protein interactions in animal cells. *Proc. Nat. Acad. Sci. U.S.A.* 88:10686-10690.

Vojtek, A.B., Hollenberg, S.M., and Cooper, J.A. 1993. Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell* 74:205-214.

Wang, Z.F., Whitfield, M.L., Ingledue, T.C.3, Dominski, A., and Marzluff, W.F. 1996. The protein that binds the 3' end of histone mRNA: A novel RNA-binding protein required for histone pre-mRNA processing. *Genes & Dev.* 10:3028-3040.

Watson, M.A., Buckholz, R., and Weiner, M.P. 1996. Vectors encoding alternative antibiotic resistance for use in the yeast two-hybrid system. *BioTechniques* 21:255-259.

West, R.W.J., Yocum, R.R., and Ptashne, M. 1984. *Saccharomyces cerevisiae* GAL1-GAL10 divergent promoter region: Location and function of the upstream activator sequence UASG. *Mol. Cell Biol.* 4:2467-2478.

Yang, M., Wu, Z., and Fields, S. 1995. Protein-peptide interactions analyzed with the yeast two-hybrid system. *Nucl. Acids Res.* 23:1152-1156.

Key Reference

Gyuris et al., 1993. See above.

Initial description of interaction trap system.

Internet Resources

<http://www.clontech.com>

<http://cmmg.biosci.wayne.edu/rfinley/lab.html>

Source of two-hybrid information, protocols, and links.

<http://www.invitrogen.com>

<http://www.origene.com>

Commercial sources for basic plasmids, strains, and libraries for interaction trap experiments.

brent@molsci.org

EA_Golemis@fcc.edu

Sources of interaction trap plasmids for specialized interactions.

<http://www.fccc.edu:80/research/labs/golemis/InteractionTrapInWork/html>

Database for false positive proteins detected in interaction trap experiments; analysis of two-hybrid usage.

<http://xanadu.mgh.harvard.edu/brentlab/homepage4.html>

Database of interaction trap protocols and related issues.

Contributed by Erica A. Golemis and Ilya Serebriiskii
Fox Chase Cancer Center
Philadelphia, Pennsylvania

Russell L. Finley, Jr. and Mikhail G. Kolonin
(hunt by interaction mating)
Wayne State University School of Medicine
Detroit, Michigan

Jeno Gyuris
Mitotix, Inc.
Cambridge, Massachusetts

Roger Brent
The Molecular Sciences Institute
Berkeley, California

Affinity Purification of Proteins Binding to GST Fusion Proteins

This unit describes the use of proteins fused to glutathione-*S*-transferase (GST fusion proteins) to affinity purify other proteins, a technique also known as GST pulldown purification. The GST fusion protein is first purified on glutathione-agarose beads as described in UNIT 16.7. The bead-bound fusion protein is then used as “bait” to test for binding to a known or suspected “test” protein which may be either purified or labeled by *in vitro* translation (UNIT 10.17). Beads with bound protein are washed, and the amount of test protein retained is determined by elution with glutathione or salt or simply by applying the beads and bound proteins to an SDS-polyacrylamide gel (UNIT 10.2; Fig. 20.2.1). The bead-bound fusion protein can also be incubated with a complex mixture of proteins. If performed on a large scale, this technique can be used to purify known or unknown proteins that interact directly or indirectly with the bait protein (Fig. 20.2.2). When a method (e.g., antibody binding) exists for detecting a particular protein, it is possible to determine if a test protein is retained from a complex mixture of proteins such as a crude cellular lysate. This GST pulldown technique can complement other methods for assessing protein-protein interactions—in *in vitro* assays such as electrophoretic mobility shift assays (EMSA; see UNIT 12.2) or coimmunoprecipitation (see UNIT 10.16), or *in vivo* assays such as the two-hybrid interaction trap (UNIT 20.1) or ubiquitin-based split-protein sensor (Johnsson and Varshavsky, 1994). The GST pulldown affinity technique is simple to perform and quite powerful. However, it is prone to artifacts due to mass-action effects, and some care should be used in interpreting positive results.

The Basic Protocol describes a strategy that can be used to isolate proteins that display affinity for (bind to) proteins fused to GST. GST fusion protein bound to agarose affinity beads is used to assay the binding of a specific test protein that has been labeled with [³⁵S]methionine by *in vitro* translation (UNIT 10.17). However, this method can be adapted for use with other types of fusion proteins—for example, His₆ (UNIT 10.11B), biotin tags, or maltose-binding protein fusions (MBP; UNIT 16.6), and these may offer particular advantages. A Support Protocol describes preparation of an *Escherichia coli* extract that is added to the reaction mixture with purified test protein to reduce nonspecific binding.

STRATEGIC PLANNING

There are three choices to be made in planning a GST fusion protein affinity purification: the selection and design of the affinity fusion to the bait protein, the assay used to detect the test protein(s), and the controls used to check for specificity.

A GST fusion protein serves well for most applications; the features of this fusion and the methods involved are described in UNIT 16.7. If either glutathione-*S*-transferase or the agarose matrix retains an interfering contaminating protein, changing either the affinity moiety (e.g., nickel, maltose, or streptavidin) or the matrix (e.g., Streptavidin Dynabeads, Dynal) may avoid the problem. If the purpose of the experiment is to determine if the bait and test proteins multimerize, it may be better to use monomeric MBP or His₆ to construct the fusion protein because GST is known to form dimers. Each system has its own advantages and disadvantages. One advantage of GST or MBP fusions is that they can increase the solubility of certain insoluble or semi-soluble proteins. A second advantage is that peptides are more readily fused to a protein (GST, MBP) than to His₆. A third advantage is the low background compared to His₆/Ni agarose. One disadvantage, however, is that the capacity of glutathione agarose beads for GST fusion proteins falls steeply with increasing size of the fusion.

Contributed by Jonathan C. Swaffield and Stephen Albert Johnston

Current Protocols in Molecular Biology (1996) 20.2.1-20.2.10

Copyright © 2000 by John Wiley & Sons, Inc.

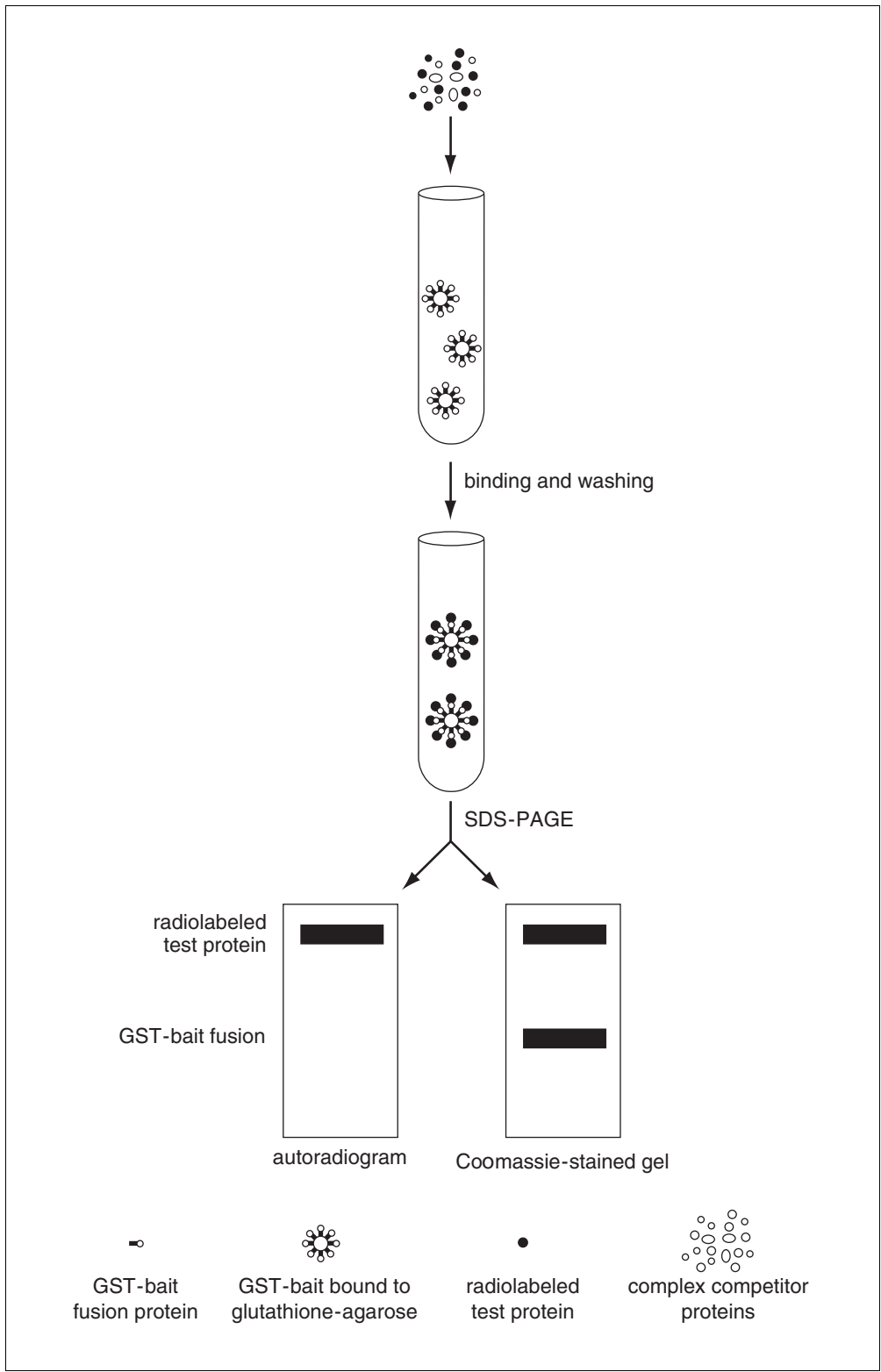


Figure 20.2.1 Schematic representation of GST pull-down purification of proteins.

The two methods commonly used to detect binding of the test protein are immunoblotting with a protein-specific antibody (UNIT 10.8) or measuring bound radioactivity of an in vitro–translated test protein (UNIT 10.17) that has been labeled with [³⁵S]methionine. Radioactive detection is preferable because it is possible to accurately quantitate the protein by phosphorimaging or scanning densitometry, and confers greater sensitivity of detection. Furthermore, after labeled proteins are separated on SDS-PAGE analytical gels, the gels can be fixed, stained for protein, and scanned as a control for loss of beads and bait protein during washing steps. With immunoblotting, accurate quantitation of the signal on a sheet of film (for chemiluminescent detection) or membrane (for the insoluble dye method) is difficult. In addition, the relatively large amount of GST fusion protein bait may exhibit nonspecific cross-reactivity with the antibodies employed, making detection difficult if the GST-bait and test proteins are similar in size. This problem can be circumvented by constructing a GST fusion protein having a protease site at the junction between GST and the bait protein sequence (see UNIT 16.4B). After binding and washing, the bait protein and any bound test protein can be released from the beads by treatment with the appropriate protease. Should nonspecific binding of the test protein to the agarose beads be a problem, protease or glutathione can be employed to elute only that portion of the test protein that is specifically bound to the bait protein or to glutathione, as opposed to the matrix.

A third method to detect binding of the target protein to the bait protein is an activity assay of the washed beads after the binding reaction. When employing an activity assay, however, it may be advantageous to elute the bait/target protein complex first. This will allow the protein to be assayed in a soluble form in the buffer of choice.

Glutathione-agarose beads with bound GST alone are often used as a negative control for nonspecific retention. Two better controls, however, are competition with a large excess of cold target protein or use of a variant of the bait protein which is known not to bind the test protein in vivo (Fig. 20.2.3). When using His₆ or biotin tags on the bait protein, a negative control with an unrelated protein may be used, but, again, it is best to use a nonbinding variant of the bait protein. It should be noted, however, that a number of naturally biotinylated proteins occur in cells—one in *E. coli* (Agarana et al., 1986), four or five in *Saccharomyces cerevisiae*, depending on growth conditions (Lim et al., 1987), and four in mammalian cells (Chandler and Ballard, 1988) and in plants (Nikolau et al., 1985). These may cause high background by binding to streptavidin beads used to immobilize biotinylated bait protein and give nonspecific bands.

GST FUSION PROTEIN–AFFINITY PURIFICATION

Bait protein fused to GST and bound to glutathione-agarose beads is incubated with radiolabeled test protein in the presence of an *Escherichia coli* protein extract. The beads and associated protein complexes are removed from the incubation mixture, washed, eluted in loading buffer, and analyzed on an SDS-PAGE gel. The resulting gel is quantitated using scanning densitometry or phosphorimaging.

Materials

- E. coli* extract in bead binding buffer (see Support Protocol)
- Test protein labeled with [³⁵S]methionine (UNIT 10.17)
- 100 to 250 μg/ml GST and GST fusion protein bound to agarose beads (UNIT 16.7), freshly prepared
- 1× SDS sample buffer (UNIT 10.2)
- Gel fixative: 10% (v/v) glacial acetic acid/45% (v/v) methanol
- Centrifuge and Beckman TLA-100 rotor or equivalent

BASIC PROTOCOL

Analysis of Protein Interactions

20.2.3

Microcentrifuge, 4°C
X-ray film *or* Storage Phosphor screen (Molecular Dynamics)
Scanning densitometer *or* phosphorimager

Additional reagents and equipment for SDS-PAGE (UNIT 10.2), Coomassie blue staining (UNIT 10.6; optional), and autoradiography (APPENDIX 3)

Prepare the extract

1. Thaw soluble *E. coli* protein extract on ice.
2. For each reaction, centrifuge 500 µl thawed extract 30 min at $175,000 \times g$ (70,000 rpm in a Beckman TLA-100 rotor), 4°C

Centrifugation removes any insoluble protein aggregates that form during storage or thawing.

3. Remove the supernatant (≥ 400 µl). Transfer 200 µl supernatant into each of two tubes, and store on ice.

Prepare test protein

4. Add 1 to 5 µl [³⁵S]methionine-labeled test protein to one tube containing 200 µl of the supernatant from step 3. Incubate 15 min on ice.

The specific activity of the [³⁵S]methionine-labeled test protein should be high enough to permit detection within a reasonable amount of time.

5. Microcentrifuge 15 min at maximum speed, 4°C.

This step removes insoluble protein aggregates from the in vitro translation mix.

6. Transfer 200 µl of the supernatant (containing labeled test protein + extract) to a clean microcentrifuge tube.

Prepare the bait protein

7. Add 20 µl GST or GST fusion protein (bait) bound to agarose beads to the other tube containing 200 µl of supernatant from the *E. coli* extract (step 3) and mix thoroughly.

From 2 to 5 µg GST bait protein is required for each reaction. It is best to include ~20 µl beads in each reaction so the bead pellet will be visible during the washes. However, if 20 µl of beads would add more than 2 to 5 µg protein, add unbound glutathione beads to provide the necessary volume. It may be necessary to minimize the amount of bait protein to decrease nonspecific interactions.

Glutathione beads bound with GST alone may be stored a number of months in PBS at 4°C, but fusion proteins are likely to be less stable and preparations should be checked for degradation before use. Excessive degradation will alter the concentration of the bait protein (see Critical Parameters), and released protein fragments may bind test protein and may not be retained on the glutathione beads during the washes, thus interfering with subsequent analysis. Although such degradation usually results in complete loss of the bait component of the fusion protein, partial degradation may leave N-terminal fragments that may or may not bind test protein. SDS-PAGE can be used to check for degradation. Proteolysis may be limited by including protease inhibitors (UNIT 10.11B) in all buffers. Unless a specific GST fusion protein is known to be stable, the binding assay should use freshly prepared beads bound with purified full-length fusion protein. Stored beads should be washed to remove any degradation products.

Perform binding assay

8. Add 200 µl well-mixed bead + extract suspension (step 7) to the tube containing test protein + extract (step 6). Incubate 1 to 2 hr at 4°C with gentle mixing.

9. Microcentrifuge the reaction 1 min at maximum speed, 4°C, to pellet the beads.
10. Remove supernatant and wash the beads three times with 1 ml bead binding buffer each time. Microcentrifuge 1 min at maximum speed, 4°C, between washes.
11. After the final wash, carefully remove all liquid.

It is important to remove all liquid to ensure that equal volumes are loaded on the analytical SDS-PAGE gel. A twisted Kimwipe or a very thin gel-loading pipet tip is useful for removing the final few microliters.

12. Add 25 µl of 1× SDS sample buffer directly to each tube and boil 5 min.

A reduction in background and an increase in the specificity of the reaction can be obtained by eluting only the GST-bait protein and any complexed target from the beads with either glutathione or protease. The eluted material is then mixed with SDS sample buffer and analyzed. Elution may be particularly advantageous when monitoring binding of the target protein by an activity assay. Bait/target complexes can be eluted into the buffer of choice and assayed away from potential interfering activities nonspecifically bound to the beads. One disadvantage of elution, however, is that elution is rarely 100% efficient, which makes quantitation of binding difficult.

Run and analyze an analytical gel

13. Run an analytical SDS-PAGE gel (see UNIT 10.2).

The amount of sample required varies with different bait/target protein pairs and the manner of detection. Loading all the sample avoids the question of loading volumes because there is no uncertainty about the ratio of loaded to unloaded sample.

14. *Optional:* Stain the gel for protein with Coomassie blue (UNIT 10.6).

The stained gel provides a visual confirmation that beads and the bait protein complexes were not lost during the washes and that all reactions contain equivalent amounts of bait protein. To verify that no material was lost during the washes, a parallel gel loaded with one-tenth the sample can be run and stained for protein to more accurately compare the GST fusion protein bands. Also, degradation of the fusion protein may lead to a decrease in the amount of bait protein, especially when using crude extracts (see Fig. 20.2.1) instead of purified labeled test protein. Such degradation must be taken into account in quantifying bound test protein.

15. Fix the gel in gel fixative 30 min at room temperature with gentle shaking.
16. Dry the gel and autoradiograph it or place it on a Storage Phosphor screen.
17. Quantitate the film using a scanning densitometer or phosphorimager.

PREPARATION OF *E. COLI* EXTRACT

It is necessary to perform the binding reactions in the presence of a complex competitor to reduce nonspecific background binding. This is best done by preparing a soluble protein extract from *Escherichia coli* in the buffer that will be used in the GST fusion protein binding reaction (see Basic Protocol).

Materials

E. coli overnight culture grown in LB medium (UNIT 1.1)
Bead binding buffer (see recipe) with and without Triton X-100 and glycerol, 4°C
Tween 20
Glycerol
Sonicator with microprobe
Centrifuge and Sorvall SS-34 rotor or equivalent

SUPPORT PROTOCOL

Analysis of Protein Interactions

20.2.5

Additional reagents and equipment for growing *E. coli* (UNIT 1.2) and quantitating proteins in solution (UNIT 10.1)

1. Harvest an overnight culture of *E. coli* grown in LB medium. Wash the cells in cold bead binding buffer.

If proteolysis of either the target or the GST-bait fusion protein proves to be a problem during the incubation period, a protease-deficient strain of E. coli may be used to prepare the complex competitor extract. However, it should be noted that such strains are often fragile and may not grow well when they contain the pLysS plasmid described in the annotation to step 3. Additional protease inhibitors also may be added to the bead binding buffer. If proteolysis continues to be a problem, alternative blocking/competing reagents—such as BLOTTO or UNI-BLOCK (Analytical Genetic Testing Center)—may be tried.

The size of the culture is dictated by the amount of extract required for the binding assays. A 1-liter culture will yield at least 10 ml of a 10 mg/ml extract.

2. Pellet the cells. Resuspend the pellet in 10 ml per liter of culture of cold bead binding buffer *without* Triton X-100 and glycerol.

Glycerol and Triton X-100 tend to cause foaming during sonication, thus reducing its effectiveness.

3. Sonicate five times, 1 min each, with a 2-min pause on ice between each cycle.

If a sonicator is unavailable, use E. coli strain BL21 (DE3) containing the pLysS plasmid (Novagen), harvest as in steps 1 and 2 but include 1% (v/v) Triton X-100, and subject the resuspended pellet to a single freeze-thaw cycle, which is sufficient to lyse the cells of this strain. DNaseI treatment in the presence of 10 mM Mg²⁺ can be used to reduce the viscosity caused by released DNA.

4. Centrifuge the sonicate 30 min at $47,800 \times g_{\text{avg}}$ (20,000 rpm in a Sorvall SS-34 rotor), 4°C.

5. Remove the supernatant and store on ice.

6. Add Triton X-100 to 1% (v/v) and glycerol to 10% (v/v) and mix thoroughly.

7. Assay the protein concentration (UNIT 10.1). Dilute the extract with cold bead binding buffer to give a final concentration of 10 mg/ml protein and use.

Undiluted extracts may also be stored in 5-ml (10 reactions) or 10-ml (20 reactions) aliquots at -80° or -20°C at least three months. After the aliquot is thawed, repeat step 4 before assaying the protein concentration and diluting to 10 mg/ml.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Bead binding buffer

50 mM potassium phosphate, pH 7.5

150 mM KCl

1 mM MgCl₂

10% (v/v) glycerol

1% (v/v) Triton X-100

Protease inhibitors (UNIT 10.11B or Protease Cocktail, Boehringer Mannheim)

Store buffer at room temperature and add protease inhibitors just before use

Also, prepare the buffer without Triton X-100 and glycerol to be used for resuspending extracts for sonication.

COMMENTARY

Background Information

The glutathione-*S*-transferase (GST) fusion–affinity purification is a simple method to present a native bait protein (as a GST fusion protein) to a test protein or complex of proteins. The assay does not involve denaturing either bait or test proteins, nor does it rely on the more restrictive assay of coimmunoprecipitation with an antibody. It is faster than using the *in vivo* two-hybrid interaction trap system (*UNIT 20.1*) to detect protein–protein interaction and can be performed in a defined medium. It is also quite sensitive—capable of detecting interactions with K_d s in the micromolar range.

The GST pulldown assay has usually been used for qualitative assessment of whether two proteins interact, but demonstration of binding *in vitro* does not necessarily indicate that the binding has *in vivo* biological relevance. Once a binding reaction is identified, variations on the basic assay can be used to further characterize the reaction. The actual bait sequence sufficient for interaction with the test protein can be identified by constructing GST fusions to different regions or fragments of the bait protein and testing them in the binding assay. If it is possible to quantitate the amount of bait protein, the binding constant can then be determined. The half-life of the binding reaction can be determined by binding the labeled test protein in an excess of bait protein: after equilibrium is reached, a large excess of cold test protein is added, and the amount of labeled test protein retained on the beads is assayed at various time points. Because it is not generally feasible to produce large amounts of labeled protein by *in vitro* translation, the test protein can alternatively be labeled with ^{32}P *in vitro* by incorporating a five-amino-acid recognition element for the catalytic domain of heart muscle kinase into the amino-terminal sequence of the protein (Blanan and Rutter, 1992). Competition assays can indicate whether two proteins use exclusive binding sites. With a limiting amount of bait protein, equimolar amounts of two test proteins are added, and the relative amounts of these proteins bound to the bait protein at equilibrium are assayed. The system can also be used to construct a binding pathway for a complex. For example, it may be determined that test protein A may only bind the bait in the presence of test protein B, whereas protein B binds the bait independently of A. This approach has been elegantly used to reconstruct the TFIID complex (Chen et al., 1994).

Many of the potential applications of the pulldown technique rely on accurate quantitation of the reactants and products. However, the sensitivity of this assay rests in the mass-action effects of using high concentrations of bait which allows detection of even weak interactions but also creates the possibility of artifacts. Paying attention to the quantitative features of the pulldown assays leads to an appreciation of both the limitations and powers of this technique.

Critical Parameters

Because of the potential for artifacts, merely showing that a test protein binds the bait protein but not GST is not sufficient to establish that this interaction takes place *in vivo*. Support for a positive binding result can come either from independent assays, such as the two-hybrid interaction trap or immunoprecipitation, or from other control affinity experiments. The most convincing demonstrations are those in which a control variant bait protein that has lost function *in vivo* also does not bind *in vitro* (see Fig. 20.2.3) and those in which the bait–test protein complex reconstitutes an *in vivo* activity.

Quantitation of the amounts of bait and test proteins relative to the question being posed is critical. If the bait is present at a high concentration, nonspecific interactions can be enhanced. If the test protein is in excess, no estimate of relative binding to the bait protein can be made. If substantially more test protein binds to the beads than to bait protein, this probably indicates a nonspecific interaction with the beads; however, test proteins may, in some cases, exhibit a genuine interaction in molar excess over bait (Fig. 20.2.2).

For quantitative purification of a test protein, however, the bait should be in a large excess and the amount of bait bound to the beads should be directly determined. Some GST fusion proteins are prone to degradation, so the beads will contain partial fragments in addition to full-length fusions. Usually, proteolysis leaves an intact GST with little or no bait protein attached, and the amount of GST versus GST-fusion protein can be directly estimated by Coomassie stain on an SDS gel. However, some fusion proteins are prone to small C-terminal deletions, and if estimation of the amount of full-length protein is critical, it should be done carefully. Degradation can result in the accumulation of N-terminal bait fragments on

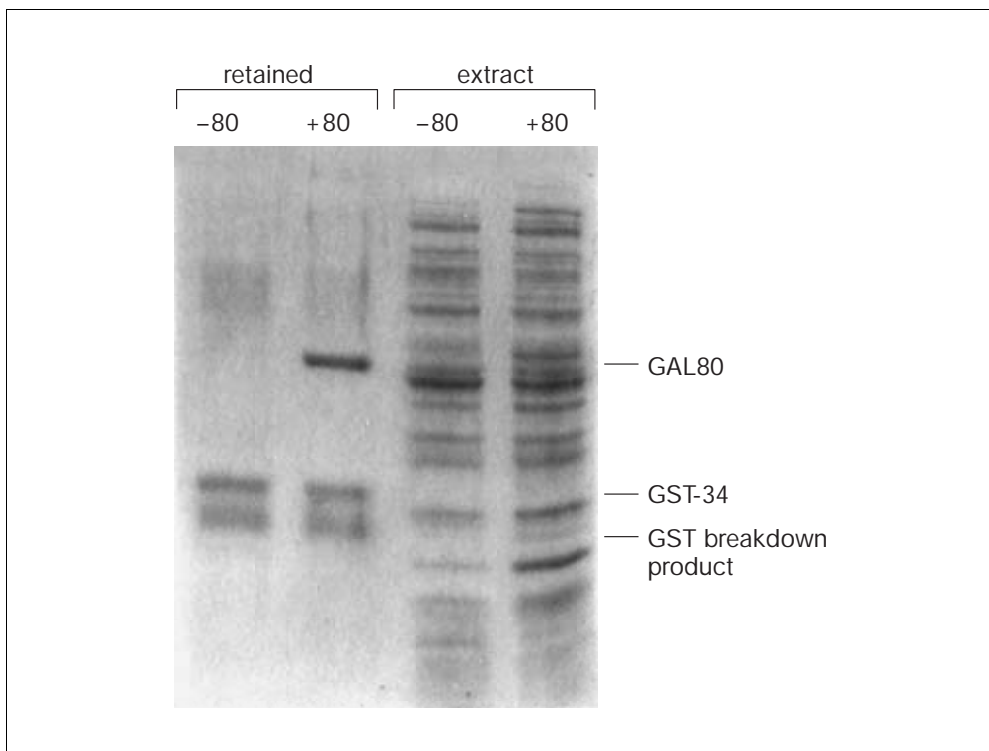


Figure 20.2.2 Purification of an interacting test protein from a crude *E. coli* extract (-80)/(+80). Crude *E. coli* extracts from control cells (-80) or cells expressing the yeast negative regulator Gal80p (+80) were incubated with a GST fusion protein containing a 34-amino-acid fragment that included the activation domain of the yeast transactivator Gal4p (GST-34). After extensive washing, the beads were boiled in 2× SDS sample buffer and analyzed by SDS-PAGE. Proteins were visualized by Coomassie blue staining. This gel shows that the retained fraction includes GST that has lost the 34-mer bait protein. Gal80p binds to GST-34 as a dimer, hence the molar excess seen in the +80 retained fraction. Photograph provided by Karsten Melcher.

the agarose beads, because these are still able to bind via the GST moiety. If the point of fusion between GST and the bait protein is attacked, GST alone is bound to the agarose beads. If GST is degraded, the fusion protein is unable to bind to the beads. If there are a substantial number of GST fusions with N-terminal bait protein fragments, quantitation of active protein will require limiting proteolysis or purifying the full-length fusion protein.

Troubleshooting

If an interaction occurs between two proteins *in vivo*, a GST pulldown assay will usually detect it. If it is suspected that two proteins interact *in vivo* but the interaction is not detected in the pulldown assay, there are two possible explanations. One is that there is no interaction. The other is that something is suboptimal or missing in the assay. The bait protein may be sterically inhibited as a fusion protein. In this case, the bait protein can be switched to the N-terminus of GST or con-

structed using another affinity tag. Another possibility is that the bait protein is not properly modified when produced as a fusion protein in *E. coli*. In this case, the fusion protein can be produced in yeast (Higgins, 1995; Romanos et al., 1995; Strausberg and Strausberg, 1995), baculovirus (see *UNITS 16.9, 16.10 & 16.11*), or mammalian cells (see *UNITS 16.12, 16.13 & 16.14*). However, if one of these alternative systems is chosen, it is better to use an affinity tag other than GST (such as maltose or biotin) because the level of glutathione in eukaryotic cells can inhibit binding to the agarose beads. Finally, the binding reaction may require one or more additional factors. Investigating this possibility will require binding the test protein out of a cellular extract or adding cellular extract to the labeled, *in vitro*-translated test protein.

If the test protein is specifically retained on the bait protein agarose but not on the GST or negative control agarose, it does not necessarily mean that the bait and test proteins interact *in vivo*. Artifacts may result from mass-action

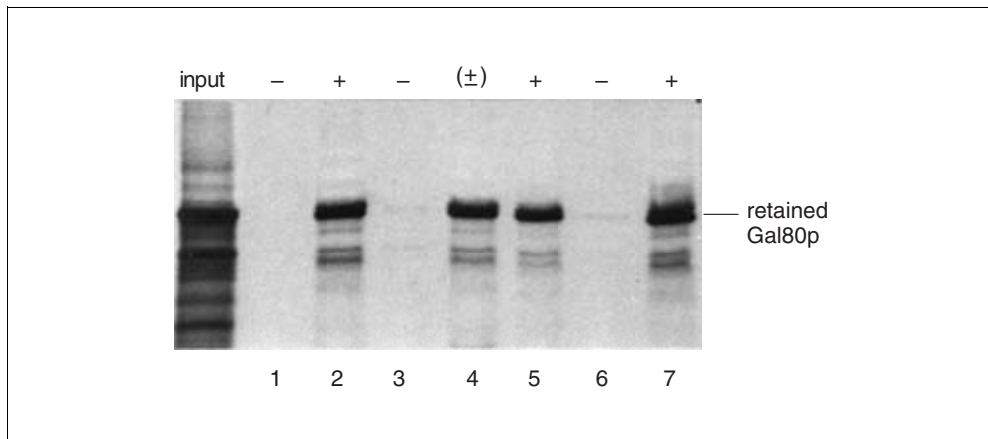


Figure 20.2.3 Replication in vitro of in vivo Gal4p/Gal80p interactions. In the presence of 10 to 20 μg bait protein and in the absence of complex competitor protein, Gal80p binds to all the mutant Gal4p activation domains tested, even those it does not repress in vivo (data not shown). However, this figure shows that when the amount of bait protein is reduced to 4 μg and 4 mg complex competitor is included, the in vitro pull-down assay parallels the in vivo interaction seen between Gal80p and several different Gal4p activation domain mutants. Lane 1, GST alone; Lane 2, GST + wild-type Gal4p activation domain; Lanes 3 to 7, GST + mutant Gal4p activation domains. A + indicates mutants with a positive in vivo Gal4p/Gal80p interaction; a – indicates mutants with negative in vivo Gal4p/Gal80p interaction; and \pm indicates mutants with low-level in vivo Gal4p/Gal80p interaction. Abbreviation: GST, glutathione-S-transferase. Photograph provided by Karsten Melcher.

effects of large amounts of bait or target protein that drive nonphysiological interactions. Artifacts may also be the consequence of taking a protein out of its normal context. A protein that normally exists in a complex may, when expressed as a GST fusion protein, expose many surfaces that are not normally exposed in the cell. These surfaces may be “sticky” and lead to a general affinity for other proteins. Artifacts can be minimized by reducing the amount of bait protein. Initial GST pull-down assays used tens or even hundreds of micrograms of bait protein. Such high concentrations can produce mass-action effects that drive any interactions with the test protein. While developing conditions to replicate in vitro the known in vivo interactions between the activation domain of the yeast transactivator Gal4p and its repressor Gal80p (Fig. 20.2.3), the authors found that the amount of bait protein needed to be reduced to 1 to 5 μg per reaction.

Inclusion of a complex protein competitor, such as an extract of soluble *E. coli* proteins, in the reaction mix also reduces artifactual interactions. This is very important for the reduction of nonspecific binding, for which a complex competitor is much more effective than a single competing protein like BSA. In the standard assay, the complex protein competitor is pre-

sent in ≥ 100 -fold excess over the GST fusion protein.

Anticipated Results

In the protocol described, with *E. coli* competitor, typically 80% of the input test protein binds to the bait if the affinity is $\leq 10^{-8}$ M. For proteins with binding constants of 10^{-6} to 10^{-8} M, typically 10% to 40% of input test protein is retained. Of course, with large amounts of bait and no competitor even weakly binding proteins can be detected.

If the interaction assay identifies a novel or unknown protein, the system can be scaled up to obtain biochemical amounts of the novel protein for further analysis and identification. After elution with protease or glutathione, the target protein should be highly enriched. Further purification can then be performed following standard protocols (see Chapter 10).

Time Considerations

One of the attractive features of this protocol is how quickly it can be performed. Using the incubation times outlined in the protocol description and labeled test protein, it is easy to determine whether the test protein binds to the GST fusion protein within 1 day. Analysis of the bound test protein by immunoblotting (*UNIT 10.8*) requires an additional day. Construction,

expression, and purification of the GST fusion protein and binding to beads (*UNIT 16.7*) requires additional time.

Literature Cited

- Agarana, C.E., Kuntz, I.D., Birkin, S., Axel, R., and Cantor, C.R. 1986. Molecular cloning and nucleotide sequence of the streptavidin gene. *Nucl. Acids Res.* 14:871-882.
- Blanar, M.A. and Rutter, W.J. 1992. Interaction cloning: Identification of a helix-loop-helix zipper protein that interacts with c-fos. *Science* 256:1014-1018.
- Chandler, C.S. and Ballard, F.J. 1988. Regulation of the breakdown rates of biotin-containing proteins in Swiss 3T3-L1 cells. *Biochem. J.* 251:749-755.
- Chen, J.-L., Attardi, L.D., Verrijzer, C.P., Yokomori, K., and Tjian, R. 1994. Assembly of recombinant TFIIID reveals differential coactivator requirements for distinct transcriptional activators. *Cell* 79:93-105.
- Higgins, D.R. 1995. Overview of protein expression in *Pichia pastoris*. In *Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 5.7.1-5.7.16. John Wiley & Sons, New York.
- Johnsson, N. and Varshavsky, A. 1994. Split ubiquitin as a sensor of protein interactions in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 91:10340-10344.
- Lim, P., Rhode, M., Morris, C.P., and Wallace, J.C. 1987. Pyruvate carboxylase in the yeast pyc mutant. *Arch. Biochem. Biophys.* 258:219-264.

- Nikolau, B.J., Wurtele, E.S., and Stumpf, P.K. 1985. Use of streptavidin to detect biotin-containing proteins in plants. *Anal. Biochem.* 149:448-453.
- Romanos, M.A., Clare, J.I., and Brown, C. 1995. Culture of yeast for the production of heterologous proteins. In *Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 5.8.1-5.8.17. John Wiley & Sons, New York.
- Strausberg, R.L. and Strausberg, S.I. 1995. Overview of protein expression in *Saccharomyces cerevisiae*. In *Current Protocols in Protein Science* (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 5.6.1-5.6.7 John Wiley & Sons, New York.

Key Reference

- Melcher, K. and Johnston, S.A. 1995. Gal4 interacts with TATA-binding protein and co-activators. *Mol. Cell Biol.* 15:2839-2848.

Describes binding from extracts, binding of in vitro-translated and labeled protein, binding-deficient mutants, correlation of in vitro and in vivo results, and determination of binding constant.

Contributed by Jonathan C. Swaffield and
Stephen Albert Johnston
University of Texas Southwestern
Medical Center
Dallas, Texas

Phage-Based Expression Cloning to Identify Interacting Proteins

Interaction cloning (also known as expression cloning) is a technique to identify and clone genes which encode proteins that interact with a protein of interest, or “bait” protein. Phage-based interaction cloning requires a gene encoding the bait protein and an appropriate expression library constructed in a bacteriophage expression vector, such as λ gt11. The gene encoding the bait protein is used to produce recombinant fusion protein in *E. coli*. The cDNA is radioactively labeled with ^{32}P . A recognition site for cyclic adenosine 3', 5'-phosphate (cAMP)-dependent protein kinase (protein kinase A; PKA) is introduced into the recombinant fusion protein to allow its enzymatic phosphorylation by PKA and [γ - ^{32}P]ATP. The procedure presented here (see Basic Protocol) involves a fusion protein consisting of bait protein and glutathione-S-transferase (GST) with a PKA site at the junction between them (the protocol can, however, be adapted to use other PKA-containing recombinant proteins). The labeled protein is subsequently used as a probe to screen a λ bacteriophage-derived cDNA expression library, which expresses β -galactosidase fusion proteins that contain in-frame gene fusions. The phages lyse cells, form plaques, and release fusion proteins that are adsorbed onto nitrocellulose membrane filters. The filters are blocked with excess nonspecific protein to eliminate nonspecific binding and probed with the radiolabeled bait protein (see Fig. 20.3.1). This procedure leads directly to the isolation of genes encoding the interacting protein, bypassing the need for purification and microsequencing or for antibody production.

NOTE: Radioactive label is used in this protocol, and appropriate precautions and shielding should be used (*APPENDIX 1F*).

STRATEGIC PLANNING

There are two important choices one must make to begin this procedure: (1) how to design the bait protein and (2) how to construct or acquire an appropriate phage-derived expression library.

Vectors for recombinant fusion protein expression that contain a PKA recognition site can be obtained commercially. Several companies now sell these vectors with various affinity tags such as GST (Pharmacia Biotech), histidine (Novagen), or calmodulin-binding protein (Stratagene). Alternatively, one can engineer the PKA recognition site (the five-amino acid sequence RRASV) into existing vectors by using synthetic DNA that encodes it.

Lambda-derived expression libraries that direct the expression of cDNAs made from many different mRNA sources are widely available. Alternatively, a library may be constructed for a particular experimental purpose; when designing a library to be used in expression cloning, several points should be considered. Libraries made from cDNA synthesized with random primers or an oligo dT-primer during first-strand synthesis (*UNITS 5.5 & 5.6*) can often be advantageous in that multiple clones representing different portions of the same protein can be identified in the screening. Analysis of coding regions present in these multiple clones could provide useful information about what region of the protein is responsible for the observed interaction. Full-length cDNA clones can subsequently be obtained from other available libraries.

Many suitable λ vectors are available for constructing cDNA expression libraries. The most widely used cDNA expression vector has been λ gt11 (Huynh et al., 1985). There are modifications of these λ vectors, however, that facilitate the recovery of the cDNA

Contributed by Julie M. Stone

Current Protocols in Molecular Biology (1997) 20.3.1-20.3.9

Copyright © 1997 by John Wiley & Sons, Inc.

inserts by avoiding the necessity of time-consuming λ phage DNA preparations. These modified vectors include those that make use of the *cre-lox* recombination for in vivo conversion of recombinant phages into plasmid DNA in *cre* recombinase-expressing host strains (e.g., λ ZipLox, Life Technologies) as well as those that employ helper phage for in vivo excision of a phagemid vector (e.g., the λ ZAP series of vectors, Stratagene). Most λ vectors available produce fusions of the cDNA inserts to β -galactosidase, because they are cloned into the *lacZ* gene. However, some λ -derived expression vectors (e.g., λ SCREEN-1, Novagen) direct expression of proteins fused to the T7 DNA polymerase promoter (gene *10* under control of the T7 promoter) and may yield higher expression of the library proteins (Margolis et al., 1992).

INTERACTION CLONING

Phage-based interaction expression cloning is a simple, rapid, and powerful technique to identify interacting proteins. A protein of interest is expressed as a recombinant fusion protein and labeled with ^{32}P at a serine residue in an engineered PKA recognition site to facilitate detection. β -galactosidase proteins that are fused in-frame to cDNA inserts in a bacteriophage λ -derived expression library are produced by the phage and adsorbed onto nitrocellulose filters. The filters are then screened with the radiolabeled protein probe to identify phage clones that express an interacting protein.

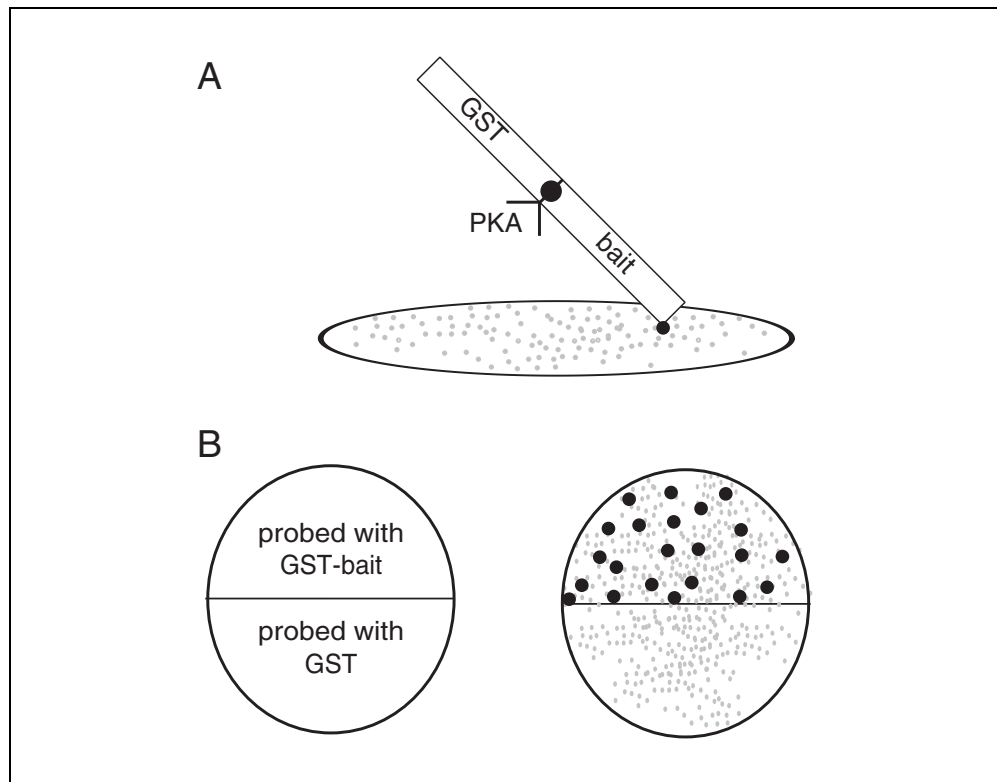


Figure 20.3.1 Schematic representation of the interaction cloning technique used to identify proteins that associate with a protein of interest (bait protein), and the expected results of a successful screen. **(A)** Expression of β -galactosidase fusion proteins from in-frame cDNA inserts of a λ gt11 *Arabidopsis* library is induced with IPTG-impregnated nitrocellulose membrane filters (indicated by an oval). Filters are probed with GST-bait fusion protein labeled with ^{32}P (•) at the PKA recognition site located at the junction of the fusion. Interacting clones are detected by autoradiography. **(B)** Representative autoradiogram of a tertiary screen of a positive plaque after the control experiment to determine whether the interaction is specific for the bait portion of the probe. The top half of the filter is probed with GST-bait while the bottom half is probed with GST alone as a control.

Materials

cAMP-dependent protein kinase (PKA; e.g., 250-U lots from Sigma)
40 mM DTT, prepared fresh
10× PKA buffer (see recipe)
10 mCi/ml [γ - 32 P]ATP (6000 mCi/mmol)
Purified glutathione-*S*-transferase (GST)-bait protein fusion protein with a PKA recognition site (UNIT 16.7), at ~0.1 to 1 μ g/ μ l concentration
Z'-KCl (see recipe), ice cold
Sephadex G-50 equilibrated in Z'-KCl
E. coli Y1090r⁻ or other appropriate host strain
LB medium containing appropriate selective antibiotic (see UNIT 1.1, APPENDIX 3F, and Table 1.4.1), 10 mM MgSO₄, and 0.2% maltose
10 mM MgSO₄
10 mM IPTG (Table 1.4.2)
150- or 100-mm LB plates (with antibiotic, if necessary; UNIT 1.1)
0.7% top agarose (UNIT 1.1), 47°C
Tris-buffered saline with Triton X-100 (TBS-T; see recipe)
India ink
HEPES blocking buffer (HBB; see recipe)
Binding buffer (BB; see recipe)
Suspension medium (SM; UNIT 6.12)
Chloroform
3-ml disposable plastic columns *or* disposable syringe and glass wool
Scintillation counter and fluid
Tabletop centrifuge or equivalent
Nitrocellulose membrane filters (137- and 82-mm disks)
22-G needle
Additional reagents and equipment for preparation and purification of recombinant glutathione-*S*-transferase fusion protein (UNIT 16.7), SDS-PAGE (optional; UNIT 10.2), autoradiography (APPENDIX 3A), titering and plating λ phage to generate plaques (UNIT 1.11 & UNIT 6.1), and purification of bacteriophage clones (UNIT 6.5)

Prepare the 32 P-labeled protein probe

1. Resuspend 250 U PKA in 25 μ l freshly prepared 40 mM DTT. Let the reconstituted enzyme stand at room temperature ~10 min before use.

It is important to use freshly prepared 40 mM DTT. PKA is extremely unstable after reconstitution; the enzyme stock solution can be stored temporarily at 4°C but retains activity for only 2 to 3 days.

2. Prepare a phosphorylation reaction mixture containing:

1 μ l 10 U/ μ l PKA (from step 1)
3 μ l 10× PKA buffer
5 μ l 10 mCi/ml (6000 mCi/mmol) [γ - 32 P]ATP
1 to 10 μ l (~1 μ g) purified GST-bait fusion protein
H₂O to 30 μ l.

Incubate 1 hr at room temperature.

A fusion protein unrelated to the bait protein, but containing the PKA recognition site, should also be expressed for use as a control to determine whether the observed interaction is specific for the bait moiety. Due to the instability of PKA, it is advisable to label both the bait protein and the unrelated control fusion protein simultaneously (in separate reactions).

3. Add 170 μ l ice-cold Z'-KCl (to stop reaction) and store on ice until use.
4. Prepare a gel filtration column by pouring Sephadex G-50 equilibrated in Z'-KCl into a 3-ml disposable plastic column (or a 3-ml syringe with a glass wool plug) for a final bed volume of \sim 3 ml. Allow the column to drain until the level of the buffer is at the top of the column bed.

The Sephadex G-50 can be swelled in water or buffer and stored at 4°C for months. In this case, five to ten column volumes of Z'-KCl can be used to equilibrate the column after it is poured.

5. Load the entire phosphorylation reaction (from step 3) onto the column and collect the effluent in a 1.5-ml microcentrifuge tube. Place a second tube under the column, add a 200- μ l aliquot of Z'-KCl, and collect the effluent again. Repeat the above loading and collecting steps an additional ten times to collect a total of twelve 200- μ l fractions.
6. Measure the Cerenkov counts with a scintillation counter to identify the fractions with the highest specific activities, and calculate cpm/ μ l.

Typically two peaks of radioactivity are observed. The first peak elutes in fractions five to nine and corresponds to labeled protein. The second peak, usually found in the last few fractions collected, corresponds to unincorporated ATP and should be discarded. The hottest fraction(s), which elute first, should be used. Fractions can be stored at 4°C for several weeks.

7. *Optional:* Analyze a small amount (1 to 2 μ l) of the 32 P-labeled protein probe by SDS-PAGE and autoradiography (UNIT 10.2 & APPENDIX 3A).

Typically two strong signals are observed: one at the predicted size of the GST-bait fusion protein and the other at the predicted molecular weight of the GST alone (28 kDa). This is due to the fusion protein's inherent susceptibility to protease cleavage at the GST-bait protein junction. The presence of labeled GST protein in the probe should not interfere with the screen; a control experiment with labeled GST will be performed at the later stages of screening.

Prepare host strain cells and dilution of bacteriophage cDNA library

8. Using serial dilution as described in UNIT 1.11, determine the titer of the bacteriophage cDNA library.
9. Grow an overnight 50-ml culture of *E. coli* Y1090^r (or other appropriate host strain; UNIT 1.4) in LB medium containing an appropriate selective antibiotic, 10 mM MgSO₄, and 0.2% maltose.
10. Centrifuge cells 10 min at 2000 \times g, room temperature, and resuspend in 25 ml of 10 mM MgSO₄.

The resuspended cells can be stored up to 1 week at 4°C before use.

Prepare the filters to screen the bacteriophage cDNA expression library

11. Soak nitrocellulose filters in 10 mM IPTG for 15 min at room temperature and air dry.

IPTG-impregnated filters can be stored in a petri plate until use.

12. Prepare eight 1.5-ml microcentrifuge tubes each containing 0.6 ml Y1090^r cells (from step 10) and \sim 40,000 pfu bacteriophage (from step 8), and incubate 15 min at 37°C.

13. Add contents of each tube to 7 ml of 0.7% top agarose at 47°C and pour onto 150-mm LB plates (with antibiotic, if necessary). Incubate plates ~3 hr at 42°C, until small plaques are visible.
14. Overlay the plates with IPTG-impregnated filters and incubate an additional 6 to 8 hr at 37°C.

Formation of plaques in the absence of induction of the lacZ gene promoter ensures that any library-encoded proteins that are deleterious to phage growth will not be expressed while the phage are forming a plaque.

Incubation in the presence of IPTG-impregnated filters is usually performed for a 6- to 8-hr period, but can proceed overnight for convenience.

15. Chill plates 15 min (or overnight) at 4°C.
16. Pierce each filter in several locations with a 22-G needle dipped in India ink to mark orientation. Remove filters from plates and wash 15 min in TBS-T at room temperature with shaking.

Screen bacteriophage cDNA expression library

17. Incubate filters 1 to 4 hr with rocking at 4°C in 100 ml HBB.

This blocking step (to reduce nonspecific binding) can also be performed overnight for convenience.

18. Incubate overnight with rocking at 4°C in BB containing $2.5\text{--}5 \times 10^5$ cpm/ml of radiolabeled fusion protein (from step 6).

All eight filters can be placed in one 150-mm plate and probed with 30 to 40 ml of solution. The plate is wrapped in Parafilm and placed in a Plexiglas box for shielding. Alternatively, the filters can be placed with solution in a heat-sealable bag. The probe solution can be stored at 4°C and used for the subsequent secondary and tertiary screenings.

19. Wash membrane filters three times, 10 min each, in 100 ml BB with shaking at room temperature. Air dry and expose to film (see APPENDIX 3).
20. Using the large end of a Pasteur pipet, take an agarose plug at the position of the positive clone. Place into a 1.5-ml microcentrifuge tube and add 1 ml SM and 1 drop of chloroform.

Agarose plugs can be stored at 4°C for months. Therefore, if many putative positive clones are identified in the primary screen, they can all be stored for future analysis if necessary.

21. Determine the titer by serial dilution (UNIT 1.11) and perform successive screening procedures to obtain purified clones as described in UNIT 6.5.

Subsequent screens are performed on 100-mm LB plates with ~2000 pfu/plate for secondary screens and ~300 to 500 pfu/plate for tertiary screens. Before purified clones are obtained (usually during the tertiary screen), a control to eliminate clones that might interact with the GST fusion portion of the probe should be performed. To do this, cut the filters in half and probe one half with the labeled GST fusion to the protein of interest and the other with labeled GST alone or an unrelated control GST fusion protein (see Fig. 20.3.1).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. All solutions should be prepared from sterile, autoclaved stock solutions, except Z'-KCl, which should be filter sterilized. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Binding buffer (BB)

20 mM HEPES·OH, pH 7.4
7.5 mM KCl
0.1 mM EDTA
2.5 mM MgCl₂
1% (w/v) nonfat dry milk

The solution can be prepared without milk and stored indefinitely at room temperature (add milk prior to use).

HEPES blocking buffer (HBB)

20 mM HEPES·OH, pH 7.4
5 mM MgCl₂
1 mM KCl
5% (w/v) nonfat dry milk

The solution can be prepared without milk and stored indefinitely at room temperature (add milk prior to use).

PKA buffer, 10×

200 mM Tris·Cl, pH 7.5
10 mM DTT
1 M NaCl
120 mM MgCl₂
Store up to 6 months at room temperature

Tris-buffered saline with Triton X-100 (TBS-T)

10 mM Tris·Cl, pH 8.0
150 mM NaCl
0.05% (v/v) Triton X-100
Store up to 6 months at room temperature

Z'-KCl

25 mM HEPES·OH, pH 7.4
12.5 mM MgCl₂
20% (w/v) glycerol
100 mM KCl
1 mg/ml BSA
1 mM DTT
Filter sterilize
Store up to 6 months at 4°C

COMMENTARY

Background Information

Historically, interacting proteins have been isolated by biochemical approaches, which required purification of the interacting protein for antibody production or microsequencing before a clone encoding the protein could be identified. Two of the molecular biological approaches described in this chapter, the yeast two-hybrid system (UNIT 20.1) and phage-based

interaction expression cloning (this unit), however, directly yield a clone encoding the interacting protein.

Bacteriophage cDNA expression libraries are commonly screened using antibodies (see UNIT 6.7 and references therein) or radiolabeled DNA probes to identify DNA-binding proteins (see UNIT 12.7 and references therein). Modifications to identify interacting proteins include

screening with the protein of interest (the bait protein) and detecting that protein with antibodies (Chapline et al., 1993). However, by screening with a radiolabeled protein probe, one avoids the additional incubations and washes which are necessary for immunodetection but that increase the likelihood of disrupting weak protein-protein interactions. ¹²⁵I-labeled protein probes have been used successfully to screen for interacting proteins (Hoeffler et al., 1991), although the technique described in this unit avoids the need to label the proteins with ¹²⁵I, and the complications of handling this isotope.

Protein probes autophosphorylated with ³²P were originally used to screen cDNA expression libraries in the isolation of proteins that interact with receptor protein kinases, a technique referred to as CORT (cloning of receptor targets; Skolnik et al., 1991; Lowenstein et al., 1992; Margolis et al., 1992). By introducing a PKA recognition site into the protein probe, the technique was made suitable for proteins that were not themselves protein kinases (Blaner and Rutter, 1992; Kaelin et al., 1992).

Phage-based interaction expression cloning as described in this unit has been used successfully to identify many interacting proteins, but may not be successful for all types of interactions. For example, many related proteins (known as A-kinase anchoring proteins, or AKAPs) have been identified by their ability to interact with the type II cAMP-dependent kinase regulatory subunit, RII (Carr and Scott, 1992). All the AKAPs can bind RII under denaturing conditions (Lester et al., 1996), illustrating that the success of interaction cloning is often dependent on the nature of the interactions and that interactions less dependent on three-dimensional structure may be favored. However, there are many examples of proteins identified by this technique that cannot interact under the denaturing conditions often used to confirm the binding (commonly referred to as overlay assays or Far Western analysis). For example, a protein domain identified by its interaction with a plant receptor-like protein kinase is capable of interacting with a denatured form of the bait protein, but is unable to interact when the interacting protein is denatured (Stone et al., 1994). This fact is consistent with the idea that immobilization on membranes denatures some proteins but not others.

Critical Parameters and Troubleshooting

The success of phage-based interaction expression cloning is inherently dependent on the

quality of the protein probe used and the extent of representation of the bacteriophage cDNA library. GST fusion proteins are often obtained in high quantity in a soluble form, avoiding the necessity of solubilizing and refolding during protein purification from *E. coli* extracts. In most cases the PKA recognition site is readily accessible, and allows production of radiolabeled protein with a high specific activity.

The strength of the protein-protein interaction is also critical. Depending on the nature of the interaction, a filter-binding technique might not be suitable. This technique may not detect weak or transient interactions, such as an enzyme/substrate interaction, and techniques such as the yeast two-hybrid system (*UNIT 20.1*) might be more appropriate.

If no positive clones are identified or if background is high, it may be helpful to add reducing agents or detergents or to alter the salt conditions of the binding and wash solutions (Vinson et al., 1988). Moreover, denaturation and renaturation of the proteins on the filters using 6 M guanidine-HCl, as described in *UNIT 12.7* (Alternate Protocol), may facilitate the recovery of clones expressing interacting proteins, because adsorption of the β -galactosidase fusion proteins to nitrocellulose filters may alter the conformation of the proteins.

For a discussion of common problems with this procedure and their diagnosis and possible solutions, see Table 20.3.1.

Anticipated Results

Depending on the nature of the interaction being sought by this technique, many interacting clones, none, or just a few may be identified. However, the technique is as simple as screening a library by hybridization and often well worth the effort. In any case, observed interactions should be confirmed by other means.

Time Considerations

The interaction cloning technique described in this unit is extremely rapid in comparison with other techniques to identify interacting proteins, such as the yeast two-hybrid system (*UNIT 20.1*). The technique does not require any unusual reagents that would not be readily available in any laboratory routinely engaged in molecular biology, other than the cAMP-dependent protein kinase. Once the appropriate recombinant fusion protein and cDNA expression library are obtained, the primary screen can be completed in a few days. Once the appropriate construct for expression of the GST-bait is obtained, purification of the recom-

Table 20.3.1 General Troubleshooting Guide for Interaction Cloning

Problem	Possible cause	Solution
High background	Insufficient washing	Use more BB wash solution
	Poor choice of nitrocellulose membrane filters	Try a different supplier for membranes
	Insufficient blocking	Block overnight in HBB
No positive plaques	Poor-quality protein probe	Check the probe by SDS-PAGE and autoradiography
	Wash conditions too stringent	Reduce wash times Vary salt and/or detergent concentration
	Poorly folded library fusion proteins	Perform 6 M guanidine-HCl denaturation/renaturation (<i>UNIT 12.7</i>)
Too many positive plaques	Wash and binding conditions insufficiently stringent	Increase salt and/or detergent in BB
	Library biased to proteins that interact with the affinity tag	Include unlabeled affinity tag in blocking, binding, and wash solutions

binant GST-bait protein can be achieved in 1 day. ³²P labeling of the bait protein requires 2 hr (several additional hours if the optional SDS-PAGE and autoradiography are performed). The initial screening of the bacteriophage cDNA library takes 3 days; day 1 to prepare the filters and induce expression of library-encoded proteins, day 2 for blocking and probing overnight with ³²P-labeled bait protein, and day 3 for washing and autoradiography. Subsequent purification of the cDNA clones should only take an additional week. There are a number of steps that can be performed either rapidly or overnight, providing a great deal of convenience and flexibility.

Literature Cited

- Blanar, M.A. and Rutter, W.J. 1992. Interaction cloning: Identification of a helix-loop-helix zipper protein that interacts with c-Fos. *Science* 256:1014-1018.
- Carr, D.W. and Scott, J.D. 1992. Blotting and band-shifting: Techniques for studying protein-protein interactions. *Trends Biochem. Sci.* 17:246-249.
- Chapline, C., Ramsay, K., Klauck, T., and Jaken, S. 1993. Interaction cloning of protein kinase C substrates. *J. Biol. Chem.* 268:6858-6861.
- Hoeffler, J.B., Lustbader, J.W., and Chen, C.Y. 1991. Identification of multiple nuclear factors that interact with cyclic AMP response element-binding protein and activation transcription factor-2 by protein interactions. *Mol. Endocrinol.* 5:256-266.
- Huynh, T.V., Young, R.A., and Davis, R.W. 1985. Constructing and screening cDNA libraries in λ gt10 and λ gt11. *In* DNA Cloning: A Practical Approach (D.M. Glover, ed.) pp. 49-78. IRL Press, Oxford.
- Kaelin, W.G.J., Krek, W., Sellers, W.R., DeCaprio, J.A., Ajchenbaum, F., Fuchs, C.S., Chittenden, T., Li, Y., Farnham, P.J., Blanar, M.A., Livingston, D.M., and Flemington, E.K. 1992. Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F-like properties. *Cell* 70:351-364.
- Lester, L.B., Coghlan, V.M., Nauert, B., and Scott, J.D. 1996. Cloning and characterization of a novel A-kinase anchoring protein: AKAP220, association with testicular peroxisomes. *J. Biol. Chem.* 271:9460-9465.
- Lowenstein, E.J., Daly, R.J., Batzer, A.G., Li, W., Margolis, B., Lammers, R., Ullrich, A., Skolnik, E.Y., Bar-Sagi, D., and Schlessinger, J. 1992. The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell* 70:431-442.
- Margolis, B., Silvennoinen, O., Comoglio, F., Roonprapunt, C., Skolnik, E., Ullrich, A., and Schlessinger, J. 1992. High-efficiency expression/cloning of epidermal growth factor-receptor-binding proteins with *src* homology 2 domains. *Proc. Natl. Acad. Sci. U.S.A.* 89:8894-8898.
- Skolnik, E.Y., Margolis, B., Mohammadi, M., Lowenstein, E., Fischer, R., Drepps, A., Ullrich, A., and Schlessinger, J. 1991. Cloning of PI3 kinase-associated p85 utilizing a novel method of expression/cloning of target proteins for receptor tyrosine kinases. *Cell* 65:83-90.

- Stone, J.M., Collinge, M.A., Smith, R.D., Horn, M.A. and Walker, J.C. 1994. Interaction of a protein phosphatase with an *Arabidopsis* serine-threonine receptor kinase. *Science* 266:793-795.
- Vinson, C.R., LaMarco, K.L., Johnson, P.F., Landschulz, W.H., and McKnight, S.L. 1988. In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. *Genes & Dev.* 2:801-806.

Key References

- Blanar and Rutter, 1992. See above.
The basic protocol described in this unit is modified directly from the Blanar and Rutter protocol.
- Huynh et al., 1985. See above.
Provides an excellent description of constructing and screening λ gt11 cDNA expression libraries.

Contributed by Julie M. Stone
University of Missouri, Columbia
Columbia, Missouri

Surface Plasmon Resonance for Measurements of Biological Interest

UNIT 20.4

Genetic manipulations, including gene knockouts and mutant screens, provide an initial hint as to the function of a gene product and indicate possible associated factors. To unravel complicated biological processes that control the development of organisms one must identify the interacting components. Yeast two-hybrid methods (UNIT 20.1; Gyuris et al., 1993; Zervos et al., 1993; Bartel and Fields, 1995) provide an *in vivo* means of identifying species that directly associate with each other. However, many biological systems are notoriously redundant and consist of networks of interacting components. For these and other reasons, it is advisable to verify the *in vivo* observation by an *in vitro* detection method where there are fewer variables and where reaction conditions can be controlled. Few techniques measure equilibrium binding because routine washing steps can cause the dissociation of the interaction under study. Moreover, most *in vitro* techniques used for the detection of intermolecular interactions require the use of antibodies or protein-labeling and subject the interaction under study to harsh and potentially denaturing conditions. These techniques only provide static “snap-shots” of a dynamic process, and kinetic parameters must be inferred from successive experiments.

An *in vitro* technique based on an optical phenomenon, called surface plasmon resonance (SPR; Liedberg et al., 1983; Daniels et al., 1988), can simultaneously detect interactions between unmodified proteins and directly measure kinetic parameters of the interaction. The use of SPR to study macromolecular interactions has recently become popular with the availability of user-friendly instruments. The most popular commercially available SPR device is the BIAcore instrument (BIAcore). This instrument consists of sensing optics, an automated sample delivery system, and a computer for instrument control, data collection, and data processing. Experiments are performed on disposable chips (Löfås and Johnsson, 1990). In practice, a ligand protein is immobilized on the chip while buffer continuously flows over the surface. The buffer flow is interrupted when sample “plugs” containing putative binding partners (analyte or target proteins) are sequentially flowed over the protein surface. The sensing apparatus monitors changes in the angle of minimum reflectance from the interface that result when a target protein associates with the ligand protein. Molecular interactions can be directly visualized (on the computer monitor) in real time as the optical response is plotted against time. This response is measured in resonance units (RUs, where 1000 RUs = 1 ng protein/mm²).

The user handbook with the BIAcore SPR apparatus contains detailed information on the use of the equipment, as well as a number of protocols and information for experimental design. Basic Protocol 1, presented here, is a modification of the standard protocol in the BIAcore user handbook, and is used with the CM-5 BIAcore carboxylated dextran chip. Basic Protocol 2 describes the particular details necessary when using NTA-SAM chips. A detailed discussion on choosing between these two kinds of chips can be found in Critical Parameters and Anticipated Results. The use of SPR provides an accurate measurement of kinetic rate constants for a ligand-target binding reaction. If a ball-park estimate of the range of the equilibrium constants is desired, or if an absolute kinetic rate constant is desired, refer to the discussion in Anticipated Results.

SPR USING BIAcore CHIPS

The following protocol is used when performing an SPR experiment on a BIAcore device with the BIAcore CM-5 chip. This is the most commonly used SPR chip and is made with carboxylated dextran. Two particular deviations from the standard BIAcore protocol are discussed within annotations (steps 3 and 6); both deviations concern ligand binding to the chip, and were designed to avoid artifactual pitfalls that are peculiar to dextran-based chips. For detailed steps in the performance of an SPR experiment, and for additional reagents and equipment information, refer to the BIAcore user handbook.

Materials

- CM-5 dextran chips (BIAcore)
- PBS (*APPENDIX 2*)
- Ligand protein
- Target protein
- Amine-coupling kit (BIAcore), containing:
 - N*-ethyl-*N'*-[(dimethylamino) propyl] carbodiimide hydrochloride (EDC)
 - N*-hydroxysuccinimide (NHS)
 - Sodium acetate buffer, low pH
 - Ethanolamine
- BIAcore SPR equipment
- BIA evaluation point-and-click software

Determine binding conditions

1. Insert a new BIAcore CM-5 dextran chip.
2. Allow the system to equilibrate with PBS as the running buffer in the constant flow system.
3. Experimentally determine buffer conditions and concentrations that minimize non-specific binding of target and ligand proteins to the carboxylated dextran.

With dextran chips, conditions must be determined for the ligand because a ligand which has nonspecifically adsorbed to the highly negatively charged dextran is likely to be denatured and can nonspecifically bind the target protein in solution.

This process can easily use up all four flow channels of two chips.

Determine regeneration conditions

4. Insert new chip.
5. Activate the carboxylates of the dextran support by injecting *N*-ethyl-*N'*-[(dimethylamino) propyl] carbodiimide hydrochloride (EDC) in the presence of *N*-hydroxysuccinimide (NHS).

These reagents are included in an amine-coupling kit from BIAcore.

6. Inject ligand protein in low-pH sodium acetate buffer such that an adequate amount of the protein is presented on the dextran. Perform several parallel experiments in which the density of ligand presented by the dextran is varied.

This can be accomplished by altering contact times and protein concentration. Proceed with steps 7 through 14, and then measure the rates of dissociation of the target protein from each of the ligand densities (step 15). The rate of dissociation should be constant and independent of the concentration of the surface density of the ligand (see Anticipated Results). The ligand density should be optimized to produce a surface presenting enough molecules to elicit a clear binding response yet dilute enough to measure a dissociation rate constant that is constant.

7. Cap any unreacted activated carboxylates by treatment with ethanolamine.
8. Inject a solution containing a known concentration of the putative target protein under buffer conditions and within the concentration range determined in step 3.

Vary the flow rate and the concentration of the analyte to minimize diffusion or “mass transport” effects. The association rate constant for a particular reaction should be just that: constant (see Anticipated Results discussion on Rate and equilibrium constants).
9. Experimentally determine regeneration conditions, i.e., conditions under which the target protein is completely dissociated from the covalently coupled ligand.

One must be careful to determine conditions for the regeneration of the surface that preserve the activity of the immobilized species.

This procedure can easily consume another chip.

Perform experimental measurements

10. Insert new chip, if necessary.
11. Perform steps 2, 5, 6, 7, and 8 (i.e., activate the dextran carboxylates, covalently attach the protein of interest, cap unreacted sites, and introduce a putative binding partner in solution).

Use the optimal buffer conditions and ligand density determined in steps 3 and 6, respectively.
12. Regenerate using reagents determined in step 9.
13. Repeat steps 8 and 12 as necessary to obtain data using a total of at least three different concentrations of the target protein.
14. Using BIA evaluation point-and-click software, measure dissociation and association rate constants. Calculate an equilibrium constant using average values of the association and dissociation rates that resulted from the three different target protein concentrations.

The kinetic parameters measured for the various concentrations of analyte in solution should be within reasonable experimental error limits, similar to the error limits obtained from one “identical” experiment to another.
15. Reverse the experimental configuration: confirm that the same kinetic parameters are measured when the previous analyte is immobilized and the previous ligand is in solution.

SPR USING NTA-SAM CHIPS

The following protocol was developed for SPR experiments using nitrilotri-acetic acid self-assembled monolayer (NTA-SAM; Bamdad et al., 1994; Sigal et al., 1996) chips. This protocol is simpler than Basic Protocol 1, as NTA-SAM chips do not require several of the control measures necessitated by dextran-based chips (see Critical Parameters discussion on NTA-SAM chips). For detailed steps in the performance of an SPR experiment, and for additional reagents and equipment information, refer to the BIAcore user handbook.

Materials

- NTA-SAM chips (3% to 5% NTA relative to an inert ethylene glycol-terminated thiol)
- PBS or HeBS (*APPENDIX 2*)
- 1 mM NaOH
- 1% (w/v) Ni(II)SO₄

**BASIC
PROTOCOL 2**

**Analysis of
Protein
Interactions**

20.4.3

Histidine-tagged ligand protein
Target protein
BIAcore SPR equipment
BIA evaluation point-and-click software

1. Insert an NTA-SAM chip.
2. Use PBS or HeBS as the running buffer.
3. Inject 10 μ l of 1 mM NaOH.
4. Inject 25 μ l of aqueous 1% Ni(II)SO₄.
5. Inject a histidine-tagged ligand protein.
6. Inject a putative target protein.
7. Repeat steps 3 through 6 with a total of at least three concentrations of the target protein.

Histidine-tagged proteins can be removed from the surface by treatment with 200-mM imidazole. Alternatively, a fresh flow cell can be used for each experiment (each chip has four flow cells).

8. Using BIA evaluation point-and-click software, measure dissociation and association rate constants. Calculate an equilibrium constant using average values of the association and dissociation rates that resulted from the three different target protein concentrations.

The kinetic parameters measured for the various concentrations of analyte in solution should be within reasonable experimental error limits, similar to the error limits obtained from one "identical" experiment to another.

COMMENTARY

Background Information

Use of SPR to analyze biomolecular interactions

Optical sensing devices for detecting and kinetically characterizing interactions among biomolecules are now in mainstream use. Complicated, multicomponent interactions can be quickly "dissected" by sequential determination of pair-wise interactions. The BIAcore SPR instrument can tolerate a wide range of buffer conditions. A broad range of affinities are measurable by SPR: from ten millimolar to picomolar. Because binding events are viewed in real time, one can alter experiments immediately in response to initial results. Researchers have eagerly adopted the new methods because they don't require time-consuming and hazardous labeling steps. Perhaps the most important advantage is that because the *mass* of a molecular species is detected by an optical response, no protein labeling is required and only microliter sample volumes are consumed. The microfluidic technology of these devices satisfies a long-standing research need—allow-

ing for both minimal sample consumption and rapid measurements.

The commercial availability of user-friendly optical sensors is a fairly new development. Until recently, the design of the devices has received little input from the bioscientific community. We are now entering a new and exciting phase in the development of this technology that is being driven by the needs of researchers to not only carry out routine experiments faster but to venture into unexplored areas. Derivatized monolayers will play an increasingly important role in state-of-the-art detection devices, as they have demonstrated clear advantages over disordered polymer-based supports (see Critical Parameters discussion on NTA-SAM chips). With optical sensing, it is now possible to design experiments to measure the difference between monovalent and bivalent binding, detect very weak interactions between proteins and protein complexes, and characterize interactions between factors that facilitate the assembly (or catalyze the disruption) of large megadalton complexes. The future evolution of these devices is only limited by the

imagination and needs of researchers in their desire to unravel the complex processes of biological systems.

There is an intrinsic danger in using any instrument as a black box. The fact that many biologists, largely untrained in theoretical physical chemistry, use the instrument to perform kinetic characterization of interactions has led to considerable skepticism among researchers regarding much of the BIAcore-generated data in the literature. An understanding of the physics of the instrument is essential for experimental design and data interpretation. Thus, this commentary treats (1) the theory behind SPR devices, (2) factors that affect the accuracy of an SPR experiment, (3) choice of SPR chips, (4) caveats for experimentation and additional control experiments, and (5) kinetic analysis and higher-order kinetics.

SPR theory

Optical phenomena involving electron oscillations in the surface of a metal have been studied for some time. Electrons move freely through a metal. However, at a metal-medium interface, dipole excitations cause the propagation of electron waves called surface plasma or plasmon waves (SPWs). Because the wave vectors of the SPWs are long, they cannot be excited by simple incident light, so direct coupling of the two waves is not possible. X-rays were originally used to excite SPWs and for this reason interest in the phenomenon was limited to physicists studying how electromagnetic energy behaves at surfaces. A major breakthrough came in the late 1960s, when researchers succeeded in exciting SPWs using simple incident light. Otto (1968a,b) and the team of Kretschmann and Raether (1968), building on the earlier work of Turbadar (1959), exploited the phenomenon of total internal reflection (TIR) to couple incident light to SPWs. This achievement made simple SPR devices possible.

The basic configuration of modern SPR instruments has not deviated significantly from the original experimental setup (see Fig. 20.4.1). Total internal reflection occurs at an interface, when light propagates from a medium of higher refractive index to one of lower refractive index at an angle of incidence greater than the critical angle. The light energy is then totally reflected off of the interface, except for an evanescent field that extends a distance of one wavelength (of the incident light) into the second medium. The magnitude and direction

of the reflected light can be monitored by an array detector. If a very thin layer of a third dielectric material, of lower refractive index than either of the others, is sandwiched between the two different dielectrics within the one wavelength distance, the incident wave will be transmitted through the thin layer and into the second medium by photon tunneling. If the sandwiched material is a metal and the wave vector of the incident light can somehow be increased (achieved by passing it through a prism) to match that of an SPW propagating through the metal, then coupling of the incident light and the SPW will occur. Energy from the incident light will be fed into the SPW and not reflected off the interface. This is important because the coupling of the two waves will produce a minimum or “dip” in the reflected light; the angle at which this occurs is a parameter exquisitely dependent upon the second medium’s dielectric properties (Wahling et al., 1979). Recall that a small portion of incident light extends into the second medium, so that the angle at which the incident wave couples to the SPW is determined by the refractive index of the second medium by Snell’s Law (assuming the first dielectric is held constant). Small changes in the refractive index of the second material cause large shifts in the angle at which coupling and the corresponding dip in reflected light occur.

In the study of protein-protein interactions, increasing the protein concentration of a solution will result in a more dense medium with an altered refractive index. One can infer how the protein content of the second material is changing over time by monitoring the change in the angle of minimum light intensity. If ligand is immobilized at the interface of the thin metal film and the second medium, then subsequent interaction with proteins injected into the second medium will cause a change in the refractive index of the interfacial region, again changing the minimum reflectance angle. It is important to appreciate that the attenuation of the “sensing wave” which projects into the second medium is severe. This allows kinetic information to be extracted from the optical response as a function of time because the sensing wave only detects proteins that are very close to the interface (where the first species is immobilized), as opposed to the total mass of protein in the solution sample. It is assumed that proteins that are so close to the immobilized species are associating.

Critical Parameters

Factors that affect the accuracy of an SPR experiment

The physics of the optical phenomenon of SPR govern the accuracy of the device. Because the sensing wave falls off very rapidly, the change in optical response is only a *linear* function of the net change in protein mass up to a maximum distance of 150 nm from the interfacial region (Stenberg et al., 1991). To reliably measure the kinetics of an interaction, the sensing apparatus should be as inert as possible. Additionally, immobilized ligands should be oriented to provide equal accessibility of binding sites to the binding partners in solution (analytes), and analytes in solution

should have equal access to the immobilized ligands.

Dextran-based chips

Virtually all of the commercially available chips that are suitable for the study of protein-protein interactions using a BIAcore device have a functionalized dextran base that is ~150 nm thick. This presents a problem, as it means that the processes under study are occurring at the very limit of accurate detection. Several other problems arise with the use of dextran-based SPR chips. The most commonly used chip is a carboxylated dextran chip, CM-5, that is available from BIAcore. Proteins are nonspecifically coupled to the carboxylates of the support by standard EDC/NHS chemistry, the

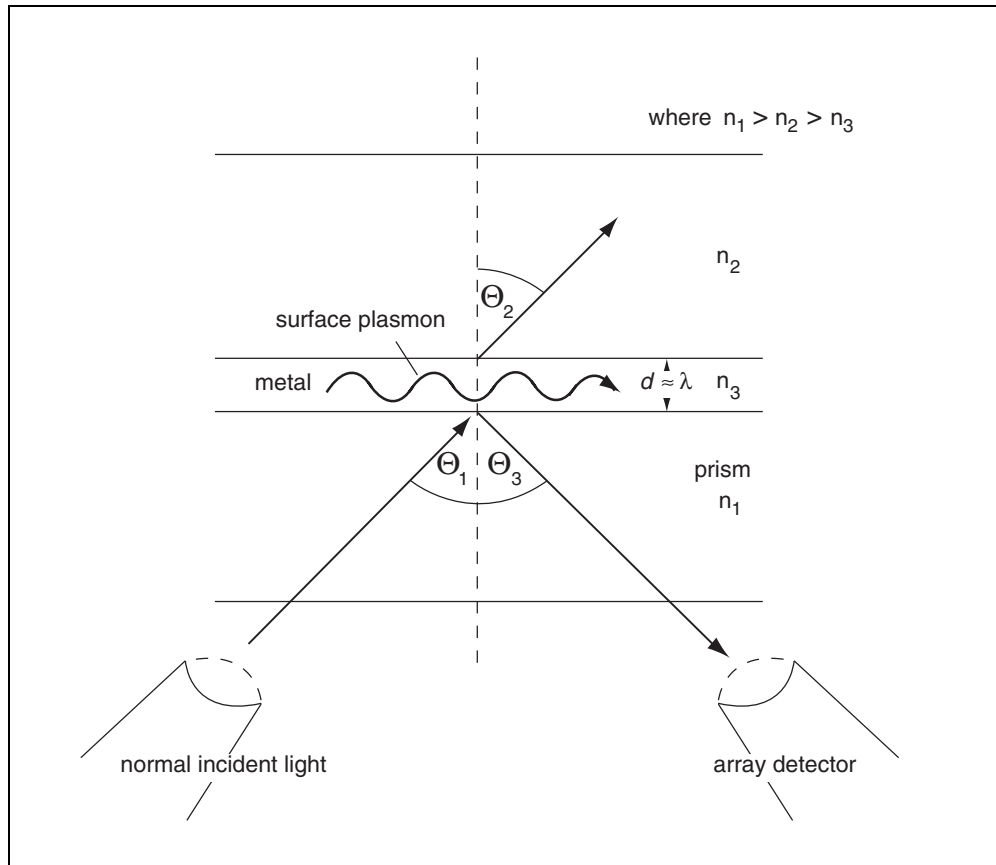


Figure 20.4.1 The physical configuration of modern SPR devices. Light passing through a dense medium (n_1) incident at an angle (Θ_1) greater than the critical angle, to a second medium of lower refractive index (n_2) is totally reflected off the interface, except for an evanescent field that extends into the second medium for a distance of one wavelength. At a certain angle (Θ_2) exquisitely dependent on the refractive index of the second medium, the incident light wave can couple to surface plasmon waves (electron oscillations) propagating through a thin metal film (n_3) sandwiched between the two media. The incident light energy is fed into the surface plasmon wave and a corresponding minimum in reflected light is observed. Proteins immobilized at (or recruited to) the metal-medium interface alter the dielectric of the region, which causes a shift in the angle of minimum reflectivity (Θ_3).

which targets primary amines on the protein surface. The carboxylated dextran carries a high net negative charge (Johnsson et al., 1991); it is thus not inert and may alter the kinetics of the system under study.

The dextran support is also problematic with respect to mutual accessibility of ligand and target proteins. Experiments indicate that very large proteins may be completely prevented from entering the dextran matrix (see Critical Parameters discussion on NTA-SAM chips). The nonspecific coupling of proteins to the support renders the ligand's active site randomly oriented and often inaccessible. Furthermore, flow over the nonuniform dextran polymer is turbulent. A large static layer separates the analyte-containing solution from the ligand-presenting dextran and hinders the diffusion of molecules into the matrix. Increasing the flow rate will help speed diffusion, but if the diffusion rate is slower than the association rate, then the diffusion rate will be the measured parameter. Although analytical methods have been proposed to model the diffusion or "mass transport" factor (Fisher and Fivash, 1994), it would be preferable to eliminate the problem rather than to model its effect into data evaluation software.

NTA-SAM chips

In 1994, a self-assembled monolayer (SAM; Nuzzo et al., 1987; Bain et al., 1989) SPR chip was developed that presents nitrilotri-acetic acid (NTA) ligands for the specific capture of histidine-tagged proteins to overcome the problems intrinsic to the dextran chip (Bamdad et al., 1994; Sigal et al., 1996). The thickness of the monolayer is 2 nm, as opposed to the 150-nm thickness of the dextran chip. It has been possible to detect and characterize interactions on the NTA-SAM that were undetectable on the dextran chip because (1) processes under study are much closer to the emanation of the sensing wave, (2) the active sites are uniformly presented to analytes in solution, (3) noncovalent coupling to the support retains greater protein viability, and (4) flow over the monolayer is near laminar with an insignificant static layer.

The first demonstration of the enhanced sensitivity of the NTA-SAM chip involved the interaction between a mutant form of a yeast holoenzyme component, Gal11p, and an inert DNA-binding domain, Gal4(58 to 100). Although there was considerable genetic evidence that the two proteins were associated, direct detection of the interaction by standard techniques was not demonstrable for a number of

years. The interaction between these two proteins could not be detected by SPR on dextran-based chips, but was easily detected and quantitated using the NTA-SAM chip (Barberis et al., 1995). Subsequently, a fusion protein of Gal11p, which could be prepared at much higher concentrations, was observed to interact with Gal4(58 to 100) using both dextran SPR chips and gel mobility shift assays (Farrell et al., 1996). It is reassuring that the k_d measured on the NTA chip, where protein concentration was low, was virtually the same as that measured on the dextran chip using the fusion protein at a much higher concentration.

The second demonstration of the enhanced sensing capabilities of the NTA-SAM chip was in detecting the interaction of yeast polymerase II with TFIIF (Bushnell et al., 1996), which had not been detectable by SPR with standard dextran chips. A third example of the superiority of monolayer-based SPR chips was the detection of the interaction between TFIIE and polymerase II using the NTA-SAM (Bushnell et al., 1996). The interaction could not be detected using standard dextran chips, presumably because of the size of the protein complexes and related diffusion problems.

To measure the sensitivity of this method, side-by-side experiments were performed comparing dextran chips to NTA-SAMs; a chimeric T-cell receptor (TCR) was probed with two monoclonal antibodies (Sigal et al., 1996). One antibody recognized a particular conformation of the protein; the other recognized a linear epitope. Although six-times more protein was attached to the dextran surface (because of the increased surface area) than to the two-dimensional NTA array, only an effective monolayer of protein on the dextran was accessible to incoming ligand; thus, the total amount of antibody that bound to each surface was identical. The ratio of the conformationally sensitive antibody to the linear epitope-antibody that bound to each TCR-derivatized surface showed that five-times more protein remained in the active conformation when it was attached to the NTA-SAM than when attached to dextran.

In addition to increased sensitivity, NTA-SAMs offer increased experimental simplicity. Because there are a finite and approximately constant number of NTA ligands on every NTA-SAM, the determination of optimal immobilization conditions (e.g., protein pH or concentration) or regeneration conditions is not necessary. For example, maximal NTA occupancy on 5% NTA-SAMs is virtually constant

and is easily achieved by injecting 25 μ l of a 25-kD protein at a concentration of 3 mg/ml; higher injection concentrations will not alter the amount of protein presented. At 3% to 5% NTA, rebinding events are minimal so there is no need to experimentally determine the ligand concentration that will minimize rebinding in order to reflect a true off-rate. It is not necessary to check for artifacts stemming from an attraction to the chip surface by performing reciprocal experiments (reverse configuration) since the NTA-SAM has been shown to be highly resistant to nonspecific binding of all proteins tested including TBP.

An NTA SPR chip for the attachment of histidine-tagged proteins is now commercially available from BIAcore, but importantly the NTA ligands are attached to dextran (Gershon and Khilko, 1995), not incorporated into a monolayer. Noncovalent coupling of the protein of interest to the dextran support is an advantage because the chemical treatment used to promote covalent coupling denatures a significant portion of the protein. However, the chip still suffers from other problems intrinsic to dextran: high background charge, random orientation of the active site relative to analyte flow, and turbulent flow dynamics with a large static layer. Moreover, this chip may introduce a new experimental artifact, which was encountered when this approach was experimented with in 1994 (Bamdad and colleagues; unpub. observ.). The NTA group chelates a Ni(II) atom in a very-high-affinity interaction. However, the disordered nature of the dextran may position two NTA groups such that two ligands simultaneously coordinate one Ni(II) atom, which is much more unstable: subtle changes in pH and protein addition could trigger rearrangements that would act to crosslink the dextran and alter its optical response, thus mimicking protein binding.

Expanded applications with NTA-SAMs

It is likely that monolayer chips will find wide use. This is because biological systems often consist of networks of proteins that interact with each other. For example, it now appears that the active transcription complex may consist of some thirty proteins (Kim et al., 1994; Koleske and Young, 1994). Future research may be increasingly focused on the study of proteins interacting with *complexes* of proteins. The nonuniform surface of dextran can interfere with the entry of large protein complexes into the matrix, whereas monolayers provide near laminar flow dynamics that allow

for the characterization of interactions between large protein complexes. On dextran, the stoichiometry of interactions between a protein in solution and an evolving protein complex assembling on the chip may not be possible, as the interstices of dextran would render variable amounts of the complex accessible at each stage.

Another problem of great interest to biologists is the study of interactions where one or more of the partners is a membrane-associated protein. Plant et al. (1995) reported the development of a lipid bilayer chip for use in SPR instruments. Hydrophobic forces caused the spontaneous attachment of a phospholipid layer to a SAM comprised of hydrophobic alkyl thiolates. The work demonstrated the feasibility of using this approach to study interactions between components that are not water soluble: cell-surface receptor-ligand binding, protein-membrane interactions, and cell-cell associations. Unfortunately, this chip has not yet been made commercially available.

Many protein-protein interactions take place when one or both proteins are bound to DNA. The only method commercially available for the study of these interactions by SPR derivatizes dextran with streptavidin and then attaches it to biotinylated DNA. DNA-binding proteins can then be bound to the presented oligonucleotides for subsequent interaction studies with analytes in solution. These studies are subject to artifacts intrinsic to the dextran-based immobilization methods. A DNA-SAM was recently developed (Bamdad, 1997a) that presents single-stranded DNA above a background of inert alkane thiolates. Double-stranded DNA was hybridized to this surface via a complimentary single stranded "tail". The method was specific for the desired DNA, accepted unmodified bacterially produced DNA, and can easily be extended to present mass arrays of heterogeneous oligonucleotides. The sense strand of the hybridized double-stranded DNA was attached to the surface by an enzyme reaction leaving the antisense strand dissociable by heat treatment. Endonuclease digestion of hybridized DNA would generate a universal acceptor surface for ligation of sample DNA cut with the same enzyme.

Caveats and additional control experiments

Obtaining reliable results from SPR experiments requires careful attention to experimental design. To optimize the veracity of SPR-obtained results, experimental techniques must be designed to minimize systematic artifacts. The

BIAcore handbook is instructive for designing a preliminary set of SPR experiments, but the following additional experiments are well advised.

Once a set of experiments has been analyzed and kinetic rate constants obtained, the user should verify that the calculated rates are approximately correct using some simple experiments. For example, analyte injected at a concentration equal to the BIAcore-calculated k_d should give rise to approximately half-maximal binding, in RUs. Similarly, if the protein in solution (at a concentration capable of eliciting maximal binding) is incubated (competitive inhibition) with ligand at a concentration equal to the calculated k_d , and is then injected over immobilized ligand, approximately half-maximal binding should result. The kinetics of an interaction should be independent of the experimental configuration. To identify artifacts in the *apparent* kinetics of an interaction, a reciprocal experiment should be performed in which the protein in solution (the analyte) is immobilized as the ligand, and the previously immobilized protein becomes the analyte in solution. Performing the experiment in reverse configuration should result in comparable rate measurements. On dextran chips, discrepancies in reciprocal rates may be an indication of an interaction in which diffusion into the dextran, rather than association, is the rate-limiting step. Since diffusion into the dextran matrix is a function of the protein's size and charge, use the rate that is obtained when the smaller or less-charged of the two proteins is used as the analyte. Verification must be done using another method.

If a higher-order interaction is suspected, additional experiments should be designed to directly prove or disprove higher-order binding models, such as dimerization dependence (see Anticipated Results discussion on Kinetics analysis and higher-order kinetics).

It is also useful to repeat binding experiments using genetic variants of one or both of the interacting proteins, particularly variants whose *in vivo* activities are known. Reduced-activity mutants often display differential affinities for functionally relevant target molecules. A panel of mutants possessing a range of *in vivo* activities would provide for excellent control experiments.

Recall that SPR only detects the mass of a species near the interfacial region of the instrument, so the nonspecific binding of a few protein aggregates cannot be distinguished from multiple 1:1 binding events. One must there-

fore be rigorous about protein homogeneity. HPLC purification or centrifugation just prior to the experiment is recommended to rid the preparations of aggregates. Inaccuracies in measured binding constants can arise from discrepancies in the assessment of the *active* concentration of protein components. If the active concentration of one of the binding partners can be adequately assessed, that protein should be used as the analyte, since only the concentration of the analyte enters into the BIAcore kinetic equations. This minimizes problems associated with protein preparation heterogeneity.

Anticipated Results

Rate and equilibrium constants

The measured rate of dissociation should be a constant and independent of the concentration of either the surface density of the ligand or the analyte in solution. A dissociation rate that decreases as the ligand density increases indicates that dissociating molecules are rebinding to nearby sites. The ligand density should be optimized to produce a surface presenting enough molecules to elicit a clear binding response yet dilute enough to measure a dissociation rate constant that is constant. When extracting a dissociation rate from sensorgrams, choose time periods early in the dissociation phase to help eliminate rebinding events.

The measured rate of association for a particular reaction should also be constant. However, if the rate of diffusion into the disordered dextran polymer is slower than the rate of association with the immobilized ligand, then the measurable parameter will be the diffusion rate, not the association rate. Because the diffusion rate is a function of the size of the protein in solution, the flow rate, and the size of the surface irregularities, it may not be possible to measure the true association rate between large analyte proteins (or protein complexes) and ligands immobilized on dextran.

The accuracy of SPR measurements critically hinges on the stringency of the experimental design. For a ball-park estimate of the range of the equilibrium constant, determine the analyte concentrations that give rise to both maximal and minimal binding to a ligand surface. To a first approximation, the equilibrium constant should be the mean of the two concentrations. If an absolute kinetic rate is desired (i.e., one that is comparable to a solution measurement rather than a rank ordering of affinities), obtain an IC_{50} by performing a series of competitive inhibition experiments.

Kinetics analysis and higher-order kinetics

The BIAcore instrument comes with user-friendly kinetic evaluation software (modeled after the analysis described in O'Shannessy et al., 1993) that is now adequate to address simple first-order binding reactions. However, it may not be accurate for higher-order reactions. Others have developed more rigorous software for the kinetic analysis of BIAcore-derived data that uses numerical methods analysis and global fit programs (Fisher et al., 1994; Myszka, 1997). Fisher and Fivash (1994) have done considerable mathematical modeling of higher-order reactions that are then compared to experimental binding curves using a Cray supercomputer and their global fit analysis software. Less rigorous inferences of higher-order reactions based on curve-matching analysis of single experiments is risky. Additionally, over-analyzing data is less reliable than performing experiments designed to test a specific higher-order binding model.

Experimental confirmation of a theorized higher-order binding model was recently accomplished using variable-density NTA-SAMs

with SPR (Bamdad, 1997b). The concentration of NTA-terminated thiol was varied from 3% to 11% relative to an inert alkane thiol (the major component), producing a panel of SPR chips capable of presenting His-tagged peptides in variable-density arrays. It had previously been shown that four or more tandem repeats of an eight-amino acid motif could stimulate transcription *in vivo* and bind to the transcription factor TATA box-binding protein (TBP) *in vitro* (Tanaka, 1996). His-tagged peptides comprised of only two tandem repeats of the same motif were separately immobilized on NTA-SAMs of various density. BIAcore experiments showed that the peptides immobilized on low-density monolayers (3.8% NTA) were not able to bind TBP from solution. However, when the average distance between the peptides was decreased by immobilization on higher-density monolayers (5.7% NTA), a high-affinity interaction (20 nM) resulted (see Fig. 20.4.2). The complete study indicated that this observation resulted from the significant kinetic difference between monovalent and bivalent binding of the target protein, TBP. Com-

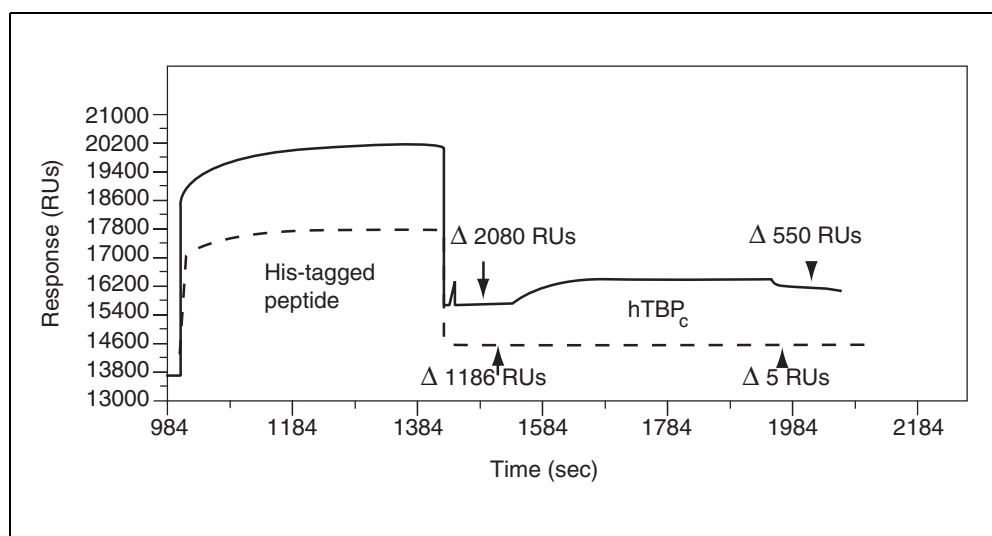


Figure 20.4.2 BIAcore sensorgram showing that binding of hTBP_c in solution to peptide motif surfaces is dependent upon density of ligand immobilization. The BIAcore SPR instrument plots changes in the angle of minimum reflectance in resonance units (RUs) as a function of time. The “square waves” represent injections of protein “plugs” that interrupt the constant buffer flow. The net change in measured RUs following a protein plug injection (arrows) indicates binding and can be correlated to a mass of protein recruited to the surface. An association constant can be derived from analysis of the initial phase of the injection, and a dissociation rate can be extracted from analysis of the system as it returns to buffer flow. Histidine-tagged fusion proteins terminated with the desired recognition motif were separately immobilized on NTA-SAMs presenting 3.8% (dashed line) or 5.7% NTA (solid line). hTBP_c in solution was injected over the surfaces in two separate experiments. An overlay of the two resultant SPR sensorgrams shows that hTBP_c cannot bind ($\Delta 5$ RUs; arrowhead) to the repeats when they are immobilized at low density (3.8% NTA; average distance = 29 Å), but bind very tightly ($\Delta 550$ RUs; arrowhead) when the peptides are positioned closer together at a higher density (5.7% NTA; average distance = 23 Å).

petitive inhibition experiments involving peptides either two or four repeats long demonstrated that there was a 250-fold difference in affinity between the two binding modes. The average distance between peptides on the chip surface, which presumably represents the distance between binding sites on TBP, was extrapolated from a statistical analysis of the lowest peptide density that elicited a large increase in target affinity. This analysis enabled the discrimination between three possible mechanistic models describing how reiterated repeats could synergistically effect the general transcription factor.

Despite the advantages presented by the ability to perform SPR experiments on a monolayer surface, it is still possible that experiments on these surfaces reflect interactions that differ significantly from those in solution due to local concentration effects and/or differences between two- and three-dimensional diffusion. A solution measurement, the IC_{50} , can be generated from a series of competitive inhibition experiments where the analyte is preincubated with varying concentrations of the ligand prior to injection of the mix over ligand-immobilized surfaces. In the above experiments, such an analysis revealed that the surface interaction was 20 times tighter than the comparable solution interaction.

Time Considerations

The major time factor in performing SPR experiments is in working out the appropriate conditions and experimental design, which can take weeks or even months. A single binding assay can be carried out in 5 to 10 min per injection. Using NTA-SAM chips, a set of experiments using three protein concentrations can be performed in 2 to 3 hr. However, to perform the same experiments with a dextran chip, it is likely to take several days to work out parameters such as the appropriate binding conditions and regeneration conditions.

Literature Cited

Bain, C.D., Evall, J., and Whitesides, G.M. 1989. Formation of monolayers by the coadsorption of thiols on gold: Variation in the head group, tail group and solvent. *J. Am. Chem. Soc.* 111:7155-7164.

Bamdad, C. 1997a. Submitted for publication.

Bamdad, C. 1997b. Submitted for publication.

Bamdad, C., Sigal, G., Strominger, J., and Whitesides, G. 1994. A mixed monolayer for the immobilization of histidine-tagged proteins. Appendix: A self-assembled monolayer for the

presentation of oligonucleotides. U.S. patent application PCL01016-94.

- Barberis, A., Pearlberg, J., Simkovich, N., Farrell, S., Reinagel, P., Bamdad, C., Sigal, G., and Ptashne, M. 1995. Contact with a component of the polymerase II holoenzyme suffices for gene activation. *Cell* 81:359-368.
- Bartel, P.L. and Fields, S. 1995. Analyzing protein-protein interactions using a two-hybrid system. *Methods Enzymol.* 254:241-263.
- Bushnell, D.A., Bamdad, C., and Kornberg, R. 1996. A minimal set of RNA polymerase II transcription protein interactions. *J. Biol. Chem.* 271:20170-20174.
- Daniels, P.B., Deacon, J.K., Eddowes, M.J., and Pedley, D.G. 1988. Surface plasmon resonance applied to immunosensing. *Sens. Actuators* 16:11-18.
- Farrell, S., Simkovich, N., Wu, Y., Barberis, A., and Ptashne, M. 1996. Gene activation by recruitment of the RNA polymerase II holoenzyme. *Genes & Dev.* 10:2359-2367.
- Fisher, R.J. and Fivash, M. 1994. Surface plasmon resonance based methods for measuring the kinetics and binding affinities of biomolecular interactions. *Curr. Opin. Biotechnol.* 5:389-395.
- Fisher, R.J., Fivash, M., Casas-Finet, J., Bladen, S., and Larson McNitt, K. 1994. Real-time BIAcore measurements of *Escherichia coli* single-stranded DNA binding (SSB) protein to polydeoxythymidylic acid reveal single-state kinetics with steric cooperativity. *Methods* 6:121-133.
- Gershon, P.D. and Khilko, S. 1995. Stable chelating linkage for reversible immobilization of oligohistidine tagged proteins in the BIAcore surface plasmon resonance detector. *J. Immunol. Methods* 183:65-76.
- Gyuris, J., Golemis, E., Chertkov, H., and Brent, R. 1993. Cdk2, a human G1 and S-phase protein phosphatase that associates with Cdk2. *Cell* 75:791-803.
- Johnsson, B., Löfås, S., and Lindquist, G. 1991. Immobilization of proteins to a carboxymethyl-dextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Anal. Biochem.* 198:268-277.
- Kim, Y.J., Bjorklund, S., Li, Y., Sayre, M.H., and Kornberg, R.D. 1994. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* 77:599-608.
- Koleske, A. and Young, R.A. 1994. An RNA polymerase II holoenzyme responsive to activators. *Nature* 368:466-469.
- Kretschmann, E. and Raether, H. 1968. *Z. Naturf.* 230:2135.
- Liedberg, B., Nylander, C., and Lundström, I. 1983. Surface plasmon resonance for gas detection and biosensing. *Sens. Actuators* 4:299-304.

- Löfås, S. and Johnsson, B. 1990. A novel hydrogel matrix on gold surfaces in surface plasmon resonance sensors for fast and efficient covalent immobilization of ligands. *J. Chem. Soc. Chem. Commun.* 1526-1528.
- Myszka, D. 1997. Kinetic analysis of macromolecular interactions using surface plasmon resonance biosensors. *Curr. Opin. Biotechnol.* 8:50-57.
- Nuzzo, R.G., Fusco, F.A., and Allara, D.L. 1987. Spontaneously organized molecular assemblies. 3. Preparation and properties of solution adsorbed monolayers of organic disulfides on gold surfaces. *J. Am. Chem. Soc.* 109:2358-2368.
- O'Shannessy, D., Brigham-Burke, M., Soneson, K., Hensley, P., and Brooks, I. 1993. Determination of rate and equilibrium binding constants for macromolecular interactions using surface plasmon resonance: Use of nonlinear least squares analysis methods. *Anal. Biochem.* 212:457-468.
- Otto, A. 1968a. *Z. Phys.* 216:398.
- Otto, A. 1968b. *Phys. Stat. Solidi.* 26:199.
- Plant, A., Brigham-Burke, M., Petrella, E., and O'Shannessy, D. 1995. Phospholipid/alkanethiol bilayers for cell-surface receptor studies by surface plasmon resonance. *Anal. Biochem.* 226:342-348.
- Sigal, G.B., Bamdad, C., Barberis, A., Strominger, J., and Whitesides, G.M. 1996. A self-assembled monolayer for the binding and study of histidine-tagged proteins by surface plasmon resonance. *Anal. Chem.* 68:490-497.
- Stenberg, E., Persson, B., and Roos, H. 1991. Quantitative determination of surface concentration of protein with surface plasmon resonance using radiolabeled proteins. *J. Colloid Interface Sci.* 143:513-526.
- Tanaka, M. 1996. Modulation of promoter occupancy by cooperative DNA binding and activation-function is a major determinant of transcriptional regulation by activators *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.* 93:4311-4315.
- Turbadar, T. 1959. *Proc. Phys. Soc. London* 73:40.
- Wahling, G., Raether, H., and Mobius, D. 1979. Studies of organic monolayers on thin silver films using the attenuated total reflection method. *Thin Solid Films* 58: 391-395.
- Zervos, A., Gyuris, J., and Brent, R. 1993. Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites. *Cell* 72:223-232.

Contributed by Cynthia Bamdad
Clinical Micro Sensors
Pasadena, CA

Detection of Protein-Protein Interactions by Coprecipitation

Coprecipitation of proteins from whole-cell extracts is a valuable approach to test for physical interactions between proteins of interest. When a precipitating antibody is used, this method is referred to as co-immunoprecipitation. Coprecipitation can be used to study interactions between known proteins under a variety of conditions and as a means of identifying components of a complex. Coprecipitation may be the single method of choice, or may be used in combination with other methods that detect protein-protein interactions, such as two-hybrid analysis and copurification schemes (UNIT 20.1), and tests of physical associations using purified proteins.

This unit describes basic approaches to immunoprecipitating tagged proteins from whole-cell extracts. The approaches described can be adapted for other systems. In a typical experiment, as described here, cells are lysed and a whole-cell extract is prepared under nondenaturing conditions (see Strategic Planning). The protein is precipitated from the lysate with a solid-phase affinity matrix, and the precipitate is tested for the presence of a second specifically associated protein (see Basic Protocol and Alternate Protocol). The approach can be used for native or epitope-tagged proteins for which antibodies are available, or for recombinant proteins that have been engineered to bind with high affinity to a molecule that can be coupled to a solid-phase matrix (see Strategic Planning). The presence of an associated protein is detected by separating the precipitated proteins by SDS-PAGE (UNIT 10.2A) and then immunoblotting (UNIT 10.8) with a second antibody that recognizes the putative associated protein. Controls to test specificity of interaction are crucial (see Strategic Planning).

For additional background reading, the user should consult UNIT 10.16 for a theoretical discussion of immunoprecipitation; see Chapter 11 for principles of antibody production and immunoassays; see UNITS 10.15, 20.2 & 20.3 for approaches to tagging proteins; and see Chapter 13 for transformation and propagation of *S. cerevisiae*. For an in-depth review of immunoprecipitation techniques, see Harlow and Lane (1988). For an in-depth review of coprecipitation and other approaches to detect protein-protein interactions, see Phizicky and Fields (1995), Toby and Golemis (2001), Auerbach et al. (2002), and Gavin and Superti-Ferga (2003).

STRATEGIC PLANNING

Detecting the Proteins in Question

The first step is to generate reagents that detect the two proteins in the coprecipitate under nondenaturing conditions. If antibodies are available that can immunoprecipitate the proteins under nondenaturing conditions, then they can be used. Alternatively, the proteins can be differentially tagged in a variety of ways to allow their detection with commercially available antibodies or other affinity reagents. The tagged proteins are then introduced into the host organism using expression vectors. All tagged proteins must be assessed for function in vivo.

A frequently used option is to add a short peptide or epitope that is recognized by a commercially available high-affinity monoclonal antibody (mAb; UNIT 10.15). The epitope is typically added at the amino or carboxyl terminus, although internal positions that do not disrupt function can also be used. Two frequently used epitopes are derived from influenza hemagglutinin protein (HA) and human c-Myc and are recognized by

high-affinity mAbs (12CA5 and 9E10, respectively; Kolodziej and Young, 1991). Others such as FLAG, the leader peptide of the gene 10 product of bacteriophage T7 (Knappik and Pluckthun, 1994; Witzgall et al., 1994), are also available. The choice of epitope may be dictated by its amino acid composition. It is often useful to insert tandem copies of the epitope in order to increase sensitivity. The number of additional tandem copies can range widely from one (Field et al., 1988) to several (e.g., three; Tyers et al., 1993) to many (e.g., nine; Feng et al., 1998).

Proteins can also be fused to small proteins or peptides that have high affinity to small molecules that can be attached to a solid support. This is a particularly valuable approach when the protein to be precipitated comigrates with immunoglobulin heavy or light chains in an SDS-polyacrylamide gel. Such alternative tagging methods include fusion to glutathione-*S*-transferase (to allow purification by a glutathione affinity matrix; UNIT 16.7), maltose-binding protein (to allow purification by a maltose affinity matrix; UNIT 16.6), polyhistidine (to allow purification by nickel affinity matrix; Bornhorst and Falke, 2000), or strep-tag (to allow purification by streptavidin matrix; Skerra and Schmidt, 2005). Tandem affinity protein (TAP) tags made of two IgG-binding domains of *Staphylococcus aureus* protein A and a calmodulin-binding peptide separated by a TEV protease cleavage site (Puig et al., 2001) have also been successfully used to copurify proteins, especially in yeast (Gavin and Superti-Furga, 2003). Other TAP tags have been described, including a TAP tag of three tandem FLAG epitopes and six histidines that may be more efficient in some organisms (Yang et al., 2006). A reference for identifying sources of commercially available antibodies and vectors encoding various tags can be found in the BioSupplyNet Source Book (2005). See Chapter 11 and Harlow and Lane (1998) for the generation and purification of specific antibodies. The American Type Culture Collection (ATCC) and European Collection of Animal Cell Cultures can be sources for hybridoma cell lines. Numerous genomic coprecipitation and two-hybrid studies have been done in a variety of organisms, so it is wise to check available databases for the current repertoire of interactions for a particular protein (e.g., Gavin and Superti-Furga, 2003, lists a number of such studies). Useful summaries of protein-protein interactions along with references can be found at Proteome Bioknowledge Library (Biobase Biological Databases, <http://www.biobase.de>; formerly Proteome, Inc., a subsidiary of Incyte), which is available by institutional or individual subscription.

Preparing Whole-Cell Extracts

The second step in a successful coprecipitation is generating whole-cell extracts that optimize the yield and activity of the proteins to be analyzed, using lysis buffer conditions that permit recognition of the proteins by the affinity matrix. The yield of total protein in a whole-cell extract is not always a reliable indicator of the relative yield and activity of specific proteins, so it is wise to verify both parameters at the onset of an experiment before proceeding with the coprecipitation. Yield and activity can be affected by a number of factors (see Chapter 10; see Harlow and Lane, 1988). Small variations in the relative amounts of salt and detergents in the lysis buffer can have large effects on yield and activity, as can the speed and efficiency of cell breakage. Both factors are particularly important for less soluble proteins that associate with macromolecular structures such as membranes or cytoskeleton. In addition, global inhibition of proteolysis through the inclusion of multiple classes of protease inhibitors may be essential.

Methods for preparing whole-cell extracts from yeast (UNIT 13.13), *E. coli* (UNITS 16.1-16.8), insect cells (UNIT 16.11), and mammalian cells (UNITS 16.12-16.18) can be found elsewhere in this manual, and specifics will not be discussed here. In general, the lysis buffer conditions are not very different from the coprecipitation conditions. It is recommended that the investigator begin by comparing small-scale extract preparations that vary the

amount of salt and nonionic detergent. As a starting point, a basic lysis buffer might contain the following components.

Basic components. Basic components include a buffering agent (such as 50 mM Tris-Cl, pH 7.5), a small amount of nonionic detergent (such as 0.1% [v/v] Triton X-100), salt (such as 100 mM NaCl), a reducing agent (such as 1 mM DTT), and 10% (v/v) glycerol as stabilizer.

Protease inhibitors. Protease inhibitor cocktails are described in *UNIT 13.13* and are also commercially available. A reasonable starting point would be to include 5 $\mu\text{g/ml}$ each chymostatin, pepstatin A, leupeptin, and antipain, as well as 1 mM phenylmethanesulfonyl fluoride and 1 mM benzamide.

Chelating agents. EGTA (~ 15 mM) is commonly included to chelate divalent metal ions that are essential for metalloprotease activity. Because EGTA also inhibits other metal-dependent enzymes, it may be omitted, combined with the addition of a needed metal ion, and/or substituted with EDTA.

Phosphatase inhibitors. If the phosphorylation state of the proteins in question is important, a mixture of phosphatase inhibitors should also be included in the lysis buffer. A starting mixture could contain 2.5 mM each meta- and ortho-vanadate, 10 mM NaF, and 10 mM β -glycerol phosphate.

Simple modifications of this initial buffer include varying the amount of NaCl (from 0 to 500 mM) and of Triton X-100 (from 0% to 1%). The investigator may choose to compare different means of breaking the cells (for example, glass-bead breakage versus liquid nitrogen/grinding methods for yeast cells; *UNIT 13.13*).

Total protein concentration in the whole-cell extract is generally assayed by using the Bio-Rad protein assay and calculating protein concentration (*UNIT 10.1A*). Extracts should be tested for the amount of each specific protein by immunoblot analysis (*UNIT 10.8*), analyzing 25 to 75 μg of total protein. In general, it is best to test for the presence of a second established protein (such as a housekeeping enzyme, cytoskeletal or ribosomal protein, or a previously defined component in the pathway being studied) as an internal control for normalization and as a positive control for the immunoblot. The amount of specific protein in the whole-cell extract is compared to the amount that is recovered by precipitation with an affinity matrix.

Control Tests for Specificity of Interaction

Controls are essential to verify that the antibodies and protein-protein interactions are specific. Proper controls are simplest to set up when the proteins are differentially tagged. In this instance, two parallel extracts are prepared from strains that contain each protein lacking the tag in the presence of the second tagged protein. An example is shown in the idealized gel in Figure 20.5.1, which includes lanes containing untagged protein 1 + tagged protein 2 and tagged protein 1 + untagged protein 2. If the antibodies are specific, untagged protein 1 will not immunoprecipitate. The presence of untagged protein 1 in the immunoprecipitate will indicate that it binds the affinity matrix nonspecifically. If the interaction between proteins 1 and 2 is specific, then tagged protein 2 will be present in the immunoprecipitate of tagged protein 1, but not in its absence.

If antibodies to native proteins are used, it is necessary to compare extracts made from strains harboring deletions of the proteins in question to test for the specificity of the antibody and the interaction. However, this is obviously possible only if the deletions do not cause inviability. If deletion mutations cannot be used, a common approach is to show that the preimmune serum or an antibody not known to be specific to either of

the proteins in question does not coprecipitate them in a parallel experiment. However, the latter two controls do not rule out the possibility that the antibody is precipitating the protein in question through an indirect association.

It is also essential to compare the amount of coprecipitated protein with the amounts of the two proteins in question in the whole-cell extract. This allows one to determine whether apparent differences in the ability of the two proteins to coprecipitate are a secondary consequence of the relative abundance of the proteins. This control is particularly important when an interaction has been established and the investigator wishes to search for regulatory changes in association apart from changes in abundance.

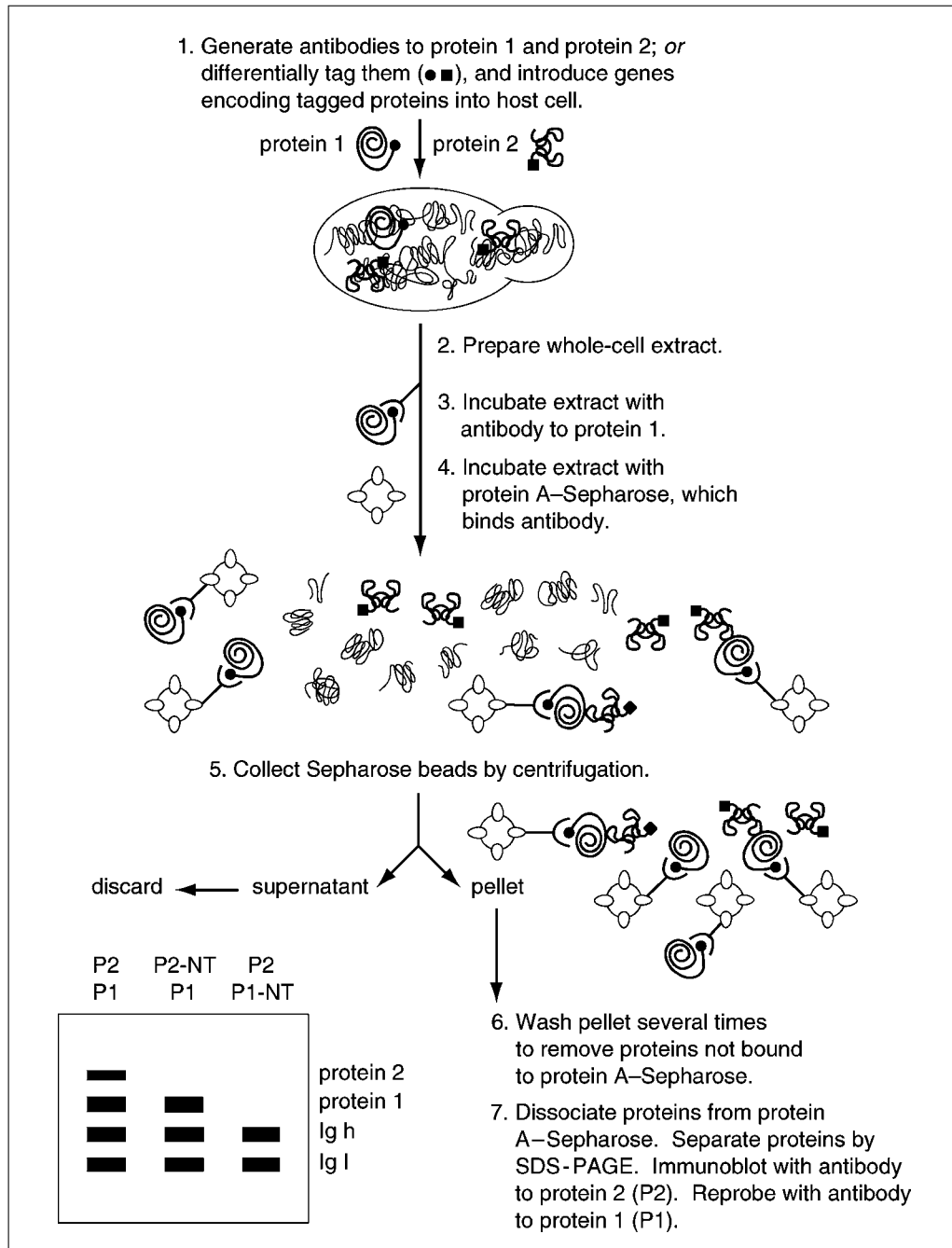


Figure 20.5.1 Flowchart for the coprecipitation of two proteins that have been differentially tagged and introduced into the host organism. Ig h and Ig l, immunoglobulin heavy and light chains; NT, no tag.

COPRECIPITATING PROTEINS WITH PROTEIN A/G–SEPHAROSE

Once the conditions of extract preparation have been established (see Strategic Planning), the next step is to test for coprecipitation of the specific proteins. This protocol describes a standard coprecipitation procedure that uses an antibody coupled to protein A–Sepharose or protein G–Sepharose. An alternative coprecipitation method that uses GST coupled to glutathione-agarose is also provided (see Alternate Protocol). It is essential to keep all buffers and tubes cold by using an ice bath and a refrigerated centrifuge. The conditions of coprecipitation match the conditions of the lysis buffer described above.

Materials

Whole-cell extract (see Strategic Planning)
Antibody against protein or epitope of interest
5 M NaCl
Co-immunoprecipitation buffer (see recipe)
Protein A/G–Sepharose slurry (see recipe)
2× sample buffer for SDS-PAGE (UNIT 10.2A)
Test tube rotator
20-ml syringe and 18-G needle
Hamilton syringe
Additional reagents and equipment for SDS-PAGE (UNIT 10.2A) and immunoblotting (UNIT 10.8)

1. Prepare duplicate samples in microcentrifuge tubes on ice:

0.5 to 1 mg whole-cell extract
1 μg antibody
5 M NaCl to equalize at 100 mM NaCl
Co-immunoprecipitation buffer to 0.5 ml final volume.

Adjust buffer by adding a divalent cation if necessary for activity of the protein in question (see Strategic Planning).

2. Invert tube gently several times and incubate on ice for 90 min with occasional tube inversion.

It is recommended that the investigator begin with a 90-min incubation. However, this incubation step can be shortened or lengthened.

3. Microcentrifuge 10 min at maximum speed, 4°C, to pellet nonspecific aggregates. Transfer supernatant to a new microcentrifuge tube.**4. Add 50 μl of protein A– or protein G–Sepharose slurry (25 to 30 μl bead volume). Be sure to evenly suspend the slurry before distributing it to the samples.**

Protein A has been used more frequently for historical reasons; however, protein G binds a broader range of Ig subtypes at higher efficiency. See UNIT 11.11 for a description of their relative binding capacities.

5. Rotate tube gently at 4°C for 30 to 60 min.

Rocking is much less efficient and should be avoided.

6. Gently pellet protein A/G–Sepharose by centrifuging 30 sec at 1000 rpm in a tabletop centrifuge, 4°C.

7. Wash pellet three times with 1 ml co-immunoprecipitation buffer. For each wash, gently invert tube three times before pelleting. After each pelleting, use a 20-ml syringe with an 18-G needle to aspirate and remove supernatant.

It may be possible to omit the costly protease inhibitors from the buffer at this stage.

8. Aspirate as much liquid as possible from the final without touching the beads and add 25 μ l of 2 \times sample buffer.

If desired, samples containing sample buffer can be frozen up to several months at -80°C prior to SDS-PAGE. In this case the buffer should be prepared with sterile stock solutions and made with 1 mM sodium azide included.

9. Prepare for SDS-PAGE analysis by boiling for 5 min, vortexing, and microcentrifuging briefly to pellet beads. Use a Hamilton syringe to load eluates onto an SDS-polyacrylamide gel, arranging duplicate samples to allow preparation of duplicate blots. Separate by electrophoresis (UNIT 10.2A).

A Hamilton syringe works well to remove the eluate from the beads during loading.

10. Immunoblot duplicate samples separately with antibodies for each of the two proteins (UNIT 10.8). Be sure to include aliquots of the whole-cell extract for comparison and as a positive control for the immunoblot.

Each immunoblot can be reprobed with the antibody to the other protein.

ALTERNATE PROTOCOL

COPRECIPTATING A GST FUSION PROTEIN

GST fusion proteins may be coprecipitated by following the co-immunoprecipitation procedure (see Basic Protocol) with the modifications outlined below. This procedure might be used when the protein in question comigrates with immunoglobulin heavy or light chain in an SDS-PAGE gel or if the antibodies being used precipitate too many cross-reacting proteins, such as the protein being tested for association. Furthermore, it is possible to dissociate the purified GST fusion from the solid-state glutathione resin under gentle conditions through the use of imidazole. The approach is also useful in that it will increase the size of the protein sufficiently that this size increase can be used as a diagnostic feature in analyzing complexes.

Additional Materials (also see Basic Protocol)

Glutathione-agarose or glutathione-Sepharose slurry (see recipe)

1. Prepare duplicate samples as described for protein A/G-Sepharose (see Basic Protocol, step 1), omitting antibody.
2. Microcentrifuge 10 min at maximum speed, 4°C , to pellet nonspecific aggregates. Transfer supernatant to new microcentrifuge tube.
3. Add 30 μ l glutathione-agarose or glutathione-Sepharose slurry (25- to 30- μ l bead volume). Be sure to evenly suspend the slurry before distributing it to the samples.
4. Rotate the sample, pellet and wash glutathione-agarose/Sepharose, and perform SDS-PAGE and immunoblot analysis (see Basic Protocol, steps 5 to 10).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Co-immunoprecipitation buffer

50 mM Tris·Cl, pH 7.5 (APPENDIX 2)

15 mM EGTA

100 mM NaCl

0.1% (w/v) Triton X-100

Store at 4°C

Immediately before use add:

1× protease inhibitor mix (see recipe)

1 mM dithiothreitol (DTT)

1 mM phenylmethylsulfonyl fluoride (PMSF; from fresh 250 mM solution in 95% ethanol)

If necessary for activity of the protein of interest, a divalent cation may need to be included in the buffer and EGTA omitted (as for extraction buffer; see Strategic Planning).

The protease inhibitor mix, PMSF, and DTT should be added fresh at the time of experimentation. The mixture without those components can be stored for months at 4°C with the addition of 1 mM sodium azide. PMSF is labile in aqueous buffer and should be added at the last minute. Protease inhibitor cocktails are also commercially available.

Since many protein-protein interactions are dependent upon phosphorylation, it may be necessary to include phosphatase inhibitors in the whole-cell extracts and the co-immunoprecipitation buffer phosphatase inhibitor solutions. A starting mix is 0.5 mM vanadate (0.25 mM each meta- and ortho-vanadate or 0.5 mM sodium vanadate, pH 7.4), 10 mM sodium fluoride (NaF), and 10 mM β-glycerol phosphate.

Glutathione-agarose or glutathione-Sepharose slurry

Swell 1.5 g glutathione-agarose or glutathione-Sepharose beads (e.g., Pierce, Sigma) in 30 ml of 50 mM Tris·Cl, pH 7.5 (APPENDIX 2), for 1 to 2 hr on ice. Pellet beads by gravity or very gentle centrifugation (1 min at 1000 rpm in a tabletop centrifuge) and then wash four times with co-immunoprecipitation buffer (see recipe) that lacks protease inhibitor mix and contains 1 mM sodium azide. Resuspend beads in 15 ml of this buffer to yield a final slurry concentration of ~100 mg/ml. Store at 4°C (stable for months).

This recipe can be scaled up or down.

Protease inhibitor mix, 1000×

Dissolve in DMSO:

5 mg/ml chymostatin

5 mg/ml pepstatin A

5 mg/ml leupeptin

5 mg/ml antipain

Store in aliquots up to 1 year at –20°C

Protein A/G–Sepharose slurry

Swell 1.5 g protein A– or protein G–Sepharose beads (e.g., Pierce, Sigma) in 30 ml of 50 mM Tris·Cl, pH 7.5 (APPENDIX 2), for 1 to 2 hr on ice. Pellet beads by gravity or very gentle centrifugation (1 min at 1000 rpm in a tabletop centrifuge) and then wash four times with co-immunoprecipitation buffer (see recipe) that lacks protease inhibitor mix and contains 1 mM sodium azide. Resuspend beads in 15 ml of this buffer to yield a final slurry concentration of ~100 mg/ml. Store at 4°C (stable for months).

This recipe can be scaled up or down.

COMMENTARY

Background Information

Coprecipitation is a powerful and simple approach to test for a physical interaction between proteins. There are many reasons to incorporate coprecipitation into a study. First, as a form of protein affinity chromatography, the method may be sensitive enough to detect weak associations that do not withstand the rigors of standard purification methods involving substantial dilution of the initial cell extract. Second, coprecipitation tests for associations between proteins within the milieu of a whole-cell extract, where the proteins are present at native concentration in a complex mixture of other cellular components. This feature makes it an important partner to two-hybrid methods and direct tests of interactions using purified proteins, because it provides a way to verify that a positive interaction reflects a true *in vivo* association. For example, nonphysiological interactions can be detected when purified proteins are present at too elevated a concentration. A falsely positive interaction between two proteins can also arise in a two-hybrid test when protein domains are inappropriately exposed due to altered folding. In addition, not all proteins are amenable to two-hybrid analysis; a negative result may mask a true association.

Nevertheless, a word of caution is in order. The ability to coprecipitate two proteins from a cellular extract is not proof that a particular interaction normally takes place *in vivo*. Additional experiments are needed to argue that a given interaction is not the result of mixing cell contents during extract preparation. Such evidence could include colocalization of the proteins or demonstration of functional relatedness.

When performing coprecipitation, it is important to precipitate from both directions (i.e., individually precipitating protein 1 and protein 2, and testing for the presence of protein 2 and protein 1, respectively). This is important in that it can provide further verification of an interaction between the proteins. It is also important because it is possible the interaction will only be detected in one direction. An inability to detect an interaction in one direction could be due to a variety of factors including obstruction of an interaction by the binding of the antibody or other affinity agent, or differences in pool size representation of each protein. For example, protein 1 may bind to many proteins besides protein 2, while most of protein 2 binds to protein 1. In this sce-

nario, it would be anticipated that detection of their association will be most efficient when protein 2 is precipitated.

Critical Parameters and Troubleshooting

Two important factors in determining how well an immunoprecipitation will work are the affinity of the antibody for the antigenic site on the protein (or the affinity of the protein tag for the affinity matrix) and the yield of the relevant proteins in the extract. For efficient immunoprecipitation, the antibody should have an affinity to the antigen in the range of 10^7 to 10^9 M^{-1} (Harlow and Lane, 1998). The easiest way to enhance detection of the immune complex is to increase the concentration of antibody and the protein being immunoprecipitated, as long as the concentration of antibody is not already saturating. This can be determined empirically. A starting point is 1 μ g antibody (Harlow and Lane, 1998), which is typically 0.5 to 1 μ l of a polyclonal antiserum, \sim 50 μ l of a hybridoma culture supernatant, or \sim 0.5 ml of ascites fluid. The reaction volume of the immunoprecipitation can be reduced to increase the concentration of the protein along with optimization of the extracts (see below). With protein tagging, it is possible to increase the number of copies of the moiety being bound by affinity matrix as a means of improving detection.

It is important to vary conditions of both the extract preparation and the coprecipitation to determine the optimal conditions. When starting from scratch, it is most prudent to use a range of lysis and precipitation conditions ranging from less to more stringent in terms of the amount of salt and non-ionic detergent to determine which conditions give the best yield of the desired proteins. When no interaction is detected, it is worthwhile to use less stringent conditions (reduced salt and nonionic detergent). In addition, it may be necessary to avoid any dilution of the whole-cell extract. This can be done by adding the protein A/G–Sepharose directly to the extract after an initial clarification centrifugation and by using smaller wash volumes. Depending on the strength and nature of the interaction, the precipitation can be performed in the presence of a mixture of detergents that includes ionic detergents (for example, 1% Triton X-100, 0.5% deoxycholate, 0.1% SDS).

In addition, it may be necessary to increase the expression levels of the proteins in question

in order to detect them readily by coprecipitation. A range of expression levels is recommended, because a level that is too high can lead to unregulated interactions (Feng et al., 1998). One can scale up the coprecipitation to at least 2 mg of whole-cell extract in the described assay (Wang et al., 2005); a starting point is 0.25 to 1 mg. Here, the limiting factor is the concentration of the extracts, which must be high enough to allow the volume of the coprecipitation mixture to remain low. Larger-scale extract preparations may be necessary to generate more concentrated extracts. Finally, in cases of failure due to low abundance of the proteins in the host organism, one can overexpress a tagged version of one of the two proteins in the same or another host (such as *E. coli*), concentrate this protein by pre-immobilization on an appropriate affinity matrix, and then incubate the affixed protein with extracts from the host organism.

The most important objective in these experiments is to generate as great a signal-to-noise ratio as possible and avoid problems of background. A variety of parameters can be changed to enhance the co-immunoprecipitation. Optimization of the precipitating antibody is one possibility. Protein A–Sepharose and protein G–Sepharose should give results comparable to anti-Ig serum. However, direct coupling of the antibody to Sepharose may lead to reduced background and more quantitative precipitation. In addition, varying the ratio of antibody to whole-cell extract and the total amount of whole-cell extract is strongly suggested to determine the optimal amount of antibody that gives the most precipitation with the least amount of background. Affinity purification of the antibody may be necessary if the antibody immunoprecipitates additional cross-reacting proteins.

Additional approaches can be taken to minimize background. First, better clarification of the cell extract can be achieved by precentrifugation at $100,000 \times g$. These extracts can be directly used for coprecipitation without an intervening freezing step, which can increase the amount of protein precipitation. Second, both the lysis buffer and the coprecipitation buffer can be supplemented with 1% BSA to reduce nonspecific binding to the affinity matrix. Third, the whole-cell extract can be preincubated with protein A/G–Sepharose to remove nonspecific proteins that bind to the solid support. Fourth, the amount of salt and detergent can be increased in both the coprecipitation and the washes to reduce nonspecific binding. Fifth, increasing the number of

washes may also help, although it may reduce the amount of specific protein that remains associated. Sixth, the expression levels of the proteins in question can be increased to generate a stronger signal that is above background binding. Alternatively, it may be possible to produce a whole-cell extract that is enriched for the proteins in question (e.g., by preparing a nuclear extract if the proteins are known to be in the nucleus). In instances where one of the proteins binds nonspecifically to Sepharose, the substitution of an agarose-based affinity matrix may help solve the problem. Finally, it may be necessary to generate a different set of reagents to precipitate the proteins in question (i.e., different antibodies and/or protein tag).

Anticipated Results

Provided suitable antibodies are available to the proteins in question and the physical interaction is stable under the coprecipitation conditions, it should be possible to detect an interaction between two proteins. This method can be scaled up to coprecipitate sufficient protein to allow detection in a polyacrylamide gel either by silver stain (~ 1 to 10 ng/band) or Coomassie blue stain (~ 0.1 to 1 mg/band), with subsequent excision of the band from the gel and identification by mass spectrometry.

Time Considerations

Once the extracts are prepared, coprecipitation can be done within 3 to 4 hr, yielding samples ready to load on a gel for SDS-PAGE and immunoblot analysis. It is possible to freeze the prepared extracts and thaw them on ice prior to setting up the coprecipitation, but higher yields may be obtained with samples that have not been frozen.

Literature Cited

- Auerbach, D., Galeuchet-Schenk, B., Hottiger, M.O., and Staglier, I. 2002. Genetic approaches to the identification of interactions between membrane proteins in yeast. *J. Recept. Signal Transduct. Res.* 22:471-481.
- BioSupplyNet Source Book. 2005. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Bornhorst, J.A. and Falke, J.J. 2000. Purification of proteins using polyhistidine affinity tags. *Methods Enzymol.* 326:245-254.
- Feng, Y., Song, L.-Y., Kincaid, E., Mahanty, S.K., and Elion, E.A. 1998. Functional binding between G β and the LIM domain of Step is required to activate the MEKK Ste11. *Curr. Biol.* 8:267-278.
- Field, J., Nikawa, J., Broek, D., MacDonald, B., Rodgers, L., Wilson, I.A., Lerner, R.A., and Wigler, M. 1988. Purification of RAS-responsive adenyl cyclase complex from

- Saccharomyces cerevisiae* by use of an epitope addition method. *Mol. Cell. Biol.* 8:2159-2165.
- Gavin, A. and Superti-Furga, G. 2003. Protein complexes and proteome organization from yeast to man. *Curr. Opin. Chem. Biol.* 7:21-27.
- Harlow, E. and Lane, D. 1988. *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Harlow, E. and Lane, D. 1998. *Using Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Knappik, A. and Pluckthun, A. 1994. An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments. *Biotechniques* 17:754-761.
- Kolodziej, P.A. and Young, R.A. 1991. Epitope tagging and protein surveillance. *Methods Enzymol.* 194:508-519.
- Phizicky, E.M. and Fields, S. 1995. Protein-protein interactions: Methods for detection and analysis. *Microbiol. Rev.* 59:94-123.
- Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Seraphin, B. 2001. The tandem affinity purification (TAP) method: A general procedure of protein complex purification. *Methods* 24:218-229.
- Skerra, A. and Schmidt, T.G.M. 2005. Use of the Strep-tag and streptavidin for detection and purification of recombinant proteins. *Methods Enzymol.* 326:271-304.
- Toby, G.G. and Golemis, E.A. 2001. Using the yeast interaction trap and other two-hybrid-based approaches to study protein-protein interactions. *Methods* 24:201-217.
- Tyers, M., Tokiwa, G., and Futcher, A.B. 1993. Comparison of the *Saccharomyces cerevisiae* G1 cyclins: Cln3 may be an upstream activator of Cln1, Cln2, and other cyclins. *EMBO J.* 11:1773-1784.
- Wang, Y., Chen, W., Simpson, D.M., and Elion, E.A. 2005. Cdc24 regulates nuclear shuttling and recruitment of the Ste5 scaffold to a heterotrimeric G protein in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 280:13084-13096.
- Witzgall, R., O'Leary, E., and Bonventre, J.V. 1994. A mammalian expression vector for the expression of GAL4 fusion proteins with an epitope tag and histidine tail. *Anal. Biochem.* 2:291-298.
- Yang, P., Sampson, H.M., and Krause, H.M. 2006. A modified tandem affinity purification strategy identifies cofactors of the *Drosophila* nuclear receptor dHNF4. *Proteomics* 6:927-935.

Key References

- BioSupplyNet Source Book, 2005. See above.
Published yearly. Instant access is available at <http://www.biosupplynet.com>.
- Harlow and Lane, 1998. See above.
General discussion of immunoprecipitation techniques and reagents.
- Phizicky and Fields 1995. See above.
General discussion of methodologies for detecting protein-protein interactions as well as their merits and drawbacks.

Contributed by Elaine A. Elion
 Harvard Medical School
 Boston, Massachusetts

Identification of Protein Interactions by Far Western Analysis

UNIT 20.6

This unit describes far western blotting, a method of identifying protein-protein interactions. In a far western blot, one protein of interest is immobilized on a solid support membrane, then probed with a non-antibody protein. Far western blots can be used to identify specific interacting proteins in a complex mixture of proteins (see Basic Protocol). They are particularly useful for examining interactions between proteins that are difficult to analyze by other methods due to solubility problems or because they are difficult to express in cells. This method is performed totally *in vitro*, and the proteins of interest can be prepared in a variety of ways. Peptides can be used to determine the effects of specific residues or post-translational modifications on protein-protein interactions (see Alternate Protocol 2). In addition, many different detection techniques, either radioactive or nonradioactive, can be used. For example, the protein probe may be detected indirectly with an antibody, rather than being labeled radioactively (see Alternate Protocol 1). Thus, techniques and reagents already in hand can frequently be adapted for use with this assay.

CAUTION: Appropriate safety precautions must be taken when working with radioactive materials. Information on proper handling and disposal of radioactive compounds can be found in *APPENDIX 3A* and may be obtained from local radiation safety officials. Specific information on handling ^{35}S -labeled compounds can be found in *UNIT 10.18*.

FAR WESTERN ANALYSIS OF A PROTEIN MIXTURE

The following is a basic method for detecting protein-protein interactions by far western blotting when one protein is contained within a simple or complex mixture of proteins. First, the protein sample is fractionated on an SDS-PAGE gel (*UNIT 10.2A*). After electrophoresis, the proteins are transferred from the gels onto a solid support membrane by electroblotting (*UNIT 10.8*). Transferred membranes may be stained with Ponceau S to facilitate location and identification of specific proteins. Nonspecific sites on the membranes are blocked with standard blocking reagents, and the membranes are then incubated with a radiolabeled non-antibody protein probe. After washing, proteins that bind to the probe are detected by autoradiography (*APPENDIX 3A*).

Materials

- Samples to be analyzed
- 1× SDS sample buffer (*UNIT 10.2A*)
- Ponceau S staining solution (see recipe)
- Blocking buffer I: 0.05% (w/v) Tween 20 in 1× PBS (see recipe for PBS); prepare fresh
- Blocking buffer II: dissolve 1 g bovine serum albumin (BSA; fraction V) in 100 ml 1× PBS (see recipe for PBS); prepare fresh
- Phosphate-buffered saline (PBS; see recipe), pH 7.9
- cDNA encoding protein of interest cloned into an *in vitro* expression vector
- In vitro* transcription/translation kit (Promega)
- 10 mCi/ml ^{35}S -methionine (1000 Ci/mmol)
- Probe purification buffer (see recipe)
- Probe dilution buffer (see recipe)

**BASIC
PROTOCOL**

**Analysis of
Protein
Interactions**

20.6.1

Polyvinylidene difluoride (PVDF) or nitrocellulose membrane for protein transfer
Microfiltration centrifuge columns (e.g., Gelman Nanosep, Pall Filtron, or
Millipore Microcon)

Additional reagents and equipment for SDS-PAGE (UNIT 10.2A), electrophoretic
transfer of proteins to a support membrane (UNIT 10.8), in vitro translation (UNIT
10.17), and autoradiography (APPENDIX 3A)

NOTE: Always handle support membranes with gloves or membrane forceps.

Prepare protein blot

1. Prepare the protein sample to be analyzed by resuspending it in 1× SDS sample buffer.

UNIT 10.2A gives instructions on preparation of samples and amount of samples to load. In general, ~50 to 100 μg can be loaded in each lane for a complex mixture of proteins. A smaller amount, i.e., 10 to 20 μg, is loaded for less complex protein samples. The amount loaded may also need to be adjusted for the size of gel. (Usually 30 μg/mm² loading surface can be resolved without smearing.)

2. Separate the samples on an SDS-polyacrylamide gel (UNIT 10.2A).
3. Transfer the proteins from the gel to a solid support membrane (e.g., PVDF or nitrocellulose) by semidry electroblotting (UNIT 10.8).

Either nitrocellulose or PVDF membranes can be used with good results. PVDF membranes are easier to handle and tend to give a slightly higher signal-to-noise ratio, probably due to increased protein retention by the membrane.

Stain with Ponceau S

4. After transfer, stain the membrane for 5 min in ~100 ml freshly diluted 1× Ponceau S staining solution. Stain the membrane in a plastic container large enough to hold the blot and use sufficient Ponceau S to cover the membrane completely.

This step is optional. When the protein samples consist of a few proteins, or when there are clearly visible bands that facilitate orientation of the blot, staining with Ponceau S can provide helpful landmarks. One can unequivocally identify interacting bands, mark the position of molecular weight standards, and trim away excess membrane more exactly.

5. Destain the membrane washing in several changes of deionized water until the proteins are clearly visible. Place light pencil marks adjacent to important protein bands to mark them for future reference. Trim away excess membrane.

The stain fades quickly so the marks must be placed immediately.

6. Destain an additional 5 min in water until the red staining fades.

Block membrane

7. Block blot for 2 hr in 200 ml blocking buffer I at room temperature with gentle agitation.
8. Decant and add 200 ml blocking buffer II. Incubate as in step 7.
9. Decant blocking buffer II and rinse the membrane briefly in 100 ml of 1× PBS.

At this point, the blot may be probed immediately or may be wrapped in plastic wrap and stored for up to 2 weeks at 4°C.

Prepare the probe

10. Following manufacturer's procedures, prepare a radiolabeled in vitro–translated probe of the protein of interest using ^{35}S methionine (also see UNIT 10.18).

The probe can be conveniently prepared during the blocking steps.

The authors routinely use the Promega TnT quick-coupled transcription/translation system for producing probes. For a small blot (e.g., $\leq 9 \times 9\text{-cm}$), one half of a standard in vitro transcription/translation reaction (i.e., 25 μl) is sufficient for the probe. For larger blots, an entire 50- μl reaction may be used. In the authors' laboratory it is considered essential that the transcription/translation lysate not be repeatedly frozen and thawed.

11. After translation, dilute the probe with 500 μl probe purification buffer, and purify by microcentrifuging 15 to 30 min at $10,000 \times g$, room temperature, in a microfiltration column. Save aliquots of the purified probe for analysis by SDS-PAGE (i.e., 2 μl), and for scintillation counting (i.e., 2 to 5 μl).

Check microfiltration column manufacturer's procedure for exact centrifugation times required by different columns. In practice, it is not always necessary to purify the probe through a microfiltration column; many probes give good signals without purification. However, if signal-to-noise ratio is low, probe purification may improve results. In addition, it is possible to quantitate the proportion of probe bound if the probe has been purified.

Bind probe

12. Preincubate blot for 10 min in 50 ml of $1\times$ probe dilution buffer (without probe) by gently agitating at room temperature.
13. Dilute the translated probe with $1\times$ probe dilution buffer in a volume sufficient to cover the membrane to be probed (typically 3 ml).

For small blots, i.e., $\leq 9 \times 9\text{-cm}$, a 50-ml conical tube makes a convenient incubation chamber. A volume of 3 ml is enough solution to cover the blot and tubes can be rotated on a mechanical rotator. In addition, the conical tube makes the radiolabeled probe easy to contain and dispose of. Larger blots can be incubated on a Nutator or orbital shaker in a heat-sealable bag, or rotated in a hybridization oven adjusted to room temperature.

14. Add the probe to the membrane and incubate 2 hr at room temperature. Rotate the tubes or agitate bags throughout the binding reaction.

Wash the membrane

15. Transfer the membrane to a plastic dish and wash the membrane with 200 ml $1\times$ PBS for 5 min, room temperature. Repeat for a total of four washes.

Background is generally quite low and extended washing does not substantially reduce background.

16. Air dry the membrane and expose to X-ray film (autoradiography) or phosphor imager screen (see APPENDIX 3A for both techniques).

Do not cover the blots with plastic wrap as this will quench the ^{35}S signal. Overnight exposure to X-ray film is usually sufficient to detect positive interactions.

DETECTING INTERACTING PROTEINS BY IMMUNOBLOTTING

In vitro–translated probes have the advantages of being quickly produced, easily detected, and quantitated to give an estimate of relative binding. In addition, mutations in the protein probe can be generated by simple cloning procedures and can provide information on binding domains and their critical residues. A disadvantage of in vitro–translated probes is the need for multiple methionine or cysteine residues to obtain a well labeled probe. For the same reasons, small peptide fragments are often not suitable for use as in

ALTERNATE PROTOCOL 1

Analysis of Protein Interactions

20.6.3

vitro–translated probes. [¹⁴C]leucine and [³H]leucine can also be used for in vitro translation of proteins; however, in the authors' laboratory [¹⁴C]leucine has not yielded probes suitable for use as far western probes.

There are many other ways to generate probes for far western blots. The protein probe may be labeled in vitro with ¹²⁵I (Schumacher and Tsomides, 1995) or enzymatically with ³²P (Kimball, 1998). Biotin-labeled probes may be detected with streptavidin-biotin detection schemes (Luna, 1996; Grulich-Henn et al., 1998; Kimball et al., 1998). Protein binding may be detected indirectly as well. If an antibody to the interacting protein is available, then an unlabeled protein probe can be bound to the blots as usual and then detected by western (immunoblot) analysis. This is especially useful when a tagged recombinant protein and antibody to the tag are available. The following procedure describes detection of an unlabeled protein probe with specific antibody. Many variations of immunoblotting exist; additional information and procedures may be found in *UNIT 10.8*.

Additional Materials (also see *Basic Protocol*)

- Recombinant protein or unlabeled in vitro translated–protein for probe
- 5% (w/v) non-fat instant dry milk in 1× TBST (see recipe for TBST)
- Primary antibody specific for protein probe
- TBST (see recipe)
- Alkaline phosphatase (AP)–conjugated secondary antibody against Ig of species from which specific antibody was obtained
- Alkaline phosphatase buffer (see recipe)
- Developing solution (see recipe)
- 100 mM EDTA, pH 8.0 (*APPENDIX 2*)

Prepare blot and probe

1. Prepare and block the blot (see *Basic Protocol*, steps 1 to 9).
2. Prepare the probe protein by diluting in vitro–translated or recombinant protein in 3 ml of 1× probe dilution buffer.

The amount of recombinant protein must be empirically determined for each protein. Various researchers have used from 0.5 to 20 μg recombinant protein/ml of probe dilution buffer.

Expose blot to probe

3. Bind the probe to blot and wash (see *Basic Protocol*, steps 12 and 14 to 15). Do not dry the membrane after washing.
4. Incubate blot in 200 ml of 5% non-fat milk in 1× TBST for 1 hr, room temperature, with gentle rotation on an orbital shaker.

Expose to antibodies

5. Dilute the primary antibody in 5% milk in 1× TBST. Incubate blot in 5 to 10 ml diluted antibody at room temperature with gentle agitation to ensure blot is evenly covered with the antibody solution.

Incubations are usually carried out in heat-sealed plastic bags or hybridization bottles to minimize the volume necessary to completely cover the blot. A volume of 5 to 10 ml of diluted antibody is sufficient to cover most blots. Appropriate antibody concentrations vary for each antibody and must be determined empirically.

6. Wash for 10 min in ≥200 ml 1× TBST by agitating on an orbital shaker. Repeat an additional two times.

7. Dilute the AP-conjugated secondary antibody in 5 to 10 ml of 5% milk in 1× TBST and incubate blot for 1 hr as in step 4.

Suppliers generally provide an estimate of appropriate dilution for the secondary antibody.

8. Wash blot six times for 5 min each in ≥200 ml TBST, with agitation.

Detect antibodies

9. Briefly, rinse blot in 50 ml alkaline phosphatase buffer.
10. Incubate blot in 20 ml developing solution for 1 to 15 min and rinse blot with 100 ml water.
11. Wash blot for 5 min with 100 ml of 100 mM EDTA, pH 8.0, to stop the development reaction. Rinse with 100 ml water, dry, and photograph.

*An example of a far western blot of proteins is shown in Figure 20.6.1. Lane 1 shows a Coomassie blue stain of a protein sample enriched for histone proteins separated on a 22% SDS-PAGE gel. Lane 2 shows a far western autoradiogram of a parallel lane probed with the yeast *Tup1* protein according to this protocol. Lane 3 shows a parallel lane probed with an unlabeled *Tup1* protein and detected with antibody specific to *Tup1* as in Alternate Protocol 2. Both protocols yield the same result—*Tup1p* interacts with H3 and H4 but not with H2A or H2B. Lanes 4 and 5 show immunoblots of parallel lanes using anti-H3 and anti-H4 antibodies to identify histones H3 and H4 unequivocally.*

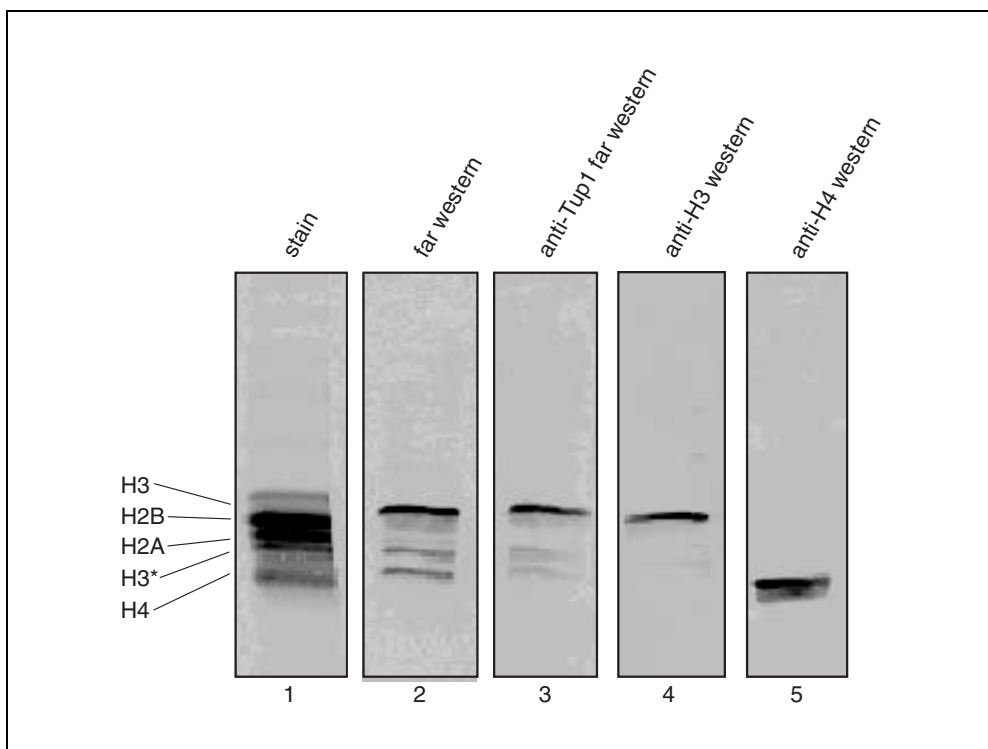


Figure 20.6.1 Far western of blotted SDS-PAGE gel. Lane 1, Coomassie blue-stained gel showing locations of histone bands. Lane 2, far western of a parallel lane using radiolabeled *in vitro*-translated probe. Lane 3, far western using unlabeled probe detected with probe-specific antibody. Lane 4, western blot using anti-histone H3 specific antibody. Lane 5, western blot using anti-histone H4 specific antibody.

**USING PEPTIDES TO IDENTIFY SPECIFIC INTERACTING SEQUENCES
IN A FAR WESTERN BLOT**

Synthetic peptides can also be used in far western analyses. The use of peptides enables the identification of specific interacting sequences. Specific post-translational modifications can be examined for their effect on protein-protein interactions. Peptides as small as 11 amino acids have been used successfully as far western targets.

Peptide far westerns differ from other far westerns only in the preparation of the blots. Peptide dilutions are prepared, then dot or slot blotted onto the support membrane. Blocking and probing of peptide blots are identical to procedures used for traditional far westerns. Duplicate blots are stained to verify that comparable quantities of different peptides have been loaded. Because Ponceau S staining is temporary, staining of duplicate blots with India ink is used to provide a permanent record for peptide blots.

In the authors' experience, only peptides that have been synthesized on MAP resins have worked well for peptide blots. MAP resins consist of branched lysine chains whose chemically active groups have been blocked. Although peptides prepared in other ways do give reproducible results, the peptide concentrations required are several orders of magnitude higher than those required for MAP peptides, making these blots very costly to perform. The reason why increased peptide is needed is unclear, but perhaps the MAP resin "presents" the peptide in such a way that it is more accessible for interaction.

Additional Materials (also see *Basic Protocol*)

Peptides

0.4% Tween 20/PBS (see recipe)

India ink solution (see recipe)

Slot or dot blot apparatus (e.g., Bio-Rad Bio-Dot SF or Schleicher & Schuell Minifold II)

1. Make dilutions of peptides between 5 ng and 5 μ g in a final volume of 100 to 200 μ l of distilled water.
2. Prepare slot or dot blotter and support membrane (PVDF or nitrocellulose) as described by the manufacturer. Load the peptide dilutions into wells. Prepare duplicate blots, one for far western and one for India ink staining. After all samples are loaded, apply vacuum to draw the peptide samples through the manifold device and onto the support membrane.
3. Block, bind, wash, and autoradiograph one blot for far western (see Basic Protocol, steps 7 to 16).
4. Incubate the second blot in 100 ml of 0.4% Tween 20/PBS for 5 min at room temperature with gentle agitation. Repeat incubation.
5. Stain blot by incubating 15 min to overnight with 100 ml India ink solution at room temperature.
6. Wash the filter for 2 hr in 4 changes of 1 \times PBS. Dry and store the membrane.

This stain is permanent.

Figure 20.6.2 shows an example of results from this alternate protocol, using peptides as a substrate for a far western. The right-hand panel is a blot stained with India ink verifying that comparable quantities of peptide were loaded on the blot. The left-hand panel is a far western of a duplicate blot demonstrating the effect of acetylation of lysine residues of histone peptides on Tup1p/histone interaction.

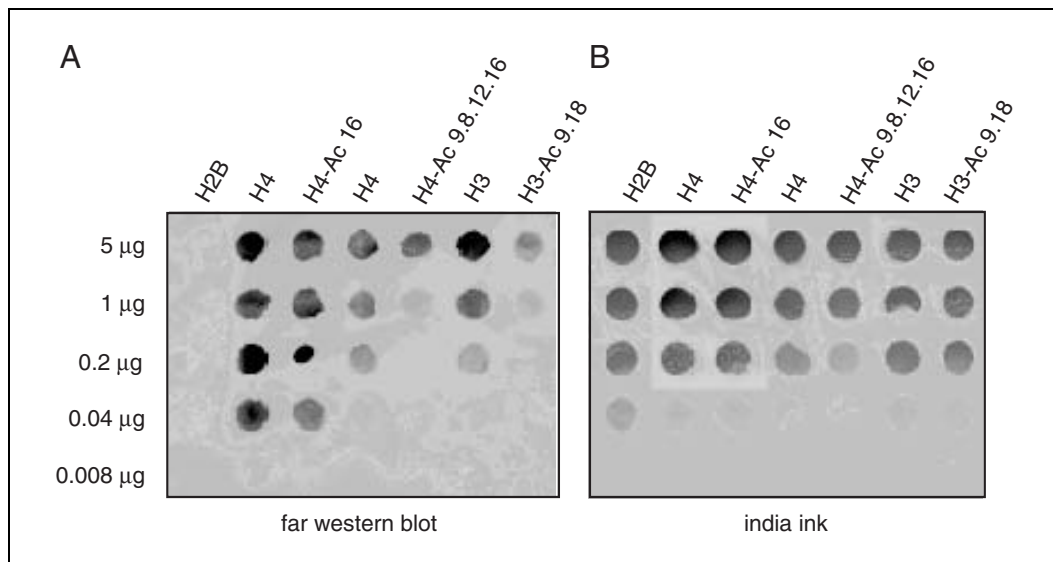


Figure 20.6.2 Far western blot of peptides (**A**) dot blotted onto PVDF membrane and (**B**) stained with India ink. Figure reproduced with permission of Cold Spring Harbor Laboratory Press.

REAGENTS AND SOLUTIONS

Use deionized or distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Alkaline phosphatase buffer

- 100 ml 1 M Tris-Cl, pH 9.5 (APPENDIX 2)
- 20 ml 5 M NaCl (APPENDIX 2)
- 5 ml 1 M MgCl₂ (APPENDIX 2)
- Add H₂O to 1 liter
- Store up to 1 year at room temperature

5-Bromo-4-chloro-3-indolyl phosphate (BCIP) stock solution

Dissolve 0.5 g of BCIP in 10 ml of 100% dimethylformamide. Store at 4°C or in small aliquots at -20°C. Discard when solution turns color.

Developing solution

Add 66 µl of NBT stock (see recipe) to 10 ml of alkaline phosphatase buffer (see recipe). Mix well. Add 33 µl of BCIP stock solution (see recipe) and mix again. Prepare fresh.

India ink solution

Add 100 µl India ink (Pelikan or Higgins) to 100 ml 0.4% Tween 20/PBS (see recipe). Prepare fresh.

Nitroblue tetrazolium chloride (NBT) stock solution

Dissolve 0.5 g of NBT in 10 ml of 70% dimethylformamide. Store at 4°C or in small aliquots up to 6 months at -20°C.

Phosphate-buffered saline (PBS), 10×

- 80 g NaCl
- 2.2 g KCl
- 9.9 g Na₂HPO₄
- 2.0 g K₂HPO₄
- Add H₂O to 1 liter

continued

Adjust pH to 7.4
Store indefinitely at room temperature
Prior to use, dilute to 1× by mixing 1 part 10× PBS with 9 parts water
Leftover 1× PBS should be stored at 4°C to discourage bacterial growth.

Ponceau S staining solution, 10×

2 g Ponceau S
30 g trichloroacetic acid
30 g sulfosalicylic acid
Add H₂O to 100 ml
Store indefinitely at room temperature
Just prior to use dilute to 1× by mixing 1 part 10× Ponceau S with 9 parts water

Probe dilution buffer, 10×

3.0 g bovine serum albumin (BSA)
10 ml normal goat serum
10 ml 10× PBS (see recipe)
H₂O to 100 ml
Store indefinitely at -20°C
Just prior to use, dilute to 1× by mixing 1 part 10× stock with 9 parts 1× PBS

Probe purification buffer

400 μl 1 M HEPES, pH 7.4,
400 μl 1 M dithiothreitol (DTT)
9.2 ml H₂O
Prepare fresh

TBST, 10×

90 g NaCl
100 ml 1 M Tris·Cl, pH 7.5 (*APPENDIX 2*)
10 g Tween 20
Add H₂O to 1 liter
For 1× TBST dilute 1 part 10× TBST with 9 parts water prior to use
Store indefinitely at room temperature

Tween 20/PBS, 0.4% (w/v)

Dissolve 0.4 g Tween 20 in 100 ml 1× PBS (see recipe). Store up to 1 week at room temperature.

COMMENTARY

Background Information

The far western blot (also called a west western and a ligand blot) has been widely used to examine the interactions of many diverse proteins. For example, it has been used to examine the interactions between the subunits of eukaryotic initiation factors (Kimball et al., 1998), to look at interactions between basic helix-loop-helix DNA-binding proteins (Chaudhary et al., 1997), and to examine the interactions of keratin intermediate filaments with desmosomal proteins (Kouklis et al., 1994). Far westerns have been particularly useful in examining interactions of histones with regulatory proteins. Far westerns have been

used to look at interactions of WD repeat proteins with histones (Edmondson et al., 1996; Palaparti et al., 1997), the interaction of Epstein-Barr virus nuclear antigen 2 with histone H1 (Grasser et al., 1993), and the interaction of histones with the *Xenopus* oocyte protein N1 (Kleinschmidt and Seiter, 1988). In addition, far westerns have been used to study receptor-ligand interactions and to screen libraries for interacting proteins (Grulich-Henn et al., 1998; Hsiao and Chang, 1999).

Sometimes, the nature of the proteins being examined is such that standard methods of studying protein-protein interactions are not possible. For example, some proteins are diffi-

cult to solubilize or to extract from cells except under conditions that disrupt protein-protein interactions and therefore, are difficult to assay by immunoprecipitation. Other proteins cannot be expressed in bacteria or yeast due to toxicity problems, thus making the production of recombinant proteins or the use of two-hybrid assays impossible. Far westerns are particularly useful in such cases. Since far westerns are performed totally *in vitro*, they circumvent these types of problems.

Another advantage of the far western blot is its flexibility. Proteins prepared in a variety of ways can be used for the assay. Cell extracts, recombinant proteins, and peptides can all be used as both probe and target proteins. For example, Palaparti et al. (1997) used cell extracts to probe a semipurified histone sample and detected bound proteins of interest with antibodies. Hsiao and Chang (1999) used phage-expressed proteins immobilized on filter lifts for a far western library screen. Unpurified *E. coli* extracts containing recombinant protein have been successfully used as probes (Fischer et al., 1997).

In addition, many different detection techniques, either radioactive or nonradioactive, can be used. Kimball and co-workers used recombinant protein probes that were labeled radioactively with kinases and recombinant proteins labeled with biotin and subsequently detected with a streptavidin-biotin detection scheme (Grulich-Henn et al., 1998; Kimball et al., 1998). Thus, techniques and reagents already in hand can frequently be adapted for use with this assay.

Finally, the far western can be modified to define the protein domains and amino acid residues that are important in protein-protein interactions. Mutagenized clones can be used to produce variants of protein probes. A single SDS-PAGE gel can be run with identical lanes and cut into strips, and a different *in vitro*-translated probe can be used for each strip. In this way, multiple variants of a protein can be tested simultaneously for their ability to interact with a target protein.

Non-SDS polyacrylamide gels can also be used to separate proteins for far westerns. For example, acid urea gels, which separate on the basis of both size and charge, have been successfully employed. Finally, peptides corresponding to specific interacting sequences can be synthesized with specific post-translational modifications to test their effects on protein-protein interactions.

Critical Parameters

Blocking nonspecific binding sites on the membranes is critical to achieving good results with far westerns. Too little blocking results in high background, while extended time in blocking solutions results in weakened or lost signal. The reason for the diminished signal is unclear, but protein renaturation apparently takes place during the blocking step, so an optimal renaturation may require limited blocking. The best time may well be different for each protein and require empirical optimization. In addition, different lots of BSA appear to result in diminished signal. Therefore, it is important to purchase high-quality BSA from a reputable manufacturer.

An important consideration is the inclusion of appropriate controls to rule out nonspecific interactions that might result in false positives. Suitable controls should be furnished for both the target proteins and the protein probe. When using a complex mixture of proteins, such as cell extracts, as target, “negative” control proteins are already present. However, when using a mixture of only a few proteins, it is important to provide a protein that does not interact to serve as a negative control for nonspecific binding. The ideal negative control should be similar to the protein of interest in charge and size. Appropriate controls should be subjected to SDS-PAGE and blotted in parallel with the samples of interest. Another important control is the use of an unprogrammed translation lysate as a probe. Translation of an unrelated protein as a control probe is also often helpful.

Troubleshooting

Precise conditions for far westerns vary from procedure to procedure and probably reflect the nature of the individual proteins being examined. Optimal conditions for each protein may need to be determined empirically. If background staining is too high, there are several possible remedies. The probe may be diluted or the sample concentration lowered. Other “blocking” reagents may be tested. The blocking reagent, nonfat dry milk, ranging in concentration from 1% to 5%, has been used successfully. Other detergents such as NP-40 and Triton X-100 are commonly employed for this procedure and may help to decrease background. Also, increasing the length of the blocking step may aid in background reduction. Most procedures that call for extended blocking times suggest incubation at 4°C.

If no specific staining is observed, confirm the quality of the radiolabeled probe by SDS-PAGE and autoradiography. Make sure the reticulocyte lysate has not been repeatedly frozen and thawed. If using a recombinant protein/antibody scheme, confirm the affinity of that interaction by immunoblotting.

Sometimes a decrease in blocking time or use of a different blocking reagent will increase signal. Some proteins may interact more readily following a denaturation/renaturation cycle. In this case, the membrane is incubated for 1 hr in PBS-buffered 7 M guanidine or 8 M urea and renatured overnight. Renature in blocking buffer I (see Basic Protocol) without agitation at 4°C with several changes of buffer. Finally, the length of the binding reaction can be increased.

Anticipated Results

Sensitivity of the far western blot is dependent on the affinity of the protein-protein interactions being investigated and on the quality of the probe. Thus, extra attention given to the preparation of a high-quality probe is almost invariably worthwhile. When using a radiolabeled probe, a positive interaction can typically be visualized after overnight exposure to X-ray film.

Time Considerations

The basic protocol can be performed in 2 days. An SDS-PAGE gel can be set up and run in 4 to 6 hr. The rest of the procedure can be performed in 8 to 10 hr. It is often convenient to set up the SDS-PAGE gels on one day and run them slowly overnight. The rest of the basic protocol can then be performed the next day. Semidry transfer requires ~2 hr, the blocking steps take ~4 hr, binding of the probe takes ~2 hr, washing, drying, and setting up the autoradiography cassette takes ~1 hr. Detection of interactions using antibodies to detect the probe protein requires an additional day. Peptide blots can be performed in one day.

It is possible to store the blots after the blocking steps for up to 2 weeks at 4°C. The blots should be stored in airtight containers or wrappings so they do not dry out. In addition, in vitro-translated probes may be produced ahead of time and stored unpurified at -20°C, although the efficiency may be reduced for some probes.

Literature Cited

Chaudhary, J., Cupp, A.S., and Skinner, M.K. 1997. Role of basic-helix-loop-helix transcription factors in Sertoli cell differentiation: Identification

of an E-box response element in the transferrin promoter. *Endocrinology* 138:667-675.

Edmondson, D.G., Smith, M.M., and Roth, S.Y. 1996. Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. *Genes & Dev.* 10:1247-1259.

Fischer, N., Kremmer, E., Lautscham, G., Mueller-Lantzsch, N., and Grasser, F. A. 1997. Epstein-Barr virus nuclear antigen 1 forms a complex with the nuclear transporter karyopherin alpha2. *J. Biol. Chem.* 272:3999-4005.

Grasser, F. A., Sauder, C., Haiss, P., Hille, A., Konig, S., Gottel, S., Kremmer, E., Leinenbach, H. P., Zeppezauer, M., and Mueller-Lantzsch, N. 1993. Immunological detection of proteins associated with the Epstein-Barr virus nuclear antigen 2A. *Virology* 195:550-560.

Grulich-Henn, J., Spiess, S., Heinrich, U., Schonberg, D., and Bettendorf, M. 1998. Ligand blot analysis of insulin-like growth factor-binding proteins using biotinylated insulin-like growth factor-I. *Horm. Res.* 49:1-7.

Hsiao, P. W., and Chang, C. 1999. Isolation and characterization of ARA160 as the first androgen receptor N-terminal-associated coactivator in human prostate cells. *J. Biol. Chem.* 274:22373-22379.

Kimball, S. R., Heinzinger, N. K., Horetsky, R. L., and Jefferson, L. S. 1998. Identification of inter-protein interactions between the subunits of eukaryotic initiation factors eIF2 and eIF2B. *J. Biol. Chem.* 273:3039-3044.

Kleinschmidt, J. A., and Seiter, A. 1988. Identification of domains involved in nuclear uptake and histone binding of protein N1 of *Xenopus laevis*. *EMBO J.* 7:1605-1614.

Kouklis, P. D., Hutton, E., and Fuchs, E. 1994. Making a connection: Direct binding between keratin intermediate filaments and desmosomal proteins. *J. Cell Biol.* 127:1049-1060.

Luna, E.J. 1996. Biotinylation of proteins in solution and on cell surfaces. *In* Current Protocols in Protein Science. (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 3.6.1-3.6.15. John Wiley & Sons, New York.

Palaparti, A., Baratz, A., and Stifani, S. 1997. The Groucho/transducin-like enhancer of split transcriptional repressors interact with the genetically defined amino-terminal silencing domain of histone H3. *J. Biol. Chem.* 272:26604-26610.

Schumacher, T.N.M. and Tsomides, T.I. 1995. In vitro radiolabeling of peptides and proteins. *In* Current Protocols in Protein Science (J.E. Coligan, B.M. Dunn, H.L. Ploegh, D.W. Speicher, and P.T. Wingfield, eds.) pp. 3.3.1-3.3.19. John Wiley & Sons, New York.

Contributed by Diane G. Edmondson and Sharon Y. Roth
M.D. Anderson Cancer Center
Houston, Texas

Two-Hybrid Dual Bait System

In characterizing a given protein's function, it is frequently desirable to identify other proteins with which it interacts. The yeast two-hybrid system, or interaction trap, is one of the most versatile methods available with which to identify/establish protein-protein interactions. This system relies on the modular nature of many transcription factors, wherein different domains (e.g., DNA binding, transcription activating) function independently. The two-hybrid system takes advantage of this functional independence of transcription factor domains by expressing the DNA binding domain and transcription activation domain as components of separate fusion proteins, which can be coalesced via noncovalent interactions to reconstitute a functional transcription factor. For example, in the two-hybrid system, a protein of interest (the bait) may be expressed as a fusion protein with a DNA binding domain. In parallel, the interacting protein (the prey) is expressed as a fusion protein with a transcription activation domain. Co-expression of the bait and prey fusion proteins in the appropriate yeast strain reconstitutes transcription activity through noncovalent bait-prey interactions.

The yeast two-hybrid system was conceived to determine the affinity of single pair protein-protein interactions (Fields and Song, 1989), but rapidly evolved into an efficient means to identify interacting protein binding pairs in both directed and nondirected screens. The latter feature enables the use of the two-hybrid system to screen a bait library against a prey library to identify interacting protein binding pairs. The two-hybrid system, therefore, lends itself to genomic and proteomic level screens. While the two-hybrid system is well suited to high-throughput screening methodology, it is an ideal system for such broad-based screening projects. More recently, the basic system has evolved to encompass a variety of adaptations, which enhance the versatility of the system by including modifying/accessory proteins and compatibility with peptide libraries and RNA-protein screens (reviewed in Serebriiskii et al., 2001).

The dual bait system (Serebriiskii et al., 1999), which is one such adaptation of the classic two-hybrid system, is the focus of this unit. The dual bait system facilitates the simultaneous comparison of two distinct baits with one prey. Briefly, in the dual bait system one protein of interest is expressed as a fusion to the DNA binding protein LexA (bait 1), while a second protein of interest is expressed as a fusion to the DNA binding protein cI (bait 2). Strains of yeast engineered for screening of these dual baits possess four separate reporter genes: *GusA* and *LYS2*, which are transcriptionally responsive to a *cI* operator (*cIop*); and *LacZ* and *LEU2*, which are transcriptionally responsive to a *LexA* operator (*lexop*). A plasmid expressing an activation domain-fused protein (prey), which can be either a defined protein interactor or a cDNA library, is also expressed to allow dual hybrid-mediated transcriptional activation. Selective interaction of the prey with one of the two baits is scored by observing transcriptional activation of either the *LexA* operator-responsive reporters or the *cI* operator-driven reporter genes—e.g., if the prey preferentially interacts with bait 1 (as shown in Fig. 20.7.1), then *LacZ* and *LEU2* will be more activated than *LYS2* and *GusA*. This selective activation can be detected by comparing yeast growth on plates lacking leucine or lysine, and by comparing quantitative results of *LacZ* and *GusA*, respectively.

By facilitating the simultaneous assay of one prey with multiple baits, the dual bait system overcomes two major problems inherent to the two-hybrid approach. First, since many proteins are members of large protein families which share considerable sequence similarity, the degree to which two-hybrid systems can differentiate between specific partners (e.g., individual members of a protein family) and less specific or nonspecific

Contributed by Elena Kotova, Ilya Serebriiskii, and Thomas Coleman

Current Protocols in Molecular Biology (2002) 20.7.1-20.7.37

Copyright © 2002 by John Wiley & Sons, Inc.

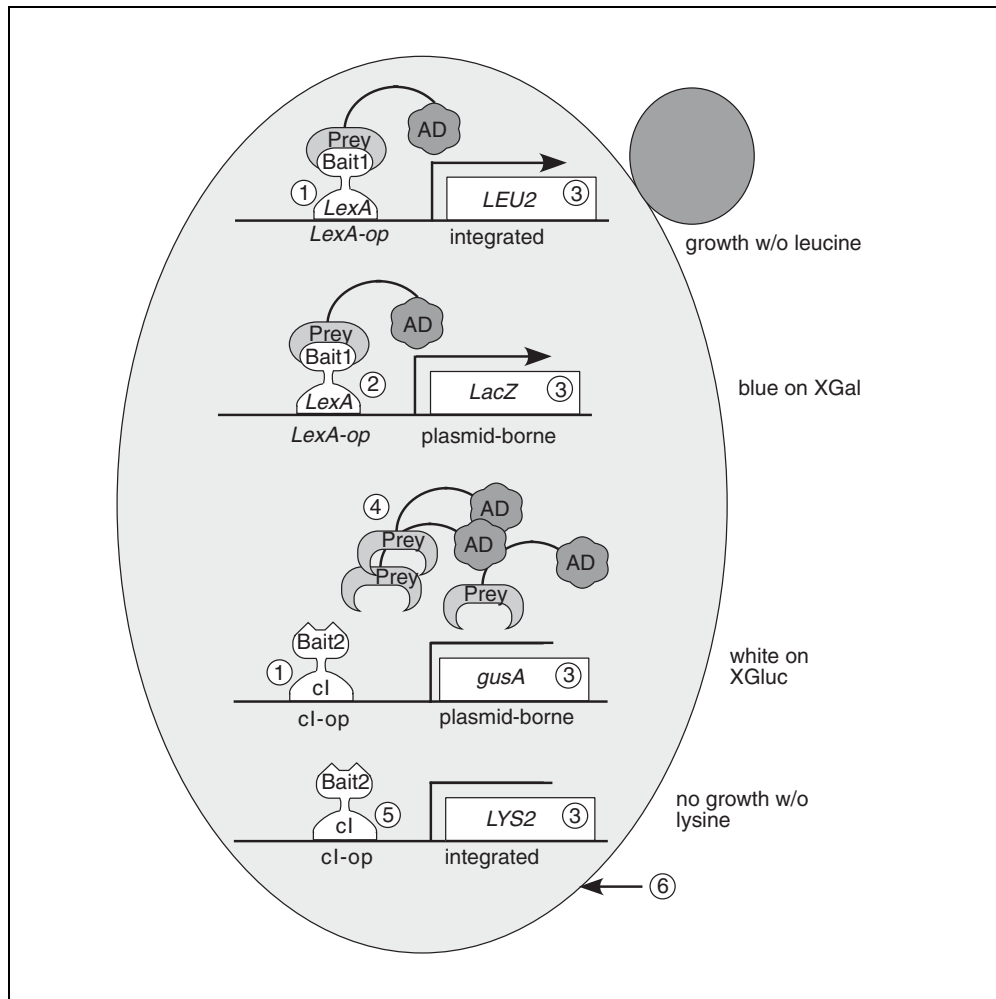


Figure 20.7.1 Outline of the dual bait system and the key control points. An activation domain (AD)-fused prey interacts with Bait1 (LexA-fused) to drive transcription of *lexA-op*-responsive reporters (*LEU2* and *LacZ*), but does not interact with Bait2 (cl-fused) and thus does not turn on transcription of *cl-op*-responsive reporters (*LYS2* and *GusA*). As shown in this figure, the cl-Bait represents a negative control for prey binding; the system can also be configured so prey interacts with only cl-bait or both baits. To maximize the utility of the system, one can adjust experimental parameters affecting expression of baits and reporter readouts. Points of flexibility in the dual bait-hybrid system are indicated in the figure by circled numbers and include: (1) varying expression level of baits, (2) varying cellular localization of baits, (3) varying sensitivity of reporters, (4) diversifying plasmid markers to broaden compatibility with different yeast two-hybrid systems and to facilitate isolation of library plasmid in *E. coli*, (5) enriching polylinkers and choice of reading frames to facilitate cloning of baits, and (6) development of robust yeast strains suitable both for bait testing and interaction mating.

partners (e.g., all of the members of the protein family) for a given bait can be an issue. This notion is of paramount importance since multiple members of a protein family are frequently co-expressed in a single cell type and assigning a single physiological interactor with a particular function can be challenging. Second, the majority of two-hybrid library screens yield one or more nonspecific interactors (see <http://www.fccc.edu/research/labs/golemis/InteractionTrapInWork.html>). These false positives are proteins with broad interaction capabilities that result in nonspecific protein binding. Sorting out the physiologically meaningful interactors from the false positives can be time consuming. The dual bait system, therefore, provides a number of significant advantages, including the ability to readily distinguish interactions specific for individual members of related protein families and specific from nonspecific protein-protein interactions.

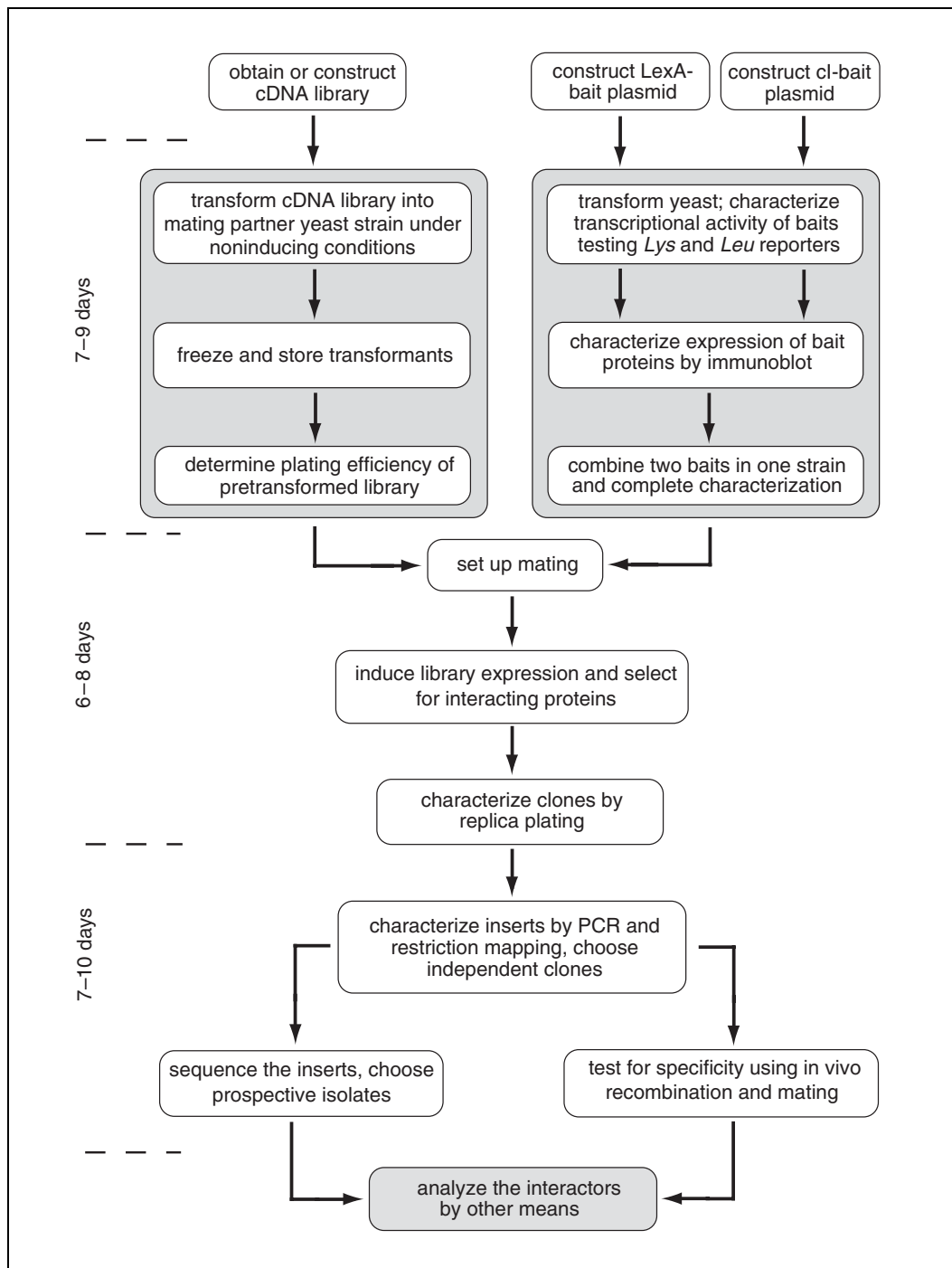


Figure 20.7.2 Flow chart of the two-hybrid screen done by interaction mating. The third stage allows some flexibility, reflected in the availability of different protocols (see Basic Protocol 4 and Alternate Protocol).

The following protocols can be used to investigate questions related to specific protein interactions that are difficult to clarify by other means. This method is a derivation of the classical interaction trap/two-hybrid system and readers are referred to Chapters 1 to 3, 10, and 13 for a more detailed description of the basic methodology related to the execution of the two-hybrid technique (e.g., preparation of yeast medium, immunoblotting). It is recommended that an investigator who is attempting to perform a dual bait screen as a first effort in the field of yeast interaction trap screens consult a basic protocol manual describing the interaction trap system (*UNIT 20.1*).

The outline of the protocol is presented in Figure 20.7.2. In the first protocol (see Basic Protocol 1), baits are transformed into the yeast reporter strains and characterized for protein expression and transcriptional activation. In parallel, a library is introduced into a yeast strain of the opposite mating type, to enable mating (see Basic Protocol 2). In the next method (see Basic Protocol 3), the bait-expressing yeast strain is mated to the library-bearing strain. Expression of library-encoded proteins is induced by growing the diploids in galactose-containing medium, and yeast containing interacting pairs of proteins are identified on the selection plates. Primary candidates are analyzed for their ability to activate the correct set of reporters selected for further study. In the final protocol (see Basic Protocol 4), selected primary candidates are further analyzed by PCR to detect redundant clones, and then the library inserts are isolated and characterized to confirm the specific interaction with the bait protein(s). Clones which satisfy the user's criteria are ready for further analysis.

Three additional protocols are provided as support material. The first (see Support Protocol 1) outlines the Xgal or Xgluc overlay technique, which provides a convenient and sensitive assay to assess β -galactosidase or β -glucuronidase activity of yeast. The second (see Support Protocol 2) describes the detection of bait protein from growth of yeast through immunoblot analysis. Establishing both the expression level and proteolytic status for each novel bait construct is critical for a meaningful interpretation of bait-prey interactions. In the last (see Support Protocol 3) a method to assess the quality of digested library plasmid is described. This protocol is recommended in order to assure minimal background in the testing specificity of interaction using homologous recombination. Together with the basic protocols, these support protocols provide the necessary steps for performing and interpreting the two-hybrid dual bait system.

BASIC PROTOCOL 1

PREPARATION OF BAITS AND LIBRARY: CHARACTERIZING BAIT PROTEINS

To utilize the dual bait system, constructs encoding fusion proteins of independent DNA-binding domains (either LexA or cI) and the two baits must be prepared. A protein to be used as a primary bait should be fused to LexA, since more options exist to optimize LexA bait performance.

Materials

- DNA encoding the proteins of interest
- LexA-fusion plasmids and controls (Table 20.7.1): pMW103 (Figs. 20.7.3 and 20.7.4), pEG202-hsRPB7 (control), pSH17-4, and pEG202-Ras (control)
- cI-fusion plasmids and controls (Table 20.7.2): pGBS10 (Figs. 20.7.4 and 20.7.5), pGBS10-Krit (control), and pGBS10-Krev (control)
- Yeast strains (Table 20.7.3): SKY191 and SKY48
- YPD plates and liquid media (*UNIT 13.1*) with and without 200 μ g/ml geneticin (G418; *APPENDIX 1K*)
- pLacGus (Fig. 20.7.6)

Table 20.7.1 LexA-Fusion Plasmids and Controls

Plasmid name	Yeast selection	<i>E. coli</i> selection	Comment/description	Contact information
<i>LexA Fusion plasmids</i>				
pMW101	<i>HIS3</i>	Cm ^R	Basic plasmids to clone bait as fusion with <i>LexA</i> . Expression is driven by the <i>ADHI</i> promoter.	RB ^a
pMW103	<i>HIS3</i>	Km ^R	Basic plasmids to clone bait as fusion with <i>LexA</i> . Expression is driven by the <i>ADHI</i> promoter.	RB ^a
pEG202	<i>HIS3</i>	Ap ^R	Basic plasmids to clone bait as fusion with <i>LexA</i> . Expression is driven by the <i>ADHI</i> promoter.	Origene ^b , MoBiTec ^c
pJK202	<i>HIS3</i>	Ap ^R	pEG202 derivative, incorporating nuclear localization sequences between <i>LexA</i> and polylinker (enhanced ability to translocate bait to nucleus)	Origene ^b
pNLexA	<i>HIS3</i>	Ap ^R	Polylinker is upstream of <i>LexA</i> , which allows fusion of <i>LexA</i> to C terminus of bait, leaving amino-terminal residues of bait unblocked	Origene ^b
pEG202I	<i>HIS3</i>	Ap ^R	pEG202 derivative (see above), which can be integrated into yeast <i>HIS3</i> gene after digestion with <i>KpnI</i> . Ensures lower levels of bait expression.	RB ^a
pGilda	<i>HIS3</i>	Ap ^R	<i>GALI</i> promoter and CEN-ARS backbone facilitate a tightly controlled, galactose-inducible bait expression; should be used if continuous presence of the bait is toxic to yeast	Origene ^b
pDD	<i>HIS3</i>	Km ^R	<i>GALI</i> promoter and CEN-ARS backbone facilitate a tightly controlled, galactose-inducible bait expression; should be used if continuous presence of the bait is toxic to yeast	R. Hopkins ^d
pHybLex/Zeo	Zeo ^R	Zeo ^R	Bait cloning vector, compatible with IT and all other two-hybrid systems. Minimal <i>ADHI</i> promoter expresses <i>LexA</i> followed by extended polylinker.	Invitrogen ^e
pCGLex/p2GLex	Zeo ^R	Zeo ^R	Gal-inducible bait vector, compatible with IT and all other two-hybrid systems. <i>GALI</i> promoter expresses <i>LexA</i> followed by extended polylinker. Both high and low copy number versions available.	J. Huang ^f
<i>Control LexA-fused baits</i>				
pEG202-Ras	<i>HIS3</i>	Ap ^R	A negative control for activation and positive control for interaction with Raf1 and RalGDS	I. Serebriiskii ^g
pEG202-hsRBP7	<i>HIS3</i>	Ap ^R	A weak positive control for activation	I. Serebriiskii ^g
pEG202-Krit1	<i>HIS3</i>	Ap ^R	A moderate positive control for activation	I. Serebriiskii ^g
pRFHM1 (control)	<i>HIS3</i>	Ap ^R	The homeodomain of bicoid cloned into pEG202 backbone (see above); the resulting nonactivating fusion is recommended as a negative control for activation and interaction assays, and as a positive control for repression assay.	Origene ^b
pSH17-4 (control)	<i>HIS3</i>	Ap ^R	<i>GAL4</i> activation domain cloned into pEG202 backbone (see above) is recommended as a positive control for transcriptional activation.	Origene ^b

^aContact R. Brent at brent@molsci.org.^bSee the Origene Web-site at <http://www.origene.com>.^cSee the MoBiTec Web-site at <http://www.mobitec-germany.com>.^dContact R. Hopkins at richardh@ichr.uwa.edu.au.^eSee the Invitrogen Web-site at <http://www.invitrogen.com>.^fFor more information, refer to Huang and Schreiber (1997) or contact S. Schreiber at sls@slsiris.harvard.edu.^gContact Ilya Serebriiskii at IG_Serebriiskii@fccc.edu.

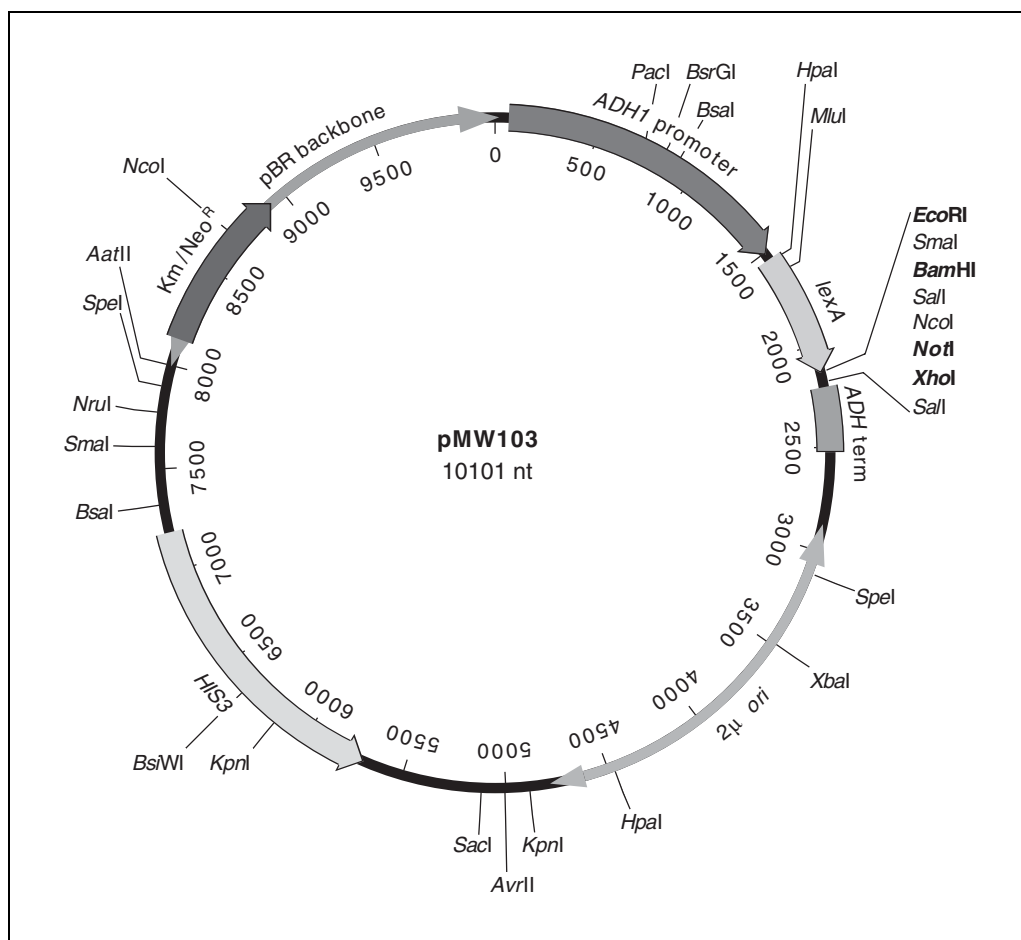


Figure 20.7.3 Plasmid map of the 10101-nt basic dual bait vector pMW103. This plasmid, a derivative of pEG202, uses the strong constitutive *ADH1* promoter to express bait proteins as fusions to the DNA binding protein LexA. The plasmid contains the *HIS3* selectable marker and the 2 μ origin of replication to allow selection and propagation in yeast, the kanamycin resistance gene and the pBR origin of replication to allow selection and propagation in *E. coli*. A detailed map of the LexA-fusion polylinker is given in Figure 20.7.4. More maps and sequences are available on the Web at <http://www.fccc.edu/research/labs/golemis/InteractionTrapInWork.html>.

90- or 100-mm complete minimal (CM) medium dropout plates (*UNIT 13.1*) with and without 350 μ g/ml G418 lacking the following nutrients and supplemented with 2% (w/v) glucose (glu) or 2% (w/v) galactose (gal) and 1% raffinose (raff):
 Glu minus uracil and histidine (Glu/CM –Ura –His)
 Glu minus lysine (Glu/CM –Lys)
 Gal and raff minus lysine (Gal-Raff/CM –Lys)
 Gal and raff minus uracil, histidine, and leucine (Gal-Raff/CM –Ura –His –Leu)
 Gal and raff minus uracil and histidine (Gal-Raff/CM –Ura –His)
 Gal and raff minus uracil, histidine, and lysine (Gal-Raff/CM –Ura –His –Lys)
 Glu minus tryptophan (Glu/CM –Trp; also in 24 \times 24-cm plates)

H₂O, sterile

96-well microtiter plate

Insert grid from a rack of 200- μ l micropipet tips

Tape

200- μ l micropipet tips, sterile

Metal frogger (e.g., Dankar Scientific) or plastic replicator (Bel-Blotter; Bel-Art Products or Fisher)

A
 pMW103 (pEG202)

EcoRI BamHI Sali NotI Sali
 GAA TTC CCG GGG ATC CGT CGA CCA TGG CGG CCG CTC GAG TCG AC

pNLexA

EcoRI NotI XhoI BamHI
 g aat tcg cgg ccg cct cga ggg atc caa ttc ATG AAA
 N S R P P R G I Q F M K

B
 cl-fusion plasmids (e.g., pGBS 10)

A = AAT frame

EcoRI SacI BglIII PvuII ApaI KpnI NotI XhoI Sali
 G AAT TCA AGC TTG AGC TCA GAT CTC AGC TGG GCC CGG TAC CGC GGC CGC CTC AGT CGA Cct gca
 N S S L S S D L S W A R Y R G R S S R P A

B = GAA frame

EcoRI SacI BglIII PvuII ApaI KpnI NotI XhoI Sali
 G AAT Ttg GAA TTC GAG CTC AGA TCT CAG CTG GGC CCG GTA CCG CGG CCG CTC GAG TCG ACC TGC
 N L E F E L R S Q L G P V P R P L E S T C

C
 pJG4-5 (aka pB42AD, displayTarget)

EcoRI XhoI
 ATG GGT GCT CCT CCA AAA AAG AAG ... CCC GAA TTC GGC CGA CTC GAG AAG CTT ...
 M G A P P K K K ... P E F G R L E K L ...

pYesTrp2

KpnI BamHI
HinDIII SacI
 ATG GGT AAG CCT ... AAG CTT GGT ACC GAG CTC GGA TCC ACT AGT AAC GGC
 M G K P ... K L G T E L G S T S N G

EcoRI NotI SphI
BstXI BstXI XhoI
 CGC CAG TGT GCT GGA ATT CTG CAG ATA TCC ATC ACA CTG GCG GCC GCT CGA GGC ATG C
 R Q C A G I L Q I S I T L A A A R G M H

Figure 20.7.4 Polylinkers of the two-hybrid basic vectors. **(A)** The LexA-fusion vectors polylinker for (top) pMW103 (also see Fig. 20.7.3) and (bottom) pNLexA. **(B)** The cl-fusion vectors polylinker. All cl-fusion plasmids with antibiotic resistance markers (Zeo or G418) share this polylinker (e.g., pGBS10; Fig. 20.7.5). Note both the A (top) and B (bottom) reading frames are shown. These are the AAT and GAA reading frames, respectively. **(C)** The AD-fusion (library) vectors polylinker for (top) pJG4-5 (also see Fig. 20.7.8), which is also known as pB42AD and displayTarget, and pYesTrp2. Only restriction sites that are available for insertion of coding sequences are shown; those shown in bold type are unique. The asterisk (*) in panel A denotes that *NcoI* is unique in pEG202 but not pMW103.

Table 20.7.2 cI-Fusion Plasmids and Controls

Plasmid name (reading frames)	Yeast selection	<i>E. coli</i> selection	Comment/description ^a
<i>cI-Fusion plasmids</i>			
pGKS3 (A, B)	<i>HIS3</i>	Ap ^R	<i>ADHI</i> promoter expresses cI followed by polylinker
pGKS4 (A, B)	<i>HIS3</i>	Km ^R	<i>ADHI</i> promoter expresses cI followed by polylinker
pHybcI/HK (B)	<i>HIS3</i>	Km ^R	<i>ADHI</i> promoter expresses cI followed by polylinker ^b
pGKS8 (A, B)	Zeo ^R	Zeo ^R	Dual purpose vector: <i>ADHI</i> promoter expresses cI followed by polylinker, while cI-responsive <i>gusA</i> reporter cassette is integrated into the same backbone.
pGKS6 (A, B, C)	Zeo ^R	Zeo ^R	<i>ADHI</i> promoter expresses cI followed by polylinker
pGKS7 (A, B)	Zeo ^R	Zeo ^R	Modified <i>ADHI</i> promoter ensures higher level of expression of cI
pGBS9 (A, B)	G418 ^R	Km ^R	<i>ADHI</i> promoter expresses cI followed by polylinker
pGBS10 (A, B)	G418 ^R	Km ^R	Modified <i>ADHI</i> promoter ensures higher level of expression of cI
pGMS11 (A)	Zeo ^R	Zeo ^R	<i>GAL</i> promoter expresses cI followed by polylinker, for use with baits whose continuous presence is toxic to yeast
pGMS12 (B)	G418 ^R	Km ^R	<i>GAL</i> promoter expresses cI followed by polylinker, for use with baits whose continuous presence is toxic to yeast
<i>Control cI-fused baits</i>			
pGKS3-Krev	<i>HIS3</i>	Ap ^R	Expresses cI-Krev fusion protein. Use as negative control for activation assay and positive control for interaction with Krit1.
pHybcI/HK-Krev	<i>HIS3</i>	Km ^R	Expresses cI-Krev fusion protein. Use as negative control for activation assay and positive control for interaction with Krit1 ^b .
pGBS9-Krev	G418 ^R	Km ^R	Expresses cI-Krev fusion protein. Use as negative control for activation assay, and positive control for interaction with Krit1.
pGBS10-Krev	G418 ^R	Km ^R	Expresses cI-Krev fusion protein. Use as negative control for activation assay, and positive control for interaction with Krit1.
pGKS6-Krev	Zeo ^R	Zeo ^R	Expresses cI-Krev fusion protein. Use as negative control for activation assay, and positive control for interaction with Krit1.
pGKS6-Krit	Zeo ^R	Zeo ^R	Expresses cI-Krit fusion protein. Use as positive control for activation assay.
pGKS7-Krit	Zeo ^R	Zeo ^R	Expresses cI-Krit fusion protein. Use as positive control for activation assay.
pGKS3-Krit	<i>HIS3</i>	Ap ^R	Expresses cI-Krit fusion protein. Use as positive control for activation assay.
pGBS9-Krit	G418 ^R	Km ^R	Expresses cI-Krit fusion protein. Use as positive control for activation assay.
pGBS10-Krit	G418 ^R	Km ^R	Expresses cI-Krit fusion protein. Use as positive control for activation assay.
pGMS12-Krit	G418 ^R	Km ^R	Expresses cI-Krit fusion protein. Use as positive control for activation assay.

^aContact Ilya Serebriiskii at IG.Serebriiskii@fccc.edu.

^bAvailable from Invitrogen (<http://www.invitrogen.com>).

Two-Hybrid Dual Bait System

20.7.8

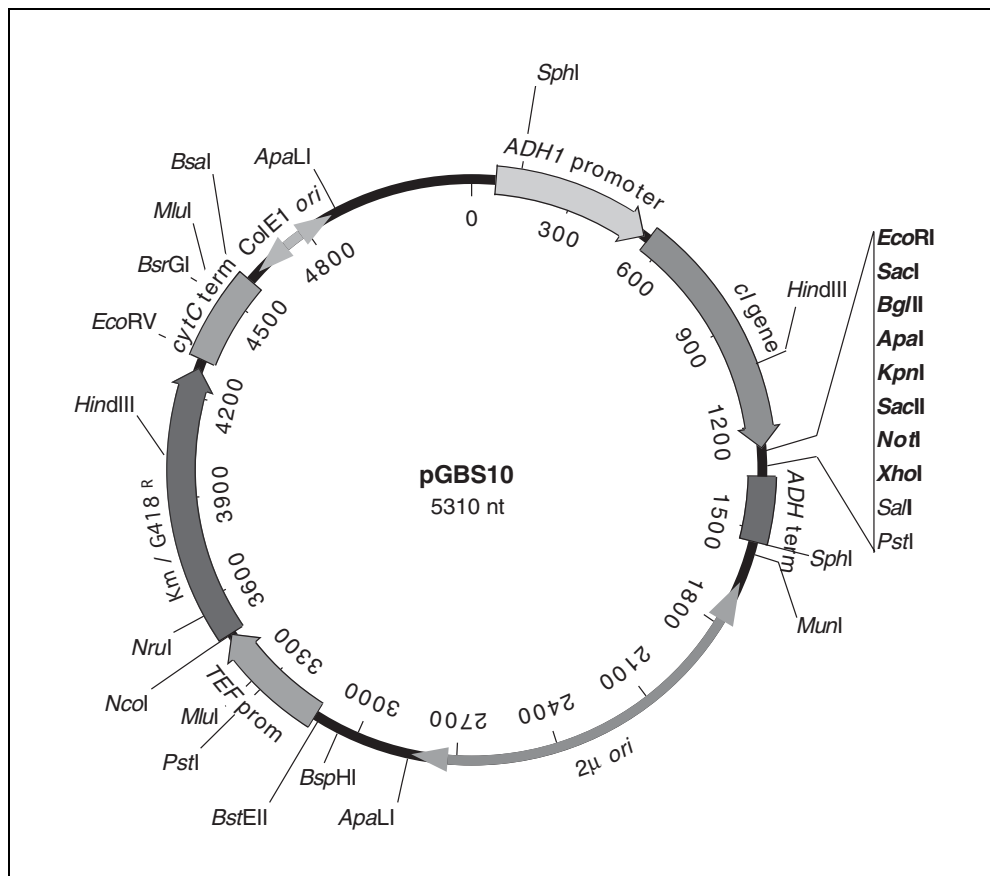


Figure 20.7.5 Plasmid map of the 5310-nt basic dual bait vector pGBS10. This plasmid uses the strong constitutive *ADH1* promoter to express bait proteins as fusions to the DNA binding protein *cI*. The plasmid contains the 2 μ origin of replication to allow propagation in yeast and the ColE1 origin of replication to allow propagation in *E. coli*. The same Tn903-encoded gene confers kanamycin resistance in *E. coli* and geneticin (G418) resistance in yeast. More maps and sequences are available on the Web at <http://www.fccc.edu/research/labs/golemis/InteractionTrapInWork.html>.

Table 20.7.3 *LEU2/LYS2* Selection Strains

Name	Genotype	No. operators	Comment/description	Contact information
SKY48	<i>MATα, trp1, his3, ura3, lexAop-LEU2, cIop-Lys2</i>	6 <i>lexA</i> , 3 <i>cI</i>	Provides a more stringent selection for interaction partners of <i>cI</i> -fused baits, and more sensitive LexA-responsive <i>LEU2</i> reporter than the one in SKY191	I. Serebriiskii ^a
SKY191	<i>MATα, trp1, his3, ura3, lexAop-LEU2, cIop-Lys2</i>	2 <i>lexA</i> , 3 <i>cI</i>	Provides a more stringent LexA-responsive <i>LEU2</i> reporter, and more sensitive <i>cI</i> -responsive <i>LYS2</i> reporter than the one in SKY48.	I. Serebriiskii ^a
SKY473	<i>MATα, his3, leu2, trp1, ura3, lexAop-LEU2, cIop-LYS2</i>	4 <i>lexA</i> , 3 <i>cI</i>	Mating partner for SKY strains; can be also used as a reporter strain itself. Sensitivity of <i>LEU2</i> reporter is intermediate between sensitivity of <i>LEU2</i> in SKY48 and SKY191. Sensitivity of <i>LYS2</i> reporter is the same as sensitivity of <i>LYS2</i> in SKY191.	I. Serebriiskii ^a

^aE-mail at IG_Serebriiskii@fccc.edu.

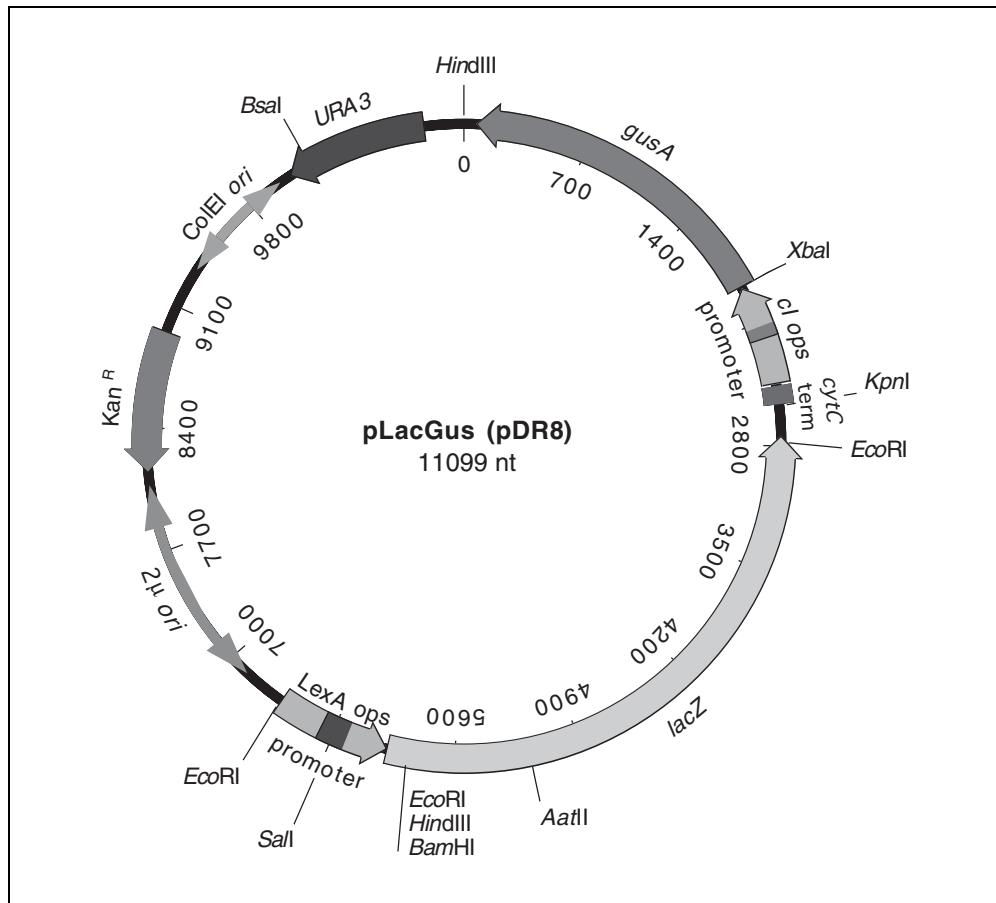


Figure 20.7.6 Plasmid map of the 11099-nt basic dual bait vector pLacGus, which is also known as pDR8. pLacGus is a dual reporter plasmid. Eight *LexA* operators upstream of the *lacZ* reporter make it the most sensitive of the *lexA*-responsive *lacZ* reporters. The *cI*-responsive *gusA* reporter (three *cI* operators) has a comparable sensitivity range. These two cassettes are separated by *cytC* terminator. The plasmid contains the *URA3* selectable marker and the 2 μ origin of replication to allow selection and propagation in yeast, and the kanamycin resistance gene and the pUC origin of replication to allow selection and propagation in *E. coli*. More maps and sequences are available on the Web at <http://www.fccc.edu/research/labs/golemis/InteractionTrapInWork.html>.

Additional reagents and equipment for subcloning DNA fragments (UNIT 3.16) or alternative cloning strategies (e.g., *in vivo* recombination; Ma et al., 1987), lithium acetate transformation of yeast (UNIT 13.7), gal overlay assay (see Support Protocol 1), and immunoblotting and immunodetection (see Support Protocol 2)

1. Using standard techniques for subcloning DNA fragments (UNIT 3.16) or alternative cloning strategies (e.g., *in vivo* recombination), insert the DNA encoding one of the two proteins of interest into the polylinker of the *LexA*-fusion plasmid pMW103 (Figs. 20.7.3 and 20.7.4A) to make an *in-frame* protein fusion to *LexA*. Call this plasmid pMW103-Bait1.

A number of modified versions of the plasmid exist which contain additional sites, altered reading frame, alternate antibiotic resistance markers, or include an in-frame nuclear-localization motif between the LexA protein and the polylinker (see Background Information in the Commentary and Table 20.7.1).

2. Similarly, clone the DNA encoding the second protein of interest into the polylinker of the *cI*-fusion plasmid pGBS10 (Figs. 20.7.4B and 20.7.5) to enable synthesis of an *in-frame* protein fusion to *cI*. Call this plasmid pGBS10-Bait2.

3. Lithium acetate transform yeast strain SKY191 (UNIT 13.7) using the following plasmids:

- pGBS10-Bait2 (step 2)
- pGBS10-Krit (positive control for activation)
- pGBS10-Krev (negative control for activation).

Plate each transformation mixture on YPD plates containing 200 µg/ml geneticin (YPD G418), and incubate two days at 30°C to select for yeast colonies containing transformed plasmid.

Incubating yeast in YPD liquid medium 6 hr to overnight at 30°C prior to plating significantly increases plating efficiency.

4. Transform SKY48 (UNIT 13.7) using the following combinations of LexA-fusion and *lexAop-LacZ* (pLacGus) plasmids:

- pMW103-Bait1 (step 1) + pLacGus
- pEG202-hsRPB7 + pLacGus (weak positive control for activation)
- pSH17-4 + pLacGus (strong positive control for activation)
- pEG202-Ras + pLacGus (negative control for activation).

Transfer the transformations to 90- to 100-mm Glu/CM –Ura –His plates without G418 to select for yeast that contain both plasmids, and incubate 2 to 3 days at 30°C.

If bait transformants either grow noticeably slower or in reduced number (compared to the controls), the bait protein is likely toxic to the yeast (see Critical Parameters and Troubleshooting, Table 20.7.11 and Table 20.7.12 for recommended modifications). Increased heterogeneity of yeast colonies compared to controls (e.g., a mix of large and small colonies) is also suggestive of bait toxicity.

If only a small number of colonies are obtained even in controls, or colonies are not apparent within three to four days, the transformation efficiency is low. (While the transformation efficiency at this particular step is not paramount, it will be critical for the library transformation). In this case, all solutions, media, and conditions must be double-checked or prepared fresh, and the transformation repeated. An efficient transformation should yield $\sim 1 \times 10^4$ transformants per microgram DNA (when two plasmids are being simultaneously transformed), and up to 1×10^5 for transformation with a single plasmid.

5. Replica plate to assess activation of *cI*- and LexA-fused baits:

- a. Add 50 to 75 µl sterile water to each well of one-half (6 × 8 wells) of a 96-well microtiter plate. Place an insert grid from a rack of 200-µl micropipet tips over the top of the microtiter plate and attach with tape.

The grid should be elevated ~1 cm from the plate and the holes in the insert grid should be aligned with the wells of the microtiter plate, as this will keep the tips in the upright position in the plate and allow their simultaneous removal. A convenient pipet-tip brand which accomplishes this is the Rainin RT series.

- b. Pick six 1- to 2-mm-diameter yeast colonies from each of the transformation plates (steps 3 and 4) using a different sterile plastic 200-µl micropipet tip for each colony. Leave the tips supported in a near-vertical position by the insert grid until all the colonies have been picked.
- c. Swirl the plate gently to ensure the yeast are mixed into suspension. Remove the sealing tape and lift the insert grid along with all of the tips.
- d. Use a metal frogger or plastic replicator to plate yeast suspensions (each spoke leaves an ~3-µl drop) on the following plates, marked for orientation:

YPD containing 200 µg/ml G418 (master plate for cI-fused baits)
Glu/CM –Lys
Gal-Raff/CM –Lys
Glu/CM –Ura –His (master plate for LexA-fused baits)
Gal-Raff/CM –Ura –His –Leu
Gal-Raff/CM –Ura –His (for Xgal overlay assay).

- e. Incubate the plates up to 4 days at 30°C, and save the YPD G418 and Glu/CM –Ura –His master plates at 4°C.

Yeast containing cI-fused bait and controls should grow on YPD G418 plates, but should not grow on any –Ura –His plates. On Glu/CM –Lys and Gal-Raff/CM –Lys plates, yeast containing the positive control (pGBS10-Krit) should exhibit growth within 4 days, while yeast containing the negative control (pGBS-Krev) should not grow.

Yeast containing LexA-fused bait and controls should not grow on either YPD G418 or CM –Lys plates, and should grow on Glu/CM –Ura –His plates. On Gal-Raff/CM –Ura –His –Leu plates, the strong positive control (pSH17-4 + pLacGus) should show detectable growth in 1 to 2 days, the weak positive control (pEG202-hsRBP7 + pLacGus) should be growing within 4 days, and the negative control (pEG202-Ras + pLacGus) should not grow. If the yeast containing the bait under test (pMW103-Bait1 + pLacGus) shows no growth in this period, it should be suitable for library screening. If it activates LEU2 reporter similar to the transformants from the pEG202-hsRBP7 reaction, it may be suitable, but is likely to have a high background in library screening (which may be slightly reduced by the use of a less sensitive strain). If it is similar to the pSH17-4 + pLacGus transformants, it must be modified before screening.

6. Approximately 24 to 30 hr after the plating, overlay the Gal-Raff/CM –Ura –His with Xgal agarose as described (see Support Protocol 1). Assess the ability of LexA-fused bait to activate transcription of *LEU2* and *LacZ* reporters.

In assessing the control transformants (step 4), strongly activating baits will be detectable as dark blue colonies in 20 to 60 min, while negative controls should be faint blue or white; an optimal bait would be either as white as the negative control or only a faint blue color.

At this step, the ability of LexA-fused bait to activate transcription is tested on both LEU2 and LacZ reporters, while potential cI self-activation is only tested for LYS2 reporters. Auxotrophic reporters are, however, the most important for the library screening, because they allow the direct selection for interaction phenotype. In addition, there is normally a good correlation between activation of the two reporters, so it is very unlikely that the bait which is not activating LYS2 will significantly activate GusA. Therefore, if no activation is detected on –Lys plates, one should proceed further; if bait causes growth on –Lys plates, it should be modified. The ability of the cI-fused bait to activate GusA reporter will be tested in step 9 below.

7. Select from the master plate at least two primary transformants for each novel bait construct and perform immunoblot analysis as described (see Support Protocol 2).

Establishing protein expression levels by immunoblot analysis is critical for several reasons. Proteins expressed at low levels, and apparently inactive in transcriptional activation assays, can be up-regulated to much higher levels under the auxotrophic selection and unexpectedly demonstrate a high background of transcriptional activation. Moreover, where proteins are proteolytically clipped, screens might inadvertently be performed with LexA fused only to the amino-terminal end of the larger intended bait.

Adding positive controls for protein expression such as pGBS10-Krev (step 3) and pEG202-Ras (step 4) is helpful.

8. Note which cI-bait colonies on the master plate express bait appropriately, and use two of these colonies to introduce LexA-fused bait and reporter plasmids. Use standard transformation procedure (UNIT 13.7), except grow the yeast in YPD liquid medium containing 200 µg/ml G418.

This transformation should combine in a single strain both of the baits and a double reporter plasmid—i.e., pMW103-Bait1 + pLacGus + pGBS10-Bait2 (clones 1 and 2).

In parallel, transform SKY191 expressing pGBS10-Krit and pGBS10-Krev to obtain pEG202-Ras + pLacGus + pGBS10-Krev and pEG202-hsRPB7 + pLacGus + pGBS10-Krit.

9. Plate each transformation mixture on Glu/CM –Ura –His G418 dropout plates, and maintain at 30°C for 2 to 4 days to select for yeast colonies containing desired plasmid combinations.
10. Replica-plate (step 5) twelve colonies of each transformant on the following set of plates:

Glu/CM –Ura –His G418 (new master plate)
Gal-Raff/CM –Ura –His G418 (Xgal overlay assay)
Gal-Raff/CM –Ura –His G418 (Xgluc overlay assay)
Gal-Raff/CM –Ura –His –Leu
Gal-Raff/CM –Ura –His –Lys

Incubate the plates up to 4 days at 30°C, and save the Glu/CM –Ura –His G418 master plate at 4°C.

11. Overlay with Xgal and Xgluc agarose as described (see Support Protocol 1) and analyze the results.

Self-activation abilities are expected to be the same as in the characterization of separate baits.

12. Reconfirm expression of the baits by immunoblot (see Support Protocol 2).

An investigator who is proficient in yeast transformations may wish to cotransform all baits and reporter plasmids into a single strain simultaneously. If the resultant strain behaves favorably, this single transformation can save time. It should be cautioned, however, that simultaneous transformation with three plasmids can be problematic and the authors recommend the step-wise strain construction and verification.

PREPARATION OF BAITS AND LIBRARY: TRANSFORMING AND CHARACTERIZING THE LIBRARY

A list of some of the libraries currently available for use with the Dual Bait system can be found at <http://www.fccc.edu/research/labs/golemis/InteractionTrapInWork.html>. Currently, the most convenient source of libraries suitable for the interaction trap is commercial, and can be viewed at the companies' web-sites.

The protocol outlined below describes the steps used to perform a screen that should saturate a cDNA library derived from a genome of mammalian complexity. Fewer plates will be required for screens with libraries derived from organisms with less complex genomes; therefore, the protocol should be scaled back accordingly.

Currently, mating the library-pretransformed strain with the desired bait strain is recommended as the most convenient strategy. The main advantage of this approach is that if the investigator wishes to use the same library to screen multiple baits, only a single large-scale transformation is required, followed by relatively easy mating steps. This approach is also useful when analyzing a toxic bait, as yeast-expressing toxic proteins can be difficult to transform with high efficiency. Finally, direct transformation in the bait strain requires media not only selective for the library plasmid, but also maintaining selective pressure to keep both baits and reporter. Large-scale transformation plating on G-418 medium will make screening much more expensive.

BASIC PROTOCOL 2

Analysis of Protein Interactions

20.7.13

As outlined in Figure 20.7.7, it is recommended to perform the large-scale library transformation (plated on multiple 24 × 24-cm plates) in parallel with several small-scale control transformations (each plated on a single 90-mm plate). These small-scale controls are critical in calculating the frequency of false positives and will be useful as positive and negative controls during interactor characterization.

Materials

Yeast strains (Table 20.7.3), fresh: SKY473
YPD liquid media (UNIT 13.1) without G418 (APPENDIX 1K)
H₂O, sterile
TE buffer (APPENDIX 2)/0.1 M lithium acetate
Library DNA in pJG4-5 (Fig. 20.7.8) or pYesTrp (Table 20.7.4)
Carrier DNA, freshly denatured
Negative control plasmids (Table 20.7.4; optional): pJG4-5 or pYesTrp2, pJG4-5-Raf1, pJG4-5-Krit, and pYesTrp2-RalGDS
40% (w/v) PEG 4000/0.1 M lithium acetate/TE buffer, pH 7.5
Dimethyl sulfoxide (DMSO)
90-mm and 24 × 24-cm Glu/CM –Trp plates (UNIT 13.1)
TE buffer (APPENDIX 2), sterile (optional)
Glycerol solution: 65% (w/v) sterile glycerol/0.1 M MgSO₄/25 mM Tris·Cl, pH 8.0 (APPENDIX 2)
Orbital shaker, 30°C
50-ml conical tubes, sterile
42°C heat block
3- to 4-mm glass balls, sterile (Thomas Scientific or Fisher)
Additional reagents and equipment for lithium acetate transformation of yeast (UNIT 13.7)

Prepare competent yeast

1. Inoculate a colony of SKY473 in ~20 ml liquid YPD medium and grow overnight on an orbital shaker at 30°C.

It is important to use fresh yeast (i.e., thawed from –70°C and streaked to a single colony less than ~7 days previously) and maintain sterile conditions throughout all subsequent procedures.

2. Dilute the overnight culture into ~300 ml YPD liquid medium such that the diluted culture has an OD₆₀₀ of ~0.15. Incubate at 30°C on an orbital shaker until the culture has reached an OD₆₀₀ of ~0.50 to 0.7.
3. Transfer the culture to six sterile 50-ml conical tubes, and centrifuge 5 min at 1000 to 1500 × g, room temperature. Gently resuspend the pellets in ~5 ml sterile water each, and combine all slurries into one of the conical tubes. Add sterile water to 50 ml and mix.

Perform transformations

4. Centrifuge cells again 5 min at 1000 to 1500 × g, room temperature. Decant water and resuspend yeast in 1.5 ml TE buffer/0.1 M lithium acetate (UNIT 13.7).
5. Mix 30 μg library DNA in pJG4-5 or pYesTrp with 1.5 mg freshly denatured carrier DNA in a microcentrifuge tube and add this mixture to the yeast. Mix gently and aliquot ~60 μl DNA/yeast suspension into each of 30 microcentrifuge tubes.
6. *Optional:* Use aliquots of competent yeast from step 4 to transform control plasmids—i.e., empty library plasmid (pJG4-5 or pYesTrp2), pJG4-5-Raf1, pJG4-5-

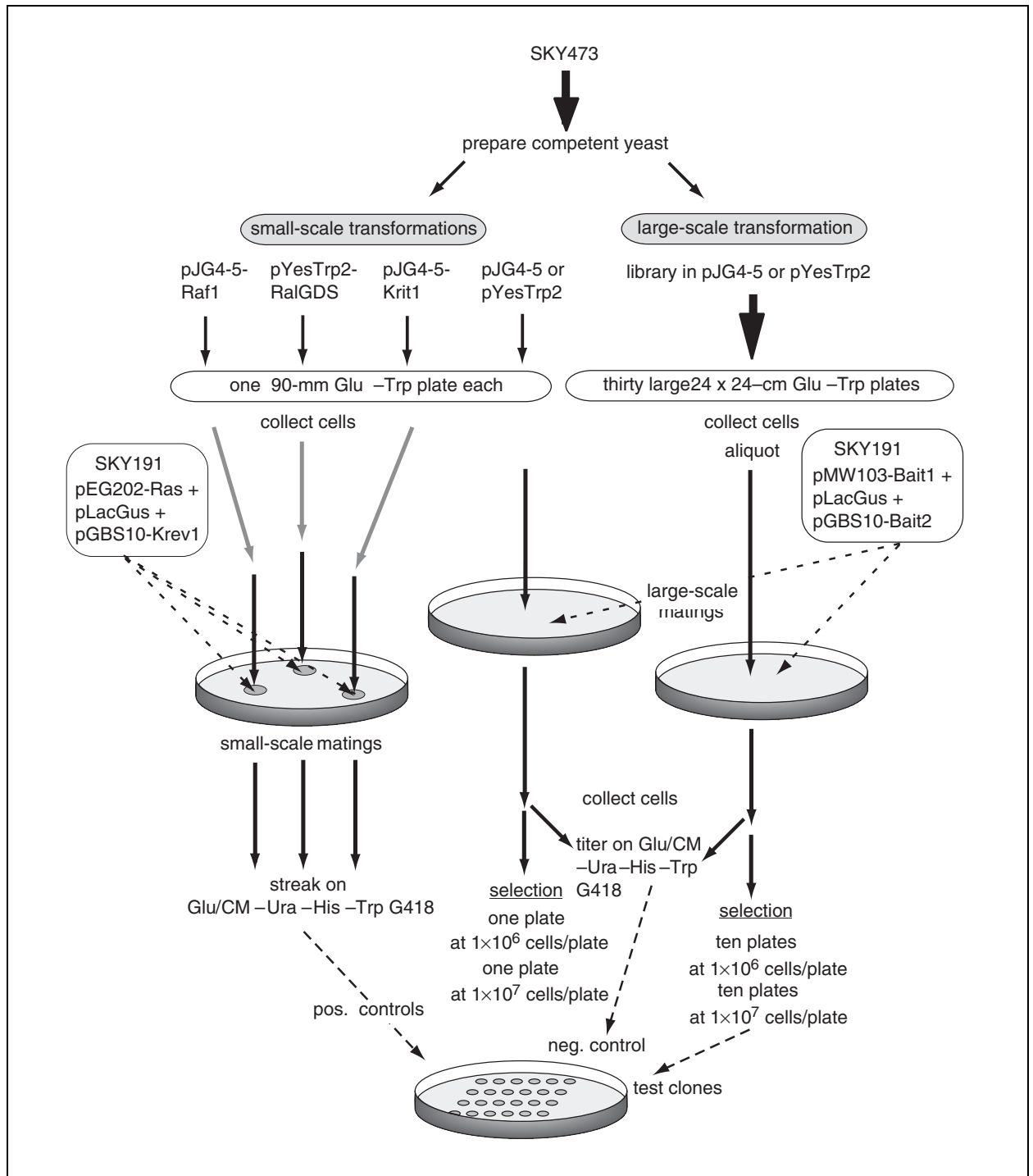


Figure 20.7.7 Detailed library screening flow chart. See Basic Protocols 2 and 3 for details.

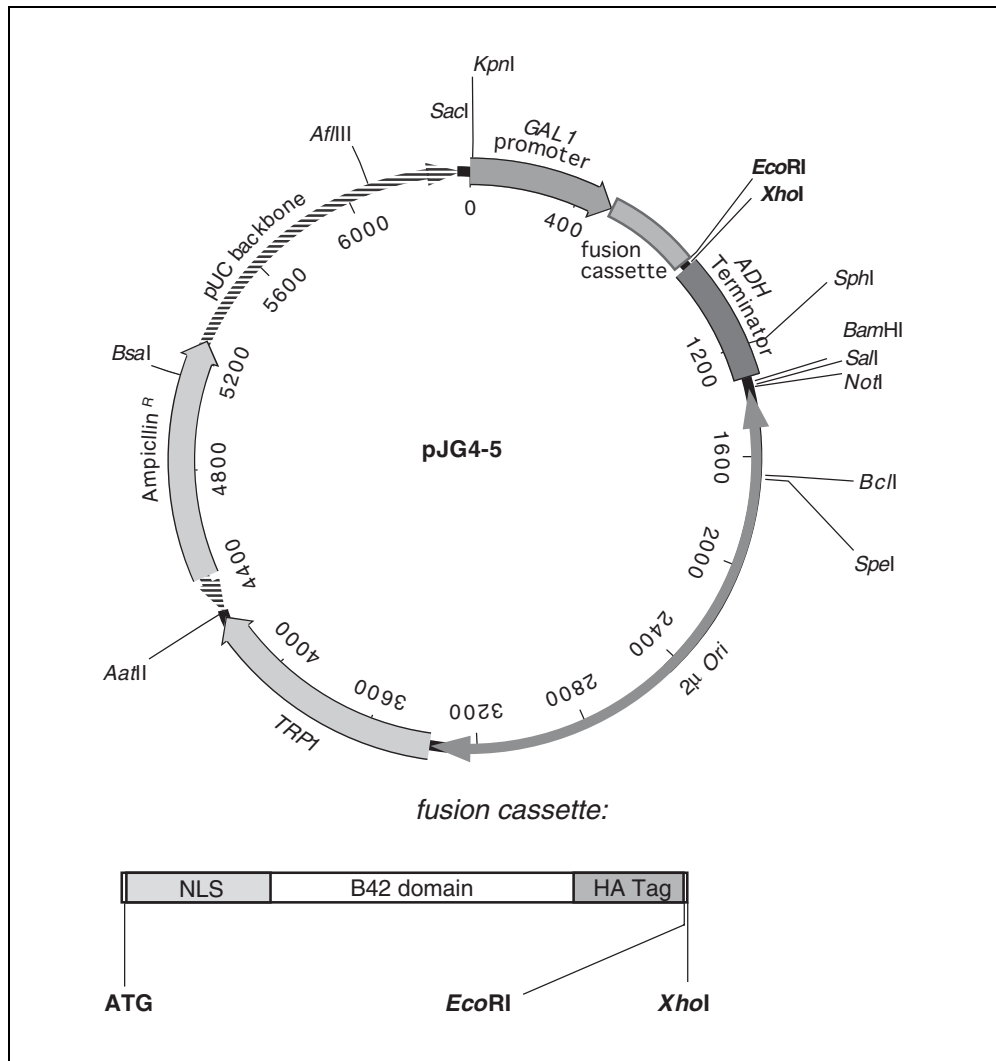


Figure 20.7.8 Plasmid map of the basic dual bait vector pJG4-5. This plasmid expresses cDNAs or other coding sequences inserted into the unique *EcoRI* and *XhoI* sites as translational fusion to a cassette consisting of the SV40 nuclear localization sequence, the acid blob B42, and the hemagglutinin (HA) epitope tag. Expression of sequences is under the control of the *GAL1* galactose-inducible promoter. The plasmid contains the *TRP1* selectable marker and the 2 μ origin of replication to allow selection and propagation in yeast, and the ampicillin resistance gene and pUC origin of replication to allow selection and propagation in *E. coli*. The fusion cassette is shown at the bottom. A detailed map of the AD-fusion polylinker is given in Figure 20.7.7. More maps and sequences are available on the Web at <http://www.fccc.edu/research/labs/golemis/InteractionTrapInWork.html>.

Krit1, and pYesTrp2-RalGDS—into SKY473. Transfer to a 100-mm Glu/CM – Trp plate and collect the transformed cells as for the library. The control strain with an empty library plasmid can be safely reamplified in liquid medium.

When using a new bait strain, it is recommended to set up a parallel mating with negative and positive controls as outlined in Figure 20.7.7. The negative control strain is the same strain used for the library (e.g., SKY473), but containing the library vector with no cDNA insert. Positive control strains are also very helpful and will be later used in mating with a positive control bait strain (e.g., pEG202-Ras + pLacGus + pGBS10-Krev).

Table 20.7.4 Activation Domain Fusion Plasmids and Controls

Plasmid name	Yeast selection	<i>E. coli</i> selection	Comment/description	Contact information
pJG4-5	<i>TRP1</i>	Ap ^R	Library construction plasmid; <i>GAL1</i> promoter provides efficient expression of a gene fused to a cassette consisting of nuclear localization sequence, transcriptional activation domain, and HA epitope tag	Origene ^a , MoBiTec ^b
pJLo	<i>TRP1</i>	Ap ^R	A derivative of pJG4-5 that has a lower copy number (CEN/ARS <i>ori</i>)	R. Hopkins ^c
pJG4-5I	<i>TRP1</i>	Ap ^R	A derivative of pJG4-5 that can be integrated into yeast <i>TRP1</i> gene after digestion with Bsu36I; designed to study interactions that occur physiologically at low protein concentrations (in combination with pEE202I)	RB ^d
pYesTrp	<i>TRP1</i>	Ap ^R	<i>GAL1</i> promoter expresses nuclear localization domain, transcriptional activation domain, V5 epitope tag, multiple cloning sites; contains <i>f1 ori</i> and T7 promoter/flanking site. Used to express cDNA libraries	Invitrogen ^e
pNB42 series	<i>TRP1</i>	Ap ^R	Allow fusion to the N terminus of an AD, leaving N-terminal residues of Prey unblocked; various multiple cloning sites. No libraries yet available	M. Brown ^f
pMW102	<i>TRP1</i>	Km ^R	Same as pJG4-5, but with altered antibiotic resistance markers; no libraries yet available	RB ^d
pMW104	<i>TRP1</i>	Cm ^R	Same as pJG4-5, but with altered antibiotic resistance markers; no libraries yet available	RB ^d
pCGB42/p2GB42	G418 ^R	Km ^R	The same Tn903-encoded gene confers kanamycin resistance in <i>E. coli</i> and geneticin (G418) resistance in yeast; both high- and low copy number versions available. Multiple cloning site.	J. Huang ^g
<i>Control activation domain-fusions</i>				
pJG4-5-Raf	<i>TRP1</i>	Ap ^R	A positive control for interaction with Ras	I. Serebriiskii ^h
pYesTrp-RalGDS	<i>TRP1</i>	Ap ^R	A positive control for interaction with Ras and Krev	I. Serebriiskii ^h
pJG4-5-Krit	<i>TRP1</i>	Ap ^R	A positive control for interaction with Krev	I. Serebriiskii ^h

^aSee the Origene homepage at <http://www.origene.com>.

^bSee the MoBiTec Website at <http://www.mobitec-germany.com>.

^cE-mail at richardh@ichr.uwa.edu.au.

^dE-mail at brent@molsci.org.

^eSee the Invitrogen Website at <http://www.invitrogen.com>.

^fSee Brown and MacGillivray (1997).

^gFor more information, refer to Huang and Schreiber (1997).

^hE-mail at IG_Serebriiskii@fcc.edu.

7. To each tube (steps 5 and 6), add 300 μ l sterile 40% (w/v) PEG 4000/0.1 M lithium acetate/TE buffer pH, 7.5. Gently invert the tubes several times (do not vortex) and incubate 30 to 60 min at 30°C.
8. To each tube, add 40 μ l DMSO and mix by inversion. Place tubes in a 42°C heat block for 10 min.
9. Evenly spread the contents of each tube onto a 24 \times 24-cm Glu/CM –Trp plate, using sterile 3- to 4-mm glass balls. Invert plates and incubate at 30°C until colonies appear (usually 2 to 4 days).

It is acceptable and efficient to keep the glass balls on the lids while incubating the plates; they will be needed to harvest the library transformants (see below). One to two dozen glass balls is sufficient.

10. Select a couple of representative transformation plates and draw a 23 \times 23-mm square (1% of plate bottom surface) over an average density spot. Count colonies in each grid section and recalculate for the whole transformation.

An efficient transformation done by this protocol should yield ~20,000 to 40,000 colonies per large plate, and represent $\sim 1 \times 10^5$ transformants per microgram library DNA.

Doing transformations in small aliquots helps reduce the likelihood of contamination, and for reasons that are not clear, provides significantly better transformation efficiency than scaled-up versions. Note, it is not recommended to use excess transforming library DNA per aliquot competent yeast cells, as each competent cell may take up multiple library plasmids, complicating subsequent analysis. Under the conditions described here, <10% of yeast will contain two or more library plasmids.

Harvest transformants

11. Inspect all transformation plates. If visible molds or other contaminants (e.g., colonies of unusual color, shape) are observed on the plates, carefully excise them and a region around them using a sterile scalpel prior to harvesting the library's transformants.
12. Pour 10 to 15 ml sterile water on each of five 24 \times 24-cm plates containing transformants. Add 3- to 4-mm sterile glass balls. Stack five of the plates on top of each other. Holding on tightly, shake the stack horizontally until all colonies are resuspended (1 to 2 min). Using a sterile pipet, tilt each plate to collect 5 ml yeast slurry and pool into sterile 50-ml conical tubes. Proceed to the next five plates, resulting in a total volume of up to 150 ml liquid spread across three 50-ml tubes.
13. Adjust the volume of each conical tube containing yeast to 50 ml with sterile TE or water, and vortex/invert to suspend cells. Centrifuge tubes ~5 min at 1000 to 1500 \times g, room temperature, and discard the supernatant. Wash cells by adding sterile water, resuspending, and centrifuging ~5 min at 1000 to 1500 \times g, room temperature. Discard the supernatant.
14. Resuspend each pellet in a small volume of sterile water and combine all slurries into one tube. Add water to 50 ml, and repeat wash (step 13). After the second wash, the cumulative pellet volume should be ~25 ml cells derived from 1.5×10^6 transformants.
15. Resuspend pellet in 1 vol glycerol solution. Freeze in 1-ml aliquots at –70°C.

The frozen aliquots are stable at least 1 year. If the intent is to proceed directly with mating, leave one aliquot unfrozen and assume viability of the culture is 100%.

SELECTING AN INTERACTOR

After the bait strain has been characterized (see Basic Protocol 1) and the library strain has been transformed and frozen in aliquots (see Basic Protocol 2), the next step is to mate the two strains. To mate the two strains, the bait strain is grown in liquid culture and then mixed with a thawed aliquot of the pretransformed library strain. The mixture is then plated on YPD medium and incubated overnight. During this time individual cells of the bait strain will fuse with individual cells of the library strain to form diploid cells. The mixture of diploids and unmated haploids is then collected, and an aliquot plated on media to select for interactors. In practice the diploid/haploid mixture is generally frozen in a few aliquots to allow titering and repeated platings at various cell densities.

The authors recommend mating any new bait strain with a control strain (see Basic Protocol 2). Mating with the control strain (containing the library vector with no cDNA insert) can be performed at the same time as the library mating, and both matings can be treated identically in the next step, selecting interactors. This control will provide a clear estimate of the frequency of cDNA-independent false positives, a frequency that is important to know when deciding how many positives to pick and characterize.

In the second part of this procedure, interactors will be selected by plating the mated cells onto selection plates lacking leucine and/or lysine. It will be important to know how many viable diploids were plated on these selection plates to gain a sense of how much of the library has been screened and to determine the false-positive frequency (see Basic Protocol 4 for discussion). To save time, the titering can be done at the same time as the selection for interactors in the protocol below. Since the titer of diploids will likely be within an order of magnitude of 1×10^8 cfu/ml, one could plate two or three different dilutions onto the selection plates to select for interactors, while at the same time plating dilutions to determine the exact titer of diploids.

Finally, the protocol describes testing for galactose-inducible transcriptional activation of both the auxotrophic and colorimetric reporters. Simultaneous activation of both reporters in a galactose-specific manner generally indicates that the transcriptional phenotype is attributable to expression of library-encoded proteins, rather than derived from mutation of the yeast.

Materials

Complete minimal (CM) medium dropout plates (or medium; *UNIT 13.1*) with and without 350 μ g/ml G418 (as necessary) lacking the following nutrients and supplemented with 2% (w/v) glucose (glu) or 2% (w/v) galactose (gal) and 1% raffinose (raff):

- Glu/CM –Ura –His
- Gal-Raff/CM –Ura –His –Trp
- Glu/CM –Ura –His –Trp
- Gal-Raff/CM –Ura –His –Trp –Leu
- Glu/CM –Ura –His –Trp –Leu
- Gal-Raff/CM –Ura –His –Trp –Lys
- Glu/CM –Ura –His –Trp –Lys
- Gal-Raff/CM –Ura –His –Trp
- Gal-Raff/CM –Lys
- Glu/CM –Ura –His
- Glu/CM –Leu (optional)
- Glu /CM –Lys (optional)
- Gal-Raff/CM –Leu (optional)
- Gal-Raff/CM –Lys (optional)
- Gal-Raff/CM –Ura –His –Trp –Lys –Leu (optional)

Bait strains containing appropriate combinations of plasmids (i.e., pMW103-Bait1 + pLacGus + pGBS10-Bait2) on Glu/CM –Ura –His G418 (see Basic Protocol 1)

Pretransformed library strain (see Basic Protocol 2)

H₂O, sterile

90- or 100-mm-diameter YPD plates (UNIT 3.1)

Glycerol solution: 65% (w/v) sterile glycerol/0.1 M MgSO₄/25 mM Tris·Cl, pH 8.0

30°C incubator with and without shaker

1.5-ml microcentrifuge tube, sterile

3- to 4-mm sterile glass balls (Thomas Scientific or Fisher)

Markers or wax pencils of different colors

Insert grid from a rack of 200- μ l pipet tips

Metal frogger (e.g., Dankar Scientific) or plastic replicator (i.e., Bel-Blotter; Bel-Art Products or Fisher)

96-well microtiter plate

Additional reagents and equipment for lithium acetate transformation of yeast (UNIT 13.7) and Support Protocol 1

Mate the bait strain and the pretransformed library

1. Inoculate 30 ml Glu/CM –Ura –His containing 350 μ g/ml G418 with the bait strain containing appropriate combinations of plasmids (i.e., pMW103-Bait1 + pLacGus + pGBS10-Bait2) from the Glu/CM –Ura –His G418 master plate. Grow in a 30°C incubator with shaking to mid- to late-log phase (i.e., OD₆₀₀ = 1.0 to 2.0).

Note that controls are to be grown in parallel with the bait strain (see step 7).

2. Collect the cells by centrifuging 5 min at 1000 \times g, room temperature. Resuspend the cell pellet in 1 ml sterile water and transfer to a sterile 1.5-ml microcentrifuge tube.

This will yield a yeast suspension of $\sim 1 \times 10^9$ cells/ml.

3. Thaw an aliquot of the pretransformed library strain at room temperature. Mix $\sim 2 \times 10^8$ cells of the bait strain (~ 200 μ l) with $\sim 1 \times 10^8$ colony-forming units (CFU) of the pretransformed library strain on a single 100-mm-diameter YPD plate and incubate at 30°C overnight. Optionally, in parallel, perform mating under the same conditions with the strain carrying the empty library plasmid (see Basic Protocol 2, step 6).
4. Add 1.5 to 2.0 ml sterile water to the surface of the YPD plate and resuspend the cells using 3- to 4-mm sterile glass balls. Transfer the suspension to a sterile tube and vortex gently 2 min. Collect the cells by centrifuging 5 min at 1000 \times g, room temperature.

5. Resuspend in 1 vol sterile glycerol solution. Distribute into 200- μ l aliquots and freeze at –80°C.

As with the frozen pretransformed library strain, the mated yeast should not be thawed and refrozen. Since only one or a few of the aliquots will be needed to represent the library, thawed aliquots can be discarded after use.

6. Titer the mated cells by plating serial dilutions on Glu/CM –Ura –His –Trp plates (unmated haploids will not grow on this medium). Count the colonies that grow after 2 to 3 days, and determine the titer of the frozen mated cells.

Titering can be performed in parallel with the selection step and is necessary to estimate how successful the mating was and therefore how many plates are needed to get full representation of the library. For full representation, it is desirable to have each primary library clone represented on the selection plate by 3 to 10 diploid yeast cells (see

explanation below). For example, if the library comprises 1×10^6 clones then a total of $\sim 5 \times 10^6$ diploid cells is to be plated on selection plates. Assuming that the titer of diploids is 5×10^7 cfu/ml, then ~ 100 μ l slurry has to be plated. However, the total number of cells plated will be much higher, since the efficiency of mating is typically 1% to 10%. An approximate titer can be estimated from the OD reading. Most likely, the same volume of the mating slurry which contains 5×10^6 cfu/ml of diploids, would contain about 1×10^8 cells. To avoid cross-feeding, only 1×10^6 – 10^7 cells can be put on one 100-mm plate. Therefore, to screen the library exhaustively, 10 to 100 plates would be needed (see Basic Protocol 4).

The authors recommend plating three to ten times more diploids than the number of colonies obtained from the transformation, based on the assumption that not all cells that contain interacting proteins plate at 100% efficiency on selective medium (Estojak et al., 1995), and because the created slurry may not be completely homogenous. While this approach may lead to multiple isolations of the same cDNA, it will increase the likelihood that all primary transformants are represented by at least one diploid on the selective plate. In fact, re-isolation of an identical cDNA among a relatively small set of “positives” can be taken as one sign of a specific interaction.

7. *Optional:* In parallel with steps 1 to 5 above, grow an overnight culture of the bait strain containing pEG202-Ras + pLacGus + pGBS10-Krev (control) in a final volume ~ 1.5 ml each. Make three spots of control bait strain by dropping ~ 5 μ l liquid culture on a YPD plate. Without waiting for the liquid to soak in, add 5 μ l of one of the three control prey strains (i.e., pJG4-5-Raf1, pJG4-5-Krit1, or pYesTrp2-RafGDS) to the same spots. Incubate overnight at 30°C and then streak all three matings onto Glu/CM –Ura –His –Trp plates containing G418.

See Figure 20.7.7 for an outline.

Screen for interacting proteins

8. Thaw an aliquot of the mated yeast (step 5). Dilute 100 μ l into 10 ml Gal-Raff/CM –Ura –His –Trp G418 liquid dropout medium and incubate 5 hr with shaking at 30°C. If the frozen culture was not previously titered, plate serial dilutions onto Glu/CM –Ura –His –Trp G418 plates. In parallel, do the same to an aliquot of the yeast mated to the control strain (step 3).
9. On the assumption that a culture at an OD₆₀₀ of 1.0 contains $\sim 1 \times 10^7$ cells/ml, plate 1×10^6 cells on ten 100-mm plates with the appropriate auxotrophic selection medium. In parallel, plate 1×10^7 cells on each of the ten additional plates with the same medium. Similarly, plate 1×10^6 cells control mating on one plate and at 1×10^7 cells/plate on another.

Plating 1×10^7 cells/plate allows screening more diploids on fewer number plates, but may result in higher levels of background growth (see below).

10. Incubate up to 6 days at 30°C. Compare selection plates seeded with lower and higher densities, which should be roughly proportional to seeding. Discard sets of plates in which disproportionately more colonies appear on the more densely seeded plates (especially sitting on the thin lawn), as this is probably background due to cross-feeding. Calculate how many plates at 1×10^6 cells per plate is needed for full representation of the calculated number of diploids. Repeat induction and plating using another frozen mating aliquot.

Depending upon the individual bait used, good candidates for positive interactors will generally produce LEU⁺ colonies during this time, with the most common appearance of colonies at 2 to 4 days. LYS⁺ colonies typically form at 3 to 5 days.

11. Inspect the plates on a daily basis. Each day, mark appearing colonies with different colors of ink or wax. At day 4 or 5, pick colonies in a microtiter plate format onto a

Glu/CM –Ura –His –Trp G418 master plate, in which colonies are grouped by day of appearance. If many apparent positives appear, pick separate master plates for colonies arising on days 2, 3, and 4, respectively.

If contamination has occurred at an earlier step, this is generally reflected by the growth of a very large number of colonies (more than 500/plate) within 24 to 48 hr after plating on selective medium. Compare the smell of the plates and morphology of the colonies to that on the titer plates. A low-power microscope can be helpful for this purpose. In the case of bacterial contamination, screening can be rescued by recreating the selective plates with 15 µg/ml tetracycline, and repeating library induction and plating.

Colonies arising later than within a week are likely to be artifactual.

12. Include positive control colonies—i.e., from matings of SKY191 bearing pEG202-Ras + pLacGus + pGBS10-Krev with SKY473 bearing pJG4-5-Raf1, pJG4-5-Krit1, and pYesTrp2-RalGDS (step 7)—on each of the master plates. Also, include a few colonies from the titer plate (step 8 above). If any colonies appeared on the control mating, include two or three representatives on one of the master plates.

Since they contain randomly chosen library plasmids, the phenotype of the colonies from the titer plates will almost certainly be negative.

The number of colonies to pick and characterize should be based on the frequency of cDNA-independent false positives that arise on the auxotrophic plates for the control mating with empty library vector. (The frequencies of true/false positives should be calculated as number of colonies per 1×10^6 diploids.) The higher the false positive frequency, the more colonies one will need to screen in order to find rare true positives. For example, if the cDNA-independent false positive frequency is one colony in 1×10^4 cfu plated and a true positive exists at a frequency of 1 in 1×10^6 cfu, then it will be necessary to pick at least 100 colonies to find this rare true positive clone.

13. Incubate the master plates at 30°C until patches/colonies form.

Confirm positive interactions

14. Invert a metal frogger or plastic replicator on a flat surface and place a master plate upside down on the spokes, making sure that the spokes and colonies are properly aligned. Remove the plate and insert the frogger or replicator into a microtiter plate containing 50 µl sterile water in each well. Let the plate sit for 5 to 10 min, shaking from time to time to resuspend the cells left on the spokes. When all yeast are resuspended, test for activation of the following genes by printing on the following plates:

Glu/CM –Ura –His –Trp G418 (master plate)
Gal-Raff/CM –Ura –His –Trp –Leu (*LEU2* growth assay)
Glu/CM –Ura –His –Trp –Leu (*LEU2* growth assay)
Gal-Raff/CM –Ura –His –Trp –Lys (*LYS2* growth assay)
Glu/CM –Ura –His –Trp –Lys (*LYS2* growth assay)
2× Glu/CM –Ura –His –Trp G418 (Xgal and Xgluc overlay assays)
2× Gal-Raff /CM –Ura –His –Trp G418 (Xgal and Xgluc overlay assays)

Repeat for each master plate.

The authors omit antibiotic for the plates used to characterize binding to LexA-fused baits, or when an alternative selection is available to maintain pressure on positive interactors for cI-fusions (as on the Lys-medium, which provides selection for growth). This antibiotic omission both accelerates the yeast growth and reduces the cost of the screen. In general, test plates for auxotrophic reporter characterization lacking only leucine or lysine would automatically keep selective pressure for the presence of the prey and the corresponding bait plasmids. Using plates with fewer dropped-out components would slightly accelerate

Table 20.7.5 Interpretation of Interaction Phenotype for Each Bait

Observed phenotype				Interpretation		Recommendation
Auxotrophic reporter		Colorimetric reporter		Traditional	Optimistic	
Glu	Gal	Glu	Gal			
–	+	–	+	Very good sign	Very good sign	Work with those clones first
(+)	+	(+)	+	Bait is mutated or its expression is up-regulated to a high background of transcriptional activation.	<i>GAL1</i> promoter is slightly leaky, both proteins are very stable, and/or interaction occurs with high affinity	Take one clone for confirmation of interaction and store the rest
–	+	–	–	Yeast mutation occurred that favors growth or transcriptional activation on galactose medium	Some bait-interactor combinations are known to preferentially activate one reporter versus another	If all other candidates winnow out, check these clones
All other phenotypes				Contamination, plasmid rearrangement, and/or mutations	Something really new	Trash

Table 20.7.6 Expected Phenotype for Control Set of Interactions^a

Bait	Prey	<i>LEU2</i>	<i>lacZ</i>	<i>LYS2</i>	<i>gusA</i>	Explanation
cI-Krev + LexA-Ras	Raf1	+	+	–	–	Interaction with LexA-fused bait only
cI-Krev + LexA-Ras	RalGDS	+	+	+	+	Interaction with both baits
cI-Krev + LexA-Ras	Krit	–	–	+	+	Interaction with cI-fused bait only
LexA-bait1 + cI-bait2	random	–	–	–	–	No interaction with any bait

^aEach feature marked as positive should be also Gal-dependent, as in Table 20.7.5, top row. Comparing the phenotype of the clones of interest to this set of controls should help to assess whether the isolated prey interacts with one or both baits.

the growth, and the potential loss of other plasmids would not influence the results of the assay on these plates. Thus, –Ura –His –Trp –Lys and –Ura –His –Trp –Leu plates can be substituted at the investigators discretion for –Lys and –Leu, respectively.

15. Incubate the plates 3 to 4 days at 30°C. After 20 to 30 hr of incubation, take out all –Ura –His –Trp plates. Retain one Glu/CM –Ura –His –Trp G418 plate as a fresh master plate, and overlay the remaining two sets with Xgal or Xgluc agarose as described in Support Protocol 1.

16. Assess growth on the –Leu and –Lys plates 48 to 72 hr after plating.

For interpretation of the results, see Table 20.7.5 and Table 20.7.6.

ANALYSIS OF PRIMARY INTERACTORS

Positives obtained during the selection step must next be evaluated for reproducible phenotype and specificity of interaction with the bait used to select them. Two strategies for analyzing positives are summarized in the flow chart in Figure 20.7.9. The rapid screen (PCR) approach is provided here, while the plasmid isolation approach is given elsewhere in the unit (see Alternate Protocol). The choice between strategies is a judgment call, and depends on whether the individual investigator prefers spending time and money doing bulk PCR (this protocol) or bulk yeast plasmid recovery (see Alternate Protocol). This protocol is generally 1 to 3 days faster.

BASIC PROTOCOL 4

Analysis of Protein Interactions

20.7.23

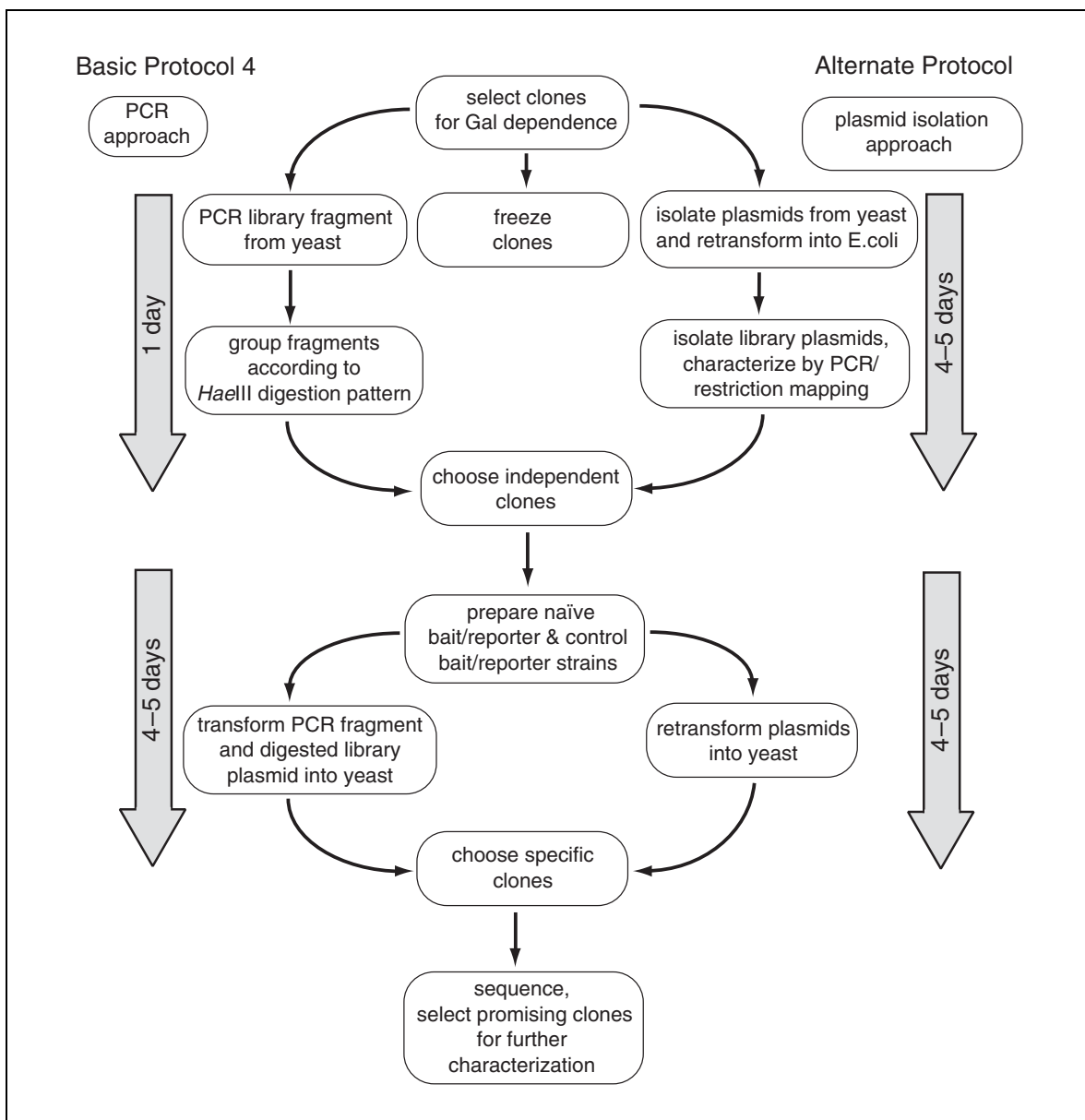


Figure 20.7.9 Detailed flow chart for characterization and second confirmation of primary positives. See Basic Protocol 4 and Alternate Protocol for details.

The next step is to determine whether isolated cDNAs will reproduce an interaction phenotype upon introduction into a fresh bait strain. This reconfirmation step eliminates a variety of false positives, including library-encoded cDNAs that interact with LexA (instead of the bait), “sticky” proteins that interact with the pBait in a nonspecific manner, and clones isolated because of mutations in the yeast strain that render nonbait mediated growth and transcriptional activation. This can be done using a PCR-recombination approach (derived from Petermann et al., 1998) in a single step, in which PCR-amplified cDNA fragments from the primary isolates and the digested library plasmid will be transformed together into a naïve bait strain. In yeast, they will undergo homologous recombination *in vivo* in up to 97% of the transformants that acquired both vector and insert. This is due to the identity between the cDNA PCR fragment and the plasmid at the sites flanking the polylinker. Colonies obtained in this protocol should be tested for

interaction phenotype. After this step, confirmed specific positive clones can be worked up through conventional plasmid purification.

To use the second step of this protocol, there must be ~0.5 to 1 µg of PCR product from the yeast colonies obtained in the first step (reamplify if necessary). Empty library plasmid should be digested with two enzymes producing incompatible ends in the polylinker region (see Support Protocol 3). Together, they will be transformed into the naïve bait strain SKY191 containing pMW103-Bait1 + pLacGus + pGBS10-Bait3 (see Basic Protocol 1, steps 8 to 12). An optional transformation in a control bait strain—SKY191 containing pEG202-Ras + pLacGus + pGBS10-Krev (see Basic Protocol 1, steps 8 to 12)—will provide additional specificity controls. Including in this transformation set Raf1 PCR fragment (from Support Protocol 3) will make the control side of this experiment impeccable.

Bait-expressing strains (see Basic Protocol 1) can only be used if they are >10 days old. Otherwise, these yeast must be replated or the whole transformation process repeated. In any case, single colonies from the fresh bait strains have to be characterized by immunoblotting to confirm the expression of bait proteins.

A major strength of this protocol is that it will identify redundant clones prior to plasmid isolation and bacterial transformation, which in some cases greatly reduces the amount of work required. However, accurate records should be maintained as to how many of each class of cDNA are obtained; and if any ambiguity is present as to whether a particular cDNA is part of a set or unique, investigators should err on the side of caution.

Materials

Glu/CM –Ura –His –Trp G418 master plate (see Basic Protocol 3)

β-glucuronidase solution: dilute crude β-glucuronidase type HP-2 from *H. pomatia* (Sigma) 1:50 in 50 mM Tris-Cl, pH 7.5 (APPENDIX 2)/10 mM EDTA; prepare fresh

Library plasmid specific primers—e.g., for JG4-5:

Forward primer (FP1): 5'-CTG AGT GGA GAT GCC TCC

Reverse primer (FP2): 5' CTG GCA AGG TAG ACA AGC CG

*Hae*III (UNIT 3.1)

SKY191 containing pMW103-Bait1 + pLacGus + pGBS10-Bait2 (see Basic Protocol 1, steps 8 to 12)

SKY 191 containing pEG202-Ras + pLacGus + pGBS10-Krev (see Basic Protocol 1, steps 8 to 12; optional)

Raf1 and Krit1 PCR fragment (see Support Protocol; optional)

Glu/CM –Ura –His –Trp G418 dropout plates (UNIT 13.1)

Metal frogger (e.g., Dankar Scientific) or plastic replicator (i.e., Bel-Blotter; Bel-Art Products or Fisher)

96-well microtiter plate

Horizontal shaker, 37°C

Tape

150- to 212-µm glass beads (e.g., Sigma G-1145)

Vortex with flat-top surface

Additional reagents and equipment for PCR (UNIT 15.1), restriction endonuclease digestion (UNIT 3.1), agarose gel electrophoresis (UNIT 2.5A), purification of DNA fragments (UNIT 2.6), lithium acetate transformation of yeast (UNIT 3.17), sequencing (Chapter 7), rapid miniprep isolation of yeast DNA (UNIT 3.6), and miniprep isolation of bacterial DNA (UNIT 1.6).

Rapid screen for interaction trap positives

1. Starting from a Glu/CM –Ura –His –Trp G418 master plate, use a metal frogger or plastic replicator to resuspend yeast in 25 μ l β -glucuronidase solution in a 96-well microtiter plate (also see Basic Protocol 3, step 14). Seal the wells using tape, and incubate 1.5 to 3.5 hr on a horizontal shaker at 37°C.

Alternatively, pipet tips can be used to transfer yeast in the wells (similar to replica plating; see Basic Protocol 1, step 5). This option is especially useful when the initial set of candidates was spread over a few master plates, and only a fraction of them has to be tested further. Do not take more than approximately the volume of one middle-sized yeast colony; otherwise, quality of isolated DNA may suffer. The master plate does not need to be absolutely fresh: plates that have been stored for a week at 4°C have been successfully used.

2. Remove the tape, add ~25 μ l (i.e., a small scoop) of 150- to 212- μ m glass balls to each well, and seal again. Attach the microtiter plate to a vortex with a flat top surface (e.g., using rubber bands) and mix vigorously for 5 min.
3. Add 100 μ l sterile water to each well. Take 1 to 2 μ l as a template for each PCR reaction. Reseal the plate with tape and keep the remainder frozen at –70°C.

*In many cases, PCR products can be obtained directly from the yeast colonies even without enzymatic treatment (e.g., by introducing a 10-min 94°C step at the beginning of the PCR program). However, the crude yeast lysate obtained in this protocol (step 2) can also be used as a source of plasmid DNA (use 1 to 2 μ l for electroporation into *E. coli*), making it worth preparing this simple lysate.*

4. Prepare master mix (25 to 30 μ l/reaction) as described in UNIT 15.1 using library-specific primers.

For example, use primer FP1 and FP2 for the JG4-5-based library or commercially available primers (Invitrogen) for pYesTrp2-based libraries.

5. Perform a PCR amplification (UNIT 15.1) in 25 to 30 μ l using the following program:

Initial step:	2 min	94°C	(denaturation)
35 cycles:	45 sec	94°C	(denaturation)
	45 sec	56°C	(annealing)
	5 sec	72°C	(extension).

Modified versions of this protocol with extended elongation times were also found to work. The variant given above has been shown to amplify 1.8-kb fragments well.

To interpret the results of the PCR, it is helpful to have the following ~1 ng control templates: empty library plasmid, yeast from the positive control colonies treated in parallel with experimental clones, and the same amounts of library plasmid and positive control yeast, mixed together. Interpretation of possible PCR outcomes is summarized in Table 20.7.7.

If the library being screened is based on JG4-5 plasmid, only clones containing Raf1 and Krit plasmids would produce products. For a pYesTrp-based library, use a RalGDS clone as positive control.

6. Load 2/3 (~20 μ l) of the PCR reaction onto a 0.7% agarose gel (UNIT 2.5A). Identify fragments which appear to be of the same size. Based on insert sizes, group obtained interactors in families (i.e., identify potential multiple isolates of the same cDNAs). Store gel in the refrigerator until ready to isolate fragments.
7. While the gel with the uncut PCR products is running, perform a restriction digest of the remainder (~10 μ l) of the PCR product with 15 U *Hae*III in a total volume of 20 μ l. Based on classification of undigested PCR products (step 6), rearrange the loading

Table 20.7.7 Analysis of PCR Reaction

Empty library plasmid	Template			Possible cause	Solution
	Control yeast	Empty library + control yeast	Clone of interest		
–	–	–	–	Bad master mix, wrong settings, or a faulty amplifier were used	Double check reaction and repeat
+	–	+	–	Not enough template	Add more template/improved
+	–	–	–	Lysed yeast inhibited PCR	Add less template
+	±	±	±	Uneven template load/digestion	Adjust template, PCR from obtained bands again

order of *HaeIII*-digested samples in such a way that the digests of the PCR products thought to be identical are run side-by-side. Load the *HaeIII* digestion products on a 1.0% agarose gel (UNIT 2.5A) and electrophoresis to optimize resolution of DNA products in the 200 to 1000 bp range.

HaeIII is recommended only because it tolerates the presence of PCR buffer well. If the PCR product is purified, any four-base cutter can be used.

This technique will generally yield distinctive and unambiguous groups of inserts, confirming whether multiple isolates of a small number of cDNAs have been obtained. As mentioned above, sometimes a single yeast will contain two or more different library plasmids. If this happens it is typically revealed by PCR. Thus, after bacterial transformation, and increased number of clones should be checked to avoid the loss of the “real” interactor.

- Purify undigested fragments from the agarose gel using standard techniques (UNIT 2.6).

In cases where a very large number of isolates have been obtained from a small number of cDNAs, the investigator may choose to sequence the PCR product directly.

PCR products obtained from this step will also be used for reconfirmation of an interaction. To get enough PCR products, reamplify if necessary.

Only the forward primer, FP1, works reliably in sequencing PCR fragments, while the reverse primer will only work in sequencing from purified plasmid (Chapter 7). In general, the TA-rich nature of the ADH terminator sequences downstream of the polylinker in the pJG4-5 vector makes it difficult to design high quality primers in this region.

Perform specificity testing

- Lithium acetate transform (UNIT 3.17) combinations of digested library plasmid and selected PCR products (optionally including Raf1 and Krit1; see Support Protocol 3) into the naïve bait strain SKY191 (pMW103-Bait1 + pLacGus + pGBS10-Bait2). Optionally, transform the same combinations in parallel into the control bait strain SKY191 (pEG202-Ras + pLacGus + pGBS1-Krev).
- Plate each transformation mix on Glu/CM –Ura –His –Trp G418 dropout plates and incubate at 30°C until colonies grow (2 to 3 days).
- Prepare a master plate for each library plasmid being tested. Each plate should contain at least ten colonies of the transformed PCR-insert/digested plasmid into each of the bait strains.

12. Test for the activity of colorimetric and auxotrophic reporters as described (see Basic Protocol 3, steps 14 and 15). At least ten to twelve colonies from each transformation should be tested.

True positives should show a LEU⁺ LacZ⁺ phenotype with pMW103-Bait1 + pLacGus + pGBS10-Bait2, but not with pEG202-Ras + pLacGus + pGBS10-Krev. Clones transformed with Raf control PCR will provide both positive and negative controls: pMW103-Bait1 + pLacGus + pGBS10-Bait2 should be negative while pEG202-Ras + pLacGus + pGBS10-Krev should be positive when assayed for β -galactosidase and growth on –Leu plates, but negative for β -glucuronidase activity and growth on –Lys plates. Vice versa, clones of pEG202-Ras + pLacGus + pGBS10-Krev transformed with Krit1 control PCR products should be blue on Xgluc, but not Xgal, and grow on –Lys, but not on –Leu dropout plates. Again, pMW103-Bait1 + pLacGus + pGBS10-Bait2 should be completely negative. The fraction of the correct clones in positive controls is indicative of the efficiency of recombination in this particular experiment. Normally, it should be between 85% and 95%.

13. Proceed with sequencing (Chapter 7) and biological characterization.

14. If recovery of yeast plasmids for archival purposes is desired, isolate yeast plasmids using traditional phenol-chloroform technique (UNIT 3.6) or a commercially available kit. Transform selected positives into *E. coli*, using 1 to 2 μ l β -glucuronidase-treated frozen yeast (see Basic Protocol 3).

Use of electroporation is highly recommended.

*An alternative protocol can be used in case PCR technology is not readily available, or in case of failure to obtain a specific PCR product using the library vector primers. It is based on plasmid isolation, its amplification in *E. coli*, and then transformation back into yeast (see Alternate Protocol).*

A database of common false positives has been compiled and made available on the World Wide Web, at <http://www.fccc.edu/research/labs/golemis/interactiontrapinwork.html>. Particularly for cDNAs only isolated a single time, or which do not appear to make biological sense in the context of the starting bait, it may be helpful to consult the database to make sure the clone has not been reported by multiple additional groups.

ALTERNATE PROTOCOL

SCREENING FOR INTERACTION TRAP POSITIVES BY YEAST PLASMID RECOVERY

This protocol can be used instead of the rapid PCR screen (see Basic Protocol 4). See outline in Figure 20.7.9 for additional details. For materials see Basic Protocol 4. Also see UNIT 1.6 for additional reagents and equipment for preparation of plasmid DNA.

1. Isolate plasmid DNA from selected yeast clones (see Basic Protocol 4, steps 1 and 2) and transform into *E. coli*.

*A transformation protocol described in UNIT 13.11 can be used. Alternatively, a number of kits for yeast minipreps are commercially available (e.g., from Clontech and others). Some companies (e.g., Qbiogene; <http://www.qbiogene.com/services/two-hybrid.html>) will isolate plasmid from the yeast cells, transform, and amplify the plasmid in *E. coli* to produce a sequencing template.*

2. Select at least two bacterial clones for each yeast clone, and prepare a small quantity of plasmid DNA (UNIT 1.6) from each bacterial clone.

*If using a specialized Ap^R LexA-fusion bait plasmid in combination with Ap^R library plasmid, it will be necessary to select specifically for transformants containing a library plasmid by the ability of the yeast TRP1 gene to complement the *E. coli* trpC mutation in a KC8 strain (see UNIT 20.1).*

3. Analyze the isolated plasmids by restriction digest (*UNIT 3.1*), to identify redundant clones and to confirm that both isolates from the same yeast clone are identical.
4. Follow the instructions given above (see Basic Protocol 4, steps 9 to 14) as described, except transform with purified library plasmids, instead of a mixture of PCR product and digested library vector.

COLORIMETRIC ASSAY OF β -GALACTOSIDASE AND β -GLUCURONIDASE ACTIVITY BY AGAROSE OVERLAY

SUPPORT PROTOCOL 1

There are a variety of the standard assays used to assess β -galactosidase and β -glucuronidase activity in yeast, including colony overlay, colony lift (filter) assays, and growth on plates containing Xgal or Xgluc. Although popular, assays in which yeast are grown in the presence of substrate yield sometimes erratic results. One of the reasons for this is that expression of some combinations of bait and prey may result in greater permeability of yeast cells, reflected by the increased degree of uptake of Xgal from the surrounding media (yeast seed density may also play a role). In addition, Xgal plates are buffered to approach pH 7.0 (optimal for β -galactosidase activity), while yeast grow optimally on media of pH 5.5. Thus, yeast grow suboptimally under these conditions.

Overlay and colony lift (filter) assays avoid these issues since yeast are first grown under optimal conditions and then exposed to substrates following permeabilization of yeast. Of these techniques, the overlay technique provides the most consistent results in the hands of the authors, in part because of the avoidance of issues such as the sometimes uneven transfer of colonies to filters. Xgal and Xgluc overlay plates are quick and extremely sensitive, and normally will give a strong signal within an hour.

Materials

Plates containing spotted yeast colonies to be assayed

Chloroform

Xgal agarose: prepare 1% low-melting agarose in 100 mM KHPO₄, pH 7.0, and heat to boil; cool to ~60°C and add Xgal to 0.25 mg/ml; prepare fresh

Xgluc agarose: prepare as described for Xgal agarose, except replace Xgal with Xgluc

1. Gently overlay each plate containing spotted yeast colonies to be assayed with a minimal amount of chloroform (CHCl₃), pipetting slowly in from the side so as not to smear colonies. Leave colonies completely submerged for 5 min, but do not cover the dish with the lid.

CAUTION: CHCl₃ is quite toxic and should neither be inhaled nor come into contact with skin. Wear gloves and work in a chemical hood. Try to minimize the amount of CHCl₃ used, just enough to cover the colonies. Try to avoid contact with the walls of the plate, as the plastic will melt.

2. *Optional:* Briefly rinse the plates with another ~5 ml chloroform, drain, and allow to dry uncovered for another 5 min at 37°C or for 10 min in a hood.
3. Overlay the plate with ~10 ml Xgal or Xgluc agarose, making sure that all yeast spots are completely covered.

Plates will be chilled after CHCl₃ evaporation, so it will be difficult to spread <7 ml of top agarose.

4. Incubate plates at 30°C and monitor for color changes. It is generally useful to check the plates after 20 minutes, and again after 1 to 3 hours.

DETECTION OF BAIT PROTEIN EXPRESSION

The technique described here works well for the vast majority of baits. If it is not clear, however, whether the degradation of the protein occurs in the cell or in the test tube. A number of more rigorous (and time-consuming) protocols are available, where yeast lysates are prepared first (*UNIT 13.13*).

Materials

Master plates containing primary transformants of each bait (see Basic Protocol 1) and positive control

Selective medium (*UNIT 13.1*)

2× Laemmli sample buffer (*UNIT 20.1*)

Antibodies to LexA and cI (Invitrogen)

Additional reagents and equipment for minigel electrophoresis (*UNIT 2.5A*)

1. Transfer colonies from the master plate containing two primary transformants of each bait into the appropriate selective liquid media using sterile technique. Grow overnight cultures at 30°C with aeration. The next morning, dilute saturated cultures into fresh tubes containing 3 to 5 ml selective medium to a starting density of OD₆₀₀ ~0.15 and grow at 30°C with aeration.
2. When the culture reaches an OD₆₀₀ of ~0.45 to 0.7 (4 to 6 hr), transfer to a 1.5-ml microcentrifuge tube and microcentrifuge 3 to 5 min at full speed. Aspirate supernatant and resuspend pellet in 50 µl of 2× Laemmli sample buffer. Vortex and then boil samples for five min.
Such samples may be analyzed immediately or stably frozen at -70°C. Alternatively, the unboiled samples can be frozen at -70°C (stable for at least 4 to 6 months), and boiled before use.
3. Pulse centrifuge to pellet debris and chill on ice. Load 5 to 15 µl per well of a minigel (*UNIT 2.5A*) and analyze by immunoblotting following standard protocols. Bait should be visualized using antibody to LexA or cI which allows comparison of expression levels of the Bait protein under test with other standard bait proteins.

PREPARATION OF HIGH-QUALITY DIGESTED LIBRARY PLASMID

This control experiment is an indicator of the degree of digestion of the library plasmid. The background level of colonies transformed with digested empty library plasmid should be minimal. If the background is high, digestion of the empty library plasmid should be improved—e.g., by adding more enzyme(s) and/or increasing the digestion time. If transformation efficiency with digested library plasmid plus Raf1 PCR product with latone is better than with the digested empty plasmid alone by 5- to 20-fold, proceed to the next steps. Uncut library plasmid is a positive control for the transformation.

Additional Materials (also see *Basic Protocol 4*)

Empty library plasmid: pJG4-5

Restriction enzymes (e.g., *EcoRI* and *XhoI*; *UNIT 3.1*)

pJG4-5-Raf1 (Table 20.7.4)

pJG4-5-Krit1 (Table 20.7.4)

SKY 48 (Table 20.7.3)

Glu/CM -Trp dropout plates (*UNIT 13.1*)

1. Digest an empty library plasmid with two enzymes producing incompatible ends in the polylinker region (e.g., *EcoRI* and *XhoI*).

Purification is not necessary.

2. PCR amplify (*UNIT 15.1*) *Raf1* and *Krit1* sequences from pJG4-5-*Raf1* and pJG4-5-*Krit1* plasmids using the FP1 and FP2 primers. Analyze and purify the PCR product if necessary. PCR products from the control reactions above can be also reused in this step.
3. Transform SKY48 with the following:
 - 50 to 100 ng digested library plasmid
 - 50 to 100 ng digested library plasmid and 0.5 to 1 μ g *Raf1* control
 - PCR product
 - 50 to 100 ng uncut library plasmid

Save the digested library plasmid and the *Raf1* and *Krit1* PCR products.
4. Plate the transformations on Glu/CM –Trp dropout plates and incubate at 30°C until colonies grow (2 to 3 days). Store at 4°C.

COMMENTARY

Background Information

The dual bait reagents are built upon the interaction trap form of the two-hybrid system (*UNIT 20.1*; Gyuris et al., 1993) and allow the expression of two independent baits in one yeast cell. This system, therefore, facilitates the efficient isolation of proteins that interact with either one or both of the two bait proteins. Using this system, one can make a direct comparison of the binding partners within a family of proteins or dissect similarities/differences in binding partners of wild-type and mutant forms of a protein.

This system offers important advantages in mutagenesis and small molecule studies. For example, the products from random mutagenesis of a protein can be screened for the persistence of an interaction with one binding partner and the simultaneous disruption of an interaction with a second binding partner. This allows identification of a mutant protein that specifically disrupts binding of discrete partners while preserving other interactions, thus selecting against mutations that cause global conformational disruptions of the mutagenized protein. Similarly, this system should provide a powerful tool for studying the promotion or disruption of protein–protein interactions using peptide or small-molecule libraries. It may be possible to identify compounds that affect the interaction of a (disease-related) mutant form of a protein with a binding partner, but not a wild-type form of the protein with the same binding partner. The authors believe this system is well suited for discovery of specific interactors with individual members of families of related proteins as well as for high-throughput studies of factors affecting protein–protein associations.

In the dual bait system (Fig. 20.7.1, top), a LexA-fused bait interacts with a galactose-inducible activation domain-fused partner to induce the expression of two reporter genes under transcriptional control of LexA operator (*lexAop*) sites: plasmid-borne *LacZ* and integrated *LEU2*. Three additional components include a *cI*-fused alternate bait, a plasmid-borne *GusA* (β -glucuronidase) reporter gene cassette, and an integrated reporter gene, in which *cI* operators direct the expression of the *LYS2* gene. Since *cI* and LexA are similar in size (237 versus 202 amino acids) and structure, and use related amino-terminal helix-turn-helix domains to bind palindromic operator sites with similar K_M 's (reviewed in Serebriiskii et al., 1999), these two systems are well matched.

A key to fully utilizing the power of these systems is the development of reagents that extend their use to as many proteins as possible, while minimizing potential false or otherwise artifactual results positive (Serebriiskii et al., 2000; <http://www.fccc.edu/research/labs/golemis/InteractionTrapInWork.html>). Based on prior experience optimizing and characterizing functionality of the interaction trap (*UNIT 20.1*; Gyuris et al., 1993; Estojak et al., 1995), several points can be noted at which it is desirable to have flexibility. These include control of bait(s) expression level, inducible expression of bait(s), and range of reporter sensitivity (Fig. 20.7.1). In addition, convenience of use is improved by the availability of bait expression variants with diverse polylinkers, plasmids with alternative *E. coli* selectable markers to aid plasmid recovery from yeast, and antibodies allowing bait detection. A whole array of reagents has been developed that should allow modulation of these

Table 20.7.8 LexA-Responsive *LacZ* Reporter Plasmids

Plasmid names	Yeast selection	<i>E. coli</i> selection	No. operators ^a	Comments/description	Contact information
pMW112	<i>URA3</i>	Km ^R	8	Transcription of the <i>lacZ</i> gene is directed by <i>lexA</i> operators	RB ^b
pMW109	<i>URA3</i>	Km ^R	2	Transcription of the <i>lacZ</i> gene is directed by <i>lexA</i> operators	RB ^b
pMW111	<i>URA3</i>	Km ^R	1	Transcription of the <i>lacZ</i> gene is directed by <i>lexA</i> operators	RB ^b
pMW107	<i>URA3</i>	Cm ^R	8	Transcription of the <i>lacZ</i> gene is directed by <i>lexA</i> operators	RB ^b
pMW108	<i>URA3</i>	Cm ^R	2	Transcription of the <i>lacZ</i> gene is directed by <i>lexA</i> operators	RB ^b
pMW110	<i>URA3</i>	Cm ^R	1	Transcription of the <i>lacZ</i> gene is directed by <i>lexA</i> operators	RB ^b
pSH18-34	<i>URA3</i>	Ap ^R	8	Transcription of the <i>lacZ</i> gene is directed by <i>lexA</i> operators	Origene ^c , Invitrogen ^d
pJK103	<i>URA3</i>	Ap ^R	2	Transcription of the <i>lacZ</i> gene is directed by <i>lexA</i> operators	Origene ^c
pRB1840	<i>URA3</i>	Ap ^R	1	Transcription of the <i>lacZ</i> gene is directed by <i>lexA</i> operators	Origene ^c
pJK101 (control)	<i>URA3</i>	Ap ^R	2	The basal activity of <i>lacZ</i> gene is under control of 2 <i>lexA</i> operators; used to monitor bait binding to operator sequences (in repression assay)	Origene ^c
pGNG1	<i>URA3</i>	Ap ^R	4	<i>lacZ</i> gene is replaced by <i>GFP</i>	MoBiTec ^e
pLexAop-lucU	<i>URA3</i>	Ap ^R	8	<i>lacZ</i> gene is replaced by luciferase gene	A. Fujita ^f

^aThe most sensitive indicator plasmids for transcriptional activation have eight operators, intermediate reporters have two, and the most stringent reporters have only one.

^bE-mail at brent@molsci.org.

^cSee the Origene homepage at <http://www.origene.com>.

^dSee the Invitrogen Web-site at <http://www.invitrogen.com>.

^eSee the MoBiTec Web-site at <http://www.mobitec-germany.com>.

^fSee Fujita et al. (1999).

properties for the dual bait. These reagents, and compatible interaction trap reagents, are summarized in Table 20.7.1, Table 20.7.2, Table 20.7.8, and Table 20.7.9. Further details of plasmid construction and composition are available at the authors' Web-site (<http://www.fccc.edu/research/labs/golemis/InteractionTrapInWork.html>).

Selection markers

A limited number of convenient selectable auxotrophic markers are available for use in *S. cerevisiae*. Of the recently described dual bait strategies (Jiang and Carlson, 1996; Inouye et al., 1997; Xu et al., 1997), most achieve the use of two distinct baits by sacrificing one of the standard bait-responsive reporters (see Table 20.7.10). It has been well documented that the most artifact-free usage of two-hybrid systems requires the use of at least two-independent

reporters for each bait. As outlined above, an optimal dual bait strategy requires introduction of a significant number of expression or reporter components into yeast (e.g., two bait-expressing elements, an activation-domain expressing element, and four reporters). In the system presented here, the amino acid biosynthetic genes *LEU2* and *LYS2* are utilized as reporters, while *TRP1* is used to select for an activation-domain fusion plasmid. The remaining convenient markers are *HIS3*, *URA3*, and a drug resistance gene, typically either Zeo^R (conferring resistance to zeocin) or G418^R (conferring resistance to G418 or neomycin). In the original design of the system, the Zeo^R pGKS8 plasmid expressed both a cI-fused bait, and a *cIop-GusA* reporter, while *HIS3* was used to express a LexA-fused protein, and *URA3* for a *lexAop-LacZ* reporter (Serebriiskii et al., 1999). The advantage of this approach is that it

Table 20.7.9 *cI*-Responsive *GusA* Reporter Plasmids

Plasmid name	Yeast selection	<i>E. coli</i> selection	No. and type of operators ^a	Comment/description
PRG64, PRG64a, PRG64b	<i>URA3</i>	Km ^R	4 <i>cI</i>	<i>cI</i> operators direct transcription of the <i>gusA</i> gene.
PRG62	<i>URA3</i>	Km ^R	2 <i>cI</i>	<i>cI</i> operators direct transcription of the <i>gusA</i> gene
PRG61	<i>URA3</i>	Km ^R	1 <i>cI</i>	<i>cI</i> operators direct transcription of the <i>gusA</i> gene
pDR8 ^b	<i>URA3</i>	Km ^R	8 <i>lexA</i> , 3 <i>cI</i>	LexA-responsive <i>LacZ</i> reporter is the same as that in pMW112 (Table 20.7.8), while <i>cI</i> -responsive <i>gusA</i> reporter has lower background level than in pGKS8 (Table 20.7.2).

^aThe most sensitive indicator plasmids for transcriptional activation have 4 operators, intermediate reporters 2, and the most stringent reporters 1 operator.

^bAvailable from Invitrogen (<http://www.invitrogen.com>).

Table 20.7.10 Two Baits Systems and Their Applications

System	DBD	Reporter		Applications ^a				
		Auxotrophic	Colorimetric	D	M	A	L	P
Differential Interaction Trap	<i>Gal4</i>	URA3	None	+	+			
	<i>LexA</i>	None	LacZ					
Double two-hybrid	<i>Gal4</i>	HIS3	None	+	+			
	<i>LexA</i>	None	LacZ					
Differential two-hybrid	<i>Gal4</i>	HIS3	LacZ	+			+	
	<i>LexA</i>	URA3	None					
Two-bait	<i>Tet^R</i>	URA3	None	+		+		
	<i>LexA</i>	None	LacZ					
Dual bait	<i>cI</i>	LYS2	<i>GusA</i>	+	+	+	+	+
	<i>LexA</i>	LEU2	<i>LacZ</i>					

^aAbbreviations: D: interaction within defined set of proteins; M: mutational analysis; A: working with aptamer/peptides; L: library screening; P: pharmacological applications.

creates a single *cI*-bait/reporter unit on a marker that can be used with any two-hybrid system (e.g., GAL4-based; Fields and Song, 1989). A disadvantage is that one must reclon a bait to a new plasmid backbone in order to modulate sensitivity of response, (e.g., by utilizing a reporter with different numbers of operators). Recent modifications, therefore, provide an option in which each bait is expressed independently on separate plasmids, while *LacZ* and *GusA* reporters are present on a common backbone. The recommended set in this protocol includes G418^R vectors for *cI*-fusion expression to be used in combination with classic *HIS3* vectors suitable for expressing *LexA*-fused bait. At user's discretion, equivalent Zeo^R vectors can be used in place of G418^R. For use with Invitrogen's Hybrid Hunter system (which utilizes Zeo^R vectors for *LexA*-bait expression) a number of *HIS3 cI*-bait vectors are available.

Note, while it is formally possible to combine Zeo^R *LexA*-baits and G418^R *cI*-baits (Huang and Schreiber, 1997), the inherent exposure to multiple antibiotics may result in the reduced yeast viability.

Bait vectors

To date, a series of vectors suitable for expressing *LexA*-fusions have been developed allowing expression of proteins at high or low levels, or inducibly following induction by galactose (reviewed in UNIT 20.1). The authors have now developed a comparable series for *cI*-fusions (Table 20.7.3; Fig. 20.7.2; Serebriiskii et al., 2002). pGKS6 (Zeo^R) and pGBS9 (G418^R) utilize the same *ADHI* promoter as in GKS8 to express moderate levels of fusion proteins. In pGKS7 (Zeo^R) and pGBS10 (G418^R), the *ADHI* promoter has been modified to enhance levels of bait expression ap-

proximately five-fold. In pGMS11 (Zeo^R) and pGMS12 (G418^R), the *ADHI* promoter has been replaced with a *GALI* promoter, allowing inducible expression. These vectors, therefore, allow the use of the dual bait system with baits whose continuous expression is toxic to yeast.

Most of the cI-fusion vectors exist in versions with polylinkers of multiple reading frames, to accommodate convenient insertion of cDNAs to synthesize fusion proteins (Table 20.7.2).

Colorimetric reporter constructs

Variation of reporter sensitivity in two-hybrid reagents is most conveniently achieved by varying the number of cognate operator sites for the DBD-fusion protein (Estojak et al., 1995). A series of *cIop-GusA* reporters containing 1, 2, 3, or 4 *cI* operators was created and evaluated for activity in parallel with a comparable series of *lexAop-LacZ* reporters with 1, 2, or 8 operators (Serebriiskii et al., 2002). These reporters demonstrate a graded response based on operator number to the cognate bait-fusion construct, while failing to respond to the negative control bait. For *lexAop-LacZ* reporters, sensitivity increases incrementally with each addition of operator (Estojak et al., 1995). For *cIop-GusA* reporters, however, constructs containing a single operator are clearly least sensitive, while constructs with 2, 3, or 4 operators display similar sensitivity. This difference may reflect differences in properties of cooperative binding between the *LexA* and *cI* proteins. Significantly, *cIop-GusA* reporters with more than one operator displayed a weak background level of *GusA* activation, while quantitation of the *cIop-GusA* reporter activation with *cI-TA* versus negative control *LexA-TA* constructs indicated that the degree of amplification was >50-fold over background, indicating a suitable dynamic range of signal for most applications. The pLacGus plasmid is an example of a plasmid which combines both colorimetric reporters on a single backbone. It utilizes eight *lexA* operators to direct *LacZ* and three *cI* operators to direct *GusA*, and provides a well balanced readout for many general applications. Other plasmids, combining *LacZ* and *GusA* reporters of various sensitivity, can be easily constructed.

Auxotrophic reporter strains

Ideally, an auxotrophic reporter yeast strain will be suitable for genomic level applications (e.g., compatible with interaction mating) and possess modulated sensitivity (e.g., variable

operator number). To date, three strains with an expanded range of response for *lexAop-LEU2* are available. These strains demonstrate high (SKY48), moderate (SKY473), or low (SKY191) sensitivity to a *LexA*-fused activator, based on the number of *lexA* operators present. Although all three strains contain the same number of *cI* operators, SKY473 and SKY191 are fortuitously more sensitive to activation by a cI-fusion. Of these strains, SKY473 has the best balance of sensitivity to both *LexA*- and *cI*-fusions. It has the mating type opposite to that of other strains, and thus can be used in combination with them to set up interaction mating.

The dual bait system has been extensively tested in a variety of applications (see Table 20.7.10), and proved to be robust and versatile (Serebriiskii et al., 2002). It was first validated in analysis of the selective interactions of Raf, Ral-GDS, and Krit1 (AD-fused components) with related Ras family GTPases—i.e., Ras (cI-fused Bait 2) and Krev (*LexA*-fused Bait 1). Raf interacts preferentially with Ras, Krit1 with Krev, and RalGDS with both. The dual bait system successfully identifies correct interactions among these proteins, and these reagents are now used as a set of controls in the described protocol. The authors have also documented the ability of the system to sort high-affinity interactors for either *LexA*- or cI-fused baits from a pool of low-affinity interactors, and to identify mutants in an activation domain-fused prey that selectively alter affinity for discrete *LexA*- versus cI-fused partners (Serebriiskii et al., 1999; Reeder et al., 2001). The new reagents described here have been used effectively to study interactions between known sets of proteins (E. Benevolenskaya and W.Kaelin, pers. comm.), and have been utilized in a number of library screening applications. In these library screens, the two baits used to screen either a bait of interest and a nonspecific control (E. Kotova, pers. comm.), or a wild-type and a mutated form of the same bait (Serebriiskii et al., 2002). Both approaches yielded specific interactive partners that were either known to be biological or have been validated (by GST pulldown or co-immunoprecipitation). Baits used in these studies range from cytoplasmic serine/threonine regulatory kinases, morphoregulatory GTPases (e.g., Cdc42, Rac), or cell cycle regulators (e.g., Cdc6, wild-type/mutant RB protein). Screens have been performed by direct library transformation, mating, or both. Significantly, substan-

Table 20.7.11 Possible Modifications to Enhance LexA-Fused Bait Performance in Specific Applications

Response ^b	Problem ^a					
	Strongly activating	Weakly activating	Not transported to the nucleus or low expression level	Continuous expression of LexA-fusion is toxic to yeast	Bait protein requires unblocked amino-terminal end for function	Bait protein expressed at high levels, unstable or interacts promiscuously
Truncate/modify bait ^c	+	+	–	+	–	+?
Use more stringent strain/reporter combination ^d	–	+	–	–	–	–
Fuse to nuclear localization sequence pJK202	–	–	+	–	–	–
Put LexA-fused protein under GAL1-inducible promoter pGilda ^e	–	–	–	+	–	+?
Fuse LexA to the carboxy terminus of the bait pNLexA ^f	–	+?	–	+?	+	+?
Integrate bait, reduce concentration pEG202I ^g	–	+?	–	+?	–	+

^aPlus sign (+) indicates the response usually helps; plus sign followed by question mark (+?) indicates the response may help, and a minus sign (–) indicates the response will not help.

^bAll of the alternative LexA bait expression vectors remain on an Amp^R selection for bacteria. If using them as is, the investigator may need to use KC8 bacteria to isolate the library plasmid after a library screen.

^cIt may be necessary to subdivide bait into two or three overlapping constructs, each of which must be tested independently.

^dUse of very stringent interaction strains may eliminate detection of biologically relevant interactions.

^eCan no longer use GAL-dependence of reporter phenotype to indicate cDNA-dependent interaction.

^fGenerally, LexA poorly tolerates attachment of the N-terminal fusion domain. Only ~60% of constructs are expressed correctly.

^gReduced bait protein concentration may lead to reduced assay sensitivity.

tially overlapping sets of interactors have been obtained both through transformation and interaction mating (B. Spodik and T. Coleman, pers. comm.), supporting the comparable nature of these approaches. Other work has indicated that these reagents will be useful in pharmaceutical screening applications, as the presence of two baits provides a significant shield against artifactual results based on potential global dysregulation of transcription (Serebriiskii et al., 2002). As with other new technologies, and particularly gene discovery screens, it will be several years until the identified interactors and general use of the system will be completely established. However, at present, the assembled reagents appear to offer an efficient, manipulable set of tools to analyze specific protein-protein interactions.

Critical Parameters and Troubleshooting

To make an ideal bait, a protein of interest should satisfy the following four criteria:

1. It should have little or no intrinsic ability to activate transcription.
2. It should be expressed at reasonably high levels.
3. It should be nontoxic for yeast.
4. It must be able to enter the yeast nucleus and bind DNA.

The first three properties are analyzed in the first protocol (see Basic Protocol 1), so that problems can be identified and potentially corrected before screening. Suggested modifications are summarized in Table 20.7.11 and Table 20.7.12. Currently, there are more options to enhance the performance of LexA-fused baits, than the cI-fused. Therefore, a pro-

Table 20.7.12 Possible Modifications to Enhance cI-fused bait Performance in Specific Applications

Response	Problem ^a			
	Strongly activating	Weakly activating	Continuous expression of cI-fusion is toxic to yeast	Bait protein expressed at high levels or interacts promiscuously
Truncate or modify bait ^b	+	+?	+?	–
Use more stringent strain/reporter combination ^c	–	+	–	–
Put cI-fused protein under GAL1-inducible promoter pGMS12 ^d	–	–	+	+?
Use low-level expression vector pGBS9 ^e	–	–	+?	+

^aPlus sign (+) indicates the response usually helps, plus sign followed by question mark (+?) indicates the response may help, and a minus sign (–) indicates the response will not help.

^bIt may be necessary to subdivide bait into two or three overlapping constructs, each of which must be tested independently.

^cUse of very stringent interaction strains may eliminate detection of biologically relevant interactions.

^dCan no longer use GAL-dependence of reporter phenotype to indicate cDNA-dependent interaction.

^eReduced bait protein concentration may lead to reduced assay sensitivity.

tein to be used as a primary bait should be fused to LexA.

The inability of the bait to go to the nucleus cannot be detected as reliably as its activation properties. Although the repression assay (described in *UNIT20.1*) was developed to assess this property, the current trend is to proceed with the screening even if the results of this assay are negative. In addition, this assay is not available for cI-fused baits, and therefore it is omitted in this protocol. However, a researcher should be aware of this possibility, and avoid (or truncate) proteins that have extensive transmembrane domains.

In screening the library, it is very important to obtain it for the tissue source in which the bait protein(s) are known to be biologically relevant. A second prerequisite is to screen a sufficient number of clones to fully represent the library. This latter goal is achieved in two steps. First, it requires an efficient transformation of the library in *S. cerevisiae*. In contrast to *E. coli*, the maximum efficiency of transformation for yeast is up to 1×10^5 cfu/ μ g DNA, so it is extremely important to optimize the transformation procedure prior to attempting library transformation. At least 1×10^6 clones should be obtained for a cDNA library derived from a genome of mammalian complexity. A second step requires an efficient mating of the pretransformed library strain with the bait strain. The number of diploids to obtain depends on the number of clones in the library and the number of primary yeast transformants. The objective is to screen a number of transfor-

ants at least equivalent to the size of the library; if the size of the library was larger than the number of transformants, then all of the yeast transformants should be screened. Under the optimal conditions, ~10% of the library strain cells will mate with the bait cells, forming diploids. Thus, a complete screen of 1×10^7 library transformants requires a single mating with about 1×10^8 cfu pretransformed library strain (and twice as much bait strain).

Colony formation on the selection plates is one of the most variable parts of the screening. For strong interactors, colonies will grow up in 2 days on –Leu plates or 3 days on –Lys plates. However, new colonies will continue to appear for up to one week if the plates are left in the incubator. Growth speed depends on many factors, including the expression levels of the library-encoded protein and the strength of interaction, so not necessarily the most rapidly growing colonies are more biologically relevant to the bait protein.

Dual bait provides a “built-in” control for the specificity of interaction. In addition, second confirmation of positives described in this protocol will test the isolated preys against two more unrelated baits. However, it is generally informative to retest positive clones on more than one bait protein. If possible, library-derived clones should be tested against a set of proteins ranging from almost identical to the original bait to clearly unrelated proteins. If the bait used in the screen showed weak transcriptional activity, it is expected to have greater difficulties with false positive background. In

this case it is advisable to choose a nonspecific control bait that can weakly activate transcription on its own in order to match the background level of transcription activation.

Anticipated Results

Depending on the properties of the bait protein(s) and the type of selection applied, screening a library of $\sim 1 \times 10^6$ primary transformants may yield from zero to hundreds of specific interactors.

Time Considerations

In general, it is good practice to move quickly through the steps of the described protocol. While plasmids will be retained in yeast for a long time, expressed protein levels will gradually drop, making the outcome of the whole screening less predictable. Under the favorable conditions, it will take about 3 weeks to complete the screening, as is outlined in Figure 20.7.2.

Literature Cited

- Brown, M.A. and MacGillivray, R.T. 1997. Vectors for expressing proteins at the amino-terminus of an activation domain for use in the yeast two-hybrid system. *Anal. Biochem.* 247:451-452
- Estojak, J., Brent, R., and Golemis, E.A. 1995. Correlation of two-hybrid affinity data with *in vitro* measurements. *Mol. Cell. Biol.* 15:5820-5829.
- Fields, S. and Song, O. 1989. A novel genetic system to detect protein-protein interaction. *Nature* 340:245-246.
- Fujita, A., Tonouchi, A., Hiroko, T., Inose, F., Nagashima, T., Satoh, R., and Tanaka, S. 1999. Hsl7p, a negative regulator of Ste20p protein kinase in the *Saccharomyces cerevisiae* filamentous growth-signaling pathway. *Proc. Natl. Acad. Sci. U.S.A.* 96:8522-8527.
- Gyuris, J., Golemis, E.A., Chertkov, H., and Brent, R. 1993. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* 75:791-803.
- Huang, J. and Schreiber, S.L. 1997. A yeast genetic system for selecting small molecule inhibitors of protein-protein interactions in nanodroplets. *Proc. Natl. Acad. Sci. U.S.A.* 94:13396-13401.
- Inouye, C., Dhillon, N., Durfee, T., Zambryski, P.C., and Thorner, J. 1997. Mutational analysis of STE5 in the yeast *Saccharomyces cerevisiae*: Application of a differential interaction trap assay for examining protein-protein interactions. *Genetics* 147:479-492.
- Jiang, R. and Carlson, M. 1996. Glucose regulates protein interactions within the yeast SNF1 protein kinase complex. *Genes Dev.* 10:3105-3115.
- Ma, H., Kunes, S., Schatz, P.J., and Botstein, D. 1987. Plasmid construction by homologous recombination in yeast. *Gene* 58:201-216.
- Petermann, R., Mossier, B.M., Aryee, D.N., and Kovar, H. 1998. A recombination-based method to rapidly assess specificity of two-hybrid clones in yeast. *Nucl. Acids Res.* 26: 2252-2253.
- Reeder, M.K., Serebriiskii, I.G., Golemis, E.A., and Chernoff, J. 2001. Analysis of small GTPase signaling pathways using Pak1 mutants that selectively couple to Cdc42. *J. Biol. Chem.* 276:40606-40613.
- Serebriiskii, I., Khazak, V., and Golemis, E.A. 1999. A two-hybrid dual bait system to discriminate specificity of protein interactions. *J. Biol. Chem.* 274:17080-17087.
- Serebriiskii, I., Estojak, J., Berman, M., and Golemis, E.A. 2000. Approaches to detecting two-hybrid false positives. *Biotechniques* 28:328-336.
- Serebriiskii, I.G., Khazak, V., and Golemis, E.A. 2001. Redefinition of the yeast two-hybrid system in dialogue with changing priorities in biological research. *BioTechniques* 30:634-655.
- Serebriiskii, I.G., Mitina, O., Pugacheva, E., Benevolenskaya, E., Kotova, E., Toby, G.G., Khazak, V., Kaelin, W.G., Chernoff, J. and Golemis, E.A. 2002. Detection of peptides, proteins, and drugs that selectively interact with protein targets. *Genome Res.* In press.
- Xu, C.W., Mendelsohn, A.R. and Brent, R. 1997. Cells that register logical relationships among proteins. *Proc. Natl. Acad. Sci. U.S.A.* 94:12473-12478.

Internet Resource

<http://www.fccc.edu/research/labs/golemis/InteractionTrapInWork.html>

Analysis of the two-hybrid usage; database of false positives detected in two-hybrid experiments; database of available libraries, strains and specialized vectors (maps, sequences) for Interaction Trap/Dual Bait system; protocols related to the two-hybrid screening and β -galactosidase assays in yeast

Contributed by Elena Kotova, Ilya Serebriiskii, and Thomas Coleman
Fox Chase Cancer Center
Philadelphia, Pennsylvania

Interaction Trap/Two-Hybrid System to Identify Loss-of-Interaction Mutant Proteins

UNIT 20.8

To understand how a specific protein functions, it can be useful to identify mutants that no longer interact with a known partner. This is especially useful for proteins that interact with multiple partners, because it allows dissection of the importance of specific protein-protein interactions. By taking advantage of the ease of the interaction trap/two-hybrid system (UNIT 20.1; Gyuris et al., 1993) and of PCR mutagenesis (UNIT 8.3; Cadwell and Joyce, 1992), it becomes possible to rapidly make and identify proteins that no longer bind to a specific partner protein. These mutants usually result from a single missense mutation, which can easily be identified.

IDENTIFYING LOSS-OF-INTERACTION MUTANT PROTEINS

BASIC
PROTOCOL

In this protocol, two-hybrid analysis in *S. cerevisiae* is used to confirm interaction of a bait, a cDNA encoding a protein of interest fused to a DNA binding domain, with a prey, a cDNA encoding an interacting protein fused to a transcription activation domain (see UNIT 20.1 for more details). The open reading frame (ORF) from the cDNA encoding the prey protein is PCR-amplified with primers designed to remove its termination codon to prepare it for fusion with green fluorescent protein (GFP; Fig. 20.8.1, top left). The ORF is mutagenized by Mn-doped PCR—also known as error-prone (EP)-PCR—to introduce an average of one mutation per DNA molecule (Fig. 20.8.1, middle left), and then PCR-ligated at its 3' end to DNA encoding GFP (Fig. 20.8.1, top right and middle). GFP expression is used as a reporter to ensure that the mutagenesis has not introduced a stop codon into the ORF. The resulting construct is then introduced back into a two-hybrid prey plasmid (pJG4-5) by recombination in *S. cerevisiae* (Fig. 20.8.1, bottom) or by standard cloning techniques (see Alternate Protocol). This mutagenized prey minilibrary is then screened for preys that do not interact with their former partner protein expressed as bait. Colonies of cells which fluoresce green and that carry missense mutant proteins that do not interact with a specific bait protein in plasmid pEG202, as shown by two-hybrid analysis, are chosen for further study. DNA derived from these colonies is sequenced and corresponding single-point mutations are introduced into the original prey plasmid by directed PCR mutagenesis. The noninteracting phenotype is then confirmed by two-hybrid analysis.

Materials

- Genes for proteins of interest
- Two-hybrid interaction trap plasmids (Table 20.1.1):
 - Prey plasmid pJG4-5
 - Bait plasmid pEG202
 - LacZ reporter plasmid pSH18-34
- Two-hybrid interaction trap yeast strain EGY48 (UNIT 20.1)
- CM dropout plates (UNIT 13.1) substituted with 2% glucose (Glc), 2% galactose (Gal), and/or 1% raffinose (Raff) as indicated:
 - Glc/CM –Ura
 - Glc/CM –Ura –His
 - Glc/CM –Ura –His, –Trp
 - Glc/CM/Xgal –Ura –His –Trp
 - Gal/Raff/CM/20 µg/ml Xgal in Z buffer –Ura, –His, –Trp: prepare as described for CM plates (UNIT 13.1) using a 1 mg/ml Xgal stock (20 µg/ml final) in Z buffer (UNIT 13.6)

Analysis of
Protein
Interactions

20.8.1

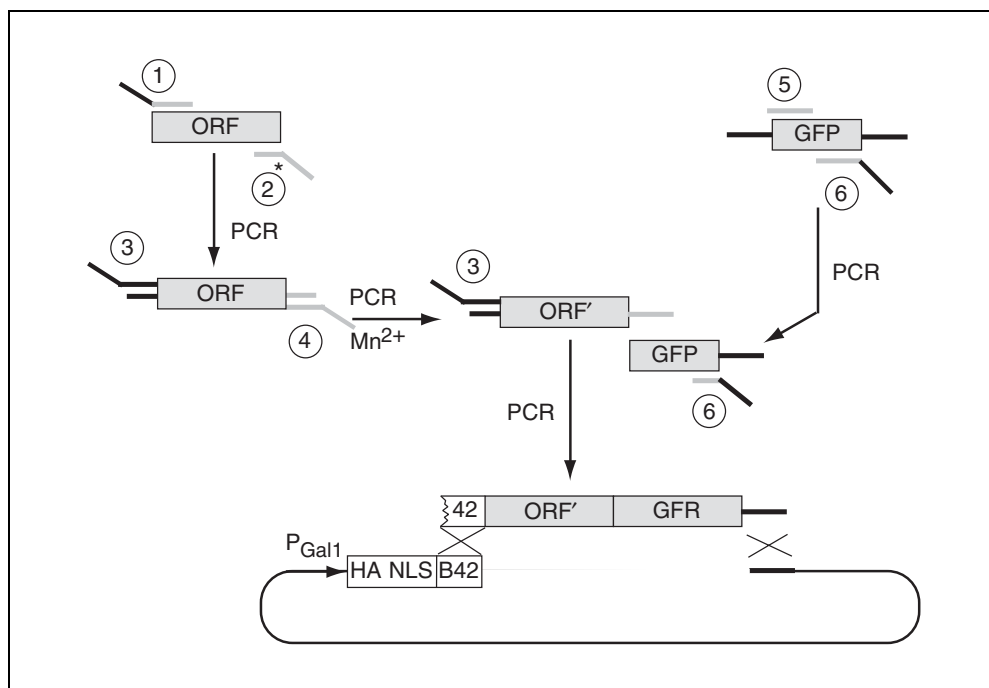


Figure 20.8.1 Interaction trap/two-hybrid system loss-of-interaction mutant proteins. Circled numbers indicate primers (see Basic Protocol). The asterisk (*) on primer 2 indicates that the stop codon has been deleted from the ORF.

- 2.5 mM (each) dNTP stock solution (UNIT 15.1)
- 2.5 U/μl high-fidelity PCR-compatible DNA polymerase (e.g., *Pfu*Turbo, Kod) and 10× buffer
- 1 U/μl *Taq* DNA polymerase (UNIT 3.5) and 10× Mg²⁺-free reaction buffer
- 2.5 mM MgCl₂
- 50 mM MnCl₂
- Plasmid carrying *EGFP* (e.g., pEGFP; BD Biosciences)
- CM –Trp liquid medium (UNIT 13.1)
- TrpC *E. coli* strain KC8 (*hsdR*, *leuB600*, *trpC9830*, *pyrF::Tn5*, *hisB463*, *lacΔX74*, *strA*, *galU*, *galK*) (UNIT 20.1; constructed by K. Struhl and available from R. Brent), competent
- M9 minimal plates (UNIT 1.1) containing 100 μg/ml amp, 40 μg/ml Leu, 40 μg/ml His, and 40 μg/ml Ura
- Commercial miniprep kit (e.g., Qiagen; optional)
- 30°C incubator
- Long-wavelength UV lamp, or fluorescence microscope equipped with fluorescein or GFP filter set
- Additional reagents and equipment for subcloning into plasmid vectors (UNIT 3.16), lithium acetate transformation (UNIT 13.7), PCR (UNIT 15.1), agarose gel electrophoresis (UNIT 2.5A), restriction digestion (UNIT 3.1), replica plating using velvet filters (UNITS 1.3 & 13.2), interaction trap (UNIT 20.1), purification of DNA by glass beads (UNIT 3.11), DNA minipreps (UNIT 2.1; optional), bacterial transformation (UNIT 1.8), sequencing DNA by dideoxy chain termination (UNIT 7.4), and site-directed PCR mutagenesis (UNIT 8.5)

Confirm interaction of bait protein with prey protein by two-hybrid analysis

Once interaction has been confirmed, the experiment can be started with step 7.

1. Insert the gene (*UNIT 3.16*) for the first protein of interest into the prey plasmid pJG4-5.

This protein will later be mutagenized (see step 10) so that it no longer interacts with its protein partner.

2. Insert the gene (*UNIT 3.16*) for the protein partner into bait plasmid pEG202.
3. Transform EGY48 with ~1 µg lacZ reporter plasmid pSH18-34 by the lithium acetate yeast protocol (*UNIT 13.7*) and select for yeast that grow on CM –Ura dropout medium.
4. Transform 1 µg of both plasmids simultaneously into interaction-trap strain EGY48 carrying pSH18-34. Select for yeast carrying all three plasmids by growing on Glc/CM –Ura –His –Trp triple dropout plates.

Selection for pSH18-34 is achieved by omitting uracil (–Ura), pEG202-derived bait plasmid by omitting histidine (–His), and pJG4-5-derived prey plasmid by omitting tryptophan (–Trp).

5. As a negative control, transform EGY48/pSH18-34 (step 3) with prey vector pJG4-5 and bait vector pEG202 and select colonies on Glc/CM –Ura –His –Trp triple dropout plates.
6. Pick colonies and replica plate onto the following fresh plates:

Glc/CM –Ura –His –Trp
Glc/CM/Xgal –Ura –His –Trp
Gal/Raff/CM/Xgal –Ura –His –Trp.

Colonies carrying interacting bait and prey should appear blue on Xgal/Gal/Raff plates and white on Xgal/Glc plates. Compare with the pJG4-5/pEG202 negative control in which colonies should appear white on all plates.

If the interaction is very strong or if the bait self-activates transcription to some extent (see Critical Parameters and Troubleshooting), the colonies may appear somewhat blue on the control Xgal/Glc plates. If the amount of blue on Xgal/Glc plates is substantial and similar to that seen on Xgal/Raff plates, it may not be possible to use this particular bait in this protocol. However, there are several ways to reduce interaction. For example, one way to reduce background is to use a less sensitive lacZ reporter plasmid, such as pJK103 (see Table 20.1.1). For a detailed discussion on how to reduce background from the bait self-activating the lacZ reporter, see UNIT 20.1, Commentary.

Remove the termination codon from the prey ORF

7. Synthesize the following PCR primers (Fig. 20.8.1) to remove the 3' translation termination codon from the prey by directed PCR mutagenesis:

- a. *Primer 1*: Design to contain 5' sequence homologous to the B42 region of pJG4-5 (for later homologous recombination) fused to ~21 bp of the prey ORF beginning at the initiation codon (ATG).

An example sequence is 5'-GATTATGCCTCTCCCGAATTC-ATG-ORF-3', where the ORF is 21 bp directly downstream of the ORFs initiation codon (ATG).

It is also possible to design an alternate primer 1 for later conventional cloning by including an appropriate restriction site such as EcoRI instead of sequence homologous to the B42 region of pJG4-5 (e.g., CTCTCTCTCTCTGAATTC-ORF-3') where the ORF is 21 bp directly downstream of the ORFs ATG.

- b. *Primer 2*: Design to contain 5' sequence homologous to the reverse complement of the sequence homologous to the 5' end of *EGFP* fused to the reverse complement of the 3' end of the ORF just upstream of, but not including, the termination codon.

An example sequence is 5'-CTCCTCGCCCTTGCTCACCAT-ORF-3', where ORF is the reverse complement of the sequence 21 bp upstream of the termination codon.

8. Set up the termination-codon-removal PCR reaction:

1 ng prey plasmid carrying ORF
0.2 μ M primer 1
0.2 μ M primer 2
2 μ l 10 \times (2.5 mM) dNTPs
2.5 μ l 10 \times polymerase buffer
1 μ l 2.5 units/ μ l high-fidelity PCR-compatible DNA polymerase (e.g., *Pfu-Turbo*; add last).

Use distilled water to bring the final reaction volume to 25 μ l.

9. Perform PCR (*UNIT 15.1*) using the following cycling conditions:

Initial step:	30 sec	94°C
20 cycles:	30 sec	94°C
	60 sec	56°C
	60 sec	72°C
Final step:	10 min	72°C.

10. Check for appropriate-sized PCR fragment on a 1% agarose gel with TBE as the buffer (*UNIT 2.5A*).

Mutagenize the prey ORF

11. Synthesize the following primers for mutagenesis of the prey ORF, with termination codon removed, by Mn-doped PCR (Fig. 20.8.1):

- a. *Primer 3*: Design to contain sequences homologous to the 5' part of primer 1 (step 7; i.e., homologous to the B42 region of pJG4-5), but without any sequence homologous to the prey ORF.

This primer is designed to extend the homology with pJG4-5 to 40 bp for later homologous recombination (see below)—e.g., 5'-CTACCCTTATGATGTGCCAGATTATGCCTCTCCCGAATTC-3'.

If the minilibrary (see below) is to be made by conventional cloning, primer 3 can be designed to be homologous only to the 5' region of alternate primer 1 (step 7), such that it does not contain any sequence homologous to the ORF (e.g., 5'-CTCTCTCTCTCTGAATTC-3').

- b. *Primer 4*: Design to be identical to the 3' section of primer 2—i.e., homologous to the reverse complement of the 5' end of the *EGFP* gene (step 7; e.g., 5'-CTCCTCGCCCTTGCTCACCAT-3').

If the region of interaction is known and defined by a contiguous region on the prey protein, then it may be useful to restrict the mutagenesis to this region by appropriate primer design.

Mn-doped PCR mutagenesis has been described by Lin-Goerke et al. (1997), and Cadwell and Joyce (1992).

12. Set up the Mn-doped PCR reaction as follows:

- 16 μ l 2.5 mM (each) dNTPs
- 2 μ l primer 3 (5 μ M final concentration)
- 2 μ l primer 4 (5 μ M final concentration)
- 5 μ l 10 \times Mg²⁺-free *Taq* reaction buffer
- 5 μ l 2.5 mM MgCl₂
- 2.5 μ l 50 mM MnCl₂
- 1 μ l 10 ng of 1.3-kb DNA fragment (step 10)
- 1 μ l 1 U/ μ l *Taq* DNA polymerase
- 15.5 μ l H₂O.

Mutagenesis occurs by increasing the misincorporation rate of Taq DNA polymerase; it is often referred to as error-prone PCR (see UNIT 8.3 for a more in depth discussion of this methodology).

13. Perform PCR (UNIT 15.1) using the following cycling conditions:

- | | | |
|---------------|--------|-------|
| Initial step: | 2 min | 94°C |
| 36 cycles: | 30 sec | 94°C |
| | 1 min | 55°C |
| | 1 min | 72°C |
| Final step: | 10 min | 72°C. |

Mn-doped PCR mutagenesis needs to be optimized in order to obtain an average of one mutation per DNA molecule. Optimization can be performed by varying the concentration of Mn²⁺ and Mg²⁺ in the protocol. It should be noted that even mutagenesis that results in as many as ten mutations per DNA molecule can be successfully used in this protocol, as the key mutation can be identified by sequencing followed by individually introducing mutations into the prey ORF and checking whether the resulting prey does not interact with the bait.

While it can be useful to vary the both the Mn²⁺ and Mg²⁺ concentrations, the authors have had some success varying just the Mg²⁺ concentration from 1 mM to 7 mM. It is suggested that several reactions be performed in parallel using various Mg²⁺ concentrations. Efficiency of mutagenesis can be later determined by sequencing resulting clones from a minilibrary (see below).

Fuse prey ORF with GFP

14. Synthesize the following primers (Fig. 20.8.1) for amplifying the full-length *EGFP* gene:

- a. *Primer 5*: Design to contain a sequence homologous to the 5' end of the *EGFP* gene.

An example sequence is 5'-ATGGTGAGCAAGGGCGAGGAG-3'

- b. *Primer 6*: Design a sequence where the 5' end is homologous to the reverse complement of the polylinker region of pJG4-5 fused to the reverse complement of the 3' end of the *EGFP* gene.

The 5' end is prepared in this fashion in order to facilitate homologous recombination (see below). An example sequence is 5'-TTGACCAAACCTCTGGCGAAGAAGTCCAAAGCT-TCTCGAGTTACTTGTACAGCTCGTCCATG-3'.

*An alternate primer 6 for conventional cloning (see Alternate Protocol) would consist of sequence creating a compatible site for cloning—e.g., *Xho*I fused to the reverse complement of the 3' end of the *EGFP* gene (5'-CTCTCTCTCTCGAGTTACTTGTACAGCTCGTCCATG-3').*

15. Set up the reaction to amplify *EGFP*:

1 ng plasmid carrying *EGFP* (e.g., pEGFP)
1 μ l primer 5 (1 μ M final)
1 μ l primer 6 (1 μ M final)
2 μ l 2.5 mM dNTPs
2.5 μ l 10 \times high-fidelity PCR-compatible DNA polymerase buffer
1 μ l 2.5 U/ μ l high-fidelity PCR-compatible DNA polymerase (e.g., *Pfu*-
Turbo; add last).
H₂O to 25 μ l.

16. Perform PCR (*UNIT 15.1*) using the following cycling conditions:

Initial step:	2 min	94°C
19 cycles:	30 sec	94°C
	1 min	56°C
	1 min	72°C
Final step:	10 min	72°C.

17. Set up PCR reaction to fuse mutated prey ORF with *EGFP*:

1 ng plasmid carrying *EGFP* (e.g., pEGFP)
1 ng mutated prey ORF DNA (step 13)
1 μ M primer 3
1 μ M primer 6
2 μ l 10 \times (2.5 mM) dNTPs
2.5 μ l 10 \times *Pfu*Turbo buffer
1 μ l *Pfu*Turbo (2.5 units/ μ l; add last)
H₂O to 50 μ l.

This reaction can be scaled up if necessary.

Primers 2 and 4 have previously been used to introduce sequence into the prey ORF such that the mutated ORF and GFP now overlap. This allows fusion of the two pieces of DNA by PCR amplification using primers that are homologous to 5' end of the mutated ORF DNA (primer 3) and the 3' end of the EGFP gene (primer 6; see Fig. 20.8.1, middle).

18. Perform PCR (*UNIT 15.1*) using the following cycling conditions:

Initial step:	2 min	94°C
19 cycles:	30 sec	94°C
	1 min	56°C
	1 min	72°C
Final step:	10 min	72°C.

At least 1 μ g DNA is needed for the minilibrary construction if using the yeast recombination method, or 100 ng DNA if using conventional cloning (see Alternate Protocol). The number of cycles or volume of the reaction should be adjusted accordingly. The yield can be checked by samples removed for gel electrophoresis or by real-time PCR.

Create prey minilibrary by recombination in yeast

See Alternate Protocol if constructing minilibrary by conventional means.

19. Make competent EGY48 carrying pSH18-34 and appropriate bait in pEG202 using the lithium acetate method (*UNIT 13.7*).

20. Plate and select on Glc/CM –Ura –His.

21. Add 1 µg mutated ORF-EGFP DNA (step 18) and 1 µg pJG4-5 linearized with *EcoRI* (UNIT 3.1) to 200 µl competent yeast (step 19).
22. Incubate 30 min at 30°C.
23. Heat shock 2 min at 37°C.
24. Plate on Glc/CM –Ura –His –Trp plates, as appropriate.
25. Determine the total number of independent transformants by counting colonies.

At least 1000 independent transformants will be needed for a successful screen. A lesser amount may be satisfactory if the library contains multiple mutations per DNA molecule.

Screen library for loss-of-interaction prey mutants

26. Replica plate the yeast transformed with the prey minilibrary onto Gal/Raff/CM/Xgal –Ura –His –Trp and Glc/CM/Xgal –Ura –His –Trp using velvet filters (UNIT 1.3 & 13.2), or by patching colonies onto the Xgal-containing plates (UNIT 20.1).
27. Incubate 24 to 48 hr at 30°C.
28. Screen the library for the following two features at the same time:
 - a. Search for mutant preys that no longer interact with the original bait by looking for white colonies on Gal/Raff/CM/Xgal –Ura –His –Trp plates.
 - b. Hunt for clones with uninterrupted ORFs by looking for colonies that remain green when viewed under a long-wavelength UV lamp or, more desirably, a fluorescence microscope equipped with a fluorescein or GFP filter set.

Green fluorescent or white colonies represent candidate loss-of-interaction mutants and should be picked for further analysis.

Identify resulting candidate mutations by sequencing

29. Grow candidate colony overnight in 5 ml CM –Trp medium.
30. Microcentrifuge yeast 20 sec to pellet.
31. Use the glass bead method (UNIT 13.11) to isolate DNA from candidates.
32. Transform *E. coli* trp⁻ strain KC8 (*hsdR*, *leuB600*, *trpC9830*, *pyrF::Tn5*, *hisB463*, *lacΔX74*, *strA*, *galU*, *galK*) with DNA (UNIT 1.8) and plate on M9 minimal plates containing 100 µg/ml amp, 40 µg/ml Leu, 40 µg/ml His, and 40 µg/ml Ura.

This works because there is enough expression of the TRP1 gene in E. coli to complement the trpC mutation in KC8.
33. Perform DNA minipreps (UNITS 1.7 or 2.1B) or use a commercial kit (e.g., Qiagen miniprep).
34. Sequence DNA by dideoxy chain termination method (UNIT 7.4).

Confirm candidate mutations by directed PCR mutagenesis and two-hybrid analysis

35. Analyze the results of sequencing the candidate DNAs:
 - a. *For single mutation:* If a candidate clone has only one missense mutation, that mutation must have resulted in the loss-of-interaction phenotype (i.e., no further analysis is necessary).
 - b. *For multiple mutations:* If a candidate clone has multiple mutations, introduce these mutations into the prey individually to determine which one actually causes the loss-of-interaction phenotype (see below).

**ALTERNATE
PROTOCOL**

36. Use site-directed PCR mutagenesis (*UNIT 8.5*) to introduce single specific mutations into the original prey ORF.
37. Individually transform preys into EGY48 carrying pSH18-34 and the original bait cloned into pEG202 (steps 19 and 20).
38. Replica-plate colonies and screen candidates for loss-of-interaction prey mutants as described above (steps 21 to 32).

CREATE PREY MINILIBRARY BY CONVENTIONAL CLONING

The following steps describe preparing minilibraries by conventional cloning. These steps are used in place of the minilibrary construction procedure described above (see Basic Protocol, steps 19 to 25).

Additional Materials (*also see Basic Protocol*)

T4 DNA ligase (*UNIT 3.14*)

LB medium and plates containing amp (*UNITS 1.1 & 1.4*)

Additional reagents and equipment for gel purification of DNA fragments (*UNIT 2.6*), multiple endonuclease digestion (*UNIT 3.2*), DNA ligation (*UNITS 3.14 & 3.16*)

19. Gel purify both cut fragments (*UNIT 2.6*). Digest mutated ORF-EGFP DNA (see Basic Protocol, step 18) and prey plasmid pJG4-5 with *EcoRI* and *XhoI* (see *UNITS 3.1 & 3.2*).

This presumes that there are no internal EcoRI and XhoI sites in the prey ORF. If these sites are present in the ORF, then alternates must be chosen that are compatible with the unique sites in the polylinker of pJG4-5.

20. Gel purify both cut fragments (*UNIT 2.6*). Ligate 100 ng cut DNA using T4 DNA ligase (*UNIT 3.14 & 3.16*).
21. Transform *E. coli* with library DNA (*UNIT 1.8*). Plate serial 1:10 dilutions of *E. coli* transformed with the library DNA onto LB plates containing amp to determine the number of independent transformants.
22. Amplify library by diluting an aliquot 1:10 into 100 ml LB medium containing amp and growing overnight at 37°C.
23. Make high-quality prey DNA from culture using a maxiprep method (*UNIT 1.7*) or commercial kit.
24. Transform competent EGY48 carrying pSH18-34 and appropriate bait in pEG202 (see Basic Protocol, step 19) with prey DNA by lithium acetate method (*UNIT 13.7*). Plate on Glc/CM –Ura –His –Trp plates, as appropriate.
25. Determine the total number of independent transformants by counting colonies.

At least 1000 independent transformants will be needed for a successful screen. A lesser amount may be satisfactory if the library contains multiple mutations per DNA molecule.

COMMENTARY

Background Information

The methodology described in this unit was used to isolate mutants of caspase 2 that no longer interacted with cyclin D3 (Mendelsohn et al., 2002). The basic techniques involved, the two-hybrid system (*UNIT 20.1*), and PCR mu-

tagenesis (*UNITS 8.3 & 8.5*) are established techniques and are easy to perform. The basic advantage of this methodology is the addition of EGFP as a reporter to identify clones that have not been truncated by termination codons introduced by PCR mutagenesis. EGFP also

serves to eliminate mutants that destabilize the prey protein in yeast.

Critical Parameters and Troubleshooting

There are two critical parameters that must be considered to use this methodology successfully. The first is that the interacting proteins must be well-behaved in the interaction trap/two-hybrid system. If the bait is self-activating, or if the interaction is so strong as to activate the lacZ reporter on the glucose plates that should greatly repress the synthesis of the prey, then steps must be taken to reduce the sensitivity of the assay. *UNIT 20.1* discusses this in more depth. The bottom line is that some proteins are not amenable to being used as baits in the two-hybrid system. Therefore, perhaps the simplest way to resolve the problem of a self-activating bait is to switch the bait and prey proteins. It is quite possible that when the prey protein is used as a bait, it will not self-activate transcription.

The second critical parameter is the condition of the *in vitro* mutagenesis. Given a specific rate of mutagenesis, the length of the ORF to be mutated is an important consideration. The same rate of mutagenesis that would cause an average of one mutation in a 100-bp ORF will cause an average of ten mutations in a 1000-bp ORF. Critically, there can be great variation from published ideal conditions for error-prone (EP)-PCR. It is strongly recommended that the magnesium and/or manganese concentration(s) be titrated to obtain the desired rate of mutagenesis on the specific template chosen. Another factor to consider is that Mn^{2+} ions may associate with the DNA and interfere with subsequent DNA manipulations. After the completion of the reaction, it may be useful to add a large excess of Mg^{2+} ions to compete out the Mn^{2+} ions, or EDTA to bind up the Mn^{2+} ions, or to subsequently purify DNA by glass-powder isolation from an agarose gel.

One factor to consider is that the screen described here for intact ORFs could be made into a selection by substituting an auxotrophic yeast marker such as *Lys2* for *EGFP* in the PCR fusion step and then transforming a *Lys2* auxotrophic mutant derived from EGY48. In this case, selection could be made on CM dropout –Ura –His –Trp –Lys plates. The absence of lysine would select for the intact ORF. Protein interaction itself is still probably best measured by screening for white colonies on plates containing Xgal, galactose, and raffinose.

Another factor to think about is that it may prove valuable to test the mutated prey against a battery of known interacting proteins as baits in the two-hybrid system. This may be especially useful if the mutated proteins are to be characterized functionally in their native biological context. Any change of phenotype would likely be the result of the loss of a specific protein-protein interaction. The proteins that no longer interact would thus represent potential key players in causing the observed phenotype. In this regard, it is important to note that it may not be the loss of the specific protein interaction selected in this procedure that has true functional significance, but rather the simultaneous loss of interaction of the mutated protein with another interacting protein that would be critical.

Furthermore, an interesting extension of this technology would be to perform a series of experiments to prove the relevance of a particular protein-protein interaction by (1) isolating a loss-of-function version of the prey protein, (2) associating this mutant with a change of phenotype, (3) isolating a mutant version of the partner protein that does interact with the initial mutant, and then (4) showing that the presence of both mutant proteins restores the original phenotype. The methodology described here could easily be modified to this task by the following modification. After isolating the loss-of-function prey mutant, insert it into the pEG202 bait vector, then insert the original bait into the pJG4-5 prey vector and follow the procedure described, but screen for blue colonies that fluoresce green. In other words, look for full-length mutant proteins that interact with the original mutant protein. After identifying such proteins, if simultaneous expression of both mutant proteins in the original biological system restores the original phenotype, then the specific protein-protein interaction between the nonmutated versions of these proteins must be critical for the function associated with the mutant phenotype.

Anticipated Results

Between one and several hundred loss-of-function mutants should be obtained from successful application of this methodology.

Time Considerations

Altogether, this methodology should require ~2 weeks to obtain and characterize loss-of-function mutants, if the cDNAs for the interacting protein pair are available and behave well in the two-hybrid system.

Literature Cited

Cadwell, R.C. and Joyce, G.F. 1992. Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* 2:28-33.

Gyuris, J., Golemis, E., Chertkov, H., and Brent, R. 1993. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* 75:791-803.

Lin-Goerke, J.L., Robbins, D.J., and Burczak, J.D. 1997. PCR-based random mutagenesis using manganese and reduced dNTP concentration. *Biotechniques* 23:409-412.

Mendelsohn, A.R., Hamer, J.D., Wang, Z.B., and Brent, R. 2002. Cyclin D3 activates caspase 2, connecting cell proliferation with cell death. *Proc. Natl. Acad. Sci. U.S.A.* 99:6871-6876.

Contributed by Andrew R. Mendelsohn
The Molecular Sciences Institute
Berkeley, California

CHAPTER 21

Chromatin Assembly and Analysis

INTRODUCTION

Many nuclear processes require that DNA be recognized by sequence-specific DNA binding proteins. The packaging of DNA into chromatin inhibits DNA binding by most proteins, and also influences the efficiency of events such as transcription, replication, recombination, and DNA repair. These considerations have led to an increased interest in characterizing chromatin structure in general and the precise structural changes that can occur over specific chromosomal domains.

The basic unit of chromatin is the nucleosome, composed of DNA and histones (small basic proteins that are highly conserved in all eukaryotes). Each nucleosome consists of two copies each of histones H2A, H2B, H3, and H4, arranged into an octamer that organizes 146 bp of DNA. The DNA is wrapped 1.65 times around the histones, and so is severely bent compared to DNA in solution. Each histone has an N-terminal “tail” that extends outside the DNA, and these tails are the sites of numerous covalent modifications that contribute significantly to the regulation of chromatin structure. The most extensively studied of these modifications is acetylation, which occurs on lysine residues.

Chromatin structure can be changed by covalently modifying the nucleosome tails or by noncovalently “remodeling” nucleosome structure. Remodeling is an ATP-dependent process that can alter nucleosome mobility and increase the access of DNA binding proteins and nucleases to the nucleosome. Both remodeling and covalent modification of nucleosomes can lead to local changes in chromatin structure that have been associated with activation or repression of transcription. These structural changes can be studied by characterizing the specific chromatin structure over a defined region of the chromosome *in vivo*, and by recapitulating chromatin structure on DNA fragments *in vitro*.

This chapter presents techniques for analyzing chromatin structure. *UNIT 21.1* describes the use of micrococcal nuclease (MNase), an important reagent for chromatin research, in characterizing nucleosome structure across any specific gene in intact cells. Because MNase cleaves preferentially in the linker DNA that lies between individual nucleosomes, MNase cleavage can be used to map the boundaries of nucleosomes and determine their spacing.

To perform a detailed analysis of chromatin structure, it is frequently necessary to assess whether the nucleosomes in a chromosomal domain are covalently modified, and to determine which other DNA binding proteins are bound nearby. The most commonly studied histone modification is the acetylation of lysine residues. Many cellular enzymes are capable of catalyzing this modification, and several of these have been implicated in transcriptional activation. To separate the differentially acetylated histone species, it is necessary to use a specialized gel system (*UNIT 21.2*). This technique can be used to quantify the amount of each acetylated species in a mixture of histones, either a crude histone fraction prepared from cells or a mixture of histones modified *in vitro*. To determine the location of covalently modified histones, or to localize any other DNA binding protein, immunoprecipitations of cross-linked chromatin can be performed (*UNIT 21.3*). This widely used technique, called Chromatin Immunoprecipitation (or ChIP), can determine

Contributed by Robert E. Kingston

Current Protocols in Molecular Biology (2004) 21.0.1-21.0.2
Copyright © 2004 by John Wiley & Sons, Inc.

**Chromatin
Assembly and
Analysis**

21.0.1

Supplement 68

the relative concentrations of any protein epitope across a region of chromatin, provided a good antisera is available. This technique has been used to analyze relative levels of covalently modified histones across broad regions of chromatin using antisera that recognize specific histone modifications, to localize sequence-specific DNA binding factors, and to analyze components of the general transcription and replication machinery.

Analysis of the precise position and structure of nucleosomes both *in vitro* and *in vivo* can be accomplished using either enzymatic cleavage reagents, such as DNase, or chemical cleavage reagents (*UNIT 21.4*). These agents produce specific patterns of cleavage that reflect the location of DNA on the surface of the histone octamer, and thus give more precise information than MNase cleavage. While these procedures are most frequently used to analyze nucleosomes formed *in vitro* (see *UNIT 21.5-21.7*), they have also been combined with PCR-based technologies to analyze detailed nucleosome structure in intact cells (as in *UNIT 21.1*).

To characterize the functional roles of chromatin structure in nuclear processes, it is necessary to use *in vitro* systems and templates that have been assembled into appropriate chromatin structure. In order to form chromatin *in vitro*, it is first necessary to have a source for core histones (*UNIT 21.5*). Either isolated histones or nucleosomal DNA can be used to transfer histones onto a specific template, although isolated histones can be used in a wider variety of protocols. *UNIT 21.5* also provides methods for isolating intact polynucleosomes from cells, which have a wide variety of uses *in vitro*.

Templates can be assembled into nucleosomes *in vitro* using either salt dialysis (*UNIT 21.6*) or an enzymatic chromatin assembly system (*UNIT 21.7*). Salt dialysis offers the advantage that no specialized proteins are needed, and the final nucleosome preparation is not contaminated by additional factors that might influence further mechanistic studies. It has the significant disadvantage that appropriate nucleosomal spacing on natural templates or on arrays of nucleosomes will not be achieved unless very specialized templates (such as 5S arrays, see *UNIT 21.6*) are used. Enzymatic assembly of nucleosomal arrays (*UNIT 21.7*) will achieve physiological spacing of nucleosomes on virtually any template, and thus is considerably more flexible than salt dialysis. Care must be taken, however, to either purify the template away from the assembly factors following assembly, or to carefully consider the potential role of the assembly factors in any process that is under study.

Many nuclear processes involve the concerted action of several different DNA binding factors. Factors might bind either sequentially or simultaneously to a given regulatory element. To determine whether factors are bound simultaneously, sequential chromatin immunoprecipitation (*UNIT 21.8*) can be used. Antisera to one factor is used to isolate DNA sequences bound by that factor and then a second immunoprecipitation is used with antisera specific to a second factor. Quantification of the amounts of specific DNA sequences that are immunoprecipitated in each step allows a determination of whether two factors can simultaneously occupy the same sequence.

The techniques described above can be used to define the proteins that bind to a specific gene of interest, but sometimes the starting point is a specific factor and the goal is to identify all of the genes that the factor interacts with. Microarray technology (see Chapter 22) can be combined with chromatin immunoprecipitation to identify target sites for a factor from a collection of a large number of DNA sequences. DNA is immunoprecipitated with antisera specific to a regulatory factor and the precipitated DNA is hybridized to a microarray (*UNIT 21.9*). In this way, binding sites for a factor can be identified on any collection of DNA sequences for which a microarray is available.

Robert E. Kingston

Micrococcal Nuclease Analysis of Chromatin Structure

Micrococcal nuclease (MNase) assays of chromatin are relatively simple procedures for obtaining information about the locations of nucleosomes along DNA strands. When nuclei or permeabilized cells are exposed to MNase in the presence of a divalent cation, the enzyme makes double-stranded cuts between nucleosome particles (see *UNIT 3.12*). Treatment of chromatin substrates with very high concentrations of MNase gives rise mostly to mononucleosome-length DNA, whereas low concentrations of the enzyme will produce one double-stranded cut every 10 to 50 nucleosomes (depending on the exact concentration). MNase can also make single-stranded DNA cuts on the histone octamer, and thus experiments to map the positions of nucleosomes are usually performed with native double-stranded DNA.

In this unit, Basic Protocols 1 and 2 describe how to permeabilize cells or prepare nuclei from tissues, respectively, and then perform MNase digestion. Basic Protocol 3 describes how to perform MNase digests of purified genomic DNA, which will serve as a control for the inherent MNase cleavage pattern over a sequence of interest. Support Protocol 1 describes the purification of DNA from chromatin digests in a manner suitable for subsequent analysis. Support Protocol 2 describes different ways of analyzing the purified DNA at the Southern blot level of resolution, and Support Protocol 3 describes a modification of ligation-mediated PCR (LM-PCR; also see *UNIT 15.3*) to map genomic, double-stranded MNase cleavages at the nucleotide level of resolution.

STRATEGIC PLANNING

In an experiment to probe chromatin structure, the chromatin is made accessible to MNase, the chromatin is digested with the enzyme for a brief period of time, the reaction is stopped by the addition of detergent and proteinase, and DNA within the chromatin is isolated free from protein. Prior to initiating such an analysis, it is important to consider the following three parameters, which will dictate a specific experimental strategy.

Chromatin Preparation Method

The quality of an MNase analysis is critically dependent upon the integrity of the chromatin substrate. Isolated nuclei are most commonly used, but there is clear evidence that chromatin components are perturbed during the nuclear isolation procedure. Permeabilized cells retain chromatin components that can be lost during nuclear isolation, and therefore permeabilization is preferable to nuclear isolation (Pfeifer and Riggs, 1991). The method of preparing a chromatin substrate is generally dictated by the source of cells: tissues must usually be disrupted mechanically, and therefore nuclear isolation is required, whereas cell suspensions or cells on tissue culture plates are uniformly accessible and therefore are suitable for cell permeabilization. The latter protocol is also far faster and simpler. The two procedures are described in Basic Protocols 1 and 2.

Type of Information Desired

The extent of MNase digestion of the sample determines the kind of inferences that can be made about the chromatin structure of the genomic region of interest. An extensive MNase digestion, which yields predominantly mononucleosomal-sized DNA fragments, can indicate the frequency with which a DNA sequence is nucleosomal among a population of templates in the reaction. A partial MNase digestion, which yields lengthy

polynucleosomal-sized DNA fragments, can indicate the positions of nucleosomes with respect to a defined point in the DNA, such as a restriction site. Each of these approaches employs a different Southern blotting or LM-PCR strategy to map MNase cleavages, described in Support Protocols 2 and 3, respectively. It is essential to plan a cleavage mapping strategy in order to determine the controls needed for the MNase digestion reactions.

Controls

Several controls are critical for the strategic planning of any MNase study. First, it is essential to analyze an MNase digestion of protein-free DNA, as described in Basic Protocol 3, because MNase cleaves DNA in a highly nonrandom manner. The range of MNase digestion of free DNA should be comparable to the range of digestion of the target sequence seen in chromatin MNase digests. Only MNase cleavages that are induced or enhanced in chromatin, compared to free DNA, can be presumed to provide conclusive information about the presence of nucleosomes and their positions (see Fig. 21.1.1 in Support Protocol 2). Second, with isolated nuclei or permeabilized cells, it is necessary to include a sample that was incubated in the MNase digestion buffer for the same amount of time and at the same temperature as the MNase-treated samples, but without added enzyme. Cleavages that occur over the genomic region of interest in this control will be due to endogenous nucleases. The Commentary following the protocols describes how to minimize the effects of these enzymes (see Critical Parameters and Troubleshooting).

In summary, a well-controlled MNase assay must include the following, preferably all on the same blot or LM-PCR gel, and in the following order: (1) MNase digestion of free DNA at several enzyme concentrations, including a no-enzyme control; (2) MNase digestion of chromatin at several enzyme concentrations, including a no-enzyme control; (3) MNase digestion of a chromatin sample from a different cell type or physiological condition, to determine the specificity of nucleosomal arrangements observed in the first chromatin sample; and (4) an internal marker lane (see Support Protocol 2).

BASIC PROTOCOL 1

MICROCOCCAL NUCLEASE DIGESTION OF CHROMATIN IN PERMEABILIZED CELLS

In this protocol, chromatin substrates are prepared by permeabilizing cultured cells using the detergent lysolecithin (either *in situ* for cells adherent on plates, or in suspension for suspension cultures). The chromatin is then digested with micrococcal nuclease, and the DNA purified so that it is suitable for Southern blot (see Support Protocol 2) or LM-PCR analysis (see Support Protocol 3). Before proceeding with the full-scale experiment, optimal conditions should be determined by performing a preliminary set of trials, both to determine the appropriate lysolecithin incubation conditions and to select a workable range of MNase that yields the desired level of digestion.

Materials

- Cells cultured on 100-mm petri dishes *or* cells in suspension
- Permeabilization solution 1, room temperature and 37°C (see recipe)
- 1 mg/ml lysolecithin (Sigma) in permeabilization solution 1 (mix fresh before use)
- Permeabilization solution 2 (see recipe), with and without MNase (see step 4 below; mix fresh before use)
- 2× TNE SK solution (see recipe)
- Lysis dilution buffer: 150 mM NaCl/5 mM EDTA
- Phase-contrast microscope
- 12-ml conical polypropylene tubes with caps (e.g., Sarstedt)

NOTE: All solutions, pipets, and pipet tips should be at room temperature so that the permeabilization reaction proceeds rapidly.

Carry out cell permeabilization

- 1a. *Adherent cells (per 100-mm plate)*: Working at room temperature (e.g., in a tissue culture hood), aspirate the medium from the cells and add 5 ml room temperature permeabilization solution 1. Do not remove cells from the plates.

Use a separate plate of cells for each MNase digestion condition, remembering to include a zero-enzyme condition (see Strategic Planning). Try to harvest cells at 70% to 90% confluence.

- 1b. *Suspension cultures*: Pellet by centrifugation under conditions appropriate for cell type and resuspend cell pellet in 5 ml permeabilization solution 1.

The remainder of the procedure is the same for suspension cells as for adherent cells, except that centrifugation is required to change solutions.

2. Aspirate medium from each plate, then treat with lysolecithin (diluted from 1 mg/ml stock in 37°C permeabilization solution 1 to 2.8 ml total volume) at room temperature as follows: either use 0.025% lysolecithin and incubate 2 min, or use 0.05% lysolecithin and incubate 1 min.

Lysolecithin has been the preferred reagent for permeabilizing cells (Zhang and Gralla, 1990; Pfeifer and Riggs, 1991), although other reagents have not been explored in depth for this purpose.

Different cell lines or the same cell line cultured under different conditions may require different concentrations of lysolecithin, or different times or temperatures (30° or 37° C instead of room temperature) for the detergent incubation (see Critical Parameters and Troubleshooting).

3. Aspirate solution from plate, then add 5 ml room temperature permeabilization solution 1 (without lysolecithin). Examine the cells by phase-contrast microscopy to be sure that cell lysis has not occurred. If it has, repeat the assay with a lower concentration of lysolecithin.

Carry out MNase digestion of permeabilized cells

4. Aspirate solution from plate, then add 2.8 ml of room temperature permeabilization solution 2 containing either 0, 7.5, 15, or 30 U MNase for a partial digestion assay, or 0, 50, 100, or 200 U MNase for an extensive digestion assay. Incubate 5 min at room temperature.

Dilute the MNase into aliquots of permeabilization solution 2 just prior to use.

5. Aspirate solution from plate, then add 2.8 ml 2× TNESK solution. Swirl around plate to complete cell lysis.

If the benchtop is not level, occasionally tilt the plate during the MNase digestion.

6. Add 2.8 ml lysis dilution buffer. Swirl around plate to complete mixing, then transfer mixture to a 12-ml conical polypropylene tube. Cap tube and incubate overnight at 37°C.

7. Proceed to purification and characterization of the DNA (see Support Protocol 1).

MICROCOCCAL NUCLEASE DIGESTION OF CHROMATIN IN ISOLATED NUCLEI

In this protocol, chromatin substrates are prepared by isolating cell nuclei, the chromatin within the nuclei is digested with micrococcal nuclease, and the DNA is purified so that it is suitable for Southern blot (UNIT 2.9A) or LM-PCR analysis (UNIT 15.3). Different nuclear isolation conditions are described for liver, kidney, and spleen, each of which exhibit different characteristics during homogenization and thus should serve as examples for a wide variety of tissue types.

BASIC PROTOCOL 2

Chromatin Assembly and Analysis

21.1.3

Materials

Animal tissues (e.g., liver, kidney, and/or spleen) from sacrificed animal(s) or biopsy
Nuclear buffers A, B, and C (see recipe)
Calcium- and magnesium-free (CMF) PBS (e.g., APPENDIX 2)
1 M NaOH
2× TNESK solution (see recipe)
0.1 M CaCl₂
Micrococcal nuclease (MNase) stock solution (see recipe)
Razor blades or scalpels
100-mm petri dishes on ice
10- and/or 15-ml tissue homogenizers with Teflon-coated pestles
Tissue grinder motor and chuck
Phase-contrast microscope
Cheesecloth
15- and 30-ml glass centrifuge tubes (e.g., Corex)
Refrigerated high-speed centrifuge with fixed-angle rotor
Ultracentrifuge with SW50.1 rotor and 1/2 × 2-in. Ultraclear tubes
Ultraviolet light spectrophotometer

NOTE: All solutions, test tubes, pipets, and pipet tips should be ice cold.

Carry out nuclear isolation

1. Immediately after sacrificing animal or obtaining biopsy, rinse isolated tissues in 25 ml nuclear buffer A in a 100-ml glass beaker on ice.

For mice, combine tissues from four mice >1 month old.

Do not freeze tissue or nuclei.

2. Transfer tissues to 100-mm plastic petri dishes on ice containing the following amounts of nuclear buffer A: liver, 8 ml buffer A; kidney, 5 ml; spleen, 4 ml. Mince tissues into chunks several millimeters in size using razor blades or scalpels in a scissors-like motion.
3. Decant each minced preparation into a separate homogenizer. For the aforementioned amounts of tissue, use a 15-ml apparatus for liver or kidney and a 10-ml apparatus for spleen. Let spleen fragments settle on ice ~10 min before decanting, allowing red blood cells to remain in the supernatant. For liver and kidney fragments, decant supernatants immediately.
4. Add 5 ml fresh nuclear buffer A to all samples. Mix, decant supernatants again, and resuspend the tissue fragments into 5 ml fresh nuclear buffer A.

IMPORTANT NOTE: *From this point on, perform all manipulations in the cold room.*

5. Homogenize tissues with five to ten strokes of a motor-driven tissue grinder, followed by an additional five sharp strokes by hand. Dilute a small portion of the homogenate into CMF PBS and view under a phase-contrast microscope. Nuclei appear as grey, evenly round blobs. If intact tissue or cells remain, perform an additional five strokes or so with the homogenizer and check once again for quantitative cell breakage.

Some tissues, such as spleen, require considerably less effort to disrupt than liver. Do not proceed until the cells are quantitatively broken.

6. Filter each homogenate through eight layers of folded cheesecloth that has been prewet with nuclear buffer A into a 30-ml Corex tube on ice. Wearing gloves, twist the cheesecloth to elute most of the liquid.

7. Layer homogenates onto cushions of 1:1 nuclear buffer A/nuclear buffer B in 15-ml Corex tubes, using the following volumes of the buffer A + B mixture: liver, 1.4 ml; kidney, 1 ml; spleen, 0.6 ml.
8. Centrifuge 15 min at $12,000 \times g$, 4°C , in a fixed-angle rotor (e.g., 10,000 rpm in 55-34 rotor).

The liver and kidney pellet should be large, with a bright red center surrounded by brown. The spleen pellet should be small and brown.
9. Discard supernatants and resuspend pellets into the following volumes of nuclear buffer B: liver and kidney, 11 ml; spleen, 3.6 ml. First add ~ 2 ml to each pellet and gently resuspend into a thick slurry, then add remaining nuclear buffer B.
10. Layer 3.6 ml of each suspension over 1.2-ml cushions of nuclear buffer B in $\frac{1}{2} \times 2$ -in. Ultraclear SW50.1 centrifuge tubes (liver and kidney in three tubes; spleen in one).
11. Spin 90 min at $120,000 \times g$ (e.g., 37,000 rpm in SW50.1 rotor) 4°C .
12. Decant by inverting tube and blotting rim on a Kimwipe tissue. Gently resuspend each pellet into the following volumes of nuclear buffer C: liver and kidney, 0.2 ml; spleen, 0.25 ml. Combine suspensions from a single tissue type into a single tube and keep on ice.
13. Dilute 5 μl of each nuclear suspension into 2 ml of 1 M NaOH and measure the OD_{260} by spectrophotometry. Dilute the nuclear suspensions with nuclear buffer C to obtain the following diluted OD_{260} values: liver, $\text{OD}_{260} = 0.25$; kidney and spleen, $\text{OD}_{260} = 0.085$.

Carry out MNase digestion of isolated nuclei

14. Remove one-fifth of each nuclear suspension and mix with an equal volume of $2 \times$ TNESK solution. Mix very well. Incubate overnight at 37°C .

This sample serves as a “quench” control for the extent of nuclear digestion prior to warming up the nuclei.
15. Divide remaining volume into the following number of reaction tubes on ice: liver and kidney, up to six; spleen, about three. Be sure to agitate samples frequently to prevent clumping.

Each reaction tube will be used for a different concentration of MNase.
16. Add 0.1 M CaCl_2 to the first tube to a final concentration of 3 mM, incubate tube 3 min at 37°C , then add equal volume of $2 \times$ TNESK solution and incubate overnight at 37°C .

This sample serves as an essential control to assess the level of endogenous nuclease. Swish the tube around in the 37°C water bath to facilitate warming.
17. Add 0.1 M CaCl_2 to the next tube to a final concentration of 3 mM and incubate tube 1.5 min at 37°C to warm nuclei.
18. Add the desired amount of MNase to each tube. For partial MNase digestions, try the following ranges: liver and kidney, 0.1 U MNase per ml reaction, 0.2 $\mu\text{g}/\text{ml}$, 0.5 $\mu\text{g}/\text{ml}$, 1 $\mu\text{g}/\text{ml}$, and 2 $\mu\text{g}/\text{ml}$; spleen, 0.5 $\mu\text{g}/\text{ml}$ and 1 $\mu\text{g}/\text{ml}$. For extensive MNase digestions, try a range of 10 to 100 U/ml. Mix, then incubate tube 1.5 min at 37°C .
19. Terminate reactions by adding an equal volume of $2 \times$ TNESK solution and shaking the tube sharply a few times to mix, then incubate overnight at 37°C .

20. Repeat steps 17 to 19 in succession for each MNase concentration to be used.

Do not add the CaCl₂ to all tubes at the same time. Do each CaCl₂ addition and MNase digestion separately.

21. Proceed to purification and characterization of DNA (see Support Protocol 1).

MICROCOCAL NUCLEASE DIGESTION OF PURIFIED GENOMIC DNA

The enzyme MNase has a nonrandom sequence preference for cleaving DNA. It is therefore essential to compare the MNase cleavage pattern of chromatin with that for free genomic DNA in order to determine how the inherently nonrandom pattern is influenced by nuclear proteins such as the histones. MNase-digested free DNA, digested to the same relative extent as DNA from a chromatin digest, should be analyzed side-by-side in the cleavage mapping studies described in Support Protocols 2 and 3 below.

Materials

0.5 to 1 mg purified genomic DNA (see Support Protocol 1)

Nuclear buffer C (see recipe)

0.5 M CaCl₂

Micrococcal nuclease (MNase)

0.25 M EGTA

Chloroform

Heating block set to 68°C

Bucket of ice

Additional reagents and equipment for agarose minigel electrophoresis (*UNIT 2.5A*)

1. Dilute 100- μ g aliquots of genomic DNA into a final volume of 300 μ l nuclear buffer C at room temperature. Set up four or five such aliquots.

IMPORTANT NOTE: *For each aliquot, carry out reaction through step 4 before starting the next.*

The reaction is carried out at room temperature, rather than 37°C, to provide better control over the digestion conditions.

2. Add CaCl₂ to 3 mM final concentration.
3. Add MNase to 0.5, 1, 2, or 3 U/ml. Mix sharply to distribute the enzyme evenly (the solution may be viscous). Incubate 3 min.
4. Aspirate solution with a micropipettor and transfer to a tube containing 50 μ l of 0.25 M EGTA, all prewarmed at 68°C. Keep tube in heat block for 10 min.
5. Transfer tube to ice bucket.
6. Electrophorese 200-ng samples on a 1.2% agarose minigel to check the extent of digestion.

It is useful to run these MNase digestion products alongside chromatin DNA samples (see Support Protocol 1), so that the ranges of digestion are similar.

7. Extract the samples of interest once with 1 vol chloroform.
8. Proceed to precipitation, quantitation, and assessment of the DNA (see Support Protocol 1, step 7 onward).

PURIFICATION AND CHARACTERIZATION OF DNA FROM CHROMATIN DIGESTIONS

This protocol has been optimized to maximize the recovery of small amounts of DNA from chromatin digests, and so that the resulting DNA is suitable for direct analysis and should not require further concentration. The protocol begins with a cell or nuclear lysate in an SDS solution that has been incubated with proteinase K for several hours or overnight (this can be used for any small-scale DNA preparation; e.g., see Basic Protocols 1 and 2).

Materials

MNase-digested cell or nuclear lysate (see Basic Protocols 1, 2, and 3)
TE buffer, pH 7.9 (APPENDIX 2)
Neutralized phenol (see recipe)
Chloroform
Ether (for permeabilized cell preparations)
5 mg/ml RNase A (for permeabilized cell preparations)
3 M sodium acetate
95% and 70% ethanol, room temperature
10- or 12-ml polypropylene tubes with tight caps (e.g., 12-ml Sarstedt tubes)
Shaker or rocking device
6-in. Pasteur pipets
30-ml Corex tubes, silanized (APPENDIX 3B; for permeabilized cell preparations)
6000- to 8000-MWCO dialysis tubing (for permeabilized cell preparations)
1- to 5- μ l glass capillary pipets
Glass capillary pipettor
Plastic wrap (e.g., Saran Wrap)
Ultraviolet light spectrophotometer
Additional reagents and equipment for dialysis (APPENDIX 3C) and agarose minigel electrophoresis (UNIT 2.5A)

Carry out organic extractions of MNase-digested DNA

1. Dilute the cell or nuclear lysate with 1 vol TE buffer, pH 7.9. If the resulting volume is ≤ 0.5 ml, perform the following steps in a 1.5-ml microcentrifuge tube; if the volume is ≥ 0.5 ml, transfer to 12-ml polypropylene tube(s).

Diluting the material will make the solution less viscous and reduce loss of DNA during the subsequent manipulations. A marked decrease in viscosity of the MNase-treated samples should be evident at this step, and indicates successful cleavage.

2. In a fume hood, add 1 vol neutralized phenol. Invert tube sharply several times, then place on a gentle shaker or rocking device for 10 to 20 min.
3. Centrifuge the samples 30 sec in a microcentrifuge for those in 1.5-ml tubes, or 5 min at $500 \times g$ (2000 rpm in 55-34 rotor) for those in 12-ml tubes. Using a 6-in. Pasteur pipet, remove the upper aqueous layer and transfer to a fresh tube.

There is little worry about shearing the DNA here, as it should be significantly digested by the MNase and the "0 MNase" samples will be digested with restriction enzymes during analysis.

4. Add 1 vol chloroform. Invert tube sharply several times, then gently shake for 10 to 20 min. Spin and obtain upper aqueous phase as in step 3. Leave the interface behind.

5. Repeat chloroform extractions for a total of two to four extractions, until the interface between the organic and aqueous phases is clear.

For samples that were treated with MNase, two or three chloroform extractions should be sufficient; for samples that did not receive MNase, the aqueous phase should be very viscous and may not clear completely even after four extractions.

For MNase digestions of isolated nuclei and purified DNA, proceed directly to step 7 of this protocol. For MNase digestions of permeabilized cells, proceed to step 6 below.

6. *For permeabilized cell preparations:* Add 1 vol ether, then invert tube sharply several times, burp the tube, and gently shake or rock for 10 to 20 min. Heat uncapped for 10 to 15 min in a 65°C water bath, shaking periodically to help evaporate the ether, then dialyze overnight in 6000- to 8000-MWCO dialysis tubing against two changes of 100 vol TE buffer, pH 7.9, at 4°C. Transfer the dialysate, which will have increased in volume by $\geq 50\%$, to a 30-ml silanized Corex tube and add RNase to 25 $\mu\text{g/ml}$. Incubate 1 hr at room temperature.

The dialysis procedure alleviates the problem of lysolecithin (which can inhibit restriction enzymes) precipitating with the DNA in later steps, and the RNase treatment digests the cytoplasmic RNA present in the permeabilized cell lysates.

7. Add $\frac{1}{10}$ vol of 3 M sodium acetate. Mix well by inversion.

Precipitate, spool, and resuspend DNA precipitate

IMPORTANT NOTE: For each sample, perform steps 8 to 12 in succession without stopping.

8. Add 2.5 vol of 95% ethanol. Invert gently 10 to 30 times, until the Schlieren patterns in the liquid are completely gone, and assess the solution to determine how to proceed.
 - a. If the DNA has not been extensively digested with MNase, a stringy white precipitate should be visible; continue with steps 9 to 13 (and omit steps 14 to 17).
 - b. If the DNA has been extensively digested with MNase, a precipitate may not be visible; proceed directly to step 14 (beginning with overnight incubation at -20°C).
 - c. With a moderate amount of MNase digestion, there may be a small stringy precipitate; spool this out by following steps 9 to 13, and proceed directly to step 14 (beginning with overnight incubation at -20°C) to recapture DNA from the remaining solution.

The DNA from MNase-digested permeabilized cells will probably be too dilute to form the stringy precipitate, in which case proceed directly to step 14.

9. Use a 1- to 5- μl glass microcapillary tube to spool out the DNA precipitate from a single tube. Use a microcapillary pipettor to expel liquid that rises into the capillary while touching the DNA to the inside of the tube to remove excess liquid.

Do this step over a sheet of plastic wrap spread on the laboratory bench. If the spooled DNA falls off the microcapillary, simply use the microcapillary to pick it up from the plastic wrap and carry on.

10. Immediately dip the DNA in a 1.5-ml microcentrifuge tube containing 70% ethanol. Swish around for a few seconds, then again blow out liquid from the tube while removing excess liquid from the DNA.

Work quickly; do not allow the DNA to dry out or it will be impossible to resuspend.

11. Blot the DNA on the inside of an empty 1.5-ml tube. Air dry for ~5 sec.
12. Place the DNA in a 1.5-ml tube containing 75 to 300 μ l TE buffer (depending on the amount of DNA spooled) for isolated nuclei or free DNA experiments, or 400 μ l TE for permeabilized cells. Discard the capillary tube.

Remember to go on to the next sample, starting at step 8. Do not add ethanol to multiple samples at the same time; if the DNA is kept as a stringy precipitate for too long it will be very difficult to resuspend.

13. Let DNA dissolve overnight at room temperature. Vortex lightly a few times, then store at 4°C. Do not freeze.

The DNA is stable for years at 4°C.

Precipitate dilute DNA remaining in solution

14. If the DNA does not precipitate at step 8, or if only a small precipitate can be spooled out, incubate the remaining mixture at -20°C overnight.
15. If using 12-ml polypropylene or 30-ml Corex tubes, centrifuge 10 min at $10,000 \times g$ in a fixed-angle rotor (e.g., 9000 rpm in 55-34 rotor); if using 1.5-ml tubes, spin 5 min in a microcentrifuge. Decant supernatant. Add 0.5 ml of 70% ethanol and gently invert the tube once. Repeat centrifuge spin for 1 min.
16. Decant supernatant. Vacuum dry pellet for 5 min. For experiments with isolated nuclei or free DNA, resuspend in 75 to 300 μ l TE buffer (depending on the size of the pellet and anticipated amount of DNA); for experiments with permeabilized cells, resuspend the DNA in 400 μ l TE. If some DNA from the sample was spooled out and resuspended in TE, use this solution to resuspend the DNA pellet, so the portions are pooled and kept concentrated. Allow solution to dissolve overnight at room temperature and then store at 4°C, as in step 13.
17. *For DNA prepared from permeabilized cells:* Repeat the precipitation starting at step 7 and resuspend the final precipitate in 75 to 150 μ l TE buffer, so that the DNA is sufficiently concentrated for subsequent analysis.

Quantitate and assess DNA

18. Dilute 4 μ l DNA solution in 400 μ l (total) TE buffer and read the OD_{260} . Calculate DNA concentration and total yield, assuming that an OD_{260} reading of 1 equals 50 $\mu\text{g/ml}$ of DNA.

If DNA is very viscous, take up 4 μ l of water in pipet tip and mark the meniscus with a felt tip pen; expel and then use marked tip to take up 4 μ l of the DNA solution.

19. Apply 0.5- μg samples of each DNA to a 1.2% agarose minigel and assess the level of endogenous nuclease and MNase cleavage of the chromatin.

A ladder of bands consistent with the nucleosomal repeat length should be evident.

NUCLEASE CLEAVAGE MAPPING STRATEGIES

Very different approaches are taken to map extensive versus partial MNase digestions, and it is important to know what patterns to look for in the final data, as described below. The methodology for running the DNA samples on agarose gels, blotting to filters, and hybridizing to radiolabeled probes involves well-established technology described in *UNITS 2.5A, 2.9A, & 2.10*, respectively. Mapping MNase cleavages at the nucleotide level of resolution requires special modifications of existing protocols and is described in Support Protocol 3.

SUPPORT PROTOCOL 2

Chromatin Assembly and Analysis

21.1.9

Cleavage Mapping of an Extensive MNase Digest

DNA from a chromatin sample that was extensively digested with MNase is typically assessed by performing gel electrophoresis (e.g., 1.2% to 1.5% agarose or 5% polyacrylamide) to resolve mononucleosomal fragments, blotting to a nylon filter, and then hybridizing to an oligonucleotide or short DNA probe corresponding to the DNA region of interest. Visualization of a mononucleosome-sized hybridization band is evidence that some fraction of the templates in the chromatin population was nucleosomal during the MNase digestion. If an oligonucleotide probe from a different part of the genome is hybridized to another portion of the original chromatin digest, then the signals from the two probes can provide information about the relative frequencies with which the two DNA sequences were nucleosomal in the original population of chromatin templates. The specific activities of the oligonucleotide probes must be comparable or accounted for. Very highly labeled probes are required to observe mononucleosome-sized bands from single-copy higher-eukaryotic gene sequences.

Cleavage Mapping of a Partial MNase Digest by Indirect End Labeling

DNA that is isolated from a chromatin sample that was only partially digested with MNase is typically assessed by an indirect end-labeling experiment (Wu, 1980; see Fig. 21.1.1). The indirect end-labeling approach is designed to map, at the Southern blot level of resolution, the positions of frequent cleavage by MNase at DNA sites in a population of chromatin templates. If multiple MNase cleavage sites occur 160 to 200 bp apart, then nucleosomes may be positioned over the sequence of interest. The variation in the nucleosome repeat length depends upon the organism and other parameters, such as the presence of linker histone in the local chromatin domain. It is also possible to map double-stranded MNase cleavages at the nucleotide level of resolution, using a modification of the LM-PCR technique (see Support Protocol 3).

For indirect end-labeling, the MNase-treated, purified DNA is digested with one or more restriction enzyme(s) that cut(s) at least 1 kb away on both sides of the region of interest (Fig. 21.1.1; see example for *EcoRI*). The genomic DNA is then electrophoresed, blotted to nylon or nitrocellulose membrane, and then hybridized to a purified DNA fragment probe that is 0.5 to 1 kb in length and that abuts one restriction-cut end of the genomic DNA target sequence (Fig. 21.1.1; *EcoRI-HindIII* fragment). Autoradiographic or phosphorimager exposures should reveal the “parent” restriction fragment from control chromatin samples that were not treated with MNase, and the presence of sub-bands in MNase-treated samples. Discrete sub-bands will appear if MNase cleaves at the same position(s) on most of the substrates in the population (Fig. 21.1.1; chromatin sample 2). The positions of cleavages can be determined by the distance in kilobases from the restriction enzyme site abutted by the probe fragment (Fig. 21.1.1, Mapping strategy). If a ladder of sub-bands is evident with the aforementioned spacing, it suggests the presence of a positioned nucleosome or a nucleosome array (Fig. 21.1.1, ovals in diagram in Mapping Strategy).

Inclusion of an internal marker control allows for a definitive mapping of the position of MNase cleavage in an indirect end-label assay. An aliquot of uncut genomic DNA (designated the internal marker control), equal to or less than the amount of the chromatin samples to be loaded onto the gel for Southern blotting, is digested to completion with the restriction enzyme(s) used for the chromatin samples. The internal marker control sample is then digested partially with restriction enzyme(s) that cleave in the vicinity of the region of interest: i.e., where nucleosome phasing is being investigated (Fig. 21.1.1; *HindIII* partial digest). The marker control sample is then loaded onto the same gel as the chromatin samples. It should yield hybridization sub-bands due to the partial restriction enzyme cleavages. These sub-bands provide the ideal standard for the mobility of

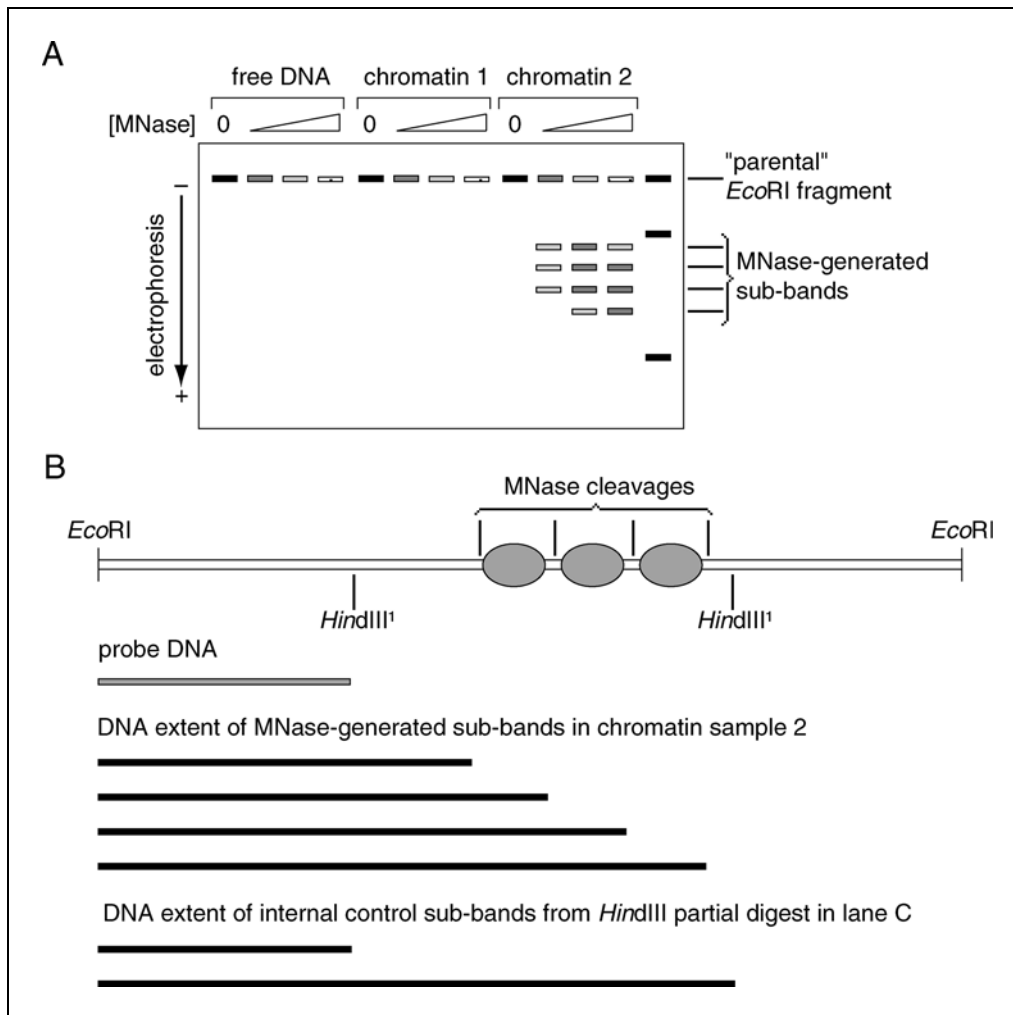


Figure 21.1.1 (A) Hypothetical Southern blot of indirect end-label analysis of MNase-treated chromatin. Expected variation in band intensities are shown with increasing MNase digestion. Lack of sub-bands in free DNA is due to lack of intrinsic MNase hypersensitivity, and lack of sub-bands in chromatin sample 1 is due to lack of nucleosome phasing in the region probed. (B) Mapping strategy and interpretation of bands depicting a positioned nucleosome array.

sub-bands in the MNase digested samples, and allow accurate mapping of the MNase cleavage sites with respect to the positions of the restriction sites from the partial digest (see Liu et al., 1988; McPherson et al., 1993).

Regarding technical details: 12 μg of genomic DNA per lane are sufficient for the Southern blots; large (e.g., 25-cm) 1.2% agarose gels run overnight at 40 to 50 V give the best resolution; and DNA probes should be labeled with ^{32}P to a specific activity of at least 0.5×10^9 cpm per microgram of DNA, preferably in the range of 10^9 cpm/ μg .

USING A MODIFIED LM-PCR PROCEDURE TO MAP DOUBLE-STRANDED MNase CLEAVAGES AT THE NUCLEOTIDE LEVEL OF RESOLUTION

The ligation-mediated PCR (LM-PCR; UNIT 15.3) procedure is used for genomic footprinting and therefore can be used to map MNase cleavages at the nucleotide level of resolution. However, conventional LM-PCR detects single-stranded genomic cleavages, whereas MNase makes single-stranded DNA cleavages on nucleosomes. In order to selectively map these double-stranded intranucleosomal MNase cleavages, modifications to the LM-PCR procedure have been developed (McPherson et al., 1993), as described

SUPPORT PROTOCOL 3

Chromatin Assembly and Analysis

21.1.11

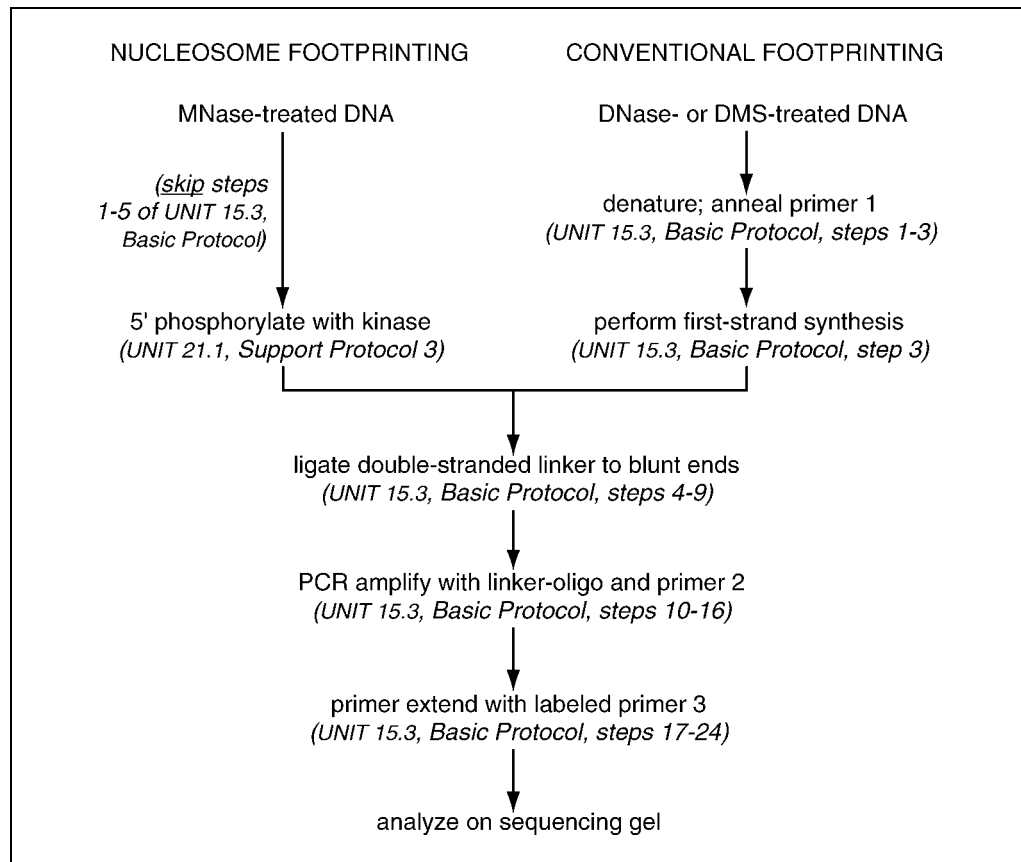


Figure 21.1.2 Nucleosome (MNase; left) versus conventional footprinting (right) by ligation-mediated PCR.

here. See Figure 21.1.2 for a comparison of nucleosome and conventional LM-PCR footprinting.

The essential change that is made in the LM-PCR protocol is to eliminate the initial primer extension step of LM-PCR and directly ligate the asymmetric linker to the MNase-cleaved double-stranded DNA. But because MNase cleavage leaves a 5' hydroxyl, it is necessary to first phosphorylate the cleaved double-stranded substrate so that the long linker strand can be ligated to it (see Fig. 15.5.3). Note that a DMS cleavage pattern with genomic DNA, generated by conventional LM-PCR and run on the same gel as the final MNase LM-PCR products, is necessary to determine the positions of MNase cleavages along the sequence.

Materials

- DNA from MNase digest of chromatin (e.g., see Basic Protocols 1 and 2; see Support Protocol 1 for purification procedure)
- DNA from MNase digest of purified genomic DNA (e.g., see Basic Protocol 3)
- 10 U/μl T4 polynucleotide kinase (e.g., New England Biolabs) and 10× buffer (supplied with enzyme)
- 10 mM ATP
- 0.5 M EDTA, pH 8 (APPENDIX 2)
- 3 M sodium acetate
- 95% and 70% ethanol, room temperature

1. Set up phosphorylation reactions as follows in 1.5-ml microcentrifuge tubes—one for each MNase enzyme digestion point, as well as for no-enzyme and MNase-treated, purified DNA controls:

5 μ l 10 \times kinase buffer
0.5 μ l 10 mM ATP
H₂ O to give 50 μ l final volume
5 μ g MNase-cleaved, purified genomic DNA.

2. Add 1 μ l of 10 U/ μ l T4 polynucleotide kinase (10 U) and incubate 1 hr at 37°C.
3. Add 1 μ l of 0.5 M EDTA and incubate 20 min at 68°C.
4. Add 5 μ l of 3 M sodium acetate, mix, then add 125 μ l of 95% ethanol, mix, and incubate 30 min at –20°C to precipitate DNA.
5. Spin tube 5 min in a microcentrifuge, decant supernatant, add 125 μ l of 70% ethanol, and gently invert once. Spin 2 min, decant supernatant, and dry briefly under vacuum.
6. Suspend the DNA in the tube in water at 0.1 to 0.2 μ g/ μ l and store at –20°C until use.

This DNA can be used directly for LM-PCR; proceed directly to the linker ligation step (step 4 in the Basic Protocol of UNIT 15.3). Use several more cycles of the PCR reaction (e.g., 22 cycles instead of 20) than would be used for conventional LM-PCR.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

MNase stock solution

Dissolve 50 activity units MNase (Worthington Biochemicals) per microliter in 5 mM Tris·Cl (pH 7.5; APPENDIX 2)/0.01 mM CaCl₂. Divide into aliquots in 0.5-ml tubes and store at –20°C.

One activity unit of MNase = 85 OD A₂₆₀ units.

Neutralized phenol

Combine 20 ml phenol with 20 ml of 100 mM Tris·Cl (pH 7.5) in a 50-ml tube. Mix by inversion. After the phases separate, the lower, organic phase will contain neutralized phenol.

Nuclear buffer A

15 mM HEPES, pH 7.5
60 mM KCl
15 mM NaCl
2 mM EDTA
0.5 mM EGTA
0.34 M sucrose
0.15 mM 2-mercaptoethanol (add just before use)
0.15 mM spermine (add just before use)
0.5 mM spermidine (add just before use)

The above was designed for homogenizing liver tissue (Hewish and Burgoyne, 1973; Kornberg et al., 1989). For other tissues, add 0.5% nonfat dry milk (e.g., Carnation).

Nuclear buffer B

15 mM HEPES, pH 7.5
60 mM KCl
15 mM NaCl
0.1 mM EDTA
0.1 mM EGTA
2.1 M sucrose
0.15 mM 2-mercaptoethanol (add just before use)
0.15 mM spermine (add just before use)
0.5 mM spermidine (add just before use)

Nuclear buffer C

15 mM HEPES, pH 7.5
60 mM KCl
15 mM NaCl
10 mM NaHSO₃
0.34 M sucrose
0.15 mM 2-mercaptoethanol (add just before use)
0.15 mM spermine (add just before use)
0.5 mM spermidine (add just before use)

Permeabilization solution 1

150 mM sucrose
80 mM KCl
35 mM HEPES, pH 7.4
5 mM K₂HPO₄
5 mM MgCl₂
0.5 mM CaCl₂

Permeabilization solution 2

150 mM sucrose
50 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
50 mM NaCl
2 mM CaCl₂

TNESK, 2×

20 mM Tris·Cl, pH 7.4 (*APPENDIX 2*)
0.2 M NaCl
2 mM EDTA
2% SDS
0.2 mg/ml proteinase K (add just before use)

COMMENTARY

Background Information

It is increasingly appreciated that chromatin structure is integral to the mechanisms of transcriptional regulation, DNA replication, and DNA repair (van Holde, 1989; Felsenfeld, 1992). The basic repeating unit of chromatin is the nucleosome core particle, which consists of an octamer of two each of the four core histone proteins, along with ~146 bp of DNA. The addition of linker histone to

the core particle creates the nucleosome and increases the amount of associated DNA by ~20 bp. Nucleosomes and nucleosome cores occur in extensive arrays that span the eukaryotic genome and elicit higher levels of chromatin compaction. Since the histone octamers may span DNA sequences that are important for transcription factors, polymerases, and DNA modification enzymes, understanding precisely where the nucleosomes occur

can provide insight into the role of chromatin structure during physiological conditions of interest.

With the modifications to standard DNA footprinting procedures described in Support Protocol 3, it is now possible to map double-stranded genomic MNase cleavages at the nucleotide level of resolution. This is useful for comparing the apparent positions of nucleosome boundaries to the positions of regulatory-protein binding sites and other genetic landmarks (see McPherson et al., 1993). This method can also be used to compare nucleosome boundaries *in vivo* with those obtained from *in vitro* chromatin reconstitution reactions (Shim et al., 1998).

Critical Parameters and Troubleshooting

For the cell permeabilization in Basic Protocol 1, which is essentially the procedure described by Pfeifer and Riggs (1991), it is best to start with trials to determine the optimal lyssolecithin concentration and incubation temperature and select a workable range of MNase that yields the desired level of digestion. The cells must be checked at various stages of the experiment to be sure that they have not lysed. If cell lysis has occurred, repeat the experiment with a lower concentration of lyssolecithin or at a lower temperature. Lyssolecithin seems to be more potent when fresh than after it has been stored for several months. Conversely, if agarose gel analysis of the bulk chromatin samples reveals a weak ladder of MNase-generated bands and a considerable amount of undigested genomic DNA, it is likely that an insufficient fraction of the cells were permeabilized, and a higher lyssolecithin concentration, longer incubation, or higher incubation temperature may be needed.

Endogenous nucleases can be a major problem when isolating nuclei from certain tissues and cell lines using Basic Protocol 2. A simple way to reduce their effect is to work quickly during the nuclear isolation and chromatin digestion procedures and keep the chromatin on ice as much as possible. It is also helpful to minimize the time taken to warm the sample for the MNase digestion and to use a sufficient concentration of enzyme that the desired amount of digestion is complete within a few minutes. The nuclear isolation buffer contains EDTA to chelate divalent cations that activate endogenous nucleases; spermine and spermidine are included to provide counter-ions that stabilize the chromatin.

To make the nuclear isolation in Basic Protocol 2 readily adaptable for formaldehyde cross-linking experiments, the original protocols of Hewish and Burgoyne (1973) and Kornberg et al. (1989) have been modified here with a Hepes buffer system instead of Tris, because formaldehyde reacts with primary amines in the latter. Spermine and spermidine must be avoided for the same reason, so for formaldehyde-based cross-linking, the cell lysis and nuclear isolations buffers should contain 2 mM MgCl₂ to stabilize the chromatin and lack both EDTA and molecules with primary amines.

For partial MNase digestion studies with indirect end-label probes, as in Support Protocol 2, sub-bands are best seen when ~25% to 75% of the parent restriction fragment is cleaved (i.e., depleted) by MNase. It will be necessary to empirically determine the appropriate concentration of MNase by varying enzyme levels in two-fold steps over the suggested range. Excessive MNase digestion can lead to the artifactual appearance of a positioned nucleosome array because short oligonucleosome fragments bounded by MNase cleavages on both sides can hybridize to the indirect end-label probe.

Note that the activity of MNase is highly dependent upon the concentration of enzyme; a two-fold difference in concentration can yield a marked difference in the average size of double-stranded DNA cleavage product.

Another important parameter is the activity state of the genetic region of interest. Silent chromatin is often highly compacted and may require a ≥ 10 -fold enzyme concentration to cleave a specific genomic sequence to the same extent as when the region is "open" or expressed. Thus, it may require a large difference in bulk chromatin MNase digestion in Basic Protocols 1 and 2 for two samples to exhibit the same extent of cleavage of a specific target gene.

For LM-PCR mapping of nucleosome boundaries, it is critical to first define the general region to be analyzed by mapping MNase cleavages at the Southern blot level of resolution. Chromatin digest samples that just begin to exhibit a laddering band pattern on a Southern blot (see Fig. 21.1.1A, first lane or two of chromatin sample 2) serve as the best substrates for the modified LM-PCR protocol; i.e., under-digested DNA rather than over-digested.

Once a set of genomic DNA samples is obtained where the extents of bulk chromatin

cleavage by MNase are known to give interesting, chromatin-specific results, it is convenient to use these samples as standards for subsequent experiments. That is, 0.5- μ g aliquots of the MNase digested chromatin samples from the known experiment can be run on a 1.2% agarose minigel alongside DNA aliquots from a new chromatin experiment from cells under similar physiological conditions. The gel can be stained with ethidium bromide and samples in the new chromatin experiment which exhibit the same extent of bulk genomic digestion, or laddering, can often be assumed to be worthy of further analysis by the aforementioned hybridization approaches.

Anticipated Results

For an extensive MNase digest, it is critical to compare the intensity of mononucleosome-sized bands hybridizing to DNA probes from different parts of the genome. For two probes of comparable specific activity, if the hybridizing bands are of similar intensity, it indicates that the same fraction of target sequences in the population were nucleosomal. Alternatively, if the hybridizing band is markedly more intense for one probe than for another, then more gene copies in the population hybridizing to the first probe were nucleosomal than in the population hybridizing to the second probe.

It is important to note that nucleosomes do not have to be phased, or positioned, over a DNA sequence in chromatin in order to give clear hybridization signals indicating that the sequence is nucleosomal. Nucleosomes that are randomly positioned over the region spanned by an oligonucleotide will, on average, yield a nucleosomal-sized band upon hybridization to extensively digested chromatin. Thus, if indirect end-labeling analysis indicates a lack of nucleosome phasing over a region, it is inappropriate to conclude that the region is free of nucleosomes. The extensive MNase digestion assay can address this issue.

For a partial MNase digest and indirect end-labeling analysis, it is critical that each sub-band generated by MNase treatment has one end resulting from restriction enzyme cleavage. If the chromatin sample generates a ladder of hybridizing sub-bands that differ in size by \sim 180 to 200 bp, it suggests the presence of a phased nucleosome array. In this case, the size of a sub-band indicates the distance between the restriction enzyme cleavage site and the position of double-stranded MNase cleavage. The hybridizing ladder of bands should not be smaller than the indirect end-labeled probe; if

it is, the band sizes are probably not generated by restriction enzyme cleavage on one side and thus fragment endpoints cannot be predicted. Ideally, internal marker controls are chosen so that the marker bands occur in the region of MNase-generated sub-bands, so that the positions of MNase cleavage can be determined with accuracy. A high level of accuracy in mapping MNase cleavage sites in chromatin, by this method, is useful for planning a primer probe strategy for LM-PCR analysis.

The clusters of MNase cleavages that define apparent linker regions between nucleosomes, by indirect end-labeling, should be evident in the LM-PCR analysis. That is, what appears as a single sub-band on a Southern blot should appear as a cluster of bands on an LM-PCR sequencing gel at about the positions that are predicted by the internal control marker on the Southern blot. Additional bands will probably be seen in the LM-PCR analysis, but if they do not occur in a cluster they may be insufficient to generate a sub-band on the Southern blot. It is worth noting that further evidence, such as may be provided by protein-DNA cross-linking experiments, should be obtained before concluding that an apparent "nucleosomal" ladder of MNase-generated bands is indeed caused by histone octamers on the DNA.

Time Considerations

Cell permeabilization (Basic Protocol 1) is much easier and quicker to perform than nuclear isolation (Basic Protocol 2). The entire cell permeabilization and MNase digestion procedure should take <1 hr. By contrast, for nuclear isolation, it can take several hours to prepare the solutions, and the procedure itself will take ≥ 4 hr, depending upon the number of tissues assayed and the difficulty encountered in releasing nuclei from the cells. One should plan an entire day for Basic Protocol 2. MNase digestion of purified genomic DNA (Basic Protocol 3) should take <1 hr to set up and perform.

When purifying DNA from chromatin digestions (Support Protocol 1), patience is a virtue. The organic extractions are time consuming and careful handling of the material is critical to prevent DNA loss. It may take 30 to 60 min just to precipitate, spool out, and resuspend the DNA; thus, this protocol could take at least several hours on the first day. After that the DNA must be allowed to dissolve overnight, and then quantitated and run on a minigel to assess quality (an additional several hours) on the second day.

When analyzing nuclease cleavages by Southern blot hybridization, any restriction digests that are done (as for a partial MNase digest) should be begun in the morning. After allowing the digests to incubate all day, the gel can be loaded and run overnight. If no restriction digests required (as with an extensive MNase digest), the gel can simply be loaded late in the day and run overnight for agarose, or run several hours, as appropriate, for polyacrylamide. On the second day, the gel should be stained with ethidium bromide, photographed, prepared for blotting, and then blotted overnight. On the third day, the blot should be baked or UV-cross-linked and pre-hybridized in the morning, the probe labeled in the afternoon, and hybridization performed overnight. On the fourth day, the blot should be washed and exposed to a phosphorimager overnight, and finally exposed to X-ray film for 2 to 10 days, depending upon the signal strength.

With nucleosome mapping by LM-PCR, eliminating steps 1 to 3 of the conventional protocol (*UNIT 15.3*) makes the procedure considerably simpler (see Fig. 20.1.2). The phosphorylation reaction and subsequent preparative work in Support Protocol 3 should take a few hours.

Literature Cited

- Felsenfeld, G. 1992. Chromatin as an essential part of the transcriptional mechanism. *Nature* 355:219-222.
- Hewish, D.R. and Burgoyne, L.A. 1973. Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochem. Biophys. Res. Comm.* 52:504-510.
- Kornberg, R.D., LaPointe, J.W., and Lorch, Y. 1989. Preparation of nucleosomes and chromatin. *Methods Enzymol.* 170:3-14.
- Liu, J.-K., Bergman, Y., and Zaret, K.S. 1988. The mouse albumin promoter and a distal upstream site are simultaneously DNaseI-hypersensitive in liver chromatin and bind similar liver-abundant factors in vitro. *Genes Dev.* 2:528-541.
- McPherson, C.E., Shim, E.-Y., Friedman, D.S., and Zaret, K.S. 1993. A tissue-specific enhancer and bound transcription factors existing in a precisely positioned nucleosome array. *Cell* 75:387-398.
- Pfeifer, G.P. and Riggs, A.D. 1991. Chromatin differences between active and inactive X chromosomes revealed by genomic footprinting of permeabilized cells using DNase I and ligation-mediated PCR. *Genes Dev.* 5:1102-1113.
- Shim, E.Y., Woodcock, C., and Zaret, K.S. 1998. Nucleosome positioning by the winged-helix transcription factor HNF3. *Genes Dev.* 12:5-10.
- van Holde, K.E. 1989. (pp. 497). Chromatin. Springer-Verlag, New York.
- Wu, C. 1980. The 5' ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I. *Nature* 286:854-860.
- Zhang, L. and Gralla, J.D. 1990. In situ nucleoprotein structure involving origin-proximal SV40 DNA control elements. *Nucl. Acids Res.* 18:1797-1803.

Contributed by Ken Zaret
Fox Chase Cancer Center
Philadelphia, Pennsylvania

Separation of Histone Variants and Post-Translationally Modified Isoforms by Triton/Acetic Acid/Urea Polyacrylamide Gel Electrophoresis

Due to their similarities in size and charge, complete resolution of histones by electrophoresis poses a considerable challenge. The addition of nonionic detergents to the traditional acetic acid/urea (AU) polyacrylamide gel electrophoresis (PAGE) system has afforded an excellent method to separate not only the different modified forms of histones, but also the primary sequence variant subtypes of selected histone species; it is widely used to separate histones with varying levels of acetylation. This unit describes the use of gels containing the nonionic detergent Triton X-100 (Triton is a registered trademark of Rohm and Haas), referred to as Triton/acetic acid/urea (TAU) polyacrylamide gels, for analysis of histones (see Basic Protocol). Also included are support protocols detailing several accessory techniques: assembly of gel plates for the TAU gel (see Support Protocol 1), preparation of histones from isolated nuclei in a solubilized form amenable to electrophoresis (see Support Protocol 2), and electrophoretic transfer of proteins from these gels to PVDF membranes (see Support Protocol 3).

TRITON/ACETIC ACID/UREA (TAU) POLYACRYLAMIDE GEL ELECTROPHORESIS FOR ANALYSIS OF HISTONES

BASIC PROTOCOL

The basic protocol describes the preparation and electrophoresis of one 16 × 22-cm (~60 ml) TAU-polyacrylamide gel (15% acrylamide/0.37% Triton X-100/6 M urea/0.9 M acetic acid; final solution volume 70 ml). The gel polymerizes overnight and is preelectrophoresed during the following day. Histones are resolved overnight and proteins can be visualized through a variety of methods (see Support Protocol 3 and *UNIT 10.6*).

Materials

- Glacial acetic acid (17.4 M)
- TEMED (*N,N,N',N'*-tetramethylethylenediamine)
- 9.5 M urea (see recipe)
- 10% (w/v) ammonium persulfate (APS; prepare fresh for each gel)
- 10% Triton X-100 (protein grade; Calbiochem)
- 63.5% acrylamide/0.4% bisacrylamide (see recipe)
- Moderately warm water (~45°C) in a beaker
- Running buffer (0.9 M acetic acid; see recipe)
- Scavenger solution (2-mercaptoethylamine; see recipe)
- 1000- and 100-ml sidearm flasks
- Gel-casting assembly (see Support Protocol 1)
- Vertical gel electrophoresis apparatus to accommodate 16 × 22-cm glass plates (see Support Protocol 1)
- Power supply capable of running in constant current mode, with two leads
- 1.5-mm-thick 16-well gel comb
- Disposable syringes with 23-G needles
- Disposable syringe with 18-G needle bent into a U shape
- Hamilton syringe
- Additional reagents and equipment for degassing gel solutions (*UNIT 10.2*)

Chromatin Assembly and Analysis

21.2.1

Prepare TAU gel

1. Combine the following in a 1000-ml sidearm flask:

3.62 ml glacial acetic acid
1.88 ml H₂O
0.329 ml TEMED.

2. Swirl gently to mix, then add:

44.21 ml 9.5 M urea
0.875 ml 10% APS.

3. Degas solution completely (*UNIT 10.2*).

4. Add 2.59 ml of 10% Triton X-100 to the flask and degas a second time.

As the detergent will tend to foam during degassing, do not degas until bubbles are completely gone, only until they are reduced somewhat. If solution is degassed excessively the Triton X-100 will begin to crystallize on the sides of the flask.

5. Place 16.5 ml of 63.5% acrylamide/0.4% bisacrylamide in a 100-ml sidearm flask. Place the flask in a beaker containing moderately warm water (~45°) immediately prior to degassing.

6. Gently degas the acrylamide solution in short bursts, swirling constantly in a warm water bath to prevent precipitation of the acrylamide. Stop when the most vigorous bubbling has ceased (do not degas too long).

7. Pour the degassed Triton X-100/urea into the flask of acrylamide and swirl gently to mix.

8. Immediately pour the gel solution into preset gel assembly set up inside a resealable plastic bag (see Support Protocol 1). Fill gel plates to within 1 to 1.5 cm of the top. Place a 1.5-mm-thick 16-well gel comb between plates and insert it until the gel solution comes halfway up the teeth of the comb.

The authors use a 16-well comb with 6 mm-wide teeth. Observe the gel closely as it polymerizes over the next 30 to 60 min, continually pushing the comb down between the plates to keep the bottom surface of the teeth in contact with the polymerizing wells.

9. Close the resealable plastic bag (with a small amount of distilled water in the bottom to prevent desiccation of the gel) and allow the gel to polymerize overnight.

Gels plates are held together with 5-cm-wide stainless steel binder clips (available from Research Products International). By standing the gel assembly upright on the bottom clips, the gel is raised above the water in the bottom of the plastic bag (see Support Protocol 1).

10. Carefully lift the polymerized gel from the plastic bag, and remove the bottom spacer. Place the assembled plates into the gel electrophoresis apparatus, with the notched plate facing the upper buffer chamber. Fasten the gel assembly to the apparatus using 5-cm-wide binder clips, positioned over the upper chamber gasket to ensure a good seal.

11. Fill both chambers with running buffer. Remove bubbles from the bottom surface of the gel using a syringe with an 18-G needle bent into a U shape. Gently remove the comb from the gel assembly, and carefully straighten any bent well walls with the needle of a disposable syringe (23-G needle). Rinse the wells with the same syringe filled with running buffer.

Electrophorese the gel

12. Connect gel to power supply with the positive (red) electrode placed uppermost. Preelectrophorese the gel at 130 V (constant voltage) in running buffer until the current no longer falls (~4 to 5 hr). During this time, periodically shut off the power supply and rinse the wells of the gel using a disposable syringe (with 23-G needle) filled with running buffer. Also, remove any bubbles that have accumulated at the bottom surface of the gel using a syringe with an 18-G needle bent into a U shape.

Removing bubbles will help to decrease the time required to reach constant current during preelectrophoresis.

13. Remove the running buffer from both chambers of the gel apparatus. Add water to the top chamber and rinse the wells. Pour the water from the chamber and remove excess water from the wells using a disposable syringe with 23-G needle, being careful not to pierce the bottom of the wells with the syringe tip.
14. Distribute scavenger solution evenly throughout the wells using a 1-ml pipettor. Slowly add fresh running buffer to the top chamber, being especially careful when the buffer begins to run into the wells; the running buffer should overlay the scavenger. Fill the lower chamber with buffer and continue electrophoresing the gel at 300 V for 3 hr.

Because the gel is polymerized using an oxidizing agent as a catalyst, the application of a "scavenger," or reducing agent, is required prior to the introduction of the sample. Scavenging the gel after preelectrophoresis inhibits oxidation of proteins during electrophoresis (see Background Information).

15. Repeat step 10.
16. Add fresh running buffer to both chambers and rinse the wells with a syringe filled with running buffer. Load samples (usually 2 to 6 μ l) into wells with a Hamilton syringe, layering the sample directly onto the bottom surface of the well. Rinse each well again with a syringe filled with running buffer immediately before loading each sample, to remove any leaching urea.

Typically ~50 to 75 μ g total histone is loaded per well when gels are to be stained with Coomassie blue. As there is no stacking gel in this system, sample loads should be kept <10 μ l.

It is helpful to load 0.1 to 0.15 ml acetic acid/urea sample buffer (see recipe) into an empty well at the edge of the gel to monitor the migration of the dye front. Often, the volume of the sample is too small for the blue dye front (used in determining when to stop electrophoresis; see below) to remain visible.

17. Electrophorese the gel at 400 V for 15 min. Lower the voltage to 200 V and run until the blue dye component of the methyl green runs off the gel (overnight, or ~15 to 17 hr), or longer (see Critical Parameters and Troubleshooting).
18. Visualize the gel by either Coomassie brilliant blue staining (see UNIT 10.6) or electrophoretic transfer and immunoblotting (see Support Protocol 3 and UNIT 10.8).

ASSEMBLY OF GEL PLATES

This support protocol represents one method of assembling gel plates for a TAU gel. Note that glass plates for TAU gels are longer (22 cm long; 20 cm separating length) than those used in typical SDS-PAGE. This is required to resolve histone primary sequence variants, as well as isoforms modified post-translationally (e.g., by phosphorylation or acetylation). Providing that the length requirement is met, other gel systems (including those that are commercially available) can be substituted.

Materials

- 95% ethanol
- 1% (w/v) agarose, melted
- 16 × 22-cm notched glass plate
- 16 × 22-cm glass plate
- Two 22-cm-long × 1.5-mm-thick Teflon spacers
- 18-cm-long × 1.5-mm-thick Teflon spacer
- Six 5-cm-wide stainless steel binder clips (Research Products International)
- Resealable plastic bag large enough to contain gel-casting assembly

1. Clean surfaces of plates and spacers with 95% ethanol. Use a clean razor blade to remove any possible debris.
2. Place the 18-cm-long spacer (bottom spacer) parallel to the bottom of one glass plate ~0.5 cm in from the edge of the plate.
3. Place the two 22-cm-long spacers (side spacers) parallel to either side of the glass plate ~0.5 cm in from the edge of the plate. Place the bottom edge of each side spacer flush with the top edge of the bottom spacer.
4. Clamp the two gel plates together with binder clips: first place two clips across the bottom of the gel, then two on either side of the gel. If necessary, reposition the spacers so that they are in their proper places. Make sure that the pressure point of the binder clips on the glass plates is directly over the spacers.

It is helpful to assemble the plates with the bottom ~5 cm hanging off the edge of the bench; that way the clips can be placed on the bottom edge of the gel without disturbing the position of the plates. Once the plates are secured with the two bottom clips, they can be picked up without disturbing the position of the plates to attach the final four clips on the sides. Often, the position of the side spacers needs to be adjusted before the side clips are added.

5. Once the clips are on the glass plates, snap the handles over onto the plates so that the gel can be stood up on the base of the binder clips.
6. Seal the sides and bottom of the gel plates with melted 1% agarose to guard against leakage of the gel solution. With a Pasteur pipet, add the agarose to the ~0.5-cm space between the outer edge of the spacers and the edge of the plates.
7. Once the agarose has set, stand the gel assembly upright on its bottom binder clips in a resealable plastic bag with a small amount of water in the bottom to prevent desiccation of the gel. The clips will hold the gel assembly above the surface of the water. Pull the opening of the bag down the sides of the gel to provide access for addition of the gel solution.

HISTONE ISOLATION FROM PREPARED NUCLEI

Histones to be subjected to TAU-PAGE can be extracted from isolated nuclei by a low-concentration acid extraction. The resulting acid-soluble nuclear proteins are then precipitated with TCA and resuspended in acetic acid/urea sample buffer at the desired dilution.

Materials

Isolated nuclei (e.g., see *UNIT 12.1*)
Concentrated sulfuric acid (H₂SO₄; 18 M H₂SO₄) diluted to 0.2 M
100% trichloroacetic acid (TCA)
11.6 M hydrochloric acid (HCl)
100% acetone
Acetic acid/urea sample buffer (see recipe)

1. Resuspend isolated nuclei in 0.2 M H₂SO₄ at a concentration of ~2 mg DNA/ml. Let stand at least several hours, or preferably overnight, at 4°C.
Sonicking the nuclei 1 to 2 sec after suspension in the acid can increase the efficiency of extraction.
2. Microcentrifuge the residual nuclei 10 min at ~12,000 × 4°C, to clarify acid-soluble nuclear proteins (including histones). Transfer supernatant to a clean microcentrifuge tube.
3. Add 1/3 vol of 100% TCA to the acid-soluble nuclear proteins (25% TCA final) and mix well to precipitate the proteins. Let stand 30 min on ice.
4. Microcentrifuge mixture 10 min at ~12,000 × g, 4°C, to pellet proteins.
5. Wash protein pellet in 1 ml of 100% acetone/0.05 M HCl, vortexing to remove any histones from the sides of the tube, and microcentrifuge as in step 4.
6. Repeat wash in step 5, this time in 100% acetone.
7. Dry protein pellet in lyophilizer or Speedvac evaporator.
8. Resuspend protein pellets in acetic acid/urea sample buffer at desired concentration (typically 5 µg total histone/µl) for electrophoresis.

ELECTROPHORETIC TRANSFER OF TAU-POLYACRYLAMIDE GELS

Unlike SDS- and AU-polyacrylamide gels, untreated TAU gels will not directly transfer to a membrane. This support protocol describes an effective method of treating TAU gels for reliable transfer (Delcuve and Davie, 1992).

Materials

Wash buffers I and II (see recipes)
CAPS transfer buffer (see recipe)
100% methanol
Ponceau S stain (Sigma; optional)
Transfer membrane (e.g., Immobilon-P PVDF membrane, Millipore)
Filter paper cut to size of gel

1. Remove gel from plates and place it in a plastic dish large enough for the entire gel to lie flat.
2. Wash the gel twice with rocking for 30 min each in 250 ml wash buffer I.

SUPPORT PROTOCOL 2

SUPPORT PROTOCOL 3

Chromatin Assembly and Analysis

21.2.5

3. Place the dish containing the gel in a fume hood and wash the gel once with rocking for 30 min in 150 ml wash buffer II.

The fume hood is necessary because of the high concentration of 2-mercaptoethanol in wash buffer II.

4. In a small plastic dish, hydrate membrane in 100% methanol for 30 sec, then wash in water for 30 sec. Equilibrate in CAPS transfer buffer.
5. Cut a piece of filter paper to the size of the gel and prewet it with transfer buffer. Assemble gel sandwich on the wet filter paper, with the membrane positioned at the anode side of the gel.
6. Place the gel in a transfer apparatus in CAPS transfer buffer and transfer proteins to membrane at 30 V overnight with cooling to 4°C.

Following transfer, proteins can be visualized by staining the membrane with Ponceau S (Sigma).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Acetic acid, 0.9 M (running buffer)

36.2 ml glacial acetic acid
H₂O to 700 ml

Acetic acid/urea sample buffer

1.8 ml 10 M urea (5.8 M final)
0.160 ml glacial acetic acid (0.9 M final)
1 ml 50% glycerol/0.6% methyl green (see recipe; 16% glycerol/0.2% methyl green final)
0.15 ml 2-mercaptoethanol (4.8% final)

63.5% acrylamide/0.4% bisacrylamide

63.5 g acrylamide
0.4 g bisacrylamide
Dissolve in H₂O to 100 ml

CAPS, pH 10, 500 mM

27.66 g 3-(cyclohexylamino)-1-propanesulfonic acid (CAPS)
Adjust pH with 10 M NaOH
Add H₂O to 250 ml final
Destaining solution
800 ml 100% methanol (20% final)
400 ml glacial acetic acid (10% final)
H₂O to 4 liters

50% (v/v) glycerol/0.6% (w/v) methyl green

0.006 g methyl green (Sigma)
0.5 ml glycerol
0.5 ml H₂O

Scavenger solution

0.326 g 2-mercaptoethylamine (Sigma; 2.12 M final)
0.3 g urea (2.5 M final)
0.1 ml glacial acetic acid (0.87 M final)
1.9 ml H₂O

Transfer buffer

250 ml 500 mM CAPS, pH 10 (see recipe; 25 mM final)

1 liter 100% methanol (20% final)

H₂O to 5 liters

Degas buffer before adding to transfer apparatus.

Urea solutions

9.5 M urea: Dissolve 28.5 g urea in H₂O to 50 ml.

10 M urea: Dissolve 30.03 g urea in H₂O to 50 ml.

Wash buffer I

1.4 ml glacial acetic acid (50 mM final)

2.5 g SDS (0.5% final)

H₂O to 500 ml

Wash buffer II

1.135 g Tris base (Sigma; 62.5 mM Tris final,)

3.45 g SDS (2.3% final)

1.5 ml 14.3 M 2-mercaptoethanol (1% final)

H₂O to 150 ml

Mix all components. Working in fume hood and using a hand-held pH meter, adjust pH to 6.8 with HCl.

COMMENTARY

Background Information

The fundamental repeating unit of chromatin, the nucleosome, contains two molecules each of the core histones (H2A, H2B, H3, H4), along with a single molecule of linker histone (H1, H5, H1^o, etc.). Histones are positively charged, low-molecular-weight proteins that are subject to numerous post-translational modifications, including acetylation, methylation, phosphorylation, poly(ADP-ribosylation), and ubiquitination (van Holde, 1988). Such modifications are thought to play a role in modulating chromatin structure, and have been linked to the regulation of gene activity (Wolffe, 1995; Mizzen and Allis, 1998). In addition, histones are found in a variety of primary sequence variants (e.g., H3.1, H3.2, H3.3), whose patterns of expression are developmentally regulated (van Holde, 1988; Wolffe, 1995). Through sequence variant heterogeneity and selective post-translational modification, an extraordinary degree of variability in nucleosome composition can be generated.

It is possible to resolve the five major histones by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), but their modified forms cannot be resolved by this method. Acetic acid/urea (AU)–PAGE, as described by Panyim and Chalkley (1969), can separate the post-translationally modified iso-

forms of histones, but the level of detailed resolution of primary sequence variants that can be achieved by this method is limited. The introduction of nonionic detergents such as Triton X-100 to the AU-PAGE system by Zweidler (1978) was an innovative step permitting the simultaneous resolution of histone variants and isoforms modified post-translationally (e.g., by acetylation or phosphorylation; Franklin and Zweidler, 1977).

The separation of histones in these gels is improved by a differential reduction in the electrophoretic mobility that is a direct result of the presence of nonionic detergents. The detergents cause the formation of mixed micelles between the detergent and the hydrophobic regions of the histones, a process that is extremely sensitive to minor differences in the hydrophobicity of each protein (Zweidler, 1978). The urea in the gel acts as a denaturing agent; the acetic acid helps separate histones on the basis of charge. As with the standard acid/urea system, TAU gels can resolve the difference in charge caused by a single functional group (Fig. 21.2.1).

An additional characteristic of this gel technique that contributes to its separation capabilities is the requirement for a “scavenger,” or reducing agent. Because the gels are polymerized using an oxidizing agent as a catalyst, scavenging after preelectrophoresis and prior

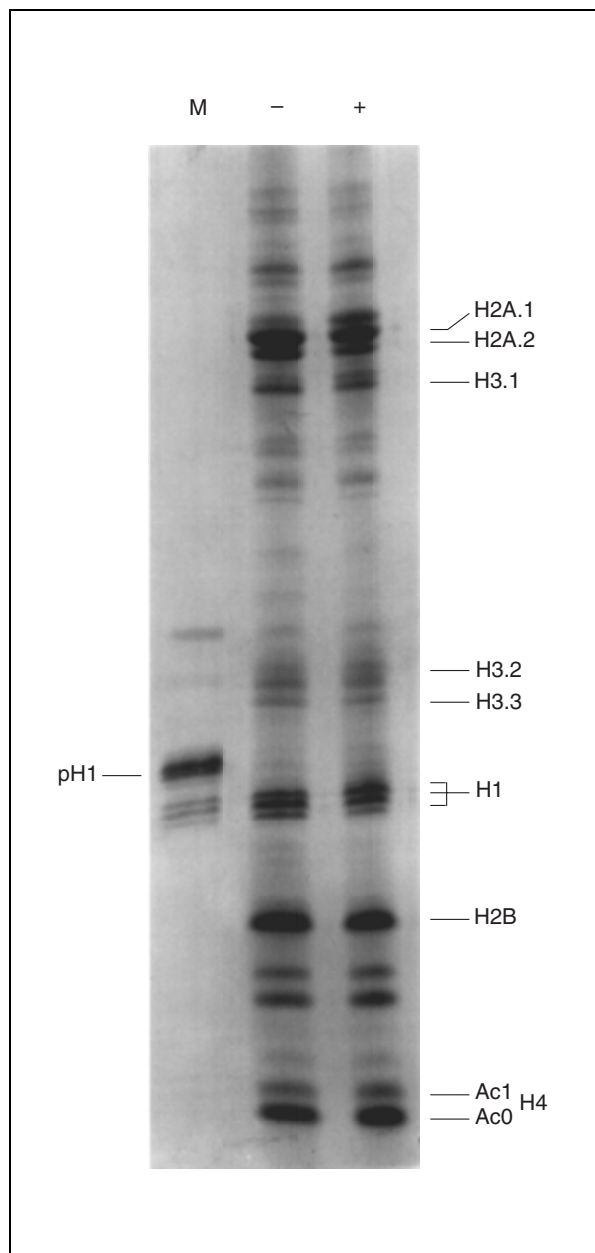


Figure 21.2.1 HeLa cells were synchronized with a double thymidine block procedure (2.7 mM thymidine; Peterson and Anderson, 1964; Knehr et al., 1995) and released into S phase for 3 hr. Released cells were then incubated 60 min in the absence (-) or presence (+) of okadaic acid (500 nM; Calbiochem), an inhibitor of cellular phosphatases. Acid-soluble nuclear proteins and isolated histone H1 were subjected to electrophoresis in a TAU gel and analyzed by Coomassie blue staining. Lane M contains hyperphosphorylated H1 (pH1) from cells arrested in M phase by treatment with Colcemid (demecolcine; 0.06 $\mu\text{g}/\text{ml}$). The un- (Ac 0) and monoacetylated (Ac 1) isoforms of histone H4 are indicated.

to the introduction of the sample is required to prevent oxidation of the proteins during electrophoresis, which could change their hydrophobic properties and alter their migration (Zweidler, 1978).

The components of TAU-polyacrylamide gels make electrophoretic transfer of the resolved proteins for immunochemical analysis troublesome. In particular, the presence of a nonionic detergent in the gel impairs the migration of the proteins out of the gel and on to the transfer membrane. The electrophoretic transfer method presented in this unit was developed by Delcuve and Davie (1992) based on the method originally presented for TAU gel transfer by Waterborg and Harrington (1987). First,

this method increases the efficiency of transfer by using an alkaline transfer buffer; alkaline transfer buffer has been found to greatly increase the efficiency of transfer of strongly basic proteins, such as histones (Delcuve and Davie, 1992). Second, this method also alleviates the difficulties presented by the presence of Triton X-100, which interferes with the binding of histones to transfer membranes (probably due to complex formation between histones and Triton; Zweidler, 1978; Waterborg and Harrington, 1987). The transfer method solves this problem by treating the gel with a variety of solutions prior to transfer. These solutions contain SDS, which displaces the Triton X-100 present in the gels and gives the

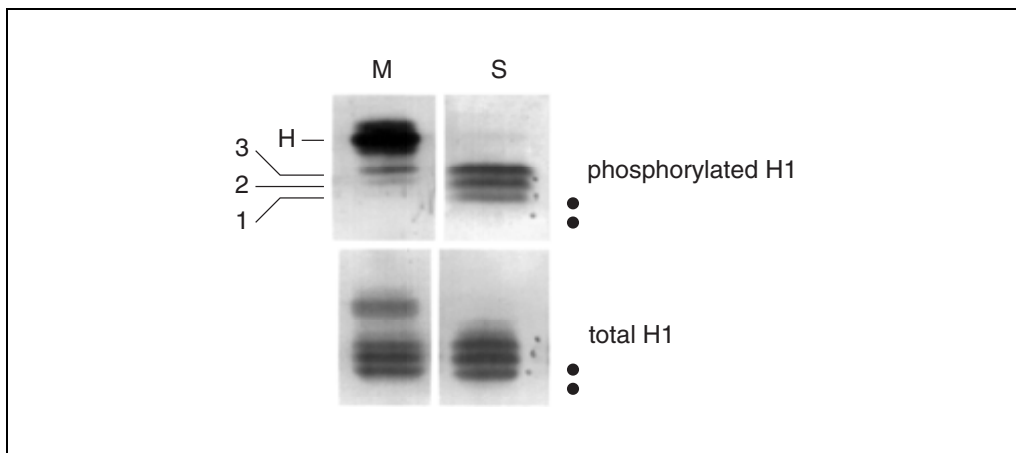


Figure 21.2.2 HeLa cells were synchronized with a double thymidine block procedure and released into S phase for 4 hr (lane S). Isolated histone H1 was subjected to electrophoresis in a TAU gel; lane M contains hyperphosphorylated H1 from cells arrested in M phase by treatment with Colcemid (demecolcine). The gel was transferred electrophoretically to Immobilon-P membrane (Millipore) and probed with antibodies specific for either phosphorylated H1 (a gift of Dr. C. David Allis) or total H1. Bound antibodies were detected by a secondary alkaline phosphatase color reaction. Bullets indicate the positions of the unphosphorylated forms of H1, as determined by staining the membrane with Ponceau S. The mono- (1), di- (2), tri- (3), and hyperphosphorylated (H) isoforms of H1 are indicated.

histones the negative charge required for transfer (Delcuve and Davie, 1992). Figure 21.2.2 shows immunoblots of histone H1 resolved in a TAU-polyacrylamide gel and transferred by this method.

Critical Parameters and Troubleshooting

Maintaining contact between the bottom surface of the comb and the interface of the gel as it polymerizes is vital to creating wells with sharp, even bottoms. A flat well surface is also important for establishing good sample migration. Because there is no stacking gel present in this system, the migration of the sample is strongly affected by how evenly it lies prior to electrophoresis, and by the sample concentration (i.e., loading volume).

Increased resolution of protein bands can also be achieved by maximizing the distance the proteins migrate through the gel—which obviously will maximize their separation. Histone H4 migrates the fastest of the five histones in a TAU gel; when analyzing other histones that migrate more slowly, it is advantageous to electrophorese the gel for an extended period of time after the blue dye component of methyl green runs off the gel. The extra electrophoresis time is determined on the basis of experience: for example, the modified forms and subtypes of histone H1 seen resolved in Figure 21.2.2 were electrophoresed in a TAU gel for an additional 5 hr after the blue dye front ran off the gel.

Anticipated Results

Electrophoretic separation of histones on a TAU-polyacrylamide gel is a highly versatile method for analyzing histones by polyacrylamide gel electrophoresis. With an appropriate electrophoresis time, this system will provide satisfactory separation of different modified forms of histones (e.g., acetylated isoforms of histone H4 or phosphorylated isoforms of histone H1). TAU gels will also resolve sequence-variant subtypes of selected histones. In Figure 21.2.1, distinct migration of the different subtypes of histone H3 and histone H2A can be observed.

Time Considerations

The time considerations for running a TAU-polyacrylamide gel are as follows: polymerization of the gel takes a minimum of several hours, but it is recommended that the process be allowed to continue overnight. Preelectrophoresis of the gel takes ~4 to 5 hr, and is usually started early in the morning following overnight polymerization. Scavenging of the gel is then allowed to proceed for 3 hr, and samples are electrophoresed overnight for ~15 to 17 hr, until the blue dye migrates off the bottom of the gel. Electrophoresis can be terminated at this time, or allowed to proceed for several hours if additional separation of more slowly migrating histones is desired.

Literature Cited

- Delcuve, G.P. and Davie, J.R. 1992. Western blotting and immunochemical detection of histones electrophoretically resolved on acid-urea-triton and sodium dodecyl sulfate-polyacrylamide gels. *Anal. Biochem.* 200:339-341.
- Franklin, S.G. and Zweidler, A. 1977. Non-allelic variants of histones 2a, 2b, and 3 in mammals. *Nature* 266:273-275.
- Knehr, M., Poppe, M., Enulescu, M., Eickelbaum, W., Stoehr, M., Schroeter, D., and Paweletz, N. 1995. A critical appraisal of synchronization methods applied to achieve maximal enrichment of HeLa cells in specific cell cycle phases. *Exp. Cell Res.* 217:546-553.
- Mizzen, C.A., and Allis, C.D. 1998. Linking histone acetylation to transcriptional regulation. *Cell. Mol. Life Sci.* 54:6-20.
- Panyim, S. and Chalkley, R. 1969. High resolution acrylamide gel electrophoresis of histones. *Arch. Biochem. Biophys.* 130:337-346.
- Peterson, D.F. and Anderson, E.C. 1964. Quantity production of synchronized mammalian cells in suspension culture. *Nature* 203:642-643.
- van Holde, K.E. 1988. Chromatin. Springer-Verlag, New York.

Waterborg, J.H. and Harrington, R.E. 1987. Western blotting of histones from acid-urea-triton- and sodium dodecyl sulfate-polyacrylamide gels. *Anal. Biochem.* 162:430-434.

Wolffe, A.P. 1995. Chromatin. Academic Press, San Diego.

Zweidler, A. 1978. Resolution of histones by polyacrylamide gel electrophoresis in presence of nonionic detergents. *Methods Cell Biol.* 17:223-233.

Key References

Zweidler, 1978. See above.

Initial description of Triton/acetic acid/urea polyacrylamide gel electrophoresis.

Delcuve and Davie, 1992. See above.

Description of the development of a method for efficiently transferring the Triton/acetic acid/urea polyacrylamide gels to membrane.

Contributed by Colleen A. Ryan and
Anthony T. Annunziato
Boston College
Chestnut Hill, Massachusetts

Chromatin Immunoprecipitation for Determining the Association of Proteins with Specific Genomic Sequences In Vivo

Chromatin immunoprecipitation (ChIP) is a powerful and widely applied technique for detecting the association of individual proteins with specific genomic regions in vivo. In this technique, live cells are treated with formaldehyde to generate protein-protein and protein-DNA cross-links between molecules in close proximity on the chromatin template in vivo. A whole-cell extract is prepared, and the cross-linked chromatin is sheared by sonication to reduce average DNA fragment size to ~500 bp. The resulting material is immunoprecipitated with an antibody against a desired protein, modified (e.g., acetylated, phosphorylated, methylated) peptide, or epitope (in situations where the protein of interest is epitope-tagged). DNA sequences that directly or indirectly cross-link with a given protein (or modified variant) are selectively enriched in the immunoprecipitated sample. Thus, the method is not restricted to sequence-specific DNA-binding proteins. Reversal of the formaldehyde cross-linking by heating permits the recovery and quantitative analysis of the immunoprecipitated DNA. The amounts of specific genomic regions in control and immunoprecipitated samples are determined individually by quantitative PCR. The fold enrichment of certain chromosomal sequences (e.g., presumed binding sites) relative to other chromosomal sequences (e.g., presumed nonbinding sites) provides quantitative information about the relative level of association of a given protein with different genomic regions. Protein association with specific genomic regions can be performed under a variety of conditions (e.g., environmental change, cell-cycle status) and/or in wild-type versus mutant strains. Furthermore, as formaldehyde inactivates cellular enzymes essentially immediately upon addition to cells, ChIP provides snapshots of protein-protein and protein-DNA interactions at a particular time point, and hence is useful for kinetic analysis of events occurring on chromosomal sequences in vivo. In addition, ChIP can be combined with microarray technology to identify the location of specific proteins on a genome-wide basis (see Commentary). This unit describes the ChIP protocol for cells of the baker's yeast *Saccharomyces cerevisiae* (see Basic Protocol 1); however, it is also applicable to other organisms, although some organism-specific modifications related to cell lysis and sonication are necessary. The ChIP protocol for mammalian cells (see Basic Protocol 2) is adapted from the yeast procedure. A protocol for eluting immunoprecipitated protein-DNA complexes is also provided (see Alternate Protocol 1). As an alternative to gel electrophoretic analysis of the PCR products, a quantitative PCR analysis in real time with SYBR Green is also provided (see Alternate Protocol 2). Finally, a procedure to map protein binding sites using ChIP and quantitative PCR scanning is provided (see Alternate Protocol 3).

CHROMATIN IMMUNOPRECIPITATION IN YEAST CELLS

In this protocol, *Saccharomyces cerevisiae* cells are cross-linked with formaldehyde, then harvested and sonicated to release the chromatin. The chromatin is then immunoprecipitated and purified, the cross-links to the formaldehyde are reversed, and the resulting DNA is analyzed by PCR and gel electrophoresis. The procedure is summarized in Figure 21.3.1.

BASIC PROTOCOL 1

Chromatin Assembly and Analysis

21.3.1

Contributed by Oscar Aparicio, Joseph V. Geisberg, Edward Sekinger, Annie Yang, Zarmik Moqtaderi, and Kevin Struhl

Current Protocols in Molecular Biology (2005) 21.3.1-21.3.33
Copyright © 2005 by John Wiley & Sons, Inc.

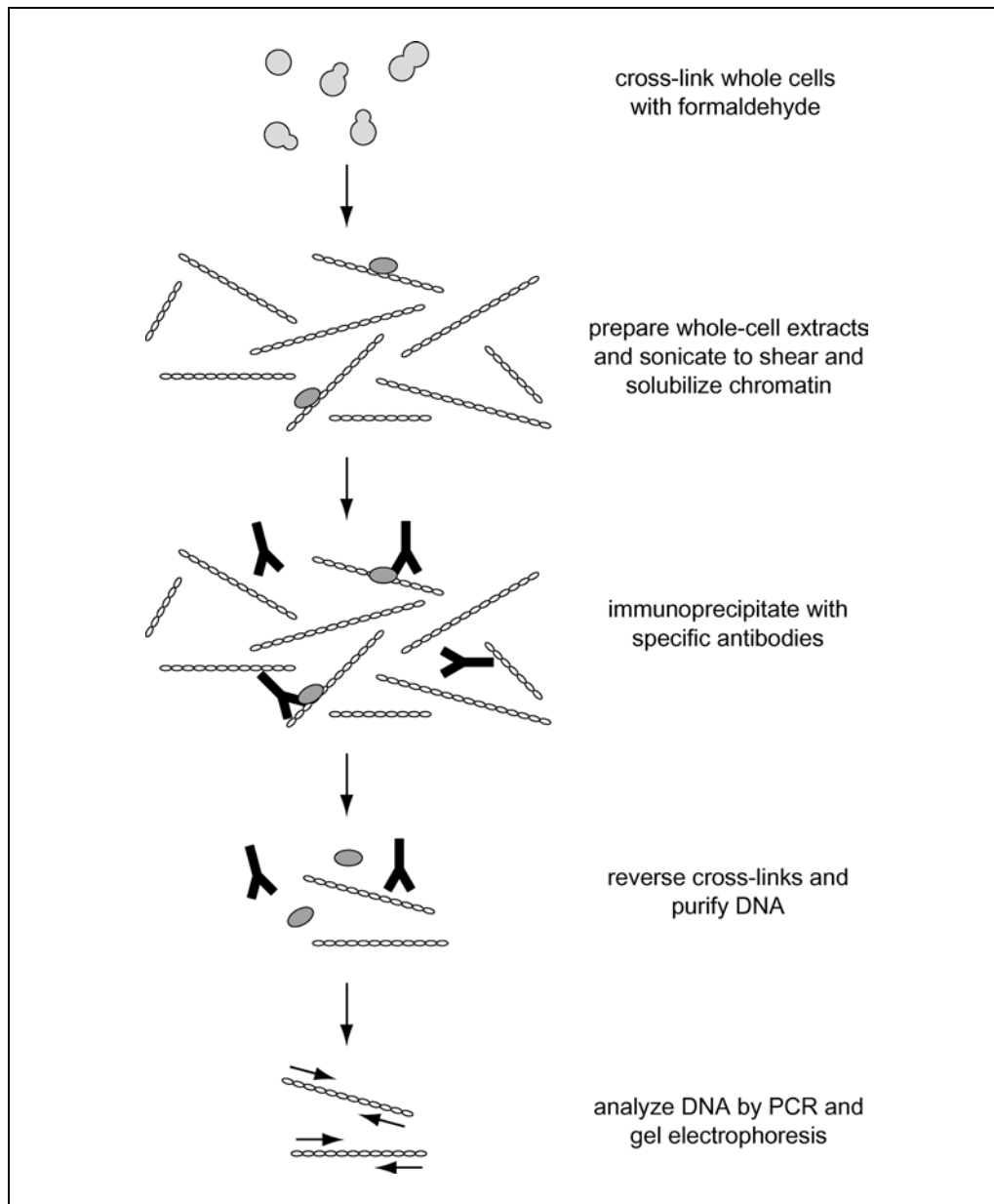


Figure 21.3.1 Scheme for chromatin immunoprecipitation from yeast whole-cell extracts.

Materials

- Saccharomyces cerevisiae* cells (Chapter 13) to be studied
- 37% formaldehyde: store up to 1 year at room temperature
- 2.5 M glycine, heat sterilized
- TBS (APPENDIX 2), ice cold
- FA lysis buffer with and without 2 mM PMSF (see recipe), ice cold
- ChIP elution buffer (see recipe)
- 20 mg/ml Pronase (Roche) in TBS; store up to 1 year at -20°C
- TE buffer, pH 7.5 (APPENDIX 2)
- 20 mg/ml DNase-free RNase A (UNIT 3.13)
- 10 \times loading buffer (UNIT 2.5A)
- Primary antibody against protein or epitope of interest
- 50% (v/v) protein A–Sepharose beads (Amersham Pharmacia Biotech) or equivalent in TBS

FA lysis buffer, room temperature
 FA lysis buffer/0.5 M NaCl
 ChIP wash buffer (see recipe)
 Primers (see Critical Parameters and Troubleshooting)
 3000 Ci/mmol [³²P]dATP
 2-ml screw-cap microcentrifuge tubes with (relatively) flat bottoms
 ~0.5-mm-diameter silica-zirconia (BioSpec; preferred) or glass beads
 Mini bead beater (BioSpec; preferred) or individual or multivortexer
 5-ml syringe
 15-ml conical tubes, disposable
 25-G needles
 Sonicator with microtip probe (e.g., Branson Sonifier 250)
 End-over-end rotator
 0.5-ml PCR tube
 Spin-X centrifuge-tube filter (e.g., Corning)
 65°C water bath
 PCR-purification spin column (Qiagen)
 Software for analyzing PCR primers and products
 Additional reagents and equipment for growth of *Saccharomyces cerevisiae* cultures (UNITS 13.1 & 13.2), phenol/chloroform extraction and ethanol precipitation (UNIT 2.1A), PCR (UNITS 15.1 & 15.7), agarose gel electrophoresis (UNIT 2.5A), and nondenaturing acrylamide gel electrophoresis (UNIT 2.7)

Cross-link protein-DNA complexes in vivo

1. For each sample, grow 200 ml *Saccharomyces cerevisiae* to OD₆₀₀ = 0.6 to 0.8 (UNITS 13.1 & 13.2).

CAUTION: Keep cultures covered or work in a fume hood to avoid noxious formaldehyde fumes.

The volumes of culture can be reduced (20 ml is a reasonable minimum) or increased depending on need. Typically, 20 to 40 ml yeast is used for an individual immunoprecipitation, so the 200-ml volume permits multiple immunoprecipitations from the same cells. This is particularly useful for experiments involving the analysis of multiple factors or for carrying out independent immunoprecipitations involving the same factor for data reproducibility.

2. Add 5.5 ml of 37% formaldehyde (1% final). Cross-link 15 to 20 min at room temperature by occasionally swirling flask or shaking slowly on a platform.
3. Add 30 ml heat-sterilized 2.5 M glycine and incubate an additional 5 min at room temperature.

Glycine stops the cross-linking by reacting with formaldehyde.

Harvest cells

4. Centrifuge cells 5 min at 2500 × g, 4°C. Discard supernatant into a chemical waste container and resuspend pellet in 50 to 200 ml ice-cold TBS. Repeat once.
5. Centrifuge cells for a third time 5 min at 2500 × g, 4°C. Discard supernatant and resuspend cells in 10 ml ice-cold FA lysis buffer.
6. Pellet cells by centrifuging in a benchtop centrifuge 5 min at 3000 rpm, 4°C. Discard supernatant.

The cells can remain on ice for a few hours while other samples are being collected so that all samples may be processed as a group from this point onward. Alternatively, the cells may be frozen in liquid nitrogen or a dry ice/ethanol bath and stored up to several months

at -80°C . This is particularly helpful if multiple samples are being generated during a time-course experiment. If cells are frozen, they must be thawed on ice before continuing with the procedure.

Lyse cells and isolate chromatin

For lysis using a mini bead beater (preferred)

- 7a. Resuspend the cell pellet in 1 ml ice-cold FA lysis buffer/2 mM PMSF. Fill three-quarters of a 2-ml flat-bottomed screw-cap microcentrifuge tube with ~ 0.5 -mm-diameter silica-zirconia or glass beads. Add cells, taking care to avoid introduction of bubbles, and screw the cap on tightly. Make sure there are no leaks.

The mini bead beater is recommended, because it is more efficient at breaking cells (multiple samples can be broken simultaneously). Silica-zirconia beads are more efficient at breaking cells than glass beads and are also recommended. To facilitate cell breakage with the mini bead beater, it is important that the final suspension nearly fill the tube. Do not break > 160 OD_{600} units of cells (i.e., $< 5 \times 10^9$ cells) in a single 2-ml tube; for larger cultures, split the cells into multiple tubes.

- 8a. Lyse cells 3 min with a mini bead beater at maximum speed. Remove sample and incubate 1 min in an ice-water bath. Repeat five times for a total breakage time of 18 min.

This step assumes breakage with silica-zirconia beads. The cell breakage time with glass beads may be longer.

For lysis using an individual or multivortexer

- 7b. Resuspend in 250 μl FA lysis buffer/2 mM PMSF. Add 350 μl silica-zirconia or glass beads to a 2-ml microcentrifuge tube with relatively flat bottom. Add cells.

When using a multivortexer (or standard vortexer), it is important to keep the volume small as this improves cell breakage. Do not break > 160 OD_{600} units of cells (i.e., $< 5 \times 10^9$ cells) in a single 2-ml tube; for larger cultures, split the cells into multiple tubes.

- 8b. Vortex continuously on an individual or multivortexer 30 min at full speed, 4°C .

Success and reproducibility of the ChIP procedure is aided by complete (or near-complete) breakage of cells. In this regard, formaldehyde-cross-linked cells are considerably harder to break than untreated cells. The use of 1.5-ml microcentrifuge tubes with conical bottoms should be avoided because the narrow shape constricts bead movement, resulting in unequal lysis among samples. The 2-ml microcentrifuge tubes have a nearly flat bottom that allows the beads to vortex vigorously. The indicated vortexing or bead-beating conditions should be tested if a different device is used.

Isolate lysate

9. Cut a 5-ml syringe ~ 1 cm below the flared opening (i.e., where the plunger is inserted) with a razor. Insert the smaller portion into a 15-ml disposable conical tube so that the flared portion of the truncated syringe rests on top of the conical tube opening, forming a microcentrifuge-tube holder.
10. Invert the sample tube and punch a hole in the bottom with a 25-G needle. Place the sample tube into the syringe/conical tube and punch a hole in the top cover with the same needle.
11. Spin the assembly in a benchtop centrifuge 1 min at 1000 rpm, 4°C . Place the conical tube on ice. Discard the 2-ml centrifuge tube containing the dry beads after confirming the sample has been transferred to the 15-ml tube.

Occasionally, beads will clog the pierced hole and prevent complete transfer of the sample. If this occurs, pierce the tube one or two more times and repeat the step in the same 15-ml tube. No additional buffer should be added.

Shear DNA

12. Transfer the sample to a standard 1.5-ml microcentrifuge tube. Microcentrifuge 15 min at maximum speed, 4°C. Discard the supernatant and add 1 ml ice-cold FA lysis buffer to the pellet.

The pellet contains the cross-linked chromatin, cell debris, and unbroken cells. The purpose of this centrifugation step is to remove soluble protein, most of which is not cross-linked to DNA, as it might contribute to nonspecific background in the subsequent immunoprecipitations step. There is no need to resuspend the pellet at this point.

13. Holding the microtip probe near the bottom of the tube to prevent foaming, sonicate the sample 30 sec at 4°C using a continuous pulse at a power output of 20%. Cool in an ice-water bath >1 min. Repeat two more times.

Take great care that the sample does not get too hot.

If a different sonication device is used, empirically determine the conditions necessary to achieve the desired level of DNA shearing. The shear size is determined as described below (see Critical Parameters and Troubleshooting).

14. Microcentrifuge 30 min at maximum speed, 4°C. Transfer the supernatant to a fresh 15-ml disposable conical tube, add 4 ml ice-cold FA lysis buffer, and gently mix by inversion. Remove 250 µl for checking DNA fragment size and freeze the remaining chromatin solution in 800-µl aliquots in liquid nitrogen.

Upon sonication, the cross-linked chromatin is solubilized and purified away from the pelleted material which contains cell debris and unbroken cells. The resulting chromatin solution constitutes the input sample for the subsequent immunoprecipitation. The frozen aliquots are stable for many months when stored at -70°C and are suitable for immunoprecipitations.

Check chromatin-fragment size

15. Add 250 µl ChIP elution buffer and 20 µl of 20 mg/ml Pronase in TBS to the 250-µl chromatin aliquot. Incubate 2 hr at 42°C, followed by 6 hr at 65°C. Phenol extract and ethanol precipitate sample (UNIT 2.1A).

While it is convenient to perform the reaction in a PCR machine overnight, it could just as easily be done in heat blocks or water baths. The same is true of the incubation described in step 26.

16. Resuspend in 30 µl TE buffer, pH 7.5, add 1 µl of 20 mg/ml DNase-free RNase A, and incubate 15 min at 37°C. Add 3 µl of 10× loading buffer and electrophoretically separate the material on a 1.5% agarose gel (UNIT 2.5A).

Fragments should be between 100 to 1000 bp, with an average length of 400 to 500 bp.

It is important to shear DNA fragments down to an average length of 400 to 500 bp. Longer fragments will increase the background and will decrease the resolution of the region to which the protein associates (see Commentary).

Immunoprecipitate

17. Incubate 800 µl chromatin solution with 10 µl primary antibody against the protein or epitope of interest and 20 µl of 50% (v/v) protein A–Sepharose beads in TBS on an end-over-end rotator 90 min at room temperature.

The actual amount of antibody needed has to be empirically determined and can vary considerably. The idea is to have an excess of antibody to efficiently precipitate at least 50% of the antigen in question. One way to assess the efficiency of antigen immunoprecipitation is to determine the amount of antigen present in the sample before and after the immunoprecipitation. An aliquot of 30 µl chromatin solution, taken before and after immunoprecipitation, is usually sufficient to visualize the protein of interest via immunoblotting and standard chemiluminescent detection (UNIT 10.8); however, the samples

have to be boiled in SDS/PAGE sample buffer for 30 min prior to loading in order to reverse the formaldehyde cross-links. The immunoprecipitation conditions can be varied (e.g., time, temperature, salt concentration, presence of detergents) if necessary.

Protein A–Sepharose beads are used here because they work well with most monoclonal and polyclonal sera derived from mouse and rabbit, respectively. In some cases, the use of other beads (e.g., protein G–Sepharose) may improve binding of some antibodies, including rat IgG.

18. Microcentrifuge beads 1 min at 3000 rpm, room temperature. Transfer 300 μ l supernatant into a 0.5-ml PCR tube labeled “INPUT.” Discard the rest of the liquid.

Wash beads

19. Resuspend beads in 700 μ l FA lysis buffer, room temperature, and transfer mixture into a Spin-X centrifuge-tube filter.

The use of Spin-X filters aids in the recovery of the beads after washes and results in better uniformity between different samples. The procedure is also substantially faster with the filters, particularly when multiple samples are processed simultaneously. Alternatively, one could use conventional microcentrifuge tubes for the washes and aspirate the supernatant with a narrow-bore pipet tip after each spin.

20. Place the filter into a 1.5-ml microcentrifuge tube and mix sample 3 min on an end-over end rotator. Microcentrifuge 2 min at 3000 rpm, room temperature. Discard the flow-through liquid at the bottom of the tube.
21. Add 700 μ l FA lysis buffer, room temperature, to the beads and repeat step 20.

Elute protein from beads

22. Wash beads for 3 min each with 700 μ l FA lysis buffer/0.5 M NaCl, 700 μ l ChIP wash buffer, and finally 700 μ l TE.

For many polyclonal antibodies, the more stringent washes in this step result in a cleaner signal, while gentle washes frequently lead to an unacceptably high background. For some antibodies (e.g., monoclonal against peptide epitopes; see Alternate Protocol 1), repeated washes with FA lysis buffer, which are gentler, might be more appropriate.

23. Place filter unit containing the beads into a new 1.5-ml microcentrifuge tube and add 100 μ l of ChIP elution buffer. Gently pipet up and down two or three times in order to dislodge beads from the filter. Incubate 10 min in a 65°C water bath.

A water bath is used instead of other heating apparatuses in order to improve heat transfer.

24. Microcentrifuge beads 2 min at 3000 rpm, room temperature. Discard filter with beads. Transfer the eluate into a 0.5-ml PCR tube labeled “IP.”

Reverse cross-links and purify DNA

25. Add 80 μ l TE and 20 μ l Pronase in TBS to the IP tube. Combine 20 μ l INPUT material (step 18), 100 μ l ChIP elution buffer, 60 μ l TE, and 20 μ l TBS into a new 0.5-ml PCR tube.

26. To reverse cross-links, place tubes into a PCR machine. Incubate 2 hr at 42°C, followed by 6 hr at 65°C. Store samples at 4°C until use.

The incubation at 42°C allows for Pronase digestion of cross-linked polypeptides, while the 65°C incubation results in a reversal of the formaldehyde cross-links.

27. Purify DNA using a Qiagen PCR-purification spin column as per manufacturer’s instructions.

This will require double loading of the spin column (i.e., 600 μ l spin through and then repeat).

Alternatively, add 20 μl of 4 M LiCl and purify by extracting with 25:24:1 phenol/chloroform/isoamyl alcohol, followed by extraction with chloroform and ethanol precipitation (UNIT 2.1). It is useful to add 2 μl of Pellet Paint (Novagen) prior to the addition of ethanol, as this aids both the ethanol precipitation and visualization of the very small pellet.

28. Resuspend in 300 μl TE and store at -20°C .

DNA pellets stored in this fashion should be stable for years.

Perform quantitative PCR

29. Design primer pairs for the desired genomic regions to be examined.

Success in obtaining high-quality data is critically dependent on good primer design (see Critical Parameters and Troubleshooting). In general, primers should be 20 to 30 bases long with a T_m of 55°C to 60°C . The design of good primers is greatly facilitated by commercially available software packages such as Oligo 6.6 (see <http://www.oligo.net>) or Primer Express 1.5 (see <http://www.appliedbiosystems.com>). Most primers require no purification or special treatment prior to PCR. Amplification products should be 75 to 350 bp; longer products should be avoided, as the amplification efficiency is substantially lower. A final primer concentration of 1 μM works well for most primers, but in some instances, improved product specificity may be obtained by lowering the final primer concentration 5 to 10 fold. Refer to UNIT 15.1 for more information.

30. Dilute INPUT DNA (obtained from step 18) in three separate tubes by a factor of 5, 10, and 20. Set up standard PCR reactions (UNITS 15.1 & 15.7) with 2 μl DNA sample, primers at 1 pmol/ μl , and total reaction volumes of 10 to 50 μl . If PCR products will be detected by radioactivity, add 1 μCi of 3000 Ci/mmol [^{32}P]dATP.

For a typical measurement, the three dilutions of input DNA are tested along with duplicate immunoprecipitated samples (or undiluted and 5-fold diluted immunoprecipitated samples). This permits an assessment of whether the assay is being performed in the linear range as well as of the reproducibility of the PCR reaction. The immunoprecipitated DNA is typically used without dilution, although it is useful to analyze different amounts to ensure that it is also in the linear range.

There are several key parameters for achieving an optimum PCR reaction. For example, it is very important to have a quality repeat pipettor that can reproducibly dispense 2 μl DNA. Pipetting inaccuracies at this stage will lead to greater well-to-well variability and poorer reproducibility among identical samples. Additionally, multiple primer pairs (up to 4 to 5) can be included in the same reaction, provided that the PCR products can be unambiguously resolved from each other by gel electrophoresis. This permits simultaneous and internally controlled analysis of multiple genomic regions in a single reaction. However, it is critical to ensure that there is no competition between the different primer pairs and PCR products. Also, comparable results are obtained when PCR reactions are performed in volumes between 10 to 50 μl ; using smaller volumes reduces the cost and facilitates loading of the reaction products on gels. See Critical Parameters and Troubleshooting for a discussion of primer choice.

Detection of PCR products by [^{32}P]label is recommended over detection by ethidium bromide or SYBR Green staining as it improves the sensitivity and extends the linear range of detection; however, it necessitates using the usual precautions in working with radioactivity.

31. Carry out hot-start PCR using the following thermal cycling parameters.

Initial step:	10 min	95°C	(denaturation)
26 cycles:	30 sec	95°C	(denaturation)
	30 sec	55°C	(annealing)
	1 min	72°C	(extension)
Final step:	4 min	72°C	(final extension).

These conditions are generally appropriate for most situations. The annealing temperature may have to be adjusted if the melting temperatures of the primers is substantially above or below 55°C. The number of cycles might also have to be adjusted in some cases if reactions are not in the linear range. See Critical Parameters and Troubleshooting for more details.

Analyze PCR products

32. Add the appropriate loading buffer to the PCR products, and analyze by electrophoresis on nondenaturing polyacrylamide (UNIT 2.7) or agarose gels (UNIT 2.5A).

The gels should be stained either with ethidium bromide or SYBR Green dyes, or analyzed by autoradiography or PhosphorImager.

33. Quantitate the relative amount of PCR products using appropriate software for the accompanying instrument.
34. Calculate the apparent immunoprecipitation efficiency for a specific fragment by dividing the amount of PCR product obtained with the immunoprecipitated DNA by the amount obtained with the input DNA.

A volume of 2 μ l immunoprecipitated DNA sample (1/150 total immunoprecipitated material) contains \sim 200 times the number of cell equivalents as 2 μ l INPUT sample that has been diluted 5-fold (1/30,000 of the original aliquot that was immunoprecipitated). Thus, if the amount of PCR product in the immunoprecipitated sample is equal to the amount of PCR product in the 5-fold diluted INPUT sample, the apparent immunoprecipitation efficiency is 0.5%. The apparent immunoprecipitation efficiency for the background signal is typically \sim 0.025% to 0.05%, and it should not be higher than 0.1%.

BASIC PROTOCOL 2

CHROMATIN IMMUNOPRECIPITATION IN MAMMALIAN CELLS

ChIP experiments in mammalian cells are performed using a modified version of the procedure developed for yeast cells (Basic Protocol 1). In this alternative procedure, immunoprecipitations are performed on material from 5×10^7 cells, which generates immunoprecipitated DNA for approximately 100 quantitative PCR reactions; this is sufficient for triplicate analysis of 33 genomic loci.

Materials

Mammalian cells growing in culture
Culture medium (see APPENDIX 3F)
37% (v/v) formaldehyde
2.5 M glycine
Tris-buffered saline (TBS; APPENDIX 2), ice cold
MC lysis buffer (see recipe), ice cold and room temperature
Liquid nitrogen
MNase buffer (see recipe), room temperature
100 mM (50 \times) phenylmethylsulfonyl fluoride (PMSF) in ethanol (prepare fresh)
20 \times protease inhibitors: 1 mini complete-EDTA free tablet (Roche) in 500 μ l
MNase buffer (prepare fresh)
Micrococcal nuclease (MNase; USB; optional)
0.2 M EGTA
20% (w/v) sodium dodecyl sulfate (SDS)
Protein A/G–Sepharose beads (Amersham Biosciences)
FA lysis buffer (with 150 mM NaCl; see recipe) with 2 mM (1 \times) PMSF and 1 \times protease inhibitors (added from 20 \times stock; see above)
Platform shaker
Refrigerated centrifuge
2-ml screw-cap microcentrifuge tubes

Probe sonicator (e.g., Branson)

15-ml conical tubes

End-over-end rotator

Additional reagents and equipment for mammalian cell culture (*APPENDIX 3F*) and checking DNA fragment size, immunoprecipitation, reversing cross-links, purification of immunoprecipitated DNA, and quantitative PCR (as for ChIP with yeast; see Basic Protocol 1)

Cross-link protein-DNA complexes in vivo

1. For each sample, grow 2×10^8 cells in appropriate medium (*APPENDIX 3F*). For cells grown in suspension, proceed to step 2. For adherent cells, harvest cells by trypsinization and inactivate trypsin by adding serum-containing medium (*APPENDIX 3F*), then resuspend in 50 ml culture medium.

The number of cells can be reduced or increased depending on need. In the procedure described here, the immunoprecipitation (IP) is performed on cross-linked chromatin from 5×10^7 cells, thereby permitting four immunoprecipitations from the culture of cells grown as described above. The minimum number of cells for a standard ChIP experiment depends on the number of immunoprecipitations to be performed and the number of genomic loci to be investigated by quantitative PCR. In general, 10^7 cells are sufficient for a ChIP experiment involving a single immunoprecipitation, but 10^6 cells are not (See Critical Parameters and Troubleshooting).

2. Add formaldehyde (as 37% solution) to cell suspension for a final concentration of 1% (v/v). Cross-link by incubating 8 min at room temperature in a fume hood, occasionally swirling the flask or shaking slowly on a platform.

The volume of cells can be varied over a wide range, but the final formaldehyde concentration should be 1%.

3. Quench the cross-linking reaction by adding 2.5 M glycine to a final concentration of 0.2 to 0.4 M. Incubate for an additional 5 min, then cool on ice.

Lyse cells and isolate chromatin

4. Centrifuge cells 5 min at $450 \times g$, 4°C , and discard supernatant into chemical waste container. Resuspend pellet in 10 to 50 ml ice-cold TBS. Centrifuge cells 5 min at $450 \times g$, 4°C . Repeat this wash one more time and discard the supernatant.

5. Wash cells three times, each time by adding 10 ml MC lysis buffer, centrifuging 5 min at $200 \times g$, room temperature, and decanting supernatant. After final wash, remove residual buffer and snap-freeze in liquid nitrogen.

MC buffer contains NP-40, a detergent that disrupts mammalian cells and generates a nuclear pellet.

6. Thaw frozen nuclear pellet in cool water, then resuspend in 1 ml MNase buffer. Bring volume to 1.5 ml with MNase buffer.

7. *Optional:* Add 10 to 100 U micrococcal nuclease (MNase) and incubate at 37°C for 10 min.

It is important to optimize the amount of MNase for the cell lines being analyzed by performing control reactions in which the amount of MNase is varied. The resulting MNase-digested DNA is analyzed by agarose gel electrophoresis. The desired DNA fragment size should average 50 to 1000 bp, and a nucleosomal ladder should be apparent, with roughly equal band intensities for DNA fragments ranging from nucleosome monomers to octamers.

Note that while MNase treatment is very useful, it is not essential for the procedure. Sonication of cross-linked chromatin from mammalian cells is often insufficient to both solubilize the chromatin (which is necessary for the subsequent immunoprecipitation)

and reduce DNA fragment size to the desired extent. The decision whether to omit MNase treatment and rely on sonication alone depends on a number of factors, including cell type and number of cells, and is difficult to predict. The combination of MNase treatment and sonication is more reliable for producing cross-linked chromatin suitable for subsequent immunoprecipitations and quantitative PCR analysis. However, MNase treatment has the potential to introduce some bias into the results, given that cleavage does not occur in a manner completely independent of DNA sequence.

- 8a. *If optional MNase reaction was performed:* Stop the MNase reaction by adding the following in the order indicated:

30 μ l 0.2 M EGTA (3 mM final)
40 μ M 100 mM (50 \times) PMSF (1 \times final)
100 μ l 20 \times protease inhibitors (1 \times final)
100 μ l 20% SDS (1% w/v final)
80 μ l 5 M NaCl (200 mM final).

Mix well.

- 8b. *If MNase treatment was not performed (i.e., if step 7 was skipped):* add all of the above components except EGTA to the pellet suspension from step 6 and proceed to step 9.

9. Transfer the sample to a 2-ml screw-cap microcentrifuge tube. Sonicate six times, each time for 1 min (power level 4, 60% duty, pulse, 4°C) with a cooling period of at least 2 min in an ice-water bath between pulses. Take great care that the sample does not get too hot.

Hold the microtip probe of the sonicator near the bottom of the microcentrifuge tube to prevent foaming of the sample. The main purpose of sonication is to solubilize the chromatin, not to reduce the size of the DNA beyond that already achieved by MNase treatment. If MNase treatment was not performed, the sonication step should yield some fragmentation of DNA, depending on the cell line.

10. Microcentrifuge 30 min at maximum speed, 4°C, to remove cellular debris. Carefully transfer the supernatant to a new microcentrifuge tube (this is the solubilized chromatin that represents the input sample for the subsequent immunoprecipitation). Remove 100 μ l, and check DNA fragment size as described for yeast (see Basic Protocol 1, steps 15 and 16). Freeze the remaining chromatin solution in 500- μ l aliquots in liquid nitrogen.

Fragments should be between 100 and 2000 base pairs (average 500 to 1000), and a nucleosomal ladder should be apparent with bands representing monomers to octamers being of approximately equal intensity. The frozen chromatin is stable for many months when stored at -70°C and is suitable for immunoprecipitations. Each 500- μ l aliquot represents material from 5×10^7 cells, a number that is appropriate for a typical immunoprecipitation.

Perform immunoprecipitation

11. Thaw chromatin on ice, and microcentrifuge 10 min at maximum speed, 4°C, to remove any precipitate. Transfer supernatant to a new 15-ml conical tube. Dilute chromatin sample 1:5 by adding 2 ml FA lysis buffer containing 1 \times (2 mM) PMSF and 1 \times protease inhibitors.

If smaller volumes are involved, the sample can be placed in a 1.5-ml microcentrifuge tube.

12. Add 50 to 100 μ l protein A/G–Sepharose beads (equilibrated in FA lysis buffer) and incubate 1 to 4 hr at 4°C on an end-over-end rotator. After this incubation, centrifuge 5 min at 300 \times g, room temperature, and transfer the supernatant to a new tube. Take 10% (250 μ l) of this supernatant (precleared chromatin) and label as INPUT.

Adjust SDS concentration (using 20% SDS stock) to 0.5% (w/v), add Pronase (from 20 mg/ml stock) to a final concentration of 1.5 $\mu\text{g}/\mu\text{l}$, then reverse cross-links as in Basic Protocol 1, steps 25 to 28.

This preclearing step removes any material that nonspecifically associates with the protein A/G–Sepharose beads. If performing multiple immunoprecipitations, it is best to preclear in a single batch and then divide the resulting material into separate aliquots for each immunoprecipitation.

13. To the remaining precleared chromatin from step 12, add $\sim 1 \mu\text{g}$ (optimal amount determined empirically) of primary antibody against the protein or epitope of interest and 50 to 100 μl of protein A/G–Sepharose beads (equilibrated in FA lysis buffer). Incubate overnight at 4°C. After incubation, centrifuge 2 min at $200 \times g$, room temperature, discard supernatant, and proceed with washing and elution steps (see Basic Protocol 1, steps 19 to 24).

As the volume of the immunoprecipitation mixture is $\sim 2.5 \text{ ml}$, the initial incubation and centrifugation are typically performed in 15-ml conical tubes, rather than the 1.5-ml microcentrifuge tubes used for the yeast procedure.

The amount of antibody needs to be empirically determined and can vary considerably. In some cases, the conditions for the immunoprecipitation and washing steps need to be varied to maximize the amount of protein-bound DNA that is immunoprecipitated while minimizing the amount of nonspecific background material that remains after washing.

14. Reverse cross-links and purify immunoprecipitated DNA as described for yeast (see Basic Protocol 1, steps 25 to 28).

Typically, 3 μl of the IP DNA is used for each PCR reaction, thereby permitting one to carry out 100 PCR reactions for each IP sample.

Analyze DNA by quantitative PCR

15. Design primer pairs for the desired genomic regions to be examined.

Success in obtaining high-quality data is critically dependent on good primer design, and obtaining suitable primers for quantitative PCR analysis is more difficult for experiments involving mammalian cells because of the increased genomic complexity. It is extremely useful to perform BLAST analysis (UNIT 19.3) on the designed primers to exclude the possibility of similar sequences elsewhere in the genome. In addition, it is critical to perform control experiments involving a range of concentrations of genomic DNA in order to determine if the primers are suitable for quantitative analysis (see Critical Parameters and Troubleshooting). In general, primers should be 20 to 30 bases long with a T_m of 55°C to 60°C. The design of good primers is greatly facilitated by commercially available software packages such as Oligo 6.6 or Primer Express 1.5. Most primers require no purification or special treatment prior to PCR. Amplification products should be 75 to 350 base pairs; longer products should be avoided, as the amplification efficiency is substantially lower. A final primer concentration of 1 μM works well for most primers, but in some instances, improved product specificity may be obtained by lowering the final primer concentration 5- to 10-fold.

16. Dilute INPUT DNA (obtained from step 12 of this protocol) in three separate tubes by factors of 300, 150, and 75. Use the immunoprecipitated DNA (prepared in steps 13 and 14) without dilution, although it is useful to analyze different amounts to ensure that is also in the linear range. Set up standard PCR reactions (UNITS 15.1 & 15.7) with primers at 1 pmol/ μl and 3 μl DNA sample.

Reactions can be performed in volumes of 10 to 50 μl , and PCR products can be detected by gel electrophoresis or on a machine that permits analysis in real time.

The number of genomic regions to be analyzed is limited by the amount of DNA in the immunoprecipitated sample (the amount of input DNA far exceeds the amount of immunoprecipitated DNA, and hence is not limiting for the PCR analysis). The above procedure is based on immunoprecipitations involving 5×10^7 cells, which generate sufficient immunoprecipitated material for 100 PCR reactions. As analysis of each genomic region requires triplicate PCR reactions, 33 genomic regions can be analyzed using the abovementioned number of cells. If material from fewer cells is used for the immunoprecipitation, this reduces the number of PCR reactions and hence the number of genomic regions that can be analyzed. It is possible to increase the number of genomic regions to be analyzed by using less immunoprecipitated material in the PCR reactions. Such an approach is best employed for genomic regions that are enriched by the immunoprecipitation, because the reproducibility of the PCR reaction is related to the number of molecules. In practice, measurements that involve 30 or more cycles of PCR amplification are more variable and should be avoided if possible.

ALTERNATE PROTOCOL 1

SPECIFIC PEPTIDE ELUTION OF PROTEIN-DNA COMPLEXES IMMUNOPRECIPITATED FROM CROSS-LINKED CHROMATIN

Peptide elution represents an alternative method for removal of immunoprecipitated protein-DNA complexes from beads. In this procedure, beads containing the immunoprecipitated complexes are incubated with high concentrations of a peptide recognized by the antibody used in the immunoprecipitation. The added peptide competes with the protein antigen of interest for binding to the antibody, and specifically liberates the protein-DNA complexes from the beads. The high specificity of peptide elution reduces the nonspecific background (typically by a factor of 2 to 4), which makes it the method of choice, particularly for applications where the expected immunoprecipitation signal is low. Peptide elution is especially useful for chromatin immunoprecipitation experiments involving proteins that are tagged with the HA or myc epitopes (in single or multiple copies); however, it would also be appropriate in cases where the antibody used for the immunoprecipitation was generated against a defined peptide sequence. Peptide elution is slightly more expensive than conventional elution, due to the cost of the peptide. In general, peptide elution should be used in conjunction with gentle washes during the immunoprecipitation procedure described below, which minimizes antigen leaching. Stringent washes, such as those employed in the main method (see Basic Protocol 1), will often result in signals that are several-fold lower, with little or no improvement in background. Finally, peptide elution may vary in quality depending on factors such as the number of epitopes in the antigen and the relative stability of the antibody:antigen interaction.

Additional Materials (also see Basic Protocol 1)

1 mg/ml peptide (e.g., myc, HA) in TBS (see APPENDIX 2 for TBS)

For this protocol, follow steps 1 to 21 of the main method (see Basic Protocol 1), replace steps 22 to 25 with the following, and continue with step 26 onwards.

22. Repeat FA lysis buffer wash (see Basic Protocol 1, steps 20 and 21) three additional times for a total of five washes.

Repeated washes with FA lysis buffer are much more gentle than the single washes with FA lysis buffer/0.5 M NaCl, ChIP wash buffer, and TE used in the Basic Protocol 1 and result in higher signal-to-background ratios.

23. Place the Spin-X centrifuge-tube filter unit containing the beads into a new 1.5-ml microcentrifuge tube and add 100 μ l of 1 mg/ml peptide (typically myc or HA) in TBS. Gently pipet up and down two or three times in order to dislodge beads from the filter. Incubate 30 min at 30°C.

24. Microcentrifuge beads 2 min at 3000, room temperature. Discard filter with beads. Transfer the eluate into a 0.5-ml PCR tube suitable for PCR labeled “IP.”
25. Add 150 μ l TE, pH 7.5, 250 μ l of ChIP elution buffer, and 20 μ l of 20 mg/ml Pronase in TBS.

ANALYSIS OF CHROMATIN IMMUNOPRECIPITATION EXPERIMENTS BY REAL-TIME QUANTITATIVE PCR WITH SYBR GREEN

*ALTERNATE
PROTOCOL 2*

Quantitative PCR (QPCR) analysis in real time with SYBR Green has several advantages over the analysis of PCR reactions by gel electrophoresis (see Basic Protocol 1, step 32). First, the method saves considerable time because no gels are involved and because quantitative values are obtained directly from the data curves and do not require densitometry or phosphor imager analysis. As a consequence, this approach permits very rapid analysis of much larger numbers of chromatin immunoprecipitation samples than can be performed with the Basic Protocol 1. Using standard 96-well instruments, it is a straightforward procedure to analyze 100 to 200 samples/day (in replicates of three) with only 1 to 2 hr of total setup time. With newer 384-well instruments and automated robotics equipment, sample throughput can be further increased to thousands per week. Second, the data generated by this procedure are more accurate and reproducible, because quantitative values are determined from continuous sampling throughout the PCR reaction rather than a single end-point determination. Furthermore, the quality and “linear range” of every PCR reaction are directly visualized. Third, the procedure is significantly safer for the researcher, as no radioactive materials or toxic acrylamide are used. The major disadvantage of this procedure is that the measurements are performed individually and hence are not internally controlled, whereas the Basic Protocol 1 permits the simultaneous analysis of multiple genomic regions in a single PCR reaction (provided the individual primer pairs function independently). As such, the Basic Protocol 1 is more useful for analyzing the same small set of genomic regions under multiple experimental conditions and for simultaneous analysis of electrophoretically distinguishable alleles of a given genomic region.

SYBR Green is a sensitive and highly selective double-stranded DNA (dsDNA)–binding dye that remains associated even at the high temperatures normally used for PCR template extension. Real-time PCR reactions involving SYBR Green are performed with standard oligonucleotide primers, and hence are much less expensive than real-time PCR reactions using fluorophore-conjugated oligonucleotides (e.g., TaqMan or Lux probes). Measurements of SYBR Green fluorescence at the polymerase extension step of PCR, when plotted against PCR cycle number, provide both a qualitative assessment of the progress of the PCR and a way to quantitate the relative amount of DNA template initially present in the reaction. Typical real-time QPCR graphs feature the plot of the \log_{10} (Net fluorescence) on the *y* axis versus the PCR cycle number on the *x* axis, and usually contain three well-defined stages: (1) baseline, (2) linear, and (3) plateau. In the baseline stage, the amount of DNA product formed is still below the sensitivity threshold of SYBR Green, so product formation is undetectable. This part of the curve is typically used as a baseline for SYBR Green signal drift. The linear part of the curve is the most important from the analytical standpoint, because it is at this stage that the rate of PCR product accumulation is both constant on a per-cycle basis and readily detectable by increased SYBR Green fluorescence. Finally, as all of the SYBR Green in the reaction becomes bound to the recently synthesized PCR products, the amount of fluorescence stays constant from cycle to cycle and the reaction reaches a plateau.

In the protocol described below, PCR is performed under special conditions that minimize the inhibitory effects of SYBR Green on *Taq* activity and maximize the linear range of product detection. After amplification is complete, raw data are stripped of outliers and

**Chromatin
Assembly and
Analysis**

21.3.13

exported in a format readable by a spreadsheet program such as Microsoft Excel. Finally, data points from replicate samples are averaged, and mean values are further manipulated and ultimately compared to some internal reference or control.

Materials

Input DNA (see Basic Protocol 1, step 28, and Basic Protocol 2, step 14)
Immunoprecipitated fragments (“IP” sample; see Basic Protocol 1, step 23, and Basic Protocol 2, step 13)
TE buffer, pH 7.5 (APPENDIX 2A)
Primers (see Critical Parameters and Troubleshooting)
2× SYBR Green *Taq* mix (see recipe)
Real-time PCR machine and corresponding software (e.g., ABI)
96-well PCR plates (ABI, cat. no. 4306737) and optical adhesive covers
Centrifuge with swinging-bucket rotor and microtiter plate adapter
Spreadsheet program (e.g., Microsoft Excel)

Set up PCR reactions

1. Dilute input DNA to an approximate equivalent of 1×10^6 cells/ml in TE buffer, pH 7.5.

If immunoprecipitations were performed as described in the Basic Protocol 1 or 2, then a 1:25 dilution of the input sample from step 28 of Basic Protocol 1 or the equivalent step of Basic Protocol 2 will result in 1:1000 overall dilution and will correspond to $\sim 5 \times 10^8$ to 1×10^9 cell equivalents.

2. If necessary, resuspend immunoprecipitated fragments in TE buffer, pH 7.5, so that the approximate cell equivalent is 1×10^9 cells/ml.

Immunoprecipitated DNA derived from the IP sample obtained by Basic Protocol 1, step 23 or the equivalent step of Basic Protocol 2 is appropriately diluted and needs no further treatment.

3. Prepare PCR primer stocks by mixing each primer pair at a final concentration of 3.3 μ M in TE buffer, pH 7.5.

It is critical to test newly obtained primer pairs for amplification specificity and performance under conditions that will be used for real-time PCR with SYBR Green (see Critical Parameters and Troubleshooting). SYBR Green can inhibit PCR reactions, and primer pairs that are appropriate for quantitative PCR analysis in the absence of SYBR Green may not work well in the presence of SYBR Green. High-quality primer pairs should result in ~ 1.9 -fold amplification/cycle (this can be determined from quantitative analysis of raw fluorescence data for each cycle, which is generally available on commercial instruments). Amplified material at the completion of the PCR should contain only one band (as assayed on high-percentage agarose or polyacrylamide gels). Specificity information may also be obtained by running dissociation curves on reactions following the conclusion of the PCR run. Typically, samples are melted for 15 min at 95°C, cooled to 60°C, and then slowly heated back up to 95°C over a period of 20 min. Plotting the first derivative of the fluorescence against the temperature allows for simple visual identification of sample heterogeneity. Some instrument-specific software packages have built-in modules for dissociation curve analysis.

4. Select and label the wells to be used in the run.

In general, individual samples should be run in triplicate. Obvious outliers occur with some frequency, generally at <5%. Triplicate analysis of samples permits removal of those outliers while still allowing for inclusion of two accurate measurements for each sample. While this reduces the number of different samples that can be run at any given time, the resulting data is much more reliable and accurate.

For each primer pair examined, the input DNA samples should be run alongside the immunoprecipitated samples. Amplification efficiencies among different primer pairs vary slightly on a per-cycle basis, but those slight variations in efficiency translate into substantially different amounts of amplified material in the cycle range used for analysis. Precise quantitation of relative binding cannot be accurately performed without primer pair-specific input signal.

Detailed instructions on the use of the real-time PCR machine and general issues, e.g., calibration and camera exposure settings, are addressed in the documentation that accompanies the instruments.

5. Program the real-time PCR machine as follows:

1 cycle:	10 min	95°C	(initial denaturation)
40 cycles:	30 sec	95°C	(denaturation)
	30 sec	53°C	(annealing)
	30 sec	72°C	(extension).

Collect the data only at 72°C.

The annealing temperature may have to be adjusted if the melting temperatures of the primers are substantially above or below 53°C. If the desired amplification product is >500 bp (this is not recommended), the extension time at 72°C should be increased to 1 min. See Critical Parameters and Troubleshooting for more details.

6. Using a small-volume automatic pipettor (20- μ l capacity), place a 2- μ l aliquot of each DNA template into the appropriate wells of a 96-well PCR plate. Gently tap the plate to allow the sample droplets to fall to the bottoms of the wells.

It is very important to have a quality repeating pipettor that can reproducibly dispense small volumes of sample into the wells. Pipetting inaccuracies at this stage will lead to greater well-to-well variability and poorer reproducibility among identical samples.

7. Using a small-volume automatic pipettor (20- μ l capacity), place a 3- μ l aliquot of primer mix (see step 3) into the relevant wells and tap the plate a few times to settle the contents.

On many real-time PCR machines, results from 10- μ l reactions are virtually indistinguishable from those of 25- and 50- μ l reactions in their accuracy and reproducibility. The use of 10- μ l reactions provides substantial savings in reagent costs. On some machines, the minimal reaction volume needed for accurate and reproducible results may be greater.

8. With a larger automatic pipettor (100- μ l capacity), add 5 μ l of 2 \times SYBR Green Taq Mix to every assayed well. Place microtiter plate into appropriate microtiter plate adapter and centrifuge 1 min at 200 \times g, room temperature, in a swinging-bucket rotor.

The 2 \times SYBR Green Taq Mix contains a variety of components that considerably reduce the inhibitory effect of SYBR Green, thereby resulting in more reproducible signals that require fewer amplification cycles. Comparable mixes containing proprietary buffers can be obtained commercially. It is critical that quantitative PCR reactions containing SYBR Green be performed under conditions of efficient amplification (e.g., 1.9-fold amplification/cycle).

9. Seal plate with clear optical adhesive covers, overlay foam compression pad with gold side facing up, and place into the real-time PCR machine. Secure lid.

The details of this step may differ, depending on the machine.

10. Start the PCR protocol (see step 5). After completion, save the run for future analysis.

DATA analysis

11. Open the file containing the real-time data according to the manufacturer's instructions for the instrument.

Although the specific protocol will depend on the software and instrument, the overall logic and approach to the analysis of real-time data is generally applicable.

12. Look at the different curves and set the baseline as needed.

Generally, the baseline should be set from cycle 3 to the cycle just prior to where the curves start increasing in a linear fashion. It is desirable to have at least 10 cycles for the calculation of the baseline, as this results in increased accuracy in the subsequent calculations of the threshold cycle.

13. Change the value in the Threshold box to be about halfway up in the linear range, and apply changes to the data set.

The threshold cycle is defined to be the PCR cycle at which the fluorescence is 10 times (10 is the default multiplier) the standard deviation obtained in the baseline calculation. When the multiplier is set to 10, the fluorescence at the threshold cycle is considered the lowest fluorescence value that is significantly above the background. In practice, this number frequently lies in the nonlinear range of many of the curves. For later calculations, it is easier to manually set the fluorescence value used to calculate the threshold cycles to 0.04. At this value, all the curves should be in the linear range and well above the baseline, allowing for far more accurate comparisons of the threshold cycles. On occasion, however, it will be necessary to adjust this value either up or down to better reflect the linear range of net fluorescence for most of the curves.

14. Manually select one group of triplicates and visually inspect their amplification plots. If curves are essentially superimposable and the threshold cycle (C_T) values are close to each other (maximal and minimal replicates within 1 cycle, preferably within 0.5 cycles), proceed to the next triplicate sample. Otherwise, remove the outlier and continue to the next triplicate.

Decisions regarding the removal of some outliers could either be straightforward or judgment calls, depending on the circumstances. In cases where two out of three curves are superimposable while the third is clearly off by more than a cycle, it is a fairly easy decision to consider the third replicate an outlier. If the curves are closer, the decision on which one to eliminate, if any, becomes much more difficult. As a general guide only, if the spread between the lowest and highest C_T values is less than 0.5, it is probably safe to average all the C_T values (see step 12). If the C_T range is <1 but >0.5 , the data are less reliable and the decision to remove any data points should probably be made on a case-by-case basis. It is highly recommended that the PCR be repeated for samples where the C_T ranges are >1 with no two curves superimposable.

15. Proceed to analyze the data for all triplicates in the manner described above. Save the results in a different file.

16. Export the data to a spreadsheet program such as Microsoft Excel by using built-in filters. The file should not contain omitted wells (see step 14) and should be in a column format containing well positions, descriptors, and C_T values for each selected well.

Final calculations are most easily handled in a spreadsheet, but could also be performed with a scientific calculator.

17. Open the exported file. Proceed to average triplicate measurements for each sample in a new column ($^{\text{AVERAGE}}C_T$).

For some samples, there may be only two measurements left as a result of the removal of outliers in step 14.

18. For each primer pair, subtract the $\text{AVERAGE } C_T \text{ (INPUT)}$ from $\text{AVERAGE } C_T \text{ (IP)}$ in a new column. This number is the $\text{NET } C_T$.

This value represents the difference in cycles between the immunoprecipitated sample and the input DNA.

19. Subtract the $\text{NET } C_T$ of a primer pair that serves as a reference or a control (CTRL) from the $\text{NET } C_T$ of the experimental (EXPT) primer pair in a new column. The resulting value is $\text{NET } C_T^{\text{EXPT}} - \text{CTRL}$. Repeat $\text{NET } C_T$ subtraction of control primer for all other experimental primers.

It is very desirable to have a control primer pair that can be used to assess the relative cross-linking efficiency at promoters of interest. Frequently, the control primer pair could be specific for a DNA region that does not bind to the immunoprecipitated protein of interest. The signal from the control primer pair could then be considered the background, and the binding efficiency of the protein to different promoter regions could be expressed as fold over background binding.

20. Evaluate the expression POWER ([mean primer slope], $-\text{NET } C_T^{\text{EXPT}} - \text{CTRL}$), where the [mean primer slope] is the base and $-\text{NET } C_T^{\text{EXPT}} - \text{CTRL}$ is the exponent. Repeat the process with other primers by using the different $\text{NET } C_T^{\text{EXPT}} - \text{CTRL}$ values calculated in step 19.

The actual value calculated in the POWER expression above is the degree of occupancy of the immunoprecipitated protein at the sequence of interest relative to that of a control (or background) DNA region.

*Perfectly efficient PCR, in which the number of amplified molecules doubles every cycle, has a primer slope of 2. As defined, this value is independent of primer pair sequence, target sequence length, and other variables that under normal circumstances may adversely affect the efficiency of amplification. In practice, however, the mean primer slope is almost always <2 and is slightly variable from primer pair to primer pair, mostly due to differences in primer sequence and other parameters (e.g. GC content and length of amplified sequence). For the majority of primers designed to amplify *S. cerevisiae* promoter sequences, the mean slope is 1.9 ± 0.06 , and this value can be safely used in the calculations above. However, it is still good practice to calculate representative slopes for every newly synthesized primer pair; any substantial deviation from $1.9 (\pm 0.06)$, especially to the downside, should be viewed suspiciously. Frequently, a slope that substantially deviates from the normal range is indicative of problems in the amplification.*

The mean primer slope for a given primer pair is most easily calculated by performing linear regression on the linear portion of the amplification plot (log net fluorescence versus PCR cycle). In order to perform this calculation, it is first necessary (if possible) to export a file containing the fluorescence values for all the wells in use at every PCR cycle (see software manuals for more information). It is then possible to use linear regression to estimate the slope within the linear range (i.e., by using the MS Excel function LINEST). Since it is rather time-consuming to calculate slopes for an entire 96-well plate one-by-one, it may well be worthwhile to write a macro (or a stand-alone program) to automate this process.

MAPPING PROTEIN BINDING SITES IN VIVO USING CHROMATIN IMMUNOPRECIPITATION FOLLOWED BY QUANTITATIVE PCR SCANNING

The chromatin immunoprecipitation procedures described in this unit are typically used to measure the association of a given protein with a general region, e.g., the promoter of a gene of interest. In fact, the technique may also be applied to relatively fine-scale mapping of binding sites by using a series of PCR primers to scan a DNA region. This method has been used to define a narrow zone of peak DNA occupancy by RNA polymerase III transcription factors at a newly described locus (Moqtaderi and Struhl, 2004). Scanning

ALTERNATE PROTOCOL 3

Chromatin Assembly and Analysis

21.3.17

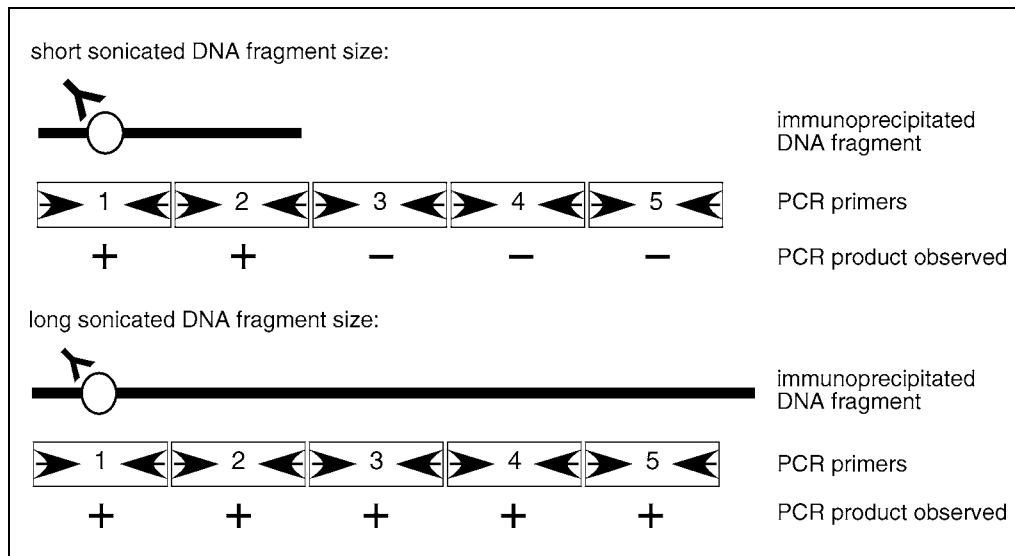


Figure 21.3.2 Effect of immunoprecipitated DNA fragment size on the distance from a binding site from which PCR signal will be observed. As DNA fragment size increases, PCR signal is observed upon amplification with primer pairs further from the binding site.

with closely spaced or even nested PCR products yields peak signal intensity over a very small area (<100 bp) and is especially valuable in distinguishing relative binding position of two proteins with fairly closely spaced recognition sequences. The primary determinants of success in a scanning experiment are the size and spacing of the quantitative PCR products, which define the optimum region of association, and the average DNA fragment size, which determines the apparent spread of the observed association peak.

First, consider the effects of DNA fragment size. Any unbound region of DNA that is contained on the same fragment as a true binding site will be co-immunoprecipitated along with the true target DNA; thus, the longer the average sonicated DNA fragment size, the higher the quantitative PCR signal at points far upstream or downstream of a true target. Figure 21.3.2 illustrates this for immunoprecipitation of a protein bound to its target sequence on either a short or a long DNA fragment. If the average size of the sonicated DNA fragments is small, then the true binding site will only rarely coexist on the same DNA molecule with unbound regions of DNA far away. These distal unbound regions will therefore almost never be co-immunoprecipitated with protein bound to the true target sequence, and they will consequently yield poor signals in the subsequent quantitative PCR. In other words, as the sonicated fragment size decreases, positive signal in a scanning experiment is focused over a narrower region of DNA. Figure 21.3.3 (A and B) shows the relative breadth of signal peaks that would be obtained using the same set of primer pairs on two different hypothetical populations of immunoprecipitated DNA differing only in average fragment size. For simplicity, an idealized population of sonicated fragments of uniform size and unbiased distribution is assumed. Consideration of how far the signal deriving from a point source might be expected to spread has been used to determine the range of histone deacetylation *in vivo* by the Sin3-Rpd3 complex (Kadosh and Struhl, 1998).

Notice that the peak quantitative PCR signal occurs in the same position (primer pair 3 in Fig. 21.3.3) regardless of sonicated fragment size. Although the larger fragment size causes signal to be observed over a wider area, it also means that there are more fragments that will both cover the binding site and be long enough to contain the complete

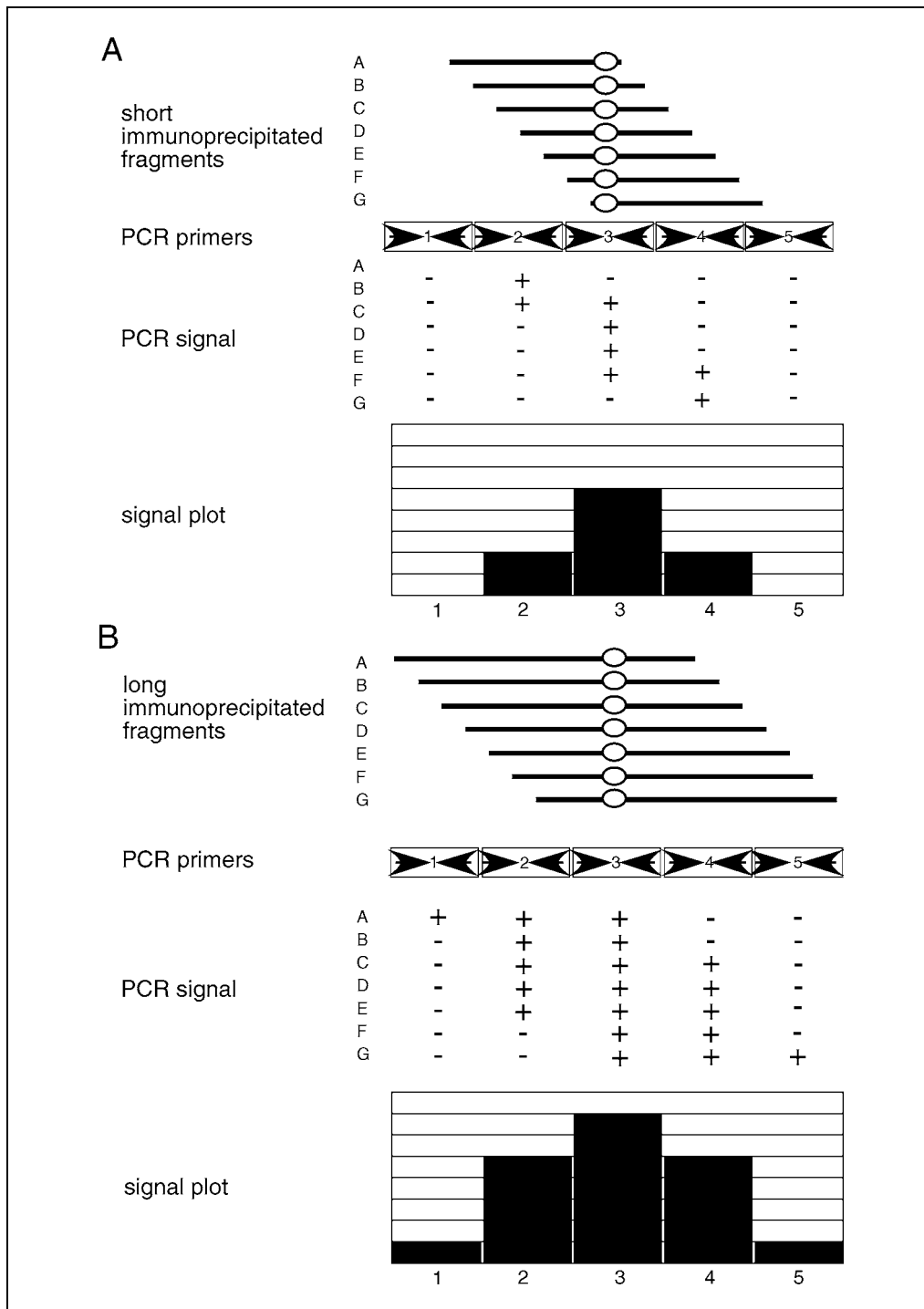


Figure 21.3.3 Distribution of PCR signal observed after ChIP in two idealized DNA populations of different sonicated fragment size. **(A)** When fragment size is short, the distribution of signal around a binding site is fairly narrow. **(B)** With a population of larger DNA fragments, the distribution of PCR signal is broader, though it will be more intense at its peak, since more immunoprecipitated fragments can be amplified by the central primer pair. Peak signal occurs in the same position regardless of fragment size.

target sequence amplified by the peak central primer pair. If, for example, all sonicated fragments are 1000 bp long, then there are 1000 possible unique DNA fragments that cover a hypothetical single-point binding site (assuming again that sonication produces an unbiased distribution of all possible 1-kb fragments). In PCR with one primer 50 bp upstream and one 50 bp downstream of this binding site, amplification can occur only

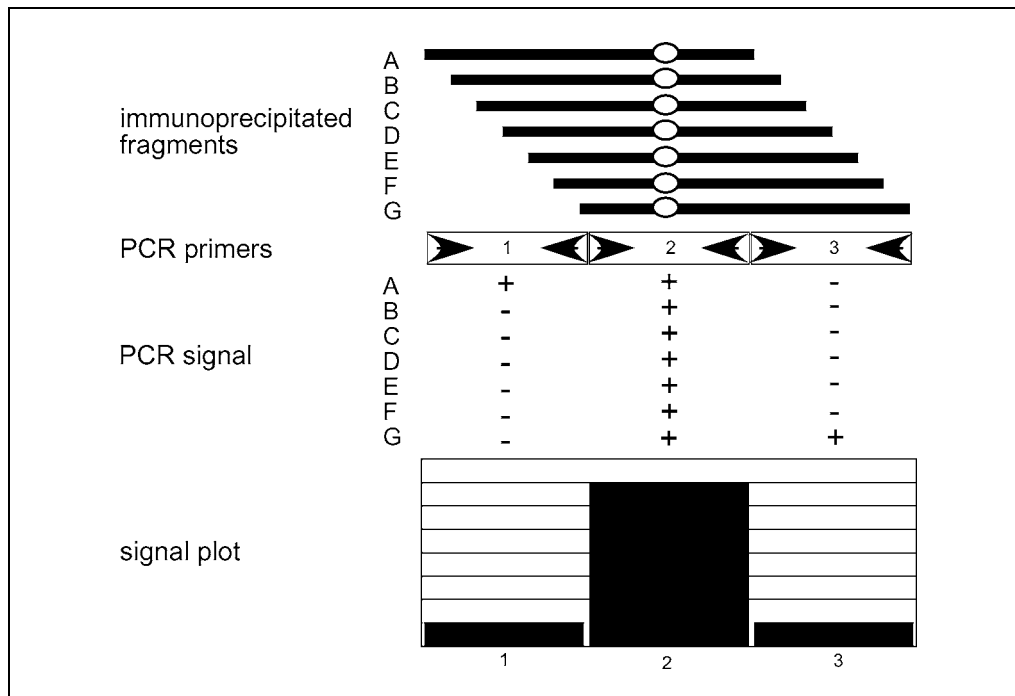


Figure 21.3.4 With fewer and larger PCR products, information about the position of the binding site within the peak PCR product is limited. Here, all that can be said is that the binding site is contained somewhere within the section of DNA amplified by the central PCR primer pair.

if the sonicated fragment contains a binding site and if this binding site is at least 50 bp from either end; these numbers are obviously approximate, since a PCR primer does not anneal to a single nucleotide. Thus, there are 900 possible fragments in an idealized pool of 1-kb fragments that can be amplified by the central PCR primer pair. If, instead of 1 kb, the fragment size were 500 bp, the number of fragments potentially detectable by the same primer pair would be only 400, yielding a lower peak, but a more focused one. The use of DNA with a smaller fragment size therefore makes it easier to exclude neighboring unbound DNA when attempting to delimit a true region of association.

Secondly, and more critically, the lower limit to the width of the signal peak is determined by the length and spacing of the PCR products. Imagine a hypothetical point-source binding site in the exact center of a 500-bp PCR product. In a series of nonoverlapping 500-bp PCR products, this product will yield the peak signal. However, to a first approximation, the most that can be said is that the binding site lies somewhere within this 500-bp region. If, instead of 500 bp, the length of the PCR products is 250 bp, then the possible position of the binding site is narrowed to a 250-bp region; the neighboring nonoverlapping 250-bp PCR products will yield a lower signal, since in some fraction of the cases the neighboring regions of DNA that they cover will not happen to be included on the same sonicated fragment as the binding site. Figure 21.3.4 shows the same hypothetical sample as in Figure 21.3.3B as it would look with fewer and longer PCR products.

When primer pairs are spaced closely enough, the symmetry of the signal contributed by PCR products adjacent to the peak will yield additional information about the position of the binding site. For example, if a series of primer pairs is designed such that the resultant PCR products overlap, it is possible to narrow the optimum binding region to an area smaller than the length of any one PCR product. This is illustrated in

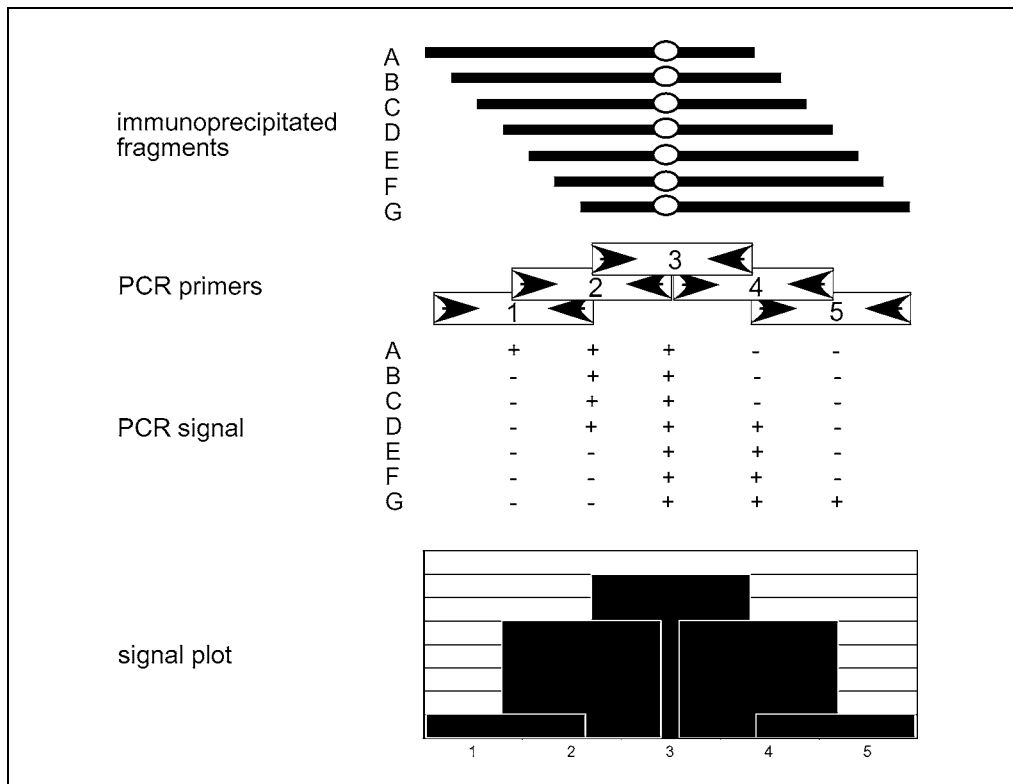


Figure 21.3.5 When the midpoint-to-midpoint distances of the PCR products are more closely spaced, the density of information and the symmetry of the signal plot make it possible to infer that the binding site is in the middle, rather than close to either end, of the DNA amplified by the central primer pair.

Figure 21.3.5 for a binding site at the exact center of PCR product 3. If the binding site is not perfectly centered over any one PCR product, the shape of the distribution will tend to the left or the right, as shown in Figure 21.3.6, panels A and B. The crucial parameter is the density of scanning information, i.e., the relative midpoint-to-midpoint spacing of the PCR products. High scanning density may be achieved with larger, overlapping PCR products or smaller, nonoverlapping PCR products.

When designing oligos for PCR scanning across a genomic region, several parameters should be considered. The size of the PCR products should be kept relatively small; 100 bp is generally a convenient compromise between achieving specificity and maintaining a reasonable limit on the number of nonoverlapping primer pairs needed to scan the whole region. The PCR primers should all produce PCR products of similar size, and the primers should all have approximately the same optimal annealing temperature, so that PCR reactions performed in parallel are equally efficient. This is most easily accomplished by using a primer-design program such as Oligo 6.6, which allows many potential primer sequences to be analyzed relatively quickly. Inevitably, the specific properties of the sequence (repeated sequence, regions of low melting temperature, the potential for primer-dimer formation) constrain the primer design such that perfectly even spacing, similar melting point, and uniform PCR product size are difficult to achieve across an entire series of primer pairs. In such situations, one should strive for well-spaced PCR product midpoints as a first consideration. This allows peak symmetry to be used most effectively in defining a binding site. In addition, for experiments meant to separate two binding sites, if there is any prediction of where the binding sites might be, it is useful to center one primer pair over each, while minimizing overlap with the neighboring binding site.

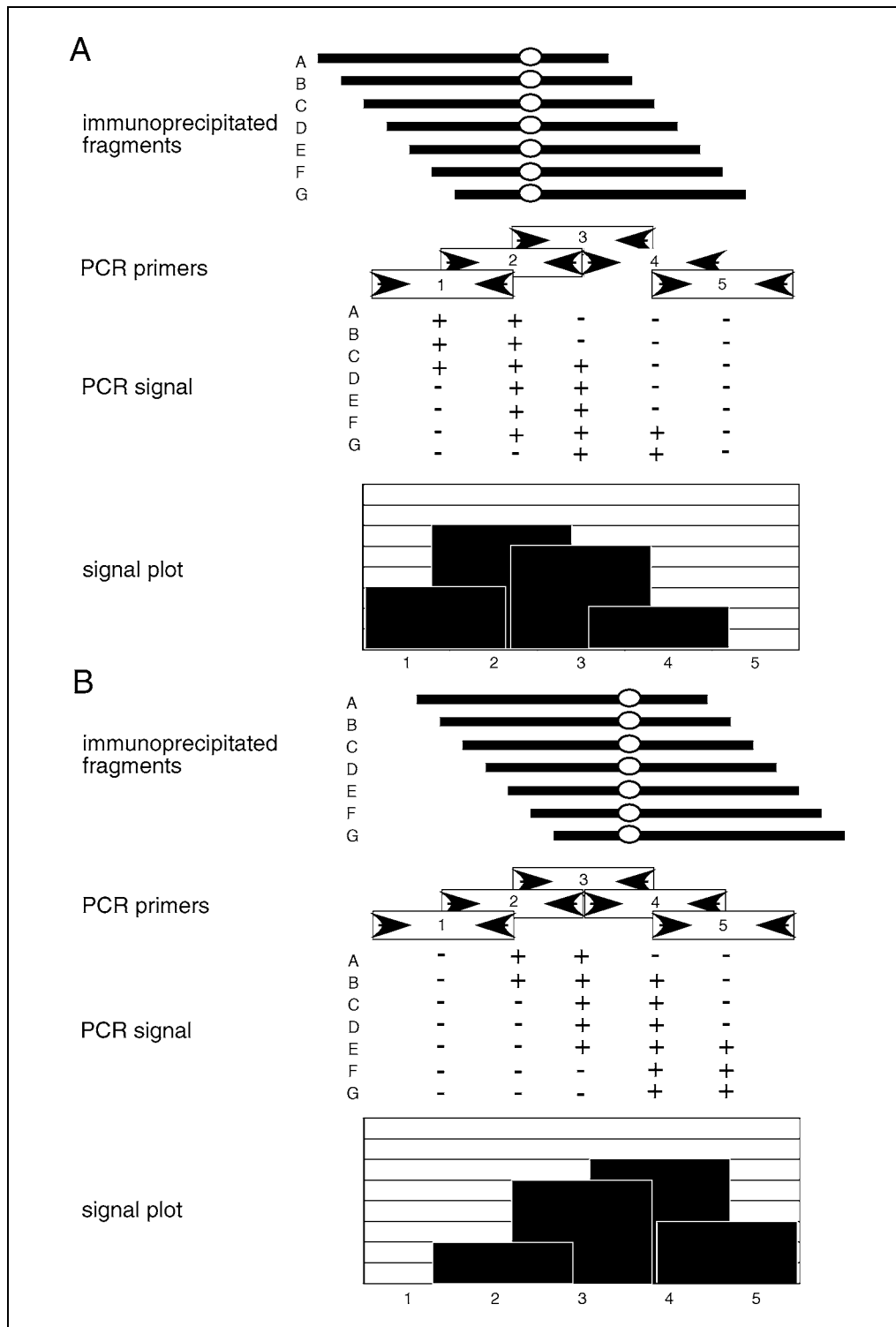


Figure 21.3.6 When the binding site is much closer to one or the other end of the fragment amplified by the central primer pair, the signal distribution will be skewed toward the neighboring PCR product. In **(A)**, peak binding occurs at the very end of the region amplified by primer pair 3 and slightly off-center with respect to that amplified by primer pair 2. In **(B)** peak binding occurs at the opposite end of the region amplified by primer pair 3 and is slightly off-center with respect to the region amplified by primer pair 4.

REAGENTS AND SOLUTIONS

Use double-distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

ChIP elution buffer

50 mM Tris·Cl, pH 7.5 (APPENDIX 2)
10 mM EDTA
1% SDS
Store up to 1 year at room temperature

ChIP wash buffer

10 mM Tris·Cl pH 8.0 (APPENDIX 2)
0.25 M LiCl
1 mM EDTA
0.5% Nonidet P-40
0.5% sodium deoxycholate
Store up to 1 year at room temperature

FA lysis buffer with and without 2 mM PMSF or 0.5 M NaCl

For FA lysis buffer:
50 mM HEPES: adjust pH to 7.5 with KOH
150 mM NaCl (or 0.5 M; see below)
1 mM EDTA
1% Triton X-100
0.1% sodium deoxycholate
0.1% SDS
Store up to 1 year at room temperature

For FA lysis buffer/2 mM PMSF add 100 mM phenylmethylsulfonyl fluoride (PMSF) in ethanol (store up to 1 year at -80°C) to a final concentration of 2 mM just before use. For FA lysis buffer/0.5 M NaCl, change the amount of NaCl added to 0.5 M. Store up to 1 year at room temperature.

MC lysis buffer

10 mM Tris·Cl pH 7.5 (APPENDIX 2A)
10 mM NaCl
3 mM MgCl_2
0.5% (v/v) NP-40
Store up to 1 year at 4°C

MNase reaction buffer

10 mM Tris·Cl pH 7.5 (APPENDIX 2A)
10 mM NaCl
3 mM MgCl_2
1 mM CaCl_2
4% (v/v) NP-40
Store buffer with above components up to 1 year at room temperature
Immediately before use, add:
1 mM PMSF (from 100 mM stock in ethanol).

SYBR Green Taq mix, 2×

12 mM Tris·Cl, pH 8.3 (*APPENDIX 2*)

50 mM KCl

5 mM MgCl₂

150 mM trehalose (Sigma)

100 mM betaine (Aldrich)

0.2% (v/v) Surfact-Amps 20 (Pierce; active ingredient, Tween 20)

0.2 mg/ml nonacetylated BSA (Sigma, B8667)

1 μM 5(6)-carboxy-X-rhodamine (ROX; Helix Research)

0.133× SYBR Green (Molecular Probes; final dilution 1:75,000)

Store solution with above components indefinitely at -80°C or up to several months at 4°C in the dark

Add the following immediately prior to use

0.5 mM each dATP, dTTP, dCTP, dGTP

0.2 U/μl hot-start *Taq* DNA polymerase (amount may need to be slightly adjusted to account for batch/activity variations among different manufacturers)

COMMENTARY

Background Information

Direct protein-DNA contacts and indirect protein-DNA interactions regulate fundamental chromosomal functions such as DNA replication, gene expression, and chromosome segregation. Thus, knowledge about the distribution of particular proteins on specific chromosomal DNA sequences can provide important insights into the mechanisms that govern chromosomal functions, structure, and organization. *In vivo* footprinting methods provide high-resolution mapping of protein-DNA interactions but cannot directly identify the chromatin-associated protein(s) responsible for the footprint. Chromatin immunoprecipitation, by contrast, is ideally suited for determining the identity of proteins associated with specific DNA sequences *in vivo*, albeit with lower resolution (≤ 1 kbp).

Two groups, Gilmour and Lis, and Solomon and Varshavsky, independently pioneered cross-linking and immunoprecipitation methods for *in vivo* chromatin analysis (Gilmour and Lis, 1984; Solomon and Varshavsky, 1985; Solomon et al., 1988; Gilmour et al., 1991). These methods exploited cross-linking to conserve *in vivo* chromatin structures and permit their isolation under the stringent conditions necessary to isolate soluble chromatin. Cross-linked protein-DNA complexes were purified by cesium chloride centrifugation (a time-consuming step) and subjected to immunoprecipitation. Their methods were distinguished primarily by the cross-linking agent: Gilmour and Lis employed UV irradiation, while Solomon and Varshavsky used formaldehyde. The biochemical characteristics of each method have been discussed ex-

tensively (Gilmour and Lis, 1984; Solomon and Varshavsky, 1985; Orlando et al., 1997, and references therein). In short, UV irradiation cross-links only protein-DNA complexes in direct contact, which limits its use. Formaldehyde reacts with primary amines on amino acids and DNA and RNA bases, reversibly forming a covalent adduct between two primary amines in close proximity to each other (≤ 2 Å). Because protein-protein adducts are formed in addition to protein-DNA adducts, chromatin-associated proteins not directly bound to DNA can be cross-linked to DNA via other proteins such as histones, significantly broadening the applicability of this procedure. Cross-linking with formaldehyde is also more easily accomplished than UV irradiation, especially with larger culture volumes.

Formaldehyde-based chromatin immunoprecipitation was simplified and adapted for use in other experimental systems, including budding yeast, where it was first used to assess the association of differentially acetylated histones with the silent mating-type loci (Dedon et al., 1991; Braunstein et al., 1993). This method involved fractionation of cell extracts to isolate a chromatin fraction before immunoprecipitation. A closely related method was used to assess the composition of the budding yeast centromere (Meluh and Koshland, 1997). The protocol presented here (see Basic Protocol 1; Fig. 21.3.1) is for a simpler procedure derived by Hecht and Grunstein in which immunoprecipitations were performed with whole-cell extracts to assess the spatial distribution of SIR proteins on telomere-proximal DNA regions (Hecht et al., 1996;

Strahl-Bolsinger et al., 1997). Basic Protocol 1 has also been applied to characterize the spatial and temporal associations of DNA replication proteins with chromatin at replication origins (Aparicio et al., 1997; Tanaka et al., 1997), the association of general transcription factors at promoters (Kuras and Struhl, 1999; Li et al., 1999), and the dynamics of DNA-binding proteins and chromatin-modifying activities associated with a cell-cycle and developmentally regulated promoter (Cosma et al., 1999). In each of these latter studies, protein association with relevant DNA sequences was examined using PCR amplification.

ChIP experiments in mammalian cells are performed using a modified version of Basic Protocol 1. The protocol for mammalian cells differs in five significant respects from the yeast protocol. First, intact mammalian cells are easily lysed by detergents, whereas breakage of yeast cells requires harsh mechanical treatment in the presence of glass beads to disrupt the cell wall. Second, formaldehyde-cross-linked chromatin in mammalian cells is generally more insoluble than comparable material from yeast cells. As a consequence, sonication is often insufficient to generate soluble material that is suitable for immunoprecipitation and that contains DNA fragments averaging the desired 500 to 1000 base pairs in length. Thus, in the mammalian cell protocol, cross-linked chromatin suitable for ChIP analysis is generated by a combination of micrococcal nuclease (MNase) treatment and sonication. Third, the immunoprecipitation procedure in the mammalian protocol includes a preclearing step to remove background material that sticks nonspecifically to the beads (this preclearing step is helpful, but not essential, and it can also be used in the yeast basic protocol). Fourth, ChIP experiments in mammalian cells are generally more limited with respect to the number of immunoprecipitations to be performed and/or the number of genomic loci to be investigated by quantitative PCR. For cultures of comparable volumes, there are ~10- to 100-fold fewer mammalian cells than yeast cells, and obtaining large numbers of mammalian cells is often limited by the expense of the growth medium and/or the inability to grow cell lines in suspension. The number of cells, and hence number of DNA molecules for individual genomic regions, is an important consideration in ChIP experiments because the sensitivity and reliability of the quantitative PCR assay is directly related to the number of molecules for a given genomic region. Finally, as mammalian genomes are 100-fold more

complex than yeast genomes, it is more difficult to generate PCR-primer pairs specific to the genomic region of interest and hence suitable for quantitative PCR analysis. It is essential to perform BLAST analyses (*UNIT 19.3*) to minimize possible cross-reactivity of primers with undesired regions of the genome, and to perform control reactions on genomic DNA to evaluate the suitability of PCR primer pairs for quantitative analysis.

ChIP can also be used to specifically follow the genomic association of mutant proteins that are unable to support cell growth (Mencia and Struhl, 2001). This involves a strain containing both an epitope-tagged version of the mutant protein and an untagged version of the wild-type protein, which supports cell growth; ChIP is performed with an antibody against the epitope. Lastly, ChIP can be combined with microarray technology to identify the location of specific proteins on a genome-wide basis (Ren et al., 2000; Iyer et al., 2001). The immunoprecipitated DNA is PCR amplified with linkers, and the resulting material hybridized to microarrays containing the complete set of intergenic regions in *Saccharomyces cerevisiae*. Such “genome-wide location” or “ChIP-chip” is particularly powerful, because it allows one to identify novel regions of protein association, without any previous knowledge.

Critical Parameters and Troubleshooting

Controls

There are two basic types of controls for a standard ChIP experiment. One control is mock immunoprecipitation to determine the specificity of an observed signal. This is accomplished by performing parallel immunoprecipitations of a given cross-linked chromatin sample with the antibody of interest and with an irrelevant antibody (or simply not providing any antibody). Alternatively, when antibodies against epitope-tagged proteins are employed (e.g., anti-HA, anti-myc), a similar comparison can be made with parallel immunoprecipitations (primary antibody included) of chromatin samples from strains expressing epitope-tagged or untagged versions of the protein of interest. However, in such experiments, it is often observed that the apparent immunoprecipitation efficiency for any irrelevant genomic region is about 2- to 3-fold higher in the experimental sample than in the control sample. This probably reflects nonspecific, and perhaps artifactual, association of the protein of interest with chromatin; hence it is not

indicative of specific protein association with the genomic region. For proteins that generally and strongly associate with the entire genome (e.g., histones), the apparent immunoprecipitation efficiency in the experimental sample will be much higher than the control sample.

The second control, which is usually more meaningful, is to examine a given pair of input and immunoprecipitated samples for multiple genomic regions. Control genomic regions (i.e., those not interacting with the protein) should all give the same background level of apparent immunoprecipitation efficiency (typically ~0.025% to 0.05%). Fragments bound by the protein of interest will give higher apparent immunoprecipitation efficiencies, and the difference can be expressed as fold enrichment over the background level. In the best cases, enrichments can be >100-fold, but even a factor of two can be meaningful if the experiment is repeated enough times and the experimentally determined error is sufficiently low. The advantage of this approach is that the identical samples are used to directly determine relative protein association to different genomic regions. Furthermore, differences in fold enrichments for different genomic regions represent relative quantitative measurements of protein association *in vivo*.

Additional controls may be used depending on the specific application. For example, where binding to a putative binding site is being tested, a mutation in the binding site is a critical control (especially if such a mutation previously was shown to eliminate binding *in vitro*). In a related manner, it might be useful to determine the protein association in mutant strains or under particular environmental conditions that are suspected to be important for the protein of interest.

Cross-linking

The extent of formaldehyde cross-linking is an important variable that in principle may be modified by changing the duration of cross-linking, the concentration of formaldehyde, or the temperature at which the cross-linking is performed. The use of 1% (final concentration) formaldehyde for 15 min at temperatures ranging from 12°C to 37°C usually works well; however, at temperatures above 30°C, background sometimes increases. Therefore, when fixation at a higher temperature is required, reducing the duration of cross-linking or the formaldehyde concentration may be helpful. Excessive cross-linking can interfere with cell breakage by bead beating and effective

fragmentation and solubilization of the DNA by sonication (see below). For some applications where protein cross-linking is particularly efficient (e.g., histones), it might be useful to decrease the cross-linking time or formaldehyde concentration. In particular, histone tails have a number of lysine residues that are likely to be modified by formaldehyde, and such modified lysines may interfere with the binding of antibodies against specific peptides corresponding to modified histones (e.g., by acetylation, phosphorylation, methylation).

Cell lysis

Although complete lysis of all cells is not absolutely necessary (and may be difficult to achieve), it is very important that lysis be as efficient as possible. Efficient lysis is important to obtain a reproducible degree of cell breakage among a group of samples to reliably compare results. Significant differences in cell lysis efficiency will result in immunoprecipitations with different ratios of antibody to chromatin, which will possibly alter immunoprecipitation efficiency. Cell breakage by a mini bead beater is generally more efficient than breakage by a multi-vortexing apparatus, although both methods work. In both cases, it is important to use flat-bottomed 2-ml microcentrifuge tubes. When using the mini bead beater, the sample and beads should nearly fill the tube, whereas for vortexing it is important to maintain a small volume. The extent of cell breakage may be monitored microscopically by comparing the number of intact cells (determined by counting on a hemacytometer) in small, diluted aliquots of the sample taken before and after vortexing. In addition, the size of the remaining pellet (unbroken cells and debris) obtained in the first centrifugation following sonication (see Basic Protocol 1, step 14) is a good general indicator of the extent of lysis. The size of this pellet should be routinely checked (by rapid visual inspection) to compare the extent of lysis among samples.

The final yield of genomic DNA in the extract is also an important indicator of the extent of cell breakage, although the DNA yield is also dependent on the solubilization of chromatin by sonication (see below). Poor or variable cell breakage may result from excessive cross-linking that toughens the yeast cell wall and other structures. The procedure for lysis of *Saccharomyces cerevisiae* is appropriate for other yeast species. However, modified procedures are necessary for breaking mammalian cells.

Sonication

Shearing DNA to a small size (~500 bp average) by sonication is the critical factor in achieving resolution between a DNA sequence where a particular protein is bound and a nearby (*cis*-)DNA sequence that does not bind that protein. In addition, fragmentation of the chromatin is essential for its solubilization from the ruptured cells. As indicated above, the ability to fragment and solubilize the chromatin depends on the extent of chromatin cross-linking. In general, more cross-linking results in larger fragment size and lower solubility, resulting in lower yield (Orlando et al., 1997). Because of the importance of this variable, the shear size of the DNA should be assessed to confirm that the desired degree of fragmentation has been achieved, and should be reassessed if fixation conditions are altered. The shear size is determined by electrophoresing DNA from step 16 of the Basic Protocol 1 on a 1.5% to 2.0% agarose gel and visualizing with ethidium bromide (UNIT 2.5A). A smear of DNA should be apparent with an average size of 500 bp and most of the DNA (>90%) should be in the size range of 100 to 1000 bp.

As an alternative to sonication, DNA fragment size can be reduced by treatment of the cross-linked chromatin with micrococcal nuclease (UNIT 3.12). Micrococcal nuclease preferentially cleaves DNA located in the linker regions between nucleosomes. By varying the concentration of micrococcal nuclease, it is possible to generate samples in which average DNA size varies. The minimal useful size is about 150 bp, which corresponds to a mononucleosome. However, cleavage to mononucleosome-sized fragments may also result in a preferential loss of certain genomic regions due to the sequence-specificity of micrococcal nuclease.

Immunoprecipitation

The success of this procedure relies on the use of an antibody that will specifically and tightly bind its target protein in the buffer and wash conditions used. In addition, antibody should be present in excess with respect to its target protein so that differences in the amounts of the protein-DNA complexes of interest will be accurately measured. Perform preliminary experiments to confirm avid immunoprecipitation and determine an approximate amount of antibody to use. Chromatin extracts should be prepared without prior cross-linking of the cells and subjected to immuno-

precipitation with varying concentrations of antibody (20 µg/ml antibody may be a good starting point). The amount of the protein of interest in the extracts before and after immunoprecipitation should be analyzed by immunoblotting (UNIT 10.8) to determine the lowest antibody concentration that depletes >90% of the protein of interest from the extract. This antibody concentration is a good starting point for the full protocol and may later be modified to maximize the signal-to-noise ratio (see Anticipated Results). With cross-linked chromatin, immunodepletion of the target protein is less efficient (~50%), presumably due to masking or modification of the epitopes, and a significant amount of the protein remains refractory to immunoprecipitation even with higher antibody concentrations. Thus, the ideal antibody concentration is ultimately determined empirically to maximize the yield of specific coprecipitated DNA while minimizing precipitation of nonspecific DNA.

Both monoclonal and polyclonal antibodies have been used in this procedure. The monoclonal antibodies 12CA5 (anti-HA), 17D09 (anti-HA), and 9E10 (anti-myc) have been used successfully in different laboratories. In general, triple-HA epitope tags work well (Hecht et al., 1996; Aparicio et al., 1997; Tanaka et al., 1997), and larger multi-myc epitope tags have also been successful (e.g., myc-9, myc-18; Tanaka et al., 1997). Protein G-Sepharose, Protein A-Sepharose, and anti-mouse immunoglobulin-coupled magnetic beads have all been used to precipitate the immune complexes, although it should be noted that certain classes of mouse and rat immunoglobulins are not strongly bound by protein A (Harlow and Lane, 1988).

For optimal results, it is critical to minimize the background level of material that inevitably comes down during the immunoprecipitation. The procedures described here work well with a diverse set of antibodies, but it might be necessary to modify the binding and elution conditions in specific cases. Peptide elution is clearly preferred over heat elution, as it is more specific and results in lower experimental backgrounds and hence higher-fold inductions. However, peptide elution is only possible for experiments using antibodies against peptides (typically for analyzing epitope-tagged proteins, but analysis of native proteins should also be possible). In performing peptide elution, it is important to add enough peptide such that the protein-DNA complexes are efficiently eluted from the beads.

Another consideration is that the epitope of interest in the chromatin-bound protein might be inaccessible to the antibody due to associated proteins or DNA structures. In such a case, one might obtain a false-negative result. Whereas the majority of a given protein may be efficiently immunoprecipitated from the cross-linked cells, the fraction that is actually cross-linked might be undetectable. The use of polyclonal antibodies (which often recognize multiple determinants within a protein) or epitope-tagged proteins (the epitope is unlikely to have a specific interaction with other proteins or DNA sequences, particularly if the epitope does not affect the biological function as determined by genetic complementation) minimizes, but does not eliminate this concern. Because of this caveat, negative results should be interpreted cautiously and alternative methods (e.g., *in vitro* DNA binding or association of the protein with bulk chromatin) should be tried. This concern is particularly relevant when a protein of interest does not appear to interact with any genomic sequence. However, if a protein selectively associates with some genomic sequences, this concern is significantly reduced—i.e., it is unlikely that epitope masking will occur at some loci, but not others.

PCR strategy

The choice of primers depends on the experimental goals. If binding to a specific site is being tested, a primer pair that flanks the site and at least one control primer pair recognizing

a DNA sequence not expected to bind the protein of interest are the minimal requirements (see Fig. 21.3.7). When choosing primers, it is important to remember that resolution between adjacent sequences is limited by the shear size of the DNA. For an average DNA size of 500 bp, a significant fraction of the DNA molecules will be in the 500 to 1000 bp range, and hence DNA sequences 1000 bp distal from the actual protein binding site may be coprecipitated. Therefore, primer pairs used as controls should amplify a region of DNA that is far enough away from the expected binding site (e.g., >1 kbp) that coprecipitation of adjacent DNA is not detected. A good strategy is to design multiple sets of primers at increasing distances from a suspected binding site. Such a strategy has also been used to probe the “spreading” and “movement” of proteins on chromatin (Hecht et al., 1996; Aparicio et al., 1997; Strahl-Bolsinger et al., 1997).

Success in obtaining high-quality quantitative data is critically dependent on good primer design! In general, primers should be 20 to 30 bases long with a T_m of 55° to 60°C. Most primers require no purification or special treatment prior to PCR. Amplification products should be 75 to 300 bp. Longer PCR products should be avoided, because the amplification efficiency is substantially lower, and DNA fragments that do not bind to both primers will not be amplified (this can be a significant problem since the size of DNA fragments in the samples averages ~500 bp and ranges between 100 to 1000 bp). A final primer concentration

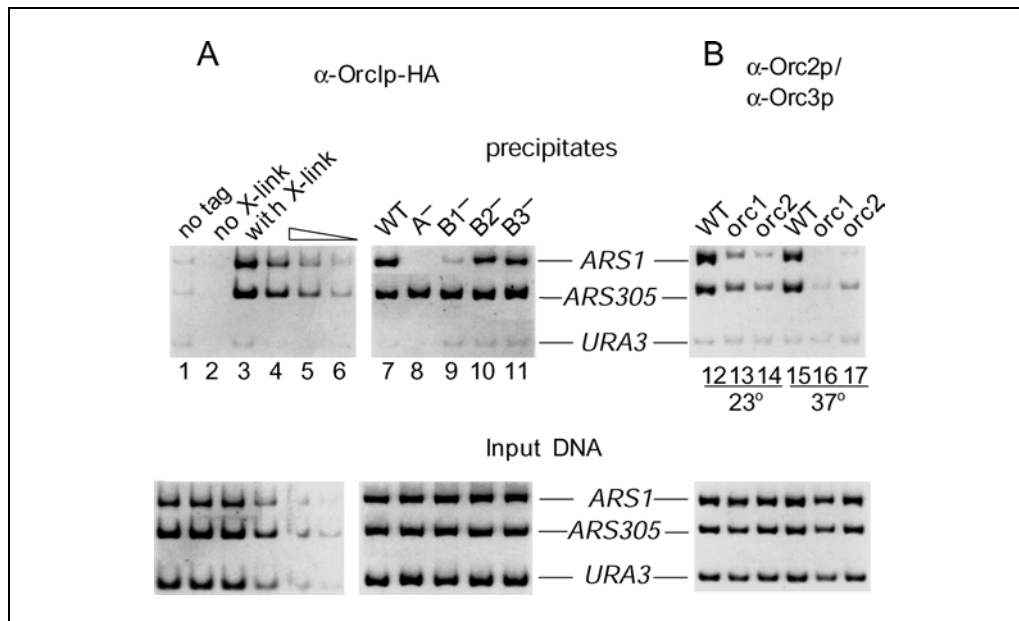


Figure 21.3.7 Anticipated results from chromatin immunoprecipitation analysis of origin recognition complex (ORC) with replication origin and nonorigin DNA sequences.

of 1 μM works well for most primers, but in some instances, improved product specificity may be obtained by lowering the final primer concentration 5 to 10 fold.

The design of good primers is greatly facilitated by commercially available software packages such as Oligo 6.6 or Primer Express 1.5. These packages allow for extensive customization of many different parameters, including T_m , oligonucleotide length, GC content, and more. While the success of each individual primer pair in the specific amplification of its target sequence is dependent on many variables, special care must be taken to minimize primer-dimers and hairpins. Finally, it is a good idea to check primers for hybridization to other genomic sequences through the use of a Web-based program such as BLAST.

Newly obtained primer pairs must be tested for amplification specificity and performance under the conditions that will be used in quantitative PCR. Primer pairs that are suitable for reactions performed by the Basic Protocol 1 might not be suitable for real-time PCR reactions using SYBR Green, because SYBR Green can inhibit *Taq* polymerase. It is particularly informative to analyze input DNA amplification by the primers in question on high-percentage agarose or polyacrylamide gels after completion of the PCR. The presence of multiple product bands indicates poor specificity, and will invariably lead to unreliable results.

For Basic Protocol 1, the best test for quality of a given primer pair is to carry out a standard curve using different dilutions of DNA. For a high-quality primer pair, the amount of PCR product should be directly proportional to the amount of DNA, with an error of less than $\pm 20\%$. The number of PCR cycles is determined empirically. Usually, 25 to 28 cycles is appropriate. More than 28 cycles can result in detection of nonspecifically precipitated sequences and/or lead to variable results due to inactivation of *Taq* polymerase. Multiple primer pairs can be used in combination if the PCR products are separable by gel electrophoresis (as many as five have been used), but some combinations interfere with efficient amplification of one or more products. It is essential to test primer pairs singly and in combination, with titrations of template DNA, to determine if this is a problem. The advantage of using multiple primer pairs is that individual reactions can generate data for multiple genomic regions in an internally controlled manner. In addition, the Basic Protocol 1 can be used to simultaneously analyze two alleles of

a given locus in an internally controlled manner, provided the individual alleles result in different-sized PCR products.

When quantitative PCR will be performed in real time using SYBR Green (see Alternate Protocol 2), high-quality primer pairs should result in ~ 1.9 -fold amplification/cycle. Such amplification efficiency can be determined from quantitative analysis of raw fluorescence data for each cycle. Amplification efficiencies < 1.8 are likely to cause problems, particularly if detection of the PCR product requires 30 cycles or more. Specificity information may also be obtained by running dissociation curves on reactions following the conclusion of the PCR run.

Quantitation

For Basic Protocol 1, PCR products are analyzed by gel electrophoresis and detected by staining with ethidium bromide or SYBR Green or by radioactivity (typically by including a small amount of [^{32}P]dATP in the reactions). DNA staining has the advantage of not having to use radioactivity, but the linear detection range is relatively limited. In this regard, SYBR Green is more sensitive than ethidium bromide and is preferred. Radioactive detection is more sensitive and has a larger linear range than detection by DNA staining. Whatever detection method is employed, it is essential to use a high-quality imager to accurately quantitate the amounts of PCR product in the various reactions.

When quantitative PCR is performed in real time using SYBR Green (Alternate Protocol 2), the linear range is directly visualized and the quality of the reactions can be directly assessed. For reactions involving a given primer, the curves should be superimposable with respect to shape, and should differ only in the number of cycles needed to reach the threshold (C_T). Amplification efficiencies should be ~ 1.9 -fold/cycle. If the curves have different shapes and/or amplification efficiencies are < 1.8 , the reactions are not equivalent and accurate quantitation is impossible.

Data interpretation

In most experiments, it is presumed that the protein of interest associates specifically with certain genomic regions and only associates nonspecifically with other genomic regions. In general, it is very difficult to distinguish true nonspecific association from experimental background of the cross-linking procedure. In this regard, immunoprecipitations with the antibody of interest generally give 2- to 3-fold higher immunoprecipitation efficiencies than

immunoprecipitations with control (or no) antibodies, but it is unclear whether this effect is physiologically meaningful or an experimental artifact.

For this reason, the best way to interpret the data for most experiments is to compare the immunoprecipitation efficiencies for different genomic regions from the same INPUT and IP samples. The immunoprecipitation efficiency is calculated by the amount of PCR product in the IP sample divided by the amount of PCR product in the INPUT sample. A typical background level for DNA fragments that do not associate with the protein of interest is 0.025% to 0.05%. However, background levels can vary, depending on the antibody used and the elution method. In general, monoclonal antibodies give lower background signals than polyclonal antibodies. Peptide elution is preferred over heat elution for the same reason, although this can only be employed for ChIP experiments involving antibodies against defined peptide epitopes. By definition, the background level should be the same for all DNA fragments that do not specifically associate with the protein of interest. In many cases, the choice of suitable negative control regions is based on expectation from other lines of evidence (e.g., the middle of protein-coding regions are unlikely to bind general transcriptional initiation factors). In cases where there is no previous knowledge, the background level can only be based on multiple regions having similar immunoprecipitation efficiencies that are roughly at the level of a typical background immunoprecipitation efficiency. In this regard, it is particularly useful to use proteins tagged with a standard epitope (e.g., HA or myc), as there is considerable information on background levels in such cases.

DNA fragments that display immunoprecipitation efficiencies significantly above the background are indicative of protein association to those genomic regions *in vivo*. Moreover, for a given pair of INPUT and IP samples, the fold enrichment of a given genomic region over the background is directly related to the level of protein association *in vivo*. It is useful to define “relative protein occupancies” for different regions by subtracting the background from the observed immunoprecipitation efficiencies. For example, if the background level is arbitrarily defined as 1 occupancy unit, fragment A that shows 6-fold enrichment over background will have 5 occupancy units, whereas fragment B that shows 21-fold enrichment over background will have 20 occupancy units. Thus, one can conclude

that the protein association with fragment B is 4-fold greater than with fragment A. Without further considerations (see below), this conclusion is only relevant for the particular pair of INPUT and IP samples because absolute immunoprecipitation efficiencies and fold enrichments can vary among repeated trials of the same experiment.

Absolute immunoprecipitation efficiencies and fold enrichments depend on multiple factors. First, the number and physical location of amino acid and nucleotide residues within the interacting protein surfaces that react with formaldehyde vary considerably among protein-protein and protein-DNA interactions. Second, proteins directly interacting with DNA can be cross-linked by a single event, whereas proteins that indirectly associate with DNA need multiple cross-linking events. In this regard, proteins directly binding DNA (e.g., specific DNA binding proteins and general transcription factors) typically give higher fold enrichments than other proteins (e.g., components of chromatin-modifying complexes). Third, some proteins might stably associate with genomic DNA sequences (maximally 100% occupancy), whereas association of other proteins might be transient. Fourth, the absolute immunoprecipitation efficiency depends on the quality of the specific antibody-antigen interaction as well as the antibody concentrations, and the fold enrichment depends on both the absolute immunoprecipitation efficiency and on the background. Thus, absolute immunoprecipitation efficiencies and fold enrichments cannot be used to compare binding characteristics of different proteins. Furthermore, ChIP experiments do not provide absolute measurements of protein occupancy on specific genomic regions or relative stoichiometry of factors on a given sequence. Despite these limitations, ChIP experiments do provide direct quantitative information on the relative levels of protein association on different genomic sequences.

As mentioned above, absolute immunoprecipitation efficiencies and fold enrichments can vary among repeated trials of the same experiment due to potential differences in overall immunoprecipitation efficiency and experimental background. To account for these experimental variations, it is useful to arbitrarily define occupancy units for a specific protein-DNA association. In the example above, one might arbitrarily define fragment A to have 4 occupancy units. In independent experiments, association with fragment A will always be defined as having 4 occupancy units, and

association with other fragments will be defined relative to that of fragment A in the same pair of INPUT and IP samples. In this manner, it is possible to accurately determine the relative level of protein association to multiple genomic regions even though the absolute immunoprecipitation efficiencies and fold enrichments might vary in different repeats of the same experiment. However, in actual practice, differences in absolute immunoprecipitation efficiencies and fold enrichments should not show significant sample-to-sample variation.

A related issue occurs when comparing the level of protein association under different physiological conditions (e.g., different growth conditions or strains). The ideal way to handle this situation is to analyze a “positive control” region that is predicted to be unaffected by the growth condition or genetic constitution. For example, in analyzing association of general transcription factors under a particular environmental condition, it would be useful to examine promoters that are regulated and promoters that are not. In some cases, such a control genomic region is not available, in which case one must rely on simple sample-to-sample reproducibility from independent trials of the same experiment.

ChIP can be used to determine the relative occupancy levels of different proteins at genomic regions (Kuras et al., 2000). Ideally, this is accomplished by performing parallel immunoprecipitations with different antibodies on the same cross-linked chromatin sample. Alternatively, immunoprecipitations can be performed on different samples (this often occurs when using epitope-tagged strains or when multiple proteins are examined). To determine occupancy ratios for two different factors (e.g., X and Y), occupancy units for X and Y are calculated independently as described above. The X:Y occupancy ratios are then calculated for all genomic regions examined. The resulting occupancy ratios are defined in arbitrary units, but the relative occupancy ratios for the different genomic regions are valid. To account for potential sample-to-sample variations among repeats of the same experiment, a given X:Y occupancy ratio should be defined for a specific genomic region and ratios at all other genomic regions calculated in relative terms. Using this rationale, it has been shown that the relative associations of TBP and the general transcription factors TFIIA and TFIIB are essentially constant at all promoters, whereas the TAF:TBP occupancy ratios vary considerably

(Kuras et al., 2000). Importantly, however, occupancy ratios determined from such experiments cannot address whether two proteins co-occupy a given genomic region or mutually compete for the same genomic region.

For some experiments, particularly those involving histone modifications, it is inappropriate to analyze the data in terms of occupancy units and specific versus nonspecific binding sites. Histones associate with essentially all genomic regions, and the level of a particular chromatin modification typically occurs in a continuum. Thus, it is very difficult to determine whether a given region is devoid of a particular modification, although information in this regard can be obtained in control immunoprecipitations using an irrelevant antibody. For these reasons, quantitative analysis of the relative level of a given histone modifications is best presented using simple immunoprecipitation efficiencies. Again, to account for sample-to-sample variations, a specific genomic region should be given an arbitrarily defined value, which is used to determine the relative levels of all other genomic regions.

Anticipated Results

Figure 21.3.7 shows the results of chromatin immunoprecipitation of protein subunits of the origin recognition complex (ORC) and relevant controls (Aparicio et al., 1997). In panel A, immunoprecipitation of Orc1p-HA was shown to specifically coprecipitate the replication origin sequences *ARS1* and *ARS305*, but not the nonorigin DNA sequence *URA3*. Enrichment of *ARS1* and *ARS305* (~0.4% precipitated relative to total) compared to *URA3* (~0.01% precipitated) was ~40-fold and depended on formaldehyde cross-linking and on the epitope-tagged Orc1p (lanes 1 to 3). Mutation of DNA sequences (A and B1) in the *ARS1* replication origin (required in vivo for origin activity and in vitro for binding of ORC to origin DNA) greatly reduced or eliminated association of Orc1p-HA with *ARS1*. The continued association with the wild-type *ARS305* origin served as an additional control (lanes 7 to 9). Analysis of the totals demonstrated that the origin and nonorigin DNA sequences were equally represented in the whole-cell extract (lower panel, “input DNA”). In panel B, origin association of ORC was tested in strains with temperature-sensitive alleles of *ORC1* or *ORC2* demonstrating loss of ORC-origin DNA binding under nonpermissive conditions (lanes 15 to 17). Nevertheless, it should not be assumed that a

mutation in a protein of interest would necessarily result in loss of its chromatin association.

Time Considerations

Basic Protocol 1 may be completed in a 2- or 3-day period. On the first day, cells are fixed with formaldehyde and harvested (1 hr). For convenience, or if preparation of the cells for cross-linking will require an extended period of time beforehand (e.g., induction of expression, cell cycle synchronization), the cells may be frozen and stored at -80°C as described (see Basic Protocol 1, step 6). Preparation of chromatin extracts (2 hr) and immunoprecipitations (primary antibody incubation, ~ 2 hr; incubation with secondary-coupled beads, ~ 1 hr; washing and elution, 2 hr) may be carried out in 1 day, after which the samples are placed at 65°C overnight to reverse the cross-links. If necessary, immunoprecipitation with the primary antibody or bead-coupled secondary antibody may be extended overnight; however, it is most efficient to perform the ≥ 6 hr cross-link reversal step overnight. On the final day, the DNA is purified (< 4 hr, including a 2-hr incubation with proteinase K), PCR amplified (< 3 hr, including 2 hr for the PCR program), and analyzed by gel electrophoresis (< 2 hr, including 1 hr of gel running time).

The entire protocol should take ~ 3 to 4 days.

For Alternate Protocol 2, setup time is < 30 min, thermal cycling takes ~ 2 hr, and data analysis takes < 30 min.

Literature Cited

- Aparicio, O.M., Weinstein, D.M., and Bell, S.P. 1997. Components and dynamics of DNA replication complexes in *S. cerevisiae*: Redistribution of MCM proteins and Cdc45p during S phase. *Cell* 91:59-69.
- Braunstein, M., Rose, A.B., Holmes, S.G., Allis, C.D., and Broach, J.R. 1993. Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. *Genes Dev.* 7:592-604.
- Cosma, M.P., Tanaka, T., and Nasmyth, K. 1999. Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell* 97:299-311.
- Dedon, P.C., Soultz, J.A., Allis, C.D., and Gorovsky, M.A. 1991. A simplified formaldehyde fixation and immunoprecipitation technique for studying protein-DNA interactions. *Anal. Biochem.* 197:83-90.
- Gilmour, D.S. and Lis, J.T. 1984. Detecting protein-DNA interactions in vivo: Distribution of RNA polymerase on specific bacterial genes. *Proc. Natl. Acad. Sci. U.S.A.* 81:4275-4279.
- Gilmour, D.S., Rougvie, A.E., and Lis, J.T. 1991. Protein-DNA cross-linking as a means to determine the distribution of proteins on DNA in vivo. *Meth. Cell Biol.* 35:369-381.
- Harlow, E. and Lane, D. 1998. *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Hecht, A., Strahl-Bolsinger, S., and Grunstein, M. 1996. Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. *Nature* 383:92-96.
- Iyer, V.R., Horak, C.E., Scafe, C.S., Botstein, D., Snyder, M., and Brown, P.O. 2001. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* 409:533-538.
- Kadosh, D. and Struhl, K. 1998. Targeted recruitment of the Sin3-Rpd3 histone deacetylase complex generates a highly localized domain of repressed chromatin in vivo. *Mol. Cell. Biol.* 18:5121-5127.
- Kuras, L. and Struhl, K. 1999. Binding of TBP to promoters *in vivo* is stimulated by activators and requires Pol II holoenzyme. *Nature* 399:609-612.
- Kuras, L., Kosa, P., Mencia, M., and Struhl, K. 2000. TAF-containing and TAF-independent forms of transcriptionally active TBP *in vivo*. *Science* 288:1244-1248.
- Li, X.-Y., Virbasius, A., Zhu, X., and Green, M.R. 1999. Enhancement of TBP binding by activators and general transcription factors. *Nature* 399:605-609.
- Meluh, P.B. and Koshland, D. 1997. Budding yeast centromere composition and assembly as revealed by *in vivo* cross-linking. *Genes Dev.* 11:3401-3412.
- Mencia, M. and Struhl, K. 2001. A region of TAF130 required for the TFIID complex to associate with promoters. *Mol. Cell. Biol.* 21:1145-1154.
- Moqtaderi, Z. and Struhl, K. 2004. Genome-wide occupancy profile of the RNA polymerase III machinery in *Saccharomyces cerevisiae* reveals loci with incomplete transcription complexes. *Mol. Cell. Biol.* 24:4118-4127.
- Orlando, V., Strutt, H., and Paro, R. 1997. Analysis of chromatin structure by *in vivo* formaldehyde cross-linking. *Methods* 11:205-214.
- Ren, B. et al. 2000. Genome-wide location and function of DNA binding proteins. *Science* 290:2306-2309.
- Solomon, M.J. and Varshavsky, A. 1985. Formaldehyde-mediated DNA-protein cross-linking: A probe for *in vivo* chromatin structures. *Proc. Natl. Acad. Sci. U.S.A.* 82:6470-6474.
- Solomon, M.J., Larsen, P.L., and Varshavsky, A. 1988. Mapping protein-DNA interactions in vivo with formaldehyde: Evidence that histone H4 is retained on a highly transcribed gene. *Cell* 53:937-947.
- Strahl-Bolsinger, S., Hecht, A., Luo, K., and Grunstein, M. 1997. SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. *Genes Dev.* 11:83-93.

Tanaka, T., Knapp, D., and Nasmyth, K. 1997. Loading of an MCM protein onto DNA replication origins is regulated by cdc6p and CDKs. *Cell* 90:649-660.

Key Reference

Hecht et al., 1996. See above.

Describes the technique from which the Basic Protocol 1 was adapted.

Orlando et al., 1997. See above.

Describes formaldehyde cross-linking and immunoprecipitation for chromatin analysis in Drosophila, and discusses various parameters of the technique.

Solomon and Varshavsky, 1985. See above.

Characterizes formaldehyde cross-linking, cross-link reversal, and sensitivity of cross-linked protein-DNA complexes to proteases and endonucleases.

Solomon et al., 1988. See above.

Describes original formaldehyde cross-linking and immunoprecipitation technique for mapping protein-DNA interactions.

Contributed by Oscar Aparicio
University of Southern California
Los Angeles, California

Joseph V. Geisberg, Edward Sekinger,
Annie Yang, Zarnik Moqtaderi, and
Kevin Struhl
Harvard Medical School
Boston, Massachusetts

DNase I and Hydroxyl Radical Characterization of Chromatin Complexes

UNIT 21.4

Purified nucleosomes are prepared with radioactively end-labeled DNA from native chromatin or by reconstitution from purified components (Wolffe and Hayes, 1993). The DNA within these complexes is then subjected to DNase I cleavage (see Basic Protocol) or hydroxyl radical cleavage (see Alternate Protocol). Cleavage products are visualized by separation on DNA sequencing gels, followed by autoradiography of the gel. Analysis of the cleavage pattern reveals details of histone-DNA interactions and the extent of nucleosome translational and rotational nucleosome positioning within the chromatin complexes.

NOTE: Radioactive label is used in this unit, and appropriate precautions and shielding should be used (*APPENDIX 1F*).

ANALYSIS OF CHROMATIN COMPLEXES BY DNase I CLEAVAGE

**BASIC
PROTOCOL**

DNase I digestion of purified chromatin complexes can be used to determine what regions of a DNA segment are intimately associated with the core histone proteins and what regions are more like naked DNA (linker DNA within the nucleosomal repeat). DNase I associates with the minor groove of the double helix and cuts DNA endonucleically, leaving a single-stranded nick end (Suck and Oefner, 1986). Thus, cleavage by this enzyme reveals where the minor groove is alternately protected by histone proteins and exposed to bulk solvent on the surface of the nucleosome. For intranucleosomal DNA, this produces a ladder of cleavage products on a sequencing gel in which the bands are spaced by 10 to 11 bp. The protocol below includes the DNase I cleavage reaction and analysis of cleavage products by sequencing gel electrophoresis, as well as an optional nucleoprotein gel electrophoresis fractionation of the nucleosomes. Note that other fractionation procedures such as sucrose gradient sedimentation (*UNIT 5.3*) may also be used.

Materials

- Nucleosome preparation, radiolabeled (see Critical Parameters)
- Nucleosome buffer (see recipe)
- 20 mM MgCl₂
- 2000 U/ml (1 mg/ml) DNase I: purchase 10,000 U lyophilized with glycine (Worthington) and reconstitute with 5 ml H₂O (store in aliquots at -20°C)
- 10× stop solution: 50 mM EDTA/0.2% (w/v) SDS (store up to 6 months at room temperature)
- 3 M and 0.3 M sodium acetate (*APPENDIX 2*)
- 95% and 70% ethanol (-20°C)
- 0.3 M sodium acetate in TE buffer, pH 8.0
- 10 mg/ml Pronase or 10 mg/ml protease K (store up to 1 year at -20°C)
- 2% (w/v) SDS
- 0.1% (w/v) SDS in TE buffer, pH 8.0
- 25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol made with buffered phenol (*UNIT 2.1A*)
- 25 mM EDTA/50% (v/v) glycerol
- 0.7% preparative agarose gel (*UNIT 2.5A*) in 1× Tris·borate buffer, pH 8.3 (see recipe)
- Running buffer for nucleoprotein gel electrophoresis: 1× Tris borate buffer, pH 8.3 (see recipe)
- TE buffer, pH 8.0

**Chromatin
Assembly and
Analysis**

100% formamide containing 1 mM EDTA and 0.005% tracking dyes

Transfer micropipets

90°C water bath

Additional reagents and equipment for phenol extraction and ethanol precipitation of DNA (UNIT 2.1A), agarose gel electrophoresis (UNIT 2.5A), autoradiography and phosphor imaging (APPENDIX 3A), and sequencing gel electrophoresis (UNIT 7.6)

NOTE: Proper sample preparation is critical for both the DNase I and hydroxyl radical cleavage procedures. See Critical Parameters for guidelines before proceeding with the experiments.

Prepare sample of chromatin complexes

1. Prepare samples containing nucleosomes at 100 to 400 ng/μl and volumes ranging from 10 μl to several hundred μl (although a convenient size is 20 to 40 μl) in nucleosome buffer.

Most nucleosomal complexes are stable under the conditions described for the nucleosome buffer (see Reagents and Solutions) although it is noted that oligonucleosomal templates will exist in a full extended conformation at these ionic strengths (Hansen, 1997). Chromatin complexes such as mononucleosomes can be separated from any naked DNA remaining after the reconstitution by sucrose or glycerol gradient sedimentation (UNIT 5.3) before the digestion. This obviates the need for nucleoprotein gel fractionation of digested complexes (see Fig. 21.4.1).

2. Equilibrate sample at room temperature.

Digestions of very concentrated material may be carried out at 37°; however, digestions usually are more easily carried out at ambient temperature.

3. If necessary, bring the Mg²⁺ concentration to 0.5 to 1.0 mM by adding 1/20 vol of 20 mM MgCl₂ stock solution.

This step is omitted if the Mg²⁺ concentration is already in molar excess over chelation agents such as EDTA (e.g., in cases where Mg²⁺ is present in the transcription buffer). If EDTA or other metal-binding species are present, adjust Mg²⁺ concentration to be 0.5 to 1.0 mM in excess of the chelation agent.

DNase I requires Mg²⁺ as a cofactor. However, note that with some oligonucleosomal complexes addition of Mg²⁺ salts must be done with care—i.e., slowly with rapid mixing—as concentrations in excess of a few mM can cause aggregation and precipitation of oligonucleosome substrates (Schwartz et al., 1996).

Digest with DNase I

4. Add 0.5 to 1 μl of a 1:50 to 1:500 dilution of the 2000 U/ml DNase I stock.

The authors have found 0.1 to 0.2 mg enzyme per mg nucleic acid to be optimal under the experimental conditions used in their laboratory. However, the optimal concentrations of DNase I must be determined empirically, since the extent of nucleolytic activity varies as a function of DNA concentration and ionic strength.

Under optimal conditions approximately 2/3 to 3/4 or more of the original DNA fragment should remain uncut by the enzyme (see Troubleshooting).

5. Incubate the digestion for 1 to 2 min.

Terminate the reaction and purify the DNA digest

The method used to terminate the reaction depends on the subsequent steps to be followed. If nucleosomes are free of contaminating DNA (i.e., have been previously purified by gradient isolation), then the gel step is not necessary. If complexes are to be fractionated by nucleoprotein gel electrophoresis, proceed to step 6c. If the complexes are relatively

pure or have been purified by gradient fractionation the digestion may be stopped and the DNA directly prepared for sequencing gel electrophoresis (step 6a).

For complexes that are relatively pure or that have been purified by gradient fractionation:

6a. Stop the reaction by adding 10× stock of stop solution to a final concentration of 1×.

Preparations of reconstituted nucleosomes containing protein and DNA concentrations <0.1 mg/ml do not require protease treatment and phenol extractions. These may be precipitated directly with ethanol/salt in the presence of SDS. The DNA pellet is then dried and analyzed directly.

7a. Adjust the solution to 0.3 M sodium acetate by adding $\frac{1}{10}$ vol of 3 M sodium acetate. Precipitate the DNA by adding 2 to 2.5 vol of -20°C 95% ethanol. Collect the DNA pellet by microcentrifuging 20 to 30 min at maximum speed, room temperature, and carefully removing the supernatant with a transfer micropipet.

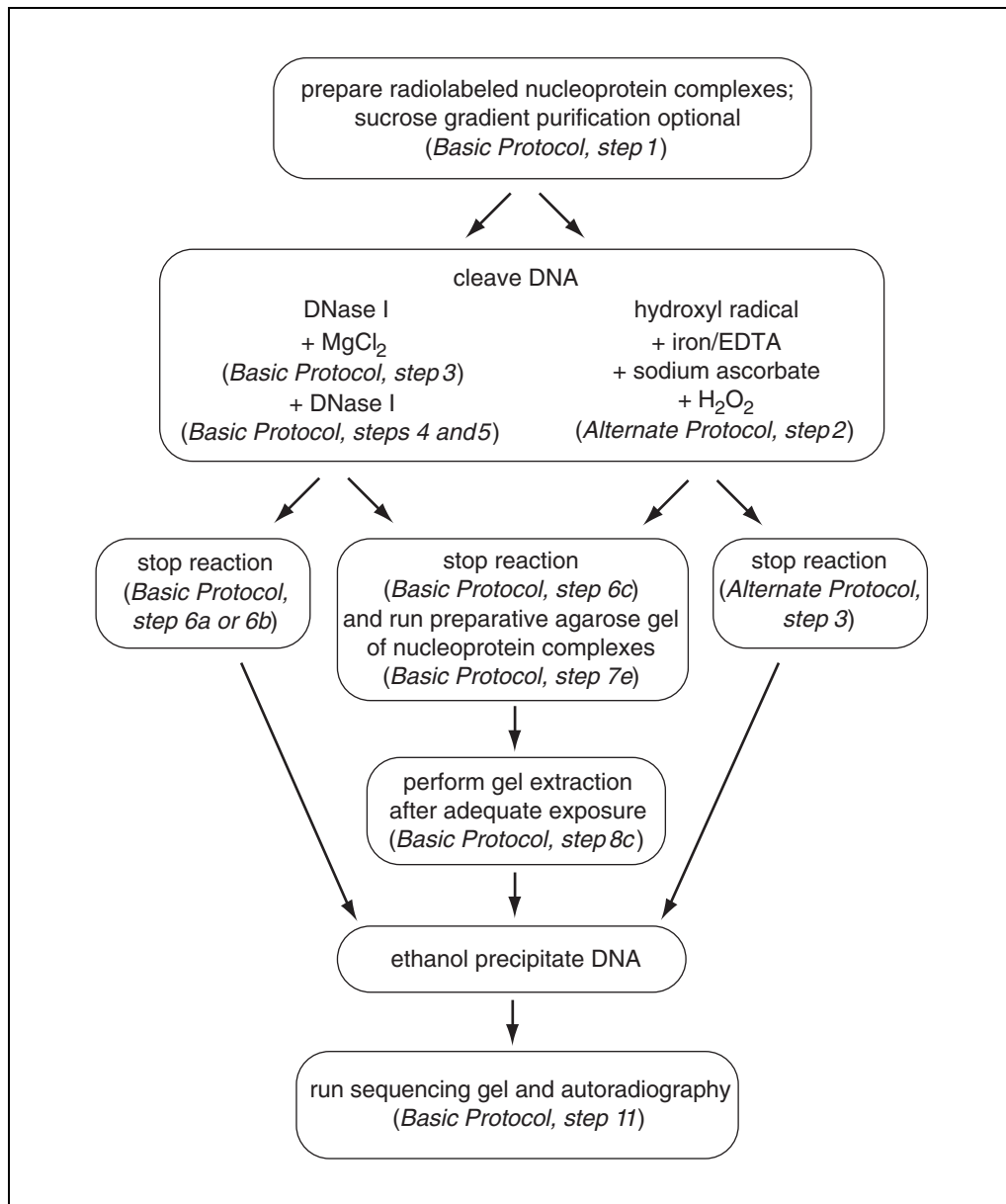


Figure 21.4.1 Flow chart for DNase I and hydroxyl radical cleavage analysis of chromatin complexes.

- 8a. Resuspend the DNA pellet in 0.3 M sodium acetate/TE buffer, pH 8.0, and reprecipitate with ethanol as before to remove additional salt and SDS. Wash the final pellet gently with 400 μ l of ice-cold 70% ethanol.

For complexes containing larger quantities of histone proteins:

- 6b. Add a 10 mg/ml stock solution of Pronase or a 10 mg/ml stock solution of protease K to ~0.25 mg/ml final concentration. Also add 1/10 vol of 2% SDS (0.2% final). Incubate 2 to 3 hr at 37° to 42°C.

Complexes containing histone protein concentrations in excess of 0.1 mg/ml, as is typical with preparations of native chromatin, should be digested with proteinase K and phenol/chloroform extracted before ethanol precipitation.

- 7b. Adjust the sample volume to 180 μ l with 0.1% SDS/TE buffer, pH 8.0. Extract with 1 vol 25:24:1 phenol/chloroform/isoamyl alcohol made with buffered phenol (*UNIT 2.1A*). Adjust the aqueous phase to 0.3 M sodium acetate by adding 20 μ l of 3 M sodium acetate, then precipitate the DNA by adding 550 μ l of -20°C 95% ethanol. Collect the DNA pellet by microcentrifuging 20 to 30 min at maximum speed, room temperature, and carefully removing the supernatant with a transfer micropipet.
- 8b. Resuspend the DNA in 0.3 M sodium acetate/TE buffer, pH 8.0, and reprecipitate with ethanol as before to remove additional salt and SDS, Wash the final pellet gently with 400 μ l of ice-cold 70% ethanol.

To separate digested nucleosomes from naked DNA (nucleoprotein gel electrophoresis):

- 6c. Stop the reaction by adding $\frac{1}{10}$ vol 25 mM EDTA/50% glycerol.
- 7c. Immediately load the sample onto a 0.7% preparative agarose gel and perform electrophoresis in 1 \times Tris-borate buffer, pH 8.3, for 3 hr with an applied voltage of 5 V/cm. When electrophoresis is complete, wrap the gel tightly in plastic wrap and expose to autoradiographic film (*APPENDIX 3A*).

Typically, all visual tracking dyes such as bromophenol blue or xylene cyanol are omitted from nucleosome samples but are loaded in adjacent wells to gauge the progress of the gel. The bromophenol blue dye is typically run approximately $\frac{1}{2}$ to $\frac{2}{3}$ of the way through a standard 10 to 12 cm horizontal agarose gel.

For the autoradiography, the authors have found that the gel must be exposed for 1.5 to 3 hr before sufficiently dark bands are obtained. It is crucial at this point to be able to precisely align the gel with the exposure after developing; typically, phosphorescent markers available from Stratagene are placed directly on the surface of the plastic wrap. Alternatively, one can simply cut one corner from the gel, wrap it in plastic wrap, then carefully trace around the edges of the gel with a marker right on the film. Care must be taken so that the film does not shift position as the gel is prepared for exposure (the gel can be fastened at the corners with transparent adhesive tape). The gel may also be aligned by preparation of radioactive "marker dye." Some old radioactive nucleotides are mixed with concentrated bromophenol blue dye and spotted into the gel with a micropipet. The dye will diffuse slightly; however the original point of placement usually can be visualized on the gel and the autoradiogram. It is also possible to prepare a "radioactive marker pen" by dipping the end of a marker in some old source radioactivity and then writing directly onto the plastic wrap.

- 8c. Cut out the digested nucleosome and naked DNA bands and isolate the radiolabeled DNA (*UNIT 2.7*). Adjust the DNA solution to 0.3 M sodium acetate by adding $\frac{1}{10}$ vol of 3 M sodium acetate. Precipitate the DNA by adding 2 to 2.5 vol of -20°C 95% ethanol. Collect the DNA pellet by microcentrifuging 20 to 30 min at maximum speed, room temperature, and carefully removing the supernatant with a transfer micropipet. Resuspend the DNA pellet in 0.3 M sodium acetate/TE buffer and

reprecipitate with ethanol as before to remove additional salt and SDS. Wash the final pellet gently with 400 μ l of ice-cold 70% ethanol.

The DNA can be recovered from the gel by electroelution of the isolated band (UNIT 2.7) or by standard “freeze and squeeze” procedures using a small filtration unit that fits into the top of a microcentrifuge tube (pore size 0.4 μ m; e.g., March Biomedical). If the latter procedure is used, the authors find it improves recovery to reextract the plug of agarose in the filter with 0.5 ml TE/0.2% SDS solution either overnight at room temperature or for 2 hr at 37°C with constant gentle mixing on a rotator.

Perform sequencing gel electrophoresis

9. Resuspend the DNA pellet from step 8a, b, or c in 10 to 20 μ l of TE buffer, pH 8.0.

The exact volume will depend on how many radioactive counts are to be used in the experiment. The 10 to 20 μ l volume allows all radioactivity to be more easily resuspended by vortexing. Note that, typically, much of the precipitated, dried DNA is stuck to the inner wall of the microcentrifuge tube and care must be taken to resuspend all the radioactive material. This can be done by vigorous vortexing and/or by rubbing the droplet of TE buffer against the inner wall of the tube. The efficiency of resuspension in the formamide dye solution is checked with a benchtop Geiger counter. Virtually all of the counts should be resuspended.

10. Briefly microcentrifuge the resuspended DNA to collect the solution at the bottom of the tube and to bring down any insoluble material from the gel isolation. Pipet approximately equivalent numbers of counts from each sample (minimally 2000 to 5000 cpm) into fresh microcentrifuge tubes (ideally this should be \sim 1/4 of the total resuspended DNA, or \sim 5 μ l). Evaporate to dryness in a Speedvac evaporator.

If all of the solution must be used, attempt to pipet the solution without disturbing the pellet of insoluble material.

11. Resuspend the DNA in \sim 3 μ l of 100% formamide containing 1 mM EDTA and 0.005% tracking dyes, denature at 90°C for 2 to 5 min, place on ice, and load into the well of a sequencing gel (UNIT 7.6). Run the gel and expose to autoradiographic film or a phosphor imager plate (APPENDIX 3A).

FOOTPRINTING WITH HYDROXYL RADICALS

Chemical nucleases such as hydroxyl radicals can be used in place of enzymatic nucleases to probe histone-DNA interactions within nucleosomes. Hydroxyl radicals can be chemically generated in solution and have many advantages over enzymatic nucleases, such as their small size and ability to cleave the DNA backbone at every position without sequence selectivity. For these reasons, much more detailed and precise mapping information can be obtained by footprinting with hydroxyl radicals than with DNase I. However, hydroxyl radical footprinting requires solutions prepared without radical scavengers and chromatin complexes prepared with virtually nick-free DNAs (see Background Information).

Additional Materials (also see Basic Protocol 1)

Glycerol

1 M Tris·Cl, pH 8.0 (APPENDIX 2)

Iron/EDTA working solution (see recipe), freshly prepared

10 mM sodium ascorbate [prepare fresh from 1 M stock stored at -70°C (stable >1 year; e.g., Aldrich)]

0.12% H_2O_2 , prepared fresh by 1:250 dilution from 30% H_2O_2 (stock solution stable up to several years at 4°C if kept free of metal contamination)

ALTERNATE PROTOCOL

Chromatin Assembly and Analysis

21.4.5

1. Add glycerol to the sample to a concentration of $\leq 0.2\%$ (v/v) and 1 M Tris-Cl, pH 8.0 to a concentration of ≤ 10 mM.

Most salts and EDTA may be present at any concentration (Tullius, 1987). Any volume may be used; typically nucleosomes are prepared in $\frac{7}{10}$ of the final reaction volume.

2. Initiate the cutting reaction by placing a drop each of the iron/EDTA working solution ($\frac{1}{10}$ vol), 10 mM sodium ascorbate ($\frac{1}{10}$ vol), and the 0.12% H_2O_2 ($\frac{1}{10}$ vol) on the inner wall of a 1.5-ml microcentrifuge tube which already contains the nucleoprotein sample ($\frac{7}{10}$ vol). Allow the three added reagents to mix briefly. Immediately after the addition of the peroxide solution, using the pipet tip, add the reagents to the nucleoprotein solution, and mix thoroughly by pipetting.

For example, for digestion of a 35 μl solution of nucleosomes 5 μl of each of these reagents would be mixed together on the side of the tube and then rapidly mixed with the sample.

3. Incubate the free radical reaction 1 to 2 min at room temperature, then quench by adding glycerol to 5% (v/v) and transfer the entire reaction volume transferred to a nucleoprotein gel (see Basic Protocol, step 7c).

Alternatively, the reaction may be adjusted to 0.3 M sodium acetate and precipitated with ice-cold ethanol (see Basic Protocol). The presence of ethanol effectively quenches the radical cleavage reaction. If considerable time (minutes) is to elapse between the cleavage reaction and ethanol precipitation, the reaction may be quenched by the adjusting the solution to 50 mM thiourea (Tullius and Dombroski, 1985).

4. Analyze the DNA cleavage products (see Basic Protocol).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Iron/EDTA solution

Stock solution: Prepare a small amount (typically 1 ml) of 100 mM ferrous ammonium sulfate hexahydrate $[\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}]$; e.g., Aldrich] containing 200 mM EDTA in water and store up to 1 year frozen at -70°C .

Working solution: Just before use, dilute stock solution 1:99 with water for a final concentration of 1 mM Fe(II)/2 mM EDTA.

See Troubleshooting for guidelines on stability of this solution.

Nucleosome buffer

10 mM Tris-Cl, pH 8.0 (APPENDIX 2)

<0.1 mM EDTA

0 to 50 mM NaCl

This is an optimal buffer for DNase I digestion of chromatin. A small amount of EDTA protects the DNA from metal-mediated damage. The amount of NaCl will depend on the preparation method and application.

Tris-borate buffer, pH 8.3, 10 \times

108 g Tris base

55 g boric acid

H_2O to ~ 800 ml

Adjust pH to 8.3

Add H_2O to 1 liter

COMMENTARY

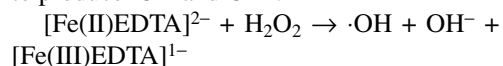
Background Information

Analysis of histone-DNA interactions within nucleosomes. The native chromatin complex within most eukaryotic nuclei is very difficult to study by biochemical means. However, as described elsewhere in this chapter, researchers have developed methods for preparation of smaller and experimentally tractable portions or forms of this complex. Often these methods involve reconstitution of the fundamental subunit of the chromatin complex, the nucleosome, from purified components (Wolffe, 1995). Once prepared, the integrity of these complexes can be ascertained by assays for protein composition, hydrodynamic behavior, DNA supercoiling, *trans*-acting factor binding, and a myriad of other biochemical and biophysical approaches (van Holde, 1989). In addition, an important and unique feature of chromatin is the distinct structure adopted by nucleosomal DNA. The DNA is wrapped into $\sim 1\frac{3}{4}$ tight 80 base pair circles on the surface of the histone octamer and contains distinct regions of helical periodicity that differ from the helical period of bare DNA (Hayes et al., 1990, 1991; Luger et al., 1997).

This unit details common but extremely powerful methods to assay DNA structure and accessibility within purified nucleosome complexes. Nucleosome-specific cleavage patterns obtained with DNase I or hydroxyl radicals typically comprise an alternating pattern of cleavage and protection with a frequency of 10 to 11 base pairs. This pattern, often referred to as a “nucleosome ladder,” reflects the helical structure of the DNA backbone, which alternately contacts the histone proteins or is oriented away from the surface of the histone octamer, toward the bulk solvent. Indeed, quantitative analysis of this pattern has been used to precisely measure the helical periodicity of DNA within the nucleosome (Hayes et al., 1990). In addition, the ladder not only indicates the region of DNA in contact with core histones but also the degree of rotational positioning of the DNA on the surface of the nucleosome(s).

DNase I digestion can reveal many details about the location and extent of histone-DNA interactions and can be used in complicated extracts containing nucleosomal complexes. However, the finer details of histone-DNA interactions and DNA structure within these complexes are best characterized by digestion of the DNA with the hydroxyl radical, an efficient chemical nuclease. Because of the small size

and lack of sequence specificity of this diffusible reactive species, very fine details of DNA conformation and protein-DNA contacts can be resolved. These features are reflected in the rates of DNA cleavage at individual nucleotide positions. Cleavage is effected by the (probable) abstraction of a hydrogen atom by a hydroxyl radical from a deoxyribose carbon center in the minor groove of the DNA. This leaves a carbon-centered radical which rapidly disintegrates to leave 5' and 3' monophosphoester termini on either side of a one nucleoside gap. Hydroxyl radicals can be efficiently produced in solution in virtually any salt conditions and in the presence of any soluble protein-DNA assembly via the one-electron reduction of hydrogen peroxide by an iron(II)-EDTA complex to produce $\cdot\text{OH}$ and OH^- :



The oxidation of Fe(II) is readily reversible by mild reducing agents such as sodium ascorbate, thus effectively recycling the iron complex until the peroxide is depleted. See Tullius et al. (1987) for a complete review of hydroxyl radical mediated DNA cleavage. Note that the hydroxyl radical cleavage procedure has the drawback that the complexes must be prepared in relatively pure form and buffer solutions must be free of radical scavengers. Common buffer components such as glycerol, Tris buffer in excess of 10 mM, and carbohydrates must be avoided.

Critical Parameters

Proper sample preparation is critical for the DNase I as well as the hydroxyl radical cleavage method. For DNase I digestions, samples of nucleosomes or other chromatin complexes should be prepared with DNA containing a unique radioactive end label or a label incorporated after the digestion. In addition, the complexes should be in appropriate buffers. A typical sample for digestion of reconstituted nucleosomes contains ~ 50 to 200 ng of cold “carrier” DNA and 2 to 10 ng (2000 to 20,000 cpm) of radioactively end-labeled specific DNA restriction fragment reconstituted with 25 to 100 ng of purified core histone proteins. Note that samples to be treated with DNase I prior to resolving nucleoprotein complexes on preparative 0.7% agarose gels should contain $>20,000$ cpm (see below).

1. DNase reactions generally require buffers lacking metal-chelation agents such as

EDTA or ATP. Alternatively, the concentration of these species should be <1 mM, since free Mg^{2+} is required for activity of the enzyme and must be added in excess of chelators.

2. For DNase I digestions, samples should contain less than ~ 100 to 150 mM ionic strength, since the activity of the enzyme decreases considerably in elevated salt solutions. For radioactively end labeled complexes, it is recommended that each sample contain at least $5,000$ cpm. Samples containing less radioactivity can be used, depending on the losses during sample preparation, but at least 1000 to 2000 cpm should be loaded in each lane of the sequencing gel.

3. For analysis of natural, random-sequence chromatin fragments, similar masses of material may be used. These are typically prepared by micrococcal nuclease digestion, and since this enzyme leaves $5'$ -OH, these complexes are directly end-labeled with T4 polynucleotide kinase and $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ (UNIT 3.10). In some cases it is advantageous to examine the DNase I digestion pattern of specific sequences within the natural chromatin by indirect end labeling via DNA blotting or LM-PCR approaches (UNIT 21.1).

Troubleshooting

This section reviews issues concerning cleavage evaluation, common problems with the hydroxyl radical cleavage reaction, and ways to optimize the cleavage. However, comments concerning background (points 2 and 3, below) and gel effects (point 6) are also relevant to experiments with DNase I.

1. *Evaluation of hydroxyl radical cleavage.* Since the hydroxyl radical yields an even pattern of cleavage at every nucleotide position, the intensity of individual bands on the gel is somewhat less than that obtained with DNase I. Thus, common problems are the lack of observable cleavage on the gel and determining if sufficient cleavage has occurred. The lack of an apparent cleavage pattern on the gel may be due to underexposure of the gel. A simple method to ensure sufficient exposure is to load a mock-treated sample containing the same number of counts in an adjacent lane. The gel is exposed to the point that the “background” within the control lane is barely visible and compared to the signal in the experimental lane.

2. *Determination of background.* The hydroxyl radical cleavage pattern within DNAs containing a high amount of inherent background cleavage may be difficult to see. The amount of background cleavage may be judged

by running a mock-treated sample as described above and directly evaluating the autoradiogram. Alternatively, after a brief exposure, the full-length band may be cut out of the sequencing gel and the total counts within this band determined by scintillation counting. The total counts within this band should be $\geq 95\%$ of the counts loaded in the lane.

3. *Background remediation.* High background may be due to several problems. Note that labeled DNA should be stored in the presence of EDTA, if possible, and at less than $\sim 10,000$ cpm per μl to avoid autoradiolysis. In addition, DNA subjected to treatment with $\text{Fe(II)}\text{-EDTA}$ should never be stored as a precipitate immediately after the digestion; the DNA should be processed at least through the rinse step. Another source of background comes from the use of partially hydrolyzed formamide or excessive heating before loading the sequencing gel. DNA should not be heated in formamide at temperatures greater than 95°C , and the formamide should be stored frozen at -20°C between uses. These problems are usually manifested as elevated background cleavage at purine residues within the DNA.

4. *Buffer components.* A lack of cleavage may be due to the presence of radioprotectants within the nucleosome solutions. Buffer components such as glycerol in excess of $\sim 0.2\%$, Tris in excess of 10 mM, and carbohydrates such as mannitol or dextrose should be avoided.

5. *Reagents.* The efficiency of radical production is greatly reduced in acidic solutions. Thus, it is important to employ sodium ascorbate. Prepare sodium ascorbate solutions from dry crystals that do not appear to have become moist or reddish in color. The authors have found that 1 M solutions of this reagent are stable for >1 year when stored at -70°C and the 10 mM working stocks are stable for >3 months when stored frozen. Likewise the working stocks of the Fe(II)EDTA complex are stable if kept frozen for extended periods. Note that the appearance of a reddish-brown color in these solutions indicates significant oxidation of the iron and formation of insoluble μ -hydroxy iron complexes, and that the stock should be discarded. Although solutions of 30% H_2O_2 are stable for several years if stored properly at 4°C , contamination of this solution with transition metals can cause a loss of peroxide. In addition, the diluted working stock solution of hydrogen peroxide is relatively unstable on the bench top and should be discarded immediately after the reaction.

6. *Band smearing*. “Smearly” bands may be due to several problems. First, a common mistake is to use too much carrier DNA in the preparation of samples. For example, an excess of tRNA added as carrier will cause a smear in the region 120 to 150 bp on the sequencing gel. Most nucleosome samples have sufficient DNA for quantitative precipitation. Another source of “fuzziness” is loading too much of the final sample into the well. This sample typically contains relatively insoluble contaminants from the agarose gel. If these are loaded into the well, the heated sample has a very viscous character and yields a poorly resolved lane. If necessary, remove 20 μ l of the 25 total resuspension in TE after a brief microcentrifuge spin of this material to sediment insoluble contaminants.

Anticipated Results

If the starting material consists of well positioned nucleosomes, the DNase cleavage should yield a pattern with maximal cleavages every 10 bp and minimal cleavages in between. This pattern will not be completely regular because DNase I shows some sequence specificity in its cleavage activity. Hydroxy radical cleavage will produce a much more regular cleavage pattern, with enhanced cleavages every 10 bp, and a gradual decrease, then increase, in between maximal cleavages. If nucleosomes are not well positioned (i.e., if the preparation contains a mixture of different positions), then no 10-bp periodicity will be observed.

Time Considerations

Performing the cleavage reactions requires <1 hr total if all starting materials (including labeled nucleosomes) are available. Preparation of the sample for electrophoresis requires several hours, and electrophoresis itself can be accomplished in several hours. The entire procedure can be completed in a full day. 1- to 2-day exposures should be sufficient to observe the signal.

Literature Cited

- Hansen, J.C. 1997. The core histone amino termini: Combinatorial interaction domains that link chromatin structure with function. *Chemtracts* 10:737-750.
- Hayes, J.J., Tullius, T.D., and Wolffe, A.P. 1990. The structure of DNA in a nucleosome. *Proc. Natl. Acad. Sci. U.S.A.* 87:7405-7409.
- Hayes, J.J., Bashkin, J., Tullius, T.D., and Wolffe, A.P. 1991. The histone core exerts a dominant constraint on the DNA in a nucleosome. *Biochemistry* 30:8434-8440.
- Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389:251-260.
- Schwartz, P.M., Felthouser, A., Fletcher, T.M., and Hansen, J.C. 1996. Reversible oligonucleosome self-association: Dependence on divalent cations and core histone tail domains. *Biochemistry* 35:4009-4015.
- Suck, D. and Oefner, C. 1986. Structure of DNase I at 2.0 Å resolution suggests a mechanism for binding to and cutting DNA. *Nature* 321:620-625.
- Tullius, T.D. 1987. Chemical snapshots of DNA: Using the hydroxyl radical to study the structure of DNA and DNA-protein complexes. *Trends Biochem. Sci.* 12:297-300.
- Tullius, T.D. and Dombroski, B.A. 1985. Iron(II) EDTA used to measure the helical twist along any DNA molecule. *Science* 230:679-681.
- Tullius, T.D., Dombroski, B.A., Churchill, M.E., and Kam, L. 1987. Hydroxyl radical footprinting: A high-resolution method for mapping protein-DNA contacts. *Methods Enzymol.* 155:537-558.
- van Holde, K.E. 1989. *Chromatin*. Springer-Verlag, New York.
- Wolffe, A.P. 1995. *Chromatin: Structure and Function*. Academic Press, London.
- Wolffe, A.P. and Hayes, J.J. 1993. Transcription factor interactions with model nucleosome templates. *Mol. Genet.* 2:314-330.

Contributed by Joseph M. Vitolo,
Christophe Thiriet, and Jeffrey J. Hayes
University of Rochester Medical Center
Rochester, New York

Isolation of Histones and Nucleosome Cores from Mammalian Cells

UNIT 21.5

In vitro analysis of DNA in chromatin is often important for understanding mechanisms of regulation of transcription and other processes that occur on DNA. The basic unit of chromatin is the nucleosome core, containing two copies each of the core histones H2A, H2B, H3, and H4 to form a histone octamer that wraps ~145 base pairs of DNA in a left-handed superhelix. In vivo, chromatin is associated with linker histones (such as H1), which facilitate the ordered packing of nucleosomes. This linker histone-containing particle is properly termed the nucleosome (or chromatosome), while the linker histone-free particle is the nucleosome core.

Long arrays of nucleosome cores (poly- or oligonucleosome cores) as well as pure populations of mono- and dinucleosome cores can be purified. These purified nucleosomes are useful for a variety of applications; for example, they can act as histone donors for the assembly of nucleosomes on defined-sequence DNA. The core histones can be purified away from chromosomal DNA and from linker histones, as well. These are useful in a variety of procedures for in vitro chromatin assembly. The method here does not denature the histones (as does the acid extraction protocol in UNIT 21.2), and may yield histones that are more active for in vitro assemblies.

Pure polynucleosome cores and histones are useful in the in vitro assembly of chromatin on defined-sequence DNA. These can be readily isolated from mammalian tissue culture cells. In Basic Protocol 1, nuclei are isolated and purified. In Basic Protocol 2, polynucleosomes lacking linker histones are isolated from these nuclei. In Basic Protocol 3, pure populations of mono- and dinucleosome cores are isolated from oligonucleosome fractions from Basic Protocol 2. In Basic Protocol 4, core histones are isolated from purified nuclei.

The large genome size of mammals makes mammalian tissue culture cells abundant sources for chromatin and histones. The core histones are highly evolutionarily conserved, and human histones are frequently used in assembly systems derived from *Drosophila* and for in vitro transcription systems from eukaryotes as simple as yeast. The assumption is that, at the level of simple hypo-acetylated nucleosome cores, histone source matters little. Thus far, this assumption appears to be warranted. However, when histone modifications, linker histones, and other nonhistone chromosomal proteins begin to be considered, the organismal, tissue, and cell-type source become increasingly important.

NOTE: Unless otherwise indicated, keep all solutions and materials on ice or at 4°C.

PREPARATION OF A WASHED NUCLEAR PELLET

The isolation of nucleosomes and free histones requires clean nuclei as starting material. In this protocol, nuclei are released from cultured cells by physical shear forces (homogenization) and nonionic detergent (NP-40; Côté et al., 1995). Several washes with a buffer containing detergent are then used to remove membranes and yield a pellet containing relatively clean nuclei. Subsequent extraction with 0.3 M KCl removes most loosely bound proteins from chromatin. This yields a nuclear pellet and what is, effectively, a nuclear extract (from which several transcription factors can be isolated). However, the NP-40 washes may make the extract less active than nuclear extracts prepared by other means. If both nuclear extract and chromatin are desired, the nuclear extract should be

**BASIC
PROTOCOL 1**

**Chromatin
Assembly and
Analysis**

21.5.1

Contributed by Gavin R. Schnitzler

Current Protocols in Molecular Biology (2000) 21.5.1-21.5.12

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 50

prepared as in *UNIT 12.1*, and the “spent” nuclear pellet from this procedure can be used to prepare chromatin (Schnitzler et al., 1998).

Materials

Mammalian tissue cell culture (e.g., HeLa cells)

PBS (*APPENDIX 2*)

Lysis buffer (see recipe)

Buffer B (see recipe)

2 M NaCl

Buffer B/0.6 M KCl/10% (v/v) glycerol

Dounce homogenizer with type B pestle

Light microscope

Additional reagents and equipment for mammalian cell culture (*APPENDIX 3F*) and quantitation of DNA (*APPENDIX 3D*)

1. Grow and harvest 3 liters of mammalian tissue culture cells (e.g., HeLa) at 1×10^6 cells/ml and wash twice in 1 liter PBS.

See APPENDIX 3F for counting cells.

2. Resuspend pellet in 20 pellet vol (~40 ml) lysis buffer, transfer to a Dounce homogenizer, and lyse cells with ~15 strokes of a type B pestle. Monitor lysis by light microscopy.

The cells should be clearly lysed when viewed by light microscopy. The NP-40 in the buffer solubilizes membranes, resulting in a clean (white or light tan) pellet.

Alternatively, the nuclear pellets remaining after nuclear extract preparation (UNIT 12.1) can be resuspended by homogenization in the same buffer. If this is done, wash as described in steps 2 to 4, and then proceed to purification of H1-depleted oligonucleosomes (see Basic Protocol 2).

3. Centrifuge 15 min at $3000 \times g$, 4°C .
4. Repeat resuspension and pelleting twice with lysis buffer and once with buffer B. Just before the final pelleting, estimate the yield of DNA in the nuclei by diluting 10- to 30-fold into 2 M NaCl and measuring A_{260} (*APPENDIX 3D*).

A few strokes with the type B pestle can be used to help resuspend the pellet.

5. Resuspend nuclei in 2 pellet vol buffer B and measure the total volume of the suspension.
6. While gently stirring, add 1 total vol of buffer B/0.6 M KCl/10% glycerol in a dropwise fashion. Continue gentle stirring for 10 min at 4°C . Homogenize gently, if necessary.
7. Pellet nuclei 30 min at $17,500 \times g$, 4°C .

Nuclear pellets can be frozen in liquid nitrogen or dry ice and kept for more than a year at -80°C .

SOLUBILIZATION AND PURIFICATION OF HISTONE H1-DEPLETED OLIGONUCLEOSOMES

BASIC
PROTOCOL 2

Oligonucleosome cores lacking histone H1 are commonly used to assemble labeled, defined-sequence mononucleosomes by the method of octamer transfer at high salt, amongst other applications. The NP-40-washed nuclear pellet is washed at 0.4 M NaCl to remove nonhistone proteins. The pellet is then solubilized in 0.65 M NaCl. At this salt concentration, the linker histones are released from the nuclei, resulting in chromatin decompaction. The resultant nuclei are much less resistant to shear forces, and vigorous homogenization fragments the chromatin to a size that allows their separation from other nuclear fragments. After dialysis to reduce the salt concentration, digestion with the endonuclease micrococcal nuclease releases even smaller fragments. These fragments are size-separated by either gradient centrifugation or gel filtration, and different size distributions are pooled and stored frozen. Size separation is done at salt concentrations (0.6 M NaCl) at which linker histones do not bind, allowing their efficient separation from oligonucleosome cores (Côté et al., 1995; Schnitzler et al., 1998).

Materials

Washed nuclear pellet (see Basic Protocol 1)
MSB (see recipe)
HSB (see recipe)
2 M NaCl
LSB (see recipe)
0.1 M CaCl₂
50 U/μl micrococcal nuclease (see recipe)
0.5 M EGTA, pH 8.0
HSB without sucrose
HSB/glycerol: HSB containing 10% and 40% (v/v) glycerol instead of sucrose
0.5% (w/v) SDS
0.5 mg/ml proteinase K
Dialysis buffer (see recipe)

Dounce homogenizer with type B pestle
6 to 8 kDa MWCO dialysis membrane
~1.6 × 58-cm Sepharose CL-6B column (Amersham Pharmacia Biotech; for gel filtration) and accessories
Refrigerated ultracentrifuge with appropriate rotor and tubes, e.g.,
Beckman SW55 rotor with polyallomer or ultraclear tubes (for gel filtration)
Beckman SW28 rotor with 1 × 3.5-in. (2.5 × 8.9-cm) polyallomer tubes (for centrifugation)
Gradient maker (for centrifugation)
21-G needle (for centrifugation)
Tubing (for centrifugation): e.g., 21-G needle infusion set with 12-in. (30.5-cm) tubing (Abbot Laboratories)
50°C water bath
6 to 8 kDa MWCO dialysis bag

Additional reagents and equipment for quantitation of DNA (APPENDIX 3D), pouring gradients (UNIT 10.10), agarose gel electrophoresis (UNIT 2.5A), and SDS-PAGE (UNIT 10.2A)

Purify oligonucleosome cores

1. Resuspend ~2 ml nuclear pellet with 40 ml MSB. Centrifuge 10 min at 10,000 × g, 4°C (e.g., 9200 rpm in Beckman JA-20 rotor).

Dounce homogenization (a few strokes with type B pestle) can be used to resuspend nuclei.

Chromatin
Assembly and
Analysis

21.5.3

2. Resuspend nuclei in 4 pellet vol of HSB and homogenize in a Dounce homogenizer with 40 to 50 strokes of a type B pestle to release oligonucleosome fragments.
3. Pellet nuclei 20 min at $10,000 \times g$, 4°C , and collect the supernatant. Dilute a sample of this supernatant into 2 M NaCl and measure A_{260} (APPENDIX 3D) to calculate the percent of chromosomal DNA that has been solubilized from the nuclei.

Typically ~3 to 4 mg of oligonucleosome cores can be expected at this point (~50% of the DNA in the nuclei).

4. Dialyze supernatant overnight at 4°C against 4 liters LSB with 6 to 8 kDa MWCO dialysis membrane. When collecting the dialysate, mix the sample (pressing on the sides of the dialysis bag) in order to collect those oligonucleosomes that precipitated during dialysis.
5. Add 0.1 M CaCl_2 to 3 mM (final) and warm the sample for 5 min at 37°C .
6. Add 50 U/ μl micrococcal nuclease to 10 U/ml (final) and incubate 5 min at 37°C .

The solution should become clearer, because digestion will resolubilize precipitated oligonucleosomes.
7. Stop digestion with 0.1 vol of 0.5 M EGTA and chill on ice.
8. While gently vortexing, add 2 M NaCl dropwise to a final concentration of 0.6 M.

For separation by gel filtration:

- 9a. Pour an $\sim 1.6 \times 58\text{-cm}$ Sepharose CL-6B column and equilibrate with several column volumes of HSB without sucrose at 12 to 24 ml/hr.
- 10a. Centrifuge polynucleosome fragments 30 min at $150,000 \times g$, 4°C , (e.g., 40,000 rpm in a Beckman SW55 rotor) to remove insoluble material.
- 11a. Load supernatant onto the column, run the column at 12 ml/hr, and collect 2-ml fractions.

If the load exceeds 10% of the column volume, rerun the column with additional material or use a larger column with proportional dimensions.

For separation by centrifugation:

- 9b. Pour two (or more) 34-ml linear gradients (UNIT 10.10) in $1 \times 3.5\text{-in.}$ polyallomer centrifuge tubes using HSB/glycerol, with 10% at the top to 40% at the bottom.

For this procedure, perform all steps and maintain all buffers at 4°C .

- 10b. Carefully layer 2 ml quenched digestion reaction on top of each gradient and centrifuge 16 hr at $100,000 \times g$, 4°C (e.g., 27,500 rpm in a Beckman SW28 rotor).

A 4-ml load onto a 32-ml gradient also gives good resolution with higher final concentration.

- 11b. Harvest the gradients by piercing the bottom edge (where it bends upward) with a 21-G needle attached to outlet tubing. Hold the needle at an $\sim 30^{\circ}$ angle, with the beveled edge up, to pierce the tube, and then lodge the needle tip near the very bottom of the tube (without puncturing the opposite end of the tube). To start the flow, cover the top with Parafilm and apply gentle pressure. Collect 1- to 1.5-ml fractions.

Analyze and store oligonucleosome cores

12. Determine DNA concentration of fractions by diluting an aliquot 10- to 40-fold into 2 M NaCl and measuring absorbance at 260 nm.

13. Measure DNA lengths of oligonucleosomes by treating aliquots of fractions (≥ 0.5 μg) with 0.5% SDS and 0.5 mg/ml proteinase K for 1 hr at 50°C, and separating by native 1.5% agarose TBE gel electrophoresis (UNIT 2.5A). Stain with 0.1 $\mu\text{g}/\text{ml}$ ethidium bromide.

Ethidium bromide staining should show a smear of overdigested DNA, a band of DNA at ~150 bp (mononucleosomes), and larger bands equal to ~150 bp plus multiples of ~200 bp (di-, tri-, tetranucleosomes and so on). Higher multiples will be nearer the column void or the bottom of the gradient.

CAUTION: Ethidium bromide is a mutagen and should be handled appropriately.

14. Analyze 5 to 20 μl of selected fractions for the presence of all four core histones (H2A, H2B, H3, and H4) and for the absence of H1 on a 15% SDS-PAGE gel (UNIT 10.2A). Stain the gel with Coomassie brilliant blue.

Molecular weights of core histones are 14.1 kDa (H2A), 13.8 kDa (H2B), 15.3 kDa (H3), and 11.3 kDa (H4). H1 (21.5 kDa) should separate from most of the DNA and run at the top of the gradient or at the end of the included volume on the column.

15. Pool fractions that contain mono- and dinucleosomes, short oligonucleosomes (3 to 6 units), and long oligonucleosomes (>6 units) total three pooled fractions). Be careful to avoid fractions where the DNA is <150 bp (overdigested) or that contain H1.

16. If DNA concentration is low (<0.5 mg/ml), concentrate by putting sample in a 6 to 8 kDa MWCO dialysis bag and surrounding with solid sucrose. Once the volume has decreased 2 to 4 fold, rinse off sucrose with water and reclip the dialysis bag to the new volume.

Adjusting the clip prevents samples from swelling back to the original volume in the next step.

17. If step 16 was bypassed, place samples in a 6 to 8 kDa MWCO dialysis bag and dialyze into 100 vol dialysis buffer at 4°C for >4 hr.

Alternatively, or in addition, samples can be concentrated up to 1 mg/ml using Centriprep-10 concentrators (Amicon).

18. Divide into aliquots, freeze on dry ice or in liquid nitrogen, and store for up to 2 years at -80°C .

The nucleosomes are also stable for many weeks at 4°C.

NOTE: When polynucleosome stocks are thawed, some precipitation may be observed. This will not interfere with the use of the polynucleosomes in salt-dilution assembly procedures, as long as they are well mixed before use to ensure proper concentration. Also, material precipitated during freezing appears to come back into solution when the stock is diluted. The addition of 20% (v/v) glycerol and rapid freezing in liquid nitrogen may help to reduce this effect.

PURIFICATION OF MONO- AND DINUCLEOSOMES

The previous protocol rarely yields pure mono- and dinucleosomes, either because fractions don't separate well or because H1 may contaminate these fractions. By redigesting pure oligonucleosomes, however, mono- and dinucleosomes with little cross-contamination can be recovered (Schnitzler et al., 1998).

Materials

Oligonucleosomes: medium or large polynucleosome fractions (see Basic Protocol 2)
100 mM CaCl₂
1 M MgCl₂
50 U/μl micrococcal nuclease (see recipe)
0.5 M EDTA
10% and 30% (v/v) glycerol gradient buffer (see recipe)
30°C water bath
Ultracentrifuge with rotor (e.g., Beckman SW55) and 0.5 × 2.5-in. (1.3 × 6.4-cm) polyallomer tubes
Gradient maker
Additional reagents and equipment for harvesting, analyzing, and concentrating fractions (see Basic Protocol 2) and for nondenaturing acrylamide electrophoresis (*UNIT 10.2B*)

1. Thaw ~1 mg oligonucleosomes (ideally ~1 to 2 mg/ml) and warm to 30°C.
2. Add 100 mM CaCl₂ to 1.5 mM final and 1 M MgCl₂ to 3.5 mM final.
3. Add 50 U/μl micrococcal nuclease to 0.1 U/μg polynucleosomes and allow digestion to proceed for 10 min at 30°C.
4. Stop the reaction by adding 0.5 M EDTA to 15 mM.
5. Pellet insoluble material by microcentrifuging 30 sec at ~10,000 × g, 4°C.
6. Pour a 4.7-ml linear gradient in a 0.5 × 2.5-in. polyallomer ultracentrifuge tube using 10% glycerol gradient buffer at the top and 30% at the bottom.
7. Layer 0.5 ml digestion reaction onto the gradient and centrifuge 18 hr at 100,000 × g, 4°C (e.g., 35,000 rpm in SW55 rotor).
8. Harvest the gradient as described (see Basic Protocol 2, step 11b), taking about twenty four 5-drop fractions.
9. Analyze fractions by agarose gel electrophoresis as described (see Basic Protocol 2, step 13).
10. Pool clean mono- and dinucleosome peaks and concentrate, if desired (see Basic Protocol 2, steps 16 and 17, but perform final dialysis in 10% glycerol gradient buffer; see recipe). Store at 4°C (>1 month) or freeze in liquid nitrogen or on dry ice and store up to 2 years at -80°C.

The integrity of mono- and dinucleosome cores can be checked by nondenaturing electrophoresis on either 5% acrylamide (UNIT 10.2B) or 1.2% agarose gels run at 4°C in 0.5-fold TBE electrophoresis buffer. Intact mononucleosomes will run much slower than the DNA they contain (~700 bp rather than 150 bp for samples treated with SDS and proteinase K).

PURIFICATION OF CORE HISTONES BY HYDROXYLAPATITE CHROMATOGRAPHY

BASIC PROTOCOL 4

A second common method for assembling mononucleosomes or polynucleosomes from defined-sequence, labeled DNA relies on the direct assembly of nucleosomes from free histones and DNA as the salt concentration is decreased from 1 or 2 M. For this and other applications, the four core histones can be readily purified by binding chromatin fragments to hydroxylapatite resin (to which DNA binds very strongly), washing away histone H1 with 0.6 M NaCl, and eluting the histones from the resin-bound DNA at 2.5 M NaCl (Côté et al., 1995; Workman et al., 1991).

Materials

Nuclear pellet (see Basic Protocol 1)

HAP buffer (see recipe), with and without 2.5 M NaCl

BioGel HTP powder (Bio-Rad), with adsorption capacity 0.6 mg DNA per g dry powder

Bio-Rad Protein Assay (optional)

2 × 15-cm column and accessories

Centriprep-10 concentrators (Amicon; optional)

Additional reagents and equipment to determining protein concentration (UNIT 10.1A) and SDS-PAGE (UNIT 10.2A)

1. Resuspend ~2 ml nuclear pellet containing ~6 mg DNA in 25 ml HAP buffer and stir gently for 10 min at 4°C.

A few strokes with a dounce homogenizer and a type B pestle can be used to aid resuspension. To avoid proteolysis, 1 μM pepstatin A and 1 μM leupeptin can be added to HAP buffer.

2. While stirring, add 10 g dry BioGel HTP powder.
3. Add just enough HAP buffer to make a slurry. Pour this into a 2 × 15-cm column and collect the eluant.
4. Wash resin with 10 vol HAP buffer (~300 ml at 30 ml/hr).

The eluant from step 3 and the first 2/3 column volume of wash contain H1, which can be further purified as described (Côté et al., 1995) in order to reconstitute chromatosomes.

5. Elute the core histones with HAP buffer containing 2.5 M NaCl, collecting 8-ml fractions.
6. Determine protein concentration of fractions by A_{230} or A_{280} readings or by protein assay, and pool the peak fractions.

Because histones are small positively charged proteins, many methods for determining protein concentration give differing results. Since nucleosome assembly protocols require careful determination of the molar ratio of histones to DNA, it is essential to use the same method for determining histone concentration as used in the assembly protocol. One reasonably good method is to measure absorbance at 230 nm. The extinction coefficient (calculated for chicken erythrocyte histones, but likely to be similar for mammalian histones) is 4.3 A_{230} units/mg histones/ml. Alternatively, a Bio-Rad or other protein assay (UNIT 10.1A) can be performed with BSA as a standard. The Bio-Rad assay tends to yield concentration estimates that are ~1.6-fold higher than the A_{230} method.

7. Measure the salt concentration of pooled fractions by conductivity readings.

The salt concentration should be ~2 M.

Chromatin Assembly and Analysis

21.5.7

8. Measure the purity of the core histones by SDS-PAGE as described (see Basic Protocol 2, step 14).
9. If necessary, concentrate the core histones to between 2 and 10 mg/ml using Centriprep-10 concentrators.
10. Divide into aliquots, freeze on dry ice, and store up to 4 years at -80°C .

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Buffer B

20 mM HEPES, pH 7.5
3 mM MgCl_2
0.2 mM EGTA
Store up to several weeks at 4°C
Immediately before use add:
3 mM 2-mercaptoethanol
0.4 mM PMSF (see recipe)
1 μM pepstatin A (see recipe)
1 μM leupeptin (see recipe)

Dialysis buffer

20 mM HEPES, pH 7.5
1 mM EDTA
Store up to several weeks at 4°C
Immediately before use add:
1 mM 2-mercaptoethanol
0.5 mM PMSF (see recipe)

Glycerol gradient buffer, 10% and 30% (v/v)

20 mM HEPES, pH 7.9
1 mM EDTA
30 mM KCl
0.1% Nonidet P-40 (NP-40)
10% or 30% (v/v) glycerol
Store up to several weeks at 4°C

This buffer can be adjusted as desired. Virtually any solution with >0.2 mM EDTA, no divalent cations, and salt from 0 to 600 mM will work well as the basis for a glycerol gradient buffer.

High-salt buffer (HSB)

20 mM HEPES, pH 7.5
0.65 M NaCl
1 mM EDTA
0.34 M sucrose
Store up to several weeks at 4°C
Immediately before use add:
1 mM 2-mercaptoethanol
0.5 mM PMSF (see recipe)

Hydroxylapatite (HAP) buffer

50 mM sodium phosphate, pH 6.8
0.6 M NaCl
Store up to several weeks at 4°C
Immediately before use add:
1 mM 2-mercaptoethanol
0.5 mM PMSF (see recipe)

Leupeptin

Prepare a 1 mM aqueous stock solution and store up to 1 year at -20°C.

Low-salt buffer (LSB)

20 mM HEPES, pH 7.5
0.1 M NaCl
1 mM EDTA
Store up to several weeks at 4°C
Immediately before use add:
1 mM 2-mercaptoethanol
0.5 mM PMSF (see recipe)

Lysis buffer

20 mM HEPES, pH 7.5
0.25 M sucrose
3 mM MgCl₂
0.5% (v/v) Nonidet P-40 (NP-40)
Store up to several weeks at 4°C
Immediately before use add:
3 mM 2-mercaptoethanol
0.4 mM PMSF (see recipe)
1 μM pepstatin A (see recipe)
1 μM leupeptin (see recipe)

Medium-salt buffer (MSB)

20 mM HEPES, pH 7.5
0.4 M NaCl
1 mM EDTA
5% (v/v) glycerol
Store up to several weeks at 4°C
Immediately before use add:
1 mM 2-mercaptoethanol
0.5 mM PMSF (see recipe)
1 μM pepstatin A (see recipe)
1 μM leupeptin (see recipe)

Micrococcal nuclease, 50 U/μl

50 U/μl micrococcal nuclease (from solid; Sigma)
50% (v/v) glycerol
50 mM Tris·Cl, pH 8 (APPENDIX 2)
0.05 mM CaCl₂
Store at -20°C (can be stable for >1 year)

Note that different suppliers define units of micrococcal nuclease differently. If Sigma enzyme is not used, be sure that the unit definition is the same, or determine the proper conversion ratio to Sigma units.

Pepstatin A

Prepare a 1 mM stock solution in methanol and store up to 2 years at -20°C .

Phenylmethylsulfonyl fluoride (PMSF)

Prepare a 100 mM stock solution in isopropanol and store up to 4 years at room temperature.

COMMENTARY

Background Information

Use of polynucleosomes and histones in chromatin assembly

A principle use of oligonucleosome cores and core histones is the *in vitro* assembly of nucleosome cores on defined-sequence DNAs. Several templates exist that allow the formation of radiolabeled mononucleosomes. Many of these contain “rotational phasing sequences,” which are essentially DNA sequences that cause the DNA to bend, promoting wrapping around the histone core with one face outwards (a single rotational phase). Frequently used mononucleosome templates are those based on 5S rDNA sequences and those with artificial phasing sequences (e.g., Kwon et al., 1994; Hayes and Lee, 1997; Luger et al., 1997; Schnitzler et al., 1998). Mononucleosomes can be formed by incubating the template DNA with excess polynucleosomes (octamer transfer method; Côté et al., 1995) or with roughly equimolar free histones at 1 or 2 M NaCl (salt assembly; Imbalzano et al., 1996). At this high salt concentration the H2A/H2B dimers dissociate and the H3/H4 tetramers readily transfer from donor polynucleosomes, or begin to assemble *de novo* onto the template DNA. As the salt concentration is reduced, either by dilution or dialysis, the H2A/H2B dimers are added and the complete nucleosome is formed. Polynucleosomes can be formed by salt assembly on any linear DNA template, although the resulting nucleosomes assume somewhat random positions (Stein, 1989). Templates that are made up of a repeated array of 5S rDNA sequences, however, result in arrays of reasonably well-positioned nucleosomes (Carruthers et al., 1999).

Although salt assembly methods are the simplest assembly systems, containing only DNA and the core histones, other systems for chromatin assembly exist. Heat-treated extracts of *Xenopus* oocytes allow assembly of poorly positioned nucleosomes at lower salt concentrations (Workman et al., 1991). Many cellular factors that facilitate such random deposition

are known and can be purified from a variety of sources (Ito et al., 1996, and references therein). Unheated *Xenopus* oocyte and *Drosophila* egg extracts contain ATP-dependent factors that result in uniform spacing of nucleosomes assembled onto templates (Sessa and Ruberti, 1990; Becker et al., 1994). The factors responsible for some of these activities have been identified, and defined *in vitro* reconstitution systems for spaced nucleosomes using pure proteins are now possible (Ito et al., 1997; Varga-Weisz et al., 1997).

Variations for histone and polynucleosome isolation

Polynucleosomes that contain linker histones can be isolated by a protocol that uses micrococcal nuclease to release chromatin fragments from nuclei (Kornberg et al., 1989). Partially purified linker histone H1 is a by-product of the chromatography step in Basic Protocol 4. This can be further purified and used to reconstitute chromatosomes (Côté et al., 1995). Treatment of cultured cells with the histone deacetylase inhibitor butyrate results in hyperacetylation of histone tails (Vettese-Dadey et al., 1995). The acetylation state of the histones can be assayed by Triton/acid/urea gels (UNIT 21.2). Nuclei from these cells can be subjected to the above protocols to yield hyperacetylated oligonucleosome cores or histones. A stepwise elution of histone from hydroxylapatite allows H2A and H2B to be separated from H3 and H4 (Simon and Felsenfeld, 1979). The N- and C-terminal tails of the histones in oligonucleosome cores can be cleaved off by treatment with trypsin. These proteolyzed arrays can be subjected to hydroxylapatite chromatography to yield tailless histones, and both the arrays and histones can be used in nucleosome assembly reactions (Guyon et al., 1999).

Critical Parameters

As with any protein isolation, everything should be kept on ice or at 4°C except where noted. Small amounts of endogenous nucleases

and residual micrococcal nuclease may be present in poly- and mononucleosome samples. To avoid slow hydrolysis during storage, samples should always be stored in buffers containing EDTA, which chelates the divalent cations required for endonuclease action.

Reducing agents (2-mercaptoethanol and DTT) and protease inhibitors (PMSF, pepstatin A, and leupeptin) used in these experiments should all be added to the buffers just before use. Other buffer components can be stored at 4°C for several weeks.

Troubleshooting

It is important to check the integrity and purity of histones in oligonucleosome and histone preparations. Each of the four core histones should stain roughly evenly using Coomassie stain. If one stains much more weakly (often H2A), protease inhibitors should be added at steps where it is an option.

Anticipated Results

The nuclear pellet from 3 liters of HeLa cells should yield 4 to 10 mg of DNA. Of this, ~1 to 2 mg of medium polynucleosomes (Basic Protocol 2) or ~3 to 8 mg of core histones (Basic Protocol 4) should be obtained. The yield of mononucleosomes (Basic Protocol 3) is generally ~20% to 30% of the mass of input polynucleosomes.

Time Considerations

Basic Protocol 1 should take ~3 hours (less if starting with spent nuclear pellet). Basic Protocol 2 will take ≥2 days to complete. The dialysis step can be reduced to ~2 to 3 hr by dialyzing into LSB with only 50 mM salt, taking samples from within the bag and monitoring salt concentration with a conductivity meter. Oligonucleosomes can be collected when the salt concentration is 100 mM. The column washes in the gel filtration method are often conveniently done overnight (washing more slowly or with a greater volume is acceptable). Depending on the experimental needs, the lengthy concentration steps may not be necessary. Be aware that nucleosomes below ~10 ng/μl may not be stable without 0.1% NP-40 or >50 μg/ml BSA as carriers.

Basic Protocol 3 takes a few hours to set up, plus an overnight centrifugation and an hour to collect fractions. Basic Protocol 4 calls for ~10 hr wash time at 0.6 M NaCl to remove all traces of H1. The flow rate during this wash can be increased (up to 60 ml/hr) without creating excessive H1 contamination.

Literature Cited

- Becker, P.B., Tsukiyama, T., and Wu, C. 1994. Chromatin assembly extracts from *Drosophila* embryos. *Methods Cell Biol.* 44:207-223.
- Carruthers, L.M., Tse, C., Walker, K.P. III, and Hansen, J.C. 1999. Assembly of defined nucleosomal and chromatin arrays from pure components. *Methods Enzymol.* 304:19-35.
- Côté, J., Utley, R.T., and Workman, J.L. 1995. Basic analysis of transcription factor binding to nucleosomes. *Methods Mol. Genet.* 6:108-128.
- Guyon, J.R., Narlikar, G.J., Sif, S., and Kingston, R.E. 1999. Stable remodeling of tailless nucleosomes by the human SWI-SNF complex. *Mol. Cell Biol.* 19:2088-2097.
- Hayes, J.J., and Lee, K.-M. 1997. In vitro reconstitution and analysis of mononucleosomes containing defined DNAs and proteins. *Methods* 12:2-9.
- Imbalzano, A.N., Schnitzler, G.R., and Kingston, R.E. 1996. Nucleosome disruption by human SWI/SNF is maintained in the absence of continued ATP hydrolysis. *J. Biol. Chem.* 271:20726-20733.
- Ito, T., Tyler, J.K., Bulger, M., Kobayashi, R., and Kadonaga, J.T. 1996. ATP-facilitated chromatin assembly with a nucleoplasmin-like protein from *Drosophila melanogaster*. *J. Biol. Chem.* 271:25041-25048.
- Ito, T., Bulger, M., Pazin, M.J., Kobayashi, R., and Kadonaga, J.T. 1997. ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell* 90:145-155.
- Kornberg, R.D., LaPointe, J.W., and Lorch, Y. 1989. Preparation of nucleosomes and chromatin. *Methods Enzymol.* 170:3-14.
- Kwon, H., Imbalzano, A.N., Khavari, P.A., Kingston, R.E., and Green, M.R. 1994. Nucleosome disruption and enhancement of activator binding by a human SWI/SNF complex. *Nature* 370:477-481.
- Luger, K., Rechsteiner, T.J., Flaus, A.J., Wayne, M., and Richmond, T.J. 1997. Characterization of nucleosome core particles containing histone proteins made in bacteria. *J. Mol. Biol.* 272:301-311.
- Schnitzler, G., Sif, S., and Kingston, R.E. 1998. Human SWI/SNF interconverts a nucleosome between its base state and a stable remodeled state. *Cell* 94:17-27.
- Sessa, G. and Ruberti, I. 1990. Assembly of correctly spaced chromatin in a nuclear extract from *Xenopus laevis* oocytes. *Nucl. Acids Res.* 18:5449-5455.
- Simon, R.H. and Felsenfeld, G. 1979. A new procedure for purifying histone pairs H2A+H2B and H3+H4 from chromatin using hydroxylapatite. *Nucl. Acids Res.* 6:689-696.
- Stein, A. 1989. Reconstitution of chromatin from purified components. *Methods Enzymol.* 170:585-603.

- Varga-Weisz, P.D., Wilm, M., Bonte, E., Dumas, K., Mann, M., and Becker, P.B. 1997. Chromatin-re-modelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature* 388:598-602.
- Vettese-Dadey, M., Adams, C.C., Côté, J., Walter, P., and Workman, J.L. 1995. Experimental analysis of transcription factor-nucleosome interactions. *Methods Mol. Genet.* 6:129-151.
- Workman, J.L., Taylor, I.C.A., Kingston, R.E., and Roeder, R.G. 1991. Control of class II gene transcription during in vitro nucleosome assembly. *Methods Cell Biol.* 35:419-447.

Key References

Côté, et al., 1995. See above.

Basic Protocols 1, 2, and 4 are modifications of protocols in this paper.

Wasserman, P.M. and Wolffe, A.P. (eds.) 1999. Chromatin. *Methods Enzymol.* Vol. 304.

This volume contains many other useful and current protocols for the assembly and analysis of chromatin.

Contributed by Gavin R. Schnitzler
Tufts University School of Medicine
Boston, Massachusetts

Assembly of Nucleosomal Templates by Salt Dialysis

UNIT 21.6

The assembly of nucleosomal templates through salt dialysis has been a useful tool to generate, conveniently and in a relatively short period of time, large quantities of homogeneous model chromatin complexes containing defined DNA sequences. The first protocol of this unit (see Basic Protocol 1) describes the assembly of mononucleosomes and nucleosomal arrays through step salt dialysis. The assembly through step salt dialysis allows rapid reconstitution of mononucleosomes and nucleosomal arrays by lowering the salt concentration of the original reconstitute by transferring it through a series of buffers. Gradient dialysis is used to reconstitute large amounts of mononucleosomes or nucleosomal arrays. The second protocol (see Basic Protocol 2) describes assembly of mononucleosomes through gradient salt dialysis; an additional basic protocol (see Basic Protocol 3) describes use of gradient dialysis to reconstitute nucleosomal arrays. Gradient dialysis is useful when step dialysis is not an option. In addition to the basic protocols, the support protocols describe methods to isolate the components of the assembly: the histones and the DNA fragment. The purification of histones from native chromatin sources is described in *UNIT 21.5* of this section. Purification of recombinant histones is described in Support Protocol 1. The protocol for preparation of the DNA template for assembly is provided in Support Protocol 2. Support Protocols 5 to 7 describe techniques to analyze the assembled nucleosomal template using gel mobility shift assays.

CAUTION: Radioactive, biological, and chemical substances require special handling; see *APPENDIX 1E* for guidelines.

STRATEGIC PLANNING

A critical aspect of nucleosome assembly is the integrity of the reconstituted nucleosomal template. Since the primary goal in nucleosome assembly by salt dialysis is to produce structurally homogeneous samples, assembly of an optimal nucleosomal template is dependent on the stoichiometry of the histones and DNA and also the positioning of the histone octamer on the DNA template. The total histone mass and DNA mass should be approximately equal to each other to generate a homogeneous population of nucleosomal template. If the histones are not sufficient, inefficient octamer assembly is the result, and, conversely, nonspecific binding and aggregation occur when the amount of DNA is not enough. The ratio between histones and DNA must be empirically adjusted by varying the ratio to yield maximum octamer-DNA complexes. The positioning of the histone octamer is also important because many experiments, such as nuclease mapping, require that the octamer adopt a single position relative to the DNA fragment. This can be accomplished by using DNA templates that carry a strong positioning sequence such as fragments that contain the *Xenopus borealis* 5S rRNA gene which has been shown to strongly direct unique positioning (see Commentary).

ASSEMBLY OF NUCLEOSOMAL TEMPLATES BY STEP SALT DIALYSIS

The method described here for the reconstitution of nucleosomes allows for large quantities of nearly homogeneous core particles in 10 hr (Hayes and Lee, 1997). Virtually any piece of DNA 147 bp or longer can be used. However to obtain nucleosomes with only one translational position, the DNA sequence should contain nucleosome positioning sequences such as that from the *Xenopus borealis* 5S rRNA gene as previously mentioned (Hayes et al., 1990).

**BASIC
PROTOCOL 1**

**Chromatin
Assembly and
Analysis**

21.6.1

Nucleosomal arrays can also be assembled by step dialysis as described in this protocol. The difference is that the relative mass ratio between histone:DNA is 1.2, which is slightly higher than for mononucleosomes. Optimal conditions should be determined empirically. The authors find that gradient dialysis (see Basic Protocol 3) gives better assembled arrays than step dialysis.

Materials

- ~1 to 2 mg/ml sonicated calf thymus DNA (~0.5 to 1 kb; Sigma)
- Radioactively labeled DNA (see Support Protocol 2)
- Purified native core histone protein fractions (Support Protocol 1)
- 5 M NaCl (APPENDIX 2)
- TE buffer, pH 8.0 (APPENDIX 2)
- TE buffer, pH 8.0 (APPENDIX 2), containing 1.2, 1.0, 0.8, and 0.6 M NaCl
- 6000 to 8000 MWCO dialysis tubing (Spectrapor; boil in Milli-Q water for 5 min and store at 4°C)
- Additional reagents and equipment for dialysis (APPENDIX 3C) and purification of nucleosomes by glycerol gradient (UNIT 21.5)

1. Prepare the following reconstitution mixture:

- ~8 µg unlabeled sonicated calf thymus DNA (as carrier DNA; add from ~1 to 2 mg/ml stock)
- 200,000 to 400,000 cpm singly labeled DNA
- 6.4 µg (for 0.8:1 histone:DNA ratio) or 0.8 µg (for 1:1 histone:DNA ratio) purified native core histone protein fractions (H2A/H2B and H3/H4)
- 160 µl 5 M NaCl (2.0 M final)
- TE buffer, pH 8.0, to 400 µl.

The unlabeled competitor DNA is necessary to accurately determine nucleic acid concentration and thereby establish the correct histone:DNA stoichiometry. It is possible to disregard the DNA amount contributed by the labeled DNA if the final unlabeled DNA concentration does not go below 50 ng/µl. The ratio of histone:DNA mass must be empirically adjusted to obtain a maximum of mononucleosomes without detectable dinucleosomes. The usual starting ratio for mononucleosome assembly is 0.8 or 1.0 parts histone to one part DNA by mass.

2. Incubate at room temperature for 30 min.
3. Place the reconstitution mixture into a 6000 to 8000 MWCO dialysis bag (APPENDIX 3C).
4. Dialyze the reconstitution mixture in 1 liter of TE buffer, pH 8.0, containing 1.2 M NaCl, for 2 hr at 4°C (APPENDIX 3C).
5. Dialyze the reconstitution mixture successively, each time for 2 hr at 4°C, against the following freshly prepared dialysis buffers:
 - TE buffer, pH 8.0, containing 1.0 M NaCl
 - TE buffer, pH 8.0, containing 0.8 M NaCl
 - TE buffer, pH 8.0, containing 0.6 M NaCl.
6. Dialyze the reconstitution mixture overnight at 4°C in TE buffer, pH 8.0, without added NaCl.

Nucleosomes at this stage can be used for gel shift experiments where EDTA does not interfere.

7. Analyze the integrity of the nucleosomes by electrophoresis (see Support Protocols 5 and 6).
8. Purify assembled nucleosome cores from naked DNA over a 10% to 30% glycerol gradient (Basic Protocol 3 in *UNIT 21.5*) run for 18.5 hr at $100,000 \times g$.

ASSEMBLY OF HIGH CONCENTRATION OF MONONUCLEOSOMES BY GRADIENT SALT DIALYSIS

**BASIC
PROTOCOL 2**

To obtain high concentrations (~ 0.5 mg/ml) of core particles, gradient dialysis is recommended over step dialysis, as high concentrations of histones and DNA form nonspecific aggregates unless the dialysis is very gradual (Luger et al., 1997a, 1999). The buffer volumes and pump flow rates mentioned below are for a 200 to 1000 μ l reconstitution mixture.

Materials

Unlabeled DNA fragment at high concentration (>2 mg/ml; Support Protocol 3)
 1 M Tris·Cl, pH 7.7 (*APPENDIX 2*)
 1 M dithiothreitol (DTT)
 0.5 M disodium EDTA, pH 8.0 (*APPENDIX 2*)
 2.5 M KCl
 0.5 M benzamidine
 Core histones (purified from HeLa nuclear pellets; *UNIT 21.5*)
 High-salt buffer (see recipe)
 Low-salt Buffer (see recipe)
 Zero-salt buffer (see recipe)
 10% and 30% glycerol gradient buffer (see recipe in *UNIT 21.5*)

Rabbit pump (Rainin) with tubes
 6000 to 8000 MWCO dialysis tubing (Spectrapor, 6.4 mm diameter; boil in Milli-Q water for 5 min and store at 4°C)
 Conductivity meter calibrated for KCl concentration

Additional reagents and equipment for analysis of reconstituted complexes by agarose gel electrophoresis (see Support Protocol 5) or polyacrylamide gel electrophoresis (see Support Protocol 6)

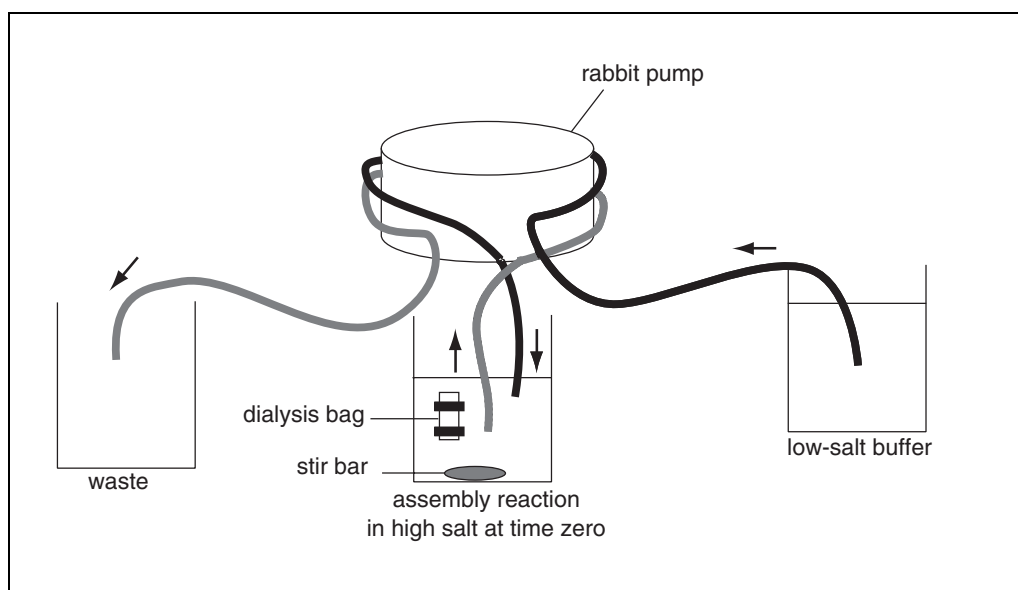


Figure 21.6.1 Setup for gradient salt dialysis (adapted from Luger et al., 1999).

**Chromatin
Assembly and
Analysis**

21.6.3

1. Prepare the following reconstitution mix (adding the histones last):

0.7 mg/ml DNA fragment
20 mM Tris-Cl, pH 7.7 (from 1 M stock)
10 mM DTT (from 1 M stock)
1 mM EDTA, pH 8.0 (from 0.5 M stock)
2 M KCl (from 2.5 M stock)
0.5 mM benzamidine (from 0.5 M stock)
0.63 mg/ml histones.

Put assembly reaction at 4°C for 30 min after mixing components.

The histones can be prediluted in 2 M NaCl to get manageable concentrations. Reconstitutions can be carried out in volumes from 200 μ l up to many milliliters.

2. Set up “rabbit” pump assembly and set flow rate at 205 to 210 μ l/min (Figure 23.6.1).
3. Transfer reaction mixture to 6000 to 8000 kDa dialysis tubing.
4. Put membrane in middle beaker containing high-salt buffer, turn on the pump and keep at 4°C for ~ 50 hr.

The published procedure (Luger et al., 1997a, 1999) is to incubate for 36 hrs so the flow rate from the pump can be adjusted a bit to reduce time. However dialyzing faster than over 36 hrs can give precipitation of histones with DNA, so it is safer to keep the dialysis rate slow.

5. Check conductivity of solution in middle beaker every 5 hr or so.

By 50 hr, should drop from 2 M to ~300 mM.

6. Dialyze into zero-salt buffer (\geq 3 hr)

Nucleosomes can also be stored in the 300 mM salt buffer if the gradient buffer used in step 8 contains KCl.

7. Check assembly on a 0.7% agarose gel (see Support Protocol 5) or a 5% polyacrylamide 0.5 \times TBE gel (see Support Protocol 6).
8. Purify nucleosomes from free DNA over a 10% to 30% glycerol gradient (UNIT 21.5) run for ~18.5 hr at 100,000 \times g.

The fractions can be stored at 4°C for up to 2 months.

BASIC PROTOCOL 3

ASSEMBLY OF NUCLEOSOMAL ARRAYS BY GRADIENT SALT DIALYSIS

The procedure has been optimized with a specific DNA fragment, (G5E4 DNA) but can be used to assemble any 5S array. The DNA used in this assembly (G5E4 DNA) is a 2546 bp fragment that has five 5S rDNA positioning sequences on either side of a ~400 bp sequence containing 5 Gal4 binding sites, the E4 promoter, and several unique restriction sites (Neely et al., 1999). The central 400 bp can accommodate two nucleosomes.

Materials

Unlabeled G5E4 DNA (Dr. Jerry Workman; jlw10@psu.edu)
End-labeled G5E4 DNA (see Support Protocol 2 for labeling)
Bovine serum albumin (BSA)
High-salt buffer (see recipe) with 2 M NaCl in place of KCl
Low-salt buffer (see recipe) with 250 mM NaCl in place of KCl

Additional reagents and equipment for assembly of mononucleosomes by gradient salt dialysis (see Basic Protocol 2) and *EcoRI* digestion to determine extent of array assembly (see Support Protocol 7)

1. Prepare the following reconstitution mix (adding the histones last):

~100 µg/ml of a mixture of 9 parts unlabeled to 1 part end-labeled (molar ratio) G5E4 DNA
20 mM Tris-Cl, pH 7.7
10 mM DTT
1 mM EDTA
2 M NaCl
0.5 mM benzamidine
~100 µg/ml histones
100 µg/ml BSA

Put assembly reaction at 4°C for 30 min after mixing components.

2. Dialyze the reconstitution mix through salt gradient (in this case, NaCl) as for mononucleosomes (see Basic Protocol, steps 2 to 6).

For arrays, final dialysis into zero-salt buffer is preferable. The high-salt and low-salt buffers are the same as above except the KCl is substituted with the same amount of NaCl.

3. Analyze the assembly by *EcoRI* digestion (see Support Protocol 7).

PURIFICATION OF RECOMBINANT CORE HISTONES FROM BACTERIA

Core histones can be purified from a variety of sources but histones obtained from native chromatin provide a simple and abundant source for many different experiments (*UNIT 21.5*). Alternately, recombinant histones that have been overexpressed in bacteria can be isolated. This has the advantage of generating mutant histones which can be used for experiments that require modified histones. e.g., histones with deleted N-terminal domains (Luger et al., 1997a).

Although the ability to obtain recombinant histones from bacteria allows manipulations of these histones for experiments that require its modified forms, a concern regarding recombinant histones is that they are often obtained individually (Luger et al., 1997a). Thus, in addition to total histone mass, the ratio of the core histones must also be adjusted empirically. This can be done by combining purified histone H3 and H4 and purifying them in one mixture and H2A and H2B in another mixture. It has been shown that H3 and H4 can form a H3/H4 complex when combined together, and the same is true for H2A and H2B (Lee et al., 1999). It is also known that these complexes can be purified stoichiometrically on an ion-exchange column (Lee et al., 1999). Once these complexes are isolated, the assembly can be done by forming tetramers of histones H3 and H4, (H3/H4)₂, onto DNA by omitting histones H2A and H2B from the reconstitution (Hayes and Lee, 1997). After identification of the tetramer complexes, several reconstitutions containing increasing amounts of H2A/H2B can be prepared, essentially titrating these proteins into the reconstitution mix until the core histone octamer-DNA complex is completely formed. The advantage of this method is that it is possible to assemble modified forms of the H2A/H2B dimer and H3/H4 tetramer complex.

SUPPORT PROTOCOL 1

Chromatin Assembly and Analysis

21.6.5

Materials

Bacteria overexpressing core histones (BL21; Novagen; Luger et al., 1997a)
Luria broth (LB), sterile
100× (0.2 M) isopropyl D-thiogalactopyranoside (IPTG) stock
TE buffer, pH 8.0 (*APPENDIX 2*) containing 1 mM PMSF (add from saturated PMSF stock in 95% ethanol)
10 mg/ml lysozyme solution
Triton X-100 detergent
Denaturing solution (see recipe)
5 M NaCl (*APPENDIX 2*)
50% (v/v) slurry of Bio-Rex 50-100 mesh chromatography resin (Bio-Rad)
TE buffer, pH 8.0, containing 0.6, 1.0, and 2.0 M NaCl and 1 mM PMSF (add from saturated PMSF stock in 95% ethanol)

Oak Ridge centrifuge tubes
Sonicator
Beckman centrifuge with JA-20 rotor (or equivalent)
6000 to 8000 MWCO dialysis tubing (Spectrapor; boil in Milli-Q water for 5 min and store at 4°C)
10-ml disposable plastic chromatography column (Bio-Rad)
15- and 50-ml conical centrifuge tubes

Additional reagents and equipment for dialysis (*APPENDIX 3C*)

Grow the bacterial culture

1. Grow the culture in Luria broth in the absence of IPTG at 37°C to an optical density of 0.6 at 595 nm wavelength light.

The overexpression of each histone in bacteria is done individually; these are later mixed together.

2. Add IPTG to a final concentration of 0.2 mg/ml and return the culture to 37°C for approximately 2 to 4 hr.
3. After 2 to 4 hr pellet the bacteria by centrifuging 15 min at 4000 × g, 4°C.
4. Decant the supernatant and resuspended the pellet in 5 to 10 ml of TE buffer, pH 8.0.

Lyse the cells and obtain histones

5. Add 10 mg/ml lysozyme to a final concentration of 0.2 mg/ml.
6. Add Triton X-100 to a final concentration of 0.2% (v/v) and incubate for 30 min at room temperature.
7. Transfer the bacteria to Oak Ridge centrifuge tubes on ice.
8. Sonicate the bacterial slurry for 2 min total (two 1-min sonications).
9. Pellet the cell debris by centrifuging 30 min at 10,000 × g, 4°C. Discard supernatant.

The pellet contains the insoluble overexpressed histone in inclusion bodies.

10. Repeat steps 8 and 9 two additional times (total of three sonications/centrifugations).

Solubilize the histone pellet

11. Dissolve pellet in 2 ml of denaturing solution for 30 min at room temperature.

The denaturing step is required to solubilize the pellet, which mainly consists of overexpressed histones in inclusion bodies.

12. Renature the histones by dialyzing away the denaturant against a total of two to three changes (1 liter each) of TE buffer, pH 8.0, at 4°C, over a period of 48 hr (see *APPENDIX 3C* for dialysis procedures), using a 6000 to 8000 MWCO dialysis membrane.

During dialysis, it is likely that there will be a significant amount of precipitation. If precipitation seems excessive (i.e., small amount of protein in supernatant), it will be necessary to purify the denatured histones on a denaturing size exclusion column and purify the histones from chromosomal DNA which is causing the precipitation (Luger et al., 1997a).

13. Spin down any precipitate for 30 min at 8000 × g (10,000 rpm in a JA-20 rotor), 4°C.
14. Collect supernatant and raise NaCl concentration to 0.5 M by adding a sufficient quantity of 5 M NaCl.

Core histone proteins must first be incubated with their partner proteins before they will bind to the chromatography matrix (i.e., H2A with H2B).

Purify histone complexes

15. Incubate 5 ml of supernatant with 1 ml of a 50% suspension of Bio-Rex 50-100 mesh beads for 2 hr at 4°C, with mixing on an end-over-end rotator.

The ratio between supernatant to beads can be raised proportionately if there is more than 5 ml of supernatant.

16. After 4 hr, collect the beads in a plastic 10-ml disposable chromatography column. Collect the flowthrough fraction in a 50-ml conical tube. Freeze at –20°C.
17. Wash the column with 2 to 3 column volumes of TE buffer, pH 8.0, containing 0.6 M NaCl. Collect each wash fraction separately in a 15-ml conical tube. Freeze at –20°C.
18. Elute the bound proteins with two separate single-column-volume elutions of TE buffer, pH 8.0, containing 2.0 M NaCl, into separate 15-ml conical tubes. Freeze at –20°C.
19. Check 10 µl of each fraction for protein content by SDS-PAGE (see Support Protocol 6).

PREPARATION OF SINGLE 5' END-LABELED DNA FOR MONONUCLEOSOME ASSEMBLY

SUPPORT PROTOCOL 2

An important consideration in nucleosome assembly is how the histone octamer positions itself relative to the DNA. In order to generate homogeneous samples with unique positioning, the use of DNA that has a well-defined nucleosome positioning sequence is required. It has been known that DNA fragments which contain 5S RNA gene sequences have been found to strongly direct translational positioning (Rhodes, 1985). DNA fragments that contain these sequences are ideal templates for assembly.

Materials

- Plasmid DNA containing sequence of interest, with two convenient restriction sites that generate a ~150-bp fragment
- Restriction endonuclease suitable for linearizing the plasmid, and corresponding endonuclease buffer
- 3 M sodium acetate (*APPENDIX 2*)
- 95% and 70% ethanol, –20°C
- Alkaline phosphatase (Boehringer Mannheim) and alkaline phosphatase buffer
- 10% (w/v) SDS

Chromatin Assembly and Analysis

21.6.7

TE buffer, pH 8.0 (*APPENDIX 2*)
10,000 U/ml T4 polynucleotide kinase (Promega) and 10× T4 polynucleotide kinase buffer
10 μCi/μl [γ -³²P]dATP (6000 Ci/mmol; NEN Life Science Products)
2.5 M ammonium acetate
Restriction endonuclease appropriate for liberating the end-labeled fragment of interest, and corresponding endonuclease buffer
6% nondenaturing polyacrylamide gel
Pestle homogenizer fitting microcentrifuge tubes (VWR)
SpinX microcentrifuge tubes (Costar)
Additional reagents and equipment for phenol extraction and ethanol precipitation of DNA (*UNIT 2.1A*), polyacrylamide gel electrophoresis of DNA (*UNIT 2.7*), and autoradiography (*APPENDIX 3A*)

Linearize and dephosphorylate the DNA

1. Treat approximately 5 μg of plasmid DNA or ~0.5 μg of a purified DNA fragment with the appropriate restriction endonuclease using the buffer supplied by the manufacturer.
2. Precipitate the DNA by adding 0.1 vol 3 M sodium acetate and 2.5 vol ice-cold 95% ethanol.
3. Microcentrifuge 15 min at maximum speed. Discard supernatant and resuspend the DNA in 20 μl of 1× phosphatase buffer. Treat with 1 U/μl alkaline phosphatase for 1 hr at 37°C.
4. Adjust the solution to 0.1% SDS (by adding 0.01 vol 10% SDS stock), then phenol extract the solution and precipitate the aqueous phase twice with ethanol and sodium acetate (*UNIT 2.1A*).

Label the 5' ends

5. Resuspend the DNA in 10 μl TE buffer, pH 8.0, and add 2.5 μl of 10× T4 polynucleotide kinase buffer.
6. Add 50 μCi of [γ -³²P]dATP and adjust volume to 24 μl with water.
7. Start the reaction by adding 10 U of T4 polynucleotide kinase and incubate for 30 min at 37°C.
8. Stop the kinase reaction by adding 200 μl of 2.5 M ammonium acetate and 700 μl of cold 95% ethanol.
9. Microcentrifuge 30 min at maximum speed, 4°C.
10. Wash the DNA pellet briefly with cold 70% ethanol and dry in a Speedvac evaporator.
11. Dissolve the DNA in 34 μl of TE buffer, pH 8.0.

Obtain specific end-labeled fragment

12. Digest the DNA fragment with a second restriction endonuclease that liberates the fragment of interest and yields fragments that can be easily separated on a gel.
13. Apply the sample to a 6% nondenaturing polyacrylamide gel (*UNIT 2.7*). Perform electrophoretic separation.

The use of chemiluminescent markers allows the alignment of the gel in the following step.

14. After separation, wrap the gel in Saran Wrap and expose to film for 1 min (autoradiography; *APPENDIX 3A*), which is sufficient to detect the specific band containing the labeled fragment.

- Excise the band of interest from the polyacrylamide gel and place into a clean microcentrifuge tube. Crush the acrylamide gel slice with a pestle homogenizer fitting the microcentrifuge tube. Add 700 μ l of TE buffer, pH 8.0.

The labeled fragment will elute overnight with passive diffusion.

- Split the sample equally into two SpinX microcentrifuge filtration devices. Microcentrifuge 30 min at maximum speed.
- Precipitate the DNA and dissolve in TE buffer, pH 8.0, adding enough TE buffer so that the labeled DNA is approximately 1000 cpm/ μ l.

PREPARATION OF DNA FOR ASSEMBLY OF HIGH CONCENTRATION UNLABELED MONONUCLEOSOMES

Large amounts of mononucleosome-length (147 to 155 bp) DNA can be prepared by digestion from a plasmid as described in Support Protocol 2, steps 1 and 2. The DNA fragment can also be made by PCR (UNIT 15.1) from the plasmid, using appropriate primers. Once large amounts of the fragment have been obtained, the fragment can be gel purified as described in the Basic Protocol 1 of UNIT 2.7. The authors find that the crush-soak method described in Basic Protocol 1 of UNIT 2.7 works well, giving > 80% recovery of the DNA.

PREPARATION OF DNA FOR NUCLEOSOMAL ARRAYS

This protocol is described for obtaining the G5E4 DNA but modifications of this protocol can be used for generating other 5S DNA fragments.

Materials

p2085-G5E4 plasmid (Dr. Jerry Workman; *jlw10@psu.edu*) or other 12S array plasmid

40 U/ μ l *Asp718* and *ClaI* restriction endonucleases and buffer B (all from Boehringer-Mannheim)

DdeI restriction endonuclease and buffer 3 (both from New England Biolabs)
65°C water bath

Additional reagents and equipment for phenol/chloroform extraction and ethanol precipitation of DNA (UNIT 2.1A), purification of restriction fragments (UNIT 2.6), and labeling with Klenow fragment (UNIT 3.5)

- Mix ~200 μ g of p2085S-G5E4 plasmid with 200 U each of *Asp718* and *ClaI* in buffer B (Boehringer-Mannheim) and bring up to a total of 160 μ l. Digest at 37°C for 3 hr, then heat for 15 min at 65°C.

High-concentration (40 U/ μ l) stocks of Asp718 and ClaI are used to avoid high glycerol concentrations in the reaction mix.

- Perform a phenol/chloroform extraction followed by a chloroform extraction and ethanol precipitation (UNIT 2.1A).
- Resuspend in ~98 μ l water. Digest with 100 U of *DdeI* using buffer 3 in total of 115 μ l for 1.5 hr at 37°C.

The DdeI digestion digests the backbone of the plasmid making it easier to separate the ~2.5 kb array template from the backbone.

- Purify the ~2.5 kb DNA fragment using the protocols in UNIT 2.6.
- Label the array template at one end using the Klenow protocol in UNIT 3.5. Use a mix of NTPs without dATP and label using [γ -³²P]dATP.

This will label the Asp718 end.

**SUPPORT
PROTOCOL 3**

**SUPPORT
PROTOCOL 4**

**Chromatin
Assembly and
Analysis**

21.6.9

ANALYSIS OF RECONSTITUTED COMPLEXES BY AGAROSE GEL ELECTROPHORESIS

The analysis of nucleosomes after reconstitution is necessary in order to determine the efficiency of nucleosome assembly, and also provides a way to gauge the integrity of the reconstituted nucleosomes. There are a number of techniques to analyze nucleosomes; combined together, they provide a reliable interpretation of the outcome of the assembly.

Materials

- Agarose (Research Genetics)
- 5× TBE buffer (see recipe)
- Labeled or unlabeled nucleosomes (see Basic Protocols 1 and 2)
- 50% (v/v) glycerol
- DNA loading dye (see recipe)
- 0.5 µg/ml ethidium bromide in 0.5× TBE buffer (see recipe for 5× TBE buffer)
- Whatman filter paper
- X-ray film (Kodak)
- Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5A*) and autoradiography (*APPENDIX 3A*)

1. Prepare a thin (~0.4 cm × 10 × 13 cm) horizontal 0.7% agarose gel containing 0.5× TBE buffer. Set up the electrophoresis apparatus and begin running the gel at ~10 V/cm.
2. Prepare the sample by mixing 9 µl (2000 cpm of labeled nucleosomes) of the final solution (or an appropriate amount adjusted to this volume) with 1 µl of 50% glycerol and loading onto the gel while the gel is running.

A naked DNA marker containing SDS and typical gel dyes may be run in adjacent lanes without detectable effects on the nucleoprotein complexes.
3. Run the gel at ~10 V/cm at room temperature until the bromophenol blue dye marker has run ~1/2 to 2/3 of the way through the gel.
4. Place the gels onto filter paper, cover with plastic wrap, and dry in gel dryer without heat for 15 min, then with heat set to 55°C for ~1 hr.
- 5a. *For labeled nucleosomes:* Expose the gel to X-ray film (autoradiography; *APPENDIX 3A*) for 1 to 12 hr, depending on the cpm loaded.

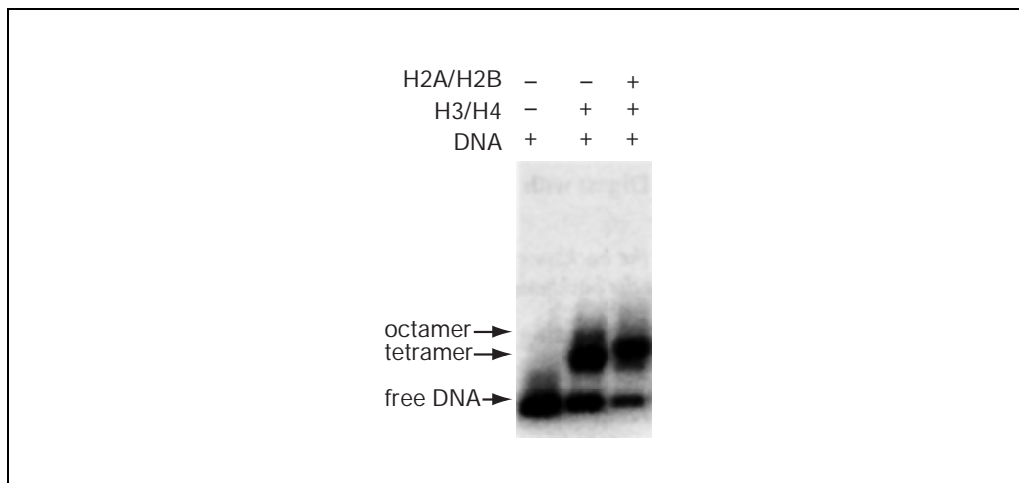


Figure 21.6.2 Gel retardation assay of reconstituted nucleosomes.

5b. *For unlabeled nucleosomes:* Stain the gel in 0.5 µg/ml ethidium bromide by gently agitating for ~20 min. Rinse with distilled water and visualize under UV light.

The results of a typical agarose gel are shown in Figure 21.6.2.

ANALYSIS OF RECONSTITUTED COMPLEXES BY POLYACRYLAMIDE GEL ELECTROPHORESIS

**SUPPORT
PROTOCOL 6**

Assembled mononucleosomes can be resolved from naked DNA fragment on 5% acrylamide, 0.5× TBE gel of dimensions ~10 cm × 8 cm × 1 mm.

Materials

40% acrylamide 29:1 (acrylamide:bis) solution (see *UNIT 2.7*)

5× TBE buffer (see recipe)

Labeled or unlabeled nucleosomes (see Basic Protocols 1 and 2)

Naked DNA template used to prepare nucleosomes, as control

50% (v/v) glycerol

DNA loading dye

0.5 µg/ml of ethidium bromide solution in 0.5× TBE buffer (see recipe for 5× TBE buffer)

X-ray film (Kodak)

Additional reagents and equipment for polyacrylamide gel electrophoresis of DNA (*UNIT 2.7*)

1. Prepare a 5% acrylamide, 0.5× TBE gel of dimensions ~10 cm × 8 cm × 1 mm (*UNIT 2.7*) using 40% acrylamide 29:1 (acrylamide:bis) solution and 5× TBE buffer. Set up the electrophoresis apparatus.
2. Load the nucleosomes by mixing 10 µl of the sample with 0.1 vol of 50% glycerol. Add loading dye containing bromophenol blue to a separate lane. As a control, add 10 µl of an appropriate amount of the naked DNA template.

Do not add dye to samples as this can disrupt histone-DNA interactions. Loading can be visualized due to the viscosity of the glycerol.

3. Run the gel at ~100 V at 4°C till the bromophenol dye is close to the bottom of the gel.
- 4a. *For labeled nucleosomes:* Dry the gel on a gel dryer for ~30 min at 80°C and expose to X-ray film (autoradiography; *APPENDIX 3A*).
- 4b. *For unlabeled nucleosomes:* Stain the gel in 0.5 µg/ml of ethidium bromide solution by gently agitating for ~20 min. Rinse with distilled water and visualize using a UV transilluminator.

The nucleosomes run significantly more slowly than the naked DNA. The ethidium fluorescence signal from the nucleosomal bands should be multiplied by 2.5 to correct for histone quenching.

The efficiency of assembly can be determined by quantifying the percentage of DNA incorporated into nucleosomes.

**Chromatin
Assembly and
Analysis**

21.6.11

***Eco*RI DIGESTION TO DETERMINE EXTENT OF G5E4 ARRAY ASSEMBLY**

Estimating the degree of array assembly is often crucial for interpretation of results from experiments using the array (Carruthers et al., 1999). The G5E4 fragment has *Eco*RI sites flanking each 5S positioning sequence but not within the central 400 bp. If all the 5S positioning sequences are saturated with histone octamers, cleavage with *Eco*RI will generate mononucleosomes and no 208-bp-length naked DNA. The central two nucleosomes, if fully assembled, will generate a dinucleosome. The percentage of naked DNA of length 208 and ~400 bp will be indicative of the degree of array assembly. If unlabeled array is being analyzed, the ethidium fluorescence signal from the nucleosomal bands should be multiplied by 2.5 to correct for histone quenching.

Materials

Nucleosomal array (see Basic Protocol 3)

*Eco*RI restriction endonuclease and corresponding buffer

25% (v/v) glycerol/10 mM disodium EDTA, pH 8.0 (see APPENDIX 2 for preparation of the EDTA)

Agarose (Research Genetics)

5× TBE buffer with increased EDTA (see recipe)

0.5 µg/ml ethidium bromide solution in 0.5× TBE buffer with regular EDTA concentration (see recipe for 5× TBE buffer)

Additional reagents and equipment for restriction endonuclease digestion (UNIT 3.1), agarose gel electrophoresis (UNIT 2.5A), and autoradiography (APPENDIX 3A)

1. Digest ~1 µg of nucleosomal array with 10 U of *Eco*RI in 20 to 100 µl for 60 min at room temperature in *Eco*RI digestion buffer.
2. Quench the reaction by adding 1/5 vol of 25% glycerol/10 mM EDTA, pH 8.0.
3. Prepare a 0.8% nondenaturing agarose gel in 0.5× TBE buffer with increased EDTA (UNIT 2.5A). Assemble the gel apparatus, load the samples, and run for 3 hr at 1 V/cm.

Do not add dye to samples as it can disrupt histone-DNA interactions. Add it in a separate lane.

- 4a. *If labeled array is being analyzed:* Place the gels onto filter paper, cover with plastic wrap, and dry in a gel dryer without heat for 15 min, then with heat set to 55°C for ~1 hr. Analyze by autoradiography (APPENDIX 3A) or using a phosphor imager.
- 4b. *If unlabeled array is being analyzed:* Stain the gel in 0.5 µg/ml ethidium bromide by gently agitating for ~20 min. Rinse with distilled water and visualize on a UV transilluminator.

The TBE buffer composition is crucial to obtain well-defined bands on the gel. Use of TAE results in smears.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Denaturing solution

6 M guanidium hydrochloride

20 mM sodium acetate

1 mM DTT

Prepare fresh before use

DNA loading dye

50% (v/v) glycerol
0.1% (w/v) bromophenol blue
0.1% (w/v) xylene cyanol

High-salt buffer

20 mM Tris·Cl, pH 7.7 (*APPENDIX 2*)
2 M KCl (2 M NaCl for nucleosome array assemblies)
1 mM EDTA
1 mM DTT
0.5 mM benzamidine
Store up to several weeks at 4°C without DTT or benzamidine

Low-salt buffer

20 mM Tris·Cl, pH 7.7 (*APPENDIX 2*)
250 mM KCl (250 mM NaCl for nucleosome assemblies)
1 mM EDTA
1 mM DTT
0.5 mM benzamidine
Store up to several weeks at 4°C without DTT or benzamidine

Tris/borate/EDTA (TBE) buffer, 5× (for Support Protocols 5 and 6)

0.45 M Tris·Cl, pH 8.3 (*APPENDIX 2*)
0.45 M boric acid
5.0 mM EDTA
Store up to several weeks at room temperature

Tris/borate/EDTA (TBE) buffer with increased EDTA, 5× (for Support Protocol 7)

0.45 M Tris·Cl, pH 8.3 (*APPENDIX 2*)
0.45 M boric acid
10 mM EDTA
Store up to several weeks at room temperature

Zero-salt buffer

20 mM Tris·Cl, pH 7.7 (*APPENDIX 2*)
1 mM EDTA
1 mM DTT
0.5 mM benzamidine
Store up to several weeks at 4°C without DTT or benzamidine

COMMENTARY**Background Information**

Chromatin is the substrate of virtually all processes within the eukaryotic nucleus that involve DNA. The structure and protein components of chromatin are increasingly being identified as integral components of these processes, which include transcription, replication, and recombination (van Holde, 1989).

The basic building block of the chromatin fiber is a 250-kDa structure known as the nucleosome. This subunit contains about 200 bp of DNA, two copies each of the four core histones, and generally a single molecule of a linker histone. The DNA within this complex

is partially wrapped in two 80-bp turns around a spool formed by the core histones. The core histones and the central 146 bp of DNA in tightest association with these proteins form a structure resistant to micrococcal nuclease digestion, referred to as the nucleosome core. The details of the organization of DNA and the proteins within the nucleosome core have been well-documented (Luger et al., 1997b). The remainder of the DNA within the nucleosome subunit (~40 bp) links nucleosome cores together to form a continuous string of these subunits. Strings of nucleosomes are further coiled or compacted into fibers approximately

Table 21.6.1 General Troubleshooting Guide For Assembly of Nucleosomal Templates by Salt Dialysis

Problem	Possible cause	Solution
Histones precipitate during renaturation dialysis.	Excess DNA present in the preparation.	Purify denatured histones away from DNA on size exclusion column before renaturation (Support Protocol 1).
When run on gel mononucleosome assembly is mainly in the wells.	Weight ratio of histones to DNA is significantly greater than 1 causing nonspecific aggregation.	Repeat assembly with lower amounts of histones and perform a titration with a few different concentrations to obtain the optimal ratio.
Complexes that run more slowly than mononucleosomes are observed on gel.	The higher complexes are likely to have more than one histone octamer per DNA fragment. This can arise from histone:DNA ratios of >1.	If the higher complexes form a significant fraction of assembly, one can perform histone titration as above. If these complexes are not a significant fraction of the total assembly, they can be separated away from standard mononucleosomes on a glycerol gradient (Basic Protocol 2).
Less than 30% of the DNA is incorporated into nucleosomes.	The histone:DNA ratio is significantly less than 1.	Perform a titration with increasing amounts of histones.
Nucleosomes look “smearly” when run on gel.	The gel temperature is high causing dissociation of nucleosomes.	Maintain gel at room temperature for agarose gel and run acrylamide gel in cold room.
<i>Eco</i> RI digestion of array gives smearly bands on agarose gel.	The running and gel buffer composition is different from recipe.	Use 45 mM Tris-borate, 1 mM EDTA buffer (see Support Protocol 7).

30 nm in diameter. Unfortunately, details of this compaction, such as the conformation of the linker DNA within the 30-nm fiber or the packing of nucleosome cores within the fiber, have not been elucidated. These fibers are organized into even higher-order structures to form the metaphase chromosome.

Efforts to recapitulate in vivo the processes which occur within the eukaryotic nucleus require the reconstruction of proper chromatin substrates. Thus, methods for the reconstitution of chromatin complexes have become widely used (Camerini-Otero, 1976). Moreover, reconstitution methods allow the construction of model chromatin complexes containing defined DNA sequences. Salt dialysis reconstitution has been useful in this respect. It is known that at 2 M salt, the histone octamer forms its canonical structure even when not bound to DNA. As the salt concentration falls below 2 M, the histone octamer dissociates into an H3/H4 and two H2A/H2B dimers. The tetramer also starts to bind to DNA at this salt concentration. At ~0.8 M salt, a single H2A/H2B dimer binds to each H3/H4 tetramer and by 0.6 M

nucleosome assembly is complete. Samples of reconstituted chromatin complexes can be prepared that exhibit less structural heterogeneity than preparations of natural chromatin complexes. This homogeneity has allowed detailed investigations of nucleosomal DNA structure and the salt-induced folding properties of nucleosomal arrays (Hayes et al., 1990, Schwarz and Hansen, 1994).

Reconstituted nucleosomes exhibit physical and structural properties identical to native complexes isolated from eukaryotic nuclei. Hydroxyl radical cleavage analysis and micrococcal nuclease digestions have shown that the DNA structure within reconstituted complexes is identical to that found in bulk preparations containing random DNA sequences (Hayes et al., 1990). Cross-linking experiments have demonstrated that the histone protein–DNA contacts within reconstituted nucleosomes are identical to that of native complexes (Lee and Hayes, 1997). Reconstituted oligonucleosomal complexes exhibit the same hydrodynamic shape and folding behavior as native complexes (Schwarz and Hansen, 1994). Further, the

stoichiometry of polypeptides within reconstituted complexes is identical to that found in native chromatin (Camerini-Otero et al., 1976).

A key element in many of these studies is the use of a well-defined nucleosome positioning sequence. Sequence-dependent variations in the B-DNA structure within these fragments cause the fragments to adopt a defined rotational and translational orientation with respect to the surface of the nucleosome core. Since the core histone-DNA contacts occur at specific locations on the surface of the octamer, these two positional variables are related. However, a single preferred rotational orientation may accommodate several translational frames related by approximately $10n$ base pairs. Therefore, reconstitutions with stably curved DNAs will exhibit a strong rotational preference but may yield a population with a heterogeneous distribution of translational frames. DNA fragments that contain 5S RNA gene sequences have been found to strongly direct translational positioning (Rhodes, 1985). A 5S DNA fragment from *Xenopus* exhibits has been found to exhibit exceptional nucleosome positioning properties. When this fragment is reconstituted with purified core histones, approximately 80% of the population of nucleosomes within the sample adopt a unique translational position (Hayes et al., 1990). The molecular basis of this translational positioning effect is not yet understood.

Critical Parameters and Troubleshooting

Critical variables and ways to overcome experimental difficulties are presented in Table 21.6.1.

Anticipated Results

The results obtained from the analysis should reveal the quality of the nucleosomal template (Support Protocol 3). If the purity of the histones and DNA is high and consideration is given to the relative stoichiometry, the assembly should generate suitable template for further experiments. The efficiency of assembly is obtained by quantifying the percentage of naked DNA assembled into nucleosomes as described in Support Protocols 5, 6, and 7. This should be between 30% and 80%.

Time Considerations

Theoretically, step salt dialysis can be completed in 10 hr. Step dialysis has the advantage of being quick and convenient to set up. However, the drawbacks are that it does not always

generate properly assembled nucleosomes when the nucleosome concentration is high (~1 mg) and when nucleosomal arrays are generated. An alternative method is to use the gradient salt dialysis method (see Basic Protocol 2). The assembly through gradient dialysis takes ~50 hr. It is a relatively long procedure but once it is set up, there are no further manipulations until the reconstitution is finished. The purification of histones from native chromatin takes roughly two full days after the necessary equipment is set up. The isolation of recombinant histones from bacteria requires 3 to 4 half days starting from the inoculation to the final chromatography steps. The analysis of the nucleosomes takes about 3 to 4 hr.

Literature Cited

- Camerini-Otero, R.D., Sollner-Webb, B., and Felsenfeld, G. 1973. The organization of histones and DNA in chromatin: evidence for an arginine-rich histone kernel. *Cell* 3:333-347.
- Carruthers, L.M., Tse, C., Walker, K.P., and Hansen, J. 1999. Assembly of defined nucleosomal and chromatin arrays from pure components. *Methods Enzymol.* 304:19-35.
- Hayes, J.J. and Lee, K.M. 1997. In vitro reconstitution and analysis of mononucleosomes containing defined DNAs and proteins. *Methods* 12:2-9.
- Hayes, J.J., Tullius, T.D., and Wolffe, A.P. 1990. The structure of DNA in a nucleosome. *Proc. Natl. Acad. Sci. U.S.A.* 87:7405-7409.
- Lee, K.M. and Hayes, J.J. 1997. The N-terminal tail of histone H2A binds to two distinct sites within the nucleosome core. *Proc. Natl. Acad. Sci. U.S.A.* 94:8959-8964.
- Lee, K.M., Chafin, D.R., and Hayes, J.J. 1999. Targeted cross-linking and DNA cleavage within model chromatin complexes. *Methods Enzymol.* 304:231-251.
- Luger, K., Rechsteiner, T.J., Flaus, A.J., Wayne, M.M., and Richmond, T.J. 1997a. Characterization of nucleosome core particles containing histone protein made in bacteria. *J. Mol. Biol.* 272:301-311.
- Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. 1997b. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389:251-260.
- Luger, K., Rechsteiner, T.J., and Richmond, T.J. 1999. Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol.* 304:3-19.
- Neely K.E., Hassan, A.H., Wallberg, A.E., Steger, D.J., Cairns, B.R., Wright, A.P. and Workman, J.L. 1999. Activation domain-mediated targeting of the SWI/SNF complex to promoters stimulates transcription from nucleosome arrays. *Mol. Cell.* 4:649-655.

- Rhodes, D. 1985. Structural analysis of a triple complex between the histone octamer, a *Xenopus* gene for 5S RNA and transcription factor IIIA. *EMBO J.* 4:3473-3482.
- Schwarz, P.M. and Hansen, J.C. 1994. Formation and stability of higher order chromatin structures.contributions of the histone octamer. *J. Biol. Chem.* 269:16284-16289.
- van Holde, K.E. 1989. Chromatin. Springer-Verlag, New York

Contributed by Kyu-Min Lee and
Geeta Narlikar
Massachusetts General Hospital
Boston, Massachusetts

Chromatin Assembly Using *Drosophila* Systems

UNIT 21.7

To study chromatin structure and activity *in vitro*, it is necessary to use appropriate methods for the preparation of chromatin substrates. The goal of chromatin assembly procedures is to prepare extended nucleosome arrays from cloned DNA templates (usually plasmid DNA) and purified chromatin proteins (core and linker histones, DNA-binding proteins). The assembled chromatin should be highly defined in its protein content and resemble bulk chromatin isolated from living cell nuclei in terms of periodicity and nucleosome positioning. In this unit we describe two systems that assemble minichromosome templates in an ATP-dependent fashion from circular plasmid DNA and purified core histones. These systems can be used to assemble minichromosomes from linear DNA (plasmid and lambda) and can also incorporate proteins other than core histones (linker histone H1, HMG17, and DNA-binding transcription factors). The products of these chromatin assembly reactions have been used directly (or after purification) in assays to study transcription, DNA replication, recombination, and repair. The first of the two systems employs a *Drosophila* embryonic extract (S-190) as a source of the assembly factors (see Basic Protocols 1 to 3). Well established and easy to troubleshoot, this protocol should be considered the first stop for researchers with limited experience in chromatin analysis. The second system utilizes purified recombinant *Drosophila* chromatin assembly factors (see Basic Protocols 4 to 6 and Alternate Protocol 1). It provides an additional advantage of strictly defined protein content of assembled chromatin. Determining the optimal ratio of core histone to DNA is important for either system; this is presented in Alternate Protocol 2. Chromatin assembly on relaxed circular DNA requires the presence of a topoisomerase activity in the reaction. An example of expression and purification of the core catalytic domain of *Drosophila* topoisomerase I is presented in the Support Protocol.

PREPARATION OF THE *DROSOPHILA* S-190 CHROMATIN ASSEMBLY EXTRACT

BASIC
PROTOCOL 1

This protocol describes the preparation of the S-190 *Drosophila* cell extract. When supplemented with core histones and ATP, the S-190 extract can mediate the *in vitro* assembly of regularly spaced arrays of nucleosomes on a DNA template.

Materials

- 0- to 6-hr *Drosophila* embryos
- Bleach wash: 50% (v/v) household bleach in distilled H₂O
- Embryo wash: 0.7% (w/v) NaCl/0.4% (v/v) Triton X-100, room temperature and 4°C
- Saline wash: 0.7% (w/v) NaCl, 4°C
- Buffer R (see recipe), 4°C
- 1 M MgCl₂ (APPENDIX 2)
- Liquid N₂
- Fine nylon mesh (e.g., Sefar America 03-80/37)
- 800-ml glass beaker
- Glass rod
- Vacuum aspirator
- 40 ml Wheaton Dounce homogenizer with “A” and “B” pestles
- 17 × 100–mm polypropylene tubes (e.g., Falcon 2059)
- Sorvall Superspeed centrifuge with SS-34 rotor (or equivalent)

Chromatin
Assembly and
Analysis

21.7.1

50-ml conical tubes
10-ml plastic syringes and 18.5-G needles
Beckman ultracentrifuge with SW 41 and SW 55 rotors
Thin-walled 14 × 89-mm (for SW 41 rotor) and 13 × 51-mm (for SW 55 rotor)
Beckman ultracentrifuge tubes

Prepare the embryos

1. Secure a fine nylon mesh in a support and place ~100 g of 0- to 6-hr *Drosophila* embryos on the mesh. Wash thoroughly with cold tap water to eliminate yeast and contaminants.

Do not freeze the embryos before extraction.

100 g of the starting material should yield ~50 ml of S-190 extract. S-190 extracts prepared from embryos older than 6 hr are significantly less active. It is important to select an appropriate mesh that is fine enough to retain the embryos but large enough to allow the buffers and yeast to flow through. A coarser mesh placed above the fine mesh can be used to separate adult fly bodies and parts as well as pieces of food agar from the embryos.

2. Soak the embryos 90 sec in 3 liters bleach wash at room temperature. Quickly rinse with 1 liter embryo wash at room temperature. Rinse extensively with distilled water. Transfer embryos to an 800-ml glass beaker on ice.

Bleach wash will remove chorions from the embryos. After this point, all subsequent steps should be carried out on ice or in a cold room, unless otherwise noted.

3. Add 500 ml of embryo wash at 4°C and suspend the embryos by stirring with a glass rod. Allow the embryos to settle to the bottom of the beaker for ~2 min. Aspirate the embryo wash using a vacuum aspirator, avoiding the settled embryos at the bottom of the beaker.

Be sure to aspirate the chorions and other floating non-embryo material at the top of the beaker.

4. Repeat step 3 with an additional 500 ml of 4°C embryo wash.
5. Repeat step 3 twice, each time with 500 ml of 4°C saline wash.

Allow the embryos to settle ~5 min in saline wash. It is important to avoid aspirating the now floating embryos in the beaker.

6. Repeat step 3 twice, each time with 500 ml of 4°C buffer R.

The embryos will require up to 10 min to settle in buffer R. Again, it is important to avoid aspirating floating embryos as much as possible. After the second aspiration, the final volume should be no more than 80 ml. While it may require some loss of the starting material, it is more desirable to have a concentrated extract than a larger volume of the extract.

Prepare the extract

7. Transfer the material to a 40 ml Wheaton Dounce homogenizer on ice. Homogenize with 15 strokes of the “B” pestle and additional 40 strokes of the “A” pestle. Repeat with any remaining material.

Be certain to homogenize with full strokes. It is important to achieve complete homogenization of the embryos. This can be verified by placing a drop of the homogenate on a glass slide and examining for the presence of intact embryos.

8. Transfer the homogenate to 17 × 100-mm polypropylene tubes on ice. Spin 5 min at 7500 × g (8000 rpm in a Sorvall SS-34 rotor), 4°C. Extract the aqueous supernatant from the middle of the tubes and pool in 50-ml conical tube(s).

To avoid the pellet at the bottom and the white lipid layer at the top of the tube, use a 10-ml syringe and an 18.5-G needle inserted through the tube just above the pellet.

- Using 1 M MgCl_2 , adjust the concentration of magnesium to 7 mM.

Note that buffer R contains 1.5 mM MgCl_2 . This step significantly reduces the amount of RNA in the final extract.

- Transfer the supernatant to thin-walled 14×89 -mm ultracentrifuge tubes. Ultracentrifuge for 2.25 hr at $200,000 \times g$ (40,000 rpm in an ultracentrifuge with Beckman SW 41 rotor), 4°C .

For smaller volumes, use 13×51 -mm tubes and spin at $192,000 \times g$ (45,000 rpm in a Beckman SW 55 rotor), 4°C .

- Remove the white upper lipid layer from each tube with a metal spatula and discard. Collect the golden-brown liquid layer and pool in a 50-ml conical tube(s) on ice. Freeze in liquid nitrogen.

It is important to avoid the beige, upper layer of the pellet, which will inhibit the assembly activity. At this point, the extract can be stored at -80°C indefinitely until the final spin. However, freezing of the extract is essential even if you proceed directly to step 12.

Handle the 50-ml polypropylene tubes gently in liquid N_2 to avoid cracking.

- Thaw the extract in a room temperature water bath. Spin 2.25 hr at $192,000 \times g$ (45,000 rpm in an ultracentrifuge with Beckman SW 55 rotor), 4°C .

Note that the SW 55 rotor consistently yields a more active extract than the SW 41 for this spin.

- Remove the white upper lipid layer (if any) and pool the extract in a 50-ml conical tube. Aliquot 1-ml fractions into microcentrifuge tubes and freeze in liquid nitrogen. Store at -80°C .

The S-190 extract will remain active for up to 3 years at -80°C .

PURIFICATION OF CORE HISTONES FROM THE *DROSOPHILA* EMBRYOS

Nuclei are isolated from the *Drosophila* embryos. Chromatin is partially digested with micrococcal nuclease to generate ~ 2000 -bp fragments. The chromatin fragments are isolated, bound to hydroxylapatite resin, and core histones are separated from the DNA by salt elution.

Materials

0- to 12-hr *Drosophila* embryos

Bleach wash: 50% (v/v) household bleach in distilled H_2O

Embryo wash: 0.7% (w/v) NaCl/0.04% (v/v) Triton X-100

Buffer B (see recipe)

Buffer A (see recipe)

1 M NaOH

0.1 M CaCl_2

200 U/ml micrococcal nuclease stock solution (see recipe)

10 mM and 500 mM EDTA

10% (w/v) SDS

5 M NaCl (APPENDIX 2)

24:1 chloroform/isoamyl alcohol

Linear 5% to 30% sucrose gradients (see recipe)

T_{50}E_4 buffer: 50 mM Tris·Cl, pH 7.9 (APPENDIX 2)/4 mM EDTA

Hydroxylapatite resin (e.g., BioGel HT gel; Bio-Rad)

HA chromatography buffer (see recipe) with 0 M, 0.35 M, and 2.5 M NaCl

Core histone storage buffer (see recipe)

BCA assay kit (Pierce)

BASIC PROTOCOL 2

Chromatin Assembly and Analysis

21.7.3

Fine nylon mesh (e.g., Sefar America 03-80/37)
Weigh boats
Yamato LH-21 homogenizer (alternatively, Potter-Elvehjem homogenizer)
500-ml centrifuge bottles for Sorvall GSA rotor (or equivalent)
Mira cloth (Calbiochem-Novabiochem)
Sorvall Superspeed centrifuge with GSA and SS-34 rotors or equivalents
Beckman ultracentrifuge with SW-28 rotor and appropriate tubes
12,000 to 15,000 and 3,500 MWCO dialysis tubing
FPLC apparatus

Additional reagents and equipment for extraction of DNA (*UNIT 2.1A*), agarose gel electrophoresis (*UNIT 2.5A*), SDS-PAGE (*UNIT 10.2*), Coomassie blue staining and destaining of proteins in gels (*UNIT 10.6*), and dialysis (*APPENDIX 3C*)

Prepare the embryos

1. Secure a fine nylon mesh in a support and place ~100 g of 0- to 12-hr *Drosophila* embryos on the mesh. Wash thoroughly with cold tap water to eliminate yeast and contaminants.

The embryos can be frozen in liquid nitrogen and stored up to 1 year at -80°C before processing.

100 g of the starting material should yield between 5 to 10 mg of core histones. See step 1 of Basic Protocol 1 for information on the nylon mesh.

2. Soak the embryos 90 sec in 3 liters bleach wash at room temperature. Quickly rinse with 1 liter of embryo wash at room temperature. Rinse extensively with distilled water. Blot embryos dry through the mesh with paper towels. Transfer to a weigh boat and weigh.

Bleach wash will remove chorions from the embryos. After this point, all subsequent steps should be carried out on ice or in a cold room, unless otherwise noted.

Isolate the nuclei

3. Transfer the embryos to a glass beaker. Resuspend in 3 ml buffer B per 1 g of embryos. Pour the suspension through a Yamato LH-21 homogenizer at 1000 rpm and filter the effluent into GSA bottles through Mira cloth.

If a Yamato homogenizer is not available, the embryos can be disrupted with six to eight strokes in a motorized Potter-Elvehjem homogenizer (serrated Teflon pestle in a glass vessel), followed by several strokes in a Wheaton Dounce homogenizer with the "B" pestle.

4. Wash the original beaker with 1 ml buffer B per 1 g of embryos and pass through the homogenizer and Mira cloth as in step 3. Repeat with 1 ml buffer B per 1 g of embryos. Centrifuge the filtrate 20 min at $10,000 \times g$ (8000 rpm in a GSA rotor), 4°C .

The final volume of the suspension before spinning should be 5 ml buffer B per 1 g of embryos.

5. Discard the supernatant. Resuspend the loose pellet of nuclei in 200 ml buffer A. Centrifuge 10 min at $10,000 \times g$ (8000 rpm in a GSA rotor), 4°C .

Avoid the yellow yolk protein pellet when resuspending the nuclei. A firmer, "yellow" yolk pellet should form at the center of the looser nuclei pellet; it is this pellet that should be avoided. A Pasteur pipet can be used to trace the outline of the yolk pellet, after which gentle resuspension of the nuclear pellet can be carried out with a battery-powered pipetting device (e.g., Pipet-Aid) and buffer A.

6. Discard the supernatant. Resuspend the nuclei in 100 ml buffer A. Centrifuge 10 min at $10,000 \times g$ (8000 rpm in a GSA rotor), 4°C .

Again, avoid the yolk protein, if any remains.

7. Discard the supernatant. Resuspend the nuclear pellet in 30 ml buffer A.

Digest a sample of the chromatin

8. Dilute 10 μl of suspended nuclei with 990 μl of 1 M NaOH. Measure the optical density at 260 nm (OD_{260}). Dilute the suspension of nuclei to 100 OD_{260} units/ml with buffer A.
9. Perform a test micrococcal nuclease digestion as follows:
 - a. In a 1.5 ml microcentrifuge tube, prewarm 1 ml of the nuclear suspension to 37°C .
 - b. Add 10 μl of 0.1 M CaCl_2 and 2 μl of 200 U/ml micrococcal nuclease stock solution.
 - c. Incubate at 37°C . At 1, 2, 4, 6, 8, 10, 15, and 20 min, remove a 100- μl aliquot and stop enzyme activity in the aliquot with 2.5 μl 500 mM EDTA.
10. To each aliquot add 30 μl water and 20 μl 10% SDS. Mix thoroughly and add 40 μl 5 M NaCl. Mix thoroughly. Extract each aliquot with 200 μl of 24:1 chloroform/isoamyl alcohol.

UNIT 2.1A describes general techniques for extraction of DNA.

11. Mix 4 μl of the aqueous phase from each aliquot with agarose gel loading buffer. Run each of the aliquots (representing time points from step 9) in a separate lane of a 1% agarose gel. Stain with ethidium bromide and photograph (see *UNIT 2.5A* for reagents and techniques used in this step).

Use a 1-kb ladder as a reference and select the time point displaying a prevalence of ~2000-bp fragments.

Perform bulk digest of the chromatin

12. Warm the nuclear suspension to 37°C and add 1/100 vol of 0.1 M CaCl_2 . Add 1/500 vol of 200 U/ml micrococcal nuclease and mix by swirling. Incubate at 37°C with occasional swirling for the time determined in steps 8 to 11. Stop digestion by adding of 1/50 vol of 500 mM EDTA and place at 4°C .
13. Centrifuge 10 min at $12,000 \times g$ (10,000 rpm in an SS-34 rotor), 4°C . Discard the supernatant. Resuspend the pellet in 10 mM EDTA for a total volume of 9 ml. Add 1 ml of 5 M NaCl and swirl 5 min.

Upon the addition of NaCl, extensive lysis should be observed and the solution should become more viscous.
14. Centrifuge 5 min at $12,000 \times g$ (10,000 rpm in an SS-34 rotor), 4°C . Save the supernatant. Dilute 10 μl into 990 μl of 1 M NaOH and measure the OD_{260} .
15. Load $\leq 500 \text{OD}_{260}$ units of supernatant onto linear gradients (18 to 20 ml each) of 5% to 30% sucrose in sucrose gradient buffer in SW 28 tubes. Ultracentrifuge 16 hr at $90,000 \times g$ (26,000 rpm in an SW 28 rotor), 4°C .
16. Cut the gradients into 2-ml fractions and run 8 μl per fraction on a 15% SDS-PAGE gel (*UNIT 10.2*). Coomassie stain and destain (*UNIT 10.6*). Pool the peak core histone containing fractions and dialyze in 12,000 to 15,000 MWCO tubing twice, each time for 2 hr against 2 liters T_{50}E_4 buffer (*APPENDIX 3C*) at 4°C .

Like fractions from each gradient tube can be pooled in 50-ml conical tubes. If any debris is suspended in sucrose-containing fractions, it can be spun down after dialysis. The fractionated digested chromatin can be stored for several months at 4°C without loss of core histone integrity. When selecting fractions for further purification, be careful to avoid histone H1-containing fractions.

Purify core histone octamers by hydroxylapatite chromatography

17. Dilute 10 µl dialyzed sample with 990 µl of 1 M NaOH and measure the OD₂₆₀. Calculate the required volume of hydroxylapatite resin as 1 ml column volume per 1.5 mg of DNA.
18. Pack an appropriately sized hydroxylapatite column on an FPLC and equilibrate in 3 column volumes HA chromatography buffer without NaCl.
19. Apply the dialyzed sample to the column and wash with an additional 3 column volumes of HA chromatography buffer without NaCl. Wash with an additional 3 column volumes of HA chromatography buffer containing 0.35 M NaCl.

The 0.35 M NaCl wash elutes DNA-binding proteins, other than histones, off the resin-bound chromatin.

20. Elute the core histones with 2 column volumes of HA chromatography buffer containing 2.5 M NaCl.

Alternatively, the core histones can be eluted with a salt gradient from 0.35 to 2.5 M NaCl in HA chromatography buffer. This separates any residual histone H1 that was collected from the sucrose gradient. A variation of this procedure can also be used to fractionate histone dimers and tetramers.

Additional information on hydroxylapatite chromatography of histones is found in UNIT 21.5.

21. Analyze 2 µl of peak fractions on a 15% SDS-PAGE gel (UNIT 10.2). Coomassie stain and destain (UNIT 10.6). Pool the peak core histone-containing fractions and dialyze in 3,500 MWCO tubing twice, each time for 2 hr against 2 liters core histone storage buffer at 4°C.
22. After dialysis, determine the histone concentration by BCA assay. Aliquot into microcentrifuge tubes and freeze in liquid nitrogen. Store at -80°C.

The core histones can be stored for several years at -80°C.

CHROMATIN ASSEMBLY WITH THE S-190 EXTRACT

The extract (from Basic Protocol 1) and core histones (from Basic Protocol 2) are combined and preincubated at room temperature to facilitate binding of histones and chaperone(s). A DNA template, ATP, and ATP regeneration system are added, and chromatin is assembled at 27°C. Chromatin is analyzed by partial micrococcal nuclease digestion.

Materials

- Buffer R (see recipe)
- S-190 extract (see Basic Protocol 1)
- Core histones (see Basic Protocol 2)
- ATP mix (see recipe)
- 0.1 M MgCl₂
- DNA template (2.5 to 25 kbp circular or linearized plasmid, or λ DNA)
- 0.1 M CaCl₂
- 200 U/ml micrococcal nuclease stock solution (see recipe)

0.5 M EDTA
10 mg/ml RNase A
Glycogen stop buffer (see recipe)
2.5 mg/ml proteinase K
50:49:1 (v/v/v) phenol/chloroform/isoamyl alcohol, equilibrated with 10 mM
Tris-Cl, pH 8.0 (see *UNIT 2.1A* for equilibration technique)
2.5 M ammonium acetate
70% and 100% ethanol
1.2% agarose gel (*UNIT 2.5A*)
TBE buffer (*APPENDIX 2*)
123 bp DNA ladder (Life Technologies)
27° and 37°C water baths

Additional reagents and equipment for extraction of DNA (*UNIT 2.1A*) and agarose gel electrophoresis (*UNIT 2.5A*)

Preincubate core histones with the extract

1. Combine 40 μ l buffer R, 30 μ l S-190 extract, and 400 ng of core histones. Incubate 30 min at room temperature.

A single reaction produces enough chromatin for one micrococcal nuclease digestion analysis (two lanes on an agarose gel). The reaction can be scaled up. Typically, the S-190 extract comprises 25% to 40% of the reaction volume and the core histones are in a 0.8:1 mass ratio with the DNA template (5 ng/ μ l). The final Mg²⁺ concentration is 7 mM.

Add ATP and the DNA template

2. Add 10 μ l ATP mix, 0.1 M MgCl₂ to 7 mM final concentration, and 500 ng DNA template. Incubate 5 hr at 27°C. If necessary, make up to a final volume of 100 μ l with buffer R.

Note that both buffer R and the S-190 extract contain MgCl₂.

Perform partial micrococcal nuclease digestions

3. Add 3 μ l of 0.1 M CaCl₂ to the assembly reaction. Divide the reaction in two equal parts (“a” and “b”). Dilute the micrococcal nuclease (200 U/ml) 1:50 and 1:150 with buffer R.

Prepare the nuclease dilutions within a few minutes before use.

4. In a controlled manner (at certain time intervals, e.g., 15 sec), add 5 μ l of the 1:150 dilution to the “a” tube and 5 μ l of the 1:50 dilution to the “b” tube. Digest 10 min at room temperature.
5. Terminate the reactions with 7 μ l of 0.5 M EDTA. Add 1 μ l of 10 mg/ml RNase A solution and incubate for 10 min at room temperature.

This step eliminates RNA present in the S-190 extract.

Deplete the chromatin samples

6. To each aliquot add 95 μ l glycogen stop buffer and 5 μ l of 2.5 mg/ml proteinase K solution. Incubate 30 min at 37°C.
7. Extract with 200 μ l of 50:49:1 phenol/chloroform/isoamyl alcohol. Precipitate the DNA from 150 μ l of the aqueous phase with 15 μ l of 2.5 M ammonium acetate and 450 μ l of 100% ethanol. Microcentrifuge for 15 min at maximum speed, room temperature. Wash the pellets with 200 μ l of 70% ethanol. Microcentrifuge 5 min at maximum speed, room temperature. Dry by exposing to air for 5 min.

UNIT 2.1A describes general techniques for extraction and precipitation of DNA.

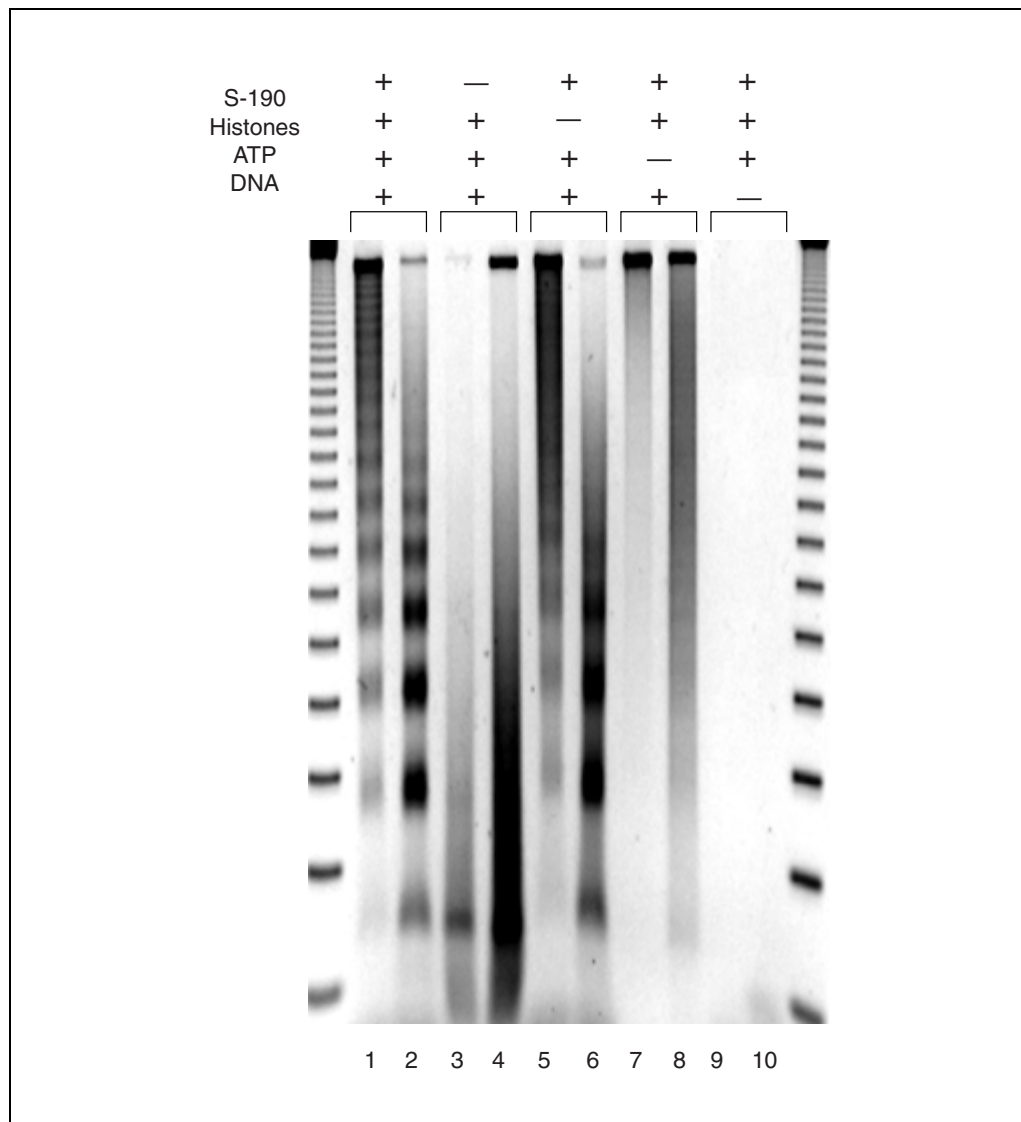


Figure 21.7.1 Chromatin assembly by the S-190 extract. Lanes 1, 2—complete system, as described in Basic Protocol 3; lanes 3, 4—no S-190 extract; lanes 5, 6—no core histones (note residual assembly activity with the endogenous histones of the extract); lanes 7, 8—no ATP; lanes 9, 10—no DNA template.

Analyze micrococcal nuclease ladders

- Resuspend the pellets in 6 μ l gel loading buffer and run on a 1.2% agarose gel in 1 \times TBE buffer. Use 123 bp DNA ladder as a marker. Stain with ethidium bromide (see *UNIT 2.5A* for the reagents and techniques used in this step).

Run the gel at ~5 to 10 V/cm. Running too quickly will cause uneven heating of the gel and tilting of the DNA bands, which will interfere with detection of the bands at the top of the ladder. Running the gel too slowly will cause excessive diffusion of the bands at the bottom of the ladder. The electrophoresis should be terminated when the bromphenol blue dye front reaches the bottom quarter of the gel. Bromphenol blue migrates at ~250 bp DNA size, while the mononucleosome band has a size of ~170 bp. For the best results, stain the gel 15 to 20 min in 2 gel volumes of 0.75 μ g/ml ethidium bromide in distilled water. Do not overstain. Destain the gel for 1 to 3 hr in distilled water with several changes. A ladder of at least 7 to 8 nucleosomal bands, counting from the bottom up, should be apparent. The repeat length should be 165 base pairs or more. An example of the type of results obtained is shown in Figure 21.7.1.

EXPRESSION AND PURIFICATION OF THE RECOMBINANT *DROSOPHILA* ACF

BASIC PROTOCOL 4

Drosophila ACF (ATP-utilizing chromatin assembly and remodeling factor) is prepared by coexpression of carboxyl-terminally FLAG-tagged Acf1 subunit with untagged ISWI subunit in baculovirus. The complex is then purified in one step by FLAG immunoaffinity chromatography and eluted with a buffer containing the FLAG peptide. This procedure typically results in a stoichiometric complex of p185 isoform of Acf1 and ISWI. In a slight alteration of this protocol, it is possible to express and purify the individual FLAG-tagged ISWI subunit.

Materials

High titer Acf1-FLAG and ISWI baculovirus stocks (Orbigen; also see *UNITS 16.9 & 16.10*)

Late log phase Sf9 cells cultured in suspension ($>2 \times 10^6$ cells/ml; also see *UNITS 16.9 & 16.10*)

Phosphate-buffered saline (PBS; *APPENDIX 2*), ice-cold

Lysis buffer F (see recipe)

1:1 (v/v) slurry of FLAG-M2 resin (Sigma-Aldrich), equilibrated in lysis buffer F

Dilution buffer F (see recipe)

Wash buffer F (see recipe)

Elution buffer F (see recipe)

Liquid nitrogen

Bovine serum albumin (BSA) standard, 2 mg/ml (Pierce, Cat. No. 23209)

Clinical centrifuge with swinging-bucket rotor, 4°C

250-ml conical centrifuge bottles appropriate for clinical centrifuge, or 50-ml conical tubes

14- and 50-ml disposable conical centrifuge tubes

15-ml Wheaton dounce homogenizer, "A" pestle

Sorvall Superspeed centrifuge with SS-34 rotor (or equivalent)

15-ml capped polypropylene tubes

Siliconized 1.5-ml polypropylene tubes (e.g., ISC BioExpress, Cat. # C-3302-1)

Additional reagents and equipment for baculovirus culture (*UNITS 16.9 & 16.10*), SDS-PAGE (*UNIT 10.2*), and staining of gels (*UNIT 10.6*)

Infect and harvest Sf9 cells

1. Amplify baculovirus stocks several days before the infection (*UNIT 16.10*).

Viruses are supplied as cell culture supernatants and are to be stored at 4°C. Titers of the original stocks may vary ($1-5 \times 10^8$ pfu per ml; see Orbigen's protocols; <http://www.orbigen.com>). To amplify the viruses, Sf9 cells are plated on 150-mm culture plates at 2×10^7 cells/plate in a total of 25 ml of appropriate serum-containing medium, and infected at multiplicity of infection (MOI) of 0.1 to 0.5 (10 to 20 μ l viral supernatant per plate). To pass a virus stock for storage, allow the infection to proceed for 60 hr and collect the medium supernatant aseptically. For high-titer stock (to be further used for expression infection), the infection is allowed to progress until cell lysis is apparent (~72 hr for His-NAP-1 virus, ~84 hr for ISWI and Acf1-FLAG). This procedure produces viral stocks with titers close to or above 10^9 pfu per ml. The medium is aspirated off the plates aseptically and stored in sterile 50-ml tubes in the dark at 4°C for up to 12 months without significant loss of titer.

2. Grow Sf9 cells (see *UNIT 16.10*) in 150 or 500 ml spinner flasks for 2 to 3 days after seeding at 0.5×10^6 cells/ml. Plate 5 to 25 plates of Sf9 cells at 2.5 to 3×10^7 cells/plate in a total of 25 ml of appropriate insect medium per plate. Allow the cells to settle

for 20 min in the tissue culture hood and infect with recombinant Acf1-FLAG and ISWI baculoviruses at an MOI of 5 to 10 each.

It is not necessary to completely replace the old medium. Typically, equal volumes (0.3 to 0.7 ml) of both high titer virus stocks are used. Occasionally, the ratio of Acf1-FLAG and ISWI viruses has to be adjusted empirically to ensure a proper stoichiometry.

3. At 44 to 46 hr subsequent to infection, aspirate the medium and wash the cells off the plates with 10 ml ice-cold PBS per plate. Centrifuge 5 min at 2000 rpm in a clinical centrifuge in 250-ml conical bottles or 50-ml tubes at 4°C.

Do not exceed the 2-day infection time, since it may result in lower yields and accumulation of complexes with partially degraded Acf1 and ISWI. From this point on, work in a cold room or on ice. Cell pellets can be frozen in liquid nitrogen and stored at -80°C for several weeks before further processing.

4. Resuspend the cell pellet in 8 ml of lysis buffer F and disrupt with a Wheaton Dounce homogenizer ("A" pestle; 3 series of 10 strokes over a 30 min period, on ice).

Cell nuclei are lysed by 500 mM salt. Both nuclei and cytosol are extracted. Cell lysis can be monitored microscopically, if desired. However, the appearance of the cell debris pellet after centrifugation, with distinct yolk and nuclear phases, may serve as an equally good indication of lysis completeness.

Prepare and analyze ACF protein

5. Pellet insoluble material by centrifuging in 14-ml disposable conical tubes 10 min at $14,500 \times g$ (11,000 rpm in an SS-34 rotor), 4°C. Combine the supernatant with 250 μ l of FLAG-M2 resin (as a 1:1 slurry equilibrated in lysis buffer F) and 7 ml dilution buffer F. Mix the slurry on a rocking platform for 3 to 4 hr at 4°C in a 15-ml capped polypropylene tube.

The salt concentration is decreased to allow efficient binding of the FLAG-tag to the M2 antibody.

6. Wash the resin four times, each time with 12 ml wash buffer F by successive cycles of centrifugation for 3 min at 2000 rpm in the clinical centrifuge at 4°C, followed by aspiration and resuspension.

Mix by inverting; avoid vigorous shaking.

7. Elute protein as follows:

- a. Add 100 μ l elution buffer F to resin pellet from step 6 and resuspend.
- b. Transfer resin to a 1.5-ml siliconized microcentrifuge tube.
- c. Incubate on ice 10 min.
- d. Microcentrifuge 30 sec at maximum speed.
- e. Transfer the supernatant to another tube (to be pooled with subsequent elutions).
- f. Continuing in the same siliconized microcentrifuge tube, repeat steps 7a, 7c, 7d, and 7e, three additional times, pooling all of the eluates together.

Insulin is thought to stabilize the recombinant protein at low concentrations.

8. Freeze the protein in liquid nitrogen in small aliquots (20 to 50 μ l) and store at -80°C.

The recombinant ACF is stable for several years and can withstand multiple (5 to 10) freeze-thaw cycles. Prepare and freeze several aliquots of wash buffer F containing 0.4 mg/ml insulin. This solution will be used further as ACF dilution buffer.

9. *Optional*: Estimate protein concentration by SDS-PAGE (UNIT 10.2) along with a range of concentrations of the BSA standard, followed by staining with Coomassie Brilliant Blue R-250 (UNIT 10.6).

Typical yields of ACF are ~5 to 10 µg per 150-mm plate.

EXPRESSION AND PURIFICATION OF THE RECOMBINANT DROSOPHILA NAP-1

**BASIC
PROTOCOL 5**

Sf9 cells are infected with dNAP-1 expressing baculovirus. The chaperone protein is purified by affinity chromatography through a 6-Histidine tag, followed by anion exchange chromatography on Source 15Q resin (Pharmacia).

Materials

Late-log-phase Sf9 cells cultured in suspension ($>2 \times 10^6$ cells/ml; also see UNITS 16.9 & 16.10)

High-titer His-NAP-1 baculovirus stock (Orbigen; also see UNITS 16.9 & 16.10)

Phosphate-buffered saline (PBS; APPENDIX 2), ice-cold

Lysis buffer H (see recipe)

Wash buffer H (see recipe)

Elution buffer H: wash buffer H (see recipe) containing 480 mM imidazole

HEGD buffer containing 0.1 M NaCl (see recipe)

NAP-1 purification buffer (see recipe) containing 0.0, 0.1, and 1.0 M NaCl

Ni-NTA agarose resin (Qiagen)

Bovine serum albumin (BSA) standard, 2 mg/ml (Pierce, Cat. No. 23209)

Source 15Q resin (Amersham Pharmacia Biotech)

20% (v/v) ethanol

8% and 15% SDS-PAGE gels (UNIT 10.2)

Liquid nitrogen

Clinical centrifuge with swinging-bucket rotor, 4°C

250-ml conical centrifuge bottles appropriate for clinical centrifuge

40-ml Wheaton Dounce homogenizer with "A" pestle

Sorvall Superspeed centrifuge with SS34 rotor (or equivalent)

15- and 50-ml conical tubes

End-over-end rotator

12,000 to 15,000 MWCO dialysis tubing

HR-5 or HR-10 FPLC column (Amersham Pharmacia Biotech)

FPLC apparatus

Siliconized 1.5 ml polypropylene tubes (e.g., ISC BioExpress, Cat. # C-3302-1)

Additional reagents and equipment for baculovirus culture and infection (UNITS 16.9-16.11), dialysis (APPENDIX 3C), and SDS-PAGE (UNIT 10.2)

Infect and harvest Sf9 cells

1. Grow Sf9 cells in 500 or 1000 ml spinner flasks to a density $>2.0 \times 10^6$ cells/ml in culture medium (UNIT 16.10). Dilute with the medium to 1.0×10^6 cells/ml. Infect with 25 ml amplified His-NAP-1 virus per 1 liter cell culture.
2. At 72 hr subsequent to infection, collect the cells by spinning 5 min at 2000 rpm in 250 ml conical tubes in a clinical centrifuge at 4°C. Resuspend each pellet in cold PBS (1/10 of the original culture volume). Repeat the centrifugation.

Exercise caution not to discard the cell pellet, which becomes quite loose after washing with PBS. From this point on, all steps should take place on ice or in a cold room, unless otherwise noted. All buffers should be at 4°C unless otherwise noted. Cell pellets can be frozen in liquid nitrogen and stored at -80°C for several weeks before further processing.

**Chromatin
Assembly and
Analysis**

21.7.11

3. Resuspend the cell pellet in lysis buffer H (1/40 original culture volume). Homogenize in a Dounce homogenizer using 40 strokes of the “A” pestle over 30 min. Centrifuge 10 min at $14,500 \times g$ (11,000 rpm in an SS-34 rotor), 4°C. Pool all supernatants in a 50-ml conical tube.

Purify NAP-1 by nickel affinity chromatography

4. Equilibrate 1 ml Ni-NTA agarose resin in lysis buffer H per 500 ml original cell culture volume. Add the cell extract and incubate for 3 to 4 hr on an end-over-end rotator. Pellet the resin by centrifuging 5 min at 2000 rpm in a clinical centrifuge, 4°C. Wash the resin twice with 100 ml lysis buffer H, then twice with 100 ml wash buffer H, resuspending the resin by inverting the tube after each addition of buffer, and pelleting the resin by centrifuging 3 min at 2000 rpm in the clinical centrifuge after each wash.
5. To elute protein, resuspend the resin in 2 ml elution buffer H by gentle vortexing. Incubate on ice for 5 min. Centrifuge 3 min at 2000 rpm in clinical centrifuge, then remove the supernatant to a fresh tube on ice. Repeat this elution cycle three more times, pooling the eluates.
6. Dialyze the eluted NAP-1 twice in 12,000 to 15,000 MWCO tubing, each time for 2 hr against 4 liters HEGD buffer containing 0.1 M NaCl.

General techniques for dialysis are described in APPENDIX 3C. This dialysis step removes phosphate from the buffer.

7. Dialyze for an additional 2 hr against 4 liters NAP-1 purification buffer containing 0.1 M NaCl.

Significant protein precipitation at this step is reduced by complete dialysis in step 6. The dialyzed protein can be frozen in liquid nitrogen and stored at -80°C before further processing.

8. Remove the precipitate by spinning in a 15-ml conical tube at $14,500 \times g$ (11,000 rpm in an SS-34 rotor), 4°C. Analyze the dialyzed NAP-1 by SDS-PAGE (UNIT 10.2) with BSA standard to estimate the amount of protein.

A 14-kDa contaminating Sf9 protein band will be apparent if analyzed on a 15% (or higher) gel. Although dNAP-1 is >95% pure after the affinity chromatography step, the contaminants appear to moderately inhibit chromatin assembly.

Purify NAP-1 to apparent homogeneity by anion-exchange chromatography

9. Using an FPLC, pack Source 15Q resin in an HR-5 or HR-10 column according to the manufacturer’s instructions. Use 1 ml of packed resin per 5 mg of NAP-1 from step 8. Equilibrate the Source 15Q column in 10 column volumes of NAP-1 purification buffer containing 0.1 M NaCl.
10. Load the NAP-1 onto the Source 15Q column. Wash the sample with 10 column volumes of NAP-1 purification buffer containing 0.2 M NaCl. Elute the protein with a 20-column-volume gradient of NAP-1 purification buffer from 0.2 M to 0.5 M NaCl.

NAP-1 should elute in two distinct peaks. The early (lower-salt) peak is inhibitory towards assembly while the later (higher salt) peak is active. The 14-kDa band (see step 8) elutes with the early (inhibitory) peak. Collect fractions of 0.25 to 0.5 column volumes.

11. Run 2 μl of each fraction on 15% SDS-PAGE gel (UNIT 10.2) to identify pure NAP-1-containing fractions. Combine peak fractions and dialyze twice (APPENDIX 3C), each time for 2 hr against 2 liters NAP-1 purification buffer containing 0.1 M NaCl.

A 15% gel will allow visualization of the 14-kDa contaminating band.

- Analyze the dialyzed NAP-1 on an 8% SDS-PAGE gel (UNIT 10.2) along with BSA standard to determine the concentration. Divide the material into 100- to 200- μ l aliquots in 1.5-ml siliconized tubes and freeze in liquid nitrogen. Store at -80°C .

The expected yield of the active NAP-1 fraction is 1 to 3 mg per 1 liter of the cell culture volume.

EXPRESSION AND PURIFICATION OF THE RECOMBINANT *DROSOPHILA* NAP-1 (NTA SUPERFLOW RESIN)

**ALTERNATE
PROTOCOL 1**

If culture volumes are large, it can be helpful to perform the affinity chromatography purification with NTA Superflow resin (Qiagen) on an FPLC, rather than using batch purification on Ni-NTA agarose.

Additional Materials (also see Basic Protocol 5)

NTA Superflow resin (Qiagen)

Superflow chromatography buffer (see recipe) containing 20 mM and 500 mM imidazole

C-10/10 or C-10/20 FPLC column (Amersham Pharmacia Biotech)

- Using an FPLC, pack an NTA Superflow column in 20% ethanol using 1 ml NTA Superflow resin per 500 ml of the original culture volume. Attach column to FPLC apparatus.
- Equilibrate the column with 5 column volumes of Superflow chromatography buffer containing 20 mM imidazole. Load the supernatant from Basic Protocol 5, step 3 onto the column.

If you plan to reuse the column, it may be helpful to briefly spin the extract in an ultracentrifuge 30 min at $80,000 \times g$ (25,000 rpm in an SW-28 rotor), 4°C , in order to reduce the particulate matter being injected on the FPLC column. The injection on the column should be very slow (overnight at 0.05 to 0.1 ml/min). If the flowthrough contains significant amounts of NAP-1, it can be used for an additional round of purification.

- Wash with 5 column volumes of Superflow chromatography buffer containing 20 mM imidazole. Elute the protein with a 10 column volume gradient from 20 mM to 500 mM imidazole in Superflow chromatography buffer.
- Run 5 μ l per fraction on a 15% SDS-PAGE gel (UNIT 10.2) and pool the protein peak. Continue as in Basic Protocol 5, step 6.

A 15% gel will allow visualization of the 14-kDa contaminating band.

CHROMATIN ASSEMBLY WITH PURIFIED RECOMBINANT *DROSOPHILA* FACTORS

**BASIC
PROTOCOL 6**

Core histones are preincubated on ice with NAP-1 to form histone-chaperone complexes. The chromatin assembly reaction is initiated at 27°C by addition of ACF, ATP, and plasmid DNA (circular supercoiled or relaxed, or linear). The extent of chromatin assembly can be monitored by analysis of circular DNA supercoiling, while the “quality” of chromatin is assayed by micrococcal nuclease digest assay.

Materials

HEG buffer (see recipe)

300 mM KCl (store in aliquots of 0.1 to 1 ml at -20°C)

PvOH/PEG solution (see recipe)

2 mg/ml BSA solution (store in aliquots of 0.1 to 1 ml at -20°C)

**Chromatin
Assembly and
Analysis**

21.7.13

0.5 to 4.0 mg/ml recombinant NAP-1 (see Basic Protocol 5)
 0.3 to 2.0 mg/ml purified *Drosophila* core histones, 0.3 to 2.0 mg/ml (see Basic Protocol 2)
 0.5 M ATP (store in aliquots of 0.1 to 1 ml at -20°C)
 0.5 M creatine phosphate (see recipe)
 5 mg/ml creatine kinase (see recipe)
 100 mM MgCl_2 (store in aliquots of 0.1 to 1 ml at -20°C)
 0.3 to 2.0 mg/ml plasmid DNA, double CsCl-purified (*UNIT 1.7*), in TE buffer (*APPENDIX 2*)
 10 \times topoisomerase I buffer (see recipe)
 Recombinant topoisomerase I working solution (see Support Protocol)
 0.002 to 0.2 mg/ml recombinant ACF (see Basic Protocol 4)
 ACF dilution buffer: wash buffer F (see recipe) containing 0.4 mg/ml recombinant human insulin (Roche)
 200 U/ml micrococcal nuclease stock solution (see recipe)
 Buffer R (see recipe)
 10 mM CaCl_2 (store in aliquots of 0.1 to 1 ml at -20°C)
 0.5 M EDTA
 10 mg/ml RNase A
 Glycogen stop buffer (see recipe)
 2.5 mg/ml proteinase K
 50:49:1 (v/v/v) phenol/chloroform/isoamyl alcohol, equilibrated with 10 mM Tris-Cl, pH 8.0 (see *UNIT 2.1A* for equilibration technique)
 2.5 M ammonium acetate
 100% ethanol

 Siliconized 1.5 ml polypropylene tubes (e.g., ISC BioExpress, Cat. # C-3302-1)
 27 $^{\circ}$ and 30 $^{\circ}\text{C}$ water baths

 Additional reagents and equipment for extraction of DNA (*UNIT 2.1A*) and agarose gel electrophoresis (*UNIT 2.5A*)

Prepare master solutions for chromatin assembly

1. Thaw all buffers and proteins.

It is recommended that all buffers be equilibrated to room temperature (for consistent pipetting). Proteins must be quick-thawed (in a room temperature water bath, mixed by tapping or extremely gentle vortexing, and transferred on ice) and quick-frozen (in liquid nitrogen) after use. ACF, NAP-1, core histones, and micrococcal nuclease (but not topoisomerase I or ISWI expressed and purified as an individual polypeptide) can withstand multiple freeze-thaw cycles (up to 10). Small working aliquots of topoisomerase I (which contains 50% glycerol) can be stored for up to 3 months at -20°C without freezing.

2. Prepare the master mix of NAP-1 and core histones (NH) by combining the following in a siliconized 1.5-ml microcentrifuge tube:

172 μl of HEG
 70 μl of 300 mM KCl
 84 μl of PvOH/PEG solution
 4.2 μl of 2 mg/ml BSA
 6.4 μl of 2 mg/ml NAP-1
 3.03 μl of 0.7 mg/ml core histones.

Vortex gently for 2 to 3 sec.

Use siliconized 1.5-ml tubes throughout the protocol. The provided recipe is calculated for six standard reactions, and should be used to perform five or fewer reactions (to allow for imprecise pipetting). A single reaction is calculated to contain 0.353333 μg DNA, 0.353333

μg core histones, 2.12 μg NAP-1, 10 units ACF (0.22 pmol), 1.4 μg BSA in 70 μl of a buffer, containing ~50 mM KCl, 3 mM ATP, and ~5 mM MgCl₂. The final volume of the reaction is made up to 70 μl with HEG. The NAP-1 to core histone mass ratio is 5:1 and should be sufficient to eliminate unbound histones in the reaction. PvOH/PEG and BSA are not obligatory—simply use or do not use them consistently in all assembly experiments. Preparing the master mix for several assembly reactions minimizes pipetting small volumes of core histones. Use a P2 Pipetman (Gilson) to pipet the core histones.

3. Pipet a 56.6-μl aliquot of the NH mix prepared in step 2 (at room temperature) into each of five siliconized 1.5-ml microcentrifuge tubes. Incubate on ice for ≥20 min to allow histone-NAP-1 binding.
4. Prepare the master mix of ATP and Mg²⁺ (AM) by combining 3 μl of 0.5 M ATP with 30 μl of 0.5 M creatine phosphate, 16.5 μl distilled water, and 25 μl 100 mM MgCl₂. Add 0.5 μl 5 mg/ml creatine kinase immediately before use.

The ATP regeneration system (creatine phosphate and creatine kinase) is optional (it is only required for assembly times >2 hr and high ACF concentrations). Keep the mix at room temperature at all times to avoid creatine phosphate precipitation. Discard the thawed aliquots of creatine kinase.

5. Prepare the DNA template by combining:

7.64 μl of plasmid DNA (at 0.42 mg/ml)
2 μl 10× topoisomerase I buffer
2.36 μl recombinant topoisomerase I working solution
8 μl distilled H₂O.

Relax the DNA for 10 min at 30°C; keep at room temperature until ready to use.

Chromatin inhibits DNA relaxation by topoisomerase I. Thus, topoisomerase I in the reaction should be in 5- to 10-fold excess over the amount that is necessary to completely relax the supercoiled plasmid after a 10-min incubation at 30°C. The purified system will efficiently assemble chromatin on linear supercoiled DNA in the absence of the topoisomerase. In the latter case, the template must contain more than 95% supercoiled DNA (achieved by two rounds of CsCl banding).

6. Prepare ACF dilution(s) in ACF dilution buffer (2 to 10 units/μl). Keep on ice.

1 unit of ACF equals 22 fmol of protein.

Assemble chromatin and analyze by micrococcal nuclease assay

7. Start 5 assembly reactions as follows:

- a. To 56.6 μl NH (step 3) add 1 μl ACF (2 to 10 units)
- b. Transfer from ice and equilibrate to room temperature.
- c. Add 10.5 μl AM master mix (step 4) and 2 μl DNA template (relaxed with excess topoisomerase I or supercoiled; step 5). Immediately vortex, gently, for 2 to 3 sec.
- d. Allow the assembly to proceed at 27°C for 1.5 to 2.5 hr.

Mix the reactions by gentle vortexing after adding each component. Microcentrifuge the tubes briefly after vortexing to collect the solution in the bottom of the tube. It is especially important to mix the reactions immediately upon the addition of the DNA template.

8. Immediately before use, prepare two dilutions of micrococcal nuclease in buffer R, 1:500 and 1:1500.

Note the difference in dilution ratios from step 3 of Basic Protocol 3. Micrococcal nuclease digests purified chromatin more efficiently in the absence of inhibitory components within the crude system.

9. Add 17.5 μl 10 mM CaCl_2 to each reaction. Divide each reaction in two equal parts (“a” and “b”). In a controlled manner (at certain time intervals, e.g., 15 sec), add 5 μl of the 1:1500 dilution prepared in step 8 to each “a” tube and 5 μl of the 1:500 dilution to each “b” tube. Allow the digestion to progress for 10 min at room temperature for every tube.

The assembly reaction can also be monitored by the DNA supercoiling assay. Stop one quarter of the 70 μl assembly reaction ($\sim 0.177 \mu\text{g}$ DNA in 17.5 μl) by addition of 3 μl 0.5 M EDTA. Deproteinase and precipitate the DNA as in step 11 below. Run, along with supercoiled and relaxed DNA samples, 1 kbp DNA ladder on a 0.8% agarose, 1 \times TBE gel until the xylene cyanol dye front reaches the bottom third of the gel. Stain and destain with ethidium bromide.

10. Prepare stop solution (ST) by mixing 55 μl of 0.5 M EDTA and 11 μl of 10 mg/ml RNaseA. Stop micrococcal digestions by adding 6 μl ST to each tube and then vortexing. Allow samples to stand 5 min at room temperature to digest contaminating RNA.

From this point on, the reactions do not have to be timed precisely.

11. Prepare proteinase solution (PR) by mixing 1.1 ml glycogen stop buffer and 55 μl of 2.5 mg/ml proteinase K solution. Add 105 μl PR to each tube, then vortex. Digest the histones and soluble proteins at 37°C for ≥ 30 min. Extract with 200 μl of 50:49:1 phenol/chloroform/isoamyl alcohol. Precipitate the DNA with 25 μl of 2.5 M ammonium acetate and 475 μl of 100% ethanol.

Do not wash the pellets with 70% ethanol. UNIT 2.1A describes general techniques for extraction and precipitation of DNA.

12. Perform agarose gel electrophoresis as described (see Basic Protocol 3, step 8, and UNIT 2.5A).

ALTERNATE PROTOCOL 2

TITRATION OF THE RATIO OF CORE HISTONES TO DNA IN THE RECOMBINANT CHROMATIN ASSEMBLY REACTION

For every new plasmid DNA and core histone preparation, the optimal core histone to DNA ratio should be determined experimentally.

For materials, see Basic Protocol 6

1. Proceed as in Basic Protocol 6, steps 1 to 6.

As an approximation for actual concentrations, use the DNA concentration as calculated from A_{260} and the histone concentration as determined by the BCA assay.

2. Start 5 assembly reactions as follows:

- a. To 56.6 μl NH (see Basic Protocol 6, step 3) add 1 μl ACF (2 to 10 units)
- b. Transfer from ice and equilibrate to room temperature.
- c. Add 10.5 μl AM master mix (see Basic Protocol 6, step 4) and 2.5, 2.22, 2.0, 1.82, or 1.67 μl DNA template (relaxed with excess topoisomerase I or supercoiled; see Basic Protocol 6, step 5). Immediately after adding DNA, vortex each tube gently for 2 to 3 sec.
- d. Allow the assembly to proceed at 27°C for 1.5 to 2.5 hr.

The assumed histone to DNA mass ratios in these reactions are 0.8:1, 0.9:1, 1:1, 1.1:1, and 1.2:1, respectively. Use a P2 Pipetman (Gilson) to pipet the DNA solution.

3. Proceed as in steps 8 to 12 of Basic Protocol 6.

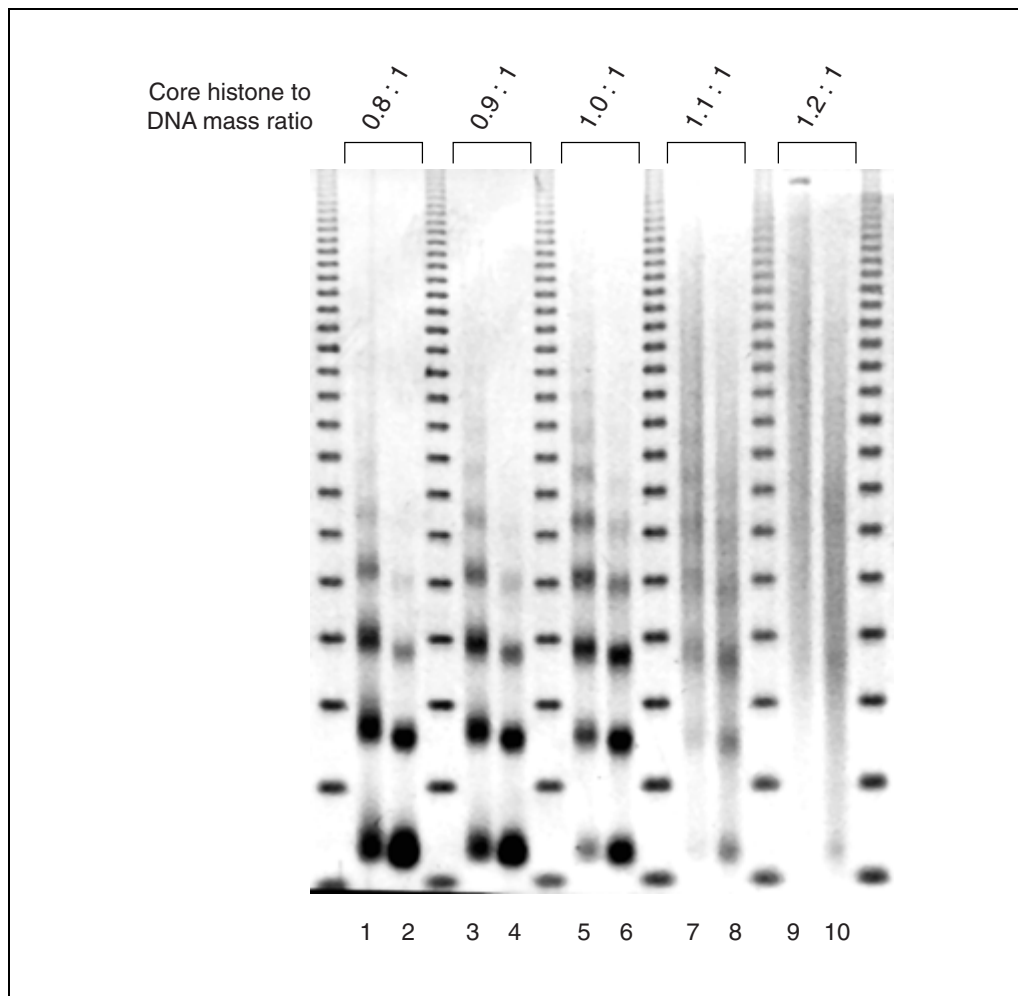


Figure 21.7.2 Chromatin assembly in the recombinant system. The reactions were performed as described in Alternate Protocol 2. The core histone to DNA ratios were initially estimated as: lanes 1, 2—0.8:1; lanes 3, 4—0.9:1; lanes 5, 6—1.0:1; lanes 7, 8—1.1:1; lanes 9, 10—1.2:1.

4. From the gel, determine the highest histone-to-DNA mass ratio that still results in an unsmeared ladder (lanes 5, 6 in Fig. 21.7.2, 1.0:1 estimated ratio).

The next lowest ratio (i.e., 0.9:1 as in lanes 3, 4 of Fig. 21.7.2) is postulated to be the optimal ratio for further chromatin assembly experiments. The effective concentration of the core histones is recalculated so that the new calculated optimal ratio is set to 1.0:1.

For instance, for the experiment in Figure 21.7.2, the original (estimated) histone concentration is multiplied 1.111 times. Thus the histone to DNA ratio in lanes 3, 4 calculated from the new (effective) histone concentration is 1.0:1.

EXPRESSION AND PURIFICATION OF THE CORE CATALYTIC DOMAIN OF THE *DROSOPHILA* TOPOISOMERASE I

The smallest active N-terminal truncation (Shaiu and Hsieh, 1998) of the *Drosophila* topoisomerase I was cloned into pET-28a expression vector (Novagen) from a cDNA that was kindly provided by Tao-shih Hsieh at Duke University. The coding sequence was subcloned into *NcoI/XhoI* restriction sites in frame with the C-terminal 6-His tag. The construct is referred to as pET-NDH6; the protein is referred to as ND423.

SUPPORT PROTOCOL

Chromatin Assembly and Analysis

21.7.17

Materials

Competent BL21(DE3) bacteria (Novagen)
PET-NDH6 plasmid (available on request from Dmitry Fyodorov;
fyodorov@ucsd.edu)
LB plates and medium (UNIT 1.1) containing 50 µg/ml kanamycin (add from 10
mg/ml kanamycin stock)
100 mM IPTG
Liquid nitrogen
Lysis buffer T (see recipe)
Ni-NTA resin (Qiagen)
Elution buffer T: lysis buffer T containing 0.5 M imidazole
Dialysis buffer T (see recipe)
Storage buffer T (see recipe)
8% SDS-PAGE gel (UNIT 10.2)
Bovine serum albumin (BSA) standard, 2 mg/ml (Pierce, Cat. No. 23209)
0.8% agarose gel (UNIT 2.5A)
10× topoisomerase I buffer (see recipe)

Sorvall Superspeed centrifuge with GSA and SS-34 rotors (or equivalents)
Microtip sonicator (e.g., Branson Sonifier 450; VWR Scientific)
10-ml polypropylene chromatography column
0.5 to 3.0 ml Slide-A-Lyzer, 10,000 MWCO (Pierce)

Additional reagents and equipment for transformation of bacteria (UNIT 1.8), growth
of bacteria in solid (UNIT 1.3) and liquid (UNIT 1.2) media, SDS-PAGE (UNIT 10.2),
and agarose gel electrophoresis (UNIT 2.5A)

Express ND423 in *E. coli*

1. Transform BL21(DE3) cells with pET-NDH6. Plate on LB plates containing 50 µg/ml kanamycin; incubate overnight at 37°C.

Transformation is performed immediately prior to the expression.

2. Pick one average size colony into 0.5 liter LB medium (containing 50 µg/ml kanamycin) and grow while shaking at 37°C for 6 to 8 hr. At a bacterial density equivalent to $A_{600} \sim 0.5$, induce by adding 100 mM IPTG to a final concentration of 0.42 mM.
3. Incubate while shaking at 30°C for 5 hr. Harvest the cells by centrifugation for 10 min at $8500 \times g$ (7000 rpm in a GSA rotor), 4°C.

Note the temperature change from 37° to 30°C in this incubation.

4. Freeze the cell pellets in liquid nitrogen.

The pellets can be stored at -80°C overnight, if desired.

Extract soluble bacterial proteins and isolate ND423 by affinity chromatography

5. Thaw and resuspend the cell pellets in 10 to 20 ml lysis buffer T (20 to 40 ml per 1 liter of original culture). Freeze again in liquid nitrogen and thaw in a water bath at room temperature.
6. Sonicate on ice with a microtip with 4 burst of 30 sec at a setting of “6.5.” Centrifuge the lysate in 40 ml polycarbonate tubes for 10 min at $19,000 \times g$ (16,000 rpm in an SS-34 rotor), 4°C.
7. Equilibrate 2 ml Ni-NTA resin with the lysis buffer T. Mix and incubate the cell lysate with the resin for 3 hr in the cold room on a rocking platform.

8. Load the resin onto a disposable 10-ml polypropylene column by gravity flow (in a cold room). Wash 3 times with 10 ml of cold lysis buffer T. Elute three times, each time with 1 ml elution buffer, discarding the void volume (the first 350 μ l).
9. Dialyze in a 10,000 MWCO Slide-A-Lyzer against 2 liters dialysis buffer T for 2 hr at 4°C. Dialyze again against 1 liter storage buffer T for 2 hr at 4°C, then freeze in 100- μ l aliquots in liquid nitrogen. Store at -80°C. Determine the concentration by visualizing on an 8% SDS-PAGE gel (UNIT 10.2) along with the BSA mass standard.

The typical yield is 1.5 to 2.0 mg protein in 1.0 to 1.2 ml.

Prepare the working stock of topoisomerase I and assay the enzymatic activity

10. Prepare the working solution by 100-fold dilution in storage buffer T containing 0.2 mg/ml human recombinant insulin (add from 50 mg/ml stock). Upon dilution, store in 100- μ l aliquots at -80°C. Keep one aliquot at -20°C for daily use.
11. Assay the activity of the protein by relaxing 0.5 μ g supercoiled plasmid DNA with different dilutions of the enzyme. Run on a 0.8% agarose gel; stain with ethidium bromide and destain (UNIT 2.5A).

1 μ l of the working stock should completely relax ~10 μ g of 3 kbp plasmid in a 100- μ l reaction in 1 \times topoisomerase buffer after 10 min incubation at 30°C.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

ATP mix

300 mM creatine phosphate
30 mM ATP
10 μ g/ml creatine kinase
Store up to 2 years at -80°C

Buffer A

15 mM Tris-Cl, pH 7.5 (APPENDIX 2)
15 mM NaCl
60 mM KCl
0.34 M sucrose
0.1% (v/v) 2-mercaptoethanol
Store solution with above components up to 24 hr at 4°C
0.5 mM spermidine (add immediately prior to use)
0.15 mM spermine (add immediately prior to use)
0.25 mM PMSF (add immediately prior to use)

Buffer B

To buffer A (see recipe above), add:
2 mM EDTA
0.5 mM EGTA
Store up to 24 hr at 4°C

Buffer R

10 mM potassium HEPES, pH 7.6
10 mM KCl
1.5 mM MgCl₂
0.5 mM EGTA

continued

10% (v/v) glycerol
Store solution with above components up to 24 hr at 4°C
10 mM β -glycerophosphate (add immediately prior to use)
1 mM DTT (add immediately prior to use)
0.2 mM PMSF (add immediately prior to use)
Aliquots of 0.1 to 1 ml can be stored at -20°C for up to 2 years.

Core histone storage buffer

10 mM potassium HEPES, pH 7.6
1 mM EDTA
10 mM KCl
10% (v/v) glycerol
Store solution with above components up to 24 hr at 4°C
1 mM DTT (add immediately prior to use)

Creatine kinase solution

5 mg/ml creatine kinase (Sigma-Aldrich)
10 mM potassium phosphate, pH 7.0 (*APPENDIX 2*)
50 mM NaCl
50% (v/v) glycerol
Store in 5- to 10- μ l aliquots up to 2 years at -80°C

Creatine phosphate, 0.5 M

106 mg/ml creatine phosphate (phosphocreatine)
20 mM potassium HEPES, pH 7.6
Adjust pH to 7.0
Store in aliquots of 0.1 to 1 ml up to 2 years at -20°C

Dialysis buffer T

HEG buffer (see recipe below) containing:
50 mM NaCl
0.01% NP-40
Store solution with above components up to 24 hr at 4°C
0.5 mM DTT (add immediately prior to use)
0.2 mM PMSF (add immediately prior to use)
0.5 mM benzamidine (add immediately prior to use)
5 mM Na₂S₂O₅ (add immediately prior to use)

Dilution buffer F

20 mM Tris-Cl, pH 7.9 (*APPENDIX 2*)
10% (v/v) glycerol
0.02% (v/v) NP-40
Store up to 24 hr at 4°C

Elution buffer F

Wash buffer F (see recipe below) containing:
0.4 mg/ml FLAG peptide (Sigma-Aldrich)
0.4 mg/ml recombinant human insulin (Roche)
Use immediately
Add FLAG peptide from 10 mg/ml stock in STE buffer and add insulin from 50 mg/ml stock in TE buffer (see *APPENDIX 2* for buffers).
Store stock solution of FLAG peptide in 20- μ l aliquots up to 2 years at -80°C. Store insulin stock solution up to 1 year at 4°C.

Glycogen stop buffer

20 mM EDTA
0.2 M NaCl
1% (w/v) SDS
0.25 mg/ml glycogen
Store up to 2 years at room temperature

HA chromatography buffer

40 mM sodium phosphate, pH 6.8 (*APPENDIX 2*)
0, 0.35, or 2.5 M NaCl
Store solution with above components up to 24 hr at 4°C
1 mM DTT (add immediately prior to use)
0.2 mM PMSF (add immediately prior to use)

HEG buffer

25 mM potassium HEPES, pH 7.6
0.1 mM EDTA
10% (v/v) glycerol
Store in aliquots of 0.1 to 1 ml up to 2 years at -20°C

HEGD buffer with 0.1 M NaCl

25 mM potassium HEPES, pH 7.6
1 mM EDTA
10% (v/v) glycerol
0.1 M NaCl
0.01% (v/v) NP-40
Store solution with above components up to 24 hr at 4°C
1 mM DTT (add immediately prior to use)
0.2 mM PMSF (add immediately prior to use)
10 mM β -glycerophosphate (add immediately prior to use)

Lysis buffer F

20 mM Tris-Cl, pH 7.9 (*APPENDIX 2*)
500 mM NaCl
20% (v/v) glycerol
4 mM MgCl₂
0.4 mM EDTA
Store solution with above components up to 24 hr at 4°C
2 mM DTT (add immediately prior to use)
20 mM β -glycerophosphate (add immediately prior to use)
0.4 mM PMSF (add immediately prior to use)
1 mM benzamidine hydrochloride (add immediately prior to use)
4 μ g/ml leupeptin (add immediately prior to use)
2 μ g/ml aprotinin (add immediately prior to use)

Lysis buffer H

50 mM sodium phosphate, pH 7.6 (*APPENDIX 2*)
0.5 M NaCl
15% (v/v) glycerol
20 mM imidazole
0.01% (v/v) NP-40
Store solution with above components up to 24 hr at 4°C
10 mM β -glycerophosphate (add immediately prior to use)
0.2 mM PMSF (add immediately prior to use)
0.5 mM benzamidine (add immediately prior to use)

Lysis buffer T

50 mM sodium phosphate, pH 7.0 (*APPENDIX 2*)
0.5 M NaCl
15% (v/v) glycerol
15 mM imidazole
0.1% (v/v) NP-40
Store solution with above components up to 24 hr at 4°C
0.2 mM PMSF (add immediately prior to use)
0.5 mM benzamidine (add immediately prior to use)
10 mM Na₂S₂O₅ (add immediately prior to use)

Micrococcal nuclease stock solution, 200 U/ml

1.56 mg/ml (200 U/ml) micrococcal nuclease (Sigma-Aldrich)
5 mM sodium phosphate, pH 7.0 (*APPENDIX 2*)
2.5 μM CaCl₂
Store in aliquots of 0.1 to 1 ml up to 1 year at -20°C

NAP-1 purification buffer

Buffer R (see recipe above) containing:
0.0, 0.1, or 1.0 M NaCl (add from 5 M NaCl stock or as solid NaCl)
0.01% NP-40 (add from 10% v/v stock)
Store up to 24 hr at 2°C

PvOH/PEG solution

HEG buffer (see recipe above) containing:
5% polyvinyl alcohol (mol. wt. 10,000, Sigma-Aldrich P-8136)
5% polyethylene glycol (mol. wt. 8,000, Sigma-Aldrich P-2139)
Store in aliquots of 0.1 to 1 ml up to 2 years at -20°C

Storage buffer T

10 mM potassium HEPES, pH 7.6
0.1 mM EDTA
50 mM NaCl
0.01% (v/v) NP-40
50% (v/v) glycerol
Store solution with above components up to 24 hr at 4°C
10 mM β-mercaptoethanol (add immediately prior to use)
0.2 mM PMSF (add immediately prior to use)
1 mM benzamidine (add immediately prior to use)
1 μg/ml leupeptin (add immediately prior to use)

Sucrose gradients, 5% to 30%

10 mM potassium HEPES, pH 7.6
1 mM EDTA
0.5 M NaCl
5% or 30% (w/v) sucrose
0.02% (w/v) NaN₃
Store solution with above components up to 1 year at 4°C
0.2 mM PMSF (add immediately prior to use)
Use a gradient maker to prepare 5% to 30% sucrose gradients in SW 28 ultracentrifuge tubes

Superflow chromatography buffer

50 mM sodium phosphate, pH 7.6 (*APPENDIX 2*)
500 mM NaCl
15% (v/v) glycerol
0.01% (v/v) NP-40
Store solution with above components up to 24 hr at 4°C
10 mM β -glycerophosphate (add immediately prior to use)
0.2 mM PMSF (add immediately prior to use)
0.5 mM benzamidine (add immediately prior to use)
2 μ g/ml leupeptin (add immediately prior to use)
2 μ g/ml aprotinin (add immediately prior to use)

Topoisomerase I buffer, 10 \times

0.5 M Tris·Cl, pH 7.5 (*APPENDIX 2*)
100 mM MgCl₂
1 mM EDTA
0.5 mg/ml BSA
5 mM DTT
Store in aliquots of 0.1 to 1 ml up to 2 years at -20°C

Wash buffer F

20 mM Tris·Cl, pH 7.9 (*APPENDIX 2*)
150 mM NaCl
15% (v/v) glycerol
2 mM MgCl₂
0.2 mM EDTA
0.01% (v/v) NP-40
Store solution with above components up to 24 hr at 4°C
1 mM DTT (add immediately prior to use)
10 mM β -glycerophosphate (add immediately prior to use)
0.2 mM PMSF (add immediately prior to use)
0.5 mM benzamidine-HCl (add immediately prior to use)
2 μ g/ml leupeptin (add immediately prior to use)
1 μ g/ml aprotinin (add immediately prior to use)

Wash buffer H

50 mM sodium phosphate, pH 7.6 (*APPENDIX 2*)
100 mM NaCl
20 mM imidazole
15% glycerol (v/v)
0.01% NP-40 (v/v)
Store solution with above components up to 24 hr at 4°C
10 mM β -glycerophosphate (add immediately prior to use)
0.2 mM PMSF (add immediately prior to use)
0.5 mM benzamidine (add immediately prior to use)

COMMENTARY

Background Information

Several procedures to assemble nucleosome arrays on circular or linear DNA in vitro have been developed (reviewed in Ito et al., 1997a). Some of these procedures employ a histone transfer process with a histone transfer vehicle (histone-binding proteins, polyglutamic acid,

RNA, or high salt concentration; *UNIT 21.6*). They yield well-defined templates, yet the nucleosomes are irregularly distributed and lack the periodicity of bulk native chromatin. The second type of these procedures utilizes crude cell extracts and yields extended periodic arrays of nucleosomes. This reaction was

achieved with cytosolic extracts from *Xenopus* oocytes (Glikin et al., 1984), HeLa cells (Banerjee and Cantor, 1990), or *Drosophila* embryos (Becker and Wu, 1992; Bulger and Kadonaga, 1994). The third type of chromatin assembly procedure is the assembly during viral DNA replication (Stillman, 1986). In this reaction, crude cell extracts (from human 293 cells) are complemented by SV40 T antigen helicase and histone chaperones CAF-1 and RCAF (Tyler et al., 1999) to achieve chromatin assembly of the newly replicated SV40 DNA.

The early realization that chromatin assembly in crude cell extracts is an ATP-dependent process has triggered the quest to purify individual protein factors that could mediate this reaction. Fractionation of *Drosophila* extracts resulted in purification of two essential assembly components, a histone chaperone NAP-1 (Bulger et al., 1995) and a SWI/SNF-like chromatin remodeling factor, ACF (Ito et al., 1997b). This work was supported by the NIH grant GM58272 to James T. Kadonaga at University of California, San Diego. ACF consists of two subunits, a SWI2/SNF2-related ATPase ISWI (p140) and a subunit termed Acf1 (p170 and p185). The chromatin assembly reaction can be reconstituted in a purified recombinant system that contains plasmid DNA, purified native or recombinant histones, recombinant NAP-1, recombinant ACF, and ATP (Ito et al., 1999; Levenstein and Kadonaga, 2002). This system combines the advantages of histone transfer procedures (well-defined reaction products) with the advantages of ATP-dependent crude extract assembly procedures (nucleosome array regularity similar to that of the native bulk chromatin).

The pair of cofactors (an ATP-dependent remodeling factor and a histone chaperone) is likely to represent the minimal required reaction system that assembles physiologically relevant chromatin. The histone chaperone NAP-1 is necessary to prevent rapid nonspecific histone-DNA aggregation and to present histones to ACF. Consistent with its role as a molecular sink that sequesters positively charged histones, NAP-1 has to be present at 1:1 or higher stoichiometry with respect to core histone polypeptides. Although NAP-1 has a higher affinity to H2A/H2B dimers, it also binds to H3/H4 tetramers and can facilitate the assembly of nucleosomes from dimers and tetramers. On the other hand, ACF performs multiple functions: core histone deposition on the DNA, establishing proper histone-DNA contacts in the nucleosome, and spacing nu-

cleosome arrays. All these tasks are carried out catalytically (substoichiometric relative to the core histones) and at the expenditure of ATP hydrolysis by ISWI.

Although histone chaperones (and specifically NAP-1) have been reported to mediate nucleosome assembly with supercoiled plasmid DNA, it is important to note that they do not deposit histones efficiently on relaxed DNA templates. Furthermore, electron microscopy of “chromatin” assembled by NAP-1 alone reveals that it is composed of large histone-DNA particles that are not canonical nucleosomes (Nakagawa et al., 2001). Finally, the replication-dependent chromatin assembly of the SV40 DNA by CAF-1 and RCAF histone chaperones may also be mediated by an ATP-dependent chromatin remodeling complex that is present in the 293 cell extract.

Critical Parameters and Troubleshooting

S-190 extract

A concentrated extract, stemming from a high ratio of embryos to buffer R, tends to be a highly active extract. Aspirating buffers too cautiously to prevent the loss of embryo material may result in a dilute extract. This can lead to low quality of assembly or require the S-190 to assume a significant portion of the reaction volume. Homogenizing embryos in a Dounce homogenizer generates sizeable resistance on down strokes as well as considerable back-pressure on up strokes. Be careful not to lose grip of the pestle during homogenization, or else the vessel may break and the embryos may be lost in the ice bucket. Be certain to use full strokes in order to homogenize all embryos. Between the first and second spins in the ultracentrifuge, the S-190 extract should be frozen in liquid nitrogen and then thawed in a room-temperature water bath. This step presumably precipitates unwanted materials and improves the activity of the assembly extract.

Core histone purification

When isolating the nuclei, the yellow yolk protein is undesirable. It will form a tight pellet at the bottom of the tube. A Pasteur pipet can be used to carefully separate the loose nuclear pellet from the yolk before resuspending. The addition of 5 M NaCl to the nuclear suspension (step 13) may not result in obvious lysis. Continue to the next step under the assumption that the lysis was successful. When selecting fractions from the sucrose gradient, avoid those

containing histone H1. H1 will be a significant band at ~40 kDa. Dialyzing the sucrose gradient fractions will result in significant swelling within the dialysis tubing. Take this into consideration when cutting the tubing size and consider using double clips to avoid mishaps. Hydroxylapatite resin crushes easily and needs to be run on the FPLC at low pressure. If the core histone volume is large, this is a good opportunity to inject the sample overnight at a very slow flow rate (0.1 ml/min or less). As an alternative to crystalline hydroxylapatite, ceramic hydroxylapatite resin is now available from Bio-Rad Laboratories (CHT Type I Ceramic Hydroxylapatite). This resin performs well at higher pressures and yields otherwise equivalent results.

S-190 chromatin assembly

Three key parameters tend to influence assembly with the S-190 extract: (1) the proportion of the extract in the reaction volume, (2) the Mg^{2+} concentration, and (3) the ratio of core histones to DNA. Variations among extracts result in a range of adequate S-190 proportions, typically falling between 25% to 40% of the reaction volume. The most desirable percentage should be determined for each extract. The optimal Mg^{2+} concentration lies between 3.5 and 7 mM. This should also be determined for each extract. The ratio of histones to DNA is affected by the volume of the S-190 used. The S-190 extract contains a residual pool of core histones (in the realm of 20% to 30% required for assembly). Typically, a 0.8:1 mass ratio yields high quality chromatin (as assayed by partial micrococcal nuclease digestion), but this ratio can be examined from 0.6:1 to 1.2:1. If low-quality agarose is used for gel electrophoresis, ladders can appear “smeared.” SeaKem LE agarose (BioWhittaker) is suggested. Complete destaining can significantly improve the appearance of nucleosomal ladders, especially those at the top of the ladder.

Purification of recombinant ACF and NAP-1

Lysing of the Sf9 cells must be thorough and complete. Low yield is likely to be a direct result of incomplete lysis. Sonication can be used in place of homogenization on the Dounce homogenizer to disrupt Sf9 cells (2 to 3 bursts of 30 sec with a microtip, on ice). The resulting extract will be cruder, and thus the affinity resins should be washed more thoroughly after the protein binding step.

Typically, purified ACF will contain equimolar amounts of p185 (Acf1-FLAG) and p140 (ISWI). However, ACF complexes with substoichiometric ISWI will also be suitable for most chromatin assembly applications. Calculate ACF molar concentration based on the concentration of the ISWI subunit. After affinity chromatography, the dNAP-1 protein is often more than 95% pure but performs poorly in assembly reactions. At this stage, it exists as a mixture of pure active NAP-1 multimers and an inhibitory NAP-1 fraction, which is contaminated with a 14-kDa Sf9 protein. It is therefore critical to perform the anion-exchange chromatography to separate active NAP-1 from this inhibitory species.

Recombinant chromatin assembly

The purified system performance is relatively harder to fine-tune than the S-190 based system. While the success in the latter is determined primarily by the quality of the extract, the purified system requires more extensive optimization of reaction parameters to achieve the best performance. The single most critical parameter in the reaction is the ratio of core histones and template DNA. Variations as low as 5% to 10% can inhibit the reaction or result in lower-quality chromatin as evidenced by the micrococcal digest analysis. Standard methods to determine DNA and protein concentrations (spectrophotometric, colorimetric) are almost never accurate enough to establish the correct histone-to-DNA ratio. Thus, it has to be determined experimentally for every pair of DNA and histone preparations. Alternate Protocol 2 provides an idea of such a titration experiment. The range and the step of this titration can be expanded, if necessary. In general, low histone content results in short ladders (2 or 3 apparent nucleosome bands at the bottom) and large nucleosome repeat (180 bp and longer). Excessive histone concentration will completely inhibit the reaction and result in histone-DNA aggregates that will not be efficiently digested by the micrococcal nuclease (see Fig. 21.7.2).

The assembly reaction is much less sensitive to the concentration of NAP-1. It should be noted, however, that it is safer to add an excess of the chaperone. If the mass ratio of NAP-1 to core histones drops below 3.8 to 4, the extra free histones that are present in the solution can form histone-DNA aggregates, which cannot serve as a substrate for ACF. The assembly reaction is tolerant to over an order of magnitude variation in ACF concentration. ACF will

assemble chromatin within a range of concentrations from 0.05 to 1 ACF protomers per 1 kbp DNA. It is recommended that the reaction be first optimized for a short plasmid template (3 kbp or less) before attempting to assemble chromatin on larger templates.

Chromatin may be purified after the assembly by gel filtration or sucrose gradient centrifugation. However, unpurified chromatin templates perform equally well in transcription assays. Several components of the assembly reaction are semi-optional. They include PvOH/PEG, BSA, ATP regeneration system, and topoisomerase I (if the DNA template is predominantly supercoiled). The reactions can be formulated and optimized accordingly (the changes will alter other reaction parameters, such as the optimal histone to DNA and NAP-1 to histone ratios). The suggested reaction conditions and micrococcal dilutions will produce nucleosome ladders with an apparent repeat length of ~165 bp. The nucleosomes can be packed more loosely or more tightly (down to 145 bp repeat length) by varying KCl concentration and the amount of histones in the assembly reaction. Tighter packed nucleosomes will require higher micrococcal concentration to be comparably digested. The purified system appears insensitive to histone modification states: it assembles chromatin equally well with purified native or unmodified bacterially expressed *Drosophila* core histones (Levenstein and Kadonaga, 2002). Moreover, the system does not exhibit specificity towards *Drosophila* core histones and can assemble chromatin with histones purified from other species.

Anticipated Results

Both ATP-dependent chromatin assembly systems described here can provide chromatin substrates for use in studies of chromatin structure, transcription, and DNA metabolism. Micrococcal nuclease analysis of the reaction products should reveal highly periodic nucleosome arrays. A DNA-binding transcription factor, when added to the assembly reaction, can position nucleosomes with respect to its binding sites. The number of negative supercoils in the assembled DNA (for circular plasmid DNA templates) can be estimated by one- or two-dimensional supercoiling analyses and may be used as a measure of the nucleosome content of the assembled minichromosomes. The assembled chromatin should migrate differently and can be separated from the unassembled DNA on sucrose gradients.

Time Considerations

Allow 2 days each for core histone purification and preparation of the S-190 extract. S-190 mediated chromatin assembly is extremely reproducible and does not require extensive troubleshooting. Typically, chromatin is assembled, then digested with micrococcal nuclease (6 to 7 hr) and analyzed by electrophoresis on the next day (3 to 4 hr). Allow 3 to 5 days for baculovirus amplification and 2 to 3 days for infection of Sf9 cells before harvesting. The recombinant ACF prep takes 1 day, and the NAP-1 prep takes 2 days. Recombinant chromatin assembly may require several pilot experiments to optimize concentrations of the components. Many researchers have reported that the quality of chromatin assembly in the recombinant system improves significantly after several trials.

Acknowledgments

The studies of the *Drosophila* chromatin assembly factors were supported by the NIH grant GM58272 to James T. Kadonaga. We thank Dr. Kadonaga for critical reading of the manuscript.

Literature Cited

- Banerjee, S. and Cantor, C.R. 1990. Nucleosome assembly of simian virus 40 DNA in a mammalian cell extract. *Mol. Cell. Biol.* 10:2863-2873.
- Becker, P.B. and Wu, C. 1992. Cell-free system for assembly of transcriptionally repressed chromatin from *Drosophila* embryos. *Mol. Cell. Biol.* 12:2241-2249.
- Bulger, M. and Kadonaga, J.T. 1994. Biochemical reconstitution of chromatin with physiological nucleosome spacing. *Methods Mol. Genet.* 5:241-262.
- Bulger, M., Ito, T., Kamakaka, R.T., and Kadonaga, J.T. 1995. Assembly of regularly-spaced nucleosome arrays by dCAF-1 and a 56 kDa histone-binding protein. *Proc. Natl. Acad. Sci. U.S.A.* 92:11726-11730.
- Glikin, G.C., Ruberti, I., and Worcel, A. 1984. Chromatin assembly in *Xenopus* oocytes: In vitro studies. *Cell* 37:33-41.
- Ito, T., Tyler, J.K., and Kadonaga, J.T. 1997a. Chromatin assembly factors: A dual function in nucleosome formation and mobilization? *Genes Cells* 2:593-600.
- Ito, T., Bulger, M., Pazin, M.J., Kobayashi, R., and Kadonaga, J.T. 1997b. ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell* 90:145-155.
- Ito, T., Levenstein, M.E., Fyodorov, D.V., Kutach, A.K., Kobayashi, R., and Kadonaga, J.T. 1999. ACF consists of two subunits, Acl1 and ISWI, that function cooperatively in the ATP-depend-

- ent catalysis of chromatin assembly. *Genes & Dev.* 13:1529-1539.
- Levenstein, M.E. and Kadonaga, J.T. 2002. Biochemical analysis of chromatin containing recombinant *Drosophila* core histones. *Submitted*.
- Nakagawa, T., Bulger, M., Muramatsu, M., and Ito, T. 2001. Multistep chromatin assembly on supercoiled plasmid DNA by nucleosome assembly protein-1 and ATP-utilizing chromatin assembly and remodeling factor. *J. Biol. Chem.* 276:27384-27391.
- Shaiu, W.L. and Hsieh, T.S. 1998. Targeting to transcriptionally active loci by the hydrophilic N-terminal domain of *Drosophila* DNA Topoisomerase I. *Mol. Cell. Biol.* 18:4358-4367.
- Stillman, B. 1986. Chromatin assembly during SV40 DNA replication in vitro. *Cell* 45:555-565.
- Tyler, J.K., Adams, C.R., Chen, S.-R., Kobayashi, R., Kamakaka, R.T., and Kadonaga, J.T. 1999. The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature* 402:555-560.

Contributed by Dmitry V. Fyodorov and
Mark E. Levenstein
University of California, San Diego
La Jolla, California

Analysis of Protein Co-Occupancy by Quantitative Sequential Chromatin Immunoprecipitation

UNIT 21.8

**BASIC
PROTOCOL**

Chromatin immunoprecipitation (ChIP; UNIT 21.3) is a widely used and powerful method for assaying individual protein-DNA interactions *in vivo*. As discussed earlier in this chapter (UNIT 21.3), ChIP experiments provide quantitative information about the relative level of binding of a given protein or proteins to different genomic regions. However, what standard ChIP experiments do not address is whether two proteins can simultaneously co-occupy a specific genomic region. For example, with conventional ChIP, the observation that two proteins associate with a given genomic region might reflect co-occupancy, or, alternatively, it might indicate that the two proteins associate with different populations of DNA molecules in a mutually exclusive fashion. There is simply no way to determine the extent to which (either positively or negatively) any two proteins influence one another when it comes to binding to the specific fragment in question via conventional ChIP.

Sequential chromatin immunoprecipitation (SeqChIP), on the other hand, is a powerful tool to address precisely this type of question: do two proteins simultaneously co-occupy a specific DNA region *in vivo*? In SeqChIP, protein-DNA complexes from the first ChIP are subjected to an additional immunoprecipitation with an antibody of a different specificity, typically to a second protein that is tested for its ability to co-occupy with the initially immunoprecipitated protein. The cross-links of these doubly immunoprecipitated protein-DNA complexes are then reversed, and the DNAs are analyzed by quantitative PCR in a manner analogous to conventional ChIP samples (see UNIT 21.3, Alternate Protocol 2). Occupancy values for both single and double immunoprecipitations are first calculated and then compared for any enrichment brought about by the second immunoprecipitation relative to the singly immunoprecipitated sample. While the following protocol has been successfully developed for and used with *Saccharomyces cerevisiae* samples, it is sufficiently general that it should be applicable for other species with minor modifications.

Materials

100 mg/ml bovine serum albumin (BSA, Fraction V; Sigma) in water (store at -20°C)

500 $\mu\text{g/ml}$ λ phage DNA (not sheared; New England Biolabs)

10 mg/ml *E. coli* tRNA in water (store at -20°C)

20 mg/ml glycogen, or Pellet Paint (Novagen) as DNA carrier

Additional reagents and equipment for ChIP (UNIT 21.3, Basic Protocol), growth of *Saccharomyces cerevisiae* (UNITS 13.1 & 13.2), extraction and purification of DNA (UNIT 2.1A), and real-time quantitative PCR (UNIT 21.3, Alternate Protocol 2)

Cross-link protein-DNA complexes *in vivo*

1. For each sample, grow 400 ml *Saccharomyces cerevisiae* to $\text{OD}_{600} = 0.6$ to 0.8 (UNITS 13.1 & 13.2).

It is necessary to use more extract for sequential immunoprecipitations (80 to 100 ml cells) than for conventional samples (20 to 40 ml cells) in order to ensure that enough material remains after the first immunoprecipitation to generate a reproducible signal following the second immunoprecipitation. The actual amount of extract required for this purpose depends on a number of factors, including cross-linking efficiency and protein

**Chromatin
Assembly and
Analysis**

21.8.1

abundance, and may have to be adjusted accordingly. This protocol doubles the culture volume (400 ml) used as compared to the single ChIP protocol (200 ml); consequently, it is helpful to split the cultures into the equivalent of two 200-ml samples for the cell lysis and sonication steps.

2. Add 11 ml of 37% formaldehyde (1% final). Cross-link for 20 min at room temperature by occasionally swirling flask or shaking slowly on a platform.
3. Add 60 ml heat-sterilized 2.5 M glycine and incubate an additional 5 min at room temperature.

Glycine stops the cross-linking by reacting with formaldehyde.

4. Split the cultures into two equal 200-ml samples and harvest cells, isolate chromatin, and shear the DNA fragments exactly as described in UNIT 21.3, Basic Protocol, steps 4 to 13.
5. Microcentrifuge 30 min at maximum speed, 4°C. Transfer the supernatants into a fresh 15-ml disposable conical tube, adjust the total sample volume to ~4.5 ml with ice-cold FA lysis buffer, and gently mix by inversion. Remove 250 µl to check DNA fragment size (see UNIT 21.3, Basic Protocol, steps 15 and 16) and freeze the remaining chromatin solution in 1-ml aliquots in liquid nitrogen up to 1 year.

Perform first immunoprecipitation

6. Incubate 1 ml chromatin solution with 20 µl primary antibody against the protein or epitope of interest and 20 µl of 50% (v/v) protein A–Sepharose beads in TBS, on an end-over-end rotator for 90 min at room temperature.

The amount of antibody used is double that in the Basic Protocol of UNIT 21.3, to reflect the increased amount of protein in the SeqChIP extract. As with conventional ChIP, the optimal amount of antibody will vary with each antigen and has to be empirically determined.

Sequential immunoprecipitations should normally be performed in both the forward and reverse directions to be able to analyze results in a quantitative fashion and to unambiguously determine the extent of co-occupancy between two proteins (see Anticipated Results).

7. Microcentrifuge beads 1 min at 3000 rpm, room temperature. Transfer 100 µl supernatant into a 0.5-ml PCR tube labeled “INPUT.” Discard the rest of the liquid.
8. Resuspend beads by pipetting up and down in 700 µl FA lysis buffer, room temperature, and transfer mixture into a Spin-X centrifuge-tube filter.

The use of Spin-X filters aids in the recovery of the beads after washes and results in better uniformity between different samples. The procedure is also substantially faster with the filters, particularly when multiple samples are processed simultaneously. Alternatively, one could use conventional microcentrifuge tubes for the washes and aspirate the supernatant with a narrow-bore pipet tip after each spin.

9. Place the filter into a 1.5-ml microcentrifuge tube and mix sample 3 min on an end-over-end rotator. Microcentrifuge 2 min at 3000 rpm, room temperature. Discard the flowthrough liquid at the bottom of the tube.
10. Add 700 µl FA lysis buffer, room temperature, to the beads and repeat step 9 two additional times.

Elute protein from beads

11. Wash beads for 3 min each with 700 µl FA lysis buffer/0.5 M NaCl, 700 µl ChIP wash buffer, and finally 700 µl TE buffer.

For many polyclonal antibodies, the more stringent washes in this step result in a cleaner signal, while gentle washes frequently lead to an unacceptably high background. For some

antibodies (e.g., monoclonal against peptide epitopes; see UNIT 21.3, Alternate Protocol 1), repeated washes with FA lysis buffer, which are gentler, might be more appropriate. Consult UNIT 21.3 for more information.

- Place filter unit containing the beads into a new 1.5-ml microcentrifuge tube and add 100 μ l of ChIP elution buffer. Gently pipet up and down two or three times to dislodge beads from the filter. Incubate 10 min in a 65°C water bath.

A water bath is used instead of other heating apparatuses to improve heat transfer.

- Microcentrifuge beads 2 min at 3000 rpm, room temperature. Discard filter with beads. Transfer 9 μ l (out of 100 μ l) of the eluate into a 0.5-ml PCR tube labeled “1st IP.”

Perform second immunoprecipitation

- Add 775 μ l FA lysis buffer (without SDS), 50 μ l of 100 mg/ml BSA, 50 μ l of 500 μ g/ml λ phage DNA, 5 μ l of 10 mg/ml *E. coli* tRNA, 10 to 20 μ l antibody, and 20 μ l of 50% protein A–Sepharose to the remaining 91 μ l of eluate obtained in the step 13. Incubate on an end-over-end rotator for 90 min at room temperature.

*In case peptide elution was performed (see Alternate Protocol 1 in UNIT 21.3), substitute FA lysis buffer containing 0.1% SDS for the non-SDS-containing buffer recommended above. It is important to keep overall SDS concentration to $\leq 0.1\%$. BSA, λ phage DNA, and *E. coli* tRNA are carriers designed to mimic the protein/nucleic acid mixture in chromatin samples and to prevent nonspecific precipitation of cross-linked protein-DNA complexes. They are also necessary for the first and second immunoprecipitations to be experimentally equivalent (see Critical Parameters and Troubleshooting). Finally, as in the case of the first immunoprecipitation, the amount of antibody required to produce optimal signal with the lowest background may have to be empirically determined.*

- Wash beads and elute protein complexes by following steps 8 to 13 (above). Transfer eluate into a 0.5-ml PCR tube labeled “SeqChIP,” taking note of the order of IPs.

For gentle washes (i.e., peptide elution), follow steps of UNIT 21.3, Alternate Protocol 1.

Reverse cross-links and purify DNA

- Add 100 μ l ChIP elution buffer to the tube marked “INPUT” and “1st IP;” add 100 μ l and 91 μ l TE to the tubes marked “SeqChIP” and “1st IP,” respectively, and 20 μ l Pronase in TBS to all three tubes.

If either the first IP and/or second IP was eluted by peptide, add 100 μ l ChIP elution buffer instead of 100 μ l TE to the relevant sample(s).

- To reverse cross-links, place tubes into a PCR machine. Incubate 2 hr at 42°C, followed by 6 hr at 65°C. Store samples at 4°C until used.

The incubation at 42°C allows for Pronase digestion of cross-linked polypeptides, while the 65°C-incubation results in a reversal of the formaldehyde cross-links.

The samples may be stored up to several days at 4°C.

- Transfer samples to new 1.5-ml microcentrifuge tubes, add 20 μ l of 4 M LiCl, and purify by extracting with 25:24:1 phenol/chloroform/isoamyl alcohol, followed by extraction with chloroform (UNIT 2.1A). Add 2 μ l of 20 mg/ml glycogen (or Pellet Paint; Novagen) as carrier, and 2.5 vol ethanol. Vortex briefly and precipitate the DNA for 1 hr at -20°C (see UNIT 2.1A).

- Microcentrifuge 30 min at maximum speed, 4°C. Discard aqueous phase and wash with 1 ml of 70% ethanol, room temperature, by gently inverting samples several times.

- Microcentrifuge 5 min at maximum speed, room temperature. Aspirate supernatant and air dry pellet for ≤ 5 min.

21. Resuspend in 150 μ l TE buffer and store at -20°C .

DNA stored in this fashion should be stable for years.

22. Analyze samples with real-time quantitative PCR (UNIT 21.3).

As signals from sequential immunoprecipitations are considerably weaker than from conventional ChIPs, it is recommended that the total number of amplification cycles be increased from 35 to 40. In some extreme circumstances where the amount of DNA to be amplified is minute, it might be beneficial to go beyond 40 cycles. Other parameters should be identical to those described in UNIT 21.3.

COMMENTARY

Background Information

Sequential chromatin immunoprecipitation, a natural extension of conventional ChIP, arose due to the inability of conventional ChIP to determine whether two different proteins can simultaneously co-occupy a given stretch of DNA *in vivo*. As described in this unit, the technique simply involves an additional immunoprecipitation with antibody to an unrelated protein, with subsequent steps essentially identical to conventional ChIP (UNIT 21.3). The resulting material can then be analyzed by either conventional (UNIT 21.3) or real-time QPCR (UNIT 21.3, Alternate Protocol 2), with real-time QPCR clearly being the preferred method due to its superior linear range, which permits simultaneous analysis of both singly and doubly immunoprecipitated samples.

SeqChIP has been successfully applied in the analysis of wide-ranging biological phenomena in a number of different organisms and cell types. For example, SeqChIP was used to help elucidate the mechanism of PPRE-dependent RXR homodimer signaling (Ijpenberg et al., 2004), to demonstrate TBP binding to nucleosomal DNA (Soutoglou and Talianidis, 2002), to identify the various protein complexes involved in mediating transcriptional activation by the estrogen receptor alpha (Metivier et al., 2003), and to study hyperosmotic stress response at the transcriptional level in yeast (Proft and Struhl, 2002). The procedure outlined in the Basic Protocol, with some modifications, can be readily adapted for use on non-*Saccharomyces cerevisiae* chromatin.

Critical Parameters and Troubleshooting

Sequential IP analysis requires substantially greater quantities of starting material (extract) than normally used for single ChIP experiments. This is due to a number of factors, including low cross-linking efficiencies (typically $<5\%$, depending on both protein and

DNA region) and the necessity to do two IPs instead of one. The problem becomes especially acute if the proteins in question are poorly cross-linked to DNA, are low in abundance, and exhibit no or partial co-occupancy. In the Basic Protocol, the amount of extract (and the corresponding purified DNA) used is four- to five-fold greater than typically used for single IPs, and is likely to be sufficient for most SeqChIP applications in yeast. Additional scale-up may be desirable in some specific instances.

Of special concern is the need to ensure that the second IP behaves identically to the first IP from a quantitative standpoint. The addition of BSA, *E. coli* tRNA, and λ phage DNA to the eluate from the first ChIP serves to ensure that the two IPs behave in a similar manner irrespective of actual order. BSA, *E. coli* tRNA, and λ phage DNA mimic the concentrated cross-linked *S. cerevisiae* chromatin that is present in the sample prior to the first IP, but which is effectively absent from the post-elution complexes. Alternatively, cross-linked chromatin from an unrelated yeast such as *Kluyveromyces lactis*, or even *E. coli* chromatin, can be used as a carrier in place of the combination of BSA/tRNA/ λ DNA. An added benefit of using non-*S. cerevisiae* yeast (or *E. coli*) chromatin is that these organisms can be engineered to serve as internal controls for the efficiency of the second immunoprecipitation. However, the actual strain construction and preparation of cross-linked chromatin from these organisms is time-consuming and provides no appreciable benefit (relative to BSA/tRNA/ λ DNA) when it comes controlling variability between 1st and 2nd IPs.

Anticipated Results

Predicted outcomes of SeqChIP experiments

SeqChIP experiments have three possible outcomes: complete co-occupancy, no co-occupancy, and partial co-occupancy (Fig. 21.8.1). Complete co-occupancy describes the scenario when two proteins, A and B, are

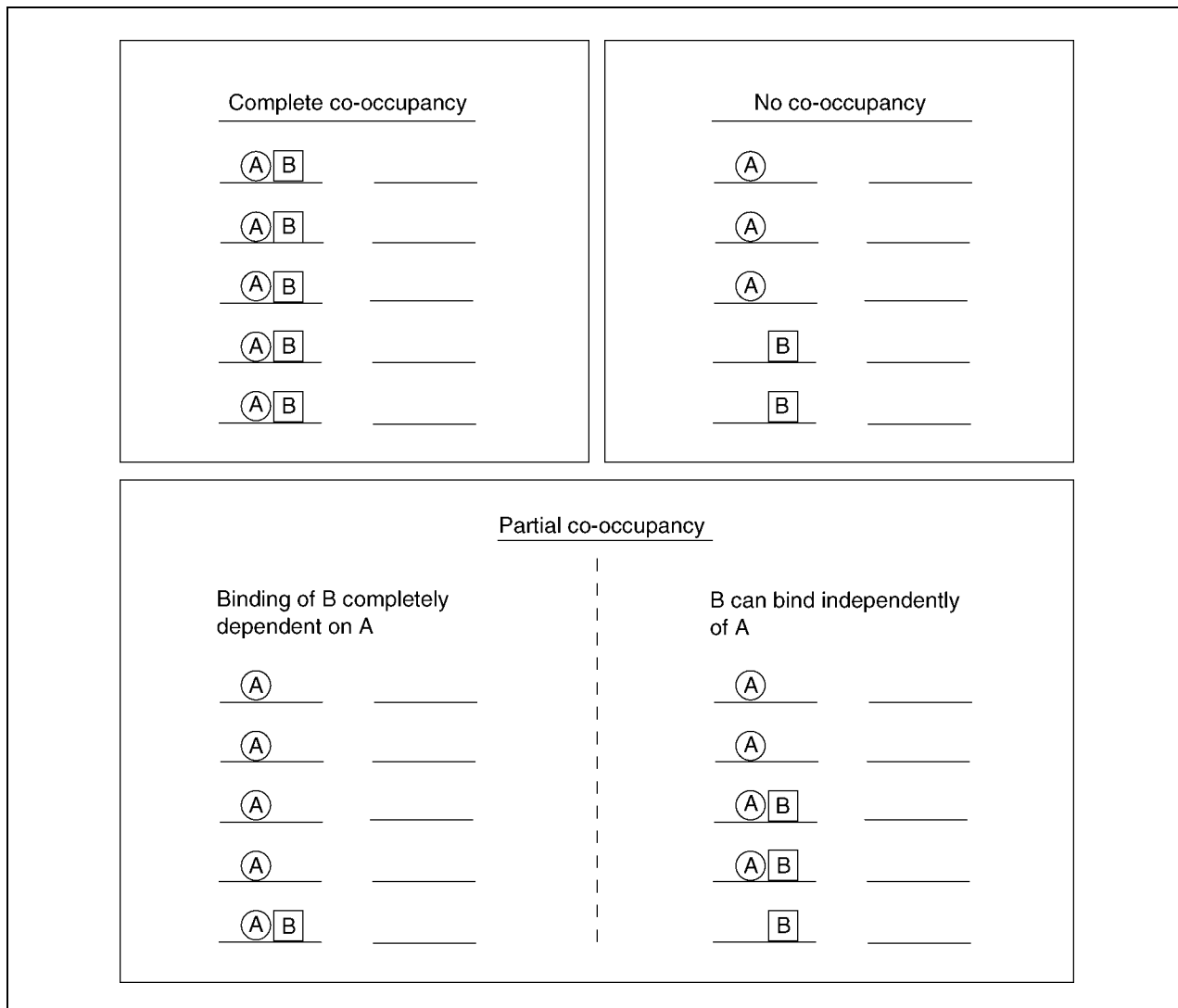


Figure 21.8.1 Schematic depiction of the three possible outcomes of a SeqChIP between proteins A and B: complete co-occupancy (top left), no co-occupancy (top right), and partial co-occupancy (bottom panel). Partial co-occupancy can be further subdivided into two categories: A is required (bottom left) or not required (bottom right) for the binding of B.

always present in unison at any given DNA molecule. No A/B co-occupancy takes place when both A and B bind to the DNA region in question, yet are confined to non-overlapping subsets of molecules and are never found at the DNA fragments at the same time. Partial co-occupancy takes place when some of the DNA molecules are bound by both A and B, while others are bound by only A and/or only B.

Partial co-occupancy can be further subdivided into two distinct subgroups, differentiated by the ability or inability of one protein (B) to bind DNA independently of the other (A) as illustrated in Figure 21.8.1. In one scenario, DNA binding of protein B always occurs in combination with (and is completely dependent of the presence of) protein A, whereas protein A can associate with DNA in the absence of protein B. An example of this factor-

dependent DNA binding can be found in the interactions between the yeast TATA-box binding protein (TBP) and TBP-associated factors (TAFs): promoter binding by TAFs is critically dependent on TBP, yet TBP can be found at many promoters that lack TAFs (Kuras et al., 2000; Li et al., 2000). In the second scenario, A and B can bind independently of one another to a given genomic region, yet the proteins may co-occupy the same stretch of DNA in some (but not all) of the cases. This scenario is likely to be valid for two DNA-binding proteins that do not interact with one another and bind to different target sequences in an enhancer or other regulatory region of a eukaryotic promoter.

Quantitative analysis of SeqChIP results

Fold-enrichment over background for both individual ChIPs and SeqChIPs can be

calculated relative to a control genomic location as described in *UNIT 21.3*. Let A, B, and AB represent the fold-enrichments for the first, second, and sequential IPs, respectively. Then, in cases of complete co-occupancy (e.g., components of the general RNA polymerase II transcription machinery), the fold-enrichment of individual ChIPs should be equal to the product of the fold-enrichment of the individual ChIPs ($A \times B = AB$; Geisberg and Struhl, 2004). In cases of no co-occupancy, no enrichment is seen over that obtained by the first immunoprecipitation ($AB \approx A$). For partial co-occupancy, the fold-enrichment typically lies somewhere in between A (no co-occupancy) and AB (full co-occupancy). Mathematically, efficiency of co-occupancy (C, in percent) can be defined as follows:

$$C = 100(AB - A)/(A \times B - A)$$

In cases of complete co-occupancy, AB typically equals the product of A and B, and C = 100. If the value of $AB = A$, then there is no enrichment over the first immunoprecipitation and C = 0. For cases of partial co-occupancy, typical experimentally obtained C values range between >0 to significantly <100. It is important to note that in order for C values to be meaningful, fold-occupancies for individual ChIPs must be substantially above background (typically greater than two-fold), and SeqChIP enrichment (the ratio AB/A must be reproducibly above background as well ($AB/A > 2$ works well). Of course, the definition of what represents experimental background is somewhat arbitrary and will vary depending on experimental setup, the number of repetitions, as well as other factors.

Order of IPs may influence SeqChIP outcomes

An interesting (and predicted) outcome in some SeqChIP experiments is that the order of sequential IPs makes a difference in cases of partial co-occupancy. In a typical experiment, there are three classes of DNA molecules of interest—those bound to A alone, B alone, or both A and B. If X is defined as the fraction of A-DNA molecules that also contain B, and Y is defined as the fraction of B-DNA that also contain A, then the expected fold-enrichments for the three different classes of molecules are given by:

- (1 - X)A for the A-alone DNA molecules
- (1 - Y)B for the B-alone DNA molecules
- (XA)(YB) for the A + B + DNA molecules

It should be noted that X and Y are independent of one another, as A and B may have different occupancy properties. Using the equations above, it follows that:

when A ChIP is first, the fold-enrichment (AB) for the SeqChIP = $(1 - X)A + (XA)(YB)$

when B ChIP is first, the fold-enrichment (BA) for the SeqChIP = $(1 - Y)B + (XA)(YB)$.

For cases of complete co-occupancy ($X = 1, Y = 1$), the fold-enrichment is the same irrespective of IP order, as the equations above simplify to $AB = BA$. Likewise, the order of IPs makes no difference in cases of no co-occupancy ($X = 0, Y = 0$), as $AB = BA = 0$. By contrast, the order of IPs does make a difference in cases of partial co-occupancy. For example, when $X = 1$ and Y is small (i.e., A always co-occupies with B, but B rarely co-occupies with A), meaningful co-occupancy is likely to be observed in only one direction. In specific terms, if $X = 1, Y = 0.1, A = 10$, and $B = 50$, then the fold-enrichments are 50 when A ChIP is first and 95 when B ChIP is first. Clearly, these two values are not equal ($AB \neq BA$). More importantly, when A ChIP is performed first, the SeqChIP value of 50 represents a five-fold enrichment relative to the value of the single IP, although it is also well below the value for predicted full co-occupancy ($A \times B = 500$). By contrast, when B ChIP is performed first, the SeqChIP value of 95 is less than two-fold greater than the value of the single IP ($B = 50$), and typically falls within the cut-off for experimental error. The theoretically predicted (and experimentally confirmed) IP-order-dependent nature of SeqChIP values requires careful interpretation of so-called negative (no co-occupancy) results: definitive demonstration of lack of co-occupancy between two factors requires that the SeqChIP be performed in both directions with identical results of no co-occupancy. On the other hand, partial (or full) co-occupancy can be effectively demonstrated if observed in one direction.

Time Considerations

Cell growth, formaldehyde cross-linking, and chromatin isolation/purification can be completed in 2 to 3 days. Sequential immunoprecipitation requires a total of 7 to 8 hr: 2 hr for antibody binding, 1 hr for washes, and 30 min for elution for each IP, with 30 min set-up time in between the two immunoprecipitations. Complexes eluted from the first IP can

be conveniently frozen at -20°C and the second immunoprecipitation can be performed at a later time for added convenience. Cross-link reversal takes ≥ 6 hr and is best performed in a thermal cycler programmed to run overnight (see UNIT 21.3, Alternate Protocol 2). DNA purification takes 2 hr and quantitative PCR takes ≤ 2 to 3 hr, depending on instrument and cycling parameters.

Literature Cited

- Geisberg, J.V. and Struhl, K. 2004. Quantitative sequential chromatin immunoprecipitation, a method for analyzing co-occupancy of proteins at genomic regions in vivo. Submitted.
- Ijpenberg, A., Tan, N.S., Gelman, L., Kersten, S., Seydoux, J., Xu, J., Metzger, D., Canaple, L., Chambon, P., Wahli, W., and Desvergne, B. 2004. In vivo activation of PPAR target genes by RXR homodimers. *EMBO J.* 23:2083-2091.
- Kuras, L., Kosa, P., Mencia, M., and Struhl, K. 2000. TAF-containing and TAF-independent forms of transcriptionally active TBP in vivo. *Science* 288:1244-1248.
- Li, X.-Y., Bhaumik, S.R., and Green, M.R. 2000. Distinct classes of yeast promoters revealed by differential TAF recruitment. *Science* 288:1242-1244.
- Metivier, R., Penot, G., Hubner, M.R., Reid, G., Brand, H., Kos, M., and Gannon, F. 2003. Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* 115:751-763.
- Proft, M. and Struhl, K. 2002. Hog1 kinase converts the Sko1-Cyc8-Tup1 repressor complex into an activator that recruits SAGA and SWI/SNF in response to osmotic stress. *Mol. Cell* 9:1307-1317.
- Soutoglou, E. and Talianidis, I. 2002. Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation. *Science* 295:1901-1904.

Contributed by Joseph V. Geisberg and
Kevin Struhl
Harvard Medical School
Boston, Massachusetts

Defining In Vivo Targets of Nuclear Proteins by Chromatin Immunoprecipitation and Microarray Analysis

UNIT 21.9

**BASIC
PROTOCOL**

Chromatin immunoprecipitation (ChIP) is an extremely useful method for determining the association of a given protein with potential DNA targets in vivo (UNIT 21.3). To identify novel regions of protein association, ChIP may be combined with microarray technology to identify the location of a specific protein in an unbiased, genome-wide manner, a technique often called “ChIP-chip” or “genome-wide location” analysis. Such an approach is particularly powerful, since it requires no previous knowledge of a protein’s likely targets. The small amount of DNA obtained by ChIP must first be amplified to yield enough material for array hybridization; the methods used are designed to ensure that the relative concentrations of all genomic regions in the original samples are faithfully maintained. Immunoprecipitated and input DNAs from standard ChIP experiments (UNIT 21.3) are subjected to primer extension using a primer with a 5′ tail consisting of a fixed sequence and a random octamer at its 3′ end; the primer-extension products are then amplified by PCR. The resulting amplified material is hybridized to microarrays of DNA fragments representing a large number of genomic regions. As described for transcriptional profiling (UNIT 22.1), DNA on the microarrays may be in the form of PCR products or oligonucleotides. The process of printing microarrays with PCR products is described in UNIT 22.3. Examples of such arrays include (1) the complete set of intergenic regions in the yeast *Saccharomyces cerevisiae*, (2) all nonrepetitive DNA sequences on entire human chromosomes, (3) numerous promoter-proximal regions of expressed mouse genes, and (4) CpG islands. This unit consists of a ChIP-chip protocol derived from experiments involving whole-genome analysis of *Saccharomyces cerevisiae* and using glass microarrays spotted with PCR products. It is applicable to the study of other organisms, although some organism- and array-specific modifications may be necessary.

Materials

- De-cross-linked DNA samples: typically yeast DNA from chromatin immunoprecipitation and input DNA control; (UNIT 21.3)
- 5× Sequenase buffer (USB)
- 80 μM primer A: GTTCCAGTCACGATCNNNNNNNNN (see UNIT 2.11 for oligonucleotide synthesis)
- 10 mg/ml BSA
- 0.1 M DTT
- 10 mM dNTP mix: 10 mM each dATP, dCTP, dGTP, and dTTP (UNIT 3.4)
- 13 U/μl Sequenase (USB) stock: dilute 1/10 immediately before use
- 10× PCR buffer (see recipe)
- Amino-allyl dNTP mix (see recipe)
- 100 μM primer B: GTTCCAGTCACGATC (see UNIT 2.11 for oligonucleotide synthesis)
- 5 U/μl *Taq* DNA polymerase
- 0.1 M sodium bicarbonate, pH 9.0
- Cy3 or Cy5 monoreactive dye (Amersham): resuspend entire tube of Cy5 or Cy3 dye in 45 μl DMSO
- 3 M sodium acetate (APPENDIX 2)
- 10 mg/ml sheared salmon sperm DNA (UNIT 20.1)
- 2.5× array hybridization buffer (see recipe)

**Chromatin
Assembly and
Analysis**

21.9.1

Array prehybridization buffer (see recipe)
 95% ethanol
 Array wash buffer (see recipe)
 0.2× SSC (APPENDIX 2)
 QIAquick PCR purification kit (Qiagen)
 Thermal cycler
 MinElute PCR purification kit (Qiagen)
 Microarrays (UNIT 22.1)
 Coplin jar or petri dish
 Plastic racks for holding microscope slides
 Plastic containers for washing microscope slides in holders
 Tabletop centrifuge with adaptors suitable for holding plastic microscope slide racks (typically the same as the adaptors meant for holding 96-well plates)
 95°C heating block or water bath
 LifterSlip coverslips for microarrays (Erie Scientific Company)
 Rubber cement
 Humidified chamber (see recipe)
 45°C oven
 Platform shaker
 Additional reagents and equipment for PCR amplification (UNIT 15.1) and separation of small DNA fragments by agarose gel electrophoresis (UNIT 2.5A)

Perform primer extension

1. Purify each DNA sample using a QIAquick PCR purification column according to the manufacturer's instructions, eluting with 42 µl water.

Use at least 50% of an immunoprecipitated sample (UNIT 21.3) from 20 ml of yeast cells—this corresponds to 10⁸ cells. For input DNA control, use 1/100 of the original chromatin sample from 20 ml cells.

2. Set up the following primer extension reaction in a 0.5-ml microcentrifuge tube:

40 µl QIAquick-purified DNA (immunoprecipitated or input-DNA control)
 12 µl 5× Sequenase buffer
 2 µl 80 µM primer A.

3. Begin the primer extension in a thermal cycler programmed as follows:

2 cycles:	2 min	94°C
	2 min	10°C
	Ramp slowly to 37°C over an 8-min period	
	8 min	37°C.

If processing many samples, it may be easier to pause the thermal cycler at 10°C long enough to add enzyme mix to each tube (see steps 4 and 5).

4. During the first 10°C step, add the following:

0.5 µl 10 mg/ml BSA
 3 µl 0.1 M DTT
 2 µl 10 mM dNTP mix
 1 µl diluted Sequenase.

It is best to premix the four components. If droplets have collected on the lid of the PCR tube, microcentrifuge briefly at maximum speed to collect the contents before adding the mix.

5. During the second 10°C step, add an additional 1 µl diluted Sequenase.
6. Upon completion of the 2 cycles, purify the DNA on a QIAquick column, eluting with 80 µl water.

Perform PCR

7. Mix the following components in a 0.5-ml microcentrifuge tube:

78 μ l DNA from step 6
10 μ l 10 \times PCR buffer
10 μ l amino-allyl dNTP mix
1 μ l 100 μ M primer B
1 μ l 5 U/ μ l *Taq* DNA polymerase.

8. Perform PCR (also see *UNIT 15.1*) using the following program in a thermal cycler:

35 cycles:	30 sec	92°C
	30 sec	40°C
	30 sec	50°C
	1 min	72°C.

If microarrays require blocking or other processing, it is convenient to do this while the PCR program is being executed.

9. Check 5 μ l of the PCR product on an agarose gel (*UNIT 2.5A*) to determine the success of amplification.

A successful amplification will yield a smear of DNA fragments with an average size of ~300 bp. A faint smear is all that is necessary.

10. Purify the remaining 95 μ l of the PCR reaction with MinElute kit according to the manufacturer's instructions, eluting in 11 μ l of 0.1 M sodium bicarbonate, pH 9.0.

For this procedure, be sure to elute using 0.1 M sodium bicarbonate, pH 9.0, not the elution buffer supplied with the MinElute kit.

Couple samples with dye

11. Mix 10 μ l eluted DNA from step 10 with 10 μ l Cy5 or Cy3 dye in DMSO. Incubate 1 hr at room temperature in darkness.

For a given array, the sample from immunoprecipitation is coupled to one dye, and the total DNA input to the other. If the hybridization will be performed on the same day, it is convenient to prehybridize the arrays (step 13) during this incubation.

12. Add 80 μ l water and 3 μ l of 3 M sodium acetate. Purify on QIAquick PCR clean-up column according to the manufacturer's instructions, washing with PE buffer (provided with QIAquick kit) three times or until all uncoupled dye has been removed and flowthrough runs clear. Elute with 32 μ l water. Store samples at -20°C until ready to perform hybridization.

The eluted dye-coupled product should be visibly pink or blue.

Prehybridize/hybridize samples to arrays

13. Incubate prepared microarrays for 1 hr at 45°C in enough array prehybridization buffer to cover the slides.

For multiple slides, it is convenient to use a Coplin jar, but for one or two slides a petri dish will work equally well.

14. Place prehybridized microarrays in microscope slide racks and briefly dip five times in water, then once in 95% ethanol. Dry by centrifuging 1 min at 800 rpm in a tabletop centrifuge.

15. Prepare hybridization solution by mixing the two dye-coupled samples representing immunoprecipitated and input DNA (30 μ l each) for a given array with 3 μ l of 10 mg/ml salmon sperm DNA and 40 μ l of 2.5 \times array hybridization buffer.

Any excess hybridization solution may be stored at -20°C for future use.

- Denature hybridization mixture at 95°C for 3 min. Transfer immediately to ice. Place a LifterSlip coverslip ridge-side-down onto the prehybridized array. Using a small plastic pipet tip, dab a small dot of rubber cement at each corner of the cover slip to prevent it from moving. Gently pipet the sample-containing hybridization solution (from step 15) onto the array under the LifterSlip coverslip, allowing capillary action to pull the liquid under the coverslip rather than squirting it all in at once.

LifterSlip coverslips have two white ridges along their edges to raise the coverslip slightly above the slide surface, allowing a larger volume of liquid to be injected. Take care not to put pressure on the coverslip when handling the array. About 50 μ l will suffice to fill the space under the cover slip. Store any leftover hybridization mixture at -20°C in case of catastrophe.

- Incubate arrays overnight at 45°C in a humidified chamber (see Reagents and Solutions).
- Remove array slides from humidified chamber, remove coverslips, and place arrays in a slide holder in a plastic wash container. Wash array twice, each time for 3 min in ~ 200 ml array wash buffer. Agitate slides on a platform shaker during washes.
- Wash arrays twice, each time for 3 min in 200 ml 0.2 \times SSC using the same technique as in step 18.

Transfer slides to a fresh rack to minimize SDS carryover between wash steps.

- Dry arrays by centrifuging 1 min at 800 rpm in a tabletop centrifuge.

Proceed quickly with the centrifugation step so as not to allow any residual wash solution to dry unevenly on the slide surface. The slides are now ready for scanning. If there is a high background on scanning, further washes may be performed and the slides re-scanned.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Amino-allyl dNTP mix

5 mM each dATP, dCTP, and dGTP
2 mM dTTP
3 mM amino-allyl dUTP (Ambion)
Store at -20°C

Array hybridization buffer, 2.5 \times

25 μ l formamide
15 μ l 20 \times SSC (APPENDIX 2)
1 μ l 10% (w/v) SDS
Store at -20°C

Array prehybridization buffer

5 \times SSC (APPENDIX 2)
0.1% (w/v) SDS
1% (w/v) BSA
Store at 4°C

Array wash buffer

0.2 \times SSC (APPENDIX 2)
0.1% (w/v) SDS
Store indefinitely at room temperature

Humidified chamber

Remove the sieve-like plastic insert from an empty hinged pipet-tip box. Place a moist paper towel in the bottom of the box and replace the plastic insert. Place arrays on top of the plastic insert, and close the hinged lid of the box before placing the box in the incubator.

The perforated insert allows the circulation of humid air within the chamber without actually allowing the arrays to sit on a wet surface.

PCR buffer, 10×

100 mM Tris·Cl, pH 8.3 (APPENDIX 2)
500 mM KCl
15 mM MgCl₂
Store at −20°C

Sequenase buffer, 5×

200 mM Tris·Cl, pH 7.5 (APPENDIX 2)
100 mM MgCl₂
250 mM NaCl
Store at −20°C

COMMENTARY

Background Information

The combination of chromatin immunoprecipitation (ChIP) and microarray technology represents a powerful approach to identifying the location of specific proteins on a genome-wide basis. ChIP and quantitative PCR (UNIT 21.3) have been used with considerable success to study the association of proteins with individually chosen potential DNA targets (Kuras et al., 1999; Li et al., 1999). In addition, microarrays are widely used for genome-wide expression profiling (DeRisi et al., 1996, 1997; Lockhart et al., 1996), as described in UNIT 22.1. In recent years, ChIP-chip using yeast whole-genome microarrays has allowed the identification of the physiological target sites of numerous DNA-binding proteins, chromatin-modifying activities, and transcription factors in an unbiased and comprehensive manner (Reid et al., 2000; Iyer et al., 2001; Lee et al., 2002; Moqtaderi and Struhl, 2004).

The small amount of DNA obtained from a typical ChIP experiment must be amplified before being hybridized to a microarray. The methods used generally involve adding fixed sequence to the ends of the immunoprecipitated DNA fragments either by ligation or by primer extension using oligonucleotides with a fixed 5' tail and a random 3' end (Bohlander et al., 1992; Eberwine, 1996). The resulting fragments may then be amplified by PCR, in the presence of either dye-coupled nucleotides

or a modified amino-allyl nucleotide that is coupled to dye in a subsequent step. The advantage of the latter method is that it eliminates unequal labeling that results from different efficiencies of incorporation of Cy3- and Cy5-coupled nucleotides during PCR.

Critical Parameters and Troubleshooting

It is important to begin with sufficient immunoprecipitated DNA of high purity. In general, half an immunoprecipitation reaction is enough, and the Qiagen purification before the initial primer extension is sufficient to ensure that the DNA is clean and of high quality. As for any ChIP experiment, be certain that the chromatin sample used in the immunoprecipitation is sufficiently sonicated; smaller fragment sizes (ideally averaging 300 bp) result in higher signal-to-background ratios. This can be especially important when trying to distinguish true occupancy at a locus from apparent occupancy at a neighboring locus on a microarray. In addition, on an array consisting of relatively large spotted PCR products (e.g., ones representing entire intergenic sequences), the true association site of any given factor is likely to be considerably smaller than the size of the spotted fragment. The contribution of the neighboring nonoccupied DNA on the same PCR fragment tends to dampen the overall positive signal of the spot, making the apparent occupancy lower for a given spot

than it would appear with smaller PCR products spanning little sequence other than true target. It is therefore especially important in a microarray experiment of this type to maximize the immunoprecipitation signal; to this end, it is often helpful to use peptide elution (UNIT 21.3, Alternate Protocol) when immunoprecipitating epitope-tagged proteins.

If the scanned array has high background fluorescence, especially background signal that appears patchy, it is often helpful to wash and scan the slide again. It is important not to allow the wash solution to dry unevenly on the glass slide—dry the slides quickly by centrifugation. Be certain to use a fresh slide rack and new plastic wash container for each wash so as to minimize carryover of SDS-containing wash solution, which can appear as a blotchy background signal.

If using a previously untested batch of microarrays, it is often useful to perform a trial hybridization using unimportant material (e.g., amplified total genomic DNA, or even leftover unhybridized material from a previous experiment) to verify the quality of a representative array. The array should consist of distinct, round spots with no unevenness of printing; missed spots or feathering of spot edges are signs of bad quality. Some of these problems may be obvious upon just a brief SYBR Green staining of the array, but occasionally problems with array quality become apparent only after overnight hybridization.

The procedure described here yields approximately twice as large a volume of hybridization mixture as is needed to cover a typical microarray. The excess material may be stored at -20°C in case any unanticipated slide-specific problems occur with the first array.

At least three replicates of an array experiment are usually performed to allow statistical determination of the validity of the results. These replicates should be fully independent, resulting from separately grown cultures and chromatin preparations. Even with multiple replicates, it will be apparent that certain spots emerge reliably as false positives—i.e., statistically meaningful positives which turn out to have no biological merit. Some of these are antibody-dependent and can be easily identified by array hybridization of sample immunoprecipitated with the same antibody, but from an untagged strain.

Anticipated Results

In general, a good yield of amplified material after the PCR step (i.e., 5 μl of the PCR

reaction should produce an easily visible smear when run on an agarose gel) is highly predictive of success in the array hybridization. It is therefore always worthwhile to run a gel to verify the success of the amplification step before continuing. If no smear of amplified DNA is visible, do not continue; any resulting array hybridization signal is likely to be very weak. Additionally, the material eluted from the Qiagen column after the subsequent dye-coupling step should be clearly and unambiguously pink or blue; it is not a good sign if seeing this color requires a great deal of imagination or fiddling about with white pieces of background paper. If a co-worker hesitates when asked to distinguish which of the two samples is the pink one, do not waste an array by proceeding with hybridization.

Time Considerations

The primer extension, amplification, and labeling steps in this procedure can be performed relatively easily in one day, allowing hybridization to be performed overnight. Washing and scanning can then be done on the morning of the second day. If the protein of interest has known DNA targets, it is often worth spending 1 day at the beginning of the procedure to determine the success of the immunoprecipitation (by quantitative PCR of known DNA targets) before embarking on amplification and array hybridization.

If the microarrays require advance preparation such as blocking or UV cross-linking, these steps may be performed during the PCR amplification. It is convenient to prehybridize the arrays during the 1-hr dye-coupling step.

Literature Cited

- Bohlander, S.K., Espinosa, R. III, Le Beau, M.M., Rowley, J.D., and Diaz, M.O. 1992. A method for the rapid sequence-independent amplification of microdissected chromosomal material. *Genomics* 13:1322-1324.
- DeRisi, J., Penland, L., Brown, P.O., Bittner, M.L., Meltzer, P.S., Ray, M., Chen, Y., Su, Y.A., and Trent, J.M. 1996. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nature Genet.* 14:457-460.
- DeRisi, J.L., Iyer, V.R., and Brown, P.O. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680-686.
- Eberwine, J. 1996. Amplification of mRNA populations using aRNA generated from immobilized oligo(dT)-T7 primed cDNA. *BioTechniques* 20:584-591.
- Iyer, V.R., Horak, C.E., Scafe, C.S., Botstein, D., Snyder, M., and Brown, P.O. 2001. Genomic

- binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* 409:533-538.
- Kuras, L. and Struhl, K. 1999. Binding of TBP to promoters in vivo is stimulated by activators and requires Pol II holoenzyme. *Nature* 399:609-613.
- Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., Zeitlinger, J., Jennings, E.G., Murray, H.L., Gordon, D.B., Ren, B., Wyrick, J.J., Tagne, J.B., Volkert, T.L., Fraenkel, E., Gifford, D.K., and Young, R.A. 2002. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* 298:799-804.
- Li, X.Y., Virbasius, A., Zhu, X., and Green, M.R. 1999. Enhancement of TBP binding by activators and general transcription factors. *Nature* 399:605-609.
- Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., and Brown, E.L. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* 14:1675-1680.
- Moqtaderi, Z. and Struhl, K. 2004. Genome-wide occupancy profile of the RNA polymerase III machinery in *Saccharomyces cerevisiae* reveals loci with incomplete transcription complexes. *Mol. Cell. Biol.* 24:4118-4127.
- Reid, J.L., Iyer, V.R., Brown, P.O., and Struhl, K. 2000. Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase. *Mol. Cell.* 6:1297-1307.

Contributed by Zarmik Moqtaderi and
Kevin Struhl
Harvard Medical School
Boston, Massachusetts

Identifying Chromosomal Targets of DNA-Binding Proteins by Sequence Tag Analysis of Genomic Enrichment (STAGE)

This unit describes a protocol for identifying the target loci of DNA-binding proteins in any genome by sequence tag analysis of genomic enrichment (STAGE). The method is conceptually similar to serial analysis of gene expression (SAGE; UNIT 25B.6), but the template for STAGE consists of DNA fragments enriched by chromatin immunoprecipitation (ChIP; UNIT 21.3). The STAGE procedure generates a library of 21-mer sequences, or tags, that represent the original ChIP sample. High-throughput sequencing of concatamers of tags, and the mapping of these tags to the genome, allows one to identify the direct binding targets of DNA-binding proteins *in vivo*.

STAGE starts with the amplification of ChIP DNA as described in the Basic Protocol. The enrichment of specific target loci by ChIP can be detected by ChIP-PCR (UNIT 21.3). If the enrichment of known target loci relative to nonspecific genomic background is not very high, a large proportion of tag sequences could be derived from this nonspecific background. An optional procedure, Subtraction STAGE (SubSTAGE; see Support Protocol), may be used to obtain better enrichment of target loci for input into the subsequent STAGE procedure.

SEQUENCE TAG ANALYSIS OF GENOMIC ENRICHMENT (STAGE)

A schematic representation of the STAGE procedure is shown in Figure 21.10.1. The steps of the STAGE procedure, in brief, are as follows. (1) ChIP DNA fragments are amplified by PCR using partially degenerate biotinylated primers, and (2) the product is digested with the restriction endonuclease *Nla*III, which cuts at 5'-CATG. (3) Biotinylated fragments are isolated using streptavidin beads and split into two pools. (4) Each pool is ligated to a different linker containing a recognition site for *Mme*I, a type IIS restriction enzyme that cuts away from its recognition site, as well as a specific primer-recognition sequence. (5) Digestion with *Mme*I releases 21-bp tags that are derived from the original ChIP DNA, attached to the linker sequence. (6) The two pools are ligated and amplified by nested PCR using primers complementary to the linkers. (7) The amplified sample is trimmed by *Nla*III digestion to generate ditags. (8) Multiple ditags are concatamerized, cloned, and sequenced. (9) Mapping each tag back to the genome sequence can identify the DNA fragments that were present in the ChIP sample and thus identify the genomic loci that are occupied by DNA-binding protein *in vivo*.

Materials

- 10 ng/ μ l chromatin immunoprecipitated (ChIP) DNA fragments (UNIT 21.3; test enrichment by PCR as described in that unit; if not sufficiently high, perform subtraction as described in Support Protocol to improve enrichment)
- 5 \times Sequenase reaction buffer (USB)
- 40 μ M round A primer (see recipe for PCR primers)
- 10 ng/ μ l sheared genomic DNA generated from the same cells used in the ChIP procedure
- TE buffer, pH 8.0 (APPENDIX 2)
- Round A dNTP mix: 3 mM each dNTP (A, C, G, and T) in TE buffer, pH 8.0
- 0.1 M dithiothreitol (DTT)
- 0.5 mg/ml bovine serum albumin (BSA)

BASIC PROTOCOL

Chromatin Assembly and Analysis

21.10.1

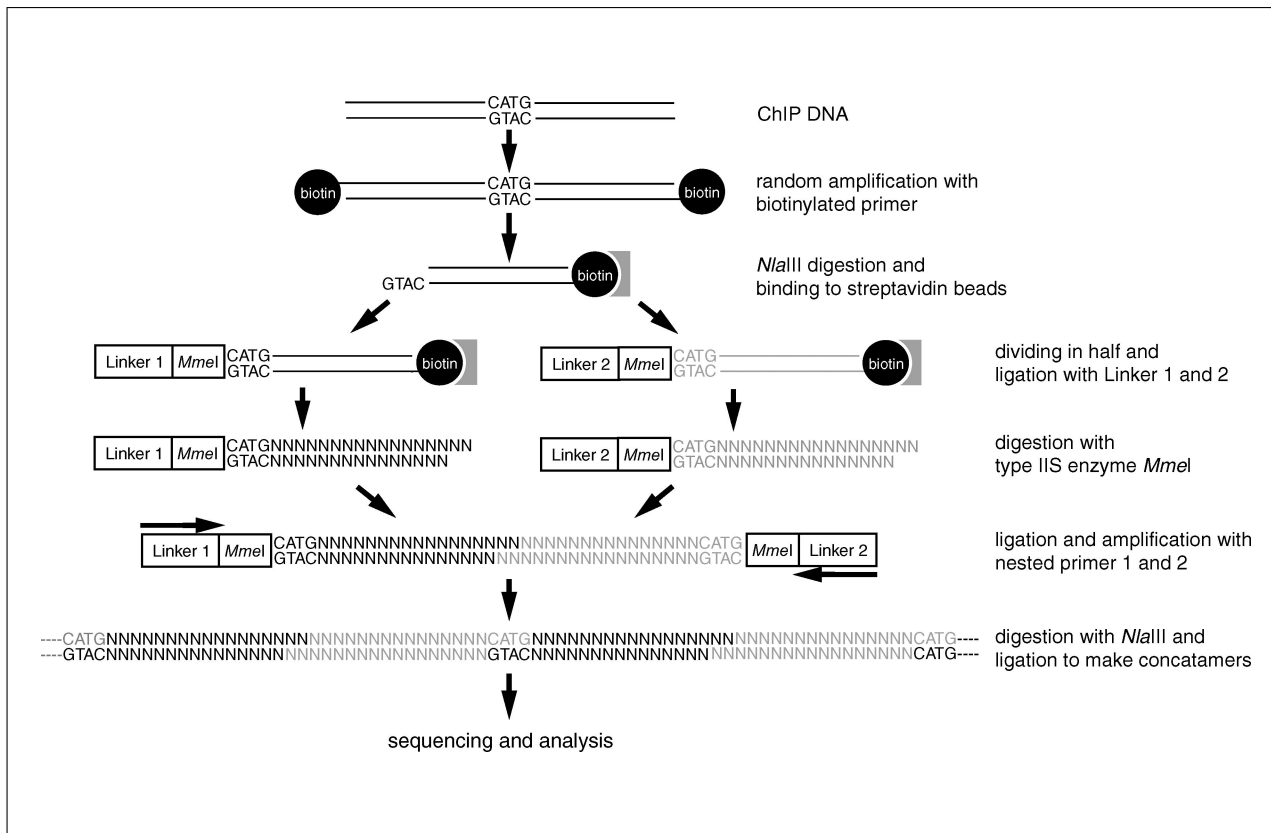


Figure 21.10.1 Scheme for sequence tag analysis of genomic enrichment (STAGE).

- 13 U/μl Sequenase (USB)
- 10× PCR buffer (see recipe)
- 25 mM MgCl₂
- 500 μM biotinylated round B primer (see recipe for PCR primers)
- Standard dNTP mix: 25 mM each dNTP (A, C, G, and T) in TE buffer, pH 8.0
- 5 U/μl *Taq* DNA polymerase
- 1% and 1.5% agarose gels (UNIT 2.5A)
- PC8 (see recipe)
- 20 mg/ml glycogen (Roche Diagnostics)
- 7.5 M ammonium acetate (Sigma)
- 100% and 70% ethanol
- LoTE buffer (see recipe)
- 100× BSA (New England Biolabs)
- 10× Buffer 4 (New England Biolabs)
- Nla*III restriction endonuclease (New England Biolabs; also see UNIT 3.1)
- Dynabeads M-280 Streptavidin slurry (Dynal)
- 1× and 2× BW buffer (see recipe)
- Annealed linkers 1 and 2: linkers 1A, 1B, 2A, and 2B (see recipe), phosphorylated and annealed as in Support Protocol 3 of UNIT 25B.6
- 10× ligase buffer (New England Biolabs)
- T4 DNA ligase (high concentration, 2000 U/μl; New England Biolabs)
- 32 mM *S*-adenosylmethionine (SAM; New England Biolabs)
- Mme*I restriction endonuclease (New England Biolabs)
- Primers 1 and 2 (see recipe for PCR primers)
- 12% (10-well) and 8% nondenaturing polyacrylamide gels prepared in TBE buffer (UNIT 2.7)

**Sequence Tag
Analysis of
Genomic
Enrichment
(STAGE)**

21.10.2

10-bp and 1-kbp DNA ladders
 6× sample loading buffer (see recipe)
 1 μg/μl pZErO-1 plasmid (Invitrogen), linearized
 10× Buffer 2 (New England Biolabs)
SphI restriction endonuclease (New England Biolabs)
 ElectroMAX DH10B electroporation-competent cells (Invitrogen; frozen at
 −80°C)
 SOC medium (see recipe)
 10-cm Zeocin-containing low-salt LB plates (see recipe)
 Dimethyl sulfoxide (DMSO)
 5 μM M13 forward and reverse primers (see recipe for PCR primers)
 0.2-ml thin-walled PCR tubes
 Thermal cycler
 0.5- and 1.5-ml snap-cap and 2-ml screw-cap microcentrifuge tubes
 Magnetic rack for Dynabead separations (Dyna)
 12°, 16°, 37°, 42°, 50°, and 65°C water baths or temperature blocks
 15- and 50-ml conical tubes
 Tabletop centrifuge with swinging-bucket rotor
 25-G needles
 Spin-X centrifuge tube filters (Costar)
 Bio-Rad GenePulser electroporator (or equivalent) and 0.1-mm disposable
 electroporation cuvettes (Bio-Rad)
 Platform shaker
 37° and 30°C incubators
 96-well PCR plate (MJ Research)
 Sterile toothpicks

Additional reagents and equipment for PCR (*UNIT 15.1*), restriction enzyme
 digestion of DNA (*UNIT 3.1*), phenol extraction and ethanol precipitation of DNA
 (*UNIT 2.1A*), agarose gel electrophoresis (*UNIT 2.5A*), nondenaturing polyacrylamide
 gel electrophoresis (*UNIT 2.7*), ligation (*UNIT 3.14*), electroporation of bacteria
 (*UNIT 1.8*), and direct DNA sequencing of PCR products (*UNIT 15.2*)

NOTE: Use aerosol-barrier pipet tips to prevent contamination of PCR reactions; see
APPENDIX 2 for additional precautions with PCR.

Perform primer extension: Round A

1. For each amplification reaction to be performed, prepare the following mixture in
 separate 0.2-ml thin-walled PCR tubes:

7 μl ChIP DNA fragments
 2 μl 5× Sequenase reaction buffer
 1 μl 40 μM round A primer.

For positive control reactions, use 7 μl of 10 ng/μl sheared genomic DNA in place of
 the ChIP DNA. For negative control reactions, use 7 μl TE buffer, pH 8.0, in place
 of the ChIP DNA.

2. Based on the total number of reactions to be performed, add the following ingredients
 to a single tube (total 5.05 μl per reaction):

1 μl 5× Sequenase reaction buffer
 1.5 μl round A dNTP mix (3 mM each dNTP)
 0.75 μl 0.1 M DTT
 1.5 μl 0.5 mg/ml BSA
 0.3 μl 13 U/μl Sequenase.

- Based on the total number of reactions to be performed, add the following ingredients to another tube (total, 1.2 μl per reaction):

0.3 μl 13 U/ μl Sequenase
0.9 μl Sequenase reaction buffer.

- Transfer the tubes from step 1 to a thermal cycler and carry out two rounds of DNA strand extension as follows:

2 cycles: 2 min 94°C
 2 min 8°C
 Ramp to 37°C over 8 min.

During the first 8°C step, add 5.05 μl of the reaction mixture from step 2 to each tube. Quickly mix the sample and reaction mixture by pipetting. During the second 8°C step, add 1.2 μl of the mixture from step 3 to each tube and mix the reaction well by pipetting.

Use the pause option on the instrument, if necessary, to add the reaction buffer and enzyme to each tube and mix.

- Dilute the product to 30 μl with TE buffer, pH 8.0, and use 15 μl of this diluted round A product for each round B reaction.

Perform PCR for biotinylation: Round B

- Mix the following components in each of two 0.2-ml thin-walled PCR tubes:

15 μl round A product
10 μl 10 \times PCR buffer
8 μl 25 mM MgCl₂
1.5 μl 500 μM biotinylated round B primer
2 μl standard dNTP mix (25 mM each dNTP)
62.5 μl H₂O
1 μl 5 U/ μl *Taq* DNA polymerase.

- Transfer the tubes to a thermal cycler and amplify the DNA templates using the following conditions:

1 cycle: 2 min 94°C (denaturation)
25 cycles: 30 sec 94°C (denaturation)
 30 sec 45°C (first annealing)
 30 sec 50°C (second annealing)
 1 min 70°C (extension)
1 cycle: 5 min 70°C (final extension).

- Analyze 5 μl of round B product by electrophoresis on a 1% agarose gel (UNIT 2.5A).

A smear of DNA averaging in size from 300 to 1000 bp should be seen from the amplified ChIP DNA sample and genomic DNA control. No smear should be seen from the amplification of the negative control (TE).

- Combine the two round B reactions (final 200 μl) in one 1.5-ml microcentrifuge tube and extract the sample with an equal volume of PC8 (also see UNIT 2.1A).

10. Precipitate with ethanol by combining the following (also see *UNIT 2.1A*):

200 μ l round B product
2 μ l 20 mg/ml glycogen
150 μ l 7.5 M ammonium acetate
600 μ l 100% ethanol.

Vortex briefly and place at -80°C for 15 min.

11. Microcentrifuge 30 min at maximum speed, room temperature. Wash by adding 500 μ l of 70% ethanol and microcentrifuging 5 min at maximum speed, room temperature. Remove ethanol and resuspend pellet in 20 μ l LoTE buffer.

Digest with *Nla*III and immobilize biotinylated fragments to magnetic beads

12. Mix the following components:

10 μ l round B product (half the total product)
2 μ l 100 \times BSA
20 μ l 10 \times Buffer 4
8 μ l *Nla*III
160 μ l LoTE buffer.

Incubate for 1.5 hr at 37°C .

13. Extract the sample with 200 μ l PC8 and precipitate with ethanol by combining the following:

200 μ l sample
2 μ l 20 mg/ml glycogen
150 μ l 7.5 M ammonium acetate
600 μ l 100% ethanol.

Vortex briefly and place at -80°C for 15 min.

14. Microcentrifuge 30 min at maximum speed, room temperature. Wash by adding 500 μ l of 70% ethanol and microcentrifuging 5 min at maximum speed, room temperature. Remove ethanol and resuspend pellet in 20 μ l LoTE buffer.

15. Place a 100- μ l aliquot of Dynabeads M-280 Streptavidin slurry in each of two microcentrifuge tubes. Immobilize the beads using a magnetic rack according to manufacturer's instructions, and remove the supernatant. Wash beads three times, each time with 500 μ l of 1 \times BW buffer, using the magnetic rack.

16. Add the following to each tube containing the washed Dynabeads:

100 μ l 2 \times BW buffer
90 μ l H_2O
10 μ l DNA sample.

Incubate 20 min at room temperature with gentle mixing.

17. Immobilize the samples in two tubes with the magnetic rack. Wash three times with 500 μ l of 1 \times BW buffer and once with LoTE buffer using the magnetic rack.

Ligate linkers to sample

18. Resuspend the beads in each tube with the following:

- 29 μl LoTE buffer (both tubes)
- 4 μl 10 \times ligase buffer (both tubes)
- 5 μl linker 1 (only in tube 1)
- 5 μl linker 2 (only in tube 2).

Prepare linkers 1 and 2 by phosphorylating linkers 1B and 2B and annealing with linkers 1A and 2A as described in Support Protocol 3 of UNIT 25B.6. Use the four oligonucleotide sequences shown in Reagents and Solutions of this unit (see recipe for linkers).

19. Incubate tubes 2 min at 50°C, then let sit 15 min at room temperature.
20. Add 2 μl high-concentration T4 DNA ligase and incubate 2 hr at 16°C with occasional mixing.
21. Place the tubes on the magnetic rack for 1 min and wash the beads in each tube three times, each time with 500 μl of 1 \times BW buffer.
22. Pool the beads from tubes 1 and 2 in a new tube, then wash once with 500 μl of 1 \times BW buffer and twice with 1 \times Buffer 4 using the magnetic rack.

Release tags by MmeI digestion and ligate tags to form ditags

23. Remove 1 \times Buffer 4 from the tube and resuspend the sample with the following mix:

- 175 μl LoTE buffer
- 20 μl 10 \times Buffer 4
- 0.3 μl 32 mM SAM
- 5 μl *MmeI*.

Incubate for 2 hr at 37°C.

24. Put the sample tube on the magnetic rack for 2 min to immobilize the beads, then transfer the supernatant to a new tube. Repeat this process three times to obtain a supernatant free of beads.
25. Extract with an equal volume of PC8 and then precipitate the sample with ethanol by combining the following:
- 200 μl sample
 - 3 μl 20 mg/ml glycogen
 - 150 μl 7.5 M ammonium acetate
 - 600 μl 100% ethanol.

Vortex briefly and place at -80°C for 15 min.

26. Microcentrifuge 30 min at maximum speed, room temperature. Wash by adding 500 μl of 70% ethanol and microcentrifuging 5 min at maximum speed, room temperature. Remove ethanol and resuspend pellet in 10 μl LoTE buffer.
27. Prepare two tubes of ligation mixture as follows:
- 5 μl sample DNA
 - 1 μl 10 \times ligase buffer
 - 1 μl high-concentration (2000 U/ μl) T4 DNA ligase (or 1 μl H₂O)
 - 3 μl H₂O.

Use ligase in one tube and water in the other as a no-ligase control. Incubate tubes overnight at 16°C.

28. Add 10 μl LoTE buffer to make final 20 μl .

Perform PCR amplification of ditags

29. Prepare 1/100 to 1/600 dilutions of sample to optimize amplification. Prepare the same dilution series for the no-ligase control.
30. Set up a series of PCR reactions using the following components:
 - 5 μ l 10 \times PCR buffer
 - 3 μ l 25 mM MgCl₂
 - 1 μ l 350 mM primer 1
 - 1 μ l 350 mM primer 2
 - 1 μ l ligation product (various dilutions)
 - 0.5 μ l standard dNTP mix (25 mM each dNTP)
 - 1 μ l 5 U/ μ l *Taq* DNA polymerase
 - 37.5 μ l H₂O.

31. Perform PCR amplification in thermal cycler with following parameters:

1 cycle:	2 min	94°C	(denaturation)
26 cycles:	40 sec	94°C	(denaturation)
	1 min	55°C	(annealing)
	1 min	70°C	(extension)
1 cycle:	5 min	70°C	(final extension).

32. Remove a 5- μ l aliquot of PCR product from each dilution and run on a 12% nondeaturing polyacrylamide/TBE gel with a 10-bp DNA ladder (*UNIT 2.7*).

The purpose of testing a dilution series of the ditag ligation is to avoid possible saturation of the PCR amplification. In the series of amplified ditag bands (120 bp in size), the band intensity will likely begin to decrease with one of the dilutions. Use this dilution factor to prepare the master mix and perform large-scale PCR amplifications in three 96-well plates. There should be no amplified ditags from the no-ligase control samples.

33. Perform a large-scale PCR reaction (three 96-well PCR plates with 50- μ l reaction volume per well) to amplify ditags using the optimal dilution and the conditions described in step 31.

Isolate ditags

34. Pool PCR reactions into two 50-ml conical tubes (~7 ml in each tube), extract with an equal volume of PC8, and precipitate by combining the following:

- 7 ml PCR sample
- 100 μ l 20 mg/ml glycogen
- 5 ml 7.5 M ammonium acetate
- 21 ml 100% ethanol.

Vortex briefly and place at -80°C for 15 min.

35. Centrifuge 30 min at 1600 \times g, room temperature, in a tabletop centrifuge with swinging-bucket rotor. Wash by adding 5 ml of 70% ethanol, centrifuging 5 min at 1600 \times g, room temperature, and removing the ethanol. Resuspend each pellet in 150 μ l LoTE buffer.
36. Pool the samples and add 60 μ l of 6 \times sample loading buffer. Load 10 μ l of the sample into each well of four 12% polyacrylamide 10-well gels and run gels at 160 V for ~2.5 hr (*UNIT 2.7*) to separate the 120-bp ditag band from the 100-bp linker-linker dimer band.

Before running all of the samples, run a test gel run using 5 μ l of pooled PCR reaction to get an idea of the maximum separation of the 120-bp ditag band and the 100-bp linker-linker dimer band.

37. Stain gels using SYBR Green or ethidium bromide (UNIT 2.7) and visualize under UV illumination.
38. Cut out only the amplified 120-bp ditag bands from the gels and place the bands in six 0.5-ml microcentrifuge tubes that have a hole pierced through the bottom by a 25-G needle.
39. Place the 0.5-ml tubes inside 2-ml microcentrifuge tubes and microcentrifuge for 3 min at maximum speed to force the gels through the holes and thereby break them up into small fragments. Discard the 0.5-ml tubes.
40. Add 500 μ l of a 5:1 mixture of LoTE buffer and 7.5 M ammonium acetate to each 2-ml tube. Vortex, then incubate 1 hr at 65°C (or overnight at 4°C).
41. Wet twelve Spin-X centrifuge tube filters by adding 5 μ l LoTE buffer and transfer the contents of each 2-ml tube to two Spin-X tubes.
42. Microcentrifuge 5 min at maximum speed and pool all the eluates in a 15-ml tube.
43. Precipitate combined eluate with ethanol by combining the following:
 - 3 ml sample
 - 50 μ l 20 mg/ml glycogen
 - 2 ml 7.5 M ammonium acetate
 - 9 ml 100% ethanol.Vortex briefly and place at -80°C for 15 min.
44. Centrifuge 30 min at $1600 \times g$, room temperature, in a tabletop centrifuge with a swinging-bucket rotor. Wash by adding 5 ml of 70% ethanol, centrifuging 5 min at $1600 \times g$, room temperature, and removing the ethanol. Resuspend sample DNA pellet in 200 μ l LoTE buffer.
45. Prepare two 1.5-ml microcentrifuge tubes as follows:
 - 100 μ l sample
 - 30 μ l 10 \times Buffer 4
 - 3 μ l 100 \times BSA
 - 20 μ l *Nla*III
 - 147 μ l H₂O.Incubate 2 hr at 37°C.
46. Extract with an equal volume of PC8, then precipitate with ethanol by combining the following:
 - 300 μ l sample
 - 5 μ l 20 mg/ml glycogen
 - 200 μ l 7.5 M ammonium acetate
 - 900 μ l 100% ethanol.Vortex briefly and place at -80°C for 15 min.
47. Microcentrifuge 30 min at maximum speed, room temperature. Wash by adding 500 μ l of 70% ethanol, microcentrifuging 5 min at maximum speed, room temperature, and removing the supernatant. Resuspend each pellet in 38 μ l LoTE buffer.

48. Pool the samples in one tube and add 15 μ l of 6 \times sample loading dye.
49. Load 10 μ l of sample into each of nine lanes of a 12% polyacrylamide gel with a 10-bp DNA ladder as a size marker in a separate lane. Run gel at 160 V for \sim 2.5 hr to separate the ditag band (\sim 40 bp) and the linker band (\sim 50 bp).
50. Stain the gel and cut out the 40-bp band from each of the nine lanes under UV illumination. Place into three 0.5-ml tubes with a hole pierced through the bottom by a 25-G needle, and break up the gels as in step 39.
51. Add 300 μ l of 5:1 mixture of LoTE buffer and 7.5 M ammonium acetate to each 2-ml tube. Vortex and incubate tubes for 2 hr at 42°C (or overnight at 4°C).
52. Wet three Spin-X centrifuge tube filters by adding 5 μ l LoTE buffer and transfer contents of each 2-ml tube to a single Spin-X tube. Microcentrifuge 5 min at maximum speed to filter the eluted DNA.
53. Extract with an equal volume of PC8, then precipitate with ethanol as follows:
 - 300 μ l sample
 - 5 μ l 20 mg/ml glycogen
 - 200 μ l 7.5 M ammonium acetate
 - 900 μ l 100% ethanol.Vortex briefly and place at -80°C for 15 min.
54. Microcentrifuge 30 min at maximum speed, room temperature. Wash by adding 500 μ l of 70% ethanol, microcentrifuging 5 min at maximum speed, room temperature, and removing the supernatant. Resuspend each pellet in 2.5 μ l cold LoTE buffer and pool (total 7.5 μ l).

Ligate ditags to form concatamers

55. Prepare the following:
 - 7.5 μ l ditag pool
 - 1 μ l 10 \times ligase buffer
 - 0.5 μ l H₂O
 - 1 μ l high-concentration (2000 U/ μ l) T4 DNA ligase.

Incubate at 16°C for 30 min to get appropriate-size concatamers for sequencing.

Concatamers that are too short will not yield sufficient numbers of tags from one sequencing reaction. Concatamers that are larger than 1.5 kb cannot be entirely sequenced in a single sequencing reaction even if sequenced from both ends. For this reason, concatamers that are 600 bp to 1.5 kb in size are most useful. In the authors' experience, a 30-min incubation is sufficient for obtaining the optimal size of concatamers. The incubation time during the ligation may be adjusted to ensure that the bulk of the concatamers are in this range.

56. After ligation, resolve the concatamers by running the sample in one lane on an 8% polyacrylamide gel, with a 1-kbp DNA ladder in a separate well. Run sample for 45 min at 200 V.
57. Stain the gel to visualize the DNA. From the smear of concatamers, cut out the portion of the gel lane that contains the 600- to 1.5-kbp size range of concatamers.
58. Place gel slices into two 0.5-ml tubes with holes and break up the gel slices by centrifugation.
59. Add 300 μ l of a 5:1 mixture of LoTE buffer and 7.5 M ammonium acetate to each 2-ml tube. Vortex and incubate tubes for 1 hr at 42°C (or overnight at 4°C).

60. Wet two Spin-X centrifuge tube filters by adding 5 μl of LoTE buffer and transfer the contents of each 2 ml tube to a Spin-X tube. Microcentrifuge 5 min at maximum speed to filter the eluted DNA.
61. Extract with an equal volume of PC8, then precipitate with ethanol as follows:
 - 300 μl sample
 - 5 μl 20 mg/ml glycogen
 - 200 7.5 M ammonium acetate
 - 900 μl 100% ethanol.Vortex briefly and place at -80°C for 15 min.
62. Microcentrifuge 30 min at maximum speed, room temperature. Wash by adding 500 μl of 70% ethanol, microcentrifuging 5 min at maximum speed, room temperature, and removing the supernatant. Resuspend each pellet in 3 μl of LoTE buffer (total 6 μl).

Ligate concatamers and vector

63. Prepare linearized plasmid by combining the following:
 - 1 μl 1 $\mu\text{g}/\mu\text{l}$ pZErO-1 vector
 - 1.5 μl 10 \times Buffer 2
 - 1.5 μl *Sph*I
 - 11 μl H₂O.Incubate 25 min at 37 $^{\circ}\text{C}$.

Do not incubate >30 min, as overdigestion will reduce cloning efficiency.
64. Add 185 μl LoTE. Extract with an equal volume of PC8 and precipitate with ethanol by combining the following:
 - 200 μl linearized plasmid
 - 3 μl 20 mg/ml glycogen
 - 150 μl 7.5 M ammonium acetate
 - 600 μl 100% ethanol.Vortex briefly and place at -80°C for 15 min.
65. Centrifuge 30 min in a tabletop centrifuge at maximum speed at room temperature. Wash with 500 μl of 70% ethanol and resuspend the pellet in 10 μl of LoTE buffer.
66. Prepare the following reaction mixture for ligation:
 - 6 μl purified concatamer (or 6 μl H₂O for vector-alone control)
 - 1 μl linearized pZErO-1 vector
 - 1 μl 10 \times ligase buffer
 - 1 μl H₂O
 - 1 μl T4 DNA ligase.Incubate overnight at 12 $^{\circ}\text{C}$.
67. Add 200 μl LoTE buffer. Extract with an equal volume of PC8, then precipitate with ethanol by combining the following:
 - 200 μl sample
 - 3 μl 20 mg/ml glycogen
 - 150 μl 7.5 M ammonium acetate
 - 600 μl 100% ethanol.Vortex briefly and place at -80°C for 15 min.

68. Microcentrifuge 30 min at maximum speed, room temperature. Wash by adding 500 μ l of 70% ethanol, microcentrifuging 5 min at maximum speed, and removing the supernatant. Resuspend pellet in 10 μ l LoTE buffer.

Transform E. coli and check insert sizes

69. Thaw ElectroMAX DH10B electroporation-competent cells on ice. Pipet 40 μ l of the competent cells and 5 μ l DNA into an electroporation cuvette.
70. Perform electroporation with Bio-Rad GenePulser electroporator at 100 Ω , 25 μ F, and 1.8 kV (also see *UNIT 1.8*). Transfer electroporated cells into a 1.5-ml microcentrifuge tube, add 1 ml SOC medium at room temperature, and shake for 20 min at 225 rpm at 37°C.
71. Plate one-fifth of the transformed cells onto each of five 10-cm Zeocin-containing low-salt LB plates. Incubate at 30°C and analyze 14 to 18 hr later.

Save all plates for future use. If insert sizes are appropriate, these may be used for further sequencing.

72. To check insert sizes, set up 50- μ l PCR reactions in a 96-well PCR plate as follows:

5 μ l 10 \times PCR buffer
3 μ l 25 mM MgCl₂
2.5 μ l DMSO
0.7 μ l standard dNTP mix (25 mM each dNTP)
1 μ l 5 μ M M13 forward primer
1 μ l 5 μ M M13 reverse primer
35.8 μ l H₂O
1 μ l 5 U/ μ l *Taq* DNA polymerase.

Use a sterile toothpick to pick each clone and transfer it to a well of the plate.

73. Perform amplification as follows:

1 cycle:	2 min	94°C	(denaturation)
25 cycles:	30 sec	94°C	(denaturation)
	1 min	55°C	(annealing)
	90 sec	72°C	(extension)
1 cycle:	5 min	72°C	(final extension).

74. Run 5 μ l of each PCR reaction on a 1.5% agarose gel and select the clones that have concatamer inserts of the appropriate size for subsequent sequencing reactions.

75. Perform sequencing reactions and analyze the sequencing data (*UNIT 15.2*).

Instead of sequencing the colony PCR product as described here, 96-well cultures can be grown from recombinant clones and frozen as glycerol stocks. These stocks can be used for subsequent PCR reactions or plasmid preparation to generate templates for sequencing.

SUBTRACTION STAGE (SubSTAGE)

Nonspecific genomic DNA carried through in the ChIP sample can also be represented as STAGE tags. Applying a subtraction step to the original ChIP DNA material increases the relative enrichment of target loci, and this increased enrichment can be verified by ChIP-PCR (*UNIT 21.3*). The Subtraction STAGE (SubSTAGE) protocol is used to subtract background from ChIP DNA with an excess of genomic DNA.

**SUPPORT
PROTOCOL**

**Chromatin
Assembly and
Analysis**

21.10.11

Additional Materials (see also *Basic Protocol*)

- Unbiotinylated round B primer (see recipe for PCR primers; omit biotinylation)
- 1 M Tris·Cl, pH 7.5 (*APPENDIX 2*)
- 0.5 M EDTA (*APPENDIX 2*)
- 1 M NaOH (*APPENDIX 2*)
- 72° and 95°C water baths or heating blocks
- Additional reagents and equipment for ChIP-PCR (*UNIT 21.3*)

Anneal ChIP DNA and genomic DNA

1. Amplify ChIP DNA samples as described in the *Basic Protocol*, steps 1 to 11, but use a normal, unbiotinylated round B primer in step 6 and perform 20 amplification cycles in step 7. Use 1 µg of this amplified ChIP DNA as a “tester” in the subtraction step below.

Here, “tester” means the amplified sample ChIP DNA and “driver” means amplified genomic DNA used to remove nonspecific DNA fragments from the tester.

2. Amplify sheared genomic DNA samples (average length, 1 kb) as described in the *Basic Protocol*, steps 1 to 11, using a biotinylated primer B and resuspending the pellet in water rather than LoTE buffer at step 11 of the *Basic Protocol*. Use a 5×, 10×, or 15× excess of this amplified genomic DNA (5 µg, 10 µg, or 15 µg) as a “driver.”

All three “drivers” may be used for subsequent steps. After finishing subtraction, enhancement of enrichment can be tested as described in step 17, below.

3. Incubate driver genomic DNA at 72°C for 15 min in a 200-µl solution containing either 5, 10, or 15 µg genomic DNA, 10 mM Tris·Cl, pH 7.5 (added from 1 M stock), 1.0 mM EDTA, and 100 mM NaOH.

The purpose of this step is to denature dsDNA fragments. Unbiotinylated ssDNA will be removed in subsequent steps.

4. Wash four 100-µl aliquots of Dynabeads M-280 Streptavidin beads three times as described in the *Basic Protocol*, step 15.
5. Add driver genomic DNA to one aliquot of prewashed beads and incubate at room temperature for 20 min with occasional mixing.
6. Immobilize driver genomic DNA using a magnetic rack and remove the supernatant.
7. Wash the beads/driver genomic DNA three times with 1× BW buffer using the magnetic rack.

This step removes unbiotinylated genomic DNA.

8. Dissolve 1 µg tester DNA in 200 µl of 1× BW buffer and add this to the complex of beads and driver genomic DNA.
9. Denature the sample by incubating at 95°C for 5 min and cool slowly to room temperature to form tester-driver heteroduplexes.
10. Add the second aliquot of prewashed beads to the reaction and incubate at room temperature for 20 min with occasional mixing.

Remove heteroduplexes

11. Immobilize the beads with tester-driver heteroduplexes using the magnetic rack and transfer the supernatant to a new tube.
12. Add the third aliquot of prewashed beads to the tube that contains the supernatant, incubate at room temperature for another 20 min, and remove the beads remaining in the supernatant using the magnetic rack. Transfer the supernatant to a new tube.

13. Add the fourth aliquot of prewashed beads, perform a third round of magnetic bead removal, and then transfer the final subtracted DNA fragments to a new tube.

Isolate subtracted DNA

14. Extract sample with an equal volume (200 μ l) of PC8.

15. Precipitate with ethanol by combining the following:

- 200 μ l subtracted ChIP DNA
- 2 μ l 20 mg/ml glycogen
- 150 μ l 7.5 M ammonium acetate
- 600 μ l 100% ethanol.

Vortex briefly and place at -80°C for 15 min.

16. Microcentrifuge 30 min at maximum speed, room temperature. Wash by adding 500 μ l of 70% ethanol, microcentrifuging 5 min at maximum speed, room temperature, and removing the ethanol. Resuspend pellet in 20 μ l LoTE buffer.
17. Test the enrichment of the original ChIP DNA and subtracted ChIP DNA by ChIP-PCR as described in *UNIT 21.3*.
18. Use this subtracted ChIP DNA as the input for STAGE when it shows better enrichment than the original ChIP DNA (see Basic Protocol, step 1).

REAGENTS AND SOLUTIONS

Use autoclaved double-distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

BW buffer, 1 \times

For 2 \times stock:

10 mM Tris \cdot Cl, pH 7.5 (*APPENDIX 2*)

1 mM EDTA

2.0 M NaCl

Store up to 1 year at room temperature

Dilute to 1 \times with H₂O just before use

Linkers

Linker 1A: 5'-TTTGGATTTGCTGGTGCAGTACAACCTAGGCTTAATATCCGACATG-3'

Linker 1B: 5'-TCGGATATTAAGCCTAGTTGTACTGCACCAGCAAATCC (aminomod C7)-3'

Linker 2A: 5'-TTTCTGCTCGAATTCAAGCTTCTAACGATGTACGTCCGACATG-3'

Linker 2B: 5'-TCGGACGTACATCGTTAGAAGCTTGAATTCGAGCAG (aminomod C7)-3'

In linkers 1B and 2B, "aminomod C7" refers to a 3' amino modification. A primary amine group is attached to the 3' end of a nucleotide through a C7 carbon-chain arm. It blocks a possible ligation reaction at the 3' terminus.

The authors recommend ordering oligonucleotides from Integrated DNA Technologies.

LoTE buffer

3 mM Tris \cdot Cl, pH 7.5 (*APPENDIX 2*)

0.2 mM EDTA, pH 7.5

Store up to 1 year at room temperature

PC8

Combine the following in the order indicated:

480 ml phenol, warmed to 65°C
320 ml 0.5 M Tris·Cl, pH 8.0 (*APPENDIX 2*)
640 ml chloroform

Shake to mix and place at 4°C. After 2 to 3 hr, shake again. After an additional 2 to 3 hr, aspirate aqueous layer. Store up to 1 year in aliquots at -20°C or up to 6 months at 4°C.

Commercially available 1:1 (v/v) phenol:chloroform mix can also be substituted, as long as the pH is preset to 8.0.

PCR buffer, 10×

500 mM KCl
100 mM Tris·Cl, pH 8.3 (*APPENDIX 2*)

PCR primers

Round A primer: 5'-GTTTCCCAGTCACGATCNNNNNNNNN-3'
Round B primer: 5'-Biotin-GTTTCCCAGTCACGATCACGATC-3'
Primer 1: 5'-GTGCTCGTGGGATTTGCTGGTGCAGTACA-3'
Primer 2: 5'-GAGCTCGTGCTGCTCGAATTCAAGCTTCT-3'
M13 forward primer: 5'-GTAAAACGACGGCCAGT-3'
M13 reverse primer: 5'-GGAAACAGCTATGACCATG-3'

The authors recommend ordering oligonucleotides from Integrated DNA Technologies.

Sample loading buffer, 6×

0.25% (w/v) bromphenol blue
0.25% (w/v) xylene cyanol FF
30% (v/v) glycerol

SOC medium

0.5% yeast extract
2% tryptone
10 mM NaCl
2.5 mM KCl
10 mM MgCl₂
10 mM MgSO₄
20 mM glucose
Store at room temperature (stable for years)

Zeocin-containing low-salt LB plates

For 1 liter:
10 g tryptone
5 g yeast extract
5 g NaCl
Adjust pH to 7.5
Add 15 g Bacto agar
Autoclave and allow to cool
Add Zeocin to 100 mg/ml

COMMENTARY

Background Information

DNA-binding proteins such as transcription factors, activators, and repressors control the expression of target genes by binding to specific genomic target loci. Mapping the chromosomal target loci of transcription factors is important for deciphering gene regulatory networks that mediate cellular responses. Conventionally, a small number of target loci of various transcription factors have been determined by chromatin immunoprecipitation (ChIP) and subsequent PCR (UNIT 21.3). More recently, methods combining ChIP and DNA microarrays have allowed researchers to define the direct chromosomal targets of transcription factors *in vivo* on a genome-wide scale (Ren et al., 2000; Iyer et al., 2001; UNIT 21.9). It is not yet feasible, however, to apply this method on a truly comprehensive scale in organisms with large and complex genomes, because whole-genome microarrays are not available for most organisms (Ren et al., 2002; Weinmann et al., 2002). Sequence tag analysis of genomic enrichment (STAGE) is a genome-wide method for identifying the chromosomal targets of DNA-binding proteins in any sequenced genome without using microarrays (Kim et al., 2005). STAGE is conceptually derived from serial analysis of gene expression (SAGE; Velculescu et al., 1995; Saha et al., 2002), but is applied to DNA isolated after ChIP.

STAGE does not rely on any assumptions about the location and distribution of these binding sites in the genome. It is therefore a useful method for identifying functional elements in the genome. The coverage, or comprehensiveness, of this method is limited in principle only by the depth of sequencing, which is a reliable, widely available, automatable, inexpensive, and high-throughput procedure. The fact that sequencing is performed on concatamerized tags makes this a truly genomic method that can be quantitative. By analogy to SAGE, the number of times one observes a given tag within a STAGE pool or library should be directly proportional to the enrichment of that sequence in the original ChIP DNA pool. Since it is not dependent on microarray resources, STAGE is applicable to any sequenced genome.

Critical Parameters and Troubleshooting

For successful STAGE, it is critical to start with high-purity ChIP material that shows high

enrichment of target loci relative to background (UNIT 21.3). The quality of the ChIP DNA material should be tested by conventional ChIP-PCR before starting the STAGE procedure.

Obtaining good separation of bands on gels and identification and isolation of the correct band away from nonspecific bands during all gel-extraction steps are very important to this procedure. In case the isolation of the right band is not successful, a second gel-purification step should be carried out.

It is possible that a proportion of STAGE tags will not match any sequence within the given genome. This can occur because of contaminating DNA in the original ChIP sample or because of PCR or sequencing errors. To minimize contamination from extraneous DNA, use fresh solutions for every step and prepare aliquots for all solutions. Use aerosol-barrier pipet tips throughout the procedure.

Anticipated Results

If the Basic Protocol works without any problems, ~90% of colonies should contain inserts of the appropriate size. The majority of inserts should contain ~25 tags, so ~50,000 STAGE tags will be obtained from 2000 sequencing reactions. The number of STAGE tags that need to be sequenced to achieve complete coverage of the targets of a protein will depend on many factors, including the actual number of targets and the quality of the ChIP reaction.

Time Considerations

Biotinylation of ChIP DNA using the two-round amplification procedure usually takes 3 to 4 hr. From this point to ditag ligation takes ~8 to 10 hr, excluding the overnight ligation step. Optimizing PCR and performing the large-scale amplification takes 6 to 8 hr. Ditag purifications, including the two gel extraction steps, take 6 to 8 hr on 2 days. Concatamer ligation, purification, and subcloning take 6 to 8 hr. Cleanup of ligation products and transformation takes 3 to 4 hr. Colony PCR and gel analysis take 5 to 6 hr. After obtaining the desired number of colonies with appropriate insert sizes, sequencing can be carried out. The laborious step of picking hundreds of *E. coli* transformants and assaying for insert sizes may be circumvented by using an automated colony picker if one is available, or by using commercial services that can carry out this part of the operation.

Literature Cited

- Iyer, V.R., Horak, C.E., Scafe, C.S., Botstein, D., Snyder, M., and Brown, P.O. 2001. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* 409:533-538.
- Kim, J., Bhing, A.A., Morgan, X.C., and Iyer, V.R. 2005. Mapping DNA-protein interactions in large genomes by sequence tag analysis of genomic enrichment. *Nat. Methods* 2:47-53.
- Ren, B., Robert, F., Wyrick, J.J., Aparicio, O., Jennings, E.G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E., Volkert, T.L., Wilson, C.J., Bell, S.P., and Young, R.A. 2000. Genome-wide location and function of DNA binding proteins. *Science* 290:2306-2309.
- Ren, B., Cam, H., Takahashi, Y., Volkert, T., Terragni, J., Young, R.A., and Dynlacht, B.D. 2002. E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. *Genes Dev.* 16:245-256.
- Saha, S., Sparks, A.B., Rago, C., Akmaev, V., Wang, C.J., Vogelstein, B., Kinzler, K.W., and Velculescu, V.E. 2002. Using the transcriptome to annotate the genome. *Nat. Biotechnol.* 20:508-512.
- Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. 1995. Serial analysis of gene expression. *Science* 270:484-487.
- Weinmann, A.S., Yan, P.S., Oberley, M.J., Huang, T.H., and Farnham, P.J. 2002. Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev.* 16:235-244.

Contributed by Jonghwan Kim and
Vishwanath R. Iyer
University of Texas at Austin
Austin, Texas

Mapping Chromatin Interactions by Chromosome Conformation Capture

This unit describes chromosome conformation capture (3C), a PCR-based method to analyze chromosome conformation (Dekker et al., 2002). This methodology is used to detect the frequencies of interactions between chromosomal fragments inside intact cells and is one of only a few techniques that allows for high-resolution analysis of chromosome organization. Specifically, 3C is widely and successfully used for detection of looping interactions between sequence elements, e.g., between gene promoters and distant regulatory elements such as enhancers (Tolhuis et al., 2002; Spilianakis and Flavell, 2004; Vakoc et al., 2005).

In this assay (depicted in Fig. 21.11.1), cells are treated with formaldehyde to induce both protein-protein and protein-DNA cross-links, which can result in covalently linking interacting chromatin segments. Cross-linked chromatin is then solubilized and digested by a restriction enzyme. The digestion is followed by ligation under dilute DNA concentrations, which strongly promotes intramolecular ligation of cross-linked fragments over intermolecular ligation of random fragments. The cross-links are then reversed and DNA is purified. The resulting 3C template contains a large number of ligation products, each representing a specific interaction between two genomic loci. The relative abundance of a particular ligation product reflects the frequency with which the two restriction fragments interact inside the cell nucleus.

Ligation products are detected and quantified by semi-quantitative PCR. A control template must also be generated to contain all possible ligation products in equal amounts. The control template is used to correct for differences in primer efficiencies. A ratio is then determined between the amount of ligation product of the 3C template versus the amount of ligation product of the control template. This value, the interaction frequency, is a measure of how frequently two chromosomal fragments interact inside the nucleus. Once a sufficient number of interaction frequencies have been obtained for a genomic region of interest, its spatial organization can be inferred.

This unit describes the protocols for generating 3C and control templates for both *Saccharomyces cerevisiae* (see Basic Protocols 1 and 2) and mammalian cells (see Basic Protocols 3 and 4). Also included in this unit is a description of how to analyze ligation products by PCR (see Basic Protocol 5).

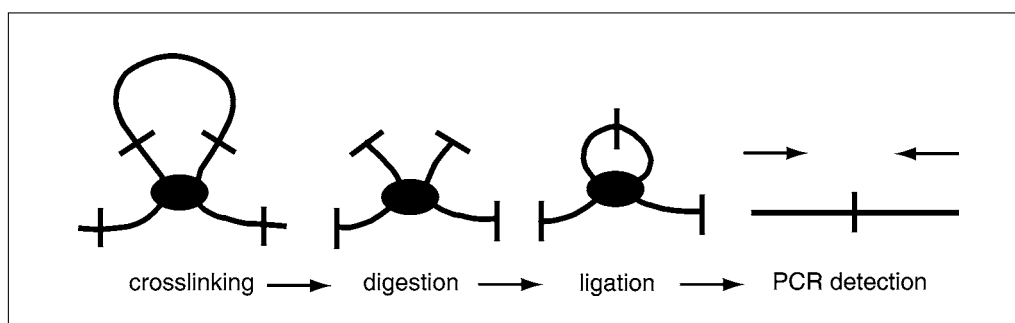


Figure 21.11.1 Schematic representation of the chromosome conformation capture (3C) technology. Cells are treated with formaldehyde to induce a cross-link resulting in covalent linkage of interacting chromatin segments (indicated by the solid center oval). Cross-linked chromatin is digested and ligated under dilute DNA concentrations. After reversal of cross-links, ligation products can be detected by semi-quantitative PCR (arrows indicate PCR primers).

STRATEGIC PLANNING

The 3C technique has proven to be exceptionally powerful and is now widely used, at least in part because it can be performed in any molecular biology laboratory and does not require any special equipment. The preparation of the assay should involve careful thought and planning. Restriction enzyme selection and primer design play pertinent roles and should be considered before the start of the experiment.

Restriction Enzyme Selection

The restriction enzyme used for digestion can vary. As mentioned previously, 3C can be used to detect looping interactions between sequence elements. It is important that the restriction enzyme cut sites are spaced evenly and frequently throughout the region of interest. For proper analysis, the putative looping elements should each be contained in a restriction fragment no larger than 10 kb and no smaller than 1 kb in size, because very large and very small fragments result in slightly higher and lower interaction frequencies, respectively. Furthermore, there should be multiple restriction fragments between the elements to obtain a high-resolution looped structure. Once a looping interaction is found, a different enzyme should be selected to confirm the results. For initial studies, a restriction enzyme that cuts every 4 kb or so is recommended. Once a looped structure is detected, a more frequently cutting enzyme can be used to obtain a more precise map.

Primer Design

Once a restriction enzyme is chosen, primers must be designed throughout the genomic region of interest. Primers are designed unidirectionally at a site that is ~80 to 150 bp 5' of the restriction enzyme cut site. Head-to-head ligation of two restriction fragments can then be amplified by two primers, resulting in a PCR product that will be 160 to 300 bp long.

Typically, primers are 28 bp in length, have a GC content of ~50%, and ideally have similar melting temperatures. Primer combinations should be tested using a control template (see Basic Protocols 2 and 4), and primer pairs that do not amplify the correct product, that amplify multiple products, that do not yield equal amounts of PCR products when using control template, or that produce high levels of primer dimers should be omitted and redesigned, if so desired. Primers should not be designed on restriction fragments > 10 kb or < 1 kb in size.

GENERATION OF 3C TEMPLATE FROM INTACT YEAST CELLS

This protocol outlines the method to obtain a 3C template from intact *Saccharomyces cerevisiae* cells. In this protocol, intact yeast cells are spheroplasted and the cell wall is digested with zymolyase. These spheroplasts are cross-linked with formaldehyde, and the chromatin is solubilized and digested with a restriction enzyme. After digestion, the cross-linked chromatin is ligated, cross-links are reversed, and DNA is purified. The resulting 3C template consists of a collection of ligation products. The abundance of a given ligation product depends on the distance between the two genomic segments (when they are located on the same chromosome) and, importantly, also depends on any specific looping interactions (see Fig. 21.11.4).

Materials

- Saccharomyces cerevisiae* cells of interest (UNIT 13.2)
- Spheroplasting buffer I (see recipe)
- 20 mg/ml zymolyase 100-T solution (see recipe)
- MES wash buffer (see recipe)
- 37% (v/v) formaldehyde
- 2.5 M glycine

BASIC PROTOCOL 1

Chromosome Conformation Capture

21.11.2

Restriction enzyme and corresponding 10× restriction enzyme buffer
 1% and 10% (w/v) sodium dodecyl sulfate (SDS; *APPENDIX 2*)
 10% (v/v) Triton X-100
 10× ligation buffer (see recipe)
 10 mg/ml bovine serum albumin (BSA)
 100 mM adenosine triphosphate (ATP)
 T4 DNA ligase
 10 mg/ml proteinase K in TE buffer, pH 8.0
 1:1 (v/v) phenol/chloroform (*UNIT 2.1A*)
 3 M sodium acetate, pH 5.2 (*APPENDIX 2*)
 100% ethanol
 TE buffer, pH 8.0 (*APPENDIX 2*)
 10 mg/ml DNase-free RNase A (*UNIT 3.13*)
 15-, 50-, and 250-ml disposable conical tubes
 Roller drum in 30°C incubator
 Refrigerated tabletop centrifuge
 1.7-ml microcentrifuge tubes
 16°, 37°, 42°, and 65°C water baths
 30-ml screw-cap centrifuge tubes

Prepare and cross-link cells

1. Obtain a 200-ml culture of *Saccharomyces cerevisiae* cells at $OD_{600} = 1$ (*UNIT 13.2*).
2. Centrifuge cells in a 250-ml disposable conical tube for 10 min at $1250 \times g$ (2500 rpm), room temperature, and remove supernatant promptly.
3. Resuspend cells in 10 ml spheroplasting buffer I and transfer to a 15-ml disposable conical tube. Add 50 μ l of 20 mg/ml zymolyase 100-T solution and gently mix tube.

The entire concentration of 20 mg/ml zymolyase 100-T will not go into solution. Make up the zymolyase solution at least 1 day prior to use. Before zymolyase is added to the cells, be sure that the solution is mixed well and in a suspension.

4. Incubate in roller drum for 40 min at 30°C.

The efficiency of cell wall digestion should be tested by cell lysis. This can be done by adding water to a small amount of cells on a glass slide while observing under a microscope. Digestion is complete when ~80% of cells burst open and exhibit hypotonic lysis within 1 to 2 min.

5. Centrifuge cells 5 min at $2460 \times g$ (3500 rpm), room temperature, in a refrigerated tabletop centrifuge. Wash cells two times in 10 ml MES wash buffer, centrifuging again after each wash.
6. Dissolve cell pellet in 10 ml MES wash buffer.
7. Add 263 μ l of 37% formaldehyde (1% final), mix thoroughly, and incubate 10 min at room temperature.
8. Add 0.5 ml of 2.5 M glycine and incubate 5 min at room temperature.

Cells can now be stored for up to 1 year at -80°C in 1-ml aliquots or can directly be digested.

Digest cross-linked cells

9. Add 50 μ l of cross-linked cells to 40 individual 1.7-ml microcentrifuge tubes.

Reactions should not be pooled. Typically, tubes containing 50 μ l of cells yield the best results, although the total amount of cells used can vary depending on need. When 40 tubes are used, the obtained 3C template is sufficient to measure ~350 interactions.

10. Wash cells three times with 100 μ l of 1 \times restriction enzyme buffer per tube. For each wash, mix by pipetting up and down, centrifuge 3 min at 18,000 \times g (14,000 rpm), room temperature, and remove the supernatant.

Restriction enzyme buffer should correspond with enzyme of choice, as recommended by the manufacturer.

11. Resuspend pellet thoroughly in 36.2 μ l of 1 \times restriction enzyme buffer per tube, add 3.8 μ l of 1% SDS per tube, and incubate for 10 min at 65°C.

This step proves to be essential in template generation. Care should be taken to ensure proper incubation temperature.

12. Add 4.4 μ l of 10% Triton X-100 per tube. Mix well by pipetting up and down.

Triton X-100 binds to SDS and will thereby effectively remove SDS, which is required for subsequent restriction digestion.

13. Add 60 U restriction enzyme per tube, mix well, and incubate reactions overnight at an appropriate temperature.

The temperature should correspond with enzyme of choice, as recommended by the manufacturer.

14. Add 8.6 μ l of 10% SDS per tube and incubate 20 min at 65°C.

SDS is added at high concentrations to inactivate the restriction enzyme.

Ligate cross-linked cells

15. Add the following to each tube and incubate for 2 hr at 16°C:

74.5 μ l of 10% Triton X-100
74.5 μ l of 10 \times ligation buffer
8 μ l of 10 mg/ml BSA
8 μ l of 100 mM ATP
596 μ l distilled water
800 Weiss units T4 DNA ligase.

Due to the large number of tubes, a master mix of all components is typically made for all ligations and distributed amongst the tubes.

Reverse cross-links and purify 3C template

16. Add 5 μ l of 10 mg/ml proteinase K in TE buffer, pH 8.0, and incubate overnight at 65°C to reverse cross-links.

17. Add an additional 5 μ l of 10 mg/ml proteinase K in TE buffer, pH 8.0, and incubate 2 hr at 42°C.

An additional proteinase K treatment improves DNA purification.

18. Combine ten reactions in 50-ml disposable conical tubes to end up with four larger pooled reactions (if starting with 40 tubes).

Do not pool more than ten reactions, as template quality will decrease.

19. Add an equal volume of 1:1 (v/v) phenol/chloroform to each of the ligation mixtures, vortex for 30 sec, and centrifuge 5 min at 2460 \times g (3500 rpm), room temperature.

At this step in the purification, the aqueous (upper) phase will appear cloudy.

20. Promptly collect the aqueous (upper) phase, taking care not to include any of the interface layer.

21. Repeat phenol/chloroform extraction and transfer aqueous phase to a 30-ml screw-cap centrifuge tube.

At this step, the aqueous phase should be clear. If the aqueous phase is still cloudy, then repeat the phenol/chloroform extraction until the aqueous phase becomes clear.

22. Add 1/10 vol of 3 M sodium acetate, pH 5.2, and vortex briefly. Precipitate DNA by adding 2.5 vol of ice-cold 100% ethanol and mixing gently. Incubate 15 min at -80°C and centrifuge 20 min at $12,000 \times g$ (10,000 rpm), 4°C .

23. Remove supernatant and let pellets dry completely before resuspending each pellet in 50 μl TE buffer, pH 8.0. Pool all four samples to obtain a 200 μl DNA solution (this is the 3C template).

A total volume of 200 μl TE buffer, pH 8.0, should be used to resuspend the pellet when 40 tubes are initially used (in step 9). If fewer tubes are used, the final volume should be adjusted accordingly.

24. Add 2 μl of 10 mg/ml DNase-free RNase A and incubate 15 min at 37°C .

25. Store 3C template up to 2 years at -20°C .

3C templates stored for 2 years at this temperature have retained original quality.

GENERATION OF CONTROL TEMPLATE FROM YEAST GENOMIC DNA

This protocol outlines the method to obtain a control template from *Saccharomyces cerevisiae*. In this protocol, yeast chromosomal DNA is isolated, digested, and randomly ligated. This control template will contain all possible ligation products in equal amounts and serves as an excellent control for primer efficiencies. The control template is used in combination with the 3C template to determine interaction frequencies (see Basic Protocol 5).

Materials

Saccharomyces cerevisiae cells of interest (UNIT 13.2)
Spheroplasting buffer II (see recipe)
Lysing buffer I (see recipe)
20 mg/ml proteinase K in TE buffer, pH 8.0
5 M potassium acetate
80% and 100% ethanol, ice cold
TE buffer, pH 8.0 (APPENDIX 2), containing 10 $\mu\text{g}/\text{ml}$ DNase-free RNase A
1:1 (v/v) phenol/chloroform (UNIT 2.1A)
100% isopropanol
Restriction enzyme and corresponding $10\times$ restriction enzyme buffer
3 M sodium acetate, pH 5.2 (APPENDIX 2)
 $10\times$ ligation buffer (see recipe)
1 mg/ml bovine serum albumin (BSA)
10 mM adenosine triphosphate (ATP)
T4 DNA ligase
0.5 M EDTA, pH 8.0 (APPENDIX 2)
250-ml disposable conical tubes
 16° , 37° , and 65°C water baths
1.7-ml microcentrifuge tubes
Refrigerated microcentrifuge
Additional reagents and equipment for quantifying DNA by absorption spectroscopy (APPENDIX 3D)

BASIC PROTOCOL 2

Chromatin Assembly and Analysis

21.11.5

Isolate yeast chromosomal DNA

1. Obtain a saturated, overnight 200-ml culture of the same *Saccharomyces cerevisiae* strain used in Basic Protocol 1.
2. Centrifuge cells in a 250-ml disposable conical tube for 10 min at $1250 \times g$ (2500 rpm), room temperature, and remove supernatant promptly.
3. Resuspend cells in 20 ml spheroplasting buffer II. Distribute 500 μl of cells to each of 40 individual 1.7-ml microcentrifuge tubes.

Reactions should not be pooled.

4. Incubate 40 min at 37°C .

Solution should appear stringy.

5. Add 100 μl lysing buffer I to each tube.
6. Add 10 μl of 20 mg/ml proteinase K in TE buffer, pH 8.0, to each tube and incubate 30 min at 65°C .

Solution should now appear clearer.

7. Add 100 μl of 5 M potassium acetate to each tube and incubate 10 min in ice water.
8. Centrifuge tubes 20 min at $18,000 \times g$ (14,000 rpm), 4°C .
9. Transfer supernatants to fresh 1.7-ml microcentrifuge tubes containing 500 μl of ice-cold 100% ethanol. Invert tubes five times and centrifuge 10 min at $18,000 \times g$ (14,000 rpm), room temperature.
10. Carefully remove supernatant and let DNA pellet dry completely.
11. Add 500 μl TE buffer, pH 8.0, containing 10 $\mu\text{g/ml}$ DNase-free RNase A and incubate 30 min at 37°C to dissolve DNA. Occasionally tap tubes gently.

Depending on amount of DNA isolated, this step may take >30 min. Incubation should continue until all DNA is dissolved.

12. Add an equal volume of 1:1 (v/v) phenol/chloroform, vortex 30 sec, and centrifuge 5 min at $1100 \times g$ (3500 rpm), room temperature.
13. Transfer aqueous (upper) phase to fresh 1.7-ml microcentrifuge tubes. Precipitate DNA by adding 500 μl of 100% isopropanol and invert tubes five times. Centrifuge 10 min at $18,000 \times g$ (14,000 rpm), room temperature.

DNA should be visible as a string-like ball after isopropanol is added.

14. Wash DNA with 500 μl of 80% ethanol and centrifuge 10 min at $18,000 \times g$ (14,000 rpm), room temperature.
15. Dry DNA pellets, then dissolve each pellet in 100 μl TE buffer, pH 8.0. Pool samples.
16. Determine DNA concentration using absorption spectroscopy (*APPENDIX 3D*).

Digest yeast genomic DNA

17. Add 10 μg genomic DNA to each of 20 individual 1.7-ml microcentrifuge tubes.

Reactions should not be pooled.

18. Digest in 400- μl reactions using 60 U of the same restriction enzyme and $10\times$ restriction enzyme buffer used in Basic Protocol 1. Incubate 3 hr at an appropriate temperature.

Restriction enzyme buffer and incubation temperature should correspond with enzyme of choice, as recommended by the manufacturer.

19. Add an equal volume of 1:1 phenol/chloroform to each reaction, vortex for 30 sec, and centrifuge 5 min at $1100 \times g$ (3500 rpm), room temperature.
20. Transfer aqueous (upper) phase to a fresh 1.7-ml microcentrifuge tube. Add 1/10 vol of 3 M sodium acetate, pH 5.2, and vortex briefly. Precipitate DNA by adding 2.5 vol of ice-cold 100% ethanol and mixing gently. Centrifuge cells for 10 min at $18,000 \times g$ (14,000 rpm), 4°C, and carefully remove supernatant.
21. Resuspend DNA in 0.5 ml of 80% ethanol to wash pellet and centrifuge tubes 10 min at $18,000 \times g$ (14,000 rpm), 4°C. Carefully remove supernatant and let pellet dry completely.
22. Dissolve each DNA pellet in 20 μ l autoclaved water.

Ligate yeast genomic DNA

23. Add the following to each tube and incubate 1 hr at 16°C:

3 μ l of 10 \times ligation buffer
 3 μ l of 1 mg/ml BSA
 3 μ l of 10 mM ATP
 800 Weiss units T4 DNA ligase.

Due to the large number of tubes, a master mix of all components is typically made for all ligations and distributed amongst the tubes.

Purify yeast genomic DNA control template

24. Make two pools of ten reactions each to result in two larger reactions, and stop the reactions by adding 6 μ l of 0.5 M EDTA, pH 8.0.
25. Add an equal volume of 1:1 phenol/chloroform, vortex for 30 sec, and centrifuge 5 min at $1100 \times g$ (3500 rpm), room temperature.
26. Transfer aqueous (upper) phase to a fresh 1.7-ml microcentrifuge tube. Add 1/10 vol of 3 M sodium acetate, pH 5.2, and vortex briefly. Precipitate DNA by adding 2.5 vol of ice-cold 100% ethanol and mixing gently. Centrifuge cells 10 min at $18,000 \times g$ (14,000 rpm), 4°C, and carefully remove supernatant.
27. Resuspend DNA in 0.5 ml of 80% ethanol to wash pellet. Centrifuge samples 10 min at $18,000 \times g$ (14,000 rpm), 4°C. Carefully remove supernatant and let pellet dry completely.
28. Dissolve each pellet in 400 μ l TE buffer, pH 8.0, and pool them to obtain a total volume of 800 μ l (this is the control template).
29. Store control template up to 3 years at -20°C .

Control templates stored for 3 years at this temperature have retained original quality.

GENERATION OF 3C TEMPLATE FROM MAMMALIAN CELLS

3C analyses in mammalian cells are performed using a modified version of the protocol developed for yeast cells (see Basic Protocol 1). In this modified version, intact mammalian cells are cross-linked and lysed. The lysed cells are then digested, ligated under dilute conditions, and the cross-links are reversed.

Materials

Mammalian cells growing in appropriate culture medium (*APPENDIX 3F*)
 37% (v/v) formaldehyde
 2.5 M glycine
 Lysing buffer II (see recipe), ice cold

**BASIC
 PROTOCOL 3**

**Chromatin
 Assembly and
 Analysis**

21.11.7

Protease inhibitor cocktail for use with mammalian cells
 Restriction enzyme and corresponding 10× restriction enzyme buffer
 1% and 10% (w/v) SDS
 10% (v/v) Triton X-100
 10× ligation buffer (see recipe)
 10 mg/ml BSA
 100 mM ATP
 T4 DNA ligase
 10 mg/ml proteinase K in TE buffer, pH 8.0
 Phenol (*UNIT 2.1A*)
 1:1 (v/v) phenol/chloroform (*UNIT 2.1A*)
 3 M sodium acetate, pH 5.2 (*APPENDIX 2*)
 70% and 100% (v/v) ethanol
 TE buffer, pH 8.0 (*APPENDIX 2*)
 Chloroform
 10 mg/ml DNase-free RNase A
 Dounce homogenizer with pestle B
 1.7-ml microcentrifuge tubes
 16°, 37°, 42°, and 65°C water baths
 15- and 50-ml disposable conical tubes
 250-ml screw-cap centrifuge bottles

Prepare and cross-link cells

1. Grow 1×10^8 cells in appropriate culture medium (*APPENDIX 3F*). Centrifuge cells 10 min at $450 \times g$ (1500 rpm), room temperature, remove supernatant, and dissolve pellet in 45 ml fresh culture medium.

The number of cells can be reduced or increased depending on need. In the procedure described here, the obtained template is sufficient to measure ~1000 interactions.

2. Add 1.35 ml of 37% formaldehyde (1% final), mix thoroughly by pipetting up and down, and incubate 10 min at room temperature.
3. Add 2.5 ml of 2.5 M glycine. Mix by pipetting up and down and incubate 5 min at room temperature.
4. Store on ice for at least 15 min.
5. Centrifuge cells 10 min at $800 \times g$ (2000 rpm), room temperature, resuspend pellet in 1 ml ice-cold lysing buffer II supplemented with 0.1 ml protease inhibitor cocktail, and incubate 15 min on ice.
6. Dounce homogenize cells on ice with pestle B by gently stroking 15 times, incubating 1 min on ice, and finally stroking 15 additional times.

Digest cross-linked cells

7. Centrifuge cells 5 min at $2500 \times g$ (5000 rpm), room temperature, in a 1.7-ml microcentrifuge tube, and wash with 0.5 ml of 1× appropriate restriction enzyme buffer. Centrifuge again and resuspend in 0.5 ml of 1× restriction enzyme buffer.
8. Distribute 25 μ l of cells to 20 individual 1.7-ml microcentrifuge tubes (each tube containing 5×10^6 cells).
9. Centrifuge 5 min at $2500 \times g$ (5000 rpm), room temperature, and resuspend each pellet in 362 μ l of 1× restriction enzyme buffer.
10. Add 38 μ l of 1% SDS to each tube and incubate 10 min at 65°C.
11. Add 44 μ l of 10% Triton X-100 to each tube. Mix by pipetting up and down.

12. Add 400 U restriction enzyme per tube. Mix well and incubate overnight at an appropriate temperature (per manufacturer's instructions).
13. Add 86 μl of 10% SDS and incubate 30 min at 65°C.

Ligate cross-linked cells

14. Transfer each reaction to a 15-ml disposable conical tube. Add the following to each tube and incubate 2 hr at 16°C:

745 μl of 10% Triton X-100
745 μl of 10 \times ligation buffer
80 μl of 10 mg/ml BSA
80 μl of 100 mM ATP
5960 μl distilled water
4000 Weiss units T4 DNA ligase.

Due to the large number of tubes, a master mix of all components is typically made for all ligations and distributed amongst the tubes.

Reverse cross-links and purify 3C template

15. Add 50 μl of 10 mg/ml proteinase K in TE buffer, pH 8.0, and incubate overnight at 65°C.
16. Add an additional 50 μl of 10 mg/ml proteinase K in TE buffer, pH 8.0, and incubate 2 hr at 42°C.
17. Transfer solutions to 50-ml disposable conical tubes, add an equal amount of phenol to each of the tubes, vortex for 30 sec, and centrifuge 5 min at 2460 $\times g$ (3500 rpm), room temperature.
18. Remove the aqueous (upper) phase, transfer to fresh 50-ml disposable conical tubes, and repeat the phenol extraction.
19. Add an equal volume of 1:1 (v/v) phenol/chloroform, vortex for 30 sec, and centrifuge 5 min at 2460 $\times g$ (3500 rpm), room temperature.
20. Pool the aqueous (upper) phases from the 20 samples into three 250-ml screw-cap centrifuge bottles. Add 1/10 vol of 3 M sodium acetate, pH 5.2, vortex briefly, and add 2.5 vol of ice-cold 100% ethanol.
21. Incubate 30 min (or overnight) at -80°C and centrifuge 20 min at 12,000 $\times g$ (10,000 rpm), 4°C.
22. Aspirate supernatant and resuspend the pellets in a total volume of 1 ml TE buffer, pH 8.0. Transfer DNA to two fresh 1.7-ml microcentrifuge tubes each containing 500 μl of DNA solution.
23. Perform phenol extraction (step 17), followed by two 1:1 phenol/chloroform extractions (step 19) for each tube.
24. Add an equal volume of chloroform to each tube, vortex for 30 sec, and centrifuge 5 min at 1100 $\times g$ (3500 rpm), room temperature. Remove the aqueous phase and transfer to fresh 1.7-ml microcentrifuge tubes.
25. Add 1/10 vol of 3 M sodium acetate, pH 5.2, and vortex briefly. Precipitate DNA by adding 2.5 vol of ice-cold 100% ethanol and mixing gently. Incubate 15 min at -20°C and centrifuge 20 min at 18,000 $\times g$ (14,000 rpm), 4°C.
26. Wash the pellets five times with 70% ethanol.

These five wash steps are necessary to remove extra salt in the template.

27. Dry the pellets and dissolve in a total volume of 1 ml TE buffer, pH 8.0 (this is the 3C template).
28. Add 1 μ l of 10 mg/ml DNase-free RNase A and incubate 15 min at 37°C.
29. Store 3C template up to 2 years at -20°C.

3C templates stored for 2 years at this temperature have retained original quality.

GENERATION OF CONTROL TEMPLATE FROM MAMMALIAN DNA

This protocol outlines the method to obtain a control template from mammalian genomic DNA. Due to the complexity of the mammalian genome, a control template generated from whole genomic DNA does not contain detectable levels of any specific ligation product. Thus, a control template is generated from the genomic region of interest only. In this protocol, a bacterial artificial chromosome (BAC) clone (*UNIT 5.9*) or set of clones containing the genomic region of interest is purified, digested, and randomly ligated. This procedure generates a control template, i.e., a collection of all possible ligation products that are present in equimolar amounts. This control template is used in combination with the 3C template to determine interaction frequencies (see Basic Protocol 5).

Materials

BAC clones (e.g., Invitrogen and CHORI; <http://bacpac.chori.org>)
TE buffer, pH 8.0 (*APPENDIX 2*)
0.8% agarose/0.5 \times TBE gel (*UNIT 2.5A*)
Molecular weight standard of known concentration
Restriction enzyme and corresponding 10 \times restriction enzyme buffer
10 mg/ml BSA
1:1 (v/v) phenol/chloroform (*UNIT 2.1A*)
3 M sodium acetate, pH 5.2 (*APPENDIX 2*)
70% and 100% (v/v) ethanol, ice cold
10 \times ligation buffer (see recipe)
100 mM ATP
T4 DNA ligase
Chloroform
1.7- and 2-ml microcentrifuge tubes
16°, 37°, and 65°C water baths
Additional reagents and equipment for recovery of DNA from PAC/BAC clones (*UNIT 5.9*), for real-time PCR if applicable (*UNIT 21.3*), and for agarose gel electrophoresis (*UNIT 2.5A*)

Select, amplify, and purify BAC clone(s)

1. Select one or more BAC clones covering the genomic region of interest. If more than one clone is needed, there should be minimal overlap and possibly no gaps between the clones.
2. Purify the BAC DNA from 500-ml overnight cultures by the modified alkaline lysis method as described in *UNIT 5.9*. Resuspend BAC DNA in 1 ml of TE buffer, pH 8.0.
3. Estimate DNA concentration by running 1 μ l of BAC DNA on a 0.8% agarose/0.5 \times TBE gel (*UNIT 2.5A*) beside a molecular weight standard of known concentration.

The concentration of the BAC DNA should be between 50 and 100 ng/ml.

*If more than one BAC clone is required to cover the genomic region of interest, the clones should be mixed in equimolar ratios before digestion. For this, the BAC preparations should be quantified by real-time quantitative PCR with SYBR green (*UNIT 21.3*) using universal primers that amplify part of the BAC vector backbone.*

Digest BAC genomic DNA clones

4. Transfer ~20 μg of BAC DNA to a 2-ml microcentrifuge tube.

If more than one BAC is required, the total amount of BAC DNA should equal 20 μg .

The total DNA volume should represent 10% of the final digestion volume. Do not exceed 200 μl of DNA in the final 2-ml digestion. If the volume is >200 μl , precipitate DNA and resuspend in a more suitable volume.

Reserve 1 μl of undigested material to run on an agarose gel alongside the digested BAC DNA and ligated BAC DNA as a control for digestion efficiency.

5. Add an appropriate amount of water, 10 \times restriction enzyme buffer, and 10 mg/ml BSA (if recommended by the manufacturer) to obtain, after addition of enzyme, a solution containing 1 \times restriction enzyme buffer and 100 ng/ μl BSA.
6. Add 40 U/ μl of restriction enzyme stock corresponding to 8.75% of the final digestion volume.
7. Incubate overnight at an appropriate temperature (per manufacturer's instructions).

BAC DNA is often difficult to digest to completion. This protocol has been optimized and standardized for digestion efficiency.
8. Add an equal volume of 1:1 (v/v) phenol/chloroform, vortex for 30 sec, and centrifuge 5 min at 18,000 $\times g$ (14,000 rpm), room temperature, separating the reaction volume in multiple tubes if necessary.
9. Transfer aqueous (upper) phase to a fresh 1.7-ml microcentrifuge tube.
10. Add 1/10 vol of 3 M sodium acetate, pH 5.2, vortex briefly.
11. Add 2.5 vol of ice-cold 100% ethanol and mix gently. Incubate at least 15 min at -20°C and centrifuge 20 min at 18,000 $\times g$ (14,000 rpm), 4°C .
12. Discard supernatant and wash pellet by resuspending in 1 ml of 70% ethanol. Centrifuge 15 min at 18,000 $\times g$ (14,000 rpm), 4°C .
13. Remove 70% ethanol wash, and briefly air dry.

Do not let the DNA dry completely as it becomes very difficult to resuspend.
14. Resuspend DNA in 161 μl water. Incubate 15 min at 37°C to dissolve the DNA completely.

Reserve a 4- μl aliquot of digested material to run on an agarose gel alongside the undigested BAC DNA and ligated BAC DNA as a control for digestion efficiency.

Ligate BAC DNA

15. Combine the following and incubate overnight at 16°C (adjust to 200 μl final volume if necessary):

157 μl of digested BAC DNA
20 μl of 10 \times ligation buffer
2 μl of 10 mg/ml BSA
2 μl of 100 mM ATP
7600 Weiss units T4 DNA ligase.

Due to the large number of tubes, a master mix of all components is typically made for all ligations and distributed amongst the tubes.

16. Incubate DNA 15 min at 65°C to inactivate the ligase.

Purify BAC genomic DNA control template

17. Add 200 μl of 1:1 phenol/chloroform, vortex for 30 sec, and centrifuge 5 min at $18,000 \times g$ (14,000 rpm), room temperature.
18. Transfer aqueous (upper) phase to a fresh 1.7-ml microcentrifuge tube without touching the interface.
19. Repeat 1:1 phenol/chloroform extraction. Transfer aqueous (upper) phase to a fresh tube.
20. Add 200 μl chloroform, vortex for 30 sec, and centrifuge 5 min at $18,000 \times g$ (14,000 rpm), room temperature.
21. Transfer aqueous (upper) phase to a fresh 1.7-ml microcentrifuge.
22. Add 1/10 vol of 3 M sodium acetate, pH 5.2, and vortex briefly. Add 2.5 vol of ice-cold 100% ethanol and mix gently. Incubate 15 min at -20°C and centrifuge 20 min at $18,000 \times g$ (14,000 rpm), 4°C .
23. Discard supernatant and resuspend the pellet in 1 ml of 70% ethanol. Centrifuge 15 min at $18,000 \times g$ (14,000 rpm), 4°C .
24. Discard supernatant and air dry slightly.
25. Resuspend pellet in TE buffer, pH 8.0, to give a final concentration of $\sim 100 \text{ ng}/\mu\text{l}$.
26. Incubate 15 min at 37°C to completely dissolve DNA (this is the control template).

Reserve 1 μl of ligated material to run on an agarose gel alongside the undigested BAC DNA and digested BAC DNA as a control for ligation efficiency.

27. Run 1 μl of uncut BAC DNA, 4 μl of cut BAC DNA, and 1 μl of ligated BAC DNA side by side on a 0.8% agarose/0.5 \times TBE gel (*UNIT 2.5A*).

There should be a smear after digestion, and part of the smear should disappear after ligation. An example is shown in Figure 21.11.2.

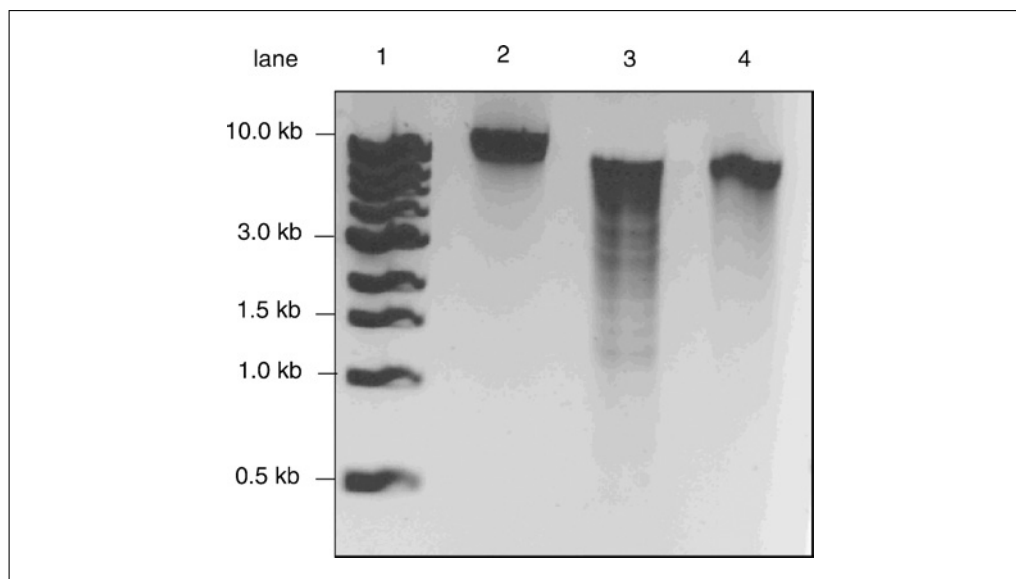


Figure 21.11.2 Gel analysis of BAC DNA during generation of mammalian control template. A 1-kb ladder is shown in lane 1 for size reference. Lane 2 is undigested BAC DNA, which runs on an agarose gel at slightly more than 10 kb. Lane 3 is digested BAC DNA, which is smear-like in appearance. Lane 4 is ligated BAC DNA, which shows the disappearance of most of the smear-like trail. Similar results are obtained for a *Saccharomyces cerevisiae* control template at the corresponding stages.

28. Store control template up to 3 years at -20°C .

Control templates stored up to 3 years at this temperature have retained original quality.

ANALYSIS OF INTERACTION FREQUENCIES USING 3C AND CONTROL TEMPLATES BY QUANTITATIVE PCR

**BASIC
PROTOCOL 5**

After generation of 3C and control templates (see Basic Protocols 1 through 4), it is necessary to roughly determine the amount of template generated. After a rough estimate of DNA concentration is determined, a titration must be performed to determine how much of the templates should be used in a PCR reaction to accurately quantify ligation products.

Materials

DNA templates (see Basic Protocols 1 through 4)
Molecular weight standard of known concentration
10 \times PCR buffer for yeast *or* mammalian templates (see recipes)
100 mM dNTPs
Primers (see Strategic Planning)
Taq DNA polymerase
50 mM MgSO_4
Automated thermal cycler
Gel documentation setup with appropriate software for quantifying PCR products
Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5A*) and PCR (*UNITS 15.1 & 15.7*)

1. Determine DNA concentration of templates by agarose gel electrophoresis by running template DNA in parallel with a molecular weight standard (*UNIT 2.5A*).

The template will run as a high-molecular-weight DNA sample. The template appearance may vary between samples and may contain smear-like trails. The appearance of the template may not be an indication of template quality, and only after titration can the quality of the template be accurately determined. Typically, the DNA concentration is ~ 0.8 to $1.4 \mu\text{g}/\mu\text{l}$ for yeast templates and ~ 200 to $250 \text{ ng}/\mu\text{l}$ for mammalian templates. Absorption spectroscopy should not be used to measure DNA concentration due to a high concentration of salt remaining in the templates after purification.

2. Make a two-fold dilution series of template DNA and include a water control.

Typically, eight to ten dilutions are made, with the highest concentrations at $\sim 250 \text{ ng}/\mu\text{l}$ for mammalian templates and $\sim 1 \mu\text{g}/\mu\text{l}$ for yeast templates.

3. Set up PCR reactions for each dilution (*UNITS 15.1 & 15.7*).

- a. For yeast, perform PCR in a 50- μl reaction containing:
 - 2 μl each template dilution (i.e., starting with 2 μg DNA)
 - 1 \times PCR buffer for yeast templates
 - 0.5 mM dNTPs
 - 0.4 μM of each primer
 - 2.5 U *Taq* DNA polymerase.
- b. For mammalian cells, perform PCR in a 25- μl reaction containing:
 - 2 μl each template dilution (i.e., starting with 500 ng DNA)
 - 1 \times PCR buffer for mammalian templates
 - 4 mM MgSO_4
 - 0.2 mM dNTPs
 - 0.4 μM of each primer
 - 1 U *Taq* DNA polymerase.

**Chromatin
Assembly and
Analysis**

21.11.13

Titration experiments are performed with two different primer combinations. Primers to use in titrations should be chosen based on efficiency (i.e., lack of primer dimers). The two primer pairs chosen should also differ in the genomic distance that separates the restriction fragments they recognize. Preferably, one pair of primers should amplify a ligation product that represents an interaction between restriction fragments that are relatively far apart (e.g., 50 to 80 kb) on the genomic map, whereas the other pair should amplify a ligation product that represents an interaction between two fragments that are located close together (i.e., 10 kb).

4. Use the following hot-start PCR parameters to program the thermal cycler:

a. For quantitative results for yeast templates:

1 cycle:	1 min	95°C
32 cycles:	1 min	95°C
	45 sec	60°C
	2 min	72°C
1 cycle:	1 min	95°C
	45 sec	60°C
	8 min	72°C

b. For quantitative results for mammalian templates:

1 cycle:	1 min	95°C
34 cycles:	1 min	95°C
	45 sec	65°C
	2 min	72°C
1 cycle:	1 min	95°C
	45 sec	65°C
	8 min	72°C

5. Analyze PCR products by agarose gel electrophoresis (UNIT 2.5A) on a 1.5% agarose gel containing 0.5× TBE and 0.5 µg/ml ethidium bromide.

An example of a titration experiment is shown in Figure 21.11.3.

A loading buffer without bromophenol blue is recommended, since this dye will run close to the PCR products and interfere with quantification. The use of Ficoll is encouraged.

6. Quantify PCR products using a gel documentation setup equipped with appropriate software and determine amount of PCR product per reaction.

An example of a quantification is shown in Figure 21.11.3.

The recommended amount of template that should be used for quantitative PCR detection of ligation products (and thus chromatin interactions) should be in the linear range of PCR amplification (e.g., as indicated in Fig. 21.11.3). This concentration should be used for all subsequent PCR reactions with template.

7. Use the proper DNA concentration determined by the PCR titration to analyze the genomic region of interest. For each primer pair, set up three reactions using the control template and three reactions using the 3C template. Follow the same PCR conditions and program as described in steps 3 and 4.

The control template and 3C template PCRs for each primer pair should always be performed during the same PCR run.

8. Analyze and quantify the PCR products using the same conditions as described in steps 5 and 6.

PCR products obtained with the 3C and control templates should be run side by side on the same gel, as there is also variation in DNA staining between gels.

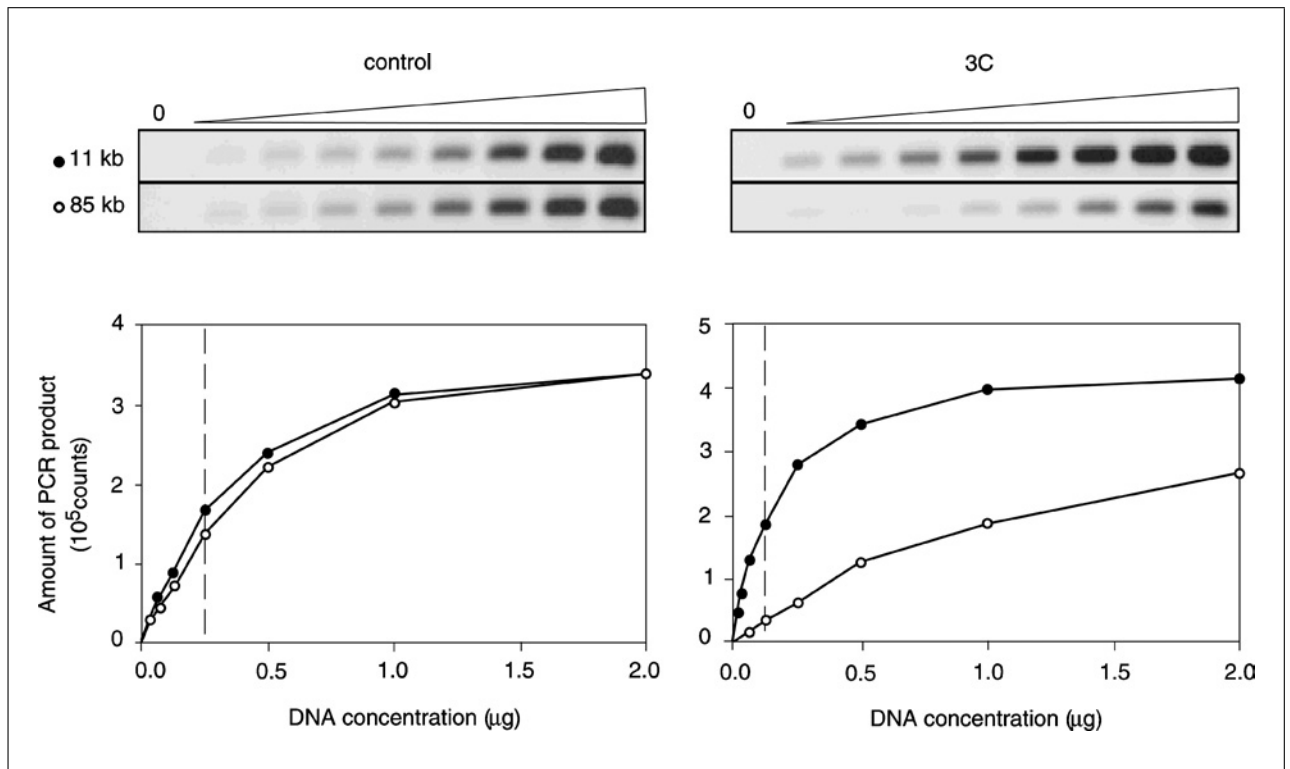


Figure 21.11.3 Titration and quantification of *Saccharomyces cerevisiae* 3C and control templates. The panels depict titration and analysis of a control template (left) and 3C template (right) using the same two primer pairs. The PCR products were run on agarose gels, PCR products were quantified, and the amount of PCR product was plotted versus DNA concentration in micrograms for the two primer pairs, one 11 kb apart in genomic distance (closed circles) and one 85 kb apart (open circles). For both primer pairs, the control template yields similar amounts of PCR product. The 3C template yields lower PCR products for the 85-kb primer pair than for the 11-kb primer pair. The linear range for PCR amplification for 3C and control templates is to the left of the dashed line. The DNA concentration for use in subsequent reactions should be in this range, but should also be high enough so that PCR products can be accurately quantified on an agarose gel. Similar results are obtained for a mammalian control template and 3C template.

- Determine the interaction frequency for each primer pair by dividing the amount of PCR product obtained using the 3C template by the amount of PCR product obtained using the control template.

As each primer pair using each template was set up in triplicate, three different values will be obtained. These values are averaged to determine the final interaction frequency for a particular primer pair.

The standard error of the mean of the three values for the interaction frequencies should not be more than 15%.

REAGENTS AND SOLUTIONS

Use deionized, autoclaved water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Ligation buffer, 10×

500 mM Tris·Cl, pH 7.5 (APPENDIX 2)
 100 mM MgCl₂
 100 mM dithiothreitol
 Store up to 1 year at −80°C

Lysing buffer I

0.25 M EDTA, pH 8.0 (*APPENDIX 2*)
0.5 M Tris base
2.5% (v/v) SDS
Make fresh on day of experiment

Lysing buffer II

10 mM Tris·Cl, pH 8.0 (*APPENDIX 2*)
10 mM NaCl
0.2% (v/v) Igepal (NP-40)
Make fresh on day of experiment

MES wash buffer

0.1 M MES
1.2 M sorbitol
1 mM EDTA, pH 8.0 (*APPENDIX 2*)
0.5 mM MgCl₂
Adjust pH to 6.4 with NaOH
Store up to 6 months at 4°C

PCR buffer for mammalian templates, 10×

600 mM Tris, adjust to pH 8.9 with H₂SO₄
180 mM (NH₄)₂SO₄
Store up to 1 year at -80°C

PCR buffer for yeast templates, 10×

100 mM Tris·Cl, pH 8.4 (*APPENDIX 2*)
500 mM KCl
22.5 mM MgCl₂
Store up to 1 year at -80°C

Spheroplasting buffer I

0.4 M sorbitol
0.4 M KCl
40 mM sodium phosphate buffer, pH 7.2 (*APPENDIX 2*)
0.5 mM MgCl₂
Store up to 6 months at 4°C

Spheroplasting buffer II

10 mM sodium phosphate buffer, pH 7.2 (*APPENDIX 2*)
10 mM EDTA, pH 8.0 (*APPENDIX 2*)
1% (v/v) 2-mercaptoethanol
100 µg/ml zymolyase 100-T (US Biological)
Store up to 6 months at 4°C

Zymolyase 100-T solution, 20 mg/ml

20 mg/ml zymolyase 100-T (US Biological)
2% (w/v) glucose
50 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
Make solution at least 1 day prior to experiment to increase the enzyme efficiency
Store up to 1 month at 4°C

COMMENTARY

Background Information

Chromatin organization has been studied at a very local level, i.e., at the level of one or a few nucleosomes. For instance, crystal structures of single nucleosomes and their constituents have been solved at high resolution (Luger et al., 1997). On the other hand, microscopic techniques have allowed analysis of whole chromosome organization. However, very little is known about chromatin conformation at a scale ranging from a few kilobases to hundreds of kilobases. This intermediate level of chromosome structure is extremely relevant for genome regulation, as it is the scale at which chromatin affects many aspects of gene regulation. The relatively new 3C technology allows for such structural analysis of chromatin, which was not possible with existing techniques (Dekker et al., 2002). 3C makes use of formaldehyde cross-linking to capture chromatin in a particular conformation. As a result, chromatin segments that interact inside of the cell will be covalently linked.

The 3C technique has been used to detect and quantify physical interactions between sequence elements located on the same chromosome or on different chromosomes in *Saccharomyces cerevisiae* (Dekker et al., 2002) as well as in higher eukaryotes (Tolhuis et al., 2002; Dekker, 2003; Spilianakis and Flavell, 2004; Spilianakis et al., 2005; Vakoc et al., 2005). Such chromatin interactions can result in the formation of chromatin loops, as in the case of enhancers and promoters, but they will also reflect more general properties of the chromatin fiber such as chromatin compaction. Together, these types of interactions can directly provide information about the spatial organization of a genomic region, but careful analysis of 3C data is required to arrive at a correct biological interpretation.

Critical Parameters and Troubleshooting

Experimental preparation

Many factors play a role in a successful 3C experiment. A 3C experiment should always involve careful thought and planning. One of the most important factors is the quality of the starting material. *Saccharomyces cerevisiae* cells should be freshly grown on the appropriate medium and a liquid culture should be inoculated from one isolated colony. Mammalian cells should be exponentially growing in the appropriate medium and exhibit healthy growth curves appropriate to the cell type.

It is also critical to start with the appropriate amount of cells indicated in Basic Protocols 1 through 4. Deviations from these conditions may result in poor or very low PCR signals.

As described in Strategic Planning, restriction enzyme selection and primer design are critical. An enzyme should be chosen based on multiple criteria. The enzyme should cut evenly and frequently throughout the region of interest. Once an enzyme is chosen, primers must be designed carefully. Primers should be designed unidirectionally, and should not be designed for large (>10 kb) or small (<1 kb) fragments, as these may result in slightly higher and lower interaction frequencies, respectively. Primers that do not amplify the correct product, amplify multiple products, do not yield equal amounts of PCR products when using control template, or produce high levels of primer dimers should not be included in the analysis. If crucial to the experiment, the primer should be redesigned.

The reagents prepared for use in a 3C experiment should be stored at the appropriate conditions, according to the recipes listed in the Reagents and Solutions section. Among the most important reagents is formaldehyde. The efficiency of cross-linking will directly affect the frequency with which ligation products are formed and will thus impact the reliability of PCR detection. Therefore, the formaldehyde used should be no more than 1 year old.

Template quality

The most commonly encountered problem during 3C is difficulty in identifying the concentration of 3C template that should be used for quantitative PCR detection of ligation products. This problem is apparent when one cannot find a range of template concentrations for which the amount of PCR product obtained is linearly related to the template concentration (see Fig. 21.11.3). If this situation occurs, it is likely due to a high concentration of salt remaining in the template after purification. This remaining salt will severely reduce the quality and intensity of the PCR signal. It is therefore necessary to remove salt from the template preparation. This can be achieved by repeating the ethanol precipitation and ethanol wash steps (as described in Basic Protocol 3), and resuspending in the appropriate volume. This extra wash step should improve template titrations.

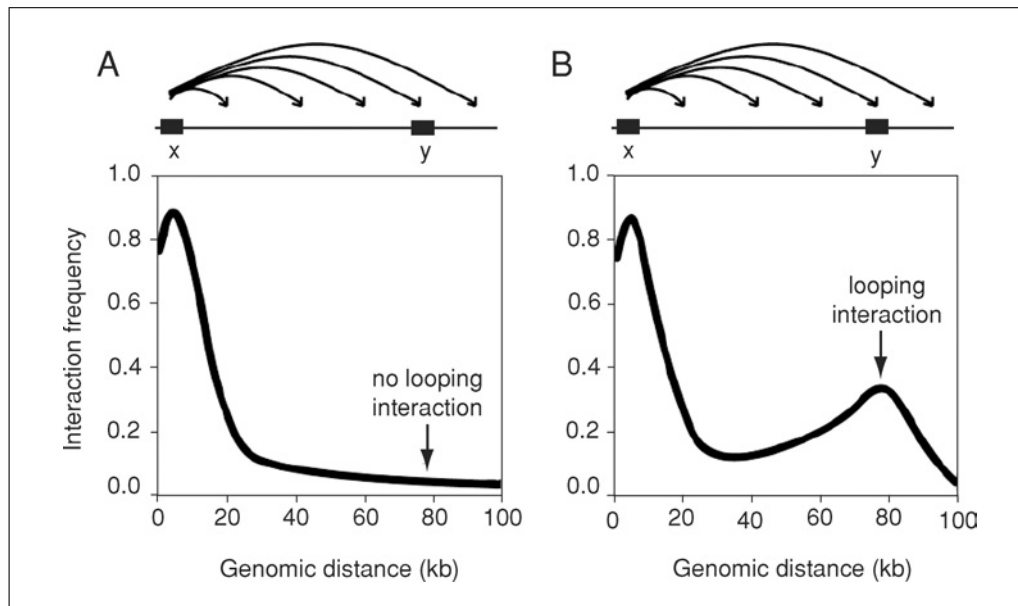


Figure 21.11.4 Hypothetical results for a 3C experiment. **(A)** A 3C dataset with no looping interaction between sequence elements *x* and *y* (black boxes). A number of interactions have been tested between sequence element *x* and sites located further away from it including sequence element *y*. Neighboring sites of *x*, those separated by a small genomic distance, exhibit high interaction frequencies. As genomic distance between sites increases, the interaction frequencies decrease progressively. A low interaction frequency is detected between sequence elements *x* and *y*. The exact shape of the curve is dependent on flexibility and the level of compaction of the chromatin fiber. **(B)** In a 3C dataset with a looping interaction between sequence elements *x* and *y*, a local peak of interaction frequencies is observed. This peak (at ~80 kb) shows that sequence elements *x* and *y* interact more frequently than expected, which is indicative of a looping interaction.

The quality of the control template is also crucial for a successful 3C experiment. As mentioned, a control template should contain equal amounts of all possible ligation products (see Fig. 21.11.3 for an example). If this is not the case, a new control template should be generated, or the reliability of the data obtained will be compromised. When generating a control template for mammalian cells, if the genomic area of interest spans multiple BAC clones, it is important that the BAC clones chosen have minimal overlaps and gaps, as these particular regions will be overrepresented or absent, respectively. If this cannot be avoided, the gap or overlap region should never be directly at a putative looping element as this may skew results.

Finally, as a general rule, sites separated by a small genomic distance (i.e., separated by <10 kb) should exhibit high interaction frequencies. The interaction frequencies should decrease as sites are tested that are further apart in genomic distance (i.e., 50 to 80 kb). If sites separated by small genomic distances do not exhibit high interaction frequencies, a new template should be generated, as something has gone awry during the course of

the experiment. Importantly, the presence of a looping interaction between two distant sequences will be apparent when they interact more frequently than one would expect for two sequences separated by that particular genomic distance (Fig. 21.11.4).

Experimental design and data analysis

The 3C technique has successfully been used to detect looping interactions between sequence elements (Dekker et al., 2002; Tolhuis et al., 2002; Spilianakis and Flavell, 2004; Vakoc et al., 2005); however, experimental design is crucial for correct interpretation of the data. When designing an experiment, it is important that multiple primer pairs are chosen to detect a loop. For example, if two putative looping elements are located 100 kb apart, it will be important to also test interactions of each of the two elements with several sites located in between them. These interactions should be less frequent than the interaction between the two looping elements (see Fig. 21.11.4). In other words, a looping interaction is detected as a local peak in interaction frequencies. Measuring only the interaction frequency between the two putative looping

elements does not allow the detection of such a local peak in interaction frequency and thus is never sufficient to confirm that a chromatin loop is formed.

As mentioned, 3C is a PCR-based method. PCR reactions themselves are a source of noise. Therefore, PCR reactions using the control template and the 3C template should be performed in triplicate as described in Basic Protocol 5. From the abundance of PCR product, three individual interaction frequencies should be determined and the values averaged to obtain one data point. Typically, the standard error of the mean of the three interaction frequencies is 15%. The average interaction frequency is used for compiling datasets. For the control template to be a true control of primer and PCR efficiency, it is imperative that the PCR reactions used to detect ligation products for both the 3C and control templates are performed during the same PCR run. PCR products obtained with the 3C and control templates should be run side by side on the same gel, as there is also variation in DNA staining between gels.

Lastly, if comparing two 3C datasets, e.g., derived from two cell strains/lines or two experimental conditions, it is important that the two datasets are normalized with respect to each other. Normalization will control for non-relevant differences between the two 3C templates, such as differences in DNA concentration. Normalization is achieved by measuring a number of interaction frequencies that are likely to be the same in both cell lines or experimental conditions. Ideally, a set of at least ten to twenty interaction frequencies is measured (relative to the control template) within a genomic region other than the region of interest to serve as an internal control. Once these interaction frequencies are determined, the log ratio between the two cell lines for each interaction is calculated. The average of all log ratios will then serve as a factor to normalize the two 3C datasets.

Anticipated Results

Figure 21.11.4 shows hypothetical results of a typical 3C experiment where interaction frequencies are plotted versus the genomic distance (in kilobases) that separates the interacting restriction fragments. Figure 21.11.4A shows a dataset exhibiting no looping interaction. Sites separated by a relatively small genomic distance exhibit high interaction frequencies. As genomic distance between sites increases, the interaction frequency decreases

progressively. The exact shape of this curve is dependent on the flexibility and level of compaction of the chromatin fiber. However, if a loop is present at a particular genomic region, a local peak of interaction frequencies will be observed, as shown in Figure 21.11.4B. This peak could be several fold higher than predicted for that particular site separation. The height of the peak can indicate the fraction of cells engaged in that particular interaction. For example, a study from Vakoc et al. (2005) observed a loop in the globin locus between the LCR and β -major gene upon GATA-1 activation that is reflected in an interaction frequency that is 3.3-fold higher than expected.

Time Considerations

Preparation of 3C templates from yeast and mammalian cells (Basic Protocols 1 and 3) may be completed in a 3- or 4-day period. Day 1 consists of preparation of cells and cross-linking (~3 hr). At this stage, cross-linked cells can be stored at -80°C or used directly for cell lysis and digestion of chromatin. Day 2 (or end of day 1) consists of a series of wash steps followed by solubilization of chromatin and addition of restriction enzyme (<4 hr). The digestion is allowed to proceed overnight. Inactivation of the restriction enzyme and the ligation reactions (~3 hr) are performed on day 3 (or day 2), followed by reversal of the cross-links, which is performed overnight. Day 4 (or day 3) consists of purification of DNA through a series of phenol/chloroform extractions and ethanol precipitations. Purification of DNA for yeast 3C templates will take up to 5 hr. A full day should be reserved for purification of DNA for mammalian 3C templates.

Preparation of the control yeast template (see Basic Protocol 2) may be completed in 2 days. Isolation of yeast genomic DNA will take 1 full day. Day 2 consists of digestion (~4 hr) followed by ligation (~1 hr) and purification of DNA (~1.5 hr). Preparation of mammalian control templates (see Basic Protocol 4) may be completed in 4 days. Day 1 consists of purification of BAC DNA. Day 2 consists of assembling digestion reactions (<1 hr). The digestion is allowed to proceed overnight. Day 3 consists of purification of digested BAC genomic DNA clones followed by assembling ligation reactions (~2 hr). The ligation is allowed to proceed overnight. Day 4 consists of purification of mammalian control template DNA (~3 hr).

Once all templates are generated, titrations must be performed as described in Basic

Protocol 5. This protocol will take ~5 hr including time for titration setup, the PCR program, and gel electrophoresis. The time needed to analyze an entire genomic region of interest will depend on the number of interactions that need to be measured.

Literature Cited

- Dekker, J. 2003. A closer look at long-range chromosomal interactions. *Trends Biochem. Sci.* 28:277-280.
- Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. 2002. Capturing chromosome conformation. *Science* 295:1306-1311.
- Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389:251-260.
- Spilianakis, C.G. and Flavell, R.A. 2004. Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat. Immunol.* 5:1017-1027.
- Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R., and Flavell, R.A. 2005. Interchromosomal associations between alternatively expressed loci. *Nature* 435:637-645.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. 2002. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* 10:1453-1465.

Vakoc, C., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J., and Blobel, G.A. 2005. Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol. Cell* 17:453-462.

Internet Resources

<http://bacpac.chori.org>

Website for information and to purchase BAC clones from various sources.

Contributed by Adriana Miele, Nele Gheldof, Tomoko M. Tabuchi, José Dostie, and Job Dekker
University of Massachusetts Medical School
Worcester, Massachusetts

Paired-End diTagging for Transcriptome and Genome Analysis

Paired-End diTagging (PET) is the process of extracting 18-base “signatures” or “tags” from each of the 5′ and 3′ termini of any contiguous DNA sequence, and ligating them into paired-end ditags (PETs) that are concatenated for enhanced sequencing efficiency. Each PET can subsequently be mapped onto the appropriate genome assembly to accurately define the location of the original fragment from which it was derived. Analysis of PETs can be used in transcriptome characterization, genome annotation, and analysis of transcription factor binding sites. There are two main approaches that can be chosen depending on the application.

GIS-PET. The gene identification signature (GIS)-PET analysis method was developed to facilitate high-efficiency transcriptome analysis and genome annotation. It captures the 5′ and 3′ terminal signatures of full-length transcripts in the form of covalently-linked DNA ditags or PETs. Essentially, GIS-PET analysis is a two-part procedure, consisting of the construction of a GIS full-length cDNA (flcDNA) library in a unique cloning vector (pGIS4) using a modified biotinylated CapTrapper approach (see Basic Protocol 1), followed by the construction of a GIS-PET library containing concatenated PETs for DNA sequencing (see Basic Protocol 2). Concatenated PETs in bacterial clones are sequenced, and the PET sequences are extracted and mapped onto assembled genomes for further analysis using a customized software suite (PET-Tool; contact Dr. Yijun Ruan at ruanyj@gis.a-star.edu.sg). For ease of visualization, the PETs are displayed on a genome browser developed in-house called “T2G Browser.” Using GIS-PET, one can verify existing transcripts, validate gene predictions, identify splice variants and novel and/or unusual transcripts, and compare gene expression between different libraries. A major advantage of this procedure is the ready availability of the flcDNA plasmid library, from which transcripts of interest can be readily obtained for downstream functional studies by simple PCR, using primers designed based on the PET sequences. The original published GIS-PET method involved the separate ligation of a 5′ adapter-mix and a 3′ adapter onto full-length cDNA before cloning into vector pGIS1. This current variation replaces pGIS1 with vector pGIS4, which already has the 3′-adapter sequences built into one of its arms, thereby reducing the number of steps required.

ChIP-PET. The same principle of paired-end ditagging can be applied to the efficient mapping of transcription factor binding sites (TFBS) identified by chromatin immunoprecipitation (ChIP). In this method, referred to as ChIP-PET, randomly sheared genomic DNA fragments that are first enriched for the TFBS of interest by ChIP (see UNIT 21.3) are inserted into a specific cloning vector (pGIS3; see Basic Protocol 3) and subjected to the same ditagging approach as in GIS-PET. Genuine TFBS are indicated by the overlapping of multiple distinct PET sequences on defined chromosomal loci.

NOTE: The terms ditag and PET are synonymous in the context of this procedure and are used interchangeably within the text.

CONSTRUCTION OF A GIS FULL-LENGTH cDNA LIBRARY

This procedure is outlined on the left side of Figure 21.12.1. In the first steps, full-length mRNA is converted to first-strand, full-length cDNA (flcDNA). For the generation of flcDNA, we decided to adapt the CapTrapper procedure first developed by Piero

BASIC PROTOCOL 1

Chromatin Assembly and Analysis

21.12.1

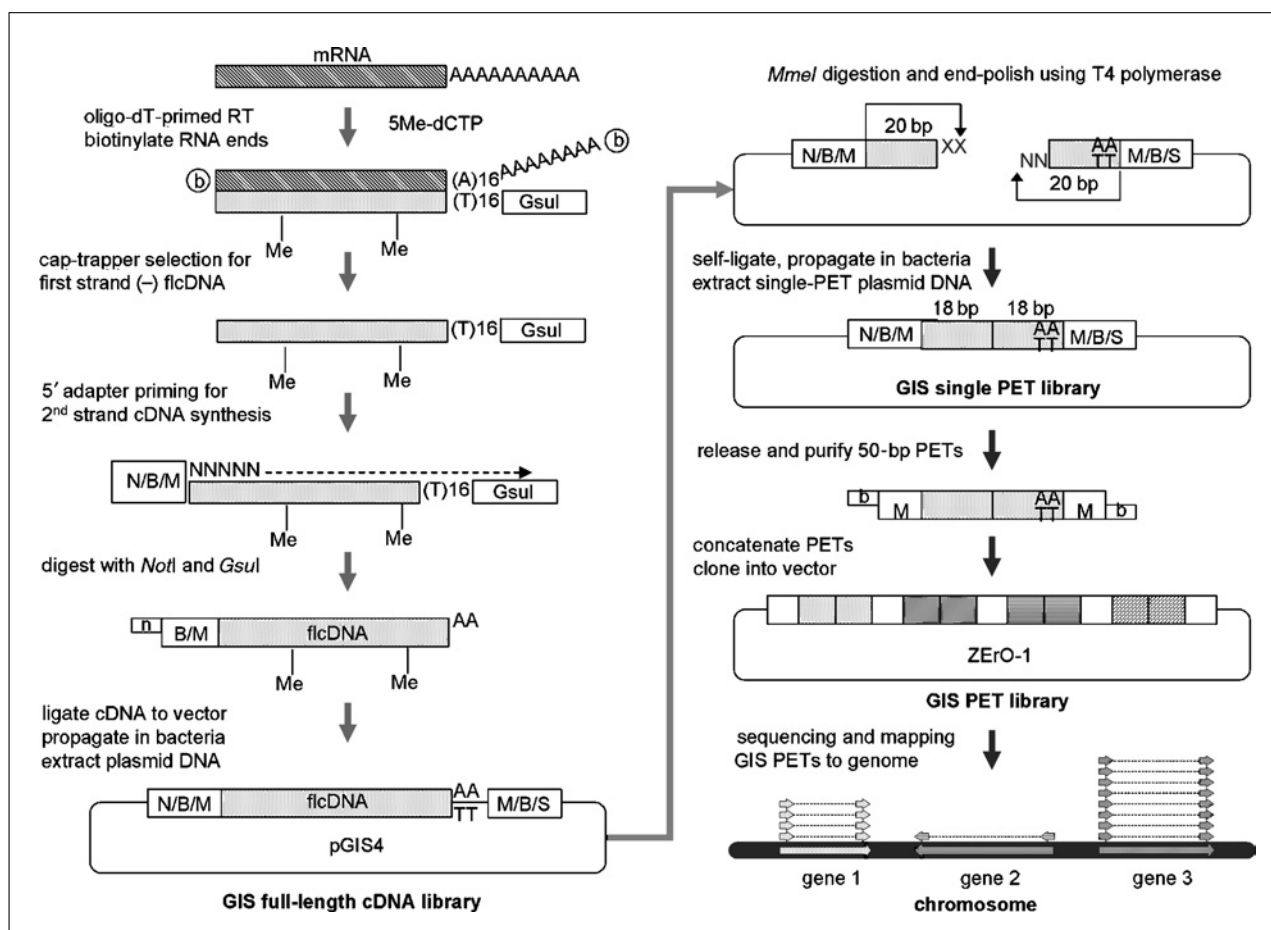


Figure 21.12.1 The current GIS-PET procedure, where the use of vector pGIS4 obviates the 3'-adapter ligation.

Carninci (Carninci and Hayashizaki, 1999). Although this procedure is lengthy and requires a significant quantity of mRNA (10 μ g), trials showed that it produces high-quality flcDNA consistently, which is critical for the success of the subsequent ditagging. In the CapTrapper procedure, the diol structures present in the cap and 3' end of mRNA are oxidized, then chemically biotinylated. RNase I digestion of single-stranded mRNA regions not protected by hybridization to the first-strand flcDNA results in the removal of the biotinylated 3' mRNA ends, as well as the biotinylated caps from prematurely terminated first-strand cDNA/mRNA hybrids. The resulting enriched flcDNA/mRNA hybrids containing biotinylated caps are then captured by binding to streptavidin-coated beads for further manipulation. The isolated (–)flcDNA is converted to double-stranded (ds) cDNA by 5'-adapter ligation using a mix of two adapters, GIS-(N)5 and GIS-(N)6, followed by primer extension. *GsuI* is used to remove the residual polyA tail, leaving an AA dinucleotide residue. This is followed by *NotI* digestion to form the 5' cohesive site, resulting in ds flcDNA that can be directionally inserted into the prepared pGIS4 vector.

The original (published) GIS-PET method involved the separate ligation of a 5' adapter-mix and a 3' adapter onto full-length cDNA, before cloning into vector pGIS1 (Ng et al., 2005). This current variation of the protocol replaces pGIS1 with vector pGIS4, which already has the 3' adapter sequences built into one of its arms, thereby reducing the number of steps required. pGIS4 was derived from pGEM-3z (Promega) using site-directed mutagenesis and adapter ligation to eliminate unwanted *MmeI* sites and to facilitate vector purification and usage. It is available upon request from the authors.

Details for oligonucleotide sequences used in GIS-PET are shown in Figure 21.12.2. Although PAGE-purified oligonucleotides are preferred, for reasons of economy, HPLC-purified oligonucleotides are usually used and have worked well.

**Paired-End
diTagging**

21.12.2

Materials

2 $\mu\text{g}/\mu\text{l}$ *GsuI*-oligo dT primer
40 U/ μl RNasin-PLUS RNase inhibitor (Promega)
PolyA RNA sample of interest (10 to 20 μg)
Nuclease-free water (Ambion)
2 \times GC-I buffer (Takara; available as part of the LA Taq kit)
Modified 10 mM dNTP (with 5-Me-dCTP instead of dCTP; see recipe)
4.9 M D-sorbitol (Sigma)
200 U/ μl Superscript II reverse transcriptase (RNase H⁻; Invitrogen)
Saturated trehalose (see recipe)
20 mg/ml proteinase K (Ambion)
25:24:1 (v/v/v) phenol/chloroform/IAA, pH 6.6 and 7.9 (Ambion)
Sodium acetate solutions: 3 M at pH 5.2 (APPENDIX 2), 1.1 M at pH 4.5, and 1 M at pH 6.1
Absolute and 70% ethanol
Biotin hydrazide (long arm; Vector Laboratories)
Sodium periodate (NaIO₄; Sigma)
10% SDS (APPENDIX 2)
5 M NaCl
Isopropanol
10 U/ μl RNaseONE ribonuclease and 10 \times RNaseONE buffer (Promega)
10 mg/ml yeast tRNA (Ambion)
M-280 streptavidin Dynabeads (Dyna)
1 \times binding buffer (see recipe)
1 \times blocking buffer (see recipe)
1 \times wash buffer (see recipe)
Alkaline hydrolysis buffer (see recipe), prepare fresh
1 M Tris-Cl, pH 7.5 (APPENDIX 2)
15 mg/ml GlycoBlue (Ambion)
LoTE buffer (optional; see recipe)
EB buffer (10 mM Tris-Cl, pH 8.5; Qiagen)
0.4 $\mu\text{g}/\mu\text{l}$ GIS-(N)5 and 0.4 $\mu\text{g}/\mu\text{l}$ GIS-(N)6 adapters (see Fig. 21.12.2; validated as in Support Protocol 2)
Solutions I (enzyme solution) and II (concatenation buffer) from DNA ligation kit (Takara version 2.1)
5 U/ μl ExTaq polymerase and 10 \times ExTaq buffer with Mg²⁺ (Takara)
2.5 mM dNTP mix (Takara)
10 \times Tango buffer (Fermentas)
32 mM S-adenosyl methionine (SAM; NEB; supplied with *MmeI* enzyme)
10 mg/ml BSA
5 U/ μl *GsuI* (Fermentas)
10 U/ μl *NotI* (NEB)
1 \times TEN buffer (see recipe)
SYBR Green I (Molecular Probes) or GelStar nucleic acid gel stain (Cambrex BioWhittaker), optional
pGIS4 cloning vector cut with *NotI/BseRI* (obtain from authors; see Support Protocol 1)
5 U/ μl T4 DNA ligase and 5 \times T4 DNA ligase buffer with PEG (Invitrogen)
Electrocompetent cells (e.g., OneShot electrocompetent TOP10 cells, Invitrogen)
SOC medium (see recipe)
LB-amp agar plates (UNIT 1.1; use Lennox formulation with 5 g/liter NaCl)
20 μM PMR011 and 20 μM PMR012 PCR primers (Fig. 21.12.2)

0.2- and 1.7-ml thin-walled PCR tubes
 Heat/chill block or thermal cycler
 1.7-ml microcentrifuge tubes
 1.5-ml siliconized microcentrifuge tubes
 1.5-ml DNA LoBind tubes, PCR clean, safe lock (Eppendorf or Sigma), optional
 Magnetic rack for washing Dynabeads (Dyna)
 Benchtop hot/cold shaking incubator (e.g., Eppendorf)
 Benchtop rotator (Grant-bio)
 Phase Lock Gel Light tubes (Eppendorf), optional
 cDNA size fractionation columns (Invitrogen)
 Spectrophotometer with 7- μ l submicro-volume cell (Amersham)
 Electroporator and cuvettes (e.g., BioRad)
 15-ml conical tubes (e.g., Falcon)
 96-well microtiter plates

Additional reagents and equipment for phenol extraction and ethanol precipitation (UNIT 2.1A), ligation (UNITS 3.14 & 3.16), agarose gel electrophoresis (UNIT 2.5A), restriction enzyme digestion (UNIT 3.1), electroporation (UNIT 9.3), PCR (UNIT 15.1), and plasmid purification (UNIT 1.6)

Reverse transcribe polyA RNA

1. Mix the following in a 0.2-ml thin-walled PCR tube:

3.5 μ l 2 μ g/ μ l *Gsu*I-oligo dT primer (7 μ g)
 1 μ l 40 U/ μ l RNasin-PLUS inhibitor
 10 to 20 μ g polyA RNA sample of interest
 Nuclease-free water to 20 μ l.

2. Heat the reaction mixture to 65°C for 10 min and cool to 37°C for 1 min. Hold at 42°C while preparing the other components.

PolyA RNA can be purified using any suitable method. The authors routinely purify polyA RNA from total RNA (extracted using Trizol reagent available from Invitrogen) using the μ MACS mRNA isolation kit (Miltenyi Biotech), as this gives the best combination of throughput and purity.

If necessary, the polyA RNA can be ethanol precipitated (without glycogen) to concentrate it prior to use. Care should be taken to not overdry the RNA pellet, as it will be exceedingly difficult to resuspend.

3. Prepare the reverse transcriptase (RT) mix in a 0.2-ml thin-walled PCR tube on ice (total 121 μ l) as follows:

75 μ l 2 \times GC-I buffer
 1 μ l RNasin-PLUS inhibitor
 4 μ l modified 10 mM dNTP (with 5-Me-dCTP instead of dCTP)
 26 μ l 4.9 M sorbitol
 15 μ l 200 U/ μ l Superscript II reverse transcriptase.

The mix should be kept on ice to ensure no RT activity. RNasin-PLUS inhibitor is preferred because of its wide-spectrum properties and its thermal stability.

4. Place 10 μ l saturated trehalose into another 0.2-ml thin-walled PCR tube and leave warming at 42°C in a thermal cycler.

5. When the oligo-dT/mRNA annealing step is complete, place the RT mix into the 42°C thermal cycler for 2 min to preheat. Add the warm trehalose to the warm RT mix (now 131 μ l), and quickly transfer the entire reaction mix into the tube containing the annealed primer/mRNA (final 151 μ l). Immediately start the incubation as follows:

40 min at 42°C
20 min at 50°C
20 min at 55°C.

6. Add 2 μ l of 20 mg/ml proteinase K to degrade all enzymes and incubate 15 min at 45°C.
7. Perform an extraction with 25:24:1 phenol/chloroform/IAA (using pH 6.6 phenol, not pH 7.9), and then re-extract with 150 μ l nuclease-free water.

Refer to UNIT 2.1A for phenol extraction and ethanol precipitation.

8. Ethanol precipitate the aqueous layer containing RNA and DNA (~300 μ l) by adding 30 μ l of 3 M sodium acetate, pH 5.2, and 825 μ l absolute ethanol. Keep 30 min at -80°C, then microcentrifuge 30 min at maximum speed, 4°C.

Do not add glycogen.

9. Wash with 70% ethanol and resuspend the sizeable white pellet in 44.5 μ l nuclease-free water.

Oxidize diol structures

10. Prepare the following stocks fresh using 1.7-ml microcentrifuge tubes:

10 mM biotin hydrazide (long arm)
100 mM sodium periodate (NaIO₄).

11. Combine the following in a 1.5-ml siliconized tube:

44.5 μ l (-)cDNA/RNA heteroduplex
3 μ l 1.1 M sodium acetate, pH 4.5
2.5 μ l fresh 100 mM NaIO₄.

Incubate on ice 45 min in the dark.

12. Add the following to the 50- μ l reaction to precipitate the (-)cDNA/RNA heteroduplex:

0.5 μ l 10% SDS
11 μ l 5 M NaCl
61 μ l isopropanol.

Keep 30 min at -80°C, then microcentrifuge 30 min at maximum speed, 4°C.

13. Wash with 70% ethanol and resuspend the small white pellet in 50 μ l nuclease-free water.

Biotinylate RNA ends

14. To the 50 μ l oxidized (-)cDNA/RNA, add the following:

5 μ l 1 M sodium acetate, pH 6.1
5 μ l 10% SDS
150 μ l 10 mM fresh biotin hydrazide.

Incubate overnight (12 to 16 hr) at room temperature in the dark.

Perform RNaseONE selection

15. Precipitate the biotinylated (–)cDNA/RNA heteroduplex (210 μ l) by adding:

- 5 μ l 5 M NaCl
- 7 μ l 1 M sodium acetate, pH 6.1
- 750 μ l absolute ethanol.

Keep 30 min at -80°C , then microcentrifuge 30 min at maximum speed, 4°C .

16. Wash one time with 70% ethanol. Resuspend the small white pellet in 170 μ l nuclease-free water.

17. Use ~ 5 U RNaseONE/ μ g starting mRNA:

- 170 μ l biotinylated (–)cDNA/RNA heteroduplex
- 20 μ l $10\times$ RNaseONE buffer
- 4.5 μ l 10 U/ μ l RNaseONE
- 5.5 μ l nuclease-free water.

Incubate 30 min at 37°C .

18. Quench the reaction by adding 4 μ l of 10 mg/ml yeast tRNA and 50 μ l of 5 M NaCl.

Bind biotinylated (–)cDNA/RNA to beads

19. During the selection procedure, wash M-280 streptavidin Dynabeads (200 μ l bead suspension per RNA sample) as follows:

- a. Wash three times with 200 μ l of $1\times$ binding buffer, room temperature.
- b. Add 200 μ l of $1\times$ binding buffer plus 0.25 $\mu\text{g}/\mu\text{l}$ yeast tRNA and incubate 30 min at 4°C on a benchtop hot/cold shaking incubator at 800 rpm. Remove wash.
- c. Wash three times with 200 μ l of $1\times$ binding buffer, room temperature.

Before beginning, be familiar with the use of the magnetic stand for washing and separating beads. Use siliconized or low-binding tubes (e.g., LoBind) for all steps involving M-280 beads.

20. Remove supernatant from beads and add the ~ 254 μ l RNaseONE-treated (–)cDNA/RNA. Rotate 30 min at room temperature on a rotator for binding to occur.

21. Wash heteroduplex-bound beads at room temperature as follows:

- two times with 200 μ l of $1\times$ binding buffer
- one time with 200 μ l of $1\times$ blocking buffer
- one time with 200 μ l of $1\times$ wash buffer
- one time with 200 μ l of 50 $\mu\text{g}/\text{ml}$ yeast tRNA.

Perform hydrolytic degradation to release minus (–) strand flcDNA

22. Remove the supernatant (wash solution) from the beads, and add 50 μ l freshly made alkaline hydrolysis buffer. Shake 10 min at 65°C on a shaker at 1400 rpm.

The alkaline hydrolysis buffer must be made fresh before use.

23. Collect the supernatant containing full-length (–)cDNA (flcDNA) into a tube containing 150 μ l of 1 M Tris·Cl, pH 7.5, for neutralization.

24. Repeat the hydrolysis and collection steps two more times, collecting all fractions into the same tube, for a final volume of 300 μ l.

25. Perform an extraction with 25:24:1 phenol/chloroform/IAA, pH 7.9, then ethanol precipitate as follows to recover the (–) f/cDNA:

300 μ l full-length (–)cDNA
30 μ l 3 M sodium acetate, pH 5.2
4 μ l 15 mg/ml GlycoBlue
800 μ l absolute ethanol.

Keep 30 min at -80°C , then microcentrifuge 30 min at maximum speed, 4°C .

In this and subsequent extractions, the use of Phase Lock Gel Light tubes greatly facilitates the extraction procedure.

26. Wash with 70% ethanol. Resuspend the barely visible pellet in 5 μ l LoTE buffer or 10 mM Tris-Cl, pH 8.5 (EB buffer).

Synthesize double-stranded cDNA for cloning

27. Perform single-stranded linker (SSL) ligation of mixed 5' adapters by combining the following in a 1.7-ml microcentrifuge tube on ice:

5 μ l (–)f/cDNA
4 μ l 0.4 $\mu\text{g}/\mu\text{l}$ GIS-(N)5 adapter
1 μ l 0.4 $\mu\text{g}/\mu\text{l}$ GIS-(N)6 adapter
10 μ l Takara solution II
20 μ l Takara solution I (ligase).

Incubate overnight (12 to 16 hr) at 16°C .

28. Set up the primer extension reaction in a 1.7-ml thin-walled PCR tube on ice:

40 μ l overnight ligation reaction
20 μ l deionized water
8 μ l 10 \times ExTaq buffer with Mg^{2+}
8 μ l 2.5 mM dNTP mix
4 μ l 5 U/ μ l ExTaq polymerase.

29. Move tube straight from the ice to a 65°C preheated thermal cycler and start incubations as follows:

5 min at 65°C
30 min at 68°C
10 min at 72°C
hold at 4°C .

30. Add 2 μ l of 20 mg/ml proteinase K and incubate 15 min at 45°C to remove any remaining DNA polymerase.

31. Adjust volume to 200 μ l with deionized water, perform phenol/chloroform extraction, pH 7.9, and then ethanol precipitate as follows:

200 μ l ds cDNA
20 μ l 3 M sodium acetate, pH 5.2
2 μ l 15 mg/ml GlycoBlue
600 μ l absolute ethanol.

Keep 30 min at -80°C , then microcentrifuge 30 min at maximum speed, 4°C .

32. Wash with 70% ethanol and resuspend ds cDNA pellet in 70 μ l water. Set aside 5 μ l to run on an agarose gel (UNIT 2.5A) to check if ds cDNA synthesis was successful.
33. To remove the polyA tail, set up the following reaction in a 1.7-ml microcentrifuge tube (total 86 μ l):

65 μ l resuspended ds fcdDNA
8.6 μ l 10 \times Tango buffer
8.6 μ l 10 \times SAM (optional, but stimulates *GsuI* activity)
1.0 μ l 10 mg/ml BSA
2 μ l 5 U/ μ l *GsuI*
0.8 μ l water.

Digest 4 hr to overnight at 30°C (not 37°C).

An older buffer formulation, TANGO/Y+, contained BSA. Since the current Tango buffer does not, BSA must be added separately, as it is required by both GsuI and NotI.

The commercially available 32 mM SAM is diluted to a 500 μ M working solution, which is referred to as 10 \times SAM. This should be prepared fresh, as it is unstable.

Incomplete GsuI digestion will lead to problems; colleagues have observed the entire oligo-dT stretch (from the oligo-dT primer used for reverse transcription) ligated to the 3' adapter/vector 3' arm. This is probably due to a combination of failed GsuI digestion (and hence no removal of the polyA tail), A-overhang after ExTaq extension, and degraded TT overhang from the 3' adapter/pGIS4 vector 3' arm. It is critical to dispense the validated, processed pGIS4 cloning vector into aliquots and store them frozen.

34. To form the cohesive 5' terminal site, add the following to the reaction (final 100 μ l):

11.4 μ l 10 \times Tango buffer
2.6 μ l 10 U/ μ l *NotI*.

Incubate 4 hr at 37°C.

*The additional Tango buffer brings the buffer concentration to \sim 2 \times final (high salt). An excess of *NotI* is used because *NotI* is only 20% to 50% active in 2 \times Tango buffer.*

35. Heat inactivate the enzymes 15 min at 65°C and place tube on ice. Proceed directly to cDNA size fractionation.

Alternatively, perform sequential digestions with each enzyme. First digest with GsuI in Buffer B (Fermentas), followed by a spin-purification step to remove enzyme and buffer salts. Then digest with NotI in NEBuffer 3 (NEB). Proceed immediately to cDNA size fractionation.

Size fractionate cDNA

36. During the restriction digest, prepare cDNA size fractionation columns (one per sample) following manufacturer's instructions. Equilibrate column to room temperature. Remove the top cap, then the bottom, and allow to drain. Add 0.8 ml of 1 \times TEN buffer and allow to drain completely. Repeat three times (this process will take 1 hr).
37. Label twenty 1.7-ml microcentrifuge tubes to be used for fraction collection.
38. Add the 100 μ l digestion reaction (adjust volume with TEN buffer) to the prepared column. Collect the entire flowthrough in collection tube 1.
39. Add 100 μ l TEN buffer and collect the flowthrough in tube 2.

40. Add another 100 μl TEN buffer and start collecting single drops, 1 drop ($\sim 35 \mu\text{l}$) per tube, for subsequent tubes 3 to 20. Allow column to drain completely before adding another 100 μl TEN buffer.
41. Measure the absorbance of each fraction, using 10 μl per fraction in a 7- μl submicro-volume cell.

Alternatively, use a Nanodrop instrument, which requires only 1 μl of sample, or use PicoGreen-based fluorimetry (Molecular Probes; see Support Protocol 3).

42. *Optional:* Remove 4 μl from each fraction and run on an agarose gel to assist in determining which fractions are suitable for cloning. Depending on yields and technique, ethidium bromide staining may be sufficient to observe the fractions; otherwise SYBR Green I or GelStar stain can be used.

Fractions 7 through 10 are usually suitable.

43. Pool selected fractions, if desired, and ethanol precipitate (step 31) to concentrate the ds f1cDNA for ligation. Wash pellet with 70% ethanol and resuspend in 6 μl EB buffer.

Ligate f1cDNA to cloning vector pGIS4

44. Set up the following reaction in 1.7-ml microcentrifuge tubes on ice (total 10 μl):

1 μl 40 ng/ μl *NotI/BseRI*-cut pGIS4
6 μl ds f1cDNA fraction(s) (a minimum of 100 ng, but preferably ≥ 200 ng)
2 μl 5 \times ligase buffer with PEG
1 μl 5 U/ μl T4 DNA ligase.

Also set up a vector self-ligation control. Incubate overnight (12 to 16 hr) at 16°C.

45. Adjust volume to 200 μl with deionized water, and perform phenol/chloroform extraction and ethanol precipitation as in step 31. Wash the pellet at least two times with 70% ethanol, and resuspend in 20 μl EB buffer.

This removes salts prior to transformation.

Transform cells

46. Add 1 μl purified ligation reaction to 25 μl electrocompetent cells (e.g., OneShot electrocompetent TOP10 cells) in prechilled 1.7-ml microcentrifuge tubes. Gently stir with the pipet tip (do not pipet up and down). Let cells stand 5 min on ice.
47. Transfer to prechilled electroporation cuvettes (0.1 cm gap) and perform electroporation (*UNIT 9.3*).
48. Add 1 ml room temperature SOC medium within 10 sec of pulsing to maximize efficiency. Transfer cells to 15-ml conical tubes and allow cells to recover 1 hr at 37°C with shaking at 200 rpm.
49. Plate 20 to 50 μl (out of 1 ml) on LB-amp agar plates for quality control (QC) screening and library efficiency calculations. Incubate overnight at 37°C.

Perform QC on f1cDNA library

50. Count the numbers of colonies and determine library efficiency taking into consideration the self-ligation background. Pick colonies (24 to 48 colonies are sufficient and convenient) for screening by PCR (*UNIT 15.1*) using primers PMR011 and PMR012. Analyze PCR products by agarose gel electrophoresis.
51. If PCR shows a satisfactory range of insert sizes, pick one to four 96-well plates of colonies for overnight culture, purify plasmids (*UNIT 1.6*), and sequence to determine

full-length efficiency (by BLAST alignment of sequences against the GenBank nr database).

PCR products typically range from ~200 to >4000 bp. There should not be a predominance of a single-sized band, which would indicate a cloning problem.

The quality of the fcdDNA library can also be assayed using a commercial cDNA Integrity Kit (KPL). This kit contains sets of primers that are designed to amplify a range of inserts from human and rodent species: abundant versus rare and long versus short. It is also important to sequence a sufficient number of clones to have an idea of the full-length success rate in the library; it is critical that the full-length rate be high to successfully perform the next part of the GIS-PET method. Using this approach, the authors consistently obtain full-length efficiencies of >90%, defined as sequences that match known transcripts at both the 5' and 3' termini (including documented splice variants of known genes).

52. Store the library in the form of a purified ligation mix up to 2 years at -20°C (or indefinitely at -80°C) until ready to construct the GIS-PET library (see Basic Protocol 2).

CONSTRUCTION OF A GIS-PET LIBRARY (DITAGGING)

This procedure is outlined on the right side of Figure 21.12.1. The ds fcdDNA clones constructed in Basic Protocol 1 contain one *MmeI* site (TCCGAC) flanking the 5' terminus of the cDNA insert, and another on the minus strand flanking the 3' end of the insert. The *MmeI* restriction enzyme is used to cleave these clones 20 bp/18 bp into the cDNA fragments from both their 5' and 3' ends, thereby forming the 5' and 3' tags. Consequently, despite the variable sizes of the cDNA inserts, the vector-plus-40-bp cDNA signature tags (linear single-PET plasmids) will be of a constant size (~ 2800 bp) that can be easily recognized using agarose gel electrophoresis, and can be purified from the unwanted cDNA fragments that are a byproduct of the *MmeI* digestion. The gel-purified linear single-PET plasmids are then blunt-ended (to remove any incompatible 3' overhangs) and self-ligated to produce "single-PET plasmids" containing the 5' and 3' GIS tags now linked in the form of PETs. The entire pool of these single-ditag plasmids thus represents the GIS single-PET library. In the final steps, plasmid DNA is extracted from the single-PET library and digested with *BamHI* to obtain the actual GIS PETs. These are gel-purified, concatenated, size-fractionated, and ligated to prepared pZErO-1 vector to yield the final GIS-PET library.

Materials

Purified ligation mix (see Basic Protocol 1)
LB medium (*UNIT 1.1*; use Lennox formulation containing 5 g/liter NaCl)
HiSpeed Plasmid Maxi kit (Qiagen) or equivalent
2 U/ μl *MmeI* (NEB) and 32 mM *S*-adenosyl methionine (SAM; supplied with *MmeI*)
10 \times NEBuffer 4 (NEB)
pGIS4 cloning vector (obtain from authors; prepare as in Support Protocol 1)
Qiaquick gel extraction kit (Qiagen)
10 \times Tango buffer (Fermentas)
0.1 M DTT
5 to 10 U/ μl T4 DNA polymerase (Promega)
10 mM dNTP mix (Invitrogen)
25:24:1 (v/v/v) phenol/chloroform/IAA, pH 7.9 (Ambion)
3 M sodium acetate, pH 5.2 (*APPENDIX 2*)
15 mg/ml GlycoBlue (Ambion)
Absolute and 70% ethanol
Solution I (enzyme solution) from DNA ligation kit (Takara version 2.1)

BASIC PROTOCOL 2

Chromatin Assembly and Analysis

21.12.11

EB buffer (10 mM Tris·Cl, pH 8.5; Qiagen)
 Electrocompetent cells (e.g., OneShot electrocompetent TOP10 cells)
 SOC medium (see recipe)
 LB-amp agar plates (*UNIT 1.1*; use Lennox formulation with 5 g/liter NaCl)
 20 U/μl *Bam*HI and 10× unique *Bam*HI buffer (NEB)
 10 mg/ml BSA
 1 M MgCl₂
 LoTE buffer (see recipe)
 Wide-range DNA ladder (50 to 10,000 bp; Takara)
 4% to 20% PAGE minigel (TBE; 1.0-mm, 10 wells; Invitrogen)
 Low DNA mass ladder (Invitrogen)
 25- and 100-bp DNA ladders (Invitrogen)
 29:1 (v/v) acryamide/bis solution (BioRad)
 5× TBE buffer (*APPENDIX 2*)
 TEMED
 10% ammonium persulfate (APS)
 SYBR Green I (Molecular Probes)
 7.5 M ammonium acetate stock solution
 5 U/μl T4 DNA ligase and 10× ligase buffer with spermidine (see recipe for ligase buffer)
 Qiaquick PCR purification kit (Qiagen)
 6× loading dye (Fermentas)
 pZerO-1 cloning vector (Invitrogen)
 5× ligase buffer with PEG (Invitrogen)
 LB agar plates (*UNIT 1.1*; use Lennox formulation with 5 g/liter NaCl) with 25 μg/ml Zeocin (Invitrogen)
 20 μM PMR011 and PMR012 PCR primers (Fig. 21.12.2)
 22 × 22-cm Q-trays (Genetix) or other large agar plates for bacterial plating
 Disposable plastic spreaders (e.g., Lazy L-spreaders, Sigma), optional
 500-ml plastic centrifuge bottles
 Floor-standing ultracentrifuge (e.g., Sorvall with an SLA3000 rotor)
 Phase Lock Gel Light tubes (Eppendorf), optional
 0.6- and 1.7-ml microcentrifuge tubes
 Electroporation cuvettes (0.1-cm gap; BioRad), prechilled
 15-ml conical tubes (e.g., Falcon)
 Hot/cold shaking incubator (Eppendorf)
 Disposable ElutaTubes (Fermentas) or equivalent
 Dark Reader blue-light transilluminator (Clare Chemical)
 21-G needles
 Microspin plastic centrifuge tube filter units (0.45-μm; Costar Spin-X or BioRad Mermaid), sterile
 Nanodrop spectrometer (optional)
 96-well microtiter plates
 Additional reagents and equipment for plasmid preparation (*UNITS 1.6 & 1.7*), restriction enzyme digestion (*UNIT 3.1*), agarose gel electrophoresis (*UNIT 2.5A*), DNA quantitation (see Support Protocol 3 and *APPENDIX 3D*), phenol extraction and ethanol precipitation (*UNIT 2.1A*), ligation (*UNITS 3.14 & 3.16*), electroporation (*UNIT 9.3*), electroelution (*UNIT 2.6*), PAGE (*UNIT 10.2A*), validation of cloning vector (see Support Protocol 1), PCR (*UNIT 15.1*)

Prepare plasmid DNA

1. Transform all of the remaining purified ligation mix and amplify once by plating an appropriate number of clones (60,000 to 100,000 cfu/tray to allow growth of each colony without excessive overcrowding) on large (22 × 22-cm) agar plates

(e.g., Genetix Q-trays). Use a maximum of 700 μl of culture per Q-tray. Incubate overnight at 37°C.

The number of colonies required is determined by the estimated transcriptome size; 1×10^6 cfu can be used as a convenient benchmark. This number is estimated to give a $10\times$ over-coverage of a transcriptome population of 100,000. (The transcriptome size itself is estimated as 100,000 because 35% to 59% of the estimated 30,000 to 40,000 human genes undergo alternative splicing to give an average of three alternative transcripts per gene.) Although only a small fraction of the expected milligram quantities of plasmid DNA obtained is subsequently used for MmeI digestion, it is still critical for the sake of proper representation/complexity to first obtain the benchmark 1×10^6 cfu.

Growth on agar (solid-phase plasmid amplification) is important, as this maintains representation better than growth in liquid culture. This is because, on a solid substrate, each clone can grow without competition from other bacterial clones.

2. Harvest the resulting bacterial colonies by manually scraping into LB medium (20 to 30 ml per Q-tray) and transfer to 500-ml plastic centrifuge bottles. Centrifuge cells 20 min at $\sim 4000 \times g$, 4°C, in a floor-standing ultracentrifuge.

Disposable plastic spreaders (e.g., Lazy L-spreaders, Sigma) are convenient for harvesting cells.

3. Prepare plasmid DNA using the HiSpeed Plasmid Maxi kit or any other preferred method.

Ten Q-trays of bacteria produce ~ 1 mg of plasmid DNA. Therefore, in theory, two Qiagen Maxi tips (per library) should be sufficient. However, to avoid clogging due to the large volumes involved, the authors usually use three to four tips per library. Suggested volumes of buffers to use with bacteria from ten Q-trays are 20 to 50 ml each of P1, P2, and P3 buffers.

Also see UNIT 1.7 for plasmid preparation.

Perform MmeI digestion (tagging)

4. Digest ~ 10 μg plasmid DNA using MmeI according to the manufacturer's recommended conditions, ensuring that the number of units of enzyme is always in less than a four-fold excess to prevent methylation-induced inhibition. Suggested reaction conditions are as follows (total 200 μl):

100 μl (10 μg) plasmid DNA
20 μl 10 \times NEBuffer 4
20 μl 10 \times SAM
12 μl 2 U/ μl MmeI
48 μl deionized water.

Incubate 4 hr to overnight at 37°C.

The 10 \times SAM should be prepared fresh, as it is unstable. The 10 \times concentration should be 500 μM , and the final SAM concentration should be ~ 50 μM .

An overnight digest usually works best. The main band of ~ 2800 bp is the desired linear single-PET plasmid.

5. Purify the entire digestion reaction on a 0.7% agarose gel (UNIT 2.5A), loading the digestion products in as few lanes as possible to facilitate excision. Run controls on the same gel, comprising uncut as well as linearized pGIS4 cloning vector, to ensure that the correct band is excised.

An ~ 2800 -bp band corresponding to linear single-PET plasmid DNA should be the dominant species in the gel. Unwanted digestion products will be present as a faint smear behind the 2800-bp band.

- Excise the 2800-bp linear single-PET plasmid DNA band and purify using a Qiaquick gel extraction kit.

The number of Qiaquick spin columns to use depends on the amount of DNA that was digested. Remember that each column has a maximum capacity of ~10 µg DNA.

- Quantify the amount of DNA recovered from the gel by spectrometry (APPENDIX 3D).

The recovery rate is consistently ~50%.

Perform self-ligation to create circular single-PET plasmids

- Set up the following reaction to create blunt ends:

50 µl (0.5 to 2.0 µg) DNA
6.0 µl 10× Tango buffer
0.3 µl 0.1 M DTT
5 U T4 DNA polymerase/µg DNA
0.6 µl 10 mM dNTP mix
Deionized water to 60.0 µl.

Incubate 5 min at 37°C, then inactivate 10 min at 75°C.

MmeI digestion (TCC(A/G)AC 20/18) results in a random 2-base 3'-overhang at the ends of both 5' and 3' signature tags. These must be removed by end-polishing using T4 DNA polymerase, leaving behind 18-bp blunt-ended tags.

Alternatively, use the Epicentre End-It blunting kit, which does not require any optimization, and works well for up to 5 µg per reaction.

- Adjust to 200 µl with deionized water, extract with 25:24:1 phenol/chloroform/IAA, pH 7.9, and perform ethanol precipitation as follows:

200 µl blunt-ended DNA
20 µl 3 M sodium acetate, pH 5.2
2 µl 15 mg/ml GlycoBlue
600 µl absolute ethanol.

Keep 30 min at –80°C, then microcentrifuge 30 min at maximum speed, 4°C.

Refer to UNIT 2.1A for phenol extraction and ethanol precipitation. The use of Phase Lock Gel Light tubes greatly facilitates the extraction procedure.

- Wash with 70% ethanol. Resuspend the DNA pellet in LoTE or EB buffer. Dilute an aliquot to ~2 ng/µl, assuming an arbitrary 100% recovery.

A very dilute DNA solution is desired (final DNA concentration of 1 to 2 ng/µl) to favor intramolecular ligation.

- Set up the self-ligation (intramolecular recircularization) reaction on ice by combining 50 µl (~100 ng) DNA with 50 µl DNA ligation solution I (from Takara ligation kit version 2.1). Incubate overnight (12 to 16 hr) at 16°C.

Alternatively, use 5× ligation buffer with PEG (Invitrogen) and 5 U T4 DNA ligase. The final DNA concentration should be adjusted to 1 to 2 ng/µl.

- Adjust volume to 200 µl with deionized water, and perform phenol/chloroform extraction and ethanol precipitation as in step 9. Wash the pellet at least two times with 70% ethanol, and resuspend in 20 µl EB buffer.

This removes salts before electroporation.

Transform cells

13. Add 1 μl purified ligation reaction to 25 μl electrocompetent cells (e.g., OneShot electrocompetent TOP10 cells) in prechilled 1.7-ml microcentrifuge tubes. Gently stir with the pipet tip (do not pipet up and down). Let cells stand 5 min on ice.
14. Transfer to prechilled electroporation cuvettes (0.1-cm gap) and perform electroporation (UNIT 9.3).
15. Add 1 ml room temperature SOC medium within 10 sec of pulsing to maximize efficiency. Transfer cells to 15-ml conical tubes and allow cells to recover 1 hr at 37°C with shaking at 200 rpm.
16. Plate 20 to 50 μl (out of 1 ml) on LB-amp agar plates for quality control (QC) screening and library efficiency calculations. Incubate overnight at 37°C.

Perform QC on GIS single-PET library

17. Plate dilutions of 1:100 to 1:1000 in a final 100- μl volume made up in SOC medium. Incubate overnight at 37°C. On the following day, examine the colonies growing on the agar.

Usually, even the 1:1000 plating is almost confluent (i.e., several thousand colonies). It is important that there are many colonies growing at this stage, as this indicates that the self-circularization step was efficient, and that the resulting single-PET library will be of sufficient complexity.

18. For large-scale plating, based on the number of colonies observed, spread enough of the remaining transformed bacterial culture on Q-trays to obtain at least the same number (or greater) of colonies as in the original f1cDNA library. Incubate overnight at 37°C.

This ensures that the library remains representative of the original sample.

Release single PETs using BamHI

19. Harvest bacterial colonies by scraping into LB medium (20 to 30 ml per Q-tray) and centrifuge 20 min at $\sim 4000 \times g$, 4°C, in a floor-standing ultracentrifuge. Prepare plasmid DNA using the HiSpeed Plasmid Maxi kit (or any other preferred method).
20. Digest ~ 100 to 500 μg plasmid DNA using *Bam*HI. Suggested reaction conditions are:

500 μg plasmid DNA ($x \mu\text{l}$)
100 μl 10 \times unique *Bam*HI buffer
10 μl 10 mg/ml BSA
50 μl 20 U/ μl *Bam*HI
Deionized water to 1 ml.

Pipet into 100- μl aliquots for enhanced digestion efficiency and incubate overnight at 37°C.

There is no upper limit to the amount of DNA being cut here. It is advisable to cut as much plasmid DNA as possible to obtain a large quantity of PETs.

*A two-fold excess of *Bam*HI is used. Better downstream results have been obtained using these conditions compared with a five-fold excess of enzyme.*

Purify BamHI-cohesive single PETs

21. To remove buffer salts and reduce volume, extract with 25:24:1 phenol/chloroform/IAA, pH 7.9, and ethanol precipitate as follows:

200 μ l digest
20 μ l 3 M sodium acetate, pH 5.2
4.5 μ l 1 M $MgCl_2$
2 μ l GlycoBlue
600 μ l absolute ethanol.

Keep 30 min at $-80^\circ C$, then microcentrifuge 30 min at maximum speed, $4^\circ C$. Wash with 70% ethanol.

If buffer salts are not removed, they will distort the subsequent gel run (especially for PAGE) and thus affect purification of the 50-bp PETs. The addition of $MgCl_2$ during precipitation enhances recovery of short fragments.

For convenience, PETs are most often run on a 2% agarose gel and then electroeluted (steps 21 to 27a). This is a lower-resolution method and does not purify PETs as efficiently as PAGE, which was used when developing the original procedure (steps 21 to 29b). The effect of these impurities is manifested as damaged sequences that “poison” the concatenation step, leading to shorter concatemers. Nonetheless, agarose gel electrophoresis is easier to set up and faster to run, and suitable if there are no problems with the single-PET library (which enables one to obtain pure PETs). It is critical to examine the quality of agarose-gel-purified PETs (e.g., by running a mini PAGE gel) before concatenation.

For agarose gels

22a. Resuspend the pellets from all the tubes in a total of 100 μ l LoTE buffer or EB buffer.

23a. Load on a 2% agarose gel along with 20 μ l wide-range DNA ladder. Run at 80 V for ~ 1.5 hr and visualize at 365 nm UV to confirm good separation of PETs from other DNA.

It is convenient to run a wide-range DNA ladder for sizing.

24a. Excise the PET band from the agarose gel and electroelute using disposable ElutaTubes (or equivalent). Follow the manufacturer’s instructions for maximum gel size per tube. Perform electroelution for 1 to 1.5 hr at 90 V.

Also see UNIT 2.6 for electroelution from agarose gels.

25a. Precipitate the harvested PETs as in step 21. Resuspend all precipitated PET DNA in 20 μ l LoTE buffer.

26a. Run 0.5 μ l and 1 μ l on a 4% to 20% gradient polyacrylamide minigel (UNIT 10.2A) together with a low DNA mass ladder and a 25-bp DNA ladder to check for purity and to quantify PETs by visual comparison to mass ladders.

27a. *Optional:* Quantify PETs using PicoGreen (see Support Protocol 3). Proceed to step 30.

Some investigators have found that visual estimation of PET DNA quantity results in subjective variability, and hence uncertainty in subsequent enzymatic digestion steps. Although not critical, PicoGreen quantitation (which is not affected by the presence of glycogen/GlycoBlue) can be useful in providing a more accurate quantitation of DNA.

For polyacrylamide gels

22b. Resuspend and pool the DNA pellets in LoTE buffer at ~ 10 μ g per 20 μ l, which is suitable for gel-loading.

23b. Cast a 15 × 15-cm, 1.5-mm thick, 10-well, 10% polyacrylamide gel according to the following recipe:

10 ml 40% acryamide/bis (29:1) solution
8 ml 5× TBE buffer
16 μl TEMED
21.6 ml deionized water
When ready to pour, add 0.4 ml of 10% APS.

Refer to UNIT 10.2A for PAGE.

24b. Load no more than ~10 to 15 μg per well, as excessive DNA loading results in self-quenching of fluorescence. Load 1 μl (1 μg) of 25-bp DNA ladder into one well. Electrophorese at 200 V until the bromphenol blue band almost reaches the bottom of the gel.

25b. Stain the gel with SYBR Green I for 30 min and visualize on a Dark Reader blue-light transilluminator.

It is preferable to use the Dark Reader transilluminator for visualization, as exposure to UV light (especially short-wavelength UV) will damage DNA. If this is not possible, at least ensure that long-wavelength (365-nm) UV light is used.

26b. Excise the 50-bp GIS-PETs and collect gel fragments into 0.6-ml microcentrifuge tubes that have been pierced at the bottom with a 21-G needle. Use DNA from two lanes per tube. Place the pierced tubes inside 1.7-ml microcentrifuge tubes and microcentrifuge 5 min at maximum speed, 4°C.

The gel pieces are conveniently shredded and collected at the bottom of each 1.7-ml microcentrifuge tube. This is known as the gel-crush method.

27b. Add 150 μl of 125:25 (v/v) LoTE buffer/ammonium acetate to each tube and elute overnight (12 to 16 hr) at 4°C and then 30 min to 2 hr at 37°C.

28b. Separate the supernatant (containing eluted 50-bp PETs) from the gel pieces with the aid of microspin plastic centrifuge tube filter units (e.g., SpinX or Mermaid). Using a 1-ml pipet tip, aspirate as much liquid and gel into each column, then microcentrifuge 10 min at maximum speed, 4°C.

29b. Pool the collected supernatant and precipitate as in step 21. Resuspend the pellets from all the tubes in a total of 20 μl LoTE buffer and proceed to quantitation by visual comparison to mass ladders or by PicoGreen fluorimetry (see Support Protocol 3).

A phenol/chloroform extraction prior to precipitation is optional.

The presence of GlycoBlue (or glycogen) at this stage precludes spectrophotometric quantitation.

Concatenate PETs

30. Set up the following ligation reaction:

200 to 1000 ng 50-bp cohesive PETs
1 μl 10× ligase buffer with spermidine
1 μl 5 U/μl T4 DNA ligase
Deionized water to 10 μl (or 20 μl if PETs are too dilute).

Incubate 2 hr to overnight at 16°C.

Addition of spermidine is important for enhancing ligase activity and favoring linear concatenation.

Overnight incubation does not appear to affect the concatenation efficiency of agarose-gel-purified PETs, which is somewhat disconcerting, since over-concatenation would be expected. Apparently there is a limit to concatenation efficiency due to impurities that “poison” the ligation reaction. In the case of PAGE-purified PETs, some optimization must be done with regard to ligation time, as over-concatenation can occur due to the higher purity of the recovered PETs. Over-concatenation will result in excessively high-molecular-weight DNA that is difficult to clone. The authors find that the >2-kb fraction (in step 36) clones poorly (high incidence of empty clones) unless special measures are taken (a much higher insert-to-vector ratio of up to 50:1; incubation of transformed cells for longer periods). Although ligation is done with this high-molecular-weight fraction as a matter of course, it is seldom actually used in the final sequencing, as the clones from the lower-molecular-weight fractions provide sufficient PET data.

31. Purify the concatemer DNA using the Qiaquick PCR purification kit, keeping in mind the 10- μ g DNA capacity of each spin column. Elute DNA with 50 μ l EB buffer. Quantitate DNA by Nanodrop (if available).

If Nanodrop is not available, quantitate using another spectrophotometer.

32. Perform a partial *Bam*HI digest as follows:

50 μ l concatemer DNA
6 μ l 10 \times *Bam*HI buffer
1 U *Bam*HI/ μ g DNA (up to a maximum of 3 U/ μ g DNA)
1 μ l 10 mg/ml BSA
Deionized water to 60 μ l.

Incubate 30 min at 37°C.

Concatenated PETs are partially redigested with BamHI to ensure the presence of cohesive termini suitable for insertion into BamHI-digested pZErO-1 vector. This step is critical to the success of concatemer cloning. A partial digestion apparently ensures the recreation of “correct” BamHI cohesive ends on the concatemer termini. Without this step, a significant fraction of the concatemers fail to insert into the vector in the next step. This BamHI redigestion is performed only on BamHI cohesive-ended concatemers.

Do not exceed the 30-min digestion time as this will destroy the concatemers. It is helpful initially to visualize the digestions at several time points on a polyacrylamide minigel.

33. Quickly add 6 μ l of 6 \times loading dye, heat 15 min at 65°C, and then chill on ice.

This step minimizes nonspecific aggregation/annealing of DNA, which might otherwise give a false picture of concatenation efficiency.

Purify concatenated PETs

34. Load the entire sample (preferably into one well to facilitate recovery) on a 4% to 20% gradient polyacrylamide minigel, flanked by suitable DNA ladders to allow sizing.

Suggested markers are the 25-bp and 100-bp ladders from Invitrogen, and the Takara wide-range DNA ladder. The use of multiple DNA ladders facilitates accurate sizing.

35. Electrophorese \sim 1 hr at 200 V or until the bromphenol blue tracking dye is at the bottom of the gel. Stain for 15 to 30 min in SYBR Green I and visualize on a Dark Reader transilluminator for band excision.

36. Excise the concatenated DNA in three separate fractions by molecular weight: low (400 to 1000 bp), medium (1000 to 2000 bp), and high (>2000 bp).

It is important when excising the high-molecular-weight fraction to avoid collecting the DNA trapped in the wells. This debris appears to poison the subsequent ligation (or is itself favorably inserted) such that no high-molecular-weight concatemers are cloned.

37. Place the gel slice of each excised size fraction into a 0.6-ml microcentrifuge tube that has been pierced at the bottom with a 21-G needle and extract concatemers by the gel-crush method (step 26b).
38. Add 200 μ l of 167:33 (v/v) LoTE buffer/ammonium acetate to each tube and elute by heating 2 hr at 65°C.
39. Separate the supernatant (containing eluted concatenated PETs) from the gel pieces with the aid of microspin filter units (step 28b).
40. Perform phenol/chloroform extraction on each eluted size fraction and then ethanol precipitate as in step 9. Wash each pellet with 70% ethanol and resuspend each in 6 μ l LoTE or EB buffer.

If there is more than one pellet per size fraction, pool pellets from the same size fraction and resuspend in a total volume of 6 μ l LoTE or EB buffer.

The excised DNA can also be recovered by electroelution using ElutaTubes.

Ligate concatenated PETs to pZErO-1 vector

41. Prepare the cloning vector by digesting 2 μ g pZErO-1 plasmid DNA with 10 U *Bam*HI for 2 hr at 37°C. Perform phenol/chloroform extraction and ethanol precipitation as in step 9, wash the pellet with 75% ethanol, and resuspend in LoTE buffer at a concentration of 33 ng/ μ l. Validate preparation before use (see Support Protocol 1).

Agarose gel purification of the vector is not necessary.

42. Set up the ligation as follows (total 10 μ l):

6 μ l concatemer DNA fraction
1 μ l *Bam*HI-digested pZErO-1 DNA
2 μ l 5 \times ligase buffer with PEG
1 μ l 5 U/ μ l T4 DNA ligase.

Also set up a vector self-ligation in parallel as a control. Incubate overnight at 16°C. Do not heat inactivate.

43. Adjust volume to 200 μ l with deionized water, and perform phenol/chloroform extraction and ethanol precipitation as in step 9. Wash the pellet at least two times with 70% ethanol, and resuspend in 20 μ l EB buffer.

This removes salts prior to transformation.

Transform cells

44. Add 1 μ l purified ligation reaction to 25 μ l electrocompetent cells (e.g., OneShot electrocompetent TOP10 cells) in prechilled 1.7-ml microcentrifuge tubes. Gently stir with a pipet tip (do not pipet up and down). Let cells stand 5 min on ice.
45. Transfer to prechilled electroporation cuvettes (0.1-cm gap) and perform electroporation.
46. Add 1 ml room temperature LB medium (not SOC) within 10 sec of pulsing to maximize efficiency. Transfer cells to 15-ml conical tubes (or equivalent) and allow cells to recover 1 hr at 37°C, shaking at 200 rpm.

Surprisingly, improved results (i.e., higher titers) have been confirmed using LB medium instead of SOC medium.

47. Plate 20 to 50 μ l (out of 1 ml) on on a small LB agar plate plus Zeocin (25 μ g/ml final concentration). Incubate overnight at 37°C.

Perform QC on GIS-PET library

48. Count the number of colonies and determine library efficiency taking into consideration the self-ligation background. Pick colonies (24 to 48 is sufficient and convenient) for PCR screening (*UNIT 15.1*) using PMR011 and PMR012 primers. Include a control PCR on the pZerO-1 vector itself. Analyze PCR products by agarose gel electrophoresis.
49. If PCR shows a satisfactory range of insert sizes, pick one to four 96-well plates of colonies for overnight culture (in low-salt LB medium plus 25 $\mu\text{g}/\text{ml}$ Zeocin), purify plasmids (*UNIT 1.6*), and sequence to determine the average number of PETs per insert.
50. Store the library in the form of a purified ligation mix at -20°C until ready to perform large-scale transformations, plasmid extractions, and sequencing of GIS-PETs.

BASIC PROTOCOL 3

CONSTRUCTION OF A ChIP-DNA PLASMID LIBRARY

Chromatin immunoprecipitation (ChIP)-PET analysis (Fig. 21.12.3) is a minor modification of the GIS-PET analysis procedure. ChIP-enriched genomic DNA (gDNA) fragments are end-polished and inserted into another unique cloning vector (pGIS3) that has the requisite restriction enzyme recognition sites built into the vector arms flanking the inserted DNA. This results in the formation of a ChIP-DNA library (analogous to the fICDNA library in GIS-PET). The ChIP-DNA library is then ditagged exactly as described in Basic Protocol 2.

This procedure assumes that the reader is familiar with the construction of a GIS fICDNA library as described in Basic Protocol 1, especially with regard to basic molecular biology procedures. The method of ChIP is not covered in this protocol (see *UNIT 21.3*). For best results, the de-linked, fragmented ChIP-enriched gDNA should show a smear between 100 and 2000 bp, centering around 500 bp. These parameters can be met by adjusting

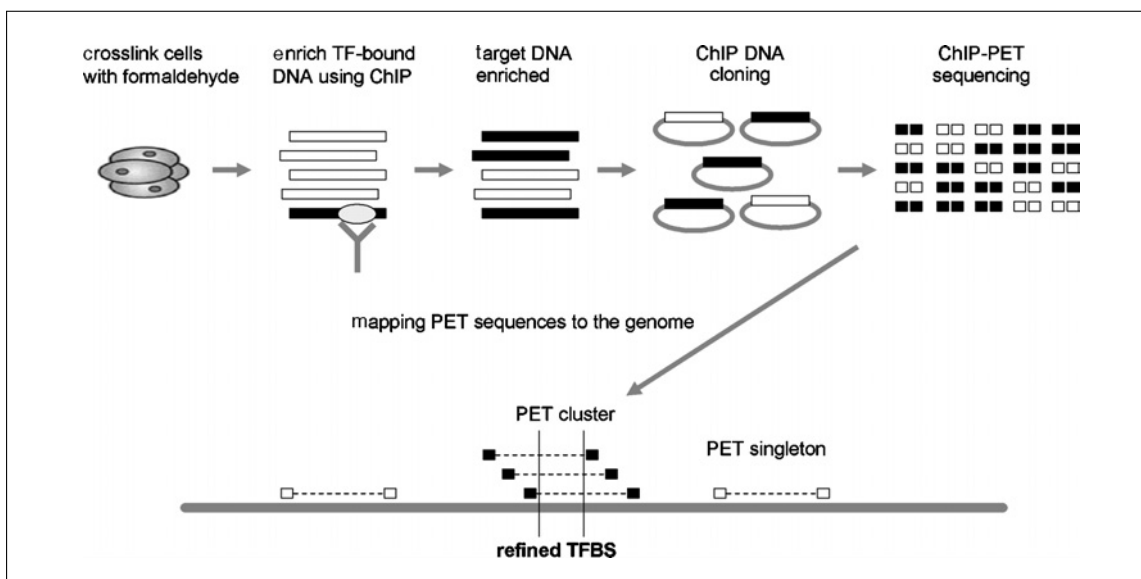


Figure 21.12.3 An overview of the ChIP-PET analysis method for mapping ChIP-enriched genomic DNA fragments.

sonication conditions. In addition, the minimum starting quantity of DNA should be 100 ng. Due to the interfering presence of glycogen in the ChIP-enriched gDNA fragments, quantitation must be done by PicoGreen fluorimetry (see Support Protocol 3).

Materials

End-It DNA end-repair kit (Epicentre) containing:

End-repair enzyme mix: T4 DNA polymerase and T4 polynucleotide kinase
End-repair 10× buffer
dNTP solution
ATP

Sample of interest: ChIP-enriched genomic DNA fragments (at least 100 ng)

25:24:1 (v/v/v) phenol/chloroform/IAA, pH 7.9 (Ambion)

3 M sodium acetate, pH 5.2 (*APPENDIX 2*)

70% and absolute ethanol

GlycoBlue (Ambion)

1× TEN buffer (see recipe)

pGIS3 cloning vector (obtain from authors; prepare as in Support Protocol 1)

5 U/μl T4 DNA ligase and 5× T4 DNA ligase buffer with PEG (Invitrogen)

EB buffer (10 mM Tris-Cl, pH 8.5; Qiagen)

20 μM PMR011 and 20 μM PMR012 PCR primers (Fig. 21.12.2)

Phase Lock Gel Light tubes (Eppendorf), optional

Heat/chill block or thermal cycler

Additional reagents and equipment for size-fractionation, transformation, and PCR screening (see Basic Protocol 1)

1. Set up an end-polishing reaction using the End-It kit (total 50 μl):

34 μl (100 to 200 ng) ChIP-enriched gDNA

5 μl end-repair 10× buffer

5 μl 2.5 mM dNTP mix

5 μl 10 mM ATP

1 μl end-repair enzyme mix.

Incubate 45 min at 25°C, then inactivate 10 min at 70°C.

The End-It kit is convenient as it simultaneously blunts and 5'-phosphorylates the DNA termini.

2. Adjust volume to 200 μl with deionized water, extract with 25:24:1 phenol/chloroform/IAA, pH 7.9, and ethanol precipitate as follows:

200 μl end-polished DNA

20 μl 3 M sodium acetate, pH 5.2

2 μl 15 mg/ml GlycoBlue

600 μl absolute ethanol.

Keep 30 min at –80°C, then microcentrifuge 30 min at maximum speed, 4°C.

Refer to UNIT 2.1A for phenol extraction and ethanol precipitation. As in Basic Protocol 1, the use of Phase Lock Gel Light tubes greatly facilitates the extraction procedure.

3. Wash with 70% ethanol. Resuspend the pellet of purified, blunted, and phosphorylated DNA in 100 μl of 1× TEN buffer for size fractionation. Size-fractionate the DNA on cDNA size fractionation columns (one per sample; see Basic Protocol 1, steps 36 to 43).

The objective of the optional electrophoresis step (see Basic Protocol 1, step 42) is to remove any DNA fragments that are too small (<100 bp), as these will be cloned preferentially, but are relatively uninformative relative to the bulk of the fragmented DNA (which should be centered between 200 and 500 bp). Fractions 7 to 12 are usually suitable.

4. To ligate DNA to the cloning vector, set up the following reactions on ice using 1.7-ml microcentrifuge tubes (total 10 μ l):

- 1 μ l 40 ng/ μ l prepared pGIS3
- 6 μ l selected ChIP DNA fractions
- 2 μ l 5 \times ligase buffer with PEG
- 1 μ l 5 U/ μ l T4 DNA ligase.

Also set up a vector self-ligation control. Incubate overnight (12 to 16 hr) at 16°C.

5. Adjust volume to 200 μ l with deionized water, and perform phenol/chloroform extraction and ethanol precipitation as in step 2.
6. Wash the pellet at least two times with 70% ethanol, and resuspend in 20 μ l EB buffer.

This removes salts prior to transformation.

7. Transform cells by electroporation and pick colonies for PCR screening using PMR011 and PMR012 primers (see Basic Protocol 1, steps 46 to 50).

A range of different-sized PCR products should be observed, corresponding in size to the inserted DNA fractions.

8. Store the library in the form of a purified ligation mix at -20°C .
9. Construct the ChiP-PET library (ditagging) as described for GIS-PET (see Basic Protocol 2).

SUPPORT PROTOCOL 1

PREPARATION AND VALIDATION OF pGIS4 AND pGIS3 CLONING VECTORS

The pGIS4 vector (Fig. 21.12.4) was created as an improvement to the published prototype pGIS1, and renders 3'-adapter ligation unnecessary. It contains an \sim 800-bp stuffer that must be excised before use by *NotI* and *BseRI* digestion, which simultaneously creates a *NotI* cohesive site to receive the 5'-adapter terminal region of the cDNA, and a TT overhang ready to receive the corresponding AA overhang at the 3' end of the prepared cDNA insert.

The pGIS3 vector (Fig. 21.12.5) was created to facilitate direct ditagging of inserted blunt-ended DNA fragments. The restriction sites important for ditagging were built into the vector backbone itself. There is no stuffer fragment in pGIS3.

Materials

- pGIS4 and pGIS3 plasmids (obtain from authors; ngwp@gis.a-star.edu.sg)
- HiSpeed plasmid maxi kit (Qiagen)
- Qiaquick gel extraction kit (Qiagen)
- 10 \times NEBuffer 2 and NEBuffer 3 (NEB)
- 10 U/ μ l *NotI* restriction enzyme (NEB)
- 4 U/ μ l *BseRI* restriction enzyme (NEB)
- 10 mg/ml BSA
- EB buffer (Qiagen)
- Compatible stuffers for validation

5 U/μl T4 DNA ligase and 5× T4 DNA ligase buffer with PEG (Invitrogen)
20 U/μl *Xho*I restriction enzyme (NEB)
70 U/μl mung bean nuclease and 10× reaction buffer (Promega)
10% SDS
Qiaquick PCR purification kit (Qiagen)
5 U/μl Antarctic phosphatase and 10× Antarctic phosphatase buffer (NEB)
TE buffer (APPENDIX 2; optional)
Nanodrop spectrometer (optional)
Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5A) and transformation (see Basic Protocol 1)

For pGIS4 plasmid DNA

- 1a. Purify pGIS4 plasmid using the HiSpeed plasmid maxi kit according to the manufacturer's instructions.
- 2a. To ensure that no contaminating bacterial genomic DNA is subsequently inserted into the prepared vector, perform an extra agarose gel purification (UNIT 2.5A) of the maxiprep to isolate supercoiled DNA.
- 3a. Recover DNA using the Qiaquick gel extraction kit according to the manufacturer's instructions (recovery is ~50%).

The maximum capacity of each Qiaquick spin tube is only 10 μg, so be sure to predetermine the number of tubes needed.

- 4a. Set up the digestion as follows:

5 μg purified supercoiled pGIS4
20 μl 10× NEBuffer 3
5 μl 10 U/μl *Not*I (five-fold excess)
12.5 μl 4 U/μl *Bse*RI (five-fold excess)
2 μl 10 mg/ml BSA
Water to 200 μl.

Incubate as two 100-μl aliquots no more than 3 hr at 37°C.

It is important to follow the conditions given here closely, as BseRI exhibits some nonspecific activity and is also only 75% active in NEBuffer 3.

- 5a. Purify by agarose gel electrophoresis. Quantitate using a Nanodrop (if available) and resuspend in EB buffer at 40 ng/μl.
- 6a. To validate, set up three test ligations as shown in Table 21.12.1. Incubate overnight at 16°C.

*This procedure is generally applicable to any cloning vector. The purpose of vector validation is to ensure that there is minimal contamination of the *Not*I/*Bse*RI-cut vector with either uncut or single-cut plasmid, which would otherwise contribute to unwanted background.*

*In this case, a compatible stuffer would be a dsDNA fragment with a cohesive 5'-*Not*I site, and a 3'-AA overhang. It is possible to simply gel extract the smaller digestion product (after the *Not*I/*Bse*RI vector digestion) and use this as an insert; however, any nonspecific cutting would be preserved in this product, giving rise to false positives in the positive control. It is far better to construct a predetermined dsDNA adapter and use this as a control insert. The authors use the adapter obtained by annealing the oligonucleotides pGIS7/4VERIF-TOP and pGIS7/4VERIF-BOT (see Fig. 21.12.2 and Support Protocol 2).*

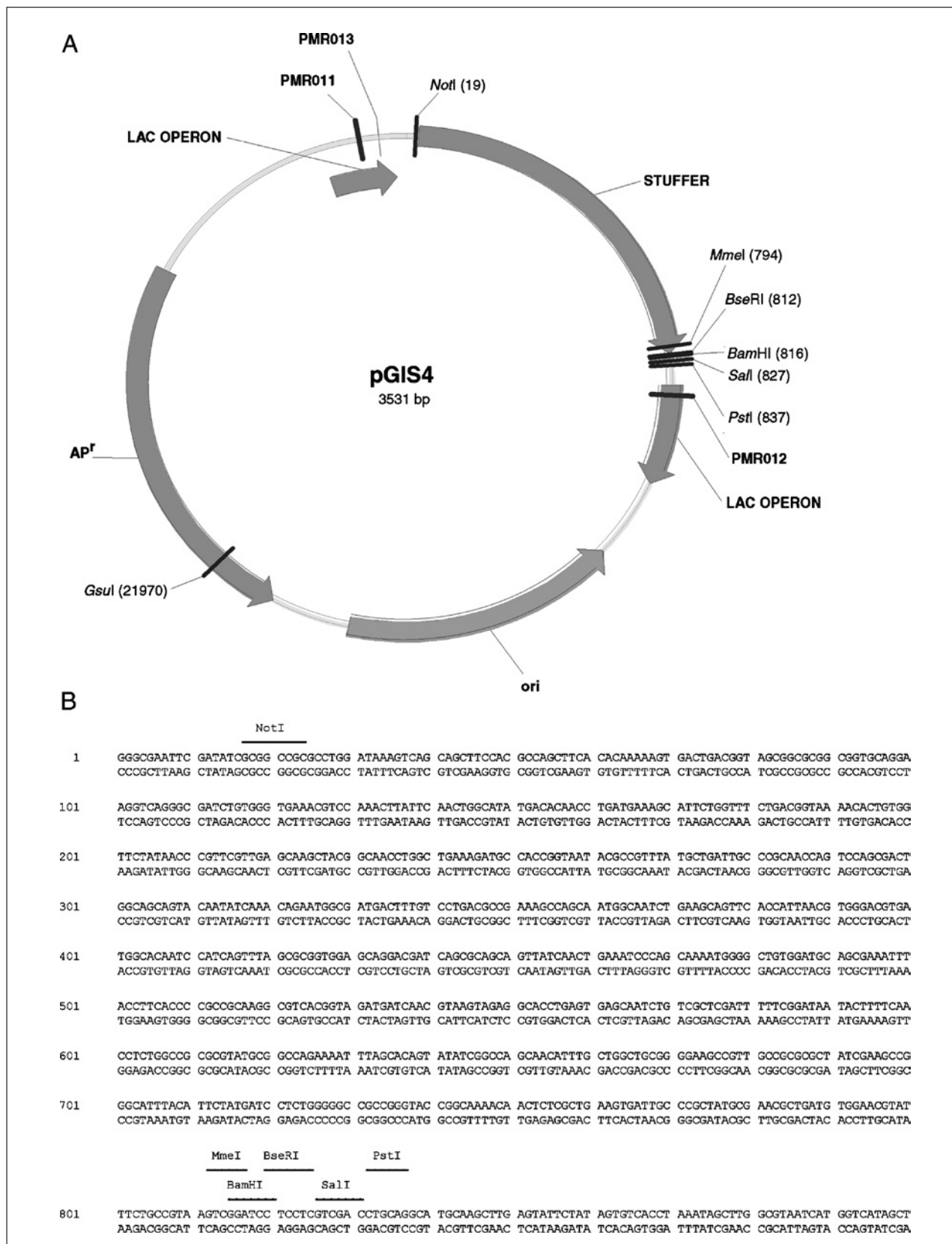


Figure 21.12.4 (A) Schematic diagram of the pGIS4 cloning vector for GIS-PET analysis. (B) Sequence details of the cloning site in pGIS4.

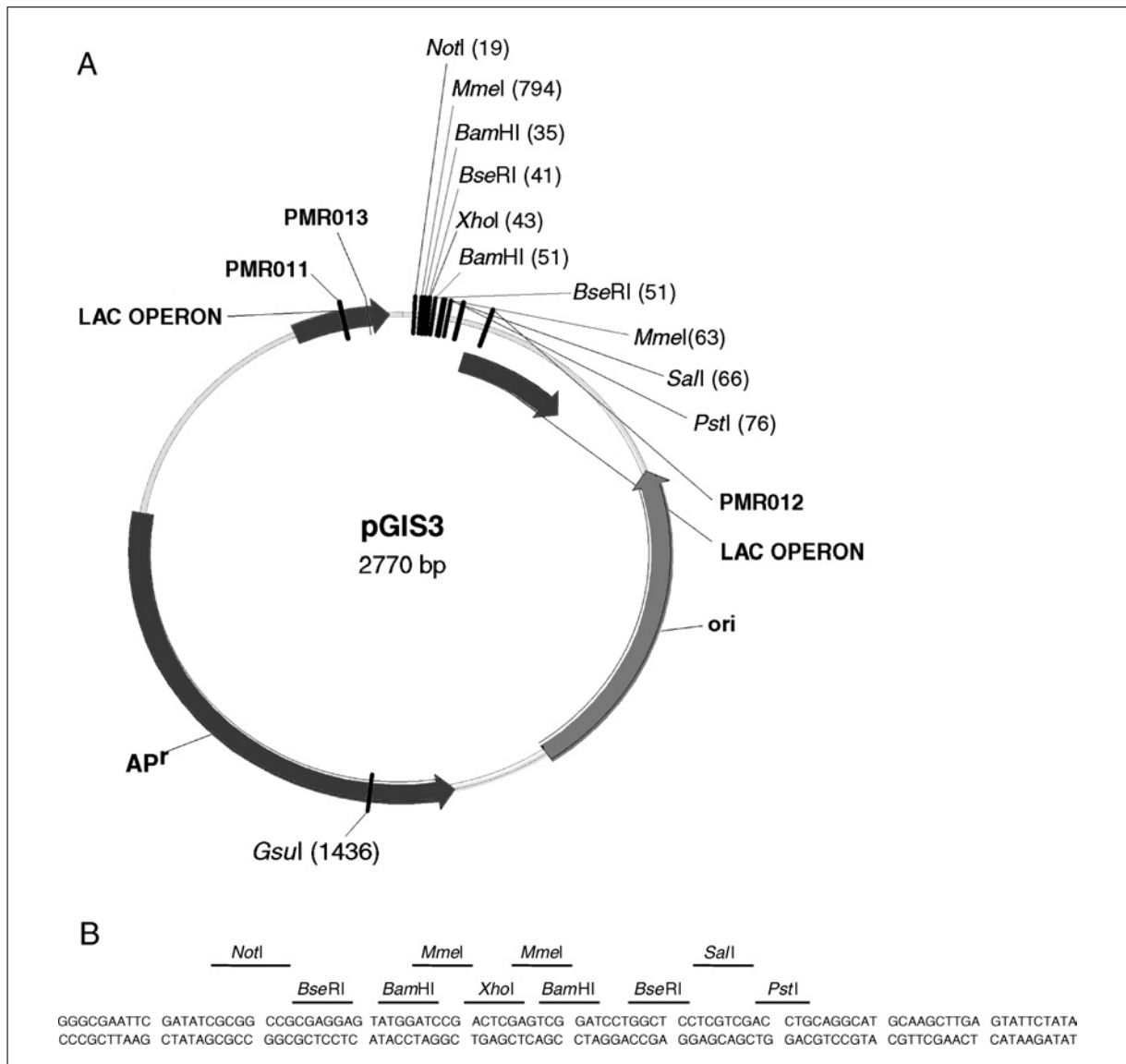


Figure 21.12.5 (A) Schematic diagram of the pGIS3 cloning vector for ChIP-PET analysis. (B) Sequence details of the cloning site in pGIS3.

Table 21.12.1 Test Ligations for Validation of Cloning Vector^a

Reagent	No ligase (uncut plasmid test)	Self-ligation (single-cut plasmid test)	Plus insert (positive control)
40 ng (<i>NotI</i> + <i>BseRI</i>) pGIS4	1	1	1
80 ng compatible stuffer	—	—	4
5× ligation buffer with PEG	2	2	2
T4 DNA ligase (5 U/μl)	—	—	1
Deionized water	7	6	2

^aVolumes (in μl) are for illustrative purposes only.

7a. Transform into bacteria (see Basic Protocol 1, steps 46 to 49) and count colonies.

There should be no or very few colonies for the uncut and single-cut ligations, and many for the positive control.

For pGIS3 plasmid DNA

1b. Purify pGIS3 plasmid DNA as in steps 1a to 3a.

2b. Perform *Xho*I digestion as follows to linearize the vector (and expose the cloning arms):

5 µg purified supercoiled pGIS3
10 µl 10× NEBuffer 2
1 µl 10 mg/ml BSA
2 µl 20 U/µl *Xho*I
Water to 100 µl.

Incubate 2 hr at 37°C.

3b. Purify by agarose gel electrophoresis.

4b. Polish off the exposed 3' overhangs using mung bean nuclease as follows:

~ 2µg *Xho*I-digested pGIS3
3 µl 10× reaction buffer
5 U mung bean nuclease per µg DNA
Deionized water to 30 µl.

Incubate 30 min at 30°C, then stop the reaction by adding 3.3 µl of 10% SDS (final 0.01%).

T4 DNA polymerase or Klenow enzyme cannot be used for this end-polishing step, as they will not remove the 5' overhangs to leave the MmeI sites directly adjacent to the inserted DNA fragment.

Mung bean nuclease is supplied at high concentration (70 U/µl) and must be diluted before use. It is typically diluted to 5 U/µl using 1× reaction buffer. It is critical to do the dilution just prior to use, as the dilute enzyme is unstable.

The mung bean nuclease polishing step can be difficult, hence the importance of the vector validation step below.

5b. Remove buffer salts and enzymes using Qiaquick PCR purification columns. Quantitate using a Nanodrop or other spectrometer.

6b. Remove 5'-phosphate groups to prevent self-circularization:

5 µg *Xho*I-cut, end-polished pGIS3
5 µl 10× Antarctic phosphatase buffer
5 µl 5 U/µl Antarctic phosphatase (use 5 U/µg target)
Deionized water to 50 µl.

Incubate exactly 30 min at 37°C; then heat inactivate 20 min at 65°C.

Volumes are only indicative; adjust as needed.

There is no need for a phenol extraction, as Antarctic phosphatase is completely inactivated by heating.

7b. If necessary, dilute prepared plasmid DNA in EB buffer (or TE buffer) to a final concentration of ≤40 ng/µl.

8b. If desired, run 2 µl (estimated to be 80 ng) of the diluted DNA on an agarose gel against known standards to confirm the quantity of DNA.

- 9b. Validate vector as in steps 6a and 7a, using 40 ng *Xho*I-digested vector. For the stuffer, use at least 80 ng of a blunt, phosphorylated stuffer fragment.

The prepared pGIS3 vector MUST be validated and sequenced to ensure that the MmeI sites were not damaged during the end-polishing step, as this would render the vector (and the entire library) useless.

PREPARATION OF ADAPTERS

Oligonucleotides are ordered pre-adjusted to 100 μ M concentration (Proligo). Since they are supplied in water, they must be adjusted to the correct concentration of Tris/NaCl/EDTA (TNE) prior to annealing.

Materials

Oligonucleotides A and B to be tested
2 \times TNE buffer (see recipe)
PAGE minigel (either 20% or 4% to 20% gradient)
*Not*I enzyme (NEB)
0.2-ml PCR tubes
Thermal cycler

1. In a 0.2-ml PCR tube, mix together:

20 μ l oligonucleotide A
20 μ l oligonucleotide B
40 μ l 2 \times TNE buffer.

2. Run in a thermal cycler using the following parameters:

hold 2 min at 95°C	
ramp from 95°C to 75°C at 0.1°C/sec	hold 2 min at 75°C
ramp from 75°C to 65°C at 0.1°C/sec	hold 2 min at 65°C
ramp from 65°C to 50°C at 0.1°C/sec	hold 2 min at 50°C
ramp from 50°C to 37°C at 0.1°C/sec	hold 2 min at 37°C
ramp from 37°C to 20°C at 0.1°C/sec	hold 2 min at 20°C
ramp from 20°C to 4°C at 0.1°C/sec	hold indefinitely at 4°C (until ready to collect).

IMPORTANT: *Keep all annealed adapters cold.*

3. Run ssDNA oligos and annealed adapters on the same PAGE minigel (either 20% or 4% to 20% gradient).
4. For GIS-(N)5 and GIS-(N)6 adapters, perform a simple *Not*I digestion, which should fragment the adapter in two, showing that the *Not*I site is present and correct.

PICOGREEN ASSAY FOR QUANTITATION OF DNA

Adapted from <http://www.molecularprobes.com/media/pis/mp07581.pdf>.

Materials

PicoGreen reagent (Invitrogen)
TE buffer (APPENDIX 2)
Calf thymus DNA (Sigma-Aldrich)
DNA samples

SUPPORT PROTOCOL 2

SUPPORT PROTOCOL 3

Chromatin Assembly and Analysis

21.12.27

Table 21.12.2 Standard Curve Preparation for PicoGreen Assay

Concentration after 50 μ l of 1 \times PicoGreen added (ng/ μ l)	1 ng/ μ l calf thymus DNA (μ l)	TE buffer (μ l)
0.1	10	40
0.05	5	45
0.025	2.5	47.5
0.01	1	49
0.005	0.5	49.5
0.0025	0.25	49.75

96-well, clear flat-bottom plate (Greiner)
Genios (Tecan) or other suitable fluorometer
Fluorometer software (e.g., Magellan)

1. Dilute 200 \times concentrated PicoGreen reagent to 1 \times in TE buffer just before the experiment. Prepare 50 μ l of 1 \times PicoGreen for each sample to be analyzed.
2. Prepare a 100 ng/ μ l stock solution of calf thymus DNA to make the standard curve. Determine the standard curve range according to the estimated concentration of DNA in the sample (see Table 21.12.2).
3. Dilute the DNA samples so that the diluted DNA concentration falls in the range of the standard curve. Prepare 50 μ l diluted DNA sample for each assay.
4. Add 50 μ l of each DNA sample to a 96-well, clear flat-bottom plate. Add 50 μ l TE buffer to a separate well as blank control.
5. Add 50 μ l of 1 \times PicoGreen into blanks and samples, protected from light. Measure the fluorescence of samples using a Tecan Genios or other suitable fluorometer.
6. Subtract the fluorescence of the reagent blank from that of each sample and determine the DNA concentration of samples from the standard curve using the fluorometer software (e.g., Magellan).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Alkaline hydrolysis buffer

50 mM NaOH
5 mM EDTA, pH 8.0
Prepare fresh before use

Binding buffer, 1 \times

2 M NaCl
50 mM EDTA, pH 8.0
Stable for years at 4 $^{\circ}$ C

Blocking buffer, 1 \times

0.4% SDS
50 μ g/ml yeast tRNA
Stable for years at -20° C

Ligase buffer with spermidine, 10×

60 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
60 mM MgCl₂
50 mM NaCl
1 mg/ml BSA
70 mM 2-mercaptoethanol
1 mM ATP
20 mM DTT
10 mM spermidine (Sigma)
Stable for years at −20°C

LoTE buffer

3 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
0.2 mM EDTA (*APPENDIX 2*)
Stable for years at room temperature

Modified 10 mM dNTP with 5-Me-dCTP

100 μl 100 mM stock solution of dATP (final 10 mM)
100 μl 100 mM stock solution of dTTP (final 10 mM)
100 μl 100 mM stock solution of dGTP (final 10 mM)
50 μl 100 mM stock solution of 5-Me-dCTP (final 5 mM)
10 μl 1 M Tris·Cl, pH 8.0 (final 10 mM)
640 μl deionized water
Stable for years at −20°C

All dNTPs are available from Roche.

Saturated trehalose (RNase-free)

Place a magnetic stir bar into a small beaker and heat 10 ml water to 42°C. Slowly add 8 g D-(+)-trehalose dihydrate powder (Sigma) and allow it to dissolve by stirring. While carefully maintaining the temperature at 42°C, continue adding trehalose until saturation is reached. Cool solution to room temperature, allowing trehalose crystals to form. At this stage, add enough diethylpyrocarbonate (DEPC) to give a 0.1% final concentration. Shake vigorously for 5 to 10 min (or stir overnight), then autoclave the saturated trehalose solution to deactivate the DEPC. Store saturated trehalose in aliquots at −20°C (stable for years).

Because the solubility of trehalose increases with temperature, it is important to maintain the temperature at not more than 42°C when making this solution. This is the temperature at which the trehalose is used in the protocol.

SOC medium

0.5% yeast extract
2% tryptone
10 mM NaCl
2.5 mM KCl
10 mM MgCl₂
10 mM MgSO₄
20 mM glucose
Store at room temperature (stable for years)

TEN buffer, 1 ×

10 mM Tris·Cl, pH 8.0 (APPENDIX 2)
0.1 mM EDTA, pH 8.0 (APPENDIX 2)
25 mM NaCl
Stable for years at 4°C

TNE buffer, 2 ×

20 mM Tris·Cl, pH 8.0 (APPENDIX 2)
0.2 mM EDTA, pH 8.0 (APPENDIX 2)
100 mM NaCl
Stable for years at 4°C

Wash buffer, 1 ×

10 mM Tris·Cl, pH 7.5 (APPENDIX 2)
0.2 mM EDTA, pH 8.0 (APPENDIX 2)
10 mM NaCl
20% glycerol
40 µg/ml yeast tRNA

If buffer is made up with yeast tRNA, then store at −20°C (stable for years). Alternatively, it is convenient to store the buffer (minus yeast tRNA) for years at 4°C, remembering to add the yeast tRNA before use (to a final concentration of 40 µg/ml).

COMMENTARY

Background Information

GIS-PET analysis

Sequence-based methods of extracting information from completed genome sequences include full-length cDNA sequencing, DNA tagging methods such as serial analysis of gene expression (SAGE; UNIT 25B.6; Velculescu et al., 1995; Saha et al., 2002), massively parallel signature sequencing (MPSS; Brenner et al., 2000), expressed sequence tag (EST) analysis (Adams et al., 1991), and more recently, terminal tag isolation procedures such as cap analysis gene expression (CAGE; Shiraki et al., 2003), 5'/3' LongSAGE (Wei et al., 2004), 5-SAGE (Hashimoto et al., 2004), and the gene identification signature-paired end ditagging (GIS-PET) analysis method (Ng et al., 2005), which is described here. These methods require extensive sample preparation and can be technically challenging, but may be implemented in any standard molecular biology laboratory, and are advantageous because they are “open” systems querying entire genomes. They allow quantitative gene expression comparisons by simple tag-counting, and are “portable,” i.e., the results are not affected by different instruments or experimenters.

The original DNA tagging methods of SAGE and MPSS, while providing an accurate and highly-efficient means to analyze tran-

scripts, were hindered by the inherent positional ambiguity of a short single tag, and the possibility of incomplete enzymatic digestion during tagging. This was partially resolved by methods that separately extracted only either 5' or 3'-terminally located tags, culminating in this present strategy of linking both terminal tags from a contiguous DNA fragment (in this case, flcDNA) into paired-end ditags (PETs) that can be mapped back onto assembled genomes to precisely locate the original sequence. The PET-based transcriptome analysis procedure called gene identification signature (GIS) analysis is as accurate as flcDNA sequencing, but potentially 30-fold more efficient (estimation based on Sanger sequencing results) for genome annotation (Ng et al., 2005). GIS-PET can also be used to characterize splice variant expression in the transcriptome; however, it is only able to differentiate variants that differ in either the first or last exons. The high-quality flcDNA library that is constructed as an intrinsic part of the procedure is an added resource, facilitating the recovery of transcripts of interest by simple PCR for downstream functional studies.

In GIS-PET, the quality of the flcDNA library is of paramount importance. During development of the GIS-PET procedure, several different approaches were tested

for generating *flcDNA*, and the CapTrapper approach (Carninci and Hayashizaki, 1999), which utilizes chemical biotinylation, produced the highest percentage of full-length clones in the library (typically >92%). The drawbacks of the CapTrapper procedure are that it is tedious, requiring multiple hands-on steps, and that it requires a relatively large quantity of starting mRNA (~10 µg). This latter requirement currently precludes the use of GIS-PET with clinical samples, which are usually present in very small quantities, although strategies to overcome this problem are currently being developed.

ChIP-PET analysis

Chromatin immunoprecipitation (ChIP; UNIT 21.3) is a powerful technique for the global identification of transcription factor binding sites *in vivo*. To map the TFBS in a high-throughput manner, the technique of ChIP-on-chip is widely used: this involves the hybridization of ChIP-enriched genomic fragments onto a series of genomic microarrays. While this has succeeded in simpler organisms such as yeast (Ren et al., 2000), the greater size and complexity of mammalian genomes has rendered a complete global survey of TFBS by ChIP-on-chip unfeasible at present, although some important progress has been made on a limited scale, such as with specific chromosomal arrays, or promoter microarrays containing CpG islands or sequences flanking transcription start sites (Horak et al., 2002; Weinmann et al., 2002; Cawley et al., 2004; Boyer et al., 2005). In another recent report, high density tiling oligo arrays covering 25% of the sequenced human genome were used to map active promoters (Kim et al., 2005).

Alternatively, immunoprecipitated DNA fragments from ChIP experiments can be directly cloned and sequenced (Weinmann et al., 2001; Hug et al., 2004). The main drawback of this is that ChIP typically presents a high background of non-specific DNA, complicating the task of identifying authentic TFBS, especially if every cloned fragment has to be sequenced. The increase in sampling efficiency offered by DNA tagging methods such as SAGE presents a possible solution, and has been applied to genome scanning for TFBS and histone modification (Impey et al., 2004; Chen and Sadowski, 2005; Kim et al., 2005; Roh et al., 2005). However, the inherent ambiguity of mapping short single tags limits the ability of such ChIP-SAGE approaches to accurately pinpoint the boundaries of each frag-

ment, thereby making it difficult to distinguish true ChIP enrichment from amplified noise generated during molecular cloning.

ChIP-PET makes use of the increased accuracy offered by mapping dual-terminal tags to the genome, while at the same time maintaining the increased efficiency of DNA tagging methods. The approach has been successfully validated in studies that identified, on a global scale, p53 TFBS within human colorectal cancer cells (Wei et al., 2006), and Oct4/Nanog binding sites in mouse embryonic stem cells (Loh et al., 2006).

Critical Parameters and Troubleshooting

***flcDNA* library**

Cloning vector quality. For library construction purposes, the quality of the cloning vector is critical, in that it should produce only a low background of no-insert clones, and a library with a high titer. For this reason, always ensure that the vector preparation is prevalidated before use, following the procedure described in Support Protocol 1. It is very important to freeze the validated vector preparation in small aliquots, as repeated freezing and thawing of the vector DNA results in degradation and, consequently, undesirable results including decreased library titers and the prevalence of clones in which the polyA tails were not removed.

Generation of *flcDNA*. Do not use glycerol at any stage where it is not specifically mentioned, as this will interfere with the CapTrapper selection process. For all steps involving single-stranded DNA or RNA, it is preferable to use a siliconized or low-binding 1.5- or 1.7-ml microcentrifuge tube to avoid loss of nucleic acid. Phase Lock Gel tubes (Eppendorf) are very convenient for the numerous phenol/chloroform/IAA extraction steps, and help to reduce loss of aqueous samples. It is important that Phase Lock Gel *not* be used after reverse-transcription, as the presence of trehalose and sorbitol increases the solution density to the point that both phases form *below* the gel layer, making recovery difficult.

Vector-insert ligations. The quantity of size-fractionated cDNA used for ligation to the pGIS vector is important; at least 100 ng per 30 to 40 ng vector is best. Additionally, one should avoid the temptation of using the smallest fractions—these contain excess adapter DNA, which can block the vector cloning sites. Finally, ensure that separate ligations are set up for different size fractions of cDNA, and avoid

pooling all size fractions together for the ligation. This ensures a more complete collection of cDNA of different sizes.

Quality of flcDNA library. Before proceeding to the ditagging procedure, it is important to assay the quality of the flcDNA library. The quality of the library can be inferred from a combination of different quality control (QC) checks, including counting total titer, determining the range of insert sizes by colony PCR, and most importantly, by sequencing a selection of clones. Check the DNA sequences for the percentage of full-length transcripts, as well as for the presence of all expected flanking restriction enzyme recognition sites. To obtain an estimated ten-fold coverage of a human transcriptome, aim for a flcDNA library with a total titer of 1×10^6 cfu.

Large-scale library plating. For the use of Q-trays, it is important to dry the plates (containing LB agar) sufficiently before use, otherwise the condensate produced during subsequent incubation will result in smearing/merging of growing colonies. Leave the plates open in a sterile hood until a pattern of “waves” appear on the agar surface. During incubation, do not stack Q-trays more than three high, to avoid excessive condensation.

Plasmid DNA maxiprep from bacteria scraped from Q-trays. Extraction of plasmid DNA from these “bacterial scrapings” is less efficient than starting with the usual liquid cultures. Expect a yield of ~ 1 mg plasmid DNA starting from 1×10^6 scraped colonies (about ten Q-trays); if yields are < 600 μg of plasmid DNA, check that enough solution volume was used at each stage, e.g., if using the Qiagen maxiprep system, ensure that 20 to 50 ml of each buffer P1, P2, and P3 is used. There is a very large quantity of precipitated flocculant debris after the P3 neutralization step, which can lead to poor recovery of the supernatant if one uses Qiafilters directly, so it would be better to preclear the precipitate by centrifugation (in 50-ml tubes, e.g., Oak Ridge, Nalgene, 30 min at 12,000 rpm in an SS34 rotor), followed by the use of Qiafilters, if necessary for a final removal of particulates before introducing the supernatant into the maxiprep tips.

GIS-PET ditagging

PET concatenation. The duration of the ligation step may need to be optimized, depending on the quality of the purified PETs obtained. Aim for a majority of concatenated DNA ~ 1000 to 2000 bp. Ensure that spermidine is added to the ligase buffer, as this

greatly increases the efficiency of the desired linear concatenation.

Cloning of BamHI-derived PET concatemers. The partial redigestion of the concatenated DNA by *Bam*HI prior to gel fractionation appears to be critical to the success of the subsequent cloning step. Follow the recommended conditions carefully, to avoid either under- or over-digestion. The authors have successfully used *Bam*HI at a specific activity of 1 to 3 U/ μg substrate DNA.

Recovery of gel-purified DNA. There is not much difference in yields between the gel-crush method of recovering DNA or electroelution using ElutaTubes. Electroelution, however, can result in significant time savings.

ChIP-PET

The general precautions listed for the GIS-PET analysis method are also applicable to ChIP-PET. In addition, it is critical that the starting material, i.e., ChIP-enriched gDNA fragments, are of a sufficient quality to generate a high-quality ChIP-DNA library. Before starting, ensure that the fragments were completely de-crosslinked, are of sufficient quantity (≥ 100 ng), and are of an appropriate size-range (centered ~ 500 bp). Poor quality starting DNA manifests itself in a ChIP-DNA library with low titer and/or excessively short inserts.

Anticipated Results

GIS-PET

Starting with ~ 10 μg of high-quality polyA RNA, one should be able to obtain a flcDNA library with a total titer of 5×10^5 to $> 2 \times 10^6$ cfu. A pool of $\sim 1 \times 10^6$ colonies scraped from ten or more Q-trays should generate ~ 1 mg of plasmid DNA. When ditagging, 10 μg of plasmid DNA from the flcDNA library should produce ~ 1 μg of purified, linearized, single-PET plasmid DNA that will be self-circularized to generate the single-PET plasmid library. This single-PET plasmid library has a typical total titer of $> 5 \times 10^6$ cfu. The final GIS-PET library in pZErO-1, constructed by cloning various size fractions of ~ 1 μg of concatenated 50-bp PETs, should have a total titer of $\sim 1 \times 10^6$ cfu, with seven to twelve PETs present per clone.

ChIP-PET

Starting with 200 ng of ChIP-enriched DNA, one should be able to generate a ChIP-DNA library with a total titer of $\sim 4 \times 10^6$ to $> 1 \times 10^7$ cfu. The other downstream results

would be similar to what is expected for GIS-PET.

Time Considerations

For GIS-PET, it will take 6 days to construct the flcDNA library starting from purified polyA RNA, and an additional 9 days to prepare the GIS-PET library for sequencing. Factor in another 4 days for preparation and validation of cloning vectors. For ChIP-PET starting from ChIP-enriched DNA fragments, it should take 3 days to construct the ChIP-DNA library, and an additional 9 days to prepare the GIS-PET library for sequencing.

Literature Cited

- Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merrill, C.R., Wu, A., Olde, B., Moreno, R.F., et al. 1991. Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science* 252:1651-1656.
- Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., Gifford, D.K., Melton, D.A., Jaenisch, R., and Young, R.A. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122:947-956.
- Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D.H., Johnson, D., Luo, S., McCurdy, S., Foy, M., Ewan, M., Roth, R., George, D., Eletr, S., Albrecht, G., Vermaas, E., Williams, S.R., Moon, K., Burcham, T., Pallas, M., DuBridge, R.B., Kirchner, J., Fearon, K., Mao, J., and Corcoran, K. 2000. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* 18:630-634.
- Carninci, P. and Hayashizaki, Y. 1999. High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303:19-44.
- Cawley, S., Bekiranov, S., Ng, H.H., Kapranov, P., Sekinger, E.A., Kampa, D., Piccolboni, A., Sementchenko, V., Cheng, J., Williams, A.J., Wheeler, R., Wong, B., Drenkow, J., Yamanaka, M., Patel, S., Brubaker, S., Tammana, H., Helt, G., Struhl, K., and Gingeras, T.R. 2004. Unbiased mapping of transcription factor binding sites along human chromosome 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* 116:499-509.
- Chen, J. and Sadowski, I. 2005. Identification of the mismatch repair genes PMS2 and MLH1 as p53 target genes by using serial analysis of binding elements. *Proc. Natl. Acad. Sci. U.S.A.* 102:4813-4818.
- Hashimoto, S., Suzuki, Y., Kasai, Y., Morohoshi, K., Yamada, T., Sese, J., Morishita, S., Sugano, S., and Matsushima, K. 2004. 5'-end SAGE for the analysis of transcriptional start sites. *Nat. Biotechnol.* 22:1146-1149.
- Horak, C.E., Mahajan, M.C., Luscombe, N.M., Gerstein, M., Weissman, S.M., and Synder, M. 2002. GATA-1 binding sites mapped in the beta-globin locus by using mammalian ChIP-chip analysis. *Proc. Natl. Acad. Sci. U.S.A.* 99:2924-2929.
- Hug, B.A., Ahmed, N., Robbins, J.A., and Lazar, M.A. 2004. A chromatin immunoprecipitation screen reveals protein kinase c-beta as a direct RUNX1 target gene. *J. Biol. Chem.* 279:825-830.
- Impey, S., McCorkle, S.R., Cha-Molstad, H., Dwyer, J.M., Yochum, G.S., Boss, J.M., McWeeney, S., Dunn, J.J., Mandel, G., and Goodman, R. 2004. Defining the CREB regulon: A genome-wide analysis of transcription factor regulatory regions. *Cell* 119:1041-1054.
- Kim, J., Bhinge, A.A., Morgan, X.C., and Iyer, V.R. 2005. Mapping DNA-protein interactions in large genomes by sequence tag analysis of genomic enrichment. *Nat. Methods* 2:47-53.
- Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., Wong, K.Y., Sung, W.K., Lipovich, L., Kuznetsov, V.A., Robson, P., Stanton, L.W., Wei, C.L., Ruan, Y., Lim, B., and Ng, H.H. 2006. The Oct4 and Nanog transcription network that regulates pluripotency in mouse embryonic stem cells. *Nat. Genetics* 38:431-440.
- Ng, P., Wei, C.L., Sung, W.K., Chiu, K.P., Lipovich, L., Ang, C.C., Shahab, A., Ridwan, A., Wong, C.H., Liu, E.T., and Ruan, Y. 2005. Gene identification signature (GIS) analysis for transcriptome characterization and genome annotation. *Nat. Methods* 2:105-111.
- Ren, B., Robert, F., Wyrick, J.J., Aparicio, O., Jennings, E.G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E., Volkert, T.L., Wilson, C.J., Bell, S.P., and Young, R.A. 2000. Genome wide location and function of DNA binding proteins. *Science* 290:2306-2309.
- Roh, T.Y., Cuddapah, S., and Zhao, K. 2005. Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes Dev.* 19:542-552.
- Saha, S., Sparks, A.B., Rago, C., Akmaev, V., Wang, C.J., Vogelstein, B., Kinzler, K.W., and Velculescu, V.E. 2002. Using the transcriptome to annotate the genome. *Nat. Biotechnol.* 20:508-512.
- Shiraki, T., Kondo, S., Katayama, S., Waki, K., Kasukawa, T., Kawaji, H., Kodzius, R., Watahiki, A., Nakamura, M., Arakawa, T., Fukuda, S., Sasaki, D., Podhajska, A., Harbers, M., Kawai, J., Carninci, P., and Hayashizaki, Y. 2003. Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. *Proc. Natl. Acad. Sci. U.S.A.* 100:15776-15781.
- Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. 1995. Serial analysis of gene expression. *Science* 270:484-487.
- Wei, C.L., Ng, P., Chiu, K.P., Wong, C.H., Ang, C.C., Lipovich, L., Liu, E.T., and Ruan, Y. 2004. 5' Long serial analysis of gene expression (LongSAGE) and 3' LongSAGE for

transcriptome characterization and genome annotation. *Proc. Natl. Acad. Sci. U.S.A.* 101:11701-11706.

Wei, C.L., Wu, Q., Vega, V.B., Chiu, K.P., Ng, P., Zhang, T., Shahab, A., Yong, H.C., Fu, Y., Weng, Z., Liu, J., Zhao, X.D., Chew, J.L., Lee, Y.L., Kuznetsov, V.A., Sung, W.K., Miller, L.D., Lim, B., Liu, E.T., Yu, Q., Ng, H.H., and Ruan, Y. 2006. A global map of p53 transcription factor binding sites in the human genome. *Cell* 124:207-219.

Weinmann, A.S., Bartley, S.M., Zhang, T., Zhang, M.Q., and Farnham, P.J. 2001. Use of chromatin immunoprecipitation to clone novel E2F target promoters. *Mol. Cell. Biol.* 21:6820-6832.

Weinmann, A.S., Pearly, S.Y., Oberley, M.J., Huang, T.H.-M., and Farnham, P.J. 2002. Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev.* 16:235-244.

Contributed by Patrick Ng, Chia-Lin Wei,
and Yijun Ruan

Cloning and Sequencing Group
Genome Institute of Singapore
Singapore

CHAPTER 22

Nucleic Acid Arrays

INTRODUCTION

Nucleic acid arrays present a near-textbook case of a technology in which an increase in throughput results in a qualitative change in the usefulness of the technique. Arrays of nucleic acids on nitrocellulose have been around for many years (consider dot blots and slot blots, used for gene-expression monitoring since the late 1970s). But in recent years, nucleic acid arrays have followed a now typical “genomic” pattern, in which, above some point where the representation of genes on the array begins to comprise a substantial fraction of the genes of a genome or of a set of important genes, the experimental value gained from the use of that array increases by more than the fractional increase in the number of genes on the array.

For insight into a variety of current and potential future applications of microarray technology in molecular biology, the reader is referred to *UNIT 22.1*. This unit also describes how and why methods that use these arrays have gone from costly—and requiring special technical sophistication—to mainstream. The Basic Protocol in the first procedural unit in this chapter details the preparation and biotin labeling of mRNA for expression-profiling analysis (*UNIT 22.2*). *UNIT 22.3* describes fluorescent cDNA microarray technology. This involves the production and spotting of cDNAs on glass slides, the extraction of mRNA, the synthesis of fluor-derivatized cDNAs by reverse transcription, to be hybridized to the microarray, the hybridization itself, and subsequent data extraction. *UNIT 22.4* provides a comparison of the various technologies that can be used for transcriptional profiling analysis including Affymetrix, NimbleGen, and spotted arrays, along with an introductory discussion about a variety of statistical considerations involved in analyzing microarray data. The discussion of the analysis of microarray data begins in *UNIT 22.4* and is continued in *UNIT 22.5*, which introduces the general concepts and mathematical procedures used to extract useful information from a comparison of multiple expression profiles, including a discussion of a variety of pattern discovery tools.

Both methods and ideology (i.e., “heuristics”) for use of DNA micorarrays are developing rapidly, and the editors will be particularly vigilant in adding new methods as those prove useful.

Roger Brent and Fred Ausubel

Nucleic acid array technology refers to the fabrication and use of arrays containing thousands of nucleic acid samples bound to solid substrates, such as glass microscope slides or silicon wafers. Because the physical area occupied by each sample is usually 50 to 200 μm in diameter, nucleic acid samples representing entire genomes, ranging in size from 3,000 to 32,000 genes, may be efficiently packaged onto a single regular microscope slide in an area easily covered by a coverslip (Fig. 22.1.1). Such “genomes on a chip” then serve as a target to which fluorescently labeled nucleic acid probes can be applied. Nucleic acid arrays, or microarrays, allow all genes of a given genome to be simultaneously monitored with respect to some experimental condition of interest. This fact has fundamentally changed the manner in which the study of genomics and gene expression can be pursued.

The majority of applications discussed in this overview relate to DNA microarrays fabricated by the mechanical deposition of nucleic acid samples onto glass. Typically, these samples are in the form of PCR products, ranging in size from 100 bps to 9 kb. However, the term “DNA microarray” may apply to several different forms of the technology, each differing in the type of nucleic acid applied and the method of application. For example, Affymetrix sells DNA arrays produced by photolithographic synthesis of individual short oligonucleotides directly on the substrate. (Fodor et al., 1991).

WHAT ARE MICROARRAYS GOOD FOR?

Gene Expression Analysis

Undoubtedly the most common use for DNA microarrays is for monitoring gene expression levels. The broad appeal of this approach stems from the fact that it can be applied to virtually any organism, tissue, or cell line from which RNA may be isolated. In a typical experiment, total RNA or mRNA is collected from two or more individuals, cultures, or conditions. The amount of RNA needed for a microarray experiment depends on many factors, such as genome complexity and message content. Most experiments use anywhere from 100 ng to as much as 20 μg of RNA. The next step is the separate conversion of the RNA samples into cDNA by reverse transcription. This is

usually accomplished by priming with randomized oligonucleotides or, in the case of organisms that produce polyadenylated messages, an oligo-dT primer. The basic principle behind these manipulations is to convert the RNA from each sample into a form that can be readily distinguished from another RNA sample. This is usually accomplished by labeling the cDNA samples with different fluorescent dyes, either during the reverse transcription process through direct incorporation by reverse transcriptase, or afterwards by chemical conjugation.

The resulting pools of cDNA are mixed together, and when the pool is hybridized to a microarray, the ratio between the intensities observed for two of the fluorophores at any given location in the array is a direct measure of the relative abundance of the corresponding cDNA transcript (see Fig. 22.1.2). By using a single reference sample as the control for a series of experimental samples collected over time, one can compare relative levels of transcript abundance among samples (see Fig. 22.1.3). This in turn allows one to identify gene expression trends. For those who wish to study global regulation of gene expression, the most significant data lies in these trends and patterns. Many instances where this methodology has been successfully put into practice can be found in the literature (DeRisi et al., 1997, Alizadeh et al., 1998, Cho et al., 1998; Chu et al., 1998, Eisen et al., 1998, Spellman et al., 1998; Amundson et al., 1999; Iyer et al., 1999; Perou et al., 1999).

It is certain that the use of microarrays to analyze gene expression will continue to increase. Aside from looking at developmental time courses, mutations, and other genetic modifications, many other novel expression experiments will undoubtedly be developed. One recent variation includes the use of expression analysis to reverse-engineer the changes that occurred during a 200-generation yeast evolution experiment (Ferea et al., 1999). In addition to well-established model systems, gene expression analysis involving microarrays may be used as a method to attack problems that were formerly too cumbersome to approach using standard molecular biology techniques. This is especially true for potentially dangerous and difficult-to-culture organisms such as *Mycobacterium tuberculosis* and

Contributed by Joseph DeRisi

Current Protocols in Molecular Biology (2000) 22.1.1-22.1.7

Copyright © 2000 by John Wiley & Sons, Inc.

Nucleic Acid
Arrays

22.1.1

Supplement 49

Plasmodium falciparum (Hayward et al., 1999, Wilson et al., 1999). The primary reason for this is that only relatively small amounts of RNA are required to determine expression profiles for entire genomes. In other words, the data return on the up-front labor investment is much more substantial than with older techniques for studying gene expression, thus enlarging the potential scope and depth of experiments. For example, consider the case of *Plasmodium*: genetic crosses are extremely difficult and transformation efficiencies are abysmal. These facts make many of the tools researchers have relied upon for cloning new genes nearly useless. Therefore, determining which genes are induced by various drugs becomes a large un-

dertaking, especially considering that the *Plasmodium* genome has yet to be sequenced completely. However, it is a relatively simple matter to collect RNA samples, so DNA microarrays (utilizing known or random genomic fragments) provide an easy means to answer these questions.

WHAT ELSE ARE NUCLEIC ACID MICROARRAYS GOOD FOR?

Despite all the recent emphasis on gene expression analysis, the number of different uses for DNA microarrays is currently limited by our imagination. The most fundamental feature of microarray analysis is that meaningful data can be derived from any set of biological

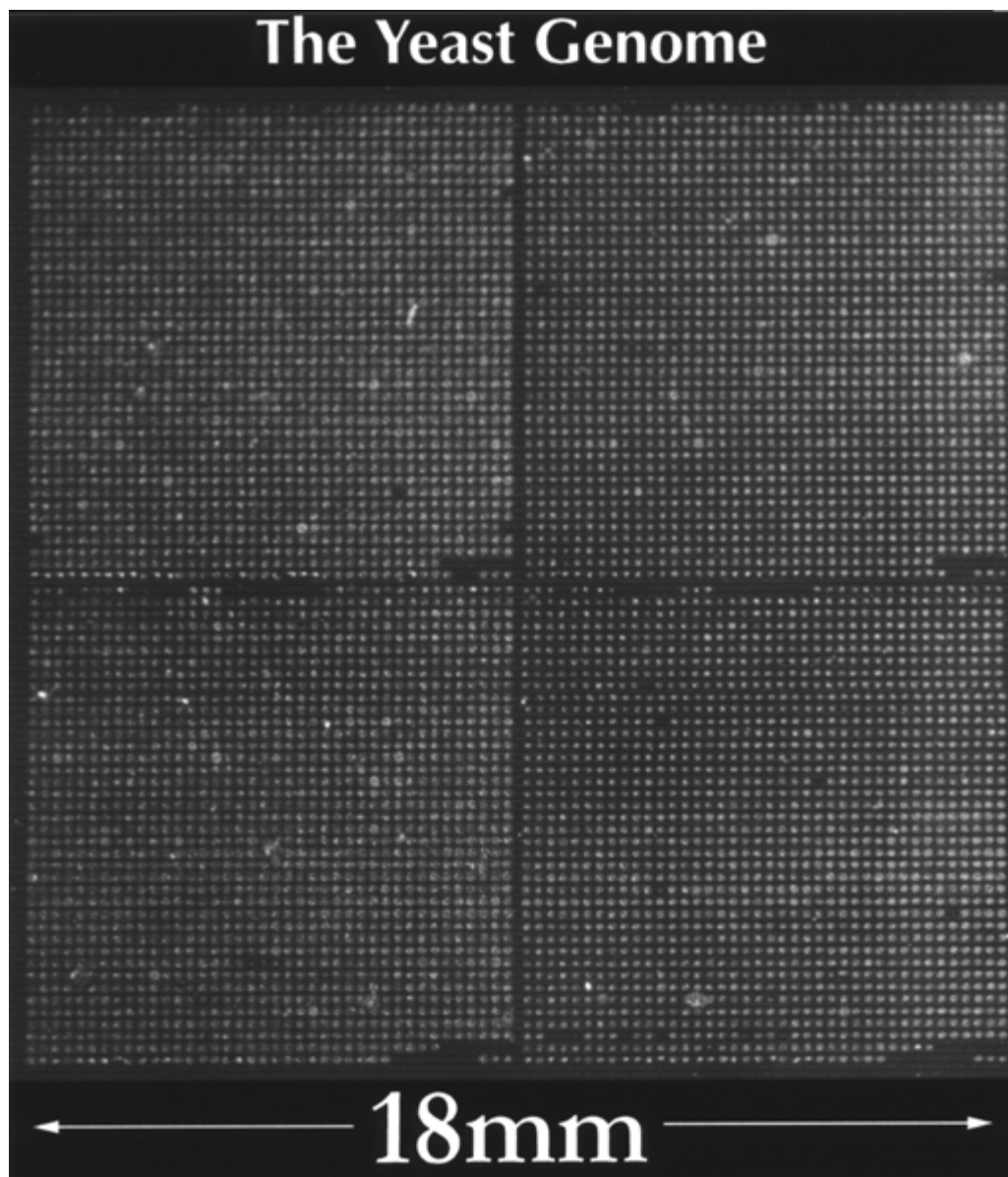


Figure 22.1.1 Yeast “genome on a chip.”

experiments that result in a differential recovery of nucleic acid material. Aside from gene expression, these uses fall into several basic categories: genetic mapping and genotyping, assessment of genome structure and copy number, polysome analysis, and assays of DNA-protein interaction. Given that there are dozens of variations in experimental design and approach within each of these categories, the following is meant only to pique the imagination and alert readers to the diverse opportunities afforded by microarray technology.

Genotyping and Genetic Mapping

Efficient genotyping of hundreds or thousands of markers for the purpose of mapping multigenic traits may be carried out using DNA microarrays. In one approach, genotyping by array analysis is accomplished by directly detecting hybridization differences caused by single-nucleotide polymorphisms (SNPs). Arrays utilizing short oligo features are well suited for the former method due to their low annealing

temperatures (Hacia et al., 1996). Another approach, genomic mismatch scanning (GMS), uses mismatch repair enzymes to recognize and selectively degrade hybrid DNA fragments between two individuals that possess SNPs (Nelson et al., 1993). The hybrid DNA is created by denaturing and reannealing fragments from two related individuals. The resulting perfectly matched DNA fragments can then be fluorescently labeled and applied to a microarray, revealing which segments are identical by descent (Cheung et al., 1998; McAllister et al., 1998). Regardless of the method used, genotyping by array analysis has the potential to increase the throughput by several orders of magnitude when compared to traditional gel electrophoresis-based methods.

Comparative Genomic Hybridization

Differences in gene copy number have been associated with various tumorigenic phenotypes. These polymorphisms, which consist of deletions and/or amplifications, can easily be

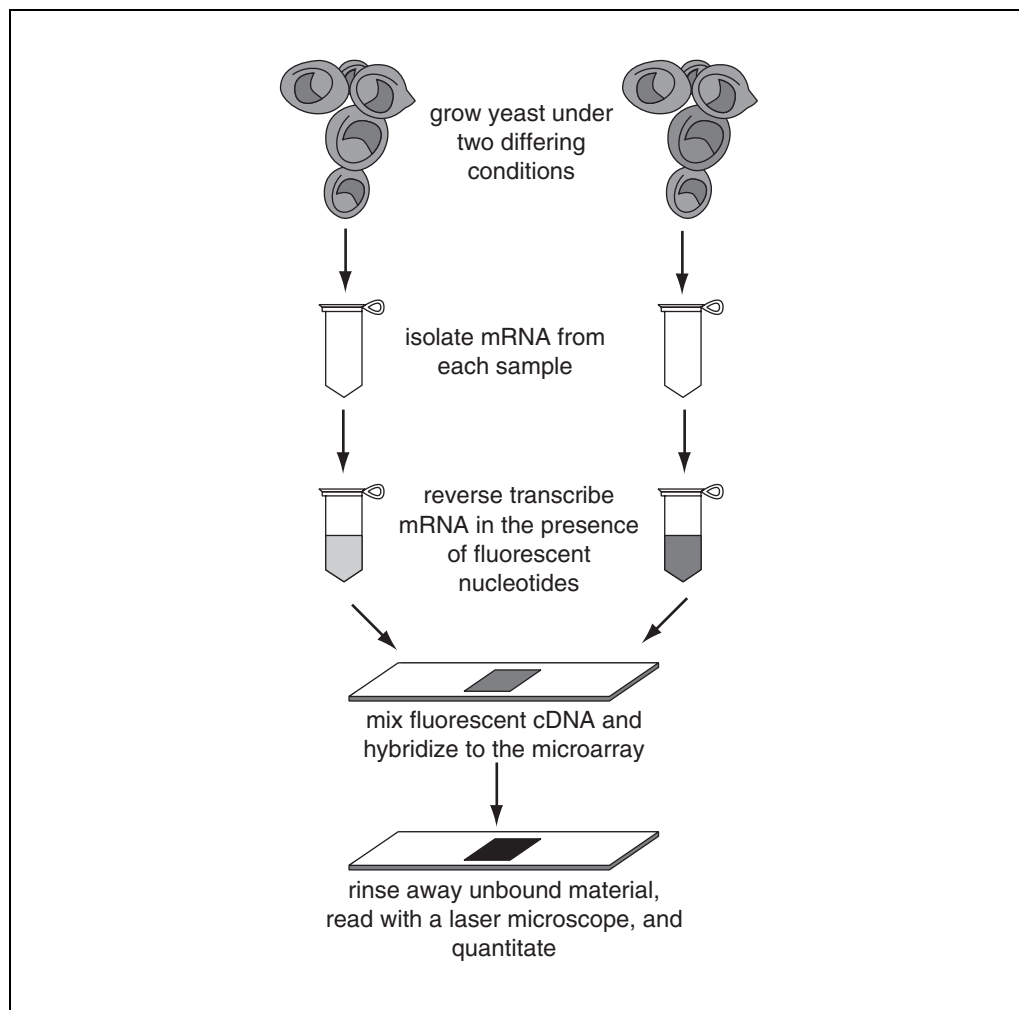


Figure 22.1.2 Scheme for a typical gene expression experiment.

assayed using DNA microarrays, which provide a viable alternative to traditional comparative genomic hybridization (CGH) techniques. The primary limitation of traditional cytogenetic CGH techniques lies in the fact that the resolution to which an amplification or deletion may be mapped is ~20 Mb (Kallioniemi et al., 1992). With array-based CGH methods, the resolution depends only on how many DNA elements are present on the microarray. Therefore, it is entirely feasible that all human genes may be represented as discrete elements on future microarrays, taking the resolution of the map to its logical limit. Indeed, with the conclusion of the human genome sequencing project, it should be feasible to detect the deletion or amplification of virtually any and all genes simultaneously for a given human cell population. Several examples already exist in the literature, and advances in the use of the technology are sure to follow (Solinas-Toldo et al., 1997; Trent et al., 1997; Pinkel et al., 1998; Behr et al., 1999; Pollack et al., 1999).

Polysome Analysis

Although transcriptional regulation is the primary focus of most microarray experiments, changes in translation can also be readily assayed. This may be accomplished by fluorescently labeling the nucleic acid portion of polysomes, which are typically isolated as low-velocity nucleoprotein fractions separated on sucrose gradients. Higher-molecular-weight material, representing messages bound by several ribosomes, may be differentially labeled

with respect to lower-molecular-weight material, representing those messages with few or no bound ribosomes. The resulting hybridization can then be used to assay simultaneously the degree to which each individual mRNA is associated with translation machinery. Tracking these associations provides a means to measure the degree to which the expression of any particular message is likely to be controlled at this level (K. Kuhn and P. Sarnow, pers. comm.). In one interesting variation on this approach, membrane-bound polysomes may also be collected and analyzed. Because membrane-bound polysomes contain messages undergoing cotranslational secretion, it is then possible to quickly identify these gene products with respect to any given experimental condition, whether it be time, cell type, or environment (Diehn et al., 1999).

DNA-Protein Interaction

The separation of nucleoprotein components for array analysis is not limited to mRNA bound to polysomes. Indeed, any method that creates a stable or covalent cross-link between the nucleic acid and the protein component may be used. Subsequent separation by immunoprecipitation (*UNIT 20.5*), filter binding, or other chromatographic procedures may then produce a nucleic acid fraction that can be directly labeled and applied to a microarray representation of the genome. Such genome-wide studies are underway for epitope-tagged DNA binding proteins and for proteins that cross-link enzymatically to DNA as part of their normal

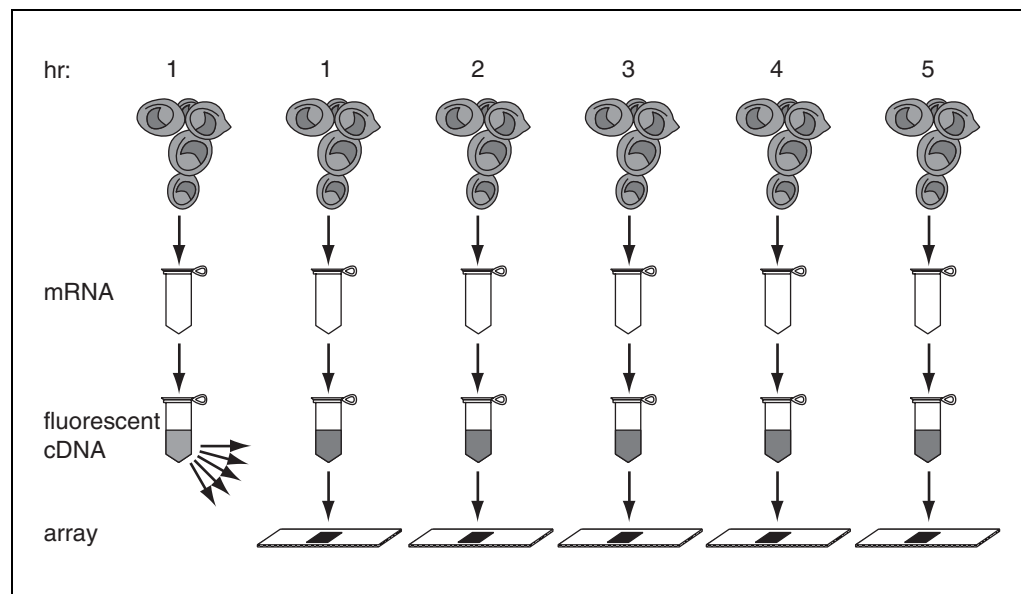


Figure 22.1.3 Monitoring levels of transcript abundance. Time zero is the reference probe for each comparison.

function. Obvious applications of this approach include the study of DNA binding proteins (Chapter 12), endonucleases, and chromatin-associated proteins (*UNIT 21.3*). These methods often produce vanishingly small amounts of nucleic acid source material, often in the sub-nanogram range. For this reason, there is likely to be much development in the area of linear DNA and RNA amplification. Indeed, amplification technology coupled to microarray analysis will allow the study of single-cell expression patterns (Eberwine, 1996).

WHAT ABOUT DATA ANALYSIS?

In addition to the many ways microarray experiments may be carried out, there are equally many ways to analyze the resulting data. Especially for gene expression, it is likely that genome-wide expression analysis will become a field of its own and that an ever-increasing number of innovative approaches will be developed to address the problem of reverse engineering the mechanisms guiding the observed expression patterns. Several methods for visualization and analysis already exist, and many of the necessary tools are readily available to be downloaded from the Internet. The method known as clustering, whereby similar expression patterns are grouped together, has been particularly successful. (Eisen et al., 1998; Bassett et al., 1999; Brown and Botstein., 1999; Iyer et al., 1999). Its success relies upon the fact that genes which participate in a common process, pathway, or function often share common regulatory mechanisms and this, in turn, results in similar expression profiles. Therefore, the functions of previously uncharacterized genes may be revealed by what other genes cluster with it over a broad range of experiments.

WHERE CAN I GET MORE INFORMATION?

The field of microarray technology is evolving rapidly, and therefore the uses and methods of this field will be continuing to change as researchers innovate and disseminate new techniques and ideas. One way to stay connected to this rapid progression is to frequent the various noncommercial web sites listed below. The sites listed contain directions for the fabrication of microarray robots, up-to-date protocol revisions, public access to genome wide expression data, free microarray software tools, forums, and database information.

www-genome.stanford.edu

Stanford Genomic Resources. One-stop shopping for genome-wide expression analysis. This site contains links to several yeast expression databases and research article companions in areas including cell-cycle analysis, sporulation, the diauxic shift, yeast evolution, and yeast clustering. Links to several human expression databases are also available here, including databases relevant to serum starvation, breast cancer expression patterns, and microarray-based CGH.

rana.stanford.edu/software

Stanford Genome Analysis Group software download area. This site provides access to several free microarray-analysis tools that run on your PC, written by Mike Eisen. They include ScanAlyze, an advanced image-analysis package, and clustering/data-visualization applications.

cmgm.stanford.edu/pbrown/mguide

The MGuide. Home of a do-it-yourself guide to building your own microarraying robot, maintained by Patrick Brown's lab at Stanford University. A step-by-step guide, technical drawings, and parts lists, as well as free software, are available for downloading here.

www.nhgri.nih.gov/DIR/LCG/15K/HTML

National Human Genome Research Institutes (NHGRI) Microarray Project. This site features database and data handling information, as well as protocols and research projects.

web.wi.mit.edu/young/expression

Genome-Wide Expression Homepage maintained by Rick Young of the Whitehead Institute. Access to various yeast transcription experiments, protocols, and experiment designs are available.

industry.ebi.ac.uk/~alan/MicroArray

Large-Scale Gene Expression and Microarray Links and Resources. A "personal not-for-profit web site" containing literally hundreds of links to microarray articles, databases, and companies, maintained by Alan Robinson, a researcher at the EMBL-European Bioinformatics Institute (EBI).

LITERATURE CITED

Alizadeh, A., Eisen, M., Botstein, D., Brown, P.O., and Staudt, L.M. 1998. Probing lymphocyte biology by genomic-scale gene expression analysis. *J. Clin. Immunol.* 18:373-379.

- Amundson, S.A., Bittner, M., Chen, Y., Trent, J., Meltzer, P., and Fornace, A.J., Jr. 1999. Fluorescent cDNA microarray hybridization reveals complexity and heterogeneity of cellular genotoxic stress responses. *Oncogene* 18:3666-3672.
- Bassett, D.E. Jr., Eisen, M.B., and Boguski, M.S. 1999. Gene expression informatics—it's all in your mine. *Nature Genet.* 21:51-55.
- Behr, M.A., Wilson, M.A. Gill, W.P., Salamon, H., Schoolnik, G.K., Rane, S., and Small, P.M. 1999. Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* 284:1520-1523.
- Brown, P.O. and Botstein, D. 1999. Exploring the new world of the genome with DNA microarrays. *Nature Genet.* 21:33-37.
- Cheung, V.G., Gregg, J.P., Gogolin-Ewens, K.J., Bandong, J., Stanley, C.A., Baker, L., Higgins, M.J., Nowak, N.J., Shows, T.B., Ewens, W.J., Nelson, S.F., and Spielman, R.S. 1998. Linkage disequilibrium mapping without genotyping. *Nature Genet.* 18:225-230.
- Cho, R.J., Campbell, M.J., Winzler, E.A., Steinmetz, L., Conway, A., Wodicka, L., Wolfsberg, T.G., Gabrielian, A.E., Landsman, D., Lockhart, D.J., and Davis, R.W. 1998. A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol. Cell* 2:65-73.
- Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P.O., and Herskowitz, I. 1998. The transcriptional program of sporulation in budding yeast. *Science* 282:699-705.
- DeRisi, J.L., Iyer, V.R., and Brown, P.O. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680-686.
- Diehn, M., Eisen, M., Brown, P., and Botstein, D. 1999. Large-scale identification of membrane-associated gene products using DNA microarrays, submitted.
- Eberwine, J. 1996. Amplification of mRNA populations using aRNA generated from immobilized oligo(dT)-T7 primed cDNA. *Biotechniques* 20:584-591.
- Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* 95:14863-14868.
- Ferea, T.L., Botstein, D., Brown, P.O., and Rosenzweig, R.F. 1999. Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 96:9721-9726.
- Fodor, S.P., Read, J.L., Pirrung, M.C., Stryer, L., Lu, A.T., and Solas, D. 1991. Light-directed, spatially addressable parallel chemical synthesis. *Science* 251:767-773.
- Hacia, J.G., Brody, L.C., Chee, M.S., Fodor, S.P., and Collins, F.S. 1996. Detection of heterozygous mutations in *BRCA1* using high density oligonucleotide arrays and two-colour fluorescence analysis. *Nature Genet.* 14:441-447.
- Hayward, R., DeRisi, J., Alfadhli, S., Kaslow, D., Brown, P., and Rathod, P. 1999. Shot-gun DNA microarrays and stage specific gene expression in *Plasmodium falciparum*, submitted.
- Iyer, V.R., Eisen, M.B., Ross, D.T., Schuler, G.T., Moore, J.C., Lee, F., Trent, J.M., Staudt, L.M., Hudson, J., Jr., Boguski, M.S., Lashkari, D., Shalon, D., Botstein, D., and Brown, P.O. 1999. The transcriptional program in the response of human fibroblasts to serum. *Science* 283:83-87.
- Kallioniemi, A., Kallioniemi, O.P., Sudar, D., Rutovitz, D., Gray, J.W., Waldman, F., and Pinkel, D. 1992. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258:818-821.
- McAllister, L., Penland, L., and Brown, P.O. 1998. Enrichment for loci identical-by-descent between pairs of mouse or human genomes by genomic mismatch scanning. *Genomics* 47:7-11.
- Nelson, S.F., McCusker, J.H., Sander, M.A., Kee, Y., Modrich, P., and Brown, P.O. 1993. Genomic mismatch scanning: A new approach to genetic linkage mapping. *Nature Genet.* 4:11-18.
- Perou, C.M., Jeffrey, S.S., van de Rijn, M., Rees, C.A., Eisen, M.B., Ross, D.T., Pergamenschikov, A., Williams, C.F., Zhu, S.X., Lee, J.C., Lashkari, D., Shalon, D., Brown, P.O., and Botstein, D. 1999. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc. Natl. Acad. Sci. U.S.A.* 96:9212-9217.
- Pinkel, D., Seagraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., Collins, C., Kuo, W.L., Chen, C., Zhai, Y., Dairkee, S.H., Ljung, B.M., Gray, J.W., and Albertson, D.G. 1998. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nature Genet.* 20:207-211.
- Pollack, J.R., Perou, C.M., Alizadeh, A.A., Eisen, M.B., Pergamenschikov, A., Williams, C.F., Jeffrey, S.S., Botstein, D., and Brown, P.O. 1999. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nature Genet.* 23:41-46.
- Solinas-Toldo, S., Lampel, S., Stilgenbauer, S., Nickolenko, J., Benner, A., Döhner, H., Cremer, T., and Lichter, P. 1997. Matrix-based comparative genomic hybridization: Biochips to screen for genomic imbalances. *Genes Chrom. Cancer* 20:399-407.
- Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., and Futcher, B. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* 9:3273-3297.
- Trent, J.M., Bittner, M., Zhang, J., Wiltshire, R., Ray, M., Su, Y., Gracia, E., Meltzer, P., De Risi, J., Penland, L., and Brown, P. 1997. Use of microgenomic technology for analysis of alterations in DNA copy number and gene expression in malignant melanoma. *Clin. Exp. Immunol.* 107(Suppl 1):33-40.

Wilson, M., DeRisi, J., Kirstensen, H., Imboden, P., Rane, S., Brown, P., and Schoolnik, G. 1999. Exploring drug-induced alterations in gene expression in *Mycobacterium tuberculosis* by microarray hybridization. *Proc. Natl. Acad. Sci. U.S.A.* (in press).

Contributed by Joseph DeRisi
University of California, San Francisco
San Francisco, California

Preparation of mRNA for Expression Monitoring

The ability to construct comprehensive gene expression profiles comprising hundreds to thousands of genes whose RNA levels are monitored simultaneously represents an exciting new capability in molecular biology. This is accomplished by hybridizing mRNA, which has been quantitatively amplified and labeled with biotin, to DNA chips that display thousands of oligonucleotides complementary to the mRNAs of interest. An overview of the entire process is shown schematically in Figure 22.2.1. The Strategic Planning section outlines considerations in starting with poly(A)⁺ versus total RNA, and also discusses oligonucleotide selection for chip design. The Basic Protocol outlines RNA amplification and labeling, which entails cDNA synthesis followed by *in vitro* transcription (IVT), and hybridization of the resulting biotinylated antisense RNA to the chip. An Alternate Protocol is included for purification of cDNA and *in vitro* transcription products on carboxy-coated magnetic beads, making the amplification reaction amenable to automation. Support Protocols 1 and 2 provide methods for quantifying the cDNA product and for producing *in vitro* transcripts of control genes.

CAUTION: Diethylpyrocarbonate (DEPC) is a suspected carcinogen and should be handled with care.

STRATEGIC PLANNING

As this is an RNA-based procedure, care should be taken at every step to avoid contamination of reagents and materials with RNases (*UNIT 4.1*). New plasticware can be used without decontamination processes if kept free of dust and handled exclusively with gloved hands. Filtered pipet tips should be used routinely to avoid contamination with the micropipettor. Any reagent that needs to be made in the laboratory, or that is not supplied RNase-free by the manufacturer, should be treated with DEPC (*UNIT 4.1*).

The amplification procedure is absolutely dependent on clean, intact starting RNA. For cultured cells, the authors use a protocol of lysis in guanidine followed by resin-based purification (such as the Qiagen RNeasy kit). The authors' results using one-step protocols (in which guanidine and phenol are combined) have been variable, presumably because inhibitors of the amplification reaction enzymes can be carried along. Tissues are immediately snap-frozen in liquid nitrogen, then ground to a powder in a dry ice-embedded mortar and pestle. RNA is then extracted in guanidine reagent in a Polytron mixer, followed by extraction with phenol (e.g., Promega RNAgents kit, Ambion Totally RNA kit).

The mRNA amplification can be performed starting with either poly(A)⁺ or total RNA. Although poly(A)⁺ RNA gives higher sensitivity because it largely eliminates mispriming on rRNA during the cDNA reaction, the ability to use total RNA makes it possible to analyze biological systems in which cells or tissues are limited. A key modification in using total RNA is elevation of the first-strand cDNA reaction temperature from 37° to 50°C. Although this will decrease cDNA yield somewhat, it dramatically decreases mispriming on rRNA. The protocol outlined here will preferentially amplify the 3' end of long mRNAs, owing to its dependence on priming with oligo(dT). It is therefore best to use either custom or commercially available chips that display probes toward the 3' end of the mRNA. In general, the authors restrict probe selection to the last 600 bases of coding sequence unless the 3'-untranslated region (3'-UT) is >800 bases, in which case some untranslated sequence is also included.

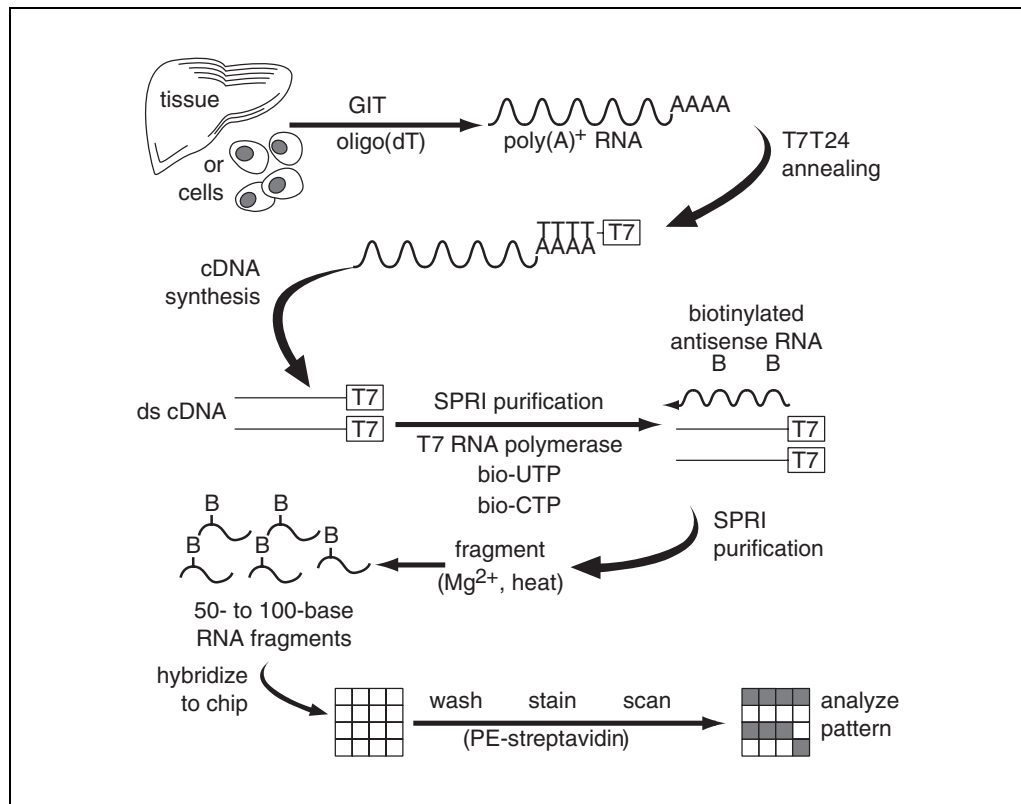


Figure 22.2.1 Chip analysis overview. Abbreviations: GIT, guanidine isothiocyanate, PE, phycoerythrin; SPRI, solid-phase reversible immobilization.

**BASIC
PROTOCOL**

**AMPLIFICATION OF mRNA FOR EXPRESSION MONITORING AND
HYBRIDIZATION TO OLIGONUCLEOTIDE ARRAY CHIPS**

The Basic Protocol details the quantitative amplification and biotin labeling of mRNA to produce antisense RNA for the purpose of gene expression monitoring. Starting RNA is first converted to double-stranded cDNA using a primer containing a T7 RNA polymerase site, so that amplified and labeled RNA can be produced directly in an *in vitro* transcription reaction. The resulting RNA is hybridized to a DNA chip.

Materials

SuperScript cDNA kit (Life Technologies), including:

- 5× First Strand Buffer
- 200 U/μl SuperScript II reverse transcriptase
- 5× Second Strand Buffer
- 10 mM dNTPs
- 10 U/μl *E. coli* ligase
- 2 U/μl *E. coli* RNase H
- 10 U/μl *E. coli* DNA polymerase
- 5 U/μl T4 DNA polymerase

T7T24 primer: 5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCCGG
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT-3' (HPLC purification is recommended)

RNase inhibitor (Life Technologies or Ambion)

RNase-free H₂O (see UNIT 4.1 for DEPC treatment of solutions; prepare from glass-distilled H₂O)

Sample RNA: poly(A)⁺ or total RNA

Sense control transcript pool (see Support Protocol 1)

**Preparation of
mRNA for
Expression
Monitoring**

22.2.2

25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol (molecular biology grade; *UNIT 2.1A*)

7.5 M ammonium acetate (*APPENDIX 2*)

Absolute ethanol

70% (v/v) ethanol in RNase-free H₂O, prechilled to -20°C

10× transcription buffer (Ambion)

10× rNTP mix (see recipe)

100 mM dithiothreitol (DTT)

10 mM Bio-11-CTP and Bio-11-UTP (Enzo Diagnostics)

2500 U/μl T7 RNA polymerase (Epicentre)

RNeasy mini columns with RLT and RPE buffers and collection tubes (Qiagen)

5× fragmentation buffer (see recipe)

20× SSPE (Bio-Whittaker)

0.5% (v/v) Triton X-100 (molecular biology grade; Sigma) in RNase-free H₂O

10 mg/ml herring sperm DNA (Promega)

500 pM Bio948 (see recipe)

20× antisense control transcript pool (see Support Protocol 1)

6× SSPET: 6× SSPE containing 0.005% (v/v) Triton X-100

Thermocycler (e.g., Perkin-Elmer 9600 PCR machine with heated lid)

0.1- to 10-μl filtered micropipet tips (Continental)

Lyophilizer

Small, thin-walled PCR tubes

GeneChip (Affymetrix)

1- to 200-μl filtered gel-loading micropipet tips (Fisher)

Rotisserie-type rotator (Appropriate Technical Resources)

50°C oven

Additional reagents and equipment for quantitation of cDNA (see Support Protocol 2), for quantitation of DNA by spectrophotometry (*APPENDIX 3D*), and for washing, staining, and scanning the GeneChip (see manufacturer's instructions)

NOTE: Many buffers and enzymes are supplied with the SuperScript II cDNA kit. Kit enzymes that are limiting may be ordered separately from Life Technologies.

NOTE: All temperature-controlled reactions are performed in an appropriate thermocycler.

Perform first-strand synthesis

1. Set up a linked program on a thermocycler as follows:

10 min	70°C
65 min	37°C (or 50°C for total RNA)
150 min	15.8°C
indefinitely	4°C (hold).

2. Prepare 10 μl first-strand reagent cocktail for each sample RNA by combining the following reagents, using filtered pipet tips:

- 4 μl 5× First Strand Buffer
- 200 pmol T7T24 primer
- 1 μl RNase inhibitor
- 1 μl 200 U/μl SuperScript II reverse transcriptase
- RNase-free H₂O to 10 μl.

3. Combine each sample RNA with sense control transcript pool. Use 5 μl sense control transcript pool per 1 μg sample poly(A)⁺ RNA or 1 μl control pool per 10 to 20 μg total RNA. Lyophilize to reduce the volume per tube to $<10 \mu\text{l}$. Adjust volume to 10 μl with RNase-free H₂O. Prewarm first-strand reagent cocktail in hand during this time.

Do not dry completely, as this will make it very difficult to resuspend the RNA.

4. Transfer each RNA to a small, thin-walled PCR tube and place in the thermocycler. Begin linked program (step 1).
5. After the 70°C step has completed, allow 2 min for solution temperature to reach 37°C (or 50°C), and add the prewarmed (in hand) first-strand reagent cocktail (10 μl per tube).
6. Hold at the appropriate reaction temperature for the remaining 60 min.

Perform second-strand synthesis

7. Prepare 130 μl second-strand reagent cocktail per tube:

91 μl RNase-free H₂O
30 μl 5 \times Second Strand Buffer
3 μl 10 mM dNTPs
1 μl 10 U/ μl *E. coli* ligase
1 μl 2 U/ μl *E. coli* RNase H
4 μl 10 U/ μl *E. coli* DNA polymerase.

This cocktail can be prepared in advance and kept on ice for up to 90 min.

8. Add 130 μl second-strand reagent cocktail to the tube (total 150 μl) once the thermocycler has shifted to 15.8°C. Mix by pipetting up and down after addition of cocktail.
9. Incubate at 15.8°C for ≥ 2 hr.
10. Add 2 μl of 5 U/ μl T4 DNA polymerase and incubate an additional 5 min.
11. Remove samples and place on ice.

Purify cDNA

12. Add 150 μl of 25:24:1 phenol/chloroform/isoamyl alcohol, vortex, and centrifuge 5 min at 13,000 $\times g$, room temperature.

Many methods are appropriate for removing enzymes and unincorporated nucleotides. Phenol/chloroform extraction followed by ethanol precipitation works well and is outlined here. See Alternate Protocol for a second method for purifying cDNA and IVT products.

13. Carefully remove aqueous top phase to new tube, avoiding contamination with interface.
14. Add 70 μl of 7.5 M ammonium acetate and 0.5 ml absolute ethanol and mix. Centrifuge 20 min at 13,000 $\times g$, room temperature.
15. Remove supernatant. Wash pellet with 0.5 ml of cold (-20°C) 70% ethanol. Vortex well.

The pellet should be vortexed well with the 70% ethanol to prevent inhibition of the in vitro transcription reaction by organic contamination.

16. Centrifuge 10 min at 13,000 $\times g$, room temperature, remove supernatant, and allow pellet to air dry for several minutes.

17. Resuspend the cDNA pellet in 25 μ l RNase-free H₂O and quantitate cDNA (see Support Protocol 2).

Perform in vitro transcription (IVT)

18. For each cDNA, set up one reaction as follows:

100 ng cDNA
6 μ l 10 \times transcription buffer
6 μ l 10 \times rNTP mix
3 μ l 100 mM DTT
2.4 μ l 10 mM Bio-11-UTP
2.4 μ l 10 mM Bio-11-CTP
2 μ l RNase inhibitor
2 μ l 2500 U/ μ l T7 RNA polymerase
RNase-free H₂O to 60 μ l.

The cocktail can be prepared in advance and kept on ice for up to 3 hr. It should be prewarmed to room temperature before adding to cDNA to avoid precipitates.

19. Incubate at 37°C for 8 hr to overnight.

If desired, the IVT reaction can be stored up to 48 hr at -80°C before purification.

Purify IVT RNA

20. Bring volume of IVT reaction to 100 μ l with RNase-free H₂O, add 350 μ l RLT buffer, and mix.
21. Add 250 μ l absolute ethanol and mix.
22. Apply each sample to an RNeasy mini column placed in a collection tube. Microcentrifuge 15 sec at >8000 \times g, room temperature.

Qiagen RNeasy columns work well for removal of unincorporated nucleotides. It is very important to remove the nucleotides in order to accurately quantitate the IVT RNA. This protocol follows the manufacturer's protocol with the following changes: no 2-mercaptoethanol is used in the RLT buffer, and final elution is done twice using 50 μ l RNase-free H₂O each time.

23. Transfer column to a new collection tube. Add 500 μ l RPE buffer.

The RPE working solution is prepared as described in the kit directions; the supplied stock is diluted with 4 vol absolute ethanol.

24. Centrifuge 15 sec at >8000 \times g, room temperature.
25. Discard flowthrough and replace column on the same collection tube. Add 500 μ l RPE buffer and microcentrifuge 2 min at maximum speed.
26. Transfer column to a new collection tube. Add 50 μ l RNase-free H₂O and centrifuge 1 min at >8000 \times g. Repeat elution step, pooling the second eluate with the first.
27. Quantitate RNA yield spectrophotometrically at 260 nm (APPENDIX 3D).

Fragment RNA

28. Bring volume of 10 μ g IVT RNA to 24 μ l with RNase-free H₂O.
29. Add 6 μ l of 5 \times fragmentation buffer. Mix carefully and incubate 35 min at 95°C in the thermocycler. Allow to cool to room temperature.

Fragmented RNA can be stored up to 1 year at -80°C. Thaw at 37°C for 5 min before preparing hybridization solution.

Hybridize RNA to chip

30. For each reaction, prepare 170 μ l hybridization master mix as follows:

- 51 μ l 20 \times SSPE
- 1.7 μ l 0.5% Triton X-100
- 1.7 μ l 10 mg/ml herring sperm DNA
- 20 μ l 500 pM Bio948
- 10 μ l 20 \times antisense control transcript pool
- 85.6 μ l RNase-free H₂O.

31. Add 170 μ l hybridization master mix to 30 μ l of each fragmented IVT RNA.
32. Heat to 99°C for 10 min in the thermocycler, then move to 37°C for \geq 5 min.
33. Microcentrifuge 5 min at maximum speed.
34. Insert a filtered micropipet tip into the upper septum of a GeneChip to provide a vent. Fill the chip from the bottom septum with the 200 μ l hybridization solution using a filtered gel-loading micropipet tip. Remove the vent and cover both septa with transparent tape.
Use one GeneChip for each IVT RNA.
35. Incubate in a 40°C oven overnight (16 to 18 hr) on a rotisserie-type rotator running at \sim 60 rpm.
36. Transfer the chip to a 50°C oven and continue rotating the chip for exactly one hour.
37. Remove chip from the oven and insert a filtered micropipet tip into the upper septum to vent the chip.
38. Using a micropipettor with plunger fully depressed, insert a 200- μ l gel-loading pipet tip into the lower septum. Holding the chip vertically, slowly draw out all of the hybridization solution and store it in a microcentrifuge tube at -20° C.
39. Fill the chip with 6 \times SSPET.
40. Wash and stain the chip with phycoerythrin according to the manufacturer's protocols. Scan the chip as soon as possible.

If the chip cannot be washed right away, seal the two septa with transparent tape and store at 4°C for up to a few hours.

Once the chips have been stained with phycoerythrin they should be kept wrapped in foil to avoid photobleaching. Although it is recommended that they be scanned as soon as possible, if necessary, they can be stored at 4°C, wrapped in foil, for several hours before scanning.

SUPPORT PROTOCOL 1

IN VITRO TRANSCRIPTION OF CONTROL GENES AND PREPARATION OF TRANSCRIPT POOLS

It is essential to include labeled antisense control RNAs of known concentration in each hybridization reaction to normalize chip-to-chip variation and to allow construction of a standard curve for converting hybridization intensity to mRNA frequency. It is also very helpful to add unlabeled sense transcripts to monitor performance of the amplification protocol. This protocol describes the preparation of these two control pools.

Preparation of mRNA for Expression Monitoring

22.2.6

Additional Materials (also see Basic Protocol)

Plasmids (Table 22.2.1; ATCC #87482 to 87490)
2500 U/μl T3 RNA polymerase (Enzo Diagnostics)
25 mM 4rNTP mix: 25 mM each rGTP, rCTP, rATP, and UTP (Ultrapure; Pharmacia Biotech) in RNase-free H₂O

Additional reagents and equipment for purifying IVT products (see Alternate Protocol)

Prepare sense and antisense transcripts

1. Prepare linearized plasmid templates according to Table 22.2.1.

See UNIT 3.1 and manufacturer's instructions for additional information on reaction conditions.

2. Purify plasmid DNA by phenol/chloroform extraction and ethanol precipitation (see Basic Protocol, steps 12 to 16). Resuspend at ~0.1 mg/ml in RNase-free H₂O and quantitate DNA (see Support Protocol 2).
3. For labeled antisense transcripts, prepare four tubes containing a master mix that includes all components except plasmid DNA:

6 μl 10× transcription buffer
6 μl 10× rNTP mix
3 μl 100 mM DTT
2.4 μl 10 mM Bio-11-UTP
2.4 μl 10 mM Bio-11-CTP
2 μl RNase inhibitor
2 μl 2500 U/μl T7 RNA polymerase
RNase-free H₂O to 60 μl.

The 60-μl volume must include the 100 ng DNA that will be added in step 5.

The master mix can be prepared in advance and held on ice for up to 3 hr, but should be warmed to room temperature before adding to plasmid DNA to avoid precipitates.

Table 22.2.1 Preparation of Plasmid Template Controls

Name ^a	ATCC #	Transcript size (kb)	Sense RNA		Antisense RNA	
			Linearize with	Polymerize with	Linearize with	Polymerize with
pGIBS-LYS ^b	87482	1.0	<i>NotI</i>	T3		
pGIBS-PHE ^b	87483	1.3	<i>NotI</i>	T3		
pGIBS-THR ^b	87484	2.0	<i>NotI</i>	T3		
pGIBS-TRP ^b	87485	2.5	<i>NotI</i>	T3		
pGIKS-BioB	87487	1.1			<i>XhoI</i>	T7
pGIKS-BioC	87488	0.8			<i>XhoI</i>	T7
pGIKS-BioD	87489	0.7			<i>XhoI</i>	T7
pGIKS-CRE	87490	1.0			<i>XhoI</i>	T7

^aAbbreviations: BioB, BioC, and BioD are cloned fragments from the *E. coli bioB*, *bioC*, and *BioD* genes, respectively. LYS, PHE, THR, and TRP are fragments from the *Bacillus subtilis lysA*, *pheA*, *thrBC*, and *trpEDCF* genes, respectively. CRE is a fragment from the Cre recombinase derived from *E. coli* bacteriophage P1. pGIBS and pGIKS are derived from the Bluescript KS II vector (Stratagene).

^bpGIBS-LYS, -PHE, -THR, -TRP, and DAP contain a 40-nucleotide synthetic poly(A) tract at the 3' end of the respective genomic fragments derived from *B. subtilis*. Sense IVT transcripts derived from *NotI*-linearized plasmid templates will contain the artificial poly(A) tail. Plasmids linearized with *Bam*HI prior to T3 IVT will generate sense transcripts without the synthetic poly(A) tract.

4. For unlabeled sense transcripts, prepare four tubes of master mix as in step 3, but substitute 25 mM 4rNTP mix for the 10× rNTP mix, eliminate the biotinylated nucleotides, and substitute T3 for T7 RNA polymerase.
5. Add 100 ng of each linearized plasmid DNA to a separate tube and incubate at 37°C for 8 hr to overnight.

The transcripts can be stored for up to 48 hr at –80°C, if desired, before purification.

6. Purify IVT products as described (see Alternate Protocol).
7. Quantitate carefully by absorbance at 260 nm (*APPENDIX 3D*).
8. Dilute unlabeled sense transcripts to 200 nM stock solutions in RNase-free H₂O for long-term storage (up to 1 year) at –80°C.
9. Fragment biotin-labeled antisense transcripts for standard chip controls (BioB, BioC, BioD, CRE) as described (see Basic Protocol, steps 18 and 19).
10. Dilute fragmented antisense transcripts to 20 nM stock solutions in 6× SSPET with 0.1 mg/ml herring sperm DNA for long-term storage (up to 1 year) at –80°C.

Prepare control transcript pools

11. Prepare 20× antisense control transcript pool by combining the following transcripts in 6× SSPET with 0.1 mg/ml herring sperm DNA. Store up to 6 months at –80°C.

30 pM fragmented BioB transcript
100 pM fragmented BioC transcript
500 pM fragmented BioD transcript
2 nM fragmented CRE transcript.

12. Prepare sense control transcript pool by combining the following transcripts in RNase-free H₂O. Store up to 6 months at –80°C in aliquots of 50 to 100 µl.

10 pM LYS transcript
30 pM PHE transcript
90 pM THR transcript
180 pM TRP transcript.

ALTERNATE PROTOCOL

SOLID-PHASE REVERSIBLE IMMOBILIZATION (SPRI) PURIFICATION OF cDNA AND IVT PRODUCTS

Because of the labor-intensive nature of the mRNA amplification protocol, the procedure has been adapted for automation. Although the assembly of the reactions is easily converted to liquid handling robotics, purification of final products should be free of organic solvents and the requirement for centrifuges. A magnetic bead-based purification protocol of cDNA and IVT products has been developed to meet this requirement. This protocol was adapted from a protocol for the automated preparation of DNA sequencing templates (DeAngelis et al., 1995). It is equivalent to the purification methods described in the Basic Protocol (steps 12 to 17 or steps 20 to 27).

Materials

Carboxy-coated magnetic beads (PerSeptive BioSystems for cDNA purification;
Bangs Laboratories for IVT purification)

0.5 M EDTA (*APPENDIX 2*)

Sample to be purified: cDNA (see Basic Protocol, step 11) or IVT RNA (see Basic Protocol, step 19, or see Support Protocol 1, step 6)

2.5 M NaCl/20% (w/v) PEG 8000 (molecular biology grade; RNase free)

70% (v/v) ethanol in RNase-free H₂O
10 mM Tris acetate, pH 7.8 (RNase free)
Magnetic stand (CPG)

Additional reagents and equipment for determining concentration of cDNA (see Support Protocol 2) or RNA (*APPENDIX 3D*)

Purification of cDNA

- 1a. Aliquot 10 μ l PerSeptive carboxy-coated magnetic beads per 150- μ l cDNA reaction, dispensing the total volume of beads into a single microcentrifuge tube.
- 2a. Place the tube on a magnetic stand and allow the beads to separate to the side of the tube. Carefully remove supernatant with a micropipet.
- 3a. Add 0.5 M EDTA equal to the starting volume and resuspend the beads by gentle vortexing or agitation. Replace tube on magnetic stand, wait for beads to separate, and remove supernatant. Repeat this washing procedure two more times.
- 4a. Resuspend beads in 0.5 M EDTA equal to the starting volume.
- 5a. To each tube of cDNA, add 150 μ l of 2.5 M NaCl/20% PEG 8000 and 10 μ l beads and mix by gentle vortexing or agitation. Incubate 10 min at room temperature.
- 6a. Place tubes on magnetic stand and allow beads to separate to the side of the tube (~2 min for the original separation, faster for the washes).
- 7a. Draw off the supernatant, then wash the beads twice with 150 μ l of 70% ethanol. Remove as much of the final ethanol wash as possible and allow to air dry for 2 min.
- 8a. Elute RNA by adding 25 μ l of 10 mM Tris acetate, pH 7.8, and incubating 5 min at room temperature.
- 9a. Place tube on magnetic stand and save supernatant.
- 10a. Determine cDNA concentration by PicoGreen fluorescence (see Support Protocol 2).

Purification of IVT RNA

- 1b. Aliquot 20 μ l Bangs Laboratories carboxy-coated magnetic beads per 60- μ l IVT reaction, dispensing the total volume of beads into a single microcentrifuge tube.
- 2b. Place the tube on a magnetic stand and allow the beads to separate to the side of the tube. Carefully remove supernatant with a micropipet.
- 3b. Add 0.5 M EDTA equal to the starting volume and resuspend the beads by gentle vortexing or agitation. Replace tube on magnetic stand, wait for beads to separate, and remove supernatant. Repeat this washing procedure two more times.
- 4b. Resuspend beads in 1.25 M NaCl/10% PEG equal to the starting volume.
- 5b. To each tube of IVT RNA, add 60 μ l of 2.5 M NaCl/20% PEG 8000 and 20 μ l beads and mix by gentle vortexing or agitation. Incubate 10 min at room temperature.
- 6b. Place tubes on magnetic stand and allow beads to separate to the side of the tube (~2 min for the original separation, faster for the washes).
- 7b. Draw off the supernatant, then wash the beads twice with 150 μ l of 70% ethanol. Remove as much of the final ethanol wash as possible and allow to air dry for 3 min.

- 8b. Elute RNA by adding 25 μ l of 10 mM Tris acetate, pH 7.8, and incubating 5 min at room temperature.
- 9b. Place tube on magnetic stand and save supernatant.
- 10b. Determine RNA concentration by absorbance at 260 nm (*APPENDIX 3D*).

**SUPPORT
PROTOCOL 2**

QUANTITATION OF cDNA

If IVT yield is low, it can be helpful to determine the amount of cDNA that is being produced. For optimal IVT yield, do not exceed 100 ng cDNA in the standard reaction. Quantitation of cDNA can be performed with Molecular Probes PicoGreen reagent.

Materials

PicoGreen dsDNA Quantitation Kit (Molecular Probes), including
100 ng/ μ l standard DNA stock solution
20 \times TE buffer
PicoGreen reagent
cDNA to be quantitated (see Basic Protocol and Alternate Protocol)
Black-walled 96-well plate (Corning)
Fluorimager (Molecular Dynamics, model FSI)

1. Prepare 1 ml of 2 μ g/ml diluted standard DNA by diluting 20 μ l of 100 ng/ μ l standard DNA stock solution in 980 μ l of 1 \times TE buffer.
2. Dilute PicoGreen reagent 1:200 (v/v) with 1 \times TE buffer. Prepare enough diluted reagent so that 100 μ l can be placed in each well.
3. Pipet 100, 50, 20, 10, 5, 2, and 0 μ l dilute standard DNA into seven wells in a black-walled 96-well plate. Bring each to a total of 100 μ l with 1 \times TE buffer.
The amounts of DNA in the standard wells are 200, 100, 40, 20, 10, 4, and 0 ng, respectively.
4. Pipet 2 μ l cDNA to be quantitated into additional wells, as needed.
5. Add 100 μ l diluted PicoGreen reagent to each well.
6. Read fluorescence of each well using a fluorimager.
7. Generate a volume report on each well according to manufacturer's instructions. Use volume number versus ng per well to generate a standard curve.
8. Calculate the concentration of cDNA in test wells.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Bio948, 500 pM

500 pM biotinylated control oligonucleotide Bio948 (5'-GTCAA-GATGCTACCGTTCAG-3')
6 \times SSPE (Bio-Whittaker)
0.1 mg/ml herring sperm DNA (Promega)
0.005% (v/v) Triton X-100 (molecular biology grade; Sigma)
Prepare in RNase-free H₂O
Store up to 1 year at -20°C

Fragmentation buffer, 5×

Dissolve 6.06 g Tris base (Sigma; molecular biology grade) in 175 ml RNase-free H₂O. Adjust pH to 8.1 with glacial acetic acid. Add 12.3 g potassium acetate (from an unopened or dedicated bottle) and 8.04 g magnesium acetate (Sigma; molecular biology grade) and adjust volume to 250 ml (final pH ~8.4). Filter sterilize with a 0.2- μ m filter. Store up to 6 months at -20°C.

rNTP mix, 10×

30 mM rGTP
15 mM rATP
12 mM rCTP
12 mM UTP

Prepare using RNase-free H₂O and Ultrapure reagents from Pharmacia Biotech. Store up to 3 months at -20°C in small aliquots (e.g., 50 to 100 μ l) to prevent multiple freeze/thaw cycles.

COMMENTARY

Background Information

The availability of thousands of gene sequences through high-throughput sequencing efforts, including the Human Genome Project, created an immediate need for highly parallel methods for assessing RNA levels of many genes simultaneously. One approach that is being practiced by an ever-expanding number of laboratories is the hybridization of labeled cellular RNA or cDNA to microarrays displaying DNA probes complementary to the genes of interest. Both cDNA (Schena et al., 1995) and oligonucleotide (Lockhart et al., 1996) probe arrays are used, and each type has its relative merits and shortcomings. The strength of cDNA arrays lies in the ability to display probes on the array prior to obtaining any sequence information, allowing the possibility of novel gene discovery. The disadvantages are that (1) quantitation is possible in a relative sense (i.e., fold-induction) but not an absolute sense (i.e., number of mRNA molecules per million), and (2) it is difficult to discern expression of different members of multigene families due to cross-hybridization. Construction of oligonucleotide arrays, on the other hand, requires that DNA sequence information be available for every gene to be monitored, precluding use of the arrays for novel gene discovery. This limitation is partially mitigated, however, by the recent availability of chips monitoring tens of thousands of expressed sequence tags (ESTs) with no known function. Advantages of oligonucleotide arrays include (1) the possibility of absolute quantitation, by virtue of averaging across multiple probes for every gene being monitored, and (2) the ability to inde-

pendently follow the expression of closely related gene family members through judicious probe selection.

Critical Parameters

The most critical parameter affecting the success of the protocol is the quality of the starting RNA. It is absolutely essential to start with RNA that is intact and free from inhibitors of the enzymes in the amplification procedure. For sample comparisons, e.g., control versus treated sample, it is necessary that the RNA samples be of similar quality and, preferably, that samples are hybridized to arrays within the same lot to minimize chip-to-chip variation.

Troubleshooting

The problems that can arise during expression monitoring procedures are detailed in Table 22.2.2.

Anticipated Results

The yield of amplified *in vitro* transcription (IVT) product reflects the quality and quantity of the starting RNA and, in general, is an indicator of the success of the subsequent chip hybridization. A good amplification reaction from 1 μ g poly(A)⁺ or 10 μ g total RNA can yield 50 to 60 μ g IVT product. Successful hybridizations can be obtained with lesser yields, but yields of \leq 10 μ g, when the quantity of starting RNA was not limiting, are generally predictive of nonoptimal IVT product.

Time Considerations

The entire protocol is generally carried out over a three-day period, starting with purified

Table 22.2.2 Troubleshooting Guide for Expression Monitoring

Problem	Possible cause	Action
Little or no cDNA yield	Beginning RNA quality was poor	Visualize size range of starting RNA on a gel. There should be a smear extending well above 5 kb, and not much <500 bases.
	RNA or reactants were RNase contaminated	Check RNA by incubating an aliquot at room temperature several hours and running on gel as above. Reactants can be incubated with test RNA and run on gel for visualization of degradation. Include RNase inhibitor in cDNA reaction.
	Enzyme inhibitors present in RNA preparation	Check signals for members of the sense control transcript pool. If 5' end signals are poor and there is no evidence for RNase contamination, reisolate sample RNA using a protocol with a more stringent purification method.
Little or no IVT yield	Too little cDNA template	Quantitate cDNA, use ≥ 50 ng in IVT reaction. If cDNA yield is low, refer to above section (little or no cDNA yield).
	Inhibitor of polymerase present	Residual phenol/chloroform or dNTPs can inhibit polymerase reaction. Reprecipitate cDNA, being careful to wash pellet with 70% ethanol.
	Wrong polymerase used	Be sure polymerase matches the incorporated promoter
	DTT inactive	Make fresh DTT stock
	Limiting nucleotide concentration	If making unlabeled RNA, be sure to substitute 25 mM 4rNTP solution for the $10 \times$ rNTP mix used with biotin nucleotides.
Blank or low signal on GeneChip	Poor RNA quality	Even when cDNA and IVT yields seem adequate, poor RNA quality can be a cause of poor chip data. Look at housekeeping gene data for ratio of 5' and 3' signal. With oligo(dT)-primed cDNA, expect 5' signal of β -actin and GAPDH ^a to be at least half of the 3' signal.
	Stringency too high	Check incubator temperature and buffer solutions
	Poor staining of biotin with streptavidin conjugate, or bleaching of phycoerythrin	SA-PE ^a reagents can be variable. Try a different lot, or increase the amount of SA-PE used in the staining solution. Store SA-PE in the dark. Wrap chips in aluminum foil after staining to prevent bleaching before scanning.
	Scanning at wrong wavelength (even edge controls will be very light or absent)	Ensure that scanner is set at 560 nm and rescan
	Hybridization inhibitor in solution	Try using the same hybridization solution on a second chip of the same lot. If signal improves, it may be necessary to incorporate a prehybridization step.
	Scanner needs adjustment	Call for service; try another scanner if available
	Defective lot of GeneChips	Rehybridize saved hybridization solution to a lot of chips that has previously given strong signals. If signal improves, and hybridization inhibitor is not indicated, contact manufacturer.

continued

Table 22.2.2 Troubleshooting Guide for Expression Monitoring, continued

Problem	Possible cause	Action
A few very bright features per gene, the rest very quiet	Failure of fragmentation reaction	Check 5× fragmentation buffer and 94°C temperature control unit. If additional IVT RNA is available, try repeating the fragmentation reaction and check on a gel. Size should be ~20 to 50 bases.
High background values	Stringency too low	Try rewashing the chips at increased temperature and rescanning
	Contaminant	Recommend hybridization solutions be made with glass-distilled water. Some deionization systems can be problematic. If this is suspected, glass-distilled water is commercially available.
Sporadic pattern of round dots all over chip	SA-PE ^a aggregates	Manufacturer recommends the solution be vortexed and then microcentrifuged 2 min prior to taking aliquot
Round area in middle of chip has very low signal compared to rest of chip	Volume of hybridization solution is too low	Need a minimum of 180 µl. If hybridization solution cannot be remade, add 6× SSPET, 0.1 mg/ml herring sperm DNA to the saved hybridization solution to bring the volume to 200 µl and rehybridize to new chip.
	Rotator stopped during hybridization	Rehybridize saved solution to new chip, checking that rotator is functioning properly.

^aGAPDH, glyceraldehyde 3-phosphate dehydrogenase; SA-PE, streptavidin-phycoerythrin.

poly(A)⁺ or total RNA, and the time estimates are for six samples or less being processed simultaneously. On the first day, first- and second-strand cDNA synthesis reactions are carried out (3.5 hr), followed by cDNA purification (1 hr). IVT reactions are then set up and incubated overnight. On the second day, the IVT product is purified (1.5 hr) and quantitated. The RNA is fragmented (40 min), then added to hybridization reagents and loaded on the chip for overnight hybridization. On the third day, the hybridization is continued at elevated temperature (1 hr), then the chips are washed and stained on a fluidics station (40 min/4 chips). Scanning each chip takes ~25 min.

Literature Cited

- DeAngelis, M.M., Wang, D.G., and Hawkins, T.L. 1995. Solid-phase reversible immobilization for the isolation of PCR products. *Nucl. Acids Res.* 23:4742-4743.
- Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., and Brown, E.L. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nature Biotechnol.* 14:1675-1680.
- Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270:467-470.

Key References

- Lockhart et al., 1996. See above.
This paper established that, in addition to its role as a resequencing tool, the oligonucleotide array could be used as a powerful methodology for transcriptional profiling.
- Schena et al., 1995. See above.
A seminal paper on the potential of robotics and microfabrication to accelerate cDNA analysis using a highly parallel format.
- Wodicka, L., Dong, H., Mittmann, M., Ho, M.H., and Lockhart, D.J. 1997. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nature Biotechnol.* 15:1359-1367.
Demonstration that oligonucleotide arrays can be used to simultaneously monitor expression of all genes of a eukaryotic organism.

Contributed by Michael C. Byrne,
Maryann Z. Whitley,
and Maximillian T. Follettie
Genetics Institute
Cambridge, Massachusetts

Profiling Human Gene Expression with cDNA Microarrays

The study of the timing and extent of expression of mRNAs coding for particular genes is one of the central research areas of molecular biology. The literature is full of compelling examples of the utility of knowing that a particular message appears, disappears, or is highly modulated at the time a particular natural or pathological process is affecting a cell. The tools available for making such determinations have steadily advanced. Early efforts utilized kinetic studies (Bishop et al., 1974) of how rapidly the various component populations of mRNA from different cells hybridized to each other to obtain an overview of the likely number and abundance of messages. Such studies were augmented by a variety of techniques that immobilized size-fractionated message populations on substrates and determined the abundance of individual species of mRNA within the pools (Alwine et al., 1977). Current methods for large surveys of gene expression patterns immobilize single cDNA representatives of many independent genes in regular arrays on substrates and quantify the relative abundance of many specific mRNAs present in the message pool of a sample simultaneously (Schena et al., 1995; Lockhart et al., 1996). The data from such surveys can be used to compare patterns of gene expression to delineate sets of genes regulated in a particular process (Spellman et al., 1998; Iyer et al., 1999) or coregulated by specific transcriptional regulators (DeRisi et al., 1997). Overall patterns of expression can be used to evaluate similarities in cell types (Khan et al., 1998).

Fluorescent cDNA microarray technology is a particularly potent variation of the latter technique. This method is especially useful for making estimates of the abundance of particular messages relative to a designated source of mRNA that serves as a reference point. Commercial support of this technology has recently reached a level where it is reasonable for departments or large laboratories to consider setting up their own cDNA array facility. This set of protocols is intended to serve as a basic introduction to making and using cDNA microarrays for those embarking on this path.

There are three fundamental types of operations required in a cDNA microarray experiment. The first (see Basic Protocol 1), cDNA amplification and printing, deals with making the cDNA microarray itself. It is necessary to collect an inventory of cDNA bacterial clones that represent the genes whose message abundance one wishes to survey. Plasmid templates are made from these clones and used as PCR substrates to produce DNA representations of the expressed-sequence tag (EST) inserts. The PCR products are characterized by agarose gel electrophoresis (see Support Protocol 1) and quantitated by fluorometry as described in Support Protocol 2. The PCR products are then purified and spotted onto poly-L-lysine-coated microscope slides. Preparation of poly-L-lysine-coated slides is described in Support Protocol 3. In the second operation (see Basic Protocol 2), RNA extraction and labeling, RNA is extracted from the cell samples to be examined, purified, and used as the substrate for reverse transcription in the presence of fluor-derivatized nucleotides. This procedure provides the tagged representations of the mRNA pools of the samples that will be hybridized to the gene-specific cDNA detectors immobilized on the microarray. The third fundamental operation (see Basic Protocol 3), hybridization and data extraction, covers the steps in which fluor-labeled cDNAs hybridize to their complements on the microarray, and the resulting localized concentrations of fluorescent molecules are detected and quantitated.

cDNA AMPLIFICATION AND PRINTING

This protocol describes the steps required to produce a cDNA microarray. Gene-specific DNA is produced by PCR amplification of purified template plasmid DNAs from cloned ESTs. The PCR product is purified by ethanol precipitation, thoroughly resuspended in 3× SSC, and printed onto a poly-L-lysine-coated slide (see Support Protocol 3).

Materials

Master set of clone-purified, sequence verified human ESTs (e.g., gf211 release, Research Genetics) in bacterial cells
LB medium (Biofluids, or see *UNIT 1.1*) containing 100 µg/ml carbenicillin (added from 100 mg/ml stock; see recipe)
70% ethanol
100% denatured ethanol
Super Broth (Biofluids) containing 100 µg/ml carbenicillin (add from 100 mg/ml carbenicillin stock; see recipe)
45% (w/v) glycerol (enzyme grade, sterilize by autoclaving and store at room temperature)
96-well alkaline lysis miniprep kit (Edge BioSystems)
Lysis buffer
RNase solution
Resuspension buffer
96-well receiving plates and filter plates
Precipitation buffer
Neutralization buffer
Wide-bore pipet tips
Deep-well plates
T low E buffer (see recipe)
10× PCR buffer
100 mM dATP
100 mM dGTP
100 mM dCTP
100 mM dTTP
1 mM PCR primer AEK M13F (5'-GTTGTAAAACGACGGCCAGTG-3')
1 mM PCR primer AEK M13R (5'-CACACAGGAAACAGCTATG-3')
Taq DNA polymerase (AmpliTaq, PE Biosystems), store at -20°C
Ethanol/acetate solution (see recipe)
20× SSC (*APPENDIX 2*)
Succinic anhydride (Sigma-Aldrich)
1-methyl-2-pyrrolidinone (Sigma-Aldrich)
1 M sodium borate, pH 8.0 (see recipe)
96-well round-bottom and V-bottom plastic cell culture plates (Corning)
Centrifuge with a horizontal microplate carrier with a depth capacity of 6.2 cm for spinning microtiter plates and filtration plates (e.g., Sorvall Super T 21, Sorvall with ST-H750 microplate carrier rotor)
96-pin multi-blot replicator (V&P Scientific)
Household 1-gallon sealable storage bags (e.g., Glad Lock)
Deep-well plates
Microporous tape sheets (e.g., AirPore Tape Sheets, Qiagen)
Platform shaker with holders for deep-well plates at 37°C
Sterile 96-well plate seals (e.g., Elkay Products)
Thin-wall 96-well PCR plates and PCR plate sealer (e.g., CycleSeal plate sealer, Robbins Scientific)

96-well thermal cycler (MJ Research)
Microtiter plate washer (e.g., Immunowash Microplate washer, Bio-Rad)
Heat sealable storage bags and heat sealer
65°C incubator
Robotic slide printer (e.g., GeneMachines, Genetic Microsystems, Genetix, Cartesian Technologies) and pens (e.g., Majer Precision Engineering, TeleChem International)
Diamond scribe for writing on slides
Slide box, plastic with no paper or cork liners (e.g., PGC Scientifics)
~24 × 34 × 5-cm Pyrex baking dish
Poly-L-lysine-coated glass slides (see Support Protocol 3)
30-slide stainless steel rack and 30-slide glass tank (Shandon/Lipshaw)
1-liter glass tank
Additional reagents and equipment for fluorometric DNA determination (see Support Protocol 2) and agarose gel electrophoresis (see Support Protocol 1)

Grow EST clones

1. Incubate sealed master plates of clone-purified, sequence verified human ESTs overnight at 37°C.

Most suppliers provide low-density bacterial cultures. Replicating directly from these dilute stocks frequently results in non-growth in the secondary culture. If making template from a plate that has previously been cultured to high density before freezing, this initial growth step should not be used, as it will reduce the viability of the cultures.

2. Prepare replicate sets of standard 96-well round-bottom plates by labeling all plates and placing 100 µl LB medium containing 100 µg/ml carbenicillin in each well.

These plates will be used as working copies.

To preserve the master set of plates, it is useful to make replicate copies of the master plate to serve as working copies when the master plate is first replicated. Check to insure that the EST clones are in a vector conferring ampicillin resistance, as is common with human IMAGE clones.

3. Centrifuge the master plates briefly for 2 min at $167 \times g$ (1000 rpm in a horizontal microplate carrier), 20° to 25°C, to remove condensation and droplets from the seals before opening.

Bacterial culture fluid on the sealers can easily be transferred from one well of the plate to others, cross-contaminating the stocks.

4. Partially fill a container with 100% ethanol, dip the 96-pin multi-blot replicator, remove from the ethanol bath, and then flame the pins.
5. Allow the 96-pin multi-blot replicator to cool briefly, then dip the replicating tool in the master plate, and then into the corresponding LB daughter plate. Repeat as necessary for each plate to be inoculated.

It is useful to color the plate corner near the A-1 well of all master and daughter plates with a marker pen before beginning the replication process, to reduce mistakes in relative orientation of the plates. The suggested plates have a notch at this corner as well.

6. Place the inoculated LB daughter plates with lids on into a 1-gallon sealable storage bag containing a moistened paper towel and grow overnight at 37°C.

Many 37°C incubators tend to dry out microtiter plate cultures. Placing the plates in a highly humidified bag avoids this problem.

7. Fill 96-well deep-well plates with 1 ml Super Broth containing 100 µg/ml carbenicillin per well.

These plates will serve as the source of culture for template preparation.

8. Using the 96-pin replicating tool, inoculate the deep-well plates directly from the freshly grown LB plates.
9. Cover the openings of the deep-well plates with microporous tape sheets and place the plastic lid over the sheet. Place the plates on a platform shaker for 24 hr at 200 rpm, 37°C.
10. Add 50 µl of 45% sterile glycerol to each well of any working plates that are to be frozen (−80°C) and subsequently used as culture sources.

Isolate plasmid templates

11. Warm the lysis buffer (from lysis miniprep kit) to 37°C until the SDS dissolves.

This buffer can be stored at room temperature.

12. Add 1 ml RNase solution to 100 ml resuspension buffer (from lysis miniprep kit) and store at 4°C.

The remaining reagents from the kit, neutralization buffer and precipitation buffers, are ready to use and should be stored at 4°C.

13. Prepare the 96-well receiving plates (from lysis miniprep kit) by adding 350 µl of 100% denatured ethanol to each well of the receiving plates. Place the filter plate on top and secure in place with tape.

Handle with care as the wells will be very full.

14. Centrifuge the bacterial cultures in the deep-well plates for 7 min at 1500 × g, 20° to 25°C, in a centrifuge equipped with a horizontal rotor for 96-well plates.
15. Invert briefly to remove the supernatant and tap out excess medium on a clean paper towel.

Do not delay or the pellets will loosen and may be lost when pouring off excess media.

16. Resuspend the pellet in 100 µl resuspension buffer. Vortex until entire pellet is re-suspended.

This step is critical.

Poor resuspension of the cells results in clumps of cells that do not lyse in subsequent steps. This reduces the yield and decreases the purity of the product.

17. Add 100 µl lysis buffer. Mix gently by rocking the plates from side; avoid shearing the bacterial chromosomal DNA.
18. Add 100 µl precipitation buffer to each well. Mix briefly.
19. Add 100 µl neutralization buffer to each well. Vortex.
20. Transfer the contents of the deep wells to the prepared filter plates/receiving plate stacks using the wide-bore pipet tips provided in the kits.
21. Centrifuge the stacked plates for 12 min at 1500 × g, 20° to 25°C.
22. Remove the stacked plates from the centrifuge and discard the filter plates. Decant the ethanol and filtrate from the receiver plate. Touch plate on clean paper towels to remove excess ethanol.

23. Add 500 μ l of 70% ethanol to each well. Decant immediately. Touch on clean paper towels to remove excess ethanol.
24. Place plates without lids in a clean drawer, cover with a clean paper towel, and allow to dry overnight.
25. Resuspend DNA pellet in 200 μ l of T low E buffer. Seal top with sterile 96-well plate seals. Rehydrate for at least 2 days at 4°C before using. Store at -20°C.

Amplify EST inserts

26. For each 96-well plate to be amplified, prepare a PCR reaction master mix containing the following ingredients:

1000 μ l 10 \times PCR buffer
 20 μ l 100 mM dATP
 20 μ l 100 mM dGTP
 20 μ l 100 mM dCTP
 20 μ l 100 mM dTTP
 5 μ l 1 mM AEK M13F primer
 5 μ l 1 mM AEK M13R primer
 100 μ l 5 U/ μ l *Taq* DNA polymerase
 8800 μ l H₂O.

27. Label 96-well PCR plates and pipet 100 μ l of PCR reaction master mix into each well. Gently tap plates to ensure that no air bubbles are trapped at the bottom of the wells.
28. Add 1 μ l of purified EST plasmid template to each well.

Mark the donor and recipient plates at the corner near the A1 well to facilitate correct orientation during transfer of the template. It is important to watch that the pipet tips are all submerged in the PCR reaction mix when delivering the template. Missing the liquid is easier when multichannel pipets are used.

29. Perform the following thermal cycling series:

1 cycle:	30 sec	96°C (initial denaturation)
25 cycles:	30 sec	94°C (denaturation)
	30 sec	55°C (annealing)
	150 sec	72°C (extension)
1 cycle:	5 min	72°C (final extension).

After PCR, plates may be stored at 4°C while quality controls are performed.

Check PCR products for purity and yield

30. If this is the first time the template for these ESTs is being amplified, analyze 2 μ l of each PCR product on a 2% agarose gel (see Support Protocol 1) with appropriate molecular-weight markers. If amplified products from this template have been previously tested, then analyze one row of wells from each plate amplified.
31. Analyze 1 μ l of amplified products from one row of wells from each amplified plate by fluorometry (see Support Protocol 2).

Purify PCR products

32. Fill 96-well V-bottom cell culture plates with 200 μ l per well of ethanol/acetate solution.

The ethanol/acetate solution used for precipitation is less acidic (i.e., pH 6) than is typically used. In this instance, more acidic solutions produce precipitates that are harder to resuspend, without improving yield.

33. Transfer 100 μ l per well of PCR product (from step 29) into corresponding 96-well V-bottom plates and mix by pipetting a volume of 75 μ l per well four times.
34. Place the plates for 1 hr at -80°C or store overnight at -20°C .
Place plates at -20°C if they are to be left for more than 1 hr; aggressive chilling produces precipitates that are hard to resuspend.
35. Thaw the plates to reduce brittleness and melt any ice that may have formed in the wells.
36. Centrifuge for 40 min at $2600 \times g$, 4°C .
37. Aspirate the supernatant from each well using a microtiter plate washer.
Settings for the depth of aspiration by the plate washer will need to be adjusted to suit the microtiter plates used. It is advisable to leave ~ 10 to 20μ l in the bottom of the well to avoid disturbing the pellet.
38. Wash pellets with 200 μ l of 70% ethanol per well using the microtiter plate washer.
39. Centrifuge plates for 40 min at $2600 \times g$, 4°C .
40. Aspirate the supernatant from each well using the microtiter plate washer.
41. Allow the plates to dry overnight in a closed drawer.
Do not dry in a Speedvac evaporator. Desiccated PCR products are hard to resuspend.

Resuspend the PCR products

42. Add 40 μ l of $3\times$ SSC per well. Seal plates with a foil sealer, taking care to achieve a tight seal over each well.
43. Place the plates in heat-sealable bags with paper towels moistened with $3\times$ SSC and seal the bag with a heat sealer.
The high external humidity within the sealed bag helps keep the volumes in the individual wells from varying.
44. Place the bags in a 65°C incubator for 2 hr, then turn off the heat in the incubator.
Allowing the plates to cool down gradually in the incubator avoids condensation on the sealers.

Check PCR products for yield and resuspension

45. Analyze 1 μ l of resuspended PCR product from one row of wells from each plate on a 2% agarose gel (see Support Protocol 1).
Adequate precipitation and resuspension will produce very intense bands, with no material failing to leave the loading well and no smear of material from the band towards the loading well.
46. Store the plates at -20°C after resuspension.

Print PCR products

The variety of printers and pens for transferring PCR products from titer plates to slides preclude highly detailed descriptions of the process. The following steps provide a general description of the processing.

47. Preclean the print pens according to the manufacturer's specification.
48. Load the printer slide deck with poly-L-lysine-coated slides (see Support Protocol 3).

49. Thaw the plates containing the purified EST PCR products and centrifuge briefly for 2 min, at $167 \times g$ (1000 rpm in a horizontal microtiter plate rotor), 20° to 25°C, to remove condensation and droplets from the seals before opening.

50. Transfer 5 to 10 μl of the purified EST PCR products to a plate that will serve as the source of solution for the printer.

Printing with quill-type pens usually requires that the volume of fluid in the print source be sufficiently low that when the pen is lowered to the bottom of the well, it is submerged in the solution to a depth of <1 mm. This keeps the pen from carrying a large amount of fluid on the outside of the pen shaft and producing variable, large spots on the first few slides printed.

51. Run a repetitive test print on the first slide.

In this operation, the pens are loaded with the DNA solution, and then the pens serially deposit this solution on the first slide in the spotting pattern specified for the print. This test is run to check the size and shape of the specified spotting pattern and its placement on the slide. It also serves to verify that the pens are loading and spotting, and that a single loading will produce as many spots as are required to deliver material to every slide in the printer.

52. If one or more of the pens is not performing at the desired level, reclean or substitute another pen and test again. If all pens are performing, carry out the full print.

53. At the end of the print, remove slides from the printer, label with the print identifier and the slide number by writing on the edge of the slide with a diamond scribe, place in a dust-free slide box, and age for 1 week.

It is useful to etch a line that outlines the printed area of the slide onto the first slide. This serves as a guide to locate the area after the slides have been processed and the salt spots washed off.

54. Place slides, printed side face up, in a Pyrex baking dish and cover with plastic wrap. Expose slides to a 450 mJ of UV irradiation in a cross-linker.

Slides should have been aged at ambient temperature in a closed slide box for one week prior to blocking.

55. Transfer slides to a 30-slide stainless steel rack and place rack into a small glass tank.

56. Dissolve 6.0 g succinic anhydride in 325 ml of 1-methyl-2-pyrrolidinone in a glass beaker by stirring with a stir bar.

This reaction adds succinate to the lysine primary amines, removing the positive charge from the surface of the slides.

CAUTION: Nitrile gloves should be worn and work carried out in a chemical fume hood while handling 1-methyl-2-pyrrolidinone (a teratogen).

57. Add 25 ml of 1 M sodium borate buffer, pH 8.0, to the beaker. Allow the solution to mix for a few seconds, then pour rapidly into glass tank with slides.

Succinic anhydride hydrolyzes quite rapidly once the aqueous buffer solution is added. To obtain quantitative passivation of the poly-L-lysine coating, it is critical that the reactive solution be brought in contact with the slides as quickly as possible.

58. Place the glass tank on a platform shaker in a fume hood for 20 min, shaking at 70 to 90 cycles/min.

Small particulates resulting from precipitation of reaction products will be visible in the fluid.

59. While the slides are incubating on the shaker, prepare a boiling water bath to denature the DNA on the slides.
60. After the slides have incubated for 20 min, transfer into the boiling water bath. Immediately turn off the heating element after submerging the slides in the bath. Allow slides to stand in the water bath for 2 min.
61. Transfer slides into a glass tank filled with 100% ethanol and incubate for 4 min.
62. Remove the slides and centrifuge for 3 min at $167 \times g$ (1000 rpm in a horizontal microtiter plate rotor) 20° to 25°C, to dry the slides.
63. Transfer slides to a clean, dust free slide box and let stand overnight before hybridizing.

**BASIC
PROTOCOL 2**

RNA EXTRACTION AND LABELING

This protocol details the methods used to extract RNA from cells and purify the RNA by a combination of phase extraction and chromatography. The protocol also describes the process of making fluorescent cDNA representations of the message pools within the isolated total RNA pools. This is accomplished by using the pure total RNA as a substrate for reverse transcription in the presence of nucleotides derivatized with either a Cy3 or a Cy5 fluorescent tag.

Materials

- Cells harvested from tissue culture, in tissue culture, or whole frozen tissue
- Phosphate-buffered saline (PBS; see recipe)
- TRIzol reagent (Life Technologies)
- Chloroform
- 100%, 75%, 70% ethanol
- RNeasy Maxi Kit (Qiagen)
 - 50-ml Maxi spin columns with collection tubes
 - RW1 buffer
 - RPE buffer
 - DEPC-treated water (provided with kit or see *UNIT 4.1*)
- 3 M sodium acetate, pH 5.2 (*APPENDIX 2*)
- 2 mg/ml anchored oligo-dT primer (anchored; 5'-TTT TTT TTT TTT TTT TTT TTT N-3'; e.g., Genosys)
- 1 mg/ml pd(T)12-18 (Amersham Pharmacia Biotech)
- 10× low-T dNTP mix (see recipe)
- 1 mM Cy 3-dUTP or Cy 5-dUTP, store -20°C, light sensitive
- RNasin RNase inhibitor (Promega)
- Superscript II RNase H⁻ Reverse Transcriptase Kit with 5× first strand buffer and 1 M DTT (Life Technologies)
- 0.5 M EDTA, pH 8.0 (*APPENDIX 2*)
- 1 N NaOH
- 1 M Tris·Cl, pH 7.5 (*APPENDIX 2*)
- TE buffer, pH 7.5 (*APPENDIX 2*)
- 1 mg/ml C₀t-1 DNA (Life Technologies)
- 2% (w/v) agarose gel (6-cm wide × 8.5-cm long, 2-mm wide wells) in TAE buffer
- 50× TAE buffer (*APPENDIX 2*)
- Tissue homogenizer (e.g., Polytron PT1200, Brinkmann Instruments)
- 15-ml round-bottom polypropylene centrifuge tubes
- 50-ml conical polypropylene centrifuge tubes

Clinical centrifuge with horizontal rotor for 50-ml conical tubes
1.5-ml microcentrifuge tubes
Centrifugal filter units (e.g., Microcon YM-100, Amicon)
0.2-ml thin-wall PCR tube with cap
Thermal cycler
Fluorescence scanner (e.g., Storm system for gel analysis, Molecular Dynamics)
Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5A)

NOTE: Use RNase-free water (e.g., DEPC-treated water, UNIT 4.1) to make up all solutions, unless indicated otherwise.

Extract and purify RNA

- 1a. *If starting with cells harvested from tissue culture:* Wash the cell pellet twice in PBS.
- 1b. *If starting with cells from tissue culture:* Add 1 ml TRIzol per 2×10^7 cells and mix by shaking.
- 1c. *If starting with tissue:* Add 100 mg of frozen tissue directly to 4 ml TRIzol, and dissociate by homogenization with a rotating blade tissue homogenizer.
2. Add 1/5 vol of chloroform, shake for 15 sec, and let stand for 3 min. Centrifuge for 15 min at $12,000 \times g$, 4°C .
3. Decant the supernatant, place in a 15-ml polypropylene tube, and record the volume.
4. Add 1 vol of 70% ethanol to the supernatant dropwise while vortexing.

This will produce a final ethanol concentration of 35%. The ethanol should be added drop by drop and allowed to mix completely with the supernatant before more ethanol is added. If a high local concentration of ethanol is produced, the RNA in that vicinity will precipitate.
5. Add the supernatant from an extraction of 2×10^7 to 1×10^8 cells to an RNeasy Maxi column that has been placed in a 50-ml conical tube.
6. Centrifuge for 5 min at $2880 \times g$, room temperature, in a clinical centrifuge with a horizontal rotor.
7. Collect the eluate and pour back onto the top of the column and centrifuge again as in step 6.

A significant amount of RNA is not captured by the column matrix in the first pass of the RNA-containing solution through the column.
8. Discard the eluate and add 15 ml RW1 buffer to the column. Centrifuge for 5 min at $2880 \times g$, 20° to 25°C .
9. Discard eluate and add 10 ml RPE buffer to the column. Centrifuge for 5 min at $2880 \times g$, 20° to 25°C .
10. Discard eluate and add another 10 ml RPE buffer to the column. Centrifuge for 10 min at $2880 \times g$, 20° to 25°C , and discard eluate.
11. Put the column in a fresh 50-ml conical polypropylene centrifuge tube and add 1 ml DEPC-treated water from the kit to the column. Let stand for 1 min and centrifuge for 5 min at $2880 \times g$, 20° to 25°C ; do not discard the eluate.
12. Add another 1 ml DEPC-treated water to the column, let stand for 1 min, and centrifuge for 10 min at $2880 \times g$, 20° to 25°C , into the same collection tube.
13. Dispense 400- μl aliquots of the column eluate into 1.5-ml microcentrifuge tubes.

14. Add 1/10 vol of 3 M sodium acetate, pH 5.2. Then add 1 ml of 100% ethanol to each tube, vortex, and let stand for 15 min at room temperature.
15. Microcentrifuge for 15 min at $12,000 \times g$, 4°C . Wash pellet two times in 75% ethanol and store in 75% ethanol at -80°C , indefinitely.

Clean up RNA

16. Microcentrifuge for 15 min at $12,000 \times g$, 4°C . Remove supernatant, air dry, and then resuspend RNA at ~ 1 mg/ml in DEPC-treated water. Determine concentration by reading A_{260} of 1 μl of this solution in 100 μl of 50 mM NaOH.
17. Concentrate to >7 mg/ml by centrifuging at $500 \times g$ on a Microcon YM-100 filter unit, checking as necessary to determine the rate of concentration.

This step removes many residual small- to medium-sized molecules that inhibit the reverse transcription reaction in the presence of fluorescently derivatized nucleotides.

18. Determine the concentration of RNA in the concentrated sample by spectrophotometry (APPENDIX 3D). Store at -80°C .

Reverse transcribe RNA

Typically, the same results are produced with either pdT or anchored pdT. Some cell types may give more signal with anchored pdT.

- 19a. *If using an anchored oligo-dT primer:* Anneal the primer to the RNA in the following 17- μl reaction (use a 0.2-ml thin-wall PCR tube so that incubations can be carried out on a thermal cycler):

<i>Component</i>	<i>For Cy5 labeling</i>	<i>For Cy3 labeling</i>
total RNA (>7 mg/ml)	150 to 200 μg	50 to 80 μg
anchored primer (2 $\mu\text{g}/\mu\text{l}$)	1 μl	1 μl
DEPC-treated H_2O	up to 17 μl	up to 17 μl .

- 19b. *If using an oligo d(T)12-18 primer:* Anneal the primer to the RNA in the following 17- μl reaction (use a 0.2-ml thin-wall PCR tube so that incubations can be carried out on a thermal cycler):

<i>Component</i>	<i>For Cy5 labeling</i>	<i>For Cy3 labeling</i>
total RNA (>7 mg/ml)	150 to 200 μg	50 to 80 μg
dT(12-18) primer (1 $\mu\text{g}/\mu\text{l}$)	1 μl	1 μl
DEPC-treated H_2O	up to 17 μl	up to 17 μl .

The incorporation rate for Cy5-dUTP is less than that of Cy3-dUTP, so more RNA is labeled to achieve a more equivalent signal from each species.

20. Heat to 65°C for 10 min and cool on ice for 2 min.
21. Prepare a master mix with the following components (total volume, 23 μl).

8 μl 5 \times first strand buffer
 4 μl 10 \times low-T dNTP mix
 4 μl 1 mM Cy5 or Cy3 dUTP
 4 μl 0.1 M DTT
 1 μl 30 U/ μl RNasin
 2 μl 200 U/ μl Superscript II.

Superscript polymerase is very sensitive to denaturation at air/liquid interfaces, so be very careful to suppress foaming in all handling of this reaction.

22. Add 23 μl of reaction mixture containing either Cy5-dUTP or Cy3-dUTP nucleotides to each sample, mix well by pipetting, and centrifuge briefly to concentrate in the bottom of the tube.
23. Incubate 30 min at 42°C, then add 2 μl more Superscript II. Make sure the enzyme is well mixed in the reaction volume and incubate for 30 to 60 min at 42°C.
24. Add 5 μl of 0.5 M EDTA, pH 8.0, to stop the reaction.

Be sure to stop the reaction with EDTA before adding NaOH, since nucleic acids precipitate in alkaline magnesium solutions.

25. Add 10 μl of 1 N NaOH, then incubate for 60 min at 65°C to hydrolyze residual RNA. Cool to room temperature.

The purity of the sodium hydroxide solution used in this step is crucial. Slight contamination or long storage in a glass vessel can produce a solution that will degrade the Cy5 dye molecule, turning the solution yellow. Some researchers achieve better results by reducing the time of hydrolysis to 30 min.

26. Neutralize by adding 25 μl of 1 M Tris-Cl, pH 7.5.
27. Desalt the labeled cDNA by adding 400 μl of TE buffer, pH 7.5, and 20 μg of human C₀t-1 DNA in a Microcon YM-100 cartridge. Pipet to mix, spin for 10 min at 500 \times g, room temperature.
28. Wash again by adding 200 μl TE buffer, pH 7.5, and concentrating in a Microcon YM-100 cartridge to about 20 to 30 μl (~8 to 10 min at 500 \times g).

Alternatively, a smaller-pore Microcon YM-30 can be used to speed the concentration step. In this case, centrifuge the first wash for ~4.5 min at 16,000 \times g and the second (200- μl wash) for ~2.5 min at 16,000 \times g.

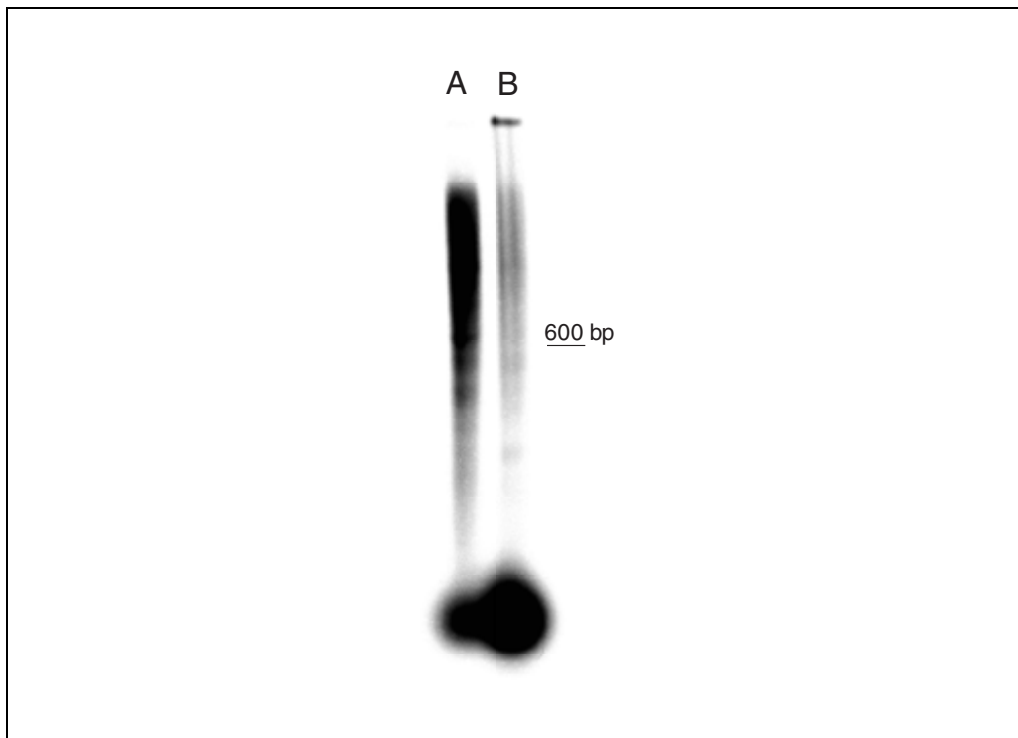


Figure 22.3.1 Fluorescence scan of a 2 Cy5 labeled cDNAs electrophoresed on a 2% agarose gel.

29. Recover by inverting the concentrator over a clean collection tube and spinning for 3 min at $500 \times g$, room temperature.

In some cases, the Cy5-labeled cDNA will form a gelatinous blue precipitate that is recovered in the concentrated volume. The presence of this material signals the presence of contaminants. The more extreme the contamination, the greater the fraction of cDNA that will be captured in this gel. Even if heat solubilized, this material tends to produce uniform nonspecific binding to the DNA targets.

When concentrating by centrifugal filtration, the time required to achieve the desired final volume is variable. Overly long spins can remove nearly all the water from the solution being filtered. When fluor-tagged nucleic acids are concentrated onto the filter in this fashion, they are very hard to remove, so it is necessary to approach the desired volume by conservative approximations of the required spin times. If control of volumes proves difficult, the final concentration can be achieved by evaporating liquid in the Speedvac evaporator. Vacuum evaporation, if not carried to dryness, does not degrade the performance of the labeled cDNA.

30. Take a 2- to 3- μ l aliquot of the Cy5-labeled cDNA for analysis, leaving 18 to 28 μ l for hybridization.

31. Run this probe on a 2% agarose gel (e.g. UNIT 2.5A) in TAE buffer.

For maximal sensitivity when running samples on a gel for fluor analysis, use loading buffer with minimal dye and do not add ethidium bromide to the gel or running buffer.

32. Scan the gel on a Molecular Dynamics Storm fluorescence scanner (setting: red fluorescence, 200- μ m resolution, 1000 V on PMT).

Successful labeling produces a dense smear of probe from 400 bp to >1000 bp, with little pile-up of low-molecular-weight transcripts (as in Fig. 22.3.1, Lane A). Weak labeling and significant levels of low-molecular-weight material indicates a poor labeling (as in Fig. 22.3.1, Lane B). A fraction of the observed low-molecular-weight material is unincorporated fluor nucleotide.

BASIC PROTOCOL 3

HYBRIDIZATION AND DATA EXTRACTION

This protocol describes the conditions for hybridizing the fluor-tagged cDNA representations of the mRNA pools of samples (see Basic Protocol 2) to the EST PCR products immobilized on the glass microarrays (see Basic Protocol 1). The format for these hybridizations is simultaneous hybridization of both labeled species to a single microarray. The amount of immobilized cDNA on the slide is in excess of the amount of labeled sample that can hybridize to it, so that the amounts of labeled cDNA that do hybridize to any given immobilized cDNA are proportional to their original relative abundance in the cellular message pool. By measuring the ratio of fluorescent intensities at each immobilized cDNA spot, one obtains a measure of relative levels of messages in one sample pool versus the other. If a number of samples are measured against a common reference, the relative amounts of the messages across all of the samples can be compared for similarities and differences.

Materials

- Glass microarrays (see Basic Protocol 1)
- Cy3- and Cy5-labeled cDNAs (see Basic Protocol 2)
- DEPC-treated water (UNIT 4.1)
- 8 mg/ml poly(dA)₄₀₋₆₀ (Amersham Pharmacia Biotech)
- 4 mg/ml yeast tRNA (see recipe)
- 10 mg/ml human C₀t-1 DNA (see recipe)
- 20 \times SSC (APPENDIX 2)

50× Denhardt's solution (APPENDIX 2)
 10% SDS
 0.5× SSC/0.01% SDS wash buffer (see recipe)
 0.06× SSC wash buffer (see recipe)
 0.2-ml thin-wall PCR tubes
 Thermal cycler
 24-mm × 50-mm glass cover slips
 Microarray hybridization chamber
 65°C water bath
 Microarray scanner
 Image analysis software

NOTE: Use RNase-free water (e.g., DEPC-treated water, UNIT 4.1) to make up all solutions, unless indicated otherwise.

Hybridize fluorescent cDNA to slide

1. Determine the volume of hybridization solution required.

The rule of thumb is to use 0.033 μl for each mm^2 of slide surface area covered by the coverslip used to cover the array. An array covered by a 24-mm by 50-mm coverslip will require 40 μl of hybridization solution.

The volume of the hybridization solution is critical. When too little solution is used, it is difficult to seat the coverslip without introducing air bubbles over some portion of the arrayed ESTs, and the coverslip will not sit at a uniform distance from the slide. If the coverslip is bowed toward the slide in the center, there will be less labeled cDNA in that area and hybridization will be nonuniform. When too much volume is applied, the coverslip will move easily during handling, leading to misplacement relative to the arrayed ESTs, and nonhybridization in some areas of the array.

2. For a 40- μl hybridization, pool the Cy3- and Cy5-labeled cDNAs into a single 0.2-ml thin-wall PCR tube and adjust the volume to 30 μl by either adding DEPC-treated water or removing water in a Speedvac evaporator.

If using a vacuum device to remove water, do not use high heat or heat lamps to accelerate evaporation because this could degrade the fluorescent dyes.

3. For a 40- μl hybridization combine the following components:

	high sample blocking	high array blocking
Cy5 + Cy3 probe	30 μl	28 μl
8 mg/ml poly(dA)	1 μl	2 μl
4 mg/ml yeast tRNA	1 μl	2 μl
10 mg/ml C_0t-1 DNA	1 μl	0 μl
20× SSC	6 μl	6 μl
50× Denhardt's solution	1 μl (optional)	2 μl
Total volume	40 μl	40 μl

Arrays and samples can vary somewhat, making it necessary to vary the composition of the hybridization cocktail. In cases where there is residual hybridization to control repeat DNA samples on the array, more C_0t-1 DNA can be used, as in the high sample blocking formulation. When there is diffuse background or a general haze on all of the array elements, more of the nonspecific blocker components can be used, as in the high array blocking formulation. The authors generally try the high sample blocking formulation first.

4. Mix the components well by pipetting, heat at 98°C for 2 min in a thermal cycler, cool quickly to 25°C, and add 0.6 μl of 10% SDS.

5. Centrifuge for 5 min at $16,000 \times g$, 20° to 25°C .

The fluor-labeled cDNAs have a tendency to form small, very fluorescent, aggregates that result in bright, punctate background on the array slide. Hard centrifugation will pellet these aggregates, allowing one to avoid introducing them to the array.

6. Apply the labeled cDNA to a 24-mm \times 50-mm glass coverslip and then touch with the inverted microarray.

Applying the hybridization mix to the array and coverslipping it is an operation that requires some dexterity to get the positioning of the coverslip and the exclusion of air bubbles just right. It is helpful to practice this operation with buffer and plain slides before attempting actual samples. The hybridization solution is added to the coverslip first, since some aggregates of fluor remain in the solution and will bind to the first surface they touch.

7. Place the slide in a microarray hybridization chamber, add 5 μl of $3\times$ SSC in the reservoir, if the chamber provides one, or else at the scribed end of the slide, and seal the chamber. Submerge the chamber in a 65°C water bath and allow the slide to hybridize for 16 to 20 hr.

There are a wide variety of commercial hybridization chambers. It is worthwhile to prepare a mock hybridization with a blank slide, load it in the chamber, and incubate it to test for leaks or drying of the hybridization fluid, either of which will cause severe fluorescent noise on the array.

Wash off unbound fluorescent cDNA

8. Remove the hybridization chamber from the water bath, cool, and carefully dry off. Unseal the chamber and remove the slide.

As there may be negative pressure in the chamber after cooling, it is necessary to remove water from around the seals so that it is not pulled into the chamber and onto the slide when the seals are loosened.

9. Place the slide, with the coverslip still affixed, into a Coplin jar filled with $0.5\times$ SSC/ 0.01% SDS wash buffer. Allow the coverslip to fall from the slide and then remove the coverslip from the jar with forceps. Allow the slide to wash for 2 to 5 min.
10. Transfer the slide to a fresh Coplin jar filled with $0.06\times$ SSC. Allow the slide to wash for 2 to 5 min.

The sequence of washes may need to be adjusted to allow for more aggressive noise removal, depending on the source of the sample RNA. Useful variations are to add a first wash that is $0.5\times$ SSC/ 0.1% SDS or to repeat the normal first wash twice.

11. Transfer the slide to a slide rack and centrifuge 3 min at low speed— $167 \times g$ (700 to 1000 rpm) in a clinical centrifuge equipped with a horizontal rotor for microtiter plates.

If the slide is simply air dried, it frequently acquires a fluorescent haze. Centrifuging off the liquids results in a lower fluorescent background. As the rate of drying can be quite rapid, it is suggested that the slide be placed in the centrifuge immediately upon removal from the Coplin jar.

Acquire a fluorescent image of the slide and extract the signal data

The particulars of adjusting a microarray imaging device will vary considerably depending upon the type used and the manufacturer. Similarly, a number of methods for identifying the signal in the image, adjusting for the level of background, and normalizing the values between the two channels can be used. This section is intended only to provide some guidelines to the general considerations of imaging.

- Adjust the photomultiplier voltage and laser power so that the brightest signals produce slightly less than the maximum possible reading (i.e., 65,535, if a 16-bit scale), and the background values (between the arrayed spots) are consistently somewhat above zero, and then collect the data for the entire image.

The aim in imaging the fluorescent signals is to use as much of the linear range of detection of the instrument as possible, and to scale both signals so that they occupy roughly the same range. Setting the device so that the assay background level is at the base of the range of values insures that the minimal detectable signal can be evaluated, and helps in the determination of whether one of the two signals is considerably stronger than the other.

- Load the captured image into the data extraction software and examine images for overall quality of hybridization, noting the uniformity and level of backgrounds, the level and distribution of signals, the relative strengths of the signals in each channel, and the comparability of signals for most of the genes.

A very good characterization of the performance of the system can be carried out with a series of four hybridizations. These basically ask whether the system accurately judges two identical samples to be identical, and checks the reproducibility of the detected differences when two samples are different. For a sample of this type of data, see Troubleshooting.

AGAROSE GEL ELECTROPHORESIS OF ESTs

Gel imaging allows a rough quantitation of product while giving an excellent characterization of the product. Band size, as well as the number of bands observed in the PCR products, contribute to understanding the final results of the hybridization. The use of gel-well formats suitable for loading from 96-well plates, as well as programmable pipettors, make this form of analysis feasible on a large scale.

Materials

2% (w/v) agarose gel in 1× TAE buffer (see *UNIT 2.5A*)

50× TAE buffer (*APPENDIX 2*)

Loading buffer (see recipe)

100-bp size standards (see recipe)

Electrophoresis apparatus with capacity for four 50-well combs, (e.g., Owl Scientific)

Disposable microtiter mixing trays (e.g., Becton Dickinson)

Programmable, 12-channel pipettor with disposable tips (e.g., Matrix Technologies)

Electrophoresis power supply

Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5A*)

Perform gel electrophoresis

- Cast a 2% agarose gel in 1× TAE buffer with 4 combs (i.e., 50-tooth) and submerge in an electrophoresis apparatus with sufficient 1× TAE buffer to just cover the surface of the gel (see *UNIT 2.5A*).
- Prepare a reservoir of loading buffer, using 12 wells of a microtiter plate.
- Program 12-channel pipettor to sequentially carry out the following steps:
 - fill with 2 μ l
 - fill with 1 μ l
 - fill with 2 μ l
 - mix a volume of 5 μ l five times
 - expel 5 μ l.

SUPPORT PROTOCOL 1

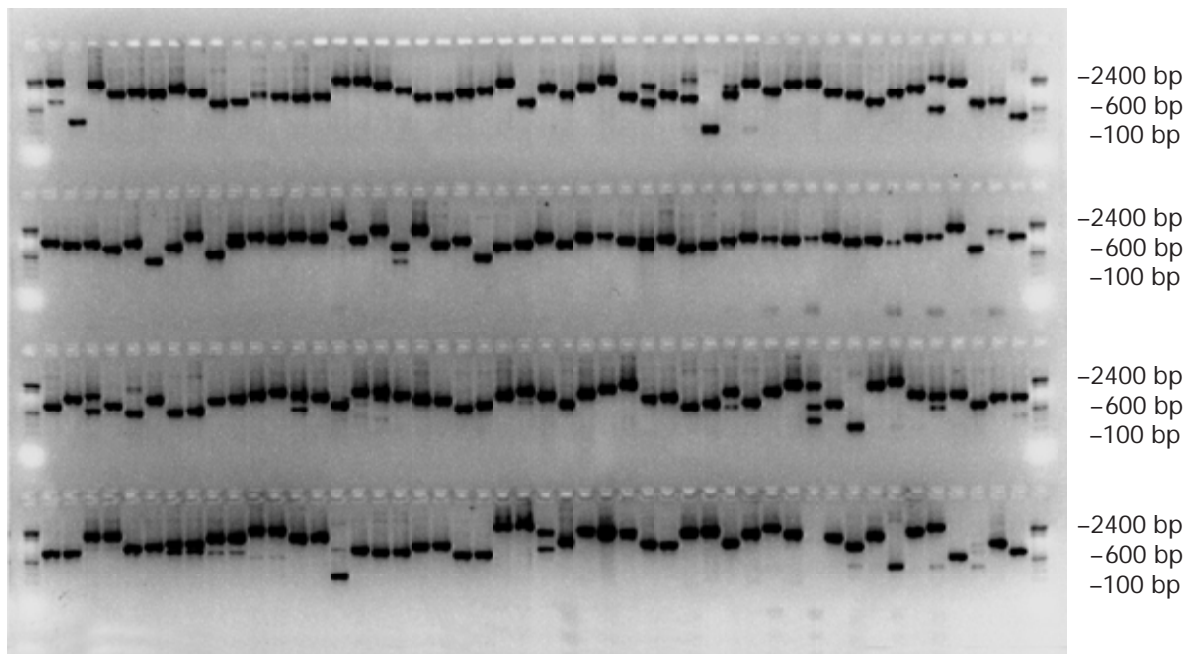


Figure 22.3.2 Ethidium bromide staining pattern of EST PCR products electrophoresed on a 2% agarose gel.

4. Place 12 disposable tips on the pipettor.
5. Load 2 μ l of PCR product from wells A1 to A12 of the PCR plate.
6. Load 1 μ l of air.
7. Load 2 μ l loading buffer from the reservoir.
8. Place tips in clean wells of disposable microtiter mixing tray and allow pipettor to mix the sample and loading dye.
9. Place the pipettor in a 50-well row so that the tip containing the PCR product from well A1 is in the second well of the row, and the other tips are in every other succeeding well.
10. Repeat the process (changing tips each time), loading PCR plate row B starting in the third well, interleaved with the A row, the C row starting at well 26, and the D row at well 27, interleaved with the C row.
11. Place 5 μ l of 100-bp size standard in wells 1 and 50.
12. Repeat this process, loading samples from rows E, F, G, and H in the second, 50-well row of gel wells, loading samples from two 96-well PCR plates per gel, or single-row samples from 16 PCR plates.

To reduce diffusion and mixing, apply voltage to the gel for a minute between loading each well strip. This will cause the DNA to enter the gel, and reduce band spreading and sample loss.

13. Apply voltage to the gel and run until the bromphenol blue (faster band) has nearly migrated to the next set of wells.

For a gel that is 14 cm in the running dimension, and 3 cm between each row of wells, apply 200 V for 15 min.

14. Take a digital photo of gel and store image for future reference.

The gels should show bands of fairly uniform brightness distributed in size between 600 to 2000 bp as in Figure 22.3.2. Further computer analysis of such images can be carried out with image-analysis packages to provide a list of the number and size of bands. Ideally this information can be made available during analysis of the data from hybridizations involving these PCR products.

FLUOROMETRIC DETERMINATION OF DNA CONCENTRATION

While it would be ideal to be able to quantify exactly each EST PCR product and spot all DNA species at equivalent concentrations, it is impractical for most laboratories to do so when thousands of ESTs must be prepared. Fortunately, it is possible to use a strategy where excess DNA is spotted, so that the exact quantities used do not produce much variation in the observed results. When using this strategy, it is necessary to track the average productivity of the PCR reactions. Fluorometry provides a simple way to obtain an approximate concentration of the double-stranded PCR product in the PCR reaction mix.

Materials

FluoReporter Blue dsDNA Quantitation Kit (Molecular Probes)
Fluor buffer (see recipe)
PCR product (see Basic Protocol 3)
TE buffer, pH 8 (APPENDIX 2)
50, 100, 250, and 500 µg/ml dsDNA reference standards (e.g., see recipe)
96-well plates for fluorescent detection (e.g., Dynex)
12-channel multipipettor
Fluorometer (e.g., PE Biosystems)
Computer equipped with Microsoft Excel software

Quantitate dsDNA

1. Label 96-well plates for fluorescence assay. With a 12-channel multipipettor, add 200 µl of fluor buffer to each well.
2. Add 1 µl PCR product from each well in a row of a PCR plate to a row of the fluorometry plate.

Samples can be added to rows A through G of the fluorometry plate.

3. In the final row of the fluorometry plate, add 1 µl TE buffer to first well and add 1 µl of each of the series of dsDNA standards—50, 100, 250, and 500 µg/ml—dsDNA to the subsequent wells. Repeat this series twice in the final row.
4. Set the fluorometer for excitation at 346 nm and emission at 460 nm. Adjust as necessary to read the plate.

If the fluorometer does not support automated analysis, export the data table to Excel.

5. Test to see that the response for the standards is linear and reproducible from the range of 0 to 500 µg/ml of dsDNA.

6. Calculate the concentration of dsDNA in the PCR reactions using the following equation after subtracting the average 0 µg/ml value from all other sample and control values:

$$[\text{dsDNA } (\mu\text{g/ml})] = [(\text{PCR sample value})/(\text{average } 100 \mu\text{g/ml value})] \times 100$$

Constantly tracking the yields of the PCRs makes it possible to rapidly detect many ways in which PCR can fail or perform poorly. This assay can also be applied after precipitation and resuspension of the PCR products to monitor overall recovery of product.

COATING SLIDES WITH POLY-L-LYSINE

Slides coated with poly-L-lysine have a surface that is both hydrophobic and positively charged. The hydrophobic character of the surface minimizes spreading of the printed spots, and the charge appears to help position the DNA on the surface in a way that makes cross-linking more efficient.

Materials

Cleaning solution (see recipe)

Poly-L-lysine solution (see recipe)

Gold Seal microscope slides (Becton Dickinson)

50-slide stainless steel rack and 50-slide glass tank (Wheaton)

25-slide plastic rack and 25-slide plastic box (Shandon Lipshaw)

Plastic slide box with no paper or cork liners (e.g., PGC Scientific)

Coat slides

1. Place slides into 50-slide racks and place racks in glass tanks with 500 ml of cleaning solution.

Gold Seal slides are highly recommended, as they have been found to have consistently low levels of autofluorescence.

It is important to wear powder-free gloves when handling the slides. Change gloves frequently, as random contact with skin and surfaces transfers grease to the gloves.

2. Place tanks on platform shaker for 2 hr at 60 rpm.
3. Pour out cleaning solution and wash in water for 3 min. Repeat wash four times.
4. Transfer slides to 25-slide plastic racks and place into small plastic boxes for coating.
5. Submerge slides in 200 ml poly-L-lysine solution per box and shake for 1 hr at 60 rpm.
6. Rinse slides three times with water and submerge slides in water for 1 min.
7. Centrifuge for 2 min at 400 × g and dry slide boxes used for coating.
8. Place slides back into slide box used for coating and let stand overnight before transferring to new slide box for storage.

This allows the coating to dry before handling.

9. Allow slides to age for two weeks on the bench, in a new slide box, before printing on them.

The coating dries slowly, becoming more hydrophobic with time.

Slide boxes used for long-term storage should be plastic and free of cork lining. The glue used to affix the cork will leach out over time and give slides stored in these types of boxes a greasy film that has a high degree of autofluorescence. Clean all glassware and racks used for slide cleaning and coating with highly purified water only. Do not use detergent.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps, unless otherwise noted. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Carbenicillin stock solution, 100 mg/ml

1 g carbenicillin (Life Technologies)
10 ml sterile water
Sterile filter with a 0.2- μ m filter
Store frozen at -20°C up to 2 months

Cleaning solution

400 ml H_2O
600 ml 100% ethanol
100 g NaOH
Dissolve NaOH in water. Add ethanol and stir until the solution clears. If the solution does not clear, add water until it does. Store up to 24 hr at room temperature.

dsDNA reference standards

50 $\mu\text{g/ml}$:
90 μl TE buffer (APPENDIX 2)
10 μl 0.5 mg/ml dsDNA (Life Technologies)

100 $\mu\text{g/ml}$:
80 μl TE buffer (APPENDIX 2)
10 μl 0.5 mg/ml dsDNA (Life Technologies)

250 $\mu\text{g/ml}$:
50 μl TE buffer (APPENDIX 2)
50 μl 0.5 mg/ml dsDNA (Life Technologies)

500 $\mu\text{g/ml}$:
0 μl TE buffer (APPENDIX 2)
100 μl 0.5 mg/ml dsDNA (Life Technologies)

It is good practice to check both the integrity (i.e., on a agarose gel) and the concentration (i.e., absorbance) of the standard before use. The reference lambda dsDNA is available from Life Technologies.

Ethanol/acetate solution

950 ml 100% ethanol
50 ml 3 M sodium acetate, pH 6.0 (see recipe)
 H_2O to 1000 ml
Store up to 2 weeks at room temperature

Fluor buffer

25 μl Hoechst 33258 solution (from FluoReporter Blue kit; Molecular Probes)
10 ml TNE buffer (from FluoReporter Blue kit; Molecular Probes)
10 ml water

Hoechst 33258 solution contains the dye at an unspecified concentration in a 1:4 mixture of DMSO: H_2O .

TNE Buffer is 10 mM Tris·Cl (pH 7.4), 2 M NaCl, 1 mM EDTA.

Human $\text{C}_0\text{t-1}$ DNA, 10 mg/ml

Add 925 μl of 100% ethanol and 75 μl of 3 M sodium acetate (pH 5.2) to 500 μl of 1 $\mu\text{g}/\mu\text{l}$ human $\text{C}_0\text{t-1}$ DNA. Centrifuge at $14,000 \times g$. Aspirate supernatant and allow pellet to air dry for 5 min. Resuspend the pellet in 50 μl DEPC-treated water (UNIT 4.1). Store up to 6 months at -20°C .

Low-T dNTP mix, 10×

25 μ l 100 mM dGTP (0.5 mM in 1 \times)
25 μ l 100 mM dATP (0.5 mM in 1 \times)
25 μ l 100 mM dCTP (0.5 mM in 1 \times)
10 μ l 100 mM dTTP (0.2 mM in 1 \times)
415 μ l DEPC-treated H₂O (*UNIT 4.1*)
Total volume, 500 μ l
Store up to 2 months at -20°C

Loading buffer

4.0 ml glycerol (enzyme grade)
0.9 ml DEPC-treated H₂O (*UNIT 4.1*)
0.1 ml 0.25% (w/v) xylene cyanol FF/0.25% (w/v) bromphenol blue
Store up to 1 month at room temperature

Phosphate-buffered saline (PBS)

8.00 g/liter NaCl
0.20 g/liter KCl
1.44 g/liter Na₂HPO₄ (anhydrous)
0.24 g/liter KH₂PO₄ (anhydrous)
Bring components to 1 liter volume with H₂O
Autoclave 20 min
Cool to room temperature
Pass through a 0.2- μ m filter
Store up to 6 months at room temperature

Poly-L-lysine solution

35 ml poly-L-lysine (0.1% w/v; Sigma)
35 ml PBS (see recipe)
280 ml H₂O
Store up to 24 hr at room temperature

Size standards, 100-bp

50 μ l 1 mg/ml DNA ladder (Life Technologies)
5 μ l 1 M Tris·Cl, pH 8.0 (*APPENDIX 2*)
5 μ l 0.5 M EDTA, pH 8.0 (*APPENDIX 2*)
440 μ l loading buffer (see recipe)
Store up to 1 month at 4°C

Sodium acetate, 3 M (pH 6.0)

Dissolve 408.24 g/liter sodium acetate trihydrate to prepare 3 M sodium acetate. Prepare 3 M acetic acid by diluting 172.4 ml glacial acetic acid to 1 liter with water. Titrate the pH of 3 M sodium acetate solution to pH 6.0 with the 3 M acetic acid solution. Filter sterilize using a 0.2- μ m filter. Store up to 6 months at room temperature.

Sodium borate, 1 M (pH 8.0)

Dissolve 61.83 g boric acid in 900 ml DEPC-treated water (*UNIT 4.1*). Adjust the pH to 8.0 with 1 N NaOH. Bring volume up to 1 liter. Sterilize with a 0.2- μ m filter and store up to 6 months at room temperature.

T low E buffer

10 ml 1 M Tris·Cl, pH 8.0 (*APPENDIX 2*)
0.2 ml 0.5 M EDTA, pH 8.0 (*APPENDIX 2*)
900 ml DEPC-treated H₂O (*UNIT 4.1*)
Autoclave and store up to 6 months at room temperature.

Wash buffer, 0.5× SSC/0.01% (v/v) SDS

Add 25 ml of 20× SSC to 974 ml DEPC water. Filter sterilize on a 0.5- μ m filter device. Add 1 ml of 10% SDS, and mix well. Store up to 2 months at room temperature.

Wash buffer (0.06× SSC)

Add 3 ml of 20× SSC (*APPENDIX 2*) to 997 ml DEPC-treated water (*UNIT 4.1*). Filter sterilize on a 0.5- μ m filter device. Store up to 2 months at room temperature.

Yeast tRNA, 4 μ g/ml

1. Resuspend yeast tRNA at 10 mg/ml (based on the supplier's quantitation of the RNA) in DEPC-treated water (*UNIT 4.1*) in a 1.5-ml polypropylene conical centrifuge tube.
2. Add 0.5 vol of buffered phenol (*UNIT 2.1A*), then vortex.
3. Add 0.5 vol chloroform, then vortex again.
4. Centrifuge 5 min at 10,000 \times g. Transfer aqueous layer to a new 1.5-ml polypropylene conical centrifuge tube.
5. Add 1 vol chloroform, and vortex. Centrifuge 5 min at 10,000 \times g.
6. Repeat chloroform extraction.
7. Transfer aqueous layer to a new 1.5-ml polypropylene conical centrifuge tube. Add 0.1 vol of 3 M sodium acetate, pH 5.2 (*APPENDIX 2*). Add 2 vol 100% ethanol.
8. Centrifuge 5 min at 10,000 \times g. Aspirate supernatant, then add 1 vol of 70% ethanol.
9. Centrifuge 5 min at 10,000 \times g. Aspirate supernatant again and allow pellet to dry.
10. Resuspend in DEPC-treated water (*UNIT 4.1*) at the original volume.
11. Determine the RNA concentration by spectrometry (*APPENDIX 3D*).
12. Dilute to 4 mg/ml and store frozen at -20°C .

COMMENTARY

Background Information

The possibility of examining the expression patterns of many genes simultaneously has been more and more enthusiastically pursued as the sequences and clones of a greater fraction of genes from model organisms have become available. A large part of the enthusiasm derives from the growing recognition of the extraordinary levels of integration and interaction between genes in all cell systems. It is clear that the same cues can provoke a wide variety of cellular responses, dependent in large part upon the proteins currently expressed in the cell being examined. By observing changes in a particular gene's expression against the backdrop of the patterns of change of other genes, it is possible to ask contextual questions about cellular function. Some of the simple but intriguing questions of this class that can be asked revolve around longstanding, basic issues. These would include questions about what activities are jointly regulated, what chains of events are involved in normal processes such

as cell division, and what expression patterns are correlated with particular processes or pathologies.

Most of the large-scale expression profiling done today is carried out either on arrays of short oligonucleotides or on arrays of ESTs. A good single-source comparative review of the strengths and weakness of the techniques, as well as example applications of both can be found in a 1999 supplement of *Nature Genetics* (Phimister, 1999; also see overview in *UNIT 22.1*). Both of these technologies are still in their early phases, and it is reasonable to expect that the methodologies for obtaining and analyzing gene expression profiles will change at a very rapid pace.

Critical Parameters

Fluorescent noise

Assays based on fluorescent staining are vulnerable to many environmental sources of noise. Nearly all dust derived from paper or

cloth fibers is intensely fluorescent across the visible spectrum. Many organic molecules are fluorescent in the spectral regions used for this type of study. It is necessary to keep exposure to dust to a minimum and to carefully monitor the purity of the reagents used to fabricate microarrays. All sources of grease in the preparation of slides must be carefully avoided, since greasy films will bind fluorescent molecules nonspecifically. For those steps where the array slides are handled, all glassware should be carefully cleaned and soaked in 1 N nitric acid before initial use, and then separately shelved and used only for arrays. Water should minimally be from either a glass still or prepared by reverse osmosis and then run through deionizing tanks and charcoal filtration tanks.

RNA handling

The routine precautions required for successful lab handling of RNA should be observed (see Chapter 4 introduction). For all DNA and RNA manipulations in these protocols use DEPC-treated water unless otherwise specified.

RNA purity

The quality of the RNA coming into the labeling will have a marked effect on the quality of the labeling and hybridization. RNA preparations judged good by standard molecular biology criteria often produce poor results. Typical problems include disperse, fine-grain noise over the entire hybridized surface and nonspecific binding of fluorescent molecules to the zones of DNA immobilization on the slide. These problems seem likely to have some roots in contaminating carbohydrate, and, as would be expected with carbohydrate, the problems are exacerbated by ethanol precipitation before and after labeling. Very impure preparations will frequently produce visible aggregates if precipitated after labeling, which are essentially resistant to solubilization. It is well known that nucleic acids form strong aggregates with carbohydrate when either dried together or when coprecipitated. This interaction is the basis for nucleic acid immobilization onto chromatography supports such as cellulose. To minimize this sort of problem, the authors recommend preparative procedures that use few or preferably no ethanol precipitations during RNA preparation and labeling.

Purity also exerts strong effects on the efficiency of labeling. Reverse transcription of RNA in the presence of nucleotides derivatized with fluorescent molecules is not an efficient

process at best, and is further impaired by impurities in the incoming template. The typical problems observed are total failure of the labeling or production of many very short reverse transcripts. The purification scheme described in this Basic Protocol 2 is designed to minimize carbohydrate and lipid carryover. It is a modification of the authors' previous method combining chromatography and phase extraction, kindly suggested by Dr. Alvydas Mikulskis of New England Nuclear.

Validation and verification

A very important step in qualifying the reliability of an array assay system is to validate experimental results. An early experiment with two samples where there is an expectation that some of the genes on the array will be differentially expressed is strongly recommended. The expected differences should be observed, and verified by northern analysis with probes made from the EST segments immobilized as reporters. Additional observed differences beyond the initial expectations should also be observed. Testing should be carried out over a range of signal strengths and a range of differential expression.

Clones and informatics

The ability to collect and manage numerous human EST clones inevitably adds clone verification and database requirements that are not easily met by public databases and repositories. It is important that the complications associated with ESTs be understood before array experiments are designed. The major obstacles in human EST usage arise from three sources. The first problem is that the categorization of the structures of human genes is not complete. Human EST sequences and the clones from which these sequences are derived are continually being deposited in publicly available repositories. The deposited sequences are then examined for similarity to each other and to known genes, and are clustered into groups of sequences thought to represent multiple examples of a single type of transcript. Organizations such as NCBI carry out this operation on a periodic basis, reforming the clusters to account for new EST sequences and name assignments and make the results publicly available (<http://www.ncbi.nlm.nih.gov/UniGene/Hs.Home.html>) as part of their UniGene project. The rate of change in assignment of an EST to a cluster remains fairly high, so that the assigned identities of many ESTs fluctuate. It is thus necessary to be able to update identity

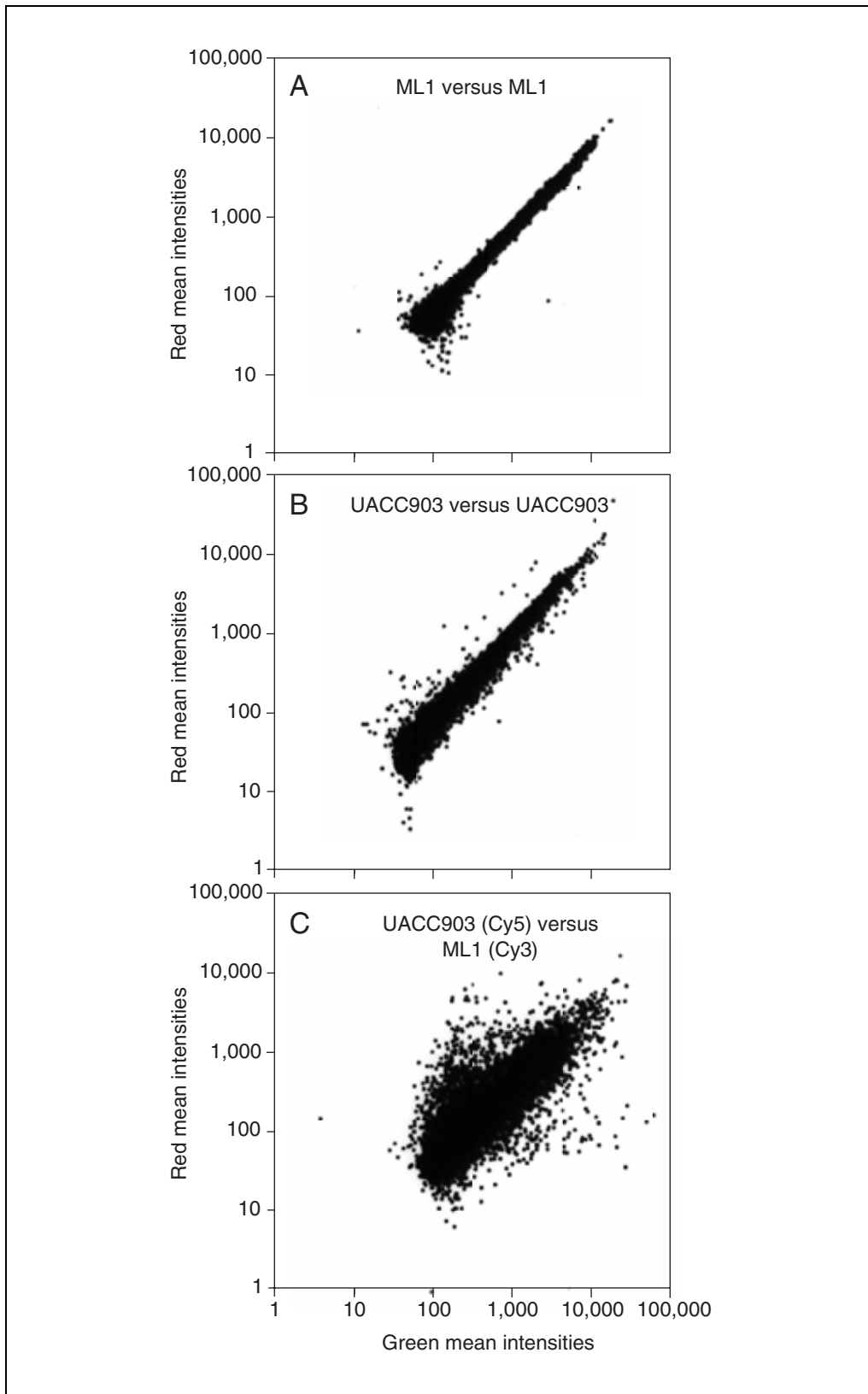


Figure 22.3.3 Scatter plots of Cy3 (green) and Cy5 (red) mean intensities observed in three hybridizations. The labeled cDNAs in the hybridizations were made from: (A) cell line ML1, Cy3 and Cy5, (B) cell line UACC903, Cy3 and Cy5, and (C) cell line UACC903, Cy5, and cell line ML1, Cy3.

and cross-referenced data, which depends on that identity, for any EST inventory.

The second obstacle derives from the fact that it is not currently possible to obtain such data for lists of ESTs from the public databases. The databases are set up to provide answers to queries on a gene by gene or EST by EST basis. Efficiently answering questions such as “What are the current chromosomal position assignments of the ESTs in a particular set?” or “Do any of the ESTs in my inventory correspond to a particular UniGene cluster?” requires a local relational database. A frequently updated minimal relational database for human ESTs based on UniGene, that can be set up and run on an

ordinary desktop computer, can be obtained to serve this purpose (<http://www.nhgri.nih.gov/DIR/LCG/15K/DATA/>).

The final serious difficulty is that many public and private EST inventories have a relatively high rate of inaccuracy. For instance, IMAGE clones available from public distributors have an ~20% chance of either not having the 3' sequence ascribed to them in the public databases or having multiple EST clones in a single well. For this reason, a number of EST suppliers are now supplying clone sets that have been clone purified and sequence verified.

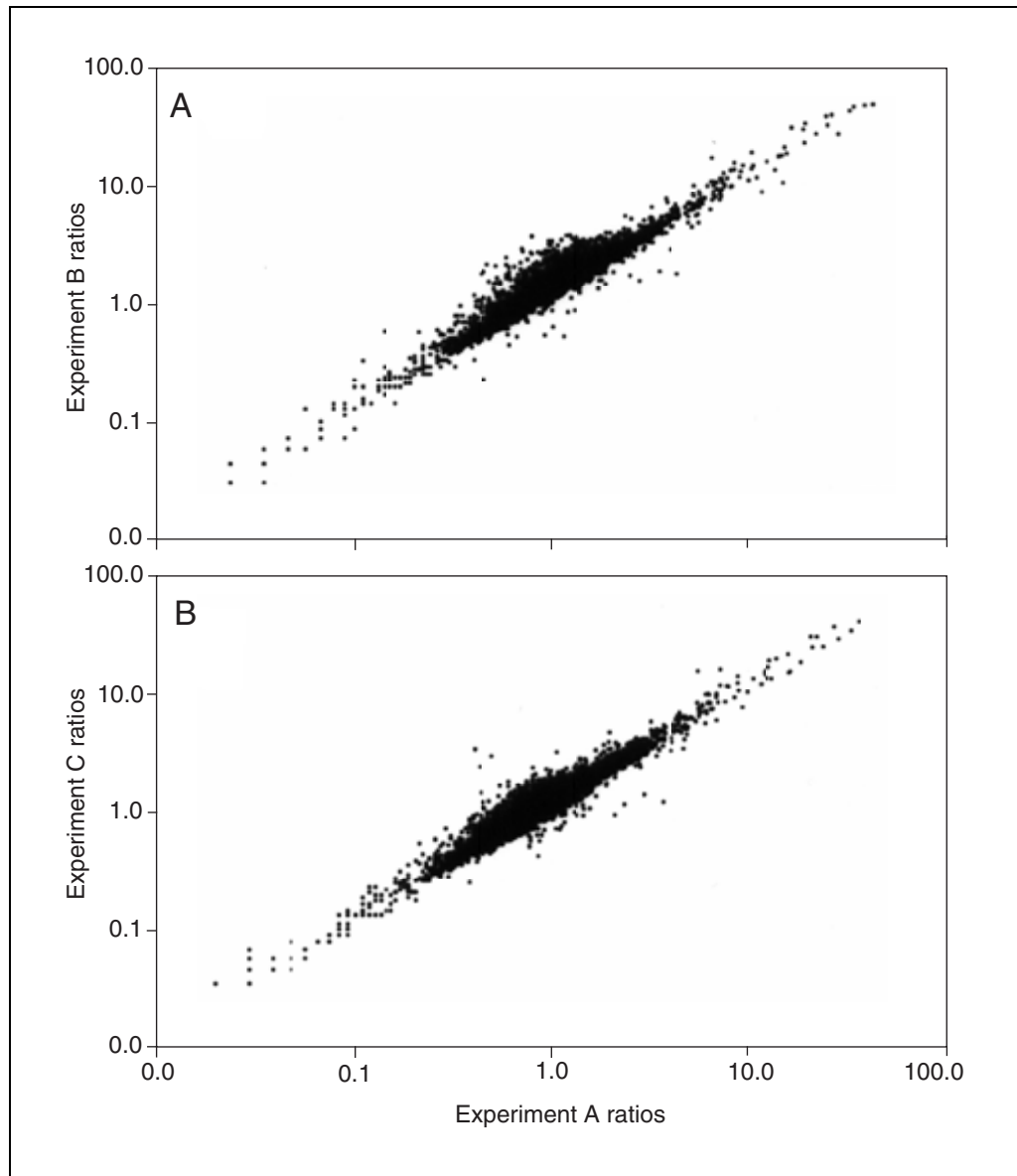


Figure 22.3.4 Scatter plots comparing the expression ratios of genes in cell line UACC903 and ML1 in three separate experiments. In panel **A**, the ratios in experiment **A** are compared to the ratios in experiment **B**. In panel **B**, the ratios in experiment **A** are compared to the ratios in experiment **C**.

Troubleshooting

One of the most illuminating diagnostics for array performance is a hybridization that reveals how tightly an array estimates that two samples from the same source produce equivalent signals. An example of a panel of three experiments that can rigorously test this capability is shown in Figure 22.3.3. Panels A and B of Figure 22.3.3 are scatter plots of the normalized mean intensities of readings from 6782 array detectors, when a single sample source is used to generate both the Cy3 and Cy5 labeled cDNA hybridized to the array. Panel C is a scatter plot of the normalized mean intensities when the two very different sample types, a melanoma cell line and a myeloid cell line used in A and B, respectively, are the sources for the labeled cDNAs in the hybridization. From such plots, one can obtain an immediate qualitative sense of the level of error in judging identity and the extent of difference in excess of error that nonidentical samples display. These parameters may obviously be more rigorously defined by statistical analysis.

After achieving satisfactory performance at the level of finding identity, it is worthwhile to examine the reproducibility with which differences are determined in multiple assays. In large measure, this is a test of the reproducibility of the printing process. A sample of this kind of testing is shown in Figure 22.3.4. Three repetitions of the experiment seen in Panel C of Figure 22.3.3 were carried out. Filtering for genes that had a minimum mean intensity of 50 in both channels developed a subset 3861 of well-detected genes. Scatter-plot comparisons of the ratios of the UACC903 (Cy5) to ML1 (Cy3) signal for genes that were well detected in all the experiments were then carried out. Panels A and B of Figure 22.3.4 show plots comparing the ratios observed in the first hybridization, experiment A, with the ratios observed in experiments B and C, respectively. Analysis at this level can easily show whether the assay is essentially working, failing in some areas of the array, or failing generally. Faults will show up as significant deviations from the diagonal. If only some portion of the genes are far from the diagonal, then their position on the array can be determined and the properties of local background and signal can be checked to see why the assay failed in those instances.

Anticipated Results

Microarray experimentation should provide a means for determining whether there is a difference in the expression level between those

genes that are present at sufficient levels in the sample to be detected by hybridization to the EST PCR products represented on the array.

Time Considerations

Time considerations for various operations described in this unit are as follows.

- Prepare replicates and cultures for template: 8 plates–5 hr.
- Prepare plasmid template from cultures: 4 plates–2.5 hr.
- Set up PCR using templates from cultures: 12 plates–2 hr.
- Prepare 200-well agarose gel: 2 gels–45 min.
- Load 2 plates of PCR samples: 1 gel–30 min.
- Ethanol precipitate and resuspend PCR products: 12 plates–5.5 hr. (with overnight incubation)
- Extract RNA: 4 samples–4 hr.
- Label RNA: 4 samples–4 hr.
- Set up hybridization: 2 slides–0.5 hr.

Acknowledgements

The authors would like to acknowledge the many excellent suggestions and kind support of our colleagues at NHGRI, Jeffrey Trent, Paul Meltzer, Yidong Chen, Javed Khan, Abdel El-kahloun, and Anthony Masiello. We would also like to thank the members of the Brown lab, especially Joe DeRisi for sharing the details of the printing methods that they developed, and the many members of the array research community who have generously shared their observations and insights with us.

Literature Cited

- Alwine, J.C., Kemp, D.J., and Stark, G.R. 1977. Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proc. Natl. Acad. Sci. USA* 74:5350-5354.
- Bishop, J.O., Morton, J.G., Rosbash, M., and Richardson, M. 1974. Three abundance classes in HeLa cell messenger RNA. *Nature* 250:199-204.
- DeRisi, J.L., Iyer, V.R., and Brown, P.O. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680-686.
- Iyer, V.R., Eisen, M.B., Ross, D.T., Schuler, G., Moore, T., Lee, J.C.F., Trent, J.M., Staudt, L.M., Hudson, J. Jr., Boguski, M.S., Lashkari, D., Shalton, D., Botstein, D., and Brown, P.O. 1999. The

- transcriptional program in the response of human fibroblasts to serum. *Science* 283:83-87.
- Khan, J., Simon, R., Bittner, M., Chen, Y., Leighton, S.B., Pohida, T., Smith, P.D., Jiang, Y., Gooden, G.C., Trent, J.M., and Meltzer, P.S. 1998. Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. *Cancer Res.* 58:5009-5013.
- Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., and Brown, E.L. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* 14:1675-1680.
- Phimister, B.E. 1999. The chipping forecast. *Nat. Genet.* 21(1 Suppl).
- Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270:467-470.
- Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D., and Futcher, B. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* 9:3273-3297.
-
- Contributed by Yuan Jiang, John Lueders,
Arthur Glatfelter, Chris Gooden,
and Michael Bittner
National Human Genome Research Institute, NIH
Bethesda, Maryland

Overview of mRNA Expression Profiling Using Microarrays

mRNA expression profiling using microarrays is a technology that allows simultaneous determination of the mRNA levels of many genes. Microarrays range from small custom arrays designed to monitor expression of a few hundred genes to very large arrays that represent tens of thousands of genes or entire genomes. Presently, there are three major applications for microarray data. One application treats microarray data as massively parallel expression assays. Typically, this data is used for identification of particular genes that undergo expression changes in response to particular treatments, in particular cell types, or in particular mutants. Such genes are often considered candidates for players in a biological process, such as response to a treatment. A second application involves treating expression profiles as descriptions of collective behaviors. The state of the cell from which the sample was prepared is collectively characterized by determining the expression levels of tens of thousands of genes. For example, an expression profile of a drug-treated cell describes the effect of the drug, and drugs with similar modes of action can be identified by comparing expression profiles of drug-treated cells. A third application which is becoming feasible is the mining of large expression profiling databases to characterize expression patterns of genes of interest over a wide range of tissue types, after various treatments, or in different mutants.

Due to the complexity of microarray data, computational tools are required for analysis. These tools must be tailored according to the type of analysis being carried out. If the goal is to identify genes that show expression-level changes between different samples, statistical tools are needed to sort genes based on the degree of confidence that they are actually differentially expressed. If the goal is to identify patterns in expression profiles that are diagnostic of particular cell states, pattern-recognition tools are needed. Another aspect of handling complex data is that in many cases subjective decisions are applied. For example, subjective criteria are used to decide how two different types of data, such as microarray data analysis results and gene annotation, should be combined for selection of candidate genes for further study. In the following discussion, an

overview of issues to consider when designing microarray experiments and analyzing the data is presented.

STRENGTHS AND WEAKNESSES OF EXPRESSION PROFILING

RNA and corresponding cDNA species have relatively homogenous chemical characteristics, and can be specifically detected based on hybridization of complementary strands. These advantages enable simultaneous monitoring of tens of thousands of genes with very good sensitivity and accuracy at reasonable cost. Another advantage of using cDNA species is that they can be amplified by PCR and/or in vitro transcription. Consequently, the amount of RNA required for expression profiling has been decreasing as methods for quantitative amplification of cDNA have improved (Wang et al., 2000; Baugh et al., 2001; Iscove et al., 2002). This feature allows expression profiling of very small amounts of tissue. Thus, microarray analysis could potentially have a very high spatial resolution (Ohshima et al., 2000). In terms of obtaining global quantitative information about a particular class of molecules, large-scale profiling technologies for other classes of molecules, such as proteins and metabolites, cannot match the performance of expression profiling.

On the other hand, the information obtained by expression profiling is simply the amount of each mRNA species present. mRNA levels do not necessarily correlate with levels of active, properly localized proteins or the amounts of metabolites that they produce. Consequently, the biological significance of observed changes in mRNA levels is open to question. This limitation must be kept in mind when interpreting the results of expression profiling experiments.

MICROARRAY TECHNOLOGIES

In its first incarnation, microarray technology involved spotting random cDNA clones for the purpose of detecting genes with differential expression levels in different samples in organisms with very limited DNA sequence information. The cDNA clones corresponding to genes that were found to be differentially expressed were then sequenced. Such an approach is quite inefficient, and other methods are now available to identify differentially-expressed genes

in organisms lacking extensive sequence information. An early example is differential display of mRNA (Liang and Pardee, 1992; UNIT 25B.3), followed more recently by methods including serial analysis of gene expression (SAGE; <http://www.sagenet.org>; UNIT 25B.6; Velculescu et al., 1995) and massively parallel signature sequencing (MPSS; http://www.lynxgen.com/wt/tert.php3?page_name=mpss; Brenner et al., 2000). Due to their capacities for covering large numbers of genes and rapid short sequence tag determination, SAGE and MPSS are also used for discovery of new mRNA species in organisms where genome information is available. Generally, the cost per sample of these methods is fairly high.

Affymetrix Arrays

These days the more common use of microarray technology is in monitoring expression of genes with known sequences. There are several different microarray technology platforms in use. Some platforms use arrays that must be purchased from commercial entities, while other types can be produced using widely available equipment. GeneChip arrays from Affymetrix (<http://www.affymetrix.com>) consist of a collection of short (typically 25-mer) oligonucleotides that are synthesized directly on the array (Lockhart et al., 1996). The oligonucleotides can be arrayed at extremely high density, so these arrays are usually used to represent entire genomes, or tens of thousands of different genes. Each gene is represented by eleven to sixteen pairs of exactly hybridizing oligonucleotides and oligonucleotides containing single-base mismatches. The signal from the mismatch oligos can be used for estimation of noise resulting from nonspecific hybridization. A single sample is hybridized to each array and the expression level for each gene is obtained by combining the signals from the multiple oligonucleotide probes. To compare data from different samples, the signals from each gene are normalized to the average signal for all the genes on the array, based on the assumption that most of the genes in the genome are not changing expression levels between any two experimental conditions. The main advantage of Affymetrix arrays is their excellent technical reproducibility due to the physical uniformity of the arrays themselves and the statistical power resulting from having multiple measurements for each gene. The disadvantages are that the design of the arrays cannot be altered to suit individual experiments, and the arrays are relatively expensive.

NimbleGen Arrays

NimbleGen arrays (<http://www.nimblegen.com>) are also high density arrays of oligonucleotides synthesized directly on a substrate (Nuwaysir et al., 2002). In contrast to Affymetrix arrays, the composition of the arrays can be changed very easily, so arrays can be tailored to particular experiments. In addition, longer oligonucleotides can be used for NimbleGen arrays. NimbleGen arrays are not available for purchase; rather, RNA samples must be sent to the company, which carries out the experiments and returns the data. NimbleGen arrays have not yet been widely used so it is difficult to evaluate the quality of the data; however, the similarity to the Affymetrix technology and published validation experiments (Nuwaysir et al., 2002) suggest high accuracy. The main advantage is the flexibility of the system—the design of individual arrays can be customized using a simple computer interface. The disadvantage is that the total cost of obtaining expression data from an RNA sample, including array production, hybridization, and scanning, is, at the time of this writing, even more expensive for a NimbleGen array than it is for a standard Affymetrix array.

Spotted Arrays

Rather than synthesizing oligonucleotides directly on a substrate, they can be spotted onto a glass slide. The required spotting robots are widely available, so this technology is suitable for production of custom arrays in academic laboratories. However, oligonucleotides cannot presently be spotted at the density required for representing tens of thousands of genes with multiple oligonucleotides per gene. Therefore, each gene is usually represented by one or two spots of an oligonucleotide of 60 to 70 bp, which can be purchased commercially or synthesized in-house by automated oligonucleotide synthesizers. The advantages of this technology are its flexibility and relatively low cost. However, there are technical challenges associated with production of high-quality data. Spotted arrays are less uniform than Affymetrix or NimbleGen arrays, making it difficult to compare data obtained from different arrays. Consequently, two-color methods are often used. In these experiments, two samples are labeled with different fluorescent dyes, and hybridized to the same array. The ratio of the two samples is used for further analysis. This method compensates for variation in the amount of oligonucleotide present in each spot although it can also result in a high level of error

(see below). Arrays on which each oligonucleotide is spotted only once or twice have inherently less statistical power than those on which each gene is represented by multiple oligonucleotides. For specialized arrays representing only a few hundred genes, the entire array can be spotted multiple times. Multiple measurements for each gene increase statistical power and result in reduced technical error.

Spotted arrays can also be produced using cDNA clones or PCR products rather than oligonucleotides. Such arrays can have high sensitivity due to the long lengths of the probes, but suffer from cross-hybridization of related genes. In addition, there is much more labor involved in the production, quality control, and tracking of the necessary PCR products or clones than in acquisition of a set of oligonucleotides. These arrays are also usually used with two-color methods.

Choosing an Array or an Alternative Technology

Choosing an appropriate microarray technology platform requires consideration of a number of questions, including: Is a satisfactory commercial array available? Is a satisfactory array available from an academic cooperative? How many genes must be monitored? What is the budget for the project? If the number of genes to be monitored is low and the number of samples is large, it might be wise to consider alternatives to microarrays. One is high-throughput quantitative RT-PCR, which offers much better sensitivity than microarrays. There are also bead-based technologies in which gene-specific probes are attached to fluorescently-labeled beads that are then counted using an instrument similar to a flow cytometer (Yang et al., 2001).

COMMON APPLICATIONS OF EXPRESSION PROFILING BY MICROARRAYS

Applications of microarray analysis fall into three major classes. One application is discovery of genes with different expression levels in particular samples. These might be genes induced in response to pathogen attack, genes expressed specifically in certain cell types, genes with altered expression after drug treatment, and so forth. Such genes might then be used as molecular markers for certain biological states, or investigated further to determine if they are causally associated with the biological process of interest. For example, one might test pathogen-induced genes for contributions

to disease resistance. To create accurate lists of differentially-expressed genes, experimental design and statistical analysis of the data are crucial. Generally, greater numbers of biological and technical replicates allow greater confidence in the accuracy of the lists of differentially-regulated genes.

Another application is the use of expression profiles as descriptions of the collective behavior of many genes for each sample (i.e., description of the cell state) or of many samples for each gene (i.e., description of the expression behavior of each gene). By comparing expression profiles of various mutants, it is possible to recognize mutants with defects similar to those of known mutants. For example, Hughes et al. (2000) discovered a yeast gene involved in sterol metabolism based on the observation that a mutation in it resulted in an expression profile similar to those of other mutants with defects in sterol metabolism. For this sort of analysis, it is critical to have powerful methods for pattern recognition to allow the investigator to recognize similarities among profiles. In contrast to conclusions about differential expression of individual genes, patterns composed of the expression values of many genes are statistically robust. Indeed, Hughes et al. (2000) showed that even when genes with high-magnitude changes are excluded from an analysis, similarity relationships among expression profiles are still quite stable. By comparing expression levels of genes in many different samples, genes that are regulated very similarly in many different situations can be identified. These genes may be involved in a common biological process (e.g., Kim et al., 2001).

Lastly, databases consisting of large amounts of expression profiling data can be used to rapidly obtain information about a gene of interest. For example, a database consisting of expression profiles from different tissues, mutants, and variously-treated samples can be mined to determine in what tissues a gene of interest is expressed, what signaling pathways control its expression, and what treatments alter its expression. The content of various microarray databases is rapidly increasing, so this sort of study will soon be possible for a variety of well-studied organisms. For this application, standardization of microarray data formats, data size and data quality are all important.

VARIATION IN MICROARRAY DATA

Two kinds of variation affect microarray data: biological and technical variation. Bio-

logical variation refers to real differences in the level of mRNA species between two samples that the investigator considers to be experimental replicates. They are caused by uncontrolled variables in experimental conditions as well as by stochastic biological events. For some samples, such as budding yeast growing in liquid culture, experimental conditions can be very tightly controlled. However, samples from two different mice, or worse yet, two different humans, are subject to many variations in experimental conditions that are beyond the investigator's control. Biological variation is also dependent on the stability of the biological process of interest in the presence of small perturbations (Tao et al., 2003). Irrespective of the source of biological variation, greater variation results in lower statistical power. The adverse effects of relatively high system variation may be overcome by increasing the number of replicates.

Technical variation results from variations in the mechanics of the microarray experiment itself (e.g., fabrication of the array, preparation of RNA from the sample, labeling, hybridization, washing, scanning). A simple way to assess the level of technical variation is to hybridize RNA from the same sample to two or more different arrays. Improvements in array technology and statistical procedures for dealing with background signal and normalization can reduce technical variation. In two-color methods, dye-swap technical replicates are highly desirable because the relative efficiencies of labeling with different dyes vary for different mRNA/cDNA species. While it is a good idea to assess technical variation for each sample in a microarray experiment, in the interests of economy, so-called "technical replicates" for experimental samples can be reduced (Churchill, 2002) or even omitted in favor of biological replicates, which include both the technical and biological variation in the system.

STATISTICS AND PRACTICAL DECISIONS

Statistics are important for many aspects of microarray data analysis. To begin with, microarray experiments should be designed based on statistical considerations, such as randomization (Churchill, 2002). For example, if mice cages for corresponding samples in biological replicates are kept in the same places, the correlation observed in gene expression might be due to the cage location rather than the intended experimental conditions. Data from a carefully designed experiment are particularly useful

when the data are shared for general use. However, for focused research, an experimental design that is not as rigorous may be used, dependent on the goals, bottlenecks, and budget of the research. For example, if the research goal is to identify a few good expression marker genes, running a microarray experiment with a small number of replicates, choosing a small number of candidate genes, and running a rigorous test for the candidate genes using quantitative RT-PCR could be more time- and cost-effective than running a microarray experiment with a large number of replicates.

Calculating Expression Values

For data analysis, there is the question of how to calculate a signal value for a probe spot based on the fluorescence intensity of a group of pixels in the scanned array image. Statistical principles are used to determine a small number of values (for example, the mean and standard deviation) that represent a larger group of values (such as the fluorescence intensity for a group of pixels). Microarray data must also be adjusted to compensate for nonspecific background signal. These calculations are usually performed using specialized software packages and require relatively little thought on the part of the typical user. For arrays on which each gene is represented by more than one probe spot, the values from each spot must be combined to yield a value for each gene. Normalization is required to allow comparisons between independent arrays with different overall signal intensities. There are multiple ways of doing this, which are often specific to particular microarray technology platforms. Comparison of different normalization methods is outside the scope of this overview. Interested readers may refer to other literature (e.g., Quackenbush, 2002; Bolstad et al., 2003).

Statistical Challenges in Comparing Expression Profiles

Once data has been obtained in the form of normalized expression values, the next common issue is determination of which genes have different expression levels in one sample relative to another. This requires calculation of the probability that the observed differences in expression values merely represent random samples taken from an underlying single population of possible values. If this probability is small, it is likely that the expression level of the gene really is different between the two samples (the underlying populations are separate), while if it is large, the gene likely does not have a different expression level in the two samples.

Significance of gene expression level differences

Determination of genes that are expressed at different levels in different samples from microarray data presents statistical challenges. Due to the high cost of carrying out the experiments, it is usually not practical to produce many independent replicates. If one assumes that the percent error (or coefficient of variation) of each gene is completely independent, analyses using small numbers of replicates have little statistical power, so only large expression differences are called statistically significant. One way to reduce the impact of this problem might be to assume the same percent errors for all genes. However, this is not a good assumption because weak signals (genes with low expression levels) typically have larger percent errors than strong signals (Jain et al., 2003). Commonly used tests for the significance of gene expression level differences, such as *t*-tests, assume that the underlying population of expression values is normally distributed; however, this assumption may not be valid, and its validity cannot be tested when the number of experimental replicates is small. It is possible to use nonparametric methods, in which no particular distribution is assumed, but these methods have less statistical power than parametric methods.

Another problem arises from the fact that microarray experiments are massively parallel assays. If 1000 genes are monitored and a 5% rejection rate (i.e., cutoff at $P = 0.05$) for each gene is applied to detect genes that are differentially expressed between two samples, on average 50 genes are falsely called differentially expressed. Multiple-testing corrections are statistical methods that increase the stringency of *P*-value cutoffs according to the number of genes monitored to reduce the frequency of such false-positives. They also increase the number of false negatives, thereby excluding genes that really are differentially expressed.

Permutation methods (SAM)

Permutation methods have been developed to respond to the statistical challenges presented by microarray data. One such method that is often used is significance analysis of microarrays (SAM; Tusher et al., 2001). In an example of SAM, an experiment with four replicate control samples and four replicate experimental samples was performed (Tusher et al., 2001). The relative difference value was calculated for each gene. The relative difference value was the difference between the mean

expression level of each gene in the experimental versus control samples, divided by the gene-specific scatter, defined as the standard deviation of repeated expression measurements. To assess the significance of this relative difference, the eight data sets were then randomly assigned to two groups of four data sets each (this is permutation), and the relative difference value for each gene was similarly calculated for the permuted groups. The expected relative difference for each gene was calculated by averaging the relative difference values of the gene from all the appropriate permuted group comparisons. The difference between the relative difference value of each gene determined in the control versus experiment comparison and the expected relative difference value of the gene was used as a measure of the likelihood that apparent expression differences were real.

Examples of Common Situations

Since there are practical limitations on the level of confidence that can be obtained from microarray data, it is important to have a clear idea of the purpose of the experiment and the effects of statistical limitations, and to apply statistical methods that are appropriate to the data set under study. The following examples illustrate some common situations.

Example 1: Identifying genes that may be important in a biological process

If genes are selected using stringent criteria to reduce false positives, very few candidate genes are obtained. If the criteria are relaxed, more genes are obtained, but these certainly include false positives. The key question is capacity for testing candidate genes. If many candidates can be tested, then it is better to use less stringent criteria and rely on downstream testing to eliminate false positives. If downstream testing is onerous, it is better to use more stringent selection criteria.

Example 2: Replicate data generated by two different individuals

The variance among replicate samples generated by different individuals is large relative to the variance among replicates generated by the same individual. Data generated by the two individuals should not be simply combined in an analysis.

Example 3: Non-normal distributions

If the distributions of values among replicate samples clearly deviate substantially from the bell shape of a normal distribution, statistical

tests that assume an underlying normal distribution should not be used. Either the values should be transformed first so that their distribution approximates a normal distribution, or a nonparametric test should be employed.

CHOICE OF STATISTICS FOR COMPARISON

When using array platforms with single-color detection methods (e.g., see Affymetrix Arrays), it is possible to compare expression values directly. This is convenient when many different combinations of data sets are used for comparison. When needed, by combining expression values from two appropriate experiments, expression ratio values can be calculated. With two-color methods commonly used with spotted arrays, only ratio values can be used. There is an advantage to using expression ratio values for comparisons. Unlike expression values, ratio values are independent of technology platform, so ratios obtained from one platform can be used for comparisons with ratios obtained from a different platform; however, when using ratio values, it is very important to consider the effect on error levels. The ratios are obtained from two measurements, each with an associated error, so the percent error in the ratio is larger than the percent error in either measurement. Ratios in which the expression level in one sample is very low will generally have large errors, since weak signals tend to have large percent errors. Deriving ratios from two different arrays rather than from a direct comparison on a single array, such as deriving the ratio of C/A from the ratios B/A obtained from one array and B/C obtained from another array, further increases errors. Therefore, it is ideal to have a direct comparison of each interesting pair of samples when using a two-color method. On the other hand, if many pair-wise comparisons are needed, using a common reference sample may be more practical despite the error problem discussed above. Factors affecting the choice of suitable reference samples have been discussed elsewhere (e.g., Churchill, 2002). For compiling standardized databases for general use, data from a single color method or data from a two-color method using the same reference sample for all the arrays is more valuable than other types of ratio data due to ease of comparison among various data sets.

Expression values and ratio values are often log-transformed before comparison. This is appropriate if log-transformation brings the dataset closer to a normal distribution, improv-

ing the accuracy of statistical tests based on the assumption of a normal distribution. In addition, when expression ratios are log-transformed, the characteristics of the transformed values are more intuitive. X-fold induction and repression have the same absolute value with opposite signs. Also, the scale is compressed for larger numbers. On a \log_2 scale, the difference between a fold change of 2 ($\log = 1$) and a fold change of 8 ($\log = 3$) is 2, rather than 6 on a linear scale, while the difference between a fold change of 10 ($\log = 3.32$) and a fold change of 16 ($\log = 4$) is 0.68, rather than 6 on a linear scale.

THE PROBLEM OF CLASSIFYING GENES INTO A SMALL NUMBER OF GROUPS

Genes are often classified into two groups such as those that are induced and those that are not. Such groups of genes are sometimes used to make arguments about how many genes are common to two groups that are induced under different conditions. Imagine that genes are divided into induced and uninduced genes using either a two-fold cutoff or a probability cutoff. Typically the cutoff values lie in the tail of the distribution of the values for all the genes. Consequently, many genes have values close to the cutoff value. For example, if the cutoff value is a two-fold change, there will be many genes with values of 1.9 or 2.1. Genes with a value of 1.9 will be judged as uninduced while those with a value of 2.1 will be judged as induced. Is there really a major difference between a 1.9-fold change and a 2.1-fold change? Among the genes that pass the cutoff, there will be large variations in the extent to which they pass, yet they are all classified in the same group. For example, a gene that shows a 2.1-fold change is grouped together with a gene that shows a 200-fold change. Are these fold-changes really similar? These problems become greater when two sets of induced genes are compared to determine the set of genes induced in common by two different treatments. The impact of these limitations on conclusions drawn from microarray data should be kept in mind.

ADVANTAGES OF COMMERCIAL SOFTWARE

Several expensive software packages for microarray analysis are commercially available. Programs that will perform most of the analytical methods implemented in these packages are freely available somewhere on the web, so why should anyone spend a lot of money on the

commercial products? First, such products have the various methods put together nicely. It is quick and easy to apply different analytical methods and combine or compare the results. Second, they have integrated database management mechanisms, which can also take care of database security and sharing issues. Third, they are designed to be highly interactive. Well-organized, highly interactive databases provide access to information that helps the investigator make subjective decisions about how to integrate different types of information. Fourth, they offer many ways to visualize the data and the results of analyses. People have excellent visual pattern recognition abilities, so visualization of data is a powerful tool aiding comprehension of large data sets. Fifth, they often offer ways to record a series of analyses applied to the data. Such a record is very important because various criteria used in analyses could be subjectively decided. In addition, such a recorded series of analyses can be used as a macro, so the same series of analyses are easily applied to different data sets.

In using sophisticated software tools, it is necessary to understand the analytical tools being applied to avoid drawing faulty conclusions. For example, when using false color to display differences in gene expression levels, the eye is drawn to the boundaries between different colors, and one tends to conclude that genes in different color groups have very different expression levels, when they may actually be quite similar.

SUMMARY

Expression profile data obtained from microarrays can be enormously useful in addressing a variety of biological questions, including genome-scale questions that were previously unapproachable. Maximizing the potential of microarray experiments requires attention to aspects of experimental design and data analysis. Experiments should be designed so as to minimize systematic errors and errors arising in the process of data analysis. Good design can also make it possible to determine how much of the variation among replicates arises from various sources, such as technical variation or biological variation. When analyzing data, it is important to understand the limitations of the analytical methods used to avoid drawing erroneous conclusions. Public databases consisting of microarray data will be very useful for a variety of large-scale analyses, provided that the data sets are of high quality. Investigators

who produce data through carefully-designed and executed experiments thus have the opportunity to benefit the larger scientific community as well as to further their own research goals.

COMPARISON OF MULTIPLE PROFILES

In an upcoming supplement, another aspect of expression profile data analysis will be considered: comparison of multiple profiles. Similarities among profiles can be used in a number of ways. For example, tumors that respond similarly to treatments can have similar expression profiles, and this can be used to determine which treatment may be most effective (Lapointe et al., 2004). Also, expression profiles of mutants can be used to predict the nature of the defects in the mutants (e.g., Hughes et al., 2000; Glazebrook et al., 2003). Detection of similarities among expression profiles is a problem of pattern-recognition among distributions of points in n -dimensional space, where each point represents an expression profile and n is the number of genes represented in each profile. While people have difficulty recognizing patterns in spaces of more than three dimensions, computer algorithms can handle as many dimensions as needed. Many computational methods for pattern recognition have developed, and several are routinely used for analysis of expression profiling data. These methods, together with methods under development, may prove to be even more useful in the future.

LITERATURE CITED

- Baugh, L.R., Hill, A.A., Brown, E.L., and Hunter, C.P. 2001. Quantitative analysis of mRNA amplification by *in vitro* transcription. *Nucl. Acids Res.* 29:e29.
- Bolstad, B.M., Irizarry, R.A., Åstrand, M., and Speed, T.P. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19:185-193.
- Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D.H., Johnson, D., Luo, S., McCurdy, S., Foy, M., Ewan, M., et al. 2000. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* 18:630-634.
- Churchill, G.A. 2002. Fundamentals of experimental design for cDNA microarrays. *Nat. Genet.* 32:490-495.
- Glazebrook, J., Chen, W., Estes, B., Chang, H.-S., Nawrath, C., Métraux, J.-P., Zhu, T., and Katagiri, F. 2003. Topology of the network integrating salicylate and jasmonate signal transduction derived from global expression phenotyping. *Plant J.* 34:217-228.

- Hughes, T.R., Marton, M.J., Jones, A.R., Roberts, C.J., Stoughton, R., Armour, C.D., Bennett, H.A., Coffey, E., Dai, H., He, Y.D., et al. 2000. Functional discovery via a compendium of expression profiles. *Cell* 102:109-126.
- Iscove, N.N., Barbara, M., Gu, M., Gibson, M., Modi, C., and Winegarden, N. 2002. Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA. *Nat. Biotechnol.* 20:940-943.
- Jain, N., Thatte, J., Braciale, T., Ley, K., O'Connell, M., and Lee, J.K. 2003. Local-pooled-error test for identifying differentially expressed genes with a small number of replicated microarrays. *Bioinformatics* 19:1945-1951.
- Kim, S.K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J.M., Eizinger, A., Wylie, B.N., and Davidson, G.S. 2001. A gene expression map for *Caenorhabditis elegans*. *Science* 293:2087-2092.
- Lapointe, J., Li, C., Higgins, J.P., van de Rijn, M., Bair, E., Montgomery, K., Ferrari, M., Egevad, L., Rayford, W., Bergerheim, U., et al. 2004. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc. Natl. Acad. Sci. U.S.A.* 101:811-816.
- Liang, P. and Pardee, A.B. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257:967-971.
- Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., and Brown, E.L. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* 14:1675-1680.
- Nuwaysir, E.F., Huang, W., Albert, T.J., Singh, J., Nuwaysir, K., Pitas, A., Richmond, T., Gorski, T., Berg, J.P., Ballin, J., et al. 2002. Gene expression analysis using oligonucleotide arrays produced by maskless photolithography. *Genome Res.* 12:1749-1755.
- Ohyama, H., Zhang, X., Kohno, Y., Alevizos, I., Posner, M., Wong, D.T., and Todd, R. 2000. Laser capture microdissection-generated target sample for high-density oligonucleotide array hybridization. *Biotechniques* 29:530-536.
- Quackenbush, J. 2002. Microarray data normalization and transformation. *Nature Genet.* 32:496-501.
- Tao, Y., Xie, Z., Chen, W., Glazebrook, J., Chang, H.-S., Han, B., Zhu, T., Zou, G., and Katagiri, F. 2003. Quantitative nature of *Arabidopsis* responses during compatible and incompatible interactions with the bacterial pathogen *Pseudomonas syringae*. *Plant Cell* 15:317-330.
- Tusher, V.G., Tibshirani, R., and Chu, G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U.S.A.* 98:5116-5121.
- Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. 1995. Serial analysis of gene expression. *Science* 270:484-487.
- Wang, E., Miller, L.D., Ohnmacht, G.A., Liu, E.T., and Marincola, F.M. 2000. High-fidelity mRNA amplification for gene profiling. *Nat. Biotechnol.* 18:457-459.
- Yang, L., Tran, D.K., and Wang, X. 2001. BADGE, BeadsArray for the Detection of Gene Expression, a high-throughput diagnostic bioassay. *Genome Res.* 11:1888-1898.

Contributed by Fumiaki Katagiri and
Jane Glazebrook
University of Minnesota
St. Paul, Minnesota

Pattern Discovery in Expression Profiling Data

Expression profile data typically consists of many profiles, each of which has many parameters. Such data are typically represented by a table with many columns and rows. This makes it extremely difficult to recognize trends in the data without the assistance of analytical tools. Recognition of major trends in the data enables the investigator to formulate new hypotheses. In this unit, some analytical tools for recognition of trends in large data collections, known as pattern discovery tools, are discussed. Pattern discovery tools are also known as multivariate analysis methods in which factors that are responsible for trends are assumed to be unknown. This unit is intended to explain pattern discovery tools to biologists, who do not necessarily have extensive mathematical knowledge. As a consequence, explanations are not mathematically rigorous. They are based mainly on viewpoints from geometry rather than viewpoints from statistics or linear algebra in hopes that many readers will find this more intuitive.

GENERAL CONCEPTS

The Same Mathematical Procedures Can Be Used to Compare Genes or Samples

There are two ways in which expression profiles can be considered. One way is to compare individual genes to each other based on their expression values in various samples. The expression values of each gene in different experimental samples are considered as parameter values comprising one profile. Another way is to compare individual samples to each other based on the expression levels of various genes. Expression values for different genes in each experimental sample are considered as parameter values comprising one profile. Once genes and samples have been assigned as profiles or parameters, the mathematical procedures used by a pattern discovery tool are the same for either comparison.

Data points in a high-dimensional Euclidian space

The principles used by computer algorithms to find genes that are similarly regulated or experimental samples with similar expression

profiles are generally based on distributions of data points in a high-dimensional Euclidian space, in which all the dimensions are perpendicular to one another. For example, when expression profiles of many genes in three different experimental samples (parameters) are compared, the profile for each gene is represented by a single data point in a 3-D linear space, in which each dimension corresponds to one experimental sample (parameter). If expression values of gene A in experimental samples x_1 , x_2 , and x_3 are 300, 500, and 220, respectively, the data point for gene A is expressed as the coordinates of $(x_1, x_2, x_3) = (300, 500, 220)$ (Fig. 22.5.1A). Thus, an expression profile of any gene in the three samples can be expressed as a single data point in this space. According to how the distance between two data points is defined (see below), the similarity between the expression profiles of the two corresponding genes can be calculated—the closer the points are, the more similar their expression profiles. If many data points form a relatively dense group, such a group can be considered as a cluster of similarly regulated genes (Fig. 22.5.1B). Thus, finding a cluster of similarly regulated genes is equivalent to recognizing closely distributed data points. Different computer algorithms use different principles to discover patterns in the distribution of data points.

For comparison of expression profiles of many genes through 100 different experimental samples, a 100-dimensional Euclidian space $(x_1, x_2, \dots, x_{100})$ is used. Again, an expression profile of a single gene can be described as one data point in the 100-D space. In other words, a profile with n parameters can be expressed as a single data point in an n -dimensional space. In the following discussion, “profile” and “data point” are used interchangeably, as are “parameter” and “dimension.” The mathematical definition of distance between two data points can be easily expanded to a high dimensional space. While a space with >3 dimensions is difficult to visualize, a computer can easily apply the same principle regardless of the number of dimensions in the space. This is why computers are valuable for analysis of data with high dimensionality (i.e., with many parameters).

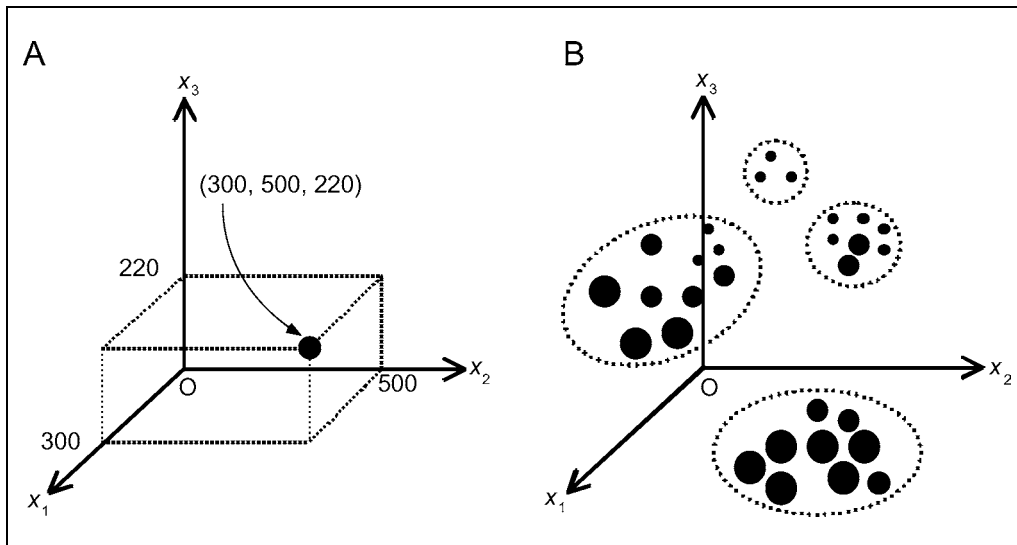


Figure 22.5.1 Representation of expression profiles in a Euclidian space. **(A)** A single point can be specified for a profile in a Euclidian space with a number of dimensions equal to the number of parameters. **(B)** Pattern discovery among expression profiles is equivalent to pattern discovery in the distribution of data points.

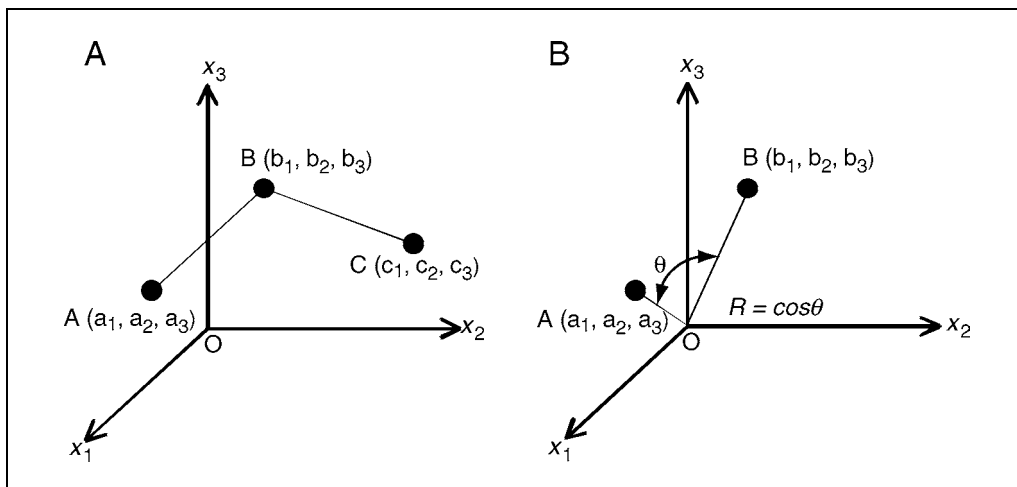


Figure 22.5.2 Distance metrics. **(A)** Euclidian distance. Euclidian distance is the length of the straight line between two points in a Euclidian space. **(B)** Uncentered Pearson correlation coefficient (R). R is the cosine of the angle made at the origin between A and B .

Distance Metrics

There are multiple ways to define the distance between two data points. Only two commonly used distance metrics are described here. One is the Euclidian distance. In a 3-D space, this is seen as the length of the straight line connecting two data points (Fig. 22.5.2B). In this figure, the Euclidian distance D between data points A and B is defined as

$$D = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2}.$$

It is easy to expand this distance metric to an n -dimensional space, where it is defined as

$$D = \sqrt{\sum_{k=1}^n (a_k - b_k)^2}.$$

The other is the uncentered Pearson correlation coefficient. This is the equivalent of the cosine of angle θ that is made between two data points at the origin (Fig. 22.5.2B). The smaller the angle is, the closer the data points

are. Since it is the cosine value, its range is from 1 (perfect correlation when $\theta = 0^\circ$) to -1 (perfect negative correlation when $\theta = 180^\circ$). When the uncentered Pearson correlation coefficient is 0, there is no correlation between two profiles and $\theta = 90^\circ$. Note that the uncentered Pearson correlation coefficient ignores the distance of a data point from the origin; it is based only on the angle between two data points. In Figure 22.5.2B, the uncentered Pearson correlation coefficient R between data points A and B is defined as

$$R = \frac{a_1b_1 + a_2b_2 + a_3b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2} \sqrt{b_1^2 + b_2^2 + b_3^2}}$$

This distance metric can be expanded to an n -dimensional space as

$$R = \frac{\sum_{k=1}^n a_k b_k}{\sqrt{\sum_{k=1}^n a_k^2} \sqrt{\sum_{k=1}^n b_k^2}}$$

If the vector expression is used,

$$R = \frac{\overline{OA} \cdot \overline{OB}}{|\overline{OA}| |\overline{OB}|},$$

which the reader may recognize as the cosine function.

The distance metric is the metric that describes to what extent two profiles are similar.

In Figure 22.5.2A, the distance between data points A and B is the same as that between data points B and C. However, from data point B, the directions to data points A and C are different. The direction represents the way in which profiles are similar. For example, the difference between data points A and B could be explained by the expression difference in sample x_1 , while the difference between data points B and C could be explained by the expression differences in samples x_2 and x_3 . Some of the pattern discovery tools described below do not consider the ways in which data points are similar, while others do.

Shapes and Amplitudes of Expression Profiles

When the number of parameters and the number of profiles are limited, one good way to visualize expression profiles is to use line graphs as shown in Figure 22.5.3A. Although line graphs are generally not recommended in ordinary biology unless the parameters represent consecutive points in a single metric, such as consecutive time points, the line graph representation strongly enhances the ability to visually compare patterns. The expression values of gene A, (300, 500, 220) are exactly four times as high as those of gene D (75, 125, 55). That is, the amplitude of the gene A profile is four times larger than the amplitude of the gene D profile. Compare two representations of these two profiles in Figures 22.5.3A and 22.5.3B. In Figure 22.5.3B, the amplitude difference is seen as the difference in the

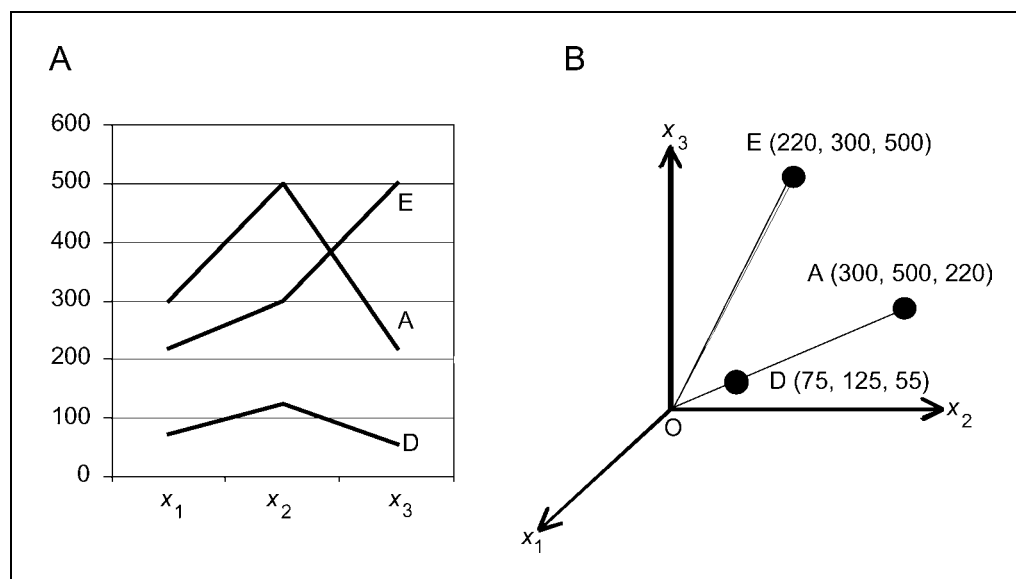


Figure 22.5.3 Shapes and amplitudes of expression profiles are represented (A) by line graphs and (B) in a Euclidian space.

distance from the origin: data point A is four times further away from the origin than data point D. The shape of the profile defined by the relative ratio of expression values is the same for genes A and D as seen in Figure 22.5.3A. The fact that the shapes of the profiles are the same is represented in Figure 22.5.3B as the same direction of data points A and D from the origin O. In other words, the vectors \vec{OA} and \vec{OD} point in the same direction. The profile of gene E (220, 300, 500) has the same amplitude as gene A, as their distances from the origin are the same. However, the shapes of their profiles are different as seen in Figure 22.5.3A. The profile shape difference is represented in Figure 22.5.3B as different directions of \vec{OA} and \vec{OE} .

The shapes and amplitudes of profiles are distinct characteristics. Generally, they should be considered separately. The reader should be aware that methods of combining these two characteristics into one metric, such as using the Euclidian distance, are subjective. For example, imagine that profile B and profile A have the same shape, but profile B is three times larger than profile A and that profile C and profile A have the same amplitude but their shapes are totally unrelated (i.e., the uncentered Pearson correlation coefficient between them is zero). Using the Euclidian distance, the distance between profiles A and B is larger than that between profiles A and C. Whether this is appropriate or not depends on subjective decisions.

What Profiles Should Be Considered the Same?

Centering and normalization

Before applying pattern discovery analysis, expression values or ratio values can be transformed. They can be log-transformed as described in the previous unit. Two other transformations are often applied—typically called “centering” and “normalization.” Centering is adjusting the mean or median of the values to 0. When an expression profile with the expression values (300, 500, 220) is centered using the mean value, it results in a profile with (−40, 160, −120). Note that centering could change both the shape and amplitude of the profile. Normalization means adjusting the Euclidian distance of each data point from the origin to 1. For example, when a profile with the expression values (300, 500, 220) is normalized, it results in a profile with (0.481, 0.802, 0.353). Note that normalization does not alter the shapes of the profiles, but could

change the amplitude of the profile. Normalization is an integral part of the uncentered Pearson correlation coefficient, and therefore, the distance of each data point from the origin, i.e., the amplitudes of profiles, is ignored in the uncentered Pearson correlation coefficient. Calculating the centered Pearson correlation coefficient is the same as applying the uncentered Pearson correlation coefficient after centering the profiles.

These transformations are applied to each profile. When the profiles for genes are being compared (samples as parameters), if any of these transformations are used, they should be applied to the set of values for each gene, not the set of values for each sample. In contrast, when the profiles for samples are being compared (genes as parameters), if any of these transformations are used, they should be applied to the set of values for each sample, not the set of values for each gene.

Statistically speaking, centering and normalization of each profile are recommended as they allow more robust statistical comparisons. However, these transformations can change the biological meaning of the profiles. Figure 22.5.4A shows profiles that become the same after centering. Figure 22.5.4B shows profiles that become the same after normalization. Figure 22.5.4C shows profiles that become the same after both centering and normalization. If both transformations are applied, centering should be performed first. The choice of transformations should be made on a case-by-case basis, depending on what profiles should be considered the same. For example, if the values are log-transformed ratio values (positive values for induced and negative values for repressed), it is not a good idea to center the values because centering could change the signs of the values. In another example, if Euclidian distance is used as the distance metric and if the amplitude of the profile is intended to be factored into the distance, normalization is not appropriate.

PATTERN DISCOVERY TOOLS

Various pattern discovery tools have been used to help find major trends in expression profile data. The principles, advantages, and disadvantages of several methods are discussed below. These analyses are generally processes of information reduction that reveal major trends in the data. For example, two drug candidates might be found to have very similar effects on expression profiles, suggesting that they have similar modes of action. However, in the course of identifying major

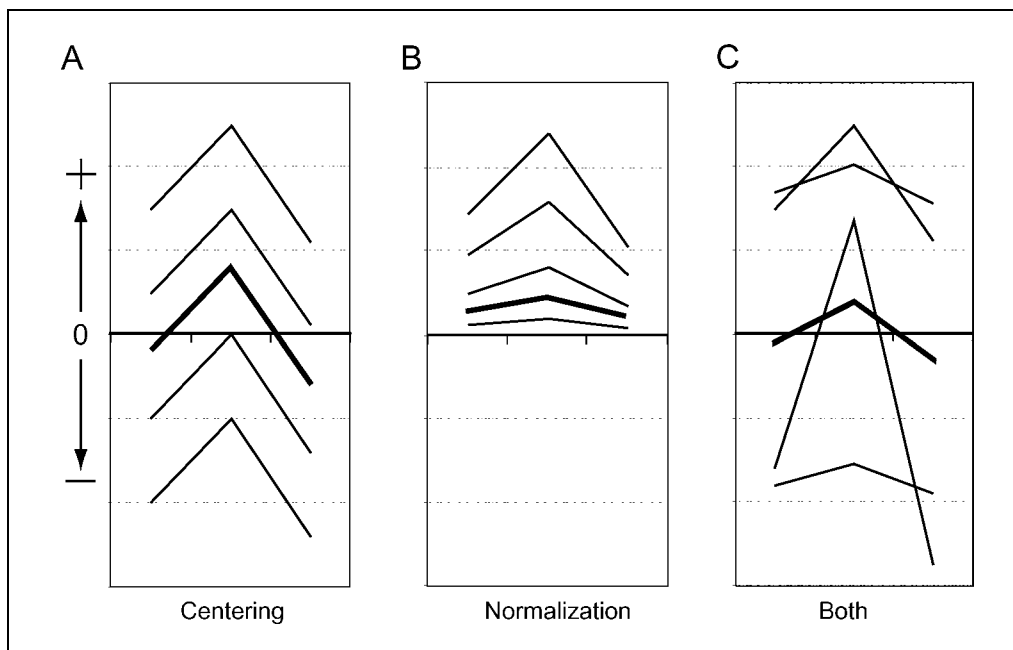


Figure 22.5.4 Effects of (A) centering, (B) normalization, and (C) both on expression profiles. In each panel, the transformation results in all the profiles being converted to the profile indicated by the thick line.

trends in the data, minor patterns are discarded, resulting in loss of some information that may actually be important. Different methods define major trends in different ways, and discard different types of minor patterns. These differences should be kept in mind when applying a pattern discovery tool to biological data.

As discussed in the beginning of the unit, large numbers of profiles and parameters create challenges for pattern discovery. Generally, the strategies used by pattern discovery tools are reducing the number of profiles by clustering similar profiles together, or reducing the number of parameters by mathematical combination. By clustering similar profiles, the number of profiles that need to be handled at once can be reduced; the profiles within a cluster can be compared or profiles representing the clusters can be compared. The number of parameters can be reduced by dimensionality reduction methods. A space with a small number of dimensions that still holds most of the information in the data can be derived by combining initial parameters appropriately. If the resulting space has three or fewer dimensions, the overall relationship among the profiles is easily visualized.

Clustering Methods

Hierarchical clustering

The version of hierarchical clustering commonly used for expression profile analysis is

agglomerative hierarchical clustering, which is a bottom-up method for generating a hierarchical tree (Eisen et al., 1998). Like a phylogenetic tree, the idea is to place similar profiles in the same branch of the tree. Among all the possible pair-wise combinations of data points, the one with the shortest distance gets connected first. The connected data points are replaced by a node, which is treated as a new data point. This process of connecting the data points is repeated until all the data points are connected. In this way, the tree (dendrogram) is built from peripheral branches towards the root (Fig. 22.5.5). There are three common ways to define nodes: (1) in average linkage, the node is the average point of all the data points represented by the node; (2) in complete linkage, the point among a set of connected data points that is furthest away from another point being considered for linkage is used as the node; and (3) in single linkage, the point among a set of connected data points that is closest to another point being considered for linkage is used as the node.

Hierarchical clustering performs well in detecting similar data points at local levels. This is very useful for detecting highly similarly regulated genes or very similar samples. However, as branches become bigger, multiple groups of data points that are fairly similar are often placed on relatively distant branches. Two factors contribute to this. First, hierarchical clustering forms a single connection at a

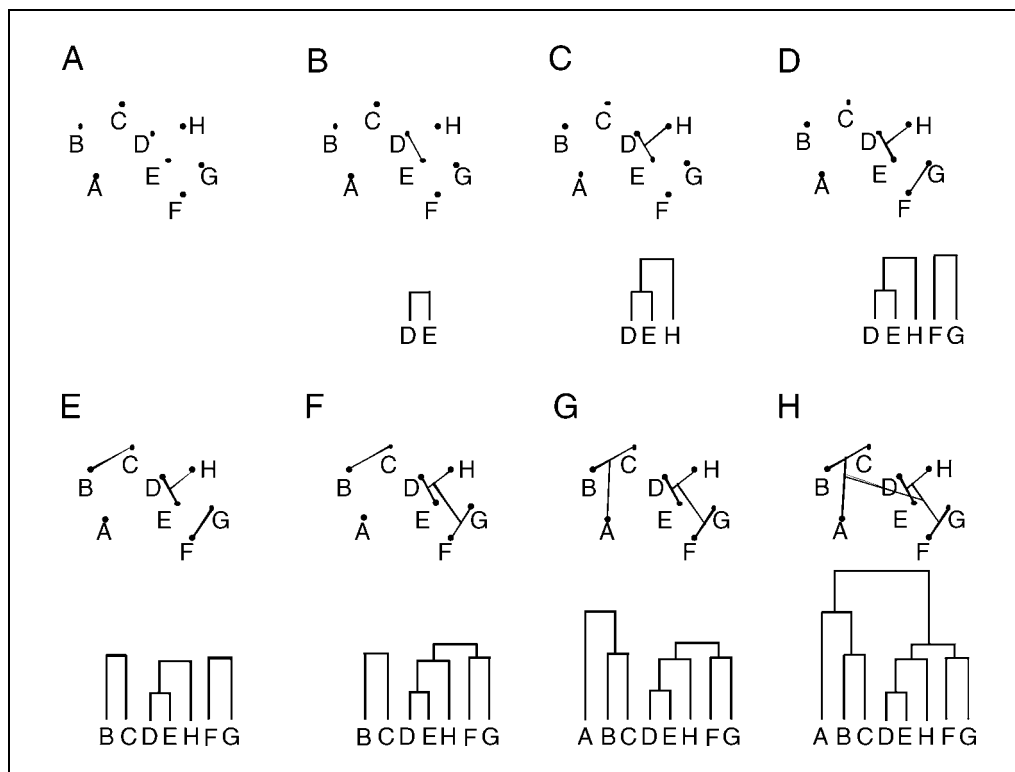


Figure 22.5.5 Agglomerative hierarchical clustering. (A through H) schematically depict sequential steps in building a hierarchical tree based on the 2-D positions of the data points shown in A. Average linkage was used to determine nodes.

time by considering only the degree of similarity and not the nature of the similarity. Second, each new connection affects the position of the node (i.e., process dependency). For example, in Figure 22.5.5, data points C and D are relatively close, but due to the effects of other data points during the clustering process, data points C and D appear to be distantly located from each other in the final dendrogram (Fig. 22.5.5H). Thus, a data point that is somewhat similar to other data points in a cluster may not be incorporated into the growing cluster until a relatively late cycle of connection forming, when the node has moved away from it.

An important advantage of hierarchical clustering is that by putting similarly behaving parameters together, it can generate a color-coded 1-D pattern for each profile with thousands of parameters, e.g., a profile with the values for thousands of genes in a sample (Eisen et al., 1998). This allows researchers to visually inspect such color-coded patterns for different samples to identify which samples are similar and which parts of the pattern (i.e., expression of which genes) are different between samples. Hierarchical clustering is a simple, powerful visualization method.

K-means clustering

This is a method that divides data points into K clusters (Herwig et al., 1999). The number of clusters, K , is given by the user. Each cluster of data points is represented by a centroid. First, K centroids are placed in the data space, and then the assignments of each data point to a particular centroid and the positions of the centroids are iteratively optimized (Fig. 22.5.6). There are several variants of K -means clustering that vary according to what metric is used for optimization of the positions of the centroids. Such a metric could be the sum of the average distance to the centroids, the total distances of all data points from their centroids, the sum of the variance over all clusters, and so on.

The results of K -means methods are highly dependent on the number and the initial positions of the centroids. Such clusters are well-separated, round-shaped (i.e., convex-shaped) clusters. K -means clustering is useful for discovery of such well-defined clusters (Xu et al., 2002).

Network-building methods

Converting similarities among data points into a network structure could help to

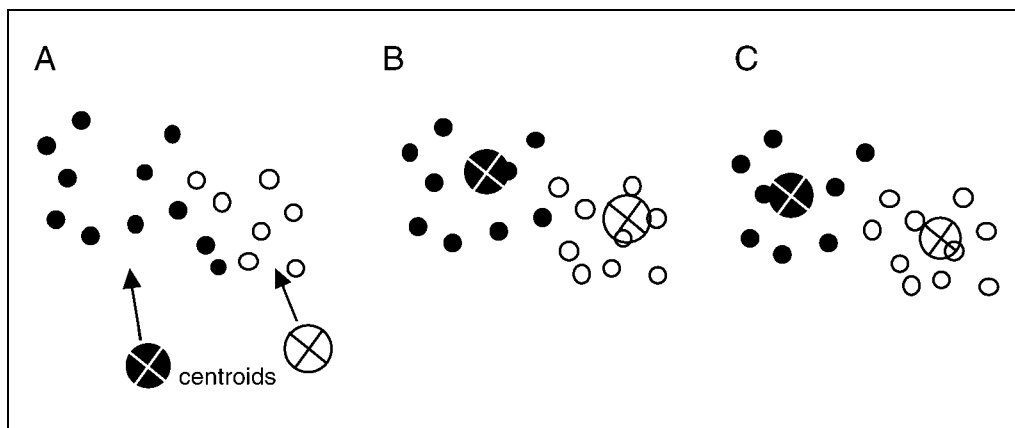


Figure 22.5.6 *K*-means clustering. **A** through **C** schematically illustrate sequential steps in the process of optimization in a 2-D space when $K = 2$. **(A)** shows the initial positions of the centroids and data point assignments to the centroids. Small circles represent data points and large circles with an "X" represent centroids. The color of each data point indicates which centroid it is assigned to. Note that during the optimization, not only the positions of the centroids but also the assignments of some data points changes.

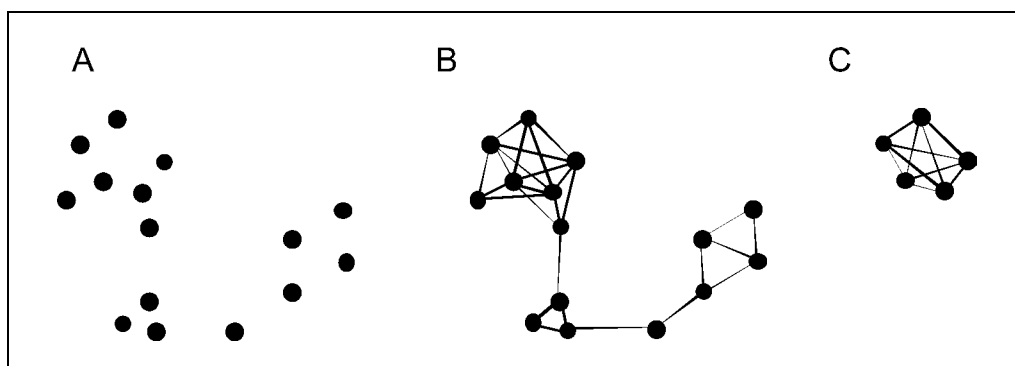


Figure 22.5.7 Network building. Among the data points in **(A)**, pairs of data points are linked when their distances are shorter than a predetermined distance **(B)**. It is clear that data points in the upper left region are highly connected to one another. For example, in the sub-network featured in **C**, each data point within the sub-network is connected to the other data points in the group.

visualize the data point distribution. For example, links can be made between all the data point pairs that have distances shorter than a certain value. When many data points are located within a small part of the space, they are highly linked to one another, and such a group can be recognized as a cluster (Fig. 22.5.7). In practice, the rule for making links between data points could be more sophisticated, such as using a cutoff for the probability value determined by a certain statistical analysis, but the basic idea is the same. Such network structures can be visualized using network visualization tools, such as Pajek (<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>); Batagelj and Mrvar, 1998) and VxInsight (<http://www.cs.sandia.gov/projects/VxInsight.html>); Davidson et al., 1998). Kim et al. (2001) provide a good example of visualization of a network built on

analysis of expression profile data. Pajek is also a network analysis tool, and can be used to define clusters according to mathematically defined network characteristics.

Dimensionality Reduction Methods

Principal component analysis (PCA), a linear dimensionality reduction method

In a linear space, when the number of data points is smaller than the number of dimensions, the number of dimensions can be reduced to the number of data points by linear transformations without any loss of information. This situation is illustrated in Figure 22.5.8. However, this number is still likely to be too large to allow the investigator to grasp the relationships among the data points. Linear dimensionality reduction methods, such as

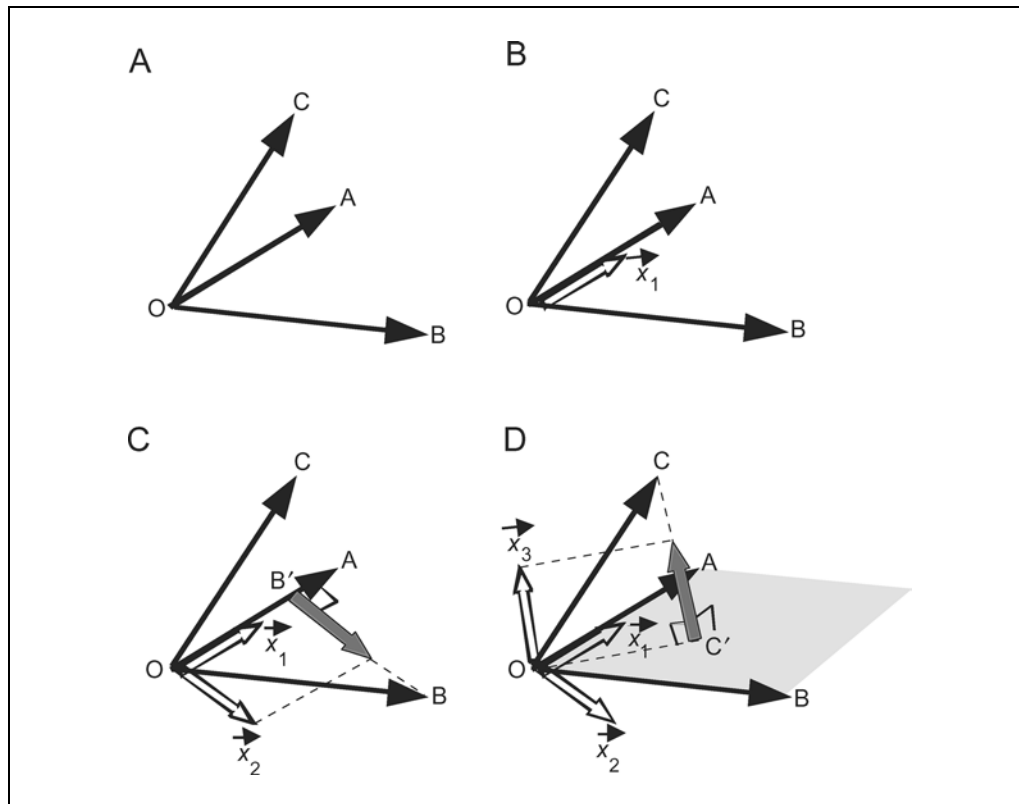


Figure 22.5.8 The number of necessary dimensions is at most the number of data points. When three data points A, B, and C are in a high-dimensional space with origin O (**A**), the first dimension \vec{x}_1 can be chosen along the line \overline{OA} (**B**). The second dimension \vec{x}_2 can be chosen parallel to $\overline{B'B}$, where B' is on \overline{OA} or its extension, and $\overline{B'B}$ is perpendicular to \overline{OA} or its extension (**C**). The third dimension \vec{x}_3 can be chosen parallel to $\overline{C'C}$, where C' is on the plane defined by \vec{x}_1 and \vec{x}_2 and $\overline{C'C}$ is perpendicular to the plane (**D**). Although it is difficult to visualize a space with a dimension higher than three, by repeating this procedure, it can be shown that the number of necessary dimensions is at most the number of data points.

PCA (Alter et al., 2000; Holter et al., 2000), are used to further reduce the number of dimensions while minimizing loss of information. Linear dimensionality reduction in data analysis is based on the belief that apparently complex data can be closely approximated by linear combinations of a relatively small number of independent parameters. Linear dimensionality reduction is equivalent to factor analysis in statistical terms, in which independent parameters are the equivalents of unknown independent factors.

Figure 22.5.9 illustrates a very simple case using a 2-D space to explain the principle of PCA. If all the data points are perfectly aligned in a straight line, one dimension along the straight line, instead of two dimensions, is sufficient to describe all the differences among all the data points (Fig. 22.5.9A). If all the data points are almost aligned in a straight line, loss of information by using only one dimension along the straight line is limited to

very small deviations of the data points from the straight line (Fig. 22.5.9B). Just using one dimension, the relationships among the data points are easily visualized. The dimension that explains the largest difference, defined as the variance along the dimension, among the data points is called the first principal component (PC) of the data. The dimension along the straight line in the Figure 22.5.9B example is the first PC of the data. The dimension that explains the largest remaining differences among the data points is called the second PC. In Figure 22.5.9B, the dimension perpendicular to the first PC is the second PC. When the initial parameter number is higher, the third and subsequent PCs can be defined similarly. Note that all the PCs are perpendicular to one another, and a Euclidian space can be defined using PCs, i.e., the data points can be expressed using PCs as the parameters. The number of PCs that should be considered can be assessed by the residual variance, i.e., how much of the

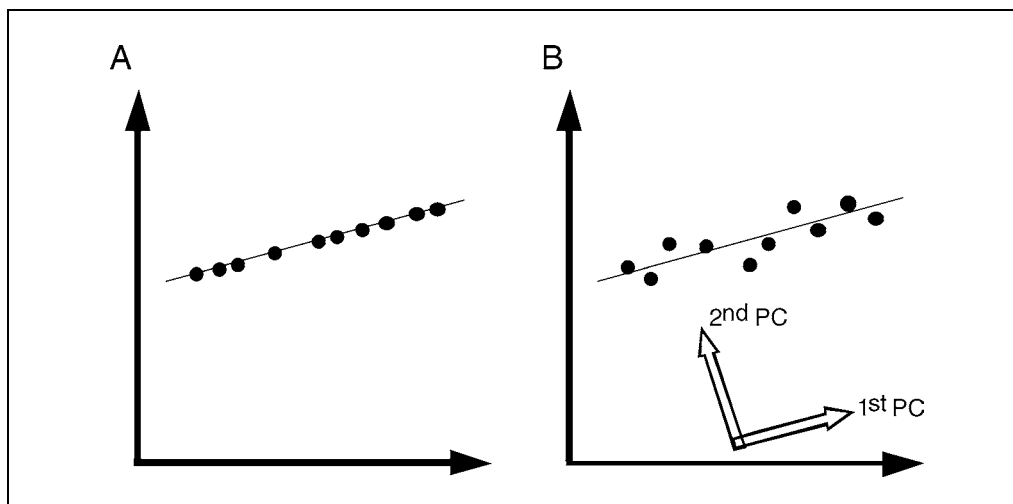


Figure 22.5.9 The idea of PCA. (A) If all the data points are aligned along a line, one dimension along the line is sufficient to describe the relationships among the data points. (B) If the data points are almost aligned along a line, the dimension along the line can describe most of the relationship information among the data points. This dimension is the one defined by the first PC. The remaining differences can be explained by adding another dimension, which is seen as the second PC. Filled circles represent data points.

difference among the data points is still not accounted for by the first x PCs. If increasing the number of PCs beyond a certain number does not reduce the residual variance much, the remaining variance may be caused by random noise. If the first two PCs are sufficient to describe most differences, all the data points can be well approximated in the 2-D space defined by the first two PCs. In this way, all the data points are plotted in the 2-D space, and the relationships among the data points can be easily visualized.

PCA and singular value decomposition (SVD) used in expression profile analysis are essentially the same. SVD is the name of a method that can be used to determine PCs.

Self-organizing maps (SOM) as a nonlinear dimensionality reduction method

Biological processes are typically nonlinear and, therefore, it is reasonable to expect that simple rules to explain major trends in expression profiles are better described in a nonlinear space. For example, imagine a group of genes that are induced by a particular chemical. When using a different concentration of the chemical, expression of different genes may hit a plateau at different concentrations. Using a linear dimensionality reduction, a description of the profiles for samples with different concentrations of the chemical using these genes as parameters could require several dimensions. However, a nonlinear method may

find that just one nonlinear dimension would be sufficient.

In PCA, a linear space with a small number of dimensions is assumed to be sufficient to describe most differences among the data points. What if the data points lie mostly on a 1-D space but it is nonlinear (Fig. 22.5.10A)? Two dimensions would be needed according to PCA. For this case, SOM (Tamayo et al., 1999) works well. Initially, nodes are placed in a straight line (1-D linear space), which is typically defined by the first PC (Fig. 22.5.10B). Then, the node positions and the assignments of data points to each node are iteratively optimized. This iterative optimization process is conceptually similar to the process used in K -means clustering. When the node positions and the data point assignments are optimized, nodes are placed in the curve, and each data point is assigned to the closest node (Fig. 22.5.10C). So, groups of data points can be mapped back to the initial positions of the nodes on the line (Fig. 22.5.10D). In other words, the curve defined by the data points can be approximately converted into a straight line by the SOM procedure.

Typically, the initial nodes in SOM are placed in a 2-D linear space (plane) defined by the first two PCs, instead of a 1-D linear space. If the data points define a 2-D nonlinear space in a high-dimensional space, SOM could closely approximate the data points in a plane. Essentially, SOM performs nonlinear dimensionality reduction.

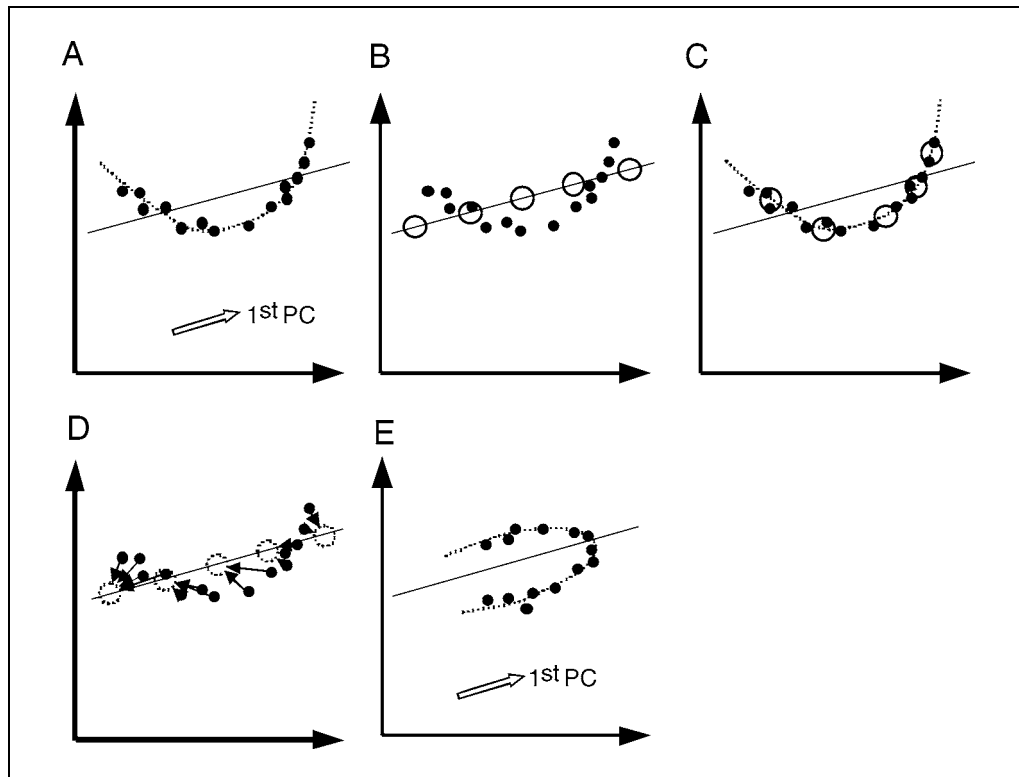


Figure 22.5.10 SOM works as a nonlinear dimensionality reduction method. (A) The data points are aligned along a curve instead of a straight line. (B) The nodes can be placed initially along the first PC. (C) The nodes are aligned along the curve after optimization. (D) Using the initial position of the nodes, each data point can be mapped onto the straight line defined by the first PC. (E) If the 1-D space defined by the data points has a high level of nonlinearity, such as in this case, SOM does not perform well as a nonlinear dimensionality reduction method. Small filled circles represent data points and large open circles represent nodes.

In principle, nodes can be initially placed in a space with three dimensions or more. However, this makes the visualization of relationships among nodes and the profiles assigned to them more difficult, especially when there are more than three dimensions. In addition, if the nonlinearity of the space defined by the data points is too high, SOM's principle of approximation starting with a linear space would not work well (Fig. 22.5.10E).

SOM is also known as a clustering method. When it is used as a clustering method, it is useful for the discovery of well-separated, round-shaped clusters (Xu et al., 2002).

When a large number of nodes are used for SOM, the data points can be finely mapped on the initial linear 2-D space. Such a fine map generated by having genes as profiles can be used as a visualization matrix for the entire profile of an experimental sample by color-coding expression values for a representative of each node (i.e., genes are assigned to nodes; Eichler et al., 2003). This SOM-based visualization method can be considered as a 2-D version of the 1-D profile visualization method

using hierarchical clustering described in the section above. The SOM-based method could work better for visualization of more information in the profile than is possible using a method based on hierarchical clustering. Note that the numbers of genes assigned to different nodes (i.e., the density of genes) in the SOM-based visualization method are not homogenous, whereas the density of genes along the axis in the hierarchical clustering-based method is homogenous. It is possible that a change that affects a large area in the SOM-based visualization matrix could be caused by a relatively small number of genes.

Nonlinear dimensionality reduction

To select important information, PCA linearly reduces dimensionality, and SOM can be interpreted as a modestly nonlinear approximation of PCA with a fixed number of reduced dimensions. What if the space defined by the data points is highly nonlinear (e.g., Fig. 22.5.10E)? In such a case, nonlinear dimensionality reduction could be used (Roweis and Saul, 2000; Tenenbaum et al., 2000). The

authors have been developing a method for microarray data analysis using nonlinear dimensionality reduction (Katagiri and Glazebrook, 2003). They have also proposed to translate the result of nonlinear dimensionality reduction into a network structure, so that network analysis methods can be used when more than three dimensions are required to explain particular data. This type of approach may prove useful.

LITERATURE CITED

- Alter, O., Brown, P.O., and Botstein, D. 2000. Singular value decomposition for genome-wide expression data processing and modeling. *Proc. Natl. Acad. Sci. U.S.A.* 97:10101-10106.
- Batagelj, V. and Mrvar, A. 1998. Pajek - Program for large network analysis. *Connections* 21:47-57.
- Davidson, G.S., Hendrickson, B., Johnson, D.K., Meyers, C.E., and Wylie, B.N. 1998. Knowledge mining with VxInsight: Discovery through interaction. *J. Intell. Inf. Syst.* 11:259-285.
- Eichler, G.S., Huang, S., and Ingber, D.E. 2003. Gene expression dynamics inspector (GEDI): For integrative analysis of expression profiles. *Bioinformatics* 19:2321-2322.
- Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* 95:14863-14868.
- Herwig, R., Poustka, A.J., Muller, C., Bull, C., Lehrach, H., and O'Brien, J. 1999. Large-scale clustering of cDNA fingerprinting data. *Genome Res.* 9:1093-1105.
- Holter, N.S., Mitra, M., Maritan, A., Cieplak, M., Banavar, J.R., and Fedoroff, N.V. 2000. Fundamental patterns underlying gene expression profiles: Simplicity from complexity. *Proc. Natl. Acad. Sci. U.S.A.* 97:8409-8414.
- Katagiri, F. and Glazebrook, J. 2003. Local context finder (LCF) reveals multidimensional relationships among mRNA expression profiles of *Arabidopsis* responding to pathogen infection. *Proc. Natl. Acad. Sci. U.S.A.* 100:10842-10847.
- Kim, S.K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J.M., Eizinger, A., Wylie, B.N., and Davidson, G.S. 2001. A gene expression map for *Caenorhabditis elegans*. *Science* 293:2087-2092.
- Roweis, S.T. and Saul, L.K. 2000. Nonlinear dimensionality reduction by locally linear embedding. *Science* 290:2323-2326.
- Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E.S., and Golub, T.R. 1999. Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. U.S.A.* 96:2907-2912.
- Tenenbaum, J.B., de Silva, V., and Langford, J.C. 2000. A global geometric framework for nonlinear dimensionality reduction. *Science* 290:2319-2323.
- Xu, Y., Olman, V., and Xu, D. 2002. Clustering gene expression data using a graph-theoretic approach: An application of minimum spanning trees. *Bioinformatics* 18:536-545.

Contributed by Fumiaki Katagiri and
Jane Glazebrook
University of Minnesota
St. Paul, Minnesota

CHAPTER 23

Manipulating the Mouse Genome

INTRODUCTION

Defining the role of a mammalian nucleic acid segment can be a difficult process. The first steps involve mRNA identification, gene cloning, and nucleic acid sequence determination. A variety of cell-free and cultured cell systems have been developed (see Chapters 8 to 22, 24) that can provide important clues to the function of a particular DNA sequence in regulating gene transcription or RNA translation. Many segments of DNA appear inert or function inappropriately in these model systems; thus, an important test of any functional model of a DNA segment is to characterize the phenotype of a viable organism carrying the sequence in its genome. Today the tools exist for manipulating DNA segments in a variety of mammals, from mice and rats to sheep and cows. However, the mouse, because of its small size (30 g) and short life cycle (about 8 weeks) from birth to the production of progeny, has been selected for the vast majority of laboratory experiments. Thus, this chapter focuses on the manipulation of the mouse genome.

Two general methods are available for the introduction and modification of mouse genomic DNA sequence. Embryonic stem cell–derived mice are produced by the introduction of embryonic stem cells into the blastocyst, which is reimplanted in a foster mother, where it goes through normal mouse development, producing a mouse pup (*UNITS 23.1–23.8*). Transgenic mice are produced by the injection of one or more transgenes (usually a DNA segment bearing its own promoter) into the pronucleus of a fertilized mouse oocyte, which, after reimplantation in a foster mother, gives birth to a transgenic mouse bearing one to several hundred copies of the transgene (*UNIT 23.9*). Transgenic mice are useful for testing the overexpression of a gene segment while embryonic stem–cell derived mice are generally useful for defining the phenotype of mice lacking a gene segment. Both types of mice can be used to express altered protein sequences; transgenic mice will usually overproduce the altered protein, while embryonic stem cell–derived mice will usually express normal levels of protein. In this chapter we describe the production of embryonic stem cell derived–mice and transgenic mice.

The production of transgenic mice involves two steps: (1) production of DNA suitable for injection and (2) injection of the DNA into a fertilized mouse oocytes and re-implantation of the injected oocyte into a foster mother. A variety of approaches are used to produce injected DNA. Perhaps the most common DNA constructs contain a eukaryotic promoter, cDNA sequence, introns and 3' untranslated region. The production of these DNA segments involves the tools that of molecular cloning (Chapters 4 and 5). Producing transgenic mice is faster than producing mice by homologous recombination because DNA is injected directly into oocytes rather than first being introduced into embryonic stem cells. However, the injection of DNA into fertilized mouse oocytes and re-implantation is technically more demanding than electroporation of DNA into cells. Nanoliter volumes containing transgene DNA are injected with a micropipette into a pro-nucleus (*UNIT 23.9*). After injection the oocyte is re-implanted into the foster mother. Despite the technical difficulties of producing transgenic mice this technology allows for rapid evaluation of the effects of over-expression of a mutationally altered or wildtype protein.

Contributed by J.G. Seidman

Current Protocols in Molecular Biology (2004) 23.0.1–23.0.2
Copyright © 2004 by John Wiley & Sons, Inc.

Manipulating
the Mouse
Genome

23.0.1

Supplement 68

The production of embryonic stem cell–derived mice can be thought of as a three-stage process that normally takes about 6 months. These steps involve the production of a targeting vector (*UNIT 23.1*), the introduction of DNA sequences into the embryonic stem cell genome by homologous recombination (*UNIT 23.5*) and eventually to the production of genetically altered mice derived from embryonic stem cells. A critical part of this procedure is the maintenance of the cells. Embryonic stem cells are cultured on feeder cells and require constant attention; these procedures are described in *UNITS 23.2 & 23.3*. A procedure describing the derivation of embryonic stem cells is described in *UNIT 23.4*. After the genetically engineered embryonic stem cells are obtained they are injected into mouse blastocysts, which are then re-implanted into a foster mother that eventually gives birth to chimeric mice composed of wild-type cells and ES-derived cells (*UNIT 23.7*). Chimeric mice are then mated to wild-type mice and eventually mice derived from the ES cell genome are obtained. The genetically engineered mice are now available for physiologic and biochemical evaluation. The production and management of a colony of genetically engineered mice requires careful planning (*UNIT 23.8*) in order to have adequate numbers of mice of appropriate ages for study.

The function of genes expressed in embryonic stem cells can be assessed directly in these cells. In *UNIT 23.5*, procedures for producing a single knockout or heterozygous cell line by homologous recombination are described. Homozygous (double-knockout) cells can be produced in tissue culture and these cells can be a useful tool for analyzing gene function in embryonic stem cells (*UNIT 23.6*). Taken together, these steps provide a powerful method for introducing a wide variety of specific changes into murine embryonic stem cells and eventually into the genome of a mouse.

J.G. Seidman

Overview of Gene Targeting by Homologous Recombination

The analysis of mutant organisms and cell lines has been important in determining the function of specific proteins. Gene targeting by homologous recombination in mammalian systems enables the production of mutants in any desired gene (Mansour, 1990; Robertson, 1991; Zimmer, 1992). Over the past two decades, gene targeting by homologous recombination has become a mainstay in gene-function investigation. This technology can be used to produce mutant mouse strains and mutant cell lines. Because most mammalian cells are diploid, they contain two copies, or alleles, of each gene encoded on a pair of autosomal (nonsex) chromosomes. In most cases, both alleles must be inactivated to produce a discernible phenotypic change in a mutant. The conversion from heterozygosity to homozygosity is accomplished by breeding in the case of mouse strains and by direct selection in cell lines.

More recent technological advances that have greatly increased the power to investigate gene function include spatial and temporal control of knockouts, mutation analysis, heterologous expression of proteins by “knock in” of coding regions into another gene, and recombineering for targeting construct production. Bacteriophage recombinase systems such as Cre with its recognition sequence, *loxP*, have played key roles in these technologies. Another recombinase system, the FLP/FRT system, can also be used (Fiering et al., 1993, 1999; Branda and Dymecki, 2004). The control can function along actual spatial coordinates when a viral gene transfer system is used, or in a cell- or tissue-specific fashion when restricted promoters are employed. Adding temporal regulation of Cre, as with the tetracycline-regulatable system (UNIT 16.14) or the modified Cre-ER fusion, allows temporal control as well.

To produce a mutant mouse strain by homologous recombination, two major elements are needed. An embryonic stem (ES) cell line capable of contributing to the germ line, and a targeting construct containing target-gene sequences with the desired mutation. ES cell lines are derived from the inner cell mass of a blastocyst-stage embryo. Maintaining ES cells in their undifferentiated state is a major task

during gene targeting (UNIT 23.3). This usually is accomplished by growing cells on a layer of feeder cells (UNIT 23.2), which supply factors, most notably leukemia inhibitory factor (LIF), to keep cells undifferentiated. The targeting construct is then transfected into cultured ES cells (see UNIT 23.5). Homologous recombination occurs in a small number of the transfected cells, resulting in introduction of the mutation in the targeting construct into the target gene. Once identified, mutant ES cell clones can be microinjected into a normal blastocyst in order to produce a chimeric mouse. Because many ES cell lines retain the ability to differentiate into every cell type present in the mouse, the chimera can have tissues, including the germ line, with contribution from both the normal blastocyst and the mutant ES cells. Breeding germ-line chimeras yields animals that are heterozygous for the mutation introduced into the ES cell, and these offspring can be interbred to produce homozygous mutant mice.

Homologous recombination can also be used to produce homozygous mutant cell lines (see UNIT 23.6). Previously, inactivation of both alleles of a gene required two rounds of homologous recombination and selection (te Riele et al., 1990; Cruz et al., 1991; Mortensen et al., 1991). Now, however, inactivation of both alleles of many genes requires only a single round of homologous recombination using a single targeting construct (Mortensen et al., 1992). The homozygous mutant cells can then be analyzed for phenotypic changes to determine the function of the gene.

SOURCE OF DNA

Although nonisogenic DNA was used for constructs early in the history of homologous recombination, isogenic DNA (from the same mouse strain) gives higher recombination rates because homologous recombination requires stretches of exact base pair matches. Bacterial artificial chromosomes (BACs) from C57BL/6 and 129 strains are now available from genome repositories, making it more convenient to obtain DNA for constructs. Manipulations of large constructs is also possible using Red/ET recombineering methodology (Copeland et al., 2001).

Contributed by Richard Mortensen

Current Protocols in Molecular Biology (2006) 23.1.1-23.1.12
Copyright © 2006 by John Wiley & Sons, Inc.

ANATOMY OF TARGETING CONSTRUCTS

Two basic configurations of constructs are used for homologous recombination—insertion constructs and replacement constructs (Fig. 23.1.1). Each can be used for different purposes in specific situations, as discussed below. The insertion construct is rarely used. It contains a region of homology to the target gene cloned as a single continuous sequence, and is linearized by cleavage of a unique restriction site within the region of homology. Homologous recombination introduces the insertion construct sequences into the homologous site of the target gene, interrupting normal target-gene structure by adding sequences. As a result, the normal gene can be regenerated from the mutated target gene by an intrachromosomal recombination event.

The replacement construct is the more commonly used construct. It contains two regions of homology to the target gene located on either side of a mutation (usually a positive selectable marker; see below). Homologous recombination proceeds by a double cross-over event that replaces the target-gene sequences with the replacement-construct sequences. Because no duplication of se-

quences occurs, the normal gene cannot be regenerated.

METHODS OF ENRICHMENT FOR HOMOLOGOUS RECOMBINANTS

Positive Selection by Drug-Resistance Gene

Nearly all constructs used for homologous recombination rely on the positive selection of a drug-resistance gene (e.g., neomycin or *neo*) that is also used to interrupt and mutate the target gene. When either insertion or replacement constructs are linearized, the drug-resistance gene is flanked by two regions of homology to the target gene. Selection of the cells using drugs (e.g., G418) eliminates the great majority of cells that have not stably incorporated the construct (see UNIT 9.5). However, in many of the surviving clones the construct has incorporated into the genome not by homologous recombination but rather through random integration. Therefore, methods to enrich for homologous recombinant clones have been developed.

Positive-Negative Selection

The most commonly used method for eliminating cells in which the construct integrated into the genome randomly, thus further

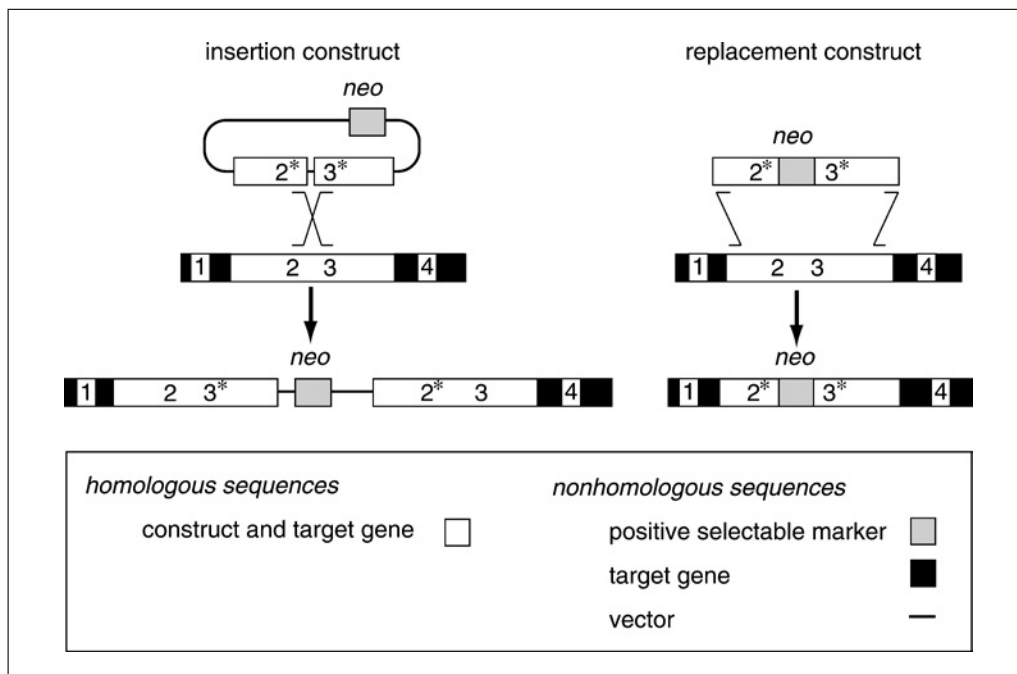


Figure 23.1.1 Two configurations of constructs used for homologous recombination. Numbers indicate target-gene sequences in the genome. An asterisk indicates homologous target-gene sequences in the construct. Replacement constructs substitute their sequences (2*, *neo*, and 3*) for the endogenous target-gene sequences (2 and 3). Insertion constructs add their sequences (2*, *neo*, and 3*) to the endogenous target gene, resulting in tandem duplication and disruption of the normal gene structure.

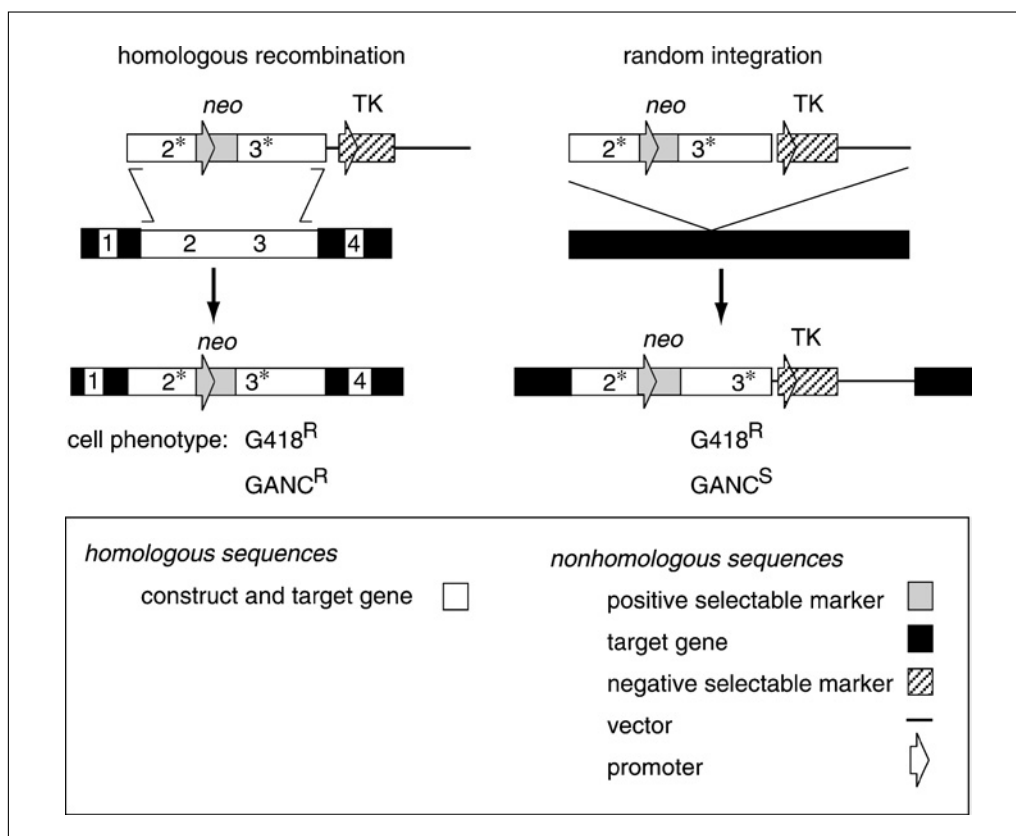


Figure 23.1.2 Enrichment for homologous recombinants by positive-negative selection using the TK gene. Homologous recombination involving cross-overs on either side of the *neo* gene results in loss of the TK gene. Random integration tends to preserve the TK gene. The presence of TK can be selected against because any cell expressing the gene will be killed by gancyclovir (GANC). Although both homologous recombinants and clones in which the construct integrated randomly are G418-resistant, only homologous recombinants are gancyclovir-resistant. The construct is shown linearized so that the plasmid vector sequences remain attached to the TK gene. This configuration helps preserve the integrity of the TK gene.

enriching for homologous recombinants, is known as positive-negative selection. It is only applicable to replacement constructs (Fig. 23.1.2; Mansour et al., 1988). In these constructs, a negative selectable marker (e.g., herpes simplex virus thymidine kinase, HSV-TK) is included outside the region of homology to the target gene. In the presence of the TK gene, the cells are sensitive to acyclovir and its analogs (e.g., gancyclovir, GANC). The HSV-TK enzyme activates these drugs, resulting in their incorporation into growing DNA, causing chain termination and cell death. During homologous recombination, sequences outside the regions of homology to the target gene are lost due to crossing over. In contrast, during random integration all sequences in the construct tend to be retained because recombination usually occurs at the ends of the construct. The presence of the TK gene can be selected against by growing the cells in gancyclovir; the homologous recombinants will be G418-resistant and gancyclovir-resistant,

whereas clones in which the construct integrated randomly will be G418-resistant and gancyclovir-sensitive. In some cases, TK is inactivated without homologous recombination; thus, the gancyclovir-resistant clones must be screened to identify the true homologous recombinants. Other markers that are lethal to cells have also been used (e.g., diphtheria toxin; Yagi et al., 1990).

Endogenous Promoters

Constructs that rely on an endogenous promoter to express the positive selectable marker can also give enrichment of homologous recombinants (Fig. 23.1.3), but can only be used if the gene of interest is expressed in the cell line. They contain the coding region of a selectable marker (e.g., *neo*) but lack a promoter for the marker. The coding sequence for the marker usually interrupts, and is in frame with, an exon of the target gene. Thus, when homologous recombination occurs, a fusion protein is produced driven by the endogenous

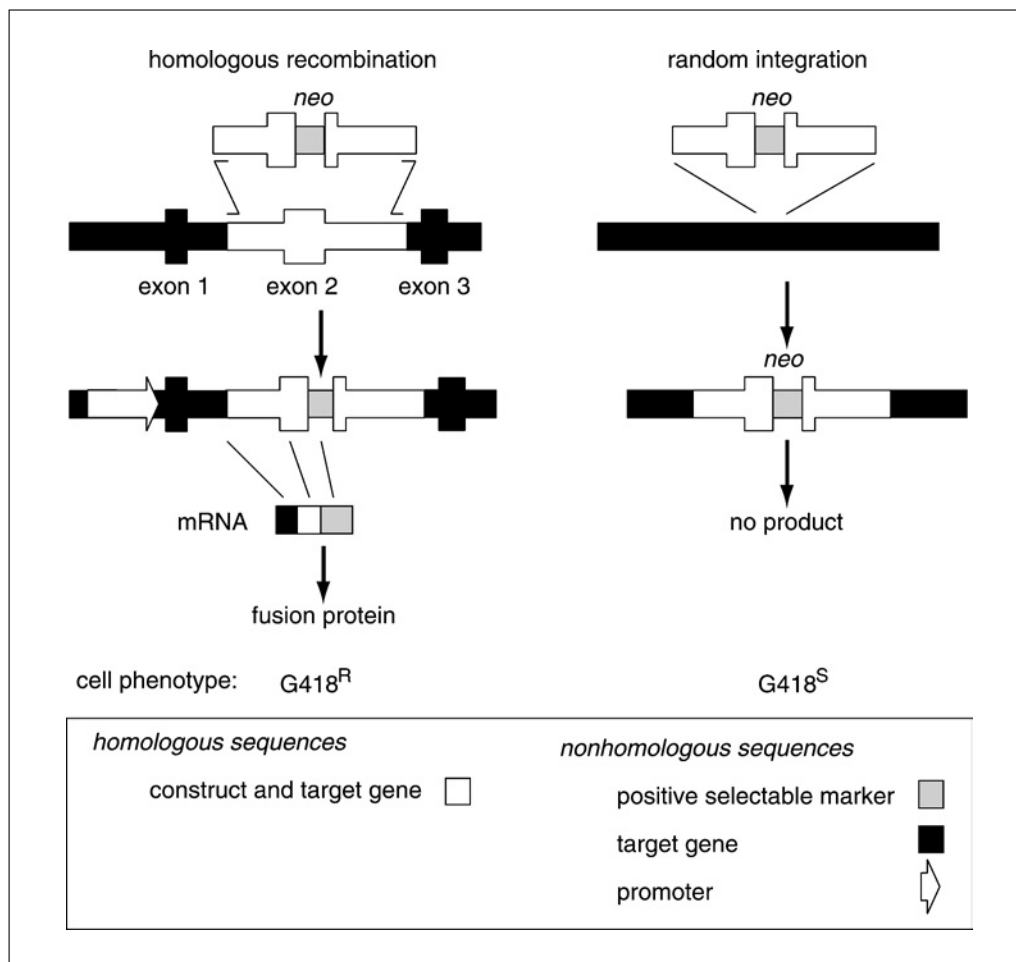


Figure 23.1.3 Enrichment for homologous recombinants using a positive selectable marker (*neo*) lacking a promoter. Clones in which integration of the construct provides an endogenous promoter to drive *neo* expression will be G418-resistant. The construct is designed so that homologous recombination will provide a promoter leading to *neo* expression, whereas random integration will most likely not provide a promoter, thus precluding *neo* expression.

target-gene promoter. In contrast, when random integration occurs, the selectable-marker protein is not usually produced. Therefore, homologous recombinants are G418-resistant, whereas cells in which the construct integrated randomly are G418-sensitive. Constructs containing a promoterless selectable marker can be generated in either replacement or insertion structure and can result in dramatic enrichment for homologous recombinants.

TYPES OF MUTATIONS

Gene Inactivation

Homologous recombination has most often been used to completely inactivate a gene (commonly termed “knockout”). Usually, an exon encoding an important region of the protein (or an exon 5' to that region) is interrupted by a positive selectable marker (e.g., *neo*), preventing the production of normal mRNA from the target gene and resulting in inactivation of the gene.

A gene may also be inactivated by creating a deletion in part of a gene, or by deleting the entire gene. By using a construct with two regions of homology to the target gene that are far apart in the genome, the sequences intervening the two regions can be deleted. Up to 15 kb have been deleted in this way; thus, many genes could be completely eliminated (Mombaerts et al., 1991). Gene inactivations may also be controlled using the *Cre/loxP* recombinase system either spatially, through cell- or tissue-specific knockout, or temporally, through control of the activity or expression of the recombinase (see *Cre/loxP* System, below).

Mutations can be introduced that have multiple purposes. Homologous recombination has been used to introduce a replacement construct containing the coding sequence of β -galactosidase in frame with the 5' end of the target gene. Downstream of the *lacZ* gene is a positive selectable marker driven by a

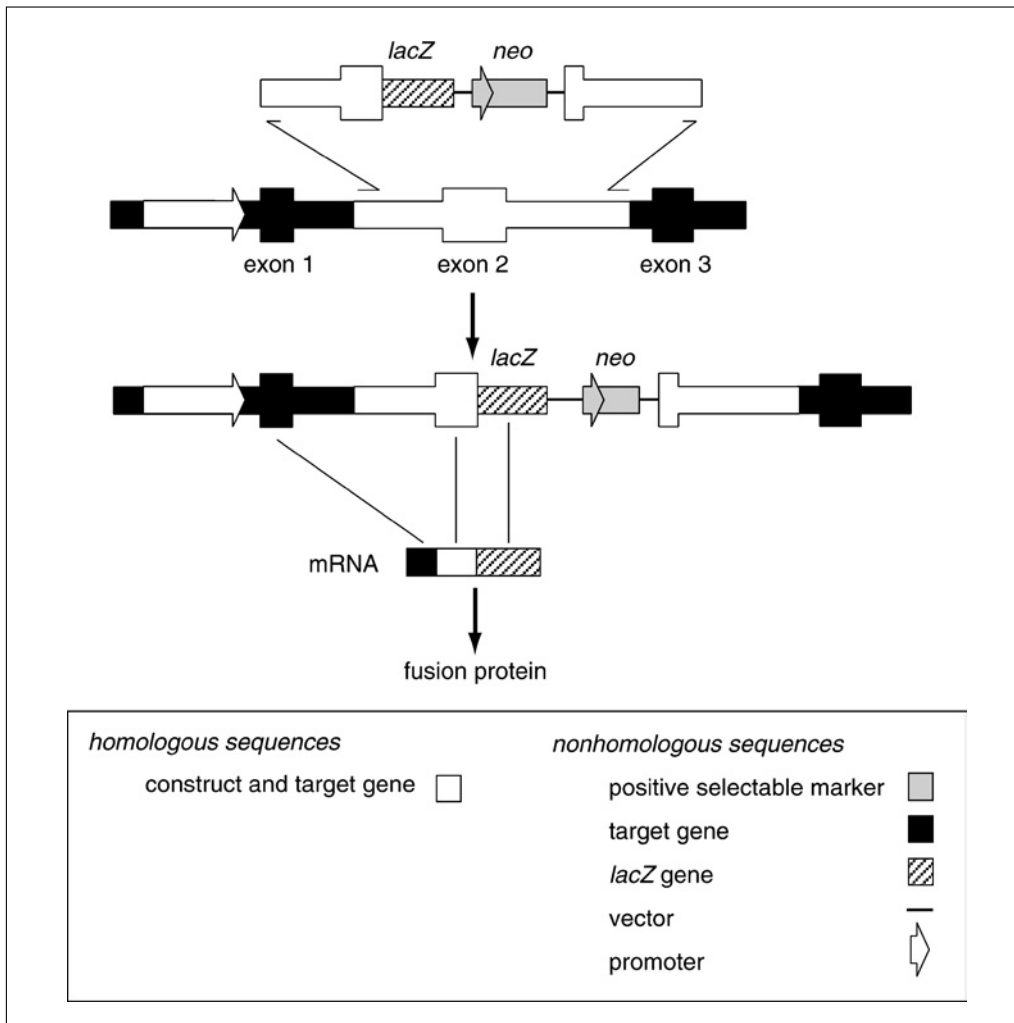


Figure 23.1.4 *LacZ* reporter construct for gene targeting. This construct has two purposes: to disrupt the target gene and to express the *lacZ* gene as a marker to monitor activity of the endogenous target gene's promoter.

heterologous promoter (Fig. 23.1.4). This construct not only disrupts target-gene function but also expresses a fusion protein with β -galactosidase activity, and thus can be used to monitor the activity of the endogenous gene's promoter in various tissues during development (Mansour et al., 1990). Targeting vectors can also be made that express β -gal without the fusion protein. Similarly, other coding sequences can be expressed from the endogenous promoter to achieve selective heterologous expression. Since this results in disruption of the gene, this approach is typically used only if the heterozygous knockout does not have a phenotype.

Subtle Gene Mutations

Homologous recombination can also be used to introduce subtle mutations in a gene. One method is analogous to a method in yeast called transplacement or allele replacement (UNIT 13.10). It is called "hit and run" (Hasty

et al., 1991) because duplications are introduced into the target gene and then removed. An insertion construct containing both positive and negative selectable markers (e.g., *neo* and TK) is used to introduce a duplication that contains a subtle mutation, such as a point mutation, into the target gene sequence (Fig. 23.1.5). After selection for integration of the construct using the positive selectable marker (e.g., G418), homologous recombinants are identified by screening. A homologous recombinant clone is cultured and then the presence of the negative selectable marker is selected against (e.g., selection against TK using gancyclovir). This selects for an intrachromosomal recombination that eliminates the target-gene duplications and the selectable markers but leaves the mutant target-gene sequences substituting for the normal target-gene sequences. Surviving clones are screened for the correct intrachromosomal rearrangements, leaving the desired mutation. A second

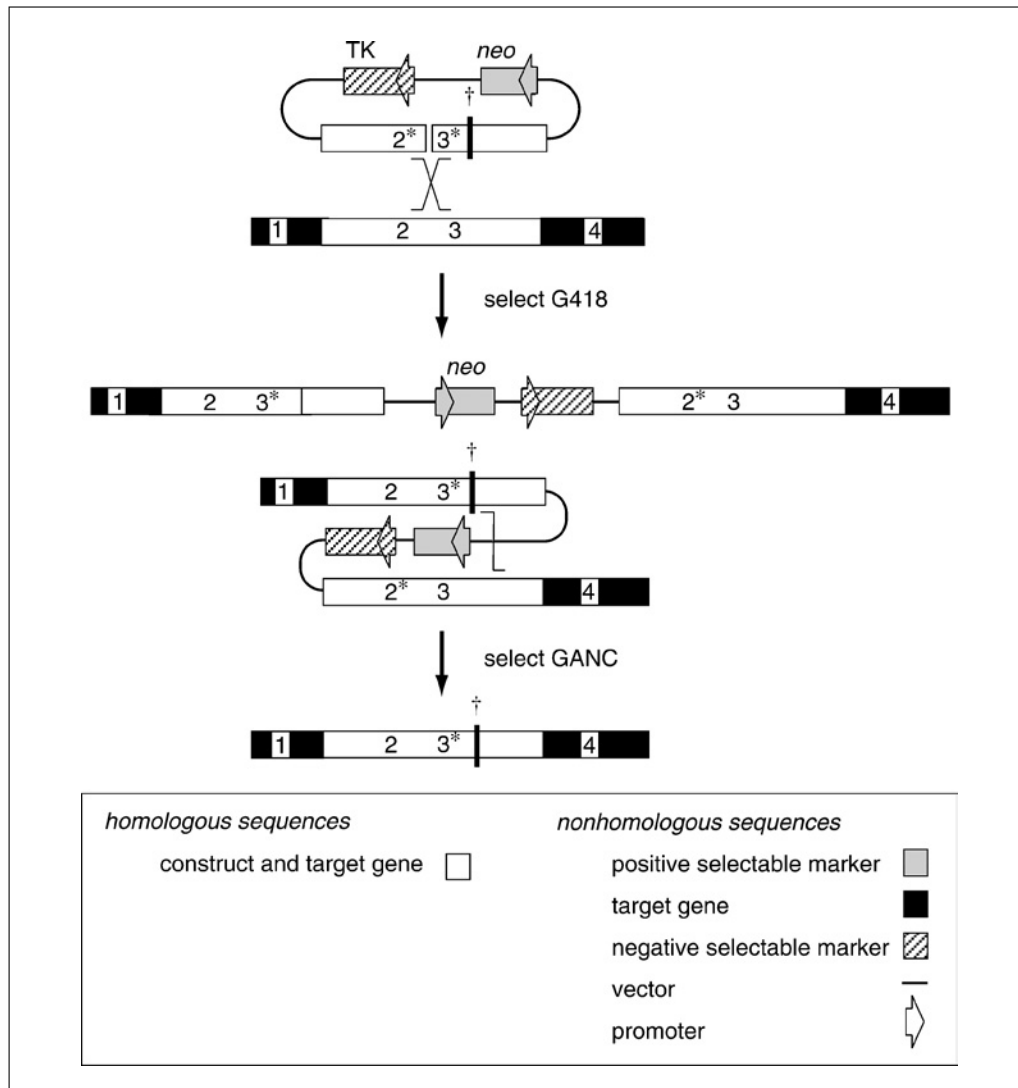


Figure 23.1.5 Allele replacement “hit and run.” Cells are cultured in G418 to select for the integration of *neo*, then homologous recombinants are identified by screening and are cultured in gancyclovir (GANC) to select against the presence of the TK gene. This strategy may yield a reconstituted gene containing the subtle mutation present in the construct (indicated by the dark bar and †). Because intrachromosomal recombination may result in the loss of the subtle mutation, its presence must be verified (e.g., by a change in restriction site).

method of introducing subtle mutations into a gene is to insert the mutation by homologous recombination and then use the *Cre/loxP* system to remove the selectable marker.

Cre/loxP SYSTEM

The *Cre/loxP* system is derived from the bacteriophage P1. The recombinase Cre acts on the DNA site *loxP*. If there are two *loxP* sites in the same orientation near each other, Cre can act to loop out the sequence between the two sites, leaving a single *loxP* site in the original DNA and a second *loxP* in a circular piece of DNA containing the intervening sequence. Therefore, a properly designed targeting construct containing *loxP* sites can be used for introducing subtle mutations or for a tem-

porally or spatially controlled knockout (for a review of the control of transgenes, see Sauer, 1993). Other recombinase systems, such as the *Flp/FRT* system, can be similarly useful (Fiering et al., 1995; Vooijs et al., 1998).

Removing the Positive Selectable Marker

Although many gene-inactivation approaches involving homologous recombination still use constructs that leave the positive selectable marker in the genomic DNA, it has become increasingly clear that this can cause a number of unanticipated effects. For example, the presence of the *neo* gene, often with its own promoter, can alter the expression of neighboring loci (Olson et al., 1996; Pham

et al., 1996). This can be a particular problem in gene clusters where neighboring genes are in the same family, since the genes affected may have similar or identical functions. As a result, slight differences in targeting constructs have led to marked differences in phenotype.

If the targeting construct includes *loxP* sites flanking the *neo* gene, then *neo* can be removed after targeting by transient expression of the Cre recombinase (as discussed in UNIT 9.5 and Fig. 23.1.3). This will leave the small *loxP* site in the genomic DNA, but the construct can be engineered so that this is in an innocuous location, such as an intron. Although theoretically even a *loxP* site could cause alterations in the expression of neighboring genes, no such cases have yet been reported. The efficiency of Cre recombination from transient expression reported in the literature varies widely, from ~2% to ~15% (Sauer and Henderson, 1989; Abuin and Bradley, 1996). This rate should be

distinguished from the efficiency of Cre recombination *in vivo*, where the expression of Cre is derived from sequences integrated into the genome and therefore will show longer-lasting expression in nearly all cases.

Introduction of Subtle Mutations Using Cre/*loxP*

The strategy described in the previous section involves introducing subtle mutations by first duplicating sequences and then screening for intrachromosomal recombination that removes the redundant sequences and leaves the mutation. A limitation to this approach is that the second homologous recombination event occurs only infrequently. A more efficient method is to use a replacement construct containing the subtle mutation and then remove the positive selectable marker, which is flanked by *loxP* sites, using the Cre recombinase system (Fig. 23.1.6). This is

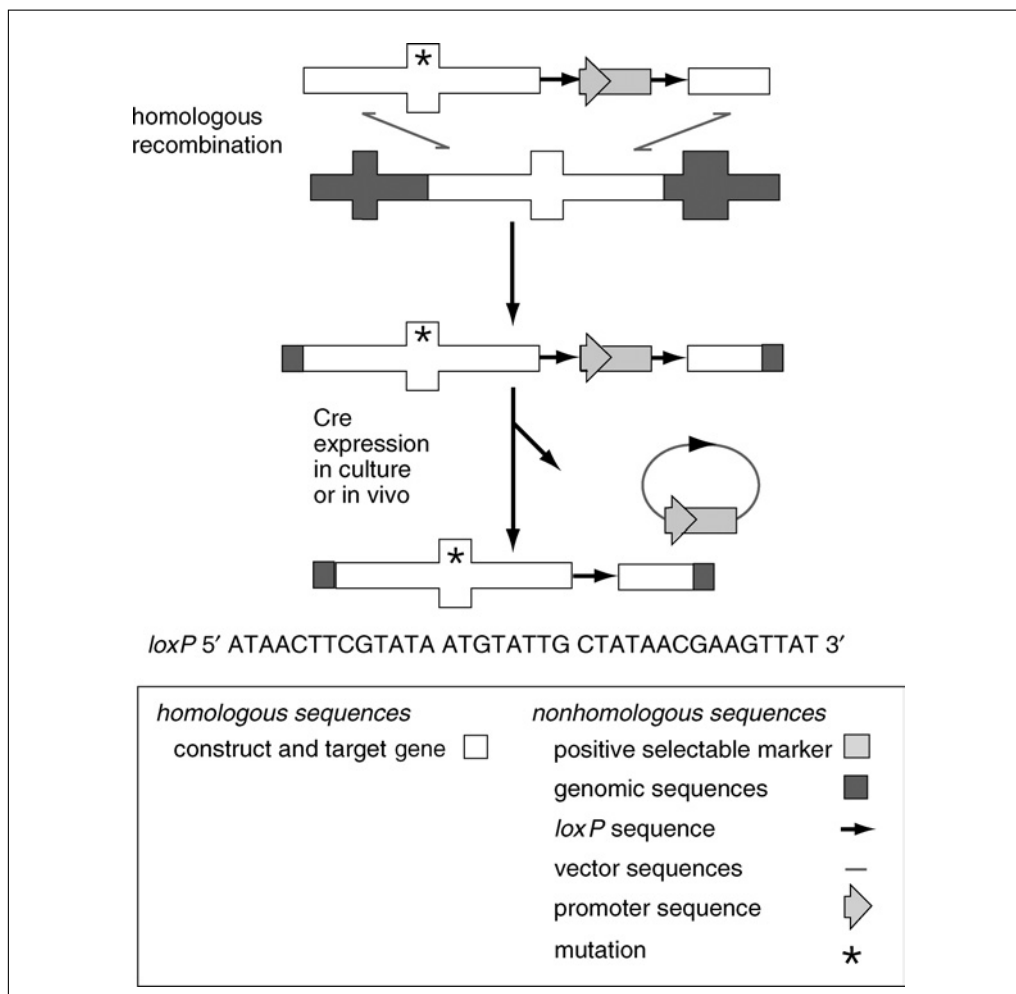


Figure 23.1.6 Using the Cre/*loxP* system to introduce subtle mutations. The subtle mutation is introduced along with the selectable marker in the targeting vector. The selectable marker is then removed by transient expression of Cre, which leaves only the subtle mutation and the small *loxP* site in a silent location.

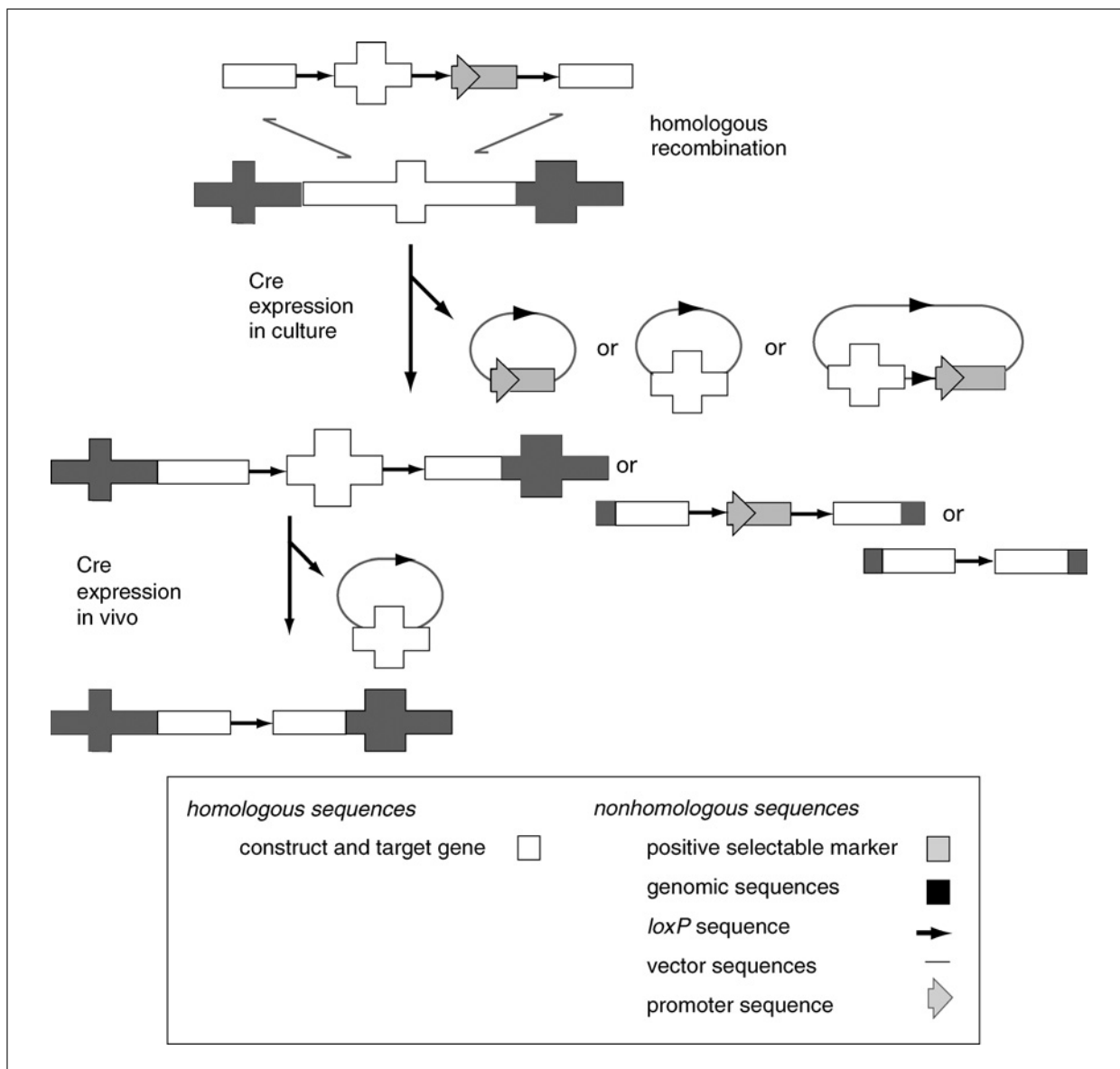


Figure 23.1.7 Conditional gene targeting using the *Cre/loxP* system. The targeting vector contains three *loxP* sites that flank the regions of the gene to be removed and the positive selectable marker *neo*. After homologous recombination is obtained, the selectable marker is excised from the gene by transient expression of Cre. The correct recombination is identified by screening by Southern analysis or PCR. The mutant ES cells are then used to produce a transgenic animal, which is finally bred to an animal line expressing Cre under either temporal or spatial control. Cre can also be ubiquitously expressed to obtain a knockout in all tissues, including the germ line.

identical in effect to removing the *neo* locus after gene inactivation, except that instead of an inactive gene, the replaced sequences contain a subtly mutated version. The selectable marker sequences can be removed by transient Cre expression in culture or by breeding animals to a Cre-deleter line (Williams-Simons and Westphal, 1999), provided that the marker does not cause lethality.

Spatial Control of Knockout

Spatially controlled targeted gene inactivations can be performed in two ways. The most

common makes use of cell-specific promoters (sometimes called tissue-specific promoters, even though tissues are actually made of a number of different cell types). This approach begins with the creation of a transgenic animal that expresses Cre in only some cells using a cell-restricted promoter. A second transgenic animal line is then created by homologous recombination that contains *loxP* sites flanking a portion of the gene that is critical for activity, typically important exons (Fig. 23.1.7). Initially there are three *loxP* sites flanking this important gene region and the selectable

marker. After homologous recombination has been verified, Cre is transiently expressed and loops out regions of DNA between pairs of *loxP* sites. The resultant colonies are screened for the desired recombination (loss of the selectable marker but retention of all regions of the gene). Depending on the frequency of recombination at the site, it may be useful to use a construct that contains a negative selectable marker (such as cytosine deaminase in the example shown in Fig. 23.5.2) between the *loxP* sites along with the positive selectable marker (also see UNIT 9.5). In this way cells that have lost the markers can be selected.

While this approach was successful for many genes, the requirement for correct partial Cre recombination led to difficulties with other genes. Although recombination in vitro could go to completion, in other instances, and for unknown reasons, the desired recombination would not take place. Fortunately, use of another recombinase with its own sequence recognition sequences has simplified this technology (Fig. 23.1.8). A construct is created that uses Cre for the conditional mutation and Flp to remove the selectable marker. The removal of the selectable marker can be accomplished by Flp expression in vitro, but also by using a

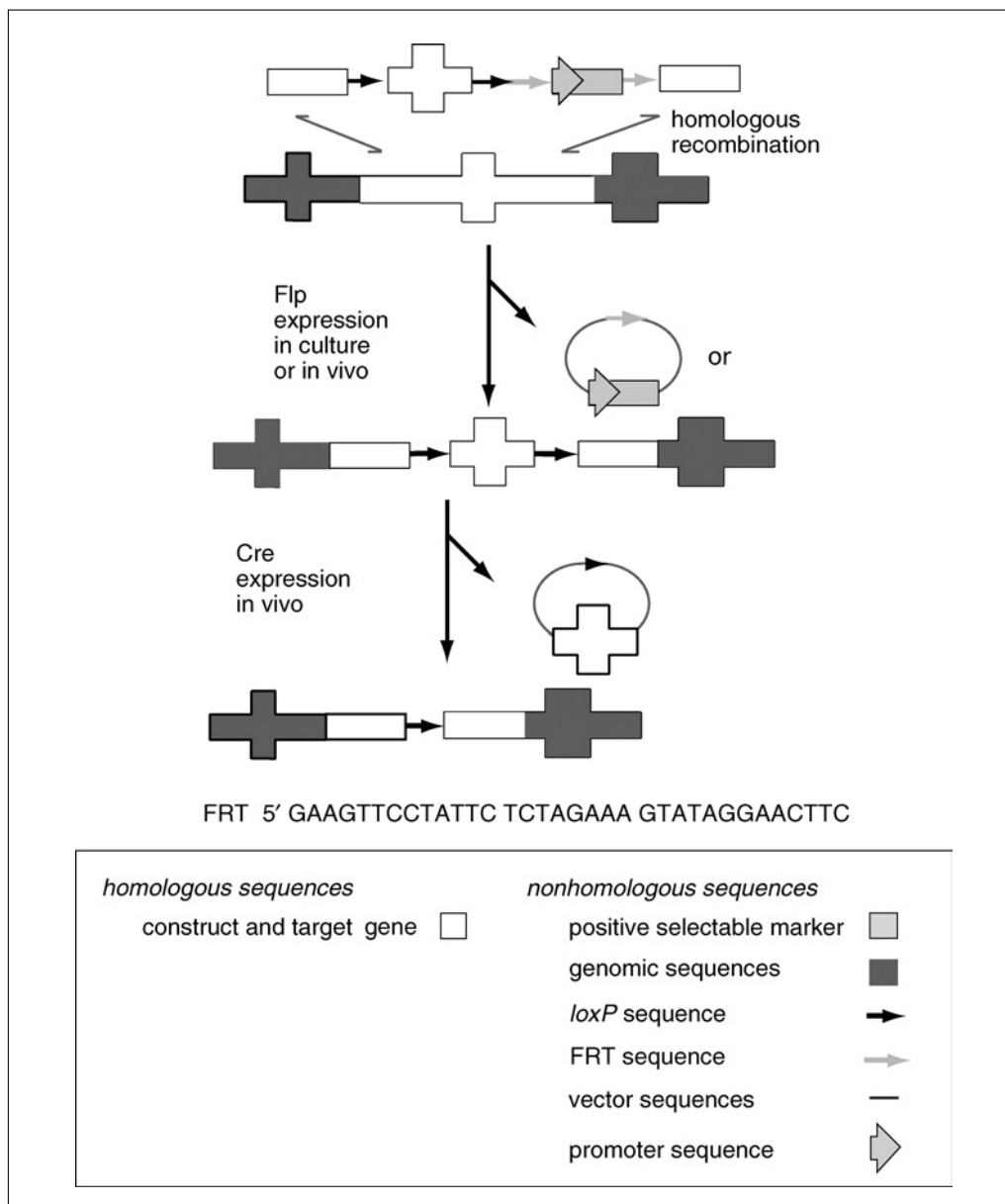


Figure 23.1.8 Alternative strategy for removal of selectable marker using Flp recombinase. The selectable marker is flanked by FRT sites for recognition and removal by the Flp recombinase allowing generation of a gene with only the *loxP* sites remaining. Otherwise the strategy is similar to that in Figure 23.1.7.

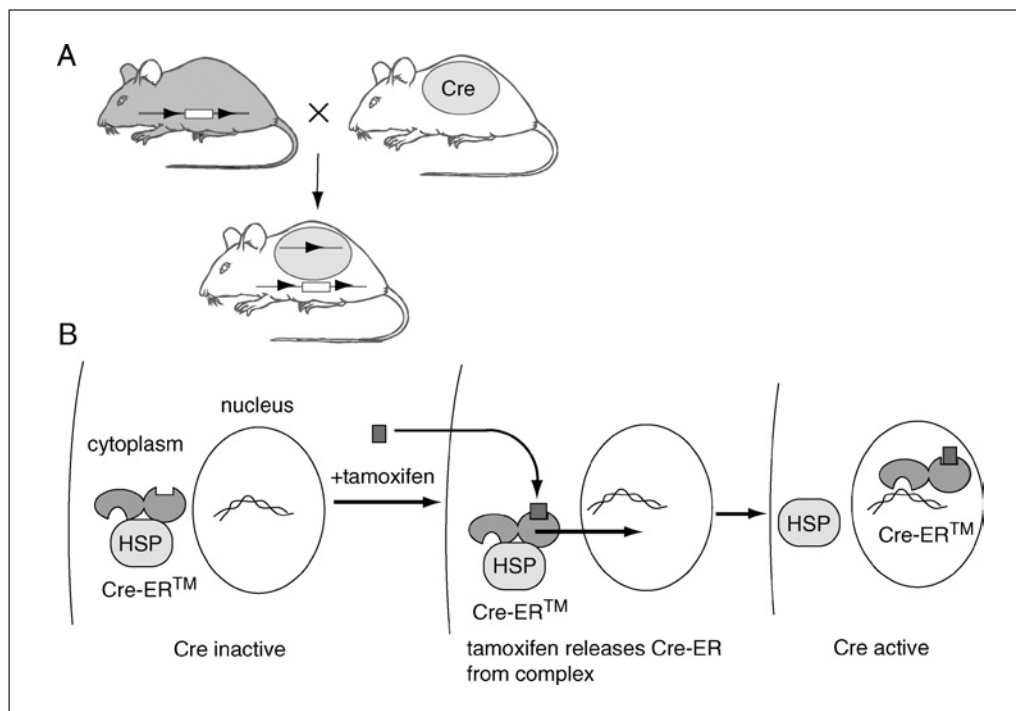


Figure 23.1.9 (A) Spatial control of Cre activity. Cre-mediated cell type–specific knockout relies on the production of a transgenic mouse expressing Cre only in the cell type of interest. By breeding these animals to an animal containing the construct with *loxP* sites flanking a critical region of the gene (called a “floxed” gene, for “flanking *lox*”), recombination occurs only in the cell types expressing Cre. The remainder of cells continue to express the target gene since the *loxP* sites typically do not alter expression. (B) Temporal control of Cre activity. In the absence of tamoxifen, the fusion protein is sequestered in the cytoplasm bound to heat-shock proteins (HSP). When tamoxifen is added, the fusion protein is released and translocates to the nucleus, where it is active for recombination. Abbreviations: ER, estrogen receptor; TM, tamoxifen.

Flp-deleter line (Rodriguez et al., 2000) after germline transmission. In vivo deletion has the advantage of minimizing the manipulation of the ES cell line, which can decrease germline transmission.

The targeted line will have normal expression of the targeted gene, since its only modification is the presence of *loxP* sites in innocuous sites (e.g., introns). When the two lines are bred together, the Cre recombinase will loop out the DNA—inactivating the gene—only in those cells where it is expressed (Fig. 23.1.9A). In this way, tissue-specific knockouts of a number of genes have been generated (Gu et al., 1994; Agah et al., 1997). The method also has the advantage that, once a transgenic line is generated with the desired restricted expression of Cre, the approach can be applied to a number of targeted lines. In addition, it is not necessary to make separate constructs for a restricted and a complete knockout, since Cre-expressing lines (called deleter lines) have been made that will produce re-arrangement in all tissues when bred to the targeted line (Schwenk et al., 1995).

Another way of spatially controlling knock-out is to use an expression system for Cre that can be applied to absolute location. In some cases, no restricted expression pattern is known for a gene that matches the desired spatial alteration; in others, the site may be particularly amenable to viral manipulation (as with an epithelial or endothelial surface) or accessible by direct injection (such as stereotactic injection of the central nervous system). By using a viral vector to express the Cre protein, it is possible to obtain knockouts that are spatially limited by the viral infection. This strategy has been applied to a number of tissues including the brain, liver, colon, and heart (Rohlmann et al., 1996; Wang et al., 1996; Agah et al., 1997; Shibata et al., 1997; van der Neut, 1997).

More refined spatial control has been achieved by using a combination of Cre and Flp in overlapping expression patterns. Effective recombination is achieved only where both recombinases are expressed (Awatramani et al., 2003).

Temporal Control of Knockout

In many cases, the phenotype of interest is in the adult animal but, because the gene is necessary for development, no adult animals are obtained. Delaying the expression of Cre activity until the animal is an adult would allow normal development, and then the knockout could be created in the adult (Rajewsky et al., 1996). This can be accomplished by using a conditional expression system (e.g., the tet-on, tet-off, or ecdysone systems; see *UNIT 16.14* and St-Onge et al., 1996) or other inducible system (such as an interferon-inducible promoter; Kuhn et al., 1995) to express Cre at the proper time. This would, however, require the construction of animals containing three transgenes. Another approach that has been used is the creation of a fusion protein with either a modified estrogen receptor (Feil et al., 1996, 1997; Zhang et al., 1996; Brocard et al., 1997) or a modified glucocorticoid receptor (Brocard et al., 1998). The estrogen receptor (ER) has been mutated so that it no longer responds to endogenous estrogens, but is activated by the administration of the partial estrogen agonist tamoxifen (Fig. 23.1.9B). Temporal control using Flp/FRT has also been achieved using a fusion of Flp with a modified estrogen receptor (Hunter et al., 2005).

LITERATURE CITED

- Abuin, A. and Bradley, A. 1996. Recycling selectable markers in mouse embryonic stem cells. *Mol. Cell. Biol.* 16:1851-1856.
- Agah, R., Frenkel, P.A., French, B.A., Michael, L.H., Overbeek, P.A., and Schneider, M.D. 1997. Gene recombination in postmitotic cells. Targeted expression of Cre recombinase provokes cardiac-restricted, site-specific rearrangement in adult ventricular muscle in vivo. *J. Clin. Invest.* 100:169-179.
- Awatramani, R., Soriano, P., Rodriguez, C., Mai, J.J., and Dymecki, S.M. 2003. Cryptic boundaries in roof plate and choroid plexus identified by intersectional gene activation. *Nat. Genet.* 35:70-75.
- Branda, C.S. and Dymecki, S.M. 2004. Talking about a revolution: The impact of site-specific recombinases on genetic analysis in mice. *Dev. Cell* 6:7-28.
- Brocard, J., Warot, X., Wendling, O., Messaddeq, N., Vonesch, J.L., Chambon, P., and Metzger, D. 1997. Spatio-temporally controlled site-specific somatic mutagenesis in the mouse. *Proc. Natl. Acad. Sci. U.S.A.* 94:14559-14563.
- Brocard, J., Feil, R., Chambon, P., and Metzger, D. 1998. A chimeric Cre recombinase inducible by synthetic, but not by natural ligands of the glucocorticoid receptor. *Nucl. Acids Res.* 26:4086-4090.
- Copeland, N.G., Jenkins, N.A., and Cour, D.L. 2001. Recombineering: A powerful new tool for mouse functional genomics. *Nat. Rev. Genet.* 2:769-779.
- Cruz, A., Coburn, C.M., and Beverley, S.M. 1991. Double targeted gene replacement for creating null mutants. *Proc. Natl. Acad. Sci. U.S.A.* 88:7170-7174.
- Feil, R., Brocard, J., Mascrez, B., LeMeur, M., Metzger, D., and Chambon, P. 1996. Ligand-activated site-specific recombination in mice. *Proc. Natl. Acad. Sci. U.S.A.* 93:10887-10890.
- Feil, R., Wagner, J., Metzger, D., and Chambon, P. 1997. Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. *Biochem. Biophys. Res. Commun.* 237:752-757.
- Fiering, S., Kim, C.G., Epner, E.M., and Groudine, M. 1993. An "in-out" strategy using gene targeting and FLP recombinase for the functional dissection of complex DNA regulatory elements: Analysis of the β -globin locus control region. *Proc. Natl. Acad. Sci. U.S.A.* 90:8469-8473.
- Fiering, S., Epner, E., Robinson, K., Zhuang, Y., Telling, A., Hu, M., Martin, D.I., Enver, T., Ley, T.J., and Groudine, M. 1995. Targeted deletion of 5'HS2 of the murine β -globin LCR reveals that it is not essential for proper regulation of the β -globin locus. *Genes Dev.* 9:2203-2213.
- Fiering, S., Bender, M.A., and Groudine, M. 1999. Analysis of mammalian cis-regulatory DNA elements by homologous recombination. *Methods Enzymol.* 306:42-66.
- Gu, H., Marth, J.D., Orban, P.C., Mossmann, H., and Rajewsky, K. 1994. Deletion of a DNA polymerase β gene segment in T cells using cell type-specific gene targeting. *Science* 265:103-106.
- Hasty, P., Ramirez, S.R., Krumlauf, R., and Bradley, A. 1991. Introduction of a subtle mutation into the Hox-2.6 locus in embryonic stem cells [published erratum appears in *Nature*, 1991, 353:94]. *Nature* 350:243-246.
- Hunter, N.L., Awatramani, R.B., Farley, F.W., and Dymecki, S.M. 2005. Ligand-activated Flpe for temporally regulated gene modifications. *Genesis* 41:99-109.
- Kuhn, R., Schwenk, F., Aguet, M., and Rajewsky, K. 1995. Inducible gene targeting in mice. *Science* 269:1427-1429.
- Mansour, S.L. 1990. Gene targeting in murine embryonic stem cells: Introduction of specific alterations into the mammalian genome. *Genet. Anal. Tech. Appl.* 7:219-227.
- Mansour, S.L., Thomas, K.R., and Capecchi, M.R. 1988. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: A general strategy for targeting mutations to nonselectable genes. *Nature* 336:348-352.
- Mansour, S.L., Thomas, K.R., Deng, C.X., and Capecchi, M.R. 1990. Introduction of a lacZ reporter gene into the mouse int-2 locus by homologous recombination. *Proc. Natl. Acad. Sci. U.S.A.* 87:7688-7692.

- Mombaerts, P., Clarke, A.R., Hooper, M.L., and Tonegawa, S. 1991. Creation of a large genomic deletion at the T-cell antigen receptor beta-subunit locus in mouse embryonic stem cells by gene targeting. *Proc. Natl. Acad. Sci. U.S.A.* 88:3084-3087.
- Mortensen, R.M., Zubiaur, M., Neer, E.J., and Seidman, J.G. 1991. Embryonic stem cells lacking a functional inhibitory G-protein subunit (alpha i2) produced by gene targeting of both alleles. *Proc. Natl. Acad. Sci. U.S.A.* 88:7036-7040.
- Mortensen, R.M., Conner, D.A., Chao, S., Geisterfer, L.A., and Seidman, J.G. 1992. Production of homozygous mutant ES cells with a single targeting construct. *Mol. Cell. Biol.* 12:2391-2395.
- Olson, E.N., Arnold, H.H., Rigby, P.W., and Wold, B.J. 1996. Know your neighbors: Three phenotypes in null mutants of the myogenic bHLH gene MRF4. *Cell* 85:1-4.
- Pham, C.T., MacIvor, D.M., Hug, B.A., Heusel, J.W., and Ley, T.J. 1996. Long-range disruption of gene expression by a selectable marker cassette. *Proc. Natl. Acad. Sci. U.S.A.* 93:13090-13095.
- Rajewsky, K., Gu, H., Kuhn, R., Betz, U.A., Muller, W., Roes, J., and Schwenk, F. 1996. Conditional gene targeting. *J. Clin. Invest.* 98:600-603.
- Robertson, E.J. 1991. Using embryonic stem cells to introduce mutations into the mouse germ line. *Biol. Reprod.* 44:238-245.
- Rodriguez, C.I., Buchholz, F., Galloway, J., Sequerra, R., Kasper, J., Ayala, R., Stewart, A.F., and Dymecki, S.M. 2000. High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. *Nat. Genet.* 25:139-140.
- Rohmann, A., Gotthardt, M., Willnow, T.E., Hammer, R.E., and Herz, J. 1996. Sustained somatic gene inactivation by viral transfer of Cre recombinase. *Nat. Biotechnol.* 14:1562-1565.
- Sauer, B. 1993. Manipulation of transgenes by site-specific recombination: Use of Cre recombinase. *Methods Enzymol.* 225:890-900.
- Sauer, B. and Henderson, N. 1989. Cre-stimulated recombination at loxP-containing DNA sequences placed into the mammalian genome. *Nucl. Acids Res.* 17:147-161.
- Schwenk, F., Baron, U., and Rajewsky, K. 1995. A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. *Nucl. Acids Res.* 23:5080-5081.
- Shibata, H., Toyama, K., Shioya, H., Ito, M., Hirota, M., Hasegawa, S., Matsumoto, H., Takano, H., Akiyama, T., Toyoshima, K., Kanamaru, R., Kanegae, Y., Saito, I., Nakamura, Y., Shiba, K., and Noda, T. 1997. Rapid colorectal adenoma formation initiated by conditional targeting of the *Apc* gene. *Science* 278:120-123.
- St-Onge, L., Furth, P.A., and Gruss, P. 1996. Temporal control of the Cre recombinase in transgenic mice by a tetracycline responsive promoter. *Nucl. Acids Res.* 24:3875-3877.
- te Riele, H., Maandag, E.R., Clarke, A., Hooper, M., and Berns, A. 1990. Consecutive inactivation of both alleles of the *pim-1* proto-oncogene by homologous recombination in embryonic stem cells. *Nature* 348:649-651.
- van der Neut, R. 1997. Targeted gene disruption: Applications in neurobiology. *J. Neurosci. Methods* 71:19-27.
- Vooijs, M., van der Valk, M., te Riele, H., and Berns, A. 1998. Flp-mediated tissue-specific inactivation of the retinoblastoma tumor suppressor gene in the mouse. *Oncogene* 17:1-12.
- Wang, Y., Krushel, L.A., and Edelman, G.M. 1996. Targeted DNA recombination in vivo using an adenovirus carrying the cre recombinase gene. *Proc. Natl. Acad. Sci. U.S.A.* 93:3932-3936.
- Williams-Simons, L. and Westphal, H. 1999. EIIaCre: Utility of a general deleter strain. *Transgenic Res.* 8:53-54.
- Yagi, T., Ikawa, Y., Yoshida, K., Shigetani, Y., Takeda, N., Mabuchi, I., Yamamoto, T., and Aizawa, S. 1990. Homologous recombination at c-fyn locus of mouse embryonic stem cells with use of diphtheria toxin A-fragment gene in negative selection. *Proc. Natl. Acad. Sci. U.S.A.* 87:9918-9922.
- Zhang, Y., Riesterer, C., Ayrall, A.M., Sablitzky, F., Littlewood, T.D., and Reth, M. 1996. Inducible site-directed recombination in mouse embryonic stem cells. *Nucl. Acids Res.* 24:543-548.
- Zimmer, A. 1992. Manipulating the genome by homologous recombination in embryonic stem cells. *Annu. Rev. Neurosci.* 15:115-137.

Contributed by Richard Mortensen
University of Michigan Medical Center
Ann Arbor, Michigan

Mouse Embryo Fibroblast (MEF) Feeder Cell Preparation

UNIT 23.2

The production of mouse mutants using homologous recombination and blastocyst-mediated transgenesis requires the maintenance of mouse embryonic stem (ES) cells in an undifferentiated state. Many investigators rely on feeder layers to prevent ES cell differentiation; feeder layers prepared from mitotically inactivated primary mouse embryo fibroblasts (MEFs) are used most commonly. This unit describes a simple method to isolate and store MEFs (Basic Protocol 1 and Support Protocol) and two common techniques for mitotic inactivation: γ -irradiation (Basic Protocol 2) and mitomycin C treatment (Alternate Protocol).

NOTE: All cell culture incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

ISOLATION OF PRIMARY MOUSE EMBRYO FIBROBLASTS

**BASIC
PROTOCOL 1**

MEFs are isolated from 12.5 to 13.5 postcoitum (p.c.) mouse embryos. The embryos are dissociated and then trypsinized to produce single-cell suspensions. After expansion, aliquots can be frozen and stored in liquid nitrogen indefinitely. Alternatively, MEFs suitable for ES cell culture may be obtained from commercial sources (see <http://www.biosupplynet.com> for a current list of suppliers). Commercial MEFs may be useful to researchers new to ES cell culture or to those who will grow ES cells on a limited scale.

Materials

Mouse embryos, 12.5 to 13.5 days postcoitum (Hogan et al., 1994)

DPBS (see recipe), sterile

Trypsin/EDTA solution (see recipe)

MEF medium (see recipe) with penicillin/streptomycin

Laminar flow hood

Inverted microscope

100-mm tissue culture dish

Dissecting forceps and fine scissors, sterilized by autoclaving or ethanol flaming

10-ml syringe and 16-G needle

100-mm tissue culture plates or 75-cm² flasks

Additional reagents and equipment for passaging and freezing MEFs (see Support Protocol)

1. Dissect mouse embryos (12.5 to 13.5 days postcoitum) into 10 to 20 ml sterile DPBS in a 100-mm tissue culture dish. Process embryos from one mouse together. Remove embryonic internal organs from the abdominal cavity using dissecting forceps. Transfer the carcass to a clean dish with fresh DPBS.

Organ removal can be done crudely. Initial dissection can be performed at the bench. Subsequent procedures should be performed in a laminar flow hood.

Any mouse strain can be used as an embryo source. However, outbred mice or F1 hybrids will usually produce more embryos per mating than inbred mice. If possible, use mice maintained in a viral-antibody-free (VAF) facility to reduce the chance of contamination. If the feeder cells will be used during the selection of antibiotic-resistant ES cells, use embryos from a transgenic mouse expressing the appropriate selectable marker. Appropriate transgenic mice may be obtained from the ES cell community or from standard mouse vendors such as The Jackson Laboratory and Taconic.

**Manipulating the
Mouse Genome**

23.2.1

Contributed by David A. Conner

Current Protocols in Molecular Biology (2000) 23.2.1-23.2.7

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 51

2. Rinse embryos again in 10 to 20 ml DPBS. Transfer the embryos to a clean 100-mm tissue culture dish containing 3 to 5 ml trypsin/EDTA solution.
3. Dissociate embryos by aspirating into a 10-ml syringe through a 16-G needle and expelling the contents. Repeat two to four times.

Too many repetitions will reduce the cell yield. The embryos should be dissociated to the extent that they can be easily aspirated into a 5- or 10-ml pipette.

4. Add trypsin/EDTA solution to 10 ml. Mix contents by trituration, return to dish, and incubate 5 to 10 min in a 37°C incubator.
5. Mix again by trituration and incubate for an additional 5 to 10 min at 37°C.
6. Transfer contents to a 50-ml conical tube and add an equal volume of MEF medium with penicillin/streptomycin. Let stand 3 to 5 min at room temperature to allow large tissue pieces to settle to the bottom.
7. Remove solution, avoiding large tissue pieces, and place in a fresh 50-ml tube. Centrifuge 5 min at 1000 × g, room temperature.
8. Remove supernatant and resuspend pellet in 10 to 50 ml fresh MEF medium with penicillin/streptomycin. Plate cells on 100-mm tissue culture plates or 75-cm² flasks, using approximately one embryo per plate or flask. Add medium to a final volume of 10 to 15 ml/plate.
9. Grow cells until confluent (2 to 5 days). Monitor cell density using an inverted microscope. Change medium after the first day and every other day thereafter.
10. Passage cells by trypsinizing (see Support Protocol, steps 1 to 4), resuspending the cell pellet in 10 to 50 ml MEF medium with penicillin/streptomycin, and plating at a dilution of 1:5 to 1:10. Add medium to a final volume of 10 to 15 ml per 75-cm² flask or 10-mm plate.

Using this dilution, fibroblasts from each original embryo are now plated on five to ten 75-cm² flasks or 100-mm plates.

11. Grow again until confluent (3 to 5 days) and freeze (see Support Protocol, steps 1 to 7) at $\sim 5 \times 10^6$ cells/ml.

A representative vial should be thawed to check for viability. It is also a good habit to check each new batch for mycoplasma contamination using a commercial service or a PCR-based screening kit (e.g., Pan Vera, Stratagene; also see Coté, 2000). In addition, MEFs made from previously untested transgenic mice should be grown in the appropriate antibiotic to ensure that they are resistant to the concentration of antibiotic used for selection.

BASIC PROTOCOL 2

MITOTIC INACTIVATION OF MEFS WITH γ -IRRADIATION

MEFs must be inactivated prior to use as a feeder layer for mouse ES cells. Mitotic inactivation prevents the dilution of ES cell lines with dividing fibroblasts. MEFs can be inactivated using γ -irradiation, as described here, or mitomycin C treatment (see Alternate Protocol). This procedure is faster and less labor intensive, but requires a convenient radiation source. Both methods produce feeder layers suitable for the maintenance of undifferentiated ES cells.

Materials

- Frozen MEF culture (see Basic Protocol 1)
- MEF medium (see recipe) without penicillin/streptomycin
- Ca²⁺- and Mg²⁺-free HBSS (see recipe)

100-mm tissue culture plates or 75-cm² flasks

150-cm² tissue culture flasks

100-mm Petri dishes

γ-Radiation source

Additional reagents and equipment for passaging, freezing, and thawing MEFs
(see Support Protocol)

1. Thaw a vial of frozen MEFs as described below (see Support Protocol, steps 8 to 10) and plate in a 75-cm² tissue culture flask or 100-mm plate. Grow until confluent (3 to 5 days). Change medium after the first day and every other day thereafter.
2. Passage cells by trypsinizing (see Support Protocol, steps 1 to 4), resuspending the cell pellet in 10 to 50 ml MEF medium without penicillin/streptomycin, and plating at a 1:10 dilution. Add medium to a final volume of 10 to 15 ml per 75-cm² flask or 100-mm plate.
3. Grow until confluent (3 to 5 days), and passage at a 1:5 to 1:10 dilution, using twenty-five to fifty 150-cm² flasks.

Primary fibroblasts undergo a limited and variable number of cell divisions. Further passaging may be possible, but the rate of cell division slows quickly. Thawing a fresh vial is usually the fastest way to generate more feeders.

4. Remove medium from confluent flasks, rinse with 15 ml Ca²⁺- and Mg²⁺-free HBSS, and trypsinize again (see Support Protocol, steps 1 to 4).

MEFs from ten to fifteen 150-cm² flasks can be processed together.

5. Resuspend pellet in 10 ml MEF medium without penicillin/streptomycin and transfer suspension in a 100-mm Petri dish.

Use a constant volume of medium regardless of the number of cells to ensure consistent results. Use a Petri dish rather than a cell culture plate to prevent adherence of the MEFs to the surface during irradiation.

6. Expose cells to 4000 rads from a γ-radiation source.
7. Dilute suspension to 50 ml with MEF medium without penicillin/streptomycin. Count the number of cells and freeze as described below (see Support Protocol, steps 4 to 7).

MITOTIC INACTIVATION OF MEFS WITH MITOMYCIN C

MEFs may be inactivated by mitomycin C treatment if a γ-radiation source is not available. Although this method is more time and labor intensive, the inactivated feeders are equally suitable for ES cell culture.

Additional Materials (also see Basic Protocol 2)

1 mg/ml mitomycin C (Sigma) stock solution, filter sterilized (store at 4°C protected from light)

1. Expand a vial of MEFs as described above (see Basic Protocol 2, steps 1 to 3).
2. Add 1 mg/ml mitomycin C stock solution to the medium to a final concentration of 10 μg/ml. Return plates to the incubator for 2 to 3 hr.
3. Rinse plates twice with 10 to 15 ml Ca²⁺- and Mg²⁺-free HBSS.
4. Trypsinize as if passaging (see Support Protocol, steps 1 to 4).

**ALTERNATE
PROTOCOL**

**Manipulating the
Mouse Genome**

23.2.3

5. Add an equal volume of MEF medium without penicillin/streptomycin.
6. Count and freeze cells (see Support Protocol, steps 4 to 7).

FREEZING AND THAWING MEFs

Freezing MEFs, particularly inactivated fibroblasts, is a great convenience. Large stocks of active or inactivated fibroblasts can be prepared at any time, obviating the need to coordinate MEF preparations with ES cell manipulations. Cells are frozen slowly in medium containing 10% (v/v) dimethyl sulfoxide and thawed rapidly.

Materials

Plates containing MEFs (see Basic Protocols 1 and 2; see Alternate Protocol)
Ca²⁺- and Mg²⁺-free HBSS (see recipe)
Trypsin/EDTA solution (see recipe)
Freezing medium (see recipe)
MEF medium (see recipe) with or without penicillin/streptomycin
Cryovials
Additional reagents and equipment for counting cells with a hemacytometer
(APPENDIX 3F)

Freeze cells

1. Remove MEF medium and rinse plates with 10 to 15 ml Ca²⁺- and Mg²⁺-free HBSS.

Washing the plates removes residual serum that will inhibit trypsin.

2. Remove HBSS and add trypsin/EDTA solution to cover the surface of the cells (e.g., 3 to 5 ml in a 75-cm² flask). Incubate 3 to 5 min at 37°C and tap the flask to release the cells.

When trypsinization is complete, tapping and rocking the plate should release the cells. Loose sheets of cells are visible to the naked eye. Incubate at 37°C for an additional 1 to 2 min if the cells do not slough off.

3. When the cell layer has loosened, add an equal volume of MEF medium and mix by trituration to produce a single-cell suspension.

MEF medium should contain penicillin/streptomycin when preparing initial MEFs (i.e., Basic Protocol 1, step 10), but not for other protocols. Addition of medium containing serum will inhibit further trypsinization.

4. Count cells using a hemacytometer (APPENDIX 3F) and then pellet by centrifuging 3 to 5 min at 1000 × g, room temperature.
5. Resuspend pellet at 3 × 10⁶ cells/ml in freezing medium and mix by trituration.
6. Dispense into cryovials in 1-ml aliquots, place cryovials in an insulated container at -80°C, and leave overnight.

Under these conditions, the cells will freeze slowly enough to maintain viability.

7. Transfer cryovials to liquid nitrogen.

Cells should not be quick-frozen in liquid nitrogen. Vials can be stored at -80°C for several months and in liquid nitrogen for years.

Thaw cells

8. Thaw cells rapidly by placing vials in a 37°C water bath.

9. Add cells to a tube with 10 ml MEF medium without penicillin/streptomycin. To remove DMSO, pellet by centrifuging 3 to 5 min at $1000 \times g$, room temperature.
10. Resuspend in a volume appropriate for the surface area of the plate (e.g., 10 to 15 ml for a 75-cm^2 flask or 100-mm plate).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

NOTE: Prepare all cell culture solutions from tissue culture–grade reagents. Use tissue culture–grade water (high resistance and endotoxin free). Sterilize all final solutions by filtration or prepare from sterile stocks. Use disposable sterile plasticware to prevent microbial and detergent contamination.

Dulbecco's phosphate-buffered saline (DPBS)

0.1 g/liter anhydrous CaCl_2
0.1 g/liter $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$
0.2 g/liter KCl
0.2 g/liter KH_2PO_4
8.0 g/liter NaCl
2.16 g/liter $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$
Adjust pH, if necessary, to 7.0 to 7.2 with 1 N HCl or 1 N NaOH
Sterilize by filtration or by autoclaving
Store at 4°C (stable indefinitely)

Freezing medium

Dulbecco's modified Eagle medium (DMEM) with high glucose (4500 mg/liter), L-glutamine, sodium pyruvate, and pyridoxine hydrochloride or pyridoxal hydrochloride (e.g., Life Technologies, JRH Biosciences, or Sigma) containing:
20 mM HEPES, pH 7.3
20% (v/v) FBS, heat inactivated for 30 min at 56°C
10% (v/v) dimethyl sulfoxide (DMSO)
Store up to 1 year at -20°C

Freezing and thawing should be avoided. Commercial freezing media based on 10% DMSO can be used. Consult UNIT 23.3 regarding the selection of FBS.

Hank's balanced salt solution (HBSS), Ca^{2+} and Mg^{2+} free

0.4 g/liter KCl
0.06 g/liter KH_2PO_4
8.0 g/liter NaCl
0.35 g/liter NaHCO_3
0.048 g/liter Na_2HPO_4
1 g/liter D-glucose
0.01 g/liter phenol red
Adjust pH, if necessary, to 7.0 to 7.4 with 1 N HCl or 1 N NaOH
Store up to 6 months at 4°C

Mouse embryo fibroblast (MEF) medium

Dulbecco's modified Eagle medium (DMEM) with high glucose (4500 mg/liter), L-glutamine, sodium pyruvate, and pyridoxine hydrochloride or pyridoxal hydrochloride (e.g., Life Technologies, JRH Biosciences, or Sigma) containing:
10% (v/v) FBS, heat inactivated for 30 min at 56°C
1× MEM nonessential amino acids (from 100× stock; Life Technologies)
2 mM L-glutamine (from 100× stock; Life Technologies)
0.1 mM 2-mercaptoethanol
20 mM HEPES, pH 7.3
1× penicillin/streptomycin when indicated (from 100× stock; Life Technologies)
Store up to 2 weeks at 4°C

Consult UNIT 23.3 regarding the selection of FBS. Penicillin/streptomycin should be included during the initial isolation of MEFs. It is not needed after Basic Protocol 1.

Trypsin/EDTA solution

Ca²⁺- and Mg²⁺-free HBSS (see recipe) containing:
2.5 g/liter porcine trypsin (0.25%)
0.38 g/liter EDTA·4H₂O
Store aliquots up to 1 year at -20°C

Aliquots can be thawed and stored at 4°C for up to 1 week. Repeated freezing and thawing should be avoided.

COMMENTARY

Background Information

Historically, ES cells have been cultured under a variety of conditions to prevent differentiation (Wurst and Joyner, 1993; *UNIT 23.3*). The most common method utilizes feeder layers prepared from mouse embryo fibroblasts or from SIM mouse embryo fibroblasts resistant to thioguanine and ouabain (STO cells; Hogan et al., 1994; Martin and Evans, 1975). Whereas MEFs are primary cultures with a limited mitotic potential, STO fibroblasts will divide indefinitely. STO cells resistant to G418 are readily available from the ES cell community. STO cells can be grown and inactivated using the protocols for MEFs. For more details see Robertson (1987), Wurst and Joyner (1993), or Hogan et al. (1994).

The advantage of MEFs is that they provide a more consistent source of feeder cells. Early passage cells with reproducible characteristics must be used because they rapidly lose their ability to divide. Long-term propagation of STO cells can lead to changes that result in characteristics that are less favorable for ES cell growth. Because of these potential problems, researchers who are new to ES cell culture and who wish to generate ES cell clones capable of germ-line transmission will probably have more success with primary MEFs.

Critical Parameters and Troubleshooting

The major problems encountered during the isolation and inactivation of MEFs are typically a result of poor aseptic technique, low cell yields during isolation, or inefficient mitotic inactivation. Special attention should be paid to the use of aseptic technique. After the initial primary cultures have been frozen, antibiotics should not be used; antibiotics can mask bad aseptic technique. Each new batch of MEFs should be checked for mycoplasma after growth without antibiotics (Coté, 2000). Inadvertent contamination will spread to ES cells and reduce their germ-line potential. Contaminated batches should be disposed of; it is best not to try to treat them. Contamination should be a rare occurrence.

Aseptic technique is also important in the preparation of media and solutions. All solutions should be prepared using tissue culture-grade reagents and water (high resistance and endotoxin free). Although they are more expensive, solutions that have been screened for toxicity and mycoplasma are available from many companies. Laboratories with little cell culture experience will probably have the greatest success with commercial, prescreened solutions.

Occasionally low cell yields are observed after the initial embryo dissociation. Plates from the initial dissociation should reach con-

fluence in 2 to 5 days. If this does not occur, it is likely due to low trypsin activity or cell lysis during dissociation with the needle and syringe. Some cells may be recovered by passaging without dilution or by concentrating the cells from several plates onto a single plate. Under these circumstances, the cells may still be used as long as they resume dividing and reach confluence. During future preparations, fresh trypsin solutions should be used and the number of times each embryo is passed through the needle should be reduced.

Occasionally evidence of cell division is observed in cultures of mitotically inactivated MEFs. When preparing inactivated feeders for the first time, check to make sure that the treatment was effective. This can be accomplished by plating out an aliquot of the inactivated feeders at a density of $\sim 6 \times 10^4$ cells/cm². The cells are then cultured for 10 to 14 days, with periodic changes of medium, and cultures are assessed for increases in cell density and foci of mitotic activity. Foci can be observed by eye as opaque splotches on the plate, but should be confirmed with a microscope. No growth should be seen on a 100-mm plate; however, feeders may still be used if there are only a few foci. If there is a dramatic change in cell density or there are many colonies of dividing cells, the preparation should be discarded and the procedure repeated with newly expanded cells. If mitomycin C was used, a fresh solution should be prepared. If γ -irradiation was used, the radiation source should be properly calibrated. Distance and shielding can affect the dose dramatically.

In addition to these concerns, it is important to remember that MEFs are primary cells with limited mitotic potential. Expanding the cells more than suggested may work, but the rate of growth will decrease. Plates can still become confluent because the cell size will increase; however, the number of cells per plate will decrease and the time to reach confluence will increase. Feeder layers made under these conditions may not be ideal for ES cell culture.

Anticipated Results

Expect 10 to 30 vials of frozen cells from each embryo. Each vial can be expanded to produce 50 to 100 vials containing 1 ml of mitotically inactive feeders. Each vial should be sufficient for one targeting experiment.

Time Considerations

Processing a single litter of embryos (six to ten) should take 1 to 2 hr. After plating the dissociated embryos, cultures should reach confluence within 2 to 5 days. Cells can be frozen at this point, but the total yield will be dramatically reduced. After passaging, the cells should be ready to freeze in 3 to 5 days. Thawed vials should reach confluence in 3 to 5 days. After each passage the cells should take 3 to 5 days to reach confluence. The cells can be inactivated at any stage, but the yield will be lower with less expansion.

Literature Cited

- Coté, R. 2000. Assessing and controlling microbial contamination in cell cultures. *In* Current Protocols in Cell Biology (J.S. Bonifacino, M. Dasso, J.B. Harford, J. Lippincott-Schwartz, and K.M. Yamada, eds.) pp. 1.5.1-1.5.18. John Wiley & Sons, New York.
- Hogan, B., Beddington, R., Constantini, F., and Lacy, E. 1994. *Manipulating the Mouse Embryo: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Martin, G.R. and Evans, M.J. 1975. Differentiation of clonal lines of teratocarcinoma cells: Formation of embryoid bodies in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 72:1441-1445.
- Robertson, E.J. 1987. Embryo-derived stem cell lines. *In* Teratocarcinomas and Embryonic Stem Cells: A Practical Approach (E.J. Robertson, ed.) pp. 71-112. IRL Press, Oxford.
- Wurst, W. and Joyner, A.L. 1993. Production of targeted embryonic stem cell clones. *In* Gene Targeting: A Practical Approach (A.L. Joyner, ed.) pp. 33-61. IRL Press, Oxford.

Key References

Hogan et al., 1994. See above.

Provides additional or alternative protocols and defines the context for use of MEFs during gene targeting in ES cells.

Internet Resources

www.biosupplynet.com

Search this Web site for "embryonic stem cell reagents" to obtain a current list of suppliers that provide medium and MEFs suitable for ES cell culture.

Contributed by David A. Conner
Harvard Medical School
Boston, Massachusetts

Mouse Embryonic Stem (ES) Cell Culture

UNIT 23.3

**BASIC
PROTOCOL**

Culturing mouse embryonic stem (ES) cells has many similarities to the culture of any adherent cell line. However, there are some special concerns that warrant elaboration because ES cells must not be allowed to differentiate if they are to be used to generate germ-line chimeras. This unit describes a common method for ES cell culture utilizing gelatinized plates, feeder layers of mitotically inactive mouse embryo fibroblasts (MEFs; UNIT 23.2), and recombinant leukemia inhibitory factor (LIF). Special attention is paid to the timing of passaging to prevent differentiation during routine culture.

NOTE: All cell culture incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

Materials

- 0.1% (w/v) gelatin solution (see recipe)
- Mitotically inactive mouse embryo fibroblasts (MEFs; UNIT 23.2)
- ES cell medium (see recipe)
- Embryonic stem (ES) cells (<http://www.biosupplynet.com>)
- Ca²⁺- and Mg²⁺-free HBSS (UNIT 23.2)
- DPBS/EDTA solution (see recipe)
- Trypsin/EDTA solution (see recipe)
- Freezing medium (UNIT 23.2)
- Cryovials

- Additional reagents and equipment for counting cells with a hemacytometer (APPENDIX 3F)

Prepare gelatin-coated plates

1. Pipet a sufficient volume of 0.1% gelatin solution onto plates to cover the bottom of the plate.
2. Leave the solution on the plate for ≥1 hr at room temperature or in a 37°C incubator.
The temperature is not critical. Plates can be prepared in advance by adding the gelatin solution and storing the plate in the incubator (to prevent drying) for up to a week.

Prepare feeder plates

3. Aspirate gelatin solution just before adding feeder cells.
4. Plate mitotically inactive MEFs (feeders) in ES cell medium on gelatinized plates at ~6 × 10⁴ viable cells/cm². Place cells in the incubator.

Either freshly inactivated MEFs or frozen inactivated MEFs can be used. Frozen cells should be thawed rapidly at 37°C and DMSO should be removed (see UNIT 23.2). Note that some cell death occurs during the freezing and thawing procedure. Frozen MEFs should thus be plated at approximately twice the density (1.2 × 10⁵ cells/cm²) of fresh MEFs (6 × 10⁴ cells/cm²) to achieve a similar density of viable feeders. If the density is too high, the cells may lift off the plate.

Feeders can be plated in either ES medium or MEF medium (see UNIT 23.2). Both have similar components, and it is often more convenient to prepare only one. The feeders attach to the plate within 30 min and begin to spread out over the next few hours.

5. Change medium to remove dead feeders the next day.

The fibroblasts should spread to cover the entire surface of the plate by the day after plating.

Feeder plates can be prepared in advance and used for routine ES cell culture within 7 to 10 days after plating. The medium should be changed every other day if they are not used immediately.

**Manipulating the
Mouse Genome**

23.3.1

Contributed by David A. Conner

Current Protocols in Molecular Biology (2000) 23.3.1-23.3.6

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 51

Culture ES cells

6. Thaw a vial of ES cells rapidly at 37°C. Remove DMSO by diluting in 10 ml ES cell medium and centrifuging 3 to 5 min at 1000 × g, room temperature.
7. Resuspend pellet in ES cell medium and plate on feeder layers prepared at least one day earlier.

The number of ES cells depends on the surface area of the feeder plate. Plate 1–3 × 10⁶ cells on a 25-cm² flask or a 60-mm plate. Plate cells frozen at a higher density on 75-cm² flasks or 100-mm plates, and fewer cells in 6- or 24-well feeder plates.

8. Change medium daily and passage cells every 2 to 3 days depending on total cell density and colony size (see Critical Parameters and Troubleshooting for more details).

Cells may require passaging the day after plating if the cell density is very high.

Passage ES cells

9. Change medium several hours before passaging to maximize plating efficiency.
10. Rinse plates with Ca²⁺- and Mg²⁺-free HBSS (e.g., 10 ml per 100-mm plate).

Washing the plates removes residual serum that will inhibit trypsin.

11. Add sufficient DPBS/EDTA solution to cover the cells and incubate 3 min at 37°C.
12. Add an equal volume of Trypsin/EDTA solution. Rock the plate to mix and return to the incubator for 1 to 2 min.
13. Tap the plate to release the cells. If the cells do not slough off, return the plate to the incubator for another minute and tap again.

Alternatively, the DPBS/EDTA step can be skipped, and the plate can be incubated for 3 to 7 min after addition of trypsin/EDTA, with periodic tapping to see if the cell layer has loosened. With either approach, the goal is to achieve a single-cell suspension with minimal trypsinization. Preincubation with DPBS/EDTA reduces the time necessary for trypsin treatment and yields the same degree of dispersion.

14. When the cell layer has loosened, add 2 to 3 vol ES cell medium and disperse the cells thoroughly by trituration.
15. Centrifuge 3 to 5 min at 1000 × g, room temperature. Resuspend the pellet in ES cell medium and plate on feeders at a 1:10 dilution.

Freeze cells

16. Trypsinize cells (steps 9 to 13).
17. Add an equal volume of ES cell medium and mix by trituration to produce a single-cell suspension.

Addition of medium containing serum will inhibit further trypsinization.

18. Count cells using a hemacytometer (*APPENDIX 3F*) and then pellet by centrifuging 3 to 5 min at 1000 × g, room temperature.
19. Resuspend pellet at 3 × 10⁶ cells/ml in freezing medium and mix by trituration.
20. Dispense into cryovials in 1-ml aliquots, place the cryovials in an insulated container at –80°C, and leave overnight.

Under these conditions, the cells will freeze slowly enough to maintain viability.

21. Transfer cryovials to liquid nitrogen.

Cells should not be quick-frozen in liquid nitrogen. Vials can be stored at –80°C for several months and in liquid nitrogen for years.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

NOTE: Prepare all cell culture solutions from tissue culture–grade reagents. Use tissue culture–grade water (high resistance and endotoxin free). Sterilize all final solutions by filtration or prepare from sterile stocks. Use disposable sterile plasticware to prevent microbial and detergent contamination.

DPBS/EDTA solution

Prepare DPBS (*UNIT 23.2*) but without calcium and magnesium
Add 0.2 g/liter EDTA·4H₂O
Sterilize by filtration or autoclaving
Store at 4°C (stable indefinitely)

ES cell medium

Dulbecco's modified Eagle medium (DMEM) with high glucose (4500 mg/liter), L-glutamine, sodium pyruvate, and pyridoxine hydrochloride or pyridoxal hydrochloride (e.g., Life Technologies, JRH Biosciences, or Sigma) containing:
15% (v/v) FBS, heat inactivated for 30 min at 56°C
1× MEM nonessential amino acids (from 100× stock; Life Technologies)
2 mM L-glutamine (from 100× stock; Life Technologies)
0.1 mM 2-mercaptoethanol
20 mM HEPES, pH 7.3
500 to 1000 U/ml murine leukemia inhibitory factor (LIF; e.g., ESGRO, Life Technologies; Chemicon)
Store up to 1 week at 4°C

The quality of FBS should be screened by assessing plating efficiency, morphology, and toxicity on ES cells plated at low density in 10%, 15%, and 30% heat-inactivated FBS. Plating efficiency should be ~10% and should be similar for all FBS concentrations. The author has had good success with FBS from Hyclone. Several companies sell FBS that has been screened for ES cell growth using the same techniques. Although it is expensive, laboratories that are new to ES cell culture may benefit from prescreened FBS. In addition, prescreened FBS provides a standard of comparison for other FBS lots.

Gelatin solution, 0.1% (w/v)

Prepare 0.1% (w/v) porcine skin gelatin (Sigma) in water and sterilize by autoclaving. Filter after cooling. Store up to 1 year at 4°C.

Trypsin/EDTA solution

Ca²⁺- and Mg²⁺-free HBSS (*UNIT 23.2*) containing:
0.5 g/liter porcine trypsin (0.05%)
0.2 g/liter EDTA·4H₂O
Store aliquots up to 1 year at –20°C

*Note that this recipe contains one-fifth the trypsin that is used for preparing MEFs (*UNIT 23.2*).*

Aliquots can be thawed and stored at 4°C for up to 1 week. Repeated freezing and thawing should be avoided.

COMMENTARY

Background Information

ES cells can be obtained from a number of investigators, from commercial sources, or by establishing new lines directly from mouse embryos (see <http://www.biosupplynet.com> for a current list of commercial suppliers). Historically, ES cells have been cultured under a variety of conditions to prevent differentiation. Lines have been successfully maintained in Buffalo rat liver (BRL) cell-conditioned medium (Smith and Hooper, 1987), in LIF-containing medium (Smith et al., 1988), and on STO or MEF feeder layers with or without LIF (Evans and Kaufman, 1981; Martin, 1981; Wurst and Joyner, 1993). ES cells may become accommodated to the culture conditions in which they were isolated; therefore, it may be advisable to modify the protocol outlined in this unit to match the growth conditions recommended for specific ES cell lines. However, most ES cell lines should do well on MEF feeder layers supplemented with LIF.

Critical Parameters and Troubleshooting

The major problems encountered during ES cell culture are contamination and differentiation. Special attention should be paid to aseptic technique. ES cells should not be cultured routinely in antibiotics. Although penicillin and streptomycin do not reduce the germ-line potential of ES cells, the antibiotics can mask bad aseptic technique. Mycoplasma contamination can reduce germ-line potential and is not eliminated by penicillin and streptomycin. Cultures should be checked periodically for mycoplasma (Coté, 2000), and contaminated cultures should be discarded. It is best not to try to treat them. Contamination should be a rare occurrence.

Aseptic technique is also important in the preparation of media and solutions. All solutions should be prepared using tissue culture-grade reagents and water (high resistance and endotoxin free). Although they are more expensive, solutions that have been screened for toxicity and mycoplasma are available from many companies. Laboratories with little cell culture experience will probably have the greatest success with commercial, prescreened solutions (<http://www.biosupplynet.com>).

ES cells grow as three-dimensional colonies, not monolayers. Undifferentiated colonies have a rounded appearance with sharp refractal edges; individual cell borders within

a colony are difficult to discern (Fig. 23.3.1A). All cultures will contain some differentiated colonies (Fig. 23.3.1B). Typically these colonies grow as monolayers with ragged edges and easily distinguishable cell borders. The appearance of ES cell colonies can change depending on total cell density and the time since passaging. To assess the quality of the culture, the cells should be observed over several days. During this period most of the colonies should be undifferentiated.

Assuming that quality reagents are used, differentiation of ES cells can be minimized by limiting the number of cell divisions and by passaging cultures at the appropriate density. It is advisable to keep track of the number of passages. New ES cell lines should be expanded with a minimum number of passages and frozen in many aliquots. A new low-passage vial should be thawed for each experiment. The passage number is relevant only because it is a rough estimate of the number of cell divisions the culture has undergone. Thus, one should not skip passaging at the appropriate density to avoid increasing the passage number. Cells should be viewed every day until their growth patterns become familiar. One day after passaging, ES cell colonies can be difficult to distinguish from cell debris on the surface of the feeder layer. By the second day, distinct colonies are apparent (Fig. 23.3.1C).

Determining the appropriate density for passaging is the most difficult aspect of ES cell culture. Both the total plate density and the individual colony size should be considered. Cells will tend to differentiate within large colonies or at very high total plate densities. For example, the cells in Figure 23.3.1C can be cultured for another day; the plate density is low and the colonies are small. The cells in Figure 23.3.1D are ready to passage; the plate density is moderate but the individual colonies are large. Although the cells seem to grow best at relatively high densities, the cells should not be allowed to become confluent; colonies should remain separated. Passaging every 2 days at a dilution of 1:8 or 1:10, when the cells are growing well, will maintain an ideal plate density and colony size. Passaging at a lower frequency is necessary when the cells are seeded at a low density. Under these conditions, it is best to pay attention to individual colony size and to passage with little dilution to disperse colonies and increase plate density.

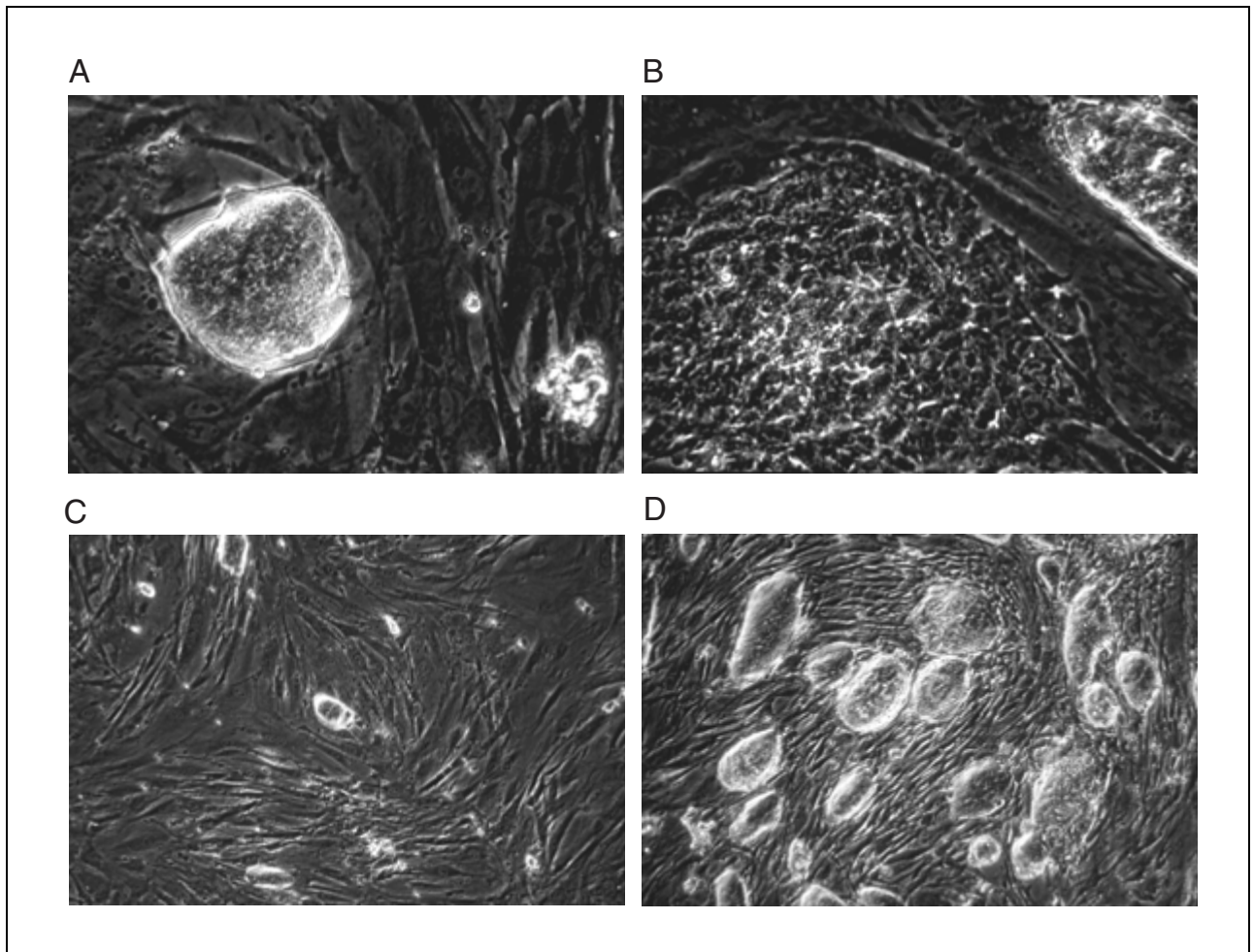


Figure 23.3.1 Morphology of ES cell colonies. Phase-contrast images of ES cell colonies on MEF feeders. **(A)** Appearance of an ideal ES cell colony that is ready to passage. **(B)** Appearance of a differentiating ES cell colony. Note that the colony has grown as a monolayer with distinct borders between cells. The edge of the colony is not yet ragged. **(C)** Appearance of ES cell colonies, at lower magnification, 2 days after plating at low density. Colonies can be distinguished from the feeder layer by their sharp refractal borders and their three-dimensional appearance. **(D)** Appearance of ES cell colonies 3 days after plating at an intermediate density. These cells are ready to passage.

Anticipated Results

When plates are grown at maximum density, there should be $2.5\text{--}4.0 \times 10^5$ cells/cm². For routine growth, the total density should be somewhat lower (e.g., $1.5\text{--}2 \times 10^5$ cells/cm²) and the cells should be passaged every 2 days.

Time Considerations

A single vial (frozen at 3.0×10^6 cells/ml) can be expanded in ~5 days to produce enough cells for a standard electroporation ($2\text{--}4 \times 10^7$ cells). A single vial can be expanded to produce 50 to 100 vials of a low-passage stock in 7 to 8 days.

Literature Cited

- Coté, R. 2000. Assessing and controlling microbial contamination in cell cultures. *In* *Current Protocols in Cell Biology* (J.S. Bonifacino, M. Dasso, J.B. Harford, J. Lippincott-Schwartz, and K.M. Yamada, eds.) pp. 1.5.1-1.5.18. John Wiley & Sons, New York.
- Evans, M.J. and Kaufman, M.H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292:154-156.
- Martin, G.R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 78:7634-7636.
- Smith, A.G. and Hooper, M.L. 1987. Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. *Dev. Biol.* 121:1-9.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* 336:688-690.

Wurst, W. and Joyner, A.L. 1993. Production of targeted embryonic stem cell clones. *In* Gene Targeting: A Practical Approach (A.L. Joyner, ed.) pp. 33-61. IRL Press, Oxford.

Key References

Hogan, B., Beddington, R., Constantini, F., and Lacy, E. 1994. Manipulating the Mouse Embryo: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Robertson, E.J. 1987. Embryo-derived stem cell lines. *In* Teratocarcinomas and Embryonic Stem Cells: A Practical Approach (E.J. Robertson, ed.) pp. 71-112. IRL Press, Oxford.

Wurst and Joyner, 1993. See above.

These three references, written by experts in the field, represent the best compilations of ES cell culture methods to date.

Internet Resources

www.biosupplynet.com

Search this Web site for "embryonic stem cell reagents" to obtain a current list of suppliers that provide medium and ES cells.

Contributed by David A. Conner
Harvard Medical School
Boston, Massachusetts

Mouse Embryonic Stem (ES) Cell Isolation

UNIT 23.4

Undifferentiated mouse embryonic stem (ES) cells are necessary for the production of mouse mutants using homologous recombination and blastocyst-mediated transgenesis. Suitable ES cell lines are available from commercial sources (UNIT 23.3) and from many investigators. However, investigators planning to use ES cells extensively may find that the isolation of new lines provides a simple and economical method for maintaining stocks of early passage ES cells. In addition, ES cells obtained from mutant mouse lines may facilitate the analysis of mutant phenotypes, particularly when the mutation causes early embryonic lethality.

ISOLATION OF MOUSE ES CELLS

ES cell isolation is straightforward, although success rates can be quite variable. ES cells are derived from the inner cell mass of blastocysts (i.e., 3.5-day-old embryos). Blastocysts are simply cultured for several days, during which time they attach to the surface of the tissue culture plate. Both trophoblast and inner cell mass cells divide after attachment. Inner cell mass outgrowths are picked, dispersed by trypsinization, and replated. Under appropriate conditions a percentage of the isolated outgrowths will continue to divide and maintain an undifferentiated ES cell morphology.

**BASIC
PROTOCOL**

Materials

- Modified ES medium (see recipe)
- Blastocysts, 3.5-day-old post-coitum embryos (Hogan et al., 1994)
- Hanks' balanced salt solution (HBSS), calcium- and magnesium-free (UNIT 23.2)
- DPBS-EDTA (UNIT 23.3)
- 0.25% (w/v) trypsin-EDTA (UNIT 23.2)
- Inverted microscope
- Gilson-style automatic pipettor with 20- μ l pipet tips
- 96-well U-bottom plate
- Additional reagents and equipment for preparing gelatin-coated plates with MEF feeder layers for embryonic stem cell culture (UNIT 23.3) and for passaging and freezing embryonic stem cells (UNIT 23.3)

NOTE: All cell culture incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

Plate out embryos

1. Prepare gelatinized 24-well tissue culture plates with mitotically inactivated MEF (mouse embryo fibroblast) feeder layers as described in UNIT 23.3.

Plates can be prepared a few hours in advance or the day before.

2. Replace medium in each well with modified ES medium prior to use. Drop individual blastocysts into separate wells and return plates to incubator.

Aim for the center of the well, otherwise the blastocysts will be difficult to observe with the microscope because of distortion near the edge of wells.

3. Observe embryos daily with an inverted microscope.

The majority of blastocysts will hatch from the zona pellucida and attach to the plate by the second or third day. Figure 23.4.1A shows an attached blastocyst approximately 48 hr after plating. The zona pellucida is not present. In this early stage, the embryo still retains its original shape.

**Manipulating the
Mouse Genome**

23.4.1

Contributed by David A. Conner

Current Protocols in Molecular Biology (2000) 23.4.1-23.4.9

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 52

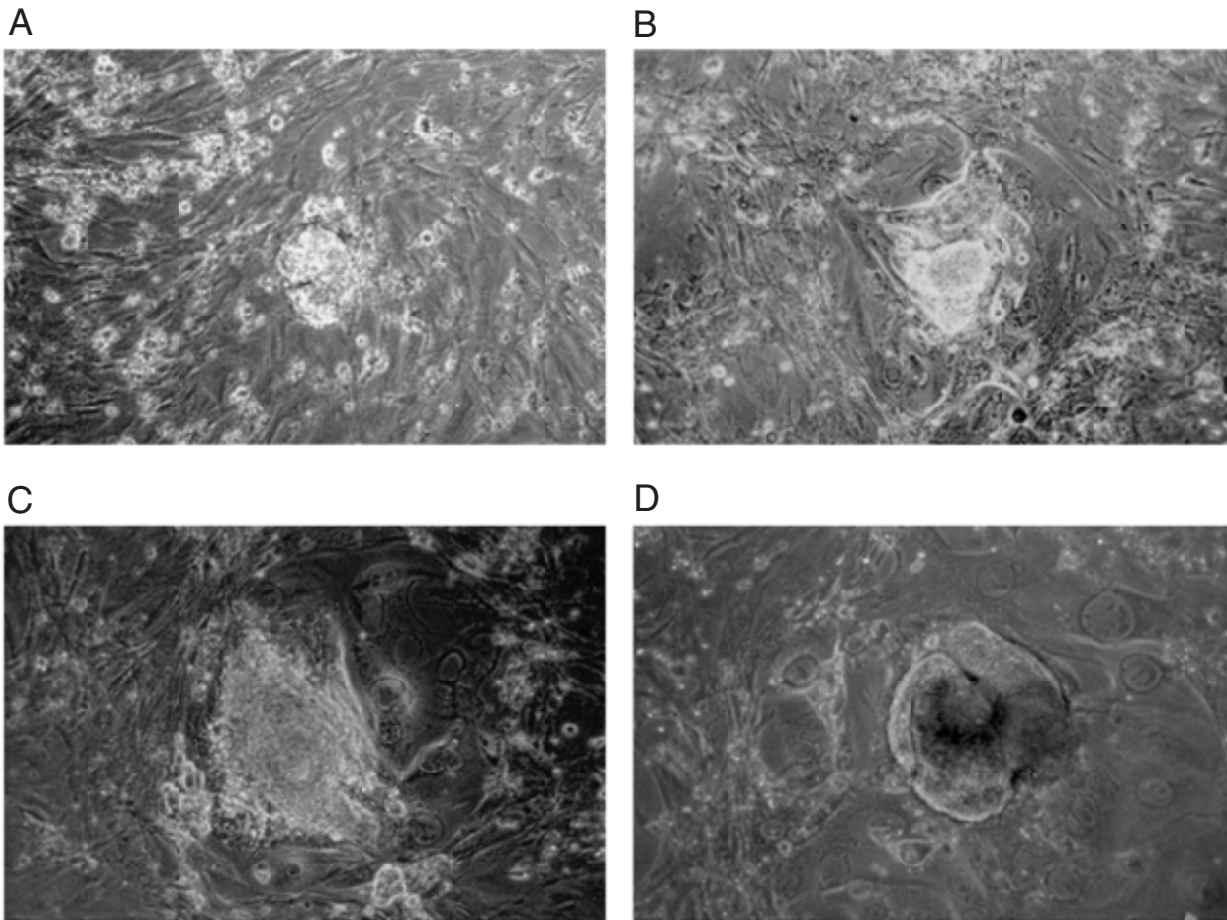


Figure 23.4.1 Morphology of plated blastocysts as shown in phase-contrast images of blastocysts at various time points after plating on MEF feeder cells. **(A)** Appearance of a blastocyst shortly after attachment to the plate ~48 hr after plating. **(B)** Typical appearance of a blastocyst 3 to 4 days after plating. Note the expansion of inner cell mass and trophoblast cells spreading out beneath. **(C)** Typical appearance of blastocyst ready to pick. **(D)** Appearance of a blastocyst that has been left on the plate 1 day too long. Note the pigmentation on the top surface indicating substantial differentiation.

4. Change medium after embryos have attached to the plate.

The trophoblast cells will spread out quickly under the embryo. Over the next several days the inner cell mass will begin to expand. Figure 23.4.1B shows the inner cell mass at an early stage of expansion. Note the large flattened trophoblast cells under the inner cell mass. The inner cell mass is expanding as a multilayered colony.

Pick inner cell mass outgrowths

Approximately 5 or 6 days after plating, the inner cell mass outgrowths will be ready to disperse. Figure 23.4.1C shows an outgrowth ready to pick. Figure 23.4.1D shows an outgrowth that has been left too long. If the outgrowth is left on the plate for too long, the cells begin to differentiate. This is typically observed as an accumulation of pigmented cells on the top of the colony. Round, loosely attached cells may also appear on the top of the outgrowth as differentiation occurs. Outgrowths should be picked before significant pigment accumulates. Outgrowths with a small amount of differentiation should still be dispersed, because they can yield decent lines.

5. Prepare gelatinized 24-well plates with feeder cells in advance with enough wells for each of embryo (UNIT 23.3).

6. Replace medium in embryo cultures 2 to 4 hr prior to picking.

This is thought to maximize plating efficiency.

7. Rinse wells with Hanks' balanced salt solution (HBSS, calcium- and magnesium-free). After aspirating, add a small volume of HBSS to wells to keep the cell layer from drying out.

Process only a few wells at a time.

8. Pick up inner cell mass outgrowth using an automatic pipettor with a 20- μ l tip and place the cell clump in a well of a 96-well U-bottom plate containing 25 μ l DPBS-EDTA in each well.

Try to transfer the cell mass in a minimal volume (i.e., $\leq 5 \mu$ l). Some investigators find that picking is more easily performed using an inverted microscope with a low-power objective or a dissecting microscope. However, the cell mass can be seen by eye at this stage. With some practice, picking by eye is less tedious than picking with a microscope. In either case, the outgrowth can be dislodged by scraping the tip of the pipet around the cell mass to tear the trophoblast and MEF layer. Then, the inner cell mass can be picked up by simultaneously nudging the clump with the pipet tip and releasing the pipet plunger to aspirate the cells. It does not matter if some fibroblasts and trophoblasts are transferred with the inner cell mass. After placing the cells in the 96-well plate, check under the microscope to verify that the picking procedure was successful.

9. Add 50 μ l of 0.25% trypsin-EDTA to each cell containing well of the 96-well plate and place in incubator for 5 to 10 min.

Do not pick too many at one time. The first outgrowth picked should not sit in PBS-EDTA for longer than 10 min prior to addition of trypsin.

10. Add 100 μ l modified ES medium to each well to inhibit trypsin. Triturate 3 to 10 times using a Gilson-style pipettor with 200- μ l tips. Transfer cells to a new 24-well feeder plate.

Try to disperse the outgrowth into several small clumps. Check with microscope to determine the degree of dispersal. Differentiation will occur if cells are not dispersed. However, do not try to achieve a single-cell suspension; trypsinization and trituration sufficient to achieve single-cell suspension usually kills cells at this stage.

Screen and expand putative ES cell lines

11. Monitor plates daily (within 2 days after plating, colonies should be visible). Mark the wells with ES cell-like colonies and observe their morphology on successive days.

Look for ES cell colonies like those depicted in Figure 23.4.2A and B. These will appear as small rounded clumps of cells with sharp refractal edges; nuclei are usually distinct, but cell borders are not (see UNIT 23.3 for more detail and pictures). Most wells will contain some embryo-derived cells. Figure 23.4.2, panels C through E, shows common morphologies of non-ES cell colonies. 5% to 15% of the wells may contain ES cell-like colonies when using embryos from the 129 mouse strain (see Background Information and see Troubleshooting regarding the effects of mouse strain on ES cell yields).

12. After 3 to 6 days, passage ES cell colonies as follows:

- a. In wells with only a few ES cell colonies, pick cells by pipet as described for the initial inner cell mass outgrowths (steps 5 to 10) and replate on gelatinized 24-well feeder plates.

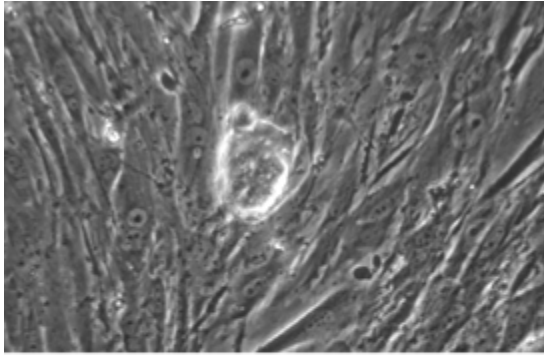
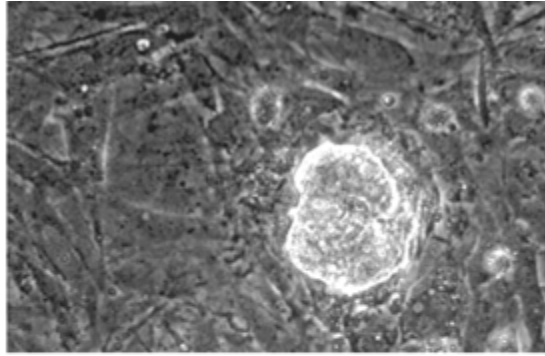
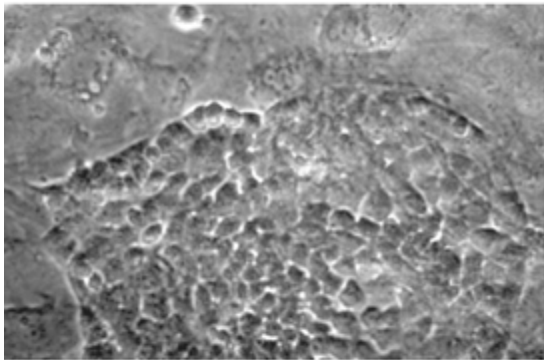
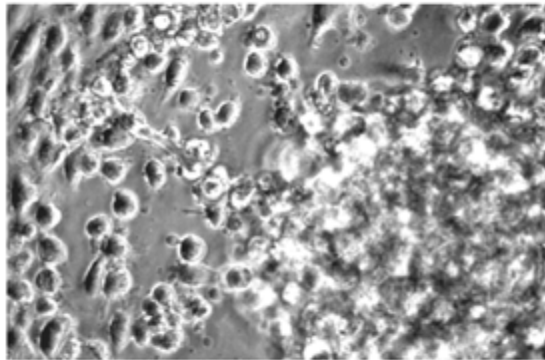
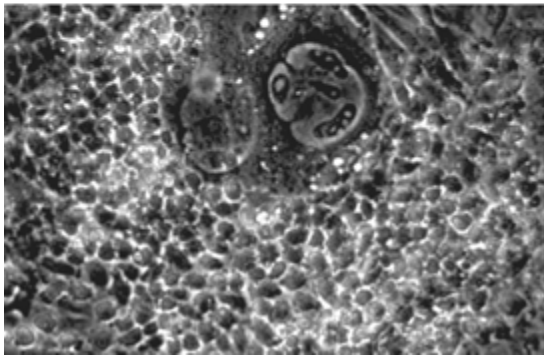
A**B****C****D****E**

Figure 23.4.2 Cell types observed after inner cell mass disaggregation. Phase-contrast images of some of the different cell morphologies seen 2 to 3 days after disaggregation. **(A, B)** Appearance of putative ES cells. **(C,D,E)** Appearance of common non-ES cell-like colonies. The cells in panel C resemble ES cells to some degree because of the distinct edge. However, unlike ES cells they are growing as a monolayer and the individual cells have distinct borders. Colonies of this type with more tightly packed cells look exactly like ES cells. After passage these misleading colonies will spread out and become easy to distinguish from true ES cells. The round, highly refractile cells in panel D look like the cells that often accumulate on the top of differentiating ES cell colonies. Note that the colony does not have a sharp, refractile edge. The colony in panel E has a diffuse border with giant trophoblast cells visible in the center.

**Mouse
Embryonic Stem
(ES) Cell Isolation**

23.4.4

- b. In wells with many ES cell colonies, trypsinize the entire well and replate on gelatinized feeder plates with a larger surface area (e.g., 6-well or 25-cm² plates) using the protocol described for passaging established ES cell lines (UNIT 23.3).

When picking individual colonies, all colonies from a single well can be combined. Deciding when to disperse the colonies depends on colony size and total plate density as described in UNIT 23.3. Passage all wells that have colonies that look remotely like ES cells and observe them daily afterwards; non-ES cells will become apparent after passage. In addition, colonies from wells with mixed morphologies should be split. ES cells will usually outgrow other cell types after several passages.

13. Expand each line that maintains an ES cell morphology until there are enough cells to freeze down 3 or more vials at a cell density per vial of 1–3 × 10⁶ cells/ml (see UNIT 23.3).

Typically, enough cells should be present on a 25-cm² flask at medium density. The protocols for expansion and freezing are described in UNITS 23.2 & 23.3. Samples of new cell lines should be analyzed for mycoplasma contamination. Investigators may want to karyotype new lines; however, most early-passage lines will have a normal complement of 40 chromosomes.

ES CELL SEX DETERMINATION

Male ES lines are used most commonly, primarily because of concerns of X chromosome instability in female ES lines (Robertson et al., 1983). This protocol describes a simple polymerase chain reaction (PCR) screen to determine the sex of an ES cell line.

Additional Materials (also see Basic Protocol)

Digestion buffer (UNIT 23.5)

Saturated NaCl

95% ethanol

10× amplification buffer (UNIT 15.1)

25 mM 4dNTP mix (UNIT 15.1)

Primers (Kunieda et al., 1992):

Set 1: SRY2: TCTTAAACTCTGAAGAAGAGAC

SRY4: GTCTTGCCTGTATGTGATGG

Set 2: NDS3: GAGTGCCTCATCTATACTTACAG

NDS4: TCTAGTTCATTGTTGAGTTGC

5 U/μl *Taq* DNA polymerase

3% (w/v) agarose gel

Molecular weight markers

1.5-ml microcentrifuge tubes

55°C incubator

Additional reagents and equipment for PCR amplification (UNIT 15.1), culture of ES cells (UNITS 23.2 and 23.3), and agarose gel electrophoresis (UNIT 2.5)

Prepare DNA

1. Plate ES cells onto gelatinized 24-well plates without feeder cells. Grow these cells for 3 to 5 passages without feeder cells. Dilute cells at least 1:10 at each split.

The cells must be diluted out to prevent contamination of the PCR reactions from any residual feeder cells. ES cells can be grown in standard ES cell medium (UNIT 23.3) without antibiotics at this stage. Some differentiation will occur in the absence of the feeder layer.

2. Remove medium. Rinse well with HBSS.

SUPPORT PROTOCOL

3. Add 300 μ l digestion buffer, transfer to a 1.5-ml microcentrifuge tube, and incubate overnight at 55°C.

Plenty of DNA can be obtained from a single well of a 24-well plate, but it is often more convenient to use a larger surface area, such as that of a 25-cm² flask.

Use 1.5 ml of digestion buffer for a 25-cm² flask and adjust the volume of subsequent solutions proportionally.

4. Add 150 μ l saturated NaCl and vortex vigorously (the solution will turn milky white). Add 2 vol of 95% ethanol (the solution will turn clear except for precipitated DNA).

Some investigators precipitate the DNA using 2 vol ethanol (or 1 vol isopropanol) without adding salt. However, the DNA pellet resuspends more easily if salt is added.

5. Resuspend DNA pellet in 50 μ l water. Determine DNA concentration by measuring the absorbance at 260 nm (APPENDIX 3D).

Perform PCR

6. Perform 2 PCR reactions with each sample: one with the SRY primers and the other with the NDS primers. Mix the following in a 0.5-ml thin-walled PCR tube for each primer set.

2.5 μ l 10 \times PCR buffer (final MgCl₂ concentration 1.5 mM)
0.2 μ l 25 mM 4dNTP mix
0.5 μ M primer 1
0.5 μ M primer 2
0.2 to 1.0 μ g DNA template
0.5 U *Taq* DNA polymerase
up to 25 μ l H₂O

7. Using the following parameters, run the PCR.

35 cycles	30 sec	94°C	(denaturation)
	30 sec	50°C	(annealing)
	60 sec	72°C	(extension)

Analyze the product

8. Run 10 μ l of the products on a 3% agarose gel with the appropriate molecular weight markers.

The SRY primers are derived from the sex-determining region of the Y chromosome. Male cells will show the 404-bp product; female cells will have no product. The NDS primers span a microsatellite dinucleotide repeat on the X chromosome and serve as a positive control; both male and female cells should show the 244-bp product.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

NOTE: Prepare all cell culture solutions from tissue culture-grade reagents. Use tissue culture-grade water (high resistance and endotoxin free). Sterilize all final solutions by filtration or prepare from sterile stocks. Use disposable sterile plasticware to prevent microbial and detergent contamination.

Modified ES medium

Dulbecco's modified Eagle medium (DMEM), with high glucose (4500 mg/liter), L-glutamine, sodium pyruvate, and pyridoxine (or pyridoxal) hydrochloride (e.g., Life Technologies, JRH Biosciences, or Sigma) containing:
20% (v/v) FBS, heat inactivated for 30 min at 56°C
1× MEM nonessential amino acids (Life Technologies)
2 mM glutamine (from 100× stock; Life Technologies)
0.1 mM 2-mercaptoethanol
20 mM HEPES, pH 7.3
500 to 1000 U/ml murine leukemia inhibitory factor (LIF; Life Technologies, Chemicon)
1× penicillin/streptomycin (from 100× stock; Life Technologies)
1× nucleoside stock (see recipe; optional)
Store at 4°C for up to 1 week

The quality of FBS should be screened by assessing plating efficiency, morphology, and toxicity on ES cells plated at low density in 10%, 15%, and 30% heat-inactivated FBS. Plating efficiency should be ~10% and should be similar for all FBS concentrations. The author has had good success with FBS from Hyclone. Several companies sell FBS that has been screened for ES cell growth using the same techniques. Although it is expensive, laboratories that are new to ES cell culture may benefit from prescreened FBS. In addition, prescreened FBS provides a standard of comparison for other FBS lots.

100× nucleoside stock

In 100 ml of tissue culture grade water dissolve:

80 mg adenine
73 mg cytidine
85 mg guanosine
24 mg thymidine
73 mg uridine

Dissolve by warming at 37°C and filter sterilize. Aliquot and freeze at –20°C. Thaw at 37°C and mix vigorously to redissolve nucleosides prior to use.

The addition of a nucleoside stock to the modified ES medium is recommended by Robertson (1987).

COMMENTARY

Background Information

Pluripotent mouse embryonic stem cells were first isolated by Evans and Kaufman (1981) and Martin (1981). The protocol outlined in this unit is essentially that of Axelrod (1984) and Robertson (1987) with some modifications. The primary differences are related to the culturing conditions including the use of MEFs as feeder layers instead of STO (SIM mouse embryo fibroblasts resistant to thioguanine and ouabain) cells and the addition of recombinant leukemia inhibitory factor (LIF) to inhibit differentiation. Evans and Kaufman (1981) used delayed blastocysts, reasoning that the increase in the number of cells in these embryos would increase the likelihood of successful stem cell isolation. Martin (1981) used immunosurgery to isolate the inner cell mass and cultured on STO feeder cell layers

with conditioned medium from embryonal carcinoma cells. Many of these ingenious tricks are no longer necessary because of a more refined understanding of the cell culture conditions required to maintain undifferentiated ES cells. Indeed, ES cell lines capable of contributing to the germ line have been isolated in media supplemented with LIF in the absence of feeder cell layers (Pease et al., 1990).

In addition to methodology, genetic background can affect the efficiency of ES cell isolation. The majority of ES cells used for gene targeting are derived from 129 substrains, in part because of the ease with which ES cells can be established from this strain. The nomenclature of the 129 strain has been revised recently; refer to Festing et al. (1999) to clarify strain names. Lines have been isolated with more difficulty from some other common

strains such as C57BL/6 (Ledermann and Burki, 1991), Balb/cJ (Noben-Trauth et al., 1996), and C3H/He (Kitani et al., 1996). The factors that influence strain permissiveness are not known.

Critical Parameters and Troubleshooting

The isolation of ES cell lines is not a complicated procedure; however, this does not mean that the procedure is always successful. Several factors are important to consider. High-quality cell culture reagents and proper aseptic technique are necessary for success. Consult *UNITS 23.2 & 23.3* for more detail on these issues. As indicated above, the strain of mouse from which the embryos are derived can have a dramatic effect on success. If a strain other than a standard 129 substrain must be used, first consult the literature for reports of successful ES cell derivation and for any strain-specific approaches. Second, increase the number of starting embryos, because the yield will probably be lower. Third, verify the technique and reagents used by isolating lines from a permissive strain. The procedure may not be successful even with these modifications. Consider more elaborate approaches such as that reported by McWhir et al. (1996).

The most common problems during isolation of ES cell lines are encountered during the initial dispersion of the inner cell mass outgrowth. Under-trypsinization or very limited trituration will result in transfer of the whole cell clump without dispersion. The large colony will differentiate rapidly. At this stage, the well can sometimes be saved by immediately picking the outgrowth a second time, trypsinizing, and replating. Over-trypsinization or extensive trituration also creates a problem. Single cells do not clone well at this stage. Dispersing the outgrowth into a single-cell suspension will usually result in cell death or differentiation. Both of these problems can be avoided by working with only a few outgrowths at one time and by monitoring the procedure with a microscope. In addition, after dispersal, the wells should be scanned carefully for several days. Single ES cell-like colonies are easy to miss, especially if they sit near the edge of the well, where the image is distorted.

Anticipated Results

With embryos derived from 129 strains (e.g., 129X1/SvJ or 129S6/SvEvTac), 5% to 15% of the inner cell mass outgrowths should give rise

to ES cell lines. Half of the resultant lines should be male.

Time Considerations

Initial inner cell mass outgrowths can be picked 5 to 6 days after plating the embryos. Within 3 to 5 days, ES cell-like clumps can be picked and dispersed; 3 to 5 days later wells containing ES cells can be expanded, and subsequently frozen 2 to 4 days later. For sex determination, aliquots of the cell lines should be expanded without feeder cells for 1 to 2 weeks (i.e., 3 to 5 passages).

Literature Cited

- Axelrod, H.R. 1984. Embryonic stem cell lines derived from blastocysts by a simplified technique. *Devel. Biol.* 101:225-228.
- Evans, M.J. and Kaufman, M.H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292:154-156.
- Festing, F.W., Simpson, E.M., Davisson, M.T., and Mobraaten, L.E. 1999. Revised nomenclature for strain 129 mice. *Mamm. Genome* 10:836.
- Hogan, B., Beddington, R., Constantini, F., and Lacy, E. 1994. *Manipulating the Mouse Embryo: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Kitani, H., Takagi, N., Atsumi, T., Kawakura, K., Imamura, K., Goto, S., Kusakabe, M., and Fukuta, K. 1996. Isolation of a germline-transmissible embryonic stem (ES) cell line from C3H/He mice. *Zool. Sci.* 13:865-871.
- Kunieda, T., Xian, M., Kobayashi, E., Imamichi, T., Moriwaki, K., and Toyoda, Y. 1992. Sexing of mouse preimplantation embryos by detection of Y chromosome-specific sequences using polymerase chain reaction. *Biol. Reprod.* 46:692-697.
- Ledermann, B. and Burki, K. 1991. Establishment of a germ-line competent C57BL/6 embryonic stem cell line. *Exp. Cell Res.* 197:254-258.
- Martin, G.R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 78:7634-7636.
- McWhir, J., Schnieke, A.E., Ansell, R., Wallace, H., Colman, A., Scott, A.R., and Kind, A.J. 1996. Selective ablation of differentiated cells permits isolation of embryonic stem cell lines from murine embryos with a non-permissive genetic background. *Nature Genet.* 14:223-226.
- Noben-Trauth, N., Kohler, G., Burki, K., and Ledermann, B. 1996. Efficient targeting of the IL-4 gene in a BALB/c embryonic stem cell line. *Transgenic Res.* 5:487-491.
- Pease, S., Braghetta, P., Gearing, D., Grail, D., and Williams, R.L. 1990. Isolation of embryonic stem (ES) cells in media supplemented with recombinant leukemia inhibitory factor (LIF). *Dev. Bio.* 141:344-352.

Robertson, E.J. 1987. Embryo-derived stem cell lines. *In* Teratocarcinomas and Embryonic Stem Cells: A Practical Approach (E.J. Robertson, ed.) pp.71-112. IRL Press, Oxford.

Robertson, E.J., Evans, M.J., and Kaufman, M.H. 1983. X-chromosome instability in pluripotential stem cell lines derived from parthenogenetic embryos. *J. Embryol. Exp. Morph.* 74:297-309.

Key References

Hogan, B., et al., 1994. See above.

Robertson, E.J. 1987. See above.

Wurst, W. and Joyner, A.L. 1993. Production of targeted embryonic stem cell clones. *In* Gene Targeting: A Practical Approach (A.L. Joyner, ed.) pp. 33-61. IRL Press, Oxford.

These three references, written by experts in the field, represent the best compilations of ES culture and isolation methods.

Internet Resources

<http://www.biosupplynet.com>

Search this web site for "embryonic stem cell reagents" to obtain a current list of suppliers that provide medium and ES cells.

Contributed by David A. Conner
Harvard Medical School
Boston, Massachusetts

Production of a Heterozygous Mutant Cell Line by Homologous Recombination (Single Knockout)

Gene targeting by homologous recombination allows the introduction of specific mutations into any cloned gene. In the method described here, the gene of interest is inactivated by interrupting its coding sequence with a positive selectable marker (e.g., *neo*). Expression of *neo* is obtained by including the phosphoglycerate kinase (PGK) promoter in the construct. To enrich for clones in which the target gene has undergone homologous recombination over those in which random integration of the construct has occurred, a negative selectable marker, herpes simplex virus thymidine kinase (HSV-TK), is included in the construct outside the region of homology to the target gene. Depending upon the target gene, it may be easier to assemble the construct by adding the *neo* and *TK* genes to the cloned target gene or by adding two fragments of the target gene to a plasmid containing the *neo* and *TK* genes (e.g., pNTK, Fig. 23.5.1).

If the Cre-*loxP* system (UNIT 23.1) is to be used for removing the selectable marker, or for tissue-specific or temporally controlled knockout, then a construct already containing *loxP* sites flanking the marker is more convenient (e.g., pTKLNL, Fig. 23.5.2A). If selection for loss of selectable marker is desired, a construct that also contains a negative selectable marker can be used (e.g., pTKLNCL, Fig. 23.5.2, panel B).

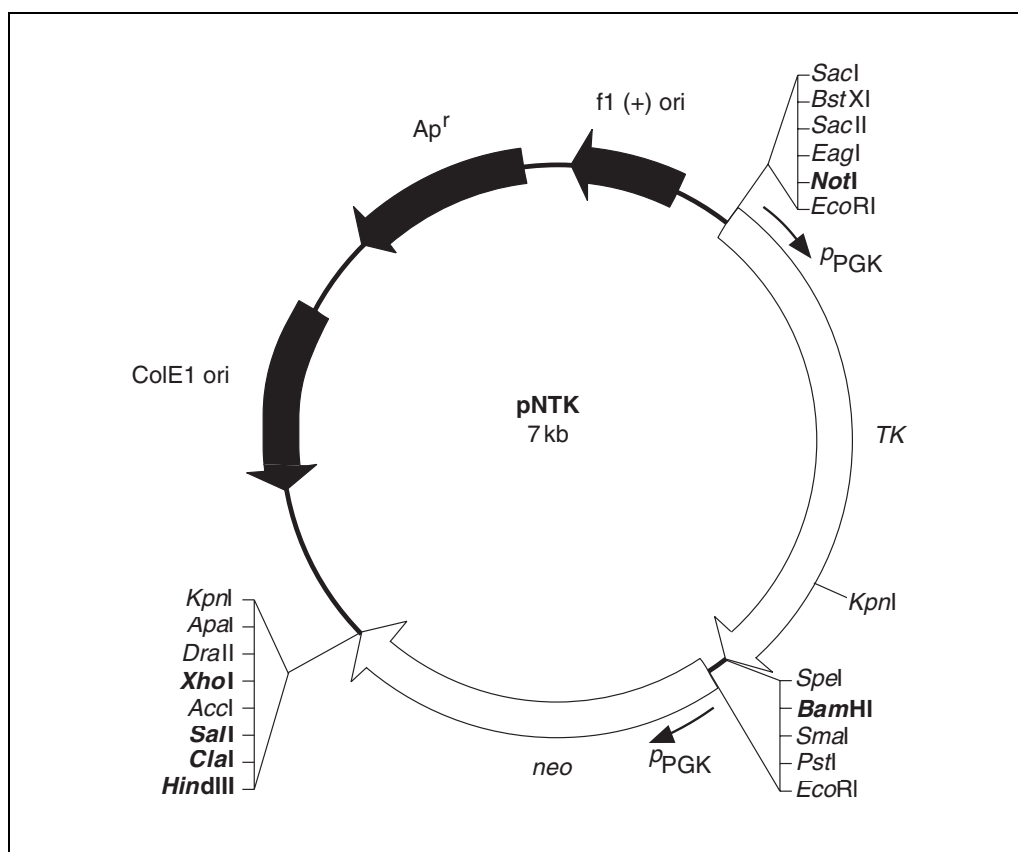


Figure 23.5.1 pNTK vector. Both the *neo* and *TK* genes are driven by a PGK promoter (p_{PGK}) that is expressed in ES cells. Unique restriction enzyme sites that are useful are indicated in bold. One genomic fragment can be cloned into the *Bam*HI site. A second genomic fragment can be cloned into the *Hind*III, *Cl*aI, *Sal*I, *Xho*I sites. A site should be preserved that will linearize the construct, leaving the majority of plasmid vector sequences attached to the *TK* gene (e.g., *Xho*I).

Contributed by Richard Mortensen

Current Protocols in Molecular Biology (2000) 23.5.1-23.5.11

Copyright © 2000 by John Wiley & Sons, Inc.

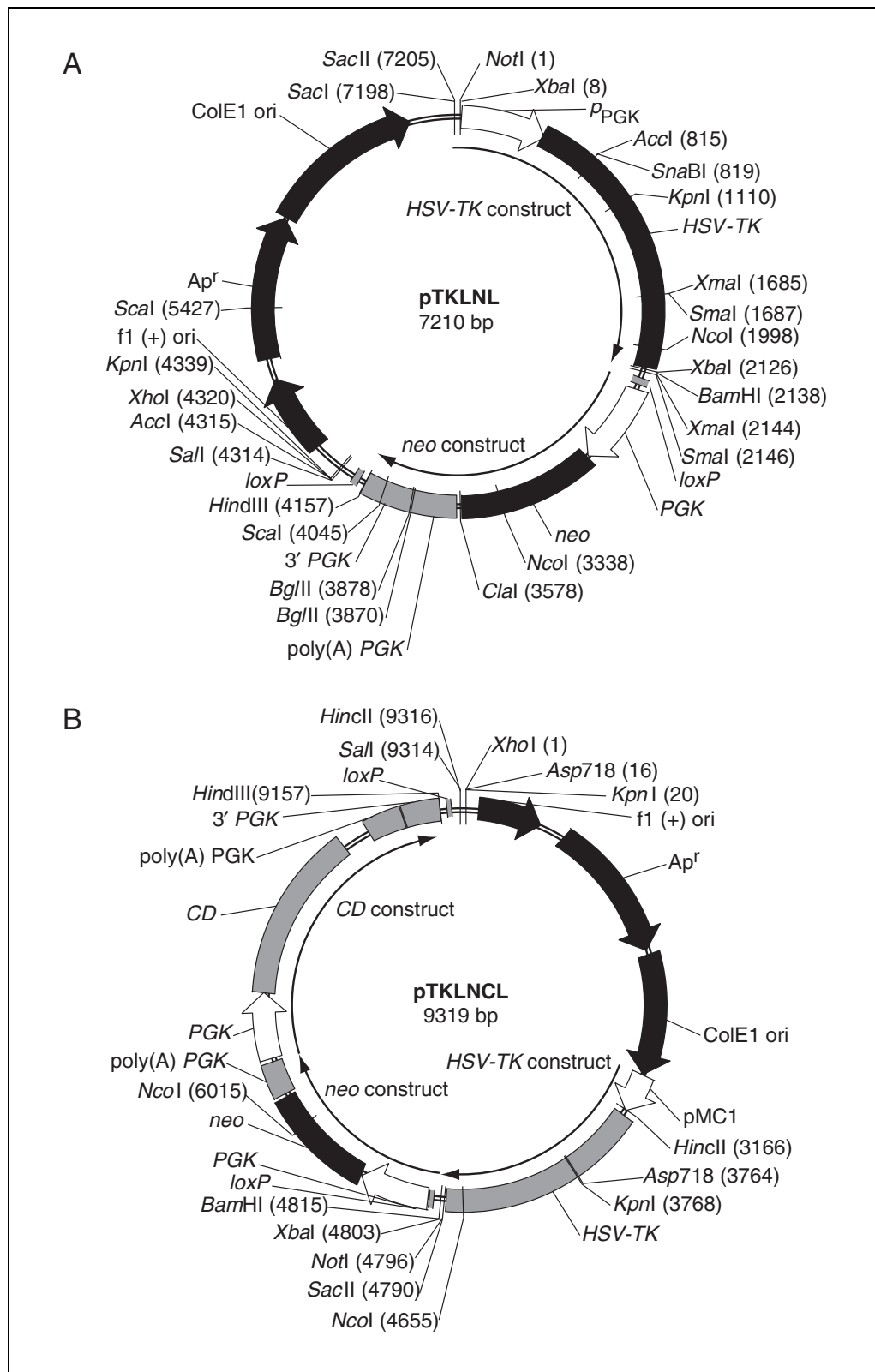


Figure 23.5.2 Constructs containing *loxP* sites surrounding a positive selectable marker, *neo* (**A**), or both a positive and a negative selectable marker, *neo* and cytosine deaminase (*CD*; **B**). Constructs can be made by insertion of homologous sequences in unique restriction sites outside the *loxP* sites. If conditional targeting constructs are desired (as in Fig. 23.1.7), a third *loxP* site can be inserted into the region of homology and then the two regions of homology inserted into the vectors. Another version of these plasmids is also available with the *TK* and *CD* reversed (Milstone et al., 1999).

STRATEGIC PLANNING

Figure 23.5.3 illustrates the production, selection, and identification of targeted gene disruption by homologous recombination. A replacement targeting construct requires the assembly of several different DNA sequences:

1. A genomic clone (preferably >10 kb) of the gene of interest, generally encoded on a bacteriophage or cosmid clone containing homologous sequences to be included in the construct. DNA isogenic to the ES cells (i.e., derived from the same animal strain) is preferred but not essential. An alternative for genes of sufficient size, and regions

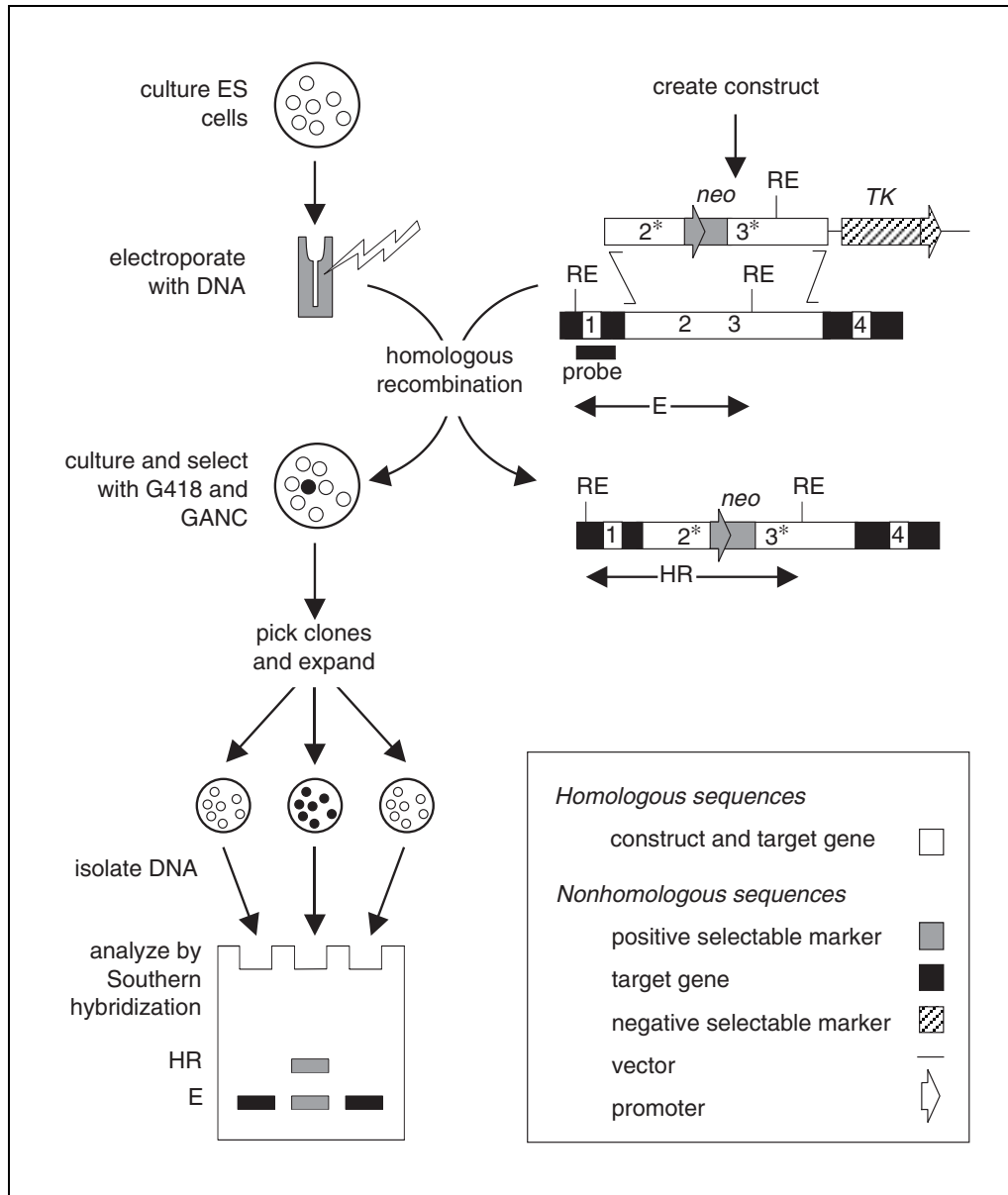


Figure 23.5.3 Production, selection, and identification of targeted gene disruption by homologous recombination. An example of a restriction enzyme site (RE) and hybridization probe that can be used to identify cells in which homologous recombination has occurred (shaded colony) is shown. The predicted size of the restriction fragment generated from an unaltered target gene (E) and a target gene that has undergone homologous recombination (HR) is shown. If equal amounts of DNA are present in the lanes of the Southern blot, the intensity of each of the two hybridizing fragments from the DNA of a homologous recombinant clone will be half of the intensity of the hybridizing fragment from unaltered clones.

for which mouse sequences are known, is long PCR (Cheng et al., 1994). This approach works for many genes as long as they have large enough introns, and is particularly convenient when the intron-exon structure of the gene is known. Long PCR may introduce mutations, which may decrease homologous recombination rates just as nonisogenic DNA can. In addition, the mutations may occur in important parts of the gene; this risk makes the approach less suitable for conditional knockouts or subtle mutations.

2. Additional cloned target-gene DNA sequences not included in the construct, to be used as a hybridization probe to identify homologous recombinants. This probe can often be isolated from the same bacteriophage or cosmid clone that provided the homologous sequences included in the construct. The hybridization probe will hybridize with either an unaltered target gene or a target gene that has undergone homologous recombination, but will not hybridize with a construct that entered the genome by random integration.
3. A positive selectable marker, such as the gene encoding neomycin phosphotransferase (*neo*) or hygromycin-B-phosphotransferase (*hyg*), which is used to disrupt the target gene. If a homozygous mutant cell line is an ultimate goal, it is recommended that the *neo* coding sequence contain the point mutation that decreases the phosphotransferase activity (Yenofsky et al., 1990). Using the PGK promoter and the wild-type *neo* gene may result in cells containing a single *neo* gene that are resistant to >10 mg/ml G418, thus precluding the use of higher G418 concentration to isolate clones containing two *neo* genes (see UNIT 23.6).
4. A negative selectable marker such as HSV-TK, which is used to enrich for ES cell clones in which homologous recombination has occurred in the target gene over clones in which random integration of the construct has occurred.

BASIC PROTOCOL

GENE TARGETING IN EMBRYONIC STEM CELLS

The basic protocol is divided into three parts. First, it outlines the assembly of a replacement targeting construct and considerations in choosing its exact structure. Second, it briefly describes the culture of embryonic stem (ES) cells and the method for introducing the construct DNA into ES cells. A more detailed description of culturing ES cells and maintaining their undifferentiated state is found in UNITS 23.2 & 23.3. Third, it outlines the method for identifying clones in which the target gene has been altered by homologous recombination. The resulting homologous recombinants are heterozygous (one allele of the target gene is altered by homologous recombination and one allele is normal) and can be used to produce transgenic murine lines or to produce homozygous mutant cell lines (in which both alleles of the target gene are altered; UNIT 23.6).

Materials

Target gene from genomic library isogenic with ES cell line (e.g., 129 SV library; Stratagene)
Plasmid vector (e.g., pNTK, available from R. Mortensen; see Fig. 23.5.1)
95% ethanol
Sterile H₂O
Embryonic stem (ES) cells (UNITS 23.2 & 23.3; ATCC)
ES/LIF medium (see recipe)
Trypsin/EDTA: 0.25% (w/v) trypsin/1 mM EDTA (20 mM HEPES, pH 7.3, optional)
ES medium (see recipe)
Electroporation buffer (see recipe)

Production of a Heterozygous Mutant Cell Line

23.5.4

G418 (UNIT 9.5)

Gancyclovir (GANC)

Freezing medium (see recipe)

Digestion buffer (see recipe)

Saturated NaCl (see recipe)

1% agarose gel (UNIT 2.5A)

Tissue culture hood

Gelatin-coated tissue culture plates (UNIT 23.3): 100-mm plates and 24-well microtiter plates

4-mm electroporation cuvettes

Pipet tips, sterilized by autoclaving

Nylon membrane

Additional reagents and equipment for subcloning DNA (UNIT 3.16), restriction enzyme digestion (UNIT 3.1), phenol/chloroform extraction of DNA (UNIT 2.1A), agarose gel electrophoresis (UNIT 2.5A), electroporation (UNIT 9.3), ES cell culture (UNITS 23.2 & 23.3 and APPENDIX 3F), stable transformation using selective medium (UNIT 9.5), and Southern blotting and hybridization (UNITS 2.9 & 2.10)

NOTE: All tissue culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise noted.

Create a replacement construct

1. Select a portion of the target gene to include in the construct and a separate portion of the target gene to use as a probe for hybridization of Southern blots to identify cells in which homologous recombination has occurred (see step 19).

It should contain a rare restriction site within an exon encoding an important region of the protein (or an exon upstream of such a region) that is ideally flanked by >1 kb of target gene DNA on each side (most constructs are made with 2 kb). The rate of homologous recombination may increase with increasing lengths of homologous DNA up to 15 kb.

2. Construct a clone in a plasmid vector (UNIT 3.16) such that the *neo* gene interrupts the gene of interest, leaving regions of homology on either side of the *neo* gene. Include a thymidine kinase (*TK*) gene in the replacement construct outside the regions of homology (Fig. 23.5.4).

*This construction can be accomplished by either adding the *neo* and *TK* sequences to the subcloned homologous sequences or by adding regions of homologous sequences to a plasmid already containing the *neo* and *TK* genes (e.g., *pNTK*, see Fig. 23.5.1).*

3. Digest the construct DNA with a restriction enzyme to linearize (UNIT 3.1).

*Linearize the construct DNA so that the plasmid vector sequences remain attached to the *TK* gene. This will help preserve the activity of *TK* gene if any loss of DNA sequence occurs during random insertion of the construct into the genome.*

4. Purify and sterilize the digested construct DNA by phenol/chloroform extraction (UNIT 2.1A).
5. Precipitate the DNA by adding 2 vol of 95% ethanol and microcentrifuging 30 sec. Using sterile technique in a tissue culture hood, remove the supernatant and allow the pellet to air dry until only slightly moist.
6. Dissolve the pellet in 100 µl sterile water. Check for complete digestion and estimate DNA concentration by electrophoresis on an agarose gel (UNIT 2.5A).

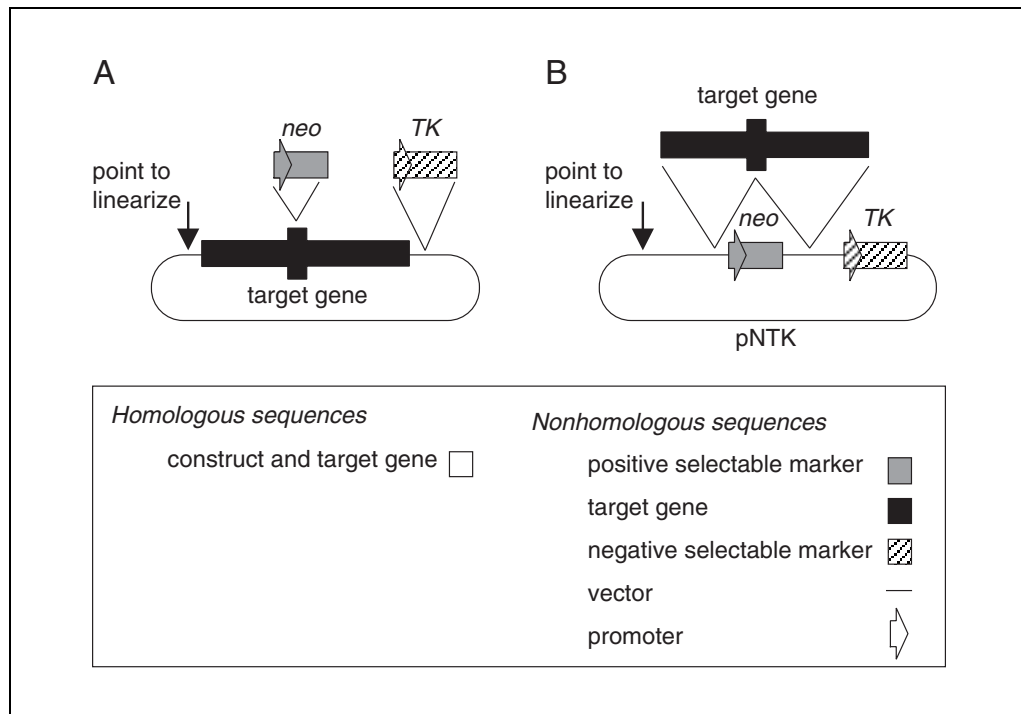


Figure 23.5.4 Two strategies to create a replacement construct. In method **A** the target gene fragment is subcloned into a plasmid vector, then $p_{PGK-neo}$ is inserted into a rare restriction enzyme site in the target-gene fragment and p_{PGK-TK} is inserted into the plasmid vector near the target gene. In method **B** the target-gene fragment is cleaved into two pieces that are subcloned into the polylinker sites of pNTK (see Fig. 23.5.1). Note that the relative orientation of homologous fragments in the construct must retain that found in the target gene.

Transfect construct and select ES cells

- Culture ES cells in ES/LIF medium (UNITS 23.2 & 23.3). Passage cells every 2 to 3 days by seeding a 100-mm gelatin-coated tissue culture plate with $1-2 \times 10^6$ cells/plate.

Leukemia inhibitory factor (LIF) prevents ES cells from differentiating.

Some investigators suggest passaging cells at a higher density if blastocyst injection of the cells (UNIT 23.4) is planned (e.g., 1.5×10^6 cells per 25-cm² flask). A detailed description of culture techniques for ES cells is found in UNITS 23.2 & 23.3.

- Harvest $\sim 5 \times 10^6$ to 1×10^7 cells by adding trypsin/EDTA and incubating for ~ 5 min until cells are freed from the plate surface. Dissociate to single cells by pipetting up and down five to ten times. Add 5 ml ES medium. Pellet cells and resuspend the cell pellet in 1 ml electroporation buffer in the same tube.

Typically, 10^7 cells can be obtained from a near-confluent 100-mm tissue culture plate.

- Add 1 pmol linearized, sterile construct DNA from step 6.
- Electroporate the mixture at 450 V and 250 μ F in a 4-mm electroporation cuvette (UNIT 9.3). Incubate 10 min at room temperature.

Many electroporation conditions can be used with ES cells.

- Plate cells in ES medium at $\sim 2 \times 10^6$ cells per 100-mm gelatin-coated tissue culture plate. Incubate 24 hr.
- Begin selection (UNIT 9.5) by changing ES medium to ES/LIF medium and adding G418 to 0.2 mg/ml and GANC to 2 μ M (final).

13. Continue incubation, changing medium daily using ES/LIF medium and adding G418 (0.2 mg/ml final) and GANC (2 μ m final), until single, isolated colonies are visible (typically 1 week after electroporation). Remove an individual colony from the plate using an autoclaved pipet tip, and place in a 35- μ l drop of trypsin/EDTA for 5 min. Pipet up and down about five times to dissociate cells. Transfer cells to a well of a gelatin-coated 24-well microtiter plate containing 1 ml ES/LIF medium.
14. Incubate until colonies are visible, but the cells are not differentiating (typically 3 to 4 days). Passage half of the cells to a well of a clean gelatin-coated 24-well microtiter plate. Add the remaining cells to 0.5 ml freezing medium and place at -70°C . Freeze overnight, then transfer to liquid nitrogen.

Undifferentiated cells grow in smooth, round colonies. Differentiated cells are flatter with distinct intercellular boundaries.

Proceed immediately to step 15 after placing half the cells in the freezer.

Screen for homologous recombinants

15. Incubate ES cells in 24-well microtiter plate (step 14) to near confluence (usually 2 to 3 days).

Because it is not critical to prevent differentiation of the ES cells at this stage, LIF can be omitted from the culture medium; however, the presence of LIF may help to maintain cell growth.

16. Add 300 μ l digestion buffer to each well. Transfer well contents to a 1.5-ml microcentrifuge tube, and incubate overnight at 55°C .

17. Add 150 μ l saturated NaCl and vortex vigorously (the solution will turn milky white). Add 2 vol of 95% ethanol (the solution will turn clear except for precipitated DNA).

Some investigators precipitate the DNA using 2 vol ethanol (or 1 vol isopropanol) without adding salt. However, the DNA pellet resuspends more easily if salt is added.

18. Resuspend DNA pellet in 50 μ l water. Determine DNA concentration by measuring the absorbance at 260 nm (*APPENDIX 3D*).
19. Digest 10 μ g DNA (or 10 μ l if DNA concentration was not determined) with the appropriate restriction enzyme (*UNIT 3.16*).
20. Fractionate the digested DNA on a 1% agarose gel (*UNIT 2.5A*). Transfer to a nylon membrane, and hybridize by Southern blotting (*UNITS 2.9 & 2.10*) to the target-gene hybridization probe chosen in step 1 to distinguish the unaltered target gene from a target gene that has undergone homologous recombination.
21. Select ES cell colonies that show two hybridizing fragments of approximately equal intensity—one fragment of the predicted size for the unaltered target gene and one fragment of the predicted size for a target gene that has undergone homologous recombination. If the two fragments are of unequal hybridization intensity, the cell population may not be clonal. Freeze cells and store in liquid nitrogen.
22. If desired remove selectable markers that are flanked by *loxP* sites by transient expression of Cre (see Support Protocol).

TRANSIENT EXPRESSION OF CRE FOR RECOMBINATION

Removal of sequences between the *lox* sites is accomplished by transient expression of Cre recombinase. If flanking *lox* sites are present in both alleles, Cre will recombine both alleles as efficiently as one. If selectable markers are between *lox* sites (as in Fig. 23.1.6), then sensitivity to selection media (e.g., containing G418) will be restored.

Additional Materials (also see *Basic Protocol*)

Cre expression plasmid using a promoter giving high expression levels in ES cells (e.g., pMC1 or pPGK)

12.5 mg/ml 5-fluorocytosine in PBS (if selecting against CD), sterile

1. Expand the homologously recombined clones obtained using the Basic Protocol by culturing and harvesting ES cells (see Basic Protocol, steps 7 and 8).
2. Prepare Cre expression plasmid DNA using the same procedure as for the original targeting vector (see Basic Protocol, steps 4 to 6).

Do not linearize the DNA, as this will increase the probability of genomic integration.

3. Use 1 to 2 pmol of this DNA to transfect the expanded ES cell clones (see Basic Protocol, step 10).

Transfecting more DNA will most likely increase expression; however, it also increases the probability of integration.

4. Plate cells at a lower density than for the original targeting (since survival is expected to be higher).

The target is a number of clones per plate that will allow convenient colony picking. If no selection is to be performed, plating at a few hundred electroporated cells per 100-mm plate is a reasonable starting point. A range of dilutions should be plated (at least to a few thousand per plate), since the exact survival is not accurately predictable and plating at low density will decrease survival. If the negative selectable marker cytosine deaminase (CD) is used, then plate at 1,000 to 10,000 cells per 100-mm plate, since higher densities will give complete killing due to a neighbor selection effect. The number of surviving colonies will depend on the frequency of recombination.

5. Continue to culture cells, replacing medium daily with fresh ES/LIF medium. If selecting against CD, include 250 µg/ml 5-fluorocytosine (from 12.5 mg/ml stock) in the medium.
6. Screen colonies loss of the selectable marker (Cre-induced recombination) by Southern analysis (see Basic Protocol, steps 14 to 21).

Although Cre-construct integration is an unlikely event, the clones for injection can also be screened for presence of Cre by reprobng the Southern blots or by PCR.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Digestion buffer

20 mM Tris·Cl, pH 8.0

10 mM NaCl

10 mM EDTA

0.5% (w/v) sodium dodecyl sulfate (SDS)

Store indefinitely at room temperature

Add 1 mg/ml proteinase K just before use

Electroporation buffer

20 mM HEPES, pH 7.3
137 mM NaCl
5 mM KCl
0.7 mM Na₂HPO₄
6 mM glucose
0.1 mM 2-mercaptoethanol (2-ME)
Store indefinitely at 4°C

ES medium

Dulbecco's minimum essential medium (DMEM), high-glucose + pyruvate formulation, containing:

15% FBS, heat-inactivated 1 hr at 56°C
0.1 mM 2-mercaptoethanol (2-ME)
20 mM HEPES, pH 7.3 (optional)

DMEM containing 4500 mg/liter D-glucose and pyruvate can be obtained from GIBCO/BRL.

A detailed description of culture conditions for embryo stem cells has been presented by Robertson (1987).

ES/LIF medium

ES medium (see recipe) containing 1000 U/ml leukemia inhibitory factor (LIF; GIBCO/BRL). Store ≤1 week at 4°C.

Conditioned medium from a CHO cell line overproducing LIF (Genetics Institute) can also be used at a dilution of 1:1000. An alternative to LIF for preventing differentiation of ES cells is to grow them on feeder layers of irradiated mouse embryo fibroblasts (MEF) in ES medium. Some investigators add 20 mM HEPES (pH 7.3) to culture medium.

Freezing medium

DMEM, high-glucose + pyruvate formulation, containing:

10% FBS (Hyclone), heat-inactivated 1 hr at 56°C
10% (v/v) dimethylsulfoxide (DMSO)
20 mM HEPES, pH 7.3 (optional)
Store at -20°C

Saturated NaCl

Add NaCl to distilled H₂O until no more dissolves (~6 M). Some solid NaCl should remain; decant solution for use. Store indefinitely at room temperature.

COMMENTARY

Background Information

See *UNIT 23.1* for an overview of gene targeting by homologous recombination.

Although homologous recombination has been used by yeast geneticists for some time, it has only recently been shown to occur in somatic mammalian cells. It was first demonstrated between exogenously introduced DNA sequences (Folger et al., 1982) and later between an exogenously introduced DNA construct and an endogenous gene (Smithies et al., 1985).

The mechanism of homologous recombination is not well understood but a number of characteristics are known. Homologous recom-

bination occurs more readily if the construct has free ends, rather than being circular (Wong and Capecchi, 1987). The rate of homologous recombination does not depend on the number of targets in the genome—at least when the target is present as tandem repeats of a dihydrofolate reductase (*DHFR*) amplified gene (Zheng and Wilson, 1990).

Initially homologous recombination in mammalian cells was studied by introducing a mutated *neo* or *TK* gene, then restoring *neo* or *TK* activity by correcting the mutation through homologous recombination (reviewed by Capecchi, 1989). This approach provided an

easy method to detect homologous recombination.

Embryonic stem (ES) cells and the related embryonic carcinoma (EC or EK) cells were first isolated and cultured in 1981 (Evans and Kaufman, 1981; Martin, 1981). They can give rise to a chimeric mouse when introduced into a normal blastocyst, which is then transferred into the uterus of a pseudopregnant foster mother. The ES cells contribute to all tissues of the chimeric mouse including the germ line (Bradley et al., 1984). Currently, ES cells, rather than EC cells, are used to produce chimeric mice, because the extent of chimerism and efficiency of germ line transmission is much higher with normal cells. Most ES cell lines used are derived from males because the karyotype of XY cells is more stable than that of XX cells and resulting chimeric male mice are easier to breed. Typically, the extent of the contribution of the ES cells to somatic tissues of the chimeric mouse is easily determined visually by choosing strains of mice for the sources of ES cells and blastocysts that have different coat colors.

Critical Parameters

The degree of homology between the construct and the target genome can have a dramatic effect on the rate of homologous recombination in two ways. First, homologous recombination requires stretches of exact DNA homology. The DNA used to construct the targeting vector must be from the same species as the cell in which the mutation is to be introduced. It should also be isogenic with the target cell (this is not absolutely required, but increases the probability of success). Because animal strains may differ just as individual outbred animals differ, there may be a mismatch of DNA on average every 500 bp. A single DNA mismatch is sufficient to dramatically decrease the rate of homologous recombination (Deng and Capecchi, 1992; teRiele et al., 1992). Mutations induced by making constructs using long PCR may similarly decrease homologous recombination rates.

Second, the rate of homologous recombination increases with increasing length of the homologous DNA sequence (within limits). The exact length of homologous DNA that gives the maximum recombination rate is controversial but may be as high as 15 kb (Deng and Capecchi, 1992; Hasty et al., 1991). Homology should also be >1 kb for the shorter arm (most constructs have used >2 kb). Further, fidelity of recombination can be lower if the

length of homology is <1 kb (Thomas et al., 1992).

The parameters outlined here are meant to be guidelines for the design of constructs. For some target genes and constructs, more or less homology may be required. Nonisogenic DNA recombines easily for some genes. If homologous recombination is not obtained initially, the same construct should be retransfected into the ES cells, as rate of homologous recombination can vary for unclear reasons. A second construct using a different region of the gene can also be used because different constructs of the same gene may recombine at different rates. If homologous recombination is occurring, but only at a low rate, screening pools of clones by the polymerase chain reaction (PCR; *UNIT 15.1*) may enable one to identify homologous recombinants (Koller et al., 1991).

Anticipated Results

In most cases, constructs prepared following the above guidelines should give rise to a number of homologous recombinant clones in a single transfection. Typically the frequency of homologous recombinant clones to surviving clones is 1:30 to 1:5; however, the range of reported results is large: the ratio may be $\geq 1:1000$, or homologous recombinants may even be in the majority.

The frequency of recombination reported in the literature varies from 1% to 16%; in the authors' hands, it has sometimes been as high as 30%.

Time Considerations

The entire procedure, from cloning to obtaining homologous clones, can take several months, depending on the difficulty in obtaining the construct. Once a construct is obtained however, transfection and selection of clones takes ~2 weeks. Screening clones takes 2 to 3 days. Addition of a *neo* removal step will add ~4 weeks for expansion of clones, transfection, and rescreeing.

Literature Cited

- Bradley, A., Evans, M., Kaufman, M.H., and Robertson, E. 1984. Formation of germ-line chimeras from embryo-derived teratocarcinoma cell lines. *Nature* 309:255-256.
- Capecchi, M.R. 1989. Altering the genome by homologous recombination. *Science* 244:1288-1292.
- Cheng, S., Fockler, C., Barnes, W., and Higuchi, R. 1994. Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc. Natl. Acad. Sci. U.S.A.* 91:5695-5699.

- Deng, C. and Capecchi, M.R. 1992. Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol. Cell. Biol.* 12:3365-3371.
- Evans, M.J. and Kaufman, M.H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292:154-156.
- Folger, K.R., Wong, E.A., Wahl, G., and Capecchi, M.R. 1982. Patterns of integration of DNA microinjected into cultured mammalian cells: Evidence for homologous recombination between injected plasmid DNA molecules. *Mol. Cell. Biol.* 2:1372-1387.
- Hasty, P., Rivera, P.J., and Bradley, A. 1991. The length of homology required for gene targeting in embryonic stem cells. *Mol. Cell. Biol.* 11:5586-5591.
- Koller, B.H., Kim, H.-S., Latour, A.M., Brigman, K., Boucher, R.C. Jr., Scambler, P., Wainwright, B., and Smithies, O. 1991. Toward an animal model of cystic fibrosis: Targeted interruption of exon 10 of the cystic fibrosis transmembrane regulator gene in embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 88:10730-10734.
- Martin, G.R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Nat. Acad. Sci. U.S.A.* 78:7634-8.
- Milstone, D.S., Bradwin, G., and Mortensen, R.M. 1999. Simultaneous Cre catalyzed recombination of two alleles to restore neomycin sensitivity and facilitate homozygous mutations. *Nuc. Acids Res.* 27:e10.
- Robertson, E.J. 1987. Embryo derived stem cell lines. In *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach* (E.J. Robertson, ed.) pp. 71-112. IRL Press, Oxford and New York.
- Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A., and Kucherlapati, R.S. 1985. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* 317:230-234.
- teRiele, H., Maandag, E.R., and Berns, A. 1992. Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. *Proc. Nat. Acad. Sci. U.S.A.* 89:5128-5132.
- Thomas, K.R., Deng, C., and Capecchi, M.R. 1992. High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors. *Mol. Cell. Biol.* 12:2919-2923.
- Wong, E.A. and Capecchi, M.R. 1987. Homologous recombination between coinjected DNA sequences peaks in early to mid-S phase. *Mol. Cell. Biol.* 7:2294-2295.
- Yenofsky, R.L., Fine, M., and Pellow, J.W. 1990. A mutant neomycin phosphotransferase II gene reduces the resistance of transformants to antibiotic selection pressure. *Proc. Nat. Acad. Sci. U.S.A.* 87:3435-3439.
- Zheng, H. and Wilson, J.H. 1990. Gene targeting in normal and amplified cell lines. *Nature* 344:170-173.

Contributed by Richard Mortensen
University of Michigan Medical School
Ann Arbor, Michigan

Production of a Homozygous Mutant Embryonic Stem Cell Line (Double Knockout)

Homozygous mutant embryonic stem (ES) cell lines have proven very useful in studying gene and protein function, and there are several reasons why this system might be preferred. Because many genes express their phenotype at the level of individual cells, using these cell lines eliminates the time and expense of producing a mutant animal. They can be isolated directly from a heterozygous mutant cell line (UNIT 23.5) without rederiving a cell line or using primary tissues from a mutant animal. In some cases, analysis of the phenotype of homozygous mutant cells can be performed more easily in culture. In addition, homozygous mutant ES cell lines may be produced even when a homozygous mutation is lethal to the animal. Finally, microinjection of homozygous mutant ES cells into blastocysts may allow investigation of the developmental potential of cells with such lethal mutations.

Selection of a homozygous cell line depends on being able to increase antibiotic concentrations such that nearly all of the cells containing a single copy of the selectable marker gene will be killed. If the wild-type *neo* gene is used in combination with a strong promoter (e.g., phosphoglycerate kinase; PGK), cells containing one *neo* gene may not be killed by the highest G418 concentrations that can be obtained. For this technique to be effective, therefore, it is recommended that a mutant *neo* gene containing a point mutation that decreases phosphotransferase activity be used in place of wild-type *neo* (Yenofsky et al., 1990). This mutant *neo* gene is readily available because it is contained in some versions of both pMC1-*neo* and p_{PGK} -*neo* (see Fig. 23.5.1). The wild-type hygromycin-B-phosphotransferase (*hyg*) gene used in combination with the PGK promoter, p_{PGK} -*hyg*, has also worked well in the method described in the basic protocol.

Alternatively, using a construct with *loxP* sites flanking the *neo* marker, such as those derived from pTKLNL or pTKLNCL, allows the *neo* to be removed from both alleles simultaneously after homozygous cells are obtained. Using this strategy, one can insert a wild-type *neo* to inactivate a gene of interest and still use high-G418 selection to isolate homozygous cells.

SELECTION FOR HOMOZYGOUS CLONES

In this protocol, a homozygous mutant ES cell line is isolated from heterozygous mutant ES cells (UNIT 23.5) by culturing the cells in increasing concentrations of G418. The more stringent selection conditions favor cells that contain, as a result of the loss of heterozygosity, two selectable marker genes (e.g., for *neo*), and therefore two mutant alleles of the target gene. Resulting colonies are harvested and screened by northern hybridization or immunoblotting to confirm that the target gene is inactive.

Materials

- Heterozygous mutant ES cell line, frozen in liquid nitrogen (UNIT 23.5)
- ES/LIF medium (UNIT 23.5)
- G418 (UNIT 9.5)
- 100-mm tissue culture plates, gelatin coated (UNIT 23.5)

Additional reagents and equipment for recovery of frozen cell lines (UNIT 11.9), ES cell culture (UNIT 23.5), northern analysis (UNIT 4.9), and immunoblotting (UNIT 10.8)

NOTE: All tissue culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise noted.

BASIC PROTOCOL

Manipulating the Mouse Genome

23.6.1

Contributed by Richard Mortensen

Current Protocols in Molecular Biology (2000) 23.6.1-23.6.3

Copyright © 2000 by John Wiley & Sons, Inc.

Supplement 52

1. Thaw heterozygous mutant ES cells. Culture in ES/LIF medium (UNIT 23.5), passaging cells every 2 to 3 days by seeding a 100-mm gelatin-coated tissue culture plate with $1-2 \times 10^6$ cells/plate.

Selection should be performed with more than one heterozygous mutant ES cell line clone because the efficiency of conversion to homozygosity varies among cell lines. In addition, more than one clone should be analyzed for the phenotype.

2. Begin the selection by plating three 100-mm gelatin-coated tissue culture plates at 10^5 cells/plate using ES/LIF medium and adding 1.0, 1.5, or 2.0 mg/ml G418 (final), respectively.

It is important to neutralize the G418 to ~pH 7.4.

Both the neo and hyg genes have been used successfully as selectable markers in mutant ES cells (R.M., unpub. observ.). To select for hyg, use ES/LIF medium containing 1.0, 1.5, or 2.0 mg/ml hygromycin-B (HPH; UNIT 9.5).

3. Incubate cells 7 to 10 days. Change the medium each day, using ES/LIF medium with the appropriate concentration of G418 added, until single colonies are detected.

If cells overgrow plates and no single colonies are obtained, repeat steps 2 and 3 using a higher G418 concentration (e.g., 2, 4, or 6 mg/ml final).

4. Screen colonies for homologous recombinants as described in UNIT 23.5 (Basic Protocol, steps 15 to 21).

The hybridizing fragment indicative of the unaltered target gene will be completely absent in homozygous mutant cells. Typically, 50% of the clones are homozygous but because the process is random, individual heterozygous clones may yield homozygous clones at a frequency of 4% to 100%.

5. Perform a northern hybridization analysis (UNIT 4.9) of mRNA or an immunoblot analysis of protein (UNIT 10.8) to confirm that the target gene is inactive. No normal mRNA or protein should be found.

6. If the constructs contain loxP sites flanking the neo gene, remove the neo by transient expression of Cre (see UNIT 23.5, Support Protocol).

Both copies of neo are removed with same efficiency as one copy.

COMMENTARY

Background Information

Comparison of the phenotype of wild-type and homozygous mutant cells should provide information concerning the function of the targeted gene, provided that the gene functions in the cell assayed. ES cells offer a unique opportunity because they are capable of differentiating into any cell type. Many cell types can be produced through manipulation of in vitro cultures, including beating cardiocytes, skeletal muscle cells, neurons, glial cells, and vascular endothelial cells. Provided that the phenotype can be analyzed with a single cell or with a few cells, these differentiated cells can then be used to analyze the phenotype of the disrupted gene.

Because the normal number of chromosomes is maintained in ES cells (Mortensen et al., 1992), this method should be applicable to the study of gene function in development. ES

cells can also differentiate within the entire organism when injected into normal blastocysts. Homozygous mutant cells offer no technical advantage over heterozygous cells if the goal is to obtain a mutant mouse line through germline transmission; however, if homozygous mutant cells are tagged by introducing a gene that has a histochemically detectable product (e.g., β -galactosidase), then the fate of the homozygous mutant cells in the animal can be determined easily. Studies of embryos derived from blastocysts injected with tagged mutant ES cells may reveal the role of the target gene during development even if the mutant gene is lethal for the animal when homozygous.

In some cases, the phenotype of the disrupted gene may be tested by introducing heterologous proteins into the cultured cells. For example, an intracellular signaling pathway

may be investigated by expression of the particular receptor that creates an intracellular response to an applied agonist.

Cells produced using this protocol are only resistant to *neo* and its analogs. Thus, further genetic manipulations of the cells, such as the targeting of other genes using homologous recombination, can be performed using other drug selections (e.g., HPH; *UNIT 9.5*).

The suitability of this technique for the production of homozygous mutant cell lines other than ES cells is less certain. Homologous recombination and spontaneous loss of heterozygosity occur in some other cell lines and these same methods have been used to produce homozygous mutant cell lines; however, the methods may not be widely applicable, as other types of immortalized cells (e.g., lymphocytes or fibroblasts) may not undergo homologous recombination of a targeting construct at the same frequency as ES cells. Furthermore, many immortalized cells are markedly polyploid or aneuploid. If more than two copies of the target gene are present, this method for producing homozygous mutant cells may not be suitable; however, because these cells often lose or gain chromosomes, some clones might be isolated with no endogenous gene remaining.

The loss of heterozygosity seems to occur by several different mechanisms that vary with cell type. The mechanism by which this conversion occurs is unknown. Some possibilities include nondisjunction, chromosomal loss and duplication, or gene conversion.

Critical Parameters and Troubleshooting

Homozygous mutant ES cells will not be obtained if the mutation is lethal for the ES cell. The critical technical factor in this method is determining the amount of G418 needed to select homozygous over heterozygous cells. Differences in G418 levels required for different constructs can arise from two sources: the strength of the promoter and the influence of the genomic site on *neo* expression. In addition, the resistance to G418 may vary, depending on the enzymatic activity of a wild-type or mutant *neo* gene (Yenofsky et al., 1990). If the heterozygous mutant ES cells are resistant to the highest concentration of antibiotic (i.e., 2 mg/ml G418), the G418 concentration should be increased (e.g., up to 6 mg/ml). The endpoint should be survival of individual colonies with death of nearly all cells.

It should be noted that this selection method allows only a minority of the homozygous

clones to survive. In one case, homozygous mutant cells constituted 25% of the cells (presumably due to the chance conversion of a heterozygous to homozygous cell early in the life of the clone). Nevertheless, <1% of the cells survived 1 mg/ml G418 (R.M., unpub. observ.).

It is also possible that loss of heterozygosity may not occur for all genomic sites, although no such genomic sites have yet been identified. Because the inability to isolate homozygous cells may lead to the erroneous conclusion that a homozygous mutation in a gene is lethal, the lethality of the mutation should be verified. A rescue experiment in which a third copy of the gene (or expressed cDNA) is introduced into the heterozygous mutant cells, followed by selection of the homozygous mutant, would distinguish between lethality and no loss of heterozygosity.

An alternative method for producing homozygous mutant cells is to disrupt the second allele of the target gene using a second targeting construct that uses a different positive selectable marker. For example, if *neo* was used in the first-round construct, the *hyg* gene can be used and cells selected in HPH as described in *UNIT 23.5*.

Anticipated Results

The hybridizing fragment indicative of the altered gene will be absent in homozygous mutant cells. Typically, 50% of the clones are homozygous, but because the process is random, individual heterozygous clones may yield homozygous clones at a frequency of 4% to 100%.

Time Considerations

Expanding and selecting clones requires ~2 to 3 weeks. Screening clones takes several days.

Literature Cited

- Mortensen, R.M., Conner, D.A., Chao, S., Geisterfer, L.A., and Seidman, J.G. 1992. Production of homozygous mutant ES cells with a single targeting construct. *Mol. Cell. Biol.* 12:2391-2395.
- Yenofsky, R.L., Fine, M., and Pellow, J.W. 1990. A mutant neomycin phosphotransferase II gene reduces the resistance of transformants to antibiotic selection pressure. *Proc. Nat. Acad. Sci. U.S.A.* 87:3435-3439.

Contributed by Richard Mortensen
University of Michigan Medical School
Ann Arbor, Michigan

Chimeric Mouse Production by Microinjection

UNIT 23.7

The culmination of the creation of a mutation in mouse embryonic stem (ES) cells is, commonly, the generation of a mouse line that can propagate the mutation. The ability to combine methods of homologous recombination in ES cells with blastocyst-mediated transgenesis has resulted in an explosion of tailored mutant mouse strains. These animals provide research tools that are virtually impossible to create using other methodologies.

The most common method used to generate chimeras involves the injection of ES cells into the blastocoel cavity of 3.5-day-old embryos. The injected embryos are surgically implanted into the uterus of a foster mother. The resultant pups are true chimeras: their tissues are derived from both the host embryo and from the ES cells. If the ES cells are able to populate the germ line, the chimera can pass an altered gene to offspring, resulting in a new mouse strain in which all cells contain an altered gene.

The procedure is technically difficult, requiring proficiency with three methods: blastocyst isolation, blastocyst injection, and embryo reimplantation. Each of these techniques requires practice in addition to knowledge of the method; unlike many protocols, simply following the methodology is not sufficient. Because of the steep learning curve and the expense of the equipment, the procedure should not be undertaken without a serious commitment.

NOTE: All protocols using live animals must first be reviewed and approved by an Institutional Animal Care and Use Committee (IACUC) or must conform to governmental regulations regarding the care and use of laboratory animals.

NOTE: All mice used should be housed in a specific pathogen free (SPF) or virus antibody free (VAF) facility.

BLASTOCYST ISOLATION

The standard method for generating chimeras with ES cells uses 3.5-day-old mouse embryos (blastocysts). Embryos at this stage have a large fluid-filled cavity into which ES cells can be placed by injection. Blastocysts have already moved out of the oviducts and are found in the uterine horns. For injections, they must be harvested prior to hatching (loss of zona pellucida) and attachment to the uterine wall. C57BL/6 mice are used most commonly to generate blastocysts for ES cell injections. This strain yields reasonable numbers of embryos. It tends to produce high-grade chimeras that can be easily distinguished by coat color when used with ES cells derived from a variety of 129 sub-strains.

Materials

- C57BL/6 female mice (3 to 4 weeks old; e.g., Taconic)
- 50 U/ml pregnant mare's serum (PMS; see recipe)
- 50 U/ml human chorionic gonadotropin (HCG; see recipe)
- C57BL/6 stud males (8 weeks to 6 months old; e.g., Taconic)
- 95% ethanol
- Injection medium (see recipe)

Tuberculin syringe (1 ml, 26-G, $\frac{3}{8}$ in.)

Surgical equipment (thoroughly washed and sterilized by autoclaving or with ethanol; individual investigators may prefer slightly different instruments):

The following are available from Biomedical Research Instruments (similar instruments are available from several suppliers):

Scissors, large (4.5 in.)

**BASIC
PROTOCOL 1**

**Manipulating the
Mouse Genome**

23.7.1

Contributed by David A. Conner

Current Protocols in Molecular Biology (2001) 23.7.1-23.7.20

Copyright © 2001 by John Wiley & Sons, Inc.

Supplement 53

Scissors, small (3.5 in.)
 Scissors, iris (3 in.)
 Forceps, small (4 in.), curved
 Forceps, large (4 in.), curved
 Forceps, toothed (3.5 in.)
 Blunted 25-G needle (see recipe) with 3, 5, or 10 ml syringe
 Watch glass (Corning, 2.5 in., washed and sterilized with ethanol prior to use)
 Dissecting microscope and fiber optic light source (stereo microscope, 0.8× to 4×
 zoom, with stand that allows illumination from above and below)
 Embryo transfer pipet (see recipe)
 Microdrops of injection medium under mineral oil (see recipe for microdrop
 cultures) in 35 × 10–mm petri dishes (Falcon)
 Light white mineral oil (Sigma)

Initiate superovulation

1. One week prior to the day of ES cell injection (see the time line, Fig. 23.7.1), inject 3 to 4 week old female C57BL/6 mice intraperitoneally (using a 1-ml tuberculin syringe with a $\frac{3}{8}$ -in., 26-G needle) with 0.1 ml of 50 U/ml PMS (5 U) between 11 a.m. and 2 p.m. to induce superovulation.

Blastocyst donors are superovulated to synchronize estrus and produce the greatest embryo yields. Mice can be mated naturally or after checking for estrus, but this requires housing

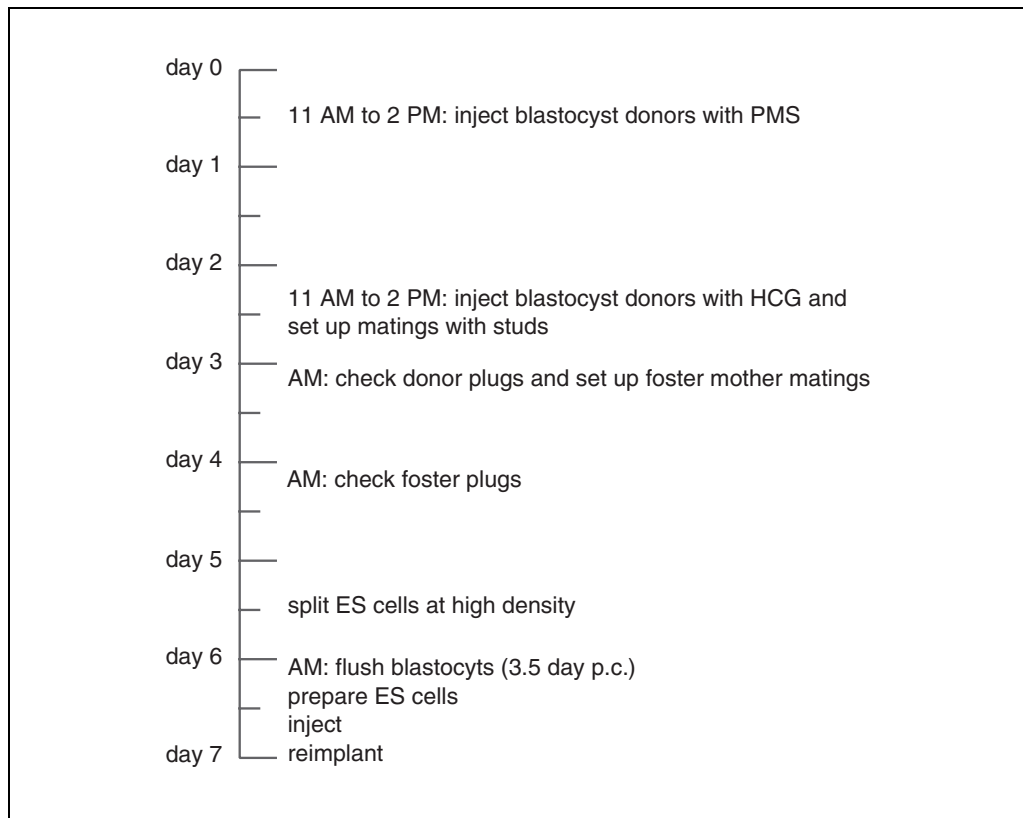


Figure 23.7.1 Time line of blastocyst injections. The major steps for a single day of injection are shown in the time line. The procedure begins 6 days before the actual day of injection with the initiation of superovulation of the donor mice. Mating is considered to occur in the middle of the dark cycle. Donors mate at the transition of the third and fourth 24-hr period. By the middle of the injection day, they are post-coital day 3.5. Fosters mate 24 hr later and are therefore post-coital day 2.5 at noon on the day of injection. ES cells should be thawed early enough in the procedure to produce high-density cultures for the day of injection—between day 0 and day 5, depending on cell density. The exact day of thawing will depend on the density at which the cells were frozen.

many more mice and will not produce more embryos. Young female mice that have not yet cycled will give the greatest embryo yields, whereas adult female C57BL/6 and even 5- to 6-week old females have more variable embryo yields. The timing of injection is important; it must be coordinated with the light/dark cycle to ensure that the females ovulate in response to the exogenous hormones prior to mating. A typical lighting cycle consists of 10 to 12 dark hours beginning at 6 to 7 p.m. and extending to 5 to 6 a.m. This represents the range of values that works well with a standard workday. For a specific animal room the timing and duration of the cycle should be held constant.

10 to 12 matings should generate enough blastocysts for 1 day of injection.

2. At 46 to 48 hr (again between 11 AM and 2 PM) after the PMS injection, inject each female intraperitoneally with 5 U of 0.1 ml of 50 U/ml HCG (5 U).

HCG replaces endogenous LH. HCG must be administered prior to the release of the endogenous hormone, which is controlled both by the timing of PMS injection and the light/dark cycle.

3. Place single superovulated females with individually caged C57BL/6 stud males.

Mice should mate in the middle of the next dark cycle. If injections are planned on consecutive days, it is a good idea to maintain a large enough colony of stud males so that no male is mated on consecutive days, otherwise embryo yields will usually drop on the second day.

4. The next morning, separate females and check for plugs.

Females plugged the next morning are considered to be 0.5-day post-coitum. Blastocysts are 3.5 days old. Most, but not all of the superovulated females should plug. Blastocyst yields drop with studs older than ~8 months. In the author's laboratory, the males are typically replaced when they are between 6 and 8 months old.

Mice typically mate in the middle of the dark cycle. A hard white or yellow mucous plug forms in the vagina after mating. The plug usually remains for 12 hr or longer. Females should be checked in the morning to avoid missing plugs. Check for plugs by picking up each female by the tail. Place her on a wire cage top and allow her to grasp the wires with her front paws. The vaginal opening can be observed by raising the hind legs off the cage top while lifting the tail. Plugs are often immediately visible. To ensure that a plug is not missed, the vagina can be probed gently with a disposable pipet tip (Gilson-style P-200). These tips are cheap, clean, and relatively blunt.

Isolate the blastocysts

5. On post-coital day 3.5 (see time line, Fig. 23.7.1), sacrifice females by cervical dislocation. Place the mouse on its back and wet the abdomen with 95% ethanol.

Work through the protocol with one mouse at a time until comfortable with the procedure. Blastocysts should be harvested before noon, when the mice are housed with a standard light/dark cycle. Later in the day, a higher percentage of the blastocysts will have hatched (lost the zona pellucida). Blastocysts without a zona cannot be used for injections.

6. Tear the skin away from the abdomen and cut through the peritoneal wall with scissors.

The skin can be cut, but with some practice the skin can be torn away quickly.

7. Remove the ovaries and uterine horns. Place them on a clean dry paper towel (see Fig. 23.7.2A for a diagram of the isolated tissues).

The ovaries and both uterine horns can be removed as a single piece. First, grasp an ovary and its associated fat pad with forceps. The ovary is easily found at the base of the kidney. Gently lift the ovary and uterine horn. As the uterine horn is lifted, cut away associated tissue and fat. Cut the cervix and continue to free the second uterine horn and ovary. The tissue will stick to the paper towel, making it less difficult to work with.

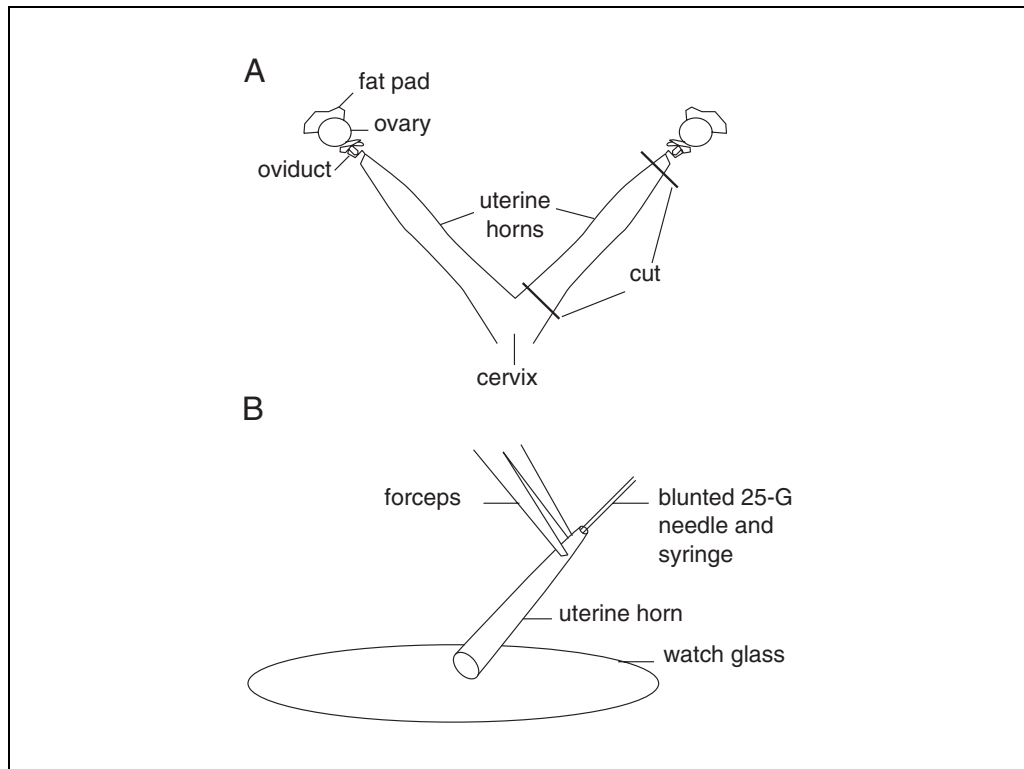


Figure 23.7.2 Blastocyst isolation. Diagram of the uterine horns during different stages of blastocyst isolation. **(A)** Appearance of the attached uterine horns immediately after removal from the mouse. At this stage fat and associated tissues can be removed easily if the uterus is placed on a clean paper towel. Note the indicated cut sites. Both horns should be cut in same locations. **(B)** Illustration of the flushing procedure. Insert the blunted 25-G needle into the opening while the uterus is still on the paper towel. Clamp down on the needle with forceps to pinch the tissue on the needle and move the assembly over a watch glass to begin flushing.

8. Trim fat and blood vessels from the uterine horns with fine scissors and forceps.

This step is not critical, but it is particularly helpful to remove debris from around the junction between the uterus and the oviduct and the uterine/cervical junction.

9. Cut the uterine horn at the cervical junction and then at the junction of the uterus and oviduct (Fig. 23.7.2A). Insert a blunted 25-G needle attached to a syringe filled with injection medium into the uterine horn (Fig. 23.7.2B). Grasp the tip of the needle and tissue with forceps. Lift the assembly over a watch glass and flush the blastocysts out of the uterine horn with 200 to 300 μ l of injection medium. Repeat for the remaining uterine horn.

The oviduct end should be cut at a point where it is wide enough to accept the needle. Blastocysts can be flushed into a tissue culture dish instead of a watch glass. The watchglass is convenient because the embryos tend to concentrate in the center due to the curvature. The volume of fluid is not critical. Observe the procedure closely to make sure the fluid is flushing the uterus; the uterine horn should expand slightly as the syringe is pushed and the fluid should come out of the other end, not dribble down the side. Blastocysts from many uterine horns can be collected in a single watch glass. The embryos can be left at room temperature during the procedure.

10. Collect blastocysts with an intact zona pellucida (Fig. 23.7.4B and Fig. 23.7.5) under a dissecting microscope using an embryo transfer pipet. Use illumination from both above and below for best visibility. Transfer the embryos to a microdrop of injection medium under oil in a petri dish and place the dish on a slide warmer set at 37°C (see Reagents and Solutions for description of microdrop cultures).

The blastocysts will remain healthy for the entire day under these conditions.

Invariably, the watchglass contains debris—air bubbles and embryos that have not developed properly. Embryos will settle to the bottom of the watch glass. Gently swirling the glass will help to concentrate the embryos in the center. Often, rocking the watch glass while focusing on the lower most layer of debris will help to distinguish embryos from debris and bubbles; embryos roll in a distinctive manner when the glass is rocked. Blastocysts are fairly easy to recognize because of the large cavity. At times it is difficult to determine if the blastocysts have a zona pellucida when viewed with a dissecting microscope. If there is any question, save the embryo; the zona pellucida is easily observed with the injection microscope. Blastocysts without zonae cannot be injected. Many poorly developed embryos are usually present. These appear as single large cells or clumps of cells enclosed in a zona pellucida and represent unfertilized embryos or embryos that were slow to develop. These cannot be injected and should be discarded. The final decision on any close calls can be made under the injection scope.

BLASTOCYST INJECTION

Conceptually, blastocyst injection is simple—blastocysts and single-cell suspensions of ES cells are placed in a microdrop under a microscope. Individual ES cells are collected by aspiration into capillary pipet. A single blastocyst is grasped by aspiration with a second pipet. The injection pipet is inserted into the blastocoel cavity and the cells are ejected into the cavity.

Materials

- Injection medium (see recipe)
- ES cell suspension (see Support Protocol 2)
- Blastocysts (see Basic Protocol 1)
- Injection chamber (Fig. 23.7.3)
- Injection apparatus (see recipe)
- Holding pipet (VacuTips, Eppendorf)
- Injection pipet (Transfer Tips, Eppendorf)

Prepare the blastocysts and injection chamber

1. Prepare an injection chamber, consisting of two 50- μ l microdrops of injection medium separated by a few millimeters and covered with mineral oil (Fig. 23.7.3A and B) in the inverted top of a Falcon 35 \times 10-mm petri dish.

The inverted top of a petri dish is used because the lip is shorter than the lip of the dish. It is less difficult to align the injection and holding pipets with the shorter lip. Not all 35-mm dishes have the short lip found on the top of the Falcon plate. Alternatively, a larger dish can be used, but this requires the addition of much more mineral oil. Use petri dishes rather than tissue culture plates to reduce cell sticking.

2. Place ES cells in the upper drop (Fig. 23.7.3A) by adding 5 or 10 μ l of the ES cell suspension to the drop.

The exact number of cells in the drop is not critical. However, too many cells will often lead to clumping and accumulation of debris on the tip of the injection pipet while too few cells will require too much searching. 5 μ l of a 1-5 \times 10⁶ cells/ml suspension is usually sufficient. With some practice there is no need to count cells. The number of cells on the plate can be adjusted by changing the volume added to the microdrop.

3. Place 10 to 15 blastocysts in the 12 o'clock position in the lower drop (Fig. 23.7.3A).

Transfer the number of embryos that can be injected in 30 to 60 min. Two drops are created to separate cell picking and injection. If the holding pipet is used in the presence of many ES cells it can get clogged with debris. By physically separating the loading and injection steps ES cells are never accidentally aspirated into the holding pipet.

BASIC PROTOCOL 2

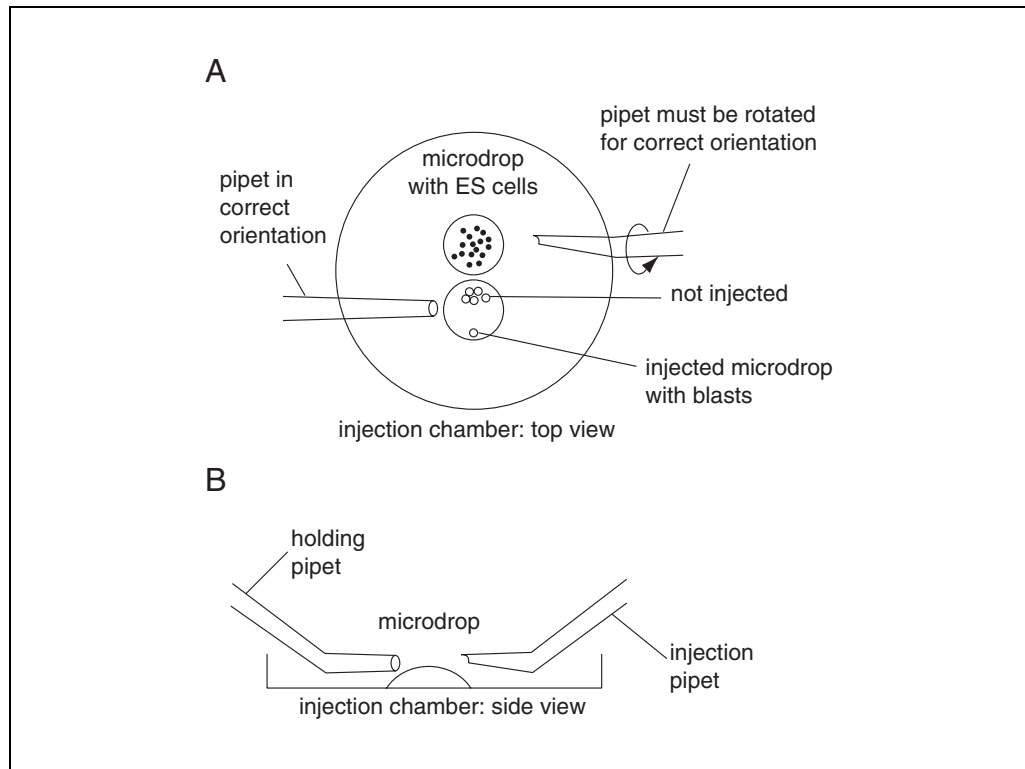


Figure 23.7.3 Injection chamber and pipet alignment. Diagram of the injection chamber with correct pipet orientation as viewed from the top and side. Note that the relative sizes are not drawn to scale. **(A)** Top view of the injection chamber. Two microdrops are illustrated with the ES cell drop at the top and the blastocyst drop below. Injected blastocysts should be moved to another part of the well to avoid confusion. Both pipet tips should appear straight when viewed from above, when they are in the correct orientation. The injection pipet on the right is not aligned properly and must be rotated in its holder in the direction indicated until it appears straight. **(B)** Side view of the injection chamber. The pipets are angled so that the bent tips are parallel to the surface of the plate.

Fine-tune the injection chamber manipulator controls

4. Fine tune as follows.

- a. Place the injection chamber on the microscope and focus under low power (4× objective) on the blastocysts.
- b. With oil-filled pipets, be sure to flush the line until there are no air bubbles. Insert the holding pipet into the pipet holder and push oil through the pipet until oil comes from the tip.
With air filled systems, the pipet can be inserted directly into the holder.
- c. Invert the holder and coarsely position the pipet tip over the microdrops. Adjust the angle of the pipet so that the tip is parallel to the surface of the plate (Fig. 23.7.3B).
- d. Lower the pipet tip to the top of the oil using the coarse manipulator controls. Bring the tip into the microscope field of view. Focus up on the pipet and then rotate the tip until it is in position.
- e. Repeat the procedure with the injection tip.

Make sure to adjust the orientation of the tips above the surface of the plate. Otherwise the tip can be broken on the surface of the injection chamber. A correctly oriented tip will appear straight when viewed from above. If the bend is visible, the pipet must be rotated in its holder until it appears straight (Fig. 23.7.3A).

- Center the tips of the two pipets in the field of view and then adjust the focus so the blastocysts are in view. Bring the pipets down slowly using the course manipulator controls until they are in focus (now the tips are in the same plane and close to the bottom of the injection chamber). Slowly back-fill the pipet with medium.

Only the tips of the pipets need to be back-filled with medium. As a rule of thumb, stop back-filling when the medium interface reaches the edge of the field of view. With oil-filled systems the tips are inverted under slight positive pressure: oil will be coming out the tips in the microdrop. The oil bubble can be removed from the tip by moving the tip between the medium and oil interface while slowly backing out the micrometer control. Immediately after back-filling there is often some movement of medium in the tip because of temperature differences of the various components. The movement should stop if the system is allowed to equilibrate for a few minutes. In oil-filled systems, air bubbles anywhere in the system can produce uncontrolled movement that will not abate with time; the bubbles must be flushed from the system.

- Change to the 10× objective and recenter the tips of the injection pipet and holding pipet. Make sure the blastocysts are also in focus so that all components are in the same plane. Repeat the procedure after switching to the 20× objective.
- Lift the tips above the surface of the plate using the fine manipulator controls and move the stage so the tips are in the microdrop containing ES cells. Focus on the ES cells and lower the injection tip until it is in the same plane of focus as the cells. Leave the holding pipet above the surface.

Whenever the stage is moved, the pipet tips should be raised to avoid breakage. The tips can be raised so they are just out of focus and appear as a hazy outline. Also make sure the tips are raised when they are near the blastocysts and the stage is moved. Otherwise the blastocysts can be inadvertently swept into the oil phase and destroyed. Note that “sweeping” with the pipet can be used intentionally to reposition groups of blastocysts in the microdrop. With practice, adjusting the vertical position of the tips will become automatic.

- Move the tip of the injection pipet to a healthy-looking ES cell (Fig. 23.7.4A) and gently load the cell by backing out the plunger of the injection pipet. Repeat the procedure until 15 to 18 cells have been collected in the injection tip. Raise the tip slightly and move the stage until the pipets are in the microdrop containing the blastocysts.

This step, although fairly easy to master, is the most time-consuming part of the procedure. Healthy looking ES cells should be round and have sharp refractal borders. Move around the microdrop by moving the stage. Do not pick up cells that are so big that they will be compressed as they enter the pipet; they may lyse. Feeder cells are larger than ES cells and will not appear round and refractal with sharp edges. Pairs of ES cells can be picked up together. The cells do not have to be packed tightly together in the pipet. If the ES cells touch the oil interface they will lyse. Do not let the oil interface pass over a cell in the pipet; the lysed debris will stick to the wall of the pipet and clog the tip often requiring pipet replacement. Do not move the tip into the oil phase when there are cells stuck near the outer mouth of the pipet; this also may result in tip clogging.

Inject the blastocysts

- Orient the blastocyst as follows.
 - Choose a blastocyst to inject and focus on it.

Figure 23.7.4, panels C to E, depict sequential steps of the injection procedure.

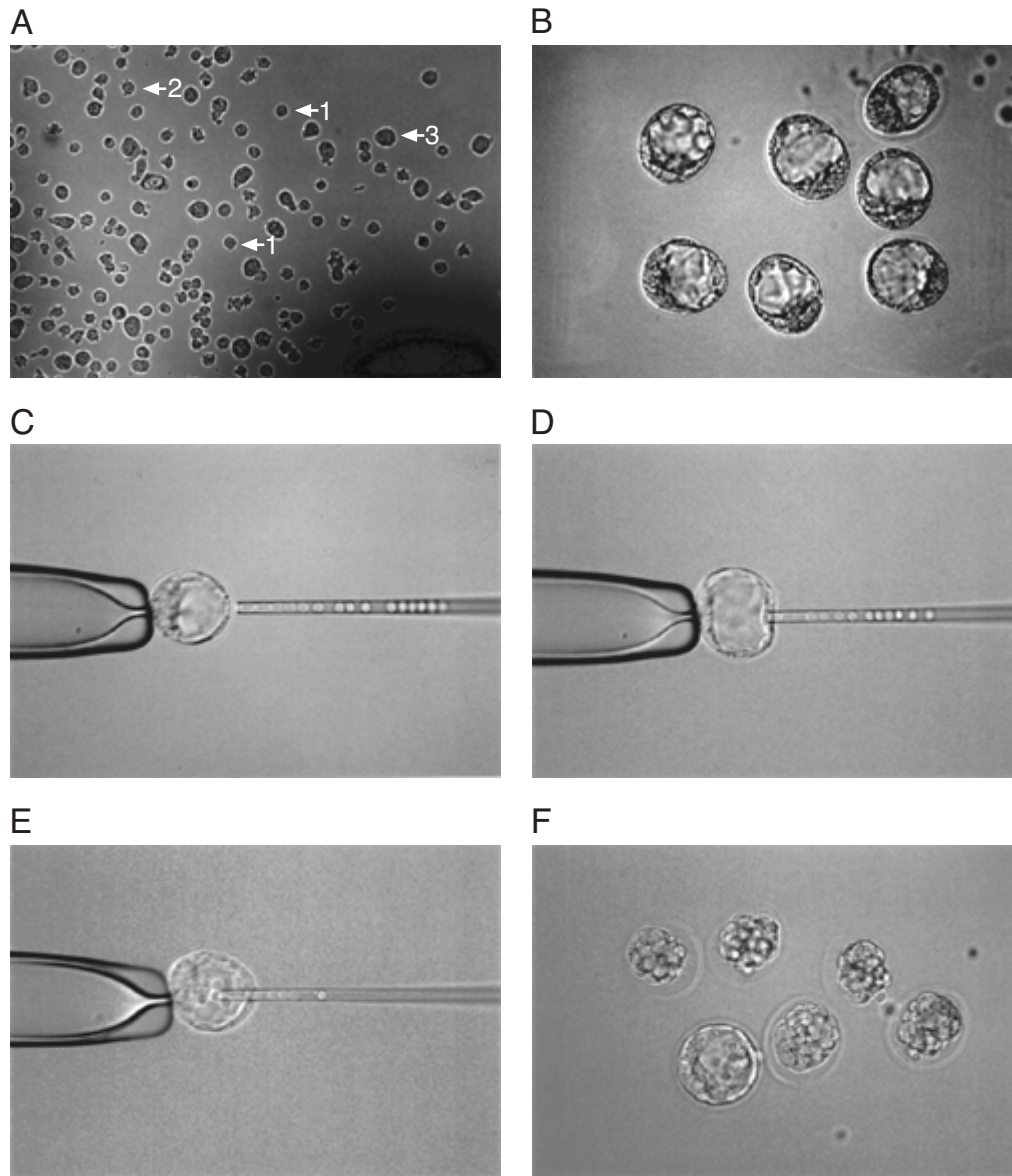


Figure 23.7.4 Blastocyst injections. Phase-contrast and Nomarski differential interference contrast micrographs of the injection procedure. When an injection pipet with a beveled tip is used, the inner cell mass should be oriented at 6 or 12 o'clock to avoid damage. **(A)** Single-cell suspension of ES cells. Pick up only those cells that are round and have sharp refractal borders (1). Avoid cells with rough borders and those beginning to "bleb" (2). The large cells (3) are fibroblasts and should be avoided. Pick cells that will fit into the injection tip without compression. **(B)** Blastocysts ready for injection. Note the large blastocoel cavities and the presence of the outer shell (zona pellucida). Blastocysts without zonae cannot be injected. **(C)** Blastocyst and pipets in correct orientation for injection. Note the tips of both pipets and the outer edge of the blastocyst are in the same focal plane. **(D)** Compressed blastocyst prior to penetration of the injection tip. Rolling of the blastocyst will be observed at this stage if the tips are not properly aligned. When using beveled injection tips (as compared to the blunt tip in the photograph) the blastocysts cannot be compressed to the same degree without penetration. **(E)** Ejection of ES cells into the blastocoel cavity. **(F)** Blastocysts at different stages of collapse and re-expansion. The blastocoel cavity will collapse almost immediately after injection and appear like a ball of cells as seen in the upper part of the photograph. Over time, most embryos will re-expand and display a characteristic blastocoel cavity.

- b. Rotate the blastocyst so the inner cell mass is at the 6 or 12 o'clock position and a joint in the trophoblast layer is adjacent to the injection tip (Fig. 23.7.5 and Fig. 23.7.4).

Note that Figure 23.7.4 illustrates the process using a blunt injection tip. In this case the tip is less likely to damage the inner cell mass, so that the embryo can be oriented with the inner cell mass at the 9 o'clock position. Tips purchased from Eppendorf are beveled, and therefore the blastocysts should be oriented with the inner cell mass in the 6 or 12 o'clock position.

Blastocysts must have a zona pellucida to be injectable and a cavity large enough to hold the entire beveled end of the injection tip. Rotate the blastocyst by "bating" it around with the injection pipet and the holding pipet. This requires a concerted 3-dimensional movement of the tip, which will take some practice. The blastocysts are oriented so that the injection tip will not damage the inner cell mass and so that the tip will penetrate a thin portion of the trophoblast layer.

10. Prepare the blastocyst for injection as follows.

- a. With the blastocyst in the correct orientation, grab the embryo with the holding pipet by applying gentle suction (Fig. 23.7.4C).

The blastocyst must be held fairly tightly. However, it is possible to apply too much suction. This will result in obvious damage to the embryo.

- b. Raise the blastocyst above the surface of the injection chamber and then lower it until the blastocyst just touches the surface.

This helps to place the holding pipet and blastocyst in the correct plane.

When the blastocyst is brought down onto the surface of the plate it will begin to roll. The beginning of the rolling motion indicates that the embryo has just contacted the surface of the plate.

- c. Focus on the outermost edge of the blastocyst. Move the injection tip until it is in focus at the 3 o'clock position (Fig. 23.7.4C and Fig. 23.7.5).

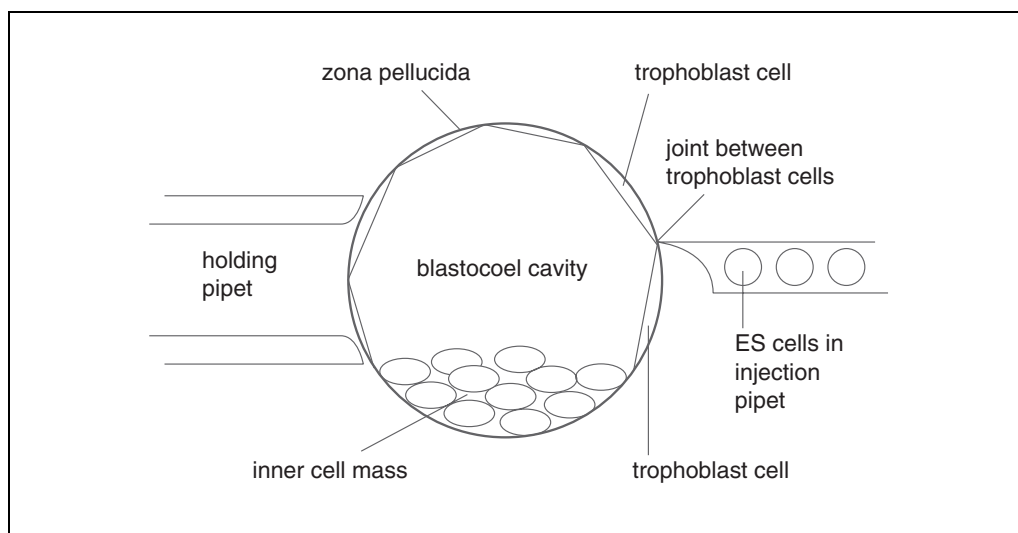


Figure 23.7.5 Blastocyst alignment for injections. Diagram of the correct blastocyst alignment prior to injection. The inner cell mass could also be oriented at the top of the field of view. The embryo is aligned to avoid damage of the inner cell mass. Note that the injection tip is aligned in opposition to a joint in the trophoblast layer. Attempts to inject blastocysts through a thick part of the trophoblast layer are often unsuccessful. The tip may not penetrate fully and the blastocyst can collapse before cells are introduced into the cavity.

The bottom of the blastocyst should be resting on the surface of the injection chamber and the tips of the injection pipet and holding pipet should be in the same plane at the equatorial edge of the blastocyst.

- d. Gently press the tip of the injection pipet into the edge of the embryo.

Before penetration, the embryo will compress slightly (Fig. 23.7.4D). If the holding pipet and injection tip are not at the ideal equatorial positions or not directly opposed, the embryo will begin to roll. Injecting at this point will usually push the embryo off the holding pipet. Instead readjust the heights of the pipet so the embryo is compressed between the two tips. Then push the injection tip into the blastocoel cavity in a quick fluid motion.

This step tends to be the most frustrating part of the whole procedure. The pipets must be in the same plane so that they trap the blastocyst between themselves and the surface of the injection plate. In addition, each new injection tip may behave slightly differently. Some will penetrate into the cavity easily even with a slow motion. Other tips will require a rapid jab.

11. With the injection tip inside the cavity, gently expel the cells (Fig. 23.7.4E). After all the cells are expelled, remove the injection pipet from the blastocyst slowly.

Positive pressure is built up inside the cavity as fluid and cells are expelled. If the tip is removed too quickly, many of the cells will spill out. If the tip is removed slowly, fluid can escape around the outside of the pipet and equalize the pressure. Often a single cell trails the pipet out. If the bevel of the tip is removed very slowly, the opening will close and prevent the escape of the last cell. However, the loss of a few cells is not critical.

12. Raise the injected blastocyst above the surface of the plate and move to the 6 o'clock position of the microdrop. Drop the blastocyst and move back to the drop containing cells to repeat the whole procedure.

The blastocysts will collapse immediately after injection (Fig. 23.7.4F) and then re-expand over the next few hours. After all the blastocysts on the plate have been injected, transfer them to a microdrop on the slide warmer. If additional blastocysts will be injected, prepare a new injection chamber or simply add fresh blastocysts to the existing microdrop. The entire procedure is tedious. On average, even an experienced injector may only inject one embryo every 3 to 5 min.

BASIC PROTOCOL 3

UTERINE TRANSFERS

Injected blastocysts must be surgically reimplanted into foster mothers to allow embryos to develop to term. Blastocysts are reimplanted into the uterine horns of 2.5-day post-coital pseudopregnant females. Approximately 17 days after reimplantation, pups will be born.

Materials

Embryos (injected blastocysts; see Basic Protocol 2)
Injection medium (see recipe)
2.5-day pseudopregnant female mice (see Support Protocol 1)
2.5% Avertin (see recipe)
95% ethanol

Embryo transfer pipets (see recipe)

Modeling clay

Surgical equipment (thoroughly washed and sterilized by autoclaving or with ethanol; individual investigators may prefer slightly different instruments)

The following are available from Biomedical Research Instruments (similar instruments are available from several suppliers):

Scissors, large (4.5 in.)

Scissors, small (3.5 in.)
Scissors, iris (3 in.)
Forceps, small (4 in.), curved
Forceps, large (4 in.), curved
Forceps, toothed (3.5 in.)
Serrefine clamps
The following are available from Fisher (similar instruments are available from several suppliers)
Wound clip applicator
Wound clips (surgical staples)
Dissecting microscope and fiber optic light source (stereo microscope, 0.8× to 4× zoom, with stand that allows illumination from above and below)
25-G needle
Suture, 5-0 Dexon II on T-31 needle (e.g., Kendall Health Care)

Prepare the embryo

1. Just prior to reimplantation, transfer embryos to a clean drop of injection medium without oil.

This helps to minimize the transfer of oil into the uterus. Reimplant all injected embryos unless the embryo is clearly destroyed. Many will have re-expanded and have a blastocoel cavity. However, embryos that have not re-expanded are not necessarily dead. Reimplantation of some embryos that will die will not affect the development of healthy embryos.

2. Load 6 to 10 embryos into each embryo transfer pipet (Fig. 23.7.6A). Place the loaded transfer pipets on the benchtop supported by a piece of modeling clay.

Air bubbles are introduced into the transfer pipet by gentle aspiration while the pipet tip is held out of medium. The primary purpose of the bubbles is to serve as a visual aid during reimplantation so it will be obvious when the medium containing the embryos has been ejected into the uterus. The medium “spacer” should be at least the same length (volume) as that taken up by the embryos and bubbles.

Reimplant the embryos

3. Anesthetize a 2.5-day pseudopregnant female by injecting intraperitoneally with 0.017 ml/g 2.5% Avertin.
4. Wet the skin on the back with 95% ethanol.
5. Using scissors, make an incision ~1 cm long in the skin on the midline of the back (Fig. 23.7.6B). Place the mouse under a dissecting microscope with overhead illumination. Slide the opening down the right or left side of the mouse until the ovaries or ovarian fat pad are visible through the peritoneal wall.

A larger incision (1.5 to 2 cm) may be helpful to those new to the procedure. In young females the ovaries (pink) and ovarian fat pad (white) are visible through the peritoneal wall. As the females age the ovaries are often obscured by fat. The ovarian fat pad is whiter than the obscuring fat, which is slightly yellow.

6. Make a small incision (~0.5 cm) in the peritoneal wall over the ovaries. Grasp the ovarian fat pad with blunt forceps and pull the ovaries and proximal end of the uterine horn out of the incision. Attach a small serrefine clamp to the ovarian fat pad and rest the clamp on the side of the animal to keep hands free (Fig. 23.7.6C).

The peritoneal incision can be made easily with fine iris scissors. Grasp the peritoneal wall with toothed forceps to make the cut. Avoid blood vessels. As indicated above, the ovarian fat pad can be distinguished from other fat in the body cavity by its distinct color.

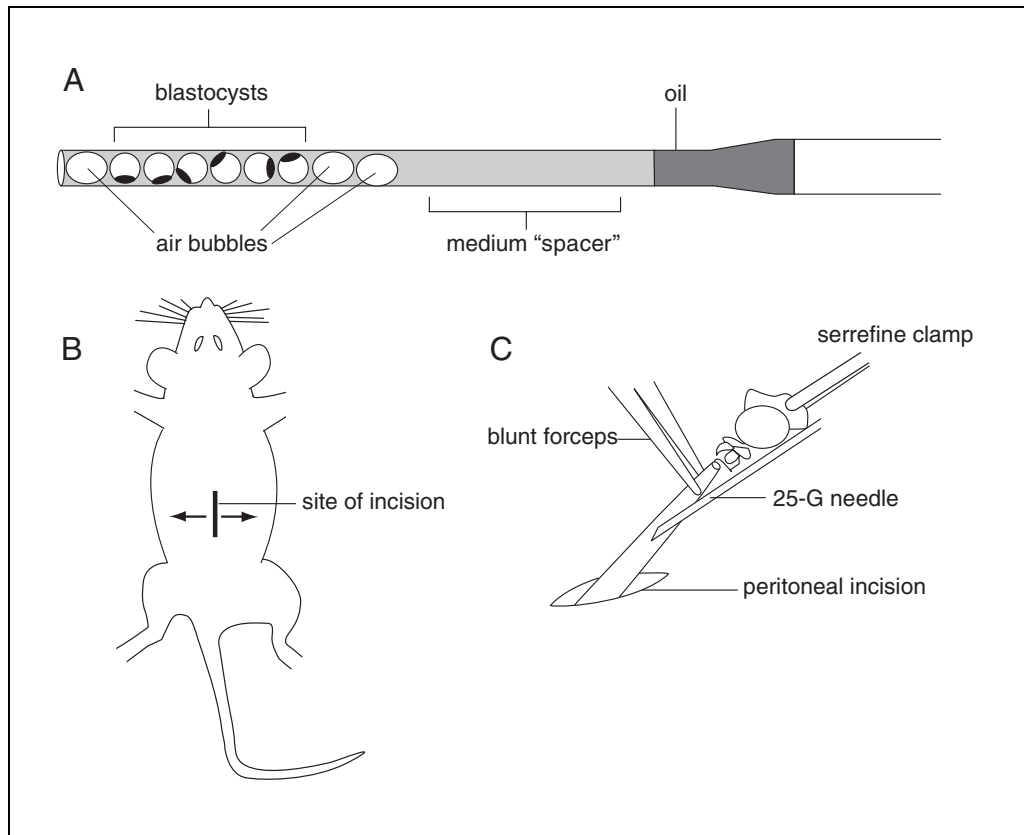


Figure 23.7.6 Reimplantation. Diagrams of various aspects of the reimplantation procedure. **(A)** A transfer pipet loaded for reimplantation. Air bubbles surround the blastocyst and act as markers that can be seen during surgery to ensure that the embryos are expelled into the uterus. Blastocysts do not have to be as tightly packed as illustrated. **(B)** Proper location of the skin incision for access to both uterine horns from a single site. The incision is on the midline of the back at the level of the last rib. The incision site can be slid to either side to gain access to either horn. Skin incisions can be made directly over the ovaries if the investigator prefers. Two incisions would be necessary to reimplant embryos in both horns. **(C)** Isolated uterine horn ready for puncture with a 25-G needle. The uterus is secured by a clamp attached to the ovarian fat pad. Blunt forceps are used to grasp the uterus near the oviduct junction as the tissue is punctured and the transfer pipet is inserted. The uterus should be held gently, to avoid damage.

7. Grasp the uterus near the oviduct junction with blunt forceps. Puncture the uterine wall with a 25-G needle (Fig. 23.7.6C). Make sure the needle enters the lumen.

The needle must enter the lumen of the uterus. Watch the surface of the uterus when the needle is removed. There is usually a landmark so the opening can be found with the transfer pipet. Typically the surface puckers slightly around the penetration site or there is a little bit of bleeding. Alternatively a preloaded pipet can be placed in the mouth pipettor prior to puncturing the uterus and held by modeling clay on the bench. After the hole is introduced in the uterus, it is possible to pick up the pipet while keeping an eye on the puncture site.

8. Insert the embryo transfer pipet into the opening and expel the embryos and air bubbles into the lumen.

Use the bubbles in the transfer pipet as a marker to ensure that the blastocysts are expelled into the lumen. The bubbles also serve as an indicator of how much of the transfer pipet has been inserted into the lumen. The tip should be inserted ~5 mm. Avoid introducing oil into the lumen. The medium/oil interface is usually visible.

9. Gently reposition the uterus and ovary in the peritoneal cavity. Use one suture to close the peritoneal wall.

10. Repeat the procedure on the other side.

Embryos do not have to be reimplanted on both sides.

11. Close the incision in the skin with surgical staples.

Staples are used rather than sutures because mice will often reopen an exterior wound that is closed with sutures. Two staples are sufficient for a 1-cm incision. Three may be necessary for a larger incision.

12. Place the mouse on a slide warmer until she begins to show signs of recovery (10 to 30 min). Place her in a clean cage.

Multiple recipients may be caged together. There are advantages and disadvantages to separating the foster mothers before the pups are born. With multiple births in the same cage, it is difficult to keep track of the successfulness of individual surgeries and the cages may become overcrowded. Alternatively, some new mothers, particularly those with small litters, often take better care of their litters when there are other new mothers in the cage. Swiss Webster foster mothers are generally very good mothers under all conditions.

PREPARATION OF PSEUDOPREGNANT FOSTER MOTHERS

Injected blastocysts must be reimplanted into foster mothers to develop to term. Foster mothers are produced by mating females with vasectomized males. After mating, the pseudopregnant female displays the hormone profile of a pregnant mouse for several days. Transfer of embryos into the oviduct or uterus at the appropriate time provides the final component necessary for the foster mother to carry a litter to term.

A variety of mouse strains can be used to produce foster mothers. The most important features of the strain are their maternal characteristics. Foster mothers that will care for litters of varied sizes and are not influenced dramatically by handling or noise are best. In general, outbred strains or F1 hybrids (offspring derived from mating 2 inbred strains) are most useful. In the author's laboratory, outbred Swiss Webster (Taconic) mice are routinely used for both foster mothers and sterile studs. Vasectomized males can be purchased directly from the vendor, obviating the need to perform the surgery in the laboratory. Vasectomized males should be test-mated to verify their sterility.

Materials

Vasectomized male mice (8 weeks to 16 months old; Taconic)

Female mice (8 weeks to 6 months old; Taconic)

1. Place female mice with individually caged vasectomized males on the fourth day of the injection time line (Fig. 23.7.1).

Two females can be mated with each male. With 2 females per male and 18 vasectomized male (36 matings), these random matings will almost always yield 2 or more pseudopregnant foster mothers each night.

2. Check for evidence of mating (plugs) the next day. Remove plugged females.

Mice typically mate in the middle of the dark cycle. A hard white or yellow mucous plug forms in the vagina after mating. The plug usually remains for 12 hr or longer. Females should be checked in the morning to avoid missing plugs. Check for plugs by picking up each female by the tail. Place her on a wire cage top and allow her to grasp the wires with her front paws. The vaginal opening can be observed by raising the hind legs off the cage top while lifting the tail. Plugs are often immediately visible. To ensure that a plug is not missed, the vagina can be probed gently with a disposable pipet tip (Gilson-style P-200). These tips are cheap, clean, and relatively blunt.

Plugged females should be separated and the date noted. The day of plugging is considered day 0.5 p.c. (post coitus). Uterine transfers are performed on 2.5 p.c. pseudopregnant females. If fosters will be required on consecutive days, the unplugged females can be left with the vasectomized studs and checked on subsequent days. Otherwise, all females should be separated. These random matings typically produce more plugged females on the second and third days than on the first day. By replacing plugged females with new females, 18 to 24 stud males can be used to produce 2 or more fosters on each of 5 consecutive days. This will provide enough fosters for a single injector for a week of injections. Mice are not always obliging; there will be some rare days when fosters are not available.

ES CELL PREPARATION

Single-cell suspensions of ES cells must be prepared for injections. The cells must be completely dispersed, but not treated so vigorously as to cause significant cell death.

Materials

Injection medium (see recipe)

Hanks' balanced salt solution (HBSS; calcium- and magnesium-free; UNIT 23.2)

Additional reagents and equipment for ES cell culture (UNIT 23.3)

1. Thaw ES cells several days before injection. Grow on feeders as described in UNIT 23.3.

Although very few cells are needed for injections, it is often convenient to use 25-cm² flasks. Split the cells at high density (1:2 or 1:3) the day before injections. This is done so the ES cells will greatly outnumber fibroblasts, and because it is easier to make single-cell suspensions from cells that have been passaged recently.

2. On the day of injection, wash the cells with calcium- and magnesium-free HBSS (5 ml for a 25-cm² flask). Remove the HBSS and add 1.5 ml DPBS/EDTA. Incubate at 37°C for 3 min, then add 1.5 ml prewarmed 0.05% trypsin/EDTA, mix, and return to the 37°C incubator.
3. After 1 to 2 min, check to see if the trypsinization is complete by tapping the flask. If the cells come off the plate the trypsinization is complete, otherwise return the flask to the incubator and check again after 1 min.

The idea is to generate a single-cell suspension with minimal damage to the cells. Preincubation with PBS-EDTA greatly reduces the time in trypsin.

4. When trypsinization is complete, add 7 ml of injection medium to the flask. Disperse the cells by pipetting the cell suspension up and down ~10 times in a 10-ml pipet, taking care to avoid introducing air bubbles. Transfer the cell suspension to a 15-ml conical tube and pellet the cells by centrifugation as described for routine ES cell culture (UNIT 23.3).
5. Resuspend the pellet in 10 ml of injection medium and centrifuge again, as before. Remove the supernatant.
6. Resuspend the cell pellet in injection medium. Keep the tube at room temperature or lower to prevent the cells from clumping.

Placing the tube on its side prevents the cells from concentrating in the conical tip of the tube and reduces the extent of clumping. A cell density of $1-3 \times 10^6$ cells/ml is most useful, although the exact density is not critical. Cells can be harvested prior to blastocyst isolation if the investigator is proficient (30 to 45 min to complete the isolation); otherwise, ES cells should be harvested just prior to the injections.

REAGENTS AND SOLUTIONS

All solutions should be prepared from tissue culture grade reagents. Water should be tissue culture grade—high resistance and endotoxin free. Sterilize all final solutions by filtration or prepare from sterile stocks. Use of disposable sterile plasticware is advised to prevent microbial and detergent contamination. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Avertin, 2.5%

Stock solution, 100%: Dissolve 5 g of 2,2,2-tribromoethanol (Aldrich) in 5 ml of *tert*-amyl alcohol (Aldrich). Make sure reagents are at room temperature before mixing. Shake the container until all the tribromoethanol is in solution.

Working solution, 2.5%: Add 2.5 ml of 100% avertin to 100 ml of phosphate buffered saline (PBS; APPENDIX 2) prewarmed to 37°C. Shake the container until the avertin is completely dissolved.

Store 100% and 2.5% avertin in the dark at 4°C. The solutions may be kept for several months. Toxic decomposition products can accumulate. Discard any questionable solutions. Purchase components in the smallest available amounts (5 g tribromomethanol and 5 ml tert-amyl alcohol). There is batch-to-batch variation. With some preparations, a dose that is 5% to 10% higher than the recommended dose of 0.017 ml/g must be administered for complete anesthesia.

Blunted 25-G needle

Prepare the blunted needle by removing the tip of the bevel on a whetstone. There is no need to remove the entire bevel. The purpose of blunting is to prevent a sharp tip from penetrating the uterine wall during blastocyst isolation.

Embryo transfer pipet

Prepare from 25- μ l disposable microcapillary pipets (e.g., Fisher). Pipets can be purchased with a mouth pipet apparatus that consists of a pipet connector, tubing, and a mouthpiece. It is often convenient to replace the tubing with a longer piece of Tygon tubing. Pipets are pulled over a small flame to produce a very fine capillary on one end. Hold the pipet over a small flame at its center. Gently rotate the capillary to ensure uniform heating. When the glass begins to turn orange from the heat, remove it from the heat and simultaneously pull the ends apart. Pull straight apart until the capillary breaks to produce two transfer pipets with flush ends. If the capillary is pulled while the glass is still in the flame, the ends will be too fine. The exact diameter of the opening is not critical (100 to 150 μ m); however, it must easily admit a blastocyst.

For maximum control, the capillary should be back-filled with mineral oil slightly beyond the point where the capillary thickens to its largest internal diameter as illustrated in Fig. 23.7.6A. A small volume of injection medium is aspirated prior to manipulating embryos. Note that there is no reason to use air bubble markers unless the pipet is being loaded for reimplantation. Practice is required to make the transfer pipets and to become comfortable moving embryos by mouth control of the pipet.

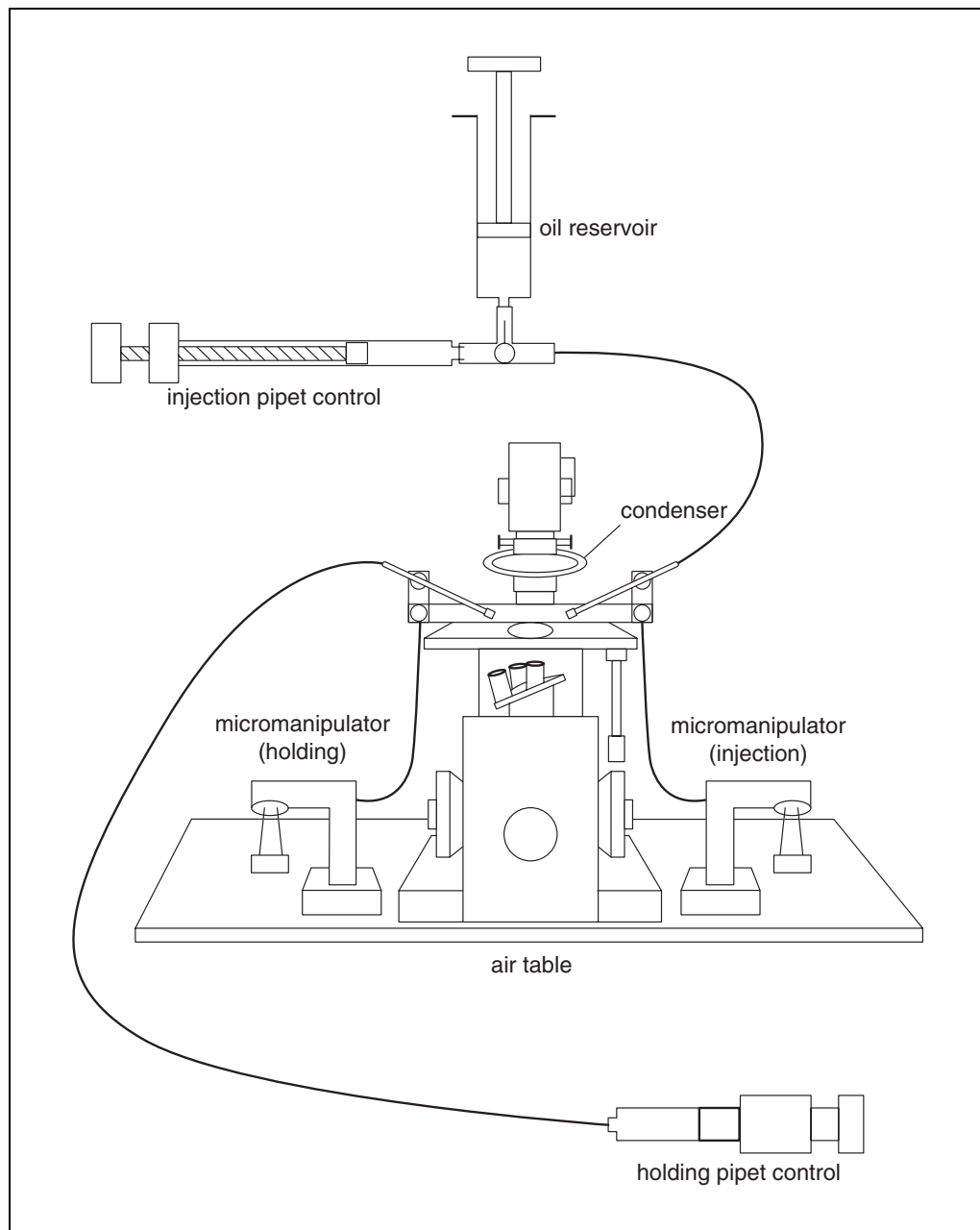


Figure 23.7.7 Injection apparatus. Diagram of a typical injection system. Note that the eyepieces are not included in the drawing. The injection pipet is illustrated as an oil-filled pipet with a syringe as an additional oil reservoir. The syringe for the injection pipet is usually placed on the left while the pipet and micromanipulator are placed on the right. The holding pipet and manipulator are placed on the left with its syringe on the right.

Injection apparatus

Many microscope configurations can be modified to work for blastocyst injections. The major components of the apparatus are illustrated in Figure 23.7.7. The basic microscope is an inverted microscope (standard microscopes can be used with a modified injection chamber) with a fixed stage. Under most circumstances the microscope must be placed on an air table to isolate it from sources of vibration. 4 \times , 10 \times , and 20 \times phase objectives are sufficient. A somewhat better image can be

continued

obtained using Nomarski differential interference contrast optics. Nomarski optics are particularly useful if the microscope will double as an oocyte injection microscope. A glass injection chamber must be used to take advantage of Nomarski optics. Joystick micromanipulators can be obtained from several sources (e.g., Nikon, Eppendorf). They should provide fine 3-dimensional control. The manipulators should also have 3-dimensional mechanical coarse controls. The pipet holders should be mounted directly to the microscope frame (preferred) or securely to the isolation table. In either case, the pipets should not move with the stage. Pipet controls consist of threaded syringes and PE tubing. The lines can be oil-filled or air-filled. The author prefers an oil-filled line, particularly for the injection pipet, because it provides slightly more control. Place a 10-ml oil-filled syringe in the line using a 3-way stopcock (Fig. 23.7.7) when using an oil-filled system. This provides easy access for refilling the system. The controls can be purchased (e.g., from Eppendorf or Nikon) or assembled. If assembling the components, use a 2.5-ml gas-tight Hamilton syringe with threaded plunger for the injection pipet and a 2-ml Gilmont Instruments syringe (VWR Scientific) for the holding pipet. This holding system can be used with or without oil. A typical set-up has the injection tip and manipulator on the right side with the syringe on the left side. This way the tip can be moved with one hand and suction can be applied with the other. The holding pipet is placed on the left side with its syringe on the right side. Air bubbles must be removed from oil-filled systems.

Injection medium

To a 1-liter packet of Dulbecco's Modified Eagle Medium (DMEM; Life Technologies), add:

20 ml 1 M HEPES

2.57 g NaCl

H₂O to 1 liter

pH should be 7.1 to 7.2

Sterilize the solution by filtration

Store in 40-ml aliquots up to 6 months at 4°C, protected from light

Prepare the final solution on the day of injection by combining:

40 ml of the DMEM solution

4.4 ml of heat-inactivated fetal bovine serum (10% v/v final)

0.5 ml of 100× penicillin/streptomycin (Life Technologies)

The final solution can be kept at room temperature for the day and then discarded.

Note that the NaCl replaces sodium bicarbonate in the DMEM. Do not add sodium bicarbonate as indicated on the package insert. With this modification and with the addition of HEPES the medium will maintain the proper pH in room air. Do not use this modified medium in a CO₂ incubator because the pH will be incorrect.

Human chorionic gonadotropin (HCG)

HCG (also known as Pregnyl) can be purchased from Organon, Inc. Store lyophilized powder at 4°C. Resuspend the lyophilized powder in sterile phosphate buffered saline (PBS; APPENDIX 2) to a final concentration of 50 U/ml. Store solution in 1-ml aliquots at -80°C for up to 6 months. Avoid multiple freeze/thaw cycles.

Microdrop cultures

Prepare by placing a small volume of injection medium (50 to 150 µl) on a 35 × 10-mm petri dish (Falcon). Cover the drop with light white mineral oil (Sigma) to prevent evaporation, and keep the dish warm on a slide warmer.

Pregnant mare's serum (PMS)

PMS (also known as Gestyl) can be purchased from Professional Compounding Centers of America (5 vials may be purchased at 1000 IU each). Prepare a sterile-filtered 1 mg/ml solution of bovine serum albumin (BSA) in phosphate buffered saline (PBS; *APPENDIX 2*). Resuspend the PMS in PBS-BSA to a final concentration of 50 U/ml. Store lyophilized powder at 4°C. Store the solution in 1-ml aliquots at -80°C for up to 6 months. Avoid multiple freeze/thaw cycles.

In the author's experience PMS purchased from this company works more consistently than that purchased from Sigma or Calbiochem.

COMMENTARY

Background Information

The methods described in this unit have their origins in studies of mammalian development and cellular differentiation. As early as 1961, the first true mouse chimeras were produced by the fusion of 8-cell embryos (Tarkowski, 1961). The potential for chimera production was enhanced with the development of injection techniques. Gardner (1968) described the first injection chimeras generated by the injection of cells from dissociated blastocysts into the blastocoel cavity of host blastocysts. Just like today, color differences were used to verify chimerism. Subsequently, Brinster (1974) demonstrated that cells that were not of immediate embryonic origin could produce chimeras. In his studies, the testicular teratocarcinoma cell line, OTT 6050, produced chimeras when injected into host blastocysts. One year later, Mince and Illmensee (1975) showed that the same teratocarcinoma cell line could produce a germ line chimera. These results were interesting primarily because of their implications for cell differentiation, but they also raised the possibility of creating genetic alterations in cell lines that could be transmitted through the mouse germ line by chimera production. However, this could not be done routinely with teratocarcinoma cell lines because they are unable to contribute to the germ line on a regular basis (Rossant and McBurney, 1983). In 1981 the final breakthrough occurred when Evans and Kaufman (1981) and Martin (1981) described the isolation of embryonic stem (ES) cells. Bradley et al. (1984) demonstrated that ES cells were capable of contributing to the germ line with much greater frequency than tumor-derived teratocarcinoma cell lines. Within a few years, genetically altered ES cells were shown to contribute to the germ line (Gossler et al., 1986; Robertson et al., 1986).

Although the injection of ES cells into blastocysts is the most common method for generating germ line chimeras, there are some alter-

natives. Chimeras can be generated by injecting ES cells under the zona pellucida of 2.5-day p.c. embryos (morula). Technically, there is no particular advantage to this over blastocyst injections. In fact, morula isolation is somewhat more difficult than blastocyst isolation because the embryos are still in the oviduct. On the other hand, aggregation chimeras are becoming a more accepted alternative. Two basic methods have been developed. Both involve the physical apposition of ES cells with morula (with zona pellucida removed). Attached ES cells are taken up by the embryo as it develops in culture to a blastocyst with relatively high efficiency. One method utilizes 10- to 15-cell clumps of ES cells that are placed in contact with 1 or 2 morulas in a microwell (Nagy and Rossant, 1993). The second method utilizes single-cell suspensions of ES cells (Hogan et al., 1994). Aggregation methods are technically less difficult and less expensive because the manipulations can be performed under a dissecting scope. However, these methods may be more sensitive to ES cell quality.

Critical Parameters and Troubleshooting

Creating chimeras by the injection of ES cells into blastocysts is a difficult procedure. Problems can occur in each of the three main steps: embryo isolation, injection, and reimplantation. Each of the steps requires both familiarity with the protocol and practice to avoid problems. It is best to practice the procedures separately before trying to combine them on a single day. Many of the problems are under the investigator's control. However, some aspects of these protocols rely on the basic biology of mice and are not entirely under the control of the investigator.

The two common problems encountered during embryo isolation are low embryo yields and poor embryo quality. A variety of factors can contribute to these problems including

housing conditions, animal age, the timing of superovulation, and reagent quality. Typically, the most consistent embryo yields will be obtained from mice that are housed in virus antibody free (VAF) or specific pathogen free (SPF) facilities with controlled temperature and humidity and consistent light/dark cycles. Animals should not be overcrowded. Mating efficiency can be reduced by noise and constant activity in the animal rooms. Animal age can dramatically affect embryo yields. The age effect will vary somewhat between strains. The C57BL/6 embryo donors that we use are purchased from Taconic. The best yields are obtained from females that are superovulated before their first estrus and from males that are between 8 weeks and ~7 months old. Note that there can be subtle strain variations between C57BL/6 mice obtained from different vendors. The mice must be superovulated at a consistent time during the day that is coordinated with the light/dark cycle. We have had the greatest success with PMS purchased from Professional Compounding Centers of America. Information about potential problems can be gleaned from the embryo yields. A large number of poorly developed (unfertilized) embryos indicates that the superovulation is working, but the males are performing badly. They may be getting too old or some may not be plugging females. For optimal yields, mating information can be kept for each male. Those that are not plugging females on a routine basis should be sacrificed. A low total embryo yield suggests a problem with the females or the superovulation procedure. No matter how much care is taken there will be some days when the mice will be disagreeable and embryo yields will be low.

The injection procedure is visual; it is fairly easy to judge success. The ES cells can be seen as they are introduced into the cavity. The health of the injected blastocysts can be assessed by their ability to re-expand. Although not all healthy injected blastocysts will re-expand, a large percentage should show some semblance of a blastocoel cavity within 1 to 2 hr after injection. However, there is no reason to postpone reimplantation for re-expansion unless there is a concern about embryo viability. Blastocysts will survive substantial damage to the trophoblast layer including penetration of both sides by the injection tip. Blastocysts are most severely damaged by partial aspiration into the holding pipet.

Reimplantation is straightforward but requires practice to acquire proficiency. The pro-

cedure should be repeated until most reimplanted unmanipulated embryos grow to term. Once this proficiency is achieved, and assuming that injections were successful, low pup yields or poor chimerism can be attributed to the quality of the ES cell line. Additional clones should be injected under these circumstances.

Anticipated Results

Pups should be born 16 to 17 days after reimplantation. Approximately 7 days after birth the coat color becomes visible. Assuming proficiency with the procedures, the number of live-born pups and the degree of chimerism are completely dependent on the quality of the ES cell line. Those new to the procedures should inject wild-type (not manipulated also known as parental) ES cells of known quality. Parental ES cells should produce high birthrates: >50% of the reimplanted embryos should survive. Eighty percent or greater should appear chimeric by coat color. With male parental ES cell lines, a significant male sex bias of the chimeras should be observed due to sex conversion of female blastocysts. A range of coat chimerism is usually observed. However, with parental cells many of the chimeras should display hair color that is entirely ES cell derived.

Results with manipulated ES cell lines can vary considerably. Assuming there is no phenotype of the altered gene, cells that have been well taken care of should produce results similar to parental cells. Typically, however, manipulated cells produce less extensive chimerism than parental cells. Birthrates are usually lower (25 to 30%) and the percent of pups that are chimeric may be lower. In extreme cases ES cell clones can be lethal (small litters with no viable chimeras) or appear not to contribute to the embryo at all (reasonable litter sizes with no apparent chimeras). Different clones with the same genetic alteration must be injected under these circumstances.

Time Considerations

Although the procedure spans 7 days, only the day of injection requires a significant time commitment. With experience blastocyst isolation should take 30 to 60 min. Injections should take 2 to 3 hr and reimplantations should take 30 to 60 min. Including cell preparation, the entire procedure should take 4 to 6 hr to produce 20 to 36 injected and reimplanted blastocysts. A novice investigator will spend a very long day trying to complete the procedure. Blastocyst isolation and reimplantation should be

mastered before attempting to complete the entire procedure.

Literature Cited

- Bradley, A., Evans, M., Kaufman, M.H., and Robertson, E. 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* 309:255-256.
- Brinster, R.L. 1974. The effect of cells transferred into the mouse blastocyst on subsequent development. *J. Exp. Med.* 140:1049-1056.
- Evans, M.J. and Kaufman, M.H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292:154-156.
- Gardner, R.L. 1968. Mouse chimeras obtained by the injection of cells into the blastocyst. *Nature* 220:596-597.
- Gossler, A., Doetschman, T., Korn, R., Serfling, E., and Kemler, R. 1986. Transgenesis by means of blastocyst-derived embryonic stem cell lines. *Proc. Natl. Acad. Sci. USA* 83:9065-9069.
- Hogan, B., Beddington, R., Constantini, F., and Lacy, E. 1994. *Manipulating the Mouse Embryo: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Martin, G.R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. USA* 78:7634-7636.
- Mince, B. and Illmensee, K. 1975. Normal genetically mosaic mice produced from malignant teratocarcinoma cells. *Proc. Natl. Acad. Sci. U.S.A.* 72:3585-3599.
- Nagy, A. and Rossant, J. 1993. Production of completely ES cell-derived fetuses. *In Gene Targeting: A Practical Approach* (A.L. Joyner, ed.). Oxford University Press, New York.

Robertson, E., Bradley, A., Kuehn, M., and Evans, M. 1986. Germ-line transmission of genes introduced into cultured pluripotential cells by a retroviral vector. *Nature* 323:445-448.

Rossant, J. and McBurney, M.W. 1983. Diploid teratocarcinoma cell lines differ in their ability to differentiate normally after blastocyst injection. *Teratocarcinoma Stem Cells, Cold Spring Harbor Conferences on Cell Proliferation* 10:625-633.

Tarkowski, A.K. 1961. Mouse chimeras developed from fused eggs. *Nature* 190:857-860.

Key References

Hogan et al., 1994. See above.

Papaioannou, V. and Johnson, R. 1993. Production of chimeras and genetically defined offspring from targeted ES cells. *In Gene Targeting: A Practical Approach* (A.L. Joyner, ed.) pp. 33-61. IRL Press, Oxford.

These two references, written by experts in the field, represent thorough compilations of transgenic methods.

Internet Resources

<http://www.biosupplynet.com>

Search this web site to obtain a current list of suppliers for materials and reagents used in the production of chimeras by blastocyst injection.

Contributed by David A. Conner
Harvard Medical School
Boston, Massachusetts

Managing a colony of mice created by gene targeting is not complicated, but does require some strategic planning. Prior to beginning matings to generate experimental animals, the investigator should consider how the background strain might affect analyses. This unit describes basic mating procedures, a simple method to take advantage of coat-color differences between strains to screen for germline transmission, and the rationales for selecting strains for the generation of experimental animals. Because of the difficulty of regenerating an inbred background after outcrossing, it is recommended that, at a minimum, the allele be maintained in an inbred background whenever possible. In addition, the unit reviews common methods for marking mice and two simple methods to isolate genomic DNA for genotyping by polymerase chain reaction (PCR) or Southern blot. Finally, strategies and software for colony data management are reviewed.

COLONY ESTABLISHMENT, EXPANSION, AND MAINTENANCE

Housing

Mice should be housed in a specific-pathogen-free (SPF) facility if possible. SPF animals will typically live longer and display greater fecundity than comparable animals housed in open facilities with rampant viral and parasitic infections. Common mouse pathogens that are excluded in SPF facilities can affect experimental results. Rooms should have a consistent light/dark cycle to facilitate mating. Minimizing human activity and noise in the room during the light phase and particularly during the dark phase will increase mating performance. Males destined for mating should be housed separately. Mature males that have mated are likely to fight if placed together, even if they were originally littermates. In most cases unrelated females may be housed together at any stage of their lives. Facility housing guidelines should be followed to ensure that mice are not overcrowded and that cages are kept clean.

Choosing Potential Founders

A knockout colony begins with chimeric founders. In most cases, the level of chimerism is estimated by the level of embryonic stem

(ES) cell contribution to coat color. Because the typical protocol involves the injection of agouti ES cells into black host embryos (C57BL/6), the degree of chimerism is estimated by the percentage of the coat color that is agouti. There is no absolute rule regarding which chimeras should be mated. Most ES cell lines are male, and thus most germline chimeras will be male—high-level ES cell contribution to male host embryos will produce male chimeras and will convert the sex of female host embryos to male. This results in a preponderance of good male chimeras. However, some female chimeras do pass. The incidence of female chimera transmission is not well documented, because most groups only mate the male chimeras.

When injections result in a number of high-level chimeras (>90% by coat color) and sex conversion, there is a high likelihood of germline transmission. In this case, only good male chimeras need to be mated. When the level of chimerism is lower and there is no evidence of sex conversion, all reasonable male and female chimeras should be mated. The possibility of injecting other targeted clones should be considered. Low-level chimeras (<50% by coat color) do pass. However, this is more common when other littermates have high levels of chimerism.

Detection of Germline Transmission

Chimeras may be crossed with specific wild-type strains to facilitate the detection of germline transmission. For example, most of the common ES cell lines are homozygous wild-type *agouti* and are typically injected into C57BL/6 embryos. Crosses of the chimeras with C57BL/6 mice will yield agouti pups when germline transmission occurs and black pups in the absence of germline transmission. When transmission rates are low, a large number of offspring can be screened visually for germline transmission with minimal effort. Outbred mice with coat color characteristics similar to C57BL/6, such as Black Swiss (Taconic), may also be used. With increased litter sizes (compared to C57BL/6) and good maternal instincts, 100 offspring can be generated in as few as 8 litters from outbred females.

Matings are initiated by placing 7- to 8-week-old (or older) chimeras with wild-type females that are at least 6 weeks old. Two females can be placed with each male. The

females can be replaced weekly; the estrus cycle is 4 to 6 days in length. Alternatively, the investigator can check for plugs (UNIT 23.7, Support Protocol 1) daily and replace the plugged females with new females. It is often worthwhile checking plugs initially to ensure that the males are mating. Some males will never mate. This can be caused by infertility due to partial sex conversion. Other males may take several weeks to begin mating. Females that were mated to different males should be housed separately unless they are individually marked. Pregnant females should be separated prior to delivery to avoid confusion over the lineage of offspring and to avoid overcrowding.

Gestation is typically 18 to 21 days depending on the strain and the litter size (Fig. 23.8.1). The ability of new mothers to care for pups depends greatly on the strain. As a general rule, outbred mothers will produce larger litters and take good care of litters even if disturbed. Inbred mothers will typically have smaller litters and may be more sensitive to environmental stress. As a precaution, pregnant females can be placed in clean cages shortly before birth. The cages should not be disturbed until pups are at least 4 days old. Activity and noise in the room should be minimized. The addition of paper towels or nestlets from which nests can be made can be helpful. Poor mothers may be encouraged to care for their offspring in the presence of another mother with a similarly aged litter, or the litter may be removed and placed with another mother with a similarly aged litter. In this case, some of the pups from the foster mother should be removed so that the litter is not too large. These problems should not be encountered when using Black Swiss and C57BL/6 females.

Coat color should become apparent within 1 week after birth. Germline transmission can occur at variable rates ranging from 0% to 100% of the offspring being ES cell-derived. As a general rule, the author would generate ~100 offspring before giving up on a chimera. Chimeras usually pass at a consistent rate: an animal that is passing at 80% is unlikely to produce subsequent litters with no ES cell-derived pups. Obviously, there is no need to mate a poor transmitter if chimeras from the same ES cell clone are passing at a higher frequency. With genes targeted on autosomes, only half of the ES cell-derived offspring will inherit the altered allele, so the pups must be genotyped even if it is clear from the coat color that germline transmission has occurred.

Generating Mice For Analysis

Most ES cell lines are derived from inbred strains (Table 23.8.1). Inbred mice and the ES cells derived from them are genetically homogeneous and homozygous at all loci. Strains are defined as inbred after 20 generations of sequential brother-to-sister matings. Animals generated to screen for germline transmission by coat color may not be ideal for experimental analysis, because of the mixed genetic background (Fig. 23.8.2). Note that in the F1 generation the mice are genetically identical (half 129 and half C57BL/6 in Figure 23.8.2). The F1 generation can be used to study heterozygous phenotypes without concern of genetic heterogeneity. When mice that are heterozygous for the targeted allele are crossed to make an F2 generation with homozygotes, they will inherit a variable complement of alleles from the two strains. This may increase the variability of a phenotype and result in the masking of subtle phenotypes. In addition, the targeted alleles and linked loci will both originate from the chimeric strain while the wild-type alleles and linked loci will come from the other strain (C57BL/6 in Fig. 23.8.2). Phenotypic differences between homozygous mutants and wild-type littermates may be due to differences in the genes surrounding the altered locus. Because of these issues, it is good practice to maintain the mutant allele in an inbred background whenever possible.

An inbred line is established by mating a germline chimera with wild-type mice of the same strain from which the ES cell was derived. Table 23.8.1 lists common ES cell lines and their strains. There is no rapid way to re-establish an inbred line after the original chimeras are gone. Heterozygotes from the first generation can be crossed to produce wild-type and homozygous mutants for study. Homozygous wild-type and homozygous mutant lines can be maintained separately if the mutants are viable and fertile, avoiding the necessity of genotyping each new litter.

In many cases an investigator is stuck with a mutant allele in a mixed background. Under these circumstances, homozygous mutants and wild-type littermates should be generated by crossing heterozygotes. Even when homozygotes are viable and fertile, separate wild-type and homozygous lines should not be used to generate experimental animals because distinct lines with different phenotypes may develop. The investigator should also consider making a congenic line (see discussion of Generating

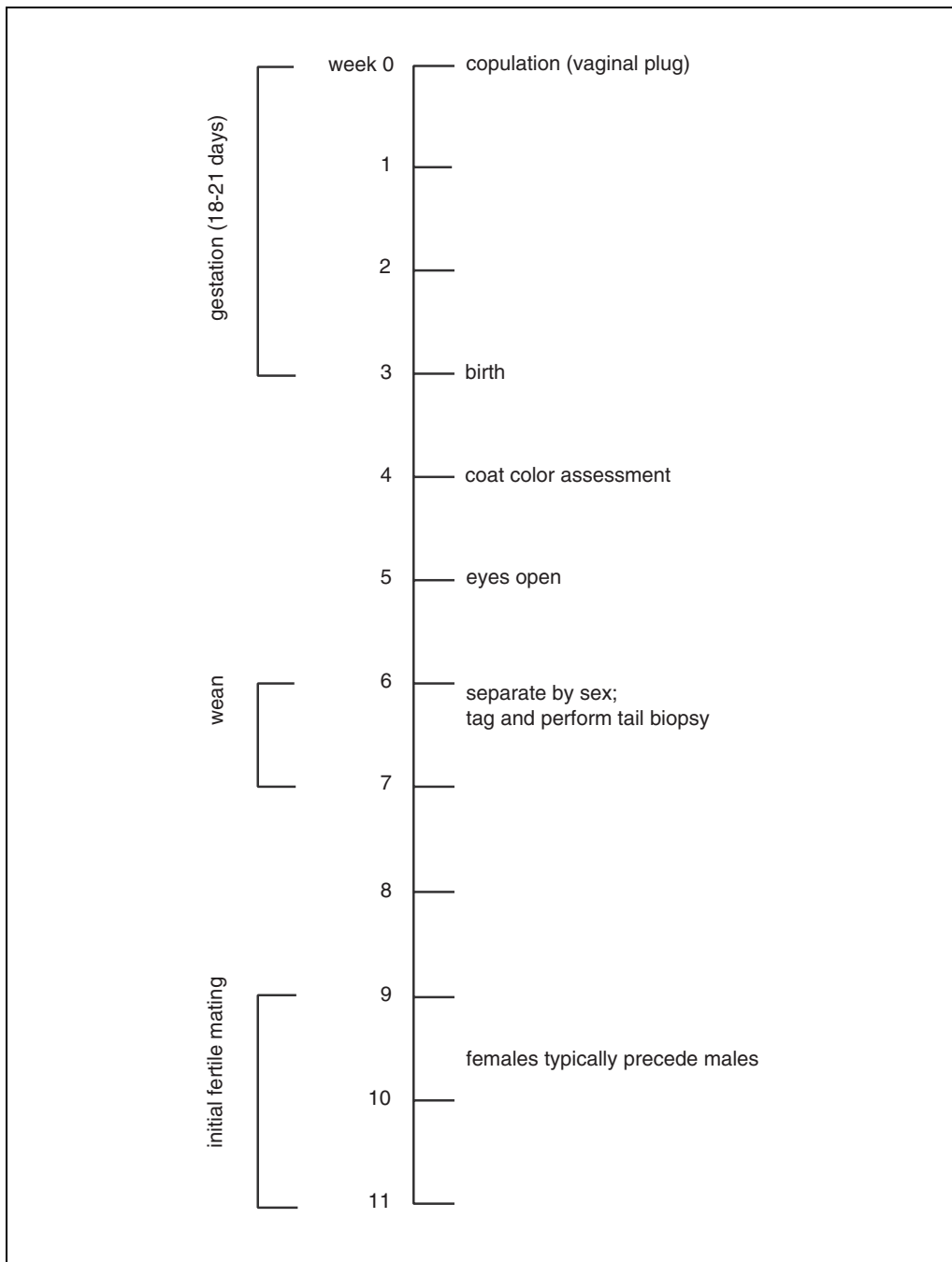


Figure 23.8.1 Mouse generation time line. Time line from copulation to sexual maturity of a typical mouse. By gestational day 14, pregnant mothers can be detected easily by their increased size. The age of a litter can be estimated from the time that the coat color is evident (~1 week after birth) or the day when the eyes open (~2 weeks after birth). Sexual maturity may occur as early as 4 weeks. However, the initial fertile mating does not usually occur until the mice are 6 to 8 weeks old. Females can often be mated at 6 to 7 weeks of age while males may not plug females until 7 to 8 weeks of age. There can be substantial differences in the generation time between strains. As a conservative estimate for calculations consider a generation to be 3 months.

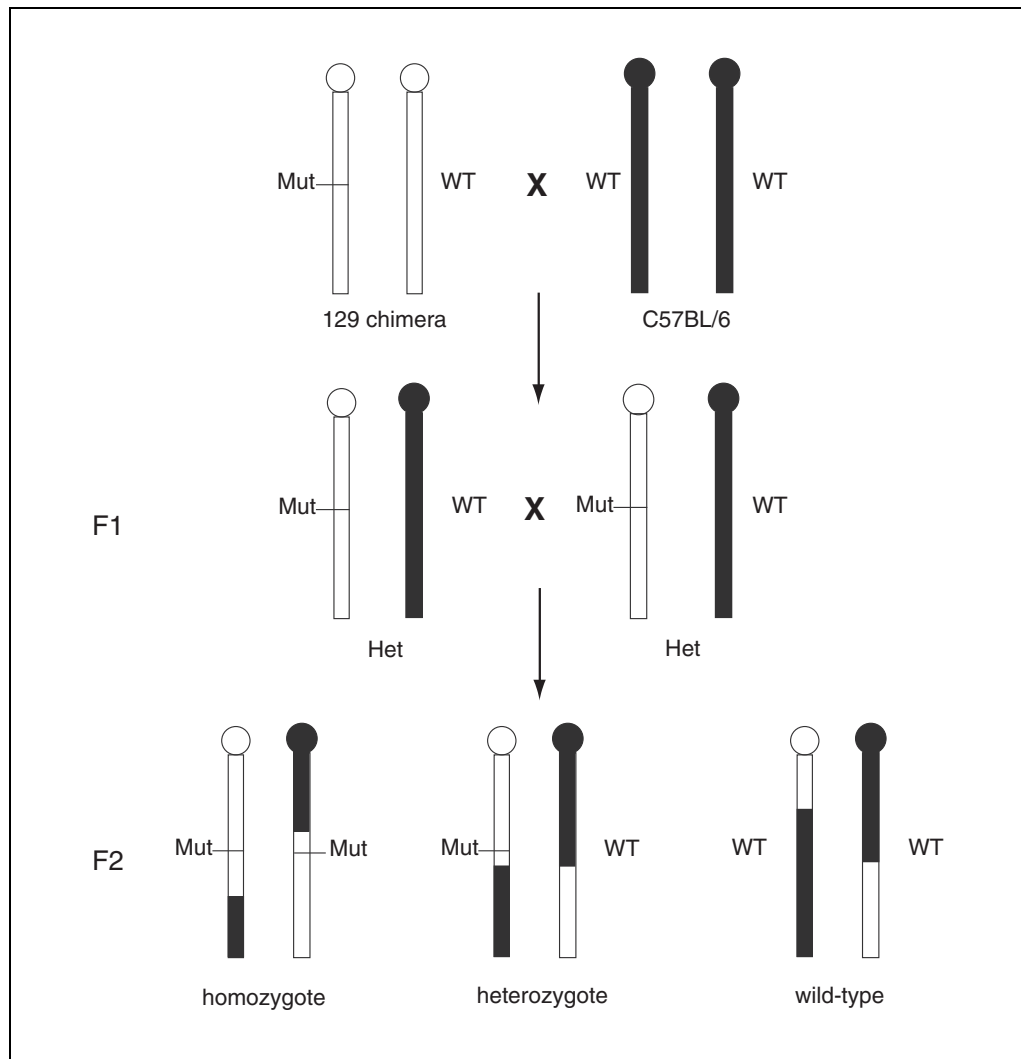


Figure 23.8.2 Genetic heterogeneity in the F2 generation. Many investigators create a mixed background by crossing the 129 chimera with C57BL/6 mice to test for germline transmission. If the ES cell line was derived from an inbred strain, members of the F1 generation are genetically identical except for the mutant locus (Mut). The F1 generation may be useful for the analysis of heterozygous phenotypes. However, the F2 generation is genetically heterogeneous because of recombination and independent assortment. Some phenotypic differences between F2 homozygotes and their wild-type littermates may be caused by the background heterogeneity and not by the mutant allele.

Congenic Lines, below) to reduce the genetic heterogeneity.

There are many circumstances when outcrossing is appropriate. First, many of the inbred lines from which ES cells were established are difficult to obtain or to work with. Poor maternal instincts and small litter sizes can make it impossible to generate a sufficient quantity of experimental animals. Chimeras can be crossed with an inbred line, such as C57BL/6, with good mating behavior. Second, certain experiments cannot be performed in specific backgrounds. For example, 129/SvEvTac mice do not develop morphine tolerance (Kolesnikov et al., 1998). Under these

circumstances the mutant allele must be moved into a responsive background if studies regarding morphine tolerance are to be pursued. In these examples where outcrossing is necessary, the investigator should consider making a congenic line (see discussion of Generating Congenic Lines, below) to reduce the genetic heterogeneity. Third, outcrosses can be used to identify genetic modifiers of the mutant allele. Variation of the mutant phenotype between inbred and outcrossed or congenic backgrounds strongly suggest the presence of genetics modifiers. The modifiers may be interesting in their own right and can be mapped (and eventually identified) or they may simply make

Table 23.8.1 Common 129 ES Cell Lines and Strains of Origin^a

Cell line	Strain (new nomenclature)	Strain (old nomenclature)	Commercial strain source ^b
RW4	129X1/SvJ	129/SvJ	Jackson
R1 ^c	129X1/SvJ x 129S1/Sv-/+ ^P + ^{Tyr-c} <i>Kitt</i> ^{Sl-J/+}	129/SvJ x 129/Sv -/+ ^P + ^{Tyr-c} + <i>Mgf</i> ^{Sl-J/+}	Jackson
J1	129S4/SvJae	129/SvJae	—
D3	129S2/SvPas	129/SvPas	—
AB1	129S7/SvEvBrd- <i>Hprt</i>	129/SvEvBrd- <i>Hprt</i>	—
TL1	129S6/SvEvTac	129/SvEvTac	Taconic
TC1	129S6/SvEvTac	129/SvEvTac	Taconic
W4	129S6/SvEv	129/SvEv	Taconic
C1	129X1/SvJ	129/SvJ	Jackson
E14TG2a	129P2/OlaHsd	129/OlaHsd	—

^aSeveral common 129-derived ES cell lines are listed with their strain of origin (Simpson et al., 1997; Threadgill et al., 1997; Festing et al., 1999). Because of recent name changes, both the old and new nomenclature has been included to avoid confusion. The nomenclature revision is based on an analysis with microsatellite markers and consists of the addition of a letter and number before the backslash. The letter P designates 129 strains derived from the original parent strain. The letter S designates a congenic strain made to introduce the steel mutation. The letter X designates a strain that was generated by an unknown outcross. The number differentiates substrains of the parental strains.

^bCommercially available sources of the mice are also listed where known. Most 129 strains are not good breeders: they produce small litters and do not care well for the pups. The 129S6/SvEvTac strain is an exception: litters average 6.3 pups and mothers tend to care for their offspring.

^cR1 ES cells were generated from a cross of two 129 substrains.

experimentation less difficult (e.g., homozygous mutants might live longer in a different background).

Generating Congenic Lines

A congenic strain is a mouse strain that is isogenic to an inbred strain, but contains a chromosomal segment from another background. Construction of a congenic strain using traditional methods will take ~3 years. Congenic strains are created by an initial outcross followed by sequential back-crosses to an inbred strain (Figure 23.8.3). For example, a C57BL/6 congenic can be generated by outcrossing a chimeric male with a wild-type C57BL/6 female. This produces an F₁ generation in which the offspring may be considered heterozygous at all loci—half the loci are derived from the chimera and half are derived from the C57BL/6. The second generation, N₂, and all subsequent generations are created by back-crosses of animals carrying the mutant allele to C57BL/6. The N₂ generation will retain only 25% of the chimera's genome. Each subsequent back-cross will reduce the original genome content by 50%. By N₁₀, only 0.1% of the chimera's original genome will remain and

the animals are considered congenic. The estimated genomic contributions are based on statistical arguments; the actual proportions in a specific mouse may vary considerably from the expected value. Genes that are tightly linked to the targeted locus will be retained in the congenic strain. The amount of the donor genome linked to the targeted locus can be estimated in centimorgans (cM) as 200/N (for N > 5), where N equals the number of back-cross generations (Silver, 1995). For the N₁₀ congenic ~20 cM surrounding the targeted locus remain in the congenic.

The production of congenics can be accelerated using a newer approach known as speed congenics (Lander and Schork, 1994; Wakeland et al., 1997). The mating strategy is identical to the traditional approach, except that the mouse containing the least amount of the donor genome is selected at each generation for subsequent matings. The relative genome contributions are determined by screening for polymorphic marker loci spread throughout the genome at generation N₂ and all subsequent generations. The contaminating donor genome contribution can be reduced to the same level seen in the N₁₀ congenic in 5 or 6 generations,

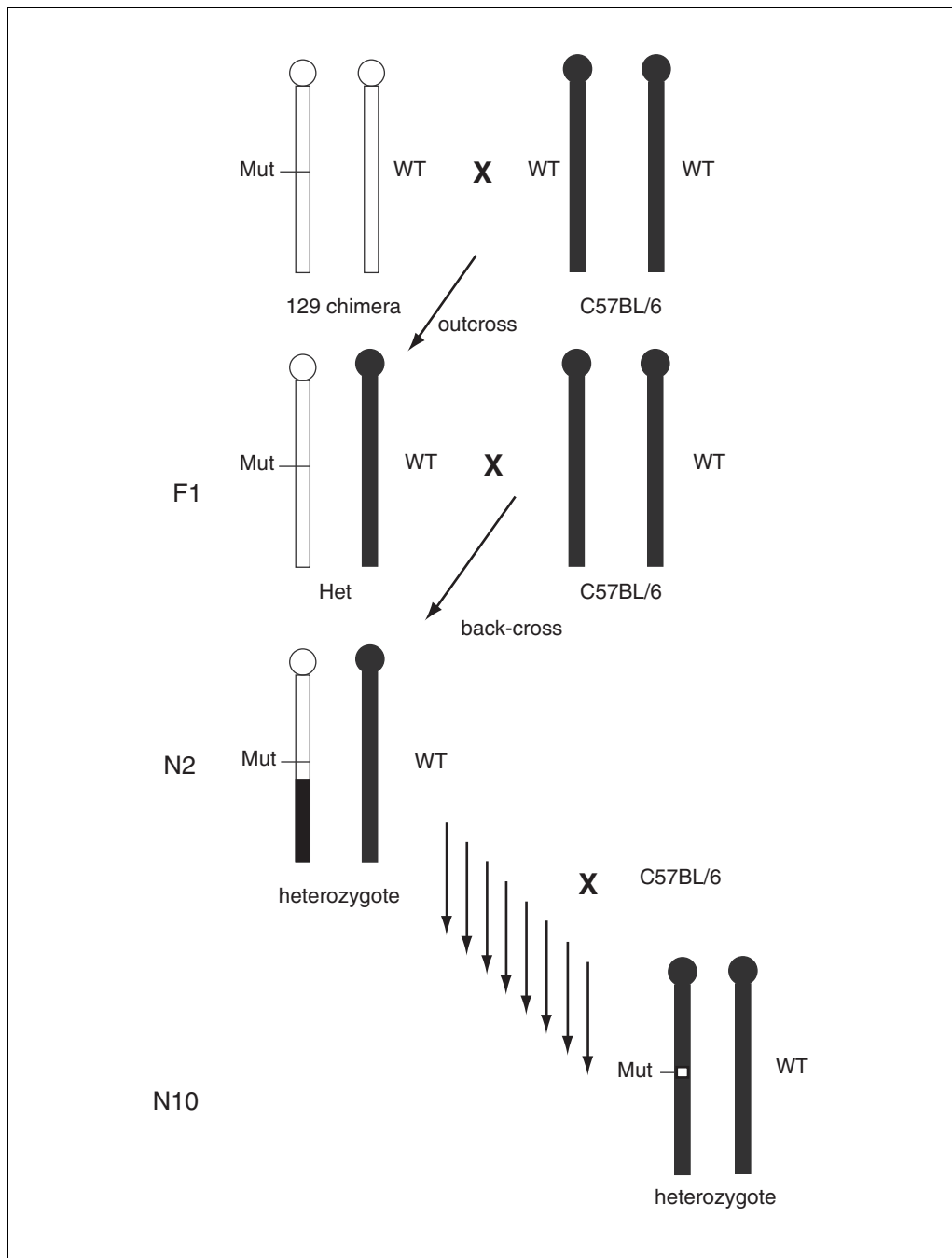


Figure 23.8.3 Generation of a congenic strain. Congenic lines are generated by an outcross to the host strain (C57BL/6) followed by successive back-crosses of mice carrying the mutant allele to the host strain. The strain is considered congenic after the tenth back-cross generation. Most loci except those tightly linked to the mutant locus are derived from the host strain and are homozygous. At this stage, brother-to-sister matings are initiated to generate experimental animals. Note that if all matings are performed with host females and donor males, the host Y chromosome will not be transferred. At least one back-cross between a donor female and host male followed by mating of a mutant male to a host female in the next back-cross is required to fix the C57BL/6 Y chromosome in the congenic line.

reducing the time required to make a congenic strain to 1.5 to 2 years. This rapid method requires larger colonies to select the “best” mouse for subsequent matings. In addition, few laboratories have the resources to perform the marker-assisted screening in house, although there are companies and core facilities that will perform the screening.

Long-Term Strain Maintenance

There are several options for maintaining mutant mouse lines when there is no immediate need to generate animals for experimentation. Most investigators do not choose to eliminate mouse lines, because of the time and effort required to regenerate them. A line can be maintained by keeping only a few mating pairs. Homozygous mutants, if viable and fertile, should be used, so that genotyping will not be necessary. Let the mating pairs breed actively. Most litters can be culled. Set up new mating pairs from the younger generations when the litter size becomes smaller or the litters become less frequent. The number of active mating pairs depends on the fecundity of the line: more pairs should be kept for lines that produce small litters or lines that have difficulty rearing pups. This may also be a good opportunity to perform back-crosses if a congenic line might be needed in the future. Alternatively, morula (2.5-day old embryos) can be frozen using a controlled-rate freezing protocol and stored in liquid nitrogen. Several of the major mouse suppliers provide morula-freezing services. There have been recent advances in sperm cryopreservation. However, this is not yet the method of choice because of the variable success of *in vitro* fertilization with frozen mouse sperm. Finally, some of the major mouse suppliers will freeze, maintain, and distribute interesting mutant mouse lines if the investigator is willing to relinquish control over their distribution.

MOUSE IDENTIFICATION

Individual animals must be identifiable when maintaining colonies of mice with different genotypes and backgrounds. A variety of methods have been used to mark individual mice including ear tagging, ear punching, toe clipping, subcutaneous transponder implantation, and tail tattooing. To a large extent, the method that a laboratory will use depends on personal preferences. However, the method must be acceptable to the institutional animal care committee. Note that there may be some variation in protocol acceptance between different institutions, so that the resident animal resource center should be consulted before choosing a protocol. None of the procedures will cause chronic pain. There are several practical factors to consider when choosing a method of identification (Table 23.8.2).

Subcutaneous transponders are the most expensive of the identification methods. The typical cost of each transponder is between \$5 and \$10. In most cases, they are not easily reused. Readers typically cost between \$500 and \$2000. Implanted animals can be identified easily and the method is permanent in most cases. However, because of the lack of standardization, transponders from different companies usually require specific readers. Subcutaneous insertion is not difficult, but may require more restraint than for the other methods: anesthesia may be necessary until the procedure becomes routine.

The cost of ear punching and toe clipping is negligible beyond the purchase of a punch or scissors. Marking the mice by these methods is simple, but may become tedious when large, continuous numbering schemes are used (see Hogan et. al., 1994 for numbering schemes). Reading the number may also be difficult when using continuous numbering schemes. Both methods are very useful for identifying neo-

Table 23.8.2 Methods of Identification

Method	Cost	Ease	Anesthesia	Comments
Ear tag	Low	Simple	No	Tags can tear out of older animals
Transponder	High	Simple	Yes/No	Lack of standardization requires transponder specific readers
Ear punch (toe clip)	Low	Simple	No	Good for identification of neonates Large numbering schemes can be difficult to read
Tail tattoo	Moderate	Requires training	Yes	Labeling with complex numbering schemes can be tedious

nates in a single litter when the mice are too small for other forms of identification. Note that neonates may also be marked with India ink for short-term identification.

Tail tattooing is permanent. Tattooing devices can be purchased for \$500 to \$1000. Marking of many animals this way can be tedious because of the technique and the requirement for anesthesia.

Ear tags are inexpensive and easy to apply. Mice should be tagged at weaning or later: tags are more likely to tear out when placed in the small ear of a younger animal. Tags should be clipped to the base of the ear, not to the edge. This will reduce the chance of tear-out. Tags may eventually tear out in older animals. Place even-numbered tags in the right ear and odd numbers in the left ear to reduce confusion when tags fall out; the torn or punctured ear will indicate whether the original tag number was even or odd.

DNA PREPARATION FOR GENOTYPE ANALYSIS

Typically, the mice are tagged for identification and tail biopsied for genotyping simultaneously at weaning (age 3 to 4 weeks). At this age, biopsies are thought to be less traumatic because the tip of the tail is still cartilaginous. Check institutional guidelines for recommended procedures: an inhalation anesthetic may be required for older animals. DNA can be isolated from tail samples in several ways depending on the method of genotype analysis. PCR screens should be used whenever possible for the sake of speed and convenience.

The following are step-by-step instructions for the abovementioned procedures.

Tail Biopsy

1. Remove 0.5 cm or less of the tip of the tail from a 3 to 4 week old pup using clean, sharp scissors. The same scissors can be used for multiple mice. The mice will not get infections at the cut site and cross-contamination of DNA samples by the scissors should not be a problem. A styptic pen can be used to stop the bleeding. However, there is very little blood loss at the cut site in normal mice.

2. Place the tail in a labeled 1.5-ml microcentrifuge tube. Samples can be collected over 1 hr or more at room temperature without affecting the quality of the final DNA. The biopsies can be frozen at -20°C or processed immediately after collection. When Southern analysis is used to detect large bands (>15 kb), samples should be processed after collection

because freezing will decrease the DNA quality, making the identification of large bands more difficult.

3. Prepare DNA using the standard protocol for Southern analysis and PCR, or the rapid protocol for PCR (see steps for each procedure below). The standard method produces DNA that is suitable for Southern analysis (*UNIT 2.9A*) or PCR (*UNIT 15.1*). It is an overnight procedure. The rapid method produces template for PCR in less than 30 min.

Standard DNA Isolation for Southern Analysis and PCR

1. Add 0.5 ml of digestion solution consisting of 100 mM Tris-Cl, pH 8.5 (*APPENDIX 2*), 5 mM EDTA, 200 mM NaCl, 0.5% Tween 20, and 1 mg/ml proteinase K to each tail biopsy. The proteinase K should be added fresh just prior to digestion.

2. Incubate overnight at 55°C . The tissue should dissociate completely leaving some undigested debris and hair at the bottom of the tube. Alternatively tails may be digested for ~ 4 hr at 65°C with intermittent agitation.

3. Mix the samples by gentle vortexing and centrifuge at top speed in a microcentrifuge at room temperature for 2 min.

4. Pour the supernatant into a new 1.5 ml microcentrifuge tube containing 0.5 ml isopropanol. Mix the tubes thoroughly by inversion and let stand at room temperature for 10 to 30 min.

5. Centrifuge at top speed in a microcentrifuge at room temperature for 2 min. Aspirate the supernatant and wash the pellet once with 70% ethanol.

8. Allow the DNA pellet to air dry and then resuspend in 100 μl TE buffer, pH 7.6 (*APPENDIX 2*). The DNA can be dissolved by gentle vortexing. If the DNA pellet is allowed to dry out completely it may become difficult to dissolve. The DNA concentration will range from 0.5 to 2 $\mu\text{g}/\mu\text{l}$, depending on the size of the tail biopsy. The DNA is suitable for restriction digestion or as a PCR template. Store the DNA at 4°C .

Rapid DNA Isolation for PCR Analysis

1. Add 0.5 ml of 0.05 M NaOH to each tail biopsy. Incubate for 10 to 20 min at 95°C .

2. Remove the samples from the heat and neutralize by adding 50 μl of a solution of 1 M Tris-Cl, pH 8 (*APPENDIX 2*), and 10 mM EDTA. Vortex to mix. The tail does not dissociate completely with this method, but DNA is released into solution. 1 or 2 μl can be used directly in a 20- μl PCR reaction (*UNIT 15.1*). The

protocol can be scaled to work in a 96-well format.

INFORMATION MANAGEMENT

Mouse colonies have a tendency to expand rapidly. Good record-keeping practices should be established immediately. Records serve several purposes. Colony sizes and expenses can be kept to a minimum when an accurate tally of mice is maintained. Phenotypic variation attributable to differences in strain background and subtle difference in fecundity are more easily detected when complete records are maintained. Paper-based records will become unwieldy for all but the smallest colonies. Fortunately, there are many options for computer-based record keeping. Table 23.8.3 lists some examples of colony management software including commercial and freeware options. Demos of most of the commercial applications can be tested prior to purchase. Many investigators may find that simple, locally constructed databases are most convenient. These can be adapted to individual lab requirements and are usually less expensive than commercial colony management software.

Whether using commercial, freeware, or locally constructed databases, it is a good practice

to record information both about individual animals and matings. Tables 23.8.4 and 23.8.5 each list a minimum set of data that should be collected for individual mice and matings, respectively. Mating data should be recorded even when there are no offspring. This type of database is easily created using Filemaker Pro. Filemaker Pro is relatively inexpensive, cross-platform (Macintosh and PC), and simple to program. Databases can be accessed through a network, allowing multiple users to log on simultaneously. Several Filemaker Pro colony databases can be downloaded from the Web site listed in Table 23.8.3 for MouSeek. These templates may be used directly or modified to individual lab requirements. In addition, handheld (Palm-compatible) versions of Filemaker Pro databases are easily created and can be synchronized with the desktop application. Handheld copies are convenient for data entry and review inside the animal facility.

Regardless of the record-keeping system, it is much more common to regret recording too little information than too much. However, a database is only as useful as the data it contains; the database must remain simple enough so that a laboratory group will actually record the information.

Table 23.8.3 Examples of Colony Management Software^a

Program	Source	Platform	Comments
Colony	Locus Technology, Inc. http://www.locusttechnology.com	Windows	Very complete packages, good for large-scale operations Options for barcode scanning and transponder interface
Big Bench Mouse	Big Bench Software http://www.bigbenchsoftware.com	Macintosh Windows	Basic management software that can be networked
Progeny	Progeny http://www.progeny2000.com	Windows	Pedigree software that can be used for colony management
Scion	Topaz Technologies http://www.topaztracks.com	Macintosh Windows	Primarily a subscription service to access remotely administered database with web browser
LAMS	Mark McKie www.hgu.mrc.ac.uk/Softdata/LAMS	Windows	Freeware for basic colony management
MouSeek	Caleb F. Davis http://Mickey.utm.edu/main/databases.html	Macintosh Windows	Freeware Filemaker Pro template; modifiable

^aSeveral software options are listed ranging from extensive commercial applications (Colony and Scion) to freeware alternatives (LAMS and MouSeek). Colony has the most options including handheld devices for editing data on site and may be most useful for managing large colonies. Big Bench Mouse is new and appears best suited for medium and small colonies. Progeny requires the user to define the relational database. This increases both flexibility and complexity. LAMS and MouSeek are inexpensive alternatives for medium and small colonies. Both are freeware, but MouSeek requires the purchase of Filemaker Pro. The Web site cited for MouSeek also provides links to two other free Filemaker Pro colony management templates. Most of the databases can be run on a network allowing multiple users to access data simultaneously.

Table 23.8.4 Mouse Data^a

Field	Type	Description
Unique ID	Number	A serial number that uniquely identifies the record; usually created automatically
Tag/ID	Text	The mouse tag number
Date of death	Date	Date of death
Cause of death	Text	Used to differentiate natural deaths from experimental deaths or those for colony culling
Mating number	Number	Unique mating number from the mating table
Genotype	Text	Genotype
Notes	Text	For any non-standard annotations
Sex	Text	Sex of the mouse

^aMother, father, strain, date of birth, and category are all defined in the corresponding mating record (Table 23.8.5).

Table 23.8.5 Mating Data^a

Field	Type	Description
Mating number	Number	Unique serial number to identify the mating record usually created automatically
Start date	Date	Date the mating was begun
Date of birth	Date	Date of birth of litter
Category	Text	Differentiates multiple targeted or transgenic lines
Strain	Text	The background strain of the mouse
Mother	Number	ID that uniquely identifies the mother—the unique number from the mouse table
Father	Number	ID that uniquely identifies the father—the unique number from the mouse table
Notes	Text	For any non-standard annotations
No. of Males	Number	Number of male offspring
No. of Females	Number	Number of female offspring

^aMouse (Table 23.8.4) and mating data tables list a minimum set of data that should be recorded in colony management software. With this basic data individual animals and their mating behavior can be tracked and genealogies can be reconstructed. A simple database in Filemaker Pro can be constructed from scratch by creating these two tables and linking them using the mating number. Each mouse has a record in the mouse table. Each mating, whether productive or not, has a record in the mating table.

LITERATURE CITED

- Festing, M.F.W., Simpson, E.M., Davisson, M.T., and Mobraaten, L.E. 1999. Revised nomenclature for strain 129 mice. *Mammalian Genome*, 10:836.
- Hogan, B., Beddington, R., Constantini, F., and Lacy, E. 1994. *Manipulating the Mouse Embryo: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Kolesnikov, Y., Jain, S., Wilson, R., and Pasternak, G.W. 1998. Lack of morphine tolerance in 129/SvEv mice: Evidence for a NMDA receptor defect. *J. Pharmacol. Exp. Ther.* 284:455-459.
- Lander, E.S. and Schork, N.J. 1994. Genetic dissection of complex traits. *Science*. 265:2037-2048.
- Silver, L.M. 1995. Creation of a congenic strain *In Mouse Genetics: Concepts and Applications*, Section 3.3.3.2 Oxford University Press, New York.
- Simpson, E.M., Linder, C.C., Sargent, E.E., Davisson, M.T., Mobraaten, L.E., and Sharp, J.J. 1997. Genetic variation among 129 substrains and its importance for targeted mutagenesis in mice. *Nature Genet.* 16:19-27.
- Threadgill, D.W., Yee, D., Matin, A., Nadeau, J.H., and Magnuson, T. 1997. Genealogy of the 129

inbred strains: 129SvJ is a contaminated inbred strain. *Mammalian Genome* 8:390-393.

Wakeland, E., Morel, L., Achey, K., Yui, M., and Longmate, J. 1997. Speed congenics: A classic technique in the fast lane (relatively speaking). *Immunol. Today* 18:472-477.

KEY REFERENCES

Banbury Conference on Genetic Background in Mice. 1997. Mutant mice and neuroscience: Recommendations concerning genetic background. *Neuron* 19:755-759.

This viewpoint article reviews many of the concerns regarding the effects of genetic heterogeneity on experimental interpretation. Suggestions for standardized strain backgrounds and mating strategies are presented.

Silver, L.M. 1995. Mouse genetics: Concepts and applications. Oxford University Press, New York.

This is a useful and thorough review of the house mouse and its use in genetics. The text is available on line at <http://www.princeton.edu/~lsilver/book/MGcontents.html>.

INTERNET RESOURCES

<http://www.biosupplynet.com>

Search this Web site for animal husbandry related items to obtain a current list of suppliers.

Ear tags and applicators

<http://www.jorvet.com>

Web site for Jorgensen Laboratories, Inc., supplier for ear tags and applicators (Tel., 800-525-5614).

<http://www.nationalband.com>

Web site for National Band and Tag Company (Tel., 859-261-2035).

Electronic transponders

<http://www.bmds.com>

Web site for BioMedic Data Systems, Inc., supplier of electronic transponders (Tel. 302-628-4100).

<http://www.minimitter.com>

Web site for MiniMitter Co., Inc. (Tel., 800-685-2999).

<http://www.stoeltingco.com>

Web site for Stoelting Co. (Tel., 630-860-9700).

Software

<http://www.filemaker.com>

Web site for Filemaker, Inc., supplier of Filemaker Pro Database Software (Tel., 800-325-2747).

Mice

<http://www.criver.com>

Web site for Charles River Laboratories, supplier of numerous mouse strains (Tel., 978-658-6000).

<http://www.taconic.com>

Web site for Taconic Farms, supplier of numerous mouse strains (Tel., 518-537-6208).

<http://www.jax.org>

Web site for The Jackson Laboratory, supplier of numerous mouse strains (Tel., 207-288-6000).

Contributed by David A. Conner
Harvard Medical School
Cambridge, Massachusetts

Transgenic Mouse Production By Zygote Injection

Transgenesis by DNA injection into fertilized eggs (zygotes) is the most common experimental method used to modify the germline of mice with specific DNA constructs. Traditionally, the resultant mice have been used for the characterization of gene-specific phenotypes and for the characterization of gene regulatory elements that define tissue and developmental-specific patterns of expression. Transgenes have also been used to rescue mutant phenotypes, to create cell lineage markers, and to ablate specific cell lineages. In addition to gain-of-function transgenes, it is possible to produce transgenes that reduce the function of specific gene products or transgenes that have inducible expression.

Unlike animals created by homologous recombination in embryonic stem (ES) cells followed by blastocyst-mediated transgenesis, every traditional transgenic founder is unique. The site of DNA integration into the genome is usually random and the number of integrated copies cannot be completely controlled. This necessitates the generation of several independent founders to ensure that the phenotype is dependent on the transgene and not a result of alterations in the function of an endogenous gene. Regardless of this caveat, transgenic mice generated by DNA microinjection represent a powerful method to study the role of specific genes in a physiological setting.

The standard protocol involves the injection of DNA into the pronucleus of fertilized eggs (0.5-day-post-coital embryos). These injected eggs are then implanted into the oviduct of a pseudopregnant foster mother. The injected DNA integrates randomly into the genome in a fraction of the implanted embryos, typically as head-to-tail concatamers of the original linearized construct. The DNA usually integrates before the first cell division; however, the integration can occur later and result in mosaic founders. In either case, founders are mated to demonstrate germline transmission and to establish a transgenic line.

This unit presents protocols for the isolation of fertilized eggs (see Basic Protocol 1), the injection of DNA into fertilized eggs (see Basic Protocol 2), and the implantation of these injected eggs into pseudopregnant foster mothers (see Basic Protocol 3). In addition, methods for preparing plasmid DNA (see Support Protocol 1) and bacterial artificial chromosome (BAC) DNA (see Support Protocol 2) suitable for microinjection are provided. The types of surgical instruments and the preparation of equipment for embryo transfer and egg injection are also described (see Support Protocol 3).

The basic protocols are technically difficult, requiring proficiency with zygote isolation, zygote injection, and embryo reimplantation. Each of these techniques requires practice. Because of the steep learning curve and the expense of the equipment, the procedure should not be undertaken without a serious commitment (also see Commentary).

NOTE: All protocols using live animals must first be reviewed and approved by an Institutional Animal Care and Use Committee (IACUC) and must follow officially approved procedures for the care and use of laboratory animals. Animals should be handled following institutional guidelines to minimize undue pain and suffering. Pay strict attention to institutional guidelines pertaining to survival surgery.

NOTE: All mice used should be housed in a specific-pathogen-free (SPF) or virus-antibody free (VAF) facility.

Contributed by David A. Conner

Current Protocols in Molecular Biology (2004) 23.9.1-23.9.29
Copyright © 2004 by John Wiley & Sons, Inc.

NOTE: All solutions and equipment coming into contact with zygotes must be sterile. Additionally, proper sterile technique must be used when isolating, manipulating, or implanting fertilized eggs.

ISOLATION OF FERTILIZED EGGS

The method for generating transgenic mice by DNA microinjection uses 0.5-day-old fertilized eggs. At the stage of isolation, fertilized eggs are present in a clump of follicular cells (cumulus mass) in an enlarged region of the oviduct (ampulla) near the infundibulum (ovarian end of the oviduct). The fertilized eggs must be removed from the oviduct and separated from the mass of follicular cells so that they can be manipulated as single eggs for microinjection. Many different strains of mice may be used as egg donors. Factors that govern the choice of strains include the quantity and quality of eggs that can be obtained, and the effect the background strain may have on the transgene phenotype. As a general rule, outbred or hybrid strains yield more eggs of better quality than inbred strains. The B6D2F1 (first generation cross between C57BL/6 and DBA) strain is a commonly used hybrid strain. A few inbred strains, such as FVB, are particularly suited for egg donors because they will produce large numbers of good quality eggs.

Materials

- 50 U/ml PMS (see recipe)
- 3- to 4-week-old FVB female mice (Taconic or The Jackson Laboratory)
- 50 U/ml HCG
- 8-week- to 1-year-old fertile stud male mice, individually caged
- 95% ethanol
- M2 medium (Sigma or Specialty Media) containing 100 U/ml penicillin and 100 µg/ml streptomycin (add antibiotics just before use)
- 1 mg/ml hyaluronidase (Sigma H-3884) in M2 medium
- M16 microdrop cultures (see recipe)
- 1-ml, $\frac{3}{8}$ -in. tuberculin syringes with 26-G needles
- Surgical equipment (see Support Protocol 3):
 - Scissors
 - Fine forceps
 - Iris scissors
- 35 × 10-mm and 60 × 10-mm petri dishes
- Dissecting microscope: stereomicroscope with fiber-optic light source, 0.8× to 4× zoom, and a stand that allows illumination from above and below
- Embryo transfer pipet (see Support Protocol 3)
- 37°C, 5% CO₂ incubator
- Additional reagents and equipment for preparing foster mothers (UNIT 23.7)

Initiate superovulation

1. Initiate superovulation by injecting 5 units of 50 U/ml PMS i.p. into 3- to 4-week-old female FVB mice using a 1-ml, $\frac{3}{8}$ -in. syringe and 26-G needle, between 11 a.m. and 2 p.m., 3 days prior to egg isolation (Table 23.9.1).

Egg donors are superovulated to synchronize estrus and produce the greatest egg yields. Young female mice that have not yet cycled usually give the greatest embryo yields. The timing of injection is important: it must be coordinated with the light/dark cycle to ensure that the females ovulate in response to the exogenous hormones prior to mating. A typical lighting cycle consists of 10 to 12 hr dark beginning at 6 to 7 p.m. and extending to 5 to 6 a.m. This represents the range of values that works well with a standard workday. For a specific animal room, the timing and duration of the cycle should be held constant. Eight matings should generate enough eggs for 1 day of injection for fertile strains such as FVB

Table 23.9.1 Timeline for Zygote Microinjection

Day	Time	Action	Comments
0	11 a.m. to 2 p.m.	Initiate superovulation in egg donors by PMS injection	Procedure begins 3 days before the actual day of injection with the initiation of superovulation of the donor mice. Mating is considered to occur in the middle of the dark cycle.
1	—	No action	—
2	46 to 48 hr after PMS injection	Inject donors with HCG and set up foster matings	Donors mate at the transition of the third and fourth 24-hr period. By the middle of the injection day, they are post-coital day 0.5. Fosters are set up to mate the day before injection at the same time the donors are placed together. Fosters will also be p.c. day 0.5 on the day they will be used for reimplantation.
3	Morning	Check donors and fosters for plugs	Donors and fosters should be checked for plugs on the morning of the day of injection.
	Morning	Isolate fertilized eggs	Fertilized eggs should be harvested early in most cases.
	Afternoon	Inject zygotes; reimplant	Injections are performed through the rest of the day followed by reimplantation

or most F1 hybrids. As a general rule, mouse strains that breed well and produce large litters will be good egg donors, while strains that are poor breeders will produce few good quality eggs.

- Forty-six to 48 hr after the PMS injection (Table 23.9.1), i.p. inject each female with 5 units of 50 U/ml HCG.

Mate superovulated females and foster mothers with stud males

- Place single superovulated females with individually caged 8-week- to 1-year-old fertile stud males. On the same day, set up foster mother matings (*UNIT 23.7*).

Stud males are usually the same strain as the donor females. Mice should mate in the middle of the next dark cycle. If injections are planned on consecutive days, it is a good idea to maintain a large enough colony of stud males so that no male is mated on consecutive days, otherwise fertilized egg yields will usually drop on the second day. A mating record for each stud male should be maintained. Males that do not plug females regularly should be replaced.

Note that foster mothers for oviduct transfers are used on the day the plug is found (0.5 day p.c.). When preparing foster mothers, set up enough matings with vasectomized males to yield at least three plugged females for each day of injection.

- The next morning (Table 23.9.1), separate fertilized females and foster mothers from the males.
- Pick up each female by the tail and place her on a wire cage top, allowing her to grasp the wires with her front paws. Observe the vaginal opening for a plug, which is hard white or yellow mucus, by raising the hind legs off the cage top while lifting the tail. If necessary, probe the vagina gently with a disposable pipet tip.

The plug usually remains for 12 hr or longer. Females should be checked in the morning to avoid missing plugs.

Females plugged the morning after the mating was initiated are considered to be 0.5 day post-coital (p.c.). Fertilized eggs used for DNA injections are considered to 0.5-days-p.c. embryos. Most, but not all, of the superovulated females should plug when using highly fertile strains. The plugging rate may be significantly lower with strains that are poor breeders. Don't forget to check for plugged foster mothers at the same time. Check plugs as early in the day as possible, as plugs may be lost as the day progresses.

A disposable pipet tip such as a Gilson-style P-200 is recommended, as these tips are cheap, clean, and relatively blunt.

Isolate reproductive tract

6. On p.c. day 0.5, sacrifice donor females by an approved method of euthanasia. Place the mouse on its back and wet the abdomen with 95% ethanol.

Work through the protocol (steps 6 to 16) with one mouse at a time until comfortable with the procedure. When the mice are housed under standard light/dark cycles, eggs should be harvested before noon. Later in the day, the cumulus mass may begin to disperse and the eggs may not be grouped together in one region of the oviduct.

7. Tear the skin away from the abdomen and cut through the peritoneal wall with scissors.

The skin can also be cut, but with some practice, the skin can be torn away quickly.

Refer to Support Protocol 3 for a description of useful surgical instruments.

8. Remove the ovaries, oviducts, and uterine horns as a single unit in the following manner (Fig. 23.9.1A and B):

- a. First, grasp an ovary by its associated fat pad with forceps.
- b. Gently lift the ovary and uterine horn.
- c. Cut the attached tissue as necessary to free the uterine horn.
- d. Cut the cervix to release the other side.
- e. Place the tissue on a clean paper towel.

The ovary is easily found at the base of the kidney.

Isolate oviducts

9. Cut the uterine horn near the oviduct/uterus junction (Fig. 23.9.1B). Discard the uterus. Collect the ovaries and oviducts in a drop of M2 medium containing 1000 U/ml penicillin and 100 μ g/ml streptomycin in a 35 \times 10–mm petri dish.

The volume of medium is not critical. The medium is just used to prevent the tissue from drying out. Typically a drop of several hundred microliters is appropriate for sixteen to twenty four ovaries obtained from eight to twelve mice. As the procedure becomes more familiar, the process can be streamlined by removing the ovary and oviducts directly without first removing them as a unit with the uterine horns.

10. Place the dish under a dissecting microscope. Under low power, cut the ovary away from the oviduct as follows (see Fig. 23.9.2A and B):
 - a. Grasp the uterine end of the oviduct with fine forceps.
 - b. Place the open blades of a pair of iris scissors around the oviducts.
 - c. Close the blades slightly, but not enough to cut the tissue.

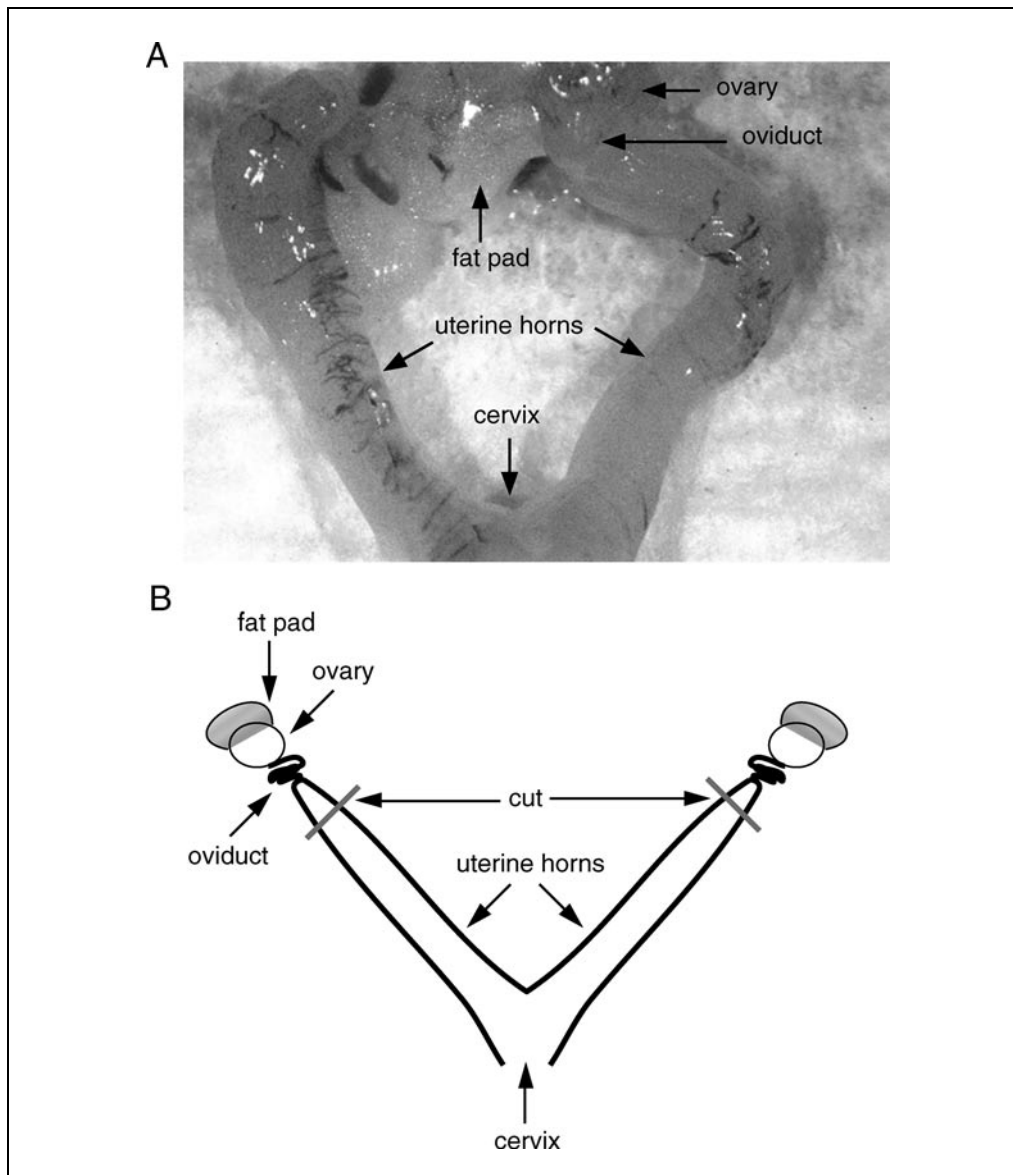


Figure 23.9.1 Uterine horns, oviduct, and ovaries during various stages of fertilized egg isolation. (A) Photograph of the uterine horns with oviduct, ovary, and fat pad still attached. Excision as a unit may be easier for the beginner than removal of the oviducts and ovaries alone. Note that the oviduct and ovary on the left side has been removed. (B) Diagram of similar structure seen in (A), with suggested cut sites to separate the oviducts and ovaries from the uterus. Grasp the ovarian fat pad with forceps to move the oviduct and ovary.

- d. Slide the blades gently toward the ovaries.
 - The blades will catch on the bulb of the ovary.*
 - e. Cut the bursa to separate the ovary from the oviduct.
 - f. Discard the ovary.
11. Place the oviduct in a fresh drop of several hundred microliters of M2 medium on a fresh 35 × 10–mm plate. Combine oviducts from eight to twelve mice.

Isolate cumulus masses

12. Release the cumulus mass by tearing the ampulla, an enlarged region of the oviduct near the ovarian end of the oviducts (Fig. 23.9.3), open with fine forceps under a dissecting microscope.

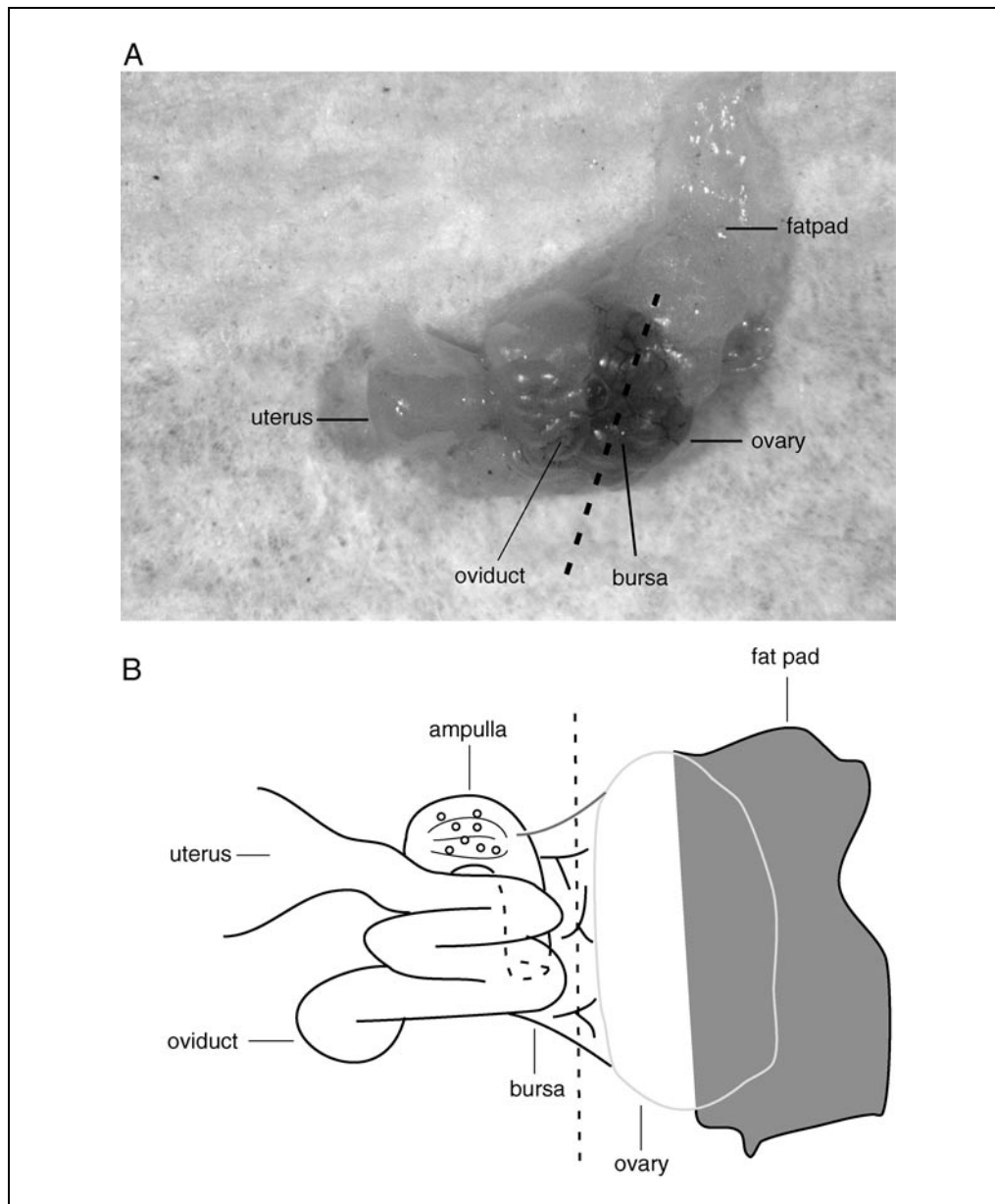


Figure 23.9.2 Oviduct and ovary after removal of the uterine horn. Note the location of the bursa, a thin transparent membrane that covers the gap between the oviduct and ovary. **(A)** The line indicates the cut site to separate the oviduct from the ovaries. **(B)** Diagram of the image seen in (A). Note the swollen region of the oviduct, the ampulla, near the ovarian end of the oviduct, which is often hidden under the folds of the oviducts. The dotted line indicates where to cut to separate the oviducts from the ovary.

At this stage of development, the fertilized eggs are present in a clump of follicular cells, known as the cumulus mass, in the ampulla. The clump is visible through the transparent wall of the ampulla.

13. After tearing open the ampulla and releasing the cumulus mass, remove the oviduct from the drop of medium. Continue the procedure until all the oviducts have been processed and the cumulus masses from eight to twelve mice are collected in a single drop of medium.

A typical cumulus mass is shown in Figure 23.9.4A.

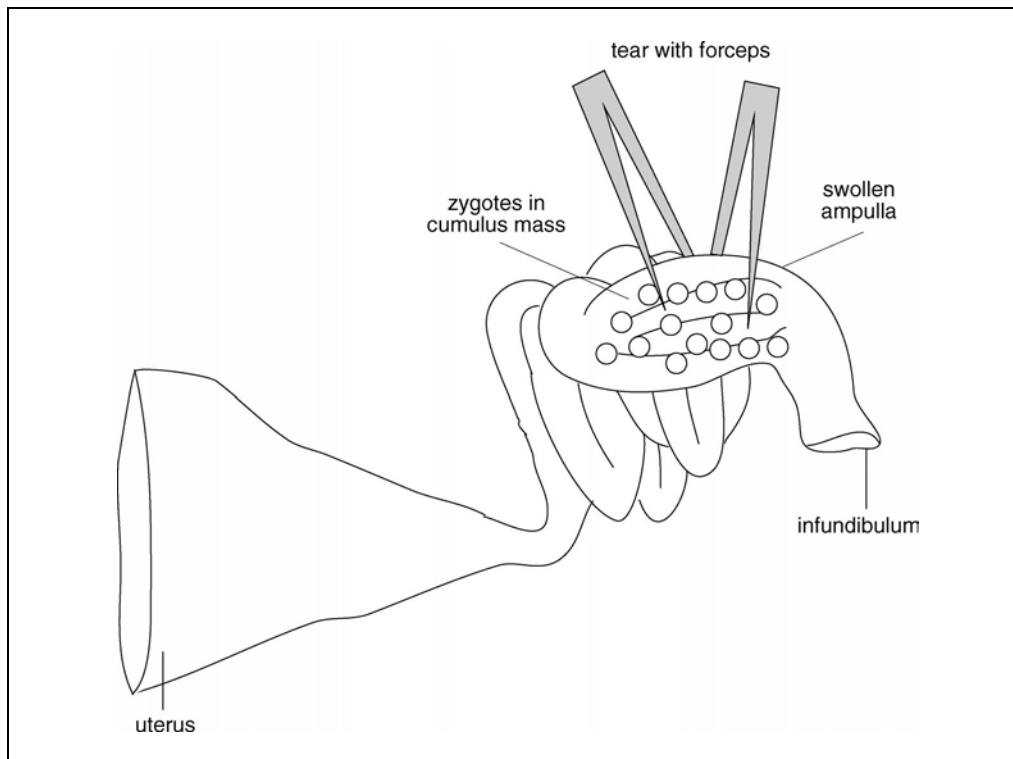


Figure 23.9.3 Diagram of the isolated oviducts. Tear the swollen ampulla of the oviduct with fine forceps as shown to release the cumulus mass containing fertilized eggs and follicle cells.

Isolate fertilized eggs

14. Add an approximately equal volume of 1 mg/ml hyaluronidase in M2 to the drop of medium containing the eggs. Gently mix the solution by swirling the pipet tip around. Place the plate under a dissecting scope and observe the cumulus mass periodically over the next few minutes.

This will release the eggs by digesting the cumulus mass.

The exact concentration of hyaluronidase is not critical. In fact, a few crystals of powdered hyaluronidase can be sprinkled on the top of the medium drop and mixed in by gentle swirling with a pipet tip. If this method is used, the digestion should be watched closely under the dissecting scope and the solution diluted and eggs removed (step 15) as soon as the cumulus mass disperses. As the digestion progresses, gently swirl a pipet tip or an embryo transfer pipet through the solution to disperse the cumulus cells.

15. After digestion is complete, add an equal volume of M2 to further dilute the hyaluronidase solution. Using an embryo transfer pipet, pick up individual eggs and transfer them to a fresh drop of M2.

Figure 23.9.4B shows the isolated zygotes.

16. Transfer the eggs through several drops of M2 to dilute the remaining hyaluronidase. Then place the eggs in a microdrop of M16 microdrop culture. Store in a 37°C, 5% CO₂ incubator for the remainder of the procedure.

There isn't a specific time limit as to how long the eggs can be stored. The eggs are harvested and then injected over the course of the afternoon. It should be a continuous procedure completed within a single day.

If a CO₂ incubator is not available, the isolated eggs can be kept in an M2 drop under oil on a slide warmer set to ~37°C. Although not ideal, these conditions are usually adequate for eggs that will be injected immediately and then reimplanted into foster mothers on the same day.

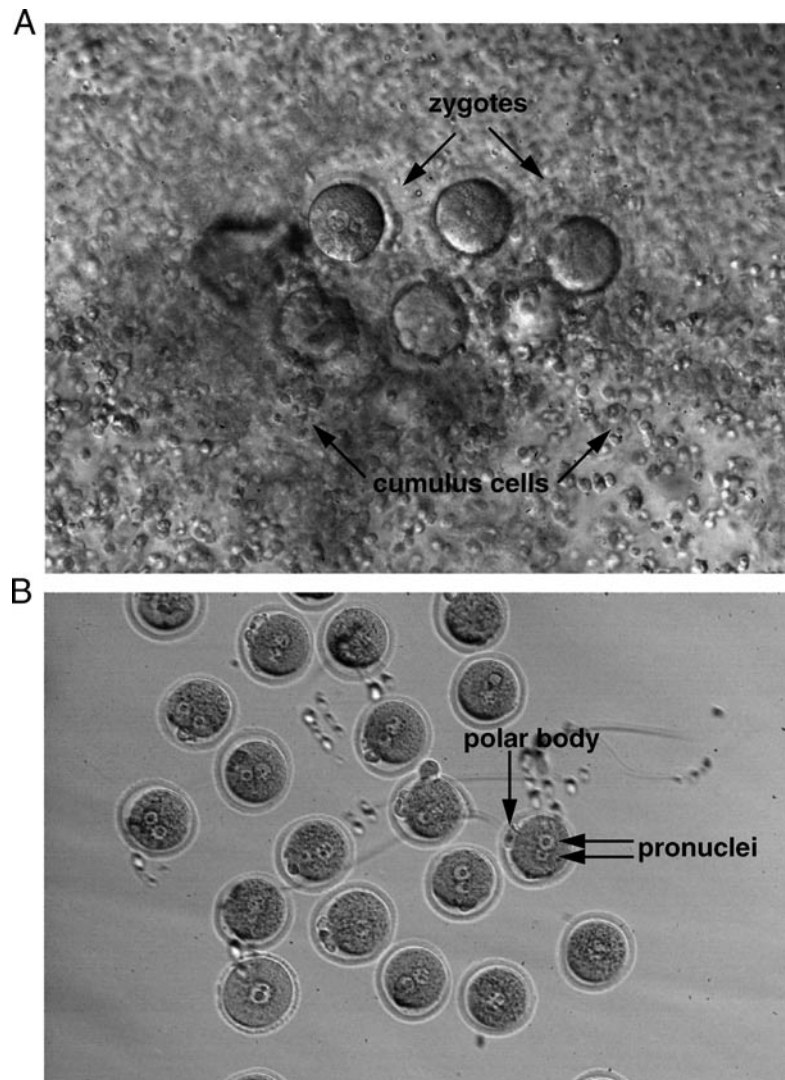


Figure 23.9.4 Zygotes in the cumulus mass and after final isolation. **(A)** Photograph of the cumulus mass immediately after release from the swollen ampulla. Multiple fertilized eggs are visible in a mass of cumulus cells. **(B)** Photograph of zygotes after hyaluronidase treatment and subsequent transfer to a new microdrop. This is a mixture of fertilized and unfertilized eggs. Note the eggs with two pronuclei indicating fertilization. These are the eggs that are candidates for injection. The polar body is also visible in most eggs.

**BASIC
PROTOCOL 2**

INJECTION OF ZYGOTES

Conceptually, zygote injection is simple. Fertilized eggs are placed in a microdrop under a microscope. Individual eggs are gently grasped by aspiration with a pipet. A second pipet filled with DNA in solution is inserted into the pronucleus of the egg. The DNA solution is injected into the pronucleus by positive pressure and the pipet tip is rapidly removed. With this procedure, 100 to 200 eggs can be injected in 1 to 2 hr.

Materials

- M2 medium (Sigma or Specialty Media) containing 100 U/ml penicillin and 100 μ g/ml streptomycin (add antibiotics just before use)
- Embryo-tested mineral oil (Sigma)
- Zygotes (see Basic Protocol 1)

**Transgenic
Mouse
Production by
Zygote Injection**

23.9.8

M16 medium (Sigma or Specialty Media) containing 100 U/ml penicillin and 100 µg/ml streptomycin
Culture slide with a single ~18-mm-diameter 0.8-mm deep depression (i.e., depression slide, VWR)
Injection apparatus with holding pipet and injection pipet loaded with DNA solution (see Support Protocol 3)
Embryo transfer pipet (see Support Protocol 3)
37°C, 5% CO₂ incubator
Transgene in injection buffer (see Support Protocols 1 or 2)

Prepare injection apparatus

1. Prepare an injection chamber consisting of ~50 µl M2 medium containing 100 U/ml penicillin and 100 µg/ml streptomycin surrounded by embryo-tested mineral oil (Fig. 23.9.5A) on a culture slide with a single ~18-mm-diameter, 0.8-mm-deep depression.

A variety of injection chambers can be used. For example, a drop culture on a petri dish can also work; however, the sides of the petri dish may limit the angle at which the pipets can be placed. A depression slide has many advantages. It has good optical qualities and the pipet tips can be aligned at a very oblique angle without having to bend the ends.

2. Place the injection chamber on the stage of the injection apparatus. Focus on the drop under low power and then place zygotes at the 12 o'clock position in the injection chamber. Transfer only as many zygotes as can be comfortably injected in 20 to 30 min.

After they have been injected, the zygotes are transferred back to a drop culture of M16 pre-equilibrated in a CO₂ incubator (step 19).

The zygotes are placed in a specific region of the medium drop to separate injected and uninjected eggs. Injected eggs are moved to the 6 o'clock position (step 18). Obviously, any preferred organization of injected and uninjected eggs is acceptable as long as they can be separated.

Align holding pipet

3. Align the holding pipet with the tip above the injection chamber at an oblique angle to the microscope stage.
4. Using the course manipulator controls, bring the tip down to the top of the medium drop. Position the tip as illustrated in Figure 23.9.5B: in the 6 o'clock position and spanning the field of view.

In this orientation, when the tip is moved toward the 12 o'clock position, the pipet can be easily "found" as it crosses the field of view while looking through the eyepiece.

5. Look through the objective under low power while focused on the eggs and move the pipet across the field of view using course adjusters.

The tip will appear as a blur because it is out of focus, but it is easy to catch the motion.

6. Once the pipet is in the field of view, focus on it, and center the tip in the field. Then refocus on the zygotes and lower the pipet until the tip is in focus.

There are many ways to align the tips. The key is to be aware of the pipet tip and the bottom of the injection chamber. Don't lower the tip in the medium unless the tip is definitely in the field of view and the microscope is focused at the bottom of the microdrop. This will prevent breakage.

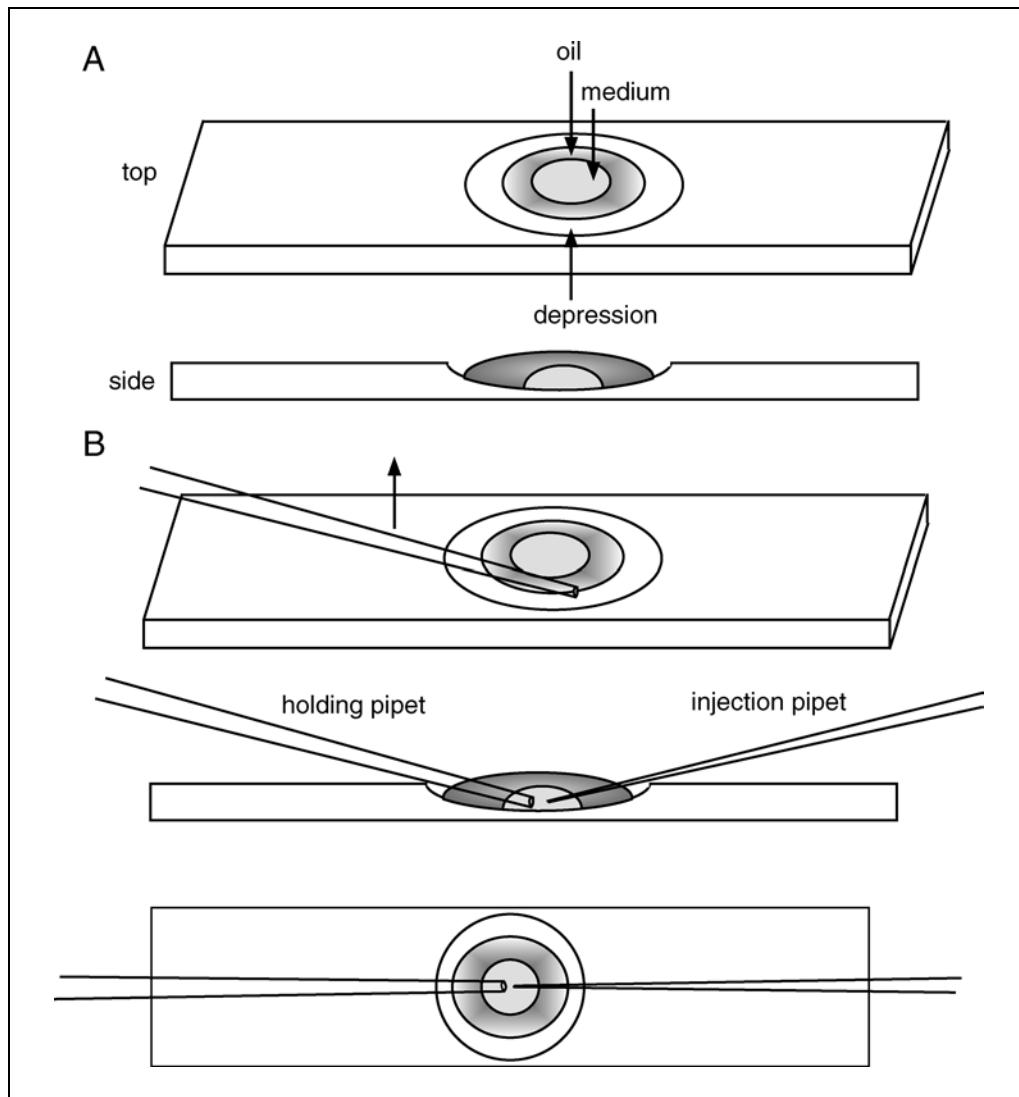


Figure 23.9.5 Injection chamber and pipet alignment. **(A)** Diagram illustrating top and side views of an injection chamber using a depression slide. **(B)** The upper diagram illustrates one method for bringing a pipet tip into alignment. Under low magnification, use the course controls to place the pipet in the 6 o'clock position with the top just at the surface of the microdrop. While looking through the microscope, focused on the zygotes in the bottom of the depression, move the pipet tip toward the 12 o'clock position. (The pipet will appear as an out-of-focus blur as it crosses the field of view.) Stop when the tip is centered in the field. Always know where the bottom of the chamber and the pipet tip are when moving the tip up and down, to avoid breakage. The middle diagram illustrates the correct orientation of both pipets for injection viewed from the side. The angle of the tips should be as oblique as possible while allowing the tip to reach the bottom of the chamber without the barrel of the pipet touching the rim of the chamber. The bottom diagram illustrates the correct orientation of both pipets for injection viewed from the top. Note that the pipet shafts are both oriented along the same straight line.

Align injection pipet

7. Repeat the procedure with the injection pipet. If the injection tip is open at the time the pipets are first aligned, make sure that the tip is under slight positive pressure before it is moved into the medium. Make sure the shafts of both pipets are oriented in the same straight line.

The final alignment is illustrated in the middle and bottom panels of Figure 23.9.5B.

If positive pressure is not applied before the tip enters the medium, it may backfill due to capillary action, possibly resulting in a full day spent injecting medium into zygotes instead

of DNA. Characteristics of the injection pipet are discussed in more detail elsewhere (see Support Protocol 3).

Prepare apparatus for injection

8. Center the pipet tips in the field of view. Focus on the zygotes and bring both tips into focus to ensure that all components are in the same vertical plane. Slowly backfill the holding pipet with a small volume of medium.
9. Change to the 10× objective and center the tips of both pipets again. Repeat the procedure after switching to the 20× objective and again with the 40× objective.

Whenever the stage is moved, the pipet tips should be raised to avoid breakage. The tips can be raised so they are just out of focus and appear as a hazy outline. Also make sure the tips are raised when they are near the zygotes and the stage is moved. Otherwise, the zygotes can be inadvertently swept into the oil phase and destroyed. Note that “sweeping” with the pipet can be used intentionally to reposition groups of eggs in the microdrop. With practice, adjusting the vertical position of the tips will become automatic. If a concave surface such as a depression slide is used for the injection chamber, take particular care with the location of the pipet tips when moving the stage. If the tips are close to the bottom of the depression slide and the stage is moved significantly, they may be broken on the higher rim of the chamber.

10. Check the flow of the injection pipet by placing the tip near an egg or small piece of debris and activating the injector.

The egg or debris should move slightly when exposed to the increased flow of DNA from the tip.

Grab egg with holding pipet

11. Look at the eggs under intermediate magnification (100× or 200×). Two pronuclei should be visible in the fertilized eggs (Fig. 23.9.4B).
12. Select a fertilized egg, and orient it so that the larger pronucleus is at the 3 o'clock position. Focus on the pronucleus and move the holding pipet up or down until it is also in focus.

The eggs can be moved and oriented in a variety of ways. They can be “batted,” rolled, or blown with the holding pipet. The shaft of the injection pipet can also be used to nudge or roll eggs; however, use of the tip of the injection pipet may result in clogging or damage and should be avoided. If a significant proportion of the eggs are not fertilized, it may be useful to separate unfertilized eggs from fertilized eggs by moving them to a different location on the drop. Note that there is often not much difference in the size of the pronuclei. Just pick the larger target if you can.

13. Bring the holding pipet into contact with the egg and gently increase the vacuum to grab it (Fig. 23.9.6A and B). Lift the egg slightly off the surface.

Inject the zygotes

14. Under high magnification (400×), refocus on the end of the holding pipet and the pronucleus. Bring the injection tip into focus at the 3 o'clock position.

When all three components are in focus, they should be in the same plane and ready for injection (Fig. 23.9.6A).

15. Bring the injection tip into light contact with the zona pellucida directly adjacent to the pronucleus. Check again to ensure the tip and pronucleus are in the same plane. Firmly push the tip into the egg and pronucleus in a firm, rapid, and continuous manner.

The path of the tip should be straight. Realignment of the tip while inside the egg will usually damage the egg. Avoid the nucleoli, as this material will often stick to the injection

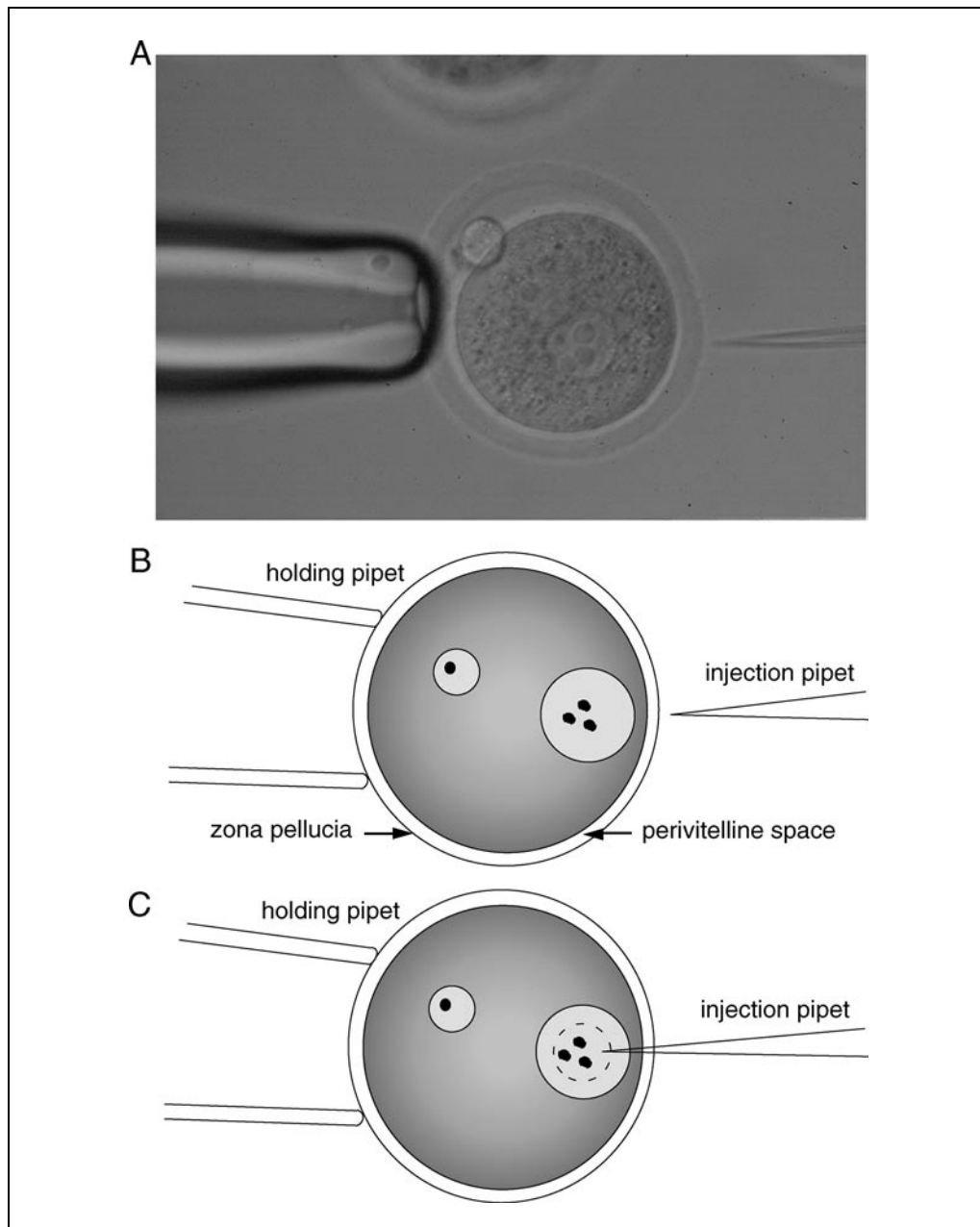


Figure 23.9.6 Alignment and injection. (A) Photograph of zygote attached to the holding pipet. The pronucleus that will be injected, the holding pipet, and the injection tip are all in focus. (B) Diagram of the same arrangement. Note the orientation of the zygote is such that the pronucleus that will be injected is adjacent to the injection pipet. (C) Diagram illustrating the injection tip after penetration. The pronucleus should swell visibly, but not lyse.

tip, resulting in clogging of the pipet and damage to the egg as the tip is pulled out with the adherent material. The egg should not roll. Rolling of the egg during injection indicates that the holding pipet, egg, and injection pipet were not properly aligned in the same plane.

16. Activate the injector. The pronucleus should swell noticeably (Fig. 23.9.6C) but not rupture.

The method of DNA injection will vary depending on the apparatus used. This is discussed in more detail in Support Protocol 3. The volume of DNA injected can be adjusted based on the swelling of the pronucleus. Use the visual feedback of the swelling pronucleus even when working with an automated injection system. Don't just assume that the equipment

is working properly. The formation of a bubble around the injection tip indicates that the egg membrane was not penetrated. Pull the tip back and try again.

17. Remove the injection tip in one quick, continuous motion when sufficient DNA has been injected.
18. Move the injected egg to the 6 o'clock position in the microdrop by moving the stage. Return to the 12 o'clock position and repeat the injection procedure.
19. After injecting all the fertilized eggs in the chamber, remove the injected eggs using an embryo transfer pipet and place them in a microdrop of M16 containing 100 U/ml penicillin and 100 µg/ml streptomycin. Transfer to a 37°C, 5% CO₂ incubator.

Typically, eggs are injected in batches that can be completed in 20 to 30 min and then collected in the incubator. In this way, several hundred eggs can be injected in the course of an afternoon without any individual egg spending more than ~30 min out of the incubator; however, the exact time that eggs spend in or out of the incubator is not critical. Fertilized eggs from most of the standard strains used for transgenics are fairly hardy. Injected eggs that have been severely damaged by the procedure will become obvious within the first 30 min, while surviving eggs will remain viable in the incubator for an entire afternoon.

20. Remove the remaining unfertilized or damaged eggs and discard.
21. Add more eggs to the injection chamber. Repeat the experiment, remembering to work with only as many eggs as can be injected in 20 to 30 min.

Expect to be able to inject 100 to 200 eggs in 1 to 2 hr with experience.

Monitor the number of eggs that survive injection within each group. If the number changes dramatically between groups, it may indicate that the injection tip is damaged or dirty and needs to be changed. Replace the injection pipet if there is debris stuck on the end or if the DNA is not flowing properly. Eggs that have died will look significantly different. The cytoplasm will expand to the edge of the zona pellucida; there will be no perivitelline space (Fig. 23.9.6B). Dead eggs will seem less dense and have a tendency to float or roll around when the microdrop is moved. The percentage of eggs that survive is dependent on many factors. Expect 60% to 80% of the eggs to survive under normal circumstances.

TRANSFER OF OVIDUCTS TO FOSTER MOTHERS

Injected zygotes must be surgically implanted into foster mothers to allow embryos to develop to term. Single-cell zygotes are reimplanted into the oviducts of 0.5-day-p.c. pseudopregnant females. Approximately 20 days after reimplantation, pups will be born.

Materials

- M2 medium (Sigma or Specialty Media) containing 100 U/ml penicillin and 100 µg/ml streptomycin
- Injected zygotes (see Basic Protocol 2)
- Modeling clay (VWR, WL6852)
- 0.5-day-p.c. Pseudopregnant females
- 2.5% avertin (see recipe)
- 70% ethanol
- Analgesics per institutional requirements
- Embryo transfer pipets (Support Protocol 3)
- 1-ml, $\frac{3}{8}$ -in. tuberculin syringes with 26-G needles
- Surgical equipment and supplies (Support Protocol 3):
 - Electric clippers
 - Scissors
 - Toothed and blunt forceps
 - Iris scissors
 - Serrefine clamp

BASIC PROTOCOL 3

Manipulating
the Mouse
Genome

23.9.13

Sutures
Surgical staples
Dissecting microscope with overhead illumination
30-G needle (VWR)
Heating element (see Support Protocol 3)
Fresh mouse cages

NOTE: Follow institutional animal care guidelines for survival surgery.

NOTE: Trying both oviduct puncture and infundibular transfer is recommended. Use whichever is least difficult.

Prepare embryos

1. Prior to implantation, transfer embryos to a fresh drop of M2 containing 100 U/ml penicillin and 100 µg/ml streptomycin but without oil.

This helps to minimize the transfer of oil into the oviduct. Inspect each egg at this stage. Transfer only those eggs which still look healthy. Healthy zygotes should have a distinct perivitelline space (Fig. 23.9.6B).

2. Load 10 to 20 embryos into each embryo transfer pipet. Place the loaded transfer pipets on the bench top, supported by a strip of modeling clay.

Zygotes for oviduct transfer should be packed fairly tightly with a minimum of space between each egg. This will minimize the total volume of fluid transferred into the oviduct.

Prepare ovaries for implantation

3. Anesthetize a 0.5-day-p.c. pseudopregnant female with 0.017 ml/g of 2.5% Avertin i.p. via a 1-ml, $\frac{3}{8}$ -in. tuberculin syringe and 25-G needle.

An analgesic (e.g., Flunixin 2.5 mg/kg SQ, Ketoprofen 5 mg/kg SQ, or buprenorphine 0.05 mg/kg) may be administered at this stage if recommended by the institution.

4. Shave the hair around the incision site (Fig. 23.9.7A) with electric clippers. Swab the area with betadine followed by 70% ethanol.
5. Using scissors, make an incision ~1-cm long in the skin on the midline of the back (Fig. 23.9.7A).
6. Place the mouse under a dissecting scope with overhead illumination. Slide the opening down the right or left side of the mouse until the ovaries or ovarian fat pad are visible through the peritoneal wall.

In young females, the ovaries (pink) and ovarian fat pad (white) are visible through the peritoneal wall. As the females age, the ovaries are often obscured by fat. The ovarian fat pad is often more white than the obscuring fat, which is slightly yellow.

7. Grasping the peritoneal wall with toothed forceps, make a small incision (~0.5 cm) in the peritoneal wall, over the ovaries, using fine iris scissors. Avoid blood vessels.
8. Grasp the ovarian fat pad with blunt forceps and pull the ovaries and proximal end of the uterine horn out of the incision. Attach a small serrefine clamp to the ovarian fat pad and rest the clamp on the side of the animal to free the hands (Fig. 23.9.7B).
9. Orient the oviduct by gently rotating the serrefine clamp so that the region just proximal to the infundibulum is visible (Fig. 23.9.8A and B).

Look for the infundibulum, which is best observed under 20× to 25× magnification, to achieve the proper orientation. This can be difficult to find. The infundibulum is under the transparent bursa and often hidden by folds of the oviduct. Once the oviduct opening is

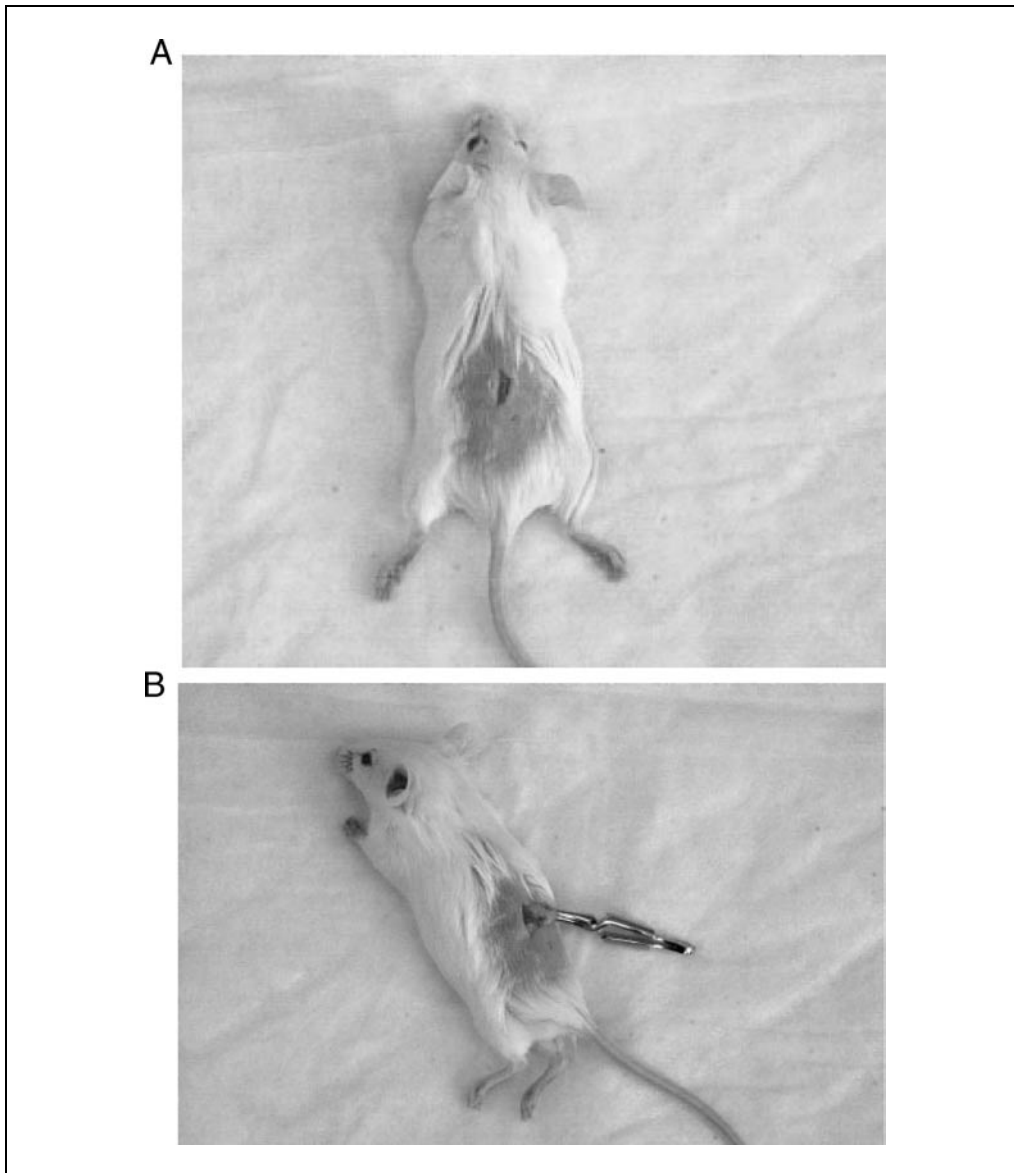


Figure 23.9.7 Oviduct Reimplantation. **(A)** Photograph showing the location of the central skin incision site. Note that the area around the site has been shaved. **(B)** Photograph showing the orientation of the mouse for reimplantation with an extruded ovary and fat pad held by a serrefine clamp.

found, follow it back past one bend to identify the site for a puncture transfer. This is the same region where the swollen ampulla is found in a superovulated female.

Implant injected zygotes

For oviduct puncture

- 10a. Grasp the region of the oviduct gently with fine forceps. Puncture the oviduct in the location indicated in Figure 23.9.8B by pushing a 30-G needle through the tissue and between the forceps tips, making sure it enters the oviduct lumen.

Minimize tissue damage with the needle and forceps. Observe the site of puncture as the needle is removed. The site is usually marked by a small indentation in the surface or by a small amount of bleeding. Observe how the tissue moves if the forceps are removed so that the opening can be found again to insert the transfer pipet.

- 11a. Insert the embryo transfer pipet into the opening and expel the embryos and air bubbles into the lumen.

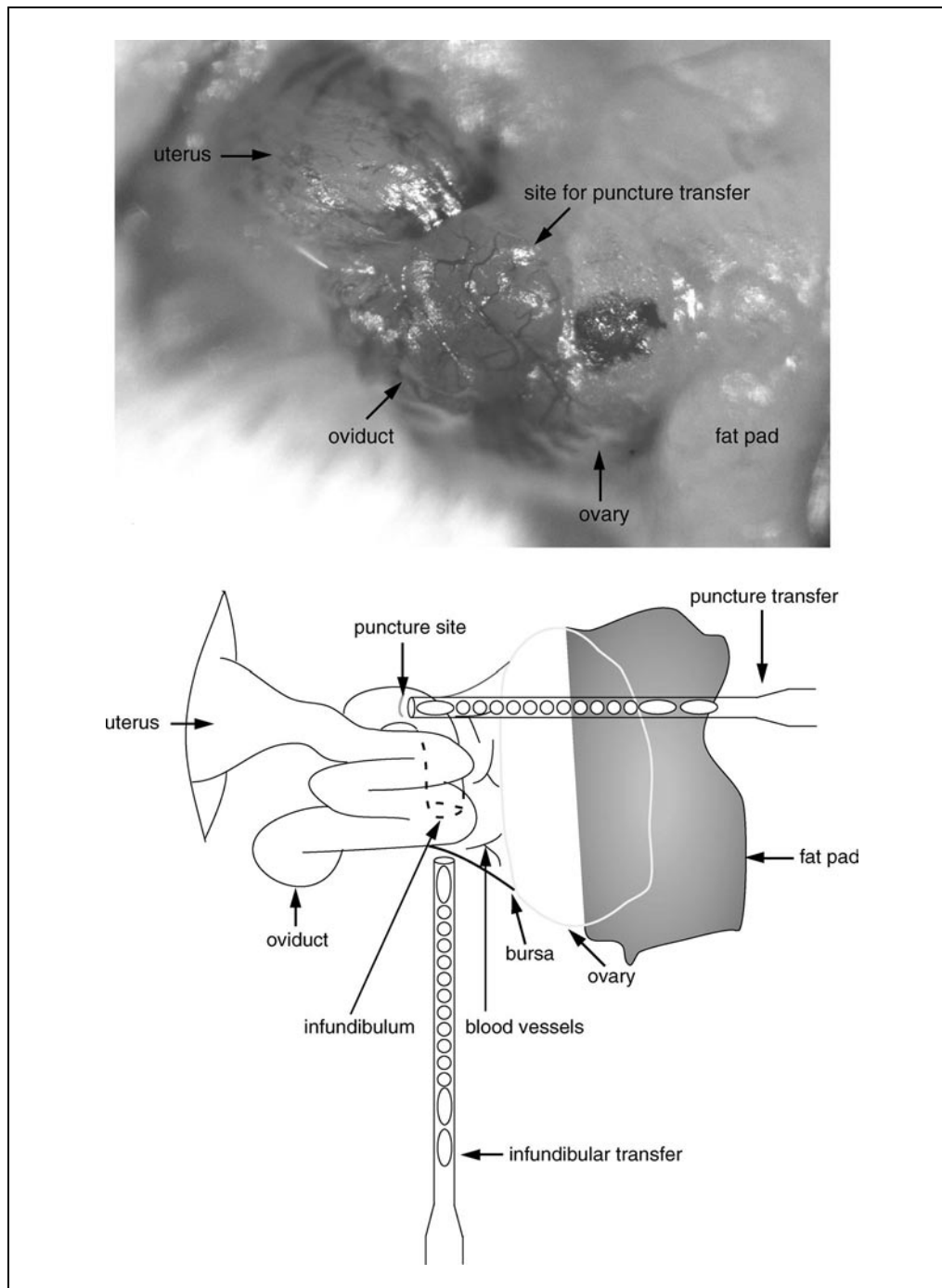


Figure 23.9.8 Reimplantation in the oviduct. (A) Photograph of the oviduct ready for reimplantation. (B) Diagram of the oviduct and alternative reimplantation sites. Note the location of the puncture site for puncture transfers and the infundibulum for infundibular transfers.

Use the bubbles in the transfer pipet as a marker to ensure that the eggs are expelled into the lumen. The bubbles also serve as an indicator of how much of the transfer pipet has been inserted into the lumen. The ejected bubbles are visible in the oviduct lumen and function as a positive indicator of a successful transfer. Avoid introducing oil into the lumen.

For infundibular transfer

- 10b. Locate the infundibulum underneath the transparent bursa and under the folds of the oviduct (Fig. 23.9.8B). Gently tear open the bursa over the infundibulum using fine forceps.

Do not tear the fine blood vessels in the bursa: localized bleeding will obscure the view.

- 11b. Place the loaded transfer pipet into the infundibulum and expel the air bubbles and eggs into the oviduct.
12. Gently reposition the uterus and ovary in the peritoneal cavity. Use one suture to close the peritoneal wall.
13. If desired, repeat the procedure on the other side.

Embryos do not need to be implanted on both sides. Several factors may affect this decision. Distributing the injected eggs in more foster mothers decreases the loss caused by a surgical death or the loss caused by poor maternal care of a litter; however, these are not common problems for an experienced investigator using outbred pseudopregnant foster mothers. Alternatively, implanting into both uterine horns reduces the number of pseudopregnant females required and thus reduces the size of the mouse colony necessary to generate the pseudopregnant females. The transfers can be completed more quickly because fewer animals need to be prepared for surgery.

14. Close the incision in the skin with surgical staples.

Staples are used rather than sutures because mice will often reopen an exterior wound that is closed with sutures.

Allow mouse to recover

15. Place the mouse on a heating element (i.e., slide warmer or heating pad) until she begins to show signs of recovery (10 to 30 min). Place her in a clean cage.

Multiple recipients may be caged together. There are advantages and disadvantages to separating the foster mothers before the pups are born. With multiple births in the same cage, it is difficult to keep track of the success of individual surgeries and the cages may become overcrowded. Alternatively, some new mothers, particularly those with small litters, often take better care of their litters when there are other new mothers in the cage. Swiss Webster or ICR foster mothers are generally very good mothers under all conditions.

16. Monitor the mouse over several days. Administer analgesics as required by the institution.

With practice, there are virtually no complications. Bleeding should be minimal and usually only at the skin incision site. The mice should display normal behavior by the next day. Incision site infections are rarely observed.

PREPARATION OF PLASMID-BASED TRANSGENE DNA

Traditionally, transgenes have been constructed in plasmid-based vectors. Small quantities of DNA are required for the actual injection; however, the DNA preparation must be very clean to avoid toxicity in the zygote. In addition, vector sequences are commonly removed from the construct prior to injection to prevent these sequences from altering its expression characteristics. There are many protocols that will provide clean transgene fragments suitable for zygote injection. This protocol is a relatively simple procedure that involves separating the transgene from the vector using a low-melt agarose gel followed by fragment purification using an Elutip-d column.

Additional Materials

- Plasmid DNA containing transgene construct (*UNIT 1.7*)
- Low-melting-point agarose (e.g., Roche)
- TAE or TBE (*APPENDIX 2A*)
- Elutip low-salt solution (see recipe), 42°C

SUPPORT PROTOCOL 1

**Manipulating
the Mouse
Genome**

23.9.17

Elutip-d column (Schleicher & Schuell)

Elutip high-salt solution (see recipe)

70% ethanol

Plasmid injection buffer (see recipe)

42° and 65°C water baths

5- and 10-ml syringes

Additional reagents and equipment for restriction enzyme analysis (UNIT 3.1), agarose gel electrophoresis (UNITS 2.5A & 2.6), and ethanol precipitation (UNIT 2.1A)

Isolate fragments on a low-melting-point agarose gel

1. Cut ~50 µg plasmid DNA with restriction enzymes that will separate the contiguous transgene from the vector (UNIT 3.1).

Because the insert will be removed by subsequent purification steps, any standard method can be used to prepare the plasmid DNA, ranging from Qiagen kits to traditional alkaline lysis protocols (UNIT 1.6); however, the preparation should be relatively free of contaminating RNA and protein.

2. Prepare a 1% gel from low-melting-point agarose (UNIT 2.6) in TAE or TBE without ethidium bromide. Load one lane with size markers. Load the digested transgene in several adjacent lanes or a single broad lane. Run the gel.

Ethidium bromide in combination with UV light can cause DNA nicking, resulting in damage to the transgene.

3. After the gel has run long enough to separate the transgene from the vector, cut the gel lengthwise to include the standard lane and a small section of the sample lane if a broad lane was used or one of the lanes loaded with transgene if multiple sample lanes were loaded.

Be gentle when handling low melt 1% gels, because they will break apart very easily.

4. Stain the fragment of the gel containing the standard lane with ethidium bromide (UNIT 2.5A).

5. View the stained gel under UV light (UNIT 2.5A). Identify the transgene fragment based on its migration relative to the standard. Cut out the region of the gel containing the fragment.

CAUTION: Remember to wear appropriate eye protection and avoid exposure of any bare skin to the UV light source.

6. Realign the ethidium-stained gel fragment with the remaining untreated portion of the gel. Using the cut portion of the gel as a template, excise the region of the untreated gel containing the transgene fragment and transfer to a microcentrifuge tube.

This fragment of low-melt agarose contains the transgene fragment isolated from the vector and has never been exposed to ethidium bromide or UV light.

Purify fragment

NOTE: Care should be taken with all the solutions from this stage on. Use high-grade water and embryo- or cell-culture-tested reagents. This will prevent embryo-toxic contaminants from ending up in the final DNA preparation.

7. Add 10 gel volumes Elutip low-salt solution and melt the gel fragment at 65°C. After the gel fragment is completely dissolved, place the tube in a 42°C water bath.

8. Prepare the Elutip-d column by washing with 3 ml Elutip high-salt solution, room temperature, followed by 3 ml Elutip low-salt solution, 42°C, per manufacturer's instructions.

Use a clean syringe for the low-salt solution to avoid transfer of the high-salt solution to the Elutip column. The DNA binds to the Elutip in low salt.

9. Run the melted gel fragment through the Elutip column followed by an additional 3 ml low-salt solution, 42°C.

Don't press the syringe plunger with too much force. The column material can be pushed out of the column if too much force is applied.

10. Elute the bound DNA with 400 µl Elutip High-Salt Solution.

11. Ethanol precipitate the DNA (UNIT 2.1A) and wash several times with 70% ethanol to remove residual salt.

These washes are important because the DNA is eluted from the column in 1 M NaCl. This concentration of salt is toxic to the zygote. Do not use a carrier (e.g., glycogen) to facilitate the precipitation as this would also be injected into the zygote.

Prepare DNA for injection

12. Dry the pellet and resuspend in 50 to 100 µl plasmid injection buffer.

This solution may be stored frozen, but avoid repeated freezing and thawing.

13. Estimate the DNA concentration by running an aliquot on a 1% agarose gel with a known amount of size standards (UNIT 2.5A).

This step is also a good check for the purity of the prep. The fragment should appear as a single sharp band. There should be no smearing of the band which might indicate degradation or shearing of the fragment. There should be no visible vector band.

14. Dilute an aliquot of the preparation to a final concentration of 1 to 3 ng/µl in plasmid injection buffer.

PREPARATION OF BAC-BASED TRANSGENE DNA

Preparation of bacterial artificial chromosome (BAC) DNA for microinjection is somewhat different than for plasmid DNA. First, due to its large size, BAC DNA is much more susceptible to shearing damage. Tubes should never be vortexed or even shaken vigorously while the DNA is in solution. Second, there is little evidence that vector sequences affect BAC transgene expression; therefore, unlike plasmid-based transgenes, the insert does not need to be separated from the vector. Third, recent observations from many laboratories have shown that the BAC transgene does not need to be linearized prior to injection to produce founders efficiently.

There are a variety of methods to prepare BAC DNA. A method using the Nucleobond system from Clontech is described below. This system relies on a column containing an anion-exchange resin. The basic method is described in detail in the product literature but is slightly modified below to combine the endotoxin-free plasmid protocol with the low copy plasmid protocol. In addition, particular points relevant to transgene preparation are highlighted.

Materials

- BAC-transfected bacteria (e.g., UNITS 1.8 & 5.9)
- LB medium (UNIT 1.1)
- Nucleobond Plasmid EF Maxi Kit (BD Biosciences):
 - S1-EF buffer
 - RNase A

SUPPORT PROTOCOL 2

**Manipulating
the Mouse
Genome**

23.9.19

S3-EF buffer
N3-EF buffer
N5-EF buffer
Nucleobond AX 500 column
Filter
2-propanol
70% ethanol
BAC injection buffer (see recipe)

Additional reagents and solutions for agarose gel electrophoresis (*UNIT 2.5A*)

NOTE: Mega or giga kits can be used to process a larger volume of overnight culture. Remember to adjust the buffer volumes accordingly if larger cultures are used.

Isolate BAC DNA–containing bacteria

1. Prepare a 500-ml culture of BAC-transfected bacteria in LB medium and incubate overnight at 37°C, unless otherwise specified for the specific BAC or host bacteriological strain.
2. Transfer the culture to centrifuge bottles and centrifuge 10 min at 6000 × g, 4°C. Discard supernatant and gently resuspend the pellet in 24 ml Nucleobond S1-EF buffer plus RNase A.

Because the BAC is not processed after purification, the endotoxin-free (EF) system is used to ensure that the final preparation will not be toxic to zygotes. When following the Nucleobond protocol, note that the BAC protocols use a larger volume of many of the solutions. Do not use bacterial pellets from volume cultures that are larger than those recommended. The protocols are also optimized for LB medium cultures. If richer medium is used, the bacteria yield will be greater but may overload the columns, resulting in dirtier preps.

Lyse bacteria and remove insoluble fraction

3. Add an equal volume of S2-EF buffer and mix gently by inverting the tube. Incubate at room temperature for 2 to 5 min.

Do not vortex and do not shake the tube. Vigorous mixing will shear the BAC and will release bacterial DNA that will contaminate the prep.

Do not incubate the solution for longer than 5 min. Prolonged exposure to alkaline conditions (S2) can lead to irreversible denaturation of the BAC.

4. Neutralize the solution by adding 24 ml S3-EF buffer. Mix gently by inversion and place on ice for 5 min.

Steps 1 through 4 are essentially identical to any standard alkaline lysis plasmid prep. A white precipitate forms with the addition of the potassium acetate solution (S3). The majority of the bacterial protein and DNA should precipitate at this stage.

5. Centrifuge the suspension 15 min at 12,000 × g, 4°C. Collect the supernatant.
6. Pour the supernatant through the prewetted filter provided in the Nucleobond Plasmid EF Maxi kit and collect filtrate.

Purify and analyze BAC DNA

7. Pour the lysate onto a pre-equilibrated Nucleobond AX 500 EF column. Collect the eluent and pour it back through the column.
8. Wash the column twice with 18 ml N3-EF buffer.
9. Elute the BAC with 15 ml of buffer N5-EF buffer.

10. Add an equal volume of room-temperature 2-propanol. Mix gently and incubate 15 to 20 min at room temperature.
11. Centrifuge the mixture 30 min at $15,000 \times g$, 4°C .
12. Wash the pellet several times with 70% ethanol to remove residual salt. Allow the sample to air dry. Resuspend the pellet in BAC injection buffer.

DNA can be stored at 4°C . Larger DNA fragments are more likely to be damaged by freezing and thawing, therefore, freezing of the BAC DNA is not recommended.

Prepare BAC DNA for injection

13. Estimate the DNA concentration by running an aliquot on a 0.8% agarose gel with a known amount of size standards (UNIT 2.5A).

This step is also a good check for the purity of the prep. The fragment should appear as a single sharp band. It may be helpful to run uncut BAC and BAC digested with a restriction enzyme (UNIT 3.1) to get a better impression of the quality of the DNA. The intact BAC will not migrate very far. BAC digested with a restriction enzyme should display clear, crisp restriction fragments. Smearing of the fragments suggests that there may have been significant sheering of the original BAC. Smearing in the lane suggests that there is bacterial genomic DNA contamination of the prep.

14. Dilute an aliquot of the preparation to a final concentration of 1 to 2 ng/ μl in BAC injection buffer.

PREPARATION OF EQUIPMENT FOR SURGERY, EMBRYO TRANSFER, AND EGG INJECTION

SUPPORT PROTOCOL 3

Surgical/dissection equipment and instruments

Surgical instruments should be washed thoroughly before use and sterilized by autoclave or glass bead sterilizer. Table 23.9.2 presents examples of useful instruments. Similar instruments can be obtained from several companies. Individual investigators may prefer slightly different tools.

Table 23.9.2 Surgical Instruments^a

Instrument	Supplier
Betadine	VWR
Electric clippers	Harvard Apparatus
Forceps, small (4 in.) curved	BRI
Forceps, large (4 in.), curved	BRI
Forceps, toothed, 3.5 in.	BRI
Forceps, fine, $4 \frac{3}{8}$ in.	BRI
Glass bead sterilizer	Fine Science Tools
Scissors, large (4.5 in.)	Biomedical Research Instruments
Scissors, small (3.5 in.)	BRI
Scissors, iris, 3 in.	BRI
Serrefine clip	BRI
Suture, 5-0 Dexon II on T-31 needle	Kendall Health Care (or equivalent)
Wound clip applicator	Fisher
Wound clips	Fisher

^aA glass bead sterilizer provides a convenient and rapid bench-top sterilizer for surgical instruments.

Warming element

Anesthetized mice can suffer from hypothermia. The embryo transfer itself can be performed rapidly enough so that the surgical platform does not need to be heated; however, during the longer recovery phase, the mouse should be warmed until it shows signs of activity. Electric or chemical heating pads are useful portable warmers. A slide warmer is a good alternative in a dedicated surgical area, because the surface can be sterilized easily and the temperature can be controlled accurately.

Embryo transfer pipet

Embryo transfer pipets are prepared from 25- μ l disposable microcapillary pipets (e.g., Fisher). Pipets can be purchased with a mouth-pipet apparatus that consists of a pipet connector, tubing, and a mouthpiece. It is often convenient to replace the tubing with a longer piece of Tygon tubing. Hold the pipet over a small flame at its center. Gently rotate the capillary to ensure uniform heating. When the glass begins to turn orange from the heat, remove it from the flame and simultaneously pull the ends apart. Pull straight apart until the capillary breaks to produce two transfer pipets with flush ends. Note that if the capillary is pulled while the glass is still in the flame, the ends will be too fine. The exact diameter of the opening is not critical (100 to 150 μ m); however, it must easily admit a zygote.

For maximum control, the capillary should be back-filled with mineral oil slightly beyond the point where the capillary thickens to its largest internal diameter. A small volume of M2 medium is then aspirated prior to manipulating embryos. Before loading injected eggs for reimplantation, put two small air bubbles in the tube before the eggs and one bubble after. These act as markers that are highly visible to ensure that all of the eggs are transferred into the oviduct. Introduce the bubbles by gently aspirating air while the tip is out of the medium. Note that there is no reason to use air bubble markers unless the pipet is being loaded for reimplantation. Practice is required to make the transfer pipets and to become comfortable moving embryos using mouth control of the pipet.

Injection and holding pipets

Injection and holding pipets can be purchased from a variety of vendors (e.g., Eppendorf, The Pipette Company, Research Instruments, Zavos Diagnostic Laboratories). The pipets can also be made, but require the purchase of a pipet puller and microforge. The procedure also requires significant practice to produce quality pipets consistently. It is recommended that an inexperienced injector purchase premade tips until familiar with the injection procedure and the characteristics of properly designed pipets.

Injection pipets are typically loaded with the DNA solution (see Support Protocols 1 and 2) through the back end using a fine microcapillary pipet. Eppendorf sells a microloader for their injection pipets. Alternatively, a loading pipet can be made by pulling a microcapillary pipet (e.g., Fisher) in a flame similar to the procedure for making an embryo transfer pipet (see above); however, the glass is heated more intensely and pulled to produce a much finer end. Only several microliters of the DNA solution needs to be loaded into the injection pipet. Make sure that the pipet is loaded so that there are no air bubbles near the tip.

Injection apparatus

The injection apparatus consists of three main components: the microscope and table, the micromanipulators, and the microinjectors.

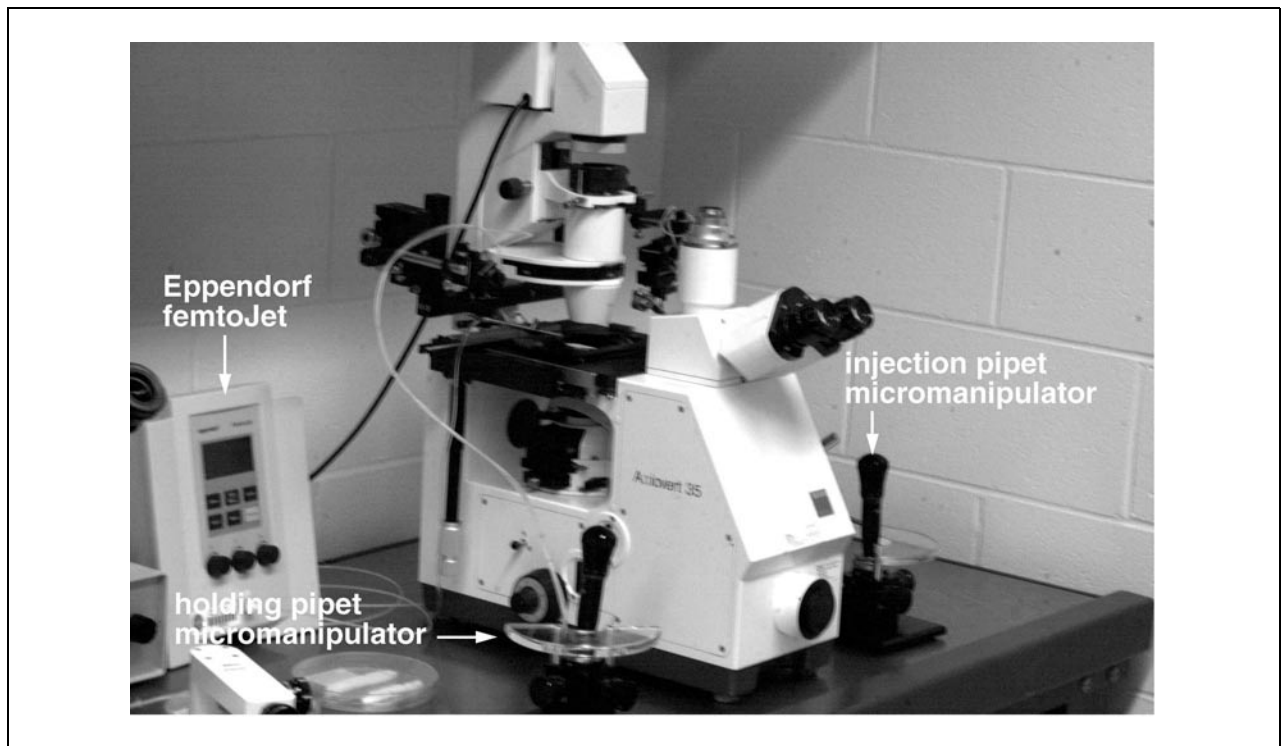


Figure 23.9.9 Injection apparatus. Photograph of a typical injection apparatus. This mixed system is based on a Zeiss Axiovert microscope with Narishige manipulators. The microinjector (not shown) for the holding pipet is a simple threaded syringe resting on the right side of the microscope connected by tubing to the pipet holder on the left side. The micromanipulator for the holding pipet is on the left side. Also on the left-hand side is an Eppendorf FemtoJet microinjector. This can be activated by a foot pedal (not shown). The entire system rests on an air isolation table.

Microscope and table

Many microscope configurations can be modified to work for zygote injections. The major components of the apparatus are illustrated in Figure 23.9.9. The basic microscope is an inverted scope with a fixed stage (e.g., Nikon, Leica Microsystems, or Zeiss). Several microscope vendors, as well as secondary vendors, market complete injection setups suitable for both zygote and blastocyst injections (e.g., Leica Microsystems, Nikon, McBain Instruments, Morrell Instrument). Under most circumstances the microscope must be placed on some type of isolation table, for example, a Micro-g (Fullam) or Kinetic Systems air table to prevent vibrations from adversely affecting the injection procedure.

It is convenient to have a range of objectives (e.g., 4 \times , 10 \times , 20 \times , and 40 \times) in combination with a standard 10 \times eyepiece. Pipets are aligned most easily under low magnification (40 \times), while the injections should be performed under high magnification (400 \times). The intermediate magnifications are useful for maintaining orientation, screening egg quality, and moving zygotes to different regions of the microdrop. Phase objectives are sufficient, but a better image can be obtained using Nomarski differential interference contrast optics. Note that a glass injection chamber must be used to take advantage of Nomarski optics. Also, consider the addition of a video camera and monitor. Although not necessary, a video system is an invaluable teaching tool.

Micromanipulators

Micromanipulators are another critical component of the injection apparatus. They are used to control the motion of the micropipets. There are several different types available. Narishige and Eppendorf controls are probably the most commonly used. These are

available in a variety of styles and are designed to attach to most standard microscopes. Manipulators that attach directly to the microscope frame are much preferred over those that rest on the microscope table. Purchase a system that allows both coarse and fine control of the pipets in all three dimensions. The coarse controls are used to set the general orientation of the pipet tips and the fine controls are used for the injections.

Microinjectors

The microinjectors are the devices used to control the action of the micropipets: aspiration or ejection. The microinjector connected to the holding pipet is basically a syringe connected by tubing to a pipet holder. The system can be oil or air filled. Some users find an oil-filled system provides slightly more control. Eppendorf and Narishige systems are probably used most commonly. The microinjector and micromanipulator for the holding pipet are placed on opposite sides of the microscope so that both controls can be manipulated simultaneously. The most common controller for the injection pipet is probably the Eppendorf FemtoJet (Fig. 23.9.9). This is basically a controllable compressor connected by tubing to a micropipet holder. Note that both microinjectors can be "homemade;" however, an inexperienced user will have less difficulty with devices designed specifically for pronuclear injection.

REAGENTS AND SOLUTIONS

Solutions should be prepared from tissue-culture or embryo-grade reagents whenever possible. Water should be tissue-culture grade (high resistance and endotoxin free) for all solutions that will come in contact with eggs. Sterilize all final solutions by filtration or prepare from sterile stocks. Use of disposable sterile plasticware is advised to prevent microbial and detergent contamination. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Avertin, 2.5%

Prepare a working stock solution (2.5% avertin) by adding 2.5 ml of 100% avertin (see recipe) to 100 ml PBS (APPENDIX 2A) prewarmed to 37°C. Shake the container until the avertin is completely dissolved. Store up to several months in the dark at 4°C.

Avertin, 100%

Dissolve 5 g 2,2,2-tribromoethanol (Aldrich) in 5 ml *tert*-amyl alcohol (Aldrich). Make sure reagents are at room temperature before mixing. Shake the container until all the tribromoethanol is in solution. Store up to several months in the dark at 4°C.

Toxic decomposition products can accumulate. Discard any questionable solutions. Purchase components in small amounts (5 ml and 5 g). There is batch-to-batch variation; therefore, with some preparations, a dose that is 5% to 10% higher than the recommended 0.017 ml/g must be administered for complete anesthesia.

BAC injection buffer

Prepare a 1000× stock solution of 30 mM spermine and 70 mM spermidine in water. Sterilize by passing through a 0.22- μ m filter. Prepare 1-ml aliquots and store up to 1 year, frozen at -20°C.

Prepare a base solution composed of 10 mM Tris-Cl, pH 7.4 (APPENDIX 2A), 0.1 mM EDTA, and 100 mM NaCl. Pass through a 0.22- μ m filter to sterilize and store up to 6 months at 4°C.

Dilute the spermine and spermidine stock solution 1:1000 in base solution just prior to use.

continued

NOTE: Reagents and water must be of the highest grade.

Although a simple plasmid injection buffer (Tris and EDTA) can be used for BAC injections, this modified buffer has been reported to increase the yield of transgenics when using large constructs (Montoliu et al., 1995). This is attributed to the stabilization of the larger constructs by the salt and polyamines.

The final concentrations are 10 mM Tris·Cl, pH 7.4, 0.1 mM EDTA, 100 mM NaCl, 30 μ M spermine, and 70 μ M spermidine.

Elutip high salt solution

1.0 M NaCl
20 mM Tris·Cl, pH 7.4 (APPENDIX 2A)
1.0 mM EDTA
Store up to 6 months at room temperature

Elutip low salt solution

0.2 M NaCl
20 mM Tris·Cl, pH 7.4 (APPENDIX 2A)
1.0 mM EDTA
Store up to 6 months at room temperature

HCG, 50 U/ml

Resuspend lyophilized human chorionic gonadotropin (HCG; also known as Pregnyl; Organon) powder in sterile PBS to a final concentration of 50 U/ml. (Store lyophilized powder up to the expiration date indicated on the vial at 4°C.) Store solution up to 6 months in 1-ml aliquots at -80°C . Avoid multiple freeze/thaw cycles.

Microdrop cultures

Prepare microdrop cultures by placing a small volume of M2 or M16 (50 to 150 μ l) on a petri dish. Cover the drop with mineral oil (Sigma, M-8410, embryo tested) to prevent evaporation. Equilibrate M16 microdrip cultures 30 min in a 37°C, 5% CO₂ incubator. Use fresh drops daily.

Note that M2 will maintain the proper pH in room air, while M16 requires 5% CO₂ to maintain proper pH.

Plasmid injection buffer

10 mM Tris·Cl, pH 7.4 (APPENDIX 2A)
0.1 mM EDTA
Pass the solution through a 0.22- μ m filter to sterilize
Store up to 6 months at 4°C

Reagents and water must be embryo or tissue culture grade.

PMS, 50 U/ml

Prepare a filter-sterilized 1 mg/ml BSA solution in PBS (APPENDIX 2A). Resuspend lyophilized pregnant mare's serum (PMS; also known as Gestyl; Professional Compounding Centers of America) in the BSA solution to a final concentration of 50 U/ml. (Store lyophilized powder up to the expiration date indicated on the vial at 4°C.) Store the solution up to 6 months in 1-ml aliquots at -80°C . Avoid multiple freeze/thaw cycles.

COMMENTARY

Background Information

The first transgenic mouse produced by pronuclear injection was described by Gordon et al. (1980). In this report, a pBR322-based plasmid containing herpes simplex virus and simian virus 40 DNA sequences was introduced into the mouse genome. Within months, several other groups reported similar success (Brinster et al., 1981; Constantini and Lacey, 1981; Harbers et al., 1981; Wagner et al., 1981a,b). These initial papers documented the fundamental aspects of a useful transgene: integration into the genome, expression in the resultant mouse, and transmission of the transgene to offspring. Pronuclear injection rapidly became the most common method for introducing specific DNA sequences into the mouse genome.

Alternative methods for the introduction of foreign DNA into the mouse genome do exist. Infection of preimplantation embryos with retroviruses had been reported during the period that pronuclear injection was gaining popularity (Jahner et al., 1985; van der Putten et al., 1985), but the method was not pursued extensively. Recent interest in the development of vectors for gene therapy has led to a reemergence of interest in the use of retroviruses for the production of transgenic mice. Lois et al. (2002) reported the germline transmission of transgenes introduced by lentiviral vector infection of single-cell embryos. The virus could be introduced by injection into the perivitelline space (Fig. 23.9.6B) or by co-incubation with denuded zygotes. The later method obviates the need for expensive injection equipment; however, there are some limitations that restrict the use of this methodology. For example, the transgene cannot exceed 10 kb, because of constraints of the lentiviral vector. In addition, the virus integrates as single segregating copies; therefore, transgenes requiring the integration of multiple linked copies for high level expression may not function effectively. DNA may also be introduced into mice by transfection into mouse embryonic stem (ES) cells followed by blastocyst-mediated transgenesis (see *UNITS* 23.5 & 23.7). This method is used primarily for the production of site-specific alterations in the mouse genome.

Historically, the vast majority of transgenes have been constructed in high-copy-number plasmids which restrict the insert size to <20 kb in most cases. Larger transgenes can be constructed in yeast artificial chromosomes

(YACs) and bacterial artificial chromosomes (BACs). YAC transgenics have been generated by pronuclear injection (Schedl et al., 1992) and by ES cell transfection followed by blastocyst-mediated transgenesis (Strauss et al., 1993). BAC transgenics created by pronuclear injection are now widely used for several reasons. Transgenes created from large genomic fragments typically display correct developmental and tissue-specific expression patterns. The availability of high quality, arrayed BAC libraries encompassing the mouse or human genome makes identification and isolation of specific clones relatively trivial. Finally, comprehensive sequence information obtained from genome sequencing projects in conjunction with new methodologies for site-specific modifications of BACs (Lee et al., 2001; Testa et al., 2004) have made the production of complex BAC transgenes a relatively simple procedure.

Critical Parameters and Troubleshooting

Creating transgenics by the injection of zygotes is a difficult procedure. Problems can occur in each of the three main steps: zygote isolation, injection, and reimplantation. To avoid problems, each of these steps requires both familiarity with the protocol and practice. It is best to practice the procedures separately before trying to combine them on a single day. Many of the problems are under the investigator's control; however, some aspects of these protocols rely on the basic biology of mice and are not entirely under the control of the investigator.

Low yield of eggs from superovulated females

Low egg yields can usually be attributed to the superovulation reagents or procedure, or to the female mouse. Check the concentration, age, and storage conditions of the PMS and HCG. Prepare new batches if there is any question. Make sure to administer the HCG 46 to 48 hr after the PMS and to coordinate the injections with the light/dark cycle. Check to make sure the light cycle has not been changed inadvertently. Verify the age of the female mice. Many strains produce their greatest egg yields before the females go through their first estrus cycle. This means that they should be used between 3 and 4 weeks of age. Females older than 2 months typically have smaller egg yields. The strain of the female mouse plays an

important role: some strains will never produce a large quantity of eggs. If a low-yield strain must be used, the number of matings should be increased. Finally, it is worth noting that any routine will have bad days because of the inherent biological variability of the system. The only way to ensure adequate egg yields for every day is to set up more matings, but this is probably not cost effective if egg yields are adequate for most days.

Low yield of fertilized eggs

Problems with the percentage of eggs that are fertilized can usually be attributed to the males or to factors that affect mating behavior. Check plugs on a regular basis. If there are few plugs, there will not be many fertilized eggs. Mating records should be kept for all stud males. Males that routinely injure females or that fail to plug regularly should be replaced. There is often a strain-specific decline in fertility as the males age, even if they plug females on a regular basis. As a general rule, males should be replaced when they reach 6 to 8 months of age, although many strains last longer. Try to stagger the ages of the stud males so that they are not all replaced at one time. Check the environment of the egg production colony. Verify the light cycle, humidity, and temperature in the room. Determine if there has been a sudden change in activity or noise in the room. Changes in these environmental factors can affect mating behavior dramatically.

Difficulty of injecting

Problems with a clogged injection needle are usually attributed to bad DNA preps; however, first make sure that the DNA concentration is correct. Some problems can be corrected by microcentrifugation for smaller transgene or filtration through 0.45- μ m syringe filters for both large plasmid and BAC transgenes. If the DNA concentration is correct and microcentrifugation or filtration does not correct the problem, a new prep of transgene DNA should be made. Note that it is not unusual to replace the injection pipet several times because of the accumulation of cellular debris on the tip. The visibility and size of the pronuclei can also affect the ease of injection. This can vary between strains and during the time the eggs are in culture. Some strains, such as FVB, have large, prominent pronuclei and are therefore particularly suited for pronuclear injections.

Poor survival of eggs

The death of a high percentage of eggs shortly after injection can be attributed to sev-

eral factors. For an experienced injector, the most common cause of rapid zygote death is a DNA prep of poor quality. Try injecting several eggs with injection buffer alone. If the majority of these eggs survive, then the DNA preparation is the most likely culprit. It may be advantageous to have a single individual prepare fragments for injection or to enforce rigid adherence to a successful transgene isolation protocol in order to ensure uniform quality of DNA preparations. When there is not a question of DNA quality and the injector is inexperienced, the technique itself is the most likely cause of problems. Monitor the volume of fluid injected; the pronucleus should swell visibly, but not lyse. Check the path of the injection pipet; the tip should pass directly into the pronucleus along a straight path and the zygote should not roll or turn during injection. Try to orient the zygote so that the pronucleus is adjacent to the injection tip to minimize damage to the egg. Withdraw the tip with a smooth and rapid motion along the path of entry; watch for the accumulation of debris on the end of the tip that may stick to DNA and proteins within the zygote.

Few pups or founders born

If there are few live-born pups, verify that the reimplantation protocol is not the problem and that the pseudopregnant females were used at the correct stage. Test the technique with zygotes that have not been manipulated. Most unmanipulated, healthy zygotes should produce pups. If the reimplantation technique is not suspect, look for other possible problems, such as the quality of the DNA prep. In some cases there may be a reasonable number of pups but a low yield of transgenics. If this is the case, check the concentration of the transgene prep: low concentrations will reduce the yield of founders. Check the injection system to ensure that there is proper compensation pressure; if there is backflow of medium into the injection tip, the founder yield will be low. Finally, consider the possibility that expression of the transgene may be toxic to the embryo.

Anticipated Results

Approximately 60% to 80% of injected zygotes should survive the injection procedure. DNA quality, DNA concentration, and injector experience are the major factors that affect zygote survival. After reimplantation, 25% to 35% of injected eggs survive to term. Pups should be born 19 to 20 days after reimplantation. About 10% to 20% of the pups can be

Table 23.9.3 Timeline of Transgenic Line Identification and Expansion

Day	Action
0	Zygote injection
20	Birth of potential founders
41	Wean, tag, and tail all pups; genotype to identify founders ^a
62	Initiate mating of founders ^b
83	Birth of first founder offspring
104	Wean, tag, and tail founder offspring ^c

^aPups are typically genotyped at weaning but they can be analyzed earlier.

^bFounders carrying the transgene are then mated after they reach sexual maturity.

^cThis generation, available 3.5 months after zygote injections, is usually the first generation that is subjected to experimentation.

expected to be transgene founders. Approximately 80% to 90% of the transgene founders are not mosaic and should transmit the transgene to half of their offspring.

Time Considerations

Although the procedure spans 4 days (Table 23.9.1), only the day of injection requires a significant time commitment. With experience, zygote isolation should take 30 to 60 min. Injection of 100 to 200 eggs should take 1 to 2 hr and reimplantations should take about 15 min for each foster mother. A novice investigator will spend a very long day trying to complete the procedure. Zygote isolation and reimplantation should be mastered before attempting to complete the entire procedure.

The length of time between zygote injection and the generation of a cohort of transgenic mice suitable for experimental studies is obviously much longer. In most cases, founder offspring, not the founders themselves, are the focus of experiments. The first generation of founder offspring are available a minimum of 3.5 months after zygote injection (Table 23.9.3).

Literature Cited

- Brinster, R.L., Chen, H.Y., Trumbauer, M.E., Senear, A.W., Warren, R., and Palmiter, R.D. 1981. Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. *Cell* 27:223-231.
- Constantini, F. and Lacy, E. 1981. Introduction of a rabbit beta-globin gene into the mouse germ line. *Nature* 294:92-94.
- Gordon, J.W., Scangos, G.A., Plotkin, D.J., Barbosa, J.A., and Ruddle, F.H. 1980. Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc. Natl. Acad. Sci. U.S.A.* 77:7380-7384.
- Harbers, K., Jahner, D., and Jaenisch, R. 1981. Microinjection of cloned retroviral genomes into mouse zygotes: Integration and expression in the animal. *Nature* 293:540-542.
- Jahner, D., Haase, K., Mulligan, R., and Jaenisch, R. 1985. Insertion of the bacterial *gpt* gene into the germ line of mice by retroviral infection. *Proc. Natl. Acad. Sci. U.S.A.* 82:6927-6931.
- Lee, E.C., Yu, D., Martinez de Velasco, J., Tessarollo, L., Swing, D.A., Court, D.L., Jenkins, N.A., and Copeland, N.G. 2001. A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* 72:56-65.
- Lois, C., Hong, E.J., Pease, S., Brown, E.J., and Baltimore, D. 2002. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* 295:868-872.
- Montoliu, L., Bock, C.-T., Schutz, G., and Zentgraf, H. 1995. Visualization of large DNA molecules by electron microscopy with polyamines: Application to the analysis of yeast endogenous and artificial chromosomes. *J. Mol. Biol.* 246:486-492.
- Schedl, A., Beermann, F., Thies, E., Montoliu, L., Kelsey, G., and Schutz, G. 1992. Transgenic mice generated by pronuclear injection of a yeast artificial chromosome. *Nucl. Acids Res.* 20:3073-3077.
- Strauss, W.M., Dausman, J., Beard, C., Johnson, C., Lawrence, J.B., and Jaenisch, R. 1993. Germ line transmission of a yeast artificial chromosome spanning the murine alpha 1(I) collagen locus. *Science* 259:1904-1907.
- Testa, G., Vintersten, K., Zhang, Y., Benes, V., Muylers, J.P., and Stewart, A.F. 2004. BAC engineering for the generation of ES cell-targeting constructs and mouse transgenes. *Methods Mol. Biol.* 256:123-139.
- van der Putten, H., Botteri, F.M., Miller, A.D., Rosenfeld, M.G., Fan, H., Evans, R.M., and Verma, I.M. 1985. Efficient insertion of genes into the mouse germ line via retroviral vectors. *Proc. Natl. Acad. Sci. U.S.A.* 82:6148-6152.

Wagner, E.F., Stewart, T.A., and Mintz, B. 1981a. The human beta-globin gene and a functional viral thymidine kinase gene in developing mice. *Proc. Natl. Acad. Sci. U.S.A.* 78:5016-5020.

Wagner, T.E., Hoppe, P.C., Jollick, J.D., Scholl, D.R., Hodinka, R.L., and Gault, J.B. 1981b. Microinjection of a rabbit beta-globin gene in zygotes and its subsequent expression in adult mice and their offspring. *Proc. Natl. Acad. Sci. U.S.A.* 78:6376-6380.

Key References

Nagy, A., Gertsenstein, M., Vinterstein, K., and Behringer, R. 2003. *Manipulating the Mouse Embryo: A Laboratory Manual*, 3rd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

This manual is a compendium of methods used to modify the mouse genome, analyze genetically altered embryos, and maintain mouse lines. It is probably the most-referenced text on transgenic mouse production.

Hammes, A. and Schedl, A. 2000. *Mouse Genetics and Transgenics: A Practical Approach* (I.J. Jackson and C.M. Abbott, eds.) pp. 217-245. Oxford University Press Inc., New York.

Voncken, J.W. 2003. *Transgenic Mouse: Methods and Protocols* (M.H. Hofker and J. van Deursen, eds.) pp. 9-34. Humana Press, Totawa, New Jersey.

The last two references are also collections of methods used to modify the mouse genome. They overlap the first reference with regard to transgenic mouse production, but provide additional protocols for the genetic analysis of mice and further characterization of transgenic animals.

Internet Resources

<http://www.biosupplynet.com>

Search this Web site to obtain a current list of suppliers for materials and reagents used in the production of transgenics by pronuclear injection.

<http://www.jax.org>

The Jackson Laboratory web page provides links to a wealth of mouse related information, including, mutant resources, trait mapping resources, and literature pertaining to mouse genetics and animal husbandry.

Contributed by David A. Conner
Harvard Medical School
Boston, Massachusetts

Transgenic Mouse Colony Management

UNIT 23.10

Managing a colony of transgenic mice created by zygote microinjection (UNIT 23.9) is similar to managing a colony of mice created by gene targeting (UNIT 23.8). However, there are important differences because of the random nature of the genetic modification that results from microinjection. This unit describes an approach for identifying founders and developing genotyping strategies. In addition, animal husbandry concepts pertaining to transgenic mice are discussed.

The time course of identifying founders and expanding and maintaining independent lines is outlined in Figure 23.10.1, starting with the day of zygote injection. Offspring from zygote injections are born 3 weeks later. The mice are weaned, tagged, and tail biopsies are taken when they are ~3 weeks old. Potential founders are identified by screening genomic DNA for the presence of the transgene. Independent lines are established by mating transgene-positive founders with wild-type mice when the founders reach sexual maturity, ~6 weeks after birth. Offspring from these matings are screened for the transgene to verify germline transmission. For the majority of transgenic experiments, the first experimentally useful animals are these offspring (F1 generation) derived from the founders. These mice reach weaning age a minimum of 4 months from the day of zygote

injection. Representative transgene-positive mice from the F1 generation of each founder line are typically screened to determine transgene copy number, a crude estimator of transgene expression levels.

COLONY ESTABLISHMENT, EXPANSION, AND MAINTENANCE

Housing

Housing considerations for transgenic mice are identical to those for mice generated by gene targeting (UNIT 23.8). In brief, mice should be housed in a specific-pathogen-free (SPF) facility if possible. SPF animals will live longer and display greater fecundity than comparable animals housed in open facilities.

Identifying Potential Founders

A transgenic colony begins with founders that have stably integrated the transgene. All offspring from zygote injections should be screened for the presence of the transgene. Each transgene-positive mouse represents a unique line because the transgene typically inserts randomly, with variable copy number. Multiple lines must be established for several reasons. First, copy number and insertion site can affect transgene expression levels, and thus a single founder may not express the transgene at experimentally useful levels. Second, random insertion of the transgene may affect the

week 0: zygote injection
week 3: birth of potential founders
week 6: wean, tag, obtain tail biopsy, screen for transgene
week 9: set up transgene-positive founder matings
week 12+: birth of first founder offspring: F1 Generation
week 15+: wean, tag, obtain tail biopsy, verify germline transmission of transgene, determine transgene copy number

Figure 23.10.1 Time line for identifying founders and establishing independent lines, from zygote injection to verification of germline transmission. Potential founders can be identified by genotyping at the time of weaning, ~6 weeks after zygote injection. Both male and female transgene-positive founders should be mated to establish independent lines. Founder matings can be initiated when the mice are ~6 weeks old, although it may take several weeks before the first successful mating occurs. Most founders will transmit the transgene to 50% of their offspring. These offspring, the F1 generation, are used to verify germline transmission and determine transgene copy number. For most investigations, the F1 generation is the first experimentally useful generation. Thus, a minimum of 15 weeks must pass between zygote injection and the weaning of the first experimentally useful animals.

Contributed by David A. Conner

Current Protocols in Molecular Biology (2005) 23.10.1-23.10.8
Copyright © 2005 by John Wiley & Sons, Inc.

Manipulating
the Mouse
Genome

23.10.1

Supplement 71

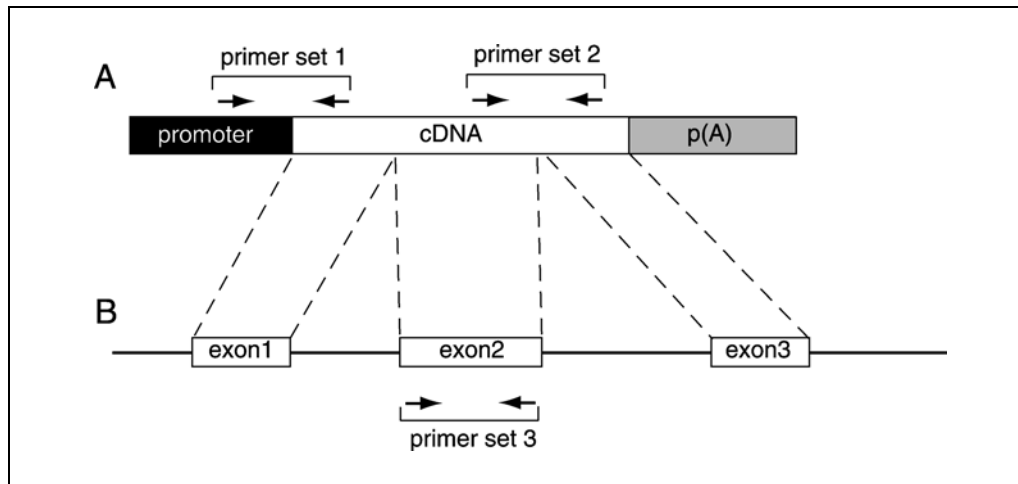


Figure 23.10.2 Genotyping primer selection. A typical transgene construct is illustrated (A), consisting of a heterologous promoter, a cDNA, and sequences containing a polyadenylation signal, p(A). The cDNA is derived from a hypothetical mouse gene (B), which consists of three exons. Three primer pairs are depicted for use as genotyping primers. Primer pair 3 would be unable to distinguish the transgene from the endogenous gene—a similar-size fragment corresponding to part of exon 2 would be amplified from both templates. Primer pair 2 could distinguish between the transgene and the endogenous gene. The primers span an intron and would amplify a longer fragment from the endogenous gene. Primer pair 1 is the most specific pair of primers. This primer pair may be the most useful in a mouse colony containing several different transgenic lines consisting of the same cDNA driven by different promoters.

function of an endogenous gene and create a phenotype that has nothing to do with expression of the transgene. With a single founder, it would be difficult to determine whether a phenotype was caused by expression of the transgene or by altered function of an endogenous gene at the site of insertion. Third, not all transgene-positive founders will transmit the transgene to their offspring. A fraction numbering 10% to 20% of the transgene founders will be mosaic and will not transmit the transgene with the predicted Mendelian frequency.

The two most common methods for identifying founders are the polymerase chain reaction (PCR; *UNIT 15.1*) and Southern blotting (*UNITS 2.9A & 2.10*). PCR is probably the best method to use at this stage because of its simplicity and sensitivity. At weaning, all offspring from the zygote injections should be marked, most commonly by ear tagging (*UNIT 23.8*). A tail biopsy should be obtained from each animal for DNA isolation (*UNIT 23.8*). Two methods, one standard and one rapid, for DNA isolation from tail biopsies are described in *UNIT 23.8*. The standard method, which yields sufficient DNA for both PCR and Southern blotting, is recommended at this stage in case PCR results need to be confirmed by Southern blotting.

DNA samples are analyzed by PCR using primers specific for the transgene. The primers

must be able to differentiate between the transgene and any highly similar endogenous genes (Fig. 23.10.2). For example, if the transgene construct contains a cDNA corresponding to an endogenous mouse gene, the primers can be designed to span an intron in the endogenous gene (primer set 2); PCR amplification of the endogenous gene would result in a larger product than the transgene product because of the inclusion of the intron. Alternatively, if the transgene construct combines a mouse cDNA with a novel promoter, one primer in the promoter and the other in the cDNA will amplify a unique product (primer set 1).

Primers should be selected to amplify relatively small products (~150 to 500 base pairs). This helps to ensure that the reaction will be robust. In addition, control primers should be included in the reaction. Control primers are typically a second pair of primers that amplify an entirely different single-copy locus present in all DNA samples. The product should be roughly the same size as the transgene-specific primer product, but still separable on a 3% to 4% agarose gel. Similar primer melting points should be selected so that transgene-specific and control primers can be run in the same reaction. Amplification of the control band verifies that the reaction conditions and the DNA samples are suitable for single-copy detection. All DNA samples should test positive with the control primers. A positive control for the

Table 23.10.1 Copy Number Standards for PCR and Southern Blotting

Transgene size (bp)	PCR single-copy standard (pg)	Southern blot single-copy standard (pg)
1000	0.067	3.3
5000	0.33	17
10000	0.66	33

transgene-specific primers can be constructed by spiking wild-type DNA with the transgene construct at a single-copy level per genome (Table 23.10.1). The reaction conditions and the primers must be optimized to ensure detection of the transgene (*UNIT 15.1*). It is important that founders not be lost at this stage because of false-negative PCR results.

Using the copy number standards indicated in Table 23.10.1, assume that the size of the haploid mouse genome is 3×10^9 base pairs. The following relationship can be applied:

$$\frac{\text{size of transgene in base pairs}}{3 \times 10^9 \text{ base pairs}} = \frac{\text{single copy mass}}{\text{genomic DNA mass}}$$

In Table 23.10.1, it is assumed that 0.2 μg of genomic DNA are used as the template in the PCR reaction and 10 μg of genomic DNA are used for digestion for the Southern blot. Therefore, for a 5000-bp transgene, the PCR reaction should be spiked with 0.33 pg of the transgenic construct to approximate the abundance of a single-copy target in 0.2 μg of genomic DNA. PCR primers should be able to detect this quantity of target DNA easily. Primer pairs should not be used if they do not work efficiently. Similarly, 17 pg of the transgenic construct should be used to spike 10 μg of genomic DNA for a single-copy standard in the Southern blot. For Southern blots, a series of several standard lanes spiked with increasing quantities of the transgene can be used to construct a standard curve to determine transgene copy number from unknown samples.

Establishing Separate Founder Lines and Determining Copy Number

Each transgene-positive animal represents the potential founder of a unique transgenic line. At this stage, all founders positive for the transgene should be crossed with wild-type mice to establish separate lines. Some founders may not pass the transgene because of mosaicism, insertion-site effects, or a transgene phenotype.

Some thought should be given to the choice of the wild-type strain used in these crosses. Background strain may affect the phenotype of a transgene. If an inbred strain, such as FVB (*UNIT 23.9*) was used as the zygote source, it may be worthwhile maintaining the transgene in the same inbred background. Recreating an inbred line after outcrossing is time-consuming. If outbred or F1 hybrid zygotes were used for the initial injection, it may be worthwhile considering the generation of congenic lines (*UNIT 23.8*) by successive crosses into a defined inbred strain, to create lines that are genetically identical except for the locus surrounding the transgene. If congenics are not created, successive inbreeding of littermates of separate founder lines derived from outbred or F1 hybrid zygotes will eventually generate inbred lines that are genetically different at many loci. This may confound phenotypic analysis.

Regardless of background-strain considerations, the first step in establishing a transgenic line is verifying that the founder can pass the transgene to its offspring. Subsequently, representative animals derived from each founder should be analyzed to determine relative transgene copy number, since there is a strong positive correlation between transgene expression levels and copy number. Copy number is usually determined by Southern blotting, whereas germline transmission can be assayed by PCR.

Mark all offspring from the first cross, typically with ear tags, and obtain a tail biopsy as described in *UNIT 23.8*. Genomic DNA should be prepared using the standard method for DNA isolation for Southern analysis and PCR (*UNIT 23.8*). Offspring from all founders should be screened for the presence of the transgene by PCR, using the same method described for founder identification.

Representative transgene-positive DNA samples from offspring of each of the founders and wild-type littermates should be analyzed by Southern blot (*UNITS 2.9A & 2.10*). Figures 23.10.3 and 23.10.4 illustrate two common combinations of restriction enzyme and probe that can be used to determine relative

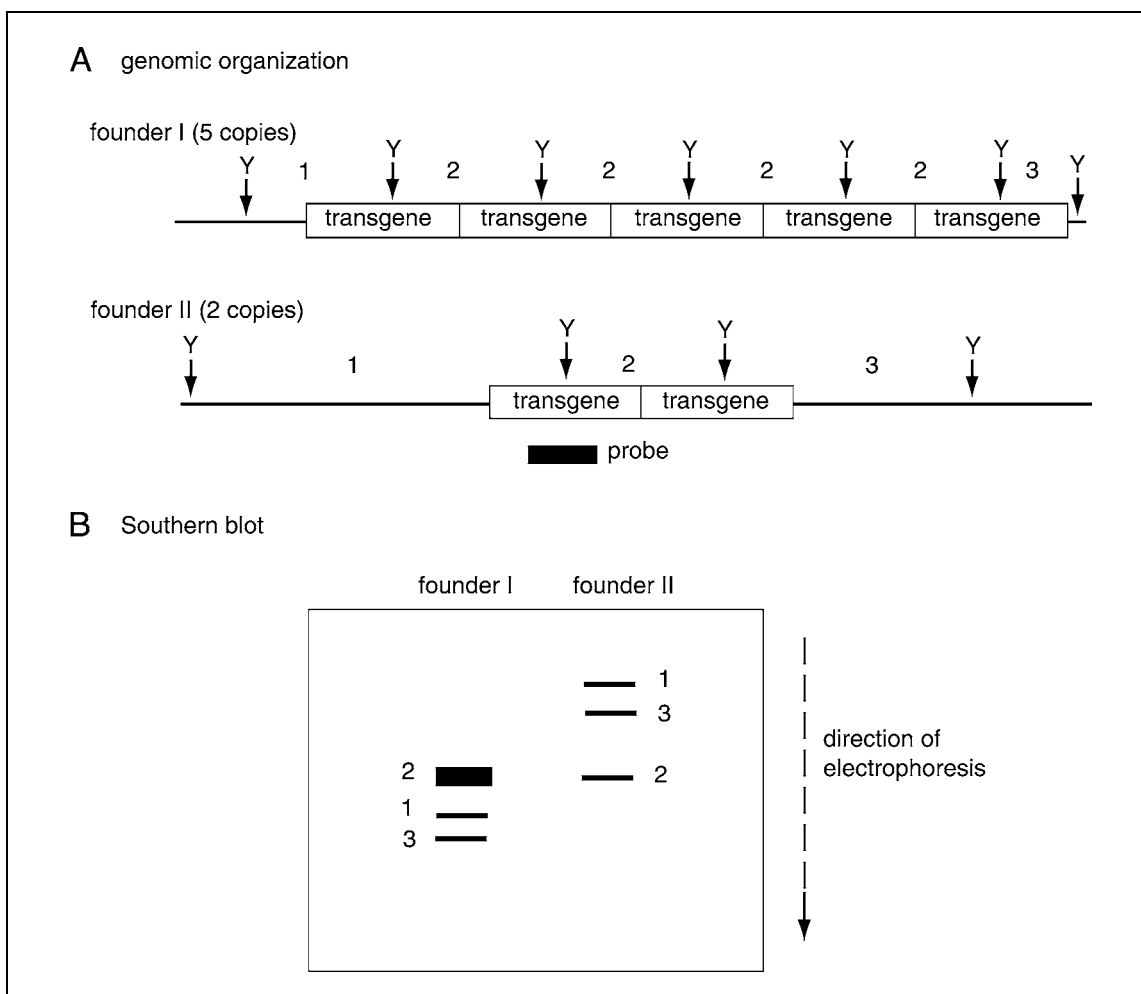


Figure 23.10.3 Southern analysis of founder lines: single site, short construct. **(A)** This panel depicts the genomic organization of two hypothetical founder lines. Founder I has five head-to-tail copies of the transgene and founder II has two head-to-tail copies of the transgene. Y represents the location of a hypothetical restriction enzyme site that cuts the transgene in a single location, and cuts outside the transgene at positions that will be determined by the site of integration. The black bar corresponds to a region of the transgene used as a probe for Southern blots. The unique fragments that are recognized by the probe are numbered 1, 2, and 3. **(B)** This panel depicts a hypothetical Southern blot of genomic DNA cut by restriction enzyme Y for founder I (first lane) and founder II (second lane) and hybridized with the probe shown in panel A. The intensity of the hybridization signal is represented by the size of the band. In the first lane, the probe hybridizes to four copies of fragment 2. Note that fragment 2 is the same size as the original transgene. The flanking fragments, 1 and 3, are present as single copies. In the second lane, the same band (2) is recognized, but in this case there is only a single copy. The flanking fragments, 1 and 3, are also present as single copies. Their sizes differ from the flanking fragments of founder I because of the difference in the site of integration. In this hypothetical example, the approximate relative difference in copy number between the two founders is determined by comparing the relative intensities of band 2, assuming that the same amount of DNA was loaded into each lane. The size of the single-copy flanking bands are unique to each insertion and can be used to identify a specific founder line. In this example, the flanking bands can also be used as single-copy internal standards to determine the absolute copy number of the transgene. Note that the flanking bands may not always be visible, because their size is dependent on the random location of restriction sites in the genome.

transgene copy number. In the first example, equivalent amounts of genomic DNA are digested with a restriction enzyme that cuts the construct at a single site. After separation on an agarose gel and transfer to a membrane (Southern blotting), the membrane is hybridized with a labeled probe that spans the restriction site. In this case, the probe will hybridize

to an internal fragment the size of the original construct because of the typical head-to-tail organization of the transgene concatamer. The relative intensity of the internal fragment when compared between different founders will be proportional to the copy number. Flanking fragments, present as single copies (Fig. 23.10.3), may also be detected. These

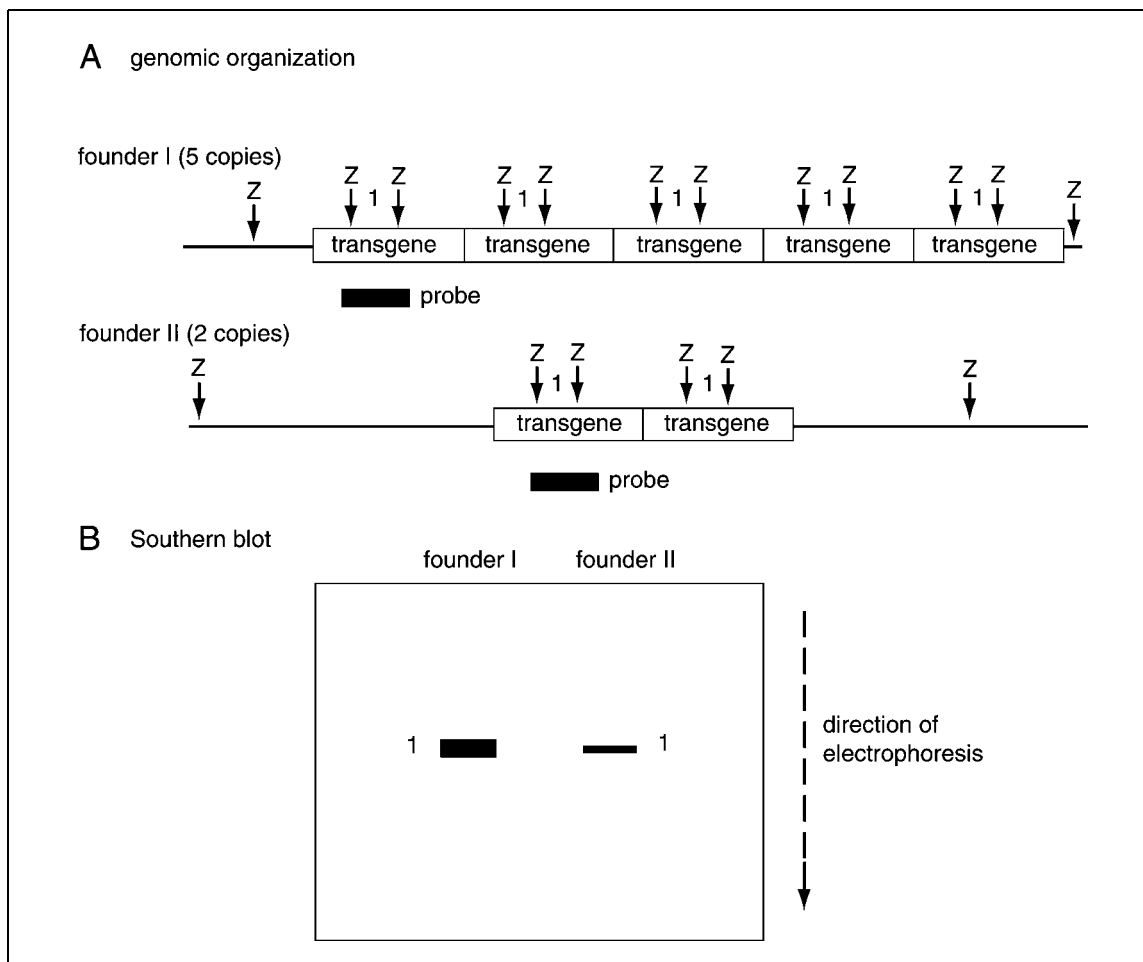


Figure 23.10.4 Southern analysis of founder lines: multiple sites, long construct. **(A)** This panel depicts the genomic organization of two hypothetical founder lines. Founder I has five head-to-tail copies of the transgene and founder II has two head-to-tail copies of the transgene. Z represents the location of a hypothetical restriction enzyme site that cuts the transgene in two locations, and cuts outside the transgene at positions that will be specific to the site of integration. The black bar corresponds to a region of the transgene that can be used as a probe for Southern blots. Only a single internal fragment (1) will be recognized by the probe. **(B)** This panel depicts a hypothetical Southern blot of genomic DNA cut by restriction enzyme Z for founder I (first lane) and founder II (second lane) and hybridized with the probe shown in panel A. The intensity of the hybridization signal is represented by the size of the band. In the first lane, the probe hybridizes to five copies of fragment 1. In the second lane, the same band (1) is recognized, but in this case there are only two copies. Note that no flanking fragments will be detected using this combination of probe and restriction enzyme. In this hypothetical example, the relative difference in copy number between the two founders is determined by comparing the relative intensities of band 1, assuming that the same amount of DNA was loaded into each lane.

can be used as internal single-copy standards to calculate the absolute copy number of the transgene.

In many cases, it may not be practical or possible to use a restriction enzyme that cuts only a single time in the construct. If the transgene is much longer than 20 kb, a restriction enzyme that cuts once will generate a fragment that is too large to resolve easily using standard Southern blotting protocols. Figure 23.10.4 illustrates an alternative method to assess copy number. In this example, a restriction enzyme that cuts two or more times within the construct is used. A probe corresponding to one

of the internal fragments is generated. The probe should hybridize to a single band in all founders; the relative hybridization intensity will be proportional to the copy number when equivalent amounts of genomic DNA are analyzed in each sample.

Copy number may also be assessed by comparing the hybridization signal from transgenic lanes with a series of control samples generated by spiking wild-type genomic DNA with known quantities of the transgene construct. The calculation used to determine the amount of transgene construct corresponding to a single copy per copy of the mouse genome is

shown in Table 23.10.1. Several samples with increasing amounts of the transgene can be used to create a standard curve. These should be run on the same blot as the unknown samples to control for variations in the specific activity of probes used in different experiments. The advantage of this approach is that a true copy number can be determined. However, in most cases it is sufficient to know the relative copy number between the different founder lines.

Expanding Founder Lines and Generating Experimental Animals

At this stage, transmitting founder lines have been identified and transgene copy number has been estimated as a crude indicator of transgene expression levels. Typically, this F1 generation represents the first experimental generation. Early experiments should usually include northern blots and immunoblots to establish that the transgene is expressed properly. In most cases, transgenic lines are maintained as heterozygotes by crossing transgene carriers with wild-type mice. Subsequent generations can be genotyped by PCR using the rapid DNA preparation (*UNIT 23.8*). Genotyping primers should be selected to be as specific as possible for the transgene. As indicated previously, without knowledge of the flanking sequence surrounding the transgene insertion site, it is not possible to develop a PCR assay that is specific for each founder line. However, it is usually possible to develop PCR primers that are specific for each transgene the laboratory pursues. For example, a laboratory may maintain multiple transgenic *lacZ* lines, each line differing by the promoter that drives the transgene. It may seem most efficient to use a pair of PCR primers that amplify part of the *lacZ* coding region; the primers can be used for all the lines. However, this is probably not a prudent approach. Accidental mixing of the lines would not be detected. In this case, PCR primers that sit in the *lacZ* coding region and in the unique promoter region of each transgene would be more likely to detect accidental mixing of the individual lines (e.g., primer set 1 in Fig. 23.10.2).

It is also common to maintain several lines from the same transgene. The lines will differ in insertion site, copy number, and, probably, expression level. Lines derived from the same transgene pose more of a problem with regard to the possibility of accidental mixing. Routine PCR genotyping will not distinguish the lines. It is a good habit to periodically genotype by Southern blotting to distinguish these lines by

copy number or by the presence of different-sized flanking fragments (Fig. 23.10.3).

In some cases, the generation of homozygous animals may be useful. Homozygous animals have a higher transgene copy number; therefore, they may express the transgene at a higher level. Homozygotes are also easier to maintain, since offspring from homozygote crosses do not need to be genotyped at every generation. However, it is a prudent practice to genotype homozygotes periodically to ensure that the lines are breeding true and that no accidental crosses have occurred. Homozygosity can be determined using semiquantitative Southern blotting. Equal amounts of genomic DNA from offspring of heterozygous crosses and from known heterozygotes are digested and analyzed by Southern blotting using the same method described for relative copy number determination. Homozygous DNA samples should have twice the hybridization signal as heterozygotes. Alternatively, quantitative PCR can be used to assess zygosity (Shitara et al., 2004). Putative homozygotes can be test-bred to wild-type mice to verify homozygosity; no wild-type offspring should be observed.

Long-Term Strain Maintenance

Transgenic mouse lines pose a slightly different problem with regard to long-term maintenance than lines created by gene targeting, although the options are the same (*UNIT 23.8*). An identical transgenic line can never be regenerated because of the random nature of the integration event. Therefore the decision to eliminate a line that is of no immediate use is permanent. There is no perfect alternative. Many investigators choose to maintain lines by keeping an active colony of a few mating pairs or a few males that can be mated periodically to produce a younger generation. This method is not labor-intensive, and is relatively inexpensive for a few lines. It is more problematic when many lines must be maintained. Two main alternatives exist, both of which can be performed by laboratories familiar with the manipulation of preimplantation embryos. Morulae can be frozen using a specialized controlled-rate freezer. Although the equipment is expensive, this method can accommodate large batches of embryos. Alternatively, morulae can be frozen by vitrification, which involves the osmotic dehydration of embryos by sequential equilibration in cryoprotectant solutions of increasing osmotic strength (Glenister and Rall, 2000). The latter method does not require expensive

equipment and is most appropriate for small batches of embryos. Several major mouse suppliers also provide morula freezing services (Taconic Farms, Charles River Laboratories). Mouse sperm cryopreservation is still not widely used because of strain-dependent variability in the success of *in vitro* fertilization with frozen sperm. However, there have been recent advances that make this method more appealing (Bath, 2002; Szczgiel et al., 2002). Current methods for sperm freezing are within the abilities of most laboratories (Sztein et al., 2001; Bath, 2002; also see <http://www.jax.org/cryo/sperm.html>). If a commercial or in-house cryogenic method of strain preservation is selected, there must be a rigorous quality-control program to ensure that frozen embryos or sperm are capable of regenerating live animals.

MOUSE IDENTIFICATION AND INFORMATION MANAGEMENT

Various methods for marking mice are reviewed in *UNIT 23.8*. The most common method is ear tagging, which is an inexpensive and quick procedure for an experienced investigator.

Mouse colonies tend to expand very rapidly, especially when several lines are maintained for each transgenic construct. It is imperative that complete records be kept, starting with the founders. A paper-based management system will be overwhelmed quickly. A variety of digital database software options are reviewed in Table 23.8.3. Information should be collected for individual animals, matings, and litters. Complete records collated in a digital database are invaluable for explaining unanticipated results. It is more common to regret keeping incomplete records than to regret keeping too much information.

TROUBLESHOOTING

Absence of Founders

The absence of founders may be a result of technical problems with zygote injections. These issues are discussed in *UNIT 23.9*. Alternatively, potential founders may be missed because of an inefficient PCR screen. New transgene-specific primers must be tested to ensure efficient amplification of a single-copy target per genome. In addition, control primers must be able to amplify a product from all DNA samples to verify the quality of the genomic DNA preparation and the efficiency of the reaction. There is also the possibility that expres-

sion of the transgene causes lethality, resulting in the death of all founders prior to birth or before weaning.

Problems with Germline Transmission

In most cases, the transgene will be passed to 50% of the offspring as expected. However, several problems may be encountered when screening the F1 generation to verify germline transmission. If transmission is not observed, the founder genotype should be verified; misidentification of the original founder is the simplest explanation. If the founder is truly transgene-positive, there are several other explanations for absent or aberrant germline transmission. A significant percentage of founders will be mosaic. These animals may never transmit the transgene or may pass the transgene at frequencies lower than expected. The transgene can integrate into the X chromosome, resulting in 100% transmission to female offspring and no transmission to male offspring of male founders. In some cases, transgenes will integrate into multiple sites. Offspring from these founders may inherit the transgene at unexpectedly high frequencies if the integration sites are not linked. In addition, these offspring may appear to have different transgene copy numbers depending on the number of separate loci that are inherited. Finally, there is always the possibility that lack of transmission of the transgene is a phenotype of the transgene expression or a result of an insertion-site mutation (see below).

Insertion-Site Mutations

Unexpected phenotypes may be a result of insertion-site mutations. Random integration of the transgene may affect the function of an endogenous gene near the site of insertion. Insertion-site mutations are more likely to have a recessive phenotype than a dominant phenotype. Therefore, the effects of these mutations may not be observed until heterozygous transgene carriers are crossed. Independent founder lines should not have the same insertion-site mutations. Investigators should suspect that a novel phenotype present only in a single founder line is a result of an insertion-site mutation, rather of transgene expression.

Transgene Silencing

In some cases, transgene expression may decrease as animals age, or in later generations (Robertson et al., 1996; Henikoff, 1998). This epigenetic phenomenon is called transgene silencing. It is most common in high-copy-number lines created with plasmid-based

transgenes. It is irreversible and unpredictable. The possibility of transgene silencing should be considered when there is a large phenotypic variation between animals of different ages or different generations.

LITERATURE CITED

- Bath, M.L. 2002. Simple and efficient in vitro fertilization with cryopreserved C57BL/6J mouse sperm. *Biol. Reprod.* 68:19-23.
- Glenister, P.H. and Rall, W.F. 2000. Cryopreservation and rederivation of embryos and gametes. *In Mouse Genetics and Transgenics: A Practical Approach* (I.J. Jackson and C.M. Abbot, eds.) pp. 27-59. Oxford University Press, New York.
- Henikoff, S. 1998. Conspiracy of silence among repeated transgenics. *BioEssays* 20:532-535.
- Robertson, G., Garrick, D., Wilson, M., Martin, D.I.K., and Whitlaw, E. 1996. Age-dependent silencing of globin transgenes in the mouse. *Nucleic Acids Res.* 24:1465-1471.
- Shitara, H., Sato, A., Hayashi, J., Mizushima, N., Yonekawa, H., and Taya, C. 2004. Simple method of zygosity identification in transgenic mice by real-time quantitative PCR. *Transgenic Res.* 13:191-194.
- Szczgiel, M.A., Kusakabe, H., Yanagimachi, R., and Whittingham, D.G. 2002. Separation of motile populations of spermatozoa prior to freezing is beneficial for subsequent fertilization in vitro: A study with various mouse strains. *Biol. Reprod.* 67:287-292.

Sztejn, J.M., Noble, K., Farley, J.S., and Mobraaten, L.E. 2001. Comparison of permeating and non-permeating cryoprotectants for mouse sperm cryopreservation. *Cryobiology* 42:28-39.

KEY REFERENCES

Silver, L.M. 1995. *Mouse Genetics: Concepts and Applications*. Oxford University Press. New York.

Useful and thorough review of the house mouse and its use in genetics. The text is available on-line at <http://www.informatics.jax.org/mgihome/resources/resources.shtml>.

INTERNET RESOURCES

<http://www.biosupplynet.com>

Search this Web site to obtain a current list of suppliers for animal husbandry-related items.

<http://www.informatics.jax.org>

The Jackson Laboratory informatics Web site provides access to a wide variety of mouse-related information.

Contributed by David A. Conner
Harvard Medical School
Cambridge, Massachusetts

Modification and Production of BAC Transgenes

The sequencing of entire genomes and assembly of physical maps of whole chromosomes has been greatly facilitated by the cloning of genomic DNA into bacterial artificial chromosome (BAC) vectors (UNIT 5.9). BACs are F-factor-based cloning systems that were designed for the specific purpose of cloning large-insert genomic DNA fragments in the 100- to 300-kb range. In vivo analysis of gene function and expression frequently relies on the production of transgenic animals by introducing DNA constructs into the mouse genome. Traditional transgenic constructs prepared in standard (pBR-type) plasmid vectors afford a maximal cloning capacity of ~20 kb. In contrast, BAC clones, owing to their larger insert size, are more likely to carry an entire transcription unit and its associated regulatory regions. As a consequence, BAC-based transgenic approaches offer improved and unique opportunities to study gene function, expression, and regulation in a genomic environment that is more consistent with the endogenous locus.

What makes the use of BACs particularly advantageous is the ease with which they can be engineered to incorporate DNA insertions, deletions, or substitutions. A variety of BAC modification protocols have recently been described that use homologous recombination-based approaches performed in *E. coli*. Most of the BAC modification methods use targeting vectors containing a recombination gene(s) (*recA*, *recE/recT*, or *lexo/beta/gam*), which transiently restores to the recombination-deficient host the ability to promote homologous recombination between the BAC of interest and the targeting DNA element (which may be a linear DNA fragment or circular vector). The flexibility in the design of the system makes it possible to introduce different types of modifications to the BAC DNA including insertion of reporter cassettes, deletion of specific gene or regulatory regions, or the substitution of one type of sequence for another.

In addition to the site-directed methods for modifying BAC DNA, other approaches utilizing in vitro transposition allow for the random integration of foreign DNA into BAC clones. This method is preferred when screening several BACs for reporter activity. In this case, the same reporter cassette can be introduced into multiple BACs.

This unit provides protocols for the modification of BAC DNA using homologous recombination in *E. coli* (see Basic Protocol 1) and in vitro transposition (see Basic Protocol 2). In addition, a simple and reliable procedure is described for the preparation of BAC DNA for pronuclear injection into fertilized mouse eggs (see Support Protocol), which is an alternative to the methodology outlined elsewhere in this manual (UNITS 2.1B & 23.9).

BAC MODIFICATION BY HOMOLOGOUS RECOMBINATION IN *E. COLI*

This protocol was developed in the laboratory of Dr. Nathaniel Heintz (Rockefeller University, New York, N.Y.) and was first published by Gong et al. (2002). It uses a shuttle vector, pLD53SC-AEB, with the following features: *recA* gene to mediate homologous recombination; R6K γ DNA replication origin for propagation in *E. coli* strains expressing the π replication factor; ampicillin resistance gene for positive selection; *sacB* gene for negative selection upon growth on plates containing sucrose; and the enhanced green fluorescent protein (eGFP) reporter gene. This BAC modification protocol is performed in three steps. The first step involves the preparation of a targeting construct using the pLD53SC-AEB shuttle vector. To introduce the eGFP reporter into a selected site within a BAC, homology arms A and B must be cloned on either side of the eGFP

BASIC PROTOCOL 1

Manipulating the Mouse Genome

23.11.1

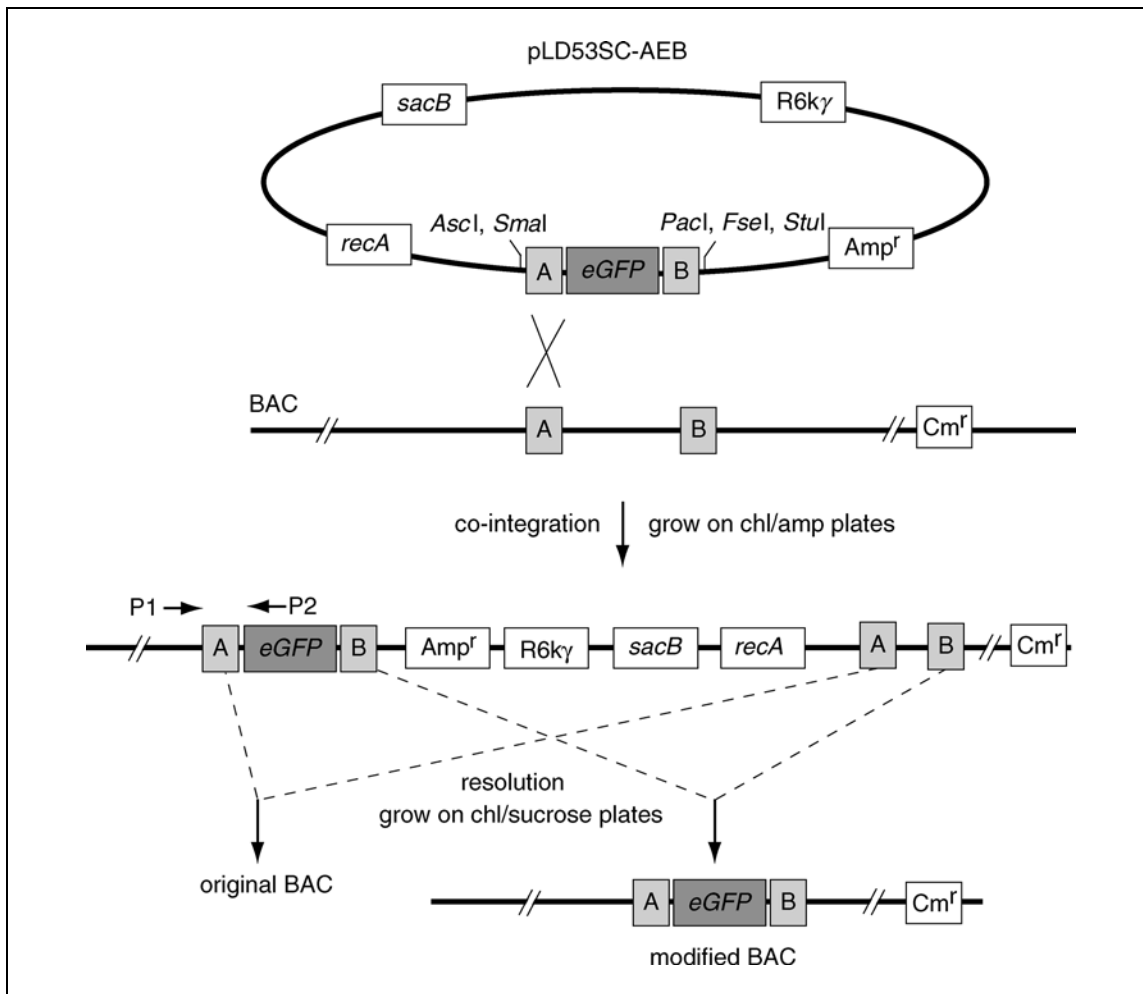


Figure 23.11.1 BAC recombination system using the pLD53SC-AEB shuttle vector. The shuttle vector contains the R6ky replication origin; ampicillin-resistance gene (*Amp^r*); and *recA*, *sacB*, and enhanced green fluorescent protein (*eGFP*) genes. Homology arms can be cloned into *Ascl* or *SmaI* (for arm A) and *PacI*, *FseI*, or *StuI* (for arm B) restriction sites. The modification occurs through a two-step homologous recombination reaction: the co-integration of the shuttle vector into the BAC target through arm A (for example) and the resolution of the co-integrant through arm B. Integration through arm B is equally possible but is not shown. P1 and P2 indicate PCR primers used to verify co-integration. Abbreviations: amp, ampicillin; chl, chloramphenicol; Cm^r, chloramphenicol resistance gene.

cDNA in the pLD53SC-AEB shuttle vector, as described in the steps below. The second step involves the integration of the shuttle vector into the BAC of choice by homologous recombination. This recombination event will occur between arm A or arm B of the shuttle vector and the corresponding DNA sequence in the BAC (Fig. 23.11.1). In the third step (resolution), a second recombination event excises the backbone of the shuttle vector and positions the intended modification within the desired target DNA. This recombination event will also occur between homology arm A or arm B and the corresponding BAC DNA, resulting in either a BAC containing the modification of interest or the original BAC (Fig. 23.11.1). The modified BAC is identified by PCR.

To incorporate other modifications into a BAC (e.g., substitutions, deletions, or point mutations), the *eGFP* sequence in the shuttle vector is simply replaced with the chosen modification, and the homology arms are cloned as described. The *eGFP* can be removed from the vector by digesting at an appropriate combination of cloning sites. For example, to delete a target sequence in the BAC, the authors typically design primers for arm A and arm B on either side of the sequence to be deleted. For arm A, the upstream primer incorporates an *Ascl* site at its 5' end and the downstream primer is phosphorylated at its

5' end. For arm B, the 5' end of the upstream primer is phosphorylated and the 3' end of the downstream primer incorporates a *PacI* site. After PCR, the 5' and 3' homology arms are digested with *AscI* and *PacI*, respectively, and ligated in a three-way reaction with the *AscI/PacI*-digested shuttle vector.

Materials

Pir2 competent cells (Invitrogen)
pLD53SC-AEB shuttle vector (from the laboratory of Dr. Nathaniel Heintz, The Rockefeller University)
LB medium (UNIT 1.1)
LB plates (UNIT 1.1) supplemented with antibiotics as follows:
30 µg/ml ampicillin
15 µg/ml chloramphenicol and 50 µg/ml ampicillin
15 µg/ml chloramphenicol and 5% (w/v) sucrose (from 50% sucrose stock solution; sterilize with 0.45-µm filter; do not boil or autoclave)
50 mg/ml ampicillin
Appropriate oligonucleotide primers for amplifying homology arms A and B
Bacteria containing BAC clone to be modified (see Internet Resources)
Restriction enzyme and buffer appropriate for cloning site in shuttle vector:
AscI or *SmaI* for homology arm A
PacI, *FseI*, or *StuI* for homology arm B
Shrimp alkaline phosphatase (Roche)
GeneClean (Qbiogene)
Primers for verifying arm insertion:
SCAB1: 5'-AAGTTGTAAGGATATGCC-3'
SCAB2: 5'-CATATCGCAATACATGCG-3'
34 mg/ml chloramphenicol
10% (v/v) glycerol, ice cold
SOC medium (UNIT 1.8), optional
Oligonucleotide primers P1 and P2 (step 27)
37° and 42°C water baths
37°C incubator, with and without shaking
Thermal cycler
1.5-ml microcentrifuge tubes, sterile and prechilled on ice
0.1-cm electroporation cuvettes (Bio-Rad), prechilled on ice
Gene Pulser (Bio-Rad)
Sterile culture tubes, such as 15-ml (17 × 120-mm) tubes (Falcon)
Additional reagents and equipment for alkaline lysis miniprep (UNIT 1.6), PCR (UNIT 15.1), agarose gel electrophoresis (UNIT 2.5A), ethanol precipitation of DNA with ammonium acetate (UNIT 2.1A), DNA ligation (UNIT 3.16), DNA sequencing (see Chapter 7), electroporation (UNIT 1.8), Southern analysis (UNITS 2.9A & 2.10), and preparation of BAC DNA (see Support Protocol)

Propagate pLD53SC-AEB shuttle vector

1. Thaw one vial of Pir2 competent cells on ice and mix with 1 to 5 µl of pLD53SC-AEB shuttle vector (1 ng/µl final) by tapping gently. Incubate 30 min on ice.
The R6Kγ DNA replication origin is not compatible for replication in most lab strains of E. coli commonly used for transformation, including DH5α and XL1. As a result, Pir2 cells must be used to propagate the shuttle vector.
2. Place vial in a 42°C water bath for 30 sec and then on ice for 2 min. Add 250 µl LB medium to the vial and incubate 1 hr at 37°C.
3. Spread 50 µl from the vial on an LB plate with 30 µg/ml ampicillin and incubate overnight at 37°C.

4. Inoculate a single colony into 5 ml LB medium with 30 µg/ml ampicillin and shake (225 rpm) overnight at 37°C.
5. Isolate DNA using standard alkaline lysis miniprep method (UNIT 1.6).

Prepare and clone homology arm A

6. Amplify arm A by PCR (UNIT 15.1) in a thermal cycler using 50 ng appropriate oligonucleotide primers (with a melting temperature, T_m) and 1 ng desired BAC DNA as template.

30 cycles: 1 min 94°C
 1 min ($T_m - 5$)°C
 1 min 72°C.

*The following points should be taken into consideration when designing PCR primers for homology arm A and B. The optimal length of homology arms is between 300 bp and 500 bp. The cloning sites available in the shuttle vector are *AscI* and *SmaI* (for homology arm A) and *PacI*, *FseI*, and *StuI* (for homology arm B). The restriction sites selected for cloning should be absent from the homology arms. UNIT 15.1 describes general considerations for the design of PCR primers.*

7. Examine the product on an agarose gel (UNIT 2.5A). If the predicted size product is obtained, digest it with the appropriate restriction enzyme (either *AscI* or *SmaI*) and remove the end fragments by ethanol precipitation with ammonium acetate (UNIT 2.1A).

Alternatively, a commercially available kit (e.g., QIAquick PCR purification kit; Qiagen) can be used to remove end fragments.

If a product of the predicted size is not obtained, the PCR should be repeated with lower annealing temperature or higher Mg^{2+} . New primers may also need to be designed.

8. Linearize the pLD53SC-AEB shuttle vector with the same restriction enzyme used to digest homology arm A and dephosphorylate it with shrimp alkaline phosphatase as per manufacturer's instructions.
9. Examine vector digestion by gel electrophoresis to confirm that majority of vector is linearized and then gel purify linearized vector using GeneClean.
10. Ligate linearized shuttle vector with digested arm A (UNIT 3.16) and transform the Pir2 cells with the ligated product as described above (steps 1 to 4).
11. Isolate DNA from colonies and verify the insert by restriction digestion or PCR. In addition, confirm the sequence of the insert (see Chapter 7).

*When verifying the presence of an insert by PCR, the following set of primers should be used: SCAB1 (upstream of *AscI* site in the vector, 5'-AAGTTGTAAGGATATGCC-3') and the downstream oligo used for the amplification of arm A. These primers can also be used for sequencing.*

Prepare and clone homology arm B

12. Prepare and insert homology arm B into the shuttle vector containing arm A as described above (steps 6 to 10). Use one of the following three cloning sites: *PacI*, *FseI*, or *StuI*.
13. Isolate DNA from colonies and verify the insert by restriction digestion or PCR. In addition, confirm the sequence of the insert.

*When verifying the presence of homology arm B by PCR, the following set of primers should be used: upstream oligo (used for the amplification of arm B) and SCAB2 (downstream of *StuI* site in the vector, 5'-CATATCGCAATACATGCG-3'). These primers can also be used for sequencing.*

This is the final step in the design of the targeting vector. When modifying several BAC clones, the target vector with homology arms A and B should be prepared on a large scale (UNIT 1.7).

Make electrocompetent cells

14. Inoculate a single colony of bacteria containing the BAC clone to be modified into 2 ml LB medium with 15 µg/ml chloramphenicol and grow overnight with shaking at 37°C.
15. Transfer 0.1 ml overnight culture into 50 ml LB medium with 15 µg/ml chloramphenicol and incubate with shaking at 37°C until the OD₆₀₀ reaches 0.5 to 0.6.

Cultures usually require 3 to 5 hr to reach this density.

16. Harvest cells by centrifuging 10 min at 1500 × g, 4°C. Resuspend pellet in 50 ml of 10% ice-cold glycerol. Repeat this step once more.
17. Resuspend pellet in 200 µl of 10% ice-cold glycerol and dispense 40-µl aliquots into sterile prechilled 1.5-ml microcentrifuge tubes. Freeze and store at –80°C.

Electrocompetent cells can be stored for up to 6 months at –80°C.

Perform co-integration

18. Thaw 40 µl electrocompetent cells on ice, mix with 1 to 2 µl targeting vector (0.5 µg/µl), and place on ice for 1 min.
19. Transfer mixture to a prechilled 0.1-cm electroporation cuvette and carry out electroporation (UNIT 1.8). On a Gene Pulser apparatus, set the capacitance to 25 µF, the voltage to 1.8 kV, and the pulse controller to 200 Ω.
20. Immediately add 1 ml SOC (or LB) medium to the cuvette and transfer to a sterile culture tube. Incubate tube 1 hr with shaking at 37°C.

LB can be used, but SOC is known to ensure maximum transformation efficiency.

21. Add 5 ml LB medium with 15 µg/ml chloramphenicol and 30 µg/ml ampicillin to the culture tube, and incubate overnight with shaking at 37°C.
22. Transfer 5 µl overnight culture to a fresh culture tube containing 5 ml LB medium with 15 µg/ml chloramphenicol and 50 µg/ml ampicillin. Incubate 8 to 14 hr with shaking at 37°C.
23. Transfer 1 µl culture to a fresh culture tube containing 5 ml LB medium with 15 µg/ml chloramphenicol and 50 µg/ml ampicillin. Incubate with shaking at 37°C overnight.
24. Make a 1:10⁴ to 1:10⁵ dilution of the overnight culture and spread 50 to 100 µl on an LB plate with 15 µg/ml chloramphenicol and 50 µg/ml ampicillin. Incubate at 37°C overnight.
25. Individually pick and inoculate a handful of colonies (5 to 10) into 1 ml LB medium containing 15 µg/ml chloramphenicol and 50 µg/ml ampicillin and streak each colony onto a chloramphenicol/ampicillin master plate. Grow overnight at 37°C.

Verify co-integration by PCR

26. Make a 1:10 dilution of the overnight culture using sterile water and lyse 5 µl diluted culture in a PCR tube by boiling for 3 min in a thermal cycler.

This will be the template for PCR.

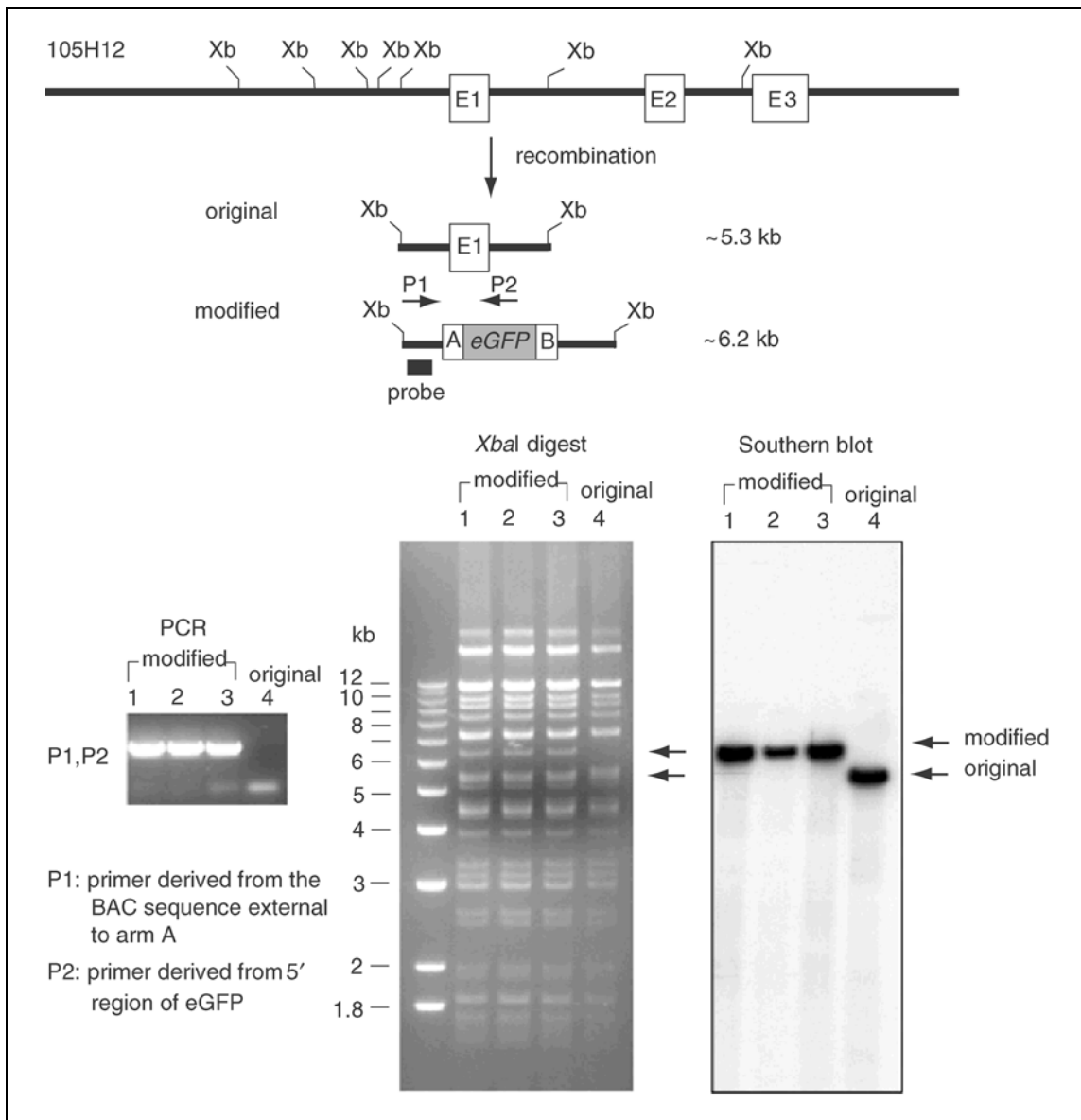


Figure 23.11.2 PCR and Southern analysis of a sonic hedgehog BAC clone modified by homologous recombination. Resolved BAC candidates were first screened by PCR using primers P1 and P2. DNA from PCR-positive clones was subsequently digested with *Xba*I (*Xb*), separated by electrophoresis on a 1% agarose gel, and transferred onto a nylon membrane. The blot was hybridized with a radiolabeled DNA probe located externally (5') to homology arm A. Lanes 1 to 3 are modified BAC clones. Lane 4 is the original BAC clone. E1, E2, and E3 are exons.

27. Use 5 μ l lysed diluted culture and 50 ng each of oligonucleotide primers P1 and P2 for PCR (Fig. 23.11.1) as follows:

30 cycles: 1 min 94°C
 1 min 55°C
 1 min 72°C.

P1 should be designed against the specific BAC sequence immediately external (5') to homology arm A. In the authors' laboratory, the *P2* primer is designed against the 5' end of the *eGFP* cDNA (5'-CCGGTGAACAGCTCCTCG-3').

Because homologous recombination can occur within arm B, screening for co-integrants can also be performed using primers that are derived from the 3' region of *eGFP* and the BAC sequence external to arm B.

The co-integration frequency may range from 30% to 70%.

Perform resolution step

28. Pick several colonies that were positive for co-integration from the master plates (step 25) and inoculate each colony into 1 ml LB medium with 15 µg/ml chloramphenicol. Incubate 1 hr with shaking at 37°C.
29. Spread 50 to 100 µl of each culture on an LB plate with 15 µg/ml chloramphenicol and 5% sucrose and incubate at 37°C overnight.

*This is the negative selection step. The *sacB* gene product, levansucrase, converts sucrose to levan, which is highly toxic to *E. coli*. Therefore, only those bacteria that have recombined out the shuttle vector are able to grow on LB plates with 5% sucrose.*

30. Pick small or intermediate-sized colonies, inoculate each colony into 2 ml LB medium with 15 µg/ml chloramphenicol, and incubate overnight with shaking at 37°C.
31. Screen for the resolution event by PCR as described (steps 26 and 27). Verify positive clones by Southern analysis (UNIT 2.9A & 2.10) using BAC DNA prepared from an overnight culture (see Support Protocol, steps 1 to 8) and a DNA probe external to homology arm A or B.

The modification frequency may range from 10% to 50%. An example of how to screen for the modified BAC by PCR, BAC DNA fingerprinting, and Southern blot analysis is shown in Figure 23.11.2.

To prepare a frozen stock of bacteria, 0.2 ml sterile glycerol is added to 0.3 ml bacterial culture, and the tube is inverted until glycerol is evenly dispersed. Frozen stocks can be stored indefinitely at -80°C.

BAC MODIFICATION BY IN VITRO TRANSPOSITION

This protocol uses a Tn7 transposon-based in vitro system (GPS; New England Biolabs) to randomly insert a *lacZ* reporter cassette into a BAC of interest. TnsABC transposases insert a transposon randomly into target DNA. Because of target immunity, only one insertion occurs per target DNA. Therefore, this system produces a population of target DNAs, each containing the transposon at a different position. This protocol uses the targeting vector pGPS1-βLacZ, which was modified from the original pGPS1 plasmid by Spitz et al. (2003; Fig. 23.11.3). This targeting vector contains transposable elements Tn7L/R; R6Kγ DNA replication origin, which allows for propagation only in *E. coli* strains expressing the π replication factor; kanamycin resistance gene; and a reporter cassette consisting of a minimal β-globin promoter, *lacZ* gene, and SV40 polyA signal.

Materials

- Pir2 competent cells (Invitrogen)
- pGPS1-βLacZ targeting vector (from the laboratory of Denis Duboule, University of Geneva, Switzerland)
- LB medium (UNIT 1.1)
- LB plates (UNIT 1.1) without antibiotics *and* supplemented with antibiotics as follows:
 - 20 µg/ml kanamycin
 - 20 µg/ml kanamycin and 15 µg/ml chloramphenicol
 - 20 mg/ml kanamycin
- E. coli* DH10B strain
- 10% (v/v) glycerol, ice cold
- BAC DNA prepared from overnight culture (see Support Protocol, steps 1 to 8)
- GPS kit (New England Biolabs), including:
 - 10× GPS buffer
 - TnsABC transposase
 - Start solution

BASIC PROTOCOL 2

SOC medium (UNIT 1.8), optional

34 mg/ml chloramphenicol

PCR primers:

β -globin: 5'-AGCCATCTATTGCTTACATTTGC-3'

lacZ: 5'-ATAGGTTACGTTGGTGTAGATGG-3'

Appropriate restriction enzyme to confirm insertion (e.g., *Mlu*I)

Primer for sequencing directed against the 3' end of the *lacZ* gene, optional

37° and 42°C water baths

37°C incubator, with and without shaking

1.5-ml microcentrifuge tubes, sterile and prechilled on ice

75°C heating block

0.1-cm electroporation cuvettes (Bio-Rad), prechilled on ice

Gene Pulser (Bio-Rad)

Thermal cycler

Additional reagents and equipment for alkaline lysis DNA miniprep (UNIT 1.6), electroporation (UNIT 1.8), PCR (UNIT 15.1), agarose gel electrophoresis (UNIT 2.5A), preparation of BAC DNA (see Support Protocol), pulsed-field gel electrophoresis (UNIT 2.5B), and DNA sequencing (see Chapter 7; optional)

Prepare shuttle vector

1. Thaw a vial of Pir2 competent cells on ice and mix with 1 to 5 μ l pGPS1- β LacZ targeting vector (1 ng/ μ l final) by tapping gently. Incubate 30 min on ice.
2. Place vial in a 42°C water bath for 30 sec and then on ice for 2 min. Add 250 μ l LB medium to the vial and incubate 1 hr at 37°C.
3. Spread 50 μ l from vial on an LB plate containing 20 μ g/ml kanamycin and incubate overnight at 37°C.
4. Inoculate a single colony into 5 ml LB medium containing 20 μ g/ml kanamycin and incubate overnight at 37°C with shaking (225 rpm). Isolate plasmid DNA by alkaline lysis miniprep (UNIT 1.6).

Make DH10B electrocompetent cells

5. Inoculate a single colony of *E. coli* DH10B strain grown on an LB plate without antibiotics into 2 ml LB medium without antibiotics and grow with shaking overnight at 37°C.
6. Transfer 0.5 ml overnight culture into 250 ml LB medium without antibiotics and incubate with shaking at 37°C until the OD₆₀₀ reaches 0.5 to 0.6.
Cultures usually require 3 to 5 hr to reach this density.
7. Harvest cells by centrifuging 10 min at 1500 \times g, 4°C. Resuspend pellet in 250 ml of 10% ice-cold glycerol. Repeat this step once.
8. Resuspend pellet in 1 ml of 10% ice-cold glycerol and dispense 40- μ l aliquots into prechilled sterile 1.5-ml microcentrifuge tubes. Freeze and store at -80°C.

Electrocompetent cells can be stored for up to 6 months at -80°C.

Perform transposition reaction

9. Mix the following in a tube by pipetting up and down two or three times.
 - 20 ng targeting vector (step 4)
 - 200 ng BAC DNA prepared from overnight culture
 - 2 μ l of 10 \times GPS buffer

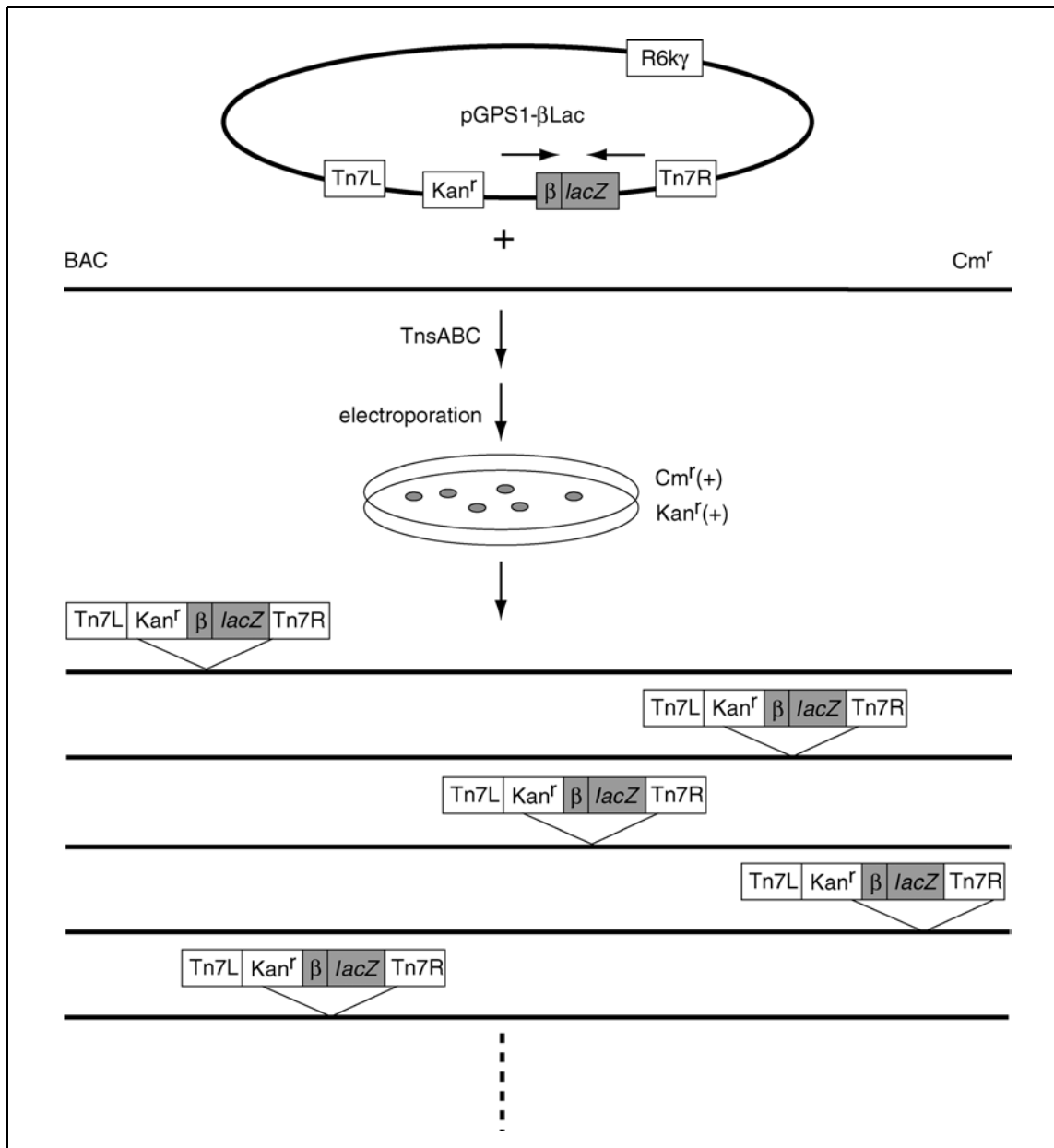


Figure 23.11.3 Overview of Tn7 transposon–mediated BAC modification. The targeting vector contains transposable elements (Tn7L/R), R6K γ DNA replication origin, kanamycin resistance gene (Kan^r), and a reporter cassette consisting of a minimal β -globin promoter (β) and *lacZ* gene (*lacZ*). The transposon cassette (Tn7L-Kan^r- β /*lacZ*-Tn7R) is inserted randomly into the BAC clone by transposases A, B, and C (TnsABC). Modified clones are selected by their ability to grow on plates containing chloramphenicol (Cm) and kanamycin.

1 μ L TnsABC transposase
 H₂O to a final volume of 18 μ L.

Incubate 10 min at 37°C.

This incubation is necessary for the assembly of the transposases, targeting vector, and BAC and must not be omitted.

10. Add 2 μ L start solution, mix well by pipetting up and down two or three times, and incubate 1 hr at 37°C.

11. Heat inactivate 10 min at 75°C.

It is essential that the heat inactivation be performed at the prescribed temperature, as the authors have found that 65°C is not adequate.

12. Thaw 40 μ l DH10B electrocompetent cells on ice. Make a 1:10 dilution of the reaction mix (step 11) with sterile water and mix with thawed cells.
13. Transfer mixture to a prechilled 0.1-cm electroporation cuvette and carry out the electroporation (*UNIT 1.8*). On a Gene Pulser apparatus, set the capacitance to 25 μ F, the voltage to 1.8 kV, and the pulse controller to 200 Ω .
14. Add 0.4 ml LB (or SOC) medium to the cuvette and transfer to a sterile culture tube. Incubate the tube with shaking at 37°C for 1 hr.

LB can be used, but SOC is known to ensure maximum transformation efficiency.

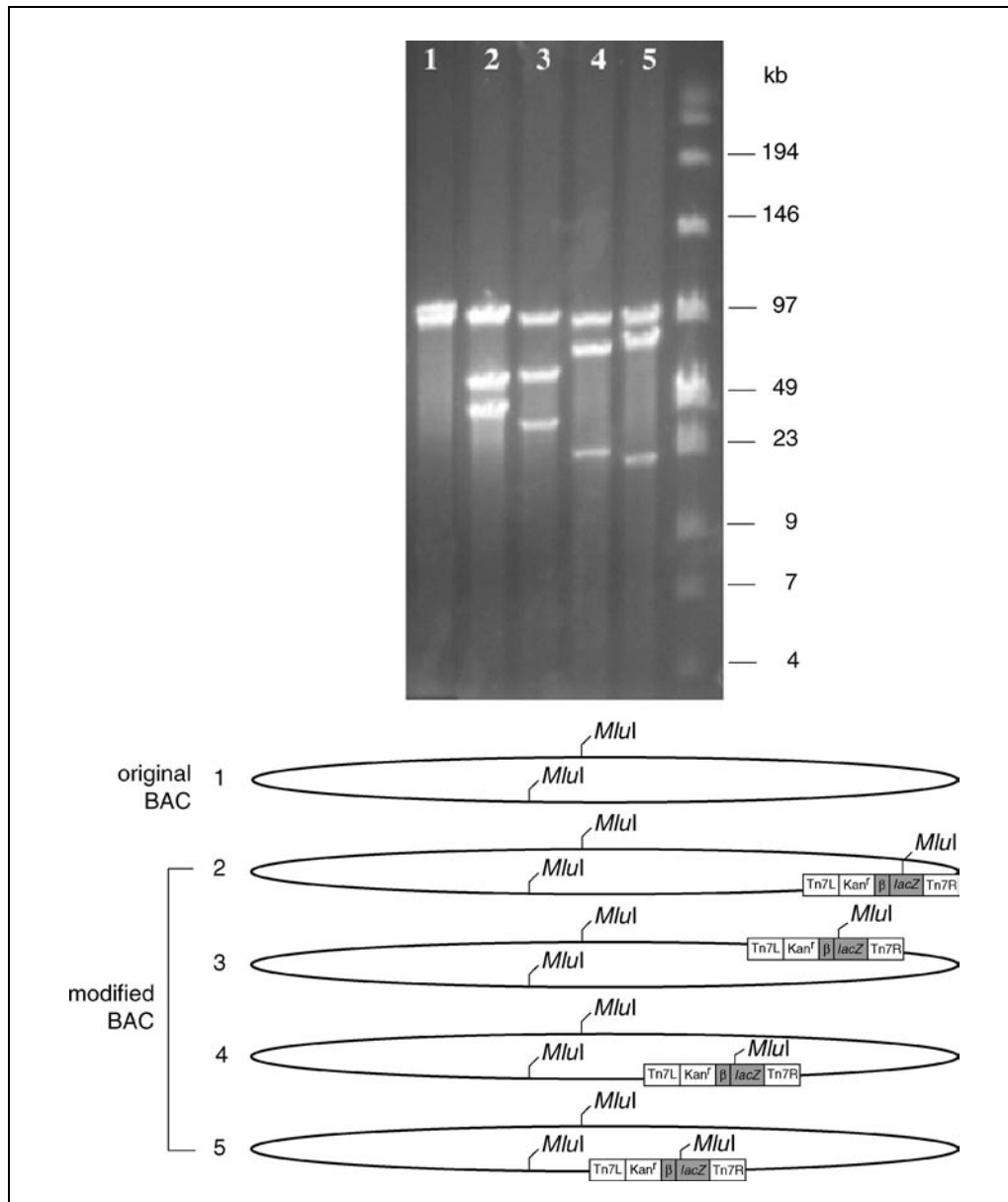


Figure 23.11.4 Restriction digest analysis of a sonic hedgehog BAC clone modified by the Tn7 transposon–based in vitro system. Modified BAC clones were first identified by PCR (data not shown). DNA from PCR-positive clones was digested with *Mlu*I and separated by pulsed-field gel electrophoresis on a 1% agarose gel. The original BAC (lane 1) contained two *Mlu*I sites and gave rise to two DNA fragments. The modified BACs gave rise to three *Mlu*I restriction fragments because of the presence of an additional *Mlu*I site in the transposon cassette. Because of the random nature of the integration event, the size of the modified *Mlu*I1 fragment varied from clone to clone.

15. Spread 100 μ l incubated cells on an LB plate with 20 μ g/ml kanamycin and 15 μ g/ml chloramphenicol and incubate at 37°C overnight.
16. Individually pick and inoculate colonies into 2 ml LB medium with 20 μ g/ml kanamycin and 15 μ g/ml chloramphenicol and incubate with shaking at 37°C overnight.

Verify insertion by PCR

17. Dilute 5 μ l overnight culture 1:10 using sterile water. Lyse 5 μ l diluted culture in a PCR tube by boiling for 3 min in a thermal cycler.
18. Use 5 μ l lysed culture as a template for PCR (*UNIT 15.1*) with 50 ng each of β -globin and *lacZ* primers and the following conditions:

30 cycles:	1 min	94°C
	1 min	55°C
	1 min	72°C.

19. Analyze product by gel electrophoresis (*UNIT 2.5A*).

The amplification product is ~350 bp.

Verify insertion by restriction digestion

20. Prepare BAC DNA from the original and modified BACs (see Support Protocol, steps 1 to 8) and digest with *MluI* or another restriction enzyme that cuts both the transposon insert and the BAC DNA.

MluI works well because it is a rare-cutting enzyme.

21. Analyze restriction fragments by pulsed-field gel electrophoresis (*UNIT 2.5B*; Fig. 23.11.4).

The authors typically run samples on a 1 \times TBE gel (using Bio-Rad PF-certified agarose) at 6 V/cm, 5- to 15-sec switch, for 20 hr at 14°C.

Verify insertion by sequencing (optional)

22. Sequence the modified BAC (see Chapter 7) using a primer directed against the 3' end of the *lacZ* gene to precisely map the transposon insertion site.

PREPARATION OF BAC DNA

Many methods are currently available to isolate BAC DNA (*UNITS 2.1B & 23.9*). This protocol describes a simple and reliable method of BAC DNA purification using phenol/chloroform extraction. BAC DNA purified by this method is suitable for many purposes including restriction digestion, sequencing, and (with the optional purification steps) pronuclear microinjection.

Materials

- Bacterial strain containing BAC of interest (see Basic Protocols 1 and 2 and see Internet Resources)
- LB medium (*UNIT 1.1*)
- Antibiotics appropriate for BAC strain of interest
- Resuspension solution: 50 mM Tris-Cl, pH 8.0 (*APPENDIX 2*)/10 mM EDTA
- 10 mg/ml RNase A
- Lysis solution: 200 mM NaOH (*APPENDIX 2*)/1% (w/v) SDS
- Neutralization solution: 3.0 M potassium acetate, pH 5.5
- 25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol, pH 6.7
- Chloroform

**SUPPORT
PROTOCOL**

75% and 100% (v/v) ethanol
TE buffer, pH 8.0 (APPENDIX 2)
3 M sodium acetate, pH 7.0 (for pronuclear microinjection only)
Microinjection buffer: 10 mM Tris·Cl, pH 7.4 (APPENDIX 2)/0.1 mM EDTA
(store ≤6 months at room temperature; for pronuclear microinjection only)
Size markers for gel electrophoresis (e.g., DNA Quanti-Ladder, OriGene Technologies; for pronuclear microinjection only)
PI-SceI and 10× reaction buffer (New England Biolabs; for pronuclear microinjection only)
37°C incubator, with shaking
Microcentrifuge, room temperature and 4°C
Pipet tip with large orifice
65°C heat block (for pronuclear microinjection only)
Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5A) and pulsed-field gel electrophoresis (see Basic Protocol 2)

NOTE: Reagents and water for pronuclear microinjection buffer must be tissue culture grade.

Prepare bacterial culture

For restriction digestion and Southern analysis

- 1a. Inoculate a single bacterial colony containing the BAC of interest into 2 ml LB medium with appropriate antibiotic(s) and incubate overnight at 37°C with shaking (225 rpm).
- 2a. Transfer 1.5 ml culture to a microcentrifuge tube, harvest cells by centrifuging 1 min at 10,000 × g, and resuspend pellet in 200 μl resuspension solution containing 2 μl of 10 mg/ml RNase A (100 μg/ml final).

For pronuclear microinjection

- 1b. Prepare an overnight culture as above, using a single colony in a volume of 10 to 20 ml LB medium with appropriate antibiotic(s).
- 2b. Harvest cells as above, but split culture into 5 to 10 tubes and process in parallel.

Isolate BAC DNA

3. Add 200 μl lysis solution and mix by gently inverting.

No incubation time is necessary.

Vortexing or vigorous shaking at this point will cause E. coli chromosomal DNA contamination.

4. Add 200 μl neutralization solution and mix by gently inverting. Centrifuge mixture 10 min at 12,000 × g, 4°C.

No incubation time is necessary after adding neutralization solution.

5. Transfer supernatant to a new microcentrifuge tube and add 500 μl of 25:24:1 phenol/chloroform/isoamyl alcohol. Vortex 5 sec and centrifuge 2 min at 12,000 × g, room temperature.

The mixture can be vortexed because supercoiled BAC DNA does not shear.

6. Transfer upper phase to a new tube and add 500 μl chloroform. Vortex 5 sec and centrifuge 2 min at 12,000 × g, room temperature.

7. Transfer upper phase to a new tube and add 900 μl of 100% ethanol. Centrifuge mixture 30 min at $12,000 \times g$, 4°C .
8. Carefully remove supernatant and wash pellet with 75% ethanol. Remove ethanol and let pellet air dry for 5 to 10 min. Resuspend pellet in 20 μl TE buffer.

This BAC DNA (1 to 2 μg) can be used for restriction digestion and subsequent Southern analysis.

Purify BAC DNA for pronuclear microinjection

9. Pool DNA samples into a single microcentrifuge tube.
10. Adjust volume to 400 μl with TE buffer and add 40 μl of 3 M sodium acetate and 800 μl of 100% ethanol.

The DNA will immediately precipitate and form an aggregate. This step will remove most, but not all, of the contaminating bacterial RNA.

11. Transfer aggregate to a new tube using a pipet tip with a large orifice. Pulse-spin aggregate down to the bottom of the tube.
12. To remove remaining trace amounts of bacterial RNA, dissolve pellet in 500 μl TE, add 1 μg RNase A, and incubate 30 min to 1 hr at room temperature.
13. Add 500 μl phenol/chloroform/isoamyl alcohol. Vortex 5 sec and centrifuge 2 min at $12,000 \times g$, room temperature.
14. Transfer upper phase to a new tube and add 500 μl chloroform. Vortex 5 sec and centrifuge 2 min at $12,000 \times g$, room temperature.
15. Transfer upper phase to a new tube and add 50 μl of 3 M sodium acetate and 900 μl of 100% ethanol. Centrifuge 30 min at $12,000 \times g$, 4°C .
16. Wash pellet with 75% ethanol, let air dry, and resuspend final pellet in 10 to 20 μl microinjection buffer.

BAC DNA can be stored at 4°C for up to 6 months. Freezing and thawing will cause shearing.

17. Run an aliquot on a 0.8% (w/v) agarose gel (UNIT 2.5A) with a known amount of size markers and estimate DNA concentration. Immediately before microinjection, dilute an aliquot of DNA to a final concentration of 0.5 to 1 ng/ μl with microinjection buffer.

*Although circular BAC DNA can be used for microinjection, the authors prefer to control the linearization site by digesting the BAC DNA with *PI-SceI* as described below. The recognition site for *PI-SceI* is extremely rare, occurring once in every 7×10^{10} base pairs of random sequence. It has been incorporated into the backbone of BAC vectors derived from CHORI.*

Linearize BAC DNA (optional)

18. Mix together the following:

2 μg BAC DNA
2 μl 10 \times reaction buffer
2 U *PI-SceI*
H₂O to a final volume of 20 μl .

Incubate ≥ 3 hr at 37°C and heat inactivate 20 min at 65°C .

19. Verify that digestion is complete by pulsed-field gel electrophoresis (see Basic Protocol 2, step 21).

20. Add the following to the tube:

80 μl H₂O
15 μl 3 M sodium acetate
300 μl 100% ethanol.

Centrifuge 30 min at 12,000 $\times g$, 4°C.

21. Wash pellet with 75% ethanol, let air dry, and resuspend DNA in 20 μl microinjection buffer.

22. Estimate DNA concentration on an agarose gel (step 17). Immediately before injection, dilute an aliquot to a final concentration of 0.5 to 1 ng/ μl using microinjection buffer.

The concentration should be ~100 ng/ μl after linearization.

Linearized BAC DNA can be stored at 4°C for up to 2 months.

COMMENTARY

Background Information

Bacterial artificial chromosomes (BACs) are a derivative of the *E. coli* F-factor-based plasmid. As an alternative to YACs (yeast artificial chromosomes), BACs have been developed for the construction of large genomic DNA libraries (Shizuya et al., 1992). Although they may propagate inserts approaching 500 kb in length, BAC insert sizes are typically 100 to 300 kb. BAC clones have several notable advantages over YACs. These include high stability and minimal insert rearrangements because of the F-factor genes *parA* and *parB*, which prevent more than one BAC from growing in an *E. coli* host. In addition, BACs are relatively easy to manipulate and propagate because they exist as supercoiled circular plasmids that are resistant to shearing. These advantages have made BACs the vector of choice for a variety of genomic-based research projects.

Despite a wide variety of applications, cloning strategies that rely on DNA restriction enzyme and DNA ligase reactions are often difficult to carry out because of limitations in the size of the DNA that can be manipulated or because of the lack of convenient restriction sites. Homologous recombination has provided an alternative approach to engineering DNA, and various systems have been developed to perform these manipulations in *E. coli*. These systems include recombination between two linear DNAs (Chartier et al., 1996), between linear and circular DNAs (Zhang et al., 1998; Lee et al., 2001), or between two circular DNAs (Gong et al., 2002). Many of the modification systems rely on targeting vectors that contain a conditional replication origin, so that recombination function can be transiently supplied.

In addition to these approaches, a new method to modify BACs was recently developed using in vitro transposition steps (Spitz et al., 2003). This protocol utilizes the Genome Priming System (GPS; New England Biolabs), which was initially developed for generating a population of DNA sequencing templates with randomly interspersed primer-binding sites. As the pGPS vector contains multiple restriction sites, it is highly conducive to cloning different types of inserts, including reporter constructs, and thus provides a rapid and convenient way to randomly introduce DNA of interest into BACs.

Critical Parameters and Troubleshooting

The recommended length of homology arms ranges from 300 bp to 500 bp. Homology arms >500 bp should also work well (although this has not been tested), but shorter ones may cause a decrease in the recombination frequency.

With electroporation, the transformation efficiency should approach $\sim 10^8$. Competent cells prepared from overgrown cultures (>0.8 OD₆₀₀) will result in a 10- to 100-fold reduction in transformation efficiency.

In transposition reactions, a starting molar ratio of $\sim 1:1$ targeting vector (pGPS1- β LacZ) to BAC is optimal. In a 20- μl reaction, 20 ng of targeting vector and 200 ng of BAC target usually works well. It is recommended to try several different ratios. Saturating amounts of the targeting vector will inhibit the reaction, however, and may result in an increased number of BAC clones carrying double insertions.

During the assembly step in the transposition reaction, TnsB binds to the Tn7 transposon

elements, and TnsA recognizes the TnsB:DNA complex. TnsC binds to the target DNA and interacts with TnsB. When these three proteins are properly bound in a two-DNA:three-protein complex, the strand transfer will occur. If the start solution is added before this assembly, the proportion of complicated products will greatly increase.

The transposition reaction must be stopped by heating at 75°C for 10 min to inactivate the transposase. Heating at 65°C for 20 min does not effectively disrupt the reaction complexes. Few or no recombinants will be obtained if this step is omitted or not executed properly.

Anticipated Results

In Basic Protocol 1, targeting can approach frequencies of 5% to 10%. If no recombinants are obtained, check the shuttle vector design (e.g., orientation of the homology arms). In Basic Protocol 2, the targeting efficiency ranges from 10% to 30%.

Time Considerations

Basic Protocol 1 can be executed in ~1 month (2 weeks for cloning homology arms into the shuttle vector, 1 day for making electrocompetent cells, 5 days for the co-integration step, and 7 days for the resolution step including confirmation of recombinants). Basic Protocol 2 can be executed in 7 to 8 days (3 days for preparing the targeting vector, 1 day for making electrocompetent cells, 3 to 4 days for the transposition step including the confirmation of recombinants). The Support Protocol can be executed in 1 day, with bacterial cultures started the evening before. BAC preparation for pronuclear microinjection may take another day or so.

Literature Cited

Chartier, C., Degryse, E., Gantzer, M., Dieterle, A., Pavirani, A., and Mehtali, M. 1996. Efficient generation of recombinant adenovirus vectors by

homologous recombination in *Escherichia coli*. *J. Virol.* 70:4805-4810.

Gong, S., Yang, X.W., Li, C., and Heintz, N. 2002. Highly efficient modification of bacterial artificial chromosomes (BACs) using novel shuttle vectors containing the R6K γ origin of replication. *Genome Res.* 12:1992-1998.

Lee, E.C., Yu, D., Martinez de Velasco, J., Tessarollo, L., Swing, D.A., Court, D.L., Jenkins, N.A., and Copeland, N.G. 2001. A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* 73:56-65.

Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. 1992. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci. U.S.A.* 15:8794-8797.

Spitz, F., Gonzalez, F., and Duboule, D. 2003. A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell* 113:405-417.

Zhang, Y., Buchholz, F., Muyrers, J.P., and Stewart, A.F. 1998. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.* 20:123-128.

Internet Resources

<http://www.genome.ucsc.edu/cgi-bin/hgGateway>

Individual BAC clones that overlap a gene or genomic region of interest can be identified using the Genome Browser Gateway of the Genome Bioinformatics Group of the University of California Santa Cruz.

<http://bacpac.chori.org/>

BACs can be ordered from the BACPAC Resources Center at the Children's Hospital Oakland Research Institute.

Contributed by Yongsu Jeong and

Douglas J. Epstein

University of Pennsylvania School
of Medicine

Philadelphia, Pennsylvania

Regulation of Transgene Expression Using Tetracycline

UNIT 23.12

**BASIC
PROTOCOL**

Transgenesis has been an efficient means of remodeling the mouse genome, and numerous gain-of-function and loss-of-function models have been and are being made. Although transgene expression was initially driven by systemic promoters, it quickly became apparent that more precise tools were needed, and the development of organ-specific or cell-type-specific promoters capable of driving high levels of transgene expression now provides investigators with the reagents needed to carry out precise transgenic expression that is targeted to a particular cell type.

Although a spectrum of different promoters have resulted in the spatial control of transgenic expression, by themselves they do not allow the ability to *reversibly* control expression at specific developmental times, and the developmental patterns of expression are dictated by the promoter construct used. For example, in the heart, the α -myosin heavy chain promoter has been widely used to drive cardiomyocyte expression, but it is only activated at 18 days of gestation and remains active in the adult (Ng et al., 1991). Thus, despite cell-type- and developmental-stage-specific expression, this approach limits the investigator to irreversible transgene activation at levels dictated by the copy number present (James and Robbins, 1997).

To circumvent these issues, inducible and reversible transgene expression systems have been developed (Shockett et al., 1995; Tsai et al., 1998). These systems normally are bigenic, with the activator and responder transgenes being driven in the same cell type in the animal. The most widely used system is based upon a transactivator fusion protein consisting of the *E. coli* tetracycline repressor linked to the strong transcriptional activation domain derived from the VP16 protein of herpes simplex virus, and both tet-on (the transgene is on in the presence of the drug) and tet-off systems have been described (Gossen et al., 1994; Baron et al., 1997). Although these systems have been used with great success in many different cell lines, their use in transgenic animals has been more limited, mainly because of the technical difficulties in obtaining a bigenic line that is not leaky, but that can be induced to produce the transgenic product at high levels and can subsequently be inactivated. This often necessitates the production of as many as 15 to 25 lines before usable ones are generated. Considering the expense and labor needed for even the preliminary analyses of the lines, this precludes the easy and widespread use of inducible systems for organ-specific or cell-type-specific expression within a transgenic context.

This protocol outlines the parameters involved in preparing a robust, reversible, and inducible system capable of driving high levels of transgenic expression in a cell-type-specific manner. Because these systems are necessarily cell-type-specific, a procedure is described that results in cardiac-specific inducible expression as an example. This protocol describes the administration of doxycycline to transgenic mice as a way of regulating gene expression. Previous units have described the cloning and production of DNA to be used in making transgenic mice (UNIT 23.10) as well as the generation of the transgenic animals (UNIT 23.9). The transgene consists of the coding region of the gene to be inducibly expressed cloned behind a modified attenuated promoter that is capable of being activated when a *trans*-acting factor driven from the second transgene binds to it. Two choices are available in selecting a doxycycline-dependent *trans*-activating factor; one is active in the absence of doxycycline (the tet-off system, tTA), while the

**Manipulating the
Mouse Genome**

23.12.1

Contributed by James Gulick and Jeffrey Robbins

Current Protocols in Molecular Biology (2005) 23.12.1-23.12.8

Copyright © 2005 by John Wiley & Sons, Inc.

Supplement 71

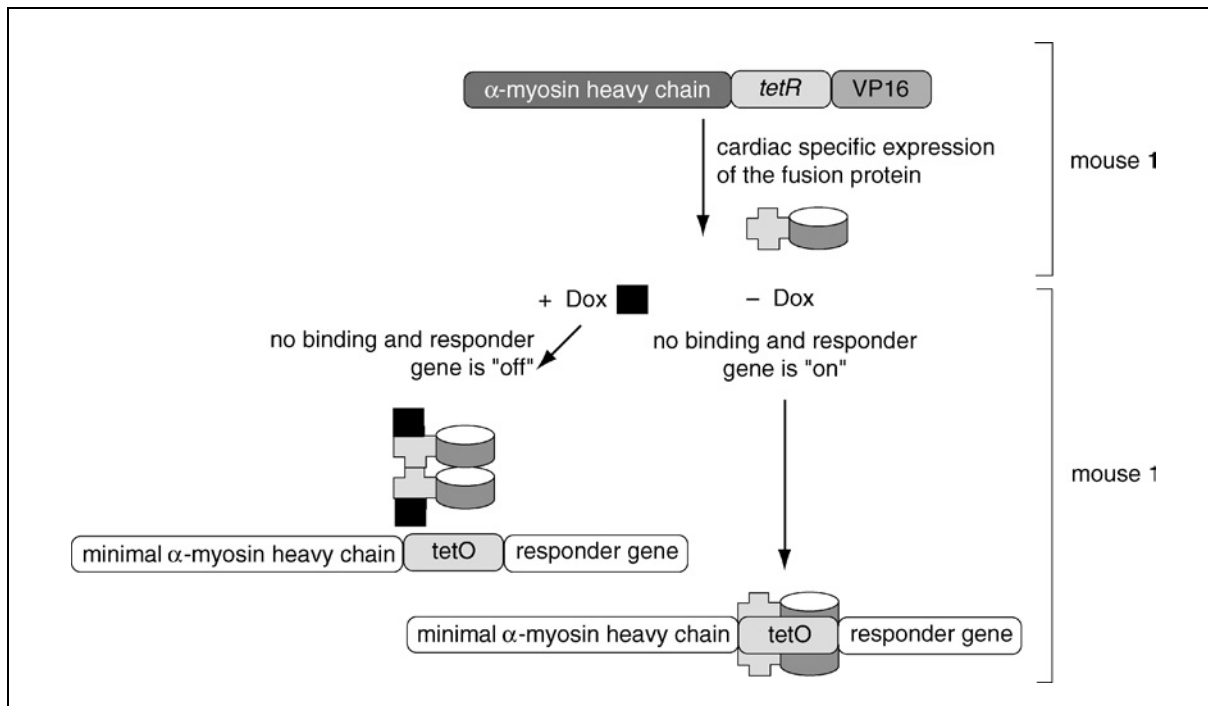


Figure 23.12.1 Cell-type-specific inducible transgene expression. The two lines of mice needed for the experiments are shown. Mouse 1 contains the cell-type-specific promoter driving the fusion tet-VP16 protein that, in the presence of tetracycline (or doxycycline, a more stable form), will not bind to the responder promoter in mouse 2. The responder promoter is not active in the absence of tet-VP16 binding.

other is active in the presence of doxycycline (the tet-on system, rtTA). This protocol specifically describes inducible cardiac-specific expression using the tet-off system. The components of the system are illustrated in Figure 23.12.1, and a flow chart for the method is illustrated in Figure 23.12.2.

Materials

- Mouse line that expresses tTA in cell- or tissue-specific manner
- Mouse line that contains the cDNA of the gene of interest cloned behind the responder promoter
- Ear clip digest buffer (see recipe)
- PCR primers for screening the tet-off (tTA) construct:
 - AGCGCATTAGAGCTGCTTAATGAGGTC (forward)
 - GTCGTAATAATGGCGGCATACTATC (reverse)
- PCR primers for the cDNA being expressed (see *UNIT 15.1* for principles of primer design)
- Drinking water containing doxycycline (0.2 to 2 g/liter; see recipe) or prepackaged, irradiated mouse chow containing doxycycline (e.g., 625 mg/kg; Harlan Teklad)
- Light-protected (dark-glass or foil-wrapped) water bottles for mice
- 4.5-cm stainless steel ear punch (Fine Science Tools)
- Additional reagents and equipment for maintenance and care of transgenic animals (*UNITS 23.8 & 23.10*), PCR (*UNIT 15.1*), northern blotting (*UNIT 4.9*), quantitative PCR (*UNIT 15.7*), and immunoblotting (*UNIT 10.8*)

Mate and maintain animals

1. Mate the mice containing the transactivator and the responder components to one another.

Transgenic animal care and maintenance is covered in UNITS 23.8 and 23.10. Ideally the mice should be housed in a virus- and pathogen-free facility.

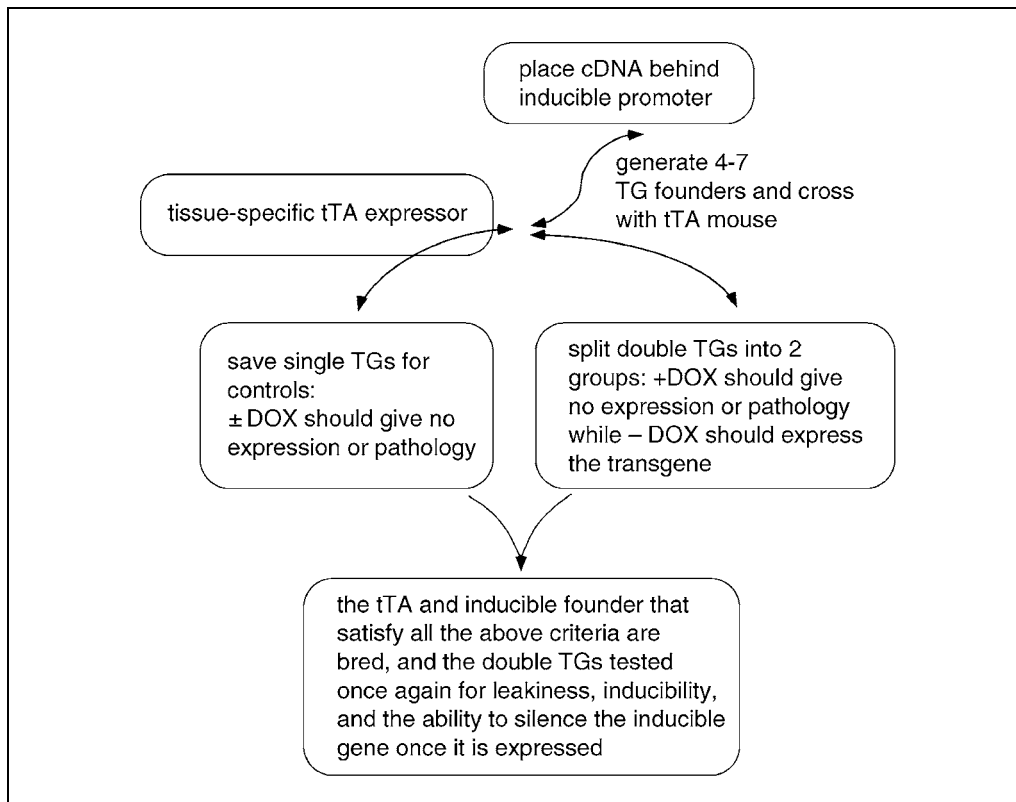


Figure 23.12.2 Flow chart for tet-off inducible transgene expression.

The myosin heavy chain promoter has been used to drive expression of the tTA in a cardiac-specific manner as previously described (Sanbe et al., 2003). For the responder side of the inducible system, a modified, attenuated myosin heavy chain promoter was used to drive expression of the specific cDNA, providing an additional level of specificity. Both constructs needed for cardiomyocyte-specific expression are available from the authors at jeff.robbs@cchmc.org.

2. If necessary, provide dams with doxycycline after mating to shut off transgene expression in gestating pups. See step 8 for administration routes and concentrations.

Some of the pups will be positive for both arms of the system. If the gene that is being expressed is expected to be embryonic-lethal, it may be necessary to shut off the transgene until later in development. Doxycycline can penetrate the placental barrier, inactivating transgene expression during gestation.

Genotype pups by PCR

3. When the pups are about 10 to 12 days of age, clip one ear using a 4.5-cm stainless steel ear punch for identification purposes. Retain the clipped ear tissue for genotyping.
4. Digest 0.2 to 2 mg of tissue from the ear clip overnight at 65°C in 100 µl ear clip digest buffer.
5. At end of incubation boil for 10 min to inactivate the proteinase K.
6. Use 1 µl of the reaction mix in a 50-µl amplification reaction (UNIT 15.1). Perform PCR to screen the tet-off construct using the following cycling program:

1 cycle:	3 min	94°C	(initial denaturation)
35 cycles:	20 sec	94°C	(denaturation)
	30 sec	65°C	(annealing)
	30 sec	72°C	(extension)
1 cycle:	10 min	72°C	(final extension).

7. Perform a separate amplification reaction to screen the cDNA being expressed, using the appropriate primers and PCR conditions.

See UNIT 15.1 for additional details on PCR amplification.

Administer doxycycline to double-transgenic animals

- 8a. *To administer doxycycline via water:* To keep transgene expression off, supply mice with drinking water containing doxycycline, in dark bottles, changing the water at least three times per week.

Dark water bottles are used, as doxycycline is light-sensitive. If not available, the water bottles can be wrapped in foil, although the mice have a tendency to shred the foil.

- 8b. *To administer doxycycline via food:* Supply commercial prepackaged, irradiated mouse chow containing the appropriate concentration of doxycycline. Change food weekly.

The authors have used chow containing 625 mg/kg doxycycline successfully, but chow containing higher dosages can be purchased if necessary. Gamma-irradiated food allows the doxycycline to be fed to mice housed in virus- and pathogen-free facilities. Since the doxycycline may break down over time, the food should be stored in a cool, dry place such as a cold room, used within 6 months of initial purchase, and changed regularly.

Analyze transgene response

9. Determine levels of RNA resulting from transcription of the transgene by northern analysis (UNIT 4.9) or quantitative PCR (UNIT 15.8).

Usually, the 3' untranslated region of the transgene consists of a viral polyadenylation signal, such as one derived from SV40, which can be probed on a northern blot. In the case of the α -myosin heavy chain responder, a human growth hormone (hGH) poly(A) signal was used, but, as this sequence is also present in the α -myosin heavy chain inducer line, a probe specific for the cDNA should be used instead.

10. Monitor levels of protein resulting from translation of the transgene by immunoblotting (UNIT 10.8), using antibodies specific for the translated protein to detect changes in protein levels.

Under certain circumstances, it may be necessary to add a tag to the cloned gene so that the translation of the protein can be verified.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Drinking water containing doxycycline

H₂O containing:
0.2 to 2 g/liter doxycycline
1% (w/v) sucrose
Use dark (or foil-wrapped) bottles in cages
Change at least three times per week
Prepare fresh solution for each change

The amount of doxycycline to be used is determined empirically. The authors have found that 0.2 g/liter works well in most cases, turning a lacZ reporter off in the tet-off system. In a few cases, 1 g/liter was necessary, possibly due to the responder lines containing relatively more copies of the transgene or due to chromosomal influences at the site of transgene integration. The authors have noted no specific adverse effects from higher doses of doxycycline except for a 10% to 15% reduction in liver mass. Adding doxycycline to the drinking water of pregnant or nursing dams at 2 g/liter will inactivate transgene expression in utero or until the pups are weaned.

Ear clip digest buffer

1 × PCR buffer (*UNIT 15.1*) containing:
0.45% (v/v) Nonidet P-40 (NP-40)
0.45% (v/v) Tween 20
10 mg/ml proteinase K
Store indefinitely at -20°C

COMMENTARY

Background Information

An overview of the technical basis for a tet-based inducible system can be found in *UNIT 16.14*. While good results can often be achieved in tissue culture with both the transactivator and responder components being driven by viral promoters, with transgenic animals, additional levels of specificity are necessary. Thus, despite the obvious advantages of reversibly altering cardiac transgene expression, the number of successful studies carried out in vivo with inducible cell-type-specific transgene expression remains limited. Current systems are plagued by a need for bi-genic breeding and by the large number of lines needed before a nonleaky, inducible line is isolated. Although a heterologous, viral-based minimal promoter such as the cytomegalovirus (CMV) promoter is often used in the responder line to drive inducible expression, the authors of this unit have found that this exacerbates the difficulties, as those promoters are often position-dependent or can be methylated and subsequently inactivated as the line is propagated from one generation to the next. The TetR-VP16 fusion protein, at least in some tissues, is not benign. Bujard and colleagues have explored the potential of creating modified proteins that exhibit decreased generalized transcriptional squelching, improved DNA-activator binding characteristics, increased sensitivity, and expanded induction range (Baron et al., 1997; Urlinger et al., 2000). Although these transactivators appear to work very well in cells, the authors of this unit have now tested a number of the constructs in transgenic mice and found that, when expressed in the heart, they cause cardiovascular disease. The single transactivator lines must therefore be tested carefully for pathology in the target tissue before being used.

This unit describes a “tet-off” system, but “tet-on” systems have also been described and, in a few instances, used successfully in transgenic animals (Russell et al., 2004). However, in the authors’ experience with the cardiovascular system, these systems are more often leaky, and also difficult to induce fully,

necessitating the creation of a large number of lines before a satisfactory one is found. Other inducible systems have also been adopted successfully for transgenic use (Wang et al., 1997), but these systems use less innocuous inducers that can have undesired pleiotropic effects on the animals. At this point, the tetracycline-based systems have clear advantages for long-term animal-based studies because the required drug treatment is minimally intrusive on the animals’ general physiology.

Critical Parameters

Activator line

The availability or creation of a stable activator line is critical. The line should produce enough activator homogeneously in the target cell population so that, in the presence of the inducer, gene expression is rapidly initiated in a cell-type-specific, homogeneous manner. The authors have found that, for the cardiac system, very low levels of expression yielding only trace amounts of protein are desirable, as higher levels invariably lead to morbidity and detectable cardiac pathology, despite the use of supposedly more benign classes of fusion proteins (Sanbe et al., 2003). Once an activator line is obtained, it should be used exclusively, and the animals should be bred only when they are young adults, so as to avoid more robust variants that might be gradually selected for as the colony ages. Although it is desirable to carry the transactivator transgene in a homozygous animal for ease of generating the double-transgenic animals necessary for the experimental cohorts, this may not be possible because of the transactivator’s potential toxicity. For the cardiac-specific transactivator line prepared in the authors’ laboratory (Sanbe et al., 2003), the homozygote tTA mice are viable, with the males being used for breeding. The mice cannot be maintained as homozygous lines, however, because the dams die shortly after giving birth. In addition, detectable cardiovascular disease is present as early as 5 months after birth in these animals, and they are normally not used for breeding after 20 weeks.

The choice of the fusion protein, either “tet-on” or “tet-off,” and the promoter that drives its expression is critical. Clearly, it would be more desirable to use a “tet-on” system in which drug administration is only necessary for transgene induction. Although the modified fusion proteins that result in a tet-on system have been described, and only a few amino acid changes are needed to convert a tet-off transactivator to a tet-on protein, in the authors’ hands these lines have been problematic. Even with carefully engineered promoters in both the activator and responder lines, they have invariably shown leaky expression or inefficient induction. Additionally, the tet-on transgene is quite toxic when expressed in the heart, even more so than the tet-off transactivator. Therefore, if the tet-on system is used in its present stage of development, the investigator should be prepared to generate a large number of lines before finding one that works satisfactorily. In contrast, the tet-off system described in this unit often requires breeding the transactivator line to only a few responder lines (see below) before a satisfactory bigenic animal is isolated.

Many cell-type-specific promoters have been described and used to make transgenic animals. The promoter driving the tTA activator should drive uniform expression in only the target cells, so that the responder transgene is activated in a specific cell background. Thus, the more restricted the expression pattern of the promoter that is used, the more precise the inducible experiment will be. The promoter should be stable and not subject to large variations in expression as a result of feedback inactivation/activation in the presence of a pathological response, as this could have the effect of modifying expression of the activator during the course of the experiment. High levels of expression are not a requirement, as only very low levels of transactivator are needed, or even desired, due to the toxic effects that may be observed. Thus, weak promoters can be used as long as they are expressed in all of the target cells and do not show a patchy expression pattern.

Responder line

The responder arm of the bigenic system normally consists of a minimal CMV promoter linked to multiple copies of the tet operator. Depending upon the target tissue, this general strategy can also present a series of significant experimental problems. For cardiac-specific inducible expression, although it was possible to obtain lines of mice that showed minimal expression levels when noninduced

and responded appropriately during induction, isolating a line that displayed those characteristics necessitated the production of many (9 to 23) lines of animals. Thus, transgenic mice obtained with such constructs only rarely yield the necessary inductive range, presumably because of the sensitivity of the minimal CMV promoter construct to chromosomal context. To circumvent these difficulties and decrease the number of lines needed before a satisfactory one is obtained, the authors redesigned this part of the system around the well-characterized α -myosin heavy chain promoter, the goal being to create a construct from the endogenous DNA that showed low background expression levels, that could be highly induced, and that was relatively insensitive to chromosomal location. This promoter has a potential advantage over the other minimal promoter constructs in that it also is highly cardiac-specific and appears to be relatively independent of chromosomal context in its expression pattern (Palermo et al., 1996). For any tissue-specific expression, the object is to create a minimally active promoter that retains these characteristics. For adult cardiomyocyte-specific expression, a successful strategy involved taking an already cell-type-specific promoter and crippling it by ablating those sites responsible for high levels of expression. Thus, for the α -myosin heavy chain promoter, three GATA sites and two thyroid-like response elements, each of which plays a role in maintaining high levels of cardiac-specific expression, were ablated (Subramaniam et al., 1993; Robbins et al., 1995). When induced, this otherwise crippled promoter drives high levels of expression. With only a limited number of lines (three to six) being produced, one or more show no expression of the transgene in the single-transgenic animal or in double-transgenic mice in the presence of doxycycline, but high levels of expression in double-transgenic animals when the drug is withdrawn. The authors believe that this general approach, using a crippled endogenous promoter rather than an ectopic viral promoter, results in a more robust system.

Mouse strains

Strain variation can play a significant role in modulating phenotype. If the activator or responder lines are in different strains, there is no guarantee that they will behave as anticipated, either because of modifier effects or because of patterning of the exogenous DNA in the different genetic background. Thus, each transgene should be placed in the same strain,

and the background maintained in the bigenic animal.

Troubleshooting

Activator lines

For cardiac-specific inducible expression, a stable and robust system has been described in this unit, which should lead to the rapid growth of the number of inducible lines for this organ system. Development of inducible transgenic systems for other cell types has also been described (e.g., unique cell types in the lung: Clark et al., 2001; Perl et al., 2003), and the list is expected to grow. However, if an activator line is not available for the desired cell type, its development can be difficult, and many different promoters and combinations with different tet-on or tet-off constructs may need to be tried before a line with the desired characteristics is obtained. If no transactivator protein is detected, the sensitivity of the immunoblot should be confirmed either with recombinant protein or with extracts obtained from previously characterized transactivator lines known to effectively induce expression. Even if no protein can be detected, the transactivator may be present in sufficient quantities to activate a responder transgene, and should be bred to an indicator line in order to determine if a biological read-out can be obtained.

The lack of production of detectable protein in the transactivator line could also be due to illicit splicing between the promoter and transactivator DNAs, so the transcript species may need to be sequenced in order to determine where the illicit splice junction lies so that it can be removed. Once again, the potential pathology due to the transactivator's expression needs to be carefully considered, with the organ system or cell type under study carefully followed over the animal's lifetime. Only lines that yield no detectable pathology in the heterozygote state should be chosen for the transactivator production line.

Responder lines

The responder lines should be carefully tested for expression in the absence of the transactivator or transactivator/inducer. If expression is observed, the investigator will need to decide whether the levels can be tolerated or additional lines need to be generated. In this cardiac-inducible system, only a limited number of lines are needed in order to generate one with the desired characteristics (Sanbe et al., 2003). In a number of cases, the lines are not leaky, but cannot be induced.

Generating additional lines invariably solves this problem. Finally, for some lines, it is impossible to silence the transgene after it has been activated, or a second cycle of transgene activation/inactivation is somewhat attenuated (J. Gulick and J. Robbins, unpub. observ.). The reasons for these phenomena are unknown, but they appear to be line-dependent, so that any difficulties encountered can usually be circumvented by generating additional lines.

Anticipated Results

In the authors' experience, the transactivator line, once established, is stable and can be repeatedly used as long as the mouse strains are kept constant. The responder line—once a stable construct is established, such as that reported for the cardiac-specific induction detailed in this unit—is also stable and quite robust in its ability to drive both strong biological effectors and structural proteins that are abundant in the cardiomyocyte. Normally, five or fewer lines will yield at least one that is not leaky and can be effectively induced and subsequently silenced.

The time course of induction and silencing will, of course, depend upon the half-life of the transgene's RNA and protein, but expression levels are usually detectable within 24 to 48 hr after the drug is withdrawn. Even long-lived RNAs and proteins can be silenced within a week or so after the drug is added in the tet-off system.

Time Considerations

These are time-consuming experiments and take considerably longer than standard, cell-type-specific transgenesis. The flow diagram (Fig. 23.12.2) assumes that a suitable activator mouse is available and that a generic responder promoter also exists. In that case, a responder line needs to be designed and made; this can be a lengthy process that is dependent upon the responder promoter's efficacy. Responder promoters that have been tested in other organ systems may very well work in a heterologous cell or tissue, and should be tried before undertaking a completely new design. Even so, the particular construct under study will need to be made and placed into the responder locus, and the mice will then have to be generated—a process that can take anywhere from 2 to 4 months. Subsequently, that line needs to be tested for transgene expression in the absence of the second transactivator transgene to confirm that leaky expression has not occurred. This can take another 2 months. Coincident with those experiments,

the double transgenics are bred and this cohort is split so that transgene expression can be induced in some of the animals. Finally, after the lines have been chosen on the basis of nonleaky expression, inducibility, and ability to silence the transgene after it has been activated, the double-transgenic cohorts need to be raised in sufficient quantities for the experiments. The complete process can easily take from 6 months to 1 year.

Literature Cited

- Baron, U., Gossen, M., and Bujard, H. 1997. Tetracycline-controlled transcription in eukaryotes: Novel transactivators with graded transactivation potential. *Nucl. Acids Res.* 25:2723-2729.
- Clark, J.C., Tichelaar, J.W., Wert, S.E., Itoh, N., Perl, A.K., Stahlman, M.T., and Whitsett, J.A. 2001. FGF-10 disrupts lung morphogenesis and causes pulmonary adenomas in vivo. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 280:L705-L715.
- Gossen, M., Bonin, A.L., Freundlieb, S., and Bujard, H. 1994. Inducible gene expression systems for higher eukaryotic cells. *Curr. Opin. Biotechnol.* 5:516-520.
- James, J. and Robbins, J. 1997. Molecular remodeling of cardiac contractile function. *Am. J. Physiol.* 273:H2105-H2118.
- Ng, W.A., Grupp, I.L., Subramaniam, A., and Robbins, J. 1991. Cardiac myosin heavy chain mRNA expression and myocardial function in the mouse heart. *Circ. Res.* 68:1742-1750.
- Palermo, J., Gulick, J., Colbert, M., Fewell, J., and Robbins, J. 1996. Transgenic remodeling of the contractile apparatus in the mammalian heart. *Circ. Res.* 78:504-509.
- Perl, A.K., Hokuto, I., Impagnatiello, M.A., Christofori, G., and Whitsett, J.A. 2003. Temporal effects of Sprouty on lung morphogenesis. *Dev. Biol.* 258:154-168.
- Robbins, J., Palermo, J., and Rindt, H. 1995. In vivo definition of a cardiac specific promoter and its potential utility in remodeling the heart. *Ann. N.Y. Acad. Sci.* 752:492-505.
- Russell, L.K., Mansfield, C.M., Lehman, J.J., Kovacs, A., Courtois, M., Saffitz, J.E., Medeiros, D.M., Valencik, M.L., McDonald, J.A., and Kelly, D.P. 2004. Cardiac-specific induction of the transcriptional coactivator peroxisome proliferator-activated receptor gamma coactivator-1alpha promotes mitochondrial biogenesis and reversible cardiomyopathy in a developmental stage-dependent manner. *Circ. Res.* 94:525-533.
- Sanbe, A., Gulick, J., Hanks, M.C., Liang, Q., Osinska, H., and Robbins, J. 2003. Reengineering inducible cardiac-specific transgenesis with an attenuated myosin heavy chain promoter. *Circ. Res.* 92:609-616.
- Shockett, P., Difilippantonio, M., Hellman, N., and Schatz, D.G. 1995. A modified tetracycline-regulated system provides autoregulatory, inducible gene expression in cultured cells and transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* 92:6522-6526.
- Subramaniam, A., Gulick, J., Neumann, J., Knotts, S., and Robbins, J. 1993. Transgenic analysis of the thyroid-responsive elements in the alpha-cardiac myosin heavy chain gene promoter. *J. Biol. Chem.* 268:4331-4336.
- Tsai, S.Y., O'Malley, B.W., DeMayo, F.J., Wang, Y., and Chua, S.S. 1998. A novel RU486 inducible system for the activation and repression of genes. *Adv. Drug Deliv. Rev.* 30:23-31.
- Urlinger, S., Baron, U., Thellmann, M., Hasan, M.T., Bujard, H., and Hillen, W. 2000. Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity. *Proc. Natl. Acad. Sci. U.S.A.* 97:7963-7968.
- Wang, Y., DeMayo, F.J., Tsai, S.Y., and O'Malley, B.W. 1997. Ligand-inducible and liver-specific target gene expression in transgenic mice. *Nat. Biotechnol.* 15:239-243.

Contributed by James Gulick and
Jeffrey Robbins
Cincinnati Children's Hospital, University
of Cincinnati
Cincinnati, Ohio

Overview of Agents from Combinatorial Nucleic Acid and Protein Libraries

UNIT 24.1

UNIT 24.1 will be published in an upcoming supplement.

Generation and
Use of
Combinatorial
Libraries

24.1.1

Design, Synthesis, and Amplification of DNA Pools for Construction of Combinatorial Pools and Libraries

This unit describes the design, synthesis, and amplification of a random sequence DNA pool. Functional nucleic acid-binding or catalytic species can be selected from these random sequence pools. In designing the DNA pool, careful consideration should be given both to the degree of randomization and the length of the random sequence region (see Strategic Planning). Following pool design, chemical synthesis on a commercial DNA synthesizer will yield a single-stranded DNA pool. The newly synthesized oligonucleotide pool can then be purified (see Basic Protocol 1). Prior to amplification, the initial complexity of the pool should be determined (see Support Protocol 1), the skewing of the pool should be determined (see Support Protocol 2), and amplification reaction conditions should be optimized (Support Protocol 3). If the nascent synthetic oligonucleotide is judged to be suitable for large-scale amplification, it can be enzymatically converted into a double-stranded DNA library (see Basic Protocol 2). Multiple copies of a single-stranded DNA pool can be derived from each double-stranded DNA library, or the library can be transcribed to yield a RNA pool or a modified RNA pool (see UNIT 24.3). Figure 24.2.1 outlines the procedure.

NOTE: The above comments and the protocols in this chapter are most relevant to the design of pools from which nucleic acids with binding activities (aptamers) or other enzymatic activities can be selected. However, many of the concepts and procedures are relevant for the design and synthesis of nucleotide pools that encode peptides of variable sequence. Considerations specific to the construction of nucleotide pools designed to encode peptides of variable sequence are given in UNIT 23.4.

STRATEGIC PLANNING

Designing the Initial DNA Pool

The nucleic acid pools used for in vitro selection experiments typically contain a randomized central core flanked by constant sequences that are required for enzymatic manipulations, such as PCR amplification, in vitro transcription, or restriction digestion (see also Fig. 24.2.2).

Since a pool is relatively expensive to synthesize, both in terms of time and cost, some effort should be devoted to pool design. There are many subtle parameters to consider that can greatly influence the outcome of a selection experiment, including the degree of randomization, pool length, and pool modularity (see Table 24.2.1 for references to selection experiments that have previously been successfully executed with different types and sizes of pools).

Type of selection and degree of randomization

Most researchers who carry out in vitro selection experiments wish to either better define or optimize a known binding site (binding-site selection), or to identify a nucleic acid that binds a particular ligand or site (aptamer selection; UNIT 24.3). Each of these tasks in turn requires the synthesis of different types of pools. The sequences and structures that contribute to known binding sites are frequently best defined by selections that start from partially randomized pools. One example of binding-site definition that started from a partially randomized pool was a selection that defined critical residues of the Rev-respon-

sive element (RRE) of HIV-1 Rev (Bartel et al., 1991). This experiment is also described in more detail below. Biased pools can also be used for the optimization of a previously isolated motif. For example, aptamers that could bind to the Rex protein of HTLV-1 were selected from a partially randomized pool based on the wild-type Rex-binding element (XBE) but in the end bound Rex 9-fold better than the XBE (Baskerville et al., 1995).

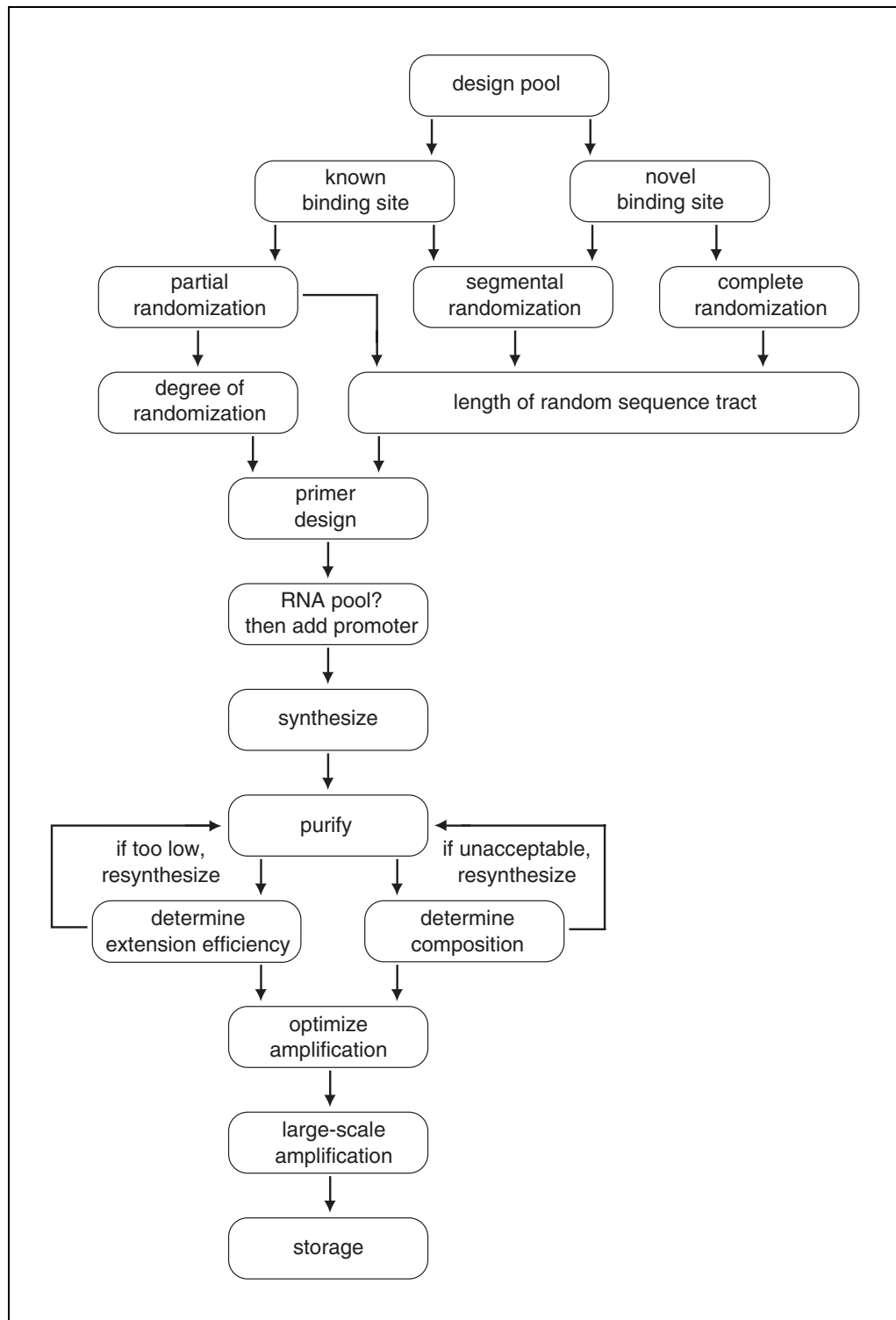


Figure 24.2.1 Flow chart outlining pool design, synthesis, and large-scale amplification.

Intermediate between partially random and completely random sequence pools are segmentally random sequence pools. In a segmentally random pool, short tracts of sequence are completely randomized. Segmental randomization thus allows all possible sequences within a short region or set of residues to be examined. Thus, if a natural binding site is known, but a portion of that binding site is suspected to be particularly important for function, then a segmentally random pool can be used to identify all possible, functional sequences within the wild-type sequence context. For example, Tuerk and Gold (1990) selected aptamers that bound T4 DNA polymerase from a pool that contained 8 random sequence positions flanked by wild-type residues. Similarly, many binding sites are known to be presented within a particular structural context, such as a stem-loop or stem-bulge structure. In these cases, a portion of the structure can be completely randomized, and all possible functional stem-loops or stem-bulges can be identified. For example, the Rev-binding element was known to form a stem-internal loop-stem structure. Giver et al. (1993) segmentally randomized only the internal loop portion of the structure and selected Rev-binding species. Many of the anti-Rev aptamers had sequences that were significantly different than the wild-type, yet were still presented in the context of a stem-internal loop-stem structure.

Partially random (doped) pool design (binding site selection)

The most important issue in the synthesis of a doped pool is the level of randomization (the probability of sequence substitution/position). As a general rule, the substitution frequency of a doped pool should roughly correspond to the number of positions thought to be required for function. For example, if 10 residues within a nucleic acid binding site are thought to be functional, then the rate of substitution might be set to yield single mutants at least half the time. If the substitution frequency is set too low, there may be too few varying residues or combinations of residues to yield information about functional sequences or structures. In contrast, if the substitution frequency is set too high, the sequence space nearest the wild-type motif will only be sparsely sampled, and many of the highly mutated molecules may be nonfunctional because their sequences will have diverged too far from the wild-type.

For example, an in vitro genetic analysis has been used to uncover the critical structural interactions between the HIV-1 Rev protein and its primary RNA binding site, the Rev-binding element (Bartel et al., 1991). The RBE had previously been mapped by deletion analysis to a short segment of HIV-1. Bartel and his co-workers assumed that the minimal RBE was smaller even than the region identified by deletion analysis, and thus decided to heavily dope a portion of a 66-nucleotide sequence at a frequency of 35% substitution/position. The initial RRE library contained $\sim 10^{13}$ molecules that had an average of 23 substitutions/template (0.35 probability substitution/position \times 66 positions = ~ 23 substitutions); less than 1 in 10^{12} molecules were completely wild-type. Following selection, a 20-nucleotide core-binding site within the 66-nucleotide pool was readily defined by sequence conservations and covarying residues. A lower substitution rate might not have precisely defined the relatively small binding site, while an even higher substitution rate might have created a mutational load that would have limited the selection of functional molecules or even have allowed the selection of novel, non-wild-type anti-Rev aptamers (Giver et al., 1993; Tuerk and MacDougall-Waugh, 1993). Conversely, if the binding site were larger than originally hypothesized, the relatively high rate of substitution might have meant that few functional molecules could have survived the selection unscathed.

The number and type of sequence substitutions, as opposed to the probable target size for mutation, can also be used to plan the synthesis of a doped sequence pool, as described by the following equations. Typically, a 1- μ mol synthesis of a 100-residue template yields

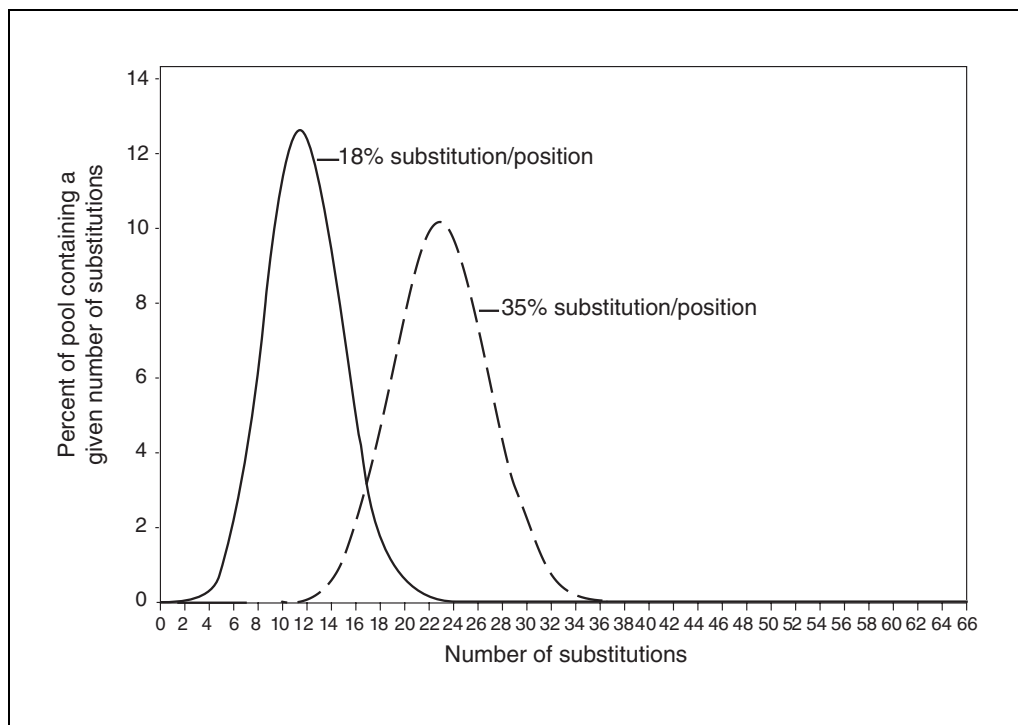


Figure 24.2.3 Comparison of substitution distributions for a 66-nucleotide pool doped to either 18% or 35%.

a pool of $\sim 10^{15}$ amplifiable molecules. Regardless of the degree of partial randomization or the precise doping strategy employed, the number of different mutational combinations is given by:

$$3^n \{L!/[n!(L-n)!]\}$$

where n is the number of sequence substitutions/template in a template of length L . For example, in the case of the 66-nucleotide RRE pool discussed earlier, there were $\sim 2.17 \times 10^9$ possible 5-residue substitutions and $\sim 1.25 \times 10^{16}$ possible 10-residue substitutions.

To calculate what fraction of a given set of substitutions are actually contained in a doped pool, the binomial probability distribution can be used:

$$P(n,L,f) = \{L!/[n!(L-n)!]\}(f^n)(1-f)^{(L-n)}$$

where P is the fraction of the template population when f is the probability of substitution/position. If primarily single-base substitutions are desired, then f should be maximized for $n = 1$; if multiple mutations (e.g., double or triple substitutions) are desired, then f should be correspondingly higher. If the doping strategy is optimized for n substitutions, then this number of substitutions will occur most frequently, “ $n - 1$ ” and “ $n + 1$ ” substitutions will occur less frequently but in roughly equal numbers, and so forth. Higher levels of sequence substitution skew the mutant frequency distribution, allowing the sampling of some regions of sequence space at the exclusion of others (Fig. 24.2.3).

Therefore, in the RRE example already cited, a pool of 1×10^{13} molecules doped at a frequency of 35% would contain few 5-residue substitutions [$1 \times 10^{13} \times P(5,66,0.35) = \sim 1.82 \times 10^6$ 5-residue substitutions out of $\sim 2.17 \times 10^9$ possible 5-residue substitutions]. In contrast, if the pool were doped at a frequency of 18%, all 5-residue substitutions would almost certainly be included [$1 \times 10^{13} \times P(5,66,0.18) = \sim 9.3 \times 10^{10}$ 5-residue substitutions].

Note that in a pool of only 1×10^{13} total molecules, neither doping scheme would yield all possible 10-residue substitutions.

Completely random pool design (aptamer selection)

Completely random sequence pools are used to initiate selection experiments when no functional nucleic acid sequence or structural motif is known in advance. There is really only one parameter to consider when designing a completely random pool: the length of the random region. While we will consider this parameter in detail below, we must first dismiss a frequent bogey of selection neophytes, the issue of complexity and representation.

Random sequence space is a vast landscape of possibilities of which only a vanishingly small fraction can be sampled by either nature or man. Assuming a 4-monomer repertoire from which pools can be constructed, there are $\sim 1.6 \times 10^{60}$ unique individual sequences in a sequence space bounded by a 100-residue template ($4^{100} = \sim 1.6 \times 10^{60}$), a quantity of nucleic acid greater than an Avogadro's number of Earth masses. While this grotesquely large value is clearly beyond the realm of experimental possibility, modern methods of chemical nucleic acid synthesis do allow the sampling of nearly as much sequence information as may be contained in the Earth's biosphere. As a back-of-the-envelope calculation, consider that there are on the order of $\sim 1 \times 10^9$ species in the biosphere, each with $\sim 1 \times 10^5$ genes. If each of these genes in turn is composed of $\sim 1 \times 10^3$ residues, then there are $\sim 1 \times 10^{17}$ residues worth of information in a biosphere. In contrast, a typical 1- μ mol synthesis of a 100-residue random sequence pool would contain 1×10^{15} molecules $\times \sim 1 \times 10^2$ residues/molecule = $\sim 1 \times 10^{17}$ unique residues or roughly 1 biosphere's worth of information. Obviously, the connection and ordering of sequence information in organisms is important as well.

Typically, a random sequence pool contains $\sim 1 \times 10^{15}$ molecules, and thus can potentially sample on the order of all possible 25-mers ($4^{25} = \sim 1.1 \times 10^{15}$). In fact, since different 25-mers can be found in different "reading frames," a slightly larger sequence space will likely be sampled. Because of this physical restriction, it is sometimes thought that random sequence pools should be no more than 25 residues in length—any longer, and only a fractional sampling would be possible, and many potential sequences would be lost. While this is true, it should be realized that longer pools do not lose any of the numerical complexity of smaller pools (except in those instances where long syntheses are extremely inefficient) and in fact gain access to some fraction of longer sequence and structural motifs as well. For example, tRNA molecules are roughly 76 nucleotides in length. It might prove more difficult to select tRNA mimics from a random sequence population containing 30 randomized residues than from a pool spanning 80 randomized residues. However, any short functional tRNA mimics present in the shorter population should also be present in equal or greater number in the longer population. In most instances, the relative completeness of the pool is not a consideration in the success of a selection. Indeed, it has been shown that functional nucleic acids are not extremely rare (for recent reviews see Gold et al., 1995; Fitzwater and Polisky, 1996) and can be isolated both from "complete" pools that span 20 random sequence positions and from very "incomplete" pools that span 90 random sequence positions.

Having dismissed considerations of complexity and representation, the one guiding principle that emerges from this analysis is that longer pools are more generally useful for selection experiments than shorter pools. However, this principle must be applied with appropriate caveats. First, aptamers derived from shorter pools are easier to analyze. Sequence and structural motifs embedded within a 30-nucleotide random sequence region are much more readily apparent than sequence and structural motifs embedded within a

90-nucleotide random sequence region, especially if the motifs are not colinear. Second, longer pools are more difficult to synthesize than shorter pools. Finally, longer pools are more likely to yield amplification or other selection artifacts than shorter pools. For example, pools that contain random regions greater than 90 nucleotides in length can form self-aggregates that precipitate from solution upon prolonged incubation, and thus require immobilization on a solid support prior to selection (Bartel and Szostak, 1993; Lorsch and Szostak, 1994). Because of these considerations, pools used for the *in vitro* selection of aptamers typically contain from 20 to 80 random sequence positions.

Longer pools are not only desirable but are likely required in selections for complex functions, such as catalysis. Pools used for the selection of ribozymes typically contain from 50 to 220 random sequence positions (for recent reviews see Gold et al., 1995; Fitzwater and Polisky, 1996). The optimal length of the random region is an active area of research (Sabeti et al., 1997) where many of the fundamental parameters remain to be defined. Practically, though, longer pools must be synthesized as oligonucleotides of 150 residues or fewer in length because of the constraints of DNA synthetic chemistry. For this reason, pools longer than 150 bases are typically generated in a modular fashion by ligating together individual, synthetic oligonucleotides (Bartel and Szostak, 1993). Segments of shorter DNAs can be stitched together by the inclusion of unique restriction sites (Bartel and Szostak, 1993). Asymmetric restriction sites, such as *Ava*I (CIYCGRG), *Ban*I (GIGYRCC), and *Sty*I (CICWWGG), are very useful for this task since they minimize intra-pool dimerization via self-ligation. Also, these enzymes are cost-effective for digesting large amounts of DNA. Alternatively, an overlapping region can be included at the 3' end of each synthetic oligonucleotide and mutually primed synthesis (e.g., UNIT 8.2) of a longer template can be carried out. After assembling pool modules, the complexity (yield) of the new, aggregate pool will need to be freshly assessed. The upper bound of the complexity of an assembled pool (e.g., 10^{11} 100-mer modules \times 10^{11} 100-mer modules) will likely be much larger than its actual complexity (e.g., 100 micrograms of ligated 200-mer, 9.12×10^{14} molecules).

Segmentally random pool design (binding site and aptamer selection)

In general, the rules governing the design of segmentally random pools are idiosyncratic, depending on experimental purpose. If the desire is to better define a known binding site, then relatively short sequence tracts (i.e., from four to ten residues) should be completely randomized. The randomization of longer sequence tracts may lead to the selection of novel binding sites rather than variants of a known binding site. The residues can either be colinear (as is the case for many DNA binding sites) or dispersed (as is the case for many RNA binding sites). If the desire is to identify a binding site within the context of a known structural element, then from four to twenty residues can be completely randomized. In this instance, the fewer the number of residues that are randomized, the more likely it will be that the selected sequences will resemble a wild-type binding site or retain an engineered structure. The greater the number of residues that are randomized, the more likely it will be that a novel aptamer sequence or structure will be discovered.

Primer design

Generally, the constant sequences at the 5' and 3' ends of a pool function as primer-binding sites and can be almost any sequence or length. Primers of 20 nucleotides in length are convenient because their melting temperatures are convenient for the PCR and they can easily be synthesized in high yields. In designing constant sequences and complementary primers, obvious artifacts associated with the PCR, such as secondary-structure formation or self-association that could lead to the production of primer dimers, should be avoided. Computer programs such as the Genetic Computer Group's PRIME or the Whitehead's PRIMER3 assist in designing constant regions. Other primer design programs include

Table 24.2.2 Representative Calculations Based on the Masses and Efficiencies for Couplings that Utilize the Canonical Tetrazole Activation Chemistry and Phosphoramidites Bearing Standard Protecting Groups

Phosphoramidite	Molecular mass (g/mol)	Mass correction	Coupling efficiency correction	Overall correction
5'-CE-dA	858	0.87	0.67	0.58
5'-CE-dC	834	0.89	0.67	0.60
5'-CE-dG	840	0.89	1.00	0.89
5'-CE-dT	745	1.00	0.83	0.83

discovered (A. Friedman, pers. comm.) that when pools and primers are synthesized on identical ports of a DNA synthesizer, there is some mixing of the molecules. The contamination is sufficient to yield a positive signal following extensive (30 to 50 PCR cycles) amplification of a no-template negative control. The unprogrammed interleaving of pools and primers can lead to extreme skewing of amplified materials, such that only a few species from the original pool may comprise a significant fraction of a subsequent amplification reaction. Therefore, pools and their cognate primers should be synthesized on different synthesizer ports and/or the machine should be extensively flushed with acetonitrile between syntheses.

Most synthesizers can be programmed for in-line, degenerate mixing of bases. While this method is useful when only a few positions must be randomized, because of the extremely fast reaction of the activated phosphoramidite with the newly deprotected 5' hydroxyl, random sequences will be skewed towards the phosphoramidite that first enters the column. Therefore, for longer pools or pools that should contain a statistically random distribution of nucleotides, it is better to manually mix the phosphoramidites off-line and use this mixture for the synthesis of degenerate sequence positions. A more stochastic distribution can be obtained by including larger amounts of A and C phosphoramidites in the mix to compensate for the faster coupling times of G and T phosphoramidites (Zon et al., 1985). Suggested ratios include a 3:3:2:2.4 molar ratio of A:C:G:T phosphoramidites (D.P. Bartel, pers. comm.), and a 1.5:1.25:1.15:1 molar ratio of A:C:G:T (see User's Manual for PE Biosystems Models 392 and 394 DNA/RNA Synthesis).

Doped pools are perhaps the most difficult to synthesize (Hermes et al., 1989; Bartel et al., 1991). Doping can be accomplished by using phosphoramidite mixtures that have been adjusted to ensure the proper level of partial randomization of a given nucleotide. For example, if a doped pool is to be synthesized in which non-wild-type residues are included at a rate of 10%/position, then for the adenosine bottle a molar ratio of 33.43:1.50:1.00:1.21 of A:C:G:T phosphoramidites should be used. These ratios were derived by first adjusting for the relative molecular mass and coupling differentials of the individual phosphoramidites and then mixing the phosphoramidite solutions on a percent

Table 24.2.3 Volumes of Acetonitrile Needed to Dissolve 1 g of Phosphoramidite

Phosphoramidite	Dissolved in X ml of acetonitrile
5'-CE-dA	11.6
5'-CE-dC	12.0
5'-CE-dG	17.8
5'-CE-dT	16.6

Table 24.2.4 Amidite Mixtures for Synthesis of Doped Pool in Which Non-Wild-Type Residues Are Included at a Rate of 10%/Position

Phosphoramidite	Mutagenesis (%)	Total volume (ml)	Volume each amidite to mix (ml)			
			A	C	G	T
A	10	10	9.00	0.33	0.33	0.33
C	10	10	0.33	9.00	0.33	0.33
G	10	10	0.33	0.33	9.00	0.33
T	10	10	0.33	0.33	0.33	9.00
A	20	10	8.00	0.67	0.67	0.67
C	20	10	0.67	8.00	0.67	0.67
G	20	10	0.67	0.67	8.00	0.67
T	20	10	0.67	0.67	0.67	8.00

volume basis to yield the desired extent of doping. This process is described in greater detail below.

To normalize the coupling of different phosphoramidites, relative correction factors that take into account different coupling efficiencies and molecular masses must be calculated. Multiplying together these correction factors gives an overall correction factor to provide equal molar coupling of each phosphoramidite. Table 24.2.2 displays representative calculations based on the masses and efficiencies for couplings that utilize the canonical tetrazole activation chemistry (UNIT 2.11; Beaucage and Caruthers, 2000) and phosphoramidites bearing standard protecting groups [cyanoethyl for the phosphates along either isobutyryl (N-2 of guanine) or benzoyl (N-6 of adenine and N-4 of cytidine) groups]. Other chemistries and protections may require the substitution of other correction factors.

Most modern synthesizers require that ~1 g of phosphoramidite be dissolved in ~20 ml of acetonitrile to be used in the coupling reaction. Applying this constraint along with the combined mass-coupling (overall) correction factor gives the volumes shown in Table 24.2.3 to dissolve 1 g of each phosphoramidite. Therefore, if equal volumes of each of these solutions are mixed, equal molar coupling should occur since the molar concentrations have been adjusted to account for both the mass and coupling differentials.

As in the example above, if a doped pool is to be synthesized in which non-wild-type residues are included at a rate of 10%/position, then the amidites should be mixed as in Table 24.2.4.

In addition to varying nucleotide composition, it is also possible to vary the length of random sequence that is synthesized. Deletions can be stochastically incorporated during a synthesis by replacing the capping step with an acetonitrile wash (Bartel et al., 1991). It is more difficult to stochastically incorporate insertions, but the lengths of segmental random sequences in a pool can be mixed. For example, in Giver et al. (1993), four columns were used to generate a pool with two random regions of 6 to 9 positions separated by a constant domain. The first column was synthesized with 6 random positions, the second with 7 random positions, etc. Following the addition of the intervening constant sequence, the synthesis was stopped, the four columns were opened, and the resins from the four columns were mixed. The mixed resins were then equally redivided into four new columns and the synthesis was resumed. The first column incorporated 6 positions, the second column 7 positions, etc. Thus, the first column

contained oligonucleotides in which the first random segment was 6, 7, 8, or 9 residues long and a second random segment that was uniformly 6 residues long. The second column contained oligonucleotides in which the first random segment was 6, 7, 8, or 9 residues long and a second random segment was uniformly 7 residues long, and so forth. Following the completion of all four syntheses, the reactions were combined to generate the final random sequence pool.

PURIFICATION OF A RANDOM SEQUENCE POOL

A newly synthesized oligonucleotide pool should be purified on a denaturing polyacrylamide gel (see e.g., *UNIT 2.12*) prior to amplification. Oligonucleotides can also be purified using an HPLC or commercially available spin columns, but HPLC purification is not recommended for ssDNA pools, due to concerns about cross-contamination. Since oligonucleotides of equivalent length but different sequence migrate at slightly varying rates (see User's Guide for PE Biosystems Expedite Nucleic Acid Synthesis System), a pool should appear as a broader band than a homogeneous sequence. In fact, because of the presence of capped failure sequences and depurinated, cleaved fragments, it is likely that the oligonucleotide product will appear even more heterogeneous.

As a general note, since sequences exist as single copies prior to amplification, individual species can be easily lost. Therefore, it is important to wash and elute the various filters, tubes, and tips described below one or more times. The eluates can then be pooled for a final precipitation and eventual amplification.

Contamination of primers or other solutions with a synthesized or isolated pool should be avoided by using aerosol barrier tips. Similarly, gel plates used during purification should be washed thoroughly to ensure that they are free of contamination with other pools or primers.

Materials

DNA pool
Ammonium hydroxide
n-butanol
TE buffer, pH 8.0 (*APPENDIX 2*)
Urea loading buffer, 2× (*UNIT 2.12*)
5 M NaCl
Ethanol

Fluorescent TLC plate (VWR), wrapped in plastic wrap
UV lamp
Razor blades
Small-bore syringes
13-ml centrifuge tubes capable of withstanding temperature extremes (Sarstedt)
90°C water bath
Rotary shaker

Additional reagents and equipment for denaturing polyacrylamide gel electrophoresis (e.g., *UNIT 2.12*)

1. After synthesis, deprotection, and cleavage from the solid support, lyophilize the oligonucleotide solution (in concentrated ammonium hydroxide) to dryness or precipitate with a 10-fold volume of *n*-butanol.

The n-butanol precipitation can occur quite quickly at room temperature for longer oligonucleotides. Shorter (<20 base) oligonucleotides may require longer or colder

incubations. To ensure more efficient recoveries of oligonucleotides it is safest to precipitate for ≥ 1 hr at -70°C .

2. Pour a denaturing polyacrylamide gel (e.g., UNIT 2.12).

To allow for good separation of near-full-length from non-full-length products, the acrylamide concentration should be chosen so that the full-length oligonucleotide will migrate approximately one-half to three-fourths of the way into the gel by the time the loading dye reaches the bottom.

3. Resuspend the lyophilized or precipitated pellet in ~ 100 to $200\ \mu\text{l}$ of water or buffer (e.g., TE buffer, pH 8.0) and add an equal volume of $2\times$ loading dye. Heat denature samples at 75°C for 5 min prior to loading. Load $\sim 20\%$ of a $1\text{-}\mu\text{mol}$ synthesis per $2\ \text{cm} \times 2\ \text{cm} \times 1.6\ \text{mm}$ well and perform electrophoretic separation.
4. Place gel on a fluorescent TLC plate that has been wrapped in plastic wrap and excise the oligonucleotide product from the gel with the aid of a UV lamp, using razor blades.

The desired oligonucleotide product is generally the darkest, shadowed band on the gel (excluding UV-absorbing material that runs at the dye front). If stepwise synthetic efficiency has been low, the product will appear as a smear instead of as a clear band. Since many of the N-1, N-2, etc. products can be converted into full-length products by the polymerase chain reaction, a fairly wide band of near full-length products can be cut from the gel. The excision should be carried out relatively quickly, since unnecessarily long UV exposure can damage the oligonucleotide product.

The full-length oligonucleotide product should be the slowest-migrating band. However, if deprotection has been incomplete, lighter bands that migrate considerably above the major fully deprotected band may be observed.

Unpolymerized acrylamide absorbs strongly at $211\ \text{nm}$ and may cause shadowing at the edges and wells of the gel. This can obscure the resolution or recovery of bands in the outer lanes.

5. Elute the oligonucleotide from the gel slices as follows.
 - a. To aid in the diffusion of the oligonucleotide from the acrylamide matrix, chop gel slabs into fine particles by forcing the gel through a small-bore syringe.
 - b. Place the crushed gel slabs in a 13-ml centrifuge tube capable of withstanding temperature extremes.
 - c. Add 3 ml of TE buffer, pH 8.0, per 0.5 ml of gel slab (typically corresponding to one to two wells), and place the sample at -80°C for 30 min or until it is frozen solid.
 - d. Quickly thaw the tube in a hot water bath and then let it soak at 90°C for 5 min. Elute the DNA overnight at room temperature on a rotary shaker.

This freeze-rapid thaw approach (Chen and Ruffner, 1996) allows ice crystals to break apart the acrylamide matrix, increasing yield and decreasing elution time. Typically, 80% of a 20-mer oligonucleotide can be recovered after 3 hr of rotary shaking, making this technique comparable to electroelution (see UNIT 2.7).

Because elution is a diffusion-controlled process, higher elution volumes or serial elutions from the same gel slice can increase the amount of DNA recovered. Longer oligonucleotides diffuse from the gel more slowly than shorter sequences. Samples of especially long synthetic DNAs and RNAs that are particularly resistant to elution with aqueous buffers may be eluted more easily in 6 vol of formamide (>5 hr at room temperature), followed by a brief elution with an aqueous buffer (~ 1 hr). Isoamyl alcohol extraction (UNIT 2.12) can be used to bring the extracts to a convenient volume for subsequent precipitation.

- Precipitate the eluted oligonucleotide pool by adjusting the salt concentration to 0.3 M using a 5 M NaCl stock solution, then adding 3 vol of ethanol. Keep at -20°C for 3 hr, then microcentrifuge at maximum speed 4°C . Lyophilize to dryness. Resuspend the synthetic pool in TE buffer, pH 8.0 (to protect against nuclease contamination or drastic pH changes).

If the volume of the eluted oligonucleotide is too large to conveniently precipitate, concentrate the sample by extracting against an equal volume of n-butanol. Remove the upper butanol layer and repeat until the aqueous volume is convenient for precipitation. About 1/5 of the aqueous layer is extracted into the organic butanol layer for every volume of butanol used. If too much butanol is used, thereby completely extracting the aqueous layer into the butanol, add more water and repeat the concentration.

DETERMINING THE POOL COMPLEXITY

The number of different molecules present in a population can affect the outcome of a selection experiment (see Troubleshooting). If the pool complexity is too low for a given application, the pool will have to be resynthesized.

Pool complexity is, in turn, a function of yield and of the number of molecules in the pool that can be fully extended by a polymerase. The overall yield of the synthesis can be calculated by determining the UV absorption of the pool. However, deletions, incompletely deprotected residues, or backbone lesions that arise during chemical synthesis decrease by 10% to 40% the fraction of molecules in a synthetic pool that can be fully extended by polymerases. For example, the rate of insertions (presumably due to DMT cleavage via tetrazole) has been measured to be as high as 0.4% per position, and the rate of deletions (presumably due to incomplete capping) has been found to be as high as 0.5% per position (A. Keefe and D. Wilson, pers. comm.). The number of usable DNA molecules that are actually present in a nascent pool can be calculated by determining the fraction of the pool that can be extended by *Taq* polymerase.

Materials

Purified ssDNA pool and labeled primers
50 mM Tris-Cl, pH 7.5 (APPENDIX 2)
10 mM MgCl_2
5 mM DTT
[γ - ^{32}P]ATP (>3000 Ci/mmol)
T4 polynucleotide kinase
1:1 phenol/chloroform (UNIT 2.1A)
Chloroform
4.0 M ammonium acetate
Taq DNA polymerase
TE buffer, pH 8.0 (APPENDIX 2)
PCR amplification buffer (see recipe)
2 \times formamide loading buffer (see recipe)
15 \times 17-cm denaturing polyacrylamide gel (UNIT 2.12)
Thermal cycler
Phosphor imager plate and phosphor imager
Additional reagents and equipment for quantitation of DNA (e.g., APPENDIX 3D), end-labeling of DNA (e.g., UNIT 3.10), phenol/chloroform and chloroform extraction of DNA (UNIT 2.1A), PCR amplification (e.g., Chapter 15), and denaturing polyacrylamide gel electrophoresis (UNIT 2.12)

SUPPORT PROTOCOL 1

1. Quantitate DNA by UV absorption assuming that A_{260} of 1.0 indicates $\sim 37 \mu\text{g/ml}$ of single stranded DNA.

Also see APPENDIX 3D.

2. Label the 5' end of the 3' PCR primer with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ by preparing the following reaction mixture.

For 30- μl reaction (volume of reaction and concentration of DNA and $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ will vary depending on application):

50 mM Tris-Cl, pH 7.5
10 mM MgCl_2
5 mM DTT
1 to 50 pmol dephosphorylated DNA, 5' ends
50 pmol (150 μCi) $[\gamma\text{-}^{32}\text{P}]\text{ATP}$
50 $\mu\text{g/ml}$ BSA
20 U T4 polynucleotide kinase

Incubate 60 min at 37°C , then stop reaction by adding 1 μl of 0.5 M EDTA. Phenol/chloroform and chloroform extract the labeled oligonucleotide (UNIT 2.1A), and precipitate by adding an equal volume of 4.0 M ammonium acetate and 2 vol ethanol. Microcentrifuge to collect the pellet, remove the supernatant, and redissolve the labeled DNA pellet in 10 μl of TE buffer, pH 8.0.

This procedure ensures that most of the unincorporated label remains in the supernatant.

3. Incubate ~ 50 pmol of labeled primer with a 2- to 5-fold molar excess of pool in a 50- μl extension reaction, under the same conditions that will be used in the final amplification, in a thermal cycler as follows (see UNIT 15.1).
 - a. Denature and anneal the primer and template DNA in PCR amplification buffer (usually 94°C for the denaturation step and $\sim 50^\circ\text{C}$ for the annealing step).
 - b. Add *Taq* or other DNA polymerase (scaled to the anticipated enzyme concentration to be used in the large-scale amplification), then increase the temperature to 72°C for 20 min.

It may be useful to take time points to determine whether the reaction has gone to completion.

- c. Finally, terminate the reaction by the addition of $2\times$ formamide loading buffer.
4. Heat the extension reaction to 90°C for 3 min and load the reaction on a $15 \times 17\text{-cm}$ denaturing polyacrylamide gel with appropriate radiolabeled size markers. Electrophorese until the dye is at or near the bottom of the gel, but do not let the radiolabeled primers run off.

It is also useful to load a separate well with an aliquot of the primer alone. Choose an acrylamide percentage that allows efficient separation of small primers from larger extended products.

5. Dry and expose the gel to a phosphor imager plate. Using a phosphor imager, quantify the control primer band and the extended product band.

There may be a smear leading up to the extended band. One should use one's best judgment in determining how much near-full-length material will be included in the quantitation. Calculate the percent extension by dividing counts of labeled, extended product by counts of labeled primer. Percent extension for a gel-purified ssDNA pool can range from 10% to 30%. The complexity of the pool is then the yield (determined in step 1) multiplied by the extension efficiency (percent extension determined above). If the complexity of the pool is insufficient for planned experiments, then the pool must be resynthesized.

DETERMINING THE POOL BIAS

Following extension, the reaction should be repeated using a cold primer and the nonradioactive double-stranded DNA pool should be amplified in a PCR reaction, cloned (e.g., using a TA cloning kit from Invitrogen), and individual members sequenced to determine the degree of partial or completely randomness. The cloning step could also be carried out following PCR optimization (see Support Protocol 3). From 20 to 30 clones should be sequenced to determine the base composition of the starting pool. The random region should be composed of roughly 25% of each base. A pool with the random region skewed toward one or more bases (>30%) should be resynthesized.

**SUPPORT
PROTOCOL 2**

SMALL-SCALE PCR OPTIMIZATION OF POOL AMPLIFICATION

To enhance yield and further avoid bias, the amplification conditions for a pool should be optimized prior to carrying out a large-scale amplification. Moreover, since amplifying a pool is costly in terms of both time and money, any optimization of the PCR should first take place on a small scale. The more involved large-scale amplification can then be carried out with confidence.

**SUPPORT
PROTOCOL 3**

Materials

dNTPs (*UNIT 3.4*)

Taq DNA polymerase (e.g., Boehringer Mannheim)

PCR amplification buffer containing 1.5 mM Mg²⁺ (see recipe)

dsDNA mass markers (e.g., Life Technologies)

4% Nu Sieve agarose gel (FMC Bioproducts)

Thermal cycler

Densitometer

Additional reagents and equipment for agarose gel electrophoresis (*UNIT 2.5*)

1. Carry out a 0.1 ml PCR reaction using 2 nM of synthetic pool oligonucleotide as template, 2 μM primers, and PCR buffer with 1.5 mM magnesium. Use the manufacturer's suggested quantity of *Taq* (e.g., 2.5 U of Boehringer Mannheim *Taq*) in a reaction containing 200 μM dNTPs. A suggested temperature regime is:

10 to 15 cycles: 2 min	95°C	(denaturation)
1 min	55°C	(dependent on primer composition)
		(annealing)
3 min	72°C	(extension).

After 10 to 15 cycles of amplification, check the length and purity of the amplified DNA on a 4% Nu Sieve agarose gel in 1× TBE buffer (*UNIT 2.5*).

A 0.1 ml reaction typically yields ~1 μg, but the amount can vary from 0.1 to 10 μg. A fuzzy band may indicate that too many cycles of PCR have been carried out. In this case, set up the reaction again and perform fewer cycles.

2. Dilute the double-stranded PCR DNA product 1:128, and repeat the PCR reaction, removing a 5- to 10-μl aliquot during the last 10 sec of the cycle-7 extension step. Serially dilute the amplified product 1:2, 1:4, ... 1:128. Electrophorese all of the samples on a large agarose gel.

Note that it is quite difficult to accurately pipet solutions at 72°C. It may therefore be desirable to pipet an amount slightly larger than that intended for use in the serial dilution.

**Generation and
Use of
Combinatorial
Libraries**

24.2.15

3. Calculate the average PCR efficiency by identifying to what extent the cycle-7 PCR reaction is the result of progressive doublings of the original synthetic DNA. Determine which dilution lanes lack detectable DNA.

The largest dilution that lacks detectable DNA is also the dilution that is a minimum estimate of the number of doublings. For example, if the 1/64 dilution is the largest dilution without detectable DNA, this implies that 6 “doublings” of the synthetic DNA yielded at least 64-fold more DNA. This is expressed as follows:

(average efficiency)^{no. of theoretical doublings (i.e., PCR cycles)} = fold increase in DNA

Thus, if 7 cycles of PCR were performed, then the average number of doublings per cycle is ~1.81 [from $(-1.81)^7 = 64$].

4. Modulate PCR conditions to enhance PCR efficiency.

If the pool's average number of doublings per cycle is <1.8, then the PCR conditions chosen may skew the representation of the pool. In that case PCR conditions should be modulated to enhance PCR efficiency. The following parameters or variables are most amenable to modification. It is best to begin the optimization with a single set of reaction conditions, modify individual parameters relative to this one reference reaction, and then combine all advantageous alterations into a single reaction. In addition, one may wish to consult UNIT 15.1.

Theoretically PCR can proceed until the primers or dNTPs are depleted. Therefore, primer and dNTP concentrations should be well above those used for the amplification of small amounts of DNA. Primer concentrations from 1 μ M to as high as 5 μ M have been used (although concentrations >5 μ M are generally not helpful). It may be useful to scan both above and below 2.5 μ M in 0.5- μ M increments.

Magnesium concentration affects both primer annealing and the fidelity of Taq (which decreases with increasing magnesium concentration). Starting at the magnesium supplied in the PCR buffer (usually 1.5 mM), scan in 1-mM increments toward 5 mM as a maximal concentration.

DNA denaturation at temperatures above 95°C is usually impractical since this greatly reduces Taq's half-life. While other thermostable polymerases can be more resistant to higher temperatures, they usually have a lower extension efficiency and are more expensive than Taq. Annealing temperatures are dependent upon both primer sequence and length. The primer annealing temperatures should already be known from the primer design process, or may be calculated via an algorithm that can be found at <http://paris.chem.yale.edu/extinct.html>. This algorithm takes into account nucleotide composition, stacking energies (according to Turner's rules), and empirical data. An annealing temperature ~5°C less than the calculated annealing temperature is a good place to begin optimization. The amplification is more efficient at a lower annealing temperature, but mispriming and secondary structural problems are more pronounced. Higher temperatures improve the specificity, but decrease the overall yield of the reaction. To determine the optimum annealing temperature for a given primer and magnesium concentration, one should scan in both directions around the annealing temperature in 5°C increments. Finally, extension temperatures are modulated by the properties of Taq, which will extend (although inefficiently) at temperatures as low as 65°C. When extending at temperatures above Taq's optimum temperature (70° to 75°C) somewhat more polymerase may be required; scanning of the enzyme quantity should be done in 2.5-U increments. However, too much Taq may be harmful to structured single-stranded nucleic acids (Lyamichev et al., 1993).

5. Confirm the results of the extension reaction described in Support Protocol 1 by the optimization method as follows. After optimizing pool PCR conditions for >1.8 average number of doublings per cycle, determine the pool complexity by performing another 0.1-ml PCR reaction with 2 nM of the original, synthetic pool oligonucleotide under the now optimized reaction conditions. After 7 or more cycles of PCR, perform agarose gel electrophoresis on serial dilutions of the PCR reaction adjacent to serial

dilutions of dsDNA mass markers. Calculate the amount of amplified DNA using either a densitometer or by estimating which dilutions are most similar. Calculate the approximate pool complexity as follows:

$$\frac{\text{g of PCR DNA after } N \text{ cycles of PCR}}{\text{g avg no. of doublings per cycle (see step 4)}} = \text{g of starting extendable ssDNA}$$

$$\frac{\text{g of starting extendable ssDNA}}{330 \text{ g/mol} \times (\text{no. of bases in full-length product})} = \text{mol of starting extendable ssDNA}$$

$$\text{mol starting extendable ssDNA} \times (6.02 \times 10^{23}) = \text{molecules of starting extendable ssDNA}$$

$$\frac{\text{molecules of starting extendable ssDNA}}{\text{starting molecules}} = \text{fraction of extendable ssDNA}$$

$$\text{fraction of extendable ssDNA} \times \text{no. of synthetic pool molecules} = \text{pool complexity}$$

PCR efficiency should be optimized to balance the average number of doublings per cycle against the total reaction volume. A pool of 1×10^{15} molecules ($\sim 1.7 \times 10^9$ mol) at a starting template concentration of 2 nM will require 0.85 L for amplification. Therefore, it is greatly desirable to amplify the pool at the highest template concentration that still gives a reasonable number of doublings per cycle. The amplification should generate at least 8 copies of pool DNA if the pool complexity is to be archived and preserved (see Basic Protocol 2).

LARGE-SCALE PCR AMPLIFICATION OF POOL DNA

Very long and complex pools often require PCR amplification on a multiple-milliliter scale. Large-scale PCR differs from conventional PCR in that it is typically conducted in water baths using 15 ml, 17 × 120-mm, screw-capped (Sarstedt) thermostable tubes to accommodate the larger volumes. Amplification reactions of up to 2.5 L have been carried out in this way. Medium-scale amplifications can sometimes be carried out in thermal cyclers that can accommodate multiple samples (e.g., 96-well PCR plates).

Materials

Purified ssDNA pool and primers

EDTA

1:1 phenol/chloroform (UNIT 2.1A)

Chloroform

4 M ammonium acetate

Ethanol

TE buffer, pH 8.0, containing 50 mM of a salt such as potassium chloride

Thermal cycler or three water baths (one must be a circulating water bath)

96-well PCR plate or 13-ml thermostable tubes (Sarstedt)

Thermometer

Styrofoam racks

Spectrophotometer or fluorometer

Additional reagents and equipment for PCR amplification (UNIT 15.1; see Support Protocol 3 for determination of conditions on a small scale) and phenol/chloroform and chloroform extraction of DNA (UNIT 2.1A)

BASIC PROTOCOL 2

Generation and
Use of
Combinatorial
Libraries

24.2.17

Plan the reaction

Since large-scale reactions are quite expensive in terms of nucleotides and enzyme, preparedness and planning for the large-scale amplification cannot be overemphasized. Primers <20 bases in length usually do not need to be gel purified and can instead be purified by precipitation.

1. After identifying the optimal PCR conditions on a small scale (see Support Protocol 3), prepare reagents for the large-scale reaction. Set aside time for the large-scale amplification, which will probably consume an entire day.

The size of the large-scale reaction will be determined in part by the amount of DNA pool to be amplified and by the number of copies of the library that are desired. For example, assume that 100 (extendable) μg of a pool are to be amplified 16-fold. Since the typical amount of DNA recovered from a 100- μl PCR reaction is 1 μg , then each 100- μl reaction should have $1 \mu\text{g}/16 = 60 \text{ ng}$ of DNA. $100 \mu\text{g total}/60 \text{ ng}/100 \mu\text{l} = 1667 \times 100 \mu\text{l}$, or a 167-ml reaction.

Choose how the amplification will be carried out

If the volume of the large-scale amplification reaction is to be $\leq 100 \text{ ml}$

- 2a. Use a commercially available thermal cycler repetitively. Set the reaction mixture up in advance, and pipet 100- μl aliquots into individual wells of a 96-well PCR plate.
- 3a. Carry out several small amplification reactions in advance to ensure that the optimized conditions determined in Support Protocol 3 work with the PCR plate format, and that amplification is uniform across the PCR plate.
- 4a. Perform thermal cycling on the entire reaction using eleven PCR plates.

For larger volumes

Reactions will be divided into aliquots in 13-ml thermostable (Sarstedt) tubes and amplified in a series of water baths. Construct floating racks by cutting off the bottom of the tubes' Styrofoam packing material. Reinforce these racks by wrapping their edges with heavy tape. Place the racks iteratively in three circulating or static water baths held at the denaturation, annealing, and elongation temperatures previously determined (see Support Protocol 3).

- 2b. Determine how long it will take for the reaction mixture in a tube to come to thermal equilibrium by constructing a temperature probe, placing a thermometer through the top of a Sarstedt tube filled with 10 ml of water. Place the probe in a rack with other, similar tubes.

Typical equilibration times range from 2 to 8 min, depending on the temperature differential. Annealing, and extension times of 5, 6, and 7 min are typical. It should be noted that these ramping temperature profiles are very slow relative to a commercial PCR machine and can yield more amplification artifacts.

- 3b. To ensure that the reaction conditions actually work as planned, fill the rack with tubes of water, a single amplification reaction, and the temperature probe. Denature the sample for 30 min, and then add *Taq* after the first annealing step. Take aliquots at each cycle to monitor the progress of the reaction.
- 4b. When reaction conditions have been confirmed, proceed with the remaining amplification reactions. Allow the final extension step to proceed for at least 20 min to ensure that all templates are completely double-stranded.

Do not be alarmed if the solution becomes cloudy; the detergent in the buffer causes the turbidity.

Amplification efficiencies of 3 to 4 doublings in 5 cycles can typically be achieved using this method.

5. Following the amplification, pool the reactions from the individual wells or tubes. Chelate the magnesium in the buffer by adding 1.1 molar equivalents of EDTA, pH 8.0.

The reactions can be left at 4°C overnight.

6. Add an equal volume of 2-butanol and extract to concentrate the reaction to a manageable volume (usually 10- to 20-fold). Mix the layers by vortexing and then separate by centrifuging 5 min at $1200 \times g$ at room temperature, then discard the upper, butanol layer. Repeat as necessary.

About one-fifth of the aqueous layer is extracted into the organic butanol layer for each volume of butanol used.

7. After concentrating the DNA, carry out a phenol/chloroform extraction, followed by two successive chloroform extractions (UNIT 2.1A).

At this point, it should be possible to easily precipitate the DNA. Be sure to temporarily save all of the organic layers in case of a mishap. Falcon tubes (50 ml) work well for these extractions, as they are conveniently sized and have a small surface area. Alternatively, a Teflon extraction funnel may be useful since nucleic acids will not stick to its surface.

8. Precipitate the DNA by adding an equal volume of 4 M ammonium acetate (final concentration, 2 M) and 2 vol ethanol in 13-ml Sarstedt tubes if possible.

If larger tubes are required, prepare a set of Beckman 250-ml high-speed centrifugation jars. Wash the jars with 15 ml of 3% hydrogen peroxide for 30 min and then rinse three times with 100 ml of distilled water to remove any residual DNases that may remain from previous use (typically bacterial cell pelleting).

9. Resuspend the amplified DNA in TE buffer, pH 8.0, containing 50 mM of a salt such as potassium chloride.

It is unwise to resuspend a double-stranded DNA pool in water, since the random segments may denature, reassort, and become transcriptionally incompetent.

If it is suspected that the pool has become denatured (for example, if a large single-stranded DNA component is seen on a nondenaturing agarose gel), simply repeat one to two cycles of PCR.

10. Quantitate the PCR DNA.

This can be done by carrying out gel electrophoresis in parallel with a DNA ladder of known concentration. The concentration can also be determined spectrophotometrically or by monitoring the change in absorbance of an intercalated fluorescent dye, Hoechst 33258 (Sigma) on a fluorometer (e.g., DyNA Quant 200, Amersham Pharmacia Biotech). These latter methods are much more quantitative (although the fluorometer method may not be accurate for sequences <100 nucleotides in length). However, these methods may not distinguish precipitated double-stranded DNA from residual, precipitated nucleotides or single-stranded primers. Determine the overall PCR efficiency and the final number of DNA molecules.

The amount of DNA obtained from large-scale amplification is often referred to in terms of the number of copies of the original synthetic pool's complexity. For example, if the starting pool had a complexity of 1×10^{15} molecules and 8×10^{15} total DNA molecules were recovered, then, on average, 8 copies of the original starting pool were obtained from the amplification. It should be noted that skewing that may arise during amplification, and sampling errors that occur during the use of the amplified pool, may cause this estimation to be grossly inaccurate; nevertheless, it is empirically useful.

11. Following large-scale amplification, store at least 4 copies of the pool at -80°C .

Because of the aforementioned sampling errors, archiving at least 4 copies worth of the pool DNA ensures the preservation of most of the pool's complexity. The amount of preserved pool complexity can be calculated using the following equation:

$$\% \text{ of the pool complexity in a given sample} = 100 \times \{1 - [(x - y)/x]^x\}$$

where x is total number of pool copies, and y is the number of pool copies archived.

Therefore, in the example given above, if 4 of the 8 copies of the pool generated through amplification are archived, then $\sim 99.6\%$ of the original starting pool's complexity is preserved. Similarly, at least 4 copies of the pool should be used whenever manipulations such as ligation, transcription, or biotinylation, are carried out, so that the original complexity is also manifest in the manipulated or synthesized copies.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Formamide loading buffer, 2 \times

Prepare in deionized formamide:

0.05% (w/v) bromphenol blue

0.05% (w/v) xylene cyanol FF

20 mM EDTA

Do not sterilize

Store at -20°C

PCR amplification buffer, 10 \times

500 mM KCl

100 mM Tris \cdot Cl, pH 8.3 (APPENDIX 2)

x mM MgCl_2

0.1% (w/v) gelatin

Store in aliquots at -20°C

This solution can be sterilized by autoclaving. Alternatively, it can be made from sterile water and stock solutions, and the sterilization omitted.

15 mM MgCl_2 is the concentration (x) used for most PCR reactions. However, the optimal concentration depends on the sequence and primer of interest and may have to be determined experimentally (also see recommendations in UNIT 15.1).

COMMENTARY

Critical Parameters

Synthesis

Depending on the size of the pool to be synthesized, the operation of the DNA synthesizer may first need to be optimized. Short pools (<80 total nucleotides in length) can be synthesized using standard protocols (see, e.g., PerSeptive Biosystems, 1998). In order to synthesize longer pools (>80 total nucleotides in length), all reagents should be fresh and special care should be taken to exclude water from the synthesis (see UNIT 2.1A). To ensure equimolar base incorporation in the random region of longer pools, the phosphoramidites must be

mixed in a skewed ratio (see Strategic Planning). Coupling efficiency should be monitored throughout the synthesis by following the trityl output (see UNIT 2.1A).

Amplification

Optimization of PCR conditions according to established protocols is vital to the success of the large-scale amplification. Cycle temperatures and times, as well as the concentrations of polymerase, primers, and dNTPs (see UNIT 15.1) should be addressed prior to the large-scale workup. Most importantly, since extremely large quantities of relatively expensive reagents (e.g., *Taq* polymerase) may be required, care

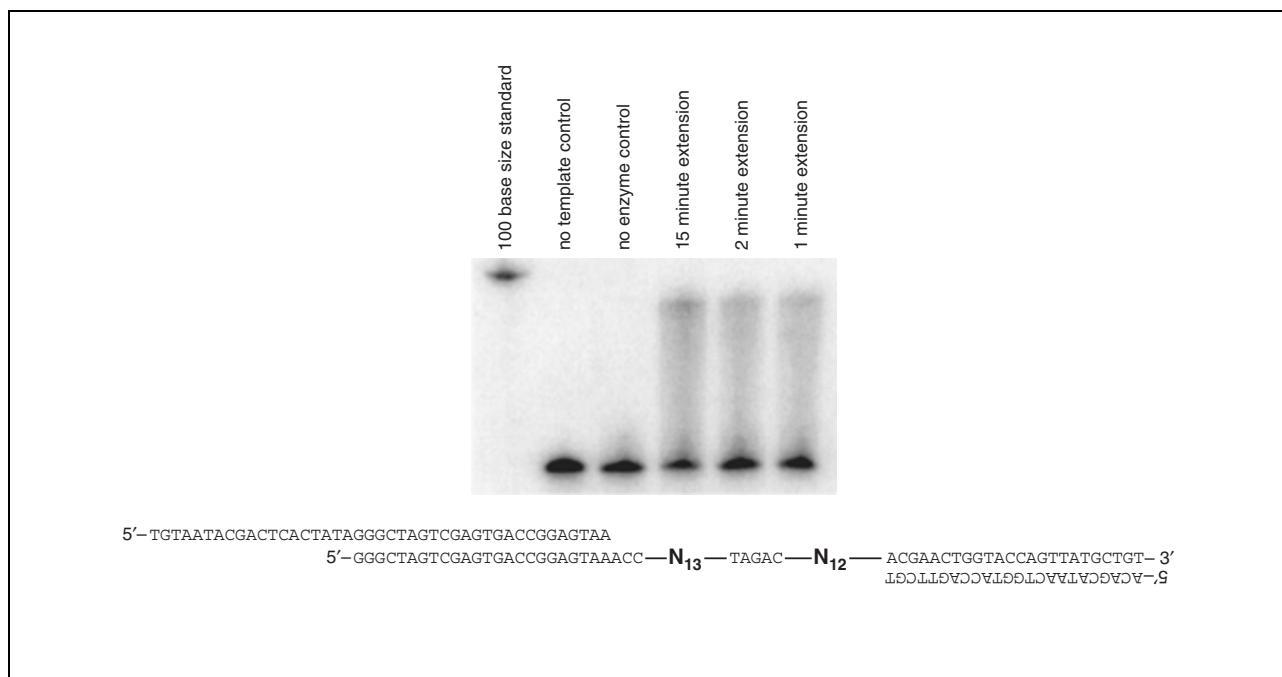


Figure 24.2.4 Typical extension reaction. The pool used (N30P) is shown below the figure of the gel. Lane 1 is a size standard, lanes 2 and 3 show control reactions, and lanes 4, 5, and 6 follow the extension efficiency after different incubation times.

should be taken to make sure that all reagents and procedures are in readiness. Different priming sequences often require distinct PCR buffers for optimal extension efficiency; the best buffer for a given pool and primer combination can be easily and systematically identified through the use of a PCR optimization kit (e.g., the PCR Optimizer Kit from Invitrogen).

Troubleshooting

The most common problem with the synthesis of a random sequence pool is the overall synthetic yield. However, researchers should carefully decide how many sequences are really necessary for their selection experiments. In selection experiments from a pool with a relatively limited potential diversity (i.e., a segmentally random pool with only 1×10^{11} possible sequences or less), even a low synthetic yield should be sufficient. However, *in vitro* selection from a pool with a very high potential diversity (i.e., a completely random pool with 1×10^{15} possible sequences or more) should use at least 1×10^{14} different sequences initially in order to adequately sample the potential sequence space. Pools that contain fewer than 1×10^{13} possible sequences should not be used.

The most likely sources of low yields and coupling efficiencies are old (i.e., water-contaminated) synthesis reagents. Thus, instead of attempting to amplify an incomplete pool, the

pool should be resynthesized with fresh reagents; the old and new pools can then be combined, if desired. If fresh synthesis reagents do not significantly raise yields, then more serious problems, such as line or valve blockage, may be the cause, and the instrument service representative should be contacted.

The second most common problem is that the base composition of a partially or completely random region is skewed. Unfortunately, skewing cannot be detected until after completion of a large-scale amplification. Fortunately, unless the degree of skewing is extreme, it should not seriously affect the outcome of a selection. Moreover, if the degree of skewing is known in advance of a selection, it can be taken into account when analyzing the results of the selection. For example, Baskerville et al. (1995) selected functional Rex-binding elements from a partially randomized pool. Despite the fact that the initial pool did not contain equimolar representation of non-wild-type bases at partially randomized positions, these authors were able to determine the relative importance of individual residues by comparing the degree of conservation or variance before and after selection. If a researcher decides that extant skewing of base ratios is unacceptable, this can only be fixed by adjustment of the randomized phosphoramidite mixture and resynthesis of the pool.

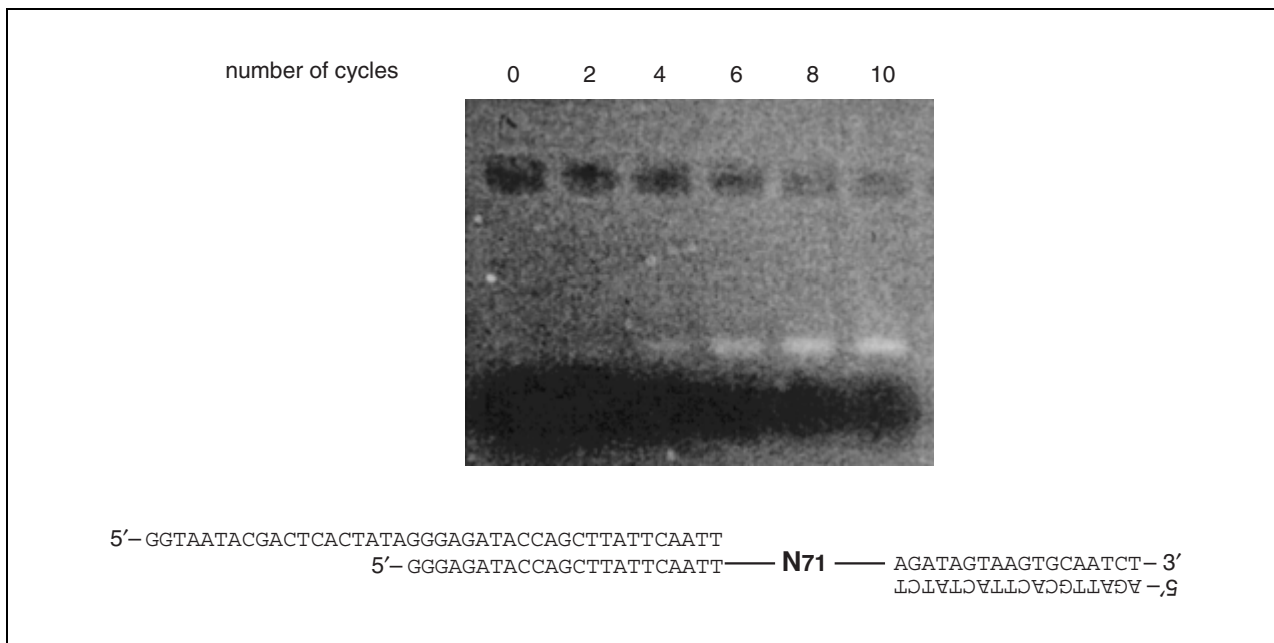


Figure 24.2.5 Cycle course. The gel follows amplification of the N71 pool after 0, 2, 4, 6, 8, and 10 cycles. The pool used in the cycle course is depicted below the figure of the gel.

The third most common problem is that the pool fails to efficiently elongate. With the proviso that the efficiency of extension may be as low as 10% of the available pool, it should not be much lower (i.e., 1% of the available pool). If extension or PCR efficiency is dauntingly low, the PCR conditions should be reexamined and optimized as described, including buffer and enzyme concentrations, temperatures, and extension times. Switching to a different thermostable polymerase, or to a combination of polymerases, will sometimes improve primer extension. If all possible PCR optimization conditions have been addressed, poor extension efficiency could reflect a problem with the synthetic DNA. For example, the pool may not have been completely deprotected or a primer binding site may have become largely depurinated during the course of a long synthesis. Although incomplete deprotection is rarely a problem, small aliquots of the pool can be further treated with ammonia, and extension and amplification can again be assessed. If additional deprotection instead yields oligonucleotide degradation, then it is likely that apurinic sites have accumulated, and the pool will have to be resynthesized.

Anticipated Results

It is apparent from the discussion earlier in this unit that there is no one correct way to design and amplify a random sequence pool.

However, by following the protocols described above, results similar to the following should be observed.

If the integrity of the nascent, synthetic pool is good, then the primer extension efficiency (described in Support Protocol 1) should be relatively high. Figure 24.2.4 shows a typical extension reaction for a pool synthesized in the authors' laboratory (N30P, a segmentally random pool). Lane 1 is an RNA size standard; lanes 2 and 3 show a control reaction with "no template" and "no enzyme," respectively. In these lanes, only the radiolabeled 3' primer (24 nucleotides in length) is visible. Lanes 4, 5, and 6 show the primer extension reaction at various incubation times. Molecules that were incapable of full extension make up the smear leading to the full-length product. By determining the number of counts in the full-length product relative to the radiolabeled primer, the extension efficiency for the N30P pool was calculated to be ~8%. Moreover, it appeared as though the extension reaction had gone to completion within 2 min.

Assuming that the nascent pool is intact and can serve as a template for the primer extension reaction, then it should be possible to amplify the pool via the polymerase chain reaction. Figure 24.2.5 shows the results of an amplification "cycle course" for a different pool (N71, with a 71-nucleotide random sequence core). An 8-ml PCR reaction was placed in a 15-ml

Falcon tube and cycled through a series of three water baths. The samples in the figure were drawn at 0, 2, 4, 6, 8, and 10 cycles. This initial PCR reaction was only a trial, and for the final, large-scale amplification of the entire pool, a 150-ml PCR reaction was distributed to 18 Falcon tubes and 7 PCR cycles were carried out. Following amplification, a portion of the N71 pool was cloned into a TA cloning vector (Invitrogen) and ten clones were sequenced. The proportions of different nucleotides in the final pool reflected almost perfect equimolar coupling efficiencies: A, 25.22%; C, 25.37%; G, 25.82%, and T, 23.58%.

Time Considerations

The amount of time required for the protocols described in this section should not be underestimated. Pool design will require at least one day, depending on the degree of background research. It is strongly recommended that pool design be discussed with one or more colleagues prior to synthesis. The synthesis of oligonucleotides <150 bases in length can be easily accomplished in one day, allowing 1 hr to ensure proper instrument setup. Pool purification and optimization of PCR conditions should take 1 to 2 additional weeks. Finally, the actual large-scale amplification and subsequent isolation of the dsDNA pool will require the researcher's undivided attention for ~2 days.

Literature Cited

- Bartel, D.P. and Szostak, J.W. 1993. Isolation of new ribozymes from a large pool of random sequences. *Science* 261:1411-1418.
- Bartel, D.P., Zapp, M.L., Green, M.R., and Szostak, J.W. 1991. HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* 67:529-536.
- Baskerville, S., Zapp, M., and Ellington, A.D. 1995. High-resolution mapping of the human T-cell leukemia virus type 1 rex-binding element by in vitro selection. *J. Virol.* 69:7559-7569.
- Beaucage, S.L. and Caruthers, M. 2000. Synthetic strategies and parameters involved in the synthesis of oligodeoxyribonucleotides according to the phosphoramidite method. *In Current Protocols in Nucleic Acid Chemistry* (S.L. Beaucage, D.E. Bergstrom, G.D. Glick, and R.A. Jones, eds.) pp. 3.3.1-3.3.20. John Wiley & Sons, New York.
- Breaker, R.R. 1997. In vitro selection of catalytic polynucleotides. *Chem. Rev.* 97:371-390.
- Chen, Z. and Ruffner, D.E. 1996. Modified crush-and-soak method for recovering oligodeoxynucleotides from polyacrylamide gel. *BioTechniques* 21:820-822.
- Conrad, R., Keranen, L.M., Ellington, A.D., and Newton, A.C. 1994. Isozyme-specific inhibition of protein kinase C by RNA aptamers. *J. Biol. Chem.* 269:32051-32054.
- Crameri, A. and Stemmer, W.P.C. 1993. 10²⁰-fold aptamer library amplification without gel purification. *Nucl. Acids Res.* 21:4410.
- Fitzwater, T. and Polisky, B. 1996. A SELEX primer. *Methods Enzymol.* 267:275-301.
- Giver, L., Bartel, D., Zapp, M., Pawul, A., Green, M., and Ellington, A.D. 1993. Selective optimization of the Rev-binding element of HIV-1. *Nucl. Acids Res.* 21:5509-5516.
- Gold, L., Polisky, B., Uhlenbeck, O., and Yarus, M. 1995. Diversity of oligonucleotide functions. *Annu. Rev. Biochem.* 64:763-797.
- Hermes, J.D., Parekh, S.M., Blacklow, S.C., Koster, H., and Knowles, J.R. 1989. A reliable method for random mutagenesis: The generation of mutant libraries using spiked oligodeoxyribonucleotide primers. *Gene* 84:143-151.
- Jaeger, J.A., Turner, D.H., and Zuker, M. 1989. Predicting optimal and suboptimal secondary structure for RNA. *Methods Enzymol.* 183:281-306.
- Jaeger, L. 1997. The new world of ribozymes. *Curr. Opin. Struct. Biol.* 7:324-335.
- Lorsch, J.R. and Szostak, J.W. 1994. In vitro evolution of new ribozymes with polynucleotide kinase activity. *Nature* 371:31-36.
- Lyamichev, V., Brow, M.A., and Dahlberg, J.E. 1993. Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerases. *Science* 260:778-783.
- Milligan, J.F., Groebe D.R., Witherell G.W., and Uhlenbeck O.C. 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucl. Acids Res.* 15:8783-8798.
- Osborne, S.E. and Ellington, A.D. 1997. Nucleic acid selection and the challenge of combinatorial chemistry. *Chem. Rev.* 97:349-370.
- PerSeptive Biosystems. 1998. Expedite Nucleic Acid Synthesis System: User's Guide. PerSeptive Biosystems, Framingham, Mass.
- Sabeti, P.C., Unrau, P.J., and Bartel, D.P. 1997. Accessing rare activities from random RNA sequences: The importance of the length of molecules in the starting pool. *Chem. Biol.* 4:767-774.
- Singer, B.S., Shtatland, T., Brown, D., and Gold, L. 1997. Libraries for genomic SELEX. *Nucl. Acids Res.* 25:781-786.
- Stromberg, R. and Stawinski, J. 2000. Synthetic strategies and parameters involved in the synthesis of oligodeoxyribo- and oligoribonucleotides according to the H-phosphonate method. *In Current Protocols in Nucleic Acid Chemistry* (S.L. Beaucage, D.E. Bergstrom, G.D. Glick, and R.A. Jones, eds.) pp. 3.4.1-3.4.11. John Wiley & Sons, New York.

- Tuerk, C. and Gold, L. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249:505-510.
- Tuerk, C. and MacDougal-Waugh, S. 1993. In vitro evolution of functional nucleic acids: High affinity RNA ligands of HIV-1 proteins. *Gene* 137:33-39.
- Zon, G., Gallo, K.A., Samson, C.J., Shao, K., Summers, M.F., and Byrd, R.A. 1985. Analytical studies of "mixed sequence" oligodeoxyribonucleotides synthesized by competitive coupling of either methyl- or β -cyanoethyl-*N,N*-diisopropylamino phosphoramidite reagents, including 2'-deoxyinosine. *Nucl. Acids Res.* 13:8181-8196.

Contributed by Jack Pollard
Mintz Levin Cohn Glovsky Ferris and Popeo
Boston, Massachusetts

Sabine D. Bell and Andrew D. Ellington
University of Texas
Austin, Texas

In Vitro Selection of RNA Aptamers to a Protein Target by Filter Immobilization

While there are multiple possible configurations for in vitro selection experiments, this unit will describe one of the most common: selection of aptamers that bind to a protein target from a single-stranded RNA pool. Aptamers generated from these types of selection experiments can potentially function as protein inhibitors, and may find applications as therapeutic or diagnostic reagents. In short, a double-stranded DNA pool (see *UNIT 24.2*) will be transcribed to generate a single-stranded RNA pool (Basic Protocol 1). The initial concentration of protein target to be used is determined by labeling an aliquot of the pool (see Support Protocol 1) and performing the binding assay as described in Support Protocol 2. Following purification, the pool is mixed with the protein target. Binding species are separated from nonbinding species by filtration (see Basic Protocol 2). RNA:protein complexes are then eluted from the filter and binding species are amplified by a combination of reverse transcription, the polymerase chain reaction (PCR), and in vitro transcription (see Basic Protocol 3). The progress of the selection will be monitored by assaying the affinity of the radiolabeled RNA pool for the protein target after several rounds of selection (see Support Protocol 3). These steps are then repeated until a significant increase in binding is observed or until the diversity of the pool has been completely plumbed. The procedure is summarized in Figure 24.3.1.

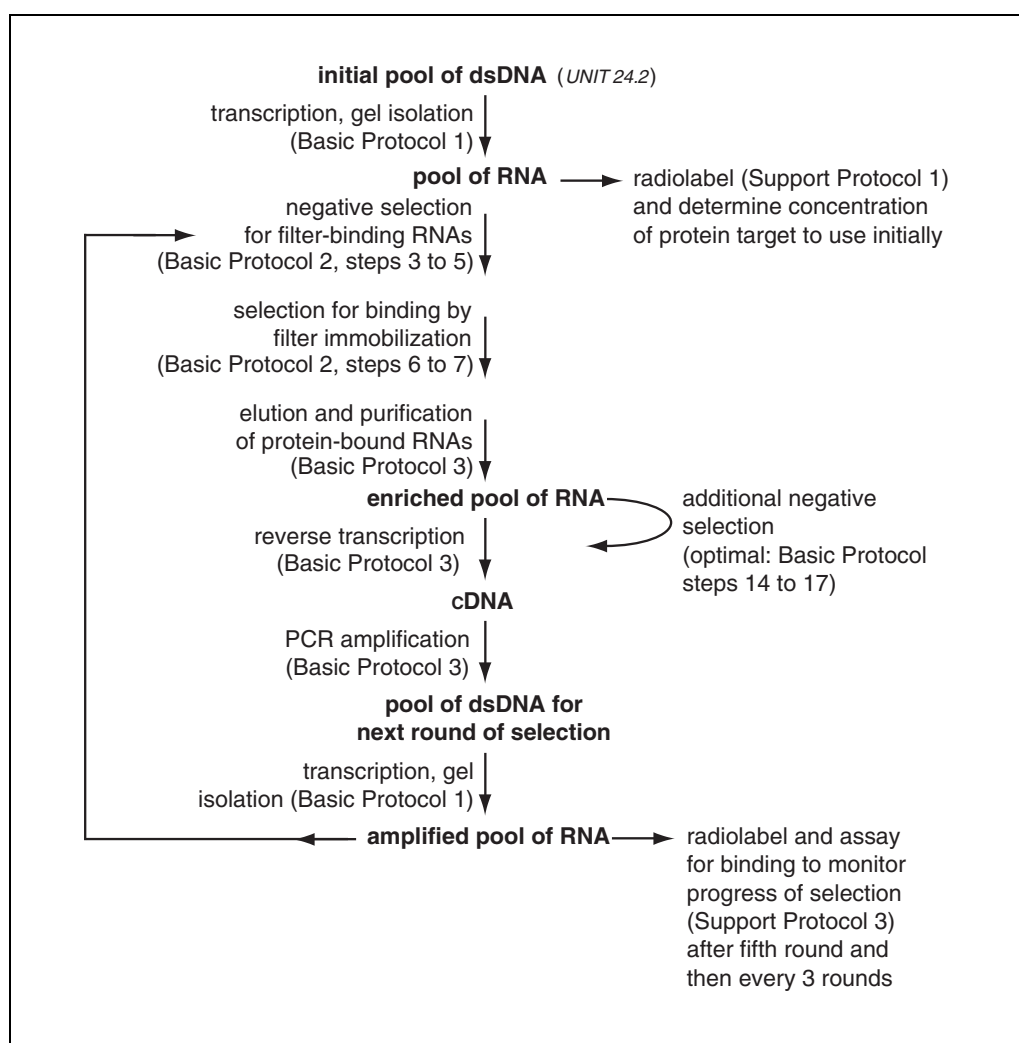


Figure 24.3.1 Steps involved in in vitro selection of RNA aptamers.

Contributed by Sulay D. Jhaveri and Andrew D. Ellington
Current Protocols in Molecular Biology (2000) 24.3.1-24.3.25
 Copyright © 2000 by John Wiley & Sons, Inc.

Generation and Use of Combinatorial Libraries

24.3.1

Supplement 52

TRANSCRIPTION AND ISOLATION OF RNA POOLS

The following protocol describes the preparation of the RNA pool to be used for selection. Starting from the dsDNA pool, the RNA is transcribed and purified by denaturing polyacrylamide gel electrophoresis. Recovery of the RNA from the gel is followed by ethanol precipitation of the RNA. Additional instructions can be found in *UNIT 3.8*. The directions provided here are specific for the isolation of nucleic acid pools. As is the case for the original amplification of DNA pools (*UNIT 24.2*), many of the procedures described here can potentially lead to the cross-contamination of different RNA selection experiments or different generations of the same selection experiment. To avoid cross-contamination, it is wise to always use barrier tips and to use disposable plastic Pasteur pipets rather than automatic micropipettors for large-volume transfers.

Materials

Double-stranded DNA pool (*UNIT 24.2*)
Transcription mix (see recipe)
RNase-free DNase (e.g., RQ1 DNase; Promega)
10% polyacrylamide denaturing gel (see recipe and *UNIT 2.12*)
2× denaturing dye (see recipe)
TBE buffer (*APPENDIX 2*)
5 M NaCl
90% and 100% ethanol
TE buffer, pH 8.0 (*APPENDIX 2*)
37° to 42°C and 65° to 75°C water baths
Fluorescent TLC plate (VWR) wrapped in plastic wrap
Spectrophotometer
Additional reagents and equipment for denaturing polyacrylamide gel electrophoresis (*UNIT 2.12*)

NOTE: All solutions and buffers should be freshly treated with DEPC (*UNIT 4.1*). Use sterile, disposable plasticware where possible. See Chapter 4 introduction and *UNIT 4.1* for guidelines on standard methods to protect against contaminating RNases.

Perform initial round of transcription

Use the double-stranded DNA pool generated in *UNIT 24.2* as a template for in vitro transcription with T7 RNA polymerase.

1. Add ~1 µg of double-stranded DNA template generated as in *UNIT 24.2* to transcription mix for a 20 µl total reaction volume. Incubate reaction overnight at 37° to 42°C.

Depending on the length and initial complexity of the pool, 1 µg of double-stranded DNA will represent ~10¹³ different sequences. The transcription reaction will yield ~10 to 50 µg of RNA, and thus from 20 to 100 copies of each sequence originally present. If more RNA is desired for initial or subsequent rounds of selection, a proportionately larger transcription reaction should be attempted.

The success of the transcription reaction can sometimes be monitored by observing the formation of a precipitate over time. This precipitate is likely a complex between magnesium and the pyrophosphate released from each polymerized ribotide. However, this rule is not absolute: many successful transcription reactions have no precipitate; some unsuccessful reactions have a precipitate.

In general, though, the authors have found that much higher yields of RNA can be obtained from commercial transcription kits (e.g., Ampliscribe from Epicentre Technologies or Megashortscript from Ambion) than from home-made in vitro transcription reactions. The use of such kits reduces the amount of time necessary to generate an adequate amount of

transcript (1 to 4 hr rather than overnight). Depending on the amount of RNA that is desired for each round of selection, such kits may also (surprisingly) represent a more cost-efficient alternative.

In some instances it will be desirable to radiolabel the RNA. For example, it is relatively easy to determine whether and how much RNA binds to a filter in the presence or absence of a protein target by radiolabeling the initial pool (see Support Protocol 1). An [α - 32 P]nucleoside triphosphate—e.g., 0.5 μ l [α - 32 P]UTP (NEN Life Science Products) in a 20- μ l total volume—can be included in the reaction mixture in addition to all the other reagents. Varying the proportion of “hot” to “cold” nucleoside triphosphates can control the specific activity of the RNA pool. Since the overall yield of the transcription reaction will generally be important, the specific activity of the nucleoside triphosphate mixture should be varied by increasing the amount of radioactive nucleotide added, rather than by decreasing the amount of unlabeled nucleotide present. Again, commercial transcription kits can be obtained that are geared towards the incorporation of labeled nucleoside triphosphates (RiboScribe, Epicentre).

2. In order to remove DNA from the transcription reaction, after the transcription incubation has been completed, add 5 to 10 U of RNase-free DNase and incubate 25 min at 37°C.

Because individual members of the double-stranded DNA library can potentially bind nonspecifically to either the target or to the selection matrix and subsequently be amplified, the DNA template should be removed from the transcription reaction according to this step, prior to proceeding with the selection.

It is essential that RNase-free DNase, such as RQ1 DNase (Promega) be used, otherwise contaminating ribonucleases may destroy the newly transcribed RNA pool. An alternative would be to add RNase inhibitors to impure DNases, but such inhibitors themselves frequently contain endogenous ribonucleases that can be released during the incubation.

Purify the RNA pool

The RNA pool should generally be purified by denaturing gel electrophoresis.

3. Prepare a 0.75-mm thick, denaturing 10% acrylamide gel (see Reagents and Solutions and, e.g., UNIT 2.12).

A 10% acrylamide concentration is convenient for the purification of RNA molecules from 45 to 70 nucleotides in length. However, the concentration of acrylamide used to separate the full-length transcript from incomplete transcripts is ultimately contingent upon the size of the RNA and should be chosen so that the RNA will migrate approximately half-way through the gel when the loading dye has reached the bottom (see UNIT 2.12).

If the RNA sample contains a significant amount of nascent structure (for example, a doped sequence population that is based on a tightly folded secondary structure) it may not fully denature. Thus, it may be advisable to warm the gel to ~55°C by first pre-running the gel at a higher voltage (300 to 400 V). The temperature of the gel can be monitored using adherent thermometers (VWR).

In some cases, very large amounts of RNA may need to be purified (for example, the initial transcription of an extremely complex DNA library may yield upwards of a milligram or more of an RNA library). In these instances, it may be desirable to purify the RNA library by either gel-filtration or ion-exchange chromatography (e.g., Qiagen RNA kit). However, the purification of the initial or subsequent pools should never be neglected, as fore-shortened amplicons can arise and overtake selected populations.

4. Fully denature the RNA pool by adding an equal volume of 2× denaturing dye, and heat the RNA-dye mix 3 min at 65° to 75°C.

Although each species in the pool has a different sequence and shape, they should migrate similarly when fully elongated.

Using a higher temperature or longer denaturing time risks hydrolysis of the RNA into smaller fragments by the high concentration of Mg^{2+} present in the transcription buffer.

5. Thoroughly rinse each well with TBE buffer using a plastic Pasteur pipet prior to loading (to remove urea, which will otherwise leach into the wells and form a barrier between the loaded sample and the gel). Load samples directly on the gel at 10 to 20 μg per 1-cm-wide lane (i.e., load a single transcription reaction in 2 to 3 lanes of the gel). Run electrophoresis for 1 to 2 hr at 150 to 250 V, until the bromophenol blue dye front reaches the bottom of the gel.

If the wells are not cleaned prior to loading, the resolution of the separation can be compromised, especially if large amounts of RNA are being isolated.

6. Visualize the RNA bands by UV shadowing on a fluorescent TLC plate covered with plastic wrap, then excise the bands. Be sure to cut with a sharp razor blade and cut only the shadowed regions that contain the bulk of the RNA.

There may be extra bands in the lane that correspond to incomplete transcripts or undigested DNA. The use of a size standard in a neighboring lane is recommended. Note, however, that the size standard should not itself be amplifiable, as cross-contamination of a single sequence with the RNA pool would drastically skew the distribution of sequences in the purified pool. Similarly, the razor blade used for excision should not have come into contact with other potentially amplifiable sequences, and should either be fresh or should have been cleaned extensively. Finally, if multiple selections are being carried out in parallel they should not be purified on the same gel.

7. Immerse the gel slices in RNase-free water at $\sim 400 \mu\text{l}$ water/ cm^2 of gel (typically, slices from 2 lanes) and incubate at 37°C overnight with agitation to elute the RNA pool.

For a quicker elution step, incubate the slices at 65° to 75°C for 1 hr. However, the amount of RNA recovered will be lower, and there is a greater risk of degradation. The gel can be macerated to increase the speed or efficiency of recovery, but in this case small fragments of acrylamide may remain in the eluant. The eluate can be filtered through an $0.45\text{-}\mu\text{m}$ nitrocellulose membrane to remove acrylamide fragments.

Collect and quantitate the RNA

8. Use a plastic Pasteur pipet to separate the RNA-containing eluate from the gel slice. Add NaCl (from 5 M stock) to a final concentration of 0.3 M and ethanol precipitate the RNA by adding 2 vol ethanol. Mix and incubate at -20°C for 30 min or -70°C for 10 min. Microcentrifuge 20 to 40 min at maximum speed, 4°C , to recover the precipitate.

Smaller RNA molecules (20 to 80 nucleotides in length) can be more efficiently precipitated with 2.5 vol of ethanol.

The authors frequently include 1 μl of a 1 mg/ml glycogen solution to increase the yield of nucleic acid precipitate and to better visualize the pellet. If the selection target binds to or interacts with glycogen, then this step should be omitted. Transfer RNA can also be used as a carrier, but will obfuscate the quantification of the pool RNA (see below).

9. Wash the RNA pellet with cold 90% ethanol and dry the pellet.

The pellet can be air dried, dried under a nitrogen or argon stream, or dried in a SpeedVac evaporator. The first method is least likely to result in cross-contamination of nucleic acid species; the last method is least likely to lead to degradation. In any event, keep the tube covered with Parafilm to avoid inadvertent nuclease contamination (poke holes in the Parafilm with a sterile needle to allow evaporation to proceed).

If the RNA pool is particularly short (≤ 50 nucleotides) use cold 95% ethanol for the wash step.

10. Resuspend the RNA pellet in 25 μ l TE buffer, pH 8.0.

To avoid disturbing the composition of the selection buffer, the pellet can also be resuspended in RNase-free water. However, the small amount of EDTA present in TE buffer will limit ribonuclease degradation of the pool, since ribonucleases frequently require a divalent metal. In some instances, though (e.g., small-volume PCR reactions), the presence of EDTA may have to be compensated for by adding more magnesium to the reaction.

11. Estimate the quantity of the RNA photometrically by measuring the absorbance at 260 nm.

Use an extinction coefficient of 0.025 ml cm^{-1} μg^{-1} (see, e.g., APPENDIX 3D). In practical terms, measure the A_{260} of a 1:500 dilution of the sample (2 μ l dissolved in 1 ml RNase-free water) and multiply the absorbance by 20 to obtain the number of $\mu\text{g}/\mu\text{l}$ in the original sample. Do not attempt to calculate concentrations using absorbance readings less than ~ 0.03 . The A_{260}/A_{280} ratio should be somewhere between 1.8 and 2.2. Ratios outside of this range make the purity of the original RNA sample suspect (with residual acrylamide being the most likely contaminant), and the sample should be reprecipitated prior to use.

RADIOLABELING THE RNA FOR USE IN AN INITIAL AFFINITY ASSAY

Radioactive RNA can be generated either by incorporation of an [α - ^{32}P]nucleoside triphosphate during transcription or by transfer of the terminal phosphate of γ - ^{32}P ATP to the 5' terminus of a dephosphorylated RNA molecule. The authors tend to prefer the latter method, despite the additional labor involved in preparation, because the specific activity of the sample is higher, less RNA is required for assays, and dissociation constants are correspondingly easier to compute.

Materials

- RNA pool (see Basic Protocol 1)
- 10 \times alkaline phosphatase buffer (Boehringer Mannheim)
- Calf alkaline phosphatase (Boehringer Mannheim)
- 1:1 phenol/chloroform (UNIT 2.1A)
- Chloroform
- 5 M NaCl
- 90% and 100% ethanol
- 10 \times PNK buffer (New England Biolabs)
- T4 polynucleotide kinase (PNK; New England Biolabs)
- 167 mCi/ml [γ - ^{32}P]ATP (7000 Ci/mmol; ICN)
- 4 M ammonium acetate
- 42 $^{\circ}$ and 75 $^{\circ}$ C water baths

NOTE: All solutions and buffers should be freshly treated with DEPC (UNIT 4.1). Use sterile, disposable plasticware where possible. See Chapter 4 introduction and UNIT 4.1 for guidelines on standard methods to protect against contaminating RNases.

Dephosphorylate the 5' triphosphate termini of the isolated RNA pool

1. Mix the following components:

- 1 μg RNA in <3.5 μl volume
- 0.5 μl 10 \times alkaline phosphatase buffer
- 1 μl (1 U) calf alkaline phosphatase
- x μl RNase-free water for a total reaction volume of 5 μl .

The RNA sample may need to be reprecipitated to obtain an adequately concentrated sample. If so, the precipitate can be resuspended directly in the reaction buffer or mixture.

SUPPORT PROTOCOL 1

Generation and Use of Combinatorial Libraries

24.3.5

Calf alkaline phosphatase is preferred over bacterial alkaline phosphatase because the activity can be heat-killed (see step 4) prior to the addition of the radiolabel.

2. Incubate at 42°C for 20 min to 2 hr.
3. Add 95 µl RNase-free water.
4. Heat denature the calf alkaline phosphatase 10 min at 75°C.
5. Perform a phenol/chloroform extraction (see Basic Protocol 2, step 10).

If the sample will be gel-isolated, this step can be omitted. If the radiolabeled sample will merely be precipitated prior to use, this step should be included.

6. Ethanol precipitate the RNA in the presence of 0.3 M NaCl and wash the pellet with 90% ethanol (see Basic Protocol 1, steps 8 and 9).

Avoid precipitating RNA in the presence of ammonium acetate since ammonium ions inhibit the T4 polynucleotide kinase used in the next step.

7. Resuspend the dried pellet in a minimal volume (3 to 10 µl) of RNase-free water.

Perform kinase reaction

8. Set up the kinase reaction as follows:

0.5 to 3 µl dephosphorylated RNA pool (from step 7)
0.5 µl 10× PNK buffer
1 µl (10 U) T4 polynucleotide kinase (PNK)
0.5 µl (83 µCi) [γ -³²P]ATP (7000 Ci/mmol, ICN)
x µl RNase-free H₂O for a total volume of 5 µl.

Only a very small amount of RNA will be used in the binding assay (~50 pM in a 100 µl reaction). Unless multiple experiments are contemplated, the specific activity of the sample can be kept quite high by using a very small amount of RNA in the kinase reaction.

9. Incubate for 1 hr at 37°C.
10. Add 95 µl RNase-free water.
If the sample will be gel-isolated, this step can be omitted.
11. Perform a phenol/chloroform extraction (see Basic Protocol 2, step 10).
If the sample will be gel-isolated, this step can be omitted.
12. Ethanol precipitate the RNA in the presence of 2.0 M ammonium acetate (i.e., by adding an equal volume of 4.0 M ammonium acetate).

The use of ammonium acetate inhibits the precipitation of nucleotides and small transcripts. However, if the RNA pool is short, the precipitation may also be inefficient. If the sample will be gel-isolated, then this step can be omitted.

13. *Optional:* In order to fully separate the radiolabeled RNA pool from unincorporated nucleotides, partially degraded transcripts, and enzymes, isolate the transcript as described in Basic Protocol 1, steps 3 to 9).

If this is done, the phenol/chloroform extractions and the final precipitation of the RNA (steps 10 to 12 of this protocol) can be omitted. The chief disadvantages of gel isolation are the time required for sample preparation and the relatively low efficiency of recovery of the radiolabeled RNA pool. However, since only a small amount of RNA pool is required for the binding assay such low yields can frequently be tolerated. The authors frequently gel isolate radiolabeled RNA pools to ensure the integrity of RNA samples prior to carrying out binding assays.

BINDING ASSAY WITH THE END-LABELED RNA POOL TO DETERMINE THE OPTIMAL PROTEIN CONCENTRATION FOR SELECTION

To determine the initial concentration of a protein target to be used in a selection experiment, it is necessary to measure the affinity of the unselected pool for the protein target. The aggregate dissociation constant of the pool:protein complex can be calculated by determining the fraction of radioactively labeled RNA that can be bound at various protein concentrations.

The radiolabeled RNA is incubated in the binding buffer and protein solutions are added. The binding reaction is filtered through a vacuum manifold containing nitrocellulose and nylon membranes and the fraction of RNA bound to the target is calculated to obtain a value for the dissociation constant. The nitrocellulose membrane will capture RNA:protein complexes, while the nylon membrane will capture all free RNA that flows through the nitrocellulose membrane.

Materials

Radiolabeled RNA pool (Support Protocol 1)
Binding buffer (see Critical Parameters)
Target protein
65° to 75°C water bath
Milliblot apparatus (Schleicher & Schuell)
Nylon transfer membrane (Hybond N+, Amersham Pharmacia Biotech)
0.45- μ m nitrocellulose transfer and immobilization membrane (Midwest Scientific)
Glass plate
Phosphorimager (Molecular Dynamics) and screen *or* X-ray film and densitometer
Graphing software (e.g., Kaleidograph from Synergy Software)

NOTE: All solutions and buffers should be freshly treated with DEPC (UNIT 4.1). Use sterile, disposable plasticware where possible. See Chapter 4 introduction and UNIT 4.1 for guidelines on standard methods to protect against contaminating RNases.

Set up binding reactions

1. Collect the RNA precipitate by centrifugation and resuspend the radiolabeled RNA in a minimal volume (i.e., 5 to 10 μ l) of RNase-free water. Dilute the RNA sample with binding buffer to a final concentration of 100 pM.

The concentration can be very roughly estimated by assuming full recovery of the RNA sample. Differences between estimated and actual concentrations are less important because the RNA sample will be limiting relative to the amount of protein sample present in the binding reaction.

The binding assay will yield 11 data points in triplicate (see below). Since each data point will be generated from a 50- μ l binding reaction, 2 ml of the RNA solution should be adequate. If the specific activity of the RNA is not high enough, a higher concentration of RNA may be used, but that will complicate the assumption that RNA is limiting and hence the calculation of the K_d .

2. To ensure that each species in the RNA pool folds into the most accessible or most stable conformation, heat the RNA pool in 25- μ l binding buffer to 65° to 75°C for 3 min and then allow the sample to cool to room temperature over ~10 min.
3. Add 25 μ l of the protein target in binding buffer to the thermally equilibrated RNA from step 2. Use ten different protein concentrations ranging from 1 μ M to 50 pM.

Also include one data point with no protein to measure the filter-binding ability of the pool itself.

The original protein solution should be sufficiently concentrated for all of the dilutions. To ensure consistency between samples, serial dilutions of the 1 μM sample can be made. The authors suggest the following concentrations: 1 μM , 333 nM, 111 nM, 37 nM, 12 nM, 4.1 nM, 1.4 nM, 460 pM, 152 pM, 51 pM, (i.e., 1 μM , and subsequent 1/3 dilutions) and a “no protein” control. For statistically significant results perform the binding assay in triplicate.

4. Incubate the binding reaction at room temperature for 15 min to 1 hr (see Critical Parameters).

Perform filter binding

5. Assemble the Milliblot apparatus (Fig. 24.3.2). Lay the nylon transfer membrane on top of the perforations in the middle section. Moisten the nylon membrane and lay the nitrocellulose membrane on top of the nylon membrane, taking care to avoid the formation of bubbles between the two membranes. Cover and tighten the brackets.

Prior to filtering the binding reactions, wash the wells that will be used with binding buffer and check for leaks. When the manifold is used in conjunction with an aspirator; turn the water faucet to a level that causes liquid to pass slowly through the membranes (i.e., 100 μl every 3 sec).

Since there are so many binding reactions, it is more convenient to use a manifold apparatus that can accommodate multiple filtrations (up to 96 slots) than to assemble 33 individual filter holders.

6. Filter the binding reactions and wash three times with binding buffer.

When pipetting onto the manifold, dispense the liquid slowly and evenly. Try to keep the membrane constantly hydrated during each wash step. Keep the pipet tip close to the membrane to avoid bubble formation, but not so close as to risk damaging the membrane.

7. Disassemble the manifold apparatus and transfer the membranes to a clean paper towel. Air dry for ~5 min. Handle membranes with a clean pair of forceps or tweezers.

The forceps or tweezers can be quickly cleaned with ethanol and RNase-free water prior to contacting the filters.

8. Transfer the membranes to a glass plate, cover with plastic wrap, and expose to a phosphor screen (e.g., Phosphorimager) or X-ray film for 4 to 12 hr.

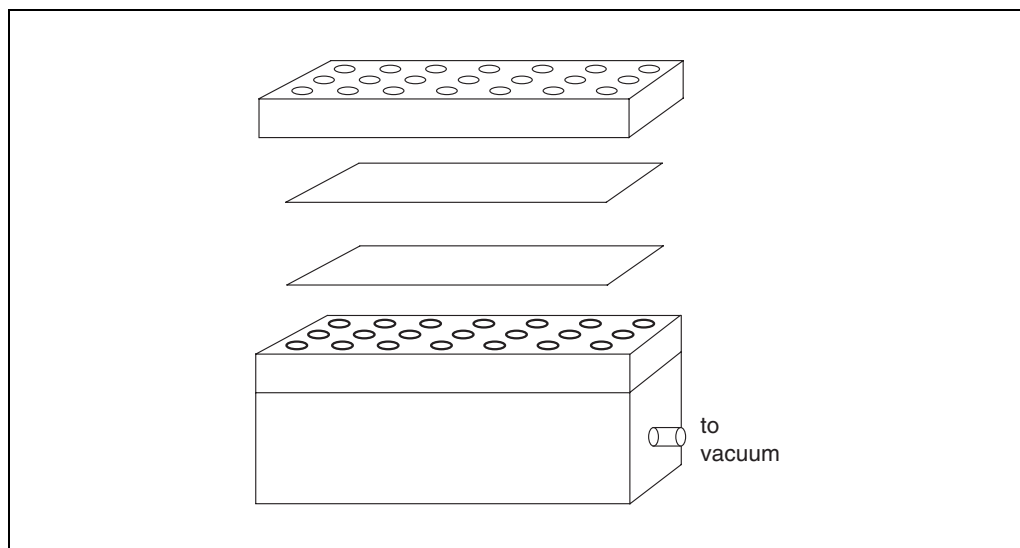


Figure 24.3.2 Assembly of the Milliblot apparatus used for binding assays.

If the samples have a very high specific activity, the exposure time can be reduced to 5 to 60 min.

9. Measure the radioactivity using the Phosphorimager or a densitometer if X-ray film was used to develop the image, and calculate the binding percentages as follows:

$$\text{Fraction bound} = \text{cpm on nitrocellulose} / (\text{cpm on nitrocellulose} + \text{cpm on nylon})$$

If X-ray film was used to develop the image, then a digitizer (densitometer) should yield similar results to those obtained with a Phosphorimager.

10. Plot the fraction bound as a function of the concentration of unbound protein. Fit the points to a curve using graphing software (e.g., Kaleidograph) and obtain a value for the aggregate parent dissociation constant. Within the Kaleidograph program, fit the curve using the equation $y = m_1 m_0 / (m_0 + m_2)$, where y = the fraction of RNA bound, m_0 = concentration of unbound protein, m_1 = the extrapolated activity of the RNA at an infinite protein concentration (maximal value of fraction bound), and m_2 = the apparent dissociation constant.

The apparent K_d is equal to the concentration of unbound protein at half the maximal value of fraction bound.

ISOLATING A FUNCTIONALLY ENRICHED POOL

In the following protocol, the RNA pool is partitioned to isolate those species that bind to the target protein and not to the filter. RNAs that are coimmobilized with the target are eluted off the filter under denaturing conditions and subsequently isolated and amplified.

Materials

RNA pool (see Basic Protocol 1)
Binding buffer (see Critical Parameters)
Elution buffer (see recipe)
1:1 phenol/chloroform (see UNIT 2.1A), ice-cold
Chloroform
Isopropanol
65° to 75°C and 100°C water baths
Filter holders (Nuclepore)
13-mm, 0.45- μm HAWP nitrocellulose disk filters (Millipore)
5-ml syringe
Vacuum manifold

NOTE: All solutions and buffers should be freshly treated with DEPC (UNIT 4.1). Use sterile, disposable plasticware where possible. See Chapter 4 introduction and UNIT 4.1 for guidelines on standard methods to protect against contaminating RNases.

Partition the pool

1. Use ~5 μg of the RNA pool (~ 10^{13} to 10^{14} sequences) for selection.

Using significantly lower quantities of RNA may affect the diversity of the population in the initial rounds of selection. Using significantly higher quantities may lead to precipitation of the nucleic acid pool. Irvine et al. (1991) have devised a formula to determine the optimum protein and RNA concentration in order to minimize the number of rounds of selection, based on the K_d of the starting pool, the desired K_d , and the fraction of free RNA molecules that partitions as nonspecific background versus the fraction of RNA molecules that forms specific RNA:protein complexes. Empirically, the concentrations of many available protein targets will be in the nanomolar range, and a 1- to 10-fold excess of the RNA pool should suffice for early rounds of selection.

BASIC PROTOCOL 2

If only a small amount of RNA pool is initially recovered from the gel, be sure to save at least some sample for the “no protein” control (see below).

2. To ensure that each species in the RNA pool folds into the most accessible or most stable conformation, heat the RNA pool in 50 to 100 μl binding buffer (see Critical Parameters for discussion on choosing a binding buffer) to 65° to 75°C for 3 min and then allow the sample to cool to room temperature over ~10 min.

Since ionic strength, monovalent and divalent cation concentrations, pH, temperature, and buffer concentrations can all influence interactions with the target, it is usually wise to keep all of these parameters constant during the early rounds of selection when productive binding species are accumulating. Hence, the binding buffer, equilibration time, and preparation of the RNA for selection should be kept uniform until a significant interaction between pool and target is observed (see Critical Parameters for discussion of stringency of selection).

Higher temperatures can be used for thermal equilibration, but the presence of divalent metal ions in the selection buffer can lead to RNA degradation.

3. Prior to the addition of the protein target, perform a negative selection to remove any filter-binding species that may be in the population. Moisten a filter disk with buffer and lock it into a filter holder (Fig. 24.3.3).

Negative selection to remove filter-binding species is an extremely important step in the selection procedure. Filter-binding species are typically more numerous in a naive RNA population than are aptamers. If filter-binding species are not efficiently sieved from the population, they will quickly accumulate to the point where it may be difficult (and likely impossible) to select protein-binding species. If the potential for accumulating filter-binding species is large (i.e., the target has a low initial affinity for a pool, or selections with DNA pools), then repeat the preselection filtration to remove any filter-binding species that may persist or carry out a post-selection filtration (see optional steps 14 to 17, below). If filter-binding species do accumulate during a selection experiment, it is usually wisest to repeat the selection starting with a different pool that can be amplified with different primers.

In addition to filter-binding species, replication parasites (see Critical Parameters for discussion on parasites) can accumulate in and overrun a selected population. A separate regime is required to avoid these selection predators.

4. Load the binding buffer onto the filter. Place the pipet tip just above the filter to avoid the formation of any bubbles. Lock a 5-ml syringe to the top of the filter holder and apply gentle pressure to force the liquid out of the filter holder and into a collecting tube.

Prior to filtering the RNA, it is important to wash the nitrocellulose filter disk with binding buffer and check for leaks in the assembled filter holder. The syringe should form a tight seal with the filter holder. The pressure applied should be just enough to force the liquid through without rupturing the membrane. The formation of foam at the bottom of the filter holder or the presence of a hissing sound when pressure is applied indicates that the pressure is too high, and the integrity of the seal or the membrane may have been breached. Test for leaks every time the filter holder is assembled in order to avoid substantial loss of sample.

5. Load the RNA solution onto the filter. Place the pipet tip just above the filter to avoid the formation of any bubbles. Lock a 5-ml syringe to the top of the filter holder and apply gentle pressure to force the liquid out of the filter holder and into a collecting tube.

Since there will still be some amount of liquid retained by the filter and filter holder, it is necessary to wash the filter with an equal amount of binding buffer to maximize the collection of non-filter-binding species. Discard the filter.

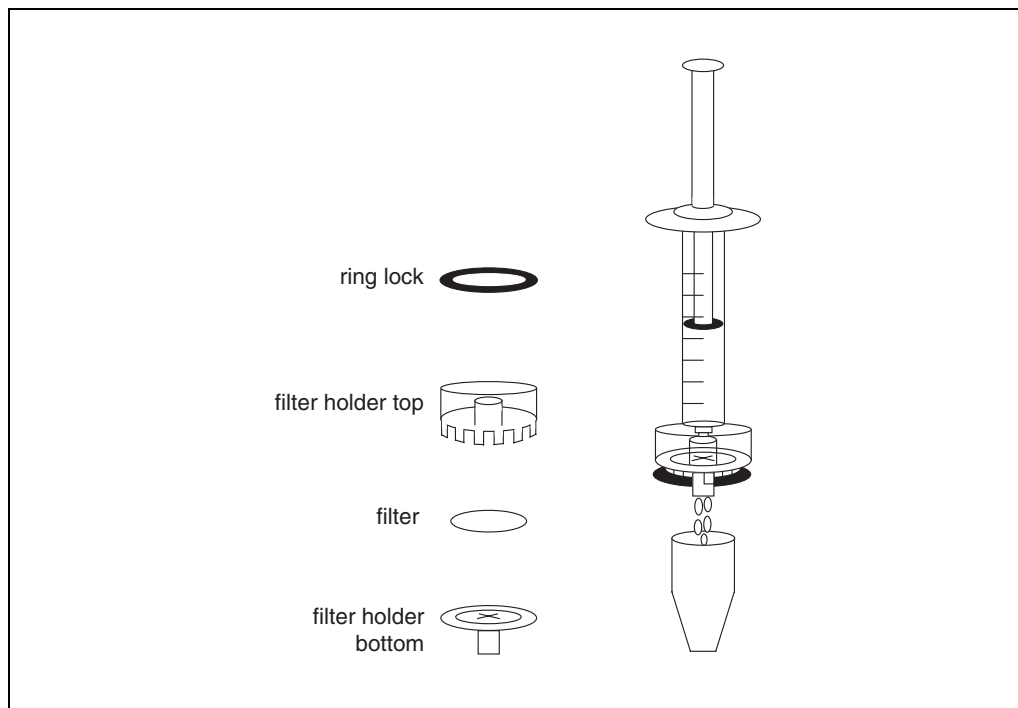


Figure 24.3.3 Components and assembly of filter holder used during selection.

6. Add the protein target and any competitors, specific and/or non-specific, to the filtrate. Allow the binding reaction to equilibrate.

In selection experiments that targeted the cytokine bFGF, the authors used an equimolar protein-to-RNA ratio for the first two rounds of selection and decreased it 10-fold after two rounds and 60-fold after another two, yielding a functionally-enriched pool after 6 rounds of selection and amplification (Table 24.3.1). The final volume of the binding reaction should be from 100 to 200 μ l. In addition, to ensure that the selected RNAs are actually binding to the target and not to the filter, a parallel binding reaction in the absence of protein can be carried out intermittently. The authors strongly suggest that “no protein” controls be scrutinized after 0, 5, 8, and 11 rounds of selection.

The choice of selection conditions is probably the second most important determinant (following the choice of target) for whether a selection experiment will succeed or fail. While general guidelines for modulating the stringency of selection can be recommended (see Critical Parameters for comments on the stringency of selection), every target and every selection are different and no precise guidelines for success can be provided. In general, the stringency of selection should be lower in the early rounds of selection and higher in the later rounds. This will give binding species an opportunity to establish themselves in the population relative to filter-binding species.

It should be noted that there is some danger of cross-contaminating the selected pool with the “no protein” control. Basically, executing the “no protein” control is identical to selecting for protein-independent (filter) binding species, hence DNA arising from the “no protein” control should be handled with care.

7. Attach the filter holder to a vacuum manifold (which is used here to maintain a constant negative pressure during filtration, so that each round of selection is similar and reproducible). Apply a negative pressure of 5 in. of Hg to the filter holder. Pipet the binding reaction directly onto the filter with the tip just above the filter, avoiding the formation of bubbles, which may lead to an uneven application of the sample to

Table 24.3.1 Progress of N30 Selection Against bFGF^a

Round	Input (RNA) nM	Input (BFGF) nM	(RNA):(bFGF)	% bound to protein	% bound to filter
1	800	760	1.05	2.1	2.3
2	800	760	1.05	—	—
3	800	76	10.5	—	—
4	800	76	10.5	6.0	4.0
5	800	13	61.5	—	—
6	800	13	61.5	17.0	0.4

^aPools were assayed in a 50- μ l reaction at a concentration of 75 nM in the presence and absence of equimolar protein.

the filter and impede the flow of liquid through the filter. Wash the filter with 3 vol of binding buffer.

Varying the strength of the vacuum, uneven application of the sample, and formation of bubbles during wash steps may result in inefficient sieving of binding from nonbinding species, and hence may reduce the efficiency of an individual round of selection. However, the selection as a whole is fairly robust with respect to changes in these parameters. In other words, even if steps are not performed perfectly, the selection can be carried forward.

Elute RNA off the filter

8. Remove the filter containing RNA:protein complexes from the filter holder using sterile forceps and place it in a 1.5-ml microcentrifuge tube. Transfer the filter quickly, in order to avoid ribonuclease contamination from the surrounding environment.
9. Add 200 μ l of elution buffer and heat for 5 min at 100°C to elute RNA molecules from the protein and filter. Remove the eluate to a tube and repeat with fresh elution buffer.

Two shorter, smaller-volume elutions will more efficiently recover intact RNA than one long, large-volume elution.

10. To remove residual peptide fragments or proteins that may have coeluted with the RNA, add an equal volume (i.e., 400 μ l) of cold, 1:1 phenol/chloroform. Vortex, then microcentrifuge 1 min at maximum speed to separate the liquid phases (the RNA should be in the top, aqueous phase). Transfer the aqueous phase to a new 1.5-ml microcentrifuge tube.

Avoid transferring phenol/chloroform with the aqueous layer, as it can interfere with subsequent enzyme reactions. Nevertheless, the aqueous phase will sometimes appear milky, especially at low temperatures, due to the presence of dissolved phenol-chloroform.

11. Extract the eluate with a similar volume of chloroform to remove any residual phenol.

Avoid transferring chloroform with the aqueous layer, as it can interfere with subsequent enzyme reactions.

12. Dilute the eluate with an equal volume (~400 μ l) of RNase-free water and add 800 μ l of isopropanol, then chill 20 min at -20°C to precipitate.

A carrier such as glycogen can be added to aid precipitation.

The elution buffer contains a high concentration of urea. Dilution with 400 μ l water and precipitation with isopropanol is necessary to avoid the formation of salt precipitates, which appear as oily, unstable droplets in the bottom of the microcentrifuge tube following

centrifugation. If such “salt pellets” appear, additional water should be added to the sample, the mixture should be homogenized, and the precipitation repeated.

13. Microcentrifuge 30 min at maximum speed, remove the supernatant, and resuspend the RNA sample in 12 ml sterile RNase-free water.

Perform an additional negative selection (optional steps)

An extremely effective method for ridding the population of filter-binding species is to carry out an additional negative selection following the selection for binding species, but prior to amplification. However, at early stages of the selection, an additional post-selection filtration step may reduce the complexity of the selected population. Therefore, it is recommended that post-selection filtration only be carried out following the second round of selection. Post-selection filtration can also be used to successfully remove filter-binding species that have begun to accumulate and overrun a selected population. However, once filter-binding species have established themselves, even a combination of pre- and post-selection filtrations may not allow specific binding species to regain a selective advantage. If a simple regime of pre- and post-filtration negative selections does not succeed in drastically reducing or eliminating established filter-binding species, the selection should be repeated with a different RNA pool that can be amplified with different primers, as recommended above.

14. Resuspend the selected RNA pellet in 50 μ l binding buffer.
15. Assemble the filter holder with a fresh filter disk as described above.
16. Filter the sample and wash as described above.
17. Discard the filter disk and ethanol precipitate the RNA filtrate.

A carrier (glycogen) can be added to improve the efficiency of precipitation.

If the binding buffer contains a high (>0.5 M) salt concentration, dilute the filtrate with an equal volume of RNase-free water and precipitate with isopropanol instead.

AMPLIFYING SELECTED BINDING SPECIES

In the following steps, RNA species that survived the positive and negative selection steps are reverse transcribed to generate a cDNA library, which is subsequently amplified by PCR. The double-stranded DNA resulting from these steps comprises the pool from which the next round of selection will begin. While the authors have found that reverse transcription and PCR steps can be combined (steps 1a to 3a) for some of our selections, this is not universally true. To obtain the highest yield of RNA and DNA products, it is frequently desirable to carry out separate reverse transcription and PCR reactions (steps 1b to 3b).

Materials

- Selected RNA pool (see Basic Protocol 2)
- TE buffer, pH 8.0, or RNase-free water
- RT-PCR mix (see recipe)
- 10 \times RT buffer (see recipe)
- 20 μ M 3'-end primer
- 20 μ M 5'-end primer
- PCR mix (see recipe)
- 4 mM dNTP mix (*UNIT 3.4*)
- AMV reverse transcriptase (USB)
- 6 \times nondenaturing dye: 0.6% bromphenol blue in TBE buffer
- 4% NuSieve agarose gel (FMC Bioproducts; also see e.g., *UNIT 2.6*)
- 10 μ g/ml ethidium bromide solution (*APPENDIX 2*)

**BASIC
PROTOCOL 3**

**Generation and
Use of
Combinatorial
Libraries**

24.3.13

TBE buffer (*APPENDIX 2*)
4 M ammonium acetate
100% ethanol
Thermal cycler (e.g., MJ Research)

Additional reagents and equipment for the polymerase chain reaction (Chapter 15)
and agarose gel electrophoresis (e.g., *UNIT 2.6*)

NOTE: All solutions and buffers should be freshly treated with DEPC (*UNIT 4.1*). Use sterile, disposable plasticware where possible. See Chapter 4 introduction and *UNIT 4.1* for guidelines on standard methods to protect against contaminating RNases.

Amplify selected binding species

To amplify selected binding species via combined RT-PCR reactions

1a. Resuspend the RNA in 12 μ l TE buffer or RNase-free water and add 4 μ l of this RNA suspension to 96 μ l RT-PCR mix.

Since only 1/3 of the total sample recovered is used for amplification this will obviously restrict the proportion of successful species that are carried into the next round of selection. This is only a potential problem in the early rounds of selection. For example, if the diversity of the RNA pool was such that each species was represented only a few times on average, then a population bottleneck is unavoidable. For this reason, it is always desirable to start with an RNA pool in which each species is represented numerous times. However, the amount of sample that is committed to amplification should probably not exceed one-half to two-thirds of the sample. If the reverse transcription or any subsequent steps are unsuccessful, then the archived RNA serves as an inviolate reservoir for proceeding forward in the selection experiment. Otherwise, one will have to return to material from an earlier round.

2a. Run the following controls in parallel with the amplification of selected RNA species in order to detect nonspecifically bound RNA species and replication parasites (see Critical Parameters for discussion of parasites).

a. *No template control:* To ensure that none of the stock solutions have been contaminated with exogenous RNA or DNA amplicons, set up a RT-PCR reaction without adding any template.

b. *No RT control:* To ensure that amplified products are in fact derived from selected RNA species and not from endogenous or cross-contaminating DNA molecules, set up a RT-PCR reaction without the reverse transcriptase.

3a. Run the RT-PCR reaction on the thermal cycler as follows:

- a. 10 min at 65°C
- b. 10 min at 50°C
- c. 45 sec at 94°C
- d. 60 sec at 50°C
- e. 90 sec at 72°C.

Repeat steps c to e six more times, then follow with:

- f. 150 sec at 72°C.

Step a allows for primers to anneal to the RNA, while step b allows the RT to generate cDNA. Steps c through e comprise the PCR cycle, and the final elongation step (f) at 72°C completes the extension of any incomplete DNA templates. The number of cycles that should initially be carried out is considered below.

It should be noted that the listed conditions have been optimized for the pool used in the selections described, the N30 pool. Different pools and primers may require very

different amplification conditions (see UNIT 15.4 for comments on primer selection and for the experimental parameters that govern reverse transcription and PCR).

If using a PCR machine without a heated bonnet, cover the amplification reaction with mineral oil (e.g., Mallinckrodt). In order to avoid the accumulation of replication parasites, it may be desirable to use one of a number of “hot-start” methods for the PCR reaction (see Chapter 15). The technically simplest of these is to add reverse transcriptase following heating to 65°C in step a, and to add Taq or another thermostable polymerase following heating to 94°C in step c.

To amplify selected binding species via separate RT and PCR reactions

1b. Resuspend the RNA in 12 µl TE buffer or RNase-free water, and set up the RT reaction as follows:

- 4 µl RNA, diluted as described above
- 2 µl 10× RT buffer
- 10 µl 20 µM 3'-end primer
- 4 µl 4 mM dNTP mix
- 0.3 µl (10 U) AMV reverse transcriptase.

Also set up “no template” control without the RNA template and “no RT” control without the reverse transcriptase.

2b. Incubate reactions 30 min at 42°C.

3b. Add 10 µl of each RT reactions to an individual tube containing 100 µl PCR mix.

Conduct the PCR reaction as follows:

- a. 45 sec at 94°C
- b. 60 sec at 50°C
- c. 90 sec at 72°C.

Repeat steps a and b six more times, then follow with:

- d. 150 sec at 72°C.

See UNIT15.1 for additional information on PCR amplification.

Check for the presence of amplified, double-stranded DNA

4. Add 1.5 µl of 6× non-denaturing dye to 5 to 10 µl of the PCR reaction. Load the sample onto a 4% NuSieve agarose gel which has been presoaked in 10 µg/ml ethidium bromide solution for 10 min (e.g., UNIT 2.6). Run the gel in TBE at 125 V for 15 min. Look for products with a hand-held UV lamp or UV light box.

An estimate of the minimal number of cycles needed to visualize a product band on the agarose gel can be roughly calculated. Consider that, of the 5 µg of RNA added to the selection, ~3% likely binds to the filter and is lost during the negative selection step. Approximately 1% of the population may bind to the target. When the selected RNA is precipitated, one-third of the sample is used for RT-PCR. Therefore:

$$(5.0 \mu\text{g})(0.97) (0.01)/3 = 0.016 \mu\text{g RNA.}$$

Assuming that every thermal cycle doubles the amount of DNA, a minimum of seven thermal cycles would be necessary to obtain 1 to 2 µg of DNA. This would imply that 0.1 to 0.2 µg could be loaded and readily visualized on the ethidium bromide-stained agarose gel. Thus, from 7 to 8 thermal cycles should initially be carried out and the products analyzed by gel electrophoresis. The authors frequently find this rough estimate to be true.

5. If no product bands are apparent, then carry out an additional 4 to 5 thermal cycles and again analyze the products by gel electrophoresis. If only faint product bands are

apparent, then one may want to accumulate additional template via an additional 2 to 4 thermal cycles.

The accumulation of double-stranded DNA is closely monitored in order to avoid “over-PCR” of the sample and the concomitant accumulation of high-molecular-weight species. DNA that has been over-amplified will look blurry and disperse following analysis by gel electrophoresis. These large DNA molecules are often the result of the 3' end of a single-stranded DNA folding back and internally priming its own extension, resulting in a long stem-loop that can be amplified by a single PCR primer (also known as single-primer artifacts). Overamplified DNA templates can also yield RNA molecules of the incorrect size following transcription. Adding 2 μ l of the RT-PCR reaction to 100 μ l of a fresh PCR mix and carrying out 2 to 3 additional thermal cycles can clean up DNA that has been over-amplified.

If one primer is more abundant or efficient than the other is, a smaller, single-stranded DNA band or bands may also be present.

The hiatus between carrying out the amplification reaction and running the agarose gel allows the sample to cool to room temperature, and can potentially result in mispriming and the accumulation of replication parasites. However, this is unavoidable and is not as serious for samples that have been partially amplified as it is for samples that are just beginning the amplification procedure. To avoid this potential problem, it is sometimes desirable to take one-fourth to one-third of the selected RNA and carry out a “ranging” RT-PCR reaction to establish the optimal number of cycles for amplification. Another one-fourth to one-third of the selected RNA can then be continuously amplified to this optimum level.

The various controls (“no protein,” “no template,” “no RT”) should be amplified in parallel with the actual sample. If specifically bound RNA is templating the accumulating amplicons, then the “No RT” sample should lag the RT-PCR reaction by at least three cycles. It is devoutly hoped that no bands will be observed in the “no template” control, but if they do arise, they should lag the RT-PCR reaction by at least five cycles. If bands do arise, a distinction should be made between full-length PCR products (indicating contaminating replicons) and smaller products (likely primer amplification artifacts). If product bands in the control lanes are as prominent as product bands in the experimental lanes, then it is necessary to check or remake reagents and go back and repeat the previous round of selection. There is one exception to this rule: in the initial rounds, it is common to see a band in the “no protein” control lane because the proportion of the population that binds to the filter is typically greater than the proportion that binds specifically to the target. However, subsequent rounds of selection should result in the diminution or disappearance of the “no protein” band.

Although we will consider methods for closely monitoring the progress of the selection experiment, observing the number of thermal cycles needed to visualize a double-stranded DNA band can loosely monitor the progress of the selection. The number of thermal cycles should be roughly proportional to the amount of RNA pool that originally binds to the protein. Therefore, if the RNA eluted from the “no protein” control requires more thermal cycles for full amplification than does the RNA selected in the presence of protein, it can be tentatively assumed that the selected RNA is binding to the protein. Occasionally, in the early rounds of selection, this may not be true, since a very small fraction of the pool will bind to the protein relative to the small fraction of the pool that adheres to the filter.

Counting PCR cycles is, however, only a very rough (and frequently inconsistent) measure of success. In fact, it is common for the number of thermal cycles required to fully amplify selected nucleic acids to vary greatly between rounds. Direct binding assays of the RNA pool (Support Protocol 3) are a much more accurate and useful gauge of the progress of a selection experiment.

6. When a product band does appear, precipitate the PCR reaction by adding an equal volume of 4 M ammonium acetate, and, to the resulting mixture, an equal volume of ethanol (i.e., 2 \times the original PCR reaction volume).

If a large amount of sample has been used for gel analysis (for example, if only ~50 μ l of the original RT-PCR reaction remains), then one may wish to return to the selected RNA reservoir and amplify a new DNA template using the already determined “optimal” number of thermal cycles.

Use amplified DNA template for the next round of selection

7. Centrifuge the sample and resuspend in 10 to 20 μ l TE buffer. Proceed with the next round of selection starting with step 1 of Basic Protocol 1.

A 100- μ l RT-PCR reaction yields ~1 to 2 μ g DNA, so approximately half of the resuspended DNA sample should be used for the next transcription reaction. The remaining DNA can serve as a long-term, archival sample.

ASSAYING THE ACCUMULATION OF BINDING SPECIES

In order to verify that the RNA pool has been or is being winnowed to those few sequences that bind the protein target with high affinity and specificity, the selected RNA pool should periodically be assayed for its ability to bind the target protein. The authors recommend an initial binding assay after five rounds of selection and amplification, then again every three additional rounds (the same recommendation that was made with regard to checking for filter-binding species; the two tests can be carried out in parallel). While the initial binding assay is carried out at a series of protein concentrations to gauge the amount of protein that should be used in the selection, the progress of the selection can be most simply monitored by radiolabeling the RNA and determining how much binds to a single, convenient concentration of the protein target.

Materials

Pool of dsDNA after n rounds of selection

Binding buffer

Target protein

13-mm, 0.45- μ m HAWP nitrocellulose disk filters (Millipore)

Filter holders (Nuclepore)

Vacuum manifold

Glass plate

Plastic wrap

Phosphoimager and screen or X-ray film

Additional reagents and equipment for purifying a radiolabeled DNA pool (see Basic Protocol 1) and performing the filter binding assay (see Support Protocol 2)

1. Generate radiolabeled RNA pool via a “hot transcription” with α -labeled nucleoside triphosphates and purify as described in the final annotation to step 1 of Basic Protocol 1.
2. Thermally equilibrate 1 μ g of the radiolabeled RNA pool after a round of selection in binding buffer as described in Support Protocol 2, step 2.
3. Add an equimolar amount of protein to the RNA pool. Incubate the binding reaction under conditions similar to those used for selection (see Support Protocol 2, steps 3 and 4).

If the amount of protein sample is limited or limiting, less protein can be used in the binding reaction. However, one should be cognizant of the fact that less than 100% binding is possible. Alternatively, less protein and less RNA sample can be used, although the diminution of both components will mean that one is assaying binding under conditions more stringent than those actually used for selection. While the volume of the binding reaction could also be diminished to conserve protein, it is difficult to uniformly apply volumes less than 30 μ l to the filter.

SUPPORT PROTOCOL 3

**Generation and
Use of
Combinatorial
Libraries**

24.3.17

4. Prior to filtration, take a small aliquot of the binding reaction (i.e., 5 μ l out of a 100 μ l binding reaction) to determine the total amount of radioactive RNA in the binding reaction. Pipet the sample onto a nitrocellulose filter disk and set the disk aside on a glass plate.
5. Filter the binding reaction and wash 3 times with 200 μ l binding buffer (see Support Protocol 2, steps 5 and 6).
6. Place the filters on the glass plate next to the initial aliquots of the binding reaction.
7. Cover the membranes with plastic wrap and expose to a phosphor screen (e.g., Phosphor imager or X-ray film for 4 to 12 hr).
8. Count the radioactivity using the Phosphorimager or a densitometer and calculate the fraction bound as follows:

Fraction bound = (cpm of filtered solution)/[(cpm of aliquot from step 4) \times (vol. of filtered solution/vol. of aliquot from step 4)].

A good result at this point would be 0.15 to 0.20 fraction bound above background (see Table 24.3.1., round 6). If binding to filter alone is too high, then filter binders are being selected and more negative selection is needed.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Denaturing dye, 2 \times

TBE buffer (APPENDIX 2) containing:

0.1% (w/v) bromophenol blue

7 M urea

Store up to 6 months at -20°C

Denaturing polyacrylamide gel, 10%

TBE buffer (APPENDIX 2) containing:

10% (w/v) acrylamide

0.5% (w/v) bisacrylamide

7 M urea

See UNIT 2.12 for full details on pouring and running the gel.

Elution buffer

7 M urea

100 mM sodium citrate

3 mM EDTA

Store up to 3 months at -20°C

Prepare with RNase-free water.

PCR mix

10 mM Tris-Cl, pH 8.4 (APPENDIX 2)

50 mM KCl

1.5 mM MgCl_2

0.2 mM each dNTP

5% (w/v) acetamide

0.05% (v/v) Nonidet P-40 (NP-40)

0.5 μ M each primer

0.2 U *Taq* DNA polymerase (Promega)

RT buffer, 10×

500 mM Tris·Cl, pH 8.0 (APPENDIX 2)
400 mM KCl
60 mM MgCl₂
Store up to 6 months at −20°C

RT-PCR mix

10 mM Tris·Cl, pH 8.4 (APPENDIX 2)
50 mM KCl
1.5 mM MgCl₂
0.2 mM each dNTP
5% (w/v) acetamide
0.05% (v/v) Nonidet P-40 (NP-40)
0.5 μM each primer
0.2 U *Taq* polymerase (Promega)
10 U AMV reverse transcriptase (USB)

Transcription mix

40 mM Tris·Cl, pH 7.9 (APPENDIX 2)
26 mM MgCl₂
0.01% (v/v) Triton X-100
2.5 mM spermidine trihydrochloride
5 mM dithiothreitol
2.5 mM each ribonucleotide triphosphate
20 U RNasin (Promega)
100 U T7 RNA polymerase (New England Biolabs)
Prepare fresh

COMMENTARY**Background Information**

Sol Spiegelman and coworkers developed a working system for the in vitro replication and evolution of small RNA molecules over 25 years ago (Mills et al., 1967; Levisohn and Spiegelman, 1969; Kramer et al., 1974). The development of more advanced (although conceptually identical) methods for in vitro evolution as described in this chapter was potentiated by advances in the chemical synthesis of oligonucleotides and the amplification of nucleic acids, such as PCR, in vitro transcription, and 3SR (Guatelli et al., 1990). The adaptation of these methods to in vitro evolution of RNA molecules was partially due to a recognition that early evolutionary events, such as the genesis of ribozymes, could be recapitulated in a test tube, and partially due to a recognition that the ability to tailor RNA binding species and catalysts might have numerous biotechnological applications. Following the publication of key papers outlining and proving selection technologies (Ellington and Szostak, 1990; Tuerk and Gold, 1990), a much wider array of selection experiments has been attempted. To date,

RNA molecules that can bind targets as small as zinc and as large as viruses and organs have been selected. RNA molecules that interact with both nucleic acid-binding proteins and non-nucleic acid binding proteins can be selected with almost equal facility from random sequence populations. These results have been thoroughly reviewed in numerous recent publications (Gold et al., 1995; Uphoff et al., 1996; Famulok and Jenne, 1998).

Critical Parameters**Choosing protein targets**

As briefly described above, a wide variety of proteins have proven to be successful targets for selection experiments, including enzymes, transcription factors, cytokines, antibodies, and viral capsids (Gold et al., 1995; Uphoff et al., 1996; Famulok and Jenna, 1998). There is no common functional theme uniting these targets, nor can many generalities be drawn regarding their biochemistry or structure. However, it is safe to say that “good” selection targets tend to fall into two classes. First, proteins that nor-

mally bind nucleic acids will also be able to extract aptamers from a random sequence pool. The notion of a nucleic acid-binding protein can to some extent be expanded to include proteins that bind nucleotides. For example, kinases and dehydrogenases bind nucleotide cofactors and have proven to be good selection targets.

Second, proteins that for whatever reason contain basic patches in their primary sequences or on their surfaces also frequently yield high-affinity aptamers. For example, many cytokines and other signal-transduction proteins bind heparin or other sulfated oligosaccharides, and can also be used to select aptamers from random sequence populations. The anti-cytokine aptamers frequently bind to the same sites as heparin (Jellinek et al., 1993). Similarly, proteins that bind phosphate or phosphomonoester or phosphodiester bonds frequently have positively charged active sites and can be used to elicit aptamers. For example, anti-phosphatase aptamers have been selected from random sequence pools (Bell et al., 1998).

This is not to say that proteins that do not fall into these categories will of necessity be poor selection targets, merely that they are not sure selection targets. For example, antibodies have frequently proven to be excellent selection targets irrespective of whether or not they bind negatively charged antigens (Keene, 1996). This likely implies that proteins that have large pockets or clefts on their surface are good selection targets. This hypothesis is further bolstered by another line of reasoning. Aptamers selected to bind proteins frequently inhibit protein function. That is, anti-antibody aptamers block interactions with antigens, anti-enzyme aptamers inhibit enzymatic activities, and so forth. This so-called “homing principle” may be due to the fact that aptamers have to not only form a surface that is chemically complementary to a target, but also must fold into a structure that properly presents the chemically complementary surface. The most informationally parsimonious way to achieve both functions is to fit into a pocket on a target, rather than to form a “grasping” structure that can enfold a surface protrusion of a target. Thus, the most common (and most highly represented) aptamers may be those that fit into surface crevices. In contrast, antibodies have a preformed structure for the presentation of chemically complementary surfaces, and thus can more easily grasp protruding epitopes and less easily fit into surface crevices.

Overall, though, researchers should be guided not so much by these considerations, but by the results of initial binding assays with their particular protein target. If the target binds to the filter (not a given, since small, acidic proteins such as the Rop protein from *E. coli* will frequently pass through the filter) and shows some affinity for a random sequence pool, then it is highly probable that there will be some sequences or structures within the pool with greatly enhanced affinities for the target.

Choosing a binding buffer

The binding buffer should promote specific binding of nucleic acids to a protein target. The first consideration in choosing a buffer is to identify conditions under which the protein is active, or at least stable. In addition, if the selected nucleic acid species are to eventually be used in a particular environment, the selection buffer should reflect this environment. For example, if the selected nucleic acids are to be expressed in a cell, then the selection buffer should be at physiological pH and contain physiological ion concentrations. Second, there are a variety of parameters that can be used to make the RNA pool more or less “sticky.” These parameters are discussed in much greater detail in the following section on the stringency of selection.

A typical binding reaction is built from one of the commonly used buffers, such as Tris-Cl, phosphate, or HEPES, which can hold the pH near 6 to 8, together with 50 to 200 mM NaCl or KCl and 1 to 10 mM MgCl₂. However, these are merely suggestions, and aptamers have in fact been selected under a variety of buffer conditions. For example, in the selection that targeted bFGF, phosphate-buffered saline was used even though it lacked divalent cations. Similarly, ribozyme selections have been carried out in which a variety of divalent metal ions are mixed, and nascent ribozyme species “decide” which combination of metals most enhance their activities (Lehman and Joyce, 1993). An equivalent strategy could be used for the selection of aptamers.

Selection matrices

Due to the tremendous ratio of matrix surface area to protein surface area, matrix-binding aptamers can quickly and easily eclipse target-binding aptamers. Proteins are likely captured on nitrocellulose or modified cellulose filters via hydrophobic interactions. Nucleic acids are, by and large, too hydrophilic or charged to be similarly captured. This distinc-

tion is the basis for most filter-binding assays. However, the nucleobases of nucleic acids obviously contain large hydrophobic surface areas, and it is easy to select nucleic acids that can present nucleobases and be captured by the filter. Selected filter-binding sequences frequently contain purine (especially guanosine) tracts presented as single-stranded loops or bulges. Interestingly, hydrophobic-binding sequences selected on one hydrophobic matrix are frequently cross-reactive with other hydrophobic matrices: i.e., microtiter plate-binding species can bind tubes and filters, filter-binding species can bind tubes and microtiter plates, and so forth.

In order to avoid filter-binding sequences, the authors have filtered RNA samples multiple times in the absence of protein, and in some cases filtered samples following selection but prior to the RT-PCR step. Matrix-binding sequences can also be avoided by altering the matrices used for selection. For example, techniques such as gel mobility shifts, immunoprecipitation, and affinity chromatography have all been successfully used to sieve pools and select target-binding aptamers (Conrad et al., 1996). If filter-binding species predominate in a population even after appropriate precautions are taken, these alternative selection techniques can be used either to rid the selected population of the filter-binding species or, better yet, to restart the selection. For example, if the immunoprecipitation of RNA:protein complexes has been worked out in advance, then immunoprecipitation can be interspersed with rounds of filter-binding.

Even though the selection of filter-binding sequences can be a problem, filter binding is still generally recommended as the technique of choice for most selections. Gel mobility shift experiments tend to be much more sensitive to parameters such as sample preparation, ionic strength, pH, and electrophoresis conditions than are filter-binding experiments. Moreover, just as filter-binding species can be inadvertently selected during filtration selection, RNA species with altered electrophoretic mobilities (e.g., dimers) can be selected during gel-mobility shift selections. Immunoprecipitation experiments require an additional protein reagent and in consequence anti-antibody rather than anti-target aptamers are frequently selected. Affinity chromatography or similar techniques generally require that very large amounts of target proteins be committed to the preparation of affinity matrices. If affinity elution is to be used, then even larger amounts of target pro-

teins will be required. Moreover, aptamers that bind to agarose matrices can be selected almost as easily as aptamers that bind to nitrocellulose or modified cellulose filters (although the two, thankfully, do not cross-bind to one another's matrices). Finally, microtiter plate panning selections encourage the accumulation of the same sorts of matrix-binding aptamers that are elicited by filter-binding selections.

Stringency of selection

Overall, most selection experiments are generally competitions between specifically and nonspecifically binding nucleic acid species. The authors tend to initially choose conservative binding conditions in hopes of promoting the early establishment of binding species in the population. While this may mean that low-affinity species are isolated from the pool along with high-affinity species, the low-affinity species can eventually be removed by increasing the stringency of selection. In essence, time (the number of cycles required to purify high-affinity species) can be traded for the assurance that filter-binding species will not accumulate and predominate.

A variety of parameters can be modulated in order to increase or decrease the stringency of a selection experiment. These parameters should initially be chosen based on the results of Support Protocol 2, which assays the affinity of the pool for the target and should be made progressively more stringent based on the results of Support Protocol 3.

The amount of protein target. The more protein there is to bind, the easier it is to capture nucleic acid binding species. Using low amounts of protein increases competition among binding species. However, the amount of protein target available to researchers is usually limited, and thus it is easier to use a set amount of protein (usually from 0.1 to 10.0 μ m per binding reaction) and to vary the RNA:protein ratio.

RNA:protein ratio. By increasing the ratio of pool to target, more binding species will compete for a smaller number of targets. Typically, after a few initial rounds with an equimolar pool-to-target ratio, the ratio is increased to between 10:1 and 100:1. This increase can be effected either by increasing the amount of RNA or by decreasing the amount of protein. Because of the underlying competition between specifically binding species and nonspecifically binding species, increasing the amount of RNA is preferable to decreasing the amount of protein. For a more detailed treatment of this

subject, see Irvine et al. (1991). However, the general conclusions of these mathematical models are similar to the empirical advice given here.

Competitors. High concentrations of non-specific, non-amplifiable competitors such as tRNA or bulk cellular RNA will compete with low-affinity binding species that adhere to basic patches on the surface of a protein. Typically, a 100-fold excess of tRNA is used. Similarly, specific competitors can be used to block the access of low-affinity binding species to a preferred site. Wild-type nucleic acid ligands can be used to block the binding sites of nucleic acid binding proteins. For example, during the selection of anti-Rev aptamers, Giver et al. (1993) included a 10-fold excess of the wild-type Rev-binding element. The anti-Rev aptamers that were obtained could bind with high affinity to the RNA-binding domain of Rev and could effectively compete with the wild-type Rev-binding element. Other ligands or substrates can also be used to block the binding or catalytic sites of non-nucleic acid binding proteins. For example, during the selection of anti-bFGF aptamers, Jelinek et al. (1993) included heparin, a natural ligand for bFGF. The anti-bFGF aptamers that were obtained could bind with high affinity to the heparin binding site and could effectively compete with heparin.

Cation concentration. Monovalent cations (such as Na⁺) and divalent cations such as Mg²⁺ stabilize the structure of RNA molecules and contribute to both specific and nonspecific binding. Decreasing monovalent and/or divalent cation concentrations therefore can therefore increase the stringency of the selection. However, it is unclear, in advance, whether specific or nonspecific binding species will be more favored by such a change. Moreover, since binding species that require a monovalent and/or divalent cation to fold into shapes that are chemically complementary to a target may be favored in the early rounds of selection, potentially high-affinity binding species may be lost by changing the binding buffer late in the selection experiment. It is better to attempt to change the buffer dependency of aptamers by partial randomization and re-selection following the initial selection experiment, rather than to attempt to change the buffer dependency during the selection.

Conversely, higher concentrations of monovalent cations (generally sodium or potassium) increase the structural integrity of folded nucleic acids by neutralizing the close approach of nucleic acid strands. However, higher mono-

valent ion concentrations also suppress electrostatic interactions with targets. Thus, paradoxically, both “low” and “high” monovalent ion concentrations can be used to increase the stringency of a selection experiments. Higher concentrations of divalent cations such as magnesium help to maintain the structural integrity of RNA molecules and potentially facilitate the formation of salt bridges between acidic residues and the phosphate backbone.

Equilibration time. Longer equilibration times give stronger binding species a greater chance to bind to the target, since weaker binding species more quickly dissociate from the target. In general, though, species with nanomolar dissociation constants or lower can be readily selected by allowing the reaction to equilibrate for 5 min or more. The authors usually allow up to 30 min for the binding reaction in order to allow for slow folding or refolding steps in the presence of the target. However, longer equilibration times may not be possible for proteins that are inherently unstable or that themselves undergo slow, buffer- or temperature-induced conformational changes.

Dilution of binding buffer. Similarly, diluting the binding reaction by 10- to 20-fold just prior to filtration will favor the selection of RNA:protein complexes with low dissociation constants over RNA:protein complexes with higher dissociation constants. Baskerville et al. (1995) have successfully used this technique to select high affinity anti-Rex aptamers.

Amount and composition of wash. Increasing the number of times a filter is washed and the volume of the buffer used for the washes should preferentially increase the retention of high-affinity binding species relative to low-affinity and nonspecific binding species. It is generally recommended that the same buffer be used for selection and for wash steps, in order to avoid changing the conditions under which aptamers are selected. However, the stringency of the selection can potentially be manipulated by changing the buffer used for the wash steps. For example, if monovalent cation concentrations are limited in the binding buffer due to requirements for the stability or activity of a protein target, a separate wash buffer that contains a higher salt concentration can be used to challenge captured RNA:protein complexes.

Parasites

Replication parasites differ from matrix-binding aptamers, but can interfere with the selection of target-binding aptamers in the

same way. Reverse transcriptase, *Taq* polymerase, and T7 RNA polymerase all have some preference for which sequences they will copy or reproduce. These preferences are generally not obvious when constant sequence nucleic acids are being synthesized. However, in selection experiments many cycles of amplification are carried out, and differences in the rates of synthesis are also proportionately amplified, leading to the selection of sequences that have no function other than to replicate optimally. For example, during the polymerase chain reaction if a primer designed to bind to a constant sequence region instead recognizes a partially complementary sequence within a random sequence region, it can set down and generate a smaller amplicon. The smaller amplicon will generally be amplified more quickly than the larger amplicon, and thus can potentially out-compete full-length species selected for binding function. Depending on the relative advantage of the replication parasite relative to an aptamer, even if the replication parasite is partially removed from the population during each selection step, enough molecules may remain to overrun the amplification reaction and displace the functionally selected aptamer. This is especially true if the amplification parasite also happens to be a filter-binding species. It is for this reason that the authors of this unit strongly recommend that DNA templates and/or RNA molecules be size-selected in each round.

The nascent reproductive differences between nucleic acid species can be grossly amplified by amplification methods that allow continuous reproduction of the nucleic acids, such as isothermal amplification or 3SR (Guatelli et al., 1990). For example, Breaker and Joyce (1994) generated an extremely robust replication parasite, RNA Z, during a selection designed to generate catalytic variants of a group II intron. Similarly, the authors have generated replication parasites of isothermal amplification reactions from completely random sequence pools (K. Marshall, pers. comm.). Interestingly, these isothermal amplification parasites were actually larger than the initial RNA species and represented recombination events between individual members of the pool. Airborne copies of these replication parasites can readily “seed” isothermal amplification reactions and overrun pool molecules that are initially present in even million-fold excess. In this respect, the replication parasites of isothermal amplification reactions resemble the midi-variants or “monsters” of Q β replicase amplification reactions, and are equally hard to

vanquish, once established. It is for this reason that the authors strongly recommend the sometimes tedious but inherently faithful regime of reverse transcription, PCR, and in vitro transcription for the amplification of RNA pools. However, successful selections have been carried out that have relied upon isothermal amplification (see, for example, Breaker et al., 1994; Wlotzka and McCaskill, 1997; Wright and Joyce, 1997), and this admonition can most confidently be challenged if the starting pool is a partially randomized binding site or ribozyme. The reason is that isothermal amplification parasites are more likely to be found in or derived from a “deep random” pool than in a pool that centers on a given functional sequence.

Anticipated Results

Table 24.3.1 shows the progression of a selection carried out in the authors’ lab against bFGF with an RNA pool with a 30 nucleotide randomized region. In order to evaluate the success of a selection experiment, it was necessary to compare the affinity of the selected pool versus the affinity of the unselected pool for the protein target (see Support Protocol 3). When assaying the pool after a round of selection, it was necessary to validate the fraction of the pool that bound to the protein by including a no protein control. If the accumulation of matrix-binding species had been evident, more stringent negative selections could have potentially been used to control or reduce their numbers.

The affinity of the RNA aptamer for the protein target cannot be anticipated. Affinity typically varies between micromolar and subnanomolar, depending presumably on the makeup of the nucleotide pool and on the targeted protein. However, it might be worth mentioning that, of the first 100 selections carried out at two commercial entities using the technology—Gilead Sciences and NeXstar—just under 80% yield aptamers with affinities under 10^{-9} M (Brody et al., 1999).

Time Considerations

The time required to go from one pool of selected DNA templates to the next is ~24 to 72 hr, depending on the researcher and the demands of the particular selection experiment. Minimally, a transcription reaction takes ~4 hr, and the ensuing DNase, heat denaturation and gel purification steps can take another 2 to 3 hr. Elution for 8 to 10 hr yields an adequate amount of RNA to be used in the subsequent binding

reaction. After precipitation and quantification of the RNA (1 hr), the preselection filtration, incubation with target, and selection steps can be performed in 2 hr. Elution of protein-RNA complexes, subsequent extractions, and another precipitation step take another 2 hr. The amount of time needed to see a DNA product varies according to the number of PCR cycles needed to amplify the pool to a certain amount, and that number is inversely related to the abundance of target-binding species that survived the selection. Nevertheless, the RT-PCR steps, followed by precipitation of the DNA templates that can be added to the transcription mix, should consume ~3 to 4 hr.

The amount of time it takes to carry out the entire selection is contingent upon the number of rounds needed to accumulate target-binding species. That number, in turn, varies depending upon the initial affinity of the unselected pool for the target and on the stringency with which each round of the selection is carried out. When additional steps such as radiolabeling and assaying unselected and selected pools are taken into account, an entire selection experiment can take up to 2 to 3 weeks. It is for this reason that the authors have recently developed automated methods for selection experiments (Cox et al., 1998) that can speed the entire process by an order of magnitude.

Literature Cited

- Baskerville, S., Zapp, M., and Ellington, A.D. 1995. High resolution mapping of the human T-cell, leukemia virus type 1 rex-binding element by in vitro selection. *J. Virol.* 69:7559-7569.
- Bell, S.D., Denu, J., Dixon, J.E., and Ellington, A.D. 1998. RNA molecules that bind to and inhibit the active site of a tyrosine phosphatase. *J. Biol. Chem.* 273:14309-14314.
- Breaker, R. and Joyce, G.F. 1994. Emergence of a replicating species from an in vitro RNA evolution reaction. *Proc. Natl. Acad. Sci. U.S.A.* 91:6093-6097.
- Breaker, R., Banerji, A., and Joyce, G.F. 1994. Continuous in vitro evolution of bacteriophage RNA polymerase promoters. *Biochemistry* 33:11980-11986.
- Brody, E.N., Willis, M.C., Smith, J.D., Jayasena, S., Zichi, D., and Gold, L. 1999. The use of aptamers in large arrays for molecular diagnostics. *Mol. Diagnon.* 4:381-388.
- Conrad, R.C., Giver, L., Tian, Y., and Ellington, A.D. 1996. In vitro selection of nucleic acid aptamers that bind proteins. *Methods Enzymol.* 267:336-367.
- Cox, J.C., Rudolph, P., and Ellington, A.D. 1998. Automated DNA selection. *Biotechnol. Prog.* 14:845-850.
- Ellington, A.D. and Szostak, J.W. 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature* 346:818-822.
- Famulok, M. and Jenne, A. 1998. Oligonucleotide libraries—variatio delectat. *Curr. Opin. Chem. Biol.* 2:320-327.
- Giver, L., Bartel, D., Zapp, M., Green, M., and Ellington, A.D. 1993. Selective optimization of the Rev-binding element of HIV-1. *Nucl. Acids Res.* 23:5509-5516.
- Gold, L., Polisky, B., Uhlenbeck, O., and Yarus, M. 1995. Diversity of oligonucleotide functions. *Annu. Rev. Biochem.* 64:763-797.
- Guatelli, J., Whitfield, K., Kwoh, D., Barringer, K.J., Richman, D., and Gingeras, T.R. 1990. Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. *Proc. Natl. Acad. Sci. U.S.A.* 87:1874-1878.
- Irvine, D., Tuerk, C., and Gold, L. 1991. SELEXION: Systematic evolution of ligands by exponential enrichment with integrated optimization by non-linear analysis. *J. Mol. Biol.* 222:739-761.
- Jellinek, D., Lynott, C., Riata, D., and Janjic, N. 1993. High affinity RNA ligands to basic fibroblast growth factor inhibit receptor binding. *Proc. Natl. Acad. Sci. U.S.A.* 90:11227-11231.
- Keene, J.D. 1996. RNA surfaces as mimetics of proteins. *Chem. Biol.* 3:505-513.
- Kramer, F.R., Mills, D.R., Cole, P.E., Nishihara, T., and Spiegelman, S. 1974. Evolution of in vitro sequence and phenotype of a mutant RNA resistant to ethidium bromide. *J. Mol. Biol.* 89:719-736.
- Lehman, N. and Joyce, G.F. 1993. Evolution in vitro of an RNA enzyme with altered metal dependence. *Nature* 361:182-185.
- Levisohn, R. and Spiegelman, S. 1969. Further extracellular Darwinian experiments with replicating RNA molecules: Diverse variants isolated under different selective conditions. *Proc. Natl. Acad. Sci. U.S.A.* 63:805-811.
- Mills, D.R., Peterson, R.L., and Spiegelman, S. 1967. An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc. Natl. Acad. Sci. U.S.A.* 58:217-224.
- Tuerk, C. and Gold, L. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249:505-510.
- Uphoff, K., Bell, S., and Ellington, A.D. 1996. In vitro selection of aptamers: The dearth of pure reason. *Curr. Opin. Struct. Biol.* 6:281-288.
- Wlotzka, B. and McCaskill, J.S. 1997. A molecular preator and its prey: Coupled isothermal amplification of nucleic acids. *Chem. Biol.* 4:25-33.
- Wright, C. and Joyce, G.F. 1997. Continuous in vitro evolution of catalytic function. *Science* 276:614-617.

Key References

Conrad et al., 1996. See above.

Conrad, R.C., Brück, F.M., Bell, S., and Ellington, A.D. 1998. In vitro selection of nucleic acid ligands. *In Nucleic Acid-Protein Interactions: A Practical Approach* (W.J. Christopher, ed.) pp. 285-315. Oxford University Press, New York.

The above two papers also provide protocols for the selection of aptamers via filter immobilization as well as by other means.

Contributed by Sulay D. Jhaveri
and Andrew D. Ellington
University of Texas
Austin, Texas

Peptide Aptamers: Dominant “Genetic” Agents for Forward and Reverse Analysis of Cellular Processes

Peptide aptamers are a new class of dominant “genetic” agents that facilitate the analysis of cellular processes in diploid and genetically intractable organisms. They are defined as protein-based recognition agents that consist of a constrained combinatorial peptide library displayed on the surface of a scaffold protein. Peptide aptamers function *in trans*, interacting with and inactivating gene products without mutating the DNA that encodes them. Combinatorial libraries of peptide aptamers contain aptamers that, in principle, can interact with almost any gene product.

The dominant combinatorial nature of peptide aptamers makes them useful as genetic agents for the reverse and forward analysis of cellular processes. Reverse analysis with peptide aptamers involves isolating aptamers that interact with a specific protein and monitoring the resulting aptamer-induced phenotype. A two-hybrid system is used to screen combinatorial libraries of peptide aptamers for those aptamers that interact with a specific protein. The isolated aptamers are then expressed within an organism to identify the aptamer-induced phenotype. Forward analysis with peptide aptamers involves expressing combinatorial libraries of aptamers within an organism and screening for aptamer-induced variations in their phenotypes. The specific protein(s) targeted by the aptamers are identified using a two-hybrid system.

This unit describes methods to construct and use thioredoxin peptide aptamers as genetic agents for the analysis of cellular processes. The interaction trap two-hybrid system (UNIT 20.1) is used to isolate peptide aptamers that interact with specific proteins (reverse analysis) and to identify the proteins targeted by aptamers (forward analysis).

Basic Protocol 1 describes the construction of a combinatorial library of thioredoxin peptide aptamers. The peptide aptamers consist of a conformationally constrained twenty–amino acid peptide displayed from the active site of thioredoxin. The peptide aptamers are subcloned into one of the pJM yeast expression vectors shown in Figure 24.4.1, depending on whether they are used for reverse or forward analysis.

Basic Protocol 2 describes a yeast-based *in vivo* screening method to obtain peptide aptamers for reverse analysis of cellular processes. Combinatorial libraries of peptide aptamers are screened for interactions with a specific protein using the interaction trap two-hybrid system (UNIT 20.1). The peptide aptamer is expressed as a fusion to a transcription activation domain, referred to as the “prey.” The target protein is expressed as a fusion to a LexA DNA binding domain, referred to as the “bait.” DNA-binding sites for the LexA fusion protein are located upstream of the two reporter genes, *Leu2* (CD8) and *lacZ*. Interaction between a peptide aptamer prey and the bait protein are detected by activation of these reporter genes.

Basic Protocol 3 describes the use of the yeast mating interaction assay to evaluate the specificity of peptide aptamers. Haploid yeast exist in two mating types (**a** or α), where opposite mating types can mate to form diploids (**a**/ α). The mating interaction assay detects aptamer/protein interactions by generating panels of aptamer preys in one mating type and panels of target bait proteins in the opposite mating type. Mating of the haploid strains forms diploid strains that carry both the bait and prey. Interactions between baits and preys are detected using the interaction trap reporters. The mating interaction assay

allows aptamer specificity to be assessed against large arrays of different but related proteins and against mutants of the same protein.

Basic Protocol 4 describes an affinity maturation strategy for enhancing the affinity of peptide aptamers to their target proteins. PCR mutagenesis is used to introduce random mutations into the variable region of a peptide aptamer. Peptide aptamers with enhanced affinity are isolated using a modified version of the interaction trap that contains a more

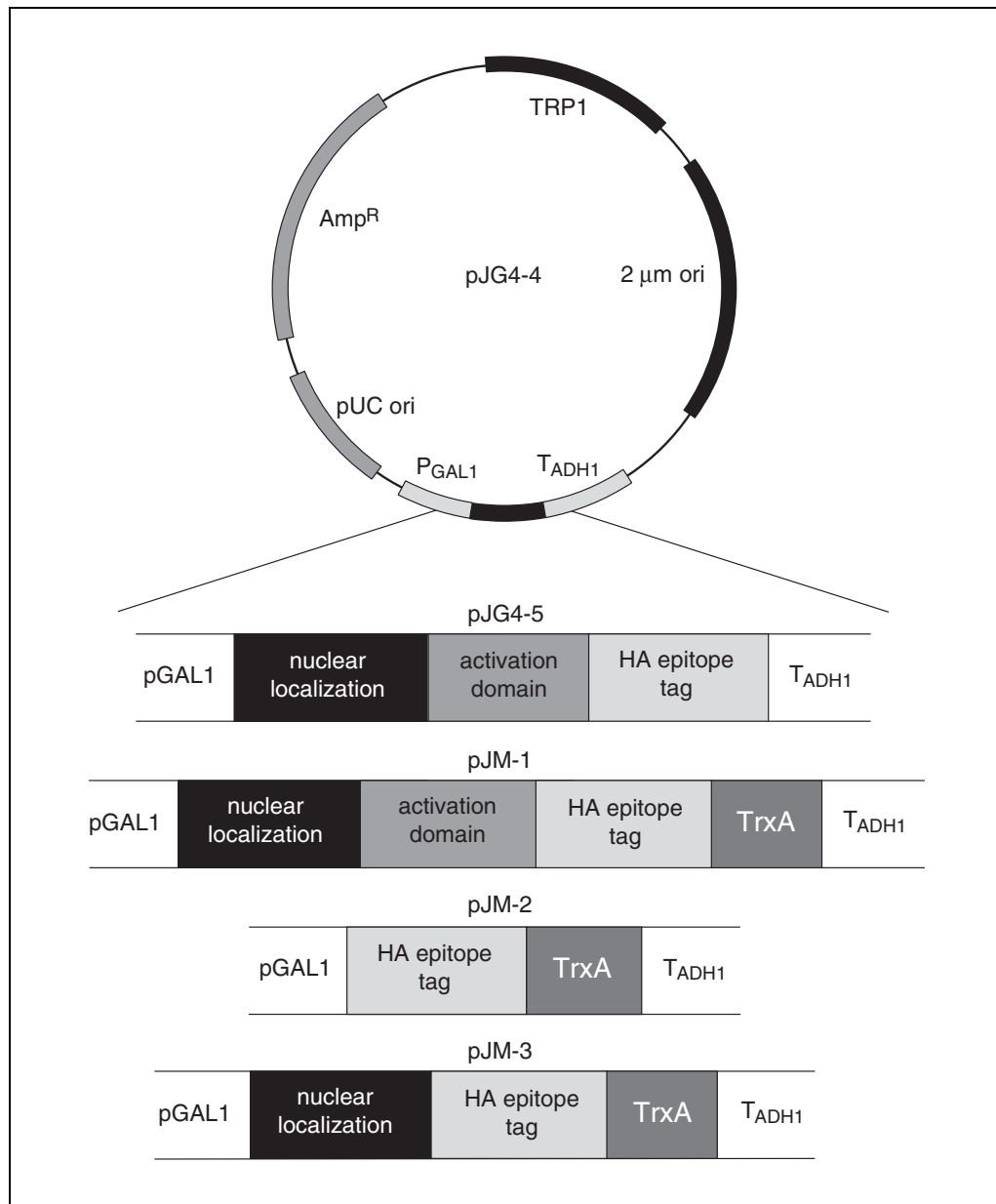


Figure 24.4.1 Expression vectors for interaction trap and genetic selection. pJG4-5 is the prey vector used in the interaction trap (*UNIT 20.1*). pJM-1 is the peptide aptamer prey vector. pJM-2 and pJM-3 are used in yeast genetic selections. These yeast-*E. coli* shuttle vectors are derivatives of pJG4-4 (Gyuris et al., 1993), and contain one of the following expression cassettes. pJG4-5: yeast GAL1 promoter (P_{GAL1}), SV40 nuclear localization signal, B42 activation domain, haemagglutinin epitope tag, *Eco*RI and *Xho*I cloning site, and yeast ADH1 transcription terminator (T_{ADH1}) (Gyuris et al., 1993). See Figure 20.1.3 for a more detailed map of pJG4-5. pJM-1: P_{GAL1}, SV40 nuclear localization signal, B42 activation domain, haemagglutinin epitope tag, *E. coli* thioredoxin (TrxA), and T_{ADH1} (Colas et al., 1996). pJM-2: P_{GAL1}, haemagglutinin epitope tag, TrxA, and T_{ADH1} (Geyer et al., 1999). pJM-3: P_{GAL1}, SV40 nuclear localization signal, haemagglutinin epitope tag, TrxA, and T_{ADH1} (Geyer et al., 1999).

stringent *lacZ* reporter. The stringency of the *lacZ* is increased by reducing the number of *LexA* operators upstream of the *lacZ* reporter gene.

Basic Protocol 5 describes a method to use peptide aptamers for the forward analysis of cellular processes. Combinatorial libraries of peptide aptamers are used as dominant genetic agents that randomly inhibit gene function. Forward analysis involves: (1) expressing combinatorial libraries of peptide aptamers in organisms, (2) isolating organisms that display aptamer-induced phenotypes, and (3) identifying peptide aptamer targets using the interaction trap.

CONSTRUCTION OF A COMBINATORIAL THIOREDOXIN PEPTIDE APTAMER LIBRARY

BASIC PROTOCOL 1

Combinatorial libraries of peptide aptamers are constructed by inserting a random twenty–amino acid peptide into the short disulfide-constrained loop (-CGPC-) in the active site of *E. coli* thioredoxin. The active site loop contains a unique *RsrII* restriction site that allows the insertion of *AvaII*-cut DNA, which encodes for random amino acids. Random peptide libraries are constructed using twenty repeats of the codon NNK, where N is A, G, C, or T and K is G or C. Using G or C in the third position of the codon reduces the number of stop codons while maintaining codons for all twenty amino acids. Depending on the application, the random peptide libraries are subcloned into one of the pJM yeast expression vectors shown in Figure 24.4.1. pJM-1 is used in the interaction trap to generate peptide aptamers against specific proteins. pJM-2 and pJM-3 are used in genetic selections to produce aptamers that alter an organism's phenotype. All of the pJM vectors use the *galI* promoter to control the expression of the peptide aptamers. The *galI* promoter induces aptamer expression in the presence of galactose and represses expression in the presence of glucose. The resulting aptamer/thioredoxin vector is transformed into *E. coli* by electroporation (also see *UNIT 9.3* for electroporation techniques).

Materials

- 5 U/μl Klenow DNA polymerase and 10× reaction buffer (New England Biolabs)
- 5 mM 4dNTP mixture: 5 mM each dTTP, dATP, dGTP, and dCTP
- 10 U/μl *AvaII* and 2 U/μl *RsrII* restriction enzymes and 10× reaction buffers (New England Biolabs)
- 10 mM Tris·Cl, pH 8 (*APPENDIX 2*)
- Nondenaturing loading buffer (see recipe)
- DNA elution buffer (see recipe)
- Thioredoxin expression vector plasmid: pJM-1, pJM-2, or pJM-3 (Fig. 24.4.1)
- 10 U/μl calf intestinal alkaline phosphatase (CIP) and 10× reaction buffer (New England Biolabs)
- 2000 U/μl T4 DNA ligase and 10× reaction buffer (New England Biolabs)
- QIAquick gel extraction kit (Qiagen)
- Ultrapure water (sterile water for irrigation preferred; Fisher Scientific)
- E. coli* MC 1061 (Bio-Rad), electroporation competent (*UNIT 9.3*)
- SOC medium (*UNIT 1.8*), prewarmed to 37°C
- LB plates and liquid medium (*UNIT 1.1*) containing 50 μg/ml ampicillin
- Large-scale plasmid preparation kit (various commercial sources, e.g., Qiagen; optional)
- DNA synthesizer
- 16° and 95°C water baths
- PCR purification column (e.g., Qiagen; optional)
- Electroporator (e.g., Bio-Rad Gene Pulser) with 0.2-cm-gap electroporation cells

Generation and Use of Combinatorial Libraries

24.4.3

Additional reagents and equipment for DNA synthesis; phenol/chloroform extraction and ethanol precipitation (*UNIT 2.1A*); polyacrylamide gel electrophoresis (*PAGE*; *UNIT 2.7*); UV shadowing and elution of DNA (*UNIT 2.7*); UV spectroscopy (*APPENDIX 3D*) or ethidium bromide dot quantitation (*UNIT 2.6*); bacterial transformation (*UNIT 1.8*); and ethidium bromide/cesium chloride gradients (optional; *UNIT 2.4*)

NOTE: Activity units of enzymes are described for enzymes obtained from New England Biolabs. Other commercial sources can be used, but units should be confirmed.

Prepare random peptide DNA cassette

1. Prepare the following 91-base random oligonucleotide and 17-base primer using an automated DNA synthesizer. Dissolve oligonucleotides separately in water to a final concentration of 1 $\mu\text{g}/\mu\text{l}$.

Oligonucleotide: 5'-GACTGACTGGTCCG(NNK)₂₀GGTCCTCAGTCAGTCAG-3', where N is A, G, C, or T and K is G or C.
Primer: 5'-CTGACTGACTGAGGACC-3'.
2. Add the following (in order) to a 1.5-ml microcentrifuge tube (final 890 μl):
 - 200 μg primer (10-fold excess)
 - 100 μg random oligonucleotide
 - 490 μl water
 - 100 μl 10 \times Klenow polymerase reaction buffer.
3. Anneal primer to random oligonucleotide by heating sample to 95°C in a water bath for 5 min. Slowly cool to room temperature (~30 min).
4. Add 90 μl of 5 mM 4dNTP mixture and 20 μl (100 U) Klenow polymerase and incubate 3 hr at 37°C.
5. Phenol/chloroform extract the mixture (*UNIT 2.1A*) and ethanol precipitate the DNA (*UNIT 2.1A*).
6. Dissolve DNA pellet in 0.8 ml water.
7. Add 100 μl of 10 \times *Ava*II reaction buffer and 100 μl (1000 U) *Ava*II. Incubate 4 hr at 37°C.
8. Repeat step 5.
9. Dissolve DNA pellet in 150 μl of 10 mM Tris-Cl, pH 8, and add 50 μl vol nondenaturing loading buffer.
10. Separate DNA on a preparative 10% nondenaturing polyacrylamide gel (*UNIT 2.7*).
11. Locate the DNA band in the gel by UV shadowing (*UNIT 2.7*) and cut out the DNA band.
12. Elute DNA from the gel by shaking in DNA elution buffer overnight (*UNIT 2.7*).
13. Ethanol precipitate the DNA and dissolve in 200 μl of 10 mM Tris-Cl, pH 8. Determine DNA concentration by UV spectroscopy (*APPENDIX 3D*), or estimate DNA concentration using ethidium bromide dot quantitation (*UNIT 2.6*).
14. Choose one of the thioredoxin expression vectors (pJM) in Figure 24.4.1 and add 12 μg of the chosen vector to 420 μl sterile water.

Prepare thioredoxin expression vector

14. Choose one of the thioredoxin expression vectors (pJM) in Figure 24.4.1 and add 12 μg of the chosen vector to 420 μl sterile water.

15. Add 50 μ l of 10 \times *Rsr*II reaction buffer and 30 μ l (60 U) *Rsr*II. Incubate overnight at 37°C.
16. Dephosphorylate *Rsr*II-cut pJM vector by adding 10 μ l (100 U) CIP and incubating 1 hr at 37°C.
17. Purify dephosphorylated, *Rsr*II-cut pJM vector using a commercially available PCR purification column or by phenol/chloroform extraction.

Ligate random peptide cassette in thioredoxin expression vector

18. Combine 8 μ g DNA cassette (step 13) and 12 μ g vector (step 17) in water to a total volume of 860 μ l.
19. Add 100 μ l of 10 \times T4 DNA ligase reaction buffer and 40 μ l (80,000 U) T4 DNA ligase. Incubate 16 hr at 16°C.
20. Purify ligated DNA using a QIAquick gel extraction kit according to manufacturer's instructions. Elute DNA from the column using 30 μ l ultrapure water.

It is important to remove as much salt, buffer, and protein from the ligated DNA as possible prior to electroporation.

Electroporate ligated DNA

21. Thaw 350 μ l electroporation-competent *E. coli* MC1061 on ice and add 30 μ l purified ligated plasmid. Transfer mixture to a 0.2-cm-gap electroporation cell.
22. Electroporate using the following conditions: 2.5 kV, 200 Ω , and 25 μ F.
23. Recover cells in 25 ml prewarmed SOC medium and incubate 1.5 hr at 37°C with gentle rocking.
24. Determine transformation efficiency by plating serial dilutions on LB plates containing 50 μ g/ml ampicillin.
25. Transfer remaining cells to 1 liter LB liquid medium containing 50 μ g/ml ampicillin and incubate overnight at 37°C.
26. Purify plasmid DNA using a commercially available large-scale plasmid preparation kit or using successive ethidium bromide/CsCl gradients (UNIT 2.4). Determine concentration and bring to 40 μ g/ml for screening (Basic Protocol 2).

ISOLATION OF PEPTIDE APTAMERS FOR SPECIFIC PROTEINS USING THE INTERACTION TRAP TWO-HYBRID SYSTEM

The interaction trap two-hybrid system (Gyuris et al., 1993; UNIT 20.1) is an established method for screening proteins for interactions with genomic and cDNA libraries (reviewed by Bai and Elledge, 1996; Finley and Brent, 1997). The interaction trap can also be extended to screen combinatorial libraries of peptide aptamers for interactions with specific proteins (Yang et al., 1995; Colas et al., 1996). The interaction trap consists of the following parts: (1) a constitutively expressed target protein fused to a *LexA* DNA-binding domain, referred to as the “bait;” (2) a galactose-induced combinatorial library of thioredoxin peptide aptamers fused to an activation domain, referred to as the “prey;” and (3) *LexA*-operator-*leu2* and *LexA*-operator-*lacZ* reporter genes for detecting interactions between the peptide aptamer prey and target protein bait. The bait protein binds to the *LexA* operators upstream of the reporters, but does not activate transcription of the reporters. Interaction between a peptide aptamer prey and target protein bait is detected by activation of reporter genes in the presence of galactose and not in the presence of

**BASIC
PROTOCOL 2**

**Generation and
Use of
Combinatorial
Libraries**

24.4.5

glucose. Figure 20.1.2 illustrates the isolation of proteins that interact with specific targets using the interaction trap.

In the first part of this protocol, the bait plasmid (pBait) is constructed by inserting DNA that encodes for the target protein into the polylinker of pEG202, in frame with *LexA*. The chimeric *LexA*-bait fusion protein is constitutively expressed using the ADH1 promoter. It is transformed into the appropriate yeast strain (EGY48) by a standard lithium acetate transformation procedure (UNIT 13.7). To be useful in the interaction trap two-hybrid system, the bait proteins must enter the nucleus, bind to the *LexA* operators, and not self-activate the *leu2* and *lacZ* reporters. After construction, pBait is characterized using protocols described elsewhere (UNIT 20.1).

The pJM-1 peptide aptamer library is used to select aptamers that bind specific protein targets using the interaction trap. pJM-1 contains a thioredoxin aptamer fused to a nuclear localization signal, a transcription activation domain, and an epitope tag under the control of the *gal1* promoter. Peptide aptamer expression is induced in the presence of galactose and repressed in the presence of glucose. A high-efficiency lithium acetate transformation procedure (Gietz and Schiestl, 1995; outlined below) is used rather than the standard procedure (UNIT 13.7) to introduce the aptamer library into the yeast strain EGY48, which contains an integrated *LexA*-operator-*leu2* reporter gene, *LexA*-operator-*lacZ* reporter plasmid and a bait plasmid. Interactions between the bait protein and the peptide aptamer prey are initially detected on galactose plates that lack leucine. Galactose induces the expression of the peptide aptamer and the absence of leucine selects for peptide aptamer/bait protein interactions that activate the *leu2* reporter. Interactions are verified by subsequently testing for galactose-dependent growth on –Leu plates and galactose-dependent blue color on Xgal plates.

The lithium acetate transformation procedure used here typically yields 10^5 to 10^6 transformants per μg of plasmid DNA. The protocol should be optimized for individual strains to achieve maximum transformation efficiency. In particular, variables such as cell concentration and heat shock time need to be optimized. The highest transformation efficiencies are obtained with $1 \mu\text{g}$ plasmid DNA per $50 \mu\text{l}$ competent yeast cells and generally do not scale up with similar efficiencies. The protocol below is designed for the transformation of $50 \mu\text{g}$ of peptide aptamer library.

Materials

- DNA encoding bait protein of interest
- Plasmid DNA: pEG202 (Fig. 20.1.3), pSH18-34 (Fig. 24.4.2)
- Yeast strain: EGY48 *ura3 trp1 his3 3LexA-operator-leu2*
- Complete minimal (CM) dropout medium (UNIT 13.1) and plates supplemented with either 2% (w/v) glucose (Glu) or 2% (w/v) galactose and 1% (w/v) raffinose (Gal/Raf):
 - Glu/CM –His,–Ura (10-cm plates and liquid medium)
 - Glu/CM –His,–Ura,–Trp (10- and 15-cm plates)
 - Glu/CM –His,–Ura,–Trp,–Leu (10-cm plates)
 - Gal/Raf/CM –His,–Ura,–Trp (liquid medium)
 - Gal/Raf/CM –His,–Ura,–Trp,–Leu (10- and 15-cm plates)
- 100 mM and 1 M lithium acetate, pH 7.5, filter sterilized
- 50% (w/v) polyethylene glycol, mol. wt. 3350 (PEG 3350; Sigma)
- 2 mg/ml single-stranded carrier DNA (sodium salt Type III from salmon testes; Sigma) TE buffer (APPENDIX 2)
- 40 $\mu\text{g}/\text{ml}$ peptide aptamer library DNA (pJM-1 aptamer plasmid; see Basic Protocol 1)

2× glycerol storage solution: 65% (v/v) glycerol, 0.1 M MgSO₄, 25 mM Tris-Cl, pH 7.4 (APPENDIX 2)
10-cm Xgal plates (UNIT 13.1)
Glu/CM –His,–Ura,–Trp, Xgal
Gal/Raf/CM –His,–Ura,–Trp, Xgal
PCR primers for thioredoxin
30° and 42°C incubators or water baths

Additional reagents and equipment for subcloning DNA (UNIT 3.16); manipulating yeast (UNIT 13.2); lithium acetate yeast transformation (UNIT 13.7); characterizing bait plasmids (UNIT 20.1); determination of cell density (UNIT 13.2) and plating efficiency (UNIT 20.1); replica plating (UNITS 1.3 & 13.2); yeast plasmid preparation (UNIT 13.11); plasmid sequencing (UNIT 7.3); *E. coli* transformation (UNIT 1.8); agarose gel electrophoresis (UNIT 2.5A); and PCR (UNIT 15.1)

Construct bait plasmid (pBait)

1. Using standard subcloning techniques (UNIT 3.16), insert DNA that codes for the bait protein into the polylinker of pEG202 to create the bait plasmid (pBait).
2. Transform pBait and pSH18-34 (*lacZ* reporter) into the interaction trap selection strain (EGY48) by lithium acetate yeast transformation (UNIT 13.7).
3. Plate transformants on Glu/CM –His,–Ura plates and place in a 30°C incubator.

Characterize bait protein

4. Confirm that the bait protein does not self-activate the reporter genes by performing plate assays for *lacZ* activation and leucine requirement (UNIT 20.1).

If the bait protein activates the leu2 and/or lacZ reporter genes, variations of the interaction trap that reduce reporter sensitivity should be tried. Yeast strains and/or plasmids containing less-sensitive leu2 and lacZ reporters reduce the background reporter output to reasonable levels. Yeast strains (Table 20.1.2) and plasmids (Fig. 24.4.2) with less sensitive reporters are described in UNIT 20.1. Truncating or separating the protein target can also eliminate transcription self-activation.

5. Confirm bait protein synthesis using the repression assay described in UNIT 20.1.

Baits that do not repress the expression of β-galactosidase in the repression assay may not be expressed correctly or may be incapable of entering the nucleus. Expression of full-length baits can be verified by immunoblotting. If full-length baits are expressed, their entry into the nucleus can be facilitated by adding a nuclear localization signal (J. Kamens, unpub. observ.). See Table 20.1.1 for description of plasmid pJK202 (a bait vector that contains a nuclear localization signal).

Transform peptide aptamer library into pBait-containing yeast

6. Inoculate 20 ml Glu/CM –His,–Ura liquid medium with transformed EGY48 (step 3) and incubate overnight at 30°C with shaking.
7. Take an OD₆₀₀ measurement and dilute to a concentration of 5 × 10⁶ cells/ml in 250 ml Glu/CM –His,–Ura.

An OD₆₀₀ of 0.1 corresponds to ~3 × 10⁶ cells/ml. This value should be confirmed for each yeast strain used (UNIT 13.2).

8. Incubate cells at 30°C with shaking until they reach an OD₆₀₀ of 0.6 to 0.8 (~5 to 6 hr).

This will yield enough yeast for 50 transformations.

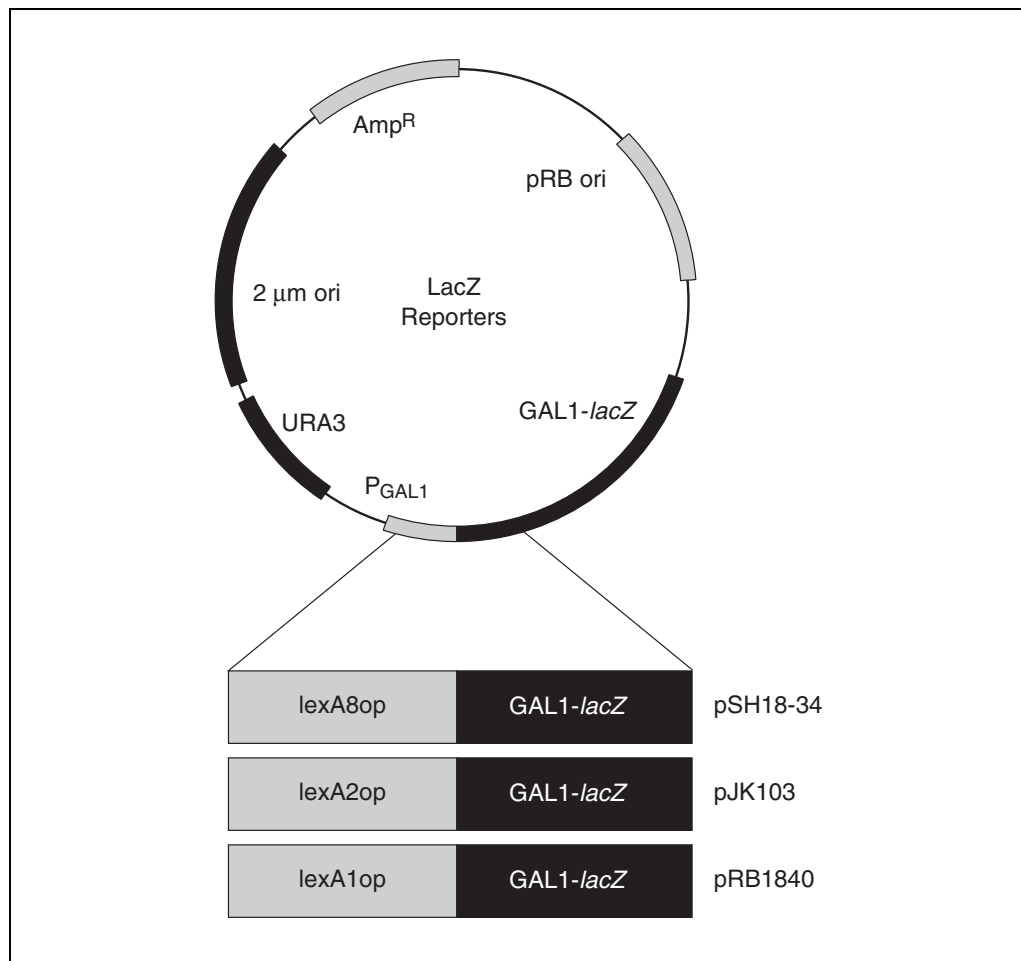


Figure 24.4.2 *lacZ* reporter plasmids. The *lacZ* reporter plasmids are derived from a plasmid that contains a wild-type GAL1 promoter fused to the *lacZ* gene (Yocum et al., 1984). *lacZ* reporters with different sensitivities are constructed by inserting different numbers of *lexA* operators into a plasmid (pLR1Δ1) that has the GAL1 upstream activating sequences (UAS_G) deleted (West et al., 1984). The *lacZ* reporters pSH18-34 (Gyuris et al., 1993), pJK103 (Kamens and Brent, 1991), and pRB1840 (Brent and Ptashne, 1985) contain eight, two, or one *lexA* operator(s). The sensitivity of the *lacZ* reporter decreases with the number of *lexA* operators.

9. Divide culture into five 50-ml conical centrifuge tubes and centrifuge 5 min at 3000 × *g*, room temperature.
10. Decant supernatant and resuspend each yeast pellet in 25 ml sterile water. Repeat centrifugation.
11. Decant supernatant and resuspend each yeast pellet in 1 ml of 100 mM lithium acetate. Transfer to a 1.5-ml microcentrifuge tube and pellet yeast by centrifuging 15 sec at 20,800 × *g*, room temperature.
12. Remove supernatant with a pipet and resuspend each yeast pellet in 350 μl of 100 mM lithium acetate (final volume ~500 μl).
13. Split the contents of each tube into ten 50-μl portions and pellet yeast by centrifuging 15 sec at 20,800 × *g*, room temperature.

14. Remove supernatant with a pipet and add the following ingredients to each sample in the order listed:

240 μ l 50% (w/v) PEG 3350
36 μ l 1 M lithium acetate
50 μ l 2 mg/ml single-stranded carrier DNA (100 μ g)
25 μ l 40 μ g/ml peptide aptamer library DNA (1 μ g).

Single-strand carrier DNA needs to be heated to 95°C for 5 min and cooled on ice prior to use.

15. Vortex the transformation mixture vigorously until the yeast pellet is completely resuspended and incubate 30 min at 30°C.
16. Heat shock 20 min at 42°C.
17. Pellet yeast by centrifuging 15 sec at 20,800 \times g, room temperature.
18. Remove supernatant with a pipet and resuspend pellet in 500 μ l sterile water.
19. Plate 48 transformations on individual 15-cm Glu/CM –His,–Ura,–Trp plates.
20. Plate 400 μ l of the two remaining transformations on 15-cm Glu/CM –His,–Ura,–Trp plates.
21. Use the remaining 100 μ l to determine the transformation efficiency. Perform a series of 10-fold dilutions in sterile water and plate on 10-cm Glu/CM –His,–Ura,–Trp plates.
22. Incubate 2 to 3 days at 30°C (until colonies are ~1 mm in diameter).

Pool transformants

23. Pool yeast from all 50 transformation plates (steps 19 and 20) in a 50-ml centrifuge tube.

See UNIT 20.1 for protocol on scraping yeast from plates.

24. Add an equal volume of 2 \times glycerol storage solution to the pooled yeast cells. Divide into 1-ml aliquots and store at –70°C.
25. Determine the plating efficiency of the frozen aliquots as described in *UNIT 20.1*.

Screen for peptide aptamers that interact with target protein

26. Inoculate ten library equivalents of the peptide aptamer library in 2 ml Gal/Raf/CM –His,–Ura,–Trp liquid medium. Incubate 4 hr at 30°C with shaking.

One library equivalent equals the total number of yeast transformants containing the peptide aptamer library, as determined in step 21.

27. Centrifuge 4 min at 3000 \times g, room temperature.
28. Remove supernatant with a pipet and resuspend yeast in 1 ml sterile water.
29. Spread yeast at a density of 10⁶ yeast cells/plate on 15-cm Gal/Raf/CM –His, –Ura, –Trp, –Leu plates.
30. Incubate at 30°C and monitor plates daily for growth.
31. Streak colonies onto 10-cm Glu/CM –His,–Ura,–Trp master plates. Incubate 1 to 2 days at 30°C.

32. Replica plate the master plates on the following indicator plates:

Glu/CM –His,–Ura,–Trp,–Leu
Gal/Raf/CM –His,–Ura,–Trp,–Leu
Glu/CM –His,–Ura,–Trp, Xgal
Gal/Raf/CM –His,–Ura,–Trp, Xgal.

33. Identify colonies that show galactose-dependent growth on –Leu plates and galactose-dependent blue color on Xgal plates.

Isolate peptide aptamers

34. Isolate the desired peptide aptamer expression plasmid (UNIT 13.11).

The plasmid preparation will contain a mixture of the three plasmids used in the interaction trap (pJM-1 aptamer plasmid, pSH18-34, and pBait).

35. Use plasmids as templates for sequencing the peptide aptamer variable regions (UNIT 7.3).

36. To separate the aptamer plasmid from pBait and pSH18-34, transform *E. coli* (UNIT 1.8) and identify the appropriate transformants by PCR (UNIT 15.1) using primers that amplify thioredoxin.

*Colonies that contain the peptide aptamer will appear as a bright band on an ethidium bromide agarose gel (UNIT 2.5A) after 20 cycles of PCR. Colonies that do not contain the peptide aptamer will appear as a faint band that is 20 base pairs shorter than the aptamer. This shorter band is due to the presence of native *E. coli* thioredoxin.*

BASIC PROTOCOL 3

DEFINING RECOGNITION SPECIFICITY WITH INTERACTION MATING

Interaction mating is a variation of the interaction trap. It allows interactions between large panels of proteins to be analyzed (Finley and Brent, 1994). Haploid yeast exist in one of two mating types (**a** or **α**). Haploid yeast that contain protein targets or related protein baits in one mating type and peptide aptamer preys in the opposite mating type can mate to form diploids that carry both the aptamers and their targets or related proteins. Interaction between the peptide aptamer prey and protein target bait is detected by the activation of two reporter genes: LexAop-LEU2 and LexAop-LacZ. Using the mating interaction assay, panels of related or mutated proteins can be assayed simultaneously for interactions with panels of peptide aptamers. See Figure 24.4.3 for schematic of the interaction mating assay.

Materials

Plasmid DNA: pBait(s) (see Basic Protocol 2), peptide aptamer preys (see Basic Protocol 2), pEG202 (Fig. 20.1.3), pJG4-5 (Fig. 24.4.1), pSH18-34 (Fig. 24.4.2)

Yeast strains:

EGY42: *Matα ura3 trp1 his3 leu2*

EGY48: *Matα ura3 trp1 his3 3LexA-operator-leu2*

10-cm complete minimal (CM) dropout plates (UNIT 13.1) supplemented with either 2% (w/v) glucose (Glu) or 2% (w/v) galactose and 1% (w/v) raffinose (Gal/Raf):

Glu/CM –Trp

Glu/CM –His,–Ura

Glu/CM –His,–Ura,–Trp,–Leu

Gal/Raf/CM –His,–Ura,–Trp,–Leu

YPD plates (UNIT 13.1)

Xgal plates (UNIT 13.1):

Glu/CM –His,–Ura,–Trp, Xgal

Gal/Raf/CM –His,–Ura,–Trp, Xgal.

30°C incubator

Additional reagents and equipment for lithium acetate yeast transformation (UNIT 13.7) and replica plating (UNITS 1.3 & 13.2)

1. Transform individual peptide aptamer prey plasmids and a control plasmid (pJG4-5) into EGY48 (*Mata*) using lithium acetate transformation (UNIT 13.7). Select transformants on 10-cm Glu/CM –Trp plates.

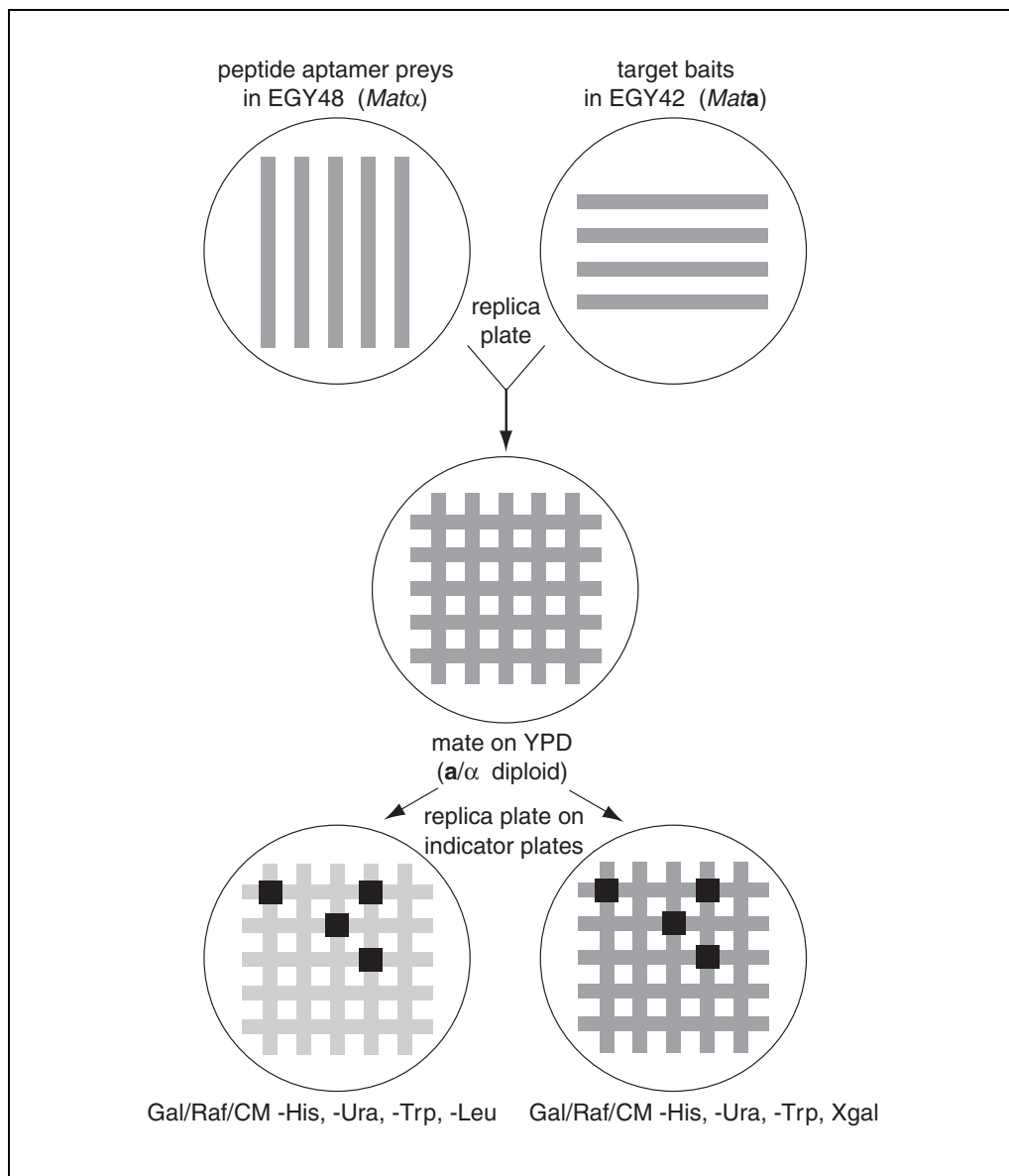


Figure 24.4.3 Mating interaction assay (Finley and Brent, 1994). Peptide aptamer preys in the yeast strain EGY48 (*Mata*) are streaked vertically on Glu/CM –Trp plates. Target protein baits and *lacZ* reporter (pSH18-34) in the yeast strain EGY42 (*Mata*) are streaked horizontally on Glu/CM –His,–Ura plates. The yeast strains are replica plated perpendicular to each other on YPD plates. The haploid strains carrying the baits and preys mate where the two strains intersect, forming (*a/α*) diploids that contain the bait, prey, and *lacZ* reporter. The YPD plates are replica plated onto the following interaction detection plates: Glu/CM –His,–Ura,–Trp,–Leu; Gal/Raf/CM –His,–Ura,–Trp,–Leu; Glu/CM –His,–Ura,–Trp, Xgal; Gal/Raf/CM –His,–Ura,–Trp, Xgal. Interacting baits and prey display galactose-dependent growth and blue color on –Leu and Xgal plates, respectively.

2. Transform individual target protein baits (pBaits) with pSH18-34 (*lacZ* reporter) and a control plasmid (pEG202) with pSH18-34 into EGY42 (*Mata*). Select transformants on 10-cm Glu/CM –His,–Ura plates.
3. Streak, in parallel lines, individual peptide aptamers and their control prey strains on 10-cm Glu/CM –Trp plates.
4. Streak, in parallel lines, individual protein targets and their control bait strains on 10-cm Glu/CM –His,–Ura plates.
5. Incubate all plates overnight at 30°C.
6. Replica plate the protein target bait and peptide aptamer prey strains on the same replica velvet by first replica plating the bait strains and then replica plating the prey strains perpendicular to the baits (see Figure 24.4.3 for schematic).
7. Transfer the yeast imprint to a 10-cm YPD plate and incubate overnight at 30°C.
8. Replica plate the YPD plate onto a replica velvet. Transfer the yeast imprint to the following indicator plates:
 - Glu/CM –His,–Ura,–Trp,–Leu
 - Gal/Raf/CM –His,–Ura,–Trp,–Leu
 - Glu/CM –His,–Ura,–Trp, Xgal
 - Gal/Raf/CM –His,–Ura,–Trp, Xgal.
9. Analyze plates for mating.

Mating occurs at the intersection of the Mata α and Mata strains. Diploid colonies should grow on the Xgal plates. Interactions between the peptide aptamer preys and protein target baits produce blue color on the galactose Xgal plates and growth on the galactose –Leu plates at the intersection of the strains.

BASIC PROTOCOL 4

AFFINITY MATURATION OF PEPTIDE APTAMERS

The binding affinity between a peptide aptamer and its protein target can be improved by mutating the peptide aptamer variable region and reselecting for aptamers that bind the target protein using a more stringent interaction trap. In this protocol, peptide aptamers are mutated by random PCR mutagenesis as described by Cadwell and Joyce, 1994. Alternatively, degenerate oligonucleotides that code for the variable region and have varying degrees of randomness can be synthesized using an automated DNA synthesizer (UNIT 2.11). The stringency of the interaction trap selection is enhanced by decreasing the number of *LexA* operators upstream of the *lacZ* reporter gene. A series of *lacZ* reporter genes containing eight, two, and one *LexA* operator(s) (Brent and Ptashne, 1985) are used to select aptamers with increased affinity toward their targets.

Materials

- 5 U/ μ l *Taq* polymerase and 10 \times buffer (Life Technologies)
- 1 M MgCl₂
- 100 mM dATP
- 100 mM dGTP
- 100 mM dCTP
- 100 mM dTTP
- 20 μ M primer 1: 5'-CCGCCGCCTGAATTCATGAGCGATAAAATTATTAC-3'
- 20 μ M primer 2: 5'-CGGGGCGATCATTTTGCACGGACC-3'
- Plasmid DNA: peptide aptamer plasmid (see Basic Protocol 2), pBait (see Basic Protocol 2), pJM-1 (Fig. 24.4.1), pRB1840 (1-LexAop-LacZ reporter plasmid; Fig. 24.4.2), and pJK103 (Fig. 24.4.2)

Mg²⁺/Mn²⁺ solution: 45 mM MgCl₂ and 5 mM MnCl₂

PCR purification column (optional; e.g., Qiagen)

Yeast strain: EGY48 *Mata ura3 trp1 his3 3LexA-operator-leu2*

Complete minimal (CM) dropout medium (*UNIT 13.1*) and plates supplemented with either 2% (w/v) glucose (Glu) or 2% (w/v) galactose and 1% (w/v) raffinose (Gal/Raf):

Glu/CM –His,–Ura (10-cm plates)

Glu/CM –His,–Ura,–Trp (10-cm plates)

Gal/Raf/CM –Ura,–His,–Trp (liquid medium)

Xgal plates (*UNIT 13.1*)

Glu/CM –His,–Ura,–Trp, Xgal (10-cm plates)

Gal/Raf/CM –His,–Ura,–Trp, Xgal (10- and 15-cm plates)

PCR tubes

Automated thermal cycler

30°C incubator

Additional reagents and equipment for agarose gel electrophoresis (optional; *UNIT 2.5A*), digesting and cloning peptide aptamer mutants (see Basic Protocol 1), lithium acetate yeast transformation (see Basic Protocol 2 and *UNIT 13.7*), determination of plating efficiency (*UNIT 20.1*), plasmid rescue (*UNIT 13.11*), and plasmid DNA sequencing (*UNIT 7.3*)

Mutagenize peptide aptamer variable region

1. Prepare PCR premixture (total 3.775 ml):

500 µl 10× *Taq* polymerase buffer

5 µl 1 M MgCl₂

10 µl 100 mM dATP

10 µl 100 mM dGTP

50 µl 100 mM dCTP

50 µl 100 mM dTTP

125 µl 20 µM primer 1

125 µl 20 µM primer 2

2.9 ml water.

2. For each sample, add the following reagents to a PCR tube:

12 µl water

1 µl peptide aptamer expression vector

10 µl Mg²⁺/Mn²⁺ solution

76 µl PCR premixture

1 µl *Taq* polymerase (5 U).

3. Amplify the reaction using the following PCR reaction program:

4 cycles: 30 sec 95°C (denaturation)

 1 min 55°C (annealing)

 1 min 72°C (extension).

4. Remove 13 µl reaction mixture and add to a new PCR tube containing:

10 µl Mg²⁺/Mn²⁺ solution

76 µl PCR premixture

1 µl *Taq* polymerase.

Amplify using the same PCR program.

5. Repeat for a total of ten rounds of amplification.

6. Purify the PCR product with a commercially available PCR purification column or by agarose gel electrophoresis (*UNIT 2.5A*).

Construct mutagenized peptide aptamer expression vector

7. Digest purified PCR product with *Ava*II and subclone it into *Rsr*II-cut pJM-1 using standard subcloning techniques (*UNIT 3.16*). Electroporate the ligated product as described above (see Basic Protocol 1, steps 20 to 26).

Select mutagenized aptamers by the interaction trap

8. Transform EGY48 with pBait and pRB1840 by standard lithium acetate yeast transformation (*UNIT 13.7*). Select transformants on 10-cm Glu/CM –His,–Ura plates.
9. Using the high-efficiency lithium acetate procedure (see Basic Protocol 2, steps 6 to 22), transform 10 to 50 µg of mutagenized peptide aptamer library into EGY48 containing pBait and pRB1840. Select transformants on 10-cm Glu/CM –His,–Ura, –Trp plates.
10. Pool transformants and determine plating efficiency as described in *UNIT 20.1*.
11. Inoculate approximately five library equivalents in 1 ml Gal/Raf/CM –His, –Ura, –Trp liquid medium. Incubate 4 hr at 30°C with shaking.

One library equivalent equals the total number of yeast transformants containing the peptide aptamer library as determined in step 9.
12. Centrifuge 4 min at 3000 × *g*, room temperature. Remove supernatant and resuspend yeast pellet in 1 ml sterile water.
13. Spread yeast on 15-cm Gal/Raf/CM –His,–Ura,–Trp, Xgal plates and incubate at 30°C until colonies appear (~2 days).
14. Streak blue colonies onto a 10-cm Glu/CM –His,–Ura,–Trp master plate and incubate 1 day at 30°C.
15. Replica plate the master plate onto 10-cm Gal/Raf/CM –His,–Ura,–Trp, Xgal and Glu/CM –His,–Ura,–Trp, Xgal plates.
16. Rescue plasmids (*UNIT 13.11*) from the galactose-dependent blue colonies and reintroduce (*UNIT 13.7*) the plasmids into the yeast strain EGY48 that contains pBait and pRB1840 to reconfirm the phenotype.
17. Rescue the plasmids from the galactose-dependent blue colonies and sequence (*UNIT 7.3*) the variable regions.

BASIC PROTOCOL 5

FORWARD ANALYSIS OF CELLULAR PROCESSES USING PEPTIDE APTAMERS

Combinatorial libraries of peptide aptamers can function as dominant agents for the forward analysis of cellular processes. Peptide aptamers function as “mutagens”, randomly inhibiting gene function and altering the phenotype of an organism. Forward analysis with peptide aptamers involves expressing combinatorial libraries in organisms and screening or selecting for aptamer-induced changes in their phenotypes. The peptide aptamer targets are subsequently identified using the interaction trap. The protein targets can be identified from panels of proteins using a mating interaction assay (Finley and Brent, 1994) or by screening for aptamer interactions against genomic or cDNA libraries using the interaction trap (*UNIT 20.1*). Currently, complete panels of proteins are not available for any organisms except yeast. As a result, panels of known proteins will need

Peptide Aptamers

24.4.14

to be combined with cDNA and genomic libraries of proteins to identify peptide aptamer targets.

The design of a genetic selection is beyond the scope of this protocol. A typical genetic selection requires the transformation of an organism selection strain with a peptide aptamer expression library containing 10^6 to 10^7 members. Peptide aptamers are expressed under the control of an inducible promoter, allowing the aptamer-induced phenotype to be confirmed by comparing the effects of the aptamer expression plasmid in the presence or absence of the inducer. The protocol described below for a genetic selection using yeast may be adapted to a variety of organisms.

Materials

Yeast strain for genetic selection

Peptide aptamer library: pJM-2 or pJM-3 (Basic Protocol 1; Fig. 24.4.1)

Complete minimal (CM) dropout liquid medium (*UNIT 13.1*) and plates supplemented with either 2% (w/v) glucose (Glu) or 2% (w/v) galactose and 1% (w/v) raffinose (Gal/Raf):

Glu/CM –Trp (10-cm plates)

Gal/Raf/CM –Trp (10-cm plates and liquid medium)

30°C incubator

Additional reagents and equipment for high-efficiency lithium acetate yeast transformation (see Basic Protocol 2), determination of plating efficiency (*UNIT 20.1*), isolation of plasmids (*UNIT 13.11*), plasmid DNA sequencing (*UNIT 7.3*), and target identification (see Support Protocol)

1. Transform 50 to 100 μg of the peptide aptamer library (in pJM-2 or pJM-3) into a yeast selection strain (10^6 to 10^7 transformants) using the high-efficiency lithium acetate transformation procedure (see Basic Protocol 2, steps 6 to 22). Plate transformants on 10-cm Glu/CM –Trp plates and incubate at 30°C until colonies are ~1 mm in diameter (~2 to 3 days).
2. Pool yeast cells and determine the plating efficiency as described in *UNIT 20.1*.
3. Inoculate ten library equivalents in 1 ml Gal/Raf/CM –Trp liquid medium. Incubate 4 hr at 30°C with shaking.

One library equivalent equals the total number of transformants containing the peptide aptamer library as determined in step 1.
4. Centrifuge culture 4 min at $3000 \times g$, room temperature. Remove supernatant with a pipet and resuspend the yeast pellet in sterile water.
5. Plate yeast on 10-cm Gal/Raf/CM –Trp selection plates and incubate under selection conditions.
6. Streak positive colonies on Glu/CM –Trp master plates.
7. Confirm galactose-dependent phenotype by replicating master plate onto Glu/CM –Trp and Gal/Raf/CM –Trp plates and incubate under selection conditions.
8. Isolate peptide aptamer expression plasmids (pJM-2 or pJM-3) from the yeast colonies that show the galactose-dependent phenotype (*UNIT 13.11*).
9. Reconfirm the peptide aptamer phenotype by transforming the isolated plasmid into the selection strain and testing for galactose-dependent phenotype.
10. Isolate the peptide aptamer expression plasmids (*UNIT 13.11*) for sequencing (*UNIT 7.3*) and target identification (see Support Protocol).

IDENTIFICATION OF PEPTIDE APTAMER TARGETS

The protein targets of the genetically selected peptide aptamers (Basic Protocol 5) can be identified using the interaction mating assay (see Basic Protocol 3) or by interaction hunts against cDNA or genomic libraries (*UNIT 20.1*). Genomic and cDNA libraries are constructed as preys since they contain many sequences capable of activating transcription in the bait configuration. As such, the peptide aptamers need to be transferred to the bait plasmid pEG202 to identify their targets in these libraries. Protocols for constructing cDNA and genomic libraries can be found in *UNITS 5.7, 5.8A & 5.8B*.

Putative peptide aptamer targets identified with either mating interaction panels or hunts should be verified using genetic tests such as: (1) immunoprecipitation to confirm the aptamer interactions *in vivo*, (2) epistasis analysis to confirm that the aptamer functions in the same area as the target protein, or (3) comparison of the phenotype(s) caused by deletion and overexpression of target protein with the phenotype caused by the aptamer.

Materials

DNA encoding thioredoxin peptide aptamer (Basic Protocol 5)

Plasmid DNA: pEG202 (Fig. 20.1.3), pSH18-34 (Fig. 24.4.2), pJG4-5 (Fig. 24.4.1)

Yeast strains:

EGY42, *Mat α ura3 trp1 his3 leu2*

EGY48, *Mata ura3 trp1 his3 3LexA-operator-leu2*

Complete minimal (CM) dropout liquid medium (*UNIT 13.1*) and plates

supplemented with either 2% (w/v) glucose (Glu) or 2% (w/v) galactose and 1% (w/v) raffinose (Gal/Raf):

Glu/CM –His,–Ura (10-cm plates)

Glu/CM –Trp (10-cm plates)

Prey library (see Table 20.1.3)

Additional reagents and equipment for PCR (*UNIT 15.1*), standard subcloning (*UNIT 3.16*), standard lithium acetate yeast transformation (*UNIT 13.7*), interaction mating (see Basic Protocol 3), interaction trap (*UNIT 20.1*)

Transfer peptide aptamers from pJM-2 or pJM-3 into pEG202

1. PCR amplify the DNA encoding the thioredoxin peptide aptamer using primers that contain restriction sites compatible with the polylinker of pEG202 and in frame with *LexA* (Fig. 20.1.3).
2. Using standard subcloning techniques (*UNIT 3.16*), insert the PCR product into pEG202 to create the peptide aptamer bait.
3. Transform the individual peptide aptamer baits and pSH18-34 (*lacZ* reporter) into EGY48 (*Mat α*) by standard lithium acetate yeast transformation (*UNIT 13.7*). At the same time transform a control plasmid (pEG202) and pSH18-34 into EGY48.
4. Select transformants on 10-cm Glu/CM –His,–Ura plates.

Identify targets

For mating interaction assay:

- 5a. Construct a panel of desired proteins by inserting coding regions of proteins into the polylinker of pJG4-5 (prey plasmid, Fig. 24.4.1).
- 6a. Transform prey plasmids and a control plasmid (pJG4-5) into EGY42 (*Mata*) by standard lithium acetate transformation. Select transformants on 10-cm Glu/CM –Trp plates.

- 7a. Mate strains containing peptide aptamer baits and target protein preys and score interactions as described (see Basic Protocol 3, steps 3 to 9).

For interaction trap library hunts:

- 5b. Transform strains containing individual peptide aptamers and pSH18-34 (step 3) with a library of genomic or cDNA preys. Follow the protocol in *UNIT 20.1* for transforming cDNA and genomic prey libraries.
- 6b. Select peptide aptamer target(s) using the interaction trap hunt protocol described in *UNIT 20.1*.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

DNA elution buffer

- 10 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
- 1 mM EDTA, pH 8 (*APPENDIX 2*)
- 50 mM NaCl
- Store up to 1 year at room temperature

Nondenaturing loading buffer

- 50 mM Tris·Cl, pH 8 (*APPENDIX 2*)
- 50 mM EDTA, pH 8 (*APPENDIX 2*)
- 50% (v/v) glycerol
- Store up to 1 year at 4°C

COMMENTARY

Background Information

Understanding cellular processes within organisms relies on forward and reverse genetic approaches to identify genetic network members and connections. In forward genetic analysis, genes are identified by isolating randomly generated mutants and mapping the genes responsible for their mutant phenotypes. Reverse genetic analysis, by contrast, involves mutating individual genes and monitoring the resulting phenotype. While both approaches are effective, they are difficult to perform, especially in diploid organisms. In diploid organisms, the identification of recessive mutations requires two generations of breeding to generate homozygotes. Consequently, genetic approaches requiring homozygous recessive mutations can only be fully applied to organisms with well-developed genetics such as phage, bacteria, yeast, *C. elegans*, and *Drosophila*.

Dominant agents that affect gene products in *trans*, instead of genes, have been developed to overcome problems associated with the analysis of recessive mutations in diploid organisms. A variety of dominant agents exist for the reverse analysis of cellular processes. These include: small molecule inhibitors (Mitchison,

1994), dominant negative proteins (Herskowitz, 1987), antibodies (Gorbsky et al., 1998), antisense RNA (Branch, 1998), ribozymes (Bramlage et al., 1998), and nucleic acid aptamers (*UNIT 24.3*; Thomas et al., 1997). These agents have improved the ability to analyze processes in diploid organisms; however, they too have limitations. For example, forward analysis requires large-scale generation of agents that are capable of inactivating the function of almost any gene product, but agents such as small molecule inhibitors and dominant negative proteins may not exist for all gene products. Similarly, although it should, in theory, be possible to generate agents such as antibodies, ribozymes, nucleic acid aptamers, and antisense RNA against almost any gene product, antibodies are not membrane permeable, and large-scale injection is tedious and impractical for most organisms. RNA agents are not very stable and it is difficult to predict sites on RNA that are exposed for inhibition by antisense RNA and ribozymes. Furthermore, agents that inhibit at the RNA level (antisense RNA and ribozymes) are affected by the stability of the protein target, which can affect the onset and/or the extent of the phenotype.

The development of combinatorial technologies for obtaining biomolecules with desired properties (*UNITS 24.2 & 24.3*; Ellington and Szostak, 1990; Scott and Smith, 1990) presents new avenues for generating “genetic” agents for characterizing genetically intractable organisms. This unit describes methods to construct combinatorial libraries of “genetic” agents referred to as peptide aptamers. Peptide aptamer libraries consist of scaffold proteins that display variable peptides constrained at both ends on their surface. They are designed to interact and interfere with the biological function of proteins. Peptide aptamers are well suited for analyzing cellular processes in diploid organisms because they act in *trans* to inhibit gene products without altering their encoding DNA. Moreover, because they are isolated from combinatorial libraries, peptide aptamers can in principle be generated to inactivate almost any gene product.

Design of intracellular peptide aptamers

Peptide aptamers are designed to interact with their protein targets through variable peptide regions displayed on the surface of a scaffold protein. To date, only a limited number of scaffold proteins have been used within organisms to display linear and constrained peptides. These include: *E. coli* thioredoxin (Colas et al., 1996), Gal4 activation domain (Yang et al., 1995), green fluorescent protein (Caponigro et al., 1998), and staphylococcal nuclease (Norman et al., 1999). A comparison of the binding constants of these aptamers shows that constrained variable regions can bind their targets between 100- and 10,000-fold better than linear peptides (Geyer and Brent, 2000). Unconstrained peptides are also known to be unstable in *E. coli* (Davidson and Sauer, 1994). Constrained peptide libraries are therefore the preferred method for displaying combinatorial peptide libraries for intracellular applications. When choosing aptamer scaffolds, they should also be small, stable, soluble, and expressed at high levels without toxicity. The scaffold should be tolerant to the addition of protein moieties such as localization sequences, epitope tags, and purification tags.

Basic Protocol 1 describes the construction of a peptide aptamer library using *E. coli* thioredoxin as the scaffold protein. Thioredoxin was first used as a scaffold protein for displaying peptides as fusions to flagellin on the surface of *E. coli* (Lu et al., 1995). Thioredoxin possesses many characteristics that make it an excellent scaffold for intracellular applications.

Structural studies on thioredoxin reveal that its active site contains a 4–amino acid loop (–CGPC–) that is constrained by the two terminal cysteines (Katti et al., 1990). This loop is tolerant to peptide insertion (LaVallie et al., 1993) and provides a site for displaying variable peptides. Thioredoxin is a small (12 Kd) cytoplasmic protein that is nontoxic when expressed at high levels (LaVallie et al., 1993). Thioredoxin is often fused to proteins to enhance their solubility (LaVallie et al., 1993). This is a useful property for expressing random sequence libraries where many of the sequences may aggregate. Thioredoxin interacts with a variety of disulfide-containing protein substrates (Wetterauer et al., 1992), suggesting that it may also contribute to the binding interactions between peptide aptamers and their protein targets.

Reverse “genetic” analysis using peptide aptamers

Reverse genetic analysis using peptide aptamers involves isolating aptamers that interact with a specific gene product and monitoring the aptamer-induced phenotype. Peptide aptamers that interact with a chosen protein are selected using yeast two-hybrid systems or a variation thereof (Chien et al., 1991; Dalton and Treisman, 1992; Durfee et al., 1993; Gyuris et al., 1993; Vojtek et al., 1993). These systems share the following features: (1) a DNA-binding domain/target protein fusion, (2) a transcription activation domain/peptide aptamer fusion, and (3) reporter gene(s) to record interactions between the peptide aptamer and protein target (see *UNIT 20.1* for a detailed description of the yeast two-hybrid system).

Basic Protocol 2 describes the interaction trap two-hybrid system as a method for obtaining peptide aptamers that interact with a selected protein target. The interaction trap is an effective method for obtaining high-affinity peptide aptamers that bind specific proteins. Aptamers obtained using the interaction trap have dissociation constants greater than the 1 μ M detection limit required to activate the interaction trap reporters (Estojak et al., 1995).

To date, the interaction trap has been used to isolate peptide aptamers against a variety of protein targets including Cdk2 (Colas et al., 1996), Ras (Xu et al., 1997), HIV-1 Rev (Cohen, 1998), and E2F (Fabbrizio et al., 1999). The dissociation and half-inhibitory constants of these aptamers range from 10^{-8} to 5×10^{-11} M.

An advantage of using the interaction trap to select peptide aptamers is that selection occurs in an intracellular environment. This increases the probability that the aptamers will retain their function when expressed in the appropriate organism. Moreover, aptamers isolated using the interaction trap function effectively under a variety of *in vivo* conditions such as cell cultures (Cohen et al., 1998; Fabbri et al., 1999) and in *Drosophila* (Kolonin and Finley, 1998).

Specificity of peptide aptamers

To be useful for genetic analysis, a peptide aptamer must interact specifically with its protein target. Peptide aptamer specificity can be evaluated by analyzing the aptamer's ability to interact with related target proteins using the interaction trap. Basic Protocol 3 describes the mating interaction assay, an extension of the interaction trap developed by Finley and Brent (1994) for determining the specificity of peptide aptamers against a large panel of related proteins. The interaction mating assay allows panels of individual aptamers to be simultaneously screened for interactions with panels of related target proteins. Using this method, Colas et al. (1996) determined the specificity of aptamers isolated against cyclin-dependent kinase 2 (Cdk2). The majority of aptamers tested were highly specific for Cdk2 and not other closely related kinases with one exception: some of the aptamers also interacted with the closely related kinase Cdk3. Their results demonstrate that aptamers can be generated against different epitopes on Cdk2, some of which are conserved between different members of the cyclin-dependent kinases.

The mating interaction assay is also used to determine the specific regions and/or amino acids that aptamers recognize on the target protein. For example, Cohen et al. (1998) showed that one of the aptamers isolated against Cdk2 (Colas et al., 1996) acts as a competitive inhibitor of the Cdk2-dependent phosphorylation of histone H1. Interaction mating with a panel of mutant Cdk2 proteins revealed that specific active site residues are required for aptamer binding, supporting the competitive inhibition mechanism. In summary, interaction mating assays using panels of related and mutated proteins can be used to classify both the specificity and binding interactions of different aptamers targeted to the same protein.

Affinity maturation

Peptide aptamer selections using the interaction trap are limited to screening $\sim 10^6$ to 10^7 unique aptamers per experiment. This is a small representation ($\sim 9 \times 10^{-6}\%$) of the entire sequence space available to aptamers containing 20-mer variable regions (20^{20} possible sequences). In addition to the small sample size, many of the aptamers will contain stop codons within the variable region. As a result, it is likely that aptamers isolated using the interaction trap do not contain the optimal binding sequences for their target proteins.

Basic Protocol 4 describes a method to obtain aptamers with increased binding affinity. The protocol involves mutating the aptamer variable region and reselecting for binding to its target protein using an interaction trap that contains a more stringent reporter gene (Cohen, 1998; Colas et al., 2000). Mutations can be introduced using mutagenic PCR or by synthesizing degenerate oligonucleotides with varying degrees of randomness (see *UNIT 2.11* for a discussion on the construction of degenerate oligonucleotides). The stringency of the interaction trap is enhanced by reducing the number of *LexA* operators upstream of the reporters. The interaction trap in *UNIT 20.1* contains eight *LexA* operators in the *LexA-lacZ* (pJG4-5) and *LexA-leu2* (EGY48 strain) reporters. These reporters are capable of detecting interactions with dissociation constants of $<1 \mu\text{M}$ (Estojak et al., 1995). Other *lacZ* reporters, developed by Brent and Ptashne (1985), contain only one (pRB1840) or two (pJK103) *LexA* operators (Fig. 24.4.2). These operators have lower affinity for the *LexA* DNA-binding domain and detect interactions with dissociation constants between 20 nM and $<1 \mu\text{M}$ (Estojak et al., 1995).

The affinity maturation described in Basic Protocol 4 has been successfully used to enhance the affinity of aptamers isolated against Cdk2 (Cohen, 1998; Colas et al., 2000). The variable region of the anti-Cdk2 aptamer was mutated by PCR and reselected for binding to a *LexA*-Cdk2 fusion using the 1-*LexA*-operator *lacZ* reporter (pRB1840). Isolated aptamers all contained the same two amino acid substitutions. The dissociation constant of the mature aptamer was reduced to 5 nM, a 20-fold decrease from the starting aptamer ($K_d = 0.1 \mu\text{M}$).

Forward "genetic" analysis with peptide aptamers

Combinatorial libraries of peptide aptamers can function as dominant agents to randomly

Table 24.4.1 Degenerate Codons for Designing Combinatorial Peptide Libraries^a

Codon ^b	Properties	Amino acids ^c	No. of codons	Stop codons
NNN	All 20 amino acids	A(4), C(2), D(2), E(2), F(2), G(4), H(2), I(3), K(2), L(6), M(1), N (2), P(4), Q(2), R(6), S(6), T(4), V(4), W(1), Y(2)	64	TAA(1), TAG(1), TGA(1)
NNS	All 20 amino acids	A(2), C(1), D(1), E(1), F(1), G(2), H(1), I(1), K(1), L(3), M(1), N(1), P(2), Q(1), R(3), S(3), T(2), V(2), W(1), Y(1)	32	TAG (1)
NNC	15 amino acids	A(1), C(1), D(1), F(1), G(1), H(1), I(1), L(1), N(1), P(1), R(1), S(2), T(1), V(1), Y(1)	16	None
NWW	Charged, hydrophobic	D(1), E(1), F(1), H(1), I(2), K(1), L(3), N(1), Q(1), V(2), Y(1)	16	TAA (1)
RVK	Charged, hydrophilic	A(2), D(1), E(1), G(2), K(1), N(1), R(1), S(1), T(2)	12	None
DVT	Hydrophilic	A(1), C(1), D(1), G(1), N(1), S(2), T(1), Y(1)	9	None
NVT	Charged, hydrophilic	A(1), C(1), D(1), G(1), H(1), N(1), P(1), R(1), S(2), T(1), Y(1)	12	None
NNT	Mixed	A(1), C(1), D(1), F(1), G(1), H(1), I(1), L(1), N(1), P(1), R(1), S(2), T(1), V(1), Y(1)	16	None
VVC	Hydrophilic	A(1), D(1), G(1), H(1), N(1), P(1), R(1), S(1), T(1)	9	None
NTT	Hydrophobic	F(1), I(1), L(1), V(1)	4	None
RST	Small side chains	A(1), G(1), S(1), T(1)	4	None
TDK	Hydrophobic	C(1), F(1), L(1), W(1), Y(1)	6	TAG (1)

^aBased on a table described by Sidhu and Weiss (2000).

^bAbbreviations: D = A, G, T; K = G, T; N = A, G, C, T; R = A, G; S = C, G; V = A, C, G; W = A, T.

^cNumbers in parentheses indicate the number of codons for each amino acid.

inactivate gene products without altering their genetic material. The forward analysis of cellular processes using peptide aptamers involves expressing libraries of peptide aptamers within cells and screening for aptamer-induced phenotypes. The protein(s) and protein interactions disrupted by the aptamers are then identified. Basic Protocol 5 describes methods for performing forward analysis of cellular processes in yeast. Methods are also described for identifying peptide aptamer target(s) using interaction trap hunts with genomic or cDNA libraries or by mating interaction assays using protein panels (Support Protocol). Combining aptamer library screening with interaction trap hunts and mating interaction assays provides a new strategy for analyzing processes in diploid organisms and in multicopy gene phenotypes.

Peptide aptamers have been used for the forward analysis of phenotypes in yeast (Caponigro et al., 1998; Geyer et al., 1999; Norman et al., 1999) and bacteria (Blum et al., 2000). In yeast, peptide aptamers were isolated that inhibited mating pheromone response (Caponigro et al., 1998; Geyer et al., 1999; Norman et al., 1999) and spindle checkpoint (Norman et al., 1999) signal transduction pathways. In bacteria, peptide aptamers were isolated that specifically inhibited thymidylate synthase or that caused growth inhibition (Blum et al., 2000). The peptide aptamer targets for forward analysis in yeast were identified using yeast two-hybrid systems. Mating interaction assays identified protein targets from panels of proteins known to be involved in the yeast pheromone response pathway (Caponigro et al., 1998; Geyer et al., 1999) or from large

Peptide Aptamers

24.4.20

panels of proteins containing almost all of the proteins in the yeast genome (Norman et al., 1999). Peptide aptamer targets were also identified using interaction trap hunts against a partial-coverage yeast genomic library (Geyer et al., 1999). Interestingly, the peptide aptamer targets identified with the mating interaction assay were not obtained with the interaction trap hunt using the partial-coverage yeast genomic library (Geyer et al., 1999). The inability of the genomic library screen to identify aptamer targets is partly due to the representation of targets in the partial-coverage library. Nevertheless, the results demonstrate a better success rate for identifying aptamer targets using mating interaction assays with arrayed panels of protein targets. Mating interaction assays have the following advantages: (1) they present protein targets as fully normalized libraries, (2) they allow reporter outputs that result from interactions to be directly compared with outputs caused by the bait alone, and (3) they allow the detection of interaction strengths independent of the differences in plating efficiencies caused by differential reporter activation (Estojak et al., 1995). Currently, protein panels that cover an organism's entire proteome are not commercially available. Consequently, the identification of targets for peptide aptamers isolated using genetic screens will consist of limited panels of known proteins complemented with cDNA or genomic libraries.

Inhibitory mechanisms of peptide aptamers

Peptide aptamers inhibit protein function by a variety of mechanisms. For example, peptide aptamers can bind to protein targets and disrupt their interactions with other proteins. They can disrupt protein interactions within cells (Xu et al., pers. comm.) and in two-hybrid assays (Geyer et al., 1999), and they can inhibit enzymes by competing with their substrates for active site binding (Cohen et al., 1998). In addition to disrupting protein interactions, peptide aptamers can also inhibit protein function by mislocalizing protein targets. Peptide aptamers modified with a localization signal can transport their target proteins into various cellular compartments (Colas et al., 2000). Peptide aptamers fused to catalytic domains can also direct the substrate specificity of enzymes. They can be used to localize enzyme activities to specific protein targets (Colas et al., 2000) or locations in the cell.

Peptide aptamers are particularly useful for the analysis of genetic networks since they can disrupt specific interactions with protein targets

that have multiple protein interactions (Geyer et al., 1999). This allows phenotypes caused by the disruption of individual interactions in a network to be observed, while leaving other interactions in the same network intact. Peptide aptamers can be isolated against allelic variants of proteins (Xu et al., 1997). Their high specificity can be used to functionally characterize variants of polymorphic proteins. In addition, controlling their expression using inducible promoters allows the penetrance and timing of the aptamer-induced phenotype to be varied. Finally, performing genetic selections with peptide aptamers targeted to different locations in the cell can provide information on the cellular location of the target protein.

Together, these properties point to the many ways in which peptide aptamers can be used to analyze cellular processes. The successful use of peptide aptamers in the reverse analysis of processes in cell cultures and in *Drosophila*, and in the forward analysis of processes in yeast, illustrates their potential as "genetic" agents in the analysis of genetically intractable organisms.

Critical Parameters and Troubleshooting

Peptide aptamer libraries

The first critical parameter to consider is the method for synthesizing peptide aptamer libraries. Preferably, peptide aptamer libraries are constructed to minimize the amount of stop codons while maintaining amino acid diversity. In general, two methods of automated DNA synthesis are used to generate DNA templates that code for combinatorial peptide libraries. The first method generates DNA templates by sequentially coupling mixtures of the four-nucleotide phosphoramidites. The second method generates DNA templates by sequentially coupling mixtures of codons.

The sequential nucleotide incorporation method uses completely random or biased mixtures of nucleosides to construct DNA templates. DNA templates constructed using equimolar mixtures of the four-nucleotide phosphoramidites contain all 64 possible codons, including 41 redundant codons and three stop codons. The completely random libraries are biased for amino acids encoded by multiple codons. In addition, the presence of stop codons produces truncated aptamers at a frequency of $3n/64$, where n is the length of the peptide library. The sequential nucleotide incorporation method is improved by restricting the nu-

cleotides that are incorporated at the third position in the codon (see Table 24.4.1 for examples of degenerate codons). The third position of a codon is responsible for most of the redundancy in the genetic code. DNA templates that contain all four nucleosides in the first two positions of the codon and only G or C at the third position consist of 32 codons, which code for 20 amino acids and one stop codon. Codons limited to G or C at the third position are biased for amino acids that are coded by multiple codons. However, the frequency of a stop codon is reduced to $n/32$, where n is the length of the peptide. The presence of stop codons in a completely random or third position–biased library limits the complexity that is obtainable with long combinatorial peptide libraries. The construction of longer peptide libraries requires the ligation of shorter DNA templates that are pre-screened to eliminate sequences that contain stop codons (Cho et al., 2000). Alternatively, combinatorial peptide libraries can be constructed that contain no stop codons, but with reduced amino acid diversity. Table 24.4.1 provides examples of degenerate codons that can be used to design peptide libraries.

The sequential codon incorporation method is used to generate DNA templates that contain 20 amino acids and no stop codons. Three strategies are used to generate codons. The first strategy involves sequentially coupling individual nucleotide phosphoramidites to generate 20 codons each of which is on a separate column (Lam et al., 1991). The beads from each column are subsequently mixed together and repacked into new columns for the synthesis of the next codon. The second strategy involves the synthesis of 20 trinucleotide phosphoramidite codons (Virnekas et al., 1994). Combinatorial peptide libraries are synthesized by coupling random or biased mixtures of the codon phosphoramidites. The third strategy combines aspects of the first two strategies and involves sequentially coupling either an A, G, C, or T phosphoramidite followed by a specific dinucleotide phosphoramidite to complete the codon (Neuner et al., 1998). After the completion of each codon, the beads from the columns are mixed and repacked into new columns for the synthesis of the next codon. The advantage of the codon incorporation method is that it generates unbiased libraries without stop codons. However, there are drawbacks to this method. For example, the bead splitting can become extremely laborious for long peptides. Also, the synthesis of dinucleotide and trinucleotide phosphoramidites is not trivial, and

these phosphoramidites are not currently commercially available.

Once the combinatorial peptide libraries are constructed and inserted into the scaffold protein, they need to be transformed into *E. coli* and amplified. Electroporation is the most efficient method for transforming high-diversity libraries into *E. coli*. DNA uptake by *E. coli* is maximized under conditions of high field strength and low current flow (see Sidhu and Weiss, 2000, for conditions to maximize transformation efficiency in *E. coli*). To reduce the current flow, the conducting species must be removed from the DNA using affinity purification columns.

The number of peptide aptamers that can be screened is generally limited by the transformation efficiencies of the organism used in the selection. In yeast, the highest transformation efficiencies are obtained using the lithium acetate transformation protocol developed by Geitz and Schiestl (1995). The diversity of peptide aptamer libraries in yeast are limited to $\sim 10^6$ to 10^7 unique aptamers. This is much lower than the 10^9 to 10^{10} libraries typically obtained in *E. coli*. Particular care should be taken to optimize the transformation efficiencies in yeast or other selected organisms. To obtain optimal transformation in yeast, it is important to perform trial transformations to optimize parameters such as heat shock time and cell density.

Screening peptide aptamers

A second critical parameter is the spontaneous reversion rate in the screen used to isolate the peptide aptamers. *UNIT 20.1* discusses critical parameters that should be taken into account when selecting peptide aptamers against specific proteins using the interaction trap. False positives that occur in either the interaction trap or other genetic screens can be eliminated more efficiently using peptide aptamers that are expressed under the control of an inducible promoter.

Identifying protein targets

A third critical parameter is the identification of proteins targeted by peptide aptamers that have been isolated based on their ability to disrupt cellular processes. In general, peptide aptamer targets are more reliably obtained from panels of known proteins rather than from genomic or cDNA libraries. Once putative peptide aptamer targets have been identified using interaction trap hunts and mating interaction assays, it is important to verify these targets

using other means. For example, immunoprecipitation can be used to confirm that aptamers form complexes with their targets under *in vivo* conditions. Genetic tests such as epistasis analysis can be used to identify the location of the aptamers relative to a known protein. Peptide aptamer targets can be deleted or overexpressed and the resulting phenotype compared to the aptamer-induced phenotype. Similarly, whole-genome transcript arrays can test whether aptamers cause the same response as known inhibitors or mutations.

Anticipated Results

In general, approximately one out of every 10^5 peptide aptamers screened using the interaction trap interacts with a given target protein (Colas et al., 1996; Xu et al., 1997; Fabbri et al., 1999). Based on results using the yeast pheromone response pathway as a model process, approximately one out of every 10^5 to 10^6 peptide aptamers can inhibit a cellular process (Geyer et al., 1999). These results apply to 20mer combinatorial peptide libraries displayed on the surface of *E. coli* thioredoxin.

Time Considerations

Basic Protocol 1: Construction of the thioredoxin peptide aptamer library and its subsequent electroporation and amplification in *E. coli* will take ~1 week.

Basic Protocol 2: Isolation of peptide aptamers that interact with a specific bait protein takes ~3-4 weeks. The bait plasmid (pBait) and the pJM-1 peptide aptamer library are constructed during the first week. During the second week the pBait and the *lacZ* reporter plasmid (pSH18-34) are transformed into EGY48. The bait protein is also assayed to determine if it self-activates the reporter genes. During the third week the peptide aptamer library is transformed into EGY48 that contains pBait and pSH18-34. The aptamers are screened for their ability to interact with the bait protein and putative interacting aptamers are obtained. A fourth week is required to isolate the aptamer plasmids from the yeast and sequence their variable regions.

Basic Protocol 3: Determination of the peptide aptamer specificity using interaction mating takes ~1-2 weeks. The time required to construct the bait proteins, which will be used to evaluate the aptamer specificity, varies depending on the number of baits chosen and difficulty in cloning the baits. Once the bait proteins are constructed it takes ~1 week to transform both the baits and *lacZ* reporter into

EGY42 and the peptide aptamer preys into EGY48. Mating EGY48 with EGY42 and scoring interactions between the peptide aptamer preys and baits takes an additional week.

Basic Protocol 4: Affinity maturation of peptide aptamers takes ~3-4 weeks. The mutagenesis of the peptide aptamer and subsequent cloning into pJM-1 (prey vector) takes ~1 week. Isolation of mutant aptamers that interact with the bait protein using the interaction trap with a more stringent *lacZ* reporter takes 3 weeks as described above in Basic Protocol 2.

Basic Protocol 5: Construction of the thioredoxin peptide aptamer library (pJM-2 or pJM-3) takes ~1 week as described in Basic Protocol 1. The time required to isolate peptide aptamers that disrupt a cellular process varies depending on the organism and selection or screen used. Before the targets of the peptide aptamers can be identified, it is necessary to transfer the thioredoxin peptide aptamers from the expression vector used in the screen (pJM-2 or pJM-3) to the bait plasmid (pEG202). This transfer takes ~1 week. Identification of the peptide aptamer target(s) using the interaction trap mating (Basic Protocol 2) or cDNA or genomic library (*UNIT 20.1*) hunts takes ~4 weeks.

References

- Bai, C. and Elledge, S.J.M. 1996. Gene identification using the yeast two-hybrid system. *Methods Enzymol.* 273:331-347.
- Blum, J.H., Dove, S.L., Hochschild, A., and Mekalanos, J.J. 2000. Isolation of peptide aptamers that inhibit intracellular processes. *Proc. Natl. Acad. Sci. U.S.A.* 97:2241-2246.
- Bramlage, B., Luzi, E., and Eckstein, F. 1998. Designing ribozymes for the inhibition of gene expression. *Trends Biotech.* 16:434-438.
- Branch, A.D. 1998. A good antisense molecule is hard to find. *Trends Biochem. Sci.* 23:45-50.
- Brent, R. and Ptashne, M. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* 43:729-736.
- Cadwell, R.C. and Joyce, G.F. 1994. Mutagenic PCR. *PCR Methods Appl.* 3:S136-S140.
- Caponigro, G., Abedi, M.R., Hurlburt, A.P., Maxfield, A., Judd, W., and Kamb, A. 1998. Transdominant genetic analysis of a growth control pathway. *Proc. Natl. Acad. Sci. U.S.A.* 95:7508-7513.
- Chien, C.T., Bartel, P.L., Sternglanz, R., and Fields, S. 1991. The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci. U.S.A.* 88:9578-9582.

- Cho, G., Keefe, A.D., Liu, R., Wilson, D.S., and Szostak, J.W. 2000. Constructing high complexity synthetic libraries of long ORFs using in vitro selection. *J. Mol. Biol.* 297:309-391.
- Cohen, B. 1998. Selection of peptide aptamers that recognize and inhibit intracellular proteins. Ph.D. Thesis, Harvard University.
- Cohen, B.A., Colas, P., and Brent, R. 1998. An artificial cell-cycle inhibitor isolated from a combinatorial library. *Proc. Natl. Acad. Sci. U.S.A.* 95:14272-14277.
- Colas, P., Cohen, B., Jessen, T., Grishina, I., McCoy, J., and Brent, R. 1996. Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2. *Nature* 380:548-550.
- Colas, P., Cohen, B., Ferrigno, P., Silver, P., and Brent, R. 2000. Targeted modification and transportation of cellular proteins. *Proc. Natl. Acad. Sci. U.S.A.* In press.
- Dalton, S. and Treisman, R. 1992. Characterization of SAP-1, a protein recruited by serum response factor to the C-FOS serum response element. *Cell* 68:597-612.
- Davidson, A.R. and Sauer, R.T. 1994. Folded protein sequences occur frequently in libraries of random amino-acid-sequences. *Proc. Natl. Acad. Sci. U.S.A.* 91:2146-2150.
- Durfee, T., Becherer, K., Chen, P.L., Yeh, S.H., Yang, Y., Kilburn, A.E., Lee, W.H. and Elledge, S.J. 1993. The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes Dev.* 7:555-569.
- Ellington, A.D. and Szostak, J.W. 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature* 346:818-822.
- Estojak, J., Brent, R., and Golemis, E.A. 1995. Correlation of two-hybrid affinity data with in vitro measurements. *Mol. Cell. Biol.* 15:5820-5829.
- Fabrizio, E., Le Cam, L., Polanowski, J., Kaczorek, M., Lamb, N., Brent, R., and Sardet, C. 1999. Inhibition of mammalian cell proliferation by genetically selected peptide aptamers that functionally antagonize E2F activity. *Oncogene* 18:4357-4363.
- Finley, R.L. Jr. and Brent, R. 1994. Interaction mapping reveals binary and ternary connections between Drosophila cell cycle regulators. *Proc. Natl. Acad. Sci. U.S.A.* 91:12980-12984.
- Finley, R.L. and Brent, R. 1997. Understanding gene and allele function with two-hybrid methods. *Annu. Rev. Genet.* 31:663-704.
- Geyer, C.R. and Brent, R. 2000. Selection of "genetic" agents from random peptide aptamer expression libraries. *Methods Enzymol.* 328:171-208.
- Geyer, C.R., Colman-Lerner, A., and Brent, R. 1999. "Mutagenesis" by peptide aptamers identifies genetic network members and pathway connections. *Proc. Natl. Acad. Sci. U.S.A.* 96:8567-8572.
- Gietz, R.D. and Schiestl, R.H. 1995. Transforming yeast with DNA. *Methods Mol. Cell. Biol.* 5:255-269.
- Gorbsky, G.J., Chen, R.H. and Murray, A.W. 1998. Microinjection of antibody to Mad2 protein into mammalian cells in mitosis induces premature anaphase. *J. Cell Biol.* 141:1193-1205.
- Gyuris, J., Golemis, E., Chertkov, H. and Brent, R. 1993. Cdi1, a human G1- and S-phase protein phosphatase that associates with Cdk2. *Cell* 75:791-803.
- Herskowitz, I. 1987. Functional inactivation of genes by dominant negative mutations. *Nature* 329:219-222.
- Kamens, J. and Brent, R. 1991. A yeast transcription assay defines distinct REL and Dorsal DNA recognition sequences. *New Biol.* 3:1005-1013.
- Katti, S.K., LeMaster, D.M. and Eklund, H. 1990. Crystal structure of thioredoxin from Escherichia coli at 1.68 Å resolution. *J. Mol. Biol.* 212:167-184.
- Kolonin, M.G. and Finley, R.L. Jr. 1998. Targeting cyclin-dependent kinases in Drosophila with peptide aptamers. *Proc. Natl. Acad. Sci. U.S.A.* 95:14266-14271.
- Lam, K.S., Salmon, S.E., Hersch, E.N., Hraby, V.J., Katzmierski, W.M., and Knapp, R.J. 1991. A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* 354:82-84.
- LaVallie, E.R., Diblasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., and McCoy, J.M. 1993. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm. *BioTechnology* 11:187-193.
- Lu, Z., Murray, K.S., Van Cleave, V., LaVallie, E.R., Stahl, M.L., and McCoy, J.M. 1995. Expression of thioredoxin random peptide libraries on the Escherichia-Coli cell-surface as functional fusions to flagellin – a system designed for exploring protein-protein interactions. *Biotechnology* 13:366-372.
- Mitchison, T.J. 1994. Towards a pharmacological genetics. *Chem. Biol.* 1:3-6.
- Neuner, P., Cortese, R., and Monaci, P. 1998. Codon-based mutagenesis using dimer-phosphoramidites. *Nucl. Acids Res.* 26:1223-1227.
- Norman, T.C., Smith, D.L., Sorger, P.K., Drees, B.L., O'Rourke, S.M., Hughes, T.R., Roberts, C.J., Friend, S.H., Fields, S., and Murray, A.W. 1999. Genetic selection of peptide inhibitors of biological pathways. *Science* 285:591-595.
- Scott, J.K. and Smith, G.P. 1990. Searching for peptide ligands with an epitope library. *Science* 249:386-390.
- Sidhu, S.S. and Weiss, G.A. 2000. Constructing phage display libraries by oligonucleotide-directed mutagenesis. In *Phage Display: A Practical Approach* (T. Clackson and H.B. Lowman, eds.) In press. Oxford University Press, Oxford.

- Thomas, M., Chedin, S., Carles, C., Riva, M., Formulaok, M., and Sentenac, A. 1997. Selective targeting and inhibition of yeast RNA polymerase II by RNA aptamers. *J. Biol. Chem.* 272:27980-27986.
- Virnekas, B., Ge, L., Pluckthun, A., Schneider, K.C., Wellnhofer, G., and Moroney, S.E. 1994. Trinucleotide phosphoramidites—Ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis. *Nucl. Acids Res.* 23:5600-5607.
- Vojtek, A.B., Hollenberg, S.M., and Cooper, J.A. 1993. Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell* 74:205-214.
- West, R.W. Jr., Yocum, R.R. and Ptashne, M. 1984. Saccharomyces cerevisiae GAL1-GAL10 divergent promoter region: Location and function of the upstream activator sequence UASG. *Mol. Cell. Biol.* 4:2467-2478.
- Wetterauer, B., Veron, M., Miginiac-Maslow, M., Decottignies, P., and Jacquot, J.P. 1992. Biochemical characterization of thioredoxin-1 from *Dictyostelium discoideum*. *Eur. J. Biochem.* 209:643-649.
- Xu, C.W., Mendelsohn, A. and Brent, R. 1997. Cells that register logical relationships among proteins. *Proc. Natl. Acad. Sci. U.S.A.* 94:12473-12478.
- Yang, M., Wu, Z., and Fields, S. 1995. Protein-peptide interactions analyzed with the yeast two-hybrid system. *Nucl. Acids Res.* 23:1152-1156.
- Yocum, R.R., Hanley, S., West, R.J., and Ptashne, M. 1984. Use of LacZ fusions to delimit regulatory elements of the inducible divergent GAL1-GAL10 promoter in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 4:1985-1998.

Key References

- Colas et al., 1996. See above.
First article to describe the interaction trap as a method to isolate thioredoxin peptide aptamers against a specific protein (Cdk2).
- Geyer et al., 1999. See above.
Describes the use of thioredoxin peptide aptamers for the forward analysis of the pheromone response pathway in yeast. Peptide aptamers were isolated

that disrupt the pathway. Peptide aptamer targets were identified using mating interaction assays that contained panels of known proteins and by using interaction trap hunts against a yeast genomic library.

- Gyuris et al., 1993. See above.
Initial description of the interaction trap.
- Finley and Brent, 1994. See above.
Initial description of the mating interaction assay.
- Kolonin and Finley, 1998. See above.
Describes the reverse analysis of a cellular process in Drosophila using peptide aptamers that bind to Drosophila Cdk's.
- Lu et al., 1995. See above.
First study to use E. coli thioredoxin as a scaffold for displaying combinatorial libraries of peptides.
- Norman et al., 1999. See above.
Describes the use of staphylococcal nuclease peptide aptamers for the forward analysis of the yeast pheromone response and the spindle checkpoint signal transduction pathways. Peptide aptamers are characterized by transcript arrays and by two-hybrid analysis using a protein panel containing almost all of the proteins in the yeast genome.
- Sidhu and Weiss, 2000. See above
Review article that describes strategies for designing combinatorial peptide libraries and efficient methods for transforming E. coli.

Internet Resources

- <http://www.umanitoba.ca/faculties/medicine/units/biochem/gietz/Trafo.html>
Web site that describes efficient protocols for transforming yeast.
- See UNIT 20.1 for Internet resources related to the interaction trap.

Contributed by C. Ronald Geyer
 University of Florida
 Gainesville, Florida

Protein Selection Using mRNA Display

mRNA display is an *in vitro* technique that may be used to search natural or synthetic DNA libraries for the functional proteins and peptides they encode. mRNA-displayed proteins are constructs in which a protein is covalently attached to the RNA that encodes it. This direct covalent association of phenotype (protein) and genotype (RNA) renders the protein directly amplifiable. This in turn allows successive cycles of selection, enrichment, and, optionally, mutagenesis, to be performed upon libraries of displayed proteins. At the end of this process, functional sequences will dominate the library; cloning and sequencing will reveal the identity of the selected functional proteins. mRNA display allows new functional proteins to be discovered without resorting to protein design.

mRNA-displayed proteins are generated by the *in vitro* translation of mRNA display templates which are mRNA molecules 3'-terminated in puromycin (Fig. 24.5.1). Puromycin is a translation inhibitor that is able to enter the ribosome during translation and form a stable covalent bond with the nascent protein. This allows a stable covalent linkage to be formed between the mRNA display template and the protein it encodes, resulting in an mRNA-displayed protein (Fig. 24.5.2).

STRATEGIC PLANNING

The first issue that needs to be addressed when embarking upon protein selection using mRNA display is the design and construction of the library at the DNA level. If the goal of the selection is the "improvement" of an existing protein aptamer or enzyme, then the starting point for the selection will be the DNA sequence encoding this protein. If the goal of the selection is to discover a new class of protein aptamers or enzymes, then the starting point for the selection will be a DNA sequence in which some or many of the positions

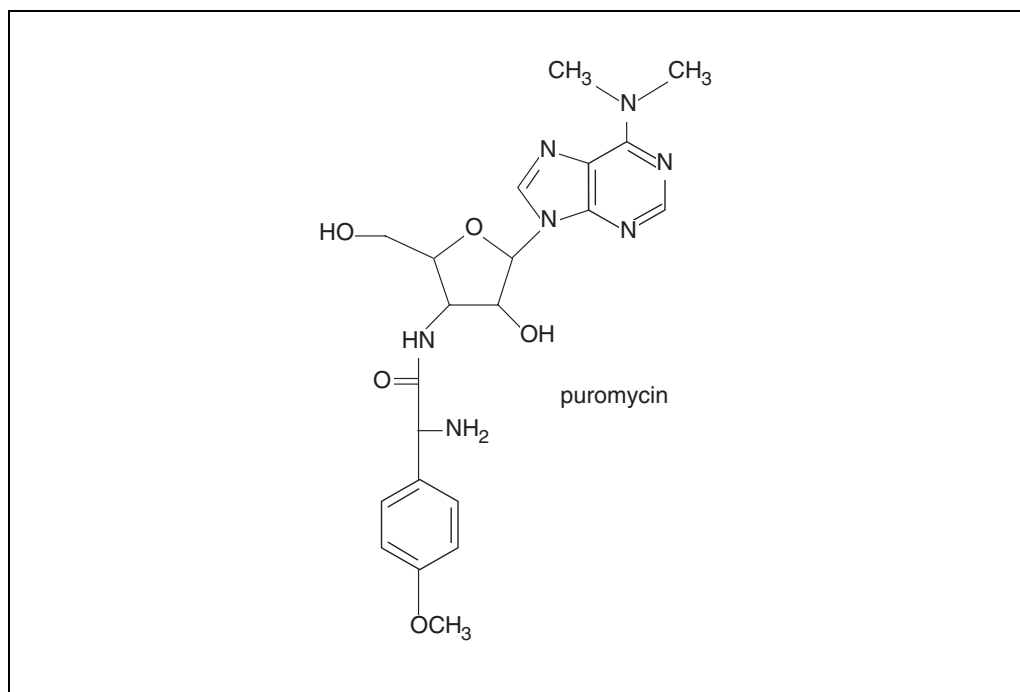


Figure 24.5.1 Puromycin is an antibiotic that functions by inhibiting translation. The molecular structure of puromycin resembles the acceptor arm of an amino-acylated tRNA. Puromycin is a nucleotide-amino acid chimera and ultimately forms the nucleic acid-protein junction in the mRNA displayed protein.

Contributed by Anthony D. Keefe
Current Protocols in Molecular Biology (2001) 24.5.1-24.5.34
Copyright © 2001 by John Wiley & Sons, Inc.

are randomized. In either case, the DNA library may originate from a fixed natural sequence (clone) and subsequently be randomized by some process such as mutagenic PCR (Cadwell and Joyce, 1992) or DNA shuffling (Stemmer, 1994). Alternatively, the DNA library may be synthetic, in which case it can be synthesized as a fixed sequence and treated as above, or it may be synthesized in a high-diversity form directly, using mixtures of nucleotide phosphoramidites on a DNA synthesizer. In a third approach, the DNA may be isolated from a natural high-diversity source, either cDNA derived from biological mRNA or genomic DNA using a diverse mixture of PCR primers, or DNA sampled directly from the environment originating from a multitude of uncultured organisms. In all of these approaches, the DNA library will ultimately need to encode terminal constant regions to permit PCR amplification. These may also encode protein affinity tags that will facilitate the purification of the resulting displayed proteins. It is also desirable to encode restriction sites close to the random-constant sequence boundary to enable these constant regions to be changed should reengineering of the library be necessary.

If the DNA library is to be synthesized in more than one piece, then a strategy of restriction and ligation of different DNA cassettes needs to be designed. This strategy ultimately

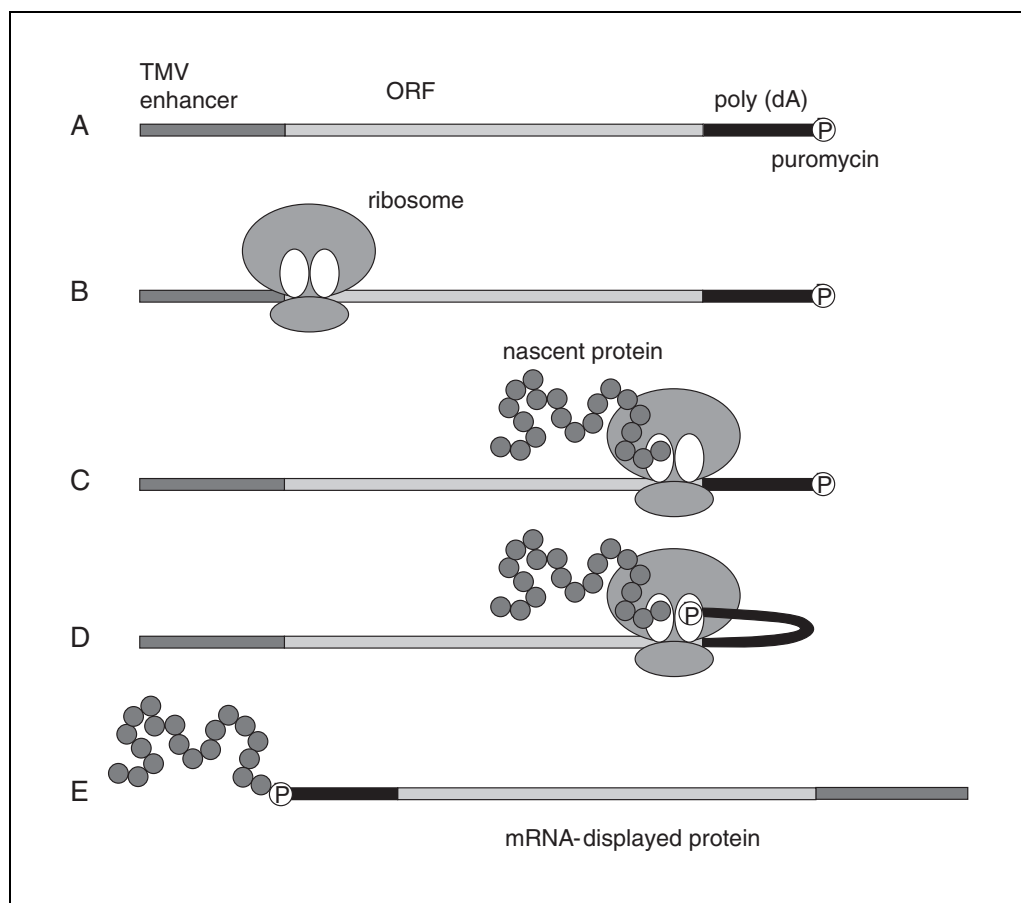


Figure 24.5.2 mRNA-displayed protein formation. **(A)** The mRNA display template consists of a Tobacco Mosaic Virus (TMV) translation enhancer sequence followed by the open reading frame encoded in RNA. This is followed by poly(dA) that is 3'-terminated with puromycin. **(B)** The ribosome initiating the translation of the mRNA display template. **(C)** The ribosome pausing at the RNA-DNA junction of the mRNA display template after it has translated the mRNA display template into protein. **(D)** The puromycin attached to the 3'-terminus of the mRNA display template entering the A site of the ribosome and forming a stable amide bond with the nascent protein. **(E)** The mRNA display template displaying the protein that it encodes after the ribosome has been released during purification.

yields the full-length library, and additionally offers the opportunities of purification and amplification of the individual DNA cassettes before they are ligated together. Amplification of the cassettes at this stage can greatly increase the library diversity in a combinatorial sense. Purification of the cassettes at this stage can decrease the proportion of cassettes that contain deletions, insertions, and stop codons. This “preselection” strategy can greatly increase the effective diversity of the DNA library since the proportion of resultant mRNA display templates, which are able to display frameshift-free proteins, will also increase in a combinatorial fashion once the DNA cassettes are ligated together. The steps in a preselection strategy are shown in Figure 24.5.3, and a detailed description is given in Cho et al. (2000).

Once the library has been designed and synthesized, the translation conditions need to be optimized for the formation of displayed proteins, and the purification strategy also needs to be optimized and subsequently piloted in a serial manner. However, the most important part of the strategic planning phase of the project is the design of the selection strategy. The first selection step needs to be designed to retain as many as possible of the functional displayed proteins that are present in the library, while discarding the great majority of those that are not functional. In this manner, the diversity of the library is taken advantage of to the maximal extent, but the diversity of the library is sufficiently reduced that the first amplification step is able to give several copies of the selected proteins for input into the next selection step. Subsequent to the first amplification step, the selection steps need to be designed to have the maximal possible reasonable discrimination between mRNA displayed proteins that exhibit the function of interest and those that do not. If at all possible, the extent of this discrimination should be assayed with positive and negative controls.

The steps within a single round of selection and amplification are shown in Figure 24.5.4.

Library Design

Library synthesis and preselection

A synthetic DNA library encoding a short open reading frame (ORF; up to ~35 amino acids) may be synthesized as a single oligonucleotide. Longer libraries of ORFs will need to be synthesized in two or more DNA cassettes that are then ligated together. Synthetic DNA has a deletion rate of ~0.5% and random regions will contain stop codons. Deletions will cause parts of the resultant proteins to be out of frame, and stop codons will prevent translated proteins from being displayed upon the mRNA that encodes them. For example, if the target ORF is 100 amino acids long, has an equal distribution of all four nucleotides at every position, and the deletion rate is 0.5%, only 0.18% will be in-frame over their entire lengths and free of stop codons. Consequently, unless the ORF is very short, one may wish to preselect the individual cassettes for being in-frame and free of stop codons. This preselection strategy is most easily accomplished by encoding different protein affinity tags close to the 3'- and 5'-termini of the cassettes. Synthesizing mRNA-displayed proteins from each individual DNA cassette, and purifying these upon the basis of the presence of each of these tags, will enrich the resultant library in those sequences that have initiated before the 5' tag, terminated after the 3' tag, and do not contain stop codons. These are likely to be in-frame over their entire sequence. The full-length DNA library should then be constructed from these preselected cassettes using RT-PCR followed by restriction and ligation. Any reduction in diversity that results from the preselection process is regained by the combinatorial ligation of the amplified DNA cassettes during the assembly of the full-length library. The steps in a preselection strategy are shown in Figure 24.5.3, and a detailed description is given in Cho et al. (2000).

Step	Product	
Synthesis and denaturing PAGE purification of DNA cassette(s), which when ligated together will encode the complete library	ssDNA cassette(s)	-----
PCR amplification of DNA cassette(s)	dsDNA cassette(s)	=====
Transcription of DNA cassette(s)	mRNA cassette(s)	
Denaturing PAGE purification of RNA cassette(s)	Purified mRNA cassette(s)	
Synthesis and denaturing PAGE of DNA splint that anneals to and aligns 3'-end of RNA and 5'-end of DNA linker to encourage ligation	DNA splint	-----
Synthesis and denaturing PAGE of DNA linker that terminates in puromycin	DNA linker	-----Ⓟ
Kinase (5'-phosphorylation) of DNA linker	Kinased DNA linker	-----Ⓟ
Splinted ligation of RNA cassette(s) to DNA linker terminated in puromycin	mRNA display template(s) of cassette(s)	
Translation of mRNA display template(s) of cassettes and high salt incubation and/or incubation at -20°C	Protein cassettes with both affinity tags displayed upon their stop codon-free and frameshift-free mRNA display templates	
	Misinitiated protein cassettes without an N-terminal affinity tag displayed upon partly untranslated mRNA display templates	
	Frameshifted protein cassettes without a C-terminal affinity tag displayed upon mRNA display templates with deletions	
	mRNA display templates with stop codons not displaying proteins and untranslated mRNA display templates	
	Free protein cassettes	

KEY:	
-----	DNA
△△△△△△△△△△	RNA
-----	protein
Ⓟ	puromycin

Figure 24.5.3 The steps that comprise the preselection process. A DNA cassette library that, when assembled into a full-length DNA library will encode a protein library, is enriched in those cassettes that are free of stop codons, insertions, and deletions. These cassettes are then used to construct the full-length library that is used for protein selection using mRNA display (*continues on next page*).

RNA polymerase promoter sequence

The library will need a transcription promoter at the 5'-end. This can be added or changed by PCR. The promoter sequences, TAATACGACTCACTATA and TTCTAATACGACTCACTATA, have both been successfully used. Transcription is most efficient if the RNA transcript starts with at least two guanines. To avoid pyrimidines (T or C) in the first few nucleotides of the transcript, it is common for the transcribed RNA sequence to commence with GGG.

Protein Selection Using mRNA Display

24.5.4

Step	Product	
Oligo(dT) cellulose purification	Protein cassettes with both affinity tags displayed upon their stop codon-free and frameshift-free mRNA display templates	
	Misinitiated protein cassettes without an N-terminal affinity tag displayed upon partly untranslated mRNA display templates	
	Frameshifted protein cassettes without a C-terminal affinity tag displayed upon mRNA display templates with deletions	
	mRNA display templates with stop codons not displaying proteins and untranslated mRNA display templates	
Reticulocyte lysate mRNA		
N-terminal protein affinity tag (such as FLAG) purification (optionally repeated)	Protein cassettes with both affinity tags displayed upon their stop codon-free and frameshift-free mRNA display templates	
	Frameshifted protein cassettes without a C-terminal affinity tag displayed upon mRNA display templates with deletions	
C-terminal protein affinity tag (such as His ₆) purification (optionally repeated)	Protein cassettes with both affinity tags displayed upon their stop codon-free and frameshift-free mRNA display templates	
Reverse transcription with the splint as primer	Protein cassettes with both affinity tags displayed upon their stop codon-free and frameshift-free and reverse transcribed mRNA display templates	
PCR amplification	dsDNA cassette(s) that are in-frame at both their 3' and 5'-ends and free of stop codons	
Restriction, native PAGE purification, ligation, native PAGE purification, optionally repeated one or more times	Full-length DNA library encoding an open reading frame substantially free of stop codons, deletions, and insertions	

KEY:
 ----- DNA protein
 ▲▲▲▲▲▲▲▲ RNA ⊕ puromycin

Figure 24.5.3 Continued.

Translation enhancer sequence

The library will need a translation enhancer before the initiating methionine codon; the truncated 5'-UTR from the Tobacco Mosaic Virus sequence (ACAATTACTATTCAATTACA) has been used successfully.

Initiating methionine

The initiating methionine (ATG) immediately follows the translation enhancer sequence.

N-terminal constant ORF sequence

It is extremely helpful to have amino acid sequences within the protein that can act as affinity tags. These are invaluable when purifying the displayed proteins. If two different affinity tags are used and these are located close to each of the termini of the expressed

Step	Product	
Synthesis and denaturing PAGE purification of full-length DNA library (skip this step if library is already dsDNA)	Full-length ssDNA library	-----
PCR amplification of DNA library (skip this step if library is already amplified)	dsDNA library	=====
Transcription of DNA library	mRNA library	ORF ΔΔΔΔΔΔΔΔΔΔΔΔΔΔΔΔ
Denaturing PAGE purification of mRNA library	Purified mRNA library	ORF ΔΔΔΔΔΔΔΔΔΔΔΔΔΔΔΔ
Synthesis and denaturing PAGE of DNA splint that anneals to and coaligns 3'-end of RNA and 5'-end of DNA linker to allow ligation	DNA splint	-----
Synthesis and denaturing PAGE of DNA linker that is 5'-terminated with puromycin	DNA linker	-----Ⓟ
Kinase (5'-phosphorylation) of DNA linker	Kinased DNA linker	-----Ⓟ
Splinted ligation of mRNA library to DNA linker	Library of mRNA display templates	ORF ΔΔΔΔΔΔΔΔΔΔΔΔΔΔΔΔ-----Ⓟ
Translation of mRNA display template and high salt incubation and/or incubation at -20°C	Protein library displayed upon mRNA display templates	ORF ΔΔΔΔΔΔΔΔΔΔΔΔΔΔΔΔ-----Ⓟ ----- -----
	Free proteins	-----
	Free mRNA display templates	ORF ΔΔΔΔΔΔΔΔΔΔΔΔΔΔΔΔ-----Ⓟ
	Reticulocyte lysate mRNA	-----
	Reticulocyte lysate	-----

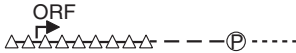

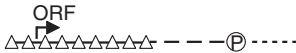
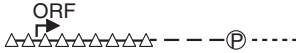

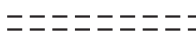
KEY:
 ----- DNA
 ΔΔΔΔΔΔΔΔΔΔ RNA
 ----- protein
 Ⓟ puromycin

Figure 24.5.4 The steps that comprise a single round of selection and amplification in a protein selection using mRNA display (*continues on next page*).

**Protein Selection
Using mRNA
Display**
24.5.6

protein, then they may be used to ensure that the expressed protein, mRNA-displayed or not, is full-length and in-frame at both ends. This double purification may optionally be performed when the DNA library is still at the individual cassette stage in order to increase the proportion of library members that are full-length, in-frame, and do not contain stop codons (see Fig. 24.5.3). FLAG and His₆-tag sequences are obvious choices.

Some constant amino acid sequence is likely to result from the ligation junctions used to construct libraries from synthetic DNA with long random regions; the identity of these

Step	Product	
Oligo(dT) cellulose purification	Protein library displayed upon mRNA display templates	
	Free mRNA display templates	
Protein affinity tag (such as FLAG or His ₆) purification	Reticulocyte lysate mRNA	
	Protein library displayed upon mRNA display templates	
Reverse transcription with the splint as primer	Protein library displayed upon reverse transcribed mRNA display templates	
Selection step	Selected members of the protein library displayed upon reverse transcribed mRNA display templates	
PCR amplification of selected fraction of library	Input dsDNA library for the next cycle of selection and amplification	
Repeat procedure from the start until the proportion of the library detected in the selected fraction has peaked or reached a plateau, at which point the library should be sequenced and individual members should be assayed for activity.	Individual functional proteins that pass the selection test with which they have been challenged	

KEY:
 - - - - - DNA
 ΔΔΔΔΔΔΔΔΔΔ RNA
 - - - - - protein
 ⊕ puromycin

Figure 24.5.4 Continued.

amino acids and their frame can be adjusted in order to avoid “inappropriate” amino acids such as several consecutive hydrophobic residues or a proline. What is considered to be inappropriate will depend upon what the library is to be used for. Regardless of how the library is to be constructed, having different restriction sites encoded within the 3’ and 5’ ends of the open reading frame will allow for the changing of one or other of the protein terminal constant sequences if design considerations change, or should reengineering be required for troubleshooting.

Long stretches of uridines in the RNA sequence should be avoided since these may anneal to the poly(dA) sequence of the puromycin-terminated linker oligonucleotide. This will interfere with the ligation used to construct the mRNA display template. Also, the double-stranded RNA-DNA, which can result from the self-annealing of the resulting

mRNA display template, will act as a substrate for the RNase H, which is present in reticulocyte lysate, resulting in degradation of the mRNA display template.

C-terminal constant ORF sequence

All of the considerations described with regard to the N-terminal constant ORF sequence also apply to the C-terminal constant ORF sequence. There are also additional considerations that are specific to the C-terminal sequence.

In the synthesis of the mRNA display template (see Fig. 24.5.4), the 3' terminus of the mRNA encoding the ORF is ligated to a short DNA oligonucleotide, which is itself 3'-terminated with puromycin (the linker). The 3' terminus of the mRNA and the 5' terminus of the DNA linker are coannealed to a short DNA oligonucleotide (the splint), which is complementary to both of them and presents the junction to be ligated as a nicked double-stranded nucleic acid. Consequently, the secondary structure of the splint, the mRNA, and the linker should be checked for self-structure likely to interfere with the assembly of the splinted nicked double-stranded complex. This can most easily be done using a computer algorithm such as MFOLD (<http://www.ibc.wustl.edu/~zucker/rna/form1.cgi>).

The mRNA-displayed protein is attached to the mRNA via its C terminus. It seems appropriate, therefore, to make the last few amino acids at the C terminus “structureless,” i.e., a stretch of glycines and serines. Incorporating extra methionines into the constant sequence will increase the signal resulting from the incorporation of ³⁵S-methionine into the protein. Extra methionines are best placed in the C-terminal constant region downstream of the C-terminal protein affinity tag. Should they result in misinitiation, then the resultant proteins will not contain either of the protein affinity tags and will not copurify with the mRNA-displayed full-length proteins. Placing out-of-frame stop codons close to the C terminus, either before or after the tag, in both the +1 and –1 frames will prevent those members of the library that are out-of-frame at the C terminus from forming mRNA-displayed proteins. This is especially useful in the context of a preselection (see Fig. 24.5.3). Incorporating a protein kinase (phosphorylation) site to allow ³²P-labeling of the protein may assist in assaying the free proteins.

Statistical appearance of different amino acids within the random region

Most DNA libraries designed for protein selection encode a wide range of amino acids in their random regions. Using a mixture of all four nucleotides at each of the three positions in the library codons will ensure that all 20 amino acids have some probability of appearing in every position of the resulting protein sequence. During the library design process it is helpful to consider the average composition of amino acids that will result from the chosen nucleotide distribution, the consequent average proportions of hydrophobic and charged amino acids, and whether these proportions are suitable for the library and its intended target. Additionally, it is useful to consider the frequency with which certain individual amino acids will appear in the random region. Cysteines can coordinate transition metal ions or form disulfide bonds, histidines can accept or donate protons or coordinate transition metal ions, and prolines may disrupt secondary structure. Other specific amino acids may be suitable for interacting with intended substrates.

Frequency of stop codon appearance in random libraries

Using a mixture of all four nucleotides at each of the three positions in the library codons in the DNA encoding the protein library will also introduce stop codons. This will reduce the proportion of expressed protein that is displayed because stop codons will cause the ribosome to release the mRNA before the terminating puromycin is able to react with the nascent peptide. By altering the proportions of the nucleotides in the DNA synthesis mixtures, the frequency of stop codons can be reduced, although this will also influence

the proportions of other amino acids in the library. Alternatively, the DNA cassettes used to construct the library can be synthesized from mixtures of nucleotides chosen only with regard to the average amino acid composition they encode. The resultant cassettes can then be “preselected” as described above in order to enrich the resultant library with those that do not contain stop codons. As shown in Fig. 24.5.3, if this procedure is done using two different affinity tags at the different termini of the ORF, then the resultant library will also be enriched in cassettes without deletions. A nucleotide distribution encoding a target amino acid composition can be iteratively approached using computer algorithms that are available on the Internet (e.g., <http://gaiberg.wi.mit.edu/cgi-bin/CombinatorialCodons>; Wolf and Kim, 1999).

Codon usage

Statistical studies of sequenced genomes have shown that, for the majority of amino acids for which more than one codon exists, certain codons appear more frequently than others. The more frequently used codons tend to end in G or C, which may relate to the extra stability that results from having three hydrogen bonds in the wobble position of the tRNA-mRNA complex. Consequently, it may be helpful to design the library so that G and C are the only nucleotides in the third position of each codon. The symmetry of the genetic code is such that the composition of the wobble position has very little effect upon the composition of the resultant protein, so this approach need not affect the amino acid composition of the protein library it encodes. However, this strategy will amplify the effect that frameshifts will have upon the resultant proteins.

Periodicity and stop codon avoidance

Some mixtures of nucleotides may result in the total omission of stop codons, $(VNN)_n$ for example (where V is a mixture of A, G, and C) does not encode any stop codons. Unfortunately, such approaches always result in the loss of some of the 20 amino acids; $(VNN)_n$ also does not encode Cys, Phe, Trp, and Tyr for example. Mixing different such codons together can give a DNA library that encodes all 20 amino acids but no stop codons. This approach, however, necessarily introduces an element of design into the library, since the statistically different codons must be placed at specified points in the sequence, most usually in a periodic fashion. Periodicity can also result in an increased tendency for protein structural units to be encoded; for example, alternate hydrophobic and hydrophilic amino acids will encourage the formation of β sheets, while alternate pairs of hydrophobic and hydrophilic amino acids will encourage the formation of α helices. Alternatively, there are nonperiodic nucleotide distributions that reduce the occurrence of stop codons to ~1%. For an example of this approach see LaBean and Kauffmann (1993).

Mutagenesis

Mutagenic procedures such as mutagenic PCR (Cadwell and Joyce, 1992) or DNA shuffling (Stemmer, 1994) may be used to generate a diverse DNA library from a less diverse DNA library or a single DNA sequence (a minimum of two homologous sequences is required for DNA shuffling). Mutagenic procedures may be used to generate the initial DNA library for an mRNA display protein selection, or to increase the diversity at any stage between cycles of selection and amplification during an mRNA display protein selection. In general, in vitro selection proceeds by the gradual loss of diversity as functional sequences are preferentially amplified and nonfunctional sequences are lost. Consequently, increasing the diversity of the library at a stage after the first selection step may appear to be a retrogressive step; however, this is not necessarily the case. Protein libraries that are generated by stochastic means, such as those generated from DNA made on a DNA synthesizer using mixtures of nucleotide phosphoramidites, sample protein sequence space extremely sparsely. For example, a member of a library of 10^{13} proteins

each of which is 72 amino acids long, with each amino acid being equally likely to appear at each position, will, on average, have a sequence with 26 differences (“mutations”) from the next most similar member of the library. If a solution to a particular problem is chanced upon using such a library, it is highly unlikely to be the optimal solution. The solution in question may be a small number of mutations away from many other superior solutions, but the initial library is extremely unlikely to have contained any of the sequences in question because the sampling was so sparse. Once a selection strategy has given one or many such nonoptimal solutions, one or more mutagenic steps will enable the exploration of the local sequence space around such solutions and, after subsequent selection, is likely to yield improved solutions closely related to one or more of the originally selected sequences.

Specific directions for the preparation of a DNA library for an mRNA display selection vary greatly depending upon the source of the DNA, the selection target, and the precise assembly strategy chosen. See *UNITS 24.2, 24.3, & 24.4* for further details. A more detailed discussion of both preselection and mRNA display library construction strategy is given in Cho et al. (2000). A generalized library construction strategy is also shown in Fig. 24.5.5. Once the DNA library has been synthesized, but before the selection has commenced, it is important to sequence some of the individual library members to ensure that the library sequence is as intended, and that the proportion of error-free sequences is appropriate for the selection strategy envisaged.

Selection

In vitro selection strategies, such as mRNA display, offer a generalized method for the discovery of functional molecules, only if the molecules in question can be enriched upon the basis of their function and subsequently directly amplified. Enrichment upon the basis of function is termed selection. The appropriate design of the selection step is absolutely crucial to the success of the project as a whole. Most in vitro selection experiments have been performed upon libraries of nucleic acids rather than proteins (Wilson and Szostak, 1999), although some protein selection experiments have been successfully performed using ribosome display (Jermutus et al., 1998) and most recently mRNA display (A.D. Keefe and J.W. Szostak, pers. commun.). Phage display (Smith and Petrenko, 1997) and the 2-hybrid system (Fields and Song, 1989; Colas et al., 1996) are similar in vivo techniques that have also been successfully used for the selection of functional peptides and proteins.

In vitro selections are divided into two main categories, (1) selections for aptamers (i.e., specific binding to a chosen target) and (2) selections for catalysts (i.e., enzymes). This is not an appropriate place for an exhaustive overview of the various approaches that have been used to discover new aptamers and catalysts, but some general points can be made.

Aptamer selections

Aptamer (specific binder) selections are in general undertaken by incubating the library with the immobilized target molecule. The target molecule immobilization is by way of covalent attachment to a solid matrix such as agarose, usually through a spacer molecule. Many immobilized target molecules are commercially available. After incubation, the flowthrough is drained away, the immobilized target molecules are washed several times, and then an elution fraction is collected by incubating the matrix and the immobilized target molecules with an elution buffer that contains the dissolved target molecule. Those library members that are contained in the elution fraction are amplified, and the process is repeated until functional molecules dominate the library. At this stage, the functional molecules are identified by cloning and sequencing. It is important to realize that in the early rounds of selection, the vast majority of the library members contained in the elution

fraction are nonspecific binders or nonbinders. Consequently, several rounds of selection and amplification will be required before the functional (specific binding) sequences dominate the library.

The composition of the selection binding and selection elution buffers is likely to influence the aptamers that are ultimately discovered using this system. It may be important to use a buffer that promotes protein folding, but discourages aggregation. The use of high concentrations of cosmotropic compounds such as ammonium sulfate will promote folding, while the use of nonionic detergents such as Triton X-100 will discourage aggregation. The oxidation potential of the buffer should also be considered. The inclusion of a reducing agent such as DTT is likely to lead to the selection of protein aptamers active under reducing conditions, while the inclusion of oxidizing agents such as glutathione disulfide is likely to lead to the selection of protein aptamers active under oxidizing conditions. It is important to ensure that the binding and elution buffers are as similar as possible. Changes in ionic strength and/or pH between these buffers will increase the proportion of nonspecific binders in the elution fraction, possibly to such an extent that the specific binders will never dominate the library and consequently will never be identified. Obviously the binding and elution buffers cannot be identical since the elution buffer contains the target molecule. In order to make the selection binding and

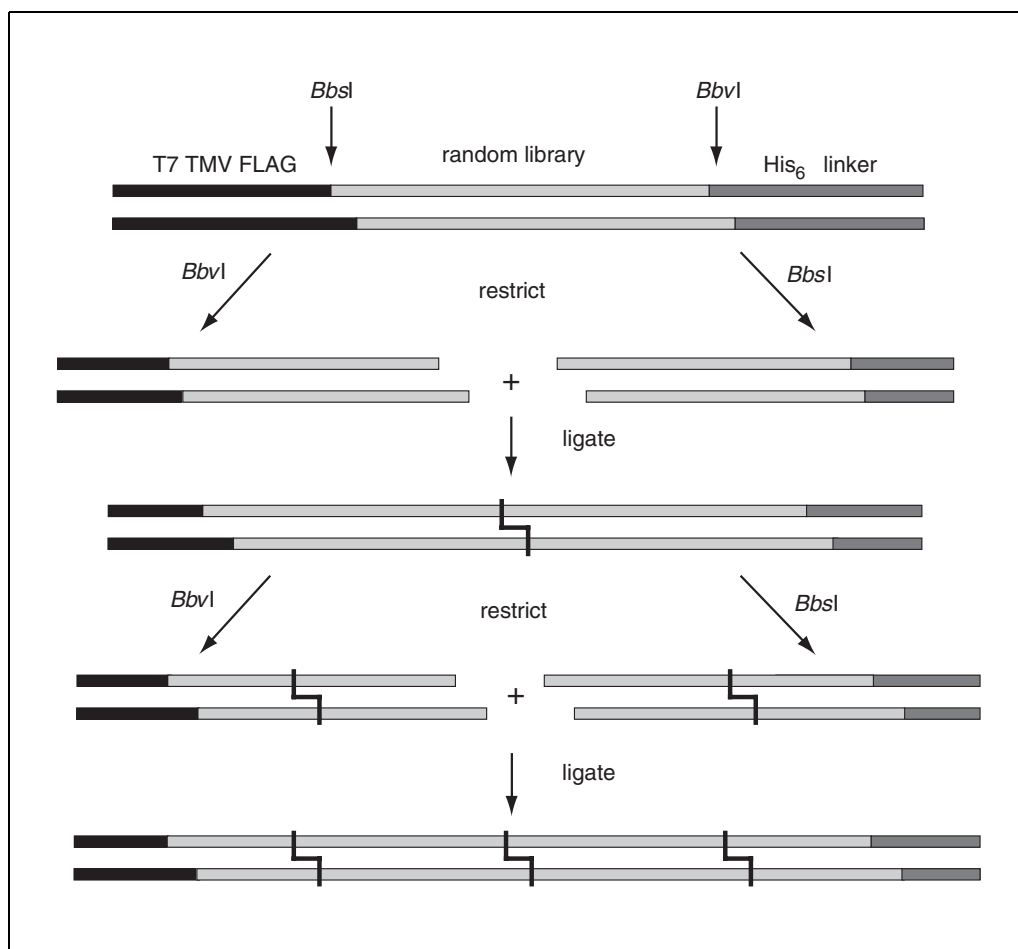


Figure 24.5.5 Assembly of a full-length mRNA display template library from DNA cassettes that result from the RT-PCR amplification of pre-selected mRNA display cassette templates. In this example, the cassettes are divided into 2 aliquots that are restricted with either *BbvI* or *BbsI*, subsequent ligation with T4 DNA ligase gives a new cassette in which the DNA between the restriction sites doubles in length while the flanking regions remain the same. The doubling of the length of this region may be repeated any number of times by repeating the restriction and ligation process.

selection elution buffers as similar as possible, it may be desirable to balance the effect of the target molecule upon the elution buffer by adding a similar molecule to the binding buffer. If the target molecule interacts with one of the buffer components, such as a nucleotide with magnesium, it may be desirable to add back an extra amount of this component to maintain the free concentration of the component in question at the concentration of that in the binding buffer. It is also possible to collect the elution fraction by disrupting the binding with denaturant or extremes of pH, although the background is likely to be higher.

Catalytic (enzyme) selections

Catalytic (enzyme) selections are a little more complex in a conceptual sense—although carefully designed, they can have lower intrinsic background rates than aptamer selections, and consequently be quicker and easier to perform in the laboratory. Enzyme selections must be performed in such a manner that those sequences which catalyze a reaction are separated from those that do not. The most obvious way to achieve this separation is to arrange the selection so that library members that catalyze the desired reaction covalently attach themselves to the substrate. If the substrate is in turn covalently attached to a tag (such as biotin), then the attachment of the tag to library members that catalyze the reaction can be used as a basis for the separation of these library members from those that do not catalyze the reaction (such as by binding to immobilized streptavidin). Alternatively the substrate may be immobilized before it is incubated with the library. Since both of these approaches effectively turn the catalysis selection into a binding selection, there will still be a background rate of isolation of sequences that do not catalyze the desired reaction. Consequently it may still be necessary to perform several rounds of selection and amplification before functional sequences dominate the library. Similar catalytic selection strategies can be envisaged in which all of the library members are immobilized, and those that successfully catalyze the desired reaction cut themselves free.

It should be noted that since the successfully selected library members are required to modify themselves in some respect, they are not acting as catalysts in the true sense of the word. However, molecules selected using such a procedure are usually easily reengineered to give true catalysts by detaching the active site part of the selected construct from the substrate part. One consequence of this limitation is that it is difficult to select for catalysts that act faster than the rate with which they can be manipulated in the laboratory, and it is not possible to select for catalysts with high turnover rates at all. Strategies in which the library member is encapsulated along with several substrate molecules may lead to systems in which the selective pressure is directly for the turnover rate.

Selection controls

The importance of selection controls cannot be emphasized strongly enough. mRNA display selection protocols in which functional library members are enriched much less than ten-fold over nonfunctional library members are unlikely to lead to the isolation of functional library members in the laboratory. Biases are present in many steps of the mRNA display amplification protocol, especially translation and protein display efficiencies, and these can overwhelm the enrichment in functional members that results from the selection step. Suitable positive controls are molecules known to catalyze or bind to the intended substrate, and need not be proteins, although the best control will usually be a similar functional protein displayed upon its reverse-transcribed mRNA display template.

Nomenclature

mRNA-displayed proteins are referred to by a variety of names in the literature, such as “RNA-protein fusions” and “profusions.”

PREPARATION AND PURIFICATION OF mRNA-DISPLAYED PROTEINS

This protocol describes the preparation of the mRNA display template from an appropriate DNA template, DNA splint, and DNA linker 3'-terminated with puromycin, the use of the mRNA display template to prepare mRNA-displayed proteins and their subsequent purification, and an example selection. The protocol steps are also shown in Figure 24.5.4. For additional details see Liu et al. (2000).

Materials

DNA library
1 M MgCl₂
100 mM nucleotide triphosphate solutions
10× transcription buffer (see recipe)
Deionized, ultrafiltered water
10 U/μl T7 RNA polymerase
Solid EDTA
Urea
0.5× TBE buffer (*APPENDIX 2*)
3 M NaCl
100% and 70% ethanol
100 mM EDTA
Puromycin-terminated DNA linker
100 mM ATP
T4 polynucleotide kinase buffer
T4 polynucleotide kinase
10× T4 DNA ligase buffer
10 U/μl T4 DNA ligase
3 M potassium acetate solution, pH 5.3
Rabbit reticulocyte lysate translation kit (e.g., Red Nova Lysate kit, Novagen)
Control RNA
12.5× methionine-free translation mix
2.5 M potassium chloride
25 mM magnesium acetate
Nuclease-free water
Rabbit reticulocyte lysate
³⁵S-methionine

Electroeluter (VWR or Schleicher & Schuell)
Denaturing PAGE gel (*UNIT 2.12*)
Gel filtration columns (Pharmacia)

Additional reagents and equipment for preparative denaturing PAGE purification (*UNIT 2.12*), determining nucleic acid concentration by spectrometry (*APPENDIX 3D*), synthesis of oligonucleotides (*UNIT 2.11*), and SDS-PAGE in Tris-tricine buffer systems (*UNIT 10.2A*)

BASIC PROTOCOL 1

Transcribe DNA

1. Make up a 1-ml transcription reaction on ice as follows. Add the T7 RNA polymerase last.

DNA library (add volume sufficient for 5 to 50 nM final concentration)
35 μ l 1 M $MgCl_2$
50 μ l each 100 mM nucleotide triphosphate (final 5 mM each NTP)
100 μ l 10 \times transcription buffer (final 1 \times)
Up to 980 μ l deionized, ultrafiltered water
20 μ l 10 U/ μ l T7 RNA polymerase (final 200 U/ml).

Incubate the transcription reaction for 3 to 16 hr at 37°C. Halt the reaction by cooling on ice, or by adding solid EDTA to a final concentration of 50 mM.

The size of the transcription reaction can be adjusted to give an appropriate amount of RNA, but care should be taken to ensure that the diversity of the DNA used is several times larger than the diversity of the displayed proteins that will ultimately result. The effect of varying the concentration of $MgCl_2$ should be explored in pilot transcriptions.

Purify RNA

2. Purify resultant RNA using denaturing PAGE (UNIT 2.12). Add solid urea to the transcription reaction to give a final concentration of 5 M and solid EDTA to give a final concentration of 50 mM, heat for 2 min at 90°C, and load onto a denaturing PAGE gel.
3. After the gel has been run, visualize by UV-shadowing and excise the band containing the purified RNA. Extract the RNA into 300 mM NaCl by passive elution or into 0.5 \times TBE buffer in an electroeluter according to the manufacturer's instructions.
4. Precipitate the RNA by adding 3 M NaCl (final concentration 300 mM) and 2.5 volumes of 100% ethanol. Cool for 20 min at -80°C or overnight at -20°C.
5. Centrifuge for 10 min at 12,000 \times g, 4°C. Decant the supernatant, wash the pellet with 70% ethanol, dry under reduced pressure, bring up to 0.5 ml with deionized, ultrafiltered water, and measure the concentration by UV-visible spectroscopy at 260 nm.

Further instructions may be found in UNIT 2.12.

This purification step separates truncated RNA molecules and PCR primers from the full-length RNA transcripts. It is important to remove the PCR primers from the transcribed RNA since they will inhibit the formation of the mRNA displayed proteins in the translation step.

Synthesize linker

6. Synthesize the linker, a DNA oligonucleotide 3'-terminated in puromycin, that is ~30 nucleotides long and "unstructured" (e.g., according to one of the following examples; see UNIT 2.11 for oligonucleotide synthesis):

Example a. AAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCP.

Poly(dA) is the most obvious choice.

Example b. AAAAAAAAAAAAAAAAAAAAAA999ACCP.

In example b, "9" is phosphoramidite spacer 9 (Glen Research) and "P" is puromycin, derived from CPG-puromycin (Glen Research). This linker may give a higher yield of displayed proteins.

Linkers much longer or shorter than 30 nucleotides will give greatly reduced yields of displayed proteins, or none at all.

The puromycin-terminated DNA oligonucleotide is gel purified by denaturing PAGE (UNIT 2.12), extracted from the gel, and precipitated as described earlier. Dissolve the DNA linker in deionized water and measure the concentration using UV-visible spectroscopy at 260 nm.

Each of the spacer 9 units result in the incorporation of a triethylene glycol phosphate ester; this adds extra flexibility to the region of the template close to the puromycin and may result in a higher proportion of the resultant mRNA display templates displaying protein.

7. Kinase (5'-phosphorylate) the DNA linker using polynucleotide kinase by making up the following 1-ml kinase reaction mixture:

300 μ l 100 μ M DNA linker (final 30 μ M)
10 μ l 100 mM ATP (final 1 mM)
100 μ l 10 \times T4 polynucleotide kinase buffer (final 1 \times)
490 μ l water
100 μ l 10 U/ μ l T4 polynucleotide kinase (final 200 U/ml).

8. Incubate the reaction mixture for 2 hr at 37°C, add 200 μ l of 100 mM EDTA, heat for 5 min at 90°C, and desalt on a gel-filtration column.

It is important to heat-denature the polynucleotide kinase to prevent it from acting in the subsequent ligation reaction. The size of the kinase reaction should be adjusted to give an appropriate amount of 5'-phosphorylated DNA linker.

Synthesize splint

9. Synthesize the splint, a DNA oligonucleotide with a sequence (reading from the 5' end) of ≥ 10 nucleotides complementary to the 3' end of the RNA library and ≥ 10 nucleotides complementary to the 5' end of the linker, usually T₁₀ (see UNIT 2.11 for oligonucleotide synthesis methods). Purify by denaturing PAGE (UNIT 2.12).
10. Extract DNA from gel as in step 3 and precipitate as in step 4.
11. Dissolve the purified splint in deionized water and measure the concentration using UV-visible spectroscopy at 260 nm (see step 5).

Prepare mRNA display template

12. Ligate the linker and RNA template with T4 DNA ligase in the presence of the splint to give the mRNA display template. Set up the following 1-ml ligation reaction:

100 μ l 100 μ M 5'-phosphorylated DNA linker (final 10 μ M)
100 μ l 100 μ M RNA library (final 10 μ M)
100 μ l 100 μ M splint (final 10 μ M)
580 μ l water.
13. Heat this mixture for 2 min at 95°C, then add 100 μ l of 10 \times T4 DNA ligase buffer (final 1 \times).
14. Vortex the resultant mixture and cool on ice for 10 minutes, allow to warm to room temperature, then add 20 μ l of 2000 U/ μ l T4 DNA ligase (final 40 U/ml).
15. Incubate the reaction for 20 min at room temperature. Add 150 μ l of 100 mM EDTA and 500 mg of solid urea, and heat for 5 min at 90°C.

Table 24.5.1 Translation Reactions for mRNA Display Proteins

Reagent	Final concentration	A	B	C	D	E	F
Control RNA	—	1 μ l	0 μ l	0 μ l	0 μ l	0 μ l	0 μ l
5 μ M mRNA display template	2/4/800 nM	0 μ l	0 μ l	0 μ l	1 μ l	2 μ l	4 μ l
5 μ M unligated RNA	400 nM	0 μ l	0 μ l	2 μ l	0 μ l	0 μ l	0 μ l
12.5 \times Met-free translation mix	1 \times	2 μ l	2 μ l	2 μ l	2 μ l	2 μ l	2 μ l
8.6 μ M labeled methionine	0.69 μ M	2 μ l	2 μ l	2 μ l	2 μ l	2 μ l	2 μ l
2.5 M KCl	100 mM	1 μ l	1 μ l	1 μ l	1 μ l	1 μ l	1 μ l
25 mM magnesium acetate	500 μ M	0.5 μ l	0.5 μ l	0.5 μ l	0.5 μ l	0.5 μ l	0.5 μ l
Water	—	8.5 μ l	9.5 μ l	7.5 μ l	8.5 μ l	7.5 μ l	5.5 μ l
2.5 \times rabbit reticulocyte lysate ^a	1 \times	10 μ l	10 μ l	10 μ l	10 μ l	10 μ l	10 μ l
Total		25 μ l	25 μ l	25 μ l	25 μ l	25 μ l	25 μ l

^aEnsure that the rabbit reticulocyte lysate is added last.

- Gel purify ligated mRNA display template by denaturing PAGE (UNIT 2.12), extract from gel (see step 3), and precipitate as in step 4, except use 3 M potassium acetate, pH 5.3, in place of 3 M sodium chloride.
- Dissolve the purified ligated mRNA display template in deionized water and measure the concentration by UV-visible spectroscopy at 260 nm.

If the template is <500 nucleotides long, it should be possible to resolve the ligated and unligated RNA on the PAGE gel, which will give some idea of the yield of the ligation reaction. Otherwise, the unresolved bands will have to be co-excised from the gel and optionally further purified using oligo(dT) cellulose as described below. It is important to perform this gel purification even if it is not possible to resolve the ligated and unligated RNA, since the presence of the splint in the translation reaction will greatly reduce the yield of displayed proteins, and RNase H in the reticulocyte lysate will cause degradation of the mRNA display template if it is annealed to the splint. It should be noted that this splinted RNA-DNA ligation is far less efficient than the ligation of sticky-ended pieces of DNA.

Translate mRNA display template and prepare mRNA displayed proteins

Before the mRNA display template is used for large-scale translation, a small-scale translation should be attempted alongside various control translations to aid the identification of the band on the protein gel that corresponds to the mRNA-displayed proteins.

- Set up the translation reactions in Table 24.5.1 on ice, adding the rabbit reticulocyte lysate last.
- Incubate for 1 hr at 30°C, then add 1.7 μ l of 1 M MgCl₂ and 7.8 μ l of 2.5 M KCl to each of the reactions and allow them to stand for 5 min at room temperature.

The reaction mixtures may be optionally stored for up to 1 week at -20°C at this point.

- Analyze the different translations using Tris-tricine SDS-PAGE as described in UNIT 10.2A, Alternate Protocol 1.

The SDS-PAGE analysis should show a number of bands in lane A, which is the control RNA supplied by the manufacturer; this is the positive control and demonstrates that the translation reaction was set up correctly. Lane B is the no-template control, and may show no bands or may show a band corresponding to tRNA charged with methionine; in either case it should show no bands with mobilities equal to those assigned to the free protein and to the displayed protein. Lane C should show a band of high mobility that can be

assigned to the free protein. Lanes D, E, and F should also show bands that can be assigned to the free protein, but also bands of much lower mobility that can be assigned to the mRNA-displayed protein. If the density of the band assigned to the displayed protein in F is of equal or lesser density to the equivalent band in E, then the mRNA display template is likely to be of high quality.

For better proof that the band assigned to the mRNA displayed proteins is done so correctly, add splint (final 1 μ M) and $MgCl_2$ (final 10 mM) to an aliquot of the translation mixture before the salt incubation, and incubate for 30 min at 37°C. RNase H within the lysate will cause the RNA part of the mRNA-displayed proteins to be digested and leave the protein displayed upon the DNA linker alone, consequently the original displayed protein band will disappear and a new band will appear of intermediate mobility between the mRNA displayed protein and the free protein.

21. Once the displayed proteins have been observed by SDS-PAGE, optimize the magnesium acetate and potassium chloride concentrations in the translation reaction. Perform a succession of translations in parallel with added magnesium acetate concentrations ranging from 0.5 mM to 2 mM and added potassium chloride concentrations ranging from 50 mM to 200 mM.

The relative proportions of the mRNA display templates that end up displaying proteins will be readily apparent when the samples are run out on a gel together, and the optimal concentrations of both magnesium acetate and potassium chloride can be chosen.

If a preselection procedure is being used to synthesize the full-length mRNA display template library, the translation magnesium acetate and potassium chloride concentrations will have to be optimized separately for each cassette used in the preselection protocol, and then again for the full-length library. Despite the fact that the 3'-terminal region of each of the cassettes and the full-length library are the same, the optimal magnesium acetate and potassium chloride concentrations for the formation of mRNA-displayed proteins are likely to be different.

22. Prepare a 1-ml translation reaction on ice as follows, adding the reticulocyte lysate last.

80 μ l 5 μ M mRNA display template (final 400 nM)
80 μ l 12.5 \times methionine-free translation mix (final 1 \times)
20 μ l 8.6 μ M [35 S]methionine (final 0.17 μ M)
2.5 M KCl (as optimized)
0.5 μ l magnesium acetate (as optimized)
Water to 600 μ l
400 μ l 2.5 \times rabbit reticulocyte lysate (final 1 \times)
Total, 1000 μ l.

23. Incubate for 1 hr at 30°C, then add 65 μ l of 1 M $MgCl_2$ and 235 μ l of 2.5 M KCl to each of the above reactions and allow them to stand for 5 min at room temperature.

The translation reaction mixtures may be optionally stored for up to a week at -20°C at this point.

One may wish to decrease the concentration of mRNA display template if there is concern about sequence-dependent bias in translation and protein display efficiencies affecting the distribution of different sequences in the library. As the concentration of the mRNA display template is reduced, the proportion of this template that ends up displaying protein increases, with a concomitant increase in the fidelity with which the mRNA library sequence distribution is represented in mRNA-displayed protein sequence distribution. This is an advisable precaution at all stages in which the library is of relatively low diversity.

**PURIFICATION AND REVERSE TRANSCRIPTION OF THE
mRNA-DISPLAYED PROTEINS**

It is extremely advisable to pilot each of the following protocol steps before attempting the large-scale treatment of translation reaction mixture containing mRNA-displayed proteins.

In order to separate the mRNA display templates that display proteins from those that do not, it is necessary to use a purification step upon the basis of a protein affinity tag. In this protocol the His₆ tag is used, although other protein affinity tags may be utilized.

Materials

- Oligo(dT) cellulose (Amersham Pharmacia Biotech)
- Oligo(dT) binding buffer (see recipe)
- 1.3-ml translation reaction mRNA displayed proteins (see Basic Protocol 1)
- Oligo(dT) wash buffer (see recipe)
- Ni-NTA agarose (Qiagen)
- Ni-NTA binding buffer (see recipe)
- 2-Mercaptoethanol
- Ni-NTA wash buffer 1 (see recipe)
- Ni-NTA wash buffer 2 (see recipe)
- Ni-NTA elution buffer (see recipe)
- 10 mg/ml salmon sperm DNA (Life Technologies)
- 1 mg/ml BSA
- 200 μM DNA splint
- 5× Superscript II reverse transcriptase buffer (NEB)
- 0.1 M DTT
- 30 μl (each) 25 mM deoxynucleotide triphosphates (final 0.5 mM)
- 200 U/ml Superscript II reverse transcriptase (NEB)
- 25 mM deoxynucleotide triphosphate solutions
- ATP-aptamer selection binding buffer (see recipe)
- ATP-aptamer selection elution buffer (see recipe)
- Chromatography columns (Bio-Rad)
- Gel filtration columns (e.g., NAP-5, Amersham Pharmacia Biotech)

For additional reagents and equipment for preparative denaturing PAGE purification (*UNIT 2.12*) and SDS-PAGE in Tris-tricine buffer systems (*UNIT 10.2A*)

Purify mRNA displayed proteins

1. Wash 20 mg of oligo(dT) cellulose repeatedly with deionized water in the chromatography column within which it will be used. Resuspend the cellulose several times and apply positive pressure to force it to drain rapidly. Finally, wash once with oligo(dT) binding buffer.

Oligo(dT) cellulose contains fine particulate matter that can drastically reduce the flow rate of aqueous solutions. These fine particles can also pass through the frit during use of the chromatography column, which will result in the loss of mRNA-displayed proteins. This step forces the finest particles through the frit, and is especially important with the use of larger amounts of oligo(dT) cellulose.

2. Dilute the 1.3-ml translation reaction containing the mRNA-displayed proteins with added KCl and MgCl₂ (from Basic Protocol 1) into 8.7 ml of oligo(dT) binding buffer and incubate with the washed oligo(dT) cellulose for 15 min at 4°C with rotation.

Retain an aliquot of the undiluted translation reaction for SDS-PAGE and scintillation counting analyses.

3. Allow the diluted translation reaction mixture and the oligo(dT) cellulose to pass through a chromatography column so that the oligo(dT) cellulose is retained on the frit, and retain the flowthrough.
4. Wash three times with 1 ml oligo(dT) binding buffer.
5. Wash once with 1 ml oligo(dT) wash buffer.
6. Elute three times with 0.5 ml deionized water.
7. Analyze the undiluted translation reaction mixture, the flowthrough, and all of the washes and elutions using Tris-tricine SDS-PAGE as described in *UNIT 10.2A*, Alternate Protocol 1, and by scintillation counting.

The volume of the oligo(dT) eluate may be reduced by lyophilization by up to a factor of 5.

The oligo(dT) cellulose purification step anneals the poly(dA) region of the mRNA display template to immobilized oligo(dT) cellulose. Consequently, mRNA display templates not displaying protein and other mRNA molecules present in the lysate will co-purify with the displayed proteins. Long stretches of adenines in the RNA region of the mRNA display template should be avoided since they will also anneal to the oligo(dT) and present a substrate for RNase H which is present in the reticulocyte lysate; this will cause mRNA display template degradation.

The oligo(dT) cellulose purification step also presents a quick approximate method for the absolute measurement of the concentration of mRNA displayed proteins in the translation reaction mixture. The scintillation counter readings give the proportion of ³⁵S-methionine that is contained in the oligo(dT) eluates by counting equal proportions of the whole translation mixture and the oligo(dT) eluate and dividing one by the other. The ratio of the intensities of the bands corresponding to mRNA-displayed proteins in these two samples on the SDS-PAGE gel gives the yield of the oligo(dT) purification.

8. Calculate the concentration of mRNA displayed proteins in the translation reaction mixture with the following equation:

$$[\text{mRNA-displayed proteins}] = Y^{-1} \times C \times [\text{methionine}] \times N^{-1}$$

where *Y* is the yield of the oligo(dT) cellulose purification determined by SDS-PAGE; *C* is the number of counts in the combined oligo(dT) elution fractions divided by the number of counts in an equal proportion of the translation reaction mixture determined by scintillation counting; [methionine] is the total concentration of hot and cold methionine in the translation reaction mixture before the high salt incubation; and *N* is the average number of methionines in a single displayed protein.

This calculation assumes that the initiating methionine is still present on the protein and that the concentration of methionine in the reticulocyte lysate is known (~5 μM before addition to the translation reaction); this last error can be reduced by adding a known amount of cold methionine to the translation reaction mixture.

A potentially more accurate method for the direct measurement of the concentration of mRNA displayed proteins is to construct the mRNA display template using a mixture of DNA linker that has been kinased (5'-phosphorylated) with labeled ATP as well as the cold kinased (5'-phosphorylated) linker. This labeled template can then be translated in the presence of only cold methionine. It is generally not possible to observe the difference in mobility on a SDS-PAGE gel between the mRNA display template displaying protein and the mRNA display template alone. The addition of a DNA oligonucleotide, complementary to a region of the RNA part of the mRNA display template close to but not right at the 3' end of the RNA, to the translation reaction mixture, as well as magnesium chloride to 10 mM, will cause the RNA part of the mRNA-displayed proteins to be digested away by RNase H, which contaminates reticulocyte lysate, leaving proteins displayed upon the ³²P-labeled DNA linker only. These may easily be resolved from the DNA linker not displaying protein using SDS-PAGE. The ratio between these two bands gives a more direct measurement of

the proportion of mRNA display template that displays protein, and using this the concentration of mRNA-displayed proteins in the translation reaction mixture may easily be calculated. The proportion of mRNA display template that ends up displaying protein can vary from <1% to 40% depending upon the sequence, the myc epitope sequence is at the upper end of this range.

Ni-NTA purification

The Ni-NTA purification is upon the basis of the His₆ tag and is only appropriate if this is present in the library sequence (see also *UNIT 10.11*).

9. Wash 100 µl of Ni-NTA agarose three times with 1 ml deionized water.
10. Mix 0.5 ml of the oligo(dT) eluate with 2× Ni-NTA binding buffer, vortex to dissolve, add 0.7 µl of 2-mercaptoethanol, incubate with the washed Ni-NTA agarose for 1 hr at 4°C with rotation.

The 2× Ni-NTA binding buffer is the solid residue obtained by evaporation to dryness of 1× Ni-NTA binding buffer.

11. Allow the Ni-NTA binding buffer and the Ni-NTA agarose to pass through a chromatography column so that the Ni-NTA agarose is retained on the frit, retain the flowthrough.
12. Perform the following washes on the chromatography column:
 - a. Wash two times with 500 µl Ni-NTA wash buffer 1.
 - b. Wash once with 500 µl of a 4:1 solution of Ni-NTA wash buffer 1:Ni-NTA wash buffer 2.
 - c. Wash once with 500 µl of a 3:2 solution of Ni-NTA wash buffer 1:Ni-NTA wash buffer 2.
 - d. Wash once with 500 µl of a 2:3 solution of Ni-NTA wash buffer 1:Ni-NTA wash buffer 2.
 - e. Wash once with 500 µl of a 1:4 solution of Ni-NTA wash buffer 1:Ni-NTA wash buffer 2.
 - f. Wash once with 500 µl Ni-NTA wash buffer 2.
 - g. Wash two times with 500 µl of a 19:1 solution of Ni-NTA wash buffer 2:Ni-NTA elution buffer.
 - h. Elute for 30 min at 4°C with rotation two times with 250 µl Ni-NTA elution buffer.
13. Analyze the starting material, the flowthrough, and all washes and elutions using Tris-tricine SDS-PAGE as described in *UNIT 10.2A*, Alternate Protocol 1, and by scintillation counting.

EDTA should be added to the eluate to give 5 mM to bind to eluted Ni²⁺.

The volume of the eluate may be reduced by lyophilization by up to a factor of 5.

If the mRNA-displayed proteins are prone to aggregation, then it may be necessary to maintain denaturing conditions throughout the Ni-NTA agarose purification process by the addition of urea or guanidinium hydrochloride to the wash and elution buffers in addition to that which is in the binding buffer.

If it is not desired to completely denature the mRNA-displayed proteins, the denaturant can be omitted from all buffers including the binding buffer; using such a native Ni-NTA agarose purification procedure is likely to result in decreased yields compared to the denaturing Ni-NTA agarose purification procedure described above.

The Ni-NTA agarose purification is upon the basis of the His₆ tag. Alternatively, other protein-affinity tags may be encoded within the protein sequence and used as a basis for

purification, such as the FLAG tag (see Support Protocol 1). The Ni-NTA purification will separate the mRNA-displayed proteins from the mRNA display templates not displaying proteins and other mRNA molecules that were not purified away in the oligo(dT) purification. The Ni-NTA agarose eluate will, however, contain contaminating free library protein if this is present in the input mixture; in this protocol this is removed in the preceding oligo(dT) purification. If it is desired to purify the free library protein, then it is best to purify the translation mixture initially upon the basis of a FLAG tag, with the Ni-NTA agarose purification subsequent to this. Additionally, a denaturing His₆ tag purification may be used to purify selected mRNA-displayed proteins away from the selection binding buffer if more than one selection step is to be used between amplification steps and the denaturing and renaturing of the mRNA-displayed proteins is desired.

Strong chelating agents such as EDTA, EGTA, and DTT must be avoided in Ni-NTA binding and wash buffers since they will compete with the immobilized NTA for complexation of the Ni²⁺, and may elute it from the agarose.

Purify on gel filtration column

14. On a NAP-5 gel filtration column, exchange the elution buffer into water by allowing 10 ml of deionized water to flow through the gel-filtration column.
15. Add 100 μ l of 10 mg/ml salmon sperm DNA and 10 μ l of 1 mg/ml BSA to 890 μ l of deionized water, vortex, and allow this to flow through the gel filtration column.
16. Allow 10 ml of deionized water flow through the gel filtration column.
17. Allow 0.5 ml of sample to flow through the gel filtration column.
18. Add 1 ml of deionized water to the column and collect the 1-ml eluate issued from bottom of column.
19. Analyze the starting material and the elution fraction using Tris-tricine SDS-PAGE as described in UNIT 10.2A, Alternate Protocol 1 and by scintillation counting.

Imidazole does not inhibit reverse transcription, so this buffer exchange is optional, it may be possible to reverse transcribe the mRNA display templates by diluting the Ni-NTA eluate directly into the reverse transcription reaction mixture. Reverse transcription may not proceed if denaturants are present in the reaction mixture.

Reverse transcribe mRNA-displayed proteins

20. Set aside a small sample of the mRNA-displayed proteins that are not reverse-transcribed for use in the no-RT control PCR amplification.
21. Make up the following reverse transcription reaction mixture on ice. Mix mRNA displayed proteins and DNA splint (functions as the RT primer) together first before adding RT buffer, and add reverse transcriptase last:

900 μ l mRNA-displayed proteins
15 μ l 200 μ M DNA splint (final 2 μ M)
300 μ l 5 \times reverse transcription buffer (final 1 \times)
150 μ l 100 mM DTT (final 10 mM)
30 μ l (each) 25 mM deoxynucleotide triphosphates (final 0.5 mM)
5 μ l water
10 μ l 200 U/ μ l Superscript II reverse transcriptase (final 1333 U/ml)
Total, 1500 μ l.

Incubate the reverse transcription reaction for 50 min at 42°C.

22. Analyze the starting material and the product of the reverse transcription using Tris-tricine SDS-PAGE as described in UNIT 10.2A, Alternate Protocol 1, and by scintillation counting.

The volume of the eluate may be reduced by lyophilization by up to a factor of 5. The reverse-transcribed mRNA-displayed proteins have greater mobility on the SDS-PAGE gel compared to those that have not been reverse transcribed. This difference in mobility provides a simple method for the accurate assay of the proportion of the mRNA displayed proteins which have been reverse transcribed. However, this change in mobility may only be observed if the cDNA-RNA association is preserved during the treatment of the sample during gel loading, which may not be the case if it is heated too strongly (much above 90°C).

It is common practice in reverse-transcription reactions to heat denature the primer and RNA template before the addition of the reverse transcriptase; this may influence the conformation of the mRNA displayed proteins, and depending on the project may not be advisable. Mixing the primer and the mRNA display template together under low-salt conditions, before the addition of the buffer, should promote their association.

The use of mRNA displayed proteins in selection experiments may yield functional RNA sequences unless the mRNA display template is reverse transcribed before the selection step. This will also reduce the likelihood that the mRNA display template will disrupt the structure of the protein that it displays. Free proteins originally selected using mRNA display may need to be incubated under reverse transcription conditions in order to achieve their active conformations.

Purify RT products

23. Exchange the buffer into selection buffer on a NAP-5 gel filtration column according to the manufacturer's instructions.
24. Allow 10 ml of selection binding buffer to flow through the gel filtration column.
25. Add 100 μ l of 10 mg/ml salmon sperm DNA and 10 μ l of 1 mg/ml BSA to 890 μ l of selection binding buffer, vortex, and allow this to flow through the gel filtration column.
26. Allow 10 ml of selection binding buffer to flow through column and 0.5 ml of sample to flow through column.
27. Add 1 ml of selection binding buffer to the top of the gel filtration column and collect the 1-ml eluate issued from bottom of column.
28. Analyze the starting material and the elution fraction using Tris-tricine SDS-PAGE as described in *UNIT 10.2A*, Alternate Protocol 1, and by scintillation counting.

The selection binding buffer used here is specific to the selection being performed.

Alternatively, a protein-folding step may accompany the buffer exchange into selection buffer. In this approach a denaturant such as guanidinium hydrochloride or urea is added directly to the mRNA-displayed proteins after reverse transcription, and this is dialyzed away over several hours into selection buffer. It is important to ensure that the denaturing conditions are not so denaturing that the association between the cDNA and the mRNA display template is broken; this may be assayed using SDS-PAGE.

BASIC PROTOCOL 3

SELECTION AND AMPLIFICATION OF THE mRNA-DISPLAYED PROTEINS

Selection protocols are highly project-dependent. The following protocol was successfully used to select ATP-binding proteins from a random sequence library and is included as an example. Cycles of selection and amplification, as described in this protocol, should be repeated until the proportion of the resultant mRNA-displayed proteins in the selected fraction is no longer increasing—typically, 8 to 12 cycles are required. At this point the selected library sequences should be determined by cloning and sequencing (see Chapters

1 and 7), and individual clones should be assayed for activity under the selection conditions both as mRNA-displayed and free proteins.

Materials

ATP agarose (Sigma)
ATP-aptamer selection binding buffer (see recipe)
Purified mRNA-displayed proteins (see Basic Protocol 2)
ATP-aptamer selection elution buffer (see recipe)
100 mM EDTA (*APPENDIX 2*)
1 M NaOH (*APPENDIX 2*)
1 M HCl
10 mg/ml salmon sperm DNA
1 mg/ml BSA
100 μ M 3' primer (specific for cDNA library)
100 μ M 5' primer (specific for cDNA library)
25 mM (each) deoxynucleotide triphosphates
10 \times PCR buffer containing 15 mM MgCl₂ (Boehringer Mannheim)
5 U/ μ l *Taq* DNA polymerase (Boehringer Mannheim)
25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol
Chloroform
1-Butanol
3 M NaCl
100% ethanol
Gel filtration columns (e.g., NAP-25, Amersham Pharmacia Biotech)
Additional reagents and equipment for butanol extraction (*UNIT 2.1A*)

Select for ATP-binding proteins

1. Wash 10 mg of ATP-agarose three times with 1 ml of deionized water followed by two times with 1 ml of ATP-aptamer selection binding buffer.
2. Incubate 1 ml of the purified mRNA-displayed proteins, from Basic Protocol 2, with the washed ATP-agarose for 1 hr at 4°C with rotation; drain for flowthrough.
3. Wash six times with 1000 μ l ATP-aptamer selection binding buffer at 4°C; allow to stand 10 min between washes.
4. Elute six times with 250 μ l ATP-aptamer selection elution buffer at 4°C; allow to stand 10 min between elutions.
5. Assay all fractions using scintillation counting.

If the proportion of mRNA-displayed proteins in the elution fraction is high (>5%), it may be helpful to perform more than one selection step between amplification steps. In the case of an aptamer selection, this will necessitate the purification of the selected mRNA-displayed proteins away from the elution buffer. This purification can be performed while preserving native conditions, and often directly in the selection buffer, upon the basis of the FLAG tag, as described in Support Protocol 1. Alternatively, this purification can be performed with a denaturing and renaturing step, optionally upon the basis of the His₆ tag under denaturing conditions, as described above.

Purify selected cDNA sequences that encode selected mRNA-displayed proteins

6. To 1.5 ml of eluted mRNA displayed proteins, add 200 μ l of 100 mM EDTA and 200 μ l of 1 M NaOH, heat for 10 min at 90°C, cool on ice, and add 200 μ l of 1 M HCl.
7. Exchange the buffer into deionized water on a NAP-25 gel filtration column according to the manufacturer's instructions.

8. Allow 25 ml of deionized water to flow through the column.
9. Add 200 μl of 10 mg/ml salmon sperm DNA and 20 μl of 1 mg/ml BSA to 1780 μl of deionized water, vortex, and allow this to flow through column.
10. Wash with 25 ml of deionized water and allow water to flow through the gel filtration column.
11. Measure the sample volume and pass through column, then add a volume of deionized water to the column such that the total volume added to the column is 2.5 ml.
12. Add 3.5 ml of deionized water to the top of the gel filtration column and collect the 3.5-ml eluate issued from the bottom of column.

With the exception of the hydrolysis step, this buffer exchange procedure may be optionally repeated after the volume of the sample has been reduced to ≤ 2.5 ml by evaporation under reduced pressure.

Amplify selected sequences by PCR

13. Amplify selected sequences by PCR (see also *UNIT 15.1*). Make up a PCR reaction mixture on ice as follows:

3500 μl selected cDNA library (from step 12)
 100 μl 100 μM 3' primer (final 2 μM)
 100 μl 100 μM 5' primer (final 2 μM)
 40 μl (each) 25 mM deoxynucleotide triphosphates (final 0.2 mM)
 500 μl 10 \times PCR buffer containing 15 mM MgCl_2 (final 1 \times)
 735 μl water
 25 μl 5 U/ μl *Taq* DNA polymerase
 Total, 5000 μl .

The number of cycles, temperatures, and durations of the incubation periods within each cycle need to be determined for the specific library being used (UNIT 15.1). The PCR amplification of DNA libraries should be piloted, and care should be exercised not to over-PCR amplify DNA libraries since they will not reanneal once denatured. If PCR is continued upon a denatured DNA library, rare sequences will be amplified to a greater extent than common sequences, which will reduce the enrichment factor of the selected functional sequences.

14. Perform a no-RT control (set aside in Basic Protocol 2, step 20) alongside this PCR reaction.

In this control a small amount of the mRNA display template that has not been reverse transcribed is used in place of the selected cDNA library. This should not give any observable product after an equivalent amount of amplification. If it does, then either the buffers are contaminated or the purification of the mRNA-displayed proteins is not stringent enough. In either case the problem must be addressed or the selection is unlikely to give the desired result. It is also often useful to perform an additional no-template control in which no template, reverse transcribed or otherwise, is added. If this gives observable product after an equivalent amount of amplification, then this is usually a sign of contaminated reagents.

Mutations will be introduced into the DNA library during PCR amplification. The mutagenic rate can be decreased by using a high-fidelity DNA polymerase such as Pfu DNA polymerase (e.g., Stratagene). The mutagenic rate can be increased by using the mutagenic PCR protocol described in Support Protocol 2. Mutagenic procedures such as mutagenic PCR may be used to increase library diversity by exploring parts of sequence space proximate to the starting sequence(s).

Purify double-stranded PCR product

15. Add 1:1 molar equivalents of 100 mM EDTA to chelate the Mg^{2+} .
16. Vortex the PCR reaction mixture with an equal volume of 25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol, centrifuge for 1 min at $10,000 \times g$, room temperature, remove and retain upper aqueous phase.
17. Re-extract aqueous phase with an equal volume of chloroform three times; centrifuge to clear on each occasion, remove and discard lower organic phase after each centrifugation.
18. 1-Butanol extract the aqueous phase to 20% of the initial volume at minimum (*UNIT 2.1A*, Support Protocol 2), remove and discard the upper 1-butanol phase.
Perform extraction in a polypropylene tube, as butanol will damage polystyrene.
19. Add 3 M NaCl to final 300 mM (include the salt that originates from the PCR buffer in calculating concentration) and 2.5 volumes of 100% ethanol.
20. Cool for 20 min at $-80^{\circ}C$ or overnight at $-20^{\circ}C$. Centrifuge for 10 min at $12,000 \times g$, $4^{\circ}C$. Decant and discard the supernatant.
21. Centrifuge the pellet for 1 min at $12,000 \times g$, remove remaining supernatant with a plastic pipet tip, make up in 30 mM NaCl, and measure the concentration by agarose gel (more explicit instructions may be found in *UNIT 2.12*).
The dsDNA library will not re-anneal if denatured, so care should be taken not to expose it to low-salt or high temperature conditions.
22. Transcribe the DNA into RNA, (see Basic Protocol 1) and repeat the entire procedure (see Basic Protocols 1, 2, and 3).

FLAG TAG PURIFICATION

The FLAG tag purification may optionally be used in place of or in addition to the His₆ tag purification in the purification of mRNA-displayed proteins. The FLAG tag purification is usually performed in addition to the His₆ tag purification during the preselection of individual cassettes during the library construction process. In this instance, the FLAG tag and His₆ tag are placed at opposite protein termini; purification upon the basis of the presence of both tags ensures that the protein is full-length and in-frame at both termini. This in turn ensures that the mRNA cassette that encodes the protein is free of insertions, deletions, and stop codons, and is suitable for the preparation of the full-length library by restriction and ligation of the resulting PCR-amplified cDNA sequences. In addition, FLAG tag purification may be used to purify selected mRNA-displayed proteins away from the selection binding buffer if more than one selection step is to be used between amplification steps and the denaturing and renaturing of the mRNA-displayed proteins is not desired.

Alternatively, FLAG tag purification may be used to purify free proteins away from reticulocyte lysate.

The FLAG purification is upon the basis of the FLAG tag sequence (DYKDDDDK) and is only appropriate if this is present in the library (see Strategic Planning).

Additional Materials (also see Basic Protocol 1)

- Anti-FLAG M2 agarose (Sigma)
- FLAG clean buffer (see recipe)
- FLAG binding buffer (see recipe)
- FLAG peptide (Sigma)

SUPPORT PROTOCOL 1

1. Wash 100 μ l of anti-FLAG M2 agarose three times with 1 ml of FLAG clean buffer, and then three times with 1 ml of FLAG binding buffer.
2. Exchange sample buffer into FLAG binding buffer according to the directions presented in Basic Protocol 1 for other buffer exchanges.

Optionally, dilute the sample buffer into the FLAG binding buffer or attempt purification directly from selection elution buffer.

3. Place 1 ml of the sample containing the mRNA-displayed proteins onto the washed anti-FLAG agarose and incubate for 1 hr at 4°C with rotation, drain, and retain flowthrough.
4. Wash the anti-FLAG agarose three times with 1 ml of FLAG binding buffer.
5. Elute from the anti-FLAG agarose two times with 0.5 ml FLAG binding buffer containing 10 μ M of the FLAG peptide, 30 min for each elution at 4°C with rotation.

If the FLAG tag purification is to be followed by a denaturing His₆ tag purification, then the elution fraction may be added directly to the 2 \times Ni-NTA binding buffer.

SUPPORT PROTOCOL 2

MUTAGENIC PCR

Mutagenic PCR may be used to increase the diversity of the DNA library that encodes the protein library. Mutagenic PCR may be used to generate the initial library, or to explore parts of sequence space proximate to the starting sequence(s). A broader discussion of the use of mutagenic PCR may be found in Cadwell and Joyce (1992).

Before the entire mutagenic protocol is enacted, it is important to pilot the PCR conditions to ensure that primer dimers are not taking over, and that the amplification per cycle is at least 1.7 to 1.8. The optimum PCR amplification conditions may be different from non-mutagenic PCR amplification performed upon the same library. One may wish to redesign the primers, since the part of the template sequence they anneal to will not be mutagenized.

Additional Materials (also see Basic Protocol 3)

- 2.5 M KCl
- 100 mM MnCl₂ solution
- 100 mM Tris·Cl, pH 8.3 (APPENDIX 2)
- 100 μ l PCR tubes (Sarstedt)

Additional reagents and equipment for agarose gel electrophoresis (UNIT 15.1)

1. Make up the following PCR reaction mixture on ice:

- 100 μ l 100 μ M 3' primer (final 2 μ M)
- 100 μ l 100 μ M 5' primer (final 2 μ M)
- 60 μ l (each) 25 mM dCTP and dTTP (final 1 mM)
- 12 μ l (each) 25 mM dATP and dGTP (final 0.2 mM)
- 30 μ l 2.5 M KCl (final 50 mM)
- 10.5 μ l 1 M MgCl₂ (final 7 mM)
- 7.5 μ l 100 mM MnCl₂ (final 0.5 mM)
- 150 μ l 100 mM Tris·Cl, pH 8.3 (final 10 mM)
- 943 μ l water
- 15 μ l 5U/ μ l *Taq* DNA polymerase
- Total, 1500 μ l.

- Pipet 16 90- μ l aliquots of PCR reaction mixture into 100- μ l PCR tubes and label them 1 to 16.

These may be stored for up to a few hours at 4°C.

- Add the DNA library or sequence to tube 1 to give 10 nM, make up to 100 μ l with PCR reaction mix.
- Perform 4 cycles of PCR amplification. During the final extension incubation, place the next-numbered tube alongside the current one in the PCR block. Before the final extension is complete but ensuring that the next-numbered tube is at the extension temperature, transfer 10 μ l of PCR reaction mixture. Retain the amplified PCR reaction mixture at 4°C.
- Repeat step 4 fourteen times. Every four transfers, analyze the PCR reaction using agarose gel electrophoresis (UNIT 15.1), quantitate the bands in successive PCR amplifications, and adjust the transfer volume in order to maintain the concentration of amplified DNA at a constant level.

It is important not to over-PCR the DNA. If PCR amplification ceases before a concentration of 100 nM is reached, then the initial DNA concentration should be reduced accordingly.

If the initial DNA was of one or a small number of known sequences, then it is possible to directly measure the average mutagenic rate by sequencing some of the individual library members from the final mutagenic PCR amplification sample. Assuming that the mutagenic rate is constant throughout the procedure allows for the direct control of the extent of mutagenesis by choosing one, or a mixture of more than one, of the successive mutagenic PCR amplification mixtures to serve as the source of the new DNA library. This sample may then be further amplified with PCR, optionally with further mutation. It is expected that the mutagenic rate will be about 0.2% per nucleotide per transfer (ten-fold amplification).

REAGENTS AND SOLUTIONS

The water used to make the following buffers should be deionized, ultrafiltered and subsequently tested for the absence of RNase by incubation with ³²P-labeled RNA and denaturing PAGE analysis. All buffers should be analyzed similarly. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

ATP-aptamer selection binding buffer

39.0 mg MgCl₂ (mol. wt. 95.2; 4.1 mM final)
2.92 g KCl (mol. wt. 74.6; 392 mM final)
476 mg HEPES (mol. wt. 238; 20 mM final)
3.07 mg glutathione (mol. wt. 307; 2 mM final)
3.06 mg glutathione disulfide (mol. wt. 612; 1 mM final)
3.72 mg EDTA·2Na⁺ (mol. wt. 372; 100 μ M final)
250 μ l Triton X-100 (0.25% final)
Bring up to 100 ml with water
Store at -20°C

Deoxygenate the buffer before the addition of the glutathione by bubbling an oxygen-free grade of an inert gas such as argon or nitrogen through it, and adjust the pH to 7.4.

ATP-aptamer selection elution buffer

285 mg ATP·2Na⁺ (mol. wt. 569; 5 mM final)
84.7 mg MgCl₂ (mol. wt. 95.2; 8.9 mM final)
2.92 g KCl (mol. wt. 74.6; 392 mM)
476 mg HEPES (mol. wt. 238; 20 mM final)

continued

3.07 mg glutathione (mol. wt. 307; 2 mM final)
3.06 mg glutathione disulfide (mol. wt. 612; 1 mM final)
3.72 mg EDTA·2Na⁺ (mol. wt. 372; 100 μM final)
0.25 g Triton X-100 (0.25% w/v final)
Bring up to 100 ml with water
Store at -20°C

Deoxygenate the buffer before the addition of the glutathione by bubbling an oxygen-free grade of an inert gas such as argon or nitrogen through it, and adjust the pH to 7.4.

FLAG binding buffer

877 mg NaCl (mol. wt. 58.4; 150 mM final)
1.19 g 50 mM HEPES (mol. wt. 238; 50 mM final)
0.25 g Triton X-100 (0.25% w/v final)
Adjust the pH to 7.4 with NaOH/HCl
Bring up to 100 ml with water
Store at -20°C

FLAG clean buffer

751 mg glycine (mol. wt. 75.1; 100 mM final)
0.25 g Triton X-100 (0.25% w/v final)
Bring up to 100 ml with water
Store at -20°C
Adjust pH to 3.5 with NaOH/HCl.

Ni-NTA binding buffer

57.4 g guanidine hydrochloride (mol. wt. 95.5; 6 M final)
2.93 g NaCl (mol. wt. 58.4; 500 mM final)
1.42 g Na₂HPO₄ (mol. wt. 142; 100 mM final)
121 mg Tris(hydroxymethyl)aminomethane (mol. wt. 121; 10 mM final)
0.25 g Triton X-100 (0.25% w/v final)
701 μl 2-mercaptoethanol (mol. wt. 78.1; 10 mM final)
Adjust the pH to 8.0 with NaOH/HCl
Bring up to 100 ml with water
Store at -20°C

In order to prepare 2× Ni-NTA binding buffer, the 1× Ni-NTA binding buffer should be evaporated to dryness under reduced pressure. Upon using the resultant 2× Ni-NTA binding buffer, the 2-mercaptoethanol will have to be added again.

Ni-NTA elution buffer

2.93 g NaCl (mol. wt. 58.4; 500 mM final)
121 mg 10 mM Tris(hydroxymethyl)aminomethane (mol. wt. 121; 10 mM final)
0.25 g Triton X-100 (0.25% w/v final)
1.70 g imidazole (mol. wt. 68.1; 250 mM final)
701 μl 2-mercaptoethanol (mol. wt. 78.1; 10 mM final)
Adjust pH to 8.0 with NaOH/HCl
Bring up to 100 ml with water
Store at -20°C

Ni-NTA wash buffer 1

48.1 g urea (mol. wt. 60.1; 8 M final)
2.93 g NaCl (mol. wt. 58.4; 500 mM final)
1.20 g NaH₂PO₄ (mol. wt. 120; 100 mM final)
121 mg Tris(hydroxymethyl)aminomethane (mol. wt. 121; 10 mM final)
0.25 g Triton X-100 (0.25% w/v final)

continued

701 μ l 2-mercaptoethanol (mol. wt. 78.1; 10 mM final)
Adjust the pH to 6.3 with NaOH/HCl
Bring up to 100 ml with water
Store at -20°C

Ni-NTA wash buffer 2

2.93 g NaCl (mol. wt. 58.4; 500 mM)
121 mg Tris(hydroxymethyl)aminomethane (mol. wt. 121; 10 mM final)
0.25 g Triton X-100 (0.25% w/v final)
701 μ l 2-mercaptoethanol (mol. wt. 78.1; 10 mM final)
Adjust the pH to 8.0 with NaOH/HCl
Bring up to 100 ml with water
Store at -20°C

Oligo(dT) binding buffer

7.46 g KCl (mol. wt. 74.6; 1 M final)
1.21 g Tris(hydroxymethyl)aminomethane (mol. wt. 121; 100 mM final)
372 mg disodium EDTA (mol. wt. 372; 10 mM final)
0.25 g Triton X-100 (0.25% w/v final)
Adjust the pH to 8.0 with NaOH/HCl
Bring up to 100 ml with water
Store at -20°C

Oligo(dT) wash buffer

746 mg KCl (mol. wt. 74.6; 100 mM final)
121 mg Tris(hydroxymethyl)aminomethane (mol. wt. 121; 10 mM final)
0.25 g Triton X-100 (0.25% w/v final)
Adjust the pH to 8.0 with NaOH/HCl
Bring up to 100 ml with water
Store at -20°C

Transcription buffer, 10 \times

255 mg spermidine trihydrochloride (mol. wt. 255; 10 mM final)
4.84 g Tris(hydroxymethyl)aminomethane (mol. wt. 121; 400 mM final)
770 mg DTT (mol. wt. 154; 50 mM final)
0.1 g Triton X-100 (0.1% w/v final)
Adjust the pH to 8.0 with NaOH/HCl
Bring up to 100 ml with water
Store at -20°C

COMMENTARY

Background Information

In vitro selection experiments were first successfully performed upon nucleic acid libraries, for reviews see Szostak and Ellington (1993), Gold et al. (1993), and Joyce (1993). Nucleic acids are the only molecular systems that are capable of being replicated directly in vitro and which also can contain more than trivial amounts of amplifiable information. Nucleic acid selections have the advantage that the target functional entity, and the information encoding the functional entity, are the same. Within nucleic acid selections, large libraries (i.e., up to 10^{17} different molecules) are sub-

jected to successive cycles of selection and amplification until functional sequences dominate the library, at which point they may be identified by cloning and sequencing. The idea of extending the in vitro selection approach to proteins was an obvious one; what was not obvious, however, was how to extract the sequence information from selected proteins in order to permit their amplification and ultimately their identification.

One way to extract the sequence information from selected proteins is to covalently attach each of them to the mRNA sequence that encodes it (Roberts and Szostak, 1997; Roberts,

1999; Roberts and Ja, 1999). In this manner the function (phenotype) and amplifiable sequence information (genotype) are part of the same molecule. Selection may be performed upon the basis of the function of the protein, while the protein may be amplified upon the basis of the mRNA that encodes it and is attached to it. The problem was how to arrange things so that proteins may be attached to the mRNA that encodes them in parallel while mixed within one reaction mixture. The molecule that makes this possible is puromycin (Fig. 24.5.1). Puromycin is an antibiotic that functions by inhibiting translation. It sufficiently closely resembles a charged tRNA that it is able to enter the A-site of the ribosome and react with the activated C-terminal of the nascent peptide. Since this reaction results in the formation of a stable amide bond, rather than the hydrolyzable ester bond that connects the tRNA to the amino acid in an amino-acyl tRNA, translation is halted. If the puromycin is already attached to the mRNA that is being translated, a stable covalent linkage results between a protein and the mRNA that encodes it and the ribosome may be purified away.

In a conceptual sense, the most similar systems to mRNA display are phage display (Smith and Petrenko, 1997) and ribosome display (Jermutus et al., 1998). All three systems may be used to search nucleic acid libraries for the functional proteins or peptides they encode. In phage display, the protein library is encoded within the phage genome and is expressed upon the surface of the phage as a fusion with the phage coat protein. Phage may be selected upon the basis of the functionality of their surface proteins, and the protein may then be amplified by allowing the selected phage to replicate. The diversity of phage display selection experiments is limited to the numbers of phage that may reasonably be transformed or packaged, which is $\sim 10^8$ to 10^9 . Recent advances have shown that it is possible to display libraries of proteins upon the surface of phage (Sche et al., 1999), but phage display has yet to be used to discover a new protein from an entirely random sequence library, as distinct from a library derived from a known folded protein. In ribosome display, paused ribosomes display the protein library in a ternary complex with the mRNA that encodes the displayed protein. In order to synthesize this ternary complex, an mRNA library is prepared without a stop codon. Upon translation of this template the ribosomes will pause at the end of the open reading frame. Because the ribosome is unable to release itself

from the message, it displays both the nascent protein and the mRNA that encodes it. These constructs may be used for in vitro selection experiments upon the basis of the function of the displayed nascent protein. Selected proteins may then be amplified using RT-PCR amplification of the associated mRNA. The ribosome display constructs are relatively large, and selections can only be performed under conditions that preserve the ribosome-mRNA-nascent protein association. Also, since the ribosome display constructs also display a single-stranded mRNA, there is a possibility that functional RNA sequences may be selected in place of the desired functional proteins. This problem can be avoided by reverse transcribing the mRNA associated with the ribosome.

Critical Parameters

After designing and synthesizing the mRNA display library, it is important to have it sequenced in order to ascertain the proportion of the library which is error-free and appropriate for the selection. If the proportion of library members with insertions, deletions, or stop codons is too high, the library may have to be resynthesized with extra purification steps incorporated at the DNA cassette stage (pre-selection, see Fig. 24.5.3). Alternatively, considerable library quality improvement may result from the careful denaturing PAGE purification of small amounts of the DNA cassettes.

The various purification steps performed after the initial synthesis of the mRNA-displayed proteins should be individually optimized, with assays performed by SDS-PAGE to ascertain that the mRNA displayed proteins are still attached to full-length mRNA. Subsequent to this, a pilot (“round zero”) purification should be performed in which the various optimized purification steps are applied sequentially to the same sample. Only upon the satisfactory completion of round zero should the large-scale translation reaction be made up for the first round of selection (“round one”).

The various purification steps that form part of each cycle of the selection of mRNA displayed proteins must be assayed by SDS-PAGE in order to confirm that the mRNA display template has not become degraded at any stage in the process.

Both positive and negative controls need to be used to assay the selection step; this should then be optimized to discriminate between the two controls to the maximum possible reasonable extent. This maximal discrimination selection protocol should be adopted after round two

or three; at this stage the absolute yield of the selection step is no longer a concern owing to the high copy number of selected sequences that have passed through ≥ 1 amplification step.

Alongside the PCR amplification that follows each selection step, a no-RT control should be performed. In this control a small amount of the mRNA display template that has not been reverse-transcribed is used in place of the selected cDNA library. This should not give any observable product after an equivalent amount of amplification. If it does, then either the buffers are contaminated or the purification of the mRNA-displayed proteins is not stringent enough. A no-template control PCR amplification will distinguish between these two possibilities. In either case, the problem must be addressed or the selection is unlikely to give the desired result.

Troubleshooting

Problems that may be encountered with this procedure are detailed in Table 24.5.2.

Anticipated Results

The results of a selection largely depend upon how many of the initial library members can perform the task for which they are being selected and how well they can perform it under the selected conditions. Ideally, the observed activity in each round of selection will exponentially rise to a high value and then plateau. Assuming that there are a relatively small number of members of the initial library with activity that causes them to be selected, it is likely that several rounds of selection and amplification will have to be performed before any significant increase in activity is observed.

Once the selection activity has peaked or reached a plateau, then the library members should be sequenced. If there were relatively large numbers of members of the initial library with activity that causes them to be selected, then the library at this stage may still be fairly diverse. Since the cycles of selection and amplification preferentially amplify the most active members, a high-diversity library at the end of the selection is likely to indicate a failed selection or a selection that is not yet finished. Successful selections are likely to yield one or a small number of families of very closely related protein sequences each of which has diverged from a single ancestral protein sequence owing to errors accumulated during the many cycles of PCR amplification to which they have been subjected. Assays of individual members of these selected families of se-

quences should yield mRNA-displayed proteins with the desired function; these are also likely to be functional as free proteins, unless the mRNA display template greatly interferes with the conformation of the displayed protein that it displays.

Early indications are that proteins selected using mRNA display may fold into multiple conformations, only some of which have the desired functionality. This behavior causes the proportion of the selected library observed to demonstrate the desired activity to rise by a factor of much less than might be expected during successive rounds of selection. The individual selected proteins behave similarly, whether mRNA displayed or not. Mutagenesis and reselection of such selected individual library members, or libraries of them, has given large families of related proteins with greatly improved characteristics in this respect (A.D. Keefe, G. Cho, and J.W. Szostak, pers. commun.).

It should always be borne in mind that selections will give a solution to the problem that is set; it is up to the experimenter to arrange the selection conditions sufficiently carefully to ensure that this solution is a consequence of the desired functionality.

The range of acceptable yields of various parts of the mRNA-displayed protein selection protocol are listed in Table 24.5.3. Observed yields falling at the lower end of these ranges may or may not be increased upon optimization.

Time Considerations

The construction of the library may take anywhere between 2 weeks and 2 months. Doing pilot preparative and purifying experiments on mRNA displayed protein may take ≥ 1 month. A single round of selection and amplification will take 2 to 4 days and the initial rounds of selection may take 1 to 2 months. Mutagenesis will take 1 to 2 weeks and subsequent rounds of selection and amplification will take 1 to 2 months. Sequencing and assays of selected proteins may take ≥ 1 month.

Literature Cited

- Cadwell, R.C. and Joyce, G.F. 1992. Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* 2:28-33.
- Cho, G., Keefe, A.D., Liu, R., Wilson, D.S., and Szostak J.W. 2000. Constructing high complexity synthetic libraries of long ORFs using in vitro selection. *J Mol. Biol.* In press.

Table 24.5.2 Troubleshooting Guide to Problems That May Be Encountered In Protein Selection Using mRNA-Displayed Proteins

Problem	Possible cause	Solution
Sequencing of initial library or library cassettes reveals many insertions and/or deletions	Synthetic DNA is of low quality	Repeat synthesis and/or careful denaturing PAGE purification to resolve n from $n+1$ and $n-1$ oligonucleotides
Sequencing of initial library or library cassettes reveals many insertions and/or deletions and/or stop codons	Synthetic DNA is of low quality and/or stop codons appear in random region as a consequence of library design	Perform “preselection” in which mRNA displayed proteins are synthesized at the cassette stage; these are purified upon the basis of the presence of both terminal tags, and the resulting cDNA is used to construct the full-length library (see Fig. 24.5.3)
mRNA-DNA ligation does not yield any/enough mRNA display template	3'-end of mRNA and/or splint have self-structure arising from internal complementarity	Redesign mRNA and/or splint sequences
	Puromycin-terminated linker was not sufficiently 5'-phosphorylated	Repeat 5'-phosphorylation, optionally with extra enzyme
	Too much salt in the ligation reaction mixture	Desalt mRNA, splint and linker
	RNA is degraded	Repeat transcription and gel purification
No mRNA-displayed proteins observed on gel	Ligation failed	See above
	mRNA display template is degraded	Repeat ligation
	No methionines present in library except initiating methionine which is degraded away in lysate	Redesign protein library with more methionines
Oligo(dT) cellulose purification low-yielding	Elution buffer not sufficiently denaturing	Further deionized water washes are needed to wash away residual salt
Ni-NTA purification low-yielding	His ₆ tag not accessible	Use more denaturing conditions for the binding step or redesign library
	Product precipitates	Add denaturant to the wash and elution buffers
	EDTA, EGTA, DTT, or other chelating agents present in the binding buffer	Redesign protocol to exclude chelating agent
Library DNA observed in no-template control PCR amplification	PCR amplification components contaminated with library DNA	Determine which PCR amplification components are contaminated and replace them
Library DNA observed in no-RT control PCR amplification	Library DNA has not been purified away from mRNA displayed proteins, or mRNA displayed protein purification buffers are contaminated	Increase the stringency of the mRNA displayed protein purification protocol, or determine which mRNA displayed protein purification components are contaminated and replace them
Activity does not rise through selection	There are no functional sequences in library	Redesign or mutagenize library and reselect
	Selection step not designed appropriately	Test selection step with positive and negative controls, redesign to maximize distinction

continued

Table 24.5.2 Troubleshooting Guide to Problems That May Be Encountered In Protein Selection Using mRNA-Displayed Proteins, continued

Problem	Possible cause	Solution
	Biases in PCR, transcription, translation, or protein display overwhelming selection bias	Adjust conditions so that biases are reduced, especially in low-yielding steps; e.g., reduce mRNA display template concentration in translation
	Immobilized target not accessible to mRNA-displayed proteins	Repeat selection with different matrix and/or linker and/or target linkage point
No families observed in sequencing data at end of selection	Not enough cycles of selection and amplification performed	Continue with cycles of selection and amplification
	There are many sequences in the selected library that are active	Assay individual selected sequences
Selected sequences not active as mRNA-displayed proteins	Not enough cycles of selection and amplification performed	Continue with cycles of selection and amplification
	There are no functional sequences in library	Redesign or mutagenize library and reselect
Selected sequences active as mRNA-displayed proteins, but not as free proteins	Assay does not treat free proteins in exactly the same manner as mRNA-displayed proteins	Repeat assay, treating free proteins in the same manner as mRNA-displayed proteins, for example include the reverse transcription step
	Selected mRNA-displayed proteins have mRNA-dependent conformations	Redesign or mutagenize library and reselect

Table 24.5.3 Results Obtained During mRNA-Displayed Protein Selection Procedure

Step	Range of acceptable yields
5'-phosphorylation of DNA linker	90%-100%
Splinted RNA-DNA ligation	20%-60%
Proportion of mRNA display template displaying protein	1%-40%
Oligo(dT) cellulose purification	30%-90%
Denaturing Ni-NTA purification	30%-90%
Anti-FLAG purification	50%-80%
Gel filtration chromatography (NAP column)	85%-100%
Reverse transcription	80%-100%
Proportion of mRNA displayed proteins in initial elution phase of aptamer selection	0.01%-1%
Proportion of mRNA displayed proteins in final elution phase of aptamer selection	3%-60%
<i>Other statistics relating to mRNA display protein selections</i>	
Number of rounds of selection until activity peaks or plateaus	8-12
Initial diversity of mRNA display library	10 ¹² -10 ¹³
Final diversity of mRNA display library	≥1-10 ⁴

Generation and Use of Combinatorial Libraries

24.5.33

- Colas, P., Cohen, B., Jessen, T., Grishina, I., McCoy, J., and Brent, R. 1996. Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2. *Nature* 380:548-550.
- Fields, S. and Song, O. 1989. A novel genetic system to detect protein-protein interactions. *Nature* 340:245-246.
- Gold, L., Allen, P., Binkley, J., Brown, D., Schneider, D., Eddy, S.R., Tuerk, C., Green, L., MacDougall, S., and Tasset, D. 1993. The shape of things to come. In *The RNA World*. (R.F. Gesteland and J.F. Atkins, eds.) pp. 497-509. Cold Spring Harbor, New York.
- Jermutus, L., Ryabova, L., and Plückthun, A. 1998. Recent advances in producing and selecting functional proteins by using cell-free translation. *Curr. Opin. Biotechnol.* 9:391-410.
- Joyce, G.F. 1993. Evolution of catalytic function. *Pure & Appl. Chem.* 65:1205-1212.
- LaBean, T.H. and Kauffman, S.A. 1993. Design of synthetic gene libraries encoding random sequence proteins with desired ensemble characteristics. *Protein Sci.* 2:1249-1254.
- Liu, R., Barrick, J., Szostak, J.W., and Roberts, R.W. 2000. Optimized synthesis of RNA-protein fusions for in vitro protein selection. *Methods Enzymol.* 317:268-293.
- Roberts, R.W. 1999. Totally in vitro protein selection using mRNA-protein fusions and ribosome display. *Curr. Opin. Chem. Biol.* 3:268-273.
- Roberts, R.W. and Ja, W.W. 1999. In vitro selection of nucleic acids and proteins: What are we learning? *Curr. Opin. Struct. Biol.* 9:521-529.
- Roberts, R.W. and Szostak, J.W. 1997. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl. Acad. Sci. U.S.A.* 94:12297-12302.
- Sche, P.P., McKenzie, K.M., White, J.D., and Austin, D.J. 1999. Display cloning: functional identification of natural product receptors using cDNA-phage cloning. *Chem. Biol.* 6:707-716.
- Smith, G.P. and Petrenko, V.A. 1997. Phage display. *Chem Rev.* 97:391-410.
- Stemmer, W.P.C. 1994. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* 370:389-391.
- Szostak, J.W. and Ellington, A.D. 1993. In vitro selection of functional RNA sequences. In *The RNA World*. (R.F. Gesteland and J.F. Atkins, eds.) pp. 551-533. Cold Spring Harbor, New York.
- Wilson, D.S. and Szostak, J.W. 1999. In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.* 68:611-647.
- Wolf, E. and Kim, P.S. 1999. Combinatorial Codons: A computer program to approximate amino acid probabilities with biased nucleotide usage. *Protein Sci.* 8:680-688.

Key References

Roberts and Szostak, 1997. See above.

The first demonstration of the formation of mRNA displayed proteins (RNA-protein fusions).

Liu et al., 2000. See above.

Describes the optimization of the synthesis and purification of mRNA displayed proteins (RNA-protein fusions).

Cho et al., 2000. See above.

Describes the use of mRNA display and in vitro selection to construct various types of high quality library for use in mRNA display protein selections.

Internet Resources

<http://gaiberg.wi.mit.edu/cgi-bin/CombinatorialCodons>

Combinatorial Codons is an extremely useful tool for the design of protein libraries; it generates a nucleotide distribution that iteratively approaches an input amino acid distribution.

<http://xanadu.mgh.harvard.edu/szostakweb/orf.html>

This site is a database of exact oligonucleotide sequences that have been successfully used in the construction of random, patterned, and structure-based mRNA-displayed protein libraries.

<http://paris.chem.yale.edu/extinct.html>

The Biopolymer Calculator is a very useful general tool for molecular biology.

<http://sun2.science.wayne.edu/%7Ejlsun2/servers/seqanal/>

A nucleic acid secondary structure prediction algorithm is given by mfold.

Contributed by Anthony D. Keefe
Massachusetts General Hospital
Boston, Massachusetts

CHAPTER 25

Discovery and Analysis of Differentially Expressed Genes in Single Cells and Cell Populations

INTRODUCTION

For decades, molecular biologists have been discovering and analyzing genes that are differentially expressed. Initially, discovery and analysis was one gene at a time. This was followed by cDNA cloning methods to identify genes that were expressed in a given tissue; however, this left the investigator with a large number of genes to screen for differential expression. A major advance was the development of subtractive cloning in the 1980s, which greatly enriched for genes that were expressed in one cell or tissue type rather than another. With the advent of PCR using thermostable DNA polymerases in the late 1980s, older methods have been refined and many new techniques have been developed that make discovery of differentially expressed genes much more facile and permit the analysis of differential gene expression at the single cell level.

This chapter consists of protocols—some of them older, some of them newer—for two kinds of methods. The first of these are amplification-based methods for analysis of individual cells and are contained within Section 25A. *UNIT 25A.1* describes the use of laser-capture microdissection of histological specimens so that one can analyze nucleic acids, in particular individual cells, using PCR. *UNIT 25A.2* describes methods for fixation of tissues and subsequent dissociation of the fixed tissue into single cells whose nucleic acids can be analyzed by PCR-based methods.

Section 25B contains molecular methods for discovery of differentially expressed genes. *UNIT 25B.1* (formerly *UNIT 5.8B*) describes production of a subtracted cDNA library while *UNIT 25B.2* (formerly *UNIT 5.9*) describes the refinement of PCR-based subtractive cDNA cloning with a support protocol for slot blot hybridization to monitor sublibraries. Subtracted cDNA libraries provide a method where cDNAs are synthesized from mRNA from the desired tissue or cell type and then sequences that are also expressed in a control tissue or cell type are removed by hybridization and selection.

UNIT 25B.3 describes a powerful application of PCR to gene discovery, differential display. This technique allows the identification and subsequent isolation of differentially expressed genes that requires no knowledge of sequences, but rather PCR amplification using arbitrary oligonucleotides and high resolution polyacrylamide gel electrophoresis. *UNITS 25B.4 & 25B.5* describe variations on differential display, restriction-mediated differential display (RMDD) and amplified fragment length polymorphism (AFLP) based transcript profiling, which make use of frequently cutting restriction enzyme sites in cDNAs and may offer advantages to the practitioner.

UNITS 25B.6 & 25B.7 contain different PCR-based approaches for determining what genes are expressed in a given cell or tissue type. *UNIT 25B.6* describes serial analysis of gene expression (SAGE). This technique generates concatemers of short cDNA sequence tags that have been ligated together. These concatemers can be cloned, sequenced, and analyzed with the aid of specialized software to identify differentially expressed genes. The unit also contains a protocol for cloning cDNA starting with a given sequence tag.

Contributed by Donald M. Coen

Current Protocols in Molecular Biology (2003) 25.0.1-25.0.2

Copyright © 2003 by John Wiley & Sons, Inc.

Discovery of
Differentially
Expressed Genes

25.0.1

Supplement 64

UNIT 25B.7 describes representational difference analysis (RDA). RDA combines PCR-mediated kinetic enrichment with subtractive hybridization to generate 0.2–2 kbp sequences that are distinct to genomic DNA or mRNA in one cell type versus another. These can then be cloned and sequenced or otherwise analyzed.

UNIT 25B.8 describes a protocol in which both PCR and reverse transcription have been optimized to permit the detection and semi-quantitative analysis of transcripts from single cells, small tissue biopsies, and microdissected samples. These protocols extend and complement those provided in *UNITS 25A.1 & 25A.2*.

Further additions to this chapter on analysis of gene expression in single cells are anticipated.

Donald M. Coen
Contributing Editor (Chapter 25)
Harvard Medical School

NUCLEIC ACID AMPLIFICATION FROM INDIVIDUAL CELLS

SECTION A

Laser Capture Microdissection

UNIT 25A.1

Mammalian tissues are histologically and biologically heterogeneous, and typically contain multiple cellular components, such as epithelial, mesenchymal (i.e., stromal), and inflammatory cells. Laser capture microdissection (LCM) offers a rapid and precise method of isolating and removing specified cells from complex tissues for subsequent analysis of their RNA, DNA, or protein content, thereby allowing assessment of the role of the cell type in the normal physiologic or disease process being studied. LCM has been utilized to study molecular changes during the neoplastic progression of specific cell types (Sgroi et al., 1999; Paweletz et al., 2000), and to understand the role of particular cell types in normal organ function (Glasow et al., 1998; Jin et al., 1999) and in various disease processes (Fend et al., 1999a; Sawyer et al., 2000). LCM has the potential to contribute to the understanding of many cellular processes, particularly processes involving multiple cell types, such as embryonic development, tissue differentiation and function, aging, and disease.

There are methods for tissue microdissection other than LCM, such as laser microbeam microdissection and laser-pressure catapulting, in which a fine laser beam is used to cut around individual or groups of cells and then laser energy is used to “catapult” the cells out of the tissue section and allow their collection (P.A.L.M. Mikrolaser Technologie; <http://www.palm-mikrolaser.com>); however, currently, Arcturus Engineering is the only manufacturer of instrumentation for LCM. Arcturus Engineering (<http://www.arctur.com>) can be contacted for details about the various LCM systems available and the current prices of instrumentation and consumables.

In this unit, protocols for the preparation of mammalian frozen tissues (see Basic Protocol 1), fixed tissues (see Basic Protocol 2), and cytologic specimens (see Basic Protocols 3 and 4) for LCM, including hematoxylin and eosin staining (H&E; see Basic Protocol 5 and UNIT 14.5), are presented, as well as a protocol for the performance of LCM utilizing the PixCell I or II Laser Capture Microdissection System manufactured by Arcturus Engineering (see Basic Protocol 6). Also provided is a protocol for tissue processing and paraffin embedding (see Support Protocol), and recipes for lysis buffers for the recovery of nucleic acids and proteins (see Reagents and Solutions). The Commentary section addresses the types of specimens that can be utilized for LCM and approaches to staining of specimens for cell visualization (see Critical Parameters). Emphasis is placed on the preparation of tissue or cytologic specimens as this is critical to effective LCM. Resources available on-line are given at the end of the unit (see Internet Resources).

PREPARATION OF FROZEN SECTIONS FOR LCM

Embedding and freezing is a way to preserve specimens and stabilize them for long-term storage and sectioning (also see UNIT 14.2). Tissue is embedded in a viscous compound, such as optimal cutting temperature (OCT; Tissue-Tek) medium, and rapidly frozen on dry ice. For long-term storage (i.e., months to years), liquid nitrogen offers the best preservation of protein and RNA. Storage at -80°C is adequate for shorter time periods (i.e., a few days to several weeks).

BASIC
PROTOCOL 1

Discovery of
Differentially
Expressed Genes

25A.1.1

Materials

Embedding medium (e.g., OCT; Tissue-Tek)
~1-cm maximum-dimension tissue samples
Cryomolds (Tissue-Tek)
Dry-ice container with lid
Aluminum foil
Microm cryostat, refrigerated to -20°C with tissue platform (chuck) and appropriate blades (Richard-Allan Scientific)
Glass slides (e.g., Gold Seal plain uncoated slides; Becton Dickinson)
No. 2 pencil or slide marker
Slide boxes (optional)

Embed tissue

1. Place a labeled empty cryomold on dry ice in a container for 1 min. Keep on dry ice during the entire embedding procedure.
2. Cover the bottom of the cryomold with ~2 to 3 mm embedding medium.
3. Place the tissue to be frozen against the bottom of the cryomold in the medium before it hardens (this may take <1 min depending on the amount of OCT used).

To facilitate cutting, the tissue should be relatively small (i.e., 1 cm in maximum dimension) and the desired cutting surface should be flush against the bottom.

4. Fill the cryomold containing the base of embedding medium and frozen tissue with more embedding medium. Cover the dry ice container and allow the embedding medium to harden (several minutes).

The medium will turn from translucent to white when frozen.

5. Wrap the resulting tissue block, still in the cryomold, in aluminum foil and keep in a -80°C freezer or in liquid nitrogen until cutting.

Tissue for RNA extraction should be frozen as quickly as possible after resection. The method described here is preferred for LCM because tissue processed in this manner is more amenable to cryostat sectioning and offers acceptable histomorphology. More rapid methods of freezing tissue, such as direct immersion into liquid nitrogen, isopentane chilled to -160°C (Sheehan and Hrapchak, 1987a), or the vapor phase of liquid nitrogen, can also be utilized; however, these methods are more technically difficult when incorporating cryostat embedding media and are more likely to result in cracking of the tissue block. Tissues that were rapidly frozen without embedding medium can be postembedded in cryostat embedding medium, but will thaw somewhat in the process. This can compromise RNA preservation and introduce undesirable histologic artifacts.

Section tissue

6. Remove the tissue block from the cryomold and attach it to the tissue platform (chuck) in the cryostat, with additional embedding media serving as the “glue” at the interface. Apply just enough embedding media to cover the surface of the chuck and quickly attach the frozen tissue block before the “glue” hardens completely.

The cutting surface should be as parallel as possible to the chuck surface.

7. Allow the block to equilibrate to the cryostat temperature (i.e., -20°C) ≥ 15 min.
8. Cut 5- to 10- μm sections onto glass slides that have been sitting at room temperature and previously labeled with identifying numbers and or letters using a no. 2 pencil or a permanent marker designed for labeling slides.

Glass slides can be plain uncoated, charged, or silanized. The properties of glass slides that allow tissue adherence are variable among different brands, even with plain uncoated slides. It is important to use slides that allow tissue sections to adhere well enough that they do not fall off during staining, but not so tightly that the tissue cannot be captured. It is likely that different brands and types of slides will have to be tried, and that slides used successfully for formalin-fixed paraffin-embedded sections may not be optimal for frozen ones. The authors have found the Becton Dickinson Gold Seal plain uncoated slides work well for LCM of frozen sections in their laboratory. It is best to begin with plain uncoated slides, and if tissue sections do not adhere well enough to allow staining, to try charged or silanized slides. Adhesives, such as Sta-On (Surgipath) can be applied directly to the slides or gelatin, or can be added to the water bath during histologic sectioning; however, these may limit the transfer efficiency of LCM.

It is important to mount the tissue as close to the center of the slide as possible. If the tissue is too far off center, the slide cannot be positioned so that the vacuum slide holder can function during microdissection.

If sections are particularly friable and thus difficult to cut, the tissue may be too cold; therefore, the time allowed for the block to equilibrate to -20°C may need to be extended. Sections should be without folds and lie as flat as possible on the slides.

Sections with $>10\text{-}\mu\text{m}$ thickness are difficult to visualize. The authors prefer sections of 5- to $6\text{-}\mu\text{m}$ thickness. Thicker sections will require a larger spot size and therefore a higher laser-energy level.

9. Keep the slides in the cryostat or on dry ice if LCM is to be performed that day. Alternatively, store in slide boxes at -80°C until needed.

The duration of preservation of RNA and protein in frozen sections at -80°C is not well documented and likely depends on the tissue and the desired analyte. Although storage over several weeks or even months at -80°C may preserve the analyte of interest well, if this has not been assessed, we recommend limiting storage of frozen sections prior to microdissection to one week.

10. Stain slides (see Basic Protocol 5) just prior to LCM.

IMPORTANT NOTE: *Do not allow the slides to dry or thaw at room temperature prior to staining and dehydration. This is critical for successful LCM. Drying and thawing causes the tissue to adhere tightly to the slide and will decrease the transfer efficiency of LCM. Additionally, it may contribute to the degradation of RNA.*

PREPARATION OF FIXED PARAFFIN-EMBEDDED SECTIONS

Paraffin embedding is a process in which fixed tissue—utilizing neutral buffered formalin (NBF) or another fixative—is infiltrated and then placed into liquefied paraffin to stabilize it for long-term storage and easy sectioning (UNIT 14.1). While fixation is performed to preserve the morphology of the tissue for histologic examination, it also effects the DNA, RNA, and protein content. Formalin fixation is the standard for morphologic preservation of tissue and has been used by most pathology laboratories for decades; however, it creates cross-links between nucleic acids and proteins, and between different proteins. This cross-linking interferes with recovery of DNA, RNA, and proteins from fixed tissue, as well as the amplification of DNA and RNA by PCR (Arnold et al., 1996; Coombs et al., 1999; Goldsworthy et al., 1999; Masuda et al., 1999); however, short lengths of DNA, up to ~ 200 bp, can be reliably amplified after extraction from formalin-fixed paraffin-embedded (FFPE) tissue. RNA is a more labile species, and formalin fixation and paraffin embedding greatly interfere with its recovery. Attempts to break cross-links and thereby improve recovery of nucleic acids and protein have been utilized with varying degrees of success (Ikeda et al., 1998; Coombs et al., 1999; Masuda et al., 1999). Optimization and standardization of methods to break the cross-links caused by formalin fixation is a goal

BASIC PROTOCOL 2

Discovery of
Differentially
Expressed Genes

25A.1.3

of many researchers. Studies have shown that, among commonly used fixatives, formalin has the worst effects on RNA, while ethanol (i.e., 70% or 95% ethanol) or ethanol-based fixatives, available from suppliers of histology-related materials (e.g., Richard-Allan Scientific), offer the best RNA preservation (Goldsworthy et al., 1999; Shibutani et al., 2000).

In this protocol, it is assumed that most researchers will procure fixed and embedded tissue from pathology laboratories or other sources and may have no control over fixation and processing of tissues; however, a suggested protocol for fixation and tissue processing (see Support Protocol) has been included in the event the researcher is prospectively collecting human or animal tissues and has some degree of control over these processes.

Materials

Paraffin-embedded tissue block mounted on appropriate microtome chuck (see Support Protocol)

Xylene

100%, 95%, and 70% ethanol

Microtome and microtome blades (disposable preferred; Richard-Allan Scientific), clean

43° to 44°C water bath

Histologic slides, plain uncoated, charged, or silanized

37° to 42°C oven (optional)

Coplin jars or other solvent containers

Section tissue

1. Cut 5- to 10- μ m sections of a paraffin-embedded tissue block mounted on an appropriate chuck on a clean microtome with a clean blade.

IMPORTANT NOTE: Careful attention should be given during sectioning and mounting of paraffin-embedded tissue to prevent carryover. Carryover contamination of one specimen from another or transfer of material from one region of a section to another can lead to spurious results. The microtome used to cut sections should be kept clean and excess paraffin and tissue fragments should be wiped from the area with a simple gauze pad. A fresh microtome blade should be used for each block and disposable blades used if possible.

Sections of 5- μ m thickness are optimal for LCM, but the thickness should be dependent on the size of the cells to be microdissected.

2. Float resulting paraffin ribbons on 43° to 44°C deionized water in a water bath to smooth out and eliminate folds and wrinkles.

The water should be changed frequently to avoid contamination of sections by tissue fragments from other tissues and to minimize growth of environmental microorganisms. The authors currently do not recommend using formalin-fixed paraffin-embedded tissue for RNA analysis; however, the authors and others have successfully performed RT-PCR on alcohol-fixed paraffin-embedded tissues. If sections will be microdissected for RNA, consideration should be given to using RNase-free water (UNIT 4.1).

Some histopathology laboratories use an adhesive in the water bath to better adhere the tissue section to the slide. As this may result in reduced LCM transfer of tissue, it is not recommended.

3. Mount sections on histologic glass slides.

Clean uncoated plain, charged, or silanized histological slides can be used. The authors have successfully performed LCM utilizing many brands of uncoated glass slides, as well as charged slides, with fixed and paraffin-embedded tissues.

It is important to mount the tissue as close to the center of the slide as possible. If the tissue is too far off center, the slide cannot be positioned so that the vacuum slide holder can function during microdissection.

- Air dry the paraffinized sections overnight or bake up to 8 hr at 37° to 42°C.

As with frozen sections, the desired result is for the tissue to remain adherent to the slide during staining, but not be so adherent as to prevent tissue transfer by LCM. Baking the slides will cause the sections to be more adherent than air drying. Relevant variables that affect LCM include the type of slide, whether the sample is air dried or baked, the duration of baking, and the type of tissue being microdissected.

Remove paraffin

- Allow the slide containing the tissue section to remain in the following solutions, in Coplin jars or other solvent containers, for the specified times in the specified order:

Xylene	5 min
Xylene	5 min
100% ethanol	30 sec
95% ethanol	30 sec
70% ethanol	30 sec

In order to proceed with histologic staining and LCM following sectioning, paraffin must be removed from the tissue sections.

If RNA is to be analyzed, consideration should be given to preparing the 95% and 70% ethanol solutions with RNase-free water. The authors routinely utilize sterile or distilled water and typically achieve good RNA recovery.

- Proceed with hematoxylin and eosin staining (see Basic Protocol 5).

PREPARATION OF CYTOLOGIC SPECIMENS FOR LCM: DIRECT SMEARS

Cellular elements in body fluids or fine-needle aspirates and cultured cells do not readily lend themselves to sectioning, but can easily be prepared for LCM by making direct smears or cytopsin preparations. The choice as to which to use will depend upon the anticipated cellularity of the sample. Highly cellular samples can be easily and rapidly prepared as direct smears and effectively utilized for LCM, whereas less cellular samples are better concentrated and prepared as cytopsin preparations. To determine if the sample requires concentration, make a direct smear as described below and examine it under the microscope. If the concentration of cells is such that the desired number of cells for LCM can be located in 1 to 4 areas each with a diameter of 0.5 cm (the appropriate diameter of the “cap” used to capture the cells of interest during LCM), the specimen does not require concentration. If however, the concentration of cells is so low that the number of desired cells is not present or the cells are so widely spaced that it will require five or more caps to obtain them, specimen concentration is recommended. For specimens contaminated with undesired blood elements (i.e., red blood cells or white cells that are not intended to be microdissected), use the protocol for cytologic smears or cytopsin preparations containing excessive blood as the contaminant (see Alternate Protocol 1). The same basic caveats apply to cytologic specimens as histologic sections—i.e., ethanol is the preferred fixative (especially for RNA analysis), the cells should never be allowed to dry on the slide prior to fixation, and the fixed and stained cells should be adequately dehydrated prior to LCM.

Materials

- High-cellularity sample: cellular fluid (e.g., fine-needle aspiration, suspended cultured cells) or fresh tissue
- 95% ethanol
- Hemocytometer cover (optional)
- Glass slides, clean
- Scalpel blade (fresh tissue)

BASIC PROTOCOL 3

Discovery of Differentially Expressed Genes

25A.1.5

- 1a. *For cellular fluid:* Place a drop of cellular fluid (i.e., fine-needle aspiration samples or cultured cells suspended in medium), no larger than 5 mm in diameter, towards the label end of a clean glass slide. Quickly utilize the edge of another glass slide, or preferably a hemacytometer cover, to thinly spread the drop (i.e., as if making a blood-smear preparation) on the slide in a single motion, relying on capillary action between the liquid and the two slides to spread the liquid in a uniform, thin-layer across the length and width of the slide. Do not apply excessive force which might result in crushing or shearing of cells.

Plain uncoated, charged, or silanized glass slides can be used.

We prefer to prepare cytologic smears with a hemocytometer cover because its width is slightly less than that of the standard glass microscopic slide and the resulting smear (i.e., cells) is not spread to, or off, the edge of the slide.

- 1b. *For fresh tissue:* Quickly sample by scraping tissue with a scalpel blade and then rapidly spread the scraped sample on a glass slide with the blade.

This is a quick and useful method of specimen preparation for tissues in which the desired cells can be readily identified cytologically, such as highly malignant cells.

2. Immediately after spreading, immerse the smear in 95% ethanol without allowing it to dry. Incubate 10 min.
3. Transfer to 70% ethanol for 30 sec.
4. Proceed to hematoxylin and eosin staining (see Basic Protocol 5).

BASIC PROTOCOL 4

PREPARATION OF CYTOLOGIC SPECIMENS FOR LCM: CYTOSPIN METHOD

Cytospin preparations can be used for any cytologic sample but are preferred for samples of low cellularity. Cytospin instrumentation allows cellular fluids to be simultaneously concentrated and placed on a glass slide. Using centrifugation, these instruments spin cell suspensions onto a microscope slide as the suspension medium is simultaneously absorbed by a blotter. The result is a monolayer of well-preserved well-displayed cells within a 6-mm² area on the slide. Another alternative for samples of low cellularity is to centrifuge the sample, decant the supernatant, and make a direct smear (see Basic Protocol 3) from the sediment. Particularly bloody specimens may benefit from the protocol provided below (see Alternate Protocol 1). To avoid RNA, DNA, or protein degradation, the cytologic samples should be processed and fixed in 95% ethanol shortly after collection. Microdissection after fixation is preferable, particularly for RNA analysis.

Materials

- Low-cellularity sample: fine-needle aspiration or cultured cells suspended in medium
- 95% and 70% ethanol
- Cytospin instrument and appropriate single sample chamber cytospin device (e.g., Shandon/Lipshaw)
- Glass slides, clean

Assemble and load cytospin devices

1. Assemble the sample chamber cytospin device with clean glass slides according to the manufacturer's instructions.

Plain uncoated, charged, or silanized glass slides can be used.

2. Load the assembled collection chamber devices into the support plate of the cytopsin instrument.

They must be secure, freely tiltable, and symmetrically distributed.

Add samples and spin

3. Pipet low-cellularity sample into sample chambers.

The optimal amount of specimen will vary with its cellularity. Samples of low cellularity will require 300 to 400 μ l per chamber; highly cellular samples will require only 100 to 200 μ l per chamber.

4. Press closure cap on each sample chamber.
5. Lock the lid of sealed head and close the cytopsin cover.
6. Program cytopsin for 3 min at 1500 rpm on high acceleration and press start.

Rapidly fix cytopsin

7. When the alarm signaling the end of the spin sounds, quickly remove the assembled collection chambers. Open the chambers and remove the slides by lifting the blotter away from the slide

This method avoids damage of cell membranes and thus smearing.

8. Quickly transfer slide into 95% ethanol without allowing the specimen to dry. Fix 10 min. Transfer slide to 70% ethanol for 30 sec.
9. Proceed to H&E staining (see Basic Protocol 5) or other stain of choice.

REMOVING BLOOD FROM SAMPLES FOR CYTOLOGIC SMEARS OR CYTOSPINS

ALTERNATE PROTOCOL 1

Particularly bloody specimens may benefit from separating red blood cells from other cellular elements, thereby concentrating the desired cells (especially epithelial cells). This can be accomplished by utilizing the Ficoll-Paque density gradient technique described here. The specimen is layered onto an undiluted Ficoll-Paque solution and centrifuged. Differential migration during centrifugation results in the formation of layers enriched in different cell types. This allows extraction of other cells in the sample from red blood cells. This method is not ideal for isolating white blood cells for microdissection as many of them separate with the red blood cells. See the Arcturus Engineering web site (<http://www.arctur.com>) for a protocol for isolating the buffy coat of blood.

Materials

Cytologic sample
Sterile saline (i.e., 0.9% w/v NaCl) or balanced salt solution
Ficoll-Paque (Pharmacia)
50-ml centrifuge tubes

Concentrate cellular components

1. Centrifuge the cytologic sample for 10 min at $350 \times g$, room temperature, in a 50-ml centrifuge tube.
2. Aspirate the supernatant with a pipet.
3. Resuspend the cell "button" in 5 to 10 ml sterile saline or balanced salt solution.

**Discovery of
Differentially
Expressed Genes**

25A.1.7

Separate cellular components

4. Add 20 ml Ficoll-Paque to a clean-50 ml centrifuge tube. Carefully pipet the cell suspension onto the Ficoll-Paque.

It is best not to mix the Ficoll-Paque with the specimen at this point.

5. Centrifuge 10 min at $350 \times g$, room temperature.

After centrifugation, the top and clearest layer contains any epithelial cells and some white blood cells. The middle layer is the Ficoll-Paque and the lowest layer is predominantly red blood cells and white blood cells.

6. Prepare the superficial cell layer as direct smears or cytopins (see Basic Protocols 3 and 4).

HEMATOXYLIN AND EOSIN STAINING

Histologic section and cytologic preparations must be stained so that the component cells can be adequately visualized for accurate identification; hematoxylin and eosin stain is commonly used for this purpose. With this stain, nuclei are black-blue and cell cytoplasm and most extracellular material are varying shades of pink. Although both hematoxylin and eosin staining solutions can be prepared from their basic components, the authors recommend purchasing prepared, ready-to-use stains.

Materials

Sample on a glass slide (see Basic Protocols 1 to 4)

70%, 95%, and 100% ethanol

Sterile, distilled, or RNase free water

Mayer's hematoxylin (Richard-Allan Scientific)

Bluing reagent (Richard-Allan Scientific)

Eosin Y

Xylene

1. *For frozen sections (optional):* Rapidly remove the sample on a glass slide from -80°C storage (see Basic Protocol 1) and immerse in or flood with 70% ethanol without allowing the slide to thaw and dry prior to contact with the ethanol. Allow the ethanol to remain in contact with the tissue for 30 sec.

Deparaffinized fixed sections (see Basic Protocol 2) as well as samples prepared by direct smear or cytopsin (see Basic Protocols 3 and 4) will already be in 70% alcohol and are ready to proceed through the following steps.

2. Allow the slide containing the tissue section to remain in the following solutions for the specified times in the specified sequence:

Sterile, distilled, or RNase-free water	10 sec
Mayer's hematoxylin	10 sec
Sterile, distilled, or RNase-free water	10 sec
Bluing reagent	15 to 30 sec
70% ethanol	15 to 30 sec
Eosin Y	15 to 30 sec
95% ethanol	30 sec
95% ethanol	30 sec
100% ethanol	30 sec
100% ethanol	30 sec to 1 min
Xylene	1 to 5 min

3. Allow the section to air dry completely and proceed to LCM (see Basic Protocol 6 and Alternate Protocol 2).

Poor LCM transfers will result if the tissue section is not fully dehydrated. This may result if the 100% ethanol becomes hydrated after repeated use. One way to check the 100% ethanol for water is to put a small amount into xylene. If there is water present, the xylene will become cloudy. The final xylene rinse also facilitates the efficiency of transfer with LCM. If a tissue section does not transfer well, repeating the dehydration with fresh 100% alcohol and/or a longer xylene rinse may help. While other staining protocols can be used, the slides should be dehydrated with graded alcohols and the final xylene step.

LASER CAPTURE MICRODISSECTION

The described procedure is for the PixCell I or II Laser Capture Microdissection System and assumes a general knowledge of the function of the components of the instrument and the software that accompanies the instrument. The general theory underlying the use of the instrument is discussed elsewhere (see Background Information). The procedure can be divided into three basic steps: slide positioning, microdissecting with the laser, and collecting the microdissected cells. Additional information about the Arcturus LCM software, including capturing and storing images, and additional instruction for LCM, can be found in the instrument users' manual and at the Arcturus Engineering web site (<http://www.arctur.com>), the National Institute of Environmental Health Sciences web site (<http://dir.niehs.nih.gov>), or from Arcturus technical support (650-962-3020).

Materials

- Glass slide with stained specimen (see Basic Protocol 5)
- Appropriate lysis buffer (e.g., DNA lysis buffer, protein lysis buffer; see recipes)
- PixCell I or II Laser Capture Microdissection System (Arcturus Engineering)
- Arcturus LCM software (Arcturus Engineering; optional)
- CapSure transfer film (Arcturus Engineering)
- 0.5-ml microcentrifuge tubes (Eppendorf)

NOTE: Wear gloves when microdissecting to avoid contamination of the LCM specimens. Clean the microscope stage and capping station before beginning the microdissection (e.g., use 95% ethanol wipes), to reduce the possibility of contamination.

Position slide (section) to be microdissected

1. Turn on the PixCell I or II Laser Capture Microdissection System. Open the Arcturus LCM software if it is to be used.

The Arcturus LCM software is not required for LCM as all adjustments of parameters can be made on the laser electronics box; however, it eases the use of the instrument and performs useful functions, such as counting the pulses of the laser ("shots") and allowing the procurement and archiving of images.

2. Place the glass slide with the stained section to be microdissected on the microscope stage. Move the joystick so that it is perpendicular to the tabletop to allow proper placement of the CapSure transfer film ("cap"). Focus the microscope to view the tissue or cells. Locate the area to be microdissected, moving the slide by hand rather than with the joystick, so that the joystick will be in proper alignment when the area to be microdissected is located.

Samples are usually stained in order to be visualized for LCM; however, LCM can be performed successfully without staining, but desired cells may not be identifiable.

The area selected should be located such that a portion of the slide covers the vacuum chuck hole and the slide spans the central hole in the stage.

BASIC PROTOCOL 6

Discovery of Differentially Expressed Genes

25A.1.9

3. Turn on the vacuum slide holder.

IMPORTANT NOTE: *The joystick should now be used to move the slide.*

4. Use the visualizer to more precisely locate the cells to be microdissected.

The light from the microscope will need to be increased when using the visualizer.

The area to be microdissected should be in the field of view.

The sections are not coverslipped; therefore, the area of interest may be difficult to visualize. All models of the PixCell System are equipped with a visualizer which acts to diffuse light and improves resolution; however, the visualizer is engaged differently on different models (see instrument user's guide).

Microdissect with the laser

5. Pick up a cap from the loaded cassette module on the right side of the microscope stage (see instrument users' guide for instructions on loading the caps into the cassette module) with the placement arm. Swing the placement arm toward the caps until the arm overrides the first cap in the cassette module. Ensure that the cassette module is engaged in the proper indent so that the first available cap is aligned with the arrow on the microscope stage. Lift the transport arm until the cap detaches from the base slide in the cassette module.

6. Without lowering the placement arm, swing the arm back toward the tissue section as far as possible, so that the arm is over the tissue. Make sure that the area to be microdissected is still in the microscopic field of view by looking through the microscope eyepieces or at the monitor. Gently lower the arm so that the cap contacts the tissue section.

If there are folds in the tissue, the cap may not make direct contact with the entire surface in the area to be microdissected, and transfer efficiency will be compromised; therefore, it is advisable to inspect the tissue before placing down the cap. If any tissue is mounded or folded, it is best not to place the cap over that area. Alternatively, the area of the tissue with folds can be scrapped off the slide using a sterile razor blade, leaving only flat portions of the tissue section. The tissue section must be dry and cannot be coverslipped for LCM transfer.

7. Enable the laser by turning the key on the laser electronics box and pushing the laser-enable button.

The laser-tracking beam should now be visible on the monitor, as well as the area to be microdissected. If it is not, try lowering the light from the microscope or raising the intensity of the tracking beam. If it is still not visible, check that the laser is enabled and that the joystick is perpendicular.

Avoid passing hands through the path of the laser when it is enabled.

8. Using the 20× objective, adjust the focus of the tissue by moving the slide via the joystick to an area of the slide without tissue. Adjust the laser spot size to 7.5 μm. Lower the light from the microscope until there is a black monitor screen, except for the tracking beam. Turn the laser focusing wheel until the tracking beam is a bright spot with a well-defined edge.

There should be no bright rings surrounding the central spot (Fig. 25A.1.1).

Always focus the laser with the 7.5-μm spot. Each tissue section and slide will need to be refocused. Once the 7.5-μm spot is focused for a particular slide, there is no need to refocus the 15-μm or 30-μm spots, as they are automatically calibrated.

9. Adjust the laser power and pulse duration settings for the particular spot size to be used as provided below:

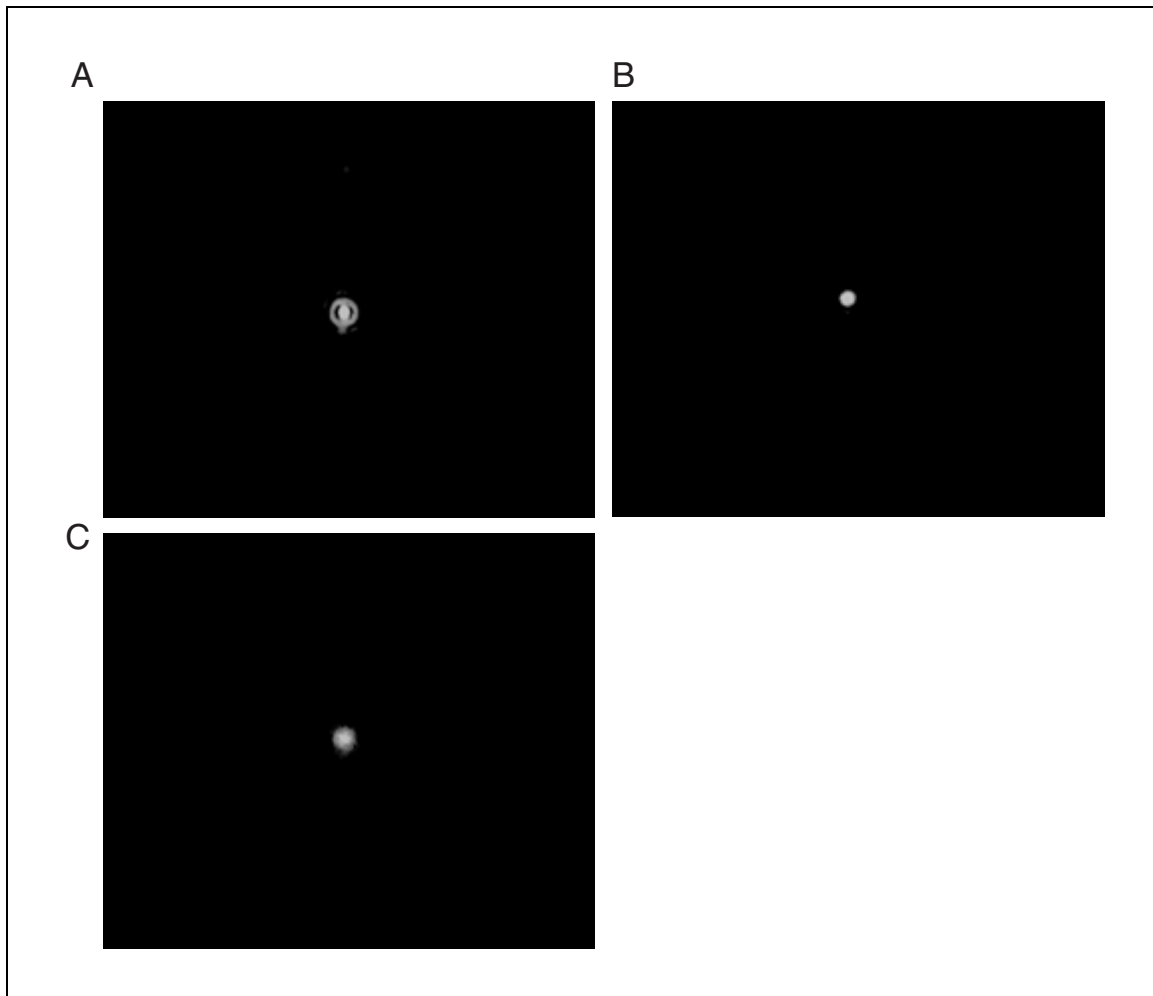


Figure 25A.1.1 Focusing the Laser Beam. **(A)** Unfocused beam, the spot of light has concentric halos of light. **(B)** Focused beam, the spot of light has a sharp border without halos of light. **(C)** Unfocused beam, the spot of light has a blurred border. A 20 \times objective and 7.5- μm spot size is used in all three pictures.

Spot size	Power	Duration
7.5 μm	40 mW	450 μsec
15 μm	25 mW	1.5 msec
30 μm	20 mW	5 msec

Laser power and duration determine the spot size. The power and duration settings given above should provide a melted area that is similar in size to the tracking beam at each of the three settings, but may require adjustment. See the user's manual for more information.

- While the tracking beam is still located in an area without tissue, fire the laser by clicking the red button on the remote thumb switch to assess the effectiveness of the laser focus and settings.

Effective melting ("wetting") of the polymer on the lower surface of the cap is indicated by a circle with a well defined black outline (see Fig. 25A.1.2).

If the edges of the circle are not well delineated, check to make sure that the tissue section where the cap is placed is flat and refocus the beam. If this fails, increase the power and/or duration gradually and as little as possible (see Troubleshooting).

- Test the effectiveness of LCM in the tissue section by moving the tracking beam to the cells to be microdissected. After targeting the cells, fire the laser. Move the slide

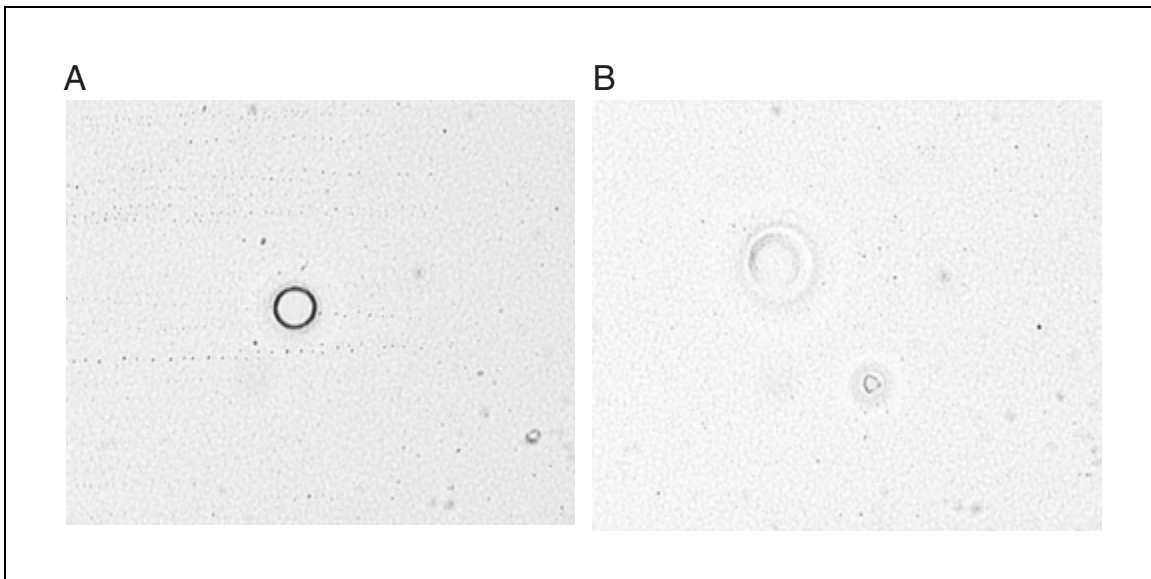


Figure 25A.1.2 Polymer Melting After Laser Firing. **(A)** An adequate and effective melt has a sharp, delineated border. **(B)** The border of an inadequate melt is blurred and indistinct.

with the joystick to another group of cells and fire the laser again. Limit the number of pulses for this test to two or three.

The delineation of the circle may be more difficult to visualize on the tissue section, but the tissue in an area of proper “wetting” should become more sharply focused because the melted polymer acts as a coverslip. Lift the placement arm and inspect the area in which the laser was fired for removal of cells (see before and after photomicrographs in Fig. 25A.1.3).

If the LCM was successful, the area where the polymer was melted should no longer be occupied by tissue and should be empty, although a small amount of cellular and stromal material may remain. The great majority of the tissue that occupied those spots should now be attached to the cap. This can be checked by releasing the vacuum slide holder, moving the slide so that a clean area without tissue is in the microscopic field of view, lowering the cap to the slide, and scanning the surface of the cap. The microdissected tissue should be visible on the cap surface. If this is not the case, there are several explanations and potential remedies (see Troubleshooting).

Avoid lifting and lowering the cap repeatedly after firing the laser and capturing some tissue. It is difficult to replace the captured tissue in the exact spot from which it came. Consequently the captured tissue, and tissue that may nonspecifically stick to the cap, will be placed on the histologic section, resulting in a layering effect which can limit contact of the cap with the tissue and compromise the effectiveness of LCM; therefore, limit the number of shots used to test the adequacy of capture, and, if the test capture is successful, avoid lifting the cap again until the microdissection is complete.

Dense, dark or thick samples may occlude the tracking beam. If this occurs, increase the intensity of the tracking beam.

12. Once LCM is achieved successfully with the test pulses, proceed to microdissect the remainder of the desired cells.

Collect microdissected cells

13. After completing the intended microdissection, lift the placement arm. Assess the completeness of the capture by inspecting the microdissected tissue and the cap as described above.

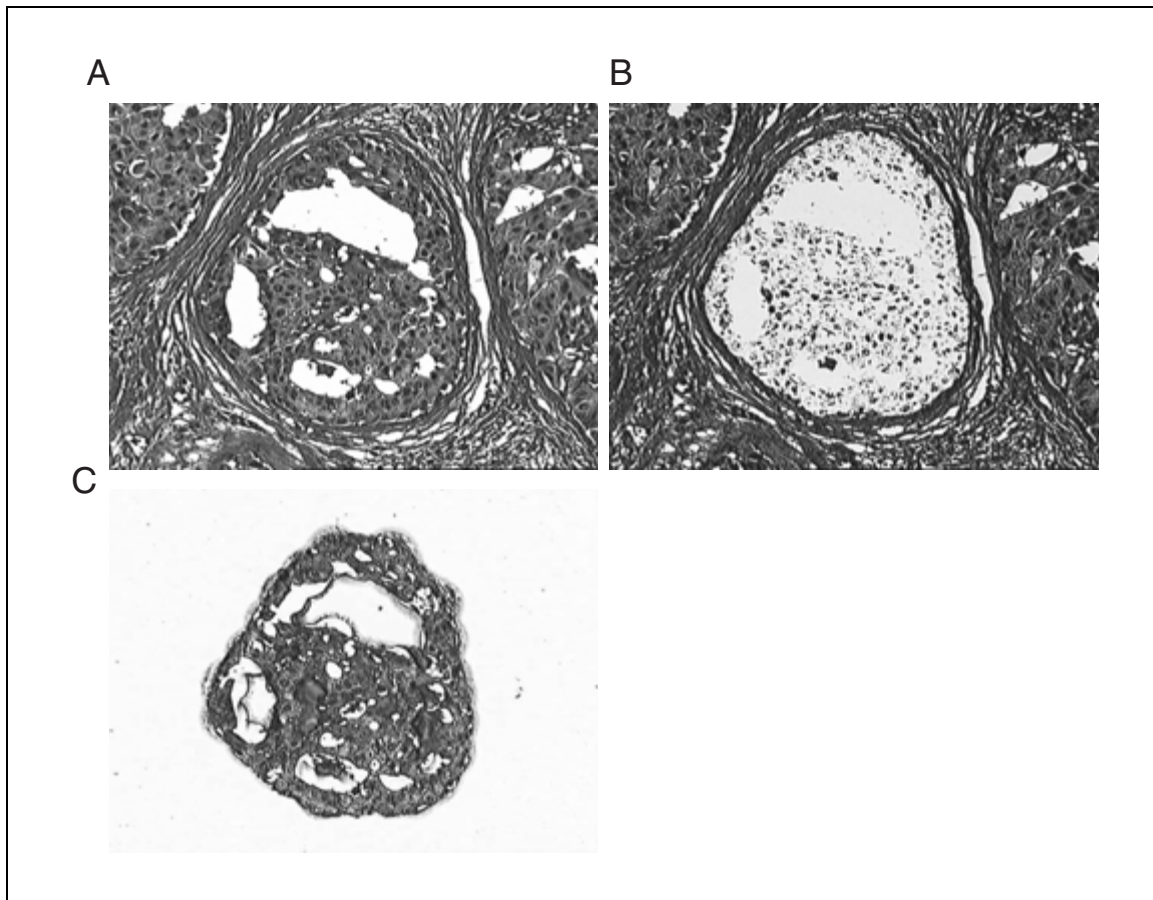


Figure 25A.1.3 LCM of ductal carcinoma in situ. (A) The area of ductal carcinoma in situ prior to LCM. (B) The same focus after LCM. (C) The microdissected focus on the transfer film (cap).

14. Swing the placement arm with the cap towards the right to the unload platform and place the cap in the designated slot. Move the placement arm, without lifting it, to the left and place in a resting position.
15. Using the cap insertion tool, pick up the cap from the unload platform by sliding the insertion tool along the guide rail until the cap is engaged in the tool. Remove the cap from the unload platform by lifting the insertion tool.

The open end of the insertion tool should face the cap.

Because tissue and cells that were not selected for capture may nonspecifically stick to the surface of the cap, it is important to remove this unwanted tissue. This can be accomplished by using the CapSure Pads (Arcturus Engineering), which have a sticky surface. If using the CapSure Pad, place the pad on the microscope stage in the path of the placement arm prior to placing the cap on the unload platform. Move the placement arm over the pad, lower the cap, and raise the pad to contact the cap. Raise the placement arm and the cap while holding the pad in place with your hand. A less costly alternative to the CapSure Pad is to use the sticky surface of Post-It Notes (3M). The Post-It Notes can be used after the cap has been removed from the unload platform. Peel a fresh Post-It Note off the pad and lower the cap, loaded into the insertion tool, to contact the sticky surface of the Post-It Note. Repeat this 2 to 3 times.

16. Using the insertion tool, insert the cap into a 0.5-ml microcentrifuge tube containing an appropriate amount of lysis buffer (e.g., DNA or protein lysis buffer), usually between 50 and 100 μ l. Press down firmly and rotate the insertion tool to ensure an even seal.

The choice of lysis or digestion buffers is dependent on the analyte and the method of analysis. The recipes supplied in this unit (see Reagents and Solutions) provide examples of lysis buffers for DNA and protein that can be used for LCM samples. Other buffer recipes can be found in many of the references provided and at the BioProtocol web site (<http://www.bioprotocol.com>); however, it is best to customize the buffer to the methodology of the specific laboratory. The authors prefer to use Trizol (Life Technologies) or Stat-60 (Tel-Test) for cell lysis and RNA stabilization prior to RNA extraction and have not provided a recipe for an RNA lysis buffer; however, other buffers containing guanidine thiocyanate and 2-mercaptoethanol can also be used.

The caps fit well in standard 0.5-ml microcentrifuge tubes. When properly seated, the cap does not sit down fully in the tube, but should be seated evenly. Capped tubes will leak if the cap is pushed all the way down into the tube so that the top portion of the cap touches the lip of the microcentrifuge tube.

17. Invert the tube so that the lysis buffer contacts the cap surface. Flick the tube to move the lysis buffer to the cap surface, if necessary.

Place on ice or refrigerate until the microdissection session is over, if this will help to preserve the analyte in the chosen lysis buffer. This sample is now ready to be processed by appropriate methods for the analyte of interest.

ALTERNATE PROTOCOL 2

LASER CAPTURE MICRODISSECTION OF SINGLE OR A SMALL NUMBER OF CELLS

Arcturus Engineering has developed a line of related consumables that are specially designed for high-sensitivity capture and extraction of a single cell or a minimal number of cells. There are three key components of the system: a preparation strip that flattens the tissue section and removes loose debris, the high-sensitivity transfer cap (HS cap) that keeps the tissue surface area adjacent to the cells being captured out of contact with the sample, and a low-volume reaction chamber that fits onto the high-sensitivity transfer caps and accepts a low volume of lysis or digestion buffer while sealing out any nonselected material from the captured cells. The HS cap has a raised ridge on the contact surface so that only the ridge actually touches the tissue section. The surface coated with polymer only contacts the tissue in the area in which the laser is fired; thus, contamination by unwanted tissue is greatly reduced.

The basic steps of LCM as described (see Basic Protocol 6) are applicable to the use of the high-sensitivity consumables, with a few modifications. The modifications to the standard LCM protocol are described briefly below. These products can be purchased as a kit from Arcturus Engineering, which includes detailed instructions on their use.

Additional Materials (also see Basic Protocol 6)

- Preparation strips (Prep Strips; Arcturus Engineering)
- High-sensitivity transfer film (HS CapSure; Arcturus Engineering)
- Tweezers, clean
- Alignment tray designed for use with the high-sensitivity system
- Low-volume reaction chamber (ExtracSure; Arcturus Engineering)

NOTE: All pipetting steps should be performed using filtered aerosol-resistant pipet tips.

Position slide

1. Prior to placing the stained sample on the glass slide on the microscope stage, apply a preparation strip (Prep Strip) to the tissue section or sample to flatten the tissue and remove loose debris.

2. Position the slide as described in the basic LCM protocol (see Basic Protocol 6, steps 1 to 4).

Microdissect

3. Pick up a high-sensitivity transfer film (HS cap; e.g., HS CapSure) from the loaded cassette module on the right side of the microscope stage (see instrument user's guide for instructions on loading the caps into the cassette module) with the placement arm and position the HS cap on the tissue to be microdissected. Enable and focus the laser as previously described (see Basic Protocol 6, steps 5 to 8).
4. Begin at a starting power of 75 mW and a pulse duration of 1 msec and make adjustments to the spot size by changing the duration setting rather than the power.

These settings are those recommended for high-sensitivity LCM.

For the smallest spot size, keep the duration and power settings low but pulse multiple times at the same target to ensure capture and transfer.

The laser activates the transfer film, which then expands down into contact with the tissue. It is preferable to capture cells as close to the center of the cap as possible. Unlike basic LCM using the standard caps, the HS caps can be repositioned as often as needed to keep the targets toward the center of the cap, because the cap surface does not contact the tissue except at the area that the laser is fired. It is important to stay within the capture ring because areas outside the ring will be excluded from the low volume reaction tube.

5. Test the effectiveness as described (Basic Protocol 6, step 11).

Collect microdissected cells

6. After completing the intended microdissection, place the HS cap on the unload platform and pick up the HS cap with the cap insertion tool.
7. Remove the HS cap from the insertion tool using clean tweezers and place the HS cap into the alignment tray so that the captured sample is facing up.
8. Using clean tweezers, position the specialized low-volume reaction chamber over the cap.

The chamber has a port for insertion of the appropriate lysis buffer (e.g., DNA or protein lysis buffer), which should be facing up.
9. Push the chamber down onto the cap until it snaps into place.
10. Pipet 10 μ l desired buffer into the fill port. Cover the port with a 0.5-ml microcentrifuge tube or thin-walled PCR tube and press down to fit securely.
11. Proceed to extraction and analysis of the desired analyte.

TISSUE FIXATION AND PARAFFIN-EMBEDDING

If the researcher can choose a fixative, one which is alcohol based (e.g., 70% ethanol) is preferable for nucleic acid and protein recovery, and provides adequate morphologic detail for most LCM uses; however, alcohol-based fixatives have been reputed to confer a shrinkage artifact in histologic sections that is undesirable to diagnostic pathologists, as it results in tissue that is difficult to section and, at low dilutions, is inadequate for long-term storage of tissues (Vardaxis et al., 1997). On the other hand, Bostwick et al. (1994) successfully utilized an alcohol-based fixative in their pathology laboratory for one year without reporting these difficulties. Fixed tissue is typically embedded in paraffin to stiffen it so that thin histologic sections can be cut. Most paraffin used in pathology laboratories melts at $\sim 60^{\circ}\text{C}$, which may accelerate formaldehyde reactions and damage RNA, DNA, and proteins; therefore, waxes or paraffins that have a lower melting point can be used, but they make softer tissue blocks that are more difficult to cut and may

SUPPORT PROTOCOL

**Discovery of
Differentially
Expressed Genes**

25A.1.15

require refrigerated storage. Tissue processing, embedding, and sectioning are generally performed in a histology laboratory by histotechnologists and generally require some degree of training and skill. The processing steps provided are suggested for utilization by histology laboratories processing tissue for LCM (<http://www.arctur.com>); however, other processing sequences may also provide good LCM results.

Materials

Fresh tissue
Fixative of choice (e.g., 70% ethanol)
Neutral buffered formalin (NBF; Richard-Allan Scientific)
70%, 80%, 95% and 100% ethanol
Xylene
Embedding paraffin

Tissue cassettes
Automated tissue processor
Embedding mold (Tissue-Tek)
Embedding center (optional; Leica)

Fix tissue

1. Place fresh tissue in a volume of fixative that is $\geq 10\times$ the tissue volume, so that the fixative surrounds the tissue on all sides.

Unfixed tissue that floats should be covered by a layer of gauze or paper towel to ensure the tissue is under the fixative. Fixation can be carried out at room temperature or 4°C.

Fixation at 4°C slows down the autolytic process and can be useful for larger specimens.

2. Fix the tissue for an appropriate amount of time.

The time required for fixation is dependent on the size of the tissue and the speed with which the fixative penetrates the tissue. Formalin and 95% ethanol penetrate at a rate of ~ 1 mm/hr. Fixation time and tissue size should be adjusted as necessary.

For any fixative used, a fixing period of 16 to 24 hr is recommended to provide complete tissue fixation; however, a fixation period of < 6 hr provides better recovery of DNA than longer fixation times (<http://www.arctur.com>).

3. *Optional:* Trim tissue sections from larger fresh or fixed tissue specimens so that they are no more than 3 mm in thickness and no larger than the dimensions of the cassette used for tissue processing. Place one section in each cassette.

Again, 1-cm maximum dimension is ideal.

Process and embed tissue

4. After the tissue sections in the cassettes are fixed, place the cassettes in the first station of an automated tissue processor. Program and load the processor.

- 5a. *For routine overnight processing:* Perform the steps in Table 25A.1.1.

After processing, the tissue will be infiltrated with paraffin.

- 5b. *For accelerated processing:* Follow the steps in Table 25A.1.2.

No difference has been found in the LCM transfer efficiency of tissues processed either way.

6. Remove the tissue from the original cassette and embed the paraffin-infiltrated tissue in additional melted paraffin in an embedding mold. Allow to cool and harden.
7. Adhere the paraffin block to a cutting platform (chuck) and remove the paraffin block from the embedding mold. The paraffin block is now ready for sectioning (see Basic Protocol 2).

Also see Sheehan and Hrapchak (1987b).

Table 25A.1.1 Routine Overnight Tissue Processing

Station	Solution	Concentration	Time (min)	Temperature (°C)
1	NBF ^a	10%	2:00	40
2	Ethanol	70%	0:30	40
3	Ethanol	80%	0:30	40
4	Ethanol	95%	0:45	40
5	Ethanol	95%	0:45	40
6	Ethanol	100%	0:45	40
7	Ethanol	100%	0:45	40
8	Ethanol	100%	0:45	40
9	Xylene	100%	0:45	40
10	Xylene	100%	0:45	40
11	Embedding paraffin	—	0:30	58
12	Embedding paraffin	—	0:30	58
13	Embedding paraffin	—	0:30	58
14	Embedding paraffin	—	0:30	58

^aIf neutral buffered formalin (NBF) is not the initial fixative, skip station 1.

Table 25A.1.2 Accelerated Tissue Processing

Station	Solution	Concentration	Time (min)	Temperature (°C)
1	Ethanol	70%	0:10	40
2	Ethanol	80%	0:10	40
3	Ethanol	95%	0:15	40
4	Ethanol	100%	0:20	40
5	Ethanol	100%	0:30	40
6	Xylene	100%	0:30	40
7	Xylene	100%	0:30	40
8	Xylene	100%	0:30	40
11	Embedding paraffin	—	0:30	60
12	Embedding paraffin	—	0:20	60
13	Embedding paraffin	—	0:30	60
14	Embedding paraffin	—	0:20	60

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

DNA lysis buffer

10 mM Tris·Cl, pH 8.0 (APPENDIX 2)

0.2% (v/v) Tween 20

100 µg/ml proteinase K

The authors use this lysis buffer for samples intended for PCR. Arcturus Engineering offers DNA extraction kits that were developed specifically for LCM specimens. The proteinase K should be stored at -20°C in aliquots, while the Tris·Cl and Tween 20 can be stored at -4°C . Once the proteinase K is thawed and added, the buffer should be used immediately.

**Discovery of
Differentially
Expressed Genes**

25A.1.17

Protein lysis buffer

10 mM Tris-Cl, pH 7.4 (APPENDIX 2)
0.1% Triton X-100
1.5 mM EDTA
10% (v/v) glycerol
Store several months at -4°C

This lysis buffer has been found to be useful for analysis of membrane-bound proteins (Simone et al., 2000). For cytoplasmic proteins, "T-Per" tissue protein extraction liquid reagent (Pierce Chemical) has been recommended (Simone et al., 2000). It has also been suggested that the addition of protease inhibitors, such as 4-(2-aminoethyl)-benzenesulfonyl fluoride (Boehringer Mannheim) to the buffer increases the yield of protein (Banks et al., 1999; Ornstein et al., 2000a).

COMMENTARY

Background Information

Technologic advances in gene sequencing and amplification techniques are allowing the identification of alterations in genes, proteins, and biochemicals that can explain the etiology and pathogenesis of many disease processes; however, the efficacy of these technologies depends on the identity and the purity of the cells being analyzed. Physical homogenization of tissues results in a mixture of many cell types—i.e., some are normal or minimally altered components, while others may be significantly diseased. Alterations detected in such homogenates cannot be localized to a particular cell type. Multiple mechanical methods for separating cells of interest from tissues have been described, especially as related to histologic sections (Sirivatanauksorn et al., 1999), but their methodology is time-consuming, extremely labor-intensive, and often imprecise. Laser Capture Microdissection (LCM) is one of the new generation of microdissection techniques that is relatively quick and precise.

LCM was conceived and first developed as a prototype research tool at the National Institute of Child Health and Human Development (NICHD) and the National Cancer Institute (NCI) of the National Institutes of Health (NIH). Arcturus Engineering and the NIH, working through a Cooperative Research and Development Agreement, developed LCM into a commercial laboratory instrument that is now utilized in many research laboratories. Other efficient microdissection techniques, such as laser pressure catapulting, have also been described (Bohm et al., 1997; Sirivatanauksorn et al., 1999).

With LCM, cells of interest are dissected from tissue sections or cytologic samples after microscopic identification with the aid of an ethylene vinyl acetate transfer film containing

a near-infrared absorbing dye. The transfer film coats a flat surface of an optically clear plastic cylinder, the "cap," with a diameter of 6 mm. The LCM system places the transfer film in contact with a histologic section and then directs an invisible infrared laser pulse onto the overlying polymer. The laser pulse is absorbed by and melts the transfer film causing it to flow around the targeted cells. The polymer rapidly cools and creates a bond between the transfer film and the targeted cells. The targeted cells can then be lifted from the section and utilized for RNA, DNA, or protein analysis (Fig. 25A.1.4). This targeting and capturing can be repeated many times on the same tissue section or cytologic sample. The temperature rise in the tissue created by the laser is limited to 90°C (Suarez-Quian et al., 1999) and is transient, lasting only a few milliseconds. Experimental results indicate that DNA, mRNA, and proteins are not degraded by the LCM process (Goldsworthy et al., 1999; Suarez-Quian et al., 1999).

Critical Parameters

LCM can be performed on solid tissues that have been either frozen or fixed under specified conditions, cytologic smears, or cytopsin preparations derived from animals or patient samples. The choice of specimen type depends on the type of tissue or cytologic specimen that is available, the physiologic or pathologic condition to be investigated, and the molecule to be analyzed (i.e., DNA, RNA, or protein). Solid tissues are typically sectioned for histologic examination, whereas cells from blood or cytologic samples, such as fine-needle aspirates, are prepared as direct smears or cytopsin. Frozen tissues have the benefit of being processed more rapidly for LCM than fixed tissue and are considered to be the most reliable source for

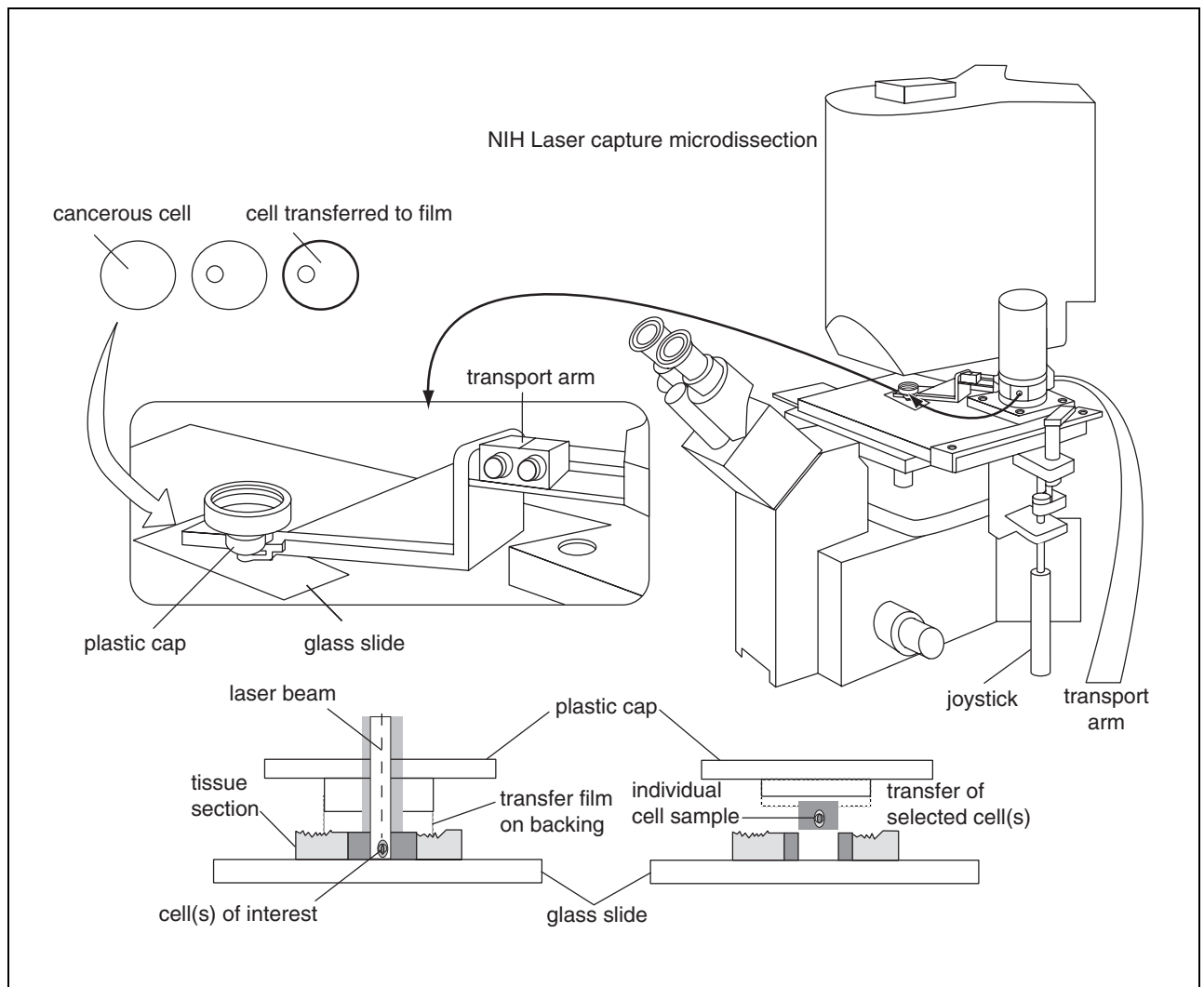


Figure 25A.1.4 LCM Instrument. Schematic of the operation of the PixCell Laser Capture Microdissection Instrument, indicating the location of the transfer arm, transfer film (“cap”), and the glass slide with the specimen to be microdissected. Also shown is a cross-section of the tissue specimen with overlying cap demonstrating the effect of laser firing. Reprinted with permission from Bonner et al. (1997).

molecular (i.e., DNA, RNA, and protein) recovery. Lengths of RNA and DNA of up to 800 base pairs have been recovered from sections prepared from frozen tissue (<http://www.arctur.com>; Dietmaier et al., 1999; Shibutani et al., 2000); however, histologic and cytologic detail are poor compared to fixed paraffin-embedded tissue and subtle diagnostic features may be difficult to discern. The most frequently utilized tissue fixative is neutral buffered formalin (NBF; i.e., 10% buffered formaldehyde) followed by paraffin embedding to allow histologic sectioning. This combination results in cross-linking and “breakage” of proteins, RNA, and DNA, which must be considered when utilizing tissues prepared in this manner.

Regardless of the preparation, cells or tissue are usually stained in order to be visualized for

LCM, although LCM can be performed successfully without staining. Hematoxylin and eosin (H&E) stain is the most commonly used stain for examination of histologic sections, and diagnostic histopathologic criteria are based on its use in veterinary and human pathology practice; therefore, it is frequently used for LCM even though hematoxylin may bind to nucleic acids causing adverse effects during PCR. Other stains such as methyl green and nuclear fast red have been recommended as alternatives, and literally hundreds of others exist in clinical practice and for research applications (Ohyama et al., 2000); however, H&E stained LCM samples have recently been shown to amplify equally as well as samples stained with methyl green, toluidine blue O, or azure B (Ehrig et al., 2001). This is likely due to the relatively small size of LCM samples,

which thus contributes only a small amount of hematoxylin to the PCR reaction mix. It is also recommended that the duration of staining with hematoxylin be minimized to decrease the concentration present. Eosin has been reported to interfere with PCR analysis utilizing the TaqMan instrument and can appear on electrophoretic gels when relatively large numbers of cells are captured for protein analysis (Banks et al., 1999; Ehrig et al., 2001). Consideration should be given to minimizing or eliminating its use when samples will be utilized for either of these assays. Specimens can also be stained immunohistochemically or with fluorescent labels prior to microdissection (Fend et al., 1999b; Murakami et al., 2000).

There are two alternative methods for specimen staining. One is to place the staining solutions into either Coplin jars or staining dishes and immerse the slides in the appropriate solutions. If this method is used, the stains should be changed frequently to prevent contamination by tissue fragments from other tissue samples or microorganisms found in the environment, and to avoid excessive dilution of the staining solutions. The second alternative, and the one that the authors prefer, is to keep the solutions in clean plastic squirt bottles and use a slide staining rack. The slide to be stained can then be placed on the staining rack and the solutions can be applied gently to the slide to cover the tissue or cells, allowed to remain the appropriate time, and then drained from the slide and replaced by the next solution. This reduces any possible contamination, minimizes dilution of solutions, and has the added advantage of using less reagents. For solutions requiring a duration of contact with the slide that is longer than 1 min (i.e., xylene), we utilize small Coplin jars. Thus, the best features of both systems may be used efficiently.

For a successful LCM transfer, the polymer film must be bonded to the targeted tissue so it forms a stronger bond than that between the tissue and the underlying glass slide; therefore, proper sample preparation is critical. It is important that the sample be well dehydrated so that the melted polymer can infiltrate intercellular spaces and create a tight bond. The final dehydration and xylene steps have been found to be absolutely crucial for successful LCM. Any moisture present in the sample during LCM will give less than optimal results. Ideally, samples should be microdissected shortly after dehydration; however, samples can be stored with desiccant after staining and dehydrated for later microdissection, although this is not rec-

ommended for recovery of RNA because of its lability. Additionally, the humidity in the laboratory will also affect the results, and protocols may need to be modified accordingly. Other factors that will affect this bond are presented below (see Troubleshooting).

Specimens, reagents, and materials for processing must be handled in a manner that will allow optimal preservation of the molecule to be analyzed; therefore, samples for RNA and DNA should be handled to minimize contamination from other tissues. Samples for RNA analysis should be processed rapidly, either as fresh-frozen material or briefly fixed in 95% ethanol. RNase-free reagents and materials should be utilized whenever possible. Also, the duration of the actual microdissection session on each stained frozen section should be limited to less than 30 min for optimal RNA preservation. Samples for protein analysis are also best processed as for RNA analysis, but reagents that include protease inhibitors can be used. DNA is more stable, and fixed or frozen tissues can be used, but samples should not be over-fixed in formalin, as DNA yield increases with prolonged fixation times (<6 hr is preferable for small samples).

Troubleshooting

If LCM fails to capture the cells (i.e., they are not released from the slide), the following steps are recommended.

1. Refocus the beam (see Basic Protocol 6).
2. Make sure the sections are flat. Wrinkles can be shaved off using sterile razor blades. Dip the section in xylene after saving the wrinkles to make sure that no contaminating debris remains on the section.
3. Change the cap. Not all caps perform equally well and the age of the caps is important. It is best not to use expired caps and to buy relatively small numbers of caps at a time so that the stock is relatively new.
4. Ensure thorough dehydration of the specimen. Place the slides in fresh xylene for 1 min or more and allow drying in a biologic safety hood for 1 to 5 min. If LCM is still not successful, pass the slides through 95% ethanol twice for 30 sec, absolute ethanol twice for 30 sec, and xylene for 1 to 5 min.
5. Process a new section and make sure that the frozen sections or cytologic specimens have not been allowed to dry on the slide prior to fixation. For formalin-fixed sections, do not bake or at least decrease the baking time.
6. Try a different brand or type of glass slide.

7. If still not successful, call the technical support at Acturus Engineering (650-962-3020). The authors also find that talking with other researchers working with LCM to be very useful.

If LCM is successful, but the cap contains contaminating debris, the following measures are recommended.

1. Make sure the slide is free of debris. It may be necessary to wash the slide in fresh changes of xylene.

2. Use a CapSure Pad or Post-It Note to remove any debris from the cap.

3. Use HS caps, which minimize contamination.

If the LCM was successful, but no RNA, DNA, or protein was identified at analysis, try the following.

1. Make sure optimum laboratory practices and conditions that are free of nucleases or proteinases have been observed.

2. Check the cap to see if the microdissected tissue has dissolved in the lysis buffer in the microcentrifuge tube.

3. Increase the number of microdissected cells.

4. An overnight incubation at 37°C can be used to lyse the cells from the cap when using DNA lysis solutions, if required. For RNA and proteins, inverting and gentle agitation should be used to dislodge the cells from the cap.

Anticipated Results

Many molecular analyses have been successfully performed on cells procured by LCM. These include genomic analyses such as loss of heterozygosity analysis, restriction fragment length polymorphism (RFLP) analysis, DNA methylation analysis, fluorescence in situ hybridization, and comparative genomic hybridization (Finkelstein et al., 1999; Guan et al., 1999; DiFrancesco et al., 2000; Jones et al., 2000; Shen et al., 2000; Slebos et al., 2000). Gene expression analysis has been accomplished from LCM samples utilizing reverse transcription PCR, construction of cDNA libraries, and differential hybridization on high-density-spotted nylon filters or glass microarrays (Peterson et al., 1998; Fend et al., 1999b; Kuecker et al., 1999; Luo et al., 1999; Sgroi et al., 1999; Garrett et al., 2000; Leethanakul et al., 2000; Ohyama et al., 2000). Successful proteomic analysis has been accomplished by coupling LCM with immunoblotting (UNIT 10.8), solid-phase sequential chemiluminescent immunometric assay, one-dimensional and two-dimensional polyacrylamide gel electrophore-

sis (PAGE; UNITS 10.2-10.4), protein chip surface enhanced laser desorption/ionization (SELDI) mass spectrometry, as well as matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (Wright et al., 1999; Natkunam et al., 2000; Ohyama et al., 2000; Ornstein et al., 2000a,b; Palmer-Toy et al., 2000; Simone et al., 2000; also see UNIT 10.21). For all these assays, the expected results will depend on the quality of preservation of the analyte of interest within the sample and upon procurement of at least the minimum number of cells required for analysis.

The number of cells captured depends on tissue thickness and type, the size of the cells, and the size of the laser spot. The number of cells procured can be estimated by counting the number of cells per spot and multiplying by the number of pulses of the laser. The transfer efficiency of the capture should also be considered and can be assessed by viewing the captured tissue on the cap and estimating the percentage of spots that contain tissue.

The number of cells required depends on the assay and whether formalin-fixed, alcohol-fixed, or frozen samples are used. A single PCR reaction (DNA analysis) can be successfully performed with a single cell; however, results are more reliable with at least 10 to 20 cells from a 10- μ m-thick, formalin-fixed, paraffin-embedded section. Such small quantities of cells may not account for the significant heterogeneity that exists even within populations of the same cell type, which should be considered when determining the number of cells to be used. For RNA analysis, fresh-frozen tissues and cytologic specimens briefly fixed in alcohol are preferred. Only a small number of cells (i.e., <50) may be required for transcripts of high copy number per cell when utilizing RT-PCR; however, the authors prefer using ≥ 1000 cells for RT-PCR. cDNA arrays require significantly more RNA, but how much will depend on the type and size of array. It is estimated that a typical mammalian cell contains ~ 20 pg total RNA/cell; therefore, to achieve 5 μ g RNA, the lower limit for some expression arrays, will require the microdissection of 2.5×10^{11} cells, a daunting task. Thus, some authors have advocated amplification of RNA or resultant cDNA prior to hybridization with these larger arrays, even though this may introduce some degree of amplification bias (Luo et al., 1999; Ohyama et al., 2000). For protein analysis, using 50,000 cells for two-dimensional PAGE analysis has been a successful starting point. For western blot analysis, the number of cells

required is at least 2000 to 3000 (<http://www.arctur.com>). Some molecular assays may require modification in order to accommodate the relatively small amount of cells obtained by LCM.

Time Considerations

The time required for LCM is highly variable and depends on the method of tissue processing and staining, the number of cells to be microdissected, and the location and number of the desired cells in each section. H&E staining (see Basic Protocol 5) requires only 10 to 15 min. Microdissecting ~5000 cells, roughly equal to 1000 shots using a 30- μ m spot size, will require 15 to 30 min, provided all the cells required are present within a single tissue section or sample. If multiple sections or samples are required to procure an adequate number of cells, the time required for staining additional sections should be added. This also assumes that the samples are well prepared and microdissected efficiently, and that the cells of interest are easy to identify and locate. Some skill is also required in operating the joystick in combination with laser firing and in being able to identify the tissue and cell type of interest.

The time required for lysis of the cells from the cap depends on the buffer and the method of sample preparation. We have found frozen tissue will be completely removed from the cap by Stat-60 after ~5 min. Formalin-fixed paraffin-embedded tissue in buffers containing proteinase K requires significantly more time and may require an overnight incubation at 37°C.

Literature Cited

- Arnold, M.M., Srivastava, S., Fredenburgh, J., Stockard, C.R., Myers, R.B., and Grizzle, W.E. 1996. Effects of fixation and tissue processing on immunohistochemical demonstration of specific antigens. *Biotechnic. Histochem.* 71:224-230.
- Banks, R.E., Dunn, M.J., Forbes, M.A., Stanley, A., Pappin, D., Naven, T., Gough, M., Harnden, P., and Selby, P.J. 1999. The potential use of laser capture microdissection to selectively obtain distinct populations of cells for proteomic analysis—preliminary findings. *Electrophoresis* 20:689-700.
- Bohm, M., Wieland, I., Schutze, K., and Rubben, H. 1997. Microbeam MOMENT: Non-contact laser microdissection of membrane-mounted native tissue. *Am. J. Pathol.* 151:63-67.
- Bonner, R.F., Emmert-Buck, M., Cole, K., Pohida, T., Chuaqui, R., Goldstein, S., and Liotta, L.A. 1997. Laser capture microdissection: Molecular analysis of tissue. *Science* 278:1481-1483.
- Bostwick, D.G., al Annouf, N., and Choi, C. 1994. Establishment of the formalin-free surgical pathology laboratory. Utility of an alcohol-based fixative. *Arch. Pathol. Lab. Med.* 118:298-302.
- Coombs, N.J., Gough, A.C., and Primrose, J.N. 1999. Optimisation of DNA and RNA extraction from archival formalin-fixed tissue. *Nucl. Acids Res.* 27:e12.
- Dietmaier, W., Hartmann, A., Wallinger, S., Heimmoller, E., Kerner, T., Endl, E., Jauch, K.W., Hofstadter, H., and Ruschoff, J. 1999. Multiple mutation analyses in single tumor cells with improved whole genome amplification. *Am. J. Pathol.* 154:83-95.
- DiFrancesco, L.M., Murthy, S.K., Luider, J., and Demetrick, D.J. 2000. Laser capture microdissection-guided fluorescence in situ hybridization and flow cytometric cell cycle analysis of purified nuclei from paraffin sections. *Modern Pathology* 13:705-711.
- Ehrig, T., Abdulkadir, S.A., Dintzis, S.M., Milbrandt, J., and Watson, M.A. 2001. Quantitative amplification of genomic DNA from histological tissue sections after staining with nuclear dyes and laser capture microdissection. *Journal of Molecular Diagnostics* 3:22-25.
- Fend, F., Quintanilla-Martinez, L., Kumar, S., Beaty, M.W., Blum, L., Sorbara, L., Jaffe, E.S., and Raffeld, M. 1999a. Composite low grade B-cell lymphomas with two immunophenotypically distinct cell populations are true biclonal lymphomas. A molecular analysis using laser capture microdissection. *Am. J. Pathol.* 154:1857-1866.
- Fend, F., Emmert-Buck, M.R., Chuaqui, R., Cole, K., Lee, J., Liotta, L.A., and Raffeld, M. 1999b. Immuno-LCM: Laser capture microdissection of immunostained frozen sections for mRNA analysis. *Am. J. Pathol.* 154:61-66.
- Finkelstein, S.D., Hasegawa, T., Colby, T., and Yousem, S.A. 1999. 11q13 allelic imbalance discriminates pulmonary carcinoids from tumors. A microdissection-based genotyping approach useful in clinical practice. *Am. J. Pathol.* 155:633-640.
- Garrett, S.H., Sens, M.A., Shukla, D., Flores, L., Somji, S., Todd, J.H., and Sens, D.A. 2000. Metallothionein isoform 1 and 2 gene expression in the human prostate: Downregulation of MT-1X in advanced prostate cancer. *Prostate* 43:125-135.
- Glasow, A., Haidan, A., Hilbers, U., Breidert, M., Gillespie, J., Scherbaum, W.A., Chrousos, G.P., and Bornstein, S.R. 1998. Expression of Ob receptor in normal human adrenals: Differential regulation of adrenocortical and adrenomedullary function by leptin. *J. Clin. Endocrinol. Metab.* 83:4459-4466.
- Goldsworthy, S.M., Stockton, P.S., Trempus, C.S., Foley, J.F., and Maronpot, R.R. 1999. Effects of fixation on RNA extraction and amplification from laser capture microdissected tissue. *Molecular Carcinogenesis* 25:86-91.

- Guan, R.J., Fu, Y., Holt, P.R., and Pardee, A.B. 1999. Association of K-ras mutations with p16 methylation in human colon cancer. *Gastroenterology* 116:1063-1071.
- Ikeda, K., Monden, T., Kanoh, T., Tsujie, M., Izawa, H., Haba, A., Ohnishi, T., Sekimoto, M., Tomita, N., Shiozaki, H., and Monden, M. 1998. Extraction and analysis of diagnostically useful proteins from formalin-fixed, paraffin-embedded tissue sections. *J. Histochem. Cytochem.* 46:397-403.
- Jin, L., Thompson, C.A., Qian, X., Kuecker, S.J., Kulig, E., and Lloyd, R.V. 1999. Analysis of anterior pituitary hormone mRNA expression in immunophenotypically characterized single cells after laser capture microdissection. *Lab. Invest.* 79:511-512.
- Jones, C., Foschini, M.P., Chaggar, R., Lu, Y.J., Wells, D., Shipley, J.M., Eusebi, V., and Lakhani, S.R. 2000. Comparative genomic hybridization analysis of myoepithelial carcinoma of the breast. *Lab. Invest.* 80:831-836.
- Kuecker, S.J., Jin, L., Kulig, E., Oudraogo, G.L., Roche, P.C., and Lloyd, R.V. 1999. Analysis of PRL, PRL-R, TGF β -R11 gene expression in normal and neoplastic breast tissues after laser capture microdissection. *Appl. Immunohist. Molec. Morph.* 7:193-200.
- Leethanakul, C., Patel, V., Gillespie, J., Pallente, M., Ensley, J.F., Kooontongkaew, S., Liotta, L.A., Emmert-Buck, M., and Gutkind, J.S. 2000. Distinct pattern of expression of differentiation and growth-related genes in squamous cell carcinomas of the head and neck revealed by the use of laser capture microdissection and cDNA arrays. *Oncogene* 19:3220-3224.
- Luo, L., Salunga, R.C., Guo, H., Bittner, A., Joy, K.C., Galindo, J.E., Xiao, H., Rogers, K.E., Wan, J.S., Jackson, M.R., and Erlander, M.G. 1999. Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nature Medicine* 5:117-122 [published erratum appears in *Nature Medicine* 5:355].
- Masuda, N., Ohnishi, T., Kawamoto, S., Monden, M., and Okubo, K. 1999. Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. *Nucl. Acids Res.* 27:4436-4443.
- Murakami, H., Liotta, L., and Star, R.A. 2000. IF-LCM: Laser capture microdissection of immunofluorescently defined cells for mRNA analysis rapid communication. *Kidney Int.* 58:1346-1353.
- Natkunam, Y., Rouse, R.V., Zhu, S., Fisher, C., and van De Rijn, M. 2000. Immunoblot analysis of CD34 expression in histologically diverse neoplasms. *Am. J. Pathol.* 156:21-27.
- Ohyama, H., Zhang, X., Kohno, Y., Alevizos, I., Posner, M., Wong, D.T., and Todd, R. 2000. Laser capture microdissection-generated target sample for high-density oligonucleotide array hybridization. *BioTechniques* 29:530-536.
- Ornstein, D.K., Englert, C., Gillespie, J.W., Paweletz, C.O., Linehan, W.M., Emmert-Buck, M.R., and Petricoin, E.F. III 2000a. Characterization of intracellular prostate-specific antigen from laser capture microdissected benign and malignant prostatic epithelium. *Clin. Cancer Res.* 6:353-356.
- Ornstein, D.K., Gillespie, J.W., Paweletz, C.P., Duray, P.H., Herring, J., Vocke, C.D., Topalian, S.L., Bostwick, D.G., Linehan, W.M., Petricoin, E.F. III, and Emmert-Buck, M.R. 2000b. Proteomic analysis of laser capture microdissected human prostate cancer and in vitro prostate cell lines. *Electrophoresis* 21:2235-2242.
- Palmer-Toy, D.E., Sarracino, D.A., Sgroi, D., LeVangie, R., and Leopold, P.E. 2000. Direct acquisition of matrix-assisted laser desorption/ionization time-of-flight mass spectra from laser capture microdissected tissues. *Clin. Chem.* 46:1513-1516.
- Paweletz, C.P., Ornstein, D.K., Roth, M.J., Bichsel, V.E., Gillespie, J.W., Calvert, V.S., Vocke, C.D., Hewitt, S.M., Duray, P.H., Herring, J., Wang, Q.H., Hu, N., Linehan, W.M., Taylor, P.R., Liotta, L.A., Emmert-Buck, M.R., and Petricoin, E.F. III. 2000. Loss of annexin I correlates with early onset of tumorigenesis in esophageal and prostate carcinoma. *Cancer Res.* 60:6293-6297.
- Peterson, L.A., Brown, M.R., Carlisle, A.J., Kohn, E.C., Liotta, L.A., Emmert-Buck, M.R., and Krizman, D.B. 1998. An improved method for construction of directionally cloned cDNA libraries from microdissected cells. *Cancer Res.* 58:5326-5328.
- Sawyer, E.J., Hanby, A.M., Ellis, P., Lakhani, S.R., Ellis, I.O., Boyle, S., and Tomlinson, I.P. 2000. Molecular analysis of phyllodes tumors reveals distinct changes in the epithelial and stromal components. *Am. J. Pathol.* 156:1093-1098.
- Sgroi, D.C., Teng, S., Robinson, G., LeVangie, R., Hudson, Jr. J.R., and Elkahoulou, A.G. 1999. In vivo gene expression profile analysis of human breast cancer progression. *Cancer Res.* 59:5656-5661.
- Sheehan, D.C. and Hrapchak, B.B. (eds.) 1987a. Specimen preparation for enzyme histochemistry *In The Theory and Practice of Histotechnology.* 2nd Edition, pp. 293-295. The C.V. Mosby Company, St. Louis, MO.
- Sheehan, D.C. and Hrapchak, B.B. (eds.) 1987b. Processing of tissue dehydrants, clearing agents, and embedding media *In The Theory and Practice of Histotechnology.* 2nd Edition, pp. 59-78. The C.V. Mosby Company, St. Louis, MO.
- Shen, C.Y., Yu, J.C., Lo, Y.L., Kuo, C.H., Yue, C.T., Jou, Y.S., Huang, C.S., Lung, J.C., and Wu, C.W. 2000. Genome-wide search for loss of heterozygosity using laser capture microdissected tissue of breast carcinoma: An implication for mutator phenotype and breast cancer pathogenesis. *Cancer Res.* 60:3884-3892.
- Shibutani, M., Uneyama, C., Miyazaki, K., Toyoda, K., and Hirose, M. 2000. Methacarn fixation: A novel tool for analysis of gene expressions in

paraffin-embedded tissue specimens. *Lab. Invest.* 80:199-208.

- Simone, N.L., Remaley, A.T., Charboneau, L., Petricoin, E.F. III, Glickman, J.W., Emmert-Buck, M.R., Fleisher, T.A., and Liotta, L.A. 2000. Sensitive immunoassay of tissue cell proteins procured by laser capture microdissection. *Am. J. Pathol.* 156:445-452.
- Sirivatanauksorn, Y., Drury, R., Crnogorac-Jurcevic, T., Sirivatanauksorn, V., and Lemoine, N.R. 1999. Laser-assisted microdissection: Applications in molecular pathology. *Journal of Pathology* 189:150-154.
- Slebos, R.J.C., Hoppin, J.A., Tolbert, P.E., Holly, E.A., Brock, J.W., Zhang, R.H., Bracci, P.M., Foley, J., Stockton, P., McGregor, L.M., Flake, G.P., and Taylor, J.A. 2000. K-ras and p53 in pancreatic cancer: Association with medical history, histopathology, and environmental exposures in a population-based study. *Cancer Epidemiol., Biomark. Prev.* 9:1223-1232.
- Suarez-Quian, C.A., Goldstein, S.R., Pohida, T., Smith, P.D., Peterson, J.I., Wellner, E., Ghany, M., and Bonner, R.F. 1999. Laser capture microdissection of single cells from complex tissues. *Biotechniques* 26:328-335.
- Vardaxis, N.J., Hoogeveen, M.M., Boon, M.E., and Hair, C.G. 1997. Sporicidal activity of chemical and physical tissue fixation methods. *J. Clin. Pathol.* 50:429-433.
- Wright G.L. Jr., Cazares, L.H., Leung, S.-M., Nasim, S., Adam, B.-L., Yip, T.-T., Schellhammer, P.F., Gong, L., and Vlahou, A. 1999. Proteinchip surface enhanced laser desorption/ionization (SELDI) mass spectrometry: A novel protein biochip technology for detection of prostate cancer biomarkers in complex protein mixtures. *Prostate Cancer and Prostatic Diseases* 2:264-276.

Key References

Sheehan and Hrapchak, 1987. See above.

This is a standard reference used by many histotechnologists for the basics of tissue processing, embedding, and sectioning.

Suarez-Quian et al., 1999. See above.

This reference provides a good overview of the mechanics and principles of LCM.

Internet Resources

<http://dir.niehs.nih.gov/dirlep/lcm/guidelines.html>

This website is maintained by the Laboratory of Experimental Pathology of the National Institute of Environmental Health Sciences and is another valuable source of protocols and general information for LCM.

<http://www.arctur.com>

This is the website of Arcturus Engineering. It is a very useful source of all LCM-related information including protocols, references, and resources. Many of the protocols that we use, including those presented here, are modifications of protocols found at this website.

<http://www.bioprotocol.com>

This website contains protocols for the performance of LCM, the preparation of tissues for LCM and for processing of microdissected tissue for DNA, RNA and protein analysis.

Contributed by Andra R. Frost, Isam-Eldin Eltoum, and Gene P. Siegal
University of Alabama at Birmingham
Birmingham, Alabama

Preparation of Single Cells from Solid Tissues for Analysis by PCR

The ability to amplify a few copies of DNA or RNA to analyzable quantities makes it technically possible to obtain detailed information regarding the DNA content and/or transcriptional pattern of a single cell (Mullis and Falona, 1987). Although in many cases, analysis at the level of the whole tissue can provide the required information, there are circumstances that necessitate acquiring data on individual cells of a particular type. A preparation of total DNA and RNA isolated from a tissue gives quantitative data but only an average profile, masking differences among individual cells. In situ analysis provides qualitative information on localization of abundant nucleic acids in specific cells, but is generally not quantitative. Thus, it can be desirable to apply quantitative assays to individual cells.

The acquisition of individual cells from blood and loosely associated tissues such as spleen is straightforward, since these organs are essentially cell suspensions. Solid tissues, however, are almost universally composed of tightly linked cells of multiple types, organized in a highly structured and functionally interactive manner (Gilbert, 1994). It is reasonable to expect that disruption of this environmental context rapidly alters the physiology of the once-partnered cells. Even in the case of easily dissociated tissues, the impact of manipulating the living tissue on the process under study must be considered. In addition, some adult cell types, most notably neurons, can be recovered only at low efficiency, with the majority bursting during the process (Pretlow and Pretlow, 1982).

This unit details a protocol for the separation of solid tissues into single-cell suspensions for subsequent analysis of nucleic acids and protein. This protocol was developed using mice, with the major focus being the analysis of the interaction of herpes simplex virus (HSV) with the neurons of the trigeminal ganglia (Sawtell, 1997). The balance between fixation and dissociation should be determined for the particular tissue of interest. It has been determined, however, that the dissociation protocol is directly useful for several other mouse tissues including liver, heart, skeletal muscle, lung, pancreas, brain, intestine, and reproductive organs. Kidney yields a combination of single cells and multicellular tubular structures. The adaptation of the method to other laboratory animals has not been fully explored. Again, the appropriate balance between fixation and dissociation would need to be determined for other species of interest. Using the approach of adjusting the volume of the fixative perfused through the animal to achieve this balance, the author's laboratory has determined that the method is directly useful for guinea pigs.

Tissues are fixed in situ by perfusion (see Basic Protocol 1), terminating cell processes and thus changes that would accompany dissociating the living tissue; their numbers can then be quantitated (see Support Protocol). Once separated, individual cells or groups of a particular cell type can then be analyzed using PCR strategies (see Basic Protocols 2 and 3; Fig. 25A.2.1). An alternative to fixing by perfusion (see Alternate Protocol 1) and a modification of the standard Percoll gradient separation to prepare *lacZ* expressing cells (see Alternate Protocol 2) are also provided. The method has broad potential and is particularly potent when the cell type of interest represents a minor population relative to other cells types in the tissue. The procedure can also be adapted to allow quantification of the number of cells within a tissue containing specific nucleic acid sequences, for example, a particular viral DNA or RNA sequence.

Contributed by N.M. Sawtell

Current Protocols in Molecular Biology (2002) 25A.2.1-25A.2.15

Copyright © 2002 by John Wiley & Sons, Inc.

Discovery of
Differentially
Expressed Genes

25A.2.1

Supplement 58

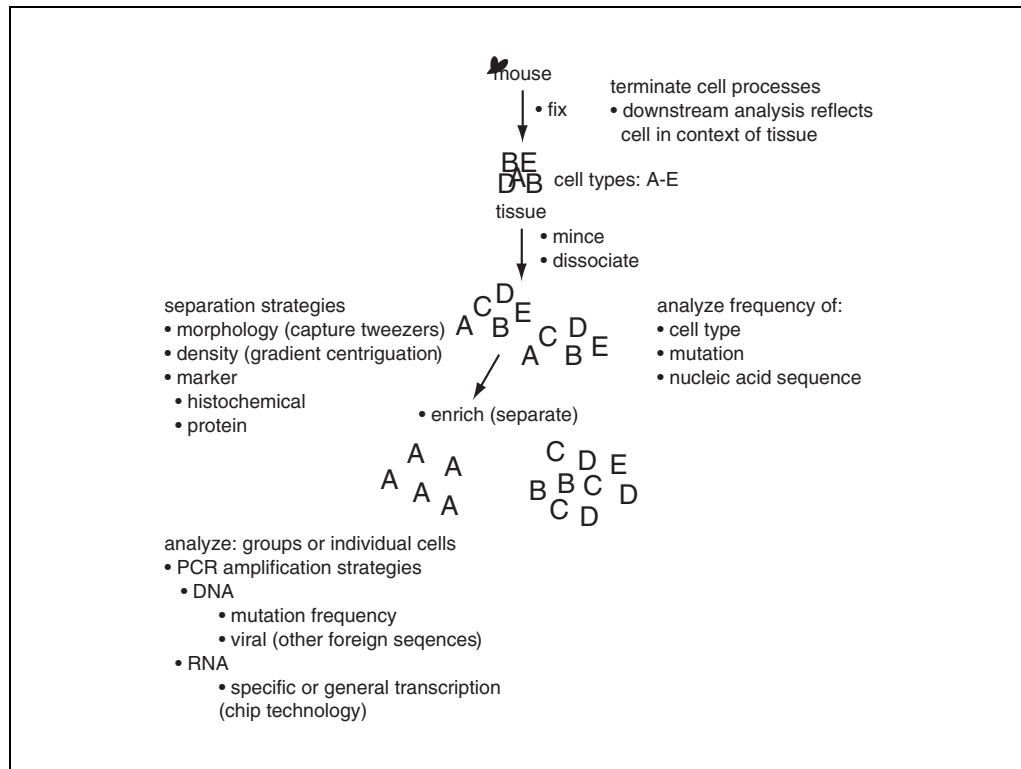


Figure 25A.2.1 Schematic representation of the preparation of single cells from solid tissue. Tissue is represented by the box in the center, and letters A to E represent different cell types within that tissue.

BASIC PROTOCOL 1

PERFUSION FIXATION AND ENRICHMENT OF SINGLE CELLS

In this protocol, the animal (here, a mouse) is perfused with Streck tissue fixative (STF), a noncrosslinking fixative. This fixative and the fixation conditions presented were determined empirically so that intracellular nucleic acids and proteins are preserved without interfering significantly with the ability of the dissociating enzymes to free the cells from the extracellular matrix. The fixed tissues of interest are dissected out, finely minced, and enzymatically separated using collagenase. The cell types of interest are then enriched using a suitable strategy.

At this point, any of a number of methods can be used to harvest the desired cell populations from the cell suspension. Percoll gradient separation is given here; however, the end application will strongly influence the procedure selected.

Materials

- Streck tissue fixative (STF; Streck Laboratories)
- Animal (e.g., mouse)
- Sodium pentobarbital
- 95% ethanol
- 0.25% (w/v) collagenase CLS I (Worthington) in Hank's balanced salt solution (HBSS; see recipe)
- Triple 0.2- μ m filtered nanopure (3 \times F) H₂O
- Percoll (Pharmacia): adjust to pH 6.0 with HCl
- Peristaltic pump (BRL CP-600 or equivalent) and appropriate tubing
- 15- and 50-ml conical tubes
- 27-G needle
- 80°C water bath

Dissecting microscope (optional)
Clean dissection tools (e.g., forceps, scalpel blades, hemostat, 25-G needles)
Glass slides: bake overnight (3 hr minimum) at 250°C
200- and 1000- μ l aerosol-resistant pipette tips
15-ml polystyrene conical tubes
9-in. Pasteur pipettes: bake overnight (3 hr minimum) at 250°C

Additional reagents and equipment for determining number of neurons recovered (see Support Protocol), and analyzing DNA or RNA from single-cell populations (see Basic Protocols 2 and 3)

NOTE: All protocols using live animals must first be reviewed and approved by an Institutional Animal Care and Use Committee (IACUC) or must conform to governmental regulations regarding the care and use of laboratory animals.

NOTE: Depending upon the final application of the cells, all materials must be DNase- and RNase-free, and free of contaminating nucleic acids which could interfere with the interpretation of downstream PCR.

Perform perfusion fixation (as performed in mice)

1. Set up perfusion equipment by placing tubing from a peristaltic pump in the bottom of a ~50 ml conical tube containing 30- to 40-ml Streck tissue fixative (STF). Attach a 27-G needle to the other end (this will be inserted into the left ventricle). Run fixative through the line.
2. Place a 50-ml conical tube containing 50 ml STF in an 80°C water bath and equilibrate to temperature.

Heat facilitates the inactivation of nucleases.

3. While the fixative is heating, anesthetize the animal by intraperitoneal injection of 80 to 100 mg/kg sodium pentobarbital. As soon as deep reflexes are fully deadened—i.e., in mice, lack of corneal reflexes (i.e., no blinking response when touched with the tip of a gloved finger) and response to pinching rear paw very firmly—place the animal ventral surface up on absorbent paper and wet the chest and abdomen with 95% ethanol.

Isoflurane can be used as an alternative anesthetic.

4. Use forceps to lift the skin and, using a scalpel, make a T-shaped incision starting over the abdomen with the vertical- and horizontal-cut centers at the base of the sternum just below the diaphragm. Cut the diaphragm along the rib line and keep the chest cavity open by clamping the base of sternum with a small hemostat, rotating it upward toward the chest.

For additional information on animal handling, see Coligan et al. (2001), Chapter 1.

5. Insert the needle at the end of the pump tubing into the left ventricle and start the pump, adjusting the flow rate to ~6 ml/min. When the right atrium becomes dilated, pierce with sharp pointed forceps to provide outflow. First pump 15 to 20 ml room-temperature STF through the animal to remove blood from the vasculature, followed by 40 to 50 ml heated (80°C) fixative. Stop the pump when fixative has been depleted.

This procedure is not difficult but requires practice. The best indicator of a successful perfusion is paling of the liver. If the liver does not begin to pale rapidly, try repositioning the needle in the ventricle, adjusting its depth and angle.

Perfusion fixation is effective because the fixative is distributed to tissues and cells via the macro- and microvasculature. Coagulation of the blood in the vessels could occur upon contact with the heated fixative, thus it is important to first remove blood with room-temperature STF.

Dissociate tissue

6. Using a dissecting microscope (if possible), dissect the tissues of interest with “clean” dissection tools and finely mince on a nuclease-free (i.e., baked) glass slide using scalpel blades or needles (e.g., 25-G needles for ganglia).

The author uses disposable instruments (e.g., unused 25-G needles, unused scalpel blades) that are discarded after use (i.e., a single dissection), which prevents the possibility of any carryover; however, it should be adequate to clean instruments in detergent (e.g., liquinox), rinse, and soak in 3% hydrogen peroxide for 2 hr, then rinse in 3×F H₂O and bake overnight (3 hr minimum) at 250°C. Any procedure for cleaning potentially contaminated instruments should be confirmed to be effective.

Visualization of the mincing procedure under a dissecting microscope is helpful. Separate instruments must be used for each tissue unit if cross contamination will present a problem in the interpretation of downstream analyses.

7. Place minced tissue into 0.25% (w/v) collagenase CLS I in HBSS and incubate in a 1.5- or 2-ml microcentrifuge tube 5 to 10 min at 37°C.

The volume of collagenase used will depend upon the amount of tissue. Six fixed mouse trigeminal ganglia (TG) are routinely digested in 1.5 ml collagenase.

The investigator must screen batches of collagenase and select a batch that is free of DNase activity. If RNA will be analyzed, a batch free of RNase must be selected (see Critical Parameters and Troubleshooting, Collagenase)

8. After collagenase treatment, facilitate dissociation by gentle trituration, first using 1000- μ l, then 200- μ l (as the tissue dissociates into smaller pieces) aerosol-resistant pipet tips.

In the author's studies, dissociation of TG is generally complete within 30 min.

Depending on the application, the requirement for complete dissociation may be less critical. It is helpful to monitor progress of dissociation by viewing a drop of the suspension under the microscope.

9. Pellet dissociated tissue by microcentrifuging 5 min at 5000 rpm, room temperature. Resuspend gently in STF at room temperature. Heat resuspended cell suspension to 70°C for 10 min. Place on ice briefly, repellet, and resuspend in triple 0.2- μ m filtered nanopure (3×F) water.

At this point the integrity of the DNA, RNA, and/or protein (depending upon what will be analyzed), should be tested. DNA can be isolated using standard proteinase K/SDS digestion followed by phenol/chloroform extractions and ethanol precipitation (UNIT 2.1A). RNA can be isolated from the cells using commercially available reagents such as Ultraspec (Biotecx). When isolating RNA, cells should be homogenized using a tissue grinder to ensure complete disruption of the cell membrane. Protein should be prepared from cells by boiling in standard Laemmli cocktail (e.g., 0.125 M Tris·Cl/4% SDS/20% glycerol/10% 2-mercaptoethanol). Integrity of nucleic acid or protein is then determined by appropriate gel electrophoresis (Chapter 10). If information about the integrity of a specific nucleic acid or protein is desired, Southern (UNIT 2.9A), northern (UNIT 4.9), and/or immunoblotting (UNIT 10.8) can then be performed, probing the membrane with the relevant labeled nucleic acid probe or antibody. One should not necessarily expect that the integrity of these cells will be as great as that from tissue culture cells or fresh tissue, but it can be more than adequate to permit qualitative analysis by RT-PCR.

Figure 25A.2.2A.1 to D.1 shows several tissues, including cerebral cortex, trigeminal ganglia, liver, and diaphragm, after fixation and dissociation.

10. Determine the number of neurons recovered (see Support Protocol).

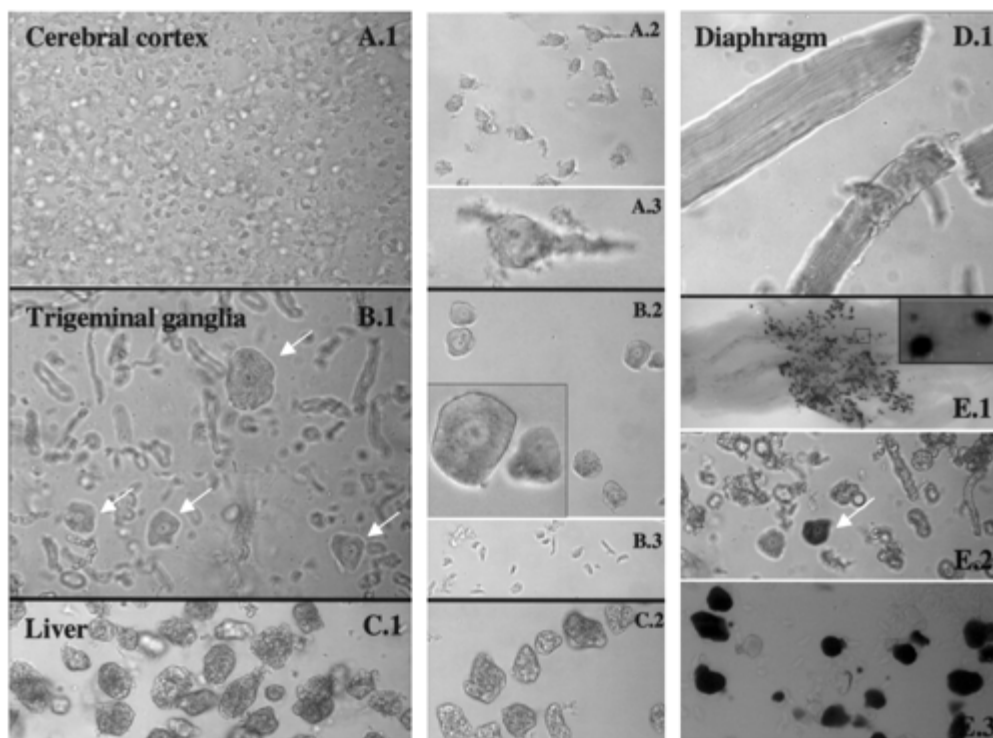


Figure 25A.2.2 Photomicrograph showing tissues after dissociation and density gradient centrifugation. Following perfusion fixation, several tissue types were removed, finely minced, and dissociated as described in this protocol. The total dissociated cell suspensions obtained from cerebral cortex (**A.1**), trigeminal ganglia (**B.1**), liver (**C.1**), and diaphragm (**D.1**) are shown. Following density gradient centrifugation, enriched populations of neurons were obtained from cerebral cortex (**A.2**, **A.3**) and trigeminal ganglia (**B.2**). Enriched populations of satellite and support cells isolated from trigeminal ganglia are shown in **B.3**. An example of “marker” based separation is shown in panels **E.1** to **E.3**. A mouse infected with a virus containing a β -galactosidase expression cassette was perfusion fixed and the trigeminal ganglia removed and stained histochemically for β -galactosidase activity using Xgal. The dark spots in panel E.1 are blue neurons, a result of the action of β -galactosidase on Xgal. The presence of this blue reaction product indicates that these neurons contain virus actively transcribing the β -galactosidase gene. The ganglia were then dissociated into single cell suspensions (E.2) and the blue neurons enriched by density gradient centrifugation (E.3). These neurons can be analyzed individually or in groups using PCR strategies.

11. Harvest desired cell populations by Percoll gradient centrifugation (steps 12 to 16) or by another suitable method.

The end application will strongly influence the procedure selected. The following steps enrich for neurons, but the protocol can also be used to enrich for other cell types.

Enrich for neurons by Percoll gradient

12. Prepare a discontinuous Percoll gradient as follows:
 - a. Mix Percoll and 3 \times F H₂O to make 40%, 50%, and 60% (v/v) Percoll solutions. Keep on ice.
 - b. Place the dissociated cell mixture on the bottom of a 15-ml polystyrene (for greater visibility) conical tube.
 - c. Using a baked 9-in. Pasteur pipette, layer 2.5 ml of 40% solution beneath the cell suspension, then carefully dispense the 50% solution under the 40% layer. Finally, carefully dispense the 60% solution beneath the 50% layer. Be sure to dispense all solutions from the tip of the pipette in a slow continuous stream.

13. Centrifuge the gradient in a benchtop centrifuge 10 min at 1800 rpm ($\sim 900 \times g$), 4°C.

The Percoll gradient resulting in the optimum separation of neurons from support cells was determined empirically.

14. Remove tube from the centrifuge and place in a stable rack, taking care not to disturb the gradient. Visually inspect the gradient. Carefully draw off the top myelin-containing layer to reduce contamination of cells banding lower on the gradient. A band of cells should be apparent at the 50%:60% interface. This band will contain highly enriched neurons.
15. Insert a baked 9-in. glass Pasteur pipette into the band of neurons and draw the banded cells into the pipette. Place the Percoll/cell mixture into a 15-ml polystyrene conical tube. Rinse by filling the tube with $3\times F$ H₂O and pelleting the cells by centrifuging in a benchtop centrifuge 10 min at 1800 rpm ($\sim 900 \times g$), 4°C.

A second gradient is not useful unless the first gradient has been overloaded.

16. Decant supernatant and resuspend the pellet in ~ 12 ml $3\times F$ H₂O. Repeat two additional times.
17. After the final rinse, decant the supernatant and resuspend the pellet in a small volume (e.g., 300 to 500 μ l) $3\times F$ H₂O. Transfer resuspended cells to a 1.5-ml centrifuge tube and examine one drop using a microscope. Determine number of neurons (see Support Protocol).

For examples of results, see Figure 25A.2.2, panels A.2 to C.2, A.3, and B.3.

Many factors will influence the separation of the cell suspension on the Percoll gradient. Thus, adjusting the gradient to give the separation desired may be required. Monitoring the distribution of cells throughout the gradient is helpful when beginning to determine optimum separation conditions.

18. Analyze DNA or RNA from single cell populations (see Basic Protocols 2 and 3).

SUPPORT PROTOCOL

DETERMINING NUMBER OF NEURONS RECOVERED

In the preceding method (see Basic Protocol 1), there are two steps (i.e., steps 10 and 17) at which evaluating the yields of the cell type of interest should be performed. The following procedure is presented for the evaluation of neurons but can be easily adapted for any cell type that can be distinguished on the basis of morphology or specific marker protein.

Materials

Cell pellet (see Basic Protocol 1)
Cresyl violet solution (see recipe)
95% and 100% ethanol
Xylene
Permount
Superfrost/Plus glass slides (Fisher) or equivalent with coverslips
Additional reagents and equipment for analyzing neuron-specific proteins (e.g., neurofilament 200 kDa peptide) by immunohistochemistry (Sawtell, 1997)

1. Resuspend cell pellet in a known volume of $3\times F$ H₂O. Mix tube well by flicking and inverting several times to ensure uniform distribution of cells. Dot five 1- μ l aliquots of the cell suspension onto a Superfrost/Plus glass slide or equivalent. Keep cell suspension thoroughly mixed during aliquoting. Dry slide thoroughly.

If more than one type of assessment is to be performed, multiple slides should be prepared.

2. Stain the slide with cresyl violet solution by overlaying the staining solution onto the slide for 5 min at room temperature, rinsing in deionized water, and dehydrating by dipping once in 95% ethanol and then twice in 100% ethanol. Clear the dehydrated slide in xylene and mount coverslip using Permount.
3. Identify neurons on the basis of morphology using a microscope. Count the number of neurons in each 1- μ l aliquot. Determine the average number of neurons per microliter and calculate the total number of neurons by multiplying the average number per microliter by the total number of microliters cell suspension.
4. Analyze an additional slide by immunohistochemistry for a neuron-specific protein, such as neurofilament 200 kDa peptide (detailed in Sawtell, 1997).

The number of neurons determined by morphology should be similar to that determined on the basis of neurofilament 200 kDa peptide staining.

NONPERFUSION FIXATION WITH STF SOLUTION

In some cases, perfusion fixation is not possible. The following procedure is an alternative to perfusion fixation for subsequent analysis of DNA.

Additional Materials (also see *Basic Protocol 1*)

Harvested tissue, fresh
HBSS (see recipe)

1. Finely mince freshly harvested tissue in a drop of STF on a glass slide.

The tube sizes and volumes given are appropriate for 30 to 40 mg of tissue. If larger amounts of tissue are used, tube sizes and volumes should be scaled up accordingly.

2. Transfer minced tissue to a 1.5 to 2-ml tube containing 1 ml STF and incubate for the desired time at room temperature.

The optimum fixation time must be determined empirically. In a preliminary experiment, divide minced tissue into several tubes and fix 5 to 15 min.

Fixation is carried out at room temperature so that subsequent dissociation is possible; therefore, this method is not recommended for separation of cells to be used for downstream analysis of RNA.

3. Following fixation, rinse minced tissue by microcentrifuging tissue 5 min at 5000 rpm, room temperature, then drawing off the supernatant and resuspending the pellet in HBSS. Repeat this process four times.
4. Treat the fixed minced tissue (Basic Protocol 1, steps 7 to 9). Examine the dissociation properties of the cells and the integrity of the nucleic acids and proteins. Select the fixation time yielding good separation and integrity.
5. Proceed as described for perfusion fixation (see Basic Protocol 1, steps 10 to 18).

PREPARATION OF *lacZ*-EXPRESSING CELLS FROM SOLID TISSUES

In this example, a procedure used in the author's laboratory, mice expressing an *E. coli* β -galactosidase expression cassette are perfusion fixed using a modification of the procedure described above (see Basic Protocol 1), to visualize *lacZ*-expressing cells.

Materials (also see *Basic Protocol 1*)

Glutaraldehyde
100 μ g/ml Xgal in Xgal buffer (see recipe)

**ALTERNATE
PROTOCOL 1**

**ALTERNATE
PROTOCOL 2**

**Discovery of
Differentially
Expressed Genes**

25A.2.7

1. Perfusion-fix the animal (see Basic Protocol 1, steps 1 to 4), except add 0.2% (w/v) glutaraldehyde to the STF and pump 20 ml of this solution through the animal at room temperature. Proceed with 80°C STF-only perfusion as described (see Basic Protocol 1, step 5).

This preserves β -galactosidase activity which does not remain active in STF alone. The author has utilized mice infected with a viral mutant containing a β -galactosidase expression cassette; however, mice containing a β -galactosidase transgene or mice in which a β -galactosidase cassette has been introduced using any gene transfer approach could also be analyzed in this way.

2. Remove tissue of interest and incubate in 100 μ g/ml Xgal in Xgal buffer at 37°C for 3 hr.

The time of incubation in the Xgal will depend on the strength of the promoter driving expression. The minimum amount of time for development should be used.

3. Inspect tissue and confirm presence of “marked” cells, then mince and dissociate the tissue (see Basic Protocol 1, steps 6 to 9).
4. Enrich cell population by Percoll gradient separation (see Basic Protocol 1, steps 10 to 17) or other suitable method.

Blue neurons are enriched in the bottom of the gradient, presumably because of increased density from the precipitated X-gal reaction product. This is shown in Fig. 25A.2.1E.1 to 3.

BASIC PROTOCOL 2

ANALYSIS OF SINGLE CELLS BY PCR

In the following section, a protocol for analyzing the dissociated enriched neurons by PCR to detect the HSV thymidine kinase gene is presented; however, this protocol can be applied to other cell types and nucleic acids as well. The goal in developing this assay was to provide a method for the quantitative assessment of the number of neurons containing the HSV genome. Because the frequency of the latent viral genome in the author’s experimental system was relatively high (20% to 30% of the total neurons in the ganglion), the analysis had to be performed on single neurons; however, depending on the frequency of the nucleic acid of interest in the cell pool being analyzed, it could be possible to perform the analysis on samples containing groups of known numbers of cells.

The primers and basic PCR conditions are essentially as reported by Katz and Coen (1990) and detailed in *UNIT 15.7*. Steps are included here for (1) aliquoting cells, (2) confirming the number of cells per tube being analyzed, and (3) eliminating any extracellular contaminating DNA. This step is critical to ensure that the DNA being amplified is actually intracellular. This is done by using DNase linked to beads. The bead cannot enter the cell, and thus the DNase is able to digest DNA in the fluid surrounding the cell, but does not destroy the DNA within the cell. In the next steps, which include a proteinase K treatment (to increase the permeability of the cell) and the PCR reaction itself, a two-part buffer system is utilized to minimize pipetting and insure maximum uniformity in the setup of samples by eliminating the need to pipet very small volumes.

Materials

- Enriched cell sample (see Basic Protocol 1 or Alternate Protocols 1 or 2)
- Triple 0.2- μ m filtered nanopure (3 \times F) H₂O
- Ponceau S solution (see recipe)
- Immobilized-DNase on PVP beads (Mobictec)
- DNase reaction buffer (see recipe)
- PCR/PK solution (see recipe)

DNA standards—e.g., cloned segments of HSV genome containing the gene being amplified (e.g., thymidine kinase)
PCR amplification solution (see recipe)
Taq DNA polymerase (Life Technologies)
200- μ l PCR tubes
Dissecting microscope
PCR Gene Amp 2400 (Perkin Elmer Cetus)
Gene screen plus nylon membrane (NEN Life Science Products)
Storage phosphor screen (Molecular Dynamics)
Imagequant software

Additional reagents and equipment for quantitating standards (UNIT 15.7; Sawtell and Thomson, 1992), PCR (UNIT 15.1), nondenaturing polyacrylamide gel electrophoresis (UNITS 2.5 & 2.7), UV-crosslinking DNA to filters (UNIT 2.9), hybridizing blots with oligonucleotides (UNITS 2.9A & 6.4), labeling oligonucleotides (UNITS 4.6, 4.8 & 15.7), and phosphorimaging (APPENDIX 3A)

Select single neurons

1. Dilute a portion of the enriched neuron sample with 3 \times F H₂O so that 1 μ l contains ~1 neuron. Add Ponceau S solution to a final volume of 1/200 and aliquot 1 μ l neuronal suspension into the bottom of a 200- μ l PCR tubes.

This dye allows easy visualization of neurons in the bottom of the PCR tube when viewed under a dissecting microscope, but does not interfere with subsequent analyses.

2. Examine each tube under a dissecting microscope and identify those containing a single neuron for use in step 3.

The number of tubes will depend upon the anticipated frequency of the DNA sequence being analyzed. In the authors studies, a typical analysis will include 200 single neuron samples.

Immobilized DNase treatment and PCR reaction

3. Resuspend immobilized-DNase on PVP beads in DNase reaction buffer so that 5 μ l contains ~100 beads. Add a 5- μ l aliquot to each PCR tube containing a single neuron. Mix gently. Incubate samples several hours or overnight at 37°C.

The purpose of the DNase treatment is to make sure that the DNA being measured is the DNA within the cell or cells in the PCR tube. While DNase treatment could be performed on cells en masse, one could not be sure that some cells were not broken during purification and aliquoting of cells.

Prior to use of the immobilized DNase, it is important to confirm that the DNase activity in the preparation is associated with the bead, and that no free DNase activity can be detected, as any free DNase could enter the cell and destroy the intracellular DNA. To test this, the immobilized DNase in activation buffer is pelleted gently (as detailed by the manufacturer) and an aliquot of the supernatant is drawn off and placed in a 1.5-ml microcentrifuge tube. The supernatant is then spiked with intact plasmid DNA of known size and incubated for ~1 hr at 37°C. Agarose gel electrophoresis (UNIT 2.5A) is then used to evaluate integrity of DNA incubated with and without supernatant. The supernatant-treated DNA should show no evidence of degradation.

When aliquoting cells at the single-cell level, many of the tubes contain no cells. Some of these samples, as well as samples spiked with HSV DNA, are utilized as controls to test the completeness of the DNase treatment (see Critical Parameters and Troubleshooting). The ability of the immobilized DNase to eliminate potential contamination should be determined by spiking a sample with the DNA sequence being amplified. It is important to spike the sample with an amount of DNA that would reflect anticipated levels of contamination. With proper technique, these levels should be extremely low and not present in every cell sample. The DNase step is an important safeguard, but not a solution for poor technique.

4. Place samples in a PCR Gene Amp 2400 or equivalent and heat to 94°C for 5 min to inactivate DNase. Reduce temperature to 50°C, add 34 µl PCR/PK solution to each sample, and incubate 3 hr.
5. Prepare standards and quantitate as described (UNIT 15.7; Sawtell and Thompson, 1992). Prepare standard dilutions representing 10,000, 1,000, 100, 10, and 0 HSV viral genomes in 6 µl.

Standards are treated identically to the cell samples with the exception of DNase treatment. To quantify other nucleic acids of interest, use appropriate standards and optimized PCR assays (UNIT 15.1).

6. Incubate samples and standards 7 min at 94°C to inactivate proteinase K.
7. Incubate at 63°C while adding 10 µl PCR amplification solution and 1.25 U *Taq* DNA polymerase per reaction (50 µl total).
8. Amplify using the following program parameters:

45 cycles:	30 sec	94°C	(denaturation)
	30 sec	55°C	(annealing)
	30 sec	72°C	(extension)
Final step:	7 min	72°C	(extension/hold).

PCR conditions should be optimized for the primer/target of interest as described in UNIT 15.1.

9. Electrophorese 5 µl each PCR product through a nondenaturing 12% polyacrylamide gel (UNITS 2.5A & 2.7), transfer to a Gene screen plus nylon membrane (UNIT 2.9A), and perform hybridization analysis (UNIT 2.9A & 6.4) using a ³²P-end-labeled oligonucleotide internal to the PCR primers (UNIT 15.7).
10. Expose blot to a storage phosphor screen (APPENDIX 3A) and analyze using Imagequant software.

BASIC PROTOCOL 3

ANALYSIS OF ENRICHED CELL POPULATIONS BY RT-PCR

Presented in this section is a protocol that can be adapted to examine either specific or general transcriptional patterns in groups of selected populations of cells harvested from solid tissues. The cells in the tissue are first stabilized by fixation, avoiding the transcriptional changes that would occur with the manipulation and dissociation of living cells. Using carefully screened reagents, it is possible to maintain the integrity of the RNA within the cells during the dissociation process so that RT-PCR analysis is possible (Sawtell, 1997). The goal of the author was to analyze the RNA contained within just a few cells using RT-PCR. In PCR analysis (see Basic Protocol 2), the integration of the pretreatment steps and the PCR reaction in a single assay tube was straightforward; however, in the case of RT-PCR, establishing the compatibility of all of the enzymatic steps required for the pretreatment, reverse transcription, and subsequent PCR without an extraction step was more challenging. The assay developed is presented below. This protocol has been successfully utilized to detect transcripts in samples of fewer than ten neurons. This approach has proven to be especially useful to examine cell type specific expression of transcripts within solid tissues. For example, the author used this approach to demonstrate that the expression of a novel stress-induced spliced form of a key transcription factor was restricted to the neurons in the trigeminal ganglion (unpub. observ.).

Primer selection will depend on the transcript of interest. The MacVector PCR primer selection program has proven to consistently yield primers that work well. If specific

transcripts are being analyzed, primers that span splice sites is a distinct advantage. If a nonspliced transcript is being amplified, it is imperative to include sufficient controls in which the reverse transcriptase has been omitted to rule out the possibility that DNA rather than RNA is being amplified. One limiting factor will be the length of product generated by the reverse transcription reaction. The author has had success using this direct fixed cell RT-PCR assay with primers to mouse genes that generate a 500-bp product.

Materials

Proteinase K solution (see recipe)
40 mM PMSF, fresh
RNase-free DNase I: 3 U RNase-free DNase I (Boehringer Mannheim)/25 mM DTT/ 2.5 U placental RNase inhibitor
8 pmols/ μ l reverse transcriptase primer
Reverse-transcription reaction mix (see recipe)
200 U/ μ l SuperScript II reverse transcriptase (Life Technologies)
PCR amplification solution (see recipe)
1.25 U *Taq* DNA polymerase (Life Technologies)
PCR tubes

Additional reagents and equipment for obtaining dissociated perfusion-fixed cells (see Basic Protocol 1) and PCR optimization (*UNIT 15.1*)

1. Obtain cells dissociated from perfusion fixed tissues as detailed above (see Basic Protocol 1) using solutions tested to be free of RNase activity.

At this point, immobilized RNase could be utilized to remove any contaminating RNA from the aliquoted cells, as detailed above (see Basic Protocol 2, step 3); however, the author has tested extensively for specific RNAs in the supernatant of washed, dissociated cells and has not detected extracellular RNA contamination. This could reflect the inability of the reverse transcription reaction to detect one or just a few template molecules. In contrast, HSV DNA could occasionally be detected in the supernatant; therefore, eliminating it was imperative.

2. Aliquot cells to be analyzed in a 1- μ l volume into PCR tubes. Add 4 μ l proteinase K solution and incubate 60 min at 50°C. After digestion, add 0.25 μ l freshly prepared 40 mM PMSF.

Preliminary analysis demonstrated the need for protease digestion of cellular proteins in isolated cells for complete DNase I digestion of genomic DNA; however, the high temperatures required for heat inactivation of this enzyme led to degradation of RNA, most likely through metal ion-catalyzed hydrolysis. Thus, following digestion with proteinase K, activity of this enzyme is selectively inhibited by adding freshly prepared PMSF.

3. Add 0.75 μ l RNase-free DNase I. Incubate 45 min at 37°C.
4. Inactivate DNase by incubating 15 min at 70°C. After this time, add 0.25 μ l of 8 pmol/ μ l (2 pmol total) reverse transcriptase primer and incubate an additional 10 min at 70°C.
5. Reduce temperature to 50°C and add 3.5 μ l reverse-transcription reaction mix, followed by an additional 0.25 μ l of 40 mM PMSF and 0.25 μ l of 200 U/ μ l (50 U) SuperScript II reverse transcriptase. Incubate 60 min at 50°C.

If the transcripts are to be detected are unspliced, samples are set up in multiples, half of which receive no reverse transcriptase. In additional controls, RNase is included with the DNase (step 3).

6. After 60 min, increase temperature to 70°C for 15 min. Add $\frac{1}{5}$ to $\frac{1}{2}$ of the cDNA sample to 47 μ l PCR reaction buffer and heat 5 min at 94°C.

7. Reduce the temperature to 5°C above the annealing temperature (*UNIT 15.1*) and add 1.25 U *Taq* DNA polymerase to each sample (50 µl total).
8. Analyze amplification products as described above (see Basic Protocol 2, steps 9 and 10)

The usefulness of the RT-PCR assay for quantification at the single-cell level has not been fully explored. Using primers to specific HSV genes, the reverse transcription reaction lacked the sensitivity required to detect the very low levels of these transcripts anticipated during viral latency. This may be due, in part, to the very high GC content of the HSV genome in general and the specific regions being reverse transcribed. Regardless, the assay can detect specific transcripts in small numbers of cells.

REAGENTS AND SOLUTIONS

Use 3×F H₂O in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Cresyl violet solution

Prepare the following in triple 0.2-µm filtered nanopure (3×F) H₂O:
0.5% (w/v) cresyl violet
10% (v/v) glacial acetic acid
Store up to 12 months at room temperature

DNase reaction buffer

Prepare the following in triple 0.2-µm filtered nanopure (3×F) H₂O:
20 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
5 mM MgCl₂
5 mM CaCl₂
Aliquot and store up to 12 months at –20°C

Hanks balanced salt solution (HBSS)

0.4 g/liter KCl
0.06 g/liter KH₂PO₄
8.00 g/liter NaCl
0.35 g/liter NaHCO₃
0.048 g/liter Na₂HPO₄
1.00 g/liter D-glucose
Sterilize by passing through three 0.2-µm filters
Aliquot and store up to 12 months at –20°C

PCR amplification solution

Prepare the following in triple 0.2-µm filtered nanopure (3×F) H₂O:
20 mM Tris·Cl, pH 8.4 (*APPENDIX 2*)
50 mM KCl
1.5 to 4.5 mM MgCl₂
5% (w/v) gelatin
200 µM each dNTP
25 to 50 pmols of each primer (*UNIT 15.7*; Katz et al., 1990)
Store up to 1 month at –20°C

While the buffer can be stored with primers and dNTPs, it is better to add them just before use. Buffer without primers or dNTPs can be stored up to 12 months at –20°C.

*The concentration of MgCl₂ will depend on specific primers utilized (*UNIT 15.1*) but will commonly range between 1.5 to 4.5 mM.*

PCR/PK solution

Prepare the following in triple 0.2- μ m filtered nanopure (3 \times F) H₂O:
20 mM Tris·Cl, pH 8.4 (*APPENDIX 2*)
50 mM KCl
1.4 to 4.5 mM MgCl₂
Aliquot and store up to 12 months at -20°C
Just before use, add 50 $\mu\text{g}/\text{ml}$ proteinase K

The concentration of MgCl₂ will depend on specific primers utilized (UNIT 15.1) but will commonly range between 1.5 to 4.5 mM.

Ponceau S solution

Prepare the following in triple 0.2- μ m filtered nanopure (3 \times F) H₂O:
0.5% (w/v) Ponceau S
1% (v/v) glacial acetic acid
Store in aliquots up to 12 months at room temperature.

Proteinase K solution

25 mM Tris·Cl, pH 8.4 (*APPENDIX 2*)
37 mM KCl
1.5 mM MgCl₂
0.3 μg proteinase K
Make fresh

Reverse-transcription reaction mix

93 mM Tris·Cl, pH 8.3 (*APPENDIX 2*)
140 mM KCl
5.5 mM MgCl₂
Store up to 12 months at -20°C
Just before use add DTT to 25 mM and dNTPs (*UNIT 3.4*) to 0.25 mM

Xgal buffer

Prepare the following in “clean” phosphate buffered saline, pH 7.4 (PBS;
APPENDIX 2):
5 mM K₃Fe(CN)₆ (potassium ferrocyanide)
5 mM K₄Fe(CN)₆·3H₂O (potassium ferricyanide)
2 mM MgCl₂
Aliquot and store up to 3 months at room temperature

COMMENTARY

Background Information

The concept of “cellular pathology” was put forth nearly 150 years ago by Virchow (1863) with the view that disturbances in structure and function of individual cells form the basis of disease. Current understanding of the interactive nature of the cells comprising an organism have substantiated this view. It is now clear that cells differentiate and function according to the summation of the molecular cues arising from many other cells in the organism (Gilbert, 1994). It follows that certain important aspects of the molecular behavior of individual cellular components can only be observed in the context of the organism.

Reported here is a strategy, contextual expression analysis (CXA), that combines the cell-specific information of in situ approaches with the analytical and quantitative potential of solution PCR (Sawtell, 1997). Cells are chemically stabilized in the context of the organism and subsequently isolated. PCR can then be utilized to gain insight into the molecular processes of a single cell among billions.

The enzymatic dissociation of living tissues has been widely used and refined for many specific tissue types (Pretlow and Pretlow, 1982). Inherent in this process are cellular molecular changes induced by disruption of context. In order to prevent these changes, tissues

are stabilized by chemical fixation prior to dissociation. Yields of even fragile adult cell types such as neurons are high. Distinct morphological features, such as brush borders of the intestinal epithelial cells as well as nuclear and cytoplasmic nucleic acid staining patterns are comparable to those in sectioned tissues (Sawtell, 1997). DNA and RNA isolated from dissociated tissues are reasonably intact. By immunohistochemical staining, the distribution of cytoskeletal proteins including actin and neurofilament 200 kDa peptide are similar in dissociated cells when compared to sectioned tissue (Sawtell, 1997).

The impetus for development of this protocol arose out of the need in the author's laboratory to identify the cell types and to quantify the number of cells in a specific solid tissue that harbored the latent HSV genome. It was also important to determine the number of viral genomes in each of those cells. The approach has proven extremely useful for this purpose (Sawtell, 1997; Thompson and Sawtell, 1997, 2000, 2001; Sawtell et al., 1998; Sawtell et al., 2001); however, this method should be widely useful for facilitating the analysis of rare cells or cellular events occurring in a complex multicellular environment.

Critical Parameters and Troubleshooting

In developing this procedure, several common fixation formulations were tested in combination with alternative digestive enzymes. For the most part, tissues either remained a solid mass or disintegrated into cellular debris; however, perfusion with STF followed by digestion with collagenase (i.e., Worthington CLSI) yielded single cell suspensions from peripheral and central nervous tissue, lung, liver, intestine, heart, pancreas, muscle, and reproductive tract. Nonetheless, optimizing the balance between fixation and dissociation for the specific tissue of interest is advisable.

Fixation

The volume of fixative perfused through the animal is critical and should be measured. The procedure can be modified for larger animals, such as guinea pigs, by increasing the volume of fixative utilized.

Several different types of fixatives were tested, including various formaldehyde based formulations. The fixative found to give the best results was Streck tissue fixative (STF). This is a noncrosslinking fixative containing diazolid-

inyl urea, 2-bromo-2-nitropropane-1,3-diol (bronopol), zinc sulfate, and sodium citrate. Why STF works in Basic Protocol 1 has not been explored; however, it is likely that absence of crosslinking in the fixed tissue is favorable for the subsequent enzymatic dissociation process.

Mincing

It is critically important to finely and uniformly mince the tissue. This allows greater and more uniform access of dissociating enzymes to the tissue.

Collagenase

The collagenase preparation used contains several collagenases, as well as caseinase, clostripain, and tryptic activities. This is a relatively crude preparation and there is lot-to-lot variation in this product. The author's laboratory has tested a number of purified enzymes with varying levels of success; however, none worked as well as the crude preparation. Batches of collagenase must be screened not only for optimum dissociation activity but also for DNase and RNase activity, depending upon final use of cells. This can be done easily by spiking an aliquot of the enzyme (5× strength) with DNA or RNA. The sample is then incubated at 37°C for 1 hr and examined by agarose gel electrophoresis for integrity of the nucleic acids. While many batches will be free of DNase, RNase-free collagenase is less common.

Contamination

Careful planning to prevent contamination control is necessary for the success of this procedure. *UNIT 15.7* discusses many of the relevant issues. It is critical that the results obtained from the PCR be related to the contents of the cell being analyzed and not contamination introduced at any point during the procedure. With all aspects of this procedure, the final use of the cells will determine the types of contamination that must be avoided. The most meticulous technique is required if downstream applications require intact RNA. The introduction of RNases at any point must be avoided.

Controls

Observe the tissue and cells of interest frequently during the dissociation process. Such observation will provide important information regarding the response of the dissociating tissue and cells to the process. Make estimates of the

number of cells expected and determine if recoveries are reasonable. There should be minimal cell loss with this method.

Examine the integrity of the nucleic acids and proteins by carrying out routine isolation and analytical procedures on the tissue after dissociation.

All standard PCR and RT-PCR controls should be included. Additional controls will be required, depending on the final application of the cells.

Anticipated Results

Single cell suspensions of tissues will be obtained. With proper attention to the quality of reagents used, specifically the collagenase, the nucleic acids within the cells will be intact. Keep in mind that the results obtained from downstream analyses will depend on the quality of the cells obtained from the dissociation step.

Time Considerations

The time required will depend on the skill level of the individual performing the task. For the skilled practitioner, mouse perfusion will require 10 to 15 min, dissection and mincing of tissue 5 to 15 min (depending on how many tissues are being dissected), tissue dissociation 30 to 45 min, rinsing and post-fixation 20 min, neuron counts ~1 hr (depending on how many slides are examined), separation of Percoll gradient and rinsing 1 hr, aliquoting individual cells one to several hours (depending on the number of tubes), PCR (including all pretreatments) 1.5 to 2 days (this is not a continuous effort), and gel electrophoresis, prehybridization, and hybridization 1 to 1.5 days (this also is not a continuous effort).

Literature Cited

- Coligan, J.E., Kruisbeek, A.M., Margulies, D.H., Shevach, E.M., and Strober, W. (eds.) 2001. *Current Protocols in Immunology*. John Wiley & Sons, New York.
- Gilbert, S. 1994. *Developmental Biology* 4th ed. Sinauer Associates, Inc., Sunderland, Mass.
- Katz, J.P., Bodin, E.T., and Coen, D.M. 1990. Quantitative polymerase chain reaction analysis of herpes simplex virus DNA in ganglia of mice infected with replication-incompetent mutants. *J. Virol.* 64:4288-4295.

Mullis, K.B. and Falona, F.A. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Meth. Enzymol.* 155:335-350.

Pretlow II, T.G. and Pretlow, T.P. (eds.) 1982. *Cell Separation. Methods and Selected Applications*. Academic Press, New York.

Sawtell, N.M. 1997. Comprehensive quantification of herpes simplex virus latency at the single cell level. *J. Virol.* 71:5423-5431.

Sawtell, N.M. and Thomson, R.L. 1992. Herpes simplex virus type 1 latency-associated transcription unit promotes anatomical site-dependent establishment and reactivation from latency. *J. Virol.* 66:2157-2169.

Sawtell, N.M., Poon, D.K., Tansky, C.S., and Thompson, R.L. 1998. The latent HSV-1 genome copy number in individual neurons is virus strain specific and correlates with reactivation. *J. Virol.* 72:5343-5350.

Sawtell, N.M., Thompson, R.L., Stanberry, L.R. and Bernstein, D.I. 2001. Early intervention with high-dose acyclovir treatment during primary herpes simplex virus infection reduces latency and subsequent reactivation in the nervous system in vivo. *J.I.D.* 184:964-971.

Thompson, R.L. and Sawtell, N.M. 1997. The herpes simplex virus type 1 latency-associated transcript gene regulates the establishment of latency. *J. Virol.* 71:5432-5440.

Thompson, R.L. and Sawtell, N.M. 2000. Replication of herpes simplex virus type 1 within the trigeminal ganglia is required for high frequency but not high viral genome copy number latency. *J. Virol.* 74:965-974.

Thompson, R.L. and Sawtell, N.M. 2001. Herpes simplex type 1 latency-associated transcript gene promotes neuronal survival. *J. Virol.* 75:6660-6675.

Virchow, R. 1863. *Cellular pathology: as based upon physiological and pathological histology*. 2nd ed. translated by F. Chance, J.B. Lippincott, Philadelphia.

Key References

Sawtell, N.M. 1997. See above.

This manuscript describes the procedure as used to quantify viral latency and includes several critical validation experiments.

Contributed by N.M. Sawtell
Children's Hospital Medical Center
Cincinnati, Ohio

MOLECULAR METHODS FOR DISCOVERY OF DIFFERENTIALLY EXPRESSED GENES

SECTION B

Production of a Subtracted cDNA Library

UNIT 25B.1

PRODUCTION OF A SUBTRACTED LIBRARY

BASIC
PROTOCOL

For some experiments, a complete cDNA library (UNIT 5.8A) is unnecessary and instead, a subtracted cDNA library is useful. A subtracted cDNA library contains cDNA clones corresponding to mRNAs present in one cell or tissue type and not present in a second type. This cDNA library is used to isolate a set of cDNA clones corresponding to a class of mRNAs, or to aid in the isolation of a cDNA clone corresponding to a particular mRNA where the screening procedure for the cDNA clone is laborious because a specific DNA or antibody probe is unavailable. A technique known as differential screening is an alternative to creating subtracted libraries (see Commentary).

In this protocol, the tissue, library, RNA, or cDNA designated with a [+] contains the target or desired sequence(s), and that which is to be subtracted from the [+] is termed [-]. Since relatively few recombinants are obtained after subtraction, this protocol is for a cDNA library constructed in the λ gt10 vector or its equivalent, which allows a high cloning efficiency and permits elimination of nonrecombinants; however, the protocol can be used to produce subtracted cDNA libraries in any vector system.

[+] cDNA with *EcoRI* ends and [-] cDNA with blunt ends are prepared. The [-] cDNA is digested with *RsaI* and *AluI* to give small blunt-ended fragments. The [+] cDNA inserts are mixed with a 50-fold excess of fragmented [-] cDNA inserts, the DNAs in the mixture are heated to melt the double-stranded DNA, and the single-stranded insert DNA is allowed to hybridize. After hybridization, annealed cDNA inserts are ligated to λ gt10 arms, packaged, and transfected.

The only [+] cDNA likely to regenerate double-stranded fragments with an *EcoRI* site at each end are those sequences for which no complementary fragments were present in the [-] cDNA. The subsequent cloning step allows the selection and amplification of these fragments.

Materials

- [+] and [-] cDNA libraries (ATCC or Stratagene)
- TE buffer (APPENDIX 2)
- EcoRI* and 10 \times *EcoRI* buffer (UNIT 3.1)
- 0.5 M EDTA, pH 8.0 (APPENDIX 2)
- 10% sucrose solution (UNIT 5.3)
- 1.5% and 2% agarose gels (UNIT 2.5A)
- TBE buffer (APPENDIX 2)
- 95% and 70% ethanol
- S1 nuclease (Sigma; UNIT 3.12) and 10 \times S1 nuclease buffer (UNIT 3.4)
- 25:24:1 phenol/chloroform/isoamyl alcohol (UNIT 2.1A)
- 3 M sodium acetate, pH 5.2 (APPENDIX 2)
- AluI* and 10 \times *AluI* buffer (UNIT 3.1)
- RsaI* (UNIT 3.1)
- Deionized formamide (Fluka, IBI, or American Bioanalytical)
- 20 \times SSC (APPENDIX 2)
- 1 M NaPO₄, pH 7.0 (see recipe)

Discovery of
Differentially
Expressed Genes

25B.1.1

Contributed by Lloyd B. Klickstein

Current Protocols in Molecular Biology (2001) 25B.1.1-25B.1.8

Copyright © 2001 by John Wiley & Sons, Inc.

Supplement 55

10% sodium dodecyl sulfate (SDS)
10 mg/ml yeast tRNA
24:1 chloroform/isoamyl alcohol
Phosphatase λ gt10 arms (Stratagene)
10 \times T4 DNA ligase buffer (UNIT 3.4)
T4 DNA ligase (measured in cohesive-end units; New England Biolabs; UNIT 3.14)
E. coli C600/hflA (Table 1.4.5)
 λ phage packaging extracts (Stratagene)
Suspension medium (SM; UNIT 1.11)

SW-28 rotor and 38-ml centrifuge tubes (Beckman) or equivalent
0.4-ml microcentrifuge tube

Additional reagents and equipment for construction of recombinant DNA libraries (UNITS 5.5 & 5.6), large-scale DNA preps from plasmids (UNIT 1.7) or phage (UNIT 1.13), sucrose gradients (UNIT 5.3), agarose gel electrophoresis (UNIT 2.5A), production and growth/maintenance of λ phage libraries (UNITS 5.8, 25B.2, and 1.9-1.13), plating and titering libraries (UNITS 6.1 & 6.2), hybridization (UNIT 6.3), and radiolabeling probes (UNIT 3.4)

Prepare the insert DNA

1. Prepare or obtain cDNA libraries from the [+] and [-] cells or tissue sources.

A major advantage of this protocol is that a subtracted library may be prepared from existing libraries, which is highly recommended. Complementary DNA libraries from many species and tissue sources are widely available and considerable time may be saved by obtaining preexisting [+] and [-] libraries to be used in this protocol.

Alternatively, prepare $\geq 1 \mu\text{g}$ [+] cDNA with EcoRI ends and $10 \mu\text{g}$ [-] cDNA with blunt ends (stop the [-] cDNA synthesis before adding linkers) from poly(A)⁺ [+] and poly(A)⁺ [-] RNA, respectively (UNITS 5.5 & 5.6). If this is done, proceed to step 13.

The protocol assumes that the [+] and [-] libraries are bacteriophage λ libraries. If the vector for either is a plasmid, only 100 μg of each is needed (scale down steps 2 and 3 by $\frac{1}{10}$) and the inserts should be purified by agarose gel electrophoresis rather than by sucrose gradient centrifugation.

2. Perform large-scale (2 to 3 liters) DNA preps of both the [+] and [-] libraries to obtain >1 mg DNA from each library. Resuspend the DNA at 1 mg/ml in TE buffer.

Digest the DNA

3. Digest 1 mg of each library DNA in a 1.5-ml microcentrifuge tube as follow (final volume 1.167 ml):

1 ml library DNA (1 mg)
0.117 ml of 10 \times EcoRI buffer
0.05 ml EcoRI (1000 U).

Mix by shaking and incubate 5 hr at 37°C. Stop the reaction by adding 40 μl of 0.5 M EDTA, pH 8.0, and incubate 10 min at 65°C. During the digestion, prepare four 10% to 40% sucrose gradients in 38-ml SW-28 tubes (UNIT 5.3). Label two tubes [+] and two tubes [-].

Internal EcoRI sites present in the cDNA inserts will be cut. If this occurs, the partial-length cDNA clone obtained through this procedure can be used to generate a probe with which to screen the initial [+] library for a full-length clone. The advent of newer vectors (e.g., λ ZAP; see Fig. 1.10.8) with cloning sites for enzymes such as NotI will nearly eliminate this difficulty.

Run the sucrose gradients

4. Mix each digest with an equal volume of 10% sucrose solution and carefully layer the digested [+] library DNA onto the two 10% to 40% sucrose gradients labeled [+]. Split the sample between the two tubes evenly. Similarly, load the [-] DNA onto the two gradients labeled [-]. Centrifuge the gradients overnight (18 to 24 hr) at 122,000 × g (26,000 rpm in an SW-28 rotor), 20°C.

The insert fragments will remain near the top of the gradient while the phage arms will migrate half the length of the tube.

5. Harvest the gradients by gently removing 0.2-ml fractions from the top of the tube with a pipettor. Place each fraction into a separate, labeled microcentrifuge tube at 4°C.

Twenty fractions per gradient are sufficient, as the insert DNA is small and barely enters the gradient under these conditions. Save the remainder of the gradient until the fractions containing the inserts have been identified, just in case!

Recover the DNA

6. Identify the tubes containing the insert DNA by analyzing 20 µl of every other fraction on a 1.5% agarose gel made in TBE buffer.
7. Precipitate the insert DNA: add 0.3 ml TE buffer and 1.0 ml of 95% ethanol to each tube, mix, and place at -20°C for 2 hr or on dry ice for 15 min.

Because the sucrose gradient buffer contains 1 M NaCl, there is sufficient NaCl in the fractions for precipitation of the DNA. The sucrose in the gradient fractions must be diluted in order to successfully precipitate the DNA. A 2- to 3-fold dilution of these low-density fractions is adequate. A greater dilution of the higher density fractions is necessary in order to obtain high recoveries of DNA following ethanol precipitation.

8. Thaw and collect the DNA by microcentrifugation at high speed for 15 min. Aspirate the supernatant and save until DNA recovery has been checked. Add 0.5 ml of 70% ethanol to each tube. Recentrifuge, aspirate the ethanol supernatants, and dry the pellets.
9. Resuspend and pool the fractions containing insert DNA from the [+] library in TE such that the final concentration is 0.2 mg/ml. Store the DNA at -20°C.
10. Resuspend and pool the insert DNA from the [-] library in 100 µl TE buffer and place on ice. Save an aliquot of 400 ng of each [+] and [-] cDNA separately, to be used in evaluating the final library produced.

Expect recoveries of >10 to 15 µg of insert DNA from 1 mg of total library DNA. The aliquots of [+] and [-] DNA, alternatively, may be radiolabeled and used as probes for differential screening of the [+] library (see Commentary).

Remove EcoRI ends from [-] DNA

11. Remove the EcoRI ends from the [-] DNA by mixing in the following order (final volume 112 µl):

100 µl [-] insert DNA (10 to 15 µg)
11 µl 10× S1 nuclease buffer
1 µl 1:500 S1 nuclease (2 U).

Mix by vortexing, briefly microcentrifuge, and incubate 30 min at 37°C.

12. Stop the reaction by adding:

5 μ l 0.5 M EDTA, pH 8.0
200 μ l TE buffer
300 μ l phenol/chloroform/isoamyl alcohol.

Vortex. Microcentrifuge 1 min to separate the phases and transfer the upper, aqueous phase to a new tube. Add 30 μ l of 3 M sodium acetate, pH 5.2, and 700 μ l ethanol. Freeze, then collect the DNA by centrifugation as in steps 7 and 8. Resuspend the washed and dried pellet in 100 μ l TE buffer.

Digest the [–] DNA with AluI and RsaI

13. Digest the S1 nuclease–treated [–] insert DNA to small fragments with *AluI* and *RsaI* by adding in the following order (final volume 121 μ l):

100 μ l [–] insert DNA (10 to 15 μ g)
12 μ l 10 \times AluI buffer
5 μ l *AluI* (50 U)
4 μ l *RsaI* (60 U).

Mix by vortexing, briefly microcentrifuge, and incubate 3 hr at 37°C. Add 5 μ l of 0.5 M EDTA, pH 8.0, and incubate 10 min at 65°C to stop the reaction. Remove and save 5 μ l of the digest for evaluation by electrophoresis.

14. Add 200 μ l TE buffer and 300 μ l phenol/chloroform/isoamyl alcohol; extract and ethanol precipitate the DNA as in step 12. Resuspend the washed and dried pellet in TE buffer at 1.0 μ g/ μ l.
15. Check the 5- μ l aliquot from step 13 by running a 2% agarose minigel (in TBE buffer) and ethidium bromide–staining. The [–] DNA fragments should be between 50 and 200 bp in length.

Hybridize the DNA

16. Hybridize the [+] insert DNA with the [–] DNA fragments. Add in the following order to a 0.4-ml microcentrifuge tube (final volume 51 μ l):

25 μ l deionized formamide (50% vol/vol final)
10 μ l [–] DNA fragments (10 μ g)
1 μ l [+] insert DNA (0.2 μ g)
12.5 μ l 20 \times SSC (5 \times final)
0.5 μ l 1 M NaPO₄, pH 7.0 (10 mM final)
0.5 μ l 0.1 M EDTA, pH 8.0 (1 mM final)
0.5 μ l 10% SDS (0.1% final)
1.0 μ l 10 mg/ml yeast tRNA (0.2 mg/ml final).

Mix by vortexing, briefly microcentrifuge, and place tube in a bath of boiling water for 5 min. Briefly microcentrifuge again and incubate 18 to 24 hr at 37°C.

The boiling step melts the DNA strands. During the hybridization step, only [+] sequences not present in the [–] DNA will find their complementary strands and regenerate clonable, double-stranded fragments with EcoRI ends. A [+] sequence also present in the [–] DNA will hybridize with at least one of the AluI/RsaI [–] fragments, forming a partially single-stranded, partially double-stranded molecule without clonable ends.

17. Add 200 μ l TE buffer and transfer the mixture to a 1.5-ml microcentrifuge tube. Wash the hybridization tube with 250 μ l TE buffer and add it to the hybridization mix (the volume is now 500 μ l). Add 500 μ l phenol/chloroform/isoamyl alcohol, vortex, and microcentrifuge 1 min to separate the phases.

- Transfer the upper, aqueous phase to a new tube. Reextract this phase with 500 μ l chloroform/isoamyl alcohol as in step 17. Recover the aqueous phase and add 50 μ l of 3 M sodium acetate, pH 5.2, and 1 ml ethanol. Precipitate as in steps 7 and 8. Resuspend the washed and dried pellet in 12 μ l TE buffer.

Chloroform/isoamyl alcohol extraction ensures removal of SDS and formamide.

Ligate the DNA

- Ligate the insert DNA to λ gt10 (not λ gt11) phage arms by adding in the following order (final volume 25 μ l):

12 μ l insert DNA
10 μ l λ gt10 phosphatased phage arms (10 μ g)
2.5 μ l 10 \times ligase buffer
0.5 μ l T4 DNA ligase (200 U).

Mix by gently pipetting up and down and incubate overnight at 12 $^{\circ}$ to 15 $^{\circ}$ C.

Package and plate the library

- Start a fresh overnight culture of *E. coli* C600\lambda phage packaging extracts according to manufacturer's instructions.

The vector λ gt10 is used here because it permits selection of recombinants when grown on the appropriate host. Ten micrograms of bacteriophage λ vector is roughly an equimolar amount of EcoRI ends with respect to the input [+] DNA, the ends of which must be considered even though only a small fraction of the [+] insert DNA is clonable after the melting and hybridization steps. The recommended 10 μ g of vector and no less than 8 to 10 packaging extracts will ensure a library of good complexity.

- Add suspension medium (SM) to the packaging mixtures and pool them in a 5-ml polypropylene tube to a final volume of 2 ml. Add two drops of chloroform, shake by hand for 3 sec, and allow the chloroform to settle.
- Plate 0.2 ml packaged phage with 3 ml fresh C600^{\circ}C.
- The following morning, count the number of plaques on a representative plate and multiply by 10 to determine the total number of recombinants in the library.

Typically, 300 to 15,000 phage per library are obtained.

- Elute the plates with SM as in UNIT 25B.2 or directly select individual plaques for screening.

Evaluate the library

- Evaluate a newly prepared subtracted library as described in UNIT 5.8A (first support protocol).

The best approach is to amplify the library and differentially screen duplicate nitrocellulose filters from a single 150-mm plate of 20,000 to 40,000 recombinants. Hybridize one lift with a total [+] cDNA probe and the other with a total [-] cDNA probe. The total [+] and [-] cDNA probes are prepared by radiolabeling some of the [+] and [-] cDNA saved from step 10. Most clones should hybridize with the [+] probe and few with the [-] probe. Evaluation by screening the library with a probe for proteins such as actin or tubulin would not be appropriate, since the expected result is no hybridization (or only a few), which may occur for a variety of reasons.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

1 M NaPO₄, pH 7.0

A: 1 M Na₂HPO₄

B: 1 M NaH₂PO₄

Add B to A until pH = 7.0

COMMENTARY

Background Information

The practical consequence of creating a subtraction library is considerable enrichment of the target cDNA clones. For example, a subtracted cDNA library was used to isolate T cell antigen receptor cDNAs. By hybridizing T cell cDNA to B cell mRNA, and selecting the single-stranded cDNA molecules by hydroxylapatite column chromatography, the T cell antigen receptor cDNAs were significantly enriched. The cDNA was then hybridized back with the T cell mRNA from which it was derived and the double-stranded RNA-DNA hybrids were selected, carried through second-strand cDNA synthesis, and the resulting cDNA was cloned (Hedrick et al., 1984). Thus, a large percentage of the clones in the subtracted library were T cell-specific. All clones in the subtracted library would have been present in a library constructed from the T cell line without the subtraction step; the objective was to obtain a library so enriched that clones derived from it could be screened by random selection.

Two major disadvantages to the approach outlined above are that poly(A)⁺ RNA from both [+] and [-] source is required and the hybridizations, hydroxylapatite columns, and library production with a very small amount of cDNA are technically difficult. A conceptually different approach, termed deletion enrichment, has been undertaken in the construction of a genomic library enriched for Y chromosome-specific sequences (Lamar and Palmer, 1984). In this case, [+] DNA (male) and an excess of [-] DNA (female) fragments were mixed, denatured, hybridized, and [+] DNA that did not hybridize to [-] was selected by a cloning step. Production and selection of a library were accomplished simultaneously.

The subtracted cDNA library protocol described here is an adaptation of the deletion enrichment method to cDNA libraries and is recommended over the other because (1) it is conceptually and practically simple, involving only standard laboratory techniques; (2) it does not involve the handling of RNA, which can be

problematic; (3) it may be performed with DNA prepared from already existing libraries, eliminating the potential time and expense involved in the preparation of fresh tissue or cells; and (4) with slight modifications, either cDNA or genomic subtracted libraries may be prepared.

A disadvantage of this approach is that if no [+] or [-] cDNA libraries are available, they must first be made or cDNA must be synthesized, requiring an extra few days to a week. A second disadvantage of this or any other subtraction protocol is that clones containing reiterated sequences—e.g., an *Alu* repeat in the 3' untranslated region—would be eliminated from the library on that basis. Thus, the representation of a clone containing a reiterated sequence would be lower than expected, with only partial-length cDNAs present after subtraction.

A good alternative to creating a subtracted library is differential screening of a library known to contain the target clone(s). In one well-characterized experiment, duplicate lifts from a lymphoid tissue cDNA library were screened with total B cell cDNA and total T cell cDNA probes. B cell-specific clones were identified as plaques that hybridized with a B cell cDNA probe and not with a T cell probe (Tedder et al., 1988). A potential drawback to the differential screening approach is that rare sequences will have very low specific probe concentrations in the mixture and thus might not hybridize to the DNA from a target plaque in a reasonable period of time (e.g., overnight).

Critical Parameters

Because the *Eco*RI ends of [+] insert cDNA must remain intact through the sucrose gradient and hybridization steps, the nuclease inhibitor and bacteriostatic agent EDTA is present in both steps. In contrast, S1 nuclease digestion destroys the *Eco*RI ends of [-] cDNA, ensuring that all clones in the final library are derived from [+] cDNA. Restriction digestion of [-] DNA with *Alu*I and *Rsa*I increases the molar ratio of [-] to [+] DNA while not increasing the

mass of [−] cDNA, which can inhibit subsequent steps. The boiling step prior to hybridization is essential to melt the cDNA to its single-stranded form. The hybridization conditions favor the annealing of fragments >50 bp and a relatively long hybridization is required to permit reannealing of rare [+] cDNA. Hybridization buffer must be diluted at least 10-fold in order to successfully phenol extract, chloroform extract, and ethanol precipitate the DNA. A full 10 μg of phage vector arms must be added in the ligation reaction. This represents only an equimolar amount of *EcoRI* ends with respect to the number of ends from the [+] cDNA, and is by no means an excess.

A major advantage of this protocol is that it may be performed with reagents, enzymes, and supplies routinely available in a typical molecular biology laboratory. Sucrose for gradients should be molecular biology grade (DNase-free). λgt10 phosphatased arms and packaging extracts should be obtained commercially unless large volume use is anticipated, in which case “homemade” arms and extracts would be more economical.

Troubleshooting

The initial *EcoRI* digestion, the sucrose gradients, and the *AluI* and *RsaI* digestion of [−] DNA are monitored by minigel electrophoresis. If difficulties such as incomplete digestion or poor separation occur here, see the commentaries of *UNITS 3.1 & 5.3*. Poor recovery of DNA is usually not a problem, since at least 10 μg of DNA is present at each precipitation. Once beyond these steps, there is no method for evaluation short of determining the titer and composition of the subtracted library. Possible adverse outcomes include too few clones, too many clones, or no enrichment for [+] clones (see Anticipated Results).

When too few clones are obtained, check the λgt10 phage arms, packaging extracts, and host bacteria by cloning a test insert. If $>1 \times 10^7$ PFU/μg test insert are obtained, the problem may be that the *EcoRI* ends of the [+] cDNA have been destroyed or there is an inhibitor of one of the later steps present in the DNA. Evaluate these possibilities by cloning [+] DNA after the sucrose step and measuring the efficiency, and by cloning a test insert with and without post-hybridization DNA added to the test insert ligation.

If too many clones are obtained, the problem is usually contamination of one or more reagents with phage, non-*E. coli* C600/*hflA* bacterial host, or failure to denature the [+] or [−]

DNA prior to hybridization. Too many clones may also be obtained if the S1 nuclease digestion of [−] DNA did not work, which can be evaluated by cloning some of the [−] DNA directly. Alternatively, check the S1 nuclease step by digesting some M13 DNA under the same conditions and monitoring the reaction by agarose gel electrophoresis.

If the subtraction did not work, duplicate filters screened with [+] and [−] total cDNA probes as described in step 20 will have roughly equal numbers of positive clones. The most likely explanation is that the S1 nuclease digestion was incomplete. Check the S1 step by cloning 100 ng of the [−] DNA; $<10^3$ PFU/μg insert is expected.

Anticipated Results

The number of clones obtained depends on the similarity of sources of [+] and [−] cDNA. For a subtracted cDNA library prepared from B cell [+] and T cell [−] cDNA, 5500 recombinants were obtained, 15% of which were immunoglobulin clones. Because of the high level of similarity between B cells and T cells, this result probably represents a minimum number to be expected. Investigators using this protocol have reported 300 to 15,000 phage per library. Twenty to fifty percent of the clones in a well-constructed library will be [+] specific; most of the remainder will be abundant cDNAs present in both [+] and [−] cDNA. Five to ten percent of clones *not* [+] specific may have inserts that are not released by *EcoRI* and probably represent aberrant ligation of [−] fragments into the vector.

Time Considerations

Once [+] and [−] total library DNA is obtained, perform the *EcoRI* digestion during the day and run the sucrose gradient overnight. On the second day harvest the gradients, perform the S1 digestion of the [−] cDNA, and store the precipitated DNA overnight. On the third day, digest the [−] DNA with *AluI* and *RsaI* and set up the hybridization overnight. The fourth day, ligate the cDNA to the λgt10 vector overnight and start an overnight culture of host cells. Package and plate over the following night. The protocol may be interrupted at any ethanol precipitation overnight or over the weekend.

Literature Cited

Hedrick, S.M., Cohen, D.I., Nielsen, E.A., and Davis, M.M. 1984. Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. *Nature (Lond.)* 308:149-153.

Lamar, E.E. and Palmer, E. 1984. Y-encoded, species-specific DNA in mice: Evidence that the Y chromosome exists in two polymorphic forms in inbred strains. *Cell* 37:171-177.

Tedder, T.F., Strueli, M., Schlossman, S.F., and Saito, H. 1988. Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen of

human B lymphocytes. *Proc. Natl. Acad. Sci. U.S.A.* 85:208-212.

Contributed by Lloyd B. Klickstein
Brigham and Women's Hospital
Boston, Massachusetts

PCR-Based Subtractive cDNA Cloning

Subtractive cloning is a powerful technique that allows isolation and cloning of mRNAs differentially expressed in two cell populations. In the generalized subtraction scheme illustrated in Figure 25B.2.1, the cell types to be compared are the [+] or tracer cells and the [-] or driver cells, where mRNAs expressed in the tracer and not the driver are isolated. Briefly, tracer nucleic acid (cDNA or mRNA) from one cell population is allowed to hybridize with an excess of complementary driver nucleic acid from a second cell population to ensure that a high percentage of the tracer forms hybrids. Hybrids that form include sequences common to both cell populations. Hybrids between the tracer and driver, and all driver sequences, are removed in the subtraction step. The unhybridized fraction is enriched for sequences that are preferentially expressed in the tracer cell population.

The method described here (see Basic Protocol) uses double-stranded cDNA (ds cDNA) as both tracer and driver and is modified from protocols devised by Sive and St. John (1988) and Wang and Brown (1991; see Background Information and Fig. 25B.2.2). Reciprocal subtractions are performed between two cell populations, A and B: that is, genes preferentially expressed in A more than in B are isolated, as are genes expressed

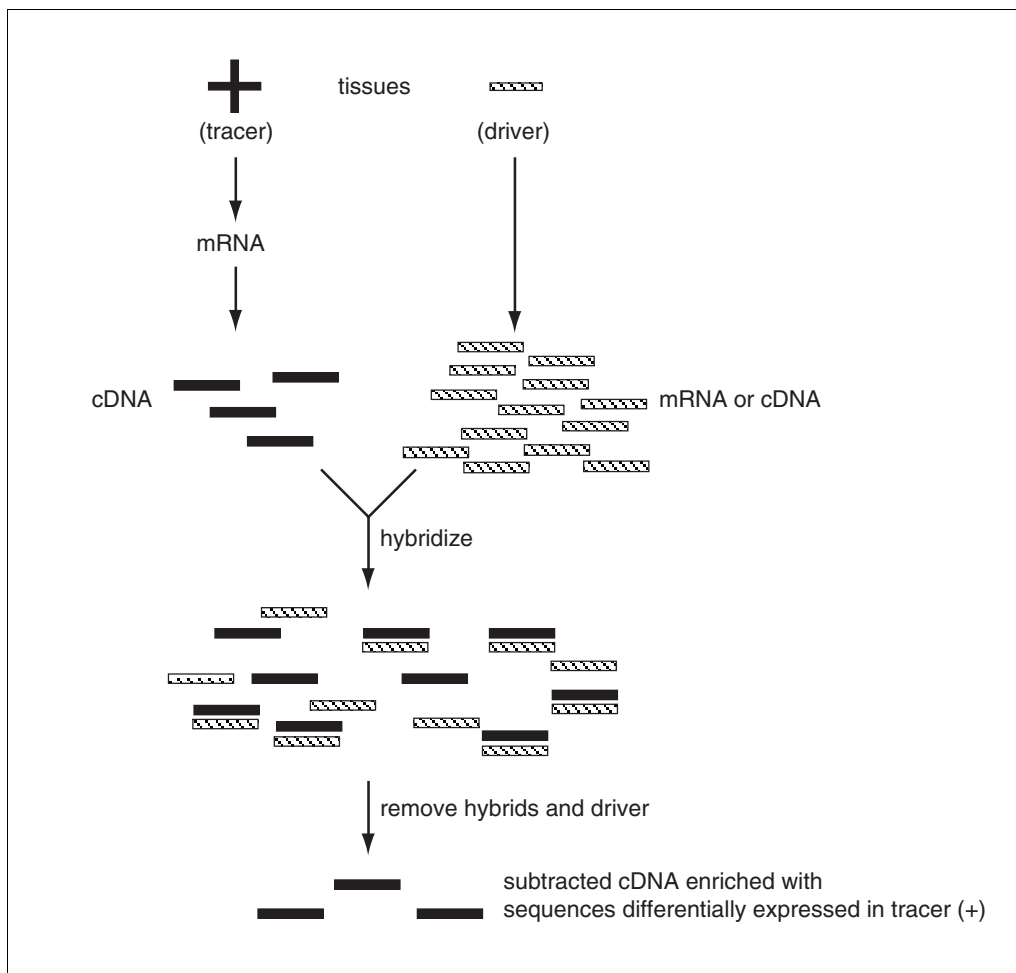


Figure 25B.2.1 Generalized subtraction scheme. Tracer cDNA from the + cell population is hybridized to >10-fold excess driver mRNA or cDNA from the - cell population. The resulting hybrids and excess driver are removed to enrich for cell type-specific sequences in the tracer. The subtraction may be repeated for further enrichment.

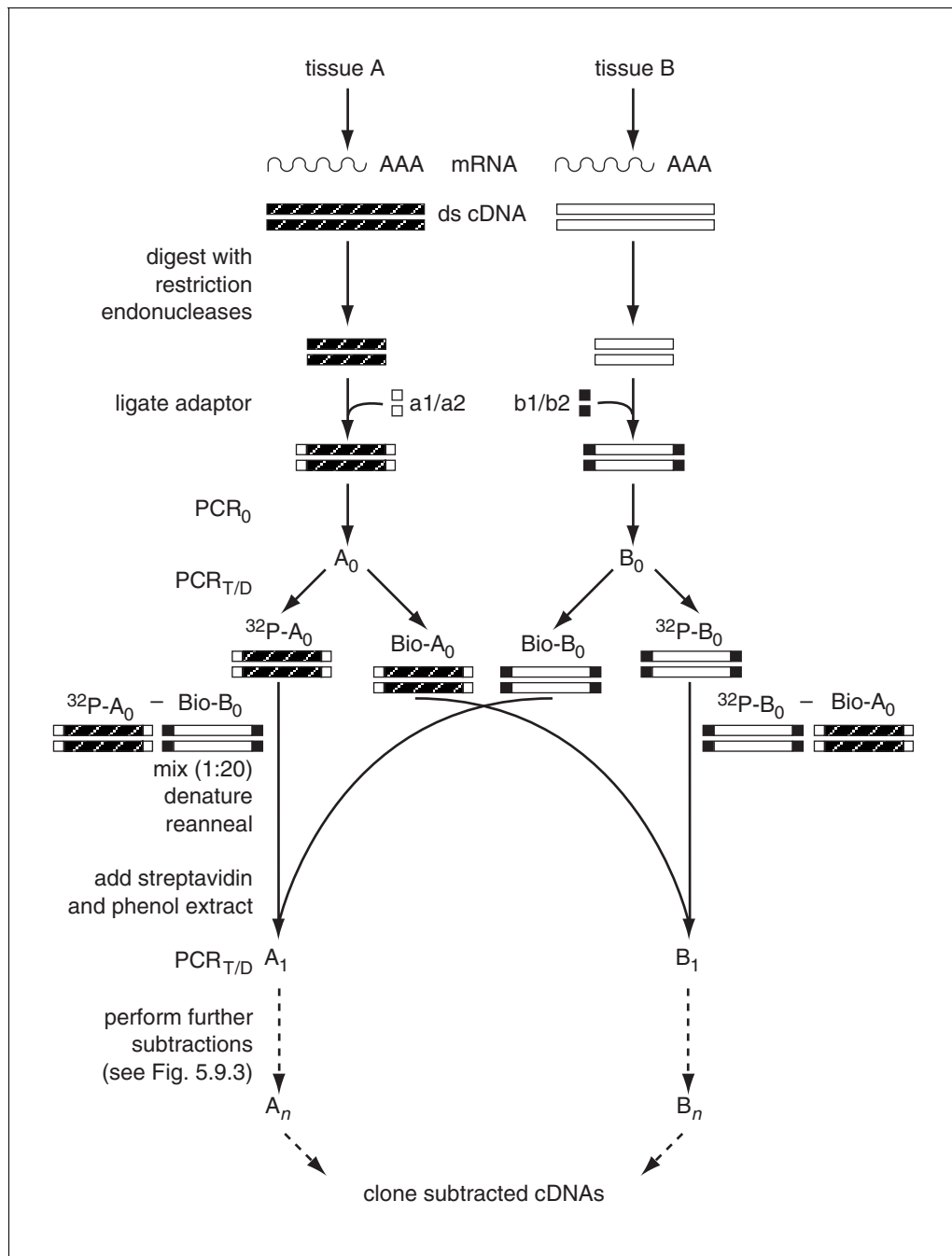


Figure 25B.2.2 Basic steps in PCR-based cDNA subtraction cloning. mRNAs purified from tissues A and B are used to synthesize double-stranded cDNA by standard methods. The resulting cDNAs are then digested with restriction endonucleases that have 4-bp recognition sequences. Two different sets of adaptors (a1/a2 and b1/b2) are ligated to the two sets of digested cDNA. The cDNAs are amplified with the appropriate primers (a1 or b1) to yield A_0 and B_0 . Two sets of subtractions are performed ($A_0 - B_0$ and $B_0 - A_0$). In each case the tracer is labeled with small amounts of [α - ^{32}P]dCTP, and the driver is labeled with bio-11-dUTP during PCR synthesis. Tracer and driver cDNAs are mixed at a ratio of 1:20, denatured, and allowed to reanneal. Driver/driver and tracer/driver hybrids are removed by treatment with streptavidin and extraction with phenol. This results in an enrichment of sequences found at greater abundance in tracer versus driver to yield A_1 and B_1 . Further subtractions are performed after another round of amplification using the appropriate cDNAs (see Fig. 5.9.3). When the subtractions are completed, the cDNAs are cloned into an appropriate vector for analysis.

preferentially in B more than in A. The method uses the polymerase chain reaction (PCR) to amplify cDNAs after each subtraction to prepare tracer and driver for the next subtraction. This makes it possible to begin with very small quantities of cells and, by performing repeated subtractions, achieve maximal enrichment of differentially expressed genes in both cell populations. The progress of subtraction is monitored by slot blot hybridization (see Support Protocol). Differentially expressed cDNA sequences are used to construct a subtracted cDNA library.

STRATEGIC CONSIDERATIONS

For this method ds cDNA, full-length (if possible) and prepared from cell types A and B using oligo(dT) as first-strand primer, is the starting material. The ds cDNA is digested by restriction endonucleases to obtain short cDNA fragments. This prevents preferential PCR amplification of naturally small cDNAs. Next, each of the two cDNA samples is ligated to different adapters and amplified by PCR to obtain a large amount of material.

In the first (and subsequent) PCR amplification step, both tracer and driver cDNAs are made for each cell type to allow subtractions in both directions. The first subtractions are A_0 tracer – B_0 driver and B_0 tracer – A_0 driver. Tracer cDNA is made partially radioactive so the success of subtractions can be monitored. Driver cDNA is biotinylated during PCR by incorporating bio-11-dUTP to provide a basis for separation of hybrids and driver. Tracer and driver are mixed, denatured, and allowed to reanneal at a driver cDNA/tracer cDNA ratio of 20:1 and a driver concentration of ≥ 2 mg/ml (or for a driver with fragment sizes of 200 bp, 15 μ mol/liter). In order to achieve this concentration, hybridizations are performed in small volumes (5 to 10 μ l). Subtractions are performed in driver excess to ensure that the reannealing rate is a function of the driver concentration only and to drive hybridization of tracer as close to completion as possible. Subtractions are performed either for a short period of time to remove sequences that are common to both A and B and abundant in both, or for much longer to remove rarer sequences that are common to both A and B (see Critical Parameters).

After annealing, tracer/driver and driver/driver hybrids are efficiently removed by addition of streptavidin (a protein that specifically and tightly binds biotin) and extraction with phenol. Biotinylated nucleic acid that has bound streptavidin is taken into the organic phase or remains at the interface (Sive and St. John, 1988). Unhybridized tracer or tracer hybrids are not removed by the streptavidin/phenol treatment because they are not biotinylated and so remain in the aqueous phase. This constitutes subtraction and enrichment for differentially expressed genes. cDNAs remaining after the first set of subtractions are termed A_1 and B_1 ; these are used for the next round of subtraction.

The subtraction sequence is shown in Figure 25B.2.3. The number of subtractions necessary depends primarily on the complexity of the cDNAs, where complexity refers to the total number of different cDNAs, or fragments of cDNA, from each cell type (Davidson, 1986). The complexity should not be confused with the number of differentially expressed cDNAs, which is only a subset of the total cDNA populations. The greater the complexity of the starting mRNA pool (or, in general, the greater the number of cell types contributing to the starting mRNA), the more subtractions will be required. Ideally, subtraction should be repeated until no more cDNA is removed after hybridization and/or until the subtracted cDNAs (A_n and B_n) do not cross-hybridize. In practice, with the scheme described here, this is usually between five and twenty subtractions.

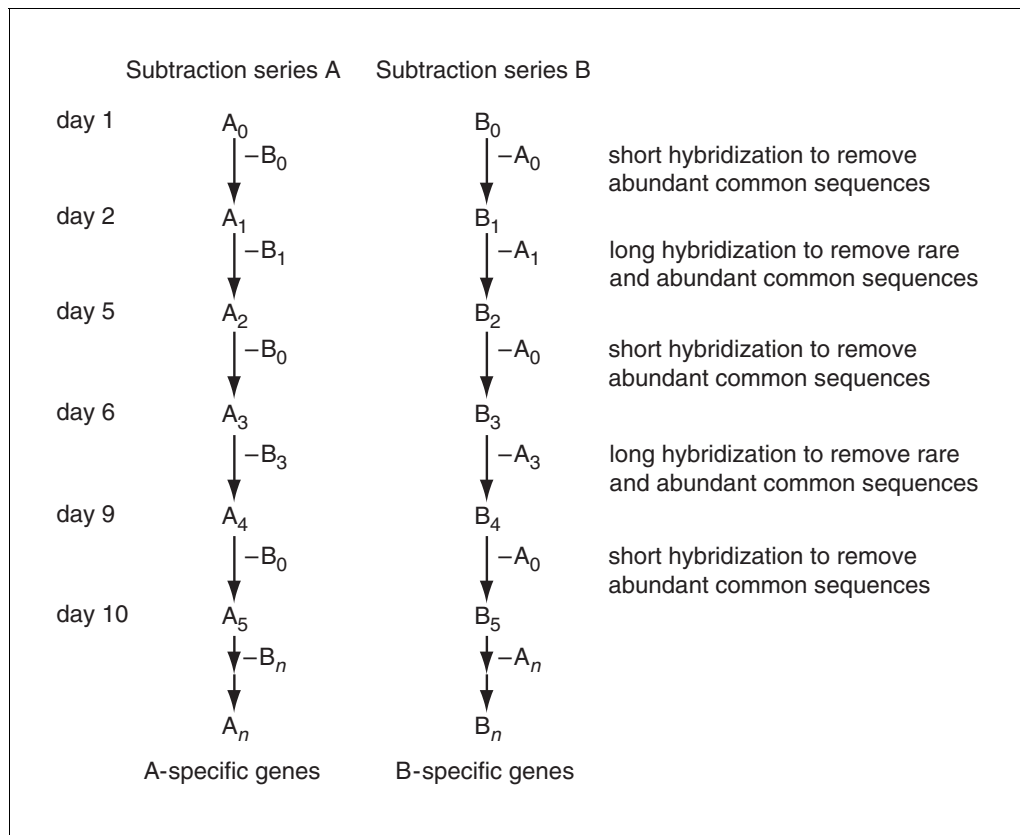


Figure 25B.2.3 Sequence of subtractions. The order of subtractions performed is outlined here for the first five subtractions. The approximate timescale and the hybridization length for each subtraction is indicated along with the primary purpose for each subtraction. Subtractions alternate between a short (2-hr) subtraction with A_0 or B_0 as driver and a long subtraction (30- to 40-hr) with A_n or B_n as driver. A_0 and B_0 are not normalized, that is, they contain an excess of abundant mRNAs or cDNAs and are therefore used to ensure that abundant common sequences are removed. Conversely, $A_1 - A_n$ and $B_1 - B_n$ are enriched for rarer sequences and therefore remove rare common sequences more efficiently than does A_0 or B_0 . The progress of the subtractions is monitored by slot blot hybridization after every three to four subtractions. When the degree of enrichment is satisfactory (>20-fold differential; that is, when A_n hybridizes to itself better than to B_n >20-fold), then the subtracted cDNAs (A_n and B_n) are cloned into appropriate vectors for clonal analysis.

BASIC PROTOCOL

CONSTRUCTION OF SUBTRACTED cDNA LIBRARIES

This protocol describes preparation of libraries of subtracted cDNA clones that represent differentially expressed genes prepared from two cell populations. Each cDNA is ligated to a specific adapter and then the two sets of cDNAs are amplified by PCR to provide large amounts of starting material. Part of the starting material is radiolabeled to provide tracer cDNA to monitor subtraction efficiency; part is biotinylated to provide driver cDNA to facilitate removal of hybrids after annealing. Tracer cDNA from cell population A is hybridized to driver cDNA from population B and vice versa. Tracer/driver and driver/driver hybrids are removed by exposure to streptavidin and phenol extraction, leaving subtracted tracer cDNAs enriched for differentially expressed genes for each population. The sequences are enriched further by repeated rounds of amplification and hybridization. The progress of subtraction is monitored by slot blot hybridization (see Support Protocol). Finally, the subtracted cDNAs are ligated into vectors and used to create libraries that can be screened for individual differentially expressed genes.

PCR-Based Subtractive cDNA Cloning

25B.2.4

Materials

Double-stranded cDNA (ds cDNA) for cell types A and B (*UNIT 5.5*)
AluI and 10× *AluI* buffer (see recipe)
RsaI
10, 15, and 75 mM ATP
10 U/μl T4 polynucleotide kinase and 10× T4 polynucleotide kinase buffer (see recipe)
Oligonucleotide primers
3 μg/μl a1: 5'-TAG TCC GAA TTC AAG CAA GAG CAC A-3'
2.5 μg/μl a2: 5'-CTC TTG CTT GAA TTC GGA CTA-3'
3 μg/μl b1: 5'-ATG CTG GAT ATC TTG GTA CTC TTC A-3'
2.5 μg/μl b2: 5'-GAG TAC CAA GAT ATC CAG CAT-3'
10 U/μl T4 DNA ligase and 10× T4 DNA ligase buffer (see recipe)
40% (w/v) polyethylene glycol 8000 (PEG 8000)
25:24 (v/v) phenol/chloroform (made with buffered phenol; *UNIT 2.1*)
Chloroform
5 U/μl *Taq* DNA polymerase and 10× *Taq* DNA polymerase buffer (see recipe)
25 mM MgCl₂
10 mM 4dNTP mix (*UNIT 3.4*)
Mineral oil, PCR-grade, sterile
800 Ci/mmol [α -³²P]dCTP (10 Ci/μl)
Driver dNTP mix (see recipe)
Ethanol
1 and 5 M NaCl
HEPES buffer (see recipe)
2× hybridization buffer for subtractions (see recipe)
Streptavidin solution (see recipe)
EcoRI and 10× *EcoRI* buffer (see recipe) or *EcoRV* and 10× *EcoRV* buffer (see recipe)
pBluescript vector cut with *EcoRI*
pBluescript vector cut with *EcoRV*
Transformation-competent bacterial strain (*UNIT 1.8*)
Radiolabeled subtraction probes (see Support Protocol)
0.5-ml PCR tubes
Sephacryl S-300 spin columns (Pharmacia Biotech)
Beckman Accuspin FR centrifuge with swinging-bucket rotor or equivalent
Thermal cycler
Anion-exchange PCR spin columns (Qiagen)
1.5-ml microcentrifuge tubes, silanized (*APPENDIX 3B*)
Hand-held Geiger counter
Heating block
Additional reagents and equipment for restriction endonuclease digestion (*UNIT 3.1*), agarose gel electrophoresis (*UNIT 2.5A*), chromatography to remove oligonucleotide fragments (*UNIT 2.6*), phenol/chloroform extraction and ethanol precipitation (*UNIT 2.1A*), anion-exchange (Qiagen) column purification of oligonucleotides (*UNIT 2.1B*), spectrophotometric quantitation of nucleic acids (*APPENDIX 3D*), hybridization of slot blots (*UNIT 2.9B* & *2.10*; also see Support Protocol), bacterial transformation (*UNIT 1.8*), plating libraries (*UNIT 6.1*), preparing replica filters (*UNIT 6.2*), hybridizing replica filters (*UNIT 6.3*), preparing minipreps of plasmid DNA (*UNIT 1.6*), and sequencing plasmid DNA (*UNIT 7.4A* & *7.4B*)

Digest ds cDNA with restriction endonucleases

Double-stranded cDNA (ds cDNA) is digested with frequent-cutting restriction endonucleases into 200- to 600-bp fragments so PCR will not be biased towards smaller fragments.

1. For each set of ds cDNA (A and B) set up two digestions (*AluI* and *AluI* + *RsaI*) as follows:

30 ng ds cDNA
3 μ l 10 \times *AluI* buffer
10 U *AluI* or 10 U *AluI* + 10 U *RsaI*
H₂O to 30.0 μ l.

Incubate overnight at 37°C to ensure complete digestion.

Any other frequent-cutting restriction endonucleases may be used, but enzymes that generate blunt ends are preferable. If an enzyme does not generate blunt-ended DNA fragments, an additional filling-in or chewing-back step is required.

This protocol starts with double-stranded cDNA, full length if possible and primed with oligo(dT), from each cell type being compared (see UNIT 5.5). Commercially available cDNA-synthesis kits from several companies (e.g., Pharmacia Biotech or Life Technologies) work well, even with ≤ 100 ng poly(A)⁺ RNA. Silanized tubes and glycogen are used during ethanol precipitation to avoid loss of cDNA. Sephacryl S-400 columns (Pharmacia Biotech) can be used to purify the synthesized cDNA, which must be cuttable and clean enough for adapter ligation. Between 10 and 100 ng cDNA is a suitable quantity for this digestion.

2. Heat inactivate restriction endonucleases by incubating the reactions ≥ 10 min at 65°C.

Some restriction endonucleases are not susceptible to heat inactivation; phenol/chloroform extraction (UNIT 2.1A) is required to remove them.

Prepare adapters

The adapters are made by annealing kinased oligonucleotide primers a1 or b1 to unphosphorylated primers a2 or b2, respectively.

3. Kinase oligonucleotides a1 and b1 using the following reaction (25 μ l per reaction):

18.0 μ l H₂O
2.5 μ l 10 mM ATP
2.5 μ l 10 \times T4 polynucleotide kinase buffer
1.5 μ l 3 μ g/ μ l oligonucleotide a1 or oligonucleotide b1
0.5 μ l 10 U/ μ l T4 polynucleotide kinase.

Incubate 60 min at 37°C.

It is important that the ligated adapters do not contain or regenerate the restriction endonuclease recognition site in case the enzymes are not totally inactivated (see Critical Parameters).

4. Heat inactivate the kinase by incubating 20 min at 65°C.
5. Add 1.5 μ l of 2.5 μ g/ μ l oligonucleotide a2 or 2.5 μ g/ μ l oligonucleotide b2 to form a1/a2 or b1/b2 adapters. Mix and microcentrifuge briefly at maximum speed. Incubate 10 min at 45°C.

The adapters can be stored at -20°C at this stage.

Ligate adapters to cDNA

Adapters are ligated onto the cDNAs and excess adapters are removed.

6. Set up ligation reactions in 0.5-ml PCR tubes for each set of cDNAs using the appropriate adapter (130 μ l per reaction):

63 μ l H₂O
13 μ l 10 \times T4 DNA ligase buffer
30 μ l 40% PEG 8000
1 μ l 15 mM ATP
10 μ l *AluI*-digested cDNA
10 μ l *AluI/RsaI*-digested cDNA
2 μ l a1/a2 adapter *or* 2 μ l b1/b2 adapter
1 μ l 10 U/ μ l T4 DNA ligase.

Mix and incubate 2 hr at 16°C.

7. Incubate reactions >10 min on ice.
8. Prepare Sephacryl S-300 spin columns according to manufacturer's instructions.
9. Add 1 μ l of 75 mM ATP and 1 μ l T4 polynucleotide kinase to each ligation reaction. Incubate 30 min at 37°C.
10. Extract the ligation reaction with 1 vol of 25:24 phenol/chloroform, then with 1 vol chloroform.
11. Centrifuge the reaction mixture through a prepared Sephacryl S-300 spin column—i.e., 2 min at 400 \times *g* in a Beckman Accuspin FR with a swinging-bucket rotor, room temperature—to remove unligated adapters.

Approximately 130 μ l ligated cDNA will come through the column.

Ligated cDNAs may also be separated from unligated adapters by agarose gel electrophoresis (UNIT 2.5A) followed by electroelution (UNIT 2.6).

Amplify ligated cDNA

Ligated cDNA is amplified by PCR to obtain large amounts of cDNA (A₀, B₀).

12. Set up a PCR mixture for each of the two sets of cDNAs (50 μ l per reaction):

35 μ l H₂O
5 μ l 10 \times *Taq* DNA polymerase buffer
3 μ l 25 mM MgCl₂
1 μ l 10 mM 4dNTP mix
0.5 μ l 2.5 μ g/ μ l oligonucleotide a2 *or* oligonucleotide b2
5 μ l 0.2 ng/ μ l ligated A cDNA *or* B cDNA
0.5 μ l 5 U/ μ l *Taq* DNA polymerase.

Add a few drops of sterile PCR-grade mineral oil to cover the reaction.

13. Amplify the cDNA using the following PCR program:

30 cycles:	1 min	94°C (denaturation)
	1 min	50°C (annealing)
	2 min	72°C (extension)
	25 sec	72°C (autoextension)

If available, use the autoextension function of the thermal cycler (e.g., Perkin-Elmer 480). Alternatively, for thermal cyclers without autoextension, increase the extension time from 2 to 4 min.

This amplification should yield ~10 μ g A₀ and B₀ cDNAs.

The reaction product can be stored overnight at 4°C or longer at -80°C.

14. Analyze 5 to 10 μl of the amplified cDNAs by agarose gel electrophoresis (UNIT 2.5A) to determine the size ranges of amplified cDNAs.

The size of amplified cDNAs should be between 150 bp and 1.5 kb with most ~250 bp.

Prepare labeled tracer and driver DNAs

Radioactive tracer DNA is required for monitoring subtraction efficiency; biotinylated driver DNA is required for removing hybrids by streptavidin binding and phenol extraction.

15. For both sets of amplified cDNAs, set up the following tracer synthesis PCR (100 μl per reaction):

77 μl H₂O
10 μl 10 \times Taq DNA polymerase buffer
6 μl 25 mM MgCl₂
2 μl 10 mM 4dNTP mix
1 μl diluted [³²P]dCTP
1 μl 2.5 $\mu\text{g}/\mu\text{l}$ oligonucleotide a2 or b2
2 μl A₀ or B₀ cDNA (~0.4 μg)
1 μl 5 U/ μl Taq DNA polymerase.

Add a few drops of sterile PCR-grade mineral oil to cover the reaction.

The amount of cDNA used for these initial A₀ and B₀ tracer synthesis reactions is 400 ng; this may be decreased but use ≥ 40 ng for the first amplification. In subsequent amplifications, use 5 to 10 ng A_n or B_n cDNA.

These reactions yield ³²P-labeled tracer cDNA ([³²P]A₀ and [³²P]B₀ in the first round and [³²P]A_n and [³²P]B_n in subsequent rounds; see Fig. 25B.2.2).

16. For both sets of amplified cDNAs, set up three or four driver synthesis PCRs (100 μl per reaction):

73.3 μl H₂O
10 μl 10 \times Taq DNA polymerase buffer
6 μl 25 mM MgCl₂
6.7 μl driver dNTP mix
1 μl 2.5 $\mu\text{g}/\mu\text{l}$ oligonucleotide a2 or b2
2 μl A₀ or B₀ cDNA (1 to 5 ng)
1 μl 5 U/ μl Taq DNA polymerase.

Add a few drops of sterile PCR-grade mineral oil to cover the reaction.

The driver dNTP mix contains 0.5 mM bio-11-dUTP and 1.0 mM dTTP. In the authors' hands this ratio of bio-11-dUTP/dTTP gives the highest overall subtraction efficiency and still allows efficient base pairing.

These reactions yield biotinylated driver cDNA (Bio-A₀ and Bio-B₀ in the first round and Bio-A_n and Bio-B_n in subsequent rounds; see Fig. 25B.2.2).

17. Use the PCR amplification program described in step 13 for tracer and driver synthesis.

18. Purify amplified cDNAs away from unincorporated nucleotides, primer, and salts using a commercial anion-exchange PCR spin column (Qiagen) as directed by the manufacturer (UNIT 2.1B).

An alternate way to purify the PCR products is by agarose gel purification (UNIT 2.5A), but care must be taken to avoid contamination with other DNAs.

19. Determine the yields by spectrophotometric quantitation of nucleic acids (*APPENDIX 3D*).

Typical amplifications yield 12 to 16 μg ^{32}P -labeled cDNA per 100- μl tracer reaction and 7 to 10 μg biotinylated cDNA per 100- μl driver reaction.

The quality and size range of the purified cDNA should be checked using agarose gel electrophoresis (UNIT 2.5A) after every third PCR amplification before proceeding to the next subtraction step. The size range should not change significantly.

Anneal tracer and driver

This is a hybridization between ^{32}P -labeled tracer and biotinylated driver cDNAs.

20. Set up two hybridization reactions ($[^{32}\text{P}]\text{A}_n - \text{Bio-B}_n$ and $[^{32}\text{P}]\text{B}_n - \text{Bio-A}_n$). Ethanol precipitate 1 μg radiolabeled tracer and 20 μg biotinylated driver DNAs in a 1.5-ml silanized microcentrifuge tube without freezing. Air dry the pellet and when just dry, resuspend in 5 μl HEPES buffer by gentle pipetting. Monitor resuspension of the pellet with a hand-held Geiger counter.

A small radioactive pellet should be clearly visible. By not freezing during ethanol precipitation, the possibility of a large salt pellet is avoided.

Resuspension of the pellet sometimes requires a little patience; warming the tube to 60°C usually helps. Also check that none of the counts (i.e., cDNA) are stuck to the pipet tip, as this can greatly reduce the subtraction efficiency. The use of silanized pipet tips may help reduce sticking.

The pellet should not be resuspended in a larger volume because this will lower the concentration of driver, and hence the reannealing rate.

21. Transfer resuspended DNA to a 0.5-ml PCR tube. Add 5 μl of 68°C 2 \times hybridization buffer for subtractions. Mix by gentle pipetting and add a few drops of sterile PCR-grade mineral oil to cover the DNA solution. Microcentrifuge briefly at maximum speed.

If a pellet is visible, the DNA has come out of solution.

22. Incubate the two tubes 10 min at 95°C and cool slowly over 1 hr to 68°C. Continue incubation 2 hr at 68°C (short hybridization).

Either a thermal cycler or a heat block may be used for this step.

Subsequent hybridizations alternate between long (30- to 40-hr) hybridizations during which both rare and abundant common sequences form hybrids, and short (2-hr) hybridizations during which only abundant common sequences form hybrids.

Remove biotinylated annealed and single-stranded DNA

Tracer/driver and driver/driver hybrids and biotinylated single-stranded driver cDNA are removed by addition of streptavidin and extraction with phenol/chloroform.

23. Mix 7 μl of 1 M NaCl with 140 μl HEPES buffer and warm to 68°C. Add to the hybridization reaction to dilute the reaction. Mix and microcentrifuge briefly at maximum speed. Cool to room temperature.
24. Remove 5 μl from each tube and save (total pre-phenol extraction counts).
25. Add 15 μl streptavidin to each tube. Vortex and incubate 5 min at room temperature.
26. Extract each tube with an equal volume 25:24 phenol/chloroform. Retain the aqueous phases and transfer to new tubes.

27. Add 10 μ l streptavidin to each tube containing aqueous phase. Mix and incubate 5 min at room temperature.
28. Extract twice with phenol/chloroform and twice with chloroform. Measure the volume of the aqueous layer for each tube.

The volume for each reaction should be ~150 μ l.

The aqueous phase contains A_n and B_n cDNA.

29. Remove 5 μ l of the aqueous layer from each tube (total post-phenol extraction counts).

Use either scintillation or Cerenkov counts of the pre- and post-phenol extraction samples to determine efficiency of subtraction. The percent tracer cDNA removed is calculated by the following equation:

$$\% \text{ tracer removed} = 100 - (\text{total post-phenol counts} \times 100 / \text{total pre-phenol counts})$$

The subtracted material can be stored at -20°C .

Perform further subtractions

Further rounds of subtraction are performed using subtracted cDNAs from the previous round as template for PCR synthesis of tracer and driver cDNAs. Additional rounds of subtraction, with alternating short and long hybridization steps, continue enriching for the differentially expressed genes.

30. Repeat the subtractions (steps 15 to 29) using A_n or B_n tracer cDNA and the appropriate driver cDNA as determined by the subtraction strategy (see Fig. 25B.2.3). Use A_0 or B_0 drivers for short (2-hr) hybridizations and A_n or B_n drivers for long (30- to 40-hr) hybridizations. Monitor the progress of subtraction by slot blot hybridization (see Support Protocol).

Between five and twenty rounds of subtraction are usually sufficient to isolate cDNAs for differentially expressed genes.

Clone subtracted cDNAs

Subtracted cDNAs are ligated into a vector and cloned to permit screening of individual clones.

31. Amplify 5 μ l of the subtracted cDNAs (A_n and B_n) using the program described in step 13. Purify PCR products with a commercial anion-exchange PCR spin column (e.g., Qiagen; UNIT 2.1B).
32. Digest the cDNAs with the appropriate restriction endonucleases that cut within the adapters (e.g., *EcoRI* and *EcoRV* for the adapters used here).

Taq DNA polymerase may survive phenol/chloroform extraction, so it may help to purify the cDNAs by treating the amplified reaction with proteinase K, extracting with phenol/chloroform, and precipitating with ethanol before digestion.

Digestion may be omitted if blunt-ended ligations are to be performed. PCR amplification often results in the addition of an extra adenosine at the 3' end; this should be removed by Klenow treatment (UNIT 3.16) if blunt-ended ligations are to be performed. Alternatively, the subtracted cDNAs may be cloned into a T-vector (UNIT 15.4).

33. Purify digested cDNAs by phenol/chloroform extraction and ethanol precipitation.
34. Ligate the DNA into an appropriate vector (UNIT 3.16; e.g., pBluescript digested with *EcoRI* or *EcoRV*).

Any convenient vector may be used (see Critical Parameters and Troubleshooting). Using a vector with blue-white selection is useful because it allows immediate assessment of the proportion of the library that contains inserts.

35. Transform the vector into a transformation-competent bacterial strain (UNIT 1.8).

If the subtractions were done to (or nearly to) completion and most of the colonies contain inserts, then it should be possible to pick colonies at random and check for differential expression. Alternatively, use the following steps to assess the quality of the library.

Assess subtracted libraries

Replica filters of the library are probed to assess for percent differentially expressed clones and to provide an indication of the success of the subtractions.

36. Plate out the subtracted library (UNIT 6.1).

It is worth titrating the library first (UNIT 1.3) to obtain individual colonies. It is also important to determine the percentage of colonies that have inserts and the sizes of the inserts (UNIT 5.8). The insert size should be ~250 bp. If the insert size is >500 bp, consider the possibility that the inserts may be double inserts.

37. Prepare four replica lifts from each primary filter (UNIT 6.2).

38. Denature, neutralize, and cross-link the lifts according to the manufacturer's instructions (also see UNIT 6.2).

39. Use subtracted probes (see Support Protocol, step 5) to hybridize the replica filters.

Comparison of filters probed with A_n versus B_n identifies those clones that are probably differentially expressed in the starting A_0 and B_0 cDNAs and also indicates what proportion of the library contains differentially expressed genes. Further rounds of subtraction may be desirable if only a small number of the clones seem to be differentially expressed. The filter probed with a common abundant gene should give very few or no positive signals if the subtractions were done to completion. Finally, probing with a known differentially expressed gene(s) gives another indication of how well the subtractions have worked. If the library evaluation suggests that no further subtractions are needed, analyze individual clones in the library.

Sort through the library

The number of differentially represented clones from the subtracted library is assessed by sequencing and/or gridding.

40. Pick 50 to 100 differentially expressed clones from the library either randomly (if the library assessment indicates most of the clones are differentially expressed) or based on a differential hybridization screen using A_n and B_n as probes. Prepare a miniprep of plasmid DNA (UNIT 1.6).

41. Sequence the inserts in each of the plasmid DNAs (UNIT 7.4A & 7.4B) and group together clones containing the same sequences.

DNA sequence analysis software such as that from DNASTar is helpful.

If most of the clones analyzed initially are the same, they should be subtracted out to reveal rarer transcripts. This is done by pooling the identified clones and using them to make driver that is then used for subtraction with A_n or B_n tracer. Alternatively, the library can be plated out and the lifts probed with mixed probe from the sequenced clones (≤ 20 sequences/mixed probe). Clones that do not hybridize have not yet been sequenced and should be analyzed. If all the clones seem to be differentially expressed but a few are particularly prevalent, then another way to reveal rare transcripts is to normalize A_n and B_n (or self-subtract them—i.e., $A_n - A_n$ and $B_n - B_n$) for a short period of time. These procedures greatly reduce the work involved in sorting through the library.

42. Determine whether the clones are truly differentially expressed in the starting tissues by RNA expression analysis—e.g., northern blot hybridization (UNIT 4.9), RNase protection assay (UNIT 4.7), quantitative RT-PCR (UNIT 15.5), or in situ hybridization (UNITS 14.3 & 14.7).

**SUPPORT
PROTOCOL**

SLOT BLOT HYBRIDIZATION TO MONITOR SUBTRACTION

After every three to four subtractions, the progress of enrichment for differentially expressed genes is monitored by slot blot hybridization (also see UNITS 2.9B & 2.10).

Additional Materials (also see Basic Protocol)

cDNA from each subtraction (see Basic Protocol, step 28)
3 M NaOH
2 M ammonium acetate, pH 7.0
Probe dNTP mix (see recipe)
Sephadex G50/80 spin column (Pharmacia Biotech) in sterile 1-ml syringe
Additional reagents and equipment for slot blotting (UNIT 2.9B) and hybridization (UNIT 2.10)

1. Denature 1200 ng cDNA from each subtraction ($A_n - B_n$ and $B_n - A_n$) by adding 0.1 vol of 3 M NaOH to cDNA and heating 30 to 60 min at 65°C.
2. Neutralize the DNA by adding 1 vol of 2 M ammonium acetate, pH 7.0.
3. Spot duplicate 100-ng aliquots of denatured and neutralized cDNA from each subtraction onto each of six or more slot blots (UNIT 2.9B).
4. Use cDNA from A_n , B_n , A_0 , B_0 , a gene expressed at high levels in both A and B, and one or more genes expressed differentially by A or B (or a gene used to spike the reaction) to prepare radiolabeled subtraction probes. Prepare a PCR mixture for each probe (50 μ l per reaction):

17.5 μ l H₂O
5 μ l 10 \times *Taq* DNA polymerase buffer
3 μ l 25 mM MgCl₂
2 μ l probe dNTP mix
20 μ l [α -³²P]dCTP
1 μ l 2.5 μ g/ μ l primer a2, primer b2, primer specific for gene expressed in both A and B, *or* primer specific for gene expressed differentially in A or B
0.5 μ l 4 ng/ μ l subtracted A_n or B_n cDNA or appropriate gene template DNA
1 μ l 5 U/ μ l *Taq* DNA polymerase.

Add a few drops of sterile PCR-grade mineral oil to cover the reaction.

5. Amplify and label the probe using the following PCR program:

30 cycles	1 min	94°C (denaturation)
	1 min	50°C (annealing)
	2 min	72°C (extension)

This reaction yields a double-stranded probe; the probes should be denatured before hybridization.

6. Purify the probe by centrifuging it through a 1-ml Sephadex G50/80 spin column, 2 min at $170 \times g$, in a Beckman Accuspin FR with a swinging-bucket rotor, room temperature.
Expect ~50 μ l eluate after centrifugation.
7. Measure incorporation by counting a 1- μ l fraction of the eluate in a scintillation counter.
Routinely, incorporation is $\sim 10^6$ cpm/ μ l eluate.
8. Hybridize each slot blot with one of the above probes (UNIT 2.10).
9. Wash the blots to high stringency (UNIT 2.10).
10. Expose filters to X-ray film or a phosphoimaging plate (APPENDIX 3A).

The A_n and B_n hybridizations are the most important because they reveal the degree to which A_n and B_n cDNAs still cross-hybridize with B_n and A_n cDNAs, respectively. In general, further subtractions are desired if the differential is <20 -fold (that is, A_n hybridizes <20 -fold better to itself than to B_n and vice versa). Probing the subtracted cDNAs with a highly expressed gene or with a differentially expressed gene gives another indication of how well the subtractions are advancing. Common abundant genes should become less abundant with increasing rounds of subtraction, and the known differentially expressed gene should become enriched in one series of cDNAs and depleted in the other. The A_0 and B_0 probes usually represent the common abundant genes and therefore behave accordingly; that is, they hybridize more strongly to cDNA from earlier rounds of subtraction and less so to later rounds. When the evaluation suggests that no more subtractions are required, then the A_n and B_n cDNAs should be cloned into an appropriate vector.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

AluI buffer, 10 \times

- 100 mM Bis-Tris propane (1,3-bis[tris(hydroxymethyl)methylamino]propane)·Cl, pH 7.0
- 100 mM MgCl₂
- 10 mM dithiothreitol (DTT; APPENDIX 2)
- Store up to 6 months at -20°C

Driver dNTP mix

- 1.5 mM each dATP, dCTP, and dGTP
- 1.0 mM dTTP
- 0.5 mM bio-11-dUTP (Enzo Diagnostics)
- Store up to 3 month at -20°C

EcoRI buffer, 10 \times

- 1 M Tris·Cl, pH 7.5 (APPENDIX 2)
- 500 mM NaCl (APPENDIX 2)
- 100 mM MgCl₂ (APPENDIX 2)
- 0.25% (v/v) Triton X-100
- Store at -20°C

EcoRV buffer, 10 \times

- 100 mM Tris·Cl, pH 7.9 (APPENDIX 2)
- 500 mM NaCl (APPENDIX 2)
- 100 mM MgCl₂ (APPENDIX 2)
- 10 mM dithiothreitol (DTT; APPENDIX 2)
- Store up to 6 months at -20°C

HEPES buffer

100 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*-2-ethanesulfonic acid), pH 7.3
1 mM EDTA (*APPENDIX 2*)
Store at -20°C

Hybridization buffer for subtractions, 2 \times

50 mM HEPES, pH 7.3
10 mM EDTA (*APPENDIX 2*)
0.2% (w/v) SDS
1.5 M NaCl (*APPENDIX 2*)
Store up to 3 months at -20°C
To avoid cloudiness, add NaCl last and warm to 68 $^{\circ}\text{C}$

Probe dNTP mix

0.5 mM each dATP, dGTP, and dTTP
0.1 mM dCTP
Store up to 3 months at -20°C

Streptavidin solution

2 $\mu\text{g}/\mu\text{l}$ streptavidin
0.15 M NaCl (*APPENDIX 2*)
HEPES buffer (see recipe)
Store up to 6 months at -20°C

T4 DNA ligase buffer, 10 \times

500 mM Tris \cdot Cl, pH 7.8 (*APPENDIX 2*)
100 mM MgCl_2 (*APPENDIX 2*)
100 mM dithiothreitol (DTT; *APPENDIX 2*)
10 mM ATP
250 $\mu\text{g}/\text{ml}$ BSA
Store up to 6 months at -20°C

T4 polynucleotide kinase buffer, 10 \times

700 mM Tris \cdot Cl, pH 7.6 (*APPENDIX 2*)
100 mM MgCl_2 (*APPENDIX 2*)
50 mM dithiothreitol (DTT; *APPENDIX 2*)
Store up to 6 months at -20°C

Taq DNA polymerase buffer, 10 \times

100 mM Tris \cdot Cl, pH 9.0 (*APPENDIX 2*)
500 mM KCl (*APPENDIX 2*)
1% (v/v) Triton X-1000
Store at -20°C

COMMENTARY**Background Information**

Early subtractive cloning involved one or two rounds of hybridization using cDNA as tracer and mRNA as driver. cDNA/mRNA hybrids were removed by binding to hydroxylapatite columns maintained at 68°C . This scheme has two major limitations that prevented subtractive cloning from becoming a routine and frequently used technique. The first was that hydroxylapatite columns are cumbersome, making it difficult to separate

single-stranded sequences from the hybrids. This problem has been largely overcome through the use of biotinylated driver sequences in combination with streptavidin treatment and phenol extractions (Sive and St. John, 1988; Sive et al., 1989), or streptavidin-conjugated magnetic beads (Uhlen, 1989; Straus and Ausubel, 1990).

A second problem with the original technique was the rapid decrease in the amount of cDNA present, making it very difficult to per-

form multiple rounds of subtraction or to clone the minute amounts of cDNA left after subtraction. Several different approaches have been used to tackle this second problem. One solution has been to construct directional phagemid libraries that can be converted into a single-stranded library; after the subtractions are performed, the remaining single-stranded plasmids are transformed into bacteria for amplification (Duguid et al., 1988; Rubenstein et al., 1990). The subtracted library can be used in further rounds of subtractions; however, the method is laborious and care must be taken to avoid contamination with the double-stranded (ds) forms of the phagemid. Contaminating ds phagemid DNA will not be subtracted away and will transform bacteria much more efficiently than single-stranded DNAs do, thus reducing the overall subtraction effect. Other methods have used cDNA attached to oligo(dT)-Latex in combination with the polymerase chain reaction (PCR). This allows the driver to be reused (Hara et al., 1993).

An alternative solution described in this protocol regenerates the cDNAs by PCR (Duguid and Dinauer, 1989; Wang and Brown, 1991). A problem with PCR is that it amplifies smaller fragments better than larger fragments and therefore selects for smaller mRNAs. Wang and Brown (1991) overcame this difficulty by cutting the original cDNAs to smaller sizes before PCR. This approach allows multiple rounds of subtractions and has allowed isolation of many genes that are differentially expressed in metamorphosis-stage *Xenopus* embryos after thyroid hormone treatment (Buckbinder and Brown, 1992). With the modified protocol detailed here, the authors have isolated many genes that delineate the early events of neural induction and anteroposterior patterning in *Xenopus* (Patel et al., unpub. observ.). The method described here is very sensitive and can isolate genes that are as little as 2- to 3-fold differentially expressed.

In the scheme described here, two cDNA populations are cross-subtracted—that is, A tracer is subtracted with B driver and B tracer with A driver. This allows isolation of genes expressed preferentially in A and genes expressed preferentially in B. Cross-subtraction has two other effects. The first is to increase the concentration of rare sequences relative to the concentration of abundant common sequences in the driver, because the latter rapidly hybridize (at low C_0t) and are removed by subtraction. This is termed normalization, as it normalizes or equalizes the concentrations of what were

initially rare and abundant common cDNAs. In practice, it is not possible to reach a truly equalized representation of sequences, but the starting concentrations of different cDNAs can vary 10,000-fold, and after normalization this can be reduced to ~10-fold (Patanjali et al., 1991; Soares et al., 1994). Normalizing the driver makes it much more efficient at removing rare common sequences than an unnormalized driver. Normalizing the driver is essential when starting with tissues that have high mRNA complexity. It is, of course, also important that some of the subtractions be performed with a driver that still contains high levels of abundant common sequences (that is, the starting cDNA population, A_0 or B_0); otherwise these abundant sequences will never be removed.

Normalization could also be achieved by subtracting the driver against itself (self-subtraction). The reason cross-subtractions are used instead is that they provide a second benefit. One of the problems with any efficient subtraction scheme is that it may remove sequences expressed only a few-fold higher in one cell population than the other, and therefore allow isolation of only those sequences that are not expressed at all in the driver. Sequences expressed with ≤ 10 -fold differential may be of great interest and can be isolated by cross-subtractions. Suppose that sequence G is present at a ratio of 1:5 in A_0/B_0 , the starting cDNAs. If B_0 is subtracted with A_0 , and vice versa, G will be removed somewhat from the resulting B_1 ; however, after the reciprocal $A_0 - B_0$ subtraction, relatively more G will be removed from the resulting A_1 than it was from B_1 because the driver (B_0) had a higher concentration of G than A_0 did. Thus, the ratio of G in A_1/B_1 will decrease, perhaps to 1:10. This enhanced relative difference in the level of G between A_1 and B_1 will be enhanced even more in subsequent cross-subtractions, to ultimately allow isolation of G as a differentially expressed clone. One problem here is that cross-subtracting can result in false positives (genes that are differentially represented in the final A_n and B_n cDNA populations, but not in the starting A_0 and B_0 cDNAs). This is a particular problem if the efficiencies of the two subtraction series ($A - B$ versus $B - A$) are different, but it can easily be checked after subtraction by asking whether a clone is differentially represented in the A_0 and B_0 starting cDNAs.

This protocol includes two modifications to the Wang and Brown (1991) method that the authors feel improve it. First, bio-11-dUTP is

incorporated into the driver as a means of biotinylating (Patel and Sive, unpub. observ.) in place of the photobiotinylation originally described (Sive and St. John, 1988) for two reasons. Incorporation of biotin during PCR amplification is extremely simple and does not require additional photobiotinylation steps. Substituting 30% of the dTTP with bio-11-dUTP in the amplification of driver nucleic acid gives maximal subtraction efficiency. With lower substitution, subtraction efficiency decreases, presumably because the density of biotin is not great enough; with greater substitution, subtraction efficiency also decreases, presumably because the biotin interferes with base-pairing (Patel and Sive, unpub. observ.). Photobiotinylated nucleic acid is rather insoluble in aqueous solutions due to a long hydrocarbon linker arm; photobiotinylated driver sometimes precipitates out of the hybridization mix. Nucleic acids with biotinylated nucleotides incorporated during PCR seem as soluble as unmodified nucleic acids and precipitation in the hybridization mix does not occur, at least in the authors' hands. Another method for incorporating biotinylated nucleotides is to use biotinylated primers for PCR (Rosenberg et al., 1994). Second, this protocol uses different adapters on the driver and tracer cDNAs. The original protocol (Wang and Brown, 1991) used the same adapters for both tracer and driver to ensure that all sequences in tracer and driver amplified to the same extent; however, this also

meant an increased risk of driver carry-over into the next round of subtraction; such carried-over driver would be amplified along with the subtracted cDNA and would decrease subtraction efficiency. Using the different primers given here, the authors have observed essentially equivalent PCR efficiency for the two cDNA pools.

Several other methods have been used to isolate genes that are differentially expressed between two or more cell populations (see Table 25B.2.1)—random sampling (in which clones are randomly selected from a cDNA library), differential hybridization (in which probes made from the mRNAs of the two tissues being compared are used to screen a cDNA library, and clones that hybridize to one probe but not to the other are isolated), and differential display (*UNIT 25B.3*; in which partially random primers are used to amplify a subset of mRNAs expressed in a given cell type; these are then separated on an acrylamide gel and the bands between different samples compared). Of all the procedures, subtractive cloning is probably the most sensitive, and it is the method of choice to isolate as complete a set of differentially expressed genes as possible. The other methods allow isolation of a small number of differentially expressed genes and may be sufficient to obtain useful markers. Random sampling of a cDNA library is useful only if the two tissues to be compared contain a widely different spectrum of mRNAs.

Table 25B.2.1 A Comparison of Differential Screening Methods

Method	Advantages	Disadvantages
Subtractive cloning	Targets rare mRNAs (<0.001%) Targets complete set of differentially expressed RNAs Requires little starting material Nonisogenic tissues can be compared	Can generally only compare two tissues at one time Procedure can be long
Differential display	Requires little starting material Can compare more than two tissues or treatments at one time Procedure is relatively short	Targets only a subset of the differentially expressed genes Generally targets medium-abundant mRNAs Can yield many false positives
Differential hybridization	Procedure is relatively easy	Targets relatively abundant mRNAs (~0.1%)
Random sampling	Simple procedure; only cDNA libraries are required	Only useful for comparing very different tissues

Critical Parameters and Troubleshooting

Some of the more common problems that arise with this procedure and their solutions are listed in Table 25B.2.2.

RNA preparation

It is essential to start with a clean preparation of RNA that is free of any salts or other substances that may inhibit reverse transcription. The RNA should not be contaminated with

even trace amounts of genomic DNA. In fact, all RNA preparations should be treated with RNase-free DNase, then checked for contaminating genomic DNA by PCR using primers for specific genes. Contaminating DNA alters the representation of the various mRNAs during the subtractions to create false positives.

Purity of oligonucleotides

The quality of the primers is crucial to the success of the procedure. It is worth purifying

Table 25B.2.2 Troubleshooting Guide for Subtractive cDNA Cloning

Problem	Probable cause(s)	Remedy
No amplified cDNAs visible on agarose gel	Failure of adapters to ligate to cDNA due to: Kinasing of wrong primer Inactive ligation buffers and/or enzymes Inhibitors in cDNA	Kinase correct primer Test and replace as necessary Repurify cDNA by phenol/chloroform extraction and ethanol precipitation using glycogen as carrier
	Failure of PCR amplification due to inactive amplification buffer and/or enzymes	Test and replace as necessary
Median size range of amplified cDNAs >500 bp in A ₀ and/or B ₀	Incomplete digestion of cDNA before amplification due to: Inhibitors in ds cDNA	Purify ds cDNA by phenol/chloroform extraction and ethanol precipitation using glycogen as carrier
	Inactive restriction buffers or enzyme	Test and replace as necessary
Low subtraction efficiency	Loss of DNA during ethanol precipitation or resuspension Incomplete resuspension of DNA before hybridization	Repeat with careful monitoring using hand-held Geiger counter Avoid complete drying of DNA pellet before resuspension Warm sample to 60°C to aid resuspension
No or few colonies after cloning into vector	Incomplete digestion of DNA	Repurify DNA; treat with proteinase K, phenol/chloroform extract, ethanol precipitate, and wash with 70% ethanol
	Inactive enzymes or buffers Poor ligation efficiency	Test and replace as necessary Repurify DNA; test and replace enzymes or buffers
	Low transformation efficiency	Test and replace competent cells
No or few differentially expressed genes	Contamination of RNA or cDNA with (genomic) DNA before ligation of adaptors	Restart with fresh RNA and treat it with DNase before reverse transcription

the primers to ensure they are full length and free of any salts or other inhibitors. The synthesized primers can be gel purified, although the authors prefer to use Nensorb Prep columns (Du Pont NEN). To use these columns, the 5'-trityl group on the primer must not be removed (see UNIT 2.11 for more information about the synthesis of oligonucleotides).

Primer design

The two different adapters used during this protocol are created by annealing a 21-mer and a 25-mer oligonucleotide. The sequences of four primers that these authors have used successfully are listed in the Basic Protocol. These primers contain sites for *EcoRI* or *EcoRV*; however, different restriction endonuclease sites or other special features for particular vectors may be desirable, so this section outlines some considerations in primer design.

First, restriction endonucleases generally require at least four bases next to their recognition sequences to work efficiently. Second, primers should contain minimal secondary structure to maximize annealing to the target sequence. Third, there should be no similarity between the primers that make up one set of adapters and those that make up the second set. This is extremely important for the success of the subtractions, and it is essential to check for any cross-annealing by testing whether a primer from one set of adapters can amplify cDNA (or a test DNA fragment) ligated to the other adapter. Fourth, in order to perform the PCR amplifications for the two sets of cDNAs (A and B) at the same time, it is important that the primers have similar melting temperatures (T_m). Fifth, the T_m should not be so high that it approaches the hybridization temperature of 68°C, so the GC content should be kept <50%, and the primers should not be excessively long (>50 bases). Standard oligonucleotide software (e.g., Oligo, Primerselect) is helpful for designing primers. Primer sequences should also be checked against a database such as GenBank for any similarities to sequences in known genes.

Restriction digestion of cDNAs

It is very important that the cDNAs of both the tracer and driver sides be cut to completion before adapter ligation. If, for example, A cDNA has not digested as well as B cDNA, PCR may be biased for smaller fragments in A but not B, resulting in false positives at the end of the procedure. The A_0 and B_0 populations

should be checked on a gel to ensure that the size ranges of amplified cDNAs are the same.

Monitoring subtractions

It is necessary to monitor efficiency of the subtractions to determine when to stop subtracting. In many cases a problem can be easily resolved on the spot rather than being discovered at the end of the subtractions so that it is necessary to start all over again. Subtraction efficiency can be monitored in the following ways. First, the cumulative percentage removal of tracer counts after the phenol extractions at each subtraction should be determined to provide an immediate and fairly accurate way of determining whether a particular subtraction has been successful and whether the subtraction should be repeated. Second, the degree to which A_n and B_n cross hybridize can be monitored by slot blotting cDNAs from each step of subtraction and probing the blots with the last set of subtracted cDNAs (A_n and B_n). Subtractions are generally stopped when a probe made from A_n hybridizes to the A_n cDNA pool ~20-fold better than it does to the B_n pool, and vice versa. Third, the removal or enrichment of a known differentially expressed gene in A_0 through A_n can be monitored by slot blot hybridization. If no such gene is available, then the original tracer may be spiked with some DNA such as β -galactosidase, which can be removed at the end. Fourth, if the subtractions are working, common abundant sequences should be progressively removed with each subtraction.

Anticipated Results

The end result of the procedure is the isolation of fragments of differentially expressed genes. The actual number of such genes obtained depends on the tissues being compared. Hence, if the two starting tissues are of very similar complexity, only a few genes may be obtained. On the other hand, if the tissues being compared contain a mixture of cell types and are very different, it is easily possible to obtain hundreds of differentially expressed genes. Abundant transcripts will be represented more frequently than rare transcripts. Additionally, each original transcript may be represented by multiple clones because the original cDNA was digested into fragments before subtraction. In theory, because the restriction endonucleases (*AluI* and *RsaI*) have 4-bp recognition sequences, digestion should produce approximately four 250-bp fragments per kilobase

original mRNA. In practice, digestion yields one to two fragments per gene. If the subtractions have been performed exhaustively, then theoretically every clone in the subtracted library should be differentially expressed. Furthermore, it should be possible to isolate genes that are 2- to 3-fold differentially expressed between two given tissues and whose abundance is as little as 5 copies mRNA/cell; however, the isolation of rare differentially expressed genes is dependent on the complexity of the starting tissues.

Time Considerations

A time schedule for this procedure is presented in Table 25B.2.3. This schedule is approximate and assumes that the procedure starts with double-stranded cDNA (see Fig. 25B.2.3).

Literature Cited

- Buckbinder, L. and Brown, D.D. 1992. Thyroid hormone-induced gene expression changes in the developing frog limb. *J. Biol. Chem.* 267:25786-25791.
- Davidson, E.H. 1986. Complexity of maternal RNA. *In Gene Activity in Early Development*, 3rd ed., pp. 50-55. Academic Press, San Diego.
- Duguid, J.R., Rohwer, R.G., and Seed, B. 1988. Isolation of cDNAs of scrapie-modulated RNAs by subtractive hybridization of a cDNA library. *Proc. Natl. Acad. Sci. U.S.A.* 85:5738-5742.
- Duguid, J.R. and Dinauer, M.C. 1989. Library subtraction of in vitro cDNA libraries to identify differentially expressed genes in scrapie infection. *Nucl. Acids Res.* 18:2789-2792.
- Hara, E., Yamaguchi, T., Tahara, H., Tsuyama, N., Tsurui, H., Ide, T., and Oda, K. 1993. DNA-DNA subtractive cDNA cloning using oligo dT-Latex and PCR: Identification of cellular genes which

Table 25B.2.3 Time Requirements for Preparation of Subtracted cDNA

Day	Procedure	Time required
1	<i>Restriction endonuclease digestion</i>	Overnight (0.5 hr to set up)
2	<i>cDNA preparation</i>	
	Adapter preparation	1.5 hr
	Adapter ligation	3.5 hr
	Amplification of ligated cDNA and checking by gel electrophoresis	5 hr
	Tracer and driver synthesis	Overnight (0.5 hr to set up)
3	<i>First (short) subtraction</i>	
	Tracer and driver purification and quantitation	1 hr
	Tracer and driver annealing	2 hr (short)
	Removal of annealed and ssDNA	3 hr
	Tracer and driver synthesis	Overnight (0.5 hr to set up)
4 to 6	<i>Second (long) subtraction</i>	
	Tracer and driver purification and quantitation	1 hr
	Tracer and driver annealing	40 hr (long)
	Removal of annealed and ssDNA	3 hr
	Tracer and driver synthesis	Overnight (0.5 hr to set up)
7 to 30	<i>Further subtractions</i>	Variable ^a
	Alternating short and long hybridizations	6 hr to 44 hr
	Slot blot hybridization to check the progress of subtraction	24 hr
31 to 34	<i>Cloning of subtracted cDNAs</i>	
	Amplification of subtracted cDNAs	8 hr
	Restriction endonuclease digestion	Overnight (0.5 hr to set up)
	Vector ligation	Overnight (0.5 hr to set up)
	Bacterial transformation and growth	Overnight (0.5 hr to set up)
35 to 38	<i>Assessment of subtracted cDNA library</i>	
	Growth of library	Overnight (0.5 hr to set up)
	Preparation of lifts	8 hr
	Hybridization with subtracted probes	24 hr

^aThe schedule for days 7 to 30 depends on the duration of the hybridization steps and the amount of progress with each subtraction.

- are overexpressed in senescent human diploid fibroblasts. *Anal. Biochem.* 214:58-64.
- Patanjali, S.R., Parimoo, S., and Weissman, S.M. 1991. Construction of a uniform abundance (normalized) cDNA library. *Proc. Natl. Acad. Sci. U.S.A.* 88:1943-1947.
- Rosenberg, M., Przybylska, M., and Straus, D. 1994. RFLP subtraction: A method for making libraries of polymorphic markers. *Proc. Natl. Acad. Sci. U.S.A.* 91:6113-6117.
- Rubenstein, J.L.R., Brice, A.E.J., Ciaranello, R.D., Denney, D., Porteus, M.H., and Usdin, T.B. 1990. Subtractive hybridization system using single-stranded phagemids with directional inserts. *Nucl. Acids Res.* 18:4833-4842.
- Sive, H.L. and St. John, T. 1988. A simple subtractive hybridization technique employing photoactivatable biotin and phenol extraction. *Nucl. Acids Res.* 16:10937.
- Sive, H.L., Hattori, K., and Weintraub, H. 1989. Progressive determination during formation of anteroposterior axis in *Xenopus laevis*. *Cell* 58:171-180.
- Soares, M.B., Bonaldo, M.F., Jelene, P., Su, L., Lawton, L., and Efstratiadis, A. 1994. Construction and characterization of a normalized cDNA library. *Proc. Natl. Acad. Sci. U.S.A.* 91:9228-9232.
- Straus, D. and Ausubel, F.M. 1990. Genome subtraction for cloning DNA corresponding to deletion mutants. *Proc. Natl. Acad. Sci. U.S.A.* 87:1889-1893.
- Uhlen, M. 1989. Magnetic separation of DNA. *Nature* 340:733-734.
- Wang, Z. and Brown, D.D. 1991. A gene expression screen. *Proc. Natl. Acad. Sci. U.S.A.* 88:11505-11509.
-
- Contributed by Mukesh Patel and Hazel Sive
Whitehead Institute for Biomedical Research
Cambridge, Massachusetts

Differential Display of mRNA by PCR

UNIT 25B.3

**BASIC
PROTOCOL**

This unit describes differential display to identify mRNA species for differentially expressed genes. DNA sequences corresponding to these mRNAs can be recovered, cloned, sequenced, and used for hybridization or library screening probes. This approach combines both the power of polymerase chain reaction (PCR) amplification and the high resolution of denaturing polyacrylamide gel electrophoresis for separation of amplified cDNA products. The basic principle is to reverse transcribe and systematically amplify the 3' termini of mRNAs with a set of anchored oligo(dT) primers and an arbitrary decamer. Figure 25B.3.1 illustrates the general strategy of differential display. Specifically, an RNA sample is reverse transcribed with each of the four sets of degenerate anchored oligo(dT) primers ($T_{12}MN$), where M can be G, A, or C and N is G, A, T, and C. Each primer set is dictated by the 3' base (N), with degeneracy in the penultimate (M) position. For example, the primer set where N = G consists of:

5'-TTTTTTTTTTTTGG-3'
5'-TTTTTTTTTTTTAG-3'
5'-TTTTTTTTTTTTCG-3'

The resulting cDNA population is PCR-amplified using the degenerate primer set, an arbitrary decamer, and radioactive nucleotide. The radioactively labeled PCR products that represent a subpopulation of mRNAs defined by the given primer set are separated on a denaturing polyacrylamide gel. By changing primer combinations, most of the RNA species in a cell may be represented. Side-by-side comparison of RNA samples from different cells allows the identification and cloning of differentially expressed genes.

Materials

Total cellular human RNA (UNIT 4.2) or poly(A)⁺ RNA (UNIT 4.5)
1 U/μl human placental RNase inhibitor
10 U/μl DNase I (RNase-free)
0.1 M Tris·Cl, pH 8.3 (APPENDIX 2)
0.5 M KCl
15 mM MgCl₂
3:1 (v/v) phenol/chloroform
3 M sodium acetate, pH 5.2 (APPENDIX 2)
100%, 70%, and 85% ethanol
Diethylpyrocarbonate (DEPC)-treated H₂O (UNIT 4.1)
10 μM each degenerate anchored oligo(dT) primer set 5'-T₁₂MN-3' (e.g., GenHunter): T₁₂MG, T₁₂MA, T₁₂MT, and T₁₂MC (M represents G, A, or C)
5× MoMuLV reverse transcriptase buffer (UNIT 15.6)
0.1 M dithiothreitol (DTT; APPENDIX 2)
250 μM and 25 μM 4dNTP mixes (UNIT 3.4)
200 U/μl Moloney murine leukemia virus (MoMuLV) reverse transcriptase
10× PCR amplification buffer (make as in UNIT 15.1, with 15 mM MgCl₂, but use only 0.1 mg/ml gelatin; store at -20°C)
10 μCi/μl [α -³²P]dATP (>2000 Ci/mmol)
2 μM arbitrary decamer (see Critical Parameters; e.g., GenHunter or Operon Technologies)
5 U/μl Taq DNA polymerase
Mineral oil
Formamide loading buffer (see recipe)
10 mg/ml glycogen (DNA-free)

**Discovery of
Differentially
Expressed Genes**

25B.3.1

Contributed by Peng Liang and Arthur B. Pardee

Current Protocols in Molecular Biology (2001) 25B.3.1-25B.3.10

Copyright © 2001 by John Wiley & Sons, Inc.

Supplement 56

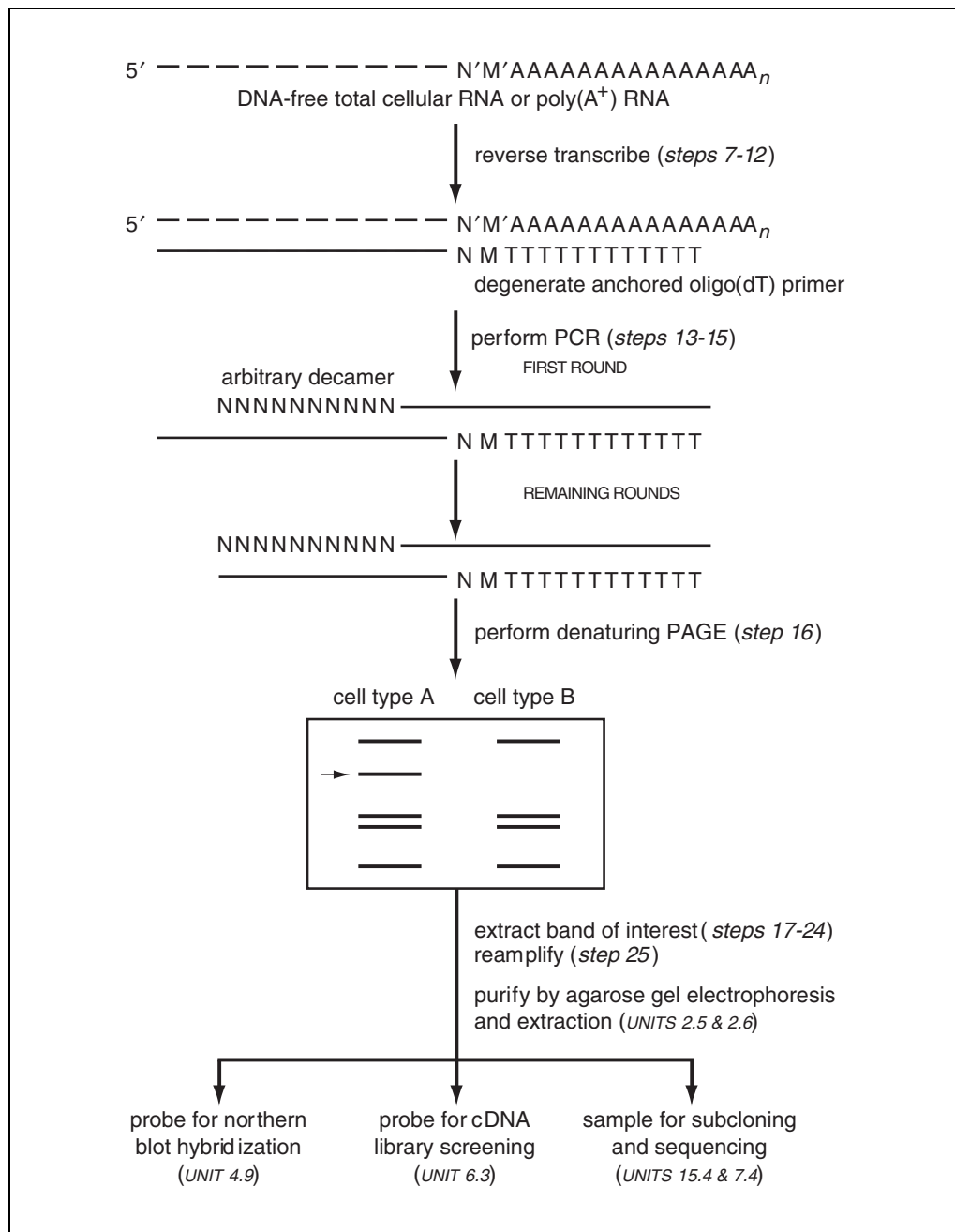


Figure 25B.3.1 Schematic representation of differential display. Diagram of gel represents results with a single primer set for two cell types, A and B. Dashed line, RNA; solid line, DNA; T₁₂MN, degenerate oligo(dT) primer; M indicates A, C, or G (degenerate); N can be A, C, G or T.

65°, 95°, 80°, and 100°C water baths

Thermal cycler

Whatman 3MM filter paper

Additional reagents and equipment for preparing total (UNIT 4.2) or poly(A)⁺ (UNIT 4.5) RNA, quantitating RNA (APPENDIX 3D), PCR (UNIT 15.1), agarose-formaldehyde gel electrophoresis (UNIT 4.9), denaturing PAGE (UNIT 7.6), autoradiography (APPENDIX 3A), agarose gel electrophoresis (UNIT 2.5A), purifying DNA from agarose gels (UNIT 2.6), analysis of RNA by northern blot analysis (UNIT 4.9), screening libraries using oligonucleotide probes (UNIT 6.3), cloning PCR products (UNIT 15.4), and dideoxy DNA sequencing (UNIT 7.4)

CAUTION: This procedure should be performed only by personnel trained in the proper use of ^{33}P isotope and in NRC licensed sites. Standard precautions to prevent excessive exposure and radioactive contamination of personnel and equipment should be followed at all times.

NOTE: Experiments involving RNA require careful technique to prevent RNA degradation (UNIT 4.1).

Remove chromosomal DNA contamination from RNA

1. Digest DNA from total cellular RNA or poly(A)⁺ RNA by mixing:

- 50 μg RNA
- 10 μl 1 U/ μl human placental RNase inhibitor
- 1 μl 10 U/ μl RNase-free DNase I
- 5 μl 0.1 M Tris·Cl, pH 8.3
- 5 μl 0.5 M KCl
- 5 μl 15 mM MgCl_2
- H_2O to 50 μl .

Incubate 30 min at 37°C.

When performing differential display, it is essential that the RNA sample be free from any genomic DNA contamination. RNA preparations isolated by various methods are often found to be contaminated with trace amounts of chromosomal DNA that results in reverse transcription-independent DNA amplification. Amounts from 15 to 100 μg of total RNA can be cleaned with this procedure.

2. Add 50 μl phenol/chloroform (3:1), vortex, and microcentrifuge 2 min at maximum speed to separate phases.

This step serves to inactivate DNase I before cDNA synthesis during reverse transcription, so vigorous mixing is important to allow complete extraction of DNase I.

3. Transfer upper phase to a clean microcentrifuge tube and add 5 μl of 3 M sodium acetate and 200 μl of 100% ethanol. Incubate 30 min at -70°C to precipitate RNA.
4. Microcentrifuge 10 min at high speed. Remove supernatant and wash pellet (precipitated RNA) once with 500 μl of 70% ethanol.
5. Dissolve RNA pellet in 20 μl DEPC-treated water and quantitate the RNA concentration accurately by measuring the A_{260} with a spectrophotometer (APPENDIX 3D).

DNA-free RNA should be stored at a concentration $>1 \mu\text{g}/\mu\text{l}$. It should not be diluted to the working concentration until immediately before reverse transcription. Diluted RNA should not be reused for differential display as diluted RNA is very unstable during storage and repeated freezing and thawing.

6. Check the integrity of the RNA to be used for differential display by performing agarose/formaldehyde gel electrophoresis (UNIT 4.9) on 3 μg of cleaned RNA. Store DNA-free RNA at -80°C until used for differential display.

For undegraded total RNA, the 28S and 18S ribosomal RNAs should be clearly visible by ethidium bromide staining.

Reverse transcribe RNA

7. For each RNA sample, label four microcentrifuge tubes G, A, T, and C—one tube for each degenerate anchored oligo(dT) primer set.
8. Dilute 1 μg DNA-free RNA (step 5) to 0.1 $\mu\text{g}/\mu\text{l}$ in DEPC-treated water and place on ice.

9. Set up reverse transcription of DNA-free total RNA *or* poly(A)⁺ RNA with each of four different degenerate anchored oligo-dT primer sets (5'-T₁₂MN-3': T₁₂MG, T₁₂MA, T₁₂MT, and T₁₂MC, where M is G, A or C) as follows:

4 μl 5× MoMuLV reverse transcriptase buffer (1× final)
2 μl 0.1 M DTT (10 mM final)
1.6 μl 250 μM 4dNTP mix (20 μM final)
0.2 μg total RNA *or* 0.1 μg poly(A)⁺ RNA
2 μl of one 10 μM degenerate anchored oligo(dT) primer set (T₁₂MN;
1 μM final)
Adjust volume to 19 μl with DEPC-treated H₂O.

There will be four reactions for each RNA sample, each made with one degenerate primer set.

10. Incubate tube 5 min at 65°C to denature the mRNA secondary structure and incubate 10 min at 37°C to allow primer annealing.
11. Add 1 μl of 200 U/μl MoMuLV reverse transcriptase to each tube, mix well, and incubate 50 min at 37°C.
12. Incubate 5 min at 95°C to inactivate the reverse transcriptase and microcentrifuge briefly at high speed to collect condensate. Place tube on ice for immediate PCR amplification *or* store at -20°C for later use (stable at least 6 months).

Perform PCR amplification

13. Prepare a 20-μl reaction mix for each primer set as follows:

10 μl H₂O
2 μl 10× amplification buffer (1× final)
1.6 μl 25 μM 4dNTP mix (2 μM final)
0.2 μl [α -³³P]dATP
2 μl 2 μM arbitrary decamer (0.2 μM final)
2 μl 10 μM degenerate anchored oligo(dT) primer set (T₁₂MN; 1 μM final)
2 μl cDNA (step 12)
0.2 μl 5 U/μl *Taq* DNA polymerase.

*To avoid pipetting errors, prepare enough PCR reaction mix without the arbitrary decamer for 5 to 10 reactions and aliquot 18 μl to each tube. Then add the arbitrary decamer. Otherwise it is difficult to pipet accurately 0.2 μl of *Taq* DNA polymerase.*

14. Pipet up and down to mix well and overlay with 25 μl mineral oil.
15. Carry out PCR in a thermal cycler using the following amplification cycles:

40 cycles:	30 sec	94°C (denaturation)
	2 min	40°C (annealing)
	30 sec	72°C (extension)
1 cycle:	5 min	72°C (extension)
Final step:	indefinitely	4°C (hold).

The 2-min incubation at 40°C is to allow sufficient time for the short primers to anneal and start extension. The short extension period at 72°C is intended to amplify only short (<600-bp) DNA products to be separated on a denaturing polyacrylamide gel.

PCR products may be stored at 4°C until used.

16. Mix 3.5 μl PCR product with 2 μl formamide loading buffer and incubate 2 min at 80°C. Load sample onto a 6% denaturing polyacrylamide gel (UNIT 7.6). Run the gel ~3 hr at 60 W until xylene cyanol runs to within 10 cm of the bottom.

Flush out the urea from the gel wells with a syringe and needle just before loading samples to obtain high-resolution differential-display cDNA patterns.

Recover differentially displayed amplified DNAs

17. Carefully remove one of the glass gel plates. Place a piece of Whatman 3MM filter paper over the gel without trapping air bubbles between filter paper and gel. Dry the gel ~1 hr at room temperature without fixing it in methanol/acetic acid.

Fixing the gel with methanol/acetic acid will make it difficult to reamplify recovered DNA because DNA is labile at acidic pH, especially at the high temperature at which the gel is normally dried.

The dried gel should be handled with gloves to prevent DNA contamination. Always store the dried gel between two sheets of clean Whatman 3MM filter paper.

18. Use either radioactive ink or needle punches to mark X-ray film and dried gel to orient them. Expose the film 24 to 48 hr at room temperature (APPENDIX 3A).
19. Develop the film, align film with gel, and indicate DNA bands of interest (those differentially displayed in different lanes) either by marking beneath the film with a clean pencil or by cutting through the film.

Typical results of differential display are shown in Figure 25B.3.2.

20. Cut out gel slice and attached Whatman 3MM filter paper with a razor blade and place in a microcentrifuge tube. Add 100 μ l H₂O and incubate 10 min at room temperature.

If more than one band is differentially expressed, extract and reamplify each one separately.

21. Cap tube tightly and boil 15 min.

Place a lid-lock on the tube to prevent it from opening while boiling.

22. Microcentrifuge 2 min at high speed to pellet gel slice and paper debris. Decant supernatant into clean tube.
23. Add 10 μ l of 3 M sodium acetate (to give 0.3 M final) and 5 μ l of 10 mg/ml glycogen (as a carrier) to supernatant. Add 400 μ l of 100% ethanol and incubate 30 min at -70° C. Microcentrifuge 10 min at high speed, 4° C.

Glycogen is soluble at ethanol concentrations <85%.

24. Rinse pellet with 500 μ l of 85% ethanol, air-dry, and dissolve the DNA in 10 μ l H₂O.

Reamplify DNA

25. Reamplify 4 μ l of the eluted DNA in a 40- μ l reaction volume using the same degenerate anchored oligo(dT) primer set and PCR conditions as in steps 13 through 15, *except* add 3.2 μ l of 250 μ M 4dNTP mix (20 μ M final) instead of 1.6 μ l of 25 μ M 4dNTP mix and omit isotope. Save the remaining recovered DNA at -20° C for future reamplification (stable indefinitely).
26. Electrophorese 30 μ l of each PCR sample on a 1.5% agarose gel and stain with 0.5 μ g/ml ethidium bromide (UNIT 2.5A). Store the remaining PCR samples at -20° C (stable for years).

Most amplified DNAs should be visible after the first reamplification. Fragment molecular weights should be checked after reamplification to ensure that they are consistent with those on the denaturing polyacrylamide gel. If a DNA is not visible after the first reamplification, 4 μ l of 1/100 dilution (in water) of the first reamplification sample may be used for a second 40-cycle amplification.

27. Extract the desired reamplified DNA band from the agarose gel (*UNIT 2.6*) and use it as a probe for northern blot analysis (*UNIT 6.3*) and cDNA library screening (*UNIT 6.3*).
Store extracted DNA at -20°C (stable for years) if it is not to be used immediately.
28. Characterize remaining PCR sample (from step 26) by subcloning (*UNIT 15.4*) and sequencing (*UNIT 7.4*).

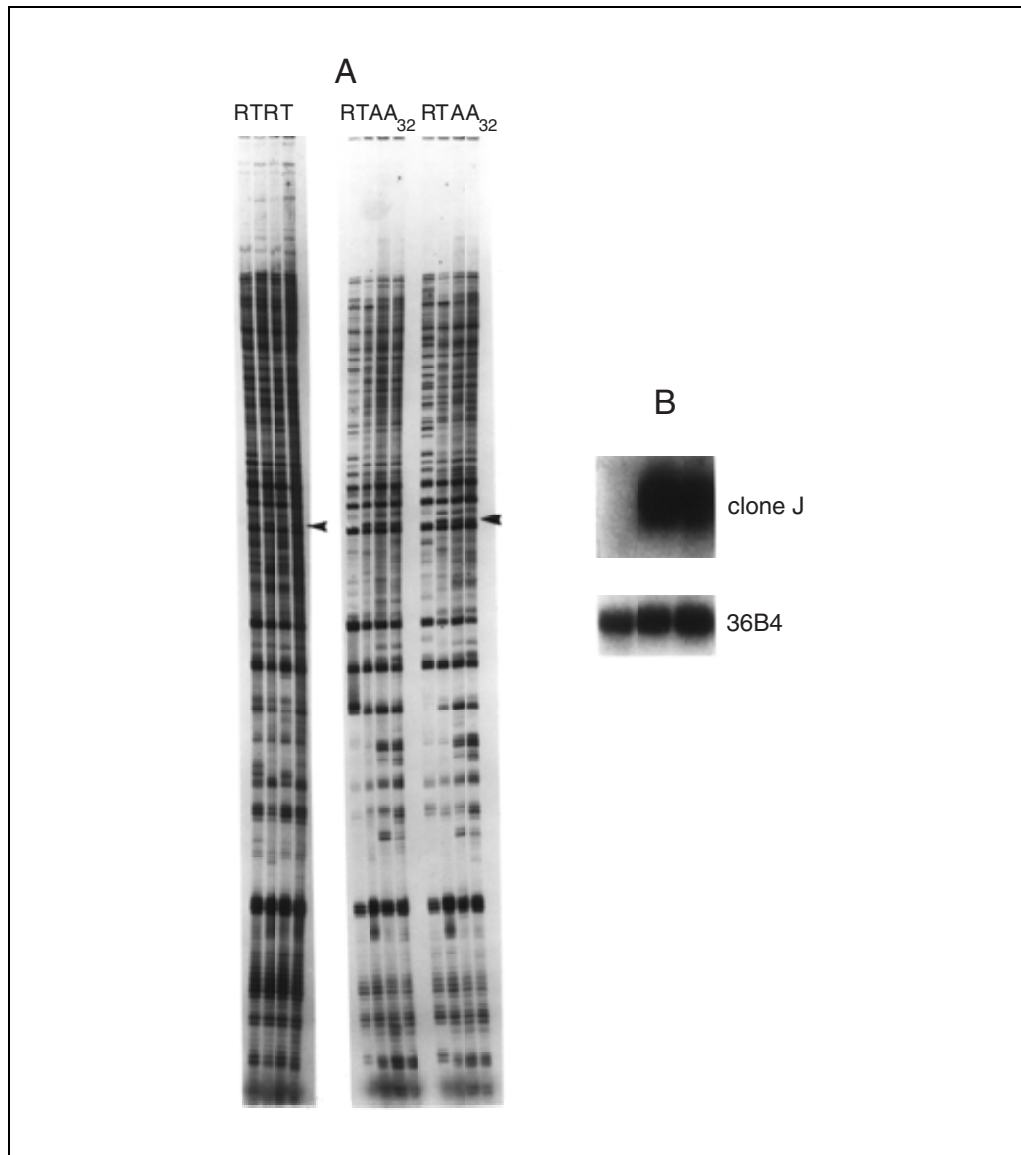


Figure 25B.3.2 Reproducibility and multiple display of mRNAs from normal versus ras/p53 mutant transformed cells. **(A)** RNA samples from normal rat embryo fibroblasts REF (R) and its ras/p53 doubly transformed derivative T101-4 cells (T) were reverse transcribed and amplified in duplicate with T12MA and OPA17 primers (left four lanes). In a separate experiment, RNA samples from REF (R), T101-4 (T), and another ras/p53 temperature-sensitive mutant transformed cell line A1-5 grown at nonpermissive temperature (A) and shifted to permissive temperature for 24 hr (A₃₂) were reverse transcribed and amplified in duplicate with T12MA and OPA17 primers (right eight lanes). An arrowhead indicates a reproducible difference between normal and transformed cells. **(B)** Northern blot analysis of this reamplified cDNA probe (named as clone J). 20 mg of total RNA from REF, T101-4, and A1-5 cells were analyzed. 36B4 was used as a probe for RNA loading control.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Formamide loading buffer

- 95% (v/v) formamide
 - 0.09% (w/v) bromphenol blue
 - 0.09% (w/v) xylene cyanol FF
- Store at 4°C

COMMENTARY

Background Information

Current methods to distinguish mRNAs in comparative studies rely largely on differential or subtractive hybridization techniques (Hedrick et al., 1984; Lee et al., 1991). Several important genes implicated in tumorigenesis have been isolated using these methods (Steege et al., 1988). Although subtraction is quite sensitive and can detect fairly rare mRNAs (see UNIT 25B.1), the method recovers genes incompletely and selects for genes in only one direction at a time during a two-way comparison between a pair of cells. The process is also laborious and time-consuming.

The differential display technique was developed with the goal of identifying differentially expressed genes, detecting individual mRNA species that are changed in different sets of mammalian cells, then recovering and cloning the cDNA (Liang and Pardee, 1993; Liang et al., 1993). This method utilizes polymerase chain reaction (PCR) amplification and denaturing polyacrylamide gel electrophoresis, two of the most commonly used molecular biological methods, and provides a sensitive, straightforward, and flexible approach to detect genes that are differentially expressed at the mRNA level.

In differential display, each RNA sample is first reverse transcribed with a degenerate anchored oligo(dT) primer set that anneals at the start of the poly(A) tails of mRNAs. Each degenerate anchored oligo(dT) primer set (e.g., T₁₂MA) will, in theory, reverse transcribe one-fourth of the total mRNA population. In combination with a decamer oligonucleotide of arbitrary sequence, which in theory can hybridize to any mRNA, cDNA fragments representing the 3' termini of mRNAs defined by both primers are amplified. Thus, this procedure allows amplification of an mRNA subpopulation without knowledge of sequence information. If any given arbitrary decamer does not actually sample all mRNAs, different de-

camers can be used to permit sampling of differential mRNA populations.

Differential display can be used for many purposes. One is to provide a picture of mRNA composition of cells by displaying subsets of mRNAs as short DNA bands. This mRNA fingerprinting is useful in the same way as are two-dimensional protein gels, for example, for observing alterations in gene expression. Secondly, these DNAs can be quickly reamplified, cloned, sequenced, and compared with sequences in data banks. Finally, reamplified cDNAs can be used as probes for northern or Southern blot hybridization and to isolate genes from genomic or cDNA libraries for further molecular characterization.

Investigations of expression genetics (Sager, 1997) has gained in preeminence. The differential display procedure is being successfully employed by many research groups to identify numerous expressed genes. Related publications have increased exponentially, and currently there are ~2000. For a cross section of results see Liang et al. (1994). Thus, differential display is a viable method for the identification of novel gene targets.

Critical Parameters and Troubleshooting

The most important, powerful application of differential display is to identify and clone differentially expressed genes in various biological systems. Because the method is based on reverse transcription-PCR (RT-PCR; UNIT 15.5), critical parameters relevant to that procedure generally apply for this protocol.

Utilization of this technique has encountered the problem of isolation of “false-positive” transcripts—i.e., PCR products that appear to be differentially expressed but which cannot be verified when subsequent northern analysis is performed using the same RNA source. PCR is highly sensitive to minor variations in experimental procedures and is noto-

riously difficult to make quantitative. In the authors' experience, success with differential display is dependent to a large degree on experimental design, great care in achieving consistency, the use of core reagent mixes, and duplicate assays, among other things.

Many modifications of the original protocol have been described, the implementation of which have resulted in enhanced fidelity and overall utility of this evolving technique.

Isolation of RNAs that are undegraded and that are free of contaminating DNA is necessary to select optimally for expressed genes (see Quality of RNA, below). A considerable number of articles propose modifications in choice of primers for both reverse transcriptase and PCR steps. Single base oligo(dT)-anchored primers reduce the number of reactions and redundancy (Liang et al., 1994). A recent study proposes primer sequences based on frequencies of gene sequences (Pesole et al., 1998). Longer arbitrary primers seem to enhance the reproducibility of the differential display patterns (Liang et al., 1994; Zhao et al., 1995). Labelling the PCR products with [³⁵S]- or [³³P]deoxynucleotides has safety advantages over [³²P] (Trentmann, 1995). Bands may be visualized nonradioactively with silver staining or fluorescence. Improved methods for cloning differential display products have also been proposed (Comes et al., 1997; Wybranietz and Laurer, 1998).

One band on a sequencing gel often contains more than one cDNA, and the contaminating band can generate a false northern signal if its mRNA is very plentiful. For avoiding false positives, cloning strategies (Zhao et al., 1996), restriction cutting (Prasher and Weissman, 1996), nested PCR reamplification (Zhang et al., 1996; Martin et al., 1998), and single-strand conformation polymorphism gels (Miele et al., 1998) can help to avoid this problem.

Direct sequencing of differentially expressed cDNAs has been reported (Wang and Feurstein, 1995). Dot blot grids are being developed to evaluate the differential display products (Martin et al., 1998).

Recently, other methods have been developed for studying expression genetics. These include representational difference analysis (Lisitsyn, 1995), serial analysis of gene expression (SAGE; Zhang et al., 1997), and dot blot analysis (Wodicka, 1997), by which differential mRNA expression is examined with high throughput mass cDNA library screening on dot blots placed on chips, together with powerful computational analysis of sequences. This

technique will in time provide massive amounts of information, although it is relatively laborious and requires special facilities.

Quality of RNA

The quality of RNA is determined by two criteria. First is the integrity of the RNA; second is the degree of chromosomal DNA contamination. The integrity of total RNA can be easily verified by agarose/formaldehyde gel electrophoresis, whereas the integrity of poly(A)⁺ RNA must be checked by northern blot hybridization using a cDNA probe for an mRNA with known molecular weight. Contamination by chromosomal DNA can be checked by performing differential display omitting the reverse transcription step. Under the PCR conditions used for differential display (i.e., low dNTP concentrations), RNA amplification is dependent on reverse transcription. Because total RNA isolated with various methods is generally found to be contaminated with DNA, it is recommended that, as a good practice, RNA samples be treated with DNase I before being used for differential display.

Design of arbitrary decamers

Generally any arbitrary decamer can be used as long it does not contain palindromic sequences and has a G+C content of 50% to 70%. The original decamer chosen for this application was from the mouse thymidine kinase gene (Liang and Pardee, 1993), but it has been used successfully to detect multiple mRNAs in cells of various species. Because the arbitrary decamers have been shown to contain up to 4-bp mismatches with the original cDNA templates and these mismatches are often clustered at the 5' end of the primers (Liang et al., 1993), the arbitrary decamers can be designed in such a way that the 3' sequences are maximally randomized while the 5' bases (up to four bases) are fixed. The G+C content of the arbitrary decamers can be increased or decreased to reflect the G+C content of the genome of the organism from which the mRNA is isolated.

False-positive difference

The intrinsic problem encountered with differential display, as with any method based on PCR, is that it is highly sensitive to minor variations. True differences in expression must be differentiated from the "noise" that is the major source of false-positive differences. If a pair of RNA samples is to be compared, the displayed (DNA) pattern differences must be reproducible. An advantage of differential dis-

play is the ability to simultaneously compare more than two relevant RNA samples (e.g., from different cell types or stages of development); multiple display thus has a built-in internal control for distinguishing “noise” from true differences. This also facilitates isolation of genes that really give useful results for the system under study.

Size of DNA probe

Short DNA probes (<150 bp) have been found to be hard to label and often fail to produce any signals in northern blot hybridizations. Therefore, it is advised that only DNA bands >150 bp be further characterized by northern blot analysis and that smaller bands be ignored.

Anticipated Results

This method should produce reproducible amplified DNA patterns. The reproducible DNA bands representing differentially expressed genes should be readily reamplifiable and usable as probes for northern blot analysis or cDNA library screening.

Time Considerations

The whole procedure from RNA to DNA samples ready to use as probes can be performed within three days. Treating RNA with DNase I and checking its integrity by gel electrophoresis takes ~2 hr. Reverse transcription takes ≤2 hr. Setting up 40 PCR samples requires 1 to 2 hr. PCR amplification requires ~4 hr but can be performed overnight. Preparing, running, and drying the denaturing polyacrylamide gel takes 1 day. Autoradiography can be as brief as overnight. Recovery, reamplification of DNA, and extraction of reamplified DNA from an agarose gel can easily fit into the third day. Northern blot analysis requires an additional 2 days.

Literature Cited

- Comes, A., Humbert, J., and Laurent, F. 1997. Rapid cloning of PCR-derived RAPD probes. *BioTechniques* 23:210-212.
- Hedrick, S.M., Cohen, D.I., Nielsen, E.A., and Davis, M.M. 1984. Isolation of cDNA clones encoding T cell specific membrane associated proteins. *Nature* 308:149-153.
- Lee, S.W., Tomasetto, C., and Sager, R. 1991. Positive selection of candidate tumor-suppressor genes by subtractive hybridization. *Proc. Natl. Acad. Sci. U.S.A.* 88:2825-2829.
- Liang, P. and Pardee, A.B. 1993. Distribution and cloning of eukaryotic mRNAs by means of dif-

ferential display: Refinements and optimization. *Nucl. Acids Res.* 21:3269-3275.

- Liang, P., Averboukh, L., and Pardee, A.B. 1993. Distribution and cloning of eukaryotic mRNAs by means of differential display: Refinements and optimization. *Nucl. Acids Res.* 21:3269-3275.
- Liang, P., Zhu, W., Zhang, X., Gui, Z., O'Connell, R.P., Averboukh, L., Want, F., and Pardee, A.B. 1994. Differential display using one-base anchored oligo-dT primers. *Nucleic Acids Res.* 22:5763-5764.
- Lisitsyn, N.A. 1995. Representational difference analysis: Finding the difference between genomes. *Trends Genet.* 11:303-307.
- Martin, K.J., Kwan, C.-P., O'Hare, M.J., Pardee, A.B., and Sager, R. 1998. Identification and verification of differential display cDNAs using gene-specific primers and hybridization arrays. *BioTechniques* 24:1018-1026.
- Miele, G., MacRae, L., McBride, D., Manson, J., and Clinton, M. 1998. Elimination of false positives generated through PCR reamplification of differential display cDNA. *BioTechniques* 25:138-144.
- Pesole, G., Liuni, S., Grillo, G., Belichard, P., Trenkle, T., Walse, J., and McClelland, M. 1998. GeneUp: A program to select short PCR primer pairs that occur in multiple members of sequence lists. *BioTechniques* 25:112-123.
- Prasher, Y. and Weissman, S.M. 1996. Analysis of differential gene expression by display of 3' end restriction fragments of cDNAs. *Proc. Natl. Acad. Sci. U.S.A.* 93:659-663.
- Sager, R. 1997. Expression genetics: Shifting the focus from DNA to RNA. *Proc. Natl. Acad. Sci. U.S.A.* 94:952-955.
- Steege, P.S., Bevilacqua, G., Kopper, L., Thorgeirson, U.P., Talmadge, J.E., Liotta L.A., and Sobel, M.E. 1988. Evidence for a novel gene associated with low tumor metastatic potential. *J. Natl. Cancer Inst.* 80:200-204.
- Trentmann, S.M., van der Dnapp, E., and Kende, H. 1995. Alternatives to ³⁵S as a label for the differential display of eukaryotic messenger RNA. *Science* 267:1186-1187.
- Wang, W. and Feurstein, G.Z. 1995. Direct sequencing of DNA isolated from mRNA differential display. *BioTechniques* 18:448-453.
- Wodicka, L., Dong, H., Mittmann, M., Ho, M.-H., and Lockhart, D.J. 1997. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nature Biotechnology* 15:1359-1367.
- Wybranietz, W. and Lauer, U. 1998. Distinct combination of purification methods dramatically improves cohesive-end subcloning of PCR products. *BioTechniques* 24:578-580.
- Zhang, H., Zhang, R., and Liang, P. 1996. Differential screening of gene expression difference enriched by differential display. *Nucleic Acids Res.* 24:2454-2455.

Zhang, L., Zhou, V.E., Velculescu, S.E., Kern, R.H., Hruban, S.R., Hamilton, B., Vogelstein, B., and Kinzler, K.W. 1997. Gene expression profiles in normal and cancer cells. *Science* 276:1268-1272.

Zhao, S., Ooi, S.L., and Pardee, A.B. 1995. New primer strategy improves precision of differential display. *BioTechniques* 18:842-850.

Zhao, S., Ooi, S.L., and Pardee, A.B. 1996. Three methods for the identification of true positive cloned cDNA fragments in differential display. *BioTechniques* 20:400-402

Key Reference

Liang et al., 1993. See above.

Uses the protocol outlined here and presents examples of data generated.

Contributed by Peng Liang
Vanderbilt-Ingram Cancer Center
Nashville, Tennessee

Arthur B. Pardee
Dana Farber Cancer Institute
Boston, Massachusetts

Restriction-Mediated Differential Display (RMDD)

Restriction-mediated differential display (RMDD) can be applied to identify differentially expressed (i.e., up- or down-regulated) genes in many eukaryotic cells or tissues by comparison of band patterns obtained from two or more different RNA preparations. As opposed to early differential display or other RNA-fingerprinting protocols based on arbitrarily primed PCR, RMDD provides very robust and reproducible results which are largely independent of the exact amount of input material or of the exact cycling conditions, respectively. Two different PCR strategies for fragment amplification, depending on the complexity of the material under investigation as well as the appropriate choice of the restriction enzyme or enzymes used, are discussed (see Strategic Planning). The first protocol describes oligo(dT)-primed conversion of total RNA into double-stranded cDNA, which is cleaved with a frequently cutting restriction enzyme, ligated to linker molecules (thus creating the “RMDD library”), and amplified with labeled selective 3'-elongated oligonucleotide primers to generate subpools of amplified fragments which represent the 3'-ends of the cDNA molecules (see Basic Protocol and Fig. 25.B4.1). A protocol outlining two-phase PCR is given as an alternative to the amplification steps used in the Basic Protocol (see Alternate Protocol). This protocol is usually chosen if the RNA samples to be analyzed are particularly complex. The final protocol describes nonradioactive fragment analysis through the use of biotinylated primers and direct-blotting electrophoresis (see Support Protocol).

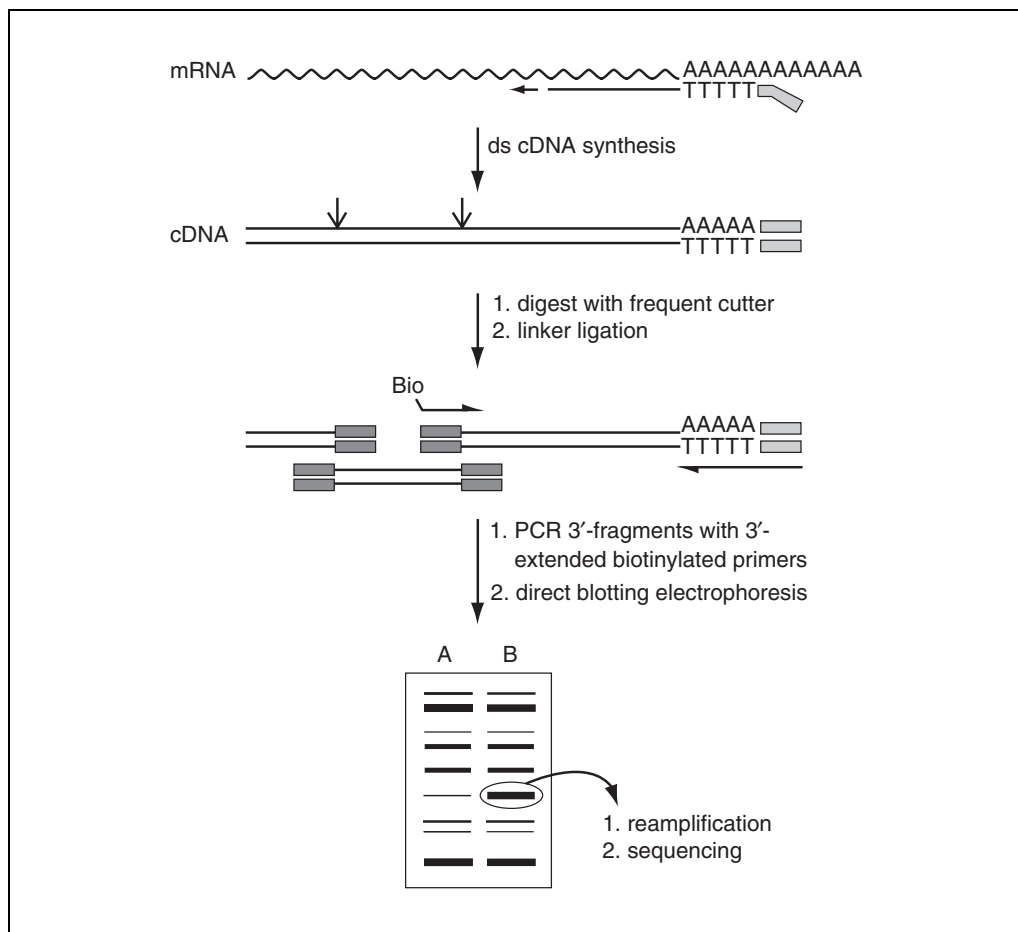


Figure 25B.4.1 Schematic of RMDD.

Contributed by Achim Fischer

Current Protocols in Molecular Biology (2001) 25B.4.1-25B.4.17

Copyright © 2001 by John Wiley & Sons, Inc.

NOTE: 5'-labeled primers are indicated by an asterisk (*). The label can be a radioactive isotope (e.g., ³²P) or a nonradioactive label such as biotin or digoxigenin. In the latter case, labels should be attached via a sufficiently long spacer to the oligonucleotide (e.g., tetraethylene glycol from Eurogentec) to ensure maximum detection sensitivity.

NOTE: The technology described in this unit is protected by certain patent rights (US 5,876,932; EP 0 743 367; JP 96/308598). Commercial application of RMDD (including in-house research projects of any company) thus requires a license. No license is required for academic use.

STRATEGIC PLANNING

The “RMDD library” contains a mixture of restriction fragments of all cDNA molecules obtained from the respective biological sample. It has been estimated that a single cell type contains ~10,000 different mRNA molecules, resulting in 10,000 different cDNA species. For successful gel display of the fragments derived from the 3' ends of these cDNA molecules, a strategy must be provided to subdivide this rather complex fragment mixture into a number of subpools, each containing a sufficiently low number (i.e., ≤50 to 100) of different fragment species. This can be easily achieved by fragment amplification employing oligonucleotide primers each carrying one additional “selective” base at the 3'-end. Theoretically, such a selective base allows primer extension by a polymerase only if it perfectly matches the corresponding base on the other strand. Combining selective primers directed against the ligated linker and against the sequence introduced by the cDNA primer thus allows subdivision of fragments into nonoverlapping subpools. The RMDD protocol (see Basic Protocol and Alternate Protocol) involves two subsequent rounds of amplification, the first employing selective primers extended by one base each and the second employing primers extended by one more base, providing a total number of $12 \times 16 = 192$ reactions to be performed for complete coverage of all generated 3'-end fragments. Two rounds are chosen, since the discrimination of a polymerase against extension of primers distinguished by a mismatch at the second last position is much less pronounced than the discrimination against extension of terminal mismatches, prohibiting use of primers carrying two selective bases at their 3'-end in a single round of PCR. However, in practice, a certain extent of “bleedthrough” can still be observed (i.e., amplification of fragments with a given selective primer, which theoretically should not take place due to a 3'-terminal mismatch of the annealed primer).

If mRNA complexity is not too high (e.g., material obtained from cell cultures or “simple” tissues of low complexity), this “bleedthrough” usually does not cause any problems; nevertheless, when working with highly complex samples (e.g., RNA isolated from mammalian brain), bleedthrough may render band patterns too crowded for reliable isolation of particular bands of interest. To reduce bleedthrough, both first and second amplification reactions can be performed in a “two-phase” manner (see the Alternate Protocol). The first phase, performed at extremely low concentrations of dNTPs (i.e., 2 μM each), involves 10 (first amplification) or 15 (second amplification) cycles and defines which products will be amplified to a detectable level. This phase exploits the fact that mismatch extension can be significantly reduced at low dNTP levels. For the second phase, dNTP concentrations are raised to “normal” levels (i.e., 200 μM each), which, after an additional 10 cycles, allows for accumulation of the desired amount of product.

The choice of the restriction enzyme used for RMDD depends on the particular organism to be analyzed, since average fragment size may vary due to differences in codon usage and G/C content. To obtain cDNA 3'-fragments in a size range optimal for gel display (i.e., most of the fragments having a size between 100 and 700 bp), an appropriate enzyme

has to be employed—e.g., *Mbo*I, as described in this protocol (see Basic Protocol)—which has proven satisfactory with RNA isolated from man, rat, mouse, corn, and *Arabidopsis*. Should another enzyme be chosen, linker and linker primer sequences will have to be modified accordingly, and the same holds true if, for the sake of more complete coverage of the transcriptome, experiments are repeated with a second enzyme. Computer analysis has demonstrated that in man and rodents roughly 80% to 85% of all transcripts contain a recognition site for *Mbo*I (unpub. observ.); therefore, 15% to 20% of transcripts would be inaccessible to analysis using this particular enzyme. Accordingly, if nearly complete coverage of transcripts is desired, a second-pass RMDD analysis might be performed with a second frequently cutting enzyme. Performing RMDD with a second enzyme, assuming both enzymes recognize 80% of cDNAs each, would provide a total coverage of 96% of all transcripts.

RMDD LIBRARY PREPARATION AND TWO-ROUND AMPLIFICATION

This protocol describes conversion of total RNA to labeled PCR products, which are ready to be displayed by gel electrophoresis.

Materials

50 µg total RNA (UNITS 4.1 & 4.2)
RNase-free water
10 µM cDNA primer CP29V: 5'-ACC TAC GTG CAG ATT TTT TTT TTT TTT TX₁-3' (X₁ = A, C, or G; equimolar amounts of all three species; see UNIT 2.11 for oligonucleotide synthesis)
100 mM RNase-free DTT (Life Technologies)
5× SuperScript buffer (Life Technologies)
10 mM RNase-free and standard dNTPs
40 U/µl RNase inhibitor (e.g., RNasin)
200 U/µl SuperScript II reverse transcriptase (Life Technologies)
5× second-strand buffer II (UNIT 5.5)
1.5 U/µl RNase H
10 U/µl *E. coli* DNA polymerase I
Phenol equilibrated with TE buffer, pH 8.0 (UNIT 2.1A)
Chloroform
20 mg/ml glycogen
28% PEG 8000/3.6 mM MgCl₂ (see recipe)
70% and 100% ethanol
10× universal buffer (Stratagene)
4 U/µl *Mbo*I restriction endonuclease (Stratagene)
3 M sodium acetate, pH 5.2 (APPENDIX 2)
10 mM ATP
0.5 µg/µl *Mbo*I-linker ML2025 (see recipe)
T4 DNA ligase and 10× buffer (Roche)
1× and 0.25× TE buffer, pH 8.0 (APPENDIX 2)
4 µM primer CP28X₁: 5'-ACC TAC GTG CAG ATT TTT TTT TTT TTT TX₁-3' (X₁ = A, C, or G; see UNIT 2.11 for oligonucleotide synthesis)
4 µM primer ML19Y₁: 5'-TGC TAA GTC TCG CGA GAT CY₁-3' (Y₁ = A, C, G, or T; see UNIT 2.11 for oligonucleotide synthesis)
10× PCR buffer (see recipe)
20 mM MgCl₂ (APPENDIX 2)
RediLoad (Research Genetics)
5 U/µl *Taq* DNA polymerase
100-bp DNA size ladder (e.g., Life Technologies)

**BASIC
PROTOCOL**

**Discovery of
Differentially
Expressed Genes**

25B.4.3

1.5% agarose gel (UNIT 2.5A)

4 μM primer CP28X₁X₂: 5'-ACC TAC GTG CAG ATT TTT TTT TTT TTT T
X₁X₂-3' (X₂ = A, C, G, or T; see UNIT 2.11 for oligonucleotide synthesis)

4 μM labeled primer *ML18Y₁Y₂: 5'-*GCT AAG TCT CGC GAG ATC Y₁Y₂-3'
(Y₂ = A, C, G, or T; see UNIT 2.11 for oligonucleotide synthesis)

Formamide buffer: 5 mM EDTA/0.1% bromophenol blue in 99% deionized formamide

22°, 37°, 42°, 65° and 75°C water bath, heat blocks, or equivalent

Thermal cycler with heated lid

96-well PCR plates (e.g., MJ Research)

Additional reagents and equipment for ethanol precipitation and phenol/chloroform extraction of DNA (UNIT 2.1A), and pouring and running (UNIT 2.5A) agarose and 6% polyacrylamide gels (UNIT 7.6)

Synthesize first-strand cDNA

1. Ethanol precipitate 50 μg total RNA (UNIT 2.1A) and dissolve in 15.5 μl RNase-free water. Add 1.5 μl of 10 μM cDNA primer CP29V, denature 5 min at 65°C (e.g., in a heat block), and cool down on ice.

It is not necessary to isolate poly(A+) RNA. Band patterns obtained with mRNA are virtually identical to those obtained with total RNA. On the other hand, mRNA isolation is a potential source of variation and should therefore be avoided.

DEPC treatment will not usually be required for RNase-free water.

2. Assemble components for first-strand synthesis on ice (29.1 μl total):

17.0 μl freshly denatured RNA with cDNA primer

3.0 μl 100 mM RNase-free DTT

6.0 μl 5 \times SuperScript buffer

1.5 μl 10 mM RNase-free dNTPs

0.6 μl 40 U/ μl RNase inhibitor (e.g., RNasin)

1.0 μl 200 U/ μl SuperScript II reverse transcriptase.

Mix well and incubate 1 hr at 42°C. Stop reaction by placing on ice.

Incubation is best done in a water bath or thermal cycler. Hot air ovens do not guarantee sufficiently quick heating of samples.

To check for possible RNA degradation in the course of first-strand synthesis due to RNase contamination, 0.5 to 1 μl of the first-strand synthesis reaction can be analyzed on a 1% standard agarose gel (UNIT 2.5A; no special RNA gel is required), watching for undegraded ribosomal RNA bands.

Synthesize second-strand cDNA

3. Assemble on ice the following components (207.2 μl total) for second-strand synthesis:

48 μl 5 \times second-strand buffer II

3.6 μl 10 mM dNTPs

148.4 μl H₂O

1.2 μl 1.5 U/ μl RNase H

6.0 μl 10 U/ μl *E. coli* DNA polymerase I.

4. Combine first-strand and second-strand synthesis reactions. Mix and incubate for 2 hr at 22°C. After completion of second-strand synthesis, inactivate DNA polymerase by heating for 20 min to 75°C.

Purify cDNA

5. Extract with 100 μ l phenol equilibrated with TE buffer, pH 8.0. Extract again with 100 μ l chloroform.

UNIT 2.1A describes the procedures for phenol-chloroform extraction of DNA.

CAUTION: Phenol and chloroform are severe health hazards. See UNIT 2.1A for precautions.

6. For size-selective PEG precipitation, carefully mix:

200 μ l phenol/chloroform-extracted ds cDNA
1.0 μ l 20 mg/ml glycogen
200 μ l 28% PEG 8000/3.6 mM MgCl₂.

Let the reaction (401 μ l total) stand at room temperature for 5 min, then microcentrifuge for 15 min at maximum speed, 10°C. Wash pellet carefully with 70% ethanol.

This precipitation step removes unincorporated cDNA primer as well as small (i.e., below ~100 nt) nucleic acid molecules. Since size-selective PEG precipitations are susceptible to minor concentration changes, it is imperative to adhere to the following guidelines:

1. *Make sure to pipet exactly 200 μ l ds cDNA. Vapor pressure of chloroform dissolved in the aqueous phase tends to displace liquid from the pipet tip, making accurate pipetting difficult. One way to overcome this problem is to repeatedly (5 to 10 times) withdraw and expel again ~50 to 100 μ l of the chloroform-saturated aqueous phase before pipetting the required 200 μ l, thus allowing the pipet to saturate with chloroform vapor.*

2. *The 28% PEG/3.6 mM MgCl₂ solution is rather viscous. Pipet slowly and carefully, again being sure to accurately transfer the required volume.*

3. *Mix carefully by first repeatedly inverting the tube, then vigorously vortexing. Due to viscosity, complete and homogeneous mixing takes a while.*

During addition of PEG solution, a white glycogen precipitate usually forms. This becomes invisible again in the course of mixing.

When washing the pellet with ethanol, detachment from the tube wall does no harm since the pellet is too large to be easily lost.

Perform restriction digest

7. Dissolve the pellet on ice in the following solution (96 μ l total):

15.0 μ l 10 \times universal buffer
81.0 μ l H₂O.

Instead of the Universal buffer supplied by Stratagene, any buffer supplied with the restriction enzyme can be used. In this case, adhere to the manufacturer's recommendations concerning dilution of buffer stock.

8. Add 4.0 μ l of 4 U/ μ l MboI and incubate 1 hr at 37°C. Inactivate the enzyme by heating 20 min at 65°C.

The choice of restriction enzyme is discussed elsewhere in this unit (see Strategic Planning).

9. Extract with 50 μ l phenol buffered with TE buffer, pH 8.0, then with 50 μ l chloroform. Add 1 μ l glycogen and 10 μ l 3 M sodium acetate, pH 5.2, followed by 2.5 vol 100% ethanol. Microcentrifuge 20 min at maximum speed and wash pellet with 70% ethanol. Air dry pellet briefly (5 to 10 min). Do not apply heat and/or vacuum, since overdrying DNA pellets might make resuspending them difficult.

Perform linker ligation

10. Dissolve pellet in ligation mix (20 μ l total), consisting of the following components:

- 1.2 μ l 10 \times ligation buffer
- 2.0 μ l 10 mM ATP
- 8.0 μ l 0.5 μ g/ μ l *Mbo*I-linker ML2025
- 7.8 μ l H₂O
- 1.0 μ l 1 U/ μ l T4 DNA ligase.

Ligate overnight at 16°C or over the weekend at 4°C.

11. Add 90 μ l water, mix, and extract with 50 μ l phenol buffered with TE buffer, pH 8.0, then with 50 μ l chloroform. For removal of unligated linkers, assemble a second PEG precipitation reaction (201 μ l total):

- 100 μ l phenol-extracted ligation products
- 1.0 μ l glycogen
- 100 μ l 28% PEG/3.6 mM MgCl₂.

Let stand at room temperature 5 min, then microcentrifuge 15 min at maximum speed, 10°C. Wash pellet carefully with 70% ethanol and resuspend in 40 μ l TE buffer, pH 8.0.

For precautions, see step 6.

Perform first-round amplification of 3'-cDNA fragments

12. Set up first-round amplification reactions by combining 1 μ l of each of the three 4 μ M CP28X₁ (X₁ = A, C, or G) primers with 1 μ l of each of the four 4 μ M ML19Y₁ (Y₁ = A, C, G, or T) primers in separate tubes on ice (12 reactions total). Assemble a master mix with all remaining components (recipe is for 1 reaction):

- 2.0 μ l template (PEG-precipitated ligation products)
- 2.0 μ l 10 \times PCR buffer
- 1.5 μ l 20 mM MgCl₂
- 0.4 μ l 10 mM dNTPs
- 2.0 μ l RediLoad
- 9.9 μ l H₂O
- 0.2 μ l 5 U/ μ l Taq DNA polymerase.

Assemble reactions and place the tubes in the wells of a thermocycler preheated to 90°C.

13. Apply the following cycling program:

- | | | |
|---------------|--------------|-------------------------|
| Initial step: | 1 min | 94°C (denaturation) |
| 25 cycles: | 20 sec | 94°C (denaturation) |
| | 30 sec | 65°C (primer annealing) |
| | 4 min | 72°C (primer extension) |
| Final step: | indefinitely | 10°C (hold/extension). |

14. Load 10 μ l of each reaction onto a 1.5% agarose gel and check for successful amplification by agarose gel electrophoresis (UNIT 2.5A). Include a 100-bp ladder as a size marker.

PCR conditions are adjusted in such a way that the amount of primers limits the amount of product. The long extension time ensures that differently sized products are simultaneously amplified essentially without a bias against the longer ones. Agarose gel electrophoresis should yield smears between ~100 bp and ~700 bp with very few (if any) discrete bands being visible. Most importantly, reactions obtained with the same primer combina-

tion, but from different RNA samples to be compared, should look essentially indistinguishable. If appearance and/or amount of material should visibly differ, probably one of the enzymatic steps prior to amplification was performed at too low an efficiency (see Troubleshooting and Table 25.B4.1).

15. Dilute reactions 1:100 with 0.25× TE buffer, pH 8.0.

Diluted reactions can be indefinitely stored at –20°C.

Perform second-round amplification of 3'-cDNA fragments

16. Set up second-round amplification mix by combining in 96-well plates each of the 12 CP28X₁X₂ primers with each of the 16^{*}ML18Y₁Y₂ primers (192 different reactions per sample; 20 µl each):

2.0 µl template (diluted first-round PCR)
2.0 µl 10 × PCR buffer
1.5 µl 20 mM MgCl₂
0.4 µl 10 mM standard dNTPs
2.0 µl 4 µM primer CP28X₁X₂ (X₂ = A, C, G, or T)
2.0 µl 4 µM labeled primer ^{*}ML18Y₁Y₂ (Y₂ = A, C, G, or T)
2.0 µl RediLoad
7.9 µl H₂O
0.2 µl 5 U/µl *Taq* DNA polymerase.

Make sure that for every reaction, X₁ and Y₁ of the second-round amplification are identical to X₁ and Y₁ of the first-round amplification. PCR can be conveniently performed in two 96-well plates per RNA sample.

It is highly preferable to use a thermocycler equipped with a hot top, obviating the need to cover reactions with oil.

Use of labeled primers instead of incorporating labeled nucleotides has the advantage that (1) only one of two complementary strands is visualized, thus limiting complexity of band patterns (usually, two complementary strands of equal length show slightly different mobility in polyacrylamide gels), and (2) label intensity does not increase with fragment length. In addition, if biotin is used as a label, incorporation of an undefined number of biotin molecules (it is not possible to replace all nucleotides of a given type by enzymatic incorporation of the biotinylated analog) into the amplified strands leads to pronounced smearing of the obtained bands due to the incremental mobility shift caused by each of the biotin groups in a DNA molecule.

17. Apply the cycle program of step 13, but for only 20 cycles. Check for successful amplification by agarose gel electrophoresis (also see step 13).

Reactions obtained with the same primer combination but from different RNA samples should again look essentially indistinguishable, whereas reactions obtained with different primer combinations usually look distinct. Other than with first-round PCR products, usually a small number of discrete bands (e.g., 1 to 5) can be observed.

18. Transfer 5 µl of each reaction into a fresh microtiter plate containing 5 µl formamide buffer per well and mix. Denature 2 min at 75°C.

Nonradioactively labeled PCR products in formamide buffer can be stored for several months at –20°C. When radioactive labeling is chosen, storage time is limited by decay of the incorporated isotope.

Label nucleotides

- 19a. *For radioactive labeling:* Load 1 to 2 µl sample into the slots of a denaturing 6% polyacrylamide gel and run as described in UNIT 25B.3, starting at Basic Protocol, step 16, of that unit.

19b. *For nonradioactive labeling and use of direct blotting electrophoresis:* See Support Protocol, below.

**ALTERNATE
PROTOCOL**

AMPLIFICATION BY TWO-PHASE PCR

Alternatively, the amplification steps (see Basic Protocol, steps 12 to 17) can be replaced by a two-phase PCR (see Strategic Planning). This procedure decreases the “bleedthrough” sometimes observed between different PCRs obtained from the same sample by “illegitimate priming” (i.e., priming with a mismatch at the primer’s 3′-ultimate base). The approach is to perform the first 10 to 15 cycles of each PCR at an extremely low nucleotide concentration (2 μ M each), which increases the bias of *Taq* polymerase against mismatch extension. After these initial cycles are finished and the product composition in each reaction has been defined, reactions are supplemented with nucleotides to a final concentration of 200 μ M each, thus allowing sufficient amounts of amplification products to be generated. The drawback is the increase in hands-on time required for pipetting.

Additional Materials (also see Basic Protocol)

0.1 mM dNTPs (freshly diluted from 10 mM dNTPs)

Perform first-round low-concentration amplification

1. Synthesize ds cDNA (see Basic Protocol, steps 1 to 11).
2. Set up first-round 2- μ M amplification reactions (12 different reactions per sample, 20 μ l each):

2.0 μ l template (PEG-precipitated ligation products; see Basic Protocol, step 11)
2.0 μ l 10 \times PCR buffer
1.5 μ l 20 mM MgCl₂
0.4 μ l 0.1 mM dNTPs (freshly diluted from 10 mM dNTPs)
2.0 μ l 4 μ M primer CP28X₁ (X₁ = A, C, or G)
2.0 μ l 4 μ M primer ML19Y₁ (Y₁ = A, C, G, or T)
9.9 μ l H₂O
5 U/ μ l 0.2 μ l *Taq* DNA polymerase.

Again, all PCR mixtures should be prepared as master mixes.

3. Carry through the same program as described (see Basic Protocol, step 13), except for 15 rather than 25 cycles.

Perform first-round normal-concentration amplification

4. Transfer reaction tubes to ice. To each tube add 20 μ l of 200 μ M amplification mix, prepared as follows:

2.0 μ l 10 \times PCR buffer
1.5 μ l 20 mM MgCl₂
0.8 μ l 10 mM dNTPs
4.0 μ l RediLoad
11.5 μ l H₂O
0.2 μ l 5 U/ μ l *Taq* DNA polymerase.

5. Repeat the program cycle as described (see Basic Protocol, step 13), performing the remaining cycles (i.e., 16 to 25).

6. Check products by agarose gel electrophoresis (see Basic Protocol, step 14 and *UNIT 2.5A*).
7. Dilute reactions 1:100 with 0.25× TE buffer.

Perform second-round low-concentration amplification

8. Using 96-well microtiter plates, set up second-round 2 μM amplification reactions (192 different reactions per sample; 20 μl each):

2.0 μl template (diluted first-round PCR)
 2.0 μl 10× PCR buffer
 1.5 μl 20 mM MgCl₂
 0.4 μl 0.1 mM dNTPs
 4.0 μl 4 μM primer CP28X₁X₂ (X₂ = A, C, G, or T)
 4.0 μl 4 μM labeled primer *ML18Y₁Y₂ (Y₂ = A, C, G, or T)
 5.9 μl H₂O
 5 U/μl 0.2 μl *Taq* DNA polymerase.

9. Transfer plates to the preheated wells of a thermal cycler and cycle as described above (see Basic Protocol, step 13), except for only 10 rather than 25 cycles.

Perform second-round normal-concentration amplification

10. Cool reaction tubes on ice and add 20 μl of 200 μM amplification mix (step 4) to each tube.
11. Repeat the program (see Basic Protocol, step 13), this time using 20 cycles (i.e., add 10 more cycles).
12. Check products by agarose gel electrophoresis (see Basic Protocol, step 14 and *UNIT 2.5A*).

Agarose gel electrophoresis can be skipped if radioactive label is used. In the latter case, adhere to the usual precautions for working with radioisotopes (APPENDIX 1F) and handle samples at a dedicated workspace only.

DIRECT BLOTTING ELECTROPHORESIS

The authors have found direct blotting electrophoresis (DBE) to be an extremely helpful technique to get high-quality display results from amplified RMDD products and to simplify physical access to bands of interest. In contrast with standard fragment analysis (see Chapter 2) based on radioactive labeling, it is not necessary, for the sake of optimal resolution of different size ranges, to perform “short” and “long” runs of each sample. In DBE, all fragments, including the largest ones, pass the whole length of the gel before being transferred to the blotting membrane, providing unsurpassed resolution of bands in the size range relevant for RMDD. Working with nonradioactive materials provides considerable convenience, and stained bands can be directly cut out of the blotting membrane for recovery and analysis.

Additional Materials (also see *Basic Protocol*)

TBE electrophoresis buffer (*APPENDIX 2*) standard and degassed (i.e., stirred under vacuum 20 min)
 Maleic buffer, pH 7.5 (see recipe)
 1.5% blocking reagent (see recipe)
 Streptavidin-alkaline phosphatase conjugate (Roche Molecular Biochemicals)
 Reaction buffer, pH 9.5 (see recipe)
 NBT/BCIP in 67% (v/v) DMSO (Roche Molecular Biochemicals)

**SUPPORT
 PROTOCOL**

**Discovery of
 Differentially
 Expressed Genes**

25B.4.9

Primers (see *UNIT 2.11* for oligonucleotide synthesis):

CP28: 5'-ACC TAC GTG CAG ATT TTT TTT TTT TTT T-3'

ML18: 5'-GCT AAG TCT CGC GAG ATC-3'

GATC 1500 Direct Blotting Electrophoresis System (GATC Biotech AG)

Direct blotting membrane (GATC Biotech AG)

10-ml syringe and 25-G needle

32-well sharkstooth comb

GELoader tips (Eppendorf) with capillary-like part cut away

Stratalinker (Stratagene)

Developing drum (e.g., GATC tube; GATC Biotech AG)

Adhesive tape

Rolling incubator accepting 18 × 35-cm tubes and capable of revolving at ~20 rpm

2-mm-thick polyethylene wrap (e.g., Neolab, Heidelberg, FRG) or material from a thick hybridization bag

T-A cloning system (e.g., Invitrogen; optional)

Additional reagents and materials for casting denaturing polyacrylamide gels (*UNIT 2.12*), agarose gel electrophoresis (*UNIT 2.5A*), and molecular cloning of PCR products (*UNIT 15.7*).

NOTE: For details concerning use of the GATC 1500 Direct Blotting Electrophoresis apparatus, consult the manufacturer's instructions.

Prepare the gel

1. Cast a denaturing 4.5% polyacrylamide gel (*UNIT 2.12*). Attach a 40- to 45-cm long piece of blotting membrane to the conveyor belt of the direct blotting electrophoresis system. Mount the gel on the apparatus and fill with the appropriate amount of TBE electrophoresis buffer, using degassed buffer in the lower chamber. Move the leading edge of the membrane 1 cm past the lower edge of the gel. Connect apparatus to a high-voltage power supply.

When choosing the direct blotting technique, all fragments, including the largest ones, pass through the whole length of the gel. Thus, a lower acrylamide concentration (i.e., 4.5% instead of 6%) is used as compared to the concentration used for standard sequencing gels.

2. Prerun (i.e., with no sample) the gel for 30 min with the power supply set to 2000 V and 30 W as limiting parameters.

Electrophorese samples and transfer to the membrane

3. Rinse gel slot with TBE buffer using a 10-ml syringe and 25-G needle, and insert a 32-well sharkstooth comb. Using GELoader tips with the capillary-like part cut away, load 1 to 1.5 μ l denatured reaction (see Basic Protocol, step 18) per well, being sure to load the whole gel within ~10 min.

Although 48-well combs are available as well, no satisfactory results could be obtained with them in the authors' laboratory.

Do not use the first and the last slot of a gel, since the corresponding lanes easily run off the membrane due to imprecise membrane alignment prior to the run.

4. Start electrophoresis with the same parameters used for prerunning. After 45 to 50 min, start the conveyor belt with an initial speed of 16 cm/hr, linearly decreasing to 10 cm/hr.

The continuous decrease in conveyor belt speed (i.e., in the blotting membrane feed rate) compensates for the nonlinear mobility of differently sized DNA molecules. The chosen parameters yield an approximate equidistant spacing of bands of different size (e.g., the

distance between a 100- and a 150-bp band is roughly the same as the distance between a 400- and a 450-bp band).

At the end of the run, the conveyor belt and membrane are wound up to the back roller. The membrane can be left wound up for drying overnight. Alternatively, it can be removed and hung up in a dust-free space.

If a size marker is desired, biotinylated Sequamark 10-bp ladder (Research Genetics) turns out to be optimal. This marker provides an accurate and easily identifiable standard for DNA fragments up to 500 bp; however, 5- to 10-fold concentration of the marker by precipitation is necessary to obtain sufficient sensitivity.

5. Air-dry the membrane overnight and fix by gentle UV irradiation in a Stratalinker with a UV dosage of $\sim 10,000 \mu\text{J}/\text{cm}^2$ (i.e., $\sim 1/10$ the “auto-cross-link” dosage).

For later recovery of bands of interest, it is important not to overfix membranes.

Rinse membrane and block nonspecific binding

6. Insert membrane into a suitable developing drum (e.g., GATC tube), fix with some adhesive tape, and rinse with 100 ml water while rotating 5 min on a suitable rolling incubator.

Any roller that accepts a tube 18 cm in diameter \times 35 cm long and is able to revolve at ~ 20 rpm will do.

7. Replace water with 150 ml maleic buffer, pH 7.5, and equilibrate membrane by rotating another 5 min. Pour buffer into a beaker and store for later use.
8. Incubate 40 to 50 min in a rolling incubator with 80 ml of 1.5% blocking reagent.

Label bands with streptavidin-alkaline phosphatase

9. Discard buffer and add 20 ml of 1.5% blocking reagent and 2 to 4 μl streptavidin-alkaline phosphatase conjugate. Incubate membrane 30 min in a rolling incubator.
10. Pour off buffer completely and wash 5 min, using the 150 ml maleic buffer set aside in step 7. Replace with 150 ml fresh maleic buffer and wash 10 min. Replace with another 150 ml maleic buffer and wash 15 min.
11. Replace with 150 ml reaction buffer, pH 9.5, and equilibrate membrane 5 min.

Develop color

12. For color development, pour off buffer and add 20 ml reaction buffer containing 400 μl NBT/BCIP stock solution. Develop under slow rotation for 2 to 3 hr.

CAUTION: NBT is a suspected carcinogen. Moreover, the DMSO in the concentrated stock solution might mediate penetrance of dissolved substances through the skin, and is itself hazardous. Wear gloves, replace contaminated gloves immediately, and carefully avoid any skin contact. Dispose of according to institutional regulations (also see APPENDIX 1H).

13. Pour off developing solution and perform three 10-min rinses with 150 ml water each.
14. Put the wet membrane between two sheets of 2-mm-thick polyethylene wrap or material from a thick hybridization bag. Inspect wet membranes visually for bands appearing significantly stronger or weaker in one lane as compared to adjacent corresponding lanes.

Polyethylene wrap is also called “tubular film” and must be thick, as thinner material makes handling of the wrapped membranes much more difficult and might not be a sufficient barrier against water vapor, allowing the membranes to dry. The material from a hybridization bag should also work, provided it is thick enough.

It is important that, after color development, membrane pieces carrying DNA to be reamplified never dry, as otherwise reamplification by PCR may become impossible.

For documentation, scanning of the wrapped wet membranes has proved to yield the most satisfactory results.

Isolate and reamplify sample band

15. Cut out “differential” bands with a scalpel and transfer to microcentrifuge tubes each containing 20 μl TE buffer, pH 8.0. Make sure that membrane pieces do not become dry during this procedure. Using the scalpel tip, immediately submerge bands in the buffer. Rinse scalpel carefully before excising the next band.

If excision is not intended to occur immediately, wet membranes can be stored 1 to 2 days at 4°C; however, during prolonged storage, wet membranes tend to become blotched. It is therefore advisable to dry membranes after at most one week. To avoid fading after drying, membranes should be kept dark (indefinitely) at room temperature.

16. For band reamplification, transfer half of the respective piece of membrane into a PCR tube containing 30 μl of the following mixture:

4.0 μl buffer from the tube in which the band was stored
3.0 μl 10 \times PCR buffer
2.25 μl 20 mM MgCl_2
0.6 μl 10 mM dNTPs
13.85 μl H_2O
3.0 μl 4 μM CP28
3.0 μl 4 μM ML18
0.3 μl 5 U/ μl *Taq* DNA polymerase.

17. Amplify under the following conditions:

Initial step:	1 min	94°C (denaturation)
20 or 25 cycles:	20 sec	94°C (denaturation)
	20 sec	65°C (annealing)
	2 min	72°C (extension)
Final step:	indefinitely	10°C (hold).

Amplification takes place for 20 cycles (strong bands) or 25 cycles (weak bands), respectively.

Do not use biotinylated primers for band reamplification. 5'-modification of oligonucleotide primers will interfere with cloning.

18. Check products by agarose gel electrophoresis (UNIT 2.5A).

Clone products

19. Clone reamplification products as described in UNIT 15.7 or by using one of the commercially available T-A cloning systems.

In the authors' laboratory, 4 to 5 clones per band are usually sequenced. Depending on band intensity, all clones may be identical, or there may be more than one sort of insert. In the latter case, choose the most frequently occurring insert for further processing.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Blocking reagent, 1.5%

Prepare stock solution by suspending blocking reagent (Roche Molecular Biochemicals) to 10% (w/v) in maleic buffer, pH 7.5 (see recipe) and autoclaving. Store frozen up to 1 year at -20°C . Immediately before use, dilute 1.5 parts (v/v) of the 10% stock with 8.5 parts (v/v) of maleic buffer.

Maleic buffer, pH 7.5

100 mM maleic acid

150 mM NaCl

200 mM NaOH

Store indefinitely at room temperature

MboI-linker ML2025

Combine:

150 μl 100 pmol/ μl ML20: 5'-TCA CAT GCT AAG TCT CGC GA-3' (see UNIT 2.11)

150 μl 100 pmol/ μl LM25: 5'-GAT CTC GCG AGA CTT AGC ATG TGA C-3' (see UNIT 2.11)

55 μl 10 \times ligation buffer

195 μl H₂O.

Mix and place in a 90°C heating block. Shut off the heating block and let cool down slowly to room temperature. The linker (~ 0.5 $\mu\text{g}/\mu\text{l}$) is now ready for use and should be stored frozen up to 1 to 2 years at -20°C .

Alternatively, a thermocycler programmed to a low cooling rate (e.g., $0.02^{\circ}\text{C}/\text{sec}$) can be used as opposed to a heating block.

PCR buffer, 10 \times

670 mM Tris·Cl, pH 8.8 (APPENDIX 2)

170 mM (NH₄)₂SO₄

1% (v/v) Tween 20

Store up to 2 years at -20°C

PEG 8000, 50%

Add exactly 10 g of PEG 8000 (Promega) to 10 g water in a 50-ml conical tube (e.g., Becton Dickinson). Close the tube and attach to the rotor of a hybridization oven with the heat turned off. Rotate at room temperature 12 hr to overnight until all flakes are completely dissolved. Store up to 1 to 2 years at -20°C .

After thawing, shake vigorously until no more "schlieren" can be observed. Wait ~ 10 to 15 min until all air bubbles introduced by shaking have come to the surface before slowly and carefully withdrawing the desired volume.

It is important to adhere to the exact 1:1 weight ratio of PEG and water.

PEG 8000, 28%/MgCl₂, 3.6 mM

Carefully mix 5.6 ml 50% PEG 8000 (see recipe) with 3.68 ml water and 720 μl of 50 mM MgCl₂ (APPENDIX 2). Store up to 2 years at -20°C .

Reaction buffer

100 mM NaCl

5 mM Tris hydrochloride

90 mM Tris base

Store up to 1 year at room temperature

COMMENTARY

Background Information

Identification of differentially expressed genes is currently one of the most promising approaches toward understanding fundamental life processes. However, due to the high complexity of mRNA composition in a living cell, as well as the broad range of relative frequencies of particular transcripts and the fact that subtle changes in the expression level of a gene can have profound biological effects, performing a sensitive, reliable, and relatively complete comparative expression analysis has remained a nontrivial task up to the present.

Probably the first methods for isolation of differentially expressed genes that found widespread acceptance were the fingerprinting techniques of differential display (e.g., Liang and Pardee, 1992; see also *UNIT 25.B3*) and RNA arbitrarily primed PCR (Welsh et al., 1992). These methods relied on the generation of arbitrarily primed amplification products, each representing a particular transcript, which were radiolabeled and separated by polyacrylamide gel electrophoresis. Resulting band patterns originating from different samples were then compared. An indisputable strength of display technology, as opposed to subtractive hybridization experiments (*UNIT 25.B2*), is the option to directly compare any desired number of different samples with each other. Moreover, no prior knowledge about the RNA to be analyzed is required, rendering these methods suitable for analysis of RNA from any source. Nevertheless, in some hands, the application of these protocols was not always satisfactory (Debouck, 1995), due to insufficient reproducibility (Malhotra et al., 1998), a high rate of isolating false positive clones (Poirier et al., 1997), a biased representation favoring abundant transcripts (Ledakis et al., 1998), and contamination of workspaces through closed tube walls by volatile sulfur compounds (Trentmann, 1995). The use of longer primers (i.e., 20-mers; Zhao et al., 1995) improved reproducibility, but not other problems.

To address these issues, arbitrarily primed PCR was replaced by amplification of linker ligated restriction fragments (Fischer, 1995; Fischer et al., 1995; Prashar and Weissman, 1996). With this approach, it is possible to generate and display exactly one fragment per cDNA, thereby clearly increasing the sensitivity of the analysis. Spiking experiments demonstrated that, following the RMDD protocol as described above, an mRNA species at a

relative concentration of 1:100,000 will usually be identifiable by a specific band. This holds true for the radioactive as well as for the non-radioactive version of the protocol—i.e., the authors could not detect any differences in the sensitivity of RMDD regardless of whether biotin or ^{33}P was used for labeling, which is due to the fact that sensitivity is not limited by the amount of amplification product used for display, but by a slight background smear which cannot be avoided when separating complex mixtures of PCR products by gel electrophoresis.

It is important to note that, due to the use of nonphosphorylated linkers, only one of the two linker strands is covalently attached to the cDNA restriction fragments upon ligation. The opposite linker strand is melted off during the initial denaturation step and can no longer serve as a primer binding site. Thus, amplification can take place only when extension of a non-linker primer (i.e., the “downstream” primer which is essentially identical to the cDNA primer) has taken place, incorporating the reverse complement of the covalently attached linker strand. As a consequence, only cDNA 3'-ends are amplified to a detectable level, whereas “internal” cDNA fragments flanked by linkers at both ends remain unamplified.

Another problem that had to be solved was band identification. “Classic” protocols rely on cutting out invisible radioactive bands from dried gels after superimposing the gel and its corresponding autoradiogram (Liang and Pardee, 1992). In addition to the uncertainty of cutting invisible bands, which may easily lead to missing the desired band, tiny splinters of the radioactive gel, which becomes quite brittle after drying, might be inhaled. On the other hand, nonradioactive in-gel detection of DNA by silver staining turned out to lack sufficient sensitivity, and also significantly reduced the dynamic range of display patterns (A. Fischer, unpub. observ.). Attempts to bypass the physical fragment isolation step by defining fragment signatures and performing database searches after fluorescent gel display on an automatic DNA sequencer (A. Fischer, unpub. observ.; Shimkets et al., 1999; Sutcliffe et al., 2000) are hampered by the unpredictable influence of base composition on the electrophoretic mobility of a DNA strand, which introduces considerable inaccuracies when fragment sizes are to be determined, and are unsuitable for organisms less well characterized molecularly

Table 25B.4.1 Troubleshooting Guide for RMDD

Problem	Possible cause	Solution
Low amount of first-round PCR product	RNase contamination	Take care to use only RNase-free solutions. Make sure RNA is not contaminated by remaining traces of RNase. Check integrity of ribosomal bands after cDNA first-strand synthesis.
	RNA preparation contaminated by inhibitors of cDNA synthesis	Use only RNA that is as pure as possible. Usually, standard purification protocols (e.g., the “classic” guanidinium method, <i>UNIT 4.2</i> , or more modern, commercially available RNA purification columns), if not overloaded, yield RNA of sufficient purity. Should problems persist, in very tenacious cases purification of RNA by CsCl density gradient centrifugation (<i>UNIT 4.2</i>) might be considered.
	Incomplete PEG precipitation	Be sure to exactly balance the amounts of DNA solution and of PEG solution.
	Inefficient ligation	Check activity of ligase or use a fresh batch. Make sure linkers fit to the fragment ends generated by the employed restriction enzyme.
Agarose gel appearance of first-round PCR products obtained with identical primer combinations between samples	Very low amounts of template DNA lead to stochastic effects in early PCR cycles (“Monte Carlo effect”; Karrer et al., 1995)	See “low amount of PCR product”
Fuzzy bands on DBE membrane	Glass plates accumulated too much silane	Immerse glass plates for 1 hr in 0.5 M NaOH
	Edges of glass plates not exactly parallel	Make sure plates are carefully aligned immediately after pouring gel
Low signal intensity after color development	Biotin label of blotted DNA not sufficiently accessible	Use biotinylated PCR primers distinguished by a TEG spacer
	Insufficient amounts of second-round PCR primers	Check primer concentration. Since primers are used at limiting concentration, inaccuracies upon determination of concentration may hamper generation of sufficient PCR product.
White vertical stripes interrupt band pattern on membrane	Air bubbles accumulated at the lower edge of the gel	Degas running buffer for lower chamber by stirring 20 min under vacuum. Insert glass plates slightly inclined.
Band reamplification fails	UV fixation too strong	Apply $\sim 1/10$ the UV dose usually chosen for fixing DNA blots (recommended dose is 10,000 $\mu\text{J}/\text{cm}^2$)
	Membrane has become dry before reamplification	Keep wet membrane between two sheets of thick polyethylene wrap until bands are cut out. After cutting out bands, immediately submerge in buffer.
False positive clones (no regulation detectable)	Reamplification product contained more than one DNA species	Sequence more than one clone per band. If several inserts are identified, choose the most frequently occurring one.

than man or mouse. Attempts to use such an approach for analysis of rat RNA resulted in an unacceptably low hit rate (i.e., <10%) of correctly identified fragments (A. Fischer, unpub. observ.). Thus, although the RMDD technique could be very well performed using ^{32}P or ^{33}P as a label, the authors developed a protocol using nonradioactive detection of biotinylated amplification products. These are transferred to a membrane by use of direct blotting electrophoresis (DBE; Beck and Pohl, 1984); visible bands are rendered directly accessible by simply cutting them out of the stained membrane, eliminating one of the most common sources of false positives. Besides providing convenient access to “differential” bands, DBE proved superior in terms of resolution power, yielding nearly equally spaced bands in the range between 100 and 1000 base pairs.

After generating a library of cDNA-derived restriction fragments ligated to linkers, sufficient resolution and sensitivity of detection for the thousands of 3'-cDNA fragments generated has to be achieved. Towards this end, RMDD primers elongated at their 3'-ends are used for amplification, each only allowing the amplification of a defined subset of fragments. Using extensions of two nucleotides on each side, 16 linker primers and 12 reverse primers (the first extension nucleotide by definition cannot be a T) are synthesized. Thus, during the subsequent PCR step, the original set of fragments is divided into $16 \times 12 = 192$ subsets, which fits exactly in two 96-well PCR plates and renders the method suitable for automation. Each of these subsets then contains an estimated 50 to 100 fragments, which can be easily resolved by denaturing polyacrylamide electrophoresis.

Critical Parameters

During efforts to identify critical steps contributing to the robustness of the RMDD protocol, the amount of amplifiable material left after enzymatic processing of input RNA was identified as a major factor causing instability of band patterns, probably due to random fluctuations during early PCR cycles (Karrer et al., 1995). The authors' RMDD protocol was therefore optimized to minimize losses during sample preparation. Toward this end, an important step was replacing spin-column chromatography for removal of unincorporated linker molecules with size-selective polyethylene glycol precipitation, allowing almost 100% recovery of the desired DNA species and reduction of the protocol to as few steps as possible. In its

current version, RMDD yields highly reproducible band patterns, independent of moderate variations of the amount of input material, in the range of at least down to 10 μg total RNA; however, it is essential that the analyzed RNA be of high purity. Otherwise, due to inhibition of enzymatic steps, the amount of linker-ligated template fragments effectively entering amplification might become too low to guarantee stable PCR results.

One should also be aware that RMDD analysis only covers those transcripts that carry a recognition site for the restriction enzyme used in an “amplifiable” distance from the poly(A) tail. For a more detailed discussion of this issue, see Strategic Planning.

Troubleshooting

For solutions to problems that may arise during these protocols, see Table 25.B4.1.

Anticipated Results

After corresponding PCRs (distinguished by identical primer extensions) from different RNA samples are run side by side on the gel, resulting band patterns are visually compared. A typical RMDD pattern shows, in each lane, bands of different sizes and intensities, each representing one particular cDNA. Patterns obtained from similar samples should very closely resemble each other, with only very few (if any) differences. Within these patterns, band intensities correlate with the original relative frequencies of the template cDNAs. This is due to the fact that in complex PCR reactions (i.e., with more than one amplification product) entry into plateau phase of amplification freezes the different amounts of synthesized products (McClelland and Welsh, 1994); therefore, if a particular cDNA is present at different amounts in the two samples, the resulting bands will show different intensities on the RMDD membrane. Differences in expression levels at least down to 2-fold will be detectable. In one instance, a band that was shown by quantitative PCR to represent a gene down-regulated 1.4-fold was isolated in the authors' laboratory. This is especially significant as the “gold standard” in transcription profiling is usually set at 2-fold up- or down-regulation; therefore, the fact that RMDD allows isolation of transcripts regulated to a lower degree, which still can be of the highest biological relevance (e.g., gradients in developmental biology), clearly contributes to its usefulness.

Time Considerations

When starting with up to six samples of precipitated RNA, the protocol, including second-round amplification with a subset of all primer combinations, can be performed within two days, including an overnight ligation step. The remaining set of second-round amplifications can be done at a rate of four to six 96-well plates per day and person. Alternatively, employment of a robotic pipetting station might be considered. Choosing the DBE variant, two membranes per day per DBE machine can be prepared, each providing space for 30 reactions. It should be noted that buffer capacity allows for using each DBE gel twice, provided that the second run starts immediately after the first run without idle electrophoresis in between. One person can then operate three to four machines per day and produce 6 to 8 membranes. In such a medium-scale setup, gels are prepared in the evening, and, with edges carefully wrapped in plastic wrap with some wetted pieces of paper towel enclosed, allowed to polymerize overnight. In the morning, gels are mounted and electrophoresis is started. During electrophoresis, the membranes of the day before are developed, the glass plates of the previous runs cleaned, and the gels for the next day are prepared.

Literature Cited

- Beck, S. and Pohl, F.M. 1984. DNA sequencing with direct blotting electrophoresis. *EMBO J.* 3:2905-2909.
- Debouck, C. 1995. Differential display or differential dismay? *Curr. Opin. Biotechnol.* 6:597-599.
- Fischer, A. 1995. Verfahren zur Genexpressionsanalyse. German patent application DE 195 18 505.6 [other members of the same patent family are given in the introduction].
- Fischer, A., Saedler, H., and Theissen, G. 1995. Restriction fragment length polymorphism-coupled domain-directed differential display: A highly efficient technique for expression analysis of multigene families. *Proc. Natl. Acad. Sci. U.S.A.* 92:5331-5335.
- Karrer, E.E., Lincoln, J.E., Hogenhout, S., Bennett, A.B., Bostock, R.M., Martineau, B., Lucas, W.J., Gilchrist, D.G., and Alexander, D. 1995. In situ isolation of mRNA from individual plant cells: Creation of cell-specific cDNA libraries. *Proc. Natl. Acad. Sci. U.S.A.* 92:3814-3818.
- Ledakis, P., Tanimura, H., and Fojo, T. 1998. Limitations of differential display. *Biochem. Biophys. Res. Commun.* 251:653-656.
- Liang, P. and Pardee, A.B. 1992. Differential display of eucaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257:967-971.
- Malhotra, K., Foltz, L., Mahoney, W.C., and Schueler, P.A. 1998. Interaction and effect of annealing temperature on primers used in differential display RT-PCR. *Nucl. Acids Res.* 26:854-856.
- McClelland, M. and Welsh, J. 1994. RNA fingerprinting by arbitrarily primed PCR. *PCR Methods Appl.* 4:S66-S81.
- Poirier, G.M., Pyati, J., Wan, J.S., and Erlander, M.G. 1997. Screening differentially expressed cDNA clones obtained by differential display using amplified RNA. *Nucl. Acids Res.* 25:913-914.
- Prashar, Y. and Weissman, S.M. 1996. Analysis of differential gene expression by display of 3' end restriction fragments of cDNAs. *Proc. Natl. Acad. Sci. U.S.A.* 93:659-663.
- Shimkets, R.A., Lowe, D.G., Tai, J.T., Sehl, P., Jin, H., Yang, R., Predki, P.F., Rothberg, B.E., Murtha, M.T., Roth, M.E., Shenoy, S.G., Windemuth, A., Simpson, J.W., Simons, J.F., Daley, M.P., Gold, S.A., McKenna, M.P., Hillan, K., Went, G.T., and Rothberg, J.M. 1999. Gene expression analysis by transcript profiling coupled to a gene database query. *Nature Biotechnol.* 17:798-803.
- Sutcliffe, J.G., Foye, P.E., Erlander, M.G., Hilbush, B.S., Bodzin, L.J., Durham, J.T., and Hassle, K.W. 2000. TOGA: An automated parsing technology for analyzing expression of nearly all genes. *Proc. Natl. Acad. Sci. U.S.A.* 97:1976-1981.
- Trentmann, S.M. 1995. Alternatives to ³⁵S as a label for the differential display of eucaryotic messenger RNA. *Science* 267:1186.
- Welsh, J., Chada, K., Dalal, S.S., Cheng, R., Ralph, D., and McClelland, M. 1992. Arbitrarily primed PCR fingerprinting of RNA. *Nucl. Acids Res.* 20:4965-4970.
- Zhao, S., Ooi, S.L., and Pardee, A.B. 1995. New primer strategy improves precision of differential display. *Biotechniques* 18:842-846, 848, 850.

Contributed by Achim Fischer
F. Hoffmann-La Roche AG
Basel, Switzerland

AFLP-Based Transcript Profiling

UNIT 25B.5

BASIC
PROTOCOL

In recent years, several techniques have been developed to analyze the transcriptome—i.e., the entirety of transcripts present in a cell, tissue, or organ. These procedures include methods based on hybridization to microarrays of known expressed sequence tag (EST)-sequences (Schena et al., 1995; De Risi et al., 1997), sequence-based approaches like SAGE (Velculescu et al., 1995; UNIT 25B.6) and random EST sequencing (Adams et al., 1991), and protocols based on display of cDNA fragment patterns on high-resolution gels (Liang and Pardee, 1992; UNITS 25B.3 & 25B.4). In the last category is transcript profiling based on amplified fragment length polymorphism (AFLP)-fingerprinting of double-stranded cDNA (Zabeau and Vos, 1993; Vos et al., 1995; Bachem et al., 1996). The protocol, illustrated in Figures 25B.5.1 and 25B.5.2, includes the following steps: (1) the isolation of poly(A)⁺ RNA from total RNA (UNIT 4.2), (2) the synthesis of a double-stranded (ds) cDNA library, (3) the preparation of template fragments by digestion of the cDNA library with a combination of two restriction enzymes and the ligation of adapters to the fragment ends, (4) the selective amplification of specific subsets of fragments, and (5) the electrophoretic analysis of these amplification products on standard denaturing polyacrylamide gels. The protocol given in this unit describes all steps in the procedure, except the isolation of total RNA; however, any of the presently used methods is acceptable (e.g., UNIT 4.2). The restriction enzyme combination (EC) used in this protocol is *TaqI*-*MseI*. This EC will target the majority of the mRNAs, and both *MseI* and *TaqI* are reliable and inexpensive. Other combinations of two 4-base cutters may also work well

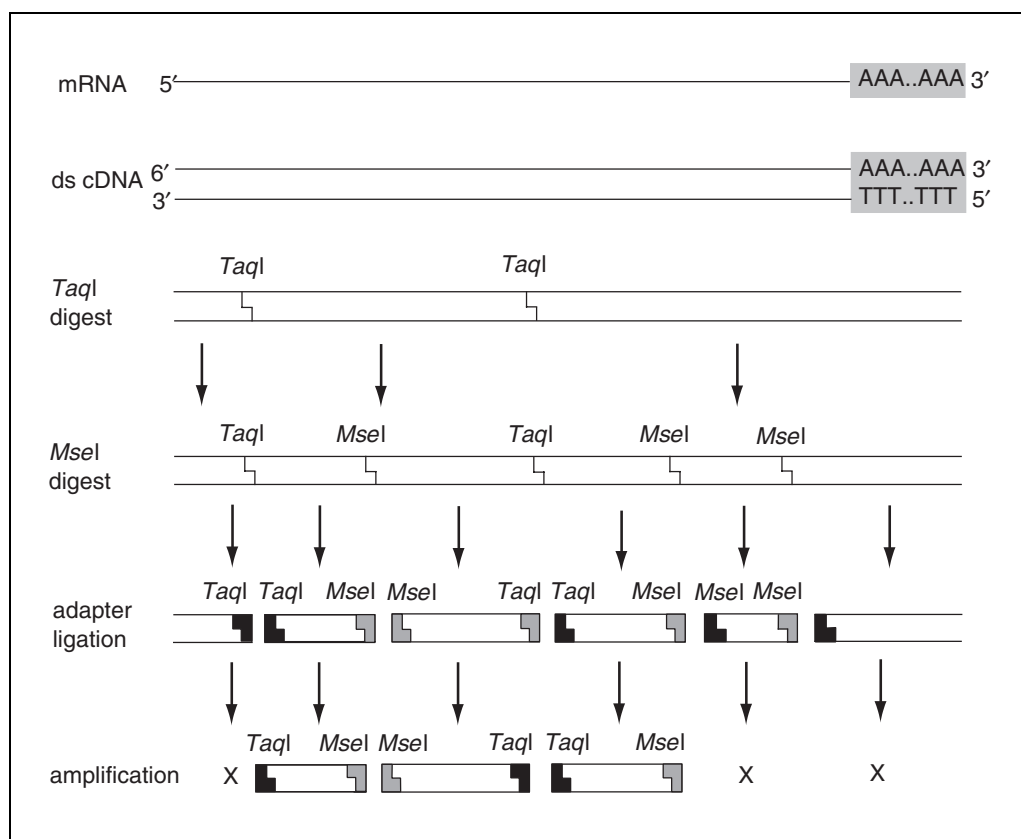


Figure 25B.5.1 Principle of the AFLP-based transcript profiling technique. The poly(A)⁺ RNA is indicated at the top with the poly(A) tail at the 3' end. The ds cDNA is shown as a double line; restriction enzyme sites with 5' overhangs are indicated. The ds *TaqI* and *MseI* adapters are depicted as small black and gray boxes respectively, attached to the protruding ends of the restriction fragments. At the bottom the "X's" illustrate the poor amplification of the *MseI*-*MseI* fragments.

Contributed by Pieter Vos and Patrick Stanssens

Current Protocols in Molecular Biology (2002) 25B.5.1-25B.5.16

Copyright © 2002 by John Wiley & Sons, Inc.

Discovery of
Differentially
Expressed Genes

25B.5.1

Supplement 57

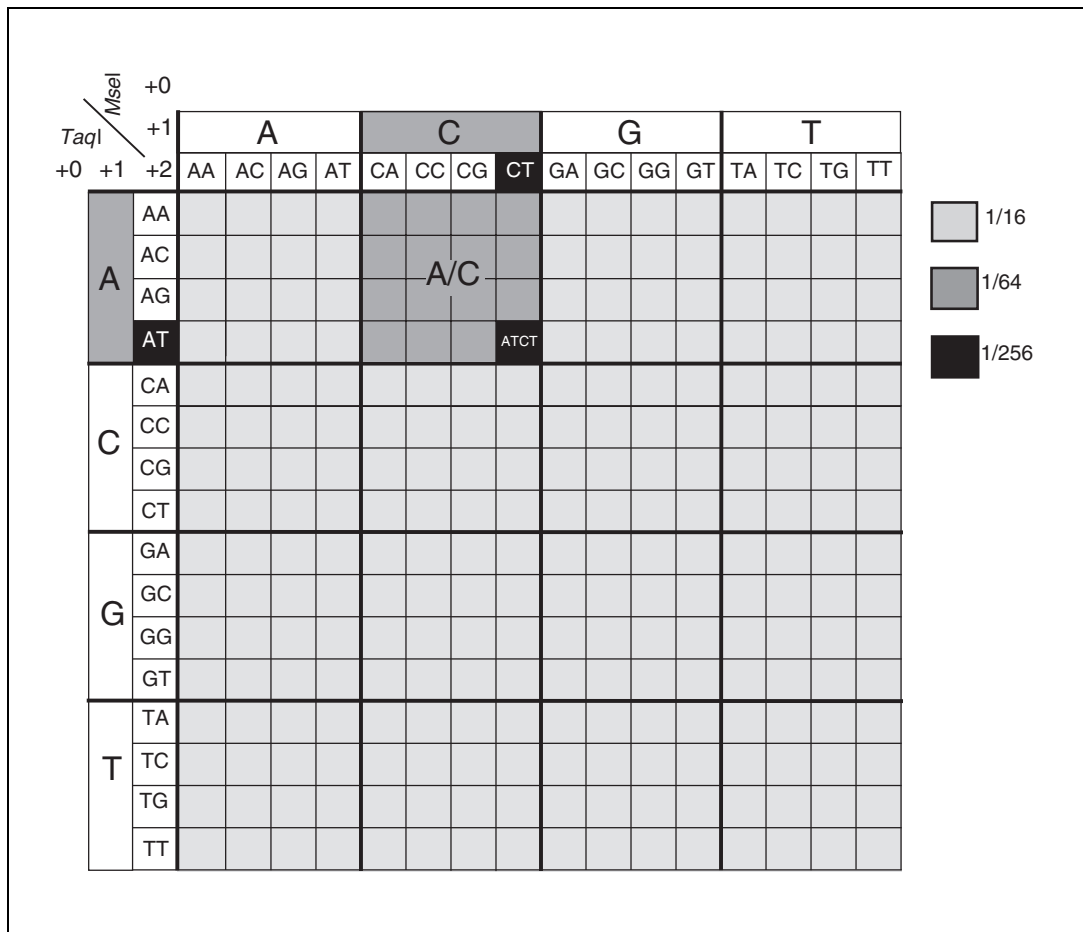


Figure 25B.5.2 Illustration of the selective amplification principle. The smallest squares indicate subsets of the transcript fragment population amplified with four selective nucleotides, two for *TaqI* and two for *MseI*, and exemplified by the black square amplified with *TaqI*-AT and *MseI*-CT. The 16 larger squares, composed of 16 of the smallest squares, indicate the transcript fragment subsets amplified using two selective nucleotides, exemplified by the dark gray square of *TaqI*-A and *MseI*-C. The total transcript fragment population is depicted by the full square, composed of 256 of the smallest squares.

but require the use of adaptors and primers that match the recognition sequences of the corresponding enzymes (see Reagents and Solutions). Restriction enzymes that cut less frequently in the cDNA are not advised since these enzymes target only a small subset of the mRNAs.

To generate specific subsets of fragments, three PCR steps are used, which minimizes mismatch amplification. When all combinations of PCR primers are used at each step, as prescribed in the protocol, this generates an expression profile consisting of 256 “fingerprints” (Fig. 25B.5.2). (One can modify the protocol to use only certain primer combinations, but this will yield fewer fingerprints and less information.) The first PCR step entails no selective nucleotides on each primer (i.e., nonselective preamplification +0/+0). The second step entails one selective nucleotide at each primer (selective preamplification +1/+1; 16 combinations). The third step entails two selective nucleotides at each primer (selective amplification +2/+2; 256 combinations).

NOTE: All solutions and materials coming into contact with RNA must be RNase free, and proper techniques should be used accordingly (see APPENDIX 2).

NOTE: AFLP is a registered trademark of Keygene N.V. and is protected by patents and patent applications of Keygene N.V.

Materials

Total RNA (*UNIT 4.2* or equivalent)
5'-biotinylated oligo-dT₂₅ (5-biotin-dT₂₅)
1× and 2× binding buffer (see recipe)
H₂O: Milli-Q purified (i.e., water deionized by passage through a five-stage Milli-Q Plus system; Millipore) or double-distilled
Streptavidin-coated magnetic beads (Dynal)
Wash buffer (see recipe)
2 mM EDTA, pH 7.5
5× first-strand buffer (see recipe)
5× second-strand buffer (see recipe)
0.1 M DTT (*APPENDIX 2*)
5 and 10 mM (each) mixture of all 4 dNTPs (Pharmacia or *UNIT 3.4*)
SuperScript II (Life Technologies)
E. coli DNA ligase (Life Technologies)
E. coli DNA polymerase I (Pharmacia Biotech)
RNase H (Pharmacia Biotech)
2× and 1× STEX (see recipe)
10 mM Tris·Cl, pH 8.0/0.1 mM EDTA (*APPENDIX 2*)
TaqI restriction endonuclease (New England Biolabs; *UNIT 3.1*)
5× RL buffer (see recipe)
MseI restriction endonuclease (New England Biolabs; *UNIT 3.1*)
50 pmol/μl *TaqI* adapter top and bottom strands (see recipe for oligonucleotides and double-stranded adapters)
50 pmol/μl *MseI* adapter top and bottom strands (see recipe for oligonucleotides and double-stranded adapters)
10 mM ATP (Pharmacia)
T4 DNA ligase (Pharmacia)
8 pmol/μl AFLP + 0 (nonselective) primers (see recipe for oligonucleotides and double-stranded adapters): *TaqI* + 0 and *MseI* + 0 primers
10× PCR buffer (see recipe)
AmpliTaq DNA polymerase (Perkin-Elmer; *UNIT 3.5*)
10 μCi/μl (~2000 Ci/mmol) [³³P-γ]ATP (Amersham)
10× T4 polynucleotide kinase buffer (see recipe)
T4 polynucleotide kinase (Pharmacia; *UNIT 3.4*)
8 pmol/μl AFLP +1 and + 2 (selective) primers (see recipe for oligonucleotides and double-stranded adapters): *TaqI* + 1 and + 2 and *MseI* + 1 and + 2 primers
AmpliTaq-Gold polymerase (Perkin-Elmer)
Loading dye (see recipe)
Repel silane (Pharmacia)
Bind silane solution, fresh: Combine 30 μl bind silane (Pharmacia Biotech) and 30 μl glacial acetic acid in 10 ml ethanol
4.5% denaturing polyacrylamide gels (see recipe)
1× TBE (see recipe)
Molecular weight standard (e.g., SequaMark 10-base ladder; Research Genetics; optional)
10% acetic acid
Microcentrifuge tubes, RNase free
Magnetic plate chamber (MPC; Dynal)
PE-9600 thermal cycler (Perkin Elmer) and PCR microtiter plate

Sequencing gel system (e.g., BioRad 38 × 50 × 0.04–cm SequiGen sequencing gel system)

PhosphorImager (Fujix BAS 2000, Molecular Dynamics STORM 824)

Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5A), analysis by denaturing polyacrylamide gel electrophoresis (UNIT 2.12), and detection of DNA by autoradiography or phosphor imaging (APPENDIX 3A)

NOTE: Suppliers and brands are generally not very critical, however in case of problems it is advised to use the suggested suppliers for at least the reverse transcriptase (SuperScript II) and *Taq* polymerases (AmpliTaq and AmpliTaq-Gold).

NOTE: When preparing AFLP amplifications, it is advisable to work with mixes of reagents as much as possible. Working with mixes facilitates assembly and is also important for the reliability and reproducibility of the reactions. In practice, the assembly of the mixes depends on the experiment—i.e., which components remain constant in a series of reactions: the template-DNA or the primer combinations (e.g., one sample with many primer combinations, many samples with one primer combination).

Isolate poly(A)⁺ RNA

1. Combine 200 μg total RNA, 600 ng 5'-biotinylated oligo-dT₂₅ (5-biotin-dT₂₅), and 300 μl 2× binding buffer in an RNase-free microcentrifuge tube. Adjust the volume to 600 μl with water. Incubate 5 min at 70°C, followed by 15 to 20 min at room temperature.

Sufficient poly(A)⁺ RNA to perform the subsequent steps (i.e., cDNA synthesis and template preparation in duplicate) is yielded from 200 μg total RNA.

2. Wash 150 μl streptavidin-coated magnetic beads with 0.5 ml of 1× binding buffer (see step 4 below for technique or use microcentrifuge). Resuspend the beads in 50 μl of the same buffer.

Mix the magnetic beads solution well before use to obtain a homogeneous suspension. Do not let the magnetic beads dry for a long period of time, as drying may lower their capacity (see Dynal, 1995).

3. Add these prewashed beads to the RNA-containing mixture (step 1) and incubate 30 min at room temperature with gentle agitation.
4. Place the microcentrifuge tube in the magnetic plate chamber (MPC) for ~30 sec and then remove as much of the supernatant as possible without disturbing the beads. Remove the tube from the MPC, add 0.5 ml wash buffer, and mix thoroughly. Repeat two more times, removing the supernatant after the final wash.

Do not allow the beads to dry out.

5. Elute poly(A)⁺ RNA by resuspending the beads in 20 μl of 2 mM EDTA and incubating 5 min at 70°C. Collect the beads with the MPC as in step 4 and transfer the supernatant to a new RNase-free microcentrifuge tube as quickly as possible without transferring any beads. Repeat once to obtain ~40 μl of poly(A)⁺ RNA solution.

For long-term storage, add 0.1 vol 2 M sodium acetate, pH 5.5 and mix. Add 3 vol 100% ethanol and store indefinitely at -20°C (UNIT 2.1A). To recover, microcentrifuge 5 min at maximum speed, remove supernatant, dry in a rotary evaporator, and resuspend in the original volume of double-distilled water or buffer.

6. Check the yield (on average ~2 μg) and quality of the isolated poly(A)⁺ RNA by performing agarose gel electrophoresis alongside molecular weight markers (UNIT

2.5A) using 5 μ l of the poly(A)⁺ RNA solution, which should appear as a faint smear from ~10 kb down (i.e., lower molecular weight) with trace rRNA bands.

It is not necessary to eliminate the rRNA contamination by extracting the mRNA from the eluate a second time.

Synthesize the ds cDNA

7. For first-strand cDNA synthesis, combine the following:

- 10 μ l poly(A)⁺ RNA (~0.5 μ g)
- 0.5 μ l 700 ng/ μ l 5-biotin-dT₂₅ (reverse transcription primer)
- 2 μ l H₂O
- 4 μ l of 5 \times first-strand buffer
- 2 μ l 0.1 M DTT
- 1 μ l 10 mM dNTPs
- 0.5 μ l of 200 U/ μ l SuperScript II (add last).

Incubate 2 hr at 42°C.

8. For second-strand synthesis, combine the following:

- 20 μ l first-strand cDNA synthesis mixture (from step 7)
- 16 μ l 5 \times second-strand buffer
- 1.5 μ l 10 mM dNTPs
- 3 μ l 0.1 M DTT
- 7.5 U *E. coli* DNA ligase
- 25 U *E. coli* DNA polymerase I
- 0.8 U RNase H
- H₂O to 80 μ l.

Incubate 1 hr at 12°C followed by 1 hr at 22°C.

The quality and yield of the resulting ds cDNA can be checked by agarose gel electrophoresis (UNIT 2.5A).

9. Wash 25 μ l streptavidin-coated beads with 100 μ l of 2 \times STEEX (see step 4 for technique). Resuspend in 80 μ l of 2 \times STEEX.

10. Add the bead suspension to the cDNA mixture and incubate 30 min at room temperature with gentle agitation.

Purification of a large number of samples using beads can be performed in 96-well format. Incubation at room temperature is done in 96-well plates with caps. Subsequently, samples are transferred to fresh microtiter plates.

11. Collect beads with the MPC (step 4), wash once with 100 μ l of 1 \times STEEX, and transfer to a fresh microcentrifuge tube. Wash twice more with 1 \times STEEX and resuspend final bead pellet in 50 μ l H₂O or 10 mM Tris-Cl, pH 8.0/0.1 mM EDTA.

Generally, 250 to 500 ng ds cDNA will be obtained from the 500 ng of input (single-stranded) poly(A)⁺ RNA.

10 mM Tris-Cl, pH 8.0/0.1 mM EDTA is also known as T₁₀E_{0.1} buffer and has a lower EDTA concentration than the TE buffer described in APPENDIX 2 of this manual.

The ds cDNA is attached to the beads and is taken into subsequent steps while attached to the beads.

Prepare the AFLP cDNA template fragments using *TaqI* and *MseI*

12. Mix the following:

20 μ l cDNA preparation (generally 100 to 200 ng cDNA)
10 U *TaqI* restriction endonuclease
8 μ l 5 \times RL buffer
Adjust volume to 40 μ l with H₂O.

Incubate 1 hr at 65°C.

13. Add the following:

10 U *MseI* restriction endonuclease enzyme
2 μ l 5 \times RL buffer
Adjust the volume to 50 μ l with H₂O.

Incubate 1 hr at 37°C.

14. Prepare the *TaqI* adapter by combining 8.5 μ g (1500 pmol) top and 8 μ g (1500 pmol) bottom strands. Adjust volume to 30 μ l with water.

*This results in a solution of 50 pmol/ μ l of double-stranded *TaqI*-adapter.*

15. Prepare the *MseI* adapter by combining 8.0 μ g (1500 pmol) of the top strand and 8.0 μ g (1500 pmol) of the bottom strand. Adjust volume to 30 μ l with water.

*This results in a solution of 50 pmol/ μ l of ds *MseI* adapter.*

*The *TaqI* and *MseI* adapters both have double-stranded parts of 14 base pairs; it appears unnecessary to perform a specific denaturation-renaturation procedure to anneal the two strands of the adapters. Note that the base-pair adjacent to the restriction site overhang is such that the recognition site is not restored upon ligation (see Background Information). Absence of 5'-phosphates prevents self-ligation of adapters.*

16. To the cDNA fragments digested with *TaqI* and *MseI* (steps 12 and 13), add 1 μ l of each adapter (50 pmol each; steps 14 and 15), 1 μ l of 10 mM ATP, 2 μ l of 5 \times RL-buffer, 1 U of T4 DNA ligase, and 10 μ l water. Incubate 2 hr at 37°C.

The cDNA is incubated for 2 hr with restriction enzymes (steps 12 and 13) followed by an additional incubation of 2 hr in the presence of DNA ligase. It is not advisable to perform the restriction digestion and ligation simultaneously. This may affect the efficiency of the DNA restriction. Longer incubation times are also not recommended, because this may affect the quality of the transcript fingerprints. After digestion and ligation of adapters, the cDNA is stored at -20°C or immediately used for the subsequent steps.

Perform nonselective preamplification of the template fragments

17. Dilute a small aliquot (2 to 5 μ l) of the template mixture (step 16) 10-fold with Tris·Cl, pH 8.0/0.1 mM EDTA. Prepare the following preamplification reactions:

5.0 μ l 1:10 diluted template mixture
1.5 μ l 8 pmol/ μ l each AFLP + 0 (nonselective) primer
2.0 μ l 5 mM dNTPs (0.2 mM final concentration of each dNTP)
5 μ l 10 \times PCR buffer
1 U AmpliTaq DNA polymerase
Adjust volume to 50 μ l with H₂O.

18. Amplify using the following temperature cycle profile on a PE-9600 thermal cycler:

20 cycles:	30 sec	94°C	(denaturation)
	60 sec	56°C	(annealing)
	60 sec	72°C	(extension).

The purpose of the nonselective preamplification reaction is to generate more starting material for the subsequent selective AFLP reactions. This is one of the advantages of the AFLP technique. Once the template-DNA is made, new starting material for selective amplifications can always be made by nonselective amplification of the template DNA, and hence, new RNA isolation will never have to be done again.

Note that the adapter strands are not phosphorylated and that, therefore, the strand which represents the primer-target is not ligated to the template DNA. Thus, a “hot-start” should never be performed (UNIT 15.1); however, during the initial heating step, Taq polymerase should elongate the staggered ends of the template replacing the adapter strands.

19. Check the preamplification by running 10 μ l of the reaction mixture on an agarose gel alongside molecular weight markers (UNIT 2.5A), which should give a visible smear of products in the size range of 50 to 500 base pairs.

Perform selective preamplification reactions using TaqI+I and MseI primers

20. Dilute 2 μ l nonselective preamplified cDNA fragments (i.e., +0/+0) 1:500 in Tris·Cl, pH 8.0/0.1 mM EDTA.
21. Prepare selective preamplification (+1/+1) reactions in a microtiter plate for a PE-9600 thermocycler in the following way:

- a. Dispense 5 μ l of 1:500 nonselective preamplification cDNA fragments into each well of the first two columns (1 and 2) of the microtiter plate.
- b. Dispense 1.5 μ l of 8 pmol/ μ l TaqI+A primer into each of wells A1 to D1, 1.5 μ l of 8 pmol/ μ l TaqI+C primer into each of wells E1 to H1, 1.5 μ l of 8 pmol/ μ l TaqI+G primer into each of wells A2 to D2, and 1.5 μ l of 8 pmol/ μ l TaqI+T primer into each of wells E2 to H2.
- c. Dispense 1.5 μ l of 8 pmol/ μ l MseI+A primer into each of wells A1, A2, E1, and E2; 1.5 μ l of 8 pmol/ μ l MseI+C primer into each of wells B1, B2, F1, and F2; 1.65 μ l of 8 pmol/ μ l MseI+G primer into wells C1, C2, G1, and G2; and 1.4 μ l of 8 pmol/ μ l MseI+T primer into each of wells D1, D2, H1, and H2.
- d. Prepare dNTP/polymerase mix by combining 32 μ l 5 mM dNTPs, 80 μ l of 10 \times PCR buffer, 16 U AmpliTaq-Gold polymerase, and adjust the volume to 672 μ l with water.
- e. Dispense 42 μ l dNTP/polymerase mix into each well of the first two columns.

The procedure above can be adjusted when more samples are processed at the same time (i.e., 3 samples occupy 48 wells of the microtiter plate and three times more TaqI+I primers, MseI+I primers, and dNTP/polymerase mix will be needed).

An individual reaction can be prepared by combining 5 μ l of 1:500 nonselective preamplification product, 1.5 μ l of 8 pmol/ μ l TaqI+I primer (12 pmol), 1.5 μ l of 8 pmol/ μ l MseI+I primer (12 pmol), 2 μ l of 5 mM dNTP, 5 μ l 10 \times PCR buffer, and 1 U AmpliTaq-Gold polymerase. The volume is adjusted to 50 μ l with water.

22. Perform AFLP amplification with the following “touch-down” temperature cycle program:

13 cycles:	30 sec	94°C	(denaturation)
	30 sec	65°–0.7°C/cycle	(annealing)
	60 sec	72°C	(extension)
23 cycles:	30 sec	94°C	(denaturation)
	30 sec	56°C	(annealing)
	60 sec	72°C	(extension).

The initial annealing step is performed at 65°C, decreasing by 0.7°C each cycle.

A stepwise amplification procedure is used to minimize mismatch amplification. A single additional selective nucleotide (one on each primer) is added per selective AFLP amplification. The most useful expression profiles consist of the 256 fingerprints obtained with all combinations of the *TaqI*+2 and *MseI*+2 primers. This implicates a series of 3 consecutive PCRs, the first with no selective nucleotides (nonselective preamplification +0/+0), the second with one selective nucleotide at both the *TaqI* and *MseI* primer (selective preamplification +1/+1), and the third with two selective nucleotides at each primer (final selective amplification +2/+2).

Similar to the nonselective preamplification, it is advisable to check 10 μl of the reaction mixtures on an agarose gel.

Label selective *TaqI* + 2 primers

23. Prepare the following phosphorylation reaction mixture:

- 2.0 μl 10 $\mu\text{Ci}/\mu\text{l}$ (~2000 Ci/mmol) [^{33}P - γ]ATP
- 1.0 μl 10 \times T4 polynucleotide kinase buffer
- 4 U T4 polynucleotide kinase
- Adjust volume to 8 μl with water.

24. To phosphorylate 16 pmol selective *TaqI*+2 primer (the amount required for 20 AFLP reactions; i.e., the amount required to perform all 16 +2/+2 reactions for a given *TaqI*+2 primer in the complete set of 256 primer combinations), combine 2.0 μl of 8 pmol/ μl selective primer (+2) and 8.0 μl phosphorylation reaction mix (step 23), yielding labeled primer at a concentration of 12.6 pmol/ μl and a final volume of 10.0 μl . Incubate 60 min at 37°C, followed by 10 min at 70°C to inactivate the kinase.

^{33}P -labeled primers are preferred because they give a better resolution of the PCR products on polyacrylamide gels. Also, the reaction products are less prone to degradation due to autoradiolysis.

*Only the *TaqI* primers should be labeled. Labeling both the *TaqI* and *MseI* primers causes each of the two strands of the AFLP fragments to be visualized on the gels, often causing "doublets" when these two strands migrate differently on the gel.*

Perform selective AFLP amplification using labeled *TaqI* + 2 and *MseI* + 2 primers

25. Dilute 2 μl of each selective preamplification product (+1/+1; step 22) 500-fold with Tris-Cl, pH 8.0/0.1 M EDTA. Prepare selective amplification (+2/+2) reactions in a microtiter plate for a PE-9600 thermocycler in the following way:

- a. Dispense 2 μl of 1:500 preamplification mixture *TaqI*+A/*MseI*+C in the first two columns of the microtiter plate.
- b. Dispense 0.5 μl labeled *TaqI*+AA primer into each of wells A1 to D1; 0.5 μl labeled *TaqI*+AC primer into each of wells E1 to H1; 0.5 μl labeled *TaqI*+AG primer into each of wells A2 to D2, and 0.5 μl labeled *TaqI*+AT primer into each of wells E2 to H2.
- c. Dispense 0.6 μl unlabeled *MseI*+CA primer into each of wells A1, A2, E1, and E2; 0.6 μl unlabeled *MseI*+CC primer into each of wells B1, B2, F1, and F2; 0.6 μl unlabeled *MseI*+CA primer into each of wells A1, A2, E1, and E2; 0.6 μl unlabeled *MseI*+CG primer into each of wells C1, C2, G1, and G2; and 0.6 μl unlabeled *MseI*+CT primer into each of wells D1, D2, H1, and H2.
- d. Prepare dNTP/polymerase mixture by combining 12.8 μl of 5 mM dNTPs, 32 μl of 10 \times PCR buffer, 6.4 U AmpliTaq-Gold polymerase, and adjusting the volume to 270.4 μl .
- e. Dispense 16.9 μl dNTP/polymerase mixture in the first two columns of the microtiter plate.

An individual reaction can be prepared by combining 2 μ l of 1:500 selective preamplification reaction product, 0.5 μ l of 1.6 pmol/ μ l (5 ng) labeled selective TaqI+2 primer, 0.6 μ l of 8 pmol/ μ l (30 ng) unlabeled selective MseI+2 primer, 0.8 μ l of 5 mM dNTPs, 2.0 μ l of 10 \times PCR buffer, and 0.4 U AmpliTaq-Gold polymerase. The volume is adjusted to 20 μ l with water.

26. Amplify the material using the “touch-down” PCR program specified in step 22.

Generally, a number of AFLP reactions will be performed in parallel and the indicated quantities of the reaction mixes should be adjusted accordingly.

Analyze amplification products by standard PAGE

27. Mix the AFLP reactions with an equal volume (20 μ l) of loading dye. Denature the AFLP reaction products by heating at 90°C for 3 min and then quickly cooling on ice.

CAUTION: Formamide is harmful—perform this step under a fume hood.

28. Treat the back plate of the sequencing gel system with 2 ml repel silane, and the front plate with 10 ml bind silane solution. Prepare 4.5% denaturing polyacrylamide gels (~100 ml).

The authors use the BioRad SequiGen sequencing gel system (38 \times 50 \times 0.04-cm), for which the parameters given in this protocol are optimized; however, other sequencing gel systems should also work well.

29. Using 1 \times TBE as the running buffer, prerun gels 0.5 hr just before loading the samples under appropriate conditions to heat the gel to ~55°C (e.g., 110-W limit for the BioRad system). Use a gel thermometer to monitor temperature.

Maintaining this temperature throughout the electrophoresis is crucial for good quality fingerprints.

30. Load either 3 μ l (for 48-lane gels) or 1.5 μ l (for 96-lane gels) of sample into each well and analyze at ~55°C. Include a molecular weight standard (e.g., SequaMark 10-base ladder) if desired.

31. After electrophoresis, disassemble the gel cassette. Fix the gel, which will stick to the front glass plate because of the silane treatments, by soaking in 10% acetic acid for 30 min. Rinse thoroughly with water and dry 10 to 20 hr at room temperature in a fume hood, or for a shorter time period at an elevated temperature (e.g., using an incubator).

CAUTION: Radioactive materials require special handling. See APPENDIX 1F and the institutional Radiation Safety Office for guidelines concerning proper handling and disposal.

Gel is dry when it is no longer “sticky.”

32. Visualize gel-fractionated cDNA AFLP fragments by autoradiography or using a phosphorimager (APPENDIX 3A).

Exposure times are reduced at least 2.5-fold using phosphorimaging technology.

REAGENTS AND SOLUTIONS

Use Milli-Q purified or double-distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Binding buffer, 2×, 1×

20 mM Tris·Cl, pH 7.5 (APPENDIX 2)

150 mM LiCl

1 mM EDTA (APPENDIX 2)

Store up to 6 months at room temperature

Dilute to 1× with Milli-Q-purified or double-distilled H₂O

Denaturing polyacrylamide gel, 4.5%

Prepare 4.5% (v/v) Sequagel ready-for-use gel mix (19:1 acrylamide/methylene bisacryl; National Diagnostics) in 7.5 M urea (Life Technologies)/0.5× TBE (see recipe) at a total volume of ~100 ml. Add 500 µl of 10% ammonium persulfate (APS), freshly made just before use, and 100 µl of TEMED (*N,N,N',N'*-tetramethylethylenediamine) immediately before casting the gel. Cast the gel according to the instructions of the gel system manufacturer, using either two 24-well (for 48-lane gels) or 48-well (for 96-lane gels) sharktooth combs to create the gel slots.

These gels are essentially normal sequencing gels (Vos and Kuiper, 1998; UNIT 7.6), with the exception that a lower percentage of polyacrylamide is used. Ready made solutions should also work well.

First-strand buffer, 5×

250 mM Tris·Cl pH 8.3 (APPENDIX 2)

15 mM MgCl₂

375 mM KCl

Store up to 6 months at -20°C

Oligonucleotides and double-stranded adapters

Adapters:

TaqI adapter top strand: 5'-CTCGTAGACTGCGTACA-3'

TaqI adapter bottom strand: 3'-CATCTGACGCATGTGC-5'

MseI adapter top strand: 5'-GACGATGAGTCCTGAG-3'

MseI adapter bottom strand: 3'-GCTACTCAGGACTCAT-5'

Nonselective primers (AFLP + 0):

TaqI + 0 primer: 5'-CTCGTAGACTGCGTACACGA-3'

MseI + 0 primer: 5'-GACGATGAGTCCTGAGTAA-3'

Selective primers (AFLP +1 and +2):

TaqI + 1 primer: 5'-GTAGACTGCGTACACGAN-3'

TaqI + 2 primer: 5'-GTAGACTGCGTACACGANN-3'

MseI + 1 primer: 5'-GATGAGTCCTGAGTAAN-3'

MseI + 2 primer: 5'-GATGAGTCCTGAGTAANN-3'

N is any nucleotide; therefore, there are a total of 1 "+ 0," 4 "+ 1," and 16 "+ 2" primers for each restriction endonuclease.

Loading dye

98% formamide, deionized and filtered (Merck)

10 mM EDTA, pH 8.0 (APPENDIX 2)

5 mM spermidine·3HCl (Sigma)

Trace amounts (i.e., ~0.5 mg/ml) of bromphenol blue and xylene cyanol

Store in small (500 µl) aliquots up to 6 months at -20°C.

PCR buffer, 10×

100 mM Tris·Cl, pH 8.3 (APPENDIX 2)
15 mM MgCl₂
500 mM KCl
Store up to 6 months at room temperature

RL buffer, 5×

50 mM Tris acetate, pH 7.5
50 mM magnesium acetate
250 mM potassium acetate
25 mM DTT
Store in small aliquots (up to 500 μl) and store up to 6 months at -20°C

Second-strand buffer, 5×

100 mM Tris·Cl, pH 7.0 (APPENDIX 2)
20 mM MgCl₂
450 mM KCl
750 μM NAD⁺
50 mM (NH₄)₂SO₄
Store in small aliquots up to 6 months at -20°C

STEX, 2×, 1×

20 mM Tris·Cl, pH 8.0 (APPENDIX 2)
2000 mM NaCl (APPENDIX 2)
2 mM EDTA (APPENDIX 2)
0.2 % (v/v) Triton X-100
Store up to 6 months at room temperature
Dilute to 1× with Milli-Q-purified or double-distilled H₂O

T4 polynucleotide kinase buffer, 10×

250 mM Tris·Cl, pH 7.5 (APPENDIX 2)
100 mM MgCl₂
50 mM DTT
Make small aliquots and store up to 6 months at -20°C

TBE, 1×

Prepare a 10× stock:
1 M Tris base
1 M boric acid
20 mM EDTA, pH 8.3
Store up to 6 months at room temperature
Dilute to 1× with water

Wash buffer

10 mM Tris·Cl, pH 7.5 (APPENDIX 2)
150 mM LiCl
1 mM EDTA
Store up to 6 months at room temperature

COMMENTARY

Background Information

At present, a variety of technologies are available for high-throughput analysis of mRNA populations in cells, tissues, and organs. These can be divided into three major classes: (1) methods based on hybridization of labeled cDNA to transcript sequences on microarrays (Skena et al., 1995; De Risi et al., 1997), (2) methods based on high-throughput sequencing of small identifier (“signature”) sequences corresponding to specific transcripts (UNIT 25B.6; Velculescu et al., 1995; Brenner et al., 2000), and (3) methods based on display of cDNA fragment patterns on high-resolution gels such as AFLP (UNITS 25B.3 & 25B.4; Liang and Pardee, 1992; the current unit).

Hybridization to microarrays of known transcript sequences is an attractive method for high-throughput transcript analysis (Skena et al., 1995; De Risi et al., 1997). The amount of data that can be obtained with this technology cannot be matched easily by any other presently known transcript analysis method. The fast growing number of gene and whole genome sequences creates a valuable resource for probe design for microarrays. One of the most attractive applications of the technology to date is the comparative analysis of gene expression between two samples for which the cDNA is differentially labeled (Welsh et al., 2001). Cross hybridization may pose a problem using microarrays, primarily because gene families are quite predominant in higher organisms; however, the use of multiple oligonucleotide probes of individual genes alleviates this problem, enabling the design of highly discriminative oligonucleotide sets (Wodicka et al., 1997).

A second category of transcript analysis technologies is represented by the SAGE technology (Serial Analysis of Gene Expression) first described by Velculescu et al. (1995; UNIT 25B.6) and the Massive Parallel Signature Sequencing (MPSS) technology first described by Brenner et al. (2000). These technologies generate small identifier or signature sequences specific for each transcript in a particular cell or tissue type, and are very well suited for transcript discovery in known genomic sequences. Gene prediction from genomic sequences is still far from perfect today, and the whole genome sequences of complex organisms suggest that the transcript repertoire may be quite complicated. The MPSS technology is commercialized by Lynx Therapeutics. SAGE

technology is described elsewhere in this book (UNIT 25B.6).

Differential display (DD) technology as first described by Liang and Pardee (1992; UNIT 25B.3) uses one random primer and an anchored oligo-d(T) primer for amplification of cDNA fragments, which are displayed on denaturing polyacrylamide gels (i.e., sequencing gels). The major difference between DD and the AFLP cDNA technology described in this unit is that AFLP cDNA profiling allows a systematic display of cDNA fragments, with each primer combination displaying a different subset of the cDNAs (Durrant et al., 2000; Van der Biezen et al., 2000; Breyne and Zabeau, 2001; Din et al., 2001; Qin et al., 2001). This, and the smaller fragments generated by AFLP, generally yield sharper and more discrete banding patterns. Another alternative to DD is restriction enzyme analysis of differentially expressed sequences. This technology makes use of restriction enzyme cleavage sites in the cDNA and yields sharp, discrete bands like AFLP (UNIT 25B.4; Fischer et al., 1995; Prashar and Weismann, 1996).

The AFLP technique allows the selective amplification of subsets of genomic restriction fragments or cDNAs, which can subsequently be displayed on DNA sequencing gels. One of the characteristics of the AFLP technique is that the reaction proceeds until the primer is depleted from the reaction mixture (Vos et al., 1995). This is different from a standard PCR, where the amplification process is inhibited in the final stage of the reaction due to competition between fragment-to-fragment, reannealing, and primer-to-template annealing. This difference is probably caused by the fact that the concentration of individual AFLP fragments is much lower compared to standard PCR due to many fragments competing for the same primer set. This characteristic of the AFLP technique is of great importance for the quantitative amplification and display of transcript fragments.

Another important characteristic of the AFLP technique is the preferential amplification of *TaqI-MseI* fragments compared to the *TaqI-TaqI* fragments and *MseI-MseI* fragments that will also result from template preparation. It is the authors' belief that the *TaqI-TaqI* fragments and *MseI-MseI* fragments amplify less efficiently because they contain inverted repeats at the fragment ends after adapter ligation. As a result, intramolecular self ligation of *TaqI-TaqI* fragments and *MseI-MseI* fragments will

Adapter (top strand) and primer sequences for <i>MseI</i> :		
	Adapter End	<i>MseI</i>
After adapter ligation	5'-GACGATGAGTCCTGAG-TAA-Internal sequence-	<i>TaqI</i> adapter-3'
<i>MseI</i> -primer+0	5'-GACGATGAGTCCTGAG-TAA-3'	
<i>MseI</i> -primer+1 (+1 = T)	5'-GACGATGAGTCCTGAG-TAA-T-3'	
<i>MseI</i> -primer+2 (+2 = C)	5'-GACGATGAGTCCTGAG-TAA-TC-3'	
Adapter (top strand) and primer sequences for <i>TaqI</i> :		
	Adapter End	<i>TaqI</i>
After adapter ligation	5'-CTCGTAGACTGCGTACA-CGA-Internal sequence-	<i>MseI</i> adapter-3'
<i>TaqI</i> primer +0	5'-CTCGTAGACTGCGTACA-CGA-3'	
<i>TaqI</i> primer +1 (+1 = A)	5'-CTCGTAGACTGCGTACA-CGA-A-3'	
<i>TaqI</i> primer +2 (+2 = G)	5'-CTCGTAGACTGCGTACA-CGA-AG-3'	

Figure 25B.5.3 Schematic of primer design.

compete with primer annealing during AFLP amplification. This hypothesis is supported by the observation that amplification of *MseI*-*MseI* fragments is efficient when two different *MseI* adapters are used for template preparation and two corresponding *MseI* primers are used for amplification (P. Vos, unpub. observ.).

In the protocol outlined in this chapter, double-stranded cDNA is restriction digested with *TaqI* and *MseI*, and adapters for these two restriction endonucleases are ligated to the resulting restriction fragments. Adapters create a target site for the AFLP primers in the subsequent amplification reactions. The adapter ligation is performed in a way that the original *TaqI* and *MseI* sites are not restored. After adapter ligation the *TaqI*-*MseI* restriction fragments have from 5' to 3' a universal sequence at the *TaqI* end (*TaqI* adapter + remnant of *TaqI* site), the original sequence between the *TaqI* and *MseI* recognition sequence, and a second (different from *TaqI*) universal sequence at the *MseI* end (*MseI* adapter + remnant of *MseI* site; Figure 25B.5.3). The primer design matches the newly created fragment ends. The use of the restriction endonuclease combination *TaqI*-*MseI* and primers containing four selective nucleotides (two selective bases for *TaqI* and two selective bases for *MseI*) divides the mixture of transcript fragments into 256 different fragment subsets. Each fragment subset will be amplified by a specific combination of *TaqI* and *MseI* primers (i.e., a primer combination), and will display a small amount (i.e., ~1/256) of the transcript fragments in a specific sample. From various experiments it is known that an AFLP fragment will be detected if at least 1/1000 part of the AFLP primer is incorporated in the AFLP

fragment (P. Vos unpub. observ.; P. Stanssens unpub. observ.); therefore, the detection sensitivity of the protocol described in this unit will generally be quite high. However, it should be noted that the detection sensitivity may vary from one primer combination to another as a result of the specific subset of transcript fragments that will be amplified within each primer combination.

In conclusion, the use of cDNA AFLP is an attractive technology for gene expression analysis and transcript discovery, particularly in organisms for which little or no sequence information is available. The technology is complementary to microarray based transcript imaging techniques that rely on prior characterization of the gene sequences.

Critical Parameters and Troubleshooting

AFLP analysis of genomic DNA is a very robust technology that has been used by numerous laboratories around the world for the past five years. Very few technical problems are generally encountered (Vos et al., 1995; Vos and Kuiper, 1998); however, the quality of the poly(A)⁺ RNA and resulting ds cDNA is critical to its success. The authors advise that the protocols for poly(A)⁺ RNA isolation from total RNA and the synthesis of ds cDNA be strictly followed.

Despite the robustness of AFLP, there are several theoretical and technical reasons why specific transcripts might not be displayed. These include (1) low transcript abundance, (2) the absence of relevant restriction enzyme sites in the transcript, and (3) features of the transcript that prevent efficient reverse transcrip-

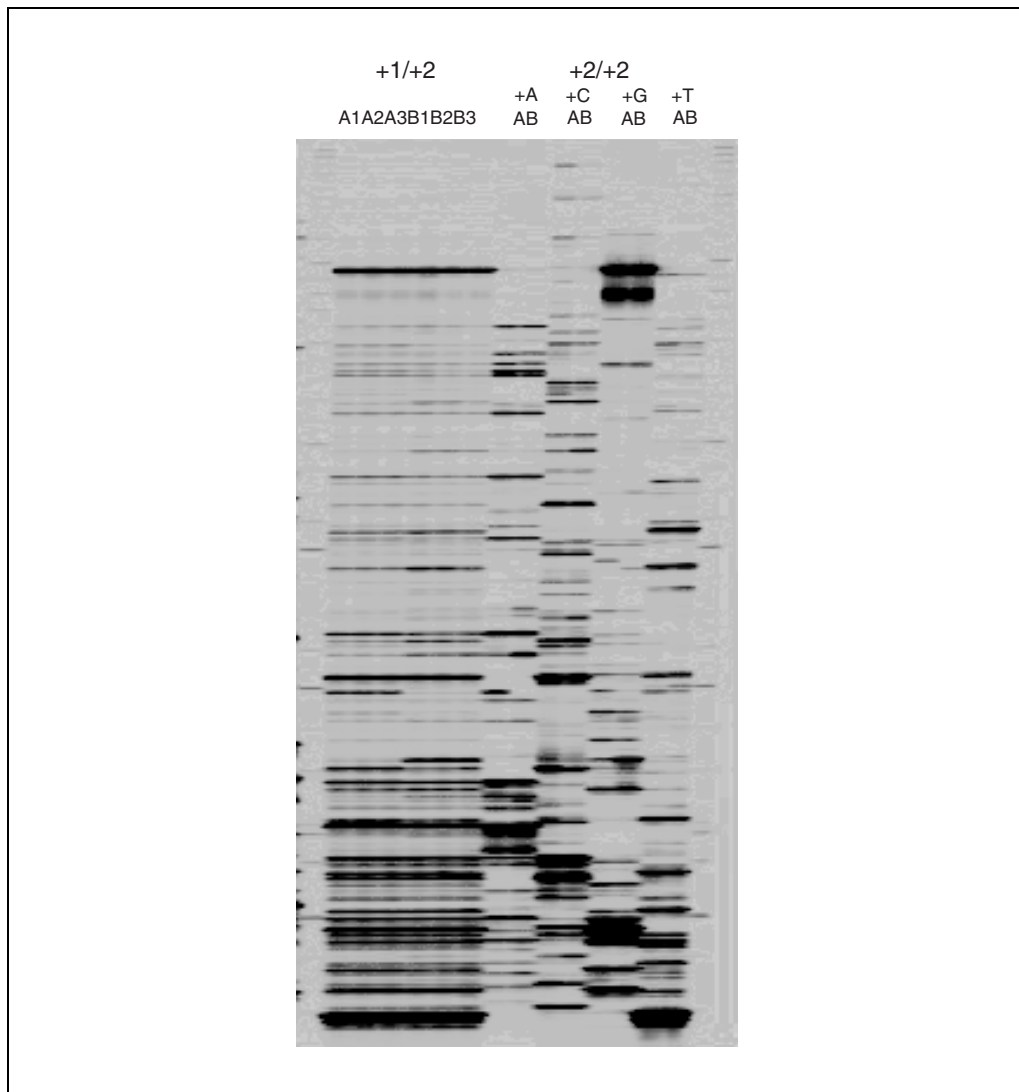


Figure 25B.5.4 cDNA fingerprint of *Aspergillus niger* that displays a very typical result for AFLP technology. Samples A1 to A3 represent three different samples which have been taken independently through the procedure of RNA isolation, cDNA synthesis, template preparation, and cDNA-AFLP reactions (notice the reproducibility). The same is true of samples B1 to B3; however, these samples were induced differently than the “A” sample sets and therefore a number of differentially expressed cDNAs are detected between the two samples. Fingerprints on the right represent +2/+2 fingerprints, and on the left corresponding +1/+2 fingerprints. The figure clearly shows that the cDNA fragments in the +2/+2 fingerprints are a subset of the cDNA fragments in the +1/+2 fingerprint.

tion (e.g., secondary structure). In the authors’ experience the major cause for this is deviation from the protocol as outlined above. Annotations to the steps highlight important considerations.

The quality of the sequence gels can simply be verified by adding a “sequence ladder” to the gel. Gels that work well for sequencing will be good for AFLP profiling as well.

Anticipated Results

All experiments carried out according to the protocol outlined above will give satisfactory

results. Typical transcript profiles show 50 to 100 cDNA AFLP fragments per lane (i.e., sample). The profiles should change completely when a different primer combination is used, with virtually none of the fragments being the same. Transcript profiles from the same individual will vary according to the tissue that is inspected and the conditions that are used (e.g., developmental stages, environmental factors, pathogenic infections). Figure 25B.5.4 displays an example of a typical experiment with the transcript profiles of various organisms.

Table 25B.5.1 Typical AFLP Experiment Time Course

Day 1	Poly(A) ⁺ RNA isolation and synthesis of ds cDNA
Day 2	AFLP cDNA template preparation and nonselective preamplification reactions
Day 3	Selective amplification, gel electrophoresis and overnight exposure of the gels to X-ray films or phosphoimaging screens
Day 4	Analysis of results

Time Considerations

Starting from total RNA, the following time considerations are expected for up to 96 samples:

1. Isolating poly(A)⁺ RNA: 1 to 2 hr depending on the number of samples.
2. Synthesizing ds cDNA from the poly(A)⁺ RNA: 5 to 6 hr.
3. Preparing the AFLP cDNA templates from the ds cDNA: 5 to 6 hr.
4. Nonselective preamplification: 3 to 4 hr (up to 96 samples).
5. Selective amplification: 3 to 4 hr (up to 96 samples).
6. Gel electrophoresis (up to 4 × 96 samples), 6 to 8 hr.

The procedure may be interrupted after each of above steps. A typical experiment time course starting from poly(A)⁺ RNA is given in Table 25B.5.1.

Literature Cited

- Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Kervallage, A.R., McCombie, W.R., and Venter, J.G. 1991. Complementary DNA sequencing: Expressed sequence tags and the human genome project. *Science* 252:1651-1655.
- Bachem, C.W.B., Van der Hoeven, R.S., De Bruijn, S.M., Vreugdenhil, D., Zabeau, M., and Visser, R.G.F. 1996. Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development. *Plant J.* 9:745-753.
- Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D.H., Johnson, D., Luo, S., McCurdy, S., Foy, M., Ewan, M., Roth, R., George, D., Eletr, S., Albrecht, G., Vermanas, E., Williams, S.R.R., Moon, K., Burcham, T., Pallas, M., DuBridg, R.B., Kirchner, J., Fearson, K., Mao, J., and Corcoran, K. 2000. Gene expression analysis by massively parallel signature sequencing on microbead arrays. *Nat. Biotechnol.* 18:630-634.
- Breyne, P. and Zabeau, M. 2001. Genome-wide expression analysis of plant cell cycle modulated genes. *Curr. Opin. Plant Biol.* 4:136-142.
- De Risi, J.L., Iyer, V.R., and Brown, P.O. 1997. Exploring the metabolic and genetic control of gene expression on a genome scale. *Science* 278:1359-1367.
- Din, R.F., Nesert, E.W., and Comai, L. 2001. Plant gene expression response to *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. U.S.A.* 98:10954-10959.
- Durrant, W.E., Rowland, O., Piedras, P., Hammond-Kosack, K.E., and Jones, J.D. 2000. cDNA-AFLP reveals a striking overlap in race-specific resistance and wound response gene expression profiles. *Plant Cell* 12:963-977.
- Dynal. 1995. Biomagnetic techniques in molecular biology. Technical Handbook, Second Edition. Dynal A.S, Oslo, Norway.
- Fischer, A., Saedler, H., and Theissen, G. 1995. Restriction fragment length polymorphism-coupled domain-directed differential display: A highly efficient technique for expression analysis of multigene families. *Proc. Natl. Acad. Sci. U.S.A.* 92:5331-5335.
- Liang, P. and Pardee, A.B. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257:967-971.
- Prashar, Y. and Weismann, S.M. 1996. Analysis of differential gene expression by display of 3' end restriction fragments of cDNAs. *Proc. Natl. Acad. Sci.* 93: 659-663.
- Qin, L., Prins, P., Jones, J.T., Popeijus, J., Smant, G., Bakker, J., and Helder, J. 2001. GenEst, a powerful bidirectional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP. *Nucl. Acids. Res.* 29: 1616-1622.
- Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270:467-470.
- Van der Biezen, E.A., Juwana, H., Parker, J.E., and Jones, J.D. 2000. cDNA-AFLP reveals a striking overlap in race-specific resistance and wound response gene expression profiles. *Plant Cell.* 12:963-977.

- Velculescu, V., Zhang, L., Vogelstein, B., and Kinzler, K.W. 1995. Serial analysis of gene expression. *Science* 270:484-487.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Res.* 23:4407-4414.
- Vos, P. and Kuiper, M. 1998. AFLP analysis. *In* DNA Markers: Protocols, Applications and Overviews (G. Caetano-Anolles and P.M. Gresshoff, eds.) pp. 115-131. John Wiley and Sons, New York.
- Welsh, J.B., Zarrinkar, P.P., Supinosos, L.M., Kern, S.G., Behling, C.A., Monk, B.J., Lockhart, D.J., Burger, S.A., and Hampton, G.M. 2001. Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. *Proc. Natl. Acad. Sci.* 98:1176-1181.
- Wodicka, L., Dong, H., Millmann, M., Ho, M.H., and Lockhart, D.J. 1997. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nature Genet.* 15:1359-1367.
- Zabeau, M. and Vos, P. 1993. Selective restriction fragment amplification: A general method for DNA fingerprinting. European Patent EP 0534858-B1.

Contributed by Pieter Vos and Patrick
Stanssens
Keygene N.V.
The Netherlands

Serial Analysis of Gene Expression (SAGE)

UNIT 25B.6

This unit provides a protocol for performing serial analysis of gene expression (SAGE). SAGE involves the generation of short fragments of DNA, or tags, from a defined point in the sequence of all cDNAs in the sample analyzed. This short tag, because of its presence in a defined point in the sequence, is typically sufficient to uniquely identify every transcript in the sample. SAGE allows one to generate a comprehensive profile of gene expression in any sample desired from as few as 100,000 cells or as little as 1 µg total RNA. SAGE also allows an investigator to readily and reliably compare data to those produced by other laboratories, making the SAGE data set increasingly useful as more data are generated and shared.

Serial analysis of gene expression (SAGE), as described in the main method (see Basic Protocol 1), involves the generation of an oligonucleotide library, with each 14-bp SAGE tag representative of a discrete cDNA. Sometimes, the gene that the SAGE tag represents cannot be readily identified. Thus, a second method (see Basic Protocol 2) describes reverse cloning the 3' end of the cognate cDNA for an unknown SAGE tag. Three additional protocols for verifying cDNA by PCR (see Support Protocol 1), optimizing ditag PCR (see Support Protocol 2), and annealing linkers (see Support Protocol 3), are also given.

MicroSAGE

SAGE library construction involves anchoring mRNA molecules via their poly(A) tails to magnetic beads. (MicroSAGE, which is described here, differs from conventional SAGE in that this anchoring at the 3' end takes place prior to cDNA synthesis rather than after cDNA synthesis.) cDNA synthesis is then conducted, and the cDNAs are cleaved with *Nla*III to completion. This results in the loss of all cDNA sequence 5' to the cleavage site, and ensures that only the 3'-most *Nla*III site is exposed at the 3' end of the cDNA. The cDNA sample is then divided into two equal pools and two sets of linkers (which contain a *Bsm*FI site, PCR primer sites, and modified 3' bases to prevent ligation to each other) are then added by ligation. *Bsm*FI is a type IIS restriction enzyme, with a cut site 15 bp 3' of the recognition site. The resulting cDNAs are then digested with *Bsm*FI, which results in the release of the linker, the *Nla*III site, and 10 to 11 bp 3' of the *Nla*III site. The resulting "tags" are then blunt-ended with the Klenow fragment of DNA polymerase I, and the two separate pools of tags are ligated together via blunt-end ligation to form "ditags." These are then amplified via the PCR primer sites incorporated into the linkers and then recleaved with *Nla*III. These cleaved ditags are purified and ligated together to form concatemers of tags, which are then subcloned into plasmid vectors to create a SAGE library. Individual clones are then sequenced, and analyzed via SAGE analysis software. SAGE software identifies and discards any sets of duplicate ditags (i.e., a given combination of any two individual tags) to control for PCR amplification bias. It can also be used to prepare a tag report, listing all tags and their abundance in a given library, or a tag comparison file, listing the tag abundances in a number of different libraries.

An overview of the microSAGE protocol is shown in Figure 25B.6.1.

**BASIC
PROTOCOL 1**

**Discovery of
Differentially
Expressed Genes**

25B.6.1

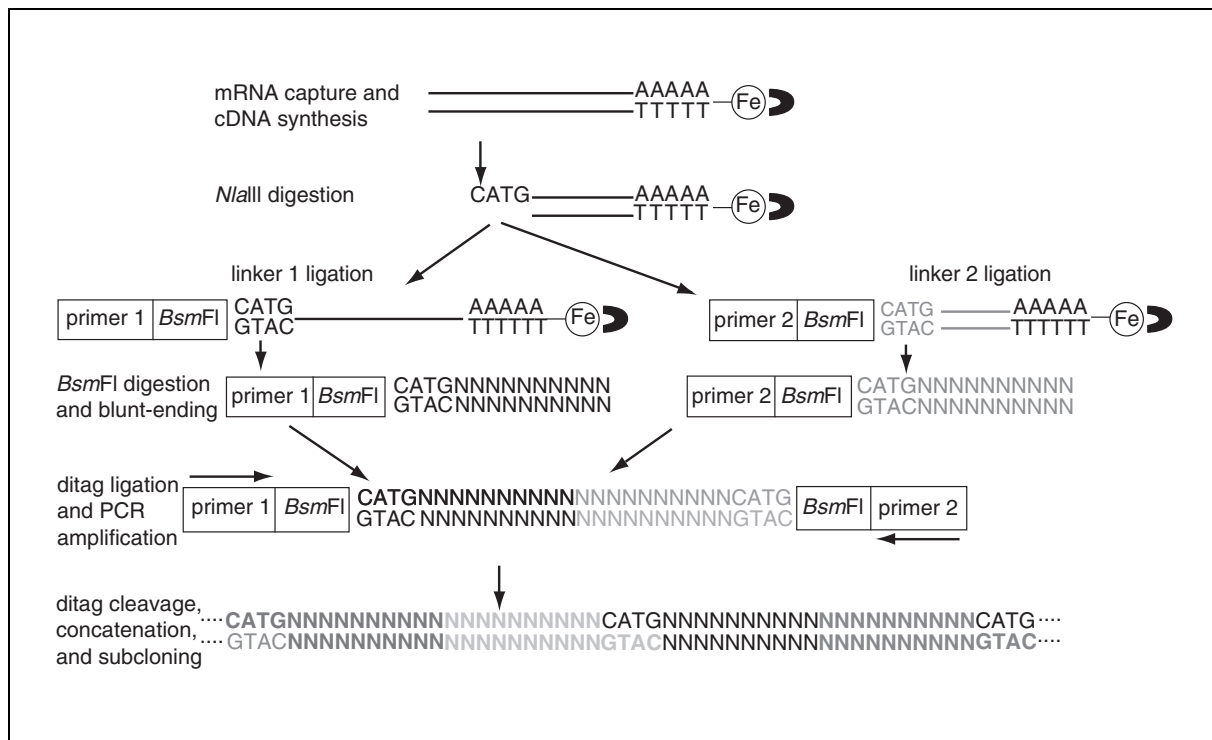


Figure 25B.6.1 The steps of a SAGE experiment.

Materials

Dynabeads mRNA DIRECT kit (Dynal Biotech):

Dynabeads oligo (dT)₂₅

Lysis/binding buffer

Washing buffer A: add 1 µl 20 mg/ml molecular-biology-grade glycogen (Roche Diagnostics) per milliliter

Washing buffer B

Cells or tissue of interest

SuperScript Choice System cDNA synthesis kit (Life Technologies):

5× first-strand buffer

DEPC-treated double-distilled water (DEPC ddH₂O)

1× first-strand buffer: dilute from 5× stock in DEPC ddH₂O

0.1 M DTT

10 mM dNTP

200 U/µl SuperScript II reverse transcriptase

5× second-strand buffer

10 U/µl *E. coli* DNA ligase

10 U/µl *E. coli* DNA polymerase I

2 U/µl *E. coli* RNase H

1× and 5× T4 DNA ligase buffer

1 U/µl T4 DNA ligase (UNIT 3.14)

0.5 M EDTA, pH 8.0 (APPENDIX 2)

1× BW buffer (see recipe)/2× BSA (New England Biolabs)/0.1% (w/v) SDS

1× BW buffer/2× BSA

1× NEBuffer 4 (New England Biolabs)/2× BSA

LoTE buffer (see recipe)

100× BSA (New England Biolabs)

10 U/µl NlaIII and 10× NEBuffer 4 (New England Biolabs; UNIT 3.1): store at -80°C

1× BW buffer/2× BSA/1% (v/v) Tween 20

Annealed linkers (see Support Protocol 3):
 5 U/μl (high-concentration) T4 DNA ligase (Life Technologies; *UNIT 3.14*)
 2 U/μl *BsmFI* (New England Biolabs; *UNIT 3.1*)
 PC8 (see recipe)
 SeeDNA (Amersham Pharmacia Biotech)
 3:1 solution of 20 mg/ml glycogen/SeeDNA (optional)
 3 M sodium acetate (*APPENDIX 2*)
 70% and 100% ethanol
 Klenow fragment of DNA polymerase I and 10× buffer (Amersham Pharmacia Biotech; *UNIT 3.5*) or Roche Buffer H
 3 mM Tris·Cl, pH 7.5 (*APPENDIX 2*)
 10× SAGE PCR amplification buffer (see recipe)
 DMSO (Sigma)
 PCR primers (see recipe):
 350 ng/μl primers 1 and 2
 350 ng/μl M13 forward and reverse primers
 5 U/μl Platinum *Taq* DNA polymerase (Life Technologies; *UNIT 3.5*)
 20 mg/ml glycogen (Roche Diagnostics)
 7.5 M ammonium acetate (Sigma)
 Dry ice/methanol bath
 5× loading buffer (*UNIT 2.7*)
 20% (w/v) polyacrylamide/TBE minigels (Novex)
 20-bp DNA ladder (GenSura)
 10,000× SYBR Green I (Roche Diagnostics)
 1× TBE (*APPENDIX 2*)
 1-kb DNA ladder
 pZErO-1 plasmid (Invitrogen)
SphI and NEBuffer 2 (New England Biolabs; *UNIT 3.1*)
 TE buffer, pH 8.0 (*APPENDIX 2*)
 SOC medium (*UNIT 1.8*)
 0.01 ng/μl pUC19 control DNA
 DH10B Electromax competent cells, −70°C (Life Technologies)
 LB medium (*UNIT 1.1*; optional)
 LB plates with 100 μg/ml ampicillin (*UNIT 1.1*)
 10-cm zeocin-containing low-salt LB plate (see recipe)
 10:1 U/μl *Taq/Pfu* polymerase (Stratagene)
 Exonuclease I (USB)
 Shrimp alkaline phosphatase (USB)
 50 mM Tris·Cl, pH 8.0 (*APPENDIX 2*)
 0.5-, 1.5-, 2.0-ml RNase-free No-stick siliconized microcentrifuge tubes (Ambion)
 Magnetic rack for 1.5-ml microcentrifuge tubes (DynaL Biotech)
 Tissue homogenizer (e.g., Polytron PT1200, Brinkmann Instruments)
 23-G needles and 1-ml syringes
 200-μl aerosol-barrier pipet tips
 16° and 65°C water baths, heat blocks, or equivalent
 96-well PCR plates
 50-ml conical tubes
 Tabletop centrifuge with swinging-bucket rotor
 Gel-loading tips
 UV box and SYBR green or UV filter
 0.5-ml microcentrifuge tubes with ~0.5-mm holes in the bottom: pierce from the inside out with a 21-G needle

Spin-X centrifuge-tube filters (Costar)
Long wavelength UV source
0.1-mm disposable micro-electroporation cuvettes (BioRad)
Bio-Rad gene pulser electroporator or equivalent
15-ml culture tubes

Additional reagents and equipment for determining integrity of cDNA by PCR (see Support Protocol 1), optimizing ditag PCR conditions (see Support Protocol 2), agarose gel electrophoresis (UNIT 2.5A), ethanol precipitation (UNIT 2.1A), and direct sequencing of PCR products (UNIT 15.2)

NOTE: Prepare Dynabeads, washing solutions, and 5× first-strand mix before thawing and collecting cells.

Prepare mRNA and synthesize cDNA

1. Thoroughly resuspend Dynabeads oligo (dT)₂₅, transfer 100 µl to a 1.5-ml RNase-free siliconized No-stick microcentrifuge tube, and place on a magnetic rack. After ~30 sec remove supernatant.

This volume of beads is much more than needed, but permits easy handling.

When removing the supernatant, always place the pipet tip at the opposite side of the tube, push the pipet tip to the bottom, and pipet very slowly, so as not to disturb the beads.

2. Resuspend beads in 500 µl lysis/binding buffer by “flicking” the tube or by gently vortexing. Leave beads in buffer until ready to add them to the cell lysate (step 4).

In this and all subsequent washing steps, add solution to the tube while keeping it on the magnetic rack in order to minimize “drying out” of the beads. Next, close the cap, remove the tube from the magnet, and resuspend the beads. Place back on the magnetic rack for ~30 sec to collect beads at the bottom before removing wash.

3. Lyse 100,000 to 1,000,000 cells (or 2 to 10 mg tissue) in 1 ml lysis/binding buffer in a 2-ml microcentrifuge tube with a tissue homogenizer for 1 min.

Before using the homogenizer, clean it thoroughly, rinse with 100% ethanol, and pulse in 1 liter DEPC ddH₂O.

If necessary, remove any cellular debris that remains following homogenization by microcentrifuging 1 min at maximum speed.

4. Immediately shear genomic DNA by pressing lysed cells through a 23-G needle attached to a 1-ml syringe into the tube containing prewashed Dynabeads (step 2), from which the buffer has been removed. Incubate 3 to 5 min at room temperature with constant agitation by hand.

Alternatively, total RNA previously isolated and stored at –80°C may be used. Total RNA (1 to 10 µg in 500 µl of lysis/binding buffer) may be added and incubated 3 to 5 min, room temperature, with constant agitation by hand. It is best to run some of the RNA on a denaturing gel to check for degradation. Visualization of sharp 28S and 18S ribosomal bands should be seen.

5. Place the tube on a magnetic rack for 2 min, then remove the supernatant.

This supernatant can be used for a genomic DNA prep if desired.

6. Wash beads by pipetting up and down several times with a 200-µl aerosol-barrier pipet tip in the following sequence:

Twice with 1 ml washing buffer A
Once with 1 ml washing buffer B
Four times with 1× first-strand buffer.

Pipetting the beads is more efficient than flicking the tubes.

7. Resuspend beads in the following first-strand synthesis mix:

54 μ l DEPC ddH₂O
18 μ l 5 \times first-strand buffer
9 μ l 0.1 M DTT
4.5 μ l 10 mM dNTP.

Heat tube 2 min at 37°C, then add 3 μ l of 200 U/ μ l SuperScript II reverse transcriptase. Incubate 1 hr at 37°C, mixing beads every 10 min by hand. Terminate reaction by placing tube on ice.

8. Add the following components of the second-strand synthesis to the first-strand reaction in the order shown on ice:

227 μ l ddH₂O, prechilled
150 μ l 5 \times second-strand buffer
15 μ l 10 mM dNTP
3 μ l 10 U/ μ l *E. coli* DNA ligase
12 μ l 10 U/ μ l *E. coli* DNA polymerase I
3 μ l 2 U/ μ l *E. coli* RNase H.

Incubate 2 hr at 16°C, mixing beads every 10 min by hand.

9. After incubation, place tubes on ice and terminate reaction by adding 100 μ l of 0.5 M EDTA, pH 8.0.

10. Wash beads one time with 0.5 ml of 1 \times BW buffer/2 \times BSA/0.1% (w/v) SDS.

The BSA appears to reduce the stickiness of the beads and improves the efficiency of the washes and the quality of the library. Extra washes with SDS can cause beads to clump severely.

11. Wash beads three times, each in 500 μ l of 1 \times BW buffer/2 \times BSA. Resuspend beads in 500 μ l of 1 \times BW buffer/2 \times BSA and heat 20 min at 75°C.

This heating step is crucial as it inactivates the nuclease activity of Poll.

12. Wash three times in 500 μ l of 1 \times BW buffer/2 \times BSA. Wash twice with 200 μ l of 1 \times NEBuffer 4/2 \times BSA, transferring to new tubes after the first wash in NEBuffer 4/BSA and saving 5 μ l of the last bead suspension.

13. Check the integrity of cDNA by PCR with primers for genes known to be in the cDNA being used for library construction using the saved 5- μ l aliquot (see Support Protocol 1).

Cleave cDNA with anchoring enzyme (NlaIII) and ligate linkers to cDNA

14. Resuspend beads in following mix:

171 μ l LoTE buffer
4 μ l 100 \times BSA
20 μ l 10 \times NEBuffer 4
5 μ l 10 U/ μ l *NlaIII*.

Incubate 1 hr at 37°C.

15. After incubation, place on a magnetic rack ~30 sec, then wash beads with the following solutions by pipetting up and down several times with a 200- μ l aerosol-barrier pipet tip:

Twice with 500 μ l 1 \times BW/2 \times BSA/1% Tween 20
Four times with 500 μ l 1 \times BW/2 \times BSA
Twice with 1 \times T4 DNA ligase buffer.

After final resuspension in ligase buffer, transfer 100 μ l of each sample into two new 1.5-ml siliconized microcentrifuge tubes.

16. Remove last wash and resuspend beads with the following:

- 5 μ l LoTE buffer (both tubes)
- 2 μ l 5 \times T4 DNA ligase buffer (both tubes)
- 3 μ l 2 ng/ μ l annealed linkers 1A and 1B (only in tube 1)
- 3 μ l 2 ng/ μ l annealed linkers 2A and 2B (only in tube 2).

17. Heat tubes 2 min at 50°C then let sit for 5 to 15 min at room temperature. Add 1 μ l of 5 U/ μ l (high-concentration) T4 DNA ligase to each tube and incubate 2 hr at 16°C. Mix beads intermittently.

Release cDNA-tags using tagging enzyme BsmFI

18. After ligation, place on a magnetic rack ~30 sec, then wash each sample two times with 500 μ l of 1 \times BW/2 \times BSA/0.1% SDS each, pooling tube 1 and tube 2 together after first wash in order to minimize loss in subsequent steps.

19. Wash four times with 500 μ l of 1 \times BW/2 \times BSA each and twice with 200 μ l of 1 \times NEBuffer 4/2 \times BSA (transfer to new tubes after first wash in NEBuffer 4/BSA).

20. Preheat the following mix 2 min at 65°C:

- 170 μ l LoTE buffer
- 20 μ l 10 \times NEBuffer 4
- 4 μ l 100 \times BSA
- 2 μ l 2 U/ μ l *BsmFI*.

Resuspend beads in the mixture and incubate 1 hr at 65°C, mixing intermittently.

21. After incubation, microcentrifuge 2 min at maximum speed, then transfer supernatant to a new 1.5-ml microcentrifuge tube. Wash beads once with 40 μ l LoTE buffer, then resuspend to a final volume of 240 μ l with LoTE buffer.

IMPORTANT NOTE: From this point on, do not use siliconized tubes.

22. Extract with 240 μ l PC8 and ethanol precipitate with SeeDNA using the following procedure:

- a. Add 4 μ l SeeDNA. Alternatively, use 4 μ l of a 3:1 solution of 20 mg/ml glycogen/SeeDNA mix.
- b. Add 0.1 vol of 3 M sodium acetate (24 μ l) and mix briefly.
- c. Add 2 vol of 100% ethanol (480 μ l) and vortex briefly.
- d. Incubate 2 min at room temperature.
- e. Microcentrifuge 5 min at maximum speed.
- f. Wash two times with 70% ethanol and microcentrifuge again after last wash. Carefully remove residual liquid with a pipet tip and resuspend pellet in 10 μ l LoTE buffer.

SeeDNA is a brightly colored carrier molecule that allows easy visualization and maximal recovery of alcohol-precipitated DNA or RNA. The glycogen/SeeDNA mixture may be used to reduce cost.

One may pause the protocol here and store the pellet overnight at -20°C.

Perform blunt-end digestion on released tags

23. Add the following mix to tags:

- 30.5 μl ddH₂O
- 5 μl 10 \times Klenow buffer (or Roche Buffer H)
- 2.5 μl 10 mM dNTPs
- 1 μl 100 \times BSA
- 1 μl Klenow fragment of DNA polymerase I.

Incubate 30 min at 37°C then add 190 μl LoTE buffer (240 μl final volume).

24. Extract with an equal volume of PC8 (240 μl). Transfer 200 μl into a ligase “+” tube and the remaining 40 μl into a ligase “-” tube.
25. Ethanol precipitate with 2 μl SeeDNA, 0.1 vol of 3 M sodium acetate, and 2 vol of 100% ethanol. Wash two times with 70% ethanol and centrifuge again after last wash. Carefully remove residual liquid with a pipet tip and air-dry 5 to 10 min. Resuspend pellet in 2 μl LoTE buffer.

Do not overdry because DNA will be lost.

Ligate tags to form ditags

26. Prepare 2 \times ligase “+” mix as follows:

- 2.5 μl 3 mM Tris·Cl, pH 7.5
- 3.0 μl 5 \times T4 DNA ligase buffer
- 2.0 μl 5 U/ μl (high-concentration) T4 DNA ligase.

Prepare a 2 \times ligase “-” mix with 4.5 μl of 3 mM Tris·Cl, pH 7.5 and 3.0 μl of 5 \times T4 DNA ligase buffer. Add 2 μl of appropriate mix to +/- ligase samples and incubate in a thermal cycler overnight (8 to 12 hr) at 16°C.

The sample may dry out in a water bath (in 4°C cold room), thus incubation in a PCR machine/thermal cycler is preferable.

27. After ligation, add 98 μl LoTE buffer, optimize PCR conditions (see Support Protocol 2), and proceed to large-scale PCR amplification.

Samples may be stored >1 year at -20°C.

Perform large-scale PCR amplification of ditags

28. Prepare a reaction mastermix for large-scale PCR (two to three 96-well PCR plates containing 50 μl reaction per well) using the following recipe for one reaction as a guide:

- 5 μl 10 \times SAGE PCR amplification buffer
- 3 μl DMSO
- 4.0 to 10 μl 10 mM dNTPs
- 1 μl 350 ng/ μl PCR primer 1
- 1 μl 350 ng/ μl PCR primer 2
- Adjust volume to 49 μl with ddH₂O
- 0.7 μl 5 U/ μl Platinum *Taq* DNA polymerase.

Aliquot 49 μl of reaction mix to each well, then add 1 μl template at appropriate dilution (see Support Protocol 2).

The authors usually use a 300-reaction PCR premix that is dispensed into 96-well plates at 50- μl per well.

The volume of dNTPs to use is determined through optimization (see Support Protocol 2).

Platinum Taq DNA polymerase is used because it allows for a room-temperature hot start reaction (the Taq DNA polymerase is complexed with an anti-Taq antibody that denatures when heated to 94°C).

29. Carry out the amplifications in a thermal cycler with the following parameters:

1 cycle:	2 min	94°C	(denaturation)
26 to 32 cycles:	30 sec	94°C	(denaturation)
	1 min	55°C	(annealing)
	1 min	70°C	(extension)
1 cycle:	5 min	70°C	(final product extension).

The number of cycles to use is determined through optimization (see Support Protocol 2).

If a thermal cycler with heated lid is not available, oil can be used to prevent evaporation (see UNIT 15.1).

The ligase “–” sample should be amplified for 35 cycles.

Do not substitute conventional hot-start PCR for use of Platinum Taq DNA polymerase. The authors have found that yields are much lower if this is done. There is no need to refrigerate the PCR mix while setting up the reactions.

Isolate ditags

30. Pool PCR reactions into a 50-ml conical tube, adjusting volume to 11.5 ml with LoTE buffer, then extract with an equal volume of PC8.

31. Precipitate with ethanol as follows:

11.5 ml samples
10 µl SeeDNA
100 µl 20 mg/ml glycogen
5.1 ml 7.5 M ammonium acetate
38.3 ml 100% ethanol.

Place in a dry ice/methanol bath for 15 min. Thaw 2 min at room temperature to fully melt the solution.

32. Vortex briefly and centrifuge 30 min in a tabletop centrifuge with swinging-bucket rotor at $\sim 3000 \times g$ (4000 rpm), room temperature.

33. Wash with 5 ml of 70% ethanol, vortex, and centrifuge an additional 5 min at $\sim 3000 \times g$, room temperature.

34. Resuspend pellet in 216 µl LoTE buffer and add 54 µl of 5× loading buffer (270 µl total).

35. Using gel-loading pipet tips, load 10 µl sample into each of 27 lanes on each of three prepoored 20% polyacrylamide/TBE minigels. Include 10 µl of a 20-bp ladder on each gel as a marker.

It is critical not to overload the gel wells, as this can lead to linker contamination and poor separation of products.

36. Electrophorese 90 min at 160 V.

The optimal distance for electrophoresis is ~ 1 cm above the bottom of the gel. The idea is to obtain maximum separation of the 102- (ditags) and 80-bp bands (linker-linker dimers) without allowing product to get too close to the edge of the gel. Depending on the apparatus and batch of TBE buffer, varying the electrophoresis time might be necessary.

37. Stain 15 min in a foil-wrapped container on a platform shaker using 2 to 5 μl of 10,000 \times SYBR Green I in 50 ml of 1 \times TBE buffer. Visualize on a UV box using a SYBR green or UV filter.

Alternatively, use long-wavelength UV. Amplified ditags should run at 102 bp while a background band (linker-linker dimers) runs at \sim 80 bp.

38. Cut out only amplified ditags from the gel, and place three cut-out bands in 0.5-ml microcentrifuge tubes (nine tubes total) which have an \sim 0.5-mm diameter hole in the bottom.

39. Place the 0.5-ml microcentrifuge tubes in 2.0-ml siliconized microcentrifuge tubes and microcentrifuge 4 min at maximum speed.

This serves to break up the acrylamide gel into small fragments at the bottom of the 2.0-ml microcentrifuge tube.

40. Discard 0.5-ml microcentrifuge tubes. Add 250 μl LoTE buffer and 50 μl of 7.5 M ammonium acetate to each 2.0-ml microcentrifuge tube.

At this point, the 2.0-ml microcentrifuge tubes can remain overnight at 4°C.

41. Vortex each tube, and incubate 15 min at 65°C. Add 5 μl LoTE buffer to the membrane of each of 18 Spin-X centrifuge-tube filters.

42. Transfer contents of each tube to two Spin-X centrifuge tube filters (i.e., nine tubes transferred to 18 Spin-X centrifuge tube filters). Microcentrifuge each SpinX filter for 5 min at maximum speed. Consolidate sets of two eluates (300 μl total) and transfer to 1.5-ml microcentrifuge tubes.

Sometimes, purified 102-bp bands do not recut well with NlaIII, which seems to be related to imperfect purification from the gel. If this is a problem, run 300 μl eluate through a Qiaquick gel extraction protocol (Qiagen). Bring the volume of the extract back up to 300 μl to proceed.

43. Ethanol precipitate eluates by adding the following:

300 μl sample
0.5 μl SeeDNA
1.5 μl glycogen
133 μl 7.5 M ammonium acetate
1000 μl 100% ethanol.

Vortex and place in a dry ice/methanol bath for 15 min. Warm 2 min at room temperature until solution has melted, then microcentrifuge 15 min at 4°C.

44. Microcentrifuge 15 min at maximum speed. Wash two times with 75% ethanol. Resuspend each DNA tube in 10 μl LoTE buffer. Pool samples into one microcentrifuge tube (90 μl total).

The total amount of DNA at this stage should be 10 to 20 μg .

45. Digest PCR products with NlaIII by adding the following:

90 μl PCR products in LoTE buffer
226 μl LoTE buffer
40 μl 10 \times NEBuffer 4
4 μl 100 \times BSA
40 μl 10 U/ μl NlaIII.

Incubate 1 hr at 37°C.

Purify the ditags

46. Extract with an equal volume of PC8. Pool aqueous phases and transfer into 1.5-ml microcentrifuge tubes. Ethanol precipitate in dry ice as follows:

200 μ l sample
66 μ l 7.5 M ammonium acetate
3 μ l SeeDNA
825 μ l 100% ethanol.

Vortex and place in dry ice/methanol bath for 15 min.

47. Warm 2 min at room temperature until solution has melted, then microcentrifuge 15 min at 4°C.
48. Wash once with cold 75% ethanol, removing ethanol traces with a gel-loading pipet tip. Resuspend pellet in 40 μ l LoTE buffer. On ice, add 10 μ l of 5 \times loading buffer (50 μ l total).
49. Load this sample into four lanes of a 20% polyacrylamide/TBE gel, load the 20-bp ladder into a separate lane, and run ~2.5 hr at 160 V. Stain as described in step 37.

Optimal electrophoresis time may vary somewhat. Be careful not to run the gel too long.

50. Cut out the 24- to 26-bp band from four lanes under long-wavelength UV illumination, and place two cut-out bands in each of two 0.5-ml microcentrifuge tubes which have an ~0.5-mm diameter hole in the bottom.
51. Microcentrifuge as described in step 39.
52. Discard the 0.5-ml microcentrifuge tubes. Add 250 μ l LoTE buffer and 50 μ l of 7.5 M ammonium acetate to each of the 2.0-ml microcentrifuge tubes. Vortex the tubes, and incubate 1 hr at 37°C.

IMPORTANT NOTE: Do not incubate at 65°C. This will cause the 26-bp ditags to denature. Longer incubations (even overnight) can be performed, but do not appear to result in significantly higher yields.

53. Use four Spin-X centrifuge-tube filters to isolate eluate as described in step 42. Ethanol precipitate in three tubes (200 μ l each) with the following:

200 μ l sample
66 μ l 7.5 M ammonium acetate
2 μ l SeeDNA
3 μ l glycogen
825 μ l 100% ethanol.

Incubate 10 min in a dry ice/methanol bath, then microcentrifuge 15 min at 4°C.

54. Wash two times with cold 75% ethanol each. Resuspend each DNA sample on ice in 2.5 μ l cold LoTE buffer and pool (7.5 μ l total).

It is critical to keep the purified ditags on ice until the ligation buffer is added. High A and T content ditags can denature at room temperature, even in LoTE buffer.

Ligate ditags to form concatemers

55. Mix the following:

7 μ l pooled purified ditags
2 μ l 5 \times T4 DNA ligase buffer
1 μ l 5 U/ μ l (high-concentration) T4 DNA ligase.

Incubate 1 to 3 hr at 16°C.

Do not ligate overnight, as this will result in long concatemers that are difficult to clone. The authors usually ligate for 2 hr with good results.

The length of ligation time depends on the quantity and purity of the ditags. Typically, several hundred nanograms of ditags are isolated and produce large concatemers when the ligation reaction is performed for 1 to 3 hr at 16°C (lower quantities or less-pure ditags will require longer ligations).

56. After completing ligation, add 2.5 μ l of 5 \times loading buffer to the ligation reaction. Heat samples 5 min at 65°C and immediately place on ice.

The heating step melts annealed sticky ends and is critical for obtaining a good yield of clonable concatemers.

57. Separate concatemers on a 10% to 12% polyacrylamide/TBE gel (UNIT 2.7). Load 1-kb DNA marker in first lane, leave one empty lane, and then load the entire concatenated sample into the third well. Run samples 45 min at 200 V.

58. Stain and visualize as described in step 37. Isolate regions of interest.

Concatemers will form a smear on the gel with a range from ~100 bp to several kilobases.

The authors usually isolate regions 600 to 1200 bp and 1200 to 2500 bp. These size ranges clone efficiently and yield ample sequencing information.

59. Place each gel piece into 0.5-ml microcentrifuge tubes which have an ~0.5-mm-diameter hole in the bottom.

60. Microcentrifuge as described in step 39.

61. Discard the 0.5-ml microcentrifuge tubes. Add 300 μ l LoTE buffer to the gel pieces in the 2.0-ml microcentrifuge tubes. Vortex each tube, and incubate 15 min at 65°C.

If desired, this incubation can be extended to overnight, but yields are not significantly increased.

Note that ammonium acetate is not required for high-molecular-weight molecules.

62. Add 5 μ l LoTE to the membrane of each four Spin-X microcentrifuge-tube filters. Transfer contents of each tube to two Spin-X microcentrifuge-tube filters (four total). Microcentrifuge each Spin-X tube 5 min at maximum speed.

63. Pool eluates from two Spin-X centrifuge tube filters into one 1.5-ml microcentrifuge tube and ethanol precipitate by adding the following:

300 μ l eluate
2 μ l SeeDNA
133 μ l 7.5 M ammonium acetate
1000 μ l 100% ethanol.

Glycogen can be substituted for SeeDNA, but the authors obtained better results when only SeeDNA was used.

64. Microcentrifuge 15 min at maximum speed. Wash two times with 70% ethanol and air dry 5 min. Resuspend purified concatemer DNA in 6 μ l LoTE buffer.

Ligate the concatemers into vector

65. Digest 1 μ g pZER0-1 plasmid with *Sph*I in a total volume of 10 μ l by adding the following:

1 μ l pZER0-1 plasmid
7 μ l ddH₂O
1 μ l 10 \times NEBuffer 2
1 μ l 10 U/ μ l *Sph*I.

Incubate 15 to 30 min at 37°C, then heat inactivate 10 min at 65°C. Do not digest >30 min.

Concatemers can be cloned and sequenced in a vector of choice. The authors currently clone concatemers into a SphI-cleaved pZErO-1.

66. Check for complete digestion on an agarose gel (UNIT 2.5A). Dilute the cut vector with 90 µl TE buffer, pH 8.0, then extract with equal volume of PC8. Ethanol precipitate (UNIT 2.1A), wash two times with 70% ethanol, and resuspend in 40 µl water or TE buffer (~25 ng/µl of vector).

The authors recommend using the linearized DNA immediately, but it may be stored for up to 2 weeks at -20°C with decreased ligation efficiency. Ligation efficiency varies beyond 2-week storage. A 2-to-5 fold increase in background is observed upon prolonged storage, due to self-ligation—i.e., no insert.

67. Mix the following ligation reaction and set up a duplicate reaction for a control with no concatemer:

6 µl purified concatemer (step 64; none in control)
1.5 µl dH₂O (7.5 µl in control)
1 µl 25 ng/µl pZErO plasmid cut with *SphI*
1 µl 10× T4 DNA ligase buffer
1.0 µl 1 U/µl T4 DNA ligase.

Incubate 2 hr at 16°C.

Consider using 3 µl concatemers and save the rest for backup.

*The manufacturer of pZErO plasmid warns that there is increased background at incubations >1 hr, which may result in breakthrough by spontaneous mutations in the *ccdB* death gene.*

68. Bring sample volume to 200 µl with LoTE buffer. Extract with an equal volume PC8, then ethanol precipitate by mixing the following:

200 µl sample
133 µl 7.5 M ammonium acetate
2 µl SeeDNA
777 µl 100% ethanol.

69. Wash four times with 70% ethanol. Microcentrifuge briefly at maximum speed, remove 70% ethanol, and air dry 5 min. Resuspend in 10 µl LoTE buffer.

Excess salt can cause arcing during electroporation and kill the cells.

Transfect DNA by electroporation

70. Place an appropriate number of 0.1-mm microelectroporation cuvettes and 1.5-ml microcentrifuge tubes on ice.

71. Place 1 ml SOC medium in an appropriate number of 15-ml culture tubes at room temperature.

72. Add 1 µl DNA from step 69 to 1.5-ml microcentrifuge tubes on ice. To determine transformation efficiency, add 1 µl of 0.01 ng/µl pUC19 control DNA to a tube labelled “control.”

Use 1 µl of the DNA for this transfection. The remainder of the sample is stored at -20°C.

73. Remove DH10B Electromax competent cells from -70°C and thaw on wet ice. When cells are thawed, mix cells by tapping gently.

74. Add 40 μl competent cells to each chilled 1.5-ml microcentrifuge tube containing DNA. Refreeze any unused cells in a dry ice/methanol bath for 5 min before returning to -70°C .
75. Pipet 40 μl of the cell/DNA mixture into a prechilled disposable microelectroporation cuvette (step 70). Perform electroporation with the Bio-Rad gene pulser electroporator at 100 Ω /25 μF /1.8 kV.
76. Transfer electroporated cells into a 15-ml culture tube and immediately add 1.0 ml SOC medium at room temperature. Shake 15 min at 225 rpm, 37°C .

The incubation time is short because, in theory, the postelectroporation incubation period is required for expression of the antibiotic resistance gene, hence increasing transformation efficiency. However, given that the doubling time of the bacteria is ~ 20 min, it is possible that the transformed bacteria may double during the incubation period, potentially skewing the library's representation of tags. With 15 min incubation prior to plating, the authors found the transformation efficiency to be 1.0×10^{10} cfu/ μg pUC19, respectable when compared with the 1-hr incubation recommended by the manufacturer that resulted in 1.5×10^{10} cfu/ μg pUC19.

77. Spread 100 μl of a 1:100 dilution of control cells (pUC19) in SOC or LB medium on LB plates containing 100 $\mu\text{g}/\text{ml}$ ampicillin.
78. Plate $\frac{1}{10}$ transfected bacteria onto each of ten 10-cm zeocin-containing low-salt LB plates. Incubate and analyze 12 to 16 hr later.

Insert-containing clones should have hundreds to thousands of colonies while no-insert control plates should have zero to tens of colonies.

Save all ten plates for each concatemer ligation reaction since, if insert size appears appropriate, these may be used for sequencing described below.

Check insert size by PCR

79. Prepare a reaction mastermix using the following recipe for one reaction as a guide:

2.5 μl 10 \times SAGE PCR amplification buffer
 1.25 μl DMSO
 1.25 μl 10 mM dNTP
 0.5 μl 350 ng/ μl M13 forward PCR primer
 0.5 μl 350 ng/ μl M13 reverse PCR primer
 18.5 μl ddH₂O
 0.5 μl 10:1 U/ μl *Taq:Pfu* DNA polymerase.

Pipet 25 μl master mix to wells of 96-well PCR plates.

*Any thermostable polymerase can be used (with the appropriate buffer), but the *Taq:Pfu* mixture works well.*

80. For each reaction, use a sterile toothpick or pipet tip to gently touch colony and then dip tip with a twirl into PCR mix.
81. Carry out the amplifications in a thermal cycler with the following parameters:

1 cycle:	2 min	95°C	(denaturation)
25 cycles:	30 sec	95°C	(denaturation)
	1 min	56°C	(annealing)
	2 min	72°C	(extension)
1 cycle:	5 min	70°C	(final product extension).

*For *Taq* DNA polymerase-based PCR amplifications, an extension time of 0.5 to 1.0 min/kb of template amplified is sufficient, but in contrast, *Pfu*-based PCR amplifications require a minimum extension time of 1 to 2 min/kb of amplified template to achieve similar target synthesis.*

82. Analyze on a 1.5% agarose gel at ~150 V (UNIT 2.5A).

For large-scale screening, use multichannel pipettors with an Owl Centipede 50-well horizontal electrophoresis system. The tips of the multichannel pipettors fit into every second well of the 50-slot comb used on the Owl Centipede rigs. Consequently, to maintain a sequential loading order for each 96-well plate, the authors prepare a separate 96-well loading plate with sample loading dye.

The authors typically get 85% to 95% of clones with inserts, of which >95% are >400 bp long. Libraries of this quality can be sequenced directly without gel screening and sorting.

Purify template and sequence amplification product

83. Use 2 μ l PCR product (the exact amount will depend on the sequencing protocol and should be optimized) for clean-up using the following:

0.1 μ l exonuclease I
0.1 μ l shrimp alkaline phosphatase
1.8 μ l 50 mM Tris·Cl, pH 8.0.

Add 2 μ l clean-up mix to 2 μ l DNA.

The exonuclease I degrades unincorporated primers while the alkaline phosphatase degrades unincorporated free nucleotides.

84. Perform reactions in 96-well plates on a thermal cycler, incubating 15 min at 37°C, then 15 min at 80°C. Add ddH₂O to 15 μ l. Sequence PCR products directly (UNIT 15.2).

Use as little as 2 μ l of diluted product for the sequencing reaction—optimize according to protocol. The authors run reactions on an ABI 3700 96 capillary machine, though any sequencing system may be used.

85. Download SAGE analysis software from SAGEnet (see Internet Resources) and follow easy-to-use instructions.

SUPPORT PROTOCOL 1

VERIFYING cDNA PRODUCTION BY PCR ANALYSIS

The PCR primers used to test efficiency of the reverse-transcription will depend on the species and tissue type from which the library is constructed. Working in mouse, the authors typically test a ubiquitously-expressed mRNA (RPS17) and a more tissue-restricted mRNA. Design primers to be 18 to 22 bp in length and have a T_m of 55° to 60°C. T_m for the two primers should not differ by more than 1° to 2°C. The PCR product should be 300 to 700 bp in length, with a 5' end not more than 1 kb from the 3' end of the mRNA. The following describes the authors' method; however, conditions will have to be optimized for each primer set (see UNIT 15.1).

Materials (also see Basic Protocol 1)

350 ng/ μ l 5' and 3' primers (e.g., Integrated DNA Technology)
Bead suspension (see Basic Protocol 1, step 13)

Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5A)

1. Prepare the following PCR mixture:

5 μ l 10 \times SAGE PCR buffer
3 μ l DMSO
4 μ l 10 mM dNTP mix
0.5 μ l 350 ng/ μ l 5' primer
0.5 μ l 350 ng/ μ l 3' primer
31.3 μ l ddH₂O
0.7 μ l 5 U/ μ l *Taq* DNA polymerase
5 μ l bead suspension.

It is possible to test smaller aliquots of bead suspension depending on the abundance of the template.

2. Perform PCR using the following program:

Initial step:	2 min	95°C	(denaturation)
30 cycles:	30 sec	95°C	(denaturation)
	1 min	53°–58°C	(annealing)
	1 min	72°C	(extension)
Final step:	5 min	70°C	(final extension).

Annealing temperature should be 2° to 3°C lower than the lowest predicted T_m for the primers.

3. Analyze 5 µl of each PCR product on a 1.5% agarose gel in TAE buffer and visualize bands by ethidium bromide staining (UNIT 2.5A).

OPTIMIZING DITAG PCR AMPLIFICATION

The following protocol gives a method for optimizing ditag PCR by varying template concentration, nucleotide concentration, and number of cycles. The optimal template concentration to use is the one which gives a high yield of the 102-bp band with the least concentration of template. A clear plateau in yield should be seen with high concentrations of template. The optimal concentration of nucleotide is simply that which gives the highest yield of the 102-bp band. If none of the PCR reactions give high yields of the 102-bp band, repeat the protocol, but run one tube for 30 cycles and one for 32 cycles. The authors have found that the optimal concentration of nucleotide can vary from batch to batch and supplier to supplier, so repeated optimization may be required.

See Basic Protocol 1 for materials.

1. Prepare serial dilutions of LoTE diluted ditag reaction (see Basic Protocol 1, step 27) at 1:3, 1:9, 1:27, 1:81, and 1:243 in LoTE buffer using 10 µl reaction and 20 µl LoTE buffer (30 µl total) at each step.
2. Prepare the following PCR reaction mixture:
 - 5 µl 10× SAGE PCR amplification buffer
 - 3 µl DMSO
 - 1 µl 350 ng/µl PCR primer 1
 - 1 µl 350 ng/µl PCR primer 2
 - 28.3 µl ddH₂O
 - 0.7 µl 5 U/µl Platinum *Taq* DNA polymerase.
3. Prepare six tubes each containing 1 µl of either stock (see Basic Protocol 1, step 27) or diluted ditag reaction (1:3, 1:9, 1:27, 1:81, or 1:243). In duplicate, add 4, 7, or 10 µl of 10 mM dNTP mix (i.e., prepare two tubes of each dilution and nucleotide concentration pair). Add sufficient double-distilled water to bring the total volume to 11 µl.
4. Perform PCR as described (see Basic Protocol 1, step 29), using 26 cycles for one of the duplicate tubes and 28 for the other.
5. Remove 10 µl from each reaction and run on a prepoured 20% polyacrylamide/TBE gel, using a 20-bp ladder as a marker (10 µl of 1:5 dilution of the marker stock solution; see Basic Protocol 1, steps 35 and 36). Stain gel and visualize as described (see Basic Protocol 1, step 37).

SUPPORT PROTOCOL 2

Discovery of
Differentially
Expressed Genes

25B.6.15

The amplified ditags should be 102 bp in size. A background band of equal or lower intensity (due to linker-linker dimers) occurs at ~80 bp. All other background bands should be of substantially lower intensity.

The ligase “-” samples should not contain any amplified product of the size of the ditags even at 35 cycles.

**BASIC
PROTOCOL 2**

REVERSE CLONING UNKNOWN SAGE TAGS (rSAGE)

SAGE is a technique that allows a generally unbiased evaluation of cellular mRNAs on a genome-wide scale, thus providing a generally more quantitative analysis than subtractive cloning or microarray approaches. Furthermore, the sequencing of 14-bp SAGE tags has a significantly higher throughput than conventional expressed sequence tag (EST) approaches; however, the cDNA that a SAGE tag represents may not be readily identifiable due to the lack of an appropriate anchored cDNA sequence or multiple potential tag to gene matches. This protocol describes an approach, reverse-SAGE (rSAGE), by which the native 3' sequence can be cloned from cDNA utilizing a variation of the original SAGE protocol and PCR primers based upon sequences in the SAGE tag. The advantage of this protocol is that the unknown gene is cloned using 3' cDNA fragments that are the most 3' sequences containing the anchoring enzyme recognition sequence. This approach provides increased specificity of cloning the appropriate cognate cDNA from an anonymous SAGE tag.

Figure 25B.6.2 summarizes this procedure. The starting material is total RNA that expresses the target gene, and as a result, the anonymous SAGE tag. Double-stranded cDNA is synthesized by mRNA priming with a biotinylated poly-dT oligonucleotide that also encodes an M13 forward priming site and an *Asc*I restriction site. The anchoring

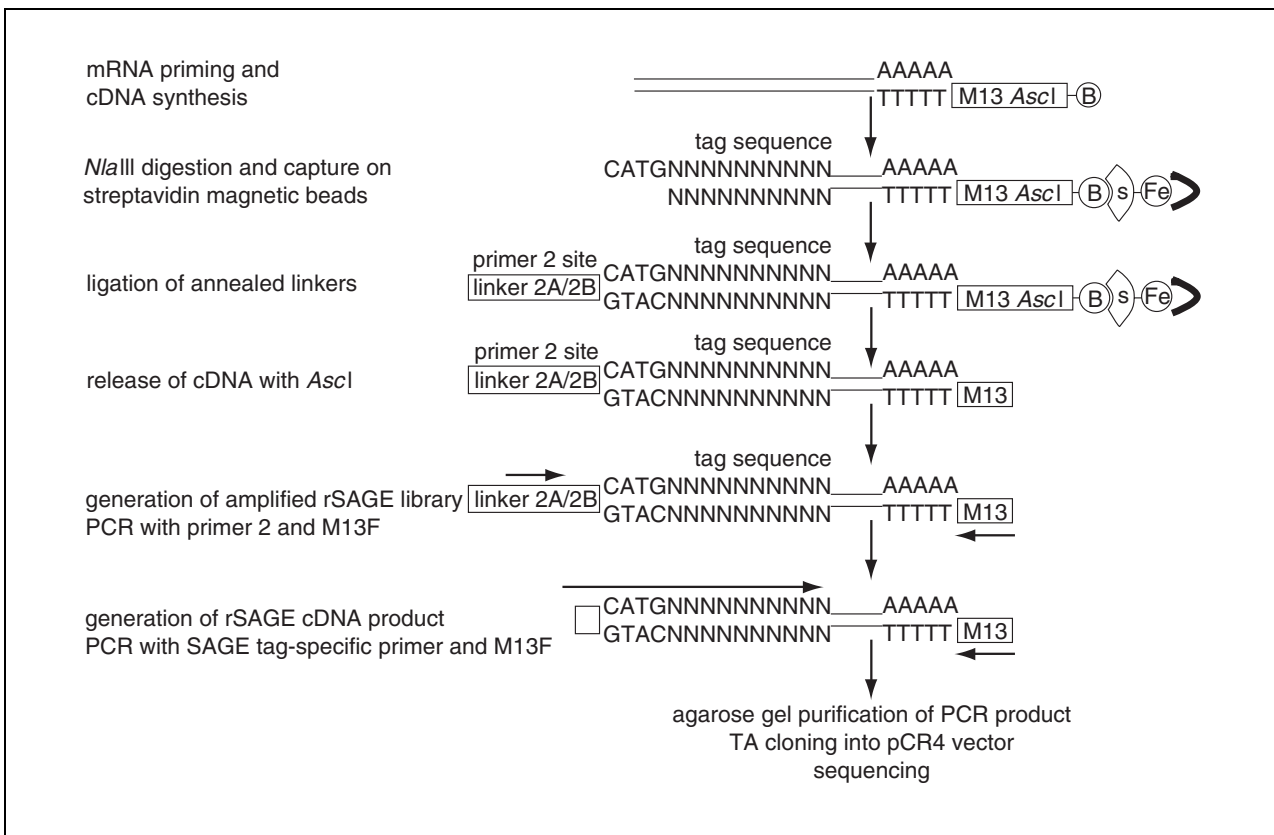


Figure 25B.6.2 Steps of an rSAGE experiment.

enzyme, *Nla*III, is used to cleave the cDNA and produce 3' cDNA fragments with *Nla*III cohesive overhangs. These 3' cDNA fragments are captured onto magnetic streptavidin Dynabeads and subsequently purified. The *Nla*III overhangs are then ligated with annealed linkers, 2A/2B, that encode a priming site for PCR primer 2, which is used for subsequent amplification. The cDNA is then released from the Dynabeads by digestion with *Asc*I restriction endonuclease. The resulting cDNA library is then amplified using PCR primer 2 and M13 forward primer (M13F). A specific rSAGE PCR product is then generated using a SAGE tag-specific primer with M13F. The SAGE tag-specific PCR product is then agarose gel purified and subsequently TA cloned into a sequencing vector.

Materials

SuperScript Choice System cDNA synthesis kit (Life Technologies):

- DEPC ddH₂O
- 5× first-strand buffer
- 0.1 M DTT
- 10 mM dNTP
- 200 U/μl SuperScript II reverse transcriptase
- 5× second-strand buffer
- 10 U/μl *E. coli* DNA ligase
- 10 U/μl *E. coli* DNA polymerase I
- 2 U/μl *E. coli* RNase H
- 5 U/μl T4 DNA polymerase
- 1× and 5× T4 DNA ligase buffer
- 1 μg/μl gel-purified BRS1 primer (see recipe)
- 0.5 M EDTA, pH 7.5 (*APPENDIX 2*)
- PC8 (see recipe)
- SeeDNA (Amersham Pharmacia Biotech)
- 7.5 M ammonium acetate (Sigma)
- 70% and 100% ethanol
- LoTE buffer (see recipe)
- 100× BSA (New England Biolabs)
- 10 U/μl *Nla*III (*UNIT 3.1*) and 10× NEBuffer 4 (New England Biolabs)
- Streptavidin Dynabeads (Dynal)
- 1× BW buffer (see recipe)
- Annealed linkers (see Support Protocol 1)
- 5 U/μl (high-concentration) T4 DNA ligase (Life Technologies; *UNIT 3.14*)
- 1× BW buffer/1× BSA
- 1× NEBuffer 4/1× BSA
- 100× BSA
- 10 U/μl *Asc*I (New England Biolabs; *UNIT 3.1*)
- 10× SAGE PCR buffer (see recipe)
- DMSO
- PCR primers (see recipe):
 - 350 ng/μl M13 forward primer
 - 350 ng/μl primer 2
- 5 U/μl Platinum Taq DNA polymerase (Life Technologies; *UNIT 3.5*)
- 4% to 20% TBE acrylamide gel (Novex)
- 1-kb ladder
- 1× SYBR green I (Roche Diagnostics) in TBE buffer (*APPENDIX 2*)
- 5 M betaine: prepare monohydrate salt (Sigma) in PCR-grade ddH₂O
- SAGE tag-specific primer (see recipe)

Qiaquick gel-extraction kit (Qiagen):

Qiaquick columns

EB Buffer

TOPO TA Cloning Kit with pCR2.1 vector (Invitrogen) *or*

TOPO TA Cloning Kit for Sequencing with pCR4-TOPO vector (Invitrogen)

16°, 50°, and 70°C water baths, heat blocks, or equivalent

1.5-ml No-stick siliconized microcentrifuge tubes (Ambion)

Magnetic rack for 1.5-ml microcentrifuge tubes (Dyna)

1.5-ml nonsiliconized nuclease-free microcentrifuge tubes

Additional reagents and equipment for preparing total RNA (UNIT 4.2) agarose-gel electrophoresis (UNIT 2.5A) and sequencing (UNIT 7.4A)

Synthesize cDNA

1. Prepare total RNA in DEPC ddH₂O using standard methods (e.g., UNIT 4.2).

Trizol (Sigma) is the preferred method in the authors' laboratory. The same RNA with which the original SAGE library was generated would be ideal (see Basic Protocol 1, steps 3 and 4).

It is advisable to also generate a control rSAGE library that will not express the genes of interest. As PCR cloning from the rSAGE library might generate more than one clonable band, PCR of a control rSAGE library would allow the researcher to discriminate and identify the likely rSAGE product representing the gene of interest.

2. Add 2 µl of 1 µg/µl gel-purified BRS1 primer to a nonsiliconized 1.5-ml microcentrifuge tube. Add 6 µl total RNA (5 to 10 µg total) and mix.
3. Heat mixture to 70°C for 10 min and quick chill on ice. Microcentrifuge briefly at room temperature. Prepare first-strand-synthesis mix as shown below:

8 µl BRS1 primer/RNA
4 µl 5× first-strand buffer
2 µl of 0.1 M DTT
1 µl of 10 mM dNTP.

4. Mix gently by vortexing and microcentrifuge briefly at room temperature. Incubate 2 min at 37°C, then add 5 µl of 200 U/µl SuperScript II reverse transcriptase and mix well. Incubate an additional 1 hr at 37°C.
5. After incubation, place tube on ice to terminate the reaction. Add the components of the second-strand-synthesis mixture to the first-strand reaction on ice in the order shown:

93 µl DEPC ddH₂O, 4°C
30 µl 5× second-strand buffer
3 µl 10 mM dNTP
1 µl 10 U/µl *E. coli* DNA ligase
4 µl 10 U/µl *E. coli* DNA polymerase I
1 µl 2 U/µl *E. coli* RNase H.

Vortex gently to mix.

6. Incubate 2 hr at 16°C. Intermittently mix by gentle flicking. Add 2 µl 5 U/µl T4 DNA polymerase and incubate 5 min at 16°C. Place tubes on ice and terminate reaction by adding 10 µl of 0.5 M EDTA, pH 7.5.

T4 DNA polymerase is used in the reverse-SAGE protocol to fill in 5' overhangs generated after second-strand synthesis.

7. Add 150 μl PC8 and vortex thoroughly. Microcentrifuge 5 min at maximum speed, room temperature. Remove and save aqueous layer ($\sim 150 \mu\text{l}$).

Unlike microSAGE, the reverse-SAGE protocol synthesizes DNA onto unbound biotinylated oligonucleotides, making purification (i.e., phenol-chloroform extraction followed by ethanol precipitation) easier. As a result, the heat denaturation and multiple wash steps in the SAGE protocol are unnecessary.

8. Ethanol precipitate aqueous layer in a fresh standard 1.5-ml microcentrifuge tube by adding the following reagents:

2 μl SeeDNA
70 μl 7.5 M ammonium acetate
500 μl 100% ethanol.

Vortex thoroughly, then microcentrifuge 20 min at maximum speed, 4°C. Wash pellet in 70% ethanol.

9. Resuspend in 20 μl LoTE buffer.

Samples may be stored at 4°C up to a week or frozen at -20°C for months. However, it is best to leave at 4°C overnight and resume the protocol the following day.

Cleave cDNA with anchoring enzyme (*NlaIII*) and ligate linkers

10. Cleave cDNA with the anchoring enzyme (*NlaIII*) using the following mixture:

20 μl cDNA (step 9)
148 μl H₂O
2 μl 100 \times BSA
20 μl 10 \times NEBuffer 4
10 μl 10 U/ μl *NlaIII*.

Mix and incubate 1 hr at 37°C.

It is best to proceed with prewashing the streptavidin-Dynabeads (step 11) during this incubation such that the beads will be ready for use in the subsequent steps.

11. Thoroughly resuspend Streptavidin Dynabeads, exercising care to avoid excessive vortexing as streptavidin may be sheared off the magnetic beads. Transfer 200 μl beads to a No-stick siliconized 1.5-ml microcentrifuge tube and place in a magnetic rack. After ~ 1 min remove supernatant. Wash beads twice in 200 μl of 1 \times BW then let stand in 200 μl of 1 \times BW until ready for use (up to several hours).

All manipulations with Dynabeads are done using siliconized microcentrifuge tubes to avoid loss of yield due to products sticking to tube walls. All other manipulations, especially ethanol precipitations, should be done in standard microcentrifuge tubes.

Dynabead washes are executed in the same fashion as done in the primary method (see Basic Protocol 1, step 2). Briefly, the beads are placed in the magnet 1 to 2 min. While the siliconized tube is still in the magnet, the buffer is gently pipetted off. The tube is then taken off the magnet and fresh buffer/wash is added to the tube and the beads are resuspended by agitation by hand or gentle vortexing. It is critical that the Dynabeads are not allowed to dry between the wash steps.

12. Ethanol precipitate cDNA from step 10 as described in step 8. Resuspend cDNA pellet in 200 μl of 1 \times BW.
13. Remove 1 \times BW from Dynabeads (step 11) and replace with 200 μl cDNA in BW. Mix gently by pipetting the mixture up and down. Incubate 15 min at room temperature with intermittent agitation by hand. Wash three times with 200 μl of 1 \times BW. Add 200 μl of 1 \times T4 DNA ligase buffer.

14. Prepare the following mix:

2 μ l 200 ng/ μ l linkers 2A and 2B (annealed)
28 μ l LoTE
8 μ l 5 \times T4 DNA ligase buffer.

Remove 1 \times ligase buffer from the Dynabeads by pipetting and add the above mixture.

15. Mix bead slurry bound with cDNA gently, but well. Heat the tube 2 min at 50°C then incubate 15 min at room temperature.

16. Add 2 μ l of 5 U/ μ l (high-concentration) T4 DNA ligase and incubate 2 hr at 16°C. Mix beads intermittently during ligation.

It is best to use annealed linkers 2A/2B that are <1 month old.

Release cDNA with *AscI*

17. After ligation, wash beads four times with 1 \times BW/1 \times BSA. Wash in 1 \times NEBuffer 4/1 \times BSA and proceed immediately to the next step.

18. Resuspend the beads by adding the following components:

85 μ l LoTE buffer
10 μ l 10 \times NEBuffer 4
2 μ l 100 \times BSA
2 μ l 10 U/ μ l *AscI*.

Mix contents gently, but well using a pipette.

19. Incubate 1 hr at 37°C, agitating intermittently by hand every 15 min.

20. After digestion, collect supernatant carefully with a magnet. Place supernatant into a fresh nonsiliconized microcentrifuge tube. Add 50 μ l LoTE to sample. Extract with PC8 as described in step 7.

21. High-concentration ethanol precipitate by combining the following:

150 μ l sample
2 μ l SeeDNA
70 μ l 7.5 M ammonium acetate
500 μ l 100% ethanol.

Microcentrifuge 20 min at full speed, room temperature. Wash with 70% ethanol and resuspend in 25 μ l LoTE.

This is the concentrated rSAGE product, which may be stored indefinitely at -20°C. Avoid repeated freeze-thaw.

Amplify rSAGE-library dilutions by PCR

22. Make several dilutions of rSAGE product in LoTE.

Usually 1 μ l of 1:25, 1:50, and 1:100 dilutions are recommended for PCR. Due to frequent variations in yield, this can vary widely. These dilutions are good starting point, however.

23. Prepare the following PCR reaction:

1 μ l rSAGE dilution
5 μ l 10 \times SAGE PCR buffer
3 μ l DMSO
3 μ l 10 mM dNTPs

1 μ l 350 ug/ μ l M13 forward primer
 1 μ l 350 ug/ μ l primer 2
 36 μ l ddH₂O
 1 μ l 5 U/ μ l Platinum *Taq* DNA polymerase.

Repeat for all dilutions.

24. Use the following PCR cycling conditions:

Initial step:	2 min	94°C	(denaturation)
25 cycles:	45 sec	94°C	(denaturation)
	1 min	57°C	(annealing)
	1 min	70°C	(extension)
1 cycle:	5 min	70°C	(fill-in)
Final step:	indefinite	4°C	(hold).

25. Analyze 10 μ l of each PCR product on a 4% to 20% Novex TBE acrylamide gel along with 1-kb ladder. Stain with 1 \times SYBR Green I in TBE buffer for 30 min and visualize under UV light.

A smear predominantly in the 200 to 500 bp range should be observed. Choose the highest rSAGE dilution that gives reliable results. The authors usually use the amplified 1:50 dilution of the rSAGE product. Amplified rSAGE libraries may be stored at -20°C for months.

PCR amplify using SAGE tag-specific primer and M13F primer

26. Prepare the following PCR mixture per reaction:

1 μ l amplified rSAGE library (step 24)
 5 μ l 10 \times SAGE PCR buffer
 2.5 μ l DMSO
 3 μ l 10 mM dNTPs
 10 μ l 5 M betaine
 1 μ l 350 ug/ μ l M13 forward primer
 1 μ l 350 ug/ μ l SAGE tag-specific primer
 25.5 μ l H₂O
 1 μ l 5 U/ μ l Platinum *Taq*.

See Critical Parameters and Troubleshooting for a discussion of SAGE tag-specific primers.

27. Amplify under the following PCR cycling conditions:

Initial step:	2 min	93°C	(denaturation)
1 cycle:	30 sec	93°C	(begin touchdown)
	1 min	60°C	
	1 min	70°C	
15 cycles:	30 sec	93°C	(touchdown cycles)
	1 min	60° - 1°C/cycle	
	1 min	70°C	
30 cycles:	30 sec	93°C	(amplification cycles)
	1 min	44°C	
	1 min	70°C	
1 cycle:	5 min	70°C	(fill-in)
Final step:	indefinite	4°C	(hold).

These PCR cycling conditions are only guidelines that happen to work well for most SAGE tag-specific primers. A prolonged touchdown is pivotal for the specificity of priming.

Optimal annealing temperatures may vary depending upon the nucleotide makeup of the SAGE tag-specific primer. Therefore, the touchdown annealing temperature should begin at least 10°C above the predicted oligonucleotide melting point (T_m). Over the 15 touchdown cycles, the annealing temperature should, by -1°C increments, settle upon the predicted SAGE-tag-specific primer's annealing temperature, where the rest of the 30 amplification cycles will proceed. It is not advisable to go below an annealing temperature of 40°C, regardless of how low the oligonucleotide T_m might be. Despite the apparent numerous amplification cycles used in this prolonged touchdown approach, the Taq polymerase remains very much active, mostly attributable to the protective effects of high-concentration betaine. See Critical Parameters and Troubleshooting for further discussion.

28. Visualize 5 µl of the PCR products on a 1.5% TBE agarose gel (UNIT 2.5A).

The expected amplicons are usually between 100 to 400 bp, sometimes larger or smaller. Sometimes multiple bands may be amplified. If a control rSAGE amplified library was constructed, the band that is more intense in the experimental rSAGE library should be selected for further characterization. Often, multiple closely sized bands are amplification products of the same cDNA, attributable to variable oligo-dT priming along the poly-A tract during reverse transcription.

29. Load 25 µl of PCR products into a 1.5% TBE agarose gel and electrophorese until individual bands can be resolved (UNIT 2.5A). Carefully excise the amplicon in the smallest agarose piece possible without sacrificing yield and place into a preweighed microcentrifuge tube.

30. Purify PCR product using the Qiaquick gel-extraction kit according to manufacturer's instructions (Qiagen). Elute Qiaquick columns with 30 µl EB Buffer. Proceed immediately to cloning using 4 µl eluant and the TOPO TA Cloning Kit or Cloning Kit for Sequencing per manufacturer's instructions.

If the only goal for the rSAGE procedure is to sequence the cDNA fragment, then the standard TA cloning vector pCR2.1 (Invitrogen) should suffice. However, if there are future plans for in vitro transcription of the cloned cDNAs, then it is advisable to use the TA cloning vector pCR4-TOPO (Invitrogen), which has both T7 and T3 RNA polymerase recognition sequences flanking the multiple cloning site.

IMPORTANT NOTE: After TOPO TA cloning, do not use the M13 forward primer for subsequent colony PCR or cycle sequencing, as the M13 forward site will be embedded in the cloned cDNA. The M13 forward primer will not discriminate between M13 forward sites in the cDNA clone and the vector.

31. Sequence TA cloning products using conventional methods (e.g., UNIT 7.4A).

SUPPORT PROTOCOL 3

PHOSPHORYLATING AND ANNEALING LINKERS

It is critical that the linkers be both annealed into double-stranded products and efficiently phosphorylated prior to ligation onto *Nla*III-digested cDNAs during SAGE-library construction. Even if linkers are ordered prephosphorylated, it is critical to test the efficiency of linker phosphorylation by self-ligation prior to SAGE library construction so as not to lose precious time and material. The following protocol details linker phosphorylation, annealing, and self ligation.

Additional Materials (also see Basic Protocol 1)

- Linkers 1A, 1B, 2A, and 2B (see recipe)
- 10× kinase buffer (New England Biolabs)
- 10 mM ATP
- 10 U/µl T4 polynucleotide kinase (New England Biolabs; UNIT 3.10)

Phosphorylate linkers

1. If linkers 1B and 2B are not already phosphorylated on their 5' ends, prepare the following mixture:

9 μ l 350 ng/ μ l linker 1B or 2B
6 μ l LoTE buffer
2 μ l 10 \times kinase buffer
2 μ l 10 mM ATP
1 μ l 10 U/ μ l T4 polynucleotide kinase.

Incubate 30 min at 37°C, then heat inactivate 15 min at 65°C.

Anneal linkers

2. Add 9 μ l of 350 ng/ μ l linker 1A to 20 μ l phosphorylated linker 1B.
3. Add 9 μ l of 350 ng/ μ l linker 2A to 20 μ l phosphorylated linker 2B.
4. Perform the following incubations on each linker pair:

2 min at 95°C
10 min at 65°C
10 min at 37°C
20 min at room temperature.

5. Dilute to 2 ng/ μ l with LoTE prior to use in SAGE-library construction.

Perform and check ligation

6. Prepare the following ligation reaction:

0.5 μ l annealed undiluted 350 ng/ml linker 1A + phosphorylated linker 1B
(step 4)
0.5 μ l annealed undiluted 350 ng/ml linker 2A + phosphorylated linker 2B
(step 4)
7 μ l H₂O
1 μ l 10 \times T4 DNA ligase buffer
1 μ l 5 U/ μ l (high-concentration) T4 DNA ligase buffer.

Incubate 4 hr at 16°C.

All linkers, whether ordered prephosphorylated or phosphorylated in house, should be checked for self-ligation.

7. Analyze product on a prepoured 20% polyacrylamide/TBE gel. Visualize as described (see Basic Protocol 1, step 37).

Phosphorylated linkers should allow linker-linker dimers (80 to 100 bp) to form after ligation, while nonphosphorylated linkers will prevent self-ligation. Only linker pairs that self-ligate >70% should be used in further steps.

REAGENTS AND SOLUTIONS

Use double-distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

BRS1 primer

5'-Biotin-CCGGGCGCGCCGTAAAACGACGGCCAG(T)₁₉-3'

Order HPLC purified from a trusted supplier. The authors recommend using Integrated DNA Technologies (IDT).

BW buffer, 1×

For 2× stock:

10 mM Tris·Cl, pH 7.5 (APPENDIX 2)

1 mM EDTA

2.0 M NaCl

Store up to 1 year at room temperature

Dilute to 1× with H₂O just before use

Linkers

Linker 1A: 5' TTTGGATTTGCTGGTGCAGTACA ACTAGGCTTAATAGGGA-CATG 3'

Linker 1B: 5' TCCCTATTAAGCCTAGTTGTACTGCACCAGCAAATCC[amino mod C7] 3'

Linker 2A: 5' TTTCTGCTCGAATTCAAGCTTCTAACGATGTACGGGGA-CATG 3'

Linker 2B: 5' TCCCCGTACATCGTTAGAAGCTTGAATTCGAGCAG[amino mod C7] 3'

The authors recommend using Integrated DNA Technologies for ordering oligonucleotides.

LoTE buffer

3 mM Tris·Cl, pH 7.5 (APPENDIX 2)

0.2 mM EDTA, pH 7.5

Store up to 1 year at room temperature

PC8

480 ml phenol, warmed to 65°C

320 ml 0.5 M Tris·Cl, pH 8.0 (APPENDIX 2)

640 ml chloroform

Add in sequence and place at 4°C. After 2 to 3 hr, shake again. After an additional 2 to 3 hr, aspirate aqueous layer. Store up to 1 year in aliquots at -20°C or 6 months at 4°C.

Commercially available 1:1 (v/v) phenol:chloroform mix can also be substituted, as long as the pH is preset to 8.0.

PCR primers

Primer 1: 5' GGATTTGCTGGTGCAGTACA 3'

Primer 2: 5' CTGCTCGAATTCAAGCTTCT 3'

M13 forward: 5' GTAAAACGACGGCCAGT 3'

M13 reverse: 5' GGAAACAGCTATGACCATG 3'

The authors recommend using Integrated DNA Technologies for ordering oligonucleotides.

SAGE PCR buffer, 10×

166 mM ammonium sulfate

670 mM Tris·Cl, pH 8.8 (APPENDIX 2)

67 mM MgCl₂

100 mM 2-mercaptoethanol

Dispense into aliquots and store up to 1 year at -20°C.

SAGE tag-specific primer

5'-GACATGXXXXXXXXXX-(10-bp SAGE tag)-3'

If the SAGE-tag-specific primer has a calculated annealing temperature below 40° C, incorporate additional bases further 5' on linker 2A (see recipe for linkers) to increase the oligonucleotide melting temperature. The full linker 2A-SAGE tag sequence is as follows:

5'-TTTCTGCTCGAATTCAAGCTTCTAACGATGTACGGGGACATGXXXXXXXXXX-(10-bp SAGE tag)-3'

The SAGE 2000 software has the ability to extract an additional base for an 11-base tag. This may be helpful, as any additional sequence-specific bases may yield a more specific product.

Zeocin-containing low-salt LB plates

For 1 liter:

10 g tryptone

5 g yeast extract

5 g NaCl

Adjust the pH to 7.5 and add 15 g bactoagar. Autoclave solution and allow to cool before adding zeocin to 100 mg/ml.

COMMENTARY

Background Information

Serial analysis of gene expression (SAGE) was first developed in 1995 (Velculescu et al., 1995), and has since been used to generate a large variety of data from normal and cancerous human tissue (Zhang et al., 1997), yeast (Velculescu et al., 1997), mouse (Virlon et al., 1999), rat (Klimaschewski et al., 2000), and even (with modifications) human oocytes (Neilson et al., 2000).

SAGE is a powerful method for providing genome-wide gene-expression data. In much the same fashion as EST libraries, SAGE utilizes cDNA “tags” which are sequenced and quantified. The 14-bp SAGE tags differ from ESTs essentially by size, allowing subsequent concatenation and high-throughput sequencing in much greater volumes. The location of the anchoring enzyme site is essentially sufficient to uniquely identify the cognate cDNA or gene. The original protocol required relatively large amounts of starting material (2 to 5 µg of polyA mRNA) and was technically quite challenging, frequently giving variable results even in experienced hands. Major improvements were made to the protocol by a number of groups (Virlon et al., 1999; Datson et al., 1999; St. Croix et al., 2000), which collectively gave rise to a version of the protocol known as microSAGE (see Basic Protocol 1), owing to the fact that over 1000-fold less starting material could be readily used for library construction. The critical modifications appear to have been anchoring the mRNA to magnetic beads prior to cDNA synthesis (rather than after cDNA synthesis via incorporation of a biotinylated oligo(dT) primer as in the original protocol) and optimization of the quantities of reagents used, in particular, the quantities of linkers. Additional improvements, such as heating the ditag concatemers prior to gel purification (Angelastro et al., 1999), have resulted in SAGE libraries with substantially higher insert frequency and larger insert size than in the original protocol. These technical improve-

ments, coupled with the drop in the cost of DNA sequencing, have combined to allow the generation of over 3.5 million human SAGE tags alone, many of which are publicly available for analysis (<http://www.ncbi.nlm.nih.gov/SAGE>).

SAGE analysis has a number of unique advantages over hybridization-based measures of global gene expression, such as microarray analysis (Chapter 22), or approaches such as subtractive hybridization (UNITS 25B.1 & 25B.2) and differential display methodologies (UNITS 25B.3-25B.5). Since very few mRNAs lack *Nla*III sites, SAGE generates a tag for virtually every cellular mRNA, providing a level of coverage unequaled by any microarray yet available for humans or mice. For these same reasons, SAGE can also serve as a tool for gene discovery and transcript annotation even in species with fully sequenced genomes. The sensitivity of SAGE is limited only by the number of tags that one has the desire or resource to sequence and, with larger numbers of tags sequenced, it becomes possible to determine relatively small (<2-fold) changes in gene expression between samples. Since individual SAGE tag levels are expressed as a percentage of total tags, it is straightforward to compare tag levels among libraries generated by other labs. As more SAGE libraries are generated and made public, these data sets can be used to generate a large-scale atlas of gene expression that is of great use to the whole scientific community. Such a resource is already available for human normal and malignant tissues at NCBI (<http://www.ncbi.nlm.nih.gov/SAGE>), and libraries from other species are available from various sources (see Internet Resources for a partial list).

The main drawbacks of SAGE analysis are the time and expense required to generate sufficient numbers of tags to examine expression of low and moderate-abundance mRNAs. The price of sequencing has dropped considerably

Table 25B.6.1 Troubleshooting for SAGE Reactions

Problem	Possible cause	Solution
<i>MicroSAGE</i>		
No PCR product with control primers following cDNA synthesis	Dynabeads inactive	Store Dynabeads at 4°C only; do not freeze
	Reverse transcriptase inactive	Replace reverse transcriptase
	RNA degraded prior to homogenization	Minimize delay between tissue harvesting and homogenization
Ditag PCR product is slightly shorter (running at ~90 bp) and will not redigest with <i>Nla</i> III	Cells insufficiently lysed	Homogenize tissue thoroughly. Use homogenization by Polytron only.
	Failure to completely remove <i>E. coli</i> DNA polymerase I following second strand cDNA synthesis	Do not omit or shorten SDS washes or 75°C heat inactivation step
PCR product in no-ligase control at 100 bp	Contamination of reagent by ditags from a previously constructed SAGE library	Use separate aliquots of LoTE buffer, ammonium acetate, and PC8 for each large-scale ditag purification. Use aerosol pipet tips.
Ditag yield low (100-bp band <80-bp band)	Ratio of linkers to cDNA too high	Reduce amount of linkers in ligation
	cDNA synthesis inefficient	See advice in steps 1-11
Ditags do not cut with <i>Nla</i> III following purification	PCR conditions not optimized	Titrate dNTP concentration and number of amplification cycles
	Quantity of starting material too low	Increase amount of starting material
Ditags do not cut with <i>Nla</i> III following purification	<i>Nla</i> III inactive	Store enzyme in aliquots at -80°C. Do not reuse thawed aliquots.
	Ditags insufficiently pure	Run Qiaquick gel extraction on eluate (see step 44)
Ditag concatemers not generated efficiently	Insufficient quantity of purified ditags used in ligation	Increase quantity of cDNA used for large-scale ditag prep and/or increase number of cycles of amplification
	Ditags used in ligation are insufficiently pure	Run preparative gel longer to more efficiently separate 100- and 80-bp bands
Ditag concatemers most concentrated at high molecular weight (3kb) and do not clone efficiently	Concatemer ligation not heated properly	Heat at 65°C and chill on ice immediately
Concatemers have a high (5%) frequency of duplicate ditags	Too many cycles of PCR used to reamplify ditags	Reduce number of cycles used. Increase amount of starting material when making cDNA

continued

Table 25B.6.1 Troubleshooting for SAGE Reactions, continued

Problem	Possible cause	Solution
<i>rSAGE procedure</i>		
No SAGE tag-specific PCR product	Degraded RNA	Use fresh RNA
	Error in generation of rSAGE amplified library	Reconstruct rSAGE library
	Poor tag specific primer design	Redesign primer (see Critical Parameters and Troubleshooting)
Multiple SAGE tag-specific PCR products	Low abundance transcript	Increase the number of amplification cycles
	Multiple splice variants or multiple gene identities for a given SAGE tag	Clone all PCR amplicons
Poor sequence quality	Nonspecific priming	Construct a control rSAGE library and select amplicons not present in control library
	Use of M13 forward primer for cycle sequencing	Start the PCR touchdown cycles at a higher temperature
		Use another universal primer on the pCR vector for sequencing—e.g., M13 reverse, T3, T7

in the past few years, but real costs still remain around \$0.25/tag. For researchers simply hoping to identify a handful of differentially expressed genes in their sample of interest, subtractive hybridization (UNIT 25B.1), differential display methodologies (UNITS 25B.3-25B.5), or even the use of commercially available microarray technology may prove more cost-effective. An additional drawback of SAGE is the requirement that a large body of cDNA/EST sequence must be available from the organism being studied in order to match SAGE tags to the specific mRNAs. This effectively limits the use of SAGE to model organisms. Another drawback of the method is the occasional failure of a SAGE tag to match a predicted gene or to be long enough to easily isolate a full-length cDNA clone. While this happens at relatively low frequency for high abundance transcripts in model organisms, it can limit the interpretation of the data in some cases.

As a result, several approaches, most of which are variations of conventional RT-PCR, have been developed to identify these unknown or anonymous SAGE tags. There has been marked improvements in strategies used to identify unknown SAGE tags by reverse-cloning cDNA fragments, collectively called reverse SAGE (rSAGE; see Basic Protocol 2). First, the cloning process is similar to the original SAGE

protocol; therefore, only cDNA pieces which are 3' to the most 3' anchoring enzyme site are used as templates for subsequent PCR amplification and subcloning (Polyak et al., 1997; also see Internet Resources, SAGEnet). Second, the use of betaine allows for a prolonged PCR touchdown that results in more specific priming.

Critical Parameters and Troubleshooting

MicroSAGE

The two key determinants of a successful SAGE library are quantity and purity of ditags. To ensure obtaining many ditags, carefully optimize the starting reaction and scale up the number of PCR reactions as desired. For certain low-yield preparations, the authors have gone as high as 700 PCR reactions of 50 μ l to generate the starting material. For purity, ensure that the 102-bp and the 80-bp bands are well separated, and be very careful not to extract any of the 80-bp band. Run the gel as long as possible and do not overload the wells (no more than 10 μ l per well, despite the large number of gels this will require). Do the same for the 26-bp cut ditag band (avoiding the 40-bp linker band).

One other problem that has been encountered occasionally is contamination of reagents fol-

lowing construction of libraries, which will result in 102-bp bands in the no-ligase control in the initial optimization PCR reactions. To avoid this, be very careful to avoid splashes and not reuse tips during the scale-up or initial purification of the 102-bp band. Make separate aliquots of LoTE buffer, PC8, ammonium acetate, and ethanol for each library during these steps to reduce the likelihood of contamination. Use aerosol-barrier tips wherever possible.

A final common cause of experimental failure is low-quality reagents. Wherever possible, order supplies from the sources specified in the protocol. The authors have most frequently observed problems with the *Nla*III enzyme and the linkers. Always store *Nla*III in aliquots at -80°C , do not reuse aliquots, and try to have the enzyme shipped on dry ice if possible. The authors order linkers prekinased, but always check via self-ligation to ensure that a sufficiently large fraction of the linkers is properly phosphorylated.

rSAGE

For the rSAGE procedure, much depends on the quality of RNA used in the sample. It would be best to use the same batch of RNA that was originally used to construct the SAGE library. As most interesting SAGE tags are those that are expressed in abundance in one RNA sample and not in another, it is advisable to make a reverse-SAGE library of such a control tissue. It is not uncommon to generate multiple PCR bands from a tag-specific rSAGE amplification. Identifying a PCR product that is specific to the experimental rSAGE library and not present (or less apparent) in the control would help in the cloning and identification process.

The most technically challenging aspect of reverse-cloning SAGE tags is the PCR of a specific cDNA with the tag-specific primer. The rSAGE-amplified library used as a template for this PCR reaction consists solely of 3'-cDNA ends which have the linker2-SAGE tag on the 5' end and an oligo dT-M13 forward sequence on the 3' end. The PCR of a specific product is difficult when the reverse primer (M13 Forward) anneals to all templates, and the forward primer (SAGE-tag specific) shares the same sequences on the 5' end. Specificity is conferred only by the last 10 bases on the forward primer, representing the unique 10-base SAGE tag. One may also choose to incorporate an additional SAGE-tag base, information that the SAGE 2000 software can extract from the raw data. The SAGE tag-specific PCR is executed with a prolonged touchdown using an automatic hot-

start *Taq* polymerase (i.e., Platinum *Taq*; Invitrogen). As a 15-cycle touchdown requires 46 denaturing cycles, betaine is used as a *Taq* polymerase protectant. The authors strongly advise against switching to a proofreading DNA polymerase, such as *Pfu* or *Vent*, in the PCR reactions. Proofreading enzymes have significant 3'-5' exonuclease activity which may digest the 3' end of the SAGE tag-specific primer. Even one-base differences may reduce the specificity of the PCR product.

Designing of SAGE tag-specific primers is a matter of much debate. Only the 3'-most ten bases of the oligonucleotide contains tag-specific sequences, and the rest of the primer at the 5' end consists of linker sequences which are shared by all the cDNAs in the amplified rSAGE library. As a result, the authors empirically use CACATG-XXXXXXXXXX as a guideline for primer design where the Xs refer to the specific sequence in the SAGE tag of interest. Only six bases are nonspecific, and the relatively low annealing temperatures allow for an extended touchdown starting at a temperature that is well above the oligonucleotide melting point. However, if the rSAGE-specific primer has an annealing temperature which is too low, there is a risk of the primers melting off the template before the extension cycle. Therefore, if the calculated T_m of the SAGE tag specific primer is below 40°C , it is advisable to incorporate more of the linker sequence to raise the melting temperature of the oligo.

In the rare case that the SAGE tag in question lies immediately 5' to the polyA tail, reverse-SAGE may yield no additional information, and the PCR product may be too small to adequately visualize on a 1.5% agarose gel.

Additional troubleshooting guidelines are presented in Table 25B.6.1.

Anticipated Results

If Basic Protocol 1 is followed closely, libraries containing >85% inserts with an average size of 30 to 50 tags (450 to 750 bp) should be routinely generated. This should enable one to obtain a SAGE data set of 50,000 tags after ~2000 individual sequencing reactions.

If the above guidelines for rSAGE (see Basic Protocol 2) are followed, one should be able to clone the cDNA, usually 75 to 400 bp, from which a given SAGE tag is generated. This cDNA fragment would stretch from the 3'-most anchoring-enzyme site to the poly-A tail. The additional sequence data can be used to BLAST genome databases (UNIT 19.3) or be used to generate primers for 5' RACE (UNIT 15.6). The cloned

fragment may also be used for northern analyses (UNIT 4.9) or in-situ hybridizations (UNIT 14.3).

Time Considerations

MicroSAGE

The time typically taken for RNA preparation through *Bsm*FI digestion is 10 to 14 hr. Blunt-ending and ditag-ditag ligation take 2 to 3 hr. Ditag amplification and PCR optimization take 2 to 3 hr and large-scale ditag amplification and purification take 6 to 8 hr/day for 2 days. Ditag digestion and purification take 6 to 8 hr. Concatemer formation, purification, and subcloning take 6 to 8 hr. Template cleanup and transformation take 4 to 6 hr. PCR of library clones and gel analysis take 4 to 5 hr.

If a high-quality SAGE library is produced, it will require ~2000 sequencing reactions to obtain 50,000 tags. This will take anywhere from an additional 1 week to 3 months, depending on the resources and sequencing capacity.

rSAGE

Generating purified double-stranded cDNA typically takes 4.5 hr. Cleaving the cDNA with the anchoring enzyme (*Nla*III), magnetic bead purification, ligating linkers to cDNA, and release of 3' cDNA fragments from magnetic beads with *Asc*I typically takes 6 to 8 hr. PCR generation of amplified rSAGE libraries takes 2.5 to 3.5 hr. SAGE tag-specific PCR takes 3.5 to 4.5 hr. TOPO-TA cloning and subsequent sequencing is user-dependent.

Literature Cited

Angelastro, J.M., Kenzelmann, M., and Muhlemann, K. 1999. Substantially enhanced cloning efficiency of SAGE (serial analysis of gene expression) by adding a heating step to the original protocol. *Nucl. Acids Res.* 27:917-918.

Datson, N.A., van der Perk-de Jong, J., van den Berg, M.P., de Kloet, E.R., and Vreugdenhil, E. 1999. MicroSAGE: A modified procedure for serial analysis of gene expression in limited amounts of tissue. *Nucl. Acids Res.* 27:1300-1307.

Klimaschewski, L., Tang, S., Vitolo, O.V., Weissman, T.A., Donlin, L.T., Shelanski, M.L., and Greene, L.A. 2000. Identification of diverse nerve growth factor-regulated genes by serial analysis of gene expression (SAGE) profiling. *Proc. Natl. Acad. Sci. U.S.A.* 97:10424-10429.

Neilson, L., Andalibi, A., Kang, D., Coutifaris, C., Strauss, J.F. 3rd, Stanton, J.A., and Green, D.P. 2000. Molecular phenotype of the human oocyte by PCR-SAGE. *Genomics* 63:13-24.

Polyak, K., Xia, Y., Zweier, J.L., Kinzler, K., and Vogelstein, B. 1997. A model for p53 induced apoptosis. *Nature* 389:300-305.

St. Croix, B., Rago, C., Velculescu, V., Traverso, G., Romans, K.E., Montgomery, E., Lal, A., Riggin, G.J., Lengauer, C., Vogelstein, B., and Kinzler, K.W. 2000. Genes expressed in human tumor endothelium. *Science* 289:1197-1202.

Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. 1995. Serial analysis of gene expression. *Science* 270:484-487.

Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Hieter, P., Vogelstein, B., and Kinzler, K.W. 1997. Characterization of the yeast transcriptome. *Cell* 88:243-251.

Virion, B., Cheval, L., Buhler, J.M., Billon, E., Doucet, A., and Elalouf, J.M. 1999. Serial microanalysis of renal transcriptomes. *Proc. Natl. Acad. Sci. U.S.A.* 96:15286-15291.

Zhang, L., Zhou, W., Velculescu, V.E., Kern, S.E., Hruban, R.H., Hamilton, S.R., Vogelstein, B., and Kinzler, K.W. 1997. Gene expression profiles in normal and cancer cells. *Science* 276:1268-1272

Internet Resources

<http://www.sagenet.org>

SAGEnet. Contains instructions for obtaining SAGE analysis software, downloadable SAGE libraries from human, mouse and yeast, and a comprehensive bibliography of SAGE papers.

<http://www.ncbi.nlm.nih.gov/SAGE>

Serial analysis of gene expression at NCBI.

<http://www.ncbi.nlm.nih.gov/CGAP>

Cancer Genome Anatomy project. Contains full downloadable predicted tag data for human, mouse, rat, zebrafish, and cow. Also contains a large number of downloadable human SAGE libraries (containing >3.5 million total tags), as well as tools for submitting SAGE data for public access and tools for searching tag abundance levels in the publicly available human SAGE data.

<http://www.umich.edu/~ehm/eSAGE>

eSAGE at University of Michigan. Helpful software for SAGE data analysis.

<http://www.invitrogen.com>

iSAGE at Invitrogen. Integrated kit and software package for conducting microSAGE. The protocol used is very similar to the one described here.

<http://arep.med.harvard.edu/labgc/adnan/projects/Utilities/mergesagetags.html>

Merge SAGE tags at Harvard Medical School. Helpful tool for merging SAGE data files and downloaded predicted tag identify files (from NCBI).

Contributed by Seth Blackshaw, Jae B. Kim, Brad St. Croix, and Kornelia Polyak
Harvard Medical School
Boston, Massachusetts

Representational Difference Analysis

UNIT 25B.7

This unit provides a protocol for performing representational difference analysis (RDA); a technique that couples subtractive hybridization to PCR-mediated kinetic enrichment for the detection of differences between two complex genomes. RDA requires the generation of representations from two pools of nearly identical DNA varying only in polymorphisms, deletions/amplifications, rearrangements, or exogenous pathogens. A representation or subset of the genome is used rather than the entire genome, since the full complexity of genomic DNA is unfavorable for hybridization to proceed to completion. In its original formulation by Lisitsyn and colleagues (1993), 2% to 15% of the genome is included in the representation, the percentage being dependent on the frequency of restriction endonuclease sites and the efficiency of PCR amplification of these restriction-generated fragments. While RDA was first developed for genomic DNA, subsequent modifications have been devised to look for differences in transcript expression.

RDA starts with the digestion of two comparison samples of DNA (see Basic Protocol 1) or cDNA (see Basic Protocol 2) with a frequently cutting restriction enzyme. Some consideration should be given to which of the two genomes is designated tester and which is designated driver. In principle, the tester should contain DNA restriction fragments not found in the driver. Specific linkers are ligated to DNA restriction fragments from each pool and amplicons are generated by PCR. Linkers are then removed from both samples and a new linker is added only to size-selected tester amplicons. These tester amplicons are mixed and melted with a large excess of driver amplicons lacking linkers. Hybridization between complementary single strands is allowed to proceed, resulting in the generation of three species of double-stranded DNA fragments: (1) both strands derived from driver DNA (lacking linkers on either strand), (2) hybrids with one strand from driver (no linker) and one from tester (with linker), and (3) both strands from tester DNA (linkers on both strands). Excess driver will soak up DNA fragments common to both samples (i.e., tester:driver), and only the DNA fragments unique to the tester (i.e., the tester:tester population) will be exponentially amplified and kinetically enriched when linker-specific primers are used. Iterative rounds of subtractive/kinetic enrichment against driver amplicons is performed until distinct difference products can be cloned.

GENOMIC REPRESENTATIONAL DIFFERENCE ANALYSIS

This protocol describes RDA for genomic DNA derived from tissues or cells. Modifications for performing cDNA RDA are discussed below (see Basic Protocol 2).

Materials

- Tester and driver DNA samples
- Phenol (Amresco; UNIT 2.1A)
- Phenol:chloroform:isoamyl alcohol (Amresco; UNIT 2.1A)
- 20 µg/µl glycogen
- TE buffer, pH 8.0 (APPENDIX 2)
- Primers/oligomers, HPLC purified (Table 25B.7.1)
- 400 U/µl T4 DNA ligase and 10× buffer (New England BioLabs; UNIT 3.14)
- 5× RDA PCR buffer (see recipe)
- dNTP chase solution: 4 mM (each) dGTP, dATP, dTTP, dCTP; store at -20°C
- 5 U/µl *Taq* DNA polymerase (Invitrogen; UNIT 3.5)
- Mineral oil
- Isopropanol
- 10 M ammonium acetate (APPENDIX 2)

**BASIC
PROTOCOL 1**

**Discovery of
Differentially
Expressed Genes**

25B.7.1

Contributed by Yuan Chang

Current Protocols in Molecular Biology (2002) 25B.7.1-25B.7.12

Copyright © 2002 by John Wiley & Sons, Inc.

Supplement 60

Table 25B.7.1 Prototypic Primers Used in RDA

Primer Type	Name ^a	Sequence ^b
Representation		
24-mers	RBgl24	5'-AGCACTCTCCAGCCTCTCACCGCA-3'
	RBam24	5'-AGCACTCTCCAGCCTCTCACCGAG-3'
	RHind24	5'-AGCACTCTCCAGCCTCTCACCGCA-3'
	RXxx24	5'-AGCACTCTCCAGCCTCTCACCGxx-3'
12-mers	RBgl12	5'-GATCTGCGGTGA-3'
	RBam12	5'-GATCCTCGGTGA-3'
	RHind12	5'-AGCTTTCGGTGA-3'
	RXxx24	5'-xxxxxx CGGTGA -3'
Odd cycle		
24-mers	OBgl24	5'-ACCGACGTCGACTATCCATGAACA-3'
	OBam24	5'-ACCGACGTCGACTATCCATGAACG-3'
	OHind24	5'-ACCGACGTCGACTATCCATGAACA-3'
	OXxx24	5'-ACCGACGTCGACTATCCATGAACx-3'
12-mers	OBgl12	5'-GATCTGTTTCATG-3'
	OBam12	5'-GATCCGTTTCATG-3'
	OHind12	5'-AGCTTTCGTTTCATG-3'
	OXxx24	5'-xxxxxx GTTTCATG -3'
Even cycle		
24-mers	EBgl24	5'-AGGCAACTGTGCTATCCGAGGGAA-3'
	EBam24	5'-AGGCAACTGTGCTATCCGAGGGAG-3'
	EHind24	5'-AGGCAACTGTGCTATCCGAGGGAGA-3'
	EXxx24	5'-AGGCAACTGTGCTATCCGAGGGAx-3'
12-mers	EBgl12	5'-GATCTTCCCTCG-3'
	EBam12	5'-GATCCTCCCTCG-3'
	EHind12	5'-AGCTTCTCCCTC-3'
	EXxx12	5'-xxxxxx TCCCTCG -3'

^aR primers are used only in making representations of the tester and driver DNAs. The O and E primers are used in odd and even iterations of the subtractive/enrichment process. These were previously designated J and N in the original protocol (Lisitsyn et al., 1993).

^bUnderscores indicate restriction sites that are variable, but limited to those comprising restriction sites (i.e., can be changed to accommodate other enzymes). Nucleotides shown in bold outline invariant core sequences of the primers. Nucleotides which are neither bold nor underscored are completely variable.

100% ethanol, ice cold

70% ethanol, room temperature

3 M sodium acetate, pH 5.2 (APPENDIX 2)

EE × 3 hybridization buffer (see recipe)

5 M NaCl

5 µg/µl glycogen in TE buffer (see APPENDIX 2 for TE buffer)

10 U/µl mung bean nuclease and 10× buffer (New England BioLabs; UNIT 3.12)

50 mM Tris·Cl, pH 8.9 (APPENDIX 2)

Thermal Cycler (Perkin-Elmer Model 480 preferred)

24-mm GF/C glass microfibre filters (Whatman)

Dialysis tubing, 6,000 to 8,000 MWCO (Spectra/Pore)

Flat blunt forceps

18-G needle

Additional reagents and equipment for restriction digestion (UNIT 3.1), agarose gel electrophoresis (UNIT 2.5A), ethanol and isopropanol precipitation (UNIT 2.1A), and quantifying DNA by absorbance spectroscopy (APPENDIX 3D), gel isolation (UNIT 2.6), and sequencing (UNIT 7.1).

NOTE: Use de-ionized, distilled water in all recipes and protocol steps, and ensure that the water is RNase/DNase free. Since minute amounts of contaminating DNA may be detected by RDA, use barrier pipet tips throughout the protocol.

Prepare amplicons and representation: Enzyme restriction of tester and driver DNA

1. Using 10 U enzyme per microgram DNA and a total volume of 200 μ l (each), separately digest 5 μ g tester and driver DNA samples with the restriction enzyme chosen for representation (UNIT 3.1). Analyze 40 μ l (1 μ g) of each reaction by electrophoresis on a 1% agarose gel (UNIT 2.5A) to confirm complete digestion. Bring volume of remaining digest to 400 μ l each with water.

This step provides three to four times the DNA needed for the preparation of amplicons; therefore, if the amount of starting DNA is a limiting factor, as little as 1 to 2 μ g DNA can be used.

BglII, BamHI, and HindIII are the enzymes which were used in the original RDA publication by Lisitsyn and colleagues (1993). Oligomer/primers compatible with each of these enzymes are listed in Table 25B.7.1. These enzymes with corresponding oligomers have been extensively and successfully used in RDA applications; however, other enzymes may be used by adapting the restriction sites adjacent to the core sequences. In particular, four-base cutters may be more appropriate for less complex genomes.

2. Extract digested tester and driver with 1 vol phenol (400 μ l each) followed by 1 vol phenol:chloroform:isoamyl alcohol (400 μ l each). Ethanol precipitate DNA (UNIT 2.1A), adding 20 μ g glycogen and microcentrifuging at 4°C to increase recovery. Dry pellets and resuspend at 0.1 μ g/ μ l in TE buffer, pH 8.0, instead of water in the final step. Confirm DNA concentration by comparison to dilution of known standards by agarose gel electrophoresis (UNIT 2.5A).
3. Resuspend HPLC-purified primers/oligomers in water at 62 pmol/ μ l, an OD₂₆₀ of 6 or 12 AU/ml for 12- and 24-mers, respectively (APPENDIX 3D).

HPLC purification of oligomers is critical for minimizing false positive RDA bands (O'Neill and Sinclair, 1997).

Ligate adapters onto driver and tester DNA

4. Mix the following in thermal cycler tubes colored differently for tester and driver DNA:

2 μ l water
3 μ l 10 \times ligase buffer
7.5 μ l 12-mer (R primer)
7.5 μ l 24-mer (R primer)
10.0 μ l (1 μ g) driver or tester DNA digest
30 μ l total volume.

The use of tubes of different colors throughout the protocol helps distinguish between driver and tester samples to avoid confusion and cross-contamination of DNA.

5. Place tubes in a thermal cycler at 55°C. Program the thermal cycler to decrease the temperature to 4°C over 1 hr.

Slow annealing allows the 12- and 24-mers to form a temporary bridging complex with cohesive ends complementary to the restriction sites on the ends of the digested DNAs.

The Perkin-Elmer Model 480 is preferred because of its larger tube capacity, but any 96-well thermal cycler may also be used.

6. Add 1 μl of 400 U/ μl T4 DNA ligase, mix by gentle pipetting, and incubate 12 to 16 hr at 14°C.

This step results in ligation of the 24-mers onto the 5' ends of the DNAs. The temperature used is below the T_m of the four base duplexes formed by the overhanging ends.

7. Transfer ligation product to 1.5-ml microcentrifuge tubes matching the colors used above (step 4). Dilute adapter-ligated tester and driver DNA to 1 ng/ μl by adding 970 μl TE buffer.

PCR-amplify driver and tester amplicons

8. Prepare two tubes of PCR mix for preparation of tester amplicon and twelve tubes for driver amplicon, each containing:

240 μl water
80 μl 5 \times RDA PCR buffer
32 μl dNTP Chase solution
8 μl 24-mer oligonucleotide (R primer)
360 μl total volume.

9. Add 40 μl diluted adapter-ligated tester or driver DNA (40 ng) to corresponding PCR tubes (two for tester and twelve for driver) and place tubes in a thermal cycler 1 to 2 min at 72°C.
10. Fill-in 3'-recessed ends of the ligated adapters by adding 3 μl of 5 U/ μl (15 U) *Taq* DNA polymerase to each tube, mix by pipetting, and overlay with 110 μl mineral oil. Incubate 5 min at 72°C.

If using a 96-well PCR machine, double the number of tubes and halve the amount of PCR mixture for each tube such that four tubes of tester and twenty-four tubes of driver amplicon are made. With the 96-well PCR machine, no mineral oil is required. Do not let the tubes cool below 72°C in steps 9 or 10.

11. Perform the following two-step PCR program:

20 cycles:	1 min	95°C	(denaturation)
	3 min	72°C	(extension)
Final step:	10 min	72°C	(extension).

Quantitate amplicons and remove linkers

12. Pipet off as much mineral oil as possible. Combine the contents of both tester PCR tubes into a single 1.5-ml microcentrifuge tube. Combine driver tubes pairwise into single microcentrifuge tubes (i.e., six driver tubes total).

For the 96-well PCR format, combine the contents of four PCR tubes into a single 1.5-ml microcentrifuge tube.

13. Extract each tube with 1 vol phenol followed by 1 vol phenol:chloroform:isoamyl alcohol, isopropanol precipitate with 20 μg glycogen, and dry the pellets (*UNIT 2.1A*).
14. Resuspend driver and tester amplicons in TE buffer at a concentration between 0.2 to 0.4 $\mu\text{g}/\mu\text{l}$ (expecting ~ 15 μg of DNA from each 0.5-ml PCR tube). Pool driver DNA into a single tube. Confirm concentrations of driver and tester DNA by agarose gel electrophoresis (*UNIT 2.5A*) against DNA standards.

Enough of the driver amplicon needs to be prepared to provide sufficient amounts of DNA such that all rounds of hybridization use aliquots that are identical and derived from the same source. Calculate the total amount of driver DNA needed for the experiment (~ 40 $\mu\text{g}/\text{round}$) and if necessary, scale up driver amplicon production or perform additional driver amplicon amplifications and pool.

15. Digest 150 µg driver DNA and 15 µg tester DNA with initially chosen restriction endonuclease (step 1) in volumes of 400 µl to remove the adapters.
16. Repeat step 2 and resuspend in 125 µl TE buffer.

Expect the concentration of tester to be ~0.1 µg/µl and that of driver to be ~1 µg/µl.
17. Dilute 2 µl resuspended driver amplicon digest with 18 µl water to an expected concentration of 0.1 µg/µl. Load 0.2, 0.4, and 0.6 µg driver and tester amplicon digests and compare with DNA standards by 2% agarose gel electrophoresis (UNIT 2.5A). Using electrophoresis results as a guide, perform final dilution with TE buffer such that the driver amplicon digest concentration is 0.5 µg/µl and the tester amplicon digest concentration is 50 ng/µl.

Change adapters on tester amplicon

18. Load 5 µg (100 µl) tester amplicon DNA digest on a 1% agarose gel (UNIT 2.5A). Electrophorese at appropriate voltage until DNA in the range from 150 to 1500 bp can be resolved.
19. With a clean razor blade, cut two full thickness slits in the running lanes, one at 150 and another at 1500 bp.
20. Soak small pieces of 24-mm GF/C glass microfibre filter and 6,000- to 8,000-MWCO dialysis tubing in water. Make a two-layer barrier of filter and dialysis tubing and cut into rectangles slightly higher and wider than the agarose lane. Using a blunt flat forceps, insert the filter/dialysis tubing barrier into each of the slits with the filters facing the loading wells.

Be sure that the entire running lane is blocked by both the filter as well as the dialysis tubing.
21. Resume electrophoresis until DNA between 150 and 1500 bp has migrated onto the filter/dialysis tubing. Stop the electrophoresis and carefully remove the DNA embedded filter/dialysis tubing from the 150-bp slit.

DNA larger than 1500 bp should be blocked from migrating past the filter/dialysis membrane in the 1500-bp slit.

In the author's hands, this method gives better recovery than gel isolation and elution.
22. Cut the lid off a 0.5-ml PCR tube and puncture a hole in the bottom with an 18-G needle so that DNA can elute. Make a collecting apparatus comprised of the PCR tube placed inside a 1.5-ml microcentrifuge tube.
23. Place the filter/dialysis membrane into the PCR tube of the collecting apparatus. Microcentrifuge the collecting apparatus 5 min at 8,000 rpm, room temperature.
24. Discard the PCR tube and filter/dialysis tubing. Bring up volume of collected liquid to 400 µl with water and extract with 1 vol phenol followed by 1 vol phenol:chloroform:isoamyl alcohol. Ethanol precipitate DNA with 20 µg glycogen and dry the pellet as described (step 2).
25. Dissolve the DNA pellet in 30 µl TE buffer, check DNA concentration by agarose gel electrophoresis against DNA standards (UNIT 2.5A), and adjust the concentration to 0.1 µg/µl.
26. Ligate 1 µg purified tester amplicon DNA digest to primer set O, as described in steps 4 to 6 above.

The R set of primers used to make the driver amplicons is never used in subsequent subtractive/kinetic enrichment rounds to prevent driver amplification as a result of uncleaved primers.

27. Dilute the pellet to a concentration of 10 ng/μl by adding 70 μl TE buffer

If using HindIII, dilute the pellet to 25 ng/μl by adding 10 μl TE buffer.

Perform subtractive/kinetic enrichment

28. In a microcentrifuge tube, combine driver and tester by mixing 80 μl driver amplicon DNA digest (0.5 μg/μl) and 40 μl diluted tester amplicon ligate (0.4 μg for representations made with most six cutters or 1 μg for *HindIII* representation).

The first hybridization is done at a tester:driver ratio of ~1:100.

29. Extract once with 120 μl phenol:chloroform:isoamyl alcohol.
30. Ethanol precipitate DNA with ammonium acetate as follows:
- Add 30 μl of 10 M ammonium acetate and mix by pipetting.
 - Add 300 μl (2 vol) ice-cold 100% ethanol.
 - Add 1 μl (20 μg) glycogen and mix by inverting.
 - Chill 10 min at -70°C.
 - Microcentrifuge 20 min at 13,000 rpm, room temperature.
 - Carefully remove the supernatant.
 - Wash the pellet with 1 ml room-temperature 70% ethanol.
 - Dry pellet.
31. Add 4 μl EE × 3 hybridization buffer to the pellet. Resuspend by pipetting, incubate 5 min at 37°C, vortex 2 min, and then microcentrifuge at maximum speed to collect the sample at the bottom.
32. Transfer resuspended DNA to a PCR tube. In another PCR tube, add 1 μl of 5 M NaCl. Place both tubes in a thermal cycler preheated to 95°C and incubate 1 min. Centrifuge the tubes briefly to collect the contents at the bottom and immediately transfer the denatured DNA to the tube containing NaCl. Mix well by pipetting and overlay with 35 μl mineral oil.
33. Incubate the tube containing DNA and NaCl in the thermal cycler for an additional 4 min at 95°C to ensure that all DNA species are denatured.
34. Set the thermal cycler to hold >20 hr at 67°C. Incubate at least 18, but not more than 48 hr, to allow the DNAs to hybridize to complementary strands.

As a result of the vast excess of driver, the majority of fragments common to both the driver and tester populations will rapidly form driver:driver or tester:driver complexes. The fragments unique (or at a relatively higher quantity) in the tester will require a significantly longer period of time to completely hybridize and form tester:tester complexes.

Perform selective amplification

35. Remove as much of the mineral oil as possible without losing the hybridizing mixture. Dilute the DNA stepwise to a concentration of 0.1 μg/μl by first adding 8 μl of 5 μg/μl glycogen in TE buffer and mixing by pipetting, then adding 23 μl TE buffer and again mixing by pipetting, and finally adding 364 μl TE buffer and vortexing.
36. To fill-in the adapter ends, make two tubes of PCR mix (not containing 24-mer):

235 μl water
80 μl 5× RDA PCR buffer
32 μl dNTP chase solution
347 μl total volume.

Add 40 μ l diluted hybridized DNA (4 μ g) to each tube. Place tubes in thermal cycler set at 72°C.

This reforms priming sites at both ends of tester:tester complexes necessary for exponential amplification of difference products.

37. Add 3 μ l *Taq* DNA polymerase, mix by pipetting, and incubate an additional 5 min.

38. Add 10 μ l 24-mer primer (O primer set), mix by pipetting, and overlay with mineral oil.

If using a 96-well thermal cycler, double the number of tubes and halve the PCR recipe in each tube. In this case, addition of mineral oil is not necessary.

39. Perform the following two-step PCR program:

10 cycles:	1 min	95°C	(denaturation)
	3 min	72°C	(extension)
Final step:	10 min	72°C	(extension).

For the OBgl 24 primer, a lower annealing temperature of 70°C is required.

40. Remove as much mineral oil as possible and combine the contents of the PCR tubes in a microcentrifuge tube. Extract and isopropanol precipitate as described (step 13), but dissolve the pellet in 40 μ l water and do not pool DNA.

41. Digest single-stranded templates with mung bean nuclease (MBN) by mixing:

- 14 μ l water
- 4 μ l 10 \times mung bean nuclease buffer
- 20 μ l amplified difference product
- 2 μ l 10 U/ μ l mung bean nuclease (MBN)
- 40 μ l total volume.

Incubate at 30°C for 30 min.

42. Add 160 μ l of 50 mM Tris·Cl, pH 8.9. Inactivate MBN by incubating 5 min at 98°C.

43. Prepare two tubes of PCR mix (360 μ l) containing the O 24-mer primer as in step 8. Add 40 μ l MBN-treated difference product in each tube and place in a thermal cycler set at 72°C.

For OBgl 24-mer use an annealing temperature of 70°C.

44. Add 3 μ l of 5 U/ μ l (15 U) *Taq* DNA polymerase to each tube, mix by pipetting, overlay with 110 μ l mineral oil, and incubate 5 min at 72°C.

Again, double the number of PCR tubes and halve the given recipe placed in each tube if using a 96-well PCR machine.

45. Perform the following two-step PCR program:

20 cycles:	1 min	95°C	(denaturation)
	3 min	72°C	(extension)
Final step:	10 min	72°C	(extension).

For the OBgl 24 primer, a lower annealing temperature of 70°C is required.

46. Run 10 μ l amplified product on a 2% agarose gel with DNA concentration standards (UNIT 2.5A). If necessary to improve the yield, perform 1 to 3 more cycles after addition of 3 μ l fresh *Taq* DNA polymerase.

The quantity of DNA should be between 0.1 to 0.3 µg.

In subsequent iterations of this step, discrete products should be observed. Alternatively, the results of the agarose gel may suggest strategies for interventional troubleshooting (see Commentary). For example, a high background may indicate either primer hydrolysis or the need for increasing the stringency of the preceding hybridization step (i.e., decreasing tester relative to driver).

Change adapter on the difference product

47. Combine the contents of the two PCR tubes in one microcentrifuge tube (four tubes for the 96-well format). Extract and isopropanol precipitate as described in step 13.
48. Dissolve the pellet in 80 µl TE buffer. Determine DNA concentration by 2% agarose gel electrophoresis (UNIT 2.5A), and adjust to 0.1 µg/µl.
49. Digest 5 µg difference product (50 µl) with 10 U/µg chosen restriction enzyme (step 1) in a total volume of 200 µl. Bring volume of digested product up to 400 µl with water.
50. Extract and ethanol precipitate as described (step 2).
51. Resuspend DNA pellet at 0.1 µg/µl in TE buffer. Take 10 µl (1 µg) DNA solution and directly ligate to primer set E in a volume of 30 µl as described in steps 4 to 6.

Changing primer sets between each round of RDA ensures that selective subtractive/kinetic enrichment of unique tester DNA restriction fragments will occur from newly ligated primer and not from uncleaved primer carried over from the previous rounds.

52. Dilute the ligated difference product to 1.25 ng/µl with TE buffer.

For HindIII representation, dilute to 2.5 ng/µl with TE buffer.

Always ligate 1 µg tester, then serially dilute the ligation product to a concentration such that 40 µl will give the appropriate amount of tester for the selected tester:driver hybridization ratio.

Perform subsequent subtractive/kinetic enrichment steps

53. For a second subtractive/kinetic enrichment, mix 40 µl (50 ng) adapter-ligated difference product (100 ng for *HindIII* representation) and 80 µl (40 µg) of driver amplicon DNA digest. Proceed through subtractive/kinetic enrichment exactly as outlined in steps 28 to 51 except substitute E for O primers/oligomers and dilute the ligated difference product to 2.5 pg/µl (10 pg/µl for *HindIII* representation).

*The second hybridization is done at a tester:driver ratio of 1:800 (1:400 for *HindIII* representations).*

54. For a third subtractive/kinetic enrichment, mix 40 µl (100 pg) difference product from the second subtractive/kinetic enrichment (400 pg for *HindIII* representation) and 80 µl (40 µg) driver amplicon DNA digest. Proceed exactly as outlined in steps 28 to 51 using O primers/oligomers.

*The third hybridization is done at a tester:driver ratio of 1:400,000 (1:200,000 for *HindIII* representations).*

55. *For HindIII:* Use 5 pg difference product from the third subtractive/kinetic enrichment (tester:driver ratio of 1:8,000,000). Again, proceed through steps 27 to 51 of the protocol, except substitute E for O primers/oligomers, and use 27 cycles in the final PCR of the selective amplification (step 44).

For HindIII representation sometimes this fourth subtractive/kinetic enrichment is needed.

56. Clone products following gel isolation (UNIT 2.6) or use shotgun cloning and subsequent sequencing (UNIT 7.1).

cDNA REPRESENTATIONAL DIFFERENCE ANALYSIS

cDNA RDA works under the same principles as RDA of genomic DNA, and requires only minor modification from the procedure described above (see Basic Protocol 1). Two RDAs may be performed at the same time with the testers and drivers reversed in order to detect both induced as well as suppressed transcripts. There are two other modifications to Basic Protocol 1. The first is the substitution of *DpnII* or its isoschizomer *Sau3AI*, as the restriction endonuclease. *DpnII* is a four-base recognition enzyme that is compatible with the *BglII* and *BamHI* primers listed in Table 25B.7.1. The second is the use of different ratios of tester to driver in the sequential hybridizations. For cDNA RDA, the ratios of 1:10, 1:100, 1:500, and 1:25000 may be used (see Table 25B.7.2 for ranges of tester:driver ratios; Pastorian et al., 2000).

Table 25B.7.2 Tester:Driver Hybridization Stringencies for cDNA RDA

Subtractive/kinetic enrichment	Range of tester:driver ratios
Round 1	1:10–1:50
Round 2	1:100–1:500
Round 3	1:1000–1:5,000
Round 4	1:10,000–1:50,000

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

***EE* × 3 hybridization buffer**

30 mM 4-(2-hydroxyethyl)-1-piperazinepropanesulfonic acid (EPPS), pH 8.0 at 20°C
3 mM EDTA
Store up to 6 months at room temperature.

***RDA PCR buffer*, 5×**

335 mM Tris·Cl, pH 8.8 at 25°C (APPENDIX 2)
20 mM MgCl₂
80 mM (NH₄)₂SO₄
50 mM 2-mercaptoethanol
0.5 mg/ml BSA
Store up to 6 months at –20°C.

COMMENTARY**Background Information**

Representational difference analysis (RDA) was first described in 1993 and has been used to detect polymorphisms between individuals, positional synteny between species, and genetic lesions in neoplasms (Lisitsyn et al., 1993; Lisitsyn and Wigler, 1995; Lowrey et al., 2000). In addition to finding genomic alterations, RDA has been successfully used to identify exogenous sequences from DNA-based infectious agents (Chang et al., 1994). While RDA

was original applied to genomic DNA, the versatility of the technique allowed minor modifications in the protocol for the examination of differences in gene expression (Hubank and Schatz, 1994; Bakin and Curran, 1999; Reick et al., 2001; Shields et al., 2001) as well as the identification of new RNA viruses (Nishizawa et al., 1997; Birkenmeyer et al., 1998).

RDA has advantages and limitations when compared to other techniques used to detect

differences in genomic content. The first generation technique of subtractive hybridization requires large amounts of starting DNA and is inefficient, usually allowing only a 1:100-fold enrichment of target sequences. This is due to the complexity of eukaryotic genomes in which hybridization of complementary sequences cannot go to completion. Therefore, only very long or abundant sequences can be isolated. RDA circumvents this problem by the incorporation of a simplification step in which only a representation of the genome is used in the analysis. The simplification process is based on restriction endonuclease digestion of the genome and is accomplished by selective amplification of digested DNA fragments with lengths amenable to PCR followed by physical size selection. This simplification is key to successful RDA, but its disadvantage is that not all of the potential differences between two genomes will be found.

Many techniques are available for scanning differential gene expression, whether to ascertain changes that occur in development and differentiation, or that are associated with disease phenotypes. These include differential display, (*UNIT 25B.3*) cDNA array, serial analysis of gene expression (SAGE; *UNIT 25B.6*), and rapid analysis of gene expression (RAGE). In a novel combination of two techniques, RDA is performed first to generate products used as hybridization probes which are then applied to cDNA microarrays (Geng et al., 1998). Consideration must be given to the strengths and weakness of each tool in individual applications. The main advantages of using RDA are that the analysis is not limited to known sequences, it is efficient, and it is affordable for even small laboratories.

Critical Parameters and Troubleshooting

General considerations

DNA RDA is dependent on the generation of different DNA restriction fragments between driver and tester after restriction endonuclease digestion. Furthermore, the extra DNA fragment(s) must be found in the tester and not the driver, and must be within the size range for standard PCR amplification. Therefore, if the targeted genetic change does not result in a unique DNA fragment after digestion, then the change cannot be detected. In the case of DNA RDA, it is critical that the two samples to be compared are extracted from tissues or cells of

nearly identical genetic background. To look for polymorphism, tissues from closely related individuals of the same gender may be used. To look for genetic changes associated with a neoplastic phenotype, tumor and normal tissue from the same individual is appropriately matched, unless the genetic change is germline. Although translocations may be identified whether the neoplastic tissue is used as the driver or tester, deletions require the neoplastic DNA to be used as driver. When the nature of the genetic change is not known, it is reasonable to perform two RDA with the samples switched from their designation as driver or tester.

Several issues arise when hunting for a microbial agent. The agent's genome must be large enough to offer a DNA fragment which when digested is big enough to PCR, and the genome must go through a DNA stage in its life cycle. RNA viruses must be pursued using cDNA RDA. Optimally, samples are acquired in a sterile manner and are free from contaminating organisms. In particular, epithelial or mucosal surfaces should be dissected off prior to DNA extraction. Diseases primarily involving such tissues are difficult to analyze by RDA unless existing microbial flora is matched. Lastly, the infected tissue should always be used as the tester, keeping in mind that the infection may be disseminated. In a related cautionary, when working with cell lines, ensure that no mycoplasma infection is present in cultures and that transformed cell lines are not generated by viral infections (i.e., herpesviruses, papillomaviruses, or adenoviruses).

The use of PCR in RDA necessitates implementation of procedures that guard against DNA contamination. If RDA is performed repetitively, all work areas and surfaces should be monitored regularly for occult adapter-ligated products. This can be done with swipe tests followed by PCR with the O and E 24-mers. PCR preparation, amplification, and analysis should be isolated from each other if possible, dedicated micropipettors should be used, and reagents should aliquoted and changed frequently.

Amplicon preparation

In both DNA and cDNA RDA, the quality of the starting material is important. Tissues or cells used to generate tester and driver DNA should be subjected to the same harvesting, storage, and DNA extraction conditions. Use methods for DNA preparation which give relatively pure DNA to ensure complete digestion.

The amount, the completion of digestion, as well as the integrity of DNA should be assessed by agarose gel electrophoresis to confirm that smears of tester and driver DNA in the initial steps prior to hybridization are comparable in both intensity and size distribution. Agarose gel electrophoresis is the preferred method for evaluating products in the protocol, since this method allows not only concentration determination, but also visualization of DNA integrity. When preparing cDNA, any standard protocol or kit may be used; however, be aware that some reverse transcriptases may contain minute amounts of contaminating vector which can give false positive results. To ensure the highest quality full-length cDNA, poly(A) RNA should be immediately subjected to reverse transcription and second-strand cDNA synthesis with no intermediate storage or precipitations.

If the amount of amplicon generated is suboptimal, several more cycles of PCR may be performed with the addition of new *Taq* DNA polymerase; however, PCR-introduced distortions of representations can be expected to be more pronounced at higher cycle numbers.

Subtractive/kinetic enrichment

In every round of subtractive hybridization, the amount of driver DNA remains constant. The amounts of tester (the product from the previous round) will diminish round by round to ultimately yield only the difference product or the differentially expressed targets. For DNA RDA, increasing stringency occurs with successive tester:driver ratios of 1:100, 1:800, 1:400,000, and 1:8,000,000. The tester:driver hybridization ratios may be modified, particularly when performing cDNA RDA to detect rare transcripts or smaller fold differences in expression between tester and driver. If no DNA products appear as bands by agarose gel electrophoresis in the later rounds of RDA, it may help to start either the particular hybridization round or the entire RDA again with a less stringent tester:driver hybridization ratio (relatively more tester DNA). If too much background smearing occurs in later rounds of RDA and primer problems have been ruled out (see below), then a more stringent tester:driver hybridization ratio in the preceding round may help.

No difference products

RDA requires the generation of restriction fragments between 200 to 1000 bp to ensure optimal PCR amplification. Because of this simplification step, a particular restriction strat-

egy may fail to find sought after differences. Therefore, if no difference products are isolated after iterative rounds of kinetic/subtractive enrichment, alternate restriction endonucleases may be tried. The placement in the tester sample of an internal control with known restriction characteristics at the beginning of an RDA experiment can be used; however, to prevent preferential amplification of the internal control, the internal standard should be spiked at sufficiently low concentrations (<1:100,000 on a weight basis).

Too many difference products or high background

HPLC purification of oligomers is critical for minimizing false positive RDA bands (O'Neill and Sinclair, 1997). Additionally, repeated thawing and freezing of primers in aqueous solution results in increased primer hydrolysis and contributes to mispriming during PCR amplification. This may result in increased false positives, increased rounds required to isolate true difference products, and excessive background smearing. Primers can be stored in lyophilized aliquots to circumvent this problem.

Anticipated Results

The RDA protocol selectively enriches for unique DNA sequences in the tester DNA sample. Upon completion of RDA, enriched populations of DNA can be visualized on agarose gel electrophoresis as a few or several distinct bands that range usually from between 150 to 800 bp. Although a significant background smear with only poorly identifiable bands may be seen at the end of the first round of kinetic/subtractive enrichment, successive rounds of enrichment should result in sharper bands with clean backgrounds. Even if discrete bands appear in the first or second rounds of RDA, three or more rounds are typically required to minimize background amplimers or stochastically amplified false positives. Authentic bands can then be cloned by a variety of different approaches. Sequence analysis of clones should reveal authentic endonuclease restriction sites at the termini of the inserts. It is not unusual to identify more than one discrete DNA fragment from each band; however, the majority of the clones should contain a single true positive difference product.

Time Considerations

Preparation of amplicons and representations requires 4 days. Each round of DNA

kinetic/subtractive enrichment also requires 4 days. If three rounds are performed, an RDA experiment exclusive of DNA and cDNA preparation or subsequent cloning can be completed in 16 days. Four rounds require 20 days. With some consideration for life's distractions, an RDA experiment can be performed in 4 weeks.

Acknowledgments

The contributor would like to thank Nikolai A. Lisitsyn, Michael Wigler, and Craig V. Byus for providing detailed laboratory protocols for RDA, Roy Bohenzky and Patrick S. Moore for helpful discussion, and Patrick S. Moore for review of the protocol.

Literature Cited

- Bakin, A.V., and Curran, T. 1999. Role of DNA 5-methylcytosine transferase in cell transformation by fos. *Science* 283:387-390.
- Birkenmeyer, L.G., Desai, S.M., Muerhoff, A.S., Leary, T.P., Simons, J.N., Montes, C.C., and Mushahwar, I.K. 1998. Isolation of a GB virus-related genome from a chimpanzee. *J. Med. Virol.* 56:44-51.
- Chang, Y., Cesarman, E., Pessin, M.S., Lee, F., Culpepper, J., Knowles, D.M., and Moore, P.S. 1994. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* 266:1865-1869.
- Geng, M., Wallrapp, C., Muller-Pillasch, F., Frohme, M., Hoheisel, J.D., and Gress, T.M. 1998. Isolation of differentially expressed genes by combining representational difference analysis (RDA) and cDNA library arrays. *Biotechniques* 25:434-438.
- Hubank, M. and Schatz, D.G. 1994. Identifying differences in mRNA expression by representational difference analysis of cDNA. *Nucleic Acids Res.* 22:5640-5648.
- Lisitsyn, N. and Wigler, M. 1995. Representational difference analysis in detection of genetic lesions in cancer. *Methods Enzymol.* 254:291-304.
- Lisitsyn, N., Lisitsyn, N., and Wigler, M. 1993. Cloning the differences between two complex genomes. *Science* 259:946-951.
- Lowrey, P.L., Shimomura, K., Antoch, M.P., Yamazaki, S., Zemenides, P.D., Ralph, M.R., Menaker, M., and Takahashi, J.S. 2000. Positional syntenic cloning and functional characterization of the mammalian circadian mutation tau. *Science* 288:483-492.
- Nishizawa, T., Okamoto, H., Konishi, K., Yoshizawa, H., Miyakawa, Y., and Mayumi, M. 1997. A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. *Biochem. Biophys. Res. Commun.* 241:92-97.
- O'Neill, M.J. and Sinclair, A.H. 1997. Isolation of rare transcripts by representational difference analysis. *Nucleic Acids Res.* 25:2681-2682.
- Pastorian, K., Hawel, L. 3rd, and Byus, C.V. 2000. Optimization of cDNA representational difference analysis for the identification of differentially expressed mRNAs. *Anal. Biochem.* 283:89-98.
- Reick, M., Garcia, J.A., Dudley, C., and McKnight, S.L. 2001. NPAS2: An analog of clock operative in the mammalian forebrain. *Science* 293:506-509.
- Shields, J.M., Der, C.J., and Powers, S. 2001. Identification of Ras-regulated genes by representational difference analysis. *Methods Enzymol.* 332:221-232.

Contributed by Yuan Chang
Hillman Cancer Center
University of Pittsburgh
Pittsburgh, Pennsylvania

Gene Expression Analysis of a Single or Few Cells

UNIT 25B.8

The need to analyze rare or even single cells is based on the dynamic nature of tissue differentiation and regeneration, the initiation and propagation of disease processes in multicellular organisms, and the functional diversity of individual cells. Gene transcription is the most important regulatory mechanism by which a phenotype and functional state of a cell is determined. Therefore, qualitative and quantitative assessment of mRNA abundance is not only a first step into the nature of biological processes but is easier to investigate in a comprehensive way than protein expression when small cell numbers are used.

In this unit, a protocol that allows a semi-quantitative analysis of gene expression of a single cell and a quantitative representation of expressed genes from >10 to 30 cells is described. This unit concentrates on the amplification procedure (see Basic Protocol 1) and less on the cDNA array hybridization. However, a basic protocol (see Basic Protocol 2) for array hybridization on nylon filters is provided because such filters are available in every laboratory without the need of additional expensive equipment. As tissue samples contain many different cell types in variable amounts, their analysis often requires microdissection of the tissue to isolate the specific cell types. Therefore, additional information on how to isolate mRNA from very small tissue samples such as biopsies and laser-microdissected material from cryosections (see Alternate Protocols 1 and 2) is given. Finally, a simple procedure to prepare the data for statistical analysis is also provided (see Basic Protocol 3).

STRATEGIC PLANNING

This unit deals with the handling of minute amounts of mRNA. Therefore, two “natural foes,” contamination and RNA degrading enzymes (RNases; see *UNIT 4.1* for additional details), will be encountered. Contamination can be reduced by working under a laminar-flow clean bench that has never been exposed to PCR-amplified DNA or cloned DNA, and that is preferably located in a room apart from laboratories where DNA is handled. It is recommended to always use filter tips for solutions and to take care not to contaminate pipets or other devices with DNA from other rooms. Unfortunately, contamination might still occur since many enzymes (in particular, reverse transcriptase) contain traces of bacterial DNA/RNA that will be co-amplified with the desired single-cell mRNA. For many assays, this bacterial DNA will not interfere, but may be a potential source of trouble. Degradation of RNA by RNases can be avoided by the use of powder-free gloves (changing them frequently) and being cautious when preparing buffers. RNase inhibitors are not added because they are frequently derived from human placenta and might therefore be contaminated with human nucleic acids. Working quickly and placing probes on ice is also recommended.

GLOBAL AMPLIFICATION OF SINGLE-CELL cDNA

This PCR-based protocol has been developed for maximal sensitivity of transcript detection. This raises the concern of exponential-error transmission, which will be discussed in detail along with the means that have been undertaken to reduce this error. However, one has to be aware that by using this method an exact quantification of the transcripts from a single cell is not possible; rather, semi-quantitative results are obtained.

**BASIC
PROTOCOL 1**

**Discovery of
Differentially
Expressed Genes**

25B.8.1

paramagnetic oligo-dT beads. While the mRNA is bound to the beads, reaction buffers can easily be changed without loss of mRNA or cDNA. This allows using optimal (i.e., high) concentrations of cDNA primers and nucleotides during cDNA synthesis without interference with the subsequent tailing reaction. To avoid loss of transcripts, do not contaminate the reaction with RNases because the mRNA holds the newly synthesized cDNA to the bead. After cDNA synthesis and before starting the tailing reaction, the unbound cDNA synthesis primers and unincorporated dNTPs have to be washed out. Tailing is performed in a KH_2PO_4 buffer that, unlike the provided potassium-cacodylate buffers, does not inhibit the subsequent PCR reaction, which is set up in the same reaction tube without discarding the tailing buffer.

Random primers were originally used because they reduced the length of an amplicon and allowed amplification of 5'-sequences. These random primers, combined with oligo-dT primers, slightly improve the results when single cells are used (CFL5 primer mix). However, when higher cell numbers (>100) are used, it appears that random primers alone work at least as well as the combination. For single cells, a random octamer increases the average fragment length, compared to a random hexamer, by ~100 to 200 bp. Due to the increasing number of commercially available oligo arrays that are restricted to the 3'-end, it might be advantageous to use oligo dT primers alone. The authors' first experiments indicate that the CFL5CT₍₂₄₎ primer should be used in this instance.

Materials

Oligo dT kit (Dynal) including:

Dynabeads Oligo (dT)₂₅

Washing buffer containing LiDS

Lysis buffer

Phosphate-buffered saline (PBS; APPENDIX 2)

5× RT buffer (Life Technologies)

0.1 M DTT (Life Technologies)

10% (v/v) Igepal

cDNA synthesis primers:

For mRNA amplification for ≥100 cells:

CFL5C6: 5'-(CCC)₅ GTC TAG ANN NNN N-3' (200 μM)

For single cells and 5' and 3' coverage:

CFI5C8: 5'-(CCC)₅ GTC TAG ANN NNN NNN-3' (200 μM)

CFI5CT: 5'-(CCC)₅ GTC TAG ATT TTT TTT TTT TTT TVN-3' (100 μM)

CFL5 primer mix: 1 vol CFI5c8 (200 μM) + 1 vol CFI5cT (100 μM)

For the use of 3'-restricted oligo arrays:

CFI5CT₍₂₄₎: 5'-(CCC)₅ GTC TAG ATT (T)₂₂ VN-3'

10 mM and 200 μM dNTPs

Reverse transcriptase (Superscript II; Life Technologies)

Igepal wash buffer (see recipe)

Tween 20 wash buffer (see recipe)

40 mM MgCl_2

2 mM dGTP

200 mM KH_2PO_4

Tailing wash buffer (see recipe)

Mineral oil

Terminal deoxynucleotide transferase (TdT; Amersham Pharmacia Biotech)

Expand Long Template (ELT) PCR system (Roche Diagnostics) including:

10× ELT buffer 1 (17.5 mM MgCl_2)

3.5 U/μl DNA polymerase mix

20% (v/v) formamide
 PCR primer, CP2: 5'- TCA-GAA-TTC-ATG-CCC-CCC-CCC-CCC-3' (24 μ M)
 1 \times PCR buffer (Sigma)
 Primers for β -actin: 5'- CTA CGT CGC CCT GGA CTT CGA GC-3' and 5'-GAT GGA GCC GCC GAT CCA CAC GG-3'
 Primers for EF-1 α : 5'- GCA GTG CAC ACA CAG AGG TGT A-3' and 5'- CTA CCG CTA GGA GGC TGA GCA A-3'
 0.75 U *Taq* DNA polymerase (Sigma)
 Magnet separation apparatus for 0.2-ml tubes (Dynal)
 0.2-ml PCR tubes
 15- to 50-ml tubes
 Roller-bottle apparatus or other rotisserie-type rotator
 Thermal cycler
 Hybridization oven or other rotator with temperature control
 Additional reagents and equipment for agarose gel electrophoresis (UNIT 2.5A)

Lyse cells and isolate mRNA

1. Wash beads two times in an equal volume of washing buffer containing LiDS using the magnet.

Dynal beads are supplied as a solution and have to be washed using the magnet prior to use. Resuspend beads in adequate volume of lysis buffer to which the cells or tissue biopsies are added.

The beads must completely adhere to the side of the tube at the site of the magnet before the supernatant is removed to avoid loss of beads. This wash procedure can take several minutes. Do not forget to prepare beads for the negative control.

2. Resuspend beads in an equal volume of lysis buffer. The amounts of lysis buffer and beads depend on the cell number. Table 25B.8.1 suggests the volumes of lysis buffer and beads to use for specific numbers of cells.

3. Pick cells in 1 \times PBS (APPENDIX 2) in the smallest possible volume. Pick single cells in a 1- to 2- μ l volume and add to the beads in lysis buffer in a 0.2-ml PCR tube.

Individual cells can be isolated from suspensions using a 2- μ l automatic pipettor and an inverted microscope.

Cell numbers >3000 in one reaction tube should be avoided because the released genomic DNA will clump the beads and prevent successful isolation of mRNA. When more cells are used, either use up to 500 μ l of lysis buffer with 50 μ l of beads, use aliquots, or isolate total RNA first by classical protocols (e.g., UNIT 4.1) and add the RNA (1 to 10 μ g total RNA) to the beads.

4. Place the 0.2-ml PCR tubes in a 15- to 50-ml tube and rotate the lysate for 30 min at 4° to 20°C (room temperature) in a roller-bottle apparatus.

Rotation ensures that the beads remain suspended.

If desired, freeze the sample after this step at -80°C. The authors have stored samples for up to 12 months without any negative effect. On continuation, resuspend the beads after thawing and rotate for 5 min.

Table 25B.8.1 Volumes of Beads and Lysis Buffer for Given Numbers of Cells

No. of cells	Oligo dT beads	Lysis buffer
1–10	10 μ l	10 μ l
11–50	30 μ l	30 μ l
51–300	50 μ l	50 μ l
>300–3000	50 μ l	50–200 μ l

Table 25B.8.2 cDNA Synthesis Mixes

No. of samples	1	2	3	4	5	6	7	8	9	10
<i>cDNA synthesis mix I^a</i>										
5× first strand buffer	2	4	6	8	10	12	14	16	18	20
0.1 M DTT	1	2	3	4	5	6	7	8	9	10
10% Igepal	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
H ₂ O	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
cDNA synthesis primers	6	12	18	24	30	36	42	48	54	60
<i>cDNA synthesis mix II^a</i>										
5× first strand buffer	2	4	6	8	10	12	14	16	18	20
0.1 M DTT	1	2	3	4	5	6	7	8	9	10
10 mM dNTP	1	2	3	4	5	6	7	8	9	10
H ₂ O	5	10	15	20	25	30	35	40	45	50
Reverse transcriptase	1	2	3	4	5	6	7	8	9	10

^aAll solution volumes are indicated in microliters.

Synthesize cDNA

5. Prepare cDNA synthesis mix I and II (see Table 25B.8.2) on ice while the beads are rotating. Add the reverse transcriptase to mix II just before use.

Never use a reverse transcriptase with RNase H activity.

6. Add an equal volume of Igepal wash buffer to the cell lysate containing the mRNA bound to the beads and place tube in the magnet. Remove supernatant after the beads have completely adhered to the tube at the site of the magnet. Resuspend beads carefully in 20 μ l Tween 20 wash buffer. Transfer to a fresh 0.2-ml tube, place in the magnet, and remove the supernatant after complete adhesion of beads to the magnet.

The multiple washing steps as well as the change of the reaction tube serve to remove the LiDS-containing buffer, since even small traces of LiDS can inhibit reverse transcription. It is very important to allow complete adhesion of the magnetic beads to the tube wall at the site of the magnet to avoid loss of cDNA. Note that collection of the supernatant and storage at -20°C may be desired because it contains the genomic DNA that can be used for additional analyses at a later time.

7. Resuspend beads in cDNA synthesis mix I and allow primers to anneal for 2 min on the bench at room temperature, then add mix II (remember to add the RT in mix II). Immediately start cDNA synthesis by placing the tubes in a hybridization oven for 45 to 60 min at 44°C with rotation.

It is important to rotate so that the beads remain suspended.

The authors tape the 0.2-ml sample tubes to pre-heated hybridization bottles.

8. Prepare tailing mix (see Table 25B.8.3).
9. Place tubes in the magnet and remove supernatant. Wash beads one time in 20 μ l tailing wash buffer. Pre-heat thermal cycler to 94°C .

After cDNA synthesis and before starting the tailing reaction, the unbound cDNA synthesis primers and unincorporated dNTPs have to be washed off. Therefore, meticulously remove all of the cDNA synthesis solutions by carefully pipetting, because dNTPs and primers will interfere with the tailing reaction.

Table 25B.8.3 Tailing Mix^a

No. of samples	1	2	3	4	5	6	7	8	9	10
40 mM MgCl ₂	1	2	3	4	5	6	7	8	9	10
1 mM DTT	1	2	3	4	5	6	7	8	9	10
2 mM dGTP	1	2	3	4	5	6	7	8	9	10
200 mM KH ₂ PO ₄	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
H ₂ O	6.5	13	19.5	26	32.5	39	45.5	52	58.5	65

^aAll solution volumes are indicated in microliters.

Table 25B.8.4 PCR Mixes for Global Amplification^a

No. of samples	1	2	3	4	5	6	7	8	9	10
<i>PCR-mix I</i>										
Roche buffer 1	4	8	12	16	20	24	28	32	36	40
20% formamide	7.5	15	22.5	30	37.5	45	52.5	60	67.5	75
H ₂ O	24	48	72	96	120	144	168	192	216	240
<i>PCR-mix II</i>										
24 μM CP2 primer	2.5	5	7.5	10	12.5	15	17.5	20	22.5	25
10 mM dNTP	1.75	3.5	5.25	7	8.75	10.5	12.25	14	15.75	17.5
Taq long template	1.5	3	4.5	6	7.5	9	10.5	12	13.5	15

^aAll solution volumes are indicated in microliters.

Tail cDNA

- Resuspend beads in tailing mix and add 40 μl mineral oil on the surface. Place the 0.2-ml tubes in the preheated thermal cycler and denature RNA-DNA hybrids for 5 min at 94°C. Immediately chill on ice.

This step serves to generate single-stranded cDNA, which is tailed with high efficiency in contrast to RNA-DNA hybrids. (After denaturation, the cDNA is no longer bound to the magnetic beads but is now found in the supernatant.) The following tailing and PCR procedure will take place with the beads in the tube.

- Add 10 to 15 U TdT, mix thoroughly, and start tailing in a thermal cycler programmed for 60 min at 37°C, then 22°C indefinitely.

Tailing is complete after 1 hr, but can be extended overnight at 22°C, whenever necessary.

- Inactivate TdT by incubating cDNA at 70°C for 5 min.

Amplify by PCR

- Prepare PCR mix I and II on ice (see Table 25B.8.4).
- After inactivation of TdT, add PCR mix I to the aqueous phase under the mineral oil. Incubate for 30 sec at 78°C.
- Add 5.5 μl mix II, then carry out the amplifications in a thermal cycler with the following parameters:

1 cycle:	30 sec	78°C
19 cycles:	15 sec	94°C
	30 sec	65°C
	2 min	68°C

continued

20 cycles:	15 sec	94°C
	30 sec	65°C
	2.5 min + 10 sec/cycle	68°C
1 cycle:	7 min	68°C
	indefinitely	4°C.

The separation of mix I and II serves for the hot-start procedure. Add the largest solution volume first, which consists of buffer and water. After 78°C has been reached, add the primers, nucleotides, and enzymes. Taq long template is one of several available mixtures of a highly processive DNA polymerase (Taq-polymerase) and a proof-reading enzyme with 3'-5' exonuclease activity (Pwo-polymerase). The exonuclease activity would degrade the single-stranded CP2 primer in absence of dNTP, which consequently has to be included in mix II. The reason for the hot start is to avoid unspecific priming and extension of the CP2 primers (that bound to the single-stranded cDNA at low temperatures) until 94°C is reached. The longer extension time in cycles 20 to 39 is due to the increased amount of product.

16. Store sample at -20°C.

Evaluate global amplification and validate genes

17. Check 3 to 5 µl of the primary PCR on a 1.5 % agarose gel for the presence of a smear in the range of 300 to 2000 bp.

18. Test amplification success by performing gene-specific PCR on at least two house-keeping genes.

For human cells, use the primers for β-actin and EF-1α (see Materials) in the conditions outlined below (see step 19). For other species, choose/design primers specific to house-keeping genes of those species.

To test amplification success, perform gene-specific PCRs for selected genes. Each gene-specific PCR should be individually optimized. For most transcripts, best results will be obtained after dilution of the primary amplifications in water (1:10). As the length of the amplified cDNA is usually <1000 bp, choosing primers that amplify sequences of 150 to 200 bp is recommended, as this size range produces the best results.

19. Make up a PCR reaction containing 2.5 ng of each cDNA in a 25-µl reaction containing 1× PCR buffer (Sigma), 200 µM dNTPs, 0.4 µM of each primer (β-actin or EF-1α), and 0.75 U Taq polymerase. Carry out the amplifications in a thermal cycler with the following parameters:

1 cycle:	2 min	94°C
	30 sec	58°C
	2 min	72°C
14 cycles:	40 sec	94°C
	30 sec	58°C
	20 sec	72°C
15-45 cycles:	40 sec	94°C
	30 sec	58°C
	30 sec	72°C
1 cycle:	40 sec	94°C
	30 sec	58°C
	2 min	72°C
	indefinitely	4°C.

The number of cycles in the main part of the amplification can be 15 to 45, depending on the transcript abundance.

20. Run PCR products on a 2% agarose gel containing 0.5 µg/ml ethidium bromide.

EXTRACTION OF mRNA FROM SMALL TISSUE BIOPSIES

This protocol is used to isolate and amplify mRNA from small biopsies that are obtained during diagnostic clinical procedures and do not undergo laser microdissection (*UNIT 25A.1*). The fresh biopsy is immediately snap-frozen in liquid nitrogen and stored in liquid nitrogen or at -80°C until lysis and mRNA preparation is performed.

Additional Materials (also see *Basic Protocol 1*)

Biopsy sample
Liquid nitrogen
Dry ice
Mortar and pestle

1. Use only a small piece of the biopsy sample with a size of 1 to 1.5 mm in diameter.
2. Using a mortar and pestle, crush the frozen tissue sample in liquid nitrogen.

Prior to using the mortar and pestle, destroy all nucleic acids by UV irradiation. Expose the internal surface of the mortar to UV light in a transilluminator or hold it close to a UV light source (254-nm wavelength) for 10 to 15 min. For the pestle, in order to expose the whole surface, it will be necessary to turn it, as only DNA lying in the direct path of the light will be destroyed by the UV irradiation (also see APPENDIX 3F for sterile technique).

Thawing of the sample must be avoided under all circumstances! Therefore, place the mortar on dry ice and frequently pour liquid nitrogen over the sample.

3. Add the powdered sample directly to 50 μl of prepared Dynal beads (see *Basic Protocol 1*, steps 1 and 2) and rotate lysate as in *Basic Protocol 1*, step 4.
4. Proceed with global amplification in *Basic Protocol 1*, steps 5 through 20.

EXTRACTION OF mRNA FROM MICRODISSECTED SAMPLES

Laser microdissection is the cleanest way to isolate selected morphologically defined cell groups from tissue sections. However, it is also possible to scratch the tissue area with a glass needle of which the tip is then broken into the lysis buffer. The authors use the PALM Laser-MicroBeam System (PALM) that first cuts the selected area by a laser beam and then catapults it into the lid of the reaction tube (see Fig. 25B.8.2). Other laser microdissection systems (see *UNIT 25A.1*) should work equally as well, as long as the isolation does not change the composition of the lysis buffer. The combination with *Basic Protocol 1* (Global Amplification) and *Basic Protocol 2* (Non-Radioactive Gene Expression Analysis on Nylon Arrays) enables quick analysis of global gene expression from 30 to 200 cells from 5- μm sections.

Materials (also see *Basic Protocol 1*)

Resectioned tissue snap-frozen in liquid nitrogen and stored at -80°C (see *Alternate Protocol 1*)
OCT embedding compound (Tissue-Tek, Miles; also see *UNIT 25A.1*)
Mayer's hematoxylin solution (Sigma)
70%, 95%, and 100% ethanol
Lysis buffer from Oligo dT kit (see *Basic Protocol 1*)
Cryostat
Slides for the PALM Laser-MicroBeam System (PALM)
PALM Laser-MicroBeam System (PALM)

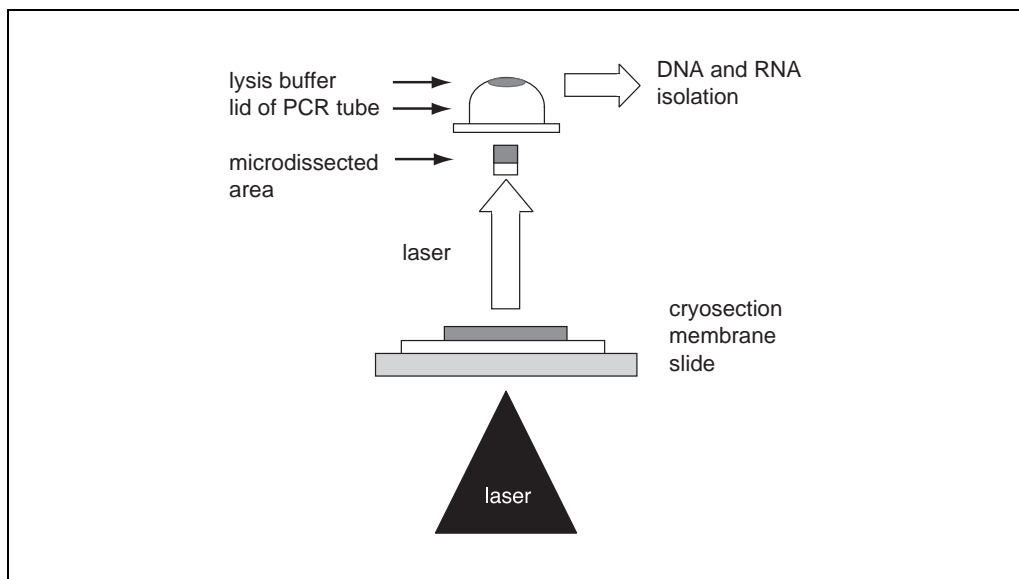


Figure 25B.8.2 Isolation of small tissue samples by laser microdissection and catapulting using the PALM system.

1. Embed the tumor sample in OCT embedding medium (*UNIT 25A.1*) and cut the sample to 5- μ m thick slices on slides for the PALM Laser-MicroBeam System using a cryostat.
2. Place the slides in Mayer's hematoxylin solution for 45 sec, in water for 5 min, and in distilled water for 1 min.
3. Dehydrate sections in 70%, 95%, and 100% ethanol for 60 sec in each concentration.
4. Dry stained tissue sections overnight at room temperature.

The slides are ready for the Laser-MicroBeam System.

For the PALM Laser-MicroBeam System, the sections have to be completely dried, otherwise the heat generated by the laser beam will be transmitted, boil the tissue, and destroy the mRNA. If using a different microdissection system, individually establish the conditions and parameters.

5. To catch the catapulted tissue area in the lid of a PCR reaction tube, pipet 5 μ l lysis buffer on the inner wall of the lid.
6. Centrifuge the lysed tissue (mRNA and DNA) at maximum speed and proceed with mRNA isolation and global amplification (see Basic Protocol 1).

NON-RADIOACTIVE GENE EXPRESSION ANALYSIS ON NYLON ARRAYS

This protocol allows one to assay the expression of many genes whose mRNAs are represented in the amplification in Basic Protocol 1 without expensive equipment. It also assesses the complexity of sequences within the amplification, which can be helpful before proceeding to more detailed analyses. Test filters may be self-prepared by spotting 5 to 50 ng of each cDNA sequence (each should have a length of 300 to 700 bp) in 1 to 2 μ l of 0.1 M NaOH on a positively charged nylon membrane. There are also several commercially available products. See Chapter 22 for methods to prepare and assay arrays on glass slides.

**BASIC
PROTOCOL 2**

**Discovery of
Differentially
Expressed Genes**

25B.8.9

Materials

Expand Long Template (ELT) PCR system (Roche Diagnostics) including:
10× ELT buffer 1 (17.5 mM MgCl₂)
3.5 U/μl DNA polymerase mix
1/7 dNTP mix (see recipe)
20% formamide
CP2 primer: 5'-TCA-GAA-TTC-ATG-CCC-CCC-CCC-CCC-3' (24 μM)
Digoxigenin-11-dUTP (Dig-UTP), alkali labile (Roche Diagnostics)
Sample
DIG Easy Hyb solution (Roche Diagnostics)
E. coli DNA
DNase I
Labeled probe
Herring sperm DNA (Invitrogen)
20× SSC
10% SDS
Development buffer 1 (see recipe)
Development buffer 2 (see recipe)
DIG Luminescent Detection Kit (Roche Diagnostics) containing:
Blocking reagent
750 U/ml anti-digoxigenin-AP (Fab fragment) antibody
11.6 mg/ml CSPD
Tween 20 (Sigma)
Development buffer 3 (see recipe)
Thermal cycler
Nylon membrane containing an array of cDNAs (either self-prepared or commercially available)
Hybridization tubes
Hybridization oven or other rotator with temperature control
1.5-ml microcentrifuge tubes
Acetate sheets
Whatman 3MM filter paper
Biomax ML film (Kodak)

Label amplifications with Dig-UTP

1. Prepare PCR master mix as in Table 25B.8.5. Pipet 49-μl aliquots in sterile PCR tubes, add 1 μl from the sample (i.e., from the PCR product obtained in Basic Protocol 1, step 15) and program the thermal cycler with the following parameters:

1 cycle:	2 min	94°C
	4 min	68°C
10 cycles:	15 sec	94°C
	4 min	68°C
2 cycles:	15 sec	94°C
	4 min + 10 sec/cycle	68°C
1 cycle:	7 min	68°C
2. Determine the concentration of the amplified DNA (see UNIT 2.6, Support Protocol).
3. Prehybridize nylon array by placing the nylon membrane containing the cDNA array in a small hybridization tube, add 6 ml DIG Easy Hyb solution supplemented with 100 μg/ml *E. coli* DNA that has been digested with DNase I to a size of 100 to 1000 bp, and prehybridize for at least 6 hr at 45°C.

Table 25B.8.5 PCR Master Mix for Non-Radioactive Gene Expression Analysis^a

No. of samples	1	2	3	4	5	6	7	8	9	10
10× ELT buffer 1	5	10	15	20	25	30	35	40	45	50
dNTP mix	1.75	3.5	5.25	7	8.75	10.5	12.25	14	15.75	17.5
20% formamide	7.5	15	22.5	30	37.5	45	52.5	60	67.5	75
24 μM CP2 primer	5	10	15	20	25	30	35	40	45	50
Dig-UTP	2.5	5	7.5	10	12.5	15	17.5	20	22.5	25
H ₂ O	28	56	84	112	140	168	196	224	252	280
3.5 U/μl DNA polymerase mix	0.5	1	1.5	2	2.5	3	3.5	4	4.5	50

^aAll solution volumes are indicated in microliters.

Be aware that several commercial membranes are heavily contaminated with bacterial and/or plasmid DNA. The additional DNA in the hybridization solution serves not only to block all non-specific binding of labeled probe but also any amplified bacterial/plasmid DNA contaminating the enzyme preparations used to generate probes. All enzyme preparations contain traces of bacterial RNA/DNA that will be amplified by the highly sensitive amplification protocol and sometimes hybridize to bacterial/plasmid DNA on the filters. Therefore, even with the additional DNA, some poor arrays might not be usable. Always test the quality of the arrays by labeling and hybridizing a probe that has been amplified by the protocol in the absence of cellular RNA (negative control), in which contaminating DNA from the enzymes can be expected to be present as in the cell samples.

Note that although some favor the opposite nomenclature for array hybridizations, the authors use the term “probe” to refer to the labeled DNA in solution.

Add probe to membrane

- Mix in a 1.5-ml microcentrifuge tube, 1 ml DIG Easy Hyb solution, 6 μg of the labeled probe from step 1, and 100 μg of herring sperm DNA. Denature 5 min at 94°C and immediately add to the prehybridization solution in the hybridization tube. Incubate with slow rotation at least 36 hr at 45°C.

It is important that the nylon membrane be completely covered with the hybridization solution before rotating. Otherwise, high non-specific backgrounds will result due to the drying of the membrane during hybridization. Therefore, adjust the amount of hybridization solution to add to the prehybridization accordingly.

Additionally, do not pour the concentrated probe directly onto the filter. This will result in high background.

Wash the membrane

- Remove the hybridization solution and wash the membrane in the bottle and in the hybridization oven at 68°C using the following regimen:

- 1 min in 2× SSC + 0.1% SDS
- 1 min in 1× SSC + 0.1% SDS
- 15 min in 0.5× SSC + 0.1% SDS
- 30 min in 0.1× SSC + 0.1% SDS (two times)

Warm all solutions to 68°C prior to use in a water bath.

The hybridization mix can be stored at –20°C and re-used for additional filters. Before re-using the hybridization mix, denature the solution for 10 min at 80°C.

- Wash the membrane in development buffer 1 for a few seconds at room temperature, then block in 25 ml development buffer 2 for 30 min with gentle agitation.

7. Dilute 2.5 μ l anti-digoxigenin-AP (Fab fragment) antibody directly into the 25 ml development buffer 2 and incubate for an additional 30 min at room temperature.
8. Pour off development buffer 2 and wash two times, 15 min each, in development buffer 1 containing 0.3% Tween 20 at room temperature.

This step will remove the unbound antibody.

Detect binding with chemiluminescent substrate

9. Prepare 50 ml of development buffer 3 and prepare 1 ml of chemiluminescent substrate by mixing 10 μ l CSPD with 990 μ l development buffer 3. Equilibrate the membrane for a few seconds in the remaining development buffer 3.
10. Place the membrane between two acetate sheets. Lift the top sheet of plastic and add 1 ml of the chemiluminescent substrate (from step 9), scattering the drops over the surface of the membrane. Carefully lower the top sheet of plastic without producing any bubbles.
11. Incubate on the bench 5 min at room temperature. Remove the membrane from the plastic sheets and place on a sheet of Whatman 3MM paper for a few seconds to remove excessive chemiluminescent substrate, then put the membrane back between two clean, dry acetate sheets.

It is important to remove any excess moisture from the membrane. This avoids the development of background during film exposures up to 60 min. However, the membrane should not completely dry out because this would exclude any further use.

12. Incubate 15 min at 37°C and place the membrane on film to be exposed.

The 37°C-incubation allows the alkaline phosphatase reaction to reach a steady state quickly. The authors recommend 15 min for the first exposure, then adjust the time according to the signal strength.

BASIC PROTOCOL 3

DATA ANALYSIS OF HYBRIDIZED cDNA ARRAYS

There are several ways to analyze and normalize the data obtained by gene-expression profiling with cDNA arrays. This protocol describes a method to measure differences of signal intensities of differentially expressed genes and to normalize the signal intensities to several housekeeping genes. See *UNIT 22.3* for other information regarding data analysis.

Materials

- Photographic step tablet (Kodak)
- Transparency scanner that can be calibrated (e.g., SNAPSCAN, Agfa)
- Labscan software or equivalent (Scanwise v. 1.2.1, Agfa)
- Array Vision software or equivalent (Clontech)
- Excel software or equivalent (Microsoft)
- SPSS software or equivalent (SPSS)

Perform intensity calibration of the scanner

1. Define the known density values from the photographic step tablet.

To analyze the signal intensity on the X-ray film, it is important to measure its optical density. Signal intensity is usually measured in units, which do not necessarily represent the same “real-world” values in different images. It is important to calibrate a scanner before measuring the optical density of the signals. Therefore, by indicating raw intensity values in an image and defining their corresponding optical density, the system can be provided with the information it needs to convert its measurements to real-world quantities.

2. Scan the photographic step tablet in the grayscale mode values that have been entered.

The Kodak No. 2 Photographic Step Tablet Standard values are provided as optical density values, starting at 0.05 and proceeding at 0.15-OD increments to 3.05 OD units. At least four calibration points are necessary to compute a calibration curve.

3. Compute a calibration curve.

At least three different curve models are available. The linear option calculates the curve with the formula: $y = ax + b$; the quadratic option with the formula: $y = ax^2 + b$; and the log linear option with the formula: $y = a \log ((255 - x)/255) + b$. The authors recommend the log linear option.

Scan the developed films

4. Scan the grayscale of the developed films in the transmission mode with a resolution of at least 600 dpi.
5. Save files as MD GEL (*.gel), MCID (*.im), BRS (*.img), TIFF (*.tif), TIFF5 (*.tif), Fujix Bas Series (*.inf), Bio-Rad PA (*.img), Packard (*.tif), or MD Dataset (*.ds).

Other data formats cannot be imported by the array vision software.

Analyze with software

6. Define a template according to the grid of the cDNA arrays that were used.
7. Import the scanned films as a data file into arrays vision.
8. Align the grid to the corresponding spots on the cDNA array.
9. Normalize the signals to the housekeeping genes present on the cDNA array.

The average of the signals of the housekeeping genes is set to a value of one and the background to a value of zero.

10. Sample the template.
11. Export the gained data to MS Excel and/or SPSS for further statistical analysis.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Development buffer 1

100 mM maleic acid
150 mM NaCl, pH 7.5
Autoclave and store up to 6 months at room temperature

Development buffer 2

100 mM maleic acid
150 mM NaCl, pH 7.5
1% blocking reagent (DIG Luminescent Detection Kit, Roche Diagnostics)
Store up to 12 months at -20°C

Development buffer 3

100 mM Tris-Cl, pH 9.5 (APPENDIX 2)
100 mM NaCl
Prepare fresh just prior to use.

dNTP mix, 1/7

10 mM dCTP
10 mM dGTP
10 mM dATP
8.4 mM dTTP
Store up to 12 months at -20°C

Igepal wash buffer

50 mM Tris·Cl, pH 8 (*APPENDIX 2*)
75 mM KCl
10 mM DTT
0.25% (v/v) Igepal CA-630 (Sigma)
Store up to 12 months at -20°C

Tailing wash buffer

50 mM potassium phosphate, pH 7 (*APPENDIX 2*)
1 mM DTT
0.25% (v/v) Igepal CA-630 (Sigma)
Store up to 12 months at -20°C

Tween 20 wash buffer

50 mM Tris·Cl, pH 8 (*APPENDIX 2*)
75 mM KCl
10 mM DTT
0.5% (v/v) Tween 20
Store up to 12 months at -20°C

COMMENTARY

Background Information

Overview of amplification methods for small amounts of mRNA

With the completion of the human genome project and the introduction of technologies such as DNA microarrays and laser microdissection, many fields in biology and medicine await the application of comprehensive gene expression analyses of specific cell types isolated from defined tissues. The first protocols for the amplification of single cell mRNA were introduced in the late 1980s and early 1990s (Belyavsky et al., 1989; Brady and Iscove, 1993) and their development as well as their technical differences and application have been recently reviewed (Brady, 2000). All protocols are based on either of two principal approaches—linear amplification by T7 RNA polymerase or PCR amplification. Both procedures have advantages and disadvantages, and the one used depends on the experimental situation.

As a general rule, PCR-based methods are easier to handle and less time consuming, although there are concerns about the quantitative reliability of measurements obtained after

exponential amplification (Brail et al., 1999). The linear amplification achieved by T7 RNA polymerase, also referred to as the Eberwine protocol (Eberwine et al., 1992; Kacharina et al., 1999), has the advantage that a failure to amplify a given transcript will not be exponentially transmitted. On the other hand, there are several publications using PCR-based protocols showing that the relative abundance of transcripts is preserved even after 50 cycles. T7 RNA polymerase-based methods have been applied to cDNA and oligonucleotide arrays, but so far, the least number of cells that could be used successfully was ~1000 (Luo et al., 1999).

The methods provided in this unit are PCR approaches, and therefore are inherently prone to exponentially propagate initial amplification errors. The authors' primary intention was to obtain a qualitative representation of a single-cell transcriptome rather than preserving the numerical ratios of transcript abundance (Klein et al., 2002). Having established the method for single cells, the authors saw that quantitative differential analysis of gene expression with higher cell numbers (100 to 1000 cells) works quite well (Zohlnhofer et al., 2001a,b). This

seems to result from the fact that all experimental steps were optimized individually and in combination, that the number of steps was kept to a minimum, which led to high-complexity transcriptomes when the amplicons derived from single cells were hybridized onto cDNA arrays. Three points seemed to be particularly important. First, random primers reduce the length of primary transcript and enable subsequent amplification within the optimal range for PCR. Second, a poly-G tail provides a much better primer binding site than a poly A or poly T tail. Third, a poly-C containing PCR primer (binding to the poly G tail) should not be combined with any other primer sequence. Therefore, the flanking region of the cDNA synthesis primer has to be a poly-C track and a single PCR primer is used. A high annealing temperature and the addition of 3% formamide provide highly specific and optimal conditions for such sequences.

Reproducibility on a single-cell level is very difficult to assess, as two individual cells cannot be assumed to be identical and in the same functional stage. To exclude intercellular variation, the cDNA of an individual cell was divided prior to amplification, and then the variation of the resulting expression patterns (which was presumably introduced by the different methodological steps) was tested. Although random priming during cDNA synthesis, labeling, and hybridization add to the total variation, overall congruence of the two halves from one cell after global PCR was remarkably high for strong and intermediate signals. The weaker the signal, the more likely it was lost in one of the two halves (Klein et al., 2002). Therefore, when single cells are analyzed, the lack of a signal is more difficult to interpret and the authors recommend using independent methods such as real-time PCR or antibody staining.

Oligo arrays have become increasingly available from commercial suppliers (Affymetrix, Clontech, Qiagen, MWG-Biotech). Most of the sequences on these arrays are selected from the 3' end of a transcript. In those cases where the 5'-3' ratio is included into the bioinformatic evaluation (Affymetrix), one should not include the random primers, as the ratio will be shifted to the 5' end. Here, initial results indicate that the CFI5CT₍₂₄₎ alone results in more quantitative results (if using the Affymetrix system, do not forget to include the T7 promoter into the oligo in the order: 5'-poly-C-flank, T7 promoter, dT₍₂₄₎-3'). In addition, if enough cells are available that allow division of the sample, it is

advisable to determine the number of cycles needed to reach the plateau of the PCR reaction. Quantification is more precise during the linear phase of PCR, i.e., just before the plateau is reached. This can be done by setting up and running the PCR with half of the cDNA, and then running a gel of 3- μ l aliquots that are taken during the PCR at various cycle numbers between 20 and 40 cycles (e.g., cycle 20, 24, 28, etc.). Then, the PCR may be set up with the other half of the cDNA, programming the thermal cycler for the ideal number of cycles.

Critical Parameters

For best results, adhere to the following rules.

High-quality enzymes are critical for amplification success. In particular, terminal deoxynucleotide transferase (TdT) and RNaseH-deficient reverse transcriptase (RT) need to be selected carefully. TdT is delivered either in cacodylate-containing or KH₂PO₄-containing storage buffers. Avoid cacodylate-containing buffers unless they can be highly diluted. Reverse transcriptase is sometimes contaminated with bacterial DNA. Therefore, check different batches of a manufacturer.

Always work under sterile conditions with filter tips and avoid RNase contamination. It is also of great importance to protect the reactions from any nucleic acid contamination because DNA/RNA molecules present in the tube will be amplified as well (reverse transcriptase also uses DNA as a template).

Always work on ice.

During all wash steps using the magnet, check that no beads are aspirated with the supernatant.

Do not allow the beads to dry out. This preserves the binding of the mRNA to the beads.

Working with more than eight samples at once is not recommended, since it increases the duration of the procedure and consequently favors RNA degradation.

Clumped beads typically result from genomic DNA. Refer to Table 25B.8.1 to adjust the bead volume to the cell number.

Perform the hot-start procedure quickly, since keeping a single-stranded cDNA at 78°C for extended times can destroy the template.

Troubleshooting

Global amplification

There is no way to check the individual steps prior to PCR amplification. Before hybridizing a sample to an array, test amplification success

Table 25B.8.6 General Troubleshooting Guide

Cause/problem	Possible solution
<i>Negative PCRs or no primary product</i>	
Inactive reagents	Because all steps are critical, be sure that all reagents have been properly stored and handled. Primers should be dispensed into aliquots prior to use in order to prevent repeated freeze/thaw cycles; do not store diluted dNTP or dGTP too long; check expiration dates of enzymes; check tailing buffer and TdT storage buffer for absence of cacodylate; formamide must be deionized. High-quality enzymes and primers are most essential.
Gene-specific PCR	Most gene-specific PCRs will work with the primary PCR products as template, but be prepared to re-test the annealing temperature for the CP2-amplified cDNA. Sometimes, gene-specific PCR works better on 1:10 to 1:1000 diluted template than on undiluted amplicons.
<i>No or weak signals on cDNA arrays</i>	
Degraded Dig-UTP	Digoxigenin is alkali-labile. Therefore, check pH of all solutions after hybridization.
Film exposure	Be sure to expose the hybridized/exposed side of the filter. Re-expose cDNA array, prolong exposure time, correct orientation of coated film.
Hybridization temperature	Control the hybridization temperature. Some hybridization buffers work at 68°C, others at 45°C, depending on the content of DNA-denaturing substances.
Denaturation of DNA	Both probe and target have to be single stranded. Check denaturation and the protocol for array preparation.
<i>Suspiciously identical results with different probes on cDNA arrays</i>	
Co-amplification, labeling, and hybridization of bacterial/plasmid DNA with cellular cDNA	<p>Control the quality of the array by hybridizing labeled <i>E. coli</i> and plasmid DNA to the array; use arrays of which the cDNAs have been amplified by insert-specific PCR or oligonucleotide arrays</p> <p>Check for possible sources of contamination in the sample; test different batches of reverse transcriptase</p> <p>If contamination is unavoidable, label the negative control and add increasing amounts of blocking DNA (i.e., <i>E. coli</i> or DNA of the most frequently used plasmids used to generate the array) until the filters are clean</p>
<i>High background of cDNA arrays</i>	
Probe concentration	Check concentration of added probe. Concentrations >1.5 µg/ml can result in high background
Addition of probe	Never add undiluted probe to the array. Direct contact with the nylon membrane will result in dark areas/spots. Dilute the labeled probe in ~1 ml hybridization buffer and be careful not to pour it directly onto the filter.
Restringency washes	Unbound or unspecifically bound probe must be entirely washed out. Check SSC concentrations and washing temperatures.
Alkaline phosphatase	Alkaline phosphatase is expressed by bacteria. Check/autoclave buffers used for developing the filters.
Filters	Nylon membranes can be stripped and re-hybridized up to six to eight times. Repeated use, however, will increase background every time.
Precipitated Fab fragments, degraded anti-digoxigenin-alkaline phosphatase	Spin down antibody solution prior to use and use the supernatant only

by gene-specific PCRs for housekeeping genes and one less abundant, but more or less consistently expressed, gene of the investigated cells. In addition, it is advisable to run a gel with 5 μ l of the primary PCR. It should show a smear ranging from 200 to 2000 bp without bands. A sample without the addition of cellular mRNA should also be checked since contamination can be detected this way. If a smear (originating from DNA contamination in the enzyme preparations) is at all present, in the negative control, it should be smaller (100 to 300 bp) and less intense. Note that individual bands sometimes result from concatamerization of the primers and do not necessarily indicate contamination.

Gene expression analysis on nylon arrays

Roche Diagnostics provides an excellent manual with the digoxigenin hybridization kit. All relevant information for non-radioactive array analysis can be found there.

A general troubleshooting guide is presented in Table 25B.8.6.

Anticipated Results

After PCR amplification, the DNA content of the sample should be measured by following the Support Protocol in *UNIT 2.6* or by optical density, an ethidium bromide plate compared with a standard, or alternative methods like Nucleic dotMetric (Genotech). The anticipated amount of cDNA is between 100 and 300 ng/ μ l.

Before hybridization, amplification success is tested by checking the primary product and by gene-specific PCR as described. Running a gel with the primary PCR product, a smear ranging from 100 to 2000 bp without bands should be observed. Using single cells sometimes results in a smaller range. A sample without cellular mRNA should be included throughout the whole experiment as a negative control. From this sample, there should be no apparent smear; however, sometimes smears can be observed when the reagents, especially enzyme preparations, contain nucleic acids. Controlling the primary amplification by gene-specific PCRs for two housekeeping genes and one constantly but less abundantly expressed gene is recommended. The negative control must be negative for all gene-specific PCRs. Gene-expression analysis on nylon arrays should result in films with low background and ~20% to 40% positive hybridization signals for >10 to 20 cells. Positive signals from single cells should range from 5% to 25% of spotted cDNAs, depending on activation stage. The housekeep-

ing genes spotted on each filter should yield strong positive signals. The negative-control spots show no signal unless sample and array are contaminated with bacterial/plasmid-derived DNA.

Time Considerations

Global amplification of cellular cDNA

The time needed depends on the incubation/reaction times and the number of samples (washing >7 samples using the magnet is time consuming). It takes ~45 to 60 min for cell lysis, mRNA capture to the beads, and washing steps. At this point, the mRNA on the beads can be frozen and stored at -80°C . The subsequent cDNA synthesis, tailing reaction, and PCR amplification must be performed without interruption. cDNA synthesis including the wash steps will take ~1.5 hr and the tailing reaction will take an additional 1.5 hr. The PCR will take 3 to 4 hr and can be run overnight.

Non-radioactive gene expression analysis on nylon arrays

It takes ~30 min to set up the labeling PCR and the PCR itself will take ~1.5 hr. Pre-hybridization of samples requires at least 6 hr when cDNA arrays are used. Hybridize the labeled probe over 2 nights when few cells were used; cDNA from higher cell numbers might be hybridized for 1 night. Non-radioactive development of filters will require ~3 hr. The exposure time of the film has to be individually evaluated, but usually two films developed at 15 and 60 min are sufficient.

Data analysis of hybridized cDNA arrays

Scanning of the films will take ~10 min per film and data analysis by array vision will take 30 to 60 min per film.

Literature Cited

- Belyavsky, A., Vinogradova, T., and Rajewsky, K. 1989. PCR-based cDNA library construction: General cDNA libraries at the level of a few cells. *Nucleic Acids Res.* 17:2919-2932.
- Brady, G. 2000. Expression profiling of single mammalian cells—small is beautiful. *Yeast* 17:211-217.
- Brady, G. and Iscove, N.N. 1993. Construction of cDNA libraries from single cells. *Methods Enzymol.* 225:611-623.
- Brail, L.H., Jang, A., Billia, F., Iscove, N.N., Klamut, H.J., and Hill, R.P. 1999. Gene expression in individual cells: Analysis using global single cell reverse transcription polymerase chain reaction (GSC RT-PCR). *Mutat.-Res.* 406:45-54.

- Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M., and Coleman, P. 1992. Analysis of gene expression in single live neurons. *Proc. Natl. Acad. Sci. U.S.A.* 89:3010-3014.
- Kacharina, J.E., Crino, P.B., and Eberwine, J. 1999. Preparation of cDNA from single cells and subcellular regions. *Methods Enzymol.* 303:3-18.
- Klein, C.A., Seidl, S., Petat-Dutter, K., Offner, S., Geigl, J.B., Schmidt-Kittler, O., Wendler, N., Passlick, B., Huber, R.M., Schlimok, G., Baeuerle, P.A., and Riethmuller, G. 2002. Combined transcriptome and genome analysis of single micrometastatic cells. *Nat. Biotechnol.* 20:387-392.
- Luo, L., Salunga, R.C., Guo, H., Bittner, A., Joy, K.C., Galindo, J.E., Xiao, H., Rogers, K.E., Wan, J.S., Jackson, M.R., and Erlander, M.G. 1999. Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nat. Med.* 5:117-122.
- Zohlhofer, D., Richter, T., Neumann, F., Nuhrenberg, T., Wessely, R., Brandl, R., Murr, A., Klein, C.A., and Baeuerle, P.A. 2001a. Transcriptome analysis reveals a role of interferon-gamma in human neointima formation. *Mol. Cell.* 7:1059-1069.
- Zohlhofer, D., Klein, C.A., Richter, T., Brandl, R., Murr, A., Nuhrenberg, T., Schomig, A., Baeuerle, P.A., and Neumann, F.J. 2001b. Gene expression profiling of human stent-induced neointima by cDNA array analysis of microscopic specimens retrieved by helix cutter atherectomy: Detection of FK506-binding protein 12 upregulation. *Circulation* 103:1396-1402.

Contributed by Christoph A. Klein, Dietlind Zohlhöfer, Karina Petat-Dutter, and Nicole Wendler
Ludwig-Maximilians-University of Munich
Munich, Germany

CHAPTER 26

Gene Silencing

INTRODUCTION

The phenomenon of gene silencing mediated by short double-stranded RNA molecules was only discovered a few years ago, and many of the underlying molecular mechanisms involved remain to be elucidated. Nevertheless, a variety of techniques have been developed that exploit the use of short interfering RNAs (siRNAs) to study gene function in a variety of eukaryotic hosts. siRNAs were discovered in the course of experiments that involved the use of antisense RNA molecules to specifically knock down the expression of a particular target gene. Now it appears that siRNA-mediated gene silencing is an important component of the eukaryotic immune response to viral infection that has been highly conserved in evolution. Short double-stranded RNA molecules also appear to play an important role in gene regulation in plants, animals, and fungi. This high level of evolutionary conservation presumably explains the broad applicability of RNA interference technology (referred to as RNAi) to modulate gene expression in both readily manipulable genetic hosts such as *Caenorhabditis elegans*, *Drosophila melanogaster*, and *Arabidopsis thaliana*, as well as in mammalian tissue culture cells.

RNAi technology has a particularly important role in genome-wide studies of gene function. Given the genome sequence of a target eukaryotic organism, RNAi libraries corresponding to every open reading frame in a genome can be generated, thereby allowing genome-wide screens of gene function. This approach is particularly powerful for *C. elegans* and *D. melanogaster*. In *C. elegans*, RNAi can be accomplished by the relatively straightforward approach of constructing an *E. coli*-based plasmid vector that directs the synthesis of a dsRNA molecule and then simply feeding this plasmid-containing *E. coli* strain to *C. elegans*. Thus, genome-wide *C. elegans* RNAi libraries are being generated which can be systematically fed to *C. elegans*, one clone at a time. Given a particular phenotype, many *C. elegans* genes affecting the phenotype can be readily identified. In *D. melanogaster*, genome-wide RNAi analysis can be carried out by exposing tissue culture cells to short ds RNAs. In contrast, identifying genes that correspond to a particular *C. elegans* or *D. melanogaster* mutant generated by traditional forward genetic approaches is often an arduous process, involving time-intensive map-based cloning strategies.

This chapter begins with an overview of the brief history of RNAi techniques and the underlying science concerning the role of siRNAs in pathogen defense and gene regulation (UNIT 26.1). The overview is followed by a set of protocols for carrying out RNAi in different hosts. The chapter currently has protocols for mammalian cells (UNIT 26.2), *C. elegans* (UNIT 26.3), *D. melanogaster* (UNIT 26.5) and for cloning small endogenous RNA (UNIT 26.4); in future supplements, we anticipate including RNAi protocols for plants and perhaps other hosts as well. Each of these protocols is highly species-specific, focusing on the most efficient methods for delivering dsRNA molecules to a particular host. As more is understood about the underlying molecular machinery involved in mediating RNAi, we anticipate that much more efficient RNAi protocols will be developed. We will therefore update the protocols to keep them current with new advances in this rapidly evolving field.

Frederick M. Ausubel

Overview of RNA Interference and Related Processes

HISTORY

The history of RNA interference (RNAi) has unfolded rapidly since 1997 with a series of discoveries from plants, fungi, and animals. Initially, the interest was directed towards development of gene silencing technology that could be applied in research, medical therapy, and crop improvement. However, as the underlying mechanism was revealed, it became apparent that RNAi manifests a novel system of genetic regulation that was only hinted at by previous data.

A starting point for this history is the 1980s when plant researchers encountered two unexpected complications with antisense technology. First, most of the plant lines with an antisense transgene did not exhibit suppression of the corresponding endogenous gene (Smith et al., 1988). Second, some of the control lines with the sense constructs exhibited coordinate suppression (cosuppression) of the transgene and the homologous endogenous gene (Napoli et al., 1990; Smith et al., 1990; van der Krol et al., 1990).

From the initial findings, it was not clear whether the antisense and cosuppression mechanisms were related, although both involved a block at the post-transcriptional level (Kooter and Mol, 1993). One school of thought was that sense transgenes were transcribed into antisense RNA by adjacent plant promoters (Grierson et al., 1991). Other views invoked different sense and antisense mechanisms. The antisense mechanisms would be a simple block on the target RNA, whereas sense suppression would be mediated through a system that could detect gene over-expression or transgene transcripts that were structurally aberrant (Jorgensen, 1995).

At about the same time, there was a similar controversy in the *Caenorhabditis elegans* field. Gene expression could be specifically suppressed by direct injection of antisense RNA; however, the process was inefficient because large amounts of antisense RNA had to be injected (Fire et al., 1991). Additionally, there was a remarkable resonance of the plant “antisense versus cosuppression” mystery, because injected sense RNA was as potent as antisense RNA (Guo and Kemphues, 1995).

An inspired analysis of the *C. elegans* phenomenon provided the key to understanding

these unexpected findings. It revealed that RNAi in injected *C. elegans* was mediated by a small amount of double-stranded (ds) RNA that contaminated the sense and antisense RNA preparations (Montgomery et al., 1998). Subsequently, it was established in plants that if dsRNA was produced, suppression of endogenous gene expression was more efficient than with sense or antisense transgenes (Waterhouse et al., 1998; Chuang and Meyerowitz, 2000). Inverted repeat transgenes were particularly efficient, but simultaneous expression of sense and antisense RNA was also effective (Waterhouse et al., 1998).

A second important discovery followed from the search for the specificity determinant of cosuppression in plants. It was reasoned that, as cosuppression is nucleotide sequence-specific and acts at the RNA level, there should be antisense RNA corresponding to the target species. Presumably this antisense RNA would guide the degradation of the cosuppressed RNA. An initial unsuccessful search for this antisense RNA focused on RNA of >100 nucleotides length. However, when the hunt was redirected to small RNAs, an antisense species of ~25 nucleotides length was discovered (Hamilton and Baulcombe, 1999). These short RNAs also corresponded to the sense strand of the cosuppression target and it seemed likely that they were derived from a dsRNA precursor. This link with RNAi was subsequently confirmed when short RNAs were associated with RNAi in *Drosophila melanogaster* (Hammond et al., 2000). The *D. melanogaster* work also confirmed the prediction that the short RNAs guide a ribonuclease complex (RISC) to its target RNA.

The discovery that short RNAs play a key role in RNAi-mediated suppression of gene expression precipitated an avalanche of discoveries that are relevant to the application of RNAi and cosuppression, as well as to understanding of the natural roles of these processes. For example, the short RNAs have been characterized in detail and are now known to exist in a double-stranded form, with two-nucleotide overhangs at each 3' end, and are known as siRNAs (Elbashir et al., 2001a). Several different proteins have been identified that are associated with siRNAs in ribonucleoproteins (Caudy et al., 2002; Mourelatos et al., 2002).

Contributed by David Baulcombe

Current Protocols in Molecular Biology (2003) 26.1.1-26.1.6
Copyright © 2003 by John Wiley & Sons, Inc.

At least some of these proteins are part of RISC (Hammond et al., 2001). It is also now understood how dsRNA is processed into the siRNAs. The processing enzyme is known as Dicer and is a member of the RNase III family with dsRNA binding regions and a conserved PAZ domain shared with members of the piwi, argonaute and zwiille family (Bernstein et al., 2001), from which it takes its name. ATP is required at several stages in the processing of dsRNA and assembly of RISC (Nykanen et al., 2001). In some instances an RNA-dependent RNA polymerase is also involved (Cogoni and Macino, 1999; Dalmay et al., 2000; Mourrain et al., 2000). It converts a single-stranded RNA into a double-stranded siRNA precursor.

NATURAL ROLES OF RNAi AND RELATED PROCESSES

In plants, a natural role of RNAi and cosuppression is as an antiviral defense system (Voinnet, 2001). Viruses produce a dsRNA replication intermediate that is the substrate for Dicer and that leads to production of siRNA. The accumulation of siRNA guides RISC to the single-stranded viral RNA so that a feedback system is established in which there is an equilibrium between the level of RISC and the rate of virus accumulation. Accordingly, viruses produce suppressor proteins of RNAi as a counter-defense system (Voinnet, 2001). It is thought that if a suppressor is weak then virus accumulation would be transient, whereas a strong suppressor would permit prolonged accumulation to a high level. An antiviral role in insects is also likely as a nodavirus encodes a protein with suppressor activity (Li et al., 2002). At present it is not known whether the mechanism has an antiviral role in vertebrates; however, the recent findings that siRNAs have antiviral activity in mammalian cells are certainly consistent with that possibility (Novina et al., 2002).

The RNAi mechanism may also provide protection against transposons or other genome perturbations. Thus, in plants and protozoans there are endogenous short RNAs from retroelements (Hamilton et al., 2002), while in *C. elegans* mutations that affect RNAi may also cause genome instability (Ketting et al., 1999). Presumably, retroelement double-stranded RNA is processed by Dicer into short RNAs that prevent expression and mobilization of the corresponding genomic elements.

Many of the proteins required for RNAi are conserved in animals, plants, and fungi

(Zamore, 2001). It is therefore likely that the mechanism evolved in primitive eukaryotes as a defense system against viruses and selfish DNA, as in modern plants. However, it seems that the RNAi mechanisms have been recruited during evolution of modern eukaryotes into other aspects of genetic regulation. Thus, in fungi, proteins similar to those required for RNAi are required for transcriptional silencing of unpaired DNA or of centromeric genes (Shiu et al., 2001; Volpe et al., 2002). In *Tetrahymena*, a PAZ-domain protein and short RNAs that may be similar to siRNAs are implicated in control of genome rearrangement during differentiation of the transcriptionally active macronucleus (Mochizuki et al., 2002). It is not known at present whether RISC is involved in these effects at the genome level. The alternative possibility is that a separate complex recruits short RNAs and targets DNA rather than RNA.

Additional roles of RNAi-related mechanisms are implied from the identification of microRNAs (miRNAs) in plants and animals (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001; Reinhart et al., 2002; Rhoades et al., 2002). These miRNAs are similar to siRNAs in that they are produced by Dicer-like enzymes and are a similar length. They may also be recruited into RISC and mediate targeted RNA degradation or translational arrest. Other effects on, for example RNA processing, are also possible (Mishra and Handa, 1998). However, unlike siRNAs that are from transposons, transgenes, or viruses, the miRNAs are transcribed from endogenous DNA. A typical dsRNA precursor of miRNA is the transcript of inverted-repeat DNA in regions between conventional genes.

The likely role of miRNAs is in the control of mRNA translation or stability. Thus, the first miRNAs to be described were *lin4* and *let7* RNAs in *C. elegans* (Ruvkun, 2001). They mediate temporal control over translation of *lin14* and *lin41* mRNAs during *C. elegans* development. More recently, an miRNA in *Arabidopsis* (*miR171*) has been identified that apparently regulates site-specific cleavage of a transcription factor mRNA (Llave et al., 2002a).

At the time of writing this overview, *lin4*, *let7*, and *miR171* were unusual in that their mRNA targets had been identified. With *miR171*, this was relatively straightforward because it is perfectly matched to the target mRNA (Llave et al., 2002a). Other plant miRNAs may have fewer than three mismatches and putative targets have been identified by computational analysis of genome sequence

(Rhoades et al., 2002); however, these targets have not been verified. In contrast, for most animal miRNAs, the putative miRNA targets cannot be easily identified because the mismatch is greater than in plants. A strategy for miRNA target identification is a major outstanding challenge.

STRATEGIES FOR RNAi AND COSUPPRESSION IN PLANTS AND ANIMALS

RNAi and cosuppression are promising new approaches to genetic analysis and manipulation. It should be possible, for example, to infer the function of an uncharacterized gene, to correct dominant genetic defects, and to develop antiviral strategies using this approach. RNAi and related procedures may have several advantages over alternative strategies or may provide solutions to currently intractable problems. For example, if a gene of interest is a member of a multigene family, it may not be possible to infer function from standard genetic knockout strategies. The expression of functionally redundant gene family members may compensate for the mutant gene. However, because the RNAi-based procedures are nucleotide-sequence rather than genetic-locus specific, they allow suppression of many similar family members with a single manipulation.

Despite the relative novelty of RNAi and cosuppression technologies, there are already many different procedures that can be used, depending on the organism and the function of the target RNA. The most cumbersome of these involve transformation of the organism to produce a dsRNA precursor of an siRNA. This approach has been used in plants and animals (Wesley et al., 2001; Giordano et al., 2002; Paddison et al., 2002; Sui et al., 2002) and has the advantage that the phenotype is stable through several generations.

More direct approaches do not require stable transformation and are particularly amenable to high-throughput applications, including genome-wide surveys of gene function. It is possible, for example, to transfect with dsRNA or siRNA targeted at a gene of interest (Elbashir et al., 2001b; Harborth et al., 2001). If chemically synthesized siRNAs are used, this approach is relatively expensive. However siRNA produced in vitro by enzymatic cleavage of dsRNA may be a less expensive alternative (Yang et al., 2002).

Direct introduction of dsRNA or siRNA is effective in cultured cells. Surprisingly, however, direct introduction of siRNA or dsRNA

to intact multicellular animals can also result in RNAi. With *C. elegans*, for example, it is necessary only to feed animals on *E. coli* that is transformed to produce the dsRNA (Timmons and Fire, 1998). Thus, by producing a set of *E. coli* strains with dsRNA targeted against each gene, it was possible to generate a resource that will greatly facilitate characterization of the *C. elegans* genome (Kamath et al., 2003). Short dsRNA or siRNA could also be used in adult mice to target transgene expression in the liver (McCaffrey et al., 2002).

In plants, RNAi can be achieved in transgenic plants expressing the dsRNA of the target. An alternative strategy involves infecting plants with virus vectors carrying endogenous gene fragments. The corresponding host RNA is suppressed in the infected plant and the symptoms resemble a loss of function mutant phenotype. This approach has been refined on the model plant *Nicotiana benthamiana* and has been used to investigate genes involved in cell wall biosynthesis and disease resistance (Burton et al., 2000; Peart et al., 2002a,b). The potential of virus-mediated RNAi has not been fully explored in animals or animal cells.

DETECTION AND CHARACTERIZATION OF siRNAs AND miRNAs

A direct procedure for detection of siRNA and miRNA is northern blotting. The procedures are similar to the standard protocol for detection of mRNA, but are modified to account for the small size of these RNAs. Thus, the RNA is fractionated by electrophoresis on a polyacrylamide gel instead of agarose and electroblotted onto a nylon membrane. The stringency of the hybridization is also reduced. Probes can be labeled cDNA, cRNA, or oligonucleotides (Hamilton and Baulcombe, 1999).

Depending on its abundance, the short RNA may be detectable in total RNA preparations. However, for rare species it is necessary to enrich for short RNA by selective precipitation or enrichment for short RNA protein complexes (Mourelatos et al., 2002). Alternatively, for rare short RNAs, RNase protection can be used instead of northern blotting. However the cost of the increased sensitivity is the loss of size resolution—i.e., rather than appearing as a discrete-sized species, the RNase protection reveals the short RNA as a smear (Sijen et al., 2001).

There are several procedures for cloning of cDNA corresponding to siRNA or miRNA

(Djikeng et al., 2001; Lagos-Quintana et al., 2001; Llave et al., 2002b). These procedures all involve addition of adaptors to the termini of the short RNA using RNA ligase followed by reverse transcription and PCR amplification. The various protocols employ different strategies to avoid circularization or concatemerization of the reaction intermediates. However, some of the protocols do allow for concatemerization of the PCR products prior to cloning so that a single sequence reaction will be informative about several short RNA species.

LITERATURE CITED

- Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409:363-366.
- Burton, R.A., Gibeaut, D.M., Bacic, A., Findlay, K., Roberts, K., Hamilton, A., Baulcombe, D.C., and Fincher, G.B. 2000. Virus-induced silencing of a plant cellulose synthase gene. *Plant Cell* 12:691-705.
- Caudy, A.A., Myers, M., Hannon, G.J., and Hammond, S.M. 2002. Fragile X-related protein and VIG associate with the RNAi machinery. *Genes Dev.* 16:2491-2496.
- Chuang, C.-H. and Meyerowitz, E.M. 2000. Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 97:4985-4990.
- Cogoni, C. and Macino, G. 1999. Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase. *Nature* 399:166-169.
- Dalmay, T., Hamilton, A.J., Rudd, S., Angell, S., and Baulcombe, D.C. 2000. An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* 101:543-553.
- Djikeng, A., Shi, H., Tschudi, C., and Ullu, E. 2001. RNA interference in *Trypanosoma brucei*: Cloning of small interfering RNAs provides evidence for retroposon-derived 24-26 nucleotide RNAs. *RNA* 7:1522-1530.
- Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. 2001a. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J.* 20:6877-6888.
- Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. 2001b. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411:494-498.
- Fire, A., Albertson, D., Harrison, S., and Moerman, D. 1991. Production of antisense RNA leads to effective and specific inhibition of gene expression in *C. elegans*. *Devel.* 113:503-514.
- Giordano, E., Rendina, R., Peluso, I., and Furia, M. 2002. RNAi triggered by symmetrically transcribed transgenes in *Drosophila melanogaster*. *Genetics* 160:637-648.
- Grierson, D., Fray, R.G., Hamilton, A.J., Smith, C.J.S., and Watson, C.F. 1991. Does co-suppression of sense genes in transgenic plants involve antisense RNA? *Trends Biotech.* 9:122-123.
- Guo, S. and Kemphues, K. 1995. Par-1, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell* 81:611-620.
- Hamilton, A.J. and Baulcombe, D.C. 1999. A novel species of small antisense RNA in post-transcriptional gene silencing. *Science* 286:950-952.
- Hamilton, A.J., Voinnet, O., Chappell, L., and Baulcombe, D.C. 2002. Two classes of short interfering RNA in RNA silencing. *EMBO J.* 21:4671-4679.
- Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G. 2000. An RNA-directed nucleic acid mediated post-transcriptional gene silencing in *Drosophila* cell extracts. *Nature* 404:293-296.
- Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. 2001. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* 293:1146-1150.
- Harborth, J., Elbashir, S.M., Bechert, K., Tuschl, T., and Weber, K. 2001. Identification of essential genes in cultured mammalian cells using small interfering RNAs. *J. Cell Sci.* 114:4557-4565.
- Jorgensen, R.A. 1995. Cosuppression, flower color patterns and metastable gene expression states. *Science* 268:686-691.
- Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D.P., Zipperlen, P., and Ahringer, J. 2003. Systemic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* 421:231-237.
- Ketting, R., Haverkamp, T., van Luenen, H., and Plasterk, R. 1999. *mut-7* of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell* 99:133-141.
- Kooter, J.M. and Mol, J.N.M. 1993. *Trans*-inactivation of gene expression in plants. *Plant Biotech.* 4:166-171.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. 2001. Identification of novel genes coding for small expressed RNAs. *Science* 294:853-858.
- Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. 2001. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294:858-862.
- Lee, R. and Ambros, V. 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294:862-864.
- Li, H., Li, W.X., and Ding, S.W. 2002. Induction and suppression of RNA silencing by an animal virus. *Science* 296:1319-1321.

- Llave, C., Xie, Z., Kasschau, K.D., and Carrington, J.C. 2002a. Cleavage of *Scarecrow-like* mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* 297:2053-2056.
- Llave, C., Kasschau, K.D., Rector, M.A., and Carrington, J.C. 2002b. Endogenous and silencing-associated small RNAs in plants. *Plant Cell* 14:1-15.
- McCaffrey, A.P., Meuse, L., Pham, T.T.T., Conklin, D.S., Hannon, G.J., and Kay, M.A. 2002. Gene expression-RNAi in adult mice. *Nature* 418:38-39.
- Mishra, K.K. and Handa, A.K. 1998. Post-transcriptional silencing of pectin methylesterase gene in transgenic tomato fruits results from impaired pre-mRNA processing. *Plant J.* 14:583-592.
- Mochizuki, K., Fine, N.A., Fujisawa, T., and Gorovsky, M.A. 2002. Analysis of a *piwi*-related gene implicates small RNAs in genome rearrangement in *Tetrahymena*. *Cell* 110:689-699.
- Montgomery, M.K., Xu, S.Q., and Fire, A. 1998. RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* 95:15502-15507.
- Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. 2002. miRNPs: A novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev.* 16:720-728.
- Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.-B., Jouette, D., Lacombe, A.-M., Nikic, S., Picault, N., Remoue, K., Sanial, M., Vo, T.-A., and Vaucheret, H. 2000. *Arabidopsis* *SGS2* and *SGS3* genes are required for post-transcriptional gene silencing and natural virus resistance. *Cell* 101:533-542.
- Napoli, C., Lemieux, C., and Jorgensen, R.A. 1990. Introduction of a chimeric chalcone synthase gene into *Petunia* results in reversible co-suppression of homologous genes *in trans*. *Plant Cell* 2:279-289.
- Novina, C.D., Murray, M.F., Dykxhoorn, D.M., Beresford, P.J., Riess, J., Lee, S.-K., Collman, R.G., Lieberman, J., Shankar, P., and Sharp, P.A. 2002. siRNA-directed inhibition of HIV-1 infection. *Nat. Med.* 8:681-686.
- Nykanen, A., Haley, B., and Zamore, P.D. 2001. ATP requirements and small interfering RNA structure in the RNAi pathway. *Cell* 107:309-321.
- Paddison, P.J., Caudy, A.A., and Hannon, G.J. 2002. Stable suppression of gene expression by RNAi in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 99:1443-1448.
- Peart, J.R., Cook, G., Feys, B.J., Parker, J.E., and Baulcombe, D.C. 2002a. An *EDS1* orthologue is required for *N*-mediated resistance against tobacco mosaic virus. *Plant J.* 29:569-579.
- Peart, J.R., Lu, R., Sadanandom, A., Malcuit, I., Moffett, P., Brice, D.C., Schauser, L., Jaggard, D.A.W., Xiao, S., Coleman, M., Dow, J.M., Jones, J.D.G., Shirasu, K., and Baulcombe, D.C. 2002b. Ubiquitin ligase-associated protein SGT1 is required for host and nonhost disease resistance in plants. *Proc. Natl. Acad. Sci. U.S.A.* 99:10865-10869.
- Rheinhardt, B.J., Weinstein, E.G., Rhoades, M., Bartel, B., and Bartel, D.P. 2002. MicroRNAs in plants. *Genes Dev.* 16:1616-1626.
- Rhoades, M.W., Reinhardt, B.J., Lim, L.P., Burge, C.B., Bartel, B., and Bartel, D.P. 2002. Prediction of plant microRNA targets. *Cell* 110:513-520.
- Ruvkun, G. 2001. Glimpses of a tiny RNA world. *Science* 294:797-799.
- Shiu, P.K.T., Raju, N.B., Zickler, D., and Metzberg, R.L. 2001. Meiotic silencing by unpaired DNA. *Cell* 107:905-916.
- Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H.A., and Fire, A. 2001. On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* 107:465-476.
- Smith, C.J.S., Watson, C.F., Ray, J., Bird, C.R., Morris, P.C., Schuch, W., and Grierson, D. 1988. Antisense RNA inhibition of polygalacturonase gene expression in transgenic tomatoes. *Nature* 334:724-726.
- Smith, C.J.S., Watson, C.F., Bird, C.R., Ray, J., Schuch, W., and Grierson, D. 1990. Expression of a truncated tomato polygalacturonase gene inhibits expression of the endogenous gene in transgenic plants. *Mol. Gen. Genet.* 224:447-481.
- Sui, G., Soohoo, C., Affar, E.B., Gay, F., Shi, Y., Forrester, W.C., and Shi, Y. 2002. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 99:5515-5520.
- Timmons, L. and Fire, A. 1998. Specific interference by ingested dsRNA. *Nature* 395:854.
- van der Krol, A.R., Mur, L.A., Beld, M., Mol, J.N.M., and Stuitji, A.R. 1990. Flavonoid genes in *petunia*: Addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell* 2:291-299.
- Voinnet, O. 2001. RNA silencing as a plant immune system against viruses. *Trends Genet.* 17:449-459.
- Volpe, T., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I.S., and Martienssen, R. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* 297:1833-1837.
- Waterhouse, P.M., Graham, H.W., and Wang, M.B. 1998. Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc. Natl. Acad. Sci. U.S.A.* 95:13959-13964.
- Wesley, S.V., Helliwell, C.A., Smith, N.A., Wang, M., Rouse, D.T., Liu, Q., Gooding, P.S., Singh, S.P., Abbott, D., Stoutjesdijk, P.A., Robinson, S.P., Gleave, A.P., Green, A.G., and Waterhouse, P. 2001. Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant J.* 27:581-590.

Yang, D., Buchholz, F., Huang, Z.D., Goga, A., Chen, C.Y., Brodsky, F.M., and Bishop, J.M. 2002. Short RNA duplexes produced by hydrolysis with *Escherichia coli* RNase III mediate effective RNAi in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 99:9942-9947.

Zamore, P.D. 2001. RNA interference: Listening to the sound of silence. *Nat. Struct. Biol.* 8:746-750.

Contributed by David Baulcombe
The Sainsbury Laboratory
John Innes Centre
Norwich, United Kingdom

For many years, the specific and efficient modulation of mammalian gene expression has been a very difficult and labor-intensive task. Particularly promising technologies for controlling gene expression are antisense oligonucleotides and ribozymes. In the case of antisense technology, there are a number of examples in which reasonably strong down-regulation of specific target genes was obtained. In most cases, however, a tedious and costly process of target-sequence optimization has been necessary to obtain satisfactory results.

The recent discovery of gene-specific silencing in mammalian cells mediated by ~20- to 25-nucleotide double-stranded RNA, better known as short interfering RNA (siRNA), allows researchers to circumvent many problems of the traditional antisense protocols. Paying attention to a few rules of oligoribonucleotide design and determining an efficient cell transfection procedure for a particular cell line are the remaining obstacles. RNA interference works in a wide variety of cell types and target sequence identification is straightforward. In most cases, it is sufficient to design one to five different siRNA molecules for each target gene.

The induction of siRNA-mediated RNA interference (RNAi) in mammalian cells can be achieved by microinjection of siRNA, transfection of siRNA using cationic lipids, electroporation of siRNA, transfection of plasmids containing siRNA expression cassettes, or infection of cells by recombinant viruses that express siRNA. This unit contains a method for siRNA transfection (see Basic Protocol) and an alternative for electroporating siRNA (see Alternate Protocol). In addition, procedures for annealing siRNA strands (see Support Protocol 1) and for cotransfecting modified reporter genes with siRNA (see Support Protocol 2) are also provided. Besides a specific experimental siRNA, all experiments have to include at least one control siRNA unrelated to the target mRNA to evaluate nonspecific or toxic effects.

NOTE: All solutions and equipment coming into contact with cells must be sterile, and proper sterile technique should be used accordingly.

NOTE: All culture incubations should be performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

NOTE: All solutions and materials coming into contact with RNA must be RNase free, and proper techniques should be used accordingly (see UNIT 4.1 and APPENDIX 2).

LIPOSOME-MEDIATED TRANSFECTION OF MAMMALIAN CELLS WITH siRNA

BASIC PROTOCOL

The liposome transfection of mammalian cells with siRNA shares some similarity with DNA transfection protocols (see UNIT 9.4 for additional information). Similar considerations apply for siRNA and DNA transfection protocols, although special transfection reagents were developed in order to achieve optimal transfection efficiencies with siRNA. For some cell lines, specific DNA transfection reagents such as Lipofectamine PLUS (Invitrogen) appear to work the best.

Materials

- Mammalian cells to be transfected (e.g., HeLa-S3)
- Complete medium (APPENDIX 3F)
- 20 μM annealed siRNA (see Support Protocol 1)

Contributed by Matthias John, Anke Geick, Philipp Hadwiger, Hans-Peter Vornlocher, and Olaf Heidenreich

Current Protocols in Molecular Biology (2003) 26.2.1-26.2.14
Copyright © 2003 by John Wiley & Sons, Inc.

Gene Silencing

26.2.1

Supplement 62

Serum-free medium (*APPENDIX 3F*)

Lipofectamine PLUS (Invitrogen)

12-well culture dishes (2-cm diameter)

1.5-ml polypropylene tubes

Additional reagents and equipment for mammalian cell tissue culture, including trypsinization (*APPENDIX 3F*)

NOTE: Lipofectamine PLUS is used in this protocol as it works well for HeLa-3 cells. Other transfection reagents can also be used by following manufacturer's instructions. Lipofectamine PLUS consists of two reagents, the transfection reagent itself (Lipofectamine) and an enhancer (PLUS).

Culture cells

1. Cultivate the mammalian cells to be transfected to late log phase in complete medium (*APPENDIX 3F*).
2. Trypsinize and count cells (*APPENDIX 3F*). Transfer to appropriate tubes and centrifuge 5 min at $300 \times g$, room temperature. Discard supernatant and wash once with complete medium. Resuspend in complete medium.
3. Seed 1 ml containing 1×10^5 cells in each well of a 12-well tissue culture plate and incubate ~24 hr.
4. Replace the medium with 500 μ l fresh medium 4 hr before the start of transfection and return to the incubator.

Perform transfection

5. For each well, mix 1.25 μ l of 20 μ M annealed siRNA with 43.75 μ l serum-free medium in a 1.5-ml polypropylene tube. Thoroughly vortex PLUS enhancer component of Lipofectamine PLUS reagent and add 5 μ l to each well.
6. In a second 1.5-ml reaction tube, dilute 2.5 μ l Lipofectamine component in 47.5 μ l serum-free medium.
7. Incubate the two tubes for 15 min at room temperature.
8. Add the Lipofectamine solution (step 6) to the tube containing the siRNA/PLUS mixture (step 5; 100 μ l total) and mix gently. Incubate 15 min at room temperature.
9. Remove the complete medium from each well and replace it with 400 μ l serum-free medium.
10. Add all of the siRNA transfection mix (100 μ l) from step 8.
Using this protocol as written (i.e., 20 μ M siRNA stock), the final concentration of siRNA is 50 nM. The best siRNA concentration varies depending on cell line or target gene and is likely between 1 to 100 nM. If the siRNA concentration is too high, Lipofectamine PLUS can have toxic effects on some cell types.
11. Incubate 2.5 hr, remove the transfection medium, and replace it with 1 ml complete medium.
12. Incubate 16 to 48 hr and analyze RNA by Northern blot (*UNIT 4.9*) or RNase protection assay (*UNIT 4.7*). Alternatively, determine resulting protein levels by immunoblot (*UNIT 10.8*).

In most cases a decrease in target RNA level is accompanied by simultaneous depletion of the corresponding protein in cell extracts. However, some proteins may have a half life of many hours in vivo and thus, the time course of RNA and protein degradation are not the same. In those, cases the cells have to be harvested at different time points. For example RNA preparation can be done after 48 hr, whereas protein extracts have to be prepared after 72 or 96 hr.

ELECTROPORATION OF MAMMALIAN CELLS WITH siRNA

ALTERNATE PROTOCOL

Many mammalian cell types can be easily and efficiently electroporated under very mild conditions resulting in cell survival rates >90%. Particularly with suspension cells, siRNA delivery by electroporation is superior to liposome-mediated transfection. Tests with fluorescently labeled siRNA and subsequent analysis by fluorescence-activated cell sorting indicate that up to 100% of Kasumi-1 cells can be transfected (Heidenreich et al., 2002). An additional advantage is that the electroporation protocol needs no particular transfection reagents and requires only a few steps.

An increasing number of laboratories are now discovering the advantages of using electroporation for siRNA transfection instead of liposome-based methods. However, very high electroporation efficiencies are only observed for particular cell lines. Suspension cells, as for example Kasumi-1 or other hematopoietic cells, are good candidates for this method. These suspension cells seem to work well with electroporation and are often difficult to transfect with liposome-forming reagents. Nonetheless, exceptions are observed among suspension cells and there is no general advantage to electroporate adherent cells.

Additional Materials (also see Basic Protocol)

4-mm electroporation cuvettes (e.g., Equibio ECU104)
Square-wave electroporator (e.g., Fischer EPI 2500)
Culture vessel

1. Cultivate, trypsinize, and wash mammalian cells as described (see Basic Protocol, steps 1 and 2). Resuspend in complete medium at a final concentration of $1\text{--}2 \times 10^7$ cells/ml.
2. Transfer 100 to 800 μl cell suspension into a 4-mm electroporation cuvette at room temperature. Add 20 μM annealed siRNA to a final concentration of 100 to 200 nM. Mix well by flicking the cuvette.
3. Place cuvette into the holder of a square-wave electroporator and electroporate at the desired voltage and time setting.

For hematopoietic cell lines such as BAF3, K562, or Kasumi-1, settings of 300 to 350 V and 10 msec are recommended. To establish the electroporation conditions, the authors routinely electroporate 10^6 cells in 100 μl medium containing 200 nM of the corresponding siRNA for a constant time of 10 msec and varying voltages ranging from 300 to 400 V. Some electroporation systems offer more variables (e.g., voltage, pulse time, resistance, impedance). The settings have to be determined for each cell line empirically. See UNIT 9.3 for more detail.

4. Remove cuvette from holder and incubate 15 min at room temperature.
5. Return the electroporated cells to a culture vessel and dilute the cell suspension to the desired cell density.

For electroporated Kasumi-1 cells, a cell density of $5 \times 10^5/\text{ml}$ is appropriate. For hematopoietic suspension cells, transfection rates of up to 100% can be obtained with a fluorescently labeled siRNA; however, for fibroblast cell lines such as NIH3T3, the authors observed rather poor transfection rates.

6. Analyze RNA (UNITS 4.7 & 4.9), protein expression (UNIT 10.8), and consequences of target gene depletion.

Decreases in RNA levels are visible within 16 hr after electroporation. Reduction of protein levels and the consequences of the protein depletion depend on the half-life of the targeted protein. In the case of electroporated Kasumi-1 cells, the authors observe ~70% reduction of endogenous gene expression.

Gene Silencing

26.2.3

ANNEALING EQUIMOLAR siRNA

A detailed discussion of how to design and where to obtain siRNA is given in the Commentary section of this unit (see Critical Parameters).

Materials

siRNA sense and antisense strands (see Critical Parameters)

Annealing buffer: 100 mM NaCl in 20 mM sodium phosphate buffer, pH 6.8 (see APPENDIX 2 for buffer)

Additional reagents and equipment for nondenaturing polyacrylamide gel electrophoresis (UNIT 2.7) and ethidium bromide staining (UNIT 4.9)

1. Combine equimolar amounts of sense and antisense strands and adjust the siRNA concentration to 20 μ M (double stranded) in a final volume of 100 to 500 μ l annealing buffer.
2. Heat a glass beaker containing 500 ml water to 90°C on a magnetic stirrer with adjustable temperature. Place tightly capped 1.5-ml reaction tubes containing the RNA in a flotation device and transfer to the water bath. After 3 min switch the heating element off.
3. Allow the water bath to slowly cool to room temperature (~3 hr) to permit the siRNA to anneal.
4. Load ~1.5 μ g of both single-stranded and double-stranded siRNA on a 10% nondenaturing polyacrylamide gel (UNIT 2.7). Visualize by ethidium bromide staining as described for northern hybridization of RNA (UNIT 4.9).

Samples of properly annealed siRNA run slower compared to sense and antisense strands and show no signs of remaining single-stranded RNA. This is apparent on an ethidium bromide-stained gel when the fluorescent signal from the single-stranded species is below the detection limit during illumination by UV light. It is important to remember, however, that ethidium bromide interaction with single-stranded RNA is weak compared to double-stranded RNA which can cause underestimation of single-stranded RNA quantities. Also bear in mind that even properly annealed siRNA tends to dissociate during the acrylamide-gel run and minute amounts of smeared single-stranded RNA will always be detectable with other staining procedures (e.g., "stains all" dye) that visualize single- and double-stranded RNA with the same efficiency. In the authors' hands, the more common ethidium bromide staining procedure is sufficient for estimating success.

5. Store the siRNA stock solutions up to several years at -80°C.

Smaller samples of siRNA can be stored at -20°C for at least 6 months and repeated freezing/thawing cycles do not decrease their quality. Do not keep solutions of highly diluted siRNA for more than an hour, as the double-stranded molecules tend to dissociate. The stability of siRNA at room temperature depends on its sequence and the composition of the solvent (e. g. serum-free medium). Keep in mind that not only RNases are harmful to RNA. In addition, the pH and salt composition of the solvent may accelerate RNA decay in solution. Generating fresh siRNA dilutions for a new experiment is rather simple and more convenient than determining the shelf life of individual siRNAs empirically.

LIPOSOME-MEDIATED COTRANSFECTION OF REPORTER GENES WITH siRNA

Developing an efficient siRNA protocol for silencing an endogenous gene in a particular cell line can be a very tedious task. In case of failure it is impossible to tell whether the designed siRNA does not work, the transfection method was inefficient, or the target RNA or protein level was monitored within the wrong time window. In addition, the expression level of the target gene can be very low. In the cases where these problems are encountered, the cotransfection of plasmid-encoded reporter genes together with siRNA allows an

estimation of RNAi efficiency in a particular cell line via the silencing of the reporter gene. The overall transfection efficiency is no longer a limiting factor as the reporter plasmids and siRNA are obviously part of the same liposome complexes and therefore, reach the same cells. In addition, the evaluation of almost any target sequence of an endogenous gene or viral gene for siRNA design is simplified by fusing it to a reporter gene.

Presented here as an example is a reporter system consisting of a generated recombinant transcript containing a short sequence derived from a human hepatitis C virus (HCV) genome ligated to the open reading frame of the firefly luciferase gene. This example was chosen as human HCV is very difficult to study in cell culture and thus the use of models and reporter systems is prevalent, meaning that the system is well defined. The widely used reporters β -galactosidase or green fluorescent protein (gfp) tolerate even longer gene fusions. The artificial sequence context and the high expression level obviously does not represent the situation of an endogenous gene; however, the detection of siRNA activity of the reporter is comparatively easy and reproducible, and thus should give an idea of overall RNAi efficiency in the cell line used.

Additional Materials (also see *Basic Protocol*)

Reference plasmid expressing β -galactosidase (e.g., pCMV beta, Clontech)

Plasmid expressing reporter gene fusion (e.g. HCV-Luciferase in pcDNA3.1+, Invitrogen)

TE buffer (*APPENDIX 2*)

5 μ M annealed siRNA (see Support Protocol 1)

GenePORTER 2 kit (Gene Therapy Systems):

DNA diluent B

GenePORTER 2 Reagent (liposome transfection reagent)

Annealing buffer: 100 mM NaCl in 20 mM sodium phosphate buffer, pH 6.8 (see *APPENDIX 2* for buffer)

96-well culture dishes

Additional reagents and equipment for β -galactosidase and firefly luciferase assays (*UNIT 9.7B*)

1. Prepare cells as described (see *Basic Protocol*, steps 1 and 2).

In the present example, the liver cancer cell line HuH-7 was cultivated in DMEM/10% FBS.

2. Seed an appropriate number of wells in a 96-well dish with 100 μ l containing 2×10^4 cells. Incubate ~24 hr.
3. Mix 3 μ g reference plasmid expressing β -galactosidase with 1 μ g of the plasmid carrying the reporter gene fusion. Adjust the total DNA concentration to 375 ng/ μ l with water or TE.
4. Mix the following solutions in reaction tubes for the identical transfection of three wells:

Tube A:

2 μ l 375 ng/ μ l plasmid mix

6.6 μ l 5 μ M annealed siRNA

16.4 μ l DNA diluent B

Tube B:

6 μ l GenePORTER reagent

19 μ l serum-free medium.

Incubate both tubes 5 min at room temperature. Add contents of tube B to tube A and incubate an additional 5 min at room temperature.

5. As a negative control, prepare one transfection mix with the relevant plasmids, but replace the 5 μ M siRNA with 6.6 μ l annealing buffer.
6. Add 16 μ l transfection mix to each cell culture well (step 2) and mix gently by pipetting.

Each well receives 250 ng plasmid and ~100 nM siRNA.

7. Analyze the enzyme activity of β -galactosidase and firefly luciferase after 16 to 48 hr incubation.

Use the protocols of UNIT 9.7B or commercially available kits to determine enzymatic activities.

COMMENTARY

Background Information

The recent demonstration of gene-specific silencing in mammalian cells by ~20- to 25-nucleotide dsRNA, better known as siRNA, has had a very stimulating effect on cell biology research. siRNA leads to a strong reduction of either cellular or viral gene expression (Elbashir et al., 2001; Novina et al., 2002). A straightforward approach combines synthetic siRNA with transfection technologies for transferring the molecules into the cells. This results in a transient decrease in expression of the target gene for 3 to 5 days. The integration of siRNA expression cassettes in the host genome may lead to a sustained suppression of target gene expression, but is more time consuming to establish.

The practical uses of siRNA applications are very appealing; however, the mechanism of RNA interference in mammalian cells is poorly understood. Parallel work that is being carried out in plants, *D. melanogaster*, and *C. elegans* allows some understanding of the underlying mechanisms of RNA interference in mammalian systems (Hannon, 2002; Sharp, 2001; Zamore, 2001).

In *Drosophila*, extended stretches of dsRNA are recognized by dicer, an RNase III-like nuclease. Dicer cleaves the dsRNA into smaller pieces of an average ~22 nucleotides in length. The short double-stranded RNAs are then bound by a multiprotein complex, RISC (RNA-induced silencing complex). The RISC complex is able to find homologous sequences among RNA or mRNAs and possesses helicase and nuclease activity. RISC uses the dsRNA as a guide for the identification of complementarity between dsRNA and target RNA. The identified target RNA is a substrate for endonucleolytic cleavage leading to its further degradation by exo- and endonucleases (see Fig. 26.2.1). This pathway may represent a cellular

defense mechanism against viral infection or transposon replication. The discovery of microRNAs (miRNA) in *C. elegans* suggests that RNA interference is also involved in the regulation of developmental genes (Schwarz and Zamore, 2002). For more information about RNA interference, as well as dicer and RISC, refer to UNIT 26.1.

In contrast to long double-stranded RNAs, which cause a general arrest of translation by activating protein kinase R and RNase L, siRNAs induce neither of these activities due to their small size (Caplen et al., 2001). Thus, the application of siRNAs is suitable for sequence-specific inhibition of single genes.

Critical Parameters

siRNA design

siRNAs are double-stranded RNAs of 21 to 25 nucleotides with overhangs of 2 to 3 nucleotides at both 3' ends (Elbashir et al., 2001). The sense strand is identical to the target mRNA, while the antisense strand is complementary to both the sense strand and the mRNA. Figure 26.2.2 shows several siRNA molecules that trigger strong effects in mammalian RNAi studies. Although the overall structural features of potent siRNA molecules have been identified, certain details have to be considered in order to optimize their inhibitory effect. The important features to consider, as discussed below, are (1) selecting the best target site, (2) length of the double-stranded region, (3) overhang on either 3' end, and (4) tolerance for chemical modifications.

Target-site selection

In contrast to the inhibitory effects of antisense oligonucleotides or ribozymes, the RNA interference process is not merely a matter of base pairing and annealing between

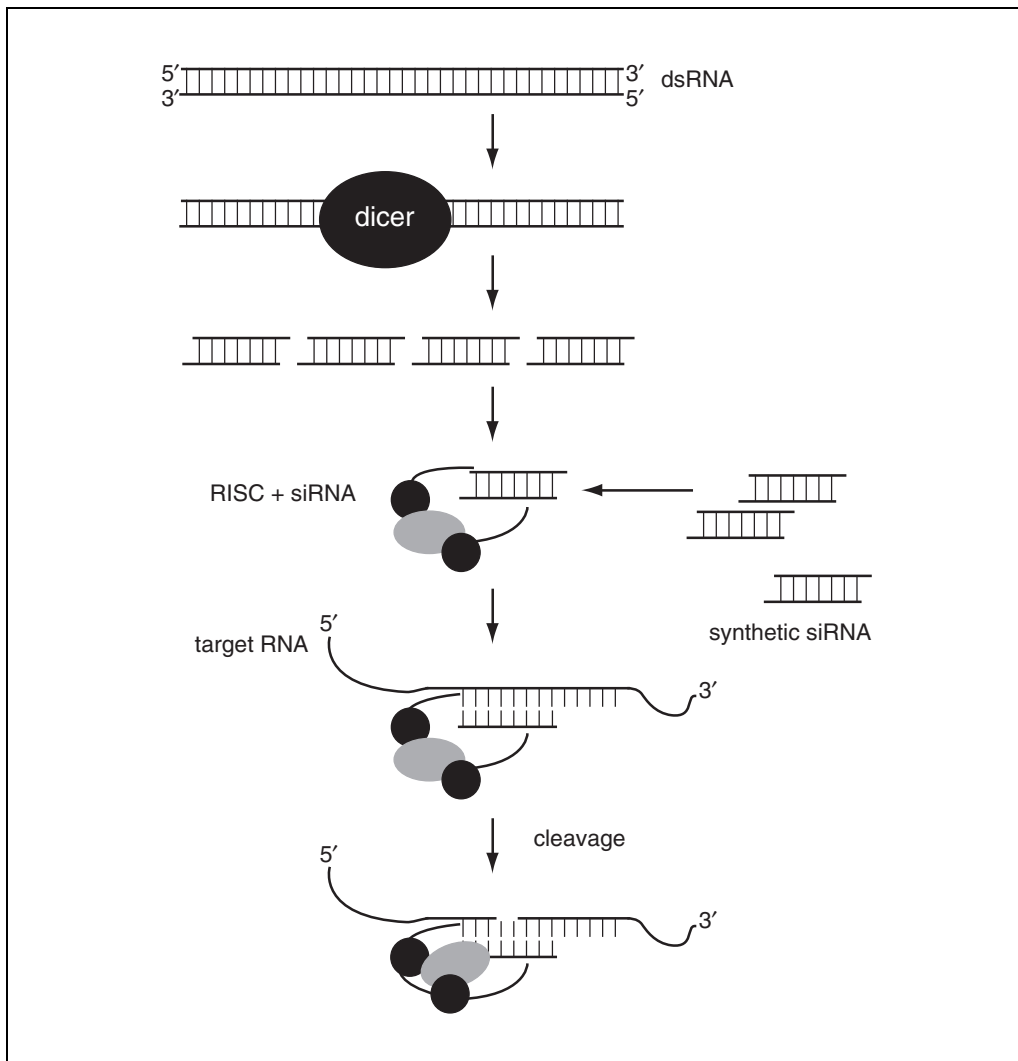


Figure 26.2.1 A model for RNA interference in mammalian cells. Long double-stranded RNAs are processed by dicer. RISC induces a target-site-specific cleavage of RNA using siRNA and dicer products as a guide. See text for more details.

target sites and the complementary antisense molecule. The basic mechanisms of RNA interference are not yet completely understood. For this reason, it is not possible to predict the sequence preference of the multienzyme complexes that are involved in the gene silencing process of mammalian cells. In spite of the limited available information, recommendations for a particular sequence motif for siRNA design are sometimes found in the literature or in technical notes of commercial siRNA suppliers.

Target sites on mRNA, noncoding RNA (e.g., rRNA), or viral RNA are also subject to RNA interference. The cognate sequence of an siRNA is either located on the untranslated 5' region of an mRNA, the open reading frame, or the 3' end of the message; however, the efficiency of individual target sites on a given RNA may vary considerably without obvious

reasons (Holen et al., 2002). A potential for stable secondary structures decreases the value of a given sequence as an siRNA target site, but is not necessarily incompatible with RNA interference. Sequences within introns should not be used as siRNA target sites as the gene silencing machinery works only in the cell cytoplasm and not in the nucleus (Zeng and Cullen, 2002).

Since siRNAs are rather short molecules, the anticipated target site may be present in more than one mammalian mRNA. At least two or three mismatches between the prospective siRNA sequence and an mRNA are necessary to completely prevent siRNA-mediated RNA cleavage. A BLAST search (*UNIT 19.3*) of the siRNA sequence against a mouse or human gene bank is highly recommended. Both mRNA and genomic sequences should be checked.

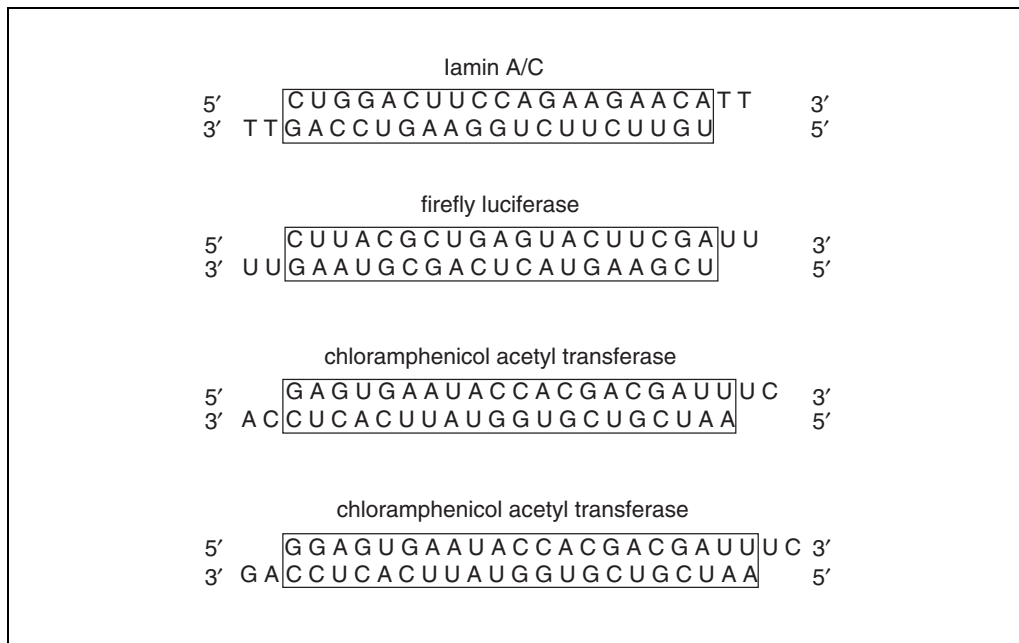


Figure 26.2.2 Examples of synthetic siRNA and their target gene. Efficient siRNA have double-stranded regions of 19 to 21 nucleotides and 2-nucleotide overhangs at the 3' end on both strands. A substitution of ribouridine bases by deoxythymidine at the 3' prime ends may stabilize the siRNA in solution, but is not necessary for efficient gene silencing in mammalian cells. The two targets of the siRNA shown at the top of the figure are lamin A/C (Elbashir et al., 2001) and firefly luciferase (Hohjoh, 2002), respectively. The two siRNAs shown at the bottom of the figure both have chloramphenicol acetyl transferase (Caplen et al., 2001) as their target.

Length of double-stranded region and overhang

siRNAs may be designed as symmetrical molecules with the sense and antisense strands having the same length. The double-stranded region may vary between 19 and 21 nucleotides. An overhang of two nucleotides in length at the 3' end of either strand is beneficial (Caplen et al., 2001; Elbashir et al., 2001; Hohjoh, 2002). Unless data about the best G+C content of the siRNA are generated, it is prudent to avoid extreme values. Therefore, the G+C content of the double-stranded region should be between 40% to 55%. Figure 26.2.2 shows some examples of siRNA that were applied in mammalian RNA interference experiments that led to strong gene silencing effects.

Tolerance for modifications of the RNA strands

Natural siRNAs are products of nuclease activities that process long double-stranded RNA. As a result, ribonucleotides are the only constituents of endogenous dicer products. However, the RNA interference machinery tolerates certain chemical modifications in synthetic siRNAs—e.g., substituting the two-nucleotide 3'-end overhang ribouridines with deoxythymidines does not interfere with siRNA

activity and probably protects the RNA against nuclease attack. The coupling of fluorophores allows the detection of siRNA by fluorescence microscopy or fluorescence-activated cell sorting. Customized siRNA with fluorophores are commercially available (e.g., Ambion, Dharmacon Research, Eurogentec). Care must be taken not to alter too many of the ribonucleotides. In particular, the core of the siRNA double-stranded region has to maintain its ribonucleotide character.

The sense strand is more tolerant of modifications without losing RNAi efficiency than the antisense strand. A covalent modification of the 3' hydroxy group at either end of the double strand or the 5' end of the sense strand is acceptable; however, blocking of the antisense 5' end destroys siRNA activity. This position has to be either a 5' hydroxyl or a phosphate group (Chiu and Rana, 2002; Martinez et al., 2002).

How to get siRNA

The World Wide Web offers many tools for selecting manufacturers of siRNA molecules or accessories (see Internet Resources). siRNA can be obtained by chemical synthesis or in vitro transcription. In addition, siRNA can be generated by intracellular expression.

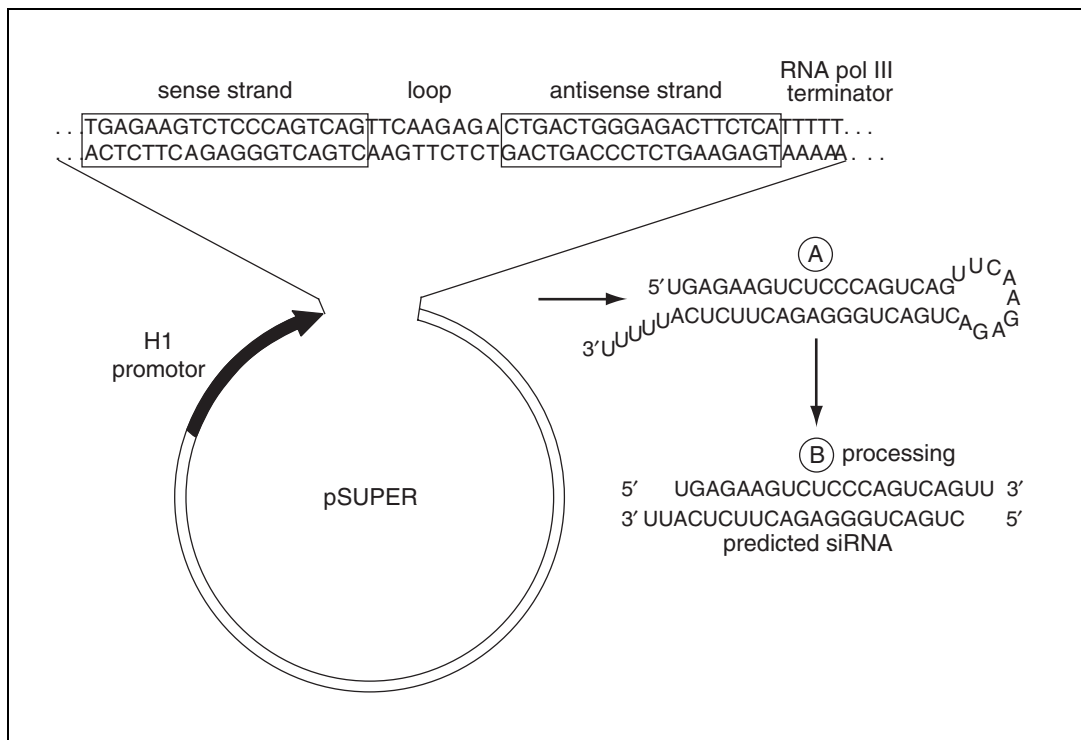


Figure 26.2.3 Model for plasmid-derived synthesis of siRNA in vivo. An RNA polymerase III promoter drives the expression of short hairpin-like RNA with a proposed 19-nucleotide duplex. The complementary sense and antisense sequence of the final siRNA are linked by a loop of 8 to 10 nucleotides (A). This hairpin RNA is probably subject to intracellular processing. The remaining double-stranded molecules (B) possess sequence homology with the target gene and perform like functional siRNA (Brummelkamp et al., 2002).

Many companies with expertise in DNA oligonucleotide production now offer customized solid-phase synthesis of RNA. RNA synthesis is more difficult and less efficient than DNA synthesis; therefore, additional purification steps of the RNA strands are recommended to obtain reproducible results. RNA purification by gel electrophoresis or HPLC and the annealing of both strands are offered by all manufacturers. Dharmacon Research synthesizes RNA with a proprietary technology that results in RNA strands of considerably higher purity and better yield than most other sources. In this case, additional purification steps are usually not necessary.

Whereas experiments with antisense DNA oligonucleotides may need up to 2 μ g DNA for transfecting cells in a single well, 50 to 100 ng siRNA are sufficient to inhibit gene expression with the same or even better efficiency. Thus, a synthesis with yields ranging from 50 to 1000 nmol provides enough material for large sets of experiments. The possibility to receive ready-to-use molecules and the option for introducing chemical modifications make synthetic siRNA an attractive choice.

Ambion offers an in vitro transcription kit for siRNA synthesis. The customer has to design some DNA oligonucleotides for cloning into an RNA expression cassette. The transcription of this unit leads to RNA strands that are hybridized and further trimmed by nucleases. This is a rather inexpensive way of getting functional siRNA and allows the probing of many different target sites at moderate costs.

A totally different approach is the intracellular synthesis of siRNA (see Fig. 26.2.3). In this case, a plasmid contains an expression cassette for the synthesis of a short-hairpin RNA molecule. An RNA polymerase III promoter, either U6 or H1, drives the expression of the RNA. This single-stranded RNA contains two stretches of complementary sequences that refold into a double-stranded region and a connecting loop, which is probably subjected to further processing (Brummelkamp et al., 2002; Paddison et al., 2002; Sui et al., 2002). Due to the constitutive expression of the siRNA, a sustained inhibition of target gene expression is achievable. The silencing plasmid can also be used for the generation of stable cell lines. Gene-silencing plasmids are commer-

cially available (e.g., Ambion or Imgenex) and come with detailed protocols for the proper setup of a hairpin-like RNA. As an alternative to hairpin RNA expression cassettes, sense and antisense strands can be expressed separately and will hybridize inside the cell to form siRNA (Lee et al., 2002).

Gene silencing by *in vivo* expression of siRNA circumvents the need for highly efficient siRNA transfection protocols. However, transient expression demands a high yield of transfected cells. The generation of permanent cell lines is a very elegant way of inhibiting target gene expression, but siRNA efficiency in gene silencing has to be verified before starting cell line selection. In addition, siRNA can be generated by intracellular expression using vectors like pSUPER (Fig. 26.2.3).

Preparation of siRNA samples

All commercial manufacturers offer synthesized siRNA in a highly purified form at a defined concentration with annealed strands. Some researchers might prefer to order individual strands or synthesize the RNA in-house. Gel purification under denaturing conditions is similar for short RNA and DNA (see *UNIT 2.12* for useful instructions for preparative denaturing polyacrylamide gel electrophoresis). After RNA purification, it is necessary to determine the concentration as accurately as possible by ultraviolet spectroscopy at 260 nm (*APPENDIX 3D*). The molar extinction coefficient of each RNA sequence has to be calculated individually, due to the limited size of 21 to 23 nucleotides. An easy-to-use program for this purpose is the biopolymer calculator. This program for computing molecular properties of RNA strands is accessible at <http://paris.chem.yale.edu/extinct.html>.

Cell transfection

There is no “gold standard” for transfecting siRNA. The best method for transferring synthetic siRNA into a given mammalian cell line has to be found empirically. High transfection efficiencies are of crucial importance for gene silencing experiments.

MacManus and Sharp (2002) summarized some of the published transfection methods in mammalian gene silencing experiments. As long as popular laboratory cell lines such as HeLa or NIH 3T3 are used, the development of siRNA transfection procedures is simple; however, this is not true for other cell lines (e.g., hepatic star cells, neuronal cells). Here, determining optimal conditions can be a tedious process.

Some companies offer siRNA and transfection kits as a package. Using a modified siRNA carrying a fluorescent marker allows the determination of transfection efficiency by fluorescence microscopy or fluorescent-activated cell sorting. Most protocols involve the use of transfection kits with either polycationic or lipid compounds which build vesicles upon contact with the siRNA. Some manufacturers offer reagents optimized for siRNA (e.g., Mirus, Invitrogen, Gene Therapy Systems); however, some cell types are more efficiently transfected by standard DNA transfection kits. Methods based on calcium phosphate complexes are not the best choice for siRNA, but the successful use of adapted protocols have also been reported (Paddison et al., 2002).

A very convenient way to render mammalian cells permeable for siRNA is electroporation. For optimum results, electric settings and siRNA concentrations have to be tested. Generally, higher siRNA concentrations are needed for electroporation as compared to lipofection in order to achieve good transfection and inhibition efficiencies.

The transfection of siRNA expression plasmids is not different from standard DNA transfection. Chapter 9 presents several units containing protocols for the successful introduction of DNA into mammalian cells and ways to select stable cell lines that contain the siRNA plasmid integrated into the genome. Methods based on calcium phosphate complexes (e.g., *UNIT 9.1*), however, are rarely used for transfecting siRNA. This is because the calcium phosphate transfection protocol often demands higher amounts of siRNA than liposome-based protocols. In addition, the authors have observed that the skills and experience of the experimenter influence the outcome of a calcium phosphate transfection protocol to a considerable extent. Nevertheless, a careful optimization of transfection conditions and siRNA concentration may result in good silencing efficiencies in certain cell lines (Paddison et al., 2002).

Troubleshooting

A guide to some of the common problems in RNAi experiments, along with their possible causes and solutions, is presented in Table 26.2.1.

Anticipated Results

Support Protocol 2 was developed for the testing of a fusion reporter gene between a short hepatitis C virus (HCV) sequence (Kolykhalov et al., 1996) and firefly luciferase. This recom-

Table 26.2.1 Troubleshooting RNA Silencing

Problem	Possible cause	Solution
No gene silencing	siRNA was not properly annealed	Repeat annealing and check on nondenaturing gel (<i>UNIT 2.7</i>)
	siRNA was degraded	Discard and use a new siRNA batch
	Low transfection efficiency	Try other transfection methods and use fluorescent siRNA for verifying transfection efficiency
		Vary siRNA concentration in the experiment
	siRNA was targeting an intron sequence	Choose a new sequence outside the intron region
	Target site was inaccessible	Try a different sequence with new siRNA
	Target site and siRNA sequence differed	Check whether the sequence information about the target site is true for the particular cell line
	Wrong time window	Analyze target gene at different time intervals
	Forbidden modification	Determine if chemical modification inactivated the siRNA
Gene silencing observed with negative control siRNA	siRNA concentrations were too high	Try lower siRNA concentrations
	siRNA affected more than one gene	Repeat BLAST search
	siRNA were toxic for the cells	Try other transfection reagents and reduce siRNA concentration

binant gene was generated by PCR and cloned into a standard mammalian expression vector. The vector was cotransfected with a β -galactosidase reference plasmid and one of three types of siRNA into the liver cell line HuH-7 (#JCRB0403, Health Science Research Resources Bank, <http://www.jhsf.or.jp>). The β -galactosidase plasmid permits correction for transfection efficiency and cell lysis differences. HCV1+2 siRNA targeted the HCV sequence, LUC1+2 was directed against a sequence within the firefly luciferase open reading frame, and C1+2 served as negative control and was not complementary either to luciferase or to β -galactosidase. Figures 26.2.4 and 26.2.5 show the luciferase reporter gene, the siRNAs used in this experiment, and the results obtained. About 48 hr after the transfection, the cells were lysed and the luciferase and β -galactosidase activities were quantitated with chemoluminescence enzyme assays (see *UNIT 9.7B*). The relative ratio between firefly luciferase and β -galactosidase was then calculated. The value for the expression of reporter plasmids in the absence of siRNA was set to 1. In the presence of specific siRNAs HCV1+2 and

LUC1+2, the expression level of the firefly gene was reduced by a factor of 5 compared to the no-siRNA control. The nonspecific control siRNA C1+2 had no inhibitory effect on the luciferase gene. Instead, a slight increase of reporter gene activity was observed. (This phenomenon is frequently observed, but at this time, a reasonable explanation for this behavior does not exist.) However, higher concentrations of negative control siRNA may lead to a nonspecific inhibitory effect. Therefore, it is necessary to determine an optimum siRNA concentration with strong inhibitory effects of the specific siRNA and only moderate or no effects for the negative siRNA control.

The siRNA-mediated inhibition of the HCV-luciferase fusion gene led to an 80% reduction. Repeating the same experiment in HeLa-S3 cells led to 95% reduction.

The HCV1+2 siRNA was as effective as LUC1+2, although the HCV target sequence possessed three particular features that are not in favor of efficient inhibition. First, the HCV sequence has a predicted tendency to fold into a stem-loop structure. Second, the HCV sequence lies upstream of the open reading frame

Gene Silencing

26.2.11

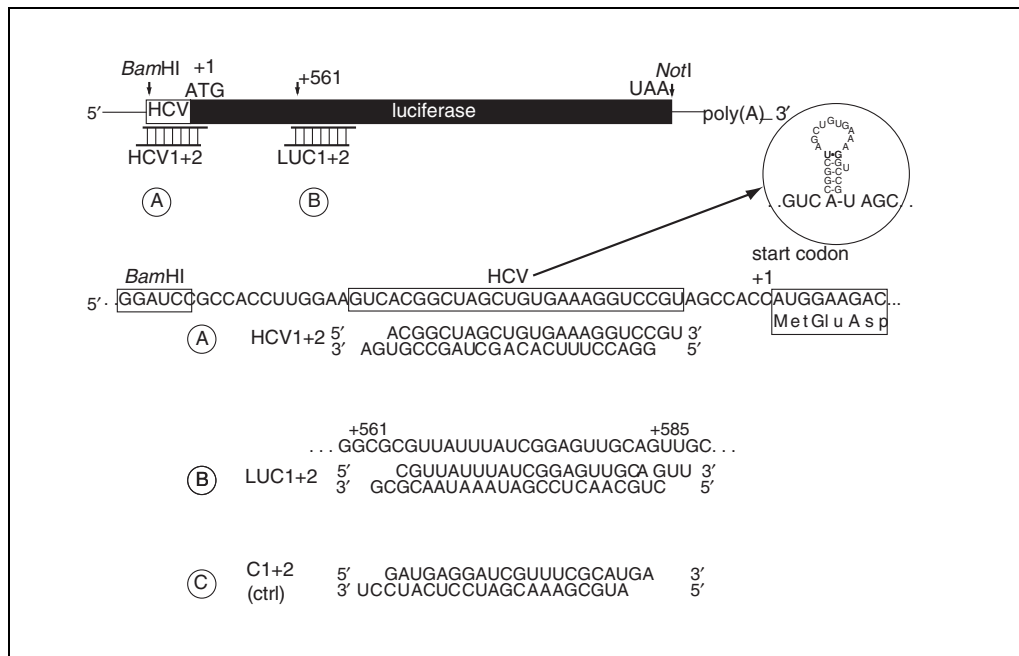


Figure 26.2.4 A recombinant fusion gene serves as a reporter system for siRNA efficiency in mammalian cells. A 26-nucleotide hepatitis C virus sequence is placed upstream of the firefly luciferase open reading frame. The viral sequence possesses the ability to form a stem-loop structure. Two siRNAs were designed to target the mRNA. HCV1+2 (A) is directed against the hepatitis C region, whereas LUC1+2 (B) is targeting the open reading frame of the fusion gene. The siRNA C1+2 (C) is not related to the reporter mRNA and serves as a negative control.

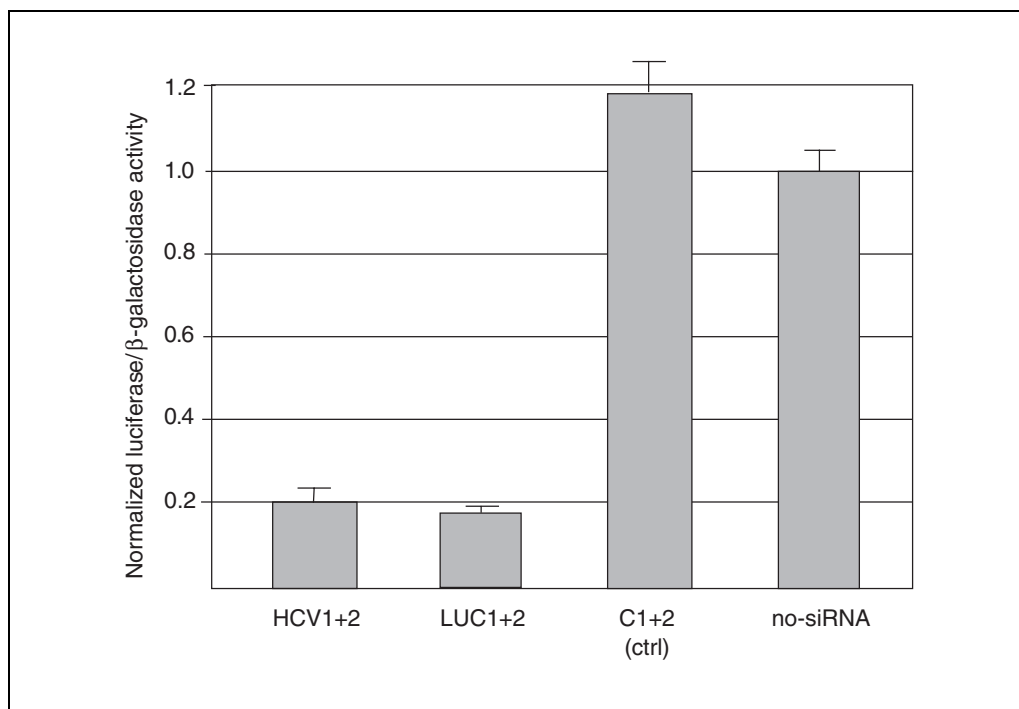


Figure 26.2.5 In liver cells, the reporter gene is expressed at a 5-fold lower level in presence of HCV1+2 and LUC1+2 compared to control siRNA or absence of siRNA. Note the slight stimulation of reporter gene activity by C1+2. Refer to Figure 26.2.4 for siRNA sequences.

start codon within the 5' untranslated region. Third, the target sequence is next to the ribosome-binding site and may compete with protein complexes for target-site binding. Nevertheless, for unknown reasons, it worked.

Time Considerations

A complete RNAi experiment can be completed within 3 days, but may last up to several weeks. Usually, the transfection of siRNA starts with seeding cells at day 1. This takes about 1 hr. On day 2, the lipofection protocol (see Basic Protocol) includes the pipetting of siRNA dilutions and the preincubation with a transfection reagent, which needs 1 hr of hands-on work. An additional 20 min to 1 hr are necessary to add the siRNA complexes to the cells. Approximately 4 hr are required for annealing the siRNA strands (see Support Protocol 1), and both annealing and electrophoresis can be performed in a single day.

For the cell electroporation protocol (see Alternate Protocol), the whole process of handling the cells, including centrifugation, electroporation, and final transfer into a culture vessel, takes <1 hr.

After transfecting synthetic siRNA, maximum effects on target gene expression are detectable after 16 hr incubation. On day 3, the harvest of transfected cells and preparation of extracts need about 2 hr for total RNA or protein.

Depending on the gene of interest, the cell line used, or the gene product, the incubation time may need to be extended to 72 hr. The authors recommend combined analyses of target RNA and protein abundance. As the half-life of proteins is generally longer than that of their corresponding mRNAs, an efficient inhibition of mRNA expression does not immediately result in a reduction of protein levels.

The expression of siRNA *in vivo* from a plasmid vector leads to a sustained inhibition of the target genes. A decrease of RNA or protein levels is still detectable after several days or weeks.

Phenotypical changes of the cells (e.g., the induction of apoptosis) may immediately follow the gene silencing effects or may need additional hours or days before becoming obvious.

Literature Cited

Brummelkamp, T.R., Bernards, R., and Agami, R. 2002. A system for stable expression of short interfering RNAs in mammalian cells. *Science* 296:550-553.

Caplen, N.J., Parrish, S., Imani, F., Fire, A., and Morgan, R.A. 2001. Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc. Natl. Acad. Sci. U.S.A.* 98:9742-9747.

Chiu, Y.-L. and Rana, T.M. 2002. RNAi in human cells: Basic structural and functional features of small interfering RNA. *Mol. Cell.* 10:549-561.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. 2001. *Nature* 411:494-498.

Hannon, G.J. 2002. RNA interference. *Nature* 418:244-251.

Heidenreich, O., Krauter, J., Riehle, H., Hadwiger, P., John, M., Heil, G., Vornlocher, HP., and Nordheim, A. 2002. AML1/MTG8 oncogene suppression by small interfering RNAs supports myeloid differentiation of t(8;21)-positive leukemic cells. *Blood*: In press. [epub ahead of print]

Hohjoh, H. 2002. RNA interference (RNAi) induction with various types of synthetic oligonucleotide duplexes in cultured human cells. *FEBS Lett.* 521:195-199.

Holen, T., Amarzguioui, M., Wiiger, M.T., Babaie, E., and Prydz, H. 2002. Positional effects of short interfering RNAs targeting the human coagulation trigger tissue factor. *Nucl. Acids. Res.* 30:1757-1766.

Kolykhalov, A. A., Feinstone, S.M., and Rice, C.M. 1996. Identification of a highly conserved sequence element at the 3' terminus of hepatitis C virus genome RNA. *J. Virol.* 70:3363-3371.

Lee, N.S., Dohjima, T., Bauer, G., Li, H., Li, M.-J., Ehsani, A., Salvaterra, P., and Rossi, J. 2002. Expression of small interfering RNAs targeted against HIV-1 rev transcript in human cells. *Nat. Biotechnol.* 20:500-505.

MacManus, M.T. and Sharp, P.A. 2002. Gene silencing in mammals by small interfering RNAs. *Nature Genet.* 3:737-747.

Martinez, J., Patkaniowska, A., Urlaub, H., Lührmann, R., and Tuschl, T. 2002. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* 110:563-574.

Novina, C.D., Murray, M.F., Dykxhoorn, D.M., Beresford, P.J., Riess, J., Lee, S.-K., Collman, R.G., Lieberman, J., Shankar, P., and Sharp, P.A. 2002. SiRNA-directed inhibition of HIV infection. *Nature Med.* 8:681-686.

Paddison, P.J., Caudy, A.A., Bernstein, E., Hannon, G.J., and Conklin, D.S. 2002. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.* 16:948-958.

Schwarz, D.S. and Zamore, P.D. 2002. Why do miRNAs live in the miRNP? *Genes Dev.* 16:1025-1031.

Sharp, P.A. 2001. RNA interference—2001. *Genes Dev.* 15:485-490.

Sui, G., Soohoo, C., Affar, E.B., Gay, F., Shi, Y., Forrester, W.C., and Shi, Y. 2002. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 99:5515-5520.

Zamore, P.D. RNA interference: Listening to the sound of silence. 2001. *Nat. Struct. Biol.* 8:746-750.

Zeng, Y. and Cullen, B.R. 2002. RNA interference in human cells is restricted to the cytoplasm. *RNA.* 8:855-860.

Key References

Elbashir, S.M., Harborth, J., Weber, K., and Tuschl, T. 2002. Analysis of gene function in somatic cells using small interfering RNAs. *Methods* 26:199-213.

The pioneers in mammalian RNA interference present a concise description of planning and performing siRNA experiments in this reference.

Internet Resources

<http://www.mpibpc.gwdg.de/abteilungen/100/105/sirna.html>

This site provides suggestions and recommendations for experiments with siRNA.

<http://www.ambion.com>

Web site with plenty of background information about RNA interference and regular literature updates.

Contributed by Matthias John, Anke Geick,
Philipp Hadwiger, and Hans-Peter
Vornlocher
Ribopharma AG
Kulmbach, Germany

Olaf Heidenreich
University of Tübingen
Tübingen, Germany

The nematode worm *C. elegans* is a useful genetic model system for a variety of reasons (see Wood, 1988; Hope, 1999). Embryonic development, from fertilization to hatching, occurs in just 12 hours at room temperature, resulting in a first-stage larva with just over 550 cells. Newly hatched larvae progress through four larval molts to reach adulthood in ~3 days, and adult hermaphrodites produce ~300 self-progeny over the next three days. Both larval and adult worms have remarkably simple anatomies which consists of an inner tube comprised of the pharynx and intestine, surrounded by a hydrostatic, fluid-filled body cavity and by an outer tube that consists of the musculature, hypodermis, and ultimately a collagenous cuticle. Remarkably, the diverse tissues that comprise the adult are specified by a total of only ~1000 somatic cells. Due to the highly stereotyped cell division patterns, the origins and fates of each cell in the embryo and adult is known. Consequently, perturbations of the developmental program can be traced to changes in the fates of individual cells.

The ability to generate mutants in *C. elegans* has made this species a powerful model genetic organism. Despite being the first metazoan with a completely sequenced genome (The *C. elegans* Sequencing Consortium, 1998), conventional methods of gene knock-out/replacement by homologous recombination remain elusive. Current methods for generating deletion mutants, while valuable, are laborious and time consuming (Jansen et al., 1997; Liu et al., 1999). Since its discovery, RNA interference (RNAi) has become an important tool for *C. elegans* research and the scientific community in general. RNAi has provided a potent, specific, rapid, and simple method for determining the null phenotype of genes in *C. elegans*, and has been the basis of several genome-wide screens (Fraser et al., 2000; Gonczy et al., 2000; Piano et al., 2000; Maeda et al., 2001; Ashrafi et al., 2003; Kamath et al., 2003). RNAi is an excellent way to examine genetic interactions. For example, the dsRNA of interest can be delivered to a battery of candidate mutants that may represent sensitized genetic backgrounds. Thus, if RNAi into wild-type animals fails to give a phenotype, one may nevertheless detect suppression or enhancement of phenotypes in such strains.

In *C. elegans*, RNAi can be induced by delivering dsRNA to animals in any of four ways: microinjection (see Basic Protocol 1 and Alternate Protocol), feeding (see Basic Protocol 2), soaking (see Basic Protocol 3), or promoter-driven expression (Tabara et al., 1999; Tavernarakis et al., 2000). Most needs can be met by at least one delivery method. Microinjection of dsRNA, while technically challenging, is nevertheless the method of choice for RNAi experiments. Indeed, while other less difficult methods, such as feeding or soaking animals in dsRNA, should be used in parallel, injection should be attempted before drawing any conclusions about the function (or lack of function) of a gene of interest. It is also possible to simultaneously inhibit two or more genes at one time by coinjecting dsRNAs within a mixture. Therefore, microinjection of dsRNA is an excellent way to assay members of a gene family for overlapping or redundant functions.

Animals exposed to dsRNA by feeding or soaking are more amenable to biochemical analyses, since these procedures may easily be scaled up. The feeding and soaking protocols can also be performed in microtiter format allowing for large numbers of RNAi experiments to be performed simultaneously (Fraser et al., 2000; Maeda et al., 2001). In promoter-driven RNAi, dsRNA is produced in vivo from a transgene (Tabara et al., 1999). This is perhaps the most technically demanding and time-consuming procedure, since it requires the generation of a transgenic line. In principal, however, the transgene can be

integrated, resulting in a stable line in which the gene of interest may be silenced in a constitutive or regulated fashion, depending on the promoter used to drive dsRNA synthesis.

In the microinjection and soaking protocols, *in vitro* preparations of dsRNA are delivered mechanically by microinjection or passively by soaking. In the feeding protocol, dsRNA is expressed in *E. coli* and ingested by animals. The protocols provided in this unit are meant to guide the complete novice at worm culture in the selection and application of these RNAi delivery methods. Towards that end, protocols are also provided for transferring and propagating (see Support Protocol 1) as well as sexing and mating (see Support Protocol 2) *C. elegans*. References for more in depth discussion of these topics as well as *C. elegans* in general are provided at the end of the unit (see Key References).

NOTE: All solutions and equipment coming into contact with organisms must be sterile, and proper sterile technique should be used accordingly.

BASIC PROTOCOL 1

MICROINJECTING WORMS WITH dsRNA

To perform RNAi by microinjection, one will need to have access to a microscope equipped with a micromanipulator for holding and maneuvering the microinjection needle. Injections need not be performed at high magnification, and differential interference contrast microscopy (Nomarski optics) is not necessary. Therefore, even if expertise in *C. elegans* microinjection is unavailable, it should be possible to master dsRNA microinjection with only a few hours of practice. Microscopes suitable for this purpose will be found in a range of laboratories.

In this procedure, healthy L4 or adult hermaphrodites are immobilized under halocarbon oil on a dry agarose pad and microinjected with dsRNA, causing the body of the animal to noticeably inflate with the dsRNA solution. Injected animals are removed from the pad using a mouth pipet and a small amount of sterile M9 buffer, and then placed onto a *C. elegans* culture dish. After a period of recovery, animals that have moved (i.e., are alive) are transferred to individual plates and allowed to lay eggs for several days. The injected animals and their progeny are examined/scored for phenotypes resulting from gene silencing. In a successful RNAi by microinjection experiment, one should aim to recover ten viable injected hermaphrodites to obtain a high degree of confidence in the results.

In addition to the standard equipment for manipulating worms (e.g., stereomicroscope, platinum wire pick, media; Hope, 1999), this procedure requires microinjection equipment, including an inverted microscope equipped with a 10× objective, micromanipulator, N₂ gas, and microinjection assembly to regulate the pressure and flow of the microinjection solution. The procedure described here utilizes a Nikon inverted microscope equipped with 10× objectives, gliding stage, and Narishige MN151 micromanipulator with fine z-axis adjustments, as well as a Tritech microinjector system for pressurizing the microinjection needle. Injections are normally carried out using the 10× objective for a total magnification of 100×. As noted above, it should be possible, with some practice, to perform microinjections with a range of different microinjection systems.

IMPORTANT NOTE: The experimenter should become familiar with the microinjection setup before proceeding with the experiment.

Materials

- 0.1 to 1 $\mu\text{g}/\mu\text{l}$ dsRNA
 - Halocarbon oil Series 700 (Halocarbon Products)
 - L4 to young-adult staged *C. elegans* (e.g., strain N2; CGC)
 - M9 buffer (see recipe)
 - NGM plates seeded with *E. coli* (see recipe)
 - Loading capillary (see recipe)
 - Microinjection needle (see recipe)
 - Inverted microscope with micromanipulator and microinjection assembly with N_2 gas supply (e.g., TRITECH Research Microinjector System)
 - Agarose microinjection pads (see recipe)
 - 30-mm Petri dish lid
 - Dissecting microscope
 - Platinum wire pick flattened at the end
- Additional reagents and materials for transferring and propagating *C. elegans* (see Support Protocol 1)

Load and mount the microinjection needle

1. Pellet insoluble material that may clog the microinjection needle by centrifuging the 0.1 to 1 $\mu\text{g}/\mu\text{l}$ dsRNA solution 5 min at $10,000 \times g$, 20°C .
2. Touch the drawn-out end of the Pasteur pipet portion of the loading capillary to the dsRNA solution, allowing a small volume to be drawn into the pipet. Deposit the RNA solution just behind the drawn-out tip of the needle by threading the loading pipet all the way into the needle and gently blowing through the mouth pipet (Fig. 26.3.1).

Needles can be loaded and stored up to several hours inside a humidifying chamber constructed from a Petri dish containing a wet piece of filter paper. It is often prudent to load two needles for each microinjection solution and store one in the humidifying chamber as a backup in case of problems.

3. Attach the loaded microinjection needle to the microinjection assembly, sliding the tubing over the back end of the needle (Fig. 26.3.2).

The tubing should fit snugly over the needle making a tight seal. If the seal is not tight enough, the N_2 gas will escape around the needle when pressure is applied and the dsRNA solution will not flow from the needle.

4. Turn on the light source of an inverted microscope. Mount the loaded needle onto the micromanipulator and roughly align the tip of the needle into the center of the light beam.

The needle should be positioned such that it has a downward angle of $\sim 15^\circ$ to 20° .

Prepare worms for microinjection

5. Working at a dissecting microscope, prepare a microinjection pad by placing a small ($\sim 10\text{-}\mu\text{l}$) drop of halocarbon oil Series 700 on the glass surface of the coverslip adjacent to the region coated with dry agarose. Place the microinjection pad onto an inverted 30-mm Petri dish lid resting on the stage of a dissecting microscope.

This will hold the microinjection pad at approximately the same height as the worms on a small plate, so that minimal focusing will be required while transferring worms to the microinjection pad.

6. Briefly flame a platinum wire pick that has been flattened at the end for picking up worms.

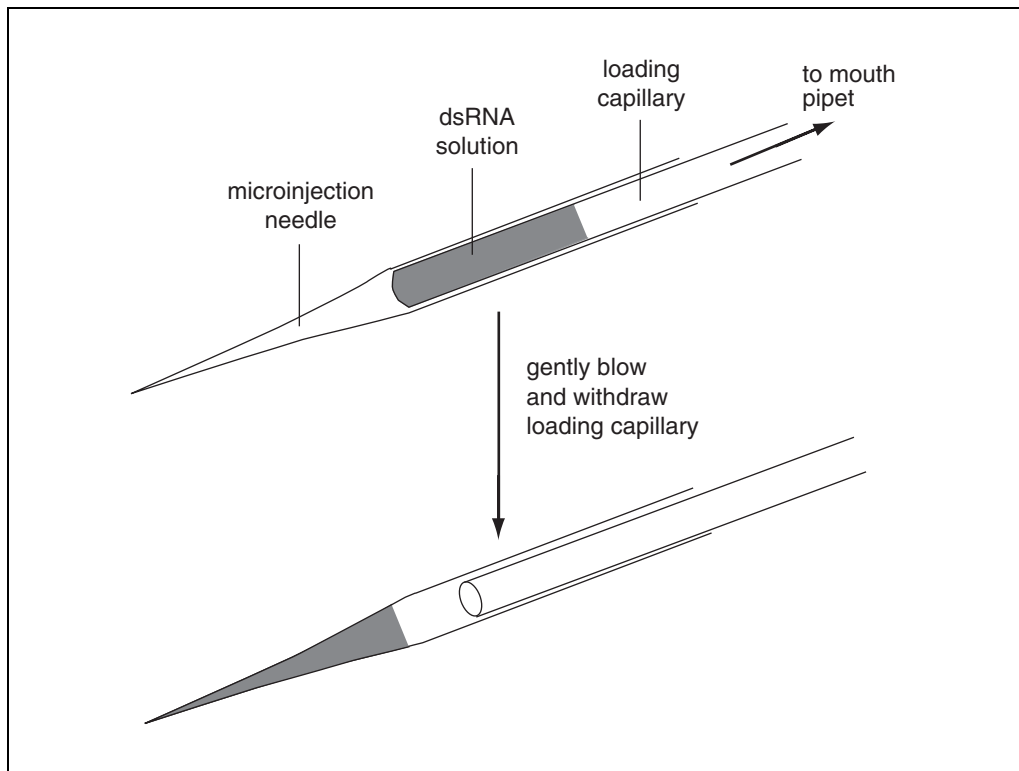


Figure 26.3.1 Loading the microinjection needle.

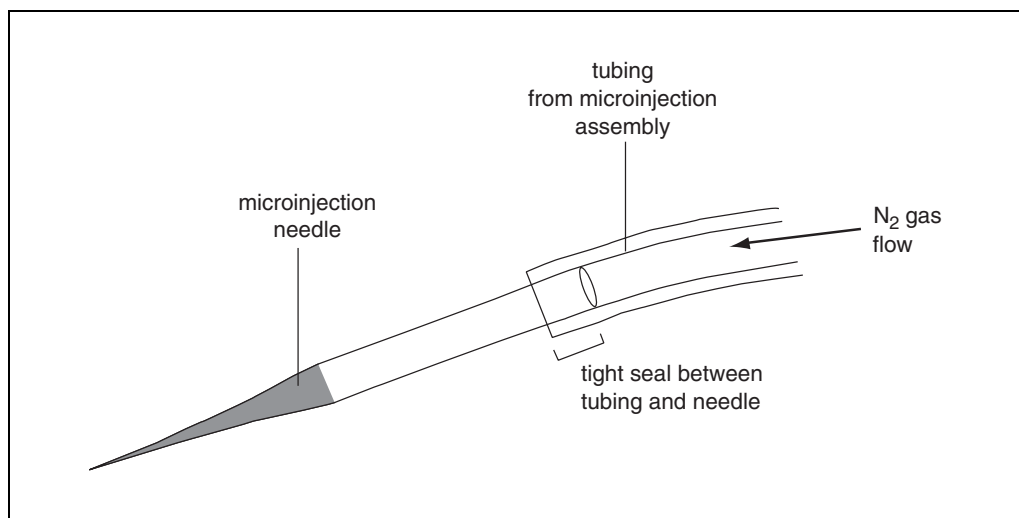


Figure 26.3.2 Attaching the microinjection needle to the microinjection assembly via tubing.

7. Touch the pick to the surface of the halocarbon oil on the pad. Using this viscous oil drop on the end of the wire, pick up several L4- to young adult-staged *C. elegans* for microinjection. Transfer the worms to a dry portion of the pad by gently touching the underside of the pick to the agarose surface of the pad and sliding the pick along to gently push the worm down to the surface of the pad, making a mental note of the configuration of the animals.

Pick up the worms by gently touching the oil-coated undersurface of the pick to the top of the worm on the Petri dish. When the pick is raised, the worm should be transferred from the plate to the underside of the pick by surface tension. Move to another worm and repeat the procedure.

Beginners should start with two or three worms, but experts will pick up as many as twenty worms on the same pick. If possible, choose worms that are either off the E. coli lawn or are on a thin area.

It is not necessary to place the worms in a row, or even to separate the worms from one another, but it is helpful if the majority of the worms are angled with the long axis of their bodies oriented in the same general direction.

8. Go back to the oil drop with the pick and obtain additional oil to spread over the worms as necessary, making sure that all of the worms are completely covered.
9. If the worms are not sticking to the microinjection pad, gently pat the worms along the length of the body to encourage them to stick to the agarose pad.

IMPORTANT NOTE: *Once the animals have been transferred to the microinjection pad, work as quickly as possible to prevent the animals from drying out.*

Use pads that are thoroughly dry to immobilize the worms. If necessary, rebake pads 1 hr in a 50°C oven. Check to make sure that the agarose side of the pad is facing up. The worms will not stick to bare glass.

Inject the dsRNA into worms

10. Place the microinjection pad with immobilized worms onto the stage of the inverted microscope and focus on the worms.
11. Move the tip of the microinjection needle into the center of the field of view, being careful to keep the needle above the worms (Fig. 26.3.3A). Maintain this needle position and only move the needle into or out of the focal plane (i.e., toward or away from the microinjection pad) using the z-axis control knob on the micromanipulator.
12. Orient the pad so that the general body axis of the animals parallels the orientation of the needle (Fig. 26.3.3).

If the needle is perpendicular to the body axis of the animal, it may completely penetrate the animal or cause it to roll over during the microinjection procedure.

13. Using the 10× objective, choose a worm at one end of the cluster of worms and position the worm directly under the needle tip by moving the stage.

Keep in mind the configuration of worms placed on the pad (step 7). Move from one end of the cluster to the other systematically.

Some investigators prefer to move the stage by hand, while others prefer to use the stage dials. Try both and use whichever method is more comfortable.

14. Focus carefully and lower the needle directly down on top of the first animal's body. With the needle pressing against the body of the animal (creating an indentation), push on the stage to gently move the worm into the needle tip (Fig. 26.3.3B to D).
15. Once the needle penetrates the cuticle, apply pressure to the needle and allow the dsRNA solution to flow into the animal.

With an ideal microinjection, the flow of the solution from the needle tip will be easily observable in both directions (i.e. towards the mouth and anus; Fig. 26.3.3E to G) and a slight increase in size near the point of injection may even be observed. If the animal rapidly increases in both diameter and length, then too much solution is being injected, and the animal may not survive. Decrease the duration of the pulse and, if necessary, give several quick pulses to deliver the dsRNA into the animal.

16. Remove the needle from the worm by pulling back on the sliding stage (Fig. 26.3.3H). Raise the needle slightly using the z-axis knob on the micromanipulator and move to the next worm.

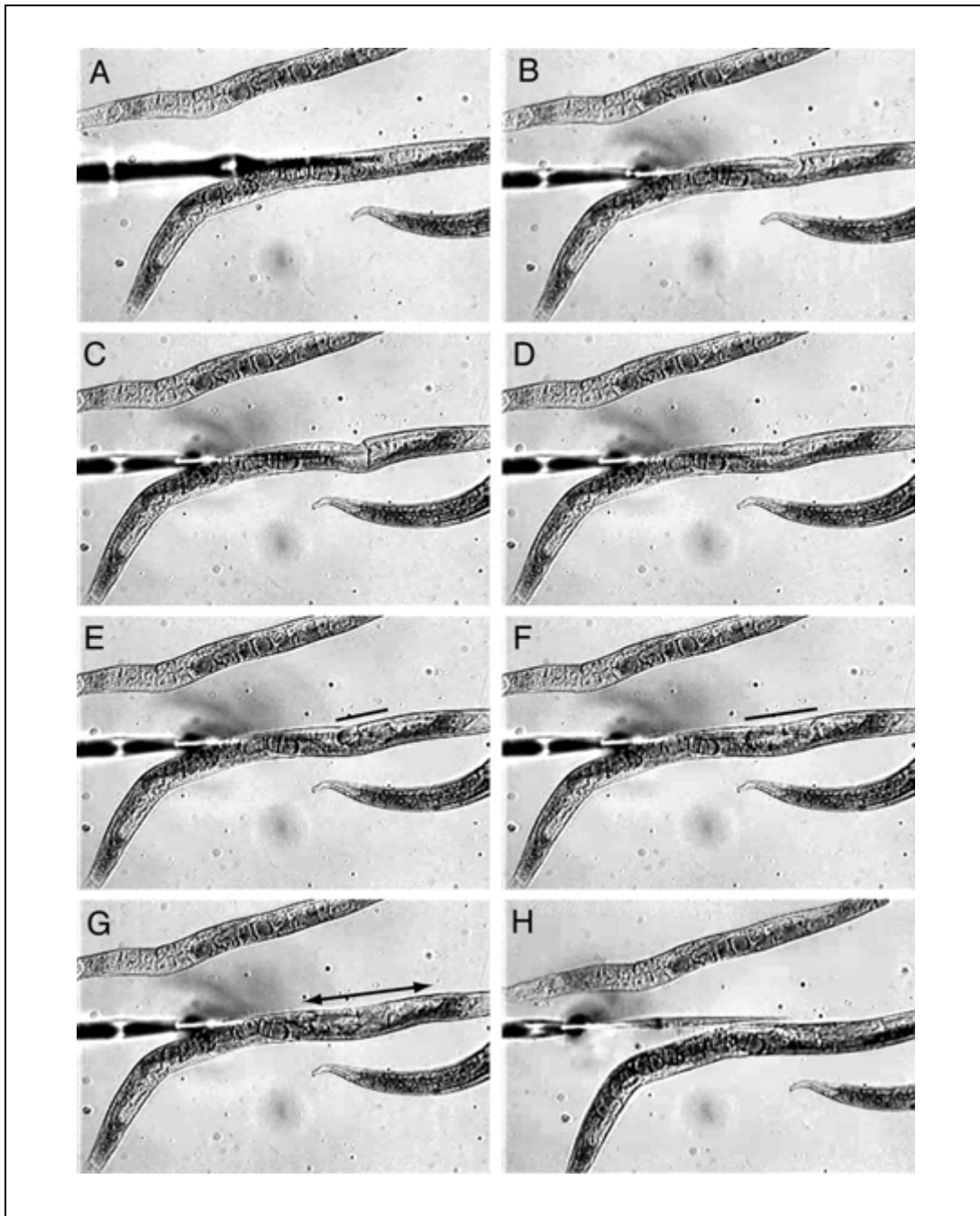


Figure 26.3.3 Overview of the microinjection technique. **(A)** The injection needle enters the field of view from the left of the image and is oriented above and roughly parallel to the body axis of the worm. **(B to D)** The needle is lowered until the tip touches the worm forming a slight indentation. The worm is pushed against the needle until the tip of the needle penetrates the cuticle. **(E to G)** The dsRNA solution is injected into the body cavity of the worm and flows in both directions away from the tip (indicated by the bars and arrows). **(H)** The needle is then backed out of the worm by moving the worm away from the needle.

- Repeat steps 13 to 16 until all of the worms are injected. Carefully remove the pad from the microinjection microscope and transfer to a dissecting microscope.

Recover injected animals for phenotype analysis

- Recover the animals by placing a small drop of M9 buffer directly under the halocarbon oil so that it contacts the worms and gently floats them off of the agarose.

The loading capillary can be used for this purpose. Simply break off the very thin drawn-out end of the loading capillary, creating an opening just slightly larger in diameter than a worm. Using the capillary, blow M9 against the worms to loosen them from the pad and suck each worm up into the capillary.

Generally, it is wise to transfer worms in small groups, approximately five animals at a time, as occasionally worms become stuck to the glass inside the capillary.

- Deposit the worms on a fresh NGM plate seeded with *E. coli*, and allow the worms to recover (i.e., incubate at 20°C) overnight (see Support Protocol 1).

Healthy worms that survive the microinjection should begin to move vigorously in the M9 buffer or shortly after transfer to the NGM plate.

- Transfer each injected animal to a fresh NGM plate (day 1) and incubate an additional 24 hr at 20°C (see Support Protocol 1). Transfer the injected animals to a fresh NGM plate once more (day 2).

- Monitor the injected animals and their progeny for phenotypes (see Hope, 1999; Wood, 1988).

Embryos that are still present on day 1 plates 24 hr after removing the adult indicates that the microinjected dsRNA produces an embryonic lethal phenotype. However, always be sure to compare the injected animals and their progeny to uninjected or control-injected animals.

*In some cases, postembryonic phenotypes can easily be determined by comparing the progeny of injected animals to the progeny of uninjected or control injected animals. These may include phenotypes such as uncoordinated (*Unc*), slow growth (*Gro*), sterility (*Ste*), larval lethal (*Lvl*), larval arrest (*Lva*), lethality (*Let*; any stage), egg-laying defective (*Egl*), high incidence of males (*Him*), dumpy (*Dpy*), small (*Sma*), long (*Lon*), blister (*Bli*), protruding vulva (*Pvl*), multivulva (*Muv*), and body morphology defects (*Bmd*).*

MICROINJECTION INTO *unc-42 rde-1* ANIMALS CROSSED TO N2 MALES TO EXAMINE ZYGOTIC RNAi PHENOTYPES

ALTERNATE PROTOCOL

If RNAi by microinjection (see Basic Protocol 1) results in embryonic arrest, this could reflect a maternal function for the targeted gene. To look specifically for zygotic activity of a targeted gene, it is possible to inject dsRNA into an RNAi-deficient (*rde* mutant) strain that has been mated to a wild-type N2 male (Fig. 26.3.4). In this scenario, maternal activity of the targeted gene will be unaffected since RNAi is not functional in the *rde* mutant hermaphrodite. However, wild-type males that successfully mate with the *rde* mutant hermaphrodite will provide RNAi activity in the cross progeny, thereby silencing zygotic expression of the targeted gene. To facilitate the identification of *rde*⁺ cross progeny, an *rde* mutant that is marked with a recessive visible mutation—e.g., uncoordinated (*Unc*)—is mated to N2 males, injected with dsRNA, and the non-*Unc* cross progeny examined for phenotypes resulting from loss of zygotic activity of the targeted gene.

Gene Silencing

26.3.7

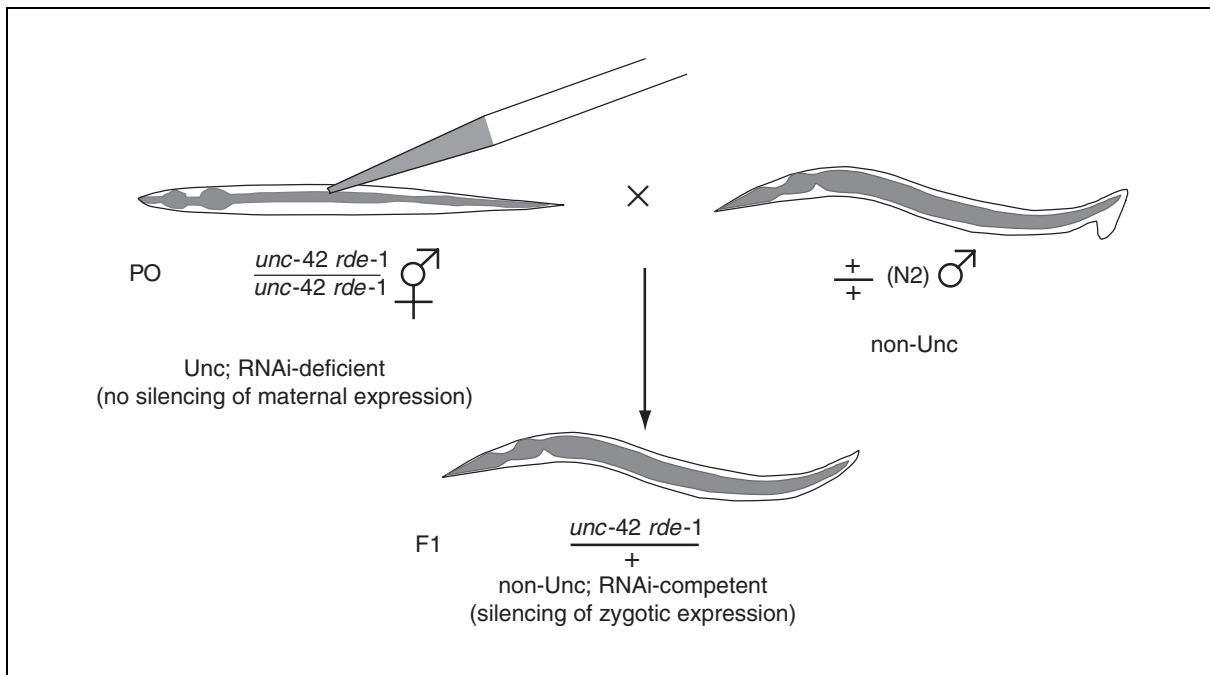


Figure 26.3.4 Schematic of the zygotic RNAi procedure.

Additional Materials (also see Basic Protocol 1)

Male N2 *C. elegans* (CGC)
unc-42 rde-1 C. elegans (WM36; CGC)
 Mating plates (see recipe)

Additional reagents and equipment for transferring and propagating (see Support Protocol 1), and mating *C. elegans*

1. Mate wild-type (N2) *C. elegans* males with an *unc-42 rde-1* mutant strain overnight (see Support Protocol 2; Hope, 1999).
2. Inject dsRNA into the mated Unc hermaphrodites (see Basic Protocol 1).
3. Transfer (see Support Protocol 1) each injected hermaphrodite to individual mating plates with males to recover (i.e., incubate at 20°C) overnight and to continue mating (see Support Protocol 2).
4. Transfer the hermaphrodite alone to a fresh NGM plate and score non-Unc cross progeny for RNAi phenotypes as described above.

**BASIC
 PROTOCOL 2**

FEEDING WORMS dsRNA

RNAi can be induced by cultivating worms on bacteria expressing dsRNA (Timmons et al., 2001). First, a cDNA corresponding to the gene of interest is cloned into a bacterial expression vector between opposing phage T7 polymerase promoter sites. The feeding vector is then transformed into the *E. coli* HT115 strain carrying the DE3 lysogen (*UNIT 1.8*), providing IPTG inducible expression of the phage T7 RNA polymerase. The HT115(DE3) strain also lacks the RNase III gene and is therefore deficient in degrading dsRNA. The RNAi food is prepared and seeded onto worm plates. Animals are placed onto the RNAi food and both the hermaphrodite and progeny are examined and scored for phenotypes.

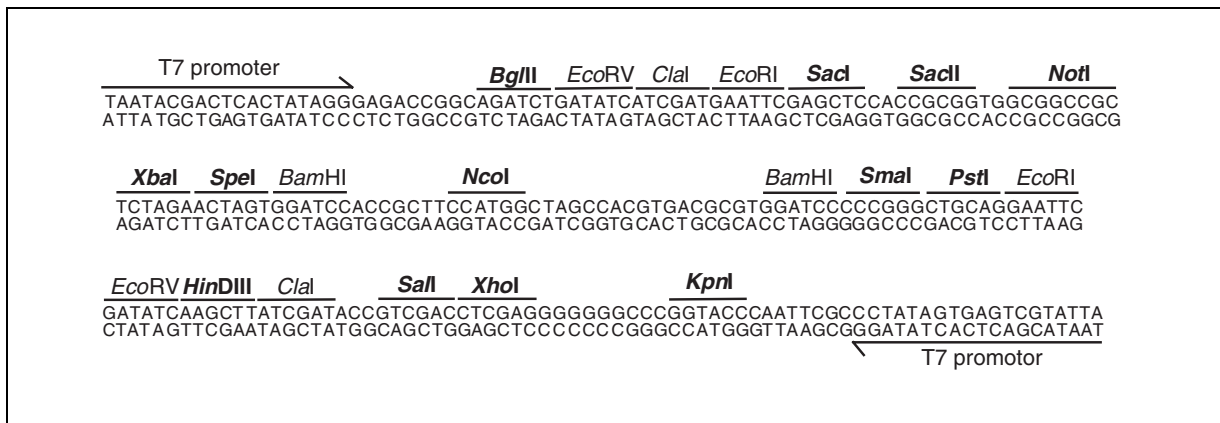


Figure 26.3.5 Multiple cloning site of the double T7 RNAi feeding vector L4440. Opposing phage T7 promoters flank the multiple cloning site of L4440. The restriction sites shown are present exclusively in the MCS with the unique sites indicated in bold.

To simultaneously target two genes by feeding, the authors have found it best to fuse two cDNA fragments, one from each gene, into a single feeding vector. Mixing of two *E. coli* strains will also work, but is more prone to error caused by unequal mixing or RNA expression in the two different feeding strains. Do not attempt to feed animals that were previously injected with dsRNA; for unknown reasons the microinjection of dsRNA (even nonspecific dsRNA) will suppress bacterial RNAi.

This protocol can be easily performed on large populations to examine strains at a biochemical level that have been depleted for a particular gene product. The procedure can also be performed in microtiter format using 12 or 24 well plates (Fraser et al., 2000).

Materials

cDNA

L4440 double T7 RNAi feeding vector (Fig. 26.3.5; A. Fire, <http://www.ciwemb.edu>)

E. coli RNAi feeding strain HT115 (CGC)

Terrific broth (TB; UNIT 1.1) containing 50 µg/ml each ampicillin and tetracycline (UNIT 1.4)

TB containing 50 µg/ml ampicillin

M9 buffer (see recipe)

M9/15% glycerol (see recipe)

NGM/amp/IPTG plates (see recipe)

Gravid/young adult, L1 larvae, or L4 larvae *C. elegans* of appropriate strain(s) (e.g., N2, CGC)

NGM/amp/IPTG plates (see recipe) seeded with HT115 containing empty L4440 RNAi feeding vector

Additional reagents and equipment for subcloning DNA fragments (UNIT 3.16), transforming *E. coli* (UNIT 1.8), growing *E. coli* (UNIT 1.2), and transferring and propagating *C. elegans* (see Support Protocol 1)

NOTE: Perform each RNAi experiment in triplicate.

Prepare the RNAi feeding strain

1. Clone a cDNA for the gene of interest into L4440 double T7 RNAi feeding vector (UNIT 3.16).

Avoid introns—the presence of introns in the feeding construct may reduce the efficiency of interference.

As alternatives to L4440, NEB offers Litmus 28i and 38i.

2. Transform the RNAi feeding construct into the *E. coli* feeding strain HT115 (UNIT 1.8).
3. Pick a colony and inoculate 2 ml terrific broth (TB) containing 50 µg/ml each ampicillin and tetracycline, and grow overnight at 37°C (UNIT 1.2).
4. Seed the starter culture into 1 liter TB containing 50 µg/ml ampicillin and incubate with shaking at 37°C for 8 to 16 hr (UNIT 1.2).

Tetracycline in this culture is unnecessary and, in fact, may reduce the efficiency of RNAi (Kamath et al., 2001).

Cultures can be scaled down; however, excess food from large cultures can be stored at –80°C, indefinitely.

Harvest the bacteria

- 5a. *For centrifugation:* Centrifuge bacteria 10 min at 800 × g, 4°C. Resuspend the bacterial pellets in M9 buffer (use 25 ml for each liter of culture) and transfer to a 50-ml conical tube.

Harvest food by centrifugation if food is needed urgently.

If time permits, allow the bacteria to precipitate passively to reduce the likelihood of contaminating the dsRNA food.

- 5b. *For passive precipitation:* Replace TB medium with M9 buffer as follows:
 - a. Allow bacteria to settle to the bottom of the culture flask overnight at 4°C. Carefully aspirate as much of the medium from the flask as possible without disturbing the bacteria, which forms a soft layer at the bottom of the flask. Swirl the flask to suspend the bacteria in the remaining volume of TB and pour into a 50-ml conical tube.
 - b. Pellet the bacteria in a clinical centrifuge 10 min at 800 × g (4500 rpm), 20°C. Remove broth by aspiration.

These pellets tend to be soft and pouring off the medium could result in loss of the bacterial pellet.
 - c. Wash the bacteria in 5 pellet volumes M9 buffer by vortexing until the clumps are dissolved.
6. Pellet the bacteria in a clinical centrifuge 10 min at 800 × g (4500 rpm), 20°C.

- 7a. *For immediate/short term use:* Resuspend washed bacteria in 5 pellet volumes of M9 buffer. Store up to 1 month at 4°C.

RNAi food is ready to use after resuspension in either buffer.

- 7b. *For storage:* Resuspend washed bacteria in M9/15% glycerol. Store in 10-ml aliquots indefinitely at –70°C.

8. Seed an appropriate number of NGM/amp/IPTG plates with one drop (~100 µl) each RNAi food.

9. Allow plates to dry overnight at room temperature.

This also allows induction of dsRNA production and consistently produces the most penetrant RNAi phenotypes by feeding.

Seeded plates can be stored in a sealed container up to several weeks at 4°C. Warm the RNAi plates to room temperature prior to use.

Cultivate worms on RNAi food and analyze phenotypes

- 10a. *Test for embryonic and post-embryonic phenotypes:* Place single gravid adult *C. elegans* onto NGM/amp/IPTG plates. As a control, place single gravid adults onto HT115(DE3) containing the empty L4440 RNAi feeding vector. After 24 hr, transfer each RNAi-fed or control-fed adult to a fresh plate of the same type and count the number of progeny (eggs and larvae) on each plate.

By placing gravid adults onto RNAi plates, one can assess the ability of the RNAi to produce embryonic as well as postembryonic phenotypes on the same plate.

- 10b. *Test for embryonic lethal phenotypes:* Place single L4 animals onto NGM/amp/IPTG (RNAi) and control—i.e., seeded with HT115(DE3) containing the empty L4440 RNAi feeding vector—plates. After 24 hr, transfer each adult to a fresh plate of the same type onto which it was seeded (i.e. RNAi or control) and count the number of progeny (eggs and larvae) on each plate.

- 10c. *Test for postembryonic phenotypes:* Place 20 newly hatched L1 larvae onto RNAi and control plates.

11. Examine the plates each day to determine whether the RNAi food produces any embryonic and/or postembryonic phenotypes.

Embryos that are still present on RNAi plates 24 hr after removing the adult indicates that the RNAi produces an embryonic lethal phenotype. However, always be sure to compare RNAi plates to control plates.

In some cases, postembryonic phenotypes can easily be determined by comparing the RNAi-fed progeny with the control-fed progeny.

SOAKING WORMS IN dsRNA

Soaking worms in dsRNA solution can also induce RNAi (Tabara et al., 1998). In this simple procedure, L4-stage animals are placed in soaking solution containing dsRNA corresponding to the gene to be silenced. Animals are soaked for up to 24 hr and recovered on Petri dishes. The exposed animals and their progeny are then examined for phenotypes.

The soaking procedure is capable of being scaled up or performed in microtiter format (Maeda et al., 2001). Another variation of the procedure is to soak L1-stage animals to examine postembryonic phenotypes.

Materials

dsRNA

Nonspecific dsRNA—i.e., dsRNA that is nonhomologous to any *C. elegans* sequence and which is comparable in length to the experimental dsRNA (e.g., *gfp*; optional)

Soaking solution (see recipe)

C. elegans (e.g., N2, CGC)

M9 buffer (see recipe)

NGM plates (see recipe) with and without seeded *E. coli* OP50 (see recipe)

65°C water bath
0.2- and 2.0-ml microcentrifuge tubes
5-ml glass pipet

Additional reagents and equipment for ethanol precipitation (*UNIT 2.1A*)

1. Ethanol precipitate (*UNIT 2.1A*) an appropriate amount of dsRNA sample and, if available, nonspecific dsRNA. Dissolve the pellet in soaking solution to a final concentration of 0.5 to 5.0 mg/ml.

It may be necessary to try several different concentrations of dsRNA to determine the concentration required to generate a phenotype.

2. Denature the RNA by heating 15 min at 65°C. Anneal the dsRNA by cooling the sample to room temperature for ~15 to 30 min.
3. Pipet 5 μ l sample dsRNA into an appropriately labeled 0.2-ml microcentrifuge tube. Do the same for the nonspecific dsRNA control, or if this is not available, use soaking solution alone.

It may also be useful to set up a positive control (i.e., a dsRNA proven to demonstrate a phenotype) to determine that the procedure is working as expected.

4. Using a 5-ml glass pipet, rinse a healthy mixed-stage population of *C. elegans* with 2 ml M9 buffer and transfer to a 2.0-ml microcentrifuge tube.
5. Pellet the animals by microcentrifuging 30 sec at 800 \times g, 20°C. Remove M9 buffer with a 1-ml pipet tip and wash again.
6. Transfer the worms to an unseeded NGM plate and let them crawl around for 5 to 10 min at 20°C to allow complete digestion of bacteria in the gut.
- 7a. *To test for embryonic lethal phenotypes:* Pick and transfer five to ten L4-stage animals to each sample or control tube (step 3). Incubate 24 hr at 20°C. Recover the soaked worms by plating them onto an NGM plate seeded with *E. coli* OP50, transfer each worm to an individual plate, and incubate 24 hr at 20°C. Transfer each worm to a fresh plate and count the number of progeny.
- 7b. *To test for postembryonic phenotypes:* Place 20 newly hatched L1 larvae into each sample or control tube and incubate 24 hr at 20°C. Recover soaked worms by plating them onto an NGM plate seeded with *E. coli* OP50. When the L1-soaked animals reach L4 stage, transfer each animal to an individual plate.
8. Monitor the treated animals and their progeny for mutant phenotypes resulting from dsRNA treatment.

SUPPORT PROTOCOL 1

PICKING AND PROPAGATING *C. ELEGANS*

C. elegans can be transferred from one plate to another using a worm pick—a short piece of platinum wire with one end flattened and the other end mounted to a Pasteur pipet. Platinum wire heats and cools rapidly and can and should be flamed before and after every pick to avoid contamination.

Label the bottom half of an NGM plate (see recipe) with the appropriate strain name and the date. Working under a dissecting microscope, pick up a dab of *E. coli* OP50 food (see recipe) on the underside of the flattened end of the worm pick. Gently touch the food on the bottom of the pick to the top of a worm. The worm should stick to the food on the worm pick by surface tension. Quickly transfer the worm to a new NGM plate, gently touching the worm to the plate and allowing the worm to swim off of the worm pick.

For those unfamiliar with *C. elegans*, propagating worms may be difficult at first, but with practice picking and maintaining worms will be routine. See Support Protocol 2, Hope (1999), and Wood (1988) for additional details on propagating, staging, cleaning, synchronizing, and mating *C. elegans*.

SEXING AND MATING *C. ELEGANS*

C. elegans comes in two sexes, hermaphrodite and male, which are determined by the number of X chromosomes. Males can be distinguished from hermaphrodites physically by the lack of hermaphrodite sex organs (e.g., vulva and uterus containing developing embryos) and by the tail, which has a barbed appearance under the dissecting microscope. Hermaphrodites (XX) produce both sperm and oocytes and can reproduce by self-fertilization. Hermaphrodites can also be cross-fertilized by *C. elegans* males (XO), which produce sperm only. If a *C. elegans* male successfully mates with a hermaphrodite, the majority of the progeny will be cross-progeny because male sperm outcompetes hermaphrodite sperm. As such, one half of the cross-progeny will be male and the other half will be hermaphrodites. For a detailed description of hermaphrodites and males and their identification, refer to Wood (1988) and Riddle et al. (1997).

Because *C. elegans* males are rare, try to set up a cross when a male is found. Males arise by spontaneous X chromosome nondisjunction at a frequency of ~0.1% (Riddle et al., 1997). This frequency can occasionally be increased by heat-shock treatment of L4 hermaphrodites at 30°C for 4 hr. Once a male stock is obtained, the population should be maintained by periodically setting up crosses.

Genetic crosses in *C. elegans* are typically set up between a phenotypically wild-type male and a hermaphrodite that is homozygous for a visible mutation—e.g., uncoordinated (Unc). This allows cross-progeny (non-Unc) to be easily distinguished from self-progeny (Unc). Multiple (e.g., ten) young healthy males and multiple (e.g., five) L4 or young adult hermaphrodites should be chosen when crossing *C. elegans* strains.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Agarose microinjection pads

Prepare injection pads by sandwiching a couple of drops of melted 2% (w/v) agarose between two 22 × 50-mm glass coverslips. Once the agarose hardens (<1 min), pop off one of the coverslips. Dry the agarose pad onto the remaining coverslip by baking at 50°C for 1 hr. A large number of microinjection pads can be made at one time and stored indefinitely at room temperature in a closed container (e.g., a Petri dish or empty coverslip box) to prevent dust from accumulating on the surface.

Baking dries the agarose to the coverslip. One of the coverslips must be removed before baking. The popped off coverslip can then be reused for another microinjection pad.

***E. coli* OP50 food**

Inoculate (UNIT 1.2) 100 ml of L broth (UNIT 6.12) with a single colony of OP50 (CGC) using sterile technique. Incubate the culture overnight at 37°C with shaking. Store the liquid culture at 4°C for several weeks

*The uracil auxotrophic *E. coli* OP50 strain is typically used as a food source for *C. elegans*, as the limited growth of *E. coli* OP50 on NGM plates provides easier observation of the worms.*

Loading capillary

The loading capillary consists of a drawn-out Pasteur pipet connected via tubing to a mouth pipet. The diameter of the drawn-out end of the Pasteur pipet must be less than that of the inner diameter of the nondrawn-out portion of the microinjection needle (see recipe) to facilitate loading.

M9 buffer with and without 15% glycerol

3.0 g monobasic potassium phosphate (KH₂PO₄)
6.0 g dibasic sodium phosphate (Na₂HPO₄)
5.0 g NaCl
1 ml 1 M magnesium sulfate (MgSO₄)
Add 150 ml glycerol (15%) if appropriate
Adjust volume to 1 liter with H₂O
If using glycerol, prepare 100-ml aliquots
Sterilize by autoclaving
Store up to 2 months at 20°C

Mating plates

Flame the end of a Pasteur pipet until it is rounded and nearly closed. Using the Pasteur pipet, withdraw some *E. coli* OP50 food (see recipe) from an overnight culture. Touch the rounded tip of the Pasteur pipet to the center of a 30-mm NGM plate (see recipe), dispensing a very small droplet of the *E. coli* OP50 food. Incubate the plates at room temperature 2 to 3 hr and store at 4°C in a sealed container for several weeks.

Mating plates are NGM plates seeded with a very thin and small (~5 mm) spot of E. coli OP50 in the center.

Alternatively, a small dab of E. coli OP50 food can be transferred from a dry seeded plate to an unseeded 30-mm NGM plate just prior to setting up a cross.

See UNIT 1.2 for more information regarding culture of E. coli.

Microinjection needles

The ideal microinjection needle (i.e., borosilicate glass capillaries, World Precision Instruments) should taper to a sharp and stiff point and can be pulled using 1.2- or 1-mm capillary blanks and a wide variety of needle-pulling devices (e.g., Narishige PN-30). After pulling the first two needles, examine them under a dissecting microscope. The needles should taper gradually and steadily to a very sharp point. Using one of the first two needles, gently touch the other needle just behind the tip while watching through the dissecting scope. If the needle tips bend easily without breaking then it is probably not stiff enough for penetrating the cuticle of the worm. A common problem is a needle that is hair-like at the tip. To correct this problem, reduce the heat setting on the needle puller. It may also be necessary to adjust the pull strength. Get advice from someone experienced in pulling capillaries if problems persist. Most needles will be closed at the tip after pulling and thus, it will be necessary to gently break the needle tip prior to microinjection. Alternatively, the needle tips can be etched with acid (Mello and Fire, 1995).

NGM plates or NGM/amp/IPTG plates

Prepare the following in a 2-liter Erlenmeyer flask:
3.0 g NaCl
17 g Agar
2.5 g Bacto-peptone
975 ml H₂O

Add a stir bar and sterilize by autoclaving. Cool to ~55°C and add 25 ml of 1 M potassium phosphate buffer, pH 6 (see recipe), 1 ml of 1 M CaCl₂, 1 ml of 5 mg/ml cholesterol in 100% ethanol, and 1 ml of 1 M MgSO₄. For NGM/amp/IPTG plates, also add 50 µg/ml ampicillin and 1 mM (final) IPTG (UNIT 1.4). Dispense into Petri dishes (~3 ml per 30-mm plate). Dry the plates at room temperature for 1 to 2 days before seeding with OP50. Store up to 1 month at 4°C in a sealed container.

NGM plates seeded with OP50

Seed each NGM plate (see recipe) with one drop (~100 µl) of *E. coli* OP50 food (see recipe). Allow the *E. coli* OP50 drop to dry into the plate and to grow overnight at room temperature or for 8 hr at 37°C (UNIT 1.2). Store seeded plates in a sealed container for several weeks at 4°C.

Potassium phosphate buffer, 1 M, pH 6.0

108.3 g monobasic potassium phosphate (KH₂PO₄)

35.6 g dibasic potassium phosphate (K₂HPO₄)

Adjust volume to 1 liter with H₂O

Adjust pH to 6.0 if necessary

Store up to 2 months at 20°C

Soaking buffer, 10×

1.55 g dibasic sodium phosphate (Na₂HPO₄)

0.75 g monobasic potassium phosphate (KH₂PO₄)

0.12 g NaCl

0.25 g ammonium chloride (NH₄Cl)

0.5 g gelatin

Adjust volume to 100 ml with H₂O

Sterilize by autoclaving

Store up to 2 months at 20°C

Soaking solution

0.5 ml 10× soaking buffer (see recipe)

15 µl 1 M spermidine (Sigma)

4.5 ml H₂O

Use immediately

COMMENTARY

Background Information

RNA interference (RNAi) is a form of sequence-specific, post-transcriptional gene silencing (PTGS) that is triggered by the introduction of dsRNA (UNIT 26.1; Fire et al., 1998). RNAi is related to a number of gene silencing phenomena in a variety of organisms. Genetic and biochemical studies are revealing a conserved set of factors required for silencing and demonstrating that the underlying mechanisms of these silencing phenomena are remarkably conserved (for reviews, see UNIT 26.1 and Hannon, 2002). The specificity and potency of RNAi has provided researchers with an efficient means of inactivating gene expression in model organisms where conventional gene replacement techniques are unavailable, such as *C. elegans*, *Drosophila*, fungi, plants, and try-

panosomes (Bosher and Labouesse, 2000). The ability to perform RNAi in human cells (Elbashir et al., 2001) has rapidly transformed basic research and holds promise as a therapy for human disease.

RNAi can inhibit gene function in worms injected with dsRNA (see Basic Protocol 1 and Alternate Protocol; Fire et al., 1998), soaked in solution containing dsRNA (see Basic Protocol 3; Tabara et al., 1998), fed bacteria expressing dsRNA (see Basic Protocol 2; Timmons et al., 2001), or engineered to express dsRNA in vivo (Tabara et al., 1999). Interference by dsRNA is sequence-specific and potent (Fire et al., 1998). Worms that are injected in the gut with ~60,000 dsRNA molecules targeting the muscle-specific *unc-22* gene exhibit an *unc-22*(-) twitching phenotype (Fire et al., 1998). Moreover, F1

progeny of injected worms also exhibit a severe twitching phenotype. These findings indicate that dsRNA, or a secondary silencing factor, is transported from one tissue to another and that interference can be inherited. The specific, potent, systemic, and heritable features of RNAi make this an ideal method for assaying the function of genes in *C. elegans*.

Not only does the ability to induce RNAi by a variety of methods allow researchers of different skill levels to apply the technique, it also allows them to characterize gene function under a variety of conditions. Typically, investigators determine whether a particular gene has an RNAi phenotype first by microinjection, simply because the preparation of dsRNA is straightforward and takes less time to prepare than RNAi food and plates (see Time Considerations). However, this does not necessarily mean that RNAi by microinjection is superior to other methods. While some genes are indeed silenced more effectively by microinjection than by feeding, the converse is also true (Kamath et al., 2001), and other genes are equally sensitive to both methods (Kamath et al., 2003; unpub. observ.). Therefore, if microinjection does not work, feeding and/or soaking should be attempted. For RNAi by feeding, Kamath et al. (2003) have generated a library of 16,757 RNAi bacterial feeding clones, representing 86% of the 19,427 predicted *C. elegans* genes.

Due to the requirement of a balancer chromosome for viability, a null mutation in an essential gene limits the ability to perform biochemistry or applications that require large populations. The ability to silence a gene in large populations of *C. elegans* using RNAi by feeding or soaking should make *C. elegans* more amenable to such applications.

RNAi is also a simple way to examine genetic interactions between genes. Sequential microinjection, microinjection or soaking with multiple dsRNAs simultaneously, exposure of mutant strains to dsRNA, as well as feeding constructs containing more than one insert have all been successfully used to determine whether two genes interact (Grishok et al., 2001; Bei et al., 2002; Dudley et al., 2002).

Despite its success, RNAi in *C. elegans* does have its limitations. RNAi does not work for every gene. In particular, pharyngeal, vulval, sperm, and neuronal genes are often difficult to silence (Fire et al., 1998). The reasons for this is not known, but could be due to inefficient transport or maintenance of trigger dsRNA, or secondary signals to these tissues. In some

cases, efficient silencing can be triggered for refractory genes by driving the expression of hairpin dsRNA from a transgene (Tavernarakis et al., 2000).

A large amount of RNAi data is being generated in *C. elegans*. The data can be accessed and searched from Wormbase (<http://www.wormbase.org>) to determine whether a particular gene exhibits an RNAi phenotype or to identify genes that show a particular phenotype. In general, the quality of the data is good, but keep in mind that a negative result does not necessarily mean that a particular gene does not exhibit an RNAi phenotype. It simply means that under the conditions tested no phenotype was observed. Investigators should try the experiment by additional methods before accepting a negative RNAi result.

A common problem with all RNAi methods is the possibility of inadvertently silencing related genes. Generally, this only occurs when a few segments of perfect nucleotide identity of at least 25 nucleotides exist within the gene. As a standard control for this problem in RNAi studies, it is always important to identify related genes in the genome and assay each of these for similar or overlapping phenotypes.

Although RNAi can be inherited (Fire et al., 1998; Grishok et al., 2000), in most instances—e.g., using BLAST (*UNIT 19.3*) or a similar search program—it is not permanent. While this may be considered as a benefit to some, to the classical geneticist this is a drawback. In order to produce a phenotype, the investigator must repeat the experiment from start to finish. Variability is typically not a problem, but a range of phenotypes is possible, so great care must be taken when assaying the animals. Because the feeding strategy (see Basic Protocol 2) requires less actual worm manipulation, it may be the more consistent protocol, provided care is taken when preparing the dsRNA food and media. When possible, it is a good idea to confirm RNAi phenotypes by obtaining a genetic mutant. Genetic mutants remain the gold standard in *C. elegans* research.

Critical Parameters

Target selection

There are few hard and fast rules for determining the region of a gene to target. Since introns and untranslated regions are not efficient targets of the RNAi machinery (Montgomery et al., 1998), the dsRNA should be designed to target the coding region of the mature mRNA. Although 500 to 1000 bp is

sufficient, if feasible, target the entire coding region of the gene of interest or several different regions. The sensitivity of worms to RNAi targeting different regions of a particular gene can be determined in a relatively short period of time by microinjection.

As mentioned above, care should be taken when performing RNAi on members of a gene family. Experiments should be designed such that the targeted region of a gene of interest does not overlap with the other member(s) of the gene family. This may be difficult if extensive homology exists, but in most cases this should not be a problem. Use BLAST (UNIT 19.3) to identify related sequences in *C. elegans* and to determine whether there is significant identity to complicate the analysis. If there is significant identity between two sequences, attempt to find and target regions that are divergent.

Obtaining cDNA clones

The laboratory of Yuji Kohara (ykohara@lab.nig.ac.jp) has generated a library of cDNAs expressed in *C. elegans*. In addition, they have generated sequence tags for a large number of these clones and assigned them to the corresponding genes. If a cDNA for a particular gene has been identified, the clone or yk (for Yuji Kohara) number will be listed at the bottom of the Gene Report Web page at Wormbase (<http://www.wormbase.org>; see Internet Resources) under the subheading Reagents: Matching cDNAs. The cDNA clones are made available to the *C. elegans* community and are provided as a small aliquot of Lambda ZAP II bacteriophage. A kit provided by Stratagene can be used to amplify the phage and to excise the pBluescript SK(-) phagemid. Each insert is flanked by the phage T3 and T7 RNA polymerase promoters, which can be used to synthesize sense and antisense RNA strands in vitro (UNIT 3.8). If a yk clone is not available, a cDNA can be generated by RT-PCR from total worm RNA or polyA-selected mRNA using sequence-specific oligonucleotides and cloned into an appropriate expression vector (e.g., pBluescript).

Transcription templates can also be generated by PCR from genomic DNA using oligonucleotides with 5' tails containing phage T7 RNA polymerase promoter sequences. In this case, try to identify and amplify the longest contiguous stretch of exon sequence possible, minimizing the inclusion of introns. The resultant PCR product will contain a portion of the gene of interest flanked by opposing T7 promoter sites. Using such a template, both sense

and antisense strands of RNA can be produced simultaneously in a single in vitro transcription reaction.

When generating clones for RNAi by feeding, inserts should be exon rich. The authors have had success with a variety of feeding constructs, targeting sequences from 300 to 2400 nt long. The optimal length of the targeted region for RNAi by feeding is likely to be gene specific and should be determined empirically; however, the parameters discussed above represent a good starting point.

dsRNA preparation

Sense and antisense RNA strands are prepared using standard protocols and kits, and then annealed to make dsRNA. Several companies now provide high yield in vitro transcription kits. The Megascript Kit from Ambion is outstanding and the instruction manual provides a lot of useful information, from designing oligonucleotides containing phage promoters to troubleshooting transcription reactions.

Efficient RNAi will be achieved for many genes over a very broad range of concentrations (0.01 to 5 mg/ml). Initially, higher concentrations (i.e., 1 mg/ml) are preferred because some genes are less sensitive to RNAi than others for unknown reasons. A small percentage (<5%) of nonspecific embryonic lethality among progeny of the injected animals is sometimes observed; however, there are no other noticeable side effects from injecting high dsRNA concentrations.

At very low concentrations of dsRNA, one may be able to induce hypomorphic phenotypes. Mixing two different RNAi feeding strains has been shown to dramatically reduce the penetrance of mutant phenotypes observed with one feeding strain alone (Kamath et al., 2001). These authors also found that a 1:1 mixture of *unc-37* RNAi food and normal OP50 (which does not express the dsRNA) completely abolished the embryonic lethality induced by *unc-37* RNAi; however, 100% of the progeny displayed hypomorphic postembryonic phenotypes, which was also observed when the RNAi food was induced with very low levels of IPTG. Therefore, adjusting the levels of dsRNA delivered to the animals can alter the penetrance and spectra of RNAi phenotypes.

Silencing multiple genes simultaneously

As discussed above, care should be taken when designing RNAi experiments so that simultaneous silencing of related genes does not

Table 26.3.1 Troubleshooting Guide for *C. elegans* RNAi Studies

Problem	Possible cause	Solution
<i>RNAi by microinjection</i>		
dsRNA solution does not flow from the needle	Needle is closed	Etch or break the needles after pulling to open
	Needle is clogged	Break the tip of the needle to restore flow Use a new needle Centrifuge dsRNA solution prior to use
Worms do not stick to agarose injection pad	Agarose is not dry enough	Bake agarose pads at least 1 hr at 50°C
No RNAi phenotype observed	dsRNA is degraded	Prepare fresh dsRNA and analyze on 1% agarose gel to check quality (<i>UNIT 2.5A</i>) Perform positive control for RNAi by injection
	Target gene is refractory to RNAi	Try another region of the gene Attempt to achieve RNAi by feeding or soaking
<i>RNAi by feeding</i>		
No RNAi phenotype observed	dsRNA production is not induced	Test the plates by performing a positive control for RNAi Prepare fresh NGM/Amp/IPTG plates
	RNAi food is contaminated	Prepare fresh RNAi food. Be sure to include both ampicillin and tetracycline in the starter culture to prevent contamination with another bacterial species. Perform all food preparation carefully.
<i>RNAi by soaking</i>		
No RNAi phenotype observed	dsRNA is degraded	Prepare fresh dsRNA and soaking solution

occur inadvertently. Nevertheless, there may be instances when it is desirable to silence multiple genes simultaneously. For instance, two related genes that function redundantly in a particular process may show no phenotype or produce an incompletely penetrant phenotype when silenced individually by RNAi. On the other hand, simultaneous RNAi targeting of both genes may dramatically produce a completely penetrant phenotype (Grishok et al., 2001). As mentioned above, RNAi is also a simple and rapid method for examining genetic interactions. In principle, each of the dsRNA delivery methods is amenable to simultaneous RNAi; however, microinjection is the most widely used technique for this purpose. Simultaneous RNAi by feeding also works, although it is somewhat limited relative to microinjection. As mentioned above, mixing two different RNAi feeding strains together dramatically reduces the penetrance of RNAi, but by engineering the

feeding vector to contain two inserts, the authors have been able to silence two genes simultaneously by double RNAi feeding without an apparent reduction in the penetrance of RNAi relative to either single RNAi food. The combined inserts so far have not exceeded 1.5 kb in length, so presumably additional genes could be targeted, for example in a triple RNAi feeding construct.

Troubleshooting

Table 26.3.1 presents a guide to some of the common problems of achieving RNAi in *C. elegans*, as well as their causes and solutions.

Anticipated Results

Regardless of the delivery method, if RNAi produces a phenotype, the animals will typically be affected for one generation—i.e., the F1 progeny of animals exposed to dsRNA. However, for RNAi by feeding, as long as the

animals are cultivated on RNAi food, they are exposed to dsRNA and the phenotype will persist. Investigators have observed persistent RNAi by microinjection into the F2 generation, although the penetrance is reduced.

In a typical RNAi microinjection experiment, early embryos developing within the uterus may be unaffected, but may carry the silencing signal into the subsequent generation. These unaffected and carrier progeny will be laid on the initial recovery plate after microinjection. After transferring an injected animal from the recovery plate to a fresh plate (~16 hr), the remaining portion of the brood should be completely affected by the RNAi.

For uncharacterized genes, it is difficult to know what to expect from an RNAi experiment; however, if the expression pattern or localization is known (e.g., from *in situ* hybridization), then an estimated guess can be made about the type of phenotype for which to look. For example, if *in situ* analysis indicates that the gene product is expressed in the germline, one might expect to see an embryonic-lethal phenotype or postembryonic-sterile phenotype. For a gene product expressed in muscle cells, one might expect defects in motility. *In situ* expression data can be searched from Wormbase or from the Nematode Expression Pattern Database (NEXTDB; see Internet Resources).

It is recommended that the novice enlist the aid of individuals with expertise in *C. elegans* biology when phenotyping animals.

Time Considerations

RNAi by microinjection

A complete microinjection experiment from injecting animals to scoring phenotypes will take 3 to 6 days. On day 0, the novice should expect to spend about 1 to 2 hr at the injection microscope performing microinjection. With practice, this time can be shortened to ~15 min for each dsRNA and strain. The amount of time to transfer animals to individual NGM plates on days 1 and 2 will vary depending on the number of worms to transfer (5 to 30 min). Expect to spend several hours counting progeny and analyzing phenotypes.

RNAi by feeding

Growth, preparation, and induction of the RNAi food will take 2 to 4 days. The remainder of the feeding experiment will take 3 to 6 days to complete. As for microinjection, plan to

spend several hours counting progeny and analyzing phenotypes.

RNAi by soaking

The amount of time it will take to perform the entire soaking procedure will be similar to that of microinjection.

Literature Cited

- Ashrafi, K., Chang, F.Y., Watts, J.L., Fraser, A.G., Kamath, R.S., Ahringer, J., and Ruvkun, G. 2003. Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. *Nature* 421:268-272.
- Bei, Y., Hogan, J., Berkowitz, L.A., Soto, M., Rocheleau, C.E., Pang, K.M., Collins, J., and Mello, C.C. 2002. SRC-1 and Wnt signaling act together to specify endoderm and to control cleavage orientation in early *C. elegans* embryos. *Dev. Cell* 3:113-125.
- Bosher, J.M. and Labouesse, M. 2000. RNA interference: Genetic wand and genetic watchdog. *Nat. Cell. Biol.* 2:E31-36.
- The *C. elegans* Sequencing Consortium. 1998. Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* 282:2012-2018.
- Dudley, N.R., Labbe, J.C., and Goldstein, B. 2002. Using RNA interference to identify genes required for RNA interference. *Proc. Natl. Acad. Sci. U.S.A.* 99:4191-4196.
- Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411:494-498.
- Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391:806-811.
- Fraser, A.G., Kamath, R.S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., and Ahringer, J. 2000. Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature* 408:325-330.
- Gonczy, P., Echeverri, C., Oegema, K., Coulson, A., Jones, S.J., Copley, R.R., Duperon, J., Oegema, J., Brehm, M., Cassin, E., Hannak, E., Kirkham, M., Pichler, S., Flohrs, K., Goessen, A., Leidel, S., Alleaume, A.M., Martin, C., Ozlu, N., Bork, P., and Hyman, A.A. 2000. Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature* 408:331-336.
- Grishok, A., Tabara, H., and Mello, C.C. 2000. Genetic requirements for inheritance of RNAi in *C. elegans*. *Science* 287:2494-2497.
- Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 106:23-34.

- Hannon, G.J. 2002. RNA interference. *Nature* 418:244-251.
- Hope, I.A., Ed. 1999. *C. elegans: A Practical Approach*. The Practical Approach Series. Oxford University Press, New York.
- Jansen, G., Hazendonk, E., Thijssen, K.L., and Plasterk, R.H. 1997. Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans*. *Nat. Genet.* 17:119-121.
- Kamath, R.S., Martinez-Campos, M., Zipperlen, P., Fraser, A.G., and Ahringer, J. 2001. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. *Genome Biol.* 2:RESEARCH0002.
- Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D.P., Zipperlen, P., and Ahringer, J. 2003. Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* 421:231-237. [comment 220-221]
- Liu, L.X., Spoerke, J.M., Mulligan, E.L., Chen, J., Reardon, B., Westlund, B., Sun, L., Abel, K., Armstrong, B., Hardiman, G., King, J., McCague, L., Basson, M., Clover, R., and Johnson, C.D. 1999. High-throughput isolation of *Caenorhabditis elegans* deletion mutants. *Genome Res.* 9:859-867.
- Maeda, I., Kohara, Y., Yamamoto, M. and Sugimoto, A. 2001. Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. *Curr. Biol.* 11:171-176.
- Mello, C. and Fire, A. 1995. DNA transformation. *Methods Cell. Biol.* 48:451-482.
- Montgomery, M.K., Xu, S., and Fire, A. 1998. RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* 95:15502-1557.
- Piano, F., Schetter, A.J., Mangone, M., Stein, L., and Kempthues, K.J. 2000. RNAi analysis of genes expressed in the ovary of *Caenorhabditis elegans*. *Curr. Biol.* 10:1619-1622.
- Riddle D.L., Blumenthal, T., Meyer, B.J., Priess, J.R. 1997. *C. ELEGANS II*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Tabara, H., Grishok, A., and Mello, C.C. 1998. RNAi in *C. elegans*: Soaking in the genome sequence. *Science* 282:430-431.
- Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. 1999. The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* 99:123-132.
- Tavernarakis, N., Wang, S.L., Dorovkov, M., Ryazanov, A., and Driscoll, M. 2000. Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. *Nat. Genet.* 24:180-183.
- Timmons, L., Court, D.L., and Fire, A. 2001. Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* 263:103-112.
- Wood, W.B. (ed.) 1988. The Nematode *Caenorhabditis elegans*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Key References

Hope, 1999. See above.

A very good source for general information on culturing and maintaining worms as well as a comprehensive microinjection protocol for germline transformation.

Kamath et al., 2001. See above.

This manuscript provides an in depth analysis of the RNAi by feeding technique as well as a comparison to RNAi by microinjection.

Riddle et al., 1997. See above.

C. ELEGANS II provides an extensive overview of worm biology.

Wood, 1988. See above.

The predecessor of C. ELEGANS II (Riddle et al., 1997). The Nematode is full of important information about the biology of C. elegans.

Internet Resources

<http://www.wormbase.org>

Wormbase is an essential tool for C. elegans biologists, and provides an abundance of useful information regarding the genome and biology of C. elegans. The curators at Wormbase have provided annotations of each of the predicted genes including RNAi phenotypes and expression data as well as links to useful sites. The curators are constantly updating and upgrading the site.

http://nematode.lab.nig.ac.jp/db/rnai_s/RNAiBySoaking.html

This site includes several protocols for RNAi by soaking as well as for preparing template DNA and dsRNA for the RNAi by soaking protocols.

<http://www.rnai.org>

The RNAi database is a searchable database of ovary or germline-enriched genes by gene or by phenotype.

<http://c.elegans.tripod.com/RNAi.htm>

The RNAi@elegans.net page has many informative links for RNAi in C. elegans and other organisms.

<http://elegans.swmed.edu>

The C. elegans WWW server.

<http://nematode.lab.nig.ac.jp/db/index.html>

NEXTDB is an expression pattern database based on in situ hybridizations using expressed sequence tags.

Contributed by Darryl Conte Jr. and

Craig C. Mello

University of Massachusetts Medical School
Worcester, Massachusetts

Cloning of Small RNA Molecules

In the past few years, it has become evident that double-stranded RNA–derived small RNAs are key regulators of gene expression in plants, animals, and most fungi. These small RNAs are involved in a variety of gene silencing processes that are now referred to as RNA silencing. The most well known of these phenomena is RNA interference (RNAi), which leads to sequence-specific degradation of mRNAs that are identical in sequence to the initiator dsRNA. The guide RNAs in RNAi are called small interfering RNAs (siRNAs), which form ribonucleoprotein complexes that mediate target mRNA cleavage.

Another group of small RNAs, the microRNAs (miRNAs), represent evolutionarily conserved small RNAs that are derived from gene products that code for short double-stranded RNA precursors. The two groups of small RNAs, miRNAs and siRNAs, share common features including their biogenesis. In plants, miRNAs act predominantly as siRNAs by guiding degradation of target mRNAs that show near perfect complementarity to the miRNAs. In contrast, animal miRNAs are believed to bind to partially complementary repetitive sequence elements within the 3′-untranslated region of the target mRNAs and repress translation. Artificial targets in animal cells fully complementary to endogenous miRNAs are subjected to degradation.

The most recent class of characterized small RNAs appears to be involved in initiating or maintaining heterochromatin and has been referred to as heterochromatic siRNAs (hsRNAs). Finally, there is emerging evidence for small RNAs participating in directed elimination of DNA sequences during macronuclei differentiation in *Tetrahymena thermophila*.

Because of their important roles in gene regulation, it is of significant interest to characterize the small regulatory RNAs of an organism at the cellular level. Cloning and sequencing of small RNAs from cultured cells or tissues is the most straightforward approach, but it is a technically demanding protocol. If only a given gene or sequence is to be characterized, northern blotting or nuclease-protection techniques may be employed.

This unit provides a detailed stepwise protocol for cloning small RNAs from tissues or cultured cells. Cloning is most successful when starting with 200 to 1000 μg of total RNA. A flowchart of the process is shown in Figure 26.4.1.

Small RNAs of the desired size range (19 to 24 nt in mammals, 19 to 29 nt in *Drosophila* or plants) are gel-isolated from total RNA and subjected to sequential 3′- and 5′-adapter ligation. The oligonucleotide adapters are needed for priming reverse transcription and for defining the orientation of the cloned small RNA. There are two protocols for adapter ligation; the Basic Protocol for ligation uses exclusively commercial reagents, while the Alternate Protocol requires a short chemical synthesis of an adenylated 3′-adapter oligonucleotide. If many ligation reactions are anticipated, the Alternate Protocol should be followed in order to save time.

CLONING OF SHORT RNA MOLECULES

In this basic adapter ligation protocol, the small RNAs are first dephosphorylated prior to ligation of the 3′ adapter to prevent circularization of the small RNAs. A 5′-phosphorylated 3′-adapter oligonucleotide with a blocked 3′-hydroxyl group is attached to the small RNAs. The ligation product is separated from excess 3′ adapter and is subsequently 5′ phosphorylated and subjected to 5′-adapter ligation. The Alternate Protocol uses an

BASIC PROTOCOL

Gene Silencing

26.4.1

Contributed by Sébastien Pfeffer, Mariana Lagos-Quintana, and Thomas Tuschl

Current Protocols in Molecular Biology (2003) 26.4.1-26.4.18

Copyright © 2003 by John Wiley & Sons, Inc.

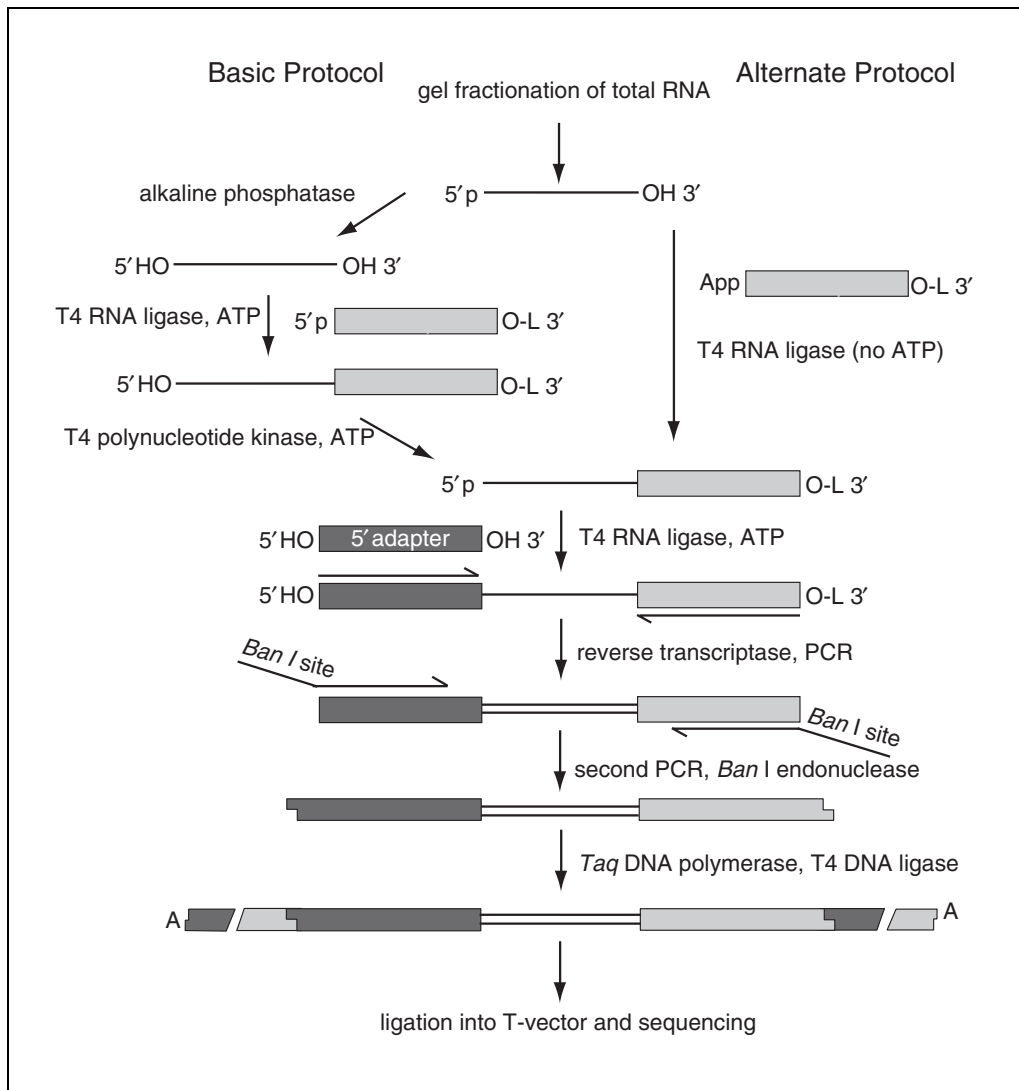


Figure 26.4.1 Schematic drawing of the Basic and Alternate Protocols small RNA cloning steps.

adenylated 3' adapter in the first ligation step in the absence of ATP, therefore bypassing the need to dephosphorylate the small RNAs to prevent circularization during T4 RNA ligation. The ligation product is joined to the 5' adapter without any further 5' phosphorylation.

The final ligation products are reverse transcribed into cDNA and then PCR amplified. To enable directional concatamerization, an additional PCR reaction is performed to introduce non-palindromic *Ban I* restriction sites at both ends of the PCR product. The *Ban I*-digested PCR products are then concatamerized using T4 DNA ligase and subsequently ligated into a commercial TOPO-TA cloning vector. Individual colonies are screened for inserts by PCR. The insert-containing PCR products are purified and submitted for sequencing. The RNA sequences are extracted manually or by using appropriately developed software tools.

Materials

- 200 to 1000 µg of total RNA isolated from cultured cells or tissues (UNIT 4.2)
- Denaturing solution (UNIT 4.2)
- Tris-buffered water-saturated phenol (UNIT 4.2), pH 7.8
- Deionized formamide

Gel-loading solution (see recipe)
 10-well 15% denaturing gel (15 × 7 × 0.15–cm, 50-ml gel volume, see *UNIT 2.12*)
 Radiolabeled RNA size markers
 0.5× TBE (*APPENDIX 2*)
 0.3 M, 0.4 M, and 5 M NaCl (RNase-free)
 Absolute ethanol
 RNase-free water
 10× phosphatase buffer (New England Biolabs)
 20 U/μl alkaline phosphatase (CIP; New England Biolabs)
 25:24:1 (v/v/v) phenol/chloroform/isoamyl alcohol
 Chloroform
 100 μM non-radioactive 5′-phosphorylated 3′-adapter oligonucleotide (see *UNIT 3.10*
 and Table 26.4.1; Fig. 26.4.2A,B)
 10 μM ³²P-labeled 3′-adapter oligonucleotide (see *UNIT 3.10*)
 10× RNA ligation buffer (see recipe)
 50% (v/v) aqueous dimethyl sulfoxide (DMSO)
 20 U/μl T4 RNA ligase (New England Biolabs)
 20-well 15% acrylamide gel (15 × 17 × 0.08–cm, 30-ml gel volume, see *UNIT 2.12*;
 see recipe)
 10× PNK buffer (New England Biolabs)
 100 mM ATP, pH 7.0
 10 U/μl T4 polynucleotide kinase (New England Biolabs)
 100 μM reverse transcription (RT) primer (see Table 26.4.1)
 0.1 M DTT
 5× first-strand buffer
 10× dNTP solution (dATP, dCTP, dGTP, dTTP, 2 mM each, pH 7.5)
 200 U/μl reverse transcriptase (Superscript II, RNase H(-) M-MLV reverse
 transcriptase, Invitrogen)
 150 mM KOH/20 mM Tris base
 150 mM HCl
 10× PCR buffer
 100 μM first PCR 5′ primer and first PCR 3′ primer (see Table 26.4.1)
 5 U/μl *Taq* polymerase
 2% (w/v) standard agarose gel (*UNIT 2.5A*)
 25-bp DNA ladder (Invitrogen)
 100 μM second PCR 5′ primer and second PCR 3′ primer (see Table 26.4.1)
 20 U/μl *Ban* I restriction endonuclease and 10× NEB buffer 4 (New England
 Biolabs)
 2000 U/μl T4 DNA ligase and 10× DNA ligation buffer (New England Biolabs)
 100-bp DNA ladder (New England Biolabs)
 2% (w/v) NuSieve (low-melt) agarose gel (Cambrex)
 10× TAE (*APPENDIX 2*)
 TOPO-TA cloning kit with TOP10 cells and pCR2.1 vector (Invitrogen)
 100 μM primer M13 (–20) F (see Table 26.4.1)
 100 μM primer M13 R (see Table 26.4.1)
 QIAquick PCR purification kit (Qiagen)
 Homogenizer
 Spectrophotometer and 1-cm quartz cuvette
 Plastic wrap
 Phosphorimaging screen and phosphorimager (see *APPENDIX 3A*)
 1.5-ml siliconized Eppendorf-style reaction tubes (Bio Plas)
 Eppendorf or Speed-Vac concentrator

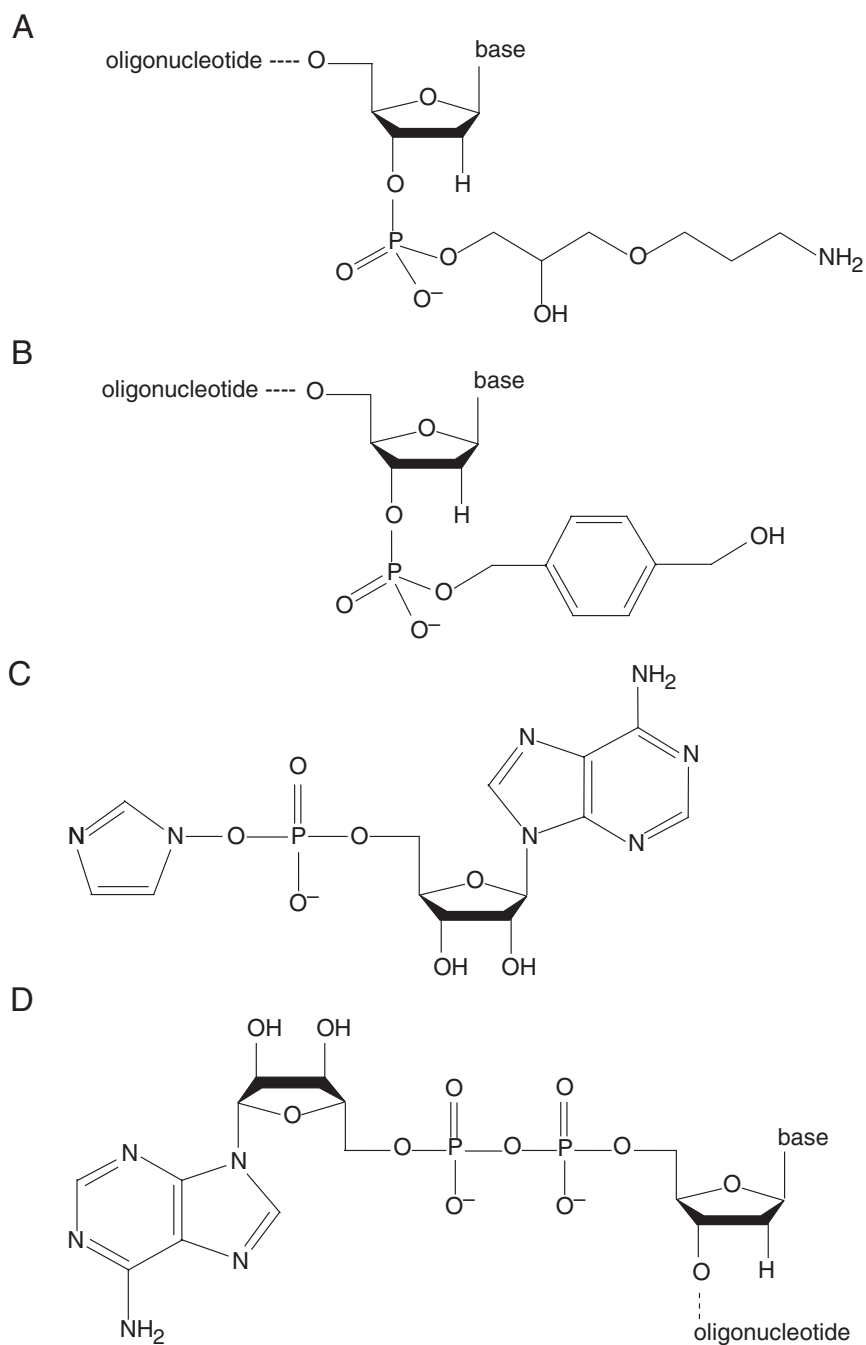


Figure 26.4.2 Structure of special reagents. **(A, B)** Structure of two alternative 3' blocks of the 3' adapter oligonucleotide after chemical synthesis and deprotection. Commercial DNA/RNA synthesis product providers use similar reagents. The C-7 linker in **(A)** was purchased from Chemgenes as 3'-TFA MODIFIER C-7 ICAA CPG and used for chemical DNA/RNA synthesis in house. The 3' blocking group after deprotection is shown in **(B)**, used during the authors' cloning experiments in 2001. The reagent used for the synthesis was dimethoxytrityl-dimethylbenzyl-succinyl-aminopropyl CPG and was a gift from Brian Sproat. Again, this 3'-modifier group was coupled to CPG for chemical DNA/RNA synthesis. **(C, D)** Reagents for the Alternate Protocol. **C**, ImpA; **D**, 3'-adenylated 3'-adapter oligonucleotide.

500- μ l PCR tubes
Thermal cycler
360-nm UV transilluminator
96-well thermocycler-compatible microtiter plates

Additional reagents and equipment for RNA extraction (UNIT 4.2), cloning and purification (Chapter 1)

Isolate total RNA

1. Extract total RNA according to the Basic Protocol of UNIT 4.2, except for the following modifications:

- a. Prepare an RNA extraction solution by pre-mixing one volume of denaturing solution with one volume acidic water-saturated phenol.
- b. In a homogenizer, pour 4 ml extraction solution (from step a, instead of the denaturing solution) per gram of tissue or cell directly onto the liquid nitrogen frozen powdered tissue or cell pellet. Follow the standard RNA isolation protocol (UNIT 4.2), but do not add any additional phenol after this homogenization step.

Do not resuspend the RNA pellet in 75% aqueous ethanol. Short RNAs are soluble to some extent in 75% ethanol in the absence of salt and a detached pellet is easily lost. The purpose of the recommended resuspension of the RNA pellet is mostly to remove residual salt. If an ethanol wash step is performed, carefully rinse the tube walls with 75% ethanol without perturbing the pellet and centrifuge immediately.

2. Dissolve the RNA pellet from 1 g of tissue in 200 μ l deionized formamide. Estimate the concentration of the total RNA after diluting with water and measuring the absorbance (A_{260}) of the RNA solution in a 1-cm quartz cuvette.

The RNA is dissolved in deionized formamide rather than water for more effective denaturation prior to gel loading.

The concentration (c) of total RNA (in μ g/ μ l) is calculated according to:

$$c = A_{260} \times f \times 0.04 \mu\text{g}/\mu\text{l}$$

where, f is the dilution factor. The yield of total RNA is ~1 mg per gram of tissue or cells and the concentration is ~5 μ g/ μ l.

Commercial RNA isolation kits with ion-exchange columns should not be used because they are not designed to isolate small RNAs.

Aqueous LiCl precipitation, which is frequently used to precipitate long RNA transcripts, must be avoided because small RNAs and tRNAs do not precipitate.

Gel-purify small RNA from total RNA

3. Dilute 600 μ g of sample RNA with gel-loading solution to a final volume of 200 μ l. Denature the RNA solution by incubating the tube for 30 sec at 90°C. Load the RNA solution in two center wells of a 10-well 15% denaturing gel (15 \times 7 \times 0.15-cm, 50 ml gel volume, see UNIT 2.12).
4. Load a cocktail of radiolabeled RNA size markers. Apply the size marker cocktail in a similar volume as the RNA sample onto the gel to ensure similar gel running behavior. To avoid spillover of marker (and subsequent cloning of marker RNA), leave one well empty between sample and size markers.

Size markers are radiolabeled individually in a 10- μ l reaction at 1 μ M concentration by T4 polynucleotide kinase using γ -³²P-ATP as described in UNIT 3.10. After gel purification of the markers, dissolve each of the marker oligoribonucleotides in 20 μ l of formamide. Use 1- μ l aliquots to prepare the size marker cocktail.

5. Run the gel for 1 to 1.5 hr at 30 W using 0.5× TBE buffer until the bromophenol blue dye of the gel-loading solution is 2 to 3 cm from the bottom of the gel.

Do not run the gel further than necessary for the separation of the size markers, because the recovery of small RNAs is more efficient if the small RNAs are concentrated in a narrow band.

6. Dismantle the gel and wrap the gel in plastic wrap (e.g., Saran wrap) to protect it from the environment. Expose the wrapped gel to a phosphorimaging screen for ~30 min (see APPENDIX 3A).

To facilitate the alignment of the gel to the phosphorimager printout, implanting three tiny radioactive gel pieces asymmetrically near the borders of the gel after removal of one of the glass plates is recommended. Radioactive gel pieces can be collected from the gel that was used to purify the size markers after initial ³²P-labeling.

7. Print out a 100%-scaled image of the gel, position the gel on top of the printout, and align to the size marker lanes. Excise the bands defined by the mobility of the RNA size markers.

8. Excise the radioactive band containing the size markers and elute the marker RNAs to run as a control. Run the control RNAs in parallel with the RNA sample for the adapter ligation steps.

The eluted markers will be used to control for RNA recovery and are subsequently needed for controlling the ligation reactions and to produce new size markers for the adapter ligation products.

9. Cut each gel slice into small pieces and transfer to a pre-weighed 1.5-ml siliconized tube. Determine the weight of the gel slices. Add 2 to 3 vol RNase-free 0.3 M NaCl and elute the small RNAs from the gel by incubating the tube at 4°C, overnight, under constant agitation (e.g., on a rocker in the cold room).

It is extremely important to carry out all RNA elution or reaction steps in siliconized reaction tubes. The walls of regular, untreated reaction tubes otherwise absorb the tiny amounts of small RNA and adapter ligation products.

10. Collect the supernatant and precipitate the small RNAs at –20°C, overnight, by adding 3 to 4 vol absolute ethanol.

Dephosphorylate small RNAs

11. Collect the small RNA pellet after ethanol precipitation in a tabletop centrifuge for 10 min at maximum speed (16,000 × g), 4°C. Remove the supernatant and collect the residual liquid at the bottom of the tube by an additional 10-sec centrifugation. Remove the residual liquid completely using a small pipet tip without disturbing the pellet.

The additional centrifugation is needed to collect all residual liquid and prevents NaCl contamination of the pellet.

12. Dry the tube for 2 min under vacuum using the Eppendorf or Speed-Vac concentrator and dissolve the pellet in 27 μl RNase-free water (dissolve the pellet in 10 μl RNase-free water for the Alternate Protocol).

All RNA precipitations in subsequent steps shall be performed similarly, but the volumes in which the pellets are dissolved may vary.

13. Add 3 μl of 10× phosphatase buffer and 0.5 μl of 20 U/μl alkaline phosphatase. Incubate the reaction mixture 30 min at 50°C.

14. Add 45.5 μ l RNase-free water and 4.5 μ l of 5 M NaCl and perform two 100- μ l 25:24:1 phenol/chloroform/isoamyl alcohol extractions followed by a final 100- μ l chloroform extraction (the upper phase always contains the RNA).
15. Precipitate the dephosphorylated RNA by adding 3 to 4 vol absolute ethanol and incubating on ice for at least 1 hr.

The precipitation may also be carried out at -20°C overnight.

Washing RNA pellets is not recommended because small pellets are frequently lost. Salt contamination is avoided by complete removal of the precipitant solution by a second short centrifugation step rather than washing the pellet.

When an RNA pellet is dissolved in a small volume of water, it is best to add the water on top of the pellet and leave the tube undisturbed for 15 min at room temperature. If the tube is vortexed too quickly after adding water, the partly dissolved pellet may be displaced to the side of the tube where it will not dissolve completely.

It is preferable to adjust RNA solutions to a 0.3 M NaCl concentration before gel elution or phenol/chloroform/isoamyl alcohol extraction while the precise volume of the solution is still known. Changes in the volume of the aqueous solution during extraction or elution due to incomplete recovery of the input solution makes it difficult to precisely adjust the NaCl concentration; imprecise NaCl concentration may affect the RNA precipitation.

Ligate phosphorylated 3' adapter to dephosphorylated small RNAs

16. Collect the RNA pellet as described in step 11. Dissolve the sample pellet and the marker pellet in 10 μ l RNase-free water.
17. Add 1 μ l of 100 μ M non-radioactive 5'-phosphorylated 3'-adapter oligonucleotide (see UNIT 3.10 and Table 26.4.1), 1 μ l of 10 μ M ^{32}P -labeled 3'-adapter oligonucleotide (see UNIT 3.10), 2 μ l of 10 \times ligation buffer, and 6 μ l of 50% aqueous DMSO.
18. Denature the RNA by incubating the tube for 30 sec at 90°C . Place the tube immediately on ice for 20 sec.
19. Add 1 μ l of T4 RNA ligase, mix gently, and incubate for 1 hr at 37°C .

Table 26.4.1 Oligonucleotides and Their Corresponding Sequences^a

Oligonucleotide	Sequence ^a
19-nt RNA marker	r(CGUACGCGGAAUACUUCGA)
24-nt RNA marker	r(CGUACGCGGAAUACUUCGAAAUGU)
33-nt RNA marker	r(CCAUCGAUAAAAAAUAUGGAGAGCUUCCCCGAAG)
3' adapter (see Basic Protocol)	rUrUrUAACCGCGAATTCCAG-L
3' adapter (see Alternate Protocol)	TTTAACCGCGAATTCCAG-L
5' adapter	ACGGAATTCCTCACTrArArA
RT primer, 1st PCR 3' primer	GACTAGCTGGAATTCGCGGTAAA
1st PCR 5' primer	CAGCCAACGGAATTCCTCACTAAA
2nd PCR 3' primer	GACTAGCTTGGTGCCGAATTCGCGGTAAA
2nd PCR 5' primer	GAGCCAACAGGCACCGAATTCCTCACTAAA
M13 -20 F primer	GTAAAACGACGGCCAG
M13 R primer	CAGGAAACAGCTATGAC

^aA, C, G, T, DNA residues; rA, rC, rG, rU, RNA residues; L, 3' hydroxyl blocking group. 3'-Adapter oligonucleotides with C3- or C7-amino modifiers as blocking groups can be purchased from commercial DNA/RNA synthesis companies.

20. Add 1 vol of gel-loading solution and load the sample in 2 wells of a 20-well 15% acrylamide gel (15 × 17 × 0.08–cm, 30-ml gel volume). Load the ligated marker sample next to the sample, but not directly adjacent to avoid contamination.
21. Run the gel for ~1 hr at 28 W using 0.5× TBE buffer until the bromophenol blue dye is ~2 cm from the bottom of the gel.
22. Dismantle and wrap the gel in plastic wrap to protect it from the environment. Record the position of the marker and the ligation products by phosphorimaging as described in steps 6 and 7.
23. Excise the ligation product band as well as the ligated size markers and elute the ligation products from the gel slices overnight at 4°C using 300 µl of 0.3 M NaCl in siliconized tubes.
24. Precipitate the RNA by addition of 3 to 4 vol absolute ethanol as described in step 14.

It is not always possible to detect the ligation product for the sample lane, especially when starting with a small amount of RNA. If a ligation band is not detected in the sample lane, refer to the marker ligation products and cut out a band of corresponding mobility.

Perform 5' phosphorylation of the ligation product

25. Collect the RNA pellet as described in step 4. Dissolve the pellet in 10 µl RNase-free water.
26. Add 2 µl of 10× PNK buffer, 0.4 µl of 100 mM ATP, 7.1 µl RNase-free water, and 0.5 µl of T4 polynucleotide kinase, and incubate for 30 min at 37°C.
27. Stop the reaction by adding 37 µl RNase-free water and 3 µl of 5 M NaCl. Perform one 50-µl 25:24:1 phenol/chloroform/isoamyl alcohol extraction followed by a 50-µl chloroform extraction. Precipitate the RNA by adding 3 to 4 vol absolute ethanol, and then incubate at least 1 hr on ice.

Ligate the 5' adapter

28. Collect the RNA pellet as described in step 11. Dissolve the pellet in 10 µl RNase-free water.
29. Add 1 µl of 100 µM 5' adapter oligonucleotide (see Table 26.4.1), 2 µl of 10× RNA ligation buffer, and 6 µl of 50% aqueous DMSO.
30. Denature the RNA by incubating the tube for 30 sec at 90°C. Place the tube immediately on ice for 20 sec.
31. Add 1 µl of T4 RNA ligase, mix gently, and incubate for 1 hr at 37°C.
32. Add 1 vol of gel-loading solution and load the sample in 2 wells of a 20-well 15% acrylamide gel (15 × 17 × 0.08–cm, 30-ml gel volume). Load the 5' adapter–ligated marker sample next to the sample, but not directly adjacent to avoid contamination.
33. Run the gel for ~1 hr at 28 W using 0.5× TBE buffer until the bromophenol blue dye is ~2 cm from the bottom of the gel. Image the gel as described in steps 6 and 7, and excise the new ligation product.
34. Elute the ligation product from the gel slices in 300 µl of 0.3 M NaCl in siliconized tubes overnight at 4°C. Add 1 µl of 100 µM reverse transcription (RT) primer (see Table 26.4.1) during the elution as a carrier to facilitate the recovery of the ligation product.

35. Precipitate the RNA and RT primer by addition of 3 to 4 vol absolute ethanol as described in step 15.

Perform reverse transcription of the final ligation product

36. Collect the RNA/RT primer pellet as described in step 11. Dissolve the pellet in 5.6 μ l RNase-free water.
37. Denature the RNA by incubating the tube for 30 sec at 90°C.
38. Add 1.5 μ l of 0.1 M DTT, 3 μ l of 5 \times first-strand buffer, 4.2 μ l of 10 \times dNTP, and incubate 3 min at 50°C. Add 0.75 μ l of reverse transcriptase and incubate 30 min at 42°C.
39. To hydrolyze the RNA, add 40 μ l of 150 mM KOH/20 mM Tris base and incubate 10 min at 90°C.
40. Neutralize the solution by adding ~40 μ l of 150 mM HCl to obtain a pH between 7.0 and 9.5.

Near the end point of titration, monitor the pH by spotting 1 μ l of the neutralized DNA solution on pH indicator paper. Be aware that HCl is volatile and KOH solutions absorb carbon dioxide so that the pH of these solutions change over time. Prior to neutralization, test the concentration of the KOH/Tris base and HCl solutions by combining equal volumes and monitoring the pH. If the pH is not within the recommended range, the pH of the subsequent PCR reaction will be affected, therefore, risking failed amplification.

Perform first PCR amplification of the cDNA

41. Combine 10 μ l of the cDNA sample from the previous step with 10 μ l of 10 \times dNTP, 10 μ l of 10 \times PCR buffer, 1 μ l of 100 μ M first PCR 5' primer (see Table 26.4.1), 1 μ l of 100 μ M first PCR 3' primer (see Table 26.4.1), 67 μ l water, and 1 μ l of *Taq* polymerase.
42. Run a PCR using the following parameters:

20 cycles:	45 sec	94°C
	1 min, 25 sec	50°C
	1 min	72°C.

43. Run 6 μ l of the PCR product on a 2% standard agarose gel containing ethidium bromide (see UNIT 2.5A). Use a 25-bp DNA ladder as a size marker, the size of the PCR product should be ~70 bp.

At this point, a double-band ~50 and 70 bp is sometimes detected on the gel. The 50-bp band is likely derived from traces of ligation product of directly joined 3' and 5' adapter oligonucleotides without a small RNA insert. The larger PCR product can be purified away from this contaminant using a 3% low-melt NuSieve agarose gel in 1 \times TAE buffer. Excise the 70-bp band, and isolate the DNA from the gel as described in step 55. Dissolve the pelleted DNA in 50 μ l of 1 \times PCR buffer.

Perform second PCR amplification to enable concatamerization

44. Combine 1 μ l of the first PCR solution with 50 μ l of 10 \times dNTP, 50 μ l of 10 \times PCR buffer, 5 μ l of 100 μ M second PCR 5' primer (see Table 26.4.1), 5 μ l of 100 μ M second PCR 3' primer (see Table 26.4.1), 384 μ l water, and 5 μ l *Taq* polymerase. For PCR, divide the reaction mixture between five 500- μ l PCR tubes.
45. Run 10 cycles of PCR as described in step 42 and examine 6 μ l of the PCR reaction as described in step 43.

46. If the PCR product is detectable, combine all the PCR reactions in one 1.5-ml Eppendorf tube, add 30 μ l of 5 M NaCl, and perform one 500- μ l 25:24:1 phenol/chloroform/isoamyl alcohol extraction followed by a 500- μ l chloroform extraction. Precipitate the DNA by adding 2 vol absolute ethanol and incubate at least 1 hr on ice.

Avoid denaturation of the second PCR product at any time during the subsequent step. The PCR product consists of a pool of distinct small sequences flanked by constant primer sequences. Denaturation and renaturation of the pool sequences may result in imperfect rehybridization and formation of DNA duplexes with internal bulges. Bulged DNA duplexes are poor substrates for the subsequent restriction endonuclease digestion or later steps.

If the PCR product is not detectable after 10 cycles, run 2 or 3 additional cycles. Do not amplify the DNA outside the exponential amplification range because it may result in the formation of bulged DNA duplexes due to re-annealing of imperfect complement strands of denatured PCR products. The optimal number of PCR cycles can be determined in a pilot experiment. If a pilot experiment is included, prepare two times the volume of the PCR reaction. Use one half for the pilot experiment while keeping the other on ice. Remove small samples between 8 and 12 PCR cycles at end of each elongation step and compare the intensity of the PCR products by running the samples on a 2% agarose gel containing ethidium bromide.

Digest second PCR product with Ban I

47. Microcentrifuge the DNA pellet 10 min at maximum speed (16,000 \times g), 4°C. Remove the supernatant completely but do not allow the pellet to dry. Dissolve the pellet in 190 μ l of 1 \times Ban I buffer. Keep 2 μ l of undigested material for further gel analysis.

To avoid DNA denaturation, do not allow the pellet to dry and resuspend the pellet in buffer rather than in water.

48. Add 10 μ l of 20 U/ μ l Ban I endonuclease and incubate for 3 hr at 37°C. Verify that the restriction digestion is complete by running 2 μ l of the digested sample and 2 μ l of the undigested sample on a 2% standard agarose gel as described in step 43.
49. Stop the reaction by adding 12 μ l of 5 M NaCl and perform a 200- μ l 25:24:1 phenol/chloroform/isoamyl alcohol extraction and one 200- μ l chloroform extraction. Precipitate the DNA by adding 2 vol absolute ethanol and incubating for at least 1 hr on ice.

Concatamerize Ban I restriction enzyme–digested DNA

50. Collect the DNA pellet as described in step 47. Dissolve the pellet in 90 μ l of 1.1 \times T4 DNA ligation buffer.
51. Add 3 μ l of 100 μ M second PCR 5' primer and 3 μ l of 100 μ M second PCR 3' primer (see Table 26.4.1). Incubate 10 min at 65°C.

This step is important to prevent religation of the 12-bp Ban I–digested fragments in the concatamerization step. At 65°C, the 12-bp fragments, not long DNA, are denatured and upon cooling will hybridize in a competitive manner to the excess PCR primers. If religation of the 12-bp Ban I–digested fragment occurs, the concatamerization step will not work.

52. Add 4 μ l of T4 DNA ligase and incubate 5 hr at 22°C.
53. Analyze 2 μ l of the reaction to check for complete concatamerization on a 2% agarose gel. Also load a 100-bp DNA ladder as a size marker.

The concatamerization product appears as a fuzzy ladder ranging in size from 60 bp (the monomer) to 1 kb.

54. Purify 50 μ l of the concatamerization products on a 2% NuSieve low-melt agarose gel containing ethidium bromide in 1 \times TAE running buffer. Load 10 to 15 μ l of sample per well of a 60-ml gel. Run the gel for ~45 min at 80 V until the marker bands are sufficiently resolved.

To avoid melting of the NuSieve gel at a high current, limit the voltage for electrophoresis to 80 V. Use TAE buffer rather than TBE gel running buffer when extracting DNA from low-melting agarose to avoid possible borate precipitates during ethanol-precipitation of the DNA.

55. Visualize the DNA in the gel using a 360-nm UV transilluminator. Excise the DNA band between 400 and 800 bp with a razor blade or scalpel and transfer the gel slices into a pre-weighed 1.5-ml reaction tube.

If the gel slices weigh >250 mg, split the agarose band into two tubes.

56. Add at least 1 vol 0.4 M NaCl to obtain a final volume of 500 μ l. Incubate the tube 10 min at 70°C to melt the gel slices and add 500 μ l of 70°C-preheated, buffered water-saturated phenol, pH 7.8. Vortex the solutions vigorously and immediately separate the phases by microcentrifuging 5 min at maximum speed, room temperature.

The agarose should accumulate at the interphase.

57. Collect the aqueous upper phase and extract the aqueous phase once again with 500 μ l of 25:24:1 phenol/chloroform/isoamyl alcohol and then with 500 μ l chloroform. Precipitate the DNA by adding 2 vol absolute ethanol and incubating for at least 1 hr on ice.

It is important not to transfer any solidified agarose with the aqueous phase during phenol extraction. Be careful to perform centrifugation steps at room temperature and not in a refrigerated centrifuge to avoid rapid solidification of the agarose and entrapment of DNA within the agarose.

Instead of a phenol extraction, the DNA may be electroeluted from the agarose gel (see UNIT 2.6, Basic Protocol 1). Avoid commercial DNA isolation kits that denature the DNA during isolation (see annotation to step 46).

Tail the ends of concatamers for T-vector ligation

58. Collect the DNA pellet as described in step 47. Dissolve the pellet in 15 μ l of a PCR mix containing 1 \times dNTPs, 1 \times PCR buffer, and 0.15 μ l *Taq* polymerase.
59. Incubate the reaction mixture for 30 min at 72°C for enzymatic 3' tailing.
60. Perform TOPO TA cloning with a pCR2.1-TOPO vector as described by the manufacturer (Invitrogen; also see Chapter 1).

Screen colonies for concatamer inserts by PCR

61. Fill the desired number of wells of a 96-well thermocycler-compatible microtiter plate with 50 μ l of a PCR solution containing 5 μ l of 10 \times PCR buffer, 5 μ l of 10 \times dNTPs, 0.5 μ l of 100 μ M primer M13 (–20) F, 0.5 μ l of 100 μ M primer M13 R (see Table 26.4.1), and 0.5 μ l *Taq* polymerase.
62. Transfer individual white colonies to the wells of the microtiter plate filled with the PCR mixture. Pick the white colonies from the agar plate using a pipet tip and resuspend the colony by pipetting up and down. Run 25 cycles of PCR as described in step 42.
63. Analyze 6 μ l of the PCR products on a 2% standard agarose gel, using the 100-bp DNA ladder as a size marker.

64. Purify the remaining PCR with the QIAquick PCR purification kit according to the manufacturer's instructions. Submit the purified PCR product for automated sequencing using M13 (-20) F or M13 R primers.

**ALTERNATE
PROTOCOL**

**CLONING OF SMALL RNA USING THE ADENYLATED 3'-ADAPTER
OLIGODEOXYNUCLEOTIDE**

This protocol is adapted from the protocol of Lau et al. (2001), which is available on the Bartel laboratory Web site at <http://web.wi.mit.edu/bartel/pub>. An adenylylated 3'-adapter oligodeoxynucleotide can be joined directly to the isolated small RNAs without prior 5' dephosphorylation in the absence of ATP (Lau et al., 2001). Without ATP, circularization of the 5' phosphorylated small RNAs cannot occur. Although the ligation reaction with adenylylated adapters is simple, it requires chemical synthesis of the 5' phosphorimidazolide of adenosine (ImpA, Fig. 26.4.2C) followed by chemical adenylation of the 5' phosphate of the 3'-adapter oligodeoxynucleotide using ImpA.

The chemical synthesis reactions in this procedure have been adapted from Mukaiyama and Hashimoto (1971) and Lohrmann and Orgel (1978).

Additional Materials (also see *Basic Protocol*)

- Adenosine 5'-monophosphoric acid (5' AMP)
- Dimethylformamide (DMF)
- Triphenylphosphine
- 2,2'-Pipyridyldisulfide
- Imidazole
- Triethylamine
- Sodium perchlorate
- Acetone
- Anhydrous ethyl ether
- Thin layer chromatography (TLC) cellulose plates with a 254-nm fluorescence indicator (cellulose-F TLC; Merck)
- 10% (v/v) saturated (NH₄)₂ SO₄
- 80% ethanol
- 1 M MgCl₂
- 5× adenylylation buffer (ATP-free; see recipe)
- 1.25 mM phosphorylated 3'-adapter oligodeoxynucleotide (see *UNIT 3.10*)
- 10 μM ³²P-labeled 3'-adapter oligonucleotide (see *UNIT 3.10*)
- 30-ml Corex tubes
- Sorvall centrifuge with SS34 rotor (or equivalent)
- Vacuum drying oven
- Glass capillaries
- 254-nm UV light
- 13-ml polypropylene tube, 95 × 16.8-mm (Sarstedt)

NOTE: All reagents can be purchased from Sigma and should be at least 99% pure.

Synthesize ImpA

1. Dissolve 174 mg (0.5 mmol) of 5' AMP in 15 ml DMF. Keep a 50-μl aliquot for TLC analysis.
2. Dissolve 262 mg (1 mmol) of triphenylphosphine, 220 mg (1 mmol) of 2,2'-dipyridyldisulfide, and 170 mg (2.5 mmol) of imidazole in 15 ml DMF and 0.9 ml (2.5 mmol) of triethylamine.

3. Add the AMP solution (from step 1) dropwise to a vigorously stirred triphenylphosphine-containing solution (from step 2). Stir for another 1.5 hr at room temperature.

Initially, a precipitate appears, then the mixture will turn yellow-green and the precipitate eventually dissolves.

Purify ImpA

4. Precipitate the ImpA by adding the reaction mixture dropwise into a vigorously stirred solution of 1.1 g (9 mmol) sodium perchlorate, 115 ml acetone, and 55 ml anhydrous ethyl ether.
5. Let the precipitate settle to the bottom of the beaker for 1 hr and decant ~150 ml of the supernatant without perturbing the precipitate.

A large glass pipet connected to a pipetting aid may also be used to aspirate off the supernatant.

6. Once the volume has been reduced to ~20 ml, resuspend the precipitate with the residual supernatant and transfer the suspension to 30-ml Corex tubes. Transfer the residual precipitate by rinsing the beaker with small volumes (5 ml) of acetone. Collect the precipitate by centrifuging 10 min at $3000 \times g$ (5000 rpm with a Sorvall SS34 rotor), room temperature.
7. Remove the supernatant and wash the pellet two times by resuspending it with 10 ml acetone and then centrifuge 5 min at $3000 \times g$ (5000 rpm with a Sorvall SS34 rotor), room temperature.
8. Resuspend the pellet in 10 ml anhydrous ethyl ether and centrifuge 20 min at $3000 \times g$ (5000 rpm with a Sorvall SS34 rotor), room temperature. Dry the pellet overnight in a vacuum oven at 40°C.
9. Store the dried powder up to 3 weeks at -20°C protected from humidity.

The yield of ImpA is ~80 mg. The molecular weight of ImpA is 412.3 g/mol.

Perform quality control of synthesized ImpA

10. Soak cellulose-F TLC plates in 10% saturated aqueous $(\text{NH}_4)_2\text{SO}_4$ and dry the TLC plates in open air for 1 hr or blow dry.
11. Dissolve 1 mg ImpA in 50 μl water and spot the sample on pre-treated TLC plates with aliquots of the AMP solution and triphenylphosphine solution from steps 1 and 2, respectively, using glass capillaries. Develop the TLC by using 80% ethanol and visualize the spots under 254-nm UV light.

The AMP should have the lowest retention factor ($R_f = 0.26$), ImpA should run faster ($R_f = 0.4$), and the components of the triphenylphosphine solution all run near the solvent front.

Perform adenylation of 3' adapter oligodeoxynucleotide

12. Dissolve 9 mg of ImpA in 420 μl of water, add 14 μl of 1 M MgCl_2 , and 80 μl of 1.25 mM phosphorylated 3'-adapter oligodeoxynucleotide (see UNIT 3.10).

The final concentrations for the chemical adenylation reaction are 200 μM phosphorylated 3'-adapter oligodeoxynucleotide (see Table 26.4.1), 50 mM ImpA (mol. wt. 412.3 g/mol), and 25 mM MgCl_2 .

13. Incubate the reaction mixture for 3 hr at 50°C. For adenylation of 5'-radiolabeled 3'-adapter oligodeoxynucleotide, reduce the reaction volume four-fold and use 5 µl of 10 µM ³²P-labeled 3'-adapter oligonucleotide (see UNIT 3.10).
14. Stop the reaction by adding 1 vol of gel-loading solution. Prepare two 20% acrylamide gels (15 × 17 × 0.15-cm, 50-ml gel volume) using a 6-well comb. Load 125 µl per well and run the gels for ~1 to 1.5 hr at 28 W using 0.5× TBE buffer until the bromophenol blue dye reaches the bottom of the gel.
15. Dismantle the gels and wrap them in plastic wrap. Place them onto a fluorescence indicator-coated silica gel plate and visualize the oligodeoxynucleotides by 254-nm UV-shadowing (see UNIT 2.12).

The adenylated oligonucleotide migrates slower and above the unreacted oligodeoxynucleotide. The yield is ~10% to 20%.

16. Excise the reaction product and elute the adenylated oligonucleotide from the gel slices overnight at 4°C (see Basic Protocol, steps 7 to 10).

If the radioactive adenylation product is not detectable by UV shadowing, use phosphorimaging to detect the bands.

To accommodate the large volume of gel slices, it is recommended to elute and precipitate the non-radioactive adenylation product in 13-ml polypropylene tubes.

17. Precipitate the adenylated oligodeoxynucleotide by adding 3 vol absolute ethanol and incubating for at least 1 hr on ice.
18. Collect the pellet as described in Basic Protocol, steps 11 and 12. When using 13-ml polypropylene tubes, centrifuge 20 min at 17,000 × g (12,000 rpm with an SS34 rotor). Dissolve the pellet in 20 µl of water and determine the concentration by UV.

Ligate adenylated 3' adapter to small RNAs

19. Dissolve the small RNA pellet from Basic Protocol, step 11, in 10 µl water. Add 4 µl of 5× adenylate ligation buffer (ATP-free), 4 µl of 200 µM adenylated 3'-adapter oligodeoxynucleotide, and 1 µl radiolabeled adenylated 3'-adapter oligodeoxynucleotide.
20. Denature the RNA by incubating the tube 30 sec at 90°C. Place the tube immediately on ice for 20 sec.
21. Add 1 µl of T4 RNA ligase, mix gently, and incubate 1 hr at 37°C.
22. Add 1 vol gel-loading solution and load the sample in 2 wells of a 20-well 15% acrylamide gel (15 × 17 × 0.08-cm, 30-ml gel volume). Load the ligated marker sample next to the sample, but not directly adjacent to avoid contamination. Run the gel for ~1 hr at 28 W in 0.5× TBE buffer until the bromophenol blue dye is ~2 cm from the bottom of the gel.
23. Dismantle the gel and wrap the gel in plastic wrap to protect it from the environment. Record the position of the marker and the ligation products by phosphorimaging as described in Basic Protocol, steps 6 and 7. Excise the ligation product band as well as the ligated size markers. Elute the ligation products from the gel slices overnight at 4°C in 300 µl of 0.3 M NaCl in siliconized tubes.
24. Precipitate the RNA by adding 3 to 4 vol absolute ethanol as described in Basic Protocol, step 10 and proceed with Basic Protocol, step 28.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Acrylamide gel solution, 15% or 20%

1× TBE (APPENDIX 2)
7.5 M urea
15% or 20% 19:1 acrylamide/bisacrylamide
Store up to 1 year at 4°C

Adenylate ligation buffer, 5×

16.5 mM DTT
41.5% glycerol
250 mM HEPES-KOH, pH 8.3
50 mM MgCl₂
50 µg/ml acetylated bovine serum albumin (Ac-BSA, Sigma)
Store up to 1 year at –20°C

Gel-loading solution

0.1 ml 0.5 M EDTA, pH 8 (APPENDIX 2)
5 mg bromophenol blue
Bring up to 9.9 ml with deionized formamide
Store up to 1 year at room temperature

RNA ligation buffer, 10×

0.1 M MgCl₂
0.1 M 2-mercaptoethanol
0.5 M Tris·Cl, pH 7.6 (APPENDIX 2)
1 mg/ml acetylated bovine serum albumin (Ac-BSA, Sigma)
2 mM ATP
Store up to 1 year at –20°C

COMMENTARY

Background Information

The first glimpse on the importance of small RNAs was given by the identification of the *lin-4* gene product (Lee et al., 1993). Mutations of this gene caused developmental defects in the nematode *Caenorhabditis elegans*. Several years later, a second 22-nt RNA, the *let-7* gene product, was also found to regulate development in *C. elegans* (Reinhart et al., 2000). These RNAs were therefore called small temporal RNAs (stRNAs) and were regarded as specific to nematode development until it was realized that *let-7* RNA was conserved in all animals with a bilateral symmetrical axis (Pasquinelli et al., 2000). Around the same time, it was also discovered that dsRNA-triggered RNA interference was guided by 21- and 22-nt small interfering RNAs (siRNAs) that mediate the target RNA degradation (Hamilton and Baulcombe, 1999; Elbashir et al., 2001). These discoveries indicated the need for the development

of methods to characterize small RNA profiles of eukaryotic cells.

The RNA cloning protocol (see Basic Protocol) was originally developed to study Dicer ribonuclease III-dependent processing patterns of long dsRNAs (30 to 500 bp) to the siRNAs in cytoplasmic lysates prepared from *D. melanogaster* embryos (Tuschl et al., 1999; Elbashir et al., 2001). When analyzing the cloned sequences, high levels of cellular 21-nt RNAs were identified. These endogenous RNAs were derived from numerous cellular genes coding for short hairpin RNAs, which represented the dsRNA precursors of miRNAs (Lagos-Quintana et al., 2001). The authors' original cloning protocol was gradually improved, incorporating modifications from the protocol used in David Bartel's laboratory (Lau et al., 2001; available at <http://web.wi.mit.edu/bartel/pub/>). The protocol is now routinely used to characterize the miRNA profile of dif-

```

cagnaacagctatgaccatgattacgccaagcttggtagcgagctcggatccactagt
aacgcccgccagtgtgctgTAACCGGAATTCGGCACCGAATTCCTCACTAAAAAAGG
CGTTGGTTGCTTATTTAACCGGAATTCGGCACCGAATTCCTCACTAAAAUGGGAACAC
CGCGUGUUGUUGGCTTTAACCGGAATTCGGCACCGAATTCCTCACTAAAAATAAGCTA
GACAACCATTTGATTTAACCGGAATTCGGCACCGAATTCCTCACTAAAAATCGGAAGGTT
GCGTGTTCCTTTAACCGGAATTCGGCACCGAATTCCTCACTAAACTGAGGTGCGGTTAT
TGCTATTTAACCGGAATTCGGCACCGAATTCCTCACTAAATATCACAGCCAGCTTTGA
GGAGCTTTAACCGGAATTCGGCACCGAATTCCTCACTAAAGGGAACACCGCGTGTGT
TGGCTTTAACCGGAATTCGGCACaagctaḡ caagggggaattctgcagatatccat
cacactggcgccgctcgagcatgcatctagagggccaattcgccgatagtgagtagc
gtatt

```

The following small RNA sequences are obtained from this clone:

```

AAAAGGCGUUGGUUGCUUA, rRNA
AUGGGAACACCGCGUGUUGUUGGC, rRNA
AUAAAGCUAGACAACCAUUGA, miR-4
AUCGGAAGGUUGCGUGUUC, tRNA
CUGAGGUGCGGUUAUUGCUA, snRNA
UAUCACAGCCAGCUUUGAGGAGC, miR-2b
GGGAACACCGCGUGUUGUUGGC, rRNA

```

Figure 26.4.3 Example of a concatamer sequence after cloning of small RNAs from *Drosophila* embryos. The lowercase letters represent the vector sequence and the uppercase letters represent concatamer sequence. Constant sequences of the adapter oligonucleotides are highlighted (dark gray, 3' adapter and light gray, 5' adapter). The *Ban* I restriction site is underlined.

ferent cell lines or tissues and requires between 1 and 2 weeks beginning with the isolation of total RNA until obtaining the sequence information (Lagos-Quintana et al., 2002, 2003).

Critical Parameters and Troubleshooting

One of the critical starting points of this protocol is high quality of the total RNA preparation. If the RNA has been prepared from cells that were not healthy or contained a high fraction of apoptotic cells, the cloning procedure will identify mostly ribosomal RNA and tRNA degradation products. It is therefore recommended to harvest the cells during their exponential growth phase. Some tissues, like pancreas or small intestine, are rich in nucleases and intrinsically difficult to handle. It is recommended that the quality of the RNA preparation be carefully examined before making the decision to start a lengthy cloning protocol. High-

quality total RNA shows defined 18S and 28S rRNA bands by agarose gel electrophoresis. Alternatively, a more modern RNA LabChip technique may be used for analyzing the quality of the total RNA. LabChip technologies (e.g., Agilent 2100 Bioanalyzer) were developed for quality control of total RNA to be used for gene profile analysis on DNA arrays.

Another difficulty is the handling of extremely low amounts of RNA products making it necessary to use surface-treated, siliconized reaction tubes. Extreme care is needed to avoid nuclease contamination throughout the protocol.

Drosophila or plant small RNAs are composed of two distinct classes of ~21 and 25 nt in length. The longer species of small RNAs has not been detected in mammals, probably because only one Dicer RNase III exists in mammals while the other species contains multiple Dicer RNase III homologs. If the longer

species of small RNA is also to be isolated, it is recommended that a 30-nt RNA size marker be included. *Drosophila* is also special because its rRNA contains an abundant 30-nt 2S rRNA not found in most other organisms. The 2S rRNA is readily detected by UV-shadowing and care has to be taken to avoid excising the small RNA band with an overlap to the 2S rRNA band.

Finally, although it is feasible to analyze the cloned sequences by hand, it is recommended that the assistance of an experienced bioinformatics specialist be sought to develop adequate analysis and archiving tools. The small RNA fractions often contain substantial amounts of degradation products of other abundant non-coding RNAs such as rRNAs, tRNAs, snRNAs, and snoRNAs, as well as messenger RNAs. A guideline to the annotation of newly identified miRNA genes has been published (Ambros et al., 2003) and a repository for miRNAs can be found at <http://www.sanger.ac.uk/Software/Rfam/index.shtml>.

Anticipated Results

Each clone carrying an insert of a concatamerized PCR product should yield 4 to 8 small RNA sequences. An example of the raw data obtained after sequencing is shown in Figure 26.4.3.

Time Considerations

The Basic Protocol is a multi-step procedure, similar to an organic synthesis of a complex molecule. If any of the many steps fails, the final product will never be obtained. See Figure 26.4.1 for a flow chart of all steps.

Isolation of total RNA, radiolabeling, and gel purification of all oligonucleotides require 1 day. Gel-purification of small RNAs from total RNA requires 1 day with an overnight elution from the gel. Dephosphorylation of small RNAs and ligation of phosphorylated 3' adapter to dephosphorylated small RNAs require 1 day with an overnight elution from the gel. 5' phosphorylation of the ligation product and ligation of the 5' adapter takes 1 day with an overnight elution from the gel. Reverse transcription of the final ligation product, PCR amplification of the cDNA, and second PCR amplification to enable concatamerization all require 1 day. *Ban* I restriction digestion of the second PCR product and concatamerization of *Ban* I restriction-digested DNA require 1 day. Tailing of the ends of concatamers for T-vector ligation and TOPO/TA cloning take 1 day. PCR screening of colonies for concatamer inserts and sequencing of positive clones require 1 day.

The protocol can be safely interrupted at every ethanol precipitation step and the ethanol solution can be stored at -20°C for several days prior to collecting the pellet.

Acknowledgements

The authors thank Heiko Manninga, Alexei Aravin, Gunter Meister, Markus Landthaler, Matthew Poy, and Yair Dorsett for critically reading the manuscript.

Literature Cited

- Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S., Marshall, M., Matzke, M., Ruvkun, G., and Tuschl, T. 2003. A uniform system for microRNA annotation. *RNA* 9:277-279.
- Elbashir, S.M., Lendeckel, W., and Tuschl, T. 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* 15:188-200.
- Hamilton, A.J. and Baulcombe, D.C. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286:950-952.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. 2001. Identification of novel genes coding for small expressed RNAs. *Science* 294:853-858.
- Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., and Tuschl, T. 2002. Identification of tissue-specific microRNAs from mouse. *Curr. Biol.* 12:735-739.
- Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A., and Tuschl, T. 2003. New microRNAs from mouse and human. *RNA* 9:175-179.
- Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. 2001. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294:858-862.
- Lee, R.C., Feinbaum, R.L., and Ambros, V. 1993. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75:843-854.
- Lohmann, R. and Orgel, L.E. 1978. Preferential formation of (2'-5')-linked internucleotide bonds in non-enzymatic reactions. *Tetrahedron* 34:853-855.
- Mukaiyama, T. and Hashimoto, M. 1971. Phosphorylation by oxidation-reduction condensation: Preparation of active phosphorylating reagents. *Bull. Chem. Soc. Japan* 44:2284.
- Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degan, B., Muller, P., Spring, J., Srinivasan, A., Fishman, M., Finnerty, J., Corbo, J., Levine, M., Leahy, P., Davidson, E., and Ruvkun, G. 2000. Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA. *Nature* 408:86-89.

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. 2000. The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403:901-906.

Tuschl, T., Zamore, P.D., Lehmann, R., Bartel, D.P., and Sharp, P.A. 1999. Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev.* 13:3191-3197.

Contributed by Sébastien Pfeffer, Mariana Lagos-Quintana, and Thomas Tuschl
The Rockefeller University
New York, New York

RNA Interference in Cultured *Drosophila* Cells

UNIT 26.5

RNA interference (RNAi) resulting in the ablation of specific proteins in *Drosophila* cell culture is a valuable tool for dissecting signal transduction pathways and elucidating the cellular functions of proteins. Since the early 1960s, several *Drosophila* cell lines representing different cell lineages have been propagated (Table 26.5.1; Schneider, 1972). Because many signal transduction pathways and the functions of conserved proteins have been maintained from *Drosophila* to mammals, the ablation of protein expression in these cell lines makes it possible to study conserved biochemical processes in a relatively simple environment. Furthermore, because *Drosophila* is amenable both to gene ablation studies and RNAi analyses, results obtained from cell culture studies can be confirmed in the whole organism. Finally, the use of double-stranded RNA (dsRNA) technology in *Drosophila* cell culture is simple, highly reproducible, and efficient. Transfection is unnecessary, because dsRNA is naturally internalized by the cells, thereby circumventing the problems associated with transfection protocols, such as variable efficiency and toxicity. The gene silencing effect is also sustainable through many cell divisions, provided the protein of interest is not necessary for viability (Clemens et al., 2000).

The ablation of protein expression in *Drosophila* cell culture is achieved by exposing the cells to dsRNA ranging from 300 to 700 bp. The efficacy of this process is best measured by immunoblot using antibodies directed against the target protein of interest. However, in the absence of available antibodies, Northern, RT-PCR, or RNase protection analyses can be employed to measure the amount of mRNA remaining. This unit contains a technique for treatment of *Drosophila* cell lines with dsRNA (see Basic Protocol). In addition, an efficient method for generating dsRNA synthesis templates that harbor T7 RNA polymerase promoters on their 5'-ends is described (see Support Protocol 1) and the procedure used for dsRNA synthesis is also provided (see Support Protocol 2).

NOTE: Care must be taken to use only RNase-free solutions and materials. Gloves should be worn to prevent RNase contamination.

TREATMENT OF *DROSOPHILA* CELLS WITH dsRNA

The advantage of using *Drosophila* cells for protein ablation studies is underscored by their ability to take up dsRNA without the use of transfection reagents. This ability is manifested by all of the *Drosophila* cell lines listed in Table 26.5.1. Furthermore, *Drosophila* cells do not require a CO₂ incubator and some can be grown in suspension cultures as well as in the more traditional tissue culture flasks. Note that although the protocol described below is optimized for S2 cells, no major modifications are necessary when using other cell lines.

Materials

Drosophila cells (e.g., S2 cells; Table 26.5.1)
S2 growth medium (see recipe)
Schneider's *Drosophila* medium (Life Technologies)
dsRNA (see Support Protocol 2)
6-well tissue culture dishes

Additional reagents and equipment for counting cells (APPENDIX 3F), assaying RNA depletion—i.e., Northern blot (UNIT 4.9), RNase protection assay (UNIT 4.7), or RT-PCR (UNIT 15.5)—and immunoblotting (UNIT 10.8)

**BASIC
PROTOCOL**

Gene Silencing

26.5.1

Table 26.5.1 *Drosophila* Cell Lines

Line	Origin	Reference	Characteristics
S2	Oregon R, dissociated embryos	Schneider (1972)	Round, semi-adherent, phagocytic, granular cytoplasm
S2-R+	Oregon R, dissociated embryos	Yanagawa et al. (1998)	Flat, adherent, phagocytic, hemocyte-like gene expression
Kc	F2 ebony × sepia, dissociated embryos	Echalier and Ohanessian (1969)	Round, clump in sheets, semiadherent, phagocytic, hemocyte-like gene expression
BG2-c6	Third instar larvae, central nervous system	Ui-Tei et al. (1994)	Semiadherent, neuronal-like processes
Clone 8	Third instar larvae, wing imaginal disc	Peel et al. (1990)	Columnar epithelial, adherent

Culture cells

1. Propagate S2 cells in S2 growth medium containing antibiotics. Passage cells at a 1:20 dilution into a new flask once they achieve a cell density $>6 \times 10^6$ cells/ml.

S2 cells that are attached to the flask are easily displaced by trituration. These cells grow well as long as they are not subcultured to a very low density (i.e., $<1 \times 10^5$ cells/ml). They are also sensitive to different sources of FBS. Life Technologies FBS designated for neuronal cells consistently produces good results; however, each new lot should be tested to ensure growth of S2 cells.

2. Count the cells (APPENDIX 3F). Pellet a sufficient number for each experimental condition (i.e., 1×10^6 cells) and an untreated control by centrifuging 10 min at $500 \times g$, 20°C. Resuspend in Schneider's *Drosophila* medium to a density of 1×10^6 cells/ml.

No serum or antibiotics should be present in the medium.

3. Seed 1 ml cells into each well of a 6-well tissue culture dish.

Plating S2 cells in serum-free media will cause them to adhere tightly to the dish.

Treat with dsRNA

4. Add 15 µg double-stranded (ds)RNA (see Support Protocol 2) and hand mix the plates ~10 sec.
5. Incubate 30 to 60 min at room temperature.
6. Add 2 ml S2 growth medium to each well.
7. Incubate the cells 2 to 6 days at room temperature to allow turnover of the targeted protein.
8. Measure depletion of the target RNA after 24 to 48 hr by Northern blot (UNIT 4.9), RNase protection assay (UNIT 4.7), or RT-PCR (UNIT 15.5).
9. If possible, measure ablation of protein expression by immunoblot analysis (UNIT 10.8).

The time course of elimination depends upon protein half-life.

PREPARATION OF PCR TEMPLATES

cDNA is used to generate a 300- to 700-bp PCR template that harbors a T7 RNA polymerase consensus binding site (5'-TTAATACGACTCACTATAGGGGAGA-3') at its 5'-termini. A discussion of where to obtain the cDNA of interest is provided in the Commentary section of this unit (see Critical Parameters). The PCR template may include coding, 5'-, or 3'-untranslated sequences. If starting with genomic DNA, care should be taken not to incorporate intervening sequences into the PCR product.

Materials

cDNA encoding the protein targeted for ablation (see Critical Parameters)
DNA polymerase, 1× buffer, and Mg²⁺ stock (if required): Applied Biosystems
Ampli*Taq* Gold (UNIT 3.5) or NEB Vent DNA polymerase
10 mM (each) dNTPs (dATP, dCTP, dGTP, dTTP; UNIT 15.1)
High Pure PCR Purification kit (Roche Molecular Biochemicals)
DEPC-treated H₂O (UNIT 4.1)

Additional reagents and equipment for purification of DNA (UNITS 2.1A & 2.1B;
optional), and agarose gel electrophoresis and ethidium bromide staining (UNIT
2.5A)

1. Design two oligonucleotides, forward and reverse, each containing the T7 RNA polymerase consensus binding sequence at their 5'-termini.

The melting temperature of these oligonucleotides should be between 52° and 58°C. A rough estimate of the melting temperature can be calculated by using 4°C for each C/G base pair and 2°C for each A/T base pair. See the Commentaries of UNITS 2.10 & 15.2 for additional discussions of melting temperature calculations.

2. Set up a typical 100-μl PCR reaction as follows:

100 ng cDNA encoding targeted protein
250 ng each oligonucleotide primer
1× polymerase buffer
0.2 mM dNTPs from a 10 mM stock
5 U DNA polymerase (e.g., Ampli*Taq* Gold or Vent)
Mg²⁺ stock (if required)
H₂O to 100 μl.

The optimal Mg²⁺ concentration must be empirically determined for individual PCR reactions (see UNIT 15.1).

3. Purify the completed PCR reaction using the High Pure PCR Product Purification kit according to manufacturer's instructions or using the procedures described in UNITS 2.1A & 2.1B.
4. Elute the PCR product from the column with 50 μl DEPC-treated H₂O.
The DNA must be at a concentration of ≥125 ng/μl to proceed to the RNA synthesis step (see Support Protocol 2).
5. Visualize the PCR product by agarose gel electrophoresis and ethidium bromide staining to determine both its size and purity (UNIT 2.5A).

PREPARATION OF DOUBLE-STRANDED RNA

The Ambion MEGAscript T7 kit is used to synthesize dsRNA from the PCR template generated above. One 20-μl reaction typically yields between 50 to 100 μg of dsRNA 700 bp in length. Properly stored, this dsRNA preparation will remain efficacious for at least two years.

Materials

MEGAscript T7 kit (Ambion):
T7 enzyme mix
10× T7 reaction buffer
75 mM ATP, CTP, GTP, and UTP solutions
Nuclease-free H₂O
DNA, purified (see Support Protocol 1)
DEPC-treated H₂O (UNIT 4.1)
3 M ammonium acetate, pH 4.0
100% ethanol

Additional reagents and equipment for determining nucleic acid concentration (APPENDIX 3D) and visualization of dsRNA by agarose gel electrophoresis and ethidium bromide staining (UNIT 2.5A)

1. With the exception of the enzyme mix, thaw the components of the MEGAscript T7 kit at room temperature.
2. Determine the concentration of purified DNA (APPENDIX 3D). Calculate the volume DNA required to yield 1 μg (y). Determine the amount of DEPC-water required to prepare a 20-μl reaction:

Vol. DEPC-treated water (x) = 8 μl – vol. DNA to yield 1 μg (y).

3. Prepare the following reaction, adding the components in order:

x μl DEPC-treated H₂O
2 μl 10× T7 reaction buffer
2 μl each ATP, GTP, CTP, and UTP solution
y μl DNA template (1 μg DNA)
2 μl enzyme mix.

Hand mix the tube gently. Incubate the reaction 2 to 6 hr at 37°C.

4. Add an equal volume (i.e., 20 μl) nuclease-free H₂O. Precipitate the RNA by adding 1/10 vol of 3 M ammonium acetate, pH 4.0, and 2.5 vol of 100% ethanol, and incubating ≥15 min at –20°C.

Template removal is unnecessary (see Critical Parameters).

5. Pellet the RNA by microcentrifuging 15 min at maximum speed, 4°C. Remove the supernatant and air dry the pellet 15 min.
6. Resuspend the RNA pellet in 50 μl DEPC-treated H₂O and determine the concentration (APPENDIX 3D).
7. *Optional:* Anneal the RNA strands by incubating 30 min at 65°C in a beaker of water. Remove beaker from heat to the bench top and allow to cool slowly (i.e., ~60 min) to room temperature.
8. Dilute the sample to a final concentration of 3 μg/μl.

Preparations are stable for a minimum of 24 months at –20°C with no loss of efficacy.

9. Visualize the dsRNA on a 1% agarose gel to check the integrity and size (UNIT 2.5A).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

S2 growth medium

Prepare Schneider's *Drosophila* medium (Life Technologies) containing 10% (v/v) FBS for neuronal cells (Life Technologies), 50 U/ml penicillin (*UNIT 1.4*), and 50 mg/ml streptomycin (*UNIT 1.4*). Store up to 2 months at 4°C.

Each new lot of FBS should be tested to ensure growth of *S2* cells.

COMMENTARY

Background Information

Double-stranded RNA (dsRNA)-mediated interference (RNAi) of gene expression has become a widely used method for reducing or ablating the synthesis of specific proteins in a variety of systems, ranging from whole animals (i.e., *C. elegans* and *Drosophila melanogaster*) to mammalian and insect cell cultures (*UNIT 25.1*; Fire et al., 1998; Kennerdell and Carthew, 1998; Ngo et al., 1998; Alvarado and Newmark, 1999). Discovery of this phenomenon originated from antisense RNA experiments, where it was noted that both sense and antisense RNAs were effective at producing the desired reduction of protein expression (Montgomery and Fire, 1998). These studies paved the way for the discovery that if complimentary strands of a specific RNA were combined, producing dsRNA, interference of protein expression was facilitated (Fire et al., 1998).

RNAi was originally used for performing reverse genetic studies in *C. elegans* (Fire et al., 1998); however, this technique was soon adopted for studies in *Drosophila* (Kennerdell and Carthew, 1998; Misquitta and Paterson, 1999) and, with some major modifications, mammalian cells (Caplen et al., 2001; Elbashir et al., 2001a). Detailed descriptions of experiments in mammalian cells and worms are found in *UNITS 26.2 & 26.3*, respectively. In *Drosophila*, injection of embryos with dsRNAs before the syncytial blastoderm stage successfully mimics previously characterized loss-of-function embryonic mutations (Kennerdell and Carthew, 1998; Misquitta and Paterson, 1999). Because embryos at this developmental stage have no cell membranes, the dsRNAs are free to diffuse throughout the embryo and become incorporated into newly forming cells. The RNAi effect persists throughout development and can be observed in the adult at low penetrance. However, unlike the case for *C. elegans*, the effect is not transmissible to progeny.

Studies in *Drosophila* cell-free systems have led to a better understanding of the mechanism of action of RNAi (Tuschl et al., 1999; Zamore et al., 2000; Elbashir et al., 2001b). Briefly, dsRNA molecules are recognized by dicer, an RNaseIII-like nuclease. Dicer cleaves the dsRNAs into ~22-bp pieces that are subsequently bound by a multiprotein complex, the RNA-induced silencing complex (RISC; Bernstein et al., 2001; Ketting et al., 2001; Knight and Bass, 2001). The RISC complex possesses helicase and nuclease activity and targets homologous sequences in mRNAs for nucleolytic attack, thereby destroying them. For more information about dicer and RISC, refer to *UNIT 26.1*.

Critical Parameters

Since dsRNA >300 bp (preferably 500 to 700 bp) is used for *Drosophila* RNAi, the selection of the dsRNA sequence is not as critical as for mammalian RNAi. Presumably, the 300-bp dsRNA will be processed into many different 22-bp pieces, several of which may be good templates for directing the attack by the RISC complex. The authors commonly prepare dsRNA from coding sequences, although 5'- and 3'-nontranslated sequences also work. However, if starting from genomic DNA, care should be taken to avoid the inclusion of intervening sequences.

Obtaining cDNA clones

Although templates for dsRNA synthesis can be generated by PCR from genomic sequences or by RT-PCR from mRNA, the authors prefer to start with ESTs (*UNIT 19.3*) provided by The Berkeley *Drosophila* Genome Project (BDGP). The BDGP has released two sets of cDNAs comprising most if not all of the predicted protein products in *Drosophila melanogaster*. These clones can be obtained from the BDGP or from Open Biosystems as bacterial stocks (see Internet Resources).

dsRNA preparation

The authors use the T7 MEGAscript kit from Ambion. They feel it is a superior product and have found the technical resources provided by Ambion to be superb. The authors find it unnecessary to remove the DNA template nor do they use any kits to clean up the reaction prior to use. However, precipitation of the dsRNA and resuspension in nuclease-free water is recommended. When treated in this manner, the dsRNA can be stored >24 months at -20°C with no loss of efficacy. It can also withstand several cycles of freezing and thawing.

Troubleshooting

Since the dsRNA used in *Drosophila* cell culture is ideally ~500 to 700 bp, target selection is not an issue. Also, since *Drosophila* cells take up the dsRNA without the need for transfection reagents, all the caveats associated with transfection are avoided. The most common problems encountered in the authors' laboratory are (1) the lack of commercial antibodies available to check the level of protein expression after dsRNA treatment and (2) that primary neuronal cultures are recalcitrant to RNAi. The first problem requires that one either produce antibodies against the protein of interest (UNIT 11.12) or determine the level of the specific mRNA remaining by Northern analysis (UNIT 4.9), RNase protection (UNIT 4.7), or RT-PCR (UNIT 15.5). It should be noted that disappearance of the requisite mRNA does not guarantee ablation of the protein, as the half-life of the protein must be taken into consideration. The second problem has not been resolved at the time of this writing.

Anticipated Results

Protein expression is generally decreased by at least 95% by 72 hr after treatment, and the effect can persist up to two weeks. Addition of more dsRNA can extend this effect indefinitely, provided that the protein of interest is not necessary to the growth and survival of the cells.

Time Considerations

S2 cells can be plated and treated in ~1 hr. The expression of the protein can then be measured after 48 to 72 hr. Routinely, the authors' experiments take 3 days, but proteins with longer half-lives may require longer incubation times.

Literature Cited

- Alvarado, A.S. and Newmark, P.A. 1999. Double-stranded RNA specifically disrupts gene expression during planarian regeneration. *Proc. Natl. Acad. Sci. U.S.A.* 96:5049-5054.
- Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409:363-366.
- Caplen, N.J., Parrish, S., Imani, F., Fire, A., and Morgan, R.A. 2001. Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc. Natl. Acad. Sci. U.S.A.* 98:9742-9747.
- Clemens, J.C., Worby, C.A., Simonson-Leff, N., Muda, M., Maehama, T., Hemmings, B.A., and Dixon, J.E. 2000. Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. *Proc. Natl. Acad. Sci. U.S.A.* 97:6499-6503.
- Echalier, G. and Ohanessian, A. 1969. Isolation, in tissue culture, of *Drosophila melanogaster* cell lines. *C.R. Acad. Sci. Hebd. Seances Acad. Sci. D* 268:1771-1773.
- Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. 2001a. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411:494-498.
- Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. 2001b. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J.* 20:6877-6888.
- Fire, A., Xu, S.Q., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391:806-811.
- Kennerdell, J.R. and Carthew, R.W. 1998. Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. *Cell* 95:1017-1026.
- Ketting, R.F., Fischer, S.E.J., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H.A. 2001. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* 15:2654-2659.
- Knight, S.W. and Bass, B.L. 2001. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* 293:2269-2271.
- Misquitta, L. and Paterson, B.M. 1999. Targeted disruption of gene function in *Drosophila* by RNA interference (RNA-i): A role for nautilus in embryonic somatic muscle formation. *Proc. Natl. Acad. Sci. U.S.A.* 96:1451-1456.
- Montgomery, M.K. and Fire, A. 1998. Double-stranded RNA as a mediator in sequence-specific genetic silencing and co-suppression. *Trends Genet.* 14:255-258.

- Ngo, H., Tschudi, C., Gull, K., and Ullu, E. 1998. Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc. Natl. Acad. Sci. U.S.A.* 95:14687-14692.
- Peel, D.J., Johnson, S.A., and Milner, M.J. 1990. The ultrastructure of imaginal disc cells in primary cultures and during cell aggregation in continuous cell lines. *Tissue Cell.* 22:749-758.
- Schneider, I. 1972. Cell lines derived from late embryonic stages of *Drosophila melanogaster*. *J. Embryol. Exp. Morphol.* 27:353-65.
- Tuschl, T., Zamore, P.D., Lehmann, R., Bartel, D.P., and Sharp, P.A. 1999. Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev.* 13:3191-3197.
- Ui-Tei, K., Nishihara, S., Sakuma, M., Togashi, S., Ueda, R., Miyata, Y., and Miyake, T. 1994. Newly established cell lines from *Drosophila* larval CNS express neural specific characteristics. *In Vitro Cell Dev. Biol. Anim.* 30A:209-216.
- Yanagawa, S., Lee, J.S., and Ishimoto, A. 1998. Identification and characterization of a novel line of *Drosophila* Schneider S2 cells that respond to wingless signaling. *J. Biol. Chem.* 273:32353-32359.
- Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. 2000. RNAi: Double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* 101:25-33.

Key Reference

Clemens, et al., 2000. See above.

A concise description of performing RNAi experiments in Drosophila tissue culture cells is contained in this reference.

Internet Resources

http://stke.sciencemag.org/cgi/content/full/OC_sigtrans;2001/95/pl1?ijkey=d2gOuU9T1u.7k&keytype=ref&siteid=sigtrans

This site provides experimental protocols for using dsRNA in Drosophila cell culture; however, the user must have a subscription to Science magazine to access it.

<http://flyrnai.org>

This site provides a description of the Drosophila RNAi Screening Center (DRSC), a facility specializing in genome wide RNAi screening. Protocols for using dsRNA in Drosophila cell culture and results from RNAi screens can also be found at this site.

<http://www.fruitfly.org>

The Berkeley Drosophila genome project (BDGP) is a consortium of the Drosophila Genome Center, funded by the National Human Genome Research Institute, National Cancer Institute, and Howard Hughes Medical Institute, supporting the work in the Gerald Rubin and Allan Spradling laboratories. Drosophila expressed sequence tags (ESTs), can be obtained from this site.

http://www.openbiosystems.com/drosophila_rnai_collection.php

Drosophila ESTs can be obtained from this company. In addition, PCR templates containing T7 promoter sequences at the 5'-ends are purchasable.

<http://www.ambion.com>

Informative Web site for RNAi experiments.

Contributed by Carolyn A. Worby and
Jack E. Dixon
University of California at San Diego
La Jolla, California

RNA silencing in plants is routinely used for assignment of gene function. It is increasingly popular because it is often quicker than genetic approaches. In plants, as in other organisms, silencing requires production of small interfering RNA (siRNA) corresponding to the target gene. Most investigators achieve this by constructing stable transgenic plants in which the transgene constructs generate double-stranded (ds) RNA, either directly or indirectly (Basic Protocol 1). The dsRNA is then processed to siRNA *in vivo* by the endogenous enzyme Dicer. Viruses can also be used to deliver the silencer sequences using a transient silencing protocol (Basic Protocol 2). Both of these basic protocols use *Agrobacterium tumefaciens* to deliver transgene constructs into plant hosts. This unit discusses practical aspects of various approaches to RNA silencing in plants. Related practical issues are also addressed elsewhere (Lu et al., 2003b). Basic Protocol 3 in this unit describes protein overexpression using viral proteins that suppress silencing.

CONSTRUCTING STABLE TRANSGENIC PLANTS THAT EXPRESS dsRNA

BASIC PROTOCOL 1

This protocol describes the application of RNAi in transgenic *Arabidopsis* plants, a typical application of RNAi technology. Introduction of dsRNA into plants by binary vectors carrying inverted cDNA repeats has been widely used to generate stable RNAi lines (i.e., stable transgenic plants) to study the functions of specific genes (Waterhouse et al., 1998; Chuang and Meyerowitz, 2000; Dalmay et al., 2000; Wesley et al., 2001; Beclin et al., 2002; Chen et al., 2003; Guo et al., 2003). Several vector systems are available to clone inverted cDNA repeats so that they can subsequently be used for *Agrobacterium tumefaciens*-mediated transformation (Table 26.6.1). As can be gleaned from Table 26.6.1, both constitutive and inducible promoters can be used. In addition, there are different methods to clone the inverted repeat (restriction enzymes or the "Gateway" recombination system). Different vectors make use of various introns inserted between the inverted repeats to serve as stuffer sequences.

In this protocol, the pHANNIBAL system is used as an example to describe the cloning of the RNAi construct and transformation of the construct into plants (Wesley et al., 2001). The pHANNIBAL vector contains a strong constitutive cauliflower mosaic virus 35S (CaMV 35S) promoter, two sets of multiple cloning sites flanking an intron from the pyruvate orthophosphate dikinase (*PDK*) gene, and the *A. tumefaciens* octopine synthase (*OCS*) terminator (Fig. 26.6.1). Sense and antisense cDNA fragments from target genes of interest are cloned into the two multiple cloning sites to form an inverted repeat. The cassette can be released by cutting at the flanking *NotI* sites and cloned into the unique *NotI* site of the *A. tumefaciens* binary vector pART27 (Gleave, 1992). The pART27 plasmid carries both the right and left T-DNA borders to integrate the flanking sequence into the plant genome and the *NPTII* gene under the control of the nopaline synthase (*NOS*) promoter and terminator to permit selection of transgenic plants using kanamycin. The resulting RNAi construct in pART27 is introduced into *A. tumefaciens* strain GV3101, which is then used to transform *Arabidopsis* plants using the floral dip method (Clough and Bent, 1998). Transgenic plants carrying the RNAi are identified by selection using kanamycin. Finally, target-gene expression is analyzed by RT-PCR (reverse transcription PCR) and northern blotting and/or immunoblotting to identify transgenic plants that exhibit an appropriate reduction in target-gene expression.

Gene Silencing

26.6.1

Contributed by Yanhai Yin, Joanne Chory, and David Baulcombe

Current Protocols in Molecular Biology (2005) 26.6.1-26.6.19

Copyright © 2005 by John Wiley & Sons, Inc.

Table 26.6.1 Vectors for RNA Silencing in Transgenic Plants^a

Vector system	Promoter	Plant selection marker	Assembly strategy	Stuffer	References/web sites/comments
ChromDB	CaMV 35S, long and short versions	<i>hyg, basta, kan</i>	Two-step ligation	GUS, CHS _a intron, rice waxy intron	http://www.chromdb.org/ A useful selection of <i>Agrobacterium</i> binary plasmid vectors with a range of plant selectable marker genes for constitutive silencing.
Hellsgate	CaMV 35S	<i>kan</i>	attP, attR	PDK intron	http://www.pi.csiro.au/tech_licensing_bioll/Main_tech_licensing_biol.htm Wesley et al. (2001) <i>Agrobacterium</i> binary plasmid vectors suitable for high-throughput applications using BP and LR. Clonase enzymes available from Invitrogen, with detailed descriptions and methods for using the Gateway cloning system from Invitrogen (http://www.invitrogen.com).
Hannibal/Kannibal	CaMV 35S	<i>kan</i>	Two-step ligation	PDK intron	http://www.pi.csiro.au/tech_licensing_bioll/Main_tech_licensing_biol.htm Wesley et al. (2001) Intermediate vectors for assembly of an inverted repeat unit in a CaMV 35S expression cassette. The cassette can be transferred into <i>Agrobacterium</i> binary plasmid vectors using flanking <i>NotI</i> sites.
pX7-RNAi	G10-90 estradiol inducible	<i>hyg</i>	Two-step ligation	<i>Arabidopsis</i> actin intron 11	Guo et al. (2003) An <i>Agrobacterium</i> binary plasmid vector. The RNAi construct is introduced into the expression cassette of a G10-90 promoter that is inactivated by an insert that can be removed by the estradiol activation of Cre recombinase. The estradiol can be applied to the whole plant to give permanent activation of the RNAi construct. It can also give localized activation of silencing when applied to specific parts of the plant.
alcR -RNAi	<i>alc</i>	<i>kan</i>	Multistep ligation	GA20 intron	Chen et al. (2003) The alcR system was developed for alcohol-inducible protein expression and has also been used for RNAi. The RNAi constructs were assembled by sequential addition of the inverted repeat arms to the stuffer sequence followed by transfer into the <i>alc</i> expression cassette and introduction into the <i>Agrobacterium</i> binary plasmid vector. Such a system could be used with any plant transformation vector. Continued expression of the induced RNAi sequence may require repeat application of the alcohol inducer. Acetaldehyde may give better induction (Junker et al., 2003).

continued

Table 26.6.1 Vectors for RNA Silencing in Transgenic Plants^a *continued*

Vector system	Promoter	Plant selection marker	Assembly strategy	Stuffer	References/web sites/comments
pJawohl3	CaMV 35S	<i>basta</i>	Two-step ligation	<i>Arabidopsis</i> WRKY factor intron	B. Ulker, V. Lipka, T.R. Rademacher, and I. Somssich (unpub. observ.) pJawohl3-RNAi. A binary vector for gene silencing in plants. EMBO database ID AF404854. This vector is a minimal <i>Agrobacterium</i> binary plasmid that is effective when maintained on media with carbenicillin in strain GV3101:pMP90RK.
pJawohl8	CaMV 35S	<i>basta</i>	attR	<i>Arabidopsis</i> WRKY factor intron	B. Ulker, V. Lipka, T.R. Rademacher, and I. Somssich (unpub. observ.) pJawohl8-RNAi. A binary vector for gene silencing in plants. EMBO database ID AF408413. This vector is a minimal <i>Agrobacterium</i> binary plasmid that is effective when maintained on media with carbenicillin in strain GV3101:pMP90RK.

^aAbbreviations: *alc*, an alcohol-inducible promoter; attP and attR, recombination sites for integration of foreign sequence with commercially available recombinase preparations; *basta*, selection for resistance to the herbicide Basta; CaMV 35S, a strong constitutive promoter used widely in plant transformation; CHS, chalcone synthase; GUS, β -glucuronidase; *hyg*, hygromycin selection; *kan*, kanamycin selection; PDK, pyruvate orthophosphate dikinase.

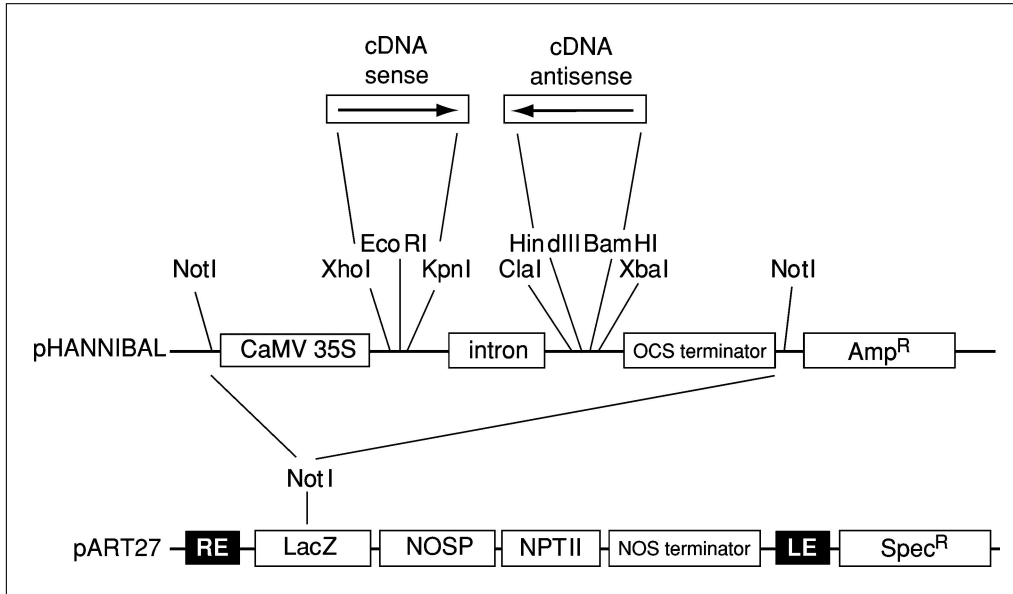


Figure 26.6.1 Vectors and cloning strategies for RNAi in transgenic plants. A target gene fragment is amplified with two sets of primers with added restriction sites and directionally cloned into two multiple cloning sites in the intermediate vector pHANNIBAL (Wesley et al., 2001). The expression of the inverted repeat of the target gene, which is separated by an intron sequence, is driven by the CaMV 35S promoter and terminated by the octopine synthase (OCS) terminator. The RNAi expression cassette is released by *NotI* and cloned into the same site of the binary vector pART27 (Gleave, 1992).

Materials

Template DNA: cDNA clone of target gene
Restriction enzymes: e.g., *KpnI*, *EcoRI*, *XhoI*, *XbaI*, *BamHI*, *HindIII*, *ClaI*, or *NotI*
(see Fig. 26.6.1)
Intermediate vector pHANNIBAL (Wesley et al., 2001) and binary vector pART27
(Gleave, 1992)
Agrobacterium tumefaciens strain GV3101 with pSa-rep (Hellens et al., 2000)
YEP medium (see recipe): with no supplements, with 50 mg/liter gentamicin, and
with 50 mg/liter gentamicin plus 50 mg/liter spectinomycin
0.15 M NaCl (autoclave and store at room temperature)
20 mM CaCl₂ (filter sterilize; store at 4°C) supplemented with 15% (v/v) glycerol
(autoclave and store at room temperature)
Liquid nitrogen
YEP plates (see recipe) supplemented with 50 mg/liter gentamicin and 50 mg/liter
spectinomycin
Arabidopsis thaliana ecotype *Columbia* (Col-0) seeds (Lehle Seeds)
0.1% (w/v) phytoagar (Caisson Laboratories; autoclave and store at room
temperature)
Arabidopsis infiltration medium (see recipe)
Silwet L-77 (Lehle Seeds)
70% ethanol containing 0.1% (v/v) Triton X-100
95% or 100% ethanol
Arabidopsis transgenic plant selection plates (see recipe) and wild-type control
plates (omit kanamycin)

100-ml and 2-liter Erlenmeyer flasks
28°C shaking incubator
High-speed refrigerated centrifuge and rotors
Equipment for growing *Arabidopsis thaliana* (Lehle Seeds):
Potting soil
3 × 5-in. plastic pots
Nylon screens
Clear covers
Environmentally controlled 22°C growth chambers and growth rooms for
Arabidopsis
22°C incubator allowing control over light/dark cycles

Rubber bands or tape
Covers of pipet-tip box or deep petri dishes as containers for *Agrobacterium*
solution
Plastic wrap (e.g., Saran wrap)
Coin envelopes for seed drying and storage
15-ml tubes
100-mm diameter filter papers, sterilized by autoclaving

Additional reagents and equipment for PCR (UNIT 15.1), restriction enzyme
digestion (UNIT 3.1), agarose gel electrophoresis (UNIT 2.5A), DNA purification
from agarose gels (UNIT 2.6), use of T4 DNA ligase (UNIT 3.14), transformation of
E. coli (UNIT 1.8), plasmid DNA minipreps (UNIT 1.6), dephosphorylation of DNA
with alkaline phosphatase (UNIT 3.10), culture of bacteria (UNIT 1.2), plant RNA
preparation (UNIT 4.3), RT-PCR (UNITS 15.5 & 15.8), northern blotting (UNIT 4.9), and
immunoblotting (UNIT 10.8)

Construct RNAi vectors for plant transformation

1. Design and synthesize two sets of primers that will amplify a region of the target gene and add appropriate restriction sites for cloning (Fig. 26.6.1).

UNIT 15.1 discusses general principles of primer design and software programs for that purpose.

The first set of primers should contain restriction sites for KpnI, EcoRI, or XhoI at the 5' end. For example, to clone the amplified cDNA into pHANNIBAL in the sense direction, the 5' primer might include a XhoI site and the 3' primer a KpnI site. The second set of primers should contain restriction sites for XbaI, BamHI, HindIII, or ClaI at the 5' end. To clone the amplified cDNA in the antisense direction, the 5' primer might include an XbaI site and the 3' primer a ClaI site. Other enzyme combinations can be selected according to the restriction sites present in the cDNA. The order of the cloning of sense or antisense fragments can also be reversed if necessary.

The authors always add 3 bp at the 5' end of the primers (before the restriction site) to facilitate restriction enzyme digestion.

2. Perform PCR to amplify the template cDNA corresponding to the target gene (UNIT 15.1).
3. Digest the PCR products with appropriate restriction enzymes (e.g., XhoI and KpnI for sense fragment; ClaI and XbaI for the antisense fragment; UNIT 3.1).
4. Perform agarose gel electrophoresis (UNIT 2.5A) on the digested PCR products. Recover the sense and antisense DNA fragments from the agarose gel (UNIT 2.6).
5. Digest the vector pHANNIBAL with one set of restriction enzymes (e.g., XhoI and KpnI). Perform agarose gel electrophoresis on the digested DNA. Recover the digested vector DNA from the agarose gel.
6. Ligate the appropriate digested cDNA fragment (in this case, the sense fragment) to the vector using T4 DNA ligase (UNIT 3.14).
7. Transform the ligation mixture into competent *E. coli* cells (UNIT 1.8).
8. Prepare minipreps of plasmid DNA (UNIT 1.6) and identify the plasmids that contain the correct insertion by restriction digestion and agarose gel electrophoresis.
9. Repeat steps 5 to 8 for the antisense cDNA fragment by digesting the plasmid DNA from step 8 that contains the correct insertion with the second set of restriction enzymes (ClaI and XbaI).
10. Digest pART27 binary vector with NotI, then treat with calf alkaline phosphatase (CIP; UNIT 3.10) to prevent self-ligation.
11. Release the RNAi cassette from the pHANNIBAL vector by digestion with NotI and clone it into the NotI site of binary vector pART27 using T4 DNA ligase.
12. Prepare minipreps of RNAi plasmid DNA.

Transform *A. tumefaciens*

13. In a 100-ml flask, inoculate a single colony of *A. tumefaciens* strain GV3101 with pSa-rep into 20 ml YEP medium supplemented with 50 mg/liter gentamicin. Incubate at 28°C overnight with shaking.

UNIT 1.2 describes basic bacterial culture techniques.

14. In a 2-liter Erlenmeyer flask, inoculate 500 ml YEP supplemented with 50 mg/liter gentamicin with 20 ml of overnight culture. Incubate at 28°C until an OD₆₀₀ of ~0.5 is obtained (which will take ~4 hr).

15. Collect cells by centrifuging 5 min at $3000 \times g$, 4°C . Remove supernatant and resuspend cells in 100 ml of 0.15 M NaCl.
16. Centrifuge 5 min at $3000 \times g$, 4°C . Remove supernatant and resuspend pellet in 10 ml ice-cold 20 mM CaCl_2 supplemented with 15% glycerol.
17. Add 1 to 2 μl of RNAi plasmid DNA (200 to 500 mg) to 200 μl of competent *A. tumefaciens* cells in a 1.5-ml microcentrifuge tube and incubate on ice for 30 min.
At this time the cells that are not needed can be frozen as 200- μl aliquots in liquid nitrogen and stored in a -80°C freezer for future use.
18. Freeze cells in liquid nitrogen for 1 min, then thaw the frozen mixture in a 37°C water bath for 1 min.
19. Add 1 ml YEP medium and incubate at 28°C for 2 hr with gentle shaking.
20. Collect cells by microcentrifuging briefly at maximum speed. Remove supernatant and resuspend cells in 100 μl YEP medium.
21. Plate the cells onto YEP plates containing 50 mg/liter gentamicin and 50 mg/liter spectinomycin and incubate at 28°C for 2 to 3 days.
22. Prepare 2-ml overnight cultures of *A. tumefaciens* from five to ten single colonies in YEP medium containing gentamicin and spectinomycin. Prepare plasmid DNA minipreps from the 2-ml overnight cultures and perform PCR (UNIT 15.1) with the primers made in step 1 to confirm that the cells contain the construct.

Grow Arabidopsis plants for transformation

23. Add ~ 200 seeds of *Arabidopsis* (ecotype Col-0) to 50 ml of 0.1% phytoagar and vernalize for 1 to 2 days at 4°C .
24. Put soil into 3×5 -in. pots and wet the soil with water (prepare 2 to 4 pots for each construct to be transformed). Put 10 to 15 of the vernalized seeds into each pot using a plastic pipet. Cover the pots with nylon screen and secure with rubber bands or tape. Cover the pots with a clear cover to keep the humidity high.
25. Grow seedlings for ~ 5 days until the first leaves appear. Remove the clear cover and grow the plants in a greenhouse or growth room at 22°C with a 16 hr light/8 hr dark cycle for ~ 4 weeks.
26. Water the plants every 5 days or as needed; when the plants start to flower, remove the emerging primary inflorescence to encourage growth of multiple secondary inflorescences.

The plants will be ready for transformation ~ 1 week after the removal of the primary inflorescence. The presence of unopened flower buds indicates the optimal time for transformation.

Transform Arabidopsis plants

27. Inoculate a single colony of *A. tumefaciens* carrying the RNAi construct into 100 ml YEP medium supplemented with 50 mg/liter gentamicin and 50 mg/liter spectinomycin. Grow for 2 days at 28°C with shaking at 200 rpm.
28. Harvest cells by centrifuging 10 min at $3000 \times g$, room temperature. Resuspend the cells in 200 ml *Arabidopsis* infiltration medium. Add 40 μl Silwet L-77 (0.02% final concentration) and mix well.
29. Pour the *A. tumefaciens* culture into a deep petri dish or the cover of a pipet-tip box. Invert the plants (inflorescences with flower buds) into the solution and incubate for 15 min at room temperature.

30. Remove plants from the *A. tumefaciens* suspension. Loosely cover plants with Saran wrap to maintain humidity and return to 22°C. Remove wrap after 24 hr. Grow plants an additional 4 weeks or until seeds are dry.
31. Harvest the seeds from each pot separately and keep them in a coin envelope for 1 week before selection.

These seeds are the first generation of transformants (T1).

Select RNAi transgenic plants

32. Sterilize 200 to 500 mg harvested T1 seeds by incubating in a 15-ml tube with 10 ml of 70% ethanol containing 0.1% Triton X-100 for 10 min with shaking. Rinse the seeds three times, each time with 10 ml of 95% or 100% ethanol. Transfer the seeds to sterilized filter papers placed in a sterile hood and let the seeds dry for 15 to 30 min.
33. Sprinkle ~20 mg of seeds (~1000 seeds) evenly onto one *Arabidopsis* transgenic plant selection plate. Seal plate with tape. Keep at 4°C for 1 to 2 days. Transfer plates to an incubator with light/dark cycles (16 hr/day) at 22°C.
34. Distinguish transgenic seedlings (green) from nontransgenic seedlings (yellow) about 2 weeks after germination. Transfer the transgenic plants to soil. Grow plants for 4 more weeks, then harvest the T2 seeds.

Characterize RNAi transgenic plants

35. Germinate 20 to 50 T2 transgenic plant lines on *Arabidopsis* transgenic plant selection plates. Also include a wild-type control plate without kanamycin.

The authors usually sterilize about 200 seeds for each line and put them in one 100 × 20-mm plate containing 25 ml of the phytoagar-containing selection medium.

36. Grow the T2 seedlings on plates for 2 to 3 weeks under light (16 hr/day). Collect 200 mg transgenic tissue (about 50 whole seedlings) and prepare total RNA (UNIT 4.3).
37. Synthesize cDNA from 2 µg RNA using RT-PCR (UNIT 15.5). Perform quantitative RT-PCR (UNIT 15.8) to quantify the expression of the target gene in wild-type control and RNAi transgenic lines.

It will be necessary to optimize the PCR cycle numbers to quantify the expression of the target gene.

38. Confirm transgenic lines with reduced target gene expression (identified by RT-PCR) by northern blotting (UNIT 4.9).

Target-gene expression can be detected by a regular northern blot (UNIT 4.9), or the presence of siRNA in the RNAi transgenic lines can be detected by a northern blot transferred from a 15% denaturing polyacrylamide gel (also described in UNIT 4.9).

39. If antibodies against the target protein are available, perform an immunoblot to confirm that the protein level of the target gene is reduced (UNIT 10.8).

VIRUS-INDUCED GENE SILENCING

This protocol describes the use of a tobacco rattle virus (TRV) vector for virus-induced gene silencing (VIGS). TRV has a bipartite RNA genome. One of the components (RNA1) encodes the RNA replication enzymes and the other component (RNA2) encodes the coat protein. In the system described here, RNA2 is engineered such that a nucleic acid sequence corresponding to the target gene of interest can be replicated as part of the viral genome. The engineered RNA2 containing the RNAi construct is then cloned into an *A. tumefaciens* binary vector. RNA1 is cloned separately into another *A. tumefaciens*

binary vector. The two different *A. tumefaciens* strains carrying the RNA1 and RNA2 constructs are then coinfecting into a target plant (*Nicotiana benthamiana*) that is highly susceptible to *A. tumefaciens*, allowing a high level of transient expression of genes on the binary vectors that are transferred from *A. tumefaciens* to the plant cells. This transient expression system is referred to as "agroinfection." VIGS develops over a period of 1 to 3 weeks following *A. tumefaciens* infection, as virus particles are produced.

It is possible to circumvent the use of *A. tumefaciens* for the inoculation; viral cDNA can be inoculated directly or it can be transcribed in vitro into infectious RNA. However, these latter procedures require isolation and repeated handling of the viral cDNA, and produce a lower inoculation efficiency than *Agrobacterium*.

In the VIGS protocol described below, it is useful to include an empty TRV2 vector as a control. The symptoms on the infected plant are slight stunting and very mild mosaicism. A useful positive control for silencing is the TRV2 vector with an insert from the endogenous *N. benthamiana* genes encoding phytoene desaturase (PDS) or magnesium chelatase. The silencing symptoms develop over 1 to 3 weeks as photobleaching (PDS) or yellow chlorosis (magnesium chelatase). It is important that the plants do not experience temperature stress above 28°C (for *N. benthamiana*).

CAUTION: Plants infected with virus vectors should be grown under appropriate containment conditions according to regulations determined by the appropriate regulatory authority and with the approval of the relevant institutional biosafety committee.

Materials

- A. tumefaciens* binary vector pTRV: TRV RNA2 vector construct (e.g., pTV00 based on TRV strain PPK20; Ratcliff et al., 2001); freely available to academic scientists from the Baulcombe laboratory (<http://www.sainsbury-laboratory.ac.uk/dcb/>)
- SOB medium (see recipe)
- 10% (v/v) glycerol, ice cold
- L medium and plates (see recipe) supplemented with 50 µg/ml kanamycin, 5 µg/ml tetracycline, and 50 µg/ml rifampicin
- A. tumefaciens* strain C58C1 carrying pBINTRA6 (GeneBank accession no. AF314165; Ratcliff et al., 2001); freely available to academic scientists from the Baulcombe laboratory (<http://www.sainsbury-laboratory.ac.uk/dcb/>)
- SOB medium (see recipe) supplemented with 50 µg/ml kanamycin, 5 µg/ml tetracycline, and 50 µg/ml rifampicin
- 10 mM MgCl₂ with 100 µM acetosyringone (3',5'-dimethoxy-4'-hydroxyacetophenone; Acros Organics) and 1 mM MES, pH 5.6
- Nicotiana benthamiana* seedlings (5 cm high, 2 weeks old)
- 2-liter Erlenmeyer flasks
- 250-ml centrifuge bottles
- Electroporator (Cell-Porator and Voltage Booster; Invitrogen) and 0.15-cm cuvettes, prechilled
- Culture tubes
- 1-ml syringes
- 20° to 24°C greenhouse or growth chamber
- Additional reagents and equipment for constructing RNAi vector and transforming *A. tumefaciens* (see Basic Protocol 1), electroporation (optional; UNIT 1.8), alkaline lysis minipreps (UNIT 1.6), PCR (UNIT 15.1), and restriction enzyme digestion (UNIT 3.1)

Construct RNAi cDNA viral vectors

1. Engineer an inverted cDNA RNAi construct in *A. tumefaciens* binary vector pTRV (see Basic Protocol 1, steps 1 to 12, but use the pTRV binary vector).

The TRV vector pTV00 has a SmaI site for insertion of blunt-ended cDNA fragments from any target gene. The fragments can be derived by suitable restriction enzyme digestion of the cloned DNA or by PCR amplification of genomic DNA or cDNA with primers that include restriction sites that leave blunt-ended DNA. The primers are designed with at least three nucleotides outside the restriction site to facilitate digestion. The ligation reaction is carried out with a two-fold molar excess of RNAi insert to TRV vector.

Transform *A. tumefaciens* with pTRV (RNA2)

Transform *A. tumefaciens* strain GV3101 with pTRV DNA by electroporation (steps 3 to 6 below) or by the CaCl₂ method (see Basic Protocol 1, steps 13 to 20). After either transformation method, proceed to selection of transformed cells as in step 7 below.

2. Add 1.0 ml of an overnight culture of *A. tumefaciens* strain GV3101 with pSa-rep to 500 ml of SOB medium in a 2-liter flask and incubate for 5 to 6 hr at 28°C with vigorous shaking.
3. When the OD₅₅₀ reaches 0.7, chill the culture on ice for 30 min, then transfer to 250-ml centrifuge bottles and centrifuge 15 min at 2600 × *g*, 4°C. Remove supernatant, then wash four times, each time by adding 250 ml 10% glycerol, centrifuging as before, and removing the supernatant. Resuspend cells in a final volume of 1.0 ml in ice-cold 10% glycerol. Use electrocompetent cells immediately or store at –70°C.
4. To a 40-μl aliquot of electrocompetent cells, add 50 ng of the engineered TRV RNA2 vector cDNA.
5. Electroporate (UNIT 1.8) using prechilled 0.15-cm cuvettes in a Cell-Porator with Voltage Booster according to conditions recommended by the manufacturer (capacitance, 330 μF; resistance, 4000 Ω; voltage, 380 V; impedance, low Ω; charge rate, fast).
6. Transfer the electroporated cells to 0.5 ml SOB medium and incubate at 28°C for 1 hr with gentle shaking (100 rpm).
7. Plate out the transformed cells on L plates supplemented with 50 μg/ml kanamycin, 5 μg/ml tetracycline, and 50 μg/ml rifampicin.
8. Prepare a 5-ml overnight culture in L medium supplemented as above.
9. Isolate plasmid DNA from *Agrobacterium* using the alkaline lysis miniprep method (UNIT 1.6). Characterize the DNA by PCR (UNIT 15.1) or restriction enzyme digestion (UNIT 3.1) to confirm the identity of the plasmid DNA in the transformed *A. tumefaciens*.

*For the restriction digestion approach, it may be necessary to reamplify the isolated DNA in *E. coli* strain DH5α because the yield from *A. tumefaciens* is low. Transform *E. coli* by calcium chloride or electroporation (UNIT 1.8) and isolate plasmid DNA by alkaline lysis miniprep (UNIT 1.6).*

Inoculate plants

10. Prepare separate 10-ml *A. tumefaciens* cultures containing pBINTRA6 (C58C1) and the TRV RNA2 vector cDNA (GV3101:pSa-rep) in SOB medium with 50 μg/ml kanamycin, 5 μg/ml tetracycline, and 50 μg/ml rifampicin. Grow overnight at 28°C.

pBINTRA6 is an RNAi Ti plasmid clone of TRV.

11. Centrifuge the cultures 20 min at 2600 × *g*, room temperature, and remove the supernatants. Resuspend each pellet individually in 10 ml of 10 mM MgCl₂ with 100 μM acetosyringone and 1 mM MES, pH 5.6.

- Mix the two suspensions in a 1:1 ratio and infiltrate the liquid into leaves of *N. benthamiana* seedlings by transferring aliquots to a 1-ml syringe and pressing the tip of the syringe against the lower surface of the leaves.

Infiltrate two to three leaves per plant. Apply as much suspension as necessary to fill most of the leaf, ~200 μ l depending on the size of the leaf. The leaf appears water-soaked in the infiltrated zone.

- Maintain the plants at 20° to 24°C in a greenhouse or growth chamber under 8 hr dark/16 hr light cycles and monitor the development of viral symptoms.

The viral symptoms appear after about 10 days and reflect the phenotype of the silenced target gene. They will vary from mild mosaic to death of the plant depending on the function of the gene that is targeted by silencing.

SUPPRESSION OF SILENCING FOR PROTEIN OVEREXPRESSION

Transient gene expression is a fast, flexible, and reproducible approach to high-level expression of useful proteins. In plants, recombinant strains of *Agrobacterium tumefaciens* can be used for transient expression of genes that have been inserted into the T-DNA region of the bacterial Ti plasmid. A bacterial culture is vacuum-infiltrated into leaves, and, upon T-DNA transfer, there is ectopic expression of the gene of interest in the plant cells. Ectopic protein expression continues for 2 or 3 days, but then ceases because of post-transcriptional gene silencing (PTGS). Virus-encoded suppressors of silencing can be used to extend the time of ectopic protein expression and to increase the level of protein-produced viral-encoded suppressor of gene silencing (Johansen and Carrington, 2001; Voinnet et al., 2003). This protocol uses the p19 silencing suppressor protein of tomato bushy stunt virus. Coexpression of p19 enhances the expression of foreign protein by up to 50-fold, so that protein purification can be achieved from as little as 100 mg of infiltrated leaf material (Voinnet et al., 2003). Silencing suppressor-enhanced expression systems have potential value in industrial production and as a research tool for isolation and biochemical characterization of a broad range of proteins, without the need for the time-consuming regeneration of stably transformed plants.

Materials

- pBin61 binary plasmid vector with *Xba*I and *Sma*I restriction sites in the CaMV 35S expression cassette (Bendahmane et al., 2000); freely available to academic scientists from the Baulcombe laboratory (<http://www.sainsbury-laboratory.ac.uk/dcb/>)
- Restriction enzymes: *Sma*I and *Xba*I
- cDNA of full open reading frame to be overexpressed
- E. coli* strain DH5 α
- L plates (see recipe) supplemented with 50 μ g/ml kanamycin
- A. tumefaciens* strain C58C1:pCH32
- L plates (see recipe) supplemented with 50 μ g/ml kanamycin, 5 μ g/ml tetracycline, and 50 μ g/ml rifampicin
- A. tumefaciens* strain C58C1:pCH32:p35S-p19 (or C58C1:pCH32:p35S-pHC-Pro) harboring expression plasmids of the p19 (or HC-pro) viral suppressors of silencing (Voinnet et al., 2003); freely available to academic scientists from the Baulcombe laboratory (<http://www.sainsbury-laboratory.ac.uk/dcb/>)
- L medium (see recipe) supplemented with 50 μ g/ml kanamycin and 5 μ g/ml tetracycline
- 10 mM MgCl₂ with 100 μ M acetosyringone (3',5'-dimethoxy-4'-hydroxyacetophenone; Acros Organics) and 1 mM MES, pH 5.6
- Nicotiana benthamiana* seedlings (5 cm high, 2 weeks old)

28°C shaking incubator
Culture tubes
1-ml syringes
20°C to 25°C greenhouse or growth chamber allowing control of light/dark cycles
Additional reagents and equipment for restriction digestion (*UNIT 3.1*), use of T4 DNA ligase (*UNIT 3.14*), transformation of *E. coli* (*UNIT 1.8*), transformation of *A. tumefaciens* (see Basic Protocol 1 or 2), and growth of bacteria (*UNIT 1.2*)

Construct overexpression clone

1. Perform a restriction digest (*UNIT 3.1*) of pBin61 using sites that cleave in the CaMV 35S expression cassette (*Sma*I, *Xba*I). Using T4 DNA ligase (*UNIT 3.14*), ligate to cDNA of the full open reading frame to be overexpressed.
2. Transform the ligation mix into *E. coli* strain DH5 α (*UNIT 1.8*) and select for transformants on L plates with 50 μ g/ml kanamycin. Confirm the clone construction by sequence analysis or restriction digestion of the cloned DNA.
3. Transform into *A. tumefaciens* strain C58C1:pCH32 using the methods described Basic Protocol 1 or 2. Plate out transformants on L plates supplemented with 50 μ g/ml kanamycin, 5 μ g/ml tetracycline, and 50 μ g/ml rifampicin.

Overexpress protein

4. Prepare separate 10-ml overnight cultures of *A. tumefaciens* strains C58C1:pCH32:p35S-p19 (or C58C1:pCH32:p35S-pHC-Pro) and C58C1:pCH32:pBin61 (protein X) in L medium supplemented with 50 μ g/ml kanamycin and 5 μ g/ml tetracycline. Grow overnight at 28°C.

UNIT 1.2 describes basic bacterial culture techniques.

5. Centrifuge the cultures 20 min at 2600 \times g, room temperature, and remove the supernatants. Resuspend each pellet individually in 10 ml of 10 mM MgCl₂ with 100 μ M acetosyringone and 1 mM MES, pH 5.6.
6. Adjust the OD₆₀₀ of each culture to 0.5 with the same solution used in step 5 and mix the suspensions in a 1:1 ratio.
7. Infiltrate the liquid into leaves of *N. benthamiana* seedlings (3 to 4 leaves per plant) by transferring aliquots to a 1-ml syringe and pressing the tip of the syringe against the lower surface of the leaves.
8. Maintain the plants at 20° to 25°C in a greenhouse or growth chamber under 8 hr dark/16 hr light cycles. Monitor protein expression over a period of up to 10 days post-infiltration in the infiltrated regions of the leaves.

The temperature should not be allowed to increase over 28°C. Expression of foreign proteins will normally be highest at 5 to 7 days post-infiltration. Signs of protein expression are entirely dependant upon the protein being expressed.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

***Arabidopsis* infiltration medium**

$\frac{1}{2} \times$ Murishige & Skoog (MS) macro and micronutrients (Caisson Laboratories)
10 μ g/liter benzylaminopurine (Sigma; add 10 μ l of a 1 mg/ml stock per liter medium)
0.02% (v/v) Silwet L-77 (Lehle Seeds)
Adjust pH to 5.7 with 1 M KOH
Prepare fresh (no need to sterilize)

Arabidopsis transgenic plant selection plates

$\frac{1}{2}$ × Murishige & Skoog (MS) macro and micronutrients (Caisson Laboratories)
1% (w/v) sucrose
0.8% (w/v) phytoagar (Caisson Laboratories)
Adjust pH to 5.7 with 1 M KOH

Autoclave, cool to 55°C, and add kanamycin to 50 mg/liter (omit for wild-type control plates). Pour into 150 × 15–mm or 100 × 20–mm petri dishes, depending on the steps in which they are to be used. Store up to 2 to 3 months at 4°C.

L medium and plates

For liquid medium:

10 g Bacto tryptone
5 g Bacto yeast extract
5 g NaCl
1 g D-glucose
H₂O to 1 liter
Autoclave
Cool to 55°C

Add appropriate antibiotics from filter-sterilized stocks
Store up to 3 months at room temperature

For plates: Prepare medium as described above, but add 1% (w/v) agar (e.g., Difco) before autoclaving. After partially cooling, and antibiotics, pour medium into plates, and allow to harden. Store up to 1 month at 4°C.

SOB medium

2% (w/v) Bacto tryptone (Difco)
0.5% (w/v) Bacto yeast extract (Difco)
10 mM NaCl
2.5 mM KCl
Store up to 2 to 3 months at room temperature
Immediately before use, add antibiotics as needed from 50 mg/ml filter-sterilized stocks

YEP medium and plates

For liquid medium:

10 g Bacto yeast extract (Difco)
10 g Bacto peptone (Difco)
5 g NaCl
H₂O to 1 liter
Adjust pH to 7.0 using 1 M NaOH
Autoclave
Cool to 55°C

Add appropriate antibiotics from filter-sterilized 50 mg/ml stocks (store stocks at –20°C)

For plates: Prepare medium as above, but add 1.5% (w/v) agar (e.g., Difco) before autoclaving. After partially cooling, and antibiotics, pour medium into plates, and allow to harden. Wrap plates in Parafilm and store up to 2 to 3 months at 4°C.

COMMENTARY

Background Information

Silencing in transgenic plants

The first reports of RNA silencing in transgenic plants involved sense or antisense transgenes. In a small proportion of the lines produced with any one construct, there was coordinate silencing of both the transgene and the endogenous gene manifested as a loss of gene function (Napoli et al., 1990; van der Krol et al., 1990). It is now known that silencing in these lines occurs because an RNA-dependent RNA polymerase (RdRP) converts the single-stranded transcript of the transgene into a dsRNA, which is then processed by a Dicer into siRNA (Dalmay et al., 2000; Beclin et al., 2002). However, it is not yet understood why there is silencing in only a subset of the lines. One of the proposed explanations implicates an RNA threshold; in non-silencing lines, the transgene RNA would be present below a threshold level that is required for activation of silencing (Lindbo et al., 1993).

An alternative explanation of silencing variation invokes a qualitative difference in the transcripts of different lines; silencing lines would be those producing aberrant RNAs that are good templates for the RdRP (English et al., 1996). There has been extensive discussion about the features that define an aberrant RNA but, as yet, no definitive information. A likely scenario is that endogenous mRNAs are associated with proteins that interfere with the RdRP. According to this idea, the transgene RNAs might have a modified ribonucleoprotein (RNP) structure that allows the RdRP to produce dsRNA. The modified RNP might arise, for example, if the transgene RNAs are incomplete at the 3' end and, consequently, lack a poly(A)-binding protein or other component of the RNP.

Until these threshold and aberrant RNA models have been tested, there will be no rational criteria for design of sense or antisense constructs giving consistent silencing. However, the requirement for an RdRP step can be circumvented by use of dsRNA constructs (Dalmay et al., 2000; Beclin et al., 2002). These constructs have inverted repeat modules in the transcribed region, separated by a short stuffer sequence, and they are coupled to either constitutive (Waterhouse et al., 1998; Chuang and Meyerowitz, 2000) or inducible (Chen et al., 2003; Guo et al., 2003) promoters. The inducible promoters have to be used if silencing of the target genes

would prevent regeneration of the transgenic plant.

Versions of the dsRNA constructs have different types of stuffer sequence between the inverted repeats. In some of the vectors, this stuffer is an intron, whereas in others it is a sequence of noncoding DNA without known biological activity. There is one report that the silencing efficiency is increased if the stuffer sequence is an intron (Smith et al., 2000). However, it has not been ruled out that there is a sequence-specific influence of the linker sequence that is unrelated to intron processing.

Constructs using the earliest generation of dsRNA vectors can be assembled using conventional restriction enzymes and ligation. More recent vectors can be assembled without restriction enzymes, using *in vitro* attP and attR recombinase systems (<http://www.invitrogen.com/>; Wesley et al., 2001). These newer vectors are useful for high-throughput applications in which hundreds or even thousands of sequences are to be targeted as part of a survey of gene function. Table 26.6.1 summarizes the features of the dsRNA vectors and indicates the source of additional information, including access procedures.

In most applications, the silencer constructs will be used as transgenes that are integrated into the genome of a plant. Although the silencing phenotype will typically be apparent in most of the transgenic lines (Chuang and Meyerowitz, 2000; Smith et al., 2000; Wesley et al., 2001), there may be some between-line variation with respect to the degree of silencing. If complete loss of function in the target gene leads to death or growth arrest, the intermediate phenotypes that are due to partial silencing may be informative.

There are some instances when the silencer constructs can be introduced directly into cells where the target gene is expressed. This approach does not require integration of the silencer construct into the genome, and is best illustrated by experiments in barley in which the silencing target encodes a component of the disease-resistance signal-transduction pathway (Schweizer et al., 2000). The silencer was an inverted-repeat construct, as described above, introduced into a cell by bombardment with DNA-coated gold particles. The silencing phenotype that developed over a few days was manifested as growth of a pathogen in cells of a pathogen-resistant plant. The cells in the bombarded tissue that received the silencer DNA

were identified by coexpression of a reporter gene together with the silencer construct.

Transient silencing can be adapted to plants other than barley by modification of the conditions for particle bombardment. Alternatively, it may be possible to use *A. tumefaciens* to deliver the silencer constructs in a procedure that is normally used for transient protein overexpression (Hamilton et al., 2002). *A. tumefaciens* cells carrying the silencer construct are grown in liquid culture and harvested, and the resuspended cells are then infiltrated into leaves. The silencer construct transfers from the *A. tumefaciens* cells into the plant cells and, after 1 or 2 days, there is accumulation of the corresponding siRNA species.

The direct introduction of silencer constructs, as opposed to stable transgenesis, has the advantage of speed. However, methods employing the former approach are only appropriate when the target gene displays a phenotype in cells located in regions of the plant where they can be bombarded with DNA-coated gold particles or infiltrated with liquid cultures of *A. tumefaciens*. These transient silencing systems are also only appropriate when the RNA silencing target encodes an unstable protein. Stable proteins would persist after the onset of silencing, and there may not be a silencing phenotype even with reduced levels of the target RNA.

Virus-induced silencing

An alternative to stable or transient transformation for delivery of silencer constructs is virus-induced gene silencing (VIGS; Baulcombe, 1999; Basic Protocol 2). A sequence tag corresponding to the silencing target is inserted into the vector, and the gene-silencing phenotype is manifested in the infected regions of the plant. With RNA virus vectors, it seems likely that the dsRNA is the viral replication intermediate and that it is processed into the siRNA that targets mRNA degradation. It is less clear how DNA geminivirus vectors generate this dsRNA. It could be that there is read-through transcription on both strands of the circular DNA to generate overlapping RNAs. Alternatively, the geminiviral RNA could be copied by the RdRP, as described above for transgenes.

VIGS was developed as an RNA-silencing approach for essential genes because it would not be possible to generate transformed plants in which such genes were silenced. However, in infected plants, it would be possible to infer the function of such genes from the characteristics of VIGS-induced growth arrest. For

example, if the target gene were essential for cell division, the growth arrest would be most apparent in the meristems and along the veins. If the target gene were required for cell expansion, these zones of cell division would be unaffected, but the elongation zones of stems and the areas between veins in the leaves would exhibit growth arrest. VIGS was also designed for high-throughput applications because it does not require the time-consuming and labor-intensive procedures for plant transformation.

The advantages of VIGS for the analysis of gene function are illustrated by its use in a survey of nearly 5000 cDNAs (Lu et al., 2003a). One of the findings from this survey was that HSP90 proteins are required for disease resistance; this is an example of how silencing can be used to investigate multigene families. Repeated genes are often not amenable to mutation studies because of functional redundancy. A second advantage of VIGS can be inferred from the stunting or growth-arrest silencing symptoms in many of the infected plants (Lu et al., 2003a). These growth-arrested plants would most likely have a lethal phenotype when silenced with a constitutive promoter transgene or in null mutants.

The most widely used VIGS vectors are tobacco rattle virus (TRV) and potato virus X (PVX). TRV is effective on *N. benthamiana*, tomato, *Arabidopsis*, and other related species (Dalmay et al., 2000; Ratcliff et al., 2001; Liu et al., 2002), whereas PVX induces strong silencing only on *N. benthamiana* and *N. clevelandii* (Ruiz et al., 1998). The first VIGS vector was tobacco mosaic virus (TMV; Kumagai et al., 1995), but it has not been widely used because it induces strong symptoms in infected plants. However, a TMV-based system has been developed that, unlike TRV and PVX, gives good VIGS on the model plant *N. tabacum* (Gossele et al., 2002). Barley stripe mosaic virus (BSMV) has been used as a VIGS vector on barley (Holzberg et al., 2002), and geminivirus vectors are effective with *N. benthamiana* (Kjemtrup et al., 1998) and *Arabidopsis* (Turnage et al., 2002). For reasons that are not fully understood, VIGS is more effective in *N. benthamiana* and related species than in other hosts. Table 26.6.2 summarizes the features of various virus vectors and indicates references or Web sites that can be used for additional information. The vectors described use simple inserts, but the silencing effect can be enhanced if the insert is in an inverted repeat configuration (Lacomme et al., 2003).

Table 26.6.2 Vectors for VIGS

Virus	Assembly strategy	Host plants	Inoculation method	Reference	Web site
PVX (pGR106-7) ^a	Ligation	<i>N. benthamiana</i>	<i>Agrobacterium</i>	Lu et al. (2003a)	http://www.sainsbury-laboratory.ac.uk
TRV (pTV00) ^b	Ligation	<i>N. benthamiana</i> , <i>Arabidopsis</i>	<i>Agrobacterium</i> or manual inoculation of partially purified virions	Ratcliff et al. (2001)	http://www.sainsbury-laboratory.ac.uk
TRV ^c	Ligation, attP	<i>N. benthamiana</i> , tomato	<i>Agrobacterium</i>	Liu et al. (2002)	http://plantfunctionalgenomics.yale.edu/Dinesh/research/NSF/VIGSindex.html
SVISS: chimeric TMV/STMV ^d	Ligation	<i>N. benthamiana</i> , <i>N. tabacum</i> ,	Manual inoculation of in vitro transcripts	Gossele et al. (2002)	—
Tomato golden mosaic geminivirus ^e	Ligation	<i>N. benthamiana</i>	Bombardment of infectious DNA	Peele et al. (2001)	—
Cabbage leaf curl geminivirus ^f	Ligation	<i>Arabidopsis</i>	Bombardment of infectious DNA	Turnage et al. (2002)	—

^aThe PVX vector system was developed originally for protein overexpression and was then adapted for silencing. Earlier versions of the vector were transcribed in vitro to produce infectious RNA or were inoculated directly as cDNA. *Agrobacterium* inoculation is now preferred because the vector constructs do not need to be isolated and purified. In some respects PVX is easier to use than TRV because its genome has only a single component, and a toothpick inoculation procedure can be used that does not require a liquid culture. A toothpick is dipped into a small colony of the *Agrobacterium* vector clone and then pushed into a leaf or petiole of the plant to be inoculated. However, PVX symptoms are more severe than those of TRV and may interfere with the silencing phenotype. The reference describes the use of this system in a forward screen of genes required for disease resistance.

^bThis vector has largely superseded PVX for *N. benthamiana* because silencing is more persistent and viral symptoms are milder. It has also been used in *Arabidopsis*, but *Agrobacterium* inoculation is not effective. The most effective strategy at present requires an inoculum of partially purified virions and, even so, a typical experiment has several plants that do not become infected.

^cThe two TRV vector systems in this table were produced in separate laboratories from the same strain of virus, PPK20. The Dinesh-Kumar version (Liu et al., 2002) is effective on tomato and *Solanum* species and generates more severe symptoms on *N. benthamiana*. This version has been adapted to clone construction using the attP recombinase system.

^dSVISS is unlike the PVX and TRV vectors in that it is effective on *N. tabacum*.

^eIf the target genes are expressed in dividing or meristematic cells, this may be the most appropriate vector. The host range has not been fully explored.

^fThis vector is probably better than TRV for *Arabidopsis*.

VIGS is effective only in the parts of the plant that are invaded by the virus. Consequently, it cannot be used for analysis of embryo and seedling characteristics or, normally, of genes expressed in fruit and seeds, because these parts of the plant are often not infected. An additional limitation of VIGS is the need for large-scale facilities for growing plants under physical containment of genetically modified plant pathogens. In the future, it may be possible to miniaturize the procedures so that plants are inoculated as they germinate and the VIGS symptoms are assayed in seedlings. Such developments will facilitate containment by reducing the time course and scale of VIGS experiments. It may also be possible to develop biological containment systems so that there will be a reduced requirement for phys-

ical containment. Such systems could involve use of a deleted vector that would be competent to survive only on a transgenic host plant that complements the missing viral gene.

Critical Parameters

RNAi construct

Sense/antisense arms ranging from 98 to 853 base pairs have been found to give efficient silencing (Wesley et al., 2001), indicating that the length of the arm sequence can be flexible. Sequences derived from ORFs (open reading frames), 5'-UTRs (untranslated regions), and 3'-UTRs did not result in significant differences in silencing (Wesley et al., 2001). However, use of highly conserved sequences in the RNAi construct can knock down the target

gene as well as its close homologs (Yin et al., 2005). According to the authors' experience, genes with sequence identity of over 80% in the inverted repeat can be silenced, although the silencing efficiency is somewhat reduced compared to that achieved with the completely matched gene. Therefore, a highly conserved region from a gene family can be used to knock down several members of the gene family simultaneously. On the other hand, if one needs to reduce the expression of a specific family member, it is important to use a unique sequence in the RNAi construct that is not conserved among the homologs.

The pHANNIBAL system uses the strong and constitutive CaMV 35S promoter, so the target gene should be silenced in most cell types throughout development. If the target gene is essential for plant growth, or if it is necessary to assess gene function in different tissues/organs or different developmental stages, several vectors with inducible promoters are also available (Chen et al., 2003; Guo et al., 2003).

Plant transformation

The floral dip method (Clough and Bent, 1998) is a very simple and efficient way to make transgenic *Arabidopsis* plants. The most important parameter is to have healthy plants with many unopened flower buds for transformation. Healthy plants can be obtained by watering plants regularly and maintaining adequate light and optimal temperatures (22°C) in the growth room. The removal of the primary inflorescence leads to many secondary inflorescences, which is also very important to get many flower buds for transformation.

VIGS

VIGS has been used for silencing of many genes in *N. benthamiana*, and has been particularly effective as a tool for the analysis of defense-related genes. In test experiments, phytoene desaturase could be silenced with inserts that were as short as 23 nucleotides (Thomas et al., 2001). However, the silencing was less extensive and more transient than with larger sequence elements, and an insert of 150 to 500 nucleotides is recommended. Larger inserts can be used, but they are genetically unstable in the virus vector and do not cause enhancement of the silencing phenotype.

If the insert is from a region that is unique to the target gene, the silencing phenotype will be highly gene-specific. However, gene families may be targeted if the insert includes regions that are similar in related genes or other members of the gene family. Until now, there has

not been a systematic analysis of the sequence similarity needed for cross-silencing of related genes, and it is recommended that sequences be chosen that are more than 90% identical to the intended target.

Anticipated Results

Plant transformation

For *Arabidopsis*, the authors routinely achieve a 0.5% to 2% transformation frequency among the progeny of the transformed line after the floral dip protocol.

Gene silencing

The pHANNIBAL vector system yields gene silencing in over 90% of the independent transgenic lines, with at least 20% reduction in target gene expression (Wesley et al., 2001). To study the loss-of-function phenotype of *BES1* and its close family members, Yin et al. (2005) cloned an inverted repeat of about 900 bp coding region from the *BES1* gene using the pHANNIBAL system and made transgenic plants. About 25% of the transgenic lines showed more than 50% reduction in the expression of both *BES1* and its closest homolog, *BZRI*, and displayed an expected dwarf phenotype.

VIGS

All plants should become infected using these methods described with all or most showing symptoms of VIGS. The VIGS symptoms appear several days post-inoculation (dpi) and become extensive throughout the infected plant by 1 to 2 weeks. At later times, they may fade. The absence of a phenotype does not necessarily rule out involvement of the target gene in the trait of interest. VIGS is never complete and it is always possible that a silencing phenotype was not observed because the target gene function was supported by the residual low level of mRNA in the virus vector-infected plants. In some instances, the silencing may not be a direct consequence of loss of target gene function; there may be pleiotropic effects of target gene silencing. A final consideration applies to genes with an embryonic-lethal silencing phenotype in transgenic silencing approaches. It would be difficult to recover such plants, particularly if the transgenic silencing constructs employed a constitutive promoter. In a VIGS assay, such genes would induce symptoms with severe effects on growth and development.

Troubleshooting

Troubleshooting for common molecular biology techniques can be found in other

sections of this book. Potential troubles specific for this protocol are discussed here.

No or very few transgenic plants are recovered on selective media

If no transgenic plants are identified, check if the right antibiotics were used. A more common problem is that low numbers of transgenic plants are identified. In this case, grow new batches of flowering plants under more optimal conditions for transformation. Growing wild-type plants every 1 to 2 weeks will ensure that there are always plants available for transformation. Addition of 5% (w/v) sucrose to the infiltration medium can also increase the transformation efficiency.

No transgenic plants show a clear reduction in target gene expression

If an RT-PCR method is used to screen the RNAi transgenic lines, the parameters for PCR, especially the cycle numbers, need to be optimized. Reactions using 20, 22, 24, 26, 28, and 30 cycles should be tested to determine the right cycle number for a specific gene. Since RNAi lines usually knock down but do not knock out target gene expression, a small reduction in target gene expression may not be quantifiable by RT-PCR. In such a case, either real-time PCR or northern blotting should be performed to detect if there is a small reduction in target gene expression. Since the expression of a transgene is dependant on the position of its insertion in the genome (position effect), it may sometimes be necessary to screen a large number of transgenic lines to identify the few lines in which the target gene is significantly silenced. In addition, as described in Critical Parameters, a silenced transgenic plant may not be recovered if a gene is critical for plant growth and development. If this is the case, the use of inducible promoters in the RNAi construct can be helpful (Chen et al., 2003; Guo et al., 2003).

VIGS: Plants do not have a silencing phenotype

The precise timing of the maximum VIGS is highly dependent on environmental conditions: any stress to the plant would interfere with the *Agrobacterium* inoculation and/or spread of the virus in the infected plant. It is therefore advisable to measure the abundance of the intended target RNA or protein to confirm directly that its expression has been silenced. It is also useful to set up tests with well-characterized vector constructs to monitor the progress of the VIGS in parallel infected plants. The phytoene desaturase vectors

are useful for this purpose because they produce photobleaching symptoms that are very obvious. However, it may also be appropriate to use vectors that are targeted against genes involved in the trait of interest. For example, to identify genes required for disease resistance, it would be appropriate to establish the progression of VIGS using a control vector that targets a known defense-related gene.

VIGS: Plants do not have the expected silencing phenotype

The VIGS phenotype might not reflect loss of function in the target gene if there is sequence similarity between the insert and an unknown mRNA. Routine use of a second, nonoverlapping insert from the same target gene can rule out this artifact. If the target gene has been correctly identified, this second insert would reproduce the original VIGS phenotype. If the target gene is a member of a multigene family, it is necessary to target conserved and nonconserved regions to determine whether the silencing phenotype was influenced by one or several family members.

Time Considerations

Transgenic plants

It takes 2 to 4 weeks to make the RNAi construct (three cloning steps) and to transform the final construct into to *Agrobacterium*. Because it takes 4 to 5 weeks before the *Arabidopsis* plants are ready for transformation, several batches of *Arabidopsis* seeds should be planted before the cloning process is started. After *Arabidopsis* transformation, it will take 4 to 5 additional weeks to harvest T1 seeds and to identify T1 transgenic plants. Another 8 weeks are needed to obtain T2 seeds for further analysis. In the next 8 weeks, the T2 seedlings or plants can be analyzed, during which homozygous transgenic lines in the T3 generation can be identified to confirm the results from the T2 plants. Even when working efficiently, the entire process takes at least 6 months from start to finish.

VIGS

VIGS symptoms appear within 7 to 10 days post-inoculation and persist for at least 3 weeks.

Literature Cited

- Baulcombe, D.C. 1999. Fast forward genetics based on virus-induced gene silencing. *Curr. Opin. Plant Biol.* 2:109-113.
- Beclin, C., Boutet, S., Waterhouse, P., and Vaucheret, H. 2002. A branched pathway for

transgene-induced RNA silencing in plants. *Curr. Biol.* 12:684-688.

- Bendahmane, A., Querci, M., Kanyuka, K. and Baulcombe, D.C. 2000. *Agrobacterium* transient expression system as a tool for isolation of disease resistance genes: Application to the *Rx2* locus in potato. *Plant J.* 21:73-81.
- Chen, S., Hofius, D., Sonnewald, U., and Bornke, F. 2003. Temporal and spatial control of gene silencing in transgenic plants by inducible expression of double-stranded RNA. *Plant J.* 36:731-740.
- Chuang, C.F. and Meyerowitz, E.M. 2000. Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 97:4985-4990.
- Clough, S.J. and Bent, A.F. 1998. Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16:735-743.
- Dalmay, T., Hamilton, A., Rudd, S., Angell, S., and Baulcombe, D.C. 2000. An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* 101:543-553.
- English, J.J., Mueller, E., and Baulcombe, D.C. 1996. Suppression of virus accumulation in transgenic plants exhibiting silencing of nuclear genes. *Plant Cell* 8:179-188.
- Gleave, A.P. 1992. A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Mol. Biol.* 20:1203-1207.
- Gossele, V., Fache, I., Meulewaeter, F., Comelissen, M., and Metzlauff, M. 2002. SVISS: A novel transient gene silencing system for gene function discovery and validation in tobacco plants. *Plant J.* 32:859-866.
- Guo, H.S., Fei, J.F., Xie, Q., and Chua, N.H. 2003. A chemical-regulated inducible RNAi system in plants. *Plant J.* 34:383-392.
- Hamilton, A.J., Voinnet, O., Chappell, L., and Baulcombe, D.C. 2002. Two classes of short interfering RNA in RNA silencing. *EMBO J.* 21:4671-4679.
- Hellens, R.P., Edwards, E.A., Leyland, N.R., Bean, S., and Mullineaux, P.M. 2000. pGreen: A versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol. Biol.* 42:819-832.
- Holzberg, S., Brosio, P., Gross, C., and Pogue, G.P. 2002. Barley stripe mosaic virus-induced gene silencing in a monocot plant. *Plant J.* 30:315-327.
- Johansen, L.K. and Carrington, J.C. 2001. Silencing on the spot: Induction and suppression of RNA silencing in the *Agrobacterium*-mediated transient expression system. *Plant Physiol.* 126:930-938.
- Junker, B.H., Chu, C.C., Sonnewald, U., Willmitzer, L., and Fernie, A.R. 2003. In plants the *alc* gene expression system responds more rapidly following induction with acetaldehyde than with ethanol. *FEBS Lett.* 535:136-140.
- Kjemtrup, S., Sampson, K.S., Peele, C.G., Nguyen, L.V., Conkling, M.A., Thompson, W.F., and Robertson, D. 1998. Gene silencing from plant DNA carried by a geminivirus. *Plant J.* 14:91-100.
- Kumagai, M.H., Donson, J., Della-Cioppa, G., Harvey, D., Hanley, K., and Grill, L.K. 1995. Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. *Proc. Natl. Acad. Sci. U.S.A.* 92:1679-1683.
- Lacomme, C., Hrubikova, K., and Hein, I. 2003. Enhancement of virus-induced gene silencing through viral-based production of inverted-repeats. *Plant J.* 34:543-553.
- Lindbo, J.A., Silva-Rosales, L., Proebsting, W.M., and Dougherty, W.G. 1993. Induction of a highly specific antiviral state in transgenic plants: Implications for regulation of gene expression and virus resistance. *Plant Cell* 5:1749-1759.
- Liu, Y.L., Schiff, M., and Dinesh-Kumar, S.P. 2002. Virus-induced gene silencing in tomato. *Plant J.* 31:777-786.
- Lu, R., Malcuit, I., Moffett, P., Ruiz, M.T., Peart, J., Wu, A.-J., Rathjen, J.P., Bendahmane, A., Day, L., and Baulcombe, D. 2003a. High throughput virus-induced gene silencing implicates heat shock protein 90 in plant disease resistance. *EMBO J.* 22:5690-5699.
- Lu, R., Montserrat Martin-Hernandez, A., Peart, J., Malcuit, I., and Baulcombe, D. 2003b. Virus-induced gene silencing in plants. *Methods* 30:296-303.
- Napoli, C., Lemieux, C., and Jorgensen, R.A. 1990. Introduction of a chimeric chalcone synthase gene into *Petunia* results in reversible co-suppression of homologous genes in trans. *Plant Cell* 2:279-289.
- Peele, C., Jordan, C.V., Muangsang, N., Turnage, M., Egelkrot, E., Eagle, P., Hanley-Bowdoin, L., and Robertson, D. 2001. Silencing of a meristematic gene using geminivirus-derived vectors. *Plant J.* 27:357-366.
- Ratcliff, F., Martin-Hernandez, A.M., and Baulcombe, D.C. 2001. Tobacco rattle virus as a vector for analysis of gene function by silencing. *Plant J.* 25:237-245.
- Ruiz, M.T., Voinnet, O., and Baulcombe, D.C. 1998. Initiation and maintenance of virus-induced gene silencing. *Plant Cell* 10:937-946.
- Schweizer, P., Pokorny, J., Schulze-Lefert, P., and Dudler, R. 2000. Double-stranded RNA interferes with gene function at the single-cell level in cereals. *Plant J.* 24:895-903.
- Smith, N.A., Singh, S.P., Wang, M.B., Stoutjesdijk, P.A., Green, A.G., and Waterhouse, P.M. 2000. Total silencing by intron-spliced hairpin RNAs. *Nature* 407:319-320.

- Thomas, C.L., Jones, L., Baulcombe, D.C., and Maule, A.J. 2001. Size constraints for targeting post-transcriptional gene silencing and for RNA-directed methylation in *Nicotiana benthamiana* using a potato virus X vector. *Plant J.* 25:417-425.
- Turnage, M.A., Muangsan, N., Peele, C.G., and Robertson, D. 2002. Geminivirus-based vectors for gene silencing in *Arabidopsis*. *Plant J.* 30:107-114.
- van der Krol, A.R., Mur, L.A., Beld, M., Mol, J.N.M., and Stuitji, A.R. 1990. Flavonoid genes in petunia: Addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell* 2:291-299.
- Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D.C. 2003. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J.* 33:949-956.
- Waterhouse, P.M., Graham, M.W., and Wang, M.B. 1998. Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc. Natl. Acad. Sci. U.S.A.* 95:13959-13964.
- Wesley, S.V., Helliwell, C.A., Smith, N.A., Wang, M.B., Rouse, D.T., Liu, Q., Gooding, P.S., Singh, S.P., Abbott, D., Stoutjesdijk, P.A., Robinson, S.P., Gleave, A.P., Green, A.G., and Waterhouse, P.M. 2001. Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant J.* 27:581-590.
- Yin, Y., Vafeados, D., Tao, Y., Yokoda, T., Asami, T., and Chory, J. 2005. A new class of transcription factors mediates brassinosteroid-regulated gene expression in *Arabidopsis*. *Cell* 120:249-259.

Contributed by Yanhai Yin

Iowa State University
Ames, Iowa

Joanne Chory
The Salk Institute for Biological Studies
La Jolla, California

David Baulcombe
John Innes Centre
Norwich, United Kingdom

CHAPTER 27

RNA-Protein Interactions

INTRODUCTION

This chapter is concerned with the analysis of RNA-protein interactions. The techniques in this chapter are analogous to those used to analyze protein-DNA interactions (Chapter 12), but the specific methods are different primarily due to the significant structural and biochemical differences between RNA and DNA and to the nature of the proteins and protein complexes that interact with these different nucleic acids.

UNIT 27.1 describes a method for fractionating and isolating RNA-protein complexes based on agarose gel electrophoresis. This method is simple, has excellent resolving power particularly for large complexes, and it minimizes disruption of the complexes. *UNIT 27.2* describes a UV cross-linking approach to identify RNA-binding proteins. The Basic Protocol involves UV cross-linking to uniformly labeled RNA, and treatment with RNase to generate proteins attached to a short oligoribonucleotide. The resulting material is analyzed by standard or two-dimensional electrophoresis, with the interacting proteins labeled by the crosslinked RNA. A modified version of this procedure incorporates immunoprecipitation, thereby allowing one to determine if a known protein is associated with a given RNA. *UNIT 27.3* describes purification of functional RNA-protein complexes using a combination of gel filtration and affinity chromatography. It is based on the specific binding of the MS2 protein to a defined RNA-hairpin structure, and is generally useful for isolating RNA-protein complexes assembled *in vitro*. Subsequent units will be concerned with specific biochemical assays for RNA-protein interactions.

Kevin Struhl

Agarose Gel Separation/Isolation of RNA-Protein Complexes

UNIT 27.1

ELECTROPHORETIC ISOLATION OF RNA-PROTEIN COMPLEXES

BASIC
PROTOCOL

Several methods are commonly used for detecting or separating RNA-protein complexes, including density-gradient sedimentation (Grabowski et al., 1985), vertical native polyacrylamide gels (Konarska and Sharp, 1986), and gel filtration (Reed, 1990). Native horizontal agarose minigels can be used for this purpose as well. Among the advantages of the agarose gel method are its simplicity, resolving power, and ability to detect complexes that are disrupted by other techniques. In addition, the use of low-melting-point agarose for fractionation readily allows for identification of the RNA species in each complex detected on the native gel.

Materials

DNA encoding desired sequence for RNA transcription (UNITS 1.5-1.7)

Reaction mixture (e.g., HeLa cell nuclear extracts, UNIT 27.3)

10× sample loading dye or 10× heparin loading dye (see recipes)

Low-melting-point agarose (Life Technologies)

0.5× TBE electrophoresis buffer (APPENDIX 2)

Gel-fixing solution: 10% acetic acid/10% methanol

Additional reagents and equipment for preparing ³²P-labeled RNA transcripts (UNIT 27.3), electrophoresis using agarose minigels (UNIT 2.5A), and autoradiography (APPENDIX 3A)

NOTE: Extreme caution should be taken to avoid RNase contamination. The experimenter should always wear gloves, and all the tubes and tips that come into contact with the sample should be certified RNase free; however, DEPC treatment (UNIT 4.1) of solutions and apparatus is not necessary.

Assembly of RNA-protein complexes

1. Prepare a ³²P-labeled RNA transcript from DNA encoding the desired sequence as described (UNIT 27.3).
2. Incubate the ³²P-labeled RNA transcript of interest in a 25- μ l reaction mixture under conditions that allow formation of RNA-protein complexes.

For example, spliceosomal complexes are assembled by incubation in HeLa cell nuclear extracts under splicing conditions (UNIT 27.3).

3. Gently mix (do not vortex) the samples with 2.5 μ l of 10× sample loading dye or 10× heparin loading dye and place on ice.

Electrophorese RNA-protein complexes on agarose gels

4. Prepare a 7 \times 8-cm horizontal low-melting-point-agarose minigel in 0.5× TBE electrophoresis buffer in a horizontal agarose minigel apparatus (UNIT 2.5A).

A thickness of ~0.5 cm is recommended (i.e., 25 ml melted agarose for a 7 \times 8-cm gel).

The optimal gel percentage should be determined for the RNA-protein complex of interest—e.g., 1.5% agarose is used for resolving spliceosomal complexes.

5. Load 10 μ l RNA-protein complex sample (step 2) in each well.

The remaining portion of each sample can be used for analysis of the RNA on a denaturing polyacrylamide gel (UNIT 2.12).

RNA-Protein
Interactions

27.1.1

Contributed by Chung-Sheng Brian Lee, Rita Das, and Robin Reed

Current Protocols in Molecular Biology (2003) 27.1.1-27.1.5

Copyright © 2003 by John Wiley & Sons, Inc.

Supplement 63

6. Run the gel at 70 V in 0.5× TBE at either room temperature or 4°C.

The resolution varies depending on temperature and must be determined for each complex. The length of time for running the gel must also be determined empirically for each type of RNA-protein complex. For the spliceosome, the 1.5% gel is run until the bromphenol blue dye reaches the bottom.

Detect RNA-protein complexes

7. Incubate the agarose gel in gel-fixing solution for 15 min.
8. Dry the gel ~1 hr under vacuum at 80°C.
9. Visualize the complexes using a phosphorimager (*APPENDIX 3A*).

For instructions regarding elution of the RNA from the gel, which can be important if multiple complexes are observed on the gel, see Support Protocol.

SUPPORT PROTOCOL

ELUTION OF RNA FROM GEL-FIXED RNA-PROTEIN COMPLEXES

If several RNA-protein complexes are detected on the native gel, it may be important to determine the identity of the RNA species in each complex. For example, in the case of the spliceosome, complexes can be resolved that contain the unspliced pre-mRNA, the excised intron, or the spliced mRNA. To do this, follow the subsequent procedure.

Materials

Gel containing bands of interest (see Basic Protocol)
2× PK buffer (see recipe)
Proteinase K
RNase-free phenol, pH 5.2
20 mg/ml glycogen, molecular-biology grade
100% ethanol
Formamide loading dye (see recipe)

Additional reagents and equipment for autoradiography (*APPENDIX 3A*)

1. Excise the bands of interest from the gel using a clean razor blade. Melt the bands by heating 5 min at 70°C.
2. Mix with an equal volume of 2× PK buffer. Add proteinase K to a final concentration of 0.25 mg/ml. Mix and incubate 10 min at 37°C.
3. Extract the samples with an equal volume of RNase-free phenol, pH 5.2. Add 1/10 vol of 20 mg/ml molecular-biology-grade glycogen (2 mg/ml final) to the aqueous phase, mix, and then add 3 vol of 100% ethanol to precipitate RNA.
4. Add 10 µl formamide loading dye to the RNA pellets and incubate 5 min at 70°C.
5. Run the RNA samples on a denaturing polyacrylamide gel (*UNIT 2.12*). Visualize the RNA species by autoradiography (*APPENDIX 3A*).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Formamide loading dye

80% formamide
10 mM EDTA
0.1% (w/v) bromphenol blue
0.1% (w/v) xylene cyanol
Store up to 1 year at -20°C

Heparin loading dye, 10×

Prepare the following in 1× TBE electrophoresis buffer (*APPENDIX 2*):

6.5 mg/ml heparin sulfate

40% (v/v) glycerol

0.5% (w/v) bromphenol blue

0.5% (w/v) xylene cyanol

Store up to one month at -20°C

PK buffer, 2×

20 mM Tris-Cl, pH 8.0 (*APPENDIX 2*)

300 mM NaCl

25 mM EDTA

2% (w/v) SDS

Store up to one year at -20°C

Sample loading dye, 10×

Prepare the following in 1× TBE electrophoresis buffer (*APPENDIX 2*)

40% (v/v) glycerol

0.5% (w/v) bromphenol blue

0.5% (w/v) xylene cyanol

Store up to 1 year at -20°C

COMMENTARY

Background Information

Electrophoresis of RNA-protein complexes on horizontal native agarose minigels was developed as a simple and rapid method for the separation of spliceosomal complexes (Das et al., 2000). In addition, the method allowed for the detection of spliceosomal complexes that are too labile to be detected by any previously developed method (Chau and Reed, 1999; Luo and Reed, 1999; Das et al., 2000). The ATP-independent E complex was identified as the earliest discrete functional spliceosomal complex using gel filtration, but could not be detected by any of the native gel systems (Reed, 1990; Michaud and Reed, 1991). However, the E complex is readily detected on native agarose minigels. The advantage of the agarose gel over gel filtration for studies of the E complex is that multiple samples can be analyzed side-by-side on the same gel. The A, B, and C spliceosomal complexes assemble sequentially after the E complex and require ATP for their assembly. In the absence of heparin, these complexes comigrate as a large complex on agarose minigels, but can be resolved when heparin loading dye is used (Das and Reed, 1999). The yeast and mammalian A, B, and C complexes, as well as the yeast E complex, can be resolved on native polyacrylamide gels (Konarska et al., 1985; Konarska and Sharp, 1986; Pikielny et al., 1986; Seraphin and Rosbash, 1989). However, these gels are more difficult to pour than the

agarose gels and often do not allow as high a level of resolution. In addition, it is much simpler to cut out complexes and identify the RNA species in the complexes using the low-melting-point native agarose gels. Finally, the spliced mRNP complex, which promotes export of mRNA, was identified as a large heparin-sensitive complex on native agarose gels (Luo and Reed, 1999). This complex, like the E complex, can be detected by gel filtration, but not by any other native-gel method. Using the spliceosomal complexes as a precedent, the native low-melting-point-agarose gels may be a good choice for analysis of other types of RNA-protein complexes.

Critical Parameters

There are several parameters that can be tested to obtain the highest possible resolution of RNA-protein complexes by native agarose gel electrophoresis. The percentage of the gel, the length of time the gel is run, the voltage for running the gel, the sample concentration, and the temperature for running the gel can all be varied to optimize the results.

Troubleshooting

With some RNA-protein complexes, a portion of the sample becomes stuck in the well of the gel. This problem can sometimes be overcome by running the gel 30 min at low voltage and then increasing the voltage for the rest of

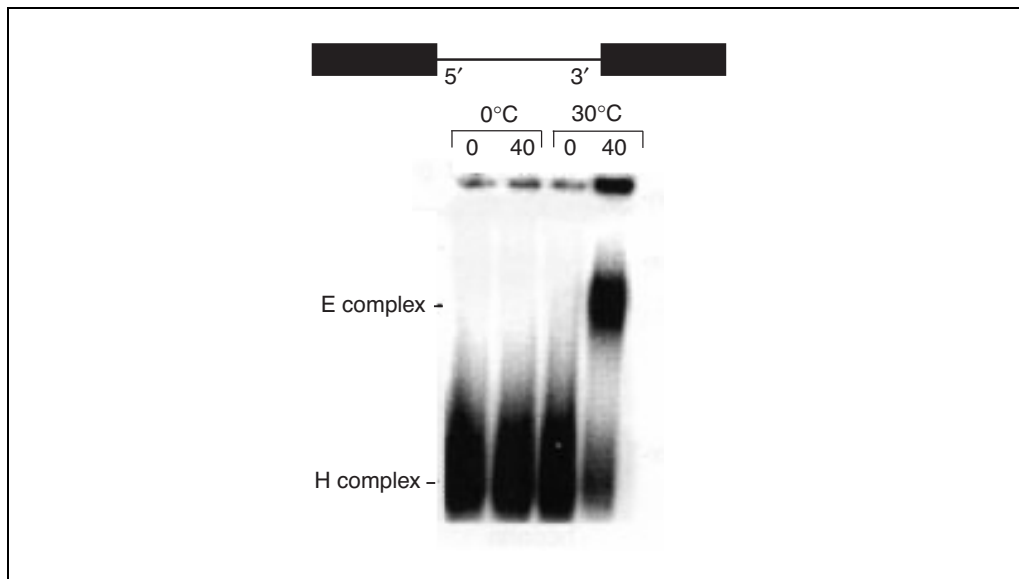


Figure 27.1.1 Detection of the earliest functional spliceosomal complex (the E complex), on a 1.5% native agarose minigel. E complex assembly requires incubation at 30°C and assembles in the absence of ATP. ³²P-labeled pre-mRNA containing one intron was incubated on ice or at 30°C for the times indicated (i.e., 0 or 40 min) in HeLa cell nuclear extracts in the absence of ATP. Aliquots of samples were gently mixed with sample loading dye (no heparin) and then loaded on the gel. The positions of the E complex and the nonspecific H complex are shown.

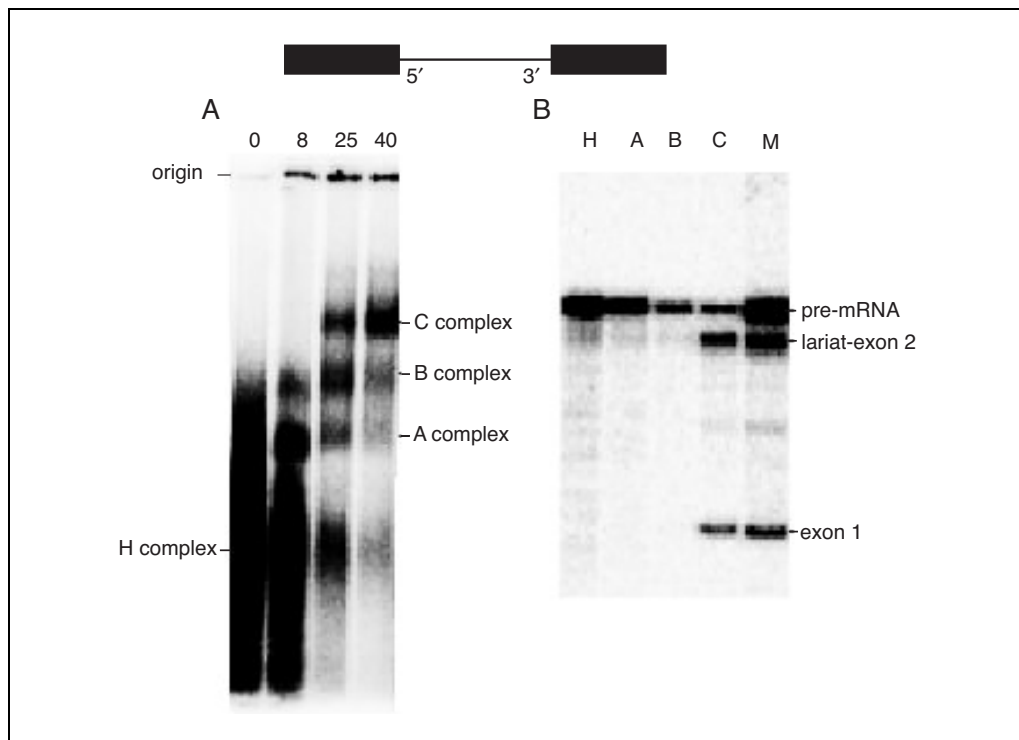


Figure 27.1.2 Resolution of the ATP-dependent spliceosomal complexes on a 2% native agarose minigel. ³²P-labeled pre-mRNA containing an AG to GG mutation at the 3' splice site (which allows assembly of the A, B, and C spliceosomal complexes, but not splicing) was incubated in nuclear extract in the presence of ATP. **(A)** At the time points indicated (i.e., 0, 8, 25, or 40 min), heparin loading dye was added to the samples before loading on a 2% agarose gel. The positions of spliceosomal complexes and nonspecific H complex are shown. **(B)** The RNA species in each complex—i.e., H (nonspecific complex), A, B, or C—was identified by excising the complexes from the agarose gel, preparing RNA, and fractionating it on an 8% denaturing gel. An aliquot of a splicing reaction incubated for 40 min was run as a marker (M). The pre-mRNA and splicing intermediates (lariat-exon 2 and exon 1) are indicated.

the running time. The sample or a portion of it may also become stuck in the well if the reaction is too concentrated. Try diluting the reaction 5- to 10-fold with 0.5× TBE electrophoresis buffer. Another strategy for eliminating the material in the well is to use the heparin loading dye. Previous studies using native polyacrylamide gels have shown that heparin dissociates nonspecific interactions and allows samples to enter the well completely. The disadvantage of heparin treatment is that it disrupts many RNA-protein complexes. The heparin can be titrated to find an optimal amount that allows sample entry without disrupting the complex of interest.

Anticipated Results

Two figures are shown to illustrate results obtained on native agarose gels when samples are treated without (Fig. 27.1.1) or with (Fig. 27.1.2) heparin.

Time Considerations

Preparation of ³²P-labeled RNA takes ~1 to 2 hr. Assembly of the RNA-protein complex varies according to the complex of interest. Separation of large RNA-protein complexes on a native agarose gel takes ~3 to 4 hr. Fixing and drying the gel takes 1.5 hr. Exposure on the phosphor imager cassette can take from 30 min to overnight depending on the efficiency of complex assembly and the specific activity of the radiolabeled RNA.

Literature Cited

- Chau, K. and Reed, R. 1999. The RNA splicing factor hSlu7 is required for correct 3' splice-site choice. *Nature* 402:207-210.
- Das, R. and Reed, R. 1999. Resolution of the mammalian E complex and the ATP-dependent

spliceosomal complexes on native agarose minigels. *RNA* 5:1504-1508.

- Das, R., Zhou, Z., and Reed, R. 2000. Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. *Molecular Cell* 5:779-787.
- Grabowski, P.J., Seiter, S.R., and Sharp, P.A. 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. *Cell* 42:345-353.
- Konarska, M.M. and Sharp, P.A. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. *Cell* 46:845-855.
- Konarska, M.M., Grabowski, P.J., Padgett, R.A., and Sharp, P.A. 1985. Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors. *Nature* 313:552-557.
- Luo, M.J. and Reed, R. 1999. Splicing is required for rapid and efficient mRNA export in metazoans. *Proc. Natl. Acad. Sci. U.S.A.* 96:14937-14942.
- Michaud, S. and Reed, R. 1991. An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes & Dev.* 5:2534-2546.
- Pikielny, C.W., Rymond, B.C., and Rosbash, M. 1986. Electrophoresis of ribonucleoproteins reveals an ordered assembly pathway of yeast splicing complexes. *Nature* 324:341-345.
- Reed, R. 1990. Protein composition of mammalian spliceosomes assembled in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 87:8031-8035.
- Seraphin, B. and Rosbash, M. 1989. Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell* 59:349-358.

Contributed by Chung-Sheng Brian Lee,
Rita Das, and Robin Reed
Harvard Medical School
Boston, Massachusetts

Identification of RNA Binding Proteins by UV Cross-Linking

UNIT 27.2

UV cross-linking is a standard method used to detect RNA-binding proteins. This method takes advantage of the ability of a photoreactive group, upon UV irradiation, to trigger the formation of a covalent bond between the RNA and closely interacting proteins. In this assay, proteins and a ^{32}P -radiolabeled RNA substrate are incubated to allow complex formation. The reaction mixture is then UV-irradiated, followed by treatment with RNase. This produces a protein attached to a short oligoribonucleotide. These protein-oligoribonucleotide products are analyzed by standard SDS-PAGE (UNIT 10.2A) or two-dimensional gel electrophoresis (UNITS 10.3 & 10.4), followed by phosphorimaging (APPENDIX 3A). Combined with immunoblotting (UNIT 10.8), immunoprecipitation (UNIT 10.16), and/or mass spectrometry (UNITS 10.21 & 10.22), the proteins can often be identified. Site-specific labeling of the RNA with ^{32}P increases the power of the UV-cross-linking assay because the RNA-protein interaction site can be defined on the RNA.

The following protocol demonstrates this method using the spliceosomal complex as an example.

CAUTION: Radioactive materials require special handling. See APPENDIX 1F and the institutional Radiation Safety Office for guidelines concerning proper handling and disposal.

NOTE: Extreme caution should be taken to avoid RNase contamination. The experimenter should always wear gloves and all the tubes and tips that come into contact with the sample should be certified RNase free; however, DEPC treatment (UNIT 4.1) of solutions and apparatuses is not necessary.

UV CROSS-LINKING USING A UNIFORMLY LABELED RADIOACTIVE RNA

BASIC
PROTOCOL

Materials

Linearized (UNIT 3.1) DNA template for transcription of the RNA of interest
200 mM and 2 M DTT (American Bioanalytical)
0.1% (v/v) Triton X-100 (Sigma)
1 mM NTPs (UNIT 3.4)
6.7 mM G(5')ppp(5')G cap (NEB)
RNasin (Amersham Pharmacia Biotech)
High specificity ^{32}P -labeled nucleotides (NEN):
10 Ci/ μl [α - ^{32}P]UTP (3000 Ci/mmol)
10 Ci/ μl [α - ^{32}P]ATP (3000 Ci/mmol)
10 Ci/ μl [α - ^{32}P]GTP (3000 Ci/mmol)
10 Ci/ μl [α - ^{32}P]CTP (3000 Ci/mmol)
50 U/ μl T7 or 20 U/ μl SP6 RNA polymerase (NEB)
5 \times transcription buffer (see recipe)
5 mg/ml DNase I (Amersham Pharmacia Biotech)
RNA phenol (American Bioanalytical)
5 M ammonium acetate
20 mg/ml glycogen, molecular-biology grade (Amersham Pharmacia Biotech)
70% and 100% ethanol
12.5 mM ATP
80 mM MgCl_2

RNA-Protein
Interactions

Contributed by Ming-Juan Luo and Robin Reed

Current Protocols in Molecular Biology (2003) 27.2.1-27.2.10

Copyright © 2003 by John Wiley & Sons, Inc.

27.2.1

Supplement 63

0.5 M creatine phosphate (prepare from di-Tris salt)
 Splicing dilution buffer (see recipe)
 HeLa cell nuclear extract (UNIT 27.3; store at -80°C)
 Sephacryl S-500 high resolution resin (Amersham Pharmacia Biotech)
 1 \times gel-filtration column buffer (see recipe)
 10 mg/ml protease-free RNase A (Amersham Pharmacia Biotech)
 100 U/ μl RNase T1 (Boehringer Mannheim)
 20% (w/v) SDS
 Acetone
 2 \times protein-gel sample buffer (see recipe) or 2 \times SDS sample buffer (UNIT 10.2A)
 30 $^{\circ}$, 37 $^{\circ}$, 40 $^{\circ}$ (for SP6 transcripts), 65 $^{\circ}$, and 100 $^{\circ}\text{C}$ water baths
 1.5-cm-diameter \times 50-cm-long column (89-ml bed volume)
 Microcentrifuge tubes with caps removed
 254-nm UV light source (e.g., Sylvania G15T8 lamp)

Additional reagents and equipment for packing gel filtration columns (UNIT 10.9),
 1-D SDS-PAGE (UNIT 10.2A) or 2-D SDS PAGE (UNITS 10.3 & 10.4), and
 autoradiography (APPENDIX 3A)

Prepare the high specificity radioactive RNA

1. Prepare a transcription mixture by combining the following:

2 μl linearized DNA template for transcription of the RNA of interest
 5 μl 200 mM DTT
 10 μl 0.1% (v/v) Triton X-100
 10 μl 1 mM NTPs
 12.5 μl of 6.7 mM G(5')ppp(5')G cap
 2 μl RNasin
 5 μl 10 Ci/ μl [α - ^{32}P]UTP
 5 μl 10 Ci/ μl [α - ^{32}P]ATP
 5 μl 10 Ci/ μl [α - ^{32}P]GTP
 5 μl 10 Ci/ μl [α - ^{32}P]CTP
 20 μl 5 \times transcription buffer (i.e., T7 or SP6 RNA polymerase buffer)
 4 μl 50 U/ μl T7 or 20 U/ μl SP6 RNA polymerase
 14.5 μl H₂O (i.e., adjust to 100 μl final).

The size of the reaction can be modified based on the desired quantity of the transcript.

The radioactive NTPs and the quantities of each radioactive NTP should be chosen based on the desired specific activity of the RNA and the RNA sequence at the protein binding site (if known). As many as 10 to 20 nucleotides can be protected from RNase digestion by the cross-linked protein and therefore remain covalently attached to the RNA. This short RNA fragment is ultimately used to detect the protein after SDS-PAGE.

- 2a. For T7 RNA polymerase: Incubate 90 min at 37 $^{\circ}\text{C}$
- 2b. For SP6 RNA polymerase: Incubate 90 min in a 40 $^{\circ}\text{C}$ water bath.
3. Add 1 μl of 5 mg/ml DNase I. Incubate 10 min in a 37 $^{\circ}\text{C}$ water bath.
4. Adjust the final volume to 200 μl with water and mix the reaction. Remove 2 μl for Cerenkov counting, labeling the tube "cpm before."
5. Extract reaction with an equal volume of RNA phenol.
6. Transfer the aqueous phase to a new microcentrifuge tube. Add 200 μl of 5 M ammonium acetate and 2.5 μl of 20 mg/ml molecular-biology-grade glycogen to the tube. Vortex to mix.

7. Add 1 ml of 100% ethanol to the tube. Vortex to mix.
8. Microcentrifuge 15 min at 14,000 rpm. Discard the radioactive supernatant and wash the pellet with 1 ml of 70% ethanol.
9. Air dry the pellet and then resuspend in 200 μ l water.
10. Transfer 2 μ l RNA to a tube labeled “cpm after” for Cerenkov counting to determine the percentage of label that was incorporated:

$$\text{Percent incorporation} = \text{cpm}_{\text{after}} / \text{cpm}_{\text{before}} \times 100$$

Normal incorporation is 50% to 80%. A 100% incorporation will yield ~13 μ g RNA for this 100- μ l reaction—i.e., RNA amount (ng) = 13 \times percent incorporation \times 100. This in vitro transcription assay can be scaled to any size desired.

11. Dilute the RNA with water to 10 ng/ μ l.

The RNA solution can be stored at -20°C for up to one month. Note, however, that RNA labeled to high specific activity degrades more rapidly due to radiolysis.

Site-specifically labeled RNA is used to determine the site where a protein cross-links.

Prepare the nucleoprotein complexes

12. Prepare the following splicing mixture on ice:

2 μ l (20 ng) ^{32}P -labeled pre-mRNA
 1 μ l 12.5 mM ATP
 1 μ l 80 mM MgCl_2
 1 μ l 0.5 M creatine phosphate
 5 μ l water
 7.5 μ l splicing dilution buffer
 7.5 μ l HeLa nuclear extract

The reaction can be scaled to any volume desired.

It is advisable to thaw the HeLa cell nuclear extract on ice while preparing the other reagents, as it is very labile.

13. Mix gently by pipetting up and down a few times. Incubate 5 min to 1.5 hr at 30°C to form the complexes of interest.

Splicing is a dynamic process that involves the formation of a series of spliceosomal complexes. By controlling the time of incubation, it is possible to accumulate different complexes. For example, under the conditions described above, the A complex is the main complex at 15-min time point, while the B and C complexes become dominant at 45 min. At 90 min, the spliced mRNP is the main species. Similar principles also apply to other nucleoprotein complexes.

14. Pack a 1.5-cm-diameter \times 50-cm-long (89-ml bed volume) column with Sephacryl S-500 high resolution resin as described (UNIT 10.9). Place in a cold room (4°C) and equilibrate with 1 \times gel-filtration column buffer, 4°C .

15. Carefully load the assembly reaction onto the column.

The maximum loading volume of the reaction is 2.5 ml.

16. Collect 1-ml fractions, eluting at a flow rate of 0.1 ml/min, 4°C .

17. Transfer 25- μ l aliquots of fractions 21 to 80 to microcentrifuge tubes with the caps removed and analyze by Cerenkov counting.

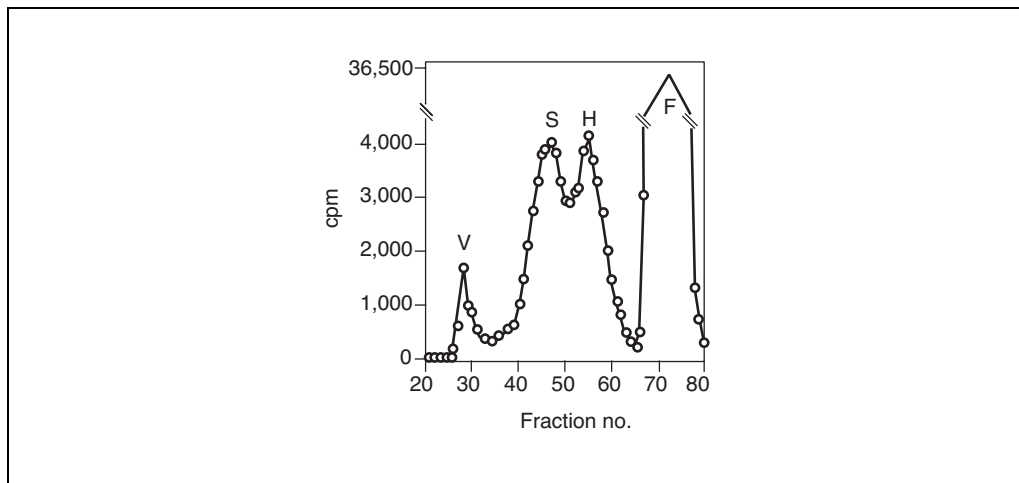


Figure 27.2.1 Isolation of spliceosomes and H complex by gel filtration. A representative gel filtration profile using a Sephacryl S500 column is shown. Peaks: V, void volume; S, spliceosomes; H, H complex; F, free counts.

18. Create an elution profile of cpm versus fraction number. Use the fractions corresponding to the desired nucleoprotein complexes for subsequent studies (see Fig. 27.2.1).

The protocol described above is for spliceosomal complexes. This protocol can also be used for preparation of other RNA-protein complexes. In this case, some parameters should be changed accordingly, including the amount of RNA used in the assembly reaction, the protein extract, the time of incubation, and the gel filtration profile.

19. Transfer 50 to 200 μl of the column fraction containing the desired ribonucleoprotein complex to a new microcentrifuge tube.

Partial purification of the nucleoprotein complexes by gel filtration is very beneficial for UV cross-linking analysis. It can greatly reduce the free nucleotides and non-specific proteins, thus increasing the cross-linking efficiency and also dramatically decreasing the background.

Cross-link the ribonucleoprotein complexes with UV light and digest the unprotected RNA

20. Place the uncovered tube containing the reaction mixture on ice, directly underneath the bulb (~5-cm from the surface) of a 254-nm UV light source. Irradiate 5 to 30 min.

The experimental setup for UV cross-linking is similar to the one described in UNIT 12.5 (Fig. 12.5.1), except that the tube containing the ribonucleoprotein complexes should be uncovered and placed on ice.

If the sample contains high levels of ATP, the time of UV exposure should be increased because ATP absorbs UV light and thus significantly reduces cross-linking efficiency.

21. Add 3 μl of 10 mg/ml protease-free RNase A and/or 3 μl of 100 U/ μl RNase T1 (see below) to each 300 μl irradiated column fraction. Incubate at 37°C for 30 min.

RNase A and RNase T1 have different cleavage specificities. RNase A digests at the 3' side of U or C nucleotides, whereas RNase T1 digests after G nucleotides. Either one of these RNases can be used based on the nucleotide sequence around the binding site of interest (if known). RNase T1 is a suspension in ammonium sulfate. Before adding it to the reaction, transfer the desired amount of RNase T1 suspension (3 μl for each 300- μl reaction) to a microcentrifuge tube. Spin at 14,000 rpm for 5 min in a microcentrifuge. Discard the supernatant. Add the irradiated reaction to the RNase T1 pellet. The activity of RNase T1 requires Mg^{2+} . If the irradiated reaction contains EDTA, MgCl_2 must be added so that 1 to 3 mM of free Mg^{2+} is present in the reaction.

Electrophorese and visualize the cross-linked proteins by phosphorimaging

22. Add 30 μ l of 20% (w/v) SDS, and 3 μ l of 2 M DTT to each 300 μ l reaction. Mix and incubate the reaction 5 min in a 65°C water bath.
23. Add 2 μ l of 20 mg/ml glycogen. Mix well.
24. Add 1.2 ml acetone and vortex to mix. Incubate the mixture 10 min at room temperature to allow efficient precipitation.
25. Centrifuge in a microcentrifuge 15 min at 14,000 rpm. Discard the supernatant and air dry the pellet.
26. Resuspend the pellet in 20 μ l of 2 \times protein-gel sample buffer or 2 \times SDS sample buffer. Incubate the samples 5 min in a boiling (100°C) water bath.
27. Electrophorese the sample through a standard SDS-PAGE gel of appropriate percentage (UNIT 10.2A) or 2-D gel (UNITS 10.3 & 10.4).

A prestained protein molecular weight marker should be included in the standard SDS-PAGE gel to determine the molecular weight of each cross-linked protein. Two-dimensional gel analysis (UNITS 10.3 & 10.4) is often used to identify these proteins, which offers further separation of the target protein from other comigrating proteins. A gel slice corresponding to a radioactively labeled protein can be subjected to mass spectrometry to reveal the identity of this protein. Please note, however, a small UV cross-linking reaction normally does not yield enough material for mass spectrometry analysis (UNITS 10.2 & 10.22). In order to increase the quantity of the target protein, crude protein extract is often mixed in with the cross-linked sample prior to electrophoresis.

28. After running the gel, cut away the region that migrates with the dye front.

The bulk of digested nucleic acid products runs with the dye front and is highly radioactive. If these products are not cut away, they can obscure signals from the cross-linked proteins migrating toward the bottom of the gel.

29. Dry the gel and place on a phosphorimaging screen. Expose overnight to visualize cross-linked proteins (APPENDIX 3A).

IMMUNOPRECIPITATION OF UV-CROSS-LINKED PROTEINS

The most obvious advantage of combining UV cross-linking, 2-D gel electrophoresis (UNITS 10.3 & 10.4), and mass spectrometry (UNITS 10.21 & 10.22) is that all proteins that bind to the RNA, either known or unknown, can be identified. However, a large amount of material is often needed to achieve this goal. In this protocol, an alternate method to determine whether a known protein is directly associated with an RNA is described. This protocol is similar to the main method (see Basic Protocol), except that after the UV-cross-linking and RNase digestion, the mixture is subjected to immunoprecipitation before electrophoresis.

Additional Materials (also see Basic Protocol)

- Protein A- or protein G-Sepharose CL-4B (Amersham Pharmacia Biotech)
- 1 \times PBS, pH 7.4 (APPENDIX 2)
- Antibody against the protein of interest
- IP100 buffer (see recipe)
- SDS
- Triton X-100
- IP150/1 M urea (see recipe)

ALTERNATE PROTOCOL

RNA-Protein Interactions

27.2.5

Couple the antibody to the beads

1. Combine 20 μ l protein A- or protein G-Sepharose CL-4B with 1 ml of 1 \times PBS, pH 7.4. Invert the tube several times. Microcentrifuge briefly at maximum speed to collect the beads in the bottom of the tube.

Protein A-Sepharose CL-4B beads are often used to couple polyclonal antibodies. Protein G-Sepharose CL-4B beads, on the other hand, are normally used with monoclonal antibodies.

2. Discard the supernatant. Add 500 μ l PBS and 20 μ l antibody against the protein of interest to the beads. Rotate 1 hr at room temperature or overnight at 4°C.

The amount of antibody used in each immunoprecipitation varies based on the quality of the antibody and the efficiency of the immunoprecipitation.

3. Microcentrifuge briefly at maximum speed to collect the beads in the bottom of the tube. Discard the supernatant. Add 1 ml PBS and invert the tube several times. Place the tube on ice to let the beads settle to the bottom.
4. Repeat step 3 four times to wash the beads. Resuspend the beads in 100 μ l IP100 buffer.

Immunoprecipitate cross-linked proteins under denaturing conditions

5. Assemble the nucleoprotein complex, partially purify the complex by gel filtration, UV-cross-link, and digest the RNA with RNase as described above (see Basic Protocol, steps 1 to 21).
6. Add SDS to 0.5% (w/v) and Triton X-100 to 1.5% (v/v). Boil the sample for 3 min.
7. Dilute the sample 10-fold with IP100 buffer. Combine the diluted sample with 20 μ l precoupled protein A- or protein G-Sepharose CL-4B (step 4). Rotate 2 to 4 hr at 4°C.
8. Wash the beads twice with IP150/1 M urea and once with IP100 as described in step 3.
9. Remove all of the buffer from the beads. Add 40 μ l of 2 \times protein-gel sample buffer or 2 \times SDS sample buffer to the beads. Boil the samples for 5 min to elute the bound proteins.
10. Electrophorese the sample through a standard SDS-PAGE of appropriate percentage (UNIT 10.2A).
11. Dry the gel and place on a phosphorimaging screen. Expose overnight to visualize cross-linked proteins (APPENDIX 3A).

The denaturing conditions used in this immunoprecipitation disrupts most of the protein-protein interactions in the ribonucleoprotein complex. Thus, only the protein recognized by the antibody is pulled down. If the radioactively labeled protein analyzed by SDS PAGE is of the correct size, it is very likely that this protein is directly associated with the RNA.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Gel-filtration column buffer, 1 \times

20 mM Tris-Cl, pH 7.6 (APPENDIX 2)
60 mM KCl
0.1% (v/v) Triton X-100
0.2 mM PMSF
0.2% (w/v) sodium azide
Store up to 3 months at 4°C

IP100 buffer

100 mM NaCl
2 mM MgCl₂
50 mM Tris·Cl, pH 7.6 (APPENDIX 2)
0.05% (v/v) NP-40
0.5 mM DTT
Store up to 3 months at 4°C

IP150/1 M urea

150 mM NaCl
2 mM MgCl₂
50 mM Tris, pH 7.6 (APPENDIX 2)
0.05% (v/v) NP-40
0.5 mM DTT
1 M urea
Store up to 3 months at 4°C

Protein-gel sample buffer, 2×

125 mM Tris·Cl, pH 7.6 (APPENDIX 2)
20% (v/v) glycerol
4% (w/v) SDS
0.008% (w/v) bromphenol blue
20 mM DTT
Store up to 1 year –20°C

Splicing dilution buffer

20 mM Tris·Cl, pH 7.6 (APPENDIX 2)
100 mM KCl
Store up to 1 year –20°C

Transcription buffer, 5×

200 mM Tris·Cl, pH 7.6 (APPENDIX 2)
30 mM MgCl₂
10 mM spermidine
Store up to 1 year –20°C

COMMENTARY**Background Information**

Cross-linking proteins to RNA with UV light is a powerful tool to identify proteins that directly bind to a particular RNA. The basis for this assay is that UV irradiation of RNA produces purine and pyrimidine free radicals. If a protein is in close proximity to the free radical (e.g., due to binding), a covalent bond can be formed, thus cross-linking the protein to the RNA. When the RNA is radioactively labeled at or close to the binding site, this label is carried with the protein, rendering it easily detectable by phosphorimaging.

In general, UV cross-linking is relatively inefficient and the formation of a photo-induced product is strongly dependent upon structural parameters, such as the proximity of reactive amino acids and bases. When a normal

RNA molecule is used in this reaction, short-wavelength UV light (254 nm) and extensive irradiation is often required, especially for those proteins that interact with the RNA weakly; however, short wavelength UV light can be damaging to the protein. To circumvent this problem, a nucleotide analog which contains a more susceptible photoreactive group is often used for the UV cross-linking assay. Depending on the nature of the photoreactive group, different long-wavelength UV lights can be chosen. Over the years, many of these analogs have been developed. They can be divided into two main categories. Analogs in the first category contain cross-linking groups directly on the nucleotide base and function essentially as zero A probes—e.g., 4-thio-UTP (Tanner et al., 1988), 5-aido-UTP (Woody et al., 1988),

5-bromo-UTP (Hanna, 1989), and 8-aido-ATP (Woody et al., 1984). Analogs in the second category contain a flexible linker between the photoreactive group and the base—e.g., 5-APAS-UTP (Hanna et al., 1989), 5-APAS-CTP (Hanna et al., 1993), and 8-APAS-ATP (Costas et al., 2000). With these analogs, proteins that are not in direct contact with the RNA can also be detected. The correct analog should be chosen based on the purpose of the experiment.

RNA containing a single ^{32}P -labeled nucleotide allows identification of proteins bound to or at the proximity of this nucleotide. Site-specific labeling of RNAs has been described in *UNIT 3.4*.

UV cross-linking between protein and nucleotides was first discovered in 1972 when Markowitz and co-workers found that UV irradiation induced the formation of protein-DNA complexes that were resistant to very stringent conditions. Since then, UV cross-linking has been used routinely to study both protein-DNA and protein-RNA interactions. Marcinovits and Molnar (1982) used UV cross-linking to study RNA-protein interactions in the 30S subunits of rat liver hnRNP. Xiang and Lee (1989) identified proteins cross-linked to RNA in the 40S ribosomal subunits of *Saccharomyces cerevisiae*. Leibold and Munro (1988) found a cytoplasmic protein bound to the 5' UTR of ferritin mRNAs. Kwon and Hecht (1991) studied translational regulation of mouse protamine 2 mRNA by analyzing proteins that cross-link to the 3' UTR of the mRNA. Ping and coworkers (1997) examined the dynamics of RNA-protein interactions in the HIV-1 Rev-RRE complex. In the pre-mRNA splicing field, this technique has also been proven to be a powerful strategy for identifying specific sites on the RNA where proteins associate within spliceosomal complexes (Reed and Chiara, 1999 and references therein). In particular, site-specific ^{32}P labeling and RNA-protein cross-linking studies have led to models for how the dynamic spliceosome assembles. For example, studies in both mammalian and yeast systems showed that a U5 snRNP protein (U5²²⁰) cross-links to a region upstream of the 5' splice site prior to the first catalytic step of splicing, and also to a region downstream of the 3' splice site after the first catalytic step, leading to the model that U5²²⁰ plays a role in aligning the 5' and 3' exons for ligation (Wyatt et al., 1992; Uman and Guthrie, 1995; Teigelkamp et al., 1995; Chiara et al., 1997). Site-specific labeling and UV cross-linking studies have also revealed that different SR proteins are arranged

in the specific 5' to 3' order on a splicing enhancer (Lynch and Maniatis, 1996).

While the UV cross-linking assay is broadly used to understand RNA-protein interactions, several key parameters in this assay, including site-specific labeling (Moore and Sharp, 1992), choices of radiolabeled nucleotide (Hartley et al., 2002), and inclusion of a photoreactive group (Costas et al., 2000), have also been intensively studied. This assay is being further developed to suit the needs of research in the RNA field.

Critical Parameters and Troubleshooting

Many critical parameters in this assay can be easily optimized, including choice of radioactive NTPs, photoreactive group substitution, conditions for assembly of the ribonucleoprotein complex, choice of gel filtration column, and time of UV irradiation. Some protein-RNA complexes are particularly sensitive to RNase, especially when the protein does not directly protect the RNA. Thus, the amount of RNase and time of incubation can be titrated down to potentially obtain a positive result.

Anticipated Results

When the RNA is body-labeled with one radioactive nucleotide (uridine is often used in this case), the specific activity of the RNA will be $\sim 6 \times 10^3$ cpm/ng. When all four radioactive nucleotides are used, the RNA will be 2.4×10^4 cpm/ng. If the RNA is site-specifically labeled at one site, one can normally obtain RNA at 3×10^3 cpm/ng.

During complex assembly in a crude protein extract, a considerable amount of the RNA can be degraded into oligonucleotides or even free nucleotides. The longer the incubation time, the more degradation will occur.

Figure 27.2.1 represents a typical gel filtration column profile of spliceosomal complexes. Spliceosomes fractionate as a large complex, whereas the nonspecific H complex, which contains mostly hnRNP proteins, runs as a small complex. Both of these complexes can be separated from most of the nuclear proteins, which co-migrate with the free ^{32}P -labeled nucleotides. When this assay is used to study other RNA-protein complexes, the column profile should be determined by analyzing the RNA species in each column fraction.

Figure 27.2.2 shows the general outcome of UV cross-linking results when different parameters are chosen in this protocol. If a body-labeled RNA is used and UV irradiation is

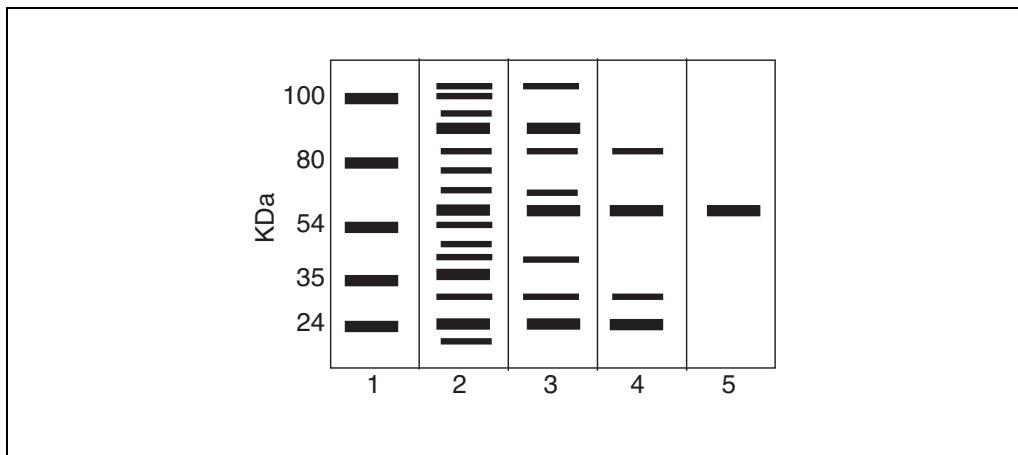


Figure 27.2.2 General outcome of autoradiograms of typical UV cross-linking experiments when different parameters (as indicated) are chosen in this protocol. Lane 1, marker; lane 2, body-labeled RNA, lane 3, site-specific-labeled RNA, lane 4, gel-filtered site-specific-labeled RNA; lane 5, gel-filtered and immunoprecipitated site-specific-labeled RNA.

carried out directly with the assembly reaction (when a crude protein extract is used for the complex assembly), multiple bands of radiolabeled proteins will appear on the gel (lane 2). If a site-specifically labeled RNA is used under similar conditions, the number of radiolabeled proteins will be fewer (lane 3); however, most of these bands are still due to nonspecific association. If a site-specific-labeled RNA is used, combined with partial purification of the ribonucleoprotein complex by gel filtration prior to UV irradiation, a much higher specificity can be achieved (lane 4). The proteins that are radiolabeled are most likely directly associated with the RNA around the labeled nucleotide. To directly ask the question of whether a protein binds to a particular site on the RNA, RNA specifically labeled at this site is often used. Immunoprecipitation with an antibody against the protein is then carried out after UV irradiation and RNase digestion. If the protein does bind to the RNA at this site, only a single band at the correct size of the protein should be detected (lane 5).

Time Considerations

In vitro transcription to prepare the radiolabeled RNA takes 1 to 2 hr. Assembly of the ribonucleoprotein complex takes 1 to 3 hr depending upon the time of incubation. Isolating the complex by gel filtration takes 14 hr. However, most of this time is not labor intensive. UV cross-linking and RNase digestion takes 30 min to 1 hr. SDS-PAGE takes 2 to 3 hr. Phosphorimaging takes between 12 hr to 3 days, depending upon the specific activity of the RNA, the efficiency of cross-linking, and the extent of RNase digestion.

Using the alternative (see Alternate Protocol), coupling the antibody to the beads takes 1 to 2 hr. Immunoprecipitation takes 3 to 5 hr.

It is advisable to make a schedule ahead of time. Transcription can be done separately from the other steps since the RNA can be stored at -20°C . If the assembly reaction is carried out in the afternoon, the gel filtration can be finished overnight. UV cross-linking, RNase digestion and the subsequent steps can be finished within the next day.

Literature Cited

- Chiara, M.D., Palandjian, L., Feld Kramer, R., and Reed, R. 1997. Evidence that U5 snRNP recognizes the 3' splice site for catalytic step II in mammals. *E.M.B.O. J.* 16:4746-4759.
- Costas, C., Yuriev, E., Meyer, K.L., Guion, T.S., and Hanna, M.M. 2000. RNA-protein cross-linking to AMP residues at internal positions in RNA with a new photocross-linking ATP analog. *Nucl. Acids Res.* 28:1849-1858.
- Hanna, M.M. 1989. Photoaffinity cross-linking methods for studying RNA-protein interactions. *Methods Enzymol.* 180:383-409.
- Hanna, M.M., Dissinger, S., Williams, B.D., and Colston, J.E. 1989. Synthesis and characterization of 5-[(4-azidophenacyl)thio]uridine 5'-triphosphate, a cleavable photo-cross-linking nucleotide analogue. *Biochemistry.* 28:5814-20.
- Hanna, M.M., Zhang, Y., Reidling, J.C., Thomas, M.J., and Jou, J. 1993. Synthesis and characterization of a new photocross-linking CTP analog and its use in photoaffinity labeling E. coli and T7 RNA polymerases. *Nucl. Acids Res.* 21:2073-2079.
- Hartley, R., Le Meuth-Metzinger, V., and Osborne, H.B. 2002. Screening for sequence-specific RNA-BPs by comprehensive UV cross-linking. *B.M.C. Mol. Biol.* 7:3:8.

- Kwon, Y.K. and Hecht, N.B. 1991. Cytoplasmic protein binding to highly conserved sequences in the 3' untranslated region of mouse protamine 2 mRNA, a translationally regulated transcript of male germ cells. *Proc. Natl. Acad. Sci. U.S.A.* 88:3584-3588.
- Leibold, E.A. and Munro, H.N. 1988. Cytoplasmic protein binds in vitro to a highly conserved sequence in the 5' untranslated region of ferritin heavy- and light-subunit mRNAs. *Proc. Natl. Acad. Sci. U.S.A.* 85:2171-2175.
- Lynch, K.W. and Maniatis, T. 1996. Assembly of specific SR protein complexes on distinct regulatory elements of the *Drosophila* doublesex splicing enhancer. *Genes Dev.* 10:2089-2101.
- Marczinovits, I. and Molnar, J. 1982. Ultraviolet light-induced cross-linking of HnRNA to infor-mofoer proteins in vitro. *Mol. Biol. Rep.* 8:111-116.
- Moore, M.J. and Sharp, P.A. 1992. Site-specific modification of pre-mRNA: The 2'-hydroxyl groups at the splice sites. *Science* 256:992-997.
- Ping, Y.H., Liu, Y., Wang, X., Neenhold, H.R., and Rana, T.M. 1997. Dynamics of RNA-protein interactions in the HIV-1 Rev-RRE complex visualized by 6-thioguanosine-mediated photocross-linking. *RNA* 3:850-860.
- Reed, R. and Chiara, M. D. 1999. Identification of RNA-protein contacts within functional ribonucleoprotein complexes by RNA site-specific labeling and UV cross-linking. *Methods* 18:3-12.
- Tanner, N.K., Hanna, M.M., and Abelson, J. 1988. Binding interactions between yeast tRNA ligase and a precursor transfer ribonucleic acid containing two photoreactive uridine analogues. *Biochemistry* 27:8852-8861.
- Teigelkamp, S., Newman, A.J., and Beggs, J.D. 1995. Extensive interactions of PRP8 protein with the 5' and 3' splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. *E.M.B.O. J.* 14:2602-2612.
- Umen, J.G. and Guthrie, C. 1995. Prp16p, Slu7p, and Prp8p interact with the 3' splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. *RNA* 1:584-597.
- Woody, A. Y., Vader, C.R., Woody, R.W., and Haley, B.E. 1984. Photoaffinity labeling of DNA-dependent RNA polymerase from *Escherichia coli* with 8-azidoadenosine 5'-triphosphate. *Biochemistry* 23:2843-2848.
- Woody, A. Y., Evans, R.K., and Woody, R.W. 1988. Characterization of a photoaffinity analog of UTP, 5-azido-UTP for analysis of the substrate binding site on *E. coli* RNA polymerase. *Biochem. Biophys. Res. Commun.* 150:917-924.
- Wyatt, J.R., Sontheimer, E.J., and Steitz, J.A. 1992. Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. *Genes Dev.* 6:2542-2553.
- Xiang, R.H. and Lee, J.C. 1989. Identification of proteins cross-linked to RNA in 40S ribosomal subunits of *Saccharomyces cerevisiae*. *Biochimie* 71:1201-1204.

Key References

Hanna, M.M. 1989. See above.

A seminal paper describing the use of several photoreactive nucleotide analogues in the UV cross-linking assay.

Hartley et al., 2002. See above.

Describes the ups and downs of the different choices of radiolabeled nucleotides in the UV cross-linking assay.

Reed and Chiara, 1999. See above.

Contains a very detailed description of the UV cross-linking assay combined with site specific labeled RNA to study RNA-protein interactions in functional spliceosomes.

Contributed by Ming-Juan Luo and Robin Reed
Harvard Medical School
Boston, Massachusetts

Purification of Functional RNA-Protein Complexes Using MS2-MBP

UNIT 27.3

**BASIC
PROTOCOL**

This protocol describes the use of MS2-MBP to purify functional RNA-protein complexes, a method also known as maltose-binding protein (MBP)-affinity purification. This approach combines gel filtration and an affinity-chromatography strategy using the bacteriophage MS2 coat protein, which binds to a specific RNA-hairpin structure. The experimental design is illustrated in Figure 27.3.1. This method has been used to isolate human spliceosomes (complexes composed of pre-mRNA, small nuclear RNAs, and proteins which are responsible for the removal of introns from pre-mRNA) assembled on a well-characterized model pre-mRNA. However, the method can be adapted to isolating other RNA-protein complexes assembled *in vitro*. For the isolation, three MS2 coat protein-binding sites (hairpins) are inserted into the RNA of interest by constructing the appropriate clone. For isolation of the spliceosome, these hairpins are inserted at the 3' end of exon 2 (Fig. 27.3.1A). A fusion protein containing the MS2 coat protein and maltose-binding protein (MS2-MBP) is bound to the hairpins (Fig. 27.3.1B). This RNA/MS2-MBP complex is then assembled into spliceosomal complexes by incubation in HeLa nuclear extracts under splicing conditions (Fig. 27.3.1C). Spliceosomes are partially purified by Sephacryl S-500 gel filtration chromatography (Fig. 27.3.1D) and then affinity selected by binding to amylose resin (Fig. 27.3.1E). After washing, spliceosomes are eluted under gentle conditions using maltose elution buffer (Fig. 27.3.1F).

Materials

E. coli strain (e.g., DH5 α) containing plasmid expressing MS2-MBP fusion protein
(UNIT 16.6)

Plasmid DNA encoding desired RNA and MS2 binding sites

Buffer A: 10 mM sodium phosphate buffer, pH 7.0 (APPENDIX 2), filtered and degassed

Buffer B: 10 mM sodium phosphate buffer, pH 7.0 (APPENDIX 2)/1 M NaCl, filtered and degassed

Dialysis buffer: 10% (v/v) glycerol in PBS (see APPENDIX 2 for PBS)

[³²P]UTP (sp. act., 800 Ci/mmol; UNIT 3.4)

Splicing dilution buffer (see recipe)

1.5% agarose gel (UNIT 2.5A)

12.5 mM ATP

80 mM MgCl₂

0.5 M creatine phosphate

10 mg/ml HeLa cell nuclear extract (UNIT 12.1)

Gel filtration buffer (see recipe)

Amylose resin (New England Biolabs)

Phosphate-buffered saline (PBS; APPENDIX 2)

Maltose elution buffer (see recipe)

HiTrap Heparin HP column (Amersham Biosciences) and FPLC system (UNIT 10.10) with gradient maker

Dialysis cassettes (MWCO 10,000; Pierce)

Centricon-50 centrifugal concentrator (Amicon)

50-ml polypropylene tubes (e.g., Falcon)

Packed 1.5/50 cm Sephacryl S-500 HR column and a gel-filtration system (UNIT 10.9)

**RNA-Protein
Interactions**

27.3.1

Contributed by Zhaolan Zhou and Robin Reed

Current Protocols in Molecular Biology (2003) 27.3.1-27.3.7

Copyright © 2003 by John Wiley & Sons, Inc.

Supplement 63

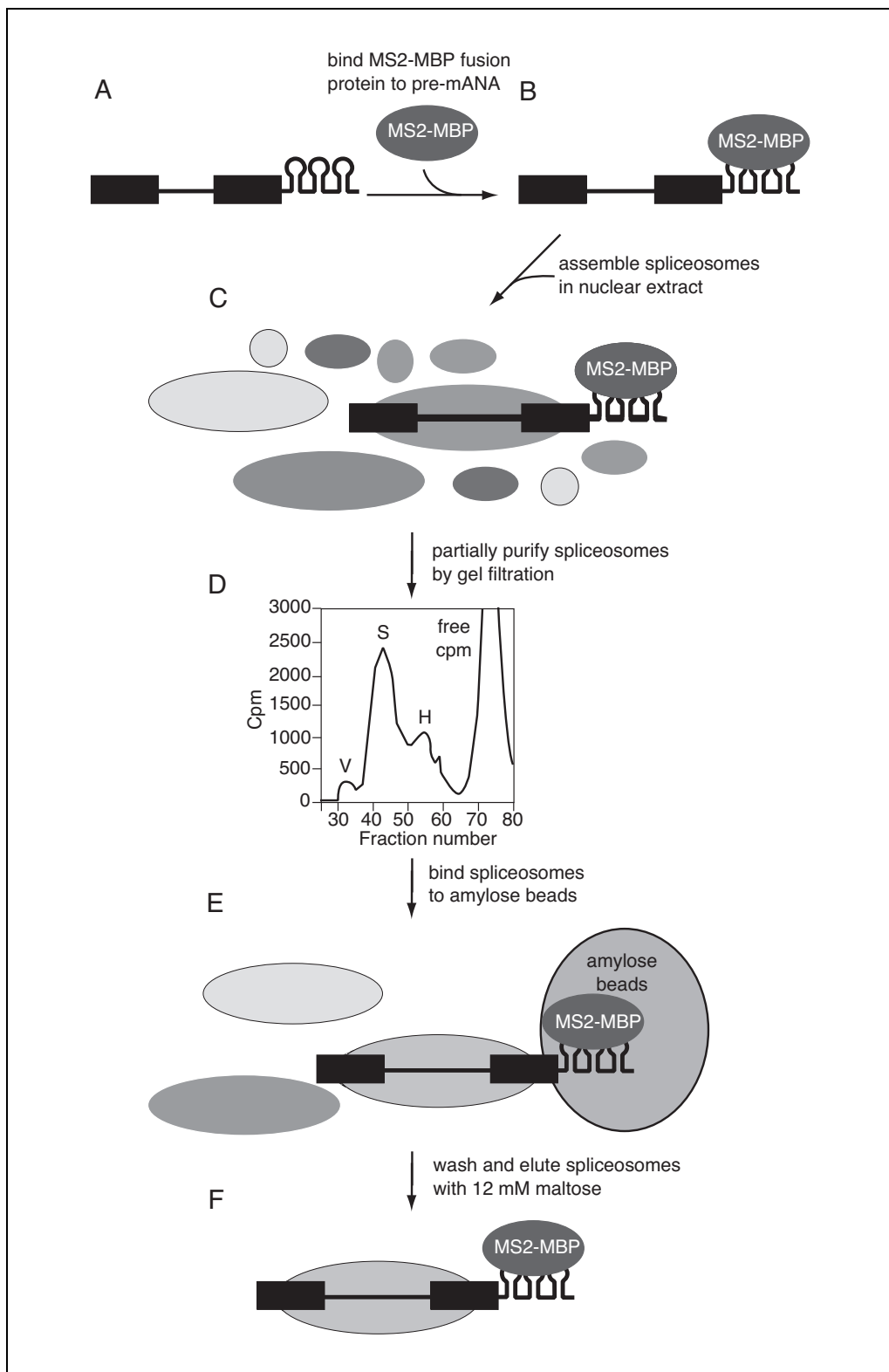


Figure 27.3.1 Experimental design of RNA-protein complex purification using MS2-MBP. **(A)** Schematic of pre-mRNA with three MS2 binding sites (hairpins) at the 3'-end of exon 2. **(B)** Pre-mRNA bound by MS2-MBP protein. **(C)** Spliceosomes are formed on pre-mRNA substrates. **(D)** Gel-filtration profile of the spliceosome purification. V indicates position of void volume; S indicates the position of spliceosomes; H indicates the position of heterogeneous RNA-protein complexes (H complex). Radiolabeled ^{32}P (cpm) in each fraction are shown in the y axis and fractions numbers are shown in the x axis. **(E)** Binding of spliceosomes to amylose beads via MS2-MBP. **(F)** Eluted spliceosomes.

Scintillation counter capable of Cerenkov counting
Data graphing software: e.g., Microsoft Excel or Cricket Graph from Computer Associates International
End-over-end rotator

Additional reagents and equipment for purification of MS2-MBP protein from *E. coli* (UNIT 16.6), dialysis (APPENDIX 3C), preparation of radiolabeled RNA (UNIT 14.3, Support Protocol 1), mobility-shift RNA-binding assay (UNIT 12.2), agarose gel electrophoresis (UNIT 2.5A), and gel-filtration chromatography (UNIT 10.9)

NOTE: All of the reagents should be RNase free (UNIT 4.1) and all procedures involving RNA-protein complex isolation should be carried out at 4°C or on ice.

Prepare MS2-MBP fusion protein and RNA substrate

1. Express and purify MS2-MBP protein as described in UNIT 16.6.

The MS2-MBP protein purified by this step binds E. coli nucleic acid and will not bind substrate RNA well. Thus, further purification of MS2-MBP via an FPLC heparin column (see step 2) is recommended.

2. Apply the MS2-MBP protein to a HiTrap Heparin column on an FPLC system per the manufacturer's recommendations.
3. Elute the MS2-MBP by a linear gradient of 0 M NaCl (buffer A) to 1 M NaCl (buffer B) in 10 mM sodium phosphate buffer, pH 7.0. Reload the flowthrough onto the column two to three times to increase the binding efficiency and final yield of MS2-MBP.

The nucleic acid will be eluted at about 100 mM NaCl and the MS2-MBP will be eluted at around 300 mM NaCl.

4. Pool the peak fractions containing MS2-MBP in a dialysis cassette (MWCO 10,000) and dialyze against dialysis buffer for 2 hr at 4°C (also see APPENDIX 3C).
5. Concentrate the protein, if necessary, using a Centricon-50 centrifugal concentrator.

The typical final concentration of MS2-MBP is 5 mg/ml. The protein can be stored at -80°C for future use.

6. Transcribe RNA substrates from the plasmid DNA bearing desired sequences and MS2 binding sites using a T7, T3, or Sp6 RNA polymerase transcription system as described in UNIT 14.3, Support Protocol 1, except label the RNA substrate with ³²P instead of ³⁵S by substituting [³²P]UTP for [³⁵S]UTP in the protocol.

The typical concentration of the RNA products is 0.2 mg/ml. The products can be stored at -80°C.

Carry out an MS2-MBP/RNA binding assay

7. Perform a binding assay similar to that described in UNIT 12.2 to check the binding efficiency of purified MS2-MBP to the substrate RNA and to optimize the molar ratio of MS2-MBP to the RNA substrate. For spliceosome purification, mix a 100-fold molar excess of MS2-MBP protein (from step 4 or 5) with pre-mRNA (i.e., the labeled RNA from step 6), or, for other complexes, prepare a series of dilutions of the MS2-MBP or the RNA, in splicing dilution buffer and incubate 30 min on ice. Assay for binding by monitoring samples on a 1.5% agarose gel.

The MS2-MBP/RNA complex will migrate more slowly than the unbound RNAs. Thus, it is recommended that the binding reaction be performed in a series of dilutions of MS2-MBP or RNA to find out the best ratio of MS2-MBP to RNA so that all of the RNA is bound by MS2-MBP.

Assembly of the RNA-protein complexes

The following steps are carried out to form spliceosomes in a 2.4-ml splicing reaction. Other RNA-protein complexes should be prepared on a large scale using appropriate cell extracts and conditions.

8. Combine 10 μl of 0.2 mg/ml radiolabeled RNA substrate (step 6) and 20 μl of 5 mg/ml MS2-MBP (step 4 or 5) in a 1.5-ml microcentrifuge tube and incubate for 30 min on ice.

The mol. wt. of the RNA used for spliceosome purification is ~132 kD; the mol. wt. for MS2-MBP is ~60 kD.

9. Add 172 μl of splicing dilution buffer to the tube, mix, and keep on ice for another 20 min.
10. Transfer the above reaction to a 50-ml polypropylene test tube and add the following reagents:

720 μl splicing dilution buffer
480 μl H₂O
96 μl 12.5 mM ATP
96 μl 80 mM MgCl₂
96 μl 0.5 M creatine phosphate
720 μl 10 mg/ml HeLa cell nuclear extract.

Gently mix the reactions and incubate at 30°C for the appropriate time.

The time is dependent upon the complex to be assembled. A pilot experiment is recommended to determine the optimal incubation time. For spliceosomal C complex formation, 40 min of incubation is necessary.

Purify the RNA-protein complexes

11. Load the mixture onto a 1.5/50 cm Sephacryl S-500 gel filtration column that has been equilibrated in gel-filtration buffer (UNIT 10.9).

This column is appropriate for the spliceosome; for other complexes it may be necessary to test a few columns to find one that gives optimal separation of the RNA-protein complex.

12. Separate the protein complex by gel filtration (UNIT 10.9, Basic Protocol 2) at a flow rate of 1.0 ml/min, collecting the first 80 1.0-ml fractions.

The gel-filtration step takes about 14 hr. Thus, it is recommended that the complexes be assembled in the afternoon and the gel filtration be run overnight.

13. While the column fractions are still in the cold room, remove 25 μl from each fraction starting at no. 25 and continuing through no. 80. Determine the cpm in each fraction by Cerenkov counting. Draw a graph of the spliceosome separation using, e.g., Excel or Cricket with the fraction number as the x axis and cpm as the y axis and determine the peak of the spliceosome as shown in Figure 27.3.1D.

The distribution of RNAs and proteins after gel filtration can be analyzed by denaturing polyacrylamide gel to check RNAs and SDS-PAGE to check proteins (UNIT 10.2A).

14. Pool fractions from the desired peak(s) containing the RNA-protein complexes.
15. Wash amylose resin in PBS by mixing 1 vol beads with 10 vol of PBS. Prepare a 50% (v/v) suspension of the amylose resin in PBS. Add 30 to 60 μl of the suspension per ml of pooled fractions (the actual bead volume will be ~15 to 30 μl /ml fractions). Incubate 4 hr to overnight at 4°C with rotation on an end-over-end rotator.

16. Centrifuge 3 min at $2000 \times g$, 4°C , in a tabletop centrifuge, then carefully remove the supernatant.
17. Wash the amylose resin (now containing the spliceosome complex) by adding 10 ml of gel-filtration buffer and rotating the tubes 10 to 20 times. Spin down the resin as described in step 16 and discard the supernatant. Repeat this washing step 3 to 5 times. Finally, transfer the washed amylose resin to a new 1.5-ml microcentrifuge tube.
18. Carefully remove all liquid from the amylose resin. Add 300 μl maltose elution buffer and rotate 30 to 60 min at 4°C .

If it is necessary to keep the complex highly concentrated, a smaller volume of maltose elution buffer can be used, but the elution efficiency will be lower.

19. Centrifuge the resin as in step 16. Retain the supernatant, which contains the purified spliceosomes.
20. Repeat the elution (steps 17 to 19) as desired to recover more complexes.

Typically, 75% of the complexes will be eluted in the first elution followed by 15% in the second elution.

REAGENTS AND SOLUTIONS

Use Milli-Q-purified water in all recipes and protocol steps. For common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

Gel filtration buffer

20 mM HEPES, pH 7.9
60 mM NaCl
0.1% (v/v) Triton X-100
0.01% (w/v) NaN_3
Store up to 2 months at 4°C

Maltose elution buffer

20 mM HEPES, pH 7.9
60 mM NaCl
10 mM 2-mercaptoethanol
12 mM maltose
1 mM PMSF (optional)
Store up to 1 week at 4°C

Splicing dilution buffer

20 mM HEPES, pH 7.9
100 mM KCl
Store up to 1 year at -20°C

COMMENTARY

Background Information

Biological machines composed of proteins and RNAs play essential biological roles. For example, the spliceosome, which is composed of pre-mRNA, five small nuclear RNAs (snRNAs), and ~145 proteins, is responsible for the removal of introns from pre-mRNA, an essential step in eukaryotic gene expression. Over the past 20 years, a great deal of progress

has been made in identifying spliceosomal components and critical RNA-RNA, RNA-protein, and protein-protein interactions. However, much less is known about the precise timing of these interactions during spliceosome assembly. In addition, little is understood about the factors that mediate the dynamics. Progress in understanding the mechanisms of pre-mRNA splicing has been hampered, in part,

because it has not been possible to isolate spliceosomes that are both highly purified and functional. Thus, it is difficult to establish the functional significance of many of the factors present in the spliceosomes.

The previously used biotin/avidin spliceosome purification method (Reed, 1990; Bennett et al., 1992) is inefficient, and requires high salt (250 mM), which dissociates some of the spliceosomal components. In addition, the bound complexes cannot be released from the affinity resin in an intact form. The MBP-MS2 method overcomes these problems, allowing efficient isolation of functional complexes in a highly pure form.

Other affinity tags were screened for their suitability for spliceosome purification. These studies revealed either high background (protein A tag) or low binding (GST or histidine tag). In contrast, significant levels of binding and low background were obtained with the maltose-binding protein (MBP). MBP affinity offers the added advantage of mild elution conditions, requiring only maltose. The MBP is fused to bacteriophage R17/MS2 coat protein, which binds to a specific hairpin RNA (Carey et al., 1983; Bardwell and Wickens, 1990; LeCuyer et al., 1995; Graveley and Maniatis, 1998). Thus, after engineering the hairpin structures into RNA, the RNA-protein complexes can be affinity-selected via MS2-MBP. A gel-filtration step that had been previously established to separate spliceosomes (Reed, 1990) was combined with MS2-MBP affinity selection. This dramatically improved the specificity of the spliceosome purification and indicated promise for isolation of other types of RNA-protein complexes. In addition, the number of MS2 binding sites and the ratio of MS2-MBP to RNA were both optimized to increase the efficiency of complex formation and purification (Zhou et al., 2002a).

Using the MS2-MBP purification method, highly purified and functional spliceosomal complexes were obtained (Das et al., 2000; Zhou et al., 2002b), the protein components were identified by mass spectrometry, and the structure of complexes were visualized by electron microscopy (Zhou et al., 2002a,b). The spliced mRNP complex that functions in nucleocytoplasmic export has also been isolated, and it retains its export activity (Zhou et al., 2000).

Critical Parameters

The efficacy of formation of the RNA-protein complex of interest is a major factor in determining the final yield of the purified com-

plexes. In the case of spliceosome purification, nuclear extract is of great importance. High-quality nuclear extract allows higher efficiency of spliceosome assembly and less RNA degradation. Another important factor is the quality of the MS2-MBP protein. Purification of MS2-MBP via a heparin column is critical for improving the binding efficiency of MS2-MBP to RNA and the final yield of the complexes. The last critical parameter is the negative control of the experiment. Generally, a similar RNA with a mutation in a critical functional site should be used in the same purification procedure to identify nonspecifically bound proteins.

Troubleshooting

The most common problem is low yield of purified RNA-protein complexes in the elution. This can be caused by insufficient binding of MS2-MBP to the RNA, inefficient binding of RNA-protein complexes to the amylose resin, poor elution, or simply inefficient formation of RNA-protein complexes. Any degradation of the RNAs or proteins will reduce or eliminate the yield as well. Thus, it is critical to perform stepwise assays to make sure that every step is working. For example, the MS2-MBP protein should be checked by SDS-PAGE to make sure it is purified and intact. A mobility gel-shift assay is necessary to optimize the binding of MS2-MBP to RNA. The RNA integrity should be monitored from the beginning of preparation of RNA, after RNA is bound by MS2-MBP, after RNA/protein complex formation, and during gel filtration and binding to the amylose beads. It is suggested that an aliquot of the reaction be saved at every step for RNA and protein quality assurance. The binding efficiency of the complex to the amylose beads and the elution efficiency of the complex can be monitored by comparing total radioactivity before and after the binding or elution.

Anticipated Results

In the case of spliceosome purification, typically about 30% of the spliceosomes will bind to the amylose resin and 80% of them will be eluted. The limiting factor in most instances is the efficiency of spliceosome assembly, which mainly depends on the quality of the nuclear extract and the sequences of the pre-mRNA substrate.

Time Considerations

Purification of MS2-MBP and preparation of substrate RNA takes one day. Assembly and purification of spliceosomes will take another

day. The time-consuming part of this protocol is the gel-filtration step, which takes ~14 hr, and it is best to schedule it to run overnight.

Literature Cited

- Bardwell, V.J. and Wickens, M. 1990. Purification of RNA and RNA-protein complexes by an R17 coat protein affinity method. *Nucl. Acids Res.* 18:6587-6594.
- Bennett, M., Michaud, S., Kingston, J. and Reed, R. 1992. Protein components specifically associated with prespliceosomes and spliceosome complexes. *Genes Dev.* 6:1986-2000.
- Carey, J., Cameron, V., de Haseth, P.L., and Uhlenbeck, O.C. 1983. Sequence-specific interaction of R17 coat protein with its ribonucleic acid binding site. *Biochemistry* 22:2601-2610.
- Das, R., Zhou, Z. and Reed, R. 2000. Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. *Mol. Cell* 5: 779-787.
- Graveley, B.R. and Maniatis, T. 1998. Arginine/Serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. *Mol. Cell* 1:765-771.
- LeCuyer, K.A., Behlen, L.S. and Uhlenbeck, O.C. 1995. Mutants of the bacteriophage MS2 coat protein that alter its cooperative binding to RNA. *Biochemistry* 34:10600-10606.
- Reed, R. 1990. Protein composition of mammalian spliceosomes assembled in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 87: 8031-8035.
- Zhou, Z., Luo, M-J., Straesser, K., Katahira, J., Hurt, E., and Reed, R. 2000. The protein Aly links pre-messenger RNA splicing to nuclear export in metazoans. *Nature* 407:401-405.
- Zhou, Z., Sim, J., Griffith, J., and Reed, R. 2002a. Purification and electron microscopic visualization of functional human spliceosomes. *Proc. Natl. Acad. Sci. U.S.A.* 99:12203-12207.
- Zhou, Z., Licklider, L.J., Gygi, S.P., and Reed, R. 2002b. Comprehensive proteomic analysis of the human spliceosome. *Nature* 419:182-185.

Contributed by Zhaolan Zhou and Robin Reed
Children's Hospital and Harvard Medical School
Boston, Massachusetts

RNA Immunoprecipitation for Determining RNA-Protein Associations In Vivo

UNIT 27.4

Chromatin immunoprecipitation (ChIP; UNIT 21.3) is a powerful and widely applied technique for detecting the association of individual proteins with specific genomic regions in vivo. RNA immunoprecipitation (RIP) is very similar, but, as the name implies, this latter method focuses on protein-RNA interactions. In RIP (as in ChIP), live cells are treated with formaldehyde to generate protein-protein, protein-DNA, and protein-RNA cross-links between molecules in close proximity. A whole-cell extract is prepared in the presence of RNase inhibitors to maintain the integrity of RNA, and the cross-linked nucleic acids are sheared by sonication to reduce average fragment size to ~500 bases. The extract is then treated with DNase I to remove DNA, and the resulting material is immunoprecipitated with an antibody against the protein of interest. RNA sequences that directly or indirectly cross-link with a given protein are selectively enriched in the immunoprecipitated sample. Reversal of the formaldehyde cross-linking by heating permits the recovery and quantitative analysis of the immunoprecipitated RNA by reverse transcription (RT) PCR analysis. The amounts of a specific RNA sequence in control and immunoprecipitated samples are determined individually by quantitative RT-PCR. The fold enrichment of certain RNA sequences (e.g., presumed binding sites) relative to other sequences (e.g., presumed nonbinding sites) provides quantitative information about the relative level of association of a given protein with different regions. Protein association with specific RNA regions can be assayed under a variety of conditions (e.g., environmental change or cell-cycle status) and/or in wild-type versus mutant strains. Furthermore, as formaldehyde inactivates cellular enzymes essentially immediately upon addition to cells, RIP provides snapshots of protein-RNA interactions at a particular time point, and is therefore useful for kinetic analysis of events occurring on RNA in vivo. In addition, like ChIP, RIP can theoretically be combined with microarray technology to identify the location of specific proteins on a “genome”-wide basis (see Commentary). This unit describes the RIP protocol for cells of the baker’s yeast *Saccharomyces cerevisiae* (see Basic Protocol). It is also applicable to other organisms, although some organism-specific modifications related to cell lysis and sonication will probably be necessary.

RNA IMMUNOPRECIPITATION IN YEAST CELLS

In this RIP protocol, *Saccharomyces cerevisiae* cells are cross-linked with formaldehyde, then harvested and sonicated to release and minimize the length of nucleic acid molecules. DNA is removed by treatment with DNase I, and the RNA is protected by the addition of RNase inhibitor throughout. The protein-RNA molecules are then immunoprecipitated and purified, the cross-links are reversed, and the resulting RNA is analyzed by RT-PCR.

Materials

- Saccharomyces cerevisiae* cells to be studied (see Chapter 13)
- 37% formaldehyde (store up to 1 year at room temperature)
- 2 M glycine, sterilized by autoclaving
- Tris-buffered saline (TBS; APPENDIX 2), ice cold
- FA lysis buffer (see recipe), ice cold and room temperature
- 40 U/ μ l RNasin (Promega)
- ~0.5-mm-diameter silica-zirconia (preferably BioSpec) or glass beads
- Ice/salt mixture in beaker for cooling
- MgCl₂

**BASIC
PROTOCOL**

**RNA-Protein
Interactions**

27.4.1

CaCl₂
 20 mg/ml RNase-free DNase I (Sigma)
 0.5 M EDTA (*APPENDIX 2*)
 Primary antibody against protein or epitope of interest
 50% (v/v) protein A–Sepharose beads (Amersham Biosciences or equivalent) in
 FA lysis buffer containing 1 mg/ml BSA
 FA lysis buffer (see recipe) containing 1 mg/ml BSA
 FA500 (see recipe)
 LiCl wash solution (see recipe)
 TE/100 mM NaCl (see recipe)
 ChIP elution buffer (see recipe)
 5 M NaCl, sterilized by autoclaving
 20 mg/ml proteinase K (Roche) in TBS/50% glycerol (store up to 1 year at –20°C)
 Acid-equilibrated 5:1 phenol/chloroform, pH 4.7 (Sigma, cat. no. P1944)
 Phase Lock Gel, Heavy (Eppendorf)
 3 M sodium acetate, pH 5.5 (*APPENDIX 2*)
 Glycogen
 Absolute ethanol, ice cold
 70% ethanol
 TE buffer, pH 7.5 (*APPENDIX 2*)
 Titan One-Tube RT-PCR kit (Roche)
 6% acrylamide/bisacrylamide (19:1) nondenaturing PAGE gel prepared in TBE
 buffer (see *UNIT 2.7*)
 SYBR Green (Molecular Probes)
 500-ml Erlenmeyer flask
 Platform rocker
 50-ml centrifuge tubes
 Refrigerated centrifuge
 1.5-ml (nonstick) microcentrifuge tubes, certified RNase-free
 FastPrep benchtop cell disruptor (Qbiogene)
 Hypodermic needle
 2-ml microcentrifuge tubes
 15-ml conical polypropylene centrifuge tubes, disposable (e.g., Falcon)
 Sonicator with microtip probe (e.g., Branson Sonifier 250) End-over-end rotator
 Spin-X microcentrifuge tube filters (Corning, available, e.g., from Sigma)
 42° (optional) and 65°C water baths
 Thermal cycler
 Additional reagents and equipment for growth of *Saccharomyces cerevisiae*
 cultures (*UNITS 13.1 & 13.2*), determining chromatin fragment size (*UNIT 21.3*),
 phenol/chloroform extraction and ethanol precipitation (*UNIT 2.1A*), primer design
 for ChIP experiments (*UNIT 21.3*), oligonucleotide synthesis (*UNIT 2.11*), PCR (*UNITS*
 15.1 & 15.7), nondenaturing polyacrylamide gel electrophoresis (*UNIT 2.7*), and
 agarose gel electrophoresis and ethidium bromide staining of gels (*UNIT 2.5A*)

NOTE: Remember that this procedure is concerned with detecting RNA; therefore, great care has to be taken to avoid its degradation during handling. After harvesting of cells, use RNase-free tubes as indicated, pipet tips with aerosol-barrier filters, and solutions prepared with nuclease-free water (Ambion). Keep samples on ice.

Cross-link protein-RNA complexes in vivo

1. In a 500-ml Erlenmeyer flask, grow 100 ml of yeast cell culture (*UNIT 13.2*) to mid-logarithmic phase in the appropriate medium (*UNIT 13.1*). Measure the cell density (*UNIT 13.2*) before formaldehyde fixation.

A concentration of $0.5\text{--}2 \times 10^7$ yeast cells/ml is usually adequate. The total volume of culture can be reduced (50 ml is a reasonable minimum) or increased depending on need. Typically, 2 to 10 ml of yeast culture ($0.2\text{--}1 \times 10^8$ cells) are used for an individual immunoprecipitation. A larger volume permits multiple immunoprecipitations from the same cells, which is particularly useful for experiments involving the analysis of multiple factors or for carrying out independent immunoprecipitations involving the same factor for data reproducibility.

2. Add 2.7 ml of 37% formaldehyde (for a final concentration of 1%). Cross-link by incubating 15 to 20 min at room temperature, occasionally swirling the flask or shaking slowly on a platform rocker.

CAUTION: Keep cultures covered or work in a fume hood to avoid noxious formaldehyde fumes.

3. Add 10 ml of sterile 2 M glycine (for a final concentration of ~ 0.2 M) and incubate an additional 5 min at room temperature.

Glycine stops the cross-linking by reacting with formaldehyde.

Harvest cells

4. Transfer cells to a 50-ml centrifuge tube and centrifuge 5 min at $2500 \times g$, 4°C . Discard supernatant into a chemical waste container and resuspend pellet in 50 ml ice-cold TBS. Repeat once.
5. Centrifuge cells a third time for 5 min at $2500 \times g$, 4°C , and discard supernatant. Supplement FA lysis buffer with RNase inhibitor by adding 1 μl of 40 U/ μl RNasin to 1 ml ice-cold FA lysis buffer. Resuspend cells in 0.5 ml of this solution and transfer to a nonstick, RNase-free microcentrifuge tube.

Lyse cells

6. Microcentrifuge cell suspension 10 sec at maximum speed, 4°C , and discard the supernatant. Resuspend the cell pellet in 400 μl ml ice-cold FA lysis buffer containing RNase inhibitor (see step 5). Add 400 μl of dry, solid silica-zirconia or glass beads.

The cells can remain on ice for a few hours while other samples are being collected, so that all samples may be processed as a group from this point onward. Alternatively, the cells may be frozen in liquid nitrogen or a dry ice/ethanol bath and stored up to several months at -80°C . This is particularly helpful if multiple samples are being generated during a time-course experiment. If cells are frozen, they must be thawed on ice before continuing with the procedure.

7. Lyse cells by applying a FastPrep cell disruptor four to five times for 30 sec each at speed 5.5 (total breakage time 2 to 3 min), removing the sample, and incubating 1 min in an ice-water bath after each treatment.

Alternatively, cell extracts can be prepared using a Mini Bead Beater (BioSpec) or by vortexing, as described in UNIT 21.3.

8. Puncture a hole in the bottom of the tube with a hypodermic needle, place the tube into a 2-ml microcentrifuge tube, and microcentrifuge 1 min at 5000 rpm.
9. Transfer liquid to a 15-ml conical polypropylene centrifuge tube and add 1 ml FA lysis buffer containing RNase inhibitor (see step 5).

Shear nucleic acids

10. Place the tube in an ice/salt mix in a beaker to maintain temperature. Holding the microtip of a Branson Digital Sonifier near the bottom of the tube to prevent foaming, sonicate the sample for one or two cycles at 50% amplitude for 15 sec. Keep the sample immersed in an ice/salt bath between and during sonications; leave the samples in the ice/salt bath to cool for at least 2 min after each sonication cycle.

Take great care that the sample does not get too hot, as this may reduce the efficiency of subsequent steps.

If a different sonication device is used, empirically determine the conditions necessary to achieve the desired level of DNA shearing.

11. Add the following to 500 μl of the sonicated extract:

MgCl₂ to 25 mM final
CaCl₂ to 5 mM final
3 μl of 40 U/ μl RNasin
6 μl of 20 mg/ml RNase-free DNase I.

Incubate at 37°C for 15 min.

12. Stop the reaction by adding 20 μl of 0.5 M EDTA (for a final concentration of 20 mM). Microcentrifuge 5 min at maximum speed and retain the supernatant. Store in aliquots (stable for many weeks) at -70°C .

These steps remove DNA so that the subsequent precipitation and detection steps focus on protein associated with RNA only. The resulting supernatant constitutes the input sample for the subsequent immunoprecipitation. The frozen aliquots are stable for many weeks when stored at -70°C and are suitable for immunoprecipitation.

13. Check RNA fragment size by electrophoresis on a polyacrylamide gel (see UNIT 21.3 for checking fragment size).

RNA pieces should be between 100 to 1000 bases, with an average length of 300 to 500 bases. It is important to shear RNA down to an average length of 400 to 500 bases. Longer fragments will increase the background and decrease the resolution of the region where the protein associates.

Immunoprecipitate

14. Dilute extract equivalent to $0.2\text{--}1 \times 10^8$ cells (typically 50 to 100 μl of the extract from step 12) to 250 μl with FA lysis buffer containing RNasin. To this diluted extract, add 1 to 5 μl primary antibody against the protein or epitope of interest and incubate overnight on an end-over-end rotator at 4°C.

The actual amount and concentration of the individual antibody has to be empirically determined and can vary considerably. The immunoprecipitation conditions can be varied (e.g., with respect to time, temperature, salt concentration, and/or presence of detergents) if necessary.

15. Microcentrifuge 15 min at maximum speed, 4°C, then transfer the supernatant to a new tube.

16. Add 20 μl of a $0.5\times$ slurry of protein A–Sepharose beads (equilibrated in FA lysis buffer containing 1 mg/ml BSA and 40 U/ μl RNasin). Incubate 1 to 2 hr at 4°C.

Protein A–Sepharose beads are used here because they work well with most monoclonal antibodies derived from mouse and polyclonal sera derived from rabbit. In some cases, the use of other beads (e.g., protein G–Sepharose) may improve binding of some antibodies, including rat IgG.

Wash beads

17. Microcentrifuge 2 min at 3000 rpm, 4°C, and remove supernatant. Resuspend beads in 700 μl FA lysis buffer, room temperature, and transfer mixture into a Spin-X centrifuge tube filter.

The use of Spin-X filters aids in the recovery of the beads after washes and results in better uniformity between different samples. The procedure is also substantially faster with the filters, particularly when multiple samples are processed simultaneously. Alternatively,

conventional microcentrifuge tubes can be used for the washes, and the supernatant can be carefully aspirated after each spin.

18. Place the filter into a 1.5-ml microcentrifuge tube. Mix sample 3 min on an end-over-end rotator, then microcentrifuge 2 min at 3000 rpm, room temperature, and discard the flow-through liquid at the bottom of the tube.
19. Add 700 μ l FA lysis buffer, room temperature, to the beads in the filter and repeat the procedures described in step 18 for a second wash.
20. Wash beads successively, each time using the technique described in step 18, with the following solutions:
 - 0.7 ml FA lysis buffer
 - 1 ml FA500
 - 0.7 ml LiCl wash
 - 0.7 ml TE/100 mM NaCl.

For some polyclonal antibodies, the more stringent washes in this step result in a cleaner signal, while gentle washes alone frequently lead to an unacceptably high background. For other antibodies, repeated washes with FA lysis buffer alone, which are gentler, might be more appropriate.

Elute protein-RNA from beads

21. Place filter unit containing the beads into a new 1.5-ml microcentrifuge tube and add 75 μ l of ChIP elution buffer as described above. Gently pipet up and down two or three times in order to dislodge beads from the filter. Incubate 10 min in a 37°C water bath.

A water bath is used instead of other heating apparatus in order to improve heat transfer.

22. Microcentrifuge beads 2 min at 3000 rpm, room temperature. Transfer eluate to another tube. Repeat elution with another 75 μ l of ChIP elution buffer. Pool eluates and discard filter with beads.

Reverse cross-links and purify RNA

23. Add 6 μ l of 5 M NaCl (for final concentration of \sim 200 mM) together with 20 μ g proteinase K (from 20 mg/ml stock). Incubate at 42°C for 1 hr (optional), then at 65°C for 1 hr.

The incubation at 42°C allows for proteinase K digestion of cross-linked polypeptides, while the 65°C incubation results in a reversal of the formaldehyde cross-links. Note that, because fractions (especially nonprecipitated inputs) contain significant amounts of DNase I, complete inactivation of the enzyme in this step (by digestion) is essential for successful RT-PCR.

24. Add 100 μ l nuclease-free water to the 150 μ l of RNA, then add an equal volume (250 μ l) of acid-equilibrated 5:1 phenol/chloroform, pH 4.8. Use Phase Lock Gel, Heavy, as recommended by the manufacturer for quick separation of the layers.

See UNIT 2.1A for additional discussion of DNA precipitation.

25. Transfer the resulting aqueous layer to a new tube and add 25 μ l 3 M sodium acetate, pH 5.5, 20 μ g glycogen, and 625 μ l ice-cold absolute ethanol. Leave at -80°C for 1 to 2 hr to allow the RNA to precipitate.
26. Microcentrifuge 30 min at maximum speed, 4°C, and remove the supernatant. Wash precipitate by adding 500 μ l ice-cold 70% ethanol and microcentrifuging 5 min at maximum speed, 4°C. Discard supernatant and allow pellet to air dry.

27. For storage, and to help avoid degradation of the resulting RNA, redissolve the pellet in 200 μ l TE buffer, pH 7.5, then add 2.5 vol of absolute ethanol, but no salt.

This mix can be stored for weeks at -80°C .

28. Remove an appropriate aliquot of the RNA solution (typically 10 to 50 μ l). Add 1 to 5 μ l of 3 M sodium acetate, pH 5.5, for a final concentration of ~ 0.3 M, then leave the mixture on dry ice for about 20 min to precipitate RNA. Microcentrifuge and wash pellet with 70% ice-cold ethanol and dry (see step 26). Redissolve pellet in 5 μ l of TE buffer, pH 7.5.

Perform quantitative RT-PCR and analyze products

29. Design primer pairs for the desired RNA regions to be examined (See *UNIT 21.3*) and synthesize the primer oligonucleotides (*UNIT 2.11*).
30. Perform RT-PCR (also see *UNITS 15.1 & 15.7*) using the Titan kit, scaling reactions down to a final volume of 25 μ l but otherwise following the manufacturers instructions. Use the following RT-PCR program:

Initial step:	30 min	50 $^{\circ}\text{C}$	(reverse transcription)
	2 min	93 $^{\circ}\text{C}$	(denaturation)
36 cycles:	20 sec	93 $^{\circ}\text{C}$	(denaturation)
	30 sec	53 $^{\circ}\text{C}$	(annealing)
	1 min	68 $^{\circ}\text{C}$	(extension)
Final step:	7 min	68 $^{\circ}\text{C}$	(final extension).

The Titan one-step kit incorporates reverse transcription and PCR into a single reaction.

For a typical measurement, input RNA is tested along with immunoprecipitated samples. Different dilutions of both input and immunoprecipitated material are tested to ensure that the RT-PCR reactions are in the linear range. See Critical Parameters and Troubleshooting for additional discussion.

It is crucial to include a control reaction in which the RT-PCR step is performed without adding reverse transcriptase, to ensure that the final product is due to RNA rather than contaminating DNA.

31. Separate PCR samples on a 6% acrylamide/bisacrylamide gel (*UNIT 2.7*) and visualize by staining with SYBR Green at a 1/10,000 dilution.

*SYBR Green staining is typically used, although similar results can be obtained by using ^{32}P -labeled nucleotides for the RT-PCR reaction. Gels containing radiolabeled PCR products are analyzed by autoradiography or phosphor imaging (*APPENDIX 3A*).*

ALTERNATE PROTOCOL

ANALYSIS BY REAL-TIME PCR

For real-time PCR analysis, RIP is performed exactly as described in the Basic Protocol, through step 28, whereupon one-step quantitative real-time RT-PCR reactions are carried out using the ABI Prism 7000 Sequence Detection System as described below. For basic considerations, see *UNIT 21.3*.

Additional Materials (also see *Basic Protocol*)

2 \times SYBR Green QPCR mix (Abgene AB-1162; contains DNA polymerase)
Multiscribe reverse transcriptase (Applied Biosystems)
ABI Prism 7000 Sequence Detection System, or equivalent
SYBR Green QPCR mix (Abgene AB-1162)
Software for analyzing PCR primers and products

1. Perform RNA immunoprecipitation and purify RNA (see Basic Protocol, steps 1 to 28)

2. Design primer pairs for the desired RNA regions to be examined (See *UNIT 21.3*) and synthesize the oligonucleotides (*UNIT 2.11*).

3. Set up the following reactions for PCR in a volume of 30 μ l:

15 μ l 2 \times SYBR Green QPCR mix
0.1 μ l 40 U/ μ l RNasin
0.15 μ l Multiscribe reverse transcriptase
10 pmol each primers
Template RNA (see Basic Protocol, step 28)
H₂O to 30 μ l.

4. Carry out RT-PCR using the following cycling program:

First step:	30 min	50°C	(reverse transcription)
	15 min	95°C	(activation)
40 cycles:	15 sec	95°	(denaturation)
	60 sec	60°	(annealing/extension).

5. Quantitate the relative amount of PCR products using appropriate software for the accompanying instrument (also see *UNIT 21.3*).

6. Calculate the apparent immunoprecipitation efficiency for a specific fragment by dividing the amount of RT-PCR product obtained in the immunoprecipitated sample by the amount obtained from the input RNA (see *UNIT 21.3*).

REAGENTS AND SOLUTIONS

Use nuclease-free, double-distilled water in all recipes and protocol steps. For common stock solutions, see *APPENDIX 2*; for suppliers, see *APPENDIX 4*.

ChIP elution buffer

100 mM Tris-Cl, pH 8 (*APPENDIX 2*)
10 mM EDTA
1% (w/v) SDS
Autoclave
Store up to 1 year at room temperature with the above ingredients
Just before use, add 40 U RNasin (Promega) per ml buffer

FA500

50 mM HEPES, pH 7.5
500 mM NaCl
1 mM EDTA
1% (v/v) Triton X-100
0.1% (w/v) sodium deoxycholate
Autoclave
Store up to 2 months at 4°C
Just before use, add 40 U RNasin (Promega) per ml buffer

FA lysis buffer

50 mM HEPES (adjust pH to 7.5 with KOH)
140 mM NaCl
1 mM EDTA
1% (v/v) Triton X-100
0.1% (w/v) sodium deoxycholate

continued

Autoclave
Store up to 1 year at room temperature with the above components
Just before use, add 100× protease inhibitors (see recipe) to 1× final and 40 U RNasin (Promega) per ml buffer

LiCl wash

10 mM Tris·Cl, pH 8 (*APPENDIX 2*)
250 mM LiCl
0.5% (v/v) NP-40
0.1% (w/v) sodium deoxycholate
1 mM EDTA
Autoclave
Store up to 2 months at 4°C
Just before use, add 40 U RNasin (Promega) per ml buffer

Protease inhibitor stock solutions, 100×

Dissolve the following in 50 ml of 100% ethanol and store up to 6 months at −20°C:
1.42 mg leupeptin
6.85 mg pepstatin A
0.85 mg phenylmethylsulfonyl fluoride (PMSF)
1.65 mg benzamidine

TE/100 mM NaCl

10 mM Tris·Cl, pH 8 (*APPENDIX 2*)
1 mM EDTA
100 mM NaCl
Store up to 2 months at 4°C

COMMENTARY

Background Information

RNA immunoprecipitation (RIP) is, in many ways, a natural and direct extension of chromatin immunoprecipitation (ChIP). Much valuable background information can be found in *UNIT 21.3*, which should be consulted before attempting the procedures described in this unit. RIP has been used for studying the interaction of proteins with RNA by Gilbert et al. (2004) and Hurt et al. (2004), based on the idea of modifying the ChIP technique and using fairly similar protocols. The paper by Gilbert et al. (2004) forms the basis for the protocols described here. In these papers, the RNA interactions of transcription-related proteins were studied, but other papers have used RIP or similar RNA immunoprecipitation techniques to study the interaction, e.g., of the basic RNAi machinery with noncoding centromeric RNA (Motamedi et al., 2004) or the interaction of proteins with tRNA (Huang et al., 2005).

The procedure described here was developed for immunoprecipitation of myc

epitope-tagged protein in yeast (Gilbert et al., 2004). However, as in ChIP, any specific antibody directed against a protein of interest can be used. The key difference between RIP and ChIP is the use of extensive DNase I digestion of the crude extracts prior to immunoprecipitation. Importantly, the inclusion or omission of this step makes it possible to perform both ChIP and RIP on the same starting cell population, which can often be desirable. Moreover, in RIP, the final detection step for the coprecipitated nucleic acid is RT-PCR, rather than PCR. Hence, although as with ChIP both the basic idea and experimental process of RIP are simple and straightforward, RIP has several unique features that are important to keep in mind in order to obtain meaningful results.

Several basic observations about RIP are worth recapitulating. First, in contrast to ChIP, the amount of input RNA required to give a suitable RT-PCR signal (in the logarithmic range) differs from transcript to transcript in the same RIP, because transcript levels can differ dramatically from gene to gene. This often

precludes the use of multiplex RT-PCR for detecting several different RNAs at the same time. Second, when ChIP is used for studying the interaction of proteins with a gene, association with the gene region before activation of the gene (where the protein is absent) and after activation of the gene (where it is present), is often examined. Such an experiment is not meaningful for RIP, because there is no RNA expressed from the gene under noninducing conditions.

Third, it has been repeatedly observed that, at least for non-sequence-specific RNA-binding (transcription-related) proteins, cross-linking is not always restricted to the portion of the transcript predicted by protein function. For example, transcription factors typically cross-link much more readily to the beginning than to the end of the transcript (Gilbert et al., 2004). This could, of course, indicate that transcription complexes are more abundant at the 5'-end of the gene, perhaps reflecting that a proportion of polymerases abort transcription during progression through the coding region (Kristjuhan and Svejstrup, 2004). However, an alternative explanation that nicely illustrates a fundamental and important difference between RIP and ChIP should be pointed out. Whereas DNA is likely to behave spatially more or less like a "linear" molecule, there is good reason to believe that RNA is spooled into a defined ribonucleoprotein (mRNP) structure during transcription (Huertas and Aguilera, 2003; Svejstrup, 2003). Such cotranscriptional RNA packing could potentially result in transcription factors apparently being cross-linked better to regions in the beginning of the transcript, because these regions (in contrast to the corresponding DNA regions) remain in close proximity to the transcribing polymerase and associated factors even as the transcription complex reaches the end of the gene. The 5'-end of the RNA molecule is thus in contact with the protein along the entire length of the gene, whereas sequences at the 3'-end of the RNA are obviously only in contact after they have been produced. Conversely, this packing of RNA also means that proteins known to specifically associate with the 5'-end of the transcript (such as CBP, which binds the cap of the finished transcript and exports the mRNA from the nucleus (Izaurralde et al., 1995; Gorlich et al., 1996) might also be expected to cross-link to regions in the 3'-end. Indeed, preliminary evidence for this has been obtained.

Finally, as in ChIP, it is important to use primers for a predicted "non-binding site" in control reactions. If mRNA-binding proteins are studied, these control primers might, for example, be designed to detect specific tRNAs or rRNAs. Likewise, if cotranscriptional association of proteins with pre-mRNA is studied, primers across an intron-exon junction can be used. Since splicing occurs cotranscriptionally (and introns are exclusively nuclear), it is then possible to study nucleus-specific RNA interactions (Gilbert et al., 2004). Obviously, considerations about the basic differences between genomic DNA and its different RNA products impact significantly on the potential uses of RIP for research.

Critical Parameters and Troubleshooting

Because RIP is designed to specifically detect RNA, it is obviously of pivotal importance to take special precautions to avoid RNA degradation throughout the procedure. Routine use of nuclease-free water, addition of RNasin to all buffers, and use of filtered pipet tips is recommended.

Cross-linking in vivo and choice of epitope

To obtain trustworthy results with RIP, it is extremely important to carry out good control reactions. The most simple and powerful control is to use the same antibody-coupled beads (and not just "specific antibody" versus "preimmune serum" or "beads alone") in two extracts that differ only in whether or not the target protein carries an epitope tag. This is because most RNA species are much more abundant than DNA and are present in both the nucleus and cytoplasm, so the risk of nonspecific cross-linking or nonspecific coimmunoprecipitation is much higher. Varying the formaldehyde cross-linking times has an impact on the efficiency and specificity of the reaction, and it is strongly advised that an initial time-course experiment be performed for each new factor studied.

Sonication and DNase I treatment

The formaldehyde cross-linked yeast cells are lysed by the use of a bead-beating procedure, which also has the effect of shearing nucleic acids. It is not required or desirable to perform extensive sonication of the extract for RIP, presumably because RNA is already substantially sheared by bead beating. Although, in the authors' laboratory, four rounds of sonication are typically optimal for ChIP, only one

or two rounds of sonication are used for a typical RIP experiment; however, this should be empirically determined.

The sonicated extracts are then treated with DNase I to remove DNA. It is important, at least initially, to perform control experiments during the RNA-detection phase in which reverse transcriptase is left out of the PCR reactions, to ensure that only RNA is detected in the experiment. It is also a good idea to perform a control reaction in which the extract is treated with RNase, again to ensure that it is indeed RNA that is being detected.

Immunoprecipitation

The authors of this unit have observed that immunoprecipitations are more specific when performed by overnight incubation, using unconjugated antibodies directly in DNase I-treated extracts. Protein-A Sepharose beads are then added the next day, but only after a high-speed centrifugation of the antibody-incubated extract to remove material that may have precipitated during the incubation. The authors typically use only 1 to 2 μ l antibody per immunoprecipitation reaction. Less material is precipitated in this way, but the signal-to-noise ratio is improved, which may be of particular importance for RIP.

Reversal of cross-links and RNA storage and detection

ChIP procedures typically include a very long incubation at 65°C to reverse formaldehyde cross-linking (UNIT 21.3). The authors of this unit have found that incubations as short as 0.5 to 1 hr are sufficient for reversal of RNA-protein interactions. After reversal of cross-links, RNA is precipitated. Problems of reproducibility with “stored” versus “fresh” RNA are sometimes encountered, but storing the RNA according to the procedure described in the Basic Protocol does help somewhat.

RT-PCR

There are several key parameters for achieving optimal amplification. For example, it is very important to have a quality repeat pipettor that can reproducibly dispense 2- μ l samples. Pipetting inaccuracies at this stage will lead to greater well-to-well variability and poorer reproducibility among identical samples.

The PCR conditions tabulated in the Basic Protocol are generally appropriate for most situations. The annealing temperature may have to be adjusted if the melting temperature of the primers is substantially above or below 55°C. The number of cycles might also have to be

adjusted in some cases if reactions are not in the linear range.

For a typical measurement, input RNA is tested along with immunoprecipitated samples. Different dilutions of both input and immunoprecipitated material are tested to ensure that the RT-PCR reactions are in the linear range. Because different RNAs are present in different copy numbers (in contrast to DNA), the dilutions needed to achieve reproducible and reliable results will differ between different RNA molecules. Typically, the input material needs to be diluted ~500-fold relative to immunoprecipitated material. Due to the variability in RNA copy numbers, simultaneous analysis of multiple RNAs in a single reaction is often not possible.

It is crucial to include a control reaction in which the RT-PCR step is performed without adding reverse transcriptase, to ensure that the final product is due to RNA rather than contaminating DNA.

Genome-wide analysis

ChIP-on-chip experiments have become a common feature of research reports over the last couple of years (see, e.g., Robert et al., 2004). Whereas RIP-on-chip has not yet been reported, there is no reason why such experiments cannot be performed (see UNIT 21.3 for details on ChIP-on-chip)

Anticipated Results

The described technique should be capable of identifying the interaction of a protein of interest with RNA, starting with $\sim 10^7$ yeast cells. Although the protocol was developed for the detection of RNA interactions of a pre-mRNA-interacting protein in the nucleus (Gilbert et al., 2004), the authors of this unit anticipate that it can be used for detecting RNA interactions (mRNA, tRNA, rRNA) of proteins in the cytoplasm, mitochondria, or nucleolus, among other locations. An example result is shown in Figure 27.4.1.

Since it is well-known that formaldehyde makes protein-protein cross-links, one cannot assume that a positive RIP result means that a given protein is actually bound to the RNA. A positive result means that the protein is “in the vicinity” of the RNA—for example, associated via other proteins in a multi-subunit complex. This is true for chromatin immunoprecipitation as well as for RNA immunoprecipitation.

Time Considerations

The basic protocol may be completed in 2 or 3 days. On the first day, cells are fixed with

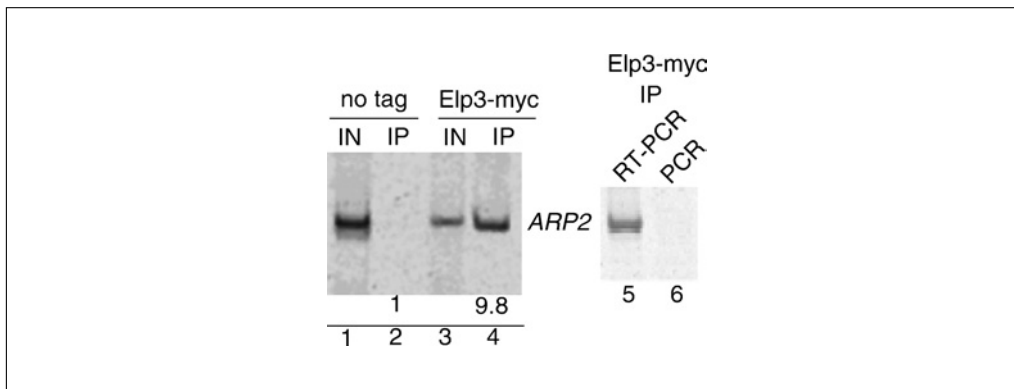


Figure 27.4.1 Results of RNA immunoprecipitation of Elp3, a subunit of the Elongator complex and relevant controls. In the left panel, immunoprecipitation of Elp3-Myc was shown to specifically co-precipitate the introns-exon junction of the *ARP2* gene. Enrichment was ~10-fold and depended on the epitope-tagged Elp3 protein (compare lanes 2 and 4). In the right panel, control experiments show that the *ARP2* product is only obtained by the use of reverse transcriptase PCR (RT-PCR), not by PCR alone (lanes 5 and 6, respectively). Reproduced from Gilbert et al. (2004) with permission from Elsevier.

formaldehyde and harvested (1 hr). Preparation of chromatin extracts (2 hr) and immunoprecipitations (primary antibody incubation, usually overnight); may be carried out on day 1, while incubation with protein A-coupled beads, ~2 hr; washing and elution, 1.5 hr, and reversal of cross-links (1-2 hours to overnight) can be done on day 2. On the final day, the DNA is purified (<4 hr), and reverse transcriptase PCR amplified (<3 hr, including 2 hr for the PCR program), and analyzed by gel electrophoresis (<2 hr, including 1 hr of gel running time). Alternatively, the samples are analyzed by reverse transcriptase real-time PCR (<3 hr, including 2 hr for the PCR program).

Literature Cited

- Gilbert, C., Kristjuhan, A., Winkler, G.S., and Svejstrup, J.Q. 2004. Elongator interactions with nascent mRNA revealed by RNA immunoprecipitation. *Mol. Cell* 14:457-464.
- Gorlich, D., Kraft, R., Kostka, S., Vogel, F., Hartmann, E., Laskey, R.A., Mattaj, I.W., and Izaurraide, E. 1996. Importin provides a link between nuclear protein import and U snRNA export. *Cell* 87:21-32.
- Huang, B., Johansson, M.J., and Bystrom, A.S. 2005. An early step in wobble uridine tRNA modification requires the Elongator complex. *RNA* 11:424-436.
- Huertas, P. and Aguilera, A. 2003. Co-transcriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell* 12:711-721.
- Hurt, E., Luo, M.J., Rother, S., Reed, R., and Strasser, K. 2004. Cotranscriptional recruitment

of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex. *Proc. Natl. Acad. Sci. U.S.A.* 101:1858-1862.

- Izaurralde, E., Lewis, J., Gamberi, C., Jarmolowski, A., McGuigan, C., and Mattaj, I.W. 1995. A cap-binding protein complex mediating U snRNA export. *Nature* 376:709-712.
- Kristjuhan, A. and Svejstrup, J.Q. 2004. Evidence for distinct mechanisms facilitating transcript elongation through chromatin in vivo. *EMBO J.* 23:4243-4252.
- Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. 2004. Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* 119:789-802.
- Robert, F., Pokholok, D.K., Hannett, N.M., Rinaldi, N.J., Chandry, M., Rolfe, A., Workman, J.L., Gifford, D.K., and Young, R.A. 2004. Global position and recruitment of HATs and HDACs in the yeast genome. *Mol. Cell* 16:199-209.
- Svejstrup, J.Q. 2003. Keeping RNA and DNA apart during transcription. *Mol. Cell* 12:538-539.

Key References

Gilbert et al., 2004. See above.

Describes the technique from which the Basic Protocol was adapted.

Contributed by Chris Gilbert and
Jesper Q. Svejstrup
Cancer Research UK
London Research Institute
Clare Hall Laboratory
Hertfordshire, U.K.

CHAPTER 28

Mammalian Cell Culture

INTRODUCTION

Cell culture technologies are a basic tool for most, if not all molecular biologists, and the use of established, immortal cell lines for production of proteins from cDNAs or other forms of genetic material has been an important and productive method for characterizing molecules identified by molecular biological techniques such as cloning, differential display, or microarray analyses (see Chapter 9, Introduction of DNA into Mammalian Cells, and Chapter 16, Protein Expression). Apart from using cell culture approaches for protein production, analysis of protein function also has made extensive use of immortal cell lines—for instance, in examining the effect of exogenous protein expression on cell signaling or transcriptional events. These approaches have been highly successful in delineating the role of proteins of interest, to the extent possible, in relatively uniform cell cultures. Moreover, a large number of different types of immortalized cell lines are available, representing material originally obtained from a large number of different tissues and species. These can be obtained through other investigators or purchased from organizations such as the American Type Culture Collection (ATCC) or commercial organizations.

However, there are limitations to the usefulness of established cells lines for studies of protein function. For instance, the ability of many of these cultures to grow without limitation is related to their origin from tumor tissues, and, consequently, these cells may have abnormal chromosomal content, more subtle genetic mutations, and abnormal protein expression. Consequently, the use of primary cell cultures obtained from normal tissues can be very desirable for studying protein function in a relatively normal cellular background. In addition, many molecular biological studies have incorporated the development of transgenic and knockout mice, and analyses of these mouse models frequently lead to questions and hypotheses that can be best explored in comparatively simpler cell culture systems. Genetically altered mice therefore serve as an invaluable source of tissue from which to isolate cells that either lack or overexpress the protein of interest in a relatively normal cellular background. Likewise, the establishment of cultures of cells obtained from human patients carrying genetic mutations provides an opportunity to study the function of mutated proteins, and, potentially, the basis for human disease. Finally, establishment of primary cell cultures is of importance for cell types in which the ability of immortalized cells to model the function of a given cell or tissue type is poor.

It is the goal of this chapter to provide protocols for the preparation of primary cell cultures from cells or tissues obtained from humans and other mammalian species; information on basic cell culture techniques, including the propagation of established, immortal cell lines can be found in *APPENDIX 3F*. To begin, this chapter presents a protocol for preparation of primary mouse embryonic fibroblasts and information on how to immortalize these fibroblasts for long-term studies (*UNIT 28.1*). This is followed by protocols for isolating lymphocytes from human blood and immortalizing these cells with Epstein-Barr virus (*UNIT 28.2*). In future supplements, we will incorporate additional units on the preparation of primary cultures from a variety of tissues from other sources, beginning with methods for preparing foreskin fibroblast cultures and continuing with protocols for establishing cultures of other specialized cell types.

Carolyn L. Smith

Contributed by Carolyn L. Smith

Current Protocols in Molecular Biology (2005) 28.0.1
Copyright © 2005 by John Wiley & Sons, Inc.

Mammalian Cell
Culture

28.0.1

Supplement 70

Preparation, Culture, and Immortalization of Mouse Embryonic Fibroblasts

UNIT 28.1

Genetically manipulated transgenic and knockout mouse lines are powerful animal models for defining the physiological function of genes during development and in organ function. Similarly, the mutant cells isolated from these mutant mouse lines are valuable tools for studying the molecular and cellular mechanisms of the mutated genes under well-defined culture conditions. The purpose of this unit is to provide convenient stepwise protocols for preparation, culture, and immortalization of mouse embryonic fibroblasts (MEFs). The complete experimental process can be divided into three major steps including: isolation of MEFs from mouse embryos (see Basic Protocol 1), culture and use of primary MEFs (see Basic Protocol 2), and immortalization of MEFs for permanent growth (see Basic Protocol 3). In addition, useful annotations and commentary relevant to the development of MEFs from mouse embryos with lethal phenotype, as well as for optimal use of MEFs in experiments are provided, based on the author's experience. Additional information for isolating MEFs for the production of mouse mutants using homologous recombination and blastocyst-mediated transgenesis can be found in *UNIT 23.2*.

NOTE: All protocols using live animals must first be reviewed and approved by an Institutional Animal Care and Use Committee (IACUC) and must follow officially approved procedures for the care and use of laboratory animals.

ISOLATION OF MOUSE EMBRYO FIBROBLASTS FROM MOUSE EMBRYOS

**BASIC
PROTOCOL 1**

Theoretically, MEFs can be isolated from mouse embryos with genetic backgrounds of any strain. However, it is recommended that wild-type control and mutant MEFs be prepared from mice with identical strain backgrounds, to reduce the risk of possible genetic background-dependent variations. If the transgenic or homozygous mutant mice are viable and fertile, wild-type and mutant embryos can be generated from wild-type and mutant mating pairs, respectively. If the homozygous mutant mice are infertile or lethal, heterozygous mating pairs should be used to produce wild-type, heterozygous, and homozygous embryos. The optimal ages of mouse embryos for MEF isolation are from embryonic day 13.5 (E13.5) to E15.5. However, if the mutant embryos are lethal at earlier stages, MEFs can be successfully isolated from mouse embryos at developmental stages as early as E8.5.

NOTE: All solutions and equipment coming into contact with cells must be sterile, and aseptic technique should be used accordingly.

Materials

- 2 to 3 male and 4 to 6 female mice (8 to 18 weeks of age)
- 2.5% Avertin (see recipe)
- 70% ethanol
- Phosphate-buffered saline, pH 7.2 (PBS; *APPENDIX 2*), sterilized by autoclaving
- 0.25% trypsin-EDTA solution (Invitrogen, cat. no. 25300-056)
- MEF culture medium (see recipe)
- Mouse ear tags (National Band and Tag)
- Cages (11 in. wide × 7 in. long × 7 in. tall) with food and water dishes
- Appropriate mouse food

**Mammalian Cell
Culture
Techniques**

Contributed by Jianming Xu

Current Protocols in Molecular Biology (2005) 28.1.1-28.1.8
Copyright © 2005 by John Wiley & Sons, Inc.

28.1.1

Supplement 70

Dissecting instruments, autoclaved:
Regular and fine dissection scissors
Two pairs autoclaved watchmaker's no. 5 forceps
Blunt fine forceps
Powder-free latex gloves
15- and 50-ml sterile screw-cap plastic conical tubes
10-cm tissue culture dishes or 75-cm² tissue culture flasks

Set up mice for mating

1. Number all mice with ear tags. Set up mouse mating cages by putting one adult male and one to two adult female mice in each cage.

Usually four to six females are required for each experiment, because not every female that has copulated can be identified or will become pregnant. Mice should be fed with appropriate food and water and housed in a temperature- and lighting-controlled animal facility.

2. Every morning after setting up mating, check all females for the presence of a copulation plug (vaginal plug) in the vagina.

The vaginal plug contains coagulated proteins from the male seminal fluid and can be easily seen in most strains. The vaginal plug usually dissolves 12 to 14 hr after mating (Hogan et al., 1994); therefore, the females should be checked before 8:00 am.

3. Transfer females with vaginal plugs into new cages. Record the age of its embryos as E0.5 on the day when the vaginal plug is seen.

Harvest fetuses

4. On day E14.5, anesthetize the plugged mouse by intraperitoneally injecting 0.5 ml of 2.5% avertin, then euthanize by cervical dislocation.

If null mutant embryos are lethal before E14.5, the plugged mice should be sacrificed earlier.

5. Place mouse on a paper towel and spray enough 70% ethanol to soak the entire mouse body. Using a pair of regular dissection scissors, cut a small incision through the skin at the centric position on the middle back and, wearing powder-free latex gloves, pull with both hands to peel the skin in both directions toward the head and tail. Lay the mouse carcass on its back on a sterile paper towel.

6. Dissect out the uterus as follows.

- a. Use a pair of sterile blunt forceps to lift the abdominal wall, then remove the abdominal wall using a pair of sterile dissection scissors.

Caution should be taken not to damage internal organs such as intestine, to avoid bacterial contamination.

The left and right uterine horns with the embryos in them should be seen.

- b. Use a pair of sterile blunt forceps to hold up the bottom part of the uterus and use a pair of sterile scissors to dissect out the entire uterus containing all embryos.
 - c. Transfer the uterus into a 50-ml tube with 30 ml sterile PBS and invert the tube several times to wash the uterus.
7. In a laminar flow cabinet, open the 50-ml tube and using blunt forceps, transfer the washed uterus with embryos into a 10-cm tissue culture dish containing 10 ml sterile PBS. Use a pair of fine watchmakers forceps to hold the uterine wall and use a pair of fine dissection scissors to cut through one side of the uterine wall to expose individual embryos.

If embryos are younger than E13.5, perform dissection procedure under a stereomicroscope.

8. Open the yolk sac with a pair of fine forceps and dissect out all fetuses intact.
9. Transfer all fetuses into a new dish containing 10 ml sterile PBS. Take each fetus and, using two pairs of fine watchmakers forceps, remove and discard the liver and heart, and also remove and discard the brain by first removing the upper part of the head. Wash the remaining part of each of the fetus in sterile PBS to remove as much blood as possible and then transfer the fetuses into a new dish.

Heterozygous parental mice produce embryos with all possible genotypes. If the genotype of individual embryos of genetically manipulated mice needs to be analyzed, handle embryos individually and match the identification number of each embryo with its biopsy for genotyping. Collect one of the hind limbs (for E13.5 and older embryos) or the yolk sac (for E12.5 and younger embryos) and extract DNA. Identify the genotype by PCR (UNIT 15.1) using allele-specific primer pairs as described in the literature (Hogan et al., 1994; Kuang et al., 2002).

Prepare MEF suspension

10. Add 2 to 3 ml of ice-cold 0.25% trypsin-EDTA to the dish. Use two pairs of fine forceps to tease all fetuses into fine pieces. Using a pipet, transfer all materials into a 15- (for up to four embryos) or a 50-ml tube (for more than four embryos), then add ice-cold 0.25% trypsin-EDTA to bring the total volume to 3 ml per embryo. Let the tube stand at 4°C overnight.

This allows trypsin, which has almost no activity at 4°C, to diffuse into the tissues.

11. On the next morning, without disturbing the tissue at the bottom of the tube, aspirate off most of the trypsin solution, leaving an amount equivalent to approximately two volumes of the tissue. Cap the tube and incubate 30 min in a 37°C water bath.
12. Add MEF culture medium to 25 ml (for a 50-ml tube) or 8 ml (for a 15-ml tube), and pipet vigorously and repeatedly up and down to break up the digested tissues into a cell suspension.
13. Allow sedimentation of the cell suspension by gravity for 1 min to let remaining clumps fall to the bottom of the tube. Transfer the supernatant cell suspension to a new tube. Add more MEF culture medium to the tube containing the remaining tissue clumps and vigorously pipet up and down for 1 min. Repeat the sedimentation and combine the supernatant cell suspension with the cell suspension already collected.
14. Mix the cell suspension and plate the cells in 10-cm tissue culture dishes or 75-cm² tissue culture flasks with 10 ml MEF culture medium. Continue with culture as in Basic Protocol 2.

The number of tissue culture plates to be used is dependent on the total number of MEFs isolated, which basically is proportional to the number and sizes of the embryos. MEFs isolated from each of the E14.5 embryos can be plated in five 10-cm dishes or 75-cm² flasks.

CULTURE AND USE OF PRIMARY MOUSE EMBRYO FIBROBLASTS

MEFs isolated from mouse embryos can be expanded in culture or cryopreserved in liquid nitrogen. They can be used as primary cultured cells or further immortalized to generate cell lines that can be maintained permanently in culture.

NOTE: All solutions and equipment coming into contact with cells must be sterile, and aseptic technique should be used accordingly.

NOTE: All culture incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

BASIC PROTOCOL 2

Mammalian Cell Culture Techniques

28.1.3

Materials

MEFs prepared from mouse embryos (see Basic Protocol 1)
Phosphate-buffered saline, pH 7.2 (PBS; APPENDIX 2), sterilized by autoclaving
MEF culture medium (see recipe)
0.05% trypsin-EDTA solution (Invitrogen, cat. no. 25300-054)
MEF freezing medium (see recipe)
Inverted phase-contrast microscope (suitable for examining the cell culture)
0.75- or 1-ml cryovials
Centrifuge and rotor with adaptors for 15- and 50-ml plastic centrifuge tubes
Liquid nitrogen tank with canes for holding cryovials
10-cm tissue culture dishes or 75-cm² tissue culture flasks
Additional reagents and equipment for mammalian cell culture and trypsinization of cells (APPENDIX 3F)

1. Incubate MEFs prepared from mouse embryos (Basic Protocol 1) overnight.
2. Aspirate medium and wash cell monolayers twice, each time with 10 ml sterile PBS, to remove all nonadherent cells. Add 10 ml fresh MEF culture medium to each plate and continue to culture cells to near confluency (as determined by examination with phase-contrast microscope).

The cells at this stage are designated as passage 1 MEFs.

APPENDIX 3F contains additional detail on mammalian cell culture.

To freeze cells for later use

- 3a. Aspirate medium and wash cell monolayers twice, each time with 5 ml sterile PBS. Add 3.5 ml 0.05% trypsin-EDTA solution to each culture plate and incubate 3 to 5 min. Add 5 ml MEF culture medium to each plate, then resuspend cells and transfer them into 15- or 50-ml sterile centrifuge tubes.

APPENDIX 3F contains additional detail on trypsinization of cells.

- 4a. Centrifuge cells 5 min at 400 × g, room temperature. Aspirate all of the supernatant.
- 5a. Add MEF freezing medium at a volume of 0.5 ml per culture plate that went into producing the pellet. Use a 5-ml pipet to resuspend cells, then place 0.5-ml aliquots of cells into cryovials.
- 6a. Freeze all vials in a –80°C freezer or in the vapor phase of the liquid nitrogen tank overnight, then transfer all vials into the liquid nitrogen tank for long-term storage.

To revive the frozen cells, transfer the cryovials containing them from the liquid nitrogen tank to the 37°C culture incubator. Incubate 10 min, then transfer all cells into a 50-ml tube containing 40 ml MEF culture medium. Mix, then divide cell suspension into four 10-cm culture plates and grow cells as described in this protocol.

To expand MEFs in culture

- 3b. After cells from step 2 have reached near confluency, aspirate medium and wash cell monolayers twice, each time with 5 ml sterile PBS.
- 4b. Add 2.5 ml of 0.05% trypsin-EDTA solution to the cells and incubate 3 to 5 min.
- 5b. Add 5 ml MEF culture medium to each plate, resuspend cells, and split into new 10-cm culture plates or 75-cm² culture flasks at a ratio of 1:5. Add sufficient MEF culture medium to bring volume to 10 ml in each plate or flask.

6b. Continue incubation.

Cells will grow to confluence in about 3 days.

Proliferation of MEFs in primary culture dramatically decreases by passage 5. If functional analysis needs to be performed with primary MEFs, experiments should be done with MEFs at passage 3, or at least no later than passage 4.

IMMORTALIZATION OF MOUSE EMBRYO FIBROBLASTS BY SERIAL PASSAGES

BASIC PROTOCOL 3

Because the life span of primary cultured MEFs is limited and isolation of primary MEFs from mouse embryos is time-consuming and labor-intensive, most laboratories wish to develop MEFs with permanent growth features. Two major approaches have been developed to immortalize primary MEFs, including serial passages of primary MEFs until they pass their growth-crisis stage and transformation of primary MEFs by overexpression of one or more oncogenes. Here, the method of serial passage is described, which is similar to the protocol used to establish the BALB/3T3 cell line (Aronson and Todaro, 1968).

NOTE: All solutions and equipment coming into contact with cells must be sterile, and aseptic technique should be used accordingly.

NOTE: All culture incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

Materials

MEFs growing in culture (see Basic Protocol 2)

MEF culture medium (see recipe)

Phosphate-buffered saline, pH 7.2 (PBS; *APPENDIX 2*), sterilized by autoclaving
25-cm² culture flasks

Additional reagents and equipment for mammalian cell culture, trypsinization of cells, and counting cells (*APPENDIX 3F*), and expansion and cryopreservation of cultured MEFs (see Basic Protocol 2)

1. Harvest MEFs from primary culture and determine of cell number by counting with a hemacytometer or an automated cell counter (*APPENDIX 3F*).
2. Mix 3.8×10^5 cells with 5 ml MEF culture medium and plate these cells in a 25-cm² culture flask. Incubate 3 days.

IMPORTANT NOTE: MEFs grow rapidly before passage 3, but their growth significantly decreases after passage 4. Therefore, it is sufficient to set up one 25-cm² flask for cells up to passage 3, but one should set up five 25-cm² flasks of MEFs after passage 3 to ensure there will be enough cells for subsequent passages.

3. After incubation, wash cell monolayer twice, each time with 5 ml PBS. Add 2.5 ml of 0.05% trypsin-EDTA solution to the cells and incubate 3 to 5 min. Add 5 ml MEF culture medium, then resuspend cells and transfer them to a 15-ml sterile centrifuge tube. Count cells (*APPENDIX 3F*) and plate again and incubate as in step 2 for the next passage.
4. Repeat the above steps for 20 to 25 passages.

Following an initial period of rapid growth, proliferation of the primary MEFs should decline from passage 5 to 14. By the 20 to 25th passage, the growth rate of MEFs should increase and the cell number also should increase 2- to 3-fold with respect to the inoculated number of cells during each 3-day transfer schedule (Fig. 28.1.1).

Mammalian Cell Culture Techniques

28.1.5

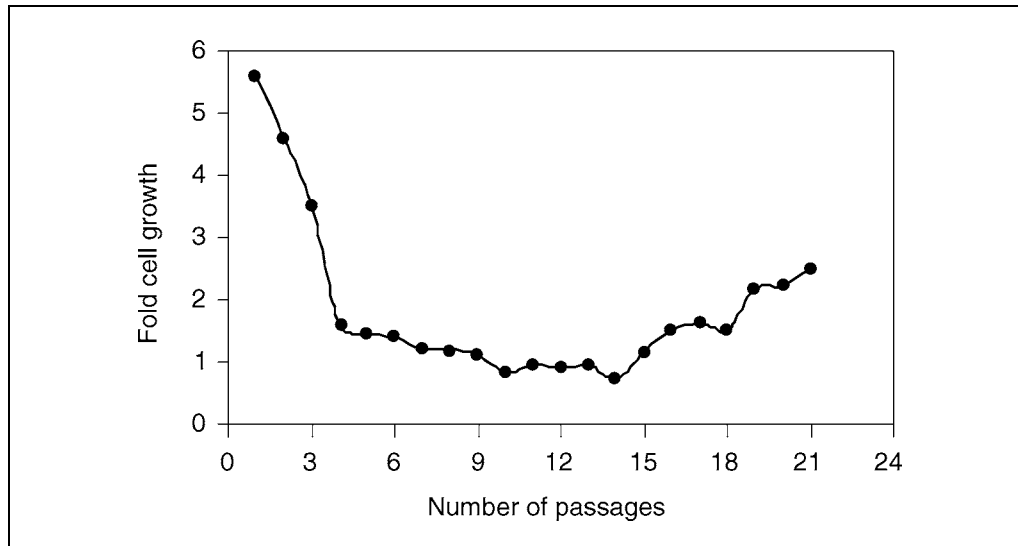


Figure 28.1.1 A representative growth curve of MEFs during serial passages. Primary MEFs were isolated from E14.5 mouse embryos with a mixed C57BL/6 and 129SvEv strain background. MEFs were cultured and transferred every 3 days as described in the protocol. The fold of cell growth was calculated by dividing the number of harvested MEFs with the number of inoculated MEFs in each passage. Data represent the average of two samples.

- After the MEFs become immortalized, expand and prepare frozen stock (see Basic Protocol 2).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Avertin, 2.5%

Mix 5 g of 2,2,2-tribromoethanol with 5 ml *tert*-amyl alcohol to make a 100% stock solution. Dilute the 100% stock solution with sterile PBS (APPENDIX 2) to a 2.5% working solution, stirring vigorously until it is completely dissolved. Wrap bottles for both the 100% stock and 2.5% working solutions in foil to avoid exposure to light and store up to 6 months at 4°C.

MEF culture medium

Dulbecco's modified Eagle medium (DMEM) containing:
 10% (v/v) fetal bovine serum (FBS)
 0.1 mM β -mercaptoethanol (tissue culture grade)
 50 U penicillin
 50 μ g/ml streptomycin
 Filter-sterilize and store up to ~3 months at 4°C.

MEF freezing medium

Dulbecco's modified Eagle medium (DMEM) containing:
 30% (v/v) fetal bovine serum (FBS)
 10% (v/v) dimethylsulfoxide (DMSO)
 Stored up to 4 to 5 months at 4°C

COMMENTARY

Background Information

MEFs growing in culture have a very high probability of developing into permanent (“immortalized”) cell lines, probably because they are susceptible to environmental stress-induced changes of chromatin stability. It was shown many years ago that MEFs could develop into permanent cell lines under appropriate culture conditions (Todaro and Green, 1963; Aaronson and Todaro, 1968). However, most of these established MEF lines were hypotetraploid (Aaronson and Todaro, 1968). It was also demonstrated that the MEF cell lines could be established from either low inoculation densities as described in Basic Protocol 3, or from a 4-fold higher inoculation density (Aaronson and Todaro, 1968). MEF cell lines developed from a low inoculation density that avoids cell-to-cell contact before transfer are able to grow at much lower inoculation densities and are extremely sensitive to contact inhibition of cell division. In addition, primary MEFs also can be transformed using viral oncoproteins such as SV40 T antigen (Todaro and Green, 1966). The transformation protocol is rapid, but the established MEF lines acquire many features similar to tumor cells.

Primary MEFs isolated by similar protocols are widely used as the feeder layer for culturing embryonic stem cells (Abbondanzo et al., 1993; Hogan et al., 1994). It is believed that MEFs can provide factors that enhance the proliferation and maintain the undifferentiated states of embryonic stem cells. In recent years, MEF cell lines established from genetically manipulated mouse embryos have become powerful tools to study the cellular function and molecular mechanisms of genes of interest through comparing mutant MEFs with wild-type MEFs (Puigserver et al., 1999; Wu et al., 2004).

Critical Parameters and Troubleshooting

To isolate MEFs (Basic Protocol 1), if embryos are at the stage of E12.5 or younger at the dissection step, there is no need to incubate the disaggregated fetal tissues with cold trypsin-EDTA solution overnight. Instead, tease the fetuses into fine pieces as described in step 10 of Basic Protocol 1 and incubate with 0.25% trypsin-EDTA (1 ml/fetus) at 37°C for 20 min, then add MEF culture medium and disaggregate the tissues by pipetting vigorously.

During the dissection of embryos and in the following processing steps, there is a high risk

for contamination by bacteria and fungi. Careful sterilization of all materials and tools and good operating skill are very helpful in reducing the risk of contamination. The MEFs at passage 1 should be examined carefully under a phase-contrast microscope. If contamination is identified, all MEFs and their containers should be treated with bleach and discarded, and the experiment should be repeated from the beginning.

Confluent MEFs in culture stop proliferating and begin to differentiate. In addition, it is known that long-term culture of MEFs at high density causes cellular transformation (Aaronson and Todaro, 1968). These autonomously transformed MEFs are able to form tumors if injected into nude mice. Therefore, both primary and immortalized MEFs in culture should never be maintained at a cell density that is over confluency.

As shown in Figure 28.1.1, the number of MEFs harvested from passages 10 to 14 after a period of 3-day culture may be even less than the number inoculated, due to extensive cell death. In this case, it is necessary to combine cells from more than one culture flask to achieve the necessary inoculation density for next passage.

Anticipated Results

MEF immortalization (Basic Protocol 3) is a process that selects for cells that acquire growth advantages. Cell-growth advantages are caused by the accumulation of genetic mutations that enhance cell survival and proliferation and inhibit apoptosis. Since multiple signaling pathways and gene networks are involved in regulation of cell survival, death, and proliferation, the immortalized MEFs must contain heterogeneous genetic mutations, and therefore their gene-expression profiles will differ from cell to cell in the population of immortalized MEFs. These internal variations may cause difficulties in comparing experimental results observed in wild-type MEFs versus mutant MEFs, even if they are developed simultaneously under same conditions. One useful approach is to express the gene of interest in the immortalized mutant MEFs and to compare these gene-expression-rescued MEFs with their parent mutant MEFs.

Time Considerations

Once the embryos have been harvested, it takes 2 to 3 hr to dissect and prepare the tissue

for overnight incubation in trypsin. The next day, it takes 1 to 2 hr to prepare the MEF primary culture. The initial primary culture takes 1 to 3 days to reach confluence. About 1 hr is required for trypsin digestion and aliquotting of passage-1 cell samples. Expansion of these confluent cells requires a 3-day incubation cycle. Each subsequent passaging requires a 3-day incubation, followed by a wash, trypsinization, and transfer to fresh plates.

Literature Cited

- Aaronson, S.A. and Todaro, G.J. 1968. Development of 3T3-like lines from Balb-c mouse embryo cultures: Transformation susceptibility to SV40. *J. Cell. Physiol.* 72:141-148.
- Abbondanzo, S.J., Gadi, I., and Stewart, C.L. 1993. Derivation of embryonic stem cell lines. *Methods Enzymol.* 225:803-855.
- Hogan, B., Beddington, R., Costantini, F., and Lacy, E. 1994. *Manipulating the Mouse Embryo: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Kuang, S.Q., Liao, L., Zhang, H., Pereira, F.A., Yuan, Y., DeMayo, F.J., Ko, L., and Xu, J. 2002. Deletion of the cancer-amplified coactivator AIB3 results in defective placentation and embryonic lethality. *J. Biol. Chem.* 277:45356-45360.
- Puigserver, P., Adelmant, G., Wu, Z., Fan, M., Xu, J., O'Malley, B. W., and Spiegelman, B.M. 1999. Activation of PPARgamma coactivator-1 through transcription factor docking. *Science* 286:1368-1371.
- Todaro, G.J., and Green, H. 1963. Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. *J. Cell Biol.* 17:299-313.
- Todaro, G.J., and Green, H. 1966. Cell growth and the initiation of transformation by SV40. *Proc. Natl. Acad. Sci. U.S.A.* 55:302-308.
- Wu, R. C., Qin, J., Yi, P., Wong, J., Tsai, S.Y., Tsai, M.J., O'Malley, B.W. 2004. Selective phosphorylations of the SRC-3/AIB1 coactivator integrate genomic responses to multiple cellular signaling pathways. *Mol. Cell* 15:937-949.

Contributed by Jianming Xu
Baylor College of Medicine
Houston, Texas

Isolation and Immortalization of Lymphocytes

A convenient renewable source for cells and genomic material from patients or patient populations is not readily available. An easy way to circumvent this problem is to utilize the intrinsic properties of Epstein-Barr virus (EBV), which are its potent ability to immortalize human B lymphocytes to produce lymphoblastoid cell lines (LCLs). Isolation of lymphocytes is relatively simple and they can be easily obtained from small volumes of blood with minimum discomfort to donors. LCLs are of great practical value as an unlimited source of stable genomic DNA and viable cells, which can be used to perform a variety of biochemical and molecular studies. The basic protocol describes simple procedures for production of EBV, isolation of lymphocytes, EBV infection of lymphocytes, and isolation of the resulting LCLs.

CAUTION: Biosafety level 2 (BSL2) practices, containment facilities, and equipment are recommended for activities utilizing known or potentially infectious clinical materials or cultures of indigenous viral agents that are associated or identified as a primary pathogen of human disease (e.g., EBV).

EPSTEIN-BARR VIRUS–MEDIATED IMMORTALIZATION OF HUMAN B LYMPHOCYTES

BASIC PROTOCOL

EBV, first discovered in cell lines derived from endemic Burkitt's lymphomas (Epstein et al., 1964), has a potent ability to transform B lymphocytes in vitro (Henle et al., 1967). In these cells EBV establishes mostly a latent infection characterized by expression of a limited number of viral proteins. The availability of cell lines that readily produce EBV and efficient methods for purification of lymphocytes along with the efficient ability of the virus to immortalize lymphocytes has resulted in EBV-induced B cell immortalization as a method of choice for generating continuously growing cells from patients or patient populations. Isolation of lymphocytes is relatively simple and they can be obtained easily from small volumes of blood. The procedure is particularly useful when cell lines from specific donors are desired.

B95-8 cells are typically used for virus production since they spontaneously produce the B95-8 strain of virus, which has been molecularly and phenotypically characterized in considerable detail. However, transforming virus can be obtained from a number of other EBV-infected cell lines but the efficiency of virus production may vary considerably between cell lines and will need to be tested empirically by the investigator.

IMPORTANT NOTE: B95-8 cells were derived from Marmoset lymphocytes infected and immortalized with Epstein-Barr virus (Miller et al., 1972). Since Marmosets are an endangered species, this cell line is no longer commercially available. However, these cells are widely distributed throughout the scientific community and should be readily available from a number of investigators. If only small quantities of virus are needed, B95-8 virus can be purchased commercially (ATCC, VR-1492).

Materials

B95-8 cells

Growth medium: RPMI-1640 medium (*APPENDIX 3F*), containing 200 mM L-glutamine, with or without fetal bovine serum (FBS; *APPENDIX 3F*)

Ficoll-Paque (Amersham Biosciences 17-1440-02) or Accuspin system-histopaque-1077 tubes (SIGMA diagnostics A 6929)

Mammalian Cell Culture

28.2.1

Anti-coagulated human blood, freshly obtained
Phosphate-buffered saline (PBS; *APPENDIX 2A*)
Cyclosporine A (Sigma-Aldrich)
Storage medium: RPMI medium containing 20% to 40% (v/v) FBS and 10% (v/v) DMSO

15- and 50-ml sterile plastic, screw-cap, conical centrifuge tubes
0.45- μ m filters
96-well flat-welled tissue culture plates, sterile
24-well sterile tissue culture plates or T-25 tissue culture flasks, sterile
Additional reagents and equipment for mammalian cell culture (*APPENDIX 3F*)

NOTE: All solutions and equipment coming into contact with cells must be sterile, and aseptic technique should be used accordingly.

NOTE: All culture incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified.

Prepare the virus

1. Split B95-8 cells to a density of approximately 1×10^6 /ml in growth medium containing 5% FBS. Incubate the cells for 9 to 12 days.

B95-8 cells are generally maintained and passed as standard suspension cells. However, a proportion of cells may also adhere to the bottom of the flask. The adherent cells can be removed by trypsinization or by gently tapping the side of the flask to knock them off the bottom.

The total number of cells will depend on the needs of the individual investigator and the amount of virus recovered. However, 0.5 ml to 1.0 ml of supernatant (see step 2) is generally sufficient to infect lymphocytes and recover LCLs.

Alternatively, more concentrated virus preparations can be obtained by growing the cells in flasks with an attached nutrient supply chamber (Integra Biosciences, Celline CL 350 or 1000). The flasks consist of a cell compartment base capable of holding from 5 to 15 ml of suspension cells separated from a larger compartment capable of holding 350 to 1000 ml of growth medium by a cellulose acetate membrane barrier. High concentrations of cells can be maintained in this environment, allowing virus particles to accumulate in the small lower chamber. 2.5×10^7 viable cells in growth medium containing 10% FBS are introduced into the cell compartment, and 975 ml of warm growth medium containing 5% FBS is introduced into the larger compartment. At day 7, the cells are removed from the lower chamber using a sterile pipet and placed into a sterile 15-ml centrifuge tube. The cells are then centrifuged for 5 to 10 min at $250 \times g$, room temperature. Growth medium is removed by aspiration and the cell pellet is resuspended with fresh growth medium containing 10% FBS. The growth medium from the upper chamber is then aspirated and replaced, this time without FBS to encourage virus production. The entire cell suspension is harvested at Day 17.

2. Remove the cells and growth medium and place in sterile 50-ml centrifuge tubes. Centrifuge 10 to 15 min at low speed ($250 \times g$), 25°C, to pellet the cells. Remove the supernatant and filter through sterile 0.45- μ m filters to remove any contaminating marmoset cells.

If the nutrient supply chamber system was used (see Step 1 annotation), the supernatant is harvested by centrifugation and sterile filtered. Final supernatant is then aliquoted and stored at -80°C .

3. Aliquot the virus-containing supernatant into cryovials or any sterile tube that can be conveniently sealed and store at -80°C .

Virus can also be stored at 4°C for several months but will lose titer. Aliquots of frozen virus will usually maintain titers for much longer periods but will lose significant titer if freeze-thawed more than once.

4. To obtain consistent and successful results, determine a rough estimate of virus titer based on its transformation efficiency (see Support Protocol).

Isolate and infect lymphocytes

5. To a 10-ml sterile tube add 2 ml of defibrinated or anticoagulant-treated human blood and an equal volume of PBS. Mix gently using a sterile pipet.

Fresh blood should be used to ensure high viability of isolated lymphocytes.

EBV can infect and immortalize marmoset B cells (this was how B98-5 cells were obtained). However, EBV is unable to infect and immortalize B cells from the Rhesus Macaque, which is a species more closely related to humans.

6. Add 3 ml Ficoll-Paque to a sterile 15-ml centrifuge tube and carefully layer the diluted blood sample (4 ml) on top of the Ficoll-Paque layer ensuring that the two solutions do not mix.

Alternatively, blood samples can be added to Accuspin system-histopaque-1077 tube (Sigma). The main advantage of using these tubes is that they have a polyethylene barrier between the Histopaque-1077 separation medium (which is located in a lower chamber) and the diluted blood. Therefore, blood added to the top chamber can be added more quickly without risk of mixing with the separation medium and there is a clear separation of blood components after spinning. Follow the manufacturer's recommendations for spin time, force, and temperature.

7. Centrifuge 15 to 30 min at $400 \times g$, 25°C .
8. Observe partitioning of the sample into 4 phases. On the bottom, there are granulocytes and erythrocytes (red), followed by a layer of Ficoll-Paque (clear). A small cloudy-white interface containing the lymphocytes rests on top of the Ficoll-Paque. A layer of plasma (red or yellow-brown) sits on top.
9. Carefully remove and discard the plasma layer. Remove the lymphocyte layer in as small a volume as possible (~ 2 ml) and transfer to a clean, sterile 15-ml centrifuge tube.
10. Add 3 volumes of PBS (6 ml) and centrifuge 10 min at $250 \times g$, 25°C . Remove and discard the supernatant, and add 6 to 8 ml fresh PBS. Resuspend the lymphocyte pellets by gently pipeting up and down.
11. Repeat two to three times. Before the final wash, count the cells using a hemacytometer.
12. After the final centrifugation, resuspend the lymphocytes with virus-containing supernatant from B95-8 cells (an amount equivalent to 10^5 transforming units; see Support Protocol 1) and add growth medium containing 10% FBS plus $1 \mu\text{g/ml}$ cyclosporine A (final concentration) to give 6×10^6 cells/ml in a total volume of 10 to 12 ml. Aliquot $100 \mu\text{l}$ cells per well into a 96-well plate (6×10^5 cells/well). Maintain cells at 37°C in a CO_2 incubator.

To inhibit or diminish regression of B-cell proliferation, cyclosporin A is added to inhibit suppressor or cytotoxic T-cell action.

Isolate and maintain lymphoblastoid cell lines

13. After 1 week in culture, feed cells with $100 \mu\text{l}$ growth medium containing 10% FBS and $1 \mu\text{g/ml}$ cyclosporine A (final concentration). Subsequently, remove one half the medium weekly and replace with fresh medium containing cyclosporine A.
14. After 2 to 3 weeks in culture, small to large clumps of cells begin to emerge and are visible microscopically and in some cases macroscopically. Expand cells by transferring them into 24-well plates or T-25 tissue culture flasks.

In general, LCLs can be passed as standard suspension cells in RPMI containing 10% FBS. Because LCLs produce a number of autocrine factors that facilitate proliferation and viability, it is recommended that cell densities be maintained at levels no lower than $1-2 \times 10^5$ cell/ml. Conversely, cells should be diluted in growth medium containing 10% FBS when cell densities reach saturation between $2-5 \times 10^6$ cells/ml.

- Freeze and store resulting lymphoblastoid cell lines (LCLs) for future use. Centrifuge exponentially growing cells 5 to 10 min at $250 \times g$, 25°C and resuspend in storage medium containing 20% to 40% FBS and 10% DMSO at a cell concentration of $5-10 \times 10^6$ cells/ml. Freeze cells at a rate of 1°C per minute to -40°C and store in liquid nitrogen.

SUPPORT PROTOCOL

TITRATION OF VIRUS ON LYMPHOCYTES

To minimize potential negative results and experimental variability it is useful to approximate the transforming activity of B95-8 virus stocks.

Materials

Lymphocytes, Ficoll-Paque purified (Basic Protocol, steps 5 to 11)

Culture medium: RPMI (APPENDIX 3F) containing 10% (v/v) FBS and 1 $\mu\text{g/ml}$ cyclosporine A

B95-8 virus stocks to be titered

96-well flat-bottomed plates

- Prepare lymphocytes as described above and resuspend at $5-6 \times 10^6$ cells/ml in culture medium containing 10% FBS plus 1 $\mu\text{g/ml}$ of cyclosporine A.
- Aliquot water into all wells in rows A and H and in columns 1 and 12 of a flat-bottomed 96-well plate to help prevent evaporation of medium in wells nearest to the edges of the plate.
- For each virus, aliquot 100 μl of cells per well in rows B to G in triplicate (i.e., three columns and six rows per virus, up to three viruses per plate).
- Aliquot 100 μl of cells in one column for a negative control. Add an additional 100 μl of culture medium to bring final volume to 200 μl .
- Dilute 200 μl virus stock 1:5 in culture medium (10^{-1}) and make five 10-fold dilutions of virus (10^{-2} to 10^{-6}) in culture medium in sterile 1.5-ml tubes (make sure to change tips between each dilution).
- Aliquot 100 μl of 10^{-6} dilution of virus into three wells in row G.
- Aliquot 100 μl of 10^{-5} dilution of virus into three wells in row F and continue until the 10^{-1} dilution is aliquoted into wells in row B. Transfer cells to a 37°C incubator and feed weekly by replacing half the medium with fresh culture medium.

Transformed cells clumps should be evident within a few days in the 10^{-1} and 10^{-2} dilutions but will take longer in the higher dilutions. The highest dilution giving rise to transformed cells after 4 to 5 weeks approximates the transforming titer of the virus stock.

COMMENTARY

Background Information

Epstein-Barr virus (EBV), a versatile lymphotropic γ -type human herpesvirus or Lymphocryptovirus (LCV), is endemic in all human populations (Rickinson and Kieff, 2001). Most people are infected during early child-

hood with no apparent clinical features. However, delayed infection occurring in adolescence or in young adults can result in mononucleosis, which is usually a benign lymphoproliferative disease. In combination with genetic, environmental, or iatrogenically

induced conditions, EBV is an important cofactor with a causal link to a wide variety of human malignancies including Burkitt's lymphoma, gastric carcinomas, nasopharyngeal carcinoma, T-cell lymphomas, Hodgkin's disease, post-transplant lymphoma-like disease (PTLD), and leiomyosarcomas.

EBV possesses an intrinsic ability to transform human B cells in culture. Cellular growth control is mediated through the concerted action of several latent cycle gene products including EBV nuclear antigen-leader protein (EBNA-LP), -2 (EBNA2), -3A and -C (EBNA3A and EBNA3C), and latent membrane protein one (LMP-1; Bornkamm and Hammerschmidt, 2001). Viral genomes are usually maintained as circular episomes but occasionally can integrate into host DNA.

The high efficiency of EBV-induced B cell immortalization (Sugden and Mark, 1977) makes it an ideal tool for generating cell lines derived from any individual (Neitzel, 1986). The applications for EBV-immortalized cell lines are many and include investigation of signaling pathways involved in B cell growth and proliferation, generation of cytotoxic T cells (CTLs) for cellular immunotherapy, and a variety of molecular and genetic studies.

Critical Parameters and Troubleshooting

Failure to obtain immortalized cell lines can be due to several problems. First, the viability of lymphocytes/mononuclear cells used for infection is important. Usually, freshly isolated cells work best. However, blood samples stored up to 4 days can be immortalized. Lymphocytes that have been frozen can also be thawed and immortalized, although this is often at lower efficiency. A second problem encountered may be the loss of infectious virus in the virus preparations or variability in the titer between different virus preparations. It is recommended that a large batch of virus be prepared, aliquoted, and stored. The virus can be titered for transformation efficiency and should retain good viability during long-term storage at -80°C . Addition of cyclosporine A increases the efficiency of LCL outgrowth by inhibiting the emergence of EBV-specific T cells (Bird et al., 1981).

Anticipated Results

Using the methods described here, multiple cell lines should emerge from an individual infection.

Time Considerations

After blood samples are obtained, preparation of lymphocytes, infection with EBV, and plating will take 1-3 hr for a few samples. The time required to obtain EBV-immortalized cell-lines is usually 2-4 weeks. Amplification of the cell lines to obtain cell numbers required for experimental manipulation may take additional time, depending on the amount of cells needed for a specific application.

Literature Cited

- Bird, A.G., McLachlan, S.M., and Britton, S. 1981. Cyclosporin A promotes spontaneous outgrowth in vitro of Epstein-Barr virus-induced B-cell lines. *Nature* 289:300-301.
- Bornkamm, G.W. and Hammerschmidt, W. 2001. Molecular virology of Epstein-Barr virus. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 356:437-459.
- Epstein, M.A., Achong, B.G., and Barr, Y.M. 1964. Virus particles in cultured lymphocytes from Burkitt's lymphoma. *Lancet* 1:702-703.
- Henle, W., Diehl, V., Kohn, G., Zur Hausen, H., and Henle, G. 1967. Herpes-type virus and chromosome marker in normal leukocytes after growth with irradiated Burkitt cells. *Science* 157:1064-1065.
- Miller, G., Shope, T., Lisco, H., Stitt, D., and Lipman, M. 1972. Epstein-Barr virus: Transformation, cytopathic changes, and viral antigens in squirrel monkey and marmoset leukocytes. *Proc. Natl. Acad. Sci. U.S.A.* 69:383-387.
- Neitzel, H. 1986. A routine method for the establishment of permanent growing lymphoblastoid cell lines. *Hum. Genet.* 73:320-326.
- Rickinson, A.B. and Kieff, E. 2001. Epstein-Barr Virus. 4th ed. In *Virology* (D.M. Knipe and P.M. Howley, Eds.), Vol. 2, pp. 2575-2627. Lippincott-Raven Publishers, Philadelphia.
- Sugden, B. and Mark, W. 1977. Clonal transformation of adult human leukocytes by Epstein-Barr virus. *J. Virol.* 23:503-508.

Internet Resources

<http://www.cdc.gov/od/ohs/biosfty/bmb14/bmb14toc.htm>

Web site for acquiring the most recent CDC handbook, *Biosafety in Microbiological and Biomedical laboratories (BMBL) 4th edition*.

Contributed by Paul D. Ling and
Helen M. Huls
Baylor College of Medicine
Houston, Texas

Establishment and Culture of Human Skin Fibroblasts

UNIT 28.3

In some instances, patient-derived skin fibroblast lines afford advantages over continuous transformed cell lines, in that biochemical abnormalities underlying a disease state may be reproduced in the patient samples. Historically, human dermal fibroblast lines established from patient biopsies have been instrumental in studies designed to elucidate the pathogenesis of several diseases, such as the characterization of the low density lipoprotein (LDL) receptor in familial hypercholesterolemia (FH; Basu et al., 1978), the function of the androgen receptor (AR) in androgen insensitivity syndrome (AIS; reviewed in Quigley et al., 1995), and the deficiency of hypoxanthine-guanine phosphoribosyl transferase (HGPT) of Lesch-Nyhan Syndrome (Migeon et al., 1968).

GENITAL SKIN BIOPSY FOR GENERATION OF FIBROBLAST CULTURES

**BASIC
PROTOCOL 1**

Biopsies to obtain material for the establishment of fibroblast cultures are invasive procedures. As such, the benefits to patients should be weighed against the risks of the procedure (albeit minor), which include the risks of bleeding from and infection at the biopsy site. In all instances, protocols describing the proposed studies must be reviewed and approved by the local institutional review committees, and informed patient consent is required. While somewhat less onerous, regulations governing the use of animal subjects also require that adaptation of these methods to animal subjects be approved by the local Institutional Animal Care and Use Committee (IACUC). As with the acquisition of any sample from patient or animal, health risks to the investigator, such as the possibility of contracting various diseases from the patient, should be taken into consideration. Sterile technique is essential at all steps. While not required, the use of gloves facilitates the implementation of sterile technique and minimizes any risks from infectious diseases.

Materials

Betadine solution

Sterile saline solution: 0.9% (w/v) NaCl in deionized, distilled water

1% sterile lidocaine solution

Biopsy culture medium (see recipe)

Scalpel and forceps, sterile (autoclaved)

50-ml conical tubes

Silver nitrate sticks (silver nitrate applicators, 75% w/w, Arzol Chemical)

4-0 Vicryl suture

1. Clean skin with Betadine solution and rinse the area with sterile saline solution.
2. Achieve local anesthesia by local subcutaneous injection of a 1% sterile lidocaine solution.
3. With sterile scalpel and forceps, take two full-thickness skin biopsies ($\sim 5 \text{ mm}^3$).

Depending on the studies to be conducted, specific sites may be preferred. For example, for studies of the AR, genital skin biopsies are employed. In males, take one from the base of the phallus and the other from the base of the scrotum; for females take one from the base of the clitoris and one adjacent to the vaginal opening (labia majora).

4. Place biopsies into 50-ml conical tubes and fill with biopsy culture medium, excluding as much air as possible from the tube.

Subsequent handling should be at room temperature (25°C).

**Mammalian Cell
Culture**

28.3.1

5. To control bleeding at the biopsy site, apply a silver nitrate stick by moistening the tip of the silver nitrate stick with sterile water or saline. Apply the stick to the site of wound (there is a slight burning sensation, usually accompanied with a skin color change which indicates sealing of the wound). Apply wet gauze to wound site. If necessary, suture skin with 4-0 Vicryl suture (one to two stitches).

Optimally, specimens should be processed quickly to establish the primary cell lines. However, biopsies may be stored in culture medium for 3 to 4 days at 4°C. If samples are to be shipped, this time limit should be taken into consideration as well as the time needed to establish the primary cell line in the receiving laboratory. Samples should be shipped in containers filled completely with culture medium at room temperature. If extremes of temperature are anticipated, shipment in an insulated container will help to preserve biopsy viability.

ESTABLISHMENT OF PRIMARY CULTURES OF SKIN FIBROBLASTS

Utilization of the primary explant method (see Basic Protocol 1) for the establishment of primary cultures from biopsies provides an easy method of generating fibroblast strains. Sterility and aseptic technique play important roles in the establishment of cell lines that are viable and free of contaminating pathogens.

Materials

Biopsy (see Basic Protocol 1)
Dulbecco's modified phosphate-buffered saline (DPBS) without calcium or magnesium (Cellgro or see recipe)
Primary culture medium (see recipe)
25-cm² tissue culture–treated flasks with 0.22- μ m filter caps
Forceps, scissors, and scalpels, sterile (autoclaved)
6-cm tissue culture dish
1-ml tuberculin syringe (disposable) with 23-G needle
37°C, 5% CO₂ humidified circulating incubator
Inverted microscope

1. Label two 25-cm² tissue culture–treated flasks for each specimen with the biopsy identifier information, distinguishing them from one another by specific designations (e.g., A and B).

Each biopsy will be split into duplicate flasks and, if possible, housed in different incubators. This provides a backup in case one of the flasks becomes contaminated or lost in the primary culture setup.

2. While working under a laminar flow hood, remove the excess biopsy culture medium present in the conical tube that contains the biopsy. Add 10 ml of DPBS (without calcium and magnesium) to completely saturate and wash the biopsy. Aspirate the DPBS and repeat two additional times for a total of three washes.

The DPBS wash volume should be adequate to remove loose fibers and contaminants (such as hair) and excess biopsy culture medium; generally 10 ml should suffice.

3. Utilizing sterile forceps, remove the biopsy to a sterile 6-cm tissue culture dish containing ~7 ml primary culture medium. With a fresh pair of sterile forceps and sterile scissors, cut the biopsy into small pieces.
4. Using either a sterile scalpel or scissors, refine the size of the biopsy pieces to a size that is approximately that of a pinhead.

- Using a 1-ml tuberculin syringe with a 23-G needle attached, take one pinhead-sized explant and place it into one of the 25-cm² flasks (the explant should be moist from the primary culture medium). Repeat this procedure, adding at least seven explants to each duplicate flask.

In adding more explants, the probability of establishing the primary cell line and having better growth in the flasks is increased.

- Allow the explants sufficient time to adhere to the primary culture flask (~15 to 20 min). Then add 12 ml primary culture medium to the flask, being careful not to detach the explants.

In steps 5 and 6, speed and efficiency are important. Once the explants have been placed into the flasks, it is important to allow enough time for the explants to firmly attach to the surface of the flask, but not so long that they will dry out. If the explants are not allowed sufficient time to adhere, they will float off when medium is added and, under these conditions, outgrowth of the fibroblasts from the explants will be unlikely. A helpful frame of reference to determine whether sufficient adherence has occurred is to assess when the residual medium has dried and formed a "halo" around the explants. Once this halo has formed, it is safe to gently add primary culture medium to the explants.

- Carefully place the flasks into a 37°C, 5% CO₂ humidified circulating incubator for human skin fibroblast cultures.

- After 5 to 7 days, remove the medium and add fresh primary culture medium to the flask.

Adding medium prior to this time may perturb the explants; optimally, every advantage for the explants to firmly attach and grow should be given.

- After the initial replacement of medium, check fibroblast cultures daily for growth, surface coverage (confluence), and contamination using an inverted microscope with a 10× objective. Replace medium approximately every 3 days until cell lines are established.

Contaminants such as bacteria, mold, and fungus—if present—will be observed within a few days.

ESTABLISHMENT AND MAINTENANCE OF SKIN FIBROBLAST CELL LINES

BASIC PROTOCOL 3

The time for outgrowth of the biopsy explants to reach confluence is generally ~2 to 3 weeks. Note that the number and size of explants per flask, the health of the explants from biopsy to implantation into a culturing vessel, and the general health of the outgrowth from the explants all contribute to the growth and time of confluence of the eventual primary fibroblast cell line. In addition, cultures established from young subjects will generally grow with greater vigor than those established from older subjects. Once the biopsy explants have covered between 75% and 90% of the culture area, a fibroblast cell line may be commenced as described.

Materials

- Explants grown in 25-cm² tissue culture–treated flasks (see Basic Protocol 2)
- Dulbecco's modified phosphate-buffered saline (DPBS) without calcium or magnesium (Cellgro or see recipe)
- 0.25% trypsin/EDTA solution: 2.5 g/liter porcine trypsin and 0.38 g/liter EDTA·4Na⁺ in Hanks' balanced salt solution (HBSS; commercially available as a 1× solution)
- Fibroblast culture medium (see recipe)
- 75-cm² tissue culture–treated flasks with 0.22-μm filter cap
- 10-ml pipets

Mammalian Cell Culture

28.3.3

1. When explant growth has covered ~75% to 90% of the culture area, commence trypsinization by removing the medium.
2. Wash culture surface of the 25-cm² flask by adding 4 ml DPBS without calcium or magnesium. Rinse and remove by aspiration.
3. Add 1 ml of 0.25% trypsin/EDTA solution to the culture surface and allow cells to detach.

It is best to observe cell detachment via 10× magnification under an inverted microscope. Fibroblasts generally retract and detach after 1 to 2 min. It is preferable to minimize the length of time that the cultures are treated, as excessive trypsinization will kill cells. Tapping the flask gently will help dislodge cells that are loosely attached.

4. Add 3 ml fibroblast culture medium to the flasks. Gently pipet the trypsin and fibroblast culture medium mixture up and down to make a homogenous suspension.
5. To a 75-cm² tissue culture–treated flask that contains 10 ml of fibroblast culture medium, add all 4 ml of the cell solution. Place the 75-cm² flask into the 37°C, 5% CO₂ incubator. Allow fibroblasts to grow to confluence.

Generally, confluence of the fibroblast is seen within 5 days, depending on the health and condition of the cells.

6. Once the cells have become confluent, trypsinize and passage cells as above, with the exception of adding 2 ml trypsin/EDTA solution followed by 4 ml fibroblast culture medium into the 75-cm² culturing vessel.

Generally, the cells may be split at a ratio of one flask of confluent cells into three new flasks (i.e., 1:3). Take 2 ml of the total 6 ml of trypsinized cell solution, as previously described, and add to a 75-cm² flask containing 10 ml fibroblast culture medium. Thus, each confluent flask is sufficient to start three new flasks. A confluent 75-cm² flask contains ~1.5 × 10⁶ cells.

Each subsequent passage from the first trypsinization to the point where the cells are frozen cryogenically should be noted. Every laboratory has a specific stringency and strictness on how many passages are allowed before frozen preservation. In general, the fewer passages before cryogenic preservation and use in subsequent functional assays, the healthier the cell lines will be. After several passages (e.g., 3 months), the fibroblast growth rate will decrease and cell lines will, therefore, take longer to reach confluence. Morphological differentiation of fibroblasts may be minimal and can be hard to distinguish even under microscopic magnification. This decrease in cell growth is the best indicator of cell line differentiation and eventual cell senescence. Hence, it is best to freeze cells from initiation to no more than nine passages.

CRYOGENIC PRESERVATION OF SKIN FIBROBLAST CELL LINES

Cryogenic preservation of fibroblasts is generally the same as cryopreservation of most continuous epithelial cell lines. The addition of DMSO to a final concentration of 10% permits the long-term preservation of cells at –80°C or in a liquid nitrogen freezer (Sherman, 1965).

Materials

- Fibroblast culture medium (see recipe)
- Dimethyl sulfoxide (DMSO), sterile
- Established fibroblast cell lines in 75-cm² tissue culture flasks (see Basic Protocol 3)
- Dulbecco's modified phosphate-buffered saline (DPBS) without calcium or magnesium (Cellgro or see recipe)
- 0.25% trypsin/EDTA solution: 2.5 g/liter porcine trypsin and 0.38 g/liter EDTA·4Na⁺ in Hanks' balanced salt solution (HBSS; commercially available as a 1× solution)

15-ml conical tube
Sterile freezing vials (1.5- to 2-ml ampules)
Cryogenic freezing container (e.g., Cryo 1°C Freezing Containers, “Mr. Frosty,”
Nalgene Labware) or liquid nitrogen freezer

1. To fibroblast culture medium, add DMSO to a concentration of 10% (v/v) to make the freezing medium.
2. Wash culture surface of the 75-cm² tissue culture flask by adding 4 ml DPBS without calcium or magnesium. Rinse and remove by aspiration.
3. Add 2 ml of 0.25% trypsin/EDTA solution to the culture surface and allow for cell dissociation.

It is best to observe cell detachment via 10× magnification under an inverted microscope. Fibroblasts generally retract and detach after 1 to 2 min of trypsinization. It is preferable to minimize the time of exposure, as excessive trypsinization will kill cells. Tapping the flask gently will help dislodge cells that are loosely attached.

4. Add 4 ml fibroblast culture medium to the flask. Gently pipet the trypsin and fibroblast culture medium mixture up and down to make a homogenous suspension.
5. Add cell suspension to a 15-ml conical tube and centrifuge 5 min at 650 to 1000 × g, 4°C.
6. Aspirate off fibroblast culture medium and resuspend cell pellet in 2 ml freezing medium. Pipet up and down to make a homogenous suspension and add 1 ml to each of two 1.5- or 2-ml freezing vials (ampules).

Freezing volume can be modified to accommodate the various sizes of freezing vials/ampules that are available. Even if the freezing volume changes, the amount of DMSO for cryogenic preservation must be maintained at 10% (v/v) of the freezing volume, with the amount of cells frozen per vial roughly equal to half of those obtained from the confluent culturing vessel from which they were derived. Freezing back approximately half of a culture in each vial assures advantageous growth and survival after future thawing since a small portion of cells do not survive the freezing and/or thawing associated with cryogenic preservation. Each vial should be appropriately labeled with identifier information and cell passage number.

7. Place vials into cryogenic carrier (e.g., Mr. Frosty) and allow to freeze at a rate of 1°C/min at –80°C overnight. Cryogenically preserve in a liquid nitrogen freezer.

If cells are preserved in liquid nitrogen, in theory they can be stored indefinitely. Cells can be stored at –80°C, however not indefinitely. Fibroblasts stored at –80°C will lose viability more rapidly than those stored in liquid nitrogen.

8. Before use in functional assays, thaw cells quickly in a 37°C water bath or by quickly thawing while rubbing the vial between the hands. Wipe vial with 70% ethanol to remove any possible contaminants derived from either the water bath or hands. Allow cells to revive growth by either of the following.
 - a. Pipet thawed volume of cells directly into 10 ml of fibroblast growth medium in a 75-cm² culturing flask and allow for cells to attach to the vessel. Replace the medium on the following day.
 - b. Place thawed cell suspension into a sterile 15-ml conical tube containing fibroblast growth medium. Centrifuge cells at 650 to 1000 × g, 4°C, to pellet the cells. Remove all of the supernatant and resuspend cell pellet in fresh 10 ml fibroblast growth medium. Transfer the cell suspension into the culture vessel. Replace medium on the third day, as for normal routine feeding.

In the first option (a), the DMSO present in the freezing medium is being diluted but not removed, and in high percentages can be toxic to growing fibroblasts. As such, cells should not be left for more than 1 day in the DMSO-containing medium, which should be changed the following day (preferably in the morning). In the second option (b), the cells are quickly removed from contact with the DMSO-containing medium (when the medium is removed by centrifugation). However, the mechanical manipulation may decrease the proportion of cells that survive the thawing process.

After cells that survived the thaw have grown to confluence, they may be passaged for several months (generally 3 months). As noted above, after several passages, cells will exhibit slowing growth rates and eventual cell senescence (see Basic Protocol 3, step 6 annotation).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Biopsy culture medium

Aseptically combine:

400 ml minimum essential medium (MEM; APPENDIX 3F)

100 ml fetal bovine serum (FBS; APPENDIX 3F)

5 ml penicillin/streptomycin solution (10,000 U/ml penicillin G and 10,000 µg/ml streptomycin; commercially available from multiple sources, e.g., Sigma)

Store up to 2 months at 4°C

This mixture gives rise to 20% FBS and 1% antibiotic in MEM.

DPBS without Ca²⁺ and Mg²⁺

0.20 g/liter KCl

0.20 g/liter KH₂PO₄

8 g/liter NaCl

1.15 g/liter NaH₂PO₄

Final pH 7.4

Autoclave

Fibroblast culture medium

Aseptically combine:

400 ml minimum essential medium (MEM; APPENDIX 3F)

100 ml fetal bovine serum (FBS; APPENDIX 3F)

5 ml penicillin/streptomycin/Fungizone solution (commercially available in 10,000 U/ml penicillin G, 10,000 µg/ml streptomycin, 25 µg/ml amphotericin B; e.g., Mediatech)

Store up to 2 months at 4°C

This mixture gives rise to 20% FBS and 1% antibiotic in MEM.

Medium should be stored at 4°C, but should be warmed in a 37°C water bath prior to addition to cells. After use, close tightly to avoid oxidation of medium by air and return for long-term storage at 4°C.

Primary culture medium

Aseptically combine:

400 ml minimum essential medium (MEM; APPENDIX 3F)

100 ml fetal bovine serum (FBS; APPENDIX 3F)

5 ml penicillin/streptomycin/Fungizone solution (commercially available in 10,000 U/ml penicillin G, 10,000 µg/ml streptomycin, 25 µg/ml amphotericin B; e.g., Mediatech)

Anti-mycoplasma reagent (amount varies per manufacturer's instructions; 1 × final concentration; e.g., Mycoplasma Removal Agent, MP Biomedicals)

Store up to 2 months at 4°C

Medium should be stored at 4°C, but should be warmed in a 37°C water bath prior to addition to cells. After use, close tightly to avoid oxidation of medium by air and return for long-term storage at 4°C.

COMMENTARY

Background Information

Although there are several ways to establish fibroblast cultures from human biopsies, the methods described have been utilized to establish a large number of cell strains that have been used successfully in a variety of in vitro studies (Avila et al., 2002). With little modification, these methods can be applied to establish fibroblast strains from other species as well (Catalano et al., 2002). Generally, the characteristic appearance of spindle-shaped cells aligned in monolayers as “whirls” (see Fig. 28.3.1) is a good indication of the health of the cultures.

There are many methods for the initiation of fibroblast cultures from human biopsies. As reviewed in Freshney (2000), many techniques are based on modification of the primary explant methods initiated by the works of Harrison (1907) and Carrel (1912), and incorporate changes permitted by improvements of

tissue culture reagents. The methods described in this unit follow a generalized modified version of these techniques as applied in the studies of fibroblast cultures in the analysis of subjects with defects of the androgen receptors (AR; Avila et al., 2002, and references therein). The nature of the fibroblast cell line and any contaminant cell types will be affected by the anatomical site from which the biopsy was obtained. Adipose cells, hair cells and follicles, and necrotic cells may be present. Adipose and hair cells will be fairly evident to inspection and are respectively removed in the wash steps and fine dissection of the biopsy into pinhead-sized explants. Thus, selection is favored towards cells that should give outgrowth, which are the fibroblasts.

General growth conditions such as 37°C, carbon dioxide exchange, humidified incubation are favorable for many cell lines and contaminants (e.g., bacteria, yeast, and mold), and

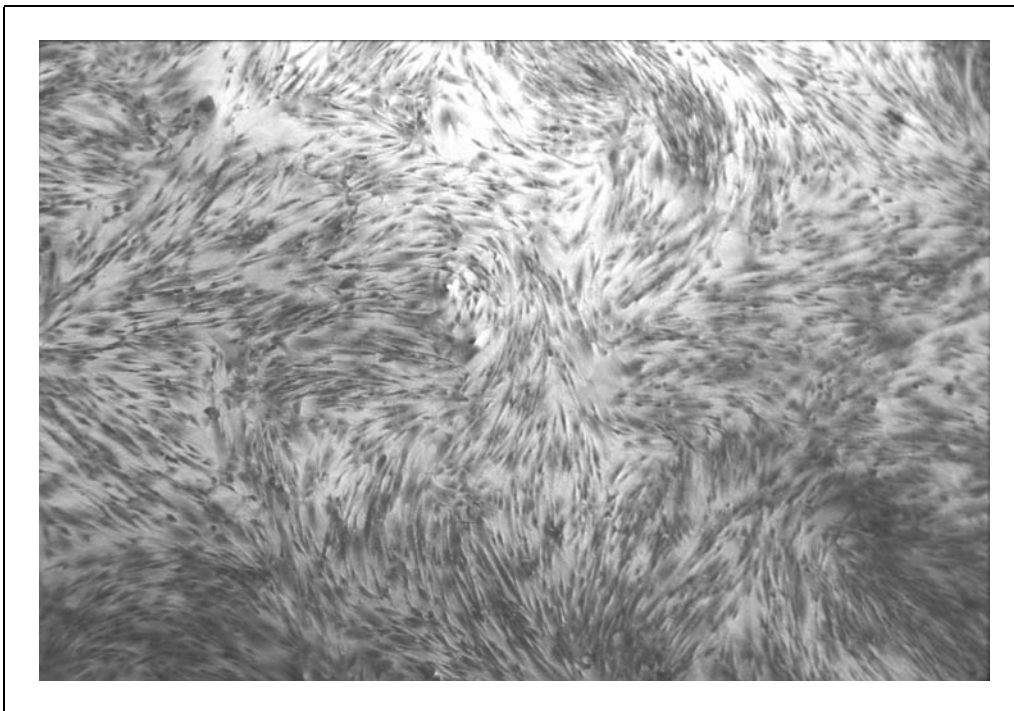


Figure 28.3.1 Healthy human fibroblasts in late exponential, early log phase growth. Fibroblast monolayers established from a genital skin biopsy and grown in MEM containing 20% FBS. Cells were stained with Giemsa stain and visualized under a phase-contrast microscope (magnification ~40×). Photo is printed in grayscale.

therefore, one should check for contaminants on a weekly—if not daily—basis. The fibroblast culture medium (MEM and 20% FBS) is favorable towards fibroblasts derived from patient biopsies and thus richer media (such as DMEM or F12) are not necessary. Fibroblasts, as is true for many cell lines, may display serum sensitivity and react favorably or negatively toward certain lots of serum, due to the varying nature of serum. However, most companies are willing for laboratories to compare their stock lots of FBS and will provide 50 ml of different lots for serum testing. There are other supplemented FBS specifically directed towards fibroblast growth, however, many fibroblast cell lines have been initiated and maintained in bovine calf serum and supplemented calf serum (e.g., Cosmic Calf Serum commercially available from Pierce/Hyclone).

Critical Parameters

Biopsies. To enhance the likelihood that healthy fibroblast cell lines are successfully established, minimize the time from biopsy to the establishment of tissue culture. Biopsies are viable for a finite period of time (typically up to 3 to 4 days). The shorter the time to culture, the greater is the likelihood that a biopsy can be successfully manipulated to establish a cell line.

Primary cultures. One of the most important and most challenging aspects in establishing the primary culture from the explant arises in permitting the adherence of the explant to the culturing vessel before adding the growth medium. As mentioned in Basic Protocol 2 (also see Time Considerations), looking for the appearance of a halo around the explant and giving a 10- to 20-min settling time should allow the explant to attach before the growth medium is added. The time it takes for the halo to dry and for the explant to attach is dependent on how much medium is on each individual explant and on the size of the explant. In general, a small explant with minimal medium will take less time to attach to the culture vessel than a larger explant that is saturated in growth medium. If a portion of the explant adheres to the vessel while another portion hangs suspended in the growth medium, the explant can still produce outgrowth from the adhered portion. Although experience will aid in the timing of this critical juncture, it is not impossible for a novice to achieve successful explant adherence with minimal loss of adherence when the growth medium is added, and to have good outgrowth from the explants.

Finally, as with any tissue culture procedures, sterile technique to keep the explants free from contaminants such as fungi, mold, and bacteria will always enhance the generation of a new cell line.

Cell lines. Although fibroblast cell lines usually do not require precise methods for maintenance, establishing a regular regimen of trypsinization and feeding will optimize the health of the established cultures. Sterile technique will protect the cell lines from contamination from microorganisms (yeast and bacteria, in particular) at all stages. Although self-evident, the use of sterile pipets and plasticware specifically designated for each cell strain will minimize possible contamination by other cell lines.

Cryogenic preservation. The cryopreservation of healthy cell lines will usually yield a high rate of viability after the cells are thawed. It should be noted, however, that the rate at which the cells are frozen at -80°C and the number of passages before cryopreservation will also affect cell viability at thawing.

Troubleshooting

There should be few complications in obtaining and manipulating the skin biopsy. Bleeding by the patient, as mentioned, should be controlled with a silver nitrate stick and closure of the site with appropriate suturing. One problem that is often encountered is from attempts to permit the attachment of the biopsy explants to the culturing vessel without sufficient drying. If there is no attachment of the explant to the culture vessel, it is unlikely that cell outgrowth will occur.

Anticipated Results

By using these protocols, healthy fibroblast cell lines can be obtained from human skin biopsies. With minor modifications, these same protocols can be applied to establish fibroblast lines from most mammalian species. In addition to providing genetic material for use in studies to investigate the genetic basis of disease, these lines may provide a physiological model system in which to explore normal or altered cellular pathways present in cells derived from the patient biopsy.

Time Considerations

Biopsies. It is best to coordinate the time of biopsy with the timing of the manipulations needed to initiate the primary culture, especially if the biopsy is to be shipped. If kept in medium at 4°C , the biopsy can be stored for 4 days.

Primary cultures. As coordinated with the time of the biopsy, biopsies should be processed as quickly as possible. In utilizing the primary explant method, the main challenge is allowing sufficient time for the explants to adhere and quickly reestablishing them in growth medium without having the explants float off.

Cell lines. Daily observance of explant growth will facilitate the transition of explants to cell lines. Timing from explant to the point of the subculture of outgrowth depends on the number of healthy explants successfully attached. With five explants, this takes ~2 to 3 weeks, depending also on the natural growth of the explants.

Cryogenic preservation. Once established, samples should be preserved as early as possible and the number of passages minimized.

Literature Cited

- Avila, D.M., Wilson, C.M., Nandi, N., Griffin, J.E., and McPhaul, M.J. 2002. Immunoreactive AR and genetic alterations in subjects with androgen resistance and undetectable AR levels in genital skin fibroblast ligand-binding assays. *J. Clin. Endocrinol. Metab.* 87:182-188.
- Basu, S.K., Goldstein, J.L., and Brown, M.S. 1978. Characterization of the low density lipoprotein receptor in membranes prepared from human fibroblasts. *J. Biol. Chem.* 253:3852-3856.
- Carrel, A. 1912. On the permanent life of tissues outside the organism. *J. Exp. Med.* 15:516-528.

Catalano, S., Avila, D.M., Marisco, S., Wilson, J.D., Glickman, S.E., and McPhaul, M.J. 2002. Virilization of the female spotted hyena cannot be explained by alterations in the amino acid sequence of the androgen receptor (AR). *Mol. Cell Endocrinol.* 194:85-94.

Freshney, R.I. 2000. *The Culture of Animal Cells: A Manual of Basic Technique*, 4th ed., pp. 158-161. Wiley-Liss, New York.

Harrison, R.G. 1907. Observations on the living developing nerve fiber. *Proc. Soc. Exp. Biol. Med.* 4:140-143.

Migeon, B.R., Der Kaloustain, V.M., Nyhan, W.L., Young, W.J., and Childs, B. 1968. X-linked hypoxanthine-guanine phosphoribosyl transferase deficiency: Heterozygote has two clonal populations. *Science* 160:425-427.

Quigley, C.A., DeBellis, A., Marchke, K.B., El-Awady, M.K., Wilson, E.M., and French, F.S. 1995. Androgen receptor defects: Historical, clinical, and molecular perspectives. *Endocr. Rev.* 16:298-300.

Sherman, J.K. 1965. Pretreatment with protective substances as a factor in freeze-thaw survival. *Cryobiology* 1:249-304.

Contributed by Jacob Villegas and
Michael McPhaul
University of Texas Southwestern
Medical Center
Dallas, Texas

CHAPTER 29

Mouse Phenotyping

INTRODUCTION

When *Current Protocols in Molecular Biology* first appeared in 1987, cloning and sequencing mammalian genes was a major goal of many laboratories. That era of gene discovery ended with the completion of the genome sequences for humans and increasing numbers of mammalian and other species. From a broad perspective, these genome sequences represent a remarkable accomplishment for humankind, of which molecular biologists can be proud. However, they are certainly not just trophies to be put on the shelf or files to be stored on hard drives, but one of the most powerful tools available in the current era of gene function.

Exploiting these tools is dependent on appropriate biological systems. “Simple” model organisms have provided numerous major insights into developmental processes, for example. Some problems, however, such as metabolic diseases and complex behaviors, are best addressed in mammalian species. Decades of physiologic, endocrinologic, and behavioral studies have focused on rodents, and this deep pool of information is now complemented by extensive genetic tools. Modern genetic studies in mice, like *Drosophila*, date to pioneering work of Haldane (Haldane, J.B.S., Sprunt, A.D., and Haldan, N.M. 1915. Reduplication in mice. *J. Genet.* 5:133-135.5). As described in Chapter 23, however, classic genetic approaches have been supplemented by powerful techniques for manipulation of the mouse genome. These extensive tools make the laboratory mouse, *Mus musculus*, the mammalian system of choice for genetic analysis.

These diverse genetic technologies produce large numbers of mice with potential phenotypes of all types. In some cases, loss of gene function results in prenatal lethality. This is perhaps the most obvious phenotype and highlights the potential importance of the gene product. It can also be problematic because it substantially complicates detailed analysis of molecular mechanisms. In many other cases, the phenotypes of gene knockouts are less severe than anticipated, often due to functional redundancy. In such circumstances, more detailed analysis often reveals strong phenotypic effects that are not immediately evident. This chapter will present a variety of approaches to analyze phenotypes of mutant mice, with an initial focus on metabolic parameters.

UNIT 29A.1 reviews approaches to genetic manipulation in the mouse, including both knockout and transgenic technologies detailed in Chapter 23 and modern approaches to analysis of quantitative traits using inbred strains. General issues faced in mouse phenotyping, such as housing and other environmental factors, are considered in *UNIT 29A.2*. *UNIT 29B.1* begins the section on metabolic regulation with analysis of energy balance. Because body weight is so tightly regulated at a variety of levels, it is surprising that many mutant mice exhibit either increased or decreased weight gain, particularly in response to high fat diets or other dietary challenges. *UNIT 29B.1* presents a detailed plan for analysis of such phenotypes. Future units in metabolic regulation will focus on glucose metabolism and lipid metabolism. Techniques for analysis of behavioral phenotypes will be the basis for a forthcoming section.

David D. Moore

Contributed by David D. Moore

Current Protocols in Molecular Biology (2006) 29.0.1
Copyright © 2006 by John Wiley & Sons, Inc.

Mouse
Phenotyping

29.0.1

Supplement 74

GENERAL CONSIDERATIONS IN MOUSE PHENOTYPING

Mice and humans share powerful genetic and biological similarities. In addition, mice are small in size, cheap to house, and have a short lifespan and large litter size, making them an excellent model for scientific investigation. Furthermore, problems of genetic and environmental heterogeneity can be minimized in the mouse, as genetic background and environmental factors can be rigorously controlled. The ability to insert or selectively delete predetermined genes in whole animals, as well as the use of genetic reference populations to study complex traits and quantitative trait loci, provide enormous power for analyzing gene function. With these successes in mouse genetics, the critical factor in the use of phenogenomic strategies is the need for an efficient and vigilant process for mouse phenotyping. It is now recognized that the phenotype of a mouse (behavioral, metabolic, and so on) is highly dependent on genetic, developmental, and environmental factors. All phenotyping experiments must begin with careful consideration of these factors, which can influence all variables and bias the resulting phenotype.

Uses of Forward and Reverse Genetics in Mice to Study Gene Function

The purpose of this chapter is to bridge the gap between molecular biology and physiology so that the novice mouse user has the required vocabulary and background information to understand the molecular origin of the mouse to be phenotyped (also see Argmann et al., 2005). This unit summarizes the diverse genetic strategies used in the mouse, and serves as a prelude to the detailed phenotyping protocols that follow. Protocols for carrying out the genetic manipulations themselves are described elsewhere in this manual, and appropriate units are cross-referenced here.

It is becoming common practice for genetically engineered mouse models (GEMMs) to be generated in collaboration with core facilities having technical expertise in this field. Furthermore, many GEMMs or genetic reference populations (GRPs) can be readily purchased from commercial suppliers. There are a number of ongoing attempts to facilitate the accessibility of available mouse strains, including an international resource that is presently being consolidated in the form of a searchable online database (<http://www.informatics.jax.org/imsr/index.jsp>). This database will regroup all mouse strains and stocks available worldwide, including inbred, mutant, and genetically engineered

strains. For additional mouse-related Web sites, see Internet Resources.

REVERSE GENETICS

Reverse genetics is a gene-driven approach that deduces, from directed mutations, the full range of phenotypes controlled by each gene. This strategy relies on engineering the mouse genome through adding, subtracting, and replacing genes to control gene activity temporally, spatially, and reversibly. Three reverse genetics approaches are described below.

Transgenics

The production of transgenic mice (UNIT 23.1) is a gain-of-function approach to manipulation of the mouse genome (Clarke, 2002; Nagy et al., 2003; Hedrich, 2004). Transgenic mice can be designed for numerous purposes including: (1) identification and characterization of specific regulatory sequences; (2) ubiquitous or tissue-specific overexpression of a transgene; and (3) cell lineage ablation achieved by coupling a tissue-specific promoter to a sequence encoding a toxin which kills cells that express it. Transgenic mice are typically generated by microinjection of the transgene construct into the male pronucleus of fertilized oocytes, which are then

reimplanted into the oviducts of foster females (UNIT 23.9). Expression of the transgene is dependent on its site of integration, with individual insertion sites typically containing multiple copies of the transgene. It is important to keep in mind that transgene copy number and expression levels of the transgene do not necessarily correlate. When chromosomal integration occurs in the fertilized oocyte, the developing animal will be fully transgenic, since all nucleated cells in the animal will contain the transgene; however, if integration occurs later, at a postzygotic stage, the animal will be a mosaic, with some cells containing the transgene and others lacking it. Founders are identified through analysis of tail biopsies, and copy number and transgene integrity are then evaluated. To establish a transgenic line, a founder is bred to mice of a defined genetic background and offspring are characterized for expression of the transgene. Generally at least two founders should be maintained for experiments. Due to the nature of random integration, no two transgenic lines will be identical. Details on the identification and analysis of founders and the establishment of transgenic lines can be found in UNIT 23.9.

Transgenics carrying large segments flanking a gene of interest isolated as bacterial artificial chromosomes (BACs) have been used to minimize the abovementioned variation due to the effects of different integration sites (UNIT 23.1). The versatility of this strategy has been significantly increased by the development of homologous recombination-based strategies to mutagenize and manipulate the BAC constructs (UNITS 1.16 & 23.11) and the use of regulated transgenics (e.g., with tetracycline; UNIT 23.12). In the GENSAT project (<http://www.ncbi.nlm.nih.gov/projects/gensat>; Gong et al., 2003), for example, an atlas of gene expression in the mouse central nervous system is based on large numbers of BAC transgenics in which the protein-coding sequences of genes expressed in the brain are replaced by an EGFP reporter.

Knockouts

Gene knockout or ablation (UNIT 23.1) is a powerful strategy to define the range of phenotypes associated with the loss of function of a gene of interest. The reverse genetic strategy of one gene—one protein—one phenotype can often aid in elaborating complex “wiring” diagrams and placing genes into biochemical pathways such as those defined by environmental stimuli, cell surface receptors, downstream targets, and changes in gene expres-

sion. There are a number of potential problems with full-knockout GEMMs, however, particularly when the phenotypes of homozygous null mutants are extreme (e.g., when they result in embryonic or perinatal lethality).

This drawback has been overcome through the increasing use of conditional GEMMs, whose genomes are subject to temporal and spatial controls (UNIT 23.1). To generate conditional knockout mice, a conventional gene-targeting strategy is used to generate homologous recombinants, whereby parts of the gene of interest are flanked by site-specific recombination sequences such that the gene retains normal function in the embryonic stem (ES) cells and offspring mice. The most commonly used site-specific recombination system is the Cre-*loxP* system from P1 bacteriophage (Branda and Dymecki, 2004). Gene-targeting constructs are generated so that *loxP* elements flank the region of the genome to be deleted. Following germline transmission, mice that carry alleles flanked by *loxP* (“floxed” alleles) are subsequently bred with a tissue-specific Cre transgenic mouse. The resulting bigenic mice carry both floxed alleles and a Cre transgene (which can be tissue specific). The Cre recombinase mediates recombination between two *loxP* sequences, thereby excising the sequence between the two *loxP* sites and hence creating a gene deletion. An added sophistication to such spatially controlled mutagenesis is the inclusion of temporal control, which can be achieved by using ligand-activated chimeric recombinases (Metzger and Chambon, 2001). These are usually chimeric proteins in which the recombinase is fused with the ligand-binding domain of a nuclear receptor, such as the estrogen or progesterone receptor, enabling gene ablation at any time by administration of the corresponding nuclear receptor ligand, e.g., tamoxifen or RU-486 (Metzger and Chambon, 2001). The expression of a transgene can also be controlled in a temporal fashion through the use of either hormonally (e.g., ecdysone) or pharmacologically (e.g., tetracycline) inducible gene-expression systems.

Gene Trapping

Gene trapping is a mutagenesis strategy that involves the transfection of ES cells with a gene-trap vector, which is based on the use of a reporter and/or selection marker to identify a trapping event. One strategy involves a promoter trap, which carries a promoterless reporter gene and a selection marker gene under the control of its own promoter. In this instance, the reporter gene is transcribed only

when the construct is correctly inserted in the promoter region of an endogenous gene. Alternatively, there are polyadenylation trap strategies, in which a constitutive promoter drives the transcription of the selection markers. Because the polyadenylation-trap vector lacks a poly(A) signal, the transcript is only stabilized as long as the gene-trap vector is integrated in the correct orientation and captures the poly(A) signal of the trapped gene. Once a library of “trapped” ES cells has been selected, insertion points are precisely determined for each line by sequencing the region neighboring the gene-trap vector (Koutnikova et al., 2003; Hedrich, 2004).

Since gene-trapping mutagenesis techniques are relatively high-throughput strategies, it is expected that they will greatly facilitate functional gene annotation. For this reason, several consortia plan to carry out functional annotation of the entire mouse genome via systematic mouse genome mutagenesis using gene-trapping as their primary strategy (Austin et al., 2004; Auwerx et al., 2004). One disadvantage associated with gene-trapping is that the impact of the mutation on protein function must be carefully determined.

FORWARD GENETICS

In contrast to the reverse genetic approach, forward genetics is centered on phenotype-driven gene discovery. The approach consists of identifying a phenotype of interest in mutant mice derived by chemical mutagenesis, radiation mutagenesis, or even spontaneous mutation, and then determining the mutation responsible for the phenotype. Prior to the introduction of technologies for manipulating the mouse genome, this was the only genetic strategy available. Continued enthusiasm for forward genetics over specifically targeted strategies is due to the random nature of the mutations, which cover the full gamut from point mutations to insertions, deletions, and rearrangements, and thus produce a full spectrum of changes to protein structure and function.

Many of the mutant mouse models used to date in forward genetic strategies have arisen from spontaneous single-gene mutagenesis events that produced striking phenotypes. The most notable examples in the metabolic field are mutations in leptin (*Lep^{ob}*; Zhang et al., 1994), leptin receptor (*Lepr^{db}*; Chen et al., 1996), and tubby (*Tubby*; Noben-Trauth et al., 1996) genes. Human homologs of many murine obesity genes have a confirmed associ-

ation with monogenic obesity in man and have helped advance understanding of the molecular pathogenesis of obesity and its complications.

Mutations can be generated by chemical means using chlorambucil or ethylnitrosourea (ENU), and also by radiation (X-ray) mutagenesis of sperm or ES cells. ENU is generally the mutagen of choice (O'Brien and Frankel, 2003; Hedrich, 2004) and is an alkylating agent that can transfer its ethyl group to oxygen or nitrogen atoms at a number of reactive sites in DNA bases. This can cause mispairing of the two strands of DNA and consequently base-pair substitutions. ENU generates point mutations in DNA at a rate of $\sim 10^{-5}$, such that $\sim 1/1000$ of the offspring carry a mutation in a given gene. Although the generation of mutants is fast with ENU, the success of this technique depends on the implementation of efficient, robust, and validated high-throughput phenotyping resources for screening the progeny of treated mice for obvious phenotypic abnormalities. One major obstacle is that the mutations occur essentially at random, and the identification of the structural change(s) in the DNA generally requires a laborious breeding strategy combined with a positional cloning approach. This task remains timely and costly, as every mutant must be back-crossed to other inbred strains to identify markers that cosegregate with the mutation.

POLYGENIC DISEASES AND GENETIC REFERENCE PANELS

Currently, there is increased focus on understanding the pathogenesis of complex diseases such as cancer, neurodegenerative disorders, and the metabolic syndrome. Monogenic GEMMs or randomly mutagenized mice that focus on an isolated genetic locus in a fixed background cannot model polygenic diseases inherited in a non-Mendelian fashion. Moreover, environmental factors can affect the manifestations of the genotype, resulting in incomplete penetrance, or mimic the contribution, resulting in phenocopies. Thus, it is desirable to have mouse models that mimic the genetic structure of human populations. Oligogenic mouse models, which are engineered with mutations in two or more genes, can be valuable in such a context and have illustrated many of the genetic principles underpinning complex diseases. Some key principles to keep in mind include: (1) that a major predisposing allele can have a modest effect alone, but play a major role in the context of a predisposing background, (2) that mutations in

different unrelated genes can yield the same phenotype (Nandi et al., 2004), and (3) that the interaction among some genetic loci can create strong positive or negative (protective) effects in the predisposition to a disease. Other genetic principles such as those underlying intertissue communication and cross-talk have been elucidated through the use of spatially and temporally controlled GEMMs.

An increasingly popular tool for the study of mouse physiology that complements the forward and reverse genetic strategies with a polygenic approach is the mapping of quantitative trait loci (QTLs) in inbred mice. Genetically homogeneous inbred strains have an inherent wealth of variation due to past spontaneous mutation events, and they can reveal subtle variation affecting many traits of biomedical interest (Paigen, 2003). The Complex Trait Consortium was formed to systematically exploit large-scale resources with the aim of partitioning strain differences in hundreds of quantitative traits to sets of biologically related genes or loci, and thereby identify genomic regions and genes associated with these traits (Threadgill et al., 2002). However, mapping QTLs based solely on the parental strain phenotypes is not an effective strategy, since many genetic contributors to complex diseases will be missed (Li et al., 2005). This is because the power and accuracy in detecting a QTL depends on multiple factors, including the genetic diversity of the parents, the density of the markers, the trait heritability, and the size of the cross. In any single or double cross, recombination events and mapping resolution are thus limited, as only loci that show allelic variation between two strains can be detected.

Two recent advances improve collection of data from resource-intensive breeding crosses. The first is the use of multiple crosses (two or more single crosses), which improves the power and resolution (Li et al., 2005). With single intercrosses, it is difficult to distinguish multiple tightly linked QTL from a single QTL with a large effect, and the QTL intervals defined are generally quite large, typically 20 to 40 cM. With multiple crosses, a computational meta-cross analysis can be performed. When the same QTL is found in more than one cross, the data can be combined from the crosses for the chromosome on which the QTL is found. If the QTL from multiple crosses are caused by the same gene, the impact of combining the cross data is equivalent to that of adding more recombination events in the QTL region (e.g. by additional mice in a cross). This will

narrow the QTL region and increase the power for resolving the QTL. For example, a QTL for HDL cholesterol was found on chromosome 4 in four different crosses, and, by combining the data from these crosses, the QTL was narrowed from 30 cM to 10 cM. (Li et al., 2005). The second is the generation of several large genetic reference panels (GRPs) of recombinant inbred (RI) mouse lines that have a higher rate of polymorphism between strains. GRPs are a set of genetically well-defined strains derived by crossing two different highly inbred progenitor strains and then inbreeding random pairs of the F2 generation (>20 generations) to produce a series of RI strains that represent an ~50:50 genetic mixture of the progenitor strains (Churchill et al., 2004; Peirce et al., 2004; Fig. 29A.1.1). Thus, the strains that make up a GRP are not mutants or engineered mice, but are normal mouse lines with a level of variation that more closely resembles that seen among humans. GRPs are also efficient because they are a renewable resource; genotyping only needs to be performed once, and genetically identical animals can then be distributed to many researchers (Churchill et al., 2004).

With recent advances in functional genomic technologies, such as the widespread use of high-density microarrays, the power to decipher complex molecular networks underlying the multifactorial nature of common disorders has been markedly increased. This is best illustrated when applied to the analysis of GRPs. The approach initially applied in yeast involves carrying out genome-wide genetic analyses of gene expression data, which are then considered to be a quantitative trait and are correlated with microsatellite markers of DNA variation (Damerval et al., 1994; Jansen and Nap, 2001; Brem et al., 2002; Schadt et al., 2003; Bystrykh et al., 2005; Chesler et al., 2005). Those genetic regions (loci) that account for variation in the levels of gene expression, termed expression quantitative trait loci (eQTL), reveal previously unknown gene-gene interactions and biochemical pathways. Compared to classical clinical quantitative trait loci (cQTL), which often include gross clinical measurements that are far removed from their causative biological processes, such eQTL approaches offer an additional level of sophistication, especially when combined with the analysis of disease traits such as cQTL. Many “omic” disciplines such as proteomics, metabolomics, and pharmacogenomics have followed in the

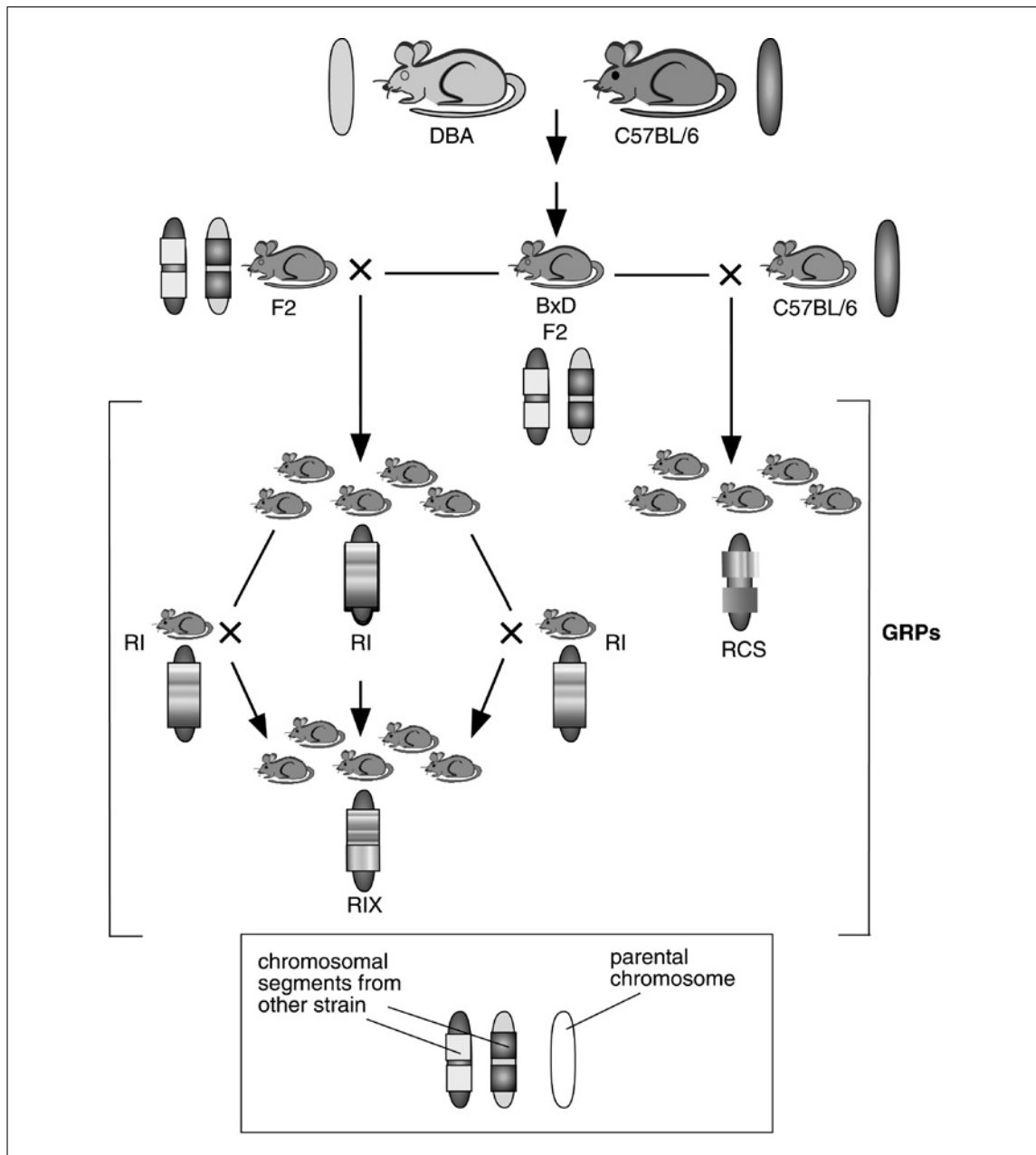


Figure 29A.1.1 GRPs can be defined as a set of genetically well-defined strains derived by systematical inbreeding, for at least 20 generations, from a single cross of two distinct inbred strains. GRPs represent an ~50:50 genetic mixture of the progenitor strains (Churchill et al., 2004; Peirce et al., 2004). Thus, the strains that make up a GRP are not mutants or engineered mice; they are normal mouse lines that have a level of variation more similar to that among humans. Approaches based on recombinant inbred (RI) strains are continually evolving to offer the variability, power, or resolution necessary for a general-purpose mapping tool, since at present there are too few GRPs with limited QTL power. Some of the sophistications include the use of recombinant congenic strains (RCS), which are variations of RI strains, whereby two parental inbred strains are initially crossed, but the resulting F1 hybrid progeny are then back-crossed with one of the parental strains prior to sib-pair matings. Instead of the mix of progenitor strain genomes approaching 50:50, the genome of RCS is representative mainly of one parent. The limitations of the current RI lines have also spurred an effort to create as many as 1000 new RI lines of mice within 7 years with the hope of covering much of the genetic variation present in natural populations. It will then be possible to generate an almost unlimited combinatorial diversity using the F1 progeny of these RI strains (RIX; Churchill et al., 2004).

footsteps of these genomic approaches. Application of these strategies to the analysis of either GEMMs or GRPs holds great promise and has clearly diversified the menu of mouse models to study human metabolic diseases.

It is becoming apparent that the best strategy for modeling complex diseases is to effectively combine data generated using relatively simple forward and reverse genetic strategies (whereby scientists drive gene elimination to precipitate disease) with global analyses of complex polygenic GRPs (where nature drives genetic variation to recreate disease). This combined strategy allows the benefits of the clear-cut results of single gene perturbations (e.g., knockout mice) to be merged with those conferred by the subtle alterations from innumerable allelic variants (GRPs).

LITERATURE CITED

- Argmann, C.A., Chambon, P., and Auwerx, J. 2005. Mouse phenogenomics: The fast track to "systems metabolism." *Cell Metab.* 2:349-360.
- Austin, C.P., Battey, J.F., Bradley, A., Bucan, M., Capecchi, M., Collins, F.S., Dove, W.F., Duyk, G., Dymecki, S., Eppig, J.T., Grieder, F.B., Heintz, N., Hicks, G., Insel, T.R., Joyner, A., Koller, B.H., Lloyd, K.C., Magnuson, T., Moore, M.W., Nagy, A., Pollock, J.D., Roses, A.D., Sands, A.T., Seed, B., Skarnes, W.C., Snoddy, J., Soriano, P., Stewart, D.J., Stewart, F., Stillman, B., Varmus, H., Varticovski, L., Verma, I.M., Vogt, T.F., von Melchner, H., Witkowski, J., Woychik, R.P., Wurst, W., Yancopoulos, G.D., Young, S.G., and Zambrowicz, B. 2004. The knockout mouse project. *Nat. Genet.* 36:921-924.
- Auwerx, J., Avner, P., Baldock, R., Ballabio, A., Balling, R., Barbacid, M., Berns, A., Bradley, A., Brown, S., Carmeliet, P., Chambon, P., Cox, R., Davidson, D., Davies, K., Duboule, D., Forejt, J., Granucci, F., Hastie, N., de Angelis, M.H., Jackson, I., Kioussis, D., Kollias, G., Lathrop, M., Lendahl, U., Malumbres, M., von Melchner, H., Muller, W., Partanen, J., Riccardi-Castagnoli, P., Rigby, P., Rosen, B., Rosenthal, N., Skarnes, B., Stewart, A.F., Thornton, J., Tocchini-Valentini, G., Wagner, E., Wahli, W., and Wurst, W. 2004. The European dimension for the mouse genome mutagenesis program. *Nat. Genet.* 36:925-927.
- Branda, C. and Dymecki, S. 2004. Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice. *Dev. Cell* 6:7-28.
- Brem, R.B., Yvert, G., Clinton, R., and Kruglyak, L. 2002. Genetic dissection of transcriptional regulation in budding yeast. *Science* 296:752-755.
- Bystrykh, L., Weersing, E., Dontje, B., Sutton, S., Pletcher, M.T., Wiltshire, T., Su, A.I., Velhenga, E., Wang, J., Manly, K.F., Lu, L., Chesler, E.J., Alberts, R., Jansen, R.C., Williams, R.W., Cooke, M.P., and de Haan, G. 2005. Uncovering regulatory pathways that affect hematopoietic stem cell function using 'genetical genomics'. *Nat. Genet.* 37:225-232.
- Chen, H., Charlat, O., Tartaglia, L.A., Woolf, E.A., Weng, X., Ellis, S.J., Lakey, N.D., Culpepper, J., Moore, K.J., Breitbart, R.E., Duyk, G.M., Tepper, R.I., and Morgenstern, J.P. 1996. Evidence that the diabetes gene encodes the leptin receptor: Identification of a mutation in the leptin receptor gene in db/db mice. *Cell* 84:491-495.
- Chesler, E.J., Lu, L., Shou, S., Qu, Y., Gu, J., Wang, J., Hsu, H.C., Mountz, J.D., Baldwin, N.E., Langston, M.A., Threadgill, D.W., Manly, K.F., and Williams, R.W. 2005. Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. *Nat. Genet.* 37:209-210.
- Churchill, G.A., Airey, D.C., Allayee, H., Angel, J.M., Attie, A.D., Beatty, J., Beavis, W.D., Belknap, J.K., Bennett, B., Berrettini, W., Bleich, A., Bogue, M., Broman, K.W., Buck, K.J., Buckler, E., Burmeister, M., Chesler, E.J., Cheverud, J.M., Clapcote, S., Cook, M.N., Cox, R.D., Crabbe, J.C., Crusio, W.E., Darvasi, A., Deschepper, C.F., Doerge, R.W., Farber, C.R., Forejt, J., Gaile, D., Garlow, S.J., Geiger, H., Gershenfeld, H., Gordon, T., Gu, J., Gu, W., de Haan, G., Hayes, N.L., Heller, C., Himmelbauer, H., Hitzemann, R., Hunter, K., Hsu, H.C., Iraqi, F.A., Ivandic, B., Jacob, H.J., Jansen, R.C., Jepsen, K.J., Johnson, D.K., Johnson, T.E., Kempermann, G., Kendziorski, C., Kotb, M., Kooy, R.F., Llamas, B., Lammert, F., Lassalle, J.M., Lowenstein, P.R., Lu, L., Lusk, A., Manly, K.F., Marcucio, R., Matthews, D., Medrano, J.F., Miller, D.R., Mittleman, G., Mock, B.A., Mogil, J.S., Montagutelli, X., Morahan, G., Morris, D.G., Mott, R., Nadeau, J.H., Nagase, H., Nowakowski, R.S., O'Hara, B.F., Osadchuk, A.V., Page, G.P., Paigen, B., Paigen, K., Palmer, A.A., Pan, H.J., Peltonen-Palotie, L., Peirce, J., Pomp, D., Pravenec, M., Prows, D.R., Qi, Z., Reeves, R.H., Roder, J., Rosen, G.D., Schadt, E.E., Schalkwyk, L.C., Seltzer, Z., Shimomura, K., Shou, S., Sillanpaa, M.J., Siracusa, L.D., Snoeck, H.W., Spearow, J.L., Svenson, K., Tarantino, L.M., Threadgill, D., Toth, L.A., Valdar, W., de Villena, F.P., Warden, C., Whatley, S., Williams, R.W., Wiltshire, T., Yi, N., Zhang, D., Zhang, M., Zou, F., and the Complex Trait Consortium. 2004. The Collaborative Cross: A community resource for the genetic analysis of complex traits. *Nat. Genet.* 36:1133-1137.
- Clarke, A.R. 2002. *Transgenesis Techniques: Principles and Protocols*, 2nd ed. Humana Press, Totowa, N.J.
- Damerval, C., Maurice, A., Josse, J.M., and de Vienne, D. 1994. Quantitative trait loci underlying gene product variation: A novel perspective for analyzing regulation of genome expression. *Genetics* 137:289-301.
- Gong, S., Zheng, C., Doughty, M.L., Losos, K., Didkovsky, N., Schambra, U.B., Nowak, N.J., Joyner, A., Leblanc, G., Hatten, M.E., and Heintz, N. 2003. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* 425:917-925.

- Hedrich, H.J. (ed.) 2004. *The Laboratory Mouse (The Handbook of Experimental Animals)*. Elsevier, New York.
- Jansen, R.C. and Nap, J.P. 2001. Genetical genomics: The added value from segregation. *Trends Genet.* 17:388-391.
- Koutnikova, H., Fayard, E., Lehmann, J., and Auwerx, J. 2003. Serial analysis of gene expression and gene trapping to identify nuclear receptor target genes. *Methods Enzymol.* 364:299-322.
- Li, R., Lyons, M.A., Wittenburg, H., Paigen, B., and Churchill, G.A. 2005. Combining data from multiple inbred line crosses improves the power and resolution of quantitative trait loci mapping. *Genetics* 169:1699-1709.
- Metzger, D. and Chambon, P. 2001. Site- and time-specific gene targeting in the mouse. *Methods* 24:71-80.
- Nagy, A., Gertsenstein, M., Vintersten, K., and Behringer, R. 2003. *Manipulating the mouse embryo*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Nandi, A., Kitamura Y., Kahn, C.R., and Accili, D. 2004. Mouse models of insulin resistance. *Physiol. Rev.* 84:623-647.
- Noben-Trauth, K., Naggert, J.K., North, M.A., and Nishina, P.M. 1996. A candidate gene for the mouse mutation tubby. *Nature* 380:534-538.
- O'Brien, T.P. and Frankel, W.N. 2003. Moving forward with chemical mutagenesis in the mouse. *J. Physiol.* 554:13-21.
- Paigen, K. 2003. Anecdotal, historical and critical commentaries on genetics. *Genetics* 163:1227-1235.
- Peirce, J.L., Lu, L., Gu, J., Silver, L.M., and Williams, R.W. 2004. A new set of BXD recombinant inbred lines from advanced intercross populations in mice. *BMC Genet.* 5:1-17.
- Schadt, E.E., Monks, S.A., Drake, T.A., Luskis, A.J., Che, N., Collnayo, V., Ruff, T.G., Milligan, S.B., Lamb, J.R., Cavet, G., Linsley, P.S., Mao, M., Stoughton, R.B., and Friend, S.H. 2003. Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422:297-302.
- Threadgill, D.W., Hunter, K.W., and Williams, R.W. 2002. Genetic dissection of complex and quantitative traits: From fantasy to reality via a community effort. *Mamm. Genome* 13:175-178.
- Zhang, Y., Proenca, R., Maffei, M., Barone, M., Leopold, L., and Friedman, J.M. 1994. Positional cloning of the mouse obese gene and its human homologue. *Nature* 372:425-432.
- <http://www.eumorphia.org>
EU Eumorphia program.
- <http://www.emma.rm.cnr.it>
European Mouse Mutant Archive (EMMA). Transgenic mouse repository.
- <http://www.informatics.jax.org>
Festing's Inbred Strain Characteristics.
- http://www.eulep.org/Necropsy_of_the_Mouse
Gross necropsy and tissue fixation.
- <http://www-mci.u-strasbg.fr/>
Institut Clinique de la Souris (ICS) or Mouse Clinical Institute.
- <http://www.informatics.jax.org/fimsr/index.jsp>
International Mouse Strain Resource. Aims to list all publicly available mouse strains.
- <http://www.informatics.jax.org>
Mouse Genetics: Concepts and Applications. L.M. Sliver's Mouse Genetics Book online.
- <http://www.informatics.jax.org/mgihome/nomen/table.shtml>
Mouse Genome Informatics.
- <http://www.biomednet.com/db/mkmd>
Mouse Knockout and Mutation Database. Searchable database for phenotypic information related to GEMMs.
- <http://www.mousephysio.com/>
MousePhysio.
- <http://www.mymouse.org>
Mymouse.org. Aims to facilitate international collaborative research among mouse users.
- <http://eulep.anat.cam.ac.uk/>
Pathbase.
- <http://www.eumorphia.org/EMPreSS/>
Standardized Operating Procedures. Covers phenotyping to necropsy.
- <http://tbase.jax.org>
TBASE. Database for transgenic and knockout mice.
- <http://www.hgmp.mrc.ac.uk/DHMHD/dysmorph.html>
The Dysmorphometric Human-Mouse Homology Database.
- <http://www.genenetwork.org/>
The Gene Network. Bioinformatic toolkit for genetic reference populations (GRPs).
- <http://www.jax.org>
The Jackson Laboratory. Mouse vendor and genetic resource center.
- <http://www.nih.gov/science/models/mouse/index.html>
Trans-NIH Mouse Initiative.
- <http://www.jax.org/phenome>
U.S. Mouse Phenome Project.
- <http://www.rodentia.com/wmc>
Whole Mouse Catalog. An information Web site.

INTERNET RESOURCES

- <http://genex.hgu.mrc.ac.uk>
Edinburgh Mouse Atlas (emap) and gene expression database (emage).
- http://www.ensembl.org/Mus_musculus
Ensembl Mouse Genome Server.
- <http://www.mgu.har.mrc.ac.uk/facilities/mutagenesis/mutabase/>
ENU Mutagenesis Program.

Contributed by Carmen A. Argmann
Institut de Génétique et de Biologie
Moléculaire et Cellulaire
Illkirch, France

Andree Dierich and Johan Auwerx
Institut de Génétique et de Biologie
Moléculaire et Cellulaire and Institut
Clinique de la Souris
Illkirch, France

Minimizing Variation Due to Genotype and Environment

The observation of different phenotypes in mouse strains that have undergone identical genetic manipulations, pharmacological interventions, and/or environmental stresses is not uncommon. This is because linking genotype directly to phenotype is often complicated by the fact that the genetic background of the mouse greatly impacts its phenotype. Thus, it is important that genetic factors such as strain background be considered in the context of any phenotypic studies, as they have an influence on almost all variables. Results obtained in one genetic background do not necessarily translate into another background. While this is perhaps the most obvious caveat that may affect the outcome of a phenotype, it is by certainly not the only one. In reality, there is a plethora of confounding factors (Table 29A.2.1), including genetic, developmental, environmental, and technical.

GENETIC BACKGROUND

Discrepancies in the outcomes of phenotyping tests reported by different laboratories evaluating the same genotype have provided a major driving force in addressing the need for standardization. For example, parallel testing of mouse behavior carried out by three independent laboratories on mice produced from the same litters and breeding stocks, in similar setups with identical protocols, led to significantly different results (Crabbe et al., 1999). In this experiment, the genotype accounted for

a significant proportion of the total variability in behavior, and historically documented strain differences were reproduced across all laboratories. For example, C57BL/6J mice are known to prefer ethanol, while DBA/2J mice avoid it. Similarly, C57BL/6J are good learners whereas DBA/2J are poor learners in hippocampus-dependent tasks (Upchurch and Wehner, 1988; Stiedl et al., 1999). These types of strain differences may significantly affect the outcome of certain phenotypic analyses.

Strain type must also be carefully chosen for metabolic/endocrine studies (addressed in Section 29B), as different strains are known to respond differently to metabolic challenges. For example, C57BL/6J mice are more susceptible to diet-induced diabetes and atherosclerosis compared to other strains (Paigen et al., 1985; Schreyer et al., 1998). C57BL/6J mice have low plasma cholesterol and low plasma triglycerides (TG), but high glucose levels in contrast to C3H mice, which generally have low blood pressure and high plasma cholesterol, TG, and glucose. Inbred 129 strains have a very low glucose production rate but high plasma cholesterol and low TG levels (Burgess et al., 2005). In addition, 129 strains are generally less active (Seburn, 2001) and are poor breeders. For additional information on mouse strain characteristics, readers should refer to the Mouse Phenome Database at <http://aretha.jax.org/pub-cgi/phenome/mpdcgi?rtn=docs/home> and to

Table 29A.2.1 Genetic, Environmental, and Experimental Factors that Influence Phenotypic Analysis

	Possible confounding factors
Animal	Age, gender, ovarian cycle, health status, treatment with drugs
Development	Intrauterine position effect, intrauterine feeding, maternal behavior
Genetics	Genetic background, gender, compensation, pleiotropy
Environment	Housing conditions, handling, stress, social hierarchy
Experimenter	Level of experience, same or different experimenters
Experimental design	Construction, standardization, illumination, time of testing (e.g., circadian rhythm)
Data acquisition and analysis	Calibration/thresholds, automated versus observed, experimenter bias, normalization, statistics

Contributed by Carmen A. Argmann and Johan Auwerx
Current Protocols in Molecular Biology (2006) 29A.2.1-29A.2.3
Copyright © 2006 by John Wiley & Sons, Inc.

Michael Festing's index of inbred strains at <http://www.informatics.jax.org/external/festing/rat/STRAINS.shtml>.

Since C57BL/6J mice have a good breeding reputation, they are frequently the preferred genetic background for a large number of congenic strains, covering both polymorphic and mutant loci. However, due to the preferred use of 129 embryonic stem (ES) cells for gene targeting and gene trapping, a large number of studies have been performed in gene-targeted mice on a non-homogeneous genetic background. Yet the genetic background of the ES cells used to generate mutant mice has been proven to influence their phenotype. If the background strain of the phenotyped mouse is different from the one used to generate the knockout, the mice must be back-crossed for at least nine generations to avoid confounding factors contributed by flanking donor chromosomal DNA (Banbury Conference on Genetic Background in Mice, 1997). Alternatively, flanking donor chromosomal DNA can be monitored by using a marker-assisted selection protocol, enabling a quicker production of congenic strains compared to traditional backcrosses (Collins et al., 2003). Overall, these strain-specific differences show why it is very important to select the most appropriate strain for a particular study.

ENVIRONMENT

The environment is also critical in the outcome of phenotypic testing and must be carefully controlled to compare results from different laboratories. In one study, mice tested in a laboratory in Edmonton were on average more active than mice of the same genetic background tested in Albany and Portland (Crabbe et al., 1999). Whereas some of environmental factors have become experimental tools with which to challenge animals (e.g., diet), many others have become standard "housekeeping" parameters, whose order and consistency must be maintained. It is vital that the potential variables in phenotypic testing be minimized or carefully noted, so that only gene-elicited, environmentally defined mutant phenotypes are measured with similar outcomes wherever and whenever testing takes place.

Typical housekeeping items include animal housing and handling conditions such as the number of animals per cage (housing density), diurnal rhythm, length of fasting, blood collection procedure (retro-orbital or tail puncture, with or without anesthesia), diet, and age and gender of the mice (Champy et al., 2004). Cage housing density is an important

contributor to stress and significantly impacts many metabolic variables. Group-housed male mice behave differently than singly housed males with respect to territorial behavior and the formation of social hierarchies (Würbel, 2001). Similarly, housing conditions can influence many metabolic parameters. Glucose levels are significantly elevated when animals are caged in pairs as compared to when they are housed in groups of four, in part due to the increased stress and reduced competition for food intake between mice housed in pairs (Champy et al., 2004). Finer housekeeping details, such as the timing of blood collection, should also be critically controlled since metabolic variables are affected by activity and food intake, which is higher at night, and hormonal variations, which show marked nycthemeral (24-hr) rhythms (Champy et al., 2004).

Behaviors can also be affected by sex and age. The estrus status affects both emotionality and cognition in mice, and DBA/2J mice are known to become deaf by adulthood (Willot et al., 1995; Wotjak, 2004, and references therein). In a recent study involving a biochemical screen for thirteen analytes pertaining to lipoprotein metabolism, electrolyte balance, and organ function in C57BL/6J mice, twelve of the analytes were affected by the age and gender of the mice tested (Zhou and Hansson, 2004).

Developmental factors also contribute to the "individuality" and/or phenotype of the mouse, or to differences that occur despite seemingly identical genetics or environmental conditions (Lathe, 2004). Activation of the immune system during gestation, intrauterine factors (position of the embryo, nutrient and endocrine exchanges, and so on), and maternal behavior (grooming and licking) have all been shown to be determinants of behavior of adult offspring. For example, mice whose mothers displayed greater nursing activity showed a higher resistance to the deteriorating effects of stress in adulthood (Lathe, 2004, and references therein). These factors, despite being beyond standardization, are nonetheless important. In the study by Crabbe et al. (1999), for example, where genetic factors as well as laboratory environments were essentially similar, these effects may offer further explanations.

CONCLUSION

It is surprising that the impact of these important confounding factors on phenotyping has only recently been appreciated. In response to these observations, major

systematic phenotyping efforts, such as the European Union's Eumorphia program (<http://www.eumorphia.org/EMPreSS>), have been designed to standardize and validate phenotyping tests, linking them to standardized operating procedures (SOPs; Green et al., 2005). Quality standards and management, not only in industry but also in academic research, should ensure that SOPs and validated test paradigms guarantee highly reproducible results within one laboratory and between different research institutions. Obviously, standardization is limited by the confounding influences of the experimenter and the different resources provided by research institutions. Nonetheless, most of the confounding variables described above can be controlled in any given laboratory by following these guidelines.

1. Reduce genetic variation by using congenic mouse lines.
2. Reduce environmental variation through careful housekeeping.
3. Use mice of the same age and gender.
4. Use standardized operating procedures (SOPs).
5. Be conscientious and minimize sources of experimenter variation.
6. Repeat the experiment several times to ensure confidence in the data.

LITERATURE CITED

- Banbury Conference on Genetic Background in Mice. 1997. Mutant mice and neuroscience: Recommendations concerning genetic background. *Neuron* 19:755-759.
- Burgess, S.C., Jeffrey, F.M., Storey, C., Milde, A., Hausler, N., Merritt, M.E., Mulder, H., Holm, C., Sherry, A.D., and Malloy, C.R. 2005. Effect of murine strain on metabolic pathways of glucose production after brief or prolonged fasting. *Am. J. Physiol. Endocrinol.* 289:E53-E61.
- Champy, M., Selloum, M., Piard, L., Zietler, V., Caradec, C., Chambon, P., and Auwerx, J. 2004. Mouse functional genomics requires standardization of mouse handling and housing conditions. *Mamm. Genome* 15:768-783.
- Collins, S.C., Wallis, R.H., Wallace, K., Bihoreau, M.T., and Gauguier, D. 2003. Marker-assisted congenic screening (MACS): A database tool for the efficient production and characterization of congenic lines. *Mamm. Genome* 14:350-356.
- Crabbe, J.C., Wahlsten, D., and Dudek, B.C. 1999. Genetics of mouse behavior: Interactions with laboratory environment. *Science* 284:1670-1672.

- Green, E.C., Gkoutos, G.V., Lad, H.V., Blake, A., Weekes, J., and Hancock, J.M. 2005. EMPReSS: European mouse phenotyping resource for standardized screens. *Bioinformatics* 21:2930-2931.
- Lathe, R. 2004. The individuality of mice. *Genes Brain Behav.* 3:317-327.
- Paigen, B., Morrow, A., Brandon, C., Mitchell, D., and Holmes, P. 1985. Variation in susceptibility to atherosclerosis among inbred strains of mice. *Atherosclerosis* 7:65-73.
- Schreyer, S.A., Wilson, D.L., and LeBoeuf, R.C. 1998. C57BL/6 mice fed high fat diets as models for diabetes-accelerated atherosclerosis. *Atherosclerosis* 136:17-24.
- Seburn, K. 2001. Mouse Phenome Database Project: Metabolic Characterization. The Jackson Laboratory, Bar Harbor, Maine.
- Stiedl, O., Radulovic, J., Lohmann, R., Birkenfeld, K., Palve, M., Kammermeier, J., Sananbenesi, F., and Spiess, J. 1999. Strain and substrain differences in context- and tone-dependent fear conditioning of inbred mice. *Behav. Brain Res.* 104:1-12.
- Upchurch, M. and Wehner, J.M. 1988. Differences between inbred strains of mice in Morris water maze performance. *Behav. Genet.* 18:55-68.
- Willott, J.F., Erway, L.C., Archer, J.R., and Harrison, D.E. 1995. Genetics of age-related hearing loss in mice. II. Strain differences and effects of caloric restriction on cochlear pathology and evoked response thresholds. *Hear Res.* 88:143-155.
- Wotjak, C.T. 2004. Of mice and men. *B.I.F. Futura* 19:158-169.
- Würbel, H. 2001. Ideal homes? Housing effects on rodent brain and behaviour. *Trends Neurosci.* 24:207-211.
- Zhou, X. and Hansson, G.K. 2004. Effect of sex and age on serum biochemical reference ranges in C57BL/6J mice. *Comp. Med.* 54:176-178.

Contributed by Carmen A. Argmann
 Institut de Génétique et de Biologie
 Moléculaire et Cellulaire
 Illkirch, France

Johan Auwerx
 Institut de Génétique et de Biologie
 Moléculaire et Cellulaire and Institut
 Clinique de la Souris
 Illkirch, France

Collection of Blood and Plasma from the Mouse

**BASIC
PROTOCOL**

When working with mice, venous blood is commonly collected from the retro-orbital plexuses or the tail vein, or by cardiac puncture (Hedrich, 2004). The choice of method will depend on the quantity of blood required, the frequency of the blood draw, the skill of the blood drawer, and local animal handling guidelines. In the authors' experience, blood collection via the tail vein is the easiest method and provides sufficient material for most analyses, including lipid analysis. Furthermore, this method is used for blood sampling during numerous types of dynamic hormone tests, such as glucose tolerance tests.

In the tail of the mouse, there are three veins and one artery (Fig. 29A.3.1). As the lateral caudal veins are thicker than the dorsal caudal vein, the former are generally better for collecting large amounts of blood. The artery can also be used; however, bleeding time is longer. In general, blood is collected from awake and unrestrained mice; in this case, the procedure is easier with two persons. Alternatively, a restraining device can be employed and the blood draw can be performed by one person.

Although sometimes used interchangeably, serum and plasma are qualitatively different. In plasma, only the cells have been removed by centrifugation. In serum, the bulk of the fibrinogen has been removed by conversion into a fibrin clot together with the platelets, which have been physically bound in the fibrin matrix or activated to form aggregates or both. Since varying amounts of other proteins can be removed in the fibrin clot through interactions, and since the clotting process can induce cellular elements to secrete certain components, the quantities of various analytes (e.g., potassium) can be different in serum versus plasma. Since plasma is prepared in the presence of an anticoagulant such as EDTA, coagulation does not occur and fibrinogen is not consumed. In general, serum or plasma samples can be used for most analytes, including basic biochemistry and lipid analysis. For lipoprotein profiles, however, plasma is generally used. The researcher

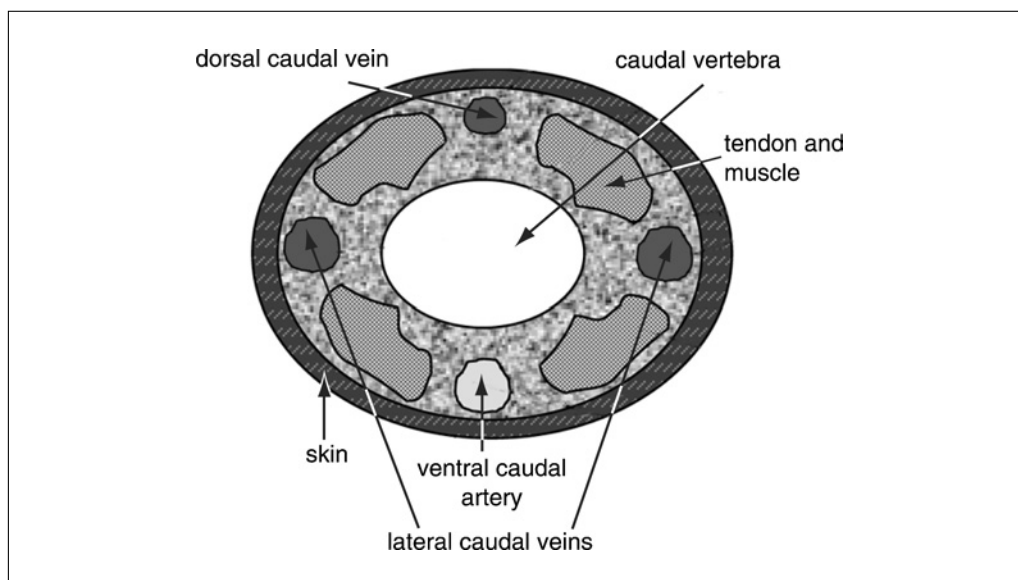


Figure 29A.3.1 A schematic drawing of a transverse section of a mouse tail showing the location of the veins and artery. The lateral caudal veins are used for venous blood collection.

should always verify for each analyte the limitations of the experimental protocols, and should use the same processing procedure when making comparisons.

NOTE: For a more complete coverage of mouse handling (e.g., the use of restraining devices) and blood collection, refer to Hoff (2000) and Coligan et al. (2006).

Materials

Mouse

Hematocrit sealing compound (e.g., <http://www.vitrex.dk>)

Alcohol-soaked cotton pads

Sterile gauze

Surgical blade, no. 10 or 11

Microhematocrit capillary tubes (with or without Na-heparin or EDTA, depending on analyte and plasma versus serum)

5-ml polypropylene centrifuge tubes

Refrigerated centrifuge

100- μ l Hamilton syringe

1. Place the mouse on the grid of a clean cage. Clean the tail with an alcohol-soaked cotton pad and then wipe it dry with sterile gauze.
2. Grip the end of the mouse's tail in one hand in such a way that the section of the tail to be cut (\sim 1 to 2 cm from tip of tail) is held steady while the mouse is moving about.
3. Using the other hand, locate the veins running on one side of the tail.

In general, if this procedure is performed on white or light-colored mice, the vein can be located visually. In black or dark-colored mice, this is more difficult and requires knowing the location of the vein.

4. Using a surgical blade held at a 90° angle with respect to the long axis of the tail, make a small cut on the side of the tail near the end.

Blood should immediately flow if the cut has been made in the right location. Care should be taken not to cut too deeply, as the tail cartilage can be affected. To improve the blood flow, the tail vein can be prewarmed by placing it in water at 40°C. A more practical and common alternative is to massage the length of the mouse's tail, as this will promote blood flow. However, for hematological analyses, massaging the tail is not recommended as this will result in leukocytosis, which increases the white blood cell count.

5. While holding and/or massaging the tail, touch a glass capillary tube to the blood droplets accumulating on the mouse's tail and allow the blood to enter the capillary tube by capillary action. To promote entry of blood into the tube via gravity, lower the opposite end of the tube so that it is below the blood droplet.

In the absence of a restraining device and with two persons performing blood collection, steps 1 to 4 should be performed by one person and the capillary tube should be manipulated by the other.

Using this procedure, at least two microhematocrit capillary tubes can be filled with blood, which corresponds to \sim 200 μ l of blood from an average-sized mouse.

For plasma, use sodium-heparinized or EDTA-containing tubes to prevent coagulation. For serum, use an untreated tube, as serum is collected in the absence of an anticoagulant.

6. Plug the capillary tubes using hematocrit sealing compound and place them in a 5-ml plastic tube on ice.
7. After blood collection, compress the cut with sterile gauze until the bleeding stops.

8. Return the mouse to its cage and monitor the animal for signs of anemia, including rapid breathing, muscle weakness, and paling of mucous membranes. Observe the site of blood collection daily for infection or irritation. When large quantities of blood are drawn, allow at least 2 weeks for recovery prior to performing another blood draw.
9. To collect plasma, centrifuge the 5-ml tube (containing the hematocrit tubes) 15 min at 3000 rpm ($1500 \times g$), 4°C .

The blood generally separates into three phases, with the bottom phase consisting of red blood cells, the top phase consisting of serum, and an interphase (referred to as the buffy coat) generally resembling a thin white layer containing mainly immune-type cells.

For serum collection, keep blood at room temperature for 1 hr in the untreated collection tubes, then centrifuge as described above.

10. Collect the plasma (or serum) from the top with a Hamilton syringe, leaving behind the buffy coat and red blood cells.

The plasma (or serum) is ready for analysis. If necessary, store at -20°C or -80°C .

COMMENTARY

Background Information

Cardiac puncture and orbital eye bleed generally yield large quantities of blood ($\sim 300 \mu\text{l}$ per draw) as compared to tail vein sampling ($\sim 150 \mu\text{l}$ per draw), assuming that an average-sized mouse is used and it is not overbled. However, these former techniques may require the sacrifice of the animal (where blood draw quantity can be even higher, $\sim 900 \mu\text{l}$), may blind the animal, and require a great deal of skill. Thus, the tail vein method provides a less invasive procedure that allows repeated blood draws from the same animal, without the confounding affects of anesthesia.

Critical Parameters

If analyses are not performed on the day of blood collection, samples should be stored at -20°C (short-term, e.g., < 2 months) or -80°C (long-term). Stability and storage depend on the different parameters to be determined. For example, free fatty acids are only stable for 2 to 3 days at -20°C .

Plasma samples can be used undiluted or diluted with deionized water if the volume is insufficient or the parameter to be measured exceeds accurate detection limits. If the blood is hemolyzed, a correction factor and/or re-sampling is suggested, as hemoglobin can interfere with certain measurements.

For plasma, a number of analyte measures favor the use of heparin over EDTA and vice versa, and this should be verified with each individual protocol before blood is drawn. For example, heparin can interfere with the assessment of plasma lipoprotein lipase activity, as heparin is a known inhibitor of this enzyme.

Because any invasive mouse technique should be considered potentially stressful, the same experimenter should perform the technique each time to minimize experimental variation due to animal handling skills. Since many factors such as centrifuge speed, clotting time, storage temperature, and time for blood processing may influence both serum and plasma preparations, carefully documented conditions of blood processing procedures should be followed to assure reproducibility of samples.

Anticipated Results

As tail vein sampling is less invasive than orbital eye bleed, multiple samplings should be well-tolerated by the mouse over the period of its lifetime. It is difficult to quantify how many samplings are possible or how much blood can be drawn at any one point in time. In general, however, the maximal volume of a single blood draw should not exceed 10% to 15% of total blood volume or 1% of the body weight of the animal. Restoration of normal blood volume takes 24 hr, but it may take up to 2 weeks for normal levels of erythrocytes and reticulocytes to be restored. Therefore, 2 weeks is the recommended minimum interval between consecutive maximal blood draws. Daily blood draws up to 1% of the total blood volume are acceptable, although the effects of stress, as well as the choice of collection site, need to be carefully considered. Blood loss exceeding 20% to 25% of total volume results in hypovolemia and cardiac failure due to decreased blood pressure, cardiac output, and oxygen delivery to vital organs.

The ratio of blood cells to plasma is typically 1:1, but a researcher should safely estimate that only ~50 μ l of plasma will be obtained from 150 μ l blood. Note that quality should not be sacrificed for quantity by removing too much plasma near the buffy coat and red blood cell layers and risking contamination. For automated analyzers such as for lipid analysis, this amount of plasma/serum is sufficient to perform a basic biochemical and lipid profile. Although the low yield of plasma from tail vein sampling is perceived as a drawback when multiple analytes are to be measured, the technique is ideal when multiple samples are required over a short period of time and the potential confounding effects of anesthesia are undesired. For example, the tail vein technique is ideal for a glucose tolerance test, which requires as many as eight blood samplings over 180 minutes but requires only 5 μ l of blood per time point.

Time Considerations

For blood draw volumes and frequencies, see Anticipated Results. Because the same experimenter should perform all blood draws, the maximum number of animals that can undergo blood draw at any time will depend on

the skills of the experimenter. An experienced blood drawer might be able to obtain samples from as many as 20 to 30 animals in 1 hr. The whole experiment should be performed within a reasonable time frame to avoid any confounding effects of changes in the animals due to diurnal rhythm and length of fasting.

Literature Cited

- Coligan, J.E., Bierer, B.E., Margulies, D.H., Shevach, E.M., and Strober, W. 2006. *Current Protocols in Immunology*. John Wiley & Sons, Hoboken, N.J.
- Hedrich, H. (ed.) 2004. *The Laboratory Mouse*. Elsevier Academic Press, New York.
- Hoff, J. 2000. Methods of blood collection in the mouse. *Lab. Animal* 29:47-53.

Contributed by Carmen A. Argmann
Institut de Genetique et de Biologie
Moleculaire et Cellulaire
Illkirch, France

Johan Auwerx
Institut de Genetique et de Biologie
Moleculaire et Cellulaire and Institut
Clinique de la Souris
Illkirch, France

METABOLIC EXPLORATION OF THE MOUSE

Endocrine and metabolic dysfunctions are among the most common diseases in developed societies. Examples of such diseases include obesity, insulin resistance, type 2 diabetes, hyperlipidemia, and atherosclerosis. Mouse models are increasingly popular tools for studying and characterizing the molecular and physiological aspects of these diseases. The characterization of endocrine and metabolic dysfunction in mice has become of utmost importance for both basic and clinical scientists. This section will describe phenotyping protocols for the evaluation of numerous metabolic parameters including energy homeostasis, serum and tissue lipids, glucose homeostasis, and select pathophysiology. Many of these protocols are based on clinical tests used to evaluate human patients with metabolic diseases.

Despite the ease of mouse manipulation, the investigator should take heed that endocrine and metabolic disturbances are often subtle in their presentation, and are thus a challenge to detect even for the best mouse physiologists. In humans, clinical history and physical examination constitute critical elements in the general diagnostic process. Although a clinical history in mice is impossible, physical examination, as with humans, can provide important hints about endocrine or metabolic dysfunction. Numerous examples illustrate this principle nicely. For example, diseases of hormone deficiency or excess are physiologic determinants of physical traits including stature, weight, complexion, hairiness, and behavior. Body aspects such as thinness or obesity can result from or be associated with abnormalities in lipid or glucose metabolism. Critical observation of mouse behavior also reveals important information. Therefore, initial metabolic exploration should begin by an in-depth clinical analysis of the mouse in its home cage using tests such as SHIRPA or the dysmorphological screen (Hrabe de Angeles et al., 2000; Nolan et al., 2000). Screening of standard blood components like glucose, lipids, calcium, and electrolytes also gives clues for the diagnosis of endocrine or metabolic dysfunction, which can be further investigated with a wide array of more specialized and detailed tests covering the whole endocrine system and metabolism in general.

The acquisition of skills in endocrine and metabolic diagnosis is facilitated by familiarity with certain simple principles and rules. In particular, endocrine and metabolic disorders are most often seen as either an excess or deficiency of a certain biologically active hormone or metabolite. Alternatively, these disorders can be the consequence of inadequate cellular responses to a particular hormone or metabolite. These principles help guide the phenotyping paradigms summarized in the following units.

Hormones and metabolites can be characterized by the pathway and location of their production, their storage and release into the serum, and their degradation and metabolism. Although hormone and metabolite levels can be regulated in all of these aspects, production rate is often the primordial factor determining their levels. It is interesting that production of most metabolites and hormones is regulated directly or indirectly by the metabolite or hormone in question through interconnected positive and negative feedback loops. This control has important implications, since plasma levels of a metabolite or hormone only make sense if the appropriate regulatory factor is taken into account (e.g., glucose and insulin). Elevation of both metabolites/hormones and their regulatory factors is often reflective of a resistant state (e.g., glucose and insulin elevation in insulin

resistance). A final interesting correlate of regulated hormone secretion is that it is the basis for various dynamic endocrine or metabolic tests for the evaluation of discrete disturbances in metabolic and endocrine function (e.g., intraperitoneal insulin tolerance test).

UNIT 29B.1

Evaluation of Energy Homeostasis

The critical factors contributing to body-weight maintenance include caloric intake and energy expenditure (Spiegelman and Flier, 2001). Alteration of either of these factors can upset the balance of the equation and result in body weight changes. Thus, the evaluation of body weight gain/loss over a period of time can be the most obvious indication of an alteration in energy homeostasis, and is typically coupled with other direct measurements such as food intake, body mass index, and fecal lipid content (see Basic Protocol 1). More sophisticated, noninvasive measurements of body composition that break down the weight in terms of body fat, lean mass, and bone mass can be obtained using dual-energy X-ray absorptiometry (DEXA; see Basic Protocol 3) or quantitative nuclear magnetic resonance (QNMR; see Alternate Protocol). To investigate the causes of energy dysregulation, several tests can be employed. Food-intake measurements such as those described in Basic Protocol 1 can indicate whether changes in caloric intake are contributing to the altered energy balance, whereas indirect calorimetry (see Basic Protocol 4) and the cold test (see Basic Protocol 2) assess the animal's expenditure of energy under basal and challenged conditions, respectively.

NOTE: All protocols using live animals must first be reviewed and approved by an Institutional Animal Care and Use Committee (IACUC) and must follow officially approved procedures for the care and use of laboratory animals.

STRATEGIC PLANNING

In vivo mouse experiments require a significant amount of strategic planning. To begin with, all in vivo experiments must be done in accordance with the appropriate animal care and use guidelines as defined by the country and institution where the work is carried out, and the time to acquire these approvals must be taken into consideration. Moreover, several parameters must be considered when designing a mouse experiment, which depend on the nature of the experiment. A minimum list of parameters to consider when planning the generation of a cohort of experimental animals for metabolic exploration includes mouse strain, gender, age, control mice, diet, housing conditions, time of experiment, and number of mice. For more information, refer to Champy et al. (2004).

For any experimental cohort, several variables need to be carefully controlled, such as the number of mice and their age, sex, and diet (Champy et al., 2004). In general, for metabolic exploration, male or female animals are tested after sexual maturity (~6 to 8 weeks), under chow, high-fat, and/or high-sucrose diet conditions. Preferably, all genotypes of animals should be derived from age-matched littermates with appropriate cohort sizes, since it is clear that there is substantial inter-animal variability. According to the authors' own experience and previously published work with C57BL/6J mice and power calculations, a group size of approximately ten to twelve animals is generally sufficient to observe statistically significant differences of the distinct interventions on metabolic and molecular parameters (Koutnikova et al., 2003; Cock et al., 2004). To be cautious and account for animals that might die in the course of these often long-term experiments, twelve animals per group are suggested; however, each individual user must determine the optimal number. In order to generate such a colony, depending on the size of litters and fertility of the animals, at least six to eight females will need to be bred at any one time. In cases where the number of animals is limiting, a consecutive

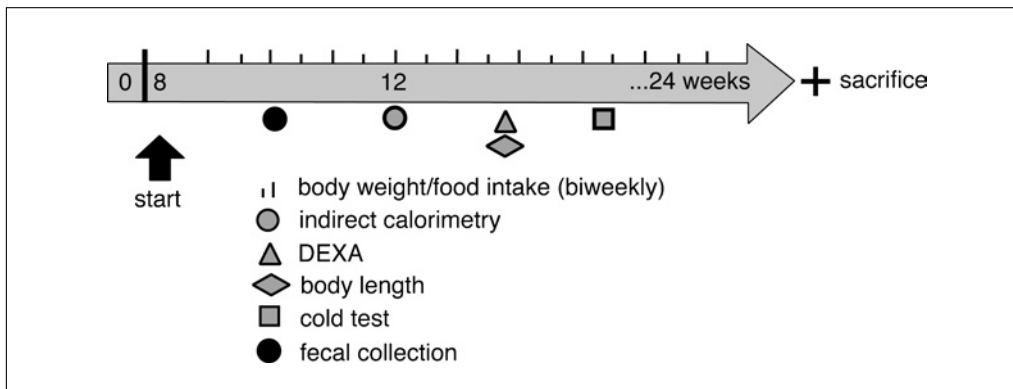


Figure 29B.1.1 A hypothetical timeline for a phenotypic analysis on a cohort of animals.

experimental strategy can be employed whereby animals are first tested under chow-fed conditions before putting them on the dietary challenge. Finally, a timeline outlining the experimental manipulations must be designed that minimizes the stress level of the animal. An example of such a timeline is provided for the procedures described in this unit (Fig. 29B.1.1); additional tests not discussed in this section can also be incorporated.

MEASUREMENT OF BODY WEIGHT, FOOD INTAKE, BODY MASS INDEX, AND FECAL LIPID CONTENT

Monitoring of body weight and food intake involves relatively simple tests that provide the first indications about metabolic homeostasis. The authors recommend initially evaluating body weight and food intake during a 6-week period, between 6 and 12 weeks of age. Body weight should be recorded twice a week at the same time of day. Care should be taken to avoid stressing the animals, since this can impact body weight gain. Twice-weekly weighing of the food given and the food remaining in the grid of the cages is a reliable method for evaluating food consumption. In general, food intake is measured when mice are housed at a density of four per cage. In the case of longer studies (e.g., aging), the frequency of weighing can be reduced to bimonthly recordings, and food intake can be measured less often, depending on results of initial testing.

Fecal lipid content is measured gravimetrically following a Folch extraction (Folch et al., 1951; Ishibashi et al., 1996). The authors recommend that feces be collected in parallel on a day when food intake is recorded, in order to measure lipid absorption. As an alternative to measuring total lipid gravimetrically, the fat extract can be partitioned according to the type of lipid (i.e., triglyceride and cholesterol) by subjecting it to enzymatic assays, thin-layer chromatography, or gas chromatography. These methods are beyond the discussion of this unit, although enzymatic assays will be included in a later unit on glucose and lipid homeostasis. A very sensitive measure of cholesterol absorption involves oral gavages of a mixture of dual-labeled [^3H]sitostanol and [^{14}C]cholesterol followed by determination of the stool sterol ratio of ^{14}C to ^3H (Schwarz et al., 2001). Whole-animal sterol synthesis rates can also be measured *in vivo* by determining the rate of incorporation of $^3\text{H}_2\text{O}$ into sterols over time (see Schwarz et al., 2001, and references therein). Finally, alternatives to the Folch lipid extraction protocol can be employed. For example, increased extraction of total fat and phospholipid can be achieved by sequential extraction with solvents chosen to optimize extraction of medium- and long-chain triglycerides and fatty acids and solvents designed to recover phospholipids (Chen and Innis, 2004).

CAUTION: Appropriate procedures for the use of radioactivity must be followed (see *APPENDIX 1F*).

BASIC PROTOCOL 1

Mouse Phenotyping

29B.1.3

Materials

Mice
Mouse food
0.1 mCi/ml [carboxyl-¹⁴C]triolein (112 mCi/mmol)
2:1 (v/v) chloroform/methanol
Container to accommodate mouse during weighing
Scale accurate to 0.01 g
Mouse cages (e.g., 24 × 38 × 20–cm) with metal floor grids (grids or screens can be custom cut to fit bottom of cage; optimal grid size is one that allows feces to pass through but is still comfortable for mouse to rest on)
Inhalation anesthesia unit (composed of a mixing and flow-controlled system for O₂/isoflurane anesthesia; e.g., TEM, <http://www.TEM.fr>)
Digital caliper (range, 0 to 20 cm; accuracy, <0.02 mm; resolution, <0.01 mm; Ted Pella, Inc.)
Analytical balance accurate to 0.0001 g
70°C vacuum oven
10-ml Erlenmeyer flasks
60°C shaking water bath
Whatman no. 1 filter paper
10-ml glass test tubes with ground glass stoppers
Clinical tabletop centrifuge
Long-stem glass Pasteur pipets
Nitrogen evaporator (e.g., N-EVAP; Organomation Associates) with dry bath at 50°C
Preweighed glass scintillation vials
β-counter and scintillation cocktail

Measure body weight

1. Place a mouse container on a scale with an accuracy of 0.01 g and calibrate to zero.
2. Place the mouse in the container and record the weight (to the same accuracy) after the measurement stabilizes.

Measure food intake

3. Place the mice in a clean standard cage (e.g., 24 × 38 × 20–cm) with a metal floor grid, optimally at four mice per cage, with free access to water.
4. Weigh a portion of food appropriate for the feeding interval (e.g., 24, 48, or 72 hr) to an accuracy of 0.001 g, and place it in the grid of the cage. Assume an average mouse consumes 3 to 5 g/day. Place a lid over the grid to prevent the food from spilling over.

For long-term food intake studies, the food quantity should be checked daily and supplementations weighed.

5. At the end of the feeding interval, remove the grid and weigh the remaining food to an accuracy of 0.001 g.
6. Calculate the amount of food consumed per mouse per day by dividing food consumed (weight of food given – weight of food remaining) by the number of animals per cage and by the number of days in the feeding interval.
7. Calculate kilocalories (kcal) of food consumed by multiplying the amount of food consumed by the caloric density of the food as provided by the manufacturer.

For example, for a chow diet of 3.37 kcal/g with 5 g food consumed, the number of kcal consumed is 16.85 kcal.

As an alternative, systems designed to specifically record and analyze liquid and food consumption can be used—e.g., Columbus Instruments Lick Counter and Feed Monitor (<http://www.colinst.com/>) and TSE Systems Drinking and Feeding Monitoring System (<http://www.tse-systems.com/>). These systems generally allow for an unlimited number of measurements to be monitored by a computer.

Measure body mass index (BMI)

8. Determine the weight of each mouse as described in steps 1 and 2.
9. Place the mouse in an inhalation anesthesia unit and anesthetize using 3.5% to 4.5% (v/v) isoflurane in O₂ for induction and 1.5% to 3.0% (v/v) isoflurane in O₂ for maintenance (Flecknell, 1993).
10. Remove the mouse from the anesthesia unit and place it in a ventral position. Gently push the back of the mouse flat and straighten the spine. Using a digital caliper, determine the nasal-to-anal length of the mouse in cm.
11. Calculate the body mass index (in g/cm²) by dividing the body weight by the square of the nasal-to-anal length.

Measure fecal lipid content

12. Place the mice in clean cages containing a metal floor grid instead of bedding.
13. Collect feces over a 24-hr period in parallel with a food intake measurement (see above) in order to determine fat balance (lipid intake and output). Dry the collected feces for 1 hr in a vacuum oven at 70°C, then clean the feces to remove remains of cage bedding and/or food pieces. Weigh total fecal output to an accuracy of 0.0001 g using an analytical balance.

Preferably, fat balance should be determined on several consecutive days to account for the variability in intestinal motility.

14. Carefully weigh a 100-mg aliquot of droppings into a 10-ml Erlenmeyer flask, and mix with 5 µl of 0.1 mCi/ml [carboxyl-¹⁴C]triolein (radioactive tracer).

To ensure that fecal aliquots are representative of the whole group, duplicate or triplicate samples should be tested.

15. Cool the sample and extract by adding 2 ml of 2:1 chloroform/methanol. Incubate 30 min at 60°C with constant agitation in a shaking water bath.
16. Pass the sample by gravity through a Whatman no. 1 filter paper into a 10-ml glass test tube. Add 2:1 chloroform/methanol up to a final volume of 4 ml.
17. Back-extract the sample by adding 1 ml water and mix well by vortexing. Induce phase separation by centrifuging 10 min at 2000 rpm in a clinical tabletop centrifuge, room temperature.

When equilibrium has been reached, there will be an upper transparent water/methanol phase, a lower chloroform phase, and an accumulation of material (“fluff”) at the interface.

18. Carefully draw off the bottom organic phase using a long-stem glass Pasteur pipet with its tip placed at the bottom of the tube, and transfer to a new tube. Evaporate the sample to dryness using a nitrogen evaporator (e.g., N-EVAP) with dry bath at 50°C. Resuspend the residue in 2 ml of 2:1 chloroform/methanol.
19. Transfer the solution into a preweighed glass scintillation vial and evaporate the solvent using a nitrogen evaporator (e.g., N-EVAP) with dry bath at 50°C. Dry the vial in a vacuum oven at 70°C for 1 hr, then carefully reweigh the vial.

20. Determine the difference in weight between the empty vial and the vial containing the dried lipid to give the amount of fecal lipid in mg.
21. To correct for loss of lipids during extraction, determine the percent recovery of radiolabeled triolein. Add scintillation cocktail to the vial and count the radioactivity in a β -counter. Adjust the weight of the fecal lipid in each sample according to the percent of triolein recovery.
22. Using the corrected amount of fecal lipid, express fecal lipid content as a percent of the weight of the starting fecal aliquot (100 mg).
23. Determine the total amount of fat excreted in the 24-hr period by multiplying the percent fecal lipid by the total fecal output (step 2).
24. Subtract the total fat excreted from the total amount of fat consumed in 24 hr (based on the fat content of the food as provided by the manufacturer) to give the amount of retained fat. Express as a percent of fat consumed.

**BASIC
PROTOCOL 2**

ASSESSMENT OF THERMOREGULATION BY THE COLD TEST

The cold test allows the evaluation of adaptive thermoregulation (Picard et al., 2002). In practice, mice are placed in individual cages without food but with free access to water. Initial body temperature is recorded by inserting a small thermoprobe into the rectum of the animal. Mice are then placed in a cold room at 4°C and rectal body temperature is recorded every hour during a 6-hr period. Results are presented as the curve of body temperature decrease over time. The authors do not normalize the cold test to body weight, but interpretation of the data should take into account whether there are differences in body weight.

Materials

Mice
 70% (v/v) isopropanol
 Mouse food
 Container to accommodate mouse during weighing
 Scale accurate to 0.1 g
 Standard mouse cages (metal grid not required)
 Thermometer with thermoprobe (accuracy <0.1°C) and rectal adapter (e.g., Bioseb; <http://www.bioseb.com>)

1. Determine the weight of the mouse as described in Basic Protocol 1, steps 1 and 2.
2. Place each test mouse individually in a clean cage, without food but with free access to water.
3. Disinfect the thermoprobe by washing it with a 70% isopropanol solution.
4. Measure the initial body temperature (T_0) of each mouse by inserting the thermoprobe into the rectum and recording the temperature on the probe once the value has stabilized (3 to 4 sec).
5. Place the individually caged mice in a cold room at 4° ± 1°C.
6. Measure and record the rectal temperature hourly, while in the cold room, over 6 hr.

The six measurements thus obtained are designated T_1 , T_2 , T_3 , T_4 , T_5 , and T_6 .

7. At the end of the experiment, place the mice into new cages in a room with normal ambient temperature (they can be regrouped if required), provide with food and water, and ensure that the animals exhibit normal behavior.

If the body temperature becomes too low (<25°C), place the cages close to a 37°C heating source for 15 to 30 min to warm the animals.

DETERMINATION OF BODY COMPOSITION BY DUAL-ENERGY X-RAY ABSORPTIOMETRY

**BASIC
PROTOCOL 3**

Bone densitometry and body composition (lean and fat content) are often measured using dual-energy X-ray absorptiometry (DEXA) analysis (Grier et al., 1996; Nagy and Clair, 2000). DEXA analysis can be performed using ultra-high-resolution densitometers such as those made by GE Medical Systems (e.g., PIXImus; <http://www.gehealthcare.com/inen/rad/bonedens/peripheral/piximus.html>) or Norland Medical Systems (pDEXA Sabre X-ray, <http://www.norland.com>) that allow for precise measurements of bone mineral and body composition from body imaging in less than 5 min in anesthetized mice.

CAUTION: Specific X-ray safety procedures should be followed.

Materials

- Anesthetic solution (see recipe)
- Mice
- PIXImus densitometer (0.18 × 0.18 mm; GE Medical Systems)
- Computer with Windows PIXImus software (e.g., version 1.4x; GE Medical Systems)
- Container to accommodate mouse during weighing
- Scale accurate to 0.1 g
- 1-ml syringe with 25-G, 0.5-mm needle
- Specimen tray for PIXImus (GE Medical Systems)

Prepare the PIXImus analyzer

1. Turn on the PIXImus analyzer at least 2 hr prior to testing.
2. Initiate the software by double-clicking the PIXImus icon.
3. Calibrate the PIXImus analyzer and perform quality control according to the operator's manual.

The field calibration procedure takes ~60 min. Each PIXImus has a "phantom," i.e., a structure with predetermined bone mineral density and fat content, which is analyzed and used for quality-control purposes.

Anesthetize the mouse

4. Determine the weight of the mouse as described in Basic Protocol 1, steps 1 and 2. Calculate and record the volume of anesthetic solution required for injection at 8 ml of anesthetic solution per kg.

When prepared as described and administered at 8 ml/kg, animals receive a dose of 200 mg/kg ketamine and 10 mg/kg xylazine.

5. Administer the appropriate volume of anesthetic solution by intraperitoneal injection using a 1-ml syringe and 25-G, 0.5-mm needle.

Perform the measurements

6. Remove the adhesive backing from the specimen tray, place the anesthetized mouse in a ventral position on the tray, and ensure that the vertebral column is straight. Wrap the mouse's tail around itself and extend the legs away from its body.

**Mouse
Phenotyping**

29B.1.7

7. Position the tray in the scanner so that the head is oriented towards the left, from the investigator's point of view.
8. Start the measurement according to the PIXImus operator's manual.

The image acquisition takes less than 5 min.

Perform the calculations and allow animals to recover

9. Following image acquisition, determine the bone mineral density (BMD, in g/cm²), bone mineral content (BMC, in mg), bone area (in cm²), and body composition (% fat and lean, and fat tissue in g) with the PIXImus software.

The specific body imaging area or the region of interest (ROI) can be manually selected within the total body image. For example, to image the spine and femur, the software positions an oval-shaped exclusion area to exclude the head and its soft tissue.

10. At the end of the experiment, remove the mouse from the plastic tray and return it to its home cage. Ensure that the test mice recover from the anesthetic and exhibit normal behavior.

Generally mice are allowed to recover on a 37°C heated surface.

**ALTERNATE
PROTOCOL**

**DETERMINATION OF BODY COMPOSITION BY QUANTITATIVE
NUCLEAR MAGNETIC RESONANCE**

As an alternative to DEXA (Basic Protocol 3), new EchoMRI quantitative magnetic resonance (QNMR) methods (e.g., Bruker Minispec NMR analyzer) can also be employed for the in vivo analysis of whole-body composition (lean and fat content) of conscious live mice. Both the DEXA and the EchoMRI methods eliminate the need for destructive chemical analysis and allow for multiple consecutive measurements throughout the lifetime of the animal.

Materials

Mice

Control/calibration sample (Bruker)

Bruker Minispec NMR analyzer (<http://www.minispec.com/mq/mice.html>), interfaced with personal computer running MS Windows 2000 or higher with Minispec software package installed

1. Turn on the NMR and then the computer.

Initially, three lights appear on the NMR indicating the preheating stage (red), the transition to normal heating (yellow), or the correct temperature (green). The NMR is ready to use ~3 hr after powering up.

2. Click the Minispec icon on the computer to load the Minispec software, then click the Daily Check icon to calibrate the machine.
3. Confirm that the control tube has been inserted into the NMR and click OK when prompted by the "Confirm minispec Daily Check sample has been inserted" message.
4. When calibration is finished, click Yes to print the results when prompted by the "minispec Daily Check OK: print out results" message.

When the calibration is not successful, an error message appears and action should be taken accordingly. Refer to troubleshooting section in the Minispec mcSeries User's Manual.

5. Weigh the mouse as described in Basic Protocol 1, steps 1 and 2. Remove the Control sample tube and insert the weighed animal into the Sample tube. Slowly insert the plunger, being careful not to drop the plunger onto the animal.

6. Insert the Sample tube, then click OK when prompted by the message “Waiting for inserting sample.”
7. Enter the animal’s identification number and the operator’s initials when the messages “Input sample ID” and “Input operator sample” appear. Click on Measure to initiate testing.

Testing takes ~3 min per animal.

8. When all testing is finished, remove the Sample tube, return the mouse to its home cage, and verify that it behaves normally. Repeat steps 5 to 8 until all animals have been tested.
9. Select the Save As option under the File menu, give the data a file name, check the “data format” radio button, and click Save. Shut down the Minispec program by selecting Exit from the File menu.

INDIRECT CALORIMETRY BY OXYMAX

Energy expenditure is evaluated through indirect calorimetry by measuring oxygen consumption with a specific apparatus such as the Oxymax (Columbus Instruments; <http://www.respirometer.com/microoxy.html>) or CaloSys (TSE Systems; <http://www.tse-systems.com/calorimetry/calorimetry.htm>; Porter, 2001). These systems monitor oxygen (O₂) and carbon dioxide (CO₂) gas volume concentrations at the inlet and outlet ports of a partially sealed chamber through which a known flow of ambient air is forcibly ventilated, either by positive pressure or by drawing the flow from the chamber. The concentration differences measured between the ports, along with the flow information, are used to compute oxygen consumption (VO₂) and carbon dioxide production (VCO₂), and the respiratory quotient (RQ) is calculated as VO₂/VCO₂. Chambers are equipped with a third port for temperature measurement.

In practice, mice are placed individually in the metabolic chambers while they have ad libitum access to food and water. Ideally, indirect calorimetry should be performed at thermoneutrality (~30°). Thermoneutrality is defined as the range of ambient temperature at which energy expenditure is lowest, no metabolic heat is required to maintain body temperature, and there is minimal oxygen consumption. At this temperature, the animal can maintain homeostasis without excessive use of energy for thermoregulation, and, hence, energy is available for maintaining optimum conditions of health and performance. When the heat load changes, the animal must expend energy to either eliminate excess heat or make heat in order to maintain internal thermal balance. Therefore, body temperature changes can affect various body functions (Gordon, 1993). If the temperature drops below 30°C, mice show an increase in metabolic rate, which can confound the outcome of the tests. The actual temperature for thermoneutrality can vary in rodents anywhere from 28° to 32°C. In general, higher temperatures should not be used. If indirect calorimetry cannot be performed at thermoneutrality, ambient temperature of (23° to 24°C) should be used. Indirect calorimetry is preferably performed for a 24-hr period while the mice have free access to food. O₂ and CO₂ measurements are taken at regular intervals during the whole experiment, and VO₂ and VCO₂ values are expressed as ml/kg/hr. As energy expenditure varies with body weight, it is important to normalize the values obtained for body mass.

The measurement of metabolic performance by calorimetry can be performed simultaneously with measurement of home cage activity and drinking and feeding behavior using modified systems—e.g., LabMaster (TSE Systems; <http://www.tse-systems.com/labmaster/labmaster.htm>) and Comprehensive Lab Animal Monitoring System (CLAMS; Columbus Instruments; <http://www.colinst.com/brief.php?id=61>). In

BASIC PROTOCOL 4

Mouse Phenotyping

29B.1.9

these types of systems, measurements of O₂ and CO₂ concentration, air flow, and temperature are recorded along with total activity (ambulatory and fine movements and rearing), food and liquid consumption, and urine and feces excretion. These systems, when available, can provide a large amount of data in a short period of time.

Materials

Mice

Container to accommodate mouse during weighing

Scale (capable of weighing 1 to 50 g to an accuracy of 0.1 g)

Respirometer: Oxymax system (Columbus Instruments;

<http://www.respirometer.com/microoxy.html>)

Computer with the Windows Oxymax software

Calibration gases: one tank with 20.5% O₂/0.8% CO₂ and a second tank with 100% N₂

1. Determine the weight of each mouse as described in Basic Protocol 1, steps 1 and 2.

As energy expenditure varies with body weight, values are normalized to body mass (see step 7).

2. In an experimentation room set to 30°C, place each mouse individually in an airtight cage (supplied with the Oxymax respirometer) with free access to food and water. Seal the cages tightly, taking care to ensure minimum leakage of gases.

3. Warm up the Oxymax respirometer for ≥2 hr, then calibrate the CO₂ and O₂ sensors according to the Oxymax operating manual prior to the start of the experiment.

The calibration of the gas sensors and flow meters is required to ensure precise measurements and involves using a gas mixture of 20.5% O₂ and 0.8% CO₂ (balance N₂) to adjust the O₂ and CO₂ sensors to reflect the contents of the calibration gas.

4. Following calibration, click the Experiment Configuration tab. Select the Chamber Configuration tab, and click on Chamber One. Enter the mouse information, including ID, weight, and sex, and then click Apply. Continue this procedure until each chamber is entered, then save the data file.

5. Ensure that the chambers are connected according to the Oxymax operating manual, and start the system by clicking Run.

Measurements of O₂ and CO₂, in units of ml/kg/hr, are acquired at regular intervals. The time interval is selected by the user, but is generally 3 min. The total duration of the recording is ~24 hr, which covers both active and resting phases.

6. At the end of the experiment, click Stop, shut off the air, and remove the mice from the chambers. Return to home cages and ensure that they exhibit normal behavior.

Chambers should be carefully washed and disinfected between experiments.

7. Normalize the values of oxygen consumption (VO₂) and carbon dioxide production (VCO₂) to body mass (ml/kg^{0.75}/hr or ml/kg^{0.67}/hr; see Critical Parameters).

Normalization is done automatically by the instrument software. In some cases (obese animals) values can be normalized to lean body mass as measured by DEXA analysis.

The respiratory exchange ratio (RER), also known as the respiratory quotient (RQ), is calculated as VO₂/VCO₂. Heat is calculated from the VO₂ and RER information using the standard formula: Heat (H) = calorific value (CV) × VO₂ × 0.001, where CV = 3.815 + 1.232 × RER. RER and heat are calculated automatically by the instrument software for each time interval. In this way, the differences in resting and active periods can be seen.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Anesthetic solution

- 1.4 ml ketamine
 - 0.35 ml 2% (w/v) xylazine
 - 3.85 ml 0.9% (w/v) NaCl (sterile)
- Prepare fresh

COMMENTARY

Background Information

The evolution of body weight gain is a critical first-line phenotyping test. Body mass and composition reflect the combined effects of three processes: energy intake, energy partitioning (storage), and energy expenditure. Body-weight evolution, to a large extent reflective of fat stores, represents the net balance between energy intake and energy expenditure. When a loss or gain of weight is evident, or when body composition is altered, a mismatch between energy intake and expenditure should be excluded. The phenotyping protocols outlined here provide a starting point for evaluating all sides of the equation and can provide a direction for subsequent molecular analysis.

Energy intake. The energy required for maintaining body function both at rest and during stress and exercise is provided in the form of carbohydrates, fat, and protein nutrients that are consumed daily (Spiegelman and Flier, 2001). In general, energy intake and energy expenditure are closely matched, such that body weight is remarkably stable even in the face of large fluctuations in caloric intake. The tight matching of energy intake and energy expenditure is governed by the central nervous system, primarily in the hypothalamus (Coppari et al., 2005). Leptin and insulin are primary effectors that activate receptors located on secondary effector neurons in the hypothalamus; these neurons express neuropeptides with long-term effects on food intake and body weight. Orexogenic neuropeptides that increase food intake, such as neuropeptide Y (NPY), are inhibited by leptin and insulin, while anorexogenic neuropeptides, such as proopiomelanocortin (POMC), are up-regulated. These neuropeptides subsequently act on local receptor populations to exert their physiological effects. As circulating insulin and leptin levels are proportional to the amount of body fat, they are critical sensors of long-term energy stores (adipose tissue).

In parallel with these more long-term regulatory signals emanating from the body's energy stores, a number of short-lived satiety factors originate within the gastrointestinal tract. For example, cholecystokinin, glucagon-like peptide, and ghrelin modulate short-term food intake by signaling to the CNS through vagal or sympathetic afferents, as well as through the circulation. Determining changes in food intake and food (lipid) absorption by measuring lipids excreted in the feces is the starting point for the evaluation of energy homeostasis.

Energy storage. The main organ used to store energy ingested in excess of immediate need is adipose tissue (Havel, 2002). In general, the energy density of adipose tissue (containing mainly fat) is >10-fold higher than that of muscle (which is principally composed of protein) or the liver (containing glycogen). The hypothalamic centers also monitor energy stores. For example, serum leptin increases with adipose tissue mass, thus informing the brain of the amount of energy stored. The role of leptin in signaling the brain about chronic changes in energy status is complemented by that of insulin, which conveys additional information about long-term changes in peripheral metabolism to the brain. The subsequent manipulation of energy stores is governed by a wide variety of enzymes, transcription factors, cytokines, and receptors. For example, changes in PPAR γ activity can alter adipocyte differentiation and, ultimately, fat stores. Techniques such as DEXA and QNMR are phenotyping tests that distinguish body composition according to lean and fat, and which can thus identify defects in the distribution or partitioning of stored/chemical energy when energy balance is challenged.

Energy expenditure. The third part of the energy balance equation is energy expenditure (Lowell and Spiegelman, 2000). Energy expenditure consists of several components, roughly separated according to obligatory basal activity (for performance of cellular

and organ functions), physical activity, and adaptive thermogenesis. Whereas energy enters the organism as food, it exits as heat and/or work on the environment. Fuel is converted step-by-step into ATP, which is then used for biological work within the cell; one side product of this reaction is heat. Energy expenditure at rest changes markedly in response to environmental temperature. Part of the response is due to shivering, but another part results from increased adaptive thermogenesis. Adaptive thermogenesis is defined as heat production in response to environmental temperature or diet. It serves to protect the organism from cold exposure or to regulate energy balance after changes in diet. In rodents, brown adipose tissue is a major site of adaptive thermogenesis. A real-life example of an extended cold test such as that described in Basic Protocol 2 is provided by animals that hibernate, such as the bear. Hibernating animals maintain a steady body temperature even though the temperature of the environment around them changes constantly. This production of heat is called non-shivering thermogenesis. The brain detects alterations in environmental temperature and diet, and, through neural circuits, activates efferent pathways to control energy expenditure. The primary efferent pathway is believed to be the sympathetic nervous system, in addition to the hypothalamic-pituitary-thyroid axis. Regulation of energy expenditure is in part achieved by tightly coupling the reactions in energy metabolism. For example, for any given molecule of fuel, a fixed amount of NADH and FADH is generated, which translates into a fixed number of protons that are pumped across the mitochondrial matrix by the electron transport chain. These protons re-enter via ATP synthase and produce a fixed number of ATP molecules that can be used to deliver a fixed amount of biological work. Because the complete combustion of food is achieved at the expense of molecular oxygen, the heat generated in these exergonic reactions can be estimated by measuring oxygen consumption using indirect calorimetry (see Basic Protocol 4). When this test is performed at rest and at thermoneutrality, it reflects mainly obligatory basal energy expenditure and modest physical activity. The cold test (see Basic Protocol 2) is one measure of adaptive thermogenesis and can provide stress that may reveal a change in energy homeostasis not seen under the experimental conditions of calorimetry. When energy expenditure is altered, as indicated by

calorimetry or cold tests, there is either an “uncoupling” of one of the reactions of cellular metabolism or a change in the amount of biological work. For example, uncoupling protein 1 (UCP1) has been shown to increase energy expenditure by uncoupling the relationship between protons entering the mitochondrial matrix and the synthesis of ATP (Lowell and Spiegelman, 2000).

Critical Parameters

Housing conditions and food intake. To determine the amount of food consumed per mouse, the total amount of food consumed per cage is divided by the number of animals per cage. Group housing can be criticized in view of issues of animal dominance as they affect food consumption. However, the authors have recently demonstrated that mice housed individually are severely stressed, which in turn significantly affects numerous parameters including blood insulin and glucose levels (Champy et al., 2004). In this instance, there is a “no win” situation regarding minimizing animal stress and maximizing the accuracy of the food intake measurement. The authors suggest that standard conditions always be used. For example, five mice cannot always be housed in one cage, especially when male littermates are being used, as males from different litters should not be housed together. Furthermore, unless the intention is to breed mice, females and males are never housed together. Thus, the number of cages and the number of mice per cage will likely change depending on the experimental cohort, and each investigator must make the best arrangement given the conditions (e.g., animal housing space, money, cohort size and distribution) available. Multiple tests can affect weight gain, so it is important to have weight values determined at times when minimal manipulations are being performed on the mouse cohorts. Comprehensive cage monitoring systems employing “metabolic cages” are an excellent alternative for automation of the measurement of food intake and various other metabolic parameters (including indirect calorimetry), thus minimizing overall animal stress and duration of the experiments. These metabolic cages can also be designed to study mouse movement and activity levels, as well as sleep patterns.

Body weight gain under calorie-dense diets. In order to accentuate differences in energy homeostasis, it is common practice to challenge mice with diets rich in caloric content. Such challenging conditions will

sometimes cause differences in body weight gain to become apparent. Unlike humans, mice in general do not like diets rich in fat and/or sucrose. This is because these diets often have a different consistency and taste from what they generally prefer. Mice usually like their food to be hard, as is the case with chow-like compositions; however, high-fat diets are generally soft. In the authors' experience, the total volume/mass of food consumed decreases when mice are switched from chow to a high-fat diet. However, despite the reduced mass of food consumed, total caloric intake increases due to the caloric density. In some instances, mice refuse to eat the new diet; to avoid this, the authors often fast mice overnight before switching to a different diet.

Body mass index. It is difficult to calculate the BMI because determining the length of a mouse is not simple. In many cases the differences will be very subtle, on the order of millimeters. In addition to measuring length by using a caliper, measurement can be obtained in parallel with DEXA and/or X-ray analysis, which require the mouse to be anaesthetized and imaged.

Fecal lipid content. The protocol described here is a rapid method for the determination of fat in feces. As the sensitivity of this assay depends on the ability to detect rather small differences in weight, care must be taken. In general, the percent recovery of radiolabeled triolein is 85% to 90%. If the recovery rate is not sufficient, an additional chloroform/methanol solvent extraction can be included and the organic phases combined.

Cold test. In general, this is a relatively simple test. Occasionally, when inserting the thermoprobe into the mouse, resistance is met due to the feces in the colon. In this instance, the lower belly of the mouse can be depressed somewhat to help with the fecal movement and excretion. As this is a fairly stressful test for mice, their health should be monitored throughout; if the body temperature of the mouse goes below 25°C, it should be removed from the cold and marked as "<25°C" for the remainder of the test. The number of mice used for the test depends on how quickly the measurements can be taken with the rectal probe and how long the investigator can stand the cold.

DEXA/QNMR. These tests can easily accommodate a large number of animals. In the case of DEXA, mice should be anesthetized in a staggered fashion, and care should be taken to ensure appropriate recovery. The actual data-acquisition steps for DEXA require

<5 min. For QNMR, mice do not need to be anesthetized, and data acquisition is faster (<2 min). In view of this advantage and the slightly better accuracy, it is expected that QNMR will rapidly become the test of choice for determining body composition (Taicher et al., 2003).

Oxymax analysis. Based on allometric scaling methods, metabolic body mass was estimated as the body weight to the power of 0.75 or 0.67 (ml/kg^{0.75}/hr or ml/kg^{0.67}/hr; Porter, 2001). Ideally, however, VO₂ should be normalized to lean body mass and not total body mass. For groups of animals with roughly the same body composition and fat mass, normalization is not required. However, in obese animals, adipose tissue consists mainly of "metabolically inactive" storage tissue, and even normalization to body weight using the power of 0.75 or 0.67 gives an incorrect estimate of metabolic body mass. In such cases, lean body mass (which reflects "metabolically active tissue") is best obtained from the analysis of body composition as determined by densitometry or an equivalent method. Taking all this into account, the authors prefer to perform indirect calorimetry in young animals, at an age when body weight differences are usually less pronounced.

In order to account for the time the animal requires to habituate to the setting of the Oxymax and overcome stress, the first hour of measurement is generally excluded from the final analysis. In some instances, indirect calorimetry measures are performed over 72 hr; however, the stress associated with continued housing of the mouse alone and on metal floor grids must also be considered. Another consideration is the number of chambers in the indirect calorimetry system, which can vary from 8 to 60 chambers, thus greatly affecting the number of mice measured at one time.

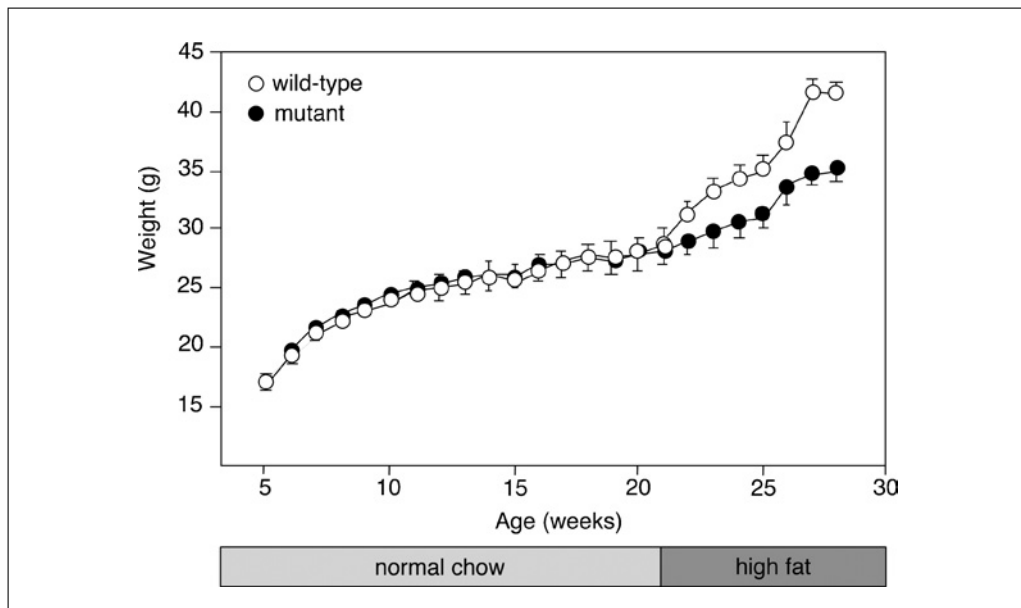
Anticipated Results

A summary of all calculations pertaining to the specific phenotyping tests is presented in Table 29B.1.1.

Body weight and food and water intake. As mice age, they will gain weight (Fig. 29B.1.2). The question is whether the mutant or treated mice gain more or less as compared to the wild-type control mice. Data are generally presented as a growth curve showing the average weight of the mouse over time. In some cases, data are expressed as percent loss or gain in weight as compared to weight at the beginning of the experiment. The length of time for the experiment is generally 6 to 12 weeks. Since

Table 29B.1.1 A Summary of the Measures Used in the Evaluation of Energy Homeostasis

	Measure	Formula (units)
Body	Body weight	(in g)
	Length	(in cm)
	Body mass index (BMI)	body weight/length (in g/cm ²)
Food intake	Food consumed	[food given – food left over (in g)]/no. mice per cage
	Energy consumed	kcal/g of diet × g food consumed (in kcal)
Fat absorption	Fecal lipid content	Retained fat = total fat consumed – excreted fat (in g)
Body composition	Fat	(in g)
	Lean	(in g)
	% Fat	[Fat (in g)/total tissue mass (in g)] × 100
	Bone mineral density (BMD)	(in g/cm ²)
	Bone mineral content (BMC)	(in g)
	Area	(in cm ²)
Cold test	Body temperature	(in °C at T ₁ , T ₂ , T ₃ , T ₄ , T ₅ , T ₆ expressed in hr)
Indirect calorimetry	VO ₂	(in ml/kg ^{0.75} /hr or ml/kg ^{0.67} /hr)
	VCO ₂	(in ml/kg ^{0.75} /hr or ml/kg ^{0.67} /hr)
	Respiratory exchange rate (RER)	VO ₂ /VCO ₂

**Figure 29B.1.2** An example of body-weight evolution in mice. A consecutive experimental strategy was employed whereby adult male wild-type and mutant mice were tested first under chow-fed conditions and then with a high-fat diet. Compared to wild-type mice, mutant mice are resistant to diet-induced obesity.

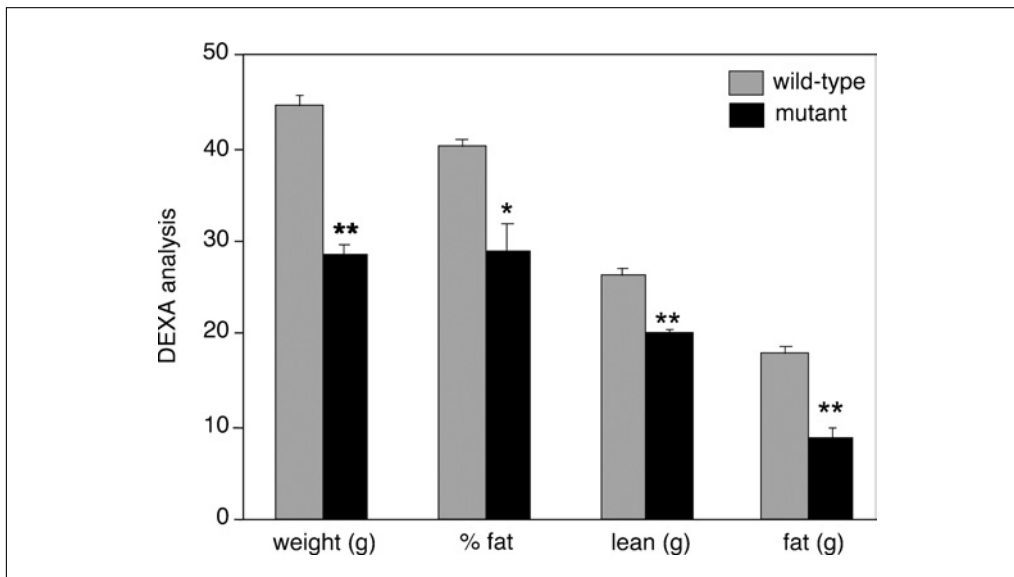


Figure 29B.1.3 An example of data obtained using DEXA analysis. Adult male wild-type and mutant mice were tested under high-fat diet conditions.

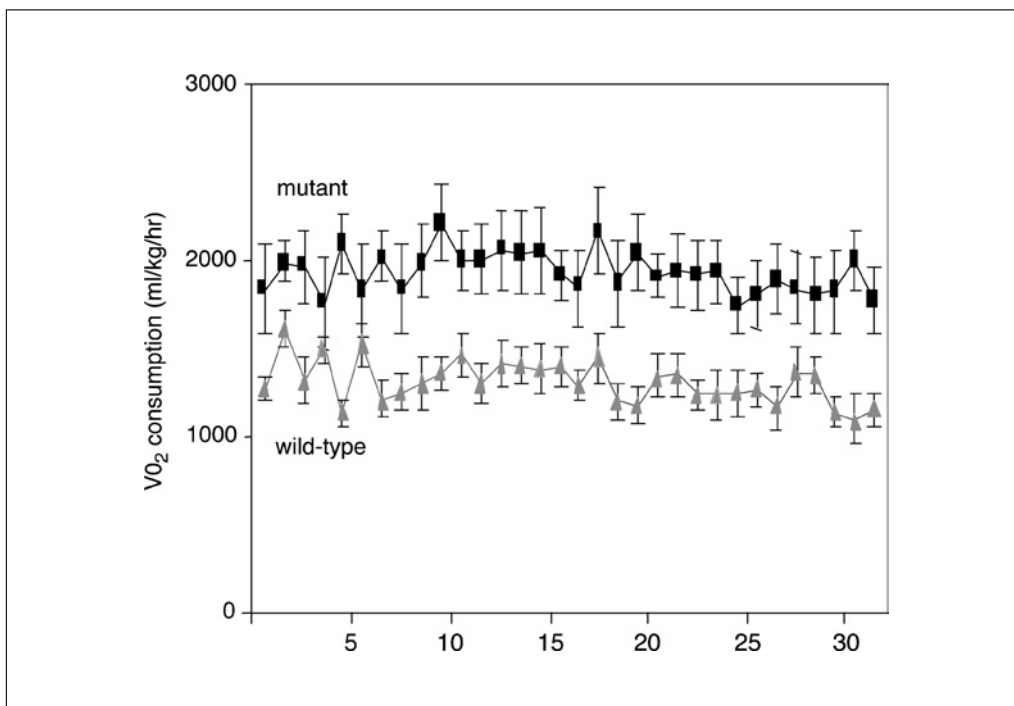


Figure 29B.1.4 An example of data obtained by Oxymax. Adult male wild-type and mutant mice were tested under high-fat diet conditions.

metabolic tests can be stressful and affect the weight, it is preferable to use weight data when the animals are undergoing a minimal number of manipulations. As stated above, calorie-rich diets can increase weight gain.

Fecal lipid content. Values of fecal lipid content in a mouse can vary greatly—anywhere from 20% to 50% of dry weight—and depend on numerous factors such as the type of diet, the genetic abnormality, and the age of the animal. When fecal lipid data are

combined with food-intake experiments, altered patterns of food consumption versus absorption can be distinguished.

DEXA and QNMR analysis. When changes in body weight are detected, parallel changes in body composition should also be found (Fig. 29B.1.3).

Indirect calorimetry. The balance of energy must be accounted for. Therefore, when there are changes in body weight, changes should either be detected in caloric intake or

loss in feces, and/or in energy expenditure (Fig. 29B.1.4).

Time Considerations

Performing *in vivo* mouse experiments is not a trivial process. Considerable planning is required to acquire ethics approvals, not to mention the time needed to generate the experimental mouse cohorts. Compared with this, the time needed to perform the actual phenotyping experiments is usually limited, especially when the protocols have already been standardized and validated. However, it is important to design and adhere to a relevant timeline of experiments, such as that proposed in Figure 29B.1.1, in order to minimize animal stress. The use of a pilot feeding study whereby mouse weight is simply monitored over several weeks is useful in order to indicate whether subsequent full metabolic phenotyping is a priority in the experimental cohort under investigation.

Literature Cited

- Champy, M., Selloum, M., Piard, L., Zietler, V., Caradec, C., Chambon, P., and Auwerx, J. 2004. Mouse functional genomics requires standardization of mouse handling and housing conditions. *Mamm. Genome* 15:768-783.
- Chen, A. and Innis, S. 2004. Assessment of phospholipid malabsorption by quantification of fecal phospholipid. *J. Pediatr. Gastroenterol. Nutr.* 39:85-91.
- Cock, T.A., Back, J., Eleftheriou, F., Karsenty, G., Kastner, P., Chan, S., and Auwerx, J. 2004. Enhanced bone formation in lipodystrophic PPAR γ (hyp/hyp) mice relocates haematopoiesis to the spleen. *EMBO Rep.* 5:1007-1012.
- Coppari, R., Ichinose, M., Lee, C.E., Pullen, A.E., Kenny, C.D., McGovern, R.A., Tang, V., Liu, S.M., Ludwig, T., Chua, S.C. Jr, Lowell, B.B., and Elmquist, J.K. 2005. The hypothalamic arcuate nucleus: A key site for mediating leptin's effects on glucose homeostasis and locomotor activity. *Cell. Metab.* 1:63-72.
- Flecknell, P.A. 1993. Anesthesia and perioperative care. *Methods Enzymol.* 225:16-33.
- Folch, J., Ascoli, I., Lees, M., Meath, J.A., and LeBaron, N. 1951. Preparation of lipid extracts from brain tissue. *J. Biol. Chem.* 191:833-841.
- Gordon, C.J. 1993. *Temperature Regulation in Laboratory Rodents.* Cambridge University Press, Cambridge, U.K.
- Grier, S., Turner, A., and Alvis, M. 1996. The use of dual-energy X-ray adsorptiometry in animals. *Invest. Radiol.* 31:50-62.
- Havel, P.J. 2002. Control of energy homeostasis and insulin action by adipocyte hormones: Leptin, acylation stimulating protein, and adiponectin. *Curr. Opin. Lipidol.* 13:51-59.
- Hrabe de Angeles, M., Flaswinkel, H., Fuchs, H., Rathkolb, B., Soewarto, D., Marschall, S., Heffner, S., Pargent, W., Wuensch, K., Jung, M., Reis, A., Richter, T., Alessandrini, F., Jakob, T., Fuchs, E., Kolb, H., Kremmer, E., Schaeble, K., Rollinski, B., Roscher, A., Peters, C., Meitinger, T., Strom, T., Steckler, T., Holsboer, F., Klopstock, T., Gekeler, F., Schindewolf, C., Jung, T., Avraham, K., Behrendt, H., Ring, J., Zimmer, A., Schughart, K., Pfeffer, K., Wolf, E., and Balling, R. 2000. Genome-wide, large-scale production of mutant mice by ENU mutagenesis. *Nat. Genet.* 25:444-447.
- Ishibashi, S., Schwarz, M., Frykman, P.K., Herz, J., and Russell, D.W. 1996. Disruption of cholesterol 7 α -hydroxylase gene in mice. I. Postnatal lethality reversed by bile acid and vitamin supplementation. *J. Biol. Chem.* 271:18017-18023.
- Koutnikova, H., Cock, T.A., Watanabe, M., Houten, S.M., Champy, M.F., Dierich, A., and Auwerx, J. 2003. Compensation by the muscle limits the metabolic consequences of lipodystrophy in PPAR gamma hypomorphic mice. *Proc. Natl. Acad. Sci. U.S.A.* 100:14457-14462.
- Lowell, B.B. and Spiegelman, B.M. 2000. Towards a molecular understanding of adaptive thermogenesis. *Nature* 404:652-660.
- Nagy, T. and Clair, A.L. 2000. Precision and accuracy of dual-energy X-ray absorptiometry for determining *in vivo* body composition of mice. *Obes. Res.* 8:392-398.
- Nolan, P.M., Peters, J., Strivens, M., Rogers, D., Hagan, J., Spurr, N., Gray, I.C., Vizor, L., Brooker, D., Whitehill, E., Washbourne, R., Hough, T., Greenaway, S., Hewitt, M., Liu, X., McCormack, S., Pickford, K., Selley, R., Wells, C., Tymowska-Lalanne, Z., Roby, P., Glenister, P., Thornton, C., Thuang, C., Stevenson, J.A., Arkell, R., Mburu, P., Hardisty, R., Kiernan, A., Erven, A., Steel, K.P., Voegeling, S., Guenet, J.L., Nickols, C., Sadri, R., Nasse, M., Isaacs, A., Davies, K., Browne, M., Fisher, E.M., Martin, J., Rastan, S., Brown, S.D., and Hunter, J. 2000. A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. *Nat. Genet.* 25:440-443.
- Picard, F., Annicotte, J., Rocchi, S., Champy, M.F., O'Malley, B.W., Chambon, P., and Auwerx, J. 2002. SRC1 and TIF2 control energy balance between white and brown adipose tissues. *Cell* 111:931-941.
- Porter, R.K. 2001. Allometry of mammalian cellular oxygen consumption. *Cell. Mol. Life Sci.* 58:815-822.
- Schwarz, M., Russell, D.W., Dietschy, J.M., and Turley, S.D. 2001. Alternate pathways of bile acid synthesis in the cholesterol 7 α -hydroxylase knockout mouse are not upregulated by either cholesterol or cholestyramine feeding. *J. Lipid Res.* 42:1594-1603.
- Spiegelman, B.M. and Flier, J.S. 2001. Obesity and the regulation of energy balance. *Cell* 104:531-543.
- Taicher, G.Z., Tinsley, F.C., Reiderman, A., and Heiman, M.L. 2003. Quantitative magnetic resonance (QMR) method for bone and whole-body-composition analysis. *Anal. Bioanal. Chem.* 377:990-1002.

Contributed by Carmen A. Armann
Institut de Génétique et de Biologie
Moléculaire et Cellulaire
Illkirch, France

Marie-France Champy
Institut Clinique de la Souris
Illkirch, France

Johan Auwerx
Institut de Génétique et de Biologie
Moléculaire et Cellulaire and Institut
Clinique de la Souris
Illkirch, France

Lipid and Bile Acid Analysis

Clinical chemistry screening to identify changes in concentrations of serum lipids, including cholesterol, free fatty acids (FFA), and triglycerides (TG), is an important part of first-line metabolic phenotyping. Whereas cholesterol is an important body constituent necessary for membrane function and synthesis of hormones and bile acids, TG and FFA are important energy reservoirs. As a consequence of the pleiotropic functions of these lipids, many diseases have been associated with altered lipid levels. For example, elevated low-density lipoprotein (LDL) cholesterol levels increase the risk of atherosclerosis, a pathology characterized by lipid accumulation in the arterial wall (Grundy, 2004). When changes in serum lipids are observed, more detailed phenotyping screens, including lipoprotein profiling and tissue lipid content, can facilitate the identification of the underlying mechanism associated with the dyslipidemia, especially when combined with other phenotypic observations. Since bile acids are the most important pathway for eliminating cholesterol from the body, analysis of serum bile acid levels as well as bile acid pool size and composition is also valuable in a comprehensive lipid phenotyping screen. In addition, serum bile acid levels are an early and sensitive indicator for liver dysfunction. In this unit, methods associated with serum and tissue lipid analysis are outlined, as are protocols for assessing bile acids.

Basic Protocol 1 describes kit-based colorimetric procedures for determining TG, total cholesterol, LDL cholesterol, HDL cholesterol, and FFA. Alternate Protocol 1 describes similar lipid analyses using an automated laboratory workstation. Basic Protocol 2 describes separation of the different lipoprotein fractions by differential ultracentrifugation, while Alternate Protocol 2 describes a comparable fractionation procedure employing fast protein liquid chromatography (FPLC). Both of these procedures yield preparations that can be quantitated via Basic Protocol 1. Basic Protocol 3 describes extraction of lipids from tissue for analysis via Basic Protocol 1. Basic Protocol 4 and Alternate Protocol 3 describe two different methods for measurement of serum bile acids, while Alternate Protocol 4 provides a method for analyzing bile acid pool size and composition. Finally, a Support Protocol is provided for lipoprotein quantitation.

STRATEGIC PLANNING

Critical considerations with respect to the general planning of *in vivo* experiments and minimizing variation due to genotype and phenotype are discussed in *UNIT 29A.2*. Additional considerations in the context of this unit include those associated with blood drawing, such as the method and timing of blood collection and the length of fasting. As mice are small animals, their blood supply is correspondingly small. In general, the blood-draw technique chosen should give the required quantity of blood while minimizing invasiveness (see tail vein method in *UNIT 29A.3*). Although there is no hard and fast rule as to the maximum amount of blood that can be collected at any single point without endangering the mouse, anything over 10% to 15% of total blood volume or 1% of the body weight of the animal can induce hypovolemic shock and/or death. As blood lipid profiles reflect both diet-derived sources and *de novo* synthesis, a critical parameter to consider is the length of fasting prior to blood draw. For general analysis, a 4-hr fast at the end of the dark cycle is sufficient, as this allows for diet-derived lipids to be processed. A 4-hr fast will also limit responses to starvation, such as FFA release. Longer (overnight) and shorter fasts are also often used; however, data must be considered in the context of the concomitant metabolic consequences. Feeding, for instance, is associated with gallbladder contraction; therefore, fasting levels of plasma bile acids tend to be lower than those in the fed state. These considerations are critical anytime comparisons are to be made.

Contributed by Carmen A. Argmann, Sander M. Houten, Marie-France Champy, and Johan Auwerx

Current Protocols in Molecular Biology (2006) 29B.2.1-29B.2.24
Copyright © 2006 by John Wiley & Sons, Inc.

Mouse
Phenotyping

29B.2.1

Supplement 75

MANUAL MEASUREMENT OF SERUM LIPIDS

Measurements of triglyceride (TG), total cholesterol, free fatty acids (FFA), and low- and high-density lipoprotein (LDL and HDL) cholesterol can be performed manually or with an automated laboratory workstation, such as the Olympus AU400 or the Cobas Integra (Roche Diagnostics), using specific colorimetric or enzymatic methods. In general, 80 to 150 μ l of plasma is sufficient for all measurements, with the lowest sample quantities being required for the automated protocols. Many of the methods are provided as kits, such that all reagents, standards, controls, and protocols are provided by the manufacturer in a standardized and user-friendly format. Several different suppliers for these kits exist, e.g., Olympus, Cobas (Roche Diagnostics), Wako, and Sigma.

The approaches are generally based on similar principles, with differences mainly in the choice of chromophore and color of the spectra. For example, cholesterol can be measured using an enzymatic color test whereby cholesterol esters (CE) are hydrolyzed by cholesterol esterase to free cholesterol, which is then oxidized by cholesterol oxidase (CHO) to cholestene-3-one. This reaction simultaneously produces hydrogen peroxide, which then is coupled with 4-aminoantipyrine and phenol in the presence of peroxidase (POD) to yield a chromophore that is measured spectrophotometrically. TGs can be measured by a series of coupled enzymatic reactions beginning with hydrolysis by lipases to give glycerol and FFA. The glycerol is then phosphorylated by glycerol kinase to produce glycerol-3 phosphate, which is oxidized by molecular oxygen in the presence of glycerol phosphate oxidase (GPO) to produce hydrogen peroxide and dihydroxyacetone phosphate. The hydrogen peroxide is then used to oxidatively couple *p*-chlorophenol and 4-aminoantipyrin catalysis by POD to give a red dye that is measured spectrophotometrically. FFA are measured using an enzymatic color test whereby acyl CoA synthase is used to convert fatty acids into acyl CoA. In a second reaction, acyl CoA oxidase catalyzes the formation of 2,3-trans-enoyl CoA and hydrogen peroxide, the later product being subsequently reacted with 4-aminophenzone and MEHA and catalyzed by POD to give a red pigment whose intensity is proportional to the concentration of FFA in the sample. HDL cholesterol can be measured using an enzyme chromogen system whereby an anti-human β -lipoprotein antibody is used to bind all non-HDL lipoproteins or apoB-containing lipoproteins (chylomicrons, VLDL, IDL, and LDL). In the first reaction, the antigen-antibody complexes formed with the apoB-containing lipoproteins block all subsequent enzyme reactions, allowing HDL cholesterol to be specifically measured in a second reaction by the CHO/POD system. LDL cholesterol can be measured following a two-step reaction in which the first step uses cholesterol esterase and cholesterol oxidase to break down all non-LDL lipoproteins in the presence of an LDL-cholesterol protecting reagent. In the second step, the protecting reagent is released from LDL, and LDL cholesterol is quantified by the CHO/POD system.

Materials

- Blood serum, plasma, or isolated lipid fraction obtained via other protocols in this unit
- Lipid analysis kits (Wako):
 - Total cholesterol kit
 - LDL cholesterol kit
 - HDL cholesterol kit
 - FFA kit
 - Free cholesterol kit
 - TG kit
- 5-ml plastic or glass tubes
- Spectrophotometer

Table 29B.2.1 Specifications for Lipid Measurements Using Wako Enzymatic Reaction Kits

Test	Sample/standard/calibrator ^a	Reagent 1	Reagent 2	Incubation time/temperature	Absorbance (nm)
Total cholesterol	20 μ l	2 ml	0	5 min, 37°C	600
Free cholesterol	50 μ l	3 ml	0	5 min, 37°C	600
Triglyceride	3 μ l	260 μ l	150 μ l	5 min, 37°C per reaction	600/700
Free fatty acids	50 μ l	1 ml	2 ml	10 min, 37°C per reaction	550
LDL cholesterol	3 μ l	270 μ l	90 μ l	5 min, 37°C per reaction	600/700
HDL cholesterol	3 μ l	270 μ l	90 μ l	5 min, 37°C per reaction	600/700

^aReaction volumes can be scaled to meet the sensitivity and/or volume capacity of spectrophotometer (e.g., 96-well microtiter plate reader).

1. Reconstitute the lipid analysis kit reagents according to the manufacturer's instructions.

Color reagents are generally stable for ~1 month following reconstitution, when stored at 2° to 10°C.

2. Prepare standards by serially diluting the stock standard solution provided at 1:1, 1:2, 1:4, and 1:8.

This should give final concentrations that cover both the lower and upper range of the linear curve. According to many manufacturers, the use of a single standard per reaction is generally sufficient after a calibration/standard curve has been established once to ensure linearity.

3. Pipet sample and standards into the appropriate tubes using the volumes indicated in Table 29B.2.1. Include a tube for a blank (deionized water).

4. Add the required amount of reagent #1 per tube and vortex. Incubate tubes for the indicated amount of time (Table 29B.2.1).

Skip step 5 and proceed to step 6 if the kit involves only a single reagent.

5. Add the required amount of reagent #2 per tube and vortex. Incubate tubes for the indicated amount of time (Table 29B.2.1).

6. Measure the absorbance of each sample (A_s) and standard (A_{std}) against the blank at the suggested wavelength(s) A_1 and A_2 (Table 29B.2.1).

7. Prepare a standard curve by plotting absorbance versus concentration for the standard. Generate an equation of the line using linear regression.

8. Calculate amount of lipid for the sample (in mg/dl) as follows:

a. *From a standard curve:* Read the concentration corresponding to the absorbance of the sample (A_s) from the calibration curve or calculate using the equation of the line $y = mx + b$, where m = slope of the curve and b = the x intercept.

b. *From a single standard:* concentration of sample (in mg/dl) = $(A_s/A_{std}) \times$ concentration of standard (in mg/dl).

c. *For dual-wavelength reactions:* Substitute A_s for $\Delta A_s [(A_{2s} - A_{2blank}) - (A_{1s} - A_{1blank})]$ and A_{std} for $\Delta A_{std} [(A_{2std} - A_{2blank}) - (A_{1std} - A_{1blank})]$.

9. Check results against the manufacturer's specifications of the linear range for the reaction.

For each parameter, a linear range is defined by the kit manufacturer. When the concentration of the parameter is outside this linear range, the results cannot be calculated accurately. If samples exceed the upper limit, they should be diluted with saline or water and the assay should be repeated. The result is then multiplied by the dilution factor. The lowest detectable level has also been estimated for each parameter; it represents the lowest measurable level of the parameter that can be distinguished from zero. It is calculated as the absolute mean plus three standard deviations of twenty replicates of an analyte-free sample.

**ALTERNATE
PROTOCOL 1**

**BLOOD LIPID ANALYSIS USING AUTOMATED LABORATORY
WORKSTATION**

Automated workstations are used increasingly for routine clinical chemistry evaluations in the mouse. Such workstations increase throughput and often reduce the sample volume required for analysis, because many measurements can be combined. Lipid analyses have been adapted for such an automated, high-throughput approach. For a comparison of automated lipid tests, see Table 29B.2.2.

Automated, high-throughput workstations are often found in laboratories that perform many assays, such as large clinical chemistry laboratories in veterinary and mouse clinics. Such high-throughput laboratories can often accommodate samples from outside laboratories. It is worthwhile to identify several laboratories with automated workstations and compare the price and quality of the services provided. One important consideration when opting to outsource tests is to verify that the tests were validated for mouse samples. Although routine cholesterol, triglyceride, and fatty acid measurements are interchangeable between species, differences exist for HDL and LDL cholesterol measurements, which often use species-specific antibodies.

This protocol uses an Olympus AU400 automated workstation, which is at present among the workstations that require the lowest sample volumes—an important advantage when analyzing mouse samples. This system is capable of analyzing serum, plasma, and urine samples, and measures the components of the samples optically. It automates the main processes, including dispensing of sample and reagent, mixing, and photometry. First, the sample probe dispenses serum or plasma in a sample cup that is set on a rack into the specified cuvette of the cuvette wheel unit (88 cuvettes are mounted on the cuvette wheel). Then the specific reagent, which has been added to the reagent refrigeration unit, is added to the sample by the reagent probe. The sample probe and reagent probe units

Table 29B.2.2 Comparison of Automated Lipid Tests

Test	Test principle	Units	Sample volume	Manufacturer/model
Cholesterol	Enzymatic color test, CHOD-PAP method	mmol/liter	3 μ l	Olympus OSR 6116
HDL cholesterol	Immuno-inhibition, enzymatic color test	mmol/liter	2 μ l	Olympus OSR 6187
LDL cholesterol	Enzymatic color test, CHOP-PAP method	mmol/liter	3 μ l	Olympus OSR 6183
Triglyceride	Enzymatic color test, GPO-PAP method	mmol/liter	3 μ l	Olympus OSR 6133
Nonesterified fatty acids (NEFA)	Enzymatic color test, AcylCoA-synthetase + ATP	mmol/liter	8 μ l	Wako

**Lipid and Bile
Acid Analysis**

29B.2.4

are equipped with a wash well for washing the probe after each sample is dispensed. The sample and reagent are mixed in the cuvette by a mixing device, and the system measures the absorbance of the sample. The system is equipped with a photometer lamp to optically measure the mixture in a cuvette. Finally, the optical density of the sample is calculated according to the absorbance measured. The data processor unit employs a personal computer to process data.

Calibration frequency using the Olympus calibrators depends on the parameters to be measured and the workflow. For most parameters described here, calibration is performed once every 5 days or when the analyzer requests a new calibration, or when the control value is outside the acceptable range. Calibrators are typically used undiluted to generate a single reference value, although manufacturer's instructions need to be followed in each case.

Materials

Calibrators for biochemistry on AU400:

Olympus OE 66300 for cholesterol and triglyceride (lyophilized human serum with chemical additives and appropriate enzymes of human, animal, and plant origin)

Olympus ODC0011 for HDL cholesterol and ODC0012 for LDL cholesterol (lyophilized human serum containing human HDL or LDL cholesterol)

Wako 994-75409 for FFA

Intralaboratory controls for biochemistry on AU400:

ODC0003 (Level 1) and ODC0004 (Level 2) Olympus System Reagent Control Serum (lyophilized preparation of stabilized human serum with adjusted levels of defined analytes)

ODC0005 LDL/HDL Cholesterol Control Serum, Levels 1/2 (lyophilized preparation of stabilized human serum with adjusted levels of defined analytes)

Interlaboratory controls: Lyphochek unassayed chemistry control, human, levels 1 and 2 (Bio-Rad C-320-10 and C-325-10)

Serum or plasma samples (80 μ l minimum)

Olympus AU400 automated laboratory workstation and operating manual (<http://www.olympus-global.com>) with appropriate sample tubes and software

Unity Plus quality control software (Bio-Rad)

4-ml, 50 \times 14-mm hemolysis tubes (Sarstedt) with bar-coded labels added in-house

NOTE: For general operation of the analyzer, refer to the operator manual of the AU400.

1. Use the yellow rack for calibration and load the blue rack at position 1 with deionized water to determine the blank value.
2. Test all parameters daily with intralaboratory control sera (Olympus ODC0003, ODC0004, and ODC0005), which consist of lyophilized human plasma with normal and pathological lipid concentrations.
 - a. Thaw and vortex the control sera before use.
 - b. Load controls according to the analyzer's display.
 - c. Manage all control data using software provided with Olympus automated workstation.

Control values must lie within three standard deviations, otherwise the measurement will have to be repeated following recalibration.

3. Test all parameters daily with interlaboratory control sera (Bio-Rad level 1 and 2). Send data monthly to Bio-Rad using Bio-Rad Unity Plus software to provide

for an interlaboratory quality control assessment by the Bio-Rad External Quality Assurance Services.

This program provides a monthly report analysis.

4. Thaw the frozen unknown plasma samples, if needed. If samples are not available in sufficient quantity and/or are very lipemic, dilute 1:2 with deionized water.
5. Aliquot samples into tubes (provided by Olympus; alternatively, use microcentrifuge tubes with lids removed), then place into matching bar-coded hemolysis tubes (bar coding added in-house). Vortex all samples.
6. Load the racks according to the working lists, starting with rack #1. Perform the analyses.
7. Calculate values from the appropriate calibrations.

Each parameter has a limit value specified by the manufacturer. The reaction is valid only within the limits of linearity. Results are generally reported as mmol/liter.

BASIC PROTOCOL 2

LIPOPROTEIN FRACTIONATION BY DIFFERENTIAL ULTRACENTRIFUGATION

When abnormalities are detected in total, HDL, or LDL cholesterol levels, it is useful to isolate and analyze the different fractions of lipoproteins. Two major techniques are used for the separation of lipoprotein fractions, differential ultracentrifugation and chromatography (see Alternate Protocol 2). Differential ultracentrifugation is a preparative method for separating lipoproteins on the basis of their hydrated density. At solvent densities higher than the lipoprotein density, lipoproteins float at a rate dependent on their density, size, and shape. The lipoprotein classes are sequentially isolated by stepwise increases in the solvent density through the addition of salts such as KBr, NaBr, or NaCl (see Fig. 29B.2.1), over a broad range of densities unequally distributed along a gradient from 0.92 to 1.25 g/ml (De Lalla and Gofman, 1954). The classification of lipoprotein particles according to buoyant density (d) includes: chylomicrons ($d < 0.94$ g/ml); very-low-density lipoproteins (VLDL; $d = 0.94$ to 1.006 g/ml); intermediate-density lipoproteins (IDL; $d = 1.006$ to 1.019 g/ml); low-density lipoproteins (LDL; $d = 1.019$ to 1.063 g/ml); and high-density lipoproteins (HDL; $d = 1.063$ to 1.21 g/ml). Their equivalent flotation rates (Sf) by analytical ultracentrifugation, at a background density of 1.063 g/ml, are >400 for chylomicrons, 20 to 400 for VLDL, 12 to 20 for IDL, and 1 to 12 for LDL. At a background density of 1.2 g/ml, HDL₂ and HDL₃ display flotation rates of 3.5 to 9 and 0 to 3.5, respectively. Once lipoprotein fractions are isolated, the particle content of cholesterol, TG, and different apolipoproteins can then be measured.

Materials

Serum or plasma: preferably 400 μ l of freshly collected sample (samples collected under the same conditions can be pooled)

Density gradient solutions (see recipe)

Potassium bromide, solid

Dialysis buffer (see recipe)

1.0-ml polycarbonate ultracentrifuge tubes (Beckman)

Optima TL-100 ultracentrifuge with TLA-100.2 rotor (Beckman) or equivalent ultracentrifuge and rotor

Digital Density Meter DMA3.5 (Anton Paar; <http://www.anton-paar.com>)

Dialysis tubing (Spectra/Por membrane 3.5, Spectrum)

Beaker for dialysis

Additional reagents and equipment for dialysis (APPENDIX 3C), lipid assays (Basic Protocol 1), and protein quantitation (Support Protocol)

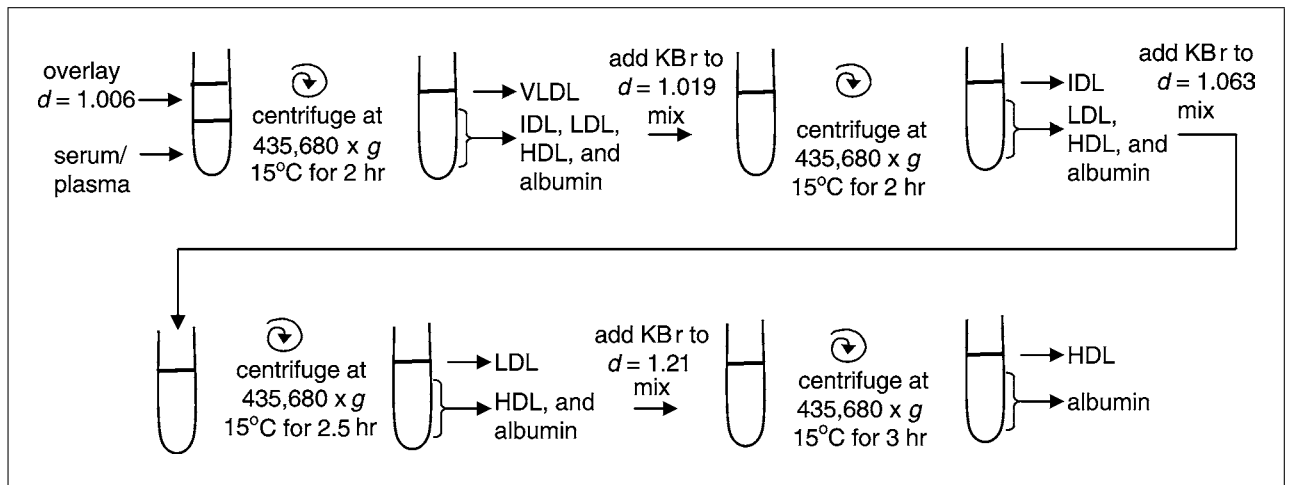


Figure 29B.2.1 Isolation of lipoprotein fractions by sequential ultracentrifugation.

Isolate VLDL ($d = 1.006$)

1. Aliquot 400 μl of serum into a 1-ml polycarbonate ultracentrifuge tube and overlay with 600 μl of the $d = 1.006$ density gradient solution.
2. Ultracentrifuge 2 hr at $435,680 \times g$, 15°C , without braking.

All centrifugations in this protocol should be done with a slow deceleration rate, as braking could disrupt the phase separation.

3. Allow the centrifuged sample to return to room temperature and transfer the top layer ($\sim 400 \mu\text{l}$), corresponding to VLDL, to a clean tube using a pipet. Set the bottom fraction aside on ice for continued fractionation.

In general the top layer is an off-white color and can be distinguished from the bottom fraction, which contains the rest of the lipoproteins.

IMPORTANT NOTE: *If the donor animals are not fasted, this top layer will contain chylomicrons as well. Chylomicrons are difficult to separate completely from VLDL; however, if the sample is centrifuged initially for 20 min, the supernatant should contain the larger and lighter particles resembling chylomicrons. The bottom layer is then processed again, as in steps 1 to 3, to isolate the VLDL.*

4. Prepare dialysis tubing by presoaking it in dialysis buffer (APPENDIX 3C). Tie a double knot at one end of the dialysis tubing, then add the VLDL fraction. Double-knot the other end close to the top of the solution to minimize the dilution of the sample. Place the dialysis tubing into a beaker and dialyze against 50 ml dialysis buffer for 24 hr at 4°C with constant mixing, changing the dialysis solution at least three times over this time interval.
5. Collect the dialyzed VLDL sample into a microcentrifuge tube. Set aside for lipid analysis.

Isolate IDL ($d = 1.019$)

6. Resuspend the 600 μl of bottom fraction in a polycarbonate tube and adjust the density from 1.006 to 1.019 by adding 24.3 μl of $d = 1.34$ density gradient solution.

This volume is determined according to the formula $(A \times Y) + (B \times X) = (A + B) \times Z$, where $A =$ volume of sample (600 μl), $B =$ volume of salt solution to add (in this case 24.3 μl), $Y =$ the initial density ($d = 1.006$), $Z =$ density of added salt solution ($d = 1.34$), and $X =$ the final density ($d = 1.019$).

7. Adjust the sample to a final volume of 1 ml by adding $d = 1.019$ density gradient solution, and mix thoroughly.

8. Ultracentrifuge the sample 2 hr at $435,680 \times g$, 15°C , without braking.
9. Warm the centrifuged sample to room temperature and transfer the top layer ($\sim 400 \mu\text{l}$), corresponding to IDL, to a clean tube using a pipet. Set the bottom fraction aside on ice for continued fractionation.
10. Dialyze as in steps 4 and 5.

Isolate LDL ($d = 1.063$)

11. Resuspend the bottom fraction ($\sim 400 \mu\text{l}$) in a polycarbonate tube, add $95.3 \mu\text{l}$ of $d = 1.34$ density gradient solution, and bring the volume to 1 ml with $d = 1.063$ density gradient solution. Mix, then ultracentrifuge 2.5 hr at $435,680 \times g$, 15°C .
12. Allow the sample to return to room temperature and transfer the top layer ($\sim 400 \mu\text{l}$), corresponding to LDL, to a clean tube using a pipet. Set the bottom fraction aside on ice for continued fractionation.
13. Dialyze as in steps 4 and 5.

Isolate HDL ($d = 1.21$)

14. Resuspend the bottom fraction ($600 \mu\text{l}$) in a microcentrifuge tube. Adjust the density to 1.21 by adding $678.5 \mu\text{l}$ of $d = 1.34$ solution, and add $21.5 \mu\text{l}$ of $d = 1.21$ density gradient solution to bring the sample to a final volume of 1.3 ml.
15. Mix the sample, then transfer 1 ml to a polycarbonate tube and ultracentrifuge 3 hr at $435,680 \times g$, 15°C .
16. Allow the sample to return to room temperature and transfer the top layer ($\sim 400 \mu\text{l}$), corresponding to HDL, to a clean tube using a pipet.

The bottom fraction corresponds to the lipoprotein-depleted plasma and can also be analyzed as a control if desired.

17. Dialyze as in steps 4 and 5.

Analyze fractions

18. Analyze dialyzed fractions for total cholesterol and TG content (see Basic Protocol 1) and/or protein content (see Support Protocol).

In general, TG are the predominant lipids in the VLDL fraction, but are very low in the LDL fraction.

To determine the concentration of the various analytes in the HDL fraction, a correction factor of 1.3 is necessary. Lipid content can be normalized to the volume of initial serum to give mg/dl, or to the protein content to give mg lipid/ μg protein.

**ALTERNATE
PROTOCOL 2**

LIPOPROTEIN ANALYSIS BY GEL-EXCLUSION CHROMATOGRAPHY

Separation of lipoprotein fractions can also be performed by gel filtration, which separates proteins on the basis of particle size (Brousseau et al., 1993). Briefly, the plasma samples are applied to a column and eluted at a constant flow rate using elution buffer. The eluted sample is then passed through an online UV detector prior to collection with a fraction collector. The fractionated samples are then subjected to manual or automated lipid measurements (e.g., cholesterol and TG) using enzymatic reagents as in Basic Protocol 1. A profile is then obtained that partitions the amount of lipid or protein according to the different lipoprotein fractions, which can be identified by their known retention times (VLDL = 33 min, LDL = 38 min, and HDL = 47 min).

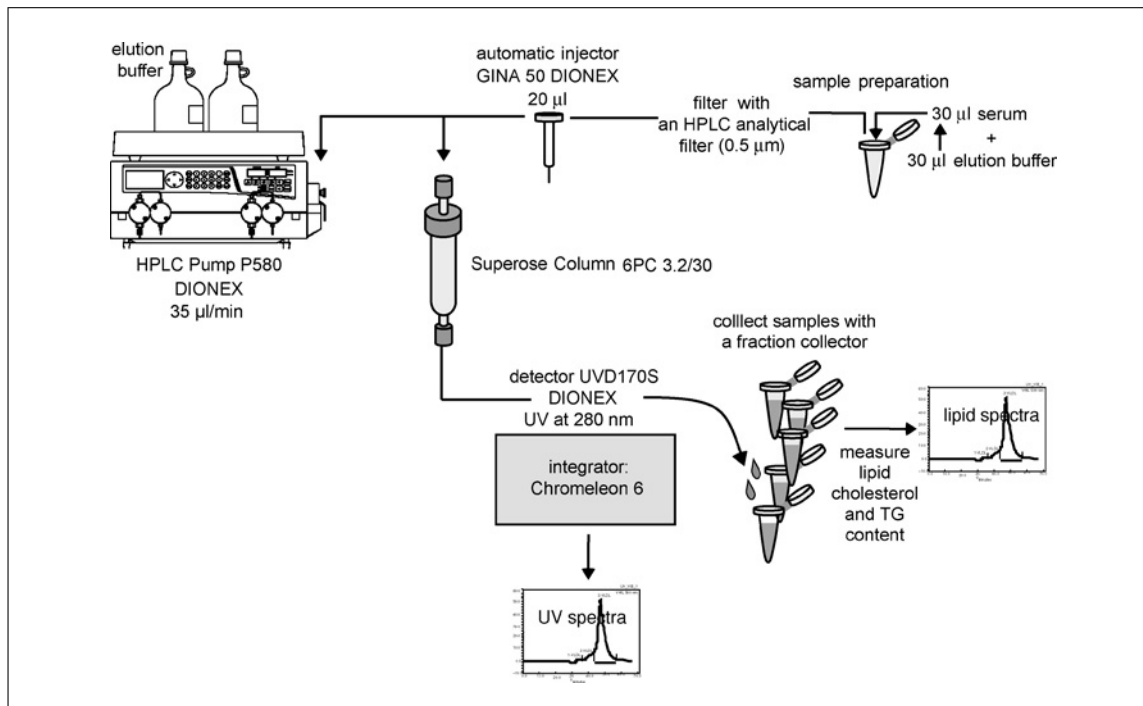


Figure 29B.2.2 FPLC schematic diagram.

Materials

Serum or plasma

Elution buffer (see recipe)

20% (v/v) ethanol

FPLC system (Fig. 29B.2.2) consisting of:

HPLC pump (P580 A, Dionex)

Multi-sample automatic injector (GINA 50, Dionex)

Column (Superose 6 PC 3.2/30, Amersham Biosciences)

UV/VIS detector (UV170S, Dionex)

Fraction collector (FRAC-100, Amersham Biosciences)

Integrator software (Chromeleon 6.4, Dionex)

0.22- or 0.45-µm HPLC analytical filter for Biocompatible UPCA

Additional reagents and equipment for lipid assays (Basic Protocol 1)

1. Set up an FPLC-type unit as shown in Figure 29B.2.2.
2. On day 1 of the experiment, install the Superose column according to the manufacturer's instructions. To equilibrate the column for first-time use or after long-term storage, flush out the storage solution (generally 20% ethanol) with 2.4 ml water at a flow rate of 20 µl/min, then equilibrate the column with elution buffer at a flow rate of 40 µl/min. Continue until the system is completely free of air bubbles and the UV signal is stable at the selected sensitivity (minimum of 6 hr, preferably overnight).
3. On day 2 of the experiment, prepare the samples by mixing 30 µl of plasma or serum (preferably plasma) with 30 µl of elution buffer. Filter the samples with either a 0.45- or 0.22-µm filter to remove aggregates.

The maximum protein capacity for the column is ~0.5 to 3 mg in a maximum concentration of 50 mg/ml. This range maximizes the resolution and detection of the protein.

4. Place the samples in the autoinjector and set the machine to inject 20 µl of sample onto the column.

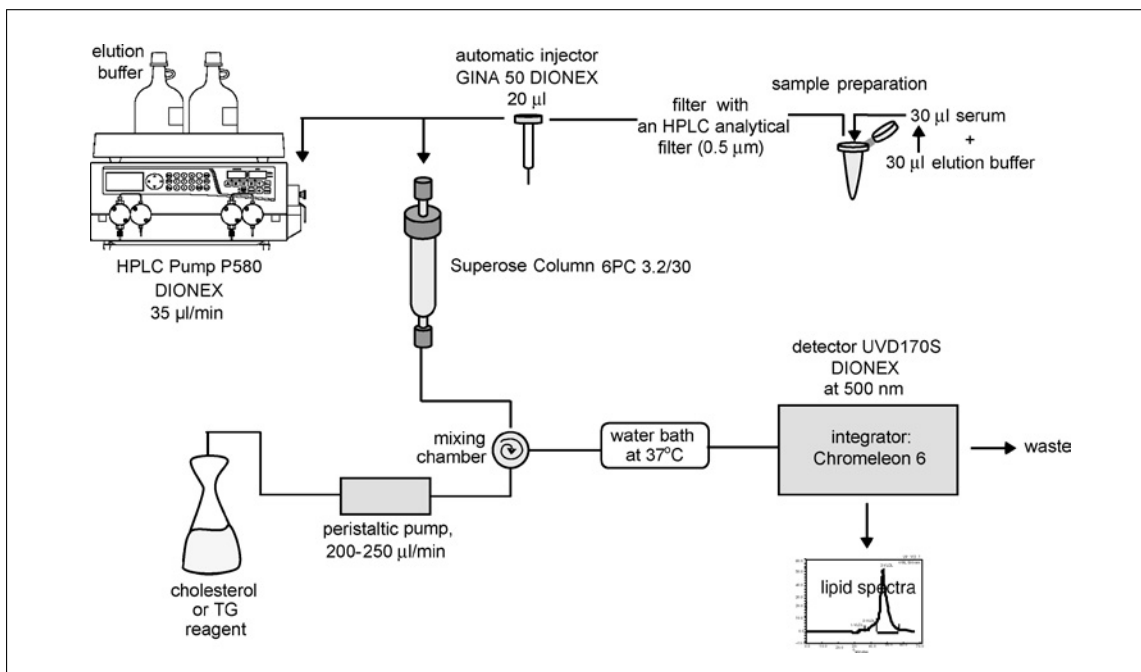


Figure 29B.2.3 Schema showing the adaptation of FPLC to include online lipid measurements.

- Inject the first sample and pass elution buffer through the column at a flow rate of 35 $\mu\text{l}/\text{min}$. Use an online fraction collector to collect ~ 30 to 40 fractions of 75 to 150 μl in microcentrifuge tubes. Measure A_{280} with the online UV detector and record the values using the Chromeleon Integrator system.

In general, the first few fractions can be discarded. Total separation time per sample is 70 min. The known retention times are 33 min for VLDL, 38 min for LDL, and 47 min for HDL.

- Prior to separating the next sample, wash column with elution buffer at a flow rate of 35 $\mu\text{l}/\text{min}$ for 10 min.
- Measure the amount of cholesterol and TG in each fraction as described in Basic Protocol 1.
- Integrate data into a lipoprotein profile using the Chromeleon software according to the manufacturer's instructions, and/or by plotting the lipid content according to fraction number or A_{280} measurement.

One possible adaptation of the FPLC technique described above is to include the capacity for online lipid measurements. A schema showing this adaptation is shown in Figure 29B.2.3; however, specific details are beyond the scope of this unit.

- At the end of the experiment, wash the column sequentially with 4.8 ml water and 4.8 ml 20% ethanol at a flow rate of 35 $\mu\text{l}/\text{min}$. Check that the UV signal is stable after each wash. To prevent drainage during storage, always fill the empty space above the filters and the rubber tubing with 20% ethanol before replacing the protective caps. Store the columns at 4°C.

Columns can be reused up to 100 times. See manufacturer's instructions for special column procedures and troubleshooting if there is a color change in the gel bed, a loss of resolution, or decreased sample recoveries.

TISSUE LIPID ANALYSIS

Lipids can be extracted from tissue (frozen or fresh) and partitioned according to type by using the method of Folch et al. (1951), which is based on a chloroform/methanol extraction. The quantity of cholesterol and TG in the tissue extracts can then be determined by the same enzymatic assays used to detect these lipids in serum. In some instances, it is of interest to distinguish between cholesterol that is esterified and free. Cholesterol ester (CE) represents the metabolically inert pool of cholesterol and is mainly found in the storage lipid droplets, as opposed to free cholesterol, which is associated with the membrane as well as with metabolic regulation. In order to obtain a measurement of CE, enzymatic assays that detect total cholesterol (esterified plus nonesterified) are used in combination with enzymatic assays that detect just free cholesterol. CE is then determined by the subtraction of the free cholesterol measurement from the total cholesterol value.

In general, the lipids that are extracted by chloroform/methanol can be dissolved in 95% ethanol or isopropanol and then quantified via the enzymatic assays. Alternatively, the chloroform/methanol extracts can be dissolved in a 1% Triton X-100 solution in chloroform; the chloroform is then evaporated away, with water being added as the final solvent prior to the measurement of the lipid content by enzymatic assay (Carr et al., 1993). This latter method can be advantageous, as some enzymatic reagents are not compatible with ethanol, and ethanol is a very volatile solvent, making it difficult to manipulate accurately.

CAUTION: Plastic tubes should be avoided due to the possibility of solvents leaching out contaminating lipids. Avoid evaporation of solvents by applying caps or corks to tubes.

Materials

Tissue (~0.5 g)
2:1 (v/v) chloroform/methanol
Methanol
1 α ,2 α -(n)-[³H]cholesteryl oleate (Amersham Biosciences)
Glycerol tri[1-¹⁴C]oleate (Amersham Biosciences)
Chloroform
0.73% (w/v) NaCl
Chloroform/methanol/NaCl (see recipe)
95% ethanol *or* 1% (v/v) Triton X-100 in chloroform
200 μ g/ml cholesterol/triglyceride standard solution: dilute cholesterol and triolein to a final concentration of 200 μ g/ml in isopropanol
14 \times 100-mm Pyrex or borosilicate glass round-bottom tubes with screw caps
Homogenizer (Polytron or equivalent)
12 \times 75-mm glass hemolysis tubes with corks
Whatman filter paper (lipid free)
Nitrogen evaporator with dry bath at ~50°C (e.g., N-EVAP from Organomation)
Scintillation vials, scintillation fluid, and β -counter with dual-channel program for ³H and ¹⁴C

Extract tissue

1. Obtain a piece of tissue with a wet weight of ~0.5 g and place in a Pyrex round-bottom tube. Add 900 μ l of 2:1 chloroform/methanol and homogenize for ~30 sec at high speed in a Polytron-type homogenizer.
2. Add 300 μ l of methanol and vortex. Add 2 μ l each of 1 α ,2 α -(n)-[³H]cholesteryl oleate and glycerol tri[1-¹⁴C]oleate to the sample. Also add the same amounts of these compounds to a scintillation vial to determine counts added.
3. Cap tubes and centrifuge 15 min at 750 \times g, room temperature.

4. Using a glass Pasteur pipet, remove the supernatant to a 12 × 75-mm glass hemolysis tube. If pieces of tissue still contaminate the sample, filter the sample into a new tube using lipid-free Whatman filter paper.
5. Add 400 µl chloroform and vortex.
6. Add 275 µl of 0.73% NaCl and vortex vigorously for ~30 sec.
7. Centrifuge samples 3 min at 750 × g, room temperature.
A phase separation should be visible.
8. Using a Pasteur pipet, remove the top phase and discard. Wash the phase interface by carefully adding 800 µl chloroform/methanol/NaCl with a Pasteur pipet, taking care not to mix the phases. Remove and discard the top phase again and wash the interface another two times in the same manner.
9. While vortexing, add a sufficient amount of methanol to the bottom phase until a single phase is obtained.
10. Evaporate the sample to dryness under a steady stream of nitrogen in a nitrogen evaporator with ~50°C dry bath.
11. Add 500 µl of 95% ethanol or 1 ml of 1% Triton X-100 in chloroform. Vortex, then evaporate to dryness as in the previous step. Resuspend sample in 500 µl deionized water.

Perform lipid assay

12. Generate a standard curve to be used in the enzymatic lipid assays, by preparing duplicate aliquots of 0 (blank), 5, 10, 20, 40, 60, 80, and 160 µl of 200 µg/ml cholesterol/triglyceride standard solution.
13. Evaporate the standards to dryness under nitrogen as in step 10. Resuspend in 500 µl of 95% ethanol or in 1 ml of 1% Triton X-100 in chloroform. For the standards in 1% Triton X-100/chloroform, dry down again and resuspend in 500 µl of deionized water.
14. Determine total cholesterol, free cholesterol, and/or TG content as in Basic Protocol 1. Retain a 25-µl aliquot of each sample and determine radioactivity by β counting.
15. Calculate the percent recovery for cholesterol and TG by dividing the number of counts extracted in the sample by the total number of counts added (step 2), and multiplying the result by 100.

$$\% \text{ recovery for cholesterol} = (\text{}^3\text{H dpm of sample} / \text{total } ^3\text{H dpm added}) \times 100\%$$

$$\% \text{ recovery for TG} = (\text{}^{14}\text{C dpm of sample} / \text{total } ^{14}\text{C dpm added}) \times 100\%$$
16. Adjust the total cholesterol, free cholesterol, and TG (step 14) by multiplying the observed values (in µg lipid/mg tissue) by the percent recovery.
17. Determine the amount of cholesterol ester (CE in µg/mg tissue) by subtracting the amount of free cholesterol from total cholesterol.

ENZYMATIC SERUM BILE ACID MEASUREMENT

Bile acids constitute a group of structurally different molecules synthesized in the liver as a breakdown product of cholesterol (Houten and Auwerx, 2004). In humans, these include the primary bile acids cholic acid (CA) and chenodeoxycholic acid (CDCA), their glycine and taurine conjugates, and the secondary bile acids deoxycholic acid (DCA) and lithocholic acid (LCA). In mice, CDCA and glycine conjugates are practically absent. Instead, muricholic acid (MCA), which is derived from CDCA, is prominent. Since

bile acid synthesis, secretion, extraction from portal blood, and re-excretion all depend on adequate liver function, quantification of serum bile acids represents a sensitive and specific test for liver function. Even more important in the context of metabolic phenotyping is the quantification of bile acid pool size, as most of the regulation of hepatic cholesterol metabolism by bile acids is attributed to the size of the bile acid pool. This is exemplified by the negative feedback mechanism by which bile acids returning to the liver control their own synthesis. However, since recent evidence suggests that the regulatory nature of the bile acid pool depends not just on its size but also on composition, the measurement of both bile acid pool size and composition (see Alternate Protocol 4) are the ultimate goal for evaluating the effect of bile acids (Houten and Auwerx, 2004). For example, a higher abundance of hydrophilic bile acids within the pool seems to induce a greater biliary lipid secretion, whereas hydrophobic bile acids seem to be less effective (Einarsson and Grundy, 1980).

In general, serum bile acids are evaluated under fasting and post-prandial conditions and can be measured using simple enzymatic colorimetric assays (manual or with an automated workstation), or more elaborate methods such as radioimmunoassays or tandem mass spectrometry (see Alternate Protocol 3; Setchell and Matsui, 1983). All bile acids contain a 3 α -hydroxyl group, which can be converted to a 3-keto group by 3 α -hydroxysteroid dehydrogenase. This enzyme simultaneously converts NAD⁺ to NADH, which can be measured spectrophotometrically using different approaches. For example, NADH can be measured by the conversion of nitroblue tetrazolium (NBT) to the formazan dye catalyzed by diaphorase (Mashige et al., 1978).

Materials

Colorimetric Total Bile Acids Assay Kit (containing reagents R1, R2, and R3; Bio-Quant, <http://www.bio-quant.com>)

10 mM taurocholic acid

Bile acid control/calibrator (high and low level; Diagnostics Chemical Limited, <http://www.dclchem.com>)

200- μ l cuvettes

Spectrophotometer

1. Reconstitute the contents of one bottle of R3 with 10 ml of reconstitution buffer (R1). Note the date of reconstitution and the expiration date as indicated by the manufacturer.
2. Prewarm R2 and the reconstituted R3 to room temperature.
3. Serially dilute 10 mM taurocholic acid with water so that the dilutions cover the upper and lower limits of the linear range (1 to 200 μ M).

The method is linear up to 200 μ M. Samples above this range should be adjusted by dilution with saline or water. This self-made taurocholic acid standard can be compared with the single standard of 35 μ M from Bio-Quant.

4. To a 200- μ l cuvette, add 150 μ l of reconstituted R3 and 20 μ l of sample, standard, quality control/calibrator (high and low level of bile acid), or blank (water). Mix well and incubate for 4 min at 37°C.
5. Add 30 μ l of R2, vortex, and immediately read the absorbance at 540 nm as A_1 .
6. Incubate 5 min and read the absorbance at 540 nm as A_2 .
7. Calculate ΔA for each sample and standard using the formula:

$$\Delta A = (A_2 - A_1).$$

8. Generate a standard curve by plotting absorbance versus concentration. Use linear regression to determine the total bile acid concentration of the sample.

Alternatively, the total bile acid concentration can be determined using a single standard (e.g., 35 μM) and the formula:

$$\text{total bile acids (in } \mu\text{M)} = \Delta A_{\text{sample}} / \Delta A_{\text{standard}} \times \text{standard (in } \mu\text{M)}.$$

However, a standard curve should be used instead of a single standard to ensure linearity of the reaction.

ALTERNATE PROTOCOL 3

SERUM BILE ACID MEASUREMENT BY HPLC TANDEM MASS SPECTROMETRY

In order to discriminate between the different bile acid molecules, a reversed-phase HPLC separation of serum can be coupled to the enzymatic detection of bile acids. For this, however, relative large volumes of plasma are needed because of the relatively low sensitivity of the enzymatic assay, which makes this somewhat impractical for measurements in the mouse. Analysis with tandem mass spectrometry (MS/MS) has many advantages. Most importantly, MS/MS gives specific structural information, which in combination with HPLC separation can unambiguously confirm the identity of a molecule. In addition, MS/MS analysis has higher sensitivity when compared with enzymatic detection methods. The sensitivity depends on the type of MS/MS used; in our method the detection limit is 0.05 $\mu\text{mol/liter}$ for conjugated bile acids. MS/MS analysis is thus considered the “gold standard” for analysis of bile acids. A disadvantage is that it requires specific equipment and the availability of stable isotope-labeled standards.

Materials

Plasma

Internal deuterium-labeled standards:

Chenodeoxycholic-2,2,4,4- d_4 acid (C/D/N Isotopes cat. no. D-2772)

Cholic-2,2,4,4- d_4 acid (C/D/N Isotopes cat. no. D-2452)

Glychenodeoxycholic-2,2,4,4- d_4 acid (C/D/N Isotopes cat. no. D-5673)

Glycocholic-2,2,4,4- d_4 acid (C/D/N Isotopes cat. no. D-3878)

Taurine conjugates (custom synthesized)

Acetonitrile (HPLC grade)

1:3 (v/v) methanol/water

1.5-ml microcentrifuge tubes

4-ml glass vials

Nitrogen evaporator with dry bath at $\sim 40^\circ\text{C}$ (e.g., N-EVAP from Organomation)

HPLC-MS/MS (e.g., Hewlett-Packard HP1100 binary pump and a Waters

Micromass Quatro II tandem mass spectrometer equipped with electrospray ionization)

Additional reagents and equipment for liquid chromatography–mass spectrometry (UNIT 10.21)

1. Premix the internal standards in water to prepare a cocktail containing 10 μM of each internal standard. Add 20 μl of this cocktail (200 pmol of each standard) to 20 μl of plasma in a 1.5-ml microcentrifuge tube.
2. Add 250 μl of acetonitrile and mix immediately to precipitate the proteins.
3. Microcentrifuge 12 min at $12,000 \times g$, 4°C .
4. Decant supernatant into a 4-ml glass vial.
5. Evaporate the solvent under a stream of N_2 with heating at 40°C .
6. Dissolve the residue in 70 μl of 1:3 methanol/water.

- Analyze the samples by HPLC-MS/MS (UNIT 10.21).
- Calculate the concentration of bile acids (in μM) by comparing the labeled stable isotopes with the specific bile acids analyzed.

BILE ACID POOL SIZE AND COMPOSITION

The bile acid pool is composed of all bile acids that are present in the liver, gallbladder, and small intestine. As only a small amount of bile acids circulate at a given time in the serum, serum bile acid levels and composition do not always accurately reflect bile acid homeostasis. For this reason, information on the bile acid pool, including of its size and relative composition, yields important additional information about whole-body cholesterol and bile acid homeostasis. Although bile acid pool size can be determined using enzymatic assays and/or radioimmunoassays, the authors recommend using HPLC-MS/MS, since this also reveals the exact composition of the bile acids present in the bile acid pool.

NOTE: All protocols using live animals must first be reviewed and approved by an Institutional Animal Care and Use Committee (IACUC) or must conform to governmental regulations regarding the care and use of laboratory animals.

Materials

- Test animal of interest
- 100% ethanol (HPLC grade)
- Recovery standard (see recipe)
- Internal deuterium-labeled standards:
 - Chenodeoxycholic-2,2,4,4- d_4 acid (C/D/N Isotopes cat. no. D-2772)
 - Cholic-2,2,4,4- d_4 acid (C/D/N Isotopes cat. no. D-2452)
 - Glycochenodeoxycholic-2,2,4,4- d_4 acid (C/D/N Isotopes cat. no. D-5673)
 - Glycocholic-2,2,4,4- d_4 acid (C/D/N Isotopes cat. no. D-3878)
 - Taurine conjugates (custom synthesized)
- Dissecting instruments
- 50-ml conical polypropylene centrifuge tubes (e.g., Falcon)
- 50-ml glass beaker
- Watch glass suitable as cover for 50-ml beaker
- Filter paper (Schleicher-Schuell, type 597)
- Glass funnels
- 50-ml volumetric flask
- HPLC-MS/MS (e.g., Hewlett-Packard HP1100 binary pump and a Waters Micro-mass Quatro II tandem mass spectrometer equipped with electrospray ionization)
- Additional reagents and equipment for liquid chromatography–mass spectrometry (UNIT 10.21)

Extract bile acids

- Remove gallbladder, liver, and the entire small intestine, with contents, from the animal. Store in a 50-ml conical polypropylene centrifuge tube at -80°C until use.

Take extra care not to tear the small intestine, as any leakage of the contents will affect the quantitative measurement. This also applies to the gallbladder and its contents.

It is not necessary to process the entire liver, since this organ normally contains only a small portion of the enterohepatic bile acids. Small pieces of liver can be removed and used for other methods, such as lipid or RNA extraction.

- Transfer the organs to a 50-ml glass beaker, rinsing the storage tube twice with 100% ethanol. Cut tissues into small pieces (1-cm long) with scissors. Add a stir bar.

The total volume of ethanol in the glass beaker should be ~ 40 ml.

ALTERNATE PROTOCOL 4

Mouse Phenotyping

29B.2.15

3. Add exactly 1 ml of recovery standard to each sample.
4. Place the beaker on a hot plate/stirrer for extraction. Cover the beaker with a watch glass. Set the heat so that the ethanol comes to a gentle boil.
5. After 5 min of boiling, remove the beaker from the hot plate and allow to cool to room temperature.
6. Filter the sample using a glass funnel and filter paper into a 50-ml volumetric flask. Rinse beaker with 100% ethanol twice. Bring sample to a volume of 50 ml with 100% ethanol.

Perform assay

7. Transfer 1 ml of ethanolic extract to a 1.5-ml microcentrifuge tube and microcentrifuge 10 min at $15,000 \times g$, room temperature.
8. Dilute supernatant 1:10 with water.
9. Premix the internal standards in water to prepare a cocktail containing 10 μM of each internal standard. Mix equal volumes of the diluted extract and internal standard cocktail (resulting mix contains 5 μM of each internal standard). Analyze the samples by HPLC-MS/MS (*UNIT 10.21*).
10. Calculate the concentration of bile acids (in μM) by comparing the labeled stable isotopes with the specific bile acids analyzed.
11. Calculate the amount of bile acids present in the enterohepatic tissues by correcting for dilution and the amount of norcholic acid added.
12. Divide the result by the body weight of the animal to calculate the amount of bile acids in $\mu\text{mol/kg}$.

SUPPORT PROTOCOL

MODIFIED LOWRY ASSAY FOR LIPOPROTEIN DETERMINATION

For the measurement of protein, several absorbance-based colorimetric methods exist. The most utilized methods rely on the reduction of copper in the presence of a chromogenic reagent. Alternatively, direct absorbance measurements of protein solutions at 280 nm can be performed (*UNIT 10.1A*). The latter technique is readily applied to online detection with FPLC-fractionated samples, whereas a modification of the Lowry procedure can be used for manual measurements. A method for the modified Lowry is provided here and readers are directed elsewhere for alternate methods (Markwell et al., 1978).

The two key modifications in this method include the addition of sodium dodecyl sulfate (SDS) and an increase in the concentration of copper tartrate reagent. The use of SDS in the alkali reagent enables the method to be used with membrane and lipoprotein samples without prior solubilization or lipid extraction. The overnight incubation in alkali is avoided, thus reducing the risk of autooxidation of the lipoprotein samples. The increase in copper tartrate reagent allows the assay of samples containing substantial amounts of sucrose or EDTA without pretreatment. Because precipitating protein quantitatively (i.e., with trichloroacetic acid) may be difficult and increases the time required, this method allows one to process large numbers of samples in a shorter period of time.

Materials

Lowry reagents A and B (see recipe)
 Folin & Ciocalteu's 2N phenol reagent (e.g., Sigma)
 1 mg/ml bovine serum albumin (BSA) standard solution (store at 4°C)
 2 N NaOH
 Spectrophotometer

1. On the day of the assay, prepare reagent 1 by mixing 1 part Lowry reagent B with 100 parts Lowry reagent A.
2. Also on the day of the assay, prepare reagent 2 by mixing Folin & Ciocalteu 2N phenol 1:1 with deionized water.
3. Generate standards by preparing duplicate aliquots of 0 (blank), 5, 10, 20, 40, 60, and 80 μl of 1 mg/ml BSA standard solution. Dilute each aliquot to 200 μl with water.
The standard dilutions should cover the linear part of the standard curve.
4. Prepare duplicate aliquots of sample (for lipoproteins, generally 5 to 10 μl is sufficient). Dilute each aliquot to 200 μl with water.
5. Add 60 μl of 2 N NaOH to each tube and vortex.
6. Add 1 ml of reagent 1 to each tube, vortex, and let incubate for 10 min at room temperature.
7. Add 100 μl of reagent 2 to each tube, vortex, and let incubate for 1 hr at room temperature.
8. Measure the absorbance at 740 nm.
9. Generate a standard curve by plotting a graph of absorbance versus standard concentration. Generate an equation of the line using linear regression. Determine protein concentrations of the unknown samples by comparison with the standard curve.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Chloroform/methanol/NaCl

3 ml chloroform
48 ml methanol
47 ml 0.58% (w/v) NaCl
Store up to 3 months at 4°C

Density gradient solutions

Prepare the following stock salt solutions:

Solution A ($d = 1.005$): dissolve 8.766 g NaCl in 1 liter of aqueous 1% (w/v) disodium EDTA for final concentration of 0.15 M NaCl.

Solution B ($d = 1.346$): dissolve 153 g NaCl and 354 g KBr in 1 liter of aqueous 1% (w/v) disodium EDTA.

Generate all density solutions needed by diluting solution B ($d = 1.346$) with solution A ($d = 1.005$) according to the following equation:

$$(A \times 1.005) + (B \times 1.346) = (A + B) \times X$$

where A = volume of solution A ($d = 1.005$); B = volume of solution B ($d = 1.346$); and X is the desired density of the mixture.

For example, to make 12.05 ml of $d = 1.063$, add 10 ml of solution A and 2.05 ml of solution B. Densities of solutions should be verified by measurement on a Digital Density Meter DMA35.

Store solutions up to several months at room temperature.

Dialysis buffer

10.8 g Na₂HPO₄·12H₂O
1.6 g NaH₂PO₄·2H₂O
35.06 g NaCl
0.4 g NaN₃
0.4 g disodium EDTA
Deionized H₂O to 4 liters
Adjust pH to 7.4 with NaOH
Store up to several months at 4°C

Elution buffer

10 mM Tris·Cl, pH 7.4 (APPENDIX 2A)
30 mM NaCl
Degas and filter through 0.22-μm filter
Store up to several months at 4°C

Lowry reagents A and B

Lowry reagent A:

2% (w/v) Na₂CO₃
0.16% (w/v) sodium tartrate
1% (w/v) SDS
Store up to several months at room temperature

Lowry reagent B:

4% (w/v) CuSO₄·5H₂O
Store up to several months at room temperature

Recovery standard

Dissolve 1 mg norcholic acid (23-nor-5β-cholanic acid-3α,7α,12α-triol) in 50 ml of 100% ethanol and adjust volume to 100 ml with Milli-Q water. Store up to 1 year at room temperature.

COMMENTARY

Background Information

Lipoproteins are supramolecular complexes consisting of lipids and proteins that function in the transport of nonpolar lipids from their sites of synthesis to their sites of utilization (Fig. 29B.2.4; Schaefer et al., 1978, Shepherd and Packard, 1989; Lee et al., 2003). These lipoproteins generally assume a spherical structure whereby the polar phospholipids, proteins, and unesterified or free cholesterol form a monolayer shell that facilitates the transport of core hydrophobic lipids, including cholesterol ester (CE) and triglycerides (TG), in the aqueous blood system. The nonpolar lipids serve numerous biological functions, acting as a transportable metabolic energy pool, as structural components for cell membranes, and as building blocks for hormones. The lipoprotein-associated proteins, or apolipoproteins (apo), have several functions including stabilizing lipid emulsions (apoA, B, C), acting as cofactors and modulators

of enzymatic reactions (apoC, and apoA-I), managing the export of lipids out of cells (apoE and apoB), assisting in the clearance of lipids by target organs and cells through specific receptor interactions (apoB, apoE and apoA), and interfering with platelet aggregation (apoH; Alaupovic, 1991). In humans, abnormalities of lipoprotein transport, synthesis, and degradation, usually of genetic etiology, often lead to disease. Most notably, the metabolic syndrome has an important lipid component including increased plasma levels of TG, apoB, and small dense LDL particles, as well as reduced apoA-I and HDL cholesterol levels. Long-term dyslipoproteinemia associated with the metabolic syndrome increases the risk of atherosclerotic vascular disease and coronary artery disease (Grundy, 2004).

Three pathways can summarize the transport of lipoproteins: exogenous cholesterol transport, forward (endogenous) cholesterol transport, and reverse cholesterol transport (Shepherd and Packard, 1989; Lee et al.,

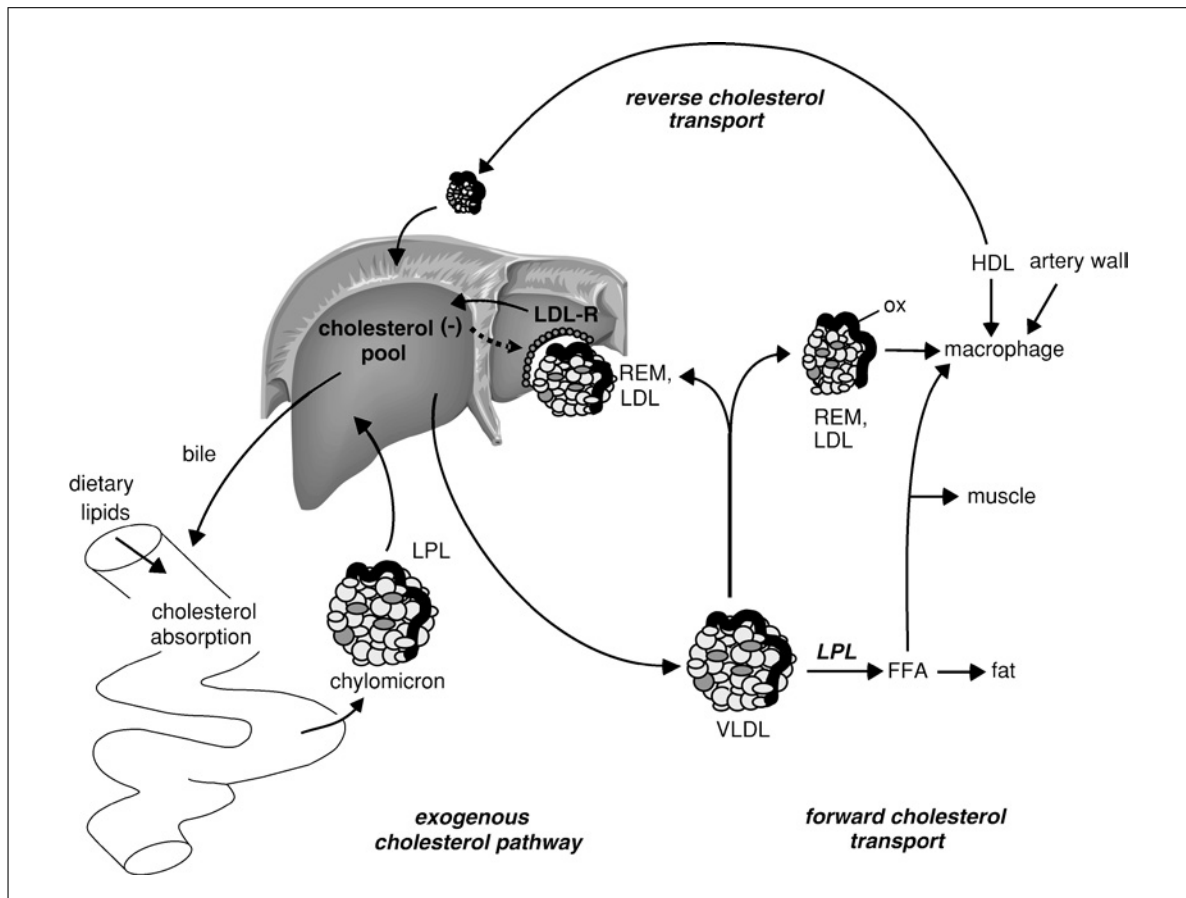


Figure 29B.2.4 A figure depicting the exogenous, forward, and reverse cholesterol pathways.

2003; Assmann and Gotto, 2004). The exogenous pathway describes the absorption and packaging of dietary-derived lipids into large TG-rich particles called chylomicrons. Chylomicrons are synthesized within the intestinal enterocytes and contain predominantly dietary-derived TGs. Lipoprotein lipase (LPL) hydrolyzes the core TG of the chylomicron, producing CE-rich chylomicron remnants, which are cleared by receptor-mediated endocytosis by the LDL receptor and/or remnant receptors. The forward cholesterol transport pathway involves the assembly and secretion of VLDL by the liver, replacing chylomicrons as the main vehicle of TG transport in the circulation in the post-absorptive state. VLDL are synthesized within hepatocytes and contain mainly endogenously derived TG. VLDL particles are acted upon by circulating lipases, including LPL, resulting in the release of free fatty acids (FFA). The CE-enriched particles—IDL or remnants and LDL—are subsequently cleared by LDL-R-mediated endocytosis in the liver or peripheral tissues including macrophages. Alternatively, these lipoproteins can be oxidized and cleared by macrophages via scavenger

receptor-mediated endocytosis. Lipoprotein uptake by arterial macrophages promotes the formation of lipid-laden cells and consequently atherosclerosis.

The reverse cholesterol transport pathway functions to balance cholesterol deposition in the periphery through the receptor-mediated uptake and retention of apoB-containing lipoproteins. This pathway involves the hepatic secretion of nascent or lipid-poor HDL that serves to remove and carry free cholesterol from the extrahepatic tissues back to the liver, where HDL is degraded and cholesterol is excreted. The ATP-binding-cassette protein ABCA1 is one membrane protein implicated in the lipidation of these nascent HDL particles, with phospholipids thus generating HDL particles that can acquire cholesterol from membrane raft domains in the peripheral tissues. The subsequent exposure of these HDL particles to enzymatic and transfer activities in the plasma transforms them into mature spherical HDL particles designed for clearance by the liver. The interested reader is referred to an extensive review on cholesterol transport pathways (Shepherd and Packard, 1989; Lee et al., 2003; Assmann and Gotto, 2004).

Once transferred back to the liver through the reverse cholesterol transport pathway, excess cholesterol can be converted into bile acids or directly secreted into bile and eliminated from the body. As such, bile formation and bile acid biosynthesis have an important function in whole-body cholesterol homeostasis (Houten and Auwerx, 2004). Bile acids are essential for the absorption of dietary lipids, including fat-soluble vitamins. In recent years, an important role for bile acids as signaling molecules has emerged (Houten et al., 2006). Different bile acids may have different signaling capacities, underpinning the importance of determining bile acid composition and pool size. Most of the bile acids are present within the enterohepatic organs. Usually, bile acids are stored in the gallbladder; however, when a meal is ingested, they flow into the duodenum and intestine. The bile acids are efficiently (95%) absorbed again by passive diffusion and active transport from the terminal ileum, then transported back to the liver via the portal vein. The liver extracts bile acids efficiently from the portal vein and transports them back into the bile duct. From here they can go to the gallbladder for storage or flow into the intestine again. This completes their enterohepatic circulation.

In general, plasma lipoproteins exist in a dynamic equilibrium; thus, what is observed by analyzing lipoprotein fractions is really a static snapshot of this equation. It is best to evaluate lipid and lipoprotein homeostasis in the context of plasma lipid measurements combined with isolated lipoprotein determinations. For example, when there is an elevation in plasma TG in the chylomicron fraction, this may be interpreted as a disorder in chylomicron clearance subsequent to a problem with LPL-mediated hydrolysis of TGs in these chylomicrons. When TG levels are elevated together with elevations in plasma cholesterol and low HDL values, this may rather indicate the presence of the metabolic syndrome with insulin resistance. More detailed measurements, such as individual apolipoprotein content, are possible through such techniques as immunoassays and isoelectric focusing gel electrophoresis. However, these types of data are not clinically useful by themselves and should always be interpreted in the context of the plasma lipid profiles and lipoprotein fraction analysis. For example, the apoB level may provide information on the number of potentially atherogenic lipoproteins, as there is one apoB molecule per particle of chylomicron, VLDL, LDL, or IDL. In humans,

the presence of small, dense LDL particles has been shown to be associated with the metabolic syndrome, and an elevated LDL apoB level may represent the presence of small dense LDL.

Finally, when analyzing lipids in mice, it is important to remember that mice are different from humans. Mouse HDL cholesterol represents between 65% and 75% of the total cholesterol, whereas the LDL cholesterol fraction is only between 10% and 20%. In contrast, human HDL represents less than 20% of the total cholesterol and HDL over 50% to 60%. This is in part due to differences in the metabolism of HDL. Mice have a significantly higher level of circulating lipases, lack the cholesterol ester transfer protein CETP, and have high levels of the phospholipid transfer protein PLTP (Tall, 1995). This is one of the reasons why it is often difficult to mimic human lipid disorders in a mouse. In order to circumvent such problems it is necessary to use transgenic mice, such as the apoE knockout, and/or challenges with environmental conditions, such as administration of a high-fat diet, which can result in more human-like lipoprotein-associated pathophysiology.

Critical Parameters

Blood lipid parameters

To make an initial assessment of lipid and lipoprotein homeostasis in mouse models, the measurement of TG, FFA, and total, LDL, and HDL cholesterol comprise the minimal set of tests required. Blood lipid levels are sensitive to numerous parameters: length of fasting, method of blood draw, time of day for sampling, and method of quantification. Therefore, animals and samples should be manipulated and measured in a consistent way so that the results are readily comparable. In general, only small differences will be observed except under spectacular conditions. When abnormalities of lipid and glucose homeostasis are suspected, a second serum analysis after a shorter (4-hr) fast, or even in absence of fasting, is advisable. Furthermore, retesting in older mice yields quick information on the evolution of biochemical and metabolic parameters.

Manual determination of lipid parameters has a number of disadvantages, since reconstituted reagents are generally good for only 3 weeks. Therefore, samples are usually collected and then batch processed for efficiency. In addition, it is more difficult to take advantage of interlaboratory control systems, which are meant for high-throughput automated

machines. Use of automated systems allows for many more blood parameters to be analyzed simultaneously and requires a smaller volume of serum per test. This can leave sufficient sample for other measurements needed for a good first metabolic screen, including measurement of liver enzymes such as AST and ALT, urea, creatinine, total protein, albumin, glucose, insulin, and total bilirubin.

In general, plasma samples, which are collected in the presence of an anticoagulant, are preferred over serum samples, especially for the automated procedure, as the fibrin in serum can cause clogging of the tiny injection needles. In most cases, either EDTA or heparin can be used as the anticoagulant. In the case of FFA measurements, however, heparin should be avoided because it may activate lipoprotein lipase, which can result in false increases in FFA levels. The range of sensitivity for each enzymatic reaction as specified by the manufacturer should be respected; when samples are too lipemic, they need to be diluted in saline and a correction factor must be applied. Additional considerations for each kit are also given by the manufacturer and should be observed. For example, because the concentration of unesterified fatty acids rises quickly in due to the lipolysis of acylated lipids, FFA determination should be performed soon after the blood draw or on freshly frozen samples. Furthermore, exceeding the recommended incubation time for FFA reactions can lead to aberrantly increased values. This is in contrast to the cholesterol reaction, which is complete in about 2 min but shows no change in absorbance when incubation is continued for another 30 min. In general, these parameters should be carefully standardized so that intra-assay variation is minimal, a task that is more challenging with the manual method.

Lipoprotein analysis

Gel-exclusion chromatography is a quicker method than sequential ultracentrifugation for the separation of lipoproteins, but there is sometimes a disparity between the fractions isolated by centrifugation and by chromatography. This disparity is at least partially due to the difficulty in deciding where the boundaries between the conventional fractions lie on the chromatographic elution curve.

Measuring the protein content of each of the isolated fractions either by UV absorption or by a modified Lowry assay (Markwell et al., 1978) allows for the estimation of particle size. This is particularly relevant in the analysis of

LDL because the protein measurement in the LDL fraction represents the quantity of apoB, as only one apoB exists per LDL particle. Thus, a higher amount of protein can reflect a greater number of particles. When cholesterol content is normalized to protein content, LDL particle density can be compared. As small dense LDL are associated with increased risk of CAD, this is a particularly important parameter (Lada and Rudel, 2004).

Tissue lipid content

Some optimization of this protocol may be required, depending on the mouse model and tissue type used. More lipids will accumulate in a high-fat-diet mouse model than in a chow-fed mouse model. Sample volumes and/or quantity of tissue used for extraction should be titrated such that the final reaction product is within the linear range of the system as stated by the manufacturer.

Bile acid measurements

Most of the bile acids of the enterohepatic circulation can be found in the gallbladder and the small intestine. The quantitative contribution of bile acids in the liver is relatively small. Therefore, it is not critical to extract the whole liver to determine bile acid pool size; however, it is critical to prevent any loss of bile from the gallbladder and contents from the small intestine, since this will affect the quantitative recovery of bile acids.

Several types of colorimetric assays are available for use in the measurement of total bile acids (see Sigma-Aldrich, Randox, BioQuant, DCL, bioMerieux). The general enzymatic principle is that the bile acids are enzymatically oxidized by the enzyme 3 α -hydroxysteroid dehydrogenase, with the concurrent reduction of NAD⁺ to NADH. However, in the fasting state and early stages of disease, levels of bile acids are low, and relatively few molecules of bile acid are available for the enzyme to act upon. Consequently, only a small amount of NAD⁺ is reduced to NADH, making it analytically difficult to directly measure the change in NADH by absorbance at 340 nm. Assay sensitivity is increased by coupling NBT to NADH, since this produces a far greater change in absorbance (now measured at 540 nm) per molecule. Post-prandial lipemia and/or hemolysis interfere with this method (Mashige et al., 1978). In some instances, ultracentrifugation and overnight refrigeration of lipemic specimens can reduce the interference.

In addition to colorimetric assays, a radioimmunoassay (RIA) can also be used to

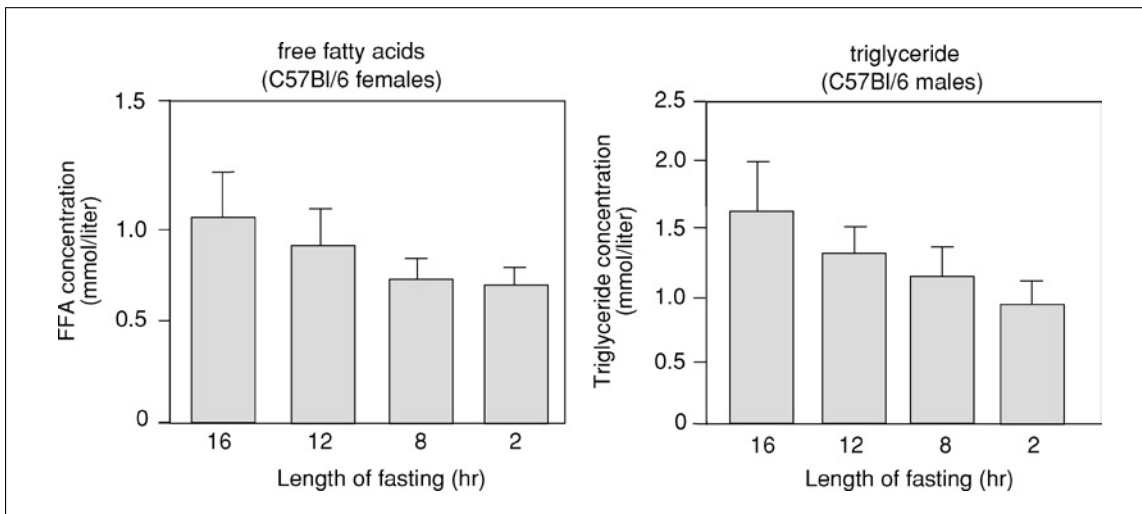


Figure 29B.2.5 Effect of length of fasting on FFA and TG plasma level in C57Bl/6 mice.

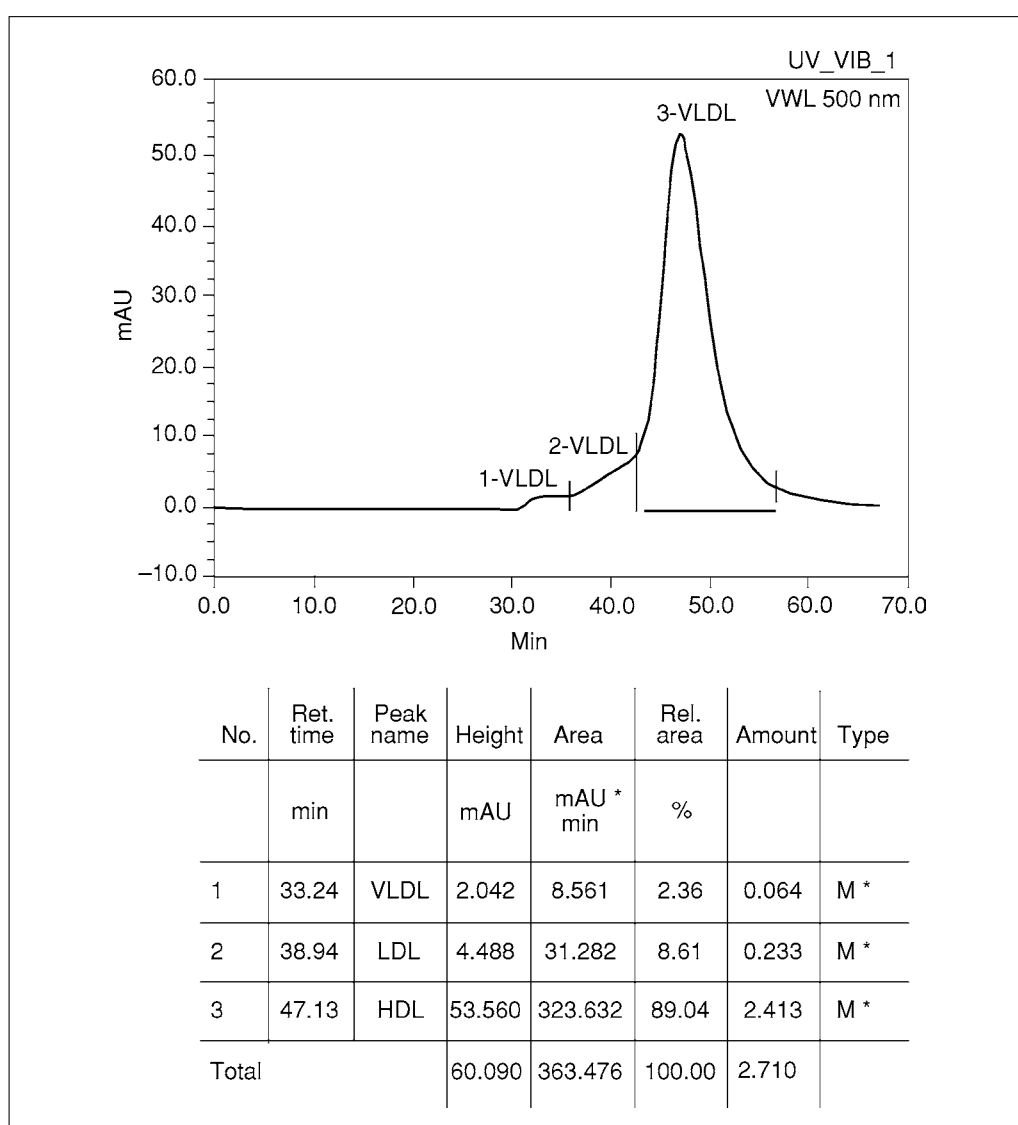


Figure 29B.2.6 Typical lipoprotein profile in plasma of C57Bl/6 mice.

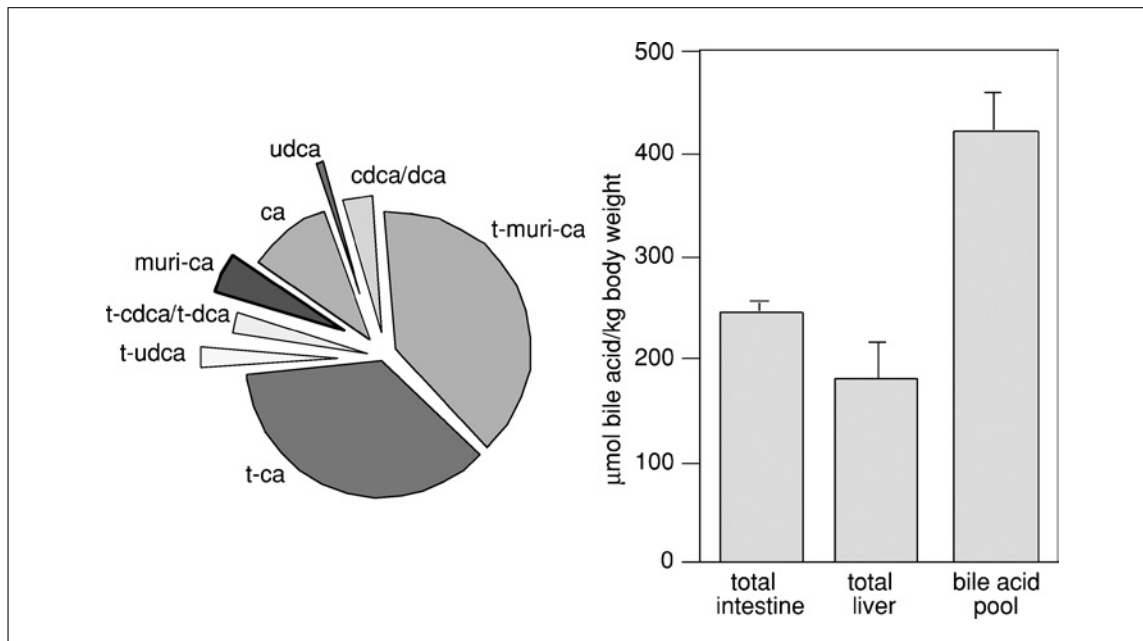


Figure 29B.2.7 A typical example of the composition and size of the bile acid pool. Abbreviations: ca, cholic acid; cdca, chenodeoxycholic acid; dca, deoxycholic acid; muri-ca, muricholic acid; udca, ursodeoxycholic acid; and t-, taurine conjugates.

measure total serum bile acid concentrations. An RIA (e.g., *UNIT 10.24*) uses a labeled iodine tracker and an antiserum against the amino acids (taurine or glycine) to which the bile acid is conjugated. Several differences have been noted between the results produced by these techniques, including different absolute values of bile acids, differential interference with hemolysis and lipemia, and differences in the volume of serum required. The authors suggest that careful consideration be given by the user to the appropriateness of the method, depending on individual needs.

Blood draw and handling

If analyses are not performed on the day of blood collection, samples should be stored at -20°C or -80°C . Tests should preferably be performed within 10 days of collection. Stability and storage depend on the different parameters to be determined. Plasma samples can be used undiluted, or diluted with deionized water if the volume is insufficient. In cases where the blood is hemolyzed, hemoglobin can sometimes interfere with certain measurements, and either a correction factor and/or resampling is suggested.

Anticipated Results

Blood lipid content

Blood lipid levels are generally reported in units of mM or mg/dl. The conversion factors for changing cholesterol, TG, and

FFA concentration from mg/dl to mM are 2.586, 1.126, and 0.035, respectively. Blood lipid values are highly influenced by genetic background and environmental factors. Classically, C57BL/6 mice have low plasma cholesterol and low plasma triglyceride (TG) but high plasma glucose levels, whereas C3H mice have high plasma cholesterol, TG, and glucose levels, and 129 inbred strains have high plasma cholesterol and low TG levels. Environmental factors such as diet and length of fasting will also impact some lipid variables (see Fig. 29B.2.5).

Lipoprotein analysis

A typical lipoprotein profile as determined by FPLC is shown in Figure 29B.2.6. Lipoprotein analysis according to differential ultracentrifugation is represented as the amount of lipoprotein (cholesterol or TG) per μg of protein or is normalized to the initial volume of serum.

Tissue lipid content

TG, CE, and free cholesterol content of tissue lipid is generally expressed as mg lipid/mg tissue.

Bile acid measurements

An illustration of bile acid composition analysis is shown in Figure 29B.2.7. In mice, most of the bile acids are conjugated to taurine. Glycine conjugates are practically absent. In the liver and gallbladder, $\sim 95\%$ of the bile

acids are conjugated, whereas in the intestine 30% of the bile acids are deconjugated by microbial enzymes. Typically 42% to 46% of the bile acids are present in the liver and gallbladder; the rest can be found in the small intestine. The hydrophobic-hydrophilic balance of the pool that results from the characteristics and proportions of the individual bile acids present within the pool is critical, as it may dictate most of the effects of bile acids on hepatic cholesterol metabolism (Mashige et al., 1978). For example, hydrophobic bile acids (e.g., CDCA) seem to increase biliary lipid secretion and inhibit cholesterol and bile acid synthesis, whereas hydrophilic bile acids (e.g., CA) are less effective.

In general, the protocol for testing bile acids entails blood sampling after a period of fasting followed by refeeding. Fasting levels of serum bile acids can be used for the diagnosis and prognosis of liver disease. Post-prandial bile acid measurements are also sensitive indicators of liver, intestinal, and gallbladder function. In this sense, bile acid measurements are considered to be a superior indicator of liver disease, compared to serum liver enzyme measures.

Literature Cited

Alaupovic, P. 1991. Apolipoprotein composition as the basis for classifying plasma lipoproteins: Characterization of Apo A- and Apo B-containing lipoprotein families. *Prog. Lipid Res.* 30:105-138.

Assmann, G. and Gotto, A.M. Jr. 2004. HDL cholesterol and protective factors in atherosclerosis. *Circulation* 109:III8-III14.

Brousseau, T., Clavey, V., Bard, J.M., and Fruchart, J.C. 1993. Sequential ultracentrifugation micro-method for separation of serum lipoproteins and assays of lipids, apolipoproteins, and lipoprotein particles. *Clin. Chem.* 39:960-964.

Carr, T.P., Andresen, C.J., and Rudel, L.L. 1993. Enzymatic determination of triglyceride, free cholesterol, and total cholesterol in tissue lipid extracts. *Clin. Biochem.* 26:39-42.

De Lalla, O. and Gofman, J.W. 1954. Ultracentrifugation as a technique for separation of lipoprotein fractions. In *Methods in Biochemical Analysis* (D. Glick, ed.) pp. 459-478. Wiley Interscience, New York.

Einarsson, K. and Grundy, S.M. 1980. Effects of feeding cholic acid and chenodeoxycholic acid on cholesterol absorption and hepatic secretion of biliary lipids in man. *J. Lipid Res.* 21:23-34.

Folch, J., Ascoli, I., Lees, M., Meath, J.A., and LeBaron, N. 1951. Preparation of lipid extracts from brain tissue. *J. Biol. Chem.* 191:833-841.

Grundy, S.M. 2004. Obesity, metabolic syndrome, and cardiovascular disease. *J. Clin. Endocrinol. Metab.* 89:2595-2600.

Houten, S.M. and Auwerx, J. 2004. The enterohepatic nuclear receptors are major regulators of the enterohepatic circulation of bile salts. *Ann. Med.* 36:482-491.

Houten, S.M., Watanabe, M., and Auwerx, J. 2006. Endocrine functions of bile acids. *EMBO J.* 25:1419-1425.

Lada, A.T. and Rudel, L.L. 2004. Associations of low density lipoprotein particle composition with atherogenicity. *Curr. Opin. Lipidol.* 15:19-24.

Lee, C., Olson, P., and Evans, R.M. 2003. Minireview: Lipid metabolism, metabolic diseases, and peroxisome proliferator-activated receptors. *Endocrinology* 144:2201-2207.

Markwell, M.A.K., Haas, S.M., Bieber, L.L., and Tolbert, N.E. 1978. A modification of the Lowry procedure to simplify protein determination in membrane and lipoprotein samples. *Anal. Biochem.* 87:206-210.

Mashige, F., Osuga, T., Tanaka, N., Imai, K., and Yamanaka, M. 1978. Continuous-flow determination of bile acids in serum, and its clinical application. *Clin. Chem.* 24:1150-1154.

Schaefer, E.J., Eisenberg, S., and Levy, R.I. 1978. Lipoprotein apoprotein metabolism. *J. Lipid Res.* 19:667-687.

Setchell, K.D. and Matsui, A. 1983. Serum bile acid analysis. *Clin. Chim. Acta* 127:1-17.

Shepherd, J. and Packard, C.J. 1989. Lipoprotein metabolism. In *Human Plasma Lipoproteins* (J.C. Fruchart and J. Shepherd, eds.) pp. 55-78. Walter de Gruyter, New York.

Tall, A. 1995. Plasma lipid transfer proteins. *Annu. Rev. Biochem.* 64:235-257.

Contributed by Carmen A. Argmann
 Institut de Genetique et de Biologie
 Moleculaire et Cellulaire
 Illkirch, France

Sander M. Houten
 Academic Medical Center
 Amsterdam, Netherlands

Marie-France Champy
 Institut Clinique de la Souris
 Illkirch, France

Johan Auwerx
 Institut de Genetique et de Biologie
 Moleculaire et Cellulaire and Institut
 Clinique de la Souris
 Illkirch, France

The development of glucose intolerance and insulin resistance (IR), characteristics of type 2 diabetes mellitus (T2DM), is a multistep process that is influenced by both genetic and environmental factors. As T2DM is currently the most common endocrine disorder in humans, the detailed understanding of these processes has been critically sought, with the mouse as the leading model. The ability to measure IR and glucose tolerance in the mouse is important both for understanding the pathogenesis of T2DM and for testing potential therapeutic interventions. When an initial blood screen in the mouse indicates altered levels of insulin and/or glucose, a number of well-established tests used to estimate glucose homeostasis and insulin sensitivity should be considered. These tests include model assessments, glucose/insulin infusion sensitivity tests, and clamps.

Because no single test is appropriate under all circumstances and there is considerable variation in the complexity and labor intensity of these various methods, care should be taken in prioritizing them. In general, the choice of method depends on the size and type of study to be undertaken, the goal of the experiment, and the technical expertise at hand. In this unit, tests for exploring glucose homeostasis and insulin sensitivity are described in order of technical invasiveness and ease, with the least invasive described first. In general, this is also the order in which the researcher should evaluate glucose homeostasis and insulin sensitivity in a mouse model, with the importance of each additional test being dependent on the result of the previous test. Thus, although the euglycemic clamp has been referred to as the gold standard for measuring IR, the relevance of performing this test must be considered if prior screening gives no indications that glucose homeostasis or IR is affected.

STRATEGIC PLANNING

The circumstances under which insulin sensitivity and/or glucose tolerance are measured highly influences their interpretation. In fact, insulin sensitivity changes not only in pathological situations but also during normal physiological stresses. To make any consistent assessment of overall insulin sensitivity, it is important to report any specific terminology or descriptors associated with the circumstances of measurement. More importantly, strategic planning is required to optimally control for the tissue-specific, metabolite-specific, and process-specific effects that can influence the conclusions drawn from an experiment. For example, when measuring the effect of IR by following the metabolic clearance of glucose, glucose-mediated glucose uptake must be taken into account, since it is greatly dependent on physical activity. Glucose is liberated from glycogen stores upon exercise, and the capacity to clear glucose from the circulation also increases. Stress and the time of day are also important determinants of IR. In general, mice have increased cortisol levels in the evening, leading to higher hepatic glucose output that can inflate IR. Similarly, increased adrenaline and cortisol levels due to stress can increase hepatic glucose output and consequently affect the assessment of IR. Furthermore, most methods used in the assessment of insulin sensitivity examine the insulin response of a single metabolite, glucose, primarily in the muscle and liver, and under fasting conditions. Because these results reflect only the conditions under which they were obtained, care should be taken not to over-interpret the data.

All these types of confounders and their potential influence on the ability to compare IR among various mouse models prompted the EU-funded Eumorphia program

(<http://www.eumorphia.org>) to standardize several of the parameters associated with measures of glucose homeostasis and IR. The standardized conditions that are presented here as individual protocols are mostly based on the Eumorphia guidelines. However, individual investigators may need to modify these protocols according to their specific needs. A summary of the methods described here is presented in Table 29B.3.1.

NOTE: The protocols described here involve mouse techniques ranging from very basic (e.g., picking up a mouse) to very specialized (e.g., catheter surgery). All personnel taking part in these techniques must be trained beforehand to obtain results while minimizing stress-induced effects on the mouse. Protocols for these techniques can be found in detail elsewhere (Hedrich, 2004). All protocols using live animals must first be reviewed and approved by an Institutional Animal Care and Use Committee (IACUC) or must conform to governmental regulations regarding the care and use of laboratory animals.

Table 29B.3.1 Summary of Methods Used to Assess Glucose Tolerance and Insulin Sensitivity^a

Test	Labor intensity	Method	Interpretation
FPG and FPI	One blood sample (50–100 μ l; tail bleed or retro-orbital)	Overnight or 4 hr fasting glucose and insulin measurements	Estimate of basal IR
HOMA	Simple calculations	Two formulas using FPG and FPI	Estimate of basal IR and β -cell function (not validated in mouse)
MTT	Two samples (50–100 μ l; tail bleed or retro-orbital)	Overnight fast, and glucose and insulin measurements \pm a standardized meal	Estimate of basal and postprandial IR; intestinal nutrient absorption taken into account
OGTT	Eight small samples (\sim 3 μ l) over 180 min (tail bleed)	Overnight fast, and glucose measurements after oral glucose load	Estimate of IR based on the evolution of glucose levels; intestinal glucose absorption taken into account
IPIST	Six small samples (\sim 3 μ l) over 90 min (tail bleed)	Overnight fast, and glucose measurements after IP insulin load	Estimate of basal insulin sensitivity based on the evolution of glucose levels
IPGTT	Eight small samples (\sim 3 μ l) over 180 min (tail bleed)	Overnight fast, and glucose measurements after IP glucose load	Estimate of IR based on the evolution of glucose levels
Euglycemic clamp	Catheter surgery, small blood samples every 5 to 10 min over 180 min (tail bleed)	Constant rate insulin infusion, variable rate glucose infusion; frequent glucose measurements	Estimate of IR based on average glucose infusion rate over the last 60 min; often regarded as the gold standard

^aFPG, fasting plasma glucose; FPI, fasting plasma insulin; IR, insulin resistance; HOMA, homeostatic model assessment; MTT, meal tolerance test; OGTT, oral glucose tolerance test; IPIST, intraperitoneal insulin sensitivity test; IPGTT, intraperitoneal glucose tolerance test.

GLUCOSE MEASUREMENT BY BLOOD GLUCOSE MONITOR

Circulating glucose is the first and foremost biochemical measure to be taken when evaluating the metabolic phenotype of mice, as changes in this vital and typically strictly regulated metabolite invariably indicate alterations in some aspect of metabolism.

The limited blood volume of the mouse imposes special requirements for the sensitivity of the technique by which any circulating metabolite or substance, including glucose, is measured. In the case of glucose, the problem is exaggerated by the fact that several of the metabolic test protocols provided in this unit require repeated glucose measurements from the same mouse over relatively short intervals. In addition, some tests (e.g., euglycemic clamp) require virtually on-line information about blood glucose levels, setting strict limits for the rapidity of the glucose measurement. Within these parameters, the best method for measuring blood glucose during mouse metabolic phenotyping experiments is usually the blood glucose monitor designed for diabetic humans. These monitors are readily available, easy to use, and relatively inexpensive. Most importantly, they provide reliable results within seconds from a small sample of blood.

This protocol outlines the use of the Accu-Chek Active blood glucose monitor from Roche Diagnostics. For specifics and troubleshooting, refer to the manufacturer's user manual. Examples of similar products include OneTouch Ultra (LifeScan), BD Logic Blood Glucose Monitor (BD), Precision Xtra (Abbott), and HemoCue Glucose (HemoCue). Other techniques for measuring blood glucose, more suited for the analysis of frozen plasma or serum samples at some later point in time, include enzymatic methods, mostly utilizing either hexokinase with glucose-6-phosphate dehydrogenase or glucose oxidase. These methods can be performed manually (e.g., Glucose (HK) assay kit; Sigma) or automatically (e.g., AU-400 automated laboratory work station with OSR6121 glucose reagent; Olympus).

Materials

Mice

Low- and high-level glucose control solutions (e.g., Accu-Chek active glucose control solutions, Roche Diagnostics; low concentration: 50 mg/dl, reference range 42 to 72 mg/dl, and high concentration: 300 mg/dl, reference range 290 to 328 mg/dl)

Blood glucose monitor and associated glucose test strips for glucose measurement (e.g., Accu-Chek Active, Roche Diagnostics)

Scalpel blade

1. Calibrate the blood glucose monitor by first inserting the code key into the slot on the left side with the monitor powered off. Then insert a test strip and verify that the code appearing on the monitor matches that on the test strip vial.

Perform this step every time a new box of test strips is opened.

2. Perform a performance check on the monitor. Insert a fresh test strip, and after the monitor is activated, press the button marked S. Carefully dispense one drop from the low-level control solution bottle and touch it to the center of the orange pad on the strip. Verify that the result, appearing in a few seconds, is within the low range indicated on the test strip vial. In cases where high glucose values are expected, repeat this procedure using a high-level control solution.

A performance check should be done every time a new box of test strips is opened, or if the performance of the monitor is in doubt. It is best to use the same lot number of strips for one experiment, as lot-to-lot variations can occur.

3. Insert a fresh test strip.

The blood sample must be applied within the next 2 min before the monitor powers off.

**BASIC
PROTOCOL 2**

4. Draw a small ($\sim 3 \mu\text{l}$) blood sample from a mouse by nicking the lateral tail vein 1 to 2 cm from the tip with a scalpel blade (see *UNIT 29A.3*).
5. Transfer the drop of blood to the orange pad on the test strip, preferably by touching the strip directly to the bleeding tail wound, or by pipeting. Record the result, which appears in 5 sec.
6. Apply direct pressure to the incision until the blood has clotted and return the mouse to its cage.

INSULIN MEASUREMENT BY ELISA

In addition to glucose, levels of circulating insulin are often of interest due to the importance of this β -cell hormone on the regulation of circulating glucose levels. In general, changes in insulin appear earlier than changes in glucose during the progression of experimental metabolic interventions.

This protocol outlines the use of the Ultrasensitive Mouse Insulin ELISA from Merckodia. This solid-phase direct sandwich ELISA uses two monoclonal antibodies raised against separate antigenic sites on the insulin molecule (for additional description of sandwich ELISAs, see *UNIT 11.2*). Insulin in the sample is immobilized by one anti-insulin antibody bound to the assay well, and is labeled by a second anti-insulin antibody conjugated with peroxidase. Unbound enzyme-conjugated antibody is removed during washing, and the bound conjugate is detected colorimetrically. For additional information, refer to the manufacturer's instructions. Other sources for sensitive ELISA kits for measuring insulin in small volumes of mouse blood or other samples include Alpco Diagnostics, Crystal Chem, and Linco Research.

CAUTION: The stop solution in this kit contains 0.5 M H_2SO_4 . Follow routine precautions for handling hazardous chemicals.

Materials

- Ultrasensitive Mouse Insulin ELISA kit (e.g., Merckodia) containing:
 - 96-well microtiter plates coated with mouse monoclonal anti-insulin antibody
 - Calibrators (i.e., standards)
 - Antibody-enzyme conjugate (concentrate) and buffer
 - Wash buffer (concentrate)
 - TMB substrate reagent (3',3',5',5'-tetramethylbenzidine; light sensitive)
 - Stop solution: 0.5 M H_2SO_4
- Mouse serum or plasma
- 0.9% (w/v) NaCl

- Microtiter plate shaker
- Microtiter plate washer (optional)
- Plate reader capable of measuring at 450 nm

1. Pipet 25 μl of calibrator 0 into each well of a 96-well microtiter plate coated with mouse monoclonal anti-insulin antibody.
2. Pipet calibrators (5 μl) and mouse serum or plasma (5 μl) into duplicate wells of the antibody-coated microtiter plate.
3. Add 50 μl of antibody-enzyme conjugate reagent to each well.

Prepare by diluting the enzyme conjugate concentrate eleven-fold with enzyme conjugate buffer, e.g., add 600 μl to 6 ml of buffer per one 96-well assay. Store up to 2 months at 2° to 8°C.

4. Incubate on a shaker for 2 hr at room temperature (18° to 25°C).

5. Aspirate the reaction volume, add 350 μ l of wash buffer to each well, and aspirate completely. Repeat five times for a total of six washes.

Prepare the wash buffer by diluting the wash buffer concentrate 21-fold with redistilled water, e.g., add 35 ml to 700 ml of redistilled water per one 96-well assay. Store up to 4 weeks at 2° to 8°C.

Washing can also be performed using an automated microtiter plate washer.

6. After the final wash, invert the plate and tap firmly against absorbent paper.
7. Add 200 μ l TMB substrate reagent to each well.
8. Incubate 30 min at room temperature.
9. Add 50 μ l stop solution.
10. Place the plate on the shaker for \sim 5 sec to ensure mixing of substrate and stop reagents.
11. Measure the absorbance at 450 nm and calculate the results according to the manufacturer's manual.

In short, for computerized analysis, plot the absorbance values (y-axis) for the calibrators against the corresponding insulin concentrations (x-axis) as a log-log graph, and fit a cubic spline regression curve (also called third-order polynomial trend line) to obtain the insulin concentrations for the samples. Sometimes a fourth-order polynomial trend line can provide more accurate results. Alternatively, plot the absorbance values on the x-axis against the insulin concentrations on the y-axis, and fit a power trend line to obtain an equation that is easier to solve than above. For manual calculation, plot the calibrators on a log-log or log-lin paper as above, and read the insulin concentrations for each sample from the curve.

12. If the insulin concentration exceeds the value of the highest calibrator, dilute the sample ten-fold with 0.9% NaCl and repeat the analysis.

HOMEOSTATIC MODEL ASSESSMENT

The homeostatic model assessment (HOMA) is a mathematical model that takes into account several parameters known to effect glucose homeostasis. It relies on experimental human data, and allows for values of insulin sensitivity and β -cell function to be obtained when simultaneous fasting plasma glucose and fasting insulin concentrations are known (Matthews et al., 1985; Levy et al., 1998; Wallace et al., 2004). The solutions for the full model have highly significant correlation to euglycemic clamp results in humans (Wallace et al., 2004). Due to their nonlinear nature, they are best calculated using a computerized program available at <http://www.dtu.ox.ac.uk/>. Two simple formulas can be used to approximate the HOMA indices, one for insulin resistance (HOMA-IR) and one for β -cell function (HOMA-%B). This protocol describes these two simple formulas. It should be noted that the validation of HOMA on mice is lacking, so these indices should be used with caution until such validation becomes available.

Materials

Mice

Heparinized tubes for blood sampling (e.g., Microvette, Sarstedt)

Mouse cages

Scalpel blade

Refrigerated microcentrifuge

Additional reagents and equipment for measuring blood glucose (see Basic Protocol 1) and insulin (see Basic Protocol 2)

BASIC PROTOCOL 3

Mouse
Phenotyping

29B.3.5

1. Fast mice 14 to 16 hr (overnight). Ensure that they have ad libitum access to drinking water.
2. Obtain ~100- μ l blood samples in heparinized tubes by nicking the lateral tail vein 1 to 2 cm from the tip with a scalpel blade (see UNIT 29A.3). Apply direct pressure to the incision until the blood clots and then return the mouse to its cage.

A blood glucose measurement can also be taken from whole blood before plasma separation, using a blood glucose monitor (see Basic Protocol 1).

3. Microcentrifuge the blood samples 15 min at $2300 \times g$ (5000 rpm), 4°C. Collect the plasma (~40 μ l) into clean tubes, and store at -20° or -80°C until glucose and insulin measurement (up to 1 month for glucose and 6 months for insulin).
4. Measure blood glucose (see Basic Protocol 1) and insulin (see Basic Protocol 2) levels.
5. Calculate the HOMA-IR and HOMA-%B results according to the following two equations:

$$\text{HOMA-IR} = \frac{\text{FPI} \times \text{FPG}}{22.5 \times 18}$$

$$\text{HOMA-\%B} = \frac{(20 \times \text{FPI})}{\frac{\text{FPG}}{18} - 3.5}$$

where FPI is the fasting plasma insulin concentration ($\mu\text{U/ml}$) and FPG is the fasting plasma glucose concentration (mg/dl).

Glucose values with mmol/liter as a unit can also be used by removing the factor 18 from the formulas (for glucose, 1 mM = 18 mg/dl). Insulin values with pmol/liter as a unit can be converted to $\mu\text{IU/ml}$ by dividing values by 6.945, or from $\mu\text{g/ml}$ to IU/ml by dividing values by 45.5 (for insulin, 1 $\mu\text{U/ml}$ = 6.945 pmol/liter and 1 IU = 45.5 μg of standard preparation).

BASIC PROTOCOL 4

MEAL TOLERANCE TEST

The meal tolerance test (MTT) is a simple method to explore glucose metabolism and insulin secretion following a meal by measuring fasting and fed glucose and insulin levels. First, animals fast for 14 to 16 hr, after which a fasting blood sample is collected by retroorbital venipuncture (Hedrich, 2004). The same animals are then fasted again 1 week later, and a calibrated meal (normally chow diet) is provided for 1 hr in the form of a specific weighed amount of food pellets. Alternatively, a standardized complex nutrient solution (e.g., Ensure Plus, Abbott Nutrition) can be administered in the beginning of the feeding period by oral gavage. A post-prandial blood sample is collected immediately after the 1-hr feeding period.

An alternative to collecting blood on separate days for the fasting and fed states is to use, instead of retroorbital venipuncture under anesthesia, blood sampling from the tail vein (UNIT 29A.3) before (fasting state) and after (fed state) the feeding period (see Basic Protocol 3). However, since stress may alter the feeding behavior of mice, this alternative is mostly useful with the administration of nutrient solution rather than with free pellet feeding.

Materials

Mice

Isoflurane inhalation anesthesia unit (composed of a mixing and flow-control system for O₂/isoflurane anesthesia; e.g., TEM, <http://www.TEM.fr>)

Non-heparinized glass capillaries, 50- to 100- μ l (e.g., Vitrex by Modulohm)

Mouse chow

Mouse cages

Refrigerated microcentrifuge

Additional reagents and equipment for measuring blood glucose (see Basic Protocol 1) and insulin (see Basic Protocol 2) levels

1. Fast mice for 14 to 16 hr (overnight). Ensure that they have access to drinking water at all times.
2. On the following day, before 8 a.m., prepare experiment record sheet, tubes for blood collection, and strips for glucose measurement. Calibrate the blood glucose monitor.
3. At 8 a.m., anesthetize a mouse lightly by isoflurane (see *UNIT 29B.1*, Basic Protocol 1) and take a 50- to 100- μ l blood sample by retroorbital venipuncture (see Hedrich, 2004) into a non-heparinized capillary. Drop a small amount (~ 3 μ l) of blood onto a test strip placed in the blood glucose monitor (see Basic Protocol 1), and collect the rest of the 50- to 100- μ l sample in a non-heparinized capillary on ice. Record the glucose measurement from the monitor.
4. Microcentrifuge the blood samples 15 min at $2300 \times g$ (5000 rpm), 4°C. Collect the plasma (~ 20 to 40 μ l) into clean tubes, and store at -20° or -80°C until glucose and insulin measurement (up to 1 month for glucose and 6 months for insulin).
5. Repeat the fasting 1 week later. On the following day, at 8 a.m., place mice individually in fresh cages with water and a preweighed meal (approximately three pellets of chow diet) for 60 min.
6. Prepare for blood sampling and glucose measurement as in step 2.
7. After the 60-min meal, weigh the remaining food.
8. Take blood samples as described in steps 3 and 4.
9. Perform glucose and insulin measurements according to Basic Protocols 1 and 2.
10. For data analysis, divide the increase in glucose and insulin values over fasting levels for each mouse by the weight of consumed food in grams.

Thus, results are reported as mg/dl \cdot g⁻¹ for glucose and ng/ml \cdot g⁻¹ for insulin. Sometimes, it is also useful to express the results as per calories consumed.

ORAL GLUCOSE TOLERANCE TEST

The oral glucose tolerance test (OGTT) measures the clearance of a standardized glucose load from the body. Since the glucose bolus enters the body via the natural route—it is absorbed from the intestinal tract—this test also takes into account intestinal aspects of glucose absorption. Animals undergo fasting for ~ 14 to 16 hr, a glucose solution is administered by oral gavage (Hedrich, 2004), and blood glucose is measured at different time points in a period of 3 hr.

**BASIC
PROTOCOL 5**

**Mouse
Phenotyping**

29B.3.7

An important consideration for this procedure is the size of the cohort to be assessed at any one time. For two trained persons, a maximum of sixteen animals is recommended. To minimize variation due to technical aspects, each person should be designated a role in the experiment, i.e., only one person/experiment gavages all mice. Therefore, careful planning is needed when dividing an experimental cohort of more than sixteen of animals.

Materials

Mice

20% (w/v) aqueous glucose solution (see recipe), sterile

Mouse cages

Animal scale

Blood glucose monitor and associated glucose test strips for glucose measurement (e.g., Accu-Chek Active, Roche Diagnostics)

1-ml syringes (e.g., Terumo) and 22-G ball-tip needle (e.g., Popper and Sons)

Scalpel blade

Timer

1. Fast mice for 14 to 16 hr (overnight). Ensure that they have access to drinking water at all times.
2. On the following day, at 8 a.m., place mice individually in fresh cages with only water (no food). Identify cages with the mouse number.
3. Record the weight of each mouse to within 0.1 g.
4. Calculate and record the volume of 20% glucose solution required for oral gavage of 2 g glucose/kg at an oral injection volume of 10 μ l/g body weight.

For a mouse with a body weight of 30 g, inject 300 μ l of glucose solution.

5. Prepare experiment record sheet, strips for glucose measurement, and a 1-ml syringe for each animal containing the calculated volume of glucose to be gavaged.

Unlike for intraperitoneal tests (see Basic Protocols 6 and 7), it is practical and more economic to use just one ball-tip needle, which is switched from syringe to syringe as needed, with brief external rinsing in 70% ethanol and distilled water between uses.

6. Calibrate the blood glucose monitor with the standard strip.
7. Measure the basal glucose concentration ($T = 0$) in each mouse by removing one mouse at a time from its cage, placing it on the top of its cage, and making a small incision over the lateral tail vein (1 to 2 cm from the tail tip) using a scalpel blade (see UNIT 29A.3). Touch a small (~ 3 μ l) blood sample directly onto the test strip placed in the blood glucose monitor (see Basic Protocol 1). Apply direct pressure to the incision until the blood clots and return the mouse to its cage.
8. After all mice have been measured for basal glucose concentrations, administer the glucose solution by oral gavage (Hedrich, 2004) to each animal at a 30- or 60-sec interval between animals, depending on the size of the experimental cohort. Start the timer with the first mouse gavaged.

In general, it takes a minimum of 30 sec for an experienced person to perform the gavage and/or take a blood measurement. All injections must be finished within 15 min in order to perform the next blood glucose measurement at $T = 15$ min.

9. At $T = 15$ min, measure the blood glucose again, starting with the first mouse gavaged and using the same time interval used for gavage, until all the mice in the

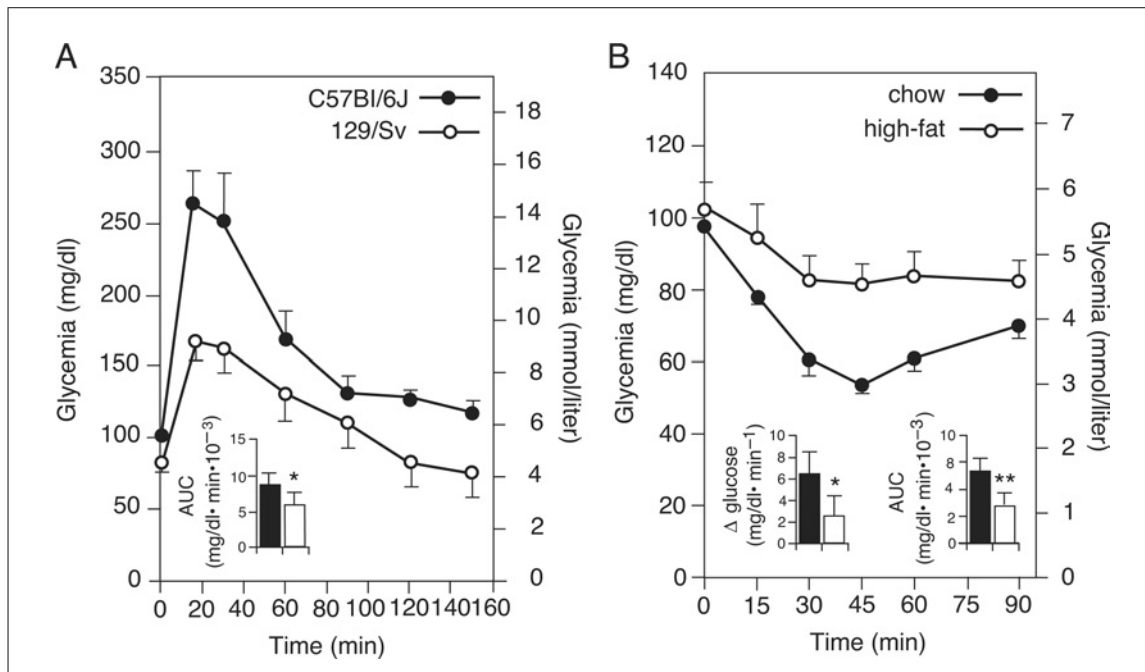


Figure 29B.3.1 Examples of the intraperitoneal insulin sensitivity test (IPIST) and intraperitoneal glucose tolerance test (IPGTT). Note that the figures represent different sets of mice. **(A)** An IPGTT showing a difference in glucose tolerance between wild-type C57Bl/6J and 129/Sv mice, the former being more insulin resistant. **(B)** An IPIST on chow-fed and high-fat-fed mice, showing decreased insulin sensitivity in the high-fat-fed group.

cohort have been measured. It is important that timing be followed as closely as possible. To restart the bleeding, remove the clot from the first incision and massage the tail if blood flow is inadequate.

10. Repeat at $T = 30, 60, 90, 120, 150,$ and 180 min after gavage.
11. At the end of the experiment, return the mice to their original cages and provide plenty of food and water. Ensure that no animal is bleeding excessively.
12. For visual representation, plot glucose values versus time (see Fig. 29B.3.1A for an example of a graph from on IPGTT).
13. For statistical analysis, calculate the area under the curve (AUC) for each mouse using the equation:

$$\text{AUC} = \sum_{n=1}^{x-1} \left((T_{n+1} - T_n) \times \left(\frac{|\text{Glc}_{T_n} - \text{Glc}_{T_{n+1}}|}{2} + \min(\text{Glc}_{T_n}, \text{Glc}_{T_{n+1}}) - \text{Glc}_{T_0} \right) \right)$$

where Glc is glucose concentration, T is time, n is the time point number, x is the total number of time points, and $\min(\text{Glc}_{T_n}, \text{Glc}_{T_{n+1}})$ is the lower of two consecutive glucose concentrations.

The rationale for this calculation is depicted in Figure 29B.3.2.

To exclude the effect of differences in fasting glucose levels between animals, AUC is usually calculated as incremental area ('above baseline'), where only the area above the fasting level for each mouse is taken into account.

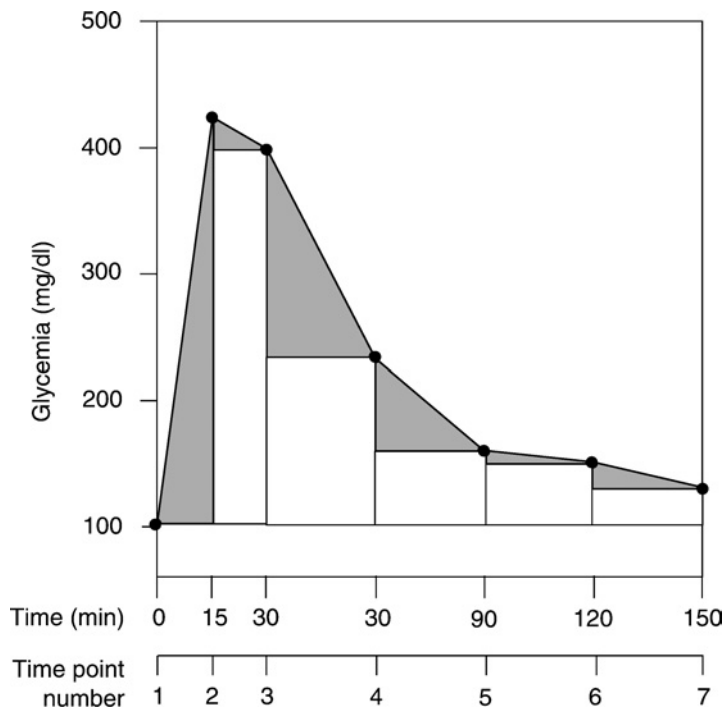


Figure 29B.3.2 Rationale for calculating the area under the curve (AUC). Areas of the triangles (gray), delimited by two consecutive time points and glucose values therein, and quadrangles (white), below each triangle but above baseline value, are summed.

**BASIC
PROTOCOL 6**

INTRAPERITONEAL INSULIN SENSITIVITY TEST

The intraperitoneal insulin sensitivity test (IPIST) measures glucose levels subsequent to a standardized insulin load. It gives an estimate of the insulin sensitivity of the animals. Animals are fasted for 14 to 16 hr, a bolus of insulin is administered intraperitoneally (i.p.), and blood glucose is measured at different time points during 90 min.

An important consideration for this procedure is the size of cohort to be assessed at any one time. For two trained persons, a maximum of sixteen animals is recommended. To minimize variation due to technical aspects, each person should be designated a role in the experiment, i.e., only one person/experiment administers insulin to all mice. Therefore, careful planning is needed when dividing an experimental cohort of more than sixteen animals.

Materials

- Mice
- Insulin solution for IPIST (see recipe)
- Mouse cages
- Animal scale
- 1-ml syringes (e.g., Terumo) and 25-G × 5/8-in. needles (e.g., Terumo)
- Blood glucose monitor and associated glucose test strips for glucose measurement (e.g., Accu-Chek Active, Roche Diagnostics)
- Scalpel blade
- Timer

1. Fast mice for 14 to 16 hr (overnight). Ensure that the animals have access to drinking water at all times.

2. On the following day, at 8 a.m., place mice individually in fresh cages with only water (no food). Identify cages with the mouse number.
3. Record the weight of each mouse to within 0.1 g.
4. Calculate and record the volume of insulin working solution required for an i.p. injection of 0.5 IU/kg in an injection volume of 3.6 $\mu\text{l/g}$ body weight.

For a mouse with a body weight of 30 g, inject 108 μl of insulin working solution.

5. Prepare the 1-ml insulin syringes and 25-G \times 5/8-in. needles with the calculated volumes for each animal in the experiment. Prepare experiment record sheet and test strips for glucose measurement.
6. Calibrate the blood glucose monitor with the standard strip.
7. Measure the basal glucose concentration ($T = 0$) in each mouse by removing one mouse at a time from its cage, placing it on the top of its cage, and making a small incision over the lateral tail vein (1 to 2 cm from the tail tip) using a scalpel blade (see UNIT 29A.3). Touch a small ($\sim 3 \mu\text{l}$) blood sample directly onto the test strip placed in the blood glucose monitor (see Basic Protocol 1). Apply direct pressure to the incision until the blood clots and return the mouse to its cage.
8. After all mice have been measured for basal glucose concentrations, administer the insulin solution by i.p. injection to each animal at a 30- or 60-sec interval between animals, depending on the size of the experimental cohort. Start the timer with the first mouse injected.

In general, it takes a minimum of 30 sec for an experienced person to perform the injection and/or take a blood measurement. All injections must be finished within 15 min in order to perform the next blood glucose measurement at $T = 15$ min.

9. At $T = 15$ min, measure the blood glucose again, starting with the first mouse injected and using the same time interval used for injections, until all the mice in the cohort have been measured. It is important that timing be followed as closely as possible. To restart the bleeding, remove the clot from the first incision and massage the tail if blood flow is inadequate.
10. Repeat at $T = 30, 60,$ and 90 min after insulin injection.

CAUTION: If at any time the blood glucose falls below ~ 36 mg/dl, the value should be confirmed. If the value is accurate, the mouse should be rescued by an injection of 0.5 to 1.0 g glucose/kg from a 20% aqueous glucose solution (see Basic Protocol 5), as very low values can be lethal to the animal.

11. At the end of the experiment, return the mice to their original cages and provide plenty of food and water. Ensure that no animal is bleeding excessively.
12. For visual representation, plot absolute glucose values versus time (Fig. 29B.3.1B).

In cases where groups differ strongly in FPG values, plotting the percent of fasting value versus time can be a viable alternative.

13. For statistical analysis, calculate the rate of decline in glucose between T_0 and T_{15} using log-transformed absolute glucose values. Also, calculate the inverse area above the curve (AUC_i) below baseline, again using absolute values, for each mouse as follows:

$$\text{AUC}_i = \sum_{n=1}^{x-1} \left((T_{n+1} - T_n) \times \left(\frac{|\text{Glc}_{T_n} - \text{Glc}_{T_{n+1}}|}{2} + \text{Glc}_{T_0} - \max(\text{Glc}_{T_n}, \text{Glc}_{T_{n+1}}) \right) \right)$$

where Glc is glucose concentration, T is time, n is the time point number, x is the total number of time points, and $\max(\text{Glc}_{T_n}, \text{Glc}_{T_{n+1}})$ is the higher of two consecutive glucose concentrations.

In some cases, it may be useful to calculate AUC_i separately for the last 85 or 60 min of the test since, after the initial direct insulin effect, counter-regulatory hormones and gluconeogenesis will start playing a role. Thus, excluding the initial phase may provide additional phenotypic information.

**BASIC
PROTOCOL 7**

INTRAPERITONEAL GLUCOSE TOLERANCE TEST

The intraperitoneal glucose tolerance test (IPGTT) measures the clearance of an intraperitoneally injected glucose load into peripheral tissues. Compared to the oral glucose tolerance test (OGTT; see Basic Protocol 5), the IPGTT does not address mechanisms participating in glucose absorption from the intestine and hence can be considered a more “artificial” test than the OGTT. This procedure is performed exactly as the OGTT, except that the 20% aqueous glucose solution is administered by i.p. injection using a 25-G \times 5/8-in. needle (e.g., Terumo) and the syringes, containing the calculated volume of glucose, should be prepared for each animal with the needles already attached and voided of dead volume. Similarly to oral administration, a 2 g glucose/kg dose is used, with an i.p. injection volume of 10 μ l/g body weight. Considerations for careful timing and cohort size apply here as well.

**BASIC
PROTOCOL 8**

HYPERINSULINEMIC-EUGLYCEMIC CLAMP

The current gold standard method for measuring IR is the euglycemic clamp, described in this protocol. In this method, glucose is “clamped” at a predetermined value (100 mg/dl or 5.5 mmol/liter for euglycemia) by titrating a variable rate of glucose infusion (glucose infusion rate or GIR) against a fixed rate of insulin infusion (resulting in hyperinsulinemia). Mice are permanently catheterized using a bifurcated catheter assembly to allow for simultaneous constant and variable infusion of insulin and glucose, respectively. The clamps are performed in awake, unrestrained, unstressed, and light-cycle-inverted mice following a 5-hr fast. Mice are acclimatized for at least 30 min to the tops of cages while their catheters are attached to a syringe-infusion pump. Baseline glucose values are measured by tail vein sampling prior to the injection of insulin. Catheter placement is assessed with a short priming dose (6 μ l/min, 1 min) of insulin prior to the constant infusion of insulin at a flow rate of 2 μ l/min equivalent to 18 mU of insulin/kg/min. Blood glucose values are monitored every 5 to 10 min throughout the test. Within the first 15 min, a decrease in blood glucose should be observed, at which point a 20% glucose (in saline) infusion is started. The GIR is varied as necessary until euglycemia ($\pm 15\%$) has been reached and maintained. The animal is “clamped” at this level, and the degree of IR is inversely related to the amount of glucose necessary to maintain the required blood glucose concentrations. The GIR at clamp (mg glucose/kg/min) is calculated as an average during the last 60 min of the experiment. When the average GIR of one animal is greater than another, it indicates better insulin sensitivity or that the clearance of glucose from the plasma is much faster.

Modifications to several parameters of this test—including the duration of fasting, the amount of insulin given at the initial infusion phase, and the inclusion of tracers—may be necessary, depending on the specific aspect of glucose homeostasis that is being studied. A validation of some of these parameters was recently summarized (Ayala et al., 2006).

The details of this procedure are based on assaying six animals at a time by two experienced experimenters. The number of mice will vary depending on the number of people

involved and the size of the experimental cohort required to obtain statistically relevant results. In general, a minimum of six animals per experimental group is required.

NOTE: Catheter surgery should be performed by an experienced veterinarian or other mouse care personnel competent in performing this procedure.

Materials

Mice

Anesthesia: ketamine and xylazine

Insulin solution for euglycemic clamp (see recipe)

20% (w/v) glucose in saline solution (see recipe), sterile

Saline: 0.9% (w/v) NaCl

70% (v/v) ethanol

Mouse femoral catheter (e.g., M-FAC, Braintree Scientific)

Mouse cages with flat-grated lid

Animal scale

1- and 2-ml syringes

Single syringe infusion pumps capable of holding a 2-ml syringe with a range of infusion rates of 5 to 20 $\mu\text{l}/\text{min}$ (e.g., Harvard Apparatus or TSE Systems)

Multi-syringe infusion pump capable of holding a minimum of six individual 1-ml syringes with a range of infusion rates of 3 to 10 $\mu\text{l}/\text{min}$ (e.g., Harvard Apparatus or TSE Systems)

Tubing (e.g., Tygon tubing, \O 0.51 mm, Biorad)

Tubing connectors (e.g., 23-G needles, Terumo)

Y-shaped canula connector (e.g., C312ICP/O, Protech International)

Retort stand with a movable arm to hold the Y-shaped canula connector and tubing in place

Timer

Blood glucose monitors and associated glucose test strips for glucose measurement (e.g., Accu-Chek Active, Roche Diagnostics)

Scalpel blade

Prepare clamp set-up

Prior to the catheter surgery and clamp, mice should be acclimatized to a light-inverted cycle for at least 1 week.

1. At least 2 to 3 days before the experiment, place a catheter in the femoral vein of all mice (under anesthesia using ketamine and xylazine), with the catheter fed underneath the skin and affixed behind the head.

The catheter should not require flushing prior to the experiment if it has been filled with saline and closed the day of the surgery.

2. After surgery, house mice individually and allow to recover for at least 48 hr, preferably long enough for them to regain their pre-operative body weight.
3. On the day of the clamp, remove the food from the mice 5 hr prior to the assay. Maintain access to water.
4. A minimum of 30 min prior to beginning the clamp, weigh the mice to within 0.1 g and place them on top of their cages to acclimatize.
5. Prepare the insulin working solution and fill six 1-ml syringes. Fill six 2-ml syringes with 20% glucose solution.
6. Place all infusion pumps on a shelf above the mice on their cages and insert the six glucose syringes vertically in the six single-infusion pumps. Attach syringes with appropriate tubing and connectors and make the connections from the pump to the

Y-shaped canula connector on the retort stand. Do not make the connection to the animal at this point. Start glucose infusion pumps at a high rate of infusion to fill the line with glucose and look for bubbles or plugs. Ensure that glucose is coming out of the end of the line at a rate comparable to the setting on the infusion pump. Ensure that the mice are not exposed to drips of glucose during the priming of the system.

7. While maintaining the glucose connections, place the six insulin syringes vertically in the multi-syringe infusion pump and connect the insulin syringes with tubing to the Y-shaped canula connector. Again, purge the lines with insulin by turning the infusion pumps on at a high rate. Ensure that insulin fills the lines and there are no bubbles or plugs.

For the general setup, see Figure 29B.3.3.

8. Stop all infusion pumps. Reset the glucose pumps to 5 $\mu\text{l}/\text{min}$ and the insulin pump to 2 $\mu\text{l}/\text{min}$.

A flow rate of 2 $\mu\text{l}/\text{min}$ is equivalent to 18 mU of insulin/kg/min based on an average mouse mass of 20 g. This amount of insulin is enough to obtain hyperinsulinemic conditions in a mouse. In general, this infusion rate of insulin is used even if the mass of each mouse is different so that only one pump is required (the validity of this approach is open for discussion).

9. Connect the infusion tubes to the mice. Ensure that the tubing is not twisted during the experiment and the mice are free to roam the top of the cage.

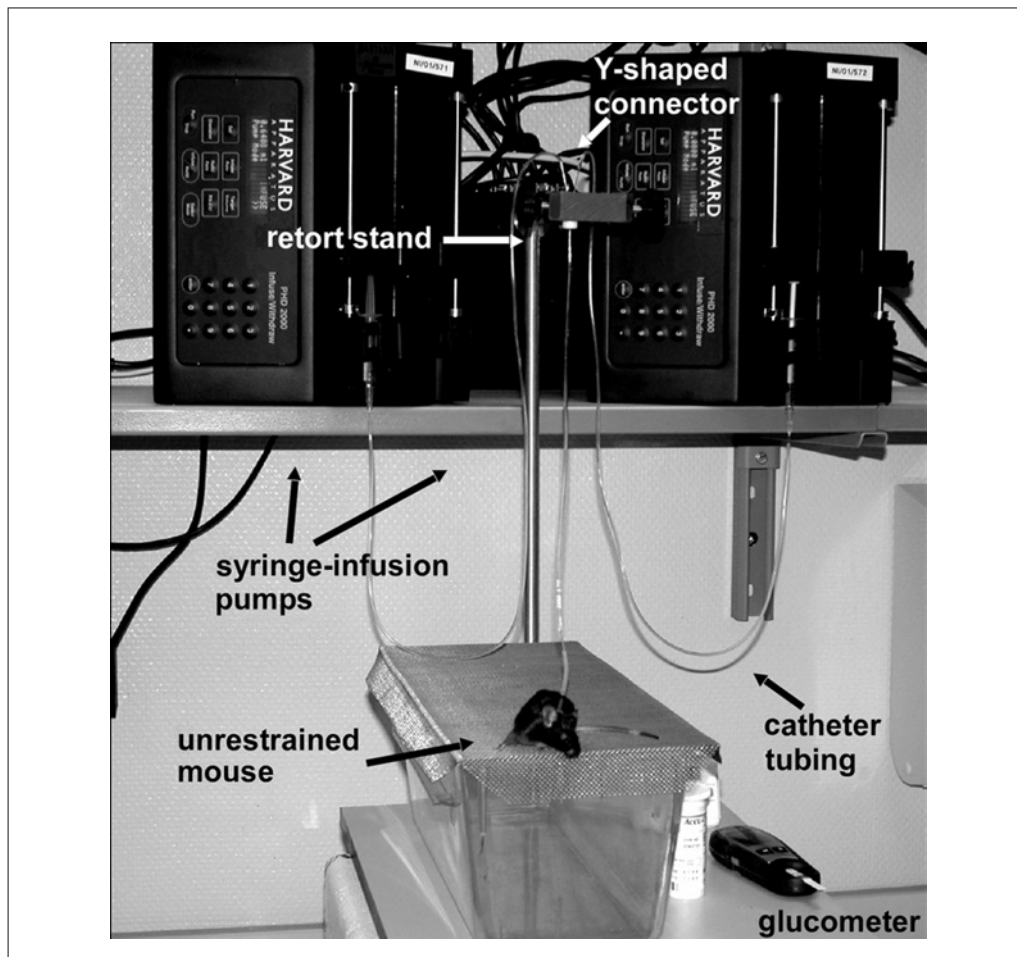


Figure 29B.3.3 An example of a set-up for a euglycemic clamp showing a single catheterized mouse.

Perform infusions and blood sampling

10. After 30 min of acclimatization, cut the tip of the mouse tail with a scalpel blade. Measure a small sample of blood with the blood glucose monitor to obtain the basal ($T = 0$) glucose value.

This value should be near euglycemia or 100 mg/dl (5.5 mmol/liter).

11. Begin the timer and the insulin infusion pump, starting at an infusion rate of 2 $\mu\text{l}/\text{min}$. After 3 min, increase the insulin infusion rate to 6 $\mu\text{l}/\text{min}$ and maintain for 1 min, and then return to 2 $\mu\text{l}/\text{min}$ for the remaining clamp (180 min). Measure the plasma glucose value at 5-min intervals.

The glucose should be lower at ~ 15 min than at time 0.

12. When the glucose is around euglycemia, begin the glucose infusion pumps, starting at an infusion rate of 5 $\mu\text{l}/\text{min}$.

It is not necessary for all animals to start glucose infusion at the same time.

13. Continue to monitor plasma glucose every 5 min and increase or decrease the individual glucose infusion rates until euglycemia has been reached and maintained. Adjust glucose infusion rates in increments of 1 to 5 $\mu\text{l}/\text{min}$ in the beginning, and then decrease the increments as levels approach euglycemia to ensure that ‘overshooting’ does not occur.

An acceptable range of plasma glucose is within $\pm 15\%$ of 100 mg/dl (5.5 mmol/liter), i.e., from 85 to 115 mg/dl (4.68 to 6.33 mmol/liter). In general, the mice should reach stable euglycemia within 60 min. From this point on, the glucose infusion rate should not change much, and the animal is thus clamped.

14. Once euglycemia is reached, continue infusing glucose and insulin until $T = 180$ min, performing glucose measurements every 10 min instead of every 5 min.

15. At the end of the experiment, disconnect the mice and inject ~ 100 μl of saline into the catheters. Disconnect the glucose and insulin lines, purge them with water, and allow them to dry.

16. Place the mice in clean cages, provide food and water, and ensure that the mice recover from this stressful experiment.

When the mice are finally euthanized, canulae can be recovered and purged with 70% ethanol for reuse.

Analyze data

17. For qualitative purposes, draw a curve that shows the progression of the clamp, i.e., plasma glucose versus time.

This is especially useful for visualizing the stabilization at euglycemia during the last 60 min of the clamp.

18. Calculate GIR (mg glucose/kg/min) for each time point where glucose was sampled using the following equation:

$$\text{GIR} = \frac{\text{Glc}}{m} \times \text{gir}$$

where Glc is glucose concentration (i.e., 20 mg/100 μl or 20%), m is animal weight (kg), and gir is the glucose infusion rate ($\mu\text{l}/\text{min}$).

If the gir was 5 $\mu\text{l}/\text{min}$ and the mouse was 20 g (0.02 kg), then $\text{GIR} = (20 \text{ mg}/100 \mu\text{l}) \div (0.02 \text{ kg}) \times (5 \mu\text{l}/\text{min}) = 50 \text{ mg}/\text{kg}/\text{min}$.

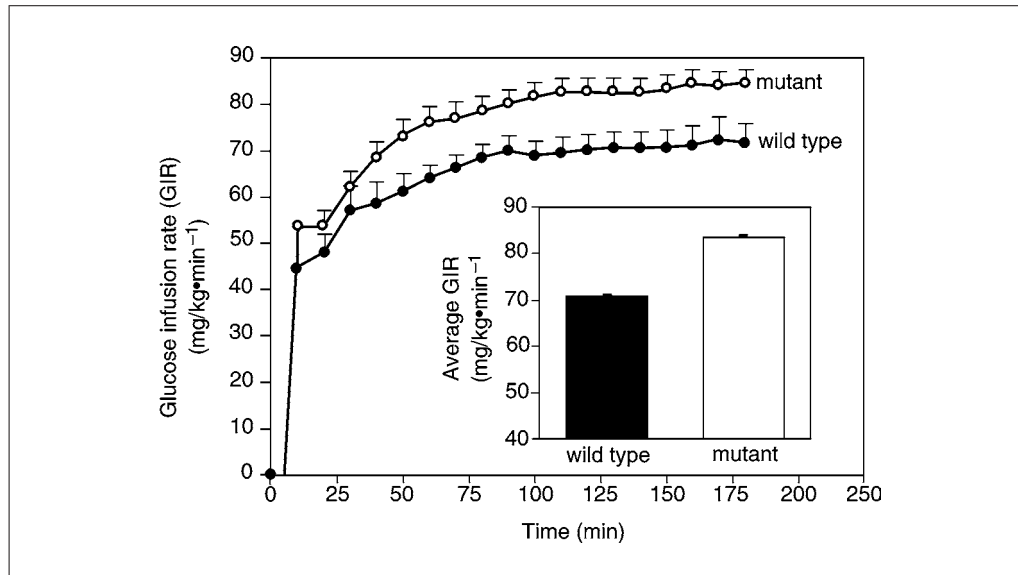


Figure 29B.3.4 Hyperinsulinemic-euglycemic clamp in wild-type and mutant mice. The evolution of glucose infusion rates (GIR), and the average thereof, show that the mutant mice are more insulin sensitive, being able to clear the infused glucose into peripheral tissues more rapidly.

19. Plot GIR versus time (see Fig. 29B.3.4).

The final result represents the GIR averaged over the last 60 min of the clamp, i.e., from 120 to 180 min, and is often depicted as a bar graph (inset in Fig. 29B.3.4).

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Glucose solution, 20% (w/v)

Dissolve 20 g D-glucose in 100 ml distilled water or saline. Filter sterilize using a 0.45- μ m filter. Store up to 1 week at 4°C or 6 months at -20°C.

Insulin solution for euglycemic clamp

Diluent stock solution: Dissolve 50 mg fatty acid-free bovine serum albumin (0.5% BSA final) in 9.9 ml sterile water, and add 100 μ l of 1 M HCl (10 mM HCl final). Filter sterilize using a 0.45- μ m filter. Prepare fresh, or store in 1.5-ml aliquots up to 6 months at -20°C.

Working solution: Add 1.15 ml stock solution into 6.048 ml saline. Add 20.1 μ l of 100 IU/ml Humulin insulin (Eli Lilly; 0.278 IU/ml final). Prepare immediately before use.

Insulin solution for IPIST

Diluent stock solution: Dissolve 50 mg fatty acid-free bovine serum albumin (0.5% BSA final) in 10 ml sterile water, and add 10 μ l of 1 M HCl (1 mM HCl final). Filter sterilize using a 0.45- μ m filter. Prepare fresh, or store in 1.5-ml aliquots up to 6 months at -20°C.

Working solution: Add 1.15 ml stock solution into 6.049 ml saline. Add 10 μ l of 100 IU/ml Humulin insulin (Eli Lilly; 0.139 IU/ml final). Prepare immediately before use.

COMMENTARY

Background Information

A principle hormone in metabolic regulation is insulin, which acts to lower blood glucose by stimulating glucose uptake by skeletal muscle and fat and other tissues, and by inhibiting glucose production by the liver. Post-prandial surges and inter-prandial declines of insulin levels tightly control glucose levels such that both hypoglycemia and hyperglycemia, which are intolerable for most tissues, are prevented. One key to glucose control is the ability of pancreatic β -cells to react quickly to changes in blood glucose by altering insulin secretion.

The motivation for investigating insulin and glucose signaling in the mouse is related to the epidemic status of T2DM. The pathophysiological events leading to T2DM in the context of the metabolic syndrome in humans include β -cell failure and resistance of target tissues to insulin. When tissues become insulin resistant, organs cannot respond properly and hyperglycemia prevails. One molecular mechanism thought to contribute to insulin resistance in obese individuals is the increased availability of free fatty acids (FFA) due both to reduced capacity to store the FFA in adipose tissue and to increased lipolysis. In the muscle, increased FFA oxidation leads to reduced glucose oxidative utilization in part through the inhibition of enzymes like pyruvate dehydrogenase, required for the oxidation of glucose, and phosphofructokinase (PFK), a key glycolytic enzyme. Glycogen synthesis is also inhibited by elevated FFA levels, thereby reducing non-oxidative glucose utilization. In the liver, FFA affects the balance between glycolysis and gluconeogenesis. This is attributed to

the tendency of FFA to increase glucose production by enhancing the expression of the key gluconeogenic enzyme phosphoenolpyruvate carboxykinase. The pancreatic β -cells initially try to compensate for the reduced insulin sensitivity by inducing insulin secretion further. Although the resulting hyperinsulinemia is a short-term solution, in the long run, the vicious cycle leads to β -cell dysfunction and severe peripheral insulin resistance (Fig. 29B.3.5).

Thus, there are several indicators of insulin resistance, including elevated fasting glucose and insulin levels and impaired glucose tolerance. These pathophysiological phenomena can be detected in the mouse using similar methods used in humans, including measurements of fasting glucose and insulin levels, glucose/insulin sensitivity tests, and more invasive clamps. After measuring circulating insulin and glucose levels at steady state, the most common test is to administer a glucose load and follow the homeostatic response, which is the culmination of an increase in insulin followed by insulin-dependent processes that lower glucose. Glucose tolerance then becomes an expression of the efficiency with which homeostatic mechanisms restore glycemia to basal levels after perturbation. Finally, the most technically challenging test is the glucose clamp, which is considered the gold standard because its interpretation is based on the least number of assumptions since glycemia is fixed under hyperinsulinemic conditions. The literature shows that correlations between these diverse measures of insulin sensitivity and glucose tolerance are often quite good.

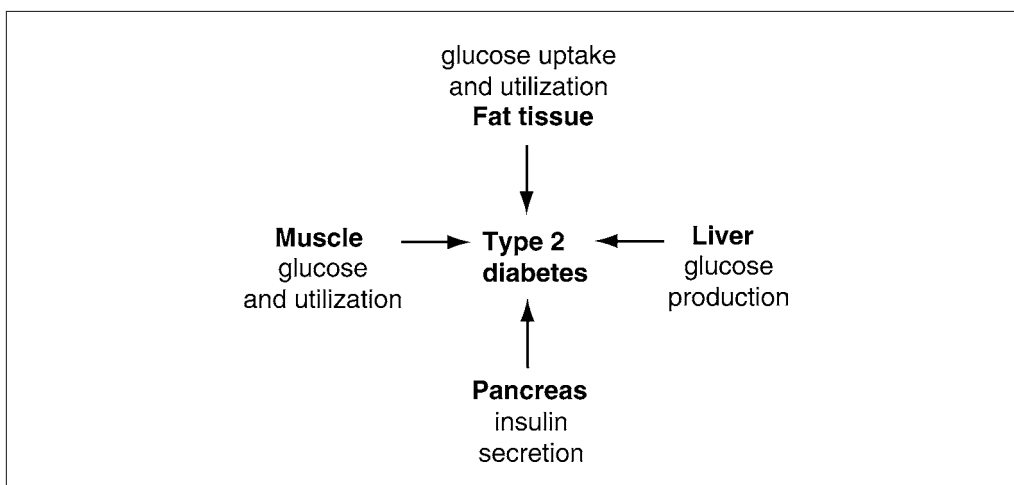


Figure 29B.3.5 Major effects of insulin resistance in different tissues, contributing to the development of type 2 diabetes mellitus (T2DM).

Critical Parameters

The validity of the results obtained from metabolic exploration phenotyping is largely dependent on methods of animal husbandry. To reduce the anxiety levels of the mice prior to and during testing, it is of vital importance that individuals performing these procedures are experienced, aware of the welfare of the mice, and familiar with the mice being tested. General laboratory procedures, and those specifically applicable to animal experimentation, should be followed throughout the protocols.

The majority of mouse metabolic studies are age, sex, and/or strain dependent (Champy et al., 2004). Therefore, it is important to keep these parameters consistent throughout a single experiment. It is also very important that all phenotyping experimentation be conducted at approximately the same time of day because physiological and biochemical parameters change throughout the day. Fasting should be done under conditions where the mice truly cannot eat, e.g., there are no fallen pieces of food pellets within the bedding material.

In tests where glucose is first administered and then measured in the blood, care should be taken not to contaminate any surfaces or materials with the glucose solution itself, as even a minute carryover into the blood samples will result in falsely high glucose measurements. Glucose syringes should be prepared cleanly and in advance, and any spillover or leakage from the injection site during the test should be carefully removed. Changing gloves after any suspected glucose contamination is recommended.

Since the changes in blood glucose (and insulin) concentrations during the tolerance tests are dynamic in nature, the number of mice that will be analyzed in any one experiment must be small enough that the sampling timetable can be precisely kept, especially during the critical initial phases where the most rapid changes are usually observed (typically T_0 to T_{30}). Although increasing the number of experimenters would facilitate the use of more mice in one experiment, for consistent results it would be best if one person obtains all blood samples. Using individual cages will aid in keeping the timepoints accurate by facilitating ease of mouse identification. Another practical method is to identify group-housed mice by marking their tails at the beginning of the test with a felt-tip pen.

Although using the composite measure AUC (or AUC_i) circumvents some of the statistical problems inherent with non-independent repeated measurements, it is still

an approximation based on a fairly limited sampling of the dynamic state. Additional timepoints at the initial phase of the tests would make AUC more accurate, but this may lead to overbleeding the mice, or compromise timely sampling. Alternate methods for calculating the AUC are available, potentially giving more precise approximation from the same number of data points, but these require sophisticated mathematical methods and specialized software beyond the scope of this unit.

Glucose measurement by blood glucose monitor. Although blood glucose monitors provide accurate readings in most situations, it is advisable to use alternate methods to confirm very low readings (<36 mg/dl or 2 mM), which are possible during IPIST or initial phases of the euglycemic clamp, and especially very high values (reported upper limit for Accu-Check glucose monitor is 600 mg/dl or 33.3 mM), which are most likely to occur during peak phase of IPGTT. In addition, some mouse models for IR or T2DM such as KK-Ay mice have high enough glucose levels to very likely require alternate methods for blood glucose analysis.

Insulin measurement by ELISA. Since the overnight fasting insulin levels of the average mouse (see Anticipated Results) are relatively close to the sensitivity limit of even the ultrasensitive insulin ELISA kits (~ 0.2 $\mu\text{g/liter}$ for the Mercodia kit), it is important to pay attention to proper sample storage conditions, minimization of freeze-thaw cycles of the samples, and good washing after the antibody incubation to obtain as many of the samples above the detection limit as possible. On the other hand, especially in some high-fat-fed diabetic mouse models, glucose-stimulated insulin levels may exceed detection limits (~ 7.0 $\mu\text{g/liter}$ for the Mercodia kit), requiring dilution and re-analysis of such samples.

Oral glucose tolerance test. Successful oral gavaging is of vital importance to this test, and should be performed by experienced personnel to minimize animal loss due to gavaging the lungs rather than the esophagus.

Meal tolerance test. The mice should not be stressed before the feeding when free pellet feeding is used to avoid modifications in feeding behavior. For example, removing the bedding material or replacing it with paper towels is not advisable, even though accurate measurement of the food remaining at the end of the feeding period may be more difficult, since small pieces of food pellets may be missed within the bedding material.

Intraperitoneal insulin sensitivity and glucose tolerance tests. It is important to perform intraperitoneal injections in a manner that avoids leakage from the injection site. Another concern, especially with obese mice, is that the injected solution may end up as an inclusion within the large amount of adipose tissue, seriously slowing down the absorption of the substance and the metabolic effects thereof. Since ways to verify this are limited, especially if further experimentation on the same mice is desired, care should be taken at the data analysis phase to exclude unreactive mice. This will be particularly difficult if the groups of mice being compared have a significantly different amount of adipose tissue in the abdominal area, or if there are experimental groups that seem to be less reactive due to physiological rather than technical reasons. Thus, solid mouse handling and injection techniques are essential for success in these tests.

For some mouse models with increased insulin sensitivity, the insulin dose suggested in the protocol may be too high, resulting in severe hypoglycemia. Such mice must either be treated with an immediate injection of ~ 0.5 mg/g body weight glucose or be euthanized. A preliminary test with a small number of mice should be performed to verify the dose-responsiveness of the mice under study. On the other hand, in some experimental setups like after high-fat feeding, increasing the insulin dose may become a necessity. From a technical standpoint, it is also important that only freshly made insulin working solutions be used. This will prevent the insulin in the dilute solution from degrading or binding excessively to the walls of the tubes and syringes. Prior to each IPIST, it is recommended that the insulin first be tested on one additional mouse to verify that a fall in glucose is achieved at the 15-min timepoint.

Euglycemic clamp. Successful catheter placement is vital in this test, and the initial bolus injection of insulin serves to test this. Decreased glucose levels within 15 to 20 min indicate successful catheterization. In addition, care should be taken throughout the entire procedure to ensure that the external catheters are not blocked by tiny air bubbles or improper connections. If bubbles do arise, eliminating them by reconnections may be attempted, but experimentation on the affected mice should be stopped if the blockage is not cleared quickly.

The timing for blood sampling (one every 5 min) is to be used only as a guide. For example, when blood glucose concentration initially starts to fall, more frequent measurements can be taken to ensure that the animal does not become too hypoglycemic before glucose infusion can rescue it; this may lead to the loss of the animal. Similarly, once the animal is clamped and the plasma glucose is obviously stable, the measurements can become less frequent, e.g., one every 10 min. Care should be taken to minimize sample blood loss from the tail bleed to prevent an excessive fall in hematocrit.

It is also important to achieve and maintain euglycemia during the last hour of the test. In non-euglycemic conditions, changes in hepatic glucose output may interfere with the interpretation of the clamp data. In some cases, insulin infusion can be administered per unit of body surface area to avoid overinsulinization of obese mice.

Because of the stressful nature of the clamp experiment, further experimentation in the animals is not recommended, and the mice should preferably be euthanized. If other procedures such as tissue collection are vital, animals should be allowed to recover for at least 1 week.

Anticipated Results

Glucose and insulin measurements. Reference values for fasting plasma glucose and insulin values for a large cohort of C57B6/J mice are presented in Table 29B.3.2. However, depending on prior treatment (e.g., fasting, diet challenges, exercise), in combination with possible genetic modifications, blood glucose and especially insulin values can vary dramatically, particularly upwards. For example, mice lacking insulin receptor specifically in the liver have over 170 times more circulating insulin in the fed state than wild-type mice in the fasting state (Michael et al., 2000). Stimulation during glucose or insulin tolerance tests leads to an even wider range of variation. In models where IR is being induced, increased insulin is usually the first observed effect, and hyperglycemia may or may not follow. As such, elevated fasting glucose level indicates basal IR, which is most often attributed to impaired suppression of hepatic glucose output (gluconeogenesis and glycogen breakdown) by insulin, especially in cases where an increase in glucose is accompanied by an increase in

Table 29B.3.2 Fasting Plasma Glucose and Insulin Values in a Large Cohort of Female and Male C57Bl/6J Mice

	Mean \pm SEM	N
<i>Glucose (mg/dl)</i>		
Female	115.5 \pm 2.7	116
Male	123.6 \pm 2.5	114
<i>Insulin (μg/liter)</i>		
Female	0.798 \pm 0.120	70
Male	0.358 \pm 0.072	78

insulin. On the other hand, low or normal insulin and high glucose levels in a fasting state is often indicative of advanced IR, including inadequate insulin secretion or even β -cell failure. Low insulin with normal glucose often indicates enhanced insulin sensitivity. In addition to conclusions drawn from these separate observations, HOMA indices for IR and β -cell function are useful for initial estimates of the phenotype and the validity of performing further, more invasive tests.

Tolerance tests in general. In all tolerance tests, excluding the IPIST, a considerable increase in glucose (and insulin, if measured) is observed in normal mice, followed by a decrease back toward basal level. Upon insulin challenge (IPIST), the reverse occurs for glucose. The amount and rate of increase and decrease in glucose (and insulin) vary widely depending on the mice and possible dietary or other challenges used. Some indication to possible ranges of outcomes is given below for each test, with a brief description of what the outcomes may indicate. In all tests, baseline (i.e., fasting) values for glucose (and insulin) should correspond to those obtained from prior blood samples from the mice in question.

Meal tolerance test. Changes in plasma glucose and insulin values upon consumption of a standardized meal not only gives information about the glucose homeostatic interplay of muscle, adipose tissue, liver, β -cells, and the brain, but also takes into account the signals that originate from the intestine upon nutrient absorption. Thus, the results of this test are best utilized when correlated with the results of other tests (e.g., IPIST and IPGTT) where intestinal absorption does not play a role. Typically, after a meal in normal mice, glucose and insulin values are expected to increase approximately two- and three-fold, respectively,

compared to the fasting values. Food intake is normally in the range of 0.1 to 0.4 g of pellet chow.

Oral glucose tolerance test. Like the meal tolerance test, an oral dose of glucose will induce some intestinal effects; however, since only glucose is given rather than a full meal containing proteins, fats, and complex carbohydrates, this test scores only for intestinal glucose absorption. Typically, blood glucose is expected to increase 1.5 to 2-fold (T_{15}), followed by a slow decline towards baseline values. The evolution of blood glucose during the test often resembles that seen in the IPGTT (Fig. 29B.3.1A), although the peak is usually somewhat delayed and blunted and the decline back to baseline slower. In insulin resistant states, peak glucose is usually higher and, at a minimum, the following decline is delayed. An increase in the composite AUC measurement is typically a strong indication of stimulated IR, especially when calculated above baseline, as this eliminates the effect of any basal IR.

Intraperitoneal insulin sensitivity test. Excluding the euglycemic clamp, the intraperitoneal insulin sensitivity test is the most direct test of basal insulin sensitivity. Upon insulin injection, glucose typically falls from the fasting level until around T_{30} , after which it stabilizes and starts to climb back towards baseline upon degradation of the injected insulin (Fig. 29B.3.1B). In cases where FPG is greatly different between groups, plotting the glucose curve as percent of baseline values may be advantageous for visual representation, although this will be inherently accounted for in the calculations for the rate of initial decline and AUC_i using absolute values. The slower rate of decline in the initial phase of the test typically indicates impaired insulin sensitivity, mostly in skeletal muscle. During the later stages

of the test, especially at the final timepoints, slower recovery from the hypoglycemic state may indicate delayed activation of counter-regulatory signals or impaired hepatic glucose output. Optimally, all mice should react to insulin by lowered blood glucose levels that stay below baseline for the duration of the test. If this is not the case, calculating AUC_i becomes complicated due to difficulties in accommodating for those values that are higher than baseline. Care should be taken when analyzing the raw data to note such 'aberrant' values and possibly remove those mice from the analysis.

Intraperitoneal glucose tolerance test. Initial peak glucose values should be at least two-fold higher than the baseline value, but can be ten-fold higher or more, greatly depending on the particulars of the experimental set-up, including the genetic background of the mice as seen in Figure 29B.3.1A. When insulin is measured in conjunction with glucose, the evolution of that curve should resemble glucose, albeit with greater potential for variation. Increased AUC in an IPGTT indicates whole-body IR, in which skeletal muscle plays a major role. A significant increase in peak glucose ($T_{15}-T_{30}$) may be indicative of β -cell dysfunction; simultaneous insulin measurement is useful to support this.

Hyperinsulinemic-euglycemic clamp. During the clamp, insulin is administered as a constant infusion and therefore the inherent variations in endogenous insulin secretion are overwhelmed. Simultaneously, the plasma glucose concentration is clamped at the normal fasting level by means of an exogenous infusion of glucose. Thus, in vivo insulin action is measured under comparable conditions of stimulus and substrate. When the steady state is achieved, the exogenous glucose infusion rate equals the amount of glucose disposed of by all tissues, thus providing a quantification of overall insulin sensitivity. An increase in the glucose infusion rate implies increased insulin sensitivity, since the animal cleared more glucose for the same insulin concentration. For an example, see Figure 29B.3.4.

Time Considerations

General time considerations for metabolic phenotyping of mice have been described in UNIT 29B.1. Thus, only those considerations concerning specific protocols in this unit are given here.

Monitoring mouse blood glucose levels using glucometers designed for diabetic persons is very rapid, giving results within seconds. This is a prerequisite for performing more elaborate tests. Performing an insulin ELISA takes the better part of 1 day, but due to the microtiterplate format, hands-on time per sample is fairly short. The tolerance tests, excluding possible prior manipulations such as feeding the mice with altered diets, have a fixed time table. The only variability comes from the number of mice in each experiment. Excluding the overnight fasting, most of the tolerance tests take only a few hours to perform. The clamp is the most time-consuming procedure due to the requirement of catheter surgery and the small number of mice that can be experimented on at one time. Each blood sampling during the tests should not take > 1 min. Analysis of the data from tolerance tests and the clamp is straightforward and fast due to the fairly small number of data points and the lack of complicated calculations. However, interpretation of the data may be a lengthy process.

Literature Cited

- Ayala, J.E., Bracy, D.P., McGuinness, O.P., and Wasserman, D.H. 2006. Considerations in the design of hyperinsulinemic-euglycemic clamps in the conscious mouse. *Diabetes* 55:390-397.
- Champy, M.F., Selloum, M., Piard, L., Zeitler, V., Caradec, C., Chambon, P., and Auwerx, J. 2004. Mouse functional genomics requires standardization of mouse handling and housing conditions. *Mamm. Genome* 15:768-783.
- Hedrich, H. (ed.) 2004. *The Laboratory Mouse*. Elsevier Academic Press, London.
- Levy, J.C., Matthews, D.R., and Hermans, M.P. 1998. Correct homeostasis model assessment (homa) evaluation uses the computer program. *Diabetes Care* 21:2191-2192.
- Matthews, D.R., Hosker, J.P., Rudenski, A.S., Naylor, B.A., Treacher, D.F., and Turner, R.C. 1985. Homeostasis model assessment: Insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 28:412-419.
- Michael, M.D., Kulkarni, R.N., Postic, C., Previs, S.F., Shulman, G.I., Magnuson, M.A., and Kahn, C.R. 2000. Loss of insulin signaling in hepatocytes leads to severe insulin resistance and progressive hepatic dysfunction. *Mol. Cell* 6:87-97.

Wallace, T.M., Levy, J.C., and Matthews, D.R.
2004. Use and abuse of homa modeling.
Diabetes Care 27:1487-1495.

Contributed by Sami Heikkinen
A.I. Virtanen Institute for Molecular
Sciences
University of Kuopio
Kuopio, Finland

Carmen A. Argmann
Institut de Genetique et de Biologie
Moleculaire et Cellulaire
Illkirch, France

Marie-France Champy
Institut Clinique de la Souris
Illkirch, France

Johan Auwerx
Institut de Genetique et de Biologie
Moleculaire et Cellulaire and Institut
Clinique de la Souris
Illkirch, France

STANDARD MEASUREMENTS, DATA, AND ABBREVIATIONS

APPENDIX 1

Common Abbreviations

APPENDIX 1A

- A₂₆₀** absorbance at 260 nm
A adenine or adenosine; one-letter code for alanine
Ab antibody
ABTS [2,2'-azino-di(3-ethylbenzothiazoline sulfonate)]
acetyl CoA acetyl coenzyme A
AcMNPV *Autographica californica* multiply-enveloped nuclear polyhedrosis virus
ADA adenosine deaminase
ADC analog-to-digital converter
ADH alcohol dehydrogenase
ADP adenosine 5'-diphosphate
ADSL asynchronous digital subscriber line
AEC 3-amino-9-ethylcarbazole
AES 3-aminopropyltriethoxysilane
AEX anion exchange
AFLP amplified fragment length polymorphism
Ag antigen
AIDS acquired immune deficiency syndrome
AK adenosine kinase
ALARA as low as reasonably achievable
ALPS autoimmune/lymphoproliferative syndrome
AM acetomethyl (moiety)
AMan anhydro-D-mannose
AMP adenosine 5'-monophosphate
AMPPD disodium 3,4-methoxy Spiro{1,2-dioxetane-3,2-tricyclo[3.3.1.1^{3,7}]decan}phenyl phosphate
AMV avian myeloblastosis virus
ANOVA analysis of variance
AP alkaline phosphatase; apyrimidinic (sites)
APH aminoglycoside phosphotransferase
APHIS Animal and Plant Health Inspection Service
aPKC atypical protein kinase C
Ap^r ampicillin resistant
APRT adenosine phosphoribosyltransferase
APS ammonium persulfate
ARS autonomous replication sequences
ASPECT augmented surface polyethylene prepared by chemical transformation
ATA aurointricarboxylic acid
ATCC American Type Culture Collection
ATP adenosine 5'-triphosphate
AUFS absorbance units, full scale
AUS *Arthrobacter ureafaciens*
β-gal β-galactosidase
BAC bacterial artificial chromosome; biospecific affinity chromatography
BAP bacterial alkaline phosphatase
BBS BES-buffered solution; borate-buffered saline
BCIP 5-bromo-4-chloro-3-indolyl phosphate
BDB bis-diazobenzidine
BES *N,N*-bis(2-hydroxyethyl)-2-aminoethanesulfonic acid
BHI brain heart infusion (medium)
biotin-11-dUTP 8-(2,4-dinitrophenyl-2,6-aminohexyl)aminoadenosine-5'-triphosphate *or* 2'-deoxyuridine-5'-triphosphate-5'-allylamin biotin
bis; bisacrylamide *N,N'*-methylene bisacrylamide
bis-Tris 2-bis(2-hydroxyethyl)amino-2-(hydroxymethyl)-1,3-propanediol
BLAST Basic Local Alignment Research Tool
Bluo-gal indoyl-β-D-galactopyranoside
BMP bitmap (file format)
Boc *t*-butyloxycarbonyl
BOP benzotriazolyl-*N*-oxy-tris(dimethylamino)phosphonium hexafluorophosphate
bp base pair
BPV bovine papilloma virus
Bq Becquerel
BrdU 5-bromodeoxyuridine
BS3 bis(sulfosuccinimidyl) suberate
BSA bovine serum albumin
BSL biosafety level
Bst *Bacillus stearothermophilus* DNA (polymerase)
C cytosine or cytidine; one-letter code for cysteine
C₁₆TAB hexadecyl trimethylammonium bromide
CA3 chromomycin A3
CAD carbamoylphosphate synthetase
CaM calmodulin
cAMP adenosine 3',5'-cyclic-monophosphate
cA-PrK cyclic AMP-dependent protein kinase
CAPS [cyclohexylamino]-1-propanesulfonic acid

Standard
Measurements,
Data, and
Abbreviations

CAT chloramphenicol acetyltransferase	CTC charge-transfer chromatography
CATH Class, Architecture, Topology, and Homologous superfamily	CTD C-terminal domain
CCD charge-coupled device	CTP cytidine 5'-triphosphate
CCR Coriell Cell Repository	CWS cell wall skeleton
CD cluster of differentiation (antigens); circular dichroism	CXA contextual expression analysis
CDC Centers for Disease Control	D dextrorotatory
cdNA complementary deoxyribonucleic acid	dA deoxyadenosine
CD-ORD circular dichroism–optical rotatory dispersion	Da Dalton
CDP cytidine 5'-diphosphate	DAB 3,3'-diaminobenzidine
CD-ROM compact disk read-only memory	DABCO 1,4-diazobicyclo-[2.2.2]octane
CDS coding sequence	DAD diode array detection
CE capillary electrophoresis	DAG diacylglycerol
CED 3' cyanoethyl protected	dAMP deoxyadenosine monophosphate
CEF chicken embryo fibroblast	DAPI 4',6-diamidino-2-phenylindole
CERN European Nuclear Research Council	d(A-T) deoxyadenylate-deoxythymidylate
CEX cation exchange	dATP deoxyadenosine triphosphate
CFA complete Freund's adjuvant	DBE direct blotting electrophoresis
CFU colony-forming unit	DBM diazobenzoyloxymethyl
CGH comparative genome hybridization	dC deoxycytosine
CHAPS 3-[(3-cholamidopropyl)-dimethylammonio]-1-propane-sulfonate	DCA dichloroacetic acid
CHEF contour-clamped homogeneous electric field	DCC dextran-coated charcoal; <i>N,N'</i> -dicyclohexylcarbodiimide
CHES 2-(<i>N</i> -cyclohexylamino)ethanesulfonic acid	dCF 2'-deoxycoformycin
CHO Chinese hamster ovary (cells)	dCMP deoxycytidine monophosphate
Ci curie	dCTP deoxycytidine triphosphate
CID charge-injection device; collision-induced dissociation	DD differential display
CIP calf intestine phosphatase	ddATP dideoxyadenosine triphosphate
CLEAR cross-linked acrylate ethoxylate resin	DDBJ DNA Data Bank of Japan
cM centimorgans	ddCTP dideoxycytidine triphosphate
CM complete minimal (medium); carboxymethyl	DDDS distributed document delivery system
CMC critical micelle concentration	ddGTP dideoxyguanosine triphosphate
CML chronic myelogenous leukemia	ddNTP dideoxynucleoside triphosphate
CMP cytidine 5'-monophosphate	ddTTP dideoxythymidine triphosphate
Cm^r chloramphenicol resistant	DEA diethylamine
CMV cytomegalovirus	DEAE diethylaminoethyl
4CN 4-chloro-1-naphthol	DEPC diethylpyrocarbonate
CNBr cyanogen bromide	DES diethylstilbestrol
Con A concanavalin A	df degrees of freedom
CORT cloning of receptor targets	dG deoxyguanosine
CPC cetylpyridinium chloride	dGTP deoxyguanosine triphosphate
CPG control pore glass	DHFR dihydrofolate reductase
cpm counts per minute	DIEA <i>N,N</i> -diisopropylethylamine
CRD cross-reacting determinant	DIGE difference gel electrophoresis
CS chemical sequencing	DiI 1,1'-dihexyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate
CSPD disodium 3-(4-methoxy Spiro[1,2-dioxetane-3,2'-(5'-chloro)tricyclo[3.3.1.1 ^{3,7}]decan]-4-yl)phenyl phosphate	DIP Database of Interacting Proteins; deletion-insertion polymorphism
CTAB cetyltrimethylammonium bromide	dITP deoxyinosine 5'-triphosphate
	DMB 1,2-diamino-4,5-methylenedioxybenzene dihydrochloride
	DMEM Dulbecco's modified Eagle (or minimum essential) medium
	DMF dimethylformamide
	DMS dimethyl sulfate
	DMSO dimethyl sulfoxide
	DMT dimethoxytrityl
	DNA deoxyribonucleic acid

Common Abbreviations

A.1A.2

DNase deoxyribonuclease
DNP 2,4-dinitrophenyl
dNTP deoxynucleoside triphosphate
DOTMA *n*-[1-(2,3-dioleoyloxy)propyl]-*N,N,N*,-trimethylammonium chloride
DPA diphenylamine
dpm disintegrations per minute
ds double-stranded
DSA *Datura stramonium* agglutinin
DSG disuccinimidyl glutarate
DSL digital subscriber line
DSS disuccinimidyl suberate
dT deoxythymidine
DTE direct transfer electrophoresis
DTT dithiothreitol
dTTP deoxythymidine triphosphate
dUMP deoxyuridine monophosphate
dUTP deoxyuridine triphosphate
DvB deleted V (Mls-like loci)
EBI European Biotechnology Information
EBV Epstein-Barr virus
EC embryonic carcinoma; Enzyme Commission
ECL enhanced chemiluminescence
ECTEOA epichlorohydrin triethanolamine
EDC *N*-ethyl-*N'*-[(dimethylamino) propyl] carbodiimide hydrochloride
EDTA ethylenediaminetetraacetic acid
EGCase endoglycoceramidase
EGFR epidermal growth factor receptor
EGTA ethylene glycol-bis(β -aminoethyl ether)-*N,N,N',N'*-tetraacetic acid
ELISA enzyme-linked immunosorbent assay
EMBL European Molecular Biology Laboratory
EMC encephalomyocarditis
EMCV encephalomyocarditis virus
EMS ethyl methanesulfonate
EMSA electrophoretic mobility shift assay
Endo endoglycosidase
EOF electroosmotic flow
EP-PCR error-prone polymerase chain reaction
ES embryonic stem (cells)
ESI electrospray ionization (mass spectrometry)
EST expressed sequence tag
ET energy transfer
exo exonuclease
F Farad
FACS fluorescence-activated cell sorting
FAQ frequently asked questions
FBS fetal bovine serum
FCS fetal calf serum
FITC fluorescein isothiocyanate
FFPE formalin-fixed paraffin-embedded (tissue)
FIGE field-inversion gel electrophoresis
FISH fluorescence in situ hybridization
Fmoc fluorenylmethyloxycarbonyl
FOA fluoroorotic acid
FPLC fast protein; fast peptide; *or* fast polynucleotide liquid chromatography
FQDN fully-qualified domain name
FRET fluorescent resonant energy transfer
FSC forward (light) scatter (in flow cytometry)
FSSP Fold classification based on Structure-Structure alignment of Proteins
FTIR Fourier transform infrared (spectroscopy)
FTP File Transfer Protocol
Fuc L-fucose
FuDR 5-fluoro-2'-deoxyuridine
g gravity (unit of centrifugal force)
G gauge; guanine or guanosine; one-letter code for glycine
GAG glycosaminoglycan
Gal D-galactose
GalNAc *N*-acetylgalactosamine
GALV gibbon ape leukemia virus
GANC gancyclovir
Gb gigabyte
GDB Genetic Data Base
GDP guanosine 5'-diphosphate
GEM gene expression monitoring
GF gel filtration (chromatography)
GFP green fluorescent protein
GIF graphics interchange (file) format
Glc D-glucose
GlcA D-glucuronic acid
GLC-FID gas-liquid chromatography with flame ionization detection
GLC-MS gas-liquid chromatography with mass spectroscopic detection
GlcN D-glucosamine
GlcNAc *N*-acetylglucosamine
GIUA; GIA D-glucuronic acid
GM-CSF granulocyte/macrophage colony stimulating factor
GMP guanosine monophosphate
GMS genomic mismatch scanning
GNA *Galanthus nivalis* agglutinin
GPI glycosyl phosphatidyl inositol
GRAIL Gene Recognition and Analysis Internet Link
GRASS Graphical Representation and Analysis of Structure Server
GS glutamine synthetase
GSS genome survey sequence
GST glutathione *S*-transferase
gTLD generic top-level domains
GTP guanosine 5'-triphosphate
GUI graphical user interface
GUS β -glucuronidase
Gy Gray (radioactivity unit)

**Standard
Measurements,
Data, and
Abbreviations**

A.1A.3

HA (influenza) hemagglutinin protein	HRPO horseradish peroxidase
HAT hypoxanthine/aminopterin/thymidine (medium)	hsiRNA heterochromatic short interfering RNA
HATU <i>O</i> -(7-azabenzotriazol-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium hexafluorophosphate	HS-TBST high-salt TBST (buffer)
HBSS Hanks' buffered salt solution	HSV herpes simplex virus
HBTU <i>O</i> -benzotriazol-1-yl- <i>N,N,N',N'</i> -tetramethyluronium hexafluorophosphate	HTG/HTGS high-throughput genome sequence
HCG human chorionic gonadotropin	HTML hypertext markup language
hCMV human cytomegalovirus	Hz hertz
HeBS HEPES-buffered saline	IAA 3- β indoleacrylic acid; indole-3-acetic acid
HEC hydroxyethylcellulose	IACUC Institutional Animal Care and Use Committee
HEPA high-efficiency particulate air (filter)	ICAT isotope coded affinity tagging
HEPES <i>N</i> -2-hydroxyethylpiperazine- <i>N'</i> -2-ethanesulfonic acid	i.d. inner diameter
HFBA heptafluorobutyric acid	IdoA; IdUA; IdA L-iduronic acid
hGH human growth hormone	IEF isoelectric focusing
HGPRT hypoxanthine-guanine phosphoribosyltransferase	IEX ion exchange
HIC hydrophobic interaction chromatography	Ig immunoglobulin
HILIC hydrophilic interaction chromatography	imm immunity region
HIV human immunodeficiency virus	IMAC immobilized metal affinity chromatography
HOAt 1-hydroxy-7-azabenzotriazole	IMPDH inosine-monophosphate dehydrogenase
HOBT 1-hydroxybenzotriazole	IODOGEN 1,3,4,6-tetrachloro-3 α ,6 α -diphenylglycouril
HPAE-PAD high-performance anion exchange chromatography with pulsed amperometric detection	IP Internet Protocol
HP-BAC high-performance biospecific/biomimetic affinity chromatography	IPTG isopropyl-1-thio- β -D-galactoside
HPCF high-performance chromatofocusing	IR infrared
HP-CTC high-performance charge transfer chromatography	IRES internal ribosomal entry site
HPH hygromycin-B-phosphotransferase	ISDN integrated services digital network
HP-HIC high-performance hydrophobic-interaction chromatography	ISH in situ hybridization
HP-HILIC high-performance hydrophilic interaction chromatography	ISP Internet service provider
HP-IEX high-performance ion-exchange chromatography	ISPCR in situ PCR
HP-IMAC high-performance immobilized metal ion affinity chromatography	IVT in vitro transcription
HPLC high-performance liquid chromatography	JIPID Japan International Protein Information Database
HP-LEC high-performance ligand-exchange chromatography	JPEG Joint Photographic Experts Group (file format)
HPMC hydroxypropyl methyl cellulose	K Michaelis constant
HP-MMC high-performance mixed mode chromatography	kb kilobase
HP-NPC high-performance normal phase chromatography	kbps kilobits per second
HPRT hypoxanthine-guanine phosphoribosyltransferase	K_d dissociation constant
HP-SEC high-performance size-exclusion chromatography	kDa kilodalton
	KEGG Kyoto Encyclopedia of Genes and Genomes
	KHz kilohertz
	KLH keyhole limpet hemocyanin
	Km^r kanamycin resistant
	L levorotatory
	LAMP lysosome-associated protein
	LAN local area network
	LB Luria-Bertani (medium)
	LC liquid chromatography
	LCL lymphoblastoid cell lines
	LCM laser capture microdissection
	LCV lymphocryptovirus

Common Abbreviations

A.1A.4

LEC ligand-exchange chromatography	MMDM Molecular Modeling Database of NCBI
LIF leukemia inhibitory factor; laser-induced fluorescence (detector)	MMLV Moloney murine leukemia virus
LMPCR ligation-mediated polymerase chain reaction	MMT monomethoxytrityl
LPA linear polyacrylamide	MMTV mouse mammary tumor virus
LRSC lissamine rhodamine	mmu millimass unit or one thousandth of a Dalton
LTR long terminal repeat	MNase micrococcal nuclease
Lumigen-PPD 4-methoxy-4-(3-phosphate phenyl)-spiro-[1,2-dioxetane-3,2(-adamantane)], disodium salt	MOI multiplicity of infection
LysoPC lysophosphatidylcholine	MolMovDB Database of Macromolecular Movement
μF microfarad	MoMuLV Moloney murine leukemia virus
M relative molecular weight	MOPS 3-(<i>N</i> -morpholino)propane sulfonic acid
mA milliampere	mp melting point
MAA <i>Maackia amurensis</i> agglutinin	MPA mycophenolic acid
MAB, mAb monoclonal antibody	MPC magnetic plate chamber
MAB maleic acid (buffer)	MPSS Massive Parallel Signature Sequencing
MALDI matrix-assisted laser desorption/ionization (mass spectrometry)	mRNA messenger ribonucleic acid
MALDI-TOF matrix-assisted laser desorption/ionization time-of-flight (mass spectroscopy)	MS mass spectroscopy
Man D-mannose	MSCV murine stem cell virus
MAP mitogen-activated protein; multiple antigenic peptide	MS/MS tandem mass spectrometry
Mb megabase, megabyte	MSX methionine sulfoximine
Mbp megabase pair	Mtv mammary tumor virus designation
MBP maltose-binding protein	MTX methotrexate
Mbps megabits per second	MUG 4-methylumbelliferyl-β-D-galactoside
MBS <i>m</i> -maleimidobenzyl- <i>N</i> -hydroxysuccinimide ester	MUP methylumbelliferyl phosphate
MBTH 3-methyl-2-benzothiazolinone hydrazone hydrochloride	MVA Modified vaccinia virus Ankara
MCA methyl celluloseacetate ether	MWCO molecular weight cutoff
MCAC metal-chelate affinity chromatography	NA not applicable
MCS multiple cloning site	NAD nicotinamide adenine dinucleotide
MDCK Madin-Darby canine kidney (cells)	Na-DOC sodium deoxycholate
MDM multiply deficient medium	NBF neutral buffered formalin
2-ME 2-mercaptoethanol	NBRF National Biomedical Research Foundation
MEF mouse embryo fibroblasts	NBT nitroblue tetrazolium
MEM minimal essential medium	NCAM neuronal cell adhesion molecule
MEMPFA MOPS/sodium chloride/magnesium sulfate/paraformaldehyde (buffer)	NCBI National Center for Biotechnology Information
MES 2-(<i>N</i> -morpholino)ethanesulfonic acid	NCI National Cancer Institute
MHz megahertz	NCS newborn calf serum
MIPS Martinsried Institute for Protein Sequences	ND not determined
miRNA microRNA	NDV Newcastle Disease Virus
Mls minor lymphocyte stimulating determinant	NGF nerve growth factor
αMM α-methyl-D-mannoside	NLM National Library of Medicine
MMC mixed-mode chromatography	neo neomycin gene (selectable marker)
mmCIF macromolecular crystallographic information file	NEPHGE nonequilibrium pH gradient electrophoresis
	Neu5Ac <i>N</i> -acetyl-D-neuraminic acid
	Neu5Gc <i>N</i> -glycolyl-D-neuraminic acid
	NHS <i>N</i> -hydroxysuccinimide
	NICHD National Institute of Child Health and Human Development
	NIH National Institutes of Health
	NK natural killer (cells)
	NLM National Library of Medicine

NMR	nuclear magnetic resonance	PIPES	piperazine- <i>N,N'</i> -bis(2-ethane sulfonic acid)
NP-40	Nonidet P-40 (detergent)	PITC	phenylisothiocyanate
NPC	normal-phase chromatography	PKC	protein kinase C
NPP	nitrophenyl phosphate	PLAP	placental alkaline phosphatase
nr	nonredundant	PMF	peptide mass fingerprint
nt	nucleotide	PMS	pregnant mare's serum
NTA	nitrilotriacetic acid	PMSF	phenylmethylsulfonyl fluoride
NTA-SAM	nitrilotriacetic acid self-assembled monolayer	PMT	photomultiplier tube
NTP	nucleoside triphosphate	PNA	peanut agglutinin
OCT	optimal cutting temperature (medium)	PNGase-F	Peptide:N-Glycosidase F
o.d.	outer diameter	PNP	purine nucleoside kinase
OD₂₆₀	optical density at 260 nm	pNP	<i>p</i> -nitrophenol
OGT	O-GlcNAc transferase	PNPP	<i>p</i> -nitrophenyl phosphate
oligo	oligonucleotide, a short, single-stranded DNA or RNA.	poly(A)	polyadenylic acid or polyadenylate
oligo(dT)	oligodeoxythymidylic acid	poly(A)⁺	polyadenylated (mRNA)
OMIM	Online Mendelian Inheritance in Man	poly(A)⁻	nonpolyadenylated (mRNA)
ONPG	<i>o</i> -nitrophenyl- β -D-galactosidase	poly(dA-dT)	poly(deoxyadenylic acid-deoxythymidylic acid)
ORC	origin recognition complex	poly(U)	polyuridylic acid or polyuridylylate
ORF	open reading frame	PP1	protein phosphatase 1
ori	origin of replication	PP2A	protein phosphatase 2A
PAC	P1-derived artificial chromosome; phenoxyacetyl	PP2B	protein phosphatase 2B
PAD	pulsed amperometric detection	PPase	thermostable pyrophosphatase
PAGE	polyacrylamide gel electrophoresis	ppm	parts per million
PAH	polyaromatic hydrocarbons	PPO	2,5-diphenyloxazole
PAP	peroxidase-anti-peroxidase (reaction)	PR	prosthetic-group removing
par	partition loci on plasmid DNA	PRF	Protein Research Foundation
PB	phosphate buffer	PRINS	primed in situ (labeling)
PBMC	peripheral blood mononuclear cells	Pristane	2,6,10,14-tetramethylpentadecane
PBS	phosphate-buffered saline	ProTherm	Thermodynamic Database for Proteins and Mutants
PCD	programmed cell death	PS	copoly(styrene-1%-divinylbenzene)
PCMB	parachloromercuric benzoate	PSF	point-spread function
PCR	polymerase chain reaction	PSI-BLAST	Position-Specific Iterated BLAST
PDB	Protein Data Bank	Pth	3-phenyl-2-thiohydrantoin
PDD	Protein Disease Database	PTP	protein tyrosine phosphatase
PDMP	1-phenyl-2-decanoylamino-3-morpholino-1-propanol	Pu	purine
PE	phycoerythrin	PVA	polyvinyl alcohol
PEEK	polyethylether ketone	PVC	polyvinyl chloride
PEG	polyethylene glycol	PVDF	polyvinylidene difluoride
PEGA	polyethylene glycol polyacrylamide	PVP	polyvinylpyrrolidone
PEI	polyethylenimine	Py	pyrimidine
PEO	polyethylene oxide	PyAOP	7-azabenzotriazol-1-yl-oxytris(pyrrolidino)phosphonium hexafluorophosphate
PFA	paraformaldehyde	PyBOP	benzotriazolyl- <i>N</i> -oxytris(pyrrolidino)phosphonium hexafluorophosphate
PFGE	pulsed-field gel electrophoresis	rA	riboadenylate
pfu	plaque-forming units	RACE	rapid amplification of cDNA ends
PG	proteoglycan	RAS	Ribi adjuvant system
PGK	phosphoglycerate kinase	RBC	red blood cell
pI	isoelectric point	RBE	relative biological effectiveness
PI	phosphatidylinositol; propidium iodide	RBS	ribosome-binding site
PI-PLC	phosphatidylinositol-specific phospholipase C	RCA I	<i>Ricinus communis</i> agglutinin
PIR	Protein Information Resource		
PITC	phenylisothiocyanate		

Common Abbreviations

A.1A.6

RCF relative centrifugal force
RDA representational difference analysis
RE restriction endonuclease
Red-gal 6-chloro-3-indoyl- β -D-galactopyranoside
RF replicative form
RFLP restriction-fragment-length polymorphisms
RIA radioimmunoassay
RIPA RadioImmunoPrecipitation Assay
RMDD restriction-mediated differential display
RNA ribonucleic acid
RNAi RNA interference
RNase ribonuclease
RP reversed phase (HPLC)
RRE Rev-responsive element
rRNA ribosomal ribonucleic acid
RT reverse transcriptase
RT-PCR reverse transcription/polymerase chain reaction
RU resonance unit
RXR retinoid X receptor
SAGE serial analysis of gene expression
SAM *S*-adenosylmethionine
Sarkosyl *N*-lauroylsarcosine
SAX strong anion exchange
SBH sequencing by hybridization
SCOP Structural Classification of Proteins
SCX strong cation exchange
SD standard deviation
SDS sodium dodecyl sulfate
SE size exclusion (chromatography)
SEAP secreted alkaline phosphatase
SEC size-exclusion chromatography
SED standard enzyme diluent
SELDI surface-enhanced laser desorption/ionization
SFFV spleen focus-forming virus
SHMT serine hydroxymethyl synthetase
ShrAP shrimp alkaline phosphatase
Sia sialic acid
siRNA short interfering RNA
SM suspension medium
SNA *Sambucus nigra* agglutinin
SNP single-nucleotide polymorphism
SPF specific pathogen free
SPPS solid-phase peptide synthesis
SPR surface plasmon resonance
SPRI solid-phase reversible immobilization
SPW surface plasma or plasmon wave
SRBC sheep red blood cells
SREBP steroid response element binding protein
ss single stranded
SSB single-stranded DNA-binding protein
SSC sodium chloride/sodium citrate (buffer); side (light) scatter (in flow cytometry)
SSCP single-stranded conformation polymorphism
SSAV simian sarcoma-associated virus
sss sheared salmon sperm
STBS suspension Tris-buffered saline
STO SIM mouse embryo fibroblasts resistant to thioguanine and ouabain
STS sequence tagged site
STZ streptozotocin
Sv Sievert (unit for radiation dosage)
T thymine or thymidine; one-letter code for threonine
TAE Tris/acetate (buffer)
Taq *Thermus aquaticus* DNA (polymerase)
TAU Triton/acetic acid/urea
TBE Tris/borate (buffer)
TBP TATA box-binding protein
TBS Tris-buffered saline
TBT TATA-binding protein
TBTU *O*-benzotriazol-1-yl-*N,N,N',N'*-tetramethyluronium tetrafluoroborate
TCA trichloroacetic acid
TCEP tris(2-carboxyethyl)phosphine
TCP Transmission Control Protocol
TCR T cell receptor
TDM trehalose dimycolate
TdT terminal deoxynucleotidyl transferase
TE Tris/EDTA (buffer)
TEA triethanolamine acetate
TEAE triethylaminoethyl
TEMED *N,N,N',N'*-tetramethylethylenediamine
TEN NaCl in TE buffer
TES *N*-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid
TFA trifluoroacetic acid
TFFH tetramethylfluoroformamidinium hexafluorophosphate
TFMSA trifluoromethanesulfonic acid
TGN trans-Golgi network
THF tetrahydrofuran
TIFF tagged-image file format
TIR total internal reflection
TK thymidine kinase
TLC thin-layer chromatography
TLD thermoluminescent dosimeter
 T_m melting (or midpoint) temperature; thermal denaturation
TMAC tetramethylammonium chloride
TMB 3,3',5,5'-tetramethylbenzidine
TMP trimethylphosphate; thymidine monophosphate
TMV Tobacco Mosaic Virus
TONPG ortho-nitrophenyl- β -D-thiogalactoside
TPCK *N*-*p*-tosyl-L-phenylalanine chloromethyl ketone
TPF tiling path format

**Standard
Measurements,
Data, and
Abbreviations**

A.1A.7

Tris tris(hydroxymethyl)aminomethane	VRC vanadyl-ribonucleoside complex
TRITC tetramethylrhodamine isothiocyanate	V₀ void volume
tRNA transfer ribonucleic acid	vol/vol; v/v volume/volume
TS thymidylate synthetase	VSG variant surface glycoprotein
TSA Tris/saline/azide (buffer)	VSV vesicular stomatitis virus
TTP thymidine 5'-triphosphate	WAIS Wide Area Information Service
TUNEL TdT-mediated dUTP biotin nick-end labeling	WAX weak anion exchange
U unit; uracil or uridine	WCX weak cation exchange
UAS upstream activating sequence	WGA wheat germ agglutinin
UDG uracil DNA glycosylase	WR Western Reserve strain (vaccinia)
UDP uridine 5'-diphosphate	WT wild-type
UDP-Gal uridine diphospho-D-galactose	wt/vol; w/v weight/volume
UF ultrafiltration	WWW World Wide Web
UMP uridine 5'-monophosphate	XBE Rex-binding element
UPHS U.S. Public Health Service	Xgal 5-bromo-4-chloro-3-indolyl- β -D-galactoside
URL uniform resource locator	XGPRT xanthine-guanine phosphoribosyl transferase
USDA United States Department of Agriculture	Xyl xylose
UTP uridine 5'-triphosphate	Xyl-A 9- β -D-xylofuranosyl adenine
UTR untranslated leader region	YAC yeast artificial chromosome
UV ultraviolet	YCp yeast centromeric plasmid
UWGCG University of Wisconsin Genetics Computer Group	YEp yeast episomal plasmid
VAF viral-antibody free	YIp yeast integrating plasmid
VAST Vector Alignment Search Tool	YNB-AA/AS yeast nitrogen base without amino acids or ammonium sulfate
Vent <i>Thermococcus litoralis</i> DNA (polymerase)	YPD yeast/peptone/dextrose (medium)
	YRp yeast replicating plasmid

Figure A.1B.1 A physical chemist's view of the cell. The data in this figure were assembled from *The Molecular Biology of the Cell* by Alberts, 1994, and represent the approximate concentrations of a variety of intracellular components.

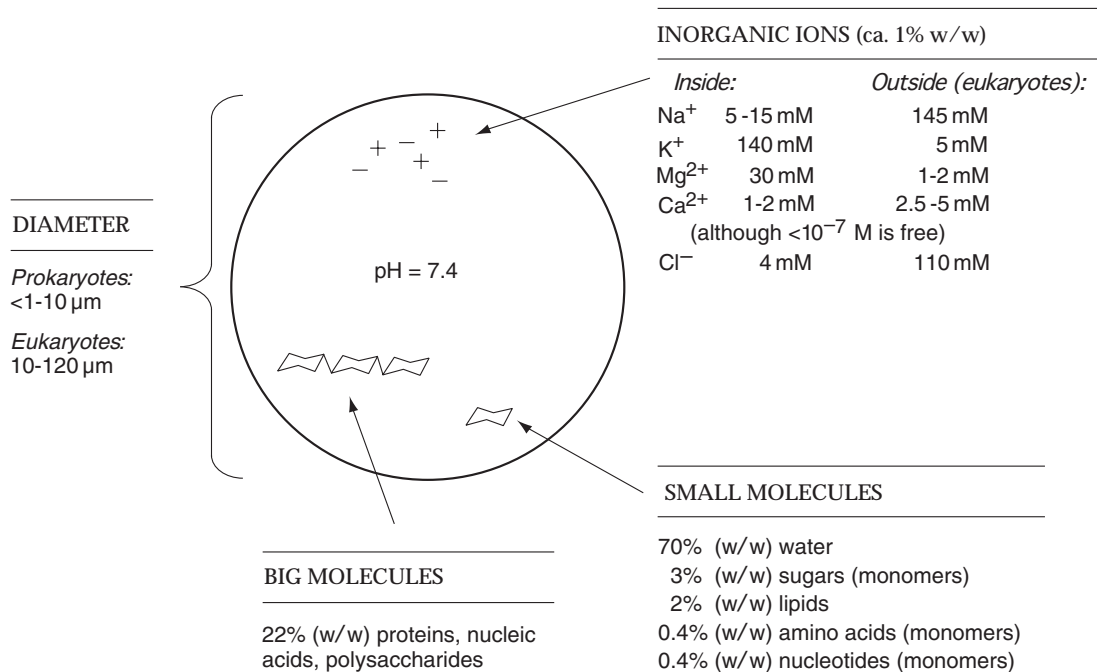


Table A.1B.1 Conversion Factors

Molecular weight (ave.) of DNA base pair: 649 Da	1 kb DNA: 333 amino acids of coding capacity
Molecular weight (ave.) of amino acid: 110 Da	$\approx 36,000$ Da
1 $\mu\text{g/ml}$ DNA: 3.08 μM phosphate	6.5×10^5 Da of double-stranded DNA (sodium salt)
1 $\mu\text{g/ml}$ of 1 kb DNA: 3.08 nM 5' ends	3.3×10^5 Da of single-stranded DNA (sodium salt)
1 μmol pBR322 (4363 bp): 2.83 g	3.4×10^5 Da of single-stranded RNA (sodium salt)
1 pmol linear pBR322 5' ends: 1.4 μg	
1 A_{260} double-stranded DNA: 50 $\mu\text{g/ml}$	10 kDa protein ≈ 91 amino acids
1 A_{260} single-stranded DNA: 37 $\mu\text{g/ml}$	≈ 273 nucleotides

Table A.1B.2 Genome Size of Various Organisms^a

Organism	Base pairs/ haploid genome	Organism	Base pairs/ haploid genome
SV40	5,243	<i>Drosophila melanogaster</i>	1.4×10^8
ΦX174	5,386	<i>Gallus domesticus</i> (chicken)	1.2×10^9
Adenovirus 2	35,937	<i>Mus musculus</i> (mouse)	2.7×10^9
Lambda	48,502	<i>Rattus norvegicus</i> (rat)	3.0×10^9
<i>Escherichia coli</i>	4.7×10^6	<i>Xenopus laevis</i>	3.1×10^9
<i>Saccharomyces cerevisiae</i>	1.5×10^7	<i>Homo sapiens</i>	3.3×10^9
<i>Dictyostelium discoideum</i>	5.4×10^7	<i>Zea mays</i>	3.9×10^9
<i>Arabidopsis thaliana</i>	7.0×10^7	<i>Nicotiana tabacum</i>	4.8×10^9
<i>Caenorhabditis elegans</i>	8.0×10^7		

^aGenome size determined either by direct sequence analysis (viruses), electrophoretic analysis (*E. coli*, *S. cerevisiae*), or a combination of DNA content per cell and hybridization kinetics. Some data are from *Gene Expression 2* by Lewin, 1980.

Characteristics of Amino Acids

PHYSICAL PROPERTIES

The physical properties of the amino acids determine the structure and function of the proteins in which they are found. Some useful details and relevant physical characteristics of the amino acids can be found in Table A.1C.1. A detailed view of the chemical structures of the amino acids, and an explanation of the role these structures play in enzymes, can be found in Figure A.1C.2. The three-dimensional structure of proteins is largely determined by the packing of their hydrophobic cores; the properties of amino acids that govern this packing are their relative hydrophobicities, which are presented in Figure A.1C.3, and their shapes and volumes, which can be assessed by referring to the space-filling models shown in Figure A.1C.4. While these figures can be useful in rationalizing amino acid functionality, it is also important to consider how natural selection views the interchangeability of amino acids, as diagrammed in Figure A.1C.5 (and Table A.1C.1).

Post-translational modifications will change the mass of a protein or peptide; values for some common mass changes are listed in Table A.1C.3. Mass changes due to some post-translational modifications are found in Table A.1C.4.

CODON USAGE (see Table A.1C.2)

While the amino acid sequence of a protein is selected for in part because of the physical properties of the amino acids themselves, a second, more subtle selection may also operate at the level of the genetic code to determine the sequence of both protein and gene. The genetic code is degenerate. Any of several codons can represent a single amino acid (up to six, in the cases of Arg, Leu, and Ser). However, the frequency with which such synonymous codons are used is not equivalent. Considerations of the bias in codon usage may be relevant to design of synthetic genes (*UNIT 8.2B*), strategies for overexpression of foreign proteins, particularly in *E. coli* (see *UNIT 16.1*), and minimizing degeneracy of oligonucleotide probes and primers (*UNIT 6.4* and *UNIT 15.1*).

The reasons for deviation from random usage seem to differ from organism to organism. For *E. coli* and other microorganisms, it is thought that the codons used more frequently correspond to abundant tRNAs, while the underrepresented codons are those associated with less abundant tRNAs. Since this bias seems particularly strong for genes encoding highly expressed proteins, it is thought to be related to maximizing translation efficiency. In higher organisms, the bias in codon usage may be more closely associated with selection pressures acting at the level of DNA. Mammalian genomes, in particular, show quite significant reductions in the frequency of the dinucleotide CpG, which is a site for methylation. In mammalian genes, codons containing this sequence can be quite strongly underrepresented.

It is worth emphasizing that the bias against particular codons is not absolute. While there may be strong trends within a particular organism, individual genes (particularly those expressed at low levels) may deviate substantially (see Sharp et al., 1988). Interestingly, codon usage is quite similar within the broad groups presented (see Wada et al., 1990, for a comparison of mammals).

Table A.1C.1 Physical Characteristics of the Amino Acids

Amino acid	3-letter code	1-letter code	Mol. wt. (g/mol)	Accessible surface area ^a	Hydrophobicity ^b	Relative mutability ^c	Surface probability ^d
Alanine	Ala	A	89.1	115	-0.40	100	62
Arginine	Arg	R	174.2	225	-0.59	65	99
Asparagine	Asn	N	132.1	160	-0.92	134	88
Aspartate	Asp	D	133.1	150	-1.31	106	85
Cysteine	Cys	C	121.2	135	0.17	20	55
Glutamate	Glu	E	147.1	190	-1.22	102	82
Glutamine	Gln	Q	146.2	180	-0.91	93	93
Glycine	Gly	G	75.1	75	-0.67	49	64
Histidine	His	H	155.2	195	-0.64	66	83
Isoleucine	Ile	I	131.2	175	1.25	96	40
Leucine	Leu	L	131.2	170	1.22	40	55
Lysine	Lys	K	146.2	200	-0.67	56	97
Methionine	Met	M	149.2	185	1.02	94	60
Phenylalanine	Phe	F	165.2	210	1.92	41	50
Proline	Pro	P	115.1	145	-0.49	56	82
Serine	Ser	S	105.1	115	-0.55	120	78
Threonine	Thr	T	119.1	140	-0.28	97	77
Tryptophan	Trp	W	204.2	255	0.50	18	73
Tyrosine	Tyr	Y	181.2	230	1.67	41	85
Valine	Val	V	117.1	155	0.91	74	46

^aAccessible surface area is in Å² and is for the amino acid as part of a polypeptide backbone (Chothia, 1976).

^bHydrophobicity is in arbitrary units and is based on the OMH scale of Sweet and Eisenberg (1983), which emphasizes the ability of amino acids to replace one another during the course of evolution.

^cRelative mutability is also in arbitrary units (with alanine set to 100) and represents the probability that an amino acid will mutate within a given time. Thus, as two closely related proteins diverge, a given tryptophan residue is only 18% as likely as a given alanine residue to mutate (Dayhoff et al., 1978).

^dSurface probability is the likelihood that 5% or more of the surface area of an amino acid will be exposed to the solution surrounding a protein (Chothia, 1976). Thus, while some portion of almost all the arginines will help make up the surface of a protein, less than half of the valines will be exposed to solution. To understand in more detail how amino acids are buried, see Rose et al. (1985); for example, although tyrosine is often found exposed to the surface of a protein, a substantial proportion of its surface area is typically buried.

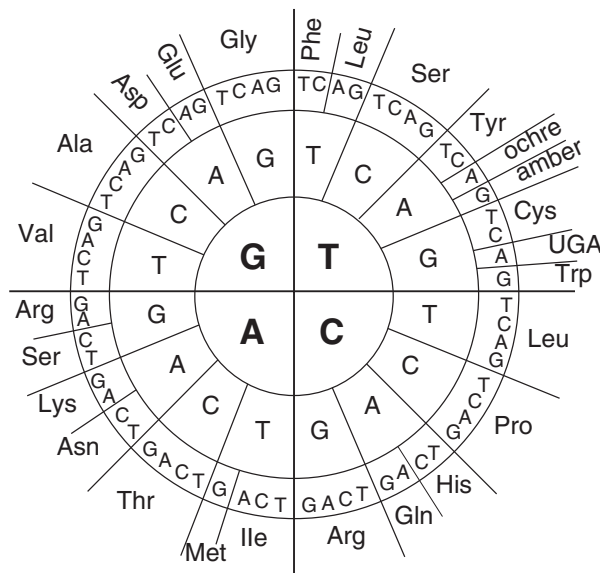


Figure A.1C.1 The genetic code. Names of amino acids and chain termination codons are on the periphery of the circle. The first base of the codon is identified in the center ring; the second base of the codon is in the middle ring; and the third base(s) of the codon is in the outer ring of the circle.

**Standard
Measurements,
Data, and
Abbreviations**

A.1C.2

Table A.1C.2 Percentage of Codon Synonymous Usage and Frequency of Codon Occurrence in Various Organisms (see description below)

AA	Codon	Mammal % freq.	Other vert. % freq.	Dicot % freq.	Monocot % freq.	Invertebrate % freq.	Yeast % freq.	Chloroplast % freq.	Yeast mito. % freq.	Gram neg. % freq.	Gram pos. % freq.
Gly	GGG	22.8	16.6	11.6	25.0	8.4	9.0	15.3	6.0	12.3	13.7
Gly	GGA	25.5	18.5	37.9	21.2	32.4	15.0	38.1	21.0	9.2	30.5
Gly	GGU	17.6	12.8	35.6	15.2	21.4	61.0	35.6	68.0	32.9	27.2
Gly	GGC	34.1	24.7	14.9	38.6	37.8	15.0	11.0	4.0	45.6	28.5
Glu	GAG	60.1	40.4	48.9	72.5	66.0	26.0	23.7	20.0	7.2	26.6
Glu	GAA	39.9	26.8	51.1	27.5	34.0	74.0	76.3	80.0	65.5	73.4
Asp	GAU	42.6	20.9	60.1	29.7	53.0	62.0	75.4	72.0	55.2	67.6
Asp	GAC	57.4	28.1	39.9	70.3	47.0	38.0	24.6	28.0	44.8	32.4
Val	GUG	48.1	29.5	28.6	38.0	41.6	15.0	15.2	8.0	34.2	21.9
Val	GUA	9.9	6.1	14.0	9.5	10.6	16.0	40.6	46.0	27.7	25.8
Val	GUU	16.4	10.0	40.0	18.8	21.1	44.0	34.3	37.0	25.9	32.2
Val	GUC	25.6	15.7	17.4	33.7	26.8	24.0	9.9	9.0	23.9	20.0
Ala	GCG	9.9	7.0	6.6	25.3	18.1	8.0	10.1	5.0	34.3	21.4
Ala	GCA	21.0	14.9	27.0	18.0	18.2	23.0	26.4	36.0	21.7	35.7
Ala	GCU	28.8	20.4	44.6	21.8	21.4	44.0	48.5	51.0	17.2	27.6
Ala	GCC	40.2	28.5	21.8	34.9	42.4	25.0	15.0	8.0	29.2	15.4
Arg	AGG	21.5	11.8	25.1	24.4	14.1	17.0	9.3	11.0	3.2	9.9
Arg	AGA	21.0	11.5	32.6	13.0	14.6	54.0	30.1	70.0	4.3	33.1
Ser	AGU	13.4	10.0	16.4	7.7	12.4	14.0	19.7	15.0	11.8	14.6
Ser	AGC	24.8	18.6	17.3	22.1	22.4	9.0	10.1	5.0	28.0	13.9
Lys	AAG	64.4	36.1	57.7	80.5	70.6	49.0	28.7	25.0	14.2	20.5
Lys	AAA	37.6	21.7	42.3	19.5	29.4	51.0	71.3	75.0	41.7	75.8
Asn	AUU	41.4	15.5	45.9	25.9	44.1	54.0	66.5	86.0	66.2	58.2
Asn	AAC	58.6	22.0	54.1	21.9	55.9	46.0	33.5	14.0	10.8	41.8
Met	AUG	100	23.4	100	20.6	100	100	100	65.0	28.7	100
Ile	AUA	13.1	16.8	18.9	13.6	15.6	20.0	24.5	35.0	15.4	18.7
Ile	AUU	33.3	14.9	46.2	27.5	33.8	50.0	51.7	86.0	42.6	51.5
Ile	AUC	53.6	24.0	34.9	58.9	50.6	30.0	23.8	14.0	50.1	29.8
Thr	ACG	11.8	6.6	8.5	21.1	24.7	12.0	9.0	6.0	24.4	20.3
Thr	ACA	26.4	14.8	29.3	18.9	21.8	26.0	30.1	46.0	11.5	47.3
Thr	ACU	23.4	13.1	35.0	20.9	18.0	38.0	40.4	38.0	18.8	21.3
Thr	ACC	38.5	21.6	27.2	39.0	35.5	25.0	20.5	10.0	45.3	11.0
Trp	UGG	100	14.4	100	13.0	100	100	100	39.0	6.8	100
End	UGA	59.6	2.6	51.4	60.3	56.2	34.0	23.1	61.0	33.2	18.9
Cys	UGU	42.7	10.0	49.3	27.0	31.5	68.0	71.4	87.0	6.6	47.8
Cys	UGC	57.3	13.4	50.7	73.0	68.5	32.0	28.6	13.0	62.0	52.2
End	UAG	17.2	0.7	18.0	21.3	15.3	17.0	15.4	7.0	9.7	15.6
End	UAA	23.2	1.0	30.6	18.4	28.5	49.0	61.5	93.0	57.2	65.6
Tyr	UAU	40.2	11.5	46.9	25.7	36.9	50.0	71.6	82.0	50.6	70.6
Tyr	UAC	59.8	17.2	53.1	74.3	63.1	50.0	28.4	18.0	49.4	29.4
Leu	UUG	12.2	11.4	26.1	15.7	16.8	36.0	21.5	10.0	11.6	15.0
Leu	UUA	5.4	5.0	11.1	4.6	7.0	27.0	31.5	72.0	9.4	29.5
Phe	UUU	40.7	15.3	49.5	30.6	34.8	53.0	59.8	60.0	47.5	68.0
Phe	UUC	59.3	22.3	50.5	69.4	65.2	47.0	40.2	40.0	52.5	32.0

A.1C.3

AA	Codon	Mammal % freq.	Other vert. % freq.	Dicot % freq.	Monocot % freq.	Invertebrate % freq.	Yeast % freq.	Chloroplast % freq.	Yeast mito. % freq.	Gram neg. % freq.	Gram pos. % freq.								
Ser	UCG	5.9	4.4	5.8	4.5	17.0	12.0	19.5	15.7	8.0	6.5	7.0	4.5	4.0	3.0	15.4	9.1	8.6	5.8
Ser	UCA	14.2	10.6	19.7	15.3	16.0	11.3	11.4	9.9	19.0	15.3	16.1	10.5	44.0	31.5	10.7	6.4	22.2	15.0
Ser	UCU	18.3	13.8	23.6	18.3	15.3	10.8	11.3	9.7	32.0	25.4	27.1	17.6	25.0	17.7	16.1	9.6	23.4	15.9
Ser	UCC	23.5	17.6	17.2	13.4	21.9	15.5	23.0	18.9	18.0	14.8	20.0	13.0	7.0	4.8	18.0	10.7	10.7	7.3
Arg	CGG	18.2	10.0	5.0	2.4	16.4	8.6	13.8	8.7	2.0	1.0	6.1	4.1	4.0	1.2	9.6	5.6	10.8	4.1
Arg	CGA	10.7	5.8	9.5	4.5	9.0	4.7	14.1	9.3	5.0	2.1	20.1	13.5	1.0	0.2	5.6	3.3	10.5	4.0
Arg	CGU	9.1	5.0	18.8	8.9	11.2	5.8	16.5	10.9	17.0	7.3	25.5	17.2	13.0	3.6	36.9	21.4	19.9	7.6
Arg	CGC	19.5	10.7	9.0	4.2	26.0	13.6	27.0	17.0	4.0	1.9	8.9	6.0	1.0	0.4	40.4	23.5	15.8	6.0
Gln	CAG	73.9	32.7	39.7	15.2	40.1	43.4	65.0	32.8	26.0	10.3	26.8	9.9	17.0	4.4	70.2	29.4	36.4	15.3
Gln	CAA	26.1	11.6	60.3	23.1	59.9	64.8	35.0	18.2	74.0	29.7	73.2	27.1	83.0	21.9	29.8	12.5	63.6	26.7
His	CAU	38.8	9.2	53.0	12.0	38.2	8.1	41.6	12.1	60.0	12.5	71.5	14.3	83.0	19.5	50.1	11.1	70.4	15.8
His	CAC	61.2	14.5	47.0	10.7	61.8	13.2	58.4	16.5	40.0	8.3	28.5	5.7	17.0	4.0	49.9	11.1	29.6	6.6
Leu	CUG	42.8	40.1	11.0	9.6	28.2	23.2	38.1	30.0	9.0	8.5	6.2	5.7	5.0	6.4	53.7	51.5	18.5	15.9
Leu	CUA	6.8	6.4	9.8	8.6	9.8	8.1	8.6	7.0	13.0	11.8	15.4	14.1	6.0	8.0	3.3	3.1	7.4	6.4
Leu	CUU	12.1	11.3	25.2	22.2	13.4	11.0	12.4	10.3	11.0	9.6	19.6	17.9	5.0	6.0	9.9	9.5	21.6	18.6
Leu	CUC	20.8	19.5	16.9	14.8	28.3	23.3	17.2	14.3	4.0	4.0	5.7	5.2	1.0	1.4	12.1	11.6	8.0	6.9
Pro	CCG	11.2	6.8	8.2	4.6	25.7	20.5	27.3	15.5	9.0	4.1	12.8	6.3	5.0	2.0	54.4	23.7	30.3	10.8
Pro	CCA	27.3	16.5	41.9	23.6	40.9	32.6	30.1	17.9	49.0	21.9	23.3	11.4	39.0	15.6	17.1	7.4	28.9	10.3
Pro	CCU	28.8	17.4	34.1	19.2	15.0	12.0	14.6	8.9	29.0	12.7	43.3	21.1	48.0	19.3	15.2	6.6	33.2	11.9
Pro	CCC	32.7	19.7	15.9	8.9	18.3	14.6	28.0	16.1	13.0	5.7	20.6	10.0	8.0	3.4	13.3	5.8	7.6	2.7

A discussion of codon randomness and the relevance of this phenomenon in molecular biology experimentation can be found on p. A.1.6. (% **Column**) Percentage of synonymous codon usage. The relative percentage of each member of the set of codons that specify a particular amino acid. Four sets that are not contiguous in this table are Arg, Ser, Leu, and chain terminator. (**Freq. Column**) Frequencies are expressed as occurrences per thousand codons for the specified organisms. The GenBank nucleic acid database as of April 23, 1990, was used as the source of gene-coding regions. This corresponds to GenBank Release 63 with the addition of three weekly updates. The feature tables were utilized to automatically extract peptide-coding regions from the database. This automated process discarded sequences that were less than ten codons in length and contained more than one chain-termination codon. The extracted coding sequences were then examined to remove all pseudogenes, nonnuclear (except for the chloroplast and mitochondrial categories), viral, rearranged, and mutant sequences. The remaining collection of extracted sequences was edited to remove duplicate entries; most significantly, only one example of each class of immunoglobulin gene was allowed per species. The codon frequencies were tabulated with the aid of the CodonFrequency program from the UWGCG package (Genetics Computer Group, Madison, WI 53711).

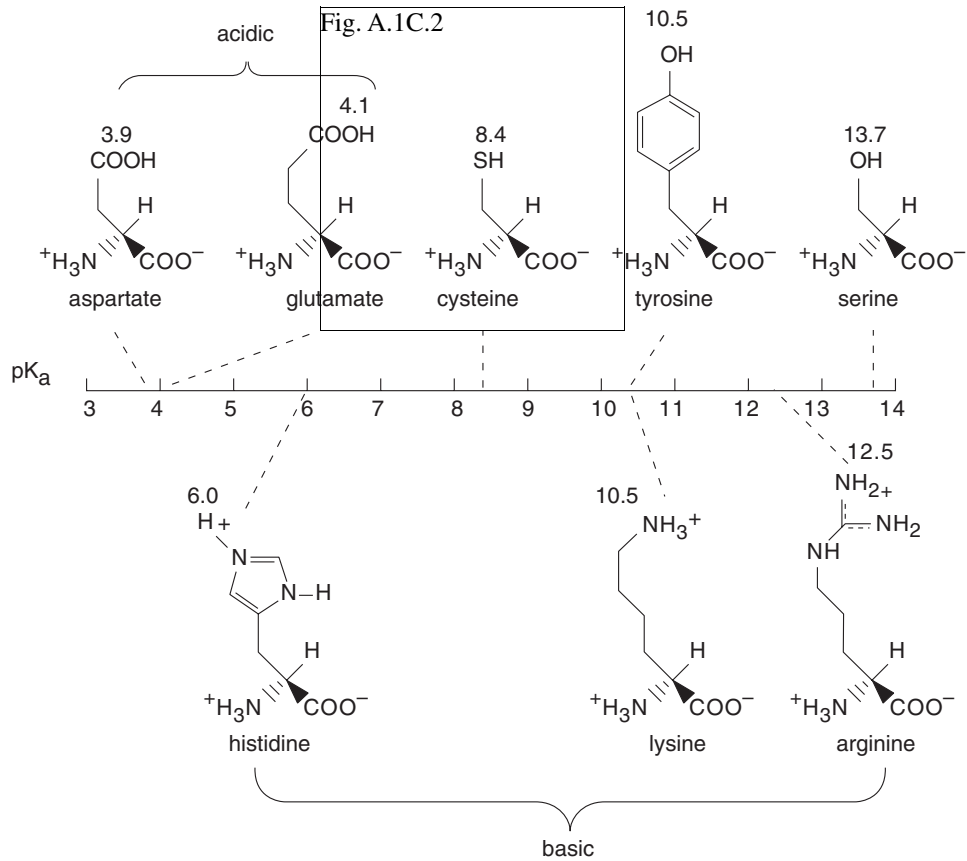
The total number of codons used for each organism category and the source of the codons are listed below expressed as a percentage of the total: **Mammal:** Total of 1,237,027 codons from cow (*Bos taurus*; 6.43%), hamster (*Cricetus sp.*; 1.48%), human (*Homo sapiens*; 48.36%), macaque (*Macaca sp.*; 0.38%), mouse (*Mus sp.*; 19.52%), rabbit (*Oryctolagus sp.*; 4.06%), rat (*Rattus sp.*; 19.17%), and sheep (*Ovis sp.*; 0.59%). **Other**

vertebrate: Total of 159, 994 codons from chicken (*Gallus sp.*; 72.24%) and *Xenopus laevis* (27.76%). **Dicot:** Total of 71, 408 codons from *Arabidopsis thaliana* (15.21%), pea (*Pisum sativum*; 13.77%), *Petunia sp.* (6.61%), lima bean (*Phaseolus vulgaris*; 11.24%), potato (*Solanum tuberosum*; 7.89%), tobacco (*Nicotiana tabacum*; 9.54%), tomato (*Lycopersicon esculentum*; 11.63%), and soybean (*Glycine max*; 24.10%). **Monocot:** Total of 45, 622 codons from barley (*Hordeum vulgare*; 21.35%), corn (*Zea mays* 47.70%), rice (*Oryza sativa*; 11.60%) and wheat (*Triticum aestivum*; 19.35%). **Invertebrate:** Total of 151, 794 codons from *Caenorhabditis elegans* (9.11%), sea urchin (*Strongylocentrotus purpuratus*; 6.01%), and fruit fly (*Drosophila melanogaster*; 84.88%). **Yeast:** Total of 216, 375 codons from *Saccharomyces cerevisiae*. **Chloroplast:** Total of 6, 866 codons from *Zea mays* chloroplast (52.18%) and *Nicotiana tabacum* chloroplast (47.82%). **Yeast mitochondrion:** Total of 4, 986 codons from *Saccharomyces cerevisiae* mitochondria (see note below). **Gram negative bacteria:** Total of 263, 904 codons from *Escherichia coli* (70.40%), *Klebsiella pneumoniae* (3.96%), *Neisseria gonorrhoeae* (1.75%), *Pseudomonas sp.* (10.97%), *Rhizobium meliloti* (2.57%), and *Salmonella typhimurium* (10.35%). **Gram positive bacteria:** Total of 38,807 codons from *Bacillus subtilis* (73.45%) and *Staphylococcus aureus* (26.55%).

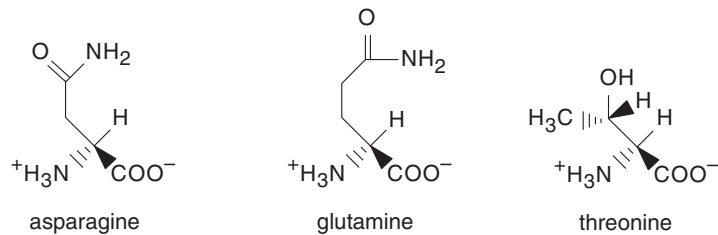
NOTE: The genetic code is universal with the exception of mitochondrial DNA. In yeast mitochondria the AUA and UGA codons that normally specify isoleucine and chain termination are used for methionine and tryptophan, respectively. These exceptions are used when calculating the yeast mitochondrion percentage of synonymous codon usage.

A

Amino acids with dissociable protons

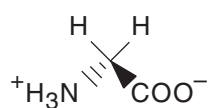


Other amino acids with polar side chains

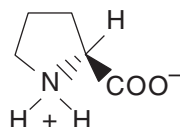


B

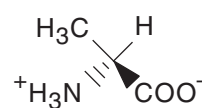
Nonpolar amino acids



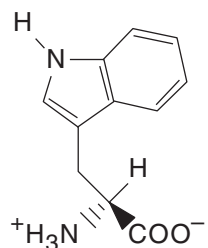
glycine



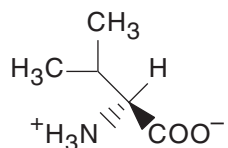
proline



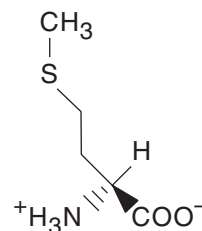
alanine



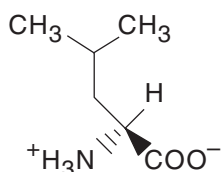
tryptophan



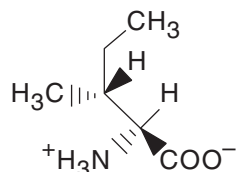
valine



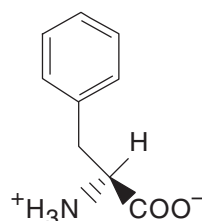
methionine



leucine



isoleucine



phenylalanine

Figure A.1C.2 Line drawings of the amino acids. The amino acids are roughly divided into three groups: amino acids with dissociable protons (**A**), other amino acids with polar side chains (**A**), and nonpolar amino acids (**B**). These groupings are designed to facilitate an understanding of enzymology and the thermodynamics of protein folding.

In this representation, hydrogens are omitted except in showing ionization or stereochemistry. In the case of arginine the delocalized positive charge is indicated by dashed double bonds. At stereocenters, bold lines indicate a group is coming out of the page toward the viewer, while hashed lines indicate that the group goes into the page away from the viewer.

Amino acids with dissociable protons are generally intimately involved in the chemistry of enzymes. Acidic and basic groups can form salt bridges to substrates or to each other. They can also act as proton donors/acceptors in mechanisms that rely on acid/base catalysis. The polar side chains of some of these amino acids (notably cysteine, serine, and histidine) can act as nucleophiles. The pK_a values for the free amino acids are shown, but these values

can markedly change when these groups are buried in proteins. The pK_a s of the α -amino groups range from 8.7 to 10.7, while the pK_a s of the α -carboxylates range from 1.8 to 2.4.

Amino acids with polar side chains (**A**) can form hydrogen bonds to substrates or to each other. Cysteine, serine, and tyrosine could also be included in this group, since the ionized forms of these amino acids do not generally perform structural roles in proteins. In general, these amino acids (and the amino acids with dissociable protons) will be found on the surfaces of proteins. Cysteine is an exception, since it is slightly hydrophobic and can often be buried as a disulfide bond.

The nonpolar amino acids (**B**) are often found in the interiors of proteins or in hydrophobic substrate-binding pockets. They interact with one another like jigsaw pieces, forming tight-fitting associations that have a density similar to that of an amino acid crystal. Proline is buried less frequently than might be expected because of its predominance in turns, which are often found on the periphery of a protein.

Standard Measurements, Data, and Abbreviations

A.1C.6

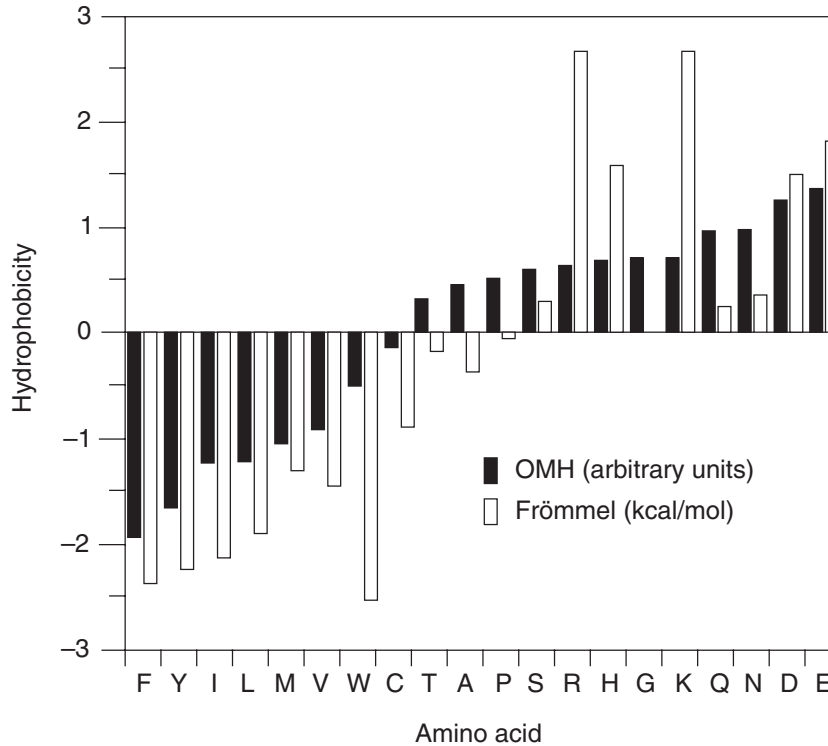


Figure A.1C.3 Amino acid hydrophobicity. The hydrophobicity of an amino acid is the degree to which it prefers a nonpolar medium, such as ethanol or the interior of a protein, to a polar medium, such as water. In this graph, the more hydrophobic amino acids “sink” below zero, while the more hydrophilic amino acids “float” above the surface.

Two scales are used. The Frömmel scale (Frömmel, 1984) represents the free energy of transfer from a hydrophobic medium to water. This value is an intrinsic property of an amino acid, separate from its role in a protein. In contrast, the OMH scale (Sweet and Eisenberg, 1983) is a measure of how likely a given amino acid will be replaced by a different hydrophobic or “buried” amino acid in a protein. In effect, this scale is how evolution views the hydrophobicity of an amino acid.

The distinction between physical and evolutionary properties is important. For example, while arginine is definitely a charged, polar amino acid (Sambrook et al., 1989), it can substitute more freely for nonpolar amino acids in the interior of a protein than glutamate (also a charged, polar amino acid) because of its long aliphatic side chain.

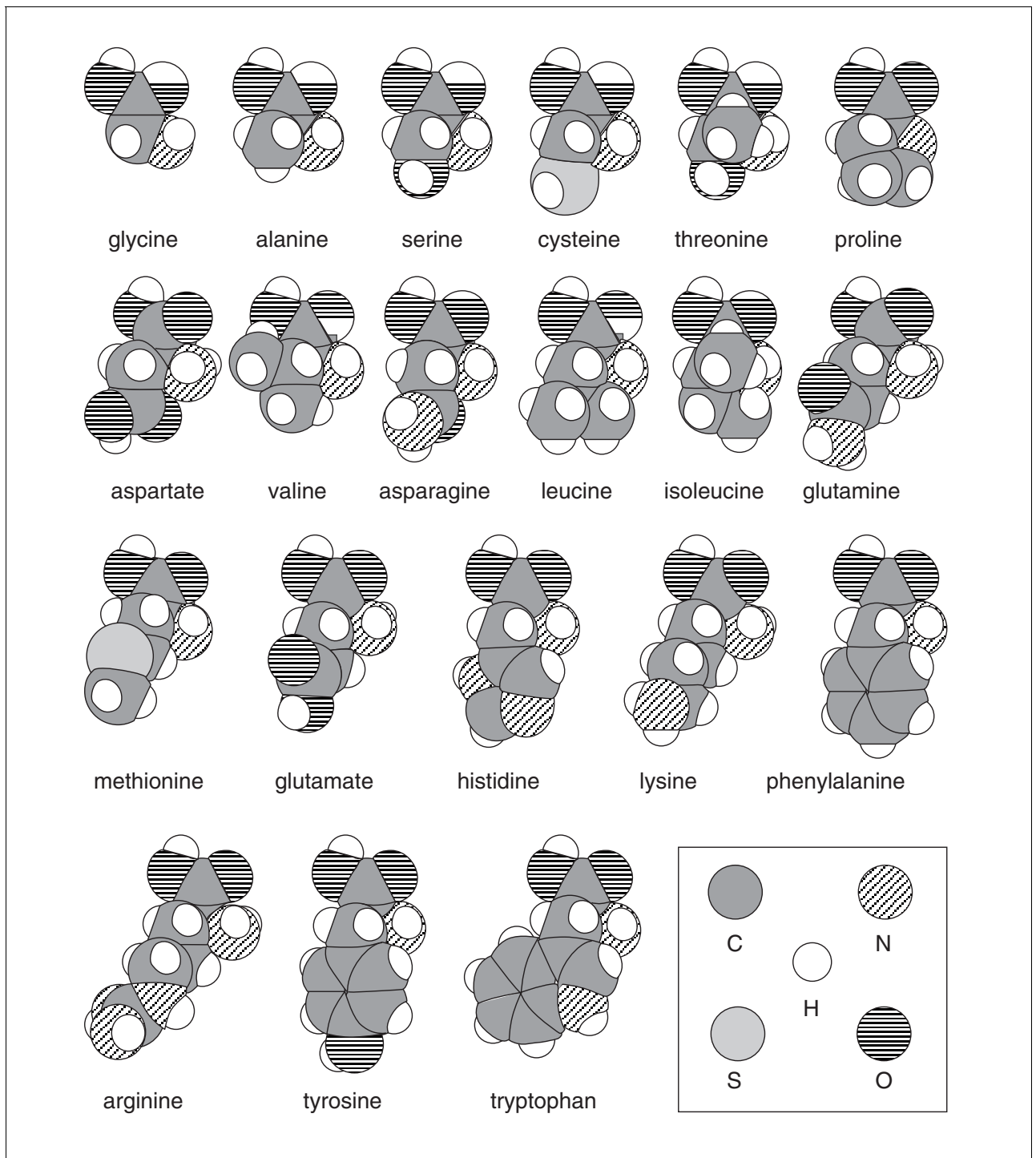


Figure A.1C.4 Space-filling representations of the amino acids. The amino acids are arranged in order of size. The conformations shown maximize the two-dimensional area but are not necessarily the most stable geometries.

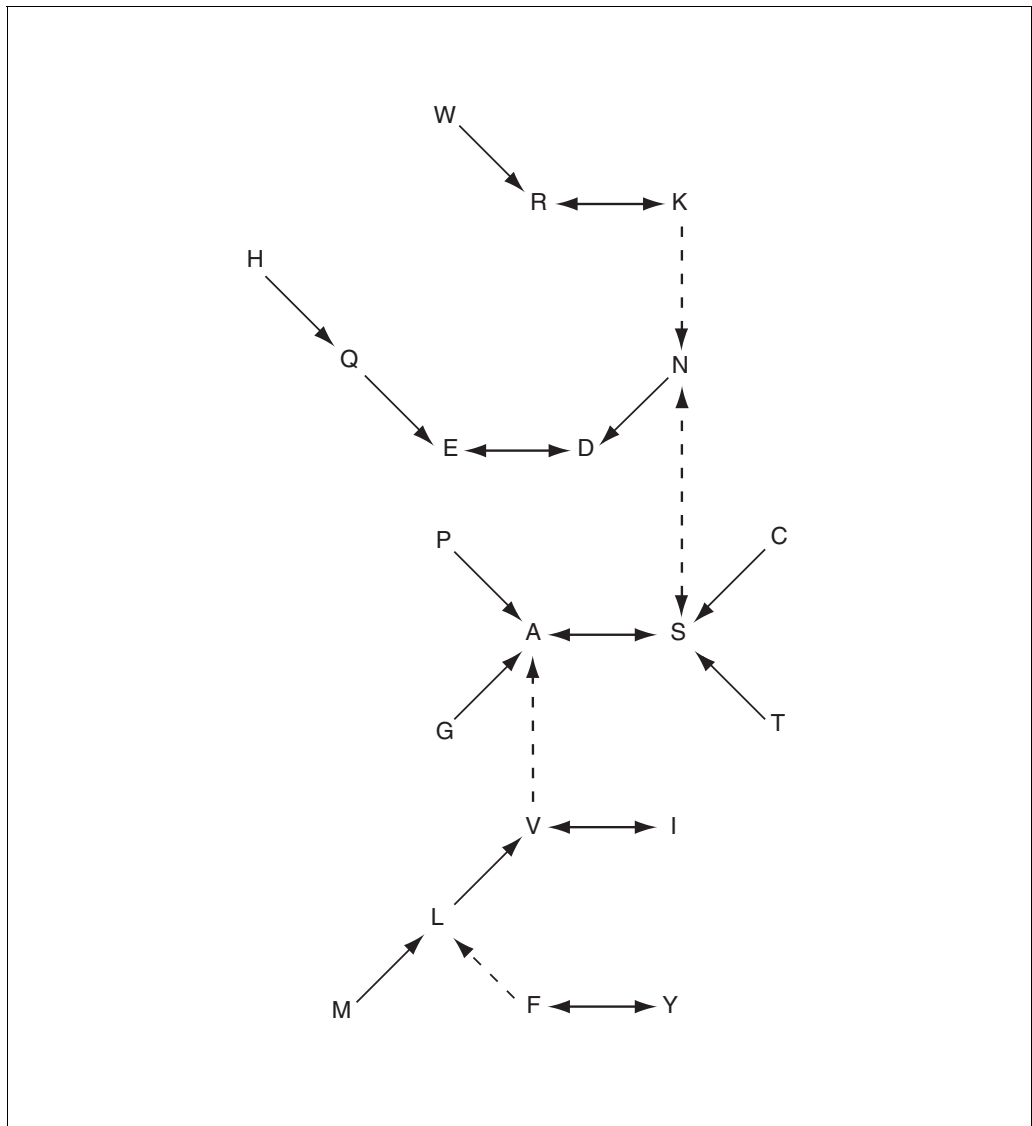


Figure A.1C.5 Mutational pathways for amino acids. In this diagram, amino acids are parsed into sets based on their ability to replace one another during the evolution of closely related proteins. Dark arrows show the *most frequent* mutational events for each of the twenty amino acids. For example, tryptophan most frequently mutates to arginine, while arginine and lysine most frequently replace one another. Dotted arrows represent the most frequent replacements *between* sets of otherwise mutationally related amino acids. Thus, while lysine mutates most frequently to arginine within the [arginine, lysine, tryptophan] set, the most likely event that will occur *outside* of this set is mutation to asparagine.

Table A.1C.3 Compositions and Masses of the Twenty Commonly Occurring Amino Acid Residues^{a,b}

Name	Composition	Monoisotopic mass	Average mass	pK _a of ionizing side chain ^c	Occurrence in proteins (%)
Alanine (Ala, A)	C ₃ H ₅ NO	71.03711	71.0788	—	8.3
Arginine (Arg, R)	C ₆ H ₁₂ N ₄ O	156.10111	156.1876	~11.5–12.5 (12)	5.7
Asparagine (Asn, N)	C ₄ H ₆ N ₂ O ₂	114.04293	114.1039	—	4.4
Aspartic acid (Asp, D)	C ₄ H ₅ NO ₃	115.02694	115.0886	3.9–4.5 (4)	5.3
Cysteine (Cys, C)	C ₃ H ₅ NOS	103.00919	103.1448	8.3–9.5 (9)	1.7
Glutamic acid (Glu, E)	C ₅ H ₇ NO ₃	129.04259	129.1155	4.3–4.5 (4.5)	6.2
Glutamine (Gln, Q)	C ₅ H ₈ N ₂ O ₂	128.05858	128.1308	—	4.0
Glycine (Gly, G)	C ₂ H ₃ NO	57.02146	57.0520	—	7.2
Histidine (His, H)	C ₆ H ₇ N ₃ O	137.05891	137.1412	6.0–7.0 (6.3)	2.2
Isoleucine (Ile, I)	C ₆ H ₁₁ NO	113.08406	113.1595	—	5.2
Leucine (Leu, L)	C ₆ H ₁₁ NO	113.08406	113.1595	—	9.0
Lysine (Lys, K)	C ₆ H ₁₂ N ₂ O	128.09496	128.1742	10.4–11.1 (10.4)	5.7
Methionine (Met, M)	C ₅ H ₉ NOS	131.04049	131.1986	—	2.4
Phenylalanine (Phe, F)	C ₉ H ₉ NO	147.06841	147.1766	—	3.9
Proline (Pro, P)	C ₅ H ₇ NO	97.05276	97.1167	—	5.1
Serine (Ser, S)	C ₃ H ₅ NO ₂	87.03203	87.0782	—	6.9
Threonine (Thr, T)	C ₄ H ₇ NO ₂	101.04768	101.1051	—	5.8
Tryptophan (Trp, W)	C ₁₁ H ₁₀ N ₂ O	186.07931	186.2133	—	1.3
Tyrosine (Tyr, Y)	C ₉ H ₉ NO ₂	163.06333	163.1760	9.7–10.1 (10.0)	3.2
Valine (Val, V)	C ₅ H ₉ NO	99.06841	99.1326	—	6.6

^aFor corresponding structures, see Figure A.1C.2 and Figure A.1C.4.

^bThe molecular mass of a normally terminated and unmodified peptide or protein may be calculated by summing the masses of the appropriate amino acid residues and adding the masses of H and OH for the N and C termini, respectively. In cases where cysteines are linked to form disulfide bridges, the mass of two hydrogen atoms should be subtracted for each disulfide bridge in the molecule. Specifically, monoisotopic masses were calculated using the atomic masses of the most abundant isotope of the elements: C = 12.0000000, H = 1.0078250, N = 14.0030740, O = 15.9949146, and S = 31.9720718. Average masses were calculated using the atomic weights of the elements: C = 12.011, H = 1.00794, N = 14.00674, O = 15.9994, and S = 32.066.

^cThese values are included for anyone wishing to make approximate isoelectric point determinations based on protein composition. Values for the terminal residues depend on the identity of the residue: α-amino, pK_a 6.8–8.2 (8.0); α-carboxyl, pK_a 3.2–4.3 (3.6). Values in parentheses are based on those given by Matthew et al. (1978) and provide a good starting point for determinations.

Table A.1C.4 Mass Changes Due to Some Post-Translational Modifications of Peptides and Proteins^a

Modification ^b	Monoisotopic mass change	Average mass change
<i>Common modifications</i>		
Pyroglutamic acid formation from Gln	-17.0265	-17.0306
Disulfide bond (cystine) formation	-2.0157	-2.0159
C-terminal amide formation from Gly	-0.9840	-0.9847
Deamidation of Asn and Gln	-0.9840	-0.9847
Methylation	14.0157	14.0269
Hydroxylation	15.9949	15.9994
Oxidation of Met	15.9949	15.9994
Proteolysis of a single peptide bond	18.0106	18.0153
Formylation	27.9949	28.0104
Acetylation	42.0106	42.0373
Carboxylation of Asp and Glu	43.9898	44.0098
Phosphorylation	79.9663	79.9799
Sulfation	79.9568	80.0642
Cysteinylation	119.0041	119.1442
Glycosylation with pentoses (Ara, Rib, Xyl)	132.0423	132.1161
Glycosylation with deoxyhexoses (Fuc, Rha)	146.0579	146.1430
Glycosylation with hexosamines (GalN, GlcN)	161.0688	161.1577
Glycosylation with hexoses (Fru, Gal, Glc, Man)	162.0528	162.1424
Modification with lipoic acid (amide bond to lysine)	188.0330	188.3147
Glycosylation with <i>N</i> -acetylhexosamines (GalNAc, GlcNAc)	203.0794	203.1950
Farnesylation	204.1878	204.3556
Myristoylation	210.1984	210.3598
Biotinylation (amide bond to lysine)	226.0776	226.2994
Modification with pyridoxal phosphate (Schiff base to lysine)	231.0297	231.1449
Palmitoylation	238.2297	238.4136
Stearoylation	266.2610	266.4674
Geranylgeranylation	272.2504	272.4741
Glycosylation with <i>N</i> -acetylneuraminic acid (sialic acid, NeuAc, NANA, SA)	291.0954	291.2579
Glutathionylation	305.0682	305.3117
Glycosylation with <i>N</i> -glycolylneuraminic acid (NeuGe)	307.0903	307.2573
5'-Adenosylation	329.0525	329.2091
Modification with 4'-phosphopantetheine	339.0780	339.3294
ADP-ribosylation (from NAD)	541.0611	541.3052
<i>Adventitious modifications</i>		
Acrylamide	71.0371	71.0788
Glutathione	304.0712	304.3038
2-Mercaptoethanol	75.9983	76.1192

^aTo obtain the molecular mass of a modified peptide or protein, the appropriate mass changes should be algebraically added to the molecular mass calculated for the unmodified molecule.

^bA more extensive list of modifications is available from the Delta mass site at <http://www.medstv.unimelb.edu.au/WWWDOCS/SVIMRDOCS/MassSpec/deltamassV2.html>.

LITERATURE CITED

- Chothia, C. 1976. The nature of the accessible and buried surfaces in proteins. *J. Mol. Biol.* 105:1-14.
- Dayhoff, M.O., Schwartz, R.M., and Orcutt, B.C. 1978. A model of evolutionary change in proteins. *In Atlas of Protein Sequence and Structure* (M. Dayhoff, ed.) Vol. 5, pp. 345-352. National Biomedical Research Foundation, Washington, D.C.
- Frömmel, C. 1984. The apolar surface area of amino acids and its empirical correlation with hydrophobic free energy. *J. Theor. Biol.* 111:247-260.
- Rose, G.D., Geselowitz, A.R., Lesser, G.J., Lee, R.H., and Zehfus, M.H. 1985. Hydrophobicity of amino acid residues in globular proteins. *Science* 229:834-838.
- Matthew, J.B., Friend, S.H., Botelho, L.H., Lehman, L.D., Hanania, G.I.H., and Gurd, F.R.N. 1979. *Biochem. Biophys. Res. Commun.* 81:416-421.
- Sambrook, J., Fritsch, E.F., and Maniatis, T.M. (eds.). 1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.
- Sharp, P.M., Cowe, E., Higgins, D.G., Shields, D.C., Wolfe, K.H., and Wright, F. 1988. Codon usage patterns in *E. coli*, *B. subtilis*, *S. cerevisiae*, *S. pombe*, *D. melanogaster*, and *H. sapiens*: A review of the considerable within-species diversity. *Nucl. Acids Res.* 16: 8207-8211.
- Sweet, R.M. and Eisenberg, D. 1983. Correlation of sequence hydrophobicities measures similarity in three-dimensional protein structure. *J. Mol. Biol.* 171:479-488.
- Wada, K.-N., Aota, S.-I., Tsuchiya, R., Ishibashi, F., Gojobori, T., and Ikemura, T. 1990. Codon usage tabulated from the GenBank genetic sequence data. *Nucl. Acids Res.* 18 (Suppl.):2367-2411.

Contributed by Andrew Ellington and J. Michael Cherry (codon usage)
Massachusetts General Hospital
Boston, Massachusetts

Characteristics of Nucleic Acids

Nucleic acids have traditionally been regarded solely as informational macromolecules with limited secondary structural features, but recent discoveries have vastly expanded the repertoire of polynucleotide structure and function. While this manual has in general concentrated on how to manipulate DNA and RNA, this section details their structural features. In addition, some experimentally useful properties of the mononucleotide building blocks are listed in Table A.1D.1. The chemical structures of the mononucleotides can be seen in Figure A.1D.1, and aspects of nucleotide stereochemistry that are important to an understanding of base pairing and secondary structure can be found in Figure A.1D.2. Although Watson-Crick pairings play a critical role in defining nucleic acid secondary structures, a wide variety of alternative base pairings can be important in higher order conformations; some of these are detailed in Figure A.1D.3. Finally, even given only Watson-Crick-style pairings, secondary structures with significantly different features can be formed. Figure A.1D.4 overviews the differences between A-, B-, and Z-form helices.

Table A.1D.1 Physical Characteristics of the Nucleotides^a

Nucleotide	Mol. wt. (g/mol)	λ_{\max} (nm)	λ_{\min} (nm)	ϵ_{\max} (mM ⁻¹ cm ⁻¹)	A_{280}/A_{260}	TLC mobility ^b		
						A	B	C
ATP	507.2	259	227	15.4	0.15	0	6	34
ADP	427.2	259	227	15.4	0.16	0	26	54
AMP	347.2	259	227	15.4	0.16	11	52	65
Adenosine ^c	267.2	260	227	14.9	0.14	—	—	—
dATP ^d	491.2	259	226	15.4	0.15	0	—	35
dAMP ^d	331.2	259	226	15.2	0.15	11	52	—
dA	251.2	260	225	15.2	0.15	—	—	—
CTP	483.2	271	249	9.0	0.97	0	11	41
CDP	403.2	271	249	9.1	0.98	0	33	64
CMP	323.2	271	249	9.1	0.98	15	64	75
Cytidine	243.2	271	250	9.1	0.93	—	—	—
dCTP ^d	467.2	272	—	9.1	0.98	0	—	43
dCMP	307.2	271	249	9.3	0.99	18	65	—
dC	227.2	271	250	9.0	0.97	—	—	—
GTP	523.2	253	223	13.7	0.66	0	5	25
GDP	443.2	253	224	13.7	0.66	0	17	45
GMP	363.2	252	224	13.7	0.66	6	40	51
Guanosine ^c	283.2	253	223	13.6	0.67	—	—	—
dGTP ^d	507.2	252	222	13.7	0.66	0	—	26
dGMP ^d	347.2	253	222	13.7	0.67	6	41	—
dG	267.2	254	223	13.0	0.68	—	—	—
UTP	484.2	262	230	10.0	0.38	0	14	49
UDP	404.2	262	230	10.0	0.39	0	41	71
UMP	324.2	262	230	10.0	0.39	20	75	80
Uridine	244.2	262	230	10.1	0.35	—	—	—
TTP ^d	482.2	267	—	9.6	0.73	0	—	52
TMP ^d	322.2	267	234	9.6	0.73	24	74	—
Thymidine ^d	242.2	267	235	9.7	0.70	—	—	—

^a Spectral data are assembled from Fasman (1975) at pH 7.0 except where footnoted otherwise.

^b TLC mobility is expressed as the percent distance a given spot migrates relative to the solvent front (*R_f*) in three different TLC systems using 0.5-mm polyethylenimine cellulose plates: "A" is 0.25 M LiCl, "B" is 1.0 M LiCl, and "C" is 1.6 M LiCl.

^c Spectral measurements taken at pH 6.0.

^d Spectral data assembled from Dawson et al. (1987).

A.1D.1

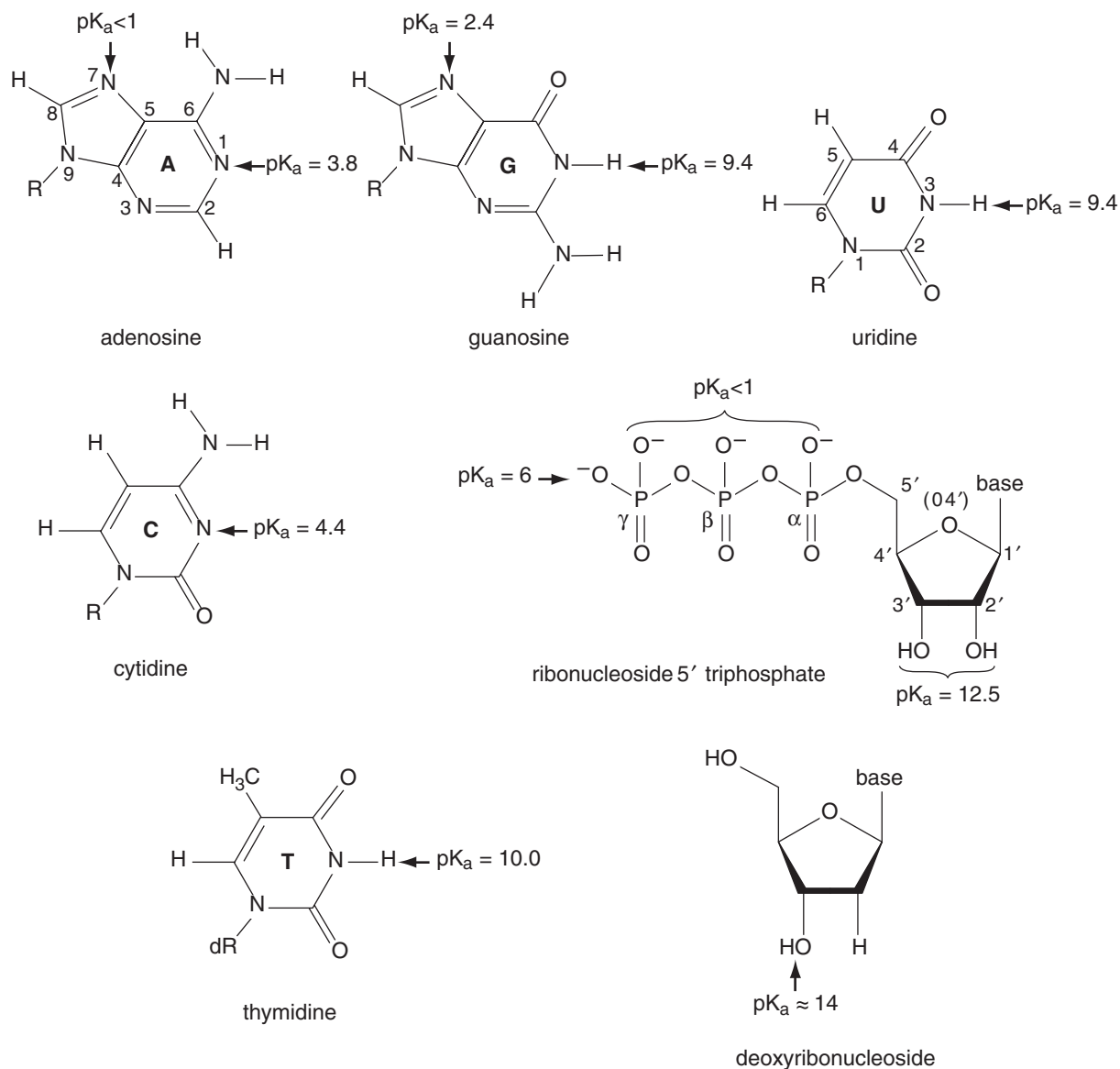


Figure A.1D.1 Line drawings of the nucleotides. The chemical structure that predominates at neutral pH is shown. Drawings of the nucleotide bases and their associated sugars, either ribose (R) or deoxyribose (dR), are shown separately. In the representations of ribose (as a nucleoside triphosphate) and deoxyribose (as a nucleotide), the bold lines indicate that this portion of the sugar is coming out of the page toward the reader. In this view, the base is found above the plane of the sugar, while the 3' hydroxyl group is found below the plane of the sugar.

The pK_a values for all groups are shown; pK_a s above 7 imply proton dissociation from the pictured structure, while pK_a s below 7 imply proton association to the pictured structure. The tautomeric form of a given base may change at different

pH values. The pK_a values given are for nucleotide mono-phosphates and were taken from Dawson et al. (1987); a fuller discussion of the chemical basis for these values can be found in a review by T'so (1974).

The small numbers adjacent to adenosine, uridine, and ribose indicate the nomenclature of the purines, pyrimidines, and sugars, respectively. Groups appended to a ring have the same numbering as the position to which they are linked; thus, the "O6" moiety of guanosine is the carbonyl oxygen bonded to C6 in the ring. Similarly, "O3" on ribose or deoxyribose indicates the oxygen of the hydroxyl group bonded to C3' in the ring. The α , β , and γ phosphates in a nucleoside triphosphate are also indicated.

A.1D.2

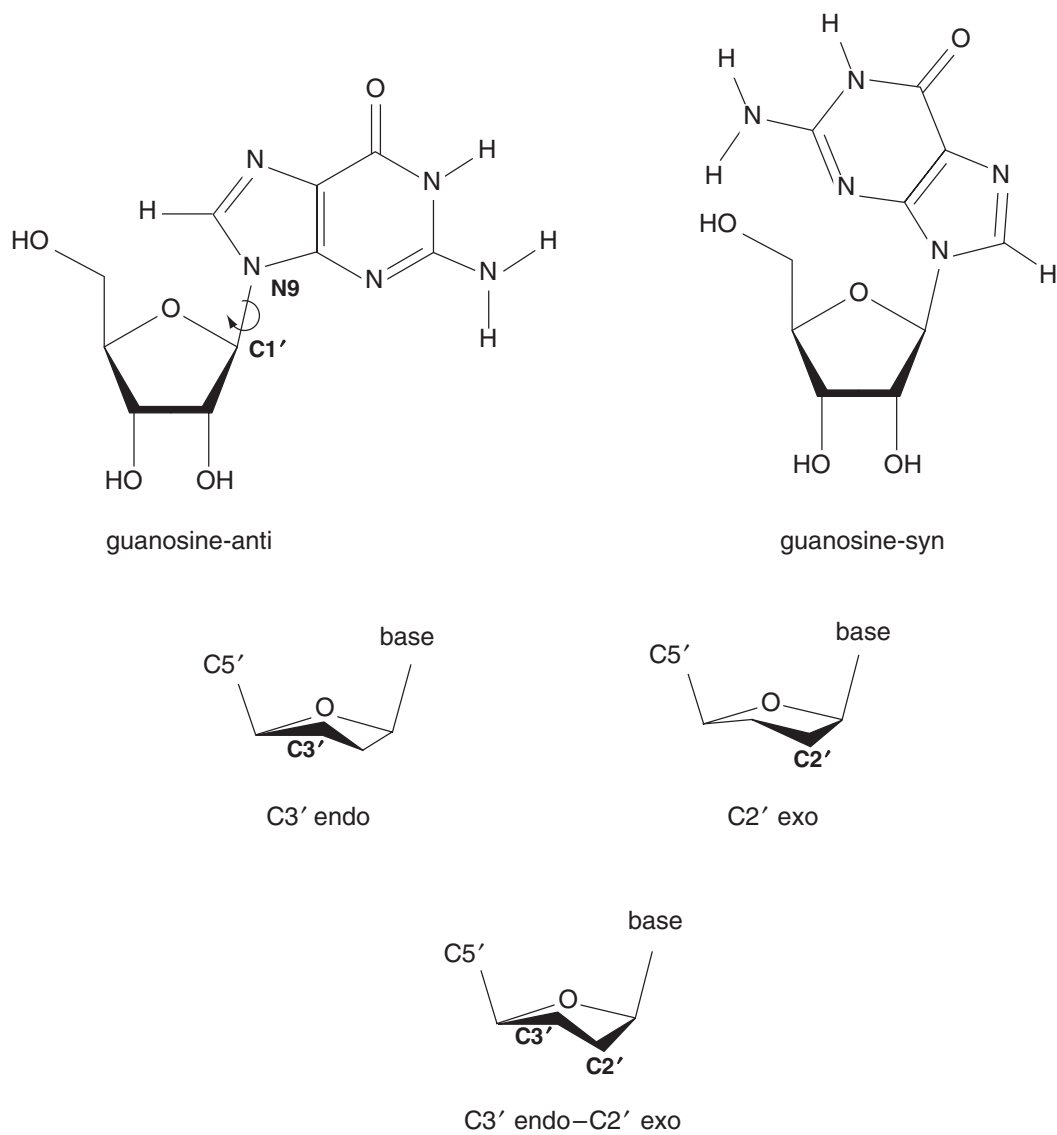
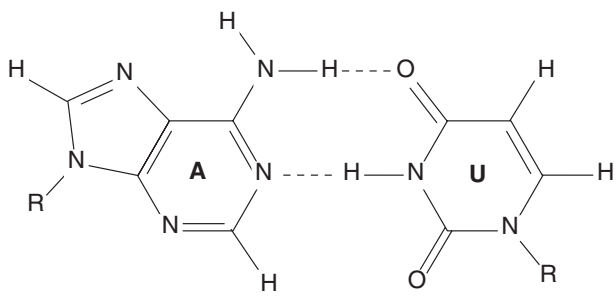
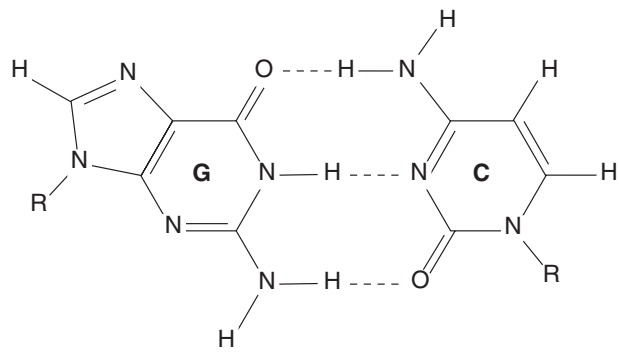


Figure A.1D.2 Nucleotide stereochemistry. Depending on the rotation about the bond between C1' of the sugar and either N1 (for pyrimidines) or N9 (for purines), a nucleotide can be described as either "anti" or "syn." Because of steric constraints, nucleotides are generally found in the "anti" configuration, with their Watson-Crick hydrogen bond donor-acceptors swung outward away from the plane of the sugar ring. However, guanosine is sometimes found in a "syn" configuration, both in polynucleotides and in solution. In this form, the bulk of the purine ring is positioned directly over the plane of the sugar. The sugar ring can also adopt

different stereochemistries. These are labeled according to which group is bent out of the plane of the ring, and in which direction. If a portion of the ring is bent "upward" toward the base, this is known as "endo," while if it is bent "downward" away from the base, this is known as "exo." In the figure, plain lines represent bonds that are within the plane of the sugar, while bold lines indicate that the bond is bent out of the plane. Hence, "C3' endo-C2' exo" describes a furanose ring in which the 2' and 3' carbons have been twisted in opposite directions and the bond connecting them crosses the plane of the ring.

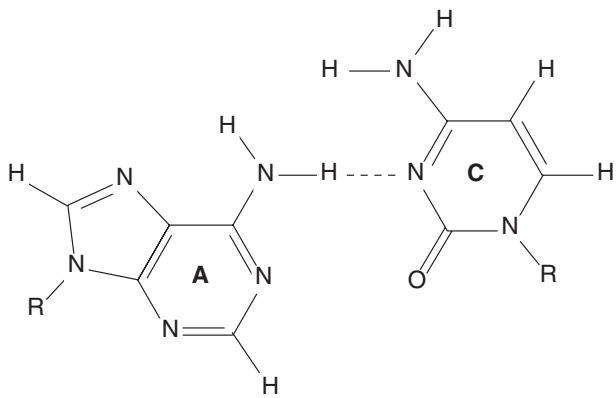


A:U

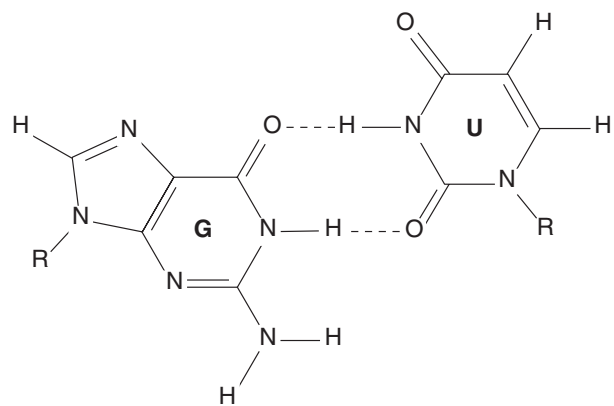


G:C

Watson-Crick pairings

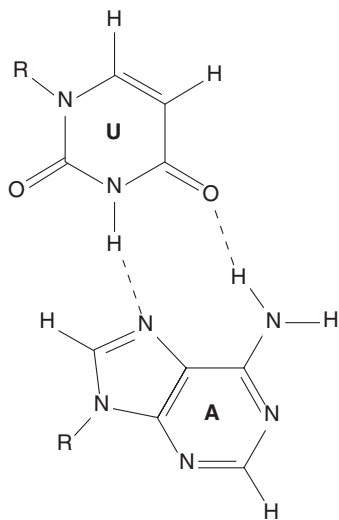


A:C

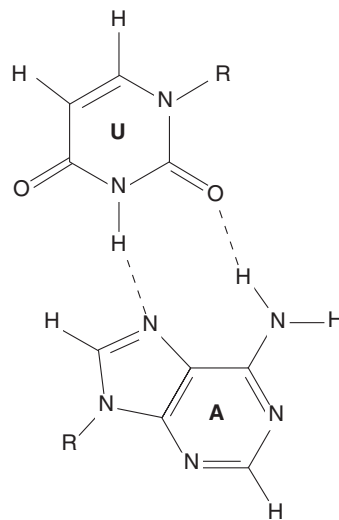


G:U

Wobble pairings



Hoogsteen
(parallel chains)



Reverse Hoogsteen
(antiparallel chains)

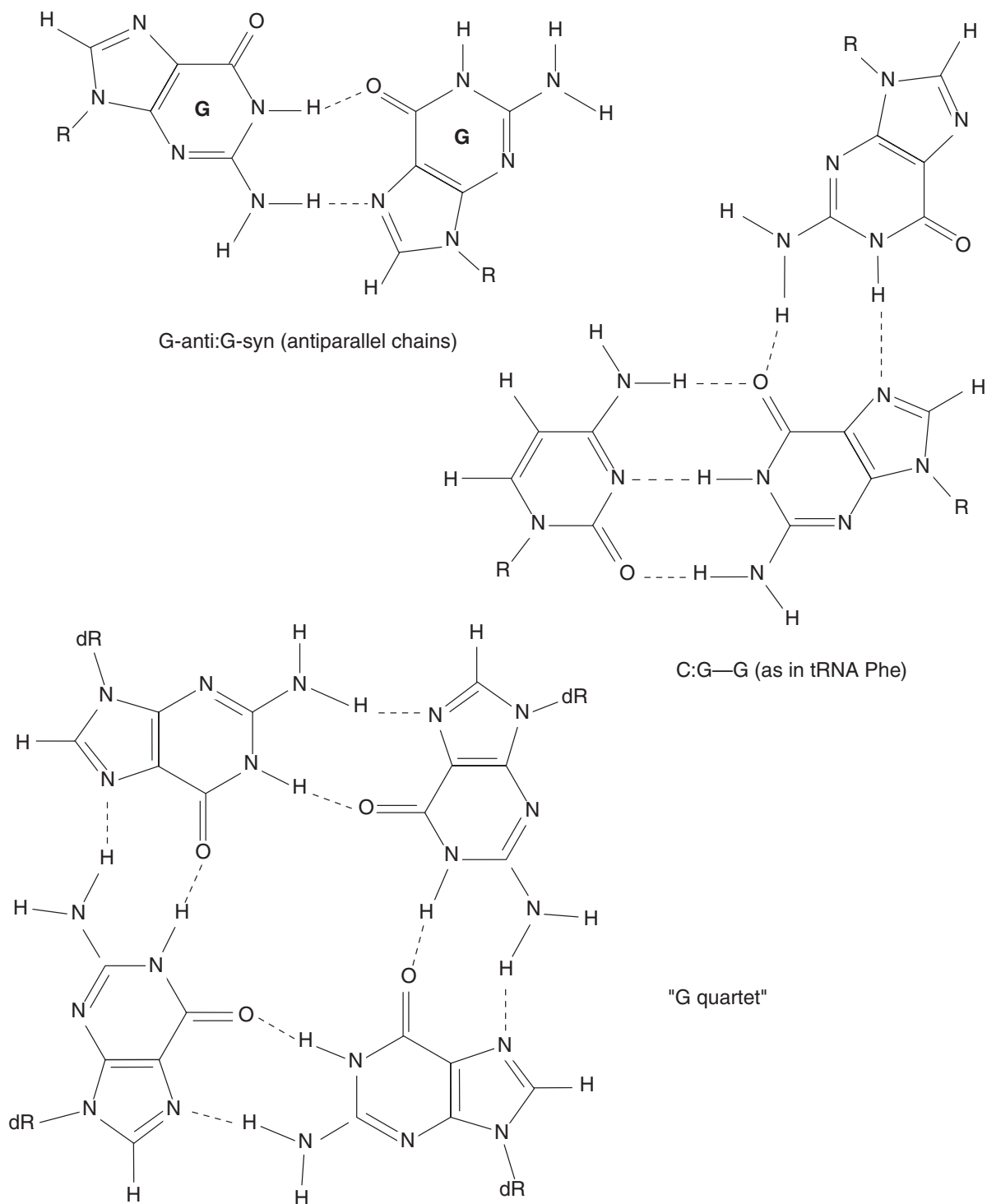


Figure A.1D.3 Base pairing schemas. The chemical structures of the nucleotide bases determine the formation of secondary and tertiary structures in nucleic acids. A wide variety of hydrogen bonding schemas (indicated by dashed lines) are possible between different bases. Watson-Crick pairings are perhaps the most widely known and are the basis of the double helical structure of complementary, anti-parallel DNA strands. Other base pairs can also be accommodated within the double helix, such as “wobble pairings,” in which the bases are slightly off-center with respect to each other. By using the N7 hydrogen bond

acceptor of the purine bases adenosine and guanosine, an even wider variety of structures becomes possible, including Hoogsteen base pairs and a G-G pairing in which one of the guanosine residues assumes a “syn” conformation. Bonds involving N7 of the purine bases allow tertiary structural interactions to occur in nucleic acids, including triple base pairs (such as those found in tRNA) and the recently described “G quartet” (Sen and Gilbert, 1988). A discussion of the structural possibilities of base pairing can be found in Saenger’s superlative book, *Principles of Nucleic Acid Structure* (1993).

A.1D.5

B DNA

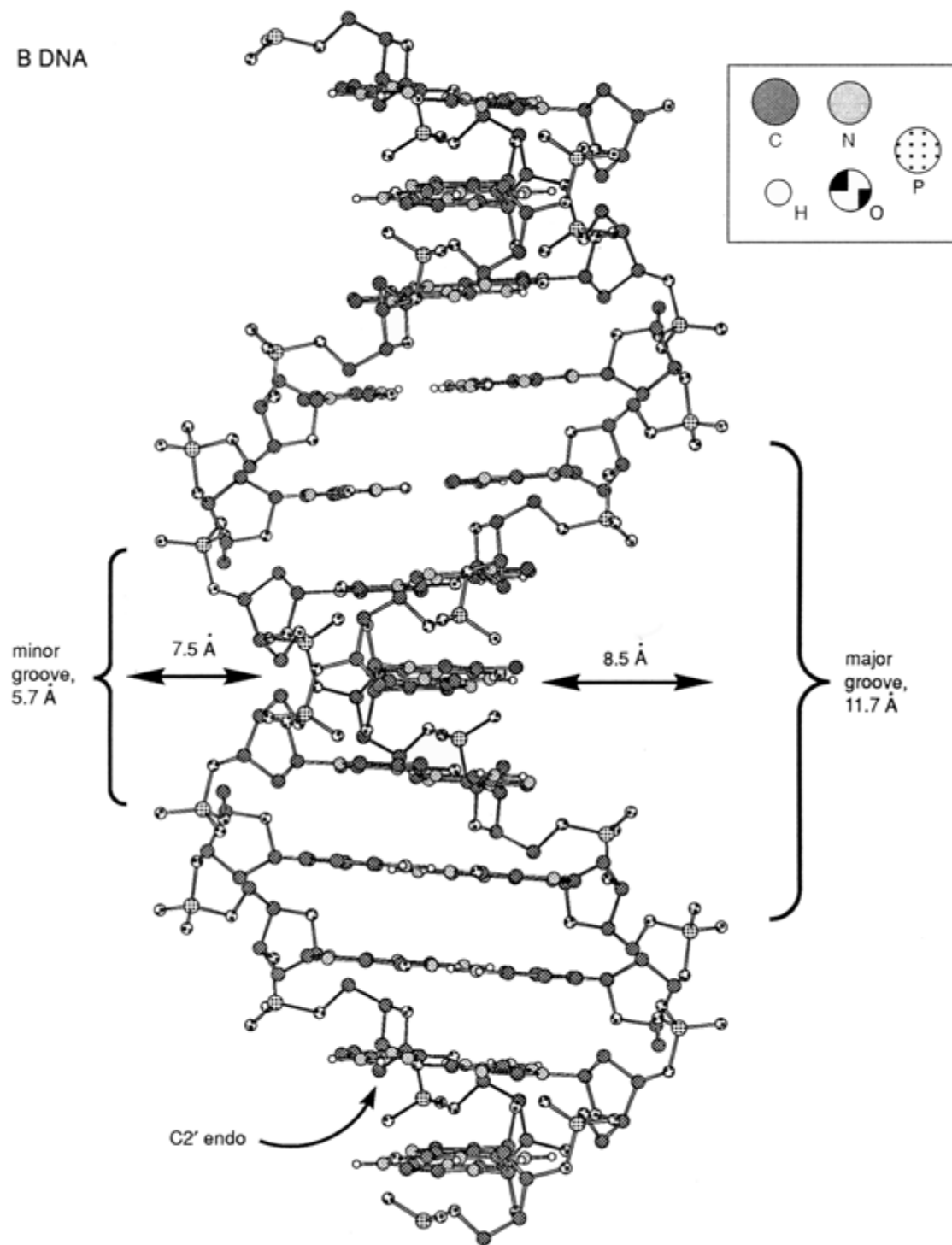
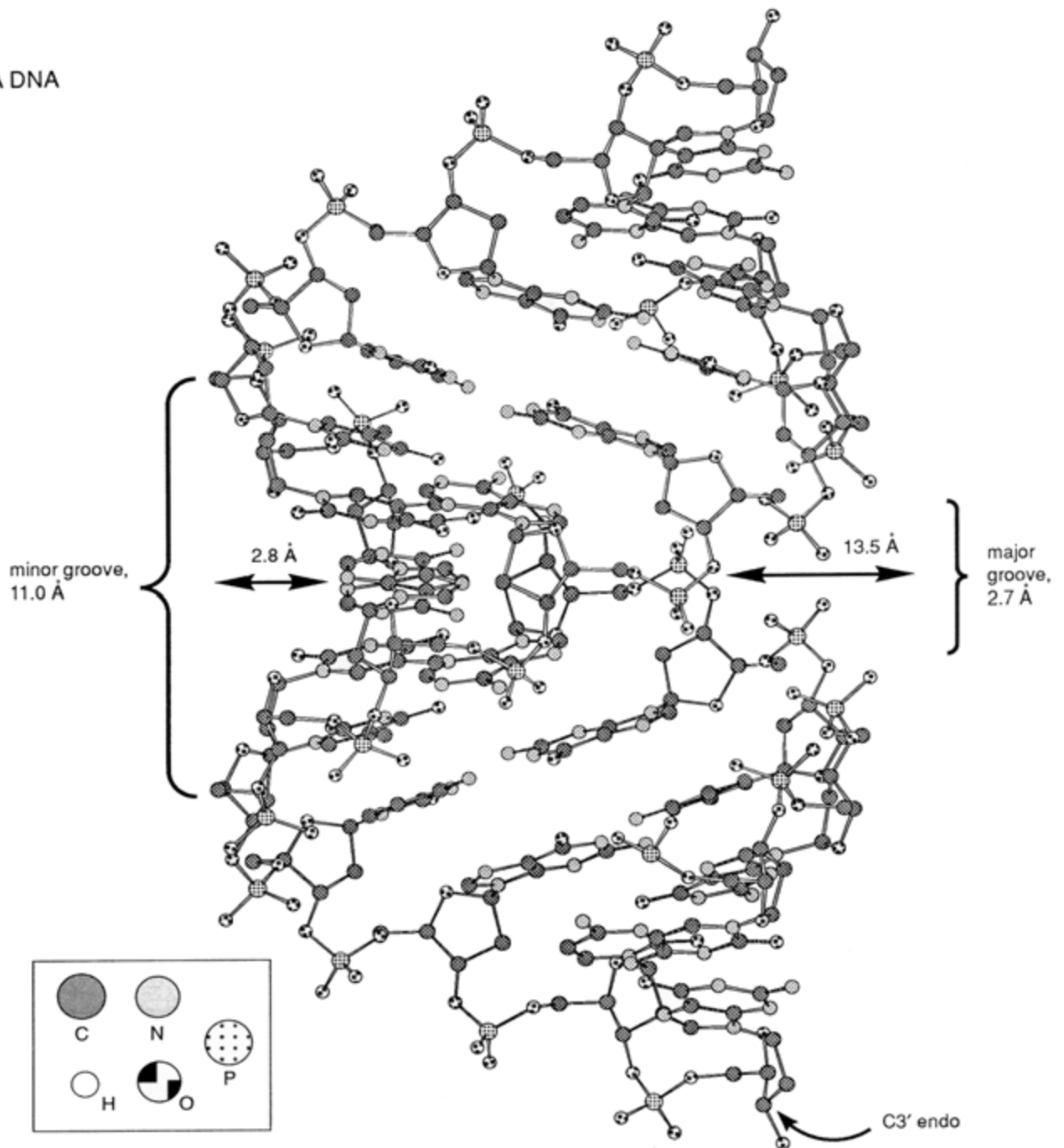


Figure A.1D.4 Nucleic acid secondary structures. The structural consequence of the ability of nucleotides to form Watson-Crick base pairs is nucleic acid double helices. In this figure, the self-complementary 12-mer CGCGAATTCGCG is shown as both A- and B-form helices. Two representations of the A helix have been shown in order to emphasize the depth of the major groove. The arrows and brackets in these figures are not drawn to scale.

While both of these helices are right-handed (in terms of anthropomorphic referents, if you were to point your thumb along a strand in a 5' to 3' manner, the twist of the helix would be the same as the curl of your right hand), their structural details are very different: B DNA has roughly 10 bases per full turn, while A DNA and A RNA have 11 to 12; the major groove of B-form helices is wide and the minor groove is narrow, while for A-form helices this is reversed; in

A.1D.6

A DNA



B-forms the base pairs are located close to the helix axis (as can be seen in end-on views), while in A-forms the base pairs are pushed out away from the long helical axis, leaving a “hole” in the middle of the polynucleotide coil (if one imagines DNA as a flat ribbon, then B DNA is twisted from its ends, while A DNA is coiled on itself).

Different helical forms are largely due to differences in sugar stereochemistry. Examples of a 2' endo deoxyribose (found in B DNA) and a 3' endo deoxyribose (found in A DNA) are indicated.

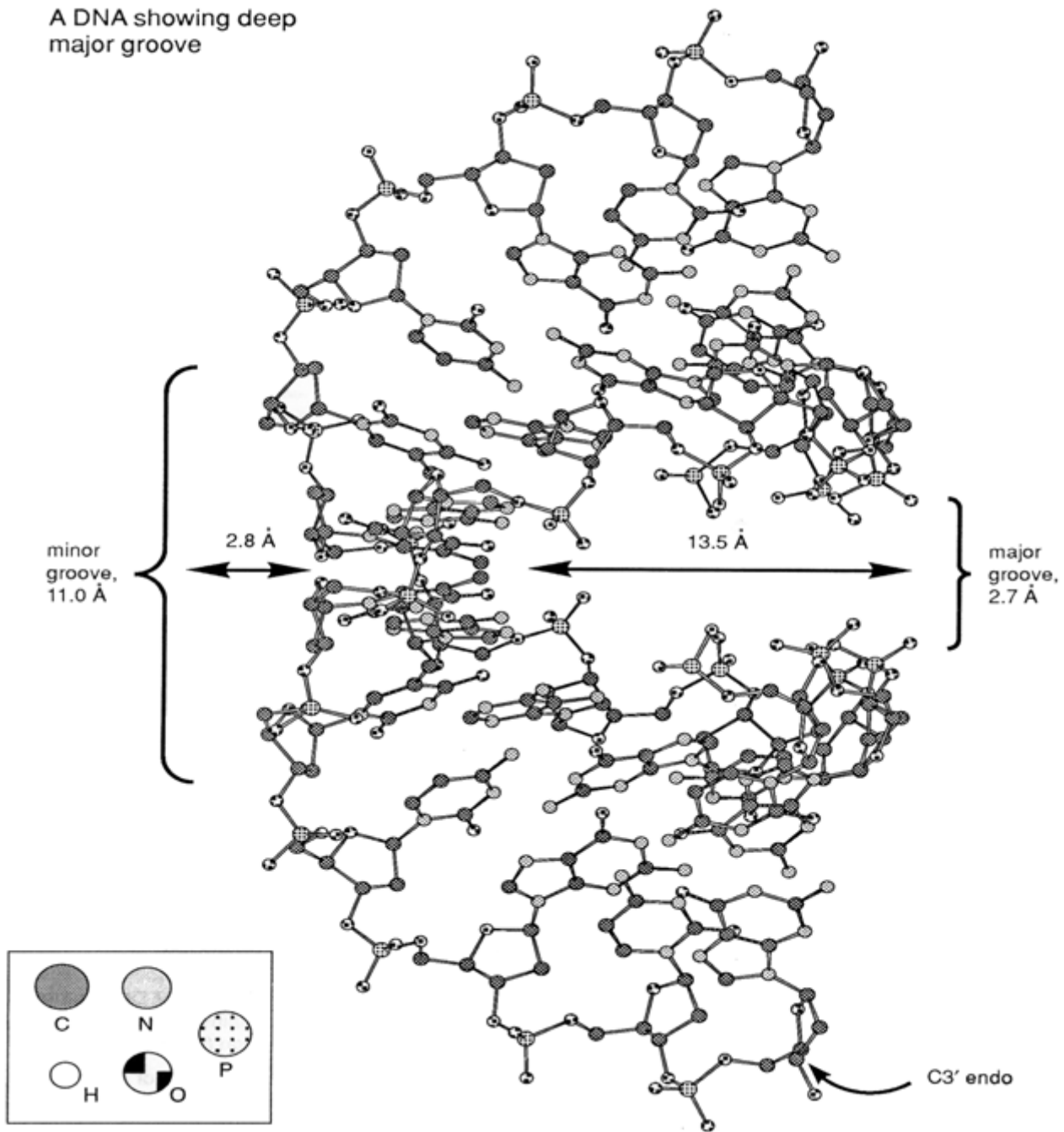
While there are a variety of other helical forms, the most

striking is that found in Z DNA. The Z DNA coil is left- rather than right-handed and contains G:C base pairs where the G is in the “syn” conformation (shown in the inset).

The uneven progression, or zigzag, of Z DNA can be more easily seen when the polynucleotide backbone is shown in isolation; the inset shows the connectivity between phosphates by 5' to 3' vector arrows. Because of its odd shape, base pairs actually protrude from what would be a cavity in A or B DNA; thus, Z DNA has a minor but no major groove. This diagram is based on the original structure of alternating C:G/G:C base pairs (Wang et al., 1979).

A.1D.7

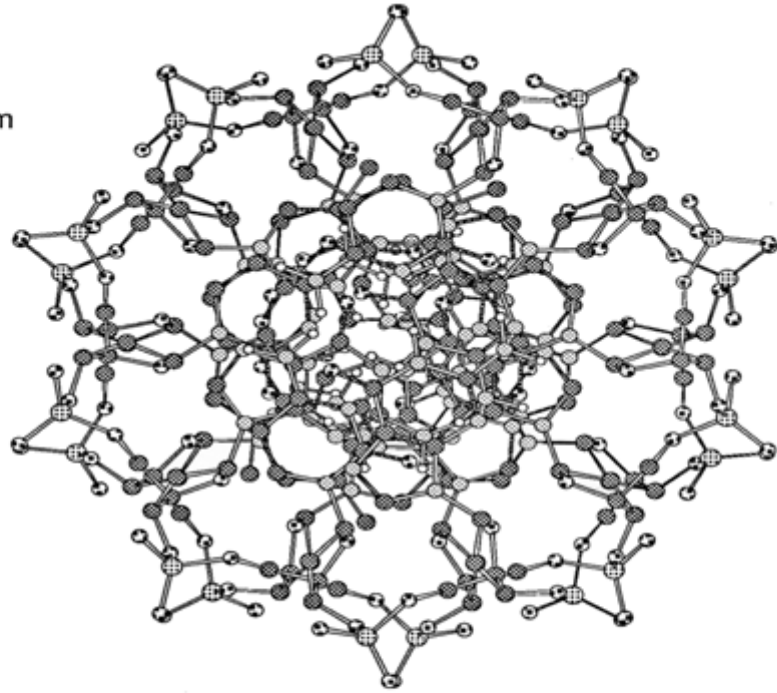
A DNA showing deep major groove



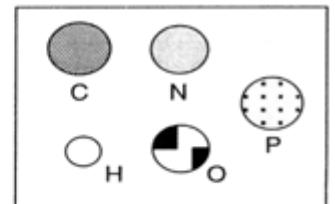
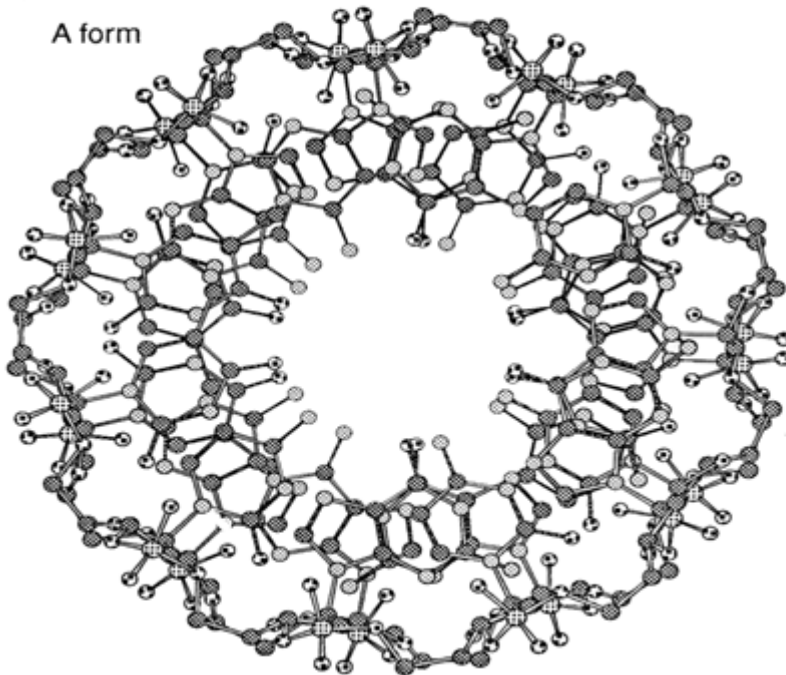
Characteristics of Nucleic Acids

A.1D.8

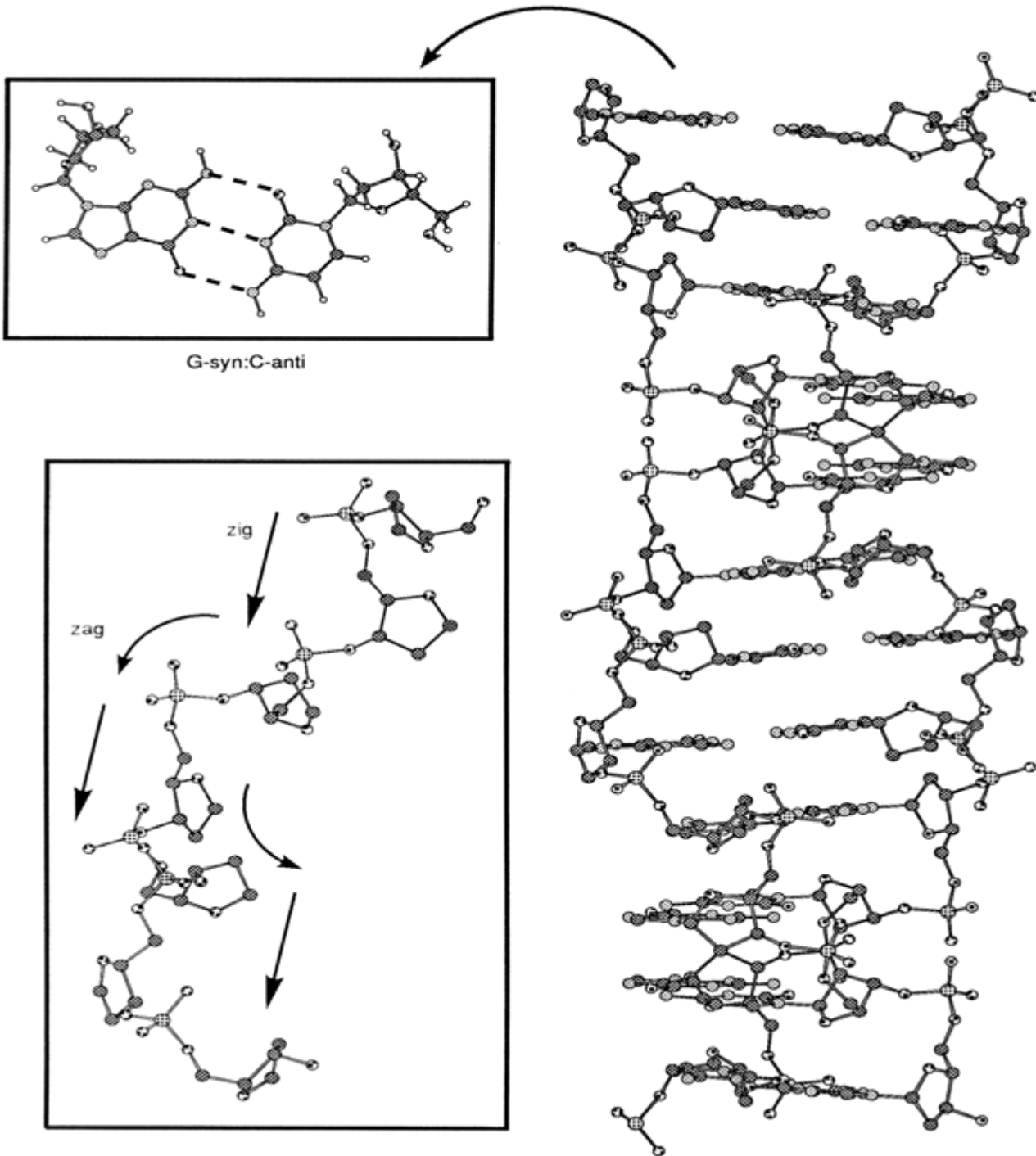
B form



A form



Z DNA



Characteristics of Nucleic Acids

A.1D.10

LITERATURE CITED

- Dawson, M.C., Elliott, D.C., Elliott, W.H., and Jones, K.M. (eds.). 1987. *Data for Biochemical Research*, 3rd ed. Clarendon Press, Oxford.
- Fasman, G. (ed.). 1975. *Handbook of Biochemistry and Molecular Biology, Vol. 1: Nucleic Acids*, 3rd ed. CRC Press, Boca Raton, Fla.
- Saenger, W. 1993. *Principles of Nucleic Acid Structure*. Springer-Verlag, New York.
- Sen, D. and Gilbert, W. 1988. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* 334:364-366.
- T'so, P.O.P. 1974. Bases, nucleosides, and nucleotides. *In Basic Principles in Nucleic Acid Chemistry, Vol. 1* (P.O.P. T'so, ed.) pp. 453-584. Academic Press, San Diego.
- Wang, A.H., Quigley, G.J., Kolpak, F.J., Crawford, J.L., van Boom, J.H., van der Marel, G., and Rich, A. 1979. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* 282:680-686.

Radioactivity

The use of radioactive materials (Tables A.1E.1 and A.1E.2) in many protocols in this manual raises the issue of safety. Because the real hazards of low levels of radiation are not known, it is generally assumed that any unavoidable exposure is too much. It is important to know applicable regulations and approved procedures on the use of isotopes, which will certainly include using gloves when handling any radioactive materials and appropriate shielding (Table A.1E.3) when handling high levels of energetic isotopes such as ^{32}P .

In addition, investigators should check themselves and the work area frequently for radioactivity with a hand-held minimonitor. It is just as important to have a clear understanding of the properties of radioisotopes and to use common sense (Table A.1E.4; Figure A.1E.1). For example, working behind an unwieldy shield can significantly increase the likelihood of a spill. One gets less exposure by forgetting about the shield for the submicrocurie amounts of ^{32}P in a final ethanol precipitation than by dropping the pellet and having to start the labeling reaction over again. The straightforward application of common sense must sometimes be tempered by consideration of the requirements of maintaining a radiation license, however. Although it may seem highly unlikely a labeling reaction will be spilled into your coffee cup, the rule against having them on the laboratory bench altogether is not a matter to be negotiated between you and your radiation safety officer. Your authority to use radioisotopes is contingent on following the rules set up by the appropriate regulatory agency. For further information, refer to the Guide for Safe Handling of Radioisotopes (APPENDIX 1F).

Table A.1E.1 Conversion Factors for Radioactivity

Measurement of Radioactivity

The SI unit for measurement of radioactivity is the Becquerel:

$$1 \text{ Becquerel (Bq)} = 1 \text{ disintegration per second}$$

The more commonly encountered unit is the Curie (Ci):

$$\begin{aligned} 1 \text{ Ci} &= 3.7 \times 10^{10} \text{ Bq} \\ &= 2.22 \times 10^{12} \text{ disintegrations per minute (dpm)} \\ 1 \text{ millicurie (mCi)} &= 3.7 \times 10^7 \text{ Bq} = 2.22 \times 10^9 \text{ dpm} \\ 1 \text{ microcurie } (\mu\text{Ci}) &= 3.7 \times 10^4 \text{ Bq} = 2.22 \times 10^6 \text{ dpm} \end{aligned}$$

Conversion factors:

$$\begin{aligned} 1 \text{ day} &= 1.44 \times 10^3 \text{ min} = 8.64 \times 10^4 \text{ sec} \\ 1 \text{ year} &= 5.26 \times 10^5 \text{ min} = 3.16 \times 10^7 \text{ sec} \\ \text{counts per minute (cpm)} &= \text{dpm} \times (\text{counting efficiency}) \end{aligned}$$

Measurement of Dose

The SI unit for energy absorbed from radiation is the Gray (Gy):

$$1 \text{ Gy} = 1 \text{ joule/kg}$$

Older units of absorbed energy are the rad (r) and Roentgen (R):

$$\begin{aligned} 1 \text{ r} &= 100 \text{ ergs/g} = 10^{-2} \text{ Gy} \\ 1 \text{ R} &= 0.877 \text{ r in air} = 0.93 - 0.98 \text{ r in water and tissue} \end{aligned}$$

The SI unit for radiation dosage is the Sievert (Sv), which takes into account the empirically determined relative biological effectiveness (RBE) of a given form of radiation:

$$\begin{aligned} \text{dosage [Sv]} &= \text{RBE} \times \text{dosage [Gy]} \\ \text{RBE} &= \frac{(\text{biological effect of a dose of standard radiation [Gy]})}{(\text{biological effect of a dose of other radiation [Gy]})} \end{aligned}$$

RBE = 1 for commonly encountered radionuclides

The older unit for dosage is the rem (Roentgen-equivalent-man):

$$1 \text{ rem} = 0.01 \text{ Sv}$$

LITERATURE CITED

- Dawson, R.M.C., Elliot, D.C., Elliott, W.H., and Jones, K.M. (eds.) 1986. Data for Biochemical Research. Alden Press, London.
- Lederer, C.M., Hollander, J.M., and Perlman, I. (eds.) 1967. Table of Radioisotopes, 6th edition. John Wiley and Sons, New York.
- Shleien, B. (ed.) 1987. Radiation Safety Manual for Users of Radioisotopes in Research and Academic Institutions. Nucleon Lectern Associates, Olney, Md.

Table A.1E.2 Physical Characteristics of Commonly Used Radionuclides^a

Nuclide	Half-life	Emission	Energy, max (MeV)	Range of emission, max	Approx. specific activity at 100% enrichment (Ci/mg)	Atom resulting from decay	Target organ
³ H	12.43 years	β	0.0186	0.42 cm (air)	9.6	³ He	Whole body
¹⁴ C	5370 years	β	0.156	21.8 cm (air)	4.4 mCi/mg	¹⁴ N	Bone, fat
³² P ^b	14.3 days	β	1.71	610 cm (air) 0.8 cm (water) 0.76 cm (Plexiglas)	285	³³ S ¹⁶ S	Bone
³³ P ^b	25.4 days	β	0.249	49 cm	156	³³ S ¹⁶ S	Bone
³⁵ S	87.4 days	β	0.167	24.4 cm (air)	43	³⁵ Cl ¹⁷ Cl	Testes
¹²⁵ I ^c	60 days	γ	0.27–0.035	0.2 mm (lead)	14.2	¹²⁵ Te ⁵² Te	Thyroid
¹³¹ I ^c	8.04 days	β γ	0.606 0.364	165 cm (air) 2.4 cm (lead)	123	¹³⁰ Xe ⁵⁴ Xe	Thyroid

^aTable compiled based on information in Lederer et al. (1967) and Shleien (1987).

^bRecommended shielding is Plexiglas; half-value layer measurement is 1 cm.

^cRecommended shielding is lead; half-value layer measurement is 0.02 mm.

Table A.1E.3 Shielding Radioactive Emission^a

β emitters

Energy (MeV)	Mass (mg)/cm ² to reduce intensity by 50%	Thickness (mm) to reduce intensity by 50%			
		Water	Glass	Lead	Plexiglas
0.1	1.3	0.013	0.005	0.0011	0.0125
1.0	48	0.48	0.192	0.042	0.38
2.0	130	1.3	0.52	0.115	1.1
5.0	400	4.0	1.6	0.35	4.2

γ emitters

Energy (MeV)	Thickness of material (cm) to attenuate a broad beam of γ-rays by a factor of 10			
	Water	Aluminum	Iron	Lead
0.5	54.6	20.3	6.1	1.8
1.0	70.0	24.4	8.2	3.8
2.0	76.0	32.0	11.0	5.9
3.0	89.0	37.0	12.0	6.4

^aFrom Dawson et al. (1986). Reprinted with permission.

Table A.1E.4 Decay factors for calculating the amount of radioactivity present at a given time after a reference date. For example, a vial containing 1.85 MBq (50 μ Ci) of an ^{35}S -labeled compound on the reference date will have the following activity 33 days later: $1.85 \times 0.770 = 1.42$ MBq; $50 \times 0.770 = 38.5$ μ Ci.

^{125}I Half-life: 60.0 days

Days	0	2	4	6	8	10	12	14	16	18
0	1.000	0.977	0.955	0.933	0.912	0.891	0.871	0.851	0.831	0.812
20	0.794	0.776	0.758	0.741	0.724	0.707	0.691	0.675	0.660	0.645
40	0.630	0.616	0.602	0.588	0.574	0.561	0.548	0.536	0.524	0.512
60	0.500	0.489	0.477	0.467	0.456	0.445	0.435	0.425	0.416	0.406
80	0.397	0.388	0.379	0.370	0.362	0.354	0.345	0.338	0.330	0.322
100	0.315	0.308	0.301	0.294	0.287	0.281	0.274	0.268	0.262	0.256
120	0.250	0.244	0.239	0.233	0.228	0.223	0.218	0.213	0.208	0.203
140	0.198	0.194	0.189	0.185	0.181	0.177	0.173	0.169	0.165	0.161
160	0.157	0.154	0.150	0.147	0.144	0.140	0.137	0.134	0.131	0.128
180	0.125	0.122	0.119	0.117	0.114	0.111	0.109	0.106	0.104	0.102
200	0.099	0.097	0.095	0.093	0.090	0.088	0.086	0.084	0.082	0.081
220	0.079	0.077	0.075	0.073	0.072	0.070	0.069	0.067	0.065	0.064
240	0.063	0.061	0.060	0.058	0.057	0.056	0.054	0.053	0.052	0.051

^{32}P Half-life: 14.3 days

Hours	0	12	24	36	48	60	72	84
0	1.000	0.976	0.953	0.930	0.908	0.886	0.865	0.844
4	0.824	0.804	0.785	0.766	0.748	0.730	0.712	0.695
8	0.679	0.662	0.646	0.631	0.616	0.601	0.587	0.573
12	0.559	0.546	0.533	0.520	0.507	0.495	0.483	0.472
16	0.460	0.449	0.439	0.428	0.418	0.408	0.398	0.389
20	0.379	0.370	0.361	0.353	0.344	0.336	0.328	0.320
24	0.312	0.305	0.298	0.291	0.284	0.277	0.270	0.264
28	0.257	0.251	0.245	0.239	0.234	0.228	0.223	0.217
32	0.212	0.207	0.202	0.197	0.192	0.188	0.183	0.179
36	0.175	0.170	0.166	0.162	0.159	0.155	0.151	0.147
40	0.144	0.140	0.137	0.134	0.131	0.127	0.124	0.121
44	0.119	0.116	0.113	0.110	0.108	0.105	0.102	0.100
48	0.098	0.095	0.093	0.091	0.089	0.086	0.084	0.082
52	0.080	0.078	0.077	0.075	0.073	0.071	0.070	0.068

^{131}I Half-life: 8.04 days

Hours	0	6	12	18	24	30	36	42	48	54	60	66
0	1.000	0.979	0.958	0.937	0.917	0.898	0.879	0.860	0.842	0.824	0.806	0.789
3	0.772	0.756	0.740	0.724	0.708	0.693	0.678	0.664	0.650	0.636	0.622	0.609
6	0.596	0.583	0.571	0.559	0.547	0.533	0.524	0.513	0.502	0.491	0.481	0.470
9	0.460	0.450	0.441	0.431	0.422	0.413	0.405	0.396	0.387	0.379	0.371	0.363
12	0.355	0.348	0.340	0.333	0.326	0.319	0.312	0.306	0.299	0.293	0.286	0.280
15	0.274	0.269	0.263	0.257	0.252	0.246	0.241	0.236	0.231	0.226	0.221	0.216
18	0.212	0.207	0.203	0.199	0.194	0.190	0.186	0.182	0.178	0.175	0.171	0.167
21	0.164	0.160	0.157	0.153	0.150	0.147	0.144	0.141	0.138	0.135	0.132	0.129
24	0.126	0.124	0.121	0.118	0.116	0.113	0.111	0.109	0.106	0.104	0.102	0.100
27	0.098	0.095	0.093	0.091	0.089	0.088	0.086	0.084	0.082	0.080	0.079	0.077
30	0.075	0.074	0.072	0.071	0.069	0.068	0.066	0.065	0.064	0.063	0.061	0.059
33	0.058	0.057	0.056	0.054	0.053	0.052	0.051	0.050	0.049	0.048	0.047	0.046
36	0.045	0.044	0.043	0.042	0.041	0.040	0.039	0.039	0.038	0.037	0.036	0.035

^{35}S Half-life: 87.4 days

Days	0	1	2	3	4	5	6	7	8	9
0	1.000	0.992	0.984	0.976	0.969	0.961	0.954	0.946	0.939	0.931
1	0.946	0.939	0.931	0.924	0.916	0.909	0.902	0.895	0.888	0.881
2	0.895	0.888	0.881	0.874	0.867	0.860	0.853	0.847	0.840	0.833
3	0.847	0.840	0.833	0.827	0.820	0.814	0.807	0.801	0.795	0.788
4	0.801	0.795	0.788	0.782	0.776	0.770	0.764	0.758	0.752	0.746
5	0.758	0.752	0.746	0.740	0.734	0.728	0.722	0.717	0.711	0.705
6	0.717	0.711	0.705	0.700	0.694	0.689	0.683	0.678	0.673	0.667
7	0.678	0.673	0.667	0.662	0.657	0.652	0.646	0.641	0.636	0.631
8	0.641	0.636	0.631	0.626	0.621	0.616	0.612	0.607	0.602	0.597
9	0.607	0.602	0.597	0.592	0.588	0.583	0.579	0.574	0.569	0.565
10	0.574	0.569	0.565	0.560	0.556	0.552	0.547	0.543	0.539	0.534
11	0.543	0.539	0.534	0.530	0.526	0.522	0.518	0.514	0.510	0.506
12	0.514	0.510	0.506	0.502	0.498	0.494	0.490	0.486	0.482	0.478

^{33}P Half-life: 25.4 days

Days	0	1	2	3	4	5	6	7	8	9
0	1.000	0.973	0.947	0.921	0.897	0.872	0.849	0.826	0.804	0.782
10	0.761	0.741	0.721	0.701	0.683	0.664	0.646	0.629	0.612	0.595
20	0.579	0.564	0.549	0.534	0.520	0.506	0.492	0.479	0.466	0.453
30	0.441	0.429	0.418	0.406	0.395	0.385	0.374	0.364	0.355	0.345
40	0.336	0.327	0.318	0.309	0.301	0.293	0.285	0.277	0.270	0.263
50	0.256	0.249	0.242	0.236	0.229	0.223	0.217	0.211	0.205	0.200
60	0.195	0.189	0.184	0.179	0.174	0.170	0.165	0.161	0.156	0.152
70	0.148	0.144	0.140	0.136	0.133	0.129	0.126	0.122	0.119	0.116
80	0.113	0.110	0.107	0.104	0.101	0.098	0.096	0.093	0.091	0.088
90	0.086	0.084	0.081	0.079	0.077	0.075	0.073	0.071	0.069	0.067
100	0.065	0.064	0.062	0.060	0.059	0.057	0.055	0.054	0.053	0.051
110	0.050	0.048	0.047	0.046	0.045	0.043	0.042	0.041	0.040	0.039
120	0.038	0.037	0.036	0.035	0.034	0.033	0.032	0.031	0.030	0.030

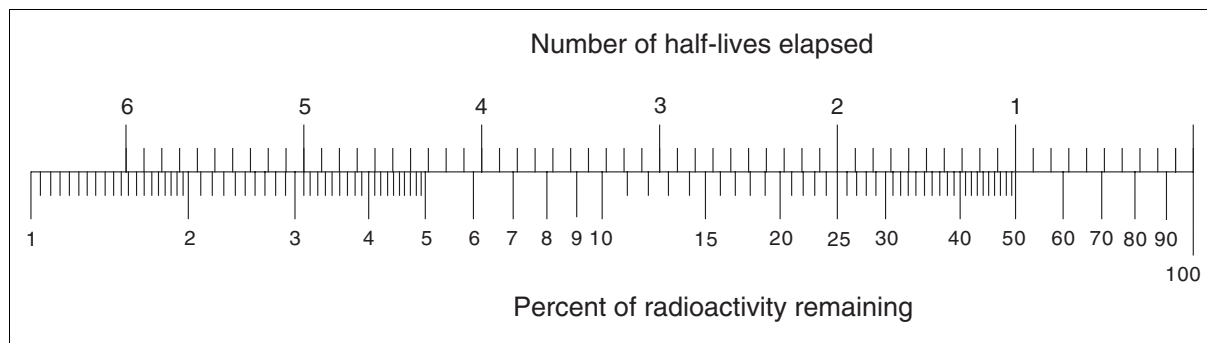


Figure A.1E.1 Correlation of loss of radioactivity with elapsing half-lives of an isotope.

Safe Use of Radioisotopes

The use of radioisotopes to label specific molecules in a defined way has greatly furthered the discovery and dissection of biochemical pathways. The development of methods to inexpensively synthesize such tagged biological compounds on an industrial scale has enabled them to be used routinely in laboratory protocols, including many detailed in this manual. Although most of these protocols involve the use of only microcurie (μCi) amounts of radioactivity, some (particularly those describing the metabolic labeling of proteins or nucleic acids within cells; see *UNITS 10.18 & 3.4*) require amounts on the order of tens of millicuries (mCi). In all cases where radioisotopes are used, depending on the quantity and nature of the isotope certain precautions must be taken to ensure the safety of the scientist. This appendix outlines a few such considerations relevant to the isotopes most frequently used in biological research.

In designing safe protocols for the use of radioactivity, the importance of common sense, based on an understanding of the general principles of isotopic decay and the importance of continuous monitoring with a hand-held radioactivity monitor (e.g., Geiger counter), cannot be overemphasized. In addition, it is also critical to take into account the rules, regulations, and limitations imposed by each specific institution. These are usually not optional considerations: an institution's license to use radioactivity normally depends on strict adherence to such rules.

Many of the protocols described have evolved (and are evolving still) over the years and through the millicuries in the Department of Molecular Biology and Virology at the Salk Institute. The authors are indebted to those who have trained them in the safe use of radioactivity, in particular to the members of the Salk Institute Radiation Safety Department. Most of the designs for the shields and other safety equipment shown in Figures A.1F.1A, A.1F.2, and A.1F.3 were created at the Salk Institute in collaboration with Dave Clarkin and Mario Tengco. Safety equipment of similar design is available from several commercial vendors, including CBS Scientific and Research Products International.

BACKGROUND INFORMATION

The Decay Process

As anyone who has taken a basic chemistry course will remember, each element is charac-

terized by its atomic number, defined as the number of orbital electrons or the number of protons in the nucleus of that atom. Isotopes of a given element exist because some atoms of each element, while by definition having the same number of protons, have a different number of neutrons and therefore a different nuclear weight. It should be noted that the number of electrons outside the nucleus remains the same for all isotopes of a given element, and so all isotopes of a given element are equivalent with respect to their chemical reactivity.

Radioactive decay occurs when subatomic particles are released from the nucleus of an atom of a heavy isotope. This often results in the conversion of an atom of one isotope to an isotope of a different element, because the original isotope's atomic number changes after decay. The subatomic particles released from naturally occurring radioisotopes are of three basic types: α and β particles and γ rays.

An α particle is essentially the nucleus of a helium atom, or two protons plus two neutrons. It is a relatively large, heavy particle that moves slowly and usually only across short distances before it encounters some other atom with which it interacts. These particles are released from isotopes with large nuclei (atomic number >82 ; e.g., plutonium or uranium); such isotopes are not commonly used in biological research.

In contrast to α particles, β particles are light, high-speed charged particles. Negatively charged β particles are essentially electrons of nuclear origin that are released when a neutron is converted to a proton. Release of a β particle thus changes the atomic number and elemental status of the isotope.

γ radiation has both particle and wave properties; its wavelength falls within the range of X-ray wavelengths. The distinction between γ rays and X rays was made when primitive X-ray machines produced X rays with a wavelength longer than those of the γ rays produced naturally by radioisotopes. Modern X-ray machines produce a much broader spectrum of wavelengths, including γ radiation; currently this sort of X-ray radiation is termed γ when it is of nuclear origin. Unlike β -particle release, the release of γ radiation by itself produces an isotopic change rather than an elemental one; however, the resultant nuclei are unstable and often decay further, releasing β particles.

The energy of all α particles and γ rays (measured in electron volts) is fixed, because they are of specific composition or wavelength.

**Standard
Measurements,
Data, and
Abbreviations**

A.1F.1

The energy of β particles, however, varies depending on the atom they originate from (and with concomitant release of neutrinos or anti-neutrinos that serve to balance the conservation of energy aspect of the decay equation). Thus there are (relatively) high-energy β particles released during the decay of ^{32}P and low-energy β particles released when tritium (^3H) decays.

Isotopic decay usually involves a chain or sequence of events rather than a single loss of a particle, because the resultant, equally unstable atoms try to achieve equilibrium. During this course of decay, secondary forms of radiation can be generated that may also pose a hazard to workers. For example, when high-energy β particles released during the decay of ^{32}P encounter the nuclei of atoms with a large atomic number, a strong interaction occurs. The β particle loses some energy in the form of a photon. Such photons are called Bremsstrahlung radiation; they are detectable using a monitor suitable for the detection of γ or X rays.

Following their release, α , β , and γ emissions (as well as secondary forms of radiation) travel varying average distances at varying average speeds, depending on their energy and the density of the material through which they are moving. The distance they actually travel before encountering either the electrons or nucleus of another atom is termed their degree of penetrance. This value is expressed as an average for each type of particle. The energy of the particles released (and therefore their potential penetrance) thus dictates what type of shielding, if any, is necessary for protection against the radiation generated by the decay of a given isotope. α , β , and γ emissions all have the potential, upon encountering an atom, to knock out its electrons, thereby creating ions. Thus, these three types of emissions are called ionizing radiation. The formation of such ions may result in the perturbation of biological processes: therein lies the danger associated with radioactivity!

Table A.1E.2 provides a summary of the factors discussed above for several isotopes commonly used in biochemical research.

Measuring Radioactivity and Individual Exposure to It

The radioactivity of a given substance is measured in terms of its ionizing activity. A Curie by definition is the amount of radioactive material that will produce 3.4×10^{10} disintegration (ions) per second. This, not coincidentally, happens to be the number of disintegrations that occur during the decay of one gram of radium

and its decay products. Exposure to such radiation is measured as the amount of energy absorbed by the recipient, which, of course, is directly related to the potential damage such radiation may cause. One rad is the dose of radiation that will cause 100 ergs of energy to be absorbed per gram of irradiated material. Another unit commonly used to measure radiation doses is the rem; this is related to the rad but takes into account a "quality factor" based on the type of ionizing radiation being received. For β particles and γ or X rays this factor is 1; therefore, rems β equal rads β . In contrast, the quality factor associated with α particles is 20, so an exposure of one rad due to α particles would be recorded as 20 rem.

The amount or dose of radiation received by materials (cells, scientists, etc.) near the source depends not only on the specific type and energy (penetrance) of the radiation being produced, but also on the subject's distance from the source, the existence of any intervening layers of attenuating material (shielding), and the length of time spent in the vicinity of the radiation source. To best measure such doses, every person working with or around radioactivity should wear an appropriate type of radiation detection badge (in addition to carrying a portable radiation monitor that can give an immediate, approximate reading.) This is normally a requirement (not an option) for compliance with an institution's license to use radioisotopes. Such badges are usually furnished by the safety department and collected at regular intervals for reading by a contracted company. At most institutions, the old-style film badges have been replaced with the more accurate TLDs (thermoluminescent dosimeters). These take advantage of chemicals such as calcium or lithium fluorides which, following exposure to ionizing radiation, will luminesce at temperatures below their normal thermal luminescence threshold. Different types of badges are sensitive to different types of radiation: always be sure to wear one that is appropriate for detecting exposure to the isotope being used! In most places pregnant women are asked to wear a more sensitive (and more expensive) dosimeter to better monitor their (and the developing fetus') exposure. Most often workers will be asked to wear a radiation detection badge on the labcoat lapel in order to measure whole-body radiation. When working with ^{32}P or ^{125}I it is also advisable to wear a ring badge to measure exposure to the unshielded extremities (fingers). The limit set for "acceptable" exposure to whole-body radiation is sev-

eral-fold less than the limit set for extremities. Nevertheless, we have found that the exposure recorded on ring badges is often significant with respect to the limit for extremities set by our institution.

What is known about the effects on humans of exposure to low levels of radiation (i.e., levels which would be received when briefly handling mCi or μ Ci amounts of radioactivity)? Not much, for the obvious reason that direct studies have not been undertaken. Accordingly, guidelines for exposure levels are set using extrapolations—either by extrapolating down from population statistics obtained following accidents or disasters (the Chernobyl meltdown, atomic bombings) or by extrapolating up from numbers obtained from animal experiments. Each form of extrapolation is subject to caveats, and given that predictions based on such extrapolations cannot be perfect, most health and safety personnel aim for radiation exposure levels said to be ALARA or “as low as reasonably achievable.” Further discussion of exposure limits and the statistics on which they are based may be found in B. Shleien’s health physics text (Shleien, 1987). Limiting exposure to radiation can be accomplished by adjusting several parameters of the exposure: the duration of exposure, distance from the source, and the density of the material (air, water, shielding) between the individual and the source.

Time is of the essence

When designing any experiment using radioactivity, every effort should be made to limit the time spent directly handling the vials or tubes containing the radioactive material. Speed should be encouraged in all manipulations, though not to the point of recklessness! Have everything needed for the experiment ready at hand before the radioactivity is introduced into the work area.

Distance helps to determine dose

When possible, experiments involving radioactivity should be performed in an area separate from the rest of the lab. Many institutions require that such work be performed in a designated “hot lab”; however, if many people in the laboratory routinely use radioisotopes, it is less than feasible to move them all into what is usually a smaller space. No matter where an individual is working, it is his or her responsibility to monitor the work area and ensure his or her own safety and the safety of those working nearby by using adequate shielding. Obvi-

ously, when handling the radioactive samples, it is necessary to work rapidly behind any required shielding. To protect bystanders, remember that the intensity of radiation from a source (moving through air) falls off in proportion to the square of the distance. Thus, if standing 1 foot away from a source for 5 min would result in an exposure of 45 units, standing 3 feet away for the same amount of time would result in an exposure $(1/3)^2$ of 45 units, or 5 units. This factor is also relevant when considering the storage of large amounts of radioactivity, particularly ^{125}I or ^{32}P , as no amount of shielding can completely eliminate radiation.

Shielding is the key to safety

As mentioned above, the energy of the particle(s) released during the decay of an isotope determines what, if any, type of shielding is appropriate. β particles released during the decay of ^{14}C and ^{35}S possess roughly ten times the energy of those released when ^3H decays. All three β particles are of relatively low energy, do not travel very far in air, and cannot penetrate solid surfaces. No barriers are necessary for shielding against this type of β radiation. The major health threat from these isotopes occurs through their accidental ingestion, inhalation, or injection.

β particles released during the decay of ^{32}P have 10-fold higher energy than those released from ^{14}C and pose a significant threat to workers. (One reported hazard is the potential for induction of cataracts in the unshielded eye.) The fact that these high-energy beta particles can potentially generate significant amounts of Bremsstrahlung radiation is the reason that low-density materials are used as the primary layer of shielding for ^{32}P β radiation. Water, glass, and plastic are suitable low-density materials (as opposed to lead). Obviously water is unsuitable as a shielding layer for work on the bench, although it does a reasonable job when samples are incubating in a water bath. Shields made from a thickness of glass sufficient to stop these particles would be extremely heavy and cumbersome (as well as dangerous if dropped). Fortunately, plastic or acrylic materials—variously called Plexiglas, Perspex, or Lucite—are available for shielding against ^{32}P β radiation. Shields as well as storage boxes constructed of various thicknesses of Plexiglas are necessary equipment in laboratories where ^{32}P is used. When mCi amounts of ^{32}P are used at one time it is necessary to also block the Bremsstrahlung radiation by adding a layer of high-density

material (such as 4 to 6 mm lead) to the outside of the Plexiglas shield (covering the side farthest from the radioactive source).

γ rays released during the decay of ^{125}I have much higher penetrance than the β particles from ^{32}P decay; this radiation must be stopped by very-high-density material, such as lead. Lead foil of varying thicknesses (2 to 6 mm) can be purchased in rolls and can be cut and molded to cover any container, or taped to a Plexiglas shield (used in this instance for support). Obviously this latter arrangement has the disadvantage that it is impossible to see what one is doing through the shield. For routine shielding of manipulations involving ^{125}I , it is useful to purchase a lead-impregnated Plexiglas shield that is transparent, albeit inevitably very heavy (as well as relatively expensive).

Although it seems logical that the use of more radioactivity necessitates the use of thicker layers of shielding, it is also true that no shielding material is capable of completely stopping all radiation. When deciding how thick is "thick enough," consult the half-value layer measurement for each type of shielding material. This number gives the thickness of a given material necessary to stop half the radiation from a source; Table A.1E.3 lists half-value layer measurements for the β particles released from isotopes commonly used in the biology laboratory. In general, 1 to 2 cm of Plexiglas and/or 0.02 mm of lead are sufficient to shield the amounts of radioactivity used in our experiments.

GENERAL PRECAUTIONS

Before going on to a discussion of specific precautions to be taken with individual isotopes, a short list of general precautions to be taken with all isotopes seems pertinent:

1. Know the rules. Be sure that each individual is authorized to use each particular isotope and uses it in an authorized work area.

2. Don the appropriate apparel. Whenever working at the lab bench, it is good safety practice to wear a labcoat for protection. Disposable paper/synthetic coats of various styles are commercially available: at \$4 each these may be conveniently thrown out if contaminated with radioactivity during an experiment, rather than held for decay as might be preferable with cloth coats costing about \$30 each. As an alternative, disposable sleeves can be purchased and worn over those of the usual cloth coat. Other necessary accessories include radiation safety badges, gloves, and protective eyewear. It's handy to wear two pairs of gloves

at once when using radioactivity; when the outer pair becomes contaminated, it is possible to strip it off and continue working without interruption.

3. Protect the work area as well as the workers. Lab benches and the bases of any shields used should be covered with some sort of disposable, preferably absorbent, paper sheet. Underpads or diapers (the kind normally used in hospitals) are convenient for this purpose.

4. Use appropriate designated equipment. It is very convenient, where use justifies the expense, to have a few adjustable pipettors dedicated for use with each particular isotope. Likewise, it is good practice to use only certain labeled centrifuges and microcentrifuge rotors for radioactive samples so that all the lab's rotors do not become contaminated. Although such equipment should be cleaned after each use, complete decontamination is often not possible. A few pipettors or a single microcentrifuge can easily be stored (and used) behind appropriate shielding. Actually, contamination of the insides and tip ends of pipettors can be greatly reduced by using tips supplied with internal aerosol barriers; these have recently become very popular for setting up PCR reactions and are available to fit a variety of pipettors. To prevent contamination of the outside of the pipettor's barrel, simply wrap the hand-grip in Parafilm, which can be discarded later. Several manufacturers sell disposable paper inserts for their microcentrifuges that protect the wall of the rotor chamber from contamination that might spin off the outside of sample tubes. Trying to fashion homemade liners of this sort is not recommended, as we have had disastrous experiences using laminated adhesive paper that "unstuck" during microcentrifuge spins. These liners would get caught by the rotor, shattering sample tubes and creating an even bigger mess!

5. Know where to dispose of radioactive waste, both liquid and solid. Most institutions require that radioactive waste be segregated by isotope. This is done not only so that appropriate shielding can be placed around waste containers, but so that some waste can be allowed to decay prior to disposal through normal trash channels. With a decreasing number of radioactive waste disposal facilities able or willing to accept radioactive waste for burial (and a concomitant increase in dumping charges from those that still do) this practice of on-site decay can save an institution thousands of dollars a year in disposal charges.

6. Label your label! It is only common courtesy (as well as common sense) to alert coworkers to the existence of anything and everything radioactive that is left where they may come in contact with it! A simple piece of tape affixed to the sample box—with the investigator's name, the amount and type of isotope, and the date written on it—should do the trick. Yellow hazard tape printed with the international symbol for radioactivity is commercially available in a variety of widths.

7. Monitor radioactivity early and often. Portable radiation detection monitors are essential equipment for every laboratory using radioactivity. No matter how much or how little radioactivity is being used, the investigator should keep a hand-held monitor nearby—and it should be on! Turn it on before touching any radioactivity to avoid contaminating the monitor's switch. Use a monitor with the appropriate detection capacity (β for ^{35}S and ^{32}P ; γ for ^{125}I) before, during, and after all procedures. The more frequently fingers and relevant equipment are monitored, the more quickly a spill or glove contamination will be detected. Such timely detection will keep both the potential mess and the cleanup time to a minimum. Because low-energy β emitters such as ^3H cannot be detected using such monitors, wipe tests of the bench and equipment used are necessary to ensure that contamination of the work area did not occur.

SPECIFIC PRECAUTIONS

The following sections describe precautions to be taken with individual isotopes in specific forms. Although the sections dealing with ^{35}S - or ^{32}P -labeling of proteins in intact cells are presented in terms of mammalian cells, most of the instructions are also pertinent (with minimal and obvious modifications) to the labeling of proteins in other cells (bacterial, insect, etc.).

Working with ^{35}S

Using ^{35}S to label cellular proteins and proteins translated in vitro

As discussed above, the β radiation generated during ^{35}S decay is not strong enough to make barrier forms of shielding necessary. The risk associated with ^{35}S comes primarily through its ingestion and subsequent concentration in various target organs, particularly the testes. Although willful ingestion of ^{35}S seems unlikely, accidental or unknowing ingestion may be more common. As reported several years ago (Meisenhelder and Hunter, 1988), ^{35}S -labeled methionine and cysteine, which are

routinely used to label proteins in intact cells and by in vitro translation, break down chemically to generate a volatile radioactive component. The breakdown occurs independent of cellular metabolism. Thus the radioactive component is generated to the same extent in stock vials as in cell culture dishes. The process seems to be promoted by freezing and thawing ^{35}S -labeled materials. The exact identity of this component is not known, although it is probably SO_2 or CH_3SH . What is known is that it dissolves readily in water and is absorbed by activated charcoal or copper.

The amount of this volatile radioactive component released, despite stabilizers added by the manufacturers, is about 1/8000 of the total radioactivity present. The amount of this radioactivity that a scientist is likely to inhale while using these compounds is presumably even smaller. Nevertheless, such a component can potentially contaminate a wide area because of its volatility, and also tends to concentrate in target organs. Thus, it is advisable to thaw vials of ^{35}S -labeled amino acids in a controlled area such as a hood equipped with a charcoal filter. This charcoal filter will become quite contaminated and should be changed every few months. If such an area is not available, the stock vial should be thawed using a needle attached to a charcoal-packed syringe to vent and trap the volatile compound.

Anyone who has ever added ^{35}S -labeled amino acids to dishes of cells for even short periods knows that the incubator(s) used for such labelings quickly become highly contaminated with ^{35}S . Such contamination is not limited to the dish itself, nor to the shelf on which the dish was placed. Rather, the radioactive component's solubility in water allows it to circulate throughout the moist atmosphere of the incubator and contaminate all the inside surfaces of the incubator. For this reason, in laboratories where such metabolic labelings are routine, it is highly convenient to designate one incubator to be used solely for working with ^{35}S -labeled samples. Such an incubator can be fitted with a large honeycomb-style filter, the size of an incubator shelf, made of pressed, activated charcoal. These filters are available from local air-quality-control companies. Such a filter will quickly become quite contaminated with radioactivity and should therefore be monitored and changed as necessary (usually about every three months if the incubator is used several times a week). The water used to humidify the incubator will also become quite "hot" (contaminated with radioactivity); keep-

ing the water in a shallow glass pan on the bottom of the incubator makes it easy to change it after every use, thus preventing contamination from accumulating. Even with the charcoal filter and water as absorbents, the shelves, fan, and inner glass door of the incubator will become contaminated, as will the tray on which the cells are carried and incubated. Routine wipe tests and cleaning when necessary will help to minimize potential spread of this contamination.

If such labelings are done infrequently or there is no "spare" incubator, dishes of cells can be placed in a box during incubation. This box should be made of plastic, which is generally more easily decontaminated than metal. Along with the dishes of cells, a small sachet made of activated charcoal wrapped loosely in tissue (Kimwipes work well) should be placed in the box. If the box is sealed, it will obviously need to be gassed with the correct mixture of CO₂; otherwise small holes can be incorporated into the box design to allow equilibration with the incubator's atmosphere. In either case, the incubator used for the labeling should be carefully monitored for radioactivity after each experiment.

Working with ³²P

μCi amounts of ³²P

The amount of ³²P-labeled nucleotide used to label nucleic acid probes for northern or Southern blotting is typically under 250 μCi, and the amount of [γ -³²P]ATP used for in vitro phosphorylation of proteins does not usually exceed 50 μCi for a single kinase reaction (or several hundred μCi per experiment). However, handling even these small amounts, given the time spent on such experiments, can result in an unacceptable level of exposure if proper shielding is not employed. With no intervening shielding, the dose rate 1 cm away from 1 mCi ³²P is 200,000 mrad/hr; the local dose rate to basal cells resulting from a skin contamination of 1 μCi/cm² is 9200 mrad/hr (Shleien, 1987). Such a skin contamination could be easily attained though careless pipetting and the resultant creation of an aerosol of radioactive microdroplets, because the concentration of a typical stock solution of labeled nucleotide may be 10 μCi/μl.

For proper protection during this sort of experiment, besides the usual personal attire (glasses, coat, safety badges, and gloves) it is necessary to use some form of Plexiglas screen between the body and the samples (see Fig.

A.1F.1A). Check the level of radiation coming through the outside of the shield with a portable monitor to be sure the thickness of the Plexiglas is adequate. Hands can be shielded from some exposure by placing the sample tubes in a solid Plexiglas rack, which is also useful for transporting samples from the bench to a centrifuge or water bath (see Fig. A.1F.1B).

Experiments of these types often include an incubation step performed at a specific temperature, usually in a water bath. Although the water surrounding the tubes or hybridization bags will effectively stop β radiation, shielding should be added over the top of the tubes (where there is no water)—e.g., a simple flat piece of Plexiglas. If the frequency of usage justifies the expense, an entire lid for the water bath can be constructed from Plexiglas. When hybridization reactions are performed in bags, care should be taken to monitor (and shield) the apparatus used to heat-seal the bags.

The waste generated during the experiments should also be shielded. It is convenient to have a temporary waste container right on the bench. Discard pipette tips and other solid waste into a beaker lined with a plastic bag and placed behind the shield. This bag can then be emptied into the appropriate shielded laboratory waste container when the experiment is done. Liquid waste can be pipetted into a disposable tube set in a stable rack behind the shield (see Fig. A.1F.1C).

When radiolabeled probes or proteins must be gel-purified, it may be necessary to shield the gel apparatus during electrophoresis if the samples are particularly hot. Be advised that the electrophoresis buffer is likely to become very radioactive if the unincorporated label is allowed to run off the bottom of the gel; check with a radiation safety officer for instructions on how to dispose of such buffer. It is also prudent to check the gel plates with a portable detection monitor after the electrophoresis is completed, because they sometimes become contaminated as well.

mCi amounts of ³²P

In order to study protein phosphorylation in intact mammalian cells, cells in tissue culture dishes are incubated in phosphate-free medium with ³²P-labeled orthophosphate for a period of several hours or overnight to label the proteins. The amount of ³²P used in such labelings can be substantial. Cells are normally incubated in 1 to 2 mCi of ³²P/ml labeling medium; for each 6-cm dish of cells, 2.5 to 5 mCi ³²P may be

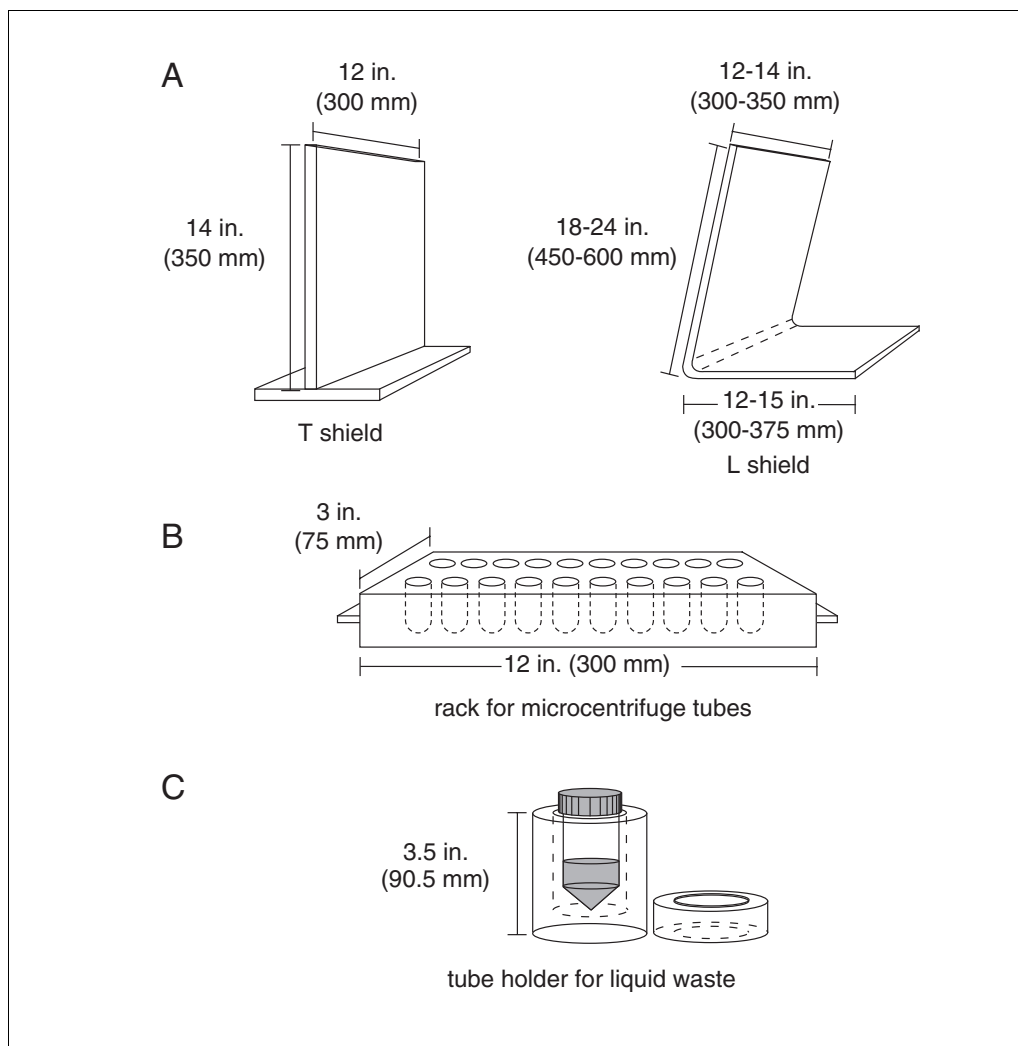


Figure A.1F.1 Plexiglas shielding for ^{32}P . (A) Two portable shields (L and T design) made of 0.5-in. Plexiglas. Either can be used to directly shield the scientist from the radioactivity he or she is using. Turned on its side, the L-shaped shield can be used to construct two sides of a cage around a temporary work area, providing shielding for other workers directly across or to the sides of the person working with ^{32}P . (B) Tube rack for samples in microcentrifuge tubes. (C) Tube holder for liquid waste collection.

used. When this figure is multiplied by the number of dishes necessary per sample, and the number of different samples in each experiment, it is clear that the amount of ^{32}P used in one experiment can easily reach 25 mCi or more. Because so much radioactivity is used in the initial labeling phase of such experiments, it is necessary for a researcher to take extra precautions in order to adequately shield him or herself and coworkers.

When adding label to dishes of cells, it is important to work as rapidly as possible. An important contribution to the speed of these manipulations is to have everything that will be needed at hand before even introducing the label into the work area. Prepare the work area, arranging shielding and covering the bench

with diapers, in advance. Set out all necessary items, including any pipettors and tips needed, a portable detection monitor, extra gloves, and a cell house (see Fig. A.1F.2A).

Work involving this much radioactivity should be done behind a Plexiglas shield at least $\frac{3}{4}$ in. (2 cm) thick; the addition of a layer of lead to the outside lower section of this shield to stop Bremsstrahlung radiation is also needed. If one shield can be dedicated to this purpose at a specific location, a sheet of lead several centimeters thick can be permanently screwed to the Plexiglas (as shown in Fig. A.1F.2B); however, this lead makes the shield extremely heavy and therefore less than portable. If space constraints do not permit the existence of such a permanent labeling station,

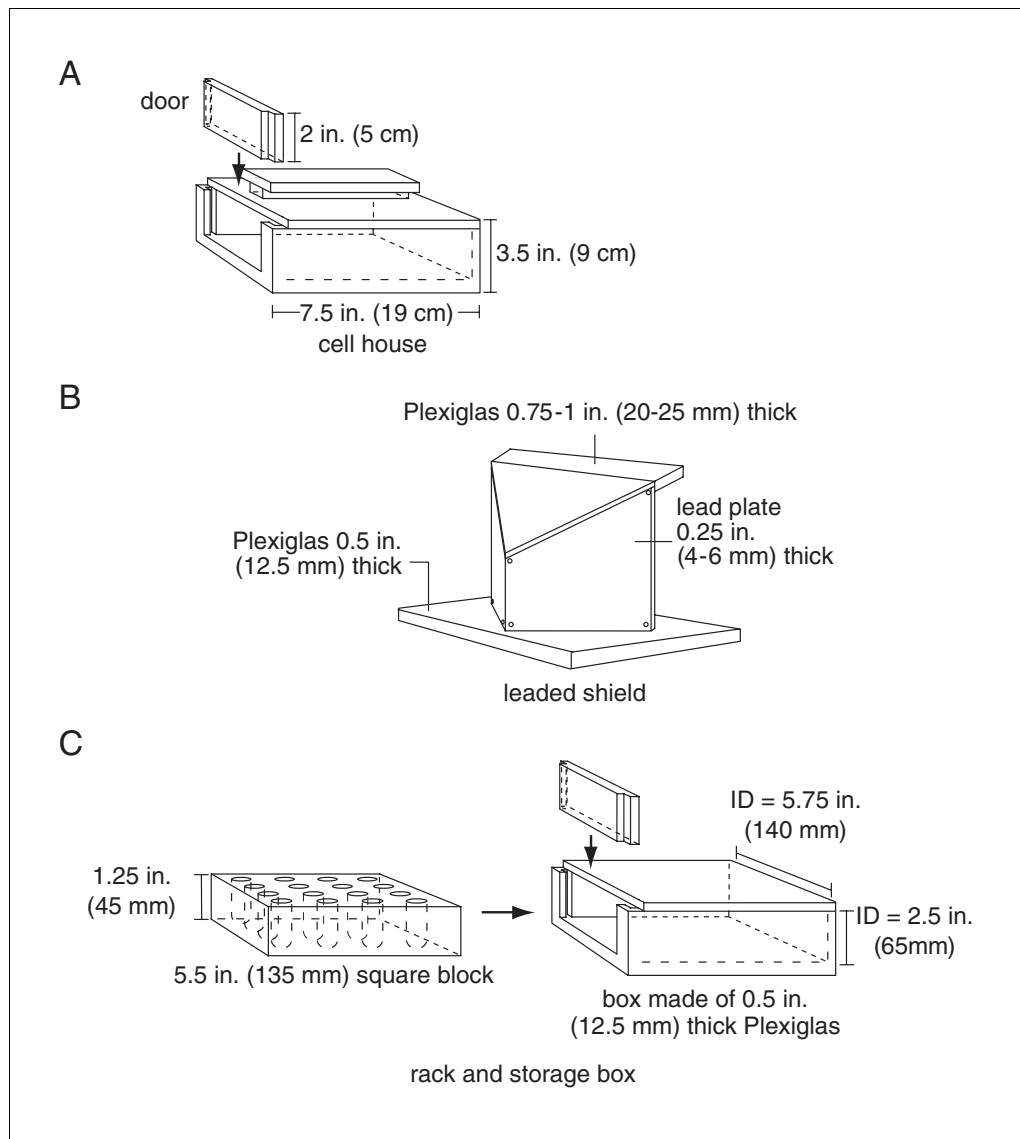


Figure A.1F.2 (A) Box for cell incubation (a “cell house”). (B) Stationary leaded shield. (C) Sample storage rack and box made of 0.5-in. Plexiglas. Abbreviations: ID, interior dimension.

a layer or two of lead foil can be taped temporarily to the outside of the Plexiglas shield.

Again, each worker should take care to shield not only him or herself, but bystanders on all sides. Handling of label should be done away from the central laboratory if possible to take maximum advantage of distance as an additional form of shielding. It is also advisable not to perform such experiments in a tissue culture room or any other room that is designed for a purpose vital to the whole laboratory. An accident involving this much ^{32}P would seriously inconvenience future work in the area, if not make it altogether uninhabitable! If care is taken to minimize the amount of time the dish of cells is open when adding the label, use of a controlled air hood to prevent fungal or bacte-

rial contamination of the cells should not be necessary.

Once the label has been added to the dishes of cells, they will also need to be shielded for transport to and from the incubator and other work areas. Plexiglas boxes that are open at one end (for insertion of the dishes) and have a handle on top (for safe carrying) make ideal “cell houses” (see Figure A.1F.2A). A Plexiglas door that slides into grooves at the open end is important to prevent dishes from sliding out if the box is tilted at all during transport. If this door is only two-thirds the height of the house wall, the open slot thus created will allow equilibration of the CO_2 level within the house with that in the incubator. Obviously, this slot will also allow a substantial stream of radiation

to pass out of the cell house, so the house should be carried and placed in the incubator with its door facing away from the worker (and others)!

Following incubation with label and any treatments or other experimental manipulations, the cells are usually lysed in some type of detergent buffer. It is probably during this lysis procedure that a worker's hands will receive their greatest exposure to radiation, because it is necessary to directly handle the dishes over a period of several minutes. It is therefore very important to streamline this procedure and use any shielding whenever possible. If the cell lysates must be made at 4°C, as required by most protocols, working on a bench in a cold room is preferable to placing the dishes on a slippery bed of ice. In either case, make the lysate using the same sort of shielding (with lead if necessary) that was used when initially adding the label. Pipet the labeling medium and any solution used to rinse unincorporated radioactivity from the cells into a small tube held in a solid Plexiglas holder (shown in Fig. A.1F.1C). The contents of this tube can later be poured into the appropriate liquid waste receptacle. If possible, it is a good practice to keep this high-specific-activity ^{32}P liquid waste separate from the lower-activity waste generated in other procedures so that it can be removed from the laboratory as soon as possible following the experiment. If it is necessary to store it in the laboratory for any time, the shielding for the waste container should also include a layer of lead.

The solid waste generated in the lysis part of these experiments (pipet tips, disposable pipettes, cell scrapers, and dishes) is very hot and should be placed immediately into some

sort of shielded container to avoid further exposure of the hands. A Plexiglas box similar in design to that in Figure A.1F.3 is convenient; placed to the side of the shield and lined with a plastic bag, it will safely hold all radioactive waste during the experiment and is light enough to be easily carried to the main laboratory waste container where the plastic bag (and its contents) can be dumped after the experiment is done. If the lid of the box protrudes an inch or so over the front wall, it can be lifted using the back of a hand, thus decreasing the possibility of contaminating it with hot gloves.

When scraping the cell lysates from the dishes, it is good practice to add them to microcentrifuge tubes that are shielded in a solid Plexiglas rack; this will help to further reduce the exposure to which the hands are subjected. At this point, the lysates are usually centrifuged at high speed ($>10,000 \times g$) to clear them of unsolubilized cell material. Use screw-cap tubes for this clarification step, as these will contain the labeled lysate more securely than flip-top tubes, which may open during centrifugation. No matter what type of tube is used, the rotor of the centrifuge often becomes contaminated, most probably due to tiny drops of lysate (aerosol) initially present on the rim of the tubes that are spun off during centrifugation. Monitor the rotor and wipe it out after each use.

The amount of ^{32}P taken up by cells during the incubation period varies considerably, depending on the growth state of the culture as well as on the cell type and its sensitivity to radiation. This makes it difficult to predict the percentage of the radioactivity initially added to the cells that is incorporated into the cell lysate; however, this figure probably does not

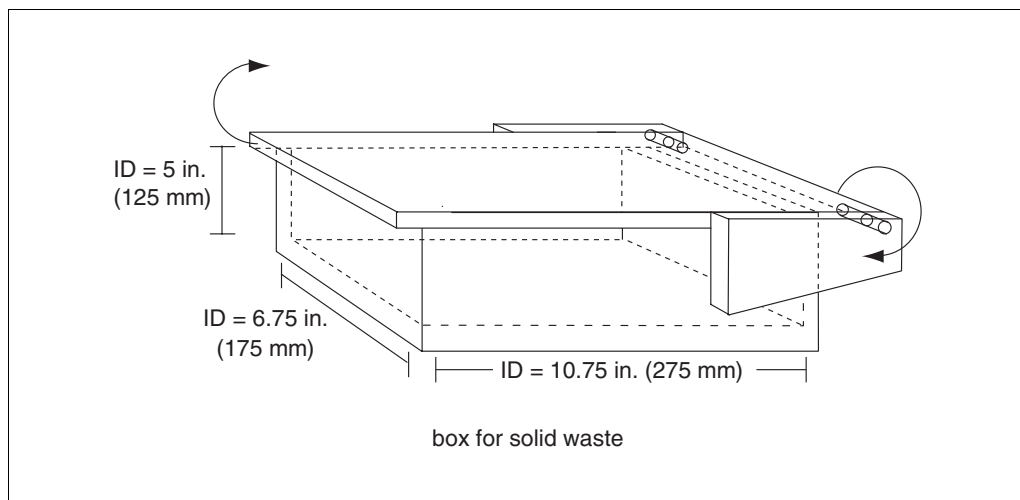


Figure A.1F.3 Box for solid waste collection made of 0.5-in. Plexiglas. Abbreviation: ID, interior dimension.

exceed 10%. Thus, the amount of radioactivity being handled decreases dramatically after lysis, making effective shielding much simpler. However, at least ten times more radioactivity than is usual in other sorts of experiments is still involved! It is easy to determine if the shielding is adequate—just use both β and γ portable monitors to measure the radiation coming through it. Again, be sure to check that people working nearby (including those across the bench) are also adequately shielded. It is sometimes necessary to construct a sort of cage of Plexiglas shields around the ice bucket that contains the lysates.

At the end of the day or the experiment, it may be necessary to store radioactive samples; in some experiments, it may be desirable to save the cell lysates. These very hot samples are best stored in tubes placed in solid Plexiglas racks that can then be put into Plexiglas boxes (see Fig. A.1F.2C). Such boxes may be of similar construction to the cell houses described above; however, they should have a door that completely covers the opening. Be sure to check the γ radiation coming through these layers and add lead outside if necessary.

Working with ^{33}P

Using ^{33}P -labeled nucleotides to label nucleic acid probes or proteins

Several of the major companies that manufacture radiolabeled biological molecules have recently introduced nucleotides labeled with ^{33}P (both α - and γ -labeled forms). ^{33}P offers a clear advantage over ^{32}P with respect to ease of handling, because the energy of the β particles it releases lies between that of ^{35}S and ^{32}P and thus its use does not require as many layers of Plexiglas and lead shielding as for ^{32}P . In fact, the β radiation emitted can barely penetrate through gloves and the surface layer of skin, so the hazard associated with exposure to even millicurie amounts of ^{33}P is thought to be insignificant (as reported in the DuPont NEN product brochure). Gel bands visualized on autoradiographs of ^{33}P -labeled compounds are sharper than bands labeled with ^{32}P because the lower-energy β radiation does not have the scatter associated with that from ^{32}P . The half-life of ^{33}P is also longer (25 days compared to 14 days for ^{32}P). Despite its higher cost, these features have led many researchers to choose ^{33}P -labeled nucleotides for use in experiments such as band/gel shift assays where discrimination of closely-spaced gel bands is important.

The best way to determine what degree of shielding is needed when using ^{33}P is to monitor the source using a portable β monitor and add layers of Plexiglas as necessary.

Working with ^{125}I

Using ^{125}I to detect immune complexes (Western blots)

^{125}I that is covalently attached to a molecule such as staphylococcal protein A is not volatile and therefore is much less hazardous than the unbound or free form. Most institutions do not insist that work with bound ^{125}I be performed in a hood, but shielding of the γ radiation is still necessary. Lead is a good high-density material for stopping these γ rays; its drawbacks are its weight and opacity. Commercially available shields for ^{125}I are made of lead-impregnated Plexiglas—though heavy, these are at least see-through. Alternatively, a piece of lead foil may be taped to a structural support, although this arrangement does not provide shielding for the head as a worker peers over the lead!

Incubations of the membrane or blot with the [^{125}I] protein A solution and subsequent washes are usually done on a shaker. For shielding during these steps, a piece of lead foil may simply be wrapped around the container. Solutions of ^{125}I can be conveniently stored for repeated use in a rack placed in a lead box.

Using ^{125}I to label proteins or peptides in vitro

Any experiments that call for the use of free, unbound ^{125}I should be done behind a shield in a hood that contains a charcoal filter to absorb the volatile iodine. Most institutions require that such experiments be done in a special hot lab to which access is limited. Ingested or inhaled iodine is concentrated in the thyroid; a portable γ monitor should therefore be used to scan the neck and throat before beginning and after completing each experiment. Similar scans should routinely be performed on all members of any laboratory in which unbound iodine is used.

DEALING WITH ACCIDENTS

Despite the best intentions and utmost caution, accidents happen! Accidents involving spills of radioactivity are particularly insidious because they can be virtually undetectable yet pose a significant threat to laboratory workers. For this reason it is best to foster a community spirit in any laboratory where radioisotopes are routinely used—a sense of cooperativity that

extends from shielding each other properly to helping each other clean up when such accidents occur.

The specific measures to be taken following an accident involving radioactivity naturally depend on the type and amount of the isotope involved, the chemical or biological hazards of the material it is associated with, and the physical parameters of the spill (i.e., where and onto what the isotope was “misplaced”). However, following any accident there are several immediate steps that should be taken:

1. Alert coworkers as well as Radiation Safety personnel to the fact that there has been an accident. This will give them the opportunity to shield themselves if necessary—and to help clean up as well!

2. Restrict access to and away from the site of the accident to ensure that any uncontained radioactive material is not spread around the laboratory. When leaving the site be sure to monitor the bottoms of the shoes as well as the rest of the body.

3. Take care of all contaminated personnel first, evacuating others if necessary. If anyone’s skin is contaminated, first use a portable monitor to identify specific areas of contamination. Then wipe these areas with a damp tissue to remove as much surface radioactivity as possible. Try to scrub only small areas at a time to keep the contamination localized. If the contamination is not easily removed with paper tissues, try a sponge or an abrasive pad, but be

careful not to break the skin! Sometimes soaking is required: do this only after all easily removed contamination is gone and keep the soaked area to a minimum. Contaminated strands of hair can be washed (or perhaps a new hairstyle may be in order).

4. When attempting to clean any contaminated equipment, floors, benches, etc., begin by soaking up any visible radioactive liquid with an absorbent material. Use a small amount of soap and water to clean the contaminated area, keeping the area wiped each time to a minimum to avoid smearing the contamination over an even greater surface. Many surfaces prove resistant to even Herculean cleaning efforts; in these instances the best that can be done is to remove all contamination possible and then shield whatever is left until the radioactivity decays sufficiently for safety.

LITERATURE CITED

- Meisenhelder, J. and Hunter, T. 1988. Radioactive protein-labelling techniques. *Nature* 335:120.
- Shleien, B. (ed.) 1987. Radiation Safety Manual for Users of Radioisotopes in Research and Academic Institutions. Nucleon Lectern Associates, Olney, Md.

Contributed by Jill Meisenhelder and
Kentaro Semba
The Salk Institute
La Jolla, California

Centrifuges and Rotors

Centrifugation runs described in this book usually specify a relative centrifugal force (RCF; measured in $\times g$), corresponding to a speed (in rpm) for a particular centrifuge and rotor model. As available equipment will vary from laboratory to laboratory, the investigator must be able to adapt these specifications to other centrifuges and rotors.

The relationship between RCF and speed (rpm) is determined by the following equation:

$$\text{RCF} = 1.12r (\text{rpm}/1000)^2$$

where r is the rotating radius between the particle being centrifuged and the axis of rotation. In most cases, an accurate conversion from speed to relative centrifugal force (or vice versa) can be obtained using the maximum value of r —or r_{max} —equal to the distance between the axis of rotation and the bottom of the centrifuge tube as it sits in the well or bucket of the rotor.

Table A.1G.1 provides r_{max} values for commonly used rotors manufactured by Du Pont (Sorvall), Beckman, Fisher, and IEC. There are situations (e.g., where an adapter is used to fit a smaller tube into a larger rotor well) where r_{max} will not accurately represent the effective rotating radius. In such cases, the manual for the rotor should be consulted to obtain the appropriate value of r .

As an alternative to use of the above equation, the nomograms in Figures A.1G.1 and A.1G.2 make it possible to determine the RCF where speed and r_{max} are known, or the speed where RCF and r_{max} are known. This is done by aligning a ruler across the two known values and reading the unknown value at the point where the ruler crosses the remaining column. Figure A.1G.1 should be used for centrifuge runs $<21,000$ rpm, while Figure A.1G.2 should be used for faster spins.

NOTE: In this manual, for spins involving microcentrifuges built to the Eppendorf standard, a shortened style of reference including only the speed (in rpm) is used. All of these instruments have approximately the same rotating radius; hence the same speed will yield the same RCF value from machine to machine. Microcentrifuge spins may also be described as at “top speed” or “maximum speed,” meaning 12,000 to 14,000 rpm, which is the maximum speed for all Eppendorf-type microcentrifuges.

CAUTION: Do not exceed maximum rotor speed! For Beckman ultracentrifuges, the maximum speed for each rotor is denoted by its name, e.g., the maximum speed of the Beckman VTi 80 rotor is 80,000 rpm. This speed refers only to centrifugation of solutions below a particular allowed density, which differs among rotors (see user manual). For centrifugation of high-density solutions, rotor maximum speed can be determined as: reduced rpm = rpm_{max}(A/B)^{1/2}, where A=allowed density and B=density of solution. A=1.7 g/ml for several vertical rotors (including VTi 80 and VTi 50), and 1.2 g/ml for several swinging-bucket rotors (including SW 55 Ti, 28, 28.1, 40 Ti, 50.1). For gradients using heavy salts such as CsCl, particularly at low temperatures, maximum rpm should be reduced to prevent precipitation (see user manual).

Table A.1G.2 describes centrifuge tube materials and their properties, including optical properties, appropriate methods for sterilization, and chemical resistances (tolerance to various media, organic solvents, and alcohols).

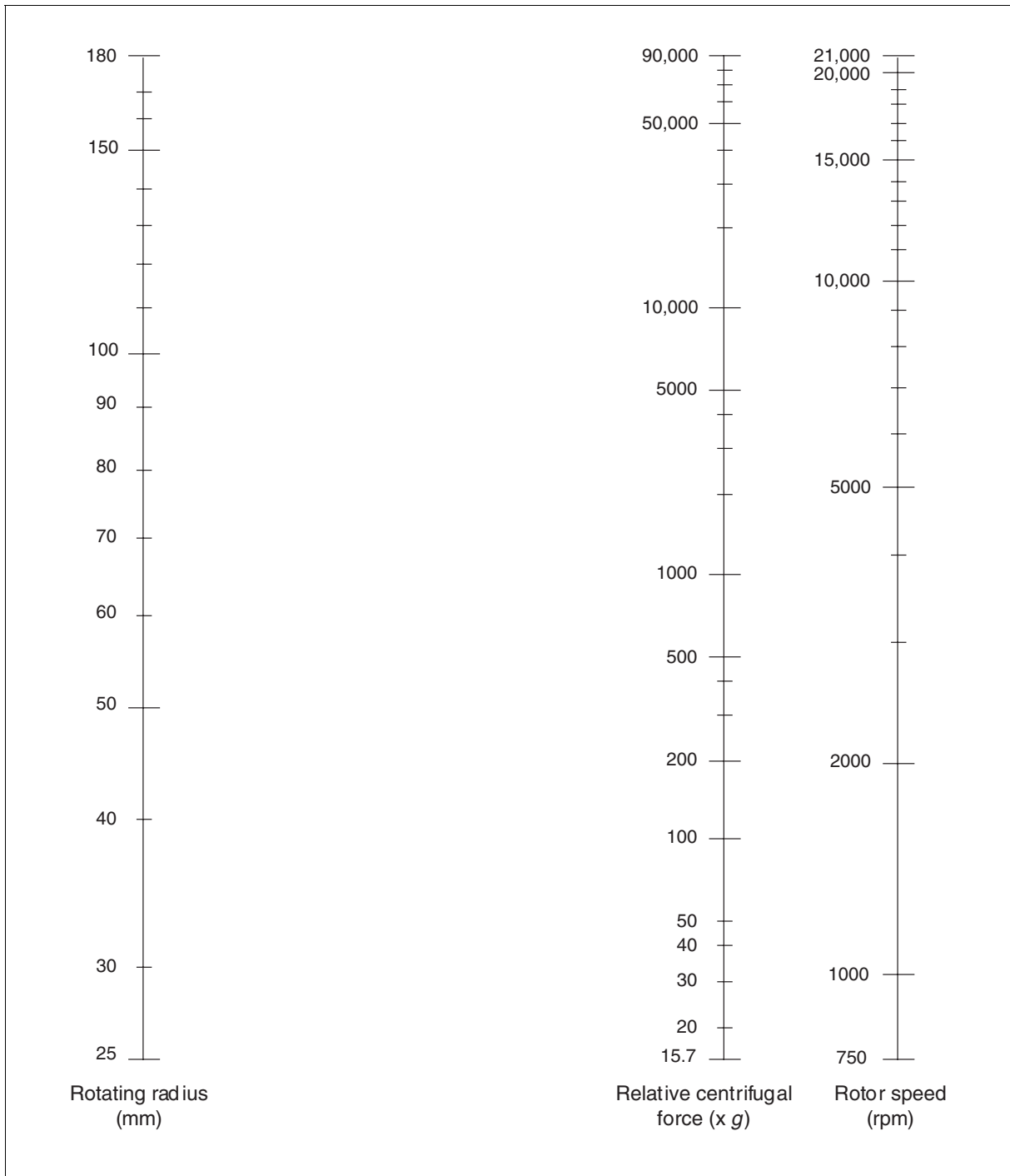


Figure A.1G.1 Nomogram for conversion of relative centrifugal force to rotor speed in low-speed centrifuge runs. To determine an unknown value in a given column, align ruler through known values in other two columns. Desired value is found at the intersection of the ruler with the column of interest. For faster centrifugations, use Figure A.1G.2. A more precise conversion can be obtained using the equation at the beginning of this appendix. See Table A.1G.1 for rotating radii of commonly used rotors.

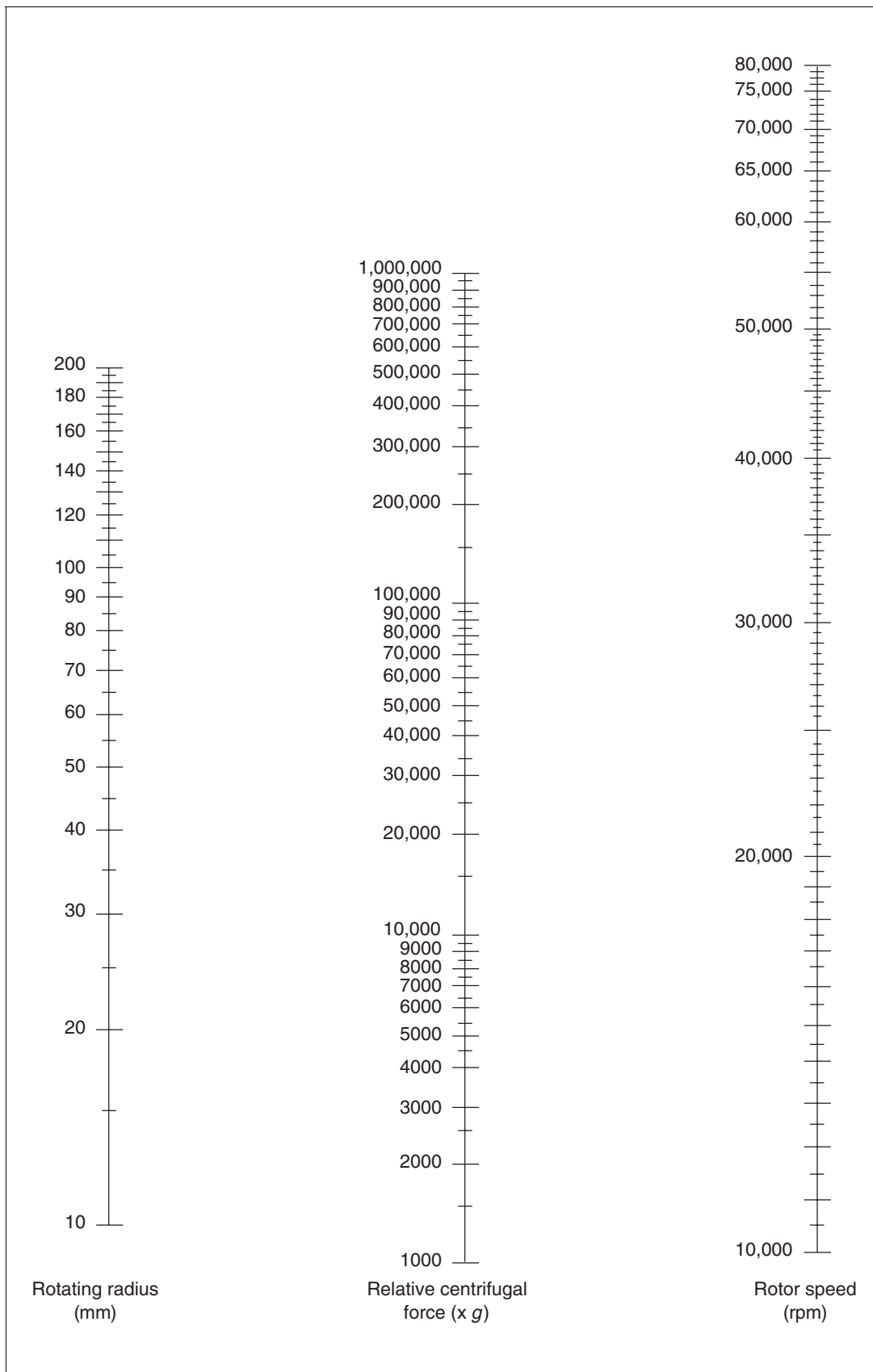


Figure A.1G.2 Nomogram for conversion of relative centrifugal force to rotor speed in high-speed centrifuge runs. For slower centrifugations and instructions for using the nomogram, use Figure A.1G.1. A more precise conversion can be obtained using the equation at the beginning of this appendix. See Table A.1G.1 for rotating radii of commonly used rotors.

A.1G.3

Table A.1G.1 Maximum Rotating Radii for Common Rotors, Grouped by Centrifuge Model

Rotor model ^a	r_{\max} (mm)	Rotor model ^a	r_{\max} (mm)
For Sorvall centrifuge models GLC-1, GLC-2, GLC-2B, GLC-3, GLC-4, RT-6000B, T-6000, T-6000B			
A/S400	140	GH-3.7 (buckets)	204
H-1000B	186	GH-3.7 (microplate carrier)	168
HL-4 with 50-ml bucket	180	GH-3.8 (buckets)	204
HL-4 with 100-ml bucket	204	GH-3.8 (microplate carrier)	168
HL-4 with Omni-Carrier	163	For Beckman TJ-6 series centrifuges	
M and A-384 (inner row)	91	TA-10	123
M and A-384 (outer row)	121	TA-24	108
SP/X and A-500 (inner row)	82	TA-24 with adapter for 10-ml tubes	123
SP/X and A-500 (outer row)	123	TH-4 (stainless steel buckets)	186
For Sorvall centrifuge models RC-3, RC-3B, RC-3C		TH-4 (100-ml tube holders)	201
H-2000B	261	TH-4 (microplate carrier)	165
H-4000 and HG-4L	230	For Beckman AccuSpin	
H-6000A	260	AA-10	123
HL-8 with Omni-Carrier	221	AA-24	108
HL-8 with 50-ml bucket	238	AA-24 with adapter for 10-ml tubes	123
HL-8 with 100-ml bucket	247	AH-4	163
HL-2 and HL-2B	166	For Beckman J6 series centrifuges	
LA/S400	140	JR-3.2	206
For Sorvall centrifuge models RC-2, RC-2B, RC-5, RC-5B, RC-5C		JS-2.9	265
GSA	145	JS-3.0	254
GS-3	151	JS-4.0	226
HB-4	147	JS-4.2	254
HS-4 with 250-ml bucket	172	JS-4.2SM	248
SA-600	129	JS-5.2	226
SE-12	93	Microplate carrier (6-bucket rotors)	214
SH-80	101	Microplate carrier (4-bucket rotors)	192
SM-24 (inner row)	91	For Beckman J2-21 series centrifuges	
SM-24 (outer row)	110	JA-10	158
SS-34	107	JA-14	137
SV-80	101	JA-17	123
SV-288	90	JA-18	132
TZ-28	95	JA-18.1 (25° angle)	112
For Sorvall ultracentrifuges		JA-18.1 (45° angle)	116
T-865	91	JA-20	108
T-865.1	87.1	JA-20.1	115
T-875	87.1	JA-21	102
T-880	84.7	JCF-Z	89
T-1270	82	JCF-Z with small pellet core	81
TFT-80.2	65.5	JE-6B	125
TFT-80.4	60.1	JS-7.5	165
For Beckman GP series centrifuges		JS-13	142
GA-10	123	JS-13.1	140
GA-24	123	JV-20	93
GA-24 with adapter for 10-ml tubes	108		

continued

Table A.1G.1 Maximum Rotating Radii for Common Rotors, Grouped by Centrifuge Model, continued

Rotor model ^a	r_{\max} (mm)	Rotor model ^a	r_{\max} (mm)
For Beckman series L7 and L8 ultracentrifuges		Type 80 Ti	84.0
SW 25.1	129.2	VAC 50	86.4
SW 28	161.0	VC 53	78.8
SW 28.1	171.3	VTi 50	86.6
SW 30	123.0	VTi 65	85.4
SW 30.1	123.0	VTi 65.2	87.9
SW 40 Ti	158.8	VTi 80	71.1
SW 41 Ti	153.1	For Beckman Airfuge ultracentrifuge	
SW 50.1	107.3	A-95	17.6
SW 55 Ti	108.5	A-100/18	14.6
SW 60 Ti	120.3	A-100/30	16.5
SW 65 Ti	89.0	A-110	14.7
Type 15	142.1	ACR-90 (2.4-ml liner)	11.8
Type 19	133.4	ACR-90 (3.5-ml liner)	13.4
Type 21	121.5	Batch rotor	14.6
Type 25	100.4	EM-90	13.0
Type 30	104.8	For Beckman TL-100 series ultracentrifuges	
Type 30.2	94.2	TLA-100	38.9
Type 35	104.0	TLA 100.1	38.9
Type 40	80.8	TLA-100.2	38.9
Type 40.3	79.5	TLA-100.3	48.3
Type 42.1	98.6	TLA-45	55.1
Type 42.2 Ti	104	TLS-55	76.4
Type 45 Ti	103	TLV-100	35.7
Type 50	70.1	Miscellaneous centrifuges and rotors^b	
Type 50 Ti	80.8	Clay Adams Dynac	— ^c
Type 50.2 Ti	107.9	Fisher Centrifuc	113
Type 50.3 Ti	79.5	Fisher Marathon 21K with	160
Type 50.4 Ti (inner row)	96.4	4-place rotor	
Type 50.4 Ti (outer row)	111.4	IEC Clinical centrifuge with	155
Type 55.2 Ti	100.3	4-place swinging-bucket	
Type 60 Ti	89.9	rotor	
Type 65	77.7	IEC general-purpose	— ^c
Type 70 Ti	91.9	centrifuge models HN,	
Type 70.1 Ti	82.0	HN-SII, and Centra-4	
Type 75 Ti	79.7		

^aSorvall centrifuges and rotors are a product of Du Pont Company Medical Products, Beckman centrifuges are a product of Beckman Instruments, IEC centrifuges are a product of International Equipment Co., Clay Adams Dynac centrifuges are a product of Becton Dickinson Labware, and Fisher centrifuges are a product of Fisher Scientific. For ordering information see *APPENDIX*.

^bThese instruments are often loosely referred to as “clinical,” “tabletop,” or “low-speed” centrifuges.

^cThese instruments accept a wide range of trunnion-ring rotors with variable rotating radii, as well as fixed-angle and swinging-bucket rotors that in turn accept a variety of adapters making it possible to spin different numbers tubes of various sizes. For instance, the commonly used IEC 958 trunnion-ring rotor may be adjusted to a radii ranging from 137 to 181 mm, depending on the trunnion-ring chosen. It is therefore necessary to consult the manual for the specific system being used to obtain an accurate speed to RCF conversion.

Table A.1G.2 Centrifuge Tube Materials and Their Properties^a

Type	Optical property	Puncturable	Sliceable	Reusable	Sterilization methods	Chemical resistances ^b
Ultra-Clear thin-walled Standard tubes Quick-Seal tubes	Transparent	Yes	Yes	Yes No	Cold sterilization only, but not with alcohol	Good tolerance to all gradient media except alkaline ones (>pH 8). Satisfactory for most weak acids and a few weak bases. Unsatisfactory for DMSO and most organic solvents, including all alcohols.
Polyallomer thin-walled Standard tubes Quick-Seal tubes	Translucent	Yes	Yes	Yes No	Can be autoclaved on a test tube rack at 121°C	Good tolerance to all gradient media, including alkaline ones. Satisfactory for most acids, many bases, many alcohols, DMSO, and some organic solvents.
Polyallomer thick-walled Tubes Bottles	Translucent	No No	No No	Yes Yes	Can be autoclaved on a test tube rack at 121°C	Good tolerance to all gradient media, including alkaline ones. Satisfactory for most acids, many bases, many alcohols, DMSO, and some organic solvents.
Polycarbonate thick-walled Tubes Bottles	Transparent	No	No	Yes	Cold sterilization recommended, but not with alcohol. Can be autoclaved at 121°C, but tube life may be reduced.	Good tolerance to all gradient media except alkaline ones (>pH 8). Satisfactory for some weak acids. Unsatisfactory for all bases, alcohols, and other organic solvents.
Cellulose propionate tubes	Transparent	No	No	Yes	Cold sterilization only, but not with alcohol	Good tolerance to all gradient media, including alkaline ones. Unsatisfactory for most acids, bases, alcohols, and other organic solvents.
Polypropylene Tubes Bottles	Translucent	No	No	Yes Yes	Can be autoclaved at 121°C	Good tolerance to all gradient media, including alkaline ones. Satisfactory for many acids, bases, and alcohols. Unsatisfactory for most organic solvents.
Stainless steel tubes	Opaque	No	No	Yes	Can be autoclaved. Dry thoroughly before storage.	Good tolerance to many organic solvents. Marginal with many gradient media and salts. Unsatisfactory for most acids and many bases.
Polyethylene tubes	Translucent	No	No	Yes	Can be autoclaved at 121°C	Good tolerance to a wide range of chemicals. Suitable for use with strong acids and bases. Unsatisfactory for most organic solvents.
Corex/Pyrex Tubes Bottles	Transparent	No	No	Yes Yes	Can be autoclaved at 121°C	Good tolerance to a wide range of gradient media. Corex has greater resistance to bases and acids.

^aTable reproduced by permission of Beckman Instruments.

^bChemical resistances are described here in general terms, and are not meant by Beckman Instruments or John Wiley & Sons to express or imply any guarantee of safety based on these recommendations or resistances. If there is any doubt about a particular solution, it should be tested under actual operating conditions to evaluate the performance of a tube material. For more detailed information regarding specific media and solvents, consult Beckman Instruments. High-vapor-pressure inflammable solvents should not be handled in close vicinity to centrifuges because of possible ignition by sparking switches, relay contacts, or motor brushes.

Safe Use of Hazardous Chemicals

Carrying out the protocols in this manual may result in exposure to toxic chemicals or carcinogenic, mutagenic, or teratogenic reagents (see Table A.1H.1). Cautionary notes and some specific guidelines are included in many instances throughout the manual; however, users must proceed with the prudence and precautions associated with good laboratory practice, under the supervision of those responsible for implementing lab safety programs at their institutions.

It is not possible in the space available to list all the precautions required for handling hazardous chemicals. Many texts have been written about laboratory safety (see Literature Cited and Key References). Obviously, all national and local laws should be obeyed, as well as all institutional regulations. Controlled substances are regulated by the Drug Enforcement Administration (<http://www.doj.gov/dea>). By law, Material Safety Data Sheets (MSDSs) must be readily available. All laboratories should have a Chemical Hygiene Plan (29CFR Part 1910.1450); institutional safety officers should be consulted as to its implementation. Help is (or should be) available from your institutional Safety Office; use it.

Chemicals must be stored properly for safety. Certain chemicals cannot be easily or safely mixed with and should not be stored near certain other chemicals, because their reaction is violently exothermic or yields a toxic product. Some examples of incompatibility are listed in Table A.1H.2. When in doubt, always consult a current MSDS for information on reactivity, handling, and storage. Chemicals should be separated into general hazard classes and stored appropriately. For example, flammable chemicals such as alcohols, ketones, aliphatic and aromatic hydrocarbons, and other materials labeled flammable should be stored in approved flammable storage cabinets, with those also requiring refrigeration being kept in explosion-proof refrigerators. Strong oxidizers must be segregated. Strong acids (e.g., sulfuric, hydrochloric, nitric, perchloric, and hydrofluoric) should be stored in a separate cabinet well removed from strong bases and from flammable organics. An exception is glacial acetic acid, which is both corrosive and flammable, and which must be stored with the flammables.

Facilities should be appropriate for working with hazardous chemicals. In particular, hazardous chemicals should be handled only in chemical fume hoods, not in laminar flow cabinets. The functioning of the fume hoods should be checked periodically. Laboratories should also be equipped with safety showers and eye-wash facilities. Again, this equipment should be tested periodically to ensure that it functions correctly. Other safety equipment may be required depending on the nature of the materials being handled. In addition, researchers should be trained in the proper procedures for handling hazardous chemicals as well as other laboratory operations—e.g., handling of compressed gases, use of cryogenic liquids, operation of high-voltage power supplies, and operation of lasers of all types.

Before starting work, know the physical and chemical hazards of the reagents used. Wear appropriate protective clothing and have a plan for dealing with spills or accidents; coming up with a good plan on the spur of the moment is very difficult. For example, have the appropriate decontaminating or neutralizing agents prepared and close at hand. Small spills can probably be cleaned up by the researcher. In the case of larger spills, the area should be evacuated and help should be sought from those experienced in and equipped for dealing with spills—e.g., the institutional Safety Office.

Protective equipment should include, at a minimum, eye protection, a lab coat, and gloves. In certain circumstances other items of protective equipment may be necessary (e.g., a face shield). Different types of gloves exhibit different resistance properties (Table

Table A.1H.1 Commonly Used Hazardous Chemicals^a

Chemical	Hazards	Remarks ^b
Acetic acid, glacial	Corrosive, flammable liquid	
Acetonitrile	Flammable liquid, teratogenic, toxic	
Acridine orange	Carcinogenic, mutagenic	See Basic Protocol 2
Acrylamide	Carcinogenic, toxic	Use dust mask; polyacrylamide gels contain residual acrylamide monomer and should be handled with gloves; acrylamide may polymerize with violence on melting at 86°C
Alcian blue 8GX		See Basic Protocol 2
Alizarin red S (monohydrate)		
<i>p</i> -Amidinophenylmethanesulfonyl fluoride (APMSF)	Enzyme inhibitor	See Basic Protocol 11
7-Aminoactinomycin D (7-AAD)	Carcinogenic	
4-(2-Aminoethyl)benzenesulfonyl fluoride (AEBSF)	Mutagenic, enzyme inhibitor	See Basic Protocol 11
Ammonium hydroxide, concentrated	Corrosive, lachrymatory, toxic	
Azure A	Mutagenic	See Basic Protocol 2
Azure B	Mutagenic	See Basic Protocol 2
Benzidine (BDB)	Carcinogenic, toxic	See Basic Protocol 1
Bisacrylamide	Toxic	
Boron dipyrromethane derivatives (BODIPY dyes)	Toxic	
Brilliant blue R	Carcinogenic, mutagenic	See Basic Protocol 2
5-Bromodeoxyuridine (BrdU)	Mutagenic, teratogenic, photosensitizing	
Cetylpyridinium chloride (CPC)	Toxic	
Cetyltrimethylammonium bromide (CTAB)	Corrosive, teratogenic, toxic	
Chloroform	Carcinogenic, teratogenic, toxic	
Chlorotrimethylsilane	Carcinogenic, corrosive, flammable liquid, toxic	Reacts violently with water; see Basic Protocol 3
Chromic/sulfuric acid cleaning solution	Carcinogenic, corrosive, oxidizer, toxic	Replace with suitable commercially available cleanser
Chromomycin A3 (CA3)	Teratogenic, toxic	
Congo red	Mutagenic, teratogenic	See Basic Protocol 2
Coomassie brilliant blue G	Mutagenic	See Basic Protocol 2
Crystal violet		See Basic Protocol 2
Cresyl violet acetate	Mutagenic	See Basic Protocol 2
Cyanides (e.g., KCN, NaCN)	Toxic	Contact with acid will liberate HCN gas; see Basic Protocol 4
Cyanines (e.g., Cy3, Cy5)	Toxic	
Cyanogen bromide (CNBr)	Toxic	See Basic Protocol 4

*continued***A.1H.2**

Table A.1H.1 Commonly Used Hazardous Chemicals^a, continued

Chemical	Hazards	Remarks ^b
2'-Deoxycoformycin (dCF, pentostatin)	Teratogenic, toxic	
4',6-Diamidino-2-phenylindole (DAPI)	Mutagenic	
Diaminobenzidine (DAB)	Carcinogenic	See Basic Protocol 1
1,4-Diazabicyclo[2,2,2]-octane (DABCO)	Toxic	Forms an explosive complex with hydrogen peroxide
Dichloroacetic acid (DCA)	Carcinogenic, corrosive, toxic	
Dichloromethane (methylene chloride)	Carcinogenic, mutagenic, teratogenic, toxic	
Diethylamine (DEA)	Corrosive, flammable liquid, toxic	
Diethylpyrocarbonate (DEPC)	Carcinogenic, toxic	
Diethyl sulfate	Carcinogenic, teratogenic, toxic	See Basic Protocol 5
Diisopropyl fluorophosphate (DFP)	Highly toxic, cholinesterase inhibitor, neurotoxin	See Basic Protocol 11
Dimethyldichlorosilane	Corrosive, flammable liquid, toxic	See Basic Protocol 3
Dimethyl sulfate (DMS)	Carcinogenic, toxic	See Basic Protocol 5
Dimethyl sulfoxide (DMSO)	Flammable liquid, toxic	Enhances absorption through skin
Diphenylamine (DPA)	Teratogenic, toxic	
2,5-Diphenyloxazole (PPO)	Toxic	
Dithiothreitol (DTT)	Toxic	
Eosin B		See Basic Protocol 2
Erythrosin B	Carcinogenic, mutagenic	See Basic Protocol 2
Ether	Flammable liquid, toxic	May form explosive peroxides on standing; do not dry with NaOH or KOH
Ethidium bromide (EB)	Mutagenic, toxic	See Basic Protocol 2 or 6
Ethyl methanesulfonate (EMS)	Carcinogenic, toxic	See Basic Protocol 5
Fluorescein and derivatives	Carcinogenic, toxic	
5-Fluoro-2'-deoxyuridine (FUdR)	Teratogenic, toxic	
Fluoroorotic acid (FOA)	Toxic	
Formaldehyde	Carcinogenic, flammable liquid, teratogenic, toxic	
Formamide	Teratogenic, toxic	
Formic acid	Corrosive, toxic	May explode when heated >180°C in a sealed tube
Glutaraldehyde	Corrosive, teratogenic, toxic	
Guanidinium thiocyanate	Toxic	
Hoechst 33258 dye	Mutagenic, toxic	
Hydrochloric acid, concentrated	Corrosive, teratogenic, toxic	

continued

**Standard
Measurements,
Data, and
Abbreviations**

A.1H.3

Table A.1H.1 Commonly Used Hazardous Chemicals^a, continued

Chemical	Hazards	Remarks ^b
Hydrogen peroxide (30%)	Carcinogenic, corrosive, mutagenic, oxidizer	Avoid bringing into contact with organic materials, which may form explosive peroxides; may decompose violently in contact with metals, salts, or oxidizable materials; see Basic Protocol 7
Hydroxylamine	Corrosive, flammable, mutagenic, toxic	Explodes in air at >70°C
3-β-Indoleacrylic acid (IAA)	Carcinogenic	
Iodine	Corrosive, toxic	See Basic Protocol 8
Iodoacetamide	Carcinogenic, mutagenic, toxic	
Janus green B	Carcinogenic, mutagenic	See Basic Protocol 2
Lead compounds	Carcinogenic, toxic	
2-Mercaptoethanol (2-ME)	Stench, toxic	
Mercury compounds	Teratogenic, toxic	See Basic Protocol 9
Methionine sulfoximine (MSX)	Teratogenic, toxic	
Methotrexate (amethopterin)	Carcinogenic, mutagenic, teratogenic, toxic	
Methylene blue	Mutagenic, toxic	See Basic Protocol 2
Methyl methanesulfonate (MMS)	Carcinogenic, toxic	See Basic Protocol 5
Mycophenolic acid (MPA)	Teratogenic, toxic	
Neutral red	Mutagenic	See Basic Protocol 2
Nigrosin, water soluble		See Basic Protocol 2
Nitric acid, concentrated	Corrosive, oxidizer, teratogenic, toxic	
Nitroblue tetrazolium (NBT)	Toxic	
Orcein, synthetic		See Basic Protocol 2
Oxonols	Toxic	
Paraformaldehyde	Toxic	
Phenol	Carcinogenic, corrosive, teratogenic, toxic	Readily absorbed through the skin
Phenylmethylsulfonyl fluoride (PMSF)	Enzyme inhibitor	See Basic Protocol 11
Phorbol 12-myristate 13-acetate (PMA)	Carcinogenic, toxic	
Phycoerythrins (PE)	Toxic	
Piperidine	Flammable liquid, teratogenic, toxic	
Potassium hydroxide, concentrated	Corrosive, toxic	Produces a highly exothermic reaction when solid is added to water
Propane sultone	Carcinogenic, toxic	See Basic Protocol 5
Propidium iodide (PI)	Mutagenic	See Basic Protocol 2 or 6
Pyridine	Flammable liquid, toxic	
Rhodamine and derivatives	Toxic	
Rose Bengal	Carcinogenic, teratogenic	See Basic Protocol 2
Safranin O	Mutagenic	See Basic Protocol 2

*continued***A.1H.4**

Table A.1H.1 Commonly Used Hazardous Chemicals^a, continued

Chemical	Hazards	Remarks ^b
Sodium azide	Carcinogenic, toxic	Adding acid liberates explosive volatile, toxic hydrazoic acid; can form explosive heavy metal azides, e.g., with plumbing fixtures— <i>do not</i> discharge down drain; see Basic Protocol 10
Sodium deoxycholate (Na-DOC)	Carcinogenic, teratogenic, toxic	
Sodium dodecyl sulfate (sodium lauryl sulfate, SDS)	Sensitizing, toxic	
Sodium hydroxide, concentrated	Corrosive, toxic	A highly exothermic reaction ensues when the solid is added to water
Sodium nitrite	Carcinogenic	
Sulfuric acid, concentrated	Corrosive, oxidizer, teratogenic, toxic	Reaction with water is very exothermic; always add concentrated sulfuric acid to water, <i>never</i> water to acid
SYTO dyes	Toxic	
Tetramethylammonium chloride (TMAC)	Toxic	
<i>N,N,N',N'</i> -Tetramethyl-ethylenediamine (TEMED)	Corrosive, flammable liquid, toxic	
Texas Red (sulforhodamine 101, acid chloride)	Toxic	
Toluene	Flammable liquid, teratogenic, toxic	
Toluidine blue O	Mutagenic, toxic	See Basic Protocol 2
<i>N</i> α- <i>p</i> -Tosyl-L-lysine chloromethyl ketone (TLCK)	Toxic, enzyme inhibitor	See Basic Protocol 11
<i>N</i> - <i>p</i> -Tosyl-L-phenylalanine chloromethyl ketone (TPCK)	Toxic, mutagenic, enzyme inhibitor	See Basic Protocol 11
Trichloroacetic acid (TCA)	Carcinogenic, corrosive, teratogenic, toxic	
Triethanolamine acetate (TEA)	Carcinogenic, toxic	
Trifluoroacetic acid (TFA)	Corrosive, toxic	
Trimethyl phosphate (TMP)	Carcinogenic, mutagenic, teratogenic	May explode on distillation
Trypan blue	Carcinogenic, mutagenic, teratogenic	See Basic Protocol 2
Xylenes	Flammable liquid, teratogenic, toxic	

^aFor extensive information on the hazards of these and other chemicals as well as cautionary details, see Bretherick (1986), O'Neil (2001), Furr (2000), Lewis (1999), Lunn and Sansone (1994a), and Bretherick et al. (1999).

^b**CAUTION:** These chemicals should be handled only in a chemical fume hood by knowledgeable workers equipped with eye protection, lab coat, and gloves. The laboratory should be equipped with a safety shower and eye wash. Additional protective equipment may be required.

A.1H.3). No gloves resist all chemicals, and no gloves resist any chemicals indefinitely. Disposable gloves labeled “exam” or “examination” are primarily for protection from biological materials (e.g., viruses, bacteria, feces, blood). They are not designed for and usually have not been tested for resistance to chemicals. Disposable gloves generally offer *extremely* marginal protection from chemical hazards in most cases and should be removed immediately upon contamination before the chemical can pass through. If possible, design handling procedures to eliminate or reduce potential for contamination. Never assume that disposable gloves will offer the same protection or even have the same

**Standard
Measurements,
Data, and
Abbreviations**

A.1H.5

Table A.1H.2 Examples of Chemical Incompatibility

Chemical	Incompatible with
Acetic acid	Aldehydes, bases, carbonates, chromic acid, ethylene glycol, hydroxides, hydroxyl compounds, metals, nitric acid, oxidizers, perchloric acid, peroxides, phosphates, permanganates, xylene
Acetone	Acids, amines, concentrated nitric and sulfuric acid mixtures, oxidizers, plastics
Acetylene	Copper, halogens, mercury, oxidizers, potassium, silver
Alkali metals, alkaline earth metals	Acids, aldehydes, carbon dioxide, carbon tetrachloride or other chlorinated hydrocarbons, halogens, ketones, plastics, sulfur, water
Ammonia (anhydrous)	Acids, aldehydes, amides, calcium hypochlorite, hydrofluoric acid, halogens, heavy metals, mercury, oxidizers, plastics, sulfur
Ammonium nitrate	Acids, alkalis, chlorates, chloride salts, flammable and combustible materials, metals, organic materials, phosphorus, reducing agents, sulfur, urea
Aniline	Acids, aluminum, dibenzoyl peroxide, oxidizers, plastics
Arsenical materials	Any reducing agent
Azides	Acids, heavy metals, oxidizers
Bromine	Acetaldehyde, alcohols, alkalis, amines, ammonia, combustible materials, ethylene, fluorine, hydrogen, ketones (e.g., acetone, carbonyls), metals, petroleum gases, sodium carbide, sulfur
Calcium oxide	Acids, ethanol, fluorine, organic materials, water
Carbon (activated)	Alkali metals, calcium hypochlorite, halogens, oxidizers
Carbon tetrachloride	Sodium
Chlorates	Acids, ammonium salts, finely divided organic or combustible materials, powdered metals, sulfur
Chlorine	Acetylene or other hydrocarbons, alcohols, ammonia, benzene, butadiene, butane, combustible materials, ethylene, flammable compounds (e.g., hydrazine), hydrogen, hydrogen peroxide, iodine, metals, methane, nitrogen, oxygen, propane (or other petroleum gases), sodium carbide, sodium hydroxide
Chlorine dioxide	Ammonia, hydrogen, hydrogen sulfide, mercury, methane, organic materials, phosphine, phosphorus, potassium hydroxide, sulfur
Chromic acid, chromic oxide	Acetic acid, acetone, alcohols, alkalis, ammonia, bases, benzene, camphor, flammable liquids, glycerin (glycerol), hydrocarbons, metals, naphthalene, organic materials, phosphorus, plastics
Copper	Acetylene, calcium, hydrocarbons, hydrogen peroxide, oxidizers
Cumene hydroperoxide	Acids (organic or inorganic)
Cyanides	Acids, alkaloids, aluminum, iodine, oxidizers, strong bases
Flammable liquids	Ammonium nitrate, chromic acid, halogens, hydrogen peroxide, nitric acid, oxidizing agents in general, oxygen, sodium peroxide
Fluorine	All other chemicals

continued

Table A.1H.2 Examples of Chemical Incompatibility, continued

Chemical	Incompatible with
Hydrocarbons (liquid or gas)	See flammable liquids
Hydrocyanic acid	Alkali, nitric acid
Hydrofluoric acid	Ammonia, metals, organic materials, plastics, silica (glass, including fiberglass), sodium
Hydrogen peroxide	All organics, most metals or their salts, nitric acid, phosphorus, sodium, sulfuric acid
Hydrogen sulfide	Acetaldehyde, fuming nitric acid, metals, oxidizers, sodium, strong bases
Hydroperoxide	Reducing agents
Hypochlorites	Acids, activated carbon
Iodine	Acetaldehyde, acetylene, ammonia, hydrogen, metals, sodium
Mercury	Acetylene, aluminum, amines, ammonia, calcium, fulminic acid, lithium, oxidizers, sodium
Nitric acid	Acids, nitrites, metals, most organics, plastics, sodium, sulfur, sulfuric acid
Nitrites	Acids
Nitroparaffins	Amines, inorganic bases
Oxalic acid	Mercury, oxidizers, silver, sodium chlorite
Oxygen	All flammable and combustible materials, ammonia, carbon monoxide, grease, metals, oil, phosphorus, polymers
Perchloric acid	All organics, bismuth and alloys, dehydrating agents, grease, hydrogen halides, iodides, paper, wood
Peroxides, organic	Acids (organic or mineral), avoid friction, store cold
Phosphorus (white)	Air, alkalis, oxygen, reducing agents
Potassium chlorate	Acids, ammonia, combustible materials, fluorine, hydrocarbons, metals, organic materials, reducing agents, sugars
Potassium perchlorate	Alcohols, combustible materials, fluorine, hydrazine, metals, organic matter, reducing agents, sulfuric acid
Potassium permanganate	Benzaldehyde, ethylene glycol, glycerin, sulfuric acid
Selenides and tellurides	Reducing agents
Silver	Acetylene, ammonium compounds, fulminic acid, oxalic acid, ozonides, peroxyformic acid, tartaric acid
Sodium	Acids, carbon dioxide, carbon tetrachloride, hydrazine, metals, oxidizers, water
Sodium nitrate	Acetic anhydride, acids, metals, organic matter, peroxyformic acid, reducing agents
Sodium peroxide	Acetic anhydride, benzaldehyde, benzene, carbon disulfide, ethyl acetate, ethyl or methyl alcohol, ethylene glycol, furfural, glacial acetic acid, glycerin, hydrogen sulfide, metals, methyl acetate, oxidizers, peroxyformic acid, phosphorus, reducing agents, sugars, water
Sulfides	Acids
Sulfuric acid	Alcohols, bases, chlorates, perchlorates, permanganates of potassium, lithium, sodium, magnesium, calcium

Table A.1H.3 Chemical Resistance of Commonly Used Gloves^{a,b}

Chemical	Neoprene gloves	Latex gloves	Butyl gloves	Nitrile gloves
*Acetaldehyde	VG	G	VG	G
Acetic acid	VG	VG	VG	VG
*Acetone	G	VG	VG	P
Ammonium hydroxide	VG	VG	VG	VG
*Amyl acetate	F	P	F	P
Aniline	G	F	F	P
*Benzaldehyde	F	F	G	G
*Benzene	P	P	P	F
Butyl acetate	G	F	F	P
Butyl alcohol	VG	VG	VG	VG
Carbon disulfide	F	F	F	F
*Carbon tetrachloride	F	P	P	G
*Chlorobenzene	F	P	F	P
*Chloroform	G	P	P	E
Chloronaphthalene	F	P	F	F
Chromic acid (50%)	F	P	F	F
Cyclohexanol	G	F	G	VG
*Dibutyl phthalate	G	P	G	G
Diisobutyl ketone	P	F	G	P
Dimethylformamide	F	F	G	G
Diocetyl phthalate	G	P	F	VG
Epoxy resins, dry	VG	VG	VG	VG
*Ethyl acetate	G	F	G	F
Ethyl alcohol	VG	VG	VG	VG
*Ethyl ether	VG	G	VG	G
*Ethylene dichloride	F	P	F	P
Ethylene glycol	VG	VG	VG	VG
Formaldehyde	VG	VG	VG	VG
Formic acid	VG	VG	VG	VG
Freon 11, 12, 21, 22	G	P	F	G
*Furfural	G	G	G	G
Glycerin	VG	VG	VG	VG
Hexane	F	P	P	G
Hydrochloric acid	VG	G	G	G
Hydrofluoric acid (48%)	VG	G	G	G
Hydrogen peroxide (30%)	G	G	G	G
Ketones	G	VG	VG	P
Lactic acid (85%)	VG	VG	VG	VG
Linseed oil	VG	P	F	VG
Methyl alcohol	VG	VG	VG	VG
Methylamine	F	F	G	G
Methyl bromide	G	F	G	F
*Methyl ethyl ketone	G	G	VG	P
*Methyl isobutylketone	F	F	VG	P
Methyl methacrylate	G	G	VG	F
Monoethanolamine	VG	G	VG	VG

continued

Table A.1H.3 Chemical Resistance of Commonly Used Gloves^{a,b}, continued

Chemical	Neoprene gloves	Latex gloves	Butyl gloves	Nitrile gloves
Morpholine	VG	VG	VG	G
Naphthalene	G	F	F	G
Naphthas, aliphatic	VG	F	F	VG
Naphthas, aromatic	G	P	P	G
*Nitric acid	G	F	F	F
Nitric acid, red and white fuming	P	P	P	P
Nitropropane (95.5%)	F	P	F	F
Oleic acid	VG	F	G	VG
Oxalic acid	VG	VG	VG	VG
Palmitic acid	VG	VG	VG	VG
Perchloric acid (60%)	VG	F	G	G
Perchloroethylene	F	P	P	G
Phenol	VG	F	G	F
Phosphoric acid	VG	G	VG	VG
Potassium hydroxide	VG	VG	VG	VG
Propyl acetate	G	F	G	F
<i>i</i> -Propyl alcohol	VG	VG	VG	VG
<i>n</i> -Propyl alcohol	VG	VG	VG	VG
Sodium hydroxide	VG	VG	VG	VG
Styrene (100%)	P	P	P	F
Sulfuric acid	G	G	G	G
Tetrahydrofuran	P	F	F	F
*Toluene	F	P	P	F
Toluene diisocyanate	F	G	G	F
*Trichloroethylene	F	F	P	G
Triethanolamine	VG	G	G	VG
Tung oil	VG	P	F	VG
Turpentine	G	F	F	VG
*Xylene	P	P	P	F

^aPerformance varies with glove thickness and duration of contact. An asterisk indicates limited use. Abbreviations: VG, very good; G, good; F, fair; P, poor (do not use).

properties as nondisposables. Select gloves carefully and always look for some evidence that they will protect against the materials being used. Inspect all gloves before every use for possible holes, tears, or weak areas. Never reuse disposable gloves. Clean reusable gloves after each use and dry carefully inside and out. Observe all common-sense precautions—e.g., do not pipet by mouth, keep unauthorized persons away from hazardous chemicals, do not eat or drink in the lab, wear proper clothing in the lab (sandals, open-toed shoes, and shorts are not appropriate).

Order hazardous chemicals only in quantities that are likely to be used in a reasonable time. Buying large quantities at a lower unit cost is no bargain if someone (perhaps you) has to pay to dispose of surplus quantities. Substitute alcohol-filled thermometers for mercury-filled thermometers, which are a hazardous chemical spill waiting to happen.

Although any number of chemicals commonly used in laboratories are toxic if used improperly, the toxic properties of a number of reagents require special mention. Chemicals that exhibit carcinogenic, corrosive, flammable, lachrymatory, mutagenic, oxidizing, teratogenic, toxic, or other hazardous properties are listed in Table A.1H.1. Chemicals listed as carcinogenic range from those accepted by expert review groups as causing cancer in humans to those for which only minimal evidence of carcinogenicity exists. No effort has been made to differentiate the carcinogenic potential of the compounds in Table A.1H.1. Oxidizers may react violently with oxidizable material (e.g., hydrocarbons, wood, and cellulose). Before using any of these chemicals, thoroughly investigate all its characteristics. Material Safety Data Sheets are readily available; they list some hazards but vary widely in quality. A number of texts describing hazardous properties are listed at the end of this unit (see Literature Cited). In particular, Sax's *Dangerous Properties of Industrial Materials*, 10th ed. (Lewis, 1999) and the *Handbook of Reactive Chemical Hazards*, 6th ed. (Bretherick et al., 1999) give comprehensive listings of known hazardous properties; however, these texts list only the known properties. Many chemicals, especially fluorochromes, have been tested only partially or not at all. Prudence dictates that, unless there is good reason for believing otherwise, all chemicals should be regarded as volatile, highly toxic, flammable human carcinogens and should be handled with great care.

Waste should be segregated according to institutional requirements, for example, into solid, aqueous, nonchlorinated organic, and chlorinated organic material, and should always be disposed of in accordance with all applicable federal, state, and local regulations. Extensive information and cautionary details along with techniques for the disposal of chemicals in laboratories have been published (Bretherick, 1986; Lunn and Sansone, 1994a; O'Neil, 2001; Furr, 2000). Some commonly used disposal procedures are outlined in Basic Protocols 1 to 11. Incorporation of these procedures into laboratory protocols can help to minimize waste disposal problems. Alternate Protocols 1 to 7 describe decontamination methods for some of the chemicals. Support Protocols 1 to 9 describe analytical techniques that are used to verify that reagents have been decontaminated; with modification, these assays may also be used to determine the concentration of a particular chemical.

DISPOSAL METHODS

A number of procedures for the disposal of hazardous chemicals are available; protocols for the disposal and decontamination of some hazardous chemicals commonly encountered in molecular biology laboratories are listed in Table A.1H.4. These procedures are necessarily brief; for full details consult the original references or a collection of these procedures (see Lunn and Sansone, 1994a).

CAUTION: These disposal methods should be carried out only in a chemical fume hood by workers equipped with eye protection, a lab coat, and gloves. Additional protective equipment may be necessary.

DISPOSAL OF BENZIDINE AND DIAMINOBENZIDINE

Benzidine and diaminobenzidine can be degraded by oxidation with potassium permanganate (Castegnaro et al., 1985; Lunn and Sansone, 1991a). This protocol presents a method for decontamination of benzidine and diaminobenzidine in bulk. This method can also be adapted to the decontamination of benzidine and diaminobenzidine spills (see Alternate Protocol 1). These compounds can also be removed from solution using horseradish peroxidase in the presence of hydrogen peroxide (see Alternate Protocol 2). Destruction and decontamination are >99%. Support Protocol 1 is used to test for the presence of benzidine and diaminobenzidine.

BASIC PROTOCOL 1

Safe Use of Hazardous Chemicals

A.1H.10

Table A.1H.4 Protocols for Disposal of Some Hazardous Chemicals

Protocol	Disposal method for
Basic Protocol 1	Benzidine and diaminobenzidine
Alternate Protocol 1	Spills of benzidine and diaminobenzidine
Alternate Protocol 2	Aqueous solutions of benzidine and diaminobenzidine
Support Protocol 1	Analysis for benzidine and diaminobenzidine
Basic Protocol 2	Biological stains
Alternate Protocol 3	Large volumes of dilute biological stains
Support Protocol 2	Analysis for biological stains
Basic Protocol 3	Silanes
Basic Protocol 4	Cyanide and cyanogen bromide
Support Protocol 3	Analysis for cyanide
Basic Protocol 5	Dimethyl sulfate, diethyl sulfate, methyl methanesulfonate, ethyl methanesulfonate, diepoxybutane, 1,3-propane sultone
Support Protocol 4	Analysis for dimethyl sulfate, diethyl sulfate, methyl methanesulfonate, ethyl methanesulfonate, diepoxybutane, 1,3-propane sultone
Basic Protocol 6	Ethidium bromide and propidium iodide
Alternate Protocol 4	Equipment contaminated with ethidium bromide
Alternate Protocol 5	Ethidium bromide in isopropanol containing cesium chloride
Alternate Protocol 6	Ethidium bromide in alcohols
Support Protocol 5	Analysis for ethidium bromide and propidium iodide
Basic Protocol 7	Hydrogen peroxide
Basic Protocol 8	Iodine
Basic Protocol 9	Mercury compounds
Alternate Protocol 7	Waste water containing mercury
Support Protocol 6	Analysis for mercury
Basic Protocol 10	Sodium azide
Support Protocol 7	Analysis for sodium azide
Support Protocol 8	Analysis for nitrite
Basic Protocol 11	Enzyme inhibitors
Support Protocol 9	Analysis for enzyme inhibitors

Materials

Benzidine or diaminobenzidine tetrahydrochloride dihydrate
 0.1 M HCl (for benzidine)
 0.2 M potassium permanganate: prepare immediately before use
 2 M sulfuric acid
 Sodium metabisulfite
 10 M potassium hydroxide (KOH)

Additional reagents and equipment for testing for the presence of aromatic amines
 (see Support Protocol 1)

1. For each 9 mg benzidine, add 10 ml of 0.1 M HCl *or* for each 9 mg diaminobenzidine tetrahydrochloride dihydrate, add 10 ml water. Stir the solution until the aromatic amine has completely dissolved.

2. For each 10 ml of solution, add 5 ml freshly prepared 0.2 M potassium permanganate and 5 ml of 2 M sulfuric acid. Allow the mixture to stand for ≥ 10 hr.
3. Add sodium metabisulfite until the solution is decolorized.
4. Add 10 M KOH to make the solution strongly basic, $\text{pH} > 12$.
CAUTION: *This reaction is exothermic.*
5. Dilute with 5 vol water and pass through filter paper to remove manganese compounds.
6. Test the filtrate for the presence of aromatic amines (i.e., benzidine or diaminobenzidine; see Support Protocol 1).
7. Neutralize the filtrate with acid and discard.

**ALTERNATE
PROTOCOL 1**

**DECONTAMINATION OF SPILLS INVOLVING BENZIDINE AND
DIAMINO BENZIDINE**

Additional Materials (also see Basic Protocol 1)

Glacial acetic acid

1:1 (v/v) 0.2 M potassium permanganate/2 M sulfuric acid: prepare immediately before use

Absorbent material (e.g., paper towels, Kimwipes)

High-efficiency particulate air (HEPA) vacuum (Fisher)

Additional reagents and equipment for testing for the presence of aromatic amines (see Support Protocol 1)

CAUTION: This procedure may damage painted surfaces and Formica.

1. Remove as much of the spill as possible using absorbent material and high-efficiency particulate air (HEPA) vacuuming.
2. Wet the surface with glacial acetic acid until all the amines are dissolved, then add an excess of freshly prepared 1:1 (v/v) 0.2 M potassium permanganate/2 M sulfuric acid to the spill area. Allow the mixture to stand ≥ 10 hr.
3. Ventilate the area and decolorize with sodium metabisulfite.
4. Mop up the liquid with paper towels. Squeeze the solution out of the towels and collect in a suitable container. Discard towels as hazardous solid waste.
5. Add 10 M KOH to make the solution strongly basic, $\text{pH} \geq 12$.
CAUTION: *This reaction is exothermic.*
6. Dilute with 5 vol water and filter through filter paper to remove manganese compounds.
7. Test the filtrate for the presence of aromatic amines (i.e., benzidine or diaminobenzidine; see Support Protocol 1).
8. Neutralize the filtrate with acid and discard it.
9. Verify complete decontamination by wiping the surface with a paper towel moistened with water and squeezing the liquid out of the towel. Test the liquid for the presence of benzidine or diaminobenzidine (see Support Protocol 1). Repeat steps 1 to 9 as necessary.

DECONTAMINATION OF AQUEOUS SOLUTIONS OF BENZIDINE AND DIAMINOBENZIDINE

ALTERNATE
PROTOCOL 2

The enzyme horseradish peroxidase catalyzes the oxidation of the amine to a radical which diffuses into solution and polymerizes. The polymers are insoluble and fall out of solution.

Additional Materials (also see Basic Protocol 1)

Aqueous solution of benzidine or diaminobenzidine
1 N HCl or NaOH
3% (v/v) hydrogen peroxide
Horseradish peroxidase (see recipe)
1:1 (v/v) 0.2 M potassium permanganate/2 M sulfuric acid
5% (w/v) ascorbic acid
Porous glass filter *or* Sorvall GLC-1 centrifuge or equivalent

Additional reagents and equipment for testing for the presence of aromatic amines (see Support Protocol 1)

1. Adjust the pH of the aqueous benzidine or diaminobenzidine solution to 5 to 7 with 1 N HCl or NaOH as required and dilute so the concentration of aromatic amines is ≤ 100 mg/liter.
2. For each liter of solution, add 3 ml of 3% hydrogen peroxide and 300 U horseradish peroxidase. Let the mixture stand 3 hr.
3. Remove the precipitate by filtering the solution through a porous glass filter or by centrifuging 5 min at room temperature in a benchtop centrifuge to pellet the precipitate.

The precipitate is mutagenic and should be treated as hazardous waste.

4. Immerse the porous glass filter in 1:1 (v/v) 0.2 M potassium permanganate/2 M sulfuric acid. Clean the filter in a conventional fashion and discard potassium permanganate/sulfuric acid solution as described for benzidine and diaminobenzidine (see Basic Protocol 1).
5. For each liter of filtrate, add 100 ml of 5% ascorbic acid.
6. Test the filtrate for the presence of aromatic amines (see Support Protocol 1).
7. Discard the decontaminated filtrate.

ANALYTICAL PROCEDURES TO DETECT BENZIDINE AND DIAMINOBENZIDINE

SUPPORT
PROTOCOL 1

Reversed-phase HPLC (Snyder et al., 1997) is used to test for the presence of aromatic amines. The limit of detection is 1 $\mu\text{g/ml}$ for benzidine and 0.25 $\mu\text{g/ml}$ for diaminobenzidine.

Materials

Decontaminated aromatic amine solution
10:30:20 (v/v/v) acetonitrile/methanol/1.5 mM potassium phosphate buffer (1.5 mM K_2HPO_4 /1.5 mM KH_2PO_4) (benzidine) *or* 75:25 (v/v) methanol/1.5 mM potassium phosphate buffer (diaminobenzidine)
250-mm \times 4.6-mm-i.d. Microsorb C-8 reversed-phase HPLC column (Varian) or equivalent

Additional reagents and equipment for reversed-phase liquid chromatography (Snyder et al., 1997)

Standard
Measurements,
Data, and
Abbreviations

A.1H.13

Analyze the decontaminated aromatic amine solution by reversed-phase HPLC using a 250-mm × 4.6-mm-i.d. Microsorb C-8 column or equivalent. To detect benzidine, elute with 10:30:20 (v/v/v) acetonitrile/methanol/1.5 mM potassium phosphate buffer at a flow rate of 1.5 ml/min and UV detection at 285 nm. To detect diaminobenzidine, elute with 75:25 (v/v) methanol/1.5 mM potassium phosphate buffer at a flow rate of 1 ml/min and UV detection at 300 nm.

DISPOSAL OF BIOLOGICAL STAINS

Biological stains (Table A.1H.5), as well as ethidium bromide and propidium iodide, can be removed from solution using the polymeric resin Amberlite XAD-16. The decontaminated solution may be disposed of as nonhazardous aqueous waste and the resin as hazardous solid waste. The volume of contaminated resin generated is much smaller than the original volume of the solution of biological stain, so the waste disposal problem is greatly reduced. The final concentration of any remaining stain should be less than the limit of detection (see Support Protocol 2 and Table A.1H.5). In each case decontamination should be >99%. This protocol describes a method for batch decontamination in which the resin is stirred in the solution to be decontaminated and removed by filtration at the end of the reaction time. Large volumes of biological stain can be decontaminated using a column (see Alternate Protocol 3). For full details refer to the original literature (Lunn and Sansone, 1991b) or a compilation (Lunn and Sansone, 1994a).

Table A.1H.5 Decontamination of Biological Stains

Compound	Time required for complete decontamination	Volume of solution (ml) decontaminated per gram resin
Acridine orange	18 hr	20
Alcian blue 8GX	10 min	500
Alizarin red S	18 hr	5
Azure A	10 min	80
Azure B	10 min	80
Brilliant blue R	2 hr	80
Congo red	2 hr	40
Coomassie brilliant blue G	2 hr	80
Cresyl violet acetate	2 hr	40
Crystal violet	30 min	200
Eosin B	30 min	40
Erythrosin B	18 hr	10
Ethidium bromide	4 hr	20
Janus green B	30 min	80
Methylene blue	30 min	80
Neutral red	10 min	500
Nigrosin	2 hr	80
Orcein	2 hr	200
Propidium iodide	2 hr	20
Rose Bengal	3 hr	20
Safranin O	1 hr	20
Toluidine blue O	30 min	80
Trypan blue	2 hr	40

Materials

Amberlite XAD-16 resin (Supelco)
100 µg/ml biological stain in water

Additional reagents and equipment for testing for the presence of biological stain
(see Support Protocol 2)

For batch decontamination of 20 ml stain

1a. Add 1 g Amberlite XAD-16 to 20 ml of 100 µg/ml biological stain in water.

For aqueous solutions having stain concentrations other than 100 µg/ml, use proportionately greater or lesser amounts of resin to achieve complete decontamination.

For solutions of erythrosin B, use 2 g Amberlite XAD-16 for 20 ml stain.

2a. Stir the mixture for at least the time indicated in Table A.1H.5.

For batch decontamination of larger volumes of stain

1b. Add 1 g Amberlite XAD-16 to the volume of a 100 µg/ml biological stain in water indicated in Table A.1H.5.

2b. Stir the mixture for at least 18 hr.

3. Remove the resin by filtration through filter paper.

4. Test the filtrate for the presence of the biological stain (see Support Protocol 2).

5. Discard the resin as hazardous solid waste and the decontaminated filtrate as liquid waste.

CONTINUOUS-FLOW DECONTAMINATION OF AQUEOUS SOLUTIONS OF BIOLOGICAL STAINS

For treating large volumes of dilute aqueous solutions of biological stains (Table A.1H.5), it is possible to put the resin in a column and run the contaminated solution through the column in a continuous-flow system (Lunn et al., 1994). Limited grinding of the Amberlite XAD-16 resin increases its efficiency.

Additional Materials (also see Basic Protocol 2)

25 µg/ml biological stain in water
Methanol (optional)

300-mm × 11-mm-i.d. glass chromatography column fitted with threaded adapters and flow-regulating valves at top and bottom nut and insert connectors, and insertion tool (Ace Glass) *or* 300-mm × 15-mm-i.d. glass chromatography column (Spectrum 124010, Fisher)

Glass wool

1.5-mm-i.d. × 0.3-mm-wall Teflon tubing

Waring blender (optional)

Rubber stopper fitted over a pencil

QG 20 lab pump (Fluid Metering)

Additional reagents and equipment for testing for the presence of biological stain
(see Support Protocol 2)

Using a slurry of Amberlite XAD-16

1a. Prepare a 300-mm × 11-mm-i.d. glass chromatography column. To prevent clogging of the column outlet, place a small plug of glass wool at the bottom of the chromatography column. Connect 1.5-mm-i.d. × 0.3-mm wall Teflon tubing to the adapters using nut and insert connectors. Attach the tubing using an insertion tool.

ALTERNATE PROTOCOL 3

Standard Measurements, Data, and Abbreviations

A.1H.15

Table A.1H.6 Breakthrough Volumes for Continuous-Flow Decontamination of Biological Stains

Compound	Breakthrough volume (ml)		
	Limit of detection	1 ppm	5 ppm
Acridine orange	465	>990	>990
Alizarin red S	120	150	240
Azure A	615	810	>975
Azure B	630	882	>1209
Cresyl violet acetate	706	>1396	>1396
Crystal violet	1020	>1630	>1630
Ethidium bromide	260	312	416
Janus green B	170	650	>870
Methylene blue	420	645	1050
Neutral red	>2480	>2480	>2480
Safranin O	365	438	584
Toluidine blue O	353	494	606

- 2a. Mix 10 g Amberlite XAD-16 and 25 ml water in a beaker and stir 5 min to wet the resin.

Using a finely ground Amberlite XAD-16 slurry

- 1b. Prepare a 300-mm × 15-mm-i.d. glass chromatography column. To prevent clogging of the column outlet, place a small plug of glass wool at the bottom of the chromatography column.
- 2b. Grind 20 g Amberlite XAD-16 with 200 ml water for exactly 10 sec in a Waring blender.
3. Pour the resin slurry into the column through a funnel. As the resin settles, tap the column with a rubber stopper fitted over a pencil to encourage even packing. Attach a QG 20 lab pump.
4. Pump the 25- μ g/ml biological stain solution through the column at 2 ml/min.

Alternatively, gravity flow coupled with periodic adjustment of the flow-regulating valve can be used.

5. Check the effluent from the column for the presence of biological stain (see Support Protocol 2). Stop the pump when stain is detected.

Table A.1H.6 lists breakthrough volumes at different detection levels for a number of biological stains.

6. Discard the decontaminated effluent and the contaminated resin appropriately.
7. Many biological stains can be washed off the resin with methanol so the resin can be reused. Discard the methanol solution of stain as hazardous organic liquid waste.

Table A.1H.7 Methods for Detecting Biological Stains^a

Compound	Reagent ^b	Procedure	Wavelength(s)(nm)	Limit of detection (ppm)
Acridine orange	DNA solution	F	ex 492, em 528	0.0032
Alcian blue 8GX		A	615	0.9
Alizarin red S	1 M KOH	A	556	0.46
Azure A		A	633	0.15
Azure B		A	648	0.13
Brilliant blue R		A	585	1.0
Congo red		A	497	0.25
Coomassie brilliant blue G		A	610	1.7
Cresyl violet acetate	pH 5 buffer	F	ex 588, em 618	0.021
Crystal violet		A	588	0.1
Eosin B		A	514	0.21
Erythrosin B		F	ex 488, em 556	0.025
Ethidium bromide		F	ex 540, em 590	0.05
Janus green B		A	660	0.6
Methylene blue		A	661	0.13
Neutral red	pH 5 buffer	A	540	0.6
Nigrosin		A	570	0.8
Orcein	1 M KOH	A	579	1.15
Propidium iodide	DNA solution	F	ex 350, em 600	0.1
Rose Bengal		F	ex 520, em 576	0.04
Safranin O		F	ex 460, em 582	0.03
Toluidine blue O		A	626	0.2
Trypan blue		A	607	0.22

^aAbbreviations: A, absorbance; em, emission; ex, excitation; F, fluorescence

^bSee Support Protocol 2

ANALYTICAL PROCEDURES TO DETECT BIOLOGICAL STAIN

Depending on the biological stain, the filtrate or eluate from the decontamination procedure can be analyzed using either UV absorption (A) or fluorescence detection (F).

Materials

Filtrate or eluate from biological stain decontamination (see Basic Protocol 2 or Alternate Protocol 3)

pH 5 buffer (see recipe)

1 M KOH solution

20 µg/ml calf thymus DNA in TBE electrophoresis buffer, pH 8.1 (APPENDIX 2A)

Test the filtrate or eluate from the biological stain decontamination procedure using the appropriate method as indicated in Table A.1H.7.

Traces of acid or base on the resin may induce color changes in the stain. Accordingly, with cresyl violet acetate or neutral red, mix aliquots of the filtrate with 1 vol pH 5 buffer before analyzing. With alizarin red S and orcein, mix aliquots of the filtrate with 1 vol of 1 M KOH solution before analyzing.

Increase the fluorescence of solutions of acridine orange, ethidium bromide, and propidium iodide by mixing an aliquot of the filtrate with an equal volume of 20 µg/ml calf thymus DNA in TBE electrophoresis buffer, pH 8.1. Let the solution stand 15 min before measuring the fluorescence.

SUPPORT PROTOCOL 2

Standard Measurements, Data, and Abbreviations

A.1H.17

**DISPOSAL OF CHLOROTRIMETHYLSILANE AND
DICHLORODIMETHYLSILANE**

Silane-containing compounds are hydrolyzed to hydrochloric acid and polymeric silicon-containing material (Patnode and Wilcock, 1946).

1. Hydrolyze silane-containing compounds by cautiously adding 5 ml silane to 100 ml vigorously stirred water in a flask. Allow the resulting suspension to settle.
2. Remove any insoluble material by filtration and discard it with the solid or liquid hazardous waste.
3. Neutralize the aqueous layer with base and discard it.

DISPOSAL OF CYANIDES AND CYANOGEN BROMIDE

Inorganic cyanides (e.g., NaCN) and cyanogen bromide (CNBr) are oxidized by sodium hypochlorite (NaOCl; e.g., Clorox) in basic solution to the much less toxic cyanate ion (Lunn and Sansone, 1985a). Destruction is >99.7%.

Materials

Cyanide (e.g., NaCN) or cyanogen bromide (CNBr)
1 M NaOH
5.25% (v/v) sodium hypochlorite (NaOCl; i.e., standard household bleach)

Additional reagents and equipment for testing for the presence of cyanide (see Support Protocol 3)

1. Dissolve cyanide in water to give a concentration ≤ 25 mg/ml *or* dissolve CNBr in water to give a concentration ≤ 60 mg/ml.

If necessary, dilute aqueous solutions so the concentration of NaCN or CNBr does not exceed the limit.

2. Mix 1 vol NaCN or CNBr solution with 1 vol 1 M NaOH and 2 vol fresh 5.25% NaOCl. Stir the mixture 3 hr.

IMPORTANT NOTE: *With age bleach may become ineffective; use of fresh bleach is strongly recommended.*

3. Test the reaction mixture for the presence of cyanide (see Support Protocol 3).
4. Neutralize the reaction mixture and discard it.

ANALYTICAL PROCEDURE TO DETECT CYANIDE

This protocol is used to detect cyanide or cyanogen bromide at ≥ 3 $\mu\text{g/ml}$.

Materials

Cyanide or cyanogen bromide decontamination reaction mixture (see Basic Protocol 4)

Phosphate buffer (see recipe)

10 mg/ml sodium ascorbate in water: prepare fresh daily

100 mg/liter NaCN in water: prepare fresh weekly

10 mg/ml chloramine-T in water: prepare fresh daily

Cyanide detection reagent (see recipe)

Sorvall GLC-1 centrifuge or equivalent

1. If necessary to remove suspended solids, centrifuge two 1-ml aliquots of the cyanide or cyanogen bromide decontamination reaction mixture 5 min in a benchtop centrifuge, room temperature. Add each supernatant to 4 ml phosphate buffer in separate tubes.
2. If an orange or yellow color appears, add 10 mg/ml freshly prepared sodium ascorbate dropwise until the mixture is colorless, but do not add more than 2 ml.
3. Add 200 μ l of 100 mg/liter NaCN to one reaction mixture (control solution).
4. Add 1 ml freshly prepared 10 mg/ml chloramine-T to each tube. Shake the tubes and let them stand 1 to 2 min.
5. Add 1 ml cyanide detection reagent, shake, and let stand 5 min.

A blue color indicates the presence of cyanide. If destruction has been complete and the analytical procedure has been carried out correctly, the treated reaction mixture should be colorless and the control solution, which contains NaCN, should be blue.

6. Centrifuge tubes 5 min, room temperature, if necessary to remove suspended solids. Measure the absorbance at 605 nm with appropriate standards and blanks.

DISPOSAL OF DIMETHYL SULFATE, DIETHYL SULFATE, METHYL METHANESULFONATE, ETHYL METHANESULFONATE, DIEPOXYBUTANE, AND 1,3-PROPANE SULFONE

**BASIC
PROTOCOL 5**

Dimethyl sulfate is hydrolyzed by base to methanol and methyl hydrogen sulfate (Lunn and Sansone, 1985b). Subsequent hydrolysis of methyl hydrogen sulfate to methanol and sulfuric acid is slow. Methyl hydrogen sulfate is nonmutagenic and a very poor alkylating agent. The other compounds can be hydrolyzed with base in a similar fashion (Lunn and Sansone, 1990a). Destruction is >99%. A method to verify that decontamination is complete is also provided (see Support Protocol 4).

NOTE: The reaction times given in the protocol should give good results; however, reaction time may be affected by such factors as the size and shape of the flask and the rate of stirring. The presence of two phases indicates that the reaction is not complete, and stirring should be continued until the reaction mixture is homogeneous.

Materials

Dimethyl sulfate, diethyl sulfate, methyl methanesulfonate, ethyl methanesulfonate, diepoxybutane, or 1,3-propane sulfone
5 M NaOH

Additional reagents and equipment for testing for the presence of dimethyl sulfate, diethyl sulfate, methyl methanesulfonate, ethyl methanesulfonate, diepoxybutane, or 1,3-propane sulfone (see Support Protocol 4)

For bulk quantities of dimethyl sulfate

- 1a. Add 100 ml dimethyl sulfate to 1 liter of 5 M NaOH. Stir the reaction mixture.
- 2a. Fifteen minutes after all the dimethyl sulfate has gone into solution, neutralize the reaction mixture with acid.

For bulk quantities of diethyl sulfate

- 1b. Add 100 ml diethyl sulfate to 1 liter of 5 M NaOH. Stir the reaction mixture 24 hr.
- 2b. Neutralize the reaction mixture with acid.

**Standard
Measurements,
Data, and
Abbreviations**

A.1H.19

For bulk quantities of methyl methanesulfonate, ethyl methanesulfonate, diepoxybutane, and 1,3-propane sultone

- 1c. Add 1 ml methyl methanesulfonate, ethyl methanesulfonate, or diepoxybutane, or 1 g of 1,3-propane sultone to 10 ml of 5 M NaOH. Stir the reaction mixture 1 hr for 1,3-propane sultone, 2 hr for methyl methanesulfonate, 22 hr for diepoxybutane, or 24 hr for ethyl methanesulfonate.
- 2c. Neutralize the reaction mixture with acid.
3. Test the reaction mixture for the presence of the original compound (see Support Protocol 4).
4. Discard the decontaminated reaction mix.

**SUPPORT
PROTOCOL 4**

ANALYTICAL PROCEDURE TO DETECT THE PRESENCE OF DIMETHYL SULFATE, DIETHYL SULFATE, METHYL METHANESULFONATE, ETHYL METHANESULFONATE, DIEPOXYBUTANE, AND 1,3-PROPANE SULTONE

This protocol is used to verify decontamination of solutions containing dimethyl sulfate, diethyl sulfate, methyl methanesulfonate, ethyl methanesulfonate, diepoxybutane, or 1,3-propane sultone. The detection limit for this assay is 40 µg/ml for dimethyl sulfate, 108 µg/ml for diethyl sulfate, 84 µg/ml for methyl methanesulfonate, 1.1 µg/ml for ethyl methanesulfonate, 360 µg/ml for diepoxybutane, and 264 µg/ml for 1,3-propane sultone.

Materials

Reaction mixture containing dimethyl sulfate, diethyl sulfate, methyl methanesulfonate, ethyl methanesulfonate, diepoxybutane, or 1,3-propane sultone

98:2 (v/v) 2-methoxyethanol/acetic acid

5% (w/v) 4-(4-nitrobenzyl)pyridine in 2-methoxyethanol

Piperidine

2-Methoxyethanol

1. Dilute an aliquot of the reaction mixture with 4 vol water.
2. Add 100 µl diluted reaction mixture to 1 ml of 98:2 (v/v) 2-methoxyethanol/acetic acid. Swirl to mix.
3. Add 1 ml of 5% (w/v) 4-(4-nitrobenzyl)pyridine in 2-methoxyethanol. Heat 10 min at 100°C, then cool 5 min in ice.
4. Add 0.5 ml piperidine and 2 ml of 2-methoxyethanol.
5. Measure the absorbance of the violet reaction mixture at 560 nm against an appropriate blank.

The absorbance of a decontaminated solution should be 0.000.

**BASIC
PROTOCOL 6**

DISPOSAL OF ETHIDIUM BROMIDE AND PROPIDIUM IODIDE

Ethidium bromide and propidium iodide in water and buffer solutions may be degraded by reaction with sodium nitrite and hypophosphorous acid in aqueous solution (Lunn and Sansone, 1987); destruction is >99.87%. This reaction may also be used to decontaminate equipment contaminated with ethidium bromide (see Alternate Protocol 4; Lunn and Sansone, 1989) and to degrade ethidium bromide in organic solvents (see Alternate Protocol 5 and Alternate Protocol 6; Lunn and Sansone, 1990b). Ethidium bromide and propidium iodide may also be removed from solution by adsorption onto Amberlite XAD-16 resin (see Basic Protocol 2).

**Safe Use of
Hazardous
Chemicals**

A.1H.20

Materials

Ethidium bromide– or propidium iodide–containing solution in water, buffer, or 1 g/ml cesium chloride

5% (v/v) hypophosphorous acid: prepare fresh daily by diluting commercial 50% reagent 1/10

0.5 M sodium nitrite: prepare fresh daily

Sodium bicarbonate

Additional reagents and equipment for testing for the presence of ethidium bromide or propidium iodide (see Support Protocol 5)

1. If necessary, dilute the ethidium bromide– or propidium iodide–containing solution so the concentration of ethidium bromide or propidium iodide is ≤ 0.5 mg/ml.
2. For each 100 ml solution, add 20 ml of 5% hypophosphorous acid solution and 12 ml of 0.5 M sodium nitrite. Stir briefly and let stand 20 hr.
3. Neutralize the reaction mixture by adding sodium bicarbonate until the evolution of gas ceases.
4. Test the reaction mixture for the presence of ethidium bromide or propidium iodide (see Support Protocol 5).
5. Discard the decontaminated reaction mixture.

DECONTAMINATION OF EQUIPMENT CONTAMINATED WITH ETHIDIUM BROMIDE

Glass, stainless steel, Formica, floor tile, and the filters of transilluminators have been successfully decontaminated using this protocol. No change in the optical properties of the transilluminator filter could be detected, even after a number of decontamination cycles.

Materials

Equipment contaminated with ethidium bromide

Decontamination solution (see recipe)

Sodium bicarbonate

Additional reagents and equipment for testing for the presence of ethidium bromide (see Support Protocol 5)

1. Wash the equipment contaminated with ethidium bromide once with a paper towel soaked in decontamination solution.
The pH of the decontamination solution is 1.8. If this would be too corrosive for the surface to be decontaminated, wash with a paper towel soaked in water instead.
2. Wash the surface five times with paper towels soaked in water using a fresh towel each time.
3. Soak all the towels 1 hr in decontamination solution.
4. Neutralize the decontamination solution by adding sodium bicarbonate until the evolution of gas ceases.
5. Test the decontamination solution for the presence of ethidium bromide (see Support Protocol 5).
6. Discard the decontamination solution and the paper towels as nonhazardous liquid and solid wastes.

**ALTERNATE
PROTOCOL 4**

**Standard
Measurements,
Data, and
Abbreviations**

A.1H.21

**DECONTAMINATION OF ETHIDIUM BROMIDE IN ISOPROPANOL
SATURATED WITH CESIUM CHLORIDE**

Materials

Ethidium bromide in isopropanol saturated with cesium chloride
Decontamination solution (see recipe)
Sodium bicarbonate

Additional reagents and equipment for testing for the presence of ethidium bromide (see Support Protocol 5)

1. If necessary, dilute the ethidium bromide in isopropanol saturated with cesium chloride so the concentration of ethidium bromide is ≤ 1 mg/ml.
2. For each volume of ethidium bromide solution, add 4 vol decontamination solution. Stir the reaction mixture 20 hr.
3. Neutralize the reaction mixture by adding sodium bicarbonate until the evolution of gas ceases.
4. Test the reaction mixture for the presence of ethidium bromide (see Support Protocol 5).
5. Discard the decontaminated solution.

**DECONTAMINATION OF ETHIDIUM BROMIDE IN ISOAMYL
ALCOHOL AND 1-BUTANOL**

Materials

Ethidium bromide in isoamyl alcohol or 1-butanol
Decontamination solution (see recipe)
Activated charcoal
Sodium bicarbonate
Separatory funnel

Additional reagents and equipment for testing for the presence of ethidium bromide

1. If necessary, dilute the ethidium bromide in isoamyl alcohol or 1-butanol so the concentration is ≤ 1 mg/ml final.
2. For each volume of ethidium bromide solution, add 4 vol decontamination solution. Stir the two-phase reaction mixture rapidly for 72 hr.
3. For each 100 ml of reaction mixture, add 2 g activated charcoal. Stir another 30 min.
4. Filter the reaction mixture.
5. Neutralize the filtrate by adding sodium bicarbonate until the evolution of gas ceases. Separate the layers using a separatory funnel.

More alcohol may tend to separate from the aqueous layer on standing.

6. Test the alcohol and aqueous layers for the presence of ethidium bromide.
7. Discard the alcohol and aqueous layers appropriately. Discard the activated charcoal as solid waste.

The aqueous layer contains 4.6% 1-butanol or 2.3% isoamyl alcohol.

ANALYTICAL PROCEDURE TO DETECT ETHIDIUM BROMIDE OR PROPIDIUM IODIDE

SUPPORT
PROTOCOL 5

This protocol is used to verify that solutions no longer contain ethidium bromide or propidium iodide. The limits of detection are 0.05 parts per million (ppm) for ethidium bromide and 0.1 ppm for propidium iodide.

Materials

Reaction mixture containing ethidium bromide or propidium iodide
TBE buffer, pH 8.1 (APPENDIX 2A)
20 µg/ml calf thymus DNA in TBE buffer, pH 8.1

1. Mix 100 µl reaction mixture containing ethidium bromide or propidium iodide with 900 µl TBE buffer, pH 8.1.
2. Add 1 ml of 20 µg/ml calf thymus DNA in TBE, pH 8.1. Prepare a blank solution (100 µl water + 900 µl TBE + 1 ml of 20 µg/ml calf thymus DNA) and control solutions containing known quantities of ethidium bromide or propidium iodide. Let the mixtures stand 15 min.
3. To detect ethidium bromide, measure the fluorescence with an excitation wavelength of 540 nm and an emission wavelength of 590 nm. To detect propidium iodide, measure the fluorescence with an excitation wavelength of 350 nm and an emission wavelength of 600 nm.

If a spectrophotofluorometer is not available, fluorescence of ethidium bromide can be qualitatively determined using a hand-held UV lamp on the long-wavelength setting (Lunn and Sansone, 1991c).

DISPOSAL OF HYDROGEN PEROXIDE

BASIC
PROTOCOL 7

Hydrogen peroxide can be reduced with sodium metabisulfite (Lunn and Sansone, 1994b).

Materials

30% hydrogen peroxide
10% (w/v) sodium metabisulfite
10% (w/v) potassium iodide
1 M HCl
1% (w/v) starch indicator solution

1. Add 5 ml of 30% hydrogen peroxide to 100 ml of 10% sodium metabisulfite. Stir the mixture at room temperature until the temperature starts to drop, indicating that the reaction is over.
2. Test for the presence of hydrogen peroxide by adding a few drops of the reaction mixture to an equal volume of 10% potassium iodide. Add a few drops of 1 M HCl to acidify the reaction mixture, then add a drop of 1% starch indicator solution.

A deep blue color indicates the presence of excess oxidant. If necessary, add more 10% sodium metabisulfite until the starch test is negative.

3. Discard the decontaminated mixture.

Standard
Measurements,
Data, and
Abbreviations

A.1H.23

REDUCTION OF IODINE

Iodine is reduced with sodium metabisulfite to iodide (Lunn and Sansone, 1994b).

Materials

Iodine
10% (w/v) sodium metabisulfite
1 M HCl
1% (w/v) starch indicator solution

1. Add 5 g iodine to 100 ml of 10% sodium metabisulfite. Stir the mixture until the iodine has completely dissolved.
2. Acidify a few drops of the reaction mixture by adding a few drops of 1 M HCl. Add 1 drop of 1% starch indicator solution.

A deep blue color indicates the presence of iodine. If reduction is not complete, add more sodium metabisulfite solution.

3. Dispose of the decontaminated solution.

DISPOSAL OF MERCURY COMPOUNDS

Solutions of mercuric acetate can be decontaminated using Dowex 50X8-100, a strongly acidic gel-type ion-exchange resin with a sulfonic acid functionality. Solutions of mercuric chloride can be decontaminated using Amberlite IRA-400(Cl), a strongly basic gel-type ion-exchange resin with a quaternary ammonium functionality. The final concentration of mercury is <3.8 ppm (Lunn and Sansone, 1994a). On a small scale it is most convenient to stir the resin in the solution to be decontaminated, but on a larger scale, or for routine use, it may be more convenient to pass the solution through a column packed with the resin. Although the volume of waste that must be disposed of is greatly reduced using this technique, a small amount of waste (i.e., the resin contaminated with mercury) remains and must be discarded appropriately. Resin can be regenerated by washing with acid, but the concentrated metal-containing solution generated by this must also be disposed of appropriately. Mercury may also be removed from laboratory waste water using a column of iron powder (see Alternate Protocol 7). Support Protocol 6 is used to detect the presence of mercury.

Materials

Solution containing ≤ 1600 ppm mercuric acetate *or* ≤ 1350 ppm mercuric chloride
Dowex 50X8-100 ion-exchange resin *or* Amberlite IRA-400(Cl) ion-exchange resin

Additional reagents and equipment to test for the presence of mercury (see Support Protocol 6)

- 1a. *For mercuric acetate:* For each 200 ml of solution containing ≤ 1600 ppm mercuric acetate, add 1 g Dowex 50X8-100 ion-exchange resin. Stir the mixture 1 hr, then filter through filter paper.
- 1b. *For mercuric chloride:* For each 200 ml of solution containing ≤ 1350 ppm mercuric chloride, add 1 g Amberlite IRA-400(Cl) ion-exchange resin. Stir the mixture 6 hr, then filter through filter paper.

The speed and efficiency of decontamination will depend on factors such as the size and shape of the flask and the rate of stirring.

3. Test the filtrate for the presence of mercury (see Support Protocol 6).
4. Discard the decontaminated filtrate and the mercury-containing resin appropriately.

DECONTAMINATION OF WASTE WATER CONTAINING MERCURY

Laboratory waste water that contains mercury is decontaminated by passing it through a column of iron powder. The mercury forms mercury amalgam and stays on the column. Some metallic mercury remains in solution but this can be removed by aeration. The final concentration of mercury is <5 ppb (Shirakashi et al., 1986).

Materials

Iron powder, 60 mesh
Waste water containing ≤ 2.5 ppm mercury
6-mm-i.d. column

1. Pack a 6-mm-i.d. column with 1 g of 60-mesh iron powder.

Use a fresh column for each treatment.

2. Pass ≤ 2 liters of water containing ≤ 2.5 ppm of mercury through the column at a flow rate of 250 ml/hr.

Solutions containing a higher concentration of mercury may also be treated, but this will result in a higher final concentration of mercury (e.g., treating a 100-ppm solution in this fashion yielded 33 ppb final).

Some iron ends up in solution and can be removed by adjusting the pH to 8. The resulting precipitated $\text{Fe}(\text{OH})_3$ can then be removed by filtration.

3. Aerate the resulting effluent to remove traces of metallic mercury and continue aeration 30 min after the last of the effluent has emerged from the column. Vent the metallic mercury removed from the solution by aeration into the fume hood or capture it in a mercury trap.

The effluent contains <5 ppb mercury. The presence of iodide or polypeptone may necessitate several treatments to reduce the mercury to an acceptable level.

ANALYTICAL PROCEDURE TO DETECT MERCURY

Atomic absorption spectroscopy with detection at 253.7 nm, a slit width of 0.7 nm, and a limit of detection of 3.8 ppm can be used to determine the concentration of mercury in solution for experiments involving ion-exchange resins. A Hiranuma mercury meter model HG-1 can be used for experiments involving iron powder.

DISPOSAL OF SODIUM AZIDE

Sodium azide can be oxidized by ceric ammonium nitrate (Manufacturing Chemists Association, 1973) to nitrogen (Mason, 1967) or by nitrous acid (National Research Council, 1983) to nitrous oxide (Mason, 1967); destruction is >99.996%. Sodium azide in buffer solution may also be degraded by the addition of sodium nitrite (Lunn and Sansone, 1994a). The reaction proceeds much more readily at low pH, but if sufficient sodium nitrite is added, it will proceed to completion even at high pH. At low pH, it may be possible to completely degrade the azide present in the buffer with less than the amount of sodium nitrite indicated. However, the reaction mixture must be carefully checked to ensure that no azide remains (see Support Protocol 7). At high pH it is possible for unreacted azide to remain in the presence of excess nitrite. Residual nitrite can be detected using Support Protocol 8.

CAUTION: Some toxic nitrogen dioxide may be produced as a by-product of these reactions, so they should always be carried out in a chemical fume hood.

ALTERNATE
PROTOCOL 7

SUPPORT
PROTOCOL 6

BASIC
PROTOCOL 10

Standard
Measurements,
Data, and
Abbreviations

A.1H.25

Materials

Sodium azide or solution containing sodium azide
Ceric ammonium nitrate
10% (w/v) potassium iodide
1 M HCl
1% (w/v) starch indicator solution
Sodium nitrite
4 M sulfuric acid

Additional reagents and equipment to test for the presence of sodium azide (see Support Protocol 7) *or* nitrite (see Support Protocol 8)

Decontamination using ceric ammonium nitrate

- 1a. For each gram of sodium azide, add 9 g ceric ammonium nitrate to 30 ml of water, and stir until it has dissolved.
- 2a. Dissolve each gram of sodium azide in 5 ml water. Add this solution to the ceric ammonium nitrate solution at the rate of 1 ml each min. Stir 1 hr.
If the reaction is carried out on a larger scale, an ice bath may be required for cooling.
- 3a. Check that the reaction is still oxidizing by adding a few drops of the reaction mixture to an equal volume of 10% potassium iodide. Acidify the mixture with 1 drop 1 M HCl and add 1 drop 1% starch indicator solution.

The deep blue color of the starch-iodine complex indicates that excess oxidant is present. If excess oxidant is not present, add more ceric ammonium nitrate.

- 4a. Test for the presence of sodium azide (see Support Protocol 8).
- 5a. Discard the decontaminated reaction mixture.

Decontamination using sodium nitrite

- 1b. For each 5 g sodium azide, dissolve 7.5 g sodium nitrite in 30 ml water.
- 2b. Dissolve each 5 g sodium azide in 100 ml water. Add the sodium nitrite solution with stirring. Slowly add 4 M sulfuric acid until the reaction mixture is acidic to litmus. Stir 1 hr.
CAUTION: It is important to add the sodium nitrite, then the sulfuric acid. Adding these reagents in reverse order will generate explosive, volatile, toxic hydrazoic acid.
If the reaction is carried out on a large scale, an ice bath may be required for cooling.
- 3b. Check that there is excess nitrous acid in the reaction. Add a few drops of the reaction mixture to an equal volume of 10% potassium iodide. Acidify the mixture with 1 drop 1 M HCl. Add 1 drop starch indicator solution.

The deep blue color of the starch-iodine complex indicates that excess nitrous acid is present. If excess nitrous acid is not present, add more sodium nitrite.

- 4b. If excess nitrous acid is present, test for the presence of sodium azide (see Support Protocol 7).
- 5b. Discard the decontaminated reaction mixture.

Decontamination of sodium azide in buffer

- 1c. If necessary, dilute the buffer solution with water so the concentration of sodium azide is ≤ 1 mg/ml.
- 2c. For each 50 ml buffer solution add 5 g sodium nitrite. Stir the reaction 18 hr.

3c. Test for the presence of sodium azide (see Support Protocol 7).

4c. Discard the decontaminated reaction solution.

ANALYTICAL PROCEDURES TO DETECT SODIUM AZIDE

Sodium azide is analyzed by reacting azide ion with 3,5-dinitrobenzoyl chloride to form 3,5-dinitrobenzoyl azide, which can be detected by reversed-phase HPLC. The limit of detection of this assay is 0.2 µg/ml sodium azide. This protocol works only in the absence of nitrite; verify that all of the nitrite has been destroyed by sulfamic acid by using the method detailed later in this unit (see Support Protocol 8).

Materials

Reaction mixture from sodium azide treated with ceric ammonium nitrate or sodium nitrite

1 M KOH

Acetonitrile

Sodium azide indicator solution (see recipe)

0.2 M HCl

20% (w/v) sulfamic acid

3,5-dinitrobenzoyl chloride

50:50 (v/v) acetonitrile/water

Sorvall GLC-1 centrifuge or equivalent

25-cm × 4.6-mm-i.d. Microsorb C-8 reversed-phase HPLC column (Varian) or equivalent

Additional reagents and equipment for reversed-phase liquid chromatography (Snyder et al., 1997)

To analyze for azide in the presence of ceric salts

1a. To a 10-ml aliquot of the reaction mixture from sodium azide treated with ceric ammonium nitrate add 40 ml water. Add 5 ml of this diluted solution to 3 ml of 1 M KOH and mix by shaking.

If <3 ml of 1 M KOH is used, precipitation of ceric salts will not be complete.

2a. Centrifuge the mixture 5 min, room temperature.

3a. Remove 2 ml supernatant and add to 1 ml acetonitrile. Add 1 drop sodium azide indicator solution, add 0.2 M HCl dropwise until the mixture turns yellow, then add 1 drop more.

To analyze for azide in the presence of nitrite

1b. To 5 ml of the reaction mixture from sodium azide treated with sodium nitrite add ≥1 ml sulfamic acid to remove excess nitrite. Let stand ≥3 min.

More sulfamic acid solution may be required for strongly basic reaction mixtures or those containing high concentrations of nitrite. Complete removal of nitrite can be checked by using a modified Griess reagent (see Support Protocol 8).

At high pH the reaction between azide and nitrite is quite slow, so the presence of excess nitrite does not mean that all the azide has been degraded.

2b. Add 1 drop sodium azide indicator solution, then basify the mixture by adding 1 M KOH until it turns purple (typically, 3 to 10 ml are required).

3b. Add 2 ml acetonitrile. Add 0.2 M HCl dropwise until the mixture turns yellow, then add 1 drop more.

If >1 ml sulfamic acid is used, add 4 ml acetonitrile.

4. Prepare a 10% (w/v) solution of 3,5-dinitrobenzoyl chloride in acetonitrile.
5. Add 50 μ l of 10% dinitrobenzoyl chloride/acetonitrile to the reaction mix (step 3a or 3b). Shake the mixture and let it stand \geq 3 min.

Longer standing times have no effect on the HPLC analysis. However, it is crucial to use freshly prepared 3,5-dinitrobenzoyl chloride solution within minutes of its preparation. It is generally most convenient to prepare all the analytical samples with the fresh solution at the beginning of the day and analyze them over the course of the day.

6. Analyze 20 μ l of each reaction mixture by reversed-phase HPLC (Snyder et al., 1997) using a mobile phase of 50:50 (v/v) acetonitrile/water with a flow rate of 1 ml/min and UV detection at 254 nm.

The peak for 3,5-dinitrobenzoyl azide elutes at ~9 min.

**SUPPORT
PROTOCOL 8**

ANALYTICAL PROCEDURE TO DETECT NITRITE

This protocol uses a modified Griess reagent to test for the presence of nitrite. The limit of detection of this assay is 0.06 μ g/ml nitrite. A similar procedure uses *N*-(1-naphthyl)-ethylenediamine (Cunniff, 1995).

Materials

- α -Naphthylamine
- 15% (v/v) aqueous acetic acid
- Sulfanilic acid solution (see recipe)
- Reaction mixture treated to remove excess nitrite (see Support Protocol 7, step 1b)

Table A.1H.8 Conditions for the Destruction of Enzyme Inhibitors

Compound	Concentration	Solvent	Solution: 1 M NaOH	Time
AEBSF	1 mM	Buffer(pH 5.0-8.0)	1:0.1	1 hr
AEBSF	20 mM	DMSO	1:10	24 hr
AEBSF	20 mM	Isopropanol	1:10	24 hr
APMSF	2.5 mM	Buffer(pH 5.0-8.0)	1:0.1	1 hr
APMSF	25 mM	DMSO	1:5	24 hr
APMSF	25 mM	50:50 isopropanol:pH 3 buffer	1:5	24 hr
APMSF	100 mM	Water	1:5	24 hr
DFP	10 mM	Buffer (pH 6.4-7.2)	1:0.2	18 hr
DFP	200 mM	DMF	1:2	18 hr
DFP	pure	—	1:25	1 hr
DFP	10 mM	Water	1:0.2	18 hr
PMSF	10 mM	Buffer (pH 5.0-8.0)	1:0.1	1 hr
PMSF	100 mM	DMSO	1:5	24 hr
PMSF	100 mM	Isopropanol	1:5	24 hr
TLCK	1 mM	Buffer (pH 5.0-8.0)	1:0.1	18 hr
TLCK	5 mM	DMSO	1:5	18 hr
TLCK	5 mM	Water	1:0.1	18 hr
TPCK	1 mM	Buffer (pH 5.0-8.0)	1:0.1	18 hr
TPCK	1 mM	DMSO	1:0.1	18 hr
TPCK	1 mM	Isopropanol	1:0.1	18 hr

1. Prepare the modified Griess reagent by boiling 0.1 g α -naphthylamine in 20 ml water until it dissolves. While the solution is still hot, pour it into 150 ml of 15% aqueous acetic acid. Add 150 ml sulfanilic acid solution.

This reagent should be stored at room temperature in a brown bottle.

CAUTION: α -Naphthylamine is a carcinogen.

2. Add 3 ml of the reaction mixture treated to remove excess nitrite to 1 ml modified Griess reagent. Let stand 6 min at room temperature.
3. Measure the absorbance at 520 nm against a suitable blank.

DISPOSAL OF ENZYME INHIBITORS

The enzyme inhibitors *p*-amidinophenylmethanesulfonyl fluoride (APMSF), 4-(2-aminoethyl)benzenesulfonyl fluoride (AEBSF), phenylmethylsulfonyl fluoride (PMSF; Lunn and Sansone, 1994c), diisopropyl fluorophosphate (DFP; Lunn and Sansone, 1994d), *N* α -*p*-tosyl-L-lysine chloromethyl ketone (TLCK), and *N*-*p*-tosyl-L-phenylalanine chloromethyl ketone (TPCK; Lunn and Sansone, 1994c) may be degraded by reaction with 1 M NaOH. Destruction is >99.8% (except TPCK >98.3%). The exact reaction conditions depend on the solvent (see Table A.1H.8). The solutions that were decontaminated are representative of those described in the literature.

Materials

Solutions of APMSF, AEBSF, PMSF, DFP, TLCK, or TPCK in buffer, DMSO, isopropanol, or water

1 M NaOH

Glacial acetic acid

Additional reagents and equipment for testing for the presence of the enzyme inhibitors (see Support Protocol 9)

1. If necessary, dilute the solutions with the same solvent so that the concentrations given in Table A.1H.8 are not exceeded.

Bulk quantities of AEBSF, PMSF, and TPCK may be dissolved in isopropanol and bulk quantities of APMSF and TLCK may be dissolved in water at the concentrations shown in Table A.1H.8. Bulk quantities of DFP (a liquid) may be mixed directly with 1 M NaOH, taking care to make sure that all the DFP has mixed thoroughly, in the ratio shown in Table A.1H.8 (e.g., 40 μ l DFP with 1 ml of 1 M NaOH).

2. Add 1 M NaOH so that the ratio of solution to 1 M NaOH is that listed in Table A.1H.8.
3. Shake to ensure complete mixing, check that the solution is strongly basic (pH \geq 12), and allow to stand for the time given in Table A.1H.8.
4. Neutralize the reaction mixture with acetic acid and test for the presence of residual enzyme inhibitor (see Support Protocol 9).
5. Discard the decontaminated reaction mixture.

Table A.1H.9 HPLC Conditions for Enzyme Inhibitors

Compound	Mobile phase	Detector	Retention time	Limit of detection
AEBSF	40:60 (v/v) acetonitrile:0.1% trifluoroacetic acid	UV 225 nm	9.5 min	0.1 µg/ml
APMSF	40:60 (v/v) acetonitrile:0.1% trifluoroacetic acid	UV 232 nm	7.7 min	0.5 µg/ml
PMSF	50:50 (v/v) acetonitrile:water	UV 220 nm	8 min	0.9 µg/ml
TLCK	40:60 (v/v) acetonitrile:0.1% trifluoroacetic acid	UV 228 nm	9.5 min	0.37 µg/ml
TPCK	48:52 (v/v) acetonitrile:10 mM pH 7 phosphate buffer	UV 228 nm	10.5 min	2 µg/ml

**SUPPORT
PROTOCOL 9****ANALYTICAL PROCEDURES TO DETECT ENZYME INHIBITORS**

DFP can be determined using a complex procedure involving the inhibition of chymotrypsin activity. For more information, refer to Lunn and Sansone (1994d). A gas chromatographic method has also been described by Degenhardt-Langelan and Kientz (1996). AEBSF, APMSF, PMSF, TLCK, and TPCK may be determined by reversed-phase HPLC (Snyder et al., 1997). The chromatographic conditions and limits of detection are shown in Table A.1H.9 (Lunn and Sansone, 1994c).

Materials

- Decontaminated enzyme inhibitor solutions
- Acetonitrile (HPLC grade)
- Water (HPLC grade)
- 0.1% (v/v) trifluoroacetic acid in water
- 10 mM phosphate buffer, pH 7
- 250-mm × 4.6 mm-i.d. Microsorb C-8 reversed-phase HPLC column (Varian) or equivalent
- Additional reagents and equipment for reversed-phase liquid chromatography (Snyder et al., 1997)

Analyze the decontaminated enzyme inhibitor solutions by reversed-phase HPLC using a 250-mm × 4.6-mm-i.d. Microsorb C-8 reversed-phase column, or equivalent, using the conditions shown in Table A.1H.9. In each case, the injection volume was 20 µl, the separation occurred at ambient temperature, and the flow rate was 1 ml/min. Check the analytical procedures by spiking an aliquot of the acidified reaction mixture with a small quantity of a dilute solution of the compound of interest.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Cyanide detection reagent

Stir 3.0 g barbituric acid in 10 ml water. Add 15 ml of 4-methylpyridine and 3 ml concentrated HCl while continuing to stir. Cool and dilute to 50 ml with water. Store at room temperature.

CAUTION: *This reaction is exothermic.*

Decontamination solution

Dissolve 4.2 g sodium nitrite (0.2 M final) and 20 ml hypophosphorous acid (3.3% w/v final) in 300 ml water. Prepare fresh.

Horseradish peroxidase

Dissolve hydrogen-peroxide oxidoreductase (EC 1.11.1.7 [Type II]; specific activity 150 to 200 purpurogallin U/mg, Sigma) in 1 g/liter sodium acetate to give 30 U/ml. Prepare fresh daily.

For small-scale reactions, a more dilute solution can be used to avoid working with inconveniently small volumes.

pH 5 buffer

2.04 g potassium hydrogen phthalate (0.05 M final)

38 ml 0.1 M potassium hydroxide (15 mM)

H₂O to 200 ml

Store at room temperature

Phosphate buffer

13.6 g monobasic potassium phosphate (KH₂PO₄; 0.1 M final)

0.28 g dibasic sodium phosphate (Na₂HPO₄; 2 mM final)

3.0 g potassium bromide (KBr; 25 mM final)

1 liter H₂O

Store at room temperature

Potassium bromide is necessary to make the assay for cyanide work correctly.

Sodium azide indicator solution

0.1 g bromocresol purple (0.4% final)

18.5 ml 0.01 M potassium hydroxide (KOH; 7.4 mM final)

H₂O to 25 ml

Store at room temperature

Sulfanilic acid solution

Dissolve 0.5 g sulfanilic acid in 150 ml of 15% (v/v) aqueous acetic acid. Use immediately.

LITERATURE CITED

- Bretherick, L. (ed.) 1986. Hazards in the Chemical Laboratory, 4th ed. Royal Society of Chemistry, London.
- Bretherick, L., Urben, P.G., and Pitt, M.J. 1999. Bretherick's Handbook of Reactive Chemical Hazards, 6th ed. Butterworth-Heinemann, London.
- Castegnaro, M., Barek, J., Dennis, J., Ellen, G., Klibanov, M., Lafontaine, M., Mitchum, R., van Roosmalen, P., Sansone, E.B., Stenson, L.A., and Vahl, M. (eds.) 1985. Laboratory Decontamination and Destruction of Carcinogens in Laboratory Wastes: Some Aromatic Amines and 4-Nitrobiphenyl. IARC Scientific Publications No. 64. International Agency for Research on Cancer, Lyon, France.
- Cunniff, P. (ed.) 1995. Official Methods of Analysis of the Association of Official Analytical Chemists, 16th ed., Ch. 4, p. 14. Association of Official Analytical Chemists, Arlington, Va.

- Degenhardt-Langelaan, C.E.A.M. and Kientz, C.E. 1996. Capillary gas chromatographic analysis of nerve agents using large volume injections. *J. Chromatogr. A*.723:210-214.
- Forsberg, K. and Keith, L.H. 1999. Chemical Protective Clothing Performance Index Book, 2nd ed. John Wiley & Sons, New York.
- Furr, A.K. (ed.) 2000. CRC Handbook of Laboratory Safety, 5th ed. CRC Press, Boca Raton, Fla.
- Lewis, R.J. Sr. 1999. Sax's Dangerous Properties of Industrial Materials, 10th ed. John Wiley & Sons, New York.
- Lunn, G. and Sansone, E.B. 1985a. Destruction of cyanogen bromide and inorganic cyanides. *Anal. Biochem.* 147:245-250.
- Lunn, G. and Sansone, E.B. 1985b. Validation of techniques for the destruction of dimethyl sulfate. *Am. Ind. Hyg. Assoc. J.* 46:111-114.
- Lunn, G. and Sansone, E.B. 1987. Ethidium bromide: Destruction and decontamination of solutions. *Anal. Biochem.* 162:453-458.
- Lunn, G. and Sansone, E.B. 1989. Decontamination of ethidium bromide spills. *Appl. Ind. Hyg.* 4:234-237.
- Lunn, G. and Sansone, E.B. 1990a. Validated methods for degrading hazardous chemicals: Some alkylating agents and other compounds. *J. Chem. Educ.* 67:A249-A251.
- Lunn, G. and Sansone, E.B. 1990b. Degradation of ethidium bromide in alcohols. *BioTechniques* 8:372-373.
- Lunn, G. and Sansone, E.B. 1991a. The safe disposal of diaminobenzidine. *Appl. Occup. Environ. Hyg.* 6:49-53.
- Lunn, G. and Sansone, E.B. 1991b. Decontamination of aqueous solutions of biological stains. *Biotech. Histochem.* 66:307-315.
- Lunn, G. and Sansone, E.B. 1991c. Decontamination of ethidium bromide spills-author's response. *Appl. Occup. Environ. Hyg.* 6:644-645.
- Lunn, G. and Sansone, E.B. 1994a. Destruction of Hazardous Chemicals in the Laboratory, 2nd ed. John Wiley & Sons, New York.
- Lunn, G. and Sansone, E.B. 1994b. Safe disposal of highly reactive chemicals. *J. Chem. Educ.* 71:972-976.
- Lunn, G. and Sansone, E.B. 1994c. Degradation and disposal of some enzyme inhibitors. Scientific note. *Appl. Biochem. Biotechnol.* 48:57-59.
- Lunn, G. and Sansone, E.B. 1994d. Safe disposal of diisopropyl fluorophosphate (DFP). *Appl. Biochem. Biotechnol.* 49:165-171.
- Lunn, G., Klausmeyer, P.K., and Sansone, E.B. 1994. Removal of biological stains from aqueous solution using a flow-through decontamination procedure. *Biotech. Histochem.* 69:45-54.
- Manufacturing Chemists Association. 1973. Laboratory Waste Disposal Manual. p. 136. Manufacturing Chemists Association, Washington, D.C.
- Mason, K.G. 1967. Hydrogen azide. *In* Mellor's Comprehensive Treatise on Inorganic and Theoretical Chemistry, Vol. VIII (Suppl. II) pp. 1-15. John Wiley & Sons, New York.
- National Research Council. 1983. Prudent Practices for Disposal of Chemicals from Laboratories, p. 88. National Academy Press, Washington, D.C.
- O'Neil, M.J. (ed.) 2001. The Merck Index, 13th ed. Merck & Co., Whitehouse Station, N.J.
- Patnode, W. and Wilcock, D.F. 1946. Methylpolysiloxanes. *J. Am. Chem. Soc.* 68:358-363.
- Shirakashi, T., Nakayama, K., Kakii, K., and Kuriyama, M. 1986. Removal of mercury from laboratory waste water with iron powder. *Chem. Abstr.* 105:213690y.
- Snyder, L.R., Kirkland, J.J., and Glajch, J.L. 1997. Practical HPLC Method Development, 2nd ed. John Wiley & Sons, New York.

KEY REFERENCES

The following are good general references for laboratory safety.

- American Chemical Society, Committee on Chemical Safety. 1995. Safety in Academic Chemistry Laboratories, 6th ed. American Chemical Society, Washington, D.C.
- Castegnaro, M. and Sansone, E.B. 1986. Chemical Carcinogens. Springer-Verlag, New York.
- DiBerardinis, L.J., First, M.W., Gatwood, G.T., and Seth, A.K. 2001. Guidelines for Laboratory Design, Health and Safety Considerations, 3rd ed. John Wiley & Sons, New York.
- Fleming, D.D., Richardson, J.H., Tulis, J.J., and Vesley, D. 1995. Laboratory Safety, Principles and Practices, 2nd ed. American Society for Microbiology, Washington, D.C.
- Freeman, N.T. and Whitehead, J. 1982. Introduction to Safety in the Chemical Laboratory. Academic Press, San Diego.

- Fuscaldò, A.A., Erlick, B.J., and Hindman, B. (eds.) 1980. *Laboratory Safety, Theory and Practice*. Academic Press, San Diego.
- Lees, R. and Smith, A.F. (eds.) 1984. *Design, Construction, and Refurbishment of Laboratories*. Ellis Horwood, Chichester, United Kingdom.
- Montesano, R., Bartsch, H., Boyland, E., Della Porta, G., Fishbein, L., Griesemer, R.A., Swan, A.B., and Tomatis, L. (eds.) 1979. *Handling Chemical Carcinogens in the Laboratory, Problems of Safety*. IARC Scientific Publications No. 33. International Agency for Research on Cancer, Lyon, France.
- National Research Council. 1995. *Prudent Practices in the Laboratory: Handling and Disposal of Chemicals*. National Academy Press, Washington, D.C.
- Occupational Health and Safety. 1993. National Safety Council, Chicago.
- Pal, S.B. (ed.) 1991. *Handbook of Laboratory Health and Safety Measures*, 2nd ed. Kluwer Academic Publishers, Hingham, Mass.
- Rosenlund, S.J. 1987. *The Chemical Laboratory: Its Design and Operation: A Practical Guide for Planners of Industrial, Medical, or Educational Facilities*. Noyes Publishers, Park Ridge, N.J.
- Young, J.A. (ed.) 1991. *Improving Safety in the Chemical Laboratory: A Practical Guide*, 2nd ed. John Wiley & Sons, New York.

INTERNET RESOURCES

<http://www.ilpi.com/msds/index.html>

Where to find MSDSs on the internet. Contains links to general sites, government and nonprofit sites, chemical manufacturers and suppliers, pesticides, and miscellaneous sites.

<http://www.OSHA.gov>

OSHA web site.

http://www.osha-slc.gov/OshStd_data/1910_1450.html

Text of OSHA Standard 29 CFR 1910.1450: Occupational Exposure to Hazardous Chemicals in Laboratories.

http://www.osha-slc.gov/OshStd_data/1910_1000_TABLE_Z-1.html

Tables of permissible exposure limits (PELs) for air contaminants.

http://www.osha-slc.gov/OshStd_data/1910_1000_TABLE_Z-2.html

Tables of PELs for toxic and hazardous substances.

<http://hazard.com/msds/index.php>

Main site for Vermont SIRI. One of the best general sites to start a search. Browse manufacturers alphabetically (for sheets not in the SIRI collection) or do a keyword search in the SIRI MSDS database. Lots of additional safety links and information.

<http://siri.uvm.edu/msds>

Alternate site for Vermont SIRI.

http://tis.eh.doe.gov/docs/osh_tr/ch5.html

DOE OSH technical reference chapter on personal protective equipment.

Contributed by George Lunn
Baltimore, Maryland

Gretchen Lawler
Purdue University
West Lafayette, Indiana

Commonly Used Detergents

Detergents are polar lipids that are soluble in water. The presence of both a hydrophobic and hydrophilic portion makes these compounds very useful for lysis of lipid membranes, solubilization of antigens, and washing of immune complexes.

TYPES OF DETERGENTS

A large variety of detergents are available (Helenius et al., 1979). For biochemical studies, they are usually categorized according to the type of hydrophilic group they contain— anionic, cationic, amphoteric, or nonionic. Tables A.11.1 and A.11.2 list commonly used members of each type. In general, nonionic and amphoteric detergents are less denaturing for proteins than ionic detergents. Sodium cholate and sodium deoxycholate are the least denaturing of the commonly used ionic detergents.

Two properties of detergents are important in their consideration for biological studies: the critical micelle concentration (CMC) and the micelle molecular weight (Table A.11.1). The CMC is the concentration at which monomers of detergent molecules combine to form micelles; each detergent micelle has a characteristic micelle molecular weight. Detergents with a high micelle molecular weight, such as nonionic detergents, are difficult to remove from samples by dialysis. The CMC and the micelle molecular weight will vary depending on the buffer, salt concentration, pH, and temperature. In general, adding salt will lower the CMC and raise the micelle size.

Table A.11.1 Physical Properties of Commonly Used Detergents^{a,b}

Detergent	mp (°C)	Molecular weight (Da)		CMC	
		Monomer	Micelle	% (w/v)	M
<i>Anionic</i>					
SDS	206	288	18,000	0.23	8.0×10^{-3}
Cholate	201	430	4,300	0.60	1.4×10^{-2}
Deoxycholate	175	432	4,200	0.21	5.0×10^{-3}
<i>Cationic</i>					
C ₁₆ TAB	230	365	62,000	0.04	1.0×10^{-3}
<i>Amphoteric</i>					
LysoPC	—	495	92,000	0.0004	7.0×10^{-6}
CHAPS	157	615	6,150	0.49	1.4×10^{-3}
Zwittergent 3-14	—	364	30,000	0.011	3.0×10^{-4}
<i>Nonionic</i>					
Octyl glucoside	105	292	8,000	0.73	2.3×10^{-2}
Digitonin	235	1,229	70,000	—	—
C ₁₂ E ₈	—	542	65,000	0.005	8.7×10^{-5}
Lubrol PX	—	582	64,000	0.006	1.0×10^{-4}
Triton X-100	—	650	90,000	0.021	3.0×10^{-4}
Nonidet P-40	—	603	90,000	0.017	3.0×10^{-4}
Tween 80	—	1,310	76,000	0.002	1.2×10^{-5}

^aReprinted with permission from IRL Press (see Jones et al., 1987).

^bAbbreviations: C₁₆TAB, hexadecyl trimethylammonium bromide; CMC, critical micelle concentration; LysoPC, lysophosphatidylcholine; mp, melting point; SDS, sodium dodecyl sulfate.

Contributed by John E. Coligan

Current Protocols in Molecular Biology (1998) A.11.1-A.11.3

Copyright © 1998 by John Wiley & Sons, Inc.

Table A.11.2 Chemical Properties of Commonly Used Detergents^{a,b}

Property	Ionic detergents							Nonionic detergents						
	SDS	CHO	DOC	C ₁₆	LYS	CHA	ZWI	OGL	DIG	C ₁₂	LUB	TNX	NP-40	T80
Strongly denaturing ^c	+	-	-	+	+/-	-	+/-	-	-	-	-	-	-	-
Dialyzable	+	+	+	+	-	+	+/-	+	-	-	-	-	-	-
Ion exchangeable ^d	+	+	+	+	-	-	-	-	-	-	-	-	-	-
Complexes ions	+	+	+	-	-	-	-	-	-	+/-	+/-	+/-	+/-	+/-
Strong A ₂₈₀	-	-	-	-	-	-	-	-	-	-	-	+	+	-
Assay interference	-	-	-	-	-	-	-	-	-	-	+/-	+/-	+/-	+/-
Cold precipitates	+	-	+	+	-	-	-	-	-	-	-	-	-	-
High cost	-	-	-	-	+	+	+	+	+	+	-	-	-	-
Availability	+	+	+	+	+	+	+/-	+	+	+/-	+	+	+	+
Toxicity	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ease of purification	+	+	+	+	+/-	+	+	-	+	-	-	-	-	-
Radiolabeled	+	+	+	-	+	-	-	+	-	+	+	+	+	+
Defined composition	+	+	+	+	+	+	+	+	-	-	-	-	-	-
Auto-oxidation	-	-	-	-	-	-	-	-	+	+	+	+	+	+

^aAdapted from IRL Press (see Jones et al., 1987).

^bAbbreviations: C₁₂, C₁₂E₈; C₁₆, hexadecyl trimethylammonium bromide; CHA, CHAPS; CHO, cholate; DIG, digitonin; DOC, deoxycholate; LUB, lubrol PX; LYS, lysophosphatidylcholine; NP-40, Nonidet P-40; OGL, octyl glucoside; SDS, sodium dodecyl sulfate; T80, Tween 80; TNX, Triton X-100; ZWI, Zwittergent 3-14.

^cDenaturing refers to disruption of secondary and tertiary protein structure.

^dIonic detergents are unsuitable for ion-exchange chromatography (UNIT 10.10).

CHOICE OF DETERGENTS

Ionic detergents are very good solubilizing agents, but they tend to denature proteins by destroying native three-dimensional structures. This denaturing ability is useful for SDS-PAGE (UNIT 10.2), but is detrimental where native structure is important, as when functional activities must be retained (antibody activity is usually retained in <0.1% SDS). Nonionic and mildly ionic detergents are less denaturing and can often be used to solubilize membrane proteins while retaining protein-protein interactions. The following detergent properties are detrimental in certain procedures:

1. Phenol-containing detergents (e.g., Triton X-100 and NP-40) have a high absorbance at 280 nm and hence interfere with protein monitoring during chromatography (most ionic detergents do not absorb at 280 nm; Brij- and Lubrol-series detergents are nonionic detergents that do not have substantial absorbance at 280 nm). Phenol-containing detergents also induce precipitation in the Folin protein assay (but they can be used with the Bradford protein assay; UNIT 10.1A). Finally, they are readily iodinated and so should not be present during radioiodination.
2. Many detergents have a very high micelle molecular weight (Table A.11.1), which makes their use in gel filtration impossible since protein sizes are insignificant relative to the micelle size. In addition, such detergents cannot be readily removed by dialysis.
3. Sodium cholate and sodium deoxycholate are insoluble below pH 7.5 or above an ionic strength of 0.1%. SDS will often crystallize below 20°C.
4. Ionic detergents interfere with nondenaturing electrophoresis and isoelectric focusing.

Detergents can be removed or exchanged for other detergents by a variety of procedures (Hjelmeland, 1979; Furth et al., 1984; Harlow and Lane, 1988). Ionic and amphoteric detergents can usually be removed by dialysis (*APPENDIX 3C*). Pierce makes Extracti-Gel D for removing a variety of detergents from protein solutions (*Pierce Immunotechnology Catalog and Handbook on Protein Modification*).

LITERATURE CITED

- Furth, A.J., Bolton, H., Potter, J., and Priddle, J.D. 1984. Separating detergents from proteins. *Methods Enzymol.* 104:318-328.
- Harlow, E. and Lane, D. 1988. Detergents. *In Antibodies: A Laboratory Manual*, pp. 687-689. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Helenius, A., McCaslin, D.R., Fries, E., and Tanford, C. 1979. Properties of detergents. *Methods Enzymol.* 56:734-749.
- Hjelmeland, L.M. 1979. Removal of detergents from membrane proteins. *Methods Enzymol.* 182:277-282.
- Jones, O.T., Earnest, J.P. and McNamee, M.G. 1987. Solubilization and reconstitution of membrane proteins. *In Biological Membranes: A Practical Approach* (J. Findlay and W.H. Evans, eds.), pp. 142-143. IRL Press, Oxford.

KEY REFERENCES

- Harlow and Lane, 1988. See above.
Provides properties of commonly used detergents and means of removing them from proteins.
- Hjelmeland, L.M. and Chrambach, A. 1984. Solubilization of functional membrane proteins. *Methods Enzymol.* 182:305-318.
Describes properties of detergents and how to use them to solubilize proteins.
- Johnstone, A. and Thorpe, R. 1982. Isolation and fractionation of lymphocytes. *In Immunochemistry in Practice*, pp. 94-101. Blackwell Scientific, Oxford.
Provides details in use of detergents to solubilize cells and membranes.
- Neugebauer, J.M. 1990. Detergents: An overview. *Methods Enzymol.* 182:239-282.
Provides details on detergent properties and how to choose one for a particular application.

Contributed by John E. Coligan
National Institute of Allergy and Infectious Diseases
Bethesda, Maryland

Common Conversion Factors

Table A.1J.1 lists some of the more common conversion factors for units of measure used throughout Current Protocols manuals, while Table A.1J.2 gives prefixes indicating powers of ten for SI units.

Table A.1J.3 gives conversions between temperatures on the Celsius (Centigrade) and Fahrenheit scales. Celsius temperatures are converted to Fahrenheit temperatures by multiplying the Celsius figure by 9, dividing by 5, and adding 32, or by multiplying the Celsius figure by 1.8 and adding 32. Fahrenheit is converted to Celsius by subtracting 32 from the Fahrenheit figure, multiplying by 5, and dividing by 9. In Table A.1J.3, the center figure represents the temperature one has read on one of the scales; the figure to the left is the conversion of that figure into Celsius if read in Fahrenheit, while that to the right represents the conversion to Fahrenheit if read in Celsius: e.g., the temperature 88 Fahrenheit converts to 31.1°C, while the temperature 88 Celsius converts to 190.4°F.

Table A.1J.1 Unit of Measurement Conversion Chart

To convert:	Into:	Use the multiplier:
amperes per square centimeter (amp/cm ²)	amperes per square inch (amp/in. ²)	6.452
	amperes per square meter (amp/m ²)	10 ⁴
amperes per square inch (amp/in. ²)	amperes per square centimeter (amp/cm ²)	0.1550
	amperes per square meter (amp/m ²)	1.55 × 10 ³
ampere-hours (amp-hr)	coulombs (C)	3.6 × 10 ³
	faradays	3.731 × 10 ⁻²
atmospheres (atm)	bar	1.01325
	millimeters of mercury (mmHg) or torr	760
	tons per square foot (tons/ft ²)	1.058
bar	atmospheres (atm)	0.9869
	dynes per square centimeter (dyn/cm ²)	10 ⁶
	kilograms per square meter (kg/m ²)	1.020 × 10 ⁴
	pounds per square foot (lb/ft ²)	2,089
	pounds per square inch (lb/in. ² or psi)	14.50
British thermal units (Btu)	ergs	1.0550 × 10 ¹⁰
	gram-calories (g-cal)	252.0
	horsepower-hours (hp-hr)	3.931 × 10 ⁴
	joules (J)	1,054.8
	kilogram-calories (kg-cal)	0.2520
	kilogram-meters (kg-m)	107.5
British thermal unit per minute (Btu/min)	kilowatt-hours (kW-hr)	2.928 × 10 ⁻⁴
	foot-pounds per second (ft-lb/sec)	12.96
	horsepower (hp)	2.356 × 10 ⁻²
	watts (W)	17.57
bushels	cubic feet (ft ³)	1.2445
	cubic inches (in. ³)	2,150.4
	cubic meters (m ³)	3.524 × 10 ⁻²
	liters	35.24
	quarts, dry	32.0

continued

Table A.1J.1 Unit of Measurement Conversion Chart, continued

To convert:	Into:	Use the multiplier:
degrees Celsius or Centigrade (°C)	degrees Fahrenheit (°F)	$(^{\circ}\text{C} \times \frac{9}{5}) + 32$
	Kelvin (K)	$(^{\circ}\text{C}) + 273.15$
degree Fahrenheit (°F)	degrees Celsius (°C)	$\frac{5}{9} \times (^{\circ}\text{F} - 32)$
	Kelvin (K)	$[\frac{5}{9} \times (^{\circ}\text{F} - 32)] + 273.15$
centimeters (cm)	feet (ft)	3.281×10^{-2}
	inches (in.)	0.3937
	kilometers (km)	10^{-5}
	meters (m)	10^{-2}
	miles	6.214×10^{-6}
	millimeters (mm)	10.0
	mils	393.7
	yards	1.094×10^{-2}
centimeters per second (cm/sec)	feet per minute (ft/min)	1.1969
	feet per second (ft/sec)	3.281×10^{-2}
	kilometers per hour (km/hr)	3.6×10^{-2}
	meters per minute (m/min)	0.6
	miles per hour (miles/hr)	2.237×10^{-2}
	miles per minute (miles/min)	3.728×10^{-4}
coulombs (C)	faradays	1.036×10^{-5}
coulombs per square centimeter (C/cm ²)	coulombs per square inch (C/in. ²)	64.52
	coulombs per square meter (C/m ²)	10^4
coulombs per square inch (C/in. ²)	coulombs per square centimeter (C/cm ²)	0.1550
	coulombs per square meter (C/m ²)	1.55×10^3
cubic centimeters (cm ³)	cubic feet (ft ³)	3.531×10^{-5}
	cubic inches (in. ³)	6.102×10^{-2}
	cubic meters (m ³)	10^{-6}
	cubic yards	1.308×10^{-6}
	gallons, U.S. liquid	2.642×10^{-4}
	liters	10^{-3}
	pints, U.S. liquid	2.113×10^{-3}
	quarts, U.S. liquid	1.057×10^{-3}
days	hours (hr)	24.0
	minutes (min)	1.44×10^3
	seconds (sec)	8.64×10^4
degrees (of angle; °)	minutes (min)	60.0
	quadrants, of angle	1.111×10^{-2}
	radians (rad)	1.745×10^{-2}
	seconds (sec)	3.6×10^4
drams	grams (g)	1.7718
	grains	27.3437
	ounces, avoirdupois (oz)	6.25×10^{-2}
dynes (dyn)	joules per centimeter (J/cm)	10^{-7}
	joules per meter (J/m) or newtons (N)	10^{-5}
	kilograms (kg)	1.020×10^{-6}
	pounds (lb)	2.248×10^{-6}

*continued***A.1J.2**

Table A.1J.1 Unit of Measurement Conversion Chart, continued

To convert:	Into:	Use the multiplier:
faradays	ampere-hours (amp-hr)	26.80
	coulombs (C)	9.649×10^{-4}
foot-pounds per minute (ft-lb/min)	British thermal units per minute (Btu/min)	1.286×10^{-3}
	foot-pounds per second (ft-lb/sec)	1.667×10^{-2}
	horsepower (hp)	3.030×10^{-5}
	kilogram-calories per minute (kg-cal/min)	3.24×10^{-4}
	kilowatts (kW)	2.260×10^{-5}
grams (g)	decigrams (dg)	10
	dekagrams (Dg)	0.1
	dynes (dyn)	980.7
	grains	15.43
	hectograms (hg)	10^{-2}
	kilograms (kg)	10^{-3}
	micrograms (μ g)	10^6
	milligrams (mg)	10^3
	ounces, avoirdupois (oz)	3.527×10^{-2}
	ounces, troy	3.215×10^{-2}
	pounds (lb)	2.205×10^{-3}
horsepower (hp)	horsepower, metric	1.014
inches (in.)	centimeters (cm)	2.540
	feet (ft)	8.333×10^{-2}
	meters (m)	2.540×10^{-2}
	miles	1.578×10^{-5}
	millimeters (mm)	25.40
	yards	2.778×10^{-2}
inches of mercury (in. Hg)	atmospheres (atm)	3.342×10^{-2}
	kilogram per square centimeter (kg/cm ²)	3.453×10^{-2}
	kilograms per square meter (kg/m ²)	345.3
	pounds per square foot (lb/ft ²)	70.73
	pounds per square inch (lb/in. ² or psi)	0.4912
joules (J)	British thermal units (Btu)	9.480×10^{-4}
	ergs	10^7
	foot-pounds (ft-lb)	0.7376
	kilogram-calories (kg-cal)	2.389×10^{-4}
	kilogram-meters (kg-m)	0.1020
	newton-meter (N-m)	1
	watt-hours (W-hr)	2.778×10^{-4}
Kelvin (K)	degrees Celsius ($^{\circ}$ C)	$K - 273.13$
	degrees Fahrenheit ($^{\circ}$ F)	$[(K - 273.13) \times \frac{9}{5}] + 32$
kilolines	maxwells (Mx)	10^3
kilometers (km)	centimeters (cm)	10^5
	feet (ft)	3,281
	inches (in.)	3.937×10^4
	meters (m)	10^3
	miles	0.6214
	yards	1,094

continued

Table A.1J.1 Unit of Measurement Conversion Chart, continued

To convert:	Into:	Use the multiplier:
kilowatts (kW)	British thermal units per minute (Btu/min)	56.92
	foot-pounds per minute (ft-lb/min)	4.426×10^4
	horsepower (hp)	1.341
	kilogram-calories per minute (kg-cal/min)	14.34
liters	bushels, U.S. dry	2.838×10^{-2}
	cubic centimeters (cm ³)	10^3
	cubic feet (ft ³)	3.531×10^{-2}
	cubic inches (in. ³)	61.02
	cubic meters (m ³)	10^{-3}
	cubic yards	1.308×10^{-3}
	gallons, U.S. liquid	0.2642
	gallons, imperial	0.21997
	kiloliter (kl)	10^{-3}
maxwells (Mx)	pints, U.S. liquid	2.113
	quarts, U.S. liquid	1.057
micrograms (μg)	webers (W)	10^{-8}
microliters (μl)	grams (g)	10^{-6}
milligrams (mg)	liters	10^{-6}
milligrams per liter (mg/liter)	grams (g)	10^{-3}
millihenries (mH)	parts per million (ppm)	1.0
milliliters (ml)	henries (H)	10^{-3}
millimeters (mm)	liters	10^{-3}
	centimeters (cm)	0.1
	feet (ft)	3.281×10^{-3}
	inches (in.)	3.937×10^{-2}
	kilometers (km)	10^{-6}
	meters (m)	10^{-3}
	miles	6.214×10^{-7}
millimeters of mercury (mmHg) or torr	atmospheres (atm)	1.316×10^{-3}
	kilograms per square meter (kg/m ²)	136.0
	pounds per square foot (lb/ft ²)	27.85
	pounds per square inch (lb/in. ² or psi)	0.1934
nepers (Np)	decibels (dB)	8.686
newtons (N)	dynes (dyn)	10^5
	kilograms, force (kg)	0.10197162
	pounds, force (lb)	4.6246×10^{-2}
ohms (Ω)	megaohms (MΩ)	10^6
	microhms (μΩ)	10^{-6}
ounces, avoirdupois	drams	16.0
	grains	437.5
	grams (g)	28.349527
	pounds (lb)	6.25×10^{-2}
	ounces, troy	0.9115
	tons, metric	2.835×10^{-5}

*continued***A.1J.4**

Table A.1J.1 Unit of Measurement Conversion Chart, continued

To convert:	Into:	Use the multiplier:	
ounces, fluid	cubic inches (in ³)	1.805	
	liters	2.957×10^{-2}	
ounces, troy	grains	480.0	
	grams (g)	31.103481	
	ounces, avoirdupois (oz)	1.09714	
	pounds, troy	8.333×10^{-2}	
pascal (P)	newton per square meter (N/m ²)	1	
pounds, force (lb)	newtons (N)	21.6237	
pounds per square foot (lb/ft ²)	atmospheres (atm)	4.725×10^{-4}	
	inches of mercury (in. Hg)	1.414×10^{-2}	
	kilograms per square meter (kg/m ²)	4.882	
	pounds per square inch (lb/in ² or psi)	6.944×10^{-3}	
pounds per square inch (lb/in. ² or psi)	atmospheres (atm)	6.804×10^{-2}	
	inches of mercury (in. Hg)	2.036	
	kilograms per square meter (kg/m ²)	703.1	
	pounds per square foot (lb/ft ²)	144.0	
quadrants, of angle	bar	6.8966×10^{-2}	
	degrees (°)	90.0	
	minutes (min)	5.4×10^3	
	radians (rad)	1.571	
quarts, dry	seconds (sec)	3.24×10^5	
	cubic inches (in. ³)	67.20	
	quarts, liquid	cubic centimeters (cm ³)	946.4
	cubic feet (ft ³)	3.342×10^{-2}	
quarts, liquid	cubic inches (in. ³)	57.75	
	cubic meters (m ³)	9.464×10^{-4}	
	cubic yards	1.238×10^{-3}	
	gallons	0.25	
	liters	0.9463	
	radians (rad)	degrees (°)	57.30
minutes (min)		3,438	
quadrants		0.6366	
seconds (sec)		2.063×10^5	
torr <i>see</i> millimeter of mercury			
watts (W)	British thermal units per hour (Btu/hr)	3.413	
	British thermal units per min (Btu/min)	5.688×10^{-2}	
	ergs per second (ergs/sec)	10^7	
webers (Wb)	maxwells (M)	10^8	
	kilolines	10^5	

Table A.1J.2 Power of Ten Prefixes for SI Units

Prefix	Factor	Abbreviation
atto	10^{-18}	a
femto	10^{-15}	f
pico	10^{-12}	p
nano	10^{-9}	n
micro	10^{-6}	μ
milli	10^{-3}	m
centi	10^{-2}	c
deci	10^{-1}	d
deca	10^1	da
hecto	10^2	h
kilo	10^3	k
myria	10^4	my
mega	10^6	M
giga	10^9	G
tera	10^{12}	T
peta	10^{15}	P
exa	10^{18}	E

Table A.1J.3 Celsius/Fahrenheit Temperature Conversion Chart

Degrees Celsius (°C)	Temperature	Degrees Fahrenheit (°F)
-17.8	0	32.0
-17.2	1	33.8
-16.7	2	35.6
-16.1	3	37.4
-15.6	4	39.2
-15.0	5	41.0
-14.4	6	42.8
-13.9	7	44.6
-13.3	8	46.4
-12.8	9	48.2
-12.2	10	50.0
-11.7	11	51.8
-11.1	12	53.6
-10.6	13	55.4
-10.0	14	57.2
-9.4	15	59.0
-8.9	16	60.8
-8.3	17	62.6
-7.8	18	64.4
-7.2	19	66.2
-6.7	20	68.0
-6.1	21	69.8
-5.6	22	71.6
-5.0	23	73.4
-4.4	24	75.2
-3.9	25	77.0
-3.3	26	78.8
-2.8	27	80.6
-2.2	28	82.4
-1.7	29	84.2
-1.1	30	86.0
-0.6	31	87.8
0.0	32	89.6
0.6	33	91.4
1.1	34	93.2
1.7	35	95.0
2.2	36	96.8
2.8	37	98.6
3.3	38	100.4
3.9	39	102.2
4.4	40	104.0
5.0	41	105.8
5.6	42	107.6
6.1	43	109.4
6.7	44	111.2
7.2	45	113.0
7.8	46	114.8
8.3	47	116.6
8.9	48	118.4
9.4	49	120.2
10.0	50	122.0
10.6	51	123.8

*continued***Useful Data****A.1J.7**

Table A.1J.3 Celsius/Fahrenheit Temperature Conversion Chart,
continued

Degrees Celsius (°C)	Temperature	Degrees Fahrenheit (°F)
11.1	52	125.6
11.7	53	127.4
12.2	54	129.2
12.8	55	131.0
13.3	56	132.8
13.9	57	134.6
14.4	58	136.4
15.0	59	138.2
15.6	60	140.0
16.1	61	141.8
16.7	62	143.6
17.2	63	145.4
17.8	64	147.2
18.3	65	149.0
18.9	66	150.8
19.4	67	152.6
20.0	68	154.4
20.6	69	156.2
21.1	70	158.0
21.7	71	159.8
22.2	72	161.6
22.8	73	163.4
23.3	74	165.2
23.9	75	167.0
24.4	76	168.8
25.0	77	170.6
25.6	78	172.4
26.1	79	174.2
26.7	80	176.0
27.2	81	177.8
27.8	82	179.6
28.3	83	181.4
28.9	84	183.2
29.4	85	185.0
30.0	86	186.8
30.6	87	188.6
31.1	88	190.4
31.7	89	192.2
32.2	90	194.0
32.8	91	195.8
33.3	92	197.6
33.9	93	199.4
34.4	94	201.2
35.0	95	203.0
35.6	96	204.8
36.1	97	206.8
36.7	98	208.4
37.2	99	210.2
37.8	100	212.0

**Common
Conversion
Factors**

A.1J.8

Compendium of Drugs Commonly Used in Molecular Biology Research

The following appendix includes an alphabetical list of drugs commonly used to examine various biological processes. Table A.1K.1 lists the drugs by activity and provides recent references. Indicated under each drug listed is its mode of action, generally including several specific experimental examples; solvent(s) used to solubilize the drug; stock and working concentrations or ranges; storage conditions; and duration of incubation with cells to achieve the desired effects. Except where indicated, the majority of drugs in this list are cell-permeant. However, despite the well characterized selectivity of many of the following drugs in vitro, the corresponding effects upon their intracellular targets may not be precisely determined directly by their extracellular concentrations, since their cell-permeation properties are not known. Therefore, several different concentrations of any particular drug, as well as alternative methods of determining drug selectivity, should be examined.

Several of these drugs are members of large families, such as those targeting protein kinases and phosphatases, as well as those that affect intracellular Ca^{2+} levels. Many of these family members have different selectivities and potencies toward similar targets, and a complete listing is not included here. The reader may consult catalogs from the following companies, which have several of these family members available: Sigma, Alexis Biochemicals (including LC Laboratories), Calbiochem, Biomol, Molecular Probes, Boehringer Mannheim, Oxford Glycosystems, and Avanti Polar Lipids.

Although not specifically indicated, many of the following drugs are hazardous and should be handled with extreme care. Material Safety Data Sheets (MSDSs) are often provided for products that are hazardous or toxic. In some cases, these products are new and have not been tested for toxicity. Thus, care should be taken to ensure the safe handling of all products.

Table A.1K.1 Biological Activity of Drugs Commonly Used In Molecular Biological Research

Drug	References
<i>DNA replication inhibitors</i>	
Aphidicolin	Borner et al., 1995. <i>Cancer Res.</i> 55:2122-2128; Debec et al., 1996. <i>J. Cell Biol.</i> 134:103-115; Jackson et al., 1995. <i>J. Cell Biol.</i> 130:755-769; Urbani et al., 1995. <i>Exp. Cell Res.</i> 219:159-168
Ara-C	Grant et al., 1994. <i>Oncol. Res.</i> 6:87-99; Tomkins et al., 1994. <i>J. Cell Sci.</i> 107:1499-1507
Camptothecin	Carettoni et al., 1994. <i>Biochem. J.</i> 299:623-629; Desai et al., 1997. <i>JBC</i> 272:24159-24164; Li et al., 1994. <i>JBC</i> 269:7051-7054
Etoposide	Kaufmann et al., 1993. <i>Cancer Res.</i> 53:3976-3985; Meyer et al., 1997. <i>J. Cell Biol.</i> 136:775-788; Shao et al., 1997. <i>JBC</i> 272:31321-31325; Terada et al., 1993. <i>J. Med. Chem.</i> 36:1689-1699; Wozniak et al., 1991. <i>J. Clin. Oncol.</i> 9:70-76
Hydroxyurea	Anand et al., 1995. <i>Cancer Lett.</i> 88:101-105; Nishijima et al., 1997. <i>J. Cell Biol.</i> 138:1105-1116; Oliver et al., 1997. <i>JBC</i> 272:10624-10630
L-Mimosine	Gilbert et al., 1995. <i>JBC</i> 270:9597-9606; Lin et al., 1996. <i>JBC</i> 271:2548-2556; Park et al., 1997. <i>J. Neurosci.</i> 17:1256-1270

continued

Contributed by Nelson B. Cole

Current Protocols in Molecular Biology (2002) A1.K.1-A1.K.26

Copyright © 2002 by John Wiley & Sons, Inc.

Table A.1K.1 Biological Activity of Drugs Commonly Used In Molecular Biological Research, continued

Drug	References
Drugs affecting the cytoskeleton	
Colcemid	Barlow et al., 1994. <i>J. Cell Biol.</i> 126:1017-1029; Bonfoco et al., 1995. <i>Exp. Cell Res.</i> 218:189-200; Goswami et al., 1994. <i>Exp. Cell Res.</i> 214:198-208
Colchicine	Bonfoco et al., 1995. <i>Exp. Cell Res.</i> 218:189-200; Lindenboim et al., 1995. <i>J. Neurochem.</i> 64:1054-1063
Cytochalasin B	Benya and Padilla, 1993. <i>Exp. Cell Res.</i> 204:268-277; Takeshita et al., 1998. <i>Cancer Lett.</i> 126:75-81; Tanaka et al., 1994. <i>Exp. Cell Res.</i> 213:242-252
Cytochalasin D	Cooper, 1987. <i>J. Cell Biol.</i> 105:1473-1478; Radhakrishna and Donaldson, 1997. <i>J. Cell Biol.</i> 139:49-61; Sasaki et al., 1995. <i>PNAS</i> 92:2026-2030; Wang et al., 1994. <i>Am. J. Physiol.</i> 267:F592-F598
Latrunculins	Wada et al., 1998. <i>J. Biochem.</i> 123:946-952; Lamaze et al., 1997. <i>JBC</i> 272:20332-20335; Spector et al., 1989. <i>Cell Motil. Cytoskeleton</i> 13:127-144
Nocodazole	Cole et al., 1996. <i>Mol. Biol. Cell.</i> 7:631-650; Liao et al., 1995. <i>J. Cell Sci.</i> 108:3473-3483; Vasquez et al., 1997. <i>Mol. Biol. Cell</i> 8:973-985
Taxol	Derry et al., 1995. <i>Biochemistry</i> 34:2203-2211; Ding et al., 1990. <i>Science</i> 248:370-372; Gallo, 1998. <i>J. Neurobiol.</i> 35:121-140; Jordan et al., 1993. <i>PNAS</i> 90:9552-9556; Mogensen and Tucker, 1990. <i>J. Cell Sci.</i> 97:101-107
Vinblastine	Dhamodharan et al., 1995. <i>Mol. Biol. Cell</i> 6:1215-1229; Lobert et al., 1998. <i>Cell Motil. Cytoskeleton</i> 39:107-121; Panda et al., 1996. <i>JBC</i> 271:29807-29812; Rai and Wolff, 1996. <i>JBC</i> 271:14707-14711; Taki et al., 1998. <i>J. Neurooncol.</i> 36:41-53; Tsukidate et al., 1993. <i>Am. J. Pathol.</i> 143:918-925
Drugs affecting intracellular Ca²⁺	
A23187	Elia et al., 1996. <i>JBC</i> 27:16111-16118; Kao, 1994. <i>Methods Cell Biol.</i> 40:155-181; Kao et al., 1990. <i>J. Cell Biol.</i> 111:183-196
BAPTA	Bissonnette et al., 1994. <i>Am. J. Physiol.</i> 267:G465-G475; Smith et al., 1992. <i>Biochem. J.</i> 288:925-929; Tsien, 1980. <i>Biochemistry</i> 19:2396-2404
Ionomycin	Aagaard-Tillery and Jelinek, 1995. <i>J. Immunol.</i> 155:3297-3307; Rock et al., 1997. <i>JBC</i> 272:33377-33383; Stewart et al., 1998. <i>J. Cell Biol.</i> 140:699-707
Thapsigargin	Kuznetsov et al., 1993. <i>JBC</i> 268:2001-2008; Lodish et al., 1992. <i>JBC</i> 267:12753-12760; Takemura et al., 1989. <i>JBC</i> 264:12266-12271; Won and Orth, 1995. <i>Endocrinology</i> 136:5399-5408; Wong et al., 1993. <i>Biochem. J.</i> 289:71-79
Drugs affecting oligosaccharide biosynthesis/processing	
Castanospermine	Ahmed et al., 1995. <i>Biochem. Biophys. Res. Commun.</i> 208:267-273; Bass et al., 1998. <i>J. Cell Biol.</i> 141:637-646; Gruters et al., 1987. <i>Nature</i> 330:74-77; Martina et al., 1998. <i>JBC</i> 273:3725-3731
Deoxymannojirimycin	Elbein et al., 1984. <i>Arch. Biochem. Biophys.</i> 235:579-588; McDowell et al., 1987. <i>Virology</i> 161:37-44; Slusarewicz and Warren, 1995. <i>Glycobiology</i> 5:154-155; Wojczyk et al., 1998. <i>Glycobiology</i> 8:121-130
Deoxynojirimycin	Labriola et al., 1995. <i>J. Cell Biol.</i> 130:771-779; Martina et al., 1998. <i>JBC</i> 273:3725-3731; Romero et al., 1985. <i>Biochem. J.</i> 226:733-740; Schlesinger et al., 1984. <i>JBC</i> 259:7597-7601
Tunicamycin	Bush et al., 1994. <i>Biochem. J.</i> 303:705-708; Kuznetsov et al., 1997. <i>JBC</i> 272:3057-3063; Lodish and Kong, 1984. <i>J. Cell Biol.</i> 98:1720-1729

continued

Table A.1K.1 Biological Activity of Drugs Commonly Used In Molecular Biological Research, continued

Drug	References
Drugs affecting the pH of intracellular organelles	
Ammonium chloride	Breuer et al., 1995. <i>JBC</i> 270:24209-24215; Davis and Mecham, 1996. <i>JBC</i> 271:3787-3794; Smith et al., 1997. <i>JBC</i> 272:5640-5646
Bafilomycin A ₁	Bowman et al., 1988. <i>PNAS</i> 85:7972-7976; Calvert and Sanders, 1995. <i>JBC</i> 270:7272-7280; Furuchi et al., 1993. <i>JBC</i> 268:27345-27348
CCCP	Arai et al., 1996. <i>Biochem. Biophys. Res. Commun.</i> 227:433-439; Babcock et al., 1997. <i>J. Cell Biol.</i> 136:833-844; Kao, 1994. <i>Methods Cell Biol.</i> 40:155-181; Simpson and Russell, 1996. <i>JBC</i> 271:33493-33501
Chloroquine	Claus et al., 1998. <i>JBC</i> 273:9842-9851; Garcia-Sainz and Mendoza-Mendoza, 1998. <i>Eur. J. Pharmacol.</i> 342:333-338; Passos and Garcia, 1998. <i>Biochem. Biophys. Res. Commun.</i> 245:155-160; Wunsch et al., 1998. <i>J. Cell Biol.</i> 140:335-345
Concanamycin B	Akifusa et al., 1998. <i>Exp. Cell Res.</i> 238:82-89; Nishihara et al., 1995. <i>Biochem. Biophys. Res. Commun.</i> 212:255-262; Yilla et al., 1993. <i>JBC</i> 268:19092-19100
Monensin	Griffiths et al., 1983. <i>J. Cell Biol.</i> 96:835-850; Hardy et al., 1997. <i>JBC</i> 272:6812-6817; Kallen et al., 1993. <i>Biochim. Biophys. Acta</i> 1166:305-308; Shiao and Vance, 1993. <i>JBC</i> 268:26085-26092
Nigericin	Sandvig et al., 1989. <i>Methods Cell Biol.</i> 32:365-382; Scorrano et al., 1997. <i>JBC</i> 272:12295-12299; Vercesi et al., 1993. <i>JBC</i> 268:8564-8568
Drugs that lead to increased intracellular cAMP levels	
8-Bromo-cyclic AMP	Boyer and Thiery, 1993. <i>J. Cell Biol.</i> 120:767-776; Hei et al., 1991. <i>Mol. Pharmacol.</i> 39:233-238; Sandberg et al., 1991. <i>Biochem. J.</i> 279:521-527
Cholera toxin	Hansen and Casanova, 1994. <i>J. Cell Biol.</i> 126:677-687; Lencer et al., 1995. <i>J. Cell Biol.</i> 131:951-962; Ma and Weiss, 1995. <i>Methods Cell Biol.</i> 49:471-485; Moss and Vaughan, 1992. <i>Curr. Top. Cell. Regul.</i> 32:49-72
Dibutyryl cyclic AMP	Cong et al., 1998. <i>JBC</i> 273:660-666; O'Malley et al., 1997. <i>J. Cell Biol.</i> 138:159-165; Sandvig et al., 1994. <i>J. Cell Biol.</i> 126:53-64; Yuan et al., 1996. <i>JBC</i> 271:27090-27098
Forskolin	Galli et al., 1995. <i>J. Neurosci.</i> 15:1172-1179; Lippincott-Schwartz et al., 1991. <i>J. Cell Biol.</i> 112:567-577; Laurenza et al., 1989. <i>Trends Pharmacol. Sci.</i> 10:442-447; Nickel et al., 1996. <i>JBC</i> 271:15870-15873
Kinase inhibitors	
Bisindolylmaleimide I (GF 109203 X)	Das and White, 1997. <i>JBC</i> 272:14914-14920; Toullec et al., 1991. <i>JBC</i> 266:15771-15781; Uberall et al., 1997. <i>JBC</i> 272:4072-4078
Calphostin C	Dubyak and Kertesy, 1997. <i>Arch. Biochem. Biophys.</i> 341:129-139; Hartzell and Rinderknecht, 1996. <i>Am. J Physiol.</i> 270:C1293-C1299; Jarvis et al., 1994. <i>Cancer Res.</i> 54:1707-1714
Chelerythrine chloride	Barg et al., 1992. <i>J. Neurochem.</i> 59:1145-1152; Herbert et al., 1990. <i>Biochem. Biophys. Res. Commun.</i> 172:993-999; Jarvis et al., 1994. <i>Cancer Res.</i> 54:1707-1714; Jia et al., 1997. <i>JBC</i> 272:4978-4984; Kandasamy et al., 1995. <i>JBC</i> 270:29209-29216
Genistein	Akiyama et al., 1987. <i>JBC</i> 262:5592-5595; Chen et al., 1997. <i>JBC</i> 272:27401-27410; Kranenburg et al., 1997. <i>J. Cell Sci.</i> 110:2417-2427
H-7	Barria et al., 1997. <i>Science</i> 276:2042-2045; Kawamoto and Hidaka, 1984. <i>Biochem. Biophys. Res. Commun.</i> 125:258-264; Wang et al., 1997. <i>JBC</i> 272:1817-1821

continued

**Standard
Measurements,
Data, and
Abbreviations**

A.1K.3

Table A.1K.1 Biological Activity of Drugs Commonly Used In Molecular Biological Research, continued

Drug	References
Herbimycin A	Cowen et al., 1996. <i>JBC</i> 271:22297-22300; Fukazawa et al., 1991. <i>Biochem. Pharmacol.</i> 42:1661-1671; Tiruppathi et al., 1997. <i>JBC</i> 272:25968-25975
KN-62	Bouvard et al., 1998. <i>J. Cell Sci.</i> 111:657-665; Enslin and Soderling, 1994. <i>JBC</i> 269:20872-20877; Wang Hy et al., 1997. <i>JBC</i> 272:1817-1821; Doroudchi et al., 1997. <i>J. Neurosci. Res.</i> 50:514-521.
LY294002	Vlahos et al., 1994. <i>JBC</i> 269:5241-5248; Vlahos et al., 1995. <i>J. Immunol.</i> 154:2413-2422
ML-7	Ohkubo et al., 1996. <i>Eur. J. Pharmacol.</i> 298:175-183; Saitoh et al., 1987. <i>JBC</i> 262:7796-7801; Watanabe et al., 1998. <i>FASEB J.</i> 12:341-348
Olomoucine	Abraham et al., 1995. <i>Biol. Cell</i> 83:105-120; Glab et al., 1994. <i>FEBS Lett.</i> 353:207-211; Howell et al., 1997. <i>Cell. Motil. Cytoskeleton</i> 38:201-214; Misteli and Warren, 1995. <i>J. Cell Sci.</i> 108:2715-2727
PD 98059	Acharya et al., 1998. <i>Cell</i> 92:183-192; Alessi et al., 1995. <i>JBC</i> 270:27489-27494; Dudley et al., 1995. <i>PNAS</i> 92:7686-7689; Karpova et al., 1997. <i>Am. J. Physiol.</i> 272:L558-L565; Waters et al., 1995. <i>JBC</i> 270:20883-20886
Piceatannol	Keely and Parise, 1996. <i>JBC</i> 271:26668-26676; Oliver et al., 1994. <i>JBC</i> 269:29697-29703; Peters et al., 1996. <i>JBC</i> 271:4755-4762
Staurosporine	Couldwell et al., 1994. <i>FEBS Lett.</i> 345: 43-46; Kiss and Deli, 1992. <i>Biochem. J.</i> 288: 853-858; Janicke et al., 1998. <i>JBC</i> 273:9357-9360; Orr et al., 1998. <i>JBC</i> 273:3803-3807
Tyrphostins	Antonyak et al., 1998. <i>JBC</i> 273:2817-2822; Austin and Shields, 1996. <i>J. Cell Biol.</i> 135:1471-1483; Gazit et al., 1989. <i>J. Med. Chem.</i> 32:2344-2352; Gohla et al., 1998. <i>JBC</i> 273:4653-4659
Wortmannin	Cross et al., 1995. <i>JBC</i> 270:25352-25355; Goeger et al., 1988. <i>Biochem. Pharmacol.</i> 37:978-981; Jones and Howell, 1997. <i>J. Cell Biol.</i> 139:339-349; Ptasznik et al., 1997. <i>J. Cell Biol.</i> 137:1127-1136
Phosphatase inhibitors	
Calyculin A	Ishihara et al., 1989. <i>Biochem. Biophys. Res. Commun.</i> 159:871-877; Murakami et al., 1994. <i>Neurosci. Lett.</i> 176: 181-184; Takeuchi et al., 1994. <i>Biochem. Biophys. Res. Commun.</i> 205:1803-1807
Microcystin-LR	Bagu et al., 1997. <i>JBC</i> 272:5087-5097; Eriksson et al., 1990. <i>Biochim. Biophys. Acta</i> 1025:60-66; Honkanen et al., 1990. <i>JBC</i> 265:19401-19404; Rabouille et al., 1995. <i>J. Cell Biol.</i> 129:605-618; Toivola et al., 1997. <i>J. Cell Sci.</i> 110:23-33
Okadaic acid	Gjertsen et al., 1994. <i>J. Cell Sci.</i> 107:3363-3377; Haystead et al., 1989. <i>Nature</i> 337:78-81; Kiguchi et al., 1994. <i>Cell Growth Differ.</i> 5:995-1004; Lucocq, 1992. <i>J. Cell Sci.</i> 103:875-880; Ohoka et al., 1993. <i>Biochem. Biophys. Res. Commun.</i> 197:916-921; Suganuma et al., 1988. <i>PNAS</i> 85:1768-1771 Takai et al., 1987. <i>FEBS Lett.</i> 217:81-84
Phenylarsine oxide	Han and Kohanski, 1997. <i>Biochem. Biophys. Res. Commun.</i> 239:316-321; Fleming et al., 1996. <i>JBC</i> 271:11009-11015; Krutetskaia et al., 1997. <i>Tsitologiya</i> 39:1116-1130; Kussmann and Przybylski, 1995. <i>Methods Enzymol.</i> 251:430-435
Sodium orthovanadate	Brown and Gordon, 1984. <i>JBC</i> 259:9580-9586; Cook et al., 1997. <i>JBC</i> 272:13309-13319; Griffith et al., 1998. <i>JBC</i> 273:10771-10776; Munoz et al., 1992. <i>JBC</i> 267:10381-10388; Seedorf et al., 1995. <i>JBC</i> 270:18953-18960; Seglen and Gordon, 1981. <i>JBC</i> 256:7699-7701; Swarup et al., 1982. <i>Biochem. Biophys. Res. Commun.</i> 107:1104-1109

continued

Table A.1K.1 Biological Activity of Drugs Commonly Used In Molecular Biological Research, continued

Drug	References
Calpain inhibitor I	Figueiredo-Pereira et al., 1994. <i>J. Neurochem.</i> 62: 1989-1994; Klafki et al., 1995. <i>Neurosci Lett.</i> 201:29-32; Milligan et al., 1996. <i>Arch. Biochem. Biophys.</i> 335:388-395
E-64	Banik et al., 1997. <i>Brain Res.</i> 1997 748:205-210; Bush et al., 1997. <i>JBC</i> 272:9086-9092; Sarin et al., 1994. <i>J. Immunol.</i> 153:862-872
Lactacystin	Choi et al., 1997. <i>JBC</i> 272:28479-28484; Dick et al., 1996. <i>JBC</i> 271:7273-7276; Fenteany et al., 1995. <i>Science</i> 268:726-731; Kim et al., 1997. <i>JBC</i> 272:11006-11010; Oda et al., 1996. <i>Biochem. Biophys. Res. Commun.</i> 219: 800-805
Leupeptin	Montenez et al., 1994. <i>Toxicol. Lett.</i> 73:201-208; Sarin et al., 1994. <i>J. Immunol.</i> 153:862-872; Wang et al., 1995. <i>JBC</i> 270:24924-24931
MG-132	Jensen et al., 1995. <i>Cell</i> 83:129-135; Lee and Goldberg, 1996. <i>JBC</i> 271:27280-27284; Meerovitch et al., 1997. <i>JBC</i> 272:6706-6713; Salceda and Caro, 1997. <i>JBC</i> 272:22642-22647
PMSF	Darby et al., 1998. <i>Biochemistry</i> 37:783-791; Turini et al., 1969. <i>J. Pharmacol. Exp. Ther.</i> 167:98-104; Weaver et al., 1993. <i>Biochem. Cell. Biol.</i> 71:488-500
Pepstatin A	Bode and Huber, 1992. <i>Eur. J. Biochem.</i> 204:433-451; Shields et al., 1991. <i>Biochem. Biophys. Res. Commun.</i> 177:1006-1012; Simon et al., 1995. <i>Biochim. Biophys. Acta</i> 1268: 143-151; Yamada et al., 1996. <i>J. Immunol.</i> 157:901-907; Wang et al., 1995. <i>JBC</i> 270:24924-24931
Protein synthesis inhibitors	
Anisomycin	Dong Chen et al., 1996. <i>JBC</i> 271:6328-6332; Kardalinou et al., 1994. <i>Mol. Cell. Biol.</i> 14:1066-1074; Sidhu and Omiecinski, 1998. <i>JBC</i> 273:4769-4775
Cycloheximide	Chow et al., 1995. <i>Exp. Cell Res.</i> 216:149-159; Cotter et al., 1992. <i>Anticancer Res.</i> 12:773-779; Waring, 1990. <i>JBC</i> 265:14476-14480
Emetine	Burhans et al., 1991. <i>EMBO J.</i> 10:4351-4360; Gabathuler et al., 1998. <i>J. Cell Biol.</i> 140:17-27; Sidhu and Omiecinski, 1998. <i>JBC</i> 273:4769-4775
Hygromycin	Gaken et al., 1992. <i>Biotechniques</i> 13:32-34; Hamada et al., 1994. <i>Curr. Genet.</i> 26:251-255; Lama and Carrasco, 1992. <i>JBC</i> 267:15932-15937
Puromycin	Chow et al., 1995. <i>Exp. Cell Res.</i> 216:149-159; Kislauskis et al., 1997. <i>J. Cell Biol.</i> 136:1263-1270; Tachibana et al., 1997. <i>EMBO J.</i> 16:4333-4339; Zhang et al., 1997. <i>Mol. Biol. Cell</i> 8:1943-1954
Transcription inhibitors	
Actinomycin D	Kuhn and Henderson, 1995. <i>JBC</i> 270:20509-20515; McGary et al., 1997. <i>JBC</i> 272:8628-8634; Wu and Yung, 1994. <i>Eur. J. Pharmacol.</i> 270:203-212
α -Amanitin	Baumann et al., 1994. <i>Protein Sci.</i> 3:750-756; Rudd and Luse, 1996. <i>JBC</i> 271:21549-21558; Seiser et al., 1995. <i>J. Biol. Chem.</i> 270:29400-29406
Other compounds	
Brefeldin A	Donaldson et al., 1991. <i>J. Cell Biol.</i> 112:579-588; Donaldson et al., 1992. <i>Nature</i> 360:350-352; Ktistakis et al., 1992. <i>Nature</i> 356:344-346; Lippincott-Schwartz et al., 1990. <i>Cell</i> 60:821-836
Cyclosporin A	Su et al., 1997. <i>J. Cell Biol.</i> 139:1533-1543; Sugimoto et al., 1997. <i>J. Biol. Chem.</i> 272:29415-29418; McKeon, 1991. <i>Cell</i> 66:823-826; Nicolli et al., 1996. <i>JBC</i> 271:2185-2192

continued

**Standard
Measurements,
Data, and
Abbreviations**

A.1K.5

Table A.1K.1 Biological Activity of Drugs Commonly Used In Molecular Biological Research, continued

Drug	References
Desferrioxamine	Ben-Shachar et al., 1995. <i>J. Neurochem.</i> 64:718-723; Bergeron et al., 1996. <i>J. Med. Chem.</i> 39:1575-1581; Dang et al., 1994. <i>Res. Commun. Mol. Pathol. Pharmacol.</i> 86: 43-57; Denicola et al., 1995. <i>Free Radical Biol. Med.</i> 19:11-19; Henderson and Kuhn, 1995. <i>JBC</i> 270:20509-20515
2-Deoxyglucose	Donaldson et al., 1991. <i>J. Cell Biol.</i> 112:579-588; Hill et al., 1998. <i>JBC</i> 273: 3308-3313; Villalba et al., 1994. <i>JBC</i> 269:2468-2476
Dithiothreitol	Braakman et al., 1992. <i>EMBO J.</i> 11:1717-1722; Lodish and Kong, 1993. <i>JBC</i> 268:20598-20605; Simons et al., 1995. <i>J. Cell Biol.</i> 130:41-49; Verde et al., 1995. <i>Eur. J. Cell Biol.</i> 67:267-274
Filipin	Liu et al., 1997. <i>JBC</i> 272:7211-7222; Schnitzer et al., 1994. <i>J. Cell Biol.</i> 127:1217-1232; Silberkang et al., 1983. <i>JBC</i> 258:8503-8511; Smart et al., 1996. <i>J. Cell Biol.</i> 134:1169-1177; Stahl and Mueller, 1995. <i>J. Cell Biol.</i> 129:335-344
Fumonisin B ₁	Merrill et al., 1993. <i>JBC</i> 268:27299-27306; Sandvig et al., 1996. <i>Mol. Biol. Cell.</i> 7:1391-1404; Spiegel and Merrill, 1996. <i>FASEB J.</i> 10:1388-1397; Wang et al., 1991. <i>JBC</i> 266:14486-14490; Wang et al., 1996. <i>PNAS</i> 93:3461-3465
Geneticin	Canaani and Berg, 1982. <i>PNAS</i> 79:5166-5170; Southern and Berg, 1982. <i>J. Mol. Appl. Genet.</i> 1:327-341; Morris et al., 1996. <i>JBC</i> 271:15468-15477.
Leptomycin B	Fornerod et al., 1997. <i>Cell</i> 90:1051-1060; Fukuda et al., 1997. <i>Nature</i> 390:308-311; Wada et al., 1998. <i>EMBO J.</i> 17:1635-1641; Wolff et al., 1997. <i>Chem. Biol.</i> 4:139-147
Lovastatin	Carel et al., 1996. <i>JBC</i> 271:30625-30630; Hancock et al., 1989. <i>Cell</i> 57:1167-1177; Jakobisiak et al., 1991. <i>PNAS</i> 88:3628-3632. Mendola and Backer, 1990. <i>Cell Growth Differ.</i> 1:499-502; Vincent et al., 1991. <i>Biochem. Biophys. Res. Commun.</i> 180:1284-1289
Lysophosphatidic acid	An et al., 1998. <i>JBC</i> 273:7906-7910; Jalink et al., 1990. <i>JBC</i> 265:12232-12239; Moolenaar, 1995. <i>JBC</i> 270:12949-12952; Zhang et al., 1997. <i>Mol. Biol. Cell.</i> 8:1415-1425
Mastoparan	Huber et al., 1997. <i>J. Cell Sci.</i> 110: 2955-2968; Klinker et al., 1996. <i>Biochem. Pharmacol.</i> 51:217-223; Konrad et al., 1995. <i>JBC</i> 270: 12869-12876; Schwaninger et al., 1992. <i>J. Cell Biol.</i> 119:1077-1096; Smith et al., 1995. <i>JBC</i> 270:18323-18328
Ouabain	Croyle et al., 1997. <i>Eur. J. Biochem.</i> 248:488-495; Peng et al., 1996. <i>JBC</i> 271: 10372-10378; Swann and Steketee, 1989. <i>J. Neurochem.</i> 52:1598-1604
PDMP	Chen et al., 1995. <i>JBC</i> 270:13291-13297; Inokuchi et al., 1987. <i>Cancer Lett.</i> 38:23-30; Maceyka and Machamer, 1997. <i>J. Cell Biol.</i> 139:1411-1418; Rosenwald et al., 1992. <i>Biochemistry</i> 31:3581-3590; Sandvig et al., 1996. <i>Mol. Biol. Cell.</i> 7:1391-1404; Uemura et al., 1990. <i>J. Biochem.</i> 108:525-530
Pertussis toxin	An et al., 1998. <i>JBC</i> 273:7906-7910; Hewlett et al., 1983. <i>Infect. Immun.</i> 40:1198-1203; Jakobs et al., 1984. <i>Eur. J. Biochem.</i> 140:177-181; Kopf and Woolkalis, 1991. <i>Methods Enzymol.</i> 195:257-266; Luttrell et al., 1997. <i>JBC</i> 272:31648-31656
Phorbol esters	Janecki et al., 1998. <i>JBC</i> 273:8790-8798; Macfarlane and O'Donnell, 1993. <i>Leukemia</i> 7:1846-1851; Nidel et al., 1983. <i>PNAS</i> 80:36-40; Nishizuka, 1986. <i>Science</i> 233:305-312; Oishi and Yamaguchi, 1994. <i>J. Cell. Biochem.</i> 55: 168-172; Schmidt et al., 1998. <i>JBC</i> 273:7413-7422; Tepper et al., 1995. <i>PNAS</i> 92:8443-8447

continued

Table A.1K.1 Biological Activity of Drugs Commonly Used In Molecular Biological Research, continued

Drug	References
Rapamycin	Brown et al., 1994. <i>Nature</i> 369:756-758 Jefferies et al., 1997. <i>EMBO J.</i> 16:3693-3704 Kozlovsky et al., 1997. <i>J. Biol. Chem.</i> 272:33367-33372 Liu et al., 1991. <i>Cell</i> 66:807-815 Sabers et al., 1995. <i>J. Biol. Chem.</i> 270:815-822
Sodium azide	Bhat et al., 1996. <i>JBC</i> 271:32551-32556; Donaldson et al., 1991. <i>J. Cell Biol.</i> 112:579-588; van Klompenburg et al., 1997. <i>EMBO J.</i> 16:4261-4266
Sodium butyrate	Calabresse et al., 1993. <i>Biochem. Biophys. Res. Commun.</i> 195:31-38; Gonzalez-Garay and Cabral, 1996. <i>J. Cell Biol.</i> 135:1525-1534; Russo et al., 1997. <i>Biochem. Biophys. Res. Commun.</i> 233:673-677; Vaziri et al., 1996. <i>JBC</i> 271:25921-25927; White et al., 1995. <i>J. Cell Sci.</i> 108:441-455
Trifluoperazine	Aussel et al., 1995. <i>JBC</i> 270:8032-8036; Massom et al., 1990. <i>Biochemistry</i> 29:671-681; Rao, 1987. <i>Biochem. Biophys. Res. Commun.</i> 148:768-775
Valinomycin	Inai et al., 1997. <i>Cell Struct. Funct.</i> 22:555-563; Loiseau et al., 1997. <i>Biochim. Biophys. Acta.</i> 1330:39-49; Orlov et al., 1994. <i>FEBS Lett.</i> 345:104-106; Szabo et al., 1997. <i>JBC</i> 272:23165-23171
W-7	de Figueiredo and Brown, 1995. <i>Mol. Biol. Cell.</i> 6:871-887; Hidaka et al., 1981. <i>PNAS</i> 78:4354-4357; Hunziker, 1994. <i>JBC</i> 269:29003-29009; Wolf and Gross, 1996. <i>JBC</i> 271:20989-20992

DRUGS COMMONLY USED IN MOLECULAR BIOLOGY

A23187

Calcium ionophore; forms stable complexes with divalent cations and increases their passage across biological membranes. Useful tool for increasing intracellular calcium concentration. The effectiveness of A23187 is dependent on the presence of extracellular calcium. Can be used as a fluorescent probe for investigating protein hydrophobicity. 4-Bromo-A23187 is a nonfluorescent derivative.

Soluble in: DMSO, methanol

Stock concentration: 100 mM (store at 4°C protected from light)

Working concentration: 0.1 to 20 μM

Duration of incubation: 2 min to 24 hr

Aggregates over time in aqueous systems.

Actinomycin D

Inhibits transcription by complexing with deoxyguanosine residues on DNA and blocking the movement of RNA polymerase. A potent inducer of apoptosis in many cell lines. However, actinomycin D has also been shown to suppress programmed cell death of PC12 cells induced by the topoisomerase II inhibitor etoposide.

Soluble in: Methanol

Stock concentration: 100 mM (store at 4°C)

Working concentration: 1 to 5 μM

Duration of incubation: 5 min to 24 hr

α-Amanitin

Acts as a potent and specific inhibitor of mRNA synthesis by binding preferentially to RNA polymerase II. At high concentrations also inhibits RNA polymerase III.

Soluble in: Methanol, water

Stock concentration: 2 to 10 mg/ml (store at 4°C protected from light)

Working concentration: 1 to 10 μg/ml (pol II) to 200 μg/ml (pol III)

Duration of incubation: 15 to 60 min

Ammonium chloride (NH₄Cl)

Permeant weak base. Used to neutralize acidic endomembrane compartments.

Inhibits synthesis of sphingoid bases.

Soluble in: Water (freely soluble)

Stock concentration: 5 M (store at 4°C)

Working concentration: 1 to 50 mM

Duration of incubation: Effective within 15 sec

Anisomycin

Inhibits protein synthesis by blocking the peptidyl transferase step during translation. Activates p54 (JNK2) and MAP kinases. May be a useful tool to study cytoplasmic signals that result in nuclear signaling and *c-fos* and *c-jun* induction. Also known to induce apoptosis in U937 cells.

Soluble in: DMSO

Stock concentration: 100 μg/ml (store at 4°C)

Working concentration: 50 ng/ml to 1 μg/ml.

Duration of incubation: 30 min to 16 hr, depending on properties studied

Aphidicolin

Cell synchronization reagent. Reversible inhibitor of DNA polymerase α and δ; blocks cell cycle at early S phase. Potentiates apoptosis induced by arabinosyl nucleosides in leukemia cell lines.

Soluble in: DMSO, methanol

Stock concentration: 2 mg/ml (store at 4°C)

Working concentration: 0.5 to 100 μg/ml.

Duration of incubation: 12 to 24 hr

Ara-C (cytosine arabinoside)

Inhibits DNA synthesis. S-phase-toxic reagent whose active metabolite (ara-CTP) is a substrate for DNA polymerases and is incorporated into DNA. Anticancer, antiviral agent that is especially effective against leukemias. Induces apoptosis in human myeloid leukemia cells and in rat sympathetic neurons.

Soluble in: Water

Stock concentration: 20 mg/ml (store at 4°C)

Working concentration: 0.1 to 1 μg/ml

Duration of incubation: >3 hr

Bafilomycin A₁

A potent and specific inhibitor of vacuolar-type H⁺-ATPases. Valuable tool for distinguishing different types of ATPases. Blocks lysosomal trafficking in macrophages.

Soluble in: DMSO

Stock concentration: 50 μM (store at -20°C protected from light)

Working concentration: 10 to 100 nM

Duration of incubation: 10 min to 2 hr

BAPTA

Ca²⁺ chelator with a 10⁵-fold greater affinity for Ca²⁺ than for Mg²⁺; can be used to control the level of both intracellular (using its membrane-permeant AM ester) and extracellular Ca²⁺.

Soluble in: DMSO

Stock concentration: 1 to 10 mM (store at –20°C in aliquots, protected from light; avoid repeated freeze-thawing)

Working concentration: (BAPTA-AM): 1 to 20 μM

Duration of incubation: 15 to 60 min at 20° to 37°C

Before incubation with BAPTA, wash cells 2 to 3 times with serum-free medium (serum may contain esterase activity). The cell-loading medium should also be free of amino acids or buffers containing primary or secondary amines that may cleave the AM esters and prevent loading.

Bisindolylmaleimide I (GF 109203X)

A highly selective cell-permeant protein kinase C (PKC) inhibitor that is structurally similar to staurosporine, but has higher selectivity. May inhibit protein kinase A at high concentrations. Acts as a competitive inhibitor for the ATP-binding site of the PKC catalytic domain. Since ATP levels are generally very high in cells, the potency of bisindolylmaleimide I is reduced accordingly in whole-cell assays.

Soluble in: DMSO

Stock concentration: 2 mM (store at ≤4°C)

Working concentration: 20 nM to 1 μM

Duration of incubation: 15 min to 6 hr

Water-soluble salts are available.

Brefeldin A

Inhibits GTP nucleotide exchange onto several members of the ARF (ADP ribosylation factor) family. Inhibits binding of the cytosolic coatamer (COPI) complex to Golgi membranes; induces the rapid redistribution of the Golgi apparatus into the ER; blocks transport out of the ER in a number of cell lines. Reversible.

Soluble in: Methanol

Stock concentration: 1 to 20 mM (store at –20°C)

Working concentration: 1 to 5 μM

Duration of incubation: 5 min to 24 hr; effects are rapid (30 sec)

8-Bromo–cyclic AMP

Cell-permeant cyclic AMP analog. Activates protein kinase A. Increased resistance to degradation by cellular phosphodiesterases as compared to cyclic AMP.

Soluble in: Water

Stock concentration: 100 mM (store at –20°C)

Working concentration: 10 to 500 μM

Duration of incubation: Up to 24 hr

Calpain inhibitor I (ALLN)

Inhibitor of calpain I, calpain II, cathepsin B, and cathepsin L. A peptide aldehyde, which inhibits neutral cysteine proteases and the proteasome. Protects against neuronal damage caused by hypoxia and ischemia. Inhibits proteolysis of IκB by the ubiquitin-proteasome complex. Inhibits cell-cycle progression at G₁/S and metaphase/anaphase in CHO cells by inhibiting cyclin B degradation. Membrane-permeant due to low molecular weight and lack of charged residues.

Soluble in: DMSO, methanol, dimethylformamide

Stock concentration: 25 mM (store at 4°C)

Working concentration: 25 to 100 μM

Duration of incubation: 1 to 18 hr

Calphostin C

Potent and highly selective inhibitor of protein kinase C (PKC; $K_i = 50$ nM). Competes with phorbol esters and diacylglycerol for binding to the PKC regulatory domain. Does not compete with Ca^{2+} or phospholipids. At higher concentrations inhibits myosin light chain kinase ($K_i > 5$ μM), protein kinase A ($K_i > 50$ μM), protein kinase G ($K_i > 25$ μM), and p60^{v-src} ($K_i > 50$ μM).

Soluble in: DMSO

Stock concentration: 1 mM (store 4°C protected from light)

Working concentration: 10 nM to 3 μM

Duration of incubation: 15 to 60 min

Brief exposure to visible light in the presence of PKC is required for PKC inhibition by calphostin C. See Table 1 (Selectivity of Selected Kinase Inhibitors) in the Biomol Catalog and Handbook, 5th ed., and Technical Note #11 from Alexis Biochemicals for additional information.

Calyculin A

Potent cell-permeant inhibitor with high specificity for the Ser/Thr protein phosphatases 1 and 2A. Calyculin A is 20 to 300 times more potent than okadaic acid as a PP-1 class phosphatase inhibitor. Stimulates contraction of smooth muscle, induces intracellular protein phosphorylation in cultured human keratinocytes, and inhibits apoptosis.

Soluble in: DMSO, ethanol

Stock concentration: 10 μM (store at -20°C protected from light and moisture)

Working concentration: 0.5 to 50 nM

Duration of incubation: 15 min to 2 hr

See Technical Note #19 from Alexis Biochemicals for additional information. May cause cell rounding.

Camptothecin

A reversible DNA topoisomerase I inhibitor. Induces breaks at replication forks by binding to and stabilizing the topoisomerase-DNA covalent complex. Causes S-phase cytotoxicity. Possesses anti-leukemic and anti-tumor properties. Inhibits *tat*-mediated transactivation of HIV-1.

Soluble in: DMSO

Stock concentration: 1 to 10 mM (store at 4°C)

Working concentration: 0.1 to 10 μM

Duration of incubation: Literature shows use from 15 min to 24 hr

Castanospermine

Inhibitor of α - and β -glucosidases; inhibitor of glycoprotein processing. Prevents calnexin and calreticulin binding to N-linked glycans on newly synthesized glycoproteins. Inhibits HIV infectivity.

Soluble in: Water

Stock concentration: 1 mM (store at 4°C)

Working concentration: 1 to 5 μM

Duration of incubation: 15 min to 3 hr

CCCP (carbonyl cyanide-*m*-chlorophenyl hydrazone)

Proton ionophore. Uncoupling agent for oxidative phosphorylation that inhibits mitochondrial function. Approximately 100 times as effective as 2,4-dinitrophenol at collapsing membrane potential. Inhibits transport processes and depresses growth.

Soluble in: DMSO, ethanol

Stock concentration: 1 to 10 mM (store at 4°C)

Working concentration: 1 to 5 μM

Duration of incubation: 5 to 15 min

Chelerythrine chloride

Potent, selective, cell-permeant inhibitor of protein kinase C ($K_i = 0.66 \mu\text{M}$). Acts on the catalytic domain. Chelerythrine shows competitive kinetics with PKC substrates, but is not competitive with ATP. Thus, the high concentration of ATP within cells should not lower the potency of chelerythrine in whole cells as compared with that seen in purified enzyme preparations. Inhibits thromboxane formation and phosphoinositide metabolism in platelets. Induces apoptosis in HL-60 cells.

Soluble in: DMSO

Stock concentration: 10 mM (store at -20°C)

Working concentration: 1 μM

Duration of incubation: 15 min to 2 hr

See Technical Note #8 from Alexis Biochemicals for additional information.

Chloroquine

Tertiary amine that accumulates within and neutralizes the pH of acidic organelles; various effects on phagosome-endosome and phagosome-lysosome fusion. Antimalarial drug that works via carrier-mediated uptake in *P. falciparum*. May activate protein kinases.

Soluble in: Water

Stock concentration: 1 to 10 mg/ml (store at room temperature)

Working concentration: 10 to 200 μg/ml

Duration of incubation: 15 min to 2 hr

Cholera toxin

Contains a single A subunit (mol. wt. = 29 kDa) and a B subunit (mol. wt. = 55 kDa) containing five B polypeptide chains. The B subunit binds to GM₁ ganglioside receptors on the surface of cells and facilitates transport of the A subunit through the membrane. The A subunit catalyzes the ADP-ribosylation of an arginine residue on the α subunit of heterotrimeric G proteins (primarily G_s), reducing its intrinsic GTPase activity. Toxicity results from activation of membrane-bound adenylate cyclase. Consequently, increased intracellular cAMP levels result in increased electrolyte transport out of the cell and water loss. Cholera toxin requires ADP-ribosylation factor (ARF) for maximal activity.

Soluble in: Water

Stock concentration: 1 mg/ml (store at 4°C, do not freeze)

Working concentration: 100 ng/ml to 2 μg/ml

Duration of incubation: 2 to 24 hr

Colcemid

Cell synchronization agent. Depolymerizes microtubules and limits microtubule formation. Low concentrations inactivate spindle dynamics. Induces apoptosis by blocking mitosis in HeLa S3 cells. Colcemid is a less toxic derivative of colchicine.

Soluble in: Ethanol

Stock concentration: 1 mM (store at or below room temperature)

Working concentration: 1 to 10 μ M

Duration of incubation: 1 to 24 hr, depending on process studied

Colchicine

Inhibitor of mitosis, used in cell-division studies. Disrupts microtubules and inhibits tubulin polymerization. Induces apoptosis in PC 12 cells and in cerebellar granule cells.

Soluble in: Ethanol

Stock concentration: 1 mM (store at or below room temperature protected from light and moisture)

Working concentration: 1 to 10 μ M

Duration of incubation: 1 to 24 hr, depending on process studied

Concanamycin B

Highly specific and sensitive inhibitor of vacuolar-type H⁺-ATPases ($K_i = 20$ pM). Related to concanamycin A (folimycin). More potent and specific than bafilomycin A₁. Inhibits acidification of organelles such as lysosomes and the Golgi apparatus. Blocks cell-surface expression of viral glycoproteins without affecting their synthesis.

Soluble in: Methanol, ethanol

Stock concentration: 10 μ M (store at -20°C protected from light)

Working concentration: 50 nM

Duration of incubation: 5 min to 1 hr

Cycloheximide

Inhibits protein synthesis in eukaryotes but not prokaryotes. Blocks the translocation step during translation. Induces apoptosis in a number of cell types. However, it inhibits DNA cleavage in rat thymocytes treated with thapsigargin and ionomycin.

Soluble in: Water, ethanol, methanol

Stock concentration: 10 mg/ml (store at or below room temperature)

Working concentration: 1 to 100 μ g/ml, depending on cell type

Duration of incubation: Effective within 10 min

To achieve >90% inhibition of protein synthesis, only 1 to 10 μ g/ml is required in CHO and HeLa cells, but 100 μ g/ml is required in COS cells.

Cyclosporin A (CsA)

Cyclic oligopeptide with immunosuppressant properties. Induces apoptosis in some cell types, while inhibiting apoptosis in others. A complex of cyclosporin A and cyclophilin inhibits protein phosphatase 2B (calcineurin) with affinity at the nanomolar level. Inhibits nitric oxide synthesis induced by interleukin-1 α , lipopolysaccharides, and TNF α .

Soluble in: Ethanol, methanol

Stock concentration: 1 to 5 mM (store at 4°C)

Working concentration: 0.1 to 10 μ M

Duration of incubation: Used anywhere from 15 min to 24 hr

Cytochalasin B

Cell-permeant fungal toxin that blocks the formation of contractile microfilaments. Shortens actin filaments by blocking monomer addition at the barbed (fast-growing)

continued

end of polymers. Inhibits cytoplasmic division, cell movement, phagocytosis, platelet aggregation, and glucose transport.

Soluble in: DMSO, ethanol

Stock concentration: 10 mM (store at -20°C protected from light)

Working concentration: 1 to 20 μM

Duration of incubation: 15 min to 2 hr

Cytochalasin D

Approximately 10-fold more potent than cytochalasin B in inhibiting actin filament function. Does not inhibit sugar transport in cells. Modulates CD4 cross-linking in T lymphocytes and increases intracellular Ca^{2+} . Exhibits antitumor activity.

Soluble in: DMSO

Stock concentration: 10 mM (store at -20°C protected from light)

Working concentration: 1 to 20 μM

Duration of incubation: 15 min to 2 hr.

Desferrioxamine (DFO)

Iron-chelating agent. Commonly used in therapy as a chelator of ferric iron in iron overload disorders. Protects against dopamine-induced cell death. Also interferes with hydroxy-radical formation. Shows an antiproliferative effect on vascular smooth muscle cells.

Soluble in: DMSO; slightly soluble in water

Stock concentration: 10 to 50 mM (store at 4°C)

Working concentration: 10 μM to 2 mM

Duration of incubation: Up to 18 hr

2-Deoxyglucose

Nonmetabolizable derivative of glucose. Competes with glucose for the GLUT-2 transporter; phosphorylation of 2-deoxyglucose by hexokinase effectively inhibits glucose flux through the glycolytic pathway. Used in combination with sodium azide or oligomycin to reduce cellular ATP levels. Blocks inhibition of IL-1 release by high glucose levels in RAW 264.7 cells.

Soluble in: Water

Stock concentration: 1 M (store at 4°C)

Working concentration: 5 to 50 mM

Duration of incubation: 15 min to 3 hr

Deoxymannojirimycin

Competitive α -mannosidase I inhibitor that blocks conversion of high mannose forms to complex oligosaccharides. Inhibits mammalian Golgi α -mannosidase I (an α -1,2-mannosidase). Other rat liver mannosidases are not significantly affected (α -1,2-specific ER mannosidase is only inhibited 2% to 5% by 100 mM deoxymannojirimycin, and Golgi α -mannosidase II is inhibited \sim 14%).

Soluble in: Water

Stock concentration: 100 mM (store at -20°C)

Working concentration: 1 to 5 mM

Duration of incubation: Anywhere from 30 min to 24 hr

Deoxynojirimycin

Specific glucosidase inhibitor. Inhibits endoplasmic reticulum trimming glucosidases I and II, which sequentially remove three glucose residues from $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ in N-linked glycan biosynthesis. Prevents calnexin and calreticulin binding to N-linked glycoproteins within the ER.

continued

**Standard
Measurements,
Data, and
Abbreviations**

A.1K.13

Soluble in: Water

Stock concentration: 100 mM (store at 4°C)

Working conditions: 1 to 5 mM.

Duration of incubation: 15 min to 24 hr

At concentrations >1 mM, deoxynojirimycin may inhibit lipid-linked oligosaccharide biosynthesis as well as trimming. In such cases, N-methyldeoxynojirimycin may be a more effective inhibitor, possibly owing to an increased ability to cross cell membranes, afforded by the N-methyl group.

Dibutyl cyclic AMP

Highly membrane-permeant cAMP analog resistant to phosphodiesterase cleavage. Constitutive activator of protein kinase A. This product releases butyrate due to intracellular and extracellular esterase action. Butyrate may have its own distinct biological effects (see sodium butyrate).

Soluble in: DMSO, ethanol

Stock concentration: 1 M (store at -20°C)

Working concentration: 100 µM to 1 mM

Duration of incubation: Anywhere from 1 to 48 hr

Dithiothreitol (DTT; Cleland's reagent)

Cell-permeant protective agent for SH groups; maintains monothiols completely in the reduced state and reduces disulfides quantitatively. DTT interferes with the folding and export of proteins located in the endoplasmic reticulum, but it does not prevent the transfer from the intermediate compartment to the Golgi complex. Reversible.

Soluble in: Water, ethanol

Stock concentration: 1 M (store at 4°C)

Working concentration: 1 to 10 mM

Duration of incubation: 1 min to several hours

E-64

Irreversible inhibitor of cysteine proteases (papain and cathepsins B and L). Has no action on cysteine residues in other proteins.

Soluble in: Water

Stock concentration: 1 mg/ml (store at -20°C)

Working concentration: 0.5 to 10 µg/ml

Duration of incubation: Up to 24 hr

Emetine

Irreversibly blocks protein synthesis by inhibiting movement of ribosomes along mRNA. Stimulates rapid and differential phosphorylation of the stress-activated protein kinase/c-Jun kinase (SAPK/JNK) pathway. Prevents apoptosis in several cell lines. In primary rat hepatocytes, the relative potency of inhibition of several protein synthesis inhibitors is in the order: emetine > anisomycin > cycloheximide > puromycin, with puromycin exhibiting only marginal inhibition at a concentration of 1 µM. In fact, 90% to 95% inhibition of protein synthesis was achieved only with emetine and anisomycin, at 10 µM concentrations. Cycloheximide and puromycin exerted only 80% and 60% inhibition, respectively, at a similar concentration.

Soluble in: Ethanol

Stock concentration: 10 to 100 mM (store at 4°C protected from light)

Working concentration: 10 to 20 µM

Duration of incubation: 15 to 30 min; can incubate 24 hr

Etoposide (VP-16)

Topoisomerase II inhibitor. Stabilizes the covalent complexes of topoisomerase II with DNA. Has major activity against a number of tumors, including germ cell neoplasms, small cell lung cancer, and malignant lymphoma. Induces apoptosis in mouse thymocytes and HL-60 cells. Activates PKC α .

Soluble in: DMSO

Stock concentration: 100 to 500 mM (store at room temperature)

Working concentration: 50 to 200 μ M

Duration of incubation: 1 to 24 hr

Filipin

A cholesterol-binding fluorochrome. Specific for unesterified cholesterol. Binds and removes cholesterol from cell surface membranes. Reversibly disassembles caveolae.

Soluble in: Methanol

Stock concentration: 500 μ g/ml (store 4°C protected from light)

Working concentration: 5 to 50 μ g/ml

Duration of incubation: 1 hr

Forskolin

Activates adenylate cyclase by interacting directly with the catalytic subunit. Leads to an increase in the intracellular concentration of cAMP. Several forskolin derivatives are available having different and improved properties. Enhances detoxification of brefeldin A.

Soluble in: DMSO, ethanol

Stock concentration: 10 to 100 mM (store at 4°C)

Working concentration: 10 μ M (to increase cAMP levels); 100 μ M (to inhibit brefeldin A)

Duration of incubation: Depending on assay, incubate cells 15 min to 12 hr

Fumonisin B₁

Inhibits sphingolipid biosynthesis via inhibition of sphingosine *N*-acyltransferase (ceramide synthase). Sphingomyelin biosynthesis is preferentially inhibited versus glycosphingolipids in neuronal cells. Inhibits the butyric acid-induced increase in transport of cell-associated Shiga toxin to the Golgi apparatus and the ER. Induces apoptosis in monkey kidney cells.

Soluble in: methanol

Stock concentration: 10 to 100 mM (store at 4°C)

Working concentration: Anywhere from 1 to 100 μ M

Duration of incubation: 15 min to 18 hr, depending on process studied

Geneticin (G418)

Aminoglycoside toxic to bacteria, yeast, higher plants, protozoa, and mammalian cells. Used for the selection and maintenance of eukaryotic cells stably transfected with the neomycin (*neo*) resistance genes from transposons Tn5 and Tn601.

Soluble in: Water or culture medium

Stock concentration: 2 mg/ml (active G418) in cell culture medium, adjust pH to ~7.4 (store at 4°C)

Working concentration: Usually 50 to 1000 μ g/ml (optimal concentration must be determined experimentally and varies with the cell type used)

Duration of incubation: >1 week

In cell types with relatively stable genomes (e.g., CHO), continuous incubation in geneticin is not generally necessary once stable cells have been selected. Expression in stable cells with down-regulated viral promoters can be enhanced with sodium butyrate.

Genistein

Inhibits protein tyrosine kinases by acting as a competitive inhibitor of ATP. Prevents EGF-stimulated tyrosine phosphorylation in A431 cells, as well as inhibiting kinases in other cultured cells.

Soluble in: DMSO

Stock concentration: 100 to 500 mM (store at -20°C)

Working concentration: 50 to 300 μM

Duration of incubation: 15 min to 1 hr

See Table II (Selectivity of Tyrosine Protein Kinase Inhibitors) in the Biomol Catalog and Handbook, 5th ed.

H-7

A broad-based, cell-permeant serine/threonine kinase inhibitor. Inhibits protein kinase C ($K_i = 6.0 \mu\text{M}$), protein kinase A ($K_i = 3.0 \mu\text{M}$), protein kinase G ($K_i = 5.8 \mu\text{M}$), and myosin light chain kinase ($K_i = 97 \mu\text{M}$). Induces apoptotic DNA fragmentation and cell death in HL-60 cells. Numerous analogs with different selectivities are available.

Soluble in: Water

Stock concentration: 100 mM (store at 4°C)

Working concentration: 10 to 100 μM

Duration of incubation: 15 min to 3 hr

See Table I (Selectivity of Selected Kinase Inhibitors) in the Biomol Catalog and Handbook, 5th ed., and Technical Note #10 from Alexis Biochemicals.

Herbimycin A

An irreversible and selective cell-permeant protein tyrosine kinase inhibitor; reacts with thiol groups. It is effective on Src, Yes, Fps, Ros, Abl, and ErbB oncogene products. Inhibits PDGF-induced phospholipase D activation in a dose-dependent manner.

Soluble in: DMSO

Stock concentration: 10 mM (store at -20°C protected from light)

Working concentration: 1 to 10 μM

Duration of incubation: 15 min to 1 hr

Hydroxyurea

Antineoplastic reagent. Blocks DNA synthesis by inhibiting ribonucleotide reductase; accumulates cells at G1/S interface.

Soluble in: Water

Stock concentration: 1 M (store at 4°C)

Working concentration: 50 μM to 1 mM

Duration of incubation: Up to 24 hr

Hygromycin

Inhibitor of both prokaryotic and eukaryotic protein synthesis; inhibits at the translocation step on 70S ribosomes and causes misreading of mRNA. The *E. coli hph* hygromycin B-resistant gene is widely used for selection of recombinant clones in a variety of cell types.

Hygromycin B is sold as an aqueous solution. Actual activity and concentration are given for each lot of product. Concentrations used must be experimentally determined.

Ionomycin

Ca^{2+} ionophore. Useful for increasing intracellular Ca^{2+} concentrations and in measurement of cytoplasmic free Ca^{2+} . More effective than A23187 and is nonfluorescent.

Soluble in: DMSO and methanol

Stock concentration: 1 mM (store at 4°C protected from light)

continued

Working concentration: Usually 1 μM

Duration of incubation: 15 to 30 min.

KN-62

Potent and selective inhibitor of Ca^{2+} /calmodulin kinase II ($K_i = 0.9 \mu\text{M}$), displaying a K_i for CaM kinase II more than 2 orders of magnitude lower than those for protein kinase C, protein kinase A, and myosin light chain kinase. A second inhibitor of CaM kinase II, KN-93, is more soluble in water and equally selective for CaM kinase II ($K_i = 0.3 \mu\text{M}$). Prevents agonist-mediated activation of $\text{Ins}(1,4,5)\text{P}_3$ 3-kinase. Inhibits differentiation of 3T3-L1 embryonic fibroblasts to adipocytes. Inactive analogs are available.

Soluble in: DMSO

Stock concentration: 10 mM (store at 4°C)

Working concentration: 2 to 10 μM

Duration of incubation: 30 min (can incubate cells up to 48 hr)

Lactacystin

A cell-permeant and irreversible proteasome inhibitor. Blocks proteasome activity by targeting the catalytic β subunit. Induces neurite outgrowth in Neuro 2A mouse neuroblastoma cells and inhibits progression of synchronized Neuro 2A cells and MG-63 human osteosarcoma cells beyond the G_1 phase of the cell cycle. Inhibits $\text{NF}\kappa\text{B}$ activation. Has revealed the role of the proteasome in the degradation of many ER proteins.

Soluble in: DMSO

Stock concentration: 10 mM (store at -20°C)

Working concentration: 10 to 20 μM

Duration of incubation: 1 to 12 hr

Latrunculin A

Inhibits actin polymerization and disrupts microfilament organization as well as microfilament-mediated processes; 10 to 100-fold more potent than cytochalasins. Whereas cytochalasins induce dissolution of F-actin and stress-fiber contraction in fibroblasts in culture, the latrunculins (A and the less potent B) cause a shortening and thickening of stress fibers. In addition, the latrunculins sequester actin monomers, whereas with the cytochalasins, actin remains in an oligomer form. Thus, the two classes of compounds may have different target sites. Reversible.

Soluble in: DMSO, ethanol

Stock concentration: 10 mg/ml (store at -20°C)

Working concentration: 0.2 to 10 $\mu\text{g/ml}$

Duration of incubation: 1 to 12 hr

Leptomycin B

A potent and specific inhibitor of the NES-dependent nuclear export of proteins; binds to the export receptor CRM1. Exhibits antifungal and antitumor effects, inhibits the nucleo-cytoplasmic translocation of the human immunodeficiency virus type 1 regulatory protein Rev, and exhibits significant antiproliferative activity.

Soluble in: Ethanol

Stock concentration: 100 μM to 1 mM (store at -20°C)

Working concentration: 10 to 100 nM

Duration of incubation: 30 min to 3 hr

Leupeptin

A reversible inhibitor of trypsin-like and cysteine proteases (including trypsin, plasmin, proteinase K, papain, thrombin, and cathepsin A and B). Inhibits activation-induced programmed cell death in T lymphocytes.

Soluble in: Water

Stock concentration: 1 to 10 mM (store at -20°C)

Working concentration: 10 to 100 μM

Duration of incubation: 15 min to several hours

Lovastatin

An antihypercholesterolemic agent and inhibitor of 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase; depletes endogenous pools of mevalonic acid, thereby blocking protein isoprenylation and cholesterol synthesis. Has a number of cellular effects. Blocks *N-ras* oncogene-induced neuronal differentiation, inhibits growth factor signaling and causes cells to arrest in late G_1 phase.

Soluble in: DMSO, ethanol

Stock concentration: 4 mg/ml (store at -20°C)

Working concentration: 20 μM

Duration of incubation: 6 to 24 hr

Lysophosphatidic acid (LPA)

Activates a number of signaling pathways and processes via heterotrimeric G proteins (primarily G_i and G_q) including: inhibition of adenylate cyclase, activation of Ras and the Raf/MAP kinase pathway, stimulation of phospholipases C and D, and stress-fiber formation through the activation of Rho.

Solubility: A stock solution may be prepared at 10 mg/ml in 95:5:5 chloroform/methanol/acetic acid (gives a clear solution). Solubility in dimethylsulfoxide (DMSO) or ethanol is limited. The sodium salt of oleoyl-LPA is reported to be readily soluble at 5 mg/ml (~ 11 mM) in calcium and magnesium-free buffers (Jalink et al., 1990). Solubilization has also been achieved (Seufferlein and Rozengurt, 1994) in phosphate-buffered saline (PBS), pH 7.4, or calcium- and magnesium-free Dulbecco's PBS (CMF-DPBS), pH 7.4 (see APPENDIX 2 for recipes), at up to 3 mM (0.14 mg/ml) in the presence of 0.1% (w/v) BSA (essentially fatty-acid free).

Storage. LPA should be stable in solution under neutral conditions. Freezer storage is recommended for solutions or aqueous preparations. Maintaining the product under an inert atmosphere (nitrogen or argon) may be appropriate for some applications.

Working concentration. Anywhere from 500 nM (1-hr incubation) to 100 μM (15-min incubation).

LY294002 (PI 3-kinase inhibitor)

Reversible inhibitor of phosphatidylinositol-3-kinase that acts on the ATP-binding site of the enzyme. Does not affect the activity of EGF receptor kinase, MAP kinase, PKC, PI4-kinase, S6 kinase, and *c-src*. Blocks proliferation of cultured rabbit aortic smooth muscle cells without inducing apoptosis. Wortmannin is more selective and more potent, but is irreversible.

Soluble in: DMSO and ethanol

Stock concentration: 1 to 10 mM (store in aliquots at -20°C)

Working concentration: 1 to 2 μM

Duration of incubation: 15 min to 3 hr

Mastoparan

Relatively cell-permeant synthetic peptide capable of directly activating pertussis toxin-sensitive G proteins by a mechanism analogous to that of G-protein-coupled receptors. Acts preferentially on G_i and G_o rather than G_s. Stimulates insulin secretion in permeabilized cells, and can increase intracellular Ca²⁺ levels. Inhibits calmodulin and activates phospholipase A₂.

Soluble in: Water

Stock concentration: 1 mM (store at –20°C)

Working concentration: 10 to 50 μM

Duration of incubation: 15 min to 1 hr

Microcystin-LR

Cyclic heptapeptide; potent inhibitor of protein phosphatases 1 and 2A (PP-1 and PP-2A). Unlike okadaic acid, microcystin-LR is equally effective on both PP-1 ($K_i = 1.7$ nM) and PP-2A ($K_i = 0.04$ nM). Has no effect on protein kinases, making it useful for reducing the effect of contaminating phosphatases in protein kinase assays. It is not cell-permeant, but can enter hepatocytes via the multispecific organic anion transporter.

Soluble in: DMSO, ethanol, methanol

Stock concentration: 1 mM (store at –20°C)

Working concentration: 1 to 5 μM in hepatic cells; 10 μM in vitro

Duration of incubation: 15 min to 1 hr

MG-132

A potent, reversible and cell-permeant proteasome inhibitor. Reduces the degradation of ubiquitin-conjugated proteins by the 26S complex without affecting its ATPase or isopeptidase activities. Has been used to implicate the proteasome in the breakdown of membrane proteins, including the CFTR, within the ER (see also lactacystin). Inhibits NFκB activation.

Soluble in: DMSO

Stock concentration: 10 to 100 mM (store at –20°C)

Working concentration: 20 to 200 μM

Duration of incubation: 30 min to 24 hr

ML-7 (MLCK inhibitor)

Potent, cell-permeant, and selective inhibitor of myosin light chain kinase ($K_i = 300$ nM). Inhibits protein kinase A ($K_i = 21$ μM) and protein kinase C ($K_i = 42$ μM) at much higher concentrations.

Soluble in: DMSO, ethanol, water

Stock concentration: 100 to 500 mM (store at 4°C protected from light)

Working concentration: 10 to 50 μM

Duration of incubation: 15 min to 1 hr

See Table 1 (Selectivity of Selected Kinase Inhibitors) in the Biomol Catalog and Handbook, 5th ed.

L-Mimosine

An inhibitor of DNA replication that may act by preventing the formation of replication forks. L-mimosine blocks the camptothecin-induced apoptosis of PC12 cells, whereas aphidicolin does not.

Soluble in: Water

Stock concentration: 10 mM (store at room temperature)

Working concentration: 25 to 400 μM

Duration of incubation: 2 to 24 hr

Monensin

Polyether antibiotic that functions as a Na⁺ ionophore. Forms stable complexes with monovalent cations that are able to cross cell membranes. Inhibits glycoprotein secretion by blocking transport through the Golgi. Neutralizes acidic endomembrane compartments. Reduces sphingomyelinase activity.

Soluble in: DMSO, methanol

Stock concentration: 2 to 30 mM (store at 4°C)

Working concentration: 1 to 30 μM

Duration of incubation: Effective within seconds; use up to 3 hr.

Nigericin

Dual antiporter ionophore that acts as a K⁺/H⁺ exchanger. Stimulates Ca²⁺ release from mitochondrial stores by disruption of membrane potential. Allows adjustment of cytoplasmic pH (when combined with K⁺ ionophore such as valinomycin).

Soluble in: Ethanol

Stock concentration: 1 mg/ml (store at 4°C)

Working concentration: 1 to 10 μM

Duration of incubation: Effective within 2 to 5 min

Nocodazole

Has specific antimicrotubular activity for mammalian cells in culture. Promotes microtubule depolymerization. Nanomolar concentrations alter microtubule dynamics and interfere with fibroblast locomotion without affecting polymer levels. Arrests cells in mitosis.

Soluble in: DMSO

Stock concentration: 10 to 30 mM (store at room temperature)

Working concentration: 50 nM (low concentrations) to 30 μM (for effective microtubule depolymerization)

Duration of incubation: For rapid depolymerization of microtubules, preincubate cells on ice with nocodazole for 15 min; use up to 24 hr.

Okadaic acid

Potent inhibitor of protein phosphatases, especially the PP-1 class ($K_i = 10\text{--}15$ nM) and PP-2A class ($K_i = 0.1$ nM), in numerous cell types. Does not affect the activity of acid or alkaline tyrosine phosphatases. It mimics the effects of insulin, enhances neurotransmitter release, causes vasodilation, and is a potent tumor promoter. Induces dispersal of the Golgi apparatus. Okadaic acid is a useful tool for studying cellular processes regulated by serine/threonine phosphorylation.

Soluble in: DMSO, ethanol, methanol

Stock concentration: 1 mM (store at -20°C protected from light)

Working concentration: 50 to 200 nM

Duration of incubation: 15 min to 2 hr

May cause cell rounding. See Technical Note #18 from Alexis Biochemicals for additional information.

Olomoucine

Adenine derivative that acts as a competitive inhibitor for ATP binding and inhibits p34^{cdc2}/cyclin B ($K_i = 7$ μM) as well as several other CDKs at low concentrations. Does not significantly affect the activity of other protein kinases at 1 mM. Inhibits DNA synthesis in IL-2 stimulated T lymphocytes. Also used to synchronize cells in G₁. Can affect microtubule dynamics at higher concentrations.

Soluble in: DMSO

Stock concentration: 100 mM (store in aliquots at -20°C)

Working concentration: Usually 10 μM (up to 100 μM)

continued

Duration of incubation: 15 min to 24 hr, depending on process studied.

See Technical Note #25 from Alexis Biochemicals for additional information.

Ouabain

Selective Na⁺/K⁺-ATPase inhibitor. Causes net influx of Ca²⁺; also initiates the rapid protein kinase C-dependent inductions of early-response genes.

Soluble in: Water

Stock concentration: 100 mM (store at –20°C protected from light)

Working concentration: 1 to 100 μM

Duration of incubation: 30 min to 24 hr

PDMP

Useful tool for studying the effects of cellular glycosphingolipid depletion. Blocks ceramide glucosylation by inhibiting UDP-glucose:ceramide glucosyltransferase (glucosylceramide synthetase). Has antitumor activity; arrests 3T3 cells at both G₁/S and G₂/M. Prevents sensitization of A431 cells to Shiga toxin. Slows the rate of both anterograde vesicular traffic and endocytosis in CHO and BHK-21 cells. Redistributes *cis*-Golgi proteins to the ER.

Soluble in: Ethanol

Stock concentration: 10 to 100 mM (store at 4°C)

Working concentration: 20 to 100 μM

Duration of incubation: 1 to 18 hr, depending on process studied

PD 98059 (MEK Inhibitor)

Potent and selective inhibitor of MAP kinase kinase (MEK or MAPK/ERK kinase). Blocks the activity of MEK, thereby inhibiting the phosphorylation and activation of MAP kinase. Inhibits cell growth and reverses the phenotype of *ras*-transformed 3T3 mouse fibroblasts and rat kidney cells. Inhibits Golgi reassembly in vitro. Cell-permeant.

Soluble in: DMSO

Stock concentration: 50 mM (store at –20°C protected from light)

Working concentration: 10 to 50 μM

Duration of incubation: 30 min to 2 hr

PMSF (phenylmethanesulfonyl fluoride)

Inhibits serine proteases like chymotrypsin, trypsin, and thrombin, as well as acetylcholinesterase and the cysteine protease papain (reversible by DTT treatment). PMSF inhibits serine proteases by sulfonating serine residues at the active site. Does not inhibit metalloproteases, most cysteine proteases, or aspartic proteases.

Soluble in: Anhydrous isopropanol at 35 mg/ml with heating, resulting in a clear to very slightly hazy, colorless to faint yellow solution, or in anhydrous (100%, not 95%) ethanol

Stock concentration: 17 mg/ml (store at room temperature)

Working concentration: 17 to 170 μg/ml

Duration of incubation: 15 min to 1 hr

PMSF is very unstable in the presence of water. The half-life of aqueous PMSF at 25°C at pH 7.0, 7.5, and 8.0 is 110, 55, and 35 min, respectively.

Pepstatin A

Inhibitor of aspartyl proteases, including pepsin, renin, cathepsin D, and HIV-1 protease. Inhibits degradation of ApoB in rat hepatocytes; inhibits cytokine-induced programmed cell death. Accelerates amyloid fibril formation in mice.

Soluble in: DMSO

Stock concentration: 10 to 35 mM (store at –20°C)

Working concentration: 50 to 100 μM in cells

Duration of incubation: 30 min to 2 hr

Pertussis toxin

Protein endotoxin that catalyzes ADP-ribosylation of GDP-bound α subunits of the G proteins G_i , G_o , and G_t . Uncouples G proteins from receptors, thereby keeping the G protein in the inactive state. Used in the study of adenylate cyclase regulation and the role of G_i proteins. Consists of an enzymatically active A protomer subunit (S-1) which possesses both NAD^+ glycohydrolase and ADP-ribosylation activities and a B oligomer subunit (S-2, S-3, S-4, and S-5) that is responsible for cell surface attachment.

Stock solution: Reconstitute commercial preparation in water. Generally, commercially available pertussis toxin (Alexis Biochemicals, Biomol, Calbiochem) contains 50 μg of protein in 10 mM sodium phosphate buffer, pH 7.0/50 mM sodium chloride after being resuspended in 0.5 ml water. It is an insoluble protein that should be shaken gently before use. Store stock solutions at 4°C. Do not freeze.

Working concentration: 50 to 100 ng/ml

Duration of incubation: 2 to 24 hr

Phenylarsine oxide (PAO)

A cell-permeant phosphotyrosine phosphatase inhibitor ($K_i = 18 \mu\text{M}$). Induces a dose-dependent increase in the free Ca^{2+} intracellular concentration in rat peritoneal macrophages, human foreskin fibroblasts, and cultured human endothelial cells, without affecting intracellular stores. Inhibits insulin activation of phosphatidylinositol 3'-kinase. Dithiol cross-linking agent.

Soluble in: DMSO and chloroform

Stock concentration: 50 mM (store at room temperature)

Working concentration: 10 to 50 μM

Duration of incubation: Anywhere from 15 sec to 2 hr

Phorbol esters

An example is phorbol myristate acetate (PMA). Extremely potent tumor promoters. Activate protein kinase C by mimicking diacylglycerols (DAGs), causing a wide range of effects in cells and tissues.

Soluble in: DMSO

Stock concentration: 1 mM (store at -20°C)

Working concentration: 50 nM to 3 μM

Duration of incubation: Cells can be incubated anywhere from 5 min to 48 hr; stable in cells; results in long-term activation of PKC. However, long-term treatment may cause down-regulation of certain PKC subtypes.

See Technical Notes #13 and #14 from Alexis Biochemicals for additional information.

Piceatannol

At low concentrations, inhibits the receptor-mediated activation of the protein tyrosine kinase Syk as compared to the Src family in mast cells and B cells. Inhibits Fc ϵ R1-mediated signaling in RBL-2H3 cells.

Soluble in: DMSO, ethanol

Stock concentration: 10 to 50 mg/ml (store at 4°C protected from light)

Working concentration: 10 to 30 $\mu\text{g}/\text{ml}$

Duration of incubation: 1 hr

Puromycin

Protein synthesis inhibitor. Causes premature release of nascent polypeptide chains by its addition to the growing chain end; structural analog of aminoacyl-tRNA.

Soluble in: Water

Stock concentration: 100 mM (store at -20°C)

Working concentration: 10 to 100 μM

Duration of incubation: 5 min to 1 hr

Rapamycin

Member of a family of macrolide immunosuppressants that binds to and inhibits the peptidylproline *cis-trans* isomerase (PPIase) activity of the immunophilin FKBP12; effectors include a large protein termed FRAP (FKBP12 rapamycin-associated protein). FKBP12-rapamycin binds to but does not inhibit the activity of the Ca^{2+} /calmodulin-dependent serine/threonine phosphatase calcineurin. Blocks signaling, leading to the activation of p70 S6 kinase.

Soluble in: DMSO, methanol, ethanol

Stock concentration: 2 mM (store at -20°C)

Working concentration: 1 to 20 nM

Duration of incubation: 30 min to 1 hr

Sodium azide (NaN_3)

Inhibits mitochondrial ATPases; generally used to deplete ATP levels within cells (often in combination with 50 mM 2-deoxyglucose).

Soluble in: Water

Stock concentration: 1 M (store at room temperature)

Working concentration: 10 to 20 mM

Duration of incubation: 15 to 90 min

Sodium butyrate

A physiologically produced short-chain fatty acid that is generally used to increase expression of transfected genes with viral promoters (inhibits histone deacetylation). Blocks serum-stimulated DNA synthesis via a G_1 block. Induces apoptosis in colon carcinoma cell lines by a p53-independent process. Interferes with signal-transduction processes, including the release of Ca^{2+} from intracellular stores.

Soluble in: Water

Stock concentration: 5 M (store at -20°C)

Working concentration: 2 to 5 mM

Duration of incubation: Usually >12 hr

Sodium orthovanadate

Broad-spectrum inhibitor of protein tyrosine phosphatases. Also inhibits other ATPases, by mimicking the γ phosphate of ATP, including Na^+/K^+ ATPase, acid and alkaline phosphatases, and adenylate cyclase. Vanadate is also a strong inhibitor of lysosomal proteolysis in hepatocytes, the effect being ascribed to a direct inhibition of lysosomal enzymes. Stimulates pp60 (*v-src*) kinase activity in intact cells. It also stimulates amino acid transport activity in skeletal muscle, in a rapid and concentration-dependent manner.

Simple aqueous solutions of $(\text{VO}_4)^{3-}$ ion involve a dozen or more ionic species, both monomeric and oligomeric, whose abundances depend upon pH and $[\text{VO}_4]^{3-}$ concentration. See Fohr et al. (1989) for directions on preparing monomeric orthovanadate. It is unclear how readily vanadate ions enter cells (likely through anion transporters). At the concentration required for maximum inhibition, vanadate may have side effects that limit its application in cell culture. Can be combined with

continued

hydrogen peroxide (forming peroxyvanadate) to facilitate cell entry (combine 100 μl 0.1 M orthovanadate, 900 μl water, and 3.3 μl 30% H_2O_2 ; use 1:100 dilution of this on cells). However, the effect of hydrogen peroxide itself should be tested.

Soluble in: Water

Stock concentration: 100 mM (store at room temperature). To ensure the presence of monomers, the solution is heated to boiling until translucent and the pH is readjusted to 10. Solutions can be divided into aliquots, stored in plastic, and frozen. The orange color observed before boiling is due to decavanadate. At pH 10 this will slowly depolymerize over several hours to the colorless monovanadate. Vanadyl, metavanadate, orthovanadate, and decavanadate will interconvert in aqueous solution without suitable precautions (i.e., control of pH, oxidation state, complexing compounds, and concentration).

Working concentration: 200 μM to 2 mM

Duration of incubation: 15 min to 2 hr

Staurosporine

A potent cell-permeant inhibitor of protein kinases, most potently protein kinase C ($K_i = 0.7$ nM), protein kinase A ($K_i = 7$ nM), and myosin light chain kinase ($K_i = 1.3$ nM). Interaction is with the ATP binding site. Induces apoptosis, but not DNA fragmentation in MCF-7 cells. Arrests normal cells at the G_1 checkpoint.

Soluble in: DMSO, methanol

Stock concentration: 1 mM (store at -20°C protected from light)

Working concentration: 10 to 200 nM

Duration of incubation: 30 min to 24 hr, depending on the assay used

See Table I (*Selectivity of Selected Kinase Inhibitors*) in the *Biomol Catalog and Handbook*, 5th ed.

Taxol (Paclitaxel)

Antitumor and antileukemic agent. Promotes assembly of microtubules and inhibits microtubule disassembly. Bundles microtubules after several hours. Similar to nocodazole, taxol can inhibit microtubule dynamics without affecting overall polymer levels at nanomolar concentrations. Blocks cells at the G_2/M stage. Induces apoptosis in several cell types.

Soluble in: DMSO, methanol

Stock concentration: 20 mM (store at -20°C protected from light)

Working concentration: 10 nM to 20 μM

Duration of incubation: Taxol works rapidly to stabilize microtubules (within several minutes), although bundling takes several hours (this may be facilitated by first depolymerizing the polymer pool with ice treatment and/or washout of low levels of nocodazole)

Thapsigargin

Potent inhibitor of sarcoplasmic reticulum (SR)/endoplasmic reticulum (ER) Ca^{2+} -ATPases. Induces IP_3 -independent release of Ca^{2+} from the endoplasmic reticulum, causing an increase in intracellular Ca^{2+} . Depletion of Ca^{2+} from intracellular stores induces stress response and defects in protein folding and processing. Induces apoptosis in rat thymocytes and in human hepatoma cells. Irreversible.

Soluble in: DMSO, ethanol

Stock concentration: 1 mM (store at -20°C in aliquots, protect from light)

Working concentration: 20 nM to 1 μM

Duration of incubation: 15 sec to 2 min produces rise in intracellular Ca^{2+} ; longer incubations may be used, depending on effect to be analyzed

See Technical Note #15 from Alexis Biochemicals for additional information.

Trifluoperazine

Calmodulin antagonist. At 10 μ M, potentiates rise in cytosolic calcium induced by agonists. Antagonizes calmodulin at higher concentrations. Inhibits IL-2 production in activated Jurkat T cells. Structurally distinct from W-7.

Soluble in: Water (dihydrochloride salt)

Stock concentration: 10 mM (store at 4°C)

Working concentration: 10 to 50 μ M

Duration of incubation: 10 min to 3 hr

Tunicamycin

Nucleoside antibiotic that inhibits N-linked glycosylation, specifically by blocking the transfer of *N*-acetylglucosamine-1-phosphate from UDP-*N*-acetylglucosamine to dolichol monophosphate; has no effect on other glycosylation forms, such as Ser/Thr-linked oligosaccharides. Causes misfolding and retention of numerous glycoproteins in the endoplasmic reticulum, which induces synthesis of ER chaperones.

Soluble in: DMSO, ethanol

Stock concentration: 10 mg/ml (store at -20°C)

Working concentration: 1 to 10 μ g/ml

Duration of incubation: Cells can be treated from 1 to 24 hr

Tyrphostins

Large family of protein tyrosine kinase inhibitors. Inhibits receptors such as EGF receptor and PDGF receptor.

Soluble in: DMSO, ethanol

Stock concentration: 20 to 100 mM (store at -20°C protected from light)

Working concentration: 10 to 150 μ M

Duration of incubation: Anywhere from 1 to 48 hr

See Technical Note #22 from Alexis Biochemicals for additional information. Also see Table II (Selectivity of Tyrosine Protein Kinase Inhibitors) in the Biomol Catalog and Handbook, 5th ed.

Valinomycin

Potassium ionophore. Decreases ATP synthesis by decreasing membrane potential at mitochondrial membranes. Reported to inhibit NGF-induced neuronal differentiation. Used with nigericin to adjust cytoplasmic pH.

Soluble in: DMSO

Stock concentration: 1 mM (store at room temperature)

Working concentration: 1 to 20 μ M

Duration of incubation: Normally 15 min to 2 hr

Vinblastine

Vinca alkaloid; antitumor drug. Inhibitor of cell proliferation that acts by disrupting spindle microtubule function. Binds tubulin and suppresses microtubule dynamics. Depolymerizes microtubules at higher concentrations. Induces apoptosis in cultured hepatocytes and human lymphoma cells. Similar, but not identical effects observed with another vinca alkaloid, vincristine.

Soluble in: Methanol

Stock concentration: 20 mM (store at 4°C protected from light)

Working concentration: <10 nM (suppresses microtubule dynamics); 100 nM to 1 μ M (depolymerizes microtubules); >10 μ M (forms non-microtubule polymers)

Duration of incubation: 30 min to 24 hr

W-7

Member of a family of calmodulin antagonists, inhibiting calcium/calmodulin regulated enzyme activity. W-7 inhibits the Ca^{2+} /calmodulin-induced activation of myosin light chain kinase ($K_i = 51 \mu\text{M}$) and phosphodiesterase ($K_i = 28 \mu\text{M}$). Inhibits membrane tubulation in cells treated with brefeldin A.

Soluble in: Water

Stock concentration: 10 to 100 mM (store at 4°C protected from light)

Working concentration: 10 to 100 μM

Duration of incubation: 30 min to 2 hr

Wortmannin

Selective and potent phosphatidylinositol 3-kinase (PI 3-kinase) inhibitor; forms covalent associations with the kinases and is, therefore, irreversible. Abolishes PDGF-mediated $\text{Ins}(3,4,5)\text{P}_3$ formation in fibroblasts. Blocks the metabolic effects of insulin in isolated rat adipocytes without affecting the insulin receptor tyrosine kinase activity. Inhibits the formation of constitutive transport vesicles from the TGN. In human fetal undifferentiated cells, wortmannin induces morphological and functional endocrine differentiation.

Soluble in: DMSO

Stock concentration: 1 to 20 mM (store at -20°C in aliquots, protected from light)

Working concentration: 10 to 100 nM

Duration of incubation: 30 min to 4 hr

At nanomolar concentrations, wortmannin is specific to PI 3-kinases, while at higher concentrations other kinases are affected. Once diluted into aqueous solutions, wortmannin is less stable and should be made fresh daily.

LITERATURE CITED

- Fohr, K.J., Scott, J., Ahnert-Hilger, G., and Gratzl, M. 1989. Characterization of the inositol 1,4,5-trisphosphate-induced calcium release from permeabilized endocrine cells and its inhibition by decavanadate and *p*-hydroxymercuribenzoate. *Biochem. J.* 262:83-89.
- Jalink, K., van Corven, E.J., and Moolenaar, W.H. 1990. Lysophosphatidic acid, but not phosphatidic acid, is a potent Ca^{2+} -mobilizing stimulus for fibroblasts. Evidence for an extracellular site of action. *J. Biol. Chem.* 265:12232-12239.
- Seufferlein, T. and Rozengurt, E. 1994. Lysophosphatidic acid stimulates tyrosine phosphorylation of focal adhesion kinase, paxillin, and p130: Signaling pathways and cross-talk with platelet-derived growth factor. *J. Biol. Chem.* 269: 9345-9351.

Contributed by Nelson B. Cole
University of Pennsylvania
Philadelphia, Pennsylvania

COMMONLY USED REAGENTS AND EQUIPMENT

BUFFERS AND STOCK SOLUTIONS

This collection describes the preparation of buffers and reagents used in the manipulation of nucleic acids and proteins (see Table A.2.1). When preparing solutions, use deionized, distilled water and reagents of the highest grade available. Sterilization—by filtration through a 0.22- μm filter or by autoclaving—is recommended for most applications. Recipes for the following can be found elsewhere in the manual: culture media (UNIT 1.1), antibiotics (Table 1.4.1), lactose analogs (Table 1.4.2), and enzyme buffers (UNIT 3.4).

CAUTION: *Handle strong acids and bases carefully.*

Table A.2.1 Molarities and Specific Gravities of Concentrated Acids and Bases

Acid/base	Molecular weight	% by weight	Molarity (approx.)	1 M solution (ml/liter)	Specific gravity
<i>Acids</i>					
Acetic acid (glacial)	60.05	99.6	17.4	57.5	1.05
Formic acid	46.03	90	23.6	42.4	1.205
		98	25.9	38.5	1.22
Hydrochloric acid	36.46	36	11.6	85.9	1.18
Nitric acid	63.01	70	15.7	63.7	1.42
Perchloric acid	100.46	60	9.2	108.8	1.54
		72	12.2	82.1	1.70
Phosphoric acid	98.00	85	14.7	67.8	1.70
Sulfuric acid	98.07	98	18.3	54.5	1.835
<i>Bases</i>					
Ammonium hydroxide	35.0	28	14.8	67.6	0.90
Potassium hydroxide	56.11	45	11.6	82.2	1.447
Potassium hydroxide	56.11	50	13.4	74.6	1.51
Sodium hydroxide	40.0	50	19.1	52.4	1.53

Acid precipitation solution

1 M HCl

0.1 M sodium pyrophosphate

Nucleic acids can also be precipitated with a 10% (w/v) solution of trichloroacetic acid (TCA); however, this recipe is cheaper, easier to prepare, and just as efficient.

Ammonium acetate, 10 M

Dissolve 385.4 g ammonium acetate in 150 ml H₂O

Add H₂O to 500 ml

BBS (BES-buffered solution), 2×

50 mM *N,N*-bis(2-hydroxyethyl)-2-aminoethanesulfonic acid (BES; Calbiochem)

280 mM NaCl

1.5 mM Na₂HPO₄, pH 6.95

800 ml H₂O

Adjust pH to 6.95 with room temperature 1 M NaOH

H₂O to 1 liter

continued

Contributed by David D. Moore

Current Protocols in Molecular Biology (1996) A.2.1-A..2.8

Copyright © 2000 by John Wiley & Sons, Inc.

Filter sterilize through a 0.45- μ m nitrocellulose filter (Nalgene)
Store in aliquots at -20°C (can be frozen and thawed repeatedly)

The pH of this solution is critical (pH 6.95 to 6.98). When a new batch of 2 \times BES buffer is prepared, its pH should be checked against a reference stock prepared (and tested) earlier.

CaCl₂, 1 M

147 g CaCl₂·2H₂O
H₂O to 1 liter

Denhardt solution, 100 \times

10 g Ficoll 400
10 g polyvinylpyrrolidone
10 g bovine serum albumin (Pentax Fraction V; Miles Laboratories)
H₂O to 500 ml
Filter sterilize and store at -20°C in 25-ml aliquots

Dithiothreitol (DTT), 1 M

Dissolve 15.45 g DTT in 100 ml H₂O
Store at -20°C

EDTA (ethylenediamine tetraacetic acid), 0.5 M (pH 8.0)

Dissolve 186.1 g Na₂EDTA·2H₂O in 700 ml H₂O
Adjust pH to 8.0 with 10 M NaOH (~50 ml)
Add H₂O to 1 liter

Ethidium bromide, 10 mg/ml

Dissolve 0.2 g ethidium bromide in 20 ml H₂O
Mix well and store at 4 $^{\circ}\text{C}$ in dark

CAUTION: *Ethidium bromide is a mutagen and must be handled carefully.*

HBSS (Hanks balanced salt solution)

5.4 mM KCl
0.3 mM Na₂HPO₄
0.4 mM KH₂PO₄
4.2 mM NaHCO₃
1.3 mM CaCl₂
0.5 mM MgCl₂
0.6 mM MgSO₄
137 mM NaCl
5.6 mM D-glucose
0.02% phenol red (optional)
Add H₂O to 1 liter and adjust pH to 7.4

HBSS can be purchased from Biofluids or Whittaker.

HBSS may be made or purchased without CaCl₂ and MgCl₂. These are optional components that usually have no effect on an experiment. In some cases, however, their presence may be detrimental to a procedure. Consult the individual protocol to see if the presence or absence of these components is recommended in the materials list.

HCl, 1 M

Mix in the following order:
913.8 ml H₂O
86.2 ml concentrated HCl

HeBS (HEPES-buffered saline) solution, 2×

16.4 g NaCl

11.9 g HEPES acid

0.21 g Na₂HPO₄

800 ml H₂O

Titrate to pH 7.05 with 5 M NaOH

Add H₂O to 1 liter

Filter sterilize through a 0.45- μ m nitrocellulose filter

Test for transfection efficiency and store at -20°C in 50-ml aliquots

An exact pH is extremely important for efficient transfection. The optimal pH range is 7.05 to 7.12.

KCl, 1 M

74.6 g KCl

H₂O to 1 liter

MgCl₂, 1 M

20.3 g MgCl₂·6H₂O

H₂O to 100 ml

MgSO₄, 1 M

24.6 g MgSO₄·7H₂O

H₂O to 100 ml

MOPS buffer

0.2 M MOPS [3-(*N*-morpholino)propanesulfonic acid], pH 7.0

0.5 M sodium acetate

0.01 M EDTA

Store in the dark and discard if it turns yellow.

NaCl, 5 M

292 g NaCl

H₂O to 1 liter

NaOH, 10 M

Dissolve 400 g NaOH in 450 ml H₂O

Add H₂O to 1 liter

PBS (phosphate-buffered saline)

10× stock solution, 1 liter:

80 g NaCl

2 g KCl

11.5 g Na₂HPO₄·7H₂O

2 g KH₂PO₄

Working solution, pH ~7.3:

137 mM NaCl

2.7 mM KCl

4.3 mM Na₂HPO₄·7H₂O

1.4 mM KH₂PO₄

Potassium acetate buffer, 0.1 M

Solution A: 11.55 ml glacial acetic acid/liter (0.2 M).

Solution B: 19.6 g potassium acetate (KC₂H₃O₂)/liter (0.2 M).

Referring to Table A.2.2 for desired pH, mix the indicated volumes of solutions A and B, then dilute with H₂O to 100 ml.

This may be made as a 5- or 10-fold concentrate by scaling up the amount of potassium acetate in the same volume. Acetate buffers show concentration-dependent pH changes, so check concentrate pH by diluting an aliquot to the final concentration.

To prepare buffers with pH intermediate between the points listed in Table A.2.2, prepare closest higher pH, then titrate with solution A.

Table A.2.2 Preparation of 0.1 M Sodium and Potassium Acetate Buffers^a

Desired pH	Solution A (ml)	Solution B (ml)
3.6	46.3	3.7
3.8	44.0	6.0
4.0	41.0	9.0
4.2	36.8	13.2
4.4	30.5	19.5
4.6	25.5	24.5
4.8	20.0	30.0
5.0	14.8	35.2
5.2	10.5	39.5
5.4	8.8	41.2
5.6	4.8	45.2

^aAdapted by permission from CRC (1975).**Potassium phosphate buffer, 0.1 M***Solution A:* 27.2 g KH₂PO₄ per liter (0.2 M).*Solution B:* 34.8 g K₂HPO₄ per liter (0.2 M).Referring to Table A.2.3 for desired pH, mix the indicated volumes of solutions A and B, then dilute with H₂O to 200 ml.*This may be made as a 5- or 10-fold concentrate by scaling up the amount of potassium phosphate in the same volume. Phosphate buffers show concentration-dependent pH changes, so check concentrate pH by diluting an aliquot to the final concentration.***Table A.2.3** Preparation of 0.1 M Sodium and Potassium Phosphate Buffers^a

Desired pH	Solution A (ml)	Solution B (ml)	Desired pH	Solution A (ml)	Solution B (ml)
5.7	93.5	6.5	6.9	45.0	55.0
5.8	92.0	8.0	7.0	39.0	61.0
5.9	90.0	10.0	7.1	33.0	67.0
6.0	87.7	12.3	7.2	28.0	72.0
6.1	85.0	15.0	7.3	23.0	77.0
6.2	81.5	18.5	7.4	19.0	81.0
6.3	77.5	22.5	7.5	16.0	84.0
6.4	73.5	26.5	7.6	13.0	87.0
6.5	68.5	31.5	7.7	10.5	90.5
6.6	62.5	37.5	7.8	8.5	91.5
6.7	56.5	43.5	7.9	7.0	93.0
6.8	51.0	49.0	8.0	5.3	94.7

^aAdapted by permission from CRC (1975).

SDS electrophoresis buffer, 5×

15.1 g Tris base
72.0 g glycine
5.0 g SDS
H₂O to 1000 ml
Dilute to 1× or 2× for working solution, as appropriate
Store up to 1 month at 0° to 4°C

Do not adjust the pH of the stock solution, as the solution is pH 8.3 when diluted.

SED (standard enzyme diluent)

20 mM Tris·Cl, pH 7.5
500 µg/ml bovine serum albumin (Pentax Fraction V)
10 mM 2-mercaptoethanol
Store up to 1 month at 4°C

Sodium acetate, 3 M

Dissolve 408 g sodium acetate·3H₂O in 800 ml H₂O
Add H₂O to 1 liter
Adjust pH to 4.8 or 5.2 (as desired) with 3 M acetic acid

Sodium acetate buffer, 0.1 M

Solution A: 11.55 ml glacial acetic acid/liter (0.2 M).
Solution B: 27.2 g sodium acetate (NaC₂H₃O₂·3H₂O)/liter (0.2 M).
Referring to Table A.2.2 for desired pH, mix the indicated volumes of solutions A and B, then dilute with H₂O to 100 ml. (See Potassium acetate buffer recipe for further details.)

Sodium phosphate buffer, 0.1 M

Solution A: 27.6 g NaH₂PO₄·H₂O per liter (0.2 M).
Solution B: 53.65 g Na₂HPO₄·7H₂O per liter (0.2 M).
Referring to Table A.2.3 for desired pH, mix the indicated volumes of solutions A and B, then dilute with H₂O to 200 ml. (See Potassium phosphate buffer recipe for further details.)

SSC (sodium chloride/sodium citrate), 20×

3 M NaCl (175 g/liter)
0.3 M Na₃citrate·2H₂O (88 g/liter)
Adjust pH to 7.0 with 1 M HCl

STE buffer

10 mM Tris·Cl, pH 7.5
10 mM NaCl
1 mM EDTA, pH 8.0

TAE (Tris/acetate/EDTA) electrophoresis buffer

<i>50× stock solution:</i>	<i>Working solution, pH ~8.5:</i>
242 g Tris base	40 mM Tris·acetate
57.1 ml glacial acetic acid	2 mM Na ₂ EDTA·2H ₂ O
37.2 g Na ₂ EDTA·2H ₂ O	
H ₂ O to 1 liter	

TBE (Tris/borate/EDTA) electrophoresis buffer

10× stock solution, 1 liter:
108 g Tris base (890 mM)
55 g boric acid (890 mM)
40 ml 0.5 M EDTA, pH 8.0 (see recipe; 20 mM)

TE (Tris/EDTA) buffer

10 mM Tris·Cl, pH 7.4, 7.5, or 8.0 (or other pH; see recipe)
1 mM EDTA, pH 8.0

TEA (triethanolamine) solution

50 mM triethanolamine, pH ~11.5
0.1% Triton X-100
0.15 M NaCl

Add Triton X-100 as a 10% stock sterilized by Millipore filtration and stored in the dark to prevent photooxidation (stock is stable 5 years at room temperature).

TEN (Tris/EDTA/NaCl) solution

40 mM Tris·Cl, pH 7.5
1 mM EDTA, pH 8.0
150 mM NaCl
Store up to 6 months at room temperature

TM buffer, 10×

100 mM Tris·Cl, pH 8.0
100 mM MgCl₂

Tris-buffered saline (TBS)

100 mM Tris·Cl, pH 7.5
0.9% (150 mM) NaCl
Store up to several months at 4°C

Tris·Cl [tris(hydroxymethyl)aminomethane], 1 M

Dissolve 121 g Tris base in 800 ml H₂O
Adjust to desired pH with concentrated HCl
Mix and add H₂O to 1 liter

Approximately 70 ml of HCl is needed to achieve a pH 7.4 solution, and approximately 42 ml for a solution that is pH 8.0.

IMPORTANT NOTE: *The pH of Tris buffers changes significantly with temperature, decreasing approximately 0.028 pH units per 1°C. Tris-buffered solutions should be adjusted to the desired pH at the temperature at which they will be used. Because the pK_a of Tris is 8.08, Tris should not be used as a buffer below pH ~7.2 or above pH ~9.0.*

TTBS (Tween 20/TBS)

0.1% Tween 20 in Tris-buffered saline (TBS; see recipe)
Store up to several months at 4°C

EQUIPMENT

Special equipment is also itemized in the materials list of each protocol. We have not attempted to list all items required for each procedure, but rather have noted those items that might not be readily available in the laboratory or that require special preparation. Listed below are standard pieces of equipment in the modern molecular biology laboratory, i.e., items used extensively in this manual and thus not included in the individual materials lists.

Autoclave

Balances, analytical and preparative

Bench protectors, plastic-backed (including “blue pads”)

Centrifuges a low-speed (20,000 rpm) refrigerated centrifuge and an ultracentrifuge (20,000 to 80,000 rpm) are required for many procedures. Vertical ultracentrifuge rotors are very convenient for preparing plasmid DNA. At least one microcentrifuge that holds stan-

dard 1.5-ml microcentrifuge tubes is essential. It is also useful to have a large-capacity, low-speed centrifuge (such as the Beckman J-6M) for spinning down large bacterial cultures and a tabletop swinging-bucket centrifuge with adapters for spinning 96-well microtiter plates.

NOTE: *Centrifuge speeds are provided as $\times g$ or as rpm (with example rotor models) throughout the manual. Readers should consult the nomograms in APPENDIX 1G to convert these speeds to their own rotor models.*

Computer (PC or Macintosh) and printer

Darkroom and developing tanks or X-Omat automatic X-ray film developer.

Dry ice

Filtration apparatus for collecting acid precipitates on nitrocellulose filters or membrane.

Flasks, glass (e.g., Erlenmeyer, Florence)

Fraction collector

Freezers and refrigerators for 4°, -20°, and -70°C incubation and storage.

Fume hood

Geiger counter

Gel dryer

Gel electrophoresis equipment at least one full-size horizontal apparatus and one horizontal minigel apparatus, two sequencing gel setups for each person engaged in large-scale sequencing projects, one vertical gel apparatus for polyacrylamide protein gels, and specialized equipment for two-dimensional protein gels as required.

Heating blocks thermostat-controlled metal heating blocks that hold test tubes and/or microcentrifuge tubes are very convenient for carrying out enzymatic reactions.

Ice maker

Incubator (37°C) for growing bacteria. We recommend an incubator large enough to hold a “tissue culture” roller drum that can be used to grow 5-ml cultures in standard 18 \times 150 mm test tubes. A convenient and durable tube roller is made by New Brunswick Scientific.

Incubator/shaker(s) an enclosed shaker (such as the New Brunswick Controlled Environment Incubator Shaker) that can spin 4-liter flasks is essential for growing 1-liter *E. coli* cultures. A rotary shaking water bath (New Brunswick R76) is useful for growing smaller cultures in flasks.

Light box for viewing autoradiograms.

Liquid nitrogen

Magnetic stirrers (with heater is useful).

Microcentrifuge, Eppendorf-type, maximum speed 12,000 to 14,000 rpm

Microcentrifuge tubes, 1.5-ml

Microwave oven to melt agar and agarose.

Mortar and pestle

Paper cutter large size, for 46 cm \times 57 cm Whatman sheets.

Paper towels

Parafilm

Pasteur pipets and bulbs

PCR machine (Perkin-Elmer or MJ Research)

pH meter

pH paper

Pipettors that use disposable tips and dispense 1 to 1000 μ l. It is best to have a set for each full-time researcher.

Plastic wrap (e.g., Saran Wrap)

Polaroid camera and UV transilluminator for taking photographs of stained gels.

Policemen, rubber or plastic

Power supplies 300-volt power supplies are sufficient for agarose gels; 2000-volt power supply required for DNA sequencing.

Racks, test tube

Radiation shield (Lucite or Plexiglas)

Radioactive ink

Radioactive waste containers, for liquid and solid waste

Refrigerator, 4°C

Safety glasses

Scalpels and blades

Scintillation counter

Scissors

Seal-A-Meal bag sealer or equivalent

Shakers, orbital and platform, room temperature or 37°C

Spectrophotometer UV and visible

Speedvac evaporator (Savant)

Tissue culture equipment CO₂ incubator, phase contrast microscope, liquid nitrogen storage container, and laminar flow hood.

UV cross-linker (e.g., Stratalinker, Stratagene)

UV light sources, long- and short-wave

UV transilluminator

Vacuum desiccator/lyophilizer

Vacuum oven

Vortex mixers

Water baths at least two with 80°C capacity

Water purification equipment or glass distillation apparatus to purify all water used in molecular biology experiments.

X-ray film cassettes and intensifying screens

LITERATURE CITED

Chemical Rubber Company (CRC). 1975. CRC Handbook of Biochemistry and Molecular Biology, Physical and Chemical Data, 3rd ed., Vol. 1. CRC Press, Boca Raton, Fla.

Contributed by David D. Moore
Massachusetts General Hospital & Harvard Medical School
Boston, Massachusetts

COMMONLY USED TECHNIQUES IN BIOCHEMISTRY AND MOLECULAR BIOLOGY

APPENDIX 3

The protocols in this appendix have been included because of their universal application in molecular biology experiments. Other commonly used protocols can be found in early chapters of this manual, for example, growing bacteria (*UNITS 1.2 & 1.3*), plasmid preps (*UNITS 1.6 & 1.7*), CsCl/ethidium bromide gradient purification (*UNITS 1.7 & 1.13*), phenol extraction and alcohol precipitation (*UNIT 2.1*), acid precipitation (*UNIT 3.4*), agarose gel electrophoresis (*UNIT 2.5*), restriction digests (*UNIT 3.1*), and radiolabeling DNA (*UNIT 3.5*). See individual chapter tables of contents for complete listings of protocols.

Detection and Quantitation of Radiolabeled Proteins and DNA in Gels and Blots

APPENDIX 3A

This unit presents procedures for visualizing and quantitating radiolabeled proteins or DNA separated by polyacrylamide gel electrophoresis (PAGE; *UNITS 2.5A & 2.7; UNIT 10.2*) or affixed to filter membranes. Autoradiography (see Basic Protocol) is the most common method by which this is accomplished, and X-ray film is the traditional recording medium. The use of autoradiography with gels requires that the gel be dried prior to being placed in contact with the film (see Support Protocol 1). The decay of radioactive materials within the dried gel or filter leaves an image on the film that reflects its distribution in the sample. Film images can be analyzed by densitometry (see Support Protocol 4) to obtain a relative quantitative measure of the amount of radioactivity in the sample.

The use of X-ray films for autoradiography, however, suffers from two drawbacks: lack of sensitivity and a limited linear range over which the image density reflects the amount of radioactivity. Lack of sensitivity can be overcome by fluorography (see Alternate Protocol 1) or by the use of intensifying screens (see Support Protocol 2), both of which enhance the radioactive signal. Ensuring that the exposure is within a linear range requires careful controls; film is often preflashed (see Support Protocol 3) to increase the linear measurement range for weakly radioactive samples, and it is important to ensure that the film not be saturated to attain strong radioactive signals. Sensitivity and linear ranges of measurement can be greatly extended by using a phosphor imaging system (see Alternate Protocol 2). Phosphor imaging also makes it much faster and easier to quantify radioactive samples.

To enhance radioactive signals, solid-state scintillation is frequently employed to convert the energy released by radioactive molecules to visible light. This is accomplished in several different ways. In fluorography (see Alternate Protocol 1), organic scintillants are incorporated into the sample to increase the proportion of emitted energy detected from low-energy β particles (e.g., from ^3H , ^{14}C , and ^{35}S). Another method uses high-density fluorescent “intensifying screens” (see Support Protocol 2), which are placed next to the sample and used to capture the excess energy of γ rays (e.g., those produced by ^{125}I) and high-energy β particles (e.g., from ^{32}P).

CAUTION: When working with radioactivity, take appropriate precautions to avoid contamination of the experimenter and the surroundings. Carry out the experiment and dispose of wastes in appropriately designated areas, following the guidelines provided by your local radiation safety officer (also see *APPENDIX 1F*).

Commonly Used
Techniques

A.3A.1

Contributed by Daniel Voytas and Ning Ke

Current Protocols in Molecular Biology (1999) A.3A.1-A.3A.10

Copyright © 1999 by John Wiley & Sons, Inc.

Supplement 48

AUTORADIOGRAPHY

In autoradiography, X-ray film is used to visualize and quantitate radioactive molecules that have been electrophoresed through agarose or polyacrylamide gels (*UNITS 2.5A & 2.7; UNIT 10.2*), hybridized to filters (e.g., immunoblots; *UNIT 6.2*), or chromatographed through paper or thin-layer plates. A photon of light or a β particle or γ ray released from a radioactive molecule “activate” silver bromide crystals on the film emulsion. This renders them capable of being reduced through the developing process to form silver metal (a “grain”). The silver grains on the film form the image.

The choice of film is critical for autoradiography. Double-coated films (e.g., Kodak X-Omat AR and Fuji RX) contain two emulsion layers on either side of a polyester support and are most commonly used for autoradiography (Laskey and Mills, 1977). Double-coated films are ideal for detecting the high-energy β particles emitted by ^{32}P and ^{125}I , since they can penetrate the polyester support and expose both emulsion layers. These films are normally used with calcium tungstate (CaWO_4) intensifying screens at reduced temperature (-70°C); they are highly sensitive to the blue light emitted by these screens. The green-light-sensitive BioMax MS film (Kodak) is a double-coated film spectrally matched to the blue- and green-light-emitting BioMax MS intensifying screen. The BioMax MS film/BioMax MS intensifying screen system normally gives greatest sensitivity to ^{32}P (four times greater than that of X-Omat AR film with CaWO_4 screens).

Single-coated films, containing one emulsion layer (e.g., Kodak BioMax MR), are optimized for direct-exposure techniques with medium-energy radioisotopes (e.g., ^{14}C , ^{35}S , and ^{33}P , but not ^3H). The majority of β particles emitted by these isotopes cannot pass through the polyester support of double-coated films, and therefore the emulsion layer on the other side of the film is useless. Even though direct exposure with single-coated films gives better clarity for medium-energy isotopes, single-coated films often require longer exposure times. Fluorography, therefore, is often used to enhance sensitivity. The blue-light-sensitive double-coated X-Omat AR film is generally used for fluorography with 2,5-diphenyloxazole (PPO) which emits at 388 nm; sodium salicylate, which emits at 420 nm; and commercial fluorographic solutions and sprays (e.g., from Amersham and DuPont NEN), which emit light in the blue end of the spectrum.

Materials

Dried gel (see Support Protocol 1) or filter (e.g., from immunoblotting; *UNIT 10.8*)

Developer: Kodak developer and replenisher, prepared according to the manufacturer’s instructions, 18° to 20°C

Fixer: Kodak fixer and replenisher, prepared according to the manufacturer’s instructions, 18° to 20°C

Metal film cassette or paper film cassette with particle-board supports and metal binder clips

Plastic wrap (e.g., Saran Wrap)

X-ray film (optionally preflashed; see Support Protocol 3)

Trays to hold film processing solutions

Clips for hanging film

1. In a darkroom illuminated with a safelight, place the sample (e.g., dried gel or filter) in the film cassette. Cover the sample with plastic wrap to prevent it from sticking to the film and contaminating the cassette with radioactivity.

The safelight should be a bulb of <15 W that is equipped with a Kodak GBX-2 red filter (or equivalent). The safelight should be located 1.2 meters from the film.

Fluorescent glow-in-the-dark ink (available at craft stores) is a convenient way to mark samples exposed to film. The ink can be spread on adhesive labels, which in turn are placed

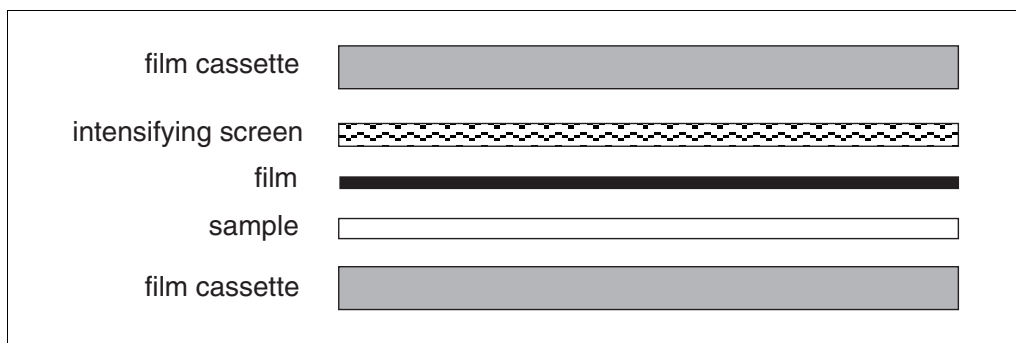


Figure A.3A.1 Autoradiography setup: intensifying screen, film, and sample in film cassette.

on the plastic wrap around the edge of the sample. If exposed to light prior to autoradiography, the ink will fluoresce and expose the film, making it possible to orient the film image on the dried gel after developing.

2. Place a sheet of X-ray film on top of the sample, then close and secure the film cassette (see Fig. A.3A.1).

If preflashed film is used for direct exposure (see Support Protocol 3), the exposed side should face the sample. Preflashed film should be used if sample is weakly radioactive or if quantitation of the radioactivity is desired. For single-coated film, the emulsion layer should face the sample.

If a paper cassette is used, particle-board supports the same size as the cassette should be placed on either side and secured with the metal binder clips. This will ensure that the sample and film are held in contact and do not shift during exposure.

3. Expose the film for the desired length of time and at the appropriate temperature.

Time of exposure will depend on the strength of the radioactivity in the sample and, in most cases, will have to be determined empirically by making multiple exposures for different lengths of time. To help estimate exposure time, a Geiger counter can often be used to detect the relative amount of radioactivity in a sample. With experience, this can help alleviate the trial and error often associated with obtaining the optimum exposure. Time of exposure and use of internal controls are particularly important if quantitative comparisons between experiments are desired.

4. After exposure, return cassette to the darkroom and remove film for developing.

If the film was exposed at -70°C , allow the cassette to come to room temperature before developing. This will avoid static discharge, which can cause black dots or stripes on the autoradiogram.

Automated film developers are also available and can be used to develop the film.

5. Immerse the film for 5 min in 18° to 20°C developer, then wash 1 min in running water at room temperature.

Shorter periods of time in developer will yield a lighter image. The amount of time in developer, therefore, can be used to roughly control intensity of the image.

6. Immerse the film for 5 min in 18° to 20°C fixer, then wash for 15 min in running water.

7. Hang the film to dry.

The orientation of the film with respect to the gel can be determined by the images of the fluorescent markers.

PREPARING GELS FOR AUTORADIOGRAPHY

Polyacrylamide gels containing radiolabeled proteins or DNA should be dried before exposure to film. Protein gels should be fixed prior to drying to prevent the gel from sticking to the film, improve the sharpness of the image, and increase sensitivity slightly. However, if the specific activity of the sample is high or the detection method is sensitive (e.g., where a phosphor imager is used; see Alternate Protocol 2), then fixing and drying the gel may not be necessary. Gel dryers are available from a number of manufacturers (e.g., Bio-Rad; most use heat and a vacuum to accelerate the drying process).

Materials

Polyacrylamide gel

Fixing solution: 10% glacial acetic acid/20% methanol (v/v) in water *or* 3% glycerol/10% glacial acetic acid/20% methanol (v/v/v) in water

Protein (for gels with $\geq 15\%$ acrylamide or thicker than 1.5 mm: 3% glycerol/10% glacial acetic acid/20% methanol (v/v/v) in water

Glass dish

Rotary shaker

Filter paper (Whatman 3MM) in sheets at least 1 to 2 cm larger than gel

UV-transparent plastic wrap (e.g., Saran Wrap)

Gel dryer with vacuum pump

1. After electrophoresis, remove the gel and the supporting glass plates from the electrophoresis apparatus and place in a glass dish. Carefully remove the upper glass plate by gently prying apart the corners with a metal spatula. Make a notch in the upper right-hand corner of the gel for orientation.

Since the gel contains radiolabeled protein or DNA, be sure to follow the necessary guidelines for handling radioactivity (see APPENDIX 1F). Everything that comes in contact with the gel is potentially radioactive.

- 2a. *For DNA gels:* Proceed to step 4.

- 2b. *For protein gels with $< 15\%$ acrylamide and < 1.5 mm thick:* Place the glass dish in a fume hood and pour enough fixing solution (without glycerol) into the dish to cover the gel. Place the dish on a rotary shaker and gently rotate until 5 min after all of the blue color from the bromphenol blue in the sample buffer (if used) has disappeared (~30 min total).

The bromphenol blue typically used in PAGE sample buffer will turn yellow as the acidic fixing solution diffuses into the gel.

During fixing, the gel will typically slide off the lower glass plate, which can be removed.

- 2c. *For protein gels with $\geq 15\%$ acrylamide or > 1.5 mm thick:* Fix gel as in step 2b, but soak 1 hr in alternative fixing solution containing 3% glycerol.

The glycerol should help prevent cracking during drying.

3. Pour off the fixing solution and rinse the gel for a few minutes with deionized water. Drain.

CAUTION: Remember that solutions which come in contact with the gel are potentially radioactive.

4. Position the gel in the center of the glass dish. Place a sheet of Whatman 3MM filter paper, at least 1 to 2 cm larger than the gel, over the gel.

The gel will stick to the filter paper, which will allow it to be lifted and turned over with the gel side facing up.

5. Cover gel with plastic wrap. Smooth the wrap with a piece of tissue paper to remove any air bubbles or wrinkles.
6. Place a piece of filter paper on the gel support of the gel dryer to prevent contamination of the dryer by radioactivity.
7. Place the filter paper/gel/plastic wrap sandwich on the filter paper in the gel dryer with the plastic sheet facing up.
8. Position the rubber sealing gasket of the gel dryer over the gel. Set the appropriate heat setting on the gel dryer (normally 80°C; 60°C if the gel contains a fluor). Apply the vacuum and allow the gel to dry (typically 2 hr for a gel of 1 mm thickness).

Removing the gel before it is completely dry can lead to cracking; it is therefore not a good idea to rush the drying process. A rough indication of whether the gel is dry can be obtained by feeling the gel under the sealing gasket. If the gel is dry, it should be warm over the entire surface.
9. Remove gel from dryer and proceed with autoradiography (see Basic Protocol).

USE OF INTENSIFYING SCREENS

Intensifying screens enhance the film image generated by radioactive molecules (Laskey and Mills, 1977; Laskey, 1980). They are used strictly in conjunction with strong β -emitting isotopes such as ^{32}P or γ -emitting isotopes such as ^{125}I . Emissions from these forms of radiation will frequently pass completely through a film, but they can be absorbed by an intensifying screen that fluoresces and exposes the film with multiple photons of light. While an intensifying screen will substantially enhance the film image as compared with direct exposure (Table A.3A.1), some loss of image resolution will occur due to light scatter. Intensifying screens are distributed by most laboratory supply companies (e.g., Fisher, Sigma, and Kodak).

As shown in Figure A.3A.1, the film should be placed between the sample and the intensifying screen. Preflashed film (see Support Protocol 3) should be used if the sample is weakly radioactive or if quantitation of the radioactivity is desired. The preflashed side of the film should be placed adjacent to the intensifying screen. For very weakly radioactive samples, a second screen can be placed on the other side of the radioactive sample (i.e., screen, then sample, then film, then screen), but this causes further light-scatter and loss of resolution. Also, the sample and sample support must be sufficiently transparent to allow light from the second screen to reach the film. The film should be exposed at -70°C to stabilize the silver bromide crystals activated by the radioactivity or the light emitted from the screen.

Table A.3A.1 Different Methods for Isotope Detection and Their Sensitivities^a

Isotope	Method ^b	Sensitivity ^c	Enhancement over direct autoradiography ^d
^{125}I	S	100	16
^{32}P	S	50	10.5
^{14}C	F	400	15
^{35}S	F	400	15
^3H	F	8000	>100

^aExposures conducted at -70°C using preexposed film.

^bS, intensifying screen; F, fluorography using PPO.

^cDefined as dpm/cm^2 required for detectable image ($A_{540} = 0.02$) in 24 hr.

^dDirect autoradiography for comparison was performed on Kodirex film (Laskey, 1980).

SUPPORT PROTOCOL 2

Commonly Used Techniques

A.3A.5

PREFLASHING (PREEXPOSING) FILM

Silver bromide crystals activated by light, β particles, or γ rays are highly unstable and quickly revert back to their stable form. The absorption of several photons increases their stability but does not ensure development; approximately five photons of light are required to obtain a 50% probability that any single silver bromide crystal will be developed during film processing. This inefficiency means that film images produced by very low levels of exposure will be disproportionately faint. Two measures can be taken to maximize efficiency and linearity of exposure at the low levels commonly encountered in ordinary use. First, the film should be preexposed to a hypersensitizing flash of light, which provides several photons per silver bromide crystal and stably activates them without providing enough exposure to cause them to become developed. This allows a linear relationship to be drawn between the film image and the amount of radioactivity in the sample. Second, film exposure should be conducted at low temperatures (-70°C) to slow the reversion of activated silver bromide crystals to their stable form (Laskey and Mills, 1975).

Film can be hypersensitized by exposure to a flash of light (<1 msec) provided by a photographic flash unit or a stroboscope before being placed onto the radioactive sample for exposure of the autoradiogram (Laskey and Mills, 1975, 1977). As the optimal light intensity required for preexposure varies with the type of film and the flash unit being used, the ideal exposure is best determined empirically as described below.

Materials

Stroboscope or flash unit (e.g., Auto 22 Electronic Flash from Vivitar or Sensitize Pre-Flash from Amersham Pharmacia Biotech)
Neutral-density filter (Kodak)
Orange filter (Wratten 22; Kodak)
X-ray film
Spectrophotometer

1. Cover the stroboscope or flash unit with the neutral-density and orange filters.

This serves to decrease the intensity of emitted light, particularly the blue wavelengths to which X-ray films are most sensitive. Filters are not required for the Amersham flash unit.

2. Place the film perpendicular to the light source at a distance of ≥ 50 cm to ensure uniform illumination.
3. Expose a series of test films for different flash lengths, then develop them (see Basic Protocol).

An uneven fog level on film can be remedied by placing a porous paper diffuser, such as a piece of Whatman no. 1 filter paper, between the film and the light source.

4. Cut the films into pieces that fit into a cuvette holder of a spectrophotometer and measure the absorbance at 540 nm.

Choose an exposure time that causes the absorbance of the preexposed film to increase by 0.15 with respect to film that was not preexposed.

FLUOROGRAPHY

Organic scintillants can be included in radioactive samples to obtain autoradiograms of weak β -emitting isotopes such as ^3H , ^{14}C , and ^{35}S . The scintillant fluoresces upon absorption of β particles from these isotopes, facilitating film exposure. Fluorographs of radioactive molecules in polyacrylamide gels have traditionally used the scintillant PPO (2,5-diphenyloxazole; Laskey and Mills, 1975). PPO, however, has largely been replaced with commercial scintillation formulations that reduce the amount of preparation time and are considerably safer to use. These scintillants (e.g., Enhance from NEN Life Science) come with complete instructions for their use. In addition, spray applicators are also available that can be used on filters or thin-layer plates. The expected levels of image enhancement obtained through fluorography are listed in Table A.3A.1. Sodium salicylate can also be used for fluorography as described below (Chamberlain, 1979). It yields levels of image enhancement comparable to those from organic scintillants, although it sometimes causes a more diffuse film image. The conditions should work for most standard sizes and thicknesses of gels.

CAUTION: Gloves should be worn at all times; sodium salicylate can elicit allergic reactions and is readily absorbed through the skin.

Materials

Polyacrylamide gel

1 M sodium salicylate, pH 5 to 7, freshly prepared

Additional reagents and equipment for fixing and drying gels (see Support Protocol 1)

1. If gel is acid-fixed, soak for 1 to 5 hr in ~ 20 vol water to prevent precipitation of salicylic acid from the sodium salicylate.
2. Soak gel 30 min in 10 vol of 1 M sodium salicylate, pH 5 to 7.

To prevent cracking of gels with $>15\%$ acrylamide or thicker than 1.5 mm, 2% (v/v) glycerol can be added to the 1 M sodium salicylate.

3. Dry the gel (see Support Protocol 1) and proceed with autoradiography (see Basic Protocol).

DENSITOMETRY

Film images obtained by autoradiographic methods can be quantified by densitometry. Densitometers work by comparing the intensity of light transmitted through a sample with the intensity of the incident light. The amount of light transmitted will be proportional to the amount of radioactivity in the gel, provided that the film has been properly preexposed (see Support Protocol 3). The linear range of correctly preexposed film is 0.1 to 1.0 absorbance units. However, if the preexposure is excessive—i.e., an increase of >0.2 absorbance units (A_{540}) treated film/untreated film—smaller amounts of radioactivity will produce disproportionately dense images. Autoradiograms that exceed an absorbance of 1.4 absorbance units (A_{540}) have saturated all available silver bromide crystals and also cannot be evaluated quantitatively.

Densitometers are available from several manufacturers (e.g., Molecular Dynamics, Bio-Rad, and UVP). Most models come with software that facilitates calculations and allows the user to define the region of the film to be measured. Procedures for the use of these machines vary and instructions are provided by the manufacturer. Densitometers are also available that measure light reflected from a sample. Reflectance densitometers are useful in instances where the sample medium is completely opaque—e.g., filters that have been probed using nonradioactive colorimetric detection assays.

PHOSPHOR IMAGING

Phosphor imaging screens can be used as an alternative to film for recording and quantifying autoradiographic images (Johnston et al., 1990). They can detect radioisotopes such as ^{32}P , ^{125}I , ^{14}C , ^{35}S , and ^3H . There are several advantages of phosphor imaging over film: (1) linear dynamic ranges are 5 orders of magnitude, compared to ~ 1.5 orders of magnitude for film (Fig. A.3A.2); (2) exposure times are 10 to 250 times faster than with film; (3) quantification is much easier and faster, and most imagers come with software to directly analyze data; (4) fluorography and gel drying are often unnecessary because of the sensitivity of phosphor imaging; and (5) phosphor screens can be reused indefinitely if handled carefully.

Phosphor imaging screens are composed of crystals of BaFBr:Eu^{+2} . When the screen is exposed to ionizing radiation such as α , β , or γ radiation, or wavelengths of light shorter than 380 nm, the electrons from Eu^{+2} are excited and then trapped in an "F-center" of the BaFBr^- complex; this results in the oxidation of Eu^{+2} to Eu^{+3} , which forms the latent image on the screen. After exposure, the latent image is released by scanning the screen with a laser (633 nm). During scanning, Eu^{3+} reverts back to Eu^{+2} , releasing a photon at 390 nm. The luminescence can then be collected and measured in relation to the position of the scanning laser beam. The result is a representation of the latent image on the storage phosphor imaging plates. The image can then be viewed on a video monitor and analyzed with the aid of appropriate software.

Some companies (e.g., Bio-Rad) offer different screens for use with different isotopes. They vary principally in the protective coating on the screen, which is optimized for low- or high-energy β particles or γ rays. No coating is typically used for weak β emitters such as tritium. More recently, screens have also been developed that measure chemiluminescence. Such screens are particularly valuable for use with many nonradioactive labeling protocols.

The protocol below is for the PhosphorImager system from Molecular Dynamics; other phosphor imaging systems are available from Bio-Rad, Imaging Research, and National Diagnostics.

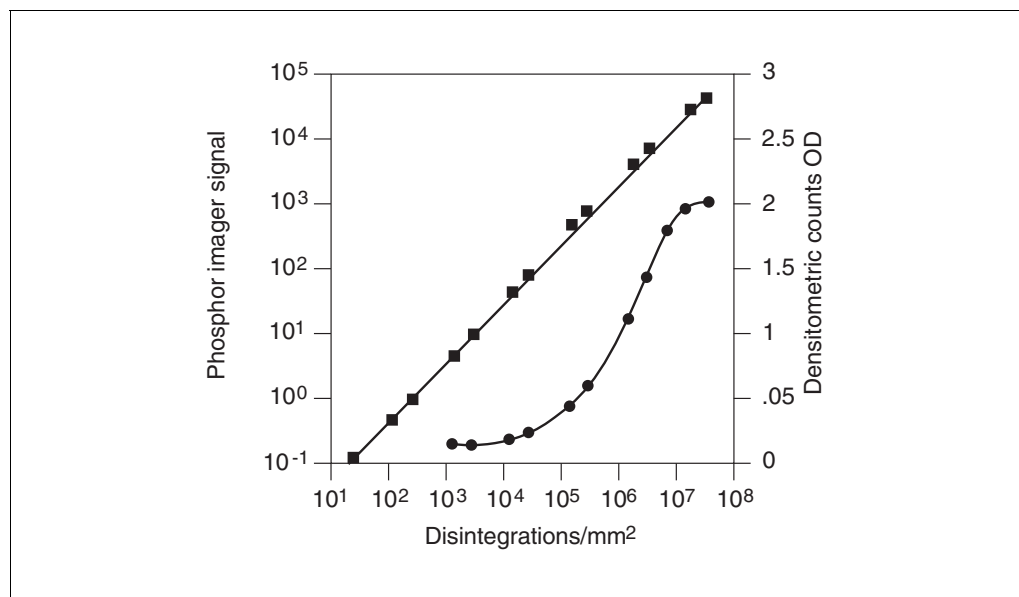


Figure A.3A.2 ^{32}P dilution series quantified on Model GS-525 phosphor imager (squares), compared to film (circles). Image courtesy of Bio-Rad, Hercules, Calif.

Materials

Gel or filter (e.g., from immunoblotting; *UNIT 10.8*)
PhosphorImager system (Molecular Dynamics) including:
ImageEraser light box
Exposure cassette with phosphor screen
Scanning software

1. Erase any latent image on the phosphor screen left by a previous user, or caused by background radiation, by exposure to visible light.

The PhosphorImager system comes with an extra-bright light box (ImageEraser) for this purpose. Standard laboratory light boxes may also be used.

2. Cover gel or filter with plastic wrap to protect the exposure cassette. Place wrapped gel or filter in the PhosphorImager cassette and close to begin exposure.

The gel does not have to be dried for this procedure. The phosphor screen is affixed to the lid of the cassette. Exposure times are typically one-tenth of the time required for film exposure.

3. After exposure, slide the screen face down into the PhosphorImager system.
4. Select the scanning area using the software supplied with the PhosphorImager and start scanning.

The blue light emitted during scanning is collected to produce the latent image.

5. Analyze and quantitate the image using the software provided.
6. Erase the phosphor screen by exposing it to visible light as in step 1.

COMMENTARY

Background Information

The ability to detect radiolabeled proteins or DNA is critical to many studies in molecular biology. More often than not, detection is coupled with the resolving power of polyacrylamide gel electrophoresis (PAGE; *UNITS 2.7 & 10.2*). Radiolabeled proteins or DNA separated on gels can be used directly to obtain an autoradiographic image. Alternatively, separated by PAGE are frequently transferred to membranes (*UNIT 10.8*) and detected using radiolabeled probes such as DNA, RNA, or antibodies. The autoradiographic image, whether generated on film or a phosphor screen, reflects the distribution of the radioactive molecules on the two-dimensional surface of the gel or filter. Molecular sizes of radiolabeled proteins or DNA species, therefore, can be determined by correlating their positions with molecular markers. Also, the density of the band images can be used to determine the relative quantities of the radiolabeled molecules in the sample.

Critical Parameters

The sensitivity of the detection device and the strength of the radioactive signal are the two

most important parameters for autoradiography. Sensitivity can be enhanced by treating samples with fluors or by using intensifying screens (Table A.3A.2). Because phosphor imaging is 10 to 250 times more sensitive than film (Johnston et al., 1990), this technology makes it possible to monitor radioactive samples that would previously have gone undetected with film.

A second important parameter is the range over which the measurement device is linear. Film must be preflashed in order for the intensity of the image to be linear with respect to the amount of radioactivity, particularly for weakly radioactive samples (Laskey and Mills, 1975, 1977). Phosphor imaging offers a much wider linear range of measurement (5 orders of magnitude compared to 1.5 for film; Johnston et al., 1990). This makes it possible to quantitate very weak or very strong radioactive samples accurately.

Troubleshooting

Cracking is one of the most common problems encountered when drying gels. This often occurs if the gel is removed from the dryer

Table A.3A.2 Film Choice and Exposure Temperature for Autoradiography

Isotope	Enhancement method	Film	Exposure temperature
³ H	Fluorography	Double-coated	-70°C
³⁵ S, ¹⁴ C, ³² P	None	Single-coated	Room temperature
³⁵ S, ¹⁴ C, ³² P	Fluorography	Double-coated	-70°C
³² P, ¹²⁵ I	CaWO ₄ intensifying screens	Double-coated	-70°C

before it has adequately dried or if drying temperatures are too high. To overcome this problem, drying times should be extended and the performance of the vacuum pump and heater unit should be checked. For protein gels, with high percentages of polyacrylamide or >1.5 mm thick, cracking can be reduced by using an alternative fixing solution containing glycerol (3% glycerol/10% glacial acetic acid/20% methanol; see Support Protocol 1).

Among the biggest problems encountered in autoradiography are images that are either too weak or too intense. Such problems can be solved by varying the exposure time. Estimating initial exposure time is difficult, since the amount of radioactivity in the sample is often unknown. A Geiger counter can offer some guidance with certain isotopes. For highly exposed film, the length of time the film sits in developer can be reduced to produce a lighter image. It is particularly important to remember that if accurate quantification of the film image is desired, film must be preflashed so that there is a linear relationship between the amount of radioactivity in the sample and the image intensity.

Artifacts such as black spots and stripes can be avoided by making sure that no moisture comes in contact with the film during developing and that films exposed at -70°C are brought to room temperature before developing. Also, it must be noted that β particles from weak isotopes such as ³H cannot penetrate plastic wrap, and plastic wraps can attenuate signals from ³⁵S and ¹⁴C up to 2-fold.

Anticipated Results

The protocols described here should yield a film image of a gel that can be quantified, stored, and photographed for publication.

Time Considerations

Fixing a protein gel will require ~45 min. Drying will take an additional 2 hr for a gel 1 mm in thickness. Incorporation of a fluor will add ~45 min to the processing time.

For protein gels >1.5 mm thick or with >15% acrylamide, an additional 30 min will be required for fixing and ~30 additional minutes will be required for drying.

The length of exposure for films in autoradiography can range from a few minutes to a few weeks, depending on the strength of the radioactivity in the sample. Most exposures last from several hours to a few days. Exposure time can be reduced >10-fold with a phosphor imager.

Literature Cited

- Chamberlain, J.P. 1979. Fluorographic detection of radioactivity in polyacrylamide gels with the water-soluble fluor, sodium salicylate. *Anal. Biochem.* 98:132-135.
- Johnston, R.F., Pickett, S.C., and Barker, D.L. 1990. Autoradiography using storage phosphor technology. *Electrophoresis* 11:355-360.
- Laskey, R.A. 1980. The use of intensifying screens or organic scintillators for visualizing radioactive molecules resolved by gel electrophoresis. *Methods Enzymol.* 65:363-371.
- Laskey, R.A. and Mills, A.D. 1975. Quantitative film detection of ³H and ¹⁴C in polyacrylamide gels by fluorography. *Eur. J. Biochem.* 56:335-341.
- Laskey, R.A. and Mills, A.D. 1977. Enhanced autoradiographic detection of ³²P and ¹²⁵I using intensifying screens and hypersensitized film. *FEBS Lett.* 82:314-316.

Contributed by Daniel Voytas and Ning Ke
Iowa State University
Ames, Iowa

Silanizing Glassware

Glassware is silanized (siliconized) to prevent adsorption of solute to the glass surface or to increase its hydrophobicity. This is particularly important when dealing with low concentrations of particularly “sticky” solutes such as single-stranded nucleic acids or proteins.

**BASIC
PROTOCOL**

Materials

Chlorotrimethylsilane *or* dichlorodimethylsilane
Vacuum pump
Desiccator, equipped with a valve

1. In a fume hood, place glassware or equipment to be silanized into desiccator along with a beaker containing 1 to 3 ml of chlorotrimethylsilane or dichlorodimethylsilane.

CAUTION: Chlorotrimethylsilane and dichlorodimethylsilane vapors are toxic and highly flammable.

2. Connect desiccator to vacuum pump until silane starts to boil, and close connection to pump (maintaining vacuum in desiccator). Leave the desiccator evacuated and closed until liquid silane is gone (~1 to 3 hr).

During the incubation the silane will evaporate, be deposited on the surface of the glassware, and polymerize. Do not leave the desiccator attached to the vacuum pump. This will suck away the silane, minimizing deposition and damaging the pump.

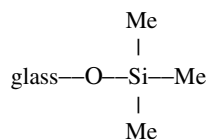
3. Open desiccator in a fume hood, and leave open for several minutes to disperse silane vapors.
4. If desired, bake or autoclave the glassware or apparatus.

Autoclaving or rinsing with water removes the reactive chlorosilane end of the dimethylsiloxane polymer generated by dichlorodimethylsilane.

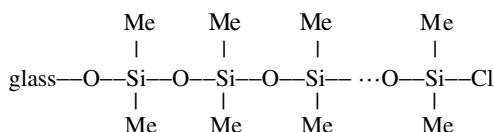
COMMENTARY

Untreated glass contains silicate and silanol groups that can act as ion-exchange and nucleophilic centers. To mask these groups and decrease the hydrophilicity of the surface, various reactive silanes are frequently used to coat the glass surface. The same or related chemistries can be used to introduce functional groups, including large molecules, onto the glass surface. On contact with a silanol (SiOH) group, chlorosilanes react to give HCl and a siloxane Si-O-Si linkage. For most applications “silanizing” or “siliconizing” a piece of glassware or equipment means introducing onto the glass surface either a short polymer of dimethylsiloxane

using dichlorodimethylsilane, or a trimethylsiloxane cap



using chlorotrimethylsilane. Polydimethylsiloxane is silicone oil; the use of dichloromethylsilane essentially gives a light coating of oil to the glassware. The polymerization of dichlorodimethylsilane is mediated by residual water adsorbed to the glass surface; chlorosilanes also react with water and alcohols to give HCl and silanol or alkoxy silane, respectively. Thus, in the polymerization reaction, the residual water reacts with the chloride endgroup to give HCl and a silanol, which then reacts with more



**Commonly Used
Techniques**

A.3B.1

dichlorodimethylsilane and more water to give polymers. Chlorotrimethylsilane just gives a trimethyl cap to the silanol or silicate group.

Items too large to fit in a desiccator can be silanized by briefly rinsing with or soaking in a solution of approximately 5% dichlorodimethylsilane in various volatile organic solvents such as chloroform or heptane. The organic solvent is removed by evaporation, depositing the dichlorodimethylsilane on the surface. This approach is particularly useful for

treating glass plates for denaturing polyacrylamide sequencing gels (see *UNIT 7.4*).

CAUTION: If a flammable solvent is used, do not bake the glassware until the solvent is completely evaporated.

Contributed by Brian Seed
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

Dialysis and Ultrafiltration

The first section describes dialysis of large- and small-volume samples using cellulose membranes with pore sizes designed to exclude molecules above a selected molecular weight (basic and alternate protocols). A support protocol describes preparation of membranes for dialysis. The second section describes protocols for ultrafiltration, and the third section describes simple methods of concentrating solutions.

DIALYSIS

Conventional dialysis separates small molecules from large molecules by allowing diffusion of only the small molecules through selectively permeable membranes. Dialysis is usually used to change the salt (small-molecule) composition of a macromolecule-containing solution. The solution to be dialyzed is placed in a sealed dialysis membrane and immersed in a selected buffer; small solute molecules then equilibrate between the sample and the dialysate. Concomitant with the movement of small solutes across the membrane, however, is the movement of solvent in the opposite direction. This can result in some sample dilution (usually <50%).

Large-Volume Dialysis

This protocol describes the use of membranes, prepared using the support protocol, for dialysis of samples in large, easily handled volumes, typically 0.1 to 500 ml.

Materials

For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

- Dialysis membrane (see support protocol)
- Clamps (Spectrapor Closures, Spectrum, or equivalent)
- Macromolecule-containing sample to be dialyzed
- Appropriate dialysis buffer

1. Remove dialysis membrane from ethanol storage solution and rinse with distilled water. Secure clamp to one end of the membrane or knot one end.

Always use gloves to handle dialysis membrane because the membrane is susceptible to cellulolytic microorganisms.

See support protocol for a discussion of commercially available membranes and preparation and storage procedures.

Clamps are less likely to leak than knots.

2. Fill membrane with water or buffer, hold the unclamped end closed, and squeeze membrane. A fine spray of liquid indicates a pinhole in the membrane; discard and try a new membrane.

Pinholes are rare but cause traumatic sample loss. Hard squeezing will cause some liquid to seep from the bag; this is normal.

3. Replace the water or buffer in dialysis membrane with the macromolecule-containing sample and clamp the open end. Again, squeeze to check the integrity of the membrane and clamps.

If dialyzing a concentrated or high-salt sample, leave some space in the clamped membrane; there will be a net flow of water into the sample, and if sufficient pressure builds up the membrane can burst.

BASIC
PROTOCOL

Commonly Used
Techniques

A.3C.1

Contributed by Louis Zumstein

Current Protocols in Molecular Biology (1998) A.3C.1-A.3C.7

Copyright © 1998 by John Wiley & Sons, Inc.

Supplement 41

4. Immerse dialysis membrane in a beaker or flask containing a large volume (relative to the sample) of the desired buffer. Dialyze for several hours at the desired temperature with gentle stirring of the buffer.

Dialysis rates are dependent on membrane pore size, sample viscosity, and ratio of membrane surface to sample volume. Temperature has little effect on dialysis rate, but low temperatures are usually chosen to enhance macromolecule stability. Common salts will equilibrate across a 15,000-MWCO membrane (see support protocol) in ~3 hr with stirring.

5. Change dialysis buffer as necessary.

Usually two dialysis buffer changes are sufficient. For example, when CsCl is removed from equilibrium density gradient-banded DNA, two equilibrations against a 1000-fold volume excess of buffer will decrease CsCl concentration 10^6 -fold, to ~5 μ M.

6. Remove dialysis membrane from the buffer. Hold the membrane vertically and remove excess buffer trapped in end of membrane outside upper clamp. Release upper clamp and remove the sample with a Pasteur pipet.

ALTERNATE PROTOCOL

Small-Volume Dialysis

For solution volumes less than ~100 μ l, the use of dialysis membrane as described above can result in unacceptable sample loss. The method described below can easily dialyze volumes as small as 10 μ l. The sample is held in a 0.5- or 1.5-ml microcentrifuge tube, with dialysis membrane covering the open end of the tube.

Additional Materials

Cork borer

1. Cut a hole in the cap of a 0.5- or 1.5-ml microcentrifuge tube using a cork borer. Be sure that there are no rough edges on the inside of the cap (which will be in contact with the dialysis membrane).

The success of this protocol is dependent on a tight-fitting cap, so make sure the cap is well-seated. On the other hand, an excessively tight-fitting cap can rip the membrane when it is inserted. Sarstedt tubes work well.

2. Place sample in the microcentrifuge tube, cover tube opening with dialysis membrane, and snap on the tube cap in which a hole has been cut.

This places a fair amount of stress on the dialysis membrane, so use a fairly thick membrane and make sure there are no rough edges on the cap.

3. To get the sample in contact with the dialysis membrane, *gently* centrifuge the microcentrifuge tube in an inverted position in a tabletop centrifuge.

4. Keep the tube inverted, immerse it in dialysis buffer, and anchor it so that it will remain inverted and immersed.

This is easily done by clamping the tube to the side of a beaker containing the dialysis buffer. For many samples, immerse and anchor a microcentrifuge tube rack in the buffer, then insert the inverted tube into the bottom of the rack. Air in the microcentrifuge tube will keep the tube pressed up into the rack.

5. Use a bent Pasteur pipet to blow out any air bubbles trapped under the cap to allow dialysis buffer to contact the membrane.

6. Stir the dialysis buffer and dialyze at an appropriate temperature for at least several hours (see first basic protocol, step 4).

7. To recover the sample, remove microcentrifuge tube from the buffer and centrifuge briefly right-side-up.

Commercially available alternatives to this method include use of individual hollow-fiber filters from Spectrum (with sample capacities of 1 to 140 μ l; see Ultrafiltration), and many different microdialysis machines, both single-sample and multisample (Spectrum, Cole-Parmer, Hofer, and others).

Selection and Preparation of Dialysis Membrane

Dialysis membranes are available in a number of thicknesses and pore sizes. Thicker membranes are tougher, but restrict solute flow and reach equilibrium more slowly. Pore size is defined by “molecular weight cutoff” (MWCO)—i.e., the size of the smallest particle that cannot penetrate the membrane. The MWCO designation should be used only as a very rough guide; if accurate MWCO information is required, it should be determined empirically (see Craig, 1967, for a discussion of parameters affecting MWCO). Knowledge of the precise MWCO is usually not required, however; it is necessary only to use a membrane with a pore size that is much smaller than the macromolecule of interest. For most plasmid and protein dialyses, a MWCO of 12,000 to 14,000 is appropriate.

Most dialysis membranes are made of derivatives of cellulose. They come in a wide variety of MWCO values, ranging from 500 to 500,000, and also vary in cleanliness, sterility, and cost. Spectrum has the most impressive inventory. The least expensive membranes come dry on rolls; these contain glycerol to keep them flexible as well as residual sulfides and traces of heavy metals from their manufacture. For most DNA dialysis applications, such as removing CsCl from gradient-purified samples (*UNIT 1.7*) and electroelution (*UNIT 2.6*), these dialysis membranes can be used directly from the roll after wetting and a brief rinse in water. If glycerol, sulfur compounds, or small amounts of heavy metals are problematic, clean membranes should be purchased or membranes should be prepared as described below. Protein dialysis should only be done with clean membranes.

Additional Materials

For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

10 mM sodium bicarbonate
10 mM Na₂EDTA, pH 8.0
20% to 50% (v/v) ethanol

1. Remove membrane from the roll and cut into usable lengths (usually 8 to 12 in.).

Always use gloves to handle dialysis membrane, as it is susceptible to a number of cellulolytic microorganisms. Membrane is available as sheets or preformed tubing.

2. Wet membrane and boil it several minutes in a large excess of 10 mM sodium bicarbonate.
3. Boil several minutes in 10 mM Na₂EDTA. Repeat.

Boiling speeds up the treatment process but is not necessary. A 30-min soak with some agitation can substitute for the boiling step.

4. Wash several times in distilled water.

5. Store at 4°C in 20% to 50% ethanol to prevent growth of cellulolytic microorganisms.

Alternatively, noxious buffers (e.g., sodium azide or sodium cacodylate) may be used for storage; however, ethanol is preferred for ease and convenience.

Key References

Craig, L.C. 1967. Techniques for the study of peptides and proteins by dialysis and diffusion. *Methods Enzymol.* 11:870-905.

McPhie, P. 1971. Dialysis. *Methods Enzymol.* 22:23-32.

Describe the theory and practice of dialysis and diffusion.

ULTRAFILTRATION

Ultrafiltration (UF) is a type of pressure-driven membrane filtration in which a pressure differential across a semipermeable membrane is used to separate solvent and small solute molecules from a solution containing larger molecules. UF usually uses membranes with MWCOs of 1000 to 1,000,000. The pressure differential is usually achieved by vacuum, pressurized inert gas, or centrifugal force. UF is frequently used to concentrate protein solutions; because buffer concentration does not change, this is one of the most gentle methods for concentrating proteins. UF is also used to separate molecules of grossly different molecular weight (e.g., plasmids or PCR products from dNTPs or primers).

The major technical difficulty with UF is concentration polarization, in which rejected large-molecule solute builds up as a gel layer that obstructs the membrane. Various stirring and pumping regimes are used to keep the solution flowing tangentially to the membrane, and this serves to dissipate the gel layer.

UF can be broadly divided into vacuum-dialysis, centrifugal-concentration, stirred-cell dialysis, and tangential-flow systems. Choice of system will depend on the nature of the solution (i.e., volume, viscosity, presence of particulates, and shear lability), concentration rate needed, and cost. In general, vacuum dialysis is cheapest, but is also slowest and has the lowest volume capacity. Centrifuge concentrators are fairly inexpensive (on a per-use basis), moderately slow, and restricted to volumes ≤ 20 ml. Stirred cells come next on the scale of increasing volume capacity, concentration rate, and cost. Tangential-flow systems can concentrate large volumes rapidly, but at substantial cost.

Vacuum dialysis can be done with common laboratory equipment; construction of a homemade apparatus is described in the basic protocol. Centrifuge concentrators are single-use disposable units. The other types of UF generally require a commercial apparatus. There is a plethora of such apparatuses; because their design and availability can change rapidly, the following discussion is meant as a general guideline to the available methods. It should be noted that most UF apparatus can also be used for large-volume dialysis.

NOTE: For a more complete listing of suppliers of ultrafiltration equipment, see the *Biotech Buyer's Guide* (ACS, 1994), published annually.

Ultrafiltration Using Vacuum Dialysis

In vacuum dialysis, a vacuum is placed on the dialysis buffer while the sample remains at atmospheric pressure. This is a very gentle method; however, as membrane polarization can be a problem, it is usually used with dilute samples. Vacuum dialysis can typically handle ≤ 20 ml sample volume.

Materials

For common stock solutions, see **APPENDIX 2**; for suppliers, see **APPENDIX 4**.

Sample
Appropriate dialysis buffer
Vacuum flask
One-hole cork stopper
Pasteur pipet
Dialysis membrane of desired MWCO

1. Connect outlet of vacuum flask to vacuum line or aspirator (see Fig. A.3C.1).
2. Prepare or purchase a one-hole cork stopper into which the large end of a Pasteur pipet fits snugly, even with lubrication. Cut off the thin end of a Pasteur pipet and smooth the end using the flame from a Bunsen burner.
3. Prepare the dialysis membrane, clamp one end, and slip the other end through the hole in the stopper. Insert the Pasteur pipet into the open end of the dialysis tubing, using glycerol as a lubricant if necessary.

A 6.4-mm-diameter membrane fits snugly over a Pasteur pipet.

4. Insert the Pasteur pipet with dialysis membrane into the stopper until firmly seated. Keep the membrane wet at all times.

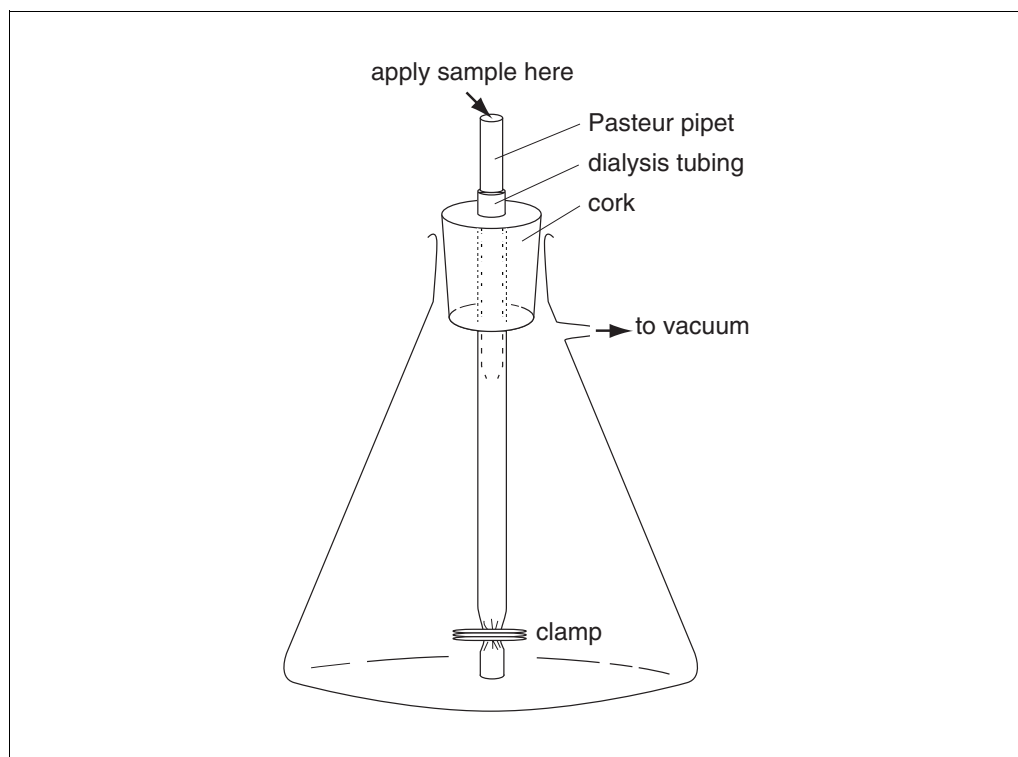


Figure A.3C.1 Vacuum dialysis apparatus (from Craig, 1967).

5. Put the stopper/Pasteur pipet/dialysis membrane assembly into the mouth of a buffer-filled vacuum flask and apply a mild vacuum. The dialysis membrane will inflate.

Check for trails of small bubbles, which indicate leaks.

6. If no leaks are detected, put the sample into the membrane via the Pasteur pipet and dialyze under vacuum until the desired sample volume is achieved. Add sample continuously or at intervals throughout dialysis.

Because vacuum pumps can rupture the membrane, the use of an aspirator with overnight stirring at 4°C is recommended. This reduces 10 to 15 ml of fairly dilute protein solution to <0.5 ml. Vacuum ≤ 3 lb/in² (150 mm Hg) will cause little membrane distortion, but 5 lb/in² will cause some change in pore size. Commercial apparatuses are also available (e.g., Spectrum); they are more expensive, but can more easily handle multiple samples.

ALTERNATE PROTOCOL 1

Ultrafiltration Using Centrifuge Concentrators

Centrifuge concentrators use centrifugal force imparted by a centrifuge to force solvent through a membrane. They are more expensive than home-made vacuum dialyzers, but are faster and more amenable to small samples and multiple samples. Centrifuge concentrators are currently available to handle sample volumes of 0.05 to 20 ml, with a huge variety of MWCOs ranging from 100 to 500,000. Polarization is avoided by spinning in fixed-angle rotors so the membrane is not at a right angle to the force; any gel that forms is pushed off to one side of the membrane. This design also prevents complete drying of the sample. Centriprep concentrators (Amicon) use a clever design to force the solution upward through the membrane, allowing any gel to pellet at the bottom of the apparatus, away from the membrane. Centrifuge concentrators are made by Amicon, Corning, Millipore, and Spectrum, among others.

ALTERNATE PROTOCOL 2

Ultrafiltration Using Stirred-Cell Apparatus

Stirred-cell units (Amicon or Spectrum) are a relatively inexpensive means for concentrating moderate volumes of dilute, nonparticulate samples. They use pressurized inert gas (usually N₂) to drive the solvent through membrane disks. A stir bar suspended above the disk keeps the solution stirred in an attempt to dissipate any gel layer. The disks are available in various MWCOs and compositions. It is possible to concentrate 3 to 400 ml with stirred-cell units. They are cheaper than a tangential-flow apparatus (see below), but have a lower concentration rate and are more prone to concentration polarization than tangential-flow systems. However, they are less prone to concentration polarization than vacuum dialysis.

ALTERNATE PROTOCOL 3

Ultrafiltration Using Tangential-Flow Systems

Like stirred-cell apparatuses, tangential-flow ultrafiltration systems are pressure-driven. However, instead of stirring the solution, these units pump the pressurized solution tangentially with respect to the membrane, greatly reducing the concentration-polarization problem. Tangential-flow systems are more expensive than stirred-cell units. Most tangential-flow systems can also be adapted for large-volume dialysis.

One design, known as thin-channel or flat-plate, uses a flat ultrafiltration membrane. The Thin Channel System (Amicon) and Minitan System (Millipore) are examples of this design. The solution is pumped across one face of the membrane through either a spiral channel or a series of linear channels. This design has a surface area similar to that of stirred cells, but can handle larger volumes (~200 ml to 2 liters). The technique is much

less susceptible to concentration polarization and can thus handle more concentrated solutions. The concentration rate is greater than that of stirred cells, as greater pressure can be achieved before concentration polarization becomes a problem. Tangential-flow systems are somewhat cheaper and simpler than the high-capacity systems described below.

Stacked-flat-plate, spiral-wound, and hollow-fiber systems (e.g., Amicon, Millipore, Spectrum) are designed for large-volume samples (1 to >1000 liters) and high concentration rates, requirements usually encountered in an industrial setting. The choice between them will normally depend more on sample characteristics than on cost (see below). Also, the filter cartridges are not cheap, so preventing sample cross-contamination usually involves extensive cleaning, rather than simply changing a membrane as with stirred-cell or lower-capacity tangential-flow systems.

The principle behind the stacked-flat-plate system (e.g., Pellicon System; Millipore) is self-explanatory. The surface area of such systems is limited, and the apparatus may be difficult to clean.

Spiral-wound systems consist of a large membrane wound on a core with a mesh separating the two faces of the membrane, thereby creating two noncontiguous spiral channels. Pressure is applied to the channel containing the input solution, and filtrate is collected from the other. This design achieves a very high surface-to-volume ratio, and thus a high concentration rate. These systems are susceptible to clogging, and are recommended only for particle-free solutions.

Hollow-fiber systems are collections of hollow fibers of ultrafiltration membrane; pressure is applied to the sample as it is pumped through the fibers. This design has the advantage of handling particulate samples (i.e., cellular material) and precipitates. They are also easy to clean, but cannot achieve high pressures, and so have lower concentration rates than spiral-channel systems.

Key References

- ACS (American Chemical Society). 1994. *Biotech Buyers' Guide 1994*. ACS, Washington, D.C.
- McGregor, W.C. 1986. Selection and use of ultrafiltration membranes. *In Membrane Separations in Biotechnology* (W.C. McGregor, ed.) pp. 1-36. Marcel Dekker, New York.
- Pohl, T. 1990. Concentration of proteins and removal of solutes. *Methods Enzymol.* 182:68-83.

CONCENTRATION USING SOLID-PHASE OR SLURRY ABSORBENTS

A cheap, low-tech alternative to dialysis and ultrafiltration is concentration with solid-phase or slurry absorbents. In this method, the sample is placed in dialysis membrane and surrounded with dry matrix or immersed in an absorbent slurry, whereupon water and small solute molecules are drawn out of the sample. This technique is slow and requires constant monitoring; otherwise, complete drying can occur, to the potential detriment of the sample. Common absorbents are polyethylene glycol (PEG) and carboxymethyl cellulose (Aquacides; Calbiochem), Sephadex beads, and polyacrylamide derivatives (Spectra/Gel, Spectrum).

Contributed by Louis Zumstein
Sennes Drug Innovations
Houston, Texas

Quantitation of DNA and RNA with Absorption and Fluorescence Spectroscopy

Reliable quantitation of nanogram and microgram amounts of DNA and RNA in solution is essential to researchers in molecular biology. In addition to the traditional absorbance measurements at 260 nm (see Basic Protocol), three more sensitive fluorescence techniques (see Alternate Protocols 1 to 3) are presented below. These four procedures cover a range from 5 to 10 ng/ml DNA to 25 pg/ml DNA (see Commentary and Table A.3D.3 therein).

Absorbance measurements are straightforward as long as any contribution from contaminants and the buffer components are taken into account. Fluorescence assays are less prone to interference than A_{260} measurements and are also simple to perform. As with absorbance measurement, a reading from the reagent blank is taken prior to adding the DNA. In instruments where the readout can be set to indicate concentration, a known concentration is used for calibration and subsequent readings are taken in $\mu\text{g/ml}$, ng/ml , or pg/ml DNA.

DETECTION OF NUCLEIC ACIDS USING ABSORPTION SPECTROSCOPY

BASIC PROTOCOL

Absorption of the sample is measured at several different wavelengths to assess purity and concentration of nucleic acids. A_{260} measurements are quantitative for relatively pure nucleic acid preparations in microgram quantities. Absorbance readings cannot discriminate between DNA and RNA; however, the ratio of A at 260 and 280 nm can be used as an indicator of nucleic acid purity. Proteins, for example, have a peak absorption at 280 nm that will reduce the A_{260}/A_{280} ratio. Absorbance at 325 nm indicates particulates in the solution or dirty cuvettes; contaminants containing peptide bonds or aromatic moieties such as protein and phenol absorb at 230 nm.

This protocol is designed for a single-beam ultraviolet to visible range (UV-VIS) spectrophotometer. If available, a double-beam spectrophotometer will simplify the measurements, as it will automatically compare the cuvette holding the sample solution to a reference cuvette that contains the blank. In addition, more sophisticated double-beam instruments will scan various wavelengths and report the results automatically.

Materials

- 1 \times TNE buffer (see recipe)
- DNA sample to be quantitated
- Calf thymus DNA standard solutions (see recipe)
- Matched quartz semi-micro spectrophotometer cuvettes (1-cm pathlength)
- Single- or dual-beam spectrophotometer (ultraviolet to visible)

1. Pipet 1.0 ml of 1 \times TNE buffer into a quartz cuvette. Place the cuvette in a single- or dual-beam spectrophotometer, read at 325 nm (note contribution of the blank relative to distilled water if necessary), and zero the instrument. Use this blank solution as the reference in double-beam instruments. For single-beam spectrophotometers, remove blank cuvette and insert cuvette containing DNA sample or standard suspended in the same solution as the blank. Take reading. Repeat this process at 280, 260, and 230 nm.

It is important that the DNA be suspended in the same solution as the blank.

Commonly Used Techniques

A.3D.1

Contributed by Sean R. Gallagher

Current Protocols in Molecular Biology (2004) A.3D.1-A.3D.12

Copyright © 2004 by John Wiley & Sons, Inc.

2. To determine the concentration (C) of DNA present, use the A_{260} reading in conjunction with one of the following equations:

$$\text{Single-stranded DNA: } C \text{ (pmol/}\mu\text{l)} = \frac{A_{260}}{10 \times S}$$

$$C \text{ (}\mu\text{g/ml)} = \frac{A_{260}}{0.027}$$

$$\text{Double-stranded DNA: } C \text{ (pmol/}\mu\text{l)} = \frac{A_{260}}{13.2 \times S}$$

$$C \text{ (}\mu\text{g/ml)} = \frac{A_{260}}{0.020}$$

$$\text{Single-stranded RNA: } C \text{ (}\mu\text{g/ml)} = \frac{A_{260}}{0.025}$$

$$\text{Oligonucleotide: } C \text{ (pmol/}\mu\text{l)} = A_{260} \times \frac{100}{1.5 N_A + 0.71 N_C + 1.20 N_G + 0.84 N_T}$$

where S represents the size of the DNA in kilobases and N is the number or residues of base A, G, C, or T.

For double- or single-stranded DNA and single-stranded RNA: These equations assume a 1-cm-pathlength spectrophotometer cuvette and neutral pH. The calculations are based on the Lambert-Beer law, $A = ECl$, where A is the absorbance at a particular wavelength, C is the concentration of DNA, l is the pathlength of the spectrophotometer cuvette (typically 1 cm), and E is the extinction coefficient. For solution concentrations given in mol/liter and a cuvette of 1-cm pathlength, E is the molar extinction coefficient and has units of $M^{-1}cm^{-1}$. If concentration units of $\mu\text{g/ml}$ are used, then E is the specific absorption coefficient and has units of $(\mu\text{g/ml})^{-1}cm^{-1}$. The values of E used here are as follows: ssDNA, $0.027 (\mu\text{g/ml})^{-1}cm^{-1}$; dsDNA, $0.020 (\mu\text{g/ml})^{-1}cm^{-1}$; ssRNA, $0.025 (\mu\text{g/ml})^{-1}cm^{-1}$. Using these calculations, an A_{260} of 1.0 indicates 50 $\mu\text{g/ml}$ double-stranded DNA, $\sim 37 \mu\text{g/ml}$ single-stranded DNA, or $\sim 40 \mu\text{g/ml}$ single-stranded RNA (adapted from Applied Biosystems, 1987).

For oligonucleotides: Concentrations are calculated in the more convenient units of pmol/ μl . The base composition of the oligonucleotide has significant effects on absorbance, because the total absorbance is the sum of the individual contributions of each base (Table A.3D.1).

3. Use the A_{260}/A_{280} ratio and readings at A_{230} and A_{325} to estimate the purity of the nucleic acid sample.

Table A.3D.1 Molar Extinction Coefficients of DNA Bases^a

Base	$\epsilon_{260 \text{ nm}}^{\text{IM}}$
Adenine	15,200
Cytosine	7,050
Guanosine	12,010
Thymine	8,400

^aMeasured at 260 nm; see Wallace and Miyada, 1987. Detailed spectrophotometric properties of nucleoside triphosphates are listed in UNIT 3.4.

Table A.3D.2 Spectrophotometric Measurements of Purified DNA^a

Wavelength (nm)	Absorbance	A_{260}/A_{280}	Conc. ($\mu\text{g/ml}$)
325	0.01	—	—
280	0.28	—	—
260	0.56	2.0	28
230	0.30	—	—

^aTypical absorbancy readings of highly purified calf thymus DNA suspended in 1× TNE buffer. The concentration of DNA was nominally 25 $\mu\text{g/ml}$.

Ratios of 1.8 to 1.9 and 1.9 to 2.0 indicate highly purified preparations of DNA and RNA, respectively. Contaminants that absorb at 280 nm (e.g., protein) will lower this ratio.

Absorbance at 230 nm reflects contamination of the sample by phenol or urea, whereas absorbance at 325 nm suggests contamination by particulates and dirty cuvettes. Light scatter at 325 nm can be magnified 5-fold at 260 nm (K. Hardy, pers. comm.).

Typical values at the four wavelengths for a highly purified preparation are shown in Table A.3D.2.

DNA DETECTION USING THE DNA-BINDING FLUOROCHROME HOECHST 33258

ALTERNATE PROTOCOL 1

Use of fluorometry to measure DNA concentration has gained popularity because it is simple and much more sensitive than spectrophotometric measurements. Specific for nanogram amounts of DNA, the Hoechst 33258 fluorochrome has little affinity for RNA and works equally well with either whole-cell homogenates or purified preparations of DNA. The fluorochrome is, however, sensitive to changes in DNA composition, with preferential binding to AT-rich regions. A fluorometer capable of an excitation wavelength of 365 nm and an emission wavelength of 460 nm is required for this assay.

Additional Materials (also see Basic Protocol)

Hoechst 33258 assay solution (working solution; see recipe)

Dedicated filter fluorometer (Hoefer TKO100) *or* scanning fluorescence spectrophotometer (Shimadzu model RF-5000 or Perkin-Elmer model LS-5B or LS-3B)

Fluorometric square glass cuvettes *or* disposable acrylic cuvettes (Sarstedt)

Teflon stir rod

1. Prepare the scanning fluorescence spectrophotometer by setting the excitation wavelength to 365 nm and the emission wavelength to 460 nm.

The dedicated filter fluorometer has fixed wavelengths at 365 and 460 nm and does not need adjustment.

2. Pipet 2.0 ml Hoechst 33258 assay solution into cuvette and place in sample chamber. Take a reading without DNA and use as background.

If the fluorometer has a concentration readout mode or is capable of creating a standard curve, set instrument to read 0 with the blank solution. Otherwise note the readings in relative fluorescence units. Be sure to take a blank reading for each cuvette used, as slight variations can cause changes in the background reading.

3. With the cuvette still in the sample chamber, add 2 μl DNA standard to the blank Hoechst 33258 assay solution. Mix in the cuvette with a Teflon stir rod or by capping and inverting the cuvette. Read emission in relative fluorescence units or set the

Commonly Used Techniques

A.3D.3

concentration readout equal to the final DNA concentration. Repeat measurements with remaining DNA standards using fresh assay solution (take background zero reading and zero instrument if needed).

If necessary, the DNA standards should be quantitated by A_{260} measurement (Basic Protocol) before being used here.

Small-bore tips designed for loading sequencing gels minimize errors of pipetting small volumes. Prerinse tips with sample and make sure no liquid remains outside the tip after drawing up the sample.

Read samples in duplicate or triplicate, with a blank reading taken each time. Unusual or unstable blank readings indicate a dirty cuvette or particulate material in the solution, respectively.

4. Repeat step 3 with unknown samples.

A dye concentration of 0.1 $\mu\text{g/ml}$ is adequate for final DNA concentrations up to ~ 500 ng/ml . Increasing the working dye concentration to 1 $\mu\text{g/ml}$ Hoechst 33258 will extend the assay's range to 15 $\mu\text{g/ml}$ DNA, but will limit sensitivity at low concentrations (5 to 10 ng/ml). Sample volumes of ≤ 10 μl can be added to the 2.0-ml aliquot of Hoechst 33258 assay solution.

**ALTERNATE
PROTOCOL 2**

**DNA AND RNA DETECTION WITH ETHIDIUM BROMIDE
FLUORESCENCE**

In contrast to the fluorochrome Hoechst 33258, ethidium bromide is relatively unaffected by differences in the base composition of DNA. Ethidium bromide is not as sensitive as Hoechst 33258 and, although capable of detecting nanogram levels of DNA, will also bind to RNA. In preparations of DNA with minimal RNA contamination or with DNA samples having an unusually high guanine and cytosine (GC) content where the Hoechst 33258 signal can be quite low, ethidium bromide offers a relatively sensitive alternative to the more popular Hoechst 33258 DNA assay. A fluorometer capable of an excitation wavelength of 302 or 546 nm and an emission wavelength of 590 nm is required for this assay.

Additional Materials (also see Basic Protocol)

Ethidium bromide assay solution (see recipe)

1. Pipet 2.0 ml ethidium bromide assay solution into cuvette and place in sample chamber. Set excitation wavelength to 302 nm or 546 nm and emission wavelength to 590 nm. Take an emission reading without DNA and use as background.

If the instrument has a concentration readout mode or is capable of creating a standard curve, set instrument to read 0 with the blank solution. Otherwise note the readings in relative fluorescence units.

The excitation wavelength of this assay can be either in the UV range (~ 302 nm) using a quartz cuvette or in the visible range (546 nm) using a glass cuvette. In both cases the emission wavelength is 590 nm.

2. Read and calibrate these samples as described in step 3 of the Hoechst 33258 assay.
3. Read emissions of the unknown samples as in step 4 of the Hoechst 33258 assay.

A dye concentration of 5 $\mu\text{g/ml}$ in the ethidium bromide assay solution is appropriate for final DNA concentrations up to 1000 ng/ml . 10 $\mu\text{g/ml}$ ethidium bromide in the ethidium bromide assay solution will extend the assay's range to 10 $\mu\text{g/ml}$ DNA, but is only used for DNA concentrations > 1 $\mu\text{g/ml}$. Sample volumes of up to 10 μl can be added to the 2.0-ml aliquot of ethidium bromide assay solution.

DNA DETECTION USING PICOGREEN dsDNA QUANTITATION REAGENT

ALTERNATE PROTOCOL 3

PicoGreen dsDNA quantitation reagent enables researchers to quantitate as little as 25 pg/ml of dsDNA (50 pg dsDNA in a 2-ml assay volume) with a standard spectrofluorometer and fluorescein excitation and emission wavelengths. This sensitivity exceeds that achieved with the Hoechst 33258-based assay (Alternate Protocol 1) by 400-fold. Using a fluorescence microplate reader, it is possible to detect as little as 250 pg/ml dsDNA (50 pg in a 200- μ l assay volume). The standard PicoGreen assay protocol is also simpler than that for Hoechst 33258 because a single concentration of the PicoGreen reagent allows detection over the full dynamic range of the assay. In order to achieve more than two orders of magnitude in dynamic range with Hoechst-based assays, two different dye concentrations are recommended. In contrast, the linear detection range of the PicoGreen assay in a standard fluorometer extends over more than four orders of magnitude in DNA concentration—from 25 pg/ml to 1000 ng/ml—with a single dye concentration (Fig. A.3D.1). Linearity is maintained in the presence of several compounds that commonly contaminate nucleic acid preparations, including salts, urea, ethanol, chloroform, detergents, proteins, and agarose.

CAUTION: No data are available addressing the mutagenicity or toxicity of PicoGreen dsDNA quantitation reagent. Because this reagent binds to nucleic acids, it should be treated as a potential mutagen and handled with appropriate care. The DMSO stock solution should be handled with particular caution, as DMSO is known to facilitate the entry of organic molecules into tissues. It is strongly recommended that double gloves be used when handling the DMSO stock solution. As with all nucleic acid reagents, solutions of PicoGreen reagent should be poured through activated charcoal before disposal. The charcoal must then be incinerated to destroy the dye.

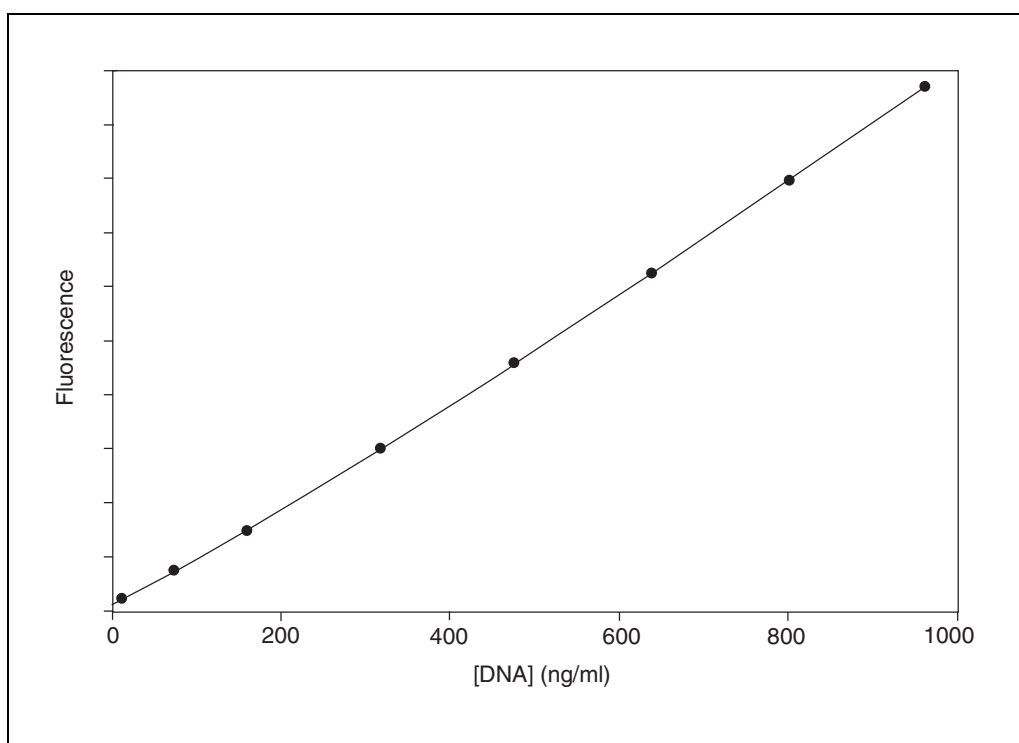


Figure A.3D.1 Dynamic range and sensitivity of the PicoGreen dsDNA quantitation assay. Calf thymus DNA was added to cuvettes containing PicoGreen dsDNA quantitation reagent diluted in 10 mM Tris-Cl/1 mM EDTA, pH 7.5 (TE). The samples were excited at 480 nm and the fluorescence emission intensity was measured at 520 nm using a spectrofluorometer. Fluorescence emission intensity was then plotted versus DNA concentration.

Commonly Used Techniques

A.3D.5

Additional Materials (also see Basic Protocol)

PicoGreen dsDNA quantitation kit containing:

PicoGreen dsDNA quantitation reagent (Component A), 1 ml solution in DMSO (store frozen up to 6 months at -20°C , protected from light)

20 \times TE (Component B), 25 ml of 200 mM Tris-Cl/20 mM EDTA, pH 7.5 (store up to 6 months at 4°C ; may be frozen for long-term storage)

Lambda DNA standard (Component C), 1 ml of 100 $\mu\text{g}/\text{ml}$ in TE (store up to 6 months at 4°C ; may be frozen for long-term storage)

Spectrofluorometer or fluorescence microplate reader

NOTE: For either the kits or the stand-alone reagent, sufficient reagent is supplied for 200 assays using an assay volume of 2 ml according to the protocol below. Note that the assay volume is dependent on the instrument used to measure fluorescence; with a microplate reader and a 96-well microplate, the assay volume is reduced to 200 μl and 2000 assays are possible. The PicoGreen reagent supplied in the kits is exactly the same as the reagent sold separately. The DMSO stock solution should be stored frozen at -20°C and protected from light. The 20 \times assay buffer and lambda DNA standard in the kits are best stored at 4°C ; however, either may be frozen for long-term storage. When properly stored, components should be stable for at least 6 months.

Prepare the reagent

1. On day of experiment, prepare an aqueous working solution of the PicoGreen reagent by making a 200-fold dilution of the concentrated DMSO solution in 1 \times TE.

Allow the PicoGreen reagent to warm to room temperature before opening the vial. Because the PicoGreen dye is an extremely sensitive detection reagent for dsDNA, it is imperative that the TE solution be free of contaminating nucleic acids. The 20 \times TE buffer included in the PicoGreen dsDNA Quantitation Kits is certified to be nucleic acid-free and DNase-free. Prepare a 1 \times TE working solution by diluting the concentrated buffer 20-fold with sterile, distilled, DNase-free water.

To prepare enough working solution to assay 20 samples in a 2-ml final volume, add 100 μl PicoGreen dsDNA quantitation reagent to 19.9 ml TE. The author recommends preparing this solution in a plastic container rather than glass, as the reagent may adsorb to glass surfaces. Protect the working solution from light by covering it with foil or placing it in the dark, as the PicoGreen reagent is susceptible to photodegradation. For best results, this solution should be used within a few hours of its preparation.

Establish the DNA standard curve

2. Prepare a 2 $\mu\text{g}/\text{ml}$ stock solution of dsDNA in 1 \times TE. Determine the DNA concentration on the basis of absorbance at 260 nm (A_{260}) in a cuvette with a 1-cm pathlength (see Basic Protocol); an A_{260} of 0.04 corresponds to 2 $\mu\text{g}/\text{ml}$ dsDNA solution.

For a standard curve, the author commonly uses bacteriophage lambda DNA (provided with kit) or calf thymus DNA, although any purified dsDNA preparation may be used. The lambda DNA standard, provided at 100 $\mu\text{g}/\text{ml}$ in the PicoGreen Kits, can simply be diluted 50-fold in 1 \times TE to make the 2 $\mu\text{g}/\text{ml}$ working solution. For example, 30 μl of the DNA standard mixed with 1.47 ml of TE will be sufficient for the standard curve described below. It is sometimes preferable to prepare the standard curve with DNA similar to the type being assayed; e.g., long or short linear DNA fragments when quantitating similar-sized restriction fragments or plasmid when quantitating plasmid DNA. However, most linear dsDNA molecules yield approximately equivalent signals, regardless of fragment length. Results have shown that the PicoGreen assay remains linear in the presence of several compounds that commonly contaminate nucleic acid preparations, although the signal intensity may be affected (Table A.3D.3). Thus, to serve as an effective control, the dsDNA solution used to prepare the standard curve should be treated the same way as the experimental samples and should contain similar levels of such compounds.

Table A.3D.3 Effects of Several Compounds that Commonly Contaminate Nucleic Acid Preparations on the Signal Intensity of the PicoGreen dsDNA Quantitation Assay

Compound	Concentration	% Signal change ^a
<i>Salts</i>		
Ammonium acetate	50 mM	3% decrease
Sodium acetate	30 mM	3% increase
Sodium chloride	200 mM	30% decrease
Zinc chloride	5 mM	8% decrease
Magnesium chloride	50 mM	33% decrease
Urea	2 M	9% increase
<i>Organic solvents</i>		
Phenol	0.1%	13% increase
Ethanol	10%	12% increase
Chloroform	2%	14% increase
<i>Detergents</i>		
Sodium dodecyl sulfate	0.01%	1% decrease
Triton X-100	0.1%	7% increase
<i>Proteins</i>		
Bovine serum albumin	2%	16% decrease
IgG	0.1%	19% increase
<i>Other compounds</i>		
Polyethylene glycol	2%	8% increase
Agarose	0.1%	4% increase

^aThe compounds were incubated at the indicated concentrations with PicoGreen reagent in the presence of 500 ng/ml calf thymus DNA. All samples were assayed in a final volume of 200 μ l in 96-well microplates using a CytoFluor microplate reader (PerSeptive Biosystems). Samples were excited at 485 nm and fluorescence intensity was measured at 520 nm.

- 3a. *For high-range standard curve:* Create a five-point standard curve from 1 ng/ml to 1 μ g/ml by combining the 2 μ g/ml stock prepared in step 2 with 1 \times TE, in disposable cuvettes (or plastic test tubes for transfer to quartz cuvettes), according to Table A.3D.4.
- 3b. *For low-range standard curve:* Prepare a 40-fold dilution of the 2 μ g/ml DNA solution to yield a 50 ng/ml DNA stock solution. Create a five-point standard curve from 25 pg/ml to 25 ng/ml by combining this 50 ng/ml stock with 1 \times TE in disposable cuvettes (or plastic test tubes for transfer to quartz cuvettes), according to Table A.3D.4.

To create the low-range standard curve, adjust the fluorometer gain to accommodate the lower fluorescence signals.
4. Add 1.0 ml of the aqueous working solution of PicoGreen reagent (prepared in step 1) to each cuvette. Mix well and incubate for 2 to 5 min at room temperature, protected from light.
5. After incubation, measure sample fluorescence using a spectrofluorometer or fluorescence microplate reader and standard fluorescein wavelengths (excitation \sim 480

Table A.3D.4 Preparing Standard Curve with the PicoGreen Reagent (Alternate Protocol 3)

Vol. (μl) DNA stock	Vol. (μl) 1× TE	Vol. (μl) PicoGreen working solution	Final DNA conc. in PicoGreen assay
<i>High-range (stock contains 2 μg/ml DNA)^a</i>			
1000	0	1000	1 μg/ml
100	900	1000	100 ng/ml
10	990	1000	10 ng/ml
1	999	1000	1 ng/ml
0	1000	1000	Blank
<i>Low range (stock contains 50 ng/ml DNA)^b</i>			
1000	0	1000	25 ng/ml
100	900	1000	2.5 ng/ml
10	990	1000	250 pg/ml
1	999	1000	25 pg/ml
0	1000	1000	Blank

^aSee Alternate Protocol 3, step 3a.^bSee Alternate Protocol 3, step 3b.

nm, emission ~520 nm). To ensure that the sample readings remain in the detection range of the fluorometer, set the instrument's gain so that the sample containing the highest DNA concentration yields a fluorescence intensity near the fluorometer's maximum. To minimize photobleaching effects, keep the time for fluorescence measurement constant for all samples.

- Subtract the fluorescence value of the reagent blank from that of each of the samples. Use corrected data to generate a standard curve of fluorescence versus DNA concentration (see Fig. A.3D.1).

Analyze samples

- Add 1.0 ml of the aqueous working solution of the PicoGreen reagent (prepared in step 1) to each sample. Incubate 2 to 5 min at room temperature, protected from light.
- Measure fluorescence of sample using instrument parameters that correspond to those used when generating standard curve (see steps 2 to 6). To minimize photobleaching effects, keep time for fluorescence measurement constant for all samples.
- Subtract fluorescence value of reagent blank from that of each sample. Determine DNA concentration of the sample from standard curve.
- If desired, repeat assay using a different dilution of the sample to confirm results.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Calf thymus DNA standard solutions

Kits containing calf thymus DNA standard for fluorometry are available (Fluorometry Reference Standard Kits, Hoefer). Premeasured, CsCl-gradient-purified DNA of defined GC content, for use in absorption and fluorometric spectroscopy, is available from Sigma (e.g., calf thymus DNA, 42% GC; *Clostridium perfringens* DNA, 26.5% GC).

Ethidium bromide assay solution

Add 10 ml of 10× TNE buffer (see recipe) to 89.5 ml H₂O. Filter through a 0.45-μm filter, then add 0.5 ml of 1 mg/ml ethidium bromide.

Add the dye after filtering, as ethidium bromide will bind to most filtration membranes.

CAUTION: *Ethidium bromide is hazardous; wear gloves and use appropriate care in handling, storage, and disposal.*

Hoechst 33258 assay solutions

Stock solution: Dissolve in H₂O at 1 mg/ml. Stable for ~6 months at 4°C.

Working solution: Add 10 ml of 10× TNE buffer (see recipe) to 90 ml H₂O. Filter through a 0.45-μm filter, then add 10 μl of 1 mg/ml Hoechst 33258.

Hoechst 33258 is a fluorochrome dye with a molecular weight of 624 and a molar extinction coefficient of $4.2 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ at 338 nm.

The dye is added after filtering because it will bind to most filtration membranes.

CAUTION: *Hoechst 33258 is hazardous; use appropriate care in handling, storage, and disposal.*

TNE buffer, 10×

100 mM Tris base

10 mM EDTA

2.0 M NaCl

Adjust pH to 7.4 with concentrated HCl

As needed, dilute with H₂O to desired concentration

COMMENTARY

Background Information

In deciding what method of nucleic acid measurement is appropriate, three issues are critical: specificity, sensitivity, and interfering substances. Properties of the four assays described in this section are listed in Table A.3D.5. The traditional method for determining the amount of DNA in solution is by measuring absorbance at 260 nm. Because many potential contaminants of DNA and RNA preparations also absorb in the UV range, absorption spectroscopy is a reliable method to assess both the purity of a preparation and the quantity of DNA or RNA present. Absorption spectroscopy does have serious limitations. Relatively large amounts of DNA are required to get accurate readings—for example, 500 ng/ml DNA is equivalent to only 0.01 A_{260} units. Furthermore, the method cannot discriminate between RNA and DNA, and UV-absorbing contaminants such as protein will cause discrepancies.

The assay using Hoechst 33258 dye (Alternate Protocol 1) is the only procedure in common use that is specific for DNA (i.e., it does not measure RNA). This assay is the method of choice for rapid measurement of low quantities of DNA, with a detection limit of ~1 ng DNA. Concentrations of DNA in both crude cell lysates and purified preparations can be quan-

tified (Labarca and Paigen, 1980). Because the assay quantifies a broad range of DNA concentrations—from 10 ng/ml to 15 μg/ml—it is useful for the measurement of both small and large amounts of DNA (e.g., in verifying DNA concentrations prior to performing electrophoretic separations and Southern blots). The Hoechst 33258 assay is also useful for measuring products of the polymerase chain reaction (PCR) synthesis.

Upon binding to DNA, the fluorescence characteristics of Hoechst 33258 change dramatically, showing a large increase in emission at ~458 nm. Hoechst 33258 is nonintercalating and apparently binds to the minor groove of the DNA, with a marked preference for AT sequences (Portugal and Waring, 1988). The fluorochrome 4',6-diamidino-2-phenylindole (DAPI; Daxhelet et al., 1989) is also appropriate for DNA quantitation, although it is not as commonly used as Hoechst 33258. DAPI is excited with a peak at 344 nm. Emission is detected at ~466 nm, similar to Hoechst 33258.

Ethidium bromide is best known for routine staining of electrophoretically separated DNA and RNA, but it can also be used to quantify both DNA and RNA in solution (Le Pecq, 1971). Unlike Hoechst 33258, ethidium bromide fluorescence is not significantly impaired

Table A.3D.5 Properties of Absorbance and Fluorescence Spectrophotometric Assays for DNA and RNA

Property	Absorbance (A_{260})	Fluorescence		
		H33258	Ethidium bromide	PicoGreen
<i>Sensitivity</i>				
DNA	1-50 $\mu\text{g/ml}$	0.01-15 $\mu\text{g/ml}$	0.1-10 $\mu\text{g/ml}$	25-1000 ng/ml
RNA	1-40 $\mu\text{g/ml}$	NA	0.2-10 $\mu\text{g/ml}$	Minimal sensitivity
<i>Ratio of signal (DNA/RNA)</i>	0.8	400	2.2	>100

by high GC content. The ethidium bromide assay (with excitation at 546 nm) is ~20-fold less sensitive than the Hoechst 33258 assay.

In addition to the advantages mentioned in the protocol itself (see Alternate Protocol 3), the PicoGreen assay protocol also minimizes the fluorescence contribution of RNA and single-stranded DNA (ssDNA). Although the Hoechst 33258-based method is not significantly affected by the presence of RNA when the assay is carried out in the recommended high-salt buffer, Hoechst 33258 does exhibit a large fluorescence enhancement with ssDNA under these conditions. Furthermore, when the Hoechst 33258-based assay is carried out in TE alone (10 mM Tris-Cl/1 mM EDTA, pH 7.5, with no NaCl added), RNA contributes a significant fluorescence signal. Using the PicoGreen dsDNA quantitation reagent as described in Alternate Protocol 3, researchers can quantitate dsDNA in the presence of equimolar concentrations of ssDNA and RNA with minimal effect on the quantitation results.

Critical Parameters

Care should be taken when handling sample cuvettes in all spectrophotometric procedures. Fluorometers use cuvettes with four optically clear faces, because excitation and emitting light enter and leave the cuvette through directly adjacent sides. Thus, fluorometric cuvettes should be held by the upper edges only. In contrast, transmission spectrophotometers use cuvettes with two opposite optical windows, with the sides frosted for easy handling. It is important to check that the optical faces of cuvettes are free of fingerprints and scratches. In addition, for accurate absorbance readings, spectrophotometer cuvettes must be perfectly matched.

Proteins in general have A_{280} readings considerably lower than nucleic acids on an equiva-

lent weight basis. Thus, even a small increase in the A_{280} relative to A_{260} (i.e., a lowering of the A_{260}/A_{280} ratio) can indicate severe protein contamination. Other commonly used buffer components absorb strongly at 260 nm and can cause interference if present in high enough concentrations. EDTA, for example, should not be present at ≥ 10 mM.

Sensitivity of the Hoechst 33258 fluorescence assay decreases with nuclease degradation, increasing GC content, or denaturation of DNA (Labarca and Paigen, 1980; Stout and Becker, 1982). Increased temperature of the assay solution and ethidium bromide contamination also decrease the Hoechst 33258 signal. Sodium dodecyl sulfate (>0.01% final concentration) also interferes with accurate readings (Cesarone et al., 1979). The pH of the assay solution is critical to sensitivity and should be pH ~7.4 (Labarca and Paigen, 1980; Stout and Becker, 1982). At a pH <6.0 or >8.0 the background becomes much higher and there is a concomitant loss of fluorescence enhancement.

High-quality double-stranded DNA is recommended, although single-stranded genomic DNA also works well with this assay. However, with very small fragments of DNA, the Hoechst 33258 dye binds to double-stranded DNA only. Thus, the assay will not work with single-stranded oligomers. Linear and circular DNA give approximately the same levels of fluorescence (Daxhelet et al., 1989). When preparing DNA standards, an attempt should be made to equalize the GC content of the standard DNA and that of the sample DNA. In most situations, salmon sperm or calf thymus DNA is suitable. An extensive list of estimated GC content for various organisms is available (Marmur and Doty, 1962). Eukaryotic cells vary somewhat in GC content but are generally in the range of 39% to 46%. Within this range, the fluorescence per microgram of DNA does not vary

Table A.3D.6 Sensitivity of the PicoGreen dsDNA Assay for Quantitating dsDNA in the Presence of Single-Stranded Nucleic Acids

[dsDNA] ^a	[RNA]	Amount of RNA relative to dsDNA	[ssDNA]	Amount of ssDNA relative to dsDNA
1 µg/ml	10 µg/ml	10×	300 ng/ml	0.3×
500 ng/ml	500 ng/ml	1×	50 ng/ml	0.1×
10 ng/ml	100 ng/ml	10×	30 ng/ml	3×
5 ng/ml	50 ng/ml	10×	15 ng/ml	3×
100 pg/ml	1 ng/ml	10×	1 ng/ml	10×
50 pg/ml	500 pg/ml	10×	500 pg/ml	10×

^aFor several concentrations of dsDNA, the concentration of RNA or ssDNA is shown that results in no more than a 10% increase in the sample's signal intensity.

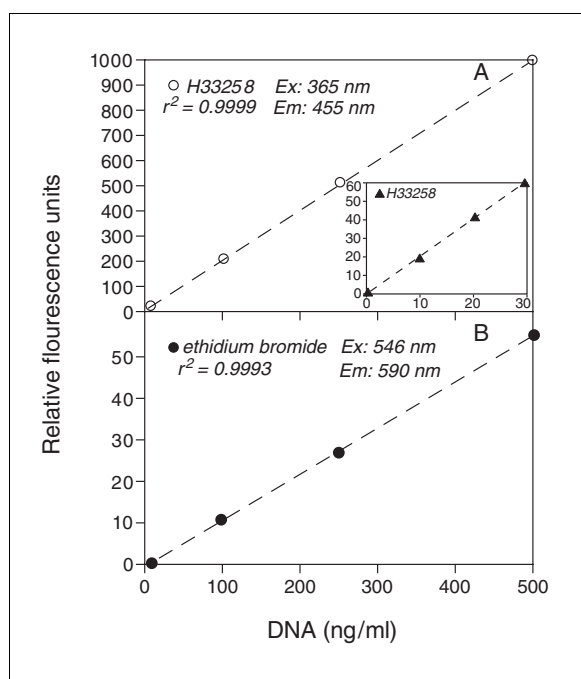


Figure A.3D.2 Fluorochrome H33258 (A) and ethidium bromide (B) DNA concentration standard curves. Assays were performed as described in alternate protocols, at indicated excitation and emission wavelengths. The concentrations of H33258 and ethidium bromide were 0.1 and 5 µg/ml, respectively. Assays contained the indicated concentrations of calf thymus DNA standards suspended in a final volume of 2.0 ml. Inset shows low DNA concentration curve for the H33258 assay. Note that, under these conditions, H33258 produces ~20 times more relative fluorescence units than ethidium bromide. A Shimadzu RF-5000 scanning fluorescence spectrometer was used for both assays.

substantially. In contrast, the GC content of prokaryotes can vary from 26% to 77%, causing considerable variation in the fluorescence signal. In these situations, the sample DNA should be first quantitated via transmission spectroscopy and compared to a readily available standard (e.g., calf thymus DNA). Future measurements would then use calf thymus as a standard, but with a correction factor for difference in fluorescence yield between the two DNA types. For further troubleshooting information, see Van Lancker and Gheysens (1986), in which the effects of interfering substances on the Hoechst 33258 assay (and several other assays) are compared.

In the ethidium bromide assay, single-stranded DNA gives approximately half the

signal as double-stranded calf thymus DNA. Ribosomal RNA also gives about half the fluorescent signal as double-stranded DNA, and RNase and DNase both severely decrease the signal. Closed circular DNA also binds less ethidium bromide than nicked or linear DNA. Further critical parameters of the ethidium bromide assay are described by Le Pecq (1971).

With PicoGreen, dsDNA can be quantitated in the presence of equimolar concentrations of single-stranded nucleic acids with minimal interference. Table A.3D.6 shows the concentrations of RNA or ssDNA that, for a given dsDNA concentration, result in less than a 10% change in the signal intensity using the PicoGreen assay protocol. Fluorescence due to PicoGreen reagent binding to RNA at high concentrations

can be eliminated by treating the sample with DNase-free RNase. The use of RNase A/RNase T1 with S1 nuclease will eliminate all single-stranded nucleic acids and ensure that the entire sample fluorescence is due to dsDNA.

Anticipated Results

The detection limit of absorption spectroscopy will depend on the sensitivity of the spectrophotometer and any UV-absorbing contaminants that might be present. The lower limit is generally ~0.5 to 1 µg nucleic acid. Typical values for a highly purified sample of DNA are shown in Table A.3D.2.

For the Hoechst 33258, ethidium bromide, and PicoGreen assays, a plot of relative fluorescence units or estimated concentration (y axis) versus actual concentration (x axis) typically produces a linear regression with a correlation coefficient (r^2) of 0.98 to 0.99 (Fig. A.3D.2). Table A.3D.3 provides a comparison of the sensitivities and specificities of the three assays.

Time Considerations

The three assays described can be performed in a short period of time. In a well-planned series of assays, 50 samples can be prepared and read comfortably in 1 hr. Although some error might be introduced, DNA samples can be sequentially added to the same cuvette containing working dye solution. The increase in fluorescence with each sample is noted and subtracted from the previous reading to give relative fluorescence or concentration of the new sample, eliminating the need to change solutions for each sample. Be certain that the final amount of DNA does not exceed the linear portion of the assay.

Literature Cited

- Applied Biosystems. 1987. User Bulletin Issue 11, Model No. 370. Applied Biosystems, Foster City, Calif.
- Cesarone, C.F., Bolognesi, C., and Santi L. 1979. Improved microfluorometric DNA determina-

tion in biological material using 33258 Hoechst. *Anal. Biochem.* 100:188-197.

Daxhelet, G.A., Coene, M.M., Hoet, P.P., and Cocito, C.G. 1989. Spectrofluorometry of dyes with DNAs of different base composition and conformation. *Anal. Biochem.* 179:401-403.

Labarca, C. and Paigen, K. 1980. A simple, rapid, and sensitive DNA assay procedure. *Anal. Biochem.* 102:344-352.

Le Pecq, J.-B. 1971. Use of ethidium bromide for separation and determination of nucleic acids of various conformational forms and measurement of their associated enzymes. *In Methods of Biochemical Analysis*, Vol. 20 (D. Glick, ed.) pp. 41-86. John Wiley & Sons, New York.

Marmur, J. and Doty, P. 1962. Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J. Molec. Biol.* 5:109-118.

Portugal, J. and Waring, M.J. 1988. Assignment of DNA binding sites for 4',6-diamidino-2-phenylindole and bisbenzimidazole (Hoechst 33258): A comparative footprinting study. *Biochem. Biophys. Acta* 949:158-168.

Stout, D.L. and Becker, F.F. 1982. Fluorometric quantitation of single-stranded DNA: A method applicable to the technique of alkaline elution. *Anal. Biochem.* 127:302-307.

Van Lancker, M. and Gheysens, L.C. 1986. A comparison of four frequently used assays for quantitative determination of DNA. *Anal. Lett.* 19:615-623.

Wallace, R.B. and Miyada, C.G. 1987. Oligonucleotide probes for the screening of recombinant DNA libraries. *In Methods of Enzymology*, Vol. 152: Guide to Molecular Cloning Techniques (S.L. Berger and A.R. Kimmel, eds.) pp. 432-442. Academic Press, San Diego.

Key References

Labarca and Paigen, 1980. See above.

Contains a detailed description of the Hoechst 33258 fluorometric DNA assay.

Contributed by Sean R. Gallagher
UVP, Inc.
Upland, California

Introduction of Restriction Enzyme Recognition Sequences by Silent Mutation

Site-directed mutagenesis has been an invaluable tool in understanding structure-function relationships. A number of techniques have been developed for introducing site-directed mutations. These techniques exploit in vitro and in vivo features of enzyme systems that act on nucleic acids (see Chapter 8) and have been immensely aided by development of the polymerase chain reaction (PCR; UNIT 15.7). Any changes in primer sequence can be easily incorporated into the DNA product of PCR; this facilitates subsequent incorporation of the changes into the gene sequence.

Sequence analysis and other strategies have identified biologically important motifs within proteins. This knowledge together with the tools of recombinant DNA technology can be used to design, construct, and express hybrid sequences that have a particular function. For example, development of multivalent recombinant vaccines requires design of a protein with immune-neutralizing domains from several pathogens expressed as a fusion protein. Alternatively, very small regulatory sequences can be introduced as part of the coding sequence to develop novel products that exert multiple effects as products of transcription and translation. Thus it is often necessary to generate gene cassette structures that encompass these motifs to facilitate manipulation of each motif for subsequent studies. Restriction enzyme recognition sequences are rarely found flanking such biologically important domains and motifs, and it is often desirable to introduce restriction sites at flanking regions. Because the restriction enzyme recognition sequence will form part of the coding sequence, it is imperative that the introduction of restriction sites does not alter the amino acid sequence of the encoded protein. The following strategy enables identification of a large number of potential sites where restriction enzyme recognition sequences can be introduced via silent mutations.

The first step for introducing a restriction enzyme recognition sequence by silent mutagenesis involves translating the sequence in all the three reading frames (Shankarappa et al., 1992a; Tables A.3E.1 and A.3E.2). Due to degeneracy of the genetic code, translation of a six-base recognition sequence yields a large number of di- and tripeptides. Next, the sequence of the protein to be modified is analyzed for the presence of any of these peptides. The presence of a peptide indicates that it is possible to introduce the corresponding restriction enzyme recognition sequence at that site by silent mutagenesis. For example, translation of the recognition sequence for *Pst*I (CTGCAG) yields a dipeptide Leu-Gln. However, translation of the same sequence in the second and third reading frames yields an additional 20 and 30 tripeptide sequences respectively. Hence, it is now possible to introduce *Pst*I recognition sequence at any of the 51 di- or tripeptide motifs by silent mutagenesis. Similarly, an analysis of the recognition sequences for *Not*I, an eight-base restriction enzyme, identifies 75 different tri- and tetrapeptide sequences that have the potential to be changed by silent mutations to incorporate the recognition sequence.

Introduction of restriction enzyme recognition sequences by silent mutations also provides a simple strategy to generate unique cohesive sites at the ends of DNA fragments (see also UNIT 3.17). It is possible to ligate several fragments containing unique overhangs at their ends, thus generating a large fragment containing all the desired mutations (Shankarappa et al., 1992b). Introduction of silent mutation can be coupled with introduction of a second mutation downstream of the restriction site. Digestion of the DNA fragment generated by PCR using such a primer will generate unique cohesive sites at the ends to facilitate subsequent ligation.

Contributed by Raj Shankarappa

Current Protocols in Molecular Biology (1996) A.3E.1-A.3E.3

Copyright © 1998 by John Wiley & Sons, Inc.

Commonly Used
Techniques

A.3E.1

Supplement 35

Table A.3E.1 Translation Sequences for Silent Mutagenesis Based on Six-Base Restriction Enzymes

Restriction enzyme ^a	Recognition sequence	First reading frame		Second reading frame			Third reading frame		
		aa 1 ^{b,c}	aa 2	aa 1	aa 2	aa 3	aa 1	aa 2	aa 3
<i>Hind</i> III	AAGCTT	K	L	XQKE	A	FLSYXCW	LSXPQRITKVAEG	S	FL
<i>Mlu</i> I	ACGCGT	T	R	YHND	A	FLSYXCW	LSXPQRITKVAEG	R	V
<i>Spe</i> I	ACTAGT	T	S	YHND	X	FLSYXCW	LSXPQRITKVAEG	L	V
<i>Bgl</i> II	AGATCT	R	S	XQKE	I	FLSYXCW	LSXPQRITKVAEG	D	L
<i>Stu</i> I	AGGCCT	R	P	XQKE	A	FLSYXCW	LSXPQRITKVAEG	G	L
<i>Bsp</i> DI/ <i>Cl</i> AI	ATCGAT	I	D	YHND	R	FLSYXCW	LSXPQRITKVAEG	S	IM
<i>Pvu</i> II	CAGCTG	Q	L	SPTA	A	VADEG	FSYCLPHRITNVADG	S	CXW
<i>Nde</i> I	CATATG	H	M	SPTA	Y	VADEG	FSYCLPHRITNVADG	I	CXW
<i>Nco</i> I	CCATGG	P	W	SPTA	M	VADEG	FSYCLPHRITNVADG	H	G
<i>Sma</i> I/ <i>Xma</i> I	CCCGGG	P	G	SPTA	R	VADEG	FSYCLPHRITNVADG	P	G
<i>Sac</i> II	CCGCGG	P	R	SPTA	A	VADEG	FSYCLPHRITNVADG	R	G
<i>Pvu</i> I	CGATCG	R	S	SPTA	I	VADEG	FSYCLPHRITNVADG	D	R
<i>Eag</i> I/ <i>Xma</i> III	CGGCCG	R	P	SPTA	A	VADEG	FSYCLPHRITNVADG	G	R
<i>Pae</i> R7I/ <i>Xho</i> I	CTCGAG	L	E	SPTA	R	VADEG	FSYCLPHRITNVADG	S	SR
<i>Pst</i> I	CTGCAG	L	Q	SPTA	A	VADEG	FSYCLPHRITNVADG	C	SR
<i>Eco</i> RI	GAATTC	E	F	XRG	I	LPHQR	LSXWPQRMTKVAEG	N	S
<i>Sac</i> I	GAGCTC	E	L	XRG	A	LPHQR	LSXWPQRMTKVAEG	S	S
<i>Eco</i> RV	GATATC	D	I	XRG	Y	LPHQR	LSXWPQRMTKVAEG	I	S
<i>Sph</i> I	GCATGC	A	C	CRSG	M	LPHQR	LSXWPQRMTKVAEG	H	A
<i>Nae</i> I	GCCGGC	A	G	CRSG	R	LPHQR	LSXWPQRMTKVAEG	P	A
<i>Nhe</i> I	GCTAGC	A	S	CRSG	X	LPHQR	LSXWPQRMTKVAEG	L	A
<i>Bam</i> HI	GGATCC	G	S	WRG	I	LPHQR	LSXWPQRMTKVAEG	D	P
<i>Nar</i> I	GGCGCC	G	A	WRG	R	LPHQR	LSXWPQRMTKVAEG	A	P
<i>Apa</i> I	GGGCCC	G	P	WRG	A	LPHQR	LSXWPQRMTKVAEG	G	P
<i>Acc</i> 65I/ <i>Kpn</i> I	GGTACC	G	T	WRG	Y	LPHQR	LSXWPQRMTKVAEG	V	P
<i>Sal</i> I	GTCGAC	V	D	CRSG	R	LPHQR	LSXWPQRMTKVAEG	S	T
<i>Apa</i> LI	GTGCAC	V	H	CRSG	A	LPHQR	LSXWPQRMTKVAEG	C	T
<i>Hpa</i> I	GTTAAC	V	N	CRSG	X	LPHQR	LSXWPQRMTKVAEG	L	T
<i>Bsp</i> EI	TCCGGA	S	G	FLIV	R	IMTNKSR	FSYCLPHRITNVADG	P	DE
<i>Nru</i> I	TCGCGA	S	R	FLIV	A	IMTNKSR	FSYCLPHRITNVADG	R	DE
<i>Xba</i> I	TCTAGA	S	R	FLIV	X	IMTNKSR	FSYCLPHRITNVADG	L	DE
<i>Bcl</i> II	TGATCA	X	S	LMV	I	IMTNKSR	FSYCLPHRITNVADG	D	HQ
<i>Bal</i> I	TGGCCA	W	P	LMV	A	IMTNKSR	FSYCLPHRITNVADG	G	HQ

^aSequences recognized by six-base restriction enzymes that do not contain degeneracy in the recognition sequence.

^bIdentity of amino acid at position 1, 2, or 3 in the translation sequence.

^cSingle-letter amino acid code (see Table A.1C.1); X indicates a stop codon.

Tables A.3E.1 and A.3E.2 list di-, tri-, and tetrapeptide translation sequences that can be used to introduce sequences recognized by common restriction enzymes that do not contain any degeneracy. A computer program, SILMUT, uses either amino acid or nucleic acid sequence to identify potential sites for silent mutations. The program is available from the author and at many anonymous ftp sites (Shankarappa et al., 1992c). When using this mutagenesis strategy, due consideration should be given to codon usage of the host organism.

**Restriction
Enzyme
Recognition by
Silent Mutation**

A.3E.2

Table A.3E.2 Translation Sequences for Silent Mutagenesis Based on Eight-Base Restriction Enzymes

Restriction enzyme ^a	Recognition sequence	First reading frame			Second reading frame			Third reading frame			
		aa 1 ^{b,c}	aa 2	aa 3	aa 1	aa 2	aa 3	aa 1	aa 2	aa 3	aa 4
<i>Swa</i> I	ATTTAAAT	I	X	IM	YHND	L	N	LSXPQRITKVAEG	F	K	FLSYXCW
<i>Sse</i> 8387I	CCTGCAGG	P	A	G	SPTA	C	R	FSYCLPHRITNVADG	L	Q	VADEG
<i>Srf</i> I	GCCCGGGC	A	R	A	CRSG	P	G	LSXWPQRMTKVAEG	P	G	LPHQR
<i>Not</i> I	GCGGCCGC	A	A	A	CRSG	G	R	LSXWPQRMTKVAEG	R	P	LPHQR
<i>Asc</i> I	GGCGCGCC	G	A	P	WRG	R	A	LSXWPQRMTKVAEG	A	R	LPHQR
<i>Pme</i> I	GTTTAAAC	V	X	T	CRSG	L	N	LSXWPQRMTKVAEG	F	K	LPHQR
<i>Pac</i> I	TTAATTAA	L	I	NK	FLIV	N	X	FSYCLPHRITNVADG	X	F	IMTNKSR

^aSequences recognized by eight-base restriction enzymes that do not contain degeneracy in the recognition sequence.

^bIdentity of amino acid at position 1, 2, 3, or 4 in the translation sequence.

^cSingle-letter amino acid code (see Table A.1C.1); X indicates a stop codon.

Literature Cited

- Shankarappa, B., Sirko, D.A., and Ehrlich, G.D. 1992a. A general method for the identification of regions suitable for site-directed silent mutagenesis. *BioTechniques* 12:382-384.
- Shankarappa, B., Balachandran, R., Gupta, P., and Ehrlich, G.D. 1992b. Introduction of multiple restriction enzyme sites by in vitro mutagenesis using the polymerase chain reaction. *PCR Method. Appl.* 1:277-278.
- Shankarappa, B., Vijayananda, K., and Ehrlich, G.D. 1992c. SILMUT: A computer program for the identification of regions compatible for the introduction of restriction enzyme sites by site-directed silent mutagenesis. *BioTechniques* 12:882-884.

Contributed by Raj Shankarappa
University of Pittsburgh School of Medicine
Pittsburgh, Pennsylvania

Techniques for Mammalian Cell Tissue Culture

Tissue culture technology has found wide application in the field of molecular biology (see, in particular, Chapter 9, Introduction of DNA into Mammalian Cells, and Chapter 16, Protein Expression). Monolayer cell cultures are utilized in cytogenetic, biochemical, and molecular laboratories for research as well as diagnostic studies. The isolation and culture of specific types of cells is described in Chapter 28 (Mammalian Cell Culture). This appendix addresses general concerns and procedures needed for successful culture of mammalian cells. In most cases, cells or tissues must be grown in culture for days or weeks to obtain sufficient numbers of cells for analysis. Maintenance of cells in long-term culture requires strict adherence to aseptic technique to avoid contamination and potential loss of valuable cell lines. The first section of this unit discusses basic principles of sterile technique.

An important factor influencing the growth of cells in culture is the choice of tissue culture medium. Many different recipes for tissue culture media are available and each laboratory must determine which medium best suits its needs. Individual laboratories may elect to use commercially prepared medium or prepare their own. Commercially available medium can be obtained as a sterile and ready-to-use liquid, in a concentrated liquid form, or in a powdered form. Besides providing nutrients for growing cells, medium is generally supplemented with antibiotics, fungicides, or both to inhibit contamination. The second section of this unit discusses medium preparation (see also the introduction to Chapter 9).

As cells grown in monolayer reach confluency, they must be subcultured or passaged. Failure to subculture confluent cells results in reduced mitotic index and eventually in cell death. The first step in subculturing is to detach cells from the surface of the primary culture vessel by trypsinization or mechanical means. The resultant cell suspension is then subdivided, or reseeded, into fresh cultures. Secondary cultures are checked for growth and fed periodically, and may be subsequently subcultured to produce tertiary cultures, etc. The time between passaging of cells varies with the cell line and depends on the growth rate.

The Basic Protocol describes subculturing of a monolayer culture grown in petri plates or flasks. Support Protocols describe freezing of monolayer cells, thawing and recovery of cells, counting cells using a hemacytometer, and preparing cells for transport. Alternate Protocols 1 and 2 describe the passaging and freezing of cells in suspension culture.

CAUTION: Radioactive, biological, and chemical substances require special handling; see *APPENDIX 1E* for guidelines.

STERILE TECHNIQUE

It is essential that sterile technique be maintained when working with cell cultures. Aseptic technique involves a number of precautions to protect both the cultured cells and the laboratory worker from infection. The laboratory worker must realize that cells handled in the lab are potentially infectious and should be handled with caution. Protective apparel such as gloves, lab coats or aprons, and eyewear should be worn when appropriate (Knutsen, 1991). Care should be taken when handling sharp objects such as

needles, scissors, scalpel blades, and glass that could puncture the skin. Sterile disposable plastic supplies may be used to avoid the risk of broken or splintered glass (Rooney and Czepulkowski, 2001).

Frequently, specimens received in the laboratory are not sterile, and cultures prepared from these specimens may become contaminated with bacteria, fungus, or yeast. The presence of microorganisms can inhibit growth, kill cell cultures, or lead to inconsistencies in test results. The contaminants deplete nutrients in the medium and may produce substances that are toxic to cells. Antibiotics and antimycotics can be used to combat potential contaminants (see Table A.3F.1). The solutions can be used to wash specimens prior to culture or to rinse contaminated cultures, and can be added to medium used for tissue culture. Antibiotics (penicillin, streptomycin, kanamycin, neomycin, or gentamycin) can be used individually or in combinations (i.e., penicillin/streptomycin/neomycin and penicillin/streptomycin/gentamycin). Likewise, fungicides such as amphotericin B (Fungizone) and mycostatin (Nystatin) can be used alone or as antibiotic/antimycotic solutions (i.e., penicillin/streptomycin/Fungizone). Particular care should be taken when using Fungizone, as it is typically very toxic to cell cultures, and adequate data are not available regarding potential adverse effects on the growth of human cells. Antibiotics and antimycotics are available from a number of vendors including Invitrogen, Sigma-Aldrich, and BioWhittaker.

All materials that come into direct contact with cultures must be sterile. Sterile disposable dishes, flasks, pipets, etc., can be obtained directly from manufacturers. Reusable glassware must be washed, rinsed thoroughly, then sterilized by autoclaving or by dry heat before reusing. With dry heat, glassware should be heated 1.5 to 2 hr at 160°C to ensure sterility. Materials that may be damaged by very high temperatures can be autoclaved 20 min at 120°C and 15 psi. All media, reagents, and other solutions that come into contact with the cultures must also be sterile; media may be obtained as a sterile liquid from the manufacturer, autoclaved if not heat-sensitive, or filter sterilized. Supplements can be added to media prior to filtration, or they can be added aseptically after filtration. Filters with 0.20- to 0.22- μm pore size should be used to remove small gram-negative bacteria from culture media and solutions.

Contamination can occur at any step in handling cultured cells. Care should be taken to maintain the sterility of petri plates, pipets, and flasks that are used for tissue culture. The use of disposable, sterile culture supplies is convenient and has virtually eliminated

Table A.3F.1 Working Concentrations of Antibiotics and Fungicides for Mammalian Cell Culture

Additive	Final concentration
Penicillin	50–100 U/ml
Streptomycin sulfate	50–100 $\mu\text{g/ml}$
Kanamycin	100 $\mu\text{g/ml}$
Gentamycin	50 $\mu\text{g/ml}$
Neomycin	50 $\mu\text{g/ml}$
Mycostatin (Nystatin)	20 $\mu\text{g/ml}$
Amphotericin B (Fungizone)	0.25 $\mu\text{g/ml}$
Pen/strep/neomycin ^a	1 \times

^aUsually supplied as 100 \times concentrated stock solution, which is diluted to 1 \times concentration in the medium.

the need to flame-sterilize instruments and vessels used for tissue culture. The initial cost of purchasing these single-use supplies outweighs the expense associated with the time and effort of washing, packaging, and sterilizing nondisposables. Good sterile technique must be followed. For example, if a sterile pipet tip should come into contact with the benchtop or other nonsterile surface, the pipet tip should be discarded and a fresh one obtained. If disposables are not an option, the necks of sterile containers, such as bottles and flasks, and the tips of pipets should be passed through a flame before the pipet is introduced into the container. After pipetting, the neck of the bottle or flask should again be flamed.

Certain instruments used for tissue culture (forceps, scissors, scalpels, and in some instances, glass bottles) may require autoclave sterilization prior to initial use. The autoclave relies on pressurized steam to destroy microorganisms. Instruments to be autoclaved should be thoroughly washed and dried, then packaged if necessary to ensure that sterility will be maintained after removal from the autoclave until use. Indicator tape or autoclave bags with indicator strips should be used to document that the items have been autoclaved. The indicator tape demonstrates that the item has been autoclaved, but does not ensure sterility.

Historically, if instruments such as forceps, tweezers, scissors, or scalpels were going to be re-used to handle several sequential specimens, they were rapidly sterilized between uses by dipping in 70% alcohol and flaming. Other methods of rapid decontamination are now available. For example, the Germinator 500 (Cellpoint Scientific) is a small benchtop unit [height, 6.76 in. (~17 cm) × width, 5.25 in. (~13 cm) × depth 5.25 in. (~13 cm)] used in tissue culture labs for decontamination of metal dissecting instruments. Instruments must be clean, dry, and free of debris prior to sterilization. The instruments are inserted into a stainless steel well [inside dimensions: diameter, 2 in. (~5 cm) × depth, 4 in. (~10 cm)] filled with glass beads that are maintained at a constant temperature of 500°F (~260°C). Dry heat decontamination occurs through heat transfer from the glass beads to the instruments. The inserted parts of small instruments are dry sterilized within ~15 sec, while larger instruments may take as long as 1 min. The Germinator 500 eliminates the need for alcohol and open flames for sterilization. Due to the lack of a method for routine monitoring of sterilization by glass beads, the Germinator 500 operating manual recommends that this equipment be used for research purposes only.

Although tissue culture work can be done on an open bench if aseptic methods are strictly enforced, many labs prefer to perform tissue culture work in a room or low-traffic area reserved specifically for that purpose. At the very least, biological safety cabinets are recommended to protect the cultures as well as the laboratory worker. In a laminar flow hood, the flow of air protects the work area from dust and contamination and acts as a barrier between the work surface and the worker. Many different styles of safety hoods are available, and the laboratory should consider the types of samples being processed and the types of potential pathogenic exposure in making a selection. Manufacturer recommendations should be followed regarding routine maintenance checks on air flow and filters. For day-to-day use, the cabinet should be turned on for at least 5 min prior to beginning work. All work surfaces both inside and outside of the hood should be kept clean and disinfected daily and after each use. A 10% household bleach solution, 70% alcohol, an iodophor, a quarternary ammonium compound, or commercially available liquid disinfectants can be used (<http://www.ehs.cornell.edu/bio/cabinets.htm>, last updated December, 2000).

Some safety cabinets are equipped with ultraviolet (UV) lights for decontamination of work surfaces. However, UV lamps are not required in biological safety cabinets, and their effectiveness has been questioned (<http://www.niehs.nih.gov/odhsb/>

[biosafe/bsc/section6.htm](http://www.biosafe/bsc/section6.htm); <http://www.ehs.cornell.edu/bio/cabinets.htm>). UV lamps may produce a false sense of security as they maintain a visible blue glow long after their germicidal effectiveness is lost. Effectiveness diminishes over time as the glass tube gradually loses its ability to transmit short UV wavelengths, and may also be reduced by dust on the glass tube, distance from the work surface, temperature, and air movement. Even when the UV output is adequate, the rays must directly strike a microorganism in order to kill it; bacteria or mold spores hidden below the surface of a material or outside the direct path of the rays will not be destroyed. Another rule of thumb is that anything that can be seen cannot be killed by UV. UV lamps will only destroy microorganisms such as bacteria, viruses, and mold spores; they will not destroy insects or other large organisms. Aside from their general ineffectiveness, UV lights pose a safety hazard, as exposure can cause damage to the eyes and skin (http://www.yale.edu/oehs/cad2_0.htm). A more reliable approach to eliminate contamination is the use of well-practiced microbiological procedures, good aseptic techniques, and standard operational procedures for working in a biological safety cabinet, including thorough decontamination with an effective disinfectant before and after use of the biological safety cabinet. The current recommendation is that work surfaces be wiped down with ethanol instead of relying on UV lamps, although some labs use the lamps in addition to ethanol wipes to decontaminate work areas. If UV lamps are used as a secondary method of disinfecting the work surface, their radiation output should be tested with a UV meter during the annual certification of the biological safety cabinet to ensure that the proper intensity of light is reaching the work surface. The radiation output should be at least $40 \mu\text{W}/\text{cm}^2$ at 254 nm when measured with a UV flux meter placed in the center of the work surface. The UV lamps should be replaced when they fall below the minimum requirements for protection (<http://www.ehs.cornell.edu/bio/cabinets.htm>).

Cultures should be visually assessed on a routine basis for evidence of contamination. Indicators in the tissue culture medium change color when contamination is present: for example, medium that contains phenol red changes to yellow because of increased acidity. Cloudiness and turbidity are also observed in contaminated cultures. Once contamination is confirmed with a microscope, infected cultures are generally discarded. Keeping contaminated cultures increases the risk of contaminating other cultures. Sometimes a contaminated cell line can be salvaged by treating it with various combinations of antibiotics and antimycotics in an attempt to eradicate the infection. In this procedure, the tissue culture medium is aspirated from the affected cultures and discarded. The cultures are rinsed with fresh “wash” medium that has been prepared by supplementing the routinely used tissue culture medium with the appropriate concentration of antibiotic and/or antimycotic (Table A.3F.1). The petri dish or flask should be gently swirled so that the cell surface is bathed in the clean medium. The wash medium can be left on the cultures for 1 to 5 min. Due to the toxicity of Fungizone and amphotericin, it is recommended that cell exposure to these agents be limited to less than 2 min. Aspirate the wash medium, replace with fresh medium and return the cultures to the incubator. Even under the best conditions, such treatment may adversely affect cell growth and it is often unsuccessful in ridding cultures of contaminants.

Mycoplasma are small (0.2 to 0.3 μm) intracellular bacteria that attach to the cell membrane, inhibit cell growth, and eventually lead to cell death. Because these parasites do not have cell walls, do not grow in colonies, and do not change the pH of the medium, they are difficult to detect visually in cultures. Mycoplasma can multiply to very high concentrations (10^7 to 10^8 organisms/ml) and adversely affect cultures by altering cell growth characteristics, inhibiting cell metabolism, disrupting nucleic acid synthesis, inducing chromosome aberrations, changing cell membrane antigenicity, and altering transfection rates and viral susceptibility (<http://www.unc.edu/depts/tcf/mycoplasma.htm>).

Mycoplasma are spread by cross-contamination from infected cultures through aerosolization during pipetting, or via the transfer of contaminated cells or contaminated reagents used in cell culture (<http://www.unc.edu/depts/tcf/mycoplasma.htm>). There are several methods of detecting mycoplasma in cell cultures and in cell culture reagents. In the direct method, both cultures and reagents can be plated on agar in order to grow and thereby reveal the presence of mycoplasma. Disadvantages of this method are that some strains of mycoplasma cannot be cultivated on agar and that the results can take from 2 to 4 weeks to obtain (<http://www.lerner.ccf.org/services/cell/cellculture.php>). Several indirect methods are available and should be selected based on the capabilities of the laboratory and the needs of the laboratory in terms of sensitivity, specificity, and time requirements. The DNA staining method uses Hoechst 33258 to highlight the A-T rich DNA of mycoplasma, so that mycoplasma appear as bright extranuclear spots in the cytoplasm. Results are available in 24 hr, but this method is not as sensitive as the direct method and can be difficult to interpret due to background bacterial/yeast/fungal contamination, excess debris, reduced or absent live cells, and broken nuclei from dead cells. The addition of an indicator such as Vero (Sigma-Aldrich) increases the sensitivity of this method. PCR primers that selectively amplify part of the mycoplasma DNA are also available. Kits include the ELISA mycoplasma detection kit from Roche Applied Science, the Mycoplasma Detection kit from Minerva Biolabs, the Myco Alert Mycoplasma Detection Assay by Cambrex (with results available in <20 min using the Cambrex instrument), and the VenorGem Mycoplasma PCR Detection Kit from Sigma-Aldrich.

Good aseptic technique and routine testing are the most effective methods for preventing mycoplasma contamination. Most standard antibiotics are not effective in treating mycoplasma contamination, and many laboratories prefer to dispose of contaminated cultures. In the event that the cultures are valuable and backup cultures are not available, it is imperative to attempt to salvage the contaminated cells. Agents such as Mynox Mycoplasma Elimination Reagent and Mycoplasma Removal Agent (MRA) are commercially available from ICN Flow, and usually require a single treatment to decontaminate cultures (<http://www.unc.edu/depts/tcf/mycoplasma.htm>).

CULTURE MEDIUM PREPARATION

Choice of tissue culture medium comes from experience. An individual laboratory must select the medium that best suits the type of cells being cultured. Chemically defined media are available in liquid or powdered form from a number of suppliers. Sterile, ready-to-use medium has the advantage of being convenient, although it is more costly than other forms. Powdered medium must be reconstituted with tissue-culture-grade water according to manufacturer's directions. Distilled or deionized water is not of sufficiently high quality for medium preparation; double- or triple-distilled water or commercially available tissue culture water should be used. The medium should be filter-sterilized and transferred to sterile bottles. Prepared medium can generally be stored ≤ 1 month in a 4°C refrigerator. Laboratories using large volumes of medium may choose to prepare their own medium from standard recipes. This may be an economical approach, but it is time-consuming and the savings may not offset the time required.

Basic media such as Eagle minimal essential medium (MEM, *UNIT 16.16*), Dulbecco's modified Eagle medium (DMEM; see recipe), Glasgow modified Eagle medium (GMEM, *UNIT 16.23*), and RPMI 1640 and Ham F10 nutrient mixture (e.g., Invitrogen) are composed of amino acids, glucose, salts, vitamins, and other nutrients. A basic medium is supplemented by addition of L-glutamine, antibiotics (typically penicillin and streptomycin sulfate), and usually serum to formulate a "complete medium." Where serum is

added, the amount is indicated as a percentage of fetal bovine serum (FBS) or other serum. Some media are also supplemented with antimycotics, nonessential amino acids, various growth factors, and/or drugs that provide selective growth conditions (see *UNIT 9.5*). Supplements should be added to medium prior to sterilization or filtration, or added aseptically just before use.

The optimum pH for most mammalian cell cultures is 7.2 to 7.4. Adjust pH of the medium as necessary after all supplements are added. Buffers such as bicarbonate and HEPES are routinely used in tissue culture medium to prevent fluctuations in pH that might adversely affect cell growth. HEPES is especially useful in solutions used for procedures that do not take place in a controlled CO₂ environment.

Fetal bovine serum (FBS; sometimes known as fetal calf serum, FCS) is the most frequently used serum supplement. Calf serum, horse serum, and human serum are also used; some cell lines are maintained in serum-free medium (Freshney, 1996). Complete medium is supplemented with 5% to 30% (v/v) serum, depending on the requirements of the particular cell type being cultured. Serum is obtained frozen, then is thawed, divided into smaller portions, and refrozen until needed.

There is considerable lot-to-lot variation in FBS. Most suppliers will provide a sample of a specific lot and reserve a supply of that lot while the serum is tested for its suitability. The suitability of a serum lot depends upon the use. Frequently the ability of serum to promote cell growth equivalent to a laboratory standard is used to evaluate a serum lot. Once an acceptable lot is identified, enough of that lot should be purchased to meet the culture needs of the laboratory for an extended period of time.

Although the use of heat-inactivated serum was once preferred, it is no longer standard, or may be standard in some labs simply because it is an established and routine practice (http://www.unc.edu/depts/tcf/tech_tipsHI.htm). Historically, heat inactivation was considered necessary to destroy complement protein due to its role in cell lysis. It has since been shown that the level of complement components in commercially available FCS was only 1% to 3% of adult levels, and that no significant cell lysis was detected with the use of commercially available FCS, even when undiluted (Triglia and Linscott, 1980). In addition, many researchers prewarm FBS to 37°C, which is enough to inactivate heat-labile complement components. Improvements in the filtration of serum products have also made the use of heat-inactivated serum nonessential for most cell types. In the past serum was filtered through 0.45- μ m or 0.22- μ m filters, raising concern that adventitious agents such as mycoplasma could persist. Today media suppliers use 0.1- μ m or 0.04- μ m porosity membranes to eliminate the need for heat inactivation (http://www.unc.edu/depts/tcf/tech_tipsHI.htm; Hyclone Labs, 1996). Not only is heat inactivation unnecessary for most cell types, it may sometimes be detrimental to cell growth. Heat inactivation not only destroys complement, it also degrades amino acids, vitamins, growth factors, and other nutrients that enhance cell growth. Furthermore, heat inactivation can increase the formation of precipitates that can be mistaken for microbial contamination (Hyclone Labs, 1996). Laboratories that continue the use of heat inactivation should conduct studies to determine whether this step is really indicated. Heat activation may be warranted when conducting immunologic studies or when culturing embryonic stem cells, insect cells, and smooth muscle cells. For cells obtained from a commercial vendor such as ATCC, serum requirements are included in the cell descriptions.

Although FBS has historically been the serum of choice, many investigators are moving towards the use of alternative types of animal serum or to serum-free media. An

increasing number of companies are offering serum-free media for a wide range of cell types. Focus on Alternatives (FOA) lists over 300 types of serum-free products along with their suppliers, applications, and the cell types supported for each product (<http://www.focusonalternatives.org.uk>, compiled in July, 2005). Although the term “serum-free” implies that a medium contains no serum, the medium may not be entirely free of serum-derived products. For example, bovine serum albumin may be used as the protein component of particular types of serum-free media (Newman, 2003). While some media are designed for culturing a particular cell type, others are general-purpose media that can support a variety of cell types.

Commercially prepared media containing L-glutamine are available, but many laboratories choose to obtain medium without L-glutamine, and then add it to a final concentration of 2 mM just before use. L-Glutamine is an unstable amino acid that, upon storage, converts to a form cells cannot use. Breakdown of L-glutamine is temperature- and pH-dependent. At 4°C, 80% of the L-glutamine remains after 3 weeks, but at near incubator temperature (35°C) only half remains after 9 days (Barch et al., 1991). To prevent degradation, 100× L-glutamine should be stored frozen in aliquots until needed.

Another option to prevent degradation is the use of a stabilized dipeptide form of L-glutamine such as GlutaMAX (Invitrogen), which is available both as a stand-alone medium and as a medium supplement. GlutaMAX media contain the dipeptide L-alanyl-L-glutamine, which is split by aminopeptidases in culture to release L-glutamine and L-alanine. This mechanism provides a controlled delivery of L-glutamine to the cells in culture. The GlutaMAX dipeptide is also available as a medium supplement to be used as a direct substitute for L-glutamine in cell culture medium preparations with minimal or no adaptation. Both GlutaMAX Media and Supplement are reported by the manufacturer to maximize cell performance, improve cell viability and growth, extend cell culture life, and minimize toxic ammonium build-up during cell culture (<http://www.invitrogen.com>).

In addition to practicing good aseptic technique, most laboratories add antimicrobial agents to medium to further reduce the risk of contamination. A combination of penicillin and streptomycin is the most commonly used antibiotic additive; kanamycin and gentamycin are used alone. Mycostatin and amphotericin B are the most commonly used fungicides (Rooney and Czepulkowski, 2001). Table A.3F.1 lists the final concentrations for the most commonly used antibiotics and antimycotics. Combining antibiotics in tissue culture media can be tricky, as some antibiotics are not compatible and one may inhibit the action of another. Furthermore, combined antibiotics may be cytotoxic at lower concentrations than the individual antibiotics. In addition, prolonged use of antibiotics may cause cell lines to develop antibiotic resistance. For this reason, some laboratories add antibiotics and/or fungicides to medium when initially establishing a culture but eliminate them from medium used in later subcultures.

Commercially available tissue culture media have been tested for sterility prior to release, and further testing within the research laboratory is generally not required. In clinical laboratories, sterility checks on tissue culture media are often performed as a quality-control monitor in compliance with requirements from the College of American Pathologists (CAP). A small aliquot from each lot of medium is incubated 48 hr at 37°C and monitored for evidence of contamination such as turbidity (infected medium will be cloudy) and color change (if phenol red is the indicator, infected medium will turn yellow). Any contaminated medium should be discarded.

TRYPsinIZING AND SUBCULTURING CELLS FROM A MONOLAYER

A primary culture is grown to confluency in a 60-mm petri plate or 25-cm² tissue culture flask containing 5 ml tissue culture medium. Cells are dispersed by trypsin treatment and then reseeded into secondary cultures. The process of removing cells from the primary culture and transferring them to secondary cultures constitutes a passage, or subculture.

Materials

Primary cultures of cells

HBSS (APPENDIX 2) without Ca²⁺ and Mg²⁺, 37°C

0.25% (w/v) trypsin/0.2% EDTA solution (see recipe), 37°C

Complete medium with serum: e.g., DMEM supplemented with 10% to 15% (v/v) fetal bovine serum (complete DMEM-10; see recipe), 37°C

Sterile Pasteur pipets

37°C warming tray or incubator

Tissue culture plasticware or glassware including pipets and 25-cm² flasks or 60-mm petri plates, sterile

NOTE: All incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified. Some media (e.g., DMEM) may require altered levels of CO₂ to maintain pH 7.4.

1. Remove all medium from primary culture with a sterile Pasteur pipet. Wash adhering cell monolayer once or twice with a small volume of 37°C HBSS without Ca²⁺ and Mg²⁺ to remove any residual FBS that may inhibit the action of trypsin.

Use a buffered salt solution that is Ca²⁺- and Mg²⁺-free to wash cells. Ca²⁺ and Mg²⁺ in the salt solution can cause cells to stick together.

If this is the first medium change, rather than discarding medium that is removed from primary culture, put it into a fresh dish or flask. The medium contains unattached cells that may attach and grow, thereby providing a backup culture.

2. Add enough 37°C trypsin/EDTA solution to culture to cover adhering cell layer.
3. Place plate on a 37°C warming tray 1 to 2 min. Tap bottom of plate on the countertop to dislodge cells. Check culture with an inverted microscope to be sure that cells are rounded up and detached from the surface.

If cells are not sufficiently detached, return plate to warming tray for an additional minute or two.

4. Add 2 ml 37°C complete medium. Draw cell suspension into a Pasteur pipet and rinse cell layer two or three times to dissociate cells and to dislodge any remaining adherent cells. As soon as cells are detached, add serum or medium containing serum to inhibit further trypsin activity that might damage cells.

If cultures are to be split 1/3 or 1/4 rather than 1/2, add sufficient medium such that 1 ml of cell suspension can be transferred into each fresh culture vessel.

5. Add an equal volume of cell suspension to fresh plates or flasks that have been appropriately labeled.

Alternatively, cells can be counted using a hemacytometer (Support Protocol 3) or Coulter counter and diluted to the desired density so a specific number of cells can be added to each culture vessel. A final concentration of $\sim 5 \times 10^4$ cells/ml is appropriate for most subcultures.

For primary cultures and early subcultures, 60-mm petri plates or 25-cm² flasks are generally used; larger petri plates or flasks (e.g., 150-mm plates or 75-cm² flasks) may be used for later subcultures.

Cultures should be labeled with at least two unique identifiers, as well as date of culture and passage number.

6. Add 4 ml fresh medium to each new culture. Incubate in a humidified 37°C, 5% CO₂ incubator.

If using 75-cm² culture flasks, add 9 ml medium per flask.

Some labs now use incubators with 5% CO₂ and 4% O₂. The low oxygen concentration is thought to simulate the in vivo environment of cells and to enhance cell growth.

7. If necessary, feed subconfluent cultures after 3 or 4 days by removing old medium and adding fresh 37°C medium.
8. Passage secondary culture when it becomes confluent by repeating steps 1 to 7, and continue to passage as necessary.

PASSAGING CELLS IN SUSPENSION CULTURE

Passaging of suspension cultures is somewhat less complicated than passaging of monolayer cultures. Because the cells are suspended in medium rather than attached to a surface, it is not necessary to disperse them enzymatically before passaging. However, before passaging, cells must be maintained in culture by feeding every 2 to 3 days until they reach confluency (i.e., until the cells clump together in the suspension and the medium appears turbid when the flask is swirled).

NOTE: All incubations are performed in a humidified 37°C, 5% CO₂ incubator unless otherwise specified. Some media (e.g., DMEM) may require altered levels of CO₂ to maintain pH 7.4.

For materials, see Basic Protocol.

1. Feed cells as follows every 2 to 3 days until the cultures are confluent:
 - a. Remove flask of suspension cells from incubator, taking care not to disturb those that have settled to the flask bottom.
 - b. Aseptically remove and discard about one-third of the medium from the flask and replace with an equal volume of prewarmed (37°C) medium. If the cells are growing rapidly, add an additional 10% medium by volume in order to maintain an optimum concentration of 1×10^6 cells/ml. Gently swirl flask to resuspend cells.
 - c. Return flask to incubator. If there is <15 ml of medium in the flask, incubate flask in horizontal position to enhance cell/medium contact.

At higher volumes of medium, the flask can be incubated in the vertical position.

If using a 25-cm² flask, there should be 20 to 30 ml of medium in the flask at confluency.

2. On the days cultures are not being fed, check by swirling flask to resuspend cells and observing color changes from pink to yellow/orange in the medium, which indicate active cell metabolism.
3. When cultures are confluent ($\sim 2.5 \times 10^6$ cells/ml), passage culture as follows:
 - a. Remove flask from incubator and swirl flask so that cells are evenly distributed in the medium.
 - b. Aseptically remove half of the volume of cell suspension and place into a fresh flask, retaining the other half of the cell suspension in the original flask.

ALTERNATE PROTOCOL 1

Commonly Used Techniques

A.3F.9

Alternatively, the entire cell suspension can be removed from the original flask and divided equally into two fresh flasks. The original flask can be discarded, or if there is concern about the need for additional cells, the original flask can be retained and fed in an attempt to salvage any residual cells.

- c. Feed each flask with 7 to 10 ml prewarmed medium and return flasks to incubator.

Some labs prefer to split the cells 1/3 or 1/4, although increasing the split ratio will result in a longer interval before subcultures reach confluency.

SUPPORT PROTOCOL 1

FREEZING HUMAN CELLS GROWN IN MONOLAYER CULTURES

It is sometimes desirable to store cell lines for future study. To preserve cells, avoid senescence, reduce the risk of contamination, and minimize effects of genetic drift, cell lines may be frozen for long-term storage. Without the use of a cryoprotective agent, freezing would be lethal to the cells in most cases. Generally, a cryoprotective agent such as dimethylsulfoxide (DMSO) is used in conjunction with complete medium for preserving cells at -70°C or lower. DMSO acts to reduce the freezing point and allows a slower cooling rate. Gradual freezing reduces the risk of ice crystal formation and cell damage.

Materials

Log-phase monolayer culture of cells in petri plate

Complete medium

Freezing medium: complete medium (e.g., DMEM or RPMI; see recipes)
supplemented with 10% to 20% (v/v) FBS and 5% to 10% (v/v) DMSO, 4°C

Benchtop clinical centrifuge with 45° fixed-angle or swinging-bucket rotor

NOTE: All incubations are performed in a humidified 37°C , 5% CO_2 incubator unless otherwise specified. Some media (e.g., DMEM) may require altered levels of CO_2 to maintain pH 7.4.

1. Trypsinize cells from plate (see Basic Protocol, steps 1 to 4).

It is best to use cells in log-phase growth for cryopreservation.

2. Transfer cell suspension to a sterile centrifuge tube and add 2 ml complete medium with serum. Centrifuge 5 min at 300 to $350 \times g$ (~ 1500 rpm in Fisher Centrifuge rotor), room temperature.

The benchtop centrifuge can be accessorized depending on the anticipated volume. A variety of brands are available, including Eppendorf, Thermo Electron, and Beckman, and can be obtained from Fisher Scientific, VWR, and other laboratory equipment vendors.

Cells from three or more dishes from the same subculture can be combined in one tube.

3. Remove supernatant and add 1 ml of 4°C freezing medium. Resuspend pellet to obtain a density of 1×10^6 cells/ml.
4. Add 4 ml of 4°C freezing medium, mix cells thoroughly, and place on wet ice.
5. Count cells using a hemacytometer (see Support Protocol 3). Dilute with more freezing medium as necessary to get a final cell concentration of 10^6 or 10^7 cells/ml.

To freeze cells from a nearly confluent 25-cm^2 flask, resuspend in roughly 3 ml freezing medium.

6. Pipet 1-ml aliquots of cell suspension into labeled 2-ml cryovials. Tighten caps on vials.

7. Place vials 1 hr to overnight in a -70°C freezer, then transfer to a liquid nitrogen storage freezer.

Keep accurate records of the identity and location of cells stored in liquid nitrogen freezers. Cells may be stored for many years and proper information is imperative for locating a particular line for future use.

FREEZING CELLS GROWN IN SUSPENSION CULTURE

Freezing cells from suspension culture is similar in principle to freezing cells from monolayer. The major difference is that suspension cultures need not be trypsinized.

NOTE: All incubations are performed in a humidified 37°C , 5% CO_2 incubator unless otherwise specified. Some media (e.g., DMEM) may require altered levels of CO_2 to maintain pH 7.4.

1. Transfer cell suspension to a centrifuge tube and spin 10 min at 300 to $350 \times g$ (-1500 rpm in Fisher Centrifuge), room temperature.
2. Remove supernatant and resuspend pellet in 4°C freezing medium at a density of 10^6 to 10^7 cells/ml.

Some laboratories freeze lymphoblastoid lines at the higher cell density because they plan to recover them in a larger volume of medium and because there may be a greater loss of cell viability upon recovery as compared to other types of cells (e.g., fibroblasts).

3. Transfer 1-ml aliquots of cell suspension into cryovials and freeze as for monolayer cultures.

THAWING AND RECOVERING HUMAN CELLS

When cryopreserved cells are needed for study, they should be thawed rapidly and plated at high density to optimize recovery.

CAUTION: Protective clothing, particularly insulated gloves and goggles, should be worn when removing frozen vials or ampules from the liquid nitrogen freezer. The room containing the liquid nitrogen freezer should be well-ventilated. Care should be taken not to spill liquid nitrogen on the skin.

NOTE: All incubations are performed in a humidified 37°C , 5% CO_2 incubator unless otherwise specified. Some media (e.g., DMEM) may require altered levels of CO_2 to maintain pH 7.4.

Materials

Cryopreserved cells stored in liquid nitrogen freezer

70% (v/v) ethanol

Complete medium (e.g., DMEM or RPMI; see recipes) containing 10% to 20% FBS (see recipe), 37°C

1. Remove vial from liquid nitrogen freezer and immediately place it into a 37°C water bath. Agitate vial continuously until medium is thawed.

The medium usually thaws in <60 sec. Cells should be thawed as quickly as possible to prevent formation of ice crystals that can cause cell lysis. Try to avoid getting water around the cap of the vial.

2. Wipe top of vial with 70% ethanol before opening.

Some labs prefer to submerge the vial in 70% ethanol and air dry before opening.

**ALTERNATE
PROTOCOL 2**

**SUPPORT
PROTOCOL 2**

**Commonly Used
Techniques**

A.3F.11

3. Transfer thawed cell suspension into a sterile centrifuge tube containing 2 ml warm complete medium containing 20% FBS. Centrifuge 10 min at 150 to 200 × g (~1000 rpm in Fisher Centrifric), room temperature. Discard supernatant.

Cells are washed with fresh medium to remove residual DMSO.

4. Gently resuspend cell pellet in small amount (~1 ml) of complete medium/20% FBS and transfer to properly labeled culture plate containing the appropriate amount of medium.

Cultures are reestablished at a higher cell density than that used for original cultures because there is some cell death associated with freezing. Generally, 1 ml cell suspension is reseeded in 5 to 20 ml medium.

5. Check cultures after ~24 hr to ensure that cells have attached to the plate.
6. Change medium after 5 to 7 days or when pH indicator (e.g., phenol red) in medium changes color. Keep cultures in medium with 20% FBS until cell line is reestablished.

If recovery rate is extremely low, only a subpopulation of the original culture may be growing; be extra careful of this when working with cell lines known to be mosaic.

SUPPORT PROTOCOL 3

DETERMINING CELL NUMBER AND VIABILITY WITH A HEMACYTOMETER AND TRYPAN BLUE STAINING

Determining the number of cells in culture is important in standardization of culture conditions and in performing accurate quantitation experiments. A hemacytometer is a thick glass slide with a central area designed as a counting chamber. Cell suspension is applied to a defined area and counted so cell density can be calculated.

The exact design of the hemacytometer may vary; the one described here is the Improved Neubauer from VWR (Fig. A.3F.1). The central portion of the slide is the counting platform, which is bordered by a 1-mm groove. The central platform is divided into two counting chambers by a transverse groove. Each counting chamber consists of a silver footplate on which is etched a 3 × 3-mm grid. This grid is divided into nine secondary squares, each 1 × 1 mm. The four corner squares and the central square are used for determining the cell count. The corner squares are further divided into 16 tertiary squares and the central square into 25 tertiary squares to aid in cell counting.

Accompanying the hemacytometer slide is a thick, even-surfaced coverslip. Ordinary coverslips may have uneven surfaces, which can introduce errors in cell counting; therefore, it is imperative that the coverslip provided with the hemacytometer be used in determining cell number.

Materials

- 70% (v/v) ethanol
- Cell suspension
- 0.4% (w/v) trypan blue or 0.4% (w/v) nigrosin, prepared in HBSS (APPENDIX 2)
- Hemacytometer with coverslip (Improved Neubauer, Baxter Scientific)
- Hand-held counter

Prepare hemacytometer

1. Clean surface of hemacytometer slide and coverslip with 70% alcohol.

Coverslip and slide should be clean, dry, and free from lint, fingerprints, and watermarks.

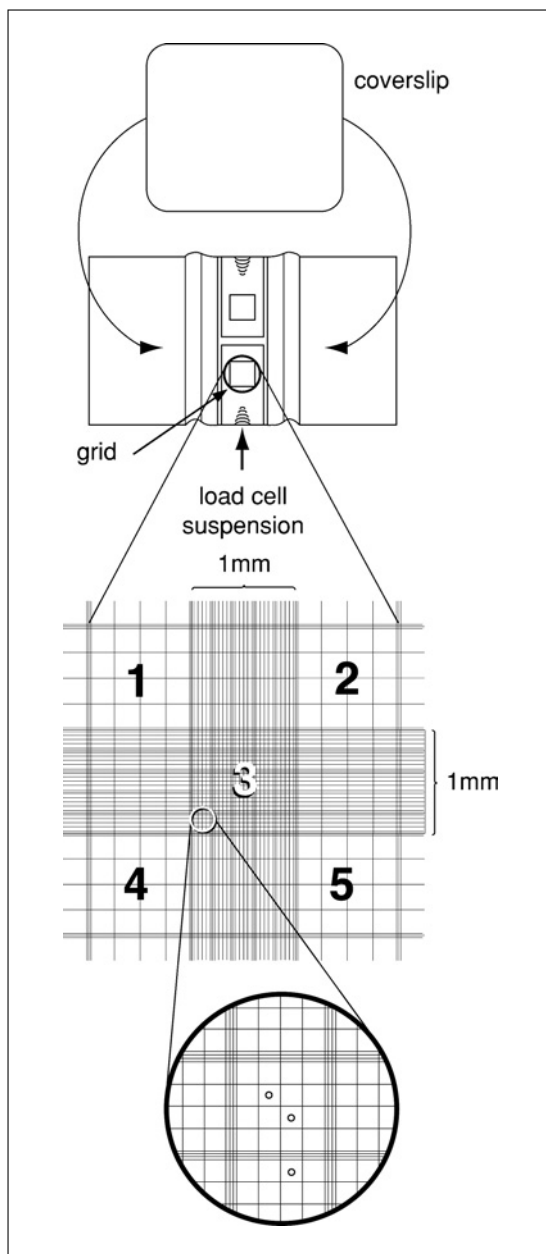


Figure A.3F.1 Hemacytometer slide (Improved Neubauer) and coverslip. Coverslip is applied to slide and cell suspension is added to counting chamber using a Pasteur pipet. Each counting chamber has a 3×3 -mm grid (enlarged). The four corner squares (1, 2, 4, and 5) and the central square (3) are counted on each side of the hemacytometer (numbers added).

2. Wet edge of coverslip slightly with tap water and press over grooves on hemacytometer. The coverslip should rest evenly over the silver counting area.

Prepare cell suspension

3. For cells grown in monolayer cultures, detach cells from surface of dish using trypsin (see Basic Protocol).
4. Dilute cells as needed to obtain a uniform suspension. Disperse any clumps.

When using the hemacytometer, a maximum cell count of 20 to 50 cells per 1-mm square is recommended.

Load hemacytometer

5. Use a sterile Pasteur pipet to transfer cell suspension to edge of hemacytometer counting chamber. Hold tip of pipet under the coverslip and dispense one drop of suspension.

Suspension will be drawn under the coverslip by capillary action.

Hemocytometer should be considered nonsterile. If cell suspension is to be used for cultures, do not reuse the pipet and do not return any excess cell suspension in the pipet to the original suspension.

6. Fill second counting chamber.

Count cells

7. Allow cells to settle for a few minutes before beginning to count. Blot off excess liquid.
8. View slide on microscope with 100× magnification.

A 10× ocular with a 10× objective = 100× magnification.

Position slide to view the large central area of the grid (section 3 in Fig. A.3F.1); this area is bordered by a set of three parallel lines. The central area of the grid should almost fill the microscope field. Subdivisions within the large central area are also bordered by three parallel lines and each subdivision is divided into sixteen smaller squares by single lines. Cells within this area should be evenly distributed without clumping. If cells are not evenly distributed, wash and reload hemacytometer.

9. Use a hand-held counter to count cells in each of the four corner and central squares (Fig. A.3F.1, squares numbered 1 to 5). Repeat counts for other counting chamber.

Five squares (four corner and one center) are counted from each of the two counting chambers for a total of ten squares counted.

Count cells touching the middle line of the triple line on the top and left of the squares. Do not count cells touching the middle line of the triple lines on the bottom or right side of the square.

Calculate cell number

10. Determine cells per ml by the following calculations:

$\text{cells/ml} = \text{average count per square} \times \text{dilution factor} \times 10^4$

$\text{total cells} = \text{cells/ml} \times \text{total original volume of cell suspension from which sample was taken.}$

The volume correction factor for the hemacytometer is 10^4 : each square is 1×1 mm and the depth is 0.1 mm.

Stain cells with trypan blue to determine cell viability

11. Determine number of viable cells by adding 0.5 ml of 0.4% trypan blue, 0.3 ml HBSS, and 0.1 ml cell suspension to a small tube. Mix thoroughly and let stand 5 min before loading hemacytometer.

Either 0.4% trypan blue or 0.4% nigrosin can be used to determine the viable cell number. Nonviable cells will take up the dye, whereas live cells will be impermeable to it.

12. Count total number of cells and total number of viable (unstained) cells. Calculate percent viable cells as follows:

$$\% \text{ viable cells} = \frac{\text{number of unstained cells}}{\text{total number of cells}} \times 100$$

13. Decontaminate coverslip and hemacytometer by rinsing with 70% ethanol and then deionized water. Air dry and store for future use.

PREPARING CELLS FOR TRANSPORT

Both monolayer and suspension cultures can easily be shipped in 25-cm² tissue culture flasks. Cells are grown to near confluency in a monolayer or to desired density in suspension. Medium is removed from monolayer cultures and the flask is filled with fresh medium. Fresh medium is added to suspension cultures to fill the flask. *It is essential that the flasks be completely filled with medium to protect cells from drying if flasks are inverted during transport.* It is also imperative that the flasks have nonvented caps. The cap is tightened and taped securely in place. The flask is sealed in a leak-proof plastic bag or other leak-proof container designed to prevent spillage in the event that the flask should become damaged. The primary container is then placed in a secondary insulated container to protect it from extreme temperatures during transport. A biohazard label is affixed to the outside of the package. Generally, cultures are transported by same-day or overnight courier.

Cells can also be shipped frozen. The vial containing frozen cells is removed from the liquid nitrogen freezer and placed immediately on dry ice in an insulated container to prevent thawing during transport.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. Suitable suppliers for media and components include Invitrogen, Cambrex, Hyclone, and Sigma-Aldrich. For common stock solutions, see APPENDIX 2; for suppliers, see APPENDIX 4.

Complete DMEM

Dulbecco's modified Eagle medium, high-glucose formulation (e.g., Invitrogen), containing:

5%, 10%, or 20% (v/v) FBS (optional; see recipe)

1% (v/v) nonessential amino acids

2 mM L-glutamine (see recipe)

100 U/ml penicillin

100 µg/ml streptomycin sulfate

Filter sterilize and store ≤1 month at 4°C

Throughout this manual, the percentage of serum (usually fetal bovine serum) used in a protocol step is indicated by a numeral hyphenated to the base medium name. Thus, "complete DMEM-10" indicates that 10% FBS is used. Absence of a numeral indicates that no serum is used. See Chapter 9 introduction for a full discussion concerning media preparation and use of serum (heat-inactivation, screening, commercial sources, etc.).

DMEM containing 4500 mg/liter D-glucose can be obtained from Invitrogen. DMEM is also known as Dulbecco's minimum essential medium.

Complete RPMI

RPMI 1640 medium (e.g., Invitrogen) containing:

2%, 5%, 10%, 15%, or 20% FBS (optional; see recipe)

2 mM L-glutamine (see recipe)

100 U/ml penicillin

100 µg/ml streptomycin sulfate

Filter sterilize and store ≤1 month at 4°C

FBS (fetal bovine serum)

Thaw purchased fetal bovine serum (shipped on dry ice and kept frozen at –20°C until needed). Store 3 to 4 weeks at 4°C. If FBS is not to be used within this time, aseptically divide into smaller aliquots and refreeze until used. Store ≤1 year at –20°C.

continued

Repeated thawing and refreezing should be avoided as it may cause denaturation of the serum.

In some cases, heat inactivation may be warranted (see Culture Medium Preparation). To inactivate FBS, heat 30 to 60 min in a 56°C water bath. Alternatively, FBS may be inactivated through radiation treatment.

L-Glutamine, 0.2 M (100×)

Thaw frozen L-glutamine, aliquot aseptically into usable portions, then refreeze. For convenience, L-glutamine can be stored in 1-ml aliquots if 100-ml bottles of medium are used, and in 5-ml aliquots if 500-ml bottles are used. Store ≤ 1 year at -20°C .

Many laboratories supplement medium with 2 mM L-glutamine—1% (v/v) of 100× stock—just prior to use.

Trypsin/EDTA solution

Prepare in sterile HBSS (APPENDIX 2) or 0.9% (w/v) NaCl:

0.25% (w/v) trypsin

0.2% (w/v) EDTA

Store ≤ 1 year (until needed) at -20°C

Most laboratories prefer to purchase trypsin/EDTA as a prepared solution, which is available from vendors including Sigma-Aldrich, Invitrogen, and Cambrex. This is a convenient and cost-effective alternative to preparing the solution within the laboratory. Trypsin/EDTA solution is available in various concentrations including 10×, 1×, and 0.25% (w/v). It is received frozen from the manufacturer and can be thawed and aseptically aliquotted into smaller volumes. Specific applications may require different concentrations of trypsin; the appropriate methods should be consulted for details.

EDTA (disodium ethylenediamine tetraacetic acid) is added as a chelating agent to bind Ca^{2+} and Mg^{2+} ions that can interfere with the action of trypsin.

COMMENTARY

Background Information

At its inception in the early twentieth century, tissue culture was applied to the study of tissue fragments in culture. New growth in culture was limited to cells that migrated out from the initial tissue fragment. Tissue culture techniques evolved rapidly, and since the 1950s culture methods have allowed the growth and study of dispersed cells in culture (Freshney, 1996). Cells dispersed from the original tissue can be grown in monolayers and passaged repeatedly to give rise to a relatively stable cell line.

Four distinct growth stages have been described for primary cells maintained in culture. First, cells adapt to the *in vitro* environment. Second, cells undergo an exponential growth phase lasting through ~ 30 passages. Third, the growth rate of cells slows, leading to a progressively longer generation time. Finally, after 40 or 50 passages, cells begin to senesce and die (Lee, 1991). It may be desirable to study a particular cell line over several months or years, so monolayer cultures can be preserved to re-

tain the integrity of the cell line. Aliquots of early-passage cell suspensions are frozen, then thawed, and cultures reestablished as needed. Freezing monolayer cultures prevents changes due to genetic drift and avoids loss of cultures due to senescence or accidental contamination (Freshney, 1996).

Cell lines are commercially available from a number of sources, including the American Type Culture Collection (ATCC; <http://www.atcc.org>) and the Human Genetic Mutant Cell Repository at the Coriell Cell Repository (CCR; <http://www.coriell.org>). These cell repositories are a valuable resource for researchers who do not have access to suitable patient populations. An up-to-date listing of available cell lines will be provided upon request. The cell lines are preserved in liquid nitrogen, and information on the characteristics are supplied by the distributor. Immortalized cell lines can be maintained in culture indefinitely, while nontransformed cells have a limited life-span *in vitro*. For human fibroblasts, about 20 population doublings occur

between the initial growth in the primary culture and the first subculture (Priest, 1997). About 50 to 60 population doublings is the standard rule for maintaining fibroblast cell lines in culture before risking the loss of cell lines through senescence, cytogenetic or biochemical changes, or contamination. If it is anticipated that these cells will be needed for future study, cells should be frozen at an early culture stage and retrieved for future use (Verma and Babu, 1995).

Critical Parameters

Use of aseptic technique is essential for successful tissue culture. Cell cultures can be contaminated at any time during handling, so precautions must be taken to minimize the chance of contamination. All supplies and reagents that come into contact with cultures must be sterile, and all work surfaces should be kept clean and free from clutter.

Cultures should be 75% to 100% confluent when selected for subculture. Growth in monolayer cultures will be adversely affected if cells are allowed to become overgrown. Passing cells too early will result in a longer lag time before subcultures are established. Following dissociation of the monolayer by trypsinization, serum or medium containing serum should be added to the cell suspension to stop further action by trypsin that might be harmful to cells.

When subculturing cells, add a sufficient number of cells to give a final concentration of $\sim 5 \times 10^4$ cells/ml in each new culture of human cell lines. Optimal concentrations are cell-type specific. Package inserts provided with commercially available cell lines contain information on recommended cell density, culture media, and appropriate culture temperatures. Cells plated at too low a density may be inhibited or delayed in entry into growth stage. Cells plated at too high a density may reach confluence before the next scheduled subculturing; this could lead to cell loss and/or cessation of proliferation. The growth characteristics for different cell lines vary. A lower cell concentration (10^4 cells/ml) may be used to initiate subcultures of rapidly growing cells, and a higher cell concentration (10^5 cells/ml) may be used to initiate subcultures of more slowly growing cells. Adjusting the initial cell concentration permits establishment of a regular, convenient schedule for subculturing—e.g., once or twice a week (Freshney, 1996).

Cells in culture will undergo changes in growth, morphology, and genetic characteristics over time. Such changes can adversely af-

fect reproducibility of laboratory results. Non-transformed cells will undergo senescence and eventual death if passaged indefinitely. The time of senescence will vary with cell line, but generally at between 40 and 50 population doublings fibroblast cell lines begin to senesce. Cryopreservation of cell lines will protect against these adverse changes and will avoid potential contamination.

Cultures selected for cryopreservation should be in log-phase growth and free from contamination. Cells should be frozen at a concentration of 10^6 to 10^7 cells/ml. Cells should be frozen gradually and thawed rapidly to prevent formation of ice crystals that may cause cells to lyse. Cell lines can be thawed and recovered after long-term storage in liquid nitrogen. The top of the freezing vial should be cleaned with 70% alcohol before opening to prevent introduction of contaminants. To aid in recovery of cultures, thawed cells should be reseeded at a higher concentration than that used for initiating primary cultures. Careful records regarding identity and characteristics of frozen cells as well as their location in the freezer should be maintained to allow for easy retrieval.

For accurate cell counting, the hemacytometer slide should be clean, dry, and free from lint, scratches, fingerprints, and watermarks. The coverslip supplied with the hemacytometer should always be used because it has an even surface and is specially designed for use with the counting chamber. Use of an ordinary coverslip may introduce errors in cell counting. If the cell suspension is too dense or the cells are clumped, inaccurate counts will be obtained. If the cell suspension is not evenly distributed over the counting chamber, the hemacytometer should be washed and reloaded.

Anticipated Results

Confluent cell lines can be successfully subcultured in the vast majority of cases. The yield of cells derived from monolayer culture is directly dependent on the available surface area of the culture vessel (Freshney, 1996). Overly confluent cultures or senescent cells may be difficult to trypsinize, but increasing the time of trypsin exposure will help dissociate resistant cells. Cell lines can be propagated to get sufficient cell populations for cytogenetic, biochemical, and molecular analyses.

It is well accepted that anyone can successfully freeze cultured cells; it is thawing and recovering the cultures that presents the problem. Cultures that are healthy and free from

contamination can be frozen and stored indefinitely. Cells stored in liquid nitrogen can be successfully thawed and recovered in over 95% of cases. Several aliquots of each cell line should be stored to increase the chance of recovery. Cells should be frozen gradually, with a temperature drop of $\sim 1^{\circ}\text{C}$ per minute, but thawed rapidly. Gradual freezing and rapid thawing prevents formation of ice crystals that might cause cell lysis.

Accurate cell counts can be obtained using the hemacytometer if cells are evenly dispersed in suspension and free from clumps. Determining the proportion of viable cells in a population will aid in standardization of experimental conditions.

Time Considerations

Establishment and maintenance of mammalian cell cultures require a regular routine for preparation of media and feeding and passaging cells. Cultures should be inspected regularly for signs of contamination and to determine if the culture needs feeding or passaging.

Literature Cited

- Barch, M.J., Lawce, H.J., and Arsham, M.S. 1991. Peripheral Blood Culture. *In* The ACT Cytogenetic Laboratory Manual, 2nd ed. (M.J. Barch, ed.) pp. 17-30. Raven Press, New York.
- Freshney, R.I. 1996. Culture of Animal Cells. A Manual of Basic Techniques, 3rd ed. Wiley-Liss, New York.
- Hyclone Labs. 1996. Heat inactivation: Are you wasting your time? *Art to Science* 15:1-5.
- Knutsen, T. 1991. The ACT Cytogenetic Laboratory Manual, 2nd ed. (M.J. Barch, ed.) pp. 563-587. Raven Press, New York.
- Lee, E.C. 1991. Cytogenetic Analysis of Continuous Cell Lines. *In* The ACT Cytogenetic Laboratory Manual, 2nd ed. (M.J. Barch, ed.) pp. 107-148. Raven Press, New York.
- Newman, C. 2003. Serum-free cell culture—the ethical, scientific and economic choice. *The Biomedical Scientist*, pp. 941-942.
- Priest, J.H. 1997. General cell culture principles and fibroblast culture. *In* The AGT Cytogenetics Laboratory Manual, 3rd ed. (M.J. Barch, T. Knutsen, and J.L. Spurbeck, eds.). Lippincott-Raven, Philadelphia.

Rooney, D.E. and Czepulskowski, B.H. (eds). 2001. Human Cytogenetics: Constitutional Analysis: A Practical Approach, 3rd ed. Oxford University Press, New York.

Triglia, R.P. and Linscott, W.D. 1980. Titers of nine complement components, conglutinin and C3b inactivator in adult and fetal bovine sera. *Mol. Immunol.* 17:741-748.

Verma, R.S. and Babu, A. 1995. Human Chromosomes: Principles and Techniques, 2nd ed. McGraw-Hill, New York.

Key Reference

Lee, 1991. See above.

Contains pertinent information on cell culture requirements, including medium preparation and sterility. Also discusses trypsinization, freezing and thawing, and cell counting.

Internet Resources

<http://www.unc.edu/depts/tcf/info.html>

This site contains troubleshooting information for treating tissue culture contamination, including discussion of contamination risks and recommendations.

<http://www.cdc.gov/od/ohs.pdffiles/BCS-3.pdf>

Primary Containment for Biohazards: Selection, Installation, and Use of Biological Safety Cabinets, 2nd ed. Sept 2001. U.S. Dept of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention and National Institutes of Health, U.S. Government Printing Office, Washington.

http://www.yale.edu/oehs/PDG_files/cad_07_01.pdf

Clean Air Device (Primary Containment Device) Program Guide, July 2001, Yale University, Office of Environmental Health and Safety, 135 College Street, New Haven, Conn.

<http://www.ehs.cornell.edu/bio/cabinets.htm>

This site contains a comprehensive discussion of the use of biological safety cabinets, including operational procedures and the advantages and disadvantages of ultraviolet lights.

Contributed by Mary C. Phelan
Molecular Pathology Laboratory Network
Maryville, Tennessee

Scientific cooperation between laboratories in the United States is no longer confined to those at the same institution, nor even within the boundaries of the country. More than ever, scientists are forming collaborations with their peers outside the U.S., leading not only to the exchange of knowledge and ideas, but also to the exchange of biological reagents. Thus, investigators need be aware that an import permit may be required to obtain purified proteins, chemically synthesized materials, cell lines, tissues, and other biological materials from abroad. Before requesting any biological material from a laboratory outside the U.S., it is the responsibility of the “receiver” to ascertain whether an import permit is required, to apply for the permit if one is determined to be necessary, and to provide the “shipper” with an official copy of the permit to be included with the shipped material.

There are two government agencies in place to issue permits for the importation of most biologicals. All materials derived from animals are subject to regulations set forth by the United States Department of Agriculture (USDA) and must be cleared by USDA inspectors at the port of arrival before entry into the United States is authorized. In general, an import permit issued by the USDA is necessary for receipt of any animal-derived material or any biological material (including proteins) that has been in contact with materials of animal origin. The USDA does not have regulatory authority over the importation of human or nonhuman primate material unless it is produced via tissue culture, due to the use of bovine material (fetal bovine serum or calf serum) in tissue culture. The U.S. Public Health Service (UPHS) has jurisdiction over human and nonhuman primate materials. According to UPHS regulations, no etiologic agent, animal host, or vector of human disease may be imported into the United States or distributed after importation without being accompanied by a permit issued by the Director of the Centers for Disease Control. Failure to obtain the appropriate import permit will result in the shipment being held at the port of entry until a permit is applied for and issued (a 4- to 6-week process) or may result in shipment confiscation and destruction by the quarantine officer present.

IMPORTING RESTRICTED BIOLOGICAL MATERIALS

The following is a list of restricted biological materials requiring an import permit: proteins, monoclonal and polyclonal antibodies, antisera, immunoglobins, recombinant products, enzymes, hormones, immunoassay components or kits, animal tissues, blood, cells, cell lines, RNA/DNA extracts, plasmids, vectors, microorganisms—e.g., fungi, bacteria, viruses—and other products that are derived from animals or have been in contact with animal-derived materials. In addition, materials received from countries shown not to be free of foot-and-mouth disease (see Title 9, code of Federal Regulations, Part 94.1) require an import permit. Note that the USDA will not permit the importation of cell cultures, monoclonal antibodies, ascites fluid, or bovine serum from countries where rinderpest and foot-and-mouth disease are present unless the imported materials are determined to be virus-free. Cell cultures being imported into the U.S. must have been grown only with fetal bovine serum of U.S., Canadian, New Zealand, or Australian origin. Material exposed to bovine products originating in the United Kingdom, Switzerland, Ireland, France, or Portugal may NOT be used, either directly or indirectly, in any animals due to the presence of bovine spongiform encephalopathy (BSE) in those countries, and the material is restricted to *in vitro* uses only.

CRITERIA FOR HANDLING IMPORTED RESTRICTED MATERIALS

An import permit is valid only for transfer of the material(s) listed on the permit (a permit can cover more than one item), and for shipment between the two laboratories listed on the permit. Therefore, for each laboratory abroad from which biological materials are to be imported, a separate import permit is required. Once biological materials have been received from abroad, the import permit still imposes regulations on how the materials are handled by the receiver. Import permits do not usually authorize direct or indirect exposure of domestic animals—work is limited to *in vitro* uses only. In addition, permits are only valid for work conducted or directed by the permit holder at his or her facilities. Materials cannot

Contributed by Paula Wolf

Current Protocols in Molecular Biology (1997) A.3G.1-A.3G.3
Copyright © 1997 by John Wiley & Sons, Inc.

Commonly Used
Techniques

A.3G.1

Supplement 40

be removed to another location, nor distributed to others, without USDA authorization. There is one exception: monoclonal antibodies produced by hybridoma cell lines may be commercially distributed for in vitro uses.

IMPORTING NONRESTRICTED BIOLOGICAL MATERIALS

Any of the restricted materials that do not contain animal- or cell culture-derived products or additives such as albumin, serum, or gelatin, and were not exposed to any infectious agents of agricultural concern, do not need an official import permit. However, an accompanying declaration is required with each shipment to indicate that the material does not contain any animal- or cell culture-derived products or additives. This information should be supplied with each shipment in a clear and concise manner, and be available for the USDA Port of Entry Inspector's review. A separate memo or letter, such as U.S. Customs declaration and invoice, should be included with the shipping documents. Note that whether a particular biological material requires an import permit cannot be decided by the investigator; the investigator must contact the USDA for a decision well in advance of making arrangements for importing the material.

PACKAGING BIOLOGICAL MATERIALS FOR IMPORTATION

Imported biological materials are subject to packaging and shipping requirements of various federal and international regulations. Proper packaging is the primary consideration and of utmost importance in the safe transportation of hazardous materials. If the biological material being shipped is infectious or radioactive, this information has to be noted on the outside of the package. In addition, if dry ice is used for shipment, that must be clearly stated on the package. It is of course advantageous to the investigator that the shipped material be packaged in the most efficient way possible, as well as in a form which will ensure its stability. If proteins are to be shipped, it may be better to send the samples lyophilized or blotted on a membrane, rather than in liquid form. This will omit the need for dry ice or ice packs in the packaging, which will save on shipping costs as well as ensure the samples will be stable if receipt of the package meets with unforeseen delays. If sending materials that require dry ice or ice packs, such as antibodies or antisera, there must be enough dry ice to allow the package to maintain the appropriate tempera-

ture during shipping for approximately 5 days. This requires utilizing much larger packaging boxes than are normally used to ship the same material within the United States. When shipping cell lines, sometimes it is better to send cells in culture, rather than frozen aliquots. To send cells in culture, the tissue culture flask containing the live cells must be filled to the very top with the appropriate medium, tightly capped, and the seal secured with Parafilm. If the tissue culture flask is not completely filled with medium, the resulting air pockets will cause the flasks to explode in flight. Thus, the packaging of any biological material for shipping is an import consideration to ensure success of the scientific collaboration!

OBTAINING AN IMPORT PERMIT APPLICATION

The Division of Veterinary Services of the USDA Animal and Plant Health Inspection Service (APHIS) administers regulatory programs to control the import/export of biological materials. Import permits can be obtained from either local USDA port offices or from APHIS directly: USDA, APHIS, VETERINARY SERVICES, National Center for Import-Export Products Program, 4700 River Road Unit 40, Riverdale, MD 20737-1231, Tel: (301) 734-7830, Fax: (301) 734-8226.

When submitting an import application, (1) obtain the import permit well in advance of the proposed shipping date; (2) inform the supplier of the USDA's requirements, and do not allow shipment of materials until an import permit has been issued and the supplier has received a copy of the permit; (3) list all the potential U.S. ports of entry; (4) specify whether the material is for in vitro or in vivo use and whether it is for commercial distribution or for research in one's own laboratory; (5) list any treatments the material has undergone, such as processing and purification steps involving pH, heat, chromatography, or other methods; and (6) indicate the source(s) of nutrient factors in culture media, cell line designation and history, whether enzymes were used, and what types of viruses (if any) are being studied in the laboratory of origin (all must be listed).

RESOURCES

APHIS Web Site

More information regarding APHIS programs and import/export can be obtained on their internet home page at <http://www.aphis.usda.gov>.

**Fax Service of Veterinary Services,
National Center for Import-Export**

By calling (301) 734-4952, the investigator can access an automated document retrieval system and arrange to have an import application faxed directly, and choose from a menu of documents that describe general instructions for completing the application, as well as information regarding the importation of biological materials from 18 distinct categories.

Contributed by Paula Wolf
Massachusetts Institute of Technology
Cambridge, Massachusetts

The design and use of a particular enzyme assay is dependent on the objectives of the particular investigation. If the focus is on enzyme purification or issues of clinical biochemistry, experimental protocols typically require optimizing conditions for all substrates and effectors—e.g., with respect to pH value, temperature, buffering agent, and ionic strength. However, optimal conditions may not be reasonable for all parameters and factors. For example, substrate instability may preclude studies at the enzyme's optimal pH or at the physiological temperature. Thus, Silverstein and Sulebele (1969) examined the exchange rates of malate dehydrogenase at 1°C to avoid problems arising from the decomposition of oxaloacetate. Likewise, in clinical studies and enzyme purification, saturating concentrations of substrate(s) are typically required, although this may not be possible or even desirable in many instances, as limitations in substrate solubility or substrate inhibition of the enzyme may discourage the use of elevated concentrations. The focus in clinical studies and enzyme purification is on providing good estimates of the activities of a specific enzyme in a number of physiological samples with a rapid, straightforward, and reproducible assay method. However, for the kineticist the objective is to characterize the initial rates of the enzyme-catalyzed reaction under a wide set of conditions. It is through these detailed studies, often carried out at subsaturating substrate concentrations, that the investigator is able to determine the true optimal parameters to be used in protein purification and clinical studies, as well as to probe the dynamics and mechanism of the reaction and its modulation by other factors. The purpose of this appendix is to provide a brief review of issues important in the design of initial-rate assay methods.

In most instances, the investigator already has some information about the kinetic properties of the enzyme, obtained from preliminary experiments as well as from the literature. It is necessary to have an approximation of the magnitude of the Michaelis constants of all the substrates in order that detailed initial rate studies can be effectively designed. Certainly, information on the identity of the substrates and products is required, along with an effective procedure for monitoring substrate depletion or product accumulation. The nature of the reac-

tion and the properties of the substrates help dictate the assay procedure used. Enzyme-assay methods include those based on spectrophotometry, radiometry, pH-stat, turbidometry, polarography, and fluorometry. Spectrophotometric and fluorometric assay protocols rely on the differences in spectral or fluorescent properties, respectively, of the product(s) relative to the substrate(s). With radiometric procedures, effective means for separating labeled substrate(s) from labeled product(s) must be available so that the amount of product formed at a given time can be accurately measured. Those reactions that utilize or generate protons (or hydrons) can be followed by pH-stat methods. If the reaction studied results in a change in solution turbidity (e.g., caused by aggregation of a macromolecular complex), then turbidometric protocols can effectively assay for the protein that causes the change in turbidity. All of these methods, as well as others, can be found in Bergmeyer (1983).

GENERAL ASPECTS OF ASSAY DESIGN

An enzyme's progress curve (i.e., a plot of product formation or substrate depletion as a function of time; see Fig. A.3H.1) consists of four phases that are of interest to the investigator. The duration of the **pre-steady-state** region is relatively short and is typically followed via rapid-flow or relaxation techniques. Discussion of this region of the progress curve is beyond the scope of this appendix, but a review of these techniques has recently appeared (Fierke and Hammes, 1995). The next portion of the curve is the **steady-state** phase, a longer phase characterized by relatively constant concentrations of each of the individual enzyme forms (e.g., the enzyme-substrate binary complexes). During the third phase of the progress curve—the **post-steady-state** region—the levels of these enzyme complexes change rapidly, until the last phase—**equilibrium**—has been reached. Each of these phases contains information of use to the investigator. The steady-state portion of the curve typically exists for a few to several hundred seconds. It is this region of the curve that is most accessible to the majority of investigators. In addition, the relatively straightforward mathematical treatment of steady-state kinetics provides an excellent foundation for the subsequent characterization

Commonly Used Techniques

A.3H.1

Contributed by R. Donald Allison

Current Protocols in Molecular Biology (1997) A.3H.1-A.3H.11

Copyright © 1997 by John Wiley & Sons, Inc.

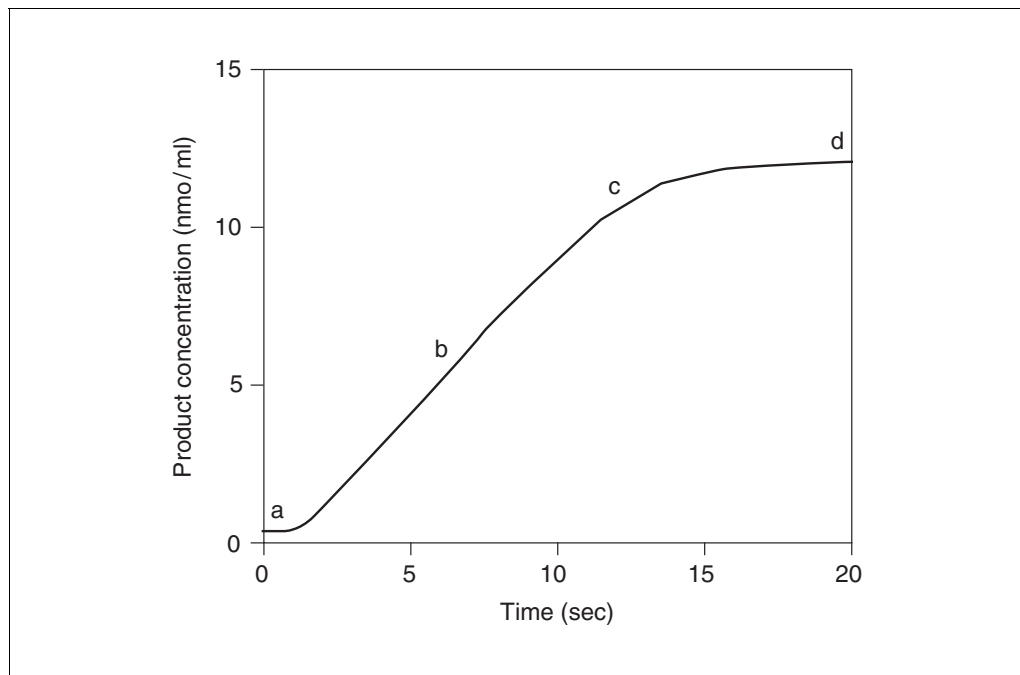


Figure A.3H.1 Typical progress curve for an enzyme-catalyzed reaction, where *a* represents the pre-steady-state region, *b* the steady-state region, *c* the post-steady-state region, and *d* the equilibrium region.

of the enzyme system. Because enzyme rate expressions can be quite complex, a number of assumptions are made to linearize the equations. The investigator should ensure that each of these factors are addressed when any assay protocol is being designed.

Maintaining Constant Reaction Conditions

The status of each experimental parameter—e.g., temperature, ionic strength, and pH—should remain constant over the time course of the assay.

Temperature

Prior to initiation of the reaction by addition of a small aliquot of enzyme or substrate, the reaction mixture should be thermally equilibrated with an accurately controlled constant-temperature water bath for several minutes. If the substrate is thermally stable, this incubation should be for ≥ 10 min or until the desired reaction temperature is reached. If the substrate is unstable at this temperature, it may be necessary either to assay at a different temperature, to correct for substrate decay with time, or to initiate the reaction by addition of a small aliquot of substrate solution that was freshly prepared and maintained in an ice bath. When the reaction is initiated by addition of either the enzyme or the substrate, the volume of the

aliquot should be small enough to minimize any alterations in temperature upon addition. Typically, there is at least a hundred-fold difference in volume between the reaction mixture and the initiating aliquot. If the enzyme and substrate are stable under the assay conditions, then both solutions should be preincubated at this temperature.

It should be emphasized that valid initial rate experiments are only obtained when good temperature control is maintained. The temperature-regulated water bath should have a variability of $\leq 0.1^\circ\text{C}$. When the reaction is initiated, the solution should be thoroughly mixed. If reaction progress is followed using an instrument such as a spectrophotometer, the sample compartment should be thermally isolated.

Another issue of concern is choice of the temperature at which enzyme activity will be measured. In 1964, the International Union of Biochemistry (IUB) recommended a temperature of 30°C . Clinical chemists have often suggested a change to 37°C . However, the final choice of temperature will be dependent on the system being investigated. For example, the activity of glutamate dehydrogenase isolated from the hyperthermophile *Pyrococcus furiosus* is negligible at both 30°C and 37°C ; the enzyme exhibits maximal activity at 97°C (Klump et al., 1992). Clearly, at some point in the investigation, the effect of temperature on

the catalytic activity will have to be characterized.

Buffer selection

In setting up the assay protocol, care should also be exercised in selecting an appropriate buffer agent. Unfortunately, many investigators choose a buffer solely on the basis of the desired pH of the reaction medium. Many excellent buffers, however, have a significant oxidation-reduction potential that can severely alter an enzyme's activity. For example, cacodylate ($pK_a = 6.27$) and other organoarsenicals have a potent oxidizing potential under acidic conditions. Borate ($pK_a = 9.23$) and other borate-based buffers form complexes with many diols, polyols, carbohydrates, and ribonucleotides; for example, in the presence of borate ion, both L- and D-serine form a transition complex and strongly inhibit γ -glutamyl transpeptidase (Tate and Meister, 1978). Hence, care should be exercised in selecting a buffer.

In determining which buffer is optimal for a given enzyme system, it is best to examine the enzyme's activity with as many buffers as possible. Buffers in whose presence the enzyme exhibits high activity should then be examined to determine if the buffering agent interacts with the protein. If pH and ionic strength are held constant, the enzyme activity should be independent of the buffer concentration. Another factor of importance in selecting the appropriate buffer is the capacity of the buffering agent to bind metal ions. Ideally, a good buffer will not form any significant complexes with metal-ion cofactors.

Metal-ion concentrations

The role of metal ions is crucial with many enzymes, particularly with systems that require nucleotides—e.g., phosphotransferases. Care must be exercised in designing initial-rate assays to ensure that the concentration of free metal ions does not vary significantly with substrate concentration. Thus, when varying the concentrations of substrates such as ATP, the concentration of free, uncomplexed Mg^{2+} should remain relatively constant. One of two methods is commonly utilized to achieve this end: either (1) maintaining the total nucleotide concentration and the metal concentration at a constant stoichiometric ratio (e.g., 10:1) or (2) maintaining a constant excess of metal ion over the total nucleotide concentration (e.g., 1.0 mM free Mg^{2+}). The second method, which has been demonstrated to be the preferable of the two (O'Sullivan and Smithers, 1979), requires

knowledge of the stability constant for the metal ion–nucleotide complex under the conditions of the assay.

Substrate Stability and Purity

It is crucial to establish the degree of stability of the substrate as well as that of any enzyme effector under the assay conditions (also see discussion of Maintaining Constant Reaction Conditions). Substrate and effector stabilities can be tested by preparing a solution of the compound at a concentration that will be used experimentally with the other components of the assay mixture, but without the enzyme. Then, samples are removed at different periods of time, usually over a period of several hours, and assayed either chemically or enzymatically (Bergmeyer, 1983). A plot of concentration as a function of time will reveal if the substrate is stable or if there is a decay. When present, this decomposition is usually a first-order process and can be described by the equation $[S]_t = [S]_0 e^{-kt}$, where $[S]_t$ and $[S]_0$ are the substrate concentrations at time t and time 0, respectively, and k is the first-order rate constant. If the decomposition is greater than first-order or composed of more than one first-order process, then the decay must be evaluated by the appropriate expression, which can be found in any basic kinetics text.

An issue related to substrate stability is substrate purity. Impurities in the substrate, effectors, or enzyme preparation can be potential sources of experimental error. Because alternative substrates and competitive inhibitors are often structurally similar to the substrate and frequently have similar chemical and physical properties, it is not surprising to find these compounds as impurities in commercial substrate preparations. Spectral and chromatographic analyses, as well as enzymatic analysis, are crucial in establishing the purity of a particular biochemical substance. If impurities are present in a reaction mixture, a kinetic analysis will give wrong estimates of the kinetic parameters (Allison and Purich, 1979). Heavy-metal ions are often found as impurities in commercial substrates. To remove these, it may be necessary to recrystallize the substrates, possibly in the presence of a chelating agent such as EDTA.

These issues of purity and stability are also pertinent for any modulators of enzyme activity, including the product(s). The products have to be stable over the time course of the assay. In some instances, the precise nature of the true product of the enzyme-catalyzed reaction may

not be known. In these cases, the investigator may follow the reaction via substrate depletion. It is frequently useful to do a kinetic analysis of freshly prepared substrate in the absence and presence of enzymatically prepared product. A decent understanding of the possible presence of product inhibition in initial-rate experiments is always useful to the protein chemist.

Enzyme Stability and Purity

Enzymes often undergo a slow inactivation. Usually, the rate of this inactivation is accelerated by dilution of the stock enzyme preparation. Efforts have to be made to eliminate or lessen the loss of enzyme activity with time. Often, multisubstrate enzymes are stabilized by the presence of a substrate. For example, tubulin:tyrosine ligase is stabilized by the presence of its protein substrate (Deans et al., 1992) and the presence of Mg^{2+} -ATP is known to stabilize a number of phosphotransferases. If the enzyme inactivation can be retarded by the presence of a substrate, the investigator should note the amount of that substrate that is present with the enzyme so that proper corrections can be made when the reaction is initiated with the full assay mixture.

Other agents—e.g., glycerol, certain salts, high-ionic-strength solutions, and EDTA—have also been used to stabilize proteins. Often, inactivation is a result of surface denaturation or adsorption of the enzyme on glass surfaces following dilution. The presence of stabilizing agents and/or use of plastic containers can lessen these effects significantly. The presence of serum albumin or another protein will frequently minimize the effects of enzyme dilution; the investigator should test several proteins when this technique is used. The stabilizing protein may also bind the substrates or the enzyme's cofactors; if this is the case, corrections have to be made.

If a stabilizing agent cannot be identified, the inactivation should be minimized by using the smallest dilution possible. If practical, the reaction may be initiated by the addition of one of the substrates. Corrections for loss of enzyme activity can also be made by establishing an activity-decay curve. However, it should be verified that the loss of activity is due to total inactivation and not simply to a conformational change to a form having less activity (i.e., a form with different kinetic parameters). This possibility could be tested by kinetic assays at significantly different times following dilution or by two reference assays, one using a saturat-

ing substrate level and the other using a saturating substrate level.

The presence of thiol reagents—e.g., 2-mercaptoethanol, dithiothreitol, or dithioerythritol—is routinely helpful in maintaining activity for those enzymes that require reduced thiol groups for activity. Care must be exercised that the reducing agent does not produce other effects. For example, γ -glutamylcysteine synthetase is inhibited by the presence of thiol reagents even though cysteine is a substrate (Seelig and Meister, 1985).

Enzyme purity is also an issue of concern. It is important in early initial-rate studies to eliminate contaminating activities and to demonstrate that any minor activity does not substantially affect the reaction being studied. For multisubstrate enzymes, this can be readily accomplished by observing the stability of each substrate, effector, and product separately in the presence of the enzyme preparation. Adenylate kinase is a common contaminant in preparations of a wide variety of phosphotransferases and other nucleotide-dependent enzymes. This contaminating enzyme can be inhibited by performing the assay at high ionic strength (Bowen and Kerwin, 1955). Addition of elevated levels of AMP will also inhibit adenylate kinase. However, many phosphotransferases are also inhibited by this mononucleotide or are themselves inhibited by high ionic strength. Purich and Fromm (1972) and Lienhard and Secemski (1973) have shown that the multisubstrate analogs P^1, P^4 -di(adenosine-5')-tetrphosphate and P^1, P^5 -di(adenosine-5')-pentaphosphate are potent and specific inhibitors of this activity.

In some instances it may be impossible to completely eliminate side reactions. Enzymes that have branched reaction pathways (i.e., multifunctional enzymes) will exhibit two pathways. Most multifunctional enzymes are transferases having multiple acceptor specificity. Glucose-6-phosphatase catalyzes not only the hydrolysis of glucose-6-phosphate but also its synthesis; thus, it may effect the transfer of phosphoryl groups from pyrophosphate, carbamoyl phosphate, or other glucose-6-phosphate molecules (Nordlie, 1982). This synthetic role of glucose-6-phosphatase is not a minor activity; at a concentration of 100 mM D-glucose, the synthesis and hydrolysis activities are approximately equal. The enzyme γ -glutamyl transpeptidase catalyzes the hydrolysis of a number of γ -glutamyl-containing molecules (including glutathione and glutathione disulfide) and also exhibits a transferase activ-

ity in which the glutamyl moiety is transferred to any of a number of amino acid acceptors, particularly cysteine, glutamine, and methionine (Allison, 1985). The hydrolase and transferase activities of γ -glutamyl transpeptidase are approximately equal in the presence of a pool of acceptor substrates at their physiological concentrations (Allison and Meister, 1981). Asparagine synthetase catalyzes the glutamine-dependent synthesis of L-asparagine and also exhibits a significant glutaminase activity (Boehlein et al., 1994). In all of these cases, it is possible to use inactivation studies to observe the parallel loss of both (or all) activities with time.

Enzyme Concentration

The duration of the steady-state phase of an enzyme-catalyzed reaction is dependent on a number of factors. Perhaps the most important of these is the relative concentration of the enzyme and substrate. The ratio of the initial substrate concentration to the total enzyme concentration should be $\geq 100:1$. With lower values, the steady-state phase will be of very short duration. It should be pointed out that this relationship holds for enzyme effectors as well. Thus, when investigating the influences of an activator or competitive inhibitor on enzyme activity, the total enzyme concentration should be much less than the concentration of the effector. This may prove difficult if the effector has a very tight affinity with the enzyme. In these cases, the protocols outlined by Williams and Morrison (1979) will prove useful.

Initial Rate and Steady-State Condition

In an experiment, the investigator typically measures the rate of an enzyme-catalyzed reaction by calculating the slope of the progress curve. This procedure is only valid if the plot of product concentration as a function of time has been demonstrated to be linear within the time frame of the assay. However, this is not the only criterion that determines initial rate and steady-state kinetics. Linearity has been observed in many non-steady-state systems (Allison and Purich, 1979).

The general rule of thumb is that initial rate conditions prevail when the substrate concentration is still within 10% of its starting value. This value is clearly acceptable for reactions that are thermodynamically quite favorable and that do not exhibit any significant product inhibition. However, if the equilibrium constant for the reaction is not large, an estimate of the

degree of substrate-to-product conversion should be made. With smaller equilibrium constants, the initial-rate assay protocol necessarily must be more sensitive. In addition, if the product has very tight affinity for the enzyme, product accumulation can lead to significant errors in initial-rate determinations even in those cases where the degree of conversion is low. A preliminary test to assess this possibility is to look at initial rates in the absence and presence of a known amount of product. Where product inhibition is a problem, product can be depleted if an auxiliary enzyme system for doing this is present. Nevertheless, product inhibition can be an important regulatory scheme for an enzyme. Any kinetic characterization of an enzyme should include accurate measurements of the dissociation constants for all products. Product-inhibition studies can also assist the investigator in establishing mechanistic details of the protein undergoing evaluation (Rudolph, 1979).

In addition to establishing that the progress curve is truly linear, demonstrating that the degree of substrate depletion in the initial time course of the assay is low, and assessing the degree of product inhibition that may be present, the investigator should also demonstrate in the early stages of the kinetic studies that the rates observed are linear with respect to the total enzyme concentration ($[E]_{\text{total}}$). If true steady-state conditions are present, a plot of initial rates versus $[E]_{\text{total}}$ (in which the enzyme concentration is varied above and below the final value to be used in the initial-rate assays) should be linear.

Substrate Concentration Ranges

A wide range of substrate concentrations should be tested in preliminary studies. Such trials will provide an estimate of the Michaelis constant (K_m) and an assessment of whether substrate inhibition occurs at elevated substrate levels. These tests will also detect the possible existence of cooperativity. Once an estimate of the Michaelis constant has been made, a more detailed kinetic study can be made, in which the substrate concentration is typically varied from ~ 0.2 to 5.0 times the value of the apparent constant. It should be emphasized that reasonable estimates of the K_m value can only be made when initial rates have been obtained from substrate concentrations both below and above the true K_m value. For a one-substrate enzyme-catalyzed reaction:

$$K_m = [A](V_{\text{max}} - v)/v$$

Commonly Used
Techniques

A.3H.5

where $[A]$ is the substrate concentration, V_{\max} is the maximal velocity of the reaction, and v is the velocity of the reaction at the substrate concentration. A series of substrate concentrations with their respective initial rate velocities will provide an estimate of the Michaelis constant. In a standard double-reciprocal plot, the horizontal intercept is numerically equal to $-1/K_m$. For multisubstrate enzymes (see below), estimates of the Michaelis constants will be obtained from slope and intercept replots of the initial rate data. In addition, appropriate statistical treatment of the raw data should be provided (Wilkinson, 1961; Cleland, 1979a) for both K_m and V_{\max} values (or k_{cat} values, equal to $V_{\max}/[E]_{\text{total}}$, if those are reported). Note that, regardless of whether V_{\max} or k_{cat} values are reported, the total protein concentration, as well as the specific activity of the enzyme used, should be reported whenever possible.

With enzyme systems utilizing two substrates, the concentrations of both substrates must be varied as described above. Typically, three solutions are prepared. One solution contains one of the substrates at its highest level. A second solution contains the second substrate at its highest level. The third solution contains the buffer, all cofactors, metal ions, auxiliary

enzymes, and any other effectors, but no substrates. This third solution is used to prepare serial dilutions of the other two stock solutions. Thus, a matrix of initial rate measurements on both substrates can be evaluated from a single experimental trial. Figure A.3H.2 illustrates a typical double-reciprocal plot for a two-substrate enzyme-catalyzed reaction (where the two substrates are A and B). In this example, Michaelis constants for the two substrates can be obtained by constructing slope and/or intercept replots of the data. For example, plotting the slopes of the five lines as a function of the reciprocal of the concentration of B will yield a straight line with a horizontal intercept numerically equal to $-1/K_B$, where K_B is the Michaelis constant for B.

For enzymes that utilize three or more substrates, the protocol described above can become unwieldy. Two procedures have been described to address the design of kinetic studies for these proteins. In the first method (Frieden, 1959), one of the three substrates is held constant at a concentration above its respective Michaelis constant that is still subsaturating. The other two are then varied as described above. This same procedure is then followed with each of the remaining substrates. Thus, multisubstrate enzymes are reduced to pseudo-

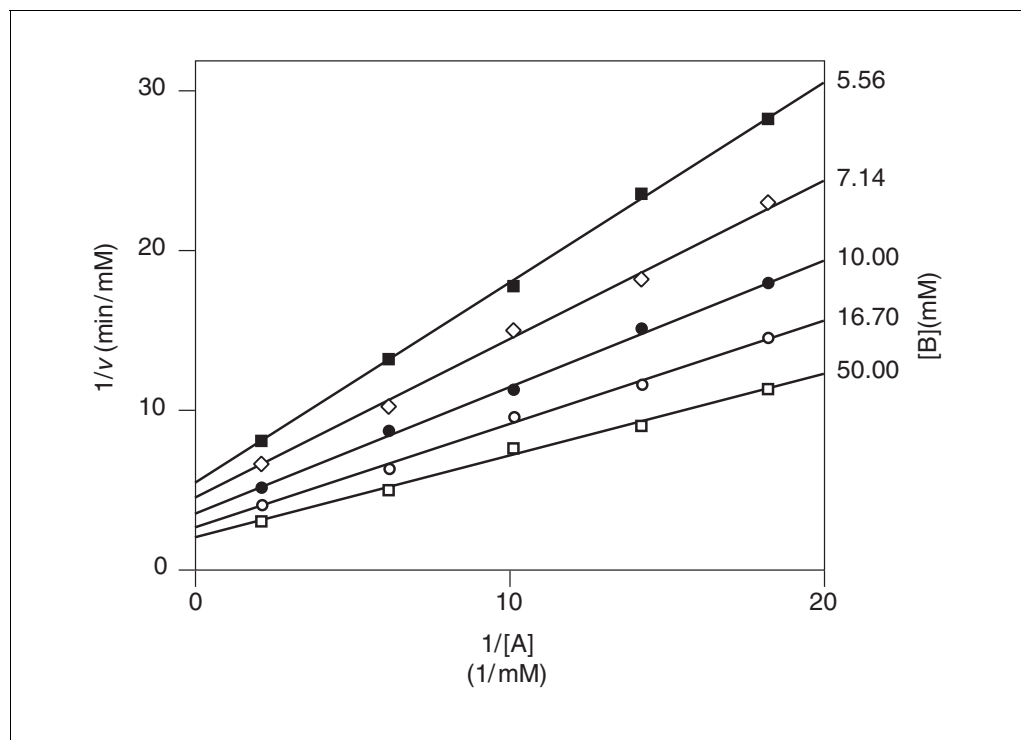


Figure A.3H.2 Double-reciprocal plot of a two-substrate enzyme-catalyzed reaction, where v = the velocity of the reaction and $[A]$ and $[B]$ represent the concentrations of substrates A and B respectively.

two-substrate systems. One distinct advantage of this procedure is that standard graphical presentations of the data thus produced (e.g., via double-reciprocal plots) are identical to that shown in Figure A.3H.2. However, the concentration value for the third substrate have to be provided. In the second procedure, the concentration of one substrate is varied while the other substrates are held at a constant ratio, although the absolute amounts are varied. This is then repeated for all remaining substrates until the concentration of each of the substrates has been varied (Fromm, 1967). One limitation with this approach is the difficulty in obtaining true Michaelis constants.

The presence of substrate inhibition can be an annoyance to an investigator who is attempting to determine the kinetic parameters of an enzyme as well as identify the binding pattern. However, the pattern of such inhibition can serve as a tool in understanding the kinetic and possibly the regulatory scheme for the enzyme. Substrate inhibition can occur as a result of a number of different causes—i.e., the formation of a dead-end complex with a particular enzyme form, the appearance of a different binding mechanism at elevated levels of substrate, the presence of an allosteric site for the substrate, or nonspecific inhibition that is due to an elevated ionic strength or another factor. When substrate inhibition occurs, in the case of multisubstrate enzymes, estimates of the K_i value (i.e., the dissociation constant for inhibition) for the substrate can be obtained from slope and intercept replots of the kinetic data. For enzymes having a single substrate, the data can be fit to an enzyme-rate expression containing a term for substrate inhibition ($v = V_{\max}[S]/\{K_m + [S] + [S]^2/K_i\}$). In addition, estimates of K_m , K_i , and V_{\max} can be obtained from double-reciprocal plots (Cleland, 1979b). Note that in such cases, double-reciprocal plots will be nonlinear, with the degree of curvature dependent on the magnitude of the K_i value. In these cases, it may be necessary to use a computer program to fit the data. Cleland (1979a) has provided a program to address substrate inhibition. In addition, one might consider utilizing a Marmasse plot (Marmasse, 1963). Although rarely used, this plot provides good estimates of K_m and K_i . A Marmasse plot is a plot of $1/v$ versus $(\alpha + 1/\alpha)$, where $\alpha = [A]/[A]_m$ and $[A]_m$ represents the concentration of substrate at which the velocity is highest (note that this is not V_{\max} , as substrate inhibition is present.) This plot, when combined with the observation that $[A]_m =$

$(K_m/K_i)^{1/2}$, will provide values for the kinetic parameters (Cleland, 1979b).

CONTINUOUS VERSUS STOP-TIME ASSAYS

As the best estimate for the initial rate of an enzyme-catalyzed reaction derives from the slope of the steady-state region of the progress curve, continuous assays are greatly preferable to single-point assays. If it is not possible to use a continuous assay, the investigator must demonstrate that the single-point assay is a true measure of the steady-state velocity of the reaction. In the single-point method, an aliquot of the reaction mixture is removed from the sample at a given time point after the reaction has been initiated (after a “blank” aliquot has been removed from the reaction mixture prior to the initiation step). Aliquots from at least four different time points are removed from the reaction solution; then the reaction is terminated and the amount of product formed (or substrate depleted) is determined. It should be ascertained that the termination procedure truly quenches the reaction. Once the steady-state relationship has been established, replicate kinetic analyses at each substrate concentration will provide sufficient data for a statistical analysis. The number of replicate samples that are assayed should also be reported.

COUPLED ENZYME ASSAYS

In many instances, continuous assays are made possible by the presence of auxiliary enzymes. Product formation is measured by having the product react further with another enzyme (or enzymes). Although such assay systems are convenient, precautions must be exercised to ensure that the initial rates observed reflect the enzyme that is undergoing investigation and not the auxiliary system. If the auxiliary system is influencing the initial-rate results, the reported values for the kinetic parameters will be in error and the double-reciprocals may exhibit considerable nonlinearity.

The product or products formed by the enzyme undergoing study clearly direct the choice of auxiliary enzyme(s). This coupling system should provide turnover in excess over that due to the enzyme being investigated. However, the degree of excess will vary depending upon the assay method, the nature of the coupling system, and the magnitude of the kinetic parameters for each auxiliary enzyme as measured under the same conditions being

observed with the target enzyme. Once this information is in hand, the amount of auxiliary enzyme needed can be calculated (McClure, 1969; Storer and Cornish-Bowden, 1974). Test runs should then be undertaken using both the calculated and higher levels of auxiliary enzymes to insure that identical rates are obtained, indicating that a sufficient excess has been achieved. The auxiliary enzyme system should also have minimal effect on cofactors and effectors required by the target enzyme.

Frequently, a small amount of product may be present in the reaction mixture as a contaminant. For example, commercial preparations of ATP often include a small amount of ADP. The contaminating product can be removed prior to initiation of the assay by incubating the reaction mixture with the auxiliary system. The system should also be checked to insure that no activity is observed in the absence of the auxiliary enzymes. Finally, care should be exercised when using commercial preparations for the auxiliary enzymes. Often these sources of proteins contain contaminating activities that may cause serious difficulties unless they are addressed.

BINDING STUDIES

Measurements of equilibrium ligand binding can frequently add to the characterization of an enzyme's binding mechanism, and also act as a tool in such procedures as molecular-receptor binding studies and radioimmunoassays. A wide variety of methods have been utilized in characterizing binding phenomena—e.g., gel filtration and equilibrium dialysis. The theory underlying this valuable tool can be found in any physical biochemistry text. Commonly, the binding data are analyzed with Scatchard plots (Scatchard, 1949) in which the number of moles of ligand bound per mole of total protein (v) are plotted versus the ratio of v to the free molar concentration of the ligand. These plots are frequently used to obtain values for the ligand dissociation constant, to identify interacting sites, and to estimate the binding capacity of the protein. Care should be exercised in this last use of the binding data, as it is easy to obtain a poor estimate of the stoichiometry of binding from Scatchard plots (Klotz, 1982, 1983). For example, a Scatchard plot of the binding of diazepam to benzodiazepine receptors is reported to provide an incorrect total number of receptor sites. Often, such problems are due to having only a few data points in the Scatchard plot at which the ligand concentration is above the half-maximum bind-

ing point. Just as in kinetic data, concentrations of the binding ligand both above and below this half-maximum point should be investigated (Klotz, 1982). Feldman (1983) has recently presented a statistical analysis associated with ligand binding data and Scatchard plots. As this paper points out, estimates of molar binding capacity should always be reported with the appropriate degree of statistical certainty. Without such statistical analyses, reports of the number of binding sites can be meaningless.

PRESENTATION OF INITIAL-RATE DATA

The description of the assay protocols used in connection with enzyme purifications, in clinical studies, and in all kinetic investigations should include a clear definition of the unit of enzyme activity. When possible, velocities should be expressed in terms of molarity changes with time, as molarity is an intensive variable. Other units, proportional to molarity, are in common use. In all cases, the conditions of the assay should be fully described, including the volume of the initiating aliquot, the specific activity, and the amount of enzyme used in each experiment. The international unit (U), defined as the conversion of 1 μ mol of substrate to product per minute, is in common use. Another unit of catalytic activity, the katal, corresponding to the transformation of one mole of substrate per second, is also in use. It has been suggested that the latter unit be restricted to catalytic activities measured in a clinical context (Dybkaer, 1979).

Initial-rate data is frequently best presented via graphical displays. These visual linearizations of kinetic information provide a straightforward evaluation of the binding sequence and mechanism, as well as a ready appraisal of the kinetic parameters, reversible inhibition, cooperativity (both allosteric and hysteretic), and many other steady-state parameters. Whenever possible, slope and/or intercept replots should also be provided in the graphical display, perhaps as insets of the initial rate data. There are three methods frequently employed for depicting kinetic data, with the double-reciprocal plot (i.e., $1/v$ versus $1/[A]$; see Fig. A.3H.2) being the most common. The other common plots are $[A]/v$ versus $[A]$ and v versus $v/[A]$. Wilkinson (1961) has discussed each of these plots in terms of statistics. When the data are unweighted, the three graphical presentations are not statistically equivalent. However, this problem is considerably mitigated by proper analysis using computer methods and data weighting

(Siano et al., 1975). As mentioned earlier, Cleland (1979a) has provided a number of computer programs for the analysis of data. Eisenthal and Cornish-Bowden (1974) have also introduced a novel method for graphical treatment of data that provides a useful tool for evaluating experimental error. An Eisenthal–Cornish-Bowden plot is based on the equation $(V_{\max}/v) - (K_m/[A]) = 1$. Instead of a particular set of data (i.e., a single initial rate velocity with its respective substrate concentration) being represented by a point in a Cartesian coordinate system, the velocity (v) is denoted along the vertical axis and the substrate concentration ($[A]$) is denoted along the horizontal axis. The two points thus plotted are then connected by a line. Thus, whereas in the other graphical procedures the velocity and substrate-concentration data are represented by points, in the Eisenthal–Cornish-Bowden plot the same data are represented by lines. These lines, obtained from a set of initial-rate experiments, will intersect about a point equivalent to $v = V_{\max}$ and $[A] = K_m$. Because median values of intersection points are used, outlier (i.e., aberrant) data will have minimal effects on estimates of the kinetic parameters. In addition, this plot is very useful in that calculations are not required and data do not have to be weighted. However, it does have limitations (Cornish-Bowden and Endrenyi, 1981, 1986). For example, the plot can become quite unwieldy with multisubstrate enzymes and in inhibition studies.

It has recently been suggested (Brand and Johnson, 1992; Johnson and Brand, 1994) that all of the graphical methods described above may be unnecessary now that personal computers and associated new methods of data analysis are available. For example, global analysis of biochemical data is a powerful tool that may prove to be quite useful in the analysis of enzyme kinetics (Beechem, 1992). However, these numerical methods have not yet been commonly applied to studies such as those of multisubstrate systems, equilibrium exchanges, or product inhibition. At this time, with proper statistical weighting of the data, multisubstrate systems are more readily appraised by one of the common graphical procedures. More importantly, these visual depictions of initial-rate data are a clear aid to any individual reviewing the kinetic experiments.

CONCLUDING REMARKS

Initial-rate assays are extremely useful tools in the detailed characterization of an enzyme. The behavior, mechanism, and regulatory prop-

erties of an enzyme can be readily assessed by straightforward assays. Such assays can frequently be completed within a day, once preliminary runs have identified the magnitude of the kinetic parameters and revealed those factors that affect activity. However, the details of the kinetic mechanism and regulatory properties of the enzyme can easily be obscured by poor assay design. The structural/functional role of an enzyme, as well as its regulation and control, can only be assessed by a reliable initial-rate assay. In addition, the data assembled from these rigorous studies permit the development of more reliable and sensitive assay protocols for protein purification and identification.

LITERATURE CITED

- Allison, R.D. 1985. γ -Glutamyl transpeptidase: Kinetics and mechanism. *Methods Enzymol.* 113:419-437.
- Allison, R.D. and Meister, A. 1981. Evidence that transpeptidation is a significant function of γ -glutamyl transpeptidase. *J. Biol. Chem.* 256:2988-2992.
- Allison, R.D. and Purich, D.L. 1979. Practical considerations in the design of initial velocity enzyme rate assays. *Methods Enzymol.* 63:3-22.
- Beechem, J.M. 1992. Global analysis of biochemical and biophysical data. *Methods Enzymol.* 210:37-54.
- Bergmeyer, H.U. (ed.) 1983. *Methods of Enzymatic Analysis*, 3rd ed., Vols. 1-4. Verlag Chemie, Deerfield Beach, Florida.
- Boehlein, S.K., Richards, N.G.J., and Schuster, S.M. 1994. Glutamine-dependent nitrogen transfer in *Escherichia coli* asparagine synthetase B. *J. Biol. Chem.* 269:7450-7457.
- Bowen, W.J. and Kerwin, T.P. 1955. The purification of myokinase with an ion exchange resin. *Arch. Biochem. Biophys.* 57:522-524.
- Brand, L. and Johnson, M.L. (eds.) 1992. Numerical computer methods. *Methods Enzymol.* Vol. 210.
- Cleland, W.W. 1979a. Statistical analysis of enzyme kinetic data. *Methods Enzymol.* 63:103-138.
- Cleland, W.W. 1979b. Substrate inhibition. *Methods Enzymol.* 63:500-513.
- Cornish-Bowden, A. and Endrenyi, L. 1981. Fitting of enzyme kinetic data without prior knowledge of weights. *Biochem. J.* 193:1005-1008.
- Cornish-Bowden, A. and Endrenyi, L. 1986. Robust regression of enzyme kinetic data. *Biochem. J.* 234:21-29.
- Deans, N.L., Allison, R.D., and Purich, D.L. 1992. Steady-state kinetic mechanism of bovine brain tubulin:tyrosine ligase. *Biochem. J.* 286:243-251.

- Dybkaer, R. 1979. International Union of Pure and Applied Chemistry and International Federation of Clinical Chemistry. IUPAC Section of Clinical Chemistry; Commission on Quantities and Units in Clinical Chemistry and IFCC Committee on Standards, Expert Panel on Quantities and Units; Approved Recommendation (1978) Quantities and units in clinical chemistry. *Clin. Chim. Acta* 96:157F-183F.
- Eisenthal, R. and Cornish-Bowden, A. 1974. The direct linear plot. A new graphical procedure for estimating enzyme kinetic parameters. *Biochem. J.* 139:715-720.
- Feldman, H.A. 1983. Statistical limits in Scatchard analysis. *J. Biol. Chem.* 258:12865-12867.
- Fierke, C.A. and Hammes, G. 1995. Transient kinetic approaches to enzyme mechanisms. *Methods Enzymol.* 249:1-32.
- Frieden, C. 1959. Glutamate dehydrogenase III. The order of substrate addition in the enzymatic reaction. *J. Biol. Chem.* 234:2391-2396.
- Fromm, H.J. 1967. The use of competitive inhibitors in studying the mechanism of action of some enzyme systems utilizing three substrates. *Biochim. Biophys. Acta* 139:221-230.
- Johnson, M.L. and Brand, L. (eds.). 1994. Numerical computer methods, part B. *Methods Enzymol.* 240:xi-xiii, 1-36, 51-170, 181-198, 311-322, 781-816.
- Klotz, I.M. 1982. Numbers of receptor sites from Scatchard graphs: Facts and fantasies. *Science* 217:1247-1249.
- Klotz, I.M. 1983. Number of receptor sites from Scatchard and Klotz graphs: A constructive critique. *Science* 220:981.
- Klump, H., Di Ruggiero, J., Kessel, M., Park, J.-B., Adams, M.W.W., and Robb, F.T. 1992. Glutamate dehydrogenase from the hyperthermophile *Pyrococcus furiosus*. *J. Biol. Chem.* 267:22681-22685.
- Lienhard, G.E. and Secemski, I.I. 1973. P^1 , P^5 -Di(adenosine-5')pentaphosphate, a potent multisubstrate inhibitor of adenylate kinase. *J. Biol. Chem.* 248:1121-1123.
- Marmasse, C. 1963. Enzyme inhibition by excess substrate. *Biochim. Biophys. Acta* 77:530-535.
- McClure, W.R. 1969. A kinetic analysis of coupled enzyme assays. *Biochem.* 8:2782-2786.
- Nordlie, R.C. 1982. Kinetic examination of enzyme mechanisms involving branched reaction pathways: A detailed consideration of multifunctional glucose-6-phosphatase. *Methods Enzymol.* 87:319-353.
- O'Sullivan, W.J. and Smithers, G.W. 1979. Stability constants for biologically important metal-ligand complexes. *Methods Enzymol.* 63:294-336.
- Purich, D.L. and Fromm, H.J. 1972. Inhibition of rabbit muscle adenylate kinase by the transition state analogue, P^1, P^4 -di(adenosine-5')tetraphosphate. *Biochim. Biophys. Acta* 276:563-567.
- Rudolph, F.B. 1979. Product inhibition and abortive complex formation. *Methods Enzymol.* 63: 411-436.
- Scatchard, G. 1949. The attraction of proteins for small molecules and ions. *Ann. N.Y. Acad. Sci.* 51:660-673.
- Seelig, G.F. and Meister, A. 1985. Glutathione biosynthesis: γ -Glutamylcysteine synthetase from rat kidney. *Methods Enzymol.* 113:379-390.
- Siano, D.B., Zyskind, J.W., and Fromm, H.J. 1975. A computer program for fitting and statistically analyzing initial rate data applied to bovine hexokinase type III isozyme. *Arch. Biochem. Biophys.* 170:587-600.
- Silverstein, E. and Sulebele, G. 1969. Catalytic mechanism of pig heart mitochondrial malate dehydrogenase studied by kinetics at equilibrium. *Biochemistry* 8:2543-2550.
- Storer, A.C. and Cornish-Bowden, A. 1974. The kinetics of coupled enzyme reactions. *Biochem. J.* 141:205-209.
- Tate, S.S. and Meister, A. 1978. Serine-borate complex as a transition-state inhibitor of γ -glutamyl transpeptidase. *Proc. Natl. Acad. Sci. U.S.A.* 75:4806-4809.
- Wilkinson, G.N. 1961. Statistical estimations in enzyme kinetics. *Biochem. J.* 80:324-332.
- Williams, J.W., and Morrison, J.F. 1979. The kinetics of reversible tight-binding inhibition. *Methods Enzymol.* 63:437-467.

KEY REFERENCES

- Bergmeyer, 1983. See above.
- Detailed collection of enzyme assay protocols and methods for a wide variety of enzymes, including descriptions of many special techniques: e.g., fluorometry, turbidimetry, luminometry, and microtechniques.*
- Dixon, M. and Webb, E.C. 1979. Enzymes, 3rd ed. Academic Press, New York.
- Classic text on the basics of enzymology; Chapter 4 deals specifically with enzyme kinetics.*
- Fromm, H.J. 1975. Initial Rate Enzyme Kinetics. Springer-Verlag, New York.
- Valuable introductory text with considerable details on experimental design.*
- Kuby, S.A. 1990. A Study of Enzymes, vol. I. CRC Press, Boca Raton, Florida
- Detailed text on steady-state kinetics. Contains useful chapters on the effects of metal cofactors and transient phase kinetics.*
- Purich, D.L. (ed.) 1979. Enzyme kinetics and mechanism, part A. *Methods Enzymol.* Vol. 63.
- Describes a variety of initial rate methods and inhibitor studies for characterizing enzyme-catalyzed reactions.*

Purich, D.L. (ed.) 1980. Enzyme kinetics and mechanism, part B. *Methods Enzymol.* Vol. 64.

Describes a number of isotopic probes for studying enzyme mechanisms and contains valuable chapters on allosterism, hysteresis, immobilized systems, and processivity.

Purich, D. L. (ed.) 1982. Enzyme kinetics and mechanism, part C. *Methods Enzymol.* Vol. 87.

Describes methods for characterizing intermediates in enzyme-catalyzed reactions and using stereochemical probes for enzyme mechanisms; includes additional initial-rate and inhibitor methods and discusses further uses of isotopic probes.

Purich, D.L. (ed.) 1995. Enzyme kinetics and mechanism, part D. *Methods in Enzymol.* 249:3-662.

Covers in detail specialized topics in enzyme kinetics including transition-state approaches, kinetic probes with site-directed mutagenesis, partition analysis, positional isotope exchange, interfacial catalysis, and hydrogen tunneling.

Contributed by R. Donald Allison
University of Florida
Gainesville, Florida

Statistics for the Molecular Biologist: Group Comparisons

In this appendix, we will consider some of the statistical tests most commonly used (and misused) in biological research. The tests discussed here are those used for comparisons among groups (e.g., *t* test and ANOVA). A number of other important areas (e.g., linear regression, correlation, and goodness-of-fit testing) are not covered. The purpose of the Appendix is to enable you to determine rapidly the most appropriate way to analyze your data, and to point out some of the most common errors to avoid. Toward this end, we have included a flow chart to provide a quick guide to choosing the right statistical test (Fig. A.3I.1). Instead of including the voluminous statistical tables necessary to perform these tests, we assume you will have access to a spreadsheet software program such as Microsoft Excel or to a statistical software package to do the actual calculations involved and to supply critical values. Clearly, this Appendix is a very superficial treatment of statistics! There are of course many statistics texts available; a few that the authors have found particularly useful are listed at the end of this Appendix (see Literature Cited and Key References).

BASIC STATISTICAL BACKGROUND

Samples and Independence

It is important to think about how you will collect your data and what statistical tests you will perform before you begin your experiment. Suppose you want to compare the immune response of normal mice with that of mice carrying a mutation of interest. You examine a sample of mice from each of these populations. For any statistical test to be valid, all subjects in a sample must come from the same population, and subjects must be selected independently of each other. For example, selecting the same mouse twice would violate the independence assumption; on the other hand, selecting mutant mice from different genetic backgrounds to form one sample would violate the assumption that all mice in a sample are coming from the same population. For some experiments, subjects may be *measured* twice (or more) by design, and this information is incorporated into the statistical test. For example, mouse immune response could be meas-

ured before and after a treatment (see discussion of Paired *t* Test), or several measurements may be made from each mouse (see discussion of Multiple Comparison Testing). However, each mouse in the sample is still chosen independently of other mice.

Hypothesis Testing

Statistical tests always test a null hypothesis, H_0 . The null hypothesis usually states that there is no difference between two (or more) populations. If the null hypothesis is true, any observed differences between samples from these populations arose by chance alone. The alternative hypothesis, H_a , states that differences observed in the samples reflect a real difference between the populations from which they were drawn. We usually hope to reject H_0 , in order to show something interesting. In our immune response experiment, for example, our H_0 would be that mutant mice have the same response as normal mice. We hope to be able to reject H_0 , and show that mutant mice have a different immune response from normal mice.

Although statistical tests can never tell us with absolute certainty whether our samples arose from two different populations, they can give us an idea about how probable our H_0 is. What we would like to know is the probability that any difference we find between samples arose by chance alone. Unfortunately, statistical tests cannot tell us this probability. What they do tell us is a related probability—i.e., the probability that we would see at least as large a difference between samples as we do, given that the null hypothesis is true. This probability is the famous *p* value.

For example, if $p = 0.50$, it means that there was a 50% chance of getting a difference between samples at least as large as we did, if the populations we sampled were really the same. This is a large probability, so we would accept the null hypothesis H_0 , and conclude that there is no difference between the two populations we sampled. The accepted standard is that we should reject H_0 if $p \leq 0.05$ —i.e., there was only a 5% chance (or less) of getting at least as large a difference between samples as we did, if the populations sampled were really the same. The value of *p* below which we will reject the H_0 is called the α value; α is usually chosen as 0.05. A difference between samples that results in a

p value of $\leq \alpha$ is called statistically significant. Note that statistical tests are not black and white; if $p = 0.05$, there is a 5% chance that we will reject the null hypothesis when it is actually true! This is called a Type I error. You might think we could avoid this error by only rejecting H_0 if the p value is even smaller; say, 0.01 or 0.001. However, if we do this, we are likely to accept H_0 when it is false! This is called a Type II error. Setting $\alpha = 0.05$ is a standard that is

generally accepted as a good compromise between Type I and Type II errors.

In practice, once you have chosen a statistical test, you calculate (or the computer calculates) the value of a statistic from your sample data using a formula. You then compare this value to a critical value in a table (or provided by the computer) to decide whether to accept or reject H_0 . The critical value (e.g., F_c for the F test, or t_c for the t test; see discussion of The

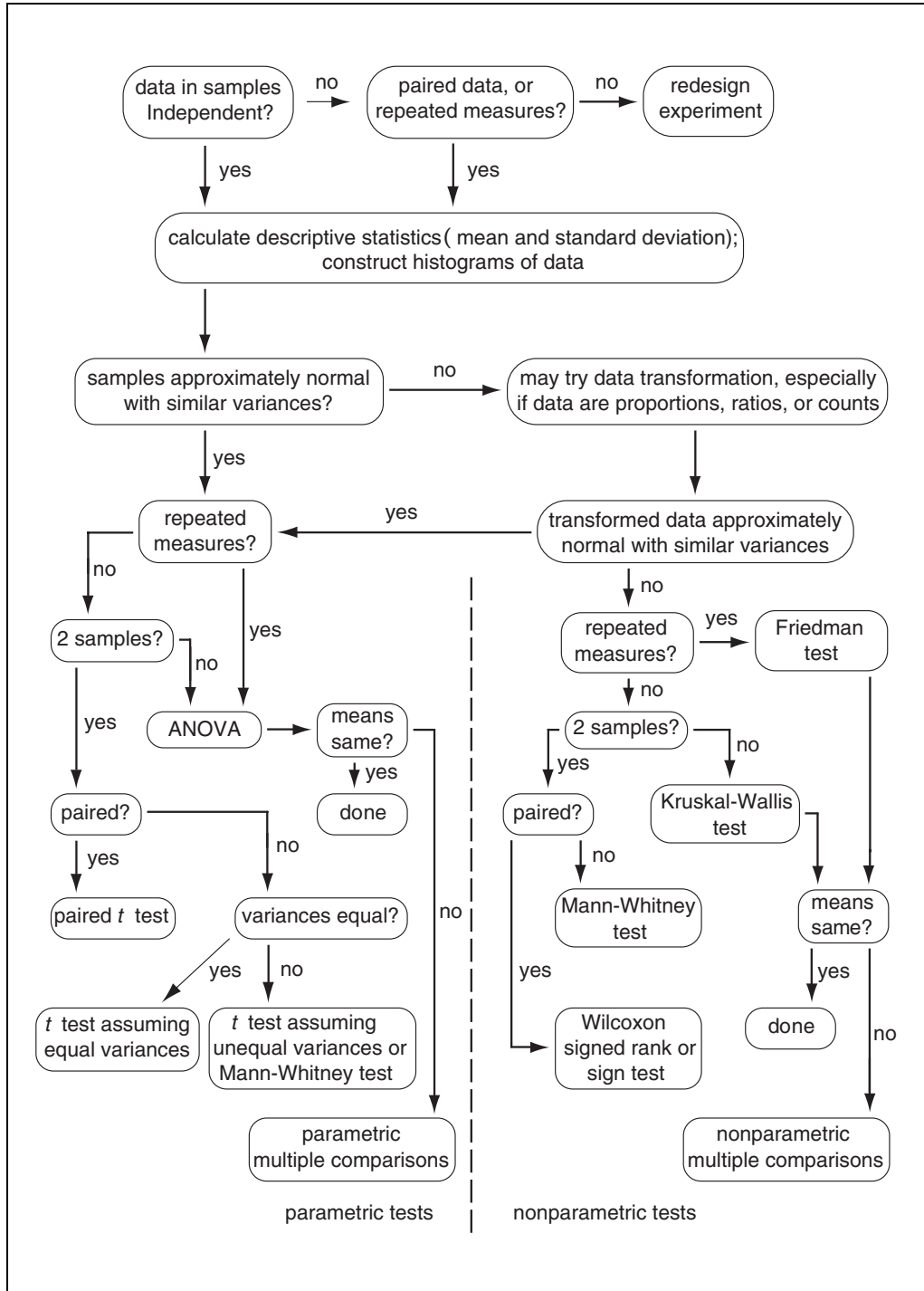


Figure A.3I.1 Flow chart for comparisons among groups.

t Test) is determined by the *p* value you have decided upon (usually, *p* = 0.05) and the degrees of freedom (df) of the data, which vary depending on the test being performed. Usually, you reject H_0 if your statistic is greater than the critical value.

Most tables (and statistical packages) give critical values with one-tailed and two-tailed levels of significance. You are doing a two-tailed test when your H_a does not specify the direction of difference between populations. You perform a one-tailed test when your H_a predicts, based, for example, on a theory that you have, that populations will differ in a particular way. For example, if you think that there may be a difference between the immune response of mutant and normal mice, but you can't predict whether the mutants will do better or worse, you should perform a two-tailed test. If you predict that the response of the mutant mice will be reduced (due to the loss of your gene of interest), you could perform a one-tailed test. There is some controversy in the literature over the use of one-tailed tests. When in doubt, perform a two-tailed test; it is always more conservative (i.e., less likely to reject the null hypothesis).

CHOOSING A TEST— EXPLORATORY DATA ANALYSIS

Parametric and Nonparametric Tests

It can be difficult to sort out what test is appropriate to apply to your data when confronted with a bewildering array of possible tests and formulas. One important distinction you need to make is between parametric and nonparametric tests. Parametric tests (such as the *t* test) assume your data follow a statistical distribution with known mathematical properties (the normal distribution, for the parametric tests we will discuss). Often, parametric tests require that the variability of all the samples you are comparing be the same. Nonparametric tests make fewer assumptions about the distribution of your data. Nonparametric tests are thus less likely than parametric tests to commit a Type I error (rejecting H_0 when it is true), but are usually less powerful (that is, more likely to commit a Type II error, and fail to detect a difference between groups). However, with data sets of reasonable size (say, 25 or more data points per group), nonparametric tests are often nearly as powerful as parametric ones. In order to determine which type of test is appropriate to your data, the first thing you need to do is examine the characteristics of that data.

What Are Parameters?

In order to perform parametric tests, we need to calculate parameters—i.e., numbers that summarize the data in some descriptive way. The statistical parameters we want to calculate are the mean (average) and amount of variability of the population (variance and standard deviation). Your data are a sample from a large population. We estimate the population parameters by sampling a subset of the population. The larger the sample size (*n*), the closer your sample statistics are to the true population parameters.

Any spreadsheet program or statistical package will allow you to determine the mean, variance, and standard deviation of your data; we show the formulas for these parameters below so that you can better interpret what such programs are telling you about the data.

The following is the equation for the sample mean \bar{x} :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Here, x_i is the value of the *i*th data point, and *n* is the number of data points in the sample.

The sample variance (s^2) is expressed as follows:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

The sample standard deviation (*s*), is just the square root of the variance:

$$s = \sqrt{s^2}$$

You can think of the standard deviation as the average deviation from the mean; the magnitude of *s* tells you how much variability there is in the data. If your data are normally distributed, then ~68% of the data should fall within 1 standard deviation of the mean and ~95% of your data should fall within 2 standard deviations of the mean.

The final parameter of interest is the standard error of the mean (s.e.m.; also called the standard error):

$$\text{s.e.m.} = \frac{s}{\sqrt{n}}$$

The standard error of the mean tells you how much the mean of a sample of size *n* taken from your population varies. That is, if you were to

take a number of different samples of size n from your population, the mean of each sample would be slightly different. If you were to plot the means of a number of such samples, they would be distributed across a range of values, but there would be far less variability in those values than in the data of the original samples. While the standard deviation gives you a description of the variability of your data, the s.e.m. is the parameter used to compare two sample means and to determine whether or not they are statistically different. For this reason, the s.e.m. is what is normally plotted as the “error bars” around the mean on a graph showing samples that are being compared. Note that the s.e.m. does not tell you much about the variability of your original data; you can see from the equation for s.e.m. that even highly variable data can give a small s.e.m. if the sample size is large enough.

Suppose we are interested in the ability of neurons to sprout neurites in response to growth factors. We want to compare a population of neurons exposed to nerve growth factor (NGF), a factor of known importance, with a population exposed to factor X. Our H_0 would be that neurite outgrowth in response to NGF is the same as neurite outgrowth in response to factor X. We hope to be able to reject H_0 , and show

Table A.3I.1 Neurite Outgrowth Data

Neurite outgrowth (μm)	
NGF	Factor X
22.5	43.2
35.4	56.9
38.2	60.2
40.4	62.4
41.3	65.7
46.7	75.8
55.4	85.4
62.5	91.3
81.2	95.3
99.2	105.2

Table A.3I.2 Sample Parameters from Neurite Outgrowth Data

	NGF	Factor X
Mean (\bar{x})	52.28	74.14
Variance (s^2)	535.00	388.26
Standard deviation (s)	23.13	19.70
Standard error (s.e.m.)	7.31	6.23

that neurons respond differently to factor X. Tables A.3I.1 and A.3I.2 show a (rather small) data set, with calculated sample parameters.

If these data are normally distributed, then ~68% of the data for NGF should fall within one standard deviation of the mean—i.e., between 29.15 and 75.41 (52.28 ± 23.13). In fact, $\frac{7}{10}$ or 70% of the data points fall within these limits. About 95% of the data should fall within two standard deviations of the mean—between 6.02 and 98.54 ($52.28 \pm 2 \times 23.13$). In fact, 90% of the data fall within these limits. The distribution of factor X data is similar. This is reasonable agreement with the normal distribution, given the small sample sizes.

Determining Whether Data Are Normally Distributed with Equal Variances

In order to determine whether to use a parametric or nonparametric test on your data, it is important to determine whether the data are normally distributed, with equal variances. Statistical tests for normality exist (e.g., the Kolmogorov-Smirnov test; Zar, 1984), and we will discuss tests for homogeneity of variance later. However, you can often find out most of what you need to know about your data by simply plotting it and looking at it.

For reasonably large data sets, the parametric tests that we will discuss are not much affected by small departures from normality (this property is called being “robust” to small departures from normality). As a rule of thumb, if you have ≥ 25 samples in each group that you are comparing, you can simply plot your data using a frequency histogram and examine it visually to see if it falls roughly into a bell-shaped curve. Our neurite outgrowth data is plotted in Figure A.3I.2. In this example, the data points are continuous numbers, so in order to make a histogram, we had to divide the data up into arbitrary “bins” (e.g., 0 to 20, 20 to 40, and so forth). The samples in our data set look approximately normally distributed, although the sample sizes are a bit small to be really sure. The variability in the two data sets also looks similar by eye. Looking back at our sample parameters, the calculated variance and standard deviation of the two data sets look reasonably similar.

If you have a small data set, it may be difficult to determine visually whether your data are normally distributed, and statistical tests also do not have much power to make the determination. If you use a parametric test in this situation, and the data are really not drawn

from a normal distribution, then you may reject your null hypothesis when it is actually true. If you use a nonparametric test, you may accept your null hypothesis when it is false. It is more conservative to use nonparametric tests in this situation; but you can see that the best solution (when possible) is to collect reasonably large data sets!

If visual inspection suggests your data are not normally distributed, or if the variance differs widely among groups, it is sometimes useful to perform a transformation (see discussion of Data Transformations). The transformed data points may then form a normal distribution with similar variances among samples, and parametric statistical tests can be performed on the transformed data points. If neither the original nor the transformed data produce approximately normally distributed data with similar variances, then nonparametric statistical tests generally should be used.

Data Transformations

The parametric tests that we will describe usually assume that the data are normally distributed, and that the variances are the same for all groups being compared (this latter property is known as “homoscedasticity”). There are a number of common situations—e.g., in which the data are expressed as proportions or counts

of randomly occurring objects or events—where these assumptions are violated, but you can still use a parametric test if you first perform a data transformation.

Data expressed as proportions of 0 to 1.00 (i.e., 0% to 100%) are not normally distributed. However, you can make them nearly normally distributed, with reasonably homogeneous variances among groups, by performing the following transformation for each data point, and then using the p' values for all of your calculations (e.g., mean, and standard deviation):

$$p' = \arcsin \sqrt{p}$$

When the data are counts of randomly occurring objects (such as the number of butterflies per square foot of a field) or events (such as the number of cars passing by a location), the data will follow a Poisson rather than a normal distribution. Group variances will then be proportional to their means. In this case, the square root transformation, shown in the following equation, should be performed. As with proportions, statistical tests are then performed on the transformed values.

$$x' = \sqrt{x + 0.5}$$

If variances among groups are very different and if the standard deviations are proportional

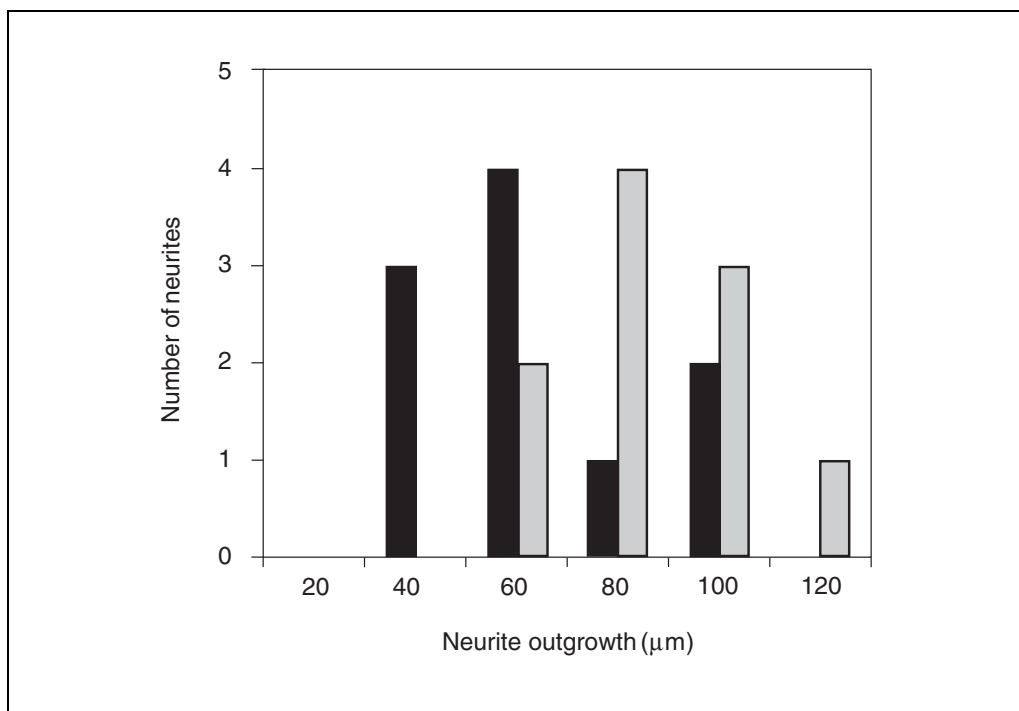


Figure A.3I.2 Neurite outgrowth histogram. Black bars correspond to results for nerve growth factor (NGF) gray bars correspond to results for “factor X.”

to the means of your groups, the logarithmic transformation can be used:

$$x' = \log(x + 1)$$

If your data are expressed as ratios, a logarithmic transformation is likely to make them more nearly normally distributed (for an example, see discussion of Ratio Paired *t* Tests).

Other types of transformations are possible (Zar, 1984). If your data still don't fit the assumptions of a parametric test after being transformed, then a nonparametric test should be used.

STATISTICAL TESTS FOR COMPARISONS BETWEEN TWO UNPAIRED GROUPS

The *t* Test

The *t* test is a parametric test that allows us to determine whether the means of two samples are statistically different. The nonparametric equivalent is the Mann-Whitney test (see discussion below; also called the Wilcoxon Rank Sum test).

Assumptions of the *t* test

1. Only two samples are being compared. If more than two samples are compared, you *must* use a test specifically designed for multiple comparisons, such as ANOVA (see discussion under Multiple Comparison Testing). Using multiple *t* tests to compare among three or more samples is invalid, and the computed *p* values will be wrong. This is one of the most common errors in the use of the *t* test.

2. Subjects in samples were chosen independently. If repeated measurements were made from individual subjects, ANOVA (or the nonparametric equivalent) should be used.

3. The two samples were not paired in any way (such as before-and-after measurements or matched cases and controls). If samples were paired, the paired *t* test (or the nonparametric Wilcoxon Signed Rank test; see discussion below) should be used.

4. The samples are approximately normally distributed. If this assumption is not met, either transform your data (see discussion of Data Transformations) or use the nonparametric Mann-Whitney test instead.

The simplest form of the *t* test assumes that both samples have the same variance. A somewhat more complicated form is necessary if the sample variances are different. Therefore, in order to determine which form of the *t* test to

use, we must first determine whether the sample variances are statistically different, by using the variance ratio test, or *F* test. Figure A.3I.3 shows a flow chart for the *t* test.

Performing the *F* test

The *F* test is summarized as follows:

$$H_0: s_1^2 = s_2^2$$

$$F = \frac{s_1^2}{s_2^2} \text{ or } F = \frac{s_2^2}{s_1^2}, \text{ whichever is larger}$$

(*F* will always be ≥ 1)

reject H_0 if $F > F_c$; accept H_0 if $F < F_c$

In the above equations, *F* is referred to as the "*F* statistic," s_1^2 and s_2^2 are the variances of sample 1 and sample 2, respectively, and F_c is the critical value.

Statistical computer packages will report the *F* statistic, along with the appropriate critical value, F_c . Note that if the variances of the two samples are very different, *F* will be large, so it makes sense to reject H_0 if $F > F_c$. If we reject H_0 , we would use the *t* test assuming unequal variances. If we accept H_0 , we would use the *t* test assuming equal variances. If the differences in variance are extreme, a data transformation may be indicated or the (nonparametric) Mann-Whitney test could be performed instead of the *t* test.

F_c is chosen based on the significance level (usually $\alpha = 0.05$), and the degrees of freedom (df) in the data. The value of df is equal to one less than the number of data points in each sample. In the case of the experiment in Table A.3I.1, which will be used as an example here, there are 10 data points in each sample, so $df = 9$ for each sample. F_c is usually written " $F_{\text{numerator df, denominator df, } \alpha}$." For our example, F_c would be written $F_{9,9,0.05(2)}$. The 2 in parentheses denotes that this is a two-tailed critical value. If the computer package reports both a one-tailed and two-tailed critical value, use the two-tailed value. You would only use a one-tailed value if you had reason to think *before* computing the variances that one variance would be larger than the other. When in doubt, use a two-tailed test.

Performing the *t* test

The *t* test with equal variances and sample sizes is summarized as follows:

$$H_0: \bar{x}_1 = \bar{x}_2$$

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2 + s_2^2}{n_1}}}$$

reject H_0 if $t > t_c$; accept H_0 if $t < t_c$

In the above equations, t is referred to as the “ t statistic” and t_c is the critical value.

Statistical packages will report t and t_c . To choose t_c from a table, the number of degrees of freedom is calculated as the sum of the number of data points in samples 1 and 2, minus 2 ($df = n_1 + n_2 - 2$). The critical value, t_c , is often written $t_{df,0.05(2)}$ for a two-tailed test. As with the F test, use a two-tailed test if both one- and two-tailed critical values are given.

Examining the formula for the t statistic, you can see that t will be large if the difference between the sample means is large. However, as the variances of the samples get larger, the t statistic gets smaller. Thus, the equation makes intuitive sense; you will be able to reject the null hypothesis if the difference between means is large and the scatter in the data is small, so that there is little overlap between the two samples.

There are several different formulas for the t test with unequal variances, which give

slightly different results; they are not presented here (see Zar, 1984). The t test with unequal variances tests the same H_0 and follows the same intuitive reasoning given above for the t test with equal variance. However, the calculation of df for this test is quite different than for the test with equal variance—they are calculated using a more complicated formula—a statistical package will calculate them, and print them for you.

Example of the t test

Let us assume that our neurite outgrowth data (see Table A.3I.1 and Table A.3I.2) fit all the assumptions for the t test, and let us perform the t test on this data set.

H_0 for the F test: variances of the two samples are equal.

$$F = 535.00/388.26 = 1.378.$$

$$F_c = F_{9,9,0.05(2)} = 4.03.$$

Thus, $F < F_c$, so we accept H_0 and conclude that the variances of the two samples are not statistically different. We proceed to the t test assuming equal variances.

H_0 for the t test: means of the two samples are equal.

We use a spreadsheet to calculate the t statistic. The computer reports:

$$t = 2.275$$

$$t_c = t_{18,0.05(2)} = 2.10.$$

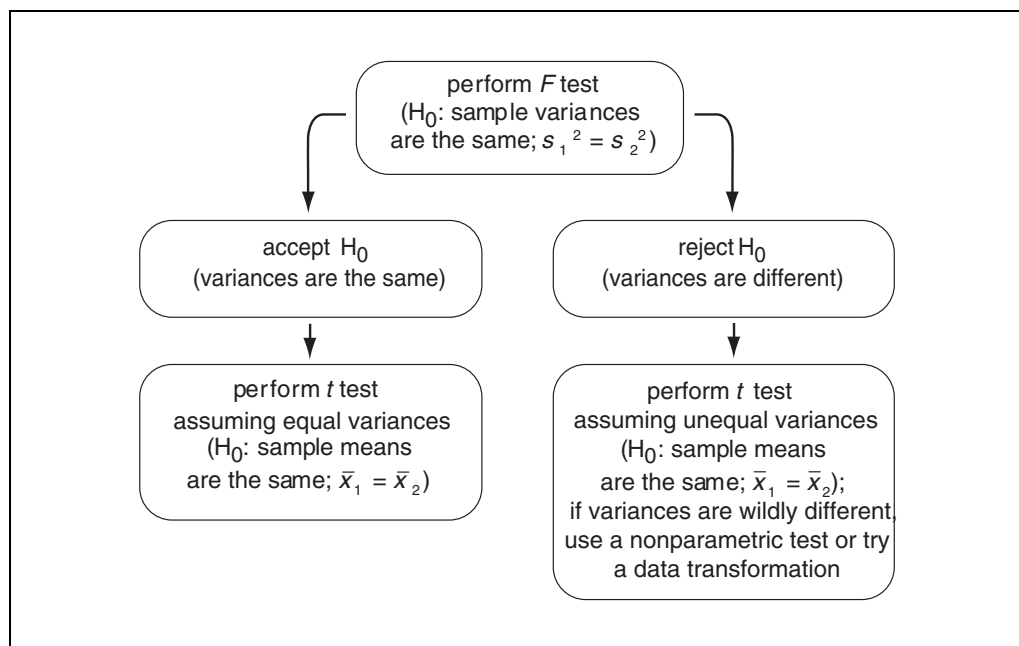


Figure A.3I.3 Flow chart for the t test. Here, we want to compare sample 1 and sample 2, with means \bar{x}_1 and \bar{x}_2 , variances s_1^2 and s_2^2 , and sample sizes n_1 and n_2 . Whichever t test you use, accepting H_0 for the t test says that the sample means are the same; rejecting H_0 for the t test says that the sample means are different.

Thus, $t > t_c$, so we reject H_0 and conclude that the means of the two samples are different, with statistical significance at the $\alpha = 0.05$ level. The calculated p value for our data is $p = 0.035$. Note that some spreadsheets print out negative values for the t statistic; in this case, simply use the absolute value.

The Mann-Whitney Test

The Mann-Whitney test (also called the Wilcoxon Rank Sum test) is the nonparametric equivalent to the t test, and is used to determine if two samples are statistically different, with no assumptions about whether the data are normally distributed.

This test is nearly as powerful as the t test. When you perform the Mann-Whitney test, you are comparing the order, or rank, of the data rather than its numerical value. To perform the Mann-Whitney test, you rank your entire data set (treating all data from both samples as a single list). For our neurite outgrowth data, the largest data point is 105.2; it has a rank of 1. The next largest is 99.2, with a rank of 2, and so on. You then sum the ranks for each individual data set.

If your two samples have similar values, the ranks would tend to go back and forth between samples. For example, the highest-ranking value might be in sample 1, the next two highest in sample 2, the next in sample 1, and so on. If you then summed the ranks from each sample, the two sums would be about equal. On the other hand, if sample 1 had mostly larger numbers than sample 2, the ranks of sample 1 would all be small numbers, and the sum of the ranks of sample 1 would be much smaller than that of sample 2. The Mann-Whitney test tells us how different the rank sums must be to indicate a significant difference between the two samples.

Assumptions of the Mann-Whitney test

Assumptions 1 to 3 of the t test (see discussion The t Test) apply also to the Mann-Whitney test. The Mann-Whitney test, however, makes no assumption about normality of data

Performing the Mann-Whitney test

The Mann-Whitney test is summarized as follows.

H_0 : There is no significant difference between the two samples (of sizes n_1 and n_2).

1. Order all the data from both samples into a single list, from greatest to smallest value, keeping track of which sample the data came from. Tied values get the average of the two

ranks they would have had if they had not tied (e.g., if two values tie for 8th largest, they would have been 8th and 9th largest, they each get a rank of 8.5).

2. Sum the ranks for each sample. Call the sums R_1 and R_2 .

3. Calculate the U and U' statistics according to the following equations:

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U' = n_1 n_2 - U$$

4. If either U or U' is greater than the critical value, $U_c (U_{0.05(2),n_1,n_2})$, in a table of U values; see, e.g., Zar, 1984), then reject H_0 .

5. With more than 20 measurements in one or both samples, the Z statistic is used, as calculated by the following equation. Only U or U' needs to be calculated in this case, not both.

$$Z = \frac{|U - \bar{U}| - 0.5}{s}$$

where

$$\bar{U} = \frac{n_1 n_2}{2} \text{ and}$$

$$s = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

The Z statistic is approximately normally distributed, even though the original data may not be at all normally distributed. For $\alpha = 0.05$, if $Z > 1.96$, H_0 is rejected.

Example of the Mann-Whitney test

Let us say that we were a bit uncomfortable assuming that our neurite outgrowth data (see Table A.3I.1 and Table A.3I.2) are really normally distributed, and we decide to perform the Mann-Whitney test instead of the t test. These data are shown ranked for the Mann-Whitney test in Table A.3I.3.

From the data in Table A.3I.3 we can make the following calculations, based on the above equations for the Mann-Whitney test:

$$U = 100 + 55 - 134 = 21$$

$$U' = 100 - 21 = 79$$

$$U_c = U_{0.05, 10, 10} = 77$$

Since $U' > U_c$, H_0 is rejected and it is concluded that the two samples are significantly different.

Table A.3I.3 Neurite Outgrowth Data Ranked for the Mann-Whitney Test

Outgrowth	Rank (NGF)	Rank (factor X)
NGF		
22.5	20	—
35.4	19	—
38.2	18	—
40.4	17	—
41.3	16	—
46.7	14	—
55.4	13	—
62.5	9	—
81.2	6	—
99.2	2	—
Factor X		
43.2	—	15
56.9	—	12
60.2	—	11
62.4	—	10
65.7	—	8
75.8	—	7
85.4	—	5
91.3	—	4
95.3	—	3
105.2	—	1
Rank sum	134	76

Note that both the t and Mann-Whitney statistics were very close to the critical values for these data. The power of the Mann-Whitney test to reject H_0 is almost as great as that of the t test.

COMPARING TWO PAIRED GROUPS: PAIRED t TEST, WILCOXON SIGNED RANK TEST, AND SIGN TEST

When to Use a Paired Analysis

A paired analysis is indicated under the following circumstances.

1. You measure a variable before and after some treatment is performed upon a subject.
2. You match subjects in pairs (e.g., for age, sex, and exposure) and then treat one of the subjects and not the other (or the other receives alternative treatment).
3. You compare relatives (e.g., siblings or parent/child).

4. You perform an experiment several times, each time with experimental and control preparations treated in parallel.

Paired t Test

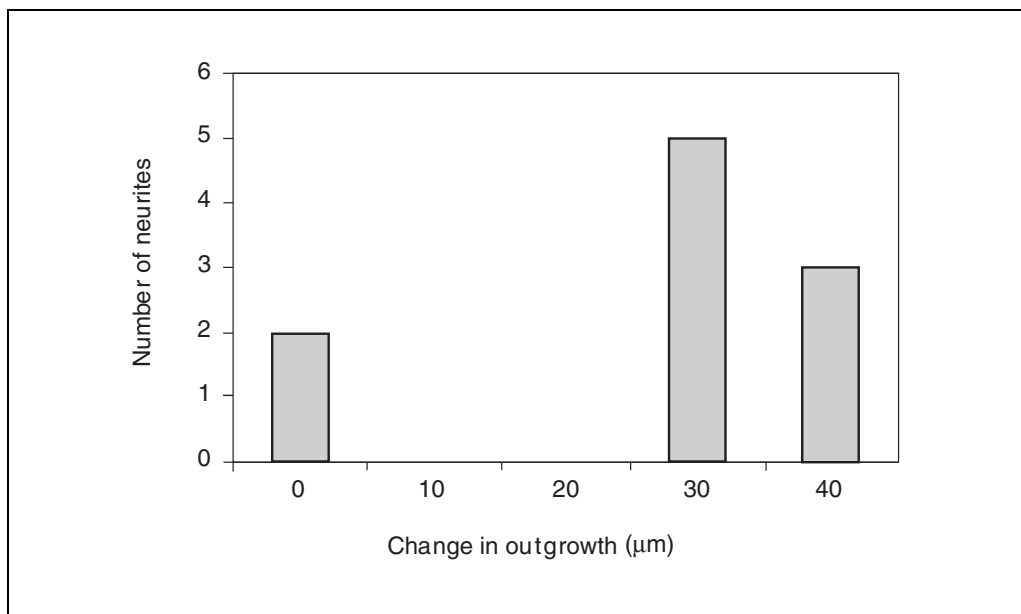
The paired t test is a parametric test used to compare two groups whose members are paired. When the assumptions of a paired t test are met, another approach is to perform a repeated-measures ANOVA (see discussion under Multiple Comparison Testing) with only two groups. However, most nonstatisticians find the paired t test easier to understand and perform.

Assumptions of the paired t test

1. Pairs must be representative of a larger population. Each pair must be selected independently of other pairs.
2. Pairs must be made *before* data is collected.
3. The calculated *differences* between the two members of each pair must be approximately normally distributed.

Table A.3I.4 Paired Neurite Outgrowth Data

Neurite outgrowth (μm)		
NGF	Factor X	Difference
22.5	43.2	20.7
35.4	56.9	21.5
38.2	65.7	27.5
40.4	62.4	22.0
41.3	75.8	34.5
46.7	85.4	38.7
55.4	95.3	39.9
62.5	60.2	-2.3
81.2	105.2	24
99.2	91.3	-7.9

**Figure A.3I.4** Change in neurite outgrowth after addition of factor X, in cells treated first with NGF.**Performing the paired *t* test**

The paired *t* test is summarized as follows.

Assume you have a set of n pairs: $x_{11}, x_{21}, x_{12}, x_{22}, \dots$ through x_{1n}, x_{2n} .

H_0 : There is no difference between members of pairs. (i.e., the mean difference between the pairs, \bar{d} , is equal to 0).

1. For each pair, calculate the difference between the pairs ($d_i = x_{1i} - x_{2i}$).

2. Calculate the mean, \bar{d} , and the s.e.m. of these differences.

3. Calculate the *t* statistic according to the following equation:

$$t = \frac{\bar{d}}{\text{s.e.m.}}$$

4. Determine t_c from a table using $df = n - 1$. Note that each pair counts as 1 in computing the df ; for the unpaired *t* test, each data point counted as 1.

5. If $t > t_c$, reject H_0 .

Example of the paired *t* test

For our neurite outgrowth data, suppose that instead of looking at two different sets of neurons treated separately with NGF and factor X, we were looking at neurons treated sequentially, first with NGF and then with factor X. We were able to keep track of each individual neurite, and want to know if neurites continued to grow out in the presence of factor X after NGF was removed. We will analyze the same

Table A.3I.5 Original and Transformed Densitometry Data from Northern Blotting of Mutant Versus Wild-Type Gene Expression^a

Mutant	Wild-type	Ratio: mutant/ wild-type	log(mutant)	log(wild-type)	log(mutant) – log(wild-type)
500	300	1.67	2.6990	2.4771	0.2218
6000	4200	1.43	3.7781	3.6232	0.1549
750	500	1.5	2.8751	2.6990	0.1761

^aValues expressed in arbitrary densitometry units.

data as before, but now it is paired, as shown in Table A.3I.4 and Figure A.3I.4.

H_0 : There is no change in neurite outgrowth when factor X is added to the medium

$$\bar{d} = 21.86$$

$$\text{s.e.m.} = 5.03$$

$$t = 21.86/5.03 = 4.35$$

$$t_c = t_{9,0.05(2)} = 2.26$$

$$t > t_c; \text{ therefore reject } H_0.$$

The p value for the paired t test is $p = 0.0018$ (the p value for the unpaired test on the same data was $p = 0.035$). Clearly, having data that are paired can give us much more power to reject H_0 .

If we are uncertain whether our differences are normally distributed, we can use a nonparametric alternative (either the Wilcoxon Signed Rank test, or the Sign test; see discussion below).

Ratio Paired t Tests

Sometimes it makes sense to compare values as ratios rather than as differences. For example, you might want to look at densitometric data from a northern blot of your favorite gene's expression in a mutant animal, compared to the signal from a wild-type animal. For each experiment, the densitometry units can be adjusted arbitrarily, and may vary widely depending on how good your probe is and how long the exposure was. However, the ratio of the mutant to wild-type readings is what is important. In this case, what you really want to know is whether the ratio of mutant to wild-type is significantly different from 1 (implying that gene X is expressed either more or less in the mutant than in the wild type).

Ratios are not normally distributed, but logarithms of ratios are quite likely to be normally distributed. The logarithm of a ratio can be written as a difference:

$$\log\left(\frac{\text{mutant}}{\text{wild-type}}\right) = \log(\text{mutant}) - \log(\text{wild-type})$$

Thus, to analyze data of this type, simply take the logarithm of each data point and perform a paired t test on this transformed data (Table A.3I.5). If the transformed data do not appear to be approximately normal, the Wilcoxon Signed Rank test can be used instead. The null hypothesis for the transformed data is that the mean difference is 0, as before; this means that for the original (untransformed) data, the null hypothesis is that the ratio of mutant to wild-type expression is 1. Rejecting the null hypothesis says that the amount of expression in mutant and wild-type animals is significantly different.

Let us perform the calculation as follows on the data in Table A.3I.5.

H_0 : Ratio of mutant to wild-type expression = 1 (there is no difference between mutant and wild-type animals).

Transform data, and perform paired t test of differences of logarithms of data as described for the paired t test.

$$\bar{d} = 0.1843$$

$$\text{s.e.m.} = 0.019755$$

$$t = 9.33$$

$$t_c = t_{2,0.05(2)} = 4.30$$

$$t > t_c$$

Thus, reject H_0 and conclude that there is a significant difference in expression between mutant and wild-type animals. (the calculated p value was 0.01). Note that this is a very significant result, even with only three pairs of data.

Wilcoxon Signed Rank Test

The Wilcoxon Signed Rank test is a nonparametric equivalent to the paired t test. It is used for data that may not be normally distributed.

Table A.3I.6 Differences Computed from Neurite Outgrowth Data, and Ranked for Use in the Wilcoxon Signed Rank Test

Difference	Rank (-)	Rank (+)
20.7	—	3
21.5	—	4
27.5	—	7
22.0	—	5
34.5	—	8
38.7	—	9
39.9	—	10
-2.3	1	—
24	—	6
-7.9	2	—
Rank Sum	3	52

Assumptions of the Signed Rank test

Assumptions 1 and 2 of the paired *t* test apply also to the Wilcoxon Signed Rank test. An additional assumption is made that the sampled population is symmetrical about the median. If the data do not fit this assumption very well, the Sign test (see discussion below) can be performed.

Performing the Signed Rank test

The Wilcoxon Signed Rank test is summarized as follows.

H_0 : There are no differences between members of pairs.

1. For each pair, calculate the difference between the members of the pair. Keep track of the sign.
2. Rank the absolute value of all differences from low to high.
3. Sum the ranks of the negative differences (call the sum T^-) and the ranks of the positive differences (call the sum T^+).
4. Reject H_0 if either T^- or T^+ is less than or equal to the critical value, T_c , as determined from a Wilcoxon Signed Rank table.

Example of the Signed Rank test

If we were concerned about the normality of our neurite outgrowth differences (Table A.3I.1 and Table A.3I.2), we could perform the Wilcoxon Signed Rank test on these data rather than the paired *t* test we did earlier. Consider the data in Table A.3I.6.

H_0 : There is no change in neurite outgrowth when factor X is added to the medium.

$$T^- = 3; T^+ = 52$$

$$T_c = 8$$

$T^- < T_c$; therefore reject H_0 .

The Sign Test

The Sign test is another nonparametric paired test that comes in especially handy when you are sure of the *direction* of a difference between the members of your pair (positive or negative), but you may not know exactly how much difference there is.

Assumptions of the Sign test

The assumptions of the Sign test are the same as those of the Wilcoxon Signed Rank test, but no symmetry assumption is made. This is a less powerful test.

Performing the Sign test

The Sign test is summarized as follows.

H_0 : There are no differences between members of pairs.

1. For each pair, record whether the first member is greater than the second (+) or less than the second (-). Ignore any pairs with equal members. C^+ is the total number of (+) pairs; C^- is the total number of (-) pairs.
2. Reject H_0 if either C^+ or C^- is less than or equal to the critical value of C for the Sign test (C_c).

Example of the Sign test

H_0 : There is no change in neurite outgrowth when factor X is added to the medium.

For our neurite outgrowth data (Table A.3I.6), $C^+ = 8$ and $C^- = 2$. For $n = 10$ and $p =$

0.05, the critical value, C_c is equal to 1. Therefore, H_0 cannot be rejected.

MULTIPLE COMPARISON TESTING

Why Not Lots of t Tests?

If we find that we want to compare multiple groups in an experiment, then the t test just will not do. Using pairwise comparisons ignores the experimental effects in our design in all but the two groups being compared. Even if it were appropriate to use multiple pairwise comparisons (and it is not), we would still run afoul of the problem of hypothesis rejection. If we used the typical $\alpha = 0.05$, that would mean that for every 20 pairwise tests we ran, we would expect to get the wrong answer for one of them. This is not a desirable situation. There are a number of methods, however, that avoid these problems by allowing us to consider our entire design in our model for analysis. The most common in current literature is the analysis of variance (ANOVA). There are classical and nonparametric methods to analyze variance differences, and we will first consider the classical, or parametric case.

Assumptions of ANOVA

As with any statistical test, there are several assumptions of the classical analysis of variance that need to be examined. Some of these are more critical than others, and we will try to make that clear. In this chapter, we are dealing only with one-way analysis of variance; additional assumptions hold for more complex ANOVA procedures. If in doubt about the ability of your own data to meet these criteria, use nonparametric alternatives, for which the assumptions are less stringent.

Independent observations

Observations must be independent even if you end up doing nonparametric analysis. Data points need to have been collected such that the magnitude of one observation does not affect another. Certain tests, such as the repeated-measures analysis of variance or Friedman's nonparametric analysis of variance, however, allow for multiple measurements on a single subject. Care should be taken to avoid interdependence of data during your experimental design because no amount of adjustment or transformation can rectify the problems inherent in nonindependent data. An example of nonindependent data would be taking three measurements of pH from a single cell culture

dish and entering them as $n = 3$ in your ANOVA. In this case, $n = 1$, and you are just as well off taking a single measurement as you are averaging the three readings.

Continuous variable

This assumption is not an extremely strong one as long as you bear several things in mind. The data need to be numerical and ordinal. Continuous data, that is, data that can take on any possible numerical value, are the kind of data for which ANOVA was designed to be used. Deviations from this will affect the sensitivity of the test. It is permissible for data to be discrete (having only a certain set of possible values, like counting numbers) if the numerical values have a fairly wide range. For ANOVA, this would mean something on the order of a 1 to 10 scale or greater. The numbers applied to observations must have some tie to a "greater than or less than" scale; they cannot simply be a classification system of your data. If you have discrete data, you should pay closer attention to the adherence of your data to the additional assumptions detailed below, or just choose a nonparametric test from the outset.

Homogeneity of variance

This is probably the strongest assumption of ANOVA, since, as we will see later, the whole test hinges on the variability within each of our experimental groups being similar. As in the case of the t test (see discussion of The t Test), the variances for each group need to be tested for similarity. The big difference here, however, is that for the greater-than-two-groups case, there is no way to continue with the test if the variances of the groups are significantly different. Instead, we need to transform the data mathematically to make the variances more similar, or move on to a nonparametric equivalent.

Checking for homogeneity of the variances should start with drawing a frequency histogram of each group and seeing how spread out the data are. It is also a good idea to calculate variances for each group and see if the numbers look close. In many cases, this is all the checking that needs be done. If you remain unsure, there is a simple test called the F_{\max} test, which is a simple matter of setting up the following ratio:

$$\frac{\text{largest group variance in the study}}{\text{smallest group variance in the study}} = F_{\max}$$

This is equivalent to using an F ratio of two groups in the t test (see discussion of The t Test),

Commonly Used Techniques

A.3I.13

but the tabled critical value to which it is compared is from a cumulative F distribution. Degrees of freedom are calculated as $n - 1$, where n is the number of observations in a single group. If n varies among groups, then the lowest value should be chosen, to be most conservative in the test. As in the case of the t test, if you fail to reject the null hypothesis, then the variances can be assumed to be roughly equivalent.

Normality

Although this is touted as an important assumption in most classical statistics, it is not as strong an assumption in practice for ANOVA as the assumption of homogeneous variances. Again, using a frequency histogram to give an idea of the shape of your distribution is often a good enough test of normality (see discussion of Determining Whether Data Are Normally Distributed with Equal Variances). If the data are distributed with a marked hump toward the average value and trailing off to either side, then ANOVA will probably work fine. Computer programs provide several different tests of normality that can also be used for this purpose.

Partitioning the Variance

It may seem odd that we wish to analyze variance when, in the long run, our question is to find the differences in means among groups in our experiment. Analysis of variance, however, takes advantage of the underlying assumption that the variation in a measurable quantity is similar within all groups in a population. One or more of these groups, however, may have a very different mean relative to the other groups. If this were the case, the total variance of the whole population would increase relative to the amount seen in each group. It is this observation that gives us the null hypothesis for ANOVA, which is that the means of all the groups are the same:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_n$$

In ANOVA, we break down the total variance into two basic components: within-group and among-group variance. Within-group or error variance is the naturally occurring variance among individuals due to inherent genetic differences and the differential effects of environment that each individual experiences. This contrasts with among-group or treatment variance, which is the variance observed among groups due to perturbations of the experiment. The treatment variance is the mathematical difference between the total variation observed in our experimental model and the natural vari-

ance measured across individuals (the error variance). Partitioning the variance into components allows us to establish the ratio of the two components of variance, defined as F_s , as shown in the following equation:

$$\frac{\text{among-group (treatment) variance}}{\text{within-group (error) variance}} = F_s$$

Testing the Null Hypothesis

After partitioning the variance into its two components (see discussion of Partitioning the Variance), we can then compare our calculated F_s to the F distribution, which is tabled in numerous statistics texts. The tables are usually organized with the critical values listed for many combinations of the numerator (treatment) versus denominator (error) degrees of freedom. Treatment degrees of freedom are the number of groups minus 1, and the error degrees of freedom are the total number of individual observations in all groups minus the number of groups. While it is unlikely that you will ever have to look up tabled F values, since computers do it for us, it is important to understand the convention used for F ratios. The F value given for a particular tested proportion is written as:

$$F_{(\text{treatment df, error df, } \alpha)} = \text{critical value from table.}$$

So, the critical value of F for an hypothetical experiment with a rejection level (α) of 0.05, four treatment groups, and a total experimental population of 48 organisms would be written as:

$$F_{(3,44,0.05)} = 2.82.$$

If α is chosen as 0.05, it is often omitted from the statistical reporting for brevity's sake. This critical value is important in our analysis because, if our calculated F ratio gives a value higher than the critical value from the table, we reject our null hypothesis. For ANOVA, the null hypothesis is that all the means for all the groups are equal. Rejecting the null hypothesis tells us only one thing: that at least one of the groups differs from the others. It does not tell us how many differ or from how many of the other groups any one might differ. For this answer we need to proceed to a method of multiple-means comparisons.

The ANOVA Table

Assuming a hypothetical experiment such as the one summarized in Table A.3I.7 and Table A.3I.8, the output from any ANOVA computer package is likely to produce an

Table A.3I.7 Hypothetical Optical Density Data for Bacterial Cultures

Control	Treatment A	Treatment B	Treatment C
0.53	0.37	0.52	0.7
0.57	0.44	0.55	0.66
0.6	0.44	0.59	0.72
0.49	0.39	0.59	0.69
0.42	0.62	0.7	0.55
0.4	0.58	0.72	0.53
0.41	0.6	0.73	0.5

Table A.3I.8 Descriptive Summary Table for Hypothetical Optical Density Data Shown in Table A.3I.7

Groups	Count	Sum	Mean	Variance
Control	7	3.42	0.488571	0.006581
Treatment A	7	3.44	0.491429	0.011081
Treatment B	7	4.4	0.628571	0.007448
Treatment C	7	4.35	0.621429	0.008381

Table A.3I.9 Single-Factor ANOVA Table for Hypothetical Optical Density Data Shown in Tables A.3I.7 and Table A.3I.8^a

Source of Variation	SS	df	MS	<i>F</i>	<i>p</i> value	<i>F</i> _c
Among groups	0.12778214	3	0.042594	5.09	0.01	3.01
Within groups	0.20094286	24	0.008373	—	—	—
Total	0.328725	27	—	—	—	—

^aAbbreviations: *F*_c, critical value of *F* statistic; SS, sum of squares; MS, mean square.

ANOVA table similar to Table A.3I.9, originally generated by Microsoft Excel. The data are hypothetical optical density readings for plasmid-transformed bacteria that have been transformed by three different experimental protocols, along with a control that has not been transformed.

The first part of the data analysis is a simple descriptive summary table (Table A.3I.8) that gives the sample size, mean, and variance for each group. Based on this table, it is fairly clear that the variances within groups are similar, but that the means for some of the groups seem to be different from others. The ANOVA table (Table A.3I.9) then gives the variance components and degrees of freedom accounted for among and within groups. The “SS” column indicates the sum of squares for each component, and the “MS” indicates mean square. Sum of squares is the summation of the square of the amount by which each observation differs from

the total mean. The mean square is the same as the variance, and is calculated by dividing the SS by the appropriate degrees of freedom. The mean square within-groups value is often called the “mean square error,” and figures prominently into tests of differences among the means.

The table also reports the calculated *F* ratio for the among versus within variance comparison. The computer then provides the exact *p* value associated with this ratio and sample size, and further indicates the critical *F* value that is necessary for rejection at the 0.05 level.

Which Means Are Different?

We have finally reached the point where we are ready to answer the question we asked, which was whether the average value of one or more of our groups differs from the others. While statisticians differ on this issue, it is the authors’ opinion that if you fail to reject the null

hypothesis of your ANOVA you should not test for significant means differences unless you planned certain comparisons *before* you ran the experiment. That having been said, it is possible that you will find significant means differences even if you fail to reject the ANOVA null hypothesis. If that is the case, however, it is worth asking how robust your initial test was—did you have large-enough sample sizes, were your variances “equal enough,” and were the distributions of your data groups normal or at least all similarly shaped? If not, you would probably be better off reworking your analysis with a nonparametric equivalent.

Assuming that you have rejected H_0 , how can you tell which means differ? Again, a first approximation is to look at the average values for each group and see which appear most different from the others. Next, you will want to proceed to a test that will provide you with a p value for those differences. For the special case of testing a control versus treatments, the Dunnett test is appropriate. Other test names you might see are Bonferroni or Tukey; computer packages have many options and you should refer to the statistical documentation of your package to determine which option best suits your data.

Since means comparisons are cumbersome to do by hand and are part of most statistical packages, we will not go into the specific calculations here. Make sure, however, that you understand how the computer is reporting the significance of your data to you. The computer will usually give you a p value. As you perform more mean-versus-mean comparisons, your acceptable α level will change, and the computer rarely keeps track of this, so it is important to know how to calculate these adjustments in α . One of these techniques is discussed below.

Planned or unplanned comparisons

The use of certain techniques in your final means comparison will depend, in part, on whether you planned before the experiment among which groups you were going to look for differences (a priori tests) or whether you waited to peruse your data first (a posteriori). This turns out to be a theoretically very complex subject and one that is not worth delving into here. A detailed treatment can be found in Sokal and Rohlf (1981).

It is worth pointing out here a common mistake in choosing significance levels for multiple means comparisons. In general, planned comparisons allow the researcher a greater rejection level (bigger α or p value) than un-

planned comparisons. The reasoning is simple: if you first look at your data, then choose your comparisons, you are introducing a bias to your analysis. To account for this, the level of rejection needs to be tightened. Even if you have planned your comparisons, however, it is easy to abuse the power of multiple means tests.

For example, declaring a priori that you are going to conduct all possible pairwise comparisons in your experiment is tantamount to using multiple t tests (see discussion of Why Not Lots of t Tests?). If you have k groups in your analysis, the number of all possible pairwise comparisons would be $k(k - 1)/2$. Since each comparison is not independent of all others, there is a need to alter the significance level of the tests. In order to use the experiment-wise error (i.e., $\alpha = p = 0.05$), the degrees of freedom for all the tests you perform should not exceed $k - 1$. So, if you have four groups (a, b, c, d), you might declare to test a versus b, c, d (1 df); a versus d (1 df); and b, c versus d (1 df). That would burn up 3 degrees of freedom at $\alpha = 0.05$. Note that in the very common case of comparing each treatment to a control, we stay within this limit as well (assume a is the control and compare it pairwise with each of the three treatments b, c, d).

If we were to go beyond this level of comparison, however, we would need to adjust the rejection level for *all* the tests we perform to account for a loss of independence. A simple and conservative technique that does this is the Bonferroni method (Sokal and Rohlf, 1987). To make an adjustment to the experiment-wise error term, α , calculate a test significance, α' , by dividing the experiment-wise error by the number of degrees of freedom used in your selected tests (df_T):

$$\alpha' = \frac{\alpha}{df_T}$$

Applying this technique to the example above, if we choose to compare the control a to each of the three treatment groups (3 df), and perform two additional internal treatment tests of b versus c, d (1 df) and c versus d (1 df), we have a total of 5 degrees of freedom, which exceeds our “allowable” number by 2. The Bonferroni method would then require that we adjust α' as follows:

$$\alpha' = \frac{0.05}{5} = 0.01$$

Therefore, for all the tests we perform, including all three of the control-versus-treatment tests, we have decreased our rejection

Table A.3I.10 Turbidity Data (see Table A.3I.7) Scored on a Scale of 1 (least turbid) to 10 (most turbid)

Control	Treatment A	Treatment B	Treatment C
5	1	4	8
5	3	4	9
5	2	6	9
5	3	5	8
6	2	6	8
5	5	7	9
5	1	6	10

Table A.3I.11 Turbidity Data (see Table A.3I.10) Ranked from 1 to N

	Control	Treatment A	Treatment B	Treatment C
	11.5	1.5	7.5	23
	11.5	5.5	7.5	26
	11.5	3.5	18.5	26
	11.5	5.5	11.5	23
	18.5	3.5	18.5	23
	11.5	11.5	21	26
	11.5	1.5	18.5	28
Rank sums	87.5	32.5	103	175
Rank averages	12.50	4.64	14.71	25.00

level to $p = 0.01$, a level that could make a big difference in the significance of our results. This should point out the importance of good experimental design and of choosing only those tests that are important to the experimental goals.

The Nonparametric ANOVA Equivalent

If the data you end up with do not meet the assumptions of ANOVA, you can still perform multiple comparisons with a nonparametric test. For a one-way ANOVA, the nonparametric equivalent most popularly used is the Kruskal-Wallis one-way analysis of variance. The assumptions of the Kruskal-Wallis test are less stringent than those of the parametric ANOVA. The samples are drawn from a population with a continuous distribution and the sampling must be random. The samples must be independent. Their values must be rankable, that is, on at least an ordinal scale. Many computer programs provide the option of calculating the Kruskal-Wallis statistic. It is simple to do by hand, and is worth outlining here.

In the hypothetical example of the optical densities for transformed bacteria strains (see

Table A.3I.7, Table A.3I.8, and Table A.3I.9), assume that the researcher, instead of using a spectrophotometer, simply used an ordinal scale of 1 to 10 to approximate the turbidity of the sample (see Table A.3I.10). (Okay, so it's not very precise science, but, who knows, maybe the spec was broken that day and the samples needed to be measured!) In this case, there is an expectation that the scores would not be distributed in such a way as to meet the criteria for a parametric ANOVA.

To proceed through the analysis, it is then necessary to find the average value for the ranks of the data in each of the groups. To do this, we will give the data across all groups a rank score. The smallest number will have the lowest rank, and so on until we have ranked all the data. Equivalent values are termed "ties" and are given the average rank of the tied observations (see Table A.3I.11).

From these ranks, the test statistic (KW) is calculated according to the following formula:

$$KW = \left[\frac{12}{N(N-1)} \sum_{j=1}^k n_j \bar{R}_j^2 \right] - 3(N+1)$$

In the above equation, N = total population size, n_j = number of observations in a group (note that groups need not be equal in size), k = number of groups, and R_j = rank average for each group.

For small samples (usually $k \leq 3$ with $n_j \leq 5$) there are special tabled values for the KW statistic. For larger samples, the resulting value is compared to a χ^2 distribution with $(k - 1)$ degrees of freedom. If the KW value is larger than the tabled value at a chosen α , then H_0 is rejected. As in the case of the parametric ANOVA, a significant result implies that at least one of the groups is different from the rest. A means comparison test needs to be done to show which group it is. For the data in Table A.3I.11, the calculation is as follows:

$$\begin{aligned} \text{KW} &= \left[\frac{12}{(28)(29)} \right] \left[7(12.5)^2 + 7(4.64)^2 \right. \\ &\quad \left. + 7(14.71)^2 + 7(25)^2 \right] - [3(28 + 1)] \\ &= 18.45 \end{aligned}$$

The critical value for $\alpha = 0.05$, with $df = (k - 1) = 3$, is 7.82. Since our KW test statistic is larger than that, our test is significant and we reject H_0 . A problem that we might have encountered with our data, however, is that we ended up with many tied ranks. There are mathematical corrections for ties that will yield a slightly higher KW statistic than the calculation without the correction. This means that if you reject H_0 without the correction, you will also reject with the correction, so it is unnecessary if your results are significant. If, on the other hand, if your significance is marginal and you are concerned about the ties in your data, you may choose to perform a ties correction. Siegel and Castellan (1988) gives one such correction. For our example, the ties correction would yield $\text{KW} = 19.0$.

Nonparametric Means Comparison

Having demonstrated a significant difference somewhere among the groups with the Kruskal-Wallis test (see discussion of The Nonparametric ANOVA Equivalent), it is then appropriate to proceed to find out where those differences are. The technique that we will apply will allow us to perform only pairwise comparisons, but it is adjusted so that all possible pairwise comparisons can be made with the same level of rejection.

Treatment versus treatment

Multiple comparisons among all treatments, as in the case of ANOVA, require attention to the error inherent in the entire experiment relative to each pairwise test we choose to do. To account for this, we adjust our test statistic by the total possible pairwise comparisons in the experiment. Whether we plan to do only some of them or all of them is irrelevant to the adjustment; however, there is a special case, outlined below, if all we wish to do is compare a single control to several treatments.

The many possible differences between pairs of observations in a data set will become approximately normal as the sample size increases, regardless of the underlying distribution of the data. This allows us to use the normal distribution probabilities to test for significant differences between groups. In effect, we can use the normal distribution to establish a confidence interval for the difference between two groups. If they fall beyond this value, we can say that their rank average values differ significantly.

To be able to use the normal approximation, we first need to adjust the value we look up in a normal probability table (z) by the number of pairwise group comparisons possible in the experiment:

$$z_{\alpha'} = z_{\alpha/k(k-1)}$$

In this equation, k is the number of groups in the experiment. For the data in Table A.3I.11, if we chose an experiment-wise $\alpha = 0.05$, this would yield $z_{\alpha'} = z_{0.05/(4)(3)} = z_{0.004}$. The numerical value of $z_{\alpha'}$ can be looked up on a normal table or derived from a computer program. In our case, $z_{0.004} = 2.65$.

Now that we have determined the correct normal approximation for our test, we apply it to determine the critical value for rejection at $\alpha = 0.05$ (not $\alpha' = 0.004$) between any two groups in the population. If the mathematical difference between the two rank averages of the groups is greater than the critical value, then we reject at $p < 0.05$. The calculation of the critical value (CV) is dependent on sample sizes:

$$\text{CV} = z_{\alpha'} \sqrt{\frac{N(N+1)}{12} \left(\frac{1}{n_u} + \frac{1}{n_v} \right)}$$

In the above equation, N is the total population size and n_u and n_v are the sample sizes of the two groups that you are comparing (they may be different sizes).

Table A.3I.12 Pairwise Comparisons of Hypothetical Turbidity Ranked Data in Table A.3I.11

Comparison	Calculation	Significance ^a
Control vs. A	$ 12.5 - 4.64 = 7.86$	NS
Control vs. B	$ 12.5 - 14.71 = 2.21$	NS
Control vs. C	$ 12.5 - 25 = 12.5$	*
A vs. B	$ 4.64 - 14.71 = 10.07$	NS
A vs. C	$ 4.64 - 25 = 20.36$	*
B vs. C	$ 14.71 - 25 = 10.29$	NS

^aNS, not significant; asterisk refers to differences significant at $p < 0.05$.

For our example (see Table A.3I.11), since all group sizes are the same, the CV will be identical for any and all of the six possible pairwise comparisons. If sample sizes are different, you need to calculate a new CV for each comparison:

$$\begin{aligned} CV &= 2.65 \sqrt{\frac{28(29)}{12} \left(\frac{1}{7} + \frac{1}{7} \right)} \\ &= 11.65 \end{aligned}$$

To determine which differences are significant, we first calculate the absolute value of the difference between all pairs of rank averages (Table A.3I.12). If the difference exceeds the critical value, then the difference can be determined to be significant at $p < 0.05$.

From this analysis it is clear that while the entire model was quite significant the differences in the data set are due mostly to the large value of Treatment C.

Control versus treatment

The test above provides a powerful nonparametric tool for comparing multiple samples. A researcher often is interested, however, in a less broad model—one that looks only at the differences between a control and two or more treatments. Following the same procedure as above, it is possible to approximate these differences with a normal distribution. A different correction is needed, however, to account for fewer, more defined comparisons. In this case the following equation applies:

$$z_{\alpha'} = z_{\alpha} / \sqrt{2(k-1)}$$

For our example, this equation would yield $z_{\alpha'} = z_{0.005/(2)(3)} = z_{0.008} = 2.41$. Because fewer comparisons are being done, this smaller z

value will lead to a smaller critical value and provide greater probability of finding a difference between the control and a treatment. Furthermore, if we expected a priori that our treatments would yield a directional result (either smaller or greater, but not both) relative to the control, we could reduce this number even more by adjusting by $\alpha/(k-1)$ rather than $\alpha/2(k-1)$.

Since we did not have this expectation in our data, we will use the two-tailed option and calculate as follows:

$$CV = z_{\alpha'} \sqrt{\frac{N(N+1)}{12} \left(\frac{1}{n_c} + \frac{1}{n_u} \right)}$$

In the equation above, N is the total population size, n_c is the sample size of the control, and n_u is the sample size of the treatment being compared.

Again, because all of our treatments have the same number of observations, we need calculate only one value for CV.

$$\begin{aligned} CV &= 2.41 \sqrt{\frac{28(29)}{12} \left(\frac{1}{7} + \frac{1}{7} \right)} \\ &= 10.60 \end{aligned}$$

This value is then compared as above to the differences between the control and treatment rank averages. We would reject any difference greater than 10.60. Despite the tighter rejection interval (compared with 11.65 from the all comparisons model) the only treatment that differs significantly from control at $p < 0.05$ is Treatment C. Note that the previously significant comparison (A versus C) is not valid in this model.

Table A.3I.13 Hypothetical Immunoprecipitation Data for Repeated-Measures ANOVA

Mouse ID	IgG precipitate (mg/ml)				
	Day 1	Day 2	Day 3	Day 4	Day 5
A	0	0	0.02	0.7	1.3
B	0.7	2.5	3.2	4.5	3
C	0.09	1.8	4	7.1	2.2
D	0.01	0.01	0.6	1.8	0.1
E	0.6	5.5	6.8	9.6	8.5
F	0.2	1.1	2.4	2.5	1.5
G	0	0.05	0.9	3	2.7
H	0.2	0.5	0.7	2.5	6

Table A.3I.14 Repeated-Measures ANOVA Table for Hypothetical Immunoprecipitation Data in Table A.3I.13^a

Source of variation	SS	df	MS	<i>F</i>	<i>p</i> value	<i>F</i> _c
Rows	121.8885	7	17.41265	8.745258	1.12E-05	2.359258
Columns	68.29064	4	17.07266	8.574505	0.00012	2.714074
Error	55.75068	28	1.991096	—	—	—
Total	245.9298	39	—	—	—	—

^aAbbreviations: *F*_c, critical value of *F* statistic; SS, sum of squares; MS, mean square.

Repeated Measures—A Special Case

Sometimes we wish to collect more than one sample from an individual organism in our study. Given that we would expect there to be less variability in samples taken from one individual than in samples taken among several individuals, it would seem that we need to adjust our analysis of variance somewhat. And indeed we do. Instead of a simple one-way analysis of variance, we perform a repeated-measures analysis of variance. This technique allows us to statistically control for the variance attributable to the inherent differences in the test subjects relative to the variance in response to the treatments. Once you determine that a difference exists somewhere in the data set, you can test for means differences in the same way as described above for one-way ANOVA.

Parametric

The same assumptions hold true for the repeated measures ANOVA as for the one-way ANOVA. For our parametric ANOVA example, let us consider the following case (see Table A.3I.13). You are producing an antibody in a strain of mice and want to know when antibody

production reaches a peak. To collect data, you perform immunoprecipitation on serum collected for 5 days from eight study animals.

As you can see, certain animals have higher values overall than others, some respond with IgG production very quickly, others lag a bit and so on. If we were to do a one-way ANOVA on these data, the variability among the animals (rather than across days) might compromise the significance of our test. To avoid this problem, the repeated-measures ANOVA calculates variances for the rows and columns separately. The final output for such an analysis would resemble Table A.3I.14, originally produced by Microsoft Excel.

In this table there are significance values reported for the rows (mice) and columns (days). For our example, both factors are clearly significant. This implies that there is a difference in the precipitation of IgG over the 5 days of measurement and a difference in IgG expression in different mice.

Nonparametric

If you have repeated measures that violate the assumptions of the parametric ANOVA,

Table A.3I.15 Hypothetical Immunoprecipitation Data in Table A.3I.13 Ranked for Friedman's Analysis of Variance

Mouse ID	IgG precipitate (ranks)				
	Day 1	Day 2	Day 3	Day 4	Day 5
A	1.5	1.5	3	4	5
B	1	2	4	5	3
C	1	2	4	5	3
D	1.5	1.5	4	5	3
E	1	2	3	5	4
F	1	2	4	5	3
G	1	2	3	5	4
H	1	2	3	4	5
Rank sums	9	15	28	38	30

then there is a nonparametric equivalent for the repeated-measures ANOVA as well. As in the turbidity example above, a common case would be when a response was scored on an ordinal scale rather than measured on a continuous one.

The nonparametric equivalent in this case is the Friedman's analysis of variance. The data are handled in a fashion similar to the Kruskal-Wallis test (see discussion of The Nonparametric ANOVA Equivalent), except that ranks are assigned only within an individual rather than across the entire data set. In this way, an individual that overall got higher responses than the others would not contribute unduly to the overall variance of the population. To illustrate, let us simply use the immunoprecipitation data in Table A.3I.13. We begin by ranking the values within each individual from 1 to 5 (see Table A.3I.15).

We then sum the ranks for each column. Under the null hypothesis, the ranks should be distributed in random order for any individual, which would result in the column rank sums being approximately equal. Clearly the rank sums are not equal, but are the differences significant? We can determine this by using the Friedman test and comparing the resulting value to a χ^2 distribution with $k - 1$ degrees of freedom for our selected α level of rejection. The calculation of the statistic (Fr) is similar to the calculation of the Kruskal-Wallis statistic:

$$Fr = \left[\frac{12}{Nk(k+1)} \sum_{j=1}^k R_j^2 \right] - 3N(k+1)$$

In the above equation, N is the number of rows (subjects), k is the number of columns

(treatments), and R_j is equal to the rank sum of each column.

For our example, this would yield the following:

$$Fr = \left[\frac{12}{(8)(5)(6)} (81 + 225 + 784 + 1444 + 900) \right] - (3)(8)(6)$$

$$Fr = 27.7$$

The χ^2 critical value for $\alpha = 0.05$ with 4 degrees of freedom is 9.48. Since our Fr statistic is greater than that, we reject the null hypothesis and conclude that there are differences in IgG production over the course of 5 days (columns), irrespective of differences inherent in the test subjects (rows).

CONCLUSION

While the methods for dealing with comparisons of data outlined above are just a small sampling of those available, we hope they add to the arsenal of tools available to biologists. By performing simple descriptive analyses of your data and following the guidelines for test selection outlined in this Appendix, you should be able to avoid many of the statistical pitfalls frequently encountered in groups comparisons.

LITERATURE CITED

- Siegel, S. and Castellan, N.J., Jr. 1988. Nonparametric Statistics for the Behavioral Sciences, 2nd ed. McGraw-Hill, New York.
- Sokal, R.R. and Rohlf, F.J. 1981. Biometry, 2nd ed. W.H. Freeman, New York.
- Sokal, R.R. and Rohlf, F.J. 1987. Introduction to Biostatistics, 2nd ed. W.H. Freeman, New York.

Zar, J.H. 1984. Biostatistical Analysis, 2nd ed. Prentice-Hall, Englewood Cliffs, New Jersey.

KEY REFERENCES

Hampton, R.E. 1994. Introductory Biological Statistics. William. C. Brown Publishers, Dubuque, Iowa.

Provides excellent outlines for many statistical tests and uses relevant, comprehensible biological examples.

Motulsky, H. 1995. Intuitive Biostatistics. Oxford University Press, New York.

A nice, nonmathematical introduction to biostatistics, covering standard topics as well as many areas important to biology that are often omitted from introductory texts.

Contributed by Elizabeth F. Ryder and
Phil Robakiewicz
Worcester Polytechnic Institute
Worcester, Massachusetts

SELECTED SUPPLIERS OF REAGENTS AND EQUIPMENT

Listed below are addresses and phone numbers of commercial suppliers who have been recommended for particular items used in our manuals because: (1) the particular brand has actually been found to be of superior quality, or (2) the item is difficult to find in the marketplace. Consequently, this compilation may not include some important vendors of biological supplies. For comprehensive listings, see *Linscott's Directory of Immunological and Biological Reagents* (Santa Rosa, CA), *The Biotechnology Directory* (Stockton Press, New York), the annual Buyers' Guide supplement to the journal *Bio/Technology*, as well as various sites on the Internet.

A.C. Daniels

72-80 Akeman Street
Tring, Hertfordshire, HP23 6AJ, UK
(44) 1442 826881
FAX: (44) 1442 826880

A.D. Instruments

5111 Nations Crossing Road #8
Suite 2
Charlotte, NC 28217
(704) 522-8415 FAX: (704) 527-5005
<http://www.us.endress.com>

A.J. Buck

11407 Cronhill Drive
Owings Mill, MD 21117
(800) 638-8673 FAX: (410) 581-1809
(410) 581-1800
<http://www.ajbuck.com>

A.M. Systems

131 Business Park Loop
P.O. Box 850
Carlsborg, WA 98324
(800) 426-1306 FAX: (360) 683-3525
(360) 683-8300
<http://www.a-msystems.com>

Aaron Medical Industries

7100 30th Avenue North
St. Petersburg, FL 33710
(727) 384-2323 FAX: (727) 347-9144
<http://www.aaronmed.com>

Abbott Laboratories

100 Abbott Park Road
Abbott Park, IL 60064
(800) 323-9100 FAX: (847) 938-7424
<http://www.abbott.com>

ABCO Dealers

55 Church Street Central Plaza
Lowell, MA 01852
(800) 462-3326 (978) 459-6101
<http://www.lomedco.com/abco.htm>

Aber Instruments

5 Science Park
Aberystwyth, Wales SY23 3AH, UK
(44) 1970 636300
FAX: (44) 1970 615455
<http://www.aber-instruments.co.uk>

ABI Biotechnologies

See Perkin-Elmer

ABI Biotechnology

See Apotex

Access Technologies

Subsidiary of Norfolk Medical
7350 N. Ridgeway
Skokie, IL 60076
(877) 674-7131 FAX: (847) 674-7066
(847) 674-7131
<http://www.norfolkaccess.com>

Accurate Chemical and Scientific

300 Shames Drive
Westbury, NY 11590
(800) 645-6264 FAX: (516) 997-4948
(516) 333-2221
<http://www.accuratechemical.com>

AccuScan Instruments

5090 Trabue Road
Columbus, OH 43228
(800) 822-1344 FAX: (614) 878-3560
(614) 878-6644
<http://www.accuscan-usa.com>

AccuStandard

125 Market Street
New Haven, CT 06513
(800) 442-5290 FAX: (877) 786-5287
<http://www.accustandard.com>

Ace Glass

1430 NW Boulevard
Vineland, NJ 08360
(800) 223-4524 FAX: (800) 543-6752
(609) 692-3333

ACO Pacific

2604 Read Avenue
Belmont, CA 94002
(650) 595-8588 FAX: (650) 591-2891
<http://www.acopacific.com>

Acros Organic

See Fisher Scientific

Action Scientific

P.O. Box 1369
Carolina Beach, NC 28428
(910) 458-0401 FAX: (910) 458-0407

AD Instruments

1949 Landings Drive
Mountain View, CA 94043
(888) 965-6040 FAX: (650) 965-9293
(650) 965-9292
<http://www.adinstruments.com>

Adaptive Biosystems

15 Ribocon Way
Progress Park
Luton, Bedfordshire LU4 9UR, UK
(44)1 582-597676
FAX: (44)1 582-581495
<http://www.adaptive.co.uk>

Adobe Systems

1585 Charleston Road
P.O. Box 7900
Mountain View, CA 94039
(800) 833-6687 FAX: (415) 961-3769
(415) 961-4400
<http://www.adobe.com>

Advanced Bioscience Resources

1516 Oak Street, Suite 303
Alameda, CA 94501
(510) 865-5872 FAX: (510) 865-4090

Advanced Biotechnologies

9108 Guilford Road
Columbia, MD 21046
(800) 426-0764 FAX: (301) 497-9773
(301) 470-3220
<http://www.abionline.com>

Advanced ChemTech

5609 Fern Valley Road
Louisville, KY 40228
(502) 969-0000
<http://www.peptide.com>

Advanced Machining and Tooling

9850 Businesspark Avenue
San Diego, CA 92131
(858) 530-0751 FAX: (858) 530-0611
<http://www.amtmfg.com>

Advanced Magnetics

See PerSeptive Biosystems

Advanced Process Supply

See Naz-Dar-KC Chicago

Advanced Separation Technologies

37 Leslie Court
P.O. Box 297
Whippany, NJ 07981
(973) 428-9080 FAX: (973) 428-0152
<http://www.astecusa.com>

Advanced Targeting Systems

11175-A Flintkote Avenue
San Diego, CA 92121
(877) 889-2288 FAX: (858) 642-1989
(858) 642-1988
<http://www.ATSBio.com>

Advent Research Materials

Eynsham, Oxford OX29 4JA, UK
(44) 1865-884440
FAX: (44) 1865-84460
<http://www.advent-rm.com>

Advet

Industrivagen 24
S-972 54 Lulea, Sweden
(46) 0920-211887
FAX: (46) 0920-13773

Aesculap

1000 Gateway Boulevard
South San Francisco, CA 94080
(800) 282-9000
<http://www.aesculap.com>

Affinity Chromatography

307 Huntingdon Road
Girton, Cambridge CB3 0JX, UK
(44) 1223 277192
FAX: (44) 1223 277502
<http://www.affinity-chrom.com>

Affinity Sensors

See Labsystems Affinity Sensors

Affymetrix

3380 Central Expressway
Santa Clara, CA 95051
(408) 731-5000 FAX: (408) 481-0422
(800) 362-2447
<http://www.affymetrix.com>

Agar Scientific

66a Cambridge Road
Stansted CM24 8DA, UK
(44) 1279-813-519
FAX: (44) 1279-815-106
<http://www.agarscientific.com>

A/G Technology

101 Hampton Avenue
Needham, MA 02494
(800) AGT-2535 FAX: (781) 449-5786
(781) 449-5774
<http://www.agtech.com>

Suppliers

A.4.1

Agen Biomedical Limited
11 Durbell Street
P.O. Box 391
Acacia Ridge 4110
Brisbane, Australia
61-7-3370-6300 FAX: 61-7-3370-6370
<http://www.agen.com>

Agilent Technologies
395 Page Mill Road
P.O. Box 10395
Palo Alto, CA 94306
(650) 752-5000
<http://www.agilent.com/chem>

Agouron Pharmaceuticals
10350 N. Torrey Pines Road
La Jolla, CA 92037
(858) 622-3000 FAX: (858) 622-3298
<http://www.agouron.com>

Agracetus
8520 University Green
Middleton, WI 53562
(608) 836-7300 FAX: (608) 836-9710
<http://www.monsanto.com>

AIDS Research and Reference
Reagent Program
U.S. Department of Health and
Human Services
625 Lofstrand Lane
Rockville, MD 20850
(301) 340-0245 FAX: (301) 340-9245
<http://www.aidsreagent.org>

AIN Plastics
249 East Sanford Boulevard
P.O. Box 151
Mt. Vernon, NY 10550
(914) 668-6800 FAX: (914) 668-8820
<http://www.tincna.com>

Air Products and Chemicals
7201 Hamilton Boulevard
Allentown, PA 18195
(800) 345-3148 FAX: (610) 481-4381
(610) 481-6799
<http://www.airproducts.com>

ALA Scientific Instruments
1100 Shames Drive
Westbury, NY 11590
(516) 997-5780 FAX: (516) 997-0528
<http://www.alascience.com>

Aladin Enterprises
1255 23rd Avenue
San Francisco, CA 94122
(415) 468-0433 FAX: (415) 468-5607

Aladdin Systems
165 Westridge Drive
Watsonville, CA 95076
(831) 761-6200 FAX: (831) 761-6206
<http://www.aladdinsys.com>

Alcide
8561 154th Avenue NE
Redmond, WA 98052
(800) 543-2133 FAX: (425) 861-0173
(425) 882-2555
<http://www.alcide.com>

Aldevron
3233 15th Street, South
Fargo, ND 58104
(877) Pure-DNA FAX: (701) 280-1642
(701) 297-9256
<http://www.aldevron.com>

Aldrich Chemical
P.O. Box 2060
Milwaukee, WI 53201
(800) 558-9160 FAX: (800) 962-9591
(414) 273-3850 FAX: (414) 273-4979
<http://www.aldrich.sial.com>

Alexis Biochemicals
6181 Cornerstone Court East, Suite 103
San Diego, CA 92121
(800) 900-0065 FAX: (858) 658-9224
(858) 658-0065
<http://www.alexis-corp.com>

Alfa Aesar
30 Bond Street
Ward Hill, MA 01835
(800) 343-0660 FAX: (800) 322-4757
(978) 521-6300 FAX: (978) 521-6350
<http://www.alfa.com>

Alfa Laval
Avenue de Ble 5 - Bazellaan 5
BE-1140 Brussels, Belgium
32(2) 728 3811
FAX: 32(2) 728 3917 or 32(2) 728 3985
<http://www.alfalaval.com>

Alice King Chatham Medical Arts
11915-17 Inglewood Avenue
Hawthorne, CA 90250
(310) 970-1834 FAX: (310) 970-0121
(310) 970-1063

Allegiance Healthcare
800-964-5227
<http://www.allegiance.net>

Allelix Biopharmaceuticals
6850 Gorway Drive
Mississauga, Ontario
L4V 1V7 Canada
(905) 677-0831 FAX: (905) 677-9595
<http://www.allelix.com>

Allentown Caging Equipment
Route 526, P.O. Box 698
Allentown, NJ 08501
(800) 762-CAGE FAX: (609) 259-0449
(609) 259-7951
<http://www.acecaging.com>

Alltech Associates
Applied Science Labs
2051 Waukegan Road
P.O. Box 23
Deerfield, IL 60015
(800) 255-8324 FAX: (847) 948-1078
(847) 948-8600
<http://www.alltechweb.com>

Alomone Labs
HaMarpeh 5
P.O. Box 4287
Jerusalem 91042, Israel
972-2-587-2202 FAX: 972-2-587-1101
US: (800) 791-3904
FAX: (800) 791-3912
<http://www.alomone.com>

Alpha Innotech
14743 Catalina Street
San Leandro, CA 94577
(800) 795-5556 FAX: (510) 483-3227
(510) 483-9620
<http://www.alphainnotech.com>

Altec Plastics
116 B Street
Boston, MA 02127
(800) 477-8196 FAX: (617) 269-8484
(617) 269-1400

Alza
1900 Charleston Road
P.O. Box 7210
Mountain View, CA 94043
(800) 692-2990 FAX: (650) 564-7070
(650) 564-5000
<http://www.alza.com>

Alzet
c/o Durect Corporation
P.O. Box 530
10240 Bubba Road
Cupertino, CA 95015
(800) 692-2990 (408) 367-4036
FAX: (408) 865-1406
<http://www.alzet.com>

Amac
160B Larrabee Road
Westbrook, ME 04092
(800) 458-5060 FAX: (207) 854-0116
(207) 854-0426

Amaresco
30175 Solon Industrial Parkway
Solon, Ohio 44139
(800) 366-1313 FAX: (440) 349-1182
(440) 349-1313

Ambion
2130 Woodward Street, Suite 200
Austin, TX 78744
(800) 888-8804 FAX: (512) 651-0190
(512) 651-0200
<http://www.ambion.com>

American Association of
Blood Banks
College of American Pathologists
325 Waukegan Road
Northfield, IL 60093
(800) 323-4040 FAX: (847) 8166
(847) 832-7000
<http://www.cap.org>

American Bio-Technologies
See Intracel Corporation

American Bioanalytical
15 Erie Drive
Natick, MA 01760
(800) 443-0600 FAX: (508) 655-2754
(508) 655-4336
<http://www.americanbio.com>

American Cyanamid
P.O. Box 400
Princeton, NJ 08543
(609) 799-0400 FAX: (609) 275-3502
<http://www.cyanamid.com>

American HistoLabs
7605-F Airpark Road
Gaithersburg, MD 20879
(301) 330-1200 FAX: (301) 330-6059

American International Chemical
17 Strathmore Road
Natick, MA 01760
(800) 238-0001 (508) 655-5805
<http://www.aicma.com>

American Laboratory Supply
See American Bioanalytical

American Medical Systems
10700 Bren Road West
Minnetonka, MN 55343
(800) 328-3881 FAX: (612) 930-6654
(612) 933-4666
<http://www.visitams.com>

American Qualex
920-A Calle Negocio
San Clemente, CA 92673
(949) 492-8298 FAX: (949) 492-6790
<http://www.americanqualex.com>

American Radiolabeled Chemicals
11624 Bowling Green
St. Louis, MO 63146
(800) 331-6661 FAX: (800) 999-9925
(314) 991-4545 FAX: (314) 991-4692
<http://www.arc-inc.com>

American Scientific Products
See VWR Scientific Products

American Society for
Histocompatibility and
Immunogenetics
P.O. Box 15804
Lenexa, KS 66285
(913) 541-0009 FAX: (913) 541-0156
http://www.swmed.edu/home_pages/ASHI/ashi.htm

American Type Culture Collection
(ATCC)
10801 University Boulevard
Manassas, VA 20110
(800) 638-6597 FAX: (703) 365-2750
(703) 365-2700
<http://www.atcc.org>

Amersham
See Amersham Pharmacia Biotech

Suppliers

A.4.2

Amersham International
Amersham Place
Little Chalfont, Buckinghamshire
HP7 9NA, UK
(44) 1494-544100
FAX: (44) 1494-544350
<http://www.apbiotech.com>

Amersham Medi-Physics
Also see Nycomed Amersham
3350 North Ridge Avenue
Arlington Heights, IL 60004
(800) 292-8514 FAX: (800) 807-2382
<http://www.nycomed-amersham.com>

Amersham Pharmacia Biotech
800 Centennial Avenue
P.O. Box 1327
Piscataway, NJ 08855
(800) 526-3593 FAX: (877) 295-8102
(732) 457-8000
<http://www.apbiotech.com>

Amgen
1 Amgen Center Drive
Thousand Oaks, CA 91320
(800) 926-4369 FAX: (805) 498-9377
(805) 447-5725
<http://www.amgen.com>

Amicon
Scientific Systems Division
72 Cherry Hill Drive
Beverly, MA 01915
(800) 426-4266 FAX: (978) 777-6204
(978) 777-3622
<http://www.amicon.com>

Amika
8980F Route 108
Oakland Center
Columbia, MD 21045
(800) 547-6766 FAX: (410) 997-7104
(410) 997-0100
<http://www.amika.com>

Amoco Performance Products
See BPAmoco

AMPI
See Pacer Scientific

Amrad
576 Swan Street
Richmond, Victoria 3121, Australia
613-9208-4000
FAX: 613-9208-4350
<http://www.amrad.com.au>

Amresco
30175 Solon Industrial Parkway
Solon, OH 44139
(800) 829-2805 FAX: (440) 349-1182
(440) 349-1199

Anachemia Chemicals
3 Lincoln Boulevard
Rouses Point, NY 12979
(800) 323-1414 FAX: (518) 462-1952
(518) 462-1066
<http://www.anachemia.com>

Ana-Gen Technologies
4015 Fabian Way
Palo Alto, CA 94303
(800) 654-4671 FAX: (650) 494-3893
(650) 494-3894
<http://www.ana-gen.com>

Analox Instruments USA
P.O. Box 208
Lunenburg, MA 01462
(978) 582-9368 FAX: (978) 582-9588
<http://www.analox.com>

Analytical Biological Services
Cornell Business Park 701-4
Wilmington, DE 19801
(800) 391-2391 FAX: (302) 654-8046
(302) 654-4492
<http://www.ABSbioreagents.com>

Analytical Genetics Testing Center
7808 Cherry Creek S. Drive, Suite 201
Denver, CO 80231
(800) 204-4721 FAX: (303) 750-2171
(303) 750-2023
<http://www.geneticid.com>

AnaSpec
2149 O'Toole Avenue, Suite F
San Jose, CA 95131
(800) 452-5530 FAX: (408) 452-5059
(408) 452-5055
<http://www.anaspec.com>

Ancare
2647 Grand Avenue
P.O. Box 814
Bellmore, NY 11710
(800) 645-6379 FAX: (516) 781-4937
(516) 781-0755
<http://www.ancare.com>

Ancell
243 Third Street North
P.O. Box 87
Bayport, MN 55033
(800) 374-9523 FAX: (651) 439-1940
(651) 439-0835
<http://www.ancell.com>

Anderson Instruments
500 Technology Court
Smyrna, GA 30082
(800) 241-6898 FAX: (770) 319-5306
(770) 319-9999
<http://www.graseby.com>

Andreas Hettich
Gartenstrasse 100
Postfach 260
D-78732 Tuttlingen, Germany
(49) 7461 705 0
FAX: (49) 7461 705-122
<http://www.hettich-centrifugen.de>

Anesthetic Vaporizer Services
10185 Main Street
Clarence, NY 14031
(719) 759-8490
<http://www.avapor.com>

Animal Identification and Marking Systems (AIMS)
13 Winchester Avenue
Budd Lake, NJ 07828
(908) 684-9105 FAX: (908) 684-9106
<http://www.animalid.com>

Annovis
34 Mount Pleasant Drive
Aston, PA 19014
(800) EASY-DNA FAX: (610) 361-8255
(610) 361-9224
<http://www.annovis.com>

Apotex
150 Signet Drive
Weston, Ontario
M9L 1T9 Canada
(416) 749-9300 FAX: (416) 749-2646
<http://www.apotex.com>

Apple Scientific
11711 Chillicothe Road, Unit 2
P.O. Box 778
Chesterland, OH 44026
(440) 729-3056 FAX: (440) 729-0928
<http://www.applesci.com>

Applied Biosystems
See PE Biosystems

Applied Imaging
2380 Walsh Avenue, Bldg. B
Santa Clara, CA 95051
(800) 634-3622 FAX: (408) 562-0264
(408) 562-0250
<http://www.aicorp.com>

Applied Photophysics
203-205 Kingston Road
Leatherhead, Surrey, KT22 7PB
UK
(44) 1372-386537

Applied Precision
1040 12th Avenue Northwest
Issaquah, Washington 98027
(425) 557-1000
FAX: (425) 557-1055
<http://www.api.com/index.html>

Appligene Oncor
Parc d'Innovation
Rue Geiler de Kaysersberg, BP 72
67402 Illkirch Cedex, France
(33) 88 67 22 67
FAX: (33) 88 67 19 45
<http://www.oncor.com/prod-app.htm>

Applikon
1165 Chess Drive, Suite G
Foster City, CA 94404
(650) 578-1396 FAX: (650) 578-8836
<http://www.applikon.com>

Appropriate Technical Resources
9157 Whiskey Bottom Road
Laurel, MD 20723
(800) 827-5931 FAX: (410) 792-2837
<http://www.atrbiotech.com>

APV Gaulin
100 S. CP Avenue
Lake Mills, WI 53551
(888) 278-4321 FAX: (888) 278-5329
<http://www.apv.com>

Aqualon
See Hercules Aqualon

Aquarium Systems
8141 Tyler Boulevard
Mentor, OH 44060
(800) 822-1100 FAX: (440) 255-8994
(440) 255-1997
<http://www.aquariumsystems.com>

Aquebogue Machine and Repair Shop
Box 2055
Main Road
Aquebogue, NY 11931
(631) 722-3635 FAX: (631) 722-3106

Archer Daniels Midland
4666 Faries Parkway
Decatur, IL 62525
(217) 424-5200
<http://www.admworld.com>

Archimica Florida
P.O. Box 1466
Gainesville, FL 32602
(800) 331-6313 FAX: (352) 371-6246
(352) 376-8246
<http://www.archimica.com>

Arcor Electronics
1845 Oak Street #15
Northfield, IL 60093
(847) 501-4848

Arcturus Engineering
400 Logue Avenue
Mountain View, CA 94043
(888) 446 7911 FAX: (650) 962 3039
(650) 962 3020
<http://www.arctur.com>

Ardais Corporation
One Ledgemont Center
128 Spring Street
Lexington, MA 02421
(781) 274-6420 (781) 274-6421
<http://www.ardais.com>

Argonaut Technologies
887 Industrial Road, Suite G
San Carlos, CA 94070
(650) 998-1350 FAX: (650) 598-1359
<http://www.argotech.com>

Ariad Pharmaceuticals
26 Landsdowne Street
Cambridge, MA 02139
(617) 494-0400 FAX: (617) 494-8144
<http://www.ariad.com>

Armour Pharmaceuticals
See Rhone-Poulenc Rorer

Suppliers

A.4.3

Aronex Pharmaceuticals
8707 Technology Forest Place
The Woodlands, TX 77381
(281) 367-1666 FAX: (281) 367-1676
<http://www.aronex.com>

Artisan Industries
73 Pond Street
Waltham, MA 02254
(617) 893-6800
<http://www.artisanind.com>

ASI Instruments
12900 Ten Mile Road
Warren, MI 48089
(800) 531-1105 FAX: (810) 756-9737
(810) 756-1222
<http://www.asi-instruments.com>

Aspen Research Laboratories
1700 Buerkle Road
White Bear Lake, MN 55140
(651) 264-6000 FAX: (651) 264-6270
<http://www.aspenresearch.com>

Associates of Cape Cod
704 Main Street
Falmouth, MA 02540
(800) LAL-TEST FAX: (508) 540-8680
(508) 540-3444
<http://www.acciusa.com>

Astra Pharmaceuticals
See AstraZeneca

AstraZeneca
1800 Concord Pike
Wilmington, DE 19850
(302) 886-3000 FAX: (302) 886-2972
<http://www.astrazeneca.com>

AT Biochem
30 Spring Mill Drive
Malvern, PA 19355
(610) 889-9300 FAX: (610) 889-9304

ATC Diagnostics
See Vysis

ATCC
See American Type Culture Collection

Athens Research and Technology
P.O. Box 5494
Athens, GA 30604
(706) 546-0207 FAX: (706) 546-7395

Atlanta Biologicals
1425-400 Oakbrook Drive
Norcross, GA 30093
(800) 780-7788 or (770) 446-1404
FAX: (800) 780-7374 or (770) 446-1404
<http://www.atlantabio.com>

Atomergic Chemical
71 Carolyn Boulevard
Farmingdale, NY 11735
(631) 694-9000 FAX: (631) 694-9177
<http://www.atomergic.com>

Atomic Energy of Canada
2251 Speakman Drive
Mississauga, Ontario
L5K 1B2 Canada
(905) 823-9040 FAX: (905) 823-1290
<http://www.aecl.ca>

ATR
P.O. Box 460
Laurel, MD 20725
(800) 827-5931 FAX: (410) 792-2837
(301) 470-2799
<http://www.atrbiotech.com>

Aurora Biosciences
11010 Torreyana Road
San Diego, CA 92121
(858) 404-6600 FAX: (858) 404-6714
<http://www.aurorabio.com>

Automatic Switch Company
A Division of Emerson Electric
50 Hanover Road
Florham Park, NJ 07932
(800) 937-2726 FAX: (973) 966-2628
(973) 966-2000
<http://www.asco.com>

Avanti Polar Lipids
700 Industrial Park Drive
Alabaster, AL 35007
(800) 227-0651 FAX: (800) 229-1004
(205) 663-2494 FAX: (205) 663-0756
<http://www.avantilipids.com>

Aventis
BP 67917
67917 Strasbourg Cedex 9, France
33 (0) 388 99 11 00
FAX: 33 (0) 388 99 11 01
<http://www.aventis.com>

Aventis Pasteur
1 Discovery Drive
Swiftwater, PA 18370
(800) 822-2463 FAX: (570) 839-0955
(570) 839-7187
<http://www.aventispasteur.com/usa>

Avery Dennison
150 North Orange Grove Boulevard
Pasadena, CA 91103
(800) 462-8379 FAX: (626) 792-7312
(626) 304-2000
<http://www.averydennison.com>

Avestin
2450 Don Reid Drive
Ottawa, Ontario
K1H 1E1 Canada
(888) AVESTIN FAX: (613) 736-8086
(613) 736-0019
<http://www.avestin.com>

AVIV Instruments
750 Vassar Avenue
Lakewood, NJ 08701
(732) 367-1663 FAX: (732) 370-0032
<http://www.avivinst.com>

Axon Instruments
1101 Chess Drive
Foster City, CA 94404
(650) 571-9400 FAX: (650) 571-9500
<http://www.axon.com>

Azon
720 Azon Road
Johnson City, NY 13790
(800) 847-9374 FAX: (800) 635-6042
(607) 797-2368
<http://www.azon.com>

BABCO
1223 South 47th Street
Richmond, CA 94804
(800) 92-BABCO FAX: (510) 412-8940
(510) 412-8930
<http://www.babco.com>

Bacharach
625 Alpha Drive
Pittsburgh, PA 15238
(800) 736-4666 FAX: (412) 963-2091
(412) 963-2000
<http://www.bacharach-inc.com>

Bachem Bioscience
3700 Horizon Drive
King of Prussia, PA 19406
(800) 634-3183 FAX: (610) 239-0800
(610) 239-0300
<http://www.bachem.com>

Bachem California
3132 Kashiwa Street
P.O. Box 3426
Torrance, CA 90510
(800) 422-2436 FAX: (310) 530-1571
(310) 539-4171
<http://www.bachem.com>

Baekon
18866 Allendale Avenue
Saratoga, CA 95070
(408) 972-8779 FAX: (408) 741-0944

Baker Chemical
See J.T. Baker

Bangs Laboratories
9025 Technology Drive
Fishers, IN 46038
(317) 570-7020 FAX: (317) 570-7034
<http://www.bangslabs.com>

Bard Parker
See Becton Dickinson

Barnstead/ThermoLyne
P.O. Box 797
2555 Kerper Boulevard
Dubuque, IA 52004
(800) 446-6060 FAX: (319) 589-0516
<http://www.barnstead.com>

Barrskogen
4612 Laverock Place N
Washington, DC 20007
(800) 237-9192 FAX: (301) 464-7347

BAS
See Bioanalytical Systems

BASF
Specialty Products
3000 Continental Drive North
Mt. Olive, NJ 07828
(800) 669-2273 FAX: (973) 426-2610
<http://www.basf.com>

Baum, W.A.
620 Oak Street
Copiague, NY 11726
(631) 226-3940 FAX: (631) 226-3969
<http://www.wabaum.com>

Bausch & Lomb
One Bausch & Lomb Place
Rochester, NY 14604
(800) 344-8815 FAX: (716) 338-6007
(716) 338-6000
<http://www.bausch.com>

Baxter
Fenwal Division
1627 Lake Cook Road
Deerfield, IL 60015
(800) 766-1077 FAX: (800) 395-3291
(847) 940-6599 FAX: (847) 940-5766
<http://www.powerfulmedicine.com>

Baxter Healthcare
One Baxter Parkway
Deerfield, IL 60015
(800) 777-2298 FAX: (847) 948-3948
(847) 948-2000
<http://www.baxter.com>

Baxter Scientific Products
See VWR Scientific

Bayer
Agricultural Division
Animal Health Products
12707 Shawnee Mission Pkwy.
Shawnee Mission, KS 66201
(800) 255-6517 FAX: (913) 268-2803
(913) 268-2000
<http://www.bayerus.com>

Bayer
Diagnostics Division (Order Services)
P.O. Box 2009
Mishawaka, IN 46546
(800) 248-2637 FAX: (800) 863-6882
(219) 256-3390
<http://www.bayer.com>

Bayer Diagnostics
511 Benedict Avenue
Tarrytown, NY 10591
(800) 255-3232 FAX: (914) 524-2132
(914) 631-8000
<http://www.bayerdiag.com>

Bayer Plc
Diagnostics Division
Bayer House, Strawberry Hill
Newbury, Berkshire RG14 1JA, UK
(44) 1635-563000
FAX: (44) 1635-563393
<http://www.bayer.co.uk>

Suppliers

A.4.4

BD Immunocytometry Systems
2350 Qume Drive
San Jose, CA 95131
(800) 223-8226 FAX: (408) 954-BDIS
<http://www.bdfacs.com>

BD Labware
Two Oak Park
Bedford, MA 01730
(800) 343-2035 FAX: (800) 743-6200
<http://www.bd.com/labware>

BD PharMingen
10975 Torreyana Road
San Diego, CA 92121
(800) 848-6227 FAX: (858) 812-8888
(858) 812-8800
<http://www.pharmingen.com>

BD Transduction Laboratories
133 Venture Court
Lexington, KY 40511
(800) 227-4063 FAX: (606) 259-1413
(606) 259-1550
<http://www.translab.com>

BDH Chemicals
Broom Road
Poole, Dorset BH12 4NN, UK
(44) 1202-745520
FAX: (44) 1202- 2413720

BDH Chemicals
See Hoefer Scientific Instruments

BDIS
See BD Immunocytometry Systems

Beckman Coulter
4300 North Harbor Boulevard
Fullerton, CA 92834
(800) 233-4685 FAX: (800) 643-4366
(714) 871-4848
<http://www.beckman-coulter.com>

Beckman Instruments
Spinco Division/Bioproductions Operation
1050 Page Mill Road
Palo Alto, CA 94304
(800) 742-2345 FAX: (415) 859-1550
(415) 857-1150
<http://www.beckman-coulter.com>

Becton Dickinson Immunocytometry & Cellular Imaging
2350 Qume Drive
San Jose, CA 95131
(800) 223-8226 FAX: (408) 954-2007
(408) 432-9475
<http://www.bdfacs.com>

Becton Dickinson Labware
1 Becton Drive
Franklin Lakes, NJ 07417
(888) 237-2762 FAX: (800) 847-2220
(201) 847-4222
<http://www.bdfacs.com>

Becton Dickinson Labware
2 Bridgewater Lane
Lincoln Park, NJ 07035
(800) 235-5953 FAX: (800) 847-2220
(201) 847-4222
<http://www.bdfacs.com>

Becton Dickinson Primary
Care Diagnostics
7 Loveton Circle
Sparks, MD 21152
(800) 675-0908 FAX: (410) 316-4723
(410) 316-4000
<http://www.bdfacs.com>

Behringwerke Diagnostika
Hoechster Strasse 70
P-65835 Liederback, Germany
(49) 69-30511 FAX: (49) 69-303-834

Bellco Glass
340 Edrudo Road
Vineland, NJ 08360
(800) 257-7043 FAX: (856) 691-3247
(856) 691-1075
<http://www.bellcoglass.com>

Bender Biosystems
See Serva

Beral Enterprises
See Garren Scientific

Berkeley Antibody
See BAbCO

Bernsco Surgical Supply
25 Plant Avenue
Hauppauge, NY 11788
(800) TIEMANN FAX: (516) 273-6199
(516) 273-0005
<http://www.bernscoco.com>

Beta Medical and Scientific
(Datesand Ltd.)
2 Ferndale Road
Sale, Manchester M33 3GP, UK
(44) 1612 317676
FAX: (44) 1612 313656

Bethesda Research Laboratories (BRL)
See Life Technologies

Biacore
200 Centennial Avenue, Suite 100
Piscataway, NJ 08854
(800) 242-2599 FAX: (732) 885-5669
(732) 885-5618
<http://www.biacore.com>

Bilaney Consultants
St. Julian's
Sevenoaks, Kent TN15 0RX, UK
(44) 1732 450002
FAX: (44) 1732 450003
<http://www.bilaney.com>

Binding Site
5889 Oberlin Drive, Suite 101
San Diego, CA 92121
(800) 633-4484 FAX: (619) 453-9189
(619) 453-9177
<http://www.bindingsite.co.uk>

BIO 101
See Qbiogene

Bio Image
See Genomic Solutions

Bioanalytical Systems
2701 Kent Avenue
West Lafayette, IN 47906
(800) 845-4246 FAX: (765) 497-1102
(765) 463-4527
<http://www.bioanalytical.com>

Biocell
2001 University Drive
Rancho Dominguez, CA 90220
(800) 222-8382 FAX: (310) 637-3927
(310) 537-3300
<http://www.biocell.com>

Biocoat
See BD Labware

BioComp Instruments
650 Churchill Road
Fredericton, New Brunswick
E3B 1P6 Canada
(800) 561-4221 FAX: (506) 453-3583
(506) 453-4812
<http://131.202.97.21>

BioDesign
P.O. Box 1050
Carmel, NY 10512
(914) 454-6610 FAX: (914) 454-6077
<http://www.biodesignofny.com>

BioDiscovery
4640 Admiralty Way, Suite 710
Marina Del Rey, CA 90292
(310) 306-9310 FAX: (310) 306-9109
<http://www.biodiscovery.com>

Bioengineering AG
Sagenrainstrasse 7
CH8636 Wald, Switzerland
(41) 55-256-8-111
FAX: (41) 55-256-8-256

Biofluids
Division of Biosource International
1114 Taft Street
Rockville, MD 20850
(800) 972-5200 FAX: (301) 424-3619
(301) 424-4140
<http://www.biosource.com>

BioFX Laboratories
9633 Liberty Road, Suite S
Randallstown, MD 21133
(800) 445-6447 FAX: (410) 498-6008
(410) 496-6006
<http://www.biofx.com>

BioGenex Laboratories
4600 Norris Canyon Road
San Ramon, CA 94583
(800) 421-4149 FAX: (925) 275-0580
(925) 275-0550
<http://www.biogenex.com>

Bioline
2470 Wrondel Way
Reno, NV 89502
(888) 257-5155 FAX: (775) 828-7676
(775) 828-0202
<http://www.bioline.com>

Bio-Logic Research & Development
1, rue de l'Europe
A.Z. de Font-Ratel
38640 CLAIX, France
(33) 76-98-68-31
FAX: (33) 76-98-69-09

Biological Detection Systems
See Cellomics or Amersham

Biomeda
1166 Triton Drive, Suite E
P.O. Box 8045
Foster City, CA 94404
(800) 341-8787 FAX: (650) 341-2299
(650) 341-8787
<http://www.biomeda.com>

BioMedic Data Systems
1 Silas Road
Seaford, DE 19973
(800) 526-2637 FAX: (302) 628-4110
(302) 628-4100
<http://www.bmds.com>

Biomedical Engineering
P.O. Box 980694
Virginia Commonwealth University
Richmond, VA 23298
(804) 828-9829 FAX: (804) 828-1008

Biomedical Research Instruments
12264 Wilkins Avenue
Rockville, MD 20852
(800) 327-9498
(301) 881-7911
<http://www.biomedinstr.com>

Bio/medical Specialties
P.O. Box 1687
Santa Monica, CA 90406
(800) 269-1158 FAX: (800) 269-1158
(323) 938-7515

BioMerieux
100 Rodolphe Street
Durham, North Carolina 27712
(919) 620-2000
<http://www.biomerieux.com>

BioMetallics
P.O. Box 2251
Princeton, NJ 08543
(800) 999-1961 FAX: (609) 275-9485
(609) 275-0133
<http://www.microplate.com>

Suppliers

A.4.5

Biomol Research Laboratories
5100 Campus Drive
Plymouth Meeting, PA 19462
(800) 942-0430 FAX: (610) 941-9252
(610) 941-0430
<http://www.biomol.com>

Bionique Testing Labs
Fay Brook Drive
RR 1, Box 196
Saranac Lake, NY 12983
(518) 891-2356 FAX: (518) 891-5753
<http://www.bionique.com>

Biopac Systems
42 Aero Camino
Santa Barbara, CA 93117
(805) 685-0066 FAX: (805) 685-0067
<http://www.biopac.com>

Bioproducts for Science
See Harlan Bioproducts for Science

Bioptechs
3560 Beck Road
Butler, PA 16002
(877) 548-3235 FAX: (724) 282-0745
(724) 282-7145
<http://www.bioptechs.com>

BIOQUANT-R&M Biometrics
5611 Ohio Avenue
Nashville, TN 37209
(800) 221-0549 (615) 350-7866
FAX: (615) 350-7282
<http://www.bioquant.com>

Bio-Rad Laboratories
2000 Alfred Nobel Drive
Hercules, CA 94547
(800) 424-6723 FAX: (800) 879-2289
(510) 741-1000 FAX: (510) 741-5800
<http://www.bio-rad.com>

Bio-Rad Laboratories
Maylands Avenue
Hemel Hempstead, Herts HP2 7TD, UK
<http://www.bio-rad.com>

Bioreclamation
492 Richmond Road
East Meadow, NY 11554
(516) 483-1196 FAX: (516) 483-4683
<http://www.bioreclamation.com>

BioRobotics
3-4 Bennell Court
Comberton, Cambridge CB3 7DS, UK
(44) 1223-264345
FAX: (44) 1223-263933
<http://www.biorobotics.co.uk>

BIOS Laboratories
See Genaisance Pharmaceuticals

Biotech Research Technologies
81 Digital Drive
Novato, CA 94949
(800) GENOME1 FAX: (415) 883-8488
(415) 883-8400
<http://www.biotechresearch.com>

BioSeptra
111 Locke Drive
Marlborough, MA 01752
(800) 752-5277 FAX: (508) 357-7595
(508) 357-7500
<http://www.bioseptra.com>

Bio-Serv
1 8th Street, Suite 1
Frenchtown, NJ 08825
(908) 996-2155 FAX: (908) 996-4123
<http://www.bio-serv.com>

BioSignal
1744 William Street, Suite 600
Montreal, Quebec
H3J 1R4 Canada
(800) 293-4501 FAX: (514) 937-0777
(514) 937-1010
<http://www.biosignal.com>

BioSoft
P.O. Box 10938
Ferguson, MO 63135
(314) 524-8029 FAX: (314) 524-8129
<http://www.biosoft.com>

Biosource International
820 Flynn Road
Camarillo, CA 93012
(800) 242-0607 FAX: (805) 987-3385
(805) 987-0086
<http://www.biosource.com>

BioSpec Products
P.O. Box 788
Bartlesville, OK 74005
(800) 617-3363 FAX: (918) 336-3363
(918) 336-3363
<http://www.biospec.com>

Biosure
See Riese Enterprises

Biosym Technologies
See Molecular Simulations

Biosys
21 quai du Clos des Roses
602000 Compiègne, France
(33) 03 4486 2275
FAX: (33) 03 4484 2297

Bio-Tech Research Laboratories
NIAID Repository
Rockville, MD 20850
<http://www.niaid.nih.gov/ncn/repos.htm>

Biotech Instruments
Biotech House
75A High Street
Kimpton, Hertfordshire SG4 8PU, UK
(44) 1438 832555
FAX: (44) 1438 833040
<http://www.biotinst.demon.co.uk>

Biotech International
11 Durbell Street
Acacia Ridge, Queensland 4110
Australia
61-7-3370-6396
FAX: 61-7-3370-6370
<http://www.avianbiotech.com>

Biotech Source
Inland Farm Drive
South Windham, ME 04062
(207) 892-3266 FAX: (207) 892-6774

Bio-Tek Instruments
Highland Industrial Park
P.O. Box 998
Winooski, VT 05404
(800) 451-5172 FAX: (802) 655-7941
(802) 655-4040
<http://www.biotek.com>

Biotech Laboratories
6023 South Loop East
Houston, TX 77033
(800) 535-6286 FAX: (713) 643-3143
(713) 643-0606
<http://www.biotech.com>

BioTherm
3260 Wilson Boulevard
Arlington, VA 22201
(703) 522-1705 FAX: (703) 522-2606

Bioventures
P.O. Box 2561
848 Scott Street
Murfreesboro, TN 37133
(800) 235-8938 FAX: (615) 896-4837
<http://www.bioventures.com>

BioWhittaker
8830 Biggs Ford Road
P.O. Box 127
Walkersville, MD 21793
(800) 638-8174 FAX: (301) 845-8338
(301) 898-7025
<http://www.biowhittaker.com>

Biozyme Laboratories
9939 Hibert Street, Suite 101
San Diego, CA 92131
(800) 423-8199 FAX: (858) 549-0138
(858) 549-4484
<http://www.biozyme.com>

Bird Products
1100 Bird Center Drive
Palm Springs, CA 92262
(800) 328-4139 FAX: (760) 778-7274
(760) 778-7200
<http://www.birdprod.com/bird>

B & K Universal
2403 Yale Way
Fremont, CA 94538
(800) USA-MICE FAX: (510) 490-3036

BLS Ltd.
Zselyi Aladar u. 31
1165 Budapest, Hungary
(36) 1-407-2602 FAX: (36) 1-407-2896
<http://www.bls-ltd.com>

Blue Sky Research
3047 Orchard Parkway
San Jose, CA 95134
(408) 474-0988 FAX: (408) 474-0989
<http://www.blueskyresearch.com>

Blumenthal Industries
7 West 36th Street, 13th floor
New York, NY 10018
(212) 719-1251 FAX: (212) 594-8828

Boblin Instruments
11 Harts Lane, Suite 1
East Brunswick, NJ 08816
(732) 254-7742 FAX: (732) 254-1577
<http://www.boblin.co.uk>

BOC Edwards
One Edwards Park
301 Ballardvale Street
Wilmington, MA 01887
(800) 848-9800 FAX: (978) 658-7969
(978) 658-5410
<http://www.bocedwards.com>

Boehringer Ingelheim
900 Ridgebury Road
P.O. Box 368
Ridgefield, CT 06877
(800) 243-0127 FAX: (203) 798-6234
(203) 798-9988
<http://www.boehringer-ingelheim.com>

Boehringer Mannheim
Biochemicals Division
See Roche Diagnostics

Boekel Scientific
855 Pennsylvania Boulevard
Feasterville, PA 19053
(800) 336-6929 FAX: (215) 396-8264
(215) 396-8200
<http://www.boekelsci.com>

Bohdan Automation
1500 McCormack Boulevard
Mundelein, IL 60060
(708) 680-3939 FAX: (708) 680-1199

BPAmoco
4500 McGinnis Ferry Road
Alpharetta, GA 30005
(800) 328-4537 FAX: (770) 772-8213
(770) 772-8200
<http://www.bpamoco.com>

Brain Research Laboratories
Waban P.O. Box 88
Newton, MA 02468
(888) BRL-5544 FAX: (617) 965-6220
(617) 965-5544
<http://www.brainresearchlab.com>

Braintree Scientific
P.O. Box 850929
Braintree, MA 02185
(781) 843-1644 FAX: (781) 982-3160
<http://www.braintreesci.com>

Brandel
8561 Atlas Drive
Gaithersburg, MD 20877
(800) 948-6506 FAX: (301) 869-5570
(301) 948-6506
<http://www.brandel.com>

Suppliers

A.4.6

Branson Ultrasonics
41 Eagle Road
Danbury, CT 06813
(203) 796-0400 FAX: (203) 796-9838
<http://www.plasticsnet.com/branson>

B. Braun Biotech
999 Postal Road
Allentown, PA 18103
(800) 258-9000 FAX: (610) 266-9319
(610) 266-6262
<http://www.bbraunbiotech.com>

B. Braun Biotech International
Schwarzenberg Weg 73-79
P.O. Box 1120
D-34209 Melsungen, Germany
(49) 5661-71-3400
FAX: (49) 5661-71-3702
<http://www.bbraunbiotech.com>

B. Braun-McGaw
2525 McGaw Avenue
Irvine, CA 92614
(800) BBRAUN-2 (800) 624-2963
<http://www.bbraunusa.com>

B. Braun Medical
Thorncliffe Park
Sheffield S35 2PW, UK
(44) 114-225-9000
FAX: (44) 114-225-9111
<http://www.bbmuk.demon.co.uk>

Brenntag
P.O. Box 13788
Reading, PA 19612-3788
(610) 926-4151 FAX: (610) 926-4160
<http://www.brenntagnortheast.com>

Bresatec
See GeneWorks

Bright/Hacker Instruments
17 Sherwood Lane
Fairfield, NJ 07004
(973) 226-8450 FAX: (973) 808-8281
<http://www.hackerinstruments.com>

Brinkmann Instruments
Subsidiary of Sybron
1 Cantiague Road
P.O. Box 1019
Westbury, NY 11590
(800) 645-3050 FAX: (516) 334-7521
(516) 334-7500
<http://www.brinkmann.com>

Bristol-Meyers Squibb
P.O. Box 4500
Princeton, NJ 08543
(800) 631-5244 FAX: (800) 523-2965
<http://www.bms.com>

Broadley James
19 Thomas
Irvine, CA 92618
(800) 288-2833 FAX: (949) 829-5560
(949) 829-5555
<http://www.broadleyjames.com>

Brookhaven Instruments
750 Blue Point Road
Holtsville, NY 11742
(631) 758-3200 FAX: (631) 758-3255
<http://www.bic.com>

Brownlee Labs
See Applied Biosystems
Distributed by Pacer Scientific

Bruel & Kjaer
Division of Spectris Technologies
2815 Colonnades Court
Norcross, GA 30071
(800) 332-2040 FAX: (770) 847-8440
(770) 209-6907
<http://www.bkhome.com>

Bruker Analytical X-Ray Systems
5465 East Cheryl Parkway
Madison, WI 53711
(800) 234-XRAY FAX: (608) 276-3006
(608) 276-3000
<http://www.bruker-axs.com>

Bruker Instruments
19 Fortune Drive
Billerica, MA 01821
(978) 667-9580 FAX: (978) 667-0985
<http://www.bruker.com>

BTX
Division of Genetronics
11199 Sorrento Valley Road
San Diego, CA 92121
(800) 289-2465 FAX: (858) 597-9594
(858) 597-6006
<http://www.genetronics.com/btx>

Buchler Instruments
See Baxter Scientific Products

Büchi Labortechnik AG
Meierseggstrasse 40
Postfach
CH-9230 Flawil 1
Switzerland
41-71-394-63 63 Fax: 41-71-394-64 64

Buckshire
2025 Ridge Road
Perkasie, PA 18944
(215) 257-0116

Burdick and Jackson
Division of Baxter Scientific Products
1953 S. Harvey Street
Muskegon, MI 49442
(800) 368-0050 FAX: (231) 728-8226
(231) 726-3171
<http://www.bandj.com/mainframe.htm>

Burleigh Instruments
P.O. Box E
Fishers, NY 14453
(716) 924-9355 FAX: (716) 924-9072
<http://www.burleigh.com>

Burns Veterinary Supply
1900 Diplomat Drive
Farmer's Branch, TX 75234
(800) 92-BURNS FAX: (972) 243-6841
<http://www.burnsvet.com>

Burroughs Wellcome
See Glaxo Wellcome

The Butler Company
5600 Blazer Parkway
Dublin, OH 43017
(800) 551-3861 FAX: (614) 761-9096
(614) 761-9095
<http://www.wabutler.com>

Butterworth Laboratories
54-56 Waldegrave Road
Teddington, Middlesex
TW11 8LG, UK
(44)(0)20-8977-0750
FAX: (44)(0)28-8943-2624
<http://www.butterworth-labs.co.uk>

Buxco Electronics
95 West Wood Road #2
Sharon, CT 06069
(860) 364-5558 FAX: (860) 364-5116
<http://www.buxco.com>

C/D/N Isotopes
88 Leacock Street
Pointe-Claire, Quebec
H9R 1H1 Canada
(800) 697-6254 FAX: (514) 697-6148

C.M.A./Microdialysis AB
73 Princeton Street
North Chelmsford, MA 01863
(800) 440-4980 FAX: (978) 251-1950
(978) 251-1940
<http://www.microdialysis.com>

Calbiochem-Novabiochem
P.O. Box 12087-2087
La Jolla, CA 92039
(800) 854-3417 FAX: (800) 776-0999
(858) 450-9600
<http://www.calbiochem.com>

California Fine Wire
338 South Fourth Street
Grover Beach, CA 93433
(805) 489-5144 FAX: (805) 489-5352
<http://www.calfinewire.com>

Caligor
846 Pelham Parkway
Pelham Manor, NY 10803
(800) 225-9906 Fax: 914-738-9538
<http://www.caligor.com>

Calorimetry Sciences
155 West 2050 North
Spanish Fork, UT 84660
(801) 794-2600 FAX: (801) 794-2700
<http://www.calscorp.com>

Caltag Laboratories
1849 Bayshore Highway, Suite 200
Burlingame, CA 94010
(800) 874-4007 FAX: (650) 652-9030
(650) 652-0468
<http://www.caltag.com>

Cambrex Corporation
1 Meadowlands Plaza
East Rutherford, NJ 07073
(201) 804-3000 FAX: (201) 804-9852
<http://www.cambrex.com>

Cambridge Electronic Design
Science Park, Milton Road
Cambridge CB4 0FE, UK
44 (0) 1223-420-186
FAX: 44 (0) 1223-420-488
<http://www.ced.co.uk>

Cambridge Isotope Laboratories
50 Frontage Road
Andover, MA 01810
(800) 322-1174 FAX: (978) 749-2768
(978) 749-8000
<http://www.isotope.com>

Cambridge Research Biochemicals
See Zeneca/CRB

Cambridge Technology
109 Smith Place
Cambridge, MA 02138
(617) 441-0600 FAX: (617) 497-8800
<http://www.camtech.com>

Camlab
Nuffield Road
Cambridge CB4 1TH, UK
(44) 122-3424222
FAX: (44) 122-3420856
<http://www.camlab.co.uk/home.htm>

Campden Instruments
Park Road
Sibley Loughborough
Leicestershire LE12 7TU, UK
(44) 1509-814790
FAX: (44) 1509-816097
<http://www.campden-inst.com/home.htm>

Captiol Equipment Corp.
900 Middlesex Turnpike
Billerica, MA 01821
(978) 663-2002 FAX: (978) 663-2626

Cappel Laboratories
See Organon Teknika Cappel

Carl Roth GmbH & Company
Schoemperlenstrasse 1-5
76185 Karlsruhe
Germany
(49) 72-156-06164
FAX: (49) 72-156-06264
<http://www.carl-roth.de>

Carl Zeiss
One Zeiss Drive
Thornwood, NY 10594
(800) 233-2343 FAX: (914) 681-7446
(914) 747-1800
<http://www.zeiss.com>

Suppliers

A.4.7

Carlo Erba Reagenti
Via Winckelmann 1
20148 Milano
Lombardia, Italy
(39) 0-29-5231
FAX: (39) 0-29-5235-904
<http://www.carloerbareagenti.com>

Carolina Biological Supply
2700 York Road
Burlington, NC 27215
(800) 334-5551 FAX: (336) 584-76869
(336) 584-0381
<http://www.carolina.com>

Carolina Fluid Components
9309 Stockport Place
Charlotte, NC 28273
(704) 588-6101 FAX: (704) 588-6115
<http://www.cfcsite.com>

Cartesian Technologies
17851 Skypark Circle, Suite C
Irvine, CA 92614
(800) 935-8007
<http://cartesiantech.com>

Cayman Chemical
1180 East Ellsworth Road
Ann Arbor, MI 48108
(800) 364-9897 FAX: (734) 971-3640
(734) 971-3335
<http://www.caymanchem.com>

CB Sciences
One Washington Street, Suite 404
Dover, NH 03820
(800) 234-1757 FAX: (603) 742-2455
<http://www.cbsci.com>

CBS Scientific
P.O. Box 856
Del Mar, CA 92014
(800) 243-4959 FAX: (858) 755-0733
(858) 755-4959
<http://www.cbssci.com>

CCR (Coriell Cell Repository)
See Coriell Institute for Medical Research

CE Instruments
Grand Avenue Parkway
Austin, TX 78728
(800) 876-6711 FAX: (512) 251-1597
<http://www.ceinstruments.com>

Cedarlane Laboratories
5516 8th Line, R.R. #2
Hornby, Ontario
L0P 1E0 Canada
(905) 878-8891 FAX: (905) 878-7800
<http://www.cedarlanelabs.com>

CEL Associates
P.O. Box 721854
Houston, TX 77272
(800) 537-9339 FAX: (281) 933-0922
(281) 933-9339
<http://www.cel-1.com>

Cel-Line Associates
See Erie Scientific

Celite World Minerals
130 Castilian Drive
Santa Barbara, CA 93117
(805) 562-0200 FAX: (805) 562-0299
<http://www.worldminerals.com/celite>

Cell Genesys
342 Lakeside Drive
Foster City, CA 94404
(650) 425-4400 FAX: (650) 425-4457
<http://www.cellgenesys.com>

Cell Signaling Technology
166B Cummings Center
Beverly, MA 01915
(877) 616-CELL FAX: (978) 867-2388
(978) 867-2488
<http://www.cellsignal.com>

Cell Systems
12815 NE 124th Street, Suite A
Kirkland, WA 98034
(800) 697-1211 FAX: (425) 820-6762
(425) 823-1010

Cellmark Diagnostics
20271 Goldenrod Lane
Germantown, MD 20876
(800) 872-5227 FAX: (301) 428-4877
(301) 428-4980
<http://www.cellmark-labs.com>

Cellomics
635 William Pitt Way
Pittsburgh, PA 15238
(888) 826-3857 FAX: (412) 826-3850
(412) 826-3600
<http://www.cellomics.com>

CellPoint Scientific
P.O. Box 10757
Gaithersburg, MD 20898
(301) 208-0696 Fax: (301) 590-1290
(800) 424-2984

Celltech
216 Bath Road
Slough, Berkshire SL1 4EN, UK
(44) 1753 534655
FAX: (44) 1753 536632
<http://www.celltech.co.uk>

Cellular Products
872 Main Street
Buffalo, NY 14202
(800) CPI-KITS FAX: (716) 882-0959
(716) 882-0920
<http://www.zeptometrix.com>

CEM
P.O. Box 200
Matthews, NC 28106
(800) 726-3331

Centers for Disease Control
1600 Clifton Road NE
Atlanta, GA 30333
(800) 311-3435 FAX: (888) 232-3228
(404) 639-3311
<http://www.cdc.gov>

CERJ
Centre d'Elevage Roger Janvier
53940 Le Genest Saint Isle
France

Cetus
See Chiron

Chance Propper
Warily, West Midlands B66 1NZ, UK
(44)(0)121-553-5551
FAX: (44)(0)121-525-0139

Charles River Laboratories
251 Ballardvale Street
Wilmington, MA 01887
(800) 522-7287 FAX: (978) 658-7132
(978) 658-6000
<http://www.criver.com>

Charm Sciences
36 Franklin Street
Malden, MA 02148
(800) 343-2170 FAX: (781) 322-3141
(781) 322-1523
<http://www.charm.com>

Chase-Walton Elastomers
29 Apsley Street
Hudson, MA 01749
(800) 448-6289 FAX: (978) 562-5178
(978) 568-0202
<http://www.chase-walton.com>

ChemGenes
Ashland Technology Center
200 Homer Avenue
Ashland, MA 01721
(800) 762-9323 FAX: (508) 881-3443
(508) 881-5200
<http://www.chemgenes.com>

Chemglass
3861 North Mill Road
Vineland, NJ 08360
(800) 843-1794 FAX: (856) 696-9102
(800) 696-0014
<http://www.chemglass.com>

Chemicon International
28835 Single Oak Drive
Temecula, CA 92590
(800) 437-7500 FAX: (909) 676-9209
(909) 676-8080
<http://www.chemicon.com>

Chem-Impex International
935 Dillon Drive
Wood Dale, IL 60191
(800) 869-9290 FAX: (630) 766-2218
(630) 766-2112
<http://www.chemimpex.com>

Chem Service
P.O. Box 599
West Chester, PA 19381-0599
(610) 692-3026 FAX: (610) 692-8729
<http://www.chemservice.com>

Chemsyn Laboratories
13605 West 96th Terrace
Lenexa, KS 66215
(913) 541-0525 FAX: (913) 888-3582
<http://www.tech.epcorp.com/ChemSyn/chemsyn.htm>

Chemunex USA
1 Deer Park Drive, Suite H-2
Monmouth Junction, NJ 08852
(800) 411-6734
<http://www.chemunex.com>

Cherwell Scientific Publishing
The Magdalen Centre
Oxford Science Park
Oxford OX44GA, UK
(44)(1) 865-784-800
FAX: (44)(1) 865-784-801
<http://www.cherwell.com>

ChiRex Cauldron
383 Phoenixville Pike
Malvern, PA 19355
(610) 727-2215 FAX: (610) 727-5762
<http://www.chirex.com>

Chiron Diagnostics
See Bayer Diagnostics

Chiron Mimotopes Peptide Systems
See Multiple Peptide Systems

Chiron
4560 Horton Street
Emeryville, CA 94608
(800) 244-7668 FAX: (510) 655-9910
(510) 655-8730
<http://www.chiron.com>

Chrom Tech
P.O. Box 24248
Apple Valley, MN 55124
(800) 822-5242 FAX: (952) 431-6345
<http://www.chromtech.com>

Chroma Technology
72 Cotton Mill Hill, Unit A-9
Brattleboro, VT 05301
(800) 824-7662 FAX: (802) 257-9400
(802) 257-1800
<http://www.chroma.com>

Chromatographie
ZAC de Moulin No. 2
91160 Saux les Chartreux
France
(33) 01-64-54-8969
FAX: (33) 01-69-0988091
<http://www.chromatographie.com>

Chromogenix
Täljegardsgatan 3
431-53 Mlndal, Sweden
(46) 31-706-20-70
FAX: (46) 31-706-20-80
<http://www.chromogenix.com>

Suppliers

A.4.8

Chrompack USA
c/o Varian USA
2700 Mitchell Drive
Walnut Creek, CA 94598
(800) 526-3687 FAX: (925) 945-2102
(925) 939-2400
<http://www.chrompack.com>

Chugai Biopharmaceuticals
6275 Nancy Ridge Drive
San Diego, CA 92121
(858) 535-5900 FAX: (858) 546-5973
<http://www.chugaibio.com>

Ciba-Corning Diagnostics
See Bayer Diagnostics

Ciba-Geigy
See Ciba Specialty Chemicals or
Novartis Biotechnology

Ciba Specialty Chemicals
540 White Plains Road
Tarrytown, NY 10591
(800) 431-1900 FAX: (914) 785-2183
(914) 785-2000
<http://www.cibasc.com>

Ciba Vision
Division of Novartis AG
11460 Johns Creek Parkway
Duluth, GA 30097
(770) 476-3937
<http://www.cvworld.com>

Cidex
Advanced Sterilization Products
33 Technology Drive
Irvine, CA 92618
(800) 595-0200 (949) 581-5799
<http://www.cidex.com/ASPnew.htm>

Cinna Scientific
Subsidiary of Molecular Research Center
5645 Montgomery Road
Cincinnati, OH 45212
(800) 462-9868 FAX: (513) 841-0080
(513) 841-0900
<http://www.mrcgene.com>

Cistron Biotechnology
10 Bloomfield Avenue
Pine Brook, NJ 07058
(800) 642-0167 FAX: (973) 575-4854
(973) 575-1700
<http://www.cistronbio.com>

Clark Electromedical Instruments
See Harvard Apparatus

Clay Adam
See Becton Dickinson Primary Care
Diagnostics

**CLB (Central Laboratory
of the Netherlands)**
Blood Transfusion Service
P.O. Box 9190
1006 AD Amsterdam, The Netherlands
(31) 20-512-9222
FAX: (31) 20-512-3332

Cleveland Scientific
P.O. Box 300
Bath, OH 44210
(800) 952-7315 FAX: (330) 666-2240
<http://www.clevelandscientific.com>

Clonetics
Division of BioWhittaker
<http://www.clonetics.com>
Also see BioWhittaker

Clontech Laboratories
1020 East Meadow Circle
Palo Alto, CA 94303
(800) 662-2566 FAX: (800) 424-1350
(650) 424-8222 FAX: (650) 424-1088
<http://www.clontech.com>

Closure Medical Corporation
5250 Greens Dairy Road
Raleigh, NC 27616
(919) 876-7800 FAX: (919) 790-1041
<http://www.closuremed.com>

CMA Microdialysis AB
73 Princeton Street
North Chelmsford, MA 01863
(800) 440-4980 FAX: (978) 251-1950
(978) 251 1940
<http://www.microdialysis.com>

Cocalico Biologicals
449 Stevens Road
P.O. Box 265
Reamstown, PA 17567
(717) 336-1990 FAX: (717) 336-1993

Coherent Laser
5100 Patrick Henry Drive
Santa Clara, CA 95056
(800) 227-1955 FAX: (408) 764-4800
(408) 764-4000
<http://www.cohr.com>

Cohu
P.O. Box 85623
San Diego, CA 92186
(858) 277-6700 FAX: (858) 277-0221
<http://www.COHU.com/cctv>

Cole-Parmer Instrument
625 East Bunker Court
Vernon Hills, IL 60061
(800) 323-4340 FAX: (847) 247-2929
(847) 549-7600
<http://www.coleparmer.com>

**Collaborative Biomedical Products
and Collaborative Research**
See Becton Dickinson Labware

Collagen Aesthetics
1850 Embarcadero Road
Palo Alto, CA 94303
(650) 856-0200 FAX: (650) 856-0533
<http://www.collagen.com>

Collagen Corporation
See Collagen Aesthetics

College of American Pathologists
325 Waukegan Road
Northfield, IL 60093
(800) 323-4040 FAX: (847) 832-8000
(847) 446-8800
<http://www.cap.org/index.cfm>

Colonial Medical Supply
504 Wells Road
Franconia, NH 03580
(603) 823-9911 FAX: (603) 823-8799
<http://www.colmedsupply.com>

Colorado Serum
4950 York Street
Denver, CO 80216
(800) 525-2065 FAX: (303) 295-1923
<http://www.colorado-serum.com>

Columbia Diagnostics
8001 Research Way
Springfield, VA 22153
(800) 336-3081 FAX: (703) 569-2353
(703) 569-7511
<http://www.columbiadiagnostics.com>

Columbus Instruments
950 North Hague Avenue
Columbus, OH 43204
(800) 669-5011 FAX: (614) 276-0529
(614) 276-0861
<http://www.columbusinstruments.com>

Compu Cyte Corp.
12 Emily Street
Cambridge, MA 02139
(800) 840-1303 FAX: (617) 577-4501
(617) 492-1300
<http://www.compucyte.com>

Compugen
25 Leek Crescent
Richmond Hill, Ontario
L4B 4B3 Canada
800-387-5045 FAX: (905) 707-2020
(905) 707-2000
<http://www.compugen.com/locations.htm>

Computer Associates International
One Computer Associates Plaza
Islandia, NY 11749
(631) 342-6000 FAX: (631) 342-6800
<http://www.cai.com>

Connaught Laboratories
See Aventis Pasteur

Connectix
2955 Campus Drive, Suite 100
San Mateo, CA 94403
(800) 950-5880 FAX: (650) 571-0850
(650) 571-5100
<http://www.connectix.com>

Contech
99 Hartford Avenue
Providence, RI 02909
(401) 351-4890 FAX: (401) 421-5072
<http://www.iol.ie/~burke/contech.html>

Continental Laboratory Products
5648 Copley Drive
San Diego, CA 92111
(800) 456-7741 FAX: (858) 279-5465
(858) 279-5000
<http://www.conlab.com>

ConvaTec
Professional Services
P.O. Box 5254
Princeton, NJ 08543
(800) 422-8811
<http://www.convatec.com>

Cooper Instruments & Systems
P.O. Box 3048
Warrenton, VA 20188
(800) 344-3921 FAX: (540) 347-4755
(540) 349-4746
<http://www.cooperinstruments.com>

Cooperative Human Tissue Network
(866) 462-2486
<http://www.chtn.ims.nci.nih.gov>

Cora Styles Needles 'N Blocks
56 Milton Street
Arlington, MA 02474
(781) 648-6289 FAX: (781) 641-7917

Coriell Cell Repository (CCR)
See Coriell Institute for Medical Research

Coriell Institute for Medical Research
Human Genetic Mutant Repository
401 Haddon Avenue
Camden, NJ 08103
(856) 966-7377 FAX: (856) 964-0254
<http://arginine.umdnj.edu>

Corion
8 East Forge Parkway
Franklin, MA 02038
(508) 528-4411 FAX: (508) 520-7583
(800) 598-6783
<http://www.corion.com>

**Corning and
Corning Science Products**
P.O. Box 5000
Corning, NY 14831
(800) 222-7740 FAX: (607) 974-0345
(607) 974-9000
<http://www.corning.com>

Costar
See Corning

Coulbourn Instruments
7462 Penn Drive
Allentown, PA 18106
(800) 424-3771 FAX: (610) 391-1333
(610) 395-3771
<http://www.coulbourninst.com>

Coulter Cytometry
See Beckman Coulter

Suppliers

A.4.9

Covance Research Products
465 Swampbridge Road
Denver, PA 17517
(800) 345-4114 FAX: (717) 336-5344
(717) 336-4921
<http://www.covance.com>

Coy Laboratory Products
14500 Coy Drive
Grass Lake, MI 49240
(734) 475-2200 FAX: (734) 475-1846
<http://www.coylab.com>

CPG
3 Borinski Road
Lincoln Park, NJ 07035
(800) 362-2740 FAX: (973) 305-0884
(973) 305-8181
<http://www.cpg-biotech.com>

CPL Scientific
43 Kingfisher Court
Hambridge Road
Newbury RG14 5SJ, UK
(44) 1635-574902
FAX: (44) 1635-529322
<http://www.cplscientific.co.uk>

CraMar Technologies
8670 Wolff Court, #160
Westminster, CO 80030
(800) 4-TOMTEC
<http://www.cramar.com>

Crescent Chemical
1324 Motor Parkway
Hauppauge, NY 11788
(800) 877-3225 FAX: (631) 348-0913
(631) 348-0333
<http://www.creschem.com>

Crist Instrument
P.O. Box 128
10200 Moxley Road
Damascus, MD 20872
(301) 253-2184 FAX: (301) 253-0069
<http://www.cristinstrument.com>

Cruachem
See Annovis
<http://www.cruachem.com>

CS Bio
1300 Industrial Road
San Carlos, CA 94070
(800) 627-2461 FAX: (415) 802-0944
(415) 802-0880
<http://www.csbio.com>

CS-Chromatographie Service
Am Parir 27
D-52379 Langerwehe, Germany
(49) 2423-40493-0
FAX: (49) 2423-40493-49
<http://www.cs-chromatographie.de>

Cuno
400 Research Parkway
Meriden, CT 06450
(800) 231-2259 FAX: (203) 238-8716
(203) 237-5541
<http://www.cuno.com>

Curtin Matheson Scientific
9999 Veterans Memorial Drive
Houston, TX 77038
(800) 392-3353 FAX: (713) 878-3598
(713) 878-3500

CWE
124 Sibley Avenue
Ardmore, PA 19003
(610) 642-7719 FAX: (610) 642-1532
<http://www.cwe-inc.com>

Cybex Computer Products
4991 Corporate Drive
Huntsville, AL 35805
(800) 932-9239 FAX: (800) 462-9239
<http://www.cybex.com>

Cygnus Technology
P.O. Box 219
Delaware Water Gap, PA 18327
(570) 424-5701 FAX: (570) 424-5630
<http://www.cygnustech.com>

Cymbus Biotechnology
Eagle Class, Chandler's Ford
Hampshire SO53 4NF, UK
(44) 1-703-267-676
FAX: (44) 1-703-267-677
<http://www.biotech@cymbus.com>

Cytogen
600 College Road East
Princeton, NJ 08540
(609) 987-8200 FAX: (609) 987-6450
<http://www.cytogen.com>

Cytogen Research and Development
89 Bellevue Hill Road
Boston, MA 02132
(617) 325-7774 FAX: (617) 327-2405

CytRx
154 Technology Parkway
Norcross, GA 30092
(800) 345-2987 FAX: (770) 368-0622
(770) 368-9500
<http://www.cytrx.com>

Dade Behring
Corporate Headquarters
1717 Deerfield Road
Deerfield, IL 60015
(847) 267-5300 FAX: (847) 267-1066
<http://www.dadebehring.com>

Dagan
2855 Park Avenue
Minneapolis, MN 55407
(612) 827-5959 FAX: (612) 827-6535
<http://www.dagan.com>

Dako
6392 Via Real
Carpinteria, CA 93013
(800) 235-5763 FAX: (805) 566-6688
(805) 566-6655
<http://www.dakousa.com>

Dako A/S
42 Produktionsvej
P.O. Box 1359
DK-2600 Glostrup, Denmark
(45) 4492-0044 FAX: (45) 4284-1822

Dakopatts
See Dako A/S

Dalton Chemical Laboratoris
349 Wildcat Road
Toronto, Ontario
M3J 2S3 Canada
(416) 661-2102 FAX: (416) 661-2108
(800) 567-5060 (in Canada only)
<http://www.dalton.com>

Damon, IEC
See Thermoquest

Dan Kar Scientific
150 West Street
Wilmington, MA 01887
(800) 942-5542 FAX: (978) 658-0380
(978) 988-9696
<http://www.dan-kar.com>

DataCell
Falcon Business Park
40 Ivanhoe Road
Finchampstead, Berkshire
RG40 4QQ, UK
(44) 1189 324324
FAX: (44) 1189 324325
<http://www.datacell.co.uk>
In the US:
(408) 446-3575 FAX: (408) 446-3589
<http://www.datacell.com>

DataWave Technologies
380 Main Street, Suite 209
Longmont, CO 80501
(800) 736-9283 FAX: (303) 776-8531
(303) 776-8214

Datex-Ohmeda
3030 Ohmeda Drive
Madison, WI 53718
(800) 345-2700 FAX: (608) 222-9147
(608) 221-1551
<http://www.us.datex-ohmeda.com>

DATU
82 State Street
Geneva, NY 14456
(315) 787-2240 FAX: (315) 787-2397
<http://www.nysaes.cornell.edu/datu>

David Kopf Instruments
7324 Elmo Street
P.O. Box 636
Tujunga, CA 91043
(818) 352-3274 FAX: (818) 352-3139

Decagon Devices
P.O. Box 835
950 NE Nelson Court
Pullman, WA 99163
(800) 755-2751 FAX: (509) 332-5158
(509) 332-2756
<http://www.decagon.com>

Decon Labs
890 Country Line Road
Bryn Mawr, PA 19010
(800) 332-6647 FAX: (610) 964-0650
(610) 520-0610
<http://www.deconlabs.com>

Decon Laboratories
Conway Street
Hove, Sussex BN3 3LY, UK
(44) 1273 739241
FAX: (44) 1273 722088

Degussa
Precious Metals Division
3900 South Clinton Avenue
South Plainfield, NJ 07080
(800) DEGUSSA FAX: (908) 756-7176
(908) 561-1100
<http://www.degussa-huls.com>

Deneba Software
1150 NW 72nd Avenue
Miami, FL 33126
(305) 596-5644 FAX: (305) 273-9069
<http://www.deneba.com>

Deseret Medical
524 West 3615 South
Salt Lake City, UT 84115
(801) 270-8440 FAX: (801) 293-9000

Devcon Plexus
30 Endicott Street
Danvers, MA 01923
(800) 626-7226 FAX: (978) 774-0516
(978) 777-1100
<http://www.devcon.com>

Developmental Studies Hybridoma Bank
University of Iowa
436 Biology Building
Iowa City, IA 52242
(319) 335-3826 FAX: (319) 335-2077
<http://www.uiowa.edu/~dshbwww>

DeVilbiss
Division of Sunrise Medical Respiratory
100 DeVilbiss Drive
P.O. Box 635
Somerset, PA 15501
(800) 338-1988 FAX: (814) 443-7572
(814) 443-4881
<http://www.sunrisemedical.com>

Dharmacon Research
1376 Miners Drive #101
Lafayette, CO 80026
(303) 604-9499 FAX: (303) 604-9680
<http://www.dharmacon.com>

DiaChem
Triangle Biomedical
Gardiners Place
West Gillibrands, Lancashire
WN8 9SP, UK
(44) 1695-555581
FAX: (44) 1695-555518
<http://www.diachem.co.uk>

Suppliers

A.4.10

Diagen

Max-Volmer Strasse 4
D-40724 Hilden, Germany
(49) 2103-892-230
FAX: (49) 2103-892-222

Diagnostic Concepts

6104 Madison Court
Morton Grove, IL 60053
(847) 604-0957

Diagnostic Developments

See DiaCheM

Diagnostic Instruments

6540 Burroughs
Sterling Heights, MI 48314
(810) 731-6000 FAX: (810) 731-6469
<http://www.diaginc.com>

Diamedix

2140 North Miami Avenue
Miami, FL 33127
(800) 327-4565 FAX: (305) 324-2395
(305) 324-2300

DiaSorin

1990 Industrial Boulevard
Stillwater, MN 55082
(800) 328-1482 FAX: (651) 779-7847
(651) 439-9719
<http://www.diasorin.com>

Diatome US

321 Morris Road
Fort Washington, PA 19034
(800) 523-5874 FAX: (215) 646-8931
(215) 646-1478
<http://www.emsdiasum.com>

Difco Laboratories

See Becton Dickinson

Digene

1201 Clopper Road
Gaithersburg, MD 20878
(301) 944-7000 (800) 344-3631
FAX: (301) 944-7121
<http://www.digene.com>

Digi-Key

701 Brooks Avenue South
Thief River Falls, MN 56701
(800) 344-4539 FAX: (218) 681-3380
(218) 681-6674
<http://www.digi-key.com>

Digitimer

37 Hydeway
Welwyn Garden City, Hertfordshire
AL7 3BE, UK
(44) 1707-328347
FAX: (44) 1707-373153
<http://www.digitimer.com>

Dimco-Gray

8200 South Suburban Road
Dayton, OH 45458
(800) 876-8353 FAX: (937) 433-0520
(937) 433-7600
<http://www.dimco-gray.com>

Dionex

1228 Titan Way
P.O. Box 3603
Sunnyvale, CA 94088
(408) 737-0700 FAX: (408) 730-9403
<http://dionex2.promptu.com>

Display Systems Biotech

1260 Liberty Way, Suite B
Vista, CA 92083
(800) 697-1111 FAX: (760) 599-9930
(760) 599-0598
<http://www.displaysystems.com>

Diversified Biotech

1208 VFW Parkway
Boston, MA 02132
(617) 965-8557 FAX: (617) 323-5641
(800) 796-9199
<http://www.divbio.com>

DNA ProScan

P.O. Box 121585
Nashville, TN 37212
(800) 841-4362 FAX: (615) 292-1436
(615) 298-3524
<http://www.dnapro.com>

DNAStar

1228 South Park Street
Madison, WI 53715
(608) 258-7420 FAX: (608) 258-7439
<http://www.dnastar.com>

DNAVIEW

Attn: Charles Brenner
<http://www.wco.com>
~cbrenner/dnaview.htm

Doall NYC

36-06 48th Avenue
Long Island City, NY 11101
(718) 392-4595 FAX: (718) 392-6115
<http://www.doall.com>

Dojindo Molecular Technologies

211 Perry Street Parkway, Suite 5
Gaithersburg, MD 20877
(877) 987-2667
<http://www.dojindo.com>

Dolla Eastern

See Doall NYC

Dolan Jenner Industries

678 Andover Street
Lawrence, MA 08143
(978) 681-8000 (978) 682-2500
<http://www.dolan-jenner.com>

Dow Chemical

Customer Service Center
2040 Willard H. Dow Center
Midland, MI 48674
(800) 232-2436 FAX: (517) 832-1190
(409) 238-9321
<http://www.dow.com>

Dow Corning

Northern Europe
Meriden Business Park
Copse Drive
Allesley, Coventry CV5 9RG, UK
(44) 1676 528 000
FAX: (44) 1676 528 001

Dow Corning

P.O. Box 994
Midland, MI 48686
(517) 496-4000
<http://www.dowcorning.com>

Dow Corning (Lubricants)

2200 West Salzburg Road
Auburn, MI 48611
(800) 248-2481 FAX: (517) 496-6974
(517) 496-6000

Dremel

4915 21st Street
Racine, WI 53406
(414) 554-1390
<http://www.dremel.com>

Drummond Scientific

500 Parkway
P.O. Box 700
Broomall, PA 19008
(800) 523-7480 FAX: (610) 353-6204
(610) 353-0200
<http://www.drummondsci.com>

Duchefa Biochemie BV

P.O. Box 2281
2002 CG Haarlem, The Netherlands
31-0-23-5319093
FAX: 31-0-23-5318027
<http://www.duchefa.com>

Duke Scientific

2463 Faber Place
Palo Alto, CA 94303
(800) 334-3883 FAX: (650) 424-1158
(650) 424-1177
<http://www.dukescientific.com>

Duke University Marine Laboratory

135 Duke Marine Lab Road
Beaufort, NC 28516-9721
(252) 504-7503 FAX: (252) 504-7648
<http://www.env.duke.edu/marinelab>

DuPont Biotechnology Systems

See NEN Life Science Products

DuPont Medical Products

See NEN Life Science Products

DuPont Merck Pharmaceuticals

331 Treble Cove Road
Billerica, MA 01862
(800) 225-1572 FAX: (508) 436-7501
<http://www.dupontmerck.com>

DuPont NEN Products

See NEN Life Science Products

Dynal

5 Delaware Drive
Lake Success, NY 11042
(800) 638-9416 FAX: (516) 326-3298
(516) 326-3270
<http://www.dynal.net>

Dynal AS

Ullemchausen 52,
0379 Oslo, Norway
47-22-06-10-00 FAX: 47-22-50-70-15
<http://www.dynal.no>

Dynalab

P.O. Box 112
Rochester, NY 14692
(800) 828-6595 FAX: (716) 334-9496
(716) 334-2060
<http://www.dynalab.com>

Dynarex

1 International Boulevard
Brewster, NY 10509
(888) DYNAREX FAX: (914) 279-9601
(914) 279-9600
<http://www.dynarex.com>

Dynatech

See Dynex Technologies

Dynex Technologies

14340 Sullyfield Circle
Chantilly, VA 22021
(800) 336-4543 FAX: (703) 631-7816
(703) 631-7800
<http://www.dynex technologies.com>

Dyno Mill

See Willy A. Bachofen

E.S.A.

22 Alpha Road
Chelmsford, MA 01824
(508) 250-7000 FAX: (508) 250-7090

E.W. Wright

760 Durham Road
Guilford, CT 06437
(203) 453-6410 FAX: (203) 458-6901
<http://www.ewwright.com>

E-Y Laboratories

107 N. Amphlett Boulevard
San Mateo, CA 94401
(800) 821-0044 FAX: (650) 342-2648
(650) 342-3296
<http://www.eylabs.com>

Eastman Kodak

1001 Lee Road
Rochester, NY 14650
(800) 225-5352 FAX: (800) 879-4979
(716) 722-5780 FAX: (716) 477-8040
<http://www.kodak.com>

ECACC

See European Collection of Animal Cell Cultures

EC Apparatus

See Savant/EC Apparatus

Suppliers**A.4.11**

Ecogen, SRL
Gensura Laboratories
Ptge. Dos de Maig
9(08041) Barcelona, Spain
(34) 3-450-2601 FAX: (34) 3-456-0607
<http://www.ecogen.com>

Ecolab
370 North Wabasha Street
St. Paul, MN 55102
(800) 35-CLEAN FAX: (651) 225-3098
(651) 352-5326
<http://www.ecolab.com>

ECO PHYSICS
3915 Research Park Drive, Suite A-3
Ann Arbor, MI 48108
(734) 998-1600 FAX: (734) 998-1180
<http://www.ecophysics.com>

Edge Biosystems
19208 Orbit Drive
Gaithersburg, MD 20879-4149
(800) 326-2685 FAX: (301) 990-0881
(301) 990-2685
<http://www.edgebio.com>

Edmund Scientific
101 E. Gloucester Pike
Barrington, NJ 08007
(800) 728-6999 FAX: (856) 573-6263
(856) 573-6250
<http://www.edsci.com>

EG&G
See Perkin-Elmer

Ekagen
969 C Industry Road
San Carlos, CA 94070
(650) 592-4500 FAX: (650) 592-4500

Elcatech
P.O. Box 10935
Winston-Salem, NC 27108
(336) 544-8613 FAX: (336) 777-3623
(910) 777-3624
<http://www.elcatech.com>

Electro Impulse, Inc.
1805 Rt. 33 (Corlies Ave.)
Neptune, NJ 07754
(732) 776-5800 Fax: 732-776-6793
<http://www.electro-impulse.com>

Electron Microscopy Sciences
321 Morris Road
Fort Washington, PA 19034
(800) 523-5874 FAX: (215) 646-8931
(215) 646-1566
<http://www.emsdiasum.com>

Electron Tubes
100 Forge Way, Unit F
Rockaway, NJ 07866
(800) 521-8382 FAX: (973) 586-9771
(973) 586-9594
<http://www.electrontubes.com>

Elicay Laboratory Products, (UK) Ltd.
4 Manborough Mews
Crockford Lane
Basingstoke, Hampshire
RG 248NA, England
(256) 811-118 FAX: (256) 811-116
<http://www.elkay-uk.co.uk>

Eli Lilly
Lilly Corporate Center
Indianapolis, IN 46285
(800) 545-5979 FAX: (317) 276-2095
(317) 276-2000
<http://www.lilly.com>

ELISA Technologies
See Neogen

Elkins-Sinn
See Wyeth-Ayerst

EMBI
See European Bioinformatics Institute

EM Science
480 Democrat Road
Gibbstown, NJ 08027
(800) 222-0342 FAX: (856) 423-4389
(856) 423-6300
<http://www.emscience.com>

EM Separations Technology
See R & S Technology

Endogen
30 Commerce Way
Woburn, MA 01801
(800) 487-4885 FAX: (617) 439-0355
(781) 937-0890
<http://www.endogen.com>

ENGEL-Loter
HSGM Heatcutting Equipment
& Machines
1865 E. Main Street, No. 5
Duncan, SC 29334
(888) 854-HSGM FAX: (864) 486-8383
(864) 486-8300
<http://www.engelgmbh.com>

Enzo Diagnostics
60 Executive Boulevard
Farmingdale, NY 11735
(800) 221-7705 FAX: (516) 694-7501
(516) 694-7070
<http://www.enzo.com>

Enzogenetics
4197 NW Douglas Avenue
Corvallis, OR 97330
(541) 757-0288

The Enzyme Center
See Charm Sciences

Enzyme Systems Products
486 Lindbergh Avenue
Livermore, CA 94550
(888) 449-2664 FAX: (925) 449-1866
(925) 449-2664
<http://www.enzymesys.com>

Epicentre Technologies
1402 Emil Street
Madison, WI 53713
(800) 284-8474 FAX: (608) 258-3088
(608) 258-3080
<http://www.epicentre.com>

Erie Scientific
20 Post Road
Portsmouth, NH 03801
(888) ERIE-SCI FAX: (603) 431-8996
(603) 431-8410
<http://www.eriesci.com>

ES Industries
701 South Route 73
West Berlin, NJ 08091
(800) 356-6140 FAX: (856) 753-8484
(856) 753-8400
<http://www.esind.com>

ESA
22 Alpha Road
Chelmsford, MA 01824
(800) 959-5095 FAX: (978) 250-7090
(978) 250-7000
<http://www.esainc.com>

Ethicon
Route 22, P.O. Box 151
Somerville, NJ 08876
(908) 218-0707
<http://www.ethiconinc.com>

Ethicon Endo-Surgery
4545 Creek Road
Cincinnati, OH 45242
(800) 766-9534 FAX: (513) 786-7080

Eurogentec
Parc Scientifique du Sart Tilman
4102 Seraing, Belgium
32-4-240-76-76 FAX: 32-4-264-07-88
<http://www.eurogentec.com>

European Bioinformatics Institute
Wellcome Trust Genomes Campus
Hinxton, Cambridge CB10 1SD, UK
(44) 1223-49444
FAX: (44) 1223-494468

**European Collection of Animal
Cell Cultures (ECACC)**
Centre for Applied Microbiology &
Research
Salisbury, Wiltshire SP4 0JG, UK
(44) 1980-612 512
FAX: (44) 1980-611 315
<http://www.camr.org.uk>

Evergreen Scientific
2254 E. 49th Street
P.O. Box 58248
Los Angeles, CA 90058
(800) 421-6261 FAX: (323) 581-2503
(323) 583-1331
<http://www.evergreensci.com>

Exalpha Biologicals
20 Hampden Street
Boston, MA 02205
(800) 395-1137 FAX: (617) 969-3872
(617) 558-3625
<http://www.exalpha.com>

Exciton
P.O. Box 31126
Dayton, OH 45437
(937) 252-2989 FAX: (937) 258-3937
<http://www.exciton.com>

Extrasynthese
ZI Lyon Nord
SA-BP62
69730 Genay, France
(33) 78-98-20-34
FAX: (33) 78-98-19-45

Factor II
1972 Forest Avenue
P.O. Box 1339
Lakeside, AZ 85929
(800) 332-8688 FAX: (520) 537-8066
(520) 537-8387
<http://www.factor2.com>

Falcon
See Becton Dickinson Labware

Febit AG
Käfertaler Strasse 190
D-68167 Mannheim
Germany
(49) 621-3804-0
FAX: (49) 621-3804-400
<http://www.febit.com>

Fenwal
See Baxter Healthcare

Filemaker
5201 Patrick Henry Drive
Santa Clara, CA 95054
(408) 987-7000 (800) 325-2747

Fine Science Tools
202-277 Mountain Highway
North Vancouver, British Columbia
V7J 3P2 Canada
(800) 665-5355 FAX: (800) 665 4544
(604) 980-2481 FAX: (604) 987-3299

Fine Science Tools
373-G Vintage Park Drive
Foster City, CA 94404
(800) 521-2109 FAX: (800) 523-2109
(650) 349-1636 FAX: (630) 349-3729

Fine Science Tools
Fahrtgasse 7-13
D-69117 Heidelberg, Germany
(49) 6221 905050
FAX: (49) 6221 600001
<http://www.finescience.com>

Finn Aqua
AMSCO Finn Aqua Oy
Teollisuustie, FIN-04300
Tuusula, Finland
358 025851 FAX: 358 0276019

Suppliers

A.4.12

Finnigan
355 River Oaks Parkway
San Jose, CA 95134
(408) 433-4800 FAX: (408) 433-4821
<http://www.finnigan.com>

Dr. L. Fischer
Lutherstrasse 25A
D-69120 Heidelberg
Germany
(49) 6221-16-0368
<http://home.eplus-online.de/electroporation>

Fisher Chemical Company
Fisher Scientific Limited
112 Colonnade Road
Nepean, Ontario K2E 7L6 Canada
(800) 234-7437 FAX: (800) 463-2996
<http://www.fisherscientific.com>

Fisher Scientific
2000 Park Lane
Pittsburgh, PA 15275
(800) 766-7000 FAX: (800) 926-1166
(412) 562-8300
<http://www3.fishersci.com>

W.F. Fisher & Son
220 Evans Way, Suite #1
Somerville, NJ 08876
(908) 707-4050 FAX: (908) 707-4099

Fitzco
5600 Pioneer Creek Drive
Maple Plain, MN 55359
(800) 367-8760 FAX: (612) 479-2880
(612) 479-3489
<http://www.fitzco.com>

5 Prime → 3 Prime
See 2000 Eppendorf-5 Prime
<http://www.5prime.com>

Flambeau
15981 Valplast Road
Middlefield, Ohio 44062
(800) 232-3474 FAX: (440) 632-1581
(440) 632-1631
<http://www.flambeau.com>

Fleisch (Rusch)
2450 Meadowbrook Parkway
Duluth, GA 30096
(770) 623-0816 FAX: (770) 623-1829
<http://ruschinc.com>

Flow Cytometry Standards
P.O. Box 194344
San Juan, PR 00919
(800) 227-8143 FAX: (787) 758-3267
(787) 753-9341
<http://www.fcstd.com>

Flow Labs
See ICN Biomedicals

Flow-Tech Supply
P.O. Box 1388
Orange, TX 77631
(409) 882-0306 FAX: (409) 882-0254
<http://www.flow-tech.com>

Fluid Marketing
See Fluid Metering

Fluid Metering
5 Aerial Way, Suite 500
Sayosett, NY 11791
(516) 922-6050 FAX: (516) 624-8261
<http://www.fmipump.com>

Fluorochrome
1801 Williams, Suite 300
Denver, CO 80264
(303) 394-1000 FAX: (303) 321-1119

Fluka Chemical
See Sigma-Aldrich

FMC BioPolymer
1735 Market Street
Philadelphia, PA 19103
(215) 299-6000 FAX: (215) 299-5809
<http://www.fmc.com>

FMC BioProducts
191 Thomaston Street
Rockland, ME 04841
(800) 521-0390 FAX: (800) 362-1133
(207) 594-3400 FAX: (207) 594-3426
<http://www.bioproducts.com>

Forma Scientific
Milcreek Road
P.O. Box 649
Marietta, OH 45750
(800) 848-3080 FAX: (740) 372-6770
(740) 373-4765
<http://www.forma.com>

Fort Dodge Animal Health
800 5th Street NW
Fort Dodge, IA 50501
(800) 685-5656 FAX: (515) 955-9193
(515) 955-4600
<http://www.ahp.com>

Fotodyne
950 Walnut Ridge Drive
Hartland, WI 53029
(800) 362-3686 FAX: (800) 362-3642
(262) 369-7000 FAX: (262) 369-7013
<http://www.fotodyne.com>

Fresenius HemoCare
6675 185th Avenue NE, Suite 100
Redwood, WA 98052
(800) 909-3872
(425) 497-1197
<http://www.freseniusht.com>

Fresenius Hemotechnology
See Fresenius HemoCare

Fuji Medical Systems
419 West Avenue
P.O. Box 120035
Stamford, CT 06902
(800) 431-1850 FAX: (203) 353-0926
(203) 324-2000
<http://www.fujimed.com>

Fujisawa USA
Parkway Center North
Deerfield, IL 60015-2548
(847) 317-1088 FAX: (847) 317-7298

Ernest F. Fullam
900 Albany Shaker Road
Latham, NY 12110
(800) 833-4024 FAX: (518) 785-8647
(518) 785-5533
<http://www.fullam.com>

Gallard-Schlesinger Industries
777 Zechendorf Boulevard
Garden City, NY 11530
(516) 229-4000 FAX: (516) 229-4015
<http://www.gallard-schlesinger.com>

Gambro
Box 7373
SE 103 91 Stockholm, Sweden
(46) 8 613 65 00
FAX: (46) 8 611 37 31

In the US: **COBE Laboratories**
225 Union Boulevard
Lakewood, CO 80215
(303) 232-6800 FAX: (303) 231-4915
<http://www.gambro.com>

Garner Glass
177 Indian Hill Boulevard
Claremont, CA 91711
(909) 624-5071 FAX: (909) 625-0173
<http://www.garnerglass.com>

Garon Plastics
16 Byre Avenue
Somerton Park, South Australia 5044
(08) 8294-5126 FAX: (08) 8376-1487
<http://www.apache.airnet.com.au/~garon>

Garren Scientific
9400 Lurline Avenue, Unit E
Chatsworth, CA 91311
(800) 342-3725 FAX: (818) 882-3229
(818) 882-6544
<http://www.garren-scientific.com>

GATC Biotech AG
Jakob-Stadler-Platz 7
D-78467 Constance, Germany
(49) 07531-8160-0
FAX: (49) 07531-8160-81
<http://www.gatc-biotech.com>

Gaussian
Carnegie Office Park
Building 6, Suite 230
Carnegie, PA 15106
(412) 279-6700 FAX: (412) 279-2118
<http://www.gaussian.com>

G.C. Electronics/A.R.C. Electronics
431 Second Street
Henderson, KY 42420
(270) 827-8981 FAX: (270) 827-8256
<http://www.arcelectronics.com>

GDB (Genome Data Base, Curation)
2024 East Monument Street, Suite 1200
Baltimore, MD 21205
(410) 955-9705 FAX: (410) 614-0434
<http://www.gdb.org>

GDB (Genome Data Base, Home)
Hospital for Sick Children
555 University Avenue
Toronto, Ontario
M5G 1X8 Canada
(416) 813-8744 FAX: (416) 813-8755
<http://www.gdb.org>

Gelman Sciences
See Pall-Gelman

Gemini BioProducts
5115-M Douglas Fir Road
Calabasas, CA 90403
(818) 591-3530 FAX: (818) 591-7084

Gen Trak
5100 Campus Drive
Plymouth Meeting, PA 19462
(800) 221-7407 FAX: (215) 941-9498
(215) 825-5115
<http://www.informagen.com>

Genaisance Pharmaceuticals
5 Science Park
New Haven, CT 06511
(800) 678-9487 FAX: (203) 562-9377
(203) 773-1450
<http://www.genaisance.com>

GENAXIS Biotechnology
Parc Technologique
10 Avenue Ampère
Montigny le Bretonneux
78180 France
(33) 01-30-14-00-20
FAX: (33) 01-30-14-00-15
<http://www.genaxis.com>

GenBank
National Center for Biotechnology
Information
National Library of Medicine/NIH
Building 38A, Room 8N805
8600 Rockville Pike
Bethesda, MD 20894
(301) 496-2475 FAX: (301) 480-9241
<http://www.ncbi.nlm.nih.gov>

Gene Codes
640 Avis Drive
Ann Arbor, MI 48108
(800) 497-4939 FAX: (734) 930-0145
(734) 769-7249
<http://www.genecodes.com>

Genemachines
935 Washington Street
San Carlos, CA 94070
(650) 508-1634 FAX: (650) 508-1644
(877) 855-4363
<http://www.genemachines.com>

Genentech
1 DNA Way
South San Francisco, CA 94080
(800) 551-2231 FAX: (650) 225-1600
(650) 225-1000
<http://www.gene.com>

Suppliers

A.4.13

General Scanning/GSI Luminomics
500 Arsenal Street
Watertown, MA 02172
(617) 924-1010 FAX: (617) 924-7327
<http://www.genescan.com>

General Valve
Division of Parker Hannifin Pneutronics
19 Gloria Lane
Fairfield, NJ 07004
(800) GVC-VALV
FAX: (800) GVC-1-FAX
<http://www.pneutronics.com>

Genespan
19310 North Creek Parkway, Suite 100
Bothell, WA 98011
(800) 231-2215 FAX: (425) 482-3005
(425) 482-3003
<http://www.genespan.com>

Gene Therapy Systems
10190 Telesis Court
San Diego, CA 92122
(858) 457-1919 FAX: (858) 623-9494
<http://www.genetherapysystems.com>

Généthon Human Genome
Research Center
1 bis rue de l'Internationale
91000 Evry, France
(33) 169-472828
FAX: (33) 607-78698
<http://www.genethon.fr>

Genetic Microsystems
34 Commerce Way
Woburn, MA 01801
(781) 932-9333 FAX: (781) 932-9433
<http://www.genticmicro.com>

Genetic Mutant Repository
See Coriell Institute for Medical Research

Genetic Research Instrumentation
Gene House
Queenborough Lane
Rayne, Braintree, Essex CM7 8TF, UK
(44) 1376 332900
FAX: (44) 1376 344724
<http://www.gri.co.uk>

Genetics Computer Group
575 Science Drive
Madison, WI 53711
(608) 231-5200 FAX: (608) 231-5202
<http://www.gcg.com>

Genetics Institute/American Home Products
87 Cambridge Park Drive
Cambridge, MA 02140
(617) 876-1170 FAX: (617) 876-0388
<http://www.genetics.com>

Genetix
63-69 Somerford Road
Christchurch, Dorset BH23 3QA, UK
(44) (0) 1202 483900
FAX: (44)(0) 1202 480289
In the US: (877) 436 3849
US FAX: (888) 522 7499
<http://www.genetix.co.uk>

Gene Tools
One Summertown Way
Philomath, OR 97370
(541) 9292-7840 FAX: (541) 9292-7841
<http://www.gene-tools.com>

Geneva Bioinformatics (GeneBio) S.A.
25 Avenue de Champel
CH—1206 Geneva, Switzerland
(41) 22-702-9900
FAX: (41) 22-702-9999
<http://www.genebio.com>

GeneWorks
P.O. Box 11, Rundle Mall
Adelaide, South Australia 5000, Australia
1800 882 555 FAX: (08) 8234 2699
(08) 8234 2644
<http://www.geneworks.com>

Genome Systems (INCYTE)
4633 World Parkway Circle
St. Louis, MO 63134
(800) 430-0030 FAX: (314) 427-3324
(314) 427-3222
<http://www.genomesystems.com>

Genomic Solutions
4355 Varsity Drive, Suite E
Ann Arbor, MI 48108
(877) GENOMIC FAX: (734) 975-4808
(734) 975-4800
<http://www.genomicsolutions.com>

Genomyx
See Beckman Coulter

Genosys Biotechnologies
1442 Lake Front Circle, Suite 185
The Woodlands, TX 77380
(281) 363-3693 FAX: (281) 363-2212
<http://www.genosys.com>

Genotech
92 Weldon Parkway
St. Louis, MO 63043
(800) 628-7730 FAX: (314) 991-1504
(314) 991-6034

GENSET
876 Prospect Street, Suite 206
La Jolla, CA 92037
(800) 551-5291 FAX: (619) 551-2041
(619) 515-3061
<http://www.genset.fr>

Gensia Laboratories Ltd.
19 Hughes
Irvine, CA 92718
(714) 455-4700 FAX: (714) 855-8210

Genta
99 Hayden Avenue, Suite 200
Lexington, MA 02421
(781) 860-5150 FAX: (781) 860-5137
<http://www.genta.com>

GENTEST
6 Henshaw Street
Woburn, MA 01801
(800) 334-5229 FAX: (888) 242-2226
(781) 935-5115 FAX: (781) 932-6855
<http://www.gentest.com>

Genra Systems
15200 25th Avenue N., Suite 104
Minneapolis, MN 55447
(800) 866-3039 FAX: (612) 476-5850
(612) 476-5858
<http://www.genra.com>

Genzyme
1 Kendall Square
Cambridge, MA 02139
(617) 252-7500 FAX: (617) 252-7600
<http://www.genzyme.com>
See also R&D Systems

Genzyme Genetics
One Mountain Road
Framingham, MA 01701
(800) 255-7357 FAX: (508) 872-9080
(508) 872-8400
<http://www.genzyme.com>

George Tiemann & Co.
25 Plant Avenue
Hauppauge, NY 11788
(516) 273-0005 FAX: (516) 273-6199

GIBCO/BRL
A Division of Life Technologies
1 Kendall Square
Grand Island, NY 14072
(800) 874-4226 FAX: (800) 352-1968
(716) 774-6700
<http://www.lifetech.com>

Gilmont Instruments
A Division of Barnant Company
28N092 Commercial Avenue
Barrington, IL 60010
(800) 637-3739 FAX: (708) 381-7053
<http://barnant.com>

Gilson
3000 West Beltline Highway
P.O. Box 620027
Middletown, WI 53562
(800) 445-7661
(608) 836-1551
<http://www.gilson.com>

Glas-Col Apparatus
P.O. Box 2128
Terre Haute, IN 47802
(800) Glas-Col FAX: (812) 234-6975
(812) 235-6167
<http://www.glascol.com>

Glaxo Wellcome
Five Moore Drive
Research Triangle Park, NC 27709
(800) SGL-AXO5 FAX: (919) 248-2386
(919) 248-2100
<http://www.glaxowellcome.com>

Glen Mills
395 Allwood Road
Clifton, NJ 07012
(973) 777-0777 FAX: (973) 777-0070
<http://www.glenmills.com>

Glen Research
22825 Davis Drive
Sterling, VA 20166
(800) 327-4536 FAX: (800) 934-2490
(703) 437-6191 FAX: (703) 435-9774
<http://www.glenresearch.com>

Glo Germ
P.O. Box 189
Moab, UT 84532
(800) 842-6622 FAX: (435) 259-5930
<http://www.glogerm.com>

Glyco
11 Pimentel Court
Novato, CA 94949
(800) 722-2597 FAX: (415) 382-3511
(415) 884-6799
<http://www.glyco.com>

Gould Instrument Systems
8333 Rockside Road
Valley View, OH 44125
(216) 328-7000 FAX: (216) 328-7400
<http://www.gould13.com>

Gralab Instruments
See Dimco-Gray

GraphPad Software
5755 Oberlin Drive #110
San Diego, CA 92121
(800) 388-4723 FAX: (558) 457-8141
(558) 457-3909
<http://www.graphpad.com>

Graseby Anderson
See Andersen Instruments
<http://www.graseby.com>

Grass Instrument
A Division of Astro-Med
600 East Greenwich Avenue
W. Warwick, RI 02893
(800) 225-5167 FAX: (877) 472-7749
<http://www.grassinstruments.com>

Greenacre and Misac Instruments
Misac Systems
27 Port Wood Road
Ware, Hertfordshire SF12 9NJ, UK
(44) 1920 463017
FAX: (44) 1920 465136

Greer Labs
639 Nuway Circle
Lenoir, NC 28645
(704) 754-5237
<http://greerlabs.com>

Greiner
Maybachstrasse 2
Postfach 1162
D-7443 Frickenhausen, Germany
(49) 0 91 31/80 79 0
FAX: (49) 0 91 31/80 79 30
<http://www.erlangen.com/greiner>

GSI Lumonics
130 Lombard Street
Oxnard, CA 93030
(805) 485-5559 FAX: (805) 485-3310
<http://www.gsilumonics.com>

Suppliers

A.4.14

GTE Internetworking
150 Cambridge Park Drive
Cambridge, MA 02140
(800) 472-4565 FAX: (508) 694-4861
<http://www.bbn.com>

GW Instruments
35 Medford Street
Somerville, MA 02143
(617) 625-4096 FAX: (617) 625-1322
<http://www.gwinst.com>

H & H Woodworking
1002 Garfield Street
Denver, CO 80206
(303) 394-3764

Hacker Instruments
17 Sherwood Lane
P.O. Box 10033
Fairfield, NJ 07004
800-442-2537 FAX: (973) 808-8281
(973) 226-8450
<http://www.hackerinstruments.com>

Haemenetics
400 Wood Road
Braintree, MA 02184
(800) 225-5297 FAX: (781) 848-7921
(781) 848-7100
<http://www.haemenetics.com>

Halocarbon Products
P.O. Box 661
River Edge, NJ 07661
(201) 242-8899 FAX: (201) 262-0019
<http://halocarbon.com>

Hamamatsu Photonic Systems
A Division of Hamamatsu
360 Foothill Road
P.O. Box 6910
Bridgewater, NJ 08807
(908) 231-1116 FAX: (908) 231-0852
<http://www.photonicsonline.com>

Hamilton Company
4970 Energy Way
P.O. Box 10030
Reno, NV 89520
(800) 648-5950 FAX: (775) 856-7259
(775) 858-3000
<http://www.hamiltoncompany.com>

Hamilton Thorne Biosciences
100 Cummings Center, Suite 102C
Beverly, MA 01915
<http://www.hamiltonthorne.com>

Hampton Research
27631 El Lazo Road
Laguna Niguel, CA 92677
(800) 452-3899 FAX: (949) 425-1611
(949) 425-6321
<http://www.hamptonresearch.com>

Harlan Bioproducts for Science
P.O. Box 29176
Indianapolis, IN 46229
(317) 894-7521 FAX: (317) 894-1840
<http://www.hbps.com>

Harlan Sera-Lab
Hillcrest, Dodgeford Lane
Belton, Loughborough
Leicester LE12 9TE, UK
(44) 1530 222123
FAX: (44) 1530 224970
<http://www.harlan.com>

Harlan Teklad
P.O. Box 44220
Madison, WI 53744
(608) 277-2070 FAX: (608) 277-2066
<http://www.harlan.com>

Harrick Scientific Corporation
88 Broadway
Ossining, NY 10562
(914) 762-0020 FAX: (914) 762-0914
<http://www.harricksci.com>

Harrison Research
840 Moana Court
Palo Alto, CA 94306
(650) 949-1565 FAX: (650) 948-0493

Harvard Apparatus
84 October Hill Road
Holliston, MA 01746
(800) 272-2775 FAX: (508) 429-5732
(508) 893-8999
<http://harvardapparatus.com>

Harvard Bioscience
See Harvard Apparatus

Haselton Biologics
See JRH Biosciences

Hazelton Research Products
See Covance Research Products

Health Products
See Pierce Chemical

Heat Systems-Ultrasonics
1938 New Highway
Farmingdale, NY 11735
(800) 645-9846 FAX: (516) 694-9412
(516) 694-9555

Heidenhain Corp
333 East State Parkway
Schaumburg, IL 60173
(847) 490-1191 FAX: (847) 490-3931
<http://www.heidenhain.com>

HEKA Instruments
33 Valley Rd.
Southboro, MA 01960
(866) 742-0606 FAX: (508) 481-8945
www.heka.com

Hellma Cells
11831 Queens Boulevard
Forest Hills, NY 11375
(718) 544-9166 FAX: (718) 263-6910
<http://www.hellmaUSA.com>

Hellma
Postfach 1163
D-79371 Müllheim/Baden, Germany
(49) 7631-1820
FAX: (49) 7631-13546
<http://www.hellma-worldwide.de>

Henry Schein
135 Duryea Road, Mail Room 150
Melville, NY 11747
(800) 472-4346 FAX: (516) 843-5652
<http://www.henryschein.com>

Heraeus Amersil
Quartzstrasse 8
Hanau D-63450, Germany
(49) 61 8135 50
(49) 61 813535 50
<http://www.heraeus-amersil.com>

Heraeus Kulzer
4315 South Lafayette Boulevard
South Bend, IN 46614
(800) 343-5336
(219) 291-0661
<http://www.kulzer.com>

Heraeus Sepatech
See Kendro Laboratory Products

Hercules Aqualon
Aqualon Division
Hercules Research Center, Bldg. 8145
500 Hercules Road
Wilmington, DE 19899
(800) 345-0447 FAX: (302) 995-4787
<http://www.herc.com/aqualon/pharma>

Heto-Holten A/S
Gydevang 17-19
DK-3450 Allerød, Denmark
(45) 48-16-62-00
FAX: (45) 48-16-62-97
Distributed by ATR

Hettich-Zentrifugen
See Andreas Hettich

Hewlett-Packard
3000 Hanover Street
Mailstop 20B3
Palo Alto, CA 94304
(650) 857-1501 FAX: (650) 857-5518
<http://www.hp.com>

HGS Hinimoto Plastics
1-10-24 Meguro-Honcho
Meguro-ku
Tokyo 152, Japan
3-3714-7226 FAX: 3-3714-4657

Hitachi Scientific Instruments
Nissei Sangyo America
8100 N. First Street
San Elsa, CA 95314
(800) 548-9001 FAX: (408) 432-0704
(408) 432-0520
<http://www.hii.hitachi.com>

Hi-Tech Scientific
Brunel Road
Salisbury, Wiltshire, SP2 7PU
UK
(44) 1722-432320
(800) 344-0724 (US only)
<http://www.hi-techsci.co.uk>

Hoechst AG
See Aventis Pharmaceutical

Hofer Scientific Instruments
Division of Amersham-Pharmacia Biotech
800 Centennial Avenue
Piscataway, NJ 08855
(800) 227-4750 FAX: (877) 295-8102
<http://www.apbiotech.com>

Hoffman-LaRoche
340 Kingsland Street
Nutley, NJ 07110
(800) 526-0189 FAX: (973) 235-9605
(973) 235-5000
<http://www.rocheUSA.com>

Holborn Surgical and Medical Instruments
Westwood Industrial Estate
Ramsgate Road
Margate, Kent CT9 4JZ UK
(44) 1843 296666
FAX: (44) 1843 295446

Honeywell
101 Columbia Road
Morristown, NJ 07962
(973) 455-2000 FAX: (973) 455-4807
<http://www.honeywell.com>

Honeywell Specialty Films
P.O. Box 1039
101 Columbia Road
Morristown, NJ 07962
(800) 934-5679 FAX: (973) 455-6045
<http://www.honeywell-specialtyfilms.com>

Hood Thermo-Pad Canada
Comp. 20, Site 61A, RR2
Summerland, British Columbia
VOH 1Z0 Canada
(800) 665-9555 FAX: (250) 494-5003
(250) 494-5002
<http://www.thermopad.com>

Horiba Instruments
17671 Armstrong Avenue
Irvine, CA 92714
(949) 250-4811 FAX: (949) 250-0924
<http://www.horiba.com>

Hoskins Manufacturing
10776 Hall Road
P.O. Box 218
Hamburg, MI 48139
(810) 231-1900 FAX: (810) 231-4311
<http://www.hoskinsmfgco.com>

Hosokawa Micron Powder Systems
10 Chatham Road
Summit, NJ 07901
(800) 526-4491 FAX: (908) 273-7432
(908) 273-6360
<http://www.hosokawamicron.com>

HT Biotechnology
Unit 4
61 Ditton Walk
Cambridge CB5 8QD, UK
(44) 1223-412583

Suppliers

A.4.15

Hugo Sachs Elektronik
Postfach 138
7806 March-Hugstetten, Germany
D-79229(49) 7665-92000
FAX: (49) 7665-920090

Human Biologics International
7150 East Camelback Road, Suite 245
Scottsdale, AZ 85251
(480) 990-2005 FAX: (480)-990-2155
<http://www.humanbiological.com>

Human Genetic Mutant Cell
Repository
See Coriell Institute for Medical Research

HVS Image
P.O. Box 100
Hampton, Middlesex TW12 2YD, UK
FAX: (44) 208 783 1223
In the US: (800) 225-9261
FAX: (888) 483-8033
<http://www.hvsimage.com>

Hybaid
111-113 Waldegrave Road
Teddington, Middlesex TW11 8LL, UK
(44) 0 1784 42500
FAX: (44) 0 1784 248085
<http://www.hybaid.co.uk>

Hybaid Instruments
8 East Forge Parkway
Franklin, MA 02028
(888)4-HYBAID FAX: (508) 541-3041
(508) 541-6918
<http://www.hybaid.com>

Hybridon
155 Fortune Boulevard
Milford, MA 01757
(508) 482-7500 FAX: (508) 482-7510
<http://www.hybridon.com>

HyClone Laboratories
1725 South HyClone Road
Logan, UT 84321
(800) HYCLONE FAX: (800) 533-9450
(801) 753-4584 FAX: (801) 750-0809
<http://www.hyclone.com>

Hyseq
670 Almanor Avenue
Sunnyvale, CA 94086
(408) 524-8100 FAX: (408) 524-8141
<http://www.hyseq.com>

IBA GmbH
1508 South Grand Blvd.
St. Louis, MO 63104
(877) 422-4624 FAX: (888) 531-6813
<http://www.iba-go.com>

IBF Biotechnics
See Sepracor

IBI (International Biotechnologies)
See Eastman Kodak
For technical service (800) 243-2555
(203) 786-5600

ICN Biochemicals
See ICN Biomedicals

ICN Biomedicals
3300 Hyland Avenue
Costa Mesa, CA 92626
(800) 854-0530 FAX: (800) 334-6999
(714) 545-0100 FAX: (714) 641-7275
<http://www.icnbiomed.com>

ICN Flow and Pharmaceuticals
See ICN Biomedicals

ICN Immunobiochemicals
See ICN Biomedicals

ICN Radiochemicals
See ICN Biomedicals

ICONIX
100 King Street West, Suite 3825
Toronto, Ontario
M5X 1E3 Canada
(416) 410-2411 FAX: (416) 368-3089
<http://www.iconix.com>

ICRT (Imperial Cancer Research
Technology)
Sardinia House
Sardinia Street
London WC2A 3NL, UK
(44) 1712-421136
FAX: (44) 1718-314991

Idea Scientific Company
P.O. Box 13210
Minneapolis, MN 55414
(800) 433-2535 FAX: (612) 331-4217
<http://www.ideascientific.com>

IEC
See International Equipment Co.

IITC
23924 Victory Boulevard
Woodland Hills, CA 91367
(888) 414-4482 (818) 710-1556
FAX: (818) 992-5185
<http://www.iitcinc.com>

IKA Works
2635 N. Chase Parkway, SE
Wilmington, NC 28405
(910) 452-7059 FAX: (910) 452-7693
<http://www.ika.net>

Ikegami Electronics
37 Brook Avenue
Maywood, NJ 07607
(201) 368-9171 FAX: (201) 569-1626

Ikemoto Scientific Technology
25-11 Hongo
3-chome, Bunkyo-ku
Tokyo 101-0025, Japan
(81) 3-3811-4181
FAX: (81) 3-3811-1960

Imagenetics
See ATC Diagnostics

Imaging Research
c/o Brock University
500 Glenridge Avenue
St. Catharines, Ontario
L2S 3A1 Canada
(905) 688-2040 FAX: (905) 685-5861
<http://www.imaging.brocku.ca>

Imclone Systems
180 Varick Street
New York, NY 10014
(212) 645-1405 FAX: (212) 645-2054
<http://www.imclone.com>

IMCO Corporation LTD., AB
P.O. Box 21195
SE-100 31
Stockholm, Sweden
46-8-33-53-09 FAX: 46-8-728-47-76
<http://www.imcocorp.se>

Imgenex Corporation
11175 Flintkote Avenue
Suite E
San Diego, CA 92121
(888) 723-4363 FAX: (858) 642-0937
(858) 642.0978
<http://www.imgenex.com>

IMICO
Calle Vivero, No. 5-4a Planta
E-28040, Madrid, Spain
(34) 1-535-3960 FAX: (34) 1-535-2780

Immunes
51 University Street
Seattle, WA 98101
(206) 587-0430 FAX: (206) 587-0606
<http://www.immunex.com>

Immunochemistry Technologies
9401 James Avenue, South
Suite 155
Bloomington, MN 55431
(800) 829-3194 FAX: (952) 888-8988
(952) 888-8788
<http://www.immunochemistry.com>

Immunocorp
1582 W. Deere Avenue
Suite C
Irvine, CA 92606
(800) 446-3063
<http://www.immunocorp.com>

Immunolex Therapeutics ApS
Grusbakken 6-8
DK-2820 Gentofte
Denmark
(45) 43-62-63-32
FAX: (45) 43-62-63-31
<http://www.immunolex.com>

Immunotech
130, av. Delattre de Tassigny
B.P. 177
13276 Marseilles Cedex 9
France
(33) 491-17-27-00
FAX: (33) 491-41-43-58
<http://www.immunotech.fr>

Imperial Chemical Industries
Imperial Chemical House
Millbank, London SW1P 3JF, UK
(44) 171-834-4444
FAX: (44)171-834-2042
<http://www.ici.com>

Inceltech
See New Brunswick Scientific

Incstar
See DiaSorin

Incyte
6519 Dumbarton Circle
Fremont, CA 94555
(510) 739-2100 FAX: (510) 739-2200
<http://www.incyte.com>

Incyte Pharmaceuticals
3160 Porter Drive
Palo Alto, CA 94304
(877) 746-2983 FAX: (650) 855-0572
(650) 855-0555
<http://www.incyte.com>

Individual Monitoring Systems
6310 Harford Road
Baltimore, MD 21214

Indo Fine Chemical
P.O. Box 473
Somerville, NJ 08876
(888) 463-6346 FAX: (908) 359-1179
(908) 359-6778
<http://www.indofinechemical.com>

Industrial Acoustics
1160 Commerce Avenue
Bronx, NY 10462
(718) 931-8000 FAX: (718) 863-1138
<http://www.industrialacoustics.com>

Inex Pharmaceuticals
100-8900 Glenlyon Parkway
Glenlyon Business Park
Burnaby, British Columbia
V5J 5J8 Canada
(604) 419-3200 FAX: (604) 419-3201
<http://www.inexpharm.com>

Ingold, Mettler, Toledo
261 Ballardvale Street
Wilmington, MA 01887
(800) 352-8763 FAX: (978) 658-0020
(978) 658-7615
<http://www.mt.com>

Innogenetics N.V.
Technologie Park 6
B-9052 Zwijnaarde
Belgium
(32) 9-329-1329 FAX: (32) 9-245-7623
<http://www.innogenetics.com>

Innovative Medical Services
1725 Gillespie Way
El Cajon, CA 92020
(619) 596-8600 FAX: (619) 596-8700
<http://www.imspure.com>

Suppliers

A.4.16

Innovative Research
3025 Harbor Lane N, Suite 300
Plymouth, MN 55447
(612) 519-0105 FAX: (612) 519-0239
<http://www.inres.com>

Innovative Research of America
2 N. Tamiami Trail, Suite 404
Sarasota, FL 34236
(800) 421-8171 FAX: (800) 643-4345
(941) 365-1406 FAX: (941) 365-1703
<http://www.innovrsrch.com>

Inotech Biosystems
15713 Crabbs Branch Way, #110
Rockville, MD 20855
(800) 635-4070 FAX: (301) 670-2859
(301) 670-2850
<http://www.inotechintl.com>

INOVISION
22699 Old Canal Road
Yorba Linda, CA 92887
(714) 998-9600 FAX: (714) 998-9666
<http://www.inovision.com>

Instech Laboratories
5209 Militia Hill Road
Plymouth Meeting, PA 19462
(800) 443-4227 FAX: (610) 941-0134
(610) 941-0132
<http://www.instechlabs.com>

Instron
100 Royall Street
Canton, MA 02021
(800) 564-8378 FAX: (781) 575-5725
(781) 575-5000
<http://www.instron.com>

Instrumentarium
P.O. Box 300
00031 Instrumentarium
Helsinki, Finland
(10) 394-5566
<http://www.instrumentarium.fi>

Instruments SA
Division Jobin Yvon
16-18 Rue du Canal
91165 Longjumeau, Cedex, France
(33)1 6454-1300
FAX: (33)1 6909-9319
<http://www.isainc.com>

Instrutech
20 Vanderverter Avenue, Suite 101E
Port Washington, NY 11050
(516) 883-1300 FAX: (516) 883-1558
<http://www.instrutech.com>

Integrated DNA Technologies
1710 Commercial Park
Coralville, IA 52241
(800) 328-2661 FAX: (319) 626-8444
<http://www.idtdna.com>

Integrated Genetics
See Genzyme Genetics

Integrated Scientific Imaging Systems
3463 State Street, Suite 431
Santa Barbara, CA 93105
(805) 692-2390 FAX: (805) 692-2391
<http://www.imagingsystems.com>

Integrated Separation Systems (ISS)
See OWL Separation Systems

IntelliGenetics
See Oxford Molecular Group

Interactiva BioTechnologie
Sedanstrasse 10
D-89077 Ulm, Germany
(49) 731-93579-290
FAX: (49) 731-93579-291
<http://www.interactiva.de>

Interchim
213 J.F. Kennedy Avenue
B.P. 1140
Montlucon
03103 France
(33) 04-70-03-83-55
FAX: (33) 04-70-03-93-60

Interfocus
14/15 Spring Rise
Falcover Road
Haverhill, Suffolk CB9 7XU, UK
(44) 1440 703460
FAX: (44) 1440 704397
<http://www.interfocus.ltd.uk>

Intergen
2 Manhattanville Road
Purchase, NY 10577
(800) 431-4505 FAX: (800) 468-7436
(914) 694-1700 FAX: (914) 694-1429
<http://www.intergenco.com>

Intermountain Scientific
420 N. Keys Drive
Kaysville, UT 84037
(800) 999-2901 FAX: (800) 574-7892
(801) 547-5047 FAX: (801) 547-5051
<http://www.bioexpress.com>

International Biotechnologies (IBI)
See Eastman Kodak

International Equipment Co. (IEC)
See Thermoquest

**International Institute for the
Advancement of Medicine**
1232 Mid-Valley Drive
Jessup, PA 18434
(800) 486-IIAM FAX: (570) 343-6993
(570) 496-3400
<http://www.iiam.org>

International Light
17 Graf Road
Newburyport, MA 01950
(978) 465-5923 FAX: (978) 462-0759

International Market Supply (I.M.S.)
Dane Mill
Broadhurst Lane
Congleton, Cheshire CW12 1LA, UK
(44) 1260 275469
FAX: (44) 1260 276007

International Marketing Services
See International Marketing Ventures

International Marketing Ventures
6301 Ivy Lane, Suite 408
Greenbelt, MD 20770
(800) 373-0096 FAX: (301) 345-0631
(301) 345-2866
<http://www.imvlimited.com>

International Products
201 Connecticut Drive
Burlington, NJ 08016
(609) 386-8770 FAX: (609) 386-8438
<http://www.mkt@ipcol.com>

Intracel Corporation
Bartels Division
2005 Sammamish Road, Suite 107
Issaquah, WA 98027
(800) 542-2281 FAX: (425) 557-1894
(425) 392-2992
<http://www.intracel.com>

Invitrogen
1600 Faraday Avenue
Carlsbad, CA 92008
(800) 955-6288 FAX: (760) 603-7201
(760) 603-7200
<http://www.invitrogen.com>

In Vivo Metric
P.O. Box 249
Healdsburg, CA 95448
(707) 433-4819 FAX: (707) 433-2407

IRORI
9640 Towne Center Drive
San Diego, CA 92121
(858) 546-1300 FAX: (858) 546-3083
<http://www.ironi.com>

Irvine Scientific
2511 Daimler Street
Santa Ana, CA 92705
(800) 577-6097 FAX: (949) 261-6522
(949) 261-7800
<http://www.irvinesci.com>

ISC BioExpress
420 North Kays Drive
Kaysville, UT 84037
(800) 999-2901 FAX: (800) 574-7892
(801) 547-5047
<http://www.bioexpress.com>

ISCO
P.O. Box 5347
4700 Superior
Lincoln, NE 68505
(800) 228-4373 FAX: (402) 464-0318
(402) 464-0231
<http://www.isco.com>

Isis Pharmaceuticals
Carlsbad Research Center
2292 Faraday Avenue
Carlsbad, CA 92008
(760) 931-9200
<http://www.isip.com>

Isolabs
See Wallac

ISS
See Integrated Separation Systems

J & W Scientific
See Agilent Technologies

J.A. Webster
86 Leominster Road
Sterling, MA 01564
(800) 225-7911 FAX: (978) 422-8959
<http://www.jawebster.com>

J.T. Baker
See Mallinckrodt Baker
222 Red School Lane
Phillipsburg, NJ 08865
(800) JTBAKER FAX: (908) 859-6974
<http://www.jtbaker.com>

**Jackson ImmunoResearch
Laboratories**
P.O. Box 9
872 W. Baltimore Pike
West Grove, PA 19390
(800) 367-5296 FAX: (610) 869-0171
(610) 869-4024
<http://www.jacksonimmuno.com>

The Jackson Laboratory
600 Maine Street
Bar Harbor, ME 04059
(800) 422-6423 FAX: (207) 288-5079
(207) 288-6000
<http://www.jax.org>

Jaece Industries
908 Niagara Falls Boulevard
North Tonawanda, NY 14120
(716) 694-2811 FAX: (716) 694-2811
<http://www.jaece.com>

Jandel Scientific
See SPSS

Janke & Kunkel
See Ika Works

Janssen Life Sciences Products
See Amersham

Janssen Pharmaceutica
1125 Trenton-Harbourton Road
Titusville, NJ 09560
(609) 730-2577 FAX: (609) 730-2116
<http://us.janssen.com>

Jasco
8649 Commerce Drive
Easton, MD 21601
(800) 333-5272 FAX: (410) 822-7526
(410) 822-1220
<http://www.jascoinc.com>

Suppliers

A.4.17

Jeffers
P.O. Box 100
Dothan, AL 36302
(800) 533-3377 Fax: 1-334-793-5179
<http://www.1800jeffers.com/ssc/>

Jena Bioscience
Loebstedter Str. 78
07749 Jena, Germany
(49) 3641-464920
FAX: (49) 3641-464991
<http://www.jenabioscience.com>

Jencons Scientific
800 Bursca Drive, Suite 801
Bridgeville, PA 15017
(800) 846-9959 Fax: (412) 257-8809
(412) 257-8861
<http://www.jencons.co.uk>

JEOL Instruments
11 Dearborn Road
Peabody, MA 01960
(978) 535-5900 Fax: (978) 536-2205
<http://www.jeol.com/index.html>

Jewett
750 Grant Street
Buffalo, NY 14213
(800) 879-7767 Fax: (716) 881-6092
(716) 881-0030
<http://www.JewettInc.com>

John's Scientific
See VWR Scientific

John Weiss and Sons
95 Alston Drive
Bradwell Abbey
Milton Keynes, Buckinghamshire
MK1 4HF UK
(44) 1908-318017
FAX: (44) 1908-318708

Johnson & Johnson Medical
2500 Arbrook Boulevard East
Arlington, TX 76004
(800) 423-4018
<http://www.jnjmedical.com>

Johnston Matthey Chemicals
Orchard Road
Royston, Hertfordshire SG8 5HE, UK
(44) 1763-253000
FAX: (44) 1763-253466
<http://www.chemicals.matthey.com>

Jolley Consulting and Research
683 E. Center Street, Unit H
Grayslake, IL 60030
(847) 548-2330 Fax: (847) 548-2984
<http://www.jolley.com>

Jordan Scientific
See Shelton Scientific

Jorgensen Laboratories
1450 N. Van Buren Avenue
Loveland, CO 80538
(800) 525-5614 Fax: (970) 663-5042
(970) 669-2500
<http://www.jorvet.com>

**JRH Biosciences and
JR Scientific**
13804 W. 107th Street
Lenexa, KS 66215
(800) 231-3735 Fax: (913) 469-5584
(913) 469-5580

Jule Bio Technologies
25 Science Park, #14, Suite 695
New Haven, CT 06511
(800) 648-1772 Fax: (203) 786-5489
(203) 786-5490
<http://hometown.aol.com/precastgel/index.htm>

K.R. Anderson
2800 Bowers Avenue
Santa Clara, CA 95051
(800) 538-8712 Fax: (408) 727-2959
(408) 727-2800
<http://www.krandonson.com>

Kabi Pharmacia Diagnostics
See Pharmacia Diagnostics

Kanthal H.P. Reid
1 Commerce Boulevard
P.O. Box 352440
Palm Coast, FL 32135
(904) 445-2000 Fax: (904) 446-2244
<http://www.kanthal.com>

Kapak
5305 Parkdale Drive
St. Louis Park, MN 55416
(800) KAPAK-57 Fax: (612) 541-0735
(612) 541-0730
<http://www.kapak.com>

Karl Hecht
Stettener Str. 22-24
D-97647 Sondheim
Rhön, Germany
(49) 9779-8080 Fax: (49) 9779-80888

Karl Storz
Köningin-Elisabeth Str. 60
D-14059 Berlin, Germany
(49) 30-30 69 09-0
FAX: (49) 30-30 19 452
<http://www.karlstorz.de>

KaVo EWL
P.O. Box 1320
D-88293 Leutkirch im Allgäu, Germany
(49) 7561-86-0 Fax: (49) 7561-86-371
<http://www.kavo.com/english/startseite.htm>

Keithley Instruments
28775 Aurora Road
Cleveland, OH 44139
(800) 552-1115 Fax: (440) 248-6168
(440) 248-0400
<http://www.keithley.com>

Kemin
2100 Maury Street, Box 70
Des Moines, IA 50301
(515) 266-2111 Fax: (515) 266-8354
<http://www.kemin.com>

Kemo
3 Brook Court, Blakeney Road
Beckenham, Kent BR3 1HG, UK
(44) 0181 658 3838
FAX: (44) 0181 658 4084
<http://www.kemo.com>

Kendall
15 Hampshire Street
Mansfield, MA 02048
(800) 962-9888 Fax: (800) 724-1324
<http://www.kendallhq.com>

Kendro Laboratory Products
31 Pecks Lane
Newtown, CT 06470
(800) 522-SPIN Fax: (203) 270-2166
(203) 270-2080
<http://www.kendro.com>

Kendro Laboratory Products
P.O. Box 1220
Am Kalkberg
D-3360 Osterod, Germany
(55) 22-316-213
FAX: (55) 22-316-202
<http://www.heraeus-instruments.de>

Kent Laboratories
23404 NE 8th Street
Redmond, WA 98053
(425) 868-6200 Fax: (425) 868-6335
<http://www.kentlabs.com>

Kent Scientific
457 Bantam Road, #16
Litchfield, CT 06759
(888) 572-8887 Fax: (860) 567-4201
(860) 567-5496
<http://www.kentscientific.com>

Kentek Corp.
1 Elm Street
Pittsfield, NH 03263
(800) 432-2323 Fax: 603-435-7441
<http://www.kentek-laser.com>

Keuffel & Esser
See Azon

Keystone Scientific
Penn Eagle Industrial Park
320 Rolling Ridge Drive
Bellefonte, PA 16823
(800) 437-2999 Fax: (814) 353-2305
(814) 353-2300 Ext 1
<http://www.keystonescientific.com>

Kimble/Kontes Biotechnology
1022 Spruce Street
P.O. Box 729
Vineland, NJ 08360
(888) 546-2531 Fax: (856) 794-9762
(856) 692-3600
<http://www.kimble-kontes.com>

Kinematica AG
Luzernerstrasse 147a
CH-6014 Littau-Luzern, Switzerland
(41) 41 2501257 Fax: (41) 41 2501460
<http://www.kinematica.ch>

Kin-Tek
504 Laurel Street
LaMarque, TX 77568
(800) 326-3627
FAX: (409) 938-3710
<http://www.kin-tek.com>

Kipp & Zonen
125 Wilbur Place
Bohemia, NY 11716
(800) 645-2065 Fax: (516) 589-2068
(516) 589-2885
<http://www.kippzonen.thomasregister.com/olc/kippzonen>

Kirkegaard & Perry Laboratories
2 Cessna Court
Gaithersburg, MD 20879
(800) 638-3167 Fax: (301) 948-0169
(301) 948-7755
<http://www.kpl.com>

Kodak
See Eastman Kodak

Kontes Glass
See Kimble/Kontes Biotechnology

Kontron Instruments AG
Postfach CH-8010
Zurich, Switzerland
41-1-733-5733 Fax: 41-1-733-5734

David Kopf Instruments
P.O. Box 636
Tujunga, CA 91043
(818) 352-3274 Fax: (818) 352-3139

Kraft Apparatus
See Glas-Col Apparatus

Kramer Scientific Corporation
711 Executive Boulevard
Valley Cottage, NY 10989
(845) 267-5050 Fax: (845) 267-5550

Kulite Semiconductor Products
1 Willow Tree Road
Leonia, NJ 07605
(201) 461-0900 Fax: (201) 461-0990
<http://www.kulite.com>

Lab-Line Instruments
15th & Bloomingdale Avenues
Melrose Park, IL 60160
(800) LAB-LINE Fax: (708) 450-5830
FAX: (800) 450-4LAB
<http://www.labline.com>

Lab Products
742 Sussex Avenue
P.O. Box 639
Seaford, DE 19973
(800) 526-0469 Fax: (302) 628-4309
(302) 628-4300
<http://www.labproductsinc.com>

LabRepco
101 Witmer Road, Suite 700
Horsham, PA 19044
(800) 521-0754 Fax: (215) 442-9202
<http://www.labrepco.com>

Suppliers

A.4.18

Lab Safety Supply

P.O. Box 1368
Janesville, WI 53547
(800) 356-0783 FAX: (800) 543-9910
(608) 754-7160 FAX: (608) 754-1806
<http://www.labsafety.com>

Lab-Tek Products

See Nalge Nunc International

Labconco

8811 Prospect Avenue
Kansas City, MO 64132
(800) 821-5525 FAX: (816) 363-0130
(816) 333-8811
<http://www.labconco.com>

Labindustries

See Barnstead/ThermoLyne

Labnet International

P.O. Box 841
Woodbridge, NJ 07095
(888) LAB-NET1 FAX: (732) 417-1750
(732) 417-0700
<http://www.nationalabnet.com>

LABO-MODERNE

37 rue Dombasle
Paris
75015 France
(33) 01-45-32-62-54
FAX: (33) 01-45-32-01-09
<http://www.laboderne.com/fr>

Laboratory of Immunoregulation

National Institute of Allergy and
Infectious Diseases/NIH
9000 Rockville Pike
Building 10, Room 11B13
Bethesda, MD 20892
(301) 496-1124

Laboratory Supplies

29 Jefry Lane
Hicksville, NY 11801
(516) 681-7711

Labscan Limited

Stillorgan Industrial Park
Stillorgan
Dublin, Ireland
(353) 1-295-2684
FAX: (353) 1-295-2685
<http://www.labscan.ie>

Labsystems

See Thermo Labsystems

Labsystems Affinity Sensors

Saxon Way, Bar Hill
Cambridge CB3 8SL, UK
44 (0) 1954 789976
FAX: 44 (0) 1954 789417
<http://www.affinity-sensors.com>

Labtronics

546 Governors Road
Guelph, Ontario
N1K 1E3 Canada
(519) 763-4930 FAX: (519) 836-4431
<http://www.labtronics.com>

Labtronix Manufacturing

3200 Investment Boulevard
Hayward, CA 94545
(510) 786-3200 FAX: (510) 786-3268
<http://www.labtronix.com>

Lafayette Instrument

3700 Sagamore Parkway North
P.O. Box 5729
Lafayette, IN 47903
(800) 428-7545 FAX: (765) 423-4111
(765) 423-1505
<http://www.lafayetteinstrument.com>

Lambert Instruments

Turfweg 4
9313 TH Leutingewolde
The Netherlands
(31) 50-5018461 FAX: (31) 50-5010034
<http://www.lambert-instruments.com>

Lampire Biological Laboratories

P.O. Box 270
Pipersville, PA 18947
(215) 795-2538 FAX: (215) 795-0237
<http://www.lampire.com>

Lancaster Synthesis

P.O. Box 1000
Windham, NH 03087
(800) 238-2324 FAX: (603) 889-3326
(603) 889-3306
<http://www.lancastersynthesis-us.com>

Lancer

140 State Road 419
Winter Springs, FL 32708
(800) 332-1855 FAX: (407) 327-1229
(407) 327-8488
<http://www.lancer.com>

LaVision GmbH

Gerhard-Gerdes-Str. 3
D-37079
Goettingen, Germany
(49) 551-50549-0
FAX: (49) 551-50549-11
<http://www.lavision.de>

Lawshe

See Advanced Process Supply

Laxotan

20, rue Leon Blum
26000 Valence, France
(33) 4-75-41-91-91
FAX: (33) 4-75-41-91-99
<http://www.latoxan.com>

LC Laboratories

165 New Boston Street
Woburn, MA 01801
(781) 937-0777 FAX: (781) 938-5420
<http://www.lclaboratories.com>

LC Packings

80 Carolina Street
San Francisco, CA 94103
(415) 552-1855 FAX: (415) 552-1859
<http://www.lcpackings.com>

LC Services

See LC Laboratories

LECO

3000 Lakeview Avenue
St. Joseph, MI 49085
(800) 292-6141 FAX: (616) 982-8977
(616) 985-5496
<http://www.leco.com>

Lederle Laboratories

See Wyeth-Ayerst

Lee Biomolecular Research

Laboratories
11211 Sorrento Valley Road, Suite M
San Diego, CA 92121
(858) 452-7700

The Lee Company

2 Pettipaug Road
P.O. Box 424
Westbrook, CT 06498
(800) LEE-PLUG FAX: (860) 399-7058
(860) 399-6281
<http://www.theleeco.com>

Lee Laboratories

1475 Athens Highway
Grayson, GA 30017
(800) 732-9150 FAX: (770) 979-9570
(770) 972-4450
<http://www.leelabs.com>

Leica

111 Deer Lake Road
Deerfield, IL 60015
(800) 248-0123 FAX: (847) 405-0147
(847) 405-0123
<http://www.leica.com>

Leica Microsystems

Imneuenheimer Feld 518
D-69120
Heidelberg, Germany
(49) 6221-41480
FAX: (49) 6221-414833
<http://www.leica-microsystems.com>

Leinco Technologies

359 Consort Drive
St. Louis, MO 63011
(314) 230-9477 FAX: (314) 527-5545
<http://www.leinco.com>

Leitz U.S.A.

See Leica

LenderKing Metal Products

8370 Jumpers Hole Road
Millersville, MD 21108
(410) 544-8795 FAX: (410) 544-5069
<http://www.lenderking.com>

Letica Scientific Instruments

Panlab s.i., c/Loreto 50
08029 Barcelona, Spain
(34) 93-419-0709 FAX: (34) 93-419-7145
<http://www.panlab-sl.com>

Lexel Laser

853 Brown Road
Fremont, CA 94539
(510)-651-0110 Fax: 510-651-1690
<http://www.lexellaser.com>

Leybold-Heraeus Trivac DZA

5700 Mellon Road
Export, PA 15632
(412) 327-5700

LI-COR

Biotechnology Division
4308 Progressive Avenue
Lincoln, NE 68504
(800) 645-4267 FAX: (402) 467-0819
(402) 467-0700
<http://www.licor.com>

Life Science Laboratories

See Adaptive Biosystems

Life Science Resources

Two Corporate Center Drive
Melville, NY 11747
(800) 747-9530 FAX: (516) 844-5114
(516) 844-5085
<http://www.astrocam.com>

Life Sciences

2900 72nd Street North
St. Petersburg, FL 33710
(800) 237-4323 FAX: (727) 347-2957
(727) 345-9371
<http://www.lifesci.com>

Life Technologies

9800 Medical Center Drive
P.O. Box 6482
Rockville, MD 20849
(800) 828-6686 FAX: (800) 331-2286
<http://www.lifetech.com>

Lifecodes

550 West Avenue
Stamford, CT 06902
(800) 543-3263 FAX: (203) 328-9599
(203) 328-9500
<http://www.lifecodes.com>

Lightnin

135 Mt. Read Boulevard
Rochester, NY 14611
(888) MIX-BEST FAX: (716) 527-1742
(716) 436-5550
<http://www.lightnin-mixers.com>

Linear Drives

Luckyn Lane, Pippis Hill
Basildon, Essex SS14 3BW, UK
(44) 1268-287070
FAX: (44) 1268-293344
<http://www.lineardrives.com>

Linscott's Directory

4877 Grange Road
Santa Rosa, CA 95404
(707) 544-9555 FAX: (415) 389-6025
<http://www.linscottsdirectory.co.uk>

Linton Instrumentation

Unit 11, Forge Business Center
Upper Rose Lane
Palgrave, Diss, Norfolk IP22 1AP, UK
(44) 1-379-651-344
FAX: (44) 1-379-650-970
<http://www.lintoninst.co.uk>

Suppliers**A.4.19**

List Biological Laboratories

501-B Vandell Way
Campbell, CA 95008
(800) 726-3213 FAX: (408) 866-6364
(408) 866-6363
<http://www.listlabs.com>

LKB Instruments

See Amersham Pharmacia Biotech

Lloyd Laboratories

604 West Thomas Avenue
Shenandoah, IA 51601
(800) 831-0004 FAX: (712) 246-5245
(712) 246-4000
<http://www.lloydinc.com>

Loctite

1001 Trout Brook Crossing
Rocky Hill, CT 06067
(860) 571-5100 FAX: (860)571-5465
<http://www.loctite.com>

Lofstrand Labs

7961 Cessna Avenue
Gaithersburg, MD 20879
(800) 541-0362 FAX: (301) 948-9214
(301) 330-0111
<http://www.lofstrand.com>

Lomir Biochemical

99 East Main Street
Malone, NY 12953
(877) 425-3604 FAX: (518) 483-8195
(518) 483-7697
<http://www.lomir.com>

LSL Biolafitte

10 rue de Temara
7810C St.-Germain-en-Laye, France
(33) 1-3061-5260
FAX: (33) 1-3061-5234

Ludl Electronic Products

171 Brady Avenue
Hawthorne, NY 10532
(888) 769-6111 FAX: (914) 769-4759
(914) 769-6111
<http://www.ludl.com>

Lumigen

24485 W. Ten Mile Road
Southfield, MI 48034
(248) 351-5600 FAX: (248) 351-0518
<http://www.lumigen.com>

Luminex

12212 Technology Boulevard
Austin, TX 78727
(888) 219-8020 FAX: (512) 258-4173
(512) 219-8020
<http://www.luminexcorp.com>

LYNX Therapeutics

25861 Industrial Boulevard
Hayward, CA 94545
(510) 670-9300 FAX: (510) 670-9302
<http://www.lynxgen.com>

Lyphomed

3 Parkway North
Deerfield, IL 60015
(847) 317-8100 FAX: (847) 317-8600

M.E.D. Associates

See Med Associates

Macherey-Nagel

6 South Third Street, #402
Easton, PA 18042
(610) 559-9848 FAX: (610) 559-9878
<http://www.macherey-nagel.com>

Macherey-Nagel

Valenciennner Strasse 11
P.O. Box 101352
D-52313 Dueren, Germany
(49) 2421-969141
FAX: (49) 2421-969199
<http://www.macherey-nagel.ch>

Mac-Mod Analytical

127 Commons Court
Chadds Ford, PA 19317
800-441-7508 FAX: (610) 358-5993
(610) 358-9696
<http://www.mac-mod.com>

Mallinckrodt Baker

222 Red School Lane
Phillipsburg, NJ 08865
(800) 582-2537 FAX: (908) 859-6974
(908) 859-2151
<http://www.mallbaker.com>

Mallinckrodt Chemicals

16305 Swingley Ridge Drive
Chesterfield, MD 63017
(314) 530-2172 FAX: (314) 530-2563
<http://www.mallichem.com>

Malven Instruments

Enigma Business Park
Groveood Road
Malven, Worcestershire
WR 141 XZ, United Kingdom

Marinus

1500 Pier C Street
Long Beach, CA 90813
(562) 435-6522 FAX: (562) 495-3120

Markson Science
c/o Whatman Labs Sales
P.O. Box 1359
Hillsboro, OR 97123
(800) 942-8626 FAX: (503) 640-9716
(503) 648-0762

Marsh Biomedical Products

565 Blossom Road
Rochester, NY 14610
(800) 445-2812 FAX: (716) 654-4810
(716) 654-4800
<http://www.biomar.com>

Marshall Farms USA

5800 Lake Bluff Road
North Rose, NY 14516
(315) 587-2295
e-mail: info@marfarms.com

Martek

6480 Dobbin Road
Columbia, MD 21045
(410) 740-0081 FAX: (410) 740-2985
<http://www.martekbio.com>

Martin Supply

Distributor of Gerber Scientific
2740 Loch Raven Road
Baltimore, MD 21218
(800) 282-5440 FAX: (410) 366-0134
(410) 366-1696

Mast Immunsystems

630 Clyde Court
Mountain View, CA 94043
(800) 233-MAST FAX: (650) 969-2745
(650) 961-5501
<http://www.mastallergy.com>

Matheson Gas Products

P.O. Box 624
959 Route 46 East
Parsippany, NJ 07054
(800) 416-2505 FAX: (973) 257-9393
(973) 257-1100
<http://www.mathesongas.com>

Mathsoft

1700 Westlake Avenue N., Suite 500
Seattle, WA 98109
(800) 569-0123 FAX: (206) 283-8691
(206) 283-8802
<http://www.mathsoft.com>

Matreya

500 Tressler Street
Pleasant Gap, PA 16823
(814) 359-5060 FAX: (814) 359-5062
<http://www.matreya.com>

Matrigel

See Becton Dickinson Labware

Matrix Technologies

22 Friars Drive
Hudson, NH 03051
(800) 345-0206 FAX: (603) 595-0106
(603) 595-0505
<http://www.matrixtechcorp.com>

MatTek Corp.

200 Homer Avenue
Ashland, Massachusetts 01721
(508) 881-6771 FAX: (508) 879-1532
<http://www.mattek.com>

Maxim Medical

89 Oxford Road
Oxford OX2 9PD
United Kingdom
44 (0)1865-865943
FAX: 44 (0)1865-865291
<http://www.maximmed.com>

Mayo Clinic

Section on Engineering
Project #ALA-1, 1982
200 1st Street SW
Rochester, MN 55905
(507) 284-2511 FAX: (507) 284-5988

McGaw

See B. Braun-McGaw

McMaster-Carr

600 County Line Road
Elmhurst, IL 60126
(630) 833-0300 FAX: (630) 834-9427
<http://www.mcmaster.com>

McNeil Pharmaceutical

See Ortho McNeil Pharmaceutical

MCNC

3021 Cornwallis Road
P.O. Box 12889
Research Triangle Park, NC 27709
(919) 248-1800 FAX: (919) 248-1455
<http://www.mcnc.org>

MD Industries

5 Revere Drive, Suite 415
Northbrook, IL 60062
(800) 421-8370 FAX: (847) 498-2627
(708) 339-6000
<http://www.mdindustries.com>

MDS Nordion

447 March Road
P.O. Box 13500
Kanata, Ontario
K2K 1X8 Canada
(800) 465-3666 FAX: (613) 592-6937
(613) 592-2790
<http://www.mds.nordion.com>

MDS Sciex

71 Four Valley Drive
Concord, Ontario
Canada L4K 4V8
(905) 660-9005 FAX: (905) 660-2600
<http://www.sciex.com>

Mead Johnson

See Bristol-Meyers Squibb

Med Associates

P.O. Box 319
St. Albans, VT 05478
(802) 527-2343 FAX: (802) 527-5095
<http://www.med-associates.com>

Medecell

239 Liverpool Road
London N1 1LX, UK
(44) 20-7607-2295
FAX: (44) 20-7700-4156
<http://www.medicell.co.uk>

Media Cybernetics

8484 Georgia Avenue, Suite 200
Silver Spring, MD 20910
(301) 495-3305 FAX: (301) 495-5964
<http://www.mediacy.com>

Mediatech

13884 Park Center Road
Herndon, VA 20171
(800) cellgro
(703) 471-5955
<http://www.cellgro.com>

Medical Systems

See Harvard Apparatus

Suppliers**A.4.20**

Medifor

647 Washington Street
Port Townsend, WA 98368
(800) 366-3710 FAX: (360) 385-4402
(360) 385-0722
<http://www.medifor.com>

MedImmune

35 W. Watkins Mill Road
Gaithersburg, MD 20878
(301) 417-0770 FAX: (301) 527-4207
<http://www.medimmune.com>

MedProbe AS

P.O. Box 2640
St. Hanshaugen
N-0131 Oslo, Norway
(47) 222 00137 FAX: (47) 222 00189
<http://www.medprobe.com>

Megazyme

Bray Business Park
Bray, County Wicklow
Ireland
(353) 1-286-1220
FAX: (353) 1-286-1264
<http://www.megazyme.com>

Melles Griot

4601 Nautilus Court South
Boulder, CO 80301
(800) 326-4363 FAX: (303) 581-0960
(303) 581-0337
<http://www.mellesgriot.com>

Menzel-Glaser

Postfach 3157
D-38021 Braunschweig, Germany
(49) 531 590080
FAX: (49) 531 509799

E. Merck

Frankfurterstrasse 250
D-64293 Darmstadt 1, Germany
(49) 6151-720

Merck

See EM Science

Merck & Company

Merck National Service Center
P.O. Box 4
West Point, PA 19486
(800) NSC-MERCK
(215) 652-5000
<http://www.merck.com>

Merck Research Laboratories

See Merck & Company

Merck Sharpe Human Health Division

300 Franklin Square Drive
Somerset, NJ 08873
(800) 637-2579 FAX: (732) 805-3960
(732) 805-0300

Merial Limited

115 Transtech Drive
Athens, GA 30601
(800) MERIAL-1 FAX: (706) 548-0608
(706) 548-9292
<http://www.merial.com>

Meridian Instruments

P.O. Box 1204
Kent, WA 98035
(253) 854-9914 FAX: (253) 854-9902
<http://www.minstrument.com>

Meta Systems Group

32 Hammond Road
Belmont, MA 02178
(617) 489-9950 FAX: (617) 489-9952

Metachem Technologies

3547 Voyager Street, Bldg. 102
Torrance, CA 90503
(310) 793-2300 FAX: (310) 793-2304
<http://www.metachem.com>

Metallhantering

Box 47172
100-74 Stockholm, Sweden
(46) 8-726-9696

MethylGene

7220 Frederick-Banting, Suite 200
Montreal, Quebec
H4S 2A1 Canada
<http://www.methylgene.com>

Metro Scientific

475 Main Street, Suite 2A
Farmingdale, NY 11735
(800) 788-6247 FAX: (516) 293-8549
(516) 293-9656

Metrowerks

980 Metric Boulevard
Austin, TX 78758
(800) 377-5416
(512) 997-4700
<http://www.metrowerks.com>

Mettler Instruments

Mettler-Toledo
1900 Polaris Parkway
Columbus, OH 43240
(800) METTLER FAX: (614) 438-4900
<http://www.mt.com>

Miami Serpentarium Labs

34879 Washington Loop Road
Punta Gorda, FL 33982
(800) 248-5050 FAX: (813) 639-1811
(813) 639-8888
<http://www.miamiserpentarium.com>

Michrom BioResources

1945 Industrial Drive
Auburn, CA 95603
(530) 888-6498 FAX: (530) 888-8295
<http://www.michrom.com>

Mickle Laboratory Engineering

Gomshall, Surrey, UK
(44) 1483-202178

Micra Scientific

A division of Eichrom Industries
8205 S. Cass Ave, Suite 111
Darien, IL 60561
(800) 283-4752 FAX: (630) 963-1928
(630) 963-0320
<http://www.micrasci.com>

MicroBrightField

74 Hegman Avenue
Colchester, VT 05446
(802) 655-9360 FAX: (802) 655-5245
<http://www.microbrightfield.com>

Micro Essential Laboratory

4224 Avenue H
Brooklyn, NY 11210
(718) 338-3618 FAX: (718) 692-4491

Micro Filtration Systems

7-3-Chome, Honcho
Nihonbashi, Tokyo, Japan
(81) 3-270-3141

Micro-Metrics

P.O. Box 13804
Atlanta, GA 30324
(770) 986-6015 FAX: (770) 986-9510
<http://www.micro-metrics.com>

Micro-Tech Scientific

140 South Wolfe Road
Sunnyvale, CA 94086
(408) 730-8324 FAX: (408) 730-3566
<http://www.microlc.com>

Microbix Biosystems

341 Bering Avenue
Toronto, Ontario
M8Z 3A8 Canada
1-800-794-6694 FAX: 416-234-1626
1-416-234-1624
<http://www.microbix.com>

MicroCal

22 Industrial Drive East
Northampton, MA 01060
(800) 633-3115 FAX: (413) 586-0149
(413) 586-7720
<http://www.microcalorimetry.com>

Microfluidics

30 Ossipee Road
P.O. Box 9101
Newton, MA 02164
(800) 370-5452 FAX: (617) 965-1213
(617) 969-5452
<http://www.microfluidicscorp.com>

Microgon

See Spectrum Laboratories

Microlase Optical Systems

West of Scotland Science Park
Kelvin Campus, Maryhill Road
Glasgow G20 0SP, UK
(44) 141-948-1000
FAX: (44) 141-946-6311
<http://www.microlase.co.uk>

Micron Instruments

4509 Runway Street
Simi Valley, CA 93063
(800) 638-3770 FAX: (805) 522-4982
(805) 552-4676
<http://www.microninstruments.com>

Micron Separations

See MSI

Micro Photonics

4949 Liberty Lane, Suite 170
P.O. Box 3129
Allentown, PA 18106
(610) 366-7103 FAX: (610) 366-7105
<http://www.microphotonics.com>

MicroTech

1420 Conchester Highway
Boothwyn, PA 19061
(610) 459-3514

Midland Certified Reagent Company

3112-A West Cuthbert Avenue
Midland, TX 79701
(800) 247-8766 FAX: (800) 359-5789
(915) 694-7950 FAX: (915) 694-2387
<http://www.mrcr.com>

Midwest Scientific

280 Vance Road
Valley Park, MO 63088
(800) 227-9997 FAX: (636) 225-9998
(636) 225-9997
<http://www.midsci.com>

Miles

See Bayer

Miles Laboratories

See Serological

Miles Scientific

See Nunc

Millar Instruments

P.O. Box 230227
6001-A Gulf Freeway
Houston, TX 77023
(713) 923-9171 FAX: (713) 923-7757
<http://www.millarinstruments.com>

MilliGen/Biosearch

See Millipore

Millipore

80 Ashbury Road
P.O. Box 9125
Bedford, MA 01730
(800) 645-5476 FAX: (781) 533-3110
(781) 533-6000
<http://www.millipore.com>

Miltenyi Biotec

251 Auburn Ravine Road, Suite 208
Auburn, CA 95603
(800) 367-6227 FAX: (530) 888-8925
(530) 888-8871
<http://www.miltenyibiotec.com>

Miltex

6 Ohio Drive
Lake Success, NY 11042
(800) 645-8000 FAX: (516) 775-7185
(516) 349-0001

Milton Roy

See Spectronic Instruments

Suppliers**A.4.21**

Mini-Instruments

15 Burnham Business Park
Springfield Road
Burnham-on-Crouch, Essex CM0 8TE, UK
(44) 1621-783282
FAX: (44) 1621-783132
<http://www.mini-instruments.co.uk>

Mini Mitter

P.O. Box 3386
Sunriver, OR 97707
(800) 685-2999 FAX: (541) 593-5604
(541) 593-8639
<http://www.minimitter.com>

Mirus Corporation

505 S. Rosa Road
Suite 104
Madison, WI 53719
(608) 441-2852 FAX: (608) 441-2849
<http://www.genetransfer.com>

Misonix

1938 New Highway
Farmingdale, NY 11735
(800) 645-9846 FAX: (516) 694-9412
<http://www.misonix.com>

Mitutoyo (MTI)

See Dolla Eastern

MJ Research

Waltham, MA 02451
(800) PELTIER FAX: (617) 923-8080
(617) 923-8000
<http://www.mjr.com>

Modular Instruments

228 West Gay Street
Westchester, PA 19380
(610) 738-1420 FAX: (610) 738-1421
<http://www.mi2.com>

Molecular Biology Insights

8685 US Highway 24
Cascade, CO 80809-1333
(800) 747-4362 FAX: (719) 684-7989
(719) 684-7988
<http://www.oligo.net>

Molecular Biosystems

10030 Barnes Canyon Road
San Diego, CA 92121
(858) 452-0681 FAX: (858) 452-6187
<http://www.mobi.com>

Molecular Devices

1312 Crossman Avenue
Sunnyvale, CA 94089
(800) 635-5577 FAX: (408) 747-3602
(408) 747-1700
<http://www.moldev.com>

Molecular Designs

1400 Catalina Street
San Leandro, CA 94577
(510) 895-1313 FAX: (510) 614-3608

Molecular Dynamics

928 East Arques Avenue
Sunnyvale, CA 94086
(800) 333-5703 FAX: (408) 773-1493
(408) 773-1222
<http://www.apbiotech.com>

Molecular Probes

4849 Pitchford Avenue
Eugene, OR 97402
(800) 438-2209 FAX: (800) 438-0228
(541) 465-8300 FAX: (541) 344-6504
<http://www.probes.com>

Molecular Research Center

5645 Montgomery Road
Cincinnati, OH 45212
(800) 462-9868 FAX: (513) 841-0080
(513) 841-0900
<http://www.mrcgene.com>

Molecular Simulations

9685 Scranton Road
San Diego, CA 92121
(800) 756-4674 FAX: (858) 458-0136
(858) 458-9990
<http://www.msi.com>

Monoject Disposable Syringes & Needles/Syrvet

16200 Walnut Street
Waukee, IA 50263
(800) 727-5203 FAX: (515) 987-5553
(515) 987-5554
<http://www.syrvet.com>

Monsanto Chemical

800 North Lindbergh Boulevard
St. Louis, MO 63167
(314) 694-1000 FAX: (314) 694-7625
<http://www.monsanto.com>

Moravek Biochemicals

577 Mercury Lane
Brea, CA 92821
(800) 447-0100 FAX: (714) 990-1824
(714) 990-2018
<http://www.moravek.com>

Moss

P.O. Box 189
Pasadena, MD 21122
(800) 932-6677 FAX: (410) 768-3971
(410) 768-3442
<http://www.mosssubstrates.com>

Motion Analysis

3617 Westwind Boulevard
Santa Rosa, CA 95403
(707) 579-6500 FAX: (707) 526-0629
<http://www.motionanalysis.com>

Mott

Farmington Industrial Park
84 Spring Lane
Farmington, CT 06032
(860) 747-6333 FAX: (860) 747-6739
<http://www.mottcorp.com>

MSI (Micron Separations)

See Osmonics

Multi Channel Systems

Markwiesenstrasse 55
72770 Reutlingen, Germany
(49) 7121-503010
FAX: (49) 7121-503011
<http://www.multichannelsystems.com>

Multiple Peptide Systems

3550 General Atomics Court
San Diego, CA 92121
(800) 338-4965 FAX: (800) 654-5592
(858) 455-3710 FAX: (858) 455-3713
<http://www.mps-sd.com>

Murex Diagnostics

3075 Northwoods Circle
Norcross, GA 30071
(707) 662-0660 FAX: (770) 447-4989

MWG-Biotech

Anzinger Str. 7
D-85560 Ebersberg, Germany
(49) 8092-82890 FAX: (49) 8092-21084
http://www.mwg_biotech.com

Myriad Industries

3454 E Street
San Diego, CA 92102
(800) 999-6777 FAX: (619) 232-4819
(619) 232-6700
<http://www.myriadindustries.com>

Nacalai Tesque

Nijo Karasuma, Nakagyo-ku
Kyoto 604, Japan
81-75-251-1723
FAX: 81-75-251-1762
<http://www.nacalai.co.jp>

Nalge Nunc International

Subsidiary of Sybron International
75 Panorama Creek Drive
P.O. Box 20365
Rochester, NY 14602
(800) 625-4327 FAX: (716) 586-8987
(716) 264-9346
<http://www.nalgenunc.com>

Nanogen

10398 Pacific Center Court
San Diego, CA 92121
(858) 410-4600 FAX: (858) 410-4848
<http://www.nanogen.com>

Nanoprobes

95 Horse Block Road
Yaphank, NY 11980
(877) 447-6266 FAX: (631) 205-9493
(631) 205-9490
<http://www.nanoprobes.com>

Narishige USA

1710 Hempstead Turnpike
East Meadow, NY 11554
(800) 445-7914 FAX: (516) 794-0066
(516) 794-8000
<http://www.narishige.co.jp>

Nasco-Fort Atkinson

P.O. Box 901
901 Janesville Ave.
Fort Atkinson, WI 53538-0901
(800) 558-9595 FAX: (920) 563-8296
<http://www.enasco.com>

National Bag Company

2233 Old Mill Road
Hudson, OH 44236
(800) 247-6000 FAX: (330) 425-9800
(330) 425-2600
<http://www.nationalbag.com>

National Band and Tag

Department X 35, Box 72430
Newport, KY 41032
(606) 261-2035 FAX: (800) 261-8247
<https://www.nationalband.com>

National Biosciences

See Molecular Biology Insights

National Diagnostics

305 Patton Drive
Atlanta, GA 30336
(800) 526-3867 FAX: (404) 699-2077
(404) 699-2121
<http://www.nationaldiagnostics.com>

National Disease Research Exchange

1880 John F. Kennedy Blvd., 11th Fl.
Philadelphia, PA 19103
(800) 222-6374
<http://www.ndri.com>

National Institute of Standards and Technology

100 Bureau Drive
Gaithersburg, MD 20899
(301) 975-NIST FAX: (301) 926-1630
<http://www.nist.gov>

National Instruments

11500 North Mopac Expressway
Austin, TX 78759
(512) 794-0100 FAX: (512) 683-8411
<http://www.ni.com>

National Labnet

See Labnet International

National Scientific Instruments

975 Progress Circle
Lawrenceville, GA 300243
(800) 332-3331 FAX: (404) 339-7173
<http://www.nationalscientific.com>

National Scientific Supply

1111 Francisco Boulevard East
San Rafael, CA 94901
(800) 525-1779 FAX: (415) 459-2954
(415) 459-6070
<http://www.nat-sci.com>

Naz-Dar-KC Chicago

Nazdar
1087 N. North Branch Street
Chicago, IL 60622
(800) 736-7636 FAX: (312) 943-8215
(312) 943-8338
<http://www.nazdar.com>

Suppliers**A.4.22**

NB Labs

1918 Avenue A
Denison, TX 75021
(903) 465-2694 FAX: (903) 463-5905
<http://www.nblabslarry.com>

NEB

See New England Biolabs

NEN Life Science Products

549 Albany Street
Boston, MA 02118
(800) 551-2121 FAX: (617) 451-8185
(617) 350-9075
<http://www.nen.com>

NEN Research Products, Dupont (UK)

Diagnostics and Biotechnology Systems
Wedgewood Way
Stevenage, Hertfordshire SG1 4QN, UK
44-1438-734831
44-1438-734000
FAX: 44-1438-734836
<http://www.dupont.com>

Neogen

628 Winchester Road
Lexington, KY 40505
(800) 477-8201 FAX: (606) 255-5532
(606) 254-1221
<http://www.neogen.com>

Neosystems

380, 11012 Macleod Trail South
Calgary, Alberta
T2J 6A5 Canada
(403) 225-9022 FAX: (403) 225-9025
<http://www.neosystems.com>

Neuralynx

2434 North Pantano Road
Tucson, AZ 85715
(520) 722-8144 FAX: (520) 722-8163
<http://www.neuralynx.com>

Neuro Probe

16008 Industrial Drive
Gaithersburg, MD 20877
(301) 417-0014 FAX: (301) 977-5711
<http://www.neuroprobe.com>

Neurocrine Biosciences

10555 Science Center Drive
San Diego, CA 92121
(619) 658-7600 FAX: (619) 658-7602
<http://www.neurocrine.com>

Nevtek

HCR03, Box 99
Burnsville, VA 24487
(540) 925-2322 FAX: (540) 925-2323
<http://www.nevtek.com>

New Brunswick Scientific

44 Talmadge Road
Edison, NJ 08818
(800) 631-5417 FAX: (732) 287-4222
(732) 287-1200
<http://www.nbsc.com>

New England Biolabs (NEB)

32 Tozer Road
Beverly, MA 01915
(800) 632-5227 FAX: (800) 632-7440
<http://www.neb.com>

New England Nuclear (NEN)

See NEN Life Science Products

New MBR

Gubelstrasse 48
CH8050 Zurich, Switzerland
(41) 1-313-0703

Newark Electronics

4801 N. Ravenswood Avenue
Chicago, IL 60640
(800) 4-NEWARK FAX: (773) 907-5339
(773) 784-5100
<http://www.newark.com>

Newell Rubbermaid

29 E. Stephenson Street
Freeport, IL 61032
(815) 235-4171 FAX: (815) 233-8060
<http://www.newellco.com>

Newport Biosystems

1860 Trainor Street
Red Bluff, CA 96080
(530) 529-2448 FAX: (530) 529-2648

Newport

1791 Deere Avenue
Irvine, CA 92606
(800) 222-6440 FAX: (949) 253-1800
(949) 253-1462
<http://www.newport.com>

Nexin Research B.V.

P.O. Box 16
4740 AA Hoeven, The Netherlands
(31) 165-503172
FAX: (31) 165-502291

NIAID

See Bio-Tech Research Laboratories

Nichiryō

230 Route 206
Building 2-2C
Flanders, NJ 07836
(877) 548-6667 FAX: (973) 927-0099
(973) 927-4001
<http://www.nichiryō.com>

Nichols Institute Diagnostics

33051 Calle Aviador
San Juan Capistrano, CA 92675
(800) 286-4NID FAX: (949) 240-5273
(949) 728-4610
<http://www.nicholsdiag.com>

Nichols Scientific Instruments

3334 Brown Station Road
Columbia, MO 65202
(573) 474-5522 FAX: (603) 215-7274
<http://home.beseen.com>
technology/lsi_technology

Nicolet Biomedical Instruments

5225 Verona Road, Building 2
Madison, WI 53711
(800) 356-0007 FAX: (608) 441-2002
(608) 273-5000
<http://nicoletbiomedical.com>

N.I.G.M.S. (National Institute of General Medical Sciences)

See Coriell Institute for Medical Research

Nikon

Science and Technologies Group
1300 Walt Whitman Road
Melville, NY 11747
(516) 547-8500 FAX: (516) 547-4045
<http://www.nikonusa.com>

Nippon Gene

1-29, Ton-ya-machi
Toyama 930, Japan
(81) 764-51-6548
FAX: (81) 764-51-6547

Noldus Information Technology

751 Miller Drive
Suite E-5
Leesburg, VA 20175
(800) 355-9541 FAX: (703) 771-0441
(703) 771-0440
<http://www.noldus.com>

Nonlinear Dynamics

See NovoDynamics

Nordion International

See MDS Nordion

North American Biologicals (NABI)

16500 NW 15th Avenue
Miami, FL 33169
(800) 327-7106 (305) 625-5305
<http://www.nabi.com>

North American Reiss

See Reiss

Northwestern Bottle

24 Walpole Park South
Walpole, MA 02081
(508) 668-8600 FAX: (508) 668-7790

NOVA Biomedical

Nova Biomedical 200
Prospect Street Waltham, MA 02454
(800) 822-0911 FAX: (781) 894-5915
<http://www.novabiomedical.com>

Novagen

601 Science Drive
Madison, WI 53711
(800) 526-7319 FAX: (608) 238-1388
(608) 238-6110
<http://www.novagen.com>

Novartis

59 Route 10
East Hanover, NJ 07936
(800)526-0175 FAX: (973) 781-6356
<http://www.novartis.com>

Novartis Biotechnology

3054 Cornwallis Road
Research Triangle Park, NC 27709
(888) 462-7288 FAX: (919) 541-8585
<http://www.novartis.com>

Nova Sina AG

Subsidiary of Airflow Lufttechnik GmbH
Kleine Heeg 21
52259 Rheinbach, Germany
(49) 02226 920-0
FAX: (49) 02226 9205-11

Novex/Invitrogen

1600 Faraday
Carlsbad, CA 92008
(800) 955-6288 FAX: (760) 603-7201
<http://www.novex.com>

Novo Nordisk Biochem

77 Perry Chapel Church Road
Franklington, NC 27525
(800) 879-6686 FAX: (919) 494-3450
(919) 494-3000
<http://www.novo.dk>

Novo Nordisk BioLabs

See Novo Nordisk Biochem

Novocastra Labs

Balliol Business Park West
Benton Lane
Newcastle-upon-Tyne
Tyne and Wear NE12 8EW, UK
(44) 191-215-0567
FAX: (44) 191-215-1152
<http://www.novocastra.co.uk>

NovoDynamics

123 North Ashley Street
Suite 210
Ann Arbor, MI 48104
(734) 205-9100 FAX: (734) 205-9101
<http://www.novodynamics.com>

Novus Biologicals

P.O. Box 802
Littleton, CO 80160
(888) 506-6887 FAX: (303) 730-1966
<http://www.novus-biologicals.com/main.html>

NPI Electronic

Hauptstrasse 96
D-71732 Tamm, Germany
(49) 7141-601534
FAX: (49) 7141-601266
<http://www.npielectronic.com>

NSG Precision Cells

195G Central Avenue
Farmingdale, NY 11735
(516) 249-7474 FAX: (516) 249-8575
<http://www.nsgpci.com>

Nu Chek Prep

109 West Main
P.O. Box 295
Elysian, MN 56028
(800) 521-7728 FAX: (507) 267-4790
(507) 267-4689

Suppliers**A.4.23**

Nuclepore

See Costar

Numonics

101 Commerce Drive
Montgomeryville, PA 18936
(800) 523-6716 FAX: (215) 361-0167
(215) 362-2766
<http://www.interactivewhiteboards.com>

NYCOMED AS Pharma

c/o Accurate Chemical & Scientific
300 Shames Drive
Westbury, NY 11590
(800) 645-6524 FAX: (516) 997-4948
(516) 333-2221
<http://www.accuratechemical.com>

Nycomed Amersham

Health Care Division
101 Carnegie Center
Princeton, NJ 08540
(800) 832-4633 FAX: (800) 807-2382
(609) 514-6000
<http://www.nycomed-amersham.com>

Nyegaard

Herserudsvagen 5254
S-122 06 Lidings, Sweden
(46) 8-765-2930

Ohmeda Catheter Products

See Datex-Ohmeda

Ohwa Tsusbo

Hiby Dai Building
1-2-2 Uchi Saiwai-cho
Chiyoda-ku
Tokyo 100, Japan
03-3591-7348 FAX: 03-3501-9001

Oligos Etc.

9775 S.W. Commerce Circle, C-6
Wilsonville, OR 97070
(800) 888-2358 FAX: (503) 6822D1635
(503) 6822D1814
<http://www.oligoetc.com>

Olis Instruments

130 Conway Drive
Bogart, GA 30622
(706) 353-6547 (800) 852-3504
<http://www.olisweb.com>

Olympus America

2 Corporate Center Drive
Melville, NY 11747
(800) 645-8160 FAX: (516) 844-5959
(516) 844-5000
<http://www.olympusamerica.com>

Omega Engineering

One Omega Drive
P.O. Box 4047
Stamford, CT 06907
(800) 848-4286 FAX: (203) 359-7700
(203) 359-1660
<http://www.omega.com>

Omega Optical

3 Grove Street
P.O. Box 573
Brattleboro, VT 05302
(802) 254-2690 FAX: (802) 254-3937
<http://www.omegafilters.com>

Omnetics Connector Corporation

7260 Commerce Circle
East Minneapolis, MN 55432
(800) 343-0025 (763) 572-0656
Fax: (763) 572-3925
<http://www.omnetics.com/main.htm>

Omni International

6530 Commerce Court
Warrenton, VA 20187
(800) 776-4431 FAX: (540) 347-5352
(540) 347-5331
<http://www.omni-inc.com>

Omnion

2010 Energy Drive
P.O. Box 879
East Troy, WI 53120
(262) 642-7200 FAX: (262) 642-7760
<http://www.omnion.com>

Omnitech Electronics

See AccuScan Instruments

Oncogene Research Products

P.O. Box 12087
La Jolla, CA 92039-2087
(800) 662-2616 FAX: (800) 766-0999
<http://www.apoptosis.com>

Oncogene Science

See OSI Pharmaceuticals

Oncor

See Intergen

Online Instruments

130 Conway Drive, Suites A & B
Bogart, GA 30622
(800) 852-3504 (706) 353-1972
(706) 353-6547
<http://www.olisweb.com>

Operon Technologies

1000 Atlantic Avenue
Alameda, CA 94501
(800) 688-2248 FAX: (510) 865-5225
(510) 865-8644
<http://www.operon.com>

Optiscan

P.O. Box 1066
Mount Waverly MDC, Victoria
Australia 3149
61-3-9538 3333 FAX: 61-3-9562 7742
<http://www.optiscan.com.au>

Optomax

9 Ash Street
P.O. Box 840
Hollis, NH 03049
(603) 465-3385 FAX: (603) 465-2291

Opto-Line Associates

265 Ballardvale Street
Wilmington, MA 01887
(978) 658-7255 FAX: (978) 658-7299
<http://www.optoline.com>

Orbigen

6827 Nancy Ridge Drive
San Diego, CA 92121
(866) 672-4436 (858) 362-2030
(858) 362-2026
<http://www.orbigen.com>

Oread BioSaftey

1501 Wakarusa Drive
Lawrence, KS 66047
(800) 447-6501 FAX: (785) 749-1882
(785) 749-0034
<http://www.oread.com>

Organomation Associates

266 River Road West
Berlin, MA 01503
(888) 978-7300 FAX: (978) 838-2786
(978) 838-7300
<http://www.organomation.com>

Organon

375 Mount Pleasant Avenue
West Orange, NJ 07052
(800) 241-8812 FAX: (973) 325-4589
(973) 325-4500
<http://www.organon.com>

Organon Teknika (Canada)

30 North Wind Place
Scarborough, Ontario
M1S 3R5 Canada
(416) 754-4344 FAX: (416) 754-4488
<http://www.organonteknika.com>

Organon Teknika Cappel

100 Akzo Avenue
Durham, NC 27712
(800) 682-2666 FAX: (800) 432-9682
(919) 620-2000 FAX: (919) 620-2107
<http://www.organonteknika.com>

Oriel Corporation of America

150 Long Beach Boulevard
Stratford, CT 06615
(203) 377-8282 FAX: (203) 378-2457
<http://www.oriel.com>

OriGene Technologies

6 Taft Court, Suite 300
Rockville, MD 20850
(888) 267-4436 FAX: (301) 340-9254
(301) 340-3188
<http://www.origene.com>

OriginLab

One Roundhouse Plaza
Northampton, MA 01060
(800) 969-7720 FAX: (413) 585-0126
<http://www.originlab.com>

Orion Research

500 Cummings Center
Beverly, MA 01915
(800) 225-1480 FAX: (978) 232-6015
(978) 232-6000
<http://www.orionres.com>

Ortho Diagnostic Systems

Subsidiary of Johnson & Johnson
1001 U.S. Highway 202
P.O. Box 350
Raritan, NJ 08869
(800) 322-6374 FAX: (908) 218-8582
(908) 218-1300

Ortho McNeil Pharmaceutical

Welsh & McKean Road
Spring House, PA 19477
(800) 682-6532
(215) 628-5000
<http://www.orthomcneil.com>

Oryza

200 Turnpike Road, Unit 5
Chelmsford, MA 01824
(978) 256-8183 FAX: (978) 256-7434
<http://www.oryzalabs.com>

OSI Pharmaceuticals

106 Charles Lindbergh Boulevard
Uniondale, NY 11553
(800) 662-2616 FAX: (516) 222-0114
(516) 222-0023
<http://www.osip.com>

Osmomics

135 Flanders Road
P.O. Box 1046
Westborough, MA 01581
(800) 444-8212 FAX: (508) 366-5840
(508) 366-8212
<http://www.osmolabstore.com>

Oster Professional Products

150 Cadillac Lane
McMinnville, TN 37110
(931) 668-4121 FAX: (931) 668-4125
<http://www.sunbeam.com>

Out Patient Services

1260 Holm Road
Petaluma, CA 94954
(800) 648-1666 FAX: (707) 762-7198
(707) 763-1581

Owens and Minor

4800 Cox Road
Glen Allen, VA 23060
(804) 747-9794 Fax: 804-270-7281
<http://www.owens-minor.com>

OWL Scientific Plastics

See OWL Separation Systems

OWL Separation Systems

55 Heritage Avenue
Portsmouth, NH 03801
(800) 242-5560 FAX: (603) 559-9258
(603) 559-9297
<http://www.owlsci.com>

Oxford Biochemical Research

P.O. Box 522
Oxford, MI 48371
(800) 692-4633 FAX: (248) 852-4466
<http://www.oxfordbiomed.com>

Oxford GlycoSystems

See Glyco

Suppliers**A.4.24**

Oxford Instruments
Old Station Way
Eynsham
Witney, Oxfordshire OX8 1TL, UK
(44) 1865-881437
FAX: (44) 1865-881944
<http://www.oxinst.com>

Oxford Labware
See Kendall

Oxford Molecular Group
Oxford Science Park
The Medawar Centre
Oxford OX4 4GA, UK
(44) 1865-784600
FAX: (44) 1865-784601
<http://www.oxmol.co.uk>

Oxford Molecular Group
2105 South Bascom Avenue, Suite 200
Campbell, CA 95008
(800) 876-9994 FAX: (408) 879-6302
(408) 879-6300
<http://www.oxmol.com>

OXIS International
6040 North Cutter Circle
Suite 317
Portland, OR 97217
(800) 547-3686 FAX: (503) 283-4058
(503) 283-3911
<http://www.oxis.com>

Oxoid
800 Proctor Avenue
Ogdensburg, NY 13669
(800) 567-8378 FAX: (613) 226-3728
<http://www.oxoid.ca>

Oxoid
Wade Road
Basingstoke, Hampshire RG24 8PW, UK
(44) 1256-841144
FAX: (4) 1256-814626
<http://www.oxoid.ca>

Oxyrase
P.O. Box 1345
Mansfield, OH 44901
(419) 589-8800 FAX: (419) 589-9919
<http://www.oxyrase.com>

Ozyme
10 Avenue Ampère
Montigny de Bretonneux
78180 France
(33) 13-46-02-424
FAX: (33) 13-46-09-212
<http://www.ozyme.fr>

PAA Laboratories
2570 Route 724
P.O. Box 435
Parker Ford, PA 19457
(610) 495-9400 FAX: (610) 495-9410
<http://www.paa-labs.com>

Pacer Scientific
5649 Valley Oak Drive
Los Angeles, CA 90068
(323) 462-0636 FAX: (323) 462-1430
<http://www.pacersci.com>

Pacific Bio-Marine Labs
P.O. Box 1348
Venice, CA 90294
(310) 677-1056 FAX: (310) 677-1207

Packard Instrument
800 Research Parkway
Meriden, CT 06450
(800) 323-1891 FAX: (203) 639-2172
(203) 238-2351
<http://www.packardinst.com>

Padgett Instrument
1730 Walnut Street
Kansas City, MO 64108
(816) 842-1029

Pall Filtron
50 Bearfoot Road
Northborough, MA 01532
(800) FILTRON FAX: (508) 393-1874
(508) 393-1800

Pall-Gelman
25 Harbor Park Drive
Port Washington, NY 11050
(800) 289-6255 FAX: (516) 484-2651
(516) 484-3600
<http://www.pall.com>

PanVera
545 Science Drive
Madison, WI 53711
(800) 791-1400 FAX: (608) 233-3007
(608) 233-9450
<http://www.panvera.com>

Parke-Davis
See Warner-Lambert

Parr Instrument
211 53rd Street
Moline, IL 61265
(800) 872-7720 FAX: (309) 762-9453
(309) 762-7716
<http://www.parrinst.com>

Partec
Otto Hahn Strasse 32
D-48161 Munster, Germany
(49) 2534-8008-0
FAX: (49) 2535-8008-90

PCR
See Archimica Florida

PE Biosystems
850 Lincoln Centre Drive
Foster City, CA 94404
(800) 345-5224 FAX: (650) 638-5884
(650) 638-5800
<http://www.pebio.com>

Pel-Freez Biologicals
219 N. Arkansas
P.O. Box 68
Rogers, AR 72757
(800) 643-3426 FAX: (501) 636-3562
(501) 636-4361
<http://www.pelfreez-bio.com>

Pel-Freez Clinical Systems
Subsidiary of Pel-Freez Biologicals
9099 N. Deerbrook Trail
Brown Deer, WI 53223
(800) 558-4511 FAX: (414) 357-4518
(414) 357-4500
<http://www.pelfreez-bio.com>

Peninsula Laboratories
601 Taylor Way
San Carlos, CA 94070
(800) 650-4442 FAX: (650) 595-4071
(650) 592-5392
<http://www.penlabs.com>

Pentex
24562 Mando Drive
Laguna Niguel, CA 92677
(800) 382-4667 FAX: (714) 643-2363
<http://www.pentex.com>

PeptoTech
5 Crescent Avenue
P.O. Box 275
Rocky Hill, NJ 08553
(800) 436-9910 FAX: (609) 497-0321
(609) 497-0253
<http://www.peprotech.com>

Peptide Institute
4-1-2 Ina, Minoh-shi
Osaka 562-8686, Japan
81-727-29-4121 FAX: 81-727-29-4124
<http://www.peptide.co.jp>

Peptide Laboratory
4175 Lakeside Drive
Richmond, CA 94806
(800) 858-7322 FAX: (510) 262-9127
(510) 262-0800
<http://www.peptidlab.com>

Peptides International
11621 Electron Drive
Louisville, KY 40299
(800) 777-4779 FAX: (502) 267-1329
(502) 266-8787
<http://www.pepnet.com>

Perceptive Science Instruments
2525 South Shore Boulevard, Suite 100
League City, TX 77573
(281) 334-3027 FAX: (281) 538-2222
<http://www.persci.com>

Perimed
4873 Princeton Drive
North Royalton, OH 44133
(440) 877-0537 FAX: (440) 877-0534
<http://www.perimed.se>

Perkin-Elmer
761 Main Avenue
Norwalk, CT 06859
(800) 762-4002 FAX: (203) 762-6000
(203) 762-1000
<http://www.perkin-elmer.com>
See also PE Biosystems

PerSeptive Bioresearch Products
See PerSeptive BioSystems

PerSeptive BioSystems
500 Old Connecticut Path
Framingham, MA 01701
(800) 899-5858 FAX: (508) 383-7885
(508) 383-7700
<http://www.pbio.com>

PerSeptive Diagnostic
See PE Biosystems
(800) 343-1346

Petersson Elektronik AB
Tallbacksvagen 51
S-756 45 Uppsala, Sweden
(46) 1830-3880 FAX: (46) 1830-3840
<http://www.bahnhof.se/~petersson>

Pfanstiehl Laboratories, Inc.
1219 Glen Rock Avenue
Waukegan, IL 60085
(800) 383-0126 FAX: (847) 623-9173
<http://www.pfanstiehl.com>

PGC Scientifics
7311 Governors Way
Frederick, MD 21704
(800) 424-3300 FAX: (800) 662-1112
(301) 620-7777 FAX: (301) 620-7497
<http://www.pgcscientifics.com>

Pharmacia Biotech
See Amersham Pharmacia Biotech

Pharmacia Diagnostics
See Wallac

Pharmacia LKB Biotech
See Amersham Pharmacia Biotech

Pharmacia LKB Biotechnology
See Amersham Pharmacia Biotech

Pharmacia LKB Nuclear
See Wallac

Pharmaderm Veterinary Products
60 Baylis Road
Melville, NY 11747
(800) 432-6673
<http://www.pharmaderm.com>

Pharmed (Norton)
Norton Performance Plastics
See Saint-Gobain Performance Plastics

PharMingen
See BD PharMingen

Phenomex
2320 W. 205th Street
Torrance, CA 90501
(310) 212-0555 FAX: (310) 328-7768
<http://www.phenomex.com>

PHLS Centre for Applied
Microbiology and Research
See European Collection of Animal
Cell Cultures (ECACC)

Suppliers

A.4.25

Phoenix Flow Systems

11575 Sorrento Valley Road, Suite 208
San Diego, CA 92121
(800) 886-3569 FAX: (619) 259-5268
(619) 453-5095
<http://www.phnxflow.com>

Phoenix Pharmaceutical

4261 Easton Road, P.O. Box 6457
St. Joseph, MO 64506
(800) 759-3644 FAX: (816) 364-4969
(816) 364-5777
<http://www.phoenixpharmaceutical.com>

Photometrics

See Roper Scientific

Photon Technology International

1 Deerpark Drive, Suite F
Monmouth Junction, NJ 08852
(732) 329-0910 FAX: (732) 329-9069
<http://www.pti-nj.com>

Physik Instrumente

Polytec PI
23 Midstate Drive, Suite 212
Auburn, MA 01501
(508) 832-3456 FAX: (508) 832-0506
<http://www.polytecpi.com>

Physitemp Instruments

154 Huron Avenue
Clifton, NJ 07013
(800) 452-8510 FAX: (973) 779-5954
(973) 779-5577
<http://www.physitemp.com>

Pico Technology

The Mill House, Cambridge Street
St. Neots, Cambridgeshire
PE19 1QB, UK
(44) 1480-396-395
FAX: (44) 1480-396-296
<http://www.picotech.com>

Pierce Chemical

P.O. Box 117
3747 Meridian Road
Rockford, IL 61105
(800) 874-3723 FAX: (800) 842-5007
FAX: (815) 968-7316
<http://www.piercenet.com>

Pierce & Warriner

44, Upper Northgate Street
Chester, Cheshire CH1 4EF, UK
(44) 1244 382 525
FAX: (44) 1244 373 212
<http://www.piercenet.com>

Pilling Weck Surgical

420 Delaware Drive
Fort Washington, PA 19034
(800) 523-2579 FAX: (800) 332-2308
<http://www.pilling-weck.com>

PixelVision

A division of Cybex Computer Products
14964 NW Greenbrier Parkway
Beaverton, OR 97006
(503) 629-3210 FAX: (503) 629-3211
<http://www.pixelvision.com>

P.J. Noyes

P.O. Box 381
89 Bridge Street
Lancaster, NH 03584
(800) 522-2469 FAX: (603) 788-3873
(603) 788-4952
<http://www.pjnoyes.com>

Plas-Labs

917 E. Chilson Street
Lansing, MI 48906
(800) 866-7527 FAX: (517) 372-2857
(517) 372-7177
<http://www.plas-labs.com>

Plastics One

6591 Merriman Road, Southwest
P.O. Box 12004
Roanoke, VA 24018
(540) 772-7950 FAX: (540) 989-7519
<http://www.plastics1.com>

Platt Electric Supply

2757 6th Avenue South
Seattle, WA 98134
(206) 624-4083 FAX: (206) 343-6342
<http://www.platt.com>

Plexon

6500 Greenville Avenue
Suite 730
Dallas, TX 75206
(214) 369-4957 FAX: (214) 369-1775
<http://www.plexoninc.com>

Polaroid

784 Memorial Drive
Cambridge, MA 01239
(800) 225-1618 FAX: (800) 832-9003
(781) 386-2000
<http://www.polaroid.com>

Polyfiltronics

136 Weymouth St.
Rockland, MA 02370
(800) 434-7659 FAX: (781) 878-0822
(781) 878-1133
<http://www.polyfiltronics.com>

Polylabo Paul Block

Parc Tertiaire de la Meinau
10, rue de la Durance
B.P. 36
67023 Strasbourg Cedex 1
Strasbourg, France
33-3-8865-8020
FAX: 33-3-8865-8039

PolyLC

9151 Rumsey Road, Suite 180
Columbia, MD 21045
(410) 992-5400 FAX: (410) 730-8340

Polymer Laboratories

Amherst Research Park
160 Old Farm Road
Amherst, MA 01002
(800) 767-3963 FAX: (413) 253-2476
<http://www.polymerlabs.com>

Polymicro Technologies

18019 North 25th Avenue
Phoenix, AZ 85023
(602) 375-4100 FAX: (602) 375-4110
<http://www.polymicro.com>

Polyphenols AS

Hanabryggene Technology Centre
Hanaveien 4-6
4327 Sandnes, Norway
(47) 51-62-0990
FAX: (47) 51-62-51-82
<http://www.polyphenols.com>

Polysciences

400 Valley Road
Warrington, PA 18976
(800) 523-2575 FAX: (800) 343-3291
<http://www.polysciences.com>

Polyscientific

70 Cleveland Avenue
Bayshore, NY 11706
(516) 586-0400 FAX: (516) 254-0618

Polytech Products

285 Washington Street
Somerville, MA 02143
(617) 666-5064 FAX: (617) 625-0975

Polytron

8585 Grovemont Circle
Gaithersburg, MD 20877
(301) 208-6597 FAX: (301) 208-8691
<http://www.polytron.com>

Popper and Sons

300 Denton Avenue
P.O. Box 128
New Hyde Park, NY 11040
(888) 717-7677 FAX: (800) 557-6773
(516) 248-0300 FAX: (516) 747-1188
<http://www.popperandsons.com>

Porphyrim Products

P.O. Box 31
Logan, UT 84323
(435) 753-1901 FAX: (435) 753-6731
<http://www.porphyrin.com>

Portex

See SIMS Portex Limited

Powderject Vaccines

585 Science Drive
Madison, WI 53711
(608) 231-3150 FAX: (608) 231-6990
<http://www.powderject.com>

Praxair

810 Jorie Boulevard
Oak Brook, IL 60521
(800) 621-7100
<http://www.praxair.com>

Precision Dynamics

13880 Del Sur Street
San Fernando, CA 91340
(800) 847-0670 FAX: (818) 899-4-45
<http://www.pdcorp.com>

Precision Scientific Laboratory

Equipment
Division of Jouan
170 Marcel Drive
Winchester, VA 22602
(800) 621-8820 FAX: (540) 869-0130
(540) 869-9892
<http://www.precisionsci.com>

Primary Care Diagnostics

See Becton Dickinson Primary
Care Diagnostics

Primate Products

1755 East Bayshore Road, Suite 28A
Redwood City, CA 94063
(650) 368-0663 FAX: (650) 368-0665
<http://www.primateproducts.com>

5 Prime → 3 Prime

See 2000 Eppendorf-5 Prime
<http://www.5prime.com>

Princeton Applied Research

PerkinElmer Instr.: Electrochemistry
801 S. Illinois
Oak Ridge, TN 37830
(800) 366-2741 FAX: (423) 425-1334
(423) 481-2442
<http://www.eggpar.com>

Princeton Instruments

A division of Roper Scientific
3660 Quakerbridge Road
Trenton, NJ 08619
(609) 587-9797 FAX: (609) 587-1970
<http://www.prinst.com>

Princeton Separations

P.O. Box 300
Aldephia, NJ 07710
(800) 223-0902 FAX: (732) 431-3768
(732) 431-3338

Prior Scientific

80 Reservoir Park Drive
Rockland, MA 02370
(781) 878-8442 FAX: (781) 878-8736
<http://www.prior.com>

PRO Scientific

P.O. Box 448
Monroe, CT 06468
(203) 452-9431 FAX: (203) 452-9753
<http://www.proscientific.com>

Professional Compounding Centers of America

9901 South Wilcrest Drive
Houston, TX 77099
(800) 331-2498 FAX: (281) 933-6227
(281) 933-6948
<http://www.pccarx.com>

Progen Biotechnik

Maass-Str. 30
69123 Heidelberg, Germany
(49) 6221-8278-0
FAX: (49) 6221-8278-23
<http://www.progen.de>

Suppliers**A.4.26**

Prolabo

A division of Merck Eurolab
54 rue Roger Salengro
94126 Fontenay Sous Bois Cedex
France
33-1-4514-8500
FAX: 33-1-4514-8616
<http://www.prolabo.fr>

Proligo

2995 Wilderness Place
Boulder, CO 80301
(888) 80-OLIGO FAX: (303) 801-1134
<http://www.proligo.com>

Promega

2800 Woods Hollow Road
Madison, WI 53711
(800) 356-9526 FAX: (800) 356-1970
(608) 274-4330 FAX: (608) 277-2516
<http://www.promega.com>

Protein Databases (PDI)

405 Oakwood Road
Huntington Station, NY 11746
(800) 777-6834 FAX: (516) 673-4502
(516) 673-3939

Protein Polymer Technologies

10655 Sorrento Valley Road
San Diego, CA 92121
(619) 558-6064 FAX: (619) 558-6477
<http://www.ppti.com>

Protein Solutions

391 G Chipeta Way
Salt Lake City, UT 84108
(801) 583-9301 FAX: (801) 583-4463
<http://www.proteinsolutions.com>

Prozyme

1933 Davis Street, Suite 207
San Leandro, CA 94577
(800) 457-9444 FAX: (510) 638-6919
(510) 638-6900
<http://www.prozyme.com>

PSI

See Perceptive Science Instruments

Pulmetrics Group

82 Beacon Street
Chestnut Hill, MA 02167
(617) 353-3833 FAX: (617) 353-6766

Purdue Frederick

100 Connecticut Avenue
Norwalk, CT 06850
(800) 633-4741 FAX: (203) 838-1576
(203) 853-0123
<http://www.pharma.com>

Purina Mills

LabDiet
P. O. Box 66812
St. Louis, MO 63166
(800) 227-8941 FAX: (314) 768-4894
<http://www.purina-mills.com>

Qbiogene

2251 Rutherford Road
Carlsbad, CA 92008
(800) 424-6101 FAX: (760) 918-9313
<http://www.qbiogene.com>

Qiagen

28159 Avenue Stanford
Valencia, CA 91355
(800) 426-8157 FAX: (800) 718-2056
<http://www.qiagen.com>

Quality Biological

7581 Lindbergh Drive
Gaithersburg, MD 20879
(800) 443-9331 FAX: (301) 840-5450
(301) 840-9331
<http://www.qualitybiological.com>

Quantitative Technologies

P.O. Box 470
Salem Industrial Park, Bldg. 5
Whitehouse, NJ 08888
(908) 534-4445 FAX: 534-1054
<http://www.qtionline.com>

Quantum Appligene

Parc d'Innovation
Rue Geller de Kayserberg
67402 Illkirch, Cedex, France
(33) 3-8867-5425
FAX: (33) 3-8867-1945
<http://www.quantum-appligene.com>

Quantum Biotechnologies

See Qbiogene

Quantum Soft

Postfach 6613
CH-8023
Zürich, Switzerland
FAX: 41-1-481-69-51
profit@quansoft.com

Quest Scientific

1177 West 21st Street
North Vancouver, BC V7R 2C7 Canada
(604) 984-7878
<http://www.quest-sci.com>

Questcor Pharmaceuticals

26118 Research Road
Hayward, CA 94545
(510) 732-5551 FAX: (510) 732-7741
<http://www.questcor.com>

Quidel

10165 McKellar Court
San Diego, CA 92121
(800) 874-1517 FAX: (858) 546-8955
(858) 552-1100
<http://www.quidel.com>

R-Biopharm

7950 Old US 27 South
Marshall, MI 49068
(616) 789-3033 FAX: (616) 789-3070
<http://www.r-biopharm.com>

R. C. Electronics

6464 Hollister Avenue
Santa Barbara, CA 93117
(805) 685-7770 FAX: (805) 685-5853
<http://www.rcelectronics.com>

R & D Systems

614 McKinley Place NE
Minneapolis, MN 55413
(800) 343-7475 FAX: (612) 379-6580
(612) 379-2956
<http://www.rndsyste.ms.com>

R & S Technology

350 Columbia Street
Peacedale, RI 02880
(401) 789-5660 FAX: (401) 792-3890
<http://www.septech.com>

RACAL Health and Safety

See 3M
7305 Executive Way
Frederick, MD 21704
(800) 692-9500 FAX: (301) 695-8200

Radiometer America

811 Sharon Drive
Westlake, OH 44145
(800) 736-0600 FAX: (440) 871-2633
(440) 871-8900
<http://www.rameusa.com>

Radiometer A/S

The Chemical Reference Laboratory
kandevej 21
DK-2700 Brnshj, Denmark
45-3827-3827 FAX: 45-3827-2727

Radionics

22 Terry Avenue
Burlington, MA 01803
(781) 272-1233 FAX: (781) 272-2428
<http://www.radionics.com>

Radnoti Glass Technology

227 W. Maple Avenue
Monrovia, CA 91016
(800) 428-1416 FAX: (626) 303-2998
(626) 357-8827
<http://www.radnoti.com>

Rainin Instrument

Rainin Road
P.O. Box 4026
Woburn, MA 01888
(800)-4-RAININ FAX: (781) 938-1152
(781) 935-3050
<http://www.rainin.com>

Rank Brothers

56 High Street
Bottisham, Cambridge
CB5 9DA UK
(44) 1223 811369
FAX: (44) 1223 811441
<http://www.rankbrothers.com>

Rapp Polymere

Ernst-Simon Strasse 9
D 72072 Tübingen, Germany
(49) 7071-763157
FAX: (49) 7071-763158
<http://www.rapp-polymere.com>

Raven Biological Laboratories

8607 Park Drive
P.O. Box 27261
Omaha, NE 68127
(800) 728-5702 FAX: (402) 593-0995
(402) 593-0781
<http://www.ravenlabs.com>

Razel Scientific Instruments

100 Research Drive
Stamford, CT 06906
(203) 324-9914 FAX: (203) 324-5568

RBI

See Research Biochemicals

Reagents International

See Biotech Source

Receptor Biology

10000 Virginia Manor Road, Suite 360
Beltsville, MD 20705
(888) 707-4200 FAX: (301) 210-6266
(301) 210-4700
<http://www.receptorbiology.com>

Regis Technologies

8210 N. Austin Avenue
Morton Grove, IL 60053
(800) 323-8144 FAX: (847) 967-1214
(847) 967-6000
<http://www.registech.com>

Reichert Ophthalmic Instruments

P.O. Box 123
Buffalo, NY 14240
(716) 686-4500 FAX: (716) 686-4545
<http://www.reichert.com>

Reiss

1 Polymer Place
P.O. Box 60
Blackstone, VA 23824
(800) 356-2829 FAX: (804) 292-1757
(804) 292-1600
<http://www.reissmfg.com>

Remel

12076 Santa Fe Trail Drive
P.O. Box 14428
Shawnee Mission, KS 66215
(800) 255-6730 FAX: (800) 621-8251
(913) 888-0939 FAX: (913) 888-5884
<http://www.remelinc.com>

Reming Bioinstruments

6680 County Route 17
Redfield, NY 13437
(315) 387-3414 FAX: (315) 387-3415

RepliGen

117 Fourth Avenue
Needham, MA 02494
(800) 622-2259 FAX: (781) 453-0048
(781) 449-9560
<http://www.repligen.com>

Research Biochemicals

1 Strathmore Road
Natick, MA 01760
(800) 736-3690 FAX: (800) 736-2480
(508) 651-8151 FAX: (508) 655-1359
<http://www.resbio.com>

Suppliers**A.4.27**

Research Corporation Technologies
101 N. Wilmot Road, Suite 600
Tucson, AZ 85711
(520) 748-4400 FAX: (520) 748-0025
<http://www.rctech.com>

Research Diagnostics
Pleasant Hill Road
Flanders, NJ 07836
(800) 631-9384 FAX: (973) 584-0210
(973) 584-7093
<http://www.researchd.com>

Research Diets
121 Jersey Avenue
New Brunswick, NJ 08901
(877) 486-2486 FAX: (732) 247-2340
(732) 247-2390
<http://www.researchdiets.com>

Research Genetics
2130 South Memorial Parkway
Huntsville, AL 35801
(800) 533-4363 FAX: (256) 536-9016
(256) 533-4363
<http://www.resgen.com>

Research Instruments
Kernick Road Perryn
Cornwall TR10 9DQ, UK
(44) 1326-372-753
FAX: (44) 1326-378-783
<http://www.research-instruments.com>

Research Organics
4353 E. 49th Street
Cleveland, OH 44125
(800) 321-0570 FAX: (216) 883-1576
(216) 883-8025
<http://www.resorg.com>

Research Plus
P.O. Box 324
Bayonne, NJ 07002
(800) 341-2296 FAX: (201) 823-9590
(201) 823-3592
<http://www.researchplus.com>

Research Products International
410 N. Business Center Drive
Mount Prospect, IL 60056
(800) 323-9814 FAX: (847) 635-1177
(847) 635-7330
<http://www.rpicorp.com>

Research Triangle Institute
P.O. Box 12194
Research Triangle Park, NC 27709
(919) 541-6000 FAX: (919) 541-6515
<http://www.rti.org>

Restek
110 Benner Circle
Bellefonte, PA 16823
(800) 356-1688 FAX: (814) 353-1309
(814) 353-1300
<http://www.restekcorp.com>

Rheodyne
P.O. Box 1909
Rohnert Park, CA 94927
(707) 588-2000 FAX: (707) 588-2020
<http://www.rheodyne.com>

Rhone Merieux
See Merial Limited

Rhone-Poulenc
2 T W Alexander Drive
P.O. Box 12014
Research Triangle Park, NC 08512
(919) 549-2000 FAX: (919) 549-2839
<http://www.Rhone-Poulenc.com>
Also see Aventis

Rhone-Poulenc Rorer
500 Arcola Road
Collegeville, PA 19426
(800) 727-6737 FAX: (610) 454-8940
(610) 454-8975
<http://www.rp-roser.com>

Rhone-Poulenc Rorer
Centre de Recherche de Vitry-Alfortville
13 Quai Jules Guesde, BP14 94403
Vitry Sur Seine, Cedex, France
(33) 145-73-85-11
FAX: (33) 145-73-81-29
<http://www.rp-roser.com>

Ribi ImmunoChem Research
563 Old Corvallis Road
Hamilton, MT 59840
(800) 548-7424 FAX: (406) 363-6129
(406) 363-3131
<http://www.ribi.com>

RiboGene
See Questcor Pharmaceuticals

Ricca Chemical
448 West Fork Drive
Arlington, TX 76012
(888) GO-RICCA FAX: (800) RICCA-93
(817) 461-5601
<http://www.riccachemical.com>

Richard-Allan Scientific
225 Parsons Street
Kalamazoo, MI 49007
(800) 522-7270 FAX: (616) 345-3577
(616) 344-2400
<http://www.rallansci.com>

Richelieu Biotechnologies
11 177 Hamon
Montral, Quebec
H3M 3E4 Canada
(802) 863-2567 FAX: (802) 862-2909
<http://www.richelieubio.com>

Richter Enterprises
20 Lake Shore Drive
Wayland, MA 01778
(508) 655-7632 FAX: (508) 652-7264
<http://www.richter-enterprises.com>

Riese Enterprises
BioSure Division
12301 G Loma Rica Drive
Grass Valley, CA 95945
(800) 345-2267 FAX: (916) 273-5097
(916) 273-5095
<http://www.biosure.com>

Ritz Safety Equipment, Inc.
1851 North Powerline Road
Pompano Beach, FL 33069
(954) 971-3176 Fax: 954-971-1272
www.vandyecrotty.com/first_aid_safety/index_safety_products.asp

Robbins Scientific
1250 Elko Drive
Sunnyvale, CA 94086
(800) 752-8585 FAX: (408) 734-0300
(408) 734-8500
<http://www.robsci.com>

Roboz Surgical Instruments
9210 Corporate Boulevard, Suite 220
Rockville, MD 20850
(800) 424-2984 FAX: (301) 590-1290
(301) 590-0055

Roche Diagnostics
9115 Hague Road
P.O. Box 50457
Indianapolis, IN 46256
(800) 262-1640 FAX: (317) 845-7120
(317) 845-2000
<http://www.roche.com>

Roche Molecular Systems
See Roche Diagnostics

Rocklabs
P.O. Box 18-142
Auckland 6, New Zealand
(64) 9-634-7696
FAX: (64) 9-634-7696
<http://www.rocklabs.com>

Rockland
P.O. Box 316
Gilbertsville, PA 19525
(800) 656-ROCK FAX: (610) 367-7825
(610) 369-1008
<http://www.rockland-inc.com>

Rockwell Laser Industries
P.O. Box 43010
7754 Camargo Road
Cincinnati, OH 45243
(513) 271-1568 Fax: 513-271-1598
www.rli.com

Rohm
Chemische Fabrik
Kirschenallee
D-64293 Darmstadt, Germany
(49) 6151-1801 FAX: (49) 6151-1802
<http://www.roehm.com>

Roper Scientific
3440 East Britannia Drive, Suite 100
Tucson, AZ 85706
(520) 889-9933 FAX: (520) 573-1944
<http://www.roperscientific.com>

Rosetta Inpharmatics
12040 115th Avenue NE
Kirkland, WA 98034
(425) 820-8900 FAX: (425) 820-5757
<http://www.rii.com>

ROTH-SOCHIEL
3 rue de la Chapelle
Lauterbourg
67630 France
(33) 03-88-94-82-42
FAX: (33) 03-88-54-63-93

Rotronic Instrument
160 E. Main Street
Huntington, NY 11743
(631) 427-3898 FAX: (631) 427-3902
<http://www.rotrotron-usa.com>

Roundy's
23000 Roundy Drive
Pewaukee, WI 53072
(262) 953-7999 FAX: (262) 953-7989
<http://www.roundys.com>

RS Components
Birchington Road
Weldon Industrial Estate
Corby, Northants NN17 9RS, UK
(44) 1536 201234
FAX: (44) 1536 405678
<http://www.rs-components.com>

Rubbermaid
See Newell Rubbermaid

SA Instrumentation
1437 Tzena Way
Encinitas, CA 92024
(858) 453-1776 FAX: (800) 266-1776
<http://www.sainst.com>

Safe Cells
See Bionique Testing Labs

Sage Instruments
240 Airport Boulevard
Freedom, CA 95076
831-761-1000 FAX: 831-761-1008
<http://www.sageinst.com>

Sage Laboratories
11 Huron Drive
Natick, MA 01760
(508) 653-0844 FAX: 508-653-5671
<http://www.sagelabs.com>

Saint-Gobain Performance Plastics
P.O. Box 3660
Akron, OH 44309
(330) 798-9240 FAX: (330) 798-6968
<http://www.nortonplastics.com>

San Diego Instruments
7758 Arjons Drive
San Diego, CA 92126
(858) 530-2600 FAX: (858) 530-2646
<http://www.sd-inst.com>

Sandown Scientific
Beards Lodge
25 Oldfield Road
Hampden, Middlesex TW12 2AJ, UK
(44) 2089 793300
FAX: (44) 2089 793311
<http://www.sandownsci.com>

Sandoz Pharmaceuticals
See Novartis

Suppliers

A.4.28

Sanofi Recherche
Centre de Montpellier
371 Rue du Professor Blayac
34184 Montpellier, Cedex 04
France
(33) 67-10-67-10
FAX: (33) 67-10-67-67

Sanofi Winthrop Pharmaceuticals
90 Park Avenue
New York, NY 10016
(800) 223-5511 FAX: (800) 933-3243
(212) 551-4000
<http://www.sanofi-synthelabo.com/us>

Santa Cruz Biotechnology
2161 Delaware Avenue
Santa Cruz, CA 95060
(800) 457-3801 FAX: (831) 457-3801
(831) 457-3800
<http://www.scbt.com>

Sarasep
(800) 605-0267 FAX: (408) 432-3231
(408) 432-3230
<http://www.transgenomic.com>

Sarstedt
P.O. Box 468
Newton, NC 28658
(800) 257-5101 FAX: (828) 465-4003
(828) 465-4000
<http://www.sarstedt.com>

Sartorius
131 Heartsland Boulevard
Edgewood, NY 11717
(800) 368-7178 FAX: (516) 254-4253
<http://www.sartorius.com>

SAS Institute
Pacific Telesis Center
One Montgomery Street
San Francisco, CA 94104
(415) 421-2227 FAX: (415) 421-1213
<http://www.sas.com>

Savant/EC Apparatus
A ThermoQuest company
100 Colin Drive
Hollbrook, NY 11741
(800) 634-8886 FAX: (516) 244-0606
(516) 244-2929
<http://www.savec.com>

Savillex
6133 Baker Road
Minnetonka, MN 55345
(612) 935-5427

Scanalytics
Division of CSP
8550 Lee Highway, Suite 400
Fairfax, VA 22031
(800) 325-3110 FAX: (703) 208-1960
(703) 208-2230
<http://www.scanalytics.com>

Schering Laboratories
See Schering-Plough

Schering-Plough
1 Giralda Farms
Madison, NJ 07940
(800) 222-7579 FAX: (973) 822-7048
(973) 822-7000
<http://www.schering-plough.com>

Schleicher & Schuell
10 Optical Avenue
Keene, NH 03431
(800) 245-4024 FAX: (603) 357-3627
(603) 352-3810
<http://www.s-und-s.de/english-index.html>

Science Technology Centre
1250 Herzberg Laboratories
Carleton University
1125 Colonel Bay Drive
Ottawa, Ontario
K1S 5B6 Canada
(613) 520-4442 FAX: (613) 520-4445
<http://www.carleton.ca/universities/stc>

Scientific Instruments
200 Saw Mill River Road
Hawthorne, NY 10532
(800) 431-1956 FAX: (914) 769-5473
(914) 769-5700
<http://www.scientificinstruments.com>

Scientific Solutions
9323 Hamilton
Mentor, OH 44060
(440) 357-1400 FAX: (440) 357-1416
<http://www.labmaster.com>

Scion
82 Worman's Mill Court, Suite H
Frederick, MD 21701
(301) 695-7870 FAX: (301) 695-0035
<http://www.scioncorp.com>

Scott Specialty Gases
6141 Easton Road
P.O. Box 310
Plumsteadville, PA 18949
(800) 21-SCOTT FAX: (215) 766-2476
(215) 766-8861
<http://www.scottgas.com>

**Scripps Clinic and Research
Foundation**
Instrumentation and Design Lab
10666 N. Torrey Pines Road
La Jolla, CA 92037
(800) 992-9962 FAX: (858) 554-8986
(858) 455-9100
<http://www.scrippsclinic.com>

SDI Sensor Devices
407 Pilot Court, 400A
Waukesha, WI 53188
(414) 524-1000 FAX: (414) 524-1009

Sefar America
111 Calumet Street
Depew, NY 14043
(716) 683-4050 FAX: (716) 683-4053
<http://www.sefaramerica.com>

Seikagaku America
Division of Associates of Cape Cod
704 Main Street
Falmouth, MA 02540
(800) 237-4512 FAX: (508) 540-8680
(508) 540-3444
<http://www.seikagaku.com>

Sellas Medizinische Gerate
Hagener Str. 393
Gevelsberg-Vogelsang, 58285
Germany
(49) 23-326-1225

Sensor Medics
22705 Savi Ranch Parkway
Yorba Linda, CA 92887
(800) 231-2466 FAX: (714) 283-8439
(714) 283-2228
<http://www.sensormedics.com>

Sensor Systems LLC
2800 Anvil Street, North
Saint Petersburg, FL 33710
(800) 688-2181 FAX: (727) 347-3881
(727) 347-2181
<http://www.vsensors.com>

SenSym/Foxboro ICT
1804 McCarthy Boulevard
Milpitas, CA 95035
(800) 392-9934 FAX: (408) 954-9458
(408) 954-6700
<http://www.sensym.com>

Separations Group
See Vydac

Sepracor
111 Locke Drive
Marlboro, MA 01752
(877)-SEPRACOR (508) 357-7300
<http://www.sepracor.com>

Sera-Lab
See Harlan Sera-Lab

Sermeter
925 Seton Court, #7
Wheeling, IL 60090
(847) 537-4747

Serological
195 W. Birch Street
Kankakee, IL 60901
(800) 227-9412 FAX: (815) 937-8285
(815) 937-8270

Seromed Biochrom
Leonorenstrasse 2-6
D-12247 Berlin, Germany
(49) 030-779-9060

Serotec
22 Bankside
Station Approach
Kidlington, Oxford OX5 1JE, UK
(44) 1865-852722
FAX: (44) 1865-373899
In the US: (800) 265-7376
<http://www.serotec.co.uk>

Serva Biochemicals
Distributed by Crescent Chemical

S.F. Medical Pharmlast
See Chase-Walton Elastomers

SGE
2007 Kramer Lane
Austin, TX 78758
(800) 945-6154 FAX: (512) 836-9159
(512) 837-7190
<http://www.sge.com>

Shandon/Lipshaw
171 Industry Drive
Pittsburgh, PA 15275
(800) 245-6212 FAX: (412) 788-1138
(412) 788-1133
<http://www.shandon.com>

Sharpint
P.O. Box 2212
Taichung, Taiwan
Republic of China
(886) 4-3206320
FAX: (886) 4-3289879
<http://www.sharpint.com.tw>

Shelton Scientific
230 Longhill Crossroads
Shelton, CT 06484
(800) 222-2092 FAX: (203) 929-2175
(203) 929-8999
<http://www.sheltonscientific.com>

Sherwood-Davis & Geck
See Kendall

Sherwood Medical
See Kendall

Shimadzu Scientific Instruments
7102 Riverwood Drive
Columbia, MD 21046
(800) 477-1227 FAX: (410) 381-1222
(410) 381-1227
<http://www.ssi.shimadzu.com>

Sialomed
See Amika

Siemens Analytical X-Ray Systems
See Bruker Analytical X-Ray Systems

Sievers Instruments
Subsidiary of Ionics
6060 Spine Road
Boulder, CO 80301
(800) 255-6964 FAX: (303) 444-6272
(303) 444-2009
<http://www.sieversinst.com>

SIFCO
970 East 46th Street
Cleveland, OH 44103
(216) 881-8600 FAX: (216) 432-6281
<http://www.sifco.com>

Sigma-Aldrich
3050 Spruce Street
St. Louis, MO 63103
(800) 358-5287 FAX: (800) 962-9591
(800) 325-3101 FAX: (800) 325-5052
<http://www.sigma-aldrich.com>

Suppliers

A.4.29

Sigma-Aldrich Canada
2149 Winston Park Drive
Oakville, Ontario
L6H 6J8 Canada
(800) 5652D1400 FAX: (800)
2652D3858
<http://www.sigma-aldrich.com>

Silenus/Amrad
34 Wadhurst Drive
Boronia, Victoria 3155 Australia
(613)9887-3909 FAX: (613)9887-3912
<http://www.amrad.com.au>

Silicon Genetics
2601 Spring Street
Redwood City, CA 94063
(866) SIG SOFT FAX: (650) 365 1735
(650) 367 9600
<http://www.sigenetics.com>

SIMS Deltec
1265 Grey Fox Road
St. Paul, Minnesota 55112
(800) 426-2448 FAX: (615) 628-7459
<http://www.deltec.com>

SIMS Portex
10 Bowman Drive
Keene, NH 03431
(800) 258-5361 FAX: (603) 352-3703
(603) 352-3812
<http://www.simsmed.com>

SIMS Portex Limited
Hythe, Kent CT21 6JL, UK
(44)1303-260551
FAX: (44)1303-266761
<http://www.portex.com>

Siris Laboratories
See Biosearch Technologies

Skatron Instruments
See Molecular Devices

SLM Instruments
See Spectronic Instruments

SLM-AMINCO Instruments
See Spectronic Instruments

Small Parts
13980 NW 58th Court
P.O. Box 4650
Miami Lakes, FL 33014
(800) 220-4242 FAX: (800) 423-9009
(305) 558-1038 FAX: (305) 558-0509
<http://www.smallparts.com>

Smith & Nephew
11775 Starkey Road
P.O. Box 1970
Largo, FL 33779
(800) 876-1261
<http://www.smith-nephew.com>

SmithKline Beecham
1 Franklin Plaza, #1800
Philadelphia, PA 19102
(215) 751-4000 FAX: (215) 751-4992
<http://www.sb.com>

Solid Phase Sciences
See Biosearch Technologies

SOMA Scientific Instruments
5319 University Drive, PMB #366
Irvine, CA 92612
(949) 854-0220 FAX: (949) 854-0223
<http://somascientific.com>

Somatix Therapy
See Cell Genesys

Sonics & Materials
53 Church Hill Road
Newtown, CT 06470
(800) 745-1105 FAX: (203) 270-4610
(203) 270-4600
<http://www.sonicsandmaterials.com>

Sonosep Biotech
See Triton Environmental Consultants

Sorvall
See Kendro Laboratory Products

Southern Biotechnology Associates
P.O. Box 26221
Birmingham, AL 35260
(800) 722-2255 FAX: (205) 945-8768
(205) 945-1774
<http://SouthernBiotech.com>

SPAFAS
190 Route 165
Preston, CT 06365
(800) SPAFAS-1 FAX: (860) 889-1991
(860) 889-1389
<http://www.spafas.com>

Specialty Media
Division of Cell & Molecular Technologies
580 Marshall Street
Phillipsburg, NJ 08865
(800) 543-6029 FAX: (908) 387-1670
(908) 454-7774
<http://www.specialtymedia.com>

Spectra Physics
See Thermo Separation Products

Spectramed
See BOC Edwards

SpectraSource Instruments
31324 Via Colinas, Suite 114
Westlake Village, CA 91362
(818) 707-2655 FAX: (818) 707-9035
<http://www.spectrasource.com>

Spectronic Instruments
820 Linden Avenue
Rochester, NY 14625
(800) 654-9955 FAX: (716) 248-4014
(716) 248-4000
<http://www.spectronic.com>

Spectrum Medical Industries
See Spectrum Laboratories

Spectrum Laboratories
18617 Broadwick Street
Rancho Dominguez, CA 90220
(800) 634-3300 FAX: (800) 445-7330
(310) 885-4601 FAX: (310) 885-4666
<http://www.spectrumlabs.com>

Spherotech
1840 Industrial Drive, Suite 270
Libertyville, IL 60048
(800) 368-0822 FAX: (847) 680-8927
(847) 680-8922
<http://www.spherotech.com>

SPSS
233 S. Wacker Drive, 11th floor
Chicago, IL 60606
(800) 521-1337 FAX: (800) 841-0064
<http://www.spss.com>

SS White Burs
1145 Towbin Avenue
Lakewood, NJ 08701
(732) 905-1100 FAX: (732) 905-0987
<http://www.sswwhiteburs.com>

Stag Instruments
16 Monument Industrial Park
Chalgrove, Oxon OX44 7RW, UK
(44) 1865-891116
FAX: (44) 1865-890562

Standard Reference Materials
Program
National Institute of Standards and
Technology
Building 202, Room 204
Gaithersburg, MD 20899
(301) 975-6776 FAX: (301) 948-3730

Starna Cells
P.O. Box 1919
Atascadero, CA 93423
(805) 466-8855 FAX: (805) 461-1575
(800) 228-4482
<http://www.starnacells.com>

Starplex Scientific
50 Steinway
Etobicoke, Ontario
M9W 6Y3 Canada
(800) 665-0954 FAX: (416) 674-6067
(416) 674-7474
<http://www.starplexscientific.com>

State Laboratory Institute of
Massachusetts
305 South Street
Jamaica Plain, MA 02130
(617) 522-3700 FAX: (617) 522-8735
<http://www.state.ma.us/dph>

Stedim Labs
1910 Mark Court, Suite 110
Concord, CA 94520
(800) 914-6644 FAX: (925) 689-6988
(925) 689-6650
<http://www.stedim.com>

Steinel America
9051 Lyndale Avenue
Bloomington, MN 55420
(800) 852 4343 FAX: (952) 888-5132
<http://www.steinellamerica.com>

Stem Cell Technologies
777 West Broadway, Suite 808
Vancouver, British Columbia
V5Z 4J7 Canada
(800) 667-0322 FAX: (800) 567-2899
(604) 877-0713 FAX: (604) 877-0704
<http://www.stemcell.com>

Stephens Scientific
107 Riverdale Road
Riverdale, NJ 07457
(800) 831-8099 FAX: (201) 831-8009
(201) 831-9800

Steraloids
P.O. Box 689
Newport, RI 02840
(401) 848-5422 FAX: (401) 848-5638
<http://www.steraloids.com>

Steris Corporation
5960 Heisley Road
Mentor, Ohio 44060
(800) 548-4873
440-354-2600
<http://www.steris.com>

Sterling Medical
2091 Springdale Road, Ste. 2
Cherry Hill, NJ 08003
(800) 229-0900 FAX: (800) 229-7854
<http://www.sterlingmedical.com>

Sterling Winthrop
90 Park Avenue
New York, NY 10016
(212) 907-2000 FAX: (212) 907-3626

Sternberger Monoclonals
10 Burwood Court
Lutherville, MD 21093
(410) 821-8505 FAX: (410) 821-8506
<http://www.sternbergermonoclonals.com>

Stoelting
502 Highway 67
Kiel, WI 53042
(920) 894-2293 FAX: (920) 894-7029
<http://www.stoelting.com>

Stovall Lifescience
206-G South Westgate Drive
Greensboro, NC 27407
(800) 852-0102 FAX: (336) 852-3507
<http://www.slsience.com>

Stratagene
11011 N. Torrey Pines Road
La Jolla, CA 92037
(800) 424-5444 FAX: (888) 267-4010
(858) 535-5400
<http://www.stratagene.com>

Strategic Applications
530A N. Milwaukee Avenue
Libertyville, IL 60048
(847) 680-9385 FAX: (847) 680-9837

Suppliers

A.4.30

Strem Chemicals

7 Mulliken Way
Newburyport, MA 01950
(800) 647-8736 FAX: (800) 517-8736
(978) 462-3191 FAX: (978) 465-3104
<http://www.strem.com>

StressGen Biotechnologies

Biochemicals Division
120-4243 Glanford Avenue
Victoria, British Columbia
V8Z 4B9 Canada
(800) 661-4978 FAX: (250) 744-2877
(250) 744-2811
<http://www.stressgen.com>

Structure Probe/SPI Supplies

(Epon-Araldite)
P.O. Box 656
West Chester, PA 19381
(800) 242-4774 FAX: (610) 436-5755
<http://www.2spi.com>

Süd-Chemie Performance Packaging

101 Christine Drive
Belen, NM 87002
(800) 989-3374 FAX: (505) 864-9296
<http://www.uniteddesiccants.com>

Sumitomo Chemical

Sumitomo Building
5-33, Kitahama 4-chome
Chuo-ku, Osaka 541-8550, Japan
(81) 6-6220-3891
FAX: (81)-6-6220-3345
<http://www.sumitomo-chem.co.jp>

Sun Box

19217 Orbit Drive
Gaithersburg, MD 20879
(800) 548-3968 FAX: (301) 977-2281
(301) 869-5980
<http://www.sunboxco.com>

Sunbrokers

See Sun International

Sun International

3700 Highway 421 North
Wilmington, NC 28401
(800) LAB-VIAL FAX: (800) 231-7861
<http://www.autosamplervial.com>

Sunox

1111 Franklin Boulevard, Unit 6
Cambridge, Ontario
N1R 8B5 Canada
(519) 624-4413 FAX: (519) 624-8378
<http://www.sunox.ca>

Supelco

See Sigma-Aldrich

SuperArray

P.O. Box 34494
Bethesda, MD 20827
(888) 503-3187 FAX: (301) 765-9859
(301) 765-9888
<http://www.superarray.com>

Surface Measurement Systems

3 Warple Mews, Warple Way
London W3 0RF, UK
(44) 20-8749-4900
FAX: (44) 20-8749-6749
<http://www.smsuk.co.uk/index.htm>

SurgiVet

N7 W22025 Johnson Road, Suite A
Waukesha, WI 53186
(262) 513-8500 (888) 745-6562
FAX: (262) 513-9069
<http://www.surgivet.com>

Sutter Instruments

51 Digital Drive
Novato, CA 94949
(415) 883-0128 FAX: (415) 883-0572
<http://www.sutter.com>

Swiss Precision Instruments

1555 Mittel Boulevard, Suite F
Wooddale, IL 60191
(800) 221-0198 FAX: (800) 842-5164

Synaptosoft

3098 Anderson Place
Decatur, GA 30033
(770) 939-4366 FAX: 770-939-9478
<http://www.synaptosoft.com>

SynChrom

See Micra Scientific

Synergy Software

2457 Perkiomen Avenue
Reading, PA 19606
(800) 876-8376 FAX: (610) 370-0548
(610) 779-0522
<http://www.synergy.com>

Synteni

See Incyte

Synthetics Industry

Lumite Division
2100A Atlantic Highway
Gainesville, GA 30501
(404) 532-9756 FAX: (404) 531-1347

Systat

See SPSS

Systems Planning and Analysis (SPA)

2000 N. Beaugard Street
Suite 400
Alexandria, VA 22311
(703) 931-3500
<http://www.spa-inc.net>

3M Bioapplications

3M Center
Building 270-15-01
St. Paul, MN 55144
(800) 257-7459 FAX: (651) 737-5645
(651) 736-4946

T Cell Diagnostics and**T Cell Sciences**

38 Sidney Street
Cambridge, MA 02139
(617) 621-1400

TAAB Laboratory Equipment

3 Minerva House
Calleva Park
Aldermaston, Berkshire RG7 8NA, UK
(44) 118 9817775
FAX: (44) 118 9817881

Taconic

273 Hover Avenue
Germantown, NY 12526
(800) TAC-ONIC FAX: (518) 537-7287
(518) 537-6208
<http://www.taconic.com>

Tago

See Biosource International

TaKaRa Biochemical

719 Alliston Way
Berkeley, CA 94710
(800) 544-9899 FAX: (510) 649-8933
(510) 649-9895
<http://www.takara.co.jp/english>

Takara Shuzo

Biomedical Group Division
Seta 3-4-1
Otsu Shiga 520-21, Japan
(81) 75-241-5100
FAX: (81) 77-543-9254
<http://www.Takara.co.jp/english>

Takeda Chemical Products

101 Takeda Drive
Wilmington, NC 28401
(800) 825-3328 FAX: (800) 825-0333
(910) 762-8666 FAX: (910) 762-6846
<http://takeda-usa.com>

TAO Biomedical

73 Manassas Court
Laurel Springs, NJ 08021
(609) 782-8622 FAX: (609) 782-8622

Tecan US

P.O. Box 13953
Research Triangle Park, NC 27709
(800) 33-TECAN FAX: (919) 361-5201
(919) 361-5208
<http://www.tecan-us.com>

Techne

University Park Plaza
743 Alexander Road
Princeton, NJ 08540
(800) 225-9243 FAX: (609) 987-8177
(609) 452-9275
<http://www.techneusa.com>

Technical Manufacturing

15 Centennial Drive
Peabody, MA 01960
(978) 532-6330 FAX: (978) 531-8682
<http://www.techmfg.com>

Technical Products International

5918 Evergreen
St. Louis, MO 63134
(800) 729-4451 FAX: (314) 522-6360
(314) 522-8671
<http://www.vibratome.com>

Technicon

See Organon Teknika Cappel

Techno-Aide

P.O. Box 90763
Nashville, TN 37209
(800) 251-2629 FAX: (800) 554-6275
(615) 350-7030
<http://www.techno-aid.com>

Ted Pella

4595 Mountain Lakes Boulevard
P.O. Box 492477
Redding, CA 96049
(800) 237-3526 FAX: (530) 243-3761
(530) 243-2200
<http://www.tedpella.com>

Tekmar-Dohrmann

P.O. Box 429576
Cincinnati, OH 45242
(800) 543-4461 FAX: (800) 841-5262
(513) 247-7000 FAX: (513) 247-7050

Tektronix

142000 S.W. Karl Braun Drive
Beaverton, OR 97077
(800) 621-1966 FAX: (503) 627-7995
(503) 627-7999
<http://www.tek.com>

Tel-Test

P.O. Box 1421
Friendswood, TX 77546
(800) 631-0600 FAX: (281)482-1070
(281)482-2672
<http://www.isotex-diag.com>

TeleChem International

524 East Weddell Drive, Suite 3
Sunnyvale, CA 94089
(408) 744-1331 FAX: (408) 744-1711
<http://www.gst.net/~telechem>

Terrachem

Mallaustrasse 57
D-68219 Mannheim, Germany
0621-876797-0 FAX: 0621-876797-19
<http://www.terrachem.de>

Terumo Medical

2101 Cottontail Lane
Somerset, NJ 08873
(800) 283-7866 FAX: (732) 302-3083
(732) 302-4900
<http://www.terumomedical.com>

Tetko

333 South Highland Manor
Briarcliff, NY 10510
(800) 289-8385 FAX: (914) 941-1017
(914) 941-7767
<http://www.tetko.com>

TetraLink

4240 Ridge Lea Road
Suite 29
Amherst, NY 14226
(800) 747-5170 FAX: (800) 747-5171
<http://www.tetra-link.com>

Suppliers**A.4.31**

TEVA Pharmaceuticals USA
1090 Horsham Road
P.O. Box 1090
North Wales, PA 19454
(215) 591-3000 FAX: (215) 721-9669
<http://www.tevapharmusa.com>

Texas Fluorescence Labs
9503 Capitol View Drive
Austin, TX 78747
(512) 280-5223 FAX: (512) 280-4997
<http://www.teflabs.com>

The Nest Group
45 Valley Road
Southborough, MA 01772
(800) 347-6378 FAX: (508) 485-5736
(508) 481-6223
<http://world.std.com/~nestgrp>

ThermoCare
P.O. Box 6069
Incline Village, NV 89450
(800) 262-4020
(775) 831-1201

Thermo Labsystems
8 East Forge Parkway
Franklin, MA 02038
(800) 522-7763 FAX: (508) 520-2229
(508) 520-0009
<http://www.finnpipette.com>

Thermometric
Spjutvagen 5A
S-175 61 Jarfalla, Sweden
(46) 8-564-72-200

Thermoquest
IEC Division
300 Second Avenue
Needham Heights, MA 02194
(800) 843-1113 FAX: (781) 444-6743
(781) 449-0800
<http://www.thermoquest.com>

Thermo Separation Products
Thermoquest
355 River Oaks Parkway
San Jose, CA 95134
(800) 538-7067 FAX: (408) 526-9810
(408) 526-1100
<http://www.thermoquest.com>

Thermo Shandon
171 Industry Drive
Pittsburgh, PA 15275
(800) 547-7429 FAX: (412) 899-4045
<http://www.thermoshandon.com>

Thermo Spectronic
820 Linden Avenue
Rochester, NY 14625
(585) 248-4000 FAX: (585) 248-4200
<http://www.thermo.com>

Thomas Scientific
99 High Hill Road at I-295
Swedesboro, NJ 08085
(800) 345-2100 FAX: (800) 345-5232
(856) 467-2000 FAX: (856) 467-3087
<http://www.wheatonsci.com/html/nt/>
Thomas.html

Thomson Instrument
354 Tyler Road
Clearbrook, VA 22624
(800) 842-4752 FAX: (540) 667-6878
(800) 541-4792 FAX: (760) 757-9367
<http://www.hplc.com>

Thorn EMI
See Electron Tubes

Thorlabs
435 Route 206
Newton, NJ 07860
(973) 579-7227 FAX: (973) 383-8406
<http://www.thorlabs.com>

Tiemann
See Bernsco Surgical Supply

TILL Photonics GmbH
Lochhamer Schlag 19
D-82166 Gräfelfing
Germany
(49) 89-895-662-0
FAX: (49) 89-895-662-101
<http://www.till-photonics.com/>

Timberline Instruments
1880 South Flatiron Court, H-2
P.O. Box 20356
Boulder, CO 80308
(800) 777-5996 FAX: (303) 440-8786
(303) 440-8779
<http://www.timberlineinstruments.com>

Tissueinformatics
711 Bingham Street, Suite 202
Pittsburgh, PA 15203
(418) 488-1100 FAX: (418) 488-6172
<http://www.tissueinformatics.com>

Tissue-Tek
A Division of Sakura Finetek USA
1750 West 214th Street
Torrance, CA 90501
(800) 725-8723 FAX: (310) 972-7888
(310) 972-7800
<http://www.sakuraus.com>

Tocris Cookson
114 Holloway Road, Suite 200
Ballwin, MO 63011
(800) 421-3701 FAX: (800) 483-1993
(636) 207-7651 FAX: (636) 207-7683
<http://www.tocris.com>

Tocris Cookson
Northpoint, Fourth Way
Avonmouth, Bristol BS11 8TA, UK
(44) 117-982-6551
FAX: (44) 117-982-6552
<http://www.tocris.com>

Tomtec
See CraMar Technologies

TopoGen
P.O. Box 20607
Columbus, OH 43220
(800) TOPOGEN
FAX: (800) ADD-TOPO
(614) 451-5810 FAX: (614) 451-5811
<http://www.topogen.com>

Toray Industries, Japan
Toray Building 2-1
Nihonbash-Muromach
2-Chome, Chuo-Ku
Tokyo, Japan 103-8666
(03) 3245-5115 FAX: (03) 3245-5555
<http://www.toray.co.jp>

Toray Industries, U.S.A.
600 Third Avenue
New York, NY 10016
(212) 697-8150 FAX: (212) 972-4279
<http://www.toray.com>

Toronto Research Chemicals
2 Brisbane Road
North York, Ontario
M3J 2J8 Canada
(416) 665-9696 FAX: (416) 665-4439
<http://www.trc-canada.com>

TosoHaas
156 Keystone Drive
Montgomeryville, PA 18036
(800) 366-4875 FAX: (215) 283-5035
(215) 283-5000
<http://www.tosoHaas.com>

Towhill
647 Summer Street
Boulder, CO 80210
(617) 542-6636 FAX: (617) 464-0804

Toxin Technology
7165 Curtiss Avenue
Sarasota, FL 34231
(941) 925-2032 FAX: (941) 925-2130
<http://www.toxintechnology.com>

Toyo Soda
See TosoHaas

Trace Analytical
3517-A Edison Way
Menlo Park, CA 94025
(650) 364-6895 FAX: (650) 364-6897
<http://www.traceanalytical.com>

Transduction Laboratories
See BD Transduction Laboratories

Transgenomic
2032 Concourse Drive
San Jose, CA 95131
(408) 432-3230 FAX: (408) 432-3231
<http://www.transgenomic.com>

Transonic Systems
34 Dutch Mill Road
Ithaca, NY 14850
(800) 353-3569 FAX: (607) 257-7256
<http://www.transonic.com>

Travenol Lab
See Baxter Healthcare

Tree Star Software
20 Winding Way
San Carlos, CA 94070
800-366-6045
<http://www.treestar.com>

Trevigen
8405 Helgerman Court
Gaithersburg, MD 20877
(800) TREVIGEN FAX: (301) 216-2801
(301) 216-2800
<http://www.trevigen.com>

Trilink Biotechnologies
6310 Nancy Ridge Drive
San Diego, CA 92121
(800) 863-6801 FAX: (858) 546-0020
<http://www.trilink.biotech.com>

Tripos Associates
1699 South Hanley Road, Suite 303
St. Louis, MO 63144
(800) 323-2960 FAX: (314) 647-9241
(314) 647-1099
<http://www.tripos.com>

Triton Environmental Consultants
120-13511 Commerce Parkway
Richmond, British Columbia
V6V 2L1 Canada
(604) 279-2093 FAX: (604) 279-2047
<http://www.triton-env.com>

Tropix
47 Wiggins Avenue
Bedford, MA 01730
(800) 542-2369 FAX: (617) 275-8581
(617) 271-0045
<http://www.tropix.com>

TSI Center for Diagnostic Products
See Intergen

2000 Eppendorf-5 Prime
5603 Arapahoe Avenue
Boulder, CO 80303
(800) 533-5703 FAX: (303) 440-0835
(303) 440-3705

Tyler Research
10328 73rd Avenue
Edmonton, Alberta
T6E 6N5 Canada
(403) 448-1249 FAX: (403) 433-0479

UBI
See Upstate Biotechnology

Ugo Basile Biological Research Apparatus
Via G. Borghi 43
21025 Comerio, Varese, Italy
(39) 332 744 574
FAX: (39) 332 745 488
<http://www.ugobasile.com>

UltraPIX
See Life Science Resources

Ultrasonic Power
239 East Stephenson Street
Freeport, IL 61032
(815) 235-6020 FAX: (815) 232-2150
<http://www.upcorp.com>

Suppliers

A.4.32

Ultrasound Advice

23 Aberdeen Road
London N52UG, UK
(44) 020-7359-1718
FAX: (44) 020-7359-3650
<http://www.ultrasoundadvice.co.uk>

UNELKO

14641 N. 74th Street
Scottsdale, AZ 85260
(480) 991-7272 FAX: (480)483-7674
<http://www.unelko.com>

Unifab Corp.

5260 Lovers Lane
Kalamazoo, MI 49002
(800) 648-9569 FAX: (616) 382-2825
(616) 382-2803

Union Carbide

10235 West Little York Road, Suite 300
Houston, TX 77040
(800) 568-4000 FAX: (713) 849-7021
(713) 849-7000
<http://www.unioncarbide.com>

United Desiccants

See Süd-Chemie Performance Packag-
ing

United States Biochemical

See USB

**United States Biological
(US Biological)**

P.O. Box 261
Swampscott, MA 01907
(800) 520-3011 FAX: (781) 639-1768
<http://www.usbio.net>

Universal Imaging

502 Brandywine Parkway
West Chester, PA 19380
(610) 344-9410 FAX: (610) 344-6515
<http://www.image1.com>

Upchurch Scientific

619 West Oak Street
P.O. Box 1529
Oak Harbor, WA 98277
(800) 426-0191 FAX: (800) 359-3460
(360) 679-2528 FAX: (360) 679-3830
<http://www.upchurch.com>

Upjohn

Pharmacia & Upjohn
<http://www.pnu.com>

Upstate Biotechnology (UBI)

1100 Winter Street, Suite 2300
Waltham, MA 02451
(800) 233-3991 FAX: (781) 890-7738
(781) 890-8845
<http://www.upstatebiotech.com>

USA/Scientific

346 SW 57th Avenue
P.O. Box 3565
Ocala, FL 34478
(800) LAB-TIPS FAX: (352) 351-2057
(3524) 237-6288
<http://www.usascientific.com>

USB

26111 Miles Road
P.O. Box 22400
Cleveland, OH 44122
(800) 321-9322 FAX: (800) 535-0898
FAX: (216) 464-5075
<http://www.usbweb.com>

USCI Bard

Bard Interventional Products
129 Concord Road
Billerica, MA 01821
(800) 225-1332 FAX: (978) 262-4805
<http://www.bardinterventional.com>

Uvex Safety

10 Thurber Blvd.
Smithfield, RI 02917
(800) 434-3411 Fax: 401-232-1830
<http://www.uvex.com>

UVP (Ultraviolet Products)

2066 W. 11th Street
Upland, CA 91786
(800) 452-6788 FAX: (909) 946-3597
(909) 946-3197
<http://www.uvp.com>

V & P Scientific

9823 Pacific Heights Boulevard, Suite T
San Diego, CA 92121
(800) 455-0644 FAX: (858) 455-0703
(858) 455-0643
<http://www.vp-scientific.com>

V-A Optical Labs

60 Red Hill Ave.
San Anselmo, CA 94960
(415) 459-1919 FAX: (415) 459-7216
<http://www.vaoptical.com/>

Valco Instruments

P.O. Box 55603
Houston, TX 77255
(800) FOR-VICI FAX: (713) 688-8106
(713) 688-9345
<http://www.vici.com>

Valpey Fisher

75 South Street
Hopkin, MA 01748
(508) 435-6831 FAX: (508) 435-5289
<http://www.valpeyfisher.com>

Value Plastics

3325 Timberline Road
Fort Collins, CO 80525
(800) 404-LUER FAX: (970) 223-0953
(970) 223-8306
<http://www.valueplastics.com>

Vanguard International

P.O. Box 308
3535 Rt. 66, Bldg. #4
Neptune, NJ 07754
(800) 922-0784 FAX: (732) 922-0557
(732) 922-4900
<http://www.vanguard1.com>

Varian Analytical Instruments

2700 Mitchell Drive
Walnut Creek, CA 94598
(800) 926-3000 FAX: (925) 945-2102
(925) 939-2400
<http://www.varianinc.com>

Varian Associates

3050 Hansen Way
Palo Alto, CA 94304
(800) 544-4636 FAX: (650) 424-5358
(650) 493-4000
<http://www.varian.com>

Vector Core Laboratory/

National Gene Vector Labs
University of Michigan
3560 E MSRB II
1150 West Medical Center Drive
Ann Arbor, MI 48109
(734) 936-5843 FAX: (734) 764-3596

Vector Laboratories

30 Ingold Road
Burlingame, CA 94010
(800) 227-6666 FAX: (650) 697-0339
(650) 697-3600
<http://www.vectorlabs.com>

Vedco

2121 S.E. Bush Road
St. Joseph, MO 64504
(888) 708-3326 FAX: (816) 238-1837
(816) 238-8840
<http://database.vedco.com>

Ventana Medical Systems

3865 North Business Center Drive
Tucson, AZ 85705
(800) 227-2155 FAX: (520) 887-2558
(520) 887-2155
<http://www.ventanamed.com>

Verity Software House

P.O. Box 247
45A Augusta Road
Topsham, ME 04086
(207) 729-6767 FAX: (207) 729-5443
<http://www.vsh.com>

Vernitron

See Sensor Systems LLC

Vertex Pharmaceuticals

130 Waverly Street
Cambridge, MA 02139
(617) 577-6000 FAX: (617) 577-6680
<http://www.vpharm.com>

Vetamac

Route 7, Box 208
Frankfort, IN 46041
(317) 379-3621

Vet Drug

Unit 8
Lakeside Industrial Estate
Colnbrook, Slough SL3 0ED, UK

VetEquip, Inc.

P.O. Box 10785
Pleasanton, CA 94588
(800)-466-6463 Fax: 925-463-1943

Vetus Animal Health

See Burns Veterinary Supply

Viamed

15 Station Road
Cross Hills, Keighley
W. Yorkshire BD20 7DT, UK
(44) 1-535-634-542
FAX: (44) 1-535-635-582
<http://www.viamed.co.uk>

Vical

9373 Town Center Drive, Suite 100
San Diego, CA 92121
(858) 646-1100 FAX: (858) 646-1150
<http://www.vical.com>

Victor Medical

2349 North Watney Way, Suite D
Fairfield, CA 94533
(800) 888-8908 FAX: (707) 425-6459
(707) 425-0294

Virion Systems

9610 Medical Center Drive, Suite 100
Rockville, MD 20850
(301) 309-1844 FAX: (301) 309-0471
<http://www.radix.net/~virion>

VirTis Company

815 Route 208
Gardiner, NY 12525
(800) 765-6198 FAX: (914) 255-5338
(914) 255-5000
<http://www.virtis.com>

Visible Genetics

700 Bay Street, Suite 1000
Toronto, Ontario
M5G 1Z6 Canada
(888) 463-6844 (416) 813-3272
<http://www.visgen.com>

Vitrocom

8 Morris Avenue
Mountain Lakes, NJ 07046
(973) 402-1443 FAX: (973) 402-1445

VTI

7650 W. 26th Avenue
Hialeah, FL 33106
(305) 828-4700 FAX: (305) 828-0299
<http://www.vticorp.com>

VWR Scientific Products

200 Center Square Road
Bridgeport, NJ 08014
(800) 932-5000 FAX: (609) 467-5499
(609) 467-2600
<http://www.vwrsp.com>

Vydac

17434 Mojave Street
P.O. Box 867
Hesperia, CA 92345
(800) 247-0924 FAX: (760) 244-1984
(760) 244-6107
<http://www.vydac.com>

Suppliers**A.4.33**

Vysis

3100 Woodcreek Drive
Downers Grove, IL 60515
(800) 553-7042 FAX: (630) 271-7138
(630) 271-7000
<http://www.vysis.com>

W&H Dentalwerk Bürmoos

P.O. Box 1
A-5111 Bürmoos, Austria
(43) 6274-6236-0
FAX: (43) 6274-6236-55
<http://www.wnhdent.com>

Wako BioProducts

See Wako Chemicals USA

Wako Chemicals USA

1600 Bellwood Road
Richmond, VA 23237
(800) 992-9256 FAX: (804) 271-7791
(804) 271-7677
<http://www.wakousa.com>

Wako Pure Chemicals

1-2, Doshomachi 3-chome
Chuo-ku, Osaka 540-8605, Japan
81-6-6203-3741 FAX: 81-6-6222-1203
<http://www.wako-chem.co.jp/egaiyo/index.htm>

Wallac

See Perkin-Elmer

Wallac

A Division of Perkin-Elmer
3985 Eastern Road
Norton, OH 44203
(800) 321-9632 FAX: (330) 825-8520
(330) 825-4525
<http://www.wallac.com>

Waring Products

283 Main Street
New Hartford, CT 06057
(800) 348-7195 FAX: (860) 738-9203
(860) 379-0731
<http://www.waringproducts.com>

Warner Instrument

1141 Dixwell Avenue
Hamden, CT 06514
(800) 599-4203 FAX: (203) 776-1278
(203) 776-0664
<http://www.warnerinstrument.com>

Warner-Lambert

Parke-Davis
201 Tabor Road
Morris Plains, NJ 07950
(973) 540-2000 FAX: (973) 540-3761
<http://www.warner-lambert.com>

Washington University Machine Shop

615 South Taylor
St. Louis, MO 63310
(314) 362-6186 FAX: (314) 362-6184

Waters Chromatography

34 Maple Street
Milford, MA 01757
(800) 252-HPLC FAX: (508) 478-1990
(508) 478-2000
<http://www.waters.com>

Watlow

12001 Lackland Road
St. Louis, MO 63146
(314) 426-7431 FAX: (314) 447-8770
<http://www.watlow.com>

Watson-Marlow

220 Ballardvale Street
Wilmington, MA 01887
(978) 658-6168 FAX: (978) 988 0828
<http://www.watson-marlow.co.uk>

Waukesha Fluid Handling

611 Sugar Creek Road
Delavan, WI 53115
(800) 252-5200 FAX: (800) 252-5012
(414) 728-1900 FAX: (414) 728-4608
<http://www.waukesha-cb.com>

WaveMetrics

P.O. Box 2088
Lake Oswego, OR 97035
(503) 620-3001 FAX: (503) 620-6754
<http://www.wavemetrics.com>

Weather Measure

P.O. Box 41257
Sacramento, CA 95641
(916) 481-7565

Weber Scientific

2732 Kuser Road
Hamilton, NJ 08691
(800) FAT-TEST FAX: (609) 584-8388
(609) 584-7677
<http://www.weberscientific.com>

Weck, Edward & Company

1 Weck Drive
Research Triangle Park, NC 27709
(919) 544-8000

Wellcome Diagnostics

See Burroughs Wellcome

Wellington Laboratories

398 Laird Road
Guelph, Ontario
N1G 3X7 Canada
(800) 578-6985 FAX: (519) 822-2849
<http://www.well-labs.com>

Wesbart Engineering

Daux Road
Billingshurst, West Sussex
RH14 9EZ, UK
(44) 1-403-782738
FAX: (44) 1-403-784180
<http://www.wesbart.co.uk>

Whatman

9 Bridewell Place
Clifton, NJ 07014
(800) 631-7290 FAX: (973) 773-3991
(973) 773-5800
<http://www.whatman.com>

Wheaton Science Products

1501 North 10th Street
Millville, NJ 08332
(800) 225-1437 FAX: (800) 368-3108
(856) 825-1100 FAX: (856) 825-1368
<http://www.algroupwheaton.com>

Whittaker Bioproducts

See BioWhittaker

Wild Heerbrugg

Juerg Dedual Gaebrisstrasse 8 CH
9056 Gais, Switzerland
(41) 71-793-2723
FAX: (41) 71-726-5957
http://www.homepage.swissonline.net/dedual/wild_heerbrugg

Willy A. Bachofen

AG Maschinenfabrik
Utengasse 15/17
CH4005 Basel, Switzerland
(41) 61-681-5151
FAX: (41) 61-681-5058
<http://www.wab.ch>

Winthrop

See Sterling Winthrop

Wolfram Research

100 Trade Center Drive
Champaign, IL 61820
(800) 965-3726 FAX: (217) 398-0747
(217) 398-0700
<http://www.wolfram.com>

World Health Organization

Microbiology and Immunology Support
20 Avenue Appia
1211 Geneva 27, Switzerland
(41-22) 791-2602
FAX: (41-22) 791-0746
<http://www.who.org>

World Precision Instruments

175 Sarasota Center Boulevard
International Trade Center
Sarasota, FL 34240
(941) 371-1003 FAX: (941) 377-5428
<http://www.wpiinc.com>

Worthington Biochemical

Halls Mill Road
Freehold, NJ 07728
(800) 445-9603 FAX: (800) 368-3108
(732) 462-3838 FAX: (732) 308-4453
<http://www.worthington-biochem.com>

WPI

See World Precision Instruments

Wyeth-Ayerst

2 Esterbrook Lane
Cherry Hill, NJ 08003
(800) 568-9938 FAX: (858) 424-8747
(858) 424-3700

Wyeth-Ayerst Laboratories

P.O. Box 1773
Paoli, PA 19301
(800) 666-7248 FAX: (610) 889-9669
(610) 644-8000
<http://www.ahp.com>

Xenotech

3800 Cambridge Street
Kansas City, KS 66103
(913) 588-7930 FAX: (913) 588-7572
<http://www.xenotechllc.com>

Xeragon

19300 Germantown Road
Germantown, MD 20874
(240) 686-7860 FAX: (240) 686-7861
<http://www.xeragon.com>

Xillix Technologies

300-13775 Commerce Parkway
Richmond, British Columbia
V6V 2V4 Canada
(800) 665-2236 FAX: (604) 278-3356
(604) 278-5000
<http://www.xillix.com>

Xomed Surgical Products

6743 Southpoint Drive N
Jacksonville, FL 32216
(800) 874-5797 FAX: (800) 678-3995
(904) 296-9600 FAX: (904) 296-9666
<http://www.xomed.com>

Yakult Honsha

1-19, Higashi-Shinbashi 1-chome
Minato-ku Tokyo 105-8660, Japan
81-3-3574-8960

Yamasa Shoyu

23-8 Nihonbashi Kakigaracho
1-chome, Chuoku
Tokyo, 103 Japan
(81) 3-479 22 0095
FAX: (81) 3-479 22 3435

Yeast Genetic Stock Center

See ATCC

Yellow Spring Instruments

See YSI

YMC

YMC Karasuma-Gojo Building
284 Daigo-Cho, Karasuma Nishihiir
Gojo-dori Shimogyo-ku
Kyoto, 600-8106, Japan
(81) 75-342-4567
FAX: (81) 75-342-4568
<http://www.ymc.co.jp>

YSI

1725-1700 Brannum Lane
Yellow Springs, OH 45387
(800) 765-9744 FAX: (937) 767-9355
(937) 767-7241
<http://www.ysi.com>

Zeneca/CRB

See AstraZeneca
(800) 327-0125 FAX: (800) 321-4745

Zivic-Miller Laboratories

178 Toll Gate Road
Zelienople, PA 16063
(800) 422-LABS FAX: (724) 452-4506
(800) MBM-RATS FAX: (724) 452-5200
<http://zivicmiller.com>

Suppliers**A.4.34**

Zymark
Zymark Center
Hopkinton, MA 01748
(508) 435-9500 FAX: (508) 435-3439
<http://www.zymark.com>

Zymed Laboratories
458 Carlton Court
South San Francisco, CA 94080
(800) 874-4494 FAX: (650) 871-4499
(650) 871-4494
<http://www.zymed.com>

Zymo Research
625 W. Katella Avenue, Suite 30
Orange, CA 92867
(888) 882-9682 FAX: (714) 288-9643
(714) 288-9682
<http://www.zymor.com>

Zynaxis Cell Science
See ChiRex Cauldron

Suppliers

A.4.35

VECTORS

The following tables collect vectors described throughout the manual, organized by particular function. Individual entries in each table provide figure citations, some specific characteristics, and restriction site data for each vector, as well as sources when available. Restriction site information was generated using MAPSORT (part of the Wisconsin Package of analysis programs; copyright 1982-1992 Genetics Computer Group, Inc., e-mail *help@gcg.com*) from sequence information compiled in part from published sequence data; therefore, although all possible care has been taken, accuracy of the listings cannot be guaranteed. A listing of the contents of each table is provided below; listings in italics represent cross-references to other tables for vectors that belong to more than one category.

NOTE: Enzymes are listed alphabetically across rows of tables, not down columns.

Table A.5.1	<i>E. coli</i> General-Purpose Cloning Vectors	A.5.3
	M13mp18	A.5.3
	pBR322	A.5.4
	pSP64	A.5.5
	pUC19	A.5.5
Table A.5.2	Yeast General-Purpose Cloning Vectors	A.5.6
	pRS303	A.5.6
	pRS304	A.5.7
	pRS305	A.5.8
	pRS306	A.5.9
	pRS313	A.5.10
	pRS314	A.5.11
	pRS315	A.5.12
	pRS316	A.5.13
	YIp5	A.5.14
Table A.5.3	Vectors for Construction of DNA Libraries in <i>E. coli</i>	A.5.15
	λEMBL3	A.5.15
	λgt10	A.5.15
	λgt11	A.5.15
	λZAP and λZAPII	A.5.16
	pJB8	A.5.16
Table A.5.4	Vectors for Construction of Artificial Chromosomes in Yeast	A.5.17
	pYAC3	A.5.17
	pYAC4	A.5.17
Table A.5.5	Vectors for Expression of Genes from Heterologous Promoters in <i>E. coli</i>	A.5.18
	<i>λZAP and λZAPII (see Table A.5.3)</i>	A.5.18
	pSKF101	A.5.18
	pSKF301	A.5.18
	<i>pSP64 (see Table A.5.1)</i>	A.5.19
	pT7-5	A.5.19
	pT7-6	A.5.20
	pT7-7	A.5.21
	pTrc99A	A.5.22
Table A.5.6	Vectors for Expression of Genes from Heterologous Promoters in Mammalian Cells^a	A.5.23
	BAG retrovirus (stable)	A.5.23
	CDM8 (transient)	A.5.23
	pEB2 (transient)	A.5.24
	pED (stable)	A.5.24

continued

Contributed by Lisa M. Albright

Current Protocols in Molecular Biology (1997) A.5.1-A.5.45

Copyright © 2003 by John Wiley & Sons, Inc.

Vectors

A.5.1

Supplement 64

Table A.5.6 Vectors for Expression of Genes from Heterologous Promoters in Mammalian Cells, continued

pEE14 (stable)	A.5.25
pMJ601 (transient)	A.5.25
pSC11 (transient)	A.5.26
pSFV1 (transient)	A.5.26
pSFV-Helper 2 (helper for pSFV1)	A.5.27
pSFV3-lacZ (control for pSFV1)	A.5.27
pTF7-5 (transient)	A.5.28
pTKgptF1s (transient)	A.5.29
pTM1 (transient)	A.5.29
pTM3 (transient)	A.5.30

Table A.5.7 Vectors Using Reporter Genes for Genetic Regulation Studies in Mammalian Cells

pβgal-Basic	A.5.31
pφGH (transient)	A.5.31
pOLUC (transient)	A.5.32
pSV2LUC (control for pOLUC)	A.5.33
pXGH5 (control for pφGH)	A.5.33
p300-CAT (transient)	A.5.34

Table A.5.8 Vectors for Expression of Fusion Proteins in *E. coli*

<i>λgt11</i> (see Table A.5.3)	A.5.35
<i>λZAP</i> and <i>λZAPII</i> (see Table A.5.3)	A.5.35
pALtrxA-781	A.5.35
pET-15b	A.5.36
pGE374	A.5.36
pGEX1	A.5.37
pGEX2T	A.5.37
pGEX3X	A.5.38
pMAL-c2	A.5.38
pMAL-p2	A.5.39
<i>pSKF301</i> (see Table A.5.5)	A.5.39
pTRXFUS	A.5.40
pTrc99A	A.5.40
pUR278	A.5.40

Table A.5.9 Vectors for Specialized Functions^a

mGP1-2 (produces T7 RNA polymerase; also pT7 series, Table A.5.5)	A.5.41
pAD1 (for screening <i>λ</i> libraries by homologous recombination)	A.5.41
pEG202 (for cloning DNA for bait protein in interactor hunt)	A.5.41
pJG4-5 (for cloning cDNA libraries of bait-protein interactors)	A.5.41
pGP1-2 (produces T7 series in Table A.5.5)	A.5.42
pNTK (for targeted gene knockout; also pSKV1, Table A.5.7)	A.5.42

Table A.5.10 Vectors of Historical Interest

Charon 4a	A.5.43
<i>λ</i> 2001	A.5.43
Wild-type <i>λ</i> bacteriophage (see also <i>λ</i> vectors, Table A.5.3)	A.5.43
Wild-type M13 bacteriophage	A.5.44
2μm circle	A.5.45

Key References

A.5.45

^aVectors for protein expression in insect cells using the baculovirus system are described in *UNIT 16.9* and listed in Fig. 16.9.1.

Table A.5.1 *E. coli* General-Purpose Cloning Vectors

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
M13mp18 (Fig. 1.14.1)	7249	Recombinant phage form colorless plaques (blue/white screen) on appropriate host	Polylinker (see Fig. 1.14.2)	Many academic and commercial labs (X02513)

COMMENTS: Member of family of vectors based on single-stranded DNA phage M13 with polylinker inserted in alpha fragment of *lacZ* gene (see Fig. 1.14.2).

ENZYMES THAT CUT ONCE:

AccI 6264	Alw44I 4742	AlwNI 2191	AvaII 5913	BamHI 6251
BglII 6436	BglIII 6934	BsaHI 6001	BseRI 2009	BsmI 1751
Bsp1407I 1020	BspGI 3641	BspHI 1298	BsrFI 5612	Bsu36I 6508
DraIII 5720	DrdI 5764	EcoRI 6230	Esp3I 5976	FspI 6426
HgiEII 6472	HincII 6265	HindIII 6281	KpnI 6246	MscI 5081
NaeI 5614	NarI 6001	PacI 4135	PstI 6273	PvuI 6407
RleAI 3006	SalI 6263	SmaI 6248	SnaBI 1269	SphI 6279
Sse8387I 6273	SstI 6240	SwaI 6785	XbaI 6257	

ENZYMES THAT CUT TWICE:

AvaI 5824 6246	BanII 5646 6240	BpmI 6513 6883	BsaBI 1152 3977	BsbI 640 5810
BsmFI 5074 6513	Bsp24I 2154 2186	ClaI 2527 6882	DsaI 2762 6616	EarI 4079 6385
EciI 1963 6604	Eco47III 2711 3040	GdiII 6035 6292	MsII 2799 6999	Pfl1108I 1257 3424
Psp1406I 4634 6773	TaqII 1438 2460	XmnI 361 2649		

ENZYMES THAT DO NOT CUT:

AatII, AflII, AgeI, ApaI, ApaBI, AscI, AvrII, BbsI, BcgI, BcgI, BclI, Bpu1102I, BsaI, BsgI, BsiI, BspEI, BssHII, Bst1107I, BstEII, BstXI, EagI, Eam1105I, Eco57I, EcoNI, EcoO109I, EcoRV, FseI, HpaI, MluI, MunI, NcoI, NheI, NotI, NruI, NsiI, NspV, PflMI, PmeI, PmlI, PshAI, Psp5II, RsrII, SacII, SapI, Scal, SexAI, SfiI, SgfI, SgrAI, SpeI, SrfI, StuI, StyI, SunI, Tth111I, XcmI, XhoI

continued

Table A.5.1 *E. coli* General-Purpose Cloning Vectors, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pBR322 (Fig. 1.5.9)	4361	Amp ^r , Tet ^r	See Fig. 1.5.9	Many academic and commercial labs (J01749)

COMMENTS: Prototype high-copy plasmid; parent of many.

ENZYMES THAT CUT ONCE:

AatII 4288	AflIII 2473	AlwNI 2889	ApoI 4359	AvaI 1425
BamHI 375	Bpu10I 1581	BsaI 3427	BsaAI 2227	BsaBI 1672
BsgI 1635	BsmI 1359	BsmBI 2116	BspEI 1664	BspLU11I 2473
BspMI 1054	Bst1107I 2246	ClaI 24	EagI 939	Eam1105I 3366
EcoNI 626	EcoRI 4359	EcoRV 187	HindIII 29	MscI 1446
NdeI 2296	NheI 229	NruI 974	PshAI 716	PstI 3611
PvuI 3736	PvuII 2066	SalI 651	SapI 2357	ScaI 3846
SgrAI 410	SphI 566	SspI 4170	StyI 1369	Tth111I 2220
VspI 3538				

ENZYMES THAT CUT TWICE:

AccI 652 2245	ApaBI 1052 2298	BanII 475 489	BsbI 2191 3909	BsrBI 2406 4207
BsrDI 3427 3601	DrdI 2168 2581	DsaI 528 1447	EarI 2357 4161	Eco57I 3021 4033
HgiEII 2298 3059	HincII 653 3907	Pfl1108I 758 3384	PflMI 1321 1370	Psp5II 1439 1481
TaqII 4052 4069	XmnI 2033 3965			

ENZYMES THAT DO NOT CUT:

AflII, *AgeI*, *ApaI*, *AscI*, *AvrII*, *BaeI*, *BclI*, *BglII*, *BmgI*, *Bpu1102I*, *BsaXI*, *BseRI*, *BsrGI*, *BssHII*, *BstEII*, *BstXI*, *Bsu36I*, *DraIII*, *DrdII*, *FseI*, *HpaI*, *KpnI*, *MluI*, *MunI*, *NcoI*, *NotI*, *NsiI*, *NspV*, *PacI*, *PmeI*, *PmlI*, *RleAI*, *RsrII*, *SacI*, *SacII*, *SexAI*, *SfiI*, *SgfI*, *SmaI*, *SnaBI*, *SpeI*, *SrfI*, *Sse8387I*, *StuI*, *SunI*, *SwaI*, *XbaI*, *XcmI*, *XhoI*

continued

Table A.5.1 *E. coli* General-Purpose Cloning Vectors, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pSP64	2999	Amp ^r	Polylinker	Promega (X65327)

COMMENTS: High copy number; contains SP6 RNA polymerase promoter reading into polylinker.

ENZYMES THAT CUT ONCE:

AatII 2223	AccI 23	AflIII 408	AlwNI 824	ApoI 52
AvaI 39	BamHI 34	BglII 1421	BpmI 1371	BsaI 1362
BsmFI 2589	Bsp1407I 2919	BspLU11I 408	BspMI 9	DrdII 2780
DsaI 2597	Eam1105I 1301	Eco47III 2633	EcoNI 2502	EcoRI 52
FspI 1523	HgiEII 994	HincII 24	HindIII 7	NaeI 2726
NheI 2757	Pfl1108I 1319	PstI 20	PvuI 1671	PvuII 232
SalI 22	SapI 292	ScaI 1781	SgrAI 2715	SmaI 41
SphI 2567	SspI 2105	SstI 50	XbaI 28	XmnI 1900

ENZYMES THAT CUT TWICE:

AvaII 1439 1661	BcgI 1806 2751	BsmBI 2291 2333	BsrDI 1362 1536	DrdI 516 2385
EarI 292 2096	EcoO109I 2277 2602	Psp1406I 1527 1900	TfiI 243 383	

ENZYMES THAT DO NOT CUT:

AflII, AgeI, ApaI, ApaBI, AscI, AvrII, BaeI, BbsI, BclI, BglIII, Bpu10I, Bpu1102I, BsaAI, BsaBI, BseRI, BsgI, BsmI, BspEI, BspGI, BssHII, Bst1107I, BstEII, BstXI, Bsu36I, ClaI, DraIII, EagI, EcoRV, FseI, HpaI, KpnI, MluI, MscI, MunI, NcoI, NdeI, NotI, NruI, NsiI, NspV, PacI, PflMI, PmeI, PmlI, PshAI, Psp5II, RleAI, RsrII, SacII, SexAI, SfiI, SgfI, SnaBI, SpeI, SrfI, Sse8387I, StuI, StyI, SunI, SwaI, Tth111I, XcmI, XhoI

pUC19 (Fig. 1.5.2)	2686	Amp ^r	Polylinker (see Fig. 1.5.2)	Many academic and commercial labs (X02514)
------------------------------	------	------------------	--------------------------------	-----------------------------------------------

COMMENTS: Very high copy number; member of family of plasmids with polylinker inserted in alpha fragment of *lacZ* gene (see Fig. 1.14.2).

ENZYMES THAT CUT ONCE:

AatII 753	AccI 252	AflIII 2560	AlwNI 2151	ApaBI 503
ApoI 284	AvaI 268	BamHI 263	BanII 282	BcgI 1168
BpmI 1603	BsaI 1606	BspLU11I 2560	BspMI 238	BsrFI 1587
Eam1105I 1672	EcoO109I 692	EcoRI 284	HincII 253	HindIII 233
KpnI 276	NarI 446	NdeI 498	Pfl1108I 1653	PstI 249
SalI 251	SapI 2677	ScaI 1191	SmaI 270	SphI 243
Sse8387I 249	SspI 867	SstI 282	XbaI 257	XmnI 1072

ENZYMES THAT CUT TWICE:

AvaII 1837 2059	BceII 403 1308	BcgI 2204 2238	BglII 251 1819	BsmBI 3 45
BsrDI 1760 1934	CjePI 984 2340	DrdI 97 914	Eco57I 1354 2366	FspI 258 1921
HgiEII 186 1392	MmeI 1021 1205	Psp1406I 1925 2298	PvuI 279 2069	PvuII 308 630
TfiI 641 781				

ENZYMES THAT DO NOT CUT:

AflII, AgeI, ApaI, AscI, AvrII, BaeI, BbsI, BclI, BglIII, Bpu10I, Bpu1102I, BsaAI, BsaBI, BseRI, BsgI, BsmI, BsmFI, Bsp1407I, BspEI, BspGI, BssHII, Bst1107I, BstEII, BstXI, Bsu36I, ClaI, DraIII, DrdII, DsaI, EagI, Eco47III, EcoNI, EcoRV, FseI, HpaI, MluI, MscI, MunI, NaeI, NcoI, NheI, NotI, NruI, NsiI, NspV, PacI, PflMI, PmeI, PmlI, PpuMI, PshAI, RleAI, RsrII, SacII, SexAI, SfiI, SgrAI, SnaBI, SpeI, SrfI, StuI, StyI, SunI, SwaI, Tth111I, XcmI, XhoI

Table A.5.2 Yeast General-Purpose Cloning Vectors

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pRS303 (Fig. 1.5.6)	4453	Yeast: HIS3 ⁺ ; <i>E. coli</i> : Amp ^r	Polylinker (see <i>COMMENTS</i> and Fig. 1.5.6)	ATCC #77138

COMMENTS: Yeast integrating plasmid containing KS polylinker from pBLUESCRIPT (unique sites are *ApaI*, *XhoI*, *Sall*, *ClaI*, *EcoRV*, *EcoRI*, *SmaI*, *BamHI*, *SpeI*, *XbaI*, *EagI*, *NotI*, *SacII*, and *SacI*) and f1(+) phage DNA replication origin from pBLUESCRIBE (Sikorski and Hieter, 1989).

ENZYMES THAT CUT ONCE:

AatII 4388	AlwNI 2989	ApaI 2173	BamHI 2109	BbsI 790
BcgI 4005	BclII 1052	BpmI 3536	BsaI 3527	BsaBI 561
BseRI 303	BsmI 716	BsmFI 565	BspGI 689	BssHIII 1312
ClaI 2146	DraIII 1647	DrdII 1652	DsaI 2081	EagI 2090
Eam1105I 3466	Eco47III 213	EcoRI 2127	EcoRV 2135	HgiEII 3159
HincII 2156	MscI 711	NdeI 694	NgoAIV 1748	NheI 1008
NotI 2090	NsiI 1283	Pfl1108I 3484	SacI 2077	SacII 2084
Sall 2154	SapI 2457	ScaI 3946	SmaI 2117	SpeI 2103
SunI 916	XbaI 2097	XhoI 2160	XmnI 4065	

ENZYMES THAT CUT TWICE:

AvaI 2115 2160	BcgI 3971 4005	BglII 899 959	BsaAI 1294 1647	BsaHI 4003 4385
BsaXI 1595 2431	BsbI 120 1554	BsmBI 3 45	BspLU11I 1203 2573	BsrDI 3527 3701
BsrFI 1748 3546	Bst1107I 1219 1245	BstXI 893 2085	FspI 1902 3688	KpnI 1126 2179
NspV 456 906	PstI 1187 2125	PvuI 1923 3836	PvuII 1952 2397	SspI 1439 4270
TaqII 4152 4169				

ENZYMES THAT DO NOT CUT:

AflII, *AgeI*, *ApaBI*, *AscI*, *AvrII*, *BaeI*, *Bpu10I*, *Bpu1102I*, *BsgI*, *BspEI*, *BspMI*, *BsrGI*, *BstEII*, *Bsu36I*, *EcoNI*, *FseI*, *HpaI*, *MluI*, *MunI*, *NarI*, *NcoI*, *NruI*, *PacI*, *PflMI*, *PmeI*, *PmlI*, *PshAI*, *Psp5II*, *RleAI*, *RsrII*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SnaBI*, *SphI*, *SrfI*, *Sse8387I*, *StuI*, *StyI*, *SwaI*, *Tth111I*, *XcmI*

continued

Table A.5.2 Yeast General-Purpose Cloning Vectors, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pRS304	4271	Yeast: TRP1 ⁺ ; <i>E. coli</i> : Amp ^r	Polylinker (see <i>COMMENTS</i>)	ATCC #77139

COMMENTS: Yeast integrating plasmid containing KS polylinker from pBLUESCRIPT (unique sites are *KpnI*, *ApaI*, *XhoI*, *SalI*, *ClaI*, *EcoRI*, *PstI*, *SmaI*, *BamHI*, *SpeI*, *EagI*, *NotI*, *SacII*, and *SacI*) and f1(+) phage DNA replication origin from pBLUESCRIBE (Sikorski and Hieter, 1989).

ENZYMES THAT CUT ONCE:

<i>AatII</i> 4206	<i>AflIII</i> 2391	<i>AlwNI</i> 2807	<i>ApaI</i> 1991	<i>ApaBI</i> 547
<i>BamHI</i> 1927	<i>BpmI</i> 3354	<i>BsaI</i> 3345	<i>BsgI</i> 880	<i>BsmFI</i> 929
<i>BspEI</i> 332	<i>BspLU11I</i> 2391	<i>BspMI</i> 900	<i>Bst1107I</i> 434	<i>Bsu36I</i> 709
<i>ClaI</i> 1964	<i>DraIII</i> 1465	<i>DrdII</i> 1470	<i>DsaI</i> 1899	<i>EagI</i> 1908
<i>Eam1105I</i> 3284	<i>EcoRI</i> 1945	<i>HgiEII</i> 2977	<i>HincII</i> 1974	<i>KpnI</i> 1997
<i>MunI</i> 607	<i>NgoAIV</i> 1566	<i>NotI</i> 1908	<i>Pfl1108I</i> 3302	<i>PmlI</i> 427
<i>PstI</i> 1943	<i>SacI</i> 1895	<i>SacII</i> 1902	<i>SalI</i> 1972	<i>SapI</i> 2275
<i>ScaI</i> 3764	<i>SmaI</i> 1935	<i>SnaBI</i> 319	<i>SpeI</i> 1921	<i>XhoI</i> 1978

ENZYMES THAT CUT TWICE:

<i>AccI</i> 433 1973	<i>AvaI</i> 1933 1978	<i>BcgI</i> 734 3823	<i>BglI</i> 1713 3404	<i>BsaXI</i> 1413 2249
<i>BsmBI</i> 3 45	<i>BsrDI</i> 3345 3519	<i>BsrFI</i> 1566 3364	<i>BstXI</i> 765 1903	<i>Eco57I</i> 2939 3951
<i>EcoO109I</i> 1987 4260	<i>EcoRV</i> 752 1953	<i>FspI</i> 1720 3506	<i>HindIII</i> 980 1957	<i>NspI</i> 41 2395
<i>PvuI</i> 1741 3654	<i>PvuII</i> 1770 2215	<i>SspI</i> 1257 4088	<i>TaqII</i> 3970 3987	<i>XbaI</i> 551 1915

ENZYMES THAT DO NOT CUT:

AflII, *AgeI*, *AscI*, *AvrII*, *BaeI*, *BbsI*, *BclI*, *BglII*, *Bpu10I*, *Bpu1102I*, *BsaBI*, *BseRI*, *BsmI*, *BspGI*, *BsrGI*, *BssHII*, *BstEII*, *Eco47III*, *EcoNI*, *FseI*, *HpaI*, *MluI*, *MscI*, *NarI*, *NcoI*, *NdeI*, *NheI*, *NruI*, *NsiI*, *NspV*, *PacI*, *PflMI*, *PmeI*, *PshAI*, *Psp5II*, *RleAI*, *RsrII*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SphI*, *SrfI*, *Sse8387I*, *StuI*, *StyI*, *SunI*, *SwaI*, *Tth111I*, *XcmI*

continued

Table A.5.2 Yeast General-Purpose Cloning Vectors, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pRS305	5504	Yeast: LEU2 ⁺ ; <i>E. coli</i> : Amp ^r	Polylinker (see <i>COMMENTS</i>)	ATCC #77140

COMMENTS: Yeast integrating plasmid containing KS polylinker from pBLUESCRIPT (unique sites are *Apa*I, *Xho*I, *Sal*I, *Hind*III, *Pst*I, *Sma*I, *Bam*HI, *Spe*I, *Xba*I, *Eag*I, *Not*I, *Sac*II, and *Sac*I) and f1(+) phage DNA replication origin from pBLUESCRIBE (Sikorski and Hieter, 1989).

ENZYMES THAT CUT ONCE:

<i>Aat</i> II 5439	<i>Afl</i> III 1438	<i>Age</i> I 1503	<i>Alw</i> NI 4040	<i>Apa</i> I 3224
<i>Bae</i> I 2157	<i>Bam</i> HI 3160	<i>Bbs</i> I 1301	<i>Bcg</i> I 5056	<i>Bsa</i> I 4578
<i>Bsa</i> AI 2698	<i>Bsm</i> FI 1602	<i>Bsp</i> MI 759	<i>Bsr</i> GI 632	<i>Bst</i> EII 1718
<i>Dra</i> III 2698	<i>Eag</i> I 3141	<i>Eam</i> 1105I 4517	<i>Hind</i> III 3190	<i>Hpa</i> I 2174
<i>Nar</i> I 2204	<i>Ngo</i> AIV 2799	<i>Not</i> I 3141	<i>Psp</i> 5II 1875	<i>Pst</i> I 3176
<i>Sac</i> I 3128	<i>Sac</i> II 3135	<i>Sal</i> I 3205	<i>Sap</i> I 3508	<i>Sca</i> I 4997
<i>Sma</i> I 3168	<i>Spe</i> I 3154	<i>Tth</i> 111I 192	<i>Xba</i> I 3148	<i>Xcm</i> I 1201
<i>Xho</i> I 3211				

ENZYMES THAT CUT TWICE:

<i>Acc</i> I 561 3206	<i>Afl</i> III 1016 3624	<i>Ava</i> I 3166 3211	<i>Bcg</i> I 5022 5056	<i>Bgl</i> I 2946 4637
<i>Bpm</i> I 902 4587	<i>Bsb</i> I 120 2605	<i>Bse</i> RI 405 2398	<i>Bsm</i> BI 3 45	<i>Bsp</i> LU11I 1016 3624
<i>Bst</i> XI 1046 3136	<i>Cl</i> aI 1603 3197	<i>Dsa</i> I 892 3132	<i>Eco</i> NI 797 1487	<i>Eco</i> RI 1116 3178
<i>Eco</i> RV 1007 3186	<i>Fsp</i> I 2953 4739	<i>Kpn</i> I 1510 3230	<i>Nsp</i> V 1633 2133	<i>Pfl</i> 1108I 369 4535
<i>Pfl</i> MI 288 1859	<i>Pvu</i> I 2974 4887	<i>Pvu</i> II 3003 3448	<i>Rle</i> AI 1542 2183	

ENZYMES THAT DO NOT CUT:

*Apa*BI, *Asc*I, *Avr*II, *Bcl*I, *Bgl*II, *Bpu*10I, *Bpu*1102I, *Bsa*BI, *Bsg*I, *Bsm*I, *Bsp*EI, *Bsp*GI, *Bss*HII, *Bst*1107I, *Bsu*36I, *Eco*47III, *Fse*I, *Mlu*I, *Msc*I, *Nco*I, *Nde*I, *Nhe*I, *Nru*I, *Nsi*I, *Pac*I, *Pme*I, *Pml*I, *Psh*AI, *Rsr*II, *Sex*AI, *Sfi*I, *Sgf*I, *Sgr*AI, *Sna*BI, *Sph*I, *Srf*I, *Sse*8387I, *Stu*I, *Sun*I, *Sw*aI

continued

Table A.5.2 Yeast General-Purpose Cloning Vectors, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pRS306	4381	Yeast: URA3 ⁺ ; <i>E. coli</i> : Amp ^r	Polylinker (see <i>COMMENTS</i>)	ATCC #77141

COMMENTS: Yeast integrating plasmid containing KS polylinker from pBLUESCRIPT (unique sites are *KpnI*, *XhoI*, *Sall*, *ClaI*, *HindIII*, *EcoRI*, *SmaI*, *BamHI*, *SpeI*, *XbaI*, *EagI*, *NotI*, *BstXI*, *SacII*, and *SacI*) and f1(+) phage DNA replication origin from pBLUESCRIBE (Sikorski and Hieter, 1989).

ENZYMES THAT CUT ONCE:

<i>AatII</i> 4316	<i>BamHI</i> 2037	<i>BbsI</i> 691	<i>BcgI</i> 3933	<i>Bpu10I</i> 1147
<i>BsaAI</i> 1575	<i>BsgI</i> 363	<i>BsmI</i> 779	<i>BsmFI</i> 556	<i>BspMI</i> 402
<i>BstXI</i> 2013	<i>ClaI</i> 2074	<i>DrdII</i> 1580	<i>EagI</i> 2018	<i>Eam1105I</i> 3394
<i>EcoRI</i> 2055	<i>HindIII</i> 2067	<i>KpnI</i> 2107	<i>NcoI</i> 622	<i>NdeI</i> 330
<i>NgoAIV</i> 1676	<i>NotI</i> 2018	<i>NsiI</i> 1244	<i>NspV</i> 682	<i>Pfl1108I</i> 3412
<i>Psp5II</i> 570	<i>RleAI</i> 1015	<i>SacI</i> 2005	<i>SacII</i> 2012	<i>Sall</i> 2082
<i>SmaI</i> 2045	<i>SpeI</i> 2031	<i>Sse8387I</i> 399	<i>StuI</i> 853	<i>XbaI</i> 2025
<i>XcmI</i> 545	<i>XhoI</i> 2088	<i>XmnI</i> 3993		

ENZYMES THAT CUT TWICE:

<i>AflIII</i> 596 2501	<i>AlwNI</i> 1098 2917	<i>ApaI</i> 795 2101	<i>AvaI</i> 2043 2088	<i>BcgI</i> 3899 3933
<i>BglII</i> 1823 3514	<i>BsaI</i> 1091 3455	<i>BsaHI</i> 3931 4313	<i>BsaXI</i> 1523 2359	<i>BspLU11I</i> 596 2501
<i>BsrFI</i> 1676 3474	<i>Bst1107I</i> 743 1248	<i>Eco57I</i> 3049 4061	<i>EcoRV</i> 605 2063	<i>FspI</i> 1830 3616
<i>HgiEII</i> 1300 3087	<i>PstI</i> 399 2053	<i>PvuI</i> 1851 3764	<i>PvuII</i> 1880 2325	<i>SapI</i> 927 2385
<i>ScaI</i> 729 3874	<i>SspI</i> 1367 4198	<i>StyI</i> 538 622		

ENZYMES THAT DO NOT CUT:

AflII, *AgeI*, *ApaBI*, *AscI*, *AvrII*, *BaeI*, *BclI*, *BglIII*, *Bpu1102I*, *BsaBI*, *BseRI*, *BspEI*, *BspGI*, *BsrGI*, *BssHII*, *BstEII*, *Bsu36I*, *Eco47III*, *EcoNI*, *FseI*, *HpaI*, *MluI*, *MscI*, *MunI*, *NarI*, *NheI*, *NruI*, *PacI*, *PflMI*, *PmeI*, *PmlI*, *PshAI*, *RsrII*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SnaBI*, *SphI*, *SrfI*, *SunI*, *SwaI*, *Tth111I*

continued

Table A.5.2 Yeast General-Purpose Cloning Vectors, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pRS313	4967	Yeast: HIS3 ⁺ ; <i>E. coli</i> : Amp ^r	Polylinker (see COMMENTS)	ATCC #77142

COMMENTS: Like pRS303 but are containing CEN/ARS cassette, making it a yeast centromere plasmid; contains KS polylinker from pBLUESCRIPT (unique sites are *Apa*I, *Xho*I, *Sal*I, *Cl*aI, *Eco*RV, *Eco*RI, *Sma*I, *Bam*HI, *Spe*I, *Xba*I, *Eag*I, *Not*I, *Sac*II, and *Sac*I) and fl(+) phage DNA replication origin from pBLUESCRIBE (Sikorski and Hieter, 1989).

ENZYMES THAT CUT ONCE:

Alw NI 2989	Apa I 2173	Bam HI 2109	Bcg I 4005	Bcl II 1052
Bpm I 3536	Bsa I 3527	Bsa BI 561	Bsa HI 4003	Bse RI 303
Bsm I 716	Bsm FI 565	Bsp GI 689	Bss HII 1312	Cl aI 2146
Dra III 1647	Drd II 1652	Dsa I 2081	Eag I 2090	Eam 1105I 3466
Eco 47III 213	Eco RI 2127	Eco RV 2135	Hgi EII 3159	Hinc II 2156
Msc I 711	Nde I 694	Ngo AIV 1748	Nhe I 1008	Not I 2090
Nsi I 1283	Pfl 1108I 3484	Pml I 4399	Psp 5II 4385	Sac I 2077
Sac II 2084	Sal I 2154	Sca I 3946	Sma I 2117	Spe I 2103
Sun I 916	Swa I 4423	Xba I 2097	Xho I 2160	

ENZYMES THAT CUT TWICE:

Ava I 2115 2160	Bbs I 790 4613	Bcg I 3971 4005	Bgl II 899 959	Bsa XI 1595 2431
Bsb I 120 1554	Bsm BI 3 45	Bsp 24I 3066 3244	Bsp LU11I 1203 2573	Bsr DI 3527 3701
Bsr FI 1748 3546	Bst 1107I 1219 1245	Bst XI 893 2085	Fsp I 1902 3688	Kpn I 1126 2179
Nsp V 456 906	Pst I 1187 2125	Pvu I 1923 3836	Pvu II 1952 2397	Sap I 2457 4561
Ssp I 1439 4270	Taq II 4152 4169	Xmn I 4065 4586		

ENZYMES THAT DO NOT CUT:

*Aat*II, *Afl*II, *Age*I, *Apa*BI, *Asc*I, *Avr*II, *Bae*I, *Bpu*10I, *Bpu*1102I, *Bsg*I, *Bsp*EI, *Bsp*MI, *Bsr*GI, *Bst*EII, *Bsu*36I, *Eco*NI, *Fse*I, *Hpa*I, *Mlu*I, *Mun*I, *Nar*I, *Nco*I, *Nru*I, *Pac*I, *Pfl*MI, *Pme*I, *Psh*AI, *Rle*AI, *Rsr*II, *Sex*AI, *Sfi*I, *Sgf*I, *Sgr*AI, *Sna*BI, *Sph*I, *Srf*I, *Sse*8387I, *Stu*I, *Sty*I, *Tth*111I, *Xcm*I

continued

Table A.5.2 Yeast General-Purpose Cloning Vectors, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pRS314	4785	Yeast: TRP1 ⁺ ; <i>E. coli</i> : Amp ^r	Polylinker (see <i>COMMENTS</i>)	ATCC #77143

COMMENTS: Like pRS304 but are containing CEN/ARS cassette, making it a yeast centromere plasmid; contains KS polylinker from pBLUESCRIPT (unique sites are *KpnI*, *ApaI*, *XhoI*, *SalI*, *Clal*, *EcoRI*, *PstI*, *SmaI*, *BamHI*, *SpeI*, *EagI*, *NotI*, *SacII*, and *SacI*) and fl(+) phage DNA replication origin from pBLUESCRIBE (Sikorski and Hieter, 1989).

ENZYMES THAT CUT ONCE:

<i>AflIII</i> 2391	<i>AlwNI</i> 2807	<i>ApaI</i> 1991	<i>ApaBI</i> 547	<i>BamHI</i> 1927
<i>BbsI</i> 4431	<i>BpmI</i> 3354	<i>BsaI</i> 3345	<i>BsgI</i> 880	<i>BsmFI</i> 929
<i>BspEI</i> 332	<i>BspLU1II</i> 2391	<i>BspMI</i> 900	<i>Bst1107I</i> 434	<i>Bsu36I</i> 709
<i>Clal</i> 1964	<i>DraIII</i> 1465	<i>DrdII</i> 1470	<i>DsaI</i> 1899	<i>EagI</i> 1908
<i>Eam1105I</i> 3284	<i>EcoRI</i> 1945	<i>HgiEII</i> 2977	<i>HincII</i> 1974	<i>KpnI</i> 1997
<i>MunI</i> 607	<i>NgoAIV</i> 1566	<i>NotI</i> 1908	<i>Pfl1108I</i> 3302	<i>Psp5II</i> 4203
<i>PstI</i> 1943	<i>SacI</i> 1895	<i>SacII</i> 1902	<i>SalI</i> 1972	<i>ScaI</i> 3764
<i>SmaI</i> 1935	<i>SnaBI</i> 319	<i>SpeI</i> 1921	<i>SwaI</i> 4241	<i>XhoI</i> 1978

ENZYMES THAT CUT TWICE:

<i>AccI</i> 433 1973	<i>AvaI</i> 1933 1978	<i>BcgI</i> 734 3823	<i>BglI</i> 1713 3404	<i>BsaXI</i> 1413 2249
<i>BsmBI</i> 3 45	<i>BsrDI</i> 3345 3519	<i>BsrFI</i> 1566 3364	<i>BstXI</i> 765 1903	<i>Eco57I</i> 2939 3951
<i>EcoRV</i> 752 1953	<i>FspI</i> 1720 3506	<i>HindIII</i> 980 1957	<i>NspI</i> 41 2395	<i>PmlI</i> 427 4217
<i>PvuI</i> 1741 3654	<i>PvuII</i> 1770 2215	<i>SapI</i> 2275 4379	<i>SspI</i> 1257 4088	<i>TaqII</i> 3970 3987
<i>XbaI</i> 551 1915				

ENZYMES THAT DO NOT CUT:

AatII, *AflII*, *AgeI*, *AscI*, *AvrII*, *BaeI*, *BclI*, *BglII*, *Bpu10I*, *Bpu1102I*, *BsaBI*, *BseRI*, *BsmI*, *BspGI*, *BsrGI*, *BssHII*, *BstEII*, *Eco47III*, *EcoNI*, *FseI*, *HpaI*, *MluI*, *MscI*, *NarI*, *NcoI*, *NdeI*, *NheI*, *NruI*, *NsiI*, *NspV*, *PacI*, *PflMI*, *PmeI*, *PshAI*, *RleAI*, *RsrII*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SphI*, *SrfI*, *Sse8387I*, *StuI*, *StyI*, *SumI*, *Tth111I*, *XcmI*

continued

Table A.5.2 Yeast General-Purpose Cloning Vectors, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pRS315	6018	Yeast: LEU2 ⁺ ; <i>E. coli</i> : Amp ^r	Polylinker (see COMMENTS)	ATCC #77144

COMMENTS: Like pRS305 but containing CEN/ARS cassette, making it a yeast centromere plasmid; contains KS polylinker from pBLUESCRIPT (unique sites are *ApaI*, *XhoI*, *Sall*, *HindIII*, *PstI*, *SmaI*, *BamHI*, *SpeI*, *XbaI*, *EagI*, *NotI*, *SacII*, and *SacI*) and f1(+) phage DNA replication origin from pBLUESCRIBE (Sikorski and Hieter, 1989).

ENZYMES THAT CUT ONCE:

<i>AflIII</i> 1438	<i>AgeI</i> 1503	<i>AlwNI</i> 4040	<i>ApaI</i> 3224	<i>BaeI</i> 2157
<i>BamHI</i> 3160	<i>BcgI</i> 5056	<i>BsaI</i> 4578	<i>BsmFI</i> 1602	<i>BspMI</i> 759
<i>BsrGI</i> 632	<i>BstEII</i> 1718	<i>DraIII</i> 2698	<i>EagI</i> 3141	<i>Eam1105I</i> 4517
<i>HindIII</i> 3190	<i>HpaI</i> 2174	<i>NarI</i> 2204	<i>NgoAIV</i> 2799	<i>NotI</i> 3141
<i>PmlI</i> 5450	<i>PstI</i> 3176	<i>SacI</i> 3128	<i>SacII</i> 3135	<i>Sall</i> 3205
<i>ScaI</i> 4997	<i>SmaI</i> 3168	<i>SpeI</i> 3154	<i>SwaI</i> 5474	<i>Tth111I</i> 192
<i>XbaI</i> 3148	<i>XcmI</i> 1201	<i>XhoI</i> 3211		

ENZYMES THAT CUT TWICE:

<i>AccI</i> 561 3206	<i>AflIII</i> 1016 3624	<i>AvaI</i> 3166 3211	<i>BbsI</i> 1301 5664	<i>BcgI</i> 5022 5056
<i>BglII</i> 2946 4637	<i>BpmI</i> 902 4587	<i>BsaAI</i> 2698 5450	<i>BsaHI</i> 2204 5054	<i>BsbI</i> 120 2605
<i>BseRI</i> 405 2398	<i>BsmBI</i> 3 45	<i>Bsp24I</i> 4117 4295	<i>BspLU11I</i> 1016 3624	<i>BstXI</i> 1046 3136
<i>ClaI</i> 1603 3197	<i>DsaI</i> 892 3132	<i>EcoNI</i> 797 1487	<i>EcoRI</i> 1116 3178	<i>EcoRV</i> 1007 3186
<i>FspI</i> 2953 4739	<i>KpnI</i> 1510 3230	<i>NspV</i> 1633 2133	<i>Pfl1108I</i> 369 4535	<i>PflMI</i> 288 1859
<i>Psp5II</i> 1875 5436	<i>PvuI</i> 2974 4887	<i>PvuII</i> 3003 3448	<i>RleAI</i> 1542 2183	<i>SapI</i> 3508 5612

ENZYMES THAT DO NOT CUT:

AatII, *ApaBI*, *AscI*, *AvrII*, *BclII*, *BglIII*, *Bpu10I*, *Bpu1102I*, *BsaBI*, *BsgI*, *BsmI*, *BspEI*, *BspGI*, *BssHII*, *Bst1107I*, *Bsu36I*, *Eco47III*, *FseI*, *MluI*, *MscI*, *NcoI*, *NdeI*, *NheI*, *NruI*, *NsiI*, *PacI*, *PmeI*, *PshAI*, *RsrII*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SnaBI*, *SphI*, *SrfI*, *Sse8387I*, *StuI*, *SunI*

continued

Table A.5.2 Yeast General-Purpose Cloning Vectors, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pRS316	4895	Yeast: URA3 ⁺ ; <i>E. coli</i> : Amp ^r	Polylinker (see <i>COMMENTS</i>)	ATCC #77145

COMMENTS: Like pRS306 but containing CEN/ARS cassette, making it a yeast centromere plasmid; contains KS polylinker from pBLUESCRIPT (unique sites are *KpnI*, *XhoI*, *SalI*, *Clal*, *HindIII*, *EcoRI*, *SmaI*, *BamHI*, *SpeI*, *XbaI*, *EagI*, *NotI*, *BstXI*, *SacII*, and *SacI*) and fl(+) phage DNA replication origin from pBLUESCRIBE (Sikorski and Hieter, 1989).

ENZYMES THAT CUT ONCE:

BamHI 2037	BcgI 3933	Bpu10I 1147	BsaHI 3931	BsgI 363
BsmI 779	BsmFI 556	BspMI 402	BstXI 2013	ClaI 2074
DrdII 1580	EagI 2018	Eam1105I 3394	EcoRI 2055	HindIII 2067
KpnI 2107	NcoI 622	NdeI 330	NgoAIV 1676	NotI 2018
NsiI 1244	NspV 682	Pfl1108I 3412	PmlI 4327	RleAI 1015
SacI 2005	SacII 2012	SalI 2082	SmaI 2045	SpeI 2031
Sse8387I 399	StuI 853	SwaI 4351	XbaI 2025	XcmI 545
XhoI 2088				

ENZYMES THAT CUT TWICE:

AflIII 596 2501	AlwNI 1098 2917	ApaI 795 2101	AvaI 2043 2088	BbsI 691 4541
BcgI 3899 3933	BglII 1823 3514	BsaI 1091 3455	BsaAI 1575 4327	BsaXI 1523 2359
BspLU11I 596 2501	BsrFI 1676 3474	Bst1107I 743 1248	Eco57I 3049 4061	EcoRV 605 2063
FspI 1830 3616	HgiEII 1300 3087	Psp5II 570 4313	PstI 399 2053	PvuI 1851 3764
PvuII 1880 2325	ScaI 729 3874	SspI 1367 4198	StyI 538 622	XmnI 3993 4514

ENZYMES THAT DO NOT CUT:

AatII, *AflII*, *AgeI*, *ApaBI*, *AscI*, *AvrII*, *BaeI*, *BclI*, *BglIII*, *Bpu1102I*, *BsaBI*, *BseRI*, *BspEI*, *BspGI*, *BsrGI*, *BssHIII*, *BstEII*, *Bsu36I*, *Eco47III*, *EcoNI*, *FseI*, *HpaI*, *MluI*, *MscI*, *MunI*, *NarI*, *NheI*, *NruI*, *PacI*, *PflMI*, *PmeI*, *PshAI*, *RsrII*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SnaBI*, *SphI*, *SrfI*, *SunI*, *Tth111I*

continued

Table A.5.2 Yeast General-Purpose Cloning Vectors, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
YIp5 (Fig. 13.4.1)	5541	Yeast: selection against URA3 ⁺ by 5-FOA resistance for gene replacement schemes; <i>E. coli</i> : Amp ^r Tet ^r	See Fig. 13.4.1	ATCC #37061 (L09157)

COMMENTS: Prototypical yeast integrating plasmid; transformation occurs by integration of plasmid into yeast genome.

ENZYMES THAT CUT ONCE:

AatII 5471	ApaI 2039	AvaI 2540	BamHI 378	BsaAI 3410
BsaBI 2853	BspEI 2845	ClaI 27	EagI 942	Eam1105I 4549
EcoNI 629	EcoRI 1	HindIII 32	MscI 2627	NcoI 1866
NheI 232	NruI 977	NsiI 2488	NspV 1926	PshAI 719
PvuI 4919	PvuII 3249	RleAI 2259	SalI 654	SgrAI 413
SmaI 2542	SphI 569	Sse8387I 1643	SspI 5353	StuI 2097
Tth111I 3403	VspI 4721	XcmI 1789		

ENZYMES THAT CUT TWICE:

AflIII 1840 3656	AlwNI 2342 4072	ApaBI 1055 3481	Bpu10I 2391 2762	BsaI 2335 4610
BsgI 1607 2816	BsmI 1362 2023	BsmBI 2283 3299	BspLU11I 1840 3656	BspMI 1057 1646
BsrBI 3589 5390	DrdI 3351 3764	Eco57I 4204 5216	EcoRV 190 1849	HgiEII 3481 4242
NdeI 1574 3479	Pfl1108I 761 4567	PflMI 1324 1373	PstI 1643 4794	SapI 2171 3540
ScaI 1973 5029	XmnI 3216 5148			

ENZYMES THAT DO NOT CUT:

AflII, AgeI, AscI, AvrII, BaeI, BclI, BglII, Bpu1102I, BsaXI, BseRI, BsrGI, BssHII, BstEII, BstXI, Bsu36I, DrdII, FseI, HpaI, KpnI, MluI, MunI, NotI, PacI, PmeI, PmlI, RsrII, SacI, SacII, SexAI, SfiI, SgfI, SnaBI, SpeI, SrfI, SunI, SwaI, XbaI, XhoI

Table A.5.3 Vectors for Construction of DNA Libraries in *E. coli*

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
λ EMBL3 (Fig. 1.10.2)	42364 ^a	Recombinant phage are Spi ⁻	<i>EcoRI</i> , <i>BamHI</i> , <i>SalI</i>	Stratagene, Promega (U02425—left arm; U02453—right arm)
<i>COMMENTS:</i> Used for cloning large (10.4-20 kbp) fragments; see Fig. 1.10.2 for restriction enzyme information.				
λ gt10 (Fig. 1.10.4)	43340 ^a	Recombinant phage are cI ⁻ , Int ⁻ , Red ⁺	<i>EcoRI</i>	Stratagene, Promega (U02447—region flanking the cloning site)
<i>COMMENTS:</i> Used for cloning 0-5 kb DNA fragments and for nucleic acid probe screening; see Fig. 1.10.4 for restriction enzyme information.				
λ gt11 ^b (Fig. 1.10.5)	43700 ^a	Recombinant phage are cI ^{ts} , Int ⁺ , Red ⁺ ; fusion proteins are Lac ⁺ on Lac ⁻ hosts	<i>EcoRI</i>	Stratagene, Promega (X04122— <i>lac</i> operon and phage junction from <i>SacI</i> to <i>KpnI</i>)

COMMENTS: Used for cloning up to 7.8-kbp DNA fragments to create cDNA libraries; screen with antibodies against native protein or nucleic acid probes; see Fig. 1.10.5 for restriction enzyme information.

^aSize information provided by Stratagene.

^bParticularly suitable for construction of cDNA libraries.

continued

Table A.5.3 Vectors for Construction of DNA Libraries in *E. coli*, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
λZAP and λZAPII^b (Fig. 1.10.8)	40820 ^a	Dephosphorylated arms select against nonrecombinants; recombinants containing fusion proteins are Lac ⁺ on appropriate host (detectable by blue/white screen using Xgal and IPTG)	<i>Xho</i> I, <i>Eco</i> RI, <i>Spe</i> I, <i>Sba</i> I, <i>Not</i> I, <i>Sac</i> I	Stratagene (X52324—pBluescript SK ⁻ only)

COMMENTS: Used for making cDNA libraries; libraries can be screened with antibody or nucleic acid probes; cloning sites are within pBluescript SK⁻ (2958 bp), which can be excised with helper phage and recircularized to generate subclones and single-stranded DNA. λZAP contains the *s100* mutation, which is not present in λZAPII. **IMPORTANT NOTE:** For λZAP restriction site information, see Fig. 1.10.8; restriction site data for pBluescript SK⁻ follow.

ENZYMES THAT CUT ONCE:

Acc I 675	Afl III 1153	Alw NI 1569	Apa I 663	Bam HI 719
Bcg I 2551	Bsa I 2107	Bsa AI 237	Bsa HI 2583	Bsb I 330
Bsp LU11I 1153	Bst XI 751	Cl aI 684	Dra III 240	Drd II 232
Dsa I 747	Eag I 738	Eam 1105I 2046	Eco O109I 660	Eco RI 701
Eco RV 697	Hgi EII 1739	Hinc II 676	Hind III 689	Kpn I 657
Ngo AIV 132	Not I 738	Nsp I 1157	Pfl 1108I 2064	Pst I 711
Sac I 759	Sac II 750	Sal I 674	Sap I 1037	Sca I 2526
Sma I 715	Spe I 725	Xba I 731	Xho I 668	Xmn I 2645

ENZYMES THAT CUT TWICE:

Alw 44I 1467 2713	Ava I 668 713	Ava II 2184 2406	Bcg I 2551 2585	Bgl II 475 2166
Bpm I 743 2116	Bsi II 1326 2710	Bsm AI 2107 2883	Bsp 24I 1678 1856	Bsr DI 2107 2281
Bsr FI 132 2126	Drd I 284 1261	Eco 57I 1701 2713	Fsp I 482 2268	Psp 1406I 2272 2645
Pvu I 503 2416	Pvu II 532 977	Rca I 1873 2881	Rsa I 655 2526	Ssp I 445 2850
Tfi I 988 1128				

ENZYMES THAT DO NOT CUT:

*Aat*II, *Afl*II, *Age*I, *Apa*BI, *Asc*I, *Avr*II, *Bae*I, *Bbs*I, *Bcl*I, *Bgl*II, *Bpu*10I, *Bpu*1102I, *Bsa*BI, *Bse*RI, *Bsg*I, *Bsm*I, *Bsm*BI, *Bsm*FI, *Bsp*EI, *Bsp*GI, *Bsp*MI, *Bsr*GI, *Bss*HII, *Bst*1107I, *Bst*EII, *Bsu*36I, *Eco*47III, *Eco*NI, *Fse*I, *Hpa*I, *Mlu*I, *Msc*I, *Mun*I, *Nar*I, *Nco*I, *Nde*I, *Nhe*I, *Nru*I, *Nsi*I, *Nsp*V, *Pac*I, *Pfl*MI, *Pme*I, *Pml*I, *Psh*AI, *Psp*5II, *Rle*AI, *Rsr*II, *Sex*AI, *Sfi*I, *Sgf*I, *Sgr*AI, *Sna*BI, *Sph*I, *Srf*I, *Sse*8387I, *Stu*I, *Sty*I, *Sun*I, *Swa*I, *Tth*111I, *Xcm*I

pJB8 (Fig. 1.10.10)	5400	Amp ^r	<i>Bam</i> HI	ATCC #37074
-------------------------------	------	------------------	---------------	-------------

COMMENTS: Representative cosmid vector bearing λ *cos* site that allows packaging of insert-containing cosmids for library construction. See Fig. 1.10.10 for restriction site information.

Table A.5.4 Vectors for Construction of Artificial Chromosomes in Yeast

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pYAC3 (Fig. 13.4.6)	11446	Yeast: URA ⁺ , TRP ⁺ , and red/white color screen using suppressible <i>ade2-1</i> ochre mutation on appropriate host; <i>E. coli</i> : Amp ^r	<i>Sna</i> BI	ATCC #37520

COMMENTS: Used to clone very large DNA fragments to create artificial linear chromosomes that are stably maintained in yeast.

ENZYMES THAT CUT ONCE:

<i>Aat</i> II 8931	<i>Afl</i> III 4487	<i>Age</i> I 423	<i>Apa</i> I 2320	<i>Bsp</i> EI 3117
<i>Bss</i> III 4719	<i>Bsu</i> 36I 9350	<i>Drd</i> II 409	<i>Hpa</i> I 10593	<i>Kpn</i> I 4913
<i>Nco</i> I 2147	<i>Nru</i> I 1258	<i>Pml</i> II 9068	<i>Psh</i> AI 1000	<i>Sal</i> I 935
<i>Sgr</i> AI 694	<i>Sma</i> I 583	<i>Sna</i> BI 552	<i>Spe</i> I 11396	<i>Sph</i> I 850
<i>Sun</i> I 5115	<i>Xba</i> I 9192	<i>Xcm</i> I 2070		

ENZYMES THAT CUT TWICE:

<i>Alw</i> NI 2623 7532	<i>Avr</i> II 10733 11183	<i>Bam</i> HI 4229 5997	<i>Bpu</i> 10I 2672 3034	<i>Bsa</i> I 2616 8070
<i>Bse</i> RI 4509 5735	<i>Bsm</i> BI 2564 6759	<i>Bsr</i> GI 9930 10100	<i>Bst</i> XI 5146 9406	<i>Cla</i> I 24 4298
<i>Drd</i> I 6811 7224	<i>Eag</i> I 1223 4319	<i>Eam</i> 1105I 5937 8009	<i>Hgi</i> EII 6941 7702	<i>Msc</i> I 2899 5324
<i>Nsi</i> II 2769 4756	<i>Pfl</i> 1108I 1042 8027	<i>Pfl</i> MI 1605 1654	<i>Rle</i> AI 2540 5900	<i>Sap</i> I 2452 7000
<i>Sse</i> 8387I 1924 9833	<i>Stu</i> I 2378 9835	<i>Tth</i> 111I 6863 11007	<i>Xho</i> I 3524 6702	

ENZYMES THAT DO NOT CUT:

*Asc*I, *Bae*I, *Bpu*1102I, *Bst*EII, *Eco*RI, *Fse*I, *Mlu*I, *Not*I, *Pme*I, *Pvu*II, *Rsr*II, *Sac*I, *Sac*II, *Sex*AI, *Sfi*I, *Sgf*I, *Srf*I, *Swa*I

pYAC4 (Figs. 6.10.1 and 13.4.6)	11454	Yeast: URA ⁺ , TRP ⁺ , and red/white color screen using suppressible <i>ade2-1</i> ochre mutation on appropriate host; <i>E. coli</i> : Amp ^r	<i>Eco</i> RI	ATCC #67379 (U01086)
-------------------------------------------	-------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------	----------------------

COMMENTS: Used to clone very large DNA fragments to create artificial linear chromosomes that are stably maintained in yeast; identical to YAC3 except for *Eco*RI site replacing *Sna*BI cloning site.

ENZYMES THAT CUT ONCE:

<i>Aat</i> II 8939	<i>Afl</i> III 4495	<i>Age</i> I 423	<i>Apa</i> I 2328	<i>Bsp</i> EI 3125
<i>Bss</i> III 4727	<i>Bsu</i> 36I 9358	<i>Drd</i> II 409	<i>Eco</i> RI 554	<i>Hpa</i> I 10601
<i>Kpn</i> I 4921	<i>Nco</i> I 2155	<i>Nru</i> I 1266	<i>Pml</i> II 9076	<i>Psh</i> AI 1008
<i>Sal</i> I 943	<i>Sgr</i> AI 702	<i>Sma</i> I 591	<i>Spe</i> I 11404	<i>Sph</i> I 858
<i>Sun</i> I 5123	<i>Xba</i> I 9200	<i>Xcm</i> I 2078		

ENZYMES THAT CUT TWICE:

<i>Alw</i> NI 2631 7540	<i>Avr</i> II 10741 11191	<i>Bam</i> HI 4237 6005	<i>Bpu</i> 10I 2680 3042	<i>Bsa</i> I 2624 8078
<i>Bse</i> RI 4517 5743	<i>Bsm</i> BI 2572 6767	<i>Bsr</i> GI 9938 10108	<i>Bst</i> XI 5154 9414	<i>Cla</i> I 24 4306
<i>Drd</i> I 6819 7232	<i>Eag</i> I 1231 4327	<i>Eam</i> 1105I 5945 8017	<i>Hgi</i> EII 6949 7710	<i>Msc</i> I 2907 5332
<i>Nsi</i> II 2777 4764	<i>Pfl</i> 1108I 1050 8035	<i>Pfl</i> MI 1613 1662	<i>Rle</i> AI 2548 5908	<i>Sap</i> I 2460 7008
<i>Sse</i> 8387I 1932 9841	<i>Stu</i> I 2386 9843	<i>Tth</i> 111I 6871 11015	<i>Xho</i> I 3532 6710	

ENZYMES THAT DO NOT CUT:

*Asc*I, *Bae*I, *Bpu*1102I, *Bst*EII, *Fse*I, *Mlu*I, *Not*I, *Pme*I, *Pvu*II, *Rsr*II, *Sac*I, *Sac*II, *Sex*AI, *Sfi*I, *Sgf*I, *Sna*BI, *Srf*I, *Swa*I

Table A.5.5 Vectors for Expression of Genes from Heterologous Promoters in *E. coli*

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
λZAP and λZAPII See Table A.5.3 and Fig. 1.10.8 for full listings.				
pSKF101 (Fig. 16.3.1)	5866	Amp ^r	See Fig. 16.3.1	Allan Shatzman, SmithKline Beecham

COMMENTS: Used for authentic gene cloning using λ pL promoter and cII ribosome binding site for expression; also known as pAS1.

ENZYMES THAT CUT ONCE:

AatII 5796	AflIII 3981	AlwNI 4397	BamHI 1882	BglII 1216
BsaI 4935	BsaAI 3735	BsgI 3141	BsmBI 3624	BspLU11I 3981
Bst1107I 3754	BstEII 552	BstXI 1819	ClaI 27	Eam1105I 4874
EcoRI 1	HindIII 32	NruI 2480	PshAI 2222	PstI 5119
PvuII 3574	SalI 2157	SapI 3865	ScaI 5354	SgrAI 1917
SphI 2072	SspI 5678	XcmI 760		

ENZYMES THAT CUT TWICE:

AccI 2158 3753	ApaBI 2558 3806	AvaI 1734 2931	BanII 1982 1996	Bpu10I 1114 3087
BsmI 1415 2865	BspEI 1478 3170	BspMI 624 2560	DrdI 3676 4089	DrdII 238 826
DsaI 2034 2953	EagI 273 2445	MscI 889 2952	MunI 1163 1852	Pfl1108I 2264 4892
PflMI 2827 2876	Psp5II 2945 2987	PvuI 1143 5244	RsaI 3789 5354	StyI 422 2875
TaqII 5560 5577	Th111I 807 3728	VspI 654 5046	XmnI 3541 5473	

ENZYMES THAT DO NOT CUT:

AflII, AgeI, ApaI, AscI, AvrII, BaeI, BclI, Bpu1102I, BseRI, BsrGI, BssHII, Bsu36I, DraIII, EcoRV, FseI, HpaI, KpnI, MluI, NcoI, NheI, NotI, NspV, PacI, PmeI, PmlI, RleAI, RsrII, SacI, SacII, SexAI, SfiI, SgfI, SmaI, SnaBI, SpeI, SrfI, Sse8387I, StuI, SunI, SwaI, XbaI, XhoI

pSKF301 (Fig. 16.3.2)	5192	Amp ^r	Polylinker (see Fig. 16.3.2)	Allan Shatzman, SmithKline Beecham
---------------------------------	------	------------------	---------------------------------	------------------------------------

COMMENTS: Similar to pSKF101, but also allows construction of protein fusions to first 81 amino acids of influenza protein NS1; also known as pMG1.

ENZYMES THAT CUT ONCE:

AatII 5120	AccI 3075	AlwNI 3721	ApaBI 1880	BamHI 994
BanII 1277	BglII 328	BsaI 4259	BsgI 2463	BsmBI 2946
BspLU11I 3305	Bst1107I 3076	BstXI 931	ClaI 1291	EagI 1588
Eam1105I 4198	EcoRV 1252	HpaI 1245	MscI 2274	NarI 2034
NcoI 1230	PmlI 1166	PshAI 1365	PstI 4443	PvuII 2896
SacI 1277	SapI 3189	ScaI 4678	SspI 5002	StuI 1259
Th111I 3050	VspI 4370	XbaI 1284	XcmI 1809	XhoI 1278

ENZYMES THAT CUT TWICE:

AflIII 1163 3305	ApoI 368 2718	Bpu10I 226 2409	BsaAI 1166 3057	BsaBI 394 2500
BsmI 527 2187	BspEI 590 2492	BspMI 1813 1882	DrdI 2998 3413	DsaI 1230 2275
Eco47III 1426 2557	EcoNI 517 904	HindIII 1012 1296	MunI 275 964	NdeI 990 1239
NsiI 175 943	Pfl1108I 1407 4216	PflMI 2149 2198	Psp5II 2267 2309	PvuI 255 4568
RsaI 3111 4678	StyI 1230 2197	TaqII 4884 4901	XmnI 2863 4797	

ENZYMES THAT DO NOT CUT:

AflII, AgeI, ApaI, AscI, AvrII, BaeI, BclI, Bpu1102I, BseRI, BsrGI, BssHII, BstEII, Bsu36I, DraIII, DrdII, EcoRI, FseI, KpnI, MluI, NheI, NotI, NruI, NspV, PacI, PmeI, RleAI, RsrII, SacI, SalI, SexAI, SfiI, SgfI, SgrAI, SmaI, SnaBI, SpeI, SphI, SrfI, Sse8387I, SunI, SwaI

continued

Table A.5.5 Vectors for Expression of Genes from Heterologous Promoters in *E. coli*, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pSP64	See Table A.5.1 and Fig. 1.5.3 for full listings.			
pT7-5 (Fig. 16.2.1)	2404	Amp ^r	Polylinker (see Fig. 16.2.1)	Stanley Tabor, Harvard Medical School

COMMENTS: Used to express genes under control of T7 RNA polymerase (see mGP1-2 and pGP1-2 in Table A.5.10); does not contain ribosome binding site or ATG codon upstream of polylinker.

ENZYMES THAT CUT ONCE:

AatII 1371	AccI 47	AflIII 533	AlwNI 949	ApaBI 358
ApoI 76	AvaI 63	BamHI 58	BanI 2337	BanII 74
BbsI 1307	BceII 1035	BcgI 1786	BpmI 2221	BsaAI 285
BsmBI 174	BspLU11I 533	BspMI 33	BsrFI 2205	ClaI 26
DraI 1712	EaeI 1897	Eam1105I 2290	EcoO109I 1310	EcoRI 76
FspI 2067	GdiII 1897	HindIII 31	NdeI 356	Pfl1108I 2271
PstI 44	PvuI 1921	PvuII 122	SacI 74	SalI 46
SapI 417	ScaI 1809	SmaI 65	SspI 1485	TfiI 508
Tth111I 278	XbaI 52	XmnI 1690		

ENZYMES THAT CUT TWICE:

AvaII 1926 2148	BcgI 1752 1786	BglII 124 2172	BsaI 89 2224	BsaHI 1368 1750
BsbI 249 1746	BsrBI 466 1448	BsrDI 2056 2230	CjePI 711 1642	DrdI 226 641
EarI 417 1491	Eco57I 1081 1624	FauI 236 246	HaeII 411 781	HgiEII 358 1119
HincII 48 1748	MmeI 748 932	NspI 170 537	Psp1406I 1688 2061	RsaI 341 1809
VspI 105 2115				

ENZYMES THAT DO NOT CUT:

AflII, AgeI, ApaI, AscI, AvrII, BaeI, BclI, BglII, Bpu10I, Bpu1102I, BsaBI, BseRI, BsgI, BsmI, BsmFI, BspEI, BspGI, BsrGI, BssHII, Bst1107I, BstEII, BstXI, Bsu36I, DraIII, DrdII, DsaI, EagI, Eco47III, EcoNI, EcoRV, FseI, HpaI, KpnI, MluI, MscI, MunI, NarI, NcoI, NgoAIV, NheI, NotI, NruI, NsiI, NspV, PacI, PflMI, PmeI, PmlI, PshAI, Psp5II, RleAI, RsrII, SacII, SexAI, SfiI, SgfI, SgrAI, SnaBI, SpeI, SphI, SrfI, Sse8387I, StuI, StyI, SunI, SwaI, XcmI, XhoI

continued

Table A.5.5 Vectors for Expression of Genes from Heterologous Promoters in *E. coli*, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pT7-6 (Fig. 16.2.1)	2208	Amp ^r	Polylinker (see Fig. 16.2.1)	Stanley Tabor, Harvard Medical School

COMMENTS: Used to express genes under control of T7 RNA polymerase (see mGP1-2 and pGP1-2 in Table A.5.12); does not contain ribosome-binding site or ATG codon upstream of polylinker.

ENZYMES THAT CUT ONCE:

AccI 59	AflIII 530	AlwNI 946	ApaBI 355	ApoI 28
AvaI 41	BamHI 46	BanI 2141	BanII 38	BceII 1032
BcgI 1590	BpmI 2025	BsaAI 282	BsaHI 1554	BsmBI 171
BspLU11I 530	BspMI 71	BsrBI 463	BsrFI 2009	DraI 1516
EaeI 1701	Eam1105I 2094	EcoRI 28	FspI 1871	GdiII 1701
HindIII 73	NdeI 353	Pfl1108I 2075	PstI 68	PvuI 1725
PvuII 119	RcaI 1250	SacI 38	SalI 58	SapI 414
ScaI 1613	SmaI 43	TfiI 505	Tth111I 275	XbaI 52
XmnI 1494				

ENZYMES THAT CUT TWICE:

AvaII 1730 1952	BcgI 1556 1590	BglII 121 1976	BsaI 86 2028	BsbI 246 1550
BsiI 703 1425	Bsp24I 1055 1233	BsrDI 1860 2034	CjePI 708 1446	DrdI 223 638
EarI 414 1295	Eco57I 1078 1428	FauI 233 243	HaeII 408 778	HgiEII 355 1116
HincII 60 1552	MmeI 745 929	NspI 167 534	Psp1406I 1492 1865	RsaI 338 1613
VspI 102 1919				

ENZYMES THAT DO NOT CUT:

AatII, AflII, AgeI, ApaI, AscI, AvrII, BaeI, BbsI, BclI, BglII, Bpu10I, Bpu1102I, BsaBI, BseRI, BsgI, BsmI, BsmFI, BspEI, BspGI, BsrGI, BssHII, Bst1107I, BstEII, BstXI, Bsu36I, ClaI, DraIII, DrdII, DsaI, EagI, Eco47III, EcoNI, EcoO109I, EcoRV, FseI, HpaI, KpnI, MluI, MscI, MunI, NarI, NcoI, NgoAIV, NheI, NotI, NruI, NsiI, NspV, PacI, PflMI, PmeI, PmlI, PshAI, Psp5II, RleAI, RsrII, SacII, SexAI, SfiI, Sgfl, SgrAI, SnaBI, SpeI, SphI, SrfI, Sse8387I, SspI, StuI, StyI, SunI, SwaI, XcmI, XhoI

continued

Table A.5.5 Vectors for Expression of Genes from Heterologous Promoters in *E. coli*, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pT7-7 (Fig. 16.2.1)	2488	Amp ^r	Polylinker (see Fig. 16.2.1)	Stanley Tabor, Harvard Medical School

COMMENTS: Used to express genes under control of T7 RNA polymerase (see mGP1-2 and pGP1-2 in Table A.5.10); contains ribosome-binding site and ATG codon upstream of polylinker.

ENZYMES THAT CUT ONCE:

AatII 1455	AccI 50	AflIII 617	AlwNI 1033	ApoI 75
AvaI 66	BamHI 61	BanI 2421	BbsI 1391	BceII 1119
BcgI 1870	BglII 2256	BglIII 434	BpmI 2305	BsaAI 359
BsmBI 248	BspLU11I 617	BspMI 36	BsrFI 2289	ClaI 26
DraI 1796	EaeI 1981	Eam1105I 2374	EcoO109I 1394	EcoRI 75
FspI 2151	GdiII 1981	HgiEII 1203	HindIII 34	NdeI 86
Pfl1108I 2355	PstI 47	PvuI 2005	SalI 49	SapI 501
ScaI 1893	SmaI 68	SspI 1569	Tth111I 352	XmnI 1774

ENZYMES THAT CUT TWICE:

AvaII 2010 2232	BcgI 1836 1870	BsaI 149 2308	BsaHI 1452 1834	BsbI 323 1830
BsrBI 550 1532	BsrDI 2140 2314	CjePI 795 1726	DrdI 300 725	EarI 501 1575
Eco57I 1165 1708	FauI 310 320	HaeII 495 865	HincII 51 1832	MmeI 832 1016
NspI 244 621	NspV 171 189	Psp1406I 1772 2145	RsaI 415 1893	TfiI 191 592
VspI 165 2199	XbaI 55 124			

ENZYMES THAT DO NOT CUT:

AflII, AgeI, ApaI, ApaBI, AscI, AvrII, BaeI, BanII, BclII, Bpu10I, Bpu1102I, BsaBI, BseRI, BsgI, BsmI, BsmFI, BspEI, BspGI, BsrGI, BssHII, Bst1107I, BstEII, BstXI, Bsu36I, DraIII, DrdII, DsaI, EagI, Eco47III, EcoNI, EcoRV, FseI, HpaI, KpnI, MluI, MscI, MunI, NarI, NcoI, NgoAIV, NheI, NotI, NruI, NsiI, PacI, PflMI, PmeI, PmlI, PshAI, Psp5II, PvuII, RleAI, RsrII, SacI, SacII, SexAI, SfiI, SgfI, SgrAI, SnaBI, SpeI, SphI, SrfI, Sse8387I, StuI, StyI, SunI, SwaI, XcmI, XhoI

continued

Table A.5.5 Vectors for Expression of Genes from Heterologous Promoters in *E. coli*, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pTrc99A (Fig. 1.5.11)	4176	Amp ^r	Polylinker (see Fig. 1.5.11)	Pharmacia Biotech (M22744)

COMMENTS: Uses hybrid *trp/lac* promoter, *lacZ* ribosome binding site, pUC18 multiple cloning site, and *rrnB* transcription terminators. Multiple cloning site is inserted behind an *NcoI* site in three reading frames to give pTrc99A, B, and C. Cloning into *NcoI* site gives expression of authentic protein; cloning into multiple cloning site gives fusion protein. *lacI^q* repressor gene is present for repression of *trp/lac* promoter.

ENZYMES THAT CUT ONCE:

AlwNI 2112	ApaI 3616	AvaI 286	BamHI 291	BclII 3419
BglII 1515	BmgI 3614	BsaAI 2771	BsaXI 4064	BsmFI 417
BspLU11I 2521	Bst1107I 2752	BstEII 3586	DrdII 3128	DsaI 265
Eam1105I 1633	EcoRI 270	EcoRV 3855	FspI 1410	HindIII 321
KpnI 286	MluI 3405	NarI 4045	NcoI 265	NdeI 2700
NsiI 2968	Pfl1108I 1614	PflMI 2987	PstI 313	PvuI 1264
RleAI 478	SacI 280	SalI 303	SapI 2638	ScaI 1152
SmaI 288	Sse8387I 313	StyI 265	Tth1111I 2777	XbaI 297

ENZYMES THAT CUT TWICE:

AccI 304 2751	AflIII 2521 3405	ApaBI 2705 3089	BanII 280 3616	BbsI 3551 3890
BcgI 1129 3697	BsaI 499 1567	BsbI 1089 2807	BsiI 964 2348	BsmBI 2878 4020
BspMI 72 316	BsrFI 1548 3091	BssHII 2944 3816	DrdI 2419 2832	EaeI 1240 4079
Eco57I 967 1979	GdiII 1240 4079	MmeI 2128 2312	PvuII 4005 4098	RcaI 793 1801
SphI 319 2966	SspI 176 828	TfiI 2547 4084	XmnI 351 1033	

ENZYMES THAT DO NOT CUT:

AatII, AflII, AgeI, AscI, AvrII, BaeI, BglII, Bpu10I, Bpu1102I, BsaBI, BseRI, BsmI, BspEI, BspGI, BsrGI, Bsu36I, ClaI, DraIII, EagI, Eco47III, EcoNI, EcoO109I, FseI, HpaI, MscI, MumI, NgoAIV, NheI, NotI, NruI, NspV, PacI, PmeI, PmlI, PshAI, Psp5II, RsrII, SacII, SexAI, SfiI, SgfI, SgrAI, SnaBI, SpeI, SrfI, StuI, SunI, SwaI, XhoI

Table A.5.6 Vectors for Expression of Genes from Heterologous Promoters in Mammalian Cells

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
BAG retrovirus	12219	Mammalian: G418 ^r ; <i>E. coli</i> : Neo ^r	Polylinker	Constance Cepko, Harvard Medical School

COMMENTS: Used for stable transduction of mitotic cells; partial listing of restriction sites below.

ENZYMES THAT CUT ONCE:

AsuII 5674 *BglIII* 4828 *BspMII* 8340 *SalI* 4156 *SfiI* 4444
SspI 2298 *SmnI* 10398

ENZYMES THAT CUT TWICE:

AatII 812 1691 *AvrII* 4491 12087 *BamHI* 1074 4146 *ClaI* 1892 6911 *DraI* 8100 10949
FspI 1210 5094 *HindIII* 4507 9647 *HpaI* 1494 2118 *MstII* 1292 9186 *NaeI* 5494 5777
NheI 31 7082 *RsrII* 5508 8343 *SpeI* 727 6722 *XmaIII* 791 4898

ENZYMES THAT DO NOT CUT:

BstEII, *EcoK*, *NotI*, *NruI*, *SacII*, *SnaBI*

CDM8 (Fig. 16.12.1)	4357	<i>E. coli</i> : suppression of amber mutations (<i>supF</i>) in <i>amp</i> and <i>tet</i> genes in strains containing P3	<i>HindIII</i> , <i>XbaI</i> , <i>XhoI</i> , <i>BstXI</i> , <i>PstI</i> , <i>NotI</i>	Brian Seed, Massachusetts General Hospital and Harvard Medical School
-------------------------------	------	-----------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------	-----------------------------------------------------------------------

COMMENTS: Used for high levels of transient expression of cloned genes in COS cells; allows production of single-stranded DNA in *E. coli*.

ENZYMES THAT CUT ONCE:

AflIII 1529 *Alw44I* 312 *ApaI* 4017 *BaeI* 3412 *BamHI* 3346
BbsI 2952 *BclI* 3735 *Bpu10I* 1481 *BsaI* 2177 *BsaBI* 3114
BsgI 3036 *BsiI* 453 *BspMI* 2619 *DraIII* 830 *DrdII* 835
EagI 2636 *EcoNI* 3681 *HgiEII* 44 *HindIII* 2190 *HpaI* 3215
KpnI 3411 *MluI* 1529 *NdeI* 1785 *NheI* 587 *NotI* 2636
NruI 1509 *PflMI* 2804 *PshAI* 1178 *Psp5II* 3389 *Psp1406I* 612
PstI 2630 *RleAI* 3181 *SacI* 2119 *SacII* 1179 *ScaI* 3503
SfiI 4285 *SnaBI* 1891 *SpeI* 1550 *Sse8387I* 2630

ENZYMES THAT CUT TWICE:

AlwNI 218 2563 *ApaBI* 2498 2564 *AvrII* 3421 4332 *Bce83I* 244 541 *BsaWI* 273 420
BsiEI 292 2639 *BsrFI* 931 2410 *BstXI* 2227 2585 *DraI* 2717 3154 *DrdI* 524 782
FokI 3844 4183 *HincII* 1535 3215 *HphI* 831 1927 *MslI* 1916 3970 *MunI* 1462 3202
NciI 250 4010 *NgoAIV* 931 2410 *NspV* 1222 1230 *SspI* 619 3083 *StuI* 3944 4331
TaqII 282 1472 *VspI* 1558 2163 *XbaI* 2195 2646 *XhoI* 2212 2592

ENZYMES THAT DO NOT CUT:

AflII, *AgeI*, *AscI*, *BcgI*, *BcgI*, *BglII*, *Bpu1102I*, *BsmBI*, *BspEI*, *BspLU11I*, *BsrDI*, *BsrGI*, *BssHII*, *Bst1107I*, *BstEII*, *Bsu36I*, *ClaI*, *Eam1105I*, *Eco47III*, *EcoRI*, *EcoRV*, *FseI*, *FspI*, *MscI*, *NarI*, *NsiI*, *NspI*, *PacI*, *Pfl1108I*, *PmeI*, *PmlI*, *PvuI*, *RcaI*, *RsrII*, *SallI*, *SapI*, *SexAI*, *SgfI*, *SgrAI*, *SmaI*, *SphI*, *SrfI*, *SunI*, *SwaI*, *Tth111I*, *XcmI*, *XmnI*

continued

Table A.5.6 Vectors for Expression of Genes from Heterologous Promoters in Mammalian Cells, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pEB2 (Fig. 16.19.1)	4602	Mammalian: BrDU ^r (TK ⁻); <i>E. coli</i> : Amp ^r	<i>EcoRI</i> , <i>StuI</i>	Bernard Moss, National Institute of Allergy and Infectious Disease

COMMENTS: Used for transient cytoplasmic expression of cloned genes under control of T7 RNA polymerase promoter; can be used to construct recombinant vaccinia virus vector by homologous recombination.

ENZYMES THAT CUT ONCE:

AatII 77	AccI 3074	AflIII 1884	AlwNI 1475	ApaBI 4377
BaeI 3649	BbsI 13	BcgI 492	BclI 3806	BpmI 927
Bpu1102I 3130	BsaAI 4441	BsbI 4477	BseRI 3989	BsmBI 4548
BspLU11I 1884	BsrFI 911	BstXI 4094	DrdII 3615	EagI 3831
Eam1105I 996	EcoRI 3051	HincII 3075	HindIII 2243	MunI 2472
NarI 4320	NdeI 4372	PacI 3069	Psp5II 2522	RleAI 4133
SalI 3073	SapI 2001	ScaI 515	StuI 3059	StyI 3152
Tth111I 4447	XbaI 2381	XmnI 396		

ENZYMES THAT CUT TWICE:

BcgI 458 492	BglII 878 4310	BsaI 930 3014	BsrDI 762 936	DrdI 1782 4502
DsaI 2431 4052	Eco57I 330 1342	FspI 773 4300	HgiEII 1302 4372	MmeI 1491 1675
NspI 1888 4560	Pfl1108I 977 3909	Psp1406I 394 767	PvuI 627 4281	PvuII 2064 4250

ENZYMES THAT DO NOT CUT:

AflII, *AgeI*, *ApaI*, *AscI*, *AvaI*, *AvrII*, *BamHI*, *BanII*, *BglIII*, *Bpu10I*, *BsgI*, *BsmI*, *BspEI*, *BspGI*, *BspMI*, *BsrGI*, *BssHIII*, *Bst1107I*, *BstEII*, *Bsu36I*, *Clal*, *DraIII*, *Eco47III*, *EcoNI*, *EcoRV*, *FseI*, *HpaI*, *KpnI*, *MluI*, *MscI*, *NcoI*, *NgoAIV*, *NheI*, *NotI*, *NruI*, *NsiI*, *NspV*, *PflMI*, *PmeI*, *PmlI*, *PshAI*, *PstI*, *RsrII*, *SacI*, *SacII*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SmaI*, *SnaBI*, *SpeI*, *SrfI*, *Sse8387I*, *SunI*, *SwaI*, *XcmI*, *XhoI*

pED (Fig. 16.23.1)	5370	CHO: DHFR; <i>E. coli</i> : Amp ^r	<i>PstI</i> , <i>EcoRI</i> (see Fig. 16.14.1)	Randal Kaufman, University of Michigan—Ann Arbor
------------------------------	------	----------------------------------------------	-----------------------------------------------	--------------------------------------------------

COMMENTS: Used for high-level expression of targeted gene by amplification in stably transfected CHO cells.

ENZYMES THAT CUT ONCE:

AflIII 1956	ApaI 1202	BcgI 4490	BclI 2656	BglII 979
BsaBI 2315	BsbI 5092	Clal 2662	DraIII 1448	EagI 2773
Eam1105I 3985	Eco47III 2848	EcoNI 2167	EcoRI 1079	HpaI 2416
KpnI 1530	MscI 986	MunI 2403	NarI 5208	NdeI 5156
NheI 2887	NspV 2759	PshAI 513	Psp5II 753	PstI 1059
RleAI 2382	SalI 1061	SapI 2976	SfiI 71	SmaI 1075
SspI 4789	StuI 22	Tth111I 2628	XbaI 1067	XhoI 1655

ENZYMES THAT CUT TWICE:

AatII 2832 4907	AvaI 1073 1655	AvrII 17 1236	BcgI 4490 4524	BpmI 480 4055
Bpu10I 757 1747	BsaI 1721 4046	BsaAI 566 1401	BspGI 429 2175	BspMI 1048 1424
BsrFI 1138 4065	BstEII 2567 2796	BstXI 1944 2564	DrdI 3200 5069	FspI 4207 5230
HgiEII 3678 5158	HincII 1063 2416	NcoI 110 2288	Pfl1108I 2213 4003	PmlI 566 1401
PvuI 4355 5251	PvuII 656 5280	SacII 379 895	SphI 205 277	

ENZYMES THAT DO NOT CUT:

AgeI, *AscI*, *BaeI*, *Bpu1102I*, *BsgI*, *BspEI*, *BsrGI*, *BssHIII*, *Bst1107I*, *Bsu36I*, *EcoRV*, *FseI*, *MluI*, *NgoAIV*, *NotI*, *NruI*, *PacI*, *PmeI*, *RsrII*, *SgfI*, *SgrAI*, *SnaBI*, *SpeI*, *SrfI*, *Sse8387I*, *SunI*, *SwaI*, *XcmI*

continued

Table A.5.6 Vectors for Expression of Genes from Heterologous Promoters in Mammalian Cells, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pEE14 (Fig. 16.23.3)	9400	CHO: methionine sulfoximine–resistance (glutamine synthetase); <i>E. coli</i> : Amp ^r	<i>Hind</i> III, <i>Xba</i> I, <i>Xma</i> I, <i>Sma</i> I, <i>Eco</i> RI, <i>Bcl</i> I	Celltech

COMMENTS: Used for expression of targeted genes in stably transfected CHO cells.

ENZYMES THAT CUT ONCE:

Acc3 8844	Acc65I 4697	Afl2 9078	Age1 4694	Aos1 1606
Apa1 3731	Asp718 4697	BamHI 249	Bcl1 9	Bfr1 9078
BsaB1 5812	BseA1 8844	BspE1 8844	BspM2 8844	Dra3 5238
Eag1 8399	EclX1 8399	Eco52I 8399	EcoR1 1	Fdi2 1606
Fsp1 1606	Hind3 9292	Kpn1 4701	Ksp1 8399	Mam1 5812
Mlu1 7220	Mro1 8844	Mst1 1606	Nde1 7978	PspA1 9332
Sac2 8399	Sal1 524	Sfi1 2615	SgrA1 284	Sma1 9334
SnaB1 8084	Spe1 7743	Sst2 8399	Swa1 5524	Xba1 9317
Xma1 9332	Xma3 8399			

ENZYMES THAT CUT TWICE:

ApaL1 805 2051	Asp700 1983 2924	Bpu1102I 5691 8565	BseP1 9207 9209	Bsg1 4028 5682
BspC1 1754 9312	BspM1 5278 9294	BssH2 9207 9209	Bst1107I 3671 6625	Cel2 5691 8565
Drd1 599 4441	Esp1 5691 8565	Esp3I 4491 8333	Mun1 105 3070	Nae1 277 3235
Ngom1 275 3233	Pvu1 1754 9312	Sno1 805 2051	Xmn1 1983 2924	

ENZYMES THAT DO NOT CUT:

Aoc1, *Asc1*, *AspA1*, *Asu2*, *BbrP1*, *Bgl2*, *BsiW1*, *BstB1*, *BstE2*, *Bsu36I*, *Ccr1*, *Csp1*, *Csp45I*, *Cvn1*, *Eco81I*, *Mst2*, *Not1*, *Nru1*, *Nsp5*, *Pac1*, *Paer7I*, *PmaC1*, *Pme1*, *Pml1*, *Rsr2*, *Sau1*, *Sce1*, *Sfu1*, *Spl1*, *Srf1*, *Sse8387I*, *Xho1*

pMJ601	7167	Mammalian: BrDU ^r (TK ⁻)/β-gal screening; <i>E. coli</i> : Amp ^r	Polylinker (see Table 16.17.1)	Bernard Moss, National Institute of Allergy and Infectious Disease
---------------	------	----------------------------------------------------------------------------------------------------	--------------------------------	--------------------------------------------------------------------

COMMENTS: Used for constructing recombinant vaccinia virus by homologous recombination; strong synthetic late promoter controls expression of cloned gene.

ENZYMES THAT CUT ONCE:

AflIII 866	AgeI 4980	ApaI 846	BamHI 848	Bpu1102I 3942
BsaI 5917	BsaAI 6960	BsbI 6399	BseRI 4780	BspEI 854
BspLU11I 3688	BssHII 2427	Bst1107I 3694	Bsu36I 1153	DraIII 2119
EagI 4622	Eam1105I 5856	Eco47III 2764	EcoRI 3933	EcoRV 2042
HindIII 818	KpnI 828	MunI 229	NarI 861	NdeI 3886
NheI 836	NspI 3692	NspV 6845	Psp5II 279	PstI 6101
SacI 2869	SacII 833	SalI 878	ScaI 6336	SmaI 874
SunI 3700	XbaI 138	XhoI 923	XmnI 6455	

ENZYMES THAT CUT TWICE:

AatII 1552 6778	AccI 879 3693	BaeI 1166 4440	BclI 2275 4597	Bpu10I 1434 7040
BspMI 1703 3575	BsrDI 5917 6091	ClaI 737 1753	DrdI 3038 5071	DrdII 2098 4406
EarI 1030 6651	EcoO109I 279 6832	FspI 1071 6078	HgiEII 1117 5549	HpaI 1355 1979
Pfl1108I 4700 5874	Psp1406I 6082 6455	RleAI 3775 4924	XcmI 3146 3891	

ENZYMES THAT DO NOT CUT:

ApaBI, *AscI*, *AvrII*, *BglII*, *BsmI*, *BspGI*, *BsrGI*, *BstEII*, *EcoNI*, *FseI*, *MscI*, *NcoI*, *NgomI*, *NotI*, *NruI*, *NsiI*, *PacI*, *PmeI*, *PmlI*, *PshAI*, *RsrII*, *SapI*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SnaBI*, *SpeI*, *SphI*, *SrfI*, *Sse8387I*, *StuI*, *StyI*, *SwaI*, *Tth1111*

continued

Table A.5.6 Vectors for Expression of Genes from Heterologous Promoters in Mammalian Cells, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pSC11 (Fig. 16.17.2)	7884	Mammalian: BrDU ^r (TK ⁻)/β-gal screening; <i>E. coli</i> : Amp ^r	<i>Sma</i> I	Bernard Moss, National Institute of Allergy and Infectious Disease

COMMENTS: Used for constructing recombinant vaccinia virus by homologous recombination. **NOTE:** Two restriction sites exist that are not indicated by sequence data: *EcoRV* at ~2700 and *Nru*I at ~2570.

ENZYMES THAT CUT ONCE:

Acc I 3355	Apa BI 7659	Bpu 10I 5613	Bpu 1102I 3105	Bsa I 930
Bsa AI 7723	Bsb I 7759	Bse RI 7275	Bss HII 4619	Bst 1107I 3356
Bsu 36I 5894	Dra III 4934	Eag I 7117	Eam 1105I 996	Eco 47III 4286
Hind III 2243	Mun I 2473	Nar I 7602	Nru I ~2570	Psp 5II 2523
Pst I 3032	Sac I 4185	Sap I 2001	Sca I 515	Sma I 6503
Sun I 3346	Th 111I 7729	Xho I 6117	Xmn I 396	

ENZYMES THAT CUT TWICE:

Aat II 77 5502	Bae I 5884 6935	Ban II 3136 4185	Bsp LU1II 1884 3358	Bsr DI 762 936
Cl aI 2981 5295	Drd II 4952 6901	Eco O109I 16 2523	Eco RV ~2700 5008	Hpa I 5071 5695
Pf I 4513 4879	Psp 1406I 394 767	Xba I 2382 6228	Xcm I 3160 3905	

ENZYMES THAT DO NOT CUT:

*Afl*II, *Age*I, *Apa*I, *Asc*I, *Avr*II, *Bgl*II, *Bsm*I, *Bsp*EI, *Bsp*GI, *Bsr*GI, *Bst*EII, *Eco*NI, *Fse*I, *Kpn*I, *Msc*I, *Nco*I, *Ngo*AIV, *Nhe*I, *Not*I, *Nsi*I, *Nsp*V, *Pac*I, *Pme*I, *Pml*I, *Psh*AI, *Rsr*II, *Sac*II, *Sal*I, *Sex*AI, *Sfi*I, *Sgf*I, *Sgr*AI, *Sna*BI, *Spe*I, *Sph*I, *Srf*I, *Sse*8387I, *Stu*I, *Sty*I, *Swa*I

pSFV1 (Fig. 16.20.1)	11033	<i>E. coli</i> : Amp ^r	<i>Bam</i> HI, <i>Sma</i> I, <i>Xma</i> I	GIBCO/BRL
--------------------------------	-------	-----------------------------------	-------------------------------------------	-----------

COMMENTS: Used to express protein-coding sequences in many kinds of animal cells; pSFV3 is identical to pSFV1 with the addition of a ribosome binding site and Met before the cloning polylinker (see Fig. 16.20.1).

ENZYMES THAT CUT ONCE:

Age I 5372	Bam HI 7400	Bc II 1106	Bgl II 6713	Bsu 36I 4917
Eco RV 278	Not I 6136	Pvu I 9705	Rsr II 2025	Sac I 1803
Sap I 8326	Sma I 7406	Spe I 8261	Sph I 10601	Sun I 513
Xba I 6638	Xho I 5304	Xmn I 9934		

ENZYMES THAT CUT TWICE:

Bsp LU1II 7815 8442	Bst 1107I 536 2867	Bst EII 1872 2312	Eco 47III 6225 10667	Eco NI 7304 10536
Hgi EII 5041 9028	Msc I 1506 7296	Pf I 682 6359	Psh AI 1134 1650	Sac II 2372 5055
Sex AI 1876 3353	Sna BI 944 1769	Stu I 1465 7605	Xcm I 2306 7298	

ENZYMES THAT DO NOT CUT:

*Apa*I, *Asc*I, *Avr*II, *Bpu*10I, *Bpu*1102I, *Bsa*BI, *Bsm*I, *Bss*HII, *Cl*aI, *Fse*I, *Nde*I, *Nru*I, *Nsi*I, *Nsp*V, *Pac*I, *Pme*I, *Sfi*I, *Sgf*I, *Srf*I, *Sse*8387I, *Swa*I

continued

Table A.5.6 Vectors for Expression of Genes from Heterologous Promoters in Mammalian Cells, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pSFV-Helper 2 (Fig. 16.20.1)	8194	<i>E. coli</i> : Amp ^r	Not applicable	GIBCO/BRL
<i>COMMENTS:</i> Codes for helper RNA which, when cotransfected with recombinant RNAs transcribed from pSFV1 or pSFV3, supports assembly of recombinant RNAs into virus particles in vivo.				
<i>ENZYMES THAT CUT ONCE:</i>				
AatII 7418	AccI 309	AflII 1022	AgeI 2730	ApaBI 3510
BclII 4122	BglII 622	BsaBI 3466	BseRI 4689	BsmI 2562
BssHIII 3136	BstXI 1416	Eco47III 7828	FseI 1992	MscI 1205
NsiI 2837	NspV 1691	PshAI 3934	RsrII 2691	SalI 308
SapI 5487	SpeI 5422	SphI 7762	Sse8387I 2663	StuI 1518
XbaI 547	XhoI 2245	XmnI 7095		
<i>ENZYMES THAT CUT TWICE:</i>				
BspEI 1119 2590	BspLU11I 4976 5603	BstEII 2073 4827	ClaI 3466 4101	Eam1105I 368 6496
EcoRI 1298 2653	EcoRV 278 3464	FspI 5207 6718	HgiEII 4454 6189	HincII 310 907
HindIII 5281 5291	NcoI 1200 2607	NdeI 2837 4941	NgoAIV 1988 7919	Pfl1108I 2276 6514
PflMI 2767 4203	PmlI 1764 2984	Psp5II 1247 3231	PvuII 684 1058	SacII 3434 3583
SgrAI 1364 7910	XcmI 1207 3335			
<i>ENZYMES THAT DO NOT CUT:</i> <i>ApaI, AscI, AvrII, BaeI, BamHI, Bpu10I, Bpu1102I, Bst1107I, Bsu36I, HpaI, KpnI, MluI, NotI, NruI, PacI, PmeI, SacI, SexAI, SfiI, SgfI, SmaI, SnaBI, SrfI, SwaI</i>				
pSFV3-lacZ (Fig. 16.20.1)	14110	<i>E. coli</i> : Amp ^r	<i>SmaI, XmaI</i>	GIBCO/BRL
<i>COMMENTS:</i> Used as control for expression in pSFV1/3 system.				
<i>ENZYMES THAT CUT ONCE:</i>				
AgeI 5372	BglII 6713	Bpu10I 7922	Bpu1102I 10430	BsaBI 8748
BssHIII 8915	ClaI 8241	NdeI 10374	NotI 6136	RsrII 2025
SapI 11403	SmaI 10501	SpeI 11338	SphI 13678	XbaI 6638
XhoI 5304	XmnI 13011			
<i>ENZYMES THAT CUT TWICE:</i>				
BamHI 7423 10495	BclII 1106 8763	BstEII 1872 2312	Bsu36I 4917 7641	EcoNI 7304 13613
EcoRV 278 8530	MscI 1506 7296	PshAI 1134 1650	Psp1406I 12638 13011	SacI 1803 9357
SacII 2372 5055	SexAI 1876 3353	SnaBI 944 1769	StuI 1465 10682	SunI 513 10188
<i>ENZYMES THAT DO NOT CUT:</i> <i>ApaI, AscI, AvrII, BsmI, FseI, NruI, NsiI, NspV, PacI, PmeI, SfiI, SgfI, SrfI, Sse8387I, SwaI</i>				
pTF7-5 (Fig. 16.19.1)	4570	Mammalian: BrDU ^r (TK ⁻); <i>E. coli</i> : Amp ^r	<i>BamHI</i>	Bernard Moss, National Institute of Allergy and Infectious Disease

continued

Table A.5.6 Vectors for Expression of Genes from Heterologous Promoters in Mammalian Cells, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
<i>COMMENTS:</i> Used for transient cytoplasmic expression of cloned genes under control of T7 RNA polymerase promoter; can be used to construct recombinant vaccinia virus vector by homologous recombination.				
<i>ENZYMES THAT CUT ONCE:</i>				
<i>AatII</i> 77	<i>AflIII</i> 1884	<i>AlwNI</i> 1475	<i>ApaBI</i> 4345	<i>BaeI</i> 3617
<i>BamHI</i> 3046	<i>BbsI</i> 13	<i>BcgI</i> 492	<i>BclII</i> 3774	<i>BpmI</i> 927
<i>Bpu1102I</i> 3098	<i>BsaAI</i> 4409	<i>BsbI</i> 4445	<i>BseRI</i> 3957	<i>BsmBI</i> 4516
<i>BspLU11I</i> 1884	<i>BsrFI</i> 911	<i>BstXI</i> 4062	<i>DrdII</i> 3583	<i>EagI</i> 3799
<i>Eam1105I</i> 996	<i>HindIII</i> 2243	<i>MunI</i> 2472	<i>NarI</i> 4288	<i>NdeI</i> 4340
<i>Psp5II</i> 2522	<i>RleAI</i> 4101	<i>SapI</i> 2001	<i>ScaI</i> 515	<i>StyI</i> 3120
<i>Tth111I</i> 4415	<i>XbaI</i> 2381	<i>XmnI</i> 396		
<i>ENZYMES THAT CUT TWICE:</i>				
<i>ApoI</i> 2481 3492	<i>BcgI</i> 458 492	<i>BglII</i> 878 4278	<i>BsaI</i> 930 3014	<i>BsrDI</i> 762 936
<i>DrdI</i> 1782 4470	<i>DsaI</i> 2431 4020	<i>Eco57I</i> 330 1342	<i>FspI</i> 773 4268	<i>HgiEII</i> 1302 4340
<i>MmeI</i> 1491 1675	<i>NspI</i> 1888 4528	<i>Pfl1108I</i> 977 3877	<i>Psp1406I</i> 394 767	<i>PvuI</i> 627 4249
<i>PvuII</i> 2064 4218				
<i>ENZYMES THAT DO NOT CUT:</i>				
<i>AccI, AflII, AgeI, ApaI, AscI, AvaI, AvrII, BanII, BglIII, Bpu10I, BsgI, BsmI, BspEI, BspGI, BspMI, BsrGI, BssHII, Bst1107I, BstEII, Bsu36I, ClaI, DraIII, Eco47III, EcoNI, EcoRI, EcoRV, FseI, HincII, HpaI, KpnI, MluI, MscI, NcoI, NgoAIV, NheI, NotI, NruI, NsiI, NspV, PacI, PflMI, PmeI, PmlI, PshAI, PstI, RsrII, SacI, SacII, SalI, SexAI, SfiI, SgfI, SgrAI, SmaI, SnaBI, SpeI, SphI, SrfI, Sse8387I, StuI, SunI, SwaI, XcmI, XhoI</i>				
pTKgptF1s	7041	Mammalian: BrDU ^r (TK ⁻) or MPA ^r (XGPRT); <i>E. coli</i> : Amp ^r	<i>EcoRI, PstI, SalI, SbaI, BamHI</i>	Bernard Moss, National Institute of Allergy and Infectious Disease

continued

Table A.5.6 Vectors for Expression of Genes from Heterologous Promoters in Mammalian Cells, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
<i>COMMENTS:</i> Used for constructing recombinant vaccinia virus by homologous recombination; vaccinia virus promoter p11 and ATG control expression of protein-coding segments; pTKgptF2s and pTKgptF3s contain one or two additional G residues, respectively, following ATG to allow cloning in all three reading frames.				
<i>ENZYMES THAT CUT ONCE:</i>				
<i>AccI</i> 4218	<i>AflIII</i> 4235	<i>AgeI</i> 5019	<i>AlwNI</i> 776	<i>ApaI</i> 5598
<i>ApaBI</i> 2478	<i>BamHI</i> 4229	<i>BanII</i> 5598	<i>BglII</i> 4664	<i>Bpu10I</i> 5308
<i>BsaI</i> 1314	<i>BsbI</i> 2373	<i>BsmBI</i> 2298	<i>BspGI</i> 5499	<i>BstXI</i> 2756
<i>ClaI</i> 6366	<i>DrdII</i> 3231	<i>EagI</i> 3011	<i>Eam1105I</i> 1253	<i>EcoRI</i> 4205
<i>EcoRV</i> 5126	<i>HpaI</i> 4242	<i>MunI</i> 6304	<i>NarI</i> 2528	<i>PflMI</i> 5753
<i>PshAI</i> 5359	<i>Psp5II</i> 6823	<i>SalI</i> 4217	<i>SapI</i> 244	<i>ScaI</i> 1733
<i>SexAI</i> 5008	<i>Sse8387I</i> 4215	<i>StyI</i> 5901	<i>Tth111I</i> 2402	
<i>ENZYMES THAT CUT TWICE:</i>				
<i>AflIII</i> 360 4771	<i>BaeI</i> 3197 5030	<i>BbsI</i> 2231 5901	<i>BcgI</i> 1758 3650	<i>BclI</i> 3036 4387
<i>BglII</i> 1373 2543	<i>BpmI</i> 1323 5299	<i>BseRI</i> 2859 5859	<i>BsgI</i> 5416 5985	<i>BsmI</i> 5788 6305
<i>BspLU11I</i> 360 4771	<i>BspMI</i> 3634 4204	<i>BsrDI</i> 1314 1488	<i>DrdI</i> 468 2350	<i>HgiEII</i> 946 2478
<i>HincII</i> 4219 4242	<i>KpnI</i> 4256 4866	<i>NdeI</i> 2476 4358	<i>NsiI</i> 3673 3946	<i>Psp1406I</i> 1479 1852
<i>PstI</i> 3645 4215	<i>PvuI</i> 1623 2571	<i>PvuII</i> 184 2600		
<i>ENZYMES THAT DO NOT CUT:</i>				
<i>AscI, Aval, AvrII, Bpu1102I, BspEI, BsrGI, BssHII, Bst1107I, BstEII, Bsu36I, DraIII, Eco47III, EcoNI, FseI, MluI, MscI, NcoI, NgoAIV, NheI, NotI, NruI, NspV, PacI, PmeI, PmlI, RsrII, SacI, SacII, SfiI, SgfI, SgrAI, SmaI, SnaBI, SpeI, SphI, SrfI, StuI, SunI, SwaI, XcmI, XhoI</i>				
pTM1 (Fig. 16.19.2)	5357	Mammalian: BrDU ^r (TK ⁻); <i>E. coli</i> : Amp ^r	Polylinker (see Fig. 16.19.2)	Bernard Moss, National Institute of Allergy and Infectious Disease
<i>COMMENTS:</i> Used for transient cytoplasmic expression of cloned genes under control of T7 RNA polymerase promoter; contains encephalomyocarditis virus untranslated leader region downstream of P _{T7} ; TK ⁻ selection can be used to isolate recombinant virus.				
<i>ENZYMES THAT CUT ONCE:</i>				
<i>AccI</i> 1488	<i>AlwNI</i> 4590	<i>ApaI</i> 968	<i>AvrII</i> 1002	<i>BaeI</i> 2063
<i>BamHI</i> 1453	<i>BclI</i> 2220	<i>BpmI</i> 4042	<i>Bpu1102I</i> 1545	<i>BsbI</i> 3058
<i>BspEI</i> 1622	<i>BspMI</i> 1190	<i>BstXI</i> 2508	<i>ClaI</i> 738	<i>EagI</i> 2245
<i>Eam1105I</i> 4111	<i>EcoRI</i> 1429	<i>HgiEII</i> 4417	<i>HincII</i> 1489	<i>KpnI</i> 1296
<i>MunI</i> 230	<i>NcoI</i> 1423	<i>NgoAIV</i> 2860	<i>PacI</i> 1483	<i>PflMI</i> 1304
<i>PmlI</i> 1167	<i>Psp5II</i> 280	<i>PstI</i> 1463	<i>RleAI</i> 2547	<i>SacI</i> 1445
<i>SalI</i> 1487	<i>SapI</i> 5116	<i>ScaI</i> 3630	<i>SmaI</i> 1437	<i>SpeI</i> 1447
<i>StuI</i> 1473	<i>XbaI</i> 139	<i>XhoI</i> 1465		
<i>ENZYMES THAT CUT TWICE:</i>				
<i>AvaI</i> 1435 1465	<i>BbsI</i> 977 1091	<i>BcgI</i> 1481 3607	<i>BsaI</i> 807 4045	<i>BsaAI</i> 1167 2965
<i>BsaHI</i> 2471 3571	<i>BseRI</i> 1241 2403	<i>BsiI</i> 3442 4826	<i>BsmI</i> 997 1036	<i>Bsp24I</i> 4305 4483
<i>BspLU11I</i> 1341 4999	<i>DraIII</i> 1214 2968	<i>DrdI</i> 3012 4897	<i>FspI</i> 2710 3888	<i>HindIII</i> 1 1075
<i>Pfl1108I</i> 2323 4092	<i>PvuI</i> 2691 3742	<i>PvuII</i> 2660 5179	<i>XmnI</i> 1063 3511	
<i>ENZYMES THAT DO NOT CUT:</i>				
<i>AatII, AflII, AgeI, ApaBI, AscI, BglII, Bpu10I, BsgI, BsmBI, BspGI, BsrGI, BssHII, Bst1107I, BstEII, Bsu36I, Eco47III, EcoNI, EcoRV, FseI, HpaI, MluI, MscI, NarI, NdeI, NheI, NotI, NruI, NsiI, NspV, PmeI, PshAI, RsrII, SacII, SexAI, SfiI, SgfI, SgrAI, SnaBI, SphI, SrfI, Sse8387I, SunI, SwaI, Tth111I, XcmI</i>				
pTM3 (Fig. 16.19.2)	7444	Mammalian: BrDU ^r (TK ⁻)/MPA ^r (gpt ⁺); <i>E. coli</i> : Amp ^r	Polylinker (see Fig. 16.19.2)	Bernard Moss, National Institute of Allergy and Infectious Disease

continued

Table A.5.6 Vectors for Expression of Genes from Heterologous Promoters in Mammalian Cells, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
<i>COMMENTS:</i> Used for transient cytoplasmic expression of cloned genes under control of T7 RNA polymerase promoter; contains encephalomyocarditis virus untranslated leader region downstream of P _{T7} ; includes guanine phosphoribosyltransferase (<i>gpt</i>) gene to permit mycophenolic acid selection for recombinant virus in addition to TK ⁻ selection.				
<i>ENZYMES THAT CUT ONCE:</i>				
<i>AatII</i> 2868	<i>AccI</i> 1488	<i>AgeI</i> 2389	<i>AlwNI</i> 6677	<i>AvrII</i> 1002
<i>BamHI</i> 1453	<i>BglII</i> 2034	<i>Bpu10I</i> 2678	<i>Bpu1102I</i> 1545	<i>BsbI</i> 5145
<i>BspGI</i> 2869	<i>BspMI</i> 1190	<i>BstXI</i> 4595	<i>EagI</i> 4332	<i>Eam1105I</i> 6198
<i>EcoRI</i> 1429	<i>EcoRV</i> 2496	<i>HgiEII</i> 6504	<i>HpaI</i> 1630	<i>NcoI</i> 1423
<i>NdeI</i> 1727	<i>NgoAIV</i> 4947	<i>PacI</i> 1483	<i>PmII</i> 1167	<i>PshAI</i> 2729
<i>Psp5II</i> 280	<i>PstI</i> 1463	<i>SacI</i> 1445	<i>SalI</i> 1487	<i>SapI</i> 7203
<i>ScaI</i> 5717	<i>SexAI</i> 2378	<i>SmaI</i> 1437	<i>SpeI</i> 1447	<i>SphI</i> 3707
<i>StuI</i> 1473	<i>XbaI</i> 139	<i>XhoI</i> 1465		
<i>ENZYMES THAT CUT TWICE:</i>				
<i>ApaI</i> 968 2968	<i>AvaI</i> 1435 1465	<i>BaeI</i> 2400 4150	<i>BcgI</i> 1481 5694	<i>BclI</i> 1756 4307
<i>BpmI</i> 2669 6129	<i>BsaI</i> 807 6132	<i>BsgI</i> 2786 3355	<i>BsiI</i> 5529 6913	<i>Bsp24I</i> 6392 6570
<i>BspEI</i> 1622 3709	<i>ClaI</i> 738 1904	<i>DraIII</i> 1214 5055	<i>DrdI</i> 5099 6984	<i>FspI</i> 4797 5975
<i>HincII</i> 1489 1630	<i>KpnI</i> 1296 2236	<i>MunI</i> 230 3674	<i>PflMI</i> 1304 3123	<i>PvuI</i> 4778 5829
<i>PvuII</i> 4747 7266	<i>XmnI</i> 1063 5598			
<i>ENZYMES THAT DO NOT CUT:</i>				
<i>AflII, ApaBI, AscI, BsmBI, BsrGI, BssHII, Bst1107I, BstEII, Bsu36I, Eco47III, EcoNI, FseI, MluI, MscI, NarI, NheI, NotI, NruI, NsiI, NspV, PmeI, RsrII, SacII, SfiI, SgfI, SgrAI, SnaBI, SrfI, Sse8387I, SunI, SwaI, Tth111I, XcmI</i>				

Table A.5.7 Vectors Using Reporter Genes for Genetic Regulation Studies in Mammalian Cells

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pβgal-Basic (Fig. 9.7.5)	7482	<i>E. coli</i> : Amp ^r	Polylinker (see Fig. 9.7.5)	Clontech (U13184)

COMMENTS: Used to clone and characterize eukaryotic promoter and enhancer sequences using *lacZ* as reporter gene; contains *f1* origin of replication for single-stranded DNA production.

ENZYMES THAT CUT ONCE:

AatII 800	BaeI 414	BamHI 4623	BclI 1523	BglII 37
Bpu10I 682	Bpu1102I 3190	BsaI 5833	BspGI 156	BssHIII 1675
Bst1107I 2942	Bsu36I 401	Eam1105I 5772	EcoNI 3578	EcoRV 1290
HindIII 47	KpnI 12	NdeI 3134	NheI 28	PshAI 4694
SalI 4629	SapI 4763	ScaI 6252	SgrAI 3389	SmaI 3
StyI 4076	SunI 2948	XhoI 33	XmnI 6371	

ENZYMES THAT CUT TWICE:

AccI 2941 4630	BmgI 2301 3710	BsaAI 6921 7254	BsaBI 1508 4391	BsbI 3340 7014
BseRI 3516 4034	BsmBI 804 1744	BspLU11I 2936 4879	BsrDI 5833 6007	BstXI 2397 3014
Clal 1001 3326	DraIII 1367 6924	DrdII 1346 6916	Eco47III 2012 4755	FspI 319 5994
HgiEII 365 5465	MunI 4479 7322	NgoAIV 4747 6816	NspI 2940 4883	Pfl1108I 4736 5790
RcaI 5599 6607	SacI 18 2117	VspI 4295 5944	XcmI 2394 3139	

ENZYMES THAT DO NOT CUT:

AflII, *AgeI*, *ApaI*, *ApaBI*, *AscI*, *AvrII*, *BspEI*, *BsrGI*, *BstEII*, *EagI*, *EcoO109I*, *EcoRI*, *FseI*, *MscI*, *NarI*, *NcoI*, *NotI*, *NruI*, *NsiI*, *NspV*, *PacI*, *PmeI*, *PmlI*, *Psp5II*, *PstI*, *RsrII*, *SacII*, *SexAI*, *SfiI*, *SgfI*, *SnaBI*, *SpeI*, *SphI*, *SrfI*, *Sse8387I*, *StuI*, *SwaI*, *Tth111I*, *XbaI*

pφGH (Fig. 9.7.3)	4817	<i>E. coli</i> : Amp ^r	<i>HindIII</i> , <i>SalI</i> , <i>HincII</i> , <i>SbaI</i>	Nichols Institute Diagnostics
--------------------------	------	-----------------------------------	---------------------------------------------------------------	-------------------------------

COMMENTS: Used for cloning promoter of interest and monitoring function in transient expression assays; reporter gene is human growth hormone (hGH) gene.

ENZYMES THAT CUT ONCE:

AflIII 2512	AgeI 2083	ApaI 1701	ApaBI 4578	ApoI 2156
AvrII 697	BamHI 1	BglII 1362	Bpu1102I 1865	BsbI 4512
BspGI 2050	BspLU11I 2512	DraIII 103	DsaI 2073	Eco47III 1296
EcoRI 2156	HincII 4808	HindIII 4791	NcoI 2073	NdeI 4576
NsiI 1856	Pfl1108I 3423	SacI 582	SalI 4806	SapI 2396
ScaI 3885	SphI 1854	StuI 1226	Tth111I 1466	XbaI 4812
XmnI 4004				

ENZYMES THAT CUT TWICE:

AccI 673 4807	BbsI 979 1717	BcgI 1798 3910	Bpu10I 581 1821	BseRI 578 581
BspMI 2077 4793	BsrFI 2083 3485	MscI 207 1913	PflMI 203 2086	Psp1406I 3631 4004
PvuI 3775 4671	SmaI 878 1534	SspI 1663 4209		

ENZYMES THAT DO NOT CUT:

AflII, *AscI*, *BaeI*, *BclI*, *BsaAI*, *BsaBI*, *BsmI*, *BspEI*, *BsrGI*, *BssHIII*, *Bst1107I*, *BstEII*, *Clal*, *EagI*, *EcoRV*, *FseI*, *HpaI*, *KpnI*, *MluI*, *MunI*, *NgoAIV*, *NheI*, *NotI*, *NruI*, *NspV*, *PacI*, *PmeI*, *PmlI*, *PshAI*, *RleAI*, *RsrII*, *SacII*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SnaBI*, *SpeI*, *SrfI*, *Sse8387I*, *SunI*, *SwaI*, *XcmI*, *XhoI*

continued

Vectors

A.5.31

Table A.5.7 Vectors Using Reporter Genes for Genetic Regulation Studies in Mammalian Cells, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pOLUC (Fig. 9.7.4)	4708	<i>E. coli</i> : Amp ^r	<i>SalI</i> , <i>BamHI</i> , <i>SmaI</i> , <i>HindIII</i>	Allan R. Brasier, University of Texas Medical Branch—Galveston

COMMENTS: Used for cloning promoter of interest and monitoring function in transient expression assays; reporter gene is luciferase; pSV2LUC (see entry this table) is used as a positive control.

ENZYMES THAT CUT ONCE:

AatII 3936	AccI 30	AlwNI 2537	BamHI 41	BglII 3134
BsaAI 252	BseRI 1653	BsmI 77	BspLU11I 2121	BstEII 705
Bsu36I 711	ClaI 1463	DsaI 4306	Eam1105I 3014	Eco47III 4342
FspI 3236	HindIII 71	NgoAIV 4433	PacI 1422	Pfl1108I 3032
Psp5II 1277	PvuI 3384	PvuII 1945	RleAI 1821	SacI 1897
SalI 29	ScaI 3494	SmaI 49	Sse8387I 27	SunI 252
XcmI 833	XmnI 3613			

ENZYMES THAT CUT TWICE:

AflIII 591 2121	AvaI 47 1154	BsaI 1912 3075	BsmBI 4004 4046	BsmFI 1388 4282
BspEI 793 1309	BspMI 16 1496	BsrBI 2054 3855	BsrDI 3075 3249	BsrGI 588 4628
DrdI 2229 4098	DrdII 203 4489	EcoNI 1715 4215	EcoRV 67 1436	HgiEII 1176 2707
HincII 31 1402	MunI 213 669	NheI 9 4466	NspV 267 1055	Psp1406I 3240 3613
PstI 27 57	SapI 900 2005	SgrAI 1526 4424	SphI 21 761	SspI 1795 3818
XbaI 35 145				

ENZYMES THAT DO NOT CUT:

AflII, *AgeI*, *ApaI*, *ApaBI*, *AscI*, *AvrII*, *BaeI*, *BclI*, *BglII*, *Bpu10I*, *Bpu1102I*, *BsaBI*, *BsgI*, *BspGI*, *BssHII*, *Bst1107I*, *BstXI*, *DraIII*, *EagI*, *FseI*, *HpaI*, *KpnI*, *MluI*, *MscI*, *NcoI*, *NdeI*, *NotI*, *NruI*, *NsiI*, *PflMI*, *PmeI*, *PmlI*, *PshAI*, *RsrII*, *SacII*, *SexAI*, *SfiI*, *SgfI*, *SnaBI*, *SpeI*, *SrfI*, *StuI*, *StyI*, *SwaI*, *Tth1111I*, *XhoI*

continued

Table A.5.7 Vectors Using Reporter Genes for Genetic Regulation Studies in Mammalian Cells, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pSV2LUC (Fig. 9.7.4)	5009	<i>E. coli</i> : Amp ^r	See Fig. 9.7.4	Allan R. Brasier, University of Texas Medical Branch—Galveston

COMMENTS: Contains SV40 promoter/enhancer sequences inserted into pOLUC for use as a positive control for luciferase expression in transient assay.

ENZYMES THAT CUT ONCE:

AatII 4237	AccI 30	AlwNI 2838	AvaI 1455	AvrII 356
BamHI 41	BsaAI 553	BsmI 378	BspLU11I 2422	BstEII 1006
Bsu36I 1012	ClaI 1764	Eam1105I 3315	Eco47III 4643	EcoRV 1737
FspI 3537	HindIII 372	NcoI 263	NgoAIV 4734	PacI 1723
Pfl1108I 3333	Psp5II 1578	PstI 27	PvuI 3685	PvuII 2246
RleAI 2122	SacI 2198	SalI 29	ScaI 3795	SfiI 309
Sse8387I 27	StuI 355	SunI 553	XcmI 1134	XmnI 3914

ENZYMES THAT CUT TWICE:

AflIII 892 2422	ApaBI 119 174	BglI 309 3435	BsaI 2213 3376	BseRI 352 1954
BsmBI 4305 4347	BspEI 1094 1610	BspMI 16 1797	BsrBI 2355 4156	BsrDI 3376 3550
BsrGI 889 4929	DrdI 2530 4399	DrdII 504 4790	DsaI 263 4607	EcoNI 2016 4516
HgiEII 1477 3008	HincII 31 1703	MunI 514 970	NheI 9 4767	NsiI 123 178
NspV 568 1356	Psp1406I 3541 3914	SapI 1201 2306	SgrAI 1827 4725	SspI 2096 4119
StyI 263 356	XbaI 35 446			

ENZYMES THAT DO NOT CUT:

AflII, *AgeI*, *ApaI*, *AscI*, *BaeI*, *BclI*, *BglII*, *Bpu10I*, *Bpu1102I*, *BsaBI*, *BsgI*, *BspGI*, *BssHII*, *Bst1107I*, *BstXI*, *DraIII*, *EagI*, *FseI*, *HpaI*, *KpnI*, *MluI*, *MscI*, *NdeI*, *NotI*, *NruI*, *PflMI*, *PmeI*, *PmlI*, *PshAI*, *RsrII*, *SacII*, *SexAI*, *SgfI*, *SmaI*, *SnaBI*, *SpeI*, *SrfI*, *Swal*, *Tth111I*, *XhoI*

pXGH5	6670	<i>E. coli</i> : Amp ^r	See Fig. 9.7.3	Nichols Institute Diagnostics
--------------	------	-----------------------------------	----------------	-------------------------------

COMMENTS: Well-characterized hGH fusion gene to characterize or optimize transfection efficiency (see pϕGH in this table); see Fig. 9.7.3 for restriction site information.

continued

Table A.5.7 Vectors Using Reporter Genes for Genetic Regulation Studies in Mammalian Cells, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
p300-CAT (Fig. 9.7.1)	4847	<i>E. coli</i> : Amp ^r	<i>HindIII</i> , <i>SphI</i> , <i>PstI</i> , <i>SalI</i> , <i>XbaI</i> , <i>BamHI</i> , <i>SmaI</i> , <i>KpnI</i> , <i>SacI</i>	Michael Gilman, Ariad Pharmaceuticals

COMMENTS: Used for cloning promoter of interest and monitoring function in transient expression assays; reporter gene is chloramphenicol acetyl transferase; backbone of p300-CAT is pBS- (Stratagene), which allows production of single-stranded DNA in *E. coli*. p300+CAT was constructed using pBS+.

ENZYMES THAT CUT ONCE:

AatII 2649	AccI 20	AflIII 4456	AlwNI 4047	ApaBI 2399
AvaI 36	BamHI 31	BbsI 1144	BcgI 3030	Bpu10I 82
BsaI 3502	BsaAI 2111	BsaBI 1459	BseRI 1102	BsgI 1228
BspEI 305	BspLU11I 4456	BspMI 6	DraIII 2111	DrdII 2116
DsaI 610	Eam1105I 3568	EcoO109I 2588	EcoRI 309	FspI 3345
HindIII 1	HpaI 1560	KpnI 44	MscI 576	MunI 1547
NcoI 610	NdeI 2394	NgoAIV 2212	Pfl1108I 3549	PstI 17
RleAI 1526	SacI 50	SalI 19	SapI 4573	SmaI 38
SphI 11	Sse8387I 17	XbaI 25	XmnI 2968	

ENZYMES THAT CUT TWICE:

AvaII 3204 3426	BanII 50 2186	BcgI 3030 3064	BglII 1874 3450	BpmI 431 3499
BsaHI 2646 3028	BsbI 2018 2460	BsrFI 2212 3483	HgiEII 2394 3874	HincII 21 1560
PflMI 541 996	PvuI 1845 3199	ScaI 726 3087	StyI 610 1144	

ENZYMES THAT DO NOT CUT:

AflII, *AgeI*, *ApaI*, *AscI*, *AvrII*, *BaeI*, *BclI*, *BglII*, *Bpu1102I*, *BsmFI*, *BspGI*, *BsrGI*, *BssHII*, *Bst1107I*, *BstEII*, *BstXI*, *Bsu36I*, *ClaI*, *EagI*, *Eco47III*, *EcoNI*, *EcoRV*, *FseI*, *MluI*, *NarI*, *NheI*, *NotI*, *NruI*, *NsiI*, *NspV*, *PacI*, *PmeI*, *PmlI*, *PshAI*, *Psp5II*, *RsrII*, *SacII*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SnaBI*, *SpeI*, *SrfI*, *StuI*, *SunI*, *Swal*, *Tth111I*, *XcmI*, *XhoI*

Table A.5.8 Vectors for Expression of Fusion Proteins in *E. coli*

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
λgt11 See Table A.5.3 and Fig. 1.10.5 for full listings.				
λZAP and λZAPII See Table A.5.3 and Fig. 1.10.8 for full listings.				
pALtrxA-781 (Fig. 16.8.1)	3551	Amp ^r	<i>Sfi</i> I, <i>Xba</i> I, <i>Sal</i> I, <i>Pst</i> I, <i>Rsr</i> II	Invitrogen; John McCoy, Genetics Institute

COMMENTS: Used for expression of protein fusions to *E. coli* thioredoxin. Peptide fusions (for antigen production) can be made in active-site loop of thioredoxin using *Rsr*II site.

ENZYMES THAT CUT ONCE:

Aat II 69	Acc I 3068	Afl III 1876	Alw NI 1467	Apo I 2100
Ava I 2578	Bpm I 919	Bsa I 922	Bsa BI 2126	Bsm I 2259
Bsp EI 2322	Bsp LU11I 1876	Bsp MI 3080	Bsr FI 903	Bst XI 2663
Cla I 2905	Dsa I 2942	Eam 1105I 988	Eco O109I 8	Hind III 3146
Msc I 3057	Mun I 2696	Nar I 3311	Nde I 2722	Nsi I 2675
Pfl 1108I 969	Pst I 3077	Pvu II 3241	Rsr II 2823	Sal I 3067
Sap I 1993	Sca I 507	Sfi I 3054	Ssp I 183	Xba I 3061

ENZYMES THAT CUT TWICE:

Ban I 1035 3310	Bsb I 2158 3431	Bsm BI 3502 3544	Bsr DI 754 928	Drd I 1774 3456
Eco NI 2249 2636	Fsp I 765 3291	Hinc II 2160 3069	Mme I 1483 1667	Nsp I 1880 3514
Psp 1406I 386 759	Pvu I 619 3272	Vsp I 813 2048	Xmn I 388 3027	

ENZYMES THAT DO NOT CUT:

*Afl*II, *Age*I, *Apa*I, *Apa*BI, *Asc*I, *Avr*II, *Bae*I, *Bam*HI, *Ban*II, *Bbs*I, *Bcl*II, *Bgl*II, *Bpu*10I, *Bpu*1102I, *Bsa*AI, *Bse*RI, *Bsg*I, *Bsm*FI, *Bsp*GI, *Bsr*GI, *Bss*HII, *Bst*1107I, *Bst*EII, *Bsu*36I, *Dra*III, *Drd*II, *Eag*I, *Eco*47III, *Eco*RI, *Eco*RV, *Fse*I, *Hpa*I, *Kpn*I, *Mlu*I, *Nco*I, *Ngo*AIV, *Nhe*I, *Not*I, *Nru*I, *Nsp*V, *Pac*I, *Pfl*MI, *Pme*I, *Pml*I, *Psh*AI, *Psp*5II, *Rle*AI, *Sac*I, *Sac*II, *Sex*AI, *Sgf*I, *Sgr*AI, *Sma*I, *Sna*BI, *Spe*I, *Sph*I, *Srf*I, *Sse*8387I, *Stu*I, *Sty*I, *Sun*I, *Swa*I, *Tth*1111I, *Xcm*I, *Xho*I

continued

Table A.5.8 Vectors for Expression of Fusion Proteins in *E. coli*, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pET-15b (Fig. 10.11B.2)	5708	Amp ^r	<i>Nde</i> I, <i>Xho</i> I, <i>Bam</i> HI	Novagen

COMMENTS: Used for expression of recombinant proteins with “His-tag” amino-terminal leader. Header is used for purification and subsequently removed by thrombin cleavage.

ENZYMES THAT CUT ONCE:

<i>Aat</i> II 5635	<i>Acc</i> I 3590	<i>Alw</i> NI 4236	<i>Apa</i> I 1427	<i>Bam</i> HI 319
<i>Bcl</i> II 1230	<i>Bgl</i> II 494	<i>Bpu</i> 10I 2926	<i>Bpu</i> 1102I 267	<i>Bsa</i> I 4774
<i>Bsa</i> AI 3572	<i>Bsm</i> I 2704	<i>Bsp</i> LU11I 3820	<i>Bsp</i> MI 2399	<i>Bss</i> HII 1627
<i>Bst</i> 1107I 3591	<i>Bst</i> EII 1397	<i>Cla</i> I 24	<i>Drd</i> II 939	<i>Eag</i> I 2284
<i>Eam</i> 1105I 4713	<i>Eco</i> NI 751	<i>Eco</i> RI 5706	<i>Hind</i> III 29	<i>Hpa</i> I 1722
<i>Mlu</i> I 1216	<i>Msc</i> I 2791	<i>Nco</i> I 389	<i>Nde</i> I 331	<i>Nru</i> I 2319
<i>Psh</i> AI 2061	<i>Pst</i> I 4958	<i>Pvu</i> I 5083	<i>Sap</i> I 3704	<i>Sca</i> I 5193
<i>Sgr</i> AI 535	<i>Sph</i> I 691	<i>Ssp</i> I 5517	<i>Th</i> 111I 3565	<i>Xba</i> I 428
<i>Xho</i> I 324				

ENZYMES THAT CUT TWICE:

<i>Afl</i> III 1216 3820	<i>Apa</i> BI 900 2397	<i>Apo</i> I 1491 5706	<i>Ava</i> I 324 2770	<i>Bsb</i> I 3536 5256
<i>Bsm</i> BI 1831 3461	<i>Bsp</i> EI 189 3009	<i>Drd</i> I 3513 3928	<i>Eco</i> 57I 4368 5380	<i>Eco</i> RV 187 1666
<i>Hgi</i> EII 814 4406	<i>Hinc</i> II 1722 5254	<i>Mme</i> I 4035 4219	<i>Pfl</i> 1108I 2103 4731	<i>Psp</i> 5II 2784 2826
<i>Xmn</i> I 3378 5312				

ENZYMES THAT DO NOT CUT:

*Afl*II, *Age*I, *Asc*I, *Avr*II, *Bae*I, *Bse*RI, *Bsr*GI, *Bsu*36I, *Dra*III, *Fse*I, *Kpn*I, *Mun*I, *Nhe*I, *Not*I, *Nsi*I, *Nsp*V, *Pac*I, *Pme*I, *Pml*I, *Rle*AI, *Rsr*II, *Sac*I, *Sac*II, *Sal*I, *Sex*AI, *Sfi*I, *Sgf*I, *Sma*I, *Sna*BI, *Spe*I, *Srf*I, *Sse*8387I, *Stu*I, *Sun*I, *Swa*I

pGE374	6605	Amp ^r	<i>Sma</i> I	George Weinstock, University of Texas Medical School—Houston
---------------	------	------------------	--------------	--------------------------------------------------------------------

COMMENTS: Used to identify and express open reading frames.

ENZYMES THAT CUT ONCE:

<i>Apa</i> BI 4545	<i>Bae</i> I 588	<i>Bcl</i> II 1697	<i>Bpu</i> 1102I 3364	<i>Bsa</i> I 5674
<i>Bsa</i> AI 4474	<i>Bsp</i> EI 3909	<i>Bsp</i> GI 4248	<i>Bss</i> HII 1849	<i>Bsu</i> 36I 575
<i>Cla</i> I 1175	<i>Dra</i> III 1541	<i>Drd</i> II 1520	<i>Eam</i> 1105I 5613	<i>Eco</i> RI 1
<i>Eco</i> RV 1464	<i>Msc</i> I 3691	<i>Nco</i> I 327	<i>Nsi</i> I 106	<i>Pfl</i> 1108I 5631
<i>Pst</i> I 5858	<i>Rle</i> AI 3197	<i>Sac</i> I 2291	<i>Sac</i> II 34	<i>Sap</i> I 4604
<i>Sca</i> I 6093	<i>Sma</i> I 354	<i>Sph</i> I 108	<i>Sty</i> I 327	<i>Sun</i> I 3122
<i>Th</i> 111I 4467	<i>Vsp</i> I 5785			

ENZYMES THAT CUT TWICE:

<i>Aat</i> II 974 6535	<i>Acc</i> I 3115 4492	<i>Bam</i> HI 347 357	<i>Ban</i> II 2291 3340	<i>Bpu</i> 10I 856 3826
<i>Bsa</i> BI 1682 3917	<i>Bsb</i> I 4438 6156	<i>Bsp</i> LU11I 3110 4720	<i>Bsp</i> MI 1125 2997	<i>Bsr</i> DI 5674 5848
<i>Bst</i> 1107I 3116 4493	<i>Bst</i> XI 2571 3188	<i>Eco</i> 47III 2186 3974	<i>Hpa</i> I 777 1401	<i>Nde</i> I 3308 4543
<i>Pfl</i> MI 1596 1962	<i>Psp</i> 5II 3684 3726	<i>Ssp</i> I 1581 6417	<i>Xmn</i> I 4280 6212	

ENZYMES THAT DO NOT CUT:

*Afl*II, *Age*I, *Apa*I, *Asc*I, *Avr*II, *Bgl*II, *Bse*RI, *Bsm*I, *Bsr*GI, *Bst*EII, *Eag*I, *Eco*NI, *Fse*I, *Hind*III, *Kpn*I, *Mun*I, *Nar*I, *Ngo*AIV, *Nhe*I, *Not*I, *Nru*I, *Nsp*V, *Pac*I, *Pme*I, *Pml*I, *Psh*AI, *Rsr*II, *Sal*I, *Sex*AI, *Sfi*I, *Sgf*I, *Sgr*AI, *Sna*BI, *Spe*I, *Srf*I, *Sse*8387I, *Stu*I, *Swa*I, *Xba*I, *Xho*I

continued

Table A.5.8 Vectors for Expression of Fusion Proteins in *E. coli*, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pGEX1 (Fig. 16.7.1)	4932	Amp ^r	<i>Bam</i> HI, <i>Sma</i> I, <i>Eco</i> RI	Amrad

COMMENTS: Allows expression of fusion proteins to glutathione-*S*-transferase.

ENZYMES THAT CUT ONCE:

<i>Aat</i> II 1208	<i>Alw</i> NI 2606	<i>Apa</i> I 3842	<i>Apa</i> BI 3315	<i>Ava</i> I 919
<i>Bae</i> I 4737	<i>Bam</i> HI 914	<i>Ban</i> II 3842	<i>Bsa</i> I 2061	<i>Bsa</i> AI 1109
<i>Bsp</i> MI 62	<i>Bss</i> HII 4042	<i>Bst</i> EII 3812	<i>Bsu</i> 36I 4724	<i>Drd</i> II 3354
<i>Dsa</i> I 4832	<i>Eam</i> 1105I 2127	<i>Eco</i> NI 268	<i>Eco</i> RI 924	<i>Eco</i> RV 4081
<i>Mlu</i> I 3631	<i>Msc</i> I 465	<i>Nar</i> I 4271	<i>Nsp</i> V 655	<i>Pfl</i> 1108I 2108
<i>Pfl</i> MI 3213	<i>Pst</i> I 1885	<i>Rle</i> AI 832	<i>Sma</i> I 921	<i>Swa</i> I 685
<i>Tth</i> 111I 1102				

ENZYMES THAT CUT TWICE:

<i>Bcl</i> II 692 3645	<i>Bgl</i> II 2009 4652	<i>Bsb</i> I 1073 1583	<i>Bsm</i> BI 998 4246	<i>Bsm</i> FI 245 4729
<i>Bsp</i> LU11I 496 3015	<i>Bsr</i> FI 2042 3317	<i>Drd</i> I 1050 2913	<i>Eco</i> 57I 1461 2473	<i>Eco</i> O109I 290 1147
<i>Fsp</i> I 1904 4642	<i>Pvu</i> I 1758 4623	<i>Sap</i> I 337 3132	<i>Sca</i> I 831 1646	<i>Ssp</i> I 166 1322
<i>Xmn</i> I 651 1527				

ENZYMES THAT DO NOT CUT:

*Acc*I, *Afl*II, *Age*I, *Asc*I, *Avr*II, *Bgl*II, *Bpu*10I, *Bpu*1102I, *Bsa*BI, *Bse*RI, *Bsm*I, *Bsp*EI, *Bsp*GI, *Bsr*GI, *Bst*1107I, *Cla*I, *Dra*III, *Eag*I, *Eco*47III, *Fse*I, *Hind*III, *Hpa*I, *Kpn*I, *Mun*I, *Nco*I, *Nde*I, *Ngo*AIV, *Nhe*I, *Not*I, *Nru*I, *Nsi*I, *Pac*I, *Pme*I, *Pml*I, *Psh*AI, *Psp*5II, *Rsr*II, *Sac*I, *Sac*II, *Sal*I, *Sex*AI, *Sfi*I, *Sgf*I, *Sgr*AI, *Sna*BI, *Spe*I, *Sph*I, *Srf*I, *Sse*8387I, *Stu*I, *Sty*I, *Sun*I, *Xba*I, *Xho*I

pGEX2T (Fig. 16.7.1)	4948	Amp ^r	<i>Bam</i> HI, <i>Sma</i> I, <i>Eco</i> RI	Pharmacia
--------------------------------	------	------------------	--------------------------------------------	-----------

COMMENTS: Allows expression of fusion proteins to glutathione-*S*-transferase; identical to pGEX1 except polylinker codes for an additional thrombin cleavage site.

ENZYMES THAT CUT ONCE:

<i>Aat</i> II 1224	<i>Alw</i> NI 2622	<i>Apa</i> I 3858	<i>Apa</i> BI 3331	<i>Ava</i> I 935
<i>Bae</i> I 4753	<i>Bam</i> HI 930	<i>Ban</i> II 3858	<i>Bsa</i> I 2077	<i>Bsa</i> AI 1125
<i>Bsp</i> MI 62	<i>Bss</i> HII 4058	<i>Bst</i> EII 3828	<i>Bsu</i> 36I 4740	<i>Dsa</i> I 4848
<i>Eam</i> 1105I 2143	<i>Eco</i> NI 268	<i>Eco</i> RI 940	<i>Eco</i> RV 4097	<i>Mlu</i> I 3647
<i>Msc</i> I 465	<i>Nar</i> I 4287	<i>Nsp</i> V 655	<i>Pfl</i> 1108I 2124	<i>Pfl</i> MI 3229
<i>Pst</i> I 1901	<i>Rle</i> AI 832	<i>Sma</i> I 937	<i>Swa</i> I 685	<i>Tth</i> 111I 1118

ENZYMES THAT CUT TWICE:

<i>Bcl</i> II 692 3661	<i>Bgl</i> II 2025 4668	<i>Bsb</i> I 1089 1599	<i>Bsm</i> BI 1014 4262	<i>Bsm</i> FI 245 4745
<i>Bsp</i> LU11I 496 3031	<i>Bsr</i> FI 2058 3333	<i>Drd</i> I 1066 2929	<i>Drd</i> II 921 3370	<i>Eco</i> 57I 1477 2489
<i>Eco</i> O109I 290 1163	<i>Fsp</i> I 1920 4658	<i>Pvu</i> I 1774 4639	<i>Sap</i> I 337 3148	<i>Sca</i> I 831 1662
<i>Ssp</i> I 166 1338	<i>Xmn</i> I 651 1543			

ENZYMES THAT DO NOT CUT:

*Acc*I, *Afl*II, *Age*I, *Asc*I, *Avr*II, *Bgl*II, *Bpu*10I, *Bpu*1102I, *Bsa*BI, *Bse*RI, *Bsm*I, *Bsp*EI, *Bsp*GI, *Bsr*GI, *Bst*1107I, *Cla*I, *Dra*III, *Eag*I, *Eco*47III, *Fse*I, *Hind*III, *Hpa*I, *Kpn*I, *Mun*I, *Nco*I, *Nde*I, *Ngo*AIV, *Nhe*I, *Not*I, *Nru*I, *Nsi*I, *Pac*I, *Pme*I, *Pml*I, *Psh*AI, *Psp*5II, *Rsr*II, *Sac*I, *Sac*II, *Sal*I, *Sex*AI, *Sfi*I, *Sgf*I, *Sgr*AI, *Sna*BI, *Spe*I, *Sph*I, *Srf*I, *Sse*8387I, *Stu*I, *Sty*I, *Sun*I, *Xba*I, *Xho*I

continued

Vectors

A.5.37

Table A.5.8 Vectors for Expression of Fusion Proteins in *E. coli*, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pGEX3X (Fig. 16.7.1)	4952	Amp ^r	<i>Bam</i> HI, <i>Sma</i> I, <i>Eco</i> RI	Pharmacia

COMMENTS: Allows expression of fusion proteins to glutathione-*S*-transferase; identical to pGEX1 except polylinker codes for an additional factor Xa cleavage site.

ENZYMES THAT CUT ONCE:

<i>Aat</i> II 1228	<i>Alw</i> NI 2626	<i>Apa</i> I 3862	<i>Apa</i> BI 3335	<i>Ava</i> I 939
<i>Bae</i> I 4757	<i>Bam</i> HI 934	<i>Ban</i> II 3862	<i>Bsa</i> I 2081	<i>Bsa</i> AI 1129
<i>Bsp</i> MI 62	<i>Bss</i> HIII 4062	<i>Bst</i> EII 3832	<i>Bsu</i> 36I 4744	<i>Drd</i> II 3374
<i>Dsa</i> I 4852	<i>Eam</i> 1105I 2147	<i>Eco</i> NI 268	<i>Eco</i> RI 944	<i>Eco</i> RV 4101
<i>Mlu</i> I 3651	<i>Msc</i> I 465	<i>Nar</i> I 4291	<i>Nsp</i> V 655	<i>Pfl</i> 1108I 2128
<i>Pfl</i> MI 3233	<i>Pst</i> I 1905	<i>Rle</i> AI 832	<i>Sma</i> I 941	<i>Swa</i> I 685
<i>Tth</i> 111I 1122				

ENZYMES THAT CUT TWICE:

<i>Bcl</i> II 692 3665	<i>Bgl</i> II 2029 4672	<i>Bsb</i> I 1093 1603	<i>Bsm</i> BI 1018 4266	<i>Bsm</i> FI 245 4749
<i>Bsp</i> LU11I 496 3035	<i>Bsr</i> FI 2062 3337	<i>Drd</i> I 1070 2933	<i>Eco</i> 57I 1481 2493	<i>Eco</i> O109I 290 1167
<i>Fsp</i> I 1924 4662	<i>Pvu</i> I 1778 4643	<i>Sap</i> I 337 3152	<i>Sca</i> I 831 1666	<i>Ssp</i> I 166 1342
<i>Xmn</i> I 651 1547				

ENZYMES THAT DO NOT CUT:

*Acc*I, *Afl*II, *Age*I, *Asc*I, *Avr*II, *Bgl*II, *Bpu*10I, *Bpu*102I, *Bsa*BI, *Bse*RI, *Bsm*I, *Bsp*EI, *Bsp*GI, *Bsr*GI, *Bst*1107I, *Cla*I, *Dra*III, *Eag*I, *Eco*47III, *Fse*I, *Hind*III, *Hpa*I, *Kpn*I, *Mun*I, *Nco*I, *Nde*I, *Ngo*AIV, *Nhe*I, *Not*I, *Nru*I, *Nsi*I, *Pac*I, *Pme*I, *Pml*I, *Psh*AI, *Psp*5II, *Rsr*II, *Sac*I, *Sac*II, *Sal*I, *Sex*AI, *Sfi*I, *Sgf*I, *Sgr*AI, *Sna*BI, *Spe*I, *Sph*I, *Srf*I, *Sse*8387I, *Stu*I, *Sty*I, *Sun*I, *Xba*I, *Xho*I

pMAL-c2 (Fig. 16.6.1)	6646	Amp ^r	<i>Xmn</i> I, <i>Eco</i> RI, <i>Bam</i> HI, <i>Xba</i> I, <i>Sal</i> I, <i>Pst</i> I, <i>Hind</i> III	New England Biolabs
---------------------------------	------	------------------	----------------------------------------------------------------------------------------------------------	---------------------

COMMENTS: Allows expression and purification of fusions to maltose binding protein; has an exact deletion of the *malE* signal sequence corresponding to bases 1581-1605 of pMAL-p2, allows production of single-stranded DNA in *E. coli*.

ENZYMES THAT CUT ONCE:

<i>Apa</i> I 642	<i>Bam</i> HI 2701	<i>Bgl</i> II 1887	<i>Bpu</i> 10I 6483	<i>Bpu</i> 1102I 2314
<i>Bsm</i> I 2148	<i>Bsp</i> GI 6064	<i>Bsp</i> LU11I 5588	<i>Bst</i> 1107I 5819	<i>Bst</i> EII 612
<i>Dra</i> III 4602	<i>Eam</i> 1105I 4205	<i>Eco</i> RI 2695	<i>Eco</i> RV 881	<i>Hind</i> III 2727
<i>Hpa</i> I 937	<i>Mlu</i> I 431	<i>Msc</i> I 6621	<i>Nar</i> I 1071	<i>Nco</i> I 2489
<i>Nde</i> I 5767	<i>Ngo</i> AIV 4703	<i>Pfl</i> 1108I 4186	<i>Pfl</i> MI 13	<i>Pst</i> I 2723
<i>Sac</i> I 2638	<i>Sal</i> I 2713	<i>Sap</i> I 5705	<i>Sca</i> I 3724	<i>Sse</i> 8387I 2723
<i>Sty</i> I 2489	<i>Sun</i> I 1818	<i>Swa</i> I 4377	<i>Tth</i> 111I 5844	<i>Xba</i> I 2707
<i>Xmn</i> I 2688				

ENZYMES THAT CUT TWICE:

<i>Acc</i> I 2714 5818	<i>Afl</i> III 431 5588	<i>Alw</i> NI 2173 5179	<i>Ava</i> I 2671 4491	<i>Bcl</i> II 445 2570
<i>Bgl</i> II 2882 4087	<i>Bsa</i> BI 2203 6395	<i>Bsm</i> BI 1046 5945	<i>Bsm</i> FI 2989 6317	<i>Bsp</i> EI 1647 6399
<i>Bss</i> HIII 842 1964	<i>Dra</i> I 3627 4377	<i>Drd</i> II 154 4607	<i>Eco</i> 47III 2438 6338	<i>Eco</i> O109I 6583 6625
<i>Psp</i> 5II 6583 6625	<i>Rca</i> I 3365 4868	<i>Rle</i> AI 1688 3050		

ENZYMES THAT DO NOT CUT:

*Aat*II, *Afl*II, *Age*I, *Asc*I, *Avr*II, *Bae*I, *Bse*RI, *Bsr*GI, *Bsu*36I, *Cla*I, *Eag*I, *Eco*NI, *Fse*I, *Kpn*I, *Mun*I, *Nhe*I, *Not*I, *Nru*I, *Nsi*I, *Nsp*V, *Pac*I, *Pme*I, *Pml*I, *Psh*AI, *Rsr*II, *Sac*II, *Sex*AI, *Sfi*I, *Sgf*I, *Sgr*AI, *Sma*I, *Sna*BI, *Spe*I, *Sph*I, *Srf*I, *Stu*I, *Xho*I

continued

Table A.5.8 Vectors for Expression of Fusion Proteins in *E. coli*, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pMAL-p2 (Fig. 16.6.1)	6721	Amp ^r	<i>XmnI</i> , <i>EcoRI</i> , <i>BamHI</i> , <i>XbaI</i> , <i>SalI</i> , <i>PstI</i> , <i>HindIII</i>	New England Biolabs

COMMENTS: Allows expression and purification of fusions to maltose binding protein; includes *malE* signal sequence, which potentially directs fusion protein to the periplasm; allows production of single-stranded DNA in *E. coli*.

ENZYMES THAT CUT ONCE:

<i>ApaI</i> 642	<i>BamHI</i> 2776	<i>BglII</i> 1962	<i>Bpu10I</i> 6558	<i>Bpu1102I</i> 2389
<i>BsmI</i> 2223	<i>BspGI</i> 6139	<i>BspLU11I</i> 5663	<i>Bst1107I</i> 5894	<i>BstEII</i> 612
<i>DraIII</i> 4677	<i>Eam1105I</i> 4280	<i>EcoRI</i> 2770	<i>EcoRV</i> 881	<i>HindIII</i> 2802
<i>HpaI</i> 937	<i>MluI</i> 431	<i>MscI</i> 6696	<i>NaeI</i> 4780	<i>NarI</i> 1071
<i>NcoI</i> 2564	<i>NdeI</i> 5842	<i>Pfl1108I</i> 4261	<i>PflMI</i> 13	<i>PstI</i> 2798
<i>SalI</i> 2788	<i>SapI</i> 5780	<i>ScaI</i> 3799	<i>Sse8387I</i> 2798	<i>SstI</i> 2713
<i>StyI</i> 2564	<i>SunI</i> 1893	<i>SwaI</i> 4452	<i>Tth111I</i> 5919	<i>XbaI</i> 2782
<i>XmnI</i> 2763				

ENZYMES THAT CUT TWICE:

<i>AccI</i> 2789 5893	<i>AflIII</i> 431 5663	<i>AlwNI</i> 2248 5254	<i>AvaI</i> 2746 4566	<i>BclI</i> 445 2645
<i>BglII</i> 2957 4162	<i>BsaBI</i> 2278 6470	<i>BsmBI</i> 1046 6020	<i>BsmFI</i> 3064 6392	<i>BspEI</i> 1722 6474
<i>BssHIII</i> 842 2039	<i>DraI</i> 3702 4452	<i>DrdII</i> 154 4682	<i>Eco47III</i> 2513 6413	<i>EcoO109I</i> 6658 6700
<i>Psp5II</i> 6658 6700	<i>RcaI</i> 3440 4943	<i>RleAI</i> 1763 3125		

ENZYMES THAT DO NOT CUT:

AatII, *AflII*, *AgeI*, *AscI*, *AvrII*, *BaeI*, *BseRI*, *BsrGI*, *Bsu36I*, *ClaI*, *EagI*, *EcoNI*, *FseI*, *KpnI*, *MunI*, *NheI*, *NotI*, *NruI*, *NsiI*, *NspV*, *PacI*, *PmeI*, *PmlI*, *PshAI*, *RsrII*, *SacII*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SmaI*, *SnaBI*, *SpeI*, *SphI*, *SrfI*, *StuI*, *XhoI*

pSKF301 See Table A.5.5 and Fig. 16.3.2 for full listings.

continued

Table A.5.8 Vectors for Expression of Fusion Proteins in *E. coli*, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pTRXFUS (Fig. 16.8.1)	3585	Amp ^r	<i>KpnI</i> , <i>BamHI</i> , <i>XbaI</i> , <i>SalI</i> , <i>PstI</i> , <i>RsrII</i>	Invitrogen; John McCoy, Genetics Institute

COMMENTS: Used for construction of protein fusions to *E. coli* thioredoxin; contains enterokinase cleavage site to release protein of interest from thioredoxin component; *RsrII* site can be used to clone short peptide-encoding DNA sequences for antigen production.

ENZYMES THAT CUT ONCE:

AatII 69	AccI 3102	AflIII 1876	AlwNI 1467	ApoI 2100
BamHI 3089	BpmI 919	BsaI 922	BsaBI 2126	BsmI 2259
BspEI 2322	BspLU11I 1876	BspMI 3114	BstXI 2663	ClaI 2905
DrdII 3058	DsaI 2942	Eam1105I 988	EcoO109I 8	HindIII 3180
KpnI 3084	MunI 2696	NarI 3345	NdeI 2722	NsiI 2675
Pfl1108I 969	PstI 3111	PvuII 3275	RsrII 2823	SalI 3101
SapI 1993	ScaI 507	SmaI 3086	SspI 183	XbaI 3095

ENZYMES THAT CUT TWICE:

AvaI 2578 3084	BglII 870 3335	BsbI 2158 3465	BsmBI 3536 3578	BsrDI 754 928
BsrFI 903 3048	DrdI 1774 3490	EcoNI 2249 2636	FspI 765 3325	HincII 2160 3103
MmeI 1483 1667	NspI 1880 3548	Psp1406I 386 759	PvuI 619 3306	VspI 813 2048
XmnI 388 3027				

ENZYMES THAT DO NOT CUT:

AflII, *AgeI*, *ApaI*, *ApaBI*, *AscI*, *AvrII*, *BaeI*, *BanII*, *BbsI*, *BclI*, *BglII*, *Bpu10I*, *Bpu1102I*, *BsaAI*, *BseRI*, *BsgI*, *BsmFI*, *BspGI*, *BsrGI*, *BssHII*, *Bst1107I*, *BstEII*, *Bsu36I*, *DraIII*, *EagI*, *Eco47III*, *EcoRI*, *EcoRV*, *FseI*, *HpaI*, *MluI*, *MscI*, *NcoI*, *NgoAIV*, *NheI*, *NotI*, *NruI*, *NspV*, *PacI*, *PflMI*, *PmeI*, *PmlI*, *PshAI*, *Psp5II*, *RleAI*, *SacI*, *SacII*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SnaBI*, *SpeI*, *SphI*, *SrfI*, *Sse8387I*, *StuI*, *StyI*, *SunI*, *SwaI*, *Tth111I*, *XcmI*, *XhoI*

pTrec99A See Table A.5.5 and Fig. 1.5.11 for full listings.

pUR278	5351	Amp ^r	(L09146)
---------------	------	------------------	----------

COMMENTS: *lacZ* fusion-protein expression vector using *lacUV5* promoter.

ENZYMES THAT CUT ONCE:

BaeI 467	BamHI 3290	BclI 1576	Bpu10I 735	Bpu1102I 3243
BsaI 4271	BsaBI 1561	BsbI 3793	BssHII 1728	Bst1107I 2995
Bsu36I 454	DraIII 1420	DrdII 1399	Eam1105I 4337	Eco47III 2065
EcoO109I 3357	EcoRV 1343	HindIII 3308	NdeI 3187	Pfl1108I 4318
PstI 4095	RleAI 3076	SacI 2170	SalI 3296	SapI 5342
ScaI 3856	SunI 3001	XbaI 3302	XmnI 3737	

ENZYMES THAT CUT TWICE:

AatII 853 3418	AccI 2994 3297	AvaI 1568 3011	BanII 2170 3219	BsmBI 857 1797
BspLU11I 2989 5225	BspMI 1004 2876	BsrDI 4103 4277	BstXI 2450 3067	ClaI 1054 3315
DrdI 2339 5123	EcoRI 3234 3339	FspI 372 4114	HgiEII 418 4643	HpaI 656 1280
NspI 2993 5229	PflMI 1475 1841	Psp1406I 3735 4108	SspI 1460 3532	XcmI 2447 3192

ENZYMES THAT DO NOT CUT:

AflII, *AgeI*, *ApaI*, *ApaBI*, *AscI*, *AvrII*, *BglII*, *BsaAI*, *BseRI*, *BsmI*, *BspEI*, *BspGI*, *BsrGI*, *BstEII*, *EagI*, *EcoNI*, *FseI*, *KpnI*, *MscI*, *MunI*, *NarI*, *NcoI*, *NgoAIV*, *NheI*, *NotI*, *NruI*, *NsiI*, *NspV*, *PacI*, *PmeI*, *PmlI*, *PshAI*, *Psp5II*, *RsrII*, *SacII*, *SexAI*, *SfiI*, *SgfI*, *SgrAI*, *SmaI*, *SnaBI*, *SpeI*, *SphI*, *SrfI*, *Sse8387I*, *StuI*, *StyI*, *SwaI*, *Tth111I*, *XhoI*

Vectors**A.5.40**

Table A.5.9 Vectors for Specialized Functions

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
mGP1-2 (Fig. 16.2.3)	9955	<i>E. coli</i> containing F or F' factor necessary for susceptibility to infection	Not applicable	Stanley Tabor, Harvard Medical School

COMMENTS: M13-based vector that produces T7 RNA polymerase; useful for production of gene products that are toxic to *E. coli* cell, in conjunction with pT7-5, pT7-6, and pT7-7 (see Table A.5.5) or other plasmids using the T7 promoter to express recombinant genes.

ENZYMES THAT CUT ONCE:

AflIII 8389	Alw44I 4743	BglII 9142	BglIII 9640	BsaHI 6002
BseRI 2023	BsgI 7085	BsmI 1752	BsmBI 5977	Bsp24I 2187
BspGI 3642	BsrFI 5613	BstXI 6994	DrdI 5765	EcoRI 6231
FspI 7262	HindIII 8987	HpaI 7971	KpnI 8741	MscI 5082
MunI 8863	NarI 6002	NgoAIV 5613	NruI 6395	PacI 4136
PstI 8982	PvuI 9113	SacI 6241	SalI 8972	SmaI 6246
SnaBI 1270	SwaI 9491			

ENZYMES THAT CUT TWICE:

AlwNI 2192 6609	AvaI 5825 6244	BaeI 1800 2141	BanII 5647 6241	Bsp24I 2155 2187
ClaI 2528 9588	EciI 1964 9310	Eco47III 2712 3041	GdiII 6036 8998	RcaI 1299 8195
RleAI 3007 8671				

ENZYMES THAT DO NOT CUT:

AatII, AgeI, ApaI, AscI, AvrII, BamHI, BclI, BsaI, BspEI, BssHII, Bst1107I, BstEII, EagI, Eam1105I, Eco57I, EcoNI, EcoO109I, EcoRV, FseI, MluI, NcoI, NheI, NotI, NsiI, NspV, PflMI, PmeI, PmlI, PshAI, Psp5II, RsrII, SacII, SapI, ScaI, SexAI, SfiI, SgfI, SgrAI, SpeI, SphI, SrfI, Sse8387I, StuI, StyI, SunI, Tth111I, XbaI, XcmI, XhoI

pAD1 (Fig. 6.12.2)	~4000	<i>E. coli</i> : Kan ^r	<i>EcoRI, SfiI, NotI, SacII, PstI, PvuII, BamHI</i>	David Kurnit, University of Michigan—Ann Arbor
------------------------------	-------	-----------------------------------	-----------------------------------------------------	------------------------------------------------

COMMENTS: Used to screen λ library by homologous recombination with sequences cloned into pAD1. See Fig. 6.12.2 for restriction site information.

pEG202 (Fig. 20.7.3)	~9800	<i>E. coli</i> : Amp ^r ; Yeast: HIS3 ⁺	<i>EcoRI, XmaI, BamHI, SalI, NcoI, NotI, XhoI, SalI, PstI</i>	Roger Brent, Massachusetts General Hospital and Harvard Medical School
--------------------------------	-------	--------------------------------------------------------------	---------------------------------------------------------------	------------------------------------------------------------------------

COMMENTS: Used to clone “bait” proteins as fusions to the DNA-binding protein *lexA*. See Fig. 20.7.4 for restriction site information.

pJG4-5	6200	<i>E. coli</i> : Amp ^r ; Yeast: TRP1 ⁺	<i>EcoRI, XhoI</i>	Roger Brent, Massachusetts General Hospital and Harvard Medical School
---------------	------	--------------------------------------------------------------	--------------------	------------------------------------------------------------------------

COMMENTS: Used to express cDNAs or other coding sequences to look for proteins that interact with “bait” protein cloned in pEG202. See Fig. 20.7.3 for restriction site information.

continued

Table A.5.9 Vectors for Specialized Functions, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
pGP1-2 (Fig. 16.2.2)	7267	<i>E. coli</i> : Kan ^r	Not applicable	Stanley Tabor, Harvard Medical School

COMMENTS: Used to produce T7 RNA polymerase by heat induction for expression of genes cloned on pT7-5, pT7-6, and pT7-7 (see Table A.5.5) or other plasmids using the T7 promoter to express recombinant genes.

ENZYMES THAT CUT ONCE:

AflIII 4636	BamHI 4059	BclI 687	BglII 1166	BsaI 1218
BspEI 6988	BspGI 1837	Bst1107I 2318	BstEII 4018	BstXI 6039
ClaI 2784	DrdII 5490	Eam1105I 1284	EcoRI 1	HpaI 5058
KpnI 4292	NdeI 4228	NheI 2327	NspV 1592	PflMI 3356
PshAI 150	PstI 1042	PvuI 3093	SacI 7266	SacII 2081
SgfI 3093	SgrAI 2241	SspI 3018	XhoI 2691	

ENZYMES THAT CUT TWICE:

AgeI 1917 2241	AlwNI 1731 6423	BanII 2748 7266	BseRI 1656 1699	BsgI 2266 5946
BsmBI 797 3109	BsmFI 3544 4952	BspMI 1045 5968	BsrBI 73 2157	BsrGI 4385 5784
Bsu36I 5620 5791	DraI 1398 1417	Eco57I 819 7049	EcoNI 3005 7064	FspI 1061 5767
MunI 853 4162	NruI 2750 6634	Psp1406I 1055 6239	SmaI 2967 7257	StuI 2377 3717
VspI 1109 3292				

ENZYMES THAT DO NOT CUT:

AatII, AflIII, Alw44I, ApaI, AscI, AvrII, BaeI, BglII, BsaAI, BsaHI, BspLU11I, BssHII, DrdI, EagI, EcoO109I, EcoRV, FseI, MluI, MscI, NarI, NcoI, NgoAIV, NotI, PacI, PmeI, PmlI, Psp5II, PvuII, RsrII, SalI, SapI, ScaI, SexAI, SfiI, SnaBI, SpeI, SphI, SrfI, Sse8387I, StyI, SunI, SwaI, Tth111I, XbaI, XcmI

pNTK (Fig. 23.5.1)	6993	Mammalian cells: G418 ^r , negative selection with gancyclovir for HSV-TK ⁻ ; <i>E. coli</i> : Amp ^r	See Fig. 23.5.1	Richard Mortensen, University of Michigan Medical Center
------------------------------	------	--------------------------------------------------------------------------------------------------------------------------------------	-----------------	----------------------------------------------------------

COMMENTS: Used for disrupting targeted cloned genes in mammalian cells by homologous recombination; negative selection with gancyclovir selects against single crossover/integration events by the plasmid. **NOTE:** Partially sequenced, so additional sites may be present; some locations approximate.

ENZYMES THAT CUT ONCE:

AatII 1534	ApaBI 729	BamHI 2127	BspLU11I 6596	DraIII 4565
Eam1105I 5708	HincII 4108	HindIII 4091	KpnI 4131	MluI 619
NarI 2800	NotI 18	NruI 885	Pfl1108I 5689	PshAI 1644
RsrII 3316	SacI 5	SacII 12	SalI 4106	ScaI 5227
SnaBI 833	SunI 669	XhoI 4112	XmnI 5108	

ENZYMES THAT CUT TWICE:

AccI 829 4107	AgeI 148 2264	Alw44I 5036 6282	BglII 3765 3773	Bpu1102I 310 2426
BsaI 1859 5642	BsmI 465 2581	BsmBI 282 2398	BstXI 13 3531	ClaI 3472 4098
DrdII 2031 4557	EcoNI 915 1218	EcoRI 29 2145	HgiEII 539 6014	NcoI 2012 3232
PvuI 4288 5339	SmaI 1701 2135	SspI 4770 4903	StuI 331 2447	Tth111I 2004 2918
XbaI 25 2115	XcmI 794 1987			

ENZYMES THAT DO NOT CUT:

AflII, AscI, AvrII, BaeI, BbsI, BclI, BsaBI, BsrGI, Bst1107I, BstEII, Bsu36I, Eco47III, FseI, HpaI, MunI, NdeI, NheI, NsiI, NspV, PacI, PflMI, PmeI, PmlI, RleAI, SexAI, SfiI, SgfI, SgrAI, SrfI, Sse8387I, SwaI

Table A.5.10 Vectors of Historical Interest

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
Charon 4a (Fig. 1.10.6)	~46,500	Size selection for packaging coupled with dephosphorylated arms; if <i>EcoRI</i> sites are used, recombinant phage are Bio ⁻ and Lac ⁻	<i>EcoRI</i> , <i>XbaI</i>	Frederick Blattner, University of Wisconsin—Madison
λ2001 (Fig. 1.10.3)	~43,500	Recombinant phage are Spi ⁻	See Fig. 1.10.3	ATCC #37331
Wild-type λ bacteriophage (Fig. 1.10.1)	48502	Infect sensitive host and grow	Not applicable	ATCC (J02459)

COMMENTS: Used to construct many earlier libraries which are still being used. See Fig. 1.10.6 for restriction site information.

COMMENTS: Used for cloning large (10.4-20 kbp) fragments; see Fig. 1.10.3 for restriction site information.

COMMENTS: Temperate bacteriophage; parent of many vectors used to construct genomic libraries.

ENZYMES THAT CUT ONCE:

ApaI 10090	NaeI 20042	NarI 45680	NheI 34679	SnaBI 12190
SunI 19323	XbaI 24508	XhoI 33498		

ENZYMES THAT CUT TWICE:

AscI 3522 16649	AvrII 24322 24396	BsaI 11418 42709	BspLU11I 628 39395	Bsu36I 26718 34319
EagI 19944 36654	Eco47III 20997 37059	KpnI 17057 18560	PmeI 8462 16296	SacI 24776 25881
SalI 32745 33244	Tth111I 11205 36123			

ENZYMES THAT DO NOT CUT:

FseI, *NotI*, *PacI*, *SfiI*, *SpeI*, *SrfI*, *SwaI*

continued

Table A.5.10 Vectors of Historical Interest, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
Wild-type M13 bacteriophage (Fig. 1.14.3)	6407	<i>E. coli</i> containing F or F' factor necessary for susceptibility to infection	Not applicable	ATCC (V00604, J02461, M10377)

COMMENTS: Parent to M13mp series and other single-stranded DNA vectors.

ENZYMES THAT CUT ONCE:

AccI 6091	Alw44I 4743	AlwNI 2192	AvaI 5825	BamHI 2220
BanII 5647	BpmI 6041	BseRI 2010	BsmI 1752	BsmFI 5091
Bsp1407I 1021	BspGI 3642	BspHI 1299	BsrFI 5613	BstYI 2220
DraIII 5721	DrdI 5765	DsaI 2763	EaeI 5080	EarI 4080
EciI 1964	HincII 6407	HpaI 6407	MscI 5082	NaeI 5615
NspBII 1633	PacI 4136	RleAI 3007	Sau96I 5724	SnaBI 1270
SwaI 5943				

ENZYMES THAT CUT TWICE:

Alw21I 4747 5469	BaeI 1800 2141	BanI 1249 5677	BsaBI 1153 3978	BsbI 640 5811
BsiEI 1425 3882	Bsp24I 2155 2187	BspMI 1104 2267	ClaI 2528 6040	Eco47III 2712 3041
EcoRII 1013 1965	MspI 2800 6157	NciI 1925 5996	Pfl1108I 1258 3425	Psp1406I 4635 5931
TaqII 1439 2461	XmnI 361 2650			

ENZYMES THAT DO NOT CUT:

AatII, AflII, AgeI, ApaI, ApaBI, AscI, AvaII, AvrII, BbsI, BcgI, BcgI, BclI, BglI, BglII, Bpu1102I, BsaI, BsaHI, BsgI, BsiI, BsmBI, BspEI, BssHII, Bst1107I, BstEII, BstXI, Bsu36I, EagI, Eam1105I, Eco57I, EcoNI, EcoO109I, EcoRI, EcoRV, FseI, FspI, GdiII, HgiEII, HindIII, KpnI, MluI, MunI, NarI, NcoI, NheI, NotI, NruI, NsiI, NspV, PflMI, PmeI, PmlI, PshAI, Psp5II, PstI, PvuI, PvuII, RsrII, SacI, SacII, SalI, SapI, ScaI, SexAI, SfiI, SgfI, SgrAI, SmaI, SpeI, SphI, SrfI, Sse8387I, StuI, StyI, SunI, Tth111I, XbaI, XcmI, XhoI

continued

Table A.5.10 Vectors of Historical Interest, continued

Name (figure)	Size (bp)	Selection/screening method	Cloning sites	Sources (GenBank accession number)
2μm circle (Fig. 13.4.4)	6318	Not applicable	Not applicable	ATCC (J01347)

COMMENTS: Naturally occurring plasmid found in almost all strains of *S. cerevisiae*; extremely stable mitotically.

ENZYMES THAT CUT ONCE:

Alw21I 1426	AlwNI 1893	ApaI 4391	BanII 4391	BcgI 267
BclII 304	BsiI 1324	BsmBI 1919	Bsp24I 5028	BspMI 5328
BsrBI 2465	Bsu36I 5604	ClaI 6285	EagI 5424	GdiII 5424
HgiEII 5604	HpaI 2966	MscI 266	NcoI 5024	NdeI 340
NruI 2560	PstI 2656	SnaBI 3608	SphI 5572	StyI 5024
SwaI 6178	TaqII 1729	VspI 1665		

ENZYMES THAT CUT TWICE:

AflIII 1839 5274	AgeI 1815 5007	ApaBI 2887 5670	AvaI 3258 4764	BaeI 5388 5390
BcgI 233 267	Bpu10I 2466 4944	Bpu1102I 2524 5564	BsbI 586 4066	BsgI 1801 5286
BsmI 1565 2275	BsmFI 1887 4713	Bsp24I 5028 5060	Bsp1286I 1426 4391	BspLU11I 1839 5274
DrdII 10 4488	EaeI 264 5424	EcoRI 1 2407	EcoRV 5540 5914	HincII 219 2966
MspA1I 2031 2135	NciI 2916 4587	PleI 3411 4962	Psp5II 1333 2470	Psp1406I 1864 6159
PvuI 1758 2732	PvuII 2031 2135	RleAI 219 1269	SapI 962 1010	SspI 808 3844
StuI 1621 2490	TaqII 1729 2597	XbaI 703 3945	XcmI 396 4257	

ENZYMES THAT DO NOT CUT:

AatII, AflII, Alw44I, AscI, AvrII, BamHI, BanI, Bce83I, BglI, BglII, BsaI, BsaBI, BsaHI, BsaXI, BscGI, BspEI, BspGI, BssHII, BstEII, BstXI, BstYI, DraIII, DrdI, Eam1105I, EcoNI, FseI, KpnI, MluI, NarI, NgoAIV, NheI, NotI, PacI, Pfl1108I, PmeI, PmlI, PshAI, RsrII, SacI, SacII, SalI, ScaI, SexAI, SfiI, SgfI, SgrAI, SmaI, SrfI, Sse8387I, SunI, Tth111I, XhoI

Key References

- Genetics Computer Group. 1991. Program Manual for the Wisconsin Package, Version 7, April 1991. Genetics Computer Group, Madison, Wis.
- Sikorski, R.S. and Hieter, P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* 122:19-27.

Contributed by Lisa M. Albright